

Encyclopedia of
**BIOMATERIALS AND
BIOMEDICAL ENGINEERING**

SECOND EDITION, VOLUME 1

Edited by

GARY E. WNEK

GARY L. BOWLIN



informa
healthcare

Encyclopedia of
**BIOMATERIALS AND
BIOMEDICAL ENGINEERING**

Encyclopedia of
**BIOMATERIALS AND
BIOMEDICAL ENGINEERING**

S E C O N D E D I T I O N , V O L U M E 1

Edited by

GARY E. WNEK

*Case Western Reserve University
Cincinnati, Ohio, USA*

GARY L. BOWLIN

*Virginia Commonwealth University
Richmond, Virginia, USA*

informa

healthcare

New York London

Informa Healthcare USA, Inc.
52 Vanderbilt Avenue
New York, NY 10017

© 2008 by Informa Healthcare USA, Inc. (except as noted on the opening page of each article). Informa Healthcare is an Informa business.

No claim to original U.S. Government works
Printed in the United States of America on acid-free paper
10 9 8 7 6 5 4 3 2 1

International Standard Book Number-10: 1-4200-7953-0 (Hardcover: Volume 1)
International Standard Book Number-13: 978-1-4200-7953-1 (Hardcover: Volume 1)
International Standard Book Number-10: 1-4200-7954-9 (Hardcover: Volume 2)
International Standard Book Number-13: 978-1-4200-7954-8 (Hardcover: Volume 2)
International Standard Book Number-10: 1-4200-7955-7 (Hardcover: Volume 3)
International Standard Book Number-13: 978-1-4200-7955-5 (Hardcover: Volume 3)
International Standard Book Number-10: 1-4200-7956-5 (Hardcover: Volume 4)
International Standard Book Number-13: 978-1-4200-7956-2 (Hardcover: Volume 4)

This book contains information obtained from authentic and highly regarded sources. Reprinted material is quoted with permission, and sources are indicated. A wide variety of references are listed. Reasonable efforts have been made to publish reliable data and information, but the author and the publisher cannot assume responsibility for the validity of all materials or for the consequence of their use.

No part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access www.copyright.com (<http://www.copyright.com/>) or contact the Copyright Clearance Center, Inc. (CCC) 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Library of Congress Cataloging-in-Publication Data

Encyclopedia of biomaterials and biomedical engineering / edited by
Gary E. Wnek, Gary L. Bowlin. – 2nd ed.

p. ; cm.

Includes bibliographical references and index.

ISBN-13: 978-1-4200-7802-2 (hb : alk. paper)

ISBN-10: 1-4200-7802-X (hb : alk. paper) 1. Biomedical materials–
Encyclopedias. 2. Biomedical engineering–Encyclopedias. I. Wnek,
Gary E. II. Bowlin, Gary L.

[DNLM: 1. Biomedical and Dental Materials–Encyclopedias–English.

2. Biomedical Engineering–Encyclopedias–English. QT 13 E5292 2008]

R857.M3E53 2008

610.28'403–dc22

2008010753

For Corporate Sales and Reprint Permissions call 212-520-2700 or write to: Sales Department, 52 Vanderbilt Avenue, 16th floor, New York, NY 10017.

**Visit the Informa Web site at
www.informa.com**

**and the Informa Healthcare Web site at
www.informahealthcare.com**

Editors

Gary E. Wnek

*Case Western Reserve University
Cleveland, Ohio, U.S.A.*

Gary L. Bowlin

*Virginia Commonwealth University
Richmond, Virginia, U.S.A.*

Editorial Advisory Board

James Anderson

*Case Western Reserve University, Cleveland, Ohio,
U.S.A.*

Deon Bezuidenhout

*University of Cape Town, Cape Town,
South Africa*

Rena Bizios

*University of Texas at San Antonio, San Antonio,
Texas, U.S.A.*

Larry L. Hench

Imperial College London, London, United Kingdom

Allan S. Hoffman

*University of Washington, Seattle, Washington,
U.S.A.*

Yoshito Ikada

*Suzuka University of Medical Science, Suzuka,
Mie, Japan*

Robert Langer

*Massachusetts Institute of Technology, Cambridge,
Massachusetts, U.S.A.*

Gerald E. Miller

*Virginia Commonwealth University, Richmond,
Virginia, U.S.A.*

Arnauld Nicogossian

*George Mason University, Fairfax, Virginia,
U.S.A.*

Nikolaos Peppas

*Purdue University, West Lafayette, Indiana,
U.S.A.*

Christine E. Schmidt

University of Texas, Austin, Texas, U.S.A.

Jerome S. Schultz

*University of California, Riverside, Riverside,
California, U.S.A.*

V. Prasad Shastri

Vanderbilt University, Nashville, Tennessee, U.S.A.

Debra Trantolo

*Cambridge Scientific, Inc., Cambridge, Massa-
chusetts, U.S.A.*

Rachel L. Williams

University of Liverpool, Liverpool, United Kingdom

Nicholas Ziats

*Case Western Reserve University, Cleveland, Ohio,
U.S.A.*

List of Contributors

- Soumyadipta Acharya** | *The University of Akron, Akron, Ohio, U.S.A.*
- Clayton Adam** | *Queensland University of Technology, Brisbane, Queensland, Australia*
- Babita Agrawal** | *University of Alberta, Edmonton, Alberta, Canada*
- Carlos A. Aguilar** | *University of Texas at Austin, Austin, Texas, U.S.A.*
- Juana Maria Alfaro** | *Cordis de Mexico, Juarez, Mexico*
- Heather Ambrose** | *Virginia Commonwealth University, Richmond, Virginia, U.S.A.*
- Catherine G. Ambrose** | *University of Texas–Houston School of Medicine, Houston, Texas, U.S.A.*
- Farid Amirouche** | *University of Illinois at Chicago, Chicago, Illinois, U.S.A.*
- Javier de Ana** | *University of Michigan, Ann Arbor, Michigan, U.S.A.*
- Luke Aram** | *DePuy Orthopedics|Johnson & Johnson Company*
- James H. Arps** | *Southwest Research Institute, San Antonio, Texas, U.S.A.*
- James A. Arrowood** | *Virginia Commonwealth University, Richmond, Virginia, U.S.A.*
- Tetsuo Asakura** | *Tokyo University of Agriculture and Technology, Koganei, Tokyo, Japan*
- Anthony Atala** | *Wake Forest Institute for Regenerative Medicine, Winston-Salem, North Carolina, U.S.A.*
- Kyriacos A. Athanasiou** | *Rice University, Houston, Texas, U.S.A.*
- Jay Audett** | *ALZA Corporation, Mountain View, California, U.S.A.*
- François A. Auger** | *LOEX, Hôpital du Saint-Sacrement du CHA, and Laval University, Quebec City, Quebec, Canada*
- Konstantinos Avgoustakis** | *University of Patras, Patras, Rio, Greece*
- Stephen F. Badylak** | *University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.*
- Richard Baker** | *Royal Children’s Hospital, Melbourne, Victoria, Australia*
- Deepak Banerjee** | *Virginia Commonwealth University, Richmond, Virginia, U.S.A.*
- Peter B. Barker** | *Johns Hopkins University School of Medicine, Baltimore, Maryland, U.S.A.*
- Clive M. Baumgarten** | *Virginia Commonwealth University, Richmond, Virginia, U.S.A.*
- Roger Bayston** | *University Hospital, Nottingham, United Kingdom*
- Matthew J. Beckman** | *Virginia Commonwealth University, and Department of Orthopedic Surgery & Biochemistry, Richmond, Virginia, U.S.A.*
- Paul J. Benkeser** | *Georgia Institute of Technology and Emory University, Atlanta, Georgia, U.S.A.*
- Besim Ben-Nissan** | *University of Technology Sydney, New South Wales, Australia*
- Stanley A. Berger** | *University of California, Berkeley, California, U.S.A.*
- Earl J. Bergey** | *State University of New York at Buffalo, Buffalo, New York, U.S.A.*
- François Berthiaume** | *Massachusetts General Hospital, Harvard Medical School, and the Shriners Hospital for Children, Boston, Massachusetts, U.S.A.*
- Serena Best** | *University of Cambridge, Cambridge, U.K.*
- Kazuhisa Bessho** | *Kyoto University, Sakyo-Ku, Kyoto, Japan*
- Deon Bezuidenhout** | *University of Cape Town Medical School, and Groote Schuur Hospital, Cape Town, South Africa*
- Vineet Bhandari** | *Yale University School of Medicine and Children’s Hospital, New Haven, Connecticut, U.S.A.*
- Luc Bilodeau** | *Montreal Heart Institute, Montreal, Quebec, Canada*
- Katharina Bittner** | *BioSurgery European Marketing, Baxter BioScience, Munich, Germany*

- Matthew J. Bizzarro** / *Yale University School of Medicine and Children's Hospital, New Haven, Connecticut, U.S.A.*
- Kirby S. Black** / *CryoLife, Inc., Kennesaw, Georgia, U.S.A.*
- Michael Blakeney** / *University of London, London, U.K.*
- Torsten Blunk** / *University of Regensburg, Regensburg, Germany*
- Aldo R. Boccaccini** / *Imperial College London, London, United Kingdom*
- Aldo Boccaccini** / *Imperial College London, London, United Kingdom*
- Ivan Bojanić** / *University of Zagreb, Zagreb, Croatia*
- Jack C. Bokros** / *Medical Carbon Research Institute, Austin, Texas, U.S.A.*
- Eugene D. Boland** / *Virginia Commonwealth University, Richmond, Virginia, U.S.A.*
- Babul Borah** / *Procter & Gamble Pharmaceuticals, Mason, Ohio, U.S.A.*
- J. T. Borenstein** / *Charles Stark Draper Laboratory, Cambridge, Massachusetts, U.S.A.*
- Adele L. Boskey** / *Hospital for Special Surgery, New York, New York, U.S.A.*
- Gary L. Bowlin** / *Virginia Commonwealth University, Richmond, Virginia, U.S.A.*
- James R. Bowman, III** / *Virginia Commonwealth University, Richmond, Virginia, U.S.A.*
- Kellye D. Branch** / *Mississippi State University, Mississippi State, Mississippi, U.S.A.*
- Ricardo R. Brau** / *Massachusetts Institute of Technology, Cambridge, Massachusetts, U.S.A.*
- Robert C. Bray** / *University of Calgary, Calgary, Alberta, Canada, and McCaig Centre for Joint Injuries and Arthritis Research*
- William T. Brinkman** / *Emory University School of Medicine, Atlanta, Georgia, U.S.A.*
- William C. Broaddus** / *Virginia Commonwealth University, Richmond, Virginia, U.S.A.*
- Kelvin G. M. Brockbank** / *Organ Recovery Systems, Charleston, South Carolina, U.S.A.*
- Timothy J. Broderick** / *University of Cincinnati College of Medicine, Cincinnati, Ohio, U.S.A.*
- Anna-Liisa Brownell** / *Massachusetts General Hospital, Boston, Massachusetts, U.S.A.*
- Ian C. Bruce** / *McMaster University, Hamilton, Ontario, Canada*
- Donald M. Brunette** / *University of British Columbia, Vancouver, British Columbia, Canada*
- Gil Bub** / *McGill University, Montreal, Quebec, Canada*
- Anne-Marie Buckle** / *The University of Manchester, Manchester, England, U.K.*
- Joel D. Bumgardner** / *Mississippi State University, Mississippi State, Mississippi, U.S.A.*
- Andrew Burd** / *The Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, Hong Kong*
- Christie D. Bures** / *University of Texas at Austin, Austin, Texas, U.S.A.*
- Elisabeth H. Burger** / *Vrije Universiteit, Amsterdam, The Netherlands*
- Karen J.L. Burg** / *Clemson University, Clemson, South Carolina, U.S.A.*
- Lori L. Burrows** / *Hospital for Sick Children Research Institute, Toronto, Ontario, Canada*
- Sandra C. P. Cachinho** / *University of Liverpool, United Kingdom*
- Jinhua Cao** / *Purdue University, West Lafayette, Indiana, U.S.A.*
- E. L. Carniel** / *University of Padua, Padua, Italy*
- Marcus E. Carr, Jr.** / *Virginia Commonwealth University, Richmond, Virginia, U.S.A.*
- Patrick Carrier** / *LOEX, Hôpital du Saint-Sacrement de CHA, and Laval University, Quebec City, Quebec, Canada*
- Elliot L. Chaikof** / *Emory University School of Medicine, and Georgia Institute of Technology, Atlanta, Georgia, U.S.A.*
- Christina Chan** / *Massachusetts General Hospital, Harvard Medical School, and the Shriners Hospital for Children, Boston, Massachusetts, U.S.A.*
- W. Chang** / *Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire, U.S.A.*
- Aditya Chaubey** / *Clemson University, Clemson, South Carolina, U.S.A.*
- Babak Chehroudi** / *University of British Columbia, Vancouver, British Columbia, Canada*
- Q. Chen** / *The University of Wisconsin–Madison, Madison, Wisconsin, U.S.A.*
- Qi-Zhi Chen** / *Imperial College London, London, United Kingdom*

- Shaochen Chen** / *University of Texas at Austin, Austin, Texas, U.S.A.*
- Ruizhen Rachel Chen** / *Virginia Commonwealth University, Richmond, Virginia, U.S.A.*
- S. Chethana** / *Central Food Technological Research Institute, Mysore, Karnataka, India*
- Vikram S. Chib** / *Northwestern University, Chicago, Illinois, U.S.A.*
- Traian V. Chirila** / *Lions Eye Institute, Nedlands, Western Australia, Australia*
- Paula Chmielewski** / *Procter & Gamble Pharmaceuticals, Mason, Ohio, U.S.A.*
- Charles D. Choi** / *General Dynamics Advanced Information Systems, Fairfax, Virginia, and University of Michigan, Ann Arbor, Michigan, U.S.A.*
- P. Christian** / *University of Manchester, Manchester, Greater Manchester, U.K.*
- Seth I. Christian** / *Mississippi State University, Mississippi State, Mississippi, U.S.A.*
- Chih-Chang Chu** / *Cornell University, Ithaca, New York, U.S.A.*
- Paul Kim Ho Chu** / *City University of Hong Kong, Kowloon, Hong Kong*
- Hsi-Wei Chung** / *Mayo Clinic College of Medicine, Rochester, Minnesota, U.S.A.*
- Timothy A. M. Chuter** / *University of California San Francisco, San Francisco, California, U.S.A.*
- Graeme M. Clark** / *University of Melbourne, East Melbourne, Victoria, Australia*
- William R. Clark** / *Gambro Renal Products (Intensive Care), Indianapolis, Indiana, U.S.A.*
- Suzanne Conroy** / *AST Products, Inc., Billerica, Massachusetts, U.S.A.*
- Rory A. Cooper** / *University of Pittsburgh, and VA Pittsburgh Healthcare System, Pittsburgh, Pennsylvania, U.S.A.*
- Alastair N. Cormack** / *New York State College of Ceramics at Alfred University, Alfred, New York, U.S.A.*
- Michel Cormier** / *ALZA Corporation, Mountain View, California, U.S.A.*
- Brian N. Craig** / *Lamar University, Beaumont, Texas, U.S.A.*
- Nancy A. Crigger** / *Baylor College of Medicine, Houston, Texas, U.S.A.*
- J. K. Critser** / *University of Missouri, Columbia, Missouri, U.S.A.*
- J.N. Amritha De Croos** / *Mount Sinai Hospital, Toronto, Ontario, Canada*
- M. J. Cross** / *Australian Institute of Musculo-Skeletal Research, Sydney, New South Wales, Australia, University of Sydney, Sydney, New South Wales, Australia*
- Guy Daculsi** / *Nantes University, and INSERM EM 9903, INSERM, Nantes, France, France*
- Peter Daddona** / *ALZA Corporation, Mountain View, California, U.S.A.*
- Jawwad A. Darr** / *University College London, London, U.K.*
- Neil Davies** / *University of Cape Town, Cape Town, South Africa*
- Jeff Davis** / *University of South Carolina, Columbia, South Carolina, U.S.A.*
- Patti E. Dawson** / *CryoLife, Inc., Kennesaw, Georgia, U.S.A.*
- Richard M. Day** / *St. Mark's Hospital, London, United Kingdom*
- Geoff Dearnaley** / *Southwest Research Institute, San Antonio, Texas, U.S.A.*
- Richard E. Debski** / *University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.*
- K. C. Dee** / *Tulane University, New Orleans, Louisiana, U.S.A.*
- Stephen P. Denyer** / *University of Brighton, Brighton, and Cardiff University, Cardiff, United Kingdom*
- Tejal Desai** / *Boston University, Boston, Massachusetts, U.S.A.*
- Steven Deutsch** / *The Pennsylvania State University, University Park, Pennsylvania, U.S.A.*
- Scott L. Diamond** / *University of Pennsylvania, Philadelphia, Pennsylvania, U.S.A.*
- Robert F. Diegelmann** / *Virginia Commonwealth University, Richmond, Virginia, U.S.A.*
- James P. Dioria** / *Baxter Technology Resources, Round Lake, Illinois, U.S.A.*
- L. Di-Silvio** / *University College London, Royal National Orthopaedic Hospital Trust, Middlesex, and The Guy's, King's, and St. Thomas' Hospital Medical and Dental Institute, King's College London, London, United Kingdom*
- Charles R. Doarn** / *Virginia Commonwealth University, Richmond, Virginia, and University of Cincinnati, Cincinnati, Ohio, U.S.A.*
- P. J. Doherty** / *University of Liverpool, Liverpool, United Kingdom*

- A. J. Domb** / *The Hebrew University of Jerusalem, Jerusalem, Israel*
- Cristina Donini** / *University of Texas at Austin, Austin, Texas, U.S.A., and Plan-les-Ouates, Geneva, Switzerland*
- Simon J. Doran** / *University of Surrey, Guildford, Surrey, and Institute of Cancer Research, Sutton, Surrey, U.K.*
- Robert H. Doremus** / *Rensselaer Polytechnic Institute, Troy, New York, U.S.A.*
- Karen A. Duca** / *Virginia State University, Blacksburg, Virginia, U.S.A.*
- Thomas Dufresne** / *Procter & Gamble Pharmaceuticals, Mason, Ohio, U.S.A.*
- James C.Y. Dunn** / *University of California, Los Angeles, California, U.S.A.*
- Dominique M. Durand** / *Case Western Reserve University, Cleveland, Ohio, U.S.A.*
- Daniel Eberli** / *Wake Forest Institute for Regenerative Medicine, Winston Salem, North Carolina, U.S.A.*
- Clive Edwards** / *School of Biological Sciences, The University of Liverpool, Liverpool, U.K.*
- Glenn A. Edwards** / *University of Melbourne, Werribee, Victoria, Australia*
- Igor R. Efimov** / *Case Western Reserve University, Cleveland, Ohio, U.S.A.*
- Jennifer H. Elisseeff** / *Johns Hopkins University, Baltimore, Maryland, U.S.A.*
- Ville Ellä** / *Tampere University of Technology, Institute of Biomaterials, Tampere, Finland*
- Özgür Erdoğan** / *Çukurova University, Balcali, Adana, Turkey*
- Thomas Eschenhagen** / *University-Hospital Hamburg-Eppendorf, Hamburg, Germany*
- Paul G. Espy** / *Virginia Commonwealth University, Richmond, Virginia, U.S.A.*
- Anna Fallon** / *Georgia Institute of Technology, Atlanta, Georgia, U.S.A.*
- Anna M. Fallon** / *Georgia Institute of Technology, Atlanta, Georgia, U.S.A.*
- Sergio Fantini** / *Tufts University, Medford, Massachusetts, U.S.A.*
- Richard G. A. Faragher** / *University of Brighton, Brighton, United Kingdom*
- William J. Federspiel** / *University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.*
- Ding-Yu Fei** / *Virginia Commonwealth University, Richmond, Virginia, U.S.A.*
- Fatiha El Feninat** / *Laval University, Quebec City, Quebec, Canada*
- John Fisher** / *University of Leeds, Leeds, United Kingdom*
- Willem Flameng** / *Katholieke Universiteit Leuven, Belgium*
- Odette Fokapu** / *University of Technology of Compiègne, Compiègne, France*
- Arnold A. Fontaine** / *The Pennsylvania State University, University Park, Pennsylvania, U.S.A.*
- Kevin R. Forrester** / *University of Calgary, Calgary, Alberta, Canada*
- Kenneth R. Foster** / *University of Pennsylvania, Philadelphia, Pennsylvania, U.S.A.*
- Joanna D. Fromstein** / *University of Toronto, Toronto, Ontario, Canada*
- James G. Fujimoto** / *Massachusetts Institute of Technology, Cambridge, Massachusetts, U.S.A.*
- Hisao Fukui** / *Aichi-Gakuin University, Chikusa-ku, Nagoya, Japan*
- Keertik S. Fulzele** / *Mississippi State University, Mississippi State, Mississippi, U.S.A.*
- Ricky K.Y. Fu** / *City University of Hong Kong, Kowloon, Hong Kong*
- Edward P. Furlani** / *State University of New York at Buffalo, Buffalo, New York, U.S.A.*
- Franck Furno** / *University Hospital, Nottingham, United Kingdom, and Stockhausen GmbH and Co. KG, Krefeld, Germany*
- Simon M. Gabe** / *St. Mark's Hospital, London, United Kingdom*
- Michael Gaitan** / *National Institute of Standards and Technology, Gaithersburg, Maryland, U.S.A.*
- Bob Gale** / *ALZA Corporation, Mountain View, California, U.S.A.*
- Didier Gamet** / *University of Technology of Compiègne, Compiègne, France*
- Israel Gannot** / *National Institute of Health, NICHD, Biomedical Stochastic Physics, Bethesda, Maryland, U.S.A., and Faculty of Engineering, Tel-Aviv University, Tel-Aviv, Israel*
- Dayong Gao** / *University of Kentucky, Lexington, Kentucky, U.S.A.*
- Martin Oswaldo Mendez Garcia** / *Polytechnic University, Milan, Italy*

- Leslie A. Geddes** / *Purdue University, West Lafayette, Indiana, U.S.A.*
- E. Gentleman** / *Tulane University, New Orleans, Louisiana, U.S.A.*
- Lucie Germain** / *LOEX, Hôpital du Saint-Sacrement du CHA, and Laval University, Quebec City, Quebec, Canada*
- Salem Ghrebi** / *University of British Columbia, Vancouver, British Columbia, Canada*
- Claude J. Giasson** / *LOEX, Hôpital du Saint-Sacrement du CHA, and Laval University, Quebec City, Quebec, Canada*
- Thomas W. Gilbert** / *University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.*
- George T. Gillies** / *Virginia Commonwealth University, Richmond, and University of Virginia, Charlottesville, Virginia, U.S.A.*
- David L. Glaser** / *University of Pennsylvania Health System, Philadelphia, Pennsylvania, U.S.A.*
- Julie Glowacki** / *Brigham and Women's Hospital, Harvard Medical School, and Harvard School of Dental Medicine, Boston, Massachusetts, U.S.A.*
- Mukul S. Goel** / *University of Pennsylvania, Philadelphia, Pennsylvania, U.S.A.*
- Alan H. Goldstein** / *New York State College of Ceramics at Alfred University, Alfred, New York, U.S.A.*
- Cheryl T. Gomillion** / *Clemson University, Clemson, South Carolina, U.S.A.*
- Mark Gonzalez** / *University of Illinois at Chicago, Chicago, Illinois, U.S.A.*
- Stuart B. Goodman** / *Stanford Medical Center, Stanford, California, U.S.A.*
- Richard L. Goodwin** / *University of South Carolina, Columbia, South Carolina, U.S.A.*
- J.E. Gough** / *University of Manchester, Manchester, Greater Manchester, U.K.*
- Achim Göpferich** / *University of Regensburg, Regensburg, Germany*
- Lloyd D. Graham** / *Riverside Life Sciences Centre, New South Wales, Australia*
- Edward M. Greenfield** / *Case Western Reserve University, Cleveland, Ohio, U.S.A.*
- Howard P. Greisler** / *Loyola University Medical Center, Maywood, Illinois, U.S.A.*
- Tapio Grönfors** / *University of Kuopio, Finland*
- Kelley Grorud** / *Edgewood College, Madison, Wisconsin, U.S.A.*
- Jianjun Guan** / *University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.*
- Sylvain L. Guerin** / *LOEX, Hôpital du Saint-Sacrement du CHA, and Laval University, Quebec City, Quebec, Canada*
- N. Gurav** / *University College London, Royal National Orthopaedic Hospital Trust, Middlesex, United Kingdom*
- Peter J. Haar** / *Virginia Commonwealth University, Richmond, Virginia, U.S.A.*
- Mutaz B. Habal** / *Tampa Bay Craniofacial Center, Tampa, Florida, U.S.A.*
- Marie Haïdopoulos** / *Laval University, Quebec City, Quebec, Canada*
- Nadim James Hallab** / *Rush-Presbyterian-St. Lukes Medical Center, Chicago, Illinois, U.S.A.*
- Douglas Hamilton** / *University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.*
- Douglas W. Hamilton** / *University of British Columbia, Vancouver, British Columbia, Canada*
- Stephen R. Hanson** / *Emory University, Atlanta, Georgia, U.S.A., and Oregon Health & Science University, Beaverton, Oregon, U.S.A.*
- Brett M. Harnett** / *Virginia Commonwealth University, Richmond, Virginia, and University of Cincinnati, Cincinnati, Ohio, U.S.A.*
- Alex Hartov** / *Dartmouth College, Hanover, New Hampshire, U.S.A.*
- Julie Hasenwinkel** / *Syracuse University, Syracuse, New York, U.S.A.*
- Mineo Hashizume** / *Nara Institute of Science and Technology, Nara, Japan*
- Vasif Hasirci** / *Middle East Technical University, Ankara, Turkey*
- Tomokazu Hattori** / *Faculty of Science and Technology, Meijo University, Tempaku-ku, Nagoya, Japan*
- Zhaoming He** / *Georgia Institute of Technology, Atlanta, Georgia, U.S.A.*
- Sam L. Helgerson** / *BioScience R & D Baxter BioScience, Fremont, California, U.S.A.*
- Kristie A. Henchir** / *University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.*

- Larry L. Hench** | *Imperial College London, London, United Kingdom*
- O. M. Hess** | *University Hospital, Bern, Switzerland*
- Celia R. Hicks** | *Lions Eye Institute, Nedlands, Western Australia, Australia*
- Robert S. D. Higgins** | *Rush-Presbyterian-St. Luke's Medical Center, Chicago, Illinois, U.S.A.*
- Motohiro Hirose** | *National Institute of Advanced Industrial Science and Technology (AIST), Amagasaki, Hyogo, Japan*
- Jeffrey W. Holmes** | *Columbia University, New York, New York, U.S.A.*
- Haoqing Hou** | *The University of Akron, Akron, Ohio, U.S.A.*
- S. M. Hudson** | *North Carolina State University, Raleigh, North Carolina, U.S.A.*
- Xin-Hua Hu** | *East Carolina University, Greenville, North Carolina, U.S.A.*
- Paul Human** | *University of Cape Town Medical School, and Groote Schuur Hospital, Cape Town, South Africa*
- John A. Hunt** | *University of Liverpool, Liverpool, United Kingdom*
- Thomas L. Husted** | *University of Cincinnati College of Medicine, Cincinnati, Ohio, U.S.A.*
- Rejhan Idrizi** | *Bernard O'Brien Institute of Microsurgery, and St. Vincent's Hospital, Melbourne, Victoria, Australia*
- Yoshito Ikada** | *Suzuka University of Medical Science, Suzuka, Mie, Japan*
- Atsushi Ikeda** | *Nara Institute of Science and Technology, Nara, Japan*
- Ilko K. Ilev** | *Center for Devices and Radiological Health, Office of Science and Engineering Laboratories, Division of Physics, Rockville, Maryland, U.S.A.*
- Eileen Ingham** | *University of Leeds, Leeds, United Kingdom*
- Koji Ioku** | *National Institute of Advanced Industrial Science and Technology (AIST), Amagasaki, Hyogo, Japan*
- Kazuhiko Ishihara** | *The University of Tokyo, Tokyo, Bunkyo-ku, Japan*
- Atsuo Ito** | *National Institute of Advanced Industrial Science and Technology, Tsukuba, Ibaraki, Japan*
- Alan Ivković** | *University of Zagreb, Zagreb, Croatia*
- Hiroo Iwata** | *Kyoto University, Kyoto, Sakyo-Ku, Japan*
- Joshua Jacobs** | *Rush-Presbyterian-St. Lukes Medical Center, Chicago, Illinois, U.S.A.*
- E. Duco Jansen** | *Vanderbilt University, Nashville, Tennessee, U.S.A.*
- Nicolas M. Jedynekiewicz** | *The University of Liverpool, Liverpool, U.K.*
- Jorge H. Jimenez** | *Georgia Institute of Technology, Atlanta, Georgia, U.S.A.*
- Julian R. Jones** | *Imperial College London, London, United Kingdom*
- James W. Jones** | *Baylor College of Medicine, Houston, Texas, U.S.A.*
- Jeffrey I. Joseph** | *Jefferson Medical College of Thomas Jefferson University, Philadelphia, Pennsylvania, U.S.A.*
- Martti Juhola** | *University of Tampere, Finland*
- M. R. Kaazempur-Mofrad** | *Massachusetts Institute of Technology, Cambridge, and Massachusetts General Hospital, and Harvard Medical School, Boston, Massachusetts, U.S.A.*
- Tonya Kaltenbach** | *Stanford University School of Medicine, Stanford, California, U.S.A.*
- Rita A. Kandel** | *Mount Sinai Hospital, Toronto, Ontario, Canada*
- Leon Kane-Maguire** | *University of Wollongong, Wollongong, New South Wales, Australia*
- W. John Kao** | *University of Wisconsin-Madison, Madison, Wisconsin, U.S.A.*
- Vigneshwar Kasirajan** | *Medical College of Virginia, Richmond, Virginia, U.S.A.*
- Toshihiro Kasuga** | *Nagoya Institute of Technology, Gokiso-cho Showa-ku, Nagoya, Japan*
- Koichi Kato** | *Kyoto University, Sakyo-ku, Kyoto, Japan*
- J. Lawrence Katz** | *University of Missouri–Kansas City School of Dentistry, Kansas City, Missouri, and University of Texas–Houston School of Medicine, Houston, Texas, U.S.A.*
- Daisuke Kawagoe** | *Tohoku University, Aoba-ku, Sendai, Japan*
- Jill S. Kawalec** | *Ohio College of Podiatric Medicine, Cleveland, Ohio, U.S.A.*

- Christopher Keen** / *Virginia Commonwealth University, Richmond, Virginia, U.S.A*
- William S. Kerwin** / *University of Washington, Seattle, Washington, U.S.A.*
- Gilson Khang** / *Chonbuk National University, Jeonju, Korea*
- Antoine E. Khoury** / *Hospital for Sick Children, Toronto, Ontario, Canada*
- Jun-ichi Kikuchi** / *Nara Institute of Science and Technology, Nara, Japan*
- Hugh Kim** / *University of British Columbia, Vancouver, British Columbia, Canada*
- Kyo-Han Kim** / *Kyunpook National University, Jung-gu, Daegu, South Korea*
- Moon Suk Kim** / *Korea Research Institute of Chemical Technology, Yuseong, Daejeon, Korea*
- Yoshihiro Kiritoshi** / *The University of Tokyo, Tokyo, Bunkyo-ku, Japan*
- Jenneke Klein-Nulend** / *Vrije Universiteit, Amsterdam, The Netherlands*
- Jonathan P. Van Kleunen** / *University of Pennsylvania Health System, Philadelphia, Pennsylvania, U.S.A.*
- Kenneth R. Knight** / *Bernard O'Brien Institute of Microsurgery, and St. Vincent's Hospital, Melbourne, Victoria, Australia*
- Kenneth L. Koch** / *Wake Forest University School of Medicine, Winston-Salem, North Carolina, U.S.A.*
- Hiroko Kojima** / *National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Ibaraki, Japan*
- Frank K. Ko** / *Drexel University, Philadelphia, Pennsylvania, U.S.A.*
- Claire F. Komives** / *San Jose State University, San Jose, California, U.S.A.*
- Thomas J. Koob** / *Shriners Hospitals for Children, Tampa, Florida, U.S.A.*
- Feza Korkusuz** / *Middle East Technical University, Ankara, Turkey*
- Petek Korkusuz** / *Hacettepe University, Sıhhiye, Ankara, Turkey*
- Noriko Kotobuki** / *National Institute of Advanced Industrial Science and Technology (AIST), Amagasaki, Hyogo, Japan*
- Katherine M. Kulig** / *Massachusetts General Hospital, Boston, Massachusetts, U.S.A.*
- Rakesh Kumar** / *University of Alberta, Edmonton, Alberta, Canada*
- M.N.V. Ravi Kumar** / *National Institute of Pharmaceutical Education and Research (NIPER), Mohali, India*
- Fred M. Kusumoto** / *Lovelace Medical Center, Albuquerque, New Mexico, U.S.A.*
- Kathy E. Laber** / *Medical University of South Carolina, Charleston, South Carolina, U.S.A.*
- Roger Laham** / *Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, Massachusetts, U.S.A.*
- Andres Laib** / *Scanco Medical AG, Bassersdorf, Switzerland*
- Henry Lai** / *University of Washington, Seattle, Washington, U.S.A.*
- James C. K. Lai** / *College of Pharmacy and Biomedical Research Institute, Idaho State University, Pocatello, Idaho, U.S.A.*
- James M. Lambert** / *SiTech, LLC, Richmond, Virginia, U.S.A.*
- Philippe Lam** / *Genentech, Inc., South San Francisco, California, U.S.A.*
- James Laredo** / *Loyola University Medical Center, Maywood, Illinois, U.S.A.*
- Rolf Larsson** / *University Hospital, Uppsala, Sweden*
- Robert A. Latour, Jr.** / *Clemson University, Clemson, South Carolina, U.S.A.*
- Jennie Baier Leach** / *University of Texas, Austin, Texas, U.S.A.*
- Benjamin Lee** / *Virginia Commonwealth University, Richmond, Virginia, U.S.A.*
- Brendan Lee** / *Baylor College of Medicine, Houston, Texas, U.S.A.*
- Hai Bang Lee** / *Korea Research Institute of Chemical Technology, Yuseong, Daejeon, Korea*
- Jin Ho Lee** / *Hannam University, Daejeon, Korea*
- Susan S. Lee** / *National Eye Institute (NEI), National Institutes of Health (NIH), Bethesda, and Johns Hopkins University, Baltimore, Maryland, U.S.A.*
- Jae Sung Lee** / *Pohang University of Science and Technology, Pohang, Korea*

- Racquel LeGeros** / *New York University College of Dentistry, New York, New York, U.S.A*
- Nic D. Leipzig** / *Rice University, Houston, Texas, U.S.A.*
- Martin L. Lenhardt** / *Virginia Commonwealth University, Richmond, Virginia, U.S.A.*
- Hwa-Liang Leo** / *Georgia Institute of Technology, Atlanta, Georgia, U.S.A.*
- Catherine A. Leonard** / *University of Calgary, Calgary, Alberta, Canada*
- Solomon W. Leung** / *Idaho State University, Pocatello, Idaho, U.S.A.*
- Andrew L. Lewis** / *Biocompatibles UK Limited, Farnham, Surrey, United Kingdom*
- Gladius Lewis** / *The University of Memphis, Memphis, Tennessee, U.S.A.*
- Baruch B. Lieber** / *University of Miami, Coral Gables, Florida, U.S.A.*
- Andrea Liebmann-Vinson** / *BD Technologies, Research Triangle Park, North Carolina, U.S.A.*
- Martin Lietz** / *NMI Naturwissenschaftliches und Medizinisches Institut an der Universität Tübingen, Reutlingen, Germany*
- E. N. Lightfoot** / *Virginia State University, Blacksburg, Virginia, U.S.A.*
- Mengnai Li** / *Department of Orthopaedic Surgery, Richmond, Virginia, U.S.A.*
- Xingde Li** / *University of Washington, Seattle, Washington, U.S.A.*
- Steven J. Lindauer** / *Virginia Commonwealth University, Richmond, Virginia, U.S.A.*
- Christian Lindqvist** / *Helsinki University and University Hospital, Helsinki, Finland*
- Tung-Liang Lin** / *AST Products, Inc., Billerica, Massachusetts, U.S.A.*
- Huinan Liu** / *Purdue University, West Lafayette, Indiana, and, Brown University, Providence, Rhode Island, U.S.A.*
- Jonathan T.C. Liu** / *Stanford University School of Medicine, Stanford, California, U.S.A.*
- G. A. Livesay** / *Tulane University, New Orleans, Louisiana, U.S.A.*
- Andrew W. Lloyd** / *University of Brighton, Brighton, United Kingdom*
- Laurie E. Locascio** / *National Institute of Standards and Technology, Gaithersburg, Maryland, U.S.A.*
- Ih-Houng Loh** / *AST Products, Inc., Billerica, Massachusetts, U.S.A.*
- Zerina Lokmic** / *Bernard O'Brien Institute of Microsurgery, and St. Vincent's Hospital, Melbourne, Victoria, Australia*
- Albert Lozano-Nieto** / *The Pennsylvania State University, Lehman, Pennsylvania, U.S.A.*
- Junfeng Lu** / *University of Kentucky, Lexington, Kentucky, U.S.A.*
- Yi Lu** / *University of Texas at Austin, Austin, Texas, U.S.A.*
- Jun Qing Lu** / *East Carolina University, Greenville, North Carolina, U.S.A.*
- Ling Ma** / *International Rectifier Corporation, El Segundo, California, U.S.A.*
- Ting Ma** / *Stanford Medical Center, Stanford, California, U.S.A.*
- S. Machan** / *Australian Institute of Musculo-Skeletal Research, Sydney, New South Wales, Australia*
- Eli E. Machtei** / *Rambam Medical Center and The Faculty of Medicine, Haifa, Israel*
- Marc Madou** / *University of California, Irvine, California, and Integrated Nanosystems Research Facility (INRF), U.S.A.*
- Agnese Magnani** / *University of Siena, and University Center "Colle di Val d'Elsa," Colle di Val d'Elsa (SI), Siena, Italy*
- Luca T. Mainardi** / *Polytechnic University, Milan, Italy*
- Viraj P. Mane** / *Baylor College of Medicine, Houston, Texas, U.S.A.*
- Keefe B. Manning** / *The Pennsylvania State University, University Park, Pennsylvania, U.S.A.*
- Athanassios Mantalaris** / *Imperial College London, London, United Kingdom*
- Diego Mantovani** / *Laval University, Quebec City, Quebec, Canada*
- Hui Mao** / *Emory University School of Medicine, Atlanta, Georgia, U.S.A.*
- Gerard H. Markx** / *The University of Manchester, Manchester, England, U.K.*
- Erika J. Martin** / *Virginia Commonwealth University, Richmond, Virginia, U.S.A.*
- Shojiro Matsuda** / *Gunze Limited, Ayabe, Kyoto, Japan*
- Timothy M. Maul** / *University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.*

- Sally L. McArthur** / *University of Sheffield, Sheffield, United Kingdom*
- Gretchen J. McAuliffe** / *Cornell University, Ithaca, New York, U.S.A.*
- William McCormack** / *Baylor College of Medicine, Houston, Texas, U.S.A.*
- Keith M. McLean** / *Commonwealth Scientific & Industrial Research Organisation (CSIRO) Molecular Science, Clayton, Victoria, Australia*
- Carl McMillin** / *Synthetic Body Parts, Brecksville, Ohio, U.S.A.*
- Lito C. Mejia** / *EnduraTEC Systems Corporation, Minnetonka, Minnesota, U.S.A.*
- Ronald C. Merrell** / *Virginia Commonwealth University, Richmond, Virginia, U.S.A.*
- Aurora Messina** / *Bernard O'Brien Institute of Microsurgery, and St. Vincent's Hospital, Melbourne, Victoria, Australia*
- Bart Meuris** / *Katholieke Universiteit Leuven, Belgium*
- Evangelia Micheli-Tzanakou** / *Rutgers University, Piscataway, New Jersey, U.S.A.*
- Gerald E. Miller** / *Virginia Commonwealth University, Richmond, Virginia, U.S.A.*
- Tommy L. Miller** / *Virginia Commonwealth University, Richmond, Virginia, U.S.A.*
- Anthony J. Minisi** / *McGuire VA Medical Center, Richmond, Virginia, U.S.A.*
- Geraldine Mitchell** / *Bernard O'Brien Institute of Microsurgery, and St. Vincent's Hospital, Melbourne, Victoria, Australia*
- Teruo Miyata** / *Koken Co. Ltd., Toshima-ku, Tokyo, Japan*
- Mark P. Mooney** / *University of Pittsburgh and Carnegie Mellon University, Pittsburgh, Pennsylvania, U.S.A.*
- Susan M. Moore** / *University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.*
- Robert B. More** / *RBMore Associates, Austin, Texas, U.S.A.*
- Wayne Morrison** / *Bernard O'Brien Institute of Microsurgery, and St. Vincent's Hospital, Melbourne, Victoria, Australia*
- Alireza Moshaverinia** / *Ohio State University, Columbus, Ohio, U.S.A.*
- Shaker A. Mousa** / *Albany College of Medicine, and Albany College of Pharmacy, Albany, New York, U.S.A.*
- Ahmed S. Mousa** / *Cornell University, Ithaca, New York, U.S.A.*
- Eike H. Mrosek** / *Mayo Clinic College of Medicine, Rochester, Minnesota, U.S.A.*
- William L. Murphy** / *University of Wisconsin, Madison, Wisconsin, U.S.A.*
- Joy L. Murray** / *University of Texas, Austin, Texas, U.S.A.*
- Karthik Nagapudi** / *Emory University School of Medicine, Atlanta, Georgia, U.S.A.*
- Kwang Woo Nam** / *University of Tokyo, Tokyo, Bunkyo-ku, Japan*
- A. N. Natali** / *University of Padova, Padova, Italy*
- E. A. Nauman** / *Tulane University, New Orleans, Louisiana, U.S.A.*
- Stephen S. Navran** / *Synthecon, Inc., Houston, Texas, U.S.A.*
- Markus Neubauer** / *University of Regensburg, Regensburg, Germany*
- Daniel Newton** / *Virginia Commonwealth University, Richmond, Virginia, U.S.A.*
- S. Nguyen** / *Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire, U.S.A.*
- Paul C. Nicolson** / *Atlanta, Georgia, U.S.A.*
- Alejandro Nieponice** / *University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.*
- Mitsuo Niinomi** / *Toyohashi University of Technology, Tempaku-cho, Toyohashi, Japan*
- Vladimir P. Nikolski** / *Case Western Reserve University, Cleveland, Ohio, U.S.A.*
- Yasuharu Noishiki** / *Yokohama City University School of Medicine, Kanazawa-ku, Yokohama, Japan*
- Giandomenico Nollo** / *University of Trento, and FBK-Fondazione Bruno Kessler, Trento, Italy*
- Yukihiko Nosé** / *Baylor College of Medicine, Houston, Texas, U.S.A.*
- Robert H. Notter** / *University of Rochester School of Medicine, Rochester, New York, U.S.A.*
- Matthew O'Donnell** / *University of Michigan, Ann Arbor, Michigan, U.S.A.*
- Shawn W. O'Driscoll** / *Mayo Clinic College of Medicine, Rochester, Minnesota, U.S.A.*

- Brenda M. Ogle** / *Mayo Clinic, Rochester, Minnesota, U.S.A.*
- Hajime Ohgushi** / *National Institute of Advanced Industrial Science and Technology, Amagasaki City, Hyogo, and Tohoku University, Aoba-ku, Sendai, Japan*
- Ruth J. Okamoto** / *Washington University, St. Louis, Missouri, U.S.A.*
- Regis J. O’Keefe** / *University of Rochester, Rochester, New York, USA*
- Hisashi Okubo** / *Baylor College of Medicine, Houston, Texas, U.S.A.*
- Eric J. Okum** / *Medical College of Virginia, Richmond, Virginia, U.S.A.*
- Joo L. Ong** / *University of Tennessee Health Science Center, Memphis, Tennessee, U.S.A.*
- Richard O. C. Oreffo** / *University of Southampton, Southampton, United Kingdom*
- Zoubeida Ounaies** / *Virginia Commonwealth University, Richmond, Virginia, U.S.A.*
- Rama Padmanabhan** / *ALZA Corporation, Mountain View, California, U.S.A.*
- Rakesh K. Pai** / *University of New Mexico School of Medicine, Albuquerque, New Mexico, U.S.A.*
- Nicki Panoskaltzis** / *Imperial College London, London, United Kingdom*
- E. N. Parish** / *Australian Institute of Musculo-Skeletal Research, Sydney, New South Wales, Australia*
- Kwideok Park** / *The University of Memphis, Memphis, Tennessee, U.S.A.*
- Sang-Won Park** / *Chonnam national University, Dong-Ku, Gwang-ju, South Korea*
- Joon B. Park** / *University of Iowa, Iowa City, Iowa, U.S.A.*
- Grace E. Park** / *Purdue University, West Lafayette, Indiana, U.S.A.*
- Robert S. Parker** / *University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.*
- Cheryl A. Parzel** / *Clemson University, Clemson, South Carolina, U.S.A.*
- Nelesh Patel** / *University of Cambridge, Cambridge, U.K.*
- Ganapathi Patil** / *Central Food Technological Research Institute, Mysore, Karnataka, India*
- Keith D. Paulsen** / *Dartmouth College, Hanover, New Hampshire, U.S.A.*
- P. G. Pavan** / *University of Padova, Padova, Italy*
- Robert R. Pavlis** / *Pittsburg State University, Pittsburg, Kansas, U.S.A.*
- Kristin J. Pawlowski** / *National Aeronautics and Space Administration (NASA), Langley Research Center, Hampton, Virginia, U.S.A.*
- Jon L. Pearlman** / *University of Pittsburgh, and VA Pittsburgh Healthcare System, Pittsburgh, Pennsylvania, U.S.A.*
- Marko Pećina** / *University of Zagreb, Zagreb, Croatia*
- Nicholas A. Peppas** / *University of Texas at Austin, Austin, Texas, U.S.A.*
- Marivalda M. Pereira** / *Federal University of Minas Gerais, Belo Horizonte, MG, Brazil*
- Brad Phipps** / *ALZA Corporation, Mountain View, California, U.S.A.*
- Dominique P. Pioletti** / *Swiss Federal Institute of Technology, and Hôpital Orthopédique de la Suisse Romande, and Bone Bioengineering Group, STI/BIO-E/CRO, Lausanne, Switzerland*
- Federica M. Piras** / *University of Siena, Siena, Italy, University Center “Colle di Val d’Elsa,” Colle di Val d’Elsa (SI), Siena, Italy*
- Eija Marjut Pirhonen** / *Inion Oy Tampere, Finland*
- Roland N. Pittman** / *Virginia Commonwealth University, Richmond, Virginia, U.S.A.*
- Jeffrey L. Platt** / *Mayo Clinic, Rochester, Minnesota, U.S.A.*
- Churn K. Poh** / *University of Kentucky, Lexington, Kentucky, U.S.A.*
- Mariano A. Polack** / *Gainesville, Virginia, U.S.A.*
- Milos R. Popovic** / *University of Toronto, Toronto, Ontario, Canada*
- Dejan B. Popović** / *Aalborg University, Aalborg, Denmark, and University of Belgrade, Belgrade, Serbia*
- Alexandra Elizabeth Porter** / *University of Cambridge, Cambridge, U.K.*
- Sharon C. Presnell** / *BD Technologies, Research Triangle Park, North Carolina, U.S.A.*
- Robert L. Price** / *University of South Carolina, Columbia, South Carolina, U.S.A.*
- D. A. Puleo** / *University of Kentucky, Lexington, Kentucky, U.S.A.*
- Andrew J. Pullan** / *University of Auckland, Auckland, New Zealand*

- K. S. M. S. Raghavarao** / *Central Food Technological Research Institute, Mysore, Karnataka, India*
- P. Rahko** / *The University of Wisconsin–Madison, Madison, Wisconsin, U.S.A.*
- Eric J. Rainis** / *University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.*
- Samuel C. Ramage** / *Department of Biochemistry, Richmond, Virginia, U.S.A.*
- John A. M. Ramshaw** / *Molecular and Health Sciences, Clayton South, Victoria, Australia*
- Ihtesham U. Rehman** / *Queen Mary University of London, London, U.K.*
- Bernd H. A. Rehm** / *Massey University, Palmerston North, New Zealand*
- Gregory G. Reinholz** / *Mayo Clinic College of Medicine, Rochester, Minnesota, U.S.A.*
- Darrell H. Reneker** / *The University of Akron, Akron, Ohio, U.S.A.*
- Darwin R. Reyes** / *National Institute of Standards and Technology, Gaithersburg, Maryland, U.S.A.*
- Daniel C. Richardson** / *Stanford University, Stanford, California, U.S.A.*
- Bruce W. Richman** / *Baylor College of Medicine, Houston, Texas, U.S.A.*
- J. Paul Robinson** / *Purdue University, West Lafayette, Indiana, U.S.A.*
- Michael R. Robinson** / *National Eye Institute (NEI), National Institutes of Health (NIH), Bethesda, Maryland, U.S.A.*
- Gerson Rosenberg** / *Penn State College of Medicine, Hershey, Pennsylvania, U.S.A.*
- Jacob Rosen** / *University of Washington, Seattle, Washington, U.S.A.*
- J.M. Rosen** / *Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire, U.S.A.*
- Tobias Saam** / *University of Washington, Seattle, Washington, U.S.A.*
- E.E. Sabelman** / *Department of Neurology, Kaiser-Permanente Medical Center, Redwood City, California, U.S.A.*
- Michael S. Sacks** / *University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.*
- Chander Sadasivan** / *University of Miami, Coral Gables, Florida, U.S.A.*
- Yudhisthira Sahoo** / *State University of New York at Buffalo, Buffalo, New York, U.S.A.*
- Takashi Saito** / *Health Sciences University of Hokkaido, Ishikari-gun, Hokkaido, Japan*
- Christian Salessé** / *LOEX, Hopital du saint-Sacrement du CHA, and Lava University, Quebec City, Quebec, Canada*
- Paul T. Salo** / *University of Calgary, Calgary, Alberta, Canada*
- Susan Sandeman** / *University of Brighton, Brighton, United Kingdom*
- Archit B. Sanghvi** / *University of Texas, Austin, Texas, U.S.A.*
- Ann Marie Sastry** / *University of Michigan, Ann Arbor, Michigan, U.S.A.*
- Jan C. Schagemann** / *Mayo Clinic College of Medicine, Rochester, Minnesota, U.S.A.*
- Constantin Schizas** / *The Whittington Hospital, University College London, London, and Institute of Orthopaedics, Royal National Orthopaedic Hospital, Stanmore, United Kingdom*
- Burkhard Schlosshauer** / *NMI Naturwissenschaftliches und Medizinisches Institut an der Universität Tübingen, Reutlingen, Germany*
- Christine E. Schmidt** / *University of Texas, Austin, Texas, U.S.A.*
- Craig L. Schmidt** / *Medtronic, Inc., Brooklyn Center, Minnesota, U.S.A.*
- Jerome S. Schultz** / *University of California, Riverside, California, U.S.A.*
- Edward M. Schwarz** / *University of Rochester, Rochester, New York, USA*
- Erik R. Scott** / *Medtronic, Inc., Brooklyn Center, Minnesota, U.S.A.*
- Thomas Seelich** / *BioSurgery R & D, Baxter Bio Science, Vienna, Austria*
- Denis Semwogerere** / *Emory University, Atlanta, Georgia, U.S.A.*
- Alexander B. G. Sevy** / *Massachusetts General Hospital, Boston, Massachusetts, U.S.A.*
- Mohsen Shahinpoor** / *University of New Mexico, Albuquerque, New Mexico, U.S.A.*
- V. Prasad Shastri** / *Vanderbilt University, Nashville, Tennessee, U.S.A.*
- Min-Shyan Sheu** / *AST Products, Inc., Billerica, Massachusetts, U.S.A.*
- Kelly J. Shields** / *Virginia Commonwealth University, Richmond, Virginia, U.S.A.*
- Michael L. Shuler** / *Cornell University, Ithaca, New York, U.S.A.*

- K. Kirk Shung** / *University of Southern California, Los Angeles, California, U.S.A.*
- Michael I. Siegel** / *University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.*
- Olivier Siegrist** / *Hôpital Orthopédique de la Suisse Romande, Lausanne, Switzerland*
- David G. Simpson** / *Virginia Commonwealth University, Richmond, Virginia, U.S.A.*
- George Sines** / *University of California, Los Angeles, California, U.S.A.*
- Paul M. Skarstad** / *Medtronic, Inc., Brooklyn Center, Minnesota, U.S.A.*
- Oskar Skrinjar** / *Georgia Institute of Technology and Emory University, Atlanta, Georgia, U.S.A.*
- John D. Smart** / *University of Portsmouth, Portsmouth, and University of Brighton, Brighton, United Kingdom*
- Mari A. Smith** / *Johns Hopkins University School of Medicine, Baltimore, Maryland, U.S.A.*
- Nadine Barrie Smith** / *Pennsylvania State University, University Park, Pennsylvania, U.S.A.*
- Daniel H. Smith** / *Mississippi State University, Mississippi State, Mississippi, U.S.A.*
- Nicolas P. Smith** / *University of Auckland, Auckland, New Zealand*
- Theo H. Smit** / *Vrije Universiteit Medical Center, Amsterdam, The Netherlands*
- Roy M. Soetikno** / *Stanford University School of Medicine, Stanford, California, U.S.A.*
- Lorenzo Soletti** / *University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.*
- Ying C. Song** / *Organ Recovery Systems, Charleston, South Carolina, U.S.A.*
- Louis J. Soslowsky** / *University of Pennsylvania Health System, Philadelphia, Pennsylvania, U.S.A.*
- Reiner Spaethe** / *BioSurgery R & D, Baxter BioScience, Vienna, Austria.*
- John E. Speich** / *Virginia Commonwealth University, Richmond, Virginia, U.S.A.*
- Paulette Spencer** / *University of Missouri–Kansas City School of Dentistry, Kansas City, Missouri, U.S.A.*
- Michael J. Spivey** / *Cornell University, Ithaca, New York, U.S.A.*
- Naveen C. Srivastav** / *University of Alberta, Edmonton, Alberta, Canada*
- John J. Stankus** / *University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.*
- Robert M. Stern** / *The Pennsylvania State University, University Park, Pennsylvania, U.S.A.*
- George D. Stetten** / *University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.*
- Molly M. Stevens** / *Imperial College London, London, United Kingdom*
- Shih-Horng Su** / *AST Products, Inc., Billerica, Massachusetts, U.S.A.*
- Riitta Suuronen** / *Tampere University and University Hospital, Tampere, Finland.*
- Robert G. Svitek** / *University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.*
- M. Michael Swindle** / *Medical University of South Carolina, Charleston, South Carolina, U.S.A.*
- Michael Szycher** / *CardioTech International, Inc., Woburn, Massachusetts, U.S.A.*
- Maryam Tabrizian** / *McGill University, Montreal, Quebec, Canada*
- Nirit Tagger-Green** / *Rambam Medical Center and The Faculty of Medicine, Haifa, Israel*
- Wei Tan** / *University of Illinois at Urbana-Champaign, Urbana, Illinois, U.S.A.*
- John M. Tarbell** / *The Pennsylvania State University, University Park, Pennsylvania, and The City College of New York, New York, New York, U.S.A.*
- Daniel A. Tatosian** / *Cornell University, Ithaca, New York, U.S.A.*
- Bill Tawil** / *BioSurgery marketing, Baxter BioScience, Westlake Village, California, U.S.A.*
- Bruce Taylor** / *The University of Akron, Akron, Ohio, U.S.A.*
- Michael J. Taylor** / *Organ Recovery Systems, Charleston, South Carolina, U.S.A.*
- Todd A. Telemeco** / *Shenandoah University, Winchester, Virginia, U.S.A.*
- Louis Terracio** / *New York University, New York, New York, U.S.A.*
- Francesco Tassarolo** / *University of Trento, Trento, Italy*
- Eser Tüfekçi** / *Virginia Commonwealth University, Richmond, Virginia, U.S.A.*
- Bhavani P. Thampatty** / *University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.*
- Benjamin Thierry** / *McGill University, Montreal, Quebec, Canada*
- J. Brock Thomas** / *University of Texas at Austin, Austin, Texas, U.S.A.*

- T. Adam Thrasher** / *Toronto Rehabilitation Institute, Toronto, Ontario, Canada*
- Marc J. Torjman** / *Jefferson Medical College of Thomas Jefferson University, Philadelphia, Pennsylvania, U.S.A.*
- Eleftherios Tsiridis** / *The Whittington Hospital, University College London, London, and Institute of Orthopaedics, Royal National Orthopaedic Hospital, Stanmore, United Kingdom, and Boston University School of Medicine, Boston, MA, U.S.A.*
- Thomas Tu** / *Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, Massachusetts, U.S.A.*
- Toshimasa Uemura** / *National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Ibaraki, Japan*
- Egidijus Uzgiris** / *Rensselaer Polytechnic Institute, Troy, New York, U.S.A.*
- Joseph P. Vacanti** / *Massachusetts General Hospital, and Harvard Medical School, Boston, Massachusetts, U.S.A.*
- Bill Van Osdol** / *ALZA Corporation, Mountain View, California, U.S.A.*
- Ray Vanderby, Jr.** / *University of Wisconsin–Madison, Madison, Wisconsin, U.S.A.*
- Tomy Varghese** / *University of Wisconsin–Madison, Madison, Wisconsin, U.S.A.*
- Marcia Vasquez-Lee** / *Mississippi State University, Mississippi State, Mississippi, U.S.A.*
- Brent Vernon** / *Arizona State University, Tempe, Arizona, U.S.A.*
- Michel Vert** / *Université Montpellier 1, Montpellier Cedex 5, France*
- Ivan Vesely** / *The Cleveland Clinic Foundation, Cleveland, Ohio, U.S.A.*
- Vibeke Videm** / *The Norwegian University of Science and Technology, and Trondheim University Hospital, Trondheim, Norway*
- Kent S. Vilendrer** / *Tissue Growth Technologies, Inc., Minnetonka, Minnesota, U.S.A.*
- Jürgen Vogt** / *CIBA Vision Corp., Duluth, Georgia, U.S.A.*
- Nicola Volpi** / *Biological Chemistry Section, University of Modena and Reggio Emilia, Modena, Modena, Italy*
- David Vorp** / *University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.*
- David A. Vorp** / *McGowan Institute for Regenerative Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.*
- William R. Wagner** / *University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.*
- Gordon Wallace** / *University of Wollongong, Wollongong, New South Wales, Australia*
- B. H. Walpoth** / *University Hospital, Geneva, Switzerland*
- John R. Walsh** / *Organ Recovery Systems, Charleston, South Carolina, U.S.A.*
- Chia-Wei Wang** / *University of Michigan, Ann Arbor, Michigan, U.S.A.*
- Dong-an Wang** / *Johns Hopkins University, Baltimore, Maryland, U.S.A.*
- Paul Wang** / *Stanford University Medical Center, Stanford, California, U.S.A.*
- Yadong Wang** / *Georgia Institute of Technology, Atlanta, Georgia, U.S.A.*
- Yao Wang** / *Pfizer, Inc., Ann Arbor, Michigan, U.S.A.*
- Zhengdong Wang** / *University of Rochester School of Medicine, Rochester, New York, U.S.A.*
- Thomas D. Wang** / *Stanford University School of Medicine, Stanford, California, U.S.A.*
- James H.-C. Wang** / *University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.*
- Junji Watanabe** / *The University of Tokyo, Tokyo, Bunkyo-ku, Japan*
- Jennifer S. Wayne** / *Virginia Commonwealth University, Richmond, Virginia, U.S.A.*
- Thomas J. Webster** / *Purdue University, West Lafayette, Indiana, U.S.A.*
- Eric R. Weeks** / *Emory University, Atlanta, Georgia, U.S.A.*
- Michael Wegner** / *Arizona State University, Tempe, Arizona, U.S.A.*
- Xiaochao Wei** / *University of Rochester, Rochester, New York, USA*
- E. J. Weinberg** / *Charles Stark Draper Laboratory, and Massachusetts Institute of Technology, Cambridge, Massachusetts, U.S.A.*
- Barbara Weiser** / *University of Regensburg, Regensburg, Germany*

- Steve T. Wereley** / *Purdue University, West Lafayette, Indiana, U.S.A.*
- Jerome A. Werkmeister** / *Molecular and Health Sciences, Clayton South, Victoria, Australia*
- Jennifer L. West** / *Rice University, Houston, Texas, U.S.A.*
- Richard L. White, Jr.** / *Carolinas Medical Center, Charlotte, North Carolina, U.S.A.*
- David L. Williams** / *Mississippi State University, Mississippi State, Mississippi, U.S.A.*
- Rachel L. Williams** / *University of Liverpool, Liverpool, United Kingdom*
- Gary E. Wnek** / *Case Western Reserve, Cleveland, Ohio, U.S.A.*
- Kimberly A. Woodhouse** / *University of Toronto, Toronto, Ontario, Canada*
- Debra D. Wright** / *Michigan Technological University, Houghton, Michigan, U.S.A.*
- Daniel E. Wueste** / *Clemson University, Clemson, South Carolina, U.S.A.*
- J. H. David Wu** / *University of Rochester, Rochester, New York, U.S.A.*
- Benjamin M. Wu** / *University of California, Los Angeles, California, U.S.A.*
- Kenneth J. Wynne** / *Virginia Commonwealth University, Richmond, Virginia, U.S.A.*
- Lian Xue** / *Loyola University Medical Center, Illinois, U.S.A.*
- Hu Yang** / *Virginia Commonwealth University, Richmond, Virginia, U.S.A.*
- Yunzhi Yang** / *University of Tennessee Health Science Center, Memphis, Tennessee, U.S.A.*
- Wentao Yan** / *New York University, New York, New York, U.S.A.*
- Ioannis V. Yannas** / *Massachusetts Institute of Technology, Cambridge, Massachusetts, U.S.A.*
- Juming Yao** / *Tokyo University of Agriculture and Technology, Koganei, Tokyo, Japan*
- Martin L. Yarmush** / *Massachusetts General Hospital, Harvard Medical School, and the Shriners Hospital for Children, Boston, Massachusetts, U.S.A.*
- Ajit Yoganathan** / *Georgia Institute of Technology, Atlanta, Georgia, U.S.A.*
- Ajit P. Yoganathan** / *Georgia Institute of Technology, Atlanta, Georgia, U.S.A.*
- Michael J. Yost** / *University of South Carolina, Columbia, South Carolina, U.S.A.*
- Chun Yuan** / *University of Washington, Seattle, Washington, U.S.A.*
- Peng Yuan** / *National Institutes of Health Clinical Center, Bethesda, Maryland, U.S.A.*
- Deniz Yucel** / *Middle East Technical University, Ankara, Turkey*
- James A. Zagzebski** / *The University of Wisconsin–Madison, Madison, Wisconsin, U.S.A.*
- Pierre-Yves Zambelli** / *Hôpital Orthopédique de la Suisse Romande, Lausanne, Switzerland*
- Conrad M. Zapanta** / *Penn State College of Medicine, Hershey, Pennsylvania, U.S.A.*
- Lyle D. Zardiackas** / *University of Mississippi Medical Center, Jackson, Mississippi, U.S.A.*
- Jing Zhang** / *University of Texas at Austin, Austin, Texas, U.S.A.*
- Peter Zilla** / *University of Cape Town Medical School, and Groote Schuur Hospital, Cape Town, South Africa*
- Wolfram-Hubertus Zimmermann** / *University-Hospital Hamburg-Eppendorf, Hamburg, Germany*

Contents

Volume 1

<i>Preface</i>	xxix
Adhesives <i>J. Brock Thomas, Nicholas A. Peppas</i>	1
Allografts <i>Kirby S. Black, Patti E. Dawson</i>	8
Alumina <i>Jae Sung Lee</i>	17
Angiogenesis Inhibitors <i>Shaker A. Mousa, Ahmed S. Mousa</i>	26
Animal Models for Bone Tissue Engineering <i>Mark P. Mooney, Michael I. Siegel</i>	33
Animal Surrogate Systems for Toxicity Testing <i>Michael L. Shuler, Gretchen J. McAuliffe, Daniel A. Tatosian</i>	51
Antimicrobial/Antibiotic (Infection Resistance) Materials <i>Franck Furno, Roger Bayston</i>	61
Antimineralization Treatment <i>Bart Meuris, Willem Flameng</i>	71
Articular Cartilage Biomechanics <i>Jennifer S. Wayne</i>	85
Artificial Heart Fluid Dynamics: Positive Displacement Pumps <i>Steven Deutsch, John M. Tarbell, Keefe B. Manning, Gerson Rosenberg, Arnold A. Fontaine</i>	95
Artificial Muscles <i>Mohsen Shahinpoor</i>	108
Artificial Neural Networks: An Overview <i>Evangelia Micheli-Tzanakou</i>	119
Bioactive Glass <i>Marivalda M. Pereira, Larry L. Hench</i>	133
Bioactive Materials and Scaffolds for Tissue Engineering <i>Qi-Zhi Chen, Aldo R. Boccaccini</i>	142
Bioadhesion <i>John D. Smart</i>	152
Biocatalysis <i>Claire F. Komives, Ruizhen Rachel Chen</i>	162
Biocompatibility Testing <i>Joel D. Bumgardner, Marcia Vasquez-Lee, Keertik S. Fulzele, Daniel H. Smith, Kellye D. Branch, Seth I. Christian, David L. Williams</i>	169
Biocomposites <i>Huinan Liu, Grace E. Park, Thomas J. Webster</i>	179
Biodegradable Polymers: An Overview <i>Chih-Chang Chu</i>	195
Biofilms <i>Clive Edwards</i>	207
Biofunctional Polymers <i>Jennifer L. West</i>	218
Biologic and Synthetic Apatites <i>Besim Ben-Nissan, Racquel LeGeros</i>	225
Biological Adhesives from Nature <i>Lloyd D. Graham</i>	236
Biological Effects of Radiofrequency Electromagnetic Field <i>Henry Lai</i>	254
Biomaterials Immune Response <i>Sandra C. P. Cachinho, John A. Hunt</i>	262
Biomaterials: Protein–Surface Interactions <i>Robert A. Latour, Jr.</i>	270
Biomedical Ethics <i>Daniel E. Wueste</i>	285
Biomimetic Materials <i>Jun-ichi Kikuchi, Atsushi Ikeda, Mineo Hashizume</i>	294
Biopotential Amplifiers <i>Albert Lozano-Nieto</i>	301
Bioreactors <i>Lito C. Mejia, Kent S. Vilendrer</i>	310
Biorubber/Poly(Glycerol Sebacate) <i>Yadong Wang</i>	327
Biosensors <i>Jerome S. Schultz</i>	335
Biosynthesis and Applications of Alginates <i>Bernd H. A. Rehm</i>	350
Biphasic Calcium Phosphate (BCP) Bioceramics: Chemical, Physical, and Biological Properties <i>Guy Daculsi, Racquel LeGeros</i>	359
Blood-Material Interactions <i>Stephen R. Hanson</i>	367
Blood Purification <i>Yukihiko Nosé, Hisashi Okubo</i>	378

Blood Vessel Mechanics <i>Ruth J. Okamoto</i>	392
Bone Cement <i>Julie Hasenwinkel</i>	403
Bone Fracture Fixation <i>Eleftherios Tsiridis, Constantin Schizas</i>	413
Bone-Implant Interface <i>D. A. Puleo</i>	423
Bone Plates and Screws, Bioabsorbable <i>Riitta Suuronen, Christian Lindqvist</i>	432
Bone Remodeling <i>Elisabeth H. Burger, Jenneke Klein-Nulend, Theo H. Smit</i>	439
Burn Dressing <i>Andrew Burd</i>	448
Calcium Phosphate Ceramics: New Generation Produced in Japan <i>Atsuo Ito, Hajime Ohgushi</i>	461
Cardiac Assist Devices <i>Keefe B. Manning, Conrad M. Zapanta, John M. Tarbell</i>	470
Cardiac Bioelectricity <i>Clive M. Baumgarten</i>	480
Cardiac Catheters <i>Tommy L. Miller, Juana Maria Alfaro, Deepak Banerjee, Benjamin Lee, Anthony J. Minisi</i>	491
Cardiac Elastography, Full-Field Development <i>Tomy Varghese, Q. Chen, P. Rahko, James A. Zagzebski</i>	506
Cardiac Output, Basic Measurement Principles <i>Leslie A. Geddes</i>	514
Cardiac Pacemaker <i>Rakesh K. Pai, Fred M. Kusumoto</i>	532
Cardiac Patch Engineering <i>Michael J. Yost, Robert L. Price, David G. Simpson, Wentao Yan, Louis Terracio</i>	542
Cartilage Regeneration <i>Nic D. Leipzig, Kyriacos A. Athanasiou</i>	552
Cell Culture Assays <i>P. J. Doherty</i>	561
Cell-Material Interaction <i>Aditya Chaubey, Karen J. L. Burg</i>	568
Centrifugal Blood Pumps <i>Gerald E. Miller</i>	574
Ceramics <i>Robert H. Doremus</i>	585
Ceramics in Dentistry <i>Nicolas M. Jedyndakiewicz</i>	593
Chitosan <i>M. N. V. Ravi Kumar, S. M. Hudson</i>	604
Chondroitin Sulfates <i>Nicola Volpi</i>	618
Collagen <i>Matthew J. Beckman, Kelly J. Shields, Robert F. Diegelmann</i>	628
Collagen Fixation <i>Thomas J. Koob</i>	639
Collagen Processing <i>Michael J. Yost, Louis Terracio, Robert L. Price</i>	652
Composites <i>Chia-Wei Wang, Ann Marie Sastry</i>	659
Compression of Digital Biomedical Signals <i>Tapio Grönfors, Martti Juhola</i>	677
Computer-Assisted Surgery <i>Thomas L. Husted, Timothy J. Broderick</i>	687
Conductive Polymers <i>Gordon Wallace, Leon Kane-Maguire</i>	695
Confocal Microscopy <i>Denis Semwogerere, Eric R. Weeks</i>	705
Contact Lenses: Silicone Hydrogels <i>Paul C. Nicolson, Jürgen Vogt</i>	715
Control of Movement <i>Dejan B. Popović</i>	729
Controlled Release <i>Brent Vernon, Michael Wegner</i>	739
Corneal Implants <i>Traian V. Chirila, Celia R. Hicks</i>	747
Corrosion <i>Rachel L. Williams</i>	754
Cryopreservation of Living Cells <i>Dayong Gao, J. K. Critser</i>	762
Defibrillators <i>Paul Wang</i>	772

Volume 2

Degradable Polymer Composites <i>Debra D. Wright</i>	779
Dendrimers for Drug Delivery <i>Hu Yang, W. John Kao</i>	789
Dental Biomechanics <i>A. N. Natali, P. G. Pavan, E. L. Carniel</i>	799
Dental Implants <i>Nirit Tagger-Green, Eli E. Machtei</i>	808
Diamond and Diamond-Like Carbons <i>James H. Arps, Geoff Dearnaley</i>	822

Dielectric Properties of Tissues / <i>Kenneth R. Foster</i>	832
Digital Biomedical Signal Acquisition and Processing: Basic Topics / <i>Martin Oswaldo Mendez Garcia, Luca T. Mainardi</i>	843
Digital Biomedical Signal Processing: Advanced Topics / <i>Luca T. Mainardi</i>	854
Distal Protection Devices / <i>Thomas Tu, Roger Laham</i>	870
Drug Delivery, Controlled / <i>M. N. V. Ravi Kumar, A. J. Domb</i>	880
Elastin / <i>William T. Brinkman, Karthik Nagapudi, Elliot L. Chaikof</i>	891
Elastomers, Biodegradable / <i>John J. Stankus, Jianjun Guan, William R. Wagner</i>	897
Electric Cell-Substrate Impedance Sensing / <i>Philippe Lam</i>	908
Electrical Impedance Imaging / <i>Keith D. Paulsen, Alex Hartov</i>	915
Electroactive Polymeric Materials / <i>V. Prasad Shastri</i>	926
Electrocardiography / <i>Soumyadipta Acharya, Bruce Taylor</i>	936
Electrogastrography / <i>Robert M. Stern, Kenneth L. Koch</i>	947
Electromyography / <i>Didier Gamet, Odette Fokapu</i>	956
Electron Microscopy / <i>Robert L. Price, Jeff Davis, Michael J. Yost, Richard L. Goodwin, Louis Terracio</i>	968
Electrospinning / <i>Darrell H. Reneker, Haoqing Hou</i>	978
Endoscopy / <i>Jonathan T. C. Liu, Tonya Kaltenbach, Thomas D. Wang, Roy M. Soetikno</i>	986
Ergonomics / <i>Brian N. Craig</i>	999
Excitable Tissue, Electrical Stimulation Of / <i>Dominique M. Durand</i>	1009
Extracellular Matrix Scaffolds / <i>Stephen F. Badylak</i>	1021
Eye Tracking: Characteristics and Methods / <i>Daniel C. Richardson, Michael J. Spivey</i>	1028
Eye Tracking: Research Areas and Applications / <i>Daniel C. Richardson, Michael J. Spivey</i>	1033
Fabrics / <i>Frank K. Ko</i>	1043
Fiber Optic Biosensors / <i>Ilko K. Ilev, Israel Gannot</i>	1062
Fibrin / <i>Sam L. Helgerson, Thomas Seelich, James P. DiOrio, Bill Tawil, Katharina Bittner, Reiner Spaethe</i>	1072
Fibrin Sealants / <i>Marcus E. Carr, Jr., Erika J. Martin, Heather Ambrose</i>	1080
Finite Element Analysis / <i>Jennifer S. Wayne</i>	1091
Flow Cytometry / <i>J. Paul Robinson</i>	1100
Foreign Body Response / <i>John Hunt</i>	1111
Functional MRI: Applications / <i>Hui Mao</i>	1119
Gait Analysis / <i>Richard Baker</i>	1129
Gene-Activated Matrix / <i>Neil Davies</i>	1142
Gene Therapy / <i>Viraj P. Mane, William McCormack, Brendan Lee</i>	1150
Glucose-Responsive Hydrogels / <i>Nicholas A. Peppas, Christie D. Bures</i>	1163
Glucose Sensors / <i>Jeffrey I. Joseph, Marc J. Torjman</i>	1174
Glues / <i>Shojiro Matsuda, Yoshito Ikada</i>	1185
Glycosaminoglycans and Proteoglycans / <i>Jan C. Schagemann, Hsi-Wei Chung, Eike H. Mrosek, Shawn W. O'Driscoll, Gregory G. Reinholz</i>	1195
Gradient Surfaces: Preparation, Characterization, and Interactions with Biological Species / <i>Hai Bang Lee, Moon Suk Kim, Gilson Khang, Jin Ho Lee</i>	1210
Hard Tissue: Biomaterial Interactions / <i>Feza Korkusuz, Petek Korkusuz</i>	1229
Hard Tissue Elastic Properties / <i>J. Lawrence Katz</i>	1236
Healing of Bone and Connective Tissues / <i>William L. Murphy, Kelley Grorud, Ray Vanderby, Jr.</i>	1249
Hearing Aids / <i>Ian C. Bruce</i>	1263
Hearing Mechanisms / <i>Martin L. Lenhardt</i>	1271
Heart Biomechanics / <i>Jeffrey W. Holmes</i>	1280

Heart Valve, Bioprosthetic	Peter Zilla, Paul Human, Deon Bezuidenhout	1291
Heart Valve Failure, Bioprosthetic	Peter Zilla, Paul Human, Deon Bezuidenhout	1306
Heart Valve Mechanics	Ajit Yoganathan, Anna M. Fallon, Jorge H. Jimenez	1317
Heart Valves, Mechanical	Ajit P. Yoganathan, Zhaoming He, Hwa-Liang Leo, Anna Fallon	1329
Hematopoietic Stem Cells and Assays	Athanassios Mantalaris, Nicki Panoskaltzis, J. H. David Wu	1338
Hemocompatible Materials	Agnese Magnani, Federica M. Piras	1345
Hemodynamics, Macrocirculatory	Baruch B. Lieber, Chander Sadasivan	1356
Heparin-Binding to Improve Biocompatibility	Rolf Larsson	1368
Hip Biomechanics	Farid Amirouche, Mark Gonzalez, Luke Aram	1377
Hip Joint: Overuse Injuries	Marko Pećina, Ivan Bojanić Alan Ivković	1393
Histogenesis	Julie Glowacki	1403
Host Reactions	Kristin J. Pawlowski	1411
Hyaluronan	Jennie Baier Leach, Christine E. Schmidt	1421
Hybrid Vascular Prostheses	Jerome A. Werkmeister, Glenn A. Edwards, John A. M. Ramshaw	1432
Hydrogels	Junji Watanabe, Yoshihiro Kiritoshi, Kwang Woo Nam, Kazuhiko Ishihara	1439
Hydroxyapatite	Alexandra Elizabeth Porter, Nelesh Patel, Serena Best	1451
Hydroxyapatite Coatings	Yunzhi Yang, Kazuhisa Bessho, Kyo-Han Kim, Sang-won Park, Joo L. Ong	1464
Implant, Total Hip	Nadim James Hallab, Joshua Jacobs	1470
Implanted Medical Devices, Power Sources For	Erik R. Scott, Craig L. Schmidt, Paul M. Skarstad	1484
Infection of Medical Devices	Lori L. Burrows, Antoine E. Houry	1497
Inner Ear Implants	Graeme M. Clark	1507
Insulin Delivery	Robert S. Parker	1514
Integrins	David G. Simpson	1524
Intellectual Property Management	Michael Blakeney	1540
Interpenetrating Polymeric Networks	Nicholas A. Peppas, Jing Zhang	1549

Volume 3

Intravascular Ultrasound	Charles D. Choi, Yao Wang, Javier de Ana, Matthew O'Donnell	1559
In Vivo Cardiovascular Models	B. H. Walpoth, O. M. Hess	1566
Kidneys, Artificial	Churn K. Poh, Junfeng Lu, William R. Clark, Dayong Gao	1576
Knee Joint: Overuse Injuries	Ivan Bojanić, Alan Ivković, Marko Pećina	1597
Knee Joint Replacement	M. J. Cross, E. N. Parish, S. Machan	1607
Laboratory Animals: Ethics and Regulations for Care and Use	Kathy E. Laber, M. Michael Swindle	1616
Laser-Tissue Interaction	E. Duco Jansen	1623
Laser Transmyocardial Revascularization	James W. Jones, Nancy A. Crigger, Bruce W. Richman	1632
Ligament Repair: Animal Models	Robert C. Bray, Catherine A. Leonard, Kevin R. Forrester, Paul T. Salo	1640
Liver, Bio-Artificial	François Berthiaume, Christina Chan, Martin L. Yarmush	1649
Lung, Artificial: Basic Principles and Current Applications	William J. Federspiel, Kristie A. Henschir	1661
Lung, Artificial: Current Research and Future Directions	William J. Federspiel, Robert G. Svitek	1673
Lung Surfactants	Robert H. Notter, Zhengdong Wang	1683

Magnetic and Electrophoretic Cell Separation / <i>S. Chethana, Ganapathi Patil, K. S. M. S. Raghavarao</i>	1695
Magnetic Nanoparticles: Structure and Bioapplications / <i>Yudhisthira Sahoo, Edward P. Furlani, Earl J. Bergey</i>	1705
Magnetic Resonance Imaging (MRI) / <i>Simon J. Doran</i>	1713
Magnetic Resonance Imaging in Temperature Measurement / <i>Nadine Barrie Smith</i>	1728
Magnetic Resonance Microscopy / <i>Egidijus Uzgiris</i>	1738
Mass Transfer in Tissue Function: Roles / <i>E. N. Lightfoot, Karen A. Duca,</i>	1749
Matrix Metalloproteinases / <i>Todd A. Telemeco, James R. Bowman, III, David G. Simpson</i>	1763
Mechanical Circulatory Support / <i>Eric J. Okum, Vigneshwar Kasirajan, Robert S. D. Higgins</i>	1775
Mechano-Regulation of Fibroblast Function / <i>James H.-C. Wang, Bhavani P. Thampatty</i>	1783
Medical Imaging, 3-D / <i>Oskar Skrinjar, Paul J. Benkeser</i>	1794
Medical Robotics / <i>John E. Speich, Jacob Rosen</i>	1804
Melt Spinning / <i>Eija Marjut Pirhonen, Ville Ellä</i>	1816
Metallic Orthopedic Devices: Failure Analysis / <i>Lyle D. Zardiackas</i>	1824
Microcirculatory Oxygen Transport / <i>Roland N. Pittman</i>	1842
Microcomputed Tomography and Its Applications / <i>Thomas Dufresne, Paula Chmielewski, Babul Borah, Andres Laib</i>	1850
Microelectromechanical Systems (MEMS) Manufacturing / <i>Marc Madou</i>	1860
Micro-Particle Image Velocimetry in Biomedical Applications / <i>Jinhua Cao, Steve T. Wereley</i>	1873
Microporous Materials / <i>Julian R. Jones, Larry L. Hench</i>	1885
Modeling, Biomedical / <i>Andrew J. Pullan, Nicolas P. Smith</i>	1894
Nanoscale Neurosurgery / <i>William C. Broaddus, Peter J. Haar, George T. Gillies</i>	1903
Nerve Guides / <i>Burkhard Schlosshauer, Martin Lietz</i>	1911
Neuroprostheses / <i>Milos R. Popovic, T. Adam Thrasher</i>	1924
Neutrophil-Biomaterial Interactions / <i>Vibeke Videm</i>	1934
Nickel and Nickel Alloys / <i>Rachel L. Williams</i>	1942
Nitric Oxide / <i>Matthew J. Bizzarro, Vineet Bhandari</i>	1951
Nuclear Magnetic Resonance Spectroscopy / <i>Mari A. Smith, Peter B. Barker</i>	1960
Ocular Implants / <i>Andrew W. Lloyd, Susan Sandeman, Richard G. A. Faragher, Stephen P. Denyer</i>	1974
Ocular Implants for Drug Delivery / <i>Susan S. Lee, Peng Yuan, Michael R. Robinson</i>	1981
Optical Coherence Tomography / <i>Xingde Li, James G. Fujimoto</i>	1996
Optical Detection of Cancers / <i>Xin-Hua Hu, Jun Qing Lu</i>	2009
Optical Mapping / <i>Igor R. Efimov, Vladimir P. Nikolski, Gil Bub</i>	2019
Optics, Biomedical / <i>Sergio Fantini</i>	2029
Organic and Inorganic Matrices / <i>Adele L. Boskey</i>	2039
Orthodontic Wires / <i>Eser Tüfekçi, Steven J. Lindauer</i>	2054
Orthopedic Biomaterials / <i>J. Lawrence Katz, Catherine G. Ambrose, Carl McMillin, Paulette Spencer</i>	2061
Orthopedic Fixation Devices / <i>Jill S. Kawalec</i>	2073
Osteoclastic Bone Resorption / <i>Samuel C. Ramage, Mengnai Li, Matthew J. Beckman</i>	2081
Osteogenic Progenitor Cells of Bone / <i>Richard O. C. Oreffo</i>	2091
Osteoinductive Substances and Materials / <i>Toshimasa Uemura, Hiroko Kojima, Takashi Saito</i>	2102
Particulate Matter and Host Reactions / <i>Edward M. Greenfield</i>	2110
Peripheral Nerve Repair and Regeneration: Historical Perspective / <i>E. E. Sabelman, W. Chang, S. Nguyen, J. M. Rosen</i>	2118
Phosphorylcholine (PC) Technology / <i>Andrew L. Lewis</i>	2136
Photopolymerization / <i>Dong-an Wang, Jennifer H. Elisseeff</i>	2150

Piezoelectric Materials Zoubeida Ounaies	2163
Plasma Modification of Materials Ricky K. Y. Fu, Paul Kim Ho Chu	2174
Polyamides (Synthetic and Natural) Naveen C. Srivastav, Babita Agrawal, Rakesh Kumar	2186
Polycaprolactone J. E. Gough, P. Christian	2197
Poly(carbonate)urethanes Michael Szycher	2208
Polyelectrolyte Multilayers Darwin R. Reyes, Laurie E. Locascio, Michael Gaitan	2217
Polyepoxy Compound Fixation Yasuharu Noishiki, Teruo Miyata	2231
Poly(Glycolic Acid) Eugene D. Boland, Gary E. Wnek, Gary L. Bowlin	2241
Poly(lactic acid)s Michel Vert	2249
Poly(lactic-Co-Glycolic Acid (PLGA)) Konstantinos Avgoustakis	2259
Polymer Foams V. Prasad Shastri, Molly M. Stevens	2270
Polymers Gary E. Wnek	2275
Polymers Used in Tissue Engineering Vasif Hasirci, Deniz Yucel	2282
Polytetrafluoroethylene Robert R. Pavlis	2300
Polyurethane Biomaterials Joanna D. Fromstein, Kimberly A. Woodhouse	2304
Positron Emission Tomography (PET) Anna-Liisa Brownell	2314
Protein Adsorption Modeling Alan H. Goldstein, Alastair N. Cormack	2326

Volume 4

Protein and Cell Signaling with Biomaterials: Influence of Surfactants Solomon W. Leung, James C. K. Lai	2337
Protein and Cell Signaling with Biomaterials: Interfacial Transport James C. K. Lai, Solomon W. Leung	2346
Protein Delivery Systems Nicholas A. Peppas, Cristina Donini	2357
Pyrolytic Carbon Robert B. More, George Sines, Ling Ma, Jack C. Bokros	2370
Rapid Prototyping Shaochen Chen, Carlos A. Aguilar, Yi Lu	2382
Real Time Tomographic Reflection Vikram S. Chib, George D. Stetten	2391
Reconstructive Biomaterials Mutaz B. Habal	2395
Rotary Cell Culture Systems Stephen S. Navran	2402
Self-Assembled Monolayers Koichi Kato	2409
Shape Memory Metals Marie Haidopoulos, Fatiha El Feninat, Diego Mantovani	2418
Silicones Kenneth J. Wynne, James M. Lambert	2427
Silks Juming Yao, Tetsuo Asakura	2442
Small-Angle Light Scattering Methods for Soft Connective Tissue Structural Analysis Michael S. Sacks	2450
Spine Biomechanics Clayton Adam	2464
Stem Cells Andrea Liebmman-Vinson, Sharon C. Presnell	2473
Stent Grafts, Endovascular Timothy A. M. Chuter	2480
Stents Benjamin Thierry, Luc Bilodeau, Maryam Tabrizian	2492
Sterilization of Biomedical Materials Francesco Tessarolo, Giandomenico Nollo	2501
Stress Fractures Marko Pećina, Ivan Bojanić, Alan Ivković	2511
Supercritical Fluid Processing Ihtesham U. Rehman, Jawwad A. Darr, Alireza Moshaverinia	2522
Surface Coatings Shih-Horng Su, Suzanne Conroy, Tung-Liang Lin, Min-Shyan Sheu, Ih-Houng Loh	2531
Surface Modification Sally L. McArthur, Keith M. McLean	2540
Surface Topography and Cell Behavior Douglas W. Hamilton, Salem Ghrebi, Hugh Kim, Babak Chehroudi, Donald M. Brunette	2551
Sutures C. C. Chu	2562

Telemedicine Brett M. Harnett, Charles R. Doarn, Ronald C. Merrell	2579
Tendons and Ligaments, Mechanical Testing Of Richard E. Debski, Susan M. Moore, Eric J. Rainis	2588
Thrombosis Mukul S. Goel, Scott L. Diamond	2601
Tinnitus Devices Martin L. Lenhardt	2610
Tissue Engineering François A. Auger, Lucie Germain	2620
Tissue Engineering: AC Electrokinetics Gerard H. Markx, Anne-Marie Buckle	2628
Tissue Engineering of Bladder Anthony Atala	2636
Tissue Engineering of Blood Vessel Douglas Hamilton, David Vorp	2643
Tissue Engineering of Bone L. Di-Silvio, N. Gurav, Eleftherios Tsiridis	2652
Tissue Engineering of Bone Marrow Athanassios Mantalaris, Nicki Panoskaltzis, J. H. David Wu	2660
Tissue Engineering of Bone Marrow, Culture Systems Athanassios Mantalaris, Nicki Panoskaltzis, J. H. David Wu	2670
Tissue Engineering of Breast Cheryl T. Gomillion, Cheryl A. Parzel, Richard L. White, Jr., Karen J.L. Burg	2678
Tissue Engineering of Cardiac Muscle Wolfram-Hubertus Zimmermann, Thomas Eschenhagen	2686
Tissue Engineering of Cartilage J. N. Amritha De Croos, Rita A. Kandel	2694
Tissue Engineering of Cornea Lucie Germain, Claude J. Giasson, Patrick Carrier, Sylvain L. Guérin, Christian Salesse, François A. Auger	2707
Tissue Engineering of the Esophagus Stephen F. Badylak, Thomas W. Gilbert, Alejandro Nieponice	2719
Tissue Engineering, Fat Barbara Weiser, Markus Neubauer, Achim Göpferich, Torsten Blunk	2725
Tissue Engineering of Heart Valves Ivan Vesely	2737
Tissue Engineering of the Intestine James C. Y. Dunn, Benjamin M. Wu	2751
Tissue Engineering of Ligament E. Gentleman, G. A. Livesay, K. C. Dee, E. A. Nauman	2758
Tissue Engineering of Liver Alexander B. G. Sevy, Katherine M. Kulig, Joseph P. Vacanti	2769
Tissue Engineering, Microscale Wei Tan, Tejal Desai	2779
Tissue Engineering of Microvascular Networks J. T. Borenstein, E. J. Weinberg, M. R. Kaazempur-Mofrad, Joseph P. Vacanti	2792
Tissue Engineering of Pancreas Hiroo Iwata	2802
Tissue Engineering of Peripheral Nerve Archit B. Sanghvi, Joy L. Murray, Christine E. Schmidt	2811
Tissue Engineering of Rotator Cuff Tendons Jonathan P. Van Kleunen, Louis J. Soslowsky, David L. Glaser	2820
Tissue Engineering Scaffolds Eugene D. Boland, Paul G. Espy, Gary L. Bowlin	2828
Tissue Engineering of Skeletal Muscle Christopher Keen, Daniel Newton, Todd A. Telemeco, Gary L. Bowlin, Gary E. Wnek, Clive M. Baumgarten, David G. Simpson	2838
Tissue Engineering of Skin Ricardo R. Brau, Ioannis V. Yannas	2851
Tissue Engineering of the Small Intestine Simon M. Gabe, Richard M. Day, Aldo Boccaccini	2860
Tissue Engineering of Tendons Dominique P. Pioletti, Olivier Siegrist, Pierre-Yves Zambelli	2871
Titanium and Its Alloys Mitsuo Niinomi, Tomokazu Hattori, Toshihiro Kasuga, Hisao Fukui	2876
Transdermal Drug Delivery Brad Phipps, Michel Cormier, Bob Gale, Bill van Osdol, Jay Audett, Rama Padmanabhan, Peter Daddona	2893
Transgenic and Gene-Targeted Organs Brenda M. Ogle, Jeffrey L. Platt	2906
Transparent Ceramics: Bioactive Materials and Tissue Engineering Motohiro Hirose, Koji Ioku, Noriko Kotobuki, Daisuke Kawagoe, Hajime Ohgushi	2916
Ultra-High Molecular Weight Polyethylene (UHMWPE) Kwideok Park, Gladius Lewis, Joon B. Park	2925
Ultrasound and Tissue Interaction K. Kirk Shung	2933

Ultrasound Doppler / <i>Ding-Yu Fei, James A. Arrowood</i>	2942
Ultrasound Therapy, Bone Healing / <i>Özgür Erdoğan</i>	2951
Urinary Prostheses / <i>Daniel Eberli, Anthony Atala</i>	2961
Vascular Fluid Dynamics / <i>Stanley A. Berger</i>	2971
Vascular Grafts / <i>Deon Bezuidenhout, Peter Zilla</i>	2985
Vascular Grafts: Development Strategies / <i>James Laredo, Lian Xue, Howard P. Greisler</i>	2997
Vascular Grafts: Host Interactions / <i>James Laredo, Lian Xue, Howard P. Greisler</i>	3006
Vascular Imaging / <i>Tobias Saam, William S. Kerwin, Chun Yuan</i>	3012
Vascularization of Engineered Constructs / <i>Zerina Lokmic, Rejhan Idrizi, Aurora Messina, Kenneth R. Knight, Wayne Morrison, Geraldine Mitchell</i>	3022
Vascular Tissue Engineering / <i>Alejandro Nieponice, Timothy M. Maul, Lorenzo Soletti, David A. Vorp</i>	3033
Vitrification / <i>Kelvin G. M. Brockbank, John R. Walsh, Ying C. Song, Michael J. Taylor</i>	3046
Wear Debris / <i>John Fisher, Eileen Ingham</i>	3058
Wear Debris, Bone Resorption Animal / <i>Xiaochao Wei, Regis J. O'Keefe, Edward M. Schwarz</i>	3067
Wear Debris from Joint Arthroplasties, Biological Effects Of / <i>Ting Ma, Stuart B. Goodman</i>	3076
Wheeled Mobility: Wheelchairs and Personal Transportation / <i>Jon L. Pearlman, Rory A. Cooper</i>	3083
Xenografts / <i>Brenda M. Ogle, Jeffrey L. Platt</i>	3093
Zirconia Ceramics / <i>Mariano A. Polack</i>	3104

Preface

The field of biomedical engineering continues to apply engineering principles to analyze and solve problems in life sciences and medicine, leading to improvements in the quality of life. Biomaterials represent a central theme in a majority of the problems encountered. These fields have evolved into a very interdisciplinary arena building on traditional engineering principles that bridge advances in the areas of materials science, life sciences, nanotechnology, and cell biology, to name a few. This trend of interdisciplinary research to solve the most challenging yet compelling medical problems has been embraced in the field and is leading to the betterment of human health. It is evident that the fields of biomaterials and biomedical engineering are continually changing due to the rapid creation and advancement in technology in more traditional areas as well as rapidly developing areas (e.g., tissue engineering).

Because the scope of these medical problems is both vast and dynamic, there is a clear need for a source of technical information that has broad coverage, is current, is written at a level accessible to nonexperts, and has the ability to be updated on a regular basis. The *Encyclopedia of Biomaterials and Biomedical Engineering* (EBBE) was born with these considerations in mind in 2001, and its first edition appeared in 2004. Due to the success of that endeavor and the privilege of working with experts worldwide in preparing the second edition of the EBBE, it has blossomed into an expansive array of 298 entries. It is our hope that we have continued to present the material in such a fashion that it conveys a balanced overview to help stimulate further advancements in all areas of biomaterials and biomedical engineering.

We also believe that many thanks are in order for the efforts that led to the successful first edition and the generation of the second edition of the EBBE. We thank all the authors who have enthusiastically provided their time and expertise, and we also thank the editorial board for all of their various contributions during the development of the EBBE. Finally, we are grateful to Informa Healthcare for acquiring the EBBE, making it a priority, and providing the excellent staff needed to ensure its continued success.

It is our continued hope that the EBBE will serve as an important and continual resource for the many individuals whose work touches, and is touched by, biomaterials and biomedical engineering in order to further stimulate, create, and deliver improvements in quality of life.

Gary E. Wnek
Gary L. Bowlin

Encyclopedia of
**BIOMATERIALS AND
BIOMEDICAL ENGINEERING**

J. Brock Thomas

Department of Chemical Engineering, University of Texas at Austin, Austin, Texas, U.S.A.

Nicholas A. Peppas

*Departments of Chemical and Biomedical Engineering and Division of Pharmaceutics,
University of Texas at Austin, Austin, Texas, U.S.A.*

INTRODUCTION

An adhesive broadly refers to any substance that possesses the ability to hold two other entities together. Historically, adhesives and bioadhesives have referred to distinctly different areas of thought with very little overlap and continuity. The term *adhesive* has been used to describe the particular biomaterial or system used that is responsible for adherence. Bioadhesion has largely remained well entrained in the field of biology referring to cellular interactions (cytoadhesion) or to mucosal adhesion (mucoadhesion). Whether it is cell-to-cell, soft or hard tissue, or mucosal adhesion, the phenomenon of bioadhesion involves the formation of a bond between two biological surfaces or a biological surface and a synthetic surface. Therefore, bioadhesion can be extended to include any type of adhesion process in contact with or within the biological milieu. The term *bioadhesive* refers to the particular substance, natural or synthetic, responsible for this adhesion. Mucoadhesion is often used to describe bioadhesion involving mucus, whereas cytoadhesion refers to cell-to-cell bioadhesion.

Common examples of bioadhesives include dental restoration fixatives, skin adhesives, and bacterial colonization. Much success has been achieved in hard-tissue adhesives with the availability of many commercially available bioadhesive systems. Mucoadhesive materials have been the focus of much research over the last couple of decades in pharmaceutical science research as a novel drug-delivery approach. With mucoadhesive drug-delivery systems, the therapeutic agent can be localized at the desired site of delivery. Peptides, proteins, or actively facilitated compounds can be targeted to the upper small intestine where achieving greater absorption is possible. Similar obstacles can be overcome in other locations of the body, ophthalmic, pulmonary, and vaginal cavities, where mucus is present.

The desire to create biomimetic materials has created interest in further understanding the mechanisms involved with cellular adhesion to aid in engineering effective tissue-engineering scaffolds and biomaterials resistant to biofilm formation. The ability to create

synthetic materials exhibiting molecular structures recognized by specific cellular components has the ability to afford better cell attachment. Increased cell attachment would lead to a more favorable biological response effectively creating more biocompatible scaffolds. On the other hand, with a complete understanding of the biofilm formation process, biomaterials can be designed so as to decrease the propensity of microbial attachment and colonization. Here we provide a broad overview of the molecular and physico-chemical characteristics involved during the formation of bioadhesive bonds along with a short critical review of the progress achieved in each of the distinct thrusts of current research—hard-tissue adhesives, soft-tissue adhesives, and cytoadhesives.

MECHANISMS OF BIOADHESION

Bioadhesion is truly an interfacial phenomenon and only differs from conventional adhesion in the special properties and characteristics of the substrate(s) being adhered. The mechanistic and structural analyses of the phenomenon of bioadhesion have been carefully outlined in several comprehensive reviews.^[1–5] The mechanisms of bioadhesion are often classified into chemical and physical mechanisms with the electronic theory and adsorption theory falling under the former mechanism, and wetting, interpenetration or diffusion, and fracture theory falling under the latter.

Electronic theory describes adhesion between two different electronic structures such as a polymer and a glycoprotein network. Upon contact of these two systems, electron transfer occurs leading to a double layer of electrical charge at the interface. Adsorption theory describes adhesion that occurs owing to secondary forces such as van der Waals, hydrogen bonding, etc. Several occurrences of these types of forces are sufficient to create a substantial adhesive bond.

The wetting theory uses a surface energy analysis to determine to what extent an adhesive or biological surface is capable of spreading over its corresponding substrate to develop intimate contact. Interpenetration theory, or diffusion theory, (Fig. 1) relies on the

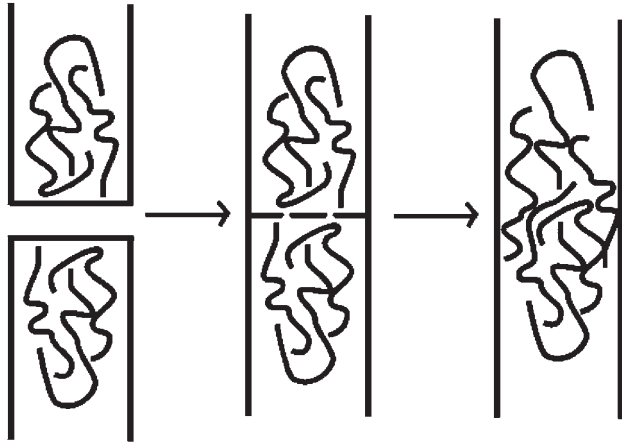


Fig. 1 Upon contact, the interface between two polymer systems disappears and chain interpenetration or diffusion occurs. This polymer interpenetration and depth is responsible for the bioadhesion that occurs.

penetration depth of the polymer chains into the biological substrate to be sufficient enough to create a semipermanent adhesive bond. Finally, fracture theory describes adhesion as the difficulty of separation of the two surfaces. In most systems, one theory alone is not sufficient to describe the process of adhesion. Usually, two or more theories more accurately portray the phenomenon that is occurring. In designing a superior bioadhesive, one must successfully employ key attributes of these theories to develop effective bioadhesion.

Hard-Tissue Adhesives

Smith^[6] offers a comprehensive review on the advances, recent trends, and future direction of hard-tissue adhesives. A great deal of interest in finding alternatives to mechanical fixatives such as screws and pins has prompted the development of efficient adhesives for skeletal and hard-tissue bonding. This area of research has had tremendous success in developing and employing bioadhesives. Dental adhesives and sealants, bone cements, and bone graft substitutes are areas where significant progress has been achieved. One key reason to achieving clinical success has been the environment and tissue involved in these particular bioadhesive processes. Great advances have been achieved in bone cements, dental cements, and enamel and dentin adhesives—thanks in great deal to the engineering of polycarboxylate and ionomer materials capable of achieving adhesion in these applications.^[7,8]

Zinc phosphate cements containing Zn and Al (Caulk Tenacin, Mizzy, Inc. Fleck's Cement) possess a high compressive strength and have been widely used since the mid-20th century in dental restoration. Addition of polycarboxylates to zinc oxide (3MTM

ESPETM DurelonTM, Caulk Tylok Plus) provided advances in adhesion owing to the presence of calcium–carboxylate ion interactions. However, a major drawback to the polycarboxylate adhesives is their poor mechanical properties. To overcome this negative attribute and achieve better optical properties, glass ionomer cements (GIC, GC Fuji PlusTM, and ESPE Ketac-Cem[®]) containing polycarboxylic acid and aluminosilicate glass were developed.^[9] These systems were further modified to include chemical polymerizable (Caulk Advance, 3MTM ESPETM VitremerTM Luting), light-cured (Fuji II LC), or dual-cured (Bisco Resinomer, Dentsply Enforce) resins, which overcome issues of drying and limited fracture resistance with the original GIC. Resins used alone (Panavia 21, J. Morita Clearfil DC Cement) have been used as adhesives that are less brittle and more fracture resistant than the originally employed zinc phosphate cements. Mechanical interlocking and electrostatic interactions are responsible for adhesion and rigidity of the material. Zinc-containing materials do not possess the desired optical properties, therefore necessitating glass ionomer systems (Fuji[®]), which are translucent.

Methyl methacrylate–styrene copolymer-based bone cements (SimplexTM P) have been widely used for securing orthopedic implants to bone since the early 1970s. Cement blends containing antibiotics such as tobramycin have recently been approved for use. Fast-setting carbonated apatite (hydroxyapatite) cements (Norian[®] SRS[®]) have recently been introduced to serve as a more biocompatible material for restoration or augmentation of the skeletal system. These materials harden to form dahllite, a complex of naturally occurring apatite found in bones, which resembles the mineral phase of new bone growth. This further remodels to bone through osteoclastic resorption and new bone formation. Owing to the inorganic composition similarities between bone and enamel and dentin, similar types of materials for dental restoration have also been applied as bone graft substitutes. Effective bioadhesives have found success in the areas of hard-tissue adhesives as evident by the large number of commercially available products in this area.

Soft-Tissue Adhesives

Skin adhesives and tissue sealants have been the traditional application of soft-tissue adhesives. Pressure sensitive adhesives (PSA) have found use in the attachment of medical appliances and to aid in the healing process of abrasions (Band-Aid[®]). These bandages typically have a high tack layer composed of an adhesive derived from polyisobutylenes, acrylates, or silicone polymers. Recent advances have led to the

development of bandage systems that are more flexible with greater adhesive capacity and able to resist water, bacteria, and exogenous material. Semipermeable membranes have been utilized in conjunction with biological fluid absorbent materials in hydrocolloid moisture-retentive wound dressings to promote the wound healing process (ConvaTec). Cyanoacrylates have been used as tissue adhesives owing to their ability to polymerize upon contact with fluids present at biological surfaces. Longer alkyl chain cyanoacrylates (2-octyl cyanoacrylate, Dermabond[®]) have become acceptable bioadhesives owing to their slower degradation, which provides a lower concentration of potentially toxic degradation by-products that could elicit an inflammatory response. Fibrin sealants have been used in wound healing, hemostasis, and tissue sealing and receive a biologically favorable response because of their biomimesis of the last stages of thrombosis (Baxter Tisseel[®], Ethicon Crosseal[™]). The fibrin-sealant systems consist of two separate containers of fibrinogen and thrombin and, when combined and applied to the treatment area, are capable of thrombus formation. The elastic clot formed adheres firmly to the bleeding site and is transparent for visual inspection. Albumin glues (Cryolife[®] Bio-Glue[®] Surgical Adhesive) have been applied at sites of surgery to control bleeding. This two-component adhesive system is composed of bovine serum albumin (BSA) and glutaraldehyde, and when mixed the glutaraldehyde cross-links the BSA molecules.

A shift in focus of soft-tissue adhesives has occurred over the last several decades to evaluate their potential use in bioadhesive controlled drug-delivery devices. Having the ability to prolong a therapeutic's residence time at a specific site of delivery will greatly enhance the drug absorption process. These mucoadhesive devices can protect the drug during the absorption process in addition to protecting it on its route to the delivery site.^[10,11] Three major categories of polymers have been used with some success as bioadhesives: hydroxyl-containing, carboxyl-containing, and charged polymers. Proposed mucoadhesives have been applied in the form of viscous liquids or pastes when used to adhere a polymer to a tissue, or as films or microparticles when used in the form of controlled-release systems adhering to the mucus. The two basic macromolecular structures that have received considerable attention are polyacrylic acid (Noveon[®] Carbopol[®]) and chitosan (Novamatrix[™] Protasan[™]).

Carbopols (Fig. 2) are cross-linked polyacrylic acid microparticles that are capable of swelling upon neutralization. Five specific grades, 934P, 971P, 974P, 71G, and Polycarbophil AA-1, have found wide application in the scientific literature. Carbopol 934P is not widely accepted owing to its polymerization in

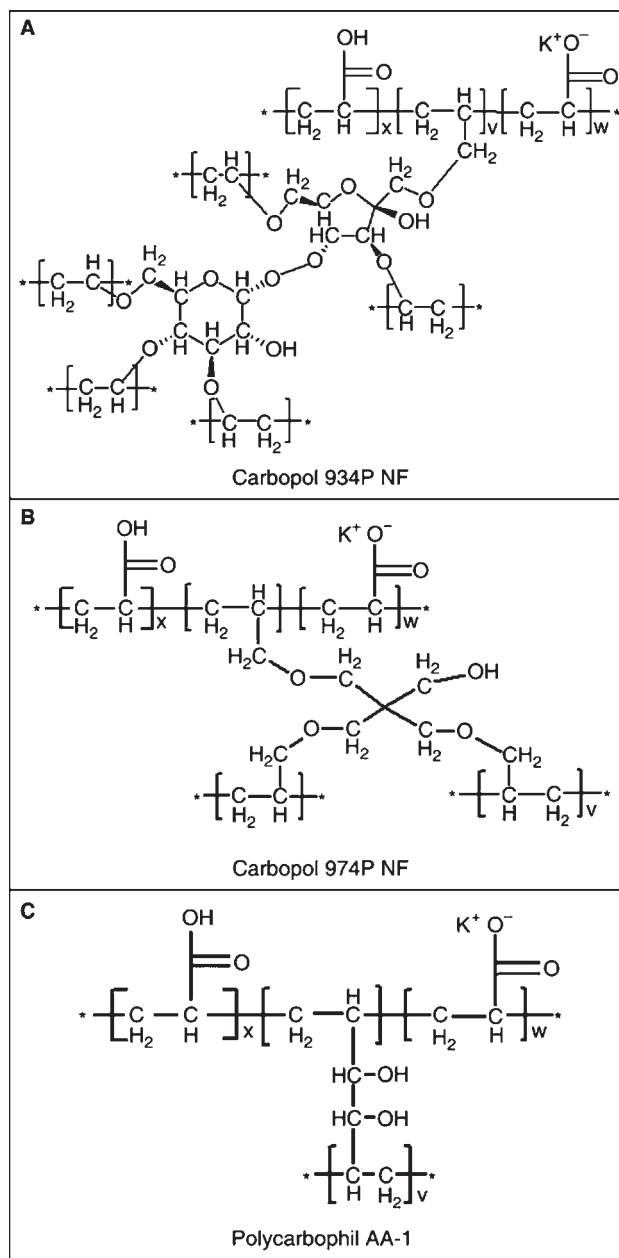


Fig. 2 Chemical structure of commonly employed poly(acrylic acid) (PAA)-based materials in mucoadhesive drug-delivery systems. 934P is cross-linked with allyl sucrose, 974P with allyl pentaerythritol, and AA-1 with divinyl glycol. A small amount of the acrylic acids are neutralized to potassium acrylate prior to polymerization.

benzene. Carbopol 971P, 974P, and 71G all possess the same basic molecular structure, with the cross-linking density being the only difference between 971P and 974P. Carbopol 71G is simply compressed granules of 971P. Polycarbophil AA-1 differs in that the polymer is cross-linked with divinyl glycol. All the later microparticles are polymerized in ethyl acetate, a benign solvent as compared to benzene. Chitosan is a linear

cationic polymer composed of D-glucosamine and N-acetyl-D-glucosamine units, with the amount of either depending on the percent deacetylation. Chitosan is derived from chitin, which originates from shellfish. These polymers along with sodium alginate and cellulose derivatives compose the first generation of bioadhesives owing to their ability to interact non-specifically with tissue surfaces.

Studies show that upon application of a polymer to a tissue, a process of swelling may occur.^[12] Shortly after the beginning of swelling, adhesion does occur but the adhesive strength is not very high. The effect of the swelling time on bioadhesion can be readily calculated by examining the time required for interpenetration (diffusion) of bioadhesive chains into the tissue in a layer of 150–200 Å. The major emphasis of recent research on mucoadhesive biomaterials is on the use of adhesion promoters, which would enhance the adhesion between synthetic polymers and mucus through additional polymer–biopolymer interpenetration (Fig. 3). The use of adhesion promoters such as linear or tethered polymer chains is a natural result of the diffusional characteristics of adhesion.^[13] We have focused our attention on the enhancement of mucoadhesion by tethered chains of poly(ethylene glycol) (PEG), grafted on a polymer backbone. Mucoadhesion depends largely on the structure of the synthetic polymer gels used in controlled release applications. Therefore, on the macroscopic level, we focus on examining the effects of various preparation conditions on the structure and morphology of PAA and related hydrogels.

The fields of biomaterials and drug delivery have evoked considerable interest and research activity in the past 25 yr. Only in the last decade have modulated release systems been considered for their potential commercial application. Prominent among these systems are ionic networks and gels used as carriers for

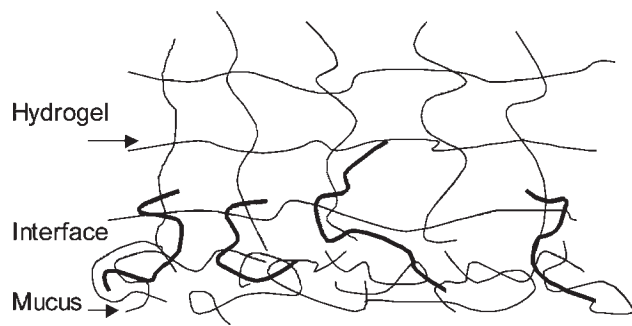


Fig. 3 The free polymer chains across the hydrogel–mucus interface act as adhesion promoters. The free chains are initially in the hydrogels. When placed in contact, the chemical potential gradient causes mutual diffusion of chains across the interface.

drug delivery. Hydrogels are hydrophilic polymer networks capable of imbibing large amounts of water. However, they are insoluble because of the presence of physical or chemical cross-links, entanglements, or crystalline regions. Hydrogels can be used in biomedical applications such as drug-delivery systems, biosensors, contact lenses, catheters, and wound dressings.^[14] Because of the presence of certain functional groups along the polymer chains, hydrogels are often sensitive to the conditions of the surrounding environment.^[15] For example, the swelling ratio of these materials may be sensitive to the temperature, pH, or ionic strength of the swelling agent or even to the presence of a magnetic field or ultraviolet light (Fig. 4).

Numerous experimental studies have been done to test the mucoadhesive ability of various polymers.^[16,17] Among the general rules for mucin-attractive polymers are the requirements of hydrogen-bond formation ability. This is reasonable based on the chemical structure of mucin components,^[18,19] including hydroxyl groups in the branched sugar chains, carboxylic, acid and sulfate groups in the terminal segments of branched chains. Consequently, polymers such as PAA are mucoadhesive; their interactions with mucus have been tested and the data have been widely cited.^[20,21]

The very interaction between polymers and mucins in the physiological environment requires more careful investigation, both theoretically and experimentally. The excess amount of water in the physiological environment (the concentration of water in the bulk phase is around 55 M) has important implications on the hydrogen-binding interaction between polymer chains. For example, in aqueous solutions of PEG, the polymer chains are surrounded by a hydration layer, with two to three water molecules/monomer unit.^[22] Thus, the hydrogen-bond formation between PEG and another polymer leads to the breakage of PEG–water

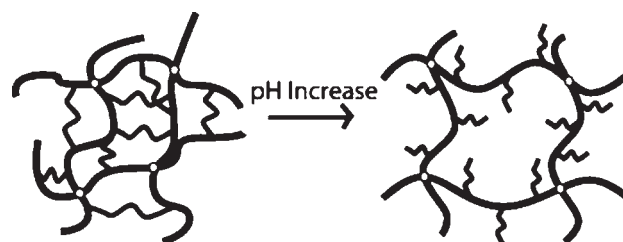


Fig. 4 Network structural changes owing to variations on environmental pH. Higher pHs disrupt the interpolymer complexes and ionic moieties deprotonize leading to extensive swelling (*left*). As pH is decreased, interactions between the tethered grafts with the protonized ionic moieties increase leading to the formation of interpolymer complexes.

hydrogen bond, and this can only occur if there is a net energy decrease. For example, the interaction between PEG and poly(vinyl alcohol) (PVA) is repulsive in water, though such polymers can form hydrogen-bond complex in the dry state.^[23,24] Recent studies of the interaction between PEG and poly(methacrylic acid) (PMAA) in water^[25] showed that the PEG–PMAA complexation diminished even when only 10% of the acid groups in PMAA were dissociated. Therefore, more careful analysis beyond simple consideration of the chemical structure is needed to determine whether there is attraction between polymers and mucins. From an experimental point of view, the current methods for mucoadhesion measurements do not provide direct evidence of the nature of polymer–mucin interactions. These methods mostly exploit homemade or modified equipment, and there is no universally accepted test method.^[26,27] This is the reason that so few comparisons of intergroup data have been made, and, except for a small number of cases, there is no consensus about which polymers are good mucoadhesives.^[26,27] Moreover, different test methods involve various macroscopic phenomena, which are not directly related to polymer–mucin interactions but are method- or sample-dependent.

Determination of adhesive polymer properties can be evaluated by using a range of experimental techniques, leading to the determination of the bond strength. In vitro techniques include mostly tests by which one can follow the destruction of the adhesive bond between a polymer and the mucus or tissue predominantly via application of shear or peeling forces. Often, the same tests are performed under application of tensile forces.

For certain controlled-release systems, there are simple experimental procedures, which can be used to determine the adhesive force, using thin disks. The experiment is carried out in a modified tensile tester using appropriate cells attached to the vertical bars of the tester. The disks or paste are placed between the two plates of the cells and a tensile force is applied at constant extension ratio and until the material breaks by molecular slippage (Figs. 5 and 6). The maximum elastic modulus is usually proportional to the adhesive strength.

The Wilhelmy plate method has been used to test the mucoadhesion of water-soluble polymers.^[20] In this method, the polymer-coated glass plate was tested against mucus. However, the contact time (the hydration time) turned out to be the essential parameter for the result analysis: too short time would cause insufficient interaction between polymers and mucus.^[16] These disadvantages are expected to cause poor correlation with the in vivo interaction.^[28] More importantly, it is barely possible to get an insight into molecular-level interactions between various polymers

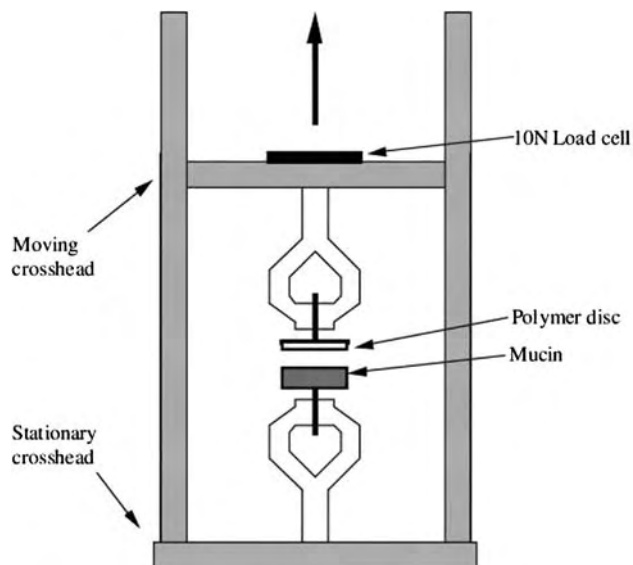


Fig. 5 Bioadhesion measurements using a tensiometric device such as an Instron. The adhesive is placed in contact with the substrate to be adhered. The adhesive is removed while measuring the force vs. elongation.

and mucin from the current macroscopic test methods.^[29] Therefore, despite the existence of numerous experimental data, we still do not have enough knowledge to molecularly design the next generation of mucoadhesive devices. We have exploited both theoretical modeling and microscopic measurement methods to provide a clearer picture of polymer–mucus interactions.

Cytoadhesives

The adhesion of cells has been described as a four-step process: cell attachment, cell spreading, organization

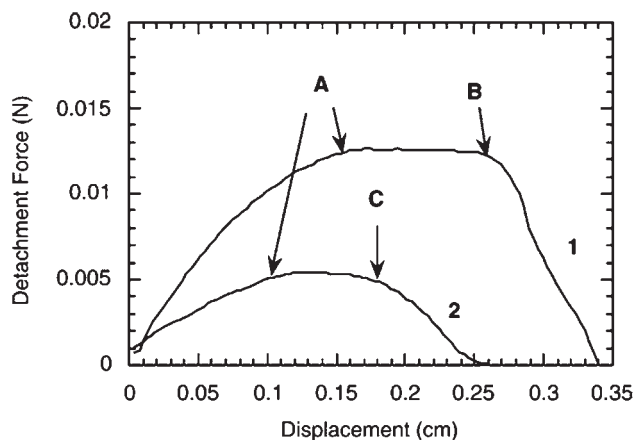


Fig. 6 Force vs. displacement curve for bioadhesives in contact with a biological substrate. (A) Represents the peak force, where (B) and (C) represent tissue detachment.

of an actin cytoskeleton, and formation of focal adhesions.^[30] Cellular adhesion is generally promoted through the presence of various adhesion proteins,^[31] with fibronectin being one of the more widely recognized proteins. Platelet adhesion is an integral step in blood coagulation and blood-material interactions.^[32] Coagulation arises from interactions of platelet glycoprotein Ib and exposed connective tissues such as collagen, with plasma von Willebrand factor (vWF) as a cofactor. Material platelet adhesion involves other platelet glycoproteins (IIb/IIIa) along with the coagulation mechanism. This material-platelet adhesion arises from IIb-IIIa interactions with plasma proteins such as fibrinogen, vWF, fibronectin, and vitronectin. Integrins, the family of cell-surface matrix receptors, serve as the mediators for cell adhesion. Arg-Gly-Asp (RGD) has often been employed on synthetic surfaces to stimulate adhesion in tissue-engineering applications so that further processes such as proliferation and migration will occur. This peptide attachment simulates the components of the extra cellular matrix (ECM), thereby imparting biomimetic properties on the biomaterial. Aside from integrin-surface receptors, proteoglycans are also responsible for cell adhesion. Carbohydrate-rich domains located at the cell's surface have the ability to interact with carbohydrate-binding lectin macromolecules.^[33]

Adhesion of bacteria (biofilm formation or biofouling) involves the interaction of bacterial fibers with fibers adsorbed to the surface being colonized.^[34] Bacteria have often been overly simplified and described as a smooth colloidal particle so that DLVO theory could be used, at least partially, to explain the adhesion process.^[35,36] Biofouling of biomaterials occurs owing to non-specific interactions of the bacterial surface dependent on polymer conformation and surface energy whereas certain bacteria (e.g., *Helicobacter pylori*) possess a surface layer of biopolymers capable of specific interactions with a biological host.^[37] This specific interaction is typically mediated by carbohydrate-binding moieties present on the bacterium's surface. Bacterial colonization of the gastrointestinal tract is a result of the ability of certain bacteria's surface binding to mucin molecules present in the mucus gel layer.

CONCLUSIONS

Bioadhesive polymers can be used in a wide range of medical applications including cytoadhesive, mucoadhesive, and tissue adhesive uses. Advances in nanostructure development and nanotechnology allow us to utilize a wide range of materials for such applications. While PAA-based, hydrophilic polymers are excellent bioadhesive candidates, addition of tethered structures,

ionizable groups, and interpenetrating structures seems to promote interdiffusion and therefore adhesion.

ARTICLES OF FURTHER INTEREST

Bioadhesion; Biofilms; Biological Adhesives from Nature; Bone Cement; Fibrin; Fibrin Sealants; Glues; Surface Coatings

REFERENCES

1. Peppas, N.A.; Buri, P.A. Surface, interface, and molecular aspects of polymer bioadhesion on soft tissues. *J. Controlled Release* **1985**, *2*, 257–275.
2. Hubbell, J.A. Biomaterials in tissue engineering. *Biotechnology* **1995**, *13*, 565–576.
3. Peppas, N.A.; Sahlin, J.J. Hydrogels as mucoadhesive and bioadhesive materials: a review. *Biomaterials* **1996**, *17*, 1553–1561.
4. Mathiowitz, C.; Chickering, D.E.; Lehr, C.M. *Bioadhesive Drug Delivery Systems*; Marcel Dekker: New York, 1999.
5. Lee, J.W.; Park, J.H.; Robinson, J.R. Bioadhesive-based dosage forms: the next generation. *J. Pharm. Sci.* **2000**, *89*, 850–866.
6. Smith, D.C. Adhesives and sealants. In *Biomaterials Science: An Introduction to Materials in Medicine*; Ratner, B.D., Hoffman, A.S., Schoen, F.J., Lemons, J.E., Eds.; Elsevier Academic Press: San Diego, 2004; 572–583.
7. Scranton, A.B.; Rangarajan, B.; Klier, J. Biomedical applications of polyelectrolytes. *Biopolymers II Adv. Polym. Sci.* **1995**, *122*, 1–54.
8. Davidson, C.L.; Feilzer, A.J. Polymerization shrinkage and polymerization shrinkage stress in polymer-based restoratives. *J. Dentistry* **1997**, *25*, 435–440.
9. Smith, D.C. Development of glass-ionomer cement systems. *Biomaterials* **1998**, *19*, 467–478.
10. Lowman, A.M.; Peppas, N.A.; Morishita, M.; Nagai, T. Novel bioadhesive complexation networks for oral protein drug delivery. In *Tailored Polymeric Materials for Controlled Delivery Systems*; McCulloch, I., Shalaby, S.W., Eds.; ACS Symposium Series: Washington, DC, 1999; Vol. 709, 156–164.
11. Lowman, A.M.; Peppas, N.A. Solute transport analysis in pH-responsive, complexing hydrogels of poly(methacrylic acid-g-ethylene glycol). *J. Biomat. Sci. Polym. Ed.* **1999**, *10*, 999–1009.
12. Mortazavi, S.A.; Smart, J.D. An investigation into the role of water-movement and mucus gel dehydration in mucoadhesion. *J. Controlled Release* **1993**, *25*, 197–203.
13. Peppas, N.A.; Bures, P.; Foss, A.; Huang, Y.; Leobandung, W. Molecular aspects of muco- and bioadhesion: tethered structures and site-specific surfaces. *Proceed. Intern. Symp. Recent Adv. Drug Deliv. Sys.* **1999**, *9*, 31–34.

14. Peppas, N.A.; Brannon-Peppas, L. Drug delivery biomaterials. In *Encyclopedia of Materials: Science and Technology*; Elsevier: Amsterdam, 2001; 2351–2355.
15. Peppas, N.A.; Bures, P.; Leobandung, L.; Ichikawa, H. Hydrogels in pharmaceutical formulations. *Eur. J. Pharm. Biopharm.* **2000**, *50*, 27–46.
16. Gu, J.M.; Robinson, J.R.; Leung, S.S. Binding of acrylic polymers to mucin/epithelial surfaces: structure-property relationships. *CRC Crit. Rev. Ther. Drug Carrier Syst.* **1988**, *5*, 21–67.
17. Peppas, N.A.; Huang, Y.; Torres-Lugo, M.; Ward, J.H.; Zhang, J. Physicochemical foundations and structural design of hydrogels in medicine and biology. *Ann. Revs. Biomed. Eng.* **2000**, *2*, 9–29.
18. Bansil, R.; Stanley, E.; LaMont, J.T. Mucin biophysics. *Ann. Rev. Physiol.* **1995**, *57*, 635–657.
19. Strous, G.J.; Dekker, J. Mucin-type glycoproteins. *Crit. Rev. Biochem. Mol. Biology* **1992**, *27*, 57–92.
20. Smart, J.D.; Kellaway, I.W.; Worthington, H.E.C. An in vitro investigation of mucosa-adhesive materials for use in controlled drug delivery. *J. Pharm. Pharmacol.* **1984**, *36*, 295–299.
21. Bures, P.; Peppas, N.A. Structural and morphological characteristics of carriers based on poly(acrylic acid). *Polym. Prepr.* **1999**, *40* (1), 345–346.
22. Kjellander, R.; Florin, E. Water structure and changes in thermal stability of the system poly(ethylene glycol)-water. *J. Chem. Soc. Faraday Trans I* **1981**, *77*, 2053–2077.
23. Inamura, I.; Jinbo, Y. Interaction between poly(vinyl alcohol) and poly(ethylene glycol) in water studied by viscosity and density. *Polym. J.* **1991**, *23*, 1143–1147.
24. Sawatari, C.; Kondo, T. Interchain hydrogen bonds in blend films of poly(vinyl alcohol) and its derivatives with poly(ethylene oxide). *Macromolecules* **1999**, *32*, 1949–1955.
25. Zeghal, M.; Auvray, L. Structure of polymer complex in water. *Europhys. Lett.* **1999**, *45*, 482–487.
26. Tobyn, M.J.; Johnson, J.R.; Dettmar, P.W. Factors affecting in vitro gastric mucoadhesion I. Test conditions and instrumental parameters. *Eur. J. Pharm. Biopharm.* **1995**, *41*, 233–241.
27. Tobyn, M.J.; Johnson, J.R.; Dettmar, P.W. Factors affecting in vitro gastric mucoadhesion II. Physical properties of polymers. *Eur. J. Pharm. Biopharm.* **1996**, *42*, 56–61.
28. Lehr, C.M. Bioadhesion technologies for the delivery of peptide and protein drugs to the gastrointestinal tract. *CRC Crit. Rev. Ther. Drug Carrier Syst.* **1994**, *11*, 119–160.
29. Schneider, J.; Tirrell, M. Direct measurement of molecular-level forces and adhesion in biological systems. In *Bioadhesive Drug Delivery Systems*; Mathiowitz, E., Chickering, D.E., Lehr, C.M., Eds.; Marcel Dekker, Inc.: New York, 1999; 223–259.
30. Drotleff, S.; Lungwitz, U.; Breunig, M.; Dennis, A.; Blunk, A.; Blunk, T.; Tessmar, J.; Gopferich, A. Biomimetic polymers in pharmaceutical and biomedical sciences. *Eur. J. Pharm. Biopharm.* **2004**, *58*, 385–407.
31. Horbett, T.A. The role of adsorbed proteins in tissue response to biomaterials. In *Biomaterials Science: An Introduction to Materials in Medicine*; Ratner, B.D., Hoffman, A.S., Schoen, F.J., Lemons, J.E., Eds.; Elsevier Academic Press: San Diego, 2004; 237–245.
32. Hanson, S.R. Blood coagulation and blood-material interactions. In *Biomaterials Science: An Introduction to Materials in Medicine*; Ratner, B.D., Hoffman, A.S., Schoen, F.J., Lemons, J.E., Eds.; Elsevier Academic Press: San Diego, 2004; 332–338.
33. Bies, C.; Lehr, C.M.; Woodley, J.F. Lectin-mediated drug targeting: history and applications. *Adv. Drug. Del. Rev.* **2004**, *56*, 425–435.
34. Costerton, B.; Cook, G.; Shirtliff, M.; Stoodley, P.; Pasmore, M. Biofilms, biomaterials, and device-related infections. In *Biomaterials Science: An Introduction to Materials in Medicine*; Ratner, B.D., Hoffman, A.S., Schoen, F.J., Lemons, J.E., Eds.; Elsevier Academic Press: San Diego, 2004; 345–354.
35. Xu, L.C.; Vadhillo-Rodriguez, V.; Logan, B.E. Residence time, loading force, pH, and ionic strength affect adhesion forces between colloids and biopolymer-coated surfaces. *Langmuir* **2005**, *21*, 7491–7500.
36. Abu-Lail, N.I.; Camesano, T.A. Role of ionic strength on the relationship of biopolymer conformation, DLVO contributions, and steric interactions to bioadhesion of *Pseudomonas putida* KT2442. *Biomacromolecules* **2003**, *4*, 1000–1012.
37. Walz, A.; Odenbreit, S.; Mahdavi, J.; Boren, T.; Ruhl, S. Identification and characterization of binding properties of *Helicobacter pylori* by glycoconjugate arrays. *Glycobiology* **2005**, *15*, 700–708.

Allografts

Kirby S. Black

Patti E. Dawson

Research and Development, CryoLife, Inc., Kennesaw, Georgia, U.S.A.

INTRODUCTION

Explorations under way for centuries have sought ideal replacement materials capable of restoring our damaged or diseased functional parts to their natural state. Early historical accounts indicate that a number of materials were fashioned to resemble what they would replace, including wood (for teeth), ivory, and animal bone and tissue. These early efforts at grafting were concerned primarily with replicating the form of the part they would replace, but did not succeed at replacing functional elements.

OVERVIEW

More recently, materials used for tissue replacement have included a much more sophisticated array of materials including stainless steels, plastics, and ceramics. In a previous work on biomaterials, homage has been paid “To those unknown and unsung researchers and clinicians who pioneered in adapting commercial metals and polymers in human needs. They replaced worn-out and injured living tissue with man-made substitutes and so took the first steps to perpetuate human life.”^[1] With this focus on man-made substitutes, engineering seems to have dominated the realm of biomaterials. The concept of biomaterials has generally been associated with synthetic materials designed to look, feel, and act like tissue. Such synthetic substitutes have been designed to better mimic the tissues they are intended to replace. Some are extremely durable; however, due to their nonbiologic nature, they do not fully integrate into the biologic systems into which they are placed. These materials either lack an important metabolic function or are actually bioincompatible. These materials, although very durable, are therefore limited in their utility by issues of bioincompatibility and by the nonrenewable nature of these replacements. The nonintegrative nature of these implants can result in short implant durability or patient incompatibility. Mechanical heart valves present one example of incompatibility between patient and graft. Even though the grafts are structurally very durable, their surface properties are thrombogenic. No available material has yet attenuated this characteristic; the

need for patients to take anticoagulants while implants are in place limits mechanical valve utility in many patient populations.

Tissue is used as an alternative to synthetic materials. Initially (and continuing today), tissues of animal origin were used as graft material. Tissue-processing methodologies such as chemical fixation have resulted in improvements in biocompatibility (by allowing cross-species implantation without rejection) and physiologic functionality (due to the intrinsic functionality of the original tissue). However, the chemical fixation of these grafts prevents recipients' cells from interacting with and routinely rebuilding and regenerating the device. Application of chemical fixation techniques to human tissue grafts (especially heart valves) has actually caused an unanticipated decrease in these grafts' functional utility.

As a consequence of the failure of wholly synthetic materials and fixation-treated animal tissues to integrate with the biological environment, newer biomaterials development has begun to incorporate the raw elements of the tissues they have been created to replace. These new generation biomaterials (many of which are currently in development) routinely incorporate the basic structural elements of human tissues—proteins, glycosaminoglycans (GAGs), minerals, and even cells. Current biomaterials development seeks to make the biomaterial more interactive with the cells and surrounding tissues. These design elements, which are becoming more common in biomaterials development, are all elemental to an allograft.

No material has yet been created that is as suitable as an allograft to replace diseased or damaged tissues when seamless integration and function are required. The particular utility of allograft tissue-based replacements is best understood in terms of the body's durability (such durability being primarily a function of normal cellular replication and cellular/matrix interactions that continuously renew and repair tissue matrices). Without reestablishment of this cell/matrix interaction, durability of tissue replacements could be greatly reduced (as has been found with chemically fixed animal tissue grafts). Therefore, the goal of a medical device for tissue replacement should include the ability of the device to interact with the patient's own cells and the potential to allow cells to rebuild and

regenerate tissue. The inclusion of cells in biomaterials created for graft construction (preseeding) is currently an area of great interest and activity. This construct could provide the same long-term durability that our natural tissues provide. The use of allograft tissue for transplantation is based on the ability of graft and host cells to interact and renew.

The term allograft (also known as homograft) indicates tissue from one species transplanted to another animal of the same species (e.g., dog to dog or human to human). A xenograft indicates tissue transplanted from one species to another (e.g., pig to monkey or pig to human). An autograft is tissue reimplanted into the donor. The most common clinical example of an autograft is the use of a patient's own veins or arteries in heart bypass grafting (Fig. 1).

Allografts can be classified into three broad categories—organs, tissues, and cells. There are several ways to make distinctions between organs and tissues, but one of the more clearly defining parameters is whether the tissue requires immediate revascularization (vascular anastomoses), or can be implanted without immediate revascularization. Organ allografts (kidneys, liver, lung, heart, etc.) all require immediate surgical revascularization for their metabolic requirements, thereby requiring reattachment by surgical anastomosis of the major arteries and veins. On the other hand, tissues and cells do not require revascularization and can be implanted without immediately establishing a direct blood supply (heart valves, blood vessels, orthopedic tissues, skin, cornea, etc.). Another distinction can be made regarding whole organs' requirement for an immediate blood supply because of their size and because they are sites of important and

continued metabolic activity (e.g., liver, kidney, pancreas). Allograft tissues are generally much smaller in dimension than whole organ allografts (and can therefore be revascularized in a short time from indirect capillary revascularization), and do not have the metabolic requirements of a whole organ. Examples of cellular allografts include blood, bone marrow, neurological cells, chondrocytes, and pancreatic islets. These do not require immediate revascularization; however, they do require implantation into a highly vascular area.

The history of the use of allografts dates back many years, if not centuries. Skin grafts were first reported in 1881.^[2] The legend of Saints Cosmas and Damien, however, dates back much further. The Roman legend tells of an old servant and admirer of these two saints. The servant fell ill from a cancer in his leg and prayed to Cosmas and Damien and fell asleep in church. The saints appeared with surgical instruments and medicines. After amputating his limb one asked, "What are we going to use to replace the diseased flesh?" Saint Damian answered, "The body of a Moor was buried today at the cemetery of St. Peter. It is still fresh and we can use it for this poor man." The limb was a different color but appeared to work well.^[2-4] Another amazing component of this story is that the saints had been martyred for their faith two centuries before the transplant.

During early experimental use of allografts it was found that if administered in large enough quantities, they could have unfortunate effects on the recipient. These effects were later understood to be one of two responses. One response is rejection, resulting in loss of tissue due to an immunological reaction to antigens in the graft tissue. The second is a sensitization that can result in a massive immune response directed back at the patient through anaphylaxis or graft versus host disease. These effects have prompted wide-ranging research efforts in development and use of various drugs that allow allograft transplantation with reduced concerns for rejection or other negative effects on the recipient.

Immunosuppressive drugs have been used since the late 1950s, starting with azothiaprine.^[4,5] These were known to have an impact on the cells of the immune system, preventing them from recognizing or responding to the allograft tissues. The field of immunosuppressives has grown rapidly and has brought many new drugs (such as cyclosporine and FK-506) to the forefront of transplantation use. These drugs are primarily associated with whole organ and cellular implants. Immunosuppressive drugs are not commonly utilized with tissue allograft implants.

Research efforts directed at preservation of allografts is another topic of interest. Whole organ grafts must still be implanted into the host within a very short

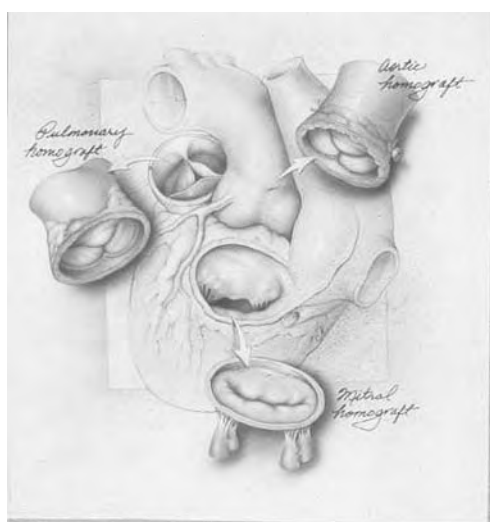


Fig. 1 Depiction of human aortic, pulmonary and mitral valves and their positions in the heart. (Figure by CryoLife, Inc.)

period of time, or the resulting ischemia will affect the organ and result in markedly reduced metabolic function. The major focus of research on whole organ transplantation revolves around development of harvest solutions and transport conditions that allow better survival of the fresh organ for longer time periods. To date, there has been little reported success in development of methods for preserving whole organs for more than 2–3 days. This short time frame does not allow for accumulation or banking of whole organs for better management of these scarce resources. Allograft tissues, on the other hand (due to their smaller size and less complex functional requirements), have been found to be less affected by longer ischemic times. They can be cryopreserved and stored for long times at liquid nitrogen temperatures, thus providing access to these allografts for a wider variety of reconstructions and offering more convenient and widespread use in surgery.

There are many parameters to examine when considering an allograft for replacement of lost or diseased tissue, including patient suitability, graft availability, functional requirements, mechanical requirements, and size. The allograft implantation can often be a challenging implant surgery, requiring the surgeon to undergo specialized training in surgical techniques and postoperative requirements. Nonetheless, the demand for allograft tissue continues to be greater than the supply.

ALLOGRAFT PROPERTIES

The wide variety of medical conditions that are currently treated with allografts demonstrates the diversity of the biomaterial properties possessed by these grafts. Consider that allografts can be optical (cornea), mechanical (bone, cartilage), physiological (organ replacement kidney, liver, lung), hydrodynamic (heart valves, vascular), electrical (nerve), barrier (skin), defense (bone marrow, lymphocytes), or chemical (oxygen carrying blood cells). The diversity of properties is achieved with a wide array of proteins, carbohydrates, lipids, and minerals. This mixture of materials makes a composite implant that can not only function immediately in the new host environment, but can also continue to integrate into the biology of the host through cellular interactions. In many cases the host takes over control of the new implant and makes adjustments to accommodate the graft into its new environment.

How this is accomplished is a question of biology and all its complexities. One of the advantages of allografts is that their complexity and diversity are already inherent in the tissue when harvested. The key is to select tissues that are free of disease and are of correct anatomical characteristics, and to maintain

them in a state that limits changes to their natural structure and capabilities. It is not required to recreate all of the intricacies of the biology of an allograft, but the donor and recipient must be carefully matched to allow the allograft maximum opportunity for its normal biological function. To further understand these diverse properties let us examine three different tissues: skin, heart valves, and orthopedic soft tissues.

Skin

Skin allografts have played a significant role in replacing skin lost in injuries such as burns. The use of skin allografts dates back to 500 B.C. in India.^[2] Gibson and Medawar studied the immunologic nature of skin allografts after using them to treat burned Royal Air Force pilots during World War II. Their work helped to lay the foundation for modern transplantation immunology.

The skin can be thought of from several different perspectives: biochemically, as an assembly of enzymes; cytologically, as a collection of cell types; anatomically, as a formed structure; and physiologically, as a functional aspect of the formed skin. Studies have been conducted that define the permeability mechanics of the skin, absorption of radiation by the skin, and heat loss through the skin.^[6] Skin is described as a membrane, a transport system, and a barrier. Many of these functions are attributable to the properties of the proteins and structural elements.

The two main structures of human skin are the epidermis and dermis. The epidermis is the outermost layer and consists of stratum corneum and differentiating keratinocytes. The epidermis is attached to the dermis through a basement membrane. The stratum corneum is the primary barrier to pathogens, toxins, gas, and fluid and is made up of flat, cornified cells. The primary components of the stratum corneum are the membranes of the dead keratinocytes. The skin is in a continual state of renewal. The keratinocytes of the lower epidermis continue to differentiate and divide, pushing up the next layer.

The underlying dermis is made of collagen fibrils and glycosaminoglycans. The viscoelastic properties of the skin are based on the unique cross-links of these two components. The vascular supply of the dermis supports the nonvascular epidermis layer. The hair follicles and sweat glands originate in the dermis and penetrate through the epidermis. These specialized cells provide sensation and evaporative heat loss capabilities. The skin is highly innervated. The skin of the fingers has unusual structures called rete ridges that allow for determination of the smoothness of a surface. These structures, known as fingerprints, allow for individual identification.

The skin also plays an important role as a first line of defense against biological pathogens like bacteria, fungi, and viruses. There is both a physical barrier (discussed previously) and an active cellular component. The immune cells of the skin include resident monocytes (T cells and macrophages). These cells may play a role as extrathymus educational centers, providing a first look at new antigens as they pass through the skin.

Skin allografts are commonly used and are well accepted clinically, and are generally provided for implant in either a fresh or cryopreserved format.

Heart Valves

The heart valve is a remarkable device. It starts to function before birth, and must continue to function within very narrow tolerances until death. In fact, the function of the heart valve is so essential that any defect can be fatal. Over a 70-year life span the aortic valve can undergo as many as 2.2 billion cycles. The valve is subjected to pressures in the order of 120–150 mmHg and must function in an environment where components of the blood could clot within seconds of a malfunction. The valve itself is a combination of at least five tissue types ranging in function from structural to blood interface barrier. The structure of the heart valve is especially interesting because it is one of the most mechanically active tissues in the body (Figs. 2 and 3).

The main function of the heart valve is to ensure that blood flows in one direction. This is essential to providing blood flow to the tissues for delivery of oxygen and nutrients and removal of CO₂ and wastes and metabolites. The leaflets of the valve provide the

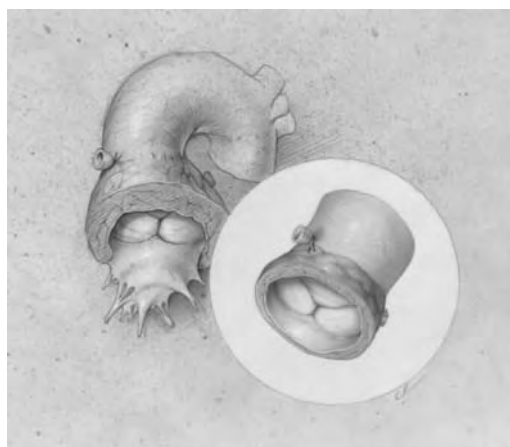


Fig. 2 Depiction of human aortic valve after completion of removal and dissection. (Figure by CryoLife, Inc.)

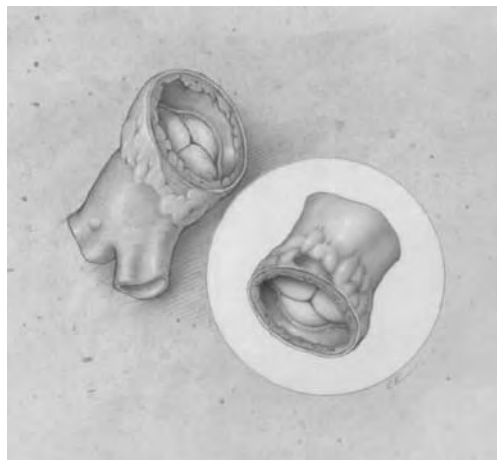


Fig. 3 Depiction of human pulmonary valve after completion of removal and dissection. (Figure by CryoLife, Inc.)

mechanism for one-way flow. There are four valves in the heart—aortic, pulmonary, mitral, and tricuspid valves.

Heart valves are a fibrous network of structural glycoprotein, collagen, elastin, glycosaminoglycans (GAGs), soluble proteins, and cells. In humans the glycosaminoglycans are made up of 65% hyaluronic acid and 35% chondroitin sulfate. The two basic structural elements of the valve are the conduit and the leaflets. The leaflet structure reflects its structural loading conditions. The aortic leaflet is made up of a three major layers—ventricularis (primarily elastin), fibrosa (primarily collagen), and spongiosa (enriched in GAGs). The leaflet motion in flow has been described as aerodynamic in nature. The ease of motion of the leaflets allows minimal opening pressures and volumes, thus maximizing cardiac output. The aortic valve conduit includes the sinus of Valsalva. These unique structures play an important role in leaflet opening and closing and in laminar flow through the valve. The flow dynamics of these structures were described by Leonardo da Vinci, among others.^[7]

The cellular makeup of the heart valve includes fibroblasts, smooth muscle cells, endothelial cells, and myofibroblasts in the leaflet. The major function of these cells has been speculated to involve repair, although they may provide for some contractile component in leaflet motion. Studies conducted at the CryoLife institution have documented the essential factors for preservation of these multicellular tissues. The maintenance of both the structural elements of the tissue matrix and retention of the cell types in significant numbers can indeed result in valvular replacements that can achieve very high long-term success rates.^[8–11] These cryopreserved valves have proven long-term durability, excellent hemodynamics with low

pressure gradients, a large effective orifice area, and low complication rates with low incidence of endocarditis or thromboembolic events.

Heart valve replacement is a procedure that can make a dramatic difference in a patient's life. Over the last four decades, heart valve replacement has become a safe and routine surgical procedure. Advances in operative procedures have dramatically reduced mortality rates associated with this procedure. The patient's quality of life is dramatically increased following replacement (Fig. 4).

The gold standard of treatment that currently provides the best chance for improved quality of life for patients requiring heart valve replacement is an allograft heart valve. Successful clinical application of long-term storage conditions to allograft heart valves was first reported in 1975 by O'Brien and coworkers, in their original series of cryopreserved heart valve implants.^[10] In this pioneering work, Dr. O'Brien compared the clinical outcomes of the viable cryopreserved valves to an earlier series of nonviable, antibiotic treated, 4°C-stored valves. The viable cryopreserved valve series demonstrated clearly superior durability, which in this series continued to be demonstrated in long-term patient follow-up. Dr. O'Brien's group in Brisbane, Australia has the longest experience with cryopreserved allograft heart valves. The reported freedom from structural valve deterioration in this series using cryopreserved aortic valves for aortic valve replacement at the 10- and 15-year time points was 93% and 80%, respectively. More recently, the same group followed this implant series out to 20 years and reported 60% freedom from structural valve deterioration.^[11]

It is now generally understood that the original donor cells do not persist even in viable allograft heart valve grafts. The preservation of the cellular content in



Fig. 4 Human aortic valve allograft prepared for implant. (Figure by CryoLife, Inc.)

these tissues has come to be accepted as a relevant measure of the ability of cryopreservation to also conserve and preserve the vital structural elements of the allograft. A cryopreserved valve at implant therefore has all the structural parts required for replacement of the patient's own heart valve, and is able to be integrated into the new host environment.

Antibiotic treatment followed by cryogenic processing and storage is presently considered the only clinically accepted method for long-term preservation of human heart valve and vascular allograft tissues.^[12-14] Studies in the CryoLife laboratory show that retention of human valve leaflet viability is clearly achievable using optimized cryopreservation methods. Table 1 shows percent viability (live cells/total cells) and cell density (total cells/mm²) in a comparison between fresh valves and cryopreserved valves. These data demonstrate that cryopreservation does not significantly affect valve cellular content or viability. Cryopreservation of allograft valves thus allows retention of the vital structural and biological elements necessary for valve replacement.

These viability studies were performed using a live/dead viability assay method that simultaneously utilizes two fluorescent nucleic acid stains, ethidium homodimer and Hoechst 33342 dye.^[15] Hoechst 33342 is a membrane-permeant blue fluorescent dye that crosses the cell wall and labels the nucleus of all cells, both viable and nonviable. The ethidium homodimer is membrane-impermeant red fluorescent dye and will label only nuclei of dead cells with permeabilized cell membranes; intact cell membranes will prevent this component from entering the nucleus. Under fluorescent light, the nuclei of cells with permeabilized, compromised cellular membranes appear red; the nuclei of live cells with intact cell membranes appear blue. Intact valve leaflets or sections of conduit wall can be incubated in this solution, mounted to microscope slides, and assessed using a fluorescent microscope for relative cell viability and cell density.

The allograft heart valve continues to demonstrate excellent long-term function in both neonate, young adult, and elderly populations.^[8,16] However, worldwide

Table 1 Human heart valve leaflet cellular viability and density assessment performed on fresh and cryopreserved valves

	Percent viable cells	Cell density cells/mm ²
Fresh	70.69	2434
Cryopreserved	75.19	1880
P value	P=0.13	P=0.08

(Data supplied by CryoLife, Inc.)

these tissues are in relatively small supply. Consequently, only a fraction of those who could benefit from these tissue valves actually receive one. Some of the most common clinical indications for allograft heart valve use are patients with active endocarditis, women of childbearing age, patients contraindicated for anticoagulation, patients with an active lifestyle or small aortic root, pulmonary valve replacement in the Ross procedure, and repair of congenital heart defects.

Since its inception in 1984, CryoLife has specialized in the ultra-low temperature preservation of human tissues for transplantation. During 18 years of tissue processing, CryoLife has documented the implant of over 76,149 cardiovascular allografts. Based on published clinical data and an in-house clinical registry, CryoLife allograft heart valves have been demonstrated to have excellent long-term durability. CryoLife's in-house registry consists of data collected at seven primary centers where a series of 2605 patients have been enrolled to date. This registry consists of 8726 patient-years of data (mean follow-up: 3.3 years). Evaluating the performance of the aortic and pulmonary allograft in the adult population, the durability of these CryoLife cryopreserved allografts (defined as freedom from explant due to structural valve deterioration) at 10 years is 88% and 94%, respectively.^[8]

Orthopedic Soft Tissues

Musculoskeletal tissues compose a large portion of the human body. The structures range from ropelike ligament to smooth bearing surfaces (like articular cartilage). Many of these tissues form the structural elements through which the human body ambulates and articulates. These soft tissues include tendons, ligaments, muscle, fascia, and articular cartilage (Fig. 5).

The tissues' structures are as diverse as their functions. The human ligament is a very specialized dense connective tissue that serves as a connection between bones, attaching at either end through structures known as insertions. Ligaments are composed of collagens, glycosaminoglycans, elastins, actins, and fibronectins. The tendon attaches a muscle to bone, and is similar in structural composition to the ligament. Tendons can be long, narrow structures (such as the flexor digitorum found in the hand), which are commonly contained in synovial sheaths. These sheaths provide a smooth surface for gliding and thicken in some areas to form pulleylike structures. These pulley systems are necessary when the tendon turns a corner, as over the knuckles of the fingers. Biomechanically, the pulleys maintain an appropriate moment arm to allow efficient use of muscle force in the movement of a digit or limb.



Fig. 5 Depiction of the human knee including placement of meniscal and articular cartilage allografts. (Figure by CryoLife, Inc.)

Articular cartilage provides a smooth lubricated surface for joint motion, allows for distribution of loads, and acts as a shock absorber. There are two main forms of articular cartilage—load-bearing and non-load-bearing. The structure of the articular cartilage includes cellular and extracellular constituents and 60–80% extracellular water. The cells are mainly chondrocytes and the extracellular matrix hyaline cartilage. The hyaline cartilage is made up of collagens, proteoglycans, and noncollagenous proteins. The articular cartilage is a unique bearing surface that is both porous and permeable and can absorb and exude fluid. The primary mechanism for smooth joint function depends on lubrication by both boundary lubrication and fluid film lubrication. This combination depends on the adsorption of the lubricating glycoprotein at the surface of the cartilage and the viscoelastic nature of the cartilage allowing a deformable bearing surface, thus allowing lubrication during both high-load/low-relative speed and low-load/high-relative speeds.

No biomaterials have yet been created that have been successful in replacing isolated tendons, cartilage, or menisci. The reliance on allografts for connective tissue grafts is therefore very strong. Although whole knee replacements are remarkably successful for whole joint surgeries in older patients, younger individuals

A

needing isolated tendon or cartilaginous grafts must rely on allografts. Orthopedic allografts are currently processed and implanted in many formats. Bone allografts are available in a wide range of configurations, from morselized freeze-dried chips to large intact bone segments. Orthopedic soft tissue allografts, which include meniscus and tendons, are currently processed using methods that yield nonviable grafts such as lyophilization and uncontrolled freezing (fresh/freezing) or by cryopreservation, which yields viable grafts^[17,18] (Fig. 6).

The first reported clinical implantation of cryopreserved meniscal allografts dates to 1989.^[19] At present, tendons, ligaments, and the fibrocartilaginous meniscus have all been successfully cryopreserved and implanted. Successful cryopreservation results in maintenance of both the collagen and proteoglycan-rich matrix and retention of viability of the cells inhabiting the matrix. Because the cellular components of these tissues are without question the most sensitive to the effects of freeze-induced injury, they represent a convenient marker for judging the success of a cryogenic system. Maintenance of intact cellular membranes resulting in viable cells after processing serves as a relevant measure to verify that ice damage and solution effects during the freezing/storage/thawing process have been sufficiently controlled, and further indicates that the other structural components of the tissue are unlikely to be damaged. Because retention of the structural components of orthopedic grafts is vital to their clinical usefulness, limiting damage to them during processing is clearly important (Fig. 7).

Studies in the CryoLife laboratory demonstrate the retention of cellular viability in orthopedic allografts following a carefully designed cryopreservation regimen.^[20,21] Using a method to measure cellular protein synthesis by labeling with tritiated glycine,^[20] human



Fig. 6 Human Achilles tendon allograft prepared for implant. (Figure by CryoLife, Inc.)



Fig. 7 Human knee meniscus allograft prepared for implant. (Figure by CryoLife, Inc.)

tendon and meniscal grafts have been assessed before and after cryopreservation. Results show that cryopreserved patellar tendons retained 86% viability and cryopreserved meniscus retained 50% viability compared to fresh. Studies utilizing the previously described live/dead viability assay^[15,22] confirmed these earlier findings, cryopreserved menisci demonstrated an average of $54\% \pm 16$ viable cells (range 35 to 81%) post cryopreservation, with storage times ranging as long as 7 years. This group of menisci was compared to a set of menisci prepared using a fresh/frozen method and showed the fresh/frozen grafts to be nonviable (average viability $3\% \pm 8$, range 0–19%) even with only short-term storage (2 weeks to 4 months).^[21] Fresh/frozen graft processing methods can vary depending on the processor. In general, all fresh/frozen processing is conducted without cryoprotectants or control of freezing temperatures, and storage is generally at temperatures above -80°C . All grafts processed by fresh/frozen methods are considered to be nonviable. As suggested earlier, the assessment of cellular viability is a relevant measure to detect damage to the tissue during processing. If cells remain viable, there is more confidence that the critical structural components of the tissue matrix are undamaged during processing and storage.

ALLOGRAFT: APPRECIATION OF THE SOURCE

This article would not be complete without a discussion of the source of organs and tissues. The loss of a loved one—leading to the donation of organs and tissues—is nothing less than tragic. The unselfish gift from the donor and donor family, however, can save

or enhance the lives of up to 50 critically ill individuals. It is the most generous gift a family can give.

The process of organ and tissue donation is complex and depends on a national network of professionals dedicated to the procurement, processing, and distribution of tissue. The national system for whole organ distribution, United Network for Organ Sharing (UNOS), is well known, with some 75,000 to 80,000 people waiting to receive the gift of life.

Approximately 5400 heart valve implants are performed each year, 60% of which are received by children under 15 years of age.^[8] The need is estimated to be twice the existing supply. The same need applies to all other allografts: Cornea, skin, vascular,^[23] and orthopedic allograft needs are all greater than their supply. This is why organ and tissue donation is a vital national need, and why thanks and support should go to all donors and donor families for the special gifts they give. The special charge to participants in the allograft processing system is to continue to advance the technology of tissue processing and preservation in order to increase the opportunity for maximum good to come from tissue donation.

CONCLUSION

Many tissues (like heart valves, blood vessels, and orthopedic tissue) undergo significant mechanical activity over their lifetimes. This requires durability, obtained through the unique blend of natural collagens, elastins, and other matrix proteins and their cellular constituents, which play a mechanical/structural role and a care/maintenance role.

As we continue to understand more about the complex biological and physiological factors from which we are formed, it becomes clear that the most appropriate way to replace diseased or damaged tissues is by developing substitutes that are as close as possible to the original tissue.

ARTICLES OF FURTHER INTEREST

Biomedical Ethics; Burn Dressing; Cartilage Regeneration; Extracellular Matrix Scaffolds; Fibrin Sealants; Heart Valve, Bioprosthetic; Tissue Engineering of Skin; Tissue Engineering of Tendons; Xenografts

REFERENCES

1. *Biomaterials in Reconstructive Surgery*; Rubin, L.R., Ed.; The C.V. Mosby Company: St. Louis, MO, 1983.
2. Achauer, B.M.; Black, K.S.; Hewitt, C.W.; Furnas, D.W. Immunosurgery. In *Clinics in Plastic Surgery An International Quarterly*; Sasaki, G.H., Krizek, T.J., Eds.; W. B. Saunders: Philadelphia, PA, April 1985; Vol. 12 (No. 2).
3. Danilevicius, Z. S.S. Cosmas and Damian: The patron saints of medicine in art. *JAMA* **1967**, *201*, 1021–1025.
4. *Tissue Engineering Intelligence Unit 2: Composite Tissue Transplantation*; Hewitt, C.W., Black, K.S., Eds.; R.G. Landes Company: Austin, TX, 1999.
5. *Cyclosporine Volume I—Biological Activity and Clinical Applications*; Kahan B. D., Ed., A Transplantation Proceedings Reprint of Supplement 1, Grune & Stratton, Inc.: Orlando, FL, 1983; Vol. XV. 1984.
6. Tregear, R.T. Physical Functions of the Skin. In *Theoretical and Experimental Biology*; Academic Press: New York, 1966; Vol. 5.
7. Thubrikar, M.J. *The Aortic Valve*; CRC Press, December 1989.
8. *CryoLife Clinical Experience*; CryoLife, Inc., 2002. Data on file.
9. *Organ and Tissue Donation for Transplantation*; Chapman, J.R., Deierhoi, M., Wight, C., Eds.; Arnold (member of Hodder Headline Group): London, 1997.
10. O'Brien, M.F.; Stafford, E.G.; Gardner, A.H.; Pohlner, P.; McGiffin, D.C.; Johnston, N.; Texas, P.; Brosnan, A.; Duffy, P. Cryopreserved Viable Allograft Aortic Valves. In *Cardiac Valve Allografts 1962–1987*; Yankah, Miller, Ross, Somerville, Yacoub, Eds.; Springer: New York, 1988; 311–321.
11. O'Brien, M.F.; Harrocks, S.; Stafford, E.G.; Gardner, M.A.; Pohlner, P.G.; Tesar, P.J.; Stephens, F. The homograft aortic valve: A 29-year, 99.3% follow up of 1022 valve replacements. *J. Heart Valve Dis.* **2001**, *10* (3), 334–345.
12. Woloszyn, D.; Johnson, D.; Yacoub, M.H. Homograft Viability, Assessment and Significance. In *Cardiac Valve Allografts: Science and Practice*; Springer: New York, 1997; 11–21.
13. American Association of Tissue Banks *Standards for Tissue Banking*; Hornicek, Woll, Kasprisin, Eds.; American Association of Tissue Banks, 2002.
14. Parker, R. An International Survey of Allograft Banks. In *Cardiac Valve Allografts: Science and Practice*; Springer: New York, 1997; 5–9.
15. Daly, C.J.; Gordon, J.F.; McGrath, J.C. The use of fluorescent nuclear dyes for the study of blood vessel structure and function: Novel applications of existing techniques. *J. Vasc. Res.* **1992**, *29*, 41–48.
16. Doty, J.R.; Salazar, J.D.; Liddicoat, J.R.; Flores, J.H.; Doty, D.B. Aortic valve replacement with cryopreserved aortic allograft: Ten-year experience. *J. Thorac. Cardiovasc. Surg.* **1998**, *115*, 371–380.
17. *CryoGraft Meniscus Allograft Reconstruction Clinical Results*; CryoLife, Inc., 2000. Data on file.
18. Carter, T.R. Meniscal allograft transplantation. *Sports Med. Arthrosc. Rev.* **1999**, *7* (1), 51–62.
19. Garrett, J.C. Meniscal transplantation: A review of 43 cases with 2- to 7-year follow-up. *Sports Med. Arthrosc. Rev.* **1993**, *1* (2), 164–167.
20. Brockbank, K.G.M.; McCaa, C.; Dawson, P.E.; Milton, J.L. Warm ischemia limits: Protein synthesis/

- viability of cryopreserved menisci. *J. Transpl. Coord.* **1991**, *1*, 121–123.
21. Dawson, P.E.; Schorr, A.J. In *Demonstration of Superior Cellular Viability in Cryopreserved Meniscus*, 22nd Annual Mtg. American Association of Tissue Banks, New Orleans, LA, 1998.
 22. Yang, H.; Acker, J.; Chen, J.; McGann, L. In situ assessment of cell viability. *Cell Transplant* **1998**, *7* (5), 443–451.
 23. *CryoVein Saphenous Vein Allograft Infringuinal Bypass Clinical Experience*; CryoLife, Inc., 2001. Data on file.

Alumina

A

Jae Sung Lee

Department of Chemical Engineering, Pohang University of Science and Technology,
Pohang, Korea

INTRODUCTION

Aluminum oxide (Al_2O_3), or alumina, has been used since 1970 as a bulk material in the manufacture of components of prosthesis, and is the inert ceramic most commonly used for surgical implants. It is found in nature as corundum in emery, topaz, amethyst, and emerald, and as the precious gemstones ruby and sapphire. However, it is from the more abundant ores such as bauxite and cryolite and from clays that the material is commercially extracted and purified.

PROCESSING ALUMINA FOR MODERN APPLICATION

The most common process for the extraction and purification of alumina is the Bayer process. The first step in the process is to mix ground bauxite into a solution of sodium hydroxide (digestion). Sodium oxide is thus a common impurity in alumina. By applying steam and pressure in tanks containing the mixture, the bauxite slowly dissolves. The released alumina reacts with sodium hydroxide to form sodium aluminate. After the contents of the tank have passed through other vessels where pressure and temperature are reduced and impurities removed, the solution of sodium aluminate is placed in a special tank where the alumina is precipitated out. The precipitate is removed from the tank, washed, and heated in a kiln to evaporate any water present. The residue is a commercially pure alumina.

For advanced ceramics uses, the alumina manufactured by these processes requires further purification. This is often achieved by recrystallization from ammonium alum. Alumina exists in many crystal phases— α , δ , γ , η , θ , ρ , and χ —depending on the heat-treatment conditions of aluminum hydroxide or aluminum oxyhydroxide. As a biomaterial, α -alumina—or its natural form, corundum—is used, which is the most thermodynamically stable form. Most alumina devices are very fine-grained polycrystalline α - Al_2O_3 produced by pressing and sintering at 1600–1800°C. A very small amount of other oxides (SiO_2 , MgO , or CaO) is used to aid sintering and limit grain growth during sintering. The single-crystal form of alumina has been successfully used to make implants;

this material may be made by feeding fine alumina powders onto the surface of a seed crystal that is slowly withdrawn from the electric arc or oxygen-hydrogen flame as the fused powders build up.

Polycrystalline α - Al_2O_3 has a rhombohedral structure ($a = 4.758 \text{ \AA}$ and $c = 12.991 \text{ \AA}$). The close packing of the aluminum and oxygen atoms within this structure leads to its good mechanical and thermal properties.

Alumina is a highly inert material and resistant to most corrosive environments, including the highly dynamic environment of the human body. Under physiological conditions, alumina is also extremely unreactive and is classed as nearly inert, eliciting little if any response from surrounding tissues and remaining essentially unchanged after many years of service. However, the body does recognize it as a foreign material and attempts to isolate it by forming a layer of nonadherent fibrous tissue around the implant where possible. Properties making alumina suited to biomaterial applications include its high degree of chemical inertness under physiological conditions, excellent wear resistance, ability to be polished to a high surface finish, and excellent hardness. These properties are exploited for implant purposes, where alumina is used as an articulating surface in hip and knee joints.^[1–5] Its ability to be polished to a high surface finish makes it an ideal candidate for this wear application, where alumina operates against materials such as ultrahigh-molecular-weight polyethylene (UHMWPE) and alumina itself. In addition, the bioinertness of alumina has led to improved devices in nearly load-free applications in ear, nose, and throat (ENT) surgery and in some soft tissue replacement. Porous alumina has also been used as a bone spacer, where sections of bone have been removed due to disease. In this application, it acts as a scaffold for bone ingrowth. Single-crystal alumina or sapphire has also been used in dental applications, although alumina's use in this application is declining with the advent of more advanced materials such as resin-based composites.

PHYSICAL PROPERTIES

The polycrystalline and single-crystal aluminas are dense, nonporous, and nearly inert. They are attached

Table 1 Physical properties of alumina bioceramics

	Commercial alumina bioceramics	Standards: ISO 6474 (2nd ed., 1994) ASTM F603-83 DIN 58 8353
Alumina content (% by weight)	>99.7	>99.5
Impurities, SiO ₂ + CaO+Na ₂ O (%)	<0.05	<0.1
Additives (%)	0.25	<0.3
Density (g/cm ³)	>3.98	>3.94
Average grain size (μm)	2.2	<4.5
Surface roughness value (μm)	0.02	–
Vickers hardness number (VHN)	2300	>2000
Compressive strength (MPa)	4500	4000
Flexural strength (MPa)	580 ^a	>250 ^b
Young's modulus (GPa)	380	–
Fracture toughness (Mpa m ^{1/2})	4	–
Impact strength (cm MPa)	>40	>40
Wear resistance ^c (mm ³ h ⁻¹)	0.001	0.01
Corrosion resistance ^c (mgm ⁻² day ⁻¹)	<0.1	<0.1

^aFour-point bending test.

^bBiaxial test.

^cWear and corrosion resistance values refer to particular test arrangements.

to surface irregularities by bone growth, by cementing the implant into the tissues, or by press-fitting into a defect (termed morphological fixation). Alumina is extremely hard and scratch-resistant, second only to diamond. It has excellent corrosion resistance in body fluids. Physical properties of alumina such as strength, fatigue resistance, and the fracture toughness of polycrystalline α -Al₂O₃ are a function of grain size, porosity, and purity.^[1] Alumina, with an average grain size of <4 μm and >99.7% purity, exhibits good flexural strength and excellent compressive strength. These and other physical properties are summarized in Table 1, for a commercially available implant material. Table 1 also lists International Standards Organization (ISO) requirements for alumina implants. Essentially the same requirements are stipulated in American Society for Testing and Materials (ASTM) F603-83 and Deutsches Institut für Normung (DIN) 58 8353. The first alumina standard for medical applications (ISO 6474) was defined in 1979; the 2nd edition, issued in 1994, addresses properties required for high-load applications.^[6] In addition to imposing more stringent requirements in density and grain size, the 1994 standard strengthens impurity requirements for SiO₂ and alkali metal oxides (mainly Na₂O). A small amount of MgO (<0.5%) is used as a grain growth inhibitor and is essential in order to achieve a fully dense sintered body with fine-grain microstructure. CaO impurities favor alumina ceramic fatigue and aging effects; SiO₂ and alkali metals impede densification and promote grain growth. Hence these

oxides must carefully be kept at very low concentration (<0.05%) for medical use. Since 1989, the main commercial products of alumina—ceramic heads and hip sockets—have exceeded the requirements of the standard. It has been shown that alumina implants that meet or exceed ISO standards have excellent resistance to dynamic and impact fatigue, and also resist subcritical crack growth.^[7] However, an increase in average grain size to >7 μm can decrease mechanical properties by about 20%.^[1] In general, the smaller the grains and the lower the porosity, the higher the strength that results.^[8] Although alumina has the highest mechanical properties among biomedical ceramics, its tensile properties are still lower than those of metallic biomaterials. Results from aging and fatigue studies show that it is essential that alumina implants be produced at the highest possible standards of quality assurance, especially if they are to be used as orthopedic prostheses in younger patients.^[9] In some cases, alumina can experience a time-dependent decrease in strength during immersion in saline in vitro and after implantation, probably due to a preferential dissolution of impurities that results in crack propagation.

WEAR

High-density alumina exhibits a high surface wettability, leading to a very low coefficient of friction when sliding against itself or polyethylene. These

excellent wetting characteristics are due to an outer surface of alumina nonsaturated with oxygen ions, which adsorbs a monomolecular film of water and biological molecules.^[2] This adsorbed layer limits the direct contact of the articulating solid surfaces. The high surface energy of alumina—a value measured at 2.64 J/m² for α -form^[10]—is demonstrated through contact angle measurements.

The low coefficient of friction and wear rates of alumina occur only when the grains are very small (<4 μm) and have a very narrow size distribution. These conditions lead to very low surface roughness values (<0.02 μm). If large grains are present, they can pull out and lead to very rapid wear of bearing surfaces, owing to local dry friction. Any roughness on alumina is inverse (due to pullout of grains), in contrast to the protruding asperities encountered with metal carbides. The concave surface roughness is an important factor contributing to the outstanding wear properties of alumina ceramics.^[2]

For alumina on load-bearing, wearing surfaces (as in hip prostheses), macrogeometry, represented by sphericity and circularity, is also important. Alumina balls must have a very high degree of sphericity (<1 μm), produced by polishing and grinding the two mating surfaces together. For example, the alumina ball and socket in a hip prosthesis are polished together and used as a matched pair. The long-term coefficient of friction of an alumina/alumina joint decreases with time and approaches the values of a normal joint (ca. 0.01 m³/h). This leads to wear on alumina-articulating surfaces being nearly 10 times less than wear on metal-polyethylene surfaces.^[1]

BIOCOMPATIBILITY

Because all materials elicit some response from living tissues, no materials implanted in a living host are inert. Alumina is classified as nearly inert because it is nontoxic and biologically inactive.^[11] Nothing goes into solution that would biochemically disturb the normal proliferation of adjacent cells. Because the surfaces of alumina implants are covered in or coated by biomolecules, within about a minute after immersion into body fluid, the implants become camouflaged against the host's immune system.^[3,5] The interface between ceramic and bone is not chemically or biologically bonded, and there is relative movement. Progressive development of a fibrous capsule occurs in soft and hard tissues with variable thickness depending on relative motion. This contact osteogenesis leads to secondary stabilization by bony ingrowth. The sequence of events in interface remodeling is essentially the same as occurs in fracture healing. Although alumina is the most biocompatible material in bulk form,

alumina particles can induce a considerable foreign-body reaction when produced in massive quantity. However, among all the wear particles derived from prosthetic components, alumina particles of very small size (0.0005–0.002 mm) exhibit the highest biocompatibility and lowest fibrogenic potential.

When alumina is used as a ceramic coating by the air-plasma-spray application, the coating is also inert and entirely compatible with bone. Ionic diffusion from alumina is considered negligible.^[12]

Biocompatibility testing of implantation materials is regulated by ISO 10993 and other standards ordained by bodies such as the European EN30993, the U.S. Food and Drug Administration (FDA), and the Japanese Ministry of Health and Welfare (MHW). These agencies share a common approach to testing for biological effects such as cytotoxicity, sensitization, genotoxicity, implantation, chronic toxicity, and carcinogenicity.^[13]

APPLICATIONS

Articulating Surfaces in Joint Replacements

Due to its ability to be polished to a high surface finish and its excellent wear resistance, alumina is often used for wear surfaces in joint replacement prostheses. Such applications include femoral heads for hip replacements and wear plates in knee replacements. More than 100,000 hip prostheses with alumina balls for the femoral head components are implanted annually.^[14] Figure 1 shows a schematic of total hip replacement (THR) surgery.^[15] In hip replacements, the alumina femoral head is used in conjunction with a metallic femoral stem and an acetabular cup made from ultrahigh-molecular-weight polyethylene (UHMWPE) or alumina itself for the opposing articulating surface. Figure 2 shows a modern ball-and-socket system for an artificial hip joint.^[16]

In joint replacements, success depends on two aspects: 1) the frictional and wear behavior of the materials; and 2) the reliability of the anchorage of the components in their bony environment. As shown in Table 2, wear rates for alumina on UHMWPE are as much as 20 times lower than those for metal on UHMWPE, producing less tribological debris and making this combination far superior.^[2] Wear debris leads to complications such as osteolysis in surrounding tissues; keeping it to a minimum is advantageous. The production of particulate wear debris from implant materials (polyethylene in particular) and subsequent osteolysis have been recognized as the major cause of long-term failure in total hip replacement. Due to their low wear rate, noncemented alumina cups press-fitted into the acetabulum of the hip

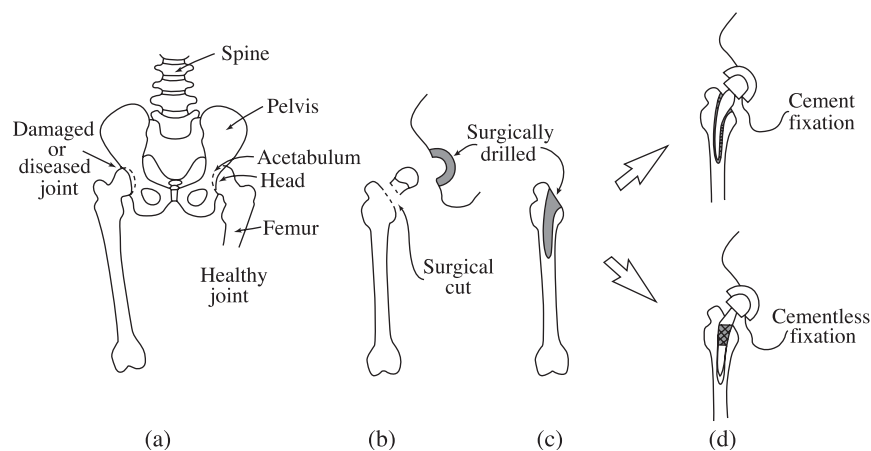


Fig. 1 Schematic of total hip replacement (THR) surgery. In general, the femoral implant stem is anchored to the bone by either a thin layer (a few mm thick) of polymethylmethacrylate (PMMA) cement or a cementless system involving a snug fit of the stem in the femoral shaft. In typical cementless fixation, the upper third of the stem is covered with a porous coating of sintered metal alloy beads or ceramics. Bone growth into the porous surface provides mechanical anchoring. (From Ref. [15].)

are widely used. The cups are stabilized by the growth of bone into grooves or around pegs.^[1] The mating femoral ball surface, bonded to a titanium alloy or cobalt-chrome alloy stem, can also be made of alumina. Typical diameters for femoral heads are 32 mm and 28 mm and lower. Well-designed tapers are used for their fixation.

An alumina/alumina ceramic bearing couple is theoretically superior to alumina/UHMWPE or metal/UHMWPE as an articulating surface for a total joint system. Use of alumina not only eliminates polyethylene from the device system, but its extremely low coefficient of friction and potential for far superior wear resistance are very advantageous. Polyethylene debris is known to be particularly responsible for the inflammatory response that leads to bone resorption and loosening of implants over time.^[17] In the case of

alumina, however, the usual biologic event is a fibrocytic reaction with very few macrophages and no giant cells^[16,18] unless there is metallic debris or a very large amount of ceramic debris generated by abnormal contact, or following a long period of component loosening.^[19] Experience with retrievals from ceramic implants indicates that ceramic debris is less reactive than polyethylene or metal debris.

However, alumina/alumina bearings have experienced early design failures due mostly to inferior material properties. Since 1985, manufacturers have introduced alumina with small grain sizes that eliminate mechanical failure under physiological loads. Improvements such as hot isotactic pressing (HIP) have led to alumina ceramics of nearly theoretical density (3.97 g/cm^3) that have much reduced grain size, fewer inclusions, and limited grain boundaries, all of which have led to a great increase in the toughness and burst strength of the material.^[12]

By the late 1960s, the Charnley hip implant system with metallic femoral stem and UHMWPE acetabular component had become the primary design for successful total hip arthroplasty (THA). Dense alumina ceramic was introduced in the 1970s in Europe as a potentially improved bearing surface for THA. Figure 3 compares long-term clinical test results in terms of the

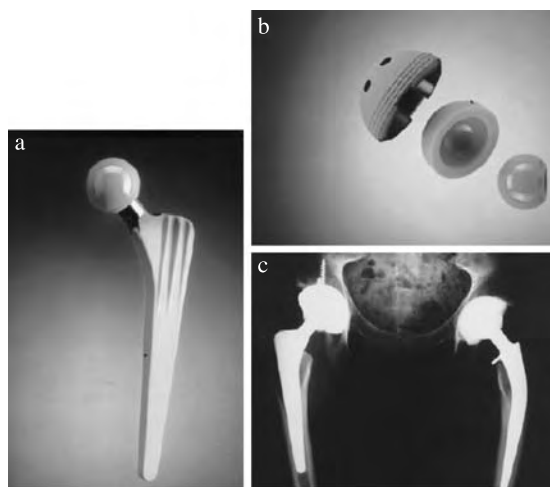


Fig. 2 Modern design of ball-and-socket bearing system for artificial hip joint. (a) Fully hydroxyapatite (HA)-coated cementless stem and alumina ball; (b) cementless titanium shell fully coated with HA, with an alumina liner; (c) one-year follow-up. (From Ref. [16].)

Table 2 Average annual wear rates of articulating surfaces in total hip prostheses

Materials	Wear rates ($\mu\text{m}/\text{year}$)
Co-Cr-Mo alloy/UHMWPE	200
Alumina/UHMWPE	20–130 ^a
Alumina/alumina	2

^a20 ($\mu\text{m}/\text{year}$) refers to observations of explanted components and includes only wear measurements. 130 ($\mu\text{m}/\text{year}$) refers to penetration measurements between head and acetabular components observed by means of radiographic analysis, and includes plastic flow in addition to wear.

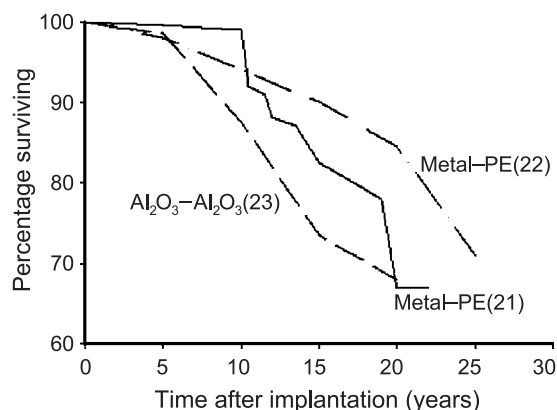


Fig. 3 Survivability curves for total hip arthroplasty. Numbers in parentheses indicate the references containing data on which the curves are based.

Kaplan-Meier survivability curves.^[20] For Charnley THAs with metal/PE bearing, a summary of the 42 published clinical studies^[21] is shown as a single curve together with a curve derived from a recent report on 25-year follow-up evaluations for 2000 consecutive primary THAs performed at one institution from 1966–1971.^[22] Such long-term data for alumina/alumina THAs are rare, and their reliability is hampered by the poor quality of early alumina ceramic materials (previously discussed). In any case, the curve in Fig. 3 for alumina/alumina bearings was constructed from a recent report on 20-year follow-up evaluations for 118 consecutive THAs between 1979 and 1980.^[23] The prostheses combined a 32-mm alumina head and an all-alumina socket. Both cemented and uncemented acetabular femoral components were used.

It is notable that the rate of alumina/alumina implant survivorship is satisfactory, yet significantly lower than that of Charnley THAs with similar duration of follow-up. However, a very low rate of acetabular surface wear (<0.025 mm/yr) was identified and little osteolysis was seen. This demonstrates that the limited amount of ceramic debris is generated by the bearing surface. Thus the failures observed in this series do not seem to be directly attributable to the alumina-on-alumina design, but rather to implant fixation (component loosening). The problem may be improved by newer implant designs such as modular, metal-backed, porous-coated, uncemented acetabular components combined with improved femoral designs. Alumina/UHMWPE prostheses are reported to have slightly higher survivability than alumina/alumina bearings, but the number of implants tested was too small (<100 implants) to have statistical significance.^[21]

In total knee replacements, alumina is used in portions coming into contact with bone, and a combination of alumina and UHMWPE is used in sliding



Fig. 4 Cementless total knee prosthesis made by a combination of alumina with UHMWPE. (From Ref. [24].)

portions. The stem is positioned at the center of the tibial plate and the load is transmitted through the stem to the cortical bone in the posterior portion of the tibia (Fig. 4). To improve the anchoring property, the tibial plate is metal-backed and both femoral and tibial components are cemented.^[24] In total knee arthroplasty, problems caused by polyethylene debris are more severe than in hip surgery because of the different curvatures of the femoral condyle and the top of the tibial plateau. The search for further improvement in materials and design therefore continues.

The cementless alumina ceramic total ankle prosthesis employs alumina ceramics for the interface with the bone and a combination of UHMWPE with alumina ceramics for the sliding part.^[25] The tibial component is designed to be loaded on the cortical bone at the anterior part of the tibia and on the cancellous and cortical bones at the posterior part.

Medical-grade alumina has a Young's modulus of 380–420 GPa, which is much higher than that of cancellous bone (0.05–0.5 GPa depending on location and age) or cortical bone (7–25 GPa). There is thus a great mismatch in elastic modulus between a bone and an implant, and the alumina implant consequently shields the bone from mechanical loading, allowing nearly all the mechanical load to be carried by the implant. Living bone must be under a certain amount of tensile load in order to remain healthy. If it is unloaded or is loaded in compression, the resulting stress shielding will cause biological changes, leading to resorption and weakening of the bone and to deterioration of the implant-bone interface.^[9] This limits the effectiveness of alumina as a bone interface material.

Bone Spacers

Porous alumina may also be used in bone spacers to replace large sections of bone that have been removed for a reason such as cancer. Bone spacers may take the

shape of concentric rings around a metallic pin, inserted up the center of the remaining native bone. The porous nature of these implants allows new bone to grow into the pores, effectively using the alumina as a scaffold for new bone formation. Here, load bearing is not the primary concern. When pore size exceeds $100\ \mu\text{m}$, bone will grow within the interconnecting pore channels near the surface and maintain its vascularity and long-term viability (termed biological fixation). In this manner, the implant serves as a structural bridge or scaffold for bone formation.^[1]

Porous ceramics for bone implants—including alumina, calcium phosphates, and hydroxyapatite—are produced by a replamineform (or hydrothermal exchange) process in which the microstructures of certain corals (*Porities* or *Goniopora*) are used to cast the porous ceramic.^[26] Porous alumina may also be made from a suitable foaming agent that evolves gases upon heating. A mixture of powdered calcium carbonate and fine alumina powder can thus be fired at $1450\text{--}1500^\circ\text{C}$ for 20 hours to obtain a foamed material whose pore size and volume fraction (33–48% of porosity) are determined by the size and concentration of the original calcium carbonate particles.

Porous alumina is weaker than the bulk form, with the material rapidly weakening as porosity increases.^[1] Because much surface area is exposed, the environment becomes much more important for porous than for nonporous, dense material. The aging of porous alumina, with subsequent decrease in strength, requires bone ingrowth to stabilize the structure of the implant.

Dental Applications

Dense alumina has been used in dental applications,^[3–5] specifically for replacement teeth. They are used immediately or shortly after extraction (and after a reshaping of the alveoli to achieve the press-fit necessary for an undisturbed period of bone healing), and later as implants in edentulous regions of the jaw. Polycrystalline alumina can be fractured when inserted as an artificial small-diameter dental root. Therefore, in many cases single-crystal alumina or sapphire is used, with bending strength ($13,000\ \text{kg}/\text{cm}^2$) three times higher than polycrystalline alumina ($3,500\ \text{kg}/\text{cm}^2$). Compressive strength ($30,000$ vs. $10,000\ \text{kg}/\text{cm}^2$) and elastic modulus (385 vs. $371 \times 10^4\ \text{kg}/\text{cm}^2$) are also better for single-crystal alumina. It has demonstrated compatibility suitable with alveolar bone and gingival epithelium.

Designs of single-crystal alumina implants have improved since 1975, and their clinical applications have widened.^[27] They have been manufactured as endodontic endosteal implants, endosteal implants, and intramucosal inserts. Porous alumina implants are

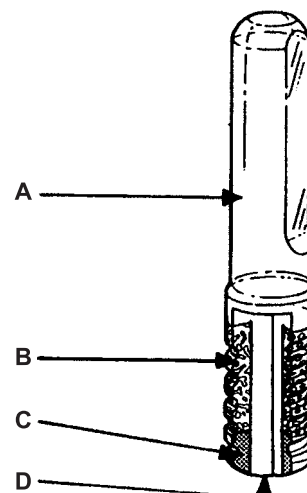


Fig. 5 Basic design for porous alumina dental implant. (A) Cervical portion; (B, C,) porous alumina layer and apex portion, made of polycrystalline alumina; (D) screw designed for self-tapping. Root core is made of single crystal alumina. (From Ref. [28].)

made of single-crystal alumina cylindrical core, around which porous polycrystalline alumina is fused at high temperatures. A basic design is shown in Fig. 5. The single-crystal core makes alumina dental implants strong enough to bear occlusal forces.

There are some disadvantages to basing the design of dental implants on alumina: It cannot be customized easily and cannot be bent like metal; it has a high modulus of elasticity compared to bone, and a wide radiolucent zone thus develops around the implants, owing to bone resorption and increased connective tissue at the implant-bone interface. However, if the implant has suitable loading of functional stress, it can maintain functional fixation with biodynamic responses of resorption and remodeling. Alumina dental implants are currently declining in popularity, being replaced by material such as dental porcelain.

Other Applications

Alumina ceramics have been clinically tested in prostheses for elbows, shoulders, wrists, and fingers with success equal to or better than other material systems.^[2,3,5] Alumina ceramics have been used in ENT applications and in maxillofacial surgery.^[28] Extended animal experiments have shown that middle ear mucosa proliferates normally on the surface of dense, pure, alumina ceramic. Since the early 1980s, total and partial ossicular replacements made of alumina ceramic have stood the test of clinical experience, showing improved success compared to plastic components. One advantage of alumina ceramic implants is the possibility of achieving a highly precise fit by careful

grinding. (The procedure requires adjustment of the implants' length and shape for each individual case.) In treating throat cancer patients, the cartilage that prevents the collapse of the trachea may need to be removed, necessitating the insertion of some support structure to keep the trachea open. Trachea-supporting rings made of alumina have shown better results than previously used polymeric rings due to alumina's higher degree of inertness.^[29] Alumina implants have been used for neurosurgical operations such as cranioplasty to repair bone defects in both the cortical convex and suboccipital regions, and to reconstruct the cellar floor and orbital wall.^[30] Orbital support plates have also proven clinically successful and are in routine use.

A soft-tissue application of dense, pure alumina ceramics is in keratoprotheses (corneal replacement). This routinely used eye prosthesis consists of a corundum single crystal as the optical part and an alumina ceramic holding ring.^[31] Small porous collars of alumina or aluminum-calcium phosphate serve as controlled drug-delivery systems to deliver hormones, vaccines, and drugs in a controlled, sustained manner over long periods of time in amounts that are effective, yet not toxic to the host.^[32]

A NEW ALUMINA MATRIX COMPOSITE

Alumina has been successfully used in orthopedic implants for the last 30 years because it is a biocompatible, inert structural material suitable for use in load-bearing joints. Because alumina has limited strength, the application remains limited to designs with sufficient wall thickness. It is relatively brittle, so can fracture; its application is therefore restricted to designs involving compressive loading or limited tensile loads. It is susceptible to slow crack growth and undergoes slight reductions in fracture strength with time, depending on the physiological environment.

Zirconia has higher strength than alumina, with excellent biocompatibility and wear resistance. However, it shows poor hydrothermal stability during sterilization and *in vivo*. Zirconia-toughened alumina (ZTA) has been used in mechanical and engineering applications for more than 20 years. New materials based on mixed zirconia and alumina have recently been proposed as biomaterials that may combine the positive effects of alumina and zirconia ceramics, while avoiding the drawbacks of zirconia. A new material—transformation-toughened and platelet-reinforced alumina (TTPA)—promises zirconia's excellent mechanical performance, but not the problems of hydrothermally-induced degradation that attend the use of alumina matrix material.^[33] The material consists of approximately 75% alumina and 25% zirconia,

combined with small amounts of additives. Chromium oxide forms a solid solution with alumina, and increases hardness and wear resistance. Controlled introduction of strontium oxide can achieve the *in situ* generation of platelet strontiumaluminate particles in the matrix during the sintering process, affording these particles excellent fixation in the alumina matrix.

The new material combines the properties of ceramic and metal, and surpasses the properties of alumina and zirconia in many aspects. It is extremely hard, wear-resistant, and well-tolerated by the body, while demonstrating good fracture toughness and flexural strength. These improved properties could redefine the design limits of implants' diameter and wall thickness. Applications like knee components and finger joints are now technically feasible. The excellent clinical performance of alumina ceramics in hip arthroplasty can thus now be transferred to applications whose existing components fall short of surgeons' expectations.

CONCLUSION

Highly pure, dense alumina is the main bioinert ceramic material used in medical prostheses and has withstood long-term applications in hip surgery and dental implantology. Although some dental implants are single-crystal sapphire, most alumina devices are very fine-grained α - Al_2O_3 produced by pressing and sintering at temperatures ranging from 1600 to 1800°C, depending on the properties of the raw materials. Its wide use in load-bearing applications is based on the combination of excellent corrosion resistance, good biocompatibility, high wear resistance, and high compressive strength. Advances in technology that have produced high-quality alumina enabled alumina/UHMWPE and alumina/alumina couplings to become viable alternatives to conventional metal/UHMWPE designs, especially for young and more active patients. Alumina for articulating surfaces in total joint replacement has the tribological properties most capable of being capitalized into the best clinical results.

However, the phenomena of slow crack growth, static and cyclic fracture, low toughness, deterioration over time in the physiological environment, and sensitivity to tensile strength are serious concerns in high-load-bearing applications. The high modulus of alumina also limits its effectiveness as bone interface material. The recent development of a new alumina matrix composite may extend the useful lifetime of artificial joints from the current 10 years to a much longer period, and may open the door for new applications. It is anticipated that the use of alumina ceramics in total joint prostheses and otologic,

maxillofacial, and dental applications will continue to grow.

ACKNOWLEDGMENTS

This work has been supported by National R & D project for Nano Science & Technology, and Research Center for Energy Conversion and Storage.

ARTICLES OF FURTHER INTEREST

Animal Models for Bone Tissue Engineering; Biologic and Synthetic Apatites; Bone-Implant Interface; Bone Remodeling; Ceramics; Ceramics in Dentistry; Composites; Dental Implants; Implant, Total Hip; Knee Joint Replacement; Orthopedic Biomaterials; Ultra-High Molecular Weight Polyethylene (UHMWPE); Wear Debris

REFERENCES

- Hench, L.L. Ceramics, Glasses, and Glass-Ceramics. In *Biomaterials Science, An Introduction to Materials in Medicine*; Ratner, B.D., Hoffman, A.S., Schoen, F.J., Lemons, J.E., Eds.; Academic Press: San Diego, 1996; 73–84.
- Hulbert, S.F. The Use of Alumina and Zirconia in Surgical Implants. In *An Introduction to Bioceramics*; Hench, L.L., Wilson, J., Eds.; World Scientific: Singapore, 1993; 25–40.
- Heimke, G. Use of Alumina Ceramics in Medicine. In *Bioceramics*; Hulbert, J.E., Hulbert, S.F., Eds.; Rose-Hulman Institute of Technology: Terre Haute, 1992; Vol. 3, 19–30.
- Park, J.B. Aluminum Oxide: Biomedical Applications. In *Concise Encyclopedia of Advanced Ceramic Materials*; Brook, R.J., Ed.; Pergamon Press: Oxford, 1991; 13–16.
- Heimke, G. Aluminum Oxide. In *Concise Encyclopedia of Medical & Dental Materials Advanced*; Williams, D., Ed.; Pergamon Press: Oxford, 1990; 28–34.
- Walter, A. On the material and tribology of alumina-alumina couplings for hip joint prosthesis. *Clin. Orthop.* **1992**, *282*, 31–46.
- Doerre, E.; Dawihl, W. Ceramic Hip Endoprosthesis. In *Mechanical Properties of Biomaterials*; Hastings, G.W., Williams, D., Eds.; Wiley: New York, 1980; 113–127.
- Park, J.B.; Lakes, R.S. *Biomaterials—An Introduction*, 2nd Ed.; Plenum: New York, 1992.
- Hench, L.L. Bioceramics: From concept to clinic. *J. Am. Ceram. Soc.* **1991**, *74*, 1487–1510.
- McHale, J.M.; Auroux, A.; Perrotta, A.J.; Navrotsky, A. Surface energies and thermodynamic phase stability in nanocrystalline aluminas. *Science* **1997**, *277*, 788–791.
- Christel, P. Biocompatibility of surgical-grade sense polycrystalline alumina. *Clin. Orthop.* **1992**, *282*, 10–18.
- Toni, A.; Sudanese, A.; Terzi, S.; Tabarroni, M.; Calista, F.; Giunti, A. Ceramics in Total Hip Arthroplasty. In *Encyclopedic Handbook of Biomaterials and BioEngineering*; Wise, D.L., Trantolo, D.J., Altobelli, D.E., Yaszemski, M.J., Gresser, J.D., Schwarz, E.R., Eds.; Marcel Dekker: New York, 1995; Vol. 2; 1501–1544. Part A.
- Black, J. *Biological Performance of Materials—Fundamentals of Biocompatibility*; Dekker: New York, 1999.
- Schackelford, J.F. *Introduction to Materials Science for Engineers*, 2nd Ed.; Macmillan: New York, 1988.
- Christel, P.; Meunier, A.; Crolet, J.-M.; Sedal, L.; Witvoet, J. Aluminum Oxide Ceramic-Titanium Alloy Materials for Total Hip Replacement. In *Perspectives on Biomaterials*; Lin, O.C.C., Chao, E.Y.S., Eds.; Elsevier: Amsterdam, 1986; 277–288.
- Bizot, P.; Nizard, R.; Lerouge, S.; Prudhommeaux, F.; Sedel, L. Ceramic/ceramic total hip arthroplasty. *J. Orthop. Sci.* **2000**, *5*, 622–627.
- Al-Saffar, N.; Revell, P.A.; Kobayashi, A. Modulation of the phenotypic and functional properties of phagocytic macrophages by wear particles from orthopaedic implants. *J. Mater. Sci., Mater. Med.* **1997**, *8*, 641–648.
- Bierbaum, B.E.; Barsoum, W.K.; Puri, L.; Headrick, J.; Gomes, S. Alumina-Alumina Ceramic Bearings. In *Bioceramics in Hip Joint Replacement*; Willman, G., Zweymueller, K., Eds.; Georg Thieme Verlag: Stuttgart, 2000; 22–26.
- Yoon, T.R.; Rowe, S.M.; Jung, S.T.; Seon, K.J.; Maloney, W.J. Osteolysis in association with a total hip arthroplasty with ceramic bearing surfaces. *J. Bone Jt. Surg., Am.* **1998**, *80*, 1458–1467.
- Kaplan, E.L.; Meier, P. Nonparametric estimation from incomplete observations. *Am. Stat. Assoc. J.* **1958**, *53*, 457–481.
- Clinical Performance of Skeletal Prostheses*; Hench, L.L., Wilson, J., Eds.; Chapman & Hall: London, 1996.
- Berry, D.J.; Harmsen, W.S.; Cabanela, M.E.; Morrey, B.F. Twenty-five-year survivorship of two thousand consecutive primary Charnley total hip replacements. *J. Bone Jt. Surg., Am.* **2002**, *84-A*, 171–177.
- Hamadouche, M.; Boutine, P.; Daussange, J.; Bolander, M.E.; Sedel, L. Alumina-on-alumina total hip arthroplasty. *J. Bone Jt. Surg., Am.* **2002**, *84-A*, 69–71.
- Oonishi, H.; Amino, H.; Ueno, M.; Yunoki, H. Concepts and Designs for Total Hip and Knee Replacement. In *Reliability and Long-Term Results of Ceramics in Orthopaedics*; Sedal, L., Willman, G., Eds.; Georg Thieme Verlag: Stuttgart, 1999; 7–28.
- Oonishi, H. Knee and Ankle Joint Replacement. In *Osseo-integrated Implants, Vol. 1, Basics, Materials, and Joint Replacements, Bioceramics*; Heimke, G., Ed.; CRC Press: Boca Raton, 1990; 171.
- White, E.; Schors, E.C. Biomaterials aspects of inter-pore-200 porous hydroxyapatite. *Dent. Clin. North Am.* **1986**, *30*, 49–67.

27. Yamagami, A.; Kawahara, H. Single-Crystal Alumina (Sapphire) for Dental Implants. In *Encyclopedic Handbook of Biomaterials and BioEngineering*; Wise, D.L., Trantolo, D.J., Altobelli, D.E., Yaszemski, M.J., Gresser, J.D., Schwarz, E.R., Eds.; Marcel Dekker: New York, 1995; Vol. 2, 1617–1638. Part B.
28. Jahnke, K.; Schrader, M.; Silberzahn, J. Osseo-integrated Implants in Otorhinolaryngology. In *Osseo-integrated Implants, Vol. II, Implants in Oral and ENT Surgery*; Heimke, G., Ed.; CRC Press: Boca Raton, 1990; 287.
29. Zollner, C.; Weerda, H.; Strutz, J. Aluminiumoxid-Keramik als Stützgerüst in der Trachealchirurgie. In *Archives of Oto-Rhino-Laryngology*; Springer-Verlag: Dotdrecht, 1983; 214–216. Supplement II. .
30. Okudera, H.; Kobayashi, S.; Takemae, T. Neurosurgical Applications of Alumina Ceramic Implants. In *Bioceramics: Proceedings of the 1st Int'l Symp. on Ceramics in Medicine*; Oonishi, H., Aoki, Sawai, K., Eds.; Ishiyaku EuroAmerica Inc.: Tokyo, 1988; 320–325.
31. Polack, F.M. Clinical results with a ceramic keratoprosthesis. *Cornea* **1983**, 2, 185–196.
32. Buykx, W.J.; Grabaretk, E.; Reeves, K.D.; Anderson, R.; Mathivanar, R.; Skalsky, M. Development of Porous Ceramics for Drug Release and Other Applications. In *Bioceramics*; Hulbert, J.E., Hulbert, S.F., Eds.; Rose-Hulman Institute of Technology: Terre Haute, 1992; Vol. 3.
33. Rack, R.; Pfaff, H.-G. A New Ceramic Material for Orthopaedics. In *Bioceramics in Hip Joint Replacement*; Willman, G., Zweymueller, K., Eds.; Georg Thieme Verlag: Stuttgart, 2000; 141–145.

Angiogenesis Inhibitors

Shaker A. Mousa

Albany College of Medicine, and Albany College of Pharmacy, Albany, New York, U.S.A.

Ahmed S. Mousa

Cornell University, Ithaca, New York, U.S.A.

INTRODUCTION

Traditional cytotoxic therapies (chemotherapy and radiation therapy) used in the management of solid tumors are associated with significant therapeutic and safety limitations that impact effective disease control and patient survival. The successful administration of cytotoxic therapy can be limited by non-specific toxicity in normal tissues, with poor tolerance resulting in subtherapeutic dosing or treatment discontinuation. Furthermore, overlapping toxicities may affect the successful administration of concurrent chemotherapy and radiation therapy. Tumor resistance is a further obstacle to effective treatment with cytotoxics. These limitations have focused attention on treatment approaches that target tumor cells while limiting damage to healthy cells. Increasing evidence suggests that agents specifically targeting cellular receptors or other elements of the tumor microenvironment may be synergistic with cytotoxics.

The field of angiogenesis modulation is at a major crossroads. A tremendous advancement in basic science in this field is providing excellent support, though there is a lack of strong clinical support at this point with the exception of the recent data on anti-vascular endothelial growth factor (VEGF) as an adjunct to chemotherapy in solid tumor. Clinical outcome should include benefit/risk ratio, cost effectiveness, and hard end points including mortality, quality of life as opposed to tumor size reduction or decrease, and microvascular density with antiangiogenesis agents.

Experimental models should be used to provide guidance, placebo effect, comparative data, and mechanistic understanding but not for expected clinical efficacy.

BACKGROUND

In normal tissue, new blood vessels are formed during tissue growth and repair and during the development of the fetus during pregnancy. In cancerous tissue, tumors cannot grow or spread (metastasize) without the development of new blood vessels. Blood vessels

supply tissues with oxygen and nutrients necessary for survival and growth.^[1]

Endothelial cells, the cells that form the walls of blood vessels, are the source of new blood vessels and have a remarkable ability to divide and migrate. The creation of new blood vessels occurs by a series of sequential steps. An endothelial cell forming the wall of an existing small blood vessel (capillary) becomes activated, secretes enzymes that degrade the extracellular matrix (the surrounding tissue), invades the matrix, and begins dividing. Eventually, strings of new endothelial cells organize into hollow tubes, creating new networks of blood vessels that make tissue growth and repair possible.^[1-3]

Most of the time, endothelial cells lie dormant, but when needed, short bursts of blood vessel growth occur in localized parts of tissues. New capillary growth is tightly controlled by a finely tuned balance between factors that activate endothelial cell growth and those that inhibit it.^[1,2]

The production of new vasculature is a key requirement for tumor growth. Several oncogenes, including growth factors and their receptors, have a role in angiogenesis. Basic fibroblast growth factor (bFGF) and VEGF are key examples.

There is evidence that epidermal growth factor receptor (EGFR) and its ligands are involved in angiogenesis. Both epidermal growth factor (EGF) and transforming growth factor (TGF)-alpha have been shown to be attractants for a rat heart vascular endothelial cell line and to induce angiogenesis *in vivo* in a hamster model. In addition, it has been demonstrated that growth factor coexpression correlates with microvessel density in invasive breast cancer. These angiogenic factors can also interact in a regulatory fashion. For example, TGF-alpha promotes expression of VEGF, and inhibition of EGFR appears to decrease the angiogenic factors VEGF, bFGF, and interleukin (IL)-8, as well as microvessel density.

About 15 proteins are known to activate endothelial cell growth and movement, including angiogenin, epidermal growth factor, estrogen, fibroblast growth factors (acidic and basic), IL-8, prostaglandin E1 and E2, tumor necrosis factor-alpha, VEGF, and granulocyte colony-stimulating factor. Some of the

known inhibitors of angiogenesis include angiostatin, endostatin, interferons, IL-1, IL-12, retinoic acid, and tissue inhibitor of metalloproteinases.^[1,2,4]

At a critical point in the growth of a tumor, the tumor sends out signals to the nearby endothelial cells to activate new blood vessel growth. Two endothelial growth factors, VEGF and bFGF, are expressed by many tumors and seem to be important in sustaining tumor growth.^[1,2]

Angiogenesis is also related to metastasis. It is generally true that tumors with higher densities of blood vessels are more likely to metastasize and are correlated with poorer clinical outcomes. Also, the shedding of cells from the primary tumor begins only after the tumor has a full network of blood vessels. In addition, both angiogenesis and metastasis require matrix metalloproteinases (MMPs), enzymes that break down the surrounding tissue (the extracellular matrix), during blood vessel and tumor invasion.^[1,3]

Strategies

Of the antiangiogenesis drugs now in clinical trial, some are designed to target specific molecules involved in new blood vessel formation. The exact mechanism of other drugs is not known, but they have been shown to be antiangiogenic by specific laboratory tests (in the test tube or in animals).

In general, four strategies are being used by investigators to design antiangiogenesis agents:

1. Block the factors that stimulate the formation of blood vessels.
2. Use natural inhibitors of angiogenesis.
3. Block molecules that allow newly forming blood vessels to invade surrounding tissue.
4. Incapacitate newly dividing endothelial cells.

Angiogenesis and Angiogenesis Inhibitors: A New Potential Anticancer Therapeutic Strategy

Tumor cells cannot grow as a mass above 2 to 3 mm³ because diffusion is insufficient for oxygen and glucose requirements, unless the tumor induces a blood supply. This mechanism of induction of a new blood supply from pre-existing vascular bed is called angiogenesis. Furthermore, tumor invasiveness and metastasis require neovascularization. In fact, recent published studies suggest that acquisition of the angiogenic phenotype is a common pathway for tumor progression, and neovascularization is linked with other molecular steps leading to tumor progression. The angiogenic process is a complex multistep cascade under the control of positive and negative soluble factors. A paracrine interaction occurs between tumor

and endothelial cells. Angiogenesis involves endothelial cell proliferation, migration, and tubule formation with associated changes in the extracellular matrix, allowing subsequent new vessel growth toward the tumor. Each of the above steps may represent a target for antiangiogenic therapy. Antiangiogenesis is to be distinguished from direct targeting and destruction of tumor vasculature.^[1,3]

BIOTECHNOLOGY-DRIVEN STRATEGIES

Kininostatin

Domain5 (D5) of high molecular weight kininogen has been shown to inhibit endothelial cell migration toward vitronectin, endothelial cell proliferation, and angiogenesis in the chicken chorioallantoic membrane assay. Structure-function studies demonstrate that D5 contains a peptide that inhibits only migration and another peptide that inhibits only proliferation. Data on a related inhibitory monoclonal antibody confirmed the role of kininogen in angiogenesis modulation.^[5]

The $\alpha 3$ Chain of Type IV Collagen Inhibits Tumor Growth

Studies from our laboratory have shown that type IV collagen (COL IV), specifically a peptide comprising residues 185–203 of the noncollagenous domain (NCI) of the $\alpha 3$ (IV) chain, has several biological activities, including inhibition of neutrophil activation, promotion of tumor cell adhesion and chemotaxis, and inhibition of tumor cell proliferation.^[6]

Anti-VEGF

VEGF is an endothelial cell-specific mitogen and an angiogenic inducer. VEGF is essential for developmental angiogenesis and is also required for female reproductive functions and bone formation. Substantial evidence also implicates VEGF in tumors and intraocular neovascular syndromes. A humanized anti-VEGF monoclonal antibody (rhuMab VEGF) Avastin demonstrated improved survival in patients with colorectal cancer when combined with chemotherapy versus chemotherapy alone.^[7] It was also shown to have clinical efficacy against lung and renal cancer.^[8,9]

Role of MMP in Cancer Metastasis

MMPs are zinc-dependent endopeptidases that degrade extracellular matrix proteins (collagens, laminin,

and fibronectin) during the process of cancer invasion and metastasis. Contrary to expectations, MMPs are produced primarily by reactive stromal and inflammatory cells surrounding tumors rather than cancer cells. Potent, orally available, hydroxamic acid-derived inhibitors of MMPs are currently in clinical trials for the treatment of lung, prostate, pancreatic, gastric, and breast cancer. These drugs have demonstrated efficacy alone and in combination with chemotherapy in preclinical trials in mice. In contrast, clinical trials in advanced stage cancer demonstrated lack of efficacy. Further trials in early stages of cancer in conjunction with chemotherapeutic agents are in progress.

2-Methoxyestradiol

2-Methoxyestradiol (2ME) is now being recognized as a potent orally active antiangiogenic and antitumor agent that has minimal if any toxicity. Despite being an endogenous estradiol metabolite, 2ME does not show any estrogenic activity in vitro or in vivo at doses at which estradiol and other metabolites are clearly estrogenic. 2ME is presently in phase I study for metastatic and recurrent breast cancer.

Small Molecule Integrin Antagonists

The role of various integrins in angiogenesis-mediated disorders was demonstrated.^[7] In addition to the well-described role of $\alpha v \beta 3$ and $\alpha v \beta 5$ integrins,^[8] a role for $\alpha 5 \beta 1$ integrin in the modulation of angiogenesis was defined.^[9] The role of $\alpha 5 \beta 1$ integrin in angiogenesis was established through the use of monoclonal antibody, cyclic peptide, and nonpeptide $\alpha 5 \beta 1$ integrin antagonists.^[9]

Ocular Angiogenesis

Age-related macular degeneration (ARMD) is the leading cause of blindness in people over the age of 55, affecting more than 10 million Americans. It is a condition in which the central portion of the retina (the macula) deteriorates. It is equally common in men and women and more common in whites than blacks. The cause is unknown, but the condition tends to run in some families. Macular degeneration affects more Americans than cataracts and glaucoma combined. In the most severe form of age-related macular degeneration (known as wet ARMD), abnormal angiogenesis occurs under the retina, resulting in irreversible loss of vision. The loss of vision is due to scarring of the retina secondary to the bleeding from the new blood vessels. Current treatments for wet ARMD utilize laser-based

therapy to destroy offending blood vessels. However, this treatment is not optimal since the laser can permanently scar the overlying retina, and the offending blood vessels often regrow. Advancing novel angiogenesis inhibitors that effectively and safely inhibit the growth of new vessels in diabetic retinopathy and ARMD, and hence restore lost vision, is the main focus of the field now.

Several antiangiogenesis compounds are currently undergoing clinical trials for treatment of these devastating eye diseases. Relevant models of ocular angiogenesis have been essential in the preclinical validation of viable drug candidates for clinical development. A variety of systems are used to test angiostatic agents; however, specific animal models may be more indicative of certain human disease targets. Rationale for and utility and limitations of these animal systems of ocular angiogenesis were discussed with regard to the preclinical pathway of a novel angiostatic steroid, anecortave acetate. Preclinical models have many limitations with regard to an expected clinical outcome, but they can be valuable in providing guidelines for clinical development.^[10]

Ocular Angiogenesis: The Key Contributions of MMP and VEGF

The inappropriate growth of blood vessels in the cornea, iris, trabecular meshwork, and retina is a leading cause of blindness in developed countries. Both MMP and VEGF are involved in the early stages of disease-related angiogenesis (neovascularization). The strategies and compounds under development for inhibiting MMPs and VEGF for preventing neovascularization and improving the outlook for patients at risk for blindness are under investigation.^[10] The role of hypoxia, extracellular matrix, growth factors, and integrins in pigmented epithelial cell biology might contribute to ocular disorders.^[10,11]

Antiangiogenic Therapy in Multiple Myeloma

In recent years, several pieces of evidence have pointed out that disease progression in multiple myeloma (MM) is characterized by increased bone marrow neovascularization. Targeting the mechanisms that control angiogenesis could thus represent an innovative therapeutic approach to MM. Thalidomide, a glutamic acid derivative with sedative properties, is able to inhibit angiogenesis in murine models; this compound has recently been demonstrated to be effective in relapsed/refractory MM, leading to a 30–40% response, coupled with mild systemic toxicity. Inhibition of angiogenesis is probably just one of the mechanisms by which thalidomide exerts

its action in MM, as immunomodulation and inhibition of cytokine production by bone marrow stroma could also be involved. At present, several studies are ongoing, aiming at testing thalidomide-based drug combinations, mainly with dexamethasone, but also with conventional chemotherapy; the results that have been obtained so far show a synergistic effect of the drug combinations, with a response rate ranging from 50 to 70% in pretreated patients. There is now growing interest in novel compounds with potential antiangiogenic effects. Among them, a promising activity both *in vitro* and in animal models has been displayed by thalidomide analogs, inhibitors of VEGF-2, and endostatin.^[12] However, those angiogenesis inhibitors such as thalidomide and anti-VEGF SU015 demonstrated increased incidence of thrombosis when used in combination with chemotherapeutic agents.^[13,14]

Anticoagulants and Angiogenesis

Many cancer patients reportedly have a hypercoagulable state, with recurrent thrombosis due to the impact of cancer cells and chemotherapy on the coagulation cascade.^[15] Analysis of biomarkers of the coagulation cascade and of vessel wall activation was performed and showed significant increases in thrombin generation and endothelial cell perturbation in a treatment cycle-dependent manner when combining angiogenesis inhibitors and chemotherapeutic agents.^[13,14] The incidence of thromboembolic events, possibly related to the particular regimen tested in this study (SU015 and chemotherapeutic agents), discourages further investigation of this regimen.^[14] This investigation, along with the increased incidences of deep vein thrombosis (DVT) in MM patients receiving thalidomide and chemotherapeutic agents,^[13] suggest the potential advantages of using an anticoagulant such as heparin or low-molecular-weight heparin (LMWH). Additionally, studies have demonstrated that unfractionated heparin (UFH) or LMWH interfere with various processes involved in tumor growth and metastasis. Clinical trials have indicated a clinically relevant effect of LMWH, compared to UFH, on the survival of cancer patients with DVT. Mechanism and efficacy of LMWH and its *in vivo* releasable tissue factor pathway inhibitor on the regulation of angiogenesis and tumor growth was documented.^[15,16] Heparin, steroids, and heparin/steroid combinations have been used in a variety of *in vitro* models and *in vivo* in animal models as effective inhibitors of angiogenesis.^[16,17] Additionally, platelet-tumor cell interactions could play a significant role in tumor metastasis.^[18]

About 20 angiogenesis inhibitors are currently being tested in human trials. Most are in early phase I

or II clinical (human) studies. [See list of angiogenesis inhibitors in clinical trials (Table 1)]. Phase I/II trials include a limited number of people to determine the safety, dosage, effectiveness, and side effects of a drug. In phase III trials, hundreds of people around the country are assigned at random to receive either the new treatment or the standard treatment (Table 1).

CRITICAL ISSUES

Antiangiogenesis is likely to develop into a novel therapeutic approach for patients with solid malignancies. Most current clinical trials evaluate antiangiogenic drugs aimed primarily against single angiogenesis stimulators. A single solid malignancy is shown to produce at least three biologically active angiogenesis stimulators (VEGF, bFGF, and IL-8). This suggests that tumor angiogenesis results from the activity of multiple rather than single angiogenesis stimulators. It was shown that a combination of antiangiogenic drugs is more effective in inhibiting tumor-induced endothelial cell growth than a single agent. These results imply that clinical antiangiogenic strategies for the treatment of solid malignancies may be most effective when multiple rather than single antiangiogenesis drugs are used.

Possible Mechanisms of Acquired Resistance to Antiangiogenic Drugs

The ultimate target of antiangiogenic drugs is the genetically stable, activated endothelial cell of a newly forming tumor blood vessel, rather than the genetically unstable tumor cell population *per se*. This led to the notion that acquired resistance to such drugs may not develop as readily, if at all. While there is some evidence that this lack of resistance development may be the case for some direct-acting angiogenesis inhibitors, it is becoming apparent that resistance can develop over time to many types of angiogenesis inhibitors including, possibly, some direct inhibitors, especially when used as monotherapies.^[19] Possible mechanisms for such acquired or induced resistance include 1) redundancy of proangiogenic growth factors when the drug used targets a single such growth factor or its cognate endothelial cell-associated receptor tyrosine kinase; 2) the antiapoptotic/prosurvival function of growth factors such as VEGF, which, in high local concentrations, can antagonize the proapoptotic effects of various angiogenesis inhibitors; 3) epigenetic, transient upregulation, or induction, of various antiapoptotic effector molecules in host endothelial cells; and 4) heterogeneous vascular

Table 1 Angiogenesis inhibitors that have reached clinical trials

Drug	Sponsor	Trial	Mechanism
<i>Drugs that prevent new blood vessels from invading surrounding tissue</i>			
Marimastat	British Biotech—Annapolis, Md.	Phase III	Synthetic inhibitor of matrixmetalloproteinases (MMPs).
Bay 12-9566	Bayer—West Haven, Conn.	Phase III	Synthetic MMP inhibitor.
AG3340	Agouron—LaJolla, Calif.	Phase III	Synthetic MMP inhibitor.
CGS27023A	Novartis—East Hanover, N.J.	Phase I	Synthetic MMP inhibitor.
COL-3	Collagenex Pharmaceuticals—Newtown, PA.	Phase I	Antibiotic derivative that inhibits MMPs.
Vitaxin	Ixsys, Inc.—LaJolla, Calif.	Phase I	Antibody to integrin, present on endothelial cell surface.
<i>Natural inhibitors of angiogenesis</i>			
Platelet factor-4	Repligen—Cambridge, Mass.	Phase II	Inhibits endothelial cell growth.
Interleukin-12	Genetics Institute—Cambridge, Mass.	Phase I/II	Inhibits endothelial cell growth.
<i>Drugs that block factors that stimulate the formation of blood vessels</i>			
RhuMabVEGF	Genentech—South San Francisco, Calif.	Phase II/III	Monoclonal antibody to vascular endothelial growth factor (VEGF).
SU5416	Sugen, Inc.—Redwood City, Calif.	Phase I	Molecule that blocks VEGF receptor signaling.
Interferon- α	Commercially available	Phase II/III	Inhibits release of endothelial growth factor.
<i>Targeted antivascular therapy</i>			
ZD0101	Zeneca Pharmaceuticals—Wilmington, Del.	Phase I/II	Bacterial toxin that binds to new blood vessels and induces inflammatory response.
<i>Interrupts function of dividing endothelial cells</i>			
TNP-40	TAP Pharmaceuticals, Inc.—Deerfield, Ill.	Phase II	Synthetic analogue of fungal protein;inhibits endothelial cell growth.
<i>Unknown mechanism; inhibits angiogenesis in laboratory and animal assays</i>			
Thalidomide	Entremed, Inc.—Rockville, MD.	Phase II	Synthetic sedative: unknown mechanism.
CAI	National Cancer Institute—Bethesda, MD.	Phase I/II	Nonspecific inhibitor cell invasion and motility.
Squalamine	Magainin Pharmaceuticals, Inc.—Plymouth Meeting, PA.	Phase I	Extract from dogfish shark liver; inhibits sodium-hydrogen exchanger, NHE3.
Suramin	Parke-Davis—Morris Plains, N.J.	Phase II/III	Nonspecific multisite effects.
IM862	Cytran—Kirkland, WA.	Phase II	Unknown mechanism.

dependence of tumor cell populations. It is suggested that long-term disease control with antiangiogenic drugs can be best achieved by judicious combination therapy. In this regard, the great molecular diversity of antiangiogenic drug targets, in contrast to chemotherapy, makes this a particularly attractive therapeutic option, especially when used in combinations.^[1,19]

Standard Chemotherapy Versus Angiogenesis Inhibitors

Several differences between standard chemotherapy and antiangiogenesis therapy result from the fact that angiogenesis inhibitors target dividing endothelial cells rather than tumor cells. Antiangiogenic drugs are not likely to cause bone marrow suppression,

gastrointestinal symptoms, or hair loss. Also, since antiangiogenic drugs may not necessarily kill tumors, but rather hold them in check indefinitely, the end point of early clinical trials may be different than for standard therapies. Rather than looking only for tumor response, it may be appropriate to evaluate increases in survival and/or time to disease progression.

Drug resistance is a major problem with chemotherapy agents. This is because most cancer cells are genetically unstable, are more prone to mutations, and are therefore likely to produce drug-resistant cells. Since angiogenic drugs target normal endothelial cells, which are not genetically unstable, drug resistance may not develop. So far, resistance has not been a major problem in long-term animal studies or in clinical trials.

Finally, antiangiogenic therapy may prove useful in combination with therapy directly aimed at tumor cells. Because each therapy is aimed at a different cellular target, the hope is that the combination will prove more effective.

Antiangiogenesis Agents and Thrombosis

Antiangiogenesis agents might mediate endothelial cell dysfunction, which could be associated with increased incidence of thrombosis.^[12–14,20]

Diagnostic Imaging of Angiogenesis

Angiogenesis is a novel field that promises to provide new venues for blood flow in patients with severe ischemic heart disease and severe peripheral vascular disease, and promises to control malignancies by controlling blood supply. However, it became clear early on that a novel field necessitates novel outcome measures that follow its beneficial effects. There is a growing need for non-invasive methods to serially assess the status of the coronary, peripheral, and tumor vasculature. Biomarkers might reflect the prognosis of the diseases as well as the effectiveness of the treatment.^[21]

These might include the following: magnetic resonance imaging and monitoring of angiogenesis, nuclear perfusion imaging of angiogenesis, and other modalities.

ARTICLES OF FURTHER INTEREST

Bioactive Materials and Scaffolds for Tissue Engineering; Cardiac Patch Engineering; Controlled Release; Drug Delivery, Controlled; Host Reactions; Matrix

Metalloproteinases; Thrombosis; Tissue Engineering of Microvascular Networks

A

REFERENCES

1. Mousa, S.A. Mechanisms of Angiogenesis in Vascular Disorders: Potential Therapeutic Targets. In *Angiogenesis Inhibitors and Stimulators: Potential Therapeutic Implications*; Mousa, S.A., Ed.; Landes, R.G.: Georgetown, TX, 2000; 1–12. Chapter 1.
2. Pavlakovic, H.; Havers, W.; Schweigerer, L. Multiple angiogenesis stimulators in a single malignancy: Implications for anti-angiogenic tumor therapy. *Angiogenesis* **2001**, *4* (4), 259–262.
3. Ranieri, G.; Gasparini, G. Angiogenesis and angiogenesis inhibitors: A new potential anticancer therapeutic strategy. *Curr. Drug Targets Immune, Endocrinol. Metab. Disord.* **2001**, *1* (3), 241–253.
4. Marneros, A.G.; Olsen, B.R. The role of collagen-derived proteolytic fragments in angiogenesis. *Matrix Biol.* **Sep 2001**, *20* (5–6), 337–345.
5. Colman, R.W.; Jameson, B.A.; Lin, Y.; Mousa, S. Inhibition of angiogenesis by kininogen domain 5. *Blood* **2000**, *95* (2), 543–550.
6. Scialdone, M.A.; Mousa, S.A.; Shuey, S.W. Angiogenesis-Inhibitory Tripeptides, Compositions and their Methods of Use. U.S. Patent 20030027769, February 2003.
7. Mousa, S. Anti-integrins: Novel therapeutic developments. *Exp. Opin. Invest. Drugs* **2001**, *10* (2), 387–391.
8. Van Waes, C.; Enamorado, I.; Hecht, D.; Sulica, L.; Chen, Z.; Batt, G.; Mousa, S. Effects of the novel α v integrin antagonist SM256 and cis-platinum on growth of murine squamous cell carcinoma PAMLY8. *Int. J. Oncol.* **2000**, *16* (6), 1189–1195.
9. Kim, S.; Mousa, S.; Varner, J. Requirement of integrin α 5 β 1 and its ligand fibronectin in angiogenesis. *Am. J. Pathol.* **2000**, *156*, 1345–1362.
10. Campochiaro, P.A.; Bora, N. Ocular Angiogenesis. In *Angiogenesis Inhibitors and Stimulators: Potential Therapeutic Implications*; Mousa, S.A., Ed.; Landes, R.G.: Georgetown, TX, 2000; 110–119.
11. Mousa, S.; Lorelli, W.P.; Campochiaro, P.A. Role of hypoxia and extracellular matrix-integrin binding in the modulation of angiogenic growth factors secretion by retinal pigmented epithelial cells. *J. Cell. Biochem.* **1999**, *74* (1), 135–143.
12. Tosi, P.; Tura, S. Antiangiogenic therapy in multiple myeloma. *Acta Haematol.* **2001**, *106* (4), 208–213.
13. Bennett, C.L.; Schumock, G.T.; Desai, A.A.; Kwaan, H.C.; Raisch, D.W.; Newlin, R.; Stadler, W. Thalidomide-associated deep vein thrombosis and pulmonary embolism. *Am. J. Med.* **2002**, *113* (7), 603–606.
14. Kuenen, B.C.; Rosen, L.; Smit, E.F.; Parson, M.R.; Levi, M.; Ruijter, R.; Huisman, H.; Kedde, M.A.; Noordhuis, P.; van der Vijgh, W.J.; Peter, G.J.; Cropp, G.F.; Scigalla, P.; Hoekman, K.; Pinedo, H.M.; Giaccone, G. Dose-finding and pharmacokinetic study of cisplatin, gemcitabine, and SU5416 in patients with solid tumors. *J. Clin. Oncol.* **Mar 15, 2002**, *20* (6), 1657–1667.

15. Mousa, S. Angiogenesis, coagulation activation, and malignant dissemination. *Sem. Thromb. Haemost.* **2002**, *28* (1), 45–52.
16. Mousa, S.; Mohamed, S. Anti-angiogenesis efficacy of the LMWH, tinzaparin and TFPI. *Blood* **1999**, *94* (10), 82. Abstract 22a.
17. Jung, S.P.; Siegrist, B.; Wade, M.R.; Anthony, C.T.; Woltering, E.A. Inhibition of human angiogenesis with heparin and hydrocortisone. *Angiogenesis* **2001**, *4* (3), 175–186.
18. McCarty, O.J.T.; Mousa, S.; Bray, P.; Konstantopoulos, K. Immobilized platelet support human colon carcinoma cell tethering, rolling and firm adhesion under dynamic flow conditions. *Blood* **2000**, *96* (5), 1789–1797.
19. Kerbel, R.S.; Yu, J.; Tran, J.; Man, S.; Vilorio-Petit, A.; Klement, G.; Coomber, B.L.; Rak, J. Possible mechanisms of acquired resistance to anti-angiogenic drugs: Implications for the use of combination therapy approaches. *Cancer Metastasis Rev.* **2001**, *20* (1–2), 79–86.
20. Kaushal, V.; Kohli, M.; Zangari, M.; Fink, L.; Mehta, P. Endothelial dysfunction in antiangiogenesis-associated thrombosis. *J. Clin. Oncol.* **2002**, *20* (13), 3042–3043.
21. Chen, Z.; Malhotra, P.; Thomas, G.; Ondrey, F.; Duffey, D.; Smith, C.; Enamorada, D.; Yeh, N.; Kroog, G.; Rudy, S.; McCullagh, L.; Mousa, S. Expression of proinflammatory and proangiogenic cytokines in patients with head and neck cancer. *Clin. Cancer Res.* **1999**, *5*, 1369–1379.

Animal Models for Bone Tissue Engineering

Mark P. Mooney

Departments of Oral Medicine and Pathology, Anthropology, and Surgery, Division of Plastic and Reconstructive Surgery, and Orthodontics and Cleft Palate-Craniofacial Center, University of Pittsburgh, and Department of Biomedical Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania, U.S.A.

Michael I. Siegel

Department of Anthropology, University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.

INTRODUCTION

Animal models of bone pathologies and wound healing have contributed significantly to the preclinical evaluation of the functional efficacy and biocompatibility of new bone tissue engineering methodologies.^[1–7] However, choosing an appropriate animal model for such evaluations is a critical first step in the experimental design process and in the extrapolation of results to the clinical setting.^[1,8–19] The scientific and ethical implications of choosing an appropriate animal model is a complicated procedure and often a heated topic of debate, and should continue to be guided in part by the three “R”s—replacement, reduction, and refinement.^[9,16–27]

In this entry we will: 1) discuss a number of practical criteria to facilitate the animal model choice for bone tissue engineering studies; 2) review the advantages and disadvantages of various animal models for basic bone biology and the testing of bone replacement materials (BRMs) in critical-sized defects (CSDs); and 3) identify animal models with clinically relevant pathologies that may be used to develop specific bone tissue engineering paradigms and strategies.^[1,10,11,15] This discussion is intended as an aid to less experienced researchers and students designing experimental studies, and to tissue engineers and clinicians who must critically evaluate the results obtained from published animal studies for validity and potential extrapolation to the human clinical condition.

WHAT MAKES AN ANIMAL MODEL “APPROPRIATE”?

When choosing an animal model for bone tissue engineering studies a number of factors should be taken into consideration.^[1,10–12,27,28] These include: 1) animal model appropriateness; 2) potential and expected extrapolation to the clinical setting; 3) genetic homogeneity of the specific animal model; 4) available data concerning skeletal anatomy, bone physiology and

biomechanical properties, and osseous wound healing; 5) cost and availability of the model; 6) generalizability of the results across species; 7) ease and adaptability of the model to experimental and laboratory manipulations; 8) ecological considerations; and 9) ethical and societal implications.

Nonhuman primates are phylogenetically closer to humans than other mammalian groups, which may fulfill many of the considerations listed above and make extrapolation of findings to humans theoretically “better.” However, not all bone tissue engineering experimental manipulations require nonhuman primate models. As we have suggested previously, the choice of an appropriate animal model for bone tissue engineering should be based, in part, on the criteria initially suggested by Smith and Reynolds for determining appropriate animal models in toxicological studies.^[3,10,15,29–32] Smith, on the basis of comparative studies of drug disposition across a large number of mammalian taxa, reported great variation in drug metabolism within the various primate groups.^[31] He suggested that choosing the best animal model should be based on similar physiologic pathways between the model and the human condition, and not necessarily on phyletic affinity. Smith also suggested that the level of hypothesis testing and the expected extrapolation of the results to humans be of major concern in choosing the best animal model.^[31] Based on these practical criteria, we propose that the choice of an appropriate model for bone tissue engineering studies also be linked to the level of hypothesis testing, the expected extrapolation to the specific human clinical condition under study, and the expected commercial market of the tissue-engineered construct.^[1]

The three main strategies taken to engineer bony tissue involve: 1) osteoconduction, which utilizes a scaffold as a mechanical barrier to exclude osteoinhibitory tissues and protect osteogenic tissues in a bony defect; 2) osteoinduction, which utilizes the release of bioactive molecules that bind only to specific host cells with receptors for the molecules and stimulate cell migration and osteogenesis to repair a bony defect;

Table 1 Levels of hypothesis testing in bone tissue engineering paradigms

Level of hypothesis testing	Bone tissue engineering paradigms		
	Osteoconduction	Osteoinduction	Osteogenic cell transplantation
Basic bone cell biology models	In vitro or in vivo manipulation of cell or tissue response to trauma using mechanical or environmental factors	In vitro or in vivo manipulation of cell or tissue response to trauma using growth factors, cytokines, or genes	In vitro or in vivo manipulation of autologous or isohistogenic donor osteogenic cell response to host bone trauma
General clinical models	Repair of CSDs using a passive mechanical scaffold	Repair of CSDs using resorbable scaffolds seeded with growth factors, cytokines, or genes	Repair of CSDs using resorbable scaffolds seeded with osteoblasts or stem cells
Specific clinical models	Repair of bony defects in specific clinical conditions (e.g., periodontitis, diabetes, cleft palate, osteoporosis) using a passive mechanical scaffold	Repair of bony defects in specific clinical conditions (e.g., periodontitis, diabetes, cleft palate, osteoporosis) using resorbable scaffolds seeded with growth factors, cytokines, or genes	Repair of bony defects in specific clinical conditions (e.g., periodontitis, diabetes, cleft palate, osteoporosis) using resorbable scaffolds seeded with osteoblasts or stem cells

(From Ref. [3].)

and 3) osteogenic cell transplantation, which involves the seeding of osteogenic cells from a donor source on a synthetic bone construct that is subsequently implanted into the defect (Table 1).^[33–35] Animal model appropriateness and “goodness-of-fit” will vary across these paradigms as the need for extrapolation to the human condition increases and can be divided into three levels of hypothesis testing: 1) “generic” animal models of basic bone cell biology; 2) phylogenetically “closer” models with comparable anatomy and bony wound healing sequelae; and 3) “fitting” the appropriate animal model to various clinical conditions.

Generic Animal Models and Basic Bone Cell Biology

Studies involving in vitro manipulations of genes, molecules, cells, and tissue are viewed as the most basic. Humans share primitive developmental and regulatory genetic mechanisms that produce homologous cortical and trabecular bony structures, similar bone physiology and wound healing sequelae, and even comparable bone pathophysiology and disease pathogenesis and progression with many other taxonomic orders.^[2,4,13,34,36–48] In general, osteogenic cells from normal and genetically engineered rodents and lagomorphs have been used extensively to elucidate a basic understanding of cell–gene, cell–factor, and cell–cell interactions (see other entries in this volume) as an aid to designing bone tissue engineering strategies (Table 1).^[1,2,35,42,49–52] The data derived from studies using

these generic animal models are as valid as those derived from carnivore or primate models, given that no attempt is made to grossly extrapolate beyond this level of hypothesis testing.

Phylogenetically “Closer” Animal Models and General Clinical Issues

In vivo bone tissue engineering studies at this level of hypothesis testing typically attempt to model clinical problems and complications associated with general skeletal trauma and postoperative bony wound healing. These studies typically involve manipulations of general clinical skeletal models (e.g., long bone, vertebral body, and craniofacial fractures, oncologic resections, elective craniofacial and maxillofacial procedures) and present more difficult problems for choosing an appropriate animal model. Depending on the level of extrapolation to the human condition, not all animal models may be appropriate at this level of hypothesis testing. The commonly used generic animal models (such as rodents and rabbits) differ somewhat in their basic bone microarchitecture and biomechanics, functional gross skeletal anatomy, and healing of CSDs compared to phylogenetically closer models (such as some carnivores and nonhuman primates).^[5,10–13,15,28,36,37,48,53–56] Such interspecific differences should be taken into account before choosing an appropriate model at this level of hypothesis testing.

Basic bone microarchitecture and biomechanics

Appropriate animal models at this level of hypothesis testing should have similar normal human bone microanatomy and systemic physiological processes. The human skeleton shows significant osteoporotic changes during immobilization, normal aging, and/or menopause as a consequence of systemic and environmental factors.^[13,37] In young individuals, the central cancellous bone network is a highly connected, very strong, plate–strut network. In older individuals, osteoclastic activity changes this network to a more fragile, disconnected, strut–strut network that is relatively weak.^[37] Because of these endocrine-mediated, age-related changes, 40–45% of the cancellous bone is lost in the vertebral bodies, long bones, and femoral neck with a concomitant loss of approximately 80–90% strength and a reduced vertebral cortical shell thickness of approximately 200–400 μm . Peak cortical bone mass, connectivity, and thickness are also reduced owing to both cortico-endosteal and intracortical (Haversian) bone remodeling.^[13,37] These changes are accelerated and more pronounced during the development of osteoporosis (age related, menopausal, or immobilization related) and result in strength changes of the human femur during compression and tension, which ranges from 90 to 143 MPa and from 90 to 167 MPa, respectively. Age-related changes in Young's modulus are also seen, which ranges from approximately 5 to 15.5 GPa during compression and from 4 to 17 GPa during tension.^[53] Such age- or disease-related changes in cancellous bone loss, cortical bone thinning, and decreased bone strength and Young's modulus are responsible, in part, for the clinical problems associated with long bone, femoral neck, and vertebral body fractures. Such clinical complications have stimulated an impressive bone tissue engineering initiative to address these issues.^[1,2]

Rodents and rabbits are probably the most commonly used animal models at this level of hypothesis testing.^[10,13,28,37,53] The advantages of these models include: 1) standardization of experimental conditions; 2) genetically specific or mutant strains can be acquired; 3) relatively inexpensive to house and maintain; 4) relatively shorter life spans and fast bone turnover rates; 5) extensive documentation of bone metabolism and skeletal effect of diet; 6) similar lamellar bone architecture; 7) similar cancellous bone thinning and fragility, as well as remodeling rates and sites; and 8) ovariectomy mimics human skeletal aging and menopause.^[12,13,28,34,37] However, there are also a number of disadvantages of the rodent and rabbit models, which should be taken into account before designing bone tissue engineering studies utilizing these models. These include: 1) different skeletal loading patterns; 2) open epiphyses at various growth plates to the age of

12–24 mo; 3) minimal intracortical remodeling; 4) rodents lack Haversian canal systems; 5) rodents also have hemopoietic bone marrow at most skeletal sites, which increases bone turnover rates compared to primates; 6) they have a smaller proportion of cancellous bone to total bone mass; 7) they do not experience a natural menopause and do not show impaired osteoblast function during late stages of estrogen deficiency; and 8) their relatively small size for the testing of prosthetic devices, repeated bone biopsies, or blood testing.^[12,13,28,37] The rodent model also shows age-related changes in femoral head strength, which ranges from approximately 1 to 5 MPa, and in Young's modulus, which ranges from approximately 12 to 405 GPa, during compression.^[53] If one can keep the disadvantages of the small animal models in mind during data interpretation, the rodent and rabbit models are excellent for the initial testing of novel bone tissue engineering paradigms (Table 1).

In contrast, larger animal models (nonhuman primates, dogs, cats, sheep, goats, and swine) show more similar microarchitecture, bone physiology, and biomechanical properties to humans than do the rodent and rabbit models, especially with regard to intracortical bone remodeling.^[12,13,28,37,57] While each large animal model has its own advantages and disadvantages, nonhuman primates are thought to be the model of choice at this level of hypothesis testing, based on very similar, anatomical, gastrointestinal/dietary, endocrine, bone metabolism, and circadian rhythm factors.^[10,12,13,28,37,57–59] However, cost, availability, and genetic heterogeneity of nonhuman primates can be problematic, which necessitates the utilization of other large animal models. In general, the advantages of large animal models over small animal models include: 1) well developed Haversian and trabecular bone remodeling; 2) greater skeletal surface areas and volumes; 3) similar skeletal disuse atrophy results; 4) highly localized bone fragility associated with stress shielding by orthopedic implants; and 5) similar postmenopausal-related osteopenia (with the exception of the dog model). Depending on the large animal model, there are also a number of disadvantages, which include: 1) relative high cost and maintenance expense; 2) USDA housing and space requirements; 3) relatively long life spans; 4) varying gastrointestinal and dietary (trace mineral) requirements, which can affect skeletal calcium:phosphorus ratios; 5) problems with generating large, homogenous samples for statistical testing; 6) potential reservoirs for a host of zoonotic diseases; 7) emotional attachment; and 8) varying ethical concerns.^[12,13,28,37] It is also difficult to create the disproportionate loss of bone strength (characteristic in humans) in many of the large animal models, with some large animal models having a vertebral cortical shell of 1500–3000 μm thickness.^[37] The large animal

models also show age-related changes in femur strength during compression and tension, which ranges from approximately 136 to 195 MPa and 93 to 172 MPa, respectively. They also showed age-related changes in Young's modulus during compression and tension, which ranges from approximately 19 to 27 GPa and 5 to 25 GPa, respectively.^[53]

Gross anatomical, functional, and growth considerations

There are also a number of gross anatomical differences in the long bones, vertebrae, and craniofacial complex of different animal models that need to be considered before determining animal model appropriateness (Figs. 1–4). Such gross skeletal differences are due, in part, to differing endocrine, biomechanical, functional (i.e., locomotor, postural, and masticatory), and dietary influences.^[12,13,28,37,53,57–59] These factors can govern the regulation of bone mass and three-dimensional structure, as well as modulate the “set point” for bone adaptation and bony wound healing.^[28]

While only humans and some birds (emus) are habitually bipedal, most nonhuman primates and rodents maintain upright postures during feeding and grooming.^[41,54,60] Such comparable functional loading produces similarly elongated femoral necks and heads (compare Figs. 1A and 1D), similarly shaped vertebral bodies, vertebral laminar thicknesses, and oriented spinous and mammillary processes of the lumbar vertebrae (compare Figs. 2A and 2B with Figs. 2G and 2H). In contrast, more habitually quadrupedal species (rabbits and dogs in this case) show more robust femoral shafts, thicker and shorter femoral necks (Figs. 1B and 1C), vertically compressed vertebral bodies, relatively thinner vertebral laminae,

and differently oriented bony vertebral processes (Figs. 2C–2F) compared to monkeys and rats. Thus, rats may be an appropriate, less expensive model compared to nonprimates for modeling human femoral or vertebral body morphology if postural loading on tissue engineered constructs is the main concern.

Craniofacial skeletal tissue loss due to congenital defects, disease, and injury is a major clinical problem.^[35] The human craniofacial skeleton is a complex region to model and postural, masticatory, dental, visceral, and neural factors all contribute to its unique adult shape.^[10,15,61] The more common generic animal models (mice and rats) are monophyodonts, exhibiting only a single set of the molar and incisor classes, while other generic models (rabbits and guinea pigs, for instance) have succedaneous dentition but maintain continually erupting incisors.^[8,54,60,62,63] These generic animal models also have large maxillary and mandibular diastema between incisors and premolars/molars with limited alveolar bone in this area (Fig. 3). They are also dolichocephalic (i.e., long, narrow headed) from limited brain cerebralization, with relatively long, narrow, and flat cranial vault bones (Figs. 3 and 4) compared to the human condition.^[54,64] Rats and mice are omnivorous while rabbits and guinea pigs are herbivorous. These functional differences are reflected in the enlarged attachments sites of the masseter m. on the ascending ramus and the reduced coronoid processes seen in the mandibles of rabbits and guinea pigs (Fig. 3). These models are also relatively small, which may affect the design and feasibility of testing synthetic constructs and orthopedic devices for implantation and osteointegration.

Larger, generic animal models have been utilized to overcome some of the disadvantages of the smaller models discussed above. These models include sheep, goats, and swine (both large and miniature versions)



Fig. 1 Cleaned and dried left femora from various animal models. Note the relatively longer femoral necks in the rat and monkey compared to the rabbit and dog, which reflect more habitual upright posture. (H, femoral head; N, femoral neck; S, femoral shaft.)

A

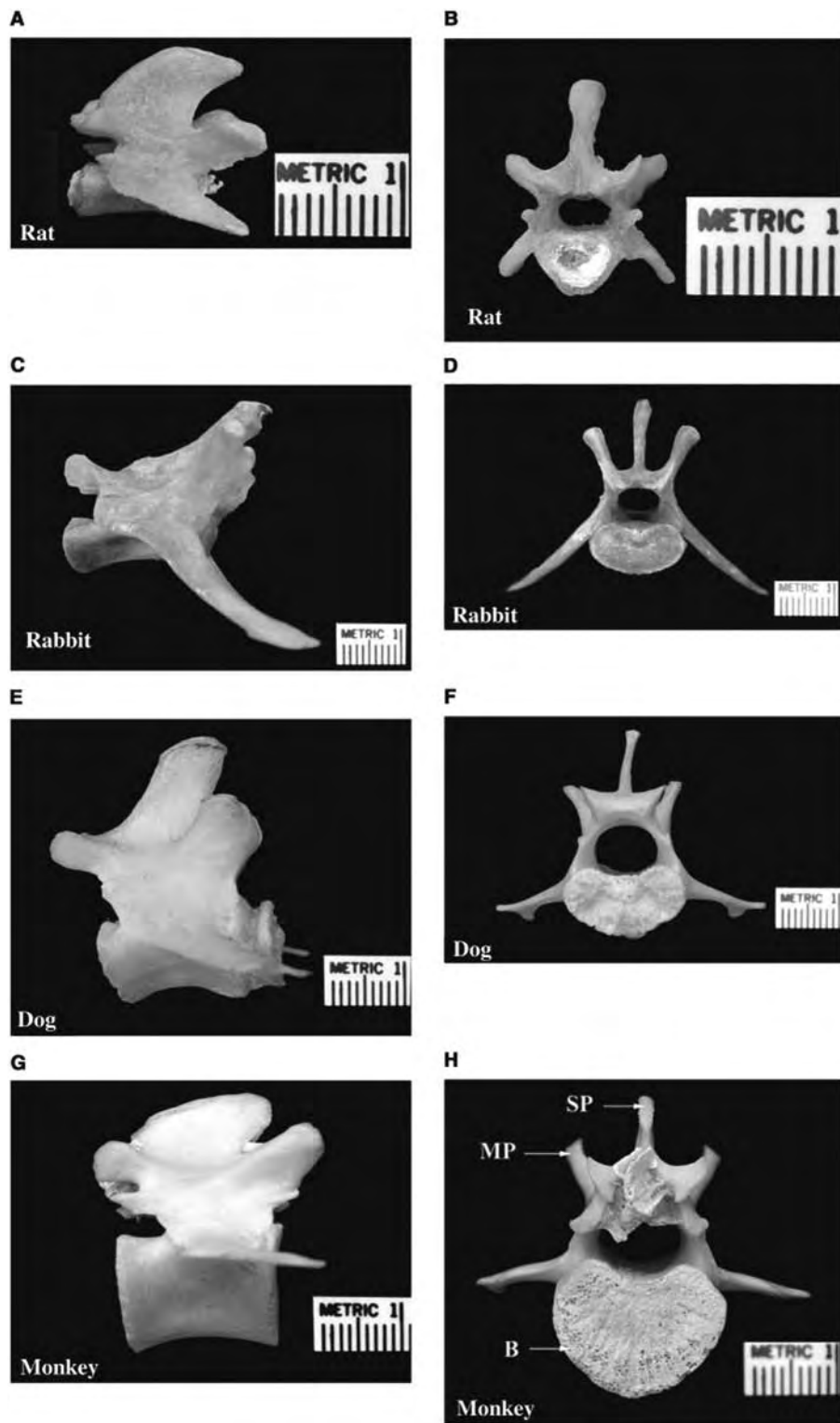


Fig. 2 Cleaned and dried second lumbar (L2) vertebrae from various animal models. Note the more oval-shaped vertebral bodies and broader spinous processes in the rat and monkey compared to the rabbit and dog, which reflect more habitual upright posture and vertebral loading. (B, vertebral body; MP, mammillary process; SP, spinous process.)

and have been well described as models for osteoporosis.^[12-14,28,36,37,53] All three models have varying dental formulas and the goat and sheep have maxillary and mandibular diastemas (Fig. 3).^[65] They are also dolichocephalic (i.e., long, narrow headed) from anteroposterior brain growth and limited superior

cerebralization.^[54,64] Their cranial bones are relatively long, narrow, and flat, but they are still much larger than those of rodents and lagomorphs (Fig. 3). The goat also has horns, which limits the accessibility and surface area of the cranial vault bones for surgical manipulations.^[66,67] Sheep and goats are herbivorous

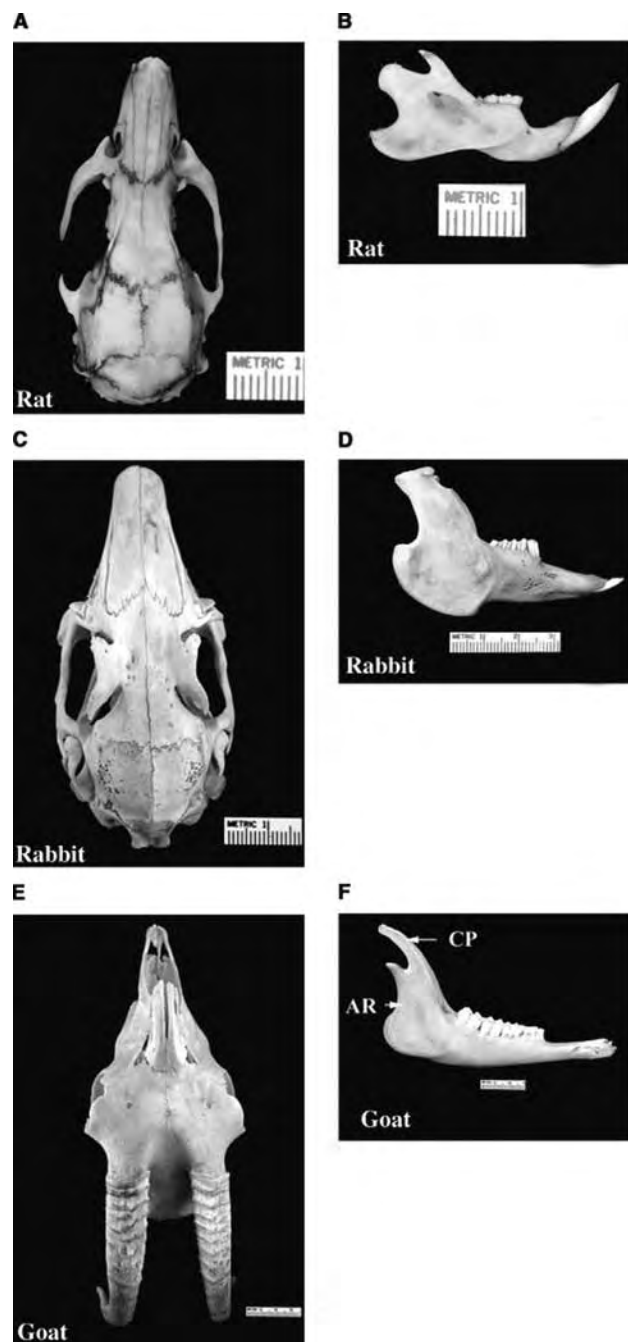


Fig. 3 Cleaned and dried skulls and mandibles from the generic animal models. Note the reduced coronoid processes in the rabbit and goat compared to the rat, which reflects differing masticatory function. (AR, ascending ramus; CP, coronoid process.)

and ruminants while pigs are omnivorous. These functional differences are reflected in the enlarged attachment sites of the masseter m. on the ascending mandibular ramus and zygomatic arch, the reduced temporal fossae for the attachment site of the temporalis m. seen in the skulls and jaws of sheep and goats compared to swine (Figs. 3 and 4). Although

omnivorous, the pig also has a very strong masticatory apparatus with an enlarged ascending mandibular ramus for masseter m. attachment. These larger, generic animal models are more suitable to studies that need greater bony surface area, more alveolar bone, or increased jaw basal bone for testing compared to the smaller, generic animal models. However, these models are much more expensive to purchase and house, bone turnover rate and osseous healing are relatively slow, and they are less genetically homogenous compared to the smaller, generic animal models, although the minpig model overcomes some of these disadvantages.

In contrast, the phylogenetically “closer” models (carnivores and nonhuman primates) exhibit more similar craniofacial structures to humans than do the generic models (Figs. 3 and 4). They possess succedaneous dentition with varying eruption times of both the deciduous and permanent teeth, some with very large canine teeth.^[8,10,54,65,68–70] They are relatively more brachycephalic (i.e., short, wide headed) from increased brain corticalization and gyrification.^[15,54,61,64,71–74] Their cranial bones are relatively shorter, wider, and more curved (Fig. 3 and 4) compared to the generic animal models, which makes them more similar to the human condition. While, cats and dogs are carnivores, laboratory-reared species usually have a much softer, standardized diet than their wild-reared counterparts. Cats and dogs have enlarged coronoid processes for temporalis m. attachment and more inferiorly located mandibular condyles compared to primates, which affords them strong lever arms, reduced horizontal excursion, and increased vertical power. In contrast, some nonhuman primates are omnivorous, while others are herbivorous, although laboratory-reared nonhuman primates also have a standardized, relatively soft diet. Their dental formulae and masticatory apparatus are also closest to those of humans (Figs. 3 and 4). Based on these striking similarities in systemic and oral skeletal physiology and anatomy, nonhuman primates have often become the model of choice of many investigators.^[10,12,28,37,57–59] However, the decreasing availability of wild-caught and laboratory-reared primates, their increased cost, and the social and ethical concerns of using nonhuman primates, necessitate the development and utilization of alternative models.^[9–11,16–25,27,75]

Early surgical intervention of craniofacial and dentofacial deformities have also recently entailed the development of strategies to engineer bone and soft tissue constructs that grow along with the individual.^[35,52] The testing of these growing constructs requires animal models that show similar human craniofacial growth patterns.^[3,10,15,29,30] Appropriate animal models at this level of hypothesis testing would be

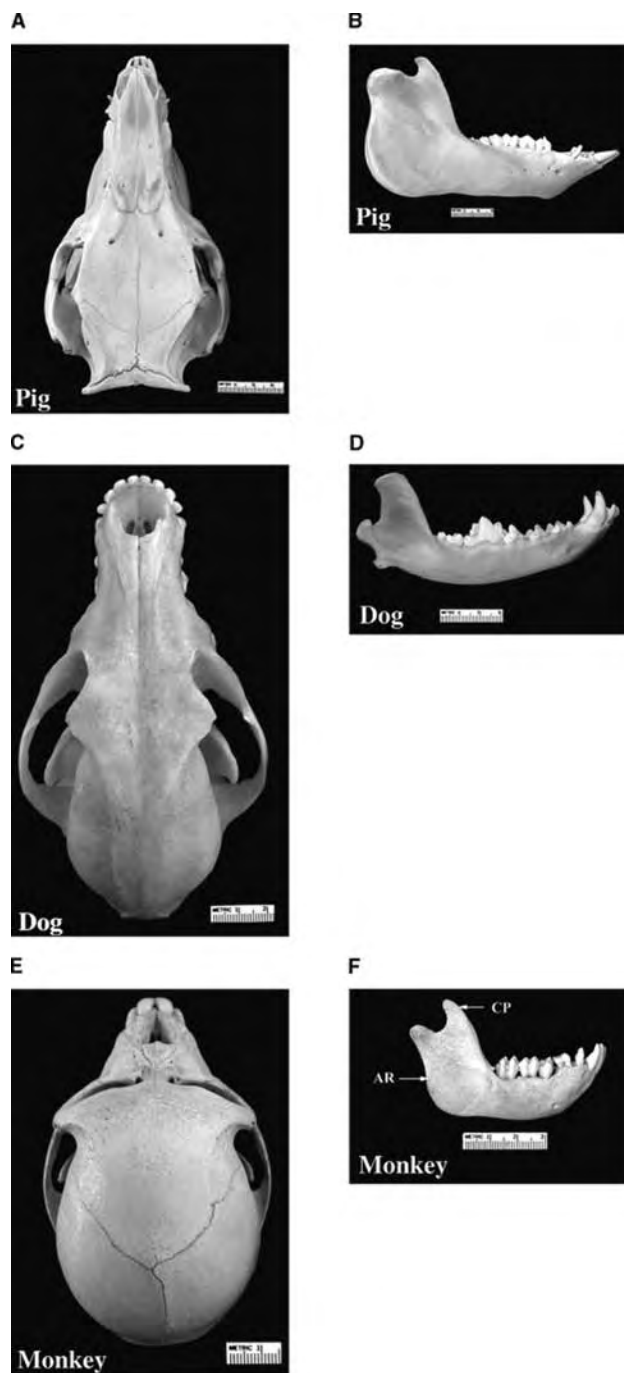


Fig. 4 Cleaned and dried skulls and mandibles from the phylogenetically “closer” animal models. Note the overall differences in the shape of the skull, dentition, and mandibles compared to the generic animal models, which reflect further changes in brain morphology and differing masticatory function.

those that exhibit similar regional growth patterns to humans. In general, human regional craniofacial growth patterns are characterized by a high degree of cranial base flexion (kyphosis), nasomaxillary reduction, anteroinferior displacement of the midface,

and significant vertical and transverse increases in both the neurocranium and the upper face.^[54,61,76] The human craniofacial growth pattern differs from the primitive adult mammalian condition of a long, slender dolichocephalic craniofacial skeleton to a broad, vertically flattened brachycephalic cranium without a prominent snout (Figs. 3 and 4) (for an excellent review see Enlow).^[61] Thus, experimental studies designed to manipulate regional craniofacial growth patterns should choose an animal model based on similar regional growth vectors and patterns, and not necessarily only on phyletic affinity.

In general, under conditions of modeling specific human craniofacial growth patterns, a single animal model may not be appropriate for all regions. For example, short-faced animals (felids and short-faced nonhuman primates) show similar midfacial and nasal capsule growth patterns to humans, while juvenile canids show similar mandibular growth patterns to humans.^[7,29,30] Thus, if the objectives of the study are to improve the human condition, the use of a primate model simply on the basis of phyletic affinity is not justified and an understanding of the relationship between the growth patterns in humans and those in the animal under consideration is a prerequisite to appropriate animal model choice.^[10,15,29,30]

Critical-sized defects and bony wound healing

The most common replacement therapy for bony tissue loss, especially in the craniofacial complex, is autologous bone grafting. However, limited supply and donor site morbidity have stimulated the search for BRMs.^[47,48,55,56,77–84] Bone tissue engineering aims to restore function and/or replace the damaged or diseased bony tissue through the principles of osteogenesis, osteoinduction, and osteoconduction (Table 1). As new BRMs are designed and developed, a consistent, standardized testing paradigm using CSDs was proposed.^[11,85]

A CSD is defined as a defect that heals by less than 10% bony regeneration during the lifetime of the individual.^[11] In contrast, fibrous nonunions can form through a combination of fibrous and bony healing that does not return bony continuity and function, but may eventually heal. Critical-sized defects are thought to result from a number of factors, including: 1) an uneven gradient of soluble osseous growth factors in the wound, which fail to reach the center of the defect and uniformly promote osteogenesis; 2) uneven biomechanical stress and loading, especially toward the center of the defect, which reduces uniform bone cell migration, adhesion, and subsequent osteoid and bone matrix formation; and/or 3) the differential and more rapid migration of fibroblasts (compared to

A

osteoblasts) into the wound site during healing, which results in a higher prevalence of fibrous tissue in the center of the defect.^[11,43,85-88] Thus, the CSD model fails to heal only because it exceeds the body's ability to regenerate adequate amounts of bone, fast enough.

Critical sized-defects make an excellent model to test bone tissue engineering paradigms.^[11,43,85] However, there are a number of confounding variables that can influence the bony wound healing of CSDs and should be taken into account before choosing an appropriate animal model. These include: 1) species-specific differences in bony wound healing; 2) skeletal age of the individual; and 3) anatomic location of the CSD.^[11,85,89]

Species-specific differences in bony wound healing have been reported.^[11,85,89] Spontaneous osseous wound healing was seen to decrease in higher-order species.^[11,89] As noted by Hollinger and Kleinschmidt, "The defect size to volume ratio in the rat is 1.96, while in the human the ratio is 0.12. If the human could regenerate as much bone as the rat it could regenerate bone to fill a defect 23 cm in diameter" (Ref. [11], p. 61). This nonallometric relationship is also seen in Table 2. The size of the CSD in the rodent calvaria ranges from 5 to 8 mm in diameter while in larger animal models the CSD size ranges from 15 to 25 mm in diameter.

There are also significant age-related changes in the healing of CSDs, which also vary across taxonomic levels. Bony defects in skeletally immature animals of most species heal at a faster rate than in skeletally mature individuals. If BRMs are tested in immature animals, falsely high expectations of the osteoconductive properties of the material could be generated.^[11,90] Hollinger and Kleinschmidt suggest that biological age, as determined by radiographic confirmation of epiphyseal closure, be used in conjunction with chronological age and body weight to ascertain skeletal maturity, prior to CSD creation.^[11] Rats are the exception to this rule because they grow constantly throughout their life cycle.^[11,60] While in most animals, the initial size of the CSD decreases with age, the initial size of the CSD in rats stays constant, regardless of age.^[11] However, there has recently been a push to engineer bone and soft tissue constructs that grow along with the individual.^[35,52] The testing of these growing constructs may require the use of CSDs in growing animal models, information on which is currently sparse in the literature.

Anatomic location of the CSD can also affect healing rate and interpretation. The craniofacial complex is much more vascular and more richly innervated than the long bones and vertebral column. Even within species, CSD size varies across these anatomic

locations and CSD sizes in the skull and mandible are relatively smaller than those produced elsewhere (Table 2). Thus, anatomic location and CSD size need to be taken into consideration before choosing an appropriate animal model.

Table 2 reviews studies of selected CSD sizes by anatomic location and species. As can be seen in the table, the majority of the CSD studies have been performed on the calvaria, mandible, and femur across a wide variety of animal species (Figs. 3 and 4). These regions are easily accessible for surgery and represent biomechanically loaded (mandible and femur) or unloaded (calvaria) regions to simulate varying clinical conditions. Animal model choice will vary depending on the paradigm being tested (i.e., size of scaffold, duration of healing, length of construct bioactivity, and so on).

Fitting the Appropriate Animal Model to Specific Human Clinical Conditions

As tissue engineering paradigms become more reproducible and sophisticated, we see bone tissue engineers trying to tackle osseous wound healing problems in specific, clinically compromised, subpopulations of patients (i.e., diabetic, osteoporotic, arthritic, geriatric, congenital anomalies, etc.). Addressing osseous wound healing problems in these clinical situations will require specific animal models of the pathology or disease state. Table 3 presents a selected list of animal models of specific bone pathologies and bone metabolic disease states. A number of these models have been developed for specific purposes (i.e., to study the onset and treatment of postmenopausal osteoporosis, inflammatory arthritis, periodontitis induced bone loss, etc.), but these models can also be adapted to study the various bone tissue engineering paradigms. Osteoconduction, osteoinduction, or cell transplant may be significantly altered in these human clinical groups and animals that model these specific clinical conditions will be instrumental in developing new treatment strategies to regenerate new bone and improve their quality of life.

Animal model choice at this level of hypothesis testing should not necessarily be based only on phyletic affinity but should also take into account the etiology of the particular bone pathology or disease state. Although the expectation is that phyletic affinity and similarity of etiology, disease progression, and pathophysiology are reasonably related, phyletic affinity should not be the exclusive criterion, but should be used in conjunction with similar biomechanical, anatomical, and wound healing factors.^[3,13,17-19,36,39,58,59,139]



Table 2 Selected CSD studies by animal model and anatomic region

Anatomic region	Animal model (Refs.)									
	Mouse	Rat	Rabbit	Cat	Dog	Sheep	Pig	Horse	Nonhuman primates	Goat
Calvaria	5 mm round ^[91]	8 mm round ^[92-94]	1.5 mm round ^[89,95-97]	25 mm round ^[98]	20 mm round ^[90,99]	20 mm round ^[100]			15 mm round (rhesus baboons), 25 mm round (baboons) ^[11,101,102,103]	14 mm round ^[66,67]
Palate			5 mm seg. ^[103,104]	5 mm ^[105]	5 mm seg. ^[106,107]					
Alveolus	3 mm seg. ^[108]	4 mm seg. ^[109]	5 mm seg., 10 mm seg. ^[88,104,110-112]						10-15 mm ^[11,113]	
Mandible	4 mm round ^[114]	4 mm seg. ^[115,116]	5 mm seg. ^[117]		17-45 mm seg. ^[11,118]	35 mm seg. ^[119]	15 mm × 15 mm seg. ^[120]		15 mm ^[11]	
Zygomatic arch			5 mm seg. ^[86,87]							
Nasal bones			20 mm × 8 mm seg. ^[121]							
Femur			6 mm seg., 5 mm seg. ^[122-124]	20 mm seg. ^[125]	21 mm seg. ^[126]		12 cm ³ (volume defect) ^[127]			
Fibula		6 mm seg. ^[128]								
Tibia		1 mm round ^[129]		15 mm seg. (1.5 × diam of shaft) ^[130]						4 mm seg. ^[131]
Metatarsal										
Radius			6 mm seg., 20 mm seg. ^[81,132-134]		25 mm seg. ^[135]					
Ulna				25 mm seg. (1.5 × diam of shaft) ^[136,137]	25 mm seg. ^[138]					

seg., segmental; diam, diameter. (From Ref. [3].)

Table 3 Selected animal model studies of bone pathologies and bone metabolic disease states

Condition	Technique	Species	Comments	Reference
<i>Inflammatory arthritis models</i>				
Osteoarthritis	Naturally occurring	Various species	Good comparative review of pathogenesis and models	[140]
Type II collagen arthritis	Immune response to injection of type II bovine collagen	DBA/LACJ mouse	Acute inflammatory response in extremity joints and loss of bone mineral density	[141,142]
Spondyloarthropathies	Immune response to injection of bacterial flora	HLA-B27 Transgenic rat	Arthritis, loss of bone mineral density, and tarsal joint ankylosis	[143]
Rheumatoid arthritis	Autoimmune diseases	Rat, mouse	Inflammatory response and loss of bone mineral density	[144]
<i>Osteonecrotic and iatrogenic models</i>				
Femoral head osteonecrosis	Surgically created via ischemic and/or cryogenic insults	Emu, dog, sheep	Developed early and late stage human pathology. Emu is a good bipedal loading model	[41,145]
Osteomyelitis	Injections with various pathogens <i>Staphylococcus aureus</i> , <i>Bacteroides fragilis</i> , <i>Escherichia coli</i>	Rat, rabbit, dog, guinea pig, chicken	Decreased bone mineral density, increased risk of fracture, and increased cortical and trabecular bone loss	[38,146–148]
Disuse atrophy and underloading-induced osteopenia	Limb immobilization, denervation, microgravity	Rat, dog, monkey	Trabecular and cortical bone loss, decreases in bone formation, increases in bone resorption	[36,149]
Age-induced osteopenia	Natural aging	Mouse, rat, monkey	Reduction in skeletal mass and osteopenia	[36]
Seasonal-induced osteopenia	Natural aging postmenopausal	Monkey	Age-related decline and skeletal bone mineral content in males and females	[57,150]
Reproductive cycle-induced osteopenia	Orchidectomy	Rat	Cancellous bone osteopenia and reduced bone turnover	[36]
Diet-induced osteopenia	Antler formation, diet, changes, hibernation	Reindeer, deer caribou, bats	Experience transient mineral deficit and reduced bone turnover	[36]
	Pregnancy, lactation, egg laying	Sheep, rat, dog, chicken, turtle	Altered calcium storage and release with subsequent osteopenia	[36]
	Calcium and phosphorus dietary manipulations	Cat	Develop bone disease and decreases, bone mineral density	[12,151]
	Vitamins C and D deficiency	Rat, rabbit	Osteomalacia, rickets, osteopenia	[36]

Alcohol-induced osteopenia	Pre- and postnatal alcohol exposure	Rat	Osteopenia and reduced skeletal growth	[152,153]
Glucocorticoid-induced osteopenia	Glucocorticoid exposure	Rat, rabbit, dog, sheep	Varies from species to species but decreases bone mass	[36]
Bone metastasis-induced osteopenia	Injection of a breast cancer cell line	Mouse	Causes osteolytic bone metastasis and osteoclasia	[154]
	Spontaneous mammary tumor model	Mouse	Causes osteolytic bone metastasis and osteoclasia	[154]
<i>Oral bone loss models</i>				
Periodontitis	Naturally occurring in laboratory-reared models	Baboon, dog, monkey, marmoset	Isolated vertical bone loss in maxillary and mandibular alveolar and basal bone	[58,59,155–160]
	Oral infection by <i>Porphyromonas gingivalis</i>	Mouse, rat, rhesus monkey, baboon, dog, sheep	Reliably produced maxillary and mandibular alveolar and basal bone loss	[161]
<i>Experimentally induced, estrogen-deficient bone loss models</i>				
Post-menopausal osteoporosis and osteopenia	Various procedures	Various models	Good discussions and critiques of various techniques and models	[36,37,139,149,162]
	Ovariectomy	Sheep	Bone loss in mandible, proximal femur and vertebral body	[12,14,163]
	Ovariectomy	Nonhuman primates	Decreased bone mineral density and reduced vertebral cancellous bone volume	[12,28,164]
	Ovariectomy	Ferret	Increased skeletal bone loss	[165]
	Ovariectomy	Dog	Similar bone microarchitecture but does not develop osteoporosis	[12,166,167]
	Ovariectomy	Rat	Developed cortical thinning and significant bone loss	[12,168–170]
	Ovariectomy	Guinea pig	Showed no effect on bone	[12,171]
	Ovariectomy and calcium-restricted diet	Sinclair S-1 minipig	Showed bone removal and loss comparable to humans	[12,172]
Perturbed bone metabolism models				
Genetic models of various metabolic bone diseases	Knockout and over expression models	Transgenic mice	Skeletal phenotypes range from osteopetrosis to osteopenia	[13,39]

(Continued)



Table 3 Selected animal model studies of bone pathologies and bone metabolic disease states (*Continued*)

Condition	Technique	Species	Comments	Reference
Diabetic osteopenia	Various experimental and transgenic insulin deficiency models	Mouse, rat	Decreased bone mass and increased prevalence of skeletal fractures	[173–175]
Growth hormone resistant dwarfism (Laron syndrome)	Congenital models of GH receptor mutation and GH resistance	Sex-linked dwarf chicken, guinea pig, miniature poodle, cattle, and pigs	Severe skeletal growth retardation and dysmorphic facial features	[176,177]
Calvarial hyperostosis and craniosynostosis	Various congenital experimental, and transgenic models	Various species	Features clinical manifestations ranging from microcephaly to craniofacial dysostoses with digital involvement	[15]
Facial clefting	Various congenital, experimental, and transgenic models	Various species	Features clinical manifestations ranging from minimal alveolar and palatal clefts to median and oblique and facial clefts	[178]

(From Ref. [3].)

CONCLUSIONS

Animal models of bone pathologies and wound healing have contributed significantly to the preclinical evaluation of the functional efficacy and biocompatibility of new bone tissue engineering methodologies. Each animal model has its own utility as well as its own limitations for the study of human bone tissue engineering. We have proposed a number of practical criteria for choosing and evaluating various animal models and suggested that appropriate animal model choice should also be linked to levels of hypothesis testing and/or expected extrapolation to the human clinical condition. The continued intelligent and conservative use of animal models should be based on such objective criteria and not necessarily on phyletic affinity alone.

ARTICLES OF FURTHER INTEREST

Animal Surrogate Systems for Toxicity Testing; Bioactive Glass; Biocompatibility Testing; Biomedical Ethics; Bone Remodeling; Bone-Implant Interface; Cell-Material Interaction; Hard Tissue: Biomaterials Interaction; Orthopedic Fixation Devices; Osteoclast Bone Resorption; Osteoinductive Substances and Materials; Tissue Engineering of Bone; Titanium and Its Alloys

REFERENCES

1. Goldstein, S.A. Tissue engineering: functional assessment and clinical outcome. *Ann. N.Y. Acad. Sci.* **2002**, *961*, 183–92.
2. Vacanti, C.A.; Bonassar, L.J. An overview of tissue engineered bone. *Clin. Orthop.* **1999**, *367* (suppl.), S375–S381.
3. Mooney, M.P.; Siegel, M.I. Animal models for bone tissue engineering of critical-sized defects (CSDs), bone pathologies, and orthopedic disease states. In *Fundamentals of Bone Tissue Engineering*; Hollinger, J.O., Einhorn, T.A., Doll, B.A., Sfeir, C., Eds.; CRC Press: Boca Raton, FL, 2005; 217–244.
4. Davey, R.A.; MacLean, H.E.; McManus, J.F.; Findlay, D.M.; Zajac, J.D. Genetically modified animal models as tools for studying bone and mineral metabolism. *J. Bone Miner. Res.* **2004**, *19*, 882–892.
5. Buma, P.; Schreurs, W.; Verdonchot, N. Skeletal tissue engineering—from in vitro studies to large animal models. *Biomaterials* **2004**, *25*, 1487–1495.
6. Fini, M.; Giavarsi, G.; Torricelli, P.; Borsari, V.; Giardino, R.; Nicolini, A.; Carpi, A. Osteoporosis and biomaterial osteointegration. *Biomed. Pharmacother.* **2004**, *58*, 487–493.
7. Hunziker, E.B. Tissue engineering of bone and cartilage. From the preclinical model to the patient. *Novartis Found. Symp.* **2003**, *249*, 70–78.
8. Navia, J.M. *Animal Models in Dental Research*; University Alabama Press: Birmingham, 1977.
9. Prichard, R.W. The need for new animal models—a philosophic approach. *J. Am. Vet. Med. Assoc.* **1978**, *173*, 1208–1209.
10. Siegel, M.I.; Mooney, M.P. Appropriate animal models for craniofacial biology. *Cleft Palate Craniofac. J.* **1990**, *27*, 18–25.
11. Hollinger, J.O.; Kleimmschmidt, J.C. The critical size defect as an experimental model to test bone repair methods. *J. Craniofac. Surg.* **1990**, *1*, 60–68.
12. Newman, E.; Turner, A.S.; Wark, J.D. The potential of sheep for the study of osteopenia: current status and comparison with other animal models. *Bone* **1995**, *16* (4 suppl.), 277S–284S.
13. Turner, R.T.; Maran, A.; Lotinun, S.; Hefferan, T.; Evans, G.L.; Zhang, M.; Sibonga, J.D. Animal models for osteoporosis. *Rev. Endocr. Metab. Disord.* **2001**, *2* (1), 117–127.
14. Turner, A.S. The sheep as a model for osteoporosis in humans. *Vet. J.* **2002**, *163* (3), 232–239.
15. Mooney, M.P.; Siegel, M.I.; Opperman, L.A.. Animal models of craniosynostosis: experimental, congenital, and transgenic. In *Understanding Craniofacial Anomalies: The Etiopathogenesis of Craniosynostosis and Facial Clefting*; Mooney, M.P., Siegel, M.I., Eds.; John W. Wiley and Sons: New York, 2002; 209–250.
16. Murray, R. Animal models for orthopaedic disease—who benefits? *Vet. J.* **2002**, *163* (3), 230–231.
17. Committee on Guidelines for the Use of Animals in Neuroscience and Behavioral Research, National Research Council of the National Academy of Science. *Guidelines for the Care and Use of Mammals in Neuroscience and Behavioral Research*; The National Academies Press: Washington, DC, 2003.
18. Bird, S.; Parlee, M.B. Of mice and men (and women and children): scientific and ethical implications of animal models. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **2000**, *24*, 1291–1227.
19. Baumans, V. Use of animals in experimental research: an ethical dilemma? *Gene Ther.* **2004**, *11*, S64–S66.
20. Rowsey, J.J. Responsibilities in animal experimentation. *Ophthalmic Surg.* **1988**, *19*, 161–162.
21. Fernandes, D. Animal experimentation: necessary or not? *Cleft Palate Craniofac. J.* **1989**, *26*, 258–259.
22. Loeb, J.M.; Hendee, W.R.; Smith, S.J.; Schwartz, M.R. Human vs animal rights: in defense of animal research. *J. Am. Med. Assoc.* **1989**, *262*, 2716–2720.
23. Roach, H.I.; Shearer, J.R.; Archer, C. The choice of an experimental model: a guide for research workers. *J. Bone Joint Surg.* **1989**, *71-B*, 549–553.
24. Taylor, I.; Baum, M.; Cooper, A.; Johnston, I.D. Dilemmas facing surgical research in the '90s. *Ann. R. Coll. Surg. Engl.* **1991**, *73*, 70–72.
25. Cramer, M. Experiments using animals. *Plast. Reconstr. Surg.* **1998**, *102*, 926–927.
26. An, Y.H.; Friedman, R.J. Animal selections in orthopaedic research. In *Animal Models in Orthopaedic Research*; An, Y.H., Friedman, R.J., Eds.; CRC Press: London, 1999; 39–57.

27. Davidson, M.K.; Lindsey, J.R.; Davis, J.K. Requirements and selection of an animal model. *Isr. J. Med. Sci.* **1987**, *23*, 551–555.
28. Rodgers, J.B.; Monier-Faugere, M.C.; Malluche, H. Animal models for the study of bone loss after cessation of ovarian function. *Bone* **1993**, *14* (3), 369–377.
29. Losken, A.; Mooney, M.P.; Siegel, M.I. Comparative analysis of mandibular growth patterns in seven animal models. *J. Oral Maxillofac. Surg.* **1992**, *50*, 490–495.
30. Losken, A.; Mooney, M.P.; Siegel, M.I. Comparative cephalometric study of nasal cavity growth patterns in seven animal models. *Cleft Palate Craniofac. J.* **1994**, *31*, 17–23.
31. Smith, G.C. Value of nonhuman primates in predicting disposition of drugs in man. *Ann. N.Y. Acad. Sci.* **1969**, *162*, 600–603.
32. Reynolds, H.H. Nonhuman primates in the study of toxicological effects on the central nervous system: a review. *Ann. N.Y. Acad. Sci.* **1969**, *162*, 604–609.
33. Langer, R.; Vacanti, J.P. Tissue engineering. *Science* **1993**, *260*, 920–926.
34. Putnam, A.J.; Mooney, D.J. Tissue engineering using synthetic extracellular matrices. *Nat. Med.* **1996**, *2*, 824–826.
35. Alsberg, E.; Hill, E.E.; Mooney, D.J. Craniofacial tissue engineering. *Crit. Rev. Oral Biol. Med.* **2001**, *12*, 64–75.
36. Miller, S.C.; Bowman, B.M.; Jee, W.S. Available animal models of osteopenia—small and large. *Bone* **1995**, *17* (4 suppl.), 117S–123S.
37. Mosekilde, L. Assessing bone quality—animal models in preclinical osteoporosis research. *Bone* **1995**, *17* (4 suppl.), 343S–352S.
38. Cremieux, A.C.; Carbon, C. Experimental models of bone and prosthetic joint infections. *Clin. Infect. Dis.* **1997**, *25* (6), 1295–1302.
39. McCauley, L.K. Transgenic mouse models of metabolic bone disease. *Curr. Opin. Rheumatol.* **2001**, *13* (4), 316–325.
40. Green, D.; Walsh, D.; Mann, S.; Oreffo, R.O. The potential of biomimesis in bone tissue engineering: lessons from the design and synthesis of invertebrate skeletons. *Bone* **2002**, *30* (6), 810–815.
41. Conzemi, M.G.; Brown, T.D.; Zhang, Y.; Robinson, R.A. A new animal model of femoral head osteonecrosis: one that progresses to human-like mechanical failure. *J. Orthop. Res.* **2002**, *20* (2), 303–309.
42. Krebsbach, P.H.; Gehron-Robey, P. Dental and skeletal stem cells: potential cellular therapeutics for craniofacial regeneration. *J. Dent. Ed.* **2002**, *66*, 766–773.
43. Rodriguez-Merchan, E.C.; Forriol, F. Nonunion: general principles and experimental data. *Clin. Orthop. Relat. Res.* **2004**, *419*, 4–12.
44. Blin-Wakkach, C.; Bernard, F.; Carle, G.F. Osteoporosis, from mouse to man. *Med. Sci. (Paris)* **2004**, *20*, 61–67.
45. Allen, M.J. Biochemical markers of bone metabolism in animals: uses and limitations. *Vet. Clin. Pathol.* **2003**, *32*, 101–113.
46. Plaza, S.M.; Lamson, D.W. Vitamin K2 in bone metabolism and osteoporosis. *Altern. Med. Rev.* **2005**, *10*, 24–35.
47. Salgado, A.J.; Coutinho, O.P.; Reis, R.L. Bone tissue engineering: state of the art and future trends. *Macromol. Biosci.* **2004**, *4*, 743–765.
48. Kruyt, M.C.; van Gaalen, S.M.; Oner, F.C.; Verbout, A.J.; de Bruijn, J.D.; Bheret, W.J.A. Bone tissue engineering and spinal fusion: the potential of hybrid constructs by combining osteoprogenitor cells and scaffolds. *Biomaterials* **2004**, *25*, 1463–1473.
49. Evans, C.H.; Ghivizzani, S.C.; Robbins, P.D. Orthopedic gene therapy. *Clin. Orthop. Relat. Res.* **2004**, *429*, 316–329.
50. Boyan, B.D.; Lohmann, C.H.; Romero, J.; Schwartz, Z. Bone and cartilage tissue engineering. *Clin. Plast. Surg.* **1999**, *26*, 629–645.
51. Goldstein, S.A.; Patil, P.V.; Moalli, M.R. Perspectives on tissue engineering of bone. *Clin. Orthop.* **1999**, *367* (suppl.), S419–S423.
52. Alsberg, E.; Anderson, K.W.; Albeiruti, A.; Rowley, J.A.; Mooney, D.J. Engineering growing tissues. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99* (19), 12,025–12,030.
53. Athanasiou, K.A. Fundamentals of biomechanics in tissue engineering of bone. *Tissue Eng.* **2000**, *6* (4), 361–381.
54. Young, J.Z. *The Life of Mammals*, 3rd Ed.; Clarendon Press: Oxford, 1975.
55. Torricelli, P.; Fini, M.; Giavaresi, G.; Giardino, R. In vitro models to test orthopedic biomaterials in view of their clinical applications in osteoporotic bone. *Int. J. Artif. Organs* **2004**, *27*, 658–663.
56. Khan, S.N.; Lane, J.M. Spinal fusion surgery: animal models for tissue-engineered bone constructs. *Biomaterials* **2004**, *25*, 1475–1485.
57. Jerome, C.P.; Peterson, P.E. Nonhuman primate models in skeletal research. *Bone* **2001**, *29* (1), 1–6.
58. Aufdemorte, T.B.; Boyan, B.D.; Fox, W.C.; Miller, D. Diagnostic tools and biologic markers: animal models in the study of osteoporosis and oral bone loss. *J. Bone Miner. Res.* **1993**, *8* (suppl. 2), S529–S534.
59. Aufdemorte, T.B.; Fox, W.C.; Miller, D.; Buffum, K.; Holt, G.R.; Carey, K.D. A non-human primate model for the study of osteoporosis and oral bone loss. *Bone* **1993**, *14*, 581–586.
60. Farris, E.F.; Griffith, J.Q. *The Rat in Laboratory Investigation*; Hafner Press: New York, 1949.
61. Enlow, D.H. *Handbook of Facial Growth*; Saunders: New York, 1990.
62. Greene, E.C. *Anatomy of the Rat*; Hafner Publishing Co.: New York, 1968.
63. Cooper, G.; Schiller, S. *Anatomy of the Guinea Pig*; Harvard University Press: Cambridge, MA, 1975.
64. Kier, E.L. Phylogenetic and ontogenetic changes of the brain relevant to the evolution of the skull. In *Development of the Basicranium*; Bosma, J., Ed.; DHEW/NIH Publication #76-989; Bethesda, MD, 1976; 468–499.
65. Getty, R. *Sissons and Grossman's the Anatomy of the Domestic Animals*; WB Saunders: Philadelphia, PA, 1975.

66. Merckx, M.A.; Maltha, J.C.; Freihofer, H.P.; Kuijpers-Jagtman, A.M. Incorporation of three types of bone block implants in the facial skeleton. *Biomaterials* **1999**, *20* (7), 639–645.
67. Merckx, M.A.; Maltha, J.C.; Freihofer, H.P.; Kuijpers-Jagtman, A.M. Incorporation of particulated bone implants in the facial skeleton. *Biomaterials* **1999**, *20* (21), 2029–2035.
68. Hartman, C.F.; Strauss, W.L. *The Anatomy of the Rhesus Monkey*; Hafner Publishing Co.: New York, 1933.
69. Elliot, R. *Reighard and Jennings' Anatomy of the Cat*; Holt, Rinehart, and Winston: New York, 1963.
70. Swindler, D.R.; Wood, C.D. *An Atlas of Primate Gross Anatomy: Baboon, Chimpanzee, and Man*; Robert, E., Ed.; Krieger Publishing: Malabar, FL, 1982.
71. Sirianni, J. Nonhuman primates as models for human craniofacial growth. In *Nonhuman Primate Models for Human Growth and Development*; Watts, E.S., Ed.; A.R. Liss: New York, 1985; 95–124.
72. Enlow, D.H.; McNamara, J. The neurocranial basis for facial form and pattern. *Angle Orthod.* **1973**, *43*, 256–271.
73. Sirianni, J.E.; Swindler, D.R. A review of the post-natal craniofacial growth in old world monkeys and apes. *Yrbk. Phys. Anthropol.* **1979**, *22*, 80–104.
74. Siebert, J.R.; Swindler, D.R. Evolutionary changes in the midface and mandible: establishing the primate form. In *Understanding Craniofacial Anomalies: The Etiopathogenesis of Craniosynostosis and Facial Clefting*; Mooney, M.P., Siegel, M.I., Eds.; John W. Wiley and Sons: New York, 2002; 345–378.
75. Geddes, A.D. Animal models of bone disease. In *Principles of Bone Biology*; Bilezikians, J.P., Raisz, L.G., Rodan, G.A., Eds.; Academic Press: San Diego, 1996; 1343–1354.
76. Enlow, D.H.; Azuma, M. Functional growth boundaries in the human and mammalian face. *Birth Defect* **1975**, *11*, 217–230.
77. Whang, K. Engineering bone regeneration with bioabsorbable scaffolds with novel microarchitecture. *Tissue Eng.* **1999**, *5*, 35–51.
78. Service, R.F. Bone remodeling and repair. *Science* **2000**, *289*, 1421–1640.
79. Murphy, W.L. Sustained release of vascular endothelial growth factor from mineralized polylactide-co-glycolic scaffolds for tissue engineering. *Biomaterials* **2000**, *21*, 2521–2527.
80. Shea, L.D. Engineered bone development from pre-osteoblast cell line on three-dimensional scaffolds. *Tissue Eng.* **2000**, *6*, 605–617.
81. Mackenzie, D.J. Recombinant human acidic fibroblast growth factor and fibrin carrier regenerates bone. *Plast. Reconstr. Surg.* **2001**, *107* (4), 989–996.
82. Hutmacher, D.W. Scaffold design and fabrication technologies for engineering tissues—state of the art future perspectives. *J. Biomater. Sci. Polym. Ed.* **2001**, *12*, 102–124.
83. Einhorn, T.; Lee, C.A. Bone regeneration: new findings and potential clinical applications. *J. Am. Acad. Orthop. Surg.* **2001**, *9*, 157–165.
84. Jansen, J.A. Laboratory animal models for the manufacture of tissue substitutes. *Ned. Tijdschr. Tandheelkd.* **2004**, *111*, 490–493.
85. Schmitz, J.P.; Hollinger, J.O. The critical size defect as an experimental model for craniomandibulofacial nonunions. *Clin. Orthop.* **1986**, *205*, 299–308.
86. Mundell, R.; Mooney, M.P.; Siegel, M.I.; Losken, A. Osseous guided tissue regeneration using a collagen barrier membrane. *J. Oral Maxillofac. Surg.* **1993**, *51*, 1004–1012.
87. Mooney, M.P.; Mundell, R.D.; Stetzer, K.; Ochs, M.W.; Milch, E.A.; Buckley, M.J.; Siegel, M.I. The effects of guided tissue regeneration and fixation technique on osseous wound healing in rabbit zygomatic arch osteotomies. *J. Craniofac. Surg.* **1996**, *7*, 46–53.
88. Stetzer, K.M.; Cooper, G.; Gassner, R.; Kapucu, R.; Mundell, R.; Mooney, M.P. The effects of fixation type and guided tissue regeneration on maxillary osteotomy healing in rabbits. *J. Oral Maxillofac. Surg.* **2002**, *60*, 427–436.
89. Frame, J.W. A convenient animal model for testing bone substitution materials. *J. Oral Surg.* **1980**, *38*, 176–180.
90. Prolo, D.J.; Pedrotti, P.W.; Bures, K.P.; Oklund, S. Superior osteogenesis in transplanted allogenic canine skull following chemical sterilization. *Clin. Orthop.* **1982**, *108*, 203–207.
91. Lee, J.Y.; Musgrave, D.; Pelinkovic, D.; Fukushima, K.; Cummins, J.; Usas, A.; Robbins, P.; Fu, F.H.; Huard, J. Effect of bone morphogenetic protein-2-expressing muscle-derived cells on healing of critical-sized bone defects in mice. *J. Bone Joint Surg. Am.* **2001**, *83-A*, 1032–1039.
92. Takagi, K.; Urist, M.R. The role of bone marrow induced repair of femoral massive diaphyseal defects. *Clin. Orthop.* **1982**, *171*, 224–231.
93. Hollinger, J.O.; Schmitz, J.P.; Yaskovich, R.; Long, M.M.; Prasad, K.U.; Urry, D.W. A synthetic polypentapeptide of elastin for initiating calcification. *Calcif. Tissue Int.* **1988**, *42*, 231–236.
94. Winn, S.R.; Schmitt, J.M.; Buck, D.; Hu, Y.; Grainger, D.; Hollinger, J.O. Tissue-engineered bone biomimetic to regenerate calvarial critical-sized defects in athymic rats. *J. Biomed. Mater. Res.* **1999**, *45* (4), 414–421.
95. Clokie, C.M.; Moghadam, H.; Jackson, M.T.; Sandor, G.K. Closure of critical sized defects with allogenic and alloplastic bone substitutes. *J. Craniofac. Surg.* **2002**, *13* (1), 111–121.
96. Mooney, M.P.; Burrows, A.M.; Smith, T.D.; Losken, H.W.; Opperman, L.A.; Dechant, J.; Kreithen, A.M.; Kapucu, R.; Cooper, G.M.; Ogle, R.C.; Siegel, M.I. Correction of coronal suture synostosis using suture and dura mater allografts in rabbits with familial craniosynostosis. *Cleft Palate Craniofac. J.* **2001**, *38*, 72–91.
97. Vesala, A.L.; Kallioinen, M.; Tormala, P.; Kellomaki, M.; Waris, T.; Ashammakhi, N. Bone tissue engineering: treatment of cranial bone defects in rabbits using self-reinforced poly-L,D lactide 96/4 sheets. *J. Craniofac. Surg.* **2002**, *13* (5), 607–613.

98. Costantino, P.D.; Friedman, C.D.; Jones, K.; Chow, L.C.; Sisson, G.A. Experimental hydroxyapatite cement cranioplasty. *Plast. Reconstr. Surg.* **1992**, *90* (2), 174–185.
99. Urist, M.R. New advanced in bone research. *West. J. Med.* **1984**, *141*, 71.
100. Shang, Q.; Wang, Z.; Liu, W.; Shi, Y.; Cui, L.; Cao, Y. Tissue-engineered bone repair of sheep cranial defects with autologous bone marrow stromal cells. *J. Craniofac. Surg.* **2001**, *12* (6), 586–593.
101. Hollinger, J.O.; Mark, D.E.; Bach, D.E.; Reddi, A.H.; Seyfer, A.E. Calvarial bone regeneration using osteogenin. *J. Oral Maxillofac. Surg.* **1989**, *47*, 1182–1186.
102. Ripamonti, U.; Ramoshebi, L.N.; Matsaba, T.; Tasker, J.; Crooks, J.; Teare, J. Bone induction by BMPs/OPs and related family members in primates. *J. Bone Joint Surg. Am.* **2001**, *83-A* (suppl. 1, pt 2), S116–S127.
103. Bardach, J.; Roberts, D.M.; Klausner, E.C. Influence of two-flap palatoplasty on facial growth in rabbits. *Cleft Palate Craniofac. J.* **1979**, *16*, 402–411.
104. Bardach, J.; Roberts, D.M.; Yale, R.; Rosewall, D.; Mooney, M. The influence of simultaneous cleft lip and palate repair on facial growth in rabbits. *Cleft Palate Craniofac. J.* **1980**, *17* (4), 309–318.
105. Freng, A. Transversal maxillary growth in experimental submucous mid-palatal clefts. A roentgen-cephalometric study in the cat. *Scand. J. Plast. Reconstr. Surg.* **1979**, *13* (3), 409–416.
106. Bardach, J.; Mooney, M.; Bardach, E. The influence of two-flap palatoplasty on facial growth in beagles. *Plast. Reconstr. Surg.* **1982**, *69* (6), 927–936.
107. Bardach, J.; Martin, R.; Mooney, M.P.; Kelly, K.M.; Albright, J. Bone formation in the canine palate following partial resection. In *Normal and Abnormal Bone Growth: Basic and Clinical Research*; Sarnat, B.G., Ed.; W.B. Saunders: New York, 1985; 365–377.
108. Kawata, T.; Kohno, S.; Fujita, T.; Sugiyama, H.; Tokimasa, C.; Kaku, M.; Tsutsui, K.; Tanne, K. Transplantation of new autologous biomaterials into jaw cleft. *J. Int. Med. Res.* **2001**, *29*, 287–291.
109. Takano-Yamamoto, T.; Kawakami, M.; Sakuda, M. Defects of the rat premaxilla as a model of alveolar clefts for testing bone-inductive agents. *J. Oral Maxillofac. Surg.* **1993**, *51*, 887–891.
110. Verwoerd, C.D.A.; Verwoerd-Verhoef, H.L.; Urbanis, N.A.M. Skulls with facial clefts. Experimental surgery on the facial skeleton. *Acta Otolaryngol.* **1976**, *81*, 249–256.
111. Eisbach, K.J.; Bardach, J.; Klausner, E.C. The influence of primary unilateral cleft lip repair on facial growth, part II: direct cephalometry of the skull. *Cleft Palate Craniofac. J.* **1978**, *15*, 109–117.
112. Verschuere, D.; Gassner, R.; Mitchell, R.; Mooney, M.P. Le Fort I osteotomy healing in rabbits with the use of guided tissue regeneration (GTR). *Int. J. Oral Maxillofac. Surg.* **2005**, *34*, 650–655.
113. El Deeb, M.; Horswell, B.; Waite, D. A primate model for producing experimental alveolar cleft defects. *J. Oral Maxillofac. Surg.* **1985**, *43*, 523–527.
114. Alden, T.D.; Jane, J.A., Jr.; Hudson, S.B.; Helm, G.A. The use of bone morphogenetic protein gene therapy in craniofacial bone repair. *J. Craniofac. Surg.* **2000**, *11* (1), 24–30.
115. Kaban, L.B.; Glowacki, J.; Murray, J.E. Repair of experimental bony defects in rats. *Surg. Forum* **1979**, *30*, 519–521.
116. Saadeh, P.B.; Khosla, R.K.; Mehrara, B.J.; Steinbrech, D.S.; McCormick, S.A.; DeVore, D.P.; Longaker, M.T. Repair of a critical size defect in the rat mandible using allogenic type I collagen. *J. Craniofac. Surg.* **2001**, *12* (6), 573–579.
117. Kahnberg, K. Restoration of mandibular jaw defects in the rabbit by subperiosteally implanted Teflon[®] mantle leaf. *Int. J. Oral Surg.* **1979**, *8*, 449–456.
118. Hollinger, J.O.; Schmitz, J.P. Restoration of bone discontinuities in dogs using a biodegradable implant. *J. Maxillofac. Surg.* **1987**, *45*, 594–600.
119. Schliephake, H.; Knebel, J.W.; Aufderheide, M.; Tauscher, M. Use of cultivated osteoprogenitor cells to increase bone formation in segmental mandibular defects: an experimental pilot study in sheep. *Int. J. Oral Maxillofac. Surg.* **2001**, *30* (6), 531–537.
120. Henkel, K.O.; Gerber, T.; Dorfling, P.; Hartel, J.; Jonas, L.; Gundlach, K.K.; Bienengraber, V. Stimulating regeneration of bone defects by implantation of bioceramics and autologous osteoblast transplantation. *Mund Kiefer Gesichtschir.* **2002**, *6* (2), 59–65.
121. Lindsey, W.H.; Franz, D.A.; Toung, J.S.; London, S.D.; Ogle, R.O. A nasal critical-size defect: an experimental model for the evaluation of facial osseous repair techniques. *Arch. Otolaryngol. Head Neck Surg.* **1998**, *124* (8), 912–915.
122. Einhorn, T.A.; Lane, J.M.; Burstein, A.H.; Kopman, C.R.; Vigorita, V.J. The healing of segmental bone defects induced by demineralized bone matrix. *J. Bone Joint Surg.* **1984**, *66A*, 274–279.
123. Lieberman, J.R.; Daluiski, A.; Stevenson, S.; Wu, L.; McAllister, P.; Lee, Y.P.; Kabo, J.M.; Finerman, G.A.; Berk, A.J.; Witte, O.N. The effect of regional gene therapy with bone morphogenetic protein-2-producing bone-marrow cells on the repair of segmental femoral defects in rats. *J. Bone Joint Surg. Am.* **1999**, *81* (7), 905–917.
124. Lane, J.M.; Yasko, A.W.; Tomin, E.; Cole, B.J.; Waller, S.; Browne, M.; Turek, T.; Gross, J. Bone marrow and recombinant human bone morphogenetic protein-2 in osseous repair. *Clin. Orthop.* **1999**, *361*, 216–227.
125. Wheeler, D.L.; Eschbach, E.J.; Hoellrich, R.G.; Montfort, M.J.; Chamberland, D.L. Assessment of resorbable bioactive material for grafting of critical-size cancellous defects. *J. Orthop. Res.* **2000**, *18* (1), 140–148.
126. Bruder, S.P.; Kraus, K.H.; Goldberg, V.M.; Kadiyala, S. The effect of implants loaded with autologous mesenchymal stem cells on the healing of canine segmental bone defects. *J. Bone Joint Surg. Am.* **1998**, *80* (7), 985–996.
127. Lange, T.A.; Zerwekh, J.E.; Peek, R.D.; Mooney, V.; Harrison, B.H. Granular tricalcium phosphate in large

- cancellous defects. *Ann. Clin. Lab. Sci.* **1986**, *16*, 467–472.
128. Narang, R.; Laskin, D.M. Experimental osteogenesis at fracture sites and gaps. *J. Oral Surg.* **1976**, *34*, 225–231.
 129. Landry, P.S.; Marino, A.A.; Sadasivan, K.K.; Albright, J.A. Bone injury response. An animal model for testing theories of regulation. *Clin. Orthop.* **1996**, *332*, 260–273.
 130. Toombs, J.P.; Wallace, L.J.; Bjorling, D.E.; Rowland, G.N. Evaluation of Key's hypothesis in the feline tibia: an experimental model for augmented bone healing studies. *Am. J. Vet. Res.* **1985**, *46*, 513–518.
 131. Collier, M.A.; Brighton, C.T.; Norrdin, R.; Twardock, A.R.; Rendano, V.T. Direct current stimulation of bone production in the horse: preliminary study with a "gap healing" model. *Am. J. Vet. Res.* **1985**, *46*, 612–621.
 132. Ben-fu, C.; Xue-ming, T. Ultrastructural investigation of experimental non-union of fractures. *Chin. Med. J.* **1986**, *99*, 207–214.
 133. Wheeler, D.L.; Chamberland, D.L.; Schmitt, J.M.; Buck, D.C.; Brekke, J.H.; Hollinger, J.O.; Joh, S.P.; Suh, K.W. Radiomorphometry and biomechanical assessment of recombinant human bone morphogenetic protein 2 and polymer in rabbit radius osteotomy model. *J. Biomed. Mater. Res.* **1998**, *43* (4), 365–373.
 134. Hollinger, J.O.; Schmitt, J.M.; Buck, D.C.; Shannon, R.; Joh, S.P.; Zegzula, H.D.; Wozney, J. Recombinant human bone morphogenetic protein-2 and collagen for bone regeneration. *J. Biomed. Mater. Res.* **1998**, *43* (4), 356–364.
 135. Sciadini, M.F.; Dawson, J.M.; Johnson, K.D. Evaluation of bovine-derived bone protein with a natural coral carrier as a bone-graft substitute in a canine segmental defect model. *J. Orthop. Res.* **1997**, *15* (6), 844–857.
 136. Key, J.A. The effects of local calcium depot on osteogenesis and healing of fractures. *J. Bone Joint Surg. (Am.)* **1934**, *16*, 176–184.
 137. Nilsson, O.S.; Urist, M.R.; Dawson, E.G.; Schmalzried, T.P.; Finerman, G.A. Bone repair induced by bone morphogenic protein in ulnar defects in dogs. *J. Bone Joint Surg.* **1986**, *63B*, 635–642.
 138. Salkeld, S.L.; Patron, L.P.; Barrack, R.L.; Cook, S.D. The effect of osteogenic protein-1 on the healing of segmental bone defects treated with autograft or allograft bone. *J. Bone Joint Surg. Am.* **2001**, *83-A* (6), 803–816.
 139. Thompson D.D.; Simmons, H.A.; Pirie, C.M.; Ke, H.Z. FDA guidelines and animal models for osteoporosis. *Bone* **1995**, *17* (4 suppl.), 125S–133S.
 140. Oegema, T.R.; Visco, D. Animal models of osteoarthritis. In *Animal Models in Orthopaedic Research*; An, Y.H., Friedman, R.J., Eds.; CRC Press: London, 1999; 349–367.
 141. Badger, A.M.; Bradbeer, J.N.; Votta, B.; Lee, J.C.; Adams, J.L.; Griswold, D.E. Pharmacological profile of SB 203580, a selective inhibitor of cytokine suppressive binding protein/p38 kinase, in animal models of arthritis, bone resorption, endotoxin shock and immune function. *J. Pharmacol. Exp. Ther.* **1996**, *279* (3), 1453–1461.
 142. Wooley, P.H. Collagen-induced arthritis in the mouse. *Methods Enzymol.* **1988**, *162*, 361–373.
 143. Breban, M. Animal models and in vitro models for the study of aetiopathogenesis of spondyloarthropathies. *Baillieres Clin. Rheumatol.* **1998**, *12* (4), 611–626.
 144. Joe, B.; Griffiths, M.M.; Remmers, E.F.; Wilder, R.L. Animal models of rheumatoid arthritis and related inflammation. *Curr. Rheumatol. Rep.* **1999**, *1* (2), 139–148.
 145. Phillips, T.W.; Johnston, G.; Wood, P. Selection of an animal model for resurfacing hip arthroplasty. *J. Arthroplasty* **1987**, *2* (2), 111–117.
 146. Mader, J.P. Animal models of osteomyelitis. *Am. J. Med.* **1985**, *78* (suppl. 6B), 213–217.
 147. Norden, C.W. Lessons learned from animal models of osteomyelitis. *Rev. Infect. Dis.* **1988**, *10*, 103–110.
 148. Rissing, J.P. Animal models of osteomyelitis: knowledge, hypothesis, and speculation. *Infect. Dis. Clin. North Am.* **1990**, *4* (3), 377–390.
 149. Cesnjaj, M.; Stavljenic, A.; Vukicevic, S. In vivo models in the study of osteopenias. *Eur. J. Clin. Chem. Clin. Biochem.* **1991**, *29* (4), 211–219.
 150. Black, A.; Tilmont, E.M.; Handy, A.M.; Scott, W.W.; Shapses, S.A.; Ingram, D.K.; Roth, G.S.; Lane, M.A. A nonhuman primate model of age-related bone loss: a longitudinal study in male and premenopausal female rhesus monkeys. *Bone* **2001**, *28* (3), 295–302.
 151. Draper, H.H. *Advances in Nutritional Research*; Plenum Press: New York, 1985; 172–186.
 152. Riesenfeld, A.; Siegel, M.I.; Mooney, M.P.; Seroky, J.T.; Taylor, A.B. The effects of perinatal alcohol exposure and dietary calcium supplements on skeletal and dental growth in rats. *Acta Anat.* **1991**, *140*, 1–7.
 153. Turner, R.T.; Aloia, R.C.; Segel, L.D.; Hannon, K.S.; Bell, N.H. Chronic alcohol treatment results in disturbed vitamin D metabolism and skeletal abnormalities in rats. *Alcohol. Clin. Exp. Res.* **1988**, *12*, 159–162.
 154. Yoneda, T.; Michigami, T.; Yi, B.; Williams, P.J.; Niewolna, M.; Hiraga, T. Use of bisphosphonates for the treatment of bone metastasis in experimental animal models. *Cancer Treat. Rev.* **1999**, *25* (5), 293–299.
 155. Miller, D.R.; Aufdemorte, T.B.; Fox, W.C.; Waldrop, T.C.; Mealey, B.L.; Brunsvold, M.A. Periodontitis in the baboon: a potential model for human disease. *J. Periodontal Res.* **1995**, *30* (6), 404–409.
 156. Haney, J.M.; Zimmerman, G.J.; Wikesjo, U.M. Periodontal repair in dogs: evaluation of the natural disease model. *J. Clin. Periodontol.* **1995**, *22* (3), 208–213.
 157. Levy, B.M. Primates in dental research. In *Proceedings of the Third Conference on Experimental Medicine and Surgery in Primates, Medical Primatology*; Goldsmith, E.I., Moor Jankowski, J., Eds.; Karger Press: Basel, 1971; 859–869.
 158. Levy, B.M. Animal model of human disease: chronic destructive periodontitis (periodontal disease, pyorrhea alveolaris, pyorrhea). *Am. J. Pathol.* **1976**, *83*, 637–640.

159. Levy, B.M. Animal analogues for the study of dental and oral diseases. *Dev. Biol. Stand.* **1980**, *45*, 51–59.
160. Dreizen, S.; Levy, B.M. Monkey models in dental research. *J. Med. Primatol.* **1977**, *6*, 133–144.
161. Genco, C.A.; Van Dyke, T.; Amar, S. Animal models for Porphyromonas gingivalis mediated periodontal disease. *Trends Microbiol.* **1998**, *6* (11), 444–449.
162. Sietsema, W.K. Animal models of cortical porosity. *Bone* **1995**, *17* (4 suppl.), 297S–305S.
163. Lamghari, M.; Huet, H.; Laurent, A.; Berland, S.; Lopez, E. A model for evaluating injectable bone replacements in the vertebrae of sheep: radiological and histological study. *Biomaterials* **1999**, *20*, 2107–2114.
164. Miller, L.C.; Weaver, D.S.; McAlister, J.A.; Koritnik, D.R. Effects of ovariectomy on vertebral bone in the cynomolgus monkey (*Macaca fascicularis*). *Calcif. Tissue Int.* **1986**, *38*, 62–65.
165. Mackey, M.S.; Stevens, M.L.; Ebert, D.C.; Tressler, D.L.; Combs, K.S.; Lowry, C.K.; Smith, P.N.; McOsker, J.E. The ferret as a small animal model with BMU-based remodeling for skeletal research. *Bone* **1995**, *17* (4 suppl.), 191S–196S.
166. Raab-Cullen, D.M.; Akhter, M.P.; Kimmel, D.B.; Recker, R.R. On animal models for studying bone adaptation. *Calcif. Tissue Int.* **1994**, *55* (4), 317–318.
167. Kimmel, D.B. Animal models for in vivo experimentation in osteoporosis research. In *Osteoporosis*; Marcus, J., Feldman, K.F., Kelsey, E.C., Eds.; Academic Press: San Diego, 1996; 671–690.
168. Wronski, T.J.; Yen, C.F. The ovariectomized rat as an animal model for postmenopausal bone loss. *Cells Mater.* **1991**, *4* (suppl. 1), 69–74.
169. Frost, H.M.; Jee, W.S.S. On the rat model of human osteopenia and osteoporosis. *Bone Miner.* **1992**, *18*, 227–236.
170. Kalu, D.N. The ovariectomized rat model of postmenopausal bone loss. *Bone Miner.* **1991**, *15*, 175–192.
171. Vanderschueren, D.; Van Herck, E.; Suiker, A.M.; Allewaert, K.; Visser, W.J.; Geusens, P.; Bouillon, R. Bone and mineral metabolism in the adult guinea pig: long-term effects of estrogen and androgen deficiency. *J. Bone Miner. Res.* **1992**, *7*, 1407–1415.
172. Mosekilde, L.; Weisbrode, S.E.; Safron, J.A.; Stills, H.F.; Jankowsky, M.L.; Ebert, D.C.; Danielsen, C.C.; Sogaard, C.H.; Franks, A.F.; Stevens, M.L. Calcium-restricted ovariectomized Sinclair S-1 minipigs: an animal model of osteopenia and trabecular plate perforation. *Bone* **1993**, *14*, 379–382.
173. Kagel, E.M.; Majeska, R.J.; Einhorn, T.A. Effects of diabetes and steroids on fracture healing. *Curr. Opin. Orthop.* **1995**, *6* (5), 7–13.
174. Hough, F.S. Alterations of bone and mineral metabolism in diabetes mellitus, part I, an overview. *S. Afr. Med. J.* **1987**, *72* (2), 116–119.
175. Zapf, J. Growth promotion by insulin-like growth factor I in hypophysectomized and diabetic rats. *Mol. Cell Endocrinol.* **1998**, *140* (1–2), 143–149.
176. Daughaday, W.H. Animal models of abnormal GH receptor binding and postbinding mechanisms. *Pediatr. Adolesc. Endocrinol.* **1992**, *24*, 282–287.
177. Hull, K.L.; Harvey, S. Growth hormone resistance: clinical states and animal models. *J. Endocrinol.* **1999**, *163* (2), 165–172.
178. Diewert, V.M.; Lozanoff, S. Animal models of facial clefting—experimental, congenital, and transgenic. In *Understanding Craniofacial Anomalies: The Etiopathogenesis of Craniosynostosis and Facial Clefting*; Mooney, M.P., Siegel, M.I., Eds.; John W. Wiley and Sons: New York, 2002; 251–272.

Animal Surrogate Systems for Toxicity Testing

A

Michael L. Shuler

Department of Biomedical Engineering, and School of Chemical and Biomolecular Engineering, Cornell University, Ithaca, New York, U.S.A.

Gretchen J. McAuliffe

Daniel A. Tatosian

School of Chemical and Biomolecular Engineering, Cornell University, Ithaca, New York, U.S.A.

INTRODUCTION

The chance of a new drug candidate gaining approval fell to 8% in 2004, with the majority of failures happening in phase III clinical trials.^[1] Clinical trials make up the majority of the costs in drug development; a 10% improvement in predicting clinical trial failure could save as much as \$100 million per drug.^[1] Both the expense of testing new drugs and chemicals and the large number of new compounds available to be tested makes a rapid, inexpensive method for accurately assessing efficacy and toxicity in humans essential in the pharmaceutical and chemical industries. An understanding of a drug's absorption, distribution, metabolism, elimination, and toxicity (ADMET) in the human body is important in determining minimum dosages and effective delivery schemes. A non-pharmaceutical chemical's action on the body must be characterized to establish safe exposure levels and first aid procedures.

Animal studies are the primary method used to determine toxicological and pharmacological profiles, but animal experiments can take months to complete, and the cost of testing a single substance can be millions of dollars.^[2,3] Recent estimates have placed the preclinical cost of new drug development between \$335M^[2] and \$620M^[4] with approximately \$90M of that expense in preclinical safety.^[4,5] Furthermore, animal cells do not always react with chemicals and drugs in the same way as human cells do.^[3] Alternatives to animal studies include computational (in silico) models and in vitro methods using one or more living cell type.

The goal of this article is to give a brief overview of many of the animal surrogate systems used for toxicity testing, including mathematical models, single cell type models, barrier models, cocultures, three-dimensional models, and multicell type devices. The reader will see how toxicity can be predicted prior to animal or human studies. Although the use of subcellular components (DNA microarrays, antigen binding, etc.) as animal surrogates in toxicity testing is an important

emerging technology, it will not be covered in this article.

COMPUTATIONAL MODELS

The utilization of computational methods for toxicity studies involves several different length scales, ranging from molecular interactions to whole organism models. At the most local scale, molecular models are used to determine ligand–receptor binding, such as the so-called docking method. Models of larger scope include networks or pathways of interactions, such as gene network systems. Approximations of physiological behavior allow the construction of bulk drug distribution models, such as pharmacokinetic (PK) models or the more detailed, organ-based, physiologically based pharmacokinetic (PBPK) models. Additionally, responses of organisms to different doses of a chemical can be captured with pharmacodynamic (PD) equations. All of these tools are used in research labs in academia and industry to increase throughput of toxicity screens and drug candidate studies.

Chemical Structure-Related Models

Most modeling for toxicity in animals requires experimental knowledge of many parameters for a given substance. Predicting these parameters a priori from molecular structures enables researchers to perform computer simulations of toxicity experiments. The ability to make toxicity predictions without actually synthesizing a chemical is of particular interest in drug discovery, where drug candidates may be screened without requiring laboratory work. Quantitative structure–activity relationships (QSARs) methods and molecular docking techniques are two of the more popular molecular structure-based prediction approaches in the literature.

The concept of QSARs is simple: Compounds that could be considered similar in their action in vivo can

be modeled as variations in a set of descriptors. Descriptors are properties of a chemical that range from molecular weight and octanol/water partition coefficient to more complex properties such as hydrogen bonding donor/acceptor density.^[6] Quantitative structure–activity relationships have been used for the last 40 years to model various ADMET properties, and have already been proven useful in toxicity screening of new chemicals.^[7] Frequently QSARs are used to directly model toxicity for one or more related activity mechanisms, such as the risk of toxicity owing to endocrine disrupting chemicals.^[8] They can also be used to model other chemical properties such as permeability through the skin.^[7]

The use of a QSAR model requires initial training with an experimental data set. Methods for this model generation and training vary from regression techniques to the use of neural networks and self-organizing maps.^[6] Once a model is generated, it is then validated by attempting to predict the remainder of the data set. Following validation, test predictions of additional compounds are made and compared against a new data set. Accuracy of these models varies, but it is not uncommon to have close to 90% accuracy in predicting certain conditions such as carcinogenicity.^[9] This accuracy typically decreases when additional modes of chemical interaction are included.

The interaction between a protein and a chemical can be estimated using molecular simulations. Often these quantum mechanical “docking” studies are combined with mathematical models such as QSARs to help predict which chemicals should be accurately modeled by a data set. These studies require 3-D models of the test chemicals and the protein being studied, which can either be obtained by x-ray crystallography or by building a model of the protein from homologous fragments found in known structures. An example of the use of molecular modeling is a study of the active site in liver cytochrome P450 (CYP) enzymes and the derivation of complementary substrate binding templates.^[10]

Whole Organism Models

Whether data for test compounds are determined from *in vitro* experiments or estimated using computational methods such as QSARs, this information can be used to construct predictive models of organisms PK and PD. These prediction methods range in degrees of complexity from simple PK and PD models to more structured PBPK and PD–PK models.

Pharmacokinetics is the study of drug absorption, distribution, metabolism, and elimination in the body, also known as ADME.^[6] In their simplest forms, PK models fit the concentration profile over time of a

chemical in the blood to one or more exponential functions. These models allow the prediction of the maximal systemic exposure (termed C_{max}), as well as the total accumulated exposure (termed Area Under the Curve, AUC) for a given dose. As it is the amount of drug that tissues “see,” and not the amount swallowed or inhaled that produces an effect, these PK models produce a clear picture of the amount of a chemical that interacts with a tissue. However, the parameters in this basic PK model, the elimination constant and the volume of distribution, are purely mathematical in nature and have no relation to any physiological dimensions. The PK model is an adequate method for predicting plasma concentration levels in very similar animals (i.e., male rats of similar sizes), but is not useful for extrapolating to other species or routes of entry. Additionally, it does not contain explanations for the interindividual variations that cause drastic differences in drug or toxin concentration in the blood.

Pharmacodynamics is the study of the time- and concentration-dependent response of the body to a drug or toxin.^[11] Pharmacodynamics models fit a response-over-time vs. concentration profile to some function, typically using a sigmoidal or hill equation.^[12] As PD models require knowledge of the local compound concentration at the receptor or enzyme site, they greatly benefit from the incorporation of PK models. Although often treated separately, the PK and PD models are functionally linked. Examples of this include the induction of a metabolic enzyme upon exposure to a chemical, which can change the rate of metabolism and hence PKs, or the interactions of drugs that modify protein binding, influencing PDs.^[13] This interaction between PKs and PDs often leads to the use of a coupled PD–PK mathematical model.

Physiologically based pharmacokinetic models are PK models that break the body down into its separate tissues and corresponding blood flows (Fig. 1). Each tissue is treated as a separate entity with parameters describing transport, binding, and reactions. Such an approach requires a wealth of time-dependent data and is a primary reason that these models are not more prevalent. Physiologically based pharmacokinetic models can be used for extrapolation of experimental observations over animals of different body sizes, as well as for PK predictions in different species. Additionally, they allow extrapolation of results from one route of entry to another.^[13] One example of this would be the use of results from a bolus dose study in rats to predict inhalation PKs. These predictions, however, are only valid when the various reactions and distribution terms are accurately known; they cannot predict responses to unanticipated reactions. But if the interindividual differences are understood, for example, the knowledge that one

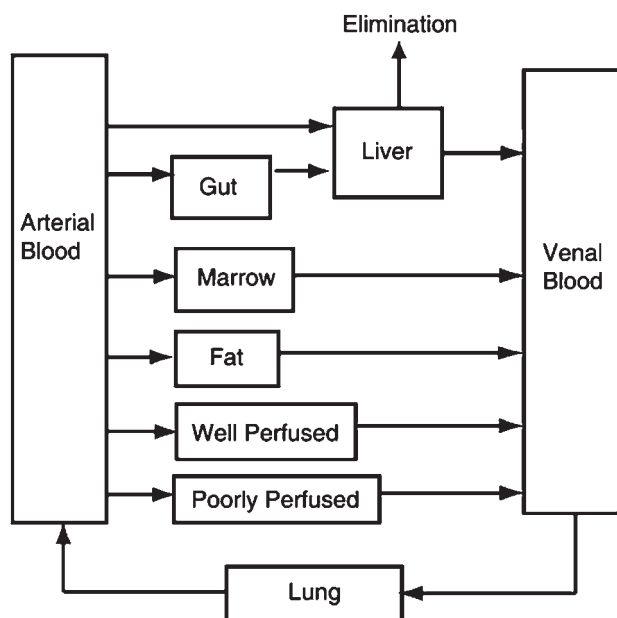


Fig. 1 A schematic representation of a PBPK. Tissues of interest, such as the gut and liver, are explicitly included. Other tissues known to have only minimal interaction can be combined into either well-perfused tissues (e.g., kidney) or slowly perfused tissues (e.g., muscle) to help maintain accurate descriptions of chemical dilution in the body.

individual has severe reduction in a liver CYP isoform activity, then the resultant differences in PKs can be easily described.

SINGLE CELL TYPE IN VITRO MODELS

The use of in vitro cultures for toxicity testing is considered one of the first steps in experimental surrogate trials. There is an enormous literature based on the use of different cells and endpoints for studies, but we will only briefly discuss them here. Two of the formats most frequently utilized for these studies are 96-well plates^[14] and flow cytometry. The 96-well plate assays use cells either attached to the bottom surface of a small well or suspended in culture medium. The cells are challenged with a chemical or treatment and then

assayed using a biomarker that produces either a luminescent (colored) or fluorescent product. This product is measured with a plate reader, and results are typically correlated with a calibration curve. Flow cytometry techniques first require exposure of the cells to a drug or toxicant in culture, and then labeling the cells with fluorescent markers for the respective attributes (i.e., viability, specific protein expression, etc.). These labeled cells are forced single-file past a laser and detector that collects light from the fluorescent tags. Results are expressed as the number of cells that were labeled with a particular condition, for example, number of live cells vs. number of dead and apoptotic cells.^[15] Table 1 lists several different cell conditions with references that have been studied utilizing various in vitro methods.

BARRIER MODELS

Epithelial cells cover most internal and external surfaces of the body and are, therefore, the first tissues that interact with foreign chemicals. Understanding how chemicals interact with and are absorbed through epithelial barriers is crucial for toxicological evaluation. Four epithelial barriers that are critical in toxicology studies are the epidermis, lung epithelium, intestinal epithelium, and blood–brain barrier. These barriers represent entry into the body via chemical contact, inhalation, ingestion, and systemic chemical contact with the brain. The models of epithelial barriers described in this section involve single cell types cultured on porous membrane supports that allow access to the apical (top) chamber and basolateral (bottom) chamber.

Skin penetration is a complex event. The stratum corneum, which is composed of the dead keratinocytes that make up the skin's outermost epidermal layer, is the main barrier to chemicals that come in contact with skin; but partitioning into the stratum corneum, diffusion within the stratum corneum, partitioning into the viable and more hydrophilic deeper epidermal layers, diffusion through the deeper epidermal layers, and metabolism of compounds in the epidermis must

Table 1 Common in vitro toxicity and metabolic assays

Toxic/functional endpoint	Study design	References
Cytotoxicity	Example: propidium iodide/trypan blue exclusion assay	[16,17]
Apoptosis	Mitochondrial H ⁺ potential (with Rhodamine 123)	[18]
Reactive oxygen species	Assays for free radical formation in cells	[19]
Genotoxicity	Micronuclei formation, single cell gel electrophoresis	[20]
Endocrine disruptors	Receptor binding assays	[21]
Metabolism	CYP metabolism	[22,23]

all be taken into consideration in epidermal toxicology testing.^[24] Poumay et al. have reconstructed a fully-differentiated epidermis by culturing normal human-derived epidermal keratinocytes on the air-liquid interface (culture medium in the basolateral chamber, air in the apical chamber) of porous membranes.^[25] The cultured epidermis exhibits basal, spinous, granular, and cornified layers.^[24] The modeled tissue showed progressive cytotoxicity in response to benzalkonium chloride, an irritant, and dinitrochlorobenzene, a sensitizer. Differences in interleukin-1 α and interleukin-8 release allowed the irritant to be distinguished from the sensitizer. Permeation studies with 17- β -estradiol indicate that the model has barrier properties similar to those seen in vivo, and the metabolism of 17- β -estradiol to estrone by the in vitro epidermis shows that the model has metabolizing properties toward hormones.^[25]

The respiratory tract is usually divided into two regions: the central conducting airways (made up of the trachea, bronchi, and bronchioles), and the peripheral alveolar regions.^[26] The airway epithelium filters air before it reaches the alveoli and prevents particulate absorption by means of a mucociliary clearance mechanism, tight junctions (protein connections between epithelial cells that reduce passive diffusion between the cells), metabolic activity, and efflux protein pumps.^[27] Calu-3 immortalized human bronchial epithelial cells grown on the air-liquid interface of porous membranes express apical cilia, mucous production, tight junctions, and several metabolic and transport systems.^[27] Mathias et al. tested the permeability of a wide range of compounds and found that Calu-3 monolayers had permeability characteristics similar to those seen in primary rabbit tracheal epithelial cultures and drug absorption rates similar to those seen in vivo in rat lungs.^[28]

The alveolar epithelium provides an immense surface area for the exchange of CO₂ and O₂.^[26] The alveolar epithelium is composed of two major cell types, Type I and Type II. Type I alveolar cells cover about 95% of the alveolar epithelium and are the main site of gas exchange, but are not able to divide, making them difficult to grow in vitro.^[26] Type II cells are more numerous even though they cover much less surface area and they secrete alveolar fluid, which keeps the surface between the air and cells moist. The Type II cells have well characterized endocytic properties and are the progenitors for Type I alveolar cells. They also express cilia, tight junctions, and CYP enzymes that metabolize foreign materials.^[29] The A549 immortalized cell line, which is derived from a human pulmonary carcinoma, expresses the differentiated biochemical properties of Type II alveolar cells in vivo, making the cell line a good model for studying alveolar metabolism and macromolecule processing.^[29] The

A549 cells do not form tight junctions, however, which make them poor candidates for the study of alveolar permeability.^[29] Currently primary cultures of Type II cells, which form Type I-like monolayers in culture, are the only model available to study absorption through the alveolar epithelium.^[26]

The epithelial lining of the small intestine is composed of several different cell types that display very tight intracellular junctions and are covered with tightly packed projections called microvilli. These microvilli serve to increase the surface area of the intestine available for absorption and contain digestive enzymes. The transport of a drug or chemical across the intestinal epithelium occurs by one or more of four different routes: passive transcellular (through the cell), passive paracellular (between cells), active (energy dependent) carrier-mediated, and transcytosis (transport across the epithelium with uptake into coated vesicles).^[30] When cultured on porous membranes, Caco-2 cells (a human colorectal adenocarcinoma cell line) differentiate into a polarized, highly functional epithelial barrier that is morphologically and biochemically very similar to the small intestinal epithelium.^[30] After a growth period of three weeks, the Caco-2 monolayers express tightly packed microvilli, tight junctions, and are capable of paracellular, transcellular, active, and transcytotic transport. Lennernäs et al.^[31] compared drug transport rates in Caco-2 monolayers with those obtained in the human jejunum (a section of the small intestine) in vivo. The researchers found that the rapidly absorbed, passively transported drugs naproxen, antipyrine, and metoprolol had comparable permeability coefficients in Caco-2 cells and in the human jejunum. The permeability coefficients of the slowly absorbed, passively transported drugs terbutaline and atenolol were 79- and 27-fold lower, respectively, in Caco-2 monolayers when compared with in vivo human jejunum results. The carrier-mediated transport rates of L-dopa, L-leucine, and D-glucose were also found to be lower in Caco-2 monolayers. The researchers concluded that Caco-2 monolayers can be used to successfully predict transport rates in humans for passively diffused drugs, but predictions of transport by carrier-mediated systems may require a scaling factor owing to the low expression of carrier proteins in the cell line.^[31]

The blood-brain barrier (BBB) is formed by the endothelial cells that line the brain capillaries. The function of the BBB is to keep blood and most compounds in blood separated from the brain and from cerebrospinal fluid.^[32] The brain capillary cells exhibit tight junctions that restrict paracellular transport, have a low rate of endocytosis, are highly metabolically active, and contain many enzymes and adenosine triphosphate (ATP)-dependent transporters that regulate the influx and efflux of compounds into the central

nervous system.^[32] Primary bovine brain microvessel endothelial cells (BMECs) grown on a porous membrane retain the characteristics of brain endothelial cells *in vivo* including morphology, BBB enzyme markers, and tight junctions.^[33] Glynn and Yazdanian used this model to study the permeability of anti-retroviral agents through the BBB. The researchers found that nevirapine was the most permeable anti-retroviral agent studied in the BBB model, making this drug a therapeutic candidate for inhibiting HIV viral replication in the brain.^[33]

COCULTURE MODELS

Cocultures of two different cell types are aimed at reproducing the cellular interactions and communication found in a tissue or organ. One example of two cell types that are often cultured together is brain capillary endothelial cells and astrocytes, a type of microglial cell. A major challenge for researchers working to create *in vitro* BBB models is culturing endothelial cells that display the extremely tight junctions seen *in vivo*. Astrocyte foot processes have close interaction with brain capillary endothelial cells in the BBB, and research has shown that *in vitro* cocultures of these cell types enhance the barrier properties of the BBB model.^[34] Ma et al. fabricated a silicon nitride membrane that is an order of magnitude thinner and at least two times as porous as commercially available membranes for coculturing primary bovine brain microvascular endothelial cells and astrocytes on opposite sides; their work was based on the hypothesis that an ultrathin, highly porous membrane would permit a high degree of direct contact between the cells without the astrocytes growing through the endothelial barrier.^[34] Fig. 2 shows astrocytes seeded on the backside of this fabricated membrane with foot processes growing through the pores. The researchers found that the cocultures could be maintained for up to two weeks, but they did not see a synergistic effect of the cocultures in terms of improved barrier properties. The researchers believe, however, that this lack of synergy is most likely because of the commercially available primary bovine brain microvascular endothelial cells losing their ability to respond to astrocyte differentiating signals; the researchers concluded that a different endothelial cell line or different endothelial primary cells may have more physiologically realistic results.^[34]

THREE-DIMENSIONAL MODELS

Three-dimensional (3-D) *in vitro* systems more closely mimic the cell shape and cellular environment found in

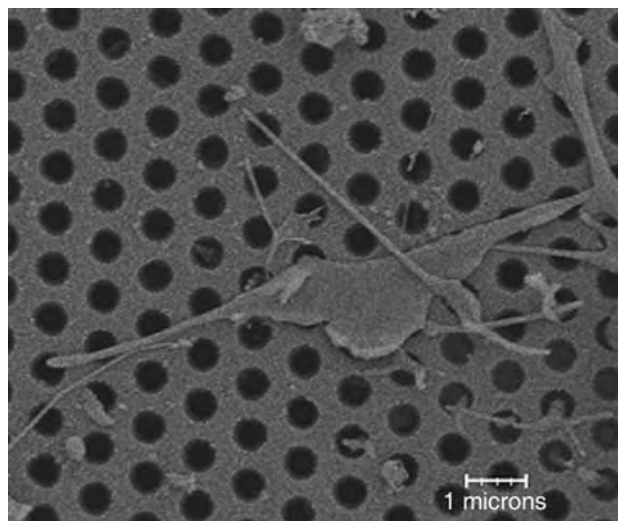


Fig. 2 An SEM image of a fabricated BBB membrane, containing 420 nm pores. Rat astrocytes were seeded on the opposite side, and foot processes are shown here growing through several pores.

in vivo.^[35] Cell shape and cellular environment can determine the gene expression and biological behavior of the cells, both of which have important implications in toxicity testing.^[35] Researchers have shown that 3-D cultures, unlike monolayer or suspension cultures, more accurately reproduce specific biochemical and morphological characteristics of the corresponding *in vivo* tissues and may stay in a functionally active, differentiated state for several weeks.^[36] The most commonly used method of 3-D culture is multicellular spheroids, which are aggregates of cells that can reach 1–2 mm in diameter.^[37] Multicellular spheroids form spontaneously in cultures with cells that cannot attach to another substrate. Normally spheroids are cultured in spinner flasks, roller bottles, or roller tubes, all of which limit cell attachment to the substrate by agitating the culture medium and allow rapid harvesting of many spheroids of approximately the same size.^[37] Stationary, nonadherent surfaces like agar or agarose in 96-well culture plates can also be used to cultivate single spheroids.^[37] The spheroids form both an extracellular matrix and a complex network of cell–cell and cell–matrix interactions. The network establishes similar morphological, functional, and mass transport features and differentiation patterns of the corresponding *in vivo* tissue for up to several weeks in culture.^[36] Castaneda and Kinne compared the effects of low concentrations of ethanol on primary rat hepatocyte spheroids and human hepatocellular carcinoma cell line (HepG2) spheroids.^[38] The primary cell spheroids represented normal liver tissue and the HepG2 spheroids represented a hepatocellular carcinoma.^[38] The researchers found that millimolar

concentrations of ethanol increased the inhibition of cell growth and apoptosis in HepG2 spheroids when compared with rat hepatocyte spheroids. This result suggests a possibility that millimolar concentrations of ethanol could be used as a treatment for hepatocellular carcinoma, because the treatment affects mainly tumor cells rather than the surrounding normal tissue.^[38]

Artificial tissues can be assembled *in vitro* by embedding primary or immortalized cells within 3-D scaffolds. Scaffolds are porous, degradable structures engineered for biocompatibility, selective permeability, and mechanical and chemical stability. They are fabricated either from natural materials such as collagen, alginate, and fibrin or synthetic polymers such as polyglycolide, polylactide, and polylactide-co-glycolide.^[39] These artificial organs could greatly accelerate the development of new drugs by creating physiologically accurate *in vitro* models for toxicity testing.^[39] Many different cell types including liver, vascular, bone, lung, intestine, kidney, and cells from mammary, salivary, and thyroid glands have been grown in 3-D matrices and shown positive trends toward regaining *in vivo* tissue function. However, few tests have yet been performed to study drug toxicity in these 3-D models.^[40] Three-dimensional, matrix-embedded cultures are also being integrated into the microcell culture analog (μ CCA), described later in this article under the heading "Microscale Designs."

Bioreactors can be combined with cells grown on 3-D scaffolds to create very high-density cultures that retain some of the metabolizing characteristics of the original tissue.^[41] Beads are often used as scaffolds to create more surface area for cell growth. Dextran-based microcarrier beads allow cells to grow on the bead surface in single or multilayers, and polylysine or alginate beads can be used to microencapsulate cells.^[41] The microcarrier or gel beads can be placed in perfusion reactors where the cells are retained. The culture medium is continuously or semicontinuously added, and spent medium together with its toxic metabolites is removed.^[41]

Choi et al.^[42] used Caco-2 cells seeded on permeable membranes together with HepG2 cells seeded onto cellulose-based microcarrier beads to create a perfusion system that mimics first pass metabolism. The results of this study showed enhanced CYP1A1/2 activity in the coculture perfusion system when compared with pure Caco-2 and HepG2 cultures. CYP1A1/2 is an enzyme responsible for breaking down toxins in the body. The researchers concluded that their *in vitro* system better modeled the *in vivo* metabolic activity of the liver and intestinal epithelium than single cell type cultures.

Hollow-fiber bioreactors are cylinders filled with bundles of hollow fibers meant to simulate the *in vivo* capillary supply.^[43] Cells are immobilized on the

external surface of the hollow fibers, which are composed of semipermeable ultrafiltration membranes. Culture medium flows through the inside of the tubes.^[41] The membrane pore size is small enough to retain cells while allowing gas, nutrients, and metabolic waste products to diffuse freely.^[43] Casciari et al.^[44] used a hollow fiber reactor to create a solid tumor model for testing tumor cell drug metabolism and resistance. The researchers successfully grew six different human tumor cell lines to confluence on 450 μ m internal diameter, polyvinylidene fluoride fibers, and then assessed doxorubicin sensitivity. The results showed that tumor cells in hollow fibers were more resistant to a four-hour doxorubicin exposure than were tumor cells in monolayer cultures.

MULTIPLE CELL DEVICES

Multiple cell or tissue types are used to bridge the gap between single cell or single tissue culture models and whole animal models for toxicity testing. The primary purpose of multiple cell types in the same culture is to investigate interactions between different cell populations, either through direct cell-cell interactions or through soluble proteins or metabolites. Of the approaches taken that combine multiple cell types to explore metabolite exchange, three will be described here.

The simplest method to study soluble protein and metabolite interactions is the use of media that is first incubated in one cell type, and then added to the second cell type. Another approach, termed the integrated discrete multiple organ cell culture (IdMOC) system, combines several cell types into the same culture container, but maintains separate cell populations within that container. The third approach, called a CCA, includes several cell cultures contained in independent culture containers with culture medium recirculating between them. These systems attempt to provide an *in vitro* experimental space that can replicate some of the cell-cell interactions not easily studied *in vivo*.

Conditioned Medium Systems

The goal of this experimental method is to mimic the metabolite exchange that occurs within the circulatory system. Medium conditioning, which is the practice of adding and incubating media in a cell culture then transferring this conditioned media into another cell culture, is a convenient method of exploring metabolic impacts in toxicity testing. An example of this method is the use of liver microsomal fractions to metabolize cyclophosphamide (CP) prior to addition to the target tissue, in this case human fibroblasts.^[45] It was

previously shown that CP is activated in the liver and then acts as a chemotherapeutic agent. Investigators utilized such a system to compare the cytotoxicity of CP when incubated directly with fibroblasts to CP when first conditioned with microsomes. The researchers found that with a dose of 0.3 mg/mL, direct exposure to the fibroblasts required several days to reach the same level of toxicity as was reached after only a few hours of exposure when the CP was pre-conditioned with microsomes. This indicates that the combination of this simple microsome-fibroblast system can begin to replicate the activation of compounds *in vivo*. However, this system is only exploratory in nature and does not include any detail of physiological quantities, such as the actual amount of enzyme or cells present.

Shared Medium Systems

Another method that enables the exchange of metabolites between different cell types is the sharing of culture media by several different cell populations. An example of this approach is the IdMOC system.^[46] This system is built by fabricating small wells, 8 mm diameter and 0.8 mm in height, inside the well of a six-well plate. Within each of the smaller wells, a different cell type is cultured. Medium inside the larger well feeds the different cell cultures. Experiments were conducted utilizing the MCF-7 tumor cell line and six other primary cell types. In response to exposure to tamoxifen, the rank order of cell toxicity as well as a therapeutic index of each primary cell type compared to the cancer cell line was generated. The system has not yet been utilized to directly investigate metabolic exchange, although such a use has been proposed.^[46]

CCA

A CCA combines several cell culture bioreactors representative of different tissues to “mimic” the variety of organs in an animal. This pseudoorgan culture system replicates the circulatory system of animals with a flow of recirculating culture medium. All bioreactor dimensions and the corresponding culture medium flow rates are scaled down from the actual organ sizes and blood flows *in vivo*. The resulting system is an experimental replicate of the corresponding, PBPK, model of the animal. The ability to experimentally study dosing conditions of compounds and the inclusion of metabolite exchange between tissues is an advantage over standard *in vitro* culture techniques. Unlike other *in vitro* systems, a CCA can mimic the dynamics of dose exposure. Because any cell type may be used, these CCA devices can be designed to mimic key target organs from mice, rats, dogs, or even

humans, just by selecting representative cell lines and designing corresponding culture vessels.

Sweeney et al.^[47] developed the first CCA device to study the toxicology of naphthalene. The system was constructed using milk dilution bottles as culture vessels. The bottles were connected with tubing and a peristaltic pump provided recirculating flow. Their device sought to mimic the exposure of naphthalene to a rat. The researchers hypothesized that naphthalene, a chemical that is only moderately toxic in its native form, becomes activated by metabolism in the liver and then can induce damage in lung tissue. To test this hypothesis, the researchers constructed a CCA device that contained the rat hepatoma cell line H4IIE, and a rat lung cell line L2. The remainder of the rat's organs were modeled as a nonmetabolizing body of distribution and included as a culture chamber devoid of cells.

Experiments with this early CCA system were able to show that naphthalene induces toxicity in lung cells after conversion by liver cells. In experiments with no liver cells present, no toxicity was observed in the lung cells. Additionally, the researchers showed that by either increasing liver cell number or increasing CYP1A1 enzyme levels, toxicity to the lung tissue increased—further supporting their original hypothesis. However, parameters in this early system such as the ratio of liquid volume to cell volume and the residence time in each cell container were far from realistic.

A revised system was constructed by Ghanem and Shuler^[48,49] to address several of the concerns from the prototype. The modified system utilized a packed-bed design, with cells cultured on microcarrier beads. This alteration allowed for time-course sampling of cells during experimentation by removal of a portion of the beads with a sterile syringe. The ratio of cells to liquid volume was increased, though the ratio was still much less than the ratio *in vivo*. Experiments with this revised CCA system did not show any toxic response to naphthalene, however, and the decreased residence time in the liver compartment was found to be the cause. The reduced residence time resulted in less conversion of naphthalene to toxic quinines and minimal toxic response.

Microscale Designs

Recently, the CCA concept has been revised by the design of a μ CCA, which is shown in Fig. 3. By creating these systems using microfabrication, the devices now can achieve near *in vivo* residence times and volume ratios. The length scale of capillaries in the body is on the order of 10 μ m. This scale can be easily reached utilizing standard photolithographic

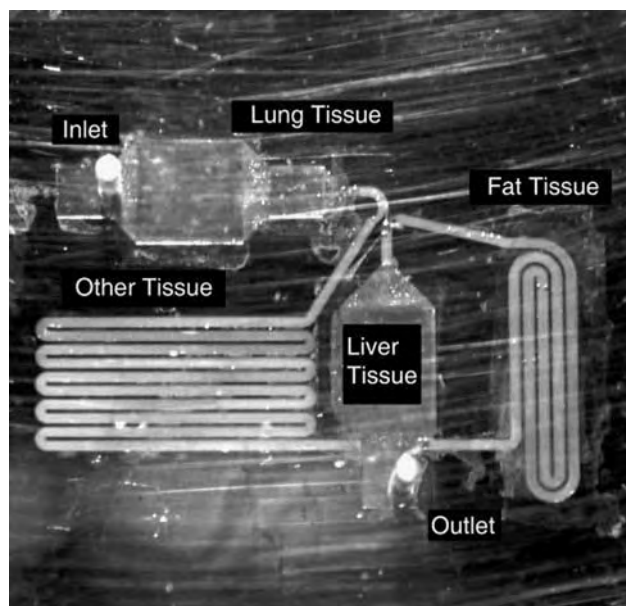


Fig. 3 An example of a microscale cell culture analog. This silicon device includes a liver tissue chamber, a lung tissue chamber, a fat tissue chamber, and an “other tissues” chamber as shown.

technology borrowed from the semiconductor industry. The first system was built to incorporate a liver cell chamber, a lung cell chamber, and an “other tissues” chamber fabricated on a silicon chip.^[50] This device utilized the same H4IIE and L2 cell lines of the previous CCA systems, but the chip was 2.5 cm × 2.5 cm in dimension and enclosed by a Plexiglas casing. Cells were attached to the culture chambers with a combination of poly-D-lysine (Sigma-Aldrich, USA) and either Matrigel (BD Biosciences, USA) for H4IIE cells or fibronectin (Chemicon International, USA) for L2 cells. Additionally, an integrated oxygen sensor was incorporated into the chip. This sensor utilized a ruthenium dye immobilized in a resin as a UV sensitive indicator of dissolved oxygen content.

Another μ CCA was designed that was similar to the first design, but divided the “other tissues” into well-perfused and slowly perfused chambers.^[51] The liver cell line H4IIE was also replaced with a human hepatoma line HepG2/C3A after finding increased CYP expression levels in this new cell model. Naphthalene was used as a study toxicant for this system, and was shown to cause a significant drop in glutathione (GSH) levels in both the liver and the lung cells after 6 hr of exposure. When no liver cells were present, there was no depletion of GSH in the lung cells, indicating that biotransformation in the liver cells produces a recycling product that injures the lung cells. Further experiments in the μ CCA with 1,2-naphthoquinone identified the reactive metabolites to be 1,2-naphthalenediol and 1,2-naphthoquinone. This discovery showed the

utility of the μ CCA system in investigating toxicity mechanisms.

To further clarify naphthalene toxicity, Viravaidya and Shuler investigated the impact of bioaccumulation, or the storage of compounds inside of fat tissue, on the response of the μ CCA system to naphthalene compounds.^[52] The rat preadipocyte cell line 3T3-L1 was used for this study because of its ability to be attached to the silicon surface. Experiments with naphthalene and 1,2-naphthoquinone showed that sequestering of naphtha compounds by the fat cells reduced the loss of GSH in both liver and lung cells when compared to a system with no fat cells. This illustrates the utility of this CCA system to explore the impact of dosing dynamics on toxicity.

CONCLUSIONS

Completing ADMET studies on new pharmaceuticals early enough to direct resources to the best drug candidates and studying the effects of chemical exposure before human illness develops has both economic and ethical value. To this end, basic toxicity predictions with single cell types and mathematical models have been relatively successful in supplementing animal data. In some cases, permeability through and metabolism by epithelial barriers can also be successfully reproduced with in vitro cell cultures. The in vitro representation of more vascularized tissues and organs such as the liver, however, is still challenging to tissue engineers and these tissue-engineering models are often incomplete. As advances in tissue engineering are made on in vitro tissues and organs, pharmaceutical and chemical ADMET studies will continue to improve. In vitro models that interconnect different tissues, particularly when used in conjunction with the corresponding PBPK, offer the possibility of making more realistic models that can act as surrogates for animals and humans in toxicity testing. Developments in μ CCAs, because they combine PBPK models with advances in microfabrication and cell culture, make microscale ADMET studies of pharmaceuticals and chemicals a promising frontier.^[53]

ACKNOWLEDGMENTS

The authors would like to acknowledge support from the New York State Office of Science, Technology and Academic Research (NYSTAR) program, the National Science Foundation (BES 0342985), and the Nanobiotechnology Center (NBTC), a STC Program of the National Science Foundation under Agreement No. ECS-9876771.

ARTICLES OF FURTHER INTEREST

Biocompatibility Testing; Biomaterials Immune Response, Biomedical Ethics; Cell Culture Assays; Cell-Material Interaction; Foreign Body Response; Host Reactions; Infection of Medical Devices; Particulate Matter and Host Reactions

REFERENCES

- Agres, T. Finding blockbusters a struggle. *Drug Discov. Devel.* **2005**, *8* (3), 16–19.
- DiMasi, J.A.; Hansen, R.W.; Grabowski, H.G. The price of innovation: new estimates of drug development costs. *J. Health Econ.* **2003**, *22* (2), 151–185.
- Davila, J.C.; Rodriguez, R.J.; Melchert, R.B.; Acosta D., Jr. Predictive value of in vitro model systems in toxicology. *Annu. Rev. Pharmacol. Toxicol.* **1998**, *38*, 63–96.
- Boston Consulting Group. In *How Genomics and Genetics are Transforming the Biopharmaceutical Industry*; Boston Consulting Group: Boston, MA, 2001.
- Rawlins, M.D. Cutting the cost of drug development? *Nat. Rev. Drug Discov.* **2004**, *3* (4), 360–364.
- van de Waterbeemd, H.; Gifford, E. ADMET in silico modelling: towards prediction paradise? *Nat. Rev. Drug. Discov.* **2003**, *2* (3), 192–204.
- Barratt, M.D. Prediction of toxicity from chemical structure. *Cell Biol. Toxicol.* **2000**, *16* (1), 1–13.
- Jacobs, M.N. In silico tools to aid risk assessment of endocrine disrupting chemicals. *Toxicology* **2004**, *205* (1–2), 43–53.
- Dearden, J.C. In silico prediction of drug toxicity. *J. Comput. Aided Mol. Des.* **2003**, *17* (2–4), 119–127.
- Ekins, S.; Wrighton, S.A. Application of in silico approaches to predicting drug–drug interactions. *J. Pharmacol. Toxicol. Methods* **2001**, *45* (1), 65–69.
- Kuh, H.J.; Jang, S.H.; Wientjes, M.G.; Au, J.L. Computational model of intracellular pharmacokinetics of paclitaxel. *J. Pharmacol. Exp. Ther.* **2000**, *293* (3), 761–770.
- Jonker, D.M.; Visser, S.A.; van der Graaf, P.H.; Voskuyl, R.A.; Danhof, M. Towards a mechanism-based analysis of pharmacodynamic drug–drug interactions in vivo. *Pharmacol. Ther.* **2005**, *106* (1), 1–18.
- Clark, L.H.; Setzer, R.W.; Barton, H.A. Framework for evaluation of physiologically-based pharmacokinetic models for use in safety or risk assessment. *Risk Anal.* **2004**, *24* (6), 1697–1717.
- Verkman, A.S. Drug discovery in academia. *Am. J. Physiol. Cell Physiol.* **2004**, *286* (3), C465–C474.
- Strebel, A.; Harr, T.; Bachmann, F.; Wernli, M.; Erb, P. Green fluorescent protein as a novel tool to measure apoptosis and necrosis. *Cytometry* **2001**, *43* (2), 126–133.
- Eisenbrand, G.; Pool-Zobel, B.; Baker, V.; Balls, M.; Blaauboer, B.J.; Boobis, A.; Carere, A.; Kevekordes, S.; Lhuguenot, J.C.; Pieters, R.; Kleiner, J. Methods of in vitro toxicology. *Food Chem. Toxicol.* **2002**, *40* (2–3), 193–236.
- Xu, J.J.; Diaz, D.; O'Brien, P.J. Applications of cytotoxicity assays and pre-lethal mechanistic assays for assessment of human hepatotoxicity potential. *Chem. Biol. Interact.* **2004**, *150* (1), 115–128.
- Bedner, E.; Li, X.; Gorczyca, W.; Melamed, M.R.; Darzynkiewicz, Z. Analysis of apoptosis by laser scanning cytometry. *Cytometry* **1999**, *35* (3), 181–195.
- Brandes, R.P.; Janiszewski, M. Direct detection of reactive oxygen species ex vivo. *Kidney Int.* **2005**, *67* (5), 1662–1664.
- Knasmuller, S.; Mersch-Sundermann, V.; Kevekordes, S.; Darroudi, F.; Huber, W.W.; Hoelzl, C.; Bichler, J.; Majer, B.J. Use of human-derived liver cell lines for the detection of environmental and dietary genotoxicants; current state of knowledge. *Toxicology* **2004**, *198* (1–3), 315–328.
- Combes, R.D. The case for taking account of metabolism when testing for potential endocrine disruptors in vitro. *Altern. Lab. Anim.* **2004**, *32* (2), 121–135.
- Zucco, F.; De Angelis, I.; Testai, E.; Stamatii, A. Toxicology investigations with cell culture systems: 20 years after. *Toxicol. In Vitro* **2004**, *18* (2), 153–163.
- Gomez-Lechon, M.J.; Donato, T.; Ponsoda, X.; Castell, J.V. Human hepatic cell cultures: in vitro and in vivo drug metabolism. *Altern. Lab. Anim.* **2003**, *31* (3), 257–265.
- Ponec, M. Skin constructs for replacement of skin tissues for in vitro testing. *Adv. Drug Deliv. Rev.* **2002**, *54* (Suppl. 1), S19–S30.
- Poumay, Y.; Dupont, F.; Marcoux, S.; Leclercq-Smekens, M.; Herin, M.; Coquette, A. A simple reconstructed human epidermis: preparation of the culture model and utilization in in vitro studies. *Arch. Dermatol. Res.* **2004**, *296* (5), 203–211.
- Forbes, I.I. Human airway epithelial cell lines for in vitro drug transport and metabolism studies. *Pharm. Sci. Technol. Today* **2000**, *3* (1), 18–27.
- Florea, B.I.; Cassara, M.L.; Junginger, H.E.; Borchard, G. Drug transport and metabolism characteristics of the human airway epithelial cell line Calu-3. *J. Control Release* **2003**, *87* (1–3), 131–138.
- Mathias, N.R.; Timoszyk, J.; Stetsko, P.I.; McGill, J.R.; Smith, R.L.; Wall, D.A. Permeability characteristics of Calu-3 human bronchial epithelial cells: in vitro-in vivo correlation to predict lung absorption in rats. *J. Drug Target* **2002**, *10* (1), 31–40.
- Foster, K.A.; Oster, C.G.; Mayer, M.M.; Avery, M.L.; Audus, K.L. Characterization of the A549 cell line as a type II pulmonary epithelial cell model for drug metabolism. *Exp. Cell Res.* **1998**, *243*, 359–366.
- Artursson, P.; Palm, K.; Luthman, K. Caco-2 monolayers in experimental and theoretical predictions of drug transport. *Adv. Drug Deliv. Rev.* **2001**, *46* (1–3), 27–43.
- Lennernäs, H.; Palm, K.; Fagerholm, U.; Artursson, P. Comparison between active and passive drug transport in human intestinal epithelial (Caco-2) cells in vitro and human jejunum in vivo. *Int. J. Pharm.* **1996**, *127*, 103–107.
- Lundquist, S.; Renftel, M. The use of in vitro cell culture models for mechanistic studies and as permeability

- screens for the blood-brain barrier in the pharmaceutical industry—background and current status in the drug discovery process. *Vascul. Pharmacol.* **2002**, *38*, 355–364.
33. Glynn, S.L.; Yazdanian, M. In vitro blood-brain barrier permeability of nevirapine compared to other antiretroviral agents. *J. Pharm. Sci.* **1998**, *87* (3), 306–310.
 34. Ma, S.H.; Lepak, L.A.; Hussain, R.J.; Shain, W.; Shuler, M.L. An endothelial and astrocyte co-culture model of the blood-brain barrier utilizing an ultra-thin, nanofabricated silicon nitride membrane. *Lab On a Chip* **2005**, *5* (1), 74–85.
 35. Mueller-Klieser, W. Three-dimensional cell cultures: from molecular mechanisms to clinical applications. *Am. J. Physiol. Cell Physiol.* **1997**, *273*, C1109–C1123.
 36. Kunz-Schughart, L.A.; Freyer, J.P.; Hofstaedter, F.; Ebner, R. The use of 3-D cultures for high-throughput screening: the multicellular spheroid model. *J. Biomol. Screen.* **2004**, *9* (4), 273–285.
 37. Kunz-Schughart, L.A.; Kreutz, M.; Knuechel, R. Multicellular spheroids: a three dimensional in vitro culture system to study tumor biology. *Int. J. Exp. Pathol.* **1998**, *79*, 1–23.
 38. Castaneda, F.; Kinne, R.K.H. Cytotoxicity of millimolar concentrations of ethanol on HepG2 human tumor cell line compared to normal rat hepatocytes in vitro. *J. Cancer Res. Clin. Oncol.* **2000**, *126*, 503–510.
 39. Griffith, L.G.; Naughton, G. Tissue engineering—current challenges and expanding opportunities. *Science* **2002**, *295*, 1009–1016.
 40. Schmeichel, K.L.; Bissell, M.J. Modeling tissue-specific signaling and organ function in three dimensions. *J. Cell Sci.* **2003**, *116* (12), 2377–2388.
 41. Shuler, M.L.; Kargi, F. *Bioprocess Engineering—Basic Concepts*, 2nd Ed.; Prentice Hall PTR: Upper Saddle River, New Jersey, 2002.
 42. Choi, S.H.; Fukuda, O.; Sakoda, A.; Sakai, Y. Enhanced cytochrome P450 capacities of Caco-2 and HepG2 cells in new coculture system under the static and perfused conditions: evidence for possible organ-to-organ interactions against exogenous stimuli. *Mat. Sci. Eng. C* **2004**, *24*, 333–339.
 43. Kumar, S.; Wittmann, C.; Heinzel, E. Minibioreactors. *Biotechnol. Lett.* **2004**, *26*, 1–10.
 44. Casciari, J.J.; Hollingshead, M.G.; Alley, M.C.; Mayo, J.G.; Malspeis, L.; Miyachi, S.; Grever, M.R.; Weinstein, J.N. Growth and chemotherapeutic response of cells in a hollow-fiber in vitro solid tumor model. *J. Natl. Cancer Inst.* **1994**, *86* (24), 1846–1852.
 45. Cooper, J.T.; Goldstein, S. Toxicity testing in vitro. II. Use of a microsomal—cultured human fibroblast system to study the cytotoxicity of cyclophosphamide. *Can. J. Physiol. Pharmacol.* **1976**, *54* (4), 546–550.
 46. Li, A.P.; Bode, C.; Sakai, Y. A novel in vitro system, the integrated discrete multiple organ cell culture (IdMOC) system, for the evaluation of human drug toxicity: comparative cytotoxicity of tamoxifen towards normal human cells from five major organs and MCF-7 adenocarcinoma breast cancer cells. *Chem. Biol. Interact.* **2004**, *150* (1), 129–136.
 47. Sweeney, L.M.; Shuler, M.L.; Babish, J.G.; Ghanem, A. A cell culture analogue of rodent physiology: application to naphthalene toxicology. *Toxicol. In Vitro* **1995**, *9* (3), 307–316.
 48. Ghanem, A.; Shuler, M.L. Characterization of a perfusion reactor utilizing mammalian cells on microcarrier beads. *Biotechnol. Prog.* **2000**, *16* (3), 471–479.
 49. Ghanem, A.; Shuler, M.L. Combining cell culture analogue reactor designs and PBPK models to probe mechanisms of naphthalene toxicity. *Biotechnol. Prog.* **2000**, *16* (3), 334–345.
 50. Sin, A.; Chin, K.C.; Jamil, M.F.; Kostov, Y.; Rao, G.; Shuler, M.L. The design and fabrication of three-chamber microscale cell culture analog devices with integrated dissolved oxygen sensors. *Biotechnol. Prog.* **2004**, *20* (1), 338–345.
 51. Viravaidya, K.; Sin, A.; Shuler, M.L. Development of a microscale cell culture analog to probe naphthalene toxicity. *Biotechnol. Prog.* **2004**, *20* (1), 316–323.
 52. Viravaidya, K.; Shuler, M.L. Incorporation of 3T3-L1 cells to mimic bioaccumulation in a microscale cell culture analog device for toxicity studies. *Biotechnol. Prog.* **2004**, *20* (2), 590–597.
 53. Khamsi, R. Labs on a chip meet the stripped down rat. *Nature* **2005**, *435* (7038), 12–13.

Antimicrobial/Antibiotic (Infection Resistance) Materials

A

Franck Furno

BRIG, University Hospital, Nottingham, United Kingdom, and Research and Development, Stockhausen GmbH and Co. KG, Krefeld, Germany

Roger Bayston

BRIG, University Hospital, Nottingham, United Kingdom

INTRODUCTION

Biomaterials can be defined as substances other than food or drugs comprising therapeutic or diagnostic systems that are in contact with tissue or biological fluids. Examples include partially implanted medical devices (such as contact lenses, central venous catheters, urinary catheters, voice prostheses, etc.) and totally implanted devices (such as pacemakers, mechanical heart valves, prosthetic joints, etc.). To a greater or lesser degree, all of these are prone to infection. The infecting organisms, their sources, and the period in which the devices are at risk vary among devices and their sites of insertion. In order to reduce some of the confusion surrounding biomaterials-related infection (BRI), a categorization system has been proposed to conveniently separate devices according to their method of implantation (Table 1).^[1] Category I devices such as hydrocephalus shunts, vascular grafts, prosthetic hip and knee joints, and prosthetic heart valves are totally implanted and have a very short period of risk at and soon after the time of insertion, and require antimicrobial protection for only a few days after insertion to prevent the majority of infections. By contrast, Category II (central venous access devices), and Category III devices (urinary catheters and voice prostheses) are at risk for the duration of their use.

OVERVIEW

Polymers are the most important class of biomaterials by far. Polymers such as polyethylene, polysiloxanes (silicones), polyurethanes, and poly(vinyl chloride) (PVC) are used extensively for the fabrication of indwelling catheters; however, they provide surfaces to which contaminating microbes can adhere and cause infection. Table 2 describes the incidence of such infections,^[1] which can vary from 1% for hip prostheses to 100% for urinary tract catheters after 3 weeks' use.^[2] Mortality rates following an infection can be as high as 10% in central venous catheter infections^[3] and up to 50% in aortic vascular graft infections.^[4]

Biomaterials-related infections (BRIs) are therefore one of the leading causes of nosocomial infections and primary septicaemia, and result in major consumption of health care resources affecting all clinical disciplines. The continuing development of biomaterials and their increasing use in medicine necessitates greater understanding of the complex nature of BRI. Many of these infections originate from bacterial contamination during surgery by exogenous sources such as the patient's skin or the health care provider's hands, or from hematogenous spread (i.e., infection from another site that is transmitted via the blood stream). Biomaterials-related infections are most commonly caused by coagulase-negative staphylococci (CoNS), with *Staphylococcus epidermidis* the most frequent organisms involved.^[5] Other BRI organisms include *Staphylococcus aureus*, *Propionibacterium acnes*, streptococci, and coliforms.

To understand approaches to prevention of infection in surgical implants, it is important to appreciate the events involved in its establishment. Once inserted, the biomaterials soon become coated with components of the body fluids in which they are bathed. This coating is known as conditioning film. Depending on the body site, the surrounding fluid can be saliva, urine, tear fluid, tissue fluid, blood, etc., and the conditioning film will consist mostly of adsorbed proteins such as fibronectin and fibrinogen, which can act as receptors for bacterial attachment. Plasma proteins are adsorbed in a competitive fashion and can displace each other until stability is reached.^[6] However, the manner in which conditioning films influence microbial attachment remains unclear. Physicochemical surface properties of the biomaterial and bacteria cell surface play a major role in the adhesion process.^[7] It should also be borne in mind that some biomaterials (such as polyvinylpyrrolidone (PVP), used as a catheter coating) have been found to profoundly influence the local macrophage response to infection.^[8]

The course of infections involves three major steps: 1) microbial adhesion; 2) microbial proliferation; and 3) formation of a bacterial biofilm (Fig. 1). Adherence of bacteria to implants and biofilm formation in vivo was first described in 1972.^[9] Most organisms stick

Table 1 Categories of medical implant device and periods of risk of infection

Category and implant	Site	Period of risk ^d
<i>Category I</i> ^a		
Prosthetic hip/knee joint	Hip or knee	A, (B)
Vascular grafts	Thoracic aorta	A, (B)
	Abdominal aorta/groin	A, (B)
Hydrocephalus shunts	Brain to abdomen	A
Prosthetic heart valves	Heart	A, B
<i>Category II</i> ^b		
Central venous catheters (cuffed)	Central venous system	B
CAPD catheters	Abdominal wall/peritoneum	B
<i>Category III</i> ^c		
Urinary catheters	Urethra/bladder	B
Voice prostheses	Trachea/esophagus	B

^aCategory I are totally implanted long-term devices.

^bCategory II are partially implanted, short- to medium-term devices.

^cCategory III are not surgically implanted.

^dPeriods of risk are: A associated with contamination at surgical insertion; B associated with contamination as a consequence of presence or use of device.

avidly to not only a variety of human cells and tissues, but also to the wide range of polymers, glasses, plastics, and metals used to make medical devices. Following the initial attachment phase, the microorganisms adopt a highly characteristic metabolic state in which they begin to produce extracellular enzymes to form long-chain extracellular polymeric substances from glycoside residues, which they synthesise intracellularly. This mucoid substance is variously called extracellular slime matrix, slime, or exopolymer, and consists mainly of high-molecular mass polysaccharide, which appears to determine both the structure and cohesive strength of biofilms by promoting bacterial cell-to-cell associations and the development of thick layers of bacteria.^[10]

Biofilms are phenotypically diverse populations of bacteria, embedded in the extracellular slime matrix. Other components in the biofilm include multivalent

cations, inorganic particles, and colloidal and dissolved compounds. Biofilm development has been explored in increasing detail since the 1980s. Studies have resulted in various models that characterize the complexity of biofilm development as a process of adaptation and changing genetic regulation. The physiology and molecular nature of biofilm formation varies between organisms. Recently, essential biochemical and genetic mechanisms of biofilm formation in *Ss. epidermidis* and *S. aureus*—the main causes of BRI—have been elucidated.^[11] Biofilms play a vital role in both the persistence of infection and the failure of treatment regimens (the two key clinical features of BRI). They can be considered to be stable microenvironments in which bacteria are protected from host-defense mechanisms and most therapeutic agents.^[12,13] Bacteria that can produce defensive enzymes that protect them from antibiotics (e.g., beta-lactamases) can accumulate high

Table 2 Incidence of infection in implantable devices

Implant	Site	Infection rate
Prosthetic hip/knee joint	Hip or knee	1–3%
Vascular grafts	Thoracic aorta	1%
	Abdominal aorta/groin	3–6%
Hydrocephalus shunts	Brain to abdomen	3–20%
Prosthetic heart valves	Heart	1.5%
Central venous catheters (cuffed)	Central venous system	3–10%
		0.5%
CAPD catheters	Abdominal wall/peritoneum	1.3 per patient/yr.
Voice prostheses	Trachea/esophagus	100%

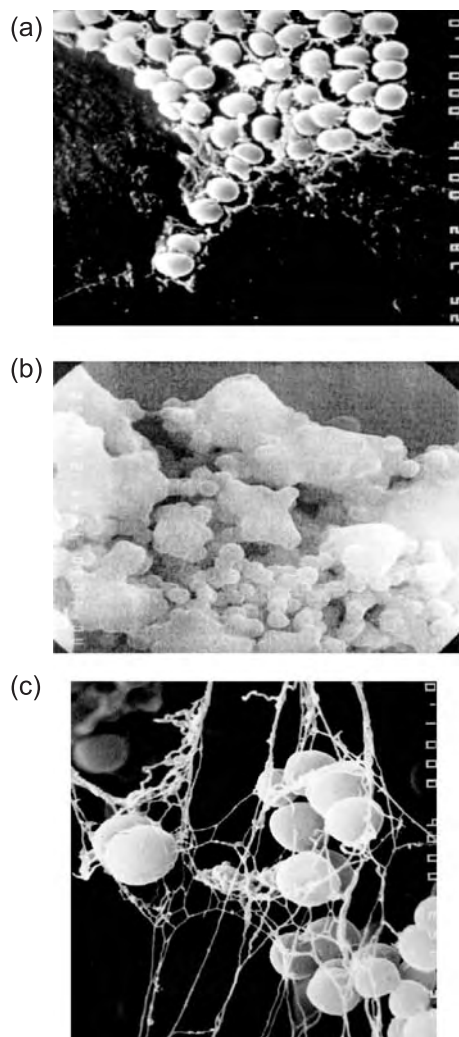


Fig. 1 (a) Plaque of proliferating bacteria on the surface of a removed catheter. Scanning electron microscopy (SEM) $\times 7800$. (b) Staphylococcal biofilm in a removed catheter, showing exopolymer production. SEM $\times 5000$. (c) Interior of staphylococcal biofilm removed catheter, showing intercellular adhesion. SEM $\times 15,000$.

concentrations of these within the microenvironment of the biofilm, whereas others can accumulate catalase or superoxide dismutase locally to protect against oxidizing ions released by phagocytes.^[14]

The concept of the biofilm phenotype explains the persistence of BRI and the need to remove the implant to eradicate infection. The clinical position is further complicated by the increasingly extensive innate resistance to antibiotics in staphylococci, restricting the choice of drugs for treatment. In addition, the bacteria in biofilms are physiologically, metabolically, and morphologically profoundly different from planktonic bacteria.^[15] Ziebuhr et al.^[16] reported that in *S. epidermidis*, such phenotypic changes were accompanied by subtle DNA rearrangements in vivo. It may take

several days to weeks before a thick mature biofilm is achieved, and sometimes detachment events occur.^[14] Active detachment is a physiologically regulated event, but only a few studies^[17] have been performed to demonstrate a biological basis for this process. Unfortunately, it is currently impossible to prevent access of bacteria to implants, even when the highest levels of hygiene are strictly adhered to. In view of the disastrous consequences of BRI, there is clearly a need to develop materials with antimicrobial activity that resist BRI for the lifetime of the device (or at least a significant part of it) and that do not provoke or select microbial resistance. Two main approaches for preventing BRI have been developed: 1) antiadhesive surfaces; and 2) materials coated or impregnated with releasable bioactive substances such as antiseptics, antibiotics, metal (or metal salt), etc.

SURFACE MODIFICATION TO PREVENT OR REDUCE BACTERIAL ADHERENCE

Extensive in-vitro investigation has focused on the search for materials and designs that diminish bacterial attachment and decrease the risk for bacterial infection. Several biomaterials approaches have been proposed, including highly biocompatible substances, biodegradable biomaterials, immobilization of biomolecules, incorporation of grafted hydrophilic outer layers to reduce protein adsorption, creation of low-surface energy (hydrophobicity) and carbon-rich materials, creation of microdomains to regulate cellular and protein adhesion, new copolymers with surface grafting to reduce friction^[18] and bacterial adhesion,^[19] as well as antimicrobial coatings and impregnation. Surface modifications may be engineered by grafting peptides or by using physical or chemical treatments to create charges (e.g., anionic groups, amino acids) on the polymer surfaces or to mimic sequences in natural molecules.^[20] Control of the charge and hydrophobic properties is likewise a pathway to influence biofilm interaction with a substratum surface. Microbial adhesion to biomaterials is mediated by specific interactions between cell surface structures and specific molecular groups on the substratum surface, or when viewed from an overall, physicochemical viewpoint by nonspecific interaction forces, including Lifshitz-van der Waals forces, electrostatic double-layer forces, solvent-dependent interactions, hydrogen bonding, hydrophobic interactions, hydration forces, acid-base interactions, and Brownian motion.^[20] Upon approach of a surface, organisms will be attracted or repelled by the surface, depending on the sum of the different nonspecific interaction forces. Lifshitz-van der Waals forces and Brownian motion usually promote adhesion, whereas electrostatic interactions can be either attractive or repulsive. Because

most organisms are negatively charged, a negatively charged substratum should theoretically exert a repulsive electrostatic force on the organisms, whereas adhesion should be promoted on positively charged surfaces. Bacterial adhesion is reported to be reduced, for example, on extremely hydrophobic surfaces.^[21,22] Gorman et al. have shown a greater adherence of *S. epidermidis* to the more hydrophilic polyurethane catheters than to silicone catheters.^[23] Price et al. were able to show that argon-plasma bombardment and silane treatment created a surface that reduced *Candida* adhesion irrespective of hydrophobicity.^[24] Titanium implants exposed in the oral cavity have been surface-modified by drying processes including ion implantation, ion plating, and oxidizing. The resulting calcium-implanted or alumina-coated surfaces have shown conclusive in vitro evidence of inhibition of adherence of oral bacteria.^[25-27] Such dry process surface modification is also useful in ensuring resistance to wear. Current ideas to prevent surface fouling by protein adsorption also involve tailoring the molecular interactions between the adsorbed proteins and the surface. It is now well established that protein adsorption is the first event following materials-tissue contact, and that the adsorption is generally nonspecific, with multiple binding sites on the proteins interacting with sites on the material surface. Protein adsorption can therefore be minimized by appropriate design of the interface, with modification of macroscopic surface properties such as wettability. Such a nonfouling surface may be obtained by coating with hydrated layers or by grafting with polyethylene oxide, phosphorylcholine (PC), heparin, and glucosamines.^[28-30] More recently, attention has been focused on a protein-based adhesion inhibitor, because proteins in the conditioning film can influence the proliferation of adhering microorganisms. For example, serum albumin reduces staphylococcal adhesion, whereas fibronectin increases it.^[31]

Although many techniques and processes can influence bacterial colonization in vitro, a physicochemical explanation of the mechanisms of inhibition by such materials is still not available. Currently no nontoxic material has been shown to completely inhibit microbial adherence. It is most important that, although some of the studies have given promising in vitro results, clinical in vivo testing has shown uniformly disappointing results. This is largely attributed to the formation of the conditioning film obliterating such surfaces.

BIOMATERIALS COATED OR IMPREGNATED WITH ANTISEPTICS AND ANTIBIOTICS

There is a great deal of confusion in the literature between coating and impregnation, with the terms being

used interchangeably. Because there are considerable implications for the use of these techniques in prevention of BRI, the distinction should be clarified. There are two ways of impregnating biomaterials. One consists of the admixture of the antimicrobial agents with the material during manufacture, and the other involves introduction of the agent into the manufactured material or device. Both should result in the antimicrobial being dispersed throughout the matrix of the material. Coatings can be applied directly to the surface of the manufactured device, or a binding agent can be utilized to fix them to the biomaterial surface. A third situation exists where a biomaterial is coated with a polymer into which an antimicrobial agent is introduced either before or after application of the surface polymer. The real difference between these lies in their efficacy. Factors that affect this are flow of blood or other fluid over the surface at risk (which removes the active agent), deposition of conditioning film on this surface (which obliterates the agent or inactivates it), and any depot effect that may create a durable antimicrobial activity. Leaving aside the issue of the conditioning film, the important site in terms of prevention of infection is the interface between the biomaterial and the fluid flow. Any antimicrobial activity (such as molecules or ions of the active agent) will be washed away from the surface by the flow, but a notional layer a few molecules thick (the Nernst layer^[32]) must remain saturated if protective activity is to be maintained. The source of active agent needed to replace that lost to fluid flow is either the coating or (in the case of impregnation) the biomaterial mass.

Because of their inherent low capacity to act as a depot, coatings are quickly depleted and their activity is short-lived (Fig. 2). If the biomaterial itself is impregnated with antimicrobial agent, the opportunity exists for long-lasting activity in the Nernst layer, and this in turn depends on the rate and mode of diffusion of the agent through the biomaterial.

If the agent is present as large aggregates and is water soluble, then so long as the aggregates are in contact with one another and with the surface, diffusion to the Nernst layer will take place by contiguous

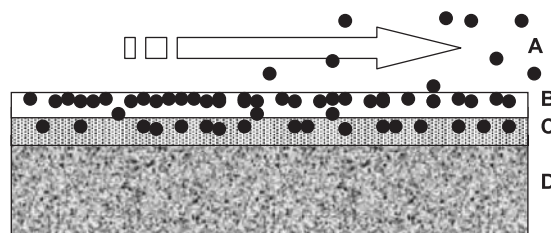


Fig. 2 Biomaterial with antimicrobial coating in fluid flow conditions. (A) Molecules of active agent lost to flow; (B) Nernst layer; (C) coating layer; (D) biomaterial.

dissolution. However, if the agent is present as molecules dispersed throughout the biomaterial (or as aggregates that are not in contact with one another), then any diffusion to replenish the Nernst layer will depend on the ability of the agent to diffuse through the biomaterial itself, and this is dependent on the Hansen parameter.^[33] Thus, unless antimicrobial agents introduced before or after manufacture into biomaterials are able to diffuse toward the surface, they will remain incarcerated and will therefore be useless. Considering the obliterative effect of the conditioning film in addition to the foregoing, it becomes clear why so little progress has been made in clinical efficacy in this field.

An early example of an antimicrobial biomaterial was antibiotic-loaded polymethylmethacrylate (PMMA) bone cement, used in joint replacement. Gentamicin (Refobacin) was added as a powder to the cement-monomer mix during preparation in the operating theater.^[34] The concentration of antibiotic was typically 2.5% by weight. There are now 18 brands of bone cement containing antibiotics, most of which contain gentamicin. Exceptions are a cement containing erythromycin and colistin, and another containing gentamicin and clindamycin. Copolymers are added to some brands but it is unclear whether this is intended to modify antibiotic release characteristics. Many studies have been conducted on the release of gentamicin from bone cement, although standardization of methodology is a major problem when comparing them. Many methods are ill conceived and inappropriate to the scientific question ostensibly being addressed. In a review of the effect of test methods, Bayston and Milner^[35] compared the release characteristics of three antibiotics (gentamicin sulphate, sodium fusidate, and diethanolamine fusidate) from two different cements and concluded that the mechanism of release (using the criteria of Flynn^[36]) depended on the drug particle size and distribution and on the physicochemical relationship between each drug and the cross-linked cement. Gentamicin thus appeared to be released predominantly by a Type I system (contiguous dissolution), whereas the other two drugs were released predominantly by a Type IIb system (diffusion through the PMMA). Each of these may have consequences for not only drug release characteristics but also for long-term mechanical properties of the cement. It is important in addition to be clear about the intended antimicrobial activity: Is it to protect the surface of the cement, or to infuse the surrounding tissues with enough antibiotic to replace levels normally achieved by systemic administration?

Gentamicin loading of PMMA has also been used in the form of beads approximately 0.8 mm in diameter strung on a steel wire. These can be used to pack an infected cavity to provide high local concentrations of

antibiotic for several days. Although these are useful in certain situations, caution should be exercised if they are used for long periods, as they have been found to provoke resistant metabolic variants of staphylococci.^[37] Attempts to produce a vascular graft protected by antibiotics date from the work of Krajicek et al. in 1969, in which antibiotics were added to the collagen used to coat the graft.^[38] A similar approach has been more recently described.^[39] However, the system most used for prevention of vascular graft infection is preinsertion soaking of the gelatin or collagen-sealed graft in a solution of rifampicin. Results of clinical studies remained controversial until the study by Braithwaite et al., who found no clinical benefit from this procedure, apparently because rifampicin is released too rapidly from the sealant.^[40]

In view of the high volume of catheter use and the high infection rates—almost one million infections annually in the United States^[41]—central venous catheters have been the subject of attempts to apply various antimicrobial processes. Elliott recently reviewed the latest use of antimicrobial agents to prevent catheter-related infections.^[42] Although some antibiotics have been found to adhere briefly to catheter biomaterial, many processes involve precoating the catheter with a binding agent. Trooskin et al.^[43] used tridodecylmethylammonium chloride (TDMAC) to bind penicillin to catheters and found that in static *in vitro* soak tests, activity could be extended to two weeks. In flow conditions this period is considerably shortened, however, and there are concerns about TDMAC toxicity. Jansen et al. used a more biocompatible polymer, PVP, to coat poly(urethane) catheters that were then soaked in teicoplanin solution.^[44] Laboratory studies demonstrated retained activity for 3–5 days, but a later clinical study showed no benefit from their use.^[45] Chlorhexidine and silver compounds have also been used, and are discussed in more detail later. A benzalkonium chloride-coated catheter has been the subject of a clinical trial in which it was shown to reduce the incidence of bacterial colonization by 45%.^[46] In order to circumvent the problems of short duration associated with hydromer coatings, Raad and Darouiche developed a catheter coated with minocycline and rifampicin, which showed enhanced *in vitro* activity compared to chlorhexidine/silver-coated catheters. When the two were compared in a clinical trial,^[47] this benefit was confirmed (the rifampicin/minocycline-coated catheters reduced infections by a factor of 12). This could be attributed to the very powerful antistaphylococcal activity of the antibiotics chosen, to their relative lipophilicity, and particularly to the fact that the catheters were coated on both inner and outer surfaces, unlike the chlorhexidine/silver catheters. Protective activity has not been shown beyond approximately 5 days, however,

and there is doubt about the antibiotics' efficacy in longer-term catheter usage. The question of selection or provocation of bacterial resistance^[48] will be answered only with continued clinical use.

The rate of silicone hydrocephalus shunt infections can be as high as 20% in high-risk groups. A process similar to that used to coat central venous catheters with PVP hydromer has been recommended for prevention of shunt infection. A laboratory study has been published^[49] in which catheter segments were soaked in various antibiotics. Results show release of antibiotics for a few days but the system has the significant disadvantage of antibiotic application on the outer surface only, and most shunt infections involve the inner surface. In another paper by the same investigators,^[50] abscess formation was noted in a rabbit model around hydromer-coated catheters, but not around noncoated silicone. This was also found when nonviable bacteria were inoculated, leading the authors to speculate that this may increase susceptibility to persistent infection. A method of true impregnation of silicone shunts has been developed^[51] and has since been explored extensively, partly because of the intimate long-term contact with neural tissue. In addition to intensive safety tests, efficacy has been studied using an *in vitro* model. Almost all infections in hydrocephalus shunts begin as microbial adherence and biofilm development inside the shunt catheters, with no contribution from the patient in terms of inflammatory or immune response. However, the flow of cerebrospinal fluid (CSF) through the catheter as well as the external tissue sink must be included in the simulation. The period of risk is confined to the time of insertion, and although bacteria can persist inside the tubing for long periods, the device is not usually subject to further microbial assault. Any assessment must allow for continuous perfusion of the catheters for several weeks to exclude later regrowth of any microbial survivors. Because many recipients of shunts have a high CSF protein level, there is concern that, unlike the situation with antimicrobial coatings, any antimicrobial protection will transcend the conditioning film. Because attached bacteria are less susceptible to antimicrobial action and are likely to survive longer than any surface protective activity in flow conditions (perhaps especially in the presence of conditioning film), these aspects have received close attention. The assessment systems have satisfied international regulatory scrutiny, especially because animal models suitable for an antimicrobial shunt are not currently available. When the duration of activity of processed shunt catheters was determined against serial staphylococcal challenge, the catheters were protected against colonization for up to 50 days, even in the presence of conditioning film.^[52] Further studies have shown that the process has no effect on bacterial

adhesion to the catheters, but that it kills all adhered bacteria in 48–52 hours (represented by the expression tK_{100}).^[53] In such devices it is essential to kill 100% of adhered bacteria. The time taken to achieve this must relate to the duration of any antimicrobial protective activity in simulated conditions in use. It is doubtful that there are many antimicrobial biomaterials currently available that could achieve tK_{100} under adequate test conditions. Clinical experience with the processed device involving over 12,000 implantations has so far shown a dramatic decrease in infection rates; quantitative results are awaited.

METALS AS ANTIMICROBIALS

Gold, silver, and copper ions are believed to have broad-spectrum antimicrobial activity and are unlikely to provoke resistance.^[20] In view of increasing bacterial resistance to antibiotics, metal coating may therefore be an effective strategy to prevent BRI. In comparison to most heavy metal ions, silver combines high antimicrobial activity with relatively low human toxicity. However, silver ion concentrations higher than 10 mg/L may be toxic to certain human cells.^[54,55] Inorganic silver salts and complexes such as argyrols, silver nitrate, carbonate, silver halides, and the metal itself are used in medicines. Nonetheless, the precise active mechanism remains unknown. A recent review of metal toxicity in yeasts sheds important light on a possible mode of antimicrobial action.^[56] The biocidal activity of silver is related to the biologically active silver ion (Ag^+) which, by binding to bacterial DNA, may inhibit a number of important transport processes (such as phosphate and succinate uptake), and can interact with cellular oxidation processes as well as the respiratory chain.^[57] The silver cation is a highly reactive chemical structure, which binds strongly to electron donor groups containing sulphur, oxygen, or nitrogen. Biological molecules generally contain all these components in the form of thio, amino, imidazole, carboxylate, and phosphate groups. Silver ions are active against a broad spectrum of nosocomial organisms, including multiresistant staphylococci and enterococci, gram-negative enteric bacilli, *Pseudomonas aeruginosa*, and *Candida albicans*, their antibacterial killing rate being directly proportional to Ag^+ concentrations. The desirable properties of such infection-resistant biomaterials have been listed by Guggenbichler et al.^[48] The gold standard should in essence thus possess a broad spectrum of antimicrobial activity and be resistant to microbial colonization. The activity should endure for the lifetime of the device, should not be reduced by contact with body fluids, and should not select drug-resistant organisms. Silver-coated biomaterials should satisfy many of these

requirements, and several methods have been developed for the surface coating of polymers with elemental silver. The techniques include vapor phase/plasma deposition or sputtering of silver coatings and ion beam-assisted deposition/ion implantation techniques using a vacuum process. The antimicrobial activity is short-lived and coatings may not affix permanently (i.e., may peel off), giving only temporary antimicrobial activity.

For the last 30 years, research groups worldwide have studied the microbiocidal activity of new types of medical devices (silver-nylon, silver-coated intramedullary pins, silver-oxide-coated Foley catheters, and silver-coated vascular prostheses) in vitro and in vivo.^[20,58–64] Most of the authors have claimed strong antimicrobial efficacy, and several products have been launched on the market. Despite promising activity in vitro, anti-infective prosthetic devices manufactured from these materials have yet to achieve any conclusive degree of success in clinical trials.^[65] Randomized clinical studies showing statistically significant efficacy of silver-coated medical devices in high-risk populations of patients are rare, deal with small numbers of patients only, or are controversial. A prospective randomized controlled clinical trial of latex urinary catheters coated with hydrogel (consisting primarily of hydrophilic polyurethane) and silver salts has found no significant benefit from their use.^[66] The largest randomized controlled clinical study with 1300 patients revealed no significant differences in infection rates between silver-coated and unmodified urinary catheters.^[60] A further concern revealed by this study was an increase in the incidence of *S. aureus* urinary tract infection (UTI) in the silver group. Another approach to the exploitation of silver in biomaterials is oligodynamic iontophoresis.^[67] This technology uses a biocompatible polymer impregnated with silver, platinum, and carbon particles. When in contact with body fluids, silver and platinum particles act as electrodes and release a steady flow of silver ions. Only in vitro experiments have been performed so far, in which sections of catheters manufactured from the biomaterial were incubated with suspensions of test bacteria. These in vitro results showed antibacterial activity superior to that of metallic silver coatings and were maintained for periods of up to 9 months. Silver-coated devices may be clinically effective only when the concentration of free silver ions can be increased and when contact with plasma proteins, chloride ions, and possible cytotoxic effects are minimized.^[61] However, silver complexed with a sulphonamide and an antiseptic (chlorhexidine) has been used to reduce central venous catheter infections. Although the papers refer to impregnated catheters, they are actually coated. A clinical benefit was shown in one trial^[12] but not in another.^[68] One important difference appeared to be

the duration of catheterization, implying that a benefit would apply only to short-term (i.e., about 5 days) catheters. Hypersensitivity has also been a problem, probably associated with chlorhexidine.^[69]

A

ALTERNATIVE APPROACHES—RECENT ADVANCES

So far, few anti-infective prosthetic devices have achieved any degree of success in clinical trials and practice. It may have been naive to imagine that incorporating antimicrobial agents into biomaterials would totally frustrate the colonization mechanisms that microbes have evolved over millions of years as a basic survival strategy in hostile habitats. Typically, metals cannot be dissolved in organic solvents, and the current methods of impregnation cannot be applied to metals such as silver. A new method has therefore been developed for impregnation of polymers with metal nanoparticles. Antibacterial activity tests are currently underway.

CONCLUSION

A comprehensive comparison of the various antibacterial biomaterials is difficult because a great diversity of methods has been used to demonstrate their activity. Whereas a few materials show promise, a gold standard biomaterial is yet to be developed. Even if we manage to produce such an anti-infective biomaterial, given time, the ingenious, rapidly evolving organisms we face will surely find ways of colonizing it. To maintain control over the problem of device-associated infection, a continuing effort to develop novel materials will be needed.

ARTICLES OF FURTHER INTEREST

Biofilms; Bone Cement; Burn Dressing; Drug Delivery, Controlled; Host Reactions; Infection of Medical Devices; Polymers; Surface Coatings; Vascular Grafts; Host Interactions

REFERENCES

1. Bayston, R. Medical Problems Due to Biofilms. In *Dental Plaque Revisited: Oral Biofilms in Health and Disease*; Newman, H.N., Wilson, M., Eds.; Boline: Cardiff, 1999.
2. Denstedt, J.D.; Wollin, T.A.; Reid, G. Biomaterials used in urology: Current issues of biocompatibility, infection and encrustation. *J. Endourol.* **1998**, *12*, 493–500.

3. Cobb, D.K.; High, K.B.; Sawyer, R.G.; Sable, C.A.; Adams, R.B.; Lindley, T.L.; Pruett, T.L.; Schwenzer, K.J.; Farr, B.M. A controlled trial of scheduled catheter replacement of central venous and pulmonary-artery catheters. *N. Engl. J. Med.* **1992**, *327*, 1062–1068.
4. Nasim, A.; Thompson, M.M.; Naylor, A.R.; Bell, P.R.F.; London, N.J.M. The impact of MRSA on vascular surgery. *Eur. J. Vasc. Endovasc. Surg.* **2001**, *22*, 211–214.
5. Bayston, R. Biofilms and Prosthetic Devices. In *Community Structure and Cooperation in Biofilms*; Allison, D.G., Gilbert, P., Lappin-Scott, H.M., Wilson, M., Eds.; Cambridge University Press: Cambridge, 2000.
6. Green, R.J.; Davies, M.C.; Roberts, C.J.; Tendler, S.J.B. Competitive protein adsorption as observed by surface plasmon resonance. *Biomaterials* **1999**, *20*, 385–391.
7. An, Y.H.; Friedman, R.J. Concise review of mechanisms of bacterial adhesion to biomaterial surfaces. *J. Biomed. Mater. Res.* **1998**, *43*, 338–348.
8. Boelens, J.J.; Dankert, J.; Murk, J.L.; Weening, J.J.; van der Poll, T.; Dingemans, K.P.; Koole, L.; Laman, J.D.; Zaat, S.A.J. Biomaterial-associated persistence of *Staphylococcus epidermidis* in pericatheter macrophages. *J. Infect. Dis.* **2000**, *181*, 1337–1349.
9. Bayston, R.; Penny, S.R. Excessive production of mucoid substance in *Staphylococcus SII A*: A possible factor in colonisation of Holter shunts. *Dev. Med. Child Neurol., Suppl.* **1972**, *27*, 25–28.
10. Christensen, B.E.; Charaklis, W.G. *Physical and Chemical Properties of Biofilms*; Characklis, K.C., Ed.; Wiley & Sons: New York, 1990; 93–130. W.G.M.
11. Mack, D.; Fischer, W.; Krokotsch, A.; Leopold, K.; Hartmann, R.; Egge, H.; Laufs, R. The intercellular adhesin involved in biofilm accumulation of *Staphylococcus epidermidis* is a linear beta-1,6-linked glucosaminoglycan: Purification and structural analysis. *J. Bacteriol.* **1996**, *178*, 175–183.
12. Maki, D.G.; Stolz, S.M.; Wheeler, S.; Mermel, L.A. Prevention of central venous catheter-related bloodstream infection by use of an antiseptic-impregnated catheter—A randomized, controlled trial. *Ann. Intern. Med.* **1997**, *127*, 257–266.
13. Raad, I.; Darouiche, R.; Hachem, R.; Sacilowski, M.; Bodey, G. Antibiotics and prevention of microbial colonization of catheters. *Antimicrob. Agents Chemother.* **1995**, *39*, 2397–2400.
14. Costerton, J.W.; Cook, G.; Lamont, R. The Community Architecture of Biofilms: Dynamic, Structures and Mechanisms. In *Dental Plaque Revisited: Oral Biofilm in Health and Disease*; Newman, H.N., Wilson, M., Eds.; Boline: Cardiff, 1999.
15. Brown, M.R.W.; Allison, D.G.; Gilbert, P. Resistance of bacterial biofilms to antibiotics: A growth related effect? *J. Antimicrob. Chemother.* **1988**, *22*, 777–780.
16. Ziebuhr, W.; Dietrich, K.; Trautmann, M.; Wilhelm, M. Chromosomal rearrangements affecting biofilm production and antibiotic resistance in a *Staphylococcus epidermidis* strain causing shunt-associated ventriculitis. *Int. J. Med. Microbiol.* **2000**, *290*, 115–120.
17. Stoodley, P.; Wilson, S.; Hall-Stoodley, L.; Boyle, J.D.; Lappin-Scott, H.M.; Costerton, J.W. Growth and detachment of cell clusters from mature mixed species biofilms. *Appl. Environ. Microbiol.* **2001**, *67*, 5608–5813.
18. Anders, C.; Gartner, R.; Steinert, V.; Voit, B.I.; Zschoche, S. Surface modification with hydrogels via macroinitiators for enhanced friction properties of biomaterials. *J. Macromol. Sci., Pure Appl. Chem.* **1999**, *A36* (7–8), 1017–1029.
19. Park, K.D.; Kim, Y.S.; Han, D.K.; Kim, Y.H.; Lee, E.H.B.; Suh, H.; Choi, K.S. Bacterial adhesion on PEG modified polyurethane surfaces. *Biomaterials* **1998**, *19* (7–9), 851–859.
20. Schierholz, J.M.; König, D.P.; Beuth, J.; Pulverer, G. Silver coating of medical devices—A review. *J. Hosp. Infect.* **1998**, *40*, 257–262.
21. Jucker, B.A.; Harma, H.; Hug, S.J.; Zehnder, A.B. Adsorption of bacterial surface polysaccharides on mineral oxides is mediated by hydrogen bonds. *Colloids Surf., B* **1997**, *9*, 331–343.
22. Gottenbos, B.; Grijpma, D.W.; van der Mei, H.C.; Feijen, J.; Busscher, H.J. Antimicrobials effects of positively charged surfaces on adhering gram-positive and gram-negative bacteria. *J. Mater. Sci., Mater. Med.* **1999**, *10*, 853–855.
23. Gorman, S.P.; Jones, D.S.; Mawhiney, W.M.; McGovern, J.G.; Adair, C.G. Conditioning fluid influences on the surface properties of silicone and polyurethane peritoneal catheters: Implications for infection. *J. Mater. Sci., Mater. Med.* **1997**, *8*, 631–635.
24. Price, C.; Waters, M.G.J.; Williams, D.W.; Lewis, M.A.O.; Stickler, D. Surface modification of an experimental silicone rubber aimed at reducing initial candidal adhesion. *J. Biomed. Mater. Res.* **2002**, *63*, 122–128.
25. Bellanda, M.; Cassinelli, C.; Morra, M. Reduced plaque accumulation on hydrocarbon thin film deposited on restorative acrylic polymers. *J. Biomed. Mater. Res.* **1997**, *36*, 216–222.
26. Miyayama, N.; Yoshinari, M.; Oda, Y. Surface modification of titanium implants with dry process-surface characterization. *Jpn. J. Dent. Mater.* **1999**, *18*, 109–121.
27. Yoshinari, M.; Oda, Y.; Kato, T.; Okuda, K.; Hirayama, A. Influence of surface modifications to titanium on oral bacterial adhesion in vitro. *J. Biomed. Mater. Res.* **2000**, *52*, 388–394.
28. Choong, S.; Wood, S.; Fry, C.; Whitfield, H. Catheter associated urinary tract infection and encrustation. *Int. J. Antimicrob. Agents* **2001**, *17*, 305–310.
29. Angell, A.H.; Resnick, M.I. Surface interaction between glycosaminoglycans and calcium oxalate. *J. Urol.* **1989**, *141*, 1255–1258.
30. Russell, J.C. Bacteria, biofilms, and devices: The possible protective role of phosphorylcholine materials. *J. Endourol.* **2000**, *14* (1), 39–42.
31. Keogh, J.R.; Velandar, F.F.; Eaton, J.W. Albumin—binding surfaces for implantable devices. *J. Biomed. Mater. Res.* **1992**, *26*, 441–456.
32. Nernst, W.Z. Theorie der Reaktionsgeschwindigkeit in heterogenen Systemen. *Z. Phys. Chem.* **1904**, *47*, 52–55.

33. Bustamente, P.; Martin, A.; Gonzalez-Guisandez, M.A. Partial solubility parameters and solvatochromic parameters for predicting the solubility of single and multiple drugs in individual solvents. *J. Pharm. Sci.* **1993**, *82*, 635–640.
34. Buchholz, H.W.; Engelbrecht, E. Über die Depotwirkung einiger Antibiotika beim Vermischen mit dem Kunstharz Palacos. *Chirurg* **1970**, *41*, 511–515.
35. Bayston, R.; Milner, R.D.G. The sustained release of antimicrobial drugs from bone cement. *J. Bone Jt. Surg.* **1982**, *64B*, 460–464.
36. Flynn, G.L. Influence of Physico-Chemical Properties of Drugs and System of Release of Drugs from Inert Matrices. In *Controlled Release of Biologically Active Agents. Advances in Experimental Biology and Medicine*; Tanquary, C., Lacey, R.E., Eds.; Plenum Pubs: New York, 1974; Vol. 47, 73–78.
37. von Eiff, C.; Lindner, N.; Proctor, R.A.; Winkelmann, W.; Peters, G. Auftreten von Gentamicin-resistenten Small Colony Variants von *S. aureus* nach Einsetzen von Gentamicin-ketten bei Osteomyelitis als Mögliche Ursache für Rezidive. *Z. Orthop.* **1998**, *136*, 268–271.
38. Krajicek, M.; Dvorak, J.; Chvapil, M. Infection-resistant synthetic vascular substitutes. *J. Cardiovasc. Surg.* **1969**, *10*, 454.
39. Chervu, A.; Moore, W.S.; Chvapil, M.; Henderson, T. Efficacy and duration of antistaphylococcal activity comparing three antibiotics bonded to Dacron vascular grafts with a collagen release system. *J. Vasc. Surg.* **1991**, *13*, 897–901.
40. Braithwaite, B.D.; Davies, B.; Heather, B.P.; Earnshaw, J.J. Early results of a randomized trial of rifampicin-bonded grafts for extra-anatomic vascular reconstruction. Joint Vascular Research Group. *Br. J. Surg.* **1998**, *85*, 1378–1381.
41. Widmer, A.F. Central Venous Catheters. In *Catheter-Related Infections*; Seifert, H., Jansen, B., Farr, B., Eds.; Marcel Dekker, Inc.: New York, 1997.
42. Elliott, T.S. Role of antimicrobial central venous catheters for the prevention of associated infections. *J. Antimicrob. Chemother.* **1999**, *43*, 441–446.
43. Trooskin, S.Z.; Donetz, A.P.; Harvey, R.A.; Greco, R.J. Prevention of catheter sepsis by antibiotic bonding. *Surgery* **1985**, *97*, 547–551.
44. Jansen, B.; Jansen, S.; Peters, G.; Pulverer, G. In vitro efficacy of a central venous catheter (hydrocath) loaded with teicoplanin to prevent bacterial colonisation. *J. Hosp. Infect.* **1992**, *22*, 93–107.
45. Bach, A.; Darby, D.; Böttiger, B.; Böhrer, H.; Motsch, J.; Martin, E. Retention of the antibiotic teicoplanin on a hydromer-coated central venous catheter to prevent bacterial colonization in postoperative surgical patients. *Intensive Care Med.* **1996**, *22*, 1066–1069.
46. Moss, H.A.; Tebbs, S.E.; Faroqui, M.H.; Herbst, T.; Isaac, J.L.; Brown, J.; Elliott, T.S. A central venous catheter coated with benzalkonium chloride for the prevention of catheter-related microbial colonization. *Eur. J. Anaesthesiol.* **2000**, *17*, 680–687.
47. Raad, I.; Darouiche, R.; Dupuis, J.; Abi-Said, D.; Gabrielli, A.; Hachem, R.; Wall, R.M.; Harris, R.; Jones, J.; Buzaid, A.; Robertson, C.; Shenaq, S.; Curling, P.; Burke, T.; Ericsson, C. Central venous catheters coated with minocycline and rifampicin for the prevention of catheter-related colonization and bloodstream infections. A randomized double-blind trial. *Ann. Intern. Med.* **1997**, *127*, 267–274.
48. Guggenbichler, J.-P.; Böswald, M.; Lugauer, S.; Krall, T. A new technology of microdispersed silver in polyurethane induces antimicrobial activity in central venous catheters. *Infection* **1999**, *27* (Suppl. 1), S16–S23.
49. Boelens, J.J.; Tan, W.-F.; Dankert, J.; Zaat, S.A.J. Antibacterial activity of antibiotic-soaked polyvinylpyrrolidone-grafted silicon (sic) elastomer hydrocephalus shunts. *J. Antimicrob. Chemother.* **2000**, *45*, 221–224.
50. Boelens, J.J.; Zaat, S.A.J.; Meeldijk, J.; Dankert, J. Subcutaneous abscess formation around catheters induced by viable and nonviable *Staphylococcus epidermidis* as well as by small amounts of bacterial cell wall components. *J. Biomed. Mater. Res.* **2000**, *50*, 546–556.
51. Bayston, R.; Grove, N.; Siegel, J.; Lawellin, D.; Barsham, S. Prevention of hydrocephalus shunt catheter colonisation in vitro by impregnation with antimicrobials. *J. Neurol. Neurosurg. Psychiatry* **1989**, *52*, 605–609.
52. Bayston, R.; Lambert, E. Duration of protective activity of cerebrospinal fluid shunt catheters impregnated with antimicrobial agents to prevent bacterial catheter-related infections. *J. Neurosurg.* **1997**, *87*, 247–251.
53. Bayston, R.; Ashraf, W.; Bhundia, C. Mode of action of an antimicrobial shunt material. *Eur. J. Pediatr. Surg.* **2002**, *12*, S56.
54. Rungby, J.; Hultman, P.; Ellermann-Eriksen, S. Silver affects viability and structure of cultured mouse peritoneal macrophages and peroxidative capacity of whole mouse liver. *Arch. Toxicol.* **1987**, *59*, 408–412.
55. Van de Belt, H.; Neut, D.; van Horn, J.R.; van der Mei, H.C.; Schenk, W.; Busscher, H.J. Or not to treat? *Nat. Med.* **1999**, *5*, 358–359.
56. Avery, S.V. Metal toxicity in yeasts and the role of oxidative stress. *Adv. Appl. Microbiol.* **2001**, *49*, 111–142. *in press*.
57. Modak, K.; Fox, C. Binding of silver sulfadiazine in the cellular components of *Pseudomonas aeruginosa*. *Biochem. Pharmacol.* **1973**, *22*, 2392–2404.
58. Liedberg, H.; Lunderberg, T. Silver alloy coated catheters reduce catheter-associated bacteremia. *Br. J. Urol.* **1990**, *65*, 379–381.
59. Lundeberg, T. Prevention of catheter-associated urinary tract infections by use of silver-impregnated catheters. *Lancet* **1986**, *2*, 1031.
60. Riley, D.K.; Classen, D.C.; Stevens, L.E.; Burke, J.P. A large randomised clinical trial of silver-impregnated urinary catheter: Lack of efficacy and staphylococcal superinfection. *Am. J. Med.* **1995**, *98*, 349–358.
61. Schierholz, J.M.; Wachol-Drebeck, L.; Lucas, L.; Pulverer, G. Activity of silver ions in biological fluids. *Zbl. Bakteriologie* **1998**, *287*, 411–420.
62. Johnson, J.R.; Robert, P.L.; Olsen, R.J. Prevention of catheters-associated urinary tract infection with a

- silver oxide-coated urinary catheter: Clinical and microbiologic correlates. *J. Infect. Dis.* **1990**, *162*, 1145–1150.
63. Furr, J.R.; Russell, A.D.; Turner, T.D.; Andrews, A. Antibacterial activity of Actisorb plus, Actisorb and silver nitrate. *J. Hosp. Infect.* **1994**, *27*, 201–208.
 64. Stickler, D.J. Biomaterials to prevent nosocomial infections: Is silver the gold standard? *Curr. Opin. Infect. Dis.* **2000**, *13* (4), 389–393.
 65. Darouiche, R.O. Anti-infective efficacy of silver-coated medical prostheses. *Clin. Infect. Dis.* **1999**, *29*, 1371–1377.
 66. Thibon, P.; Le Coutour, X.; Leroyer, R.; Fabry, J. Randomised multicentre trial of the effects of a catheter coated with hydrogel and silver salts on the incidence of hospital-acquired urinary tract infections. *J. Hosp. Infect.* **2000**, *45*, 117–124.
 67. Milder, F. Device-related nosocomial infection: Reducing infection with antimicrobial materials and coatings. *Med. Device Technol.* **1999**, *10*, 34–39.
 68. Logghe, C.; van Ossel, C.; Hoore, C.; Ezzedine, H.; Wauters, G.; Haxhe, J.J. Evaluation of chlorhexidine and silver-sulfadiazine impregnated central venous catheters for the prevention of bloodstream infection in leukaemic patients: A randomized controlled trial. *J. Hosp. Infect.* **1997**, *37*, 145–156.
 69. Center for Devices and Radiological Health. *Potential Hypersensitivity Reactions to Chlorhexidine-Impregnated Medical Devices*; FDA Public Health Notice, March 11, 1998.

Antimineralization Treatment

A

Bart Meuris
Willem Flameng

Department of Cardiovascular Diseases, Katholieke Universiteit Leuven, Belgium

INTRODUCTION

Surgical replacement of heart valves with artificial devices revolutionized the care of patients with valvular heart disease. In past decades, countless models of substitute heart valves were developed and investigated, and the technology of their design and manufacture has progressively improved. Nevertheless, an ideal valve replacement device has not been achieved.^[1]

Heart valve substitutes are of two principal types: mechanical prosthetic valves with components manufactured of nonbiologic material (e.g., polymer, metal, carbon) or tissue valves that are constructed, at least in part, of either human or animal tissue. At least 275,000 valves are implanted worldwide each year, of which it is estimated approximately 55% are mechanical and 45% are tissue, with a clear trend toward an increasingly greater usage of tissue valves over the past several years.

Heart valve substitutes are of two principal types: mechanical prosthetic valves with components manufactured of nonbiologic material (e.g., polymer, metal, carbon) or tissue valves that are constructed, at least in part, of either human or animal tissue. At least 275,000 valves are implanted worldwide each year, of which it is estimated approximately 55% are mechanical and 45% are tissue, with a clear trend toward an increasingly greater usage of tissue valves over the past several years.

Tissue valves have been used since the early 1960s, when aortic valves obtained fresh from human cadavers were transplanted to other individuals. Less than a decade later, chemically preserved tissue valves (termed bioprostheses) were first implanted. The term bioprosthesis usually denotes a valve that is chemically treated and mounted on a stent. Today, stented glutaraldehyde-preserved porcine aortic valves and bovine pericardial bioprosthetic valves are used extensively. Although tissue valve substitutes have certain advantages over mechanical valves, their durability is limited.

Mechanical and tissue valves differ in their major causes of failure. Specifically, mechanical prosthetic valves are associated with a substantial risk of systemic thromboemboli and thrombotic occlusion, largely owing to nonphysiologic surfaces and flow abnormalities created by one or more rigid occluders. To minimize this risk, chronic anticoagulation therapy is required in all mechanical valve recipients. However, systemic anticoagulation renders patients vulnerable to potentially serious hemorrhagic complications. Thus, the combined risk of thromboembolic complications and hemorrhage constitutes the principal disadvantage of mechanical prosthetic valves. In contrast with mechanical valves, tissue valves maintain a low rate of thromboembolism without anticoagulation. Their central pattern of flow approximates that of natural heart valves and their cusps are composed of valvular or non-valvular animal tissue, which is typically more thrombo-resistant and biocompatible than synthetic materials.

However, valve failure with structural dysfunction owing to progressive tissue deterioration, including calcification and non-calcific damage, is a serious disadvantage that undermines the attractiveness of tissue valve substitutes and limits their use in younger patients.^[2] Currently, most clinical heart centers use bioprosthetic valves in patients 65–70 years old or older. The life span of current bioprostheses in this age group is around 12–15 years, meaning that these patients will have reached the 8th or 9th decade of life at the moment of bioprosthetic valve failure or have died from other causes. When the cutoff age of 70 years is respected, the percentage of patients presenting for rereplacement of a dysfunctional, degenerated bioprosthesis is extremely low. Implanting bioprosthetic valves in younger patients results in a certain occurrence of valve degeneration within a time frame of several years. Valve calcification in children under the age of 15 can occur within 2 years. Unfortunately, it is precisely the younger patient who would benefit the most from the avoidance of lifelong anticoagulation. Therefore, further research toward the creation of a tissue valve with lifelong durability, even in young patients, is absolutely necessary. In order to accomplish this goal, accurate and reliable insights into the mechanisms and factors that control and influence degeneration and calcification in current bioprostheses must be obtained.

Bioprosthetic heart valves fail as a result of calcification and non-calcific deterioration. Non-calcific damage relates to the occurrence of valve dysfunction in areas not affected by deposition of calcific aggregates. Cuspal tears or perforations in uncalcified parts, pannus overgrowth, leaflet retraction, and dehiscence are examples of non-calcific damage.^[3] Calcification and non-calcific deterioration may occur independently of one another or they may be synergistic. Calcification of bioprosthetic valves has determinants related to 1) host metabolism; 2) implant structure and chemistry; and 3) mechanical factors, with several complex and poorly understood interactions.^[4]

BIOMINERALIZATION OF SOFT TISSUE

The Biomineralization Process

Calcification of biological tissue occurs in two different phases: nucleation and growth. Nucleation is the more

important and critical phase, because calcium and phosphate concentrations in normal body fluids are insufficient to allow spontaneous precipitation. Both in normal bone formation and in pathological calcification, a nucleating site must be present. A nucleating site is a localized microenvironment in which inorganic calcium and phosphate can interact in optimal circumstances and concentrations, allowing spontaneous precipitation and formation of an initial amorphous calcium phosphate crystal. Once nucleation occurs, the concentration of calcium and phosphate in body fluids is sufficient for crystal propagation and growth. Calcium and phosphate in extracellular fluids are in a metastable state: both ions normally do not precipitate, but when certain conditions are met (e.g., higher localized concentrations or the presence of an initial crystal), precipitation can and will take place.

Calcific deposits in mineralized tissue (either physiologic or pathologic) essentially are calcium phosphate salts, presenting in the form of hydroxyapatite. Hydroxyapatite is $\text{Ca}_5(\text{PO}_4)_3(\text{OH})$, but is usually written as $\text{Ca}_{10}(\text{PO}_4)_6(\text{OH})_2$ to denote that the crystal cell unit comprises two molecules. Although mammalian extracellular fluids are metastable toward hydroxyapatite, the crystal normally does not spontaneously precipitate. Hydroxyapatite requires a substrate that serves as a nucleator for crystal precipitation in bone and teeth as well as in ectopic calcification. Once the initial precipitation occurs, further crystal growth and expansion are just a matter of time. Hydroxyapatite is formed from its preliminary compounds, namely amorphous calcium phosphate and other ions in a different crystalline form such as octa-calcium phosphate, $\text{Ca}_8\text{H}_2(\text{PO}_4)_6$.^[5]

Bone Ossification

In bone formation, mineralization is an energy-requiring process. Living osteogenic cells generate small (100 nm) membranous cell fragments, so-called matrix vesicles, by pinching off vesicles from the outer plasma membrane. This reflects a stage in programmed cell death (apoptosis). These matrix vesicles serve as the initial site of bone calcification in all skeletal tissues. Initial hydroxyapatite mineral crystals are generated at the level of matrix vesicle membranes. This first phase of physiological mineralization is controlled by phosphatases (including alkaline phosphatase) and several Ca^{2+} -binding molecules, which are abundantly present in matrix vesicles. The second phase of biomineralization begins with the breakdown of matrix vesicle membranes, exposing preformed hydroxyapatite to the extracellular fluid, after which crystal proliferation can proceed, governed by extracellular conditions. Both phases of mineralization are under

cellular control: the first phase is initiated by osteogenic cells generating calcifiable matrix vesicles, releasing them into sites of intended calcification. The second phase is controlled by cells regulating extracellular ionic conditions and matrix composition. The whole process implies a variety of cellular elements including mesenchymal cells, osteoblasts, and osteoclasts, and an endless list of mediators such as phosphatases, Ca^{2+} -binding proteins, bone-morphogenetic proteins, and chemoattractants either stimulating or inhibiting crystal growth.

Pathologic Calcification of Living Soft Tissue

Pathological or ectopic mineralization of soft tissues, including articular cartilage and cardiovascular tissues, leads to morbidity and mortality. Recent findings suggest that the mechanisms and factors regulating physiological bone mineralization may be identical or at least similar to those regulating ectopic mineralization.^[6] Pathological mineralization is also initiated by matrix vesicles, membrane-enclosed particles released from the plasma membrane of mineralization-competent cells. In addition, several regulators (activators and inhibitors) of physiological mineralization have been identified and characterized, and evidence indicates that the same factors also contribute to the regulation of ectopic mineralization.

Bioprosthetic Tissue Calcification

In calcification of biological heart valves, the end product—the calcific deposit—is similar to that in bone. Calcific deposits in calcified bioprostheses are calcium phosphate salts, presenting in the form of hydroxyapatite.^[7,8] Unfortunately, the similarity stops there. When considering apatite formation and growth in either physiological mineralization (bone and teeth) or pathological mineralization (arteries, kidneys, muscles), one must deal with energy-requiring mechanisms involving living and functioning tissue cells that perform their functions, excrete mediators, and/or chemoattract other cell types. At the time of implantation, the tissue of a glutaraldehyde-fixed bioprosthesis is not a living physiological environment and energy-requiring processes do not take place. At the time of implantation, a bioprosthetic valve is devoid of living cells. The fixation of tissue with glutaraldehyde does not allow many living cells to enter, stay alive after entering, or perform their functions. The key elements providing physiologic or pathologic soft tissue mineralization are most likely not valid in bioprosthetic tissue calcification.

Terminology

Ectopic calcification is defined as inappropriate biomineralization occurring in soft tissue. In patients with chronic renal failure, a systemic mineral imbalance is associated with widespread ectopic calcification, referred to as metastatic calcification. In the absence of a systemic mineral imbalance, ectopic calcification is typically termed dystrophic calcification. Often, these sites provide evidence of tissue alteration or necrosis. Dystrophic mineralization is commonly observed in soft tissues as a result of injury, disease, and aging. Calcification of bioprosthetic valves in a normocalcemic host can be categorized as dystrophic calcification.

Extrinsic calcification of a tissue valve refers to calcific deposits occurring outside the actual valvular tissue. This may take place in cells adherent to the valve, formed neointima, thrombi, or vegetation. Although all these phenomena are relevant to global valve function and long-term behavior, the main focus of interest is the intrinsic valve calcification, occurring within the substance of the tissue of which the valve is constructed.

TISSUE VALVES: CELLS, FIBERS, AND GROUND SUBSTANCE

Cells and Cell Remnants

Prelevation, subsequent processing and fixation of tissue harvested from an animal, results in irreversible cell death. Cell death leads to the presence of cellular debris such as dead or disrupted cells, membranous cellular degradation products, remnants from the cell nucleus, free cell organelles, and fragments from the endoplasmic reticulum. These cellular remnants demonstrate ultrastructural similarities to matrix vesicles^[9] and have been identified as a potential nucleation site for initial mineralization.^[7,9-11] The mechanisms responsible for this cell-mediated mineralization are uncertain, but a few hypotheses have been formulated.^[7,8,10-12]

1. Intact living connective tissue cells have an intracellular free calcium concentration of around 10^{-7} M, whereas the extracellular free calcium concentration is 10^{-3} M, leading to a 10,000-fold gradient across the plasma membrane. In dead cells and in cells modified by aldehyde cross-linking, Ca^{2+} enters the cell passively, whereas the mechanisms for Ca^{2+} removal (membrane-bound, energy-requiring Ca^{2+} pumps) are no longer active. The intracellularly sequestered Ca^{2+} in mitochondria and endoplasmic reticulum gains free access to the cytoplasm, a mechanism that is certainly important for the abundant smooth muscle cell

portion in the aortic wall portion. The next step is a reaction with intracellular phosphate (PO_4^-). Phosphate is abundantly present throughout the cell, but is highly prevalent within organellar and plasma membranes in the form of phospholipids, proteolipids, and enzymatic systems. The massively released free Ca^{2+} in the cell binds to phosphate and leads to the formation of an early calcium phosphate-based mineral.^[11,12] Hereby, the initial step toward hydroxyapatite formation is set.

2. Phospholipids in membranes not only carry negatively charged phosphates, but also their phosphate groups can be set free through the action of activated phosphatases (triggered through free Ca^{2+} ions). Free inorganic phosphate easily reacts with Ca^{2+} to form an amorphous calcium phosphate precipitate.
3. Not only does calcium have an affinity for phospholipids, but also many membrane-bound proteins and enzymes have a high affinity for Ca^{2+} . These large proteins can attract and bind Ca^{2+} ions.
4. In pathologic cartilage ossification, a non-energy-requiring Ca^{2+} trap has been postulated through the formation of calcium-phospholipidic complexes that can nucleate hydroxyapatite from a metastable calcium phosphate solution.^[13] This non-energy-requiring process may also play a role in bioprosthetic tissue calcification.

Cells, cell remnants, membrane fragments, and organelles are all observed sites of primary nucleation of bioprosthetic valve mineralization.^[7,8] The small, early crystal nuclei will progressively accumulate additional mineral compounds, eventually growing to macroscopic calcific deposits. The role of phospholipids in the initiation of calcification is not to be underestimated. In vivo hydroxyapatite formation induced by phospholipids is demonstrated both in normal physiologic bone formation processes and in pathologic soft tissue mineralization. In commercially available tissue heart valves, a significant correlation exists between cuspal phospholipid levels and calcium content. Consequently, it is not surprising that tissue treatments aimed at removing cells and membrane remnants, thereby reducing phospholipid levels, are successful in reducing calcification, at least in the cusps. However, it is not evident to what extent these treatments effectively also prevent collagen- or elastin-related calcification. Even if cellular remnants are efficiently removed, there is evidence for host cell infiltration in implanted bioprosthetic valves. These cells may become entrapped in a cytotoxic glutaraldehyde-fixed matrix, eventually die, and in that way produce renewed cell debris.

Collagen

Collagen fibers are present in all tissues used for bioprosthetic heart valves. They provide an architectural framework and are important for the mechanical strength required for adequate valve function. The most abundant type of collagen in valvular material is collagen type I.

Collagen fibers have the ability to calcify. There is evidence of calcium deposition in isolated type I collagen samples and in acellular collagen preparations.^[12,14] In both porcine and pericardial cusps, early phases of mineral depositions are described to occur along collagen fibers, whereas calcification along cell debris was minimal in these studies. Calcific crystal nucleation in bioprosthetic tissue seems to be able to occur by different independent mechanisms in cells and in collagen. In porcine aortic wall calcification, the contribution of collagen fiber calcification is probably minimal.^[4]

Collagen fibers are built from consecutive collagen molecules (tropocollagen). In between the 300-nm-long tropocollagen molecules, small 40-nm gaps exist, also known as “hole zones.” Different protective proteoglycans reside in these holes. In bone, this area (the hole zone) is the initial site of crystallization of the collagen fiber at the time the fiber is incorporated in the apatite crystals originating from matrix vesicles. A few possible mechanisms may cause early calcific deposits to occur near the collagen fiber.

1. In bioprosthetic tissue, matrix metalloproteinases (MMPs) and other proteolytic enzymes are able to degrade the protective proteoglycans attached to collagen, thereby exposing sites in collagen prone to calcification.
2. Repetitive mechanical stress can damage the structural integrity of collagen. In damaged collagen fibers, the zones in between consecutive tropocollagen molecules become larger, facilitating calcium deposition.
3. Collagen has a certain affinity for calcium-binding proteins containing γ -carboxyglutamic acid (Gla). These Gla-containing proteins may either originate directly from blood or are secreted locally by infiltrated host lymphocytes and macrophages, infiltrated in the bioprosthesis. Gla proteins are known for their ability to attract and bind calcium.

Elastin

Elastic fibers are abundant structural proteins in the porcine aortic wall and are functionally important components of porcine aortic cusps. Ultrastructurally, elastic fibers consist of a core composed of cross-linked

elastin molecules and an external coating of microfibrils and proteoglycans located at the periphery of the amorphous component. Elastin has particular properties that set it apart from other matrix components. Glutaraldehyde does not chemically react with elastin. Elastin is intermolecularly cross-linked to form desmosine and isodesmosine residues and contains very few lysine groups to react with glutaraldehyde. Radioactive [³H]glutaraldehyde studies confirmed very low reactivity between glutaraldehyde and elastin.^[15]

Calcification of the elastic fiber in humans is well studied. The interest in elastic fiber calcification comes from observations in atherosclerotic arterial wall calcification and in senile aortic wall mineralization that can occur independent of atherosclerosis. Several possible mechanisms and pathways have been described.

1. The external coating of the elastic fiber protects elastin from calcification. Disruption of this protective coating exposes the elastin core to the extracellular environment. Elastin possesses a unique structure that facilitates calcium binding: the elastin core has several binding sites for large cations, such as Ca^{2+} .
2. Matrix metalloproteinases such as MMP-9 and MMP-2 bind and degrade insoluble elastin to soluble peptides. These elastin degradation products can elicit immune responses and trigger unwanted reactions in host cells, such as protease release and apoptosis. The mechanism is well described and occurs through the elastin–laminin receptor, which is present on the surface of most cells.^[16] Vascular smooth muscle cells can assume osteoblastic phenotypes and have osteogenic activity in cell culture experiments.^[17]
3. The elastic fibers in the aortic medium endure several changes with advancing age, including changes in their amino acid composition.^[18] This mechanism is crucial in mineralization of the aging aorta, even in the absence of atheromatous disease. The aged elastic fiber includes a diminished amount of structural proteins, an increasing content of polar amino acids (aspartic and glutamic acid), and introduction of carbonyl groups into amino acid residues. These polar residues are negatively charged and can attract counterions. The fiber thus becomes more susceptible to the activity of elastase; it may totally disintegrate and calcify. These age-related chemical changes in elastin occur independent of the onset and evolution of atheromatous lesions.

The role of elastic fibers in bioprosthetic valve calcification is less clear. Elastin can be damaged during the harvesting and preparation of the tissue.

Disruption of elastic fiber integrity was illustrated in bioprosthetic valves, where ultrastructural studies reported damage to the outer protective elastin coating. The lack of reactivity of glutaraldehyde with elastin leaves the elastic fiber unprotected and susceptible to the activity of degradative enzymes such as MMPs. Local delivery of MMP inhibitors can reduce calcification in subdermally implanted elastin. In porcine aortic valve cusps, degradation of elastic fibers through elastase causes an increase in radial dimensions with a corresponding decrease in radial extensibility and a threefold increase in stiffness. Elastin damage in cusps may be a factor in progressive cusp degeneration.

Ground Substance: Proteoglycans (PGs) and Glycosaminoglycans (GAGs)

Ground substance of tissues is composed largely of PGs carrying GAGs (mucopolysaccharides). Proteoglycans are a major component of the animal extracellular matrix, the “filler” substance existing between cells and fibers. They form large complexes, both to other proteoglycans and also to fibrous matrix proteins (such as collagen). Proteoglycans represent a special class of heavily glycosylated proteins, consisting of a large core protein with one or more covalently attached GAG chains. Solutions of GAGs have a slimy, mucus-like consistency that results from their high viscosity and elasticity. These GAG chains are long, linear carbohydrate polymers that are negatively charged under physiological conditions because of the occurrence of sulfate and uronic acid groups. The highly negative charged side chains of PGs result in an enormous water-binding capacity, which gives these PGs their important mechanical functions: 1) maintaining matrix hydration and volume; and 2) giving tissue the ability to resist and absorb compressive forces.

In a native porcine valve, PGs are mainly located in the spongiosa layer, which is the layer that must absorb the mechanical stress forces during valve closure. The spongiosa also enables shearing stresses between the fibrosa and ventricularis layers. In pericardium, PGs are mainly composed of dermatan and chondroitin sulfate. These highly charged units contribute to tissue hydration and elasticity but also have an important interaction with the abundant collagen fibers. Just like elastin, PGs and GAGs are not effectively stabilized by glutaraldehyde.

In bioprosthetic valves, the role of PGs and GAGs on mineralization is controversial but one thing is clear: losing GAGs is not beneficial to long-term durability of tissue. The presence of GAGs and PGs protects tissues from calcification. In vitro models demonstrated inhibited hydroxyapatite formation after

the addition of high chondroitin sulfate concentrations through a mechanism in which PGs block the active nucleating sites for further apatite crystal growth.^[19] And selective PG extraction from pericardium results in greater accumulation of calcium salts compared with non-extracted tissue, with reduced hydrothermal tissue stability.^[20] Glycosaminoglycan removal from porcine aortic valve cusps results in decreased cuspal thickness, reduced water content, diminished rehydration capacity, and important increases in the flexural rigidity of the cusp.

The mechanisms through which we lose PGs are twofold. First, preparation of bioprosthetic heart valves with the entire process of tissue harvesting, manipulation, fixation, etc., induces GAG loss.^[21] Valve cusps stored under static conditions in 0.2% glutaraldehyde for 6 mo lose over 50% of their original GAG content.^[21] Controlling GAG loss through more elaborate or organ-like preservation techniques could be beneficial. The second factor causing GAG loss is mechanical stress during cyclic functioning of the cusp.^[22] Given this possible importance of maintaining PG and GAG levels in cuspal tissue, alternative tissue fixations that are able to cross-link and fixate both collagen fibers as PGs have been investigated. Methods using products such as acyl azide, carbodiimide with or without the use of diamines, and periodate^[23] have all demonstrated good tissue preservation and reduction of calcification, at least in rat subcutaneous models. In properly cryopreserved porcine tissue, no significant alteration in the content, molecular size, or distribution of PGs occurs.

FACTORS INFLUENCING BIOPROSTHETIC TISSUE CALCIFICATION

Used Material

Amino acid analyses demonstrated that fresh bovine pericardial tissue consists of almost 90% collagen.^[12] Cells and ground substance are scarce in pericardium. Pericardial collagen is almost exclusively type I. Calcified pericardium has a reduced hydroxyproline content, indicating that part of the collagen disappears through absorption or simple breakdown. Several studies attempted to elucidate the origin of pericardial calcification. Unfortunately, the results are controversial: some results stress the appearance of initial deposits along remnants of connective tissue cells,^[12] whereas others claim this initiation occurs at the level of collagen fibrils.

Compared with pericardium, porcine aortic cusps contain more cells. Collagen accounts for approximately 40%, with 1% elastin and 5% GAGs. Collagen is both type I and type III.^[24] This may be of

importance because type III collagen mineralization has never been demonstrated in contrast with type I collagen. A fair comparison between porcine cusps and pericardial tissue regarding its tendency for calcification is difficult. Results will be influenced by the models used (subcutaneous vs circulatory), the exact tissue samples or valves used (influence of tissue handling and design), applied tissue treatments, etc. Clinical evidence points toward longer longevity for pericardial valves. Experimental results are mixed.

The porcine aortic wall portion consists of over 50% elastin and approximately 15% collagen and 2% GAGs. The rest consists of cells, with an abundance of smooth muscle cells. Experimental data indicate an important cell-mediated mechanism for aortic wall calcification,^[25] and a possible role for elastin is not to be excluded.^[15]

Glutaraldehyde

Glutaraldehyde reacts with tissue proteins to form intra- and intermolecular cross-links, resulting in a fixed, stable, non-immunogenic, and sterile tissue. Glutaraldehyde is the reason why bioprosthetic heart valves exist today.

Glutaraldehyde is a bifunctional agent that creates a series of various chemically stable cross-links, mainly between lysine (Lys) and hydroxylysine (Hyl) residues of tissue proteins. At low concentrations, glutaraldehyde produces mainly intramolecular cross-links in collagen, whereas at higher concentrations long polymeric chains of glutaraldehyde form and the reaction is reversible. After implantation of glutaraldehyde cross-linked biomaterial, large polymers of glutaraldehyde are continuously hydrolyzed, which results in the release of free monomeric glutaraldehyde.

Because collagen carries many Lys and Hyl residues, the collagen fiber is extensively cross-linked by glutaraldehyde. The nature and extent of the cross-linked network depends on fixation time, reagent concentration, and fixation temperature.^[26] Studies with radioisotopes of glutaraldehyde (³H]glutaraldehyde) clearly demonstrated a linear relation between glutaraldehyde uptake and tissue stability (resistance to enzymatic digestion).^[27]

Three major problems with glutaraldehyde have never been resolved: 1) tissue permeability: glutaraldehyde penetration in tissue decreases with increasing tissue depth; 2) heterogeneity of glutaraldehyde cross-linking with the formation of highly polymerized glutaraldehyde; and 3) liberation of free glutaraldehyde from the polymeric cross-links.

A few hypotheses have been postulated through which glutaraldehyde as such can be responsible for bioprosthetic tissue mineralization.

1. The pyridinium cross-link produced by glutaraldehyde is a quaternary amine that can cause an influx of PO_4^{3-} into the interfibrillar spaces of collagen, leading to hydroxyapatite formation.
2. Tissue valves contain residual amounts of aldehyde in the tissue after preparation, which can damage blood cells of infiltrating host cells after implantation. The damaged cells can adhere to the surface of the valve and act as a promoter of calcific changes.
3. The released residual glutaraldehyde is harmful to endothelial overgrowth, leaving the biological tissue uncovered for Ca^{2+} -rich plasma insudation.
4. Residual glutaraldehyde as well as unstable glutaraldehyde polymers retained in the interstitium have been implicated in cytotoxicity toward host cells and inflammatory reactions.
5. Glutaraldehyde fixation disables the dynamic interaction between fibrosa and ventricularis, resulting in worsened mechanical properties.

Because early clinical and experimental studies demonstrated that formaldehyde cross-linked valves have grossly inadequate durability, glutaraldehyde is now utilized virtually exclusively. Many different protocols exist concerning exposure time, exact concentrations, applied pressure, and temperature during fixation. The rate of glutaraldehyde incorporation by bovine pericardium is rapid, with nearly half maximal incorporation after only 1 hr and complete reaction occurring in less than 24 hr.^[12] Contrary to a widely held belief, some studies demonstrated beneficial effects of higher concentrations of glutaraldehyde. Several protocols using very early (within 30 min) fixation in a high concentration of glutaraldehyde of over 1% demonstrated lower degrees of tissue mineralization. The mechanisms of this protective effect are purely hypothetical but can be the result of a more efficient tissue fixation, less tissue permeability for diffusing ions, less cell death prior to fixation, or less free glutaraldehyde residuals. Until now, this high-glutaraldehyde strategy has not reached clinical practice.

Blood Contact

The clinical fate of every bioprosthesis is implantation in the blood stream. The interaction of fixed tissue with circulating blood is complex. Evidence exists that the tissue will be confronted with an insudation of lipids, proteins, cholesterol, lipoproteins, lipid-loaded macrophages, etc.^[28,29] Finally, the prosthesis will also be confronted with circulating free plasma calcium that will insudate into the tissue. The effect of blood contact on tissue mineralization is demonstrated in experimental settings.

Host Age

Host age is an undisputed factor in bioprosthetic valve calcification. Calcification is more rapid and aggressive in younger patients. Bioprosthetic valves implanted in children under the age of 15 will calcify to a fully dysfunctional valve within 2–3 years after implantation.^[30] Freedom from structural valve deterioration at 15 years in a population younger than 65 is around 65%, whereas in a population older than 70 it is over 90%. Although the relationship is well established, the mechanisms accounting for the effect of age are uncertain.

One explanation for this dramatic influence of host age is the difference in overall calcium handling and calcium metabolism, including higher levels of blood calcium, phosphate, bone proteins, and enhanced parathyroid hormone and vitamin D metabolism in young patients. Although this phenomenon limits the clinical use of bioprostheses in young patients, it gives the opportunity for preclinical experimental valve testing in young animals, producing high rates of valve calcification within short periods of time.

Mechanical Stress

The role of mechanical stress on bioprosthetic tissue calcification is complex. Implanted valves are subjected to large mechanical stresses during each cardiac cycle, with high transient stresses during valve closure and cyclic flexural and compressive stresses during valve opening. Because of these important repetitive forces, bioprosthetic valves can fail mechanically (cuspal tears) and this can occur in the absence of calcification.^[31] On the other hand, many clinical observations indicate a relation between mechanical stress and cuspal calcification: cusp calcification is more severe in mitral position compared with the aortic position and certainly compared with the tricuspid position. Experimentally, stress and valve calcification have been correlated.^[32]

Within the valve structure, mechanical stresses are the highest near the commissures and lowest at the base of the leaflets.^[33] At a pressure of 100 mm Hg applied to a closed porcine valve, the maximal stress is 160 kPa near the commissures, 70 kPa in the body of the leaflet, 50 kPa at the free edge, and only 40 kPa at the cuspal base. A perfectly similar distribution is obtained for maximal shear stresses in these different regions.^[33] Computer-based mathematical models demonstrated the concentration of stresses in the commissural region.^[34]

When this stress distribution is compared with the distribution of calcific deposits in clinically explanted (degenerated) valves, the regions of high stress coincide

with calcified areas.^[33,35] In over 50% of degenerated tissue valves, regions of calcification coincide with the location of cuspal tears and perforations, indicating that the stress to which the tissue is subjected first leads to calcification and subsequently to mechanical failure.^[35] Because cuspal tears occur mostly in the commissural region, a causative relation between the magnitude of mechanical stress and the site of calcification is more than likely.

The question remains regarding which mechanism is responsible for the relation between stress and mineralization. Three major elements may have a fundamental role: collagen, GAGs, and proteolysis.

1. In a series of *in vitro* accelerated fatigue studies with porcine valves, infrared spectroscopy revealed progressive and severe damage to the molecular structure of type I collagen, with loss of molecular helicity.^[22] Damaged collagen can serve as an initiation site for calcium deposition (*cf.* supra), and it will also lead to structural weakening of the tissue. Experimental animal studies demonstrated that in porcine cusps, calcification occurs earlier in the commissural region than at the base, with the earliest calcific deposits localized within collagen cords.^[32] In pericardial cusps, calcification is observed along the zone of leaflet flexion and initiates at the level of deformed collagen fibers.^[32] Pericardial specimens exposed to collagenase and dynamic stress load fail faster than specimens exposed to collagenase and static load.^[36]
2. Valve fatigue also leads to loss of GAGs from the cuspal extracellular matrix.^[22] The GAG levels dropped from 6.5 $\mu\text{g}/\text{mg}$ in glutaraldehyde-fixed cusps to 0.7 $\mu\text{g}/\text{mg}$ after cyclic fatigue. Loss of GAGs (*cf.* supra) leads to altered biomechanical tissue characteristics, as documented in decreased cuspal bending strength^[22] and altered 3-D cuspal geometry.^[37]
3. Dynamic stress accelerates local proteolytic activity in bovine pericardium.^[36]

In summary, there is evidence for a relation between stress and cuspal tears, between calcification and commissural regions, and between stress distribution and commissural regions. Stress can lead to either ultrastructural damage and initiation of calcification or primary tissue failure and non-calcific tears. In native, normal aortic valve cusps, the damage accumulated by cyclic stress will be repaired and resolved by living fibroblasts. Damaged material will be scavenged, collagen fibers repaired, and tissue GAG levels restored. In a bioprostheses, these processes of scavenging and repair are absent. The keystone for valvular and commissural stress distribution is valve design. Numerical computer models guided valve designers

toward altering stent height and tissue attachment, introducing flexible stents, etc., with the common goal of reducing leaflet (commissural) stress.

Non-collagenous Extracellular Proteins

In physiological bone mineralization, more than 17 different non-collagenous extracellular proteins (NCPs) have a crucial role in the mineralization process. These proteins (osteonectin, osteopontin, osteocalcin, matrix-Gla protein, bone sialoprotein II, bone acidic glycoprotein) are highly anionic proteins, contain high proportions of negatively charged aspartic acid and Gla, and have high affinity for Ca^{2+} and hydroxyapatite. Mainly Gla residues bind strongly to calcium and all of its precipitating salts. The NCPs are critical to the establishment and regulation of mineralization because type I collagen dissociated from NCPs is incapable of supporting hydroxyapatite formation. Together with the whole family of bone-morphogenetic proteins, their role was studied extensively in normal bone formation and in pathological dystrophic vascular calcification related to atherosclerosis. A wide variety of mesenchymal cells and inflammatory cells can synthesize the specific NCPs required to promote and regulate the development of a mineralized matrix. The strong chemotactic properties of NCPs may further promote cell recruitment and continued matrix protein expression. Non-collagenous extracellular proteins have been detected in atherosclerotic plaques and in calcified arteries and native valves.^[38] We are dealing with living tissues, and living fibroblasts, smooth muscle cells, and macrophages respond and interact to biohormonal stimuli. The exact role of this wide array of complex proteins in the calcification of fixed bioprosthetic valve tissue is far from elucidated.

Different NCPs are present in close anatomic proximity with zones of calcification in degenerated bioprosthetic valves. The main cell type expressing these proteins at the surface of the bioprostheses is macrophages, but infiltration with other cell types such as T lymphocytes, monocytes, and mesenchymal cells was also described.^[39] All these cell types are able to synthesize osteopontin, osteonectin, and other NCPs.

Osteocalcin exhibits a high affinity for calcium, phospholipids, and hydroxyapatite. Osteocalcin can be passively absorbed from plasma but it can also be expressed in situ by infiltrating cells. The presence of NCPs such as osteocalcin and osteopontin was demonstrated in clinically explanted calcified bioprostheses and in calcified tissues from rat subdermal models, so the adsorption of these proteins may have a role.

Because many of these NCPs are vitamin-K-dependent proteins, attempts to inhibit bioprosthetic tissue calcification with high-dose warfarin therapy have been performed,^[40,41] although without success.

Enzymes: Phosphatases and Proteinases

Glutaraldehyde-treated tissues have always been regarded as resistant to enzymatic degradation, including enzymes released from dead cells, because all proteins are subject to chemical cross-linking. However, different studies revealed that a variety of enzymes, either originating from the fixed tissue itself or infiltrating from host blood or host cells, can retain their activity. Alkaline phosphatase (ALP) has an important function in the pathogenesis of mineral nucleation in bone and has a role in pathological vascular and non-vascular calcification. Alkaline phosphatase can boost local PO_4^{3-} concentrations by the hydrolysis of organic phosphates. Following standard glutaraldehyde fixation, considerable ALP activity is present in pericardial tissue. Further ultrastructural studies demonstrate enzymatic reaction products along membranes of endothelial cells and interstitial fibroblasts, the sites of early calcific deposits in bioprosthetic valves.^[42,43] The widely held view that glutaraldehyde eliminates all metabolic activities of bioprosthetic tissue might be inaccurate and examination of the role of ALP and other phosphatases may stimulate approaches for inhibiting calcification.

Next to ALP, there are the MMPs, a family of Zn^{2+} -dependent endopeptidases, that contribute to tissue remodeling and repair. The MMPs are produced by several cell types including fibroblasts, macrophages, endothelial cells, and eosinophils. Concerning their possible involvement in tissue degradation (with subsequent calcification), MMPs such as collagenase and elastase were actively present in glutaraldehyde-fixed aortic cusp and wall tissue.^[44] Tissue proteases existed both in soluble form and covalently immobilized to the cross-linked matrix. Both damaged collagen and damaged elastin not only lead to tissue failure, but also enhance collagen- and elastin-mediated calcification. Further research is required to establish whether the removal of these enzymatic activities or their inhibition using specific tissue inhibitors of matrix proteinases can enhance the longevity of biomaterials.

Altered Calcium Metabolism in the Host

The non-normocalcemic host (chronic renal failure, hyperparathyroidism) will reveal early and rapid calcification of an implanted bioprosthesis. Disturbances of mineral metabolism such as hyperphosphatemia and hypercalcemia contribute to tissue calcification. In ectopic calcification, mechanisms include not only passive precipitation of elevated ions but also actively induced changes in vascular smooth muscle cell behavior toward an osteoblast-like phenotype. Specific calcium-regulatory proteins may act locally or systemically. Dysregulations of calcification inhibitors, including fetuin A, matrix Gla

protein, osteoprotegerin, and pyrophosphates, may also be pathophysiologically relevant factors.

Immune Responses

Unlike the situation in cryopreserved homografts, in which the immune response toward the valve is important, the role of immunologic reactions in bioprosthetic tissue calcification is probably limited. Although 1) experimental animals can be sensitized to both fresh and cross-linked bioprosthetic valve tissues;^[45] 2) antibodies to valve components can be detected in some patients following valve dysfunction;^[46] and 3) failed tissue valves often have mononuclear inflammation, no definite role has been demonstrated for circulating macromolecules or immunocompetent cells. In particular, the presence of mononuclear cells in a failed tissue valve does not equate to immunologic rejection. In all papers attempting to demonstrate immune-mediated injury, no evidence indicates that the mononuclear inflammatory cells cause functional valve degeneration. Experiments in congenitally athymic mice (lacking normal T-cell function) demonstrated similar calcification compared with calcium levels in normal mice in subcutaneously implanted fragments.^[47] The morphologic characteristics of the calcification were the same in both groups, with dystrophic mineralization of the spongiosa predominating. Thus, normal T-lymphocyte function is not necessary for porcine bioprosthetic calcification, and immunologic processes do not contribute to this process.^[40,48]

Infection

Clinical experience demonstrated that infection of a bioprosthetic valve leads to rapid calcification. Experimentally, tissue valves affected by endocarditis calcify extensively. Whether this is still intrinsic tissue mineralization or predominantly extrinsic calcific deposits in vegetations is unclear. The mechanisms involved are unclear but the severe inflammatory and infectious processes that take place at the level of an infected valve, with the secretion of several mediators, lytic enzymes, and chemoattractants, must be detrimental for tissue stability and integrity.

ANTICALCIFICATION TREATMENTS

Analogous to any new or modified drug or device, a potential antimineralization treatment must meet rigorous efficacy and safety requirements. Investigations of an anticalcification strategy must demonstrate both the effectiveness of the therapy and the absence of adverse effects. The treatment should not impede valve

performance such as hemodynamics and durability. Adverse effects in this setting could include systemic or local toxicity, tendency toward thrombosis on infection, induction of immunologic effects, or structural degradation, with either immediate loss of mechanical properties or premature deterioration and failure. There are several instances in which an anti-mineralization treatment contributed to unacceptable degradation of the tissue.

Non-Glutaraldehyde-Based Tissue Fixation

Implanting xenogeneic biological tissue requires a chemical process that stabilizes the tissue and protects it from immunologic and enzymatic attack. The processes developed to date were based on cross-linking with chemical agents, of which glutaraldehyde is indisputably the most widely used. The controversial status of glutaraldehyde and possible links with tissue calcification pointed research groups toward new chemical agents capable of inducing cross-linking of the side chains of amino acid residues other than Lys and Hyl. However, the search for an alternative cross-linker is still open.^[49]

Possible alternative cross-linking methodologies include glycerol,^[50] carbodiimide,^[51] acyl azide,^[52,53] epoxy compounds,^[54] genipin,^[55] glutaraldehyde acetals,^[56] poly(ethylene glycol),^[57] and poly(acrylamide) hydrogels,^[58] with numerous modifications and variations. All these tissue treatments aim at achieving a stable, cross-linked, and non-immunogenic tissue through a different type of biochemical bond than that induced by glutaraldehyde. Physical methods such as ultraviolet irradiation and dehydrothermal freeze drying have been investigated.

Despite an abundance of experimental reports claiming tissue stability, anticalcification potential, and possible future clinical application in heart valves, none of these alternative cross-linkers has made a clinical breakthrough. An exception must be made for dye-mediated photooxidation (DMPO). In DMPO, methylene blue and light are used to create cross-links in biological tissue without the use of aldehyde chemistry.^[59] The tissue is placed in a phosphate-buffered saline solution containing the photoactive methylene blue dye. Light of an appropriate wavelength is used to activate the dye catalyst. Oxygen reacts with the dye to form highly reactive oxygen free radicals, which in turn interact with specific amino acid side chains to form new cross-links.^[59,60] Both photooxidized pericardial and porcine tissue has been investigated extensively, with promising results toward biostability, softness, pliability, and calcification resistance.^[59,61] Unfortunately, a negative clinical experience with design-related valve failures in a stented prosthesis with photooxidized

pericardium, blunted further developments in this promising tissue treatment.^[62]

Surfactants (Detergents)

Detergents act by two mechanisms: first by removing acidic phospholipids and second by reducing phospholipid insudation as a result of the hydrophobic barrier preventing their penetration into valve tissue. Possible detergents include sodium dodecyl sulfate (SDS), polysorbate 80 (Tween- 80), Triton X-100, and *N*-lauryl sarcosine. Not all detergents can be used to treat valvular tissue because some (Triton X-100 and *N*-lauryl sarcosine) induce alterations that decrease the durability of the valves,^[63] despite the fact that they are the most efficient detergents. Substances such as SDS and Tween 80 are applied in currently clinically used bioprostheses. Both treatments demonstrated successful antimineralization properties in subcutaneous models,^[64,65] whereas the results with SDS in circulatory models are controversial.^[66] Tween 80 seems to perform better.

Alcohols

Delipidating agents such as methanol/chloroform and ethanol act by removing phospholipids. In pericardial tissue, selective lipid extraction using chloroform/methanol had no effect on tissue stability and demonstrated less calcium accumulation compared with glutaraldehyde-treated controls in rat models.^[67] After accelerated wear testing, however, the delipidated samples calcified in the zones that underwent mechanical stress.^[68] The use of delipidating agents require long-term studies to test the influence of such treatments on tissue properties. Ethanol affects multiple components of valvular tissue.^[69] Ethanol pretreatment induced conformational changes in cuspal collagen (documented through infrared spectroscopy), reduced *in vitro* cholesterol uptake by cusps, increased resistance to collagenase digestion, and reduced water content of cusps.^[69] Cuspal glutaraldehyde content was not changed by ethanol pretreatment. Although these changes can result in significant reductions in calcification potential in rats, the question remains of whether these alterations in ultrastructure will result in deteriorated mechanical properties. Ethanol in sheep circulatory studies prevented cusp calcification.^[66]

Amino-oleic Acid

Although exact mechanisms of action are unknown, 2- α -amino-oleic acid (AOA) is able to mitigate cusp calcification.^[70,71] It is hypothesized that AOA

covalently binds, as a result of a Schiff base reaction, to free aldehyde groups of tissue-bound glutaraldehyde. Next to blocking or reducing the free aldehyde groups, reduced calcium diffusion through AOA-treated leaflets is thought to be part of the AOA mechanism. The AOA inhibition of cusp calcification in rat subdermal implants is exposure time dependent with maximum treatment time (120 hr), resulting in the lowest calcium levels. AOA has poor results as a chemical treatment on the aortic wall portion, certainly in stentless valves.^[66,72] A possible explanation is limited penetration and binding of AOA into the aortic wall tissue, with AOA levels being consistently lower than that of the cusps after the same exposure times.^[72]

Cations

The use of trivalent cations like Al^{3+} and Fe^{3+} originates from the occurrence of osteomalacia in renal dialysis patients. Patients in chronic renal failure accumulate high Al^{3+} levels, with disturbed bone mineralization as a consequence. This observation led to strategies using metallic salts containing Al^{3+} and Fe^{3+} as a tissue treatment to prevent calcification. Tissue incubation in solutions such as AlCl_3 and FeCl_3 demonstrated markedly inhibited calcification in rat models.^[73] The action mechanism is not clear but morphological studies localized both cations at the level of devitalized connective tissue cells, indicating that the mode of action is likely related to the high affinity of these cations for membrane-associated and other intracellular PO_4^{3-} loci.^[73] Al^{3+} and Fe^{3+} can act by forming complexes with phosphate, thereby displacing calcium. This is in the short term, however, because the low solubility product of calcium salts and the calcium concentration in blood could lead to the removal of Al^{3+} and Fe^{3+} in the medium or long term. Pretreatment of tissue with AlCl_3 also inhibited ALP activity, which will result in less mineralization.^[74] Unfortunately, in a circulatory sheep model of mitral valve replacement, Al^{3+} treatment results in higher cusp calcification, necessitating a differential pretreatment of the cusp and aortic wall in order to avoid Al^{3+} -induced calcification in the leaflets.^[75,76]

From successful results in mitigating aortic wall calcification with AlCl_3 pretreatment in rat models, a possible protective effect of Al^{3+} on elastin was hypothesized.^[15] Using isolated elastin preparations, elastic-fiber-oriented calcification is indeed less after Al^{3+} treatment. Aluminum binds to elastin at similar neutral binding sites (peptide carbonyl groups) and causes structural changes in the spatial conformation of elastic fibers, demonstrated by spectroscopy.^[15] Aluminum ion binding can hypothetically block calcium from binding or interacting with elastin. It is,

however, still uncertain whether Al^{3+} will be sufficient to prevent aortic wall calcification when tested in adequate sheep models.

Diphosphonates

Diphosphonates (or biphosphonates or aminodiphosphonates) act by inhibiting calcium phosphate crystal growth and are able to bind covalently to residual glutaraldehyde. Diphosphonates can either be injected systemically^[77] or be used to pretreat the valvular tissue by incubating the tissue in it.^[78] Ethane hydroxydiphosphonate (EHB) and related compounds are drugs used in humans for pathologic calcification disorders like hypercalcemia. These drugs inhibit ALP and interfere with calcium-phosphate crystal formation, but also affect normal bone formation. Injections of EHB or diphosphonates covalently linked to the valve tissue have been investigated experimentally with reduced calcium levels compared with glutaraldehyde-fixed controls. Unfortunately, many problems related to side effects and drug delivery (effect on bone formation, limited durability of the drug, washout effect of blood) limit further clinical use.

Other Strategies and Combination Treatments

When reviewing experimental literature regarding antimineralization treatments, numerous agents and products can be retrieved. Some have demonstrated promising results and have applications in clinical products, whereas others remain experimental. These treatments include No-React[®], glutamic acid, urazole, homocysteine acid, bilinx, toluidine blue, poly (acrylamide), tannic acid, and many others. No-React is a proprietary combination treatment involving heparin applied in some clinically available products. Bilinx is a combination treatment of ethanol and Al^{3+} . The main focus is application in stentless valves to avoid both cusp- and wall-mediated calcification. Tannic acid is a plant polyphenol, known to form multiple bonds with proteins rich in proline, such as elastin.^[79] Treatment of aortic wall samples with tannic acid after standard glutaraldehyde fixation increased resistance to elastase and reduced wall calcification.

Acellularization

Given the importance of the cellular component in bioprosthesis calcification, efforts have been made not only to diminish the total phospholipid content by detergents, but also to accomplish complete acellularization. Removal of cells and cell remnants can be performed through several protocols such as sequential

freeze–thawing of tissue to reach osmotic disruption of cells, extensive rinsing in detergents, or treatment in enzymatic solutions containing, for example, DNase and RNase. One porcine acellular valve construct reached clinical use for right ventricular outflow tract reconstruction, but with the additional tissue protection using cryopreservation as in homografts.^[80] Whether these acellular, unfixed constructs will repopulate with cells is unclear, and their behavior in left-side implantations is uncertain.^[81] Experimental evidence has accumulated that acellularization as such is not the final answer. There is a need for additional tissue fixation or treatment that stimulates rapid recellularization with host cells.^[82,83]

A

Materials Other Than Porcine or Bovine Tissue

Experimental work has been performed using several other tissue sources, such as kangaroo aortic valve tissue,^[84] equine pericardial tissue,^[85] and bovine jugular vein tissue.^[86] Neither experimental results nor preliminary clinical data concerning the bovine jugular vein constructs demonstrate superior characteristics thus far.

ARTICLES OF FURTHER INTEREST

Bone–Implant Interface; Bone Remodeling; Hard Tissue: Biomaterials Interaction; Orthopedic Fixation Devices; Osteoclastic Bone Resorption; Osteoinductive Substances and Materials

REFERENCES

1. Schoen, F.J.; Levy, R.J. Founder's Award, 25th Annual Meeting of the Society for Biomaterials, Perspectives. Providence, RI, April 28–May 2, 1999. Tissue heart valves: Current challenges and future research perspectives. *J. Biomed. Mater. Res.* **1999**, *47* (4), 439–465.
2. Hammermeister, K.; Sethi, G.K.; Henderson, W.G.; Grover, F.L.; Oprian, C.; Rahimtoola, S.H. Outcomes 15 years after valve replacement with a mechanical versus a bioprosthetic valve: Final report of the Veterans Affairs randomized trial. *J. Am. Coll. Cardiol.* **2000**, *36* (4), 1152–1158.
3. Sacks, M.S.; Schoen, F.J. Collagen fiber disruption occurs independent of calcification in clinically explanted bioprosthetic heart valves. *J. Biomed. Mater. Res.* **2002**, *62* (3), 359–371.
4. Simionescu, D.T. Prevention of calcification in bioprosthetic heart valves: Challenges and perspectives. *Expert Opin. Biol. Ther.* **2004**, *4* (12), 1971–1985.
5. Samachson, J. Basic requirements for calcification. *Nature* **1969**, *221* (5187), 1247–1248.

6. Kirsch, T. Determinants of pathological mineralization. *Curr. Opin. Rheumatol.* **2006**, *18* (2), 174–180.
7. Schoen, F.J.; Levy, R.J.; Nelson, A.C.; Bernhard, W.F.; Nashef, A.; Hawley, M. Onset and progression of experimental bioprosthetic heart valve calcification. *Lab. Invest.* **1985**, *52* (5), 523–532.
8. Valente, M.; Bortolotti, U.; Thiene, G. Ultrastructural substrates of dystrophic calcification in porcine bioprosthetic valve failure. *Am. J. Pathol.* **1985**, *119* (1), 12–21.
9. Kim, K.M. Cell injury and calcification of rat aorta in vitro. *Scan. Electron Microsc.* **1984**, *4*, 1809–1818.
10. Kim, K.M. Cells, rather than extracellular matrix, nucleate apatite in glutaraldehyde-treated vascular tissue. *J. Biomed. Mater. Res.* **2002**, *59* (4), 639–645.
11. Kim, K.M. Cellular mechanism of calcification and its prevention in glutaraldehyde treated vascular tissue. *Z. Kardiol.* **2001**, *90* (Suppl. 3), 99–105.
12. Schoen, F.J.; Tsao, J.W.; Levy, R.J. Calcification of bovine pericardium used in cardiac valve bioprostheses. Implications for the mechanisms of bioprosthetic tissue mineralization. *Am. J. Pathol.* **1986**, *123* (1), 134–145.
13. Anderson, H.C. Normal and abnormal mineralization in mammals. *Trans. Am. Soc. Artif. Intern. Organs* **1981**, *27*, 702–708.
14. Levy, R.J.; Schoen, F.J.; Sherman, F.S.; Nichols, J.; Hawley, M.A.; Lund, S.A. Calcification of subcutaneously implanted type I collagen sponges. Effects of formaldehyde and glutaraldehyde pretreatments. *Am. J. Pathol.* **1986**, *122* (1), 71–82.
15. Vyavahare, N.; Ogle, M.; Schoen, F.J.; Levy, R.J. Elastin calcification and its prevention with aluminum chloride pretreatment. *Am. J. Pathol.* **1999**, *155* (3), 973–982.
16. Lee, J.S.; Basalyga, D.M.; Simionescu, A.; Isenburg, J.C.; Simionescu, D.T.; Vyavahare, N.R. Elastin calcification in the rat subdermal model is accompanied by up-regulation of degradative and osteogenic cellular responses. *Am. J. Pathol.* **2006**, *168* (2), 490–498.
17. Jian, B.; Narula, N.; Li, Q.Y.; Mohler, E.R., III; Levy, R.J. Progression of aortic valve stenosis: TGF- β 1 is present in calcified aortic valve cusps and promotes aortic valve interstitial cell calcification via apoptosis. *Ann. Thorac. Surg.* **2003**, *75* (2), 457–465.
18. LaBella, F.S.; Vivian, S.; Thornhill, D.P. Amino acid composition of human aortic elastin as influenced by age. *J. Gerontol.* **1966**, *21* (4), 550–555.
19. Chen, C.C.; Boskey, A.L. Mechanisms of proteoglycan inhibition of hydroxyapatite growth. *Calcif. Tissue Int.* **1985**, *37* (4), 395–400.
20. Jorge-Herrero, E.; Fernandez, P.; Gutierrez, M.; Castillo-Olivares, J.L. Study of the calcification of bovine pericardium: Analysis of the implication of lipids and proteoglycans. *Biomaterials* **1991**, *12* (7), 683–689.
21. Lovekamp, J.J.; Simionescu, D.T.; Mercuri, J.J.; Zubiate, B.; Sacks, M.S.; Vyavahare, N.R. Stability and function of glycosaminoglycans in porcine bioprosthetic heart valves. *Biomaterials* **2006**, *27* (8), 1507–1518.
22. Vyavahare, N.; Ogle, M.; Schoen, F.J.; Zand, R.; Gloeckner, D.C.; Sacks, M.; et al. Mechanisms of bioprosthetic heart valve failure: Fatigue causes collagen denaturation and glycosaminoglycan loss. *J. Biomed. Mater. Res.* **1999**, *46* (1), 44–50.
23. Mercuri, J.J.; Lovekamp, J.J.; Simionescu, D.T.; Vyavahare, N.R. Glycosaminoglycan-targeted fixation for improved bioprosthetic heart valve stabilization. *Biomaterials* **2007**, *28* (3), 496–503.
24. Mannschott, P.; Herbage, D.; Weiss, M.; Buffevant, C. Collagen heterogeneity in pig heart valves. *Biochim. Biophys. Acta* **1976**, *434* (1), 177–183.
25. Meuris, B.; Ozaki, S.; Herijgers, P.; Verbeken, E.; Flammeng, W. Influence of species, environmental factors, and tissue cellularity on calcification of porcine aortic wall tissue. *Semin. Thorac. Cardiovasc. Surg.* **2001**, *13* (4 Suppl. 1), 99–105.
26. Jayakrishnan, A.; Jameela, S.R. Glutaraldehyde as a fixative in bioprostheses and drug delivery matrices. *Biomaterials* **1996**, *17* (5), 471–484.
27. Golomb, G.; Schoen, F.J.; Smith, M.S.; Linden, J.; Dixon, M.; Levy, R.J. The role of glutaraldehyde-induced cross-links in calcification of bovine pericardium used in cardiac valve bioprostheses. *Am. J. Pathol.* **1987**, *127* (1), 122–130.
28. David, T.E.; Ivanov, J.; Armstrong, S.; Feindel, C.M.; Cohen, G. Late results of heart valve replacement with the Hancock II bioprosthesis. *J. Thorac. Cardiovasc. Surg.* **2001**, *121* (2), 268–278.
29. Butany, J.; Collins, M.J.; Nair, V.; Leask, R.L.; Scully, H.E.; Williams, W.G.; et al. Morphological findings in explanted Toronto stentless porcine valves. *Cardiovasc. Pathol.* **Jan 2006**, *15* (1), 41–48.
30. Fiddler, G.I.; Gerlis, L.M.; Walker, D.R.; Scott, O.; Williams, G.J. Calcification of glutaraldehyde-preserved porcine and bovine xenograft valves in young children. *Ann. Thorac. Surg.* **1983**, *35* (3), 257–261.
31. Ishihara, T.; Ferrans, V.J.; Boyce, S.W.; Jones, M.; Roberts, W.C. Structure and classification of cuspal tears and perforations in porcine bioprosthetic cardiac valves implanted in patients. *Am. J. Cardiol.* **1981**, *48* (4), 665–678.
32. Thubrikar, M.J.; Deck, J.D.; Aouad, J.; Nolan, S.P. Role of mechanical stress in calcification of aortic bioprosthetic valves. *J. Thorac. Cardiovasc. Surg.* **1983**, *86* (1), 115–125.
33. Sabbah, H.N.; Hamid, M.S.; Stein, P.D. Mechanical stresses on closed cusps of porcine bioprosthetic valves: correlation with sites of calcification. *Ann. Thorac. Surg.* **1986**, *42* (1), 93–96.
34. Vesely, I.; Krucinski, S.; Campbell, G. Micromechanics and mathematical modeling: An inside look at bioprosthetic valve function. *J. Card. Surg.* **1992**, *7* (1), 85–95.
35. Stein, P.D.; Kemp, S.R.; Riddle, J.M.; Lee, M.W.; Lewis, J.W., Jr.; Magilligan, D.J., Jr. Relation of calcification to torn leaflets of spontaneously degenerated porcine bioprosthetic valves. *Ann. Thorac. Surg.* **1985**, *40* (2), 175–180.
36. Ellsmere, J.C.; Khanna, R.A.; Lee, J.M. Mechanical loading of bovine pericardium accelerates enzymatic degradation. *Biomaterials* **1999**, *20* (12), 1143–1150.
37. Smith, D.B.; Sacks, M.S.; Pattany, P.M.; Schroeder, R. Fatigue-induced changes in bioprosthetic heart valve

- three-dimensional geometry and the relation to tissue damage. *J. Heart Valve Dis.* **1999**, *8* (1), 25–33.
38. Farzaneh-Far, A.; Proudfoot, D.; Shanahan, C.; Weissberg, P.L. Vascular and valvar calcification: Recent advances. *Heart* **2001**, *85* (1), 13–17.
 39. Srivatsa, S.S.; Harrity, P.J.; Maercklein, P.B.; Kleppe, L.; Veinot, J.; Edwards, W.D.; et al. Increased cellular expression of matrix proteins that regulate mineralization is associated with calcification of native human and porcine xenograft bioprosthetic heart valves. *J. Clin. Invest.* **1997**, *99* (5), 996–1009.
 40. Levy, R.J.; Schoen, F.J.; Levy, J.T.; Nelson, A.C.; Howard, S.L.; Oshry, L.J. Biologic determinants of dystrophic calcification and osteocalcin deposition in glutaraldehyde-preserved porcine aortic valve leaflets implanted subcutaneously in rats. *Am. J. Pathol.* **1983**, *113* (2), 143–155.
 41. Levy, R.J.; Zenker, J.A.; Bernhard, W.F. Porcine bioprosthetic valve calcification in bovine left ventricle-aorta shunts: Studies of the deposition of vitamin K-dependent proteins. *Ann. Thorac. Surg.* **1983**, *36* (2), 187–192.
 42. Maranto, A.R.; Schoen, F.J. Alkaline phosphatase activity of glutaraldehyde-treated bovine pericardium used in bioprosthetic cardiac valves. *Circ. Res.* **1988**, *63* (4), 844–848.
 43. Maranto, A.R.; Schoen, F.J. Phosphatase enzyme activity is retained in glutaraldehyde treated bioprosthetic heart valves. *ASAIO Trans.* **1988**, *34* (3), 827–830.
 44. Simionescu, A.; Simionescu, D.; Deac, R. Biochemical pathways of tissue degeneration in bioprosthetic cardiac valves. The role of matrix metalloproteinases. *ASAIO J.* **1996**, *42* (5), M561–M567.
 45. Dahm, M.; Lyman, W.D.; Schwell, A.B.; Factor, S.M.; Frater, R.W. Immunogenicity of glutaraldehyde-tanned bovine pericardium. *J. Thorac. Cardiovasc. Surg.* **1990**, *99* (6), 1082–1090.
 46. Rocchini, A.P.; Weesner, K.M.; Heidelberger, K.; Keren, D.; Behrendt, D.; Rosenthal, A. Porcine xenograft valve failure in children: an immunologic response. *Circulation* **1981**, *64* (2 Pt. 2), III162–III171.
 47. Levy, R.J.; Schoen, F.J.; Howard, S.L. Mechanism of calcification of porcine bioprosthetic aortic valve cusps: Role of T-lymphocytes. *Am. J. Cardiol.* **1983**, *52* (5), 629–631.
 48. Gong, G.; Seifter, E.; Lyman, W.D.; Factor, S.M.; Blau, S.; Frater, R.W. Bioprosthetic cardiac valve degeneration: role of inflammatory and immune reactions. *J. Heart Valve Dis.* **1993**, *2* (6), 684–693.
 49. Hendriks, M.; Everaerts, F.; Verhoeven, M. Alternative fixation of bioprostheses. *J. Long Term Eff. Med. Implants* **2001**, *11* (3-4), 163–183.
 50. Gong, G.; Ling, Z.; Seifter, E.; Factor, S.M.; Frater, R.W. Aldehyde tanning: The villain in bioprosthetic calcification. *Eur. J. Cardiothorac. Surg.* **1991**, *5* (6), 288–299.
 51. Everaerts, F.; Gillissen, M.; Torrianni, M.; Zilla, P.; Human, P.; Hendriks, M.; et al. Reduction of calcification of carbodiimide-processed heart valve tissue by prior blocking of amine groups with monoaldehydes. *J. Heart Valve Dis.* **2006**, *15* (2), 269–277.
 52. Anselme, K.; Petite, H.; Herbage, D. Inhibition of calcification in vivo by acyl azide cross-linking of a collagen-glycosaminoglycan sponge. *Matrix* **1992**, *12* (4), 264–273.
 53. Simmons, D.M.; Kearney, J.N. Evaluation of collagen cross-linking techniques for the stabilization of tissue matrices. *Biotechnol. Appl. Biochem.* **1993**, *17* (Pt. 1), 23–29.
 54. Shen, S.H.; Sung, H.W.; Tu, R.; Hata, C.; Lin, D.; Noishiki, Y.; et al. Characterization of a polyepoxy compound fixed porcine heart valve bioprosthesis. *J. Appl. Biomater.* **1994**, *5* (2), 159–162.
 55. Sung, H.W.; Chang, Y.; Chiu, C.T.; Chen, C.N.; Liang, H.C. Mechanical properties of a porcine aortic valve fixed with a naturally occurring crosslinking agent. *Biomaterials* **1999**, *20* (19), 1759–1772.
 56. Goissis, G.; Yoshioka, S.A.; Braile, D.M.; Ramirez, V.D. The chemical protecting group concept applied in crosslinking of natural tissues with glutaraldehyde acetals. *Artif. Organs* **1998**, *22* (3), 210–214.
 57. Vasudev, S.C.; Chandy, T. Polyethylene glycol-grafted bovine pericardium: a novel hybrid tissue resistant to calcification. *J. Mater. Sci. Mater. Med.* **1999**, *10* (2), 121–128.
 58. Oosthuysen, A.; Zilla, P.P.; Human, P.A.; Schmidt, C.A.; Bezuidenhout, D. Bioprosthetic tissue preservation by filling with a poly(acrylamide) hydrogel. *Biomaterials* **2006**, *27* (9), 2123–2130.
 59. Moore, M.A.; Phillips, R.E. Biocompatibility and immunologic properties of pericardial tissue stabilized by dye-mediated photooxidation. *J. Heart Valve Dis.* **1997**, *6* (3), 307–315.
 60. Moore, M.A.; Bohachevsky, I.K.; Cheung, D.T.; Boyan, B.D.; Chen, W.M.; Bickers, R.R.; et al. Stabilization of pericardial tissue by dye-mediated photooxidation. *J. Biomed. Mater. Res.* **1994**, *28* (5), 611–618.
 61. Meuris, B.; Phillips, R.; Moore, M.A.; Flameng, W. Porcine stentless bioprostheses: Prevention of aortic wall calcification by dye-mediated photo-oxidation. *Artif. Organs* **2003**, *27* (6), 537–543.
 62. Schoen, F.J. Pathologic findings in explanted clinical bioprosthetic valves fabricated from photooxidized bovine pericardium. *J. Heart Valve Dis.* **1998**, *7* (2), 174–179.
 63. Jones, M.; Eidbo, E.E.; Hilbert, S.L.; Ferrans, V.J.; Clark, R.E. The effects of anticalcification treatments on bioprosthetic heart valves implanted in sheep. *ASAIO Trans.* **1988**, *34* (4), 1027–1030.
 64. Hirsch, D.; Drader, J.; Thomas, T.J.; Schoen, F.J.; Levy, J.T.; Levy, R.J. Inhibition of calcification of glutaraldehyde pretreated porcine aortic valve cusps with sodium dodecyl sulfate: Preincubation and controlled release studies. *J. Biomed. Mater. Res.* **1993**, *27* (12), 1477–1484.
 65. Cunanan, C.M.; Cabling, C.M.; Dinh, T.T.; Shen, S.H.; Tran-Hata, P.; Rutledge, J.H., III; et al. Tissue characterization and calcification potential of commercial bioprosthetic heart valves. *Ann. Thorac. Surg.* **2001**, *71* (5 Suppl.), S417–S421.
 66. Flameng, W.; Meuris, B.; Yperman, J.; De, V.G.; Herijgers, P.; Verbeke, E. Factors influencing

- calcification of cardiac bioprostheses in adolescent sheep. *J. Thorac. Cardiovasc. Surg.* **2006**, *132* (1), 89–98.
67. Jorge-Herrero, E.; Fernandez, P.; de la, T.N.; Escudero, C.; Garcia-Paez, J.M.; Bujan, J.; et al. Inhibition of the calcification of porcine valve tissue by selective lipid removal. *Biomaterials* **1994**, *15* (10), 815–820.
 68. Jorge-Herrero, E.; Fernandez, P.; Escudero, C.; de la Torre, N.; Zurita, M.; García Páez, J.M.; Castillo-Olivares, J.L. Influence of stress on calcification of delipidated bovine pericardial tissue employed in construction of cardiac valves. *J. Biomed. Mater. Res.* **1996**, *30* (3), 411–415.
 69. Vyavahare, N.R.; Hirsch, D.; Lerner, E.; Baskin, J.Z.; Zand, R.; Schoen, F.J.; et al. Prevention of calcification of glutaraldehyde-crosslinked porcine aortic cusps by ethanol preincubation: Mechanistic studies of protein structure and water–biomaterial relationships. *J. Biomed. Mater. Res.* **1998**, *40* (4), 577–585.
 70. Gott, J.P.; Pan, C.; Dorsey, L.M.; Jay, J.L.; Jett, G.K.; Schoen, F.J.; et al. Calcification of porcine valves: A successful new method of antimineralization. *Ann. Thorac. Surg.* **1992**, *53* (2), 207–215.
 71. Gott, J.P.; Girardot, M.N.; Girardot, J.M.; Hall, J.D.; Whitlark, J.D.; Horsley, W.S.; et al. Refinement of the alpha aminooleic acid bioprosthetic valve anticalcification technique. *Ann. Thorac. Surg.* **1997**, *64* (1), 50–58.
 72. Chen, W.; Kim, J.D.; Schoen, F.J.; Levy, R.J. Effect of 2-amino oleic acid exposure conditions on the inhibition of calcification of glutaraldehyde cross-linked porcine aortic valves. *J. Biomed. Mater. Res.* **1994**, *28* (12), 1485–1495.
 73. Webb, C.L.; Schoen, F.J.; Flowers, W.E.; Alfrey, A.C.; Horton, C.; Levy, R.J. Inhibition of mineralization of glutaraldehyde-pretreated bovine pericardium by AlCl₃. Mechanisms and comparisons with FeCl₃, LaCl₃, and Ga(NO₃)₃ in rat subdermal model studies. *Am. J. Pathol.* **1991**, *138* (4), 971–981.
 74. Levy, R.J.; Schoen, F.J.; Flowers, W.B.; Staelin, S.T. Initiation of mineralization in bioprosthetic heart valves: Studies of alkaline phosphatase activity and its inhibition by AlCl₃ or FeCl₃ preincubations. *J. Biomed. Mater. Res.* **1991**, *25* (8), 905–935.
 75. Clark, J.N.; Ogle, M.F.; Ashworth, P.; Bianco, R.W.; Levy, R.J. Prevention of calcification of bioprosthetic heart valve cusp and aortic wall with ethanol and aluminum chloride. *Ann. Thorac. Surg.* **2005**, *79* (3), 897–904.
 76. Ogle, M.F.; Kelly, S.J.; Bianco, R.W.; Levy, R.J. Calcification resistance with aluminum-ethanol treated porcine aortic valve bioprostheses in juvenile sheep. *Ann. Thorac. Surg.* **2003**, *75* (4), 1267–1273.
 77. Levy, R.J.; Schoen, F.J.; Lund, S.A.; Smith, M.S. Prevention of leaflet calcification of bioprosthetic heart valves with diphosphonate injection therapy. Experimental studies of optimal dosages and therapeutic durations. *J. Thorac. Cardiovasc. Surg.* **1987**, *94* (4), 551–557.
 78. Webb, C.L.; Benedict, J.J.; Schoen, F.J.; Linden, J.A.; Levy, R.J. Inhibition of bioprosthetic heart valve calcification with aminodiphosphonate covalently bound to residual aldehyde groups. *Ann. Thorac. Surg.* **1988**, *46* (3), 309–316.
 79. Isenburg, J.C.; Simionescu, D.T.; Vyavahare, N.R. Tannic acid treatment enhances biostability and reduces calcification of glutaraldehyde fixed aortic wall. *Biomaterials* **2005**, *26* (11), 1237–1245.
 80. O'Brien, M.F.; Goldstein, S.; Walsh, S.; Black, K.S.; Elkins, R.; Clarke, D. The SynerGraft valve: A new acellular (nonglutaraldehyde-fixed) tissue heart valve for autologous recellularization first experimental studies before clinical implantation. *Semin. Thorac. Cardiovasc. Surg.* **1999**, *11* (4 Suppl. 1), 194–200.
 81. Sharp, M.A.; Phillips, D.; Roberts, I.; Hands, L. A cautionary case: The SynerGraft vascular prosthesis. *Eur. J. Vasc. Endovasc. Surg.* **2004**, *27* (1), 42–44.
 82. Meuris, B.; Verbeken, E.; Flameng, W. Prevention of porcine aortic wall calcification by acellularization: Necessity for a non-glutaraldehyde-based fixation treatment. *J. Heart Valve Dis.* **2005**, *14* (3), 358–363.
 83. Walles, T.; Puschmann, C.; Haverich, A.; Mertsching, H. Acellular scaffold implantation—No alternative to tissue engineering. *Int. J. Artif. Organs* **2003**, *26* (3), 225–234.
 84. Neethling, W.M.; van Rijn, R.S.; Van Den Heever, J.J.; Hough, J.; Hodge, A.J. Evaluation of kangaroo aortic valved conduits in a juvenile sheep model: Preliminary findings. *J. Card. Surg.* **2001**, *16* (5), 392–399.
 85. Imai, Y.; Takanashi, Y.; Hoshino, S.; Nakata, S. The equine pericardial valved conduit and current strategies for pulmonary reconstruction. *Semin. Thorac. Cardiovasc. Surg.* **1995**, *7* (3), 157–161.
 86. Herijgers, P.; Ozaki, S.; Verbeken, E.; Van, L.A.; Meuris, B.; Lesaffre, E.; et al. Valved jugular vein segments for right ventricular outflow tract reconstruction in young sheep. *J. Thorac. Cardiovasc. Surg.* **2002**, *124* (4), 798–805.

Articular Cartilage Biomechanics

A

Jennifer S. Wayne

Department of Biomedical Engineering, Virginia Commonwealth University, Richmond, Virginia, U.S.A.

INTRODUCTION

The word “cartilage” refers to a specialized, soft connective tissue found in numerous sites within the body. There are three main types of cartilages with distinct features and behaviors for their intended function: fibrous, elastic, and hyaline.^[1] Each consists of a firm matrix made up of collagenous and elastic fibers and ground substance with differences among the types dependent on its specific composition and internal arrangement of the components. Fibrous cartilage is located in such structures as the discs between vertebrae in the spinal column (intervertebral discs), symphysis pubis, and the semilunar-shaped discs in the knee that mate the femur to the tibia (menisci). Elastic cartilage, with a predominance of elastic fibers in its matrix, is found in the auricle of the external ear, the auditory tube, the epiglottis, and portions of the larynx.

The third type of cartilage is hyaline, stemming from the Greek word “hyalos” meaning “glass.” The rings of the trachea and larynx contain this cartilage. A major hyaline cartilage is articular cartilage, found covering the ends of bones in diarthrodial (i.e., freely movable) joints such as the hip, knee, and shoulder. Articular cartilage is critical to normal joint function by distributing loads over the joint surfaces, acting as a shock absorber, and aiding in joint lubrication for nearly frictionless articulation. The composition and arrangement of components in the tissue enable it to perform its functions for seven and eight decades with little to no wear. However, in many individuals and for different reasons, the articular cartilage wears away from the joint surface over time in a progressive process called arthritis, leaving a degenerating articular surface with bone exposed. This is a very painful, debilitating condition often necessitating replacement of the joint with metal and plastic parts. Much study of articular cartilage has thus been and continues to be performed to understand its normal function and to develop reparative techniques that can be implemented before the degeneration warrants an artificial replacement.

This article describes the features of articular cartilage that are important to understand how it accomplishes its normal biomechanical function. The

composition of the tissue and the arrangement of its components are first described, followed by general biomechanical behaviors that it exhibits. Then, an overview of recent mathematical models is presented as these models help elucidate how articular cartilage functions.

COMPOSITION AND STRUCTURE

Articular cartilage is a translucent, glistening layer that is white and smooth. In the adult diarthrodial joint, it is avascular, aneural, and lacks lymphatic vessels. In humans, its thickness can range from one millimeter to several millimeters. During normal joint movement, this tissue is subjected to severe loads, often of magnitude several times one’s body weight, as one joint surface rolls and slides over the other, generating complex stresses (tension, compression, and shear) in cartilage.

A strong structure–function relationship exists in all biological tissues. The components of articular cartilage and the structural arrangement of those components enable the tissue to function normally and resist tension, compression, and shear stresses. The major components are the cells, water, collagen fibrils, proteoglycan aggregates (sometimes referred to as ground substance in other tissues), and charged, mobile ions.^[2] In comparison to other tissues, the cells of articular cartilage, the chondrocytes, are sparsely populated within the matrix, accounting for less than 10% of tissue volume. However, they are metabolically active, being responsible for the maintenance of the extracellular matrix. Articular cartilage is highly hydrated. Water, as the most abundant component in the tissue, accounts for 75–85% of its wet weight. Numerous electrolytes are freely mobile within the tissue as is the majority of the water.

The organic extracellular solid matrix is composed of an intertwining mesh of collagen fibrils (predominantly Type II collagen) and proteoglycans. Collagen, the ubiquitous protein responsible for many structural features of biological tissues, makes up 60–65% of cartilage on a dry weight basis. It is biomechanically similar to rope in that it can resist being stretched in tension but does not resist compression. If a

compressive load were applied to a collagen fibril, it would tend to buckle. Thus, the collagen fibril is the prime tensile resisting component of articular cartilage.

Proteoglycans are large, negatively charged, protein-polysaccharide molecules that comprise 10–15% of articular cartilage on a dry weight basis. The immobilized negative charges, the concentration of which is the tissue's fixed charge density (FCD), repel one another and therefore try to distend the proteoglycan molecule, to separate the charges as much as possible. If the negative charges are forced into closer proximity, as occurs in compression, the response is an increased repulsion force. The proteoglycans are thus the prime compression-resisting component of articular cartilage.

The collagen network does not permit the proteoglycan molecule to expand fully but keeps it in a compressed state. Thus, the collagen network in intact articular cartilage is in a state of tension. The inherent state of compression of the proteoglycans leads the tissue to possess a swelling pressure, or propensity to swell. The magnitude of the swelling pressure can be reduced or increased by changing the concentration of ions in the solution bathing the tissue. Physiologic solution is salty, and thus cartilage is bathed in a solution where positive ions can freely move in the tissue. Some of these positive ions shield the negative charges on the proteoglycan molecule. If the solution is made hypertonic (more freely mobile ions), more negative charges are shielded by the positive, and thus the desire to swell decreases. On the contrary, if the solution is made hypotonic (fewer freely mobile ions), greater density of negative charges is left unshielded, and the desire to swell increases.

Together, the collagen and proteoglycans form an interwoven solid matrix that permits the tissue to resist shear stresses as well as tension and compression. Thus, cartilage can be viewed physically akin to a sponge, albeit with pore sizes on the order of nanometers. Water can flow within and out of the interstices of the collagen/proteoglycan matrix when the tissue is subjected to a load.

With knowledge of the components of articular cartilage, the structural organization of those components is important to further understand the tissue's function. While the articular cartilage layer covering

the bone surface is thin, it has a complex organization through its thickness. The layer can be structurally grouped into four zones: superficial tangential zone, middle/transitional zone, deep/radial zone, and calcified zone (Fig. 1). This grouping is based on orientation of the collagen fibrils and cell morphology within the zone. The superficial tangential zone typically comprises 10–20% of the thickness and is the zone that comes in to contact with the cartilage surface on the opposing bone. The collagen fibrils are aligned parallel to the articular surface in this zone and are tightly woven, with interposed chondrocytes oblong in shape. The amount of proteoglycans (content) is the lowest in this zone.

Within the middle zone, the collagen fibrils are transitioning from the parallel arrangement in the superficial tangential zone to the perpendicular orientation in the deep zone. Here, the fibrils can be viewed as having a random orientation but are not as dense as in the superficial tangential zone. The chondrocytes have a more rounded shape and are randomly distributed. The deep zone has collagen fibrils in a perpendicular orientation, crossing the tidemark of uncalcified to calcified cartilage and anchor the tissue to underlying bone. The chondrocytes maintain the rounded shape as in the middle zone but are now arranged in columns. The final zone is that of calcified cartilage where the tissue is transitioning from the uncalcified soft articular cartilage to the calcified hard bone.

BASIC BIOMECHANICAL BEHAVIOR

Stress and strain are two concepts that are important to describe the mechanical behavior of materials and structures. Stress is defined as the force or load on a material divided by the cross-sectional area over which that load is distributed. Strain is a unitless measure of shape changes and can be calculated as deformation divided by the original length of the material. For example, when tension is applied to a rope of a given length (L_0), it will stretch, and the strain it experiences can be calculated by the amount of stretch divided by L_0 . Shear stress induces shear strain, which is a

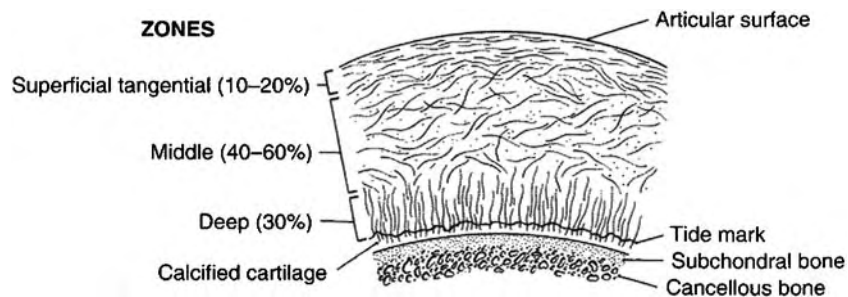


Fig. 1 Structural organization of collagen Type II fibrils within a cartilage layer. (From Ref. [3].)

measure of how an angle changes within a material, for example from 90° .

In general, biological soft tissues exhibit viscoelastic features.^[4] Viscoelasticity is a phenomenon that is the combination of viscous (dashpot-like) and elastic (spring-like) behaviors. A viscous substance is one that flows (continues to deform with time) under the action of an applied load. An elastic material deforms its shape to a certain extent under the action of an applied load but returns to its original shape once the load is removed. A viscoelastic solid therefore deforms when under load but does so in a time-dependent fashion. Three important characteristics of viscoelasticity are creep, stress relaxation, and hysteresis behaviors (Fig. 2).

In creep, one applies (controls) a specific load (or stress) to the material and monitors how the material changes shape (deforms or strains). If the material is elastic, an immediate jump in strain is exhibited for an applied jump in stress. The strain then remains constant over time as long as the applied stress does not change. Viscoelastic solids also tend to exhibit an immediate jump in strain for an applied jump in stress; however, the strain does not remain constant but exponentially rises until a plateau is reached at some later time, while the applied stress remains at a constant level.

In stress relaxation, one applies (controls) a specific deformation (or strain) to the material and monitors how much load (or stress) the materials responds with. If the material is elastic, an immediate jump in stress is exhibited for an applied jump in strain. The stress then remains constant over time as long as the applied strain does not change. Viscoelastic solids also exhibit an immediate jump in stress for an applied jump in strain; however, the stress does not remain constant but exponentially decreases until a nonzero plateau is reached at some later time, while the applied strain remains at a constant level.

In hysteresis, a stress or strain is applied to a material, most commonly in a linear ramp fashion, and then removed in a similar fashion. The loading and unloading curves are identical for an elastic material (in this case, a linear material). A viscoelastic solid however exhibits a different curve in loading vs. unloading. The loading curve is always above the unloading curve to indicate that there are viscous losses associated with the loading/unloading.

Biological soft tissues also present a nonlinear stress–strain response (Fig. 3). The onset of loading is termed the “toe region,” in which the material begins to resist the load by increasing its modulus. At a subsequent point, the curve becomes linear and remains so until failure commences in the matrix of the tissue. The slope of the curve, at any given strain value, provides a measure of the modulus or stiffness of the

tissue at that strain. The slope in the linear region is typically called the tangent modulus. A collagenous tissue subjected to a tensile test exhibits this nonlinear stress–strain behavior. The three regions of the curve can be explained by what is occurring to the collagen during the tensile loading. At the onset of stretching, the collagen fibrils are crimped or wavy such that the initial resistance (modulus) offered to the stretch results from straightening of this crimp, and a gradual rise in response is seen in the toe region. Once the crimp has been removed from the collagen, further tensile loading stretches the collagen itself, giving the highest modulus the tissue has to offer in the linear region. At a later point, the collagen fibrils begin to fail, as the slope of the stress begins to reduce in the failure region, until the tissue has been torn into two.

Another phenomenon that is important in understanding biological tissue behavior is anisotropy.^[5] The engineering term “anisotropy” depicts how materials respond when a load is applied at a particular point from different coordinate directions, for example in an XYZ rectangular Cartesian coordinate system. A material is said to be “isotropic” if the mechanical behavior, for example modulus, is the same regardless of whether a load is applied in the X, Y, or Z directions. If the behavior is dependent on the direction in which the load is applied, the material is said to be “anisotropic.”

There are different levels of anisotropy, with three commonly implemented in descriptions of biological materials. An isotropic material, as indicated, exhibits the same regardless of which direction the stress is applied. If the response along one direction results in different mechanical behavior than the other two directions, the material is said to be “transversely isotropic.” If the behavior differs along each coordinate direction, the material is called “orthotropic.” Reinforced concrete, i.e., rebar, would be considered transversely isotropic with the greatest tensile resistance (modulus) in the direction of the reinforcing bars. Composite materials can also be isotropic or anisotropic, depending on how the fibers are laid down within the matrix.

Articular cartilage exhibits the phenomena just described, creep, stress-relaxation, nonlinear stress–strain curve, and different degrees of anisotropy in various types of testing conditions. Because articular cartilage endures complex stresses (tension, compression, and shear) during its normal function, numerous experiments have been performed in tension, compression, and shear (Fig. 4) to determine its biomechanical behavior and how this behavior enables the function. Experimental data is often curve fit with a particular mathematical model to determine the properties of the test tissue. Specimens of articular cartilage have been tested either in situ (i.e., an unviolated articular

A

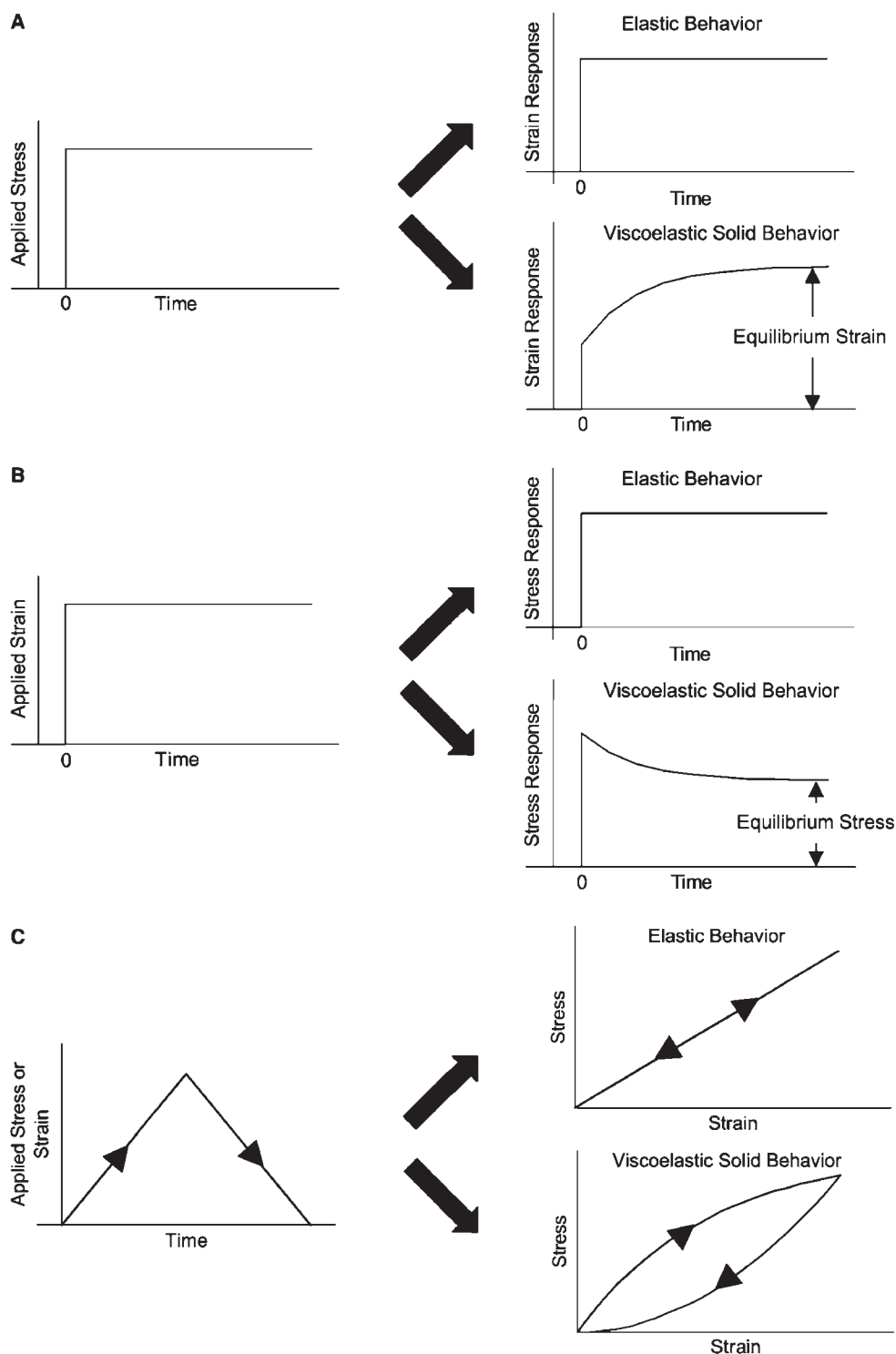


Fig. 2 Depiction of viscoelastic creep, stress relaxation, and hysteresis. In creep (A), the load or stress is controlled (here, applied instantaneously and then held constant) while the response of deformation or strain changes with time until an equilibrium (steady-state) is reached. In stress relaxation (B), the deformation or strain is controlled (here, applied instantaneously and then held constant) while the response of load or stress changes with time until an equilibrium is reached. In contrast, an elastic material does not exhibit a time-dependent response for a constant applied stress or strain. In hysteresis (C), a stress or strain is applied to a material, most commonly in a linear ramp fashion, and then removed. A viscoelastic material exhibits different loading and unloading curves, with the loading curve always residing on top, whereas the curves are identical in an elastic material (in this case, linear).

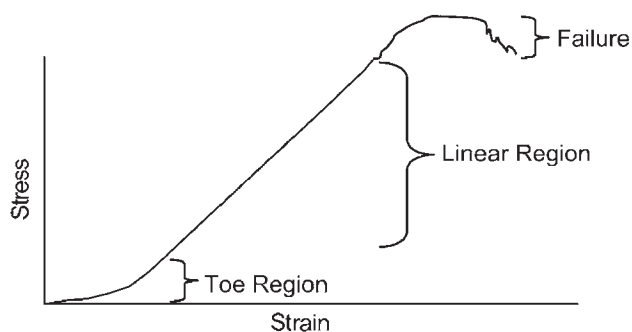


Fig. 3 Nonlinear stress–strain behavior typical of biological soft tissues.

surface) or as a portion removed from an articular surface (explants). Explants may also be full thickness (all zones) or sectioned into a specific zone to uncover zonal behavior.

As articular cartilage experiences significant compressive stresses, evaluating its response in compression is of great importance. There are three common types of compression tests performed on an articular cartilage—indentation, confined compression, and unconfined compression. Indentation tests are typically carried out on cartilage intact with underlying bone. A stress (or strain) is applied to the articular surface via a small indenter and the resulting response, strain (or stress), is monitored over time. The creep (or stress-relaxation) is then observed. The indenter may be flat-ended or spherical, porous, or impermeable. The stress and strain fields that result in the tissue under indentation loading are quite complex.

Both confined compression and unconfined compression tests require preparation of a cartilage explant, usually in the shape of a cylindrical plug. In confined compression, the tissue may remain intact with the underlying bone or the bone may be removed, whereas in an unconfined compression, the bone is typically removed. For the confined compression test, the cartilage specimen is placed within a confining chamber, with radial dimensions to match that of the cylindrical plug explant such that a snug fit is achieved. A stress (or strain) is then applied to the entire surface area of the explant via a porous, permeable platen and the strain (or stress) response is monitored. In this testing configuration, deformation may only occur in one direction, that of the applied stress, because the confining chamber restricts any radial/lateral expansion. This test and the strains that develop are considered to be one-dimensional, in the direction of the applied stress. Unconfined compression is similar to confined compression in that a stress (or strain) is applied over the entire surface area of the explant. However, the compression platen is made impermeable, and radial/lateral expansion is permitted. Tissue

strain in an unconfined compression test is both axial (along the direction of applied stress) and radial.

Compressive stresses applied to articular surfaces also generate tensile stresses in cartilage owing to the Poisson effect so that tensile tests are performed to determine tissue behavior in this mode of loading. Explants must be cut from the surface to create strips of cartilage to which uniaxial tension may be applied. The nonlinear stress–strain curve depicted in Fig. 3 is exhibited by biological soft tissues subjected to a tensile test. The shear test configuration differs from either the compression or the tension tests just described. In the latter, the stress is applied normal (perpendicular) to the surface area of the specimen, whereas in shear, the stress is applied parallel to the surface. Cartilage explants are removed from the joint, often keeping the bone intact to serve as a way to rigidly hold the specimen during the shear test. Shear stresses are typically induced to the cartilage explant by either a linear or a rotational (torsional) displacement applied to the surface. Creep and stress–relaxation may be exhibited in shear, but the magnitude of the response differs substantially from the axial modes of loading (compression, tension). In general, materials have lower modulus and strength in shear than in axial loading.

The viscoelastic response of articular cartilage found from these different types of experimental tests is important to its function. It helps the tissue absorb rapid loading, like a shock absorber, while controlling the amount of strain experienced. Two primary sources cause the viscoelastic behavior within the cartilage. Fluid movement within cartilage induces a frictional drag on the collagen/proteoglycan solid matrix and thus the flow through cartilage's matrix occurs in a time-dependent fashion.^[6] The collagen/proteoglycan matrix has inherent viscoelastic characteristics as well,^[7] resulting from the internal friction as the components slide relative to one another and each other,^[4] particularly the proteoglycans.

The three uncalcified zones in articular cartilage exhibit different properties. This inhomogeneity (differing from point to point, or zone to zone) results from the different composition and organization found in the zones. Each zone may also yield a different response to applied loads in different directions (at the same point), particularly in tension owing to the organization of the collagen fibrils.^[8] This organization results in the tissue exhibiting transversely isotropic or orthotropic behavior in the superficial tangential zone. The preferred direction, i.e., the direction having the different behavior, is the direction along which the collagen fibrils are oriented. When a tensile stress is applied along the fibril direction, a greater resistance to the stress is exhibited. The tensile tangent modulus in this direction has been estimated to be 5–20 MPa,

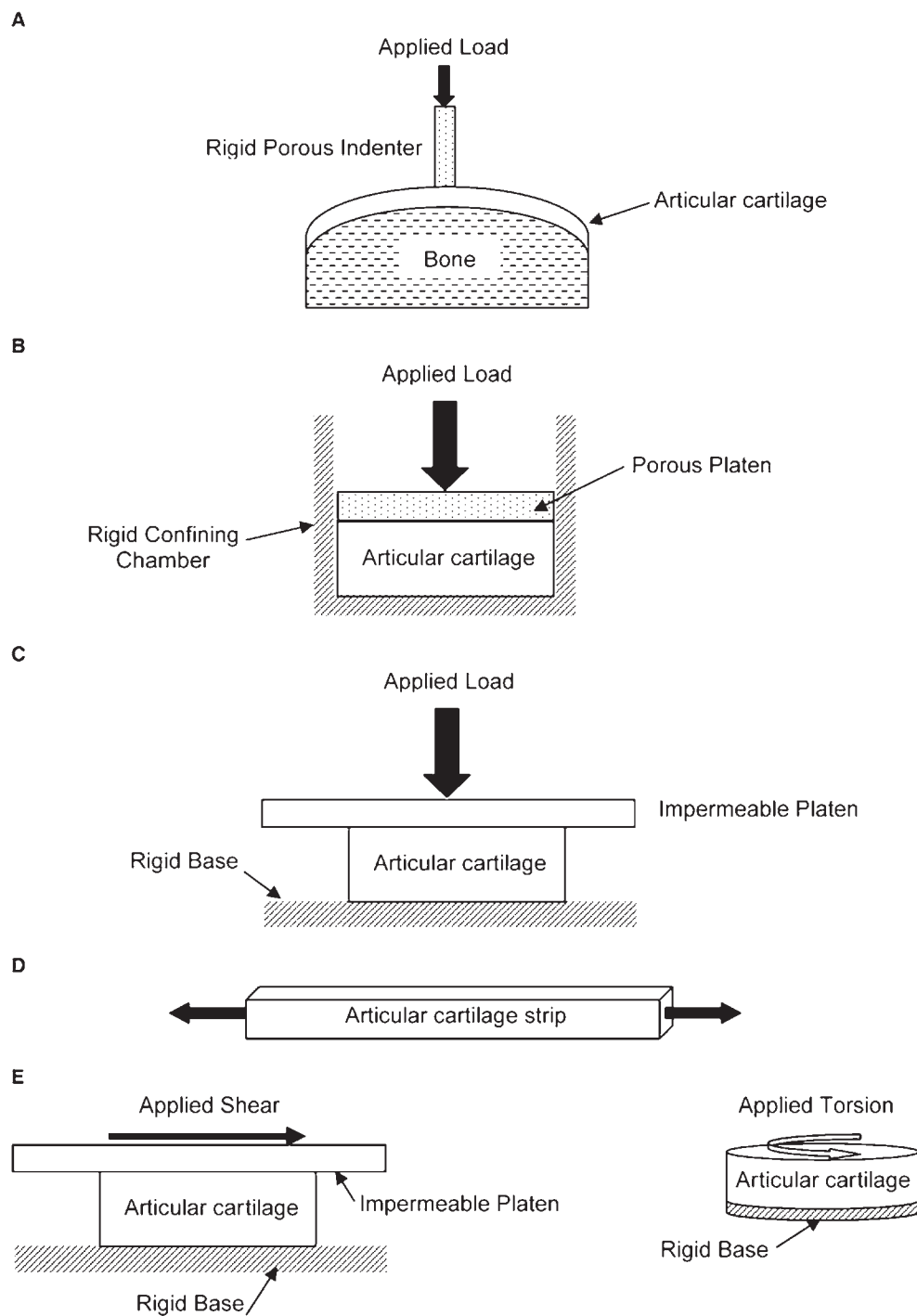


Fig. 4 Common experimental testing configurations to assess cartilage behavior in the laboratory: (A) indentation configuration; (B) confined compression configuration; (C) unconfined compression configuration; (D) tensile test; and (E) shear test, as induced by a linear (left image) or rotational (right image) displacement.

whereas the modulus perpendicular to this direction is significantly lower. In the middle zone, the behavior is more isotropic, with a tensile modulus in the order of 5 MPa. The deep zone could be considered transversely isotropic because of the radially oriented collagen fibrils when subjected to tensile loads, with a modulus in the order of 1 MPa.

In compression, each zone can also exhibit a degree of anisotropy but to a much lesser extent than when subjected to tensile stress.^[9] This results from the proteoglycans acting as the main compression-resisting component, with more isotropic behavior than the collagen acting in tension. There is the possibility that the isotropic middle zone can begin exhibiting more

transversely isotropic behavior as the tissue is compressed from the surface, because the compression redirects the once random orientation of the collagen fibrils to begin aligning more parallel to the articular surface and offering more resistance to further radial expansion. Compressive moduli have been reported in the order of 0.4–0.8 MPa.^[6]

Properties of articular cartilage can vary from joint to joint in a given individual, vary from species to species in a given joint, as well as vary from location to location in a given joint. Some of the differences can be explained by the magnitude of the stresses that the tissue must be capable of resisting, from a structure–function point of view. Certain joints are more highly loaded than others because they bear a great percentage of body weight, for example, the hip and knee in the human. Size/weight of the species also affects the magnitude of stress experienced by the articular cartilage. Finally, joint mechanics dictates that certain locations on an articular surface in a given joint are weight bearing while others are nonweight bearing.

MATHEMATICAL MODELS

A strong structure–function relationship exists in biological tissues where the structure (i.e., types and amounts of components and organization of those components) is best suited for the specific function, and vice-versa, the intended function requires a specific structure. This is akin to an engineer designing a physical structure for a particular application and selects components (I-beams, columns, trusses) made of appropriate materials with appropriate properties (i.e., modulus, strength) to resist the stresses that the structure is expected to experience. The complex stresses (tension, compression, and shear) that articular cartilage endures during its normal function requires that certain components are present and that these components are organized in a particular fashion. For example, articular cartilage is subjected to large magnitude compressive stresses; the biological component for soft tissues that resists compressive stresses is the proteoglycan aggregate. Thus, cartilage has a high proteoglycan content. Additionally, fluid movement is believed critical to normal function, and thus viewing the tissue as porous and permeable aids in understanding function.

Knowledge of articular cartilage behavior and function can come from experimental studies as described previously. More in depth knowledge of how the behavior underlies the function is best realized through implementation of mathematical models. Numerous models have been applied to articular cartilage biomechanics, including solid elasticity (considering the tissue to be one phase, a solid that behaves

elastically), solid viscoelasticity (considering the tissue again to be one phase, a solid that has viscoelastic characteristics), and multiphase models. Multiphase models view the tissue as consisting of distinct phases, typically some combination of water, solid (collagen and proteoglycan), and electrolytes, and mathematically modeling each phase individually as well as the interaction among the phases to describe overall tissue behavior and function.

The biphasic theory^[6] considers articular cartilage to be a mixture of two phases—a porous, permeable solid phase (the collagen/proteoglycan solid matrix acting as an elastic material) and an interstitial fluid phase (water as an ideal fluid). This mixture modeling approach has been extremely helpful in depicting the role that fluid movement plays in cartilage behavior and in understanding the tissue’s function on an articular surface. In this theory, the viscoelastic response of cartilage results from the frictional drag of the fluid on the solid matrix as it flows within the tissue. Both the solid and the fluid phases share the support of stresses experienced by the tissue. The elastic stress in the solid arises from strains induced on the matrix. Fluid stress arises from pressurizing the water in the tissue. A gradient in fluid pressure induces fluid to flow within, into, and out of the tissue.

As a mixture of elastic solid and fluid, the time-dependent, viscoelastic behavior of articular cartilage in compression can be explained in the following manner. Under the applied jump in stress as seen in creep (Fig. 2A), the solid may deform elastically via the Poisson effect (which is a stress applied in one direction to strain a material in the same direction that results in a concomitant strain in the other two directions) to yield the immediate jump in strain, but the majority of the applied stress is initially borne by pressurizing the fluid phase within the tissue. A pressure gradient therefore exists between the high fluid pressure in the tissue and the low fluid pressure outside of the tissue such that fluid flow within and out of the cartilage will occur. As fluid leaves the tissue, the solid matrix compresses (i.e., experiences strain), increasing the solid stress while at the same time reducing the fluid pressure. This increase in solid stress while fluid stress is decreasing can be viewed as stress transfer then between the phases, from fluid to solid. Stress transfer continues until fluid flow ceases (pressure gradient is zero), and all the applied stress is borne by the solid matrix. Equilibrium is then said to have occurred, and the creep curve plateaus.

Similarly, fluid flow is responsible for the observed stress relaxation behavior under the action of an applied compressive strain, as depicted in Fig. 2B. The imposed strain compresses the solid matrix and forces fluid to flow out of the tissue, resulting in an immediate high resistance (large stress response). No more fluid

leaves the tissue after this initial strain because the amount of compression remains constant over time. However, the subsequent stress relaxing (reducing) with time is the result of fluid redistribution within the tissue such that a uniform strain exists throughout the tissue when equilibrium is reached. At equilibrium, fluid pressure is zero, and the nonzero equilibrium stress is supported by the strain generated in the solid matrix.

The coupling of fluid to solid behavior in a mixture model requires an additional property to be specified to govern the interaction between the two phases. The solid matrix is again characterized by a property, elastic modulus, as described previously. The flow of fluid through the permeable solid is governed by a property called permeability. This property captures the resistance of the solid/fluid coupling to fluid flow. A high permeability means that fluid can easily flow in the tissue (low frictional drag), whereas a low permeability indicates that flow is difficult to occur (high frictional drag).

Permeability can be inhomogeneous (differing from point to point) as well as anisotropic (differing in different directions at the same point). It is also strain dependent, that is, it is not a constant but decreases in value as compressive strain occurs in the tissue. This decrease with compaction/compression is straightforward to consider; as the tissue compacts, the pores through which the fluid must flow become smaller. Additionally, the FCD of cartilage increases when the tissue is compacted (same number of negative charges on the proteoglycans now occupying a smaller volume), thus increasingly retarding the movement of the positive ions in the fluid flowing out. Higher friction therefore develops from these two phenomena as fluid flows out of the tissue, leading to the decreased permeability. Normal permeability values for articular cartilage are in the order of $1-5 \times 10^{-15} \text{ m}^4/\text{N sec}$.

Use of the biphasic theory to describe cartilage behavior has led to hypotheses about how cartilage fulfills its functions and resists the demanding stress environment in which it must operate without failing. One of the tissue's main functions is to distribute loads over the joint surface. Distribution of a given applied load over a joint surface is dependent on both area of contact and how stress is transmitted from the articular surface down to the bone. The fact that the tissue deforms under load that increases its area of contact and the congruency between joint surfaces helps distribute the load over the surface. Additionally, the stress of fluid pressurization owing to the applied load, as predicted by theory, can be transmitted through the tissue to an underlying bone over a larger region than that because of contact alone, particularly at early times of loading when tissue deformation is limited.

Consideration of the solid/fluid interaction permits a hypothesis into the longevity of cartilage. The ability of the fluid phase to withstand stresses implies that the solid matrix experiences less stress as the two phases share in supporting the applied stresses. This helps to protect the solid matrix from experiencing excessive stresses and strains that can result in failure. The longer that the fluid remains pressurized, the longer the solid stress and strain remain low. Cartilage's inherent low permeability helps to maintain fluid pressurization. The superficial tangential zone also helps to maintain fluid pressurization longer; the tight weave of collagen fibrils in this anisotropic zone results in an inherently low permeability. The lower proteoglycan content of the zone makes it easier to compress, which can further reduce the low permeability of this zone. Thus, the fluid is hindered from flowing out of the tissue through the superficial tangential zone, thus maintaining a higher fluid pressurization for longer times.

Articular cartilage also helps in lubrication within diarthrodial joints, which increases the tissue's longevity by minimizing wear. Diarthrodial joints are referred to as synovial joints because a capsule, lined with a synovial membrane, encapsulates the joint. The synovial membrane secretes a viscous substance called synovial fluid that consists of a gel of proteins, electrolytes, and water. Sophisticated lubrication mechanisms are at work in these joints,^[10] which do not permit elaboration of them here. However, responses of articular cartilage to applied stresses may enhance joint lubrication. The tissue deforms over time, increasing contact area and congruency between the joint surfaces, which may maintain a layer of the lubricant between opposing articular surfaces to separate the surfaces for longer periods of time. This is referred to as "elastohydrodynamic" ("squeeze-film") lubrication. Fluid flow out of the tissue from interstitial fluid pressurization may increase the amount of lubricant available for articulation; this is often referred to as "weeping" lubrication. Some areas of the joint surface may permit fluid flow into the articular cartilage, creating an ultrafiltration of the synovial fluid. This concentrated layer of the synovial fluid gel with increased viscosity lubricating the surfaces is termed as "boosted" lubrication. A pressurized layer of fluid separating articular surfaces diminishes solid-to-solid contact and thus friction is also reduced. Finally, there is evidence that a specific molecule, "lubricin," is adsorbed onto the articular surface to act as a bound lubricant.

The biphasic theory has been implemented into a computational tool, called finite element analysis, which has permitted more extensive investigations into the functional behavior of articular cartilage.^[11,12] Recently, a similar, poroelastic model has been made available in a commercially available finite element software package.^[13]

Extension of the mixture approach to mathematical modeling has been performed to further define the role of other cartilage features to tissue behavior. In addition to fluid flow as the source of viscoelastic response, the solid matrix has been extended from linearly elastic to viscoelastic to yield a biphasic poroviscoelastic model.^[14] With this, the relative impact of the two main sources of viscoelasticity in cartilage can be observed. The different responses (i.e., modulus) in tension (owing to collagen fibrils) vs. compression (owing to the proteoglycans) have given rise to a modeling of this tension/compression nonlinearity in the solid matrix via a theory called “conewise” linear elasticity.^[15] A different approach to modeling the tension/compression nonlinearity is to consider the collagen fibrils as a distinct solid from the proteoglycans, i.e., two solid phases, rather than the collagen/proteoglycan matrix as one homogeneous solid. This has been implemented in a fibril-network reinforced biphasic model^[16] whereby collagen is simulated as linear springs to resist tension, with a collagen modulus, while the distinct proteoglycans resist the compression, with its own modulus. One of the results is that the stresses experienced by different components (collagen vs. proteoglycan) can be different, even though the strains are the same. A recent modification is to relax the linear assumption of collagen in tension and permit a nonlinear stress–strain relationship to simulate collagen behavior. This is achieved by incorporating a strain dependence on the collagen modulus.^[17] Finally, incorporating a third phase, an ion species, has given rise to the triphasic theory whereby the importance of the FCD, swelling pressure, and flow of ions (resulting in a streaming potential as fluid/ions flow out of cartilage) can be described.^[18] A different approach to depicting the influence of ion flow has yielded an electrokinetic model.^[19]

The mathematical models have grown increasingly complex with the expectation that each modification more appropriately reflects the complex behavior of articular cartilage. With this is the need to accurately reflect the properties used in the models, from inhomogeneities to anisotropy to strain dependence. Additionally, certain features may be more influential in governing behavior in certain modes of loading than in others. It is thus important to fully understand the assumptions inherent in different models so that conclusions can be accurately drawn.

CONCLUSIONS

This article has provided a basis for understanding the biomechanics of articular cartilage and how it functions in a diarthrodial joint. Several book chapters and numerous journal publications have been written with

much details and can be used for further reference (see for example, Ref. [10]). Articular cartilage continues to challenge the scientific community to continue developing new ways to more accurately reflect behavior and discovering more in depth as to how it accomplishes its function. Understanding normal articular cartilage biomechanics then has application to determine the changes that occur when degeneration begins on an articular surface, why it is progressive, and what methods may be used to arrest the progression with repair techniques.

ARTICLES OF FURTHER INTEREST

Cartilage Regeneration; Chondroitin Sulfates; Glycosaminoglycans and Proteoglycans; Hip Biomechanics; Knee Joint: Overuse Injuries; Tissue Engineering of Cartilage; Wear Debris from Joint Arthroplasties, Biological Effects Of

REFERENCES

1. Jacob, S.W.; Francone, C.A.; Lossow, W.J. *Structure and Function in Man*; W.B. Saunders Company: Philadelphia, 1978.
2. Maroudas, A. Physicochemical properties of articular cartilage. In *Adult Articular Cartilage*; Freeman, M.A.R., Ed.; Pitman Medical: England, 1979; 215–290.
3. Mow, V.C.; Lai, W.M. Some surface characteristics of articular cartilage. I. A scanning electron microscopy study and a theoretical model for the dynamic interaction of synovial fluid and articular cartilage. *J. Biomech.* **1974**, *7* (5), 449–456.
4. Fung, Y.C. *Biomechanics—Mechanical Properties of Living Tissues*, 2nd Ed.; Springer-Verlag: New York, 1993.
5. Spencer, A.J.M. *Continuum Mechanics*; Longman Scientific & Technical: England, 1980.
6. Mow, V.C.; Kuei, S.C.; Lai, W.M.; Armstrong, C.G. Biphasic creep and stress relaxation of articular cartilage in compression: theory and experiments. *J. Biomech. Eng.* **1980**, *102*, 73–84.
7. Hayes, W.C.; Bodine, A.J. Flow-independent viscoelastic properties of articular cartilage matrix. *J. Biomech.* **1978**, *11* (8–9), 407–419.
8. Woo, S.L.-Y.; Akeson, W.H.; Jemcott, G.F. Measurements of nonhomogeneous directional mechanical properties of articular cartilage in tension. *J. Biomech.* **1976**, *9*, 785–791.
9. Chen, A.C.; Bae, W.C.; Schinagl, R.M.; Sah, R.L. Depth- and strain-dependent mechanical and electromechanical properties of full-thickness bovine articular cartilage in confined compression. *J. Biomech.* **2001**, *34* (1), 1–12.
10. Mow, V.C.; Hung, C.T. Biomechanics of articular cartilage. In *Basic Biomechanics of the Musculoskeletal*

- System*, 3rd Ed.; Nordin, M., Frankel, V.H., Eds.; Lippincott Williams & Wilkins: Maryland, 2001; 60–100.
11. Wayne, J.S.; Woo, S.L.-Y.; Kwan, M.K. Application of the U-P finite element method to the study of articular cartilage. *J. Biomech. Eng.* **1991**, *113*, 397–403.
 12. Spilker, R.L.; Suh, J.K.; Mow, V.C. Effects of friction on the unconfined compressive response of articular cartilage: a finite element analysis. *J. Biomech. Eng.* **1990**, *112* (2), 138–146.
 13. Wu, J.; Herzog, W.; Epstein, M. Evaluation of the finite element software ABAQUS for biomechanical modeling of biphasic tissues. *J. Biomech.* **1998**, *31*, 165–169.
 14. Mak, A.F. The apparent viscoelastic behavior of articular cartilage—the contributions from the intrinsic matrix viscoelasticity and interstitial fluid flows. *J. Biomech. Eng.* **1986**, *108* (2), 123–130.
 15. Soltz, M.A.; Ateshian, G.A. A conewise linear elasticity mixture model for the analysis of tension-compression nonlinearity in articular cartilage. *J. Biomech. Eng.* **2000**, *122* (6), 576–586.
 16. Soulhat, J.; Buschmann, M.D.; Shirazi-Adl, A. A fibril-network-reinforced biphasic model of cartilage in unconfined compression. *J. Biomech. Eng.* **1999**, *121* (3), 340–347.
 17. Wilson, W.; van Donkelaar, C.C.; van Rietbergen, C.; Ito, K.; Huijskes, R. Stresses in the local collagen network of articular cartilage: a poroviscoelastic fibril-reinforced finite element study. *J. Biomech.* **2004**, *37* (3), 357–366.
 18. Lai, W.M.; Hou, J.S.; Mow, V.C. A triphasic theory for the swelling and deformation behaviors of articular cartilage. *J. Biomech. Eng.* **1991**, *113* (3), 245–258.
 19. Frank, E.H.; Grodzinsky, A.J. Cartilage electro-mechanics—II. A continuum model of cartilage electrokinetics and correlation with experiments. *J. Biomech.* **1987**, *20* (6), 629–639.

Artificial Heart Fluid Dynamics: Positive Displacement Pumps

Steven Deutsch

Department of Bioengineering, The Pennsylvania State University, University Park, Pennsylvania, U.S.A.

John M. Tarbell

Department of Biomedical Engineering, The City College of New York, New York, New York, U.S.A.

Keefe B. Manning

Department of Bioengineering, The Pennsylvania State University, University Park, Pennsylvania, U.S.A.

Gerson Rosenberg

Penn State Milton S. Hershey Medical Center, Penn State College of Medicine, Hershey, Pennsylvania, U.S.A.

Arnold A. Fontaine

Department of Bioengineering, The Pennsylvania State University, University Park, Pennsylvania, U.S.A.

INTRODUCTION

Although the use of mechanical circulatory support was postulated as early as 1812 by Legallois,^[1] it was not until 1961 that the first clinical left heart bypass was performed by Hall et al.^[2] The first artificial heart was implanted by Cooley^[3] in 1969. Thrombus formation and hemolysis are fundamental problems which persist in spite of the use of anticoagulant and platelet-inhibiting agents. In the early 1970s, Phillips et al.^[4] began the use of visualization techniques to study the fluid motion in transparent models of blood pumps. Changes in blood pump geometry and valve type and geometry were shown to influence thrombus formation. Techniques such as hot film anemometry, laser Doppler anemometry, and particle image velocimetry (PIV) are now a routine part of the development of artificial heart pumps. We note that the current review is an abbreviated version of “Experimental Fluid Mechanics of Pulsatile Artificial Blood Pumps,” written by the same authors in the Annual Review of Fluid Mechanics^[5] (Permission granted from Annual Reviews, Deutsch et al. 2006).

General Description

The LionHeart™ Left Ventricular Assist System, shown in Fig. 1, illustrates one end product of experimentation discussed here. In pulsatile pumps, the flow is driven either pneumatically or by a pusher plate against a segmented polyurethane blood sac. Where measurement access to the ventricle is required, the blood sac is

replaced by a diaphragm of the same material so that the interior of the model is exposed. This is a good representation of pusher plate devices, where only the pusher plate side of the sac moves. Generally, the device is cylindrical, with ports for the inlet and outlet artificial heart valves that are joined tangentially to the body. For an adult device under physiologic conditions, the mean aortic (outlet) pressure is 100 mm Hg (120/80), the mean atrial (inlet) pressure is 10 mm Hg (20/0), and the cardiac output is 5 l/min. The beat rate is 72 beats/min (bpm) and the percentage of the cycle in outlet flow (systolic duration) is 30%–50%. Physiological conditions can vary widely and automatic control of the pump cycle is normally through monitoring of the end diastolic volume, diastole being the filling portion of the cycle. Mehta et al.^[6] provides a description of a typical, fully implantable device.

Rosenberg et al.^[7] describe a mock circulatory loop for testing blood pumps. Inlet and outlet compliance chambers simulate the atrial and aortic compliance of the native cardiovascular system while a parallel plate resistor downstream of the aortic compliance simulates the systemic resistance of the circulation. A reservoir between the systemic resistance and atrial compliance controls the preload to the pump. Pressure waveforms are measured in the compliance chambers and flow waveforms at the inlet and outlet ports. The variable compliance and resistance are used to set the fixed flow conditions. Beat rate and systolic duration are also set through an appropriate drive system.

Blood is a shear-thinning, viscoelastic fluid^[8] that is often taken as Newtonian at sufficiently high shear

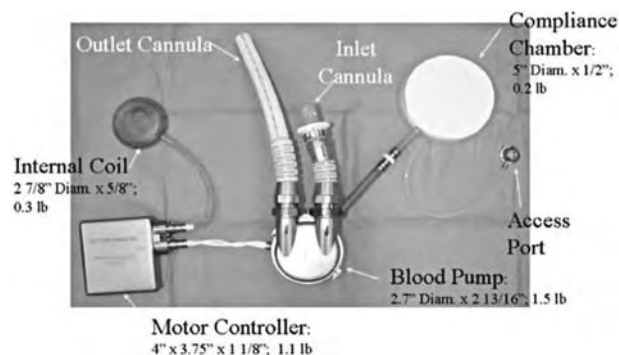


Fig. 1 The LionHeart™ left ventricular assist system.

rates (above 500 s^{-1}). The hematocrit (relative volume of red blood cells) greatly affects the magnitude and relative importance of the viscous and elastic components of the complex viscosity.^[9] The high shear rate kinematic viscosity asymptote for normal hematocrit blood (40%) is about 3.5 centistokes (cs) and solutions of glycerin and water (40/60) or mineral oils are often taken as blood analogs.^[10] Optical access to the fluid for velocity measurements can be important and Baldwin et al.^[11] among others, used a solution of 79% saturated aqueous sodium iodide, 20% pure glycerol, and 1% water by volume to produce a fluid with a kinematic viscosity of 3.8 cs and an index of refraction (matching Plexiglas) of 1.49 at 25°C . Brookshier and Tarbell^[12] and Long et al.^[13] developed Xanthan gum/glycerin solutions that simulate blood viscoelasticity well; sodium iodide may be added to adjust the index of refraction.

Heart valves, which maintain unidirectional flow, play a major role in the mechanical environment of the artificial heart. They are generally chosen for their durability. With the device shown in Fig. 1, Bjork-Shiley tilting disc valves were used. Mechanical heart valves (MHVs) are not specifically designed for mechanical blood pumps, and their efficiency can be compromised. Yoganathan et al.^[14] gives a survey of MHVs and their fluid mechanics.

Hemolysis—the destruction of red blood cells, and thrombosis—clot formation, must be avoided in artificial blood pumps to achieve long-term clinical success. The relationship of these events to the fluid mechanics, velocity, shear and wall shear rates, and turbulence is the major impetus for flow studies in blood pumps. Neither phenomenon is completely understood. A hemolysis potential curve from the National Heart, Lung, and Blood Institute,^[15] which relates shear stress and exposure time to red cell, white cell, and platelet lysis has been available since 1985. Because blood cells are viscoelastic, they can tolerate high stresses for short exposure times without hemolysis. For example, an exposure time of more than 0.1 ms at a shear stress of $10,000 \text{ dyn/cm}^2$ will

produce red cell lysis, as will 1500 dyn/cm^2 for times over 100 s. Nevaril et al.^[16] concluded that prolonged exposure to laminar shear stress on the order of 1500 dyn/cm^2 could cause lysis of red cells, and Sallam and Hwang^[17] showed that sustained turbulent stresses above 4000 dyn/cm^2 created by a submerged jet would cause hemolysis. Baldwin et al.^[11] concluded on the basis of these and other published studies that stress levels above $1500\text{--}4000 \text{ dyn/cm}^2$ were undesirable. Platelet activation and the initiation of the clotting process may occur at still lower stresses. Thrombosis formation has long been thought to be a function of, among other factors, (low) wall shear stress and blood residence time (see Wootton and Ku,^[18] for example). Hubbell and McIntire^[19] reported that the wall shear rate should be above 500 s^{-1} (18 dyn/cm^2 for a viscosity of 3.5 centipoise [cp]) to prevent clot formation on segmented polyurethane (the blood sac material). Daily et al.^[20] pointed out that “the thrombogenicity of assist devices can be attributed to (1) the coagulability of the blood, (2) the properties of the blood contacting surfaces, and (3) fluid dynamic factors.” It is often not easy to separate these. Early fluid mechanics studies were through flow visualization,^[4,21] single-component laser Doppler anemometry,^[22] hot film anemometry in conjunction with dye washout,^[23] and pulsed Doppler ultrasound.^[24,25] Jarvis et al.^[26] used human blood in a 100 cc artificial ventricle to measure hemolysis directly through quantification of plasma-free hemoglobin and found that the degree of hemolysis was a function of the operating conditions. Baldwin et al.^[27] did extensive measurements of wall shear stress inside a ventricle using flush-mounted hot film anemometry probes. Francischelli et al.^[28] used a fiber optic system to look at residence times for an analog fluid doped with fluorescein dye. Flow visualization continues to be useful for qualitative assessment. More recent flow visualization studies are by Hochareon et al.^[10] Mussivand et al.^[29] and Woodward et al.^[30] for example.

IN VITRO MEASUREMENTS

Baldwin et al.^[11,31–33] published what is still the most thorough study of artificial heart fluid mechanics. They used two-component laser Doppler anemometry (LDA) to make mean and turbulence measurements at some 135 locations within the ventricle and 10 locations at each of the outlet and inlet flow tracts at normal physiologic conditions. A standard four-beam, two-component system was used in backscatter with counter signal processors to perform the measurements. The measurement ellipsoids had a diameter of roughly $65 \mu\text{m}$ and a length of 1.13 mm.

The beat cycle was divided into 8 time windows which varied from 20 to 100 ms, as a function of data rate (as described by Baldwin et al.^[33]), with 40 ms used for most cycle times and locations. Coincident data occurring during any time of interest was placed in the appropriate time window file. Mean and fluctuating velocities and Reynolds stresses were calculated from 250 ensembles at each time window and location. Baldwin et al.^[33] estimated that 95% of the Reynolds stresses would be within 20% of the (converged) values obtained for 4096 ensembles. The Reynolds stresses are not invariant to coordinate rotations, so that data was presented, in principle axes, as the maximum Reynolds normal and shear stress.^[33] A problem inherent to turbulence measurements of this nature is that the beat-to-beat variation of the flow will appear as a “pseudo turbulence” that cannot be separated out. Setting a single “coincidence time” in these unsteady flows may also lead to errors in the stress. In addition, we note, as do the authors, that it is not clear how the Reynolds stresses are related to the damage of red blood cells—roughly $3 \times 8\text{-}\mu\text{m}$, biconcave disks. Perhaps the case can be made, as the authors do, that the turbulent dissipation will increase as the Reynolds stress to the $3/2$, so that the Kolmogorov scale, proportional to the stress to the $-3/8$, will be smaller as the stress increases and therefore more dangerous to the red cells. Some estimates by Baldwin et al.^[11] suggest that the small-scale structure of regurgitant jets through the closed valves is the order of $5\ \mu\text{m}$, as discussed below.

We reproduce the mean velocity map of the chamber flow in Fig. 2. Mean velocities in the chamber are not available at 300 and 400 ms into the cycle (during systole), as the beams are blocked by the pusher plate. The highest velocities in the chamber are in the major orifices of the aortic valve (1.9 m/s) and of the mitral valve (1.2 m/s) in early systole and early diastole, respectively. The inlet jet through the major orifice helps to produce a rotational pattern in the chamber that persists into early systole (0–500 ms). The authors note that this rotational pattern appears to provide good “washing” of the chamber. Other experiments, with sac-type artificial hearts, note quite similar flow patterns (see, for example^[34]). However, Baldwin et al.^[11] demonstrate that the minor orifice of the mitral valve does not show significant inflow during diastole (400–700 ms), and that this may be a result of the rotational motion “clipping” the incoming flow. Of great interest are the large retrograde fluid velocities, through the “closed” valves, in the near wall regions of the aortic valve during diastole and mitral valve during systole.

The major Reynolds normal stresses are shown in Fig. 3. Major Reynolds shear stresses are half these values and are rotated 45° clockwise from the principle stress axis. The authors note that major normal stresses

do not exceed $1000\ \text{dyn/cm}^2$ in the chamber and $2000\ \text{dyn/cm}^2$ in the aortic outflow tract. The outflow values are similar to those observed by Yoganathan et al.^[35] with this valve. Much larger Reynolds stresses were found in the regurgitant (retrograde) jets. The mitral valve regurgitant jet is stronger than that of the aortic valve because of the larger pressure gradient across it during systole than across the aortic valve during diastole. Velocities as high as 4.4 m/s and normal stresses as large as $20,000\ \text{dyn/cm}^2$ were observed.

Baldwin et al.^[11] concludes by asking whether “artificial heart fluid mechanics can be improved.” They base this on the rather innocuous fluid mechanics of the pumping chamber and the relatively minor ways in which the geometry, with respect to the size and shape of the natural heart, may be changed. They find that the near valve flow is of concern. Maymir et al.^[36,37] continued the study of regurgitant jets, in particular, the influence of occluder to housing valve gap width. Meyer et al.^[38,39] extended the work by using a three-component laser Doppler system for three additional MHVs the Medtronic-Hall tilting disc, the Carbomedics, and St. Jude bileaflet designs and report turbulent jets with large sustained Reynolds stress even for the bileaflet valves.

An additional concern with using MHVs is the recognition^[40–42] that they cavitate. Cavitation is the formation of bubbles from gaseous nuclei in the fluid due to a drop in local pressure.^[43] Although Zapanta et al.^[44] showed valve cavitation in vivo in an artificial heart, the problem is not just associated with the use of MHVs in the artificial heart, but with the general use of these valves. There are several serious potential problems associated with cavitation—hemolysis and thrombosis initiation, valve leaflet damage, and the formation of stable gas bubbles that may find their way to the cranial circulation.

Although cavitation-induced pitting of explanted valves has been observed,^[45] significant valve leaflet damage is rare. Lamson et al.^[46] show that the index of hemolysis is a strong function of cavitation intensity and cavitation duration. There have been reports (for example, Dauzat et al.)^[47] of gaseous emboli in the cranial circulation, detected by Doppler ultrasound, for MHV recipients. Bachmann et al.^[48] Biancucci et al.^[49] and Lin et al.^[50] suggest that these emboli might be the aftermath of cavitation growth and collapse. There is no comprehensive current review, but much is described in the work of Graf et al.^[51] Zapanta et al.^[44] Chandran et al.^[52] and Bachmann et al.^[48]

The growing need for long-term pediatric, circulatory assist has resulted in a NIH program to develop such an assist device by 2009. The required output of the device is about 1 l/min. The simple geometric scaling of the pumps is described by Bachmann et al.^[53]

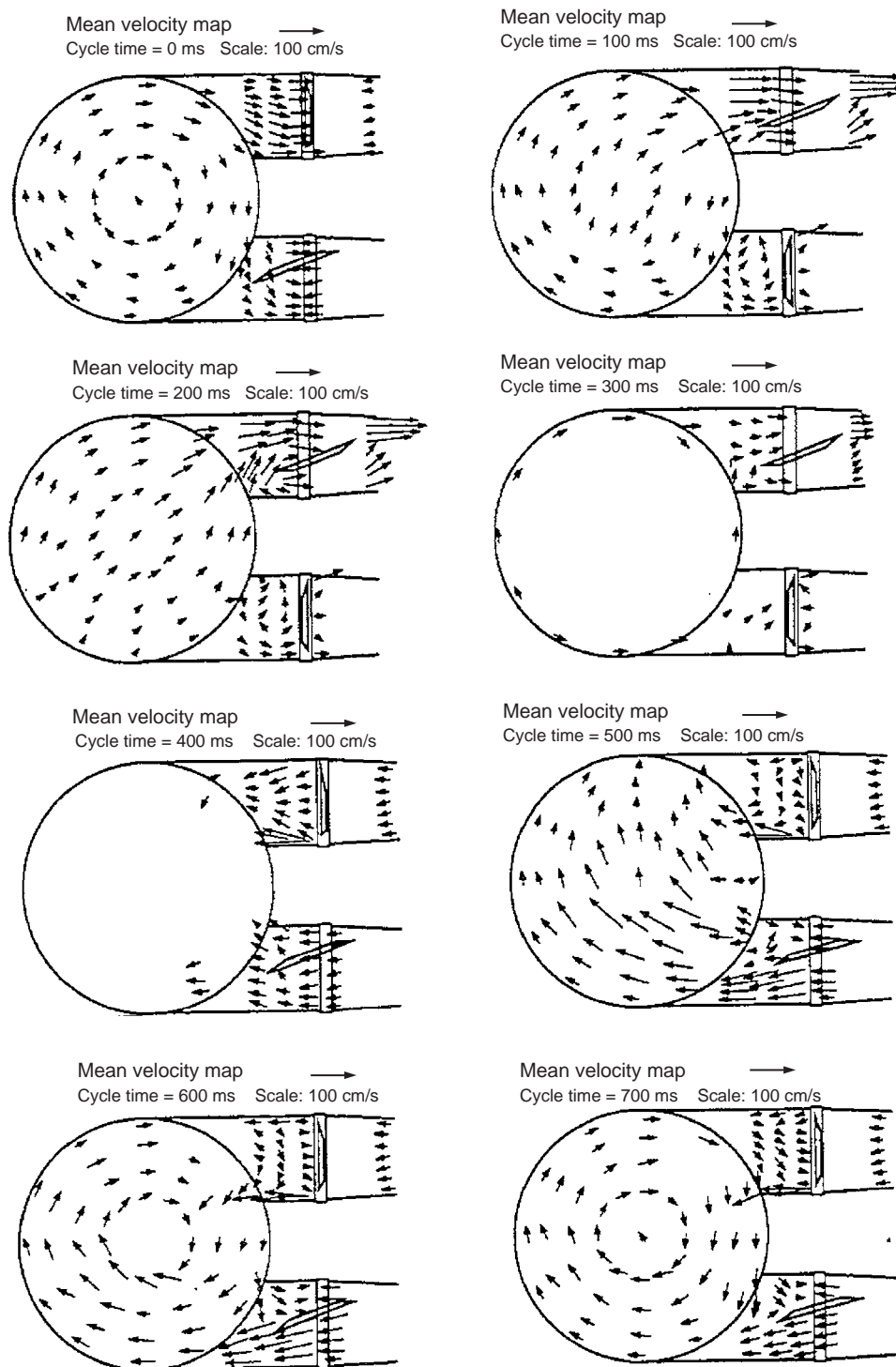


Fig. 2 Mean (ensemble averaged) velocity maps of a 70-cc device at 8 times during the cardiac cycle. Time zero is the onset of systole; diastole begins at 400 ms; and the cycle duration is 800 ms. Arrow lengths are proportional to mean velocity magnitude (see scale) and point in the direction of the mean velocity vector. The aortic (ejecting) port is located at the top and the mitral (filling) port is located at the bottom.

Source: Reprinted with permission from ASME (see Ref. [11]).

For example, to reduce the volume from 70 to 15 cc, one might reduce all linear dimensions by the cube root of the ratio of volumes. Assuming that the non-Newtonian nature of blood does not introduce any

additional parameters, the “global” fluid dynamics of the system is described by the Reynolds (Re) and Strouhal (St) numbers. In a study of 73 healthy subjects ranging in age from 5 days to 84 years, Gharib

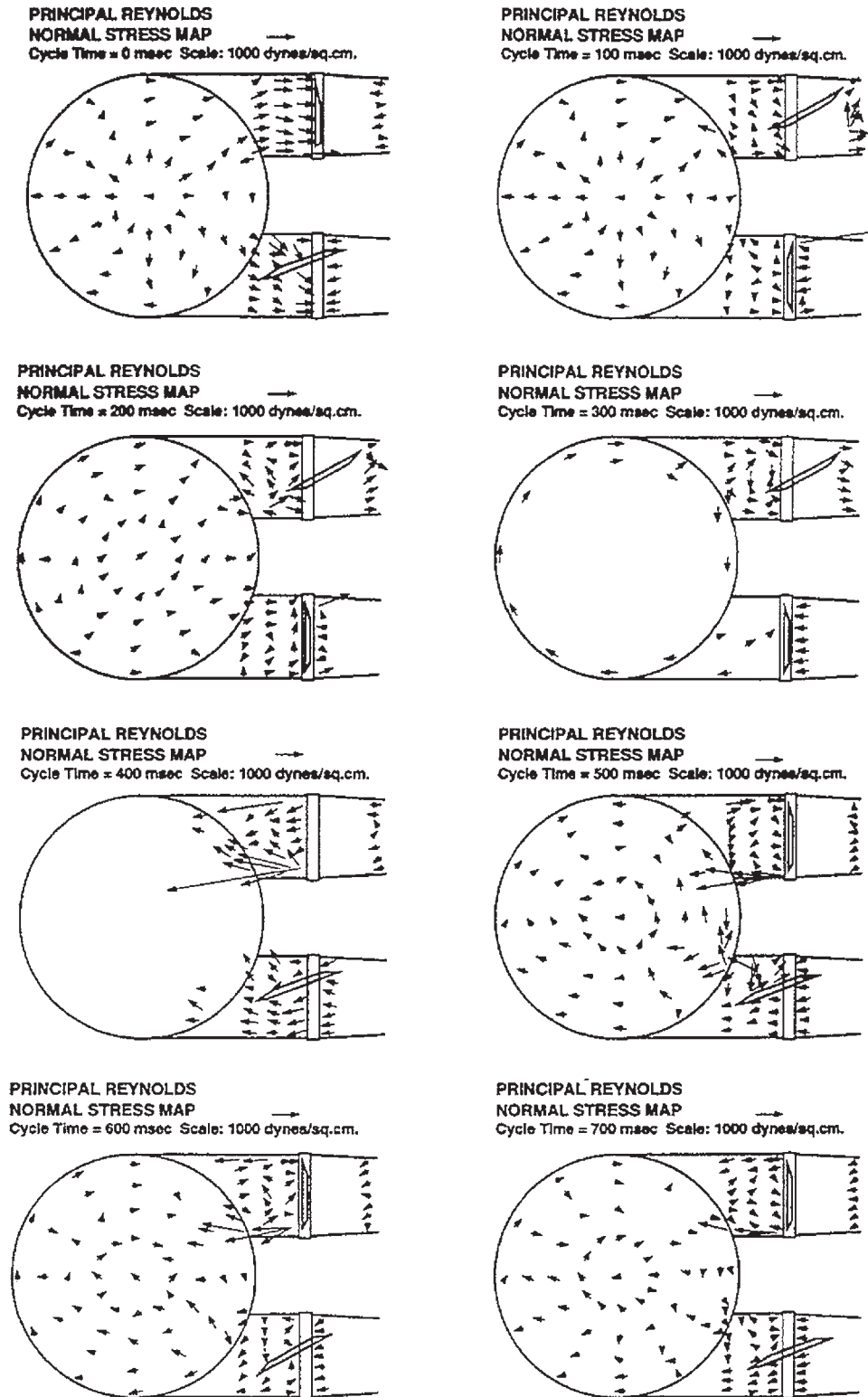


Fig. 3 Major Reynolds normal stress maps of a 70-cc device at 8 times during the cardiac cycle. Time zero is the onset of systole; diastole begins at 400 ms; and the cycle duration is 800 ms. Arrow lengths are proportional to normal stress magnitude (see scale) and point in the direction of the major axis of the principal stress axes. The aortic (ejecting) port is located at the top and the mitral (filling) port is located at the bottom.
Source: Reprinted with permission from ASME (see Ref. [11]).



Fig. 4 The 15-cc Penn State pediatric ventricular assist device (circa 1995). The device measured 55 mm for the chamber outer diameter, 35 mm for the overall chamber thickness, 82 mm for the overall pump length with cannula connectors, and weighed 57 g.

et al.^[54] found that the Strouhal number remained fairly constant at 4–7. Bachmann et al.^[53] showed that geometrically similar pumps have constant Strouhal number.

Daily et al.^[20] and Bachmann et al.^[53] have both studied a roughly 15 cc pediatric assist device, shown in Fig. 4. Reynolds and Strouhal numbers for the devices, taken from Bachmann et al.^[53] are given in Table 1.

Table 1 Comparison of the Reynolds and Strouhal numbers for the 70, 50, and 15-cc artificial blood pumps

Pump size	Re	St
Penn state 70-cc device	2482	8.3
Penn state 50-cc device	1054	4.5
Penn state 15-cc device	1567	45.3
Yonsei 34-cc device	1500	7.7
Toyobo 20-cc device	988	6.3
MEDHOS-HIA 10-cc device	655	7.4
Berlin heart 12-cc device	785	8.8

The large increase in *St* for the 15 cc device is a result of undersizing the inlet port.

Daily et al.^[20] provided both PIV maps and clinical studies of the device that focused on the choice of MHVs—handmade ball and cage valves (which were initially used clinically) vs bileaflet valves. The PIV maps compared valve types for a single instant of diastole and a single instant of systole. They reported that for the bileaflet valve the inlet jet penetrated more deeply into the chamber and was more coherent; the diastolic rotational motion was formed sooner and the amount of fluid entrained by the outlet jet was greater. In addition, the pressure drop and mean energy loss through the ball and cage valves was much greater than that through the bileaflet. Moreover, they reported that animal experiments of the device with handmade ball and cage valves showed thrombus formation in the device—something rarely seen in larger pumps. Initial experiments with the bileaflet valves showed no such thrombus formation.

Bachmann et al.^[53] used a TSI Inc. two-component LDA system to measure mean and turbulence quantities in a pediatric ventricle with handmade ball and cage valves at normal physiologic conditions. By using beam expansion, they reduced each measurement volume to a roughly $200 \times 30 \mu\text{m}^2$ ellipse. At each of 75 locations, 250 ensembles were measured at distances from the wall opposite the pusher plate of 0.1, 0.3, 0.6, and 1.0 mm. The data reduction follows Baldwin et al.^[11] Both a sodium iodide solution and a Xanthan gum viscoelastic solution were used. The wall shear rate was estimated from the velocity measurement 0.1 mm from the wall, using the no-slip condition. A gray-scale contour map of the average wall shear stress over the filling portion of the cycle is shown for each fluid in Fig. 5 (the mitral port is located on the right side of the device). Large regions of very low wall shear stress are apparent. A similar plot for the wall shear stresses averaged over the ejection portion of the cycle is shown in Fig. 6. Again, we observe large regions of very low shear stresses. Differences between the results for each fluid, particularly on the inlet side of the model, are striking. Bachmann et al.^[53] compare the characteristics of the Penn State pediatric pump with other small pumps that have shown some clinical promise. These include pumps described by Park and Kim,^[55] who use Carbomedics bileaflet valves; by Taenaka et al.^[56] and Takano et al.^[57] who use Bjork-Shiley tilting disk valves; and by Konertz et al.^[58,59] who use polyurethane trileaflet valves. A consequence of using commercially available valves is a larger inlet length scale and reduced Strouhal number. Comparisons among the pumps, adapted from Bachmann et al.^[53] are shown in Table 1.

The 70–100 cc ventricles described earlier are too large, as the basis for implantable artificial hearts and

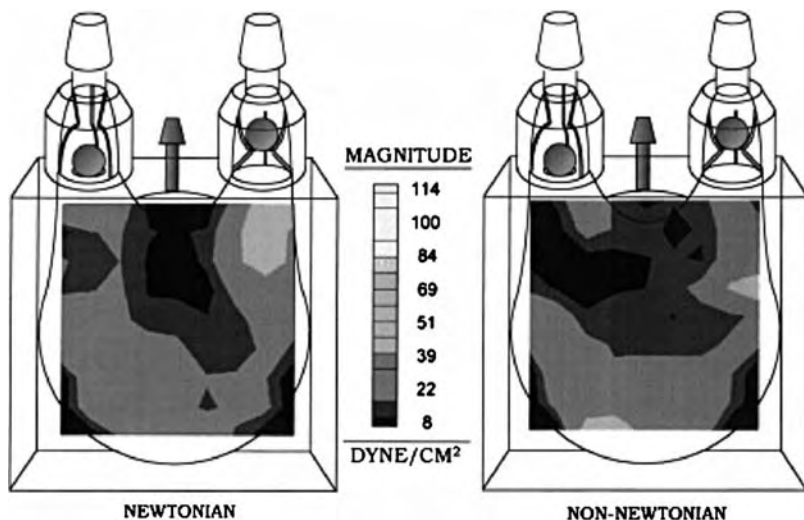


Fig. 5 The diagrams of a pediatric device design show wall shear stresses averaged over the filling portion of the cardiac cycle for both Newtonian (left) and non-Newtonian fluids (right).

blood pumps, to be used for much of the adult population. The development of smaller blood pumps that do not sacrifice cardiac output is a continuing research area. Hochareon et al.^[10,60-62] and Oley et al.^[63] recently presented a study of the mean velocity and wall shear stress in a 50 cc device using high-speed video and PIV. Hochareon et al.^[10] examined the opening pattern of the diaphragm using high-speed video. They determined that the opening pattern of the diaphragm, as it affected the diastolic jet and subsequent rotational motion, was a critical aspect of the overall flow. By comparison against flow visualization of the sac motion in a clinically approved 70-cc device, they also showed that the diaphragm motion was a good representative of the whole sac motion. Jin and Clark^[64] reported a similar study.

Hochareon et al.^[60-62] made PIV measurements in the transparent 50 cc pump model as a function of pump cycle time. The Reynolds and Strouhal number

are included in Table 1. All measurements were in the plane of the pusher plate. In this design, however, the inlet valve is rotated 30° from the pusher plate direction so that the light sheet is not aligned with the maximum jet velocity. The blood analog fluid was mineral oil. The pump was run at physiological conditions. A standard, planar TSI, Inc. PIV system was used to acquire 200 images at each condition. The light sheet was estimated at 0.5 mm thick and was initially centered 5 mm from the front edge. Cross-correlation of the images was performed by the TSI, Inc., Insight™ software. The final interrogation window size was 16 × 16 pixels. Both a global area and eight local areas (medial and lateral walls of the mitral and aortic ports and walls of the chamber body) were investigated. Resolution was 85 μm/pixel and 25 μm/pixel for the global and local maps, respectively. Components of the velocity gradient were calculated as central differences and wall shear rates estimated from the velocity point

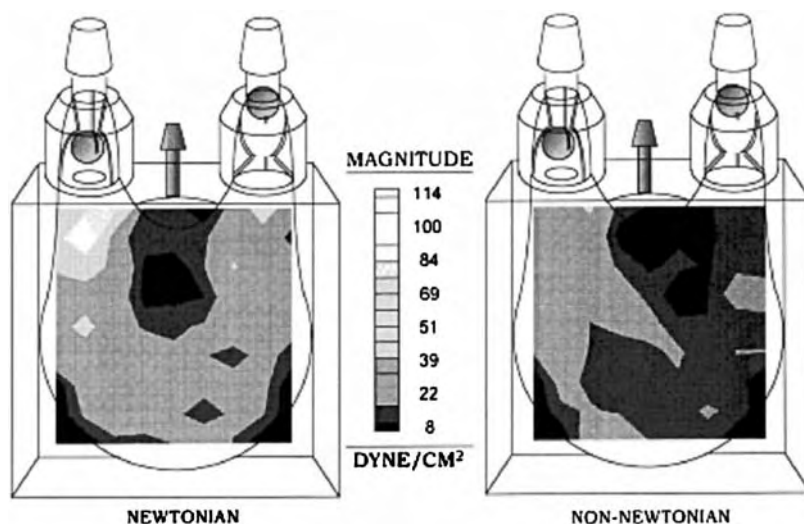


Fig. 6 The diagrams of a pediatric device design show wall shear stresses averaged over the ejection portion of the cardiac cycle for both Newtonian (left) and non-Newtonian fluids (right).

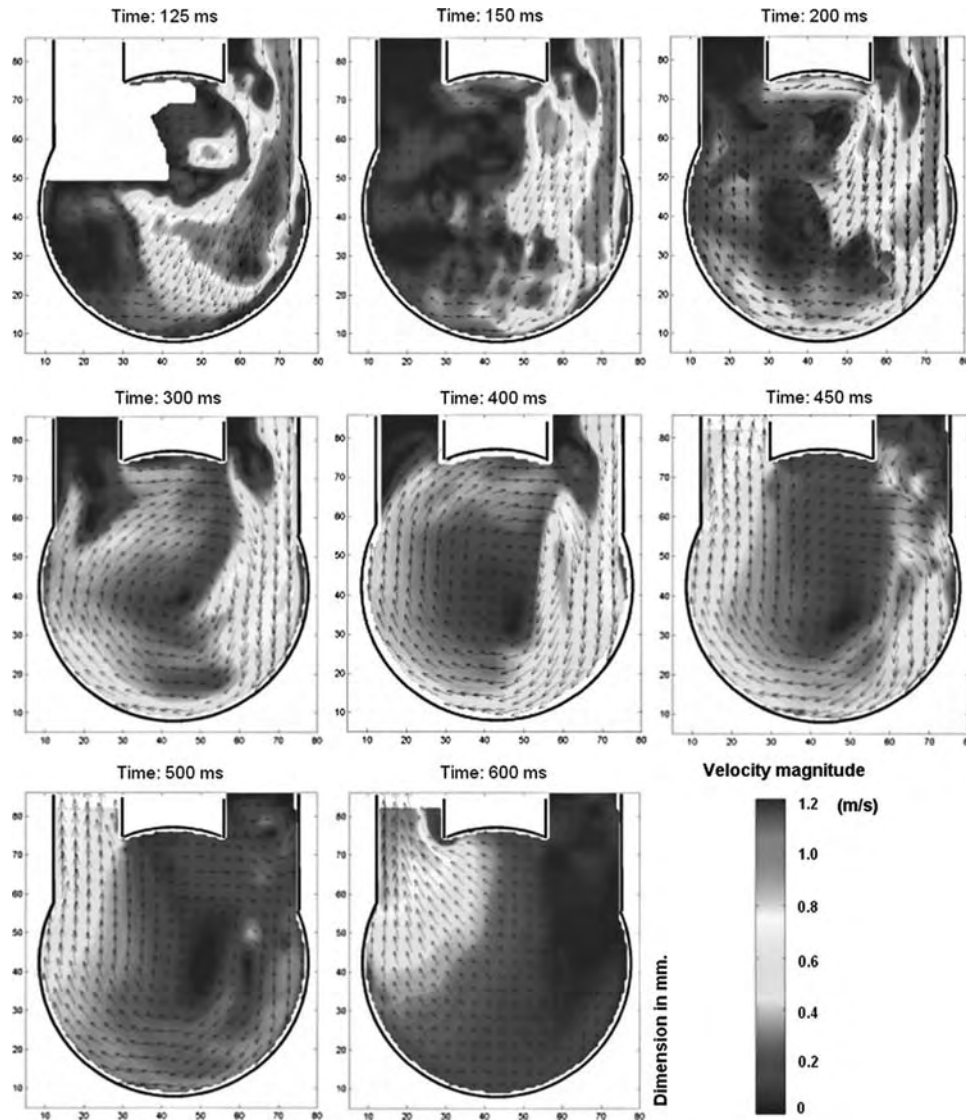


Fig. 7 The PIV velocity maps during early diastole (125 and 150 ms), middle to late diastole (200–400 ms), and systole (450–600 ms) for the 50-cc Penn State ventricular assist device. (Time reference is from the onset of diastole.) Source: Reprinted with permission from ASME (see Refs. [60–62]).

nearest the wall. The authors did not attempt to use PIV to estimate the turbulence levels.

Global flow maps are shown in Fig. 7. Note that diastole starts at 0 ms and systole at 430 ms with the mitral port on the right side of the chamber. The flow is again dominated by the diastolic jet and subsequent large-scale rotation. Peak velocities are the order of those observed by Baldwin et al.^[11] in a 70-cc device. The authors use vorticity maps to highlight the growth of the wall boundary layers. The local flow field near the mitral port at 200 and 400 ms is reproduced in Fig. 8. The associated wall shear rates shown in Fig. 9 never exceed some 3000 s^{-1} . The secondary inflow jet through the minor orifice of the mitral port had not been previously studied.

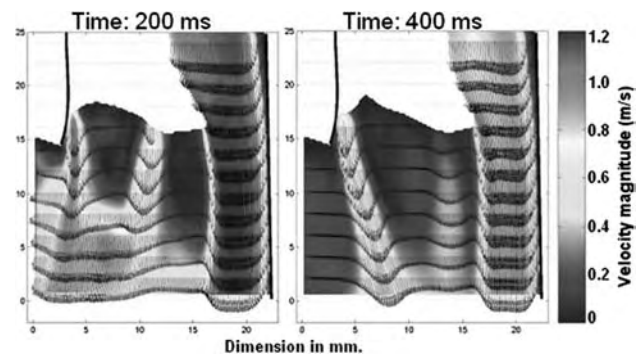


Fig. 8 The velocity maps of the mitral port at 200 ms and 400 ms for the 50-cc Penn State ventricular assist device. Source: Reprinted with permission from ASME (see Refs. [60–62]).

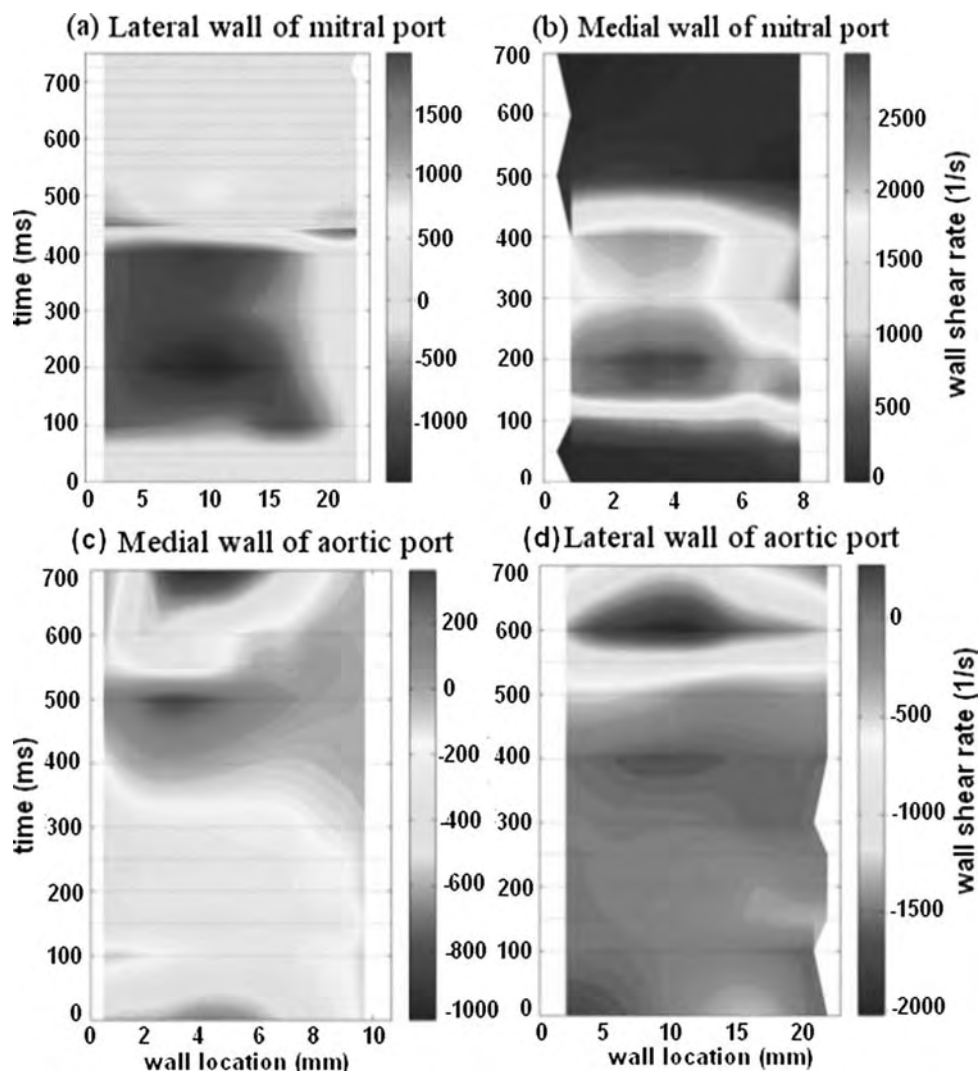


Fig. 9 The inlet port's average wall shear rate in time series in the beat cycle from (1) the lateral wall of the mitral port (the fully open valve tip position is at wall location approximately 16 mm), and (2) the medial wall of the mitral port (the fully open valve tip position is at wall location approximately 8 mm), of the 50-cc Penn State ventricular assist device. Source: Reprinted with permission from ASME (see Refs. [60–62]).

The local flow and vorticity fields near the bottom of the device are shown in Fig. 10. Shear rates for this region are $0\text{--}250\text{ s}^{-1}$. The authors note such low shear rates over the entire cycle are of concern. Similar shear rates are observed at the upper wall region between the valve ports. A rough summary of shear rates in the device is reproduced in Fig. 11. In general, the wall shear rates observed in the 50 cc device are much lower than those observed by Baldwin et al.^[27] in a 100-cc ventricle. Hochareon et al.^[61] developed refined methods to estimate the wall shear stress from PIV measurements in the artificial ventricle. Issues include the improvement of wall location estimates and the position of the velocity vector in the irregular measurement volumes nearest the wall. The influence of the size of the interrogation region was studied by

simulations. Hochareon et al.^[62] used the refined method for determining wall shear rate to obtain more extensive data in the bottom region of the 50 cc device. Yamanaka et al.^[65] performed an in vivo study of clot deposition in the 50-cc heart implanted in calves, which shows good correlation with regions of persistent low wall shear. Much more work correlating wall shear and clot formation is warranted.

Oley et al.^[63] recently completed a PIV study of the effect of beat rate and systolic duration on the global flow characteristics in the same 50-cc device. Shorter diastolic times produced a stronger inlet jet and an earlier and stronger diastolic rotation. However, the stronger the diastolic rotation, the larger the separated flow region on the inlet side of the aortic valve. The authors note that the relatively rapid acquisition of

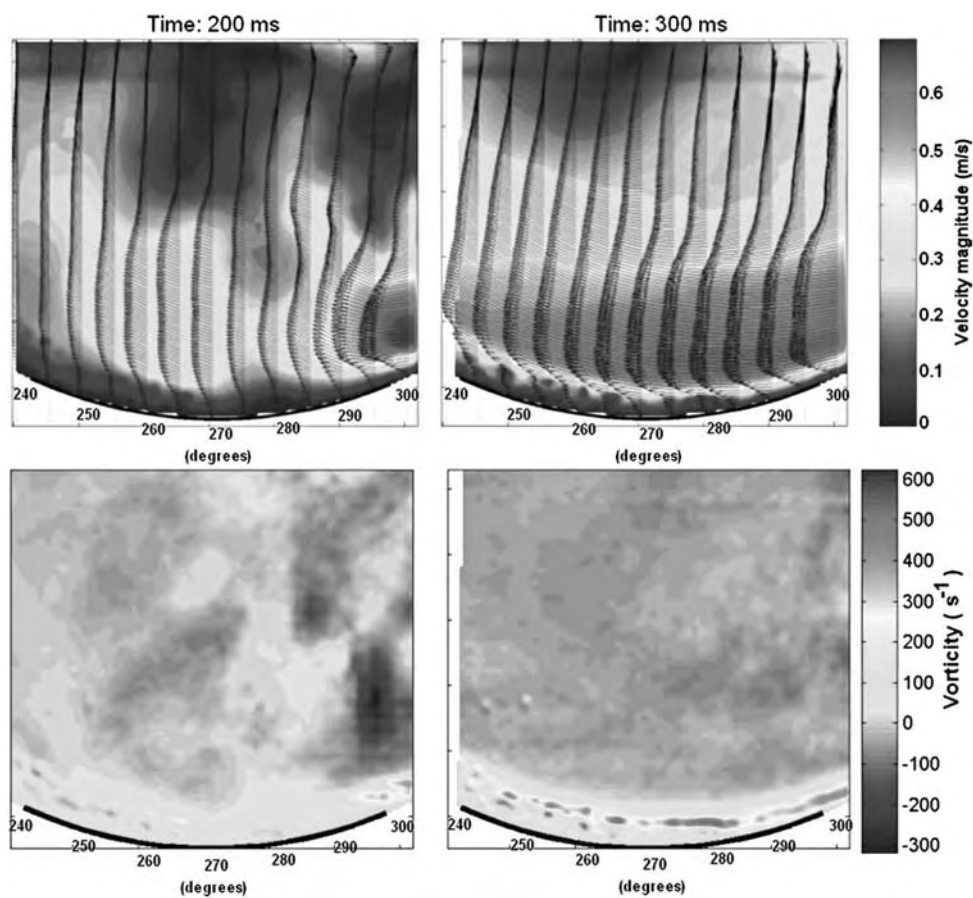


Fig. 10 The velocity and vorticity maps of the bottom wall from time 200 and 300 ms for the 50-cc Penn State ventricular assist device.

Note: The size of area is $30 \times 30 \text{ mm}^2$.

Source: Reprinted with permission from ASME (see Refs. [60–62]).

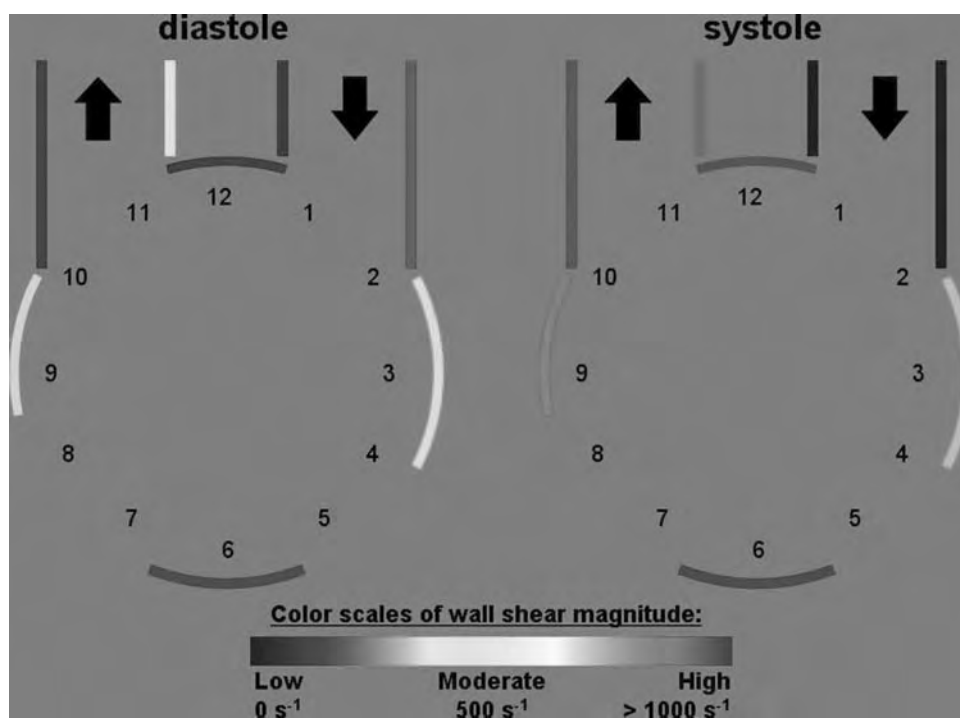


Fig. 11 Qualitative summary of wall shear rates within the 50-cc Penn state ventricular assist device during diastole and systole.

Source: Reprinted with permission from ASME (see Refs. [60–62]).

whole-flow field data, using PIV may permit experiments to play a more active role in the design process for artificial devices.

CONCLUSIONS

For artificial ventricles suitable for large adults ($>/= 70$ cc), clot formation within the ventricle is not generally observed. The major problems are associated with the valves, both with the high stresses in the regurgitant jets and with the influence of cavitation. Activation of the clotting cycle is likely, although the clots do not adhere to the surface of the pump. Smaller pumps show some thrombus deposition in addition to the valve related problems. Maintaining the inlet Strouhal number near physiologic values is sensible, but clot deposition has been observed in a 50-cc device with a physiologic Strouhal number of about 4.

Analysis of the fluid mechanics of artificial blood pumps has helped to drive design modifications over the years. For example, measurements of stagnation led our group to move the inlet port of the 100-cc pump from the center to the periphery of the device, near wall stasis led us to redesign the inlet port to be more parallel to pump body in our 50-cc device, and extensive measurements helped determine that ball and cage valves would not be suitable for the pediatric device (15 cc). Details of the local fluid mechanics, particularly of the wall shear stresses, will be critical to the successful design of the smaller pumps. Oley et al.^[63] note that the relatively rapid acquisition of whole-flow field data using PIV will be useful in this regard, but we note that the motion of the formed blood elements and their interaction with the artificial materials are a parallel part of the problem not yet addressed by experiment. Computation of the flow field and motion of the formed elements would be extremely useful, but the problems facing a successful computation are formidable. The flow and species motion are unsteady with valve-induced turbulence (at modest Reynolds number) through some of the cycle. The fluid is shear thinning and viscoelastic; the flow is driven by a flexible sac. Work in this important area seems likely to continue for a long time.

Finally, a good deal of effort is currently directed toward the development and testing of rotary blood pumps including axial and centrifugal flow assist devices.^[66]

ACKNOWLEDGMENTS

We gratefully acknowledge the support of 30 years of continuous National Institutes of Health funding

from NHLBI Grants HL13426, HL20356, HL48652, HL62076, RR15930, HV48191, and HV88105. We also appreciate the dedication and hard work from the faculty, engineers, graduate students, technicians, undergraduate students, and support staff at both the University Park and Hershey campuses of Pennsylvania State University during this research endeavor.



ARTICLES OF FURTHER INTEREST

Tissue Engineering of Heart Valves; Thrombosis

REFERENCES

1. LeGallois, M.; Nancrede, N.C.; Nancrede, J.G. *Experiments on the Principle of Life*; M. Thomas: Philadelphia, PA, 1813.
2. Hall, D.P.; Moreno, J.R.; Dennis, C.; Senning, A. An experimental study of prolonged left heart bypass without thoracotomy. *Ann. Surg.* **1962**, *156*, 190–196.
3. Cooley, D.A. First human implantation of cardiac prosthesis for staged total replacement of the heart. *ASAIO Trans.* **1969**, *15*, 252.
4. Phillips, W.M.; Brighton, J.A.; Pierce, W.S. Artificial heart evaluation using flow visualization techniques. *ASAIO Trans.* **1972**, *18*, 194–199.
5. Deutsch, S.; Tarbell, J.M.; Manning, K.B.; Rosenberg, G.; Fontaine, A.A.; et al. Experimental fluid mechanics of pulsatile artificial blood pumps. *Annu. Rev. Fluid Mech.* **2006**, *38*, 65–86.
6. Mehta, S.M.; Pae, W.E.; Rosenberg, G.; et al. The LionHeart LVD-2000: a completely implanted left ventricular assist device for chronic circulatory support. *Soc. Thorac. Surg.* **2001**, *71*, S156–S161.
7. Rosenberg, G.; Phillips, W.M.; Landis, D.; Pierce, W.S. Design and evaluation of The Pennsylvania State University mock circulatory system. *ASAIO J.* **1981**, *4*, 41–49.
8. Cokelet, G.R. The rheology and tube flow of blood. In *Handbook of Bioengineering*; Skalak, R., Chien, S., Eds.; McGraw Hill: New York, 1987; 14.1–14.17.
9. Thurston, G.B. Viscoelastic properties of blood and blood analogs. In *Advances in Hemodynamics and Hemorheology*; How, T.V., Ed.; JAI: Greenwich, Ct, 1996; 1–34.
10. Hochareon, P.; Manning, K.B.; Fontaine, A.A.; Deutsch, S.; Tarbell, J.M.; et al. Diaphragm motion affects flow patterns in an artificial heart. *Artif. Organs* **2003**, *27*, 1102–1109.
11. Baldwin, J.T.; Deutsch, S.; Geselowitz, D.B.; Tarbell, J.M. LDA measurements of mean velocity and Reynolds stress fields within an artificial heart ventricle. *J. Biomech. Eng.* **1994**, *116*, 190–200.
12. Brookshier, K.A.; Tarbell, J.M. Evaluation of a transparent blood analog fluid: aqueous Xanthan gum/glycerin. *Biorheology* **1993**, *30*, 107–116.
13. Long, J.A.; Undar, A.; Manning, K.B.; Deutsch, S.; et al. The viscoelasticity of pediatric blood and its

- implications for the testing of a pulsatile pediatric blood pump. *ASAIO J.* **2005**, *51*, 563–566.
14. Yoganathan, A.P.; He, Z.; Jones, S.C. Fluid mechanics of heart valves. *Annu. Rev. Biomed. Eng.* **2004**, *6*, 331–362.
 15. National Heart, Lung, and Blood Institute Working Group. *Guidelines for blood materials interaction*. NIH Pub. 85-2185.78 Natl Heart Lung Blood Inst.; Bethesda, MD: 1985.
 16. Nevaril, C.; Hellums, J.; Alfrey, C.J.; Lynch, E. Physical effects in red blood cell trauma. *Am. Inst. Chem. Eng. J.* **1969**, *15*, 707–711.
 17. Sallam, A.M.; Hwang, N.H.C. Human red blood cell hemolysis in a turbulent shear flow: contribution of Reynolds shear stresses. *Biorheology* **1984**, *21*, 783–797.
 18. Wootton, D.M.; Ku, D.N. Fluid mechanics of vascular systems, diseases, and thrombosis. *Annu. Rev. Biomed. Eng.* **1999**, *1*, 299–329.
 19. Hubbell, J.A.; McIntire, L.V. Visualization and analysis of mural thrombogenesis on collagen, polyurethane and nylon. *Biomaterials* **1986**, *7*, 354–363.
 20. Daily, B.B.; Pettitt, T.W.; Sutura, S.P.; Pierce, W.S. Pierce-Donachy pediatric VAD: progress in development. *Ann. Thorac. Surg.* **1996**, *61*, 437–443.
 21. Lenker, J.A. *Flow studies in artificial hearts and LVAD: an application of flow visualization analysis*. PhD thesis, Pennsylvania State Univ; 1978.
 22. Phillips, W.M.; Furkay, S.S.; Pierce, W.S. Laser Doppler anemometer studies in unsteady ventricular flows. *ASAIO Trans.* **1979**, *25*, 56–60.
 23. Affeld, A. The state of the art of the Berlin total artificial heart—technical aspects. In *Assisted Circulation*; Unger, F., Ed.; Springer: New York, 1979; 307–333.
 24. Mann, K.A.; Deutsch, S.; Tarbell, J.M.; Geselowitz, D.B.; Rosenberg, G.; Pierce, W.S. An experimental study of Newtonian and non-Newtonian Flow dynamics in a ventricular assist device. *J. Biomech. Eng.* **1987**, *109*, 139–147.
 25. Tarbell, J.M.; Gunshinan, J.P.; Geselowitz, D.B.; Rosenberg, G.; Shung, K.K.; Pierce, W.S. Pulsed ultrasonic Doppler velocity measurements inside a left ventricular assist device. *J. Biomech. Eng.* **1986**, *108*, 232–238.
 26. Jarvis, P.; Tarbell, J.M.; Frangos, J.A. An in vitro analysis of an artificial heart. *ASAIO Trans.* **1991**, *37*, 27–32.
 27. Baldwin, J.T.; Tarbell, J.M.; Deutsch, S.; Geselowitz, D.B.; Rosenberg, G. Hot-film wall shear probe measurements inside a ventricular assist device. *J. Biomech. Eng.* **1988**, *110*, 326–333.
 28. Francischelli, D.E.; Tarbell, J.M.; Geselowitz, D.B. Local blood residence times in the Penn State artificial heart. *Artif. Organs* **1991**, *15*, 218–224.
 29. Mussivand, T.; Navarro, R.; Chen, J.F. Flow visualization in an artificial heart using diffuse and planar lighting. *ASAIO Trans.* **1988**, *34*, 317–321.
 30. Woodward, J.; Shaffer, F.; Schaub, R.; Lund, L.; Borovetz, H. Optimal management of a ventricular assist device. *ASAIO J.* **1992**, *38*, M216–M219.
 31. Baldwin, J.T.; Tarbell, J.M.; Deutsch, S.; Geselowitz, D.B. Mean flow velocity patterns within a ventricular assist device. *ASAIO Trans.* **1989**, *35*, 429–433.
 32. Baldwin, J.T.; Deutsch, S.; Geselowitz, D.B.; Tarbell, J.M. Estimation of Reynolds stresses within the Penn State left ventricular assist device. *ASAIO Trans.* **1990**, *3*, M274–M278.
 33. Baldwin, J.T.; Deutsch, S.; Petrie, H.L.; Tarbell, J.M. Determination of principal Reynolds stresses in pulsatile flows after elliptical filtering of discrete velocity measurements. *J. Biomech. Eng.* **1993**, *115*, 396–403.
 34. Jin, W.; Clark, C. Experimental investigation of unsteady-flow behavior within a sac-type ventricular assist device (VAD). *J. Biomech.* **1993**, *26*, 697–707.
 35. Yoganathan, A.P.; Woo, Y.R.; Sung, H.W. Turbulent shear stress measurements in the vicinity of aortic heart valve prostheses. *J. Biomech.* **1986**, *19*, 422–433.
 36. Maymir, J.C.; Deutsch, S.; Meyer, R.; Geselowitz, D.B.; Tarbell, J.M. Effects of tilting disk valve gap width on regurgitant flow through an artificial heart mitral valve. *Artif. Organs* **1997**, *21*, 1014–1025.
 37. Maymir, J.C.; Deutsch, S.; Meyer, R.; Geselowitz, D.B.; Tarbell, J.M. Mean velocity and Reynolds stress measurements in the regurgitant jets of tilting disk valves in an artificial heart environment. *Ann. Biomed. Eng.* **1988**, *26*, 146–156.
 38. Meyer, R.S.; Deutsch, S.; Maymir, J.C.; et al. Three component LDV measurements in the regurgitant flow region of a Bjork-Shiley monostrut mitral valve. *Ann. Biomed. Eng.* **1997**, *25*, 1081–1091.
 39. Meyer, R.S.; Deutsch, S.; Bachmann, C.B.; Tarbell, J.M. Laser Doppler velocimetry and flow visualization studies in the regurgitant leakage flow region of three mechanical heart valves. *Artif. Organs* **2001**, *25*, 292–299.
 40. Leuer, L. In vitro evaluation of drive parameters and valve selection for the total artificial heart. *Proceedings of International Symposium of Artificial Organs, Biomedical Engineering and Transplantation*, Salt Lake City, Utah: 1986; Univ. Utah: Salt Lake City, 1986.
 41. Quijano, R. Edwards-Duromedics dysfunctional analysis. *Proceedings of 6th Conference on Cardiosimulation*, Monte Carlo, 1988.
 42. Walker, W. *Cavitation in pulsatile blood pumps*; ASME: New York, 1974; 148–150.
 43. Young, F.R. *Cavitation*; McGraw-Hill: London, 1989.
 44. Zapanta, C.M.; Stinebring, D.R.; Sneckenberger, D.S.; et al. In vivo observation of cavitation on prosthetic heart valves. *ASAIO J.* **1996**, *42*, M550–M555.
 45. Kafesian, R.; Howanec, M.; Ward, G.D.; Diep, L.; Wagstaff, L.S.; Rhee, R. Cavitation damage of pyrolytic carbon in mechanical heart valves. *J. Heart Valve Dis.* **1994**, *3*, 52–57.
 46. Lamson, T.C.; Rosenberg, G.; Deutsch, S.; et al. Relative blood damage in the three phases of a prosthetic heart valve flow cycle. *ASAIO J.* **1993**, *39*, M626–M633.
 47. Dazat, M.; Deklunder, G.; Aldis, A.; Rabinovitch, M.; Burte, F.; Bret, P.M. Gas bubble emboli detected by transcranial Doppler sonography in patients with prosthetic heart valves: a preliminary report. *J. Ultrasound Med.* **1994**, *13*, 129–135.
 48. Bachmann, C.; Kini, V.; Deutsch, S.; et al. Mechanisms of cavitation and the formation of stable bubbles on the

- Bjork-Shiley monostrut prosthetic heart valve. *J. Heart Valve Dis.* **2001**, *11*, 105–113.
49. Biancucci, B.; Deutsch, S.; Geselowitz, D.B.; Tarbell, J.M. In vitro studies of gas bubble formation by mechanical heart valves. *J. Heart Valve Dis.* **1999**, *8*, 186–196.
 50. Lin, H.Y.; Biancucci, B.; Deutsch, S.; Fontaine, A.A.; Tarbell, J.M. Observation and quantification of gas bubble formation on a mechanical heart valve. *J. Biomech. Eng.* **2000**, *122*, 304–309.
 51. Graf, T.; Reul, H.; Detlefs, C.; Wilmes, R.; Rau, G. Causes and formation of cavitation in mechanical heart valves. *J. Heart Valve Dis.* **1994**, *1*, S49–S64.
 52. Chandran, K.B.; Aluri, S. Mechanical valve closing dynamics: relationship between velocity of closing, pressure transients, and cavitation initiation. *Ann. Biomed. Eng.* **1997**, *25*, 926–938.
 53. Bachmann, C.; Hugo, G.; Rosenberg, G.; et al. Fluid dynamics of a pediatric ventricular assist device. *Artif. Organs* **2000**, *24*, 362–372. [Erratum *Artif. Organs*; *24*:989].
 54. Gharib, M.; Rambod, E.; Shiota, T.; et al. Dynamic filling characteristics of the left ventricle of the heart. Presented at 3rd International Symposium Biofluid Mechanics, Munich, VDI Verlag: Dusseldorf, 1994; 343–345
 55. Park, Y.; Kim, S. Development and animal study of a pediatric ventricular assist device. *Yonsei Med. J.* **1988**, *39*, 154–158.
 56. Taenaka, Y.; Takano, H.; Noda, H.; Kinoshita, M. A pediatric ventricular assist device: its development and experimental evaluation of hemodynamic effects on postoperative heart failure of congenital heart diseases. *Artif. Organs* **1990**, *14*, 49–56.
 57. Takano, H.; Nakatani, T. Ventricular assist systems: experience in Japan with Toyobo pump and Zeon pump. *Ann. Thorac. Surg.* **1996**, *61*, 317–322.
 58. Konertz, W.; Holger, H.; Schneider, M.; et al. Clinical experience with the MEDOSHIA–VAD system in infants and children: a preliminary report. *Ann. Thorac. Surg.* **1997**, *63*, 1138–1144.
 59. Konertz, W.; Reul, H. Mechanical circulatory support in children. *Artif. Organs* **1997**, *20*, 657–658.
 60. Hochareon, P.; Manning, K.B.; Fontaine, A.A.; Tarbell, J.M.; Deutsch, S. Wall shear rate estimation within the 50 cc Penn State artificial heart using particle image velocimetry. *J. Biomech. Eng.* **2004**, *126*, 430–437.
 61. Hochareon, P.; Manning, K.B.; Fontaine, A.A.; Tarbell, J.M.; Deutsch, S. Fluid dynamic analysis of the 50 cc Penn State artificial heart under physiological operating conditions using particle image velocimetry. *J. Biomech. Eng.* **2004**, *126*, 585–593.
 62. Hochareon, P.; Manning, K.B.; Fontaine, A.A.; et al. Correlation of in vivo clot deposition with the flow characteristics in the 50 cc Penn State artificial heart: a preliminary study. *ASIAO J.* **2004**, *50*, 537–542.
 63. Oley, L.A.; Manning, K.B.; Fontaine, A.A.; Deutsch, S. Off design considerations of the 50 cc Penn State ventricular assist device. *Artif. Organs* **2005**, *29*, 378–386.
 64. Jin, W.; Clark, C. Experimental investigation of the pumping diaphragm within a sac-type pneumatically driven ventricular assist device. *J. Biomech.* **1994**, *27*, 43–55.
 65. Yamanaka, H.; Rosenberg, G.; Weiss, W.J.; et al. A multiscale surface evaluation of thrombosis in left ventricular assist systems. *ASAIO J.* **2003**, *49*, 222.
 66. Reul, H. *Overview of Rotary Blood Pump Designs*; Presented at Amer. Soc. Artif. Inter. Organs, 49th: Washington, DC, 2003.

Artificial Muscles

Mohsen Shahinpoor

Artificial Muscle Research Institute (AMRI), School of Engineering and School of Medicine, University of New Mexico, Albuquerque, New Mexico, U.S.A.

INTRODUCTION

The first attempts to create artificial muscles date to the pioneering work of Kuhn and his student Katchalsky and his students in the late 1940s and early 1950s in connection with pH-activated muscles or simply pH muscles.^[1–6] Note also the other means of activating artificial muscles such as thermal,^[7–9] pneumatic,^[10] optical,^[11] and magnetic.^[12] Chemically stimulated polymers were discovered more than half a century ago when it was shown that collagen filaments could reversibly contract or expand when dipped in acidic or alkaline solutions, respectively.^[1] This early work pioneered the development of synthetic polymers that mimic biological muscles.^[3,5] However, electrical stimulation has always remained the best means of artificial muscle material actuation and sensing. Shahinpoor^[4,5,10,13–19] was one of the pioneers in making electrically active, in sensing and actuation, ionic polymer–conductor composites (IPCCs) and ionic polymer–metal composites (IPMCs). Zhang^[20] was able to observe a substantial piezoelectric activity in polyvinylidene fluoride trifluoroethylene (PVF2-TrFE) as early as 1998. The greatest progress in artificial muscle materials development has occurred in the last 10 years when effective materials that can induce strains exceeding 100% have emerged.^[21]

SPECIFIC ARTIFICIAL MUSCLE MATERIALS

Shape Memory Alloys (SMAs) and Polymers (SMPs)

The history of SMA and SMP artificial muscles is extensive and will not be reported here. However, the pioneering works of Liang and Rogers^[22,23] and Liang, Rogers, and Malafeev^[24] in developing SMA and SMP actuators are noteworthy. In their work, the load of a spring-biased SMA actuator was modeled as a dead weight. However, many practical applications involve varying loads, as in the cases of SMA rotatory joint actuators.^[13,25–27] A general design methodology of various types of bias force SMA actuators has been investigated by Shahinpoor and Wang.^[28] Heat-shrink polymers such as polyolefins, PVC, polypropylene,

and polyesters are also sometimes considered as heat-activated shape memory polymers and artificial muscles.

Magnetically Activated Artificial Muscles

Magnetically activated gels, so-called ferrogels, are chemically cross-linked polymer networks that are swollen by the presence of a magnetic field.^[29] Zrinyi has been a pioneer of this technology. Such a gel is a colloidal dispersion of monodomain magnetic nanoparticles. Such magnetic ferrogel materials deform in the presence of a spatially nonuniform magnetic field because its embedded nanoparticles move in reaction to the field. The ferrogel material can be activated to bend, elongate, deform, curve, or contract repeatedly, and it has a response time of less than 100 ms, which is independent of the particle size. Another group of magnetically activated artificial muscles are the magnetostrictive materials such as TERFENOL-D, which are the metallic elements terbium (TER), iron (FE), Naval Ordnance Labs (NOL), and Dysprosium-D^[30] and give rise to infinitesimal dimensional changes of the order of 0.1% or so, and magnetic shape memory materials (MSM) such as NiMnGa, which undergo gigantic martensitic-austenitic solid-phase transformation and can give rise to 6–8% dimensional change.^[31]

Electronic EAP/Ferroelectric Polymeric Artificial Muscles

The basic phenomenon is called ferroelectricity when a nonconducting crystal or dielectric material exhibits spontaneous electric polarization. These are based on the phenomenon of piezoelectricity, which is found only in noncentrosymmetric materials such as poly(vinylidene fluoride) (also known as PVDF). These polymers are partly crystalline, with an inactive amorphous phase and a Young's modulus of about 1–10 GPa. This relatively high elastic modulus offers high mechanical energy density. A large applied AC field (–200 MV/m) can induce electrostrictive (non-linear) strains greater than 1%. Sen et al.^[32] investigated the effect of mixing plasticizers (–65% wt.) with ferroelectric polymers hoping to achieve large strains

at reasonable applied fields. However, the plasticizer is also amorphous and inactive, resulting in decreased Young's modulus, permittivity, and electrostrictive strains. As large as 4% electrostrictive strains can be achieved with low-frequency driving fields having amplitudes of about $150 \text{ V}/\mu\text{m}$. As with ceramic ferroelectrics, electrostriction can be considered to be the origin of piezoelectricity in ferroelectric polymers.^[33] Unlike electrostriction, piezoelectricity is a linear effect, where not only will the material be strained when voltage is applied, a voltage signal will be induced when a stress is applied. Thus they can be used as sensors, transducers, and actuators. Depoling due to excessive loading, heating, cooling, or electric driving is a problem with these materials.

Dielectric Elastomer Electroactive Polymers (EAPs)

Polymers with low elastic stiffness modulus and high dielectric constant can be packaged with interdigitated electrodes to generate large actuation strain by subjecting them to an electric field. This dielectric elastomer EAP can be represented by a parallel-plate capacitor.^[21] The induced strain is proportional to the square of the electric field, multiplied by the dielectric constant and is inversely proportional to the elastic modulus. Dielectric elastomer EAP actuators require large electric fields ($100 \text{ V}/\mu\text{m}$) and can induce significant levels of strain (10% to 200%). Recently, Perline et al. have introduced a new class of polymers that exhibits an extremely high strain response.^[34] These acrylic-based elastomers have produced large strains of more than 200% but suffer from the fact that they require gigantic electric fields in the range of hundreds of megavolts per meter.

Liquid Crystal Elastomer (LCE) Materials

Liquid crystal elastomers were pioneered at Albert-Ludwigs Universität (Freiburg, Germany).^[35] These materials can be used to form an EAP actuator by inducing isotropic-nematic phase transition due to temperature increase via Joule heating. LCEs are composite materials that consist of monodomain nematic liquid crystal elastomers and conductive polymers that are distributed within their network structure.^[14,36,37] The actuation mechanism of these materials involves phase transition between nematic and isotropic phases over a period of less than a second. The reverse process is slower, taking about 10s, and it requires cooling, causing expansion of the elastomer to its original length. The mechanical properties of LCE materials can be controlled and optimized by effective selection of the liquid crystalline

phase, density of cross-linking, flexibility of the polymer backbone, coupling between the backbone and liquid crystal group, and the coupling between the liquid crystal group and the external stimuli.

Ionic EAP/Ionic Polymer Gels (IPGs)

Polymer gels can be synthesized to produce strong actuators having the potential of matching the force and energy density of biological muscles. These materials (e.g., polyacrylonitrile or PAN) are generally activated by a chemical reaction, changing from an acid to an alkaline environment causing the gel to become dense or swollen, respectively. This reaction can be stimulated electrically, as was shown by Shahinpoor et al.^[5,10,14–19] Current efforts are directed toward the development of thin layers and more robust techniques of placing electrodes on the surface of the composite. Progress was recently reported by researchers at the University of New Mexico using a mix of conductive and PANS fibers.^[38] The mechanism that is responsible for the chemomechanical behavior of ionic gels under electrical excitation is described by Osada and Ross-Murphy,^[39] and a model for hydrogel behavior as contractile EAP is described in Gong et al.^[40] A significant amount of research and development was conducted at the Hokkaido University, Japan, and applications using ionic gel polymers were explored. These include electrically induced bending of gels^[39,41] and electrically induced reversible volume change of gel particles.^[42]

Nonionic Polymer Gels/EAPs

Nonionic polymer gels containing a dielectric solvent can be made to swell under a DC electric field with a significant strain. Hirai and his coworkers at Shinshu University in Japan have created bending and crawling nonionic EAPs using a poly(vinyl alcohol) gel with dimethyl sulfoxide.^[43,44] A $10 \times 3 \times 2 \text{ mm}$ actuator gel was subjected to an electric field and exhibited bending at angles that were greater than 90° at a speed of 6 cm/s . This phenomenon is attributed to charge injection into the gel and a flow of solvated charges that induce an asymmetric pressure distribution in the gel. Another nonionic gel is poly(vinyl chloride) (PVC), which is generally inactive when subjected to electric fields. However, if PVC is plasticized with dioctyl phthalate (DOP), a typical plasticizer, it can maintain its shape and behave as an elastic nonionic gel.

Ionic Polymer–Metal Composites (IPMCs)

Ionic polymer–metal composite (IPMC) is an EAP that bends in response to a small electric field ($5\text{--}10 \text{ V/mm}$)

as a result of mobility of cations in the polymer network. In 1992, IPMC was realized, based on a chemical plating technique developed by Merlet, Pinneri, and coworkers in France and by Kawami and Takanake in Japan in the 1980s. The first working actuators were built by Oguro et al.^[45] in Japan and by Shahinpoor^[4] in the United States. The first working sensors of this kind were first fabricated by Shahinpoor^[4] and Sadehipour et al.^[46] in the United States. The operation as actuators is the reverse process of the charge storage mechanism associated with fuel cells^[47] and Kim et al.^[48,49] A relatively low electric field is required (five orders of magnitude smaller than the fields required for PVDF-TrFE and dielectric elastomers) to stimulate bending in IPMC, where the base polymer provides channels for mobility of positive ions in a fixed network of negative ions on interconnected clusters. In order to chemically deposit electrodes on the polymer films, metal ions (platinum, gold, palladium, or others) are dispersed throughout the hydrophilic regions of the polymer surface and are subsequently reduced to the corresponding zero-valence metal atoms.

Conductive Polymers (CPs)

Conductive polymers operate under an electric field by the reversible counterion insertion and expulsion that occurs during redox cycling.^[6,50] Oxidation and reduction occur at the electrodes, inducing a considerable volume change due mainly to the exchange of ions with an electrolyte. When a voltage is applied between the electrodes, oxidation occurs at the anode and reduction at the cathode. Ions (H^+) migrate between the electrolyte and the electrodes to balance the electric charge. Addition of the ions causes swelling of the polymer and conversely their removal results in shrinkage. As a result, the sandwich assembly bends. Conductive polymer actuators generally require small electric fields in the range of 1–5 V/ μm , and the speed increases with the voltage having relatively high mechanical energy densities of over 20 J/cm³, but with low efficiencies at the level of 1%. In recent years, several conductive polymers have been reported, including polypyrrole, polyethylenedioxythiophene, poly(p-phenylene vinylene)s, polyanilines, and polythiophenes. Operation of conductive polymers as actuators at the single-molecule level is currently being studied, taking advantage of the intrinsic electroactive property of individual polymer chains.

SMA ARTIFICIAL MUSCLES

Since SMA artificial muscle deformation is temperature-induced and rather small (4–6% linear strain), design

means must be incorporated to create larger strains to mimic biological muscles. For example, the human biceps contract more than 30%. Thus, the reader is referred to some design methodologies to create large-strain SMA muscles for biomedical applications.^[13,25–27] These papers discuss a new design for a rotatory joint actuator made with shape memory alloy contractile muscle wires. These papers also present a design of SMA rotatory joint actuators using both shape memory effect (SME) and pseudoelastic effect (PEE). The novelty of such mechanisms is the use of a pseudoelastic SMA wire as a replacement of the bias spring in a bias force type of SMA actuator. Besides the development of a general design methodology of bias force SMA actuators, these efforts by Wang and Shahinpoor intend to investigate a knee and leg muscle exerciser to be used by paraplegics or quadriplegics for exercising their knee and leg muscles. Refer to a dissertation by Guoping Wang^[28] for a complete and thorough coverage of SMA artificial muscles.

ELECTROACTIVE POLYACRYLONITRILE (PAN) FIBERS

Activated polyacrylonitrile (PAN) fibers, which are suitably annealed, cross-linked, and hydrolyzed, are known to contract and expand when ionically activated with cations and anions, respectively. The key engineering features of PAN fibers are their ability to change their length routinely more than 100% and their comparable strength to biological muscles. Described here are some recent results on electroactive PAN artificial muscles as well as a technique that allows one to electrically control the actuation of active PAN fiber bundles. Increasing the conductivity of PAN fibers by making a composite of them with a conductive medium such as platinum, gold, graphite, carbon nanotubes, and conductive polymers such as polyaniline or polypyrrole has allowed for electric activation of PAN fibers when a conductive polyacrylonitrile (C-PAN) fiber bundle is placed in a chemical electrolysis cell as an electrode. A change in concentration of cations in the vicinity of C-PAN fiber electrode leads to contraction and expansion of C-PAN fibers depending upon the applied electric field polarity. Typically close to 100% change in C-PAN length in a few seconds is observed in a weak electrolyte solution with tens of volts of DC power supply. These results indicate a great potential in developing electrically activated C-PAN muscles and linear actuators, as well as integrated pairs of antagonistic muscles and muscle sarcomere and myosin/actin assembly resembling actual biological muscles.

Activated polyacrylonitrile (PAN) fibers, which are suitably annealed, cross-linked, and hydrolyzed, are known to contract and expand when ionically activated with cations and anions, respectively. The change in length for these pH-activated fibers is typically greater than 100%, but up to 200% contraction/expansion of PAN fibers has been observed in our laboratories at AMRI. They are comparable in strength to human muscles and have the potential of becoming medically implantable electroactive contractile artificial muscle fibers.^[5,19,38,51] Increasing the conductivity of PAN by making a composite with a conductive medium (e.g., a noble metal such as gold, palladium, platinum, or graphite or a conductive polymer such as polypyrrole or polyaniline) has allowed for electrical activation of PAN artificial muscles when it is placed in an electrochemical cell. The electrolysis of water in such a cell produces hydrogen ions at a PAN anode, thus locally decreasing the cationic concentration and causing the PAN muscle to contract. Reversing the electric field allows the PAN muscle to elongate. Typically, close to 100% change in PAN muscle length in a few minutes is observed when it is placed as an electrode in a 10 mM NaCl electrolyte solution and connected to a 20-volt power supply. Recently, the response time has been reduced to a few seconds. These results indicate the potential in developing electrically active PAN artificial muscles and linear actuators.

Activated PAN contracts when exposed to protons (H^+) in an aqueous medium and elongates when exposed to hydroxyl ion (OH^-) in a strong alkaline medium. The length of activated PAN fibers can potentially more than double when going from short to long. A possible explanation for the contraction and elongation of activated PAN is the effect carboxylic acid groups have in the molecular geometry. At low cationic concentration, all acid groups are protonated, potentially contracting the polymer. Raw PAN fibers, which are composed of roughly 2000 individual strands of PAN (each about $10\mu m$ in diameter), are first annealed at $220^\circ C$ for 2 hours. The fibers can be bundled together at this point to form a PAN muscle. The PAN is then placed in a solution of boiling 1 N LiOH for 30 minutes, after which the PAN is elastic like a rubber band. At this point, the PAN can be ionically activated; flooding the activated PAN with a high concentration of cations such as H^+ induces contraction, while anionic concentration elongates the fibers. An assortment of such PAN muscles is shown in Fig. 1.

Polyacrylonitrile (PAN) has been studied for more than a half century in many institutions and now is popularly used in the textile industry as a fiber form of artificial silk. The useful properties of PAN include the insolubility, thermal stability, and resistance to



Fig. 1 An assortment of PAN artificial muscles in a variety of configurations.

swelling in most organic solvents. Such properties were thought to be due to the cross-linked nature of the polymer structure. However, a recent finding has shown that a number of strong polar solvents can dissolve PAN and thus has raised a question that its structure could be of the linear zigzag conformations^[52,53] with hydrogen bonding between hydrogen and the neighboring nitrogen of the nitrile group. Activated PAN fibers are elastic. The term “activated PAN fibers” refers to the PAN fibers acting as the elastic fibers of which length varies depending upon the ionic concentration of cations in the solution. They contract at low cationic concentration and expand at high cationic concentration. At high cationic concentration they appear to reject water out of the polymer network resulting in shrinkage due to their elastic nature. Conversely, they elongate at low cationic concentration, attracting water from outside into the polymer network. In our laboratory, it has been observed that activated PAN fibers can contract more than 200% relative to that of the expanded PAN. The use of PAN as artificial muscles is promising since they are able to convert chemical energy to mechanical motion, possibly acting as artificial sarcomeres or muscles. Other types of materials that have a capability to be electroactive are polyelectrolyte gels, such as

polyacrylamide (PAM), polyvinylalcohol-polyacrylic acid (PAA-PVA), and poly(2-acrylamido-2-methyl propane) sulphonic acid (PAMPS). Under an electric field, these gels are able to swell and de-swell, inducing large changes in the gel volume. Such changes in volume can then be converted to mechanical work. Yet one disadvantage of such polyelectrolyte gels is that they are mechanically weak. Artificial sarcomeres or muscles made from polyacrylonitrile (PAN) have much greater mechanical strength than polyelectrolyte gels, thus having a greater potential for application as artificial sarcomeres and muscles (or soft actuators).

The strength of activated PAN and its ability to change length up to 100% or more makes it an appealing material for use as linear actuators and artificial muscles. An attractive alternative is electrical activation. Shahinpoor and coworkers^[5,19,38,51] have reported considerable efforts on electrical activation of PAN artificial muscles. During the electrolysis of water, hydrogen ions are generated at the anode while hydroxyl ions are formed at the cathode in an electrochemical cell. Electrochemical reactions can then potentially be used to control the length of a PAN artificial muscle. This may be achieved by either locating a PAN muscle near an electrode where the ions are generated, or, if the conductivity of activated PAN can be increased, the PAN muscle can serve as the electrode itself.

The study reported herein takes the second approach, where platinum is deposited on PAN fibers to increase conductivity so the muscle can serve as the electrode directly. This procedure of depositing Pt on the polymer and activating it in an electrochemical cell was initially demonstrated a few years ago with a polyvinyl alcohol-polyacrylic acid copolymer, resulting in about a 5% decrease in length, with both contraction and elongation each taking about 12 minutes. However, our recent improvements^[19,38,51] to electroactive PAN muscles have reduced the activation time to a few seconds with muscle strength approaching human muscles in the range of 2 N per cm² of muscle fiber cross section.

In order to manufacture pH-active PAN fibers, the raw fibers are first annealed in air in a computer-controlled oven at an elevated temperature of 220–240°C. It is likely that the annealing process cross-links the polymer and creates a cross-linked structure of pyridine and cyano groups. The preoxidized or annealed PAN fibers show a dark brown or black color depending upon the level of cross-linking. Second, they are subsequently hydrolyzed by boiling in an alkaline solution (2 N NaOH for 20–30 minutes). Following this process, the PAN fibers become extremely pH active (pH muscles). In an approximate pH range of 2–14, the PAN fibers can contract and expand more than 200%.

Based upon the Donnan Theory of ionic equilibrium, the important forces arise from the induced osmotic pressure of free ions between activated PAN fibers and their environment, the ionic interaction of fixed ionic groups, and the charged network itself. Among them, the induced osmotic pressure of free ionic groups is the dominating force. PAN fibers can be activated electrically, by providing a conductive medium or electrode in contact with or within the PAN fibers to act like an anode electrode in a chemical cell with another cathode electrode separated from the PAN fibers. Upon applying an electric field across the electrodes, H⁺ evolves at the anode electrode via $2\text{H}_2\text{O} \Rightarrow \text{O}_2 + 4\text{H}^+ + 4e^-$, and the OH⁻ evolves at the cathode via $2\text{H}_2\text{O} + 2e^- \Rightarrow \text{H}_2 + 2\text{OH}^-$. Upon contact with H⁺ ions in the vicinity of the PAN anode electrode, the decreased pH causes the PAN fibers to contract. If the polarity of the electric field is reversed, then upon contact with OH⁻ ions in the vicinity of the cathode the increased pH causes the PAN fibers to expand.

IONIC POLYMER-CONDUCTOR COMPOSITES (IPCCs and IPMCs) AS EAPs

A large number of polyelectrolytes in a composite form with a conductive medium such as a metal (ionic polymer-conductor composites [IPCCs] or ionic polymer-metal composites [IPMCs]) can exhibit large dynamic deformation if suitable electrodes are mounted on them and placed in a time-varying electric field.^[4,5,10,13–19] Conversely, dynamic deformation of such polyelectrolytes can produce dynamic electric fields across their electrodes. A recently presented model by de Gennes, Okumura, Shahinpoor, and Kim^[54] describes the underlying principle of electrothermodynamics in such polyelectrolytes based upon internal transport phenomena and electrophoresis. It should be pointed out that IPMCs show great potential as soft robotic actuators, artificial muscles, and dynamic sensors in micro to macro size range. An effective way of manufacturing an IPMC starts with making a composite of polyelectrolytes with a metal by means of a chemical reduction process. The current state-of-the-art IPMC manufacturing technique^[5,10,14–19] incorporates two distinct preparation processes: initial manufacturing of a polyelectrolyte-metal composite, and the process of placing surface electrodes on the composite. Figure 2 shows illustrative schematics of two different preparation processes (top left and bottom left) and two top-view SEM micrographs for the platinum surface electrode (top right and bottom right).

The initial process of making a polyelectrolyte-metal composite requires an appropriate platinum salt

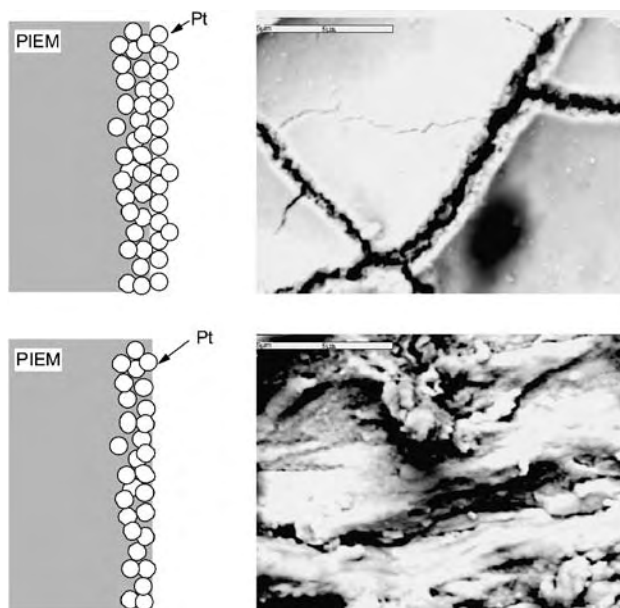
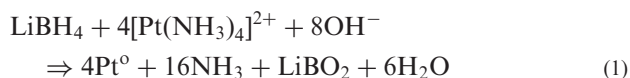


Fig. 2 Schematic diagrams showing two different preparation processes: (top left) a schematic showing initial process of making a polyelectrolyte–metal composite; (top right) its top-view SEM micrograph; (bottom left) a schematic showing the process of placing electrodes on the surface of the composite; and (bottom right) its top-view SEM micrograph where platinum deposited predominately on top of the initial Pt layer. Note that PIEM stands for perfluorinated ion exchange membrane.

such as $\text{Pt}(\text{NH}_3)_4\text{HCl}$ or $\text{Pd}(\text{NH}_3)_4\text{HCl}$ in the context of chemical reduction processes similar to the processes evaluated by a number of investigators including Takenaka et al.^[55] and Millet et al.^[56] The principle of the process of making a polyelectrolyte–metal composite is to metalize the inner surface of the material by a chemical reduction means such as LiBH_4 or NaBH_4 . The polyelectrolyte is soaked in a salt solution to allow platinum-containing cations to diffuse through via the ion-exchange process. Later, a proper reducing agent such as LiBH_4 or NaBH_4 is introduced to platinize the material.^[5,10,14–19] It has been experimentally observed that the platinum particulate layer is buried microns deep (typically 1–20 μm) within the IPMC surface

and is highly dispersed. The fabricated IPMCs can be optimized to produce a maximum force density by changing multiple process parameters. These parameters include time-dependent concentrations of the salt and the reducing agents. Applying the Taguchi technique to identify the optimum process parameters for such optimizations has been successful in the past.^[19,57] The primary reaction is,



In the subsequent process of placing electrodes on the surface of the composite, multiple reducing agents are introduced (under optimized concentrations) to carry out the reducing reaction similar to Eq. 1, in addition to the initial platinum layer formed by the initial process of making a polyelectrolyte–metal composite. This is clearly shown in Fig. 2 (bottom right), where the roughened surface disappears because of additional metals deposited on the surface. Figure 3 depicts a typical spectacular steady-state deformation mode of a strip of such ionic polymers under a step voltage.

For further references, see Refs. [45,58–70].

LIQUID CRYSTAL ELASTOMER ARTIFICIAL MUSCLES

Recently, liquid single crystal elastomers (LSCEs) have gained a lot of importance in connection with artificial muscles technologies. See Finkelmann,^[35,71] Brand and coworkers,^[72,73] Shahinpoor,^[14] Ratna,^[36] and Finkelmann and Shahinpoor^[37] where the liquid crystalline phase structure is macroscopically orientated in the network. At the liquid crystalline to isotropic phase transformation temperature, when the liquid crystalline network becomes isotropic like a conventional rubber, the dimensions of the network change. Due to the anisotropy in the phase structure of liquid crystal elastomers, the networks shorten in the direction of the optical axis or director axis. De Gennes and coworkers^[74,75] have also alluded to the fact that such nematic networks present a potential as artificial

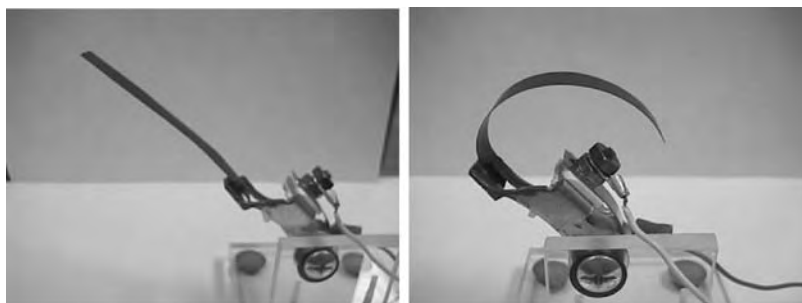


Fig. 3 Typical deformation of strips (10 mm \times 80 mm \times 0.34 mm) of ionic polymers under a step voltage of 4 volts.

muscles just like the IPMC materials mentioned before. For nematic LSCE synthesized from nematic LC side-chain polymers, the change of this length is normally more than 100% and sometimes up to 400%.^[38,51] The speed of this process is determined by the thermal heat conductivity of the network and not by material transport processes. If the sample thickness is limited, calculations indicate that the response time of the LSCE is similar to that of natural muscles and mainly determined by the relaxation behavior of the polymer chains or network strands.

The magnitude of changes in the dimensions of nematic LSCE with temperature is determined by the coupling between the nematic state of order and the conformation of the polymer main chains. In this regard, the conformation of nematic side-chain polymers is only slightly affected. For nematic main-chain polymers, where the mesogenic units are incorporated into the polymer main chain, large conformational changes should occur and be directly expressed in the thermoelastic behavior. The electrical activation of liquid crystal elastomers has also been reported by Shahinpoor^[14] and Finkelmann and Shahinpoor.^[37]

Making a composite of them with a conductor phase such as a metal or graphite powder does this. Joule heating of some new high-performance nematic LSCE developed by Finkelmann and coworkers^[35,71] that consists of a combination of LC side-chain and main-chain polymers has been successful. These materials have been synthesized previously by a hydrosilylation reaction of the monofunctional side-chain mesogen and the bifunctional liquid crystalline main-chain polyether with poly(methylhydrogensiloxane). Figure 4 displays the main chemical synthesis steps in producing liquid crystal elastomers by hydrosilylation reaction of the monofunctional side-chain mesogen. Note that a weakly cross-linked network is mechanically loaded. Under load, the hydrosilylation reaction is completed. The LSCE has a weight ratio between the mesogenic units in the main chain to those in the side chains of 57/43 and a nematic to isotropic phase transition temperature of $T_{n,i} = 104^{\circ}\text{C}$.

Thus, coelastomers containing network strands comprising LC side- and main-chain polymers exhibit exceptional thermoelastic behavior. These systems might be suitable as model systems for artificial muscles. Thus, liquid crystalline to isotropic phase transition can cause conformational changes of macromolecules, which are sufficient to obtain a large mechanical response of polymer networks.

The samples of monodomain (MD-72/28) nematic liquid side-chain crystal elastomers graphite (LSCE-G) composites under a stress of 100 kPa and contracts up to 400%.^[72] The composite of them with a conductor^[14] becomes a conductor and has an effective

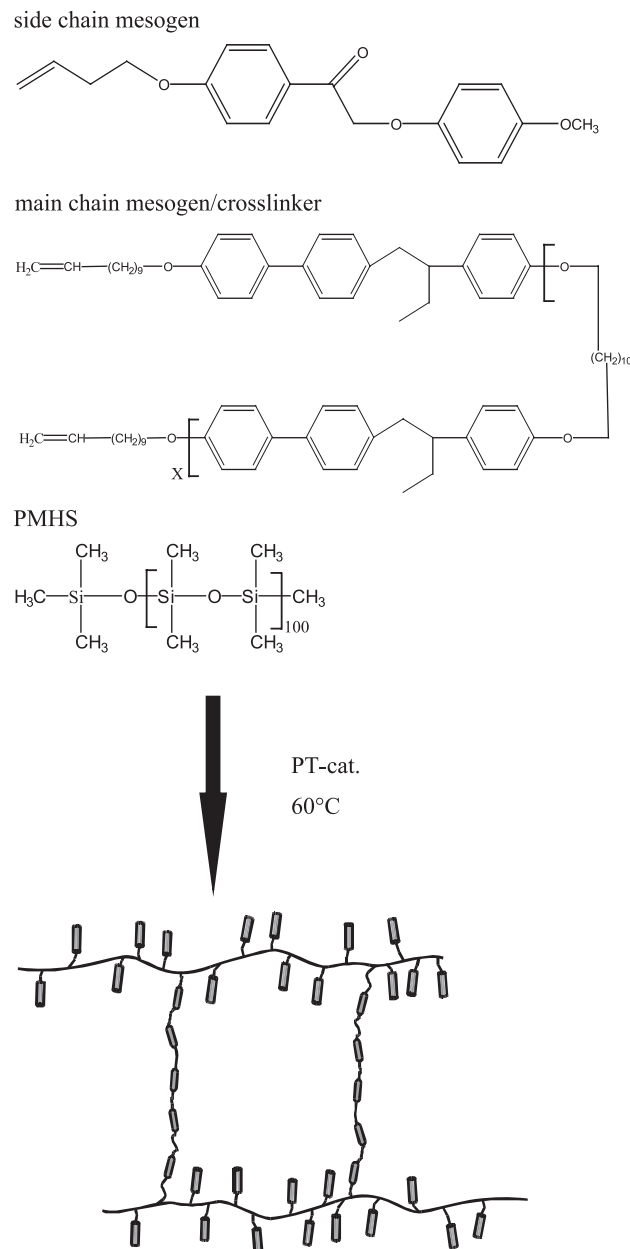


Fig. 4 Synthesis of the nematic coelastomer in a liquid single crystalline coelastomer.

resistivity of about 1.6 ohms/square on its surface and about 0.8 ohm/square across its thickness. Upon applying a DC voltage of 0.5 to 5 volts to a typical sample for about 4 seconds, under a stress of about 10 kPa, the sample contracts quickly in about a second to about 18 mm lengthwise with an average linear strain of about 25%. The sample does not contract appreciably in the transverse direction and the thickness direction. The cooling is also quick and it takes the sample about 4.4 seconds to revert back to its initial length under stretching load of 10 kPa.

CONCLUSIONS

We began this article by discussing a brief history of artificial muscles. We then discussed the shape memory alloys and polymer-type artificial muscles followed by electrical activation of PAN artificial muscles through increasing the conductivity of PAN fibers. The conductivity of PAN was increased by either depositing a coat of metal on the fibers or interweaving it with graphite fibers. Electrochemical reactions were used to generate hydrogen ions or hydroxyl ions for the contraction and elongation, respectively, of these PAN-C muscles. Therefore, by increasing the conductivity of activated PAN, a PAN artificial muscle or linear actuator could be electrically activated in an antagonistic manner (push-pull, biceps, and triceps actions) suitable for industrial and medical applications. Increasing the conductivity of PAN fibers by making a composite of them with a conductive medium such as platinum, gold, graphite, carbon nanotubes, and conductive polymers such as polyaniline or polypyrrole was shown to allow for electric activation of PAN fibers when a conductive polyacrylonitrile (C-PAN) fiber bundle is placed in a chemical electrolysis cell as an electrode. Typically close to 100% change in C-PAN fiber length in a few seconds is observed in a weak electrolyte solution with tens of volts of DC power supply. These results indicate a great potential in developing electrically activated C-PAN muscles and linear actuators, as well as integrated pairs of antagonistic muscles and muscle sarcomere and myosin/actin assembly. The history of ionic polymer-metal composites (IPMCs) in artificial muscles was also discussed. The fundamental properties and characteristics of ionic polymer (polyelectrolyte)-metal composites as biomimetic sensors, actuators, transducers, and artificial muscles were also presented. Finally, a brief discussion was presented on liquid crystal elastomer artificial muscles. It was reported that liquid crystal artificial muscles have the capability of contracting more than 400%. For additional references on artificial muscles, please refer to Refs. [76-91].

ARTICLES OF FURTHER INTEREST

Conductive Polymers; Electroactive Polymeric Materials; Piezoelectric Materials; Shape Memory Metals; Tissue Engineering of Skeletal Muscle

REFERENCES

1. Katchalsky, A. Rapid swelling and deswelling of reversible gels of polymeric acids by ionization. *Experientia* **1949**, *V*, 319-320.
2. Kuhn, W.; Hargitay, B.; Katchalsky, A.; Eisenburg, H. Reversible dilatation and contraction by changing the state of ionization of high-polymer acid networks. *Nature* **1950**, *165*, 514-516.
3. Steinberg, I.Z.; Oplatka, A.; Katchalsky, A. Mechanochemical engines. *Nature* **1966**, *210*, 568-571.
4. Shahinpoor, M. Conceptual design, kinematics and dynamics of swimming robotic structures using ionic polymeric gel muscles. *Smart Mater. Struc.* **1992**, *1* (1), 91-94.
5. Shahinpoor, M.; Bar-Cohen, Y.; Simpon, J.O.; Smith, J. Ionic polymer-metal composites (IPMC) as biomimetic sensors and structures—A review. *Smart Mater. Struc.* **1998**, *7*, 15-30.
6. Otero, T.F.; Grande, H.; Rodriguez, J. A new model for electrochemical oxidation of polypyrrole under conformational relaxation control. *J. Electr. Chem.* **1995**, *394*, 211-216.
7. Kishi, R.; Ichijo, H.; Hirasa, O. Thermo-responsive devices using poly(vinyl methyl ether) hydrogels. *J. Intell. Mater. Syst. Struct.* **1993**, *4*, 533-537.
8. Tobushi, H.; Hayashi, S.; Kojima, S. Mechanical properties of shape memory polymer of polyurethane series. *JSAE Int. J., Ser. 1* **1992**, *35* (3), 296-302.
9. Li, F.K.; Zhu, W.; Zhang, X.; Zhao, C.T.; Xu, M. Shape memory effect of ethylene-vinyl acetate copolymers. *J. Appl. Polym. Sci.* **1999**, *71* (7), 1063-1070.
10. Shahinpoor, M.; Kim, K.J. Ionic polymer-metal composites—I. Fundamentals. *Smart Mater. Struc. Int. J.* **2001**, *10*, 819-833.
11. Van der Veen, G.; Prins, W. Light-sensitive polymers, nature. *Phys. Sci.* **1971**, *230*, 70-72.
12. Zrinyi, M.; Barsi, L.; Szabo, D.; Kilian, H.G. Direct observation of abrupt shape transition in ferrogels induced by nonuniform magnetic field. *J. Chem. Phys.* **1997**, *106* (13), 5685-5692.
13. Shahinpoor, M.; Wang, G. In *Design, Modeling and Performance Evaluation of a Novel Large Motion Shape Memory Alloy Actuator*, Proc. SPIE 1995 North American Conference on Smart Structures and Materials, San Diego, CA, February 28-March 2, 1995; Vol. 2447, paper no. 31.
14. Shahinpoor, M. In *Elastically-Activated Artificial Muscles Made with Liquid Crystal Elastomers*, Proceedings of the SPIE's 7th Annual International Symposium on Smart Structures and Materials, EAPAD Conf., Bar-Cohen, Y., Ed.; 2000; Vol. 3987, 187-192.
15. Shahinpoor, M.; Kim, K.J. A solid-state soft actuator exhibiting large electromechanical effect. *Appl. Phys. Lett. (APL)* **2002**, *80* (18), 3445-3447.
16. Shahinpoor, M.; Mojarrad, M. Soft Actuators and Artificial Muscles. U.S. Patent, #6,109,852, April 13, 2000.
17. Shahinpoor, M.; Kim, K.J. The effect of surface-electrode resistance on the performance of ionic polymer-metal composites (IPMC) artificial muscles. *Smart Mater. Struc.* **2000**, *9*, 543-551.
18. Shahinpoor, M.; Kim, K.J. Novel ionic polymer-metal composites equipped with physically-loaded particulate electrode as biomimetic sensors, actuators and artificial muscles. *Actuators Sens. A, Phys., A* **2002**, *3163*, 1-8.
19. Shahinpoor, M.; Norris, I.D.; Mattes, B.R.; Kim, K.J.; Sillerud, L.O. In *Electroactive Polyacrylonitrile*

- Nanofibers as Artificial Nano-Muscles*, Proceedings of the 2002 SPIE Conference on Smart Mater. Struc., San Diego, CA, March 2002; Vol. 4695 (42), 169–173.
20. Zhang, Q.M.; Bharti, V.; Zhao, X. Giant electrostriction and relaxor ferroelectric behavior in electron-irradiated poly(vinylidene fluoride-trifluoroethylene) copolymer. *Science* **1998**, *280*, 2101–2104.
 21. Perline, R.; Kornbluh, R.; Joseph, J.P. Electrostriction of polymer dielectrics with compliant electrodes as a means of actuation. *Sens. Actuators*, **4** **1998**, *64*, 77–85.
 22. Liang, C.; Rogers, C.A. One-dimensional thermometrical constitutive relations of shape memory materials. *J. Intell. Mater. Syst. Struct.*, **1991**, *1* (2), 207–234.
 23. Liang, C.; Rogers, C.A. Design of shape memory alloy actuators. *ASME J. Mech. Des.* **1992**, *114*, 223–230.
 24. Liang, C.; Rogers, C.A.; Malafeev, E. Investigation of shape memory polymers and their hybrid composites. *J. Intell. Mater. Syst. Struct.* **1997**, *8*, 380.
 25. Wang, G.; Shahinpoor, M. Design for shape memory alloy rotatory joint actuators using shape memory effect and pseudoelastic effect. *Smart Mater. Technol.*, SPIE Publication No. **1997**, *3040*, 23–30.
 26. Wang, G.; Shahinpoor, M. Design, prototyping and computer simulation of a novel large bending actuator made with a shape memory alloy contractile wire. *Smart Mater. Struc., Int. J* **1997**, *6* (2), 214–221.
 27. Wang, G.; Shahinpoor, M. A new design for a rotatory joint actuator made with shape memory alloy contractile wire. *Int. J. Intell. Mater. Syst. Struct.* **March 1997**, *8* (3), 191–279.
 28. Wang, G. A General Design of Bias Force Shape Memory Alloy (BFSMA) Actuators and An Electrically-Controlled SMA Knee and Leg Muscle Exerciser for Paraplegics and Quadriplegics. In *Ph.D.*; Department of Mechanical Engineering, University of New Mexico: Albuquerque, NM, May 1998.
 29. Zrinyi, M.; Szabo, D.; Feher, J. In *Comparative Studies of Electro- and Magnetic Field Sensitive Polymer Gels*, Proceedings of the SPIE's 6th Annual International Symposium on Smart Structures and Materials, EAPAD Conf., SPIE Proc., Bar-Cohen, Y., Ed.; 1999; Vol. 3669, 406–413.
 30. Dapino, M.; Flatau, A.; Calkins, F. In *Statistical Analysis of Terfenol-D Material Properties*, Proceedings of SPIE 1997 Symposium on Smart Structures and Materials, San Diego, CA, March 1997; Vol. 3041, 256–267.
 31. O'Handley, R.C. Model for strain and magnetostriction in magnetic shape memory alloys. *J. Appl. Phys.* **1998**, *83*, 3263.
 32. Sen, A.; Scheinbeim, J.I.; Newman, B.A. The effect of plasticizer on the polarization of poly(vinylidene fluoride) films. *J. Appl. Phys.* **1984**, *56* (9), 2433–2439.
 33. Furukawa, T.; Seo, N. Electrostriction as the origin of piezoelectricity in ferroelectric polymers. *Jpn. J. Appl. Phys.* **1990**, *29* (4), 675–680.
 34. Pelrine, R.; Kornbluh, R.; Pei, Q.; Joseph, J. High speed electrically actuated elastomers with strain greater than 100%. *Science* **2000**, *287*, 836–839.
 35. Finkelmann, H.; Kock, H.J.; Rehage, G. Investigations on liquid crystalline siloxanes: 3. Liquid crystalline elastomer—A new type of liquid crystalline material. *Makromol. Chem., Rapid Commun.* **1981**, *2*, 317–323.
 36. Ratna, B.R.; Selinger, J.V.; Srinivasan, A.; Hong, J.; Naciri, J. In *Nematic Elastomers as Artificial Muscles*, Proceedings of the First World Congress on Biometrics and Artificial Muscles, Albuquerque Conventional Center, Albuquerque, New Mexico, USA, December 9–11, 2002, 2003; 1.
 37. Finkelmann, H.; Shahinpoor, M. In *Electrically-Controllable Liquid Crystal Elastomer-Graphite Composites Artificial Muscles*, Proceeding of SPIE 9th Annual International Symposium on Smart Structures and Materials, 2002, San Diego, California, SPIE Publication No. 4695-53.
 38. Schreyer, H.B.; Gebhart, N.; Kim, K.J.; Shahinpoor, M. Electric activation of artificial muscles containing polyacrylonitrile gel fibers. *Biomacromolecules J.*, ACS Publications **2000**, *1*, 642–647.
 39. Osada, Y.; Ross-Murphy, S. Intelligent gels. *Sci. Am.* **1993**, *268*, 82–87.
 40. Gong, J.P.; Nitta, T.; Osada, Y. Electrokinetic modeling of the contractile phenomena of polyelectrolyte gels-one dimensional capillary model. *J. Phys. Chem.* **1994**, *98*, 9583–9587.
 41. Osada, Y.; Hasebe, M. Electrically activated mechanochemical devices using polyelectrolyte gels. *Chem. Lett.* **1985**, *12*, 1285–1288.
 42. Osada, Y.; Kishi, R. Reversible volume change of microparticles in an electric field. *J. Chem. Soc.* **1989**, *85*, 662–665.
 43. Hirai, M.; Hirai, T.; Sukumoda, A.; Nemoto, H.; Amemiya, Y.; Kobayashi, K.; Ueki, T. Electrically induced reversible structural change of a highly swollen polymer gel network. *J. Chem. Soc., Faraday Trans.* **1995**, *91*, 473–477.
 44. Hirai, T.; Zheng, J.; Watanabe, M. In *Solvent-Drag Bending Motion of Polymer Gel Induced by an Electric Field*, Proceedings of the SPIE's 6th Annual International Symposium on Smart Structures and Materials, Bar-Cohen, Y., Ed.; 1999; Vol. 3669, 209–217.
 45. Oguro, K.; Kawami, Y.; Takenaka, H. Bending of an ion-conducting polymer film-electrode composite by an electric stimulus at low voltage. *Trans. J. Micromach. Soc.* **1992**, *5*, 27–30.
 46. Sadeghipour, K.; Salomon, R.; Neogi, S. Development of a novel electrochemically active membrane and smart material based vibration sensor/damper. *Smart Mater. Struc.* **1992**, *1* (2), 172–179.
 47. Heitner-Wirguin, C. Recent advances in perfluorinated ionomer membranes: Structure, properties and applications. *J. Membr. Sci.* **1996**, *120* (1), 1–33.
 48. Kim, K.J.; Shahinpoor, M. Application of Polyelectrolytes in Ionic Polymeric Sensors, Actuators, and Artificial Muscles. In *Handbook of Polyelectrolytes*; Tripathy, S., Nalwa, H.S., Eds.; Mindspring/Academic Press, 2002. in print, Review.
 49. Kim, K.J.; Shahinpoor, M. A novel method of manufacturing three-dimensional ionic polymer-metal composites (IPMC's) biomimetic sensors, actuators and artificial muscle. *Polymer* **2002**, *43/3*, 797–802.

50. Gandhi, M.R.; Murray, P.; Spinks, G.M.; Wallace, G.G. Mechanisms of electromechanical actuation in polypyrrole. *Synth. Met.* **1995**, *75*, 247–256.
51. Schreyer, H.B.; Shahinpoor, M.; Kim, K.J. In *Electrical Activation of PAN-Pt Artificial Muscles*, Proceedings of SPIE/Smart Structures and Materials/Electroactive Polymer Actuators and Devices, Newport Beach, CA, March 1999; Vol. 3669, 192–198.
52. Hu, X. Molecular structure of polyacrylonitrile fibers. *J. Appl. Polym. Sci.* **1996**, *62*, 1925–1932.
53. Umemoto, S.; Okui, N.; Sakai, T. *Contraction Behavior of Poly(acrylonitrile) Gel Fibers, Polymer Gels*; DeRossi, D., et al., Eds.; Plenum Press: New York, 1991; 257–270.
54. De Gennes, P.G.; Okumura, K.; Shahinpoor, M.; Kim, K.J. Mechanoelectric effects in ionic gels. *Europhys. Lett.* **2000**, *50* (4), 513–518.
55. Takenaka, H.; Torikai, E.; Kawami, Y.; Wakabayashi, N. Solid polymer electrolyte water electrolysis. *Int. J. Hydrogen Energy* **1982**, *7*, 397–403.
56. Millet, P.; Pinneri, M.; Durand, R. New solid polymer electrolyte composites for water electrolysis. *J. Appl. Electrochem.* **1989**, *19*, 162–166.
57. Peace, G.S. *Taguchi Methods: Hands-On Approach*; Addison-Wesley Publishing Company, Inc.: New York, 1993.
58. Adolf, D.; Shahinpoor, M.; Segalman, D.; Witkowski, W. Electrically Controlled Polymeric Gel Actuators. U.S. Patent #5,250,167, October 13, 1993.
59. Brock, D.L. *Review of Artificial Muscle Based on Contractile Polymers*; AI Memo No. 1330, MIT, Artificial Intelligence Lab. 1991.
60. De Rossi, D.E.; Chiarelli, P.; Buzzigoli, G.; Domenici, C.; Lazzari, L. Contractile behavior of electrically activated mechanochemical polymer actuators. *Trans. Am. Soc. Artif. Intern. Organs* **1986**, *32*, 157–163.
61. De Rossi, D.; Suzuki, M.; Osada, Y.; Morasso, P. Pseudomuscular gel actuators for advanced robotics. *J. Intell. Mater. Syst. Struct.* **1992**, *3*, 75–95.
62. Doi, M.; Matsumoto, M.; Hirose, Y. Deformation of ionic polymer gels by electric fields. *Macromolecules* **1992**, *25*, 5504–5511.
63. Hamlen, R.P.; Kent, C.E.; Shafer, S.N. Electrolytically activated contractile polymer. *Nature* **1965**, *206*, 1149–1150.
64. Nemat-Nasser, S.; Li, J.Y. Electromechanical response of ionic polymer-metal composites. *J. Appl. Phys.* **2000**, *87* (7), 3321–3331.
65. Oguro, K.; Kawami, Y.; Takenaka, H. Actuator Element. U.S. Patent #5,268,082, 1993.
66. Oguro, K.; Fujiwara, N.; Asaka, K.; Onishi, K.; Sewa, S. Polymer Electrolyte Actuator with Gold Electrodes, Proceedings of the SPIE's 6th Annual International Symposium on Smart Structures and Materials, SPIE Proc., 1999; Vol. 3669, 64–71.
67. Pei, Q.; Ingnas, O.; Lundstrom, I. Bending bilayer strips built from polyaniline for artificial electrochemical muscles. *Smart Mater. Struct.* **1993**, *2* (1), 1–5.
68. Segalman, D.; Witkowski, W.; Adolf, D.; Shahinpoor, M. In *Electrically Controlled Polymeric Muscles as Active Materials Used in Adaptive Structures*, Proceedings of ADPA/AIAA/ASME/SPIE Conference on Active Materials and Adaptive Structures, Alexandria, VA, Nov. 1991.
69. Tanaka, T.; Nishio, I.; Sun, S.T.; Nishio, S.U. Collapse of gels in an electric field. *Science* **1982**, *218*, 467–469.
70. Woojin, L. Polymer Gel Based Actuator: Dynamic Model of Gel for Real Time Control. In *Ph.D. Thesis*; Massachusetts Institute of Technology, 1996.
71. Finkelmann, H.; Brand, H.R. Liquid crystalline elastomers—A class of materials with novel properties. *Trends Polym. Sci.* **1994**, *2*, 222.
72. Brand, H.R.; Finkelmann, H. Physical Properties of Liquid Crystalline Elastomers. In *Handbook of Liquid Crystals Vol. 3: High Molecular Weight Liquid Crystals*; Demus, D., Goodby, J., Gray, G.W., Spiess, H.-W., Vill, V., Eds.; Wiley-VCH: Weinheim, 1998; 277–289.
73. Brandt, H.R.; Pleiner, H. Electrohydrodynamics of nematic liquid crystalline elastomers. *Physica* **1994**, *A208*, 359.
74. De Gennes, P.G.; Hebert, M.; Kant, R. Artificial muscles based on nematic gels. *Macromol. Symp.* **1997**, *113*, 39.
75. Hebert, M.; Kant, R.; de Gennes, P.G. Dynamics and thermodynamics of artificial muscles based on nematic gels. *J. Phys., I, France* **1997**, *7*, 909–918.
76. Shahinpoor, M. “Ionic polymer-conductor composites as biomimetic sensors, Robotic actuators and artificial muscle—A Review.” *Electrochimica Acta*, **2003**, *48* (14–16), 2343–2353.
77. *Advances in Polymer Science, Vols. 109 and 110, Responsive Gels, Transitions I and II*; Dusek, K., Ed.; Springer-Verlag: Berlin, 1993.
78. Li, J.Y.; Nemat-Nasser, S. Micromechanical analysis of ionic clustering in nafion perfluorinated membrane. *Mech. Mater.* **2000**, *32* (5), 303–314.
79. Liu, Z.; Calvert, P. Multilayer hydrogens and muscle-like actuators. *Adv. Mater.* **2000**, *12* (4), 288–291.
80. McGehee, M.D.; Miller, E.K.; Moses, D.; Heeger, A.J. Twenty Years of Conductive Polymers: From Fundamental Science to Applications. In *Advances in Synthetic Metal: Twenty Years of Progress in Science and Technology*; Bamier, P., Lefrant, S., Bidan, G., Eds.; Elsevier, 1999; 98–203. published.
81. Osada, Y.; Matsuda, A. Shape memory in hydrogels. *Nature* **1995**, *376*, 219.
82. Osada, Y.; Okuzaki, H.; Hori, H. A polymer gel with electrically driven motility. *Nature* **1992**, *355*, 242–244.
83. Otero, T.F.; Sansifena, J.M. Soft and wet conducting polymers for artificial muscles. *Adv. Mater.* **1998**, *10* (6), 491–494.
84. Roentgen, W.C. About the changes in shape and volume of dielectrics caused by electricity. *Ann. Phys. Chem.* **1880**, *1* (1), 771–786.
85. Sacerdote, M.P. Strains in polymers due to electricity. *J. Phys.* **1899**, *VIII* (3 Series, t), 31.
86. Shiga, T. Deformation and viscoelastic behavior of polymer gels in electric fields. *Adv. Polym. Sci.* **1997**, *134*, 131–163.

87. Shiga, T.; Kurauchi, T. Deformation of polyelectrolyte gels under the influence of electric field. *J. Appl. Polym. Sci.* **1990**, *39*, 2305–2320.
88. Shiga, T.; Hirose, Y.; Okada, A.; Kurauchi, T. Bending of ionic polymer gel caused by swelling under sinusoidally varying electric fields. *J. Appl. Polym. Sci.* **1993**, *47*, 113–119.
89. *The Applications of Ferroelectric Polymers*; Wang, T.T., Herbert, J.M., Glass, A.M., Eds.; Chapman and Hall: New York, 1988.
90. Wang, Q.; Du, X.; Xu, B.; Cross, L.E. Electro-mechanical coupling and output efficiency of piezoelectric bending actuators *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **1999**, *46* (3), 638–646.
91. Proceedings of the Fall MRS Symposium on Electroactive Polymers (EAP), Warrendale, PA, Zhang, Q.M., Furukawa, T., Bar-Cohen, Y., Scheinbeim, J., Eds.; 1999; 1–336.

Artificial Neural Networks: An Overview

A

Evangelia Micheli-Tzanakou

Department of Biomedical Engineering, Rutgers University, Piscataway, New Jersey, U.S.A.

INTRODUCTION

Neural networks have been a much publicized topic of research in recent years and are now beginning to be used in a wide range of subject areas. One of the strands of interest in neural networks is to explore possible models of biological computation. Human brains contain about 1.5×10^{11} neurons of various types, with each receiving signals through 10 to 10^4 synapses. The response of a neuron happens in about 1 to 10 msec. Yet, one can recognize an old friend's face and call him in about 0.1 sec. This is a complex pattern recognition task that must be performed in a highly parallel way, because the recognition occurs in about 100 to 1000 steps. This indicates that highly parallel systems can perform pattern recognition tasks more rapidly than current conventional sequential computers. As yet, very-large-scale integration (VLSI) technology, which is essential planar implementation with at most two or three layer cross-connections, is far from achieving these parallel connections that require 3-D interconnections.

Artificial Neural Networks

Although originally the neural networks were intended to mimic a task-specific subsystem of a mammalian or human brain, recent research has mostly concentrated on the ANNs, which are only vaguely related to the biological system. Neural networks are specified by 1) the net topology; 2) node characteristics; and 3) training or learning rules.

Topological consideration of the Artificial Neural Networks (ANNs) for different purposes is provided in review papers.^[1,2] Because our interests in the neural networks are in classification only, the feed-forward multilayer perceptron topology is considered, leaving the feedback connections to the references.

The topology describes the connection with the number of the layers and the units in each layer for feed-forward networks. Node functions are usually non-linear in the middle layers but can be linear or non-linear for output layer nodes. However, all of the units in the input layer are linear and have fan-out connections from the input to the next layer.

Each output y_j is weighted by w_{ij} and summed at the linear combiner, represented by a small circle in Fig. 1. The linear combiner thresholds its inputs before it sends them to the node function ϕ_j . The unit functions are non-linear, monotonically increasing, and bounded functions as shown at the right in Fig. 1.

Usage of Neural Networks

One use of a neural network is classification. For this purpose, each input pattern is forced adaptively to output the pattern indicators that are part of the training data; the training set consists of the input covariate x and the corresponding class labels. Feed forward networks, sometimes called multilayer perceptrons (MLP), are trained adaptively to transform a set of input signals, X , into a set of output signals, G . Feedback networks begin with the initial activity state of a feedback system. After state transitions have taken place the asymptotic final state is identified as the outcome of the computation. One use of the feedback networks is in the case of associative memories: upon being presented with a pattern near a prototype X it should output pattern X' , and as autoassociative memory or contents addressable memory by which the desired output is completed to become X .

In all cases the network learns or is trained by the repeated presentation of patterns with known required outputs (or pattern indicators). Supervised neural networks find a mapping $f: X \rightarrow G$ for a given set of input and output pairs.

Other Neural Networks

The other dichotomy of the neural networks family is unsupervised learning, that is clustering. The class information is not known or is irrelevant. The networks find groups of similar input patterns.

The neighboring code vectors in a neural network compete in their activities by means of mutual lateral interactions and develop adaptively into specific detectors of different signal patterns. Examples are the Self Organizing Map^[3] and the Adaptive Resonance Theory (ART) networks.^[4] The ART is different from other unsupervised learning networks in that it develops new clusters by itself. The network develops a

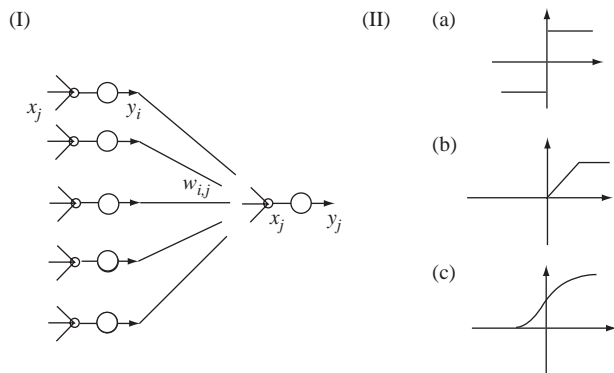


Fig. 1 (I) The linear combiner output $x_j = \sum_{i=1}^n y_i w_{ij}$ is

input to the node function ϕ_j given the output y_j .
 (II) Possible node functions: hard limiter (a), threshold (b), and sigmoid (c) non-linear function.

new code vector if there exist sufficiently different patterns. Thus, the ART is truly adaptive, whereas other methods require the number of clusters to be specified in advance.

Feed-Forward Networks

In feed-forward networks the signal flows only in the forward direction; no feedback exists for any node. This is perhaps best illustrated graphically in Fig. 2. This is the simplest topology and is good enough for most practical classification problems.^[5]

The general definition allows more than one hidden layer and also allows “skip layer” connections from

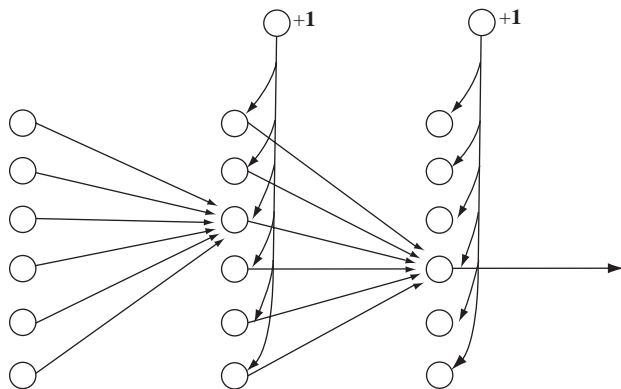


Fig. 2 A generic feed-forward network with a single hidden layer. For bias terms the constant inputs with 1 are shown and the weights of the constant inputs are the bias values, which will be learned as training proceeds.

input to output. With this skip layer, one can write a general expression for a network output y_k with one hidden layer,

$$y_k = \phi_k \left\{ b_k + \sum_{i \rightarrow k} w_{ik} x_i + \sum_{j \rightarrow k} w_{jk} \phi_j \left\{ b_j + \sum_{i \rightarrow j} w_{ij} x_i \right\} \right\}, \quad (1)$$

where the b_j and b_k represent the thresholds for each unit in the j th hidden layer and the output layer, which is the k th layer. Because the threshold values b_j , b_k are to be adaptive, it is useful to have a threshold for the weights, for constant input value of 1, as in Fig. 2. The function $\phi(\cdot)$ is almost inevitably taken to be a linear, sigmoidal ($\phi(x) = e^x / (1 + e^x)$) or threshold function ($\phi(x) = I(x > 0)$).

Rumelhart et al.^[6] demonstrated that the feed-forward multilayer perceptron networks can learn using gradient values obtained by an algorithm called error backpropagation.¹

This contribution is a remarkable advance since 1969, when Minsky and Papert^[7] claimed that the non-linear boundary required for the XOR problem can be obtained using a multilayer perceptron. The learning method was unknown at the time.

Rosenblatt^[8] introduced the one-layer single perceptron learning method called the *perceptron convergence* procedure and research on the single perceptron was active until the counter example of the XOR problem, which the single perceptron could not solve, was introduced.

In multilayer network learning the usual objective or error function to be minimized has the form of a squared error

$$E(\mathbf{w}) = \sum_{p=1}^P \|t^p - f(x^p; \mathbf{w})\|^2 \quad (2)$$

that is to be minimized with respect to \mathbf{w} , the weights in the network. Here p represents the pattern index, $p = 1, 2, \dots, P$ and t^p is the target or desired value when x^p is the input to the network. Clearly, this minimization can be obtained by any number of unconstrained optimization algorithms; gradient methods or stochastic optimization are possible candidates.

¹A comment on the terminology backpropagation is given in the next section. There, backpropagation is interpreted as a method of finding the gradient values of a feed-forward multilayer perceptron network rather than as a learning method. A pseudosteepest descent method is the learning mechanism used in the network.

The updating of weights has a form of the steepest descent method,

$$w_{ij} = w_{ij} - \eta \frac{\partial E}{\partial w_{ij}}, \quad (3)$$

where the gradient value $\partial E/\partial w_{ij}$ is calculated for each pattern being present; the error term $E(w)$ in the on-line learning is not the summation of the squared error for all P patterns.

Note that the gradient points are in the direction of maximum increasing error. In order to minimize the error it is necessary to multiply the gradient vector by -1 and by a learning rate η .

The updating method (Eq. 3) has a constant learning rate η for all weights and is independent of time. The original method of steepest descent has the time-dependent parameter η_k ; hence, η_k must be calculated as iterations progress.

Error Backpropagation

Backpropagation was first discussed by Bryson and Ho^[9], later researched by Werbos^[10] and Parker^[11], and rediscovered and popularized by Rumelhart et al.^[6,12] Each pattern is presented to the network and the input x_j and output y_j are calculated as in Fig. 3. The partial derivative of the error function with respect to weights is

$$\nabla E(t) = \left(\frac{\partial E(t)}{\partial w(t)} \dots \frac{\partial E(t)}{\partial w_n(t)} \right)^T \quad (4)$$

where, n is the number of weights and t is the time index representing the instance of the input pattern presented to the network.

The former indexing is for the on-line learning in which the gradient term of each weight does not accumulate. This is the simplified version of the gradient method that makes use of the gradient information of

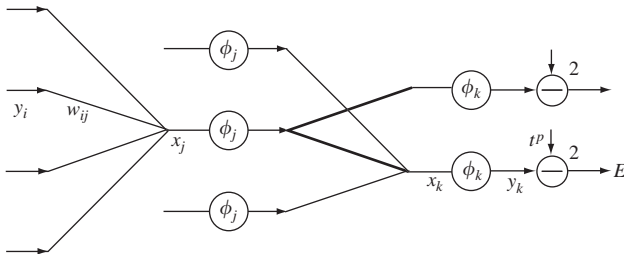


Fig. 3 Error backpropagation. The δ_j for weight w_{ij} is obtained. δ_k 's are then backpropagated via thicker weight lines, w_{jk} 's.

all training data. In other words, there are two ways to update the weights using Eq. 4.

$$w_{ij}^{(p)} \leftarrow w_{ij}^{(p)} - \eta \left(\frac{\partial E}{\partial w_{ij}} \right)^{(p)} \quad \text{temporal learning} \quad (5)$$

$$w_{ij} \leftarrow w_{ij} - \eta \sum \left(\frac{\partial E}{\partial w_{ij}} \right)^{(p)} \quad \text{epoch learning} \quad (6)$$

One way is to sum all the P patterns to get the sum of the derivatives in Eq. 6 and the other way (Eq. 5) is to update the weights for each input and output pair temporally without summation of the derivatives. The temporal learning, also called on-line learning (Eq. 5), is simple to implement in a VLSI chip, because it does not require the summation logic and storing of each weight, whereas the epoch learning in Eq. 6 does. However, the temporal learning is an asymptotic approximation version of the epoch learning, which is based on minimizing an objective function (Eq. 2).

With the help of Fig. 3, the first derivatives of E with respect to a specific weight w_{jk} can be expanded by the chain rule:

$$\frac{\partial E}{\partial w_{jk}} = \frac{\partial E}{\partial x_k} \frac{\partial x_k}{\partial w_{jk}} = \frac{\partial E}{\partial x_k} y_j = \phi'_k(x_k) \frac{\partial E}{\partial y_k} y_j \quad (7)$$

Or

$$\frac{\partial E}{\partial w_{jk}} = \frac{\partial \phi_k(x)}{\partial x_k} \frac{\partial E}{\partial y_k} y_j = \delta_k y_j. \quad (8)$$

For output units, $\partial E/\partial y_k$ is readily available, i.e., $2(y_k - t^p)$, where y_k and t^p are the network output and the desired target value for input pattern x^p . The $\phi'_k x_k$ is straightforward for the linear and logistic non-linear node functions; the hard limiter, on the other hand, is not differentiable.

For the linear node function,

$$\phi'(x) = 1 \quad \text{with} \quad y = \phi_x = x,$$

whereas for the logistic unit the first-order derivative becomes

$$\phi'(x) = y(1 - y) \quad (9)$$

$$\text{when} \quad \phi(x) = \frac{e^x}{1 + e^x} \quad (10)$$

The derivative can be written in the form

$$\frac{\partial E}{\partial w_{ij}} = \sum_p y_i^p \delta_j^p, \quad (11)$$

which has become known as the *generalized delta rule*.

The δ 's in the generalized delta rule (Eq. 11) for output nodes, therefore, become

$$\delta_k = 2y_k(1 - y_k)(y_k - t^p) \quad \text{for a logistic output unit} \quad (12a)$$

and

$$\delta_k = 2(y_p - t^p) \quad \text{for a linear output unit.} \quad (12b)$$

The interesting point in the backpropagation algorithm is that the δ 's can be computed from output to input through hidden layers across the network; δ 's for the units in earlier layers can be obtained by summing the δ 's in the higher layers. As illustrated in Fig. 3, the δ_j are obtained as

$$\begin{aligned} \delta_j &= \phi'_j(x_j) \frac{\partial E}{\partial y_j} \\ &= \phi'_j(x_j) \sum_{j \rightarrow k} w_{jk} \frac{\partial E}{\partial x_k} \\ &= \phi'_j(x_j) \sum_{j \rightarrow k} w_{jk} \delta_k \end{aligned} \quad (13)$$

The δ_k 's are available from the output nodes. As the updating or learning progresses backward, the previous (or higher) δ_k are weighted by the weights w_{jk} 's and summed to give the δ_j 's. Because Eq. 13 for δ_j only contains terms at higher layer units, it can be calculated backward from the output to the input of the network; hence, the name backpropagation.

Madaline Rule III for multilayer network with sigmoid function

Widrow took an independent path in learning as early as the 1960s.^[13,14] After some 30 years of research in adaptive filtering, Widrow and colleagues returned to neural network research^[13] and extended Madaline I with the goal of developing a new technique that could adapt multiple layers of adaptive elements using the simpler hard-limiting quantizer. The result was Madaline Rule II (or simply MR II), a multilayer linear combiner with a hard-limiting quantizer.

Andes (unpublished, 1988) modified the MR II by replacing the hard-limiting quantizer, resulting in MR III by a sigmoid function in the Madaline, i.e., a single-layer linear combiner with a hard-limiting quantizer. It was proven later that MR III is, in essence, equivalent to backpropagation. The important difference from the gradient-based backpropagation method is that the derivative of the sigmoid function is not required in this realization; thus, the analog

implementation becomes feasible with this MR III multilayer learning rule.

A comment on the terminology backpropagation

The terminology backpropagation has been used differently from what it should mean. To obtain the partial derivatives of the error function (at the system output node) with respect to the weights of the units lower than the output unit, the terms in the output unit are propagated backward, as in Eq. 13. However, the network (actually the weights) learns (or weights are updated) using the pseudo steepest descent method (Eq. 3); it is pseudo because a constant term is used, whereas the steepest descent method requires an optimal learning rate for each weight and time instance, i.e., $\eta_{ij}(k)$. The error backpropagation is indeed used to find the necessary gradient values in the updating rule. Thus, it is not a good idea to call the backpropagation a learning method: the learning method is a simple version of the steepest descent method, which is one of the classical minimizer finding algorithms. Backpropagation is an algorithm used to find the gradient of E in a feed-forward multilayer perceptron network.

Optimization machines with feed forward multilayer perceptrons

Optimization in multilayer perceptron structures can be easily realized by gradient-based optimization methods with the help of backpropagation. In the multilayer perceptron structure, the functions can be minimized/maximized via any gradient-based unconstrained optimization algorithm, such as Newton's method or the steepest descent method.

The optimization machine is functionally described as depicted in Fig. 4 and consists of two parts: gradient calculation and weight (or parameter) updating.

The gradient of E of the multilayer perceptron network is obtained by error backpropagation. If this gradient is used in an on-line fashion with the constant learning rate, as in Eq. 3, then this structure is the neural network used earlier.^[12] This on-line learning structure possesses a desirable feature in VLSI implementation of the algorithm because it is temporal: no summation over all the patterns is required but the weights are updated as the individual pattern is presented to the network. It requires little memory but sometimes the convergence is too slow.

The other branch in Fig. 4 illustrates unconstrained optimization of the non-linear function. The optimization machine obtains the gradient information as before, but various and well-developed unconstrained optimizations can be used to find the optimizer. The

unconstrained non-linear minimization is divided basically into two categories: gradient methods and stochastic optimization. The gradient methods are deterministic and use the gradient information to find the direction for the minimizer. Stochastic optimization methods such as the ALgorithm Of Pattern Extraction (ALOPEx) are discussed in another section of this chapter and in other references.^[15–17]

Comparisons of ALOPEX with backpropagation are reported by Zahner and Micheli-Tzanakou.^[15]

Justification for gradient methods for non-linear function approximation

Getting stuck in local minima is a well-known problem for gradient methods. However, the weights (or the dimensionality of the weight space in neural networks) are usually much larger than the dimensionality of the input space $X \subset R^P$ that is preferred when searching for optimization. The redundant degrees of freedom used to find a better minimizer justify the gradient methods used in neural networks.

Another justification for use of the gradient method in optimization may result from the approximation of highly nonlinear functions using Taylor expansion,^[5] where the first- and second-order approximations, i.e., a quadratic approximation to the non-linear function, are used. The quadratic function in a covariate x has a unique minimum or maximum.

Training Methods for Feed-Forward Networks

Two basic methods exist to train the feed-forward networks: gradient-based learning and stochastic learning. Training or learning is essentially an unconstrained optimization problem. A number of algorithm optimizations can be applied to the function approximated by the network in a structured way defined by the network topology.

In gradient-based methods, the most popular learning involves the steepest descent/ascent method with the error backpropagation algorithm to obtain the required gradient of the minimizing/maximizing error function with respect to the weights in the network.^[6,12] Another method using the gradient

information is Newton's method, which is basically used for zero finding of a non-linear function. The function optimization problem is the same as the zero finding of the first derivative of a function; hence, Newton's method is valid.

All the deterministic (as opposed to stochastic) minimization techniques are based on one or both of the steepest descent and Newton's methods. The objective function to be optimized is usually limited to a certain class in network optimization. The square of the error $\|t - \hat{y}\|^2$ and the information theoretic measuring the Kullback–Leibler distance are objective functions used in feed-forward networks. This is because of the limitation in calculating the gradient values of the network utilized by the error backpropagation algorithm.

The recommended method of optimization according to Broyden, Fletcher, Goldfarb, and Shannon (BFGS) is the well-known Hessian matrix update in Newton's method of unconstrained optimization,^[18] which requires gradient values. For the optimization machine in Fig. 4, the feed-forward network with backpropagation provides the gradients and the Hessian approximation is obtained using the BFGS method.

The other dichotomy of the minimization of an unconstrained non-linear multivariate function is grouped into the so-called stochastic optimization. The representative algorithms are simulated annealing,^[19] Boltzmann machine learning,^[20] and ALOPEX.^[21,22] Simulated annealing^[19] has been successfully used in combinatoric optimization problems such as the traveling salesman problem, VLSI wiring, and VLSI placement problems. An application of feed-forward network learning was reported,^[23] with the weights being constrained to be integers or discrete values rather than a continuum of the weight space.

Boltzmann machine learning, by Hinton and Sejnowski,^[20] is similar to simulated annealing except that the acceptance of randomly chosen weights is possible even when the energy state is decreased. In simulated annealing, the weights yielding the decreased energy state are always accepted, but in the Boltzmann machine probability is used in accepting the increased energy states.

The simulated annealing and the Boltzmann machine learning (a general form of Hopfield network^[24] for the associative memory application) algorithms are mainly

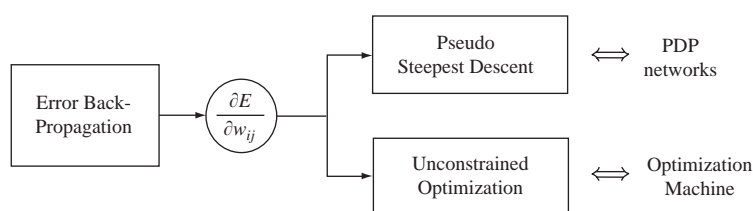


Fig. 4 Functional diagram for an optimization machine.

A

for combinatoric optimization problems with binary states of the units and the weights. Extension from binary to M -ary in the states of the weights has been reported for classification problems^[24] in simulated annealing training of the feed-forward perceptrons.

The ALOPEX was originally used for construction of visual receptive fields,^[21,22,25,26] but with some modifications it was later applied to the learning of any type of network not restricted to multilayer perceptrons. The ALOPEX is a random walk process in each parameter in which the direction of the constant jump is decided by the correlation between the weight changes and the energy changes.

The ALOPEX

The ALOPEX process is an optimization procedure that has been successfully demonstrated in a wide variety of applications. Originally developed for receptive field mapping in the visual pathway of frogs, ALOPEX's usefulness, combined with its flexible form, increased the scope of its applications to a wide range of optimization problems. Since its development by Tzanakou and Harth in 1973,^[21,22] ALOPEX has been applied to real-time noise reduction,^[27] and pattern recognition,^[28] in multilayer neural network training, among other applications.

Optimization procedures in general attempt to maximize or minimize a function $F(\cdot)$. The function $F(\cdot)$ is called the cost function and its value depends on many parameters or variables. When the number of parameters is large, finding the set that corresponds to the optimal solution is exceedingly difficult. If N were small, then one could perform an exhaustive search of the entire parameter space to find the "best" solution. As N increases, intelligent algorithms are needed to quickly locate the solution. Only an exhaustive search can guarantee that a global optimum is found; however, near-optimal solutions are acceptable because of the tremendous speed improvement over exhaustive search methods.

Backpropagation, described earlier, being a gradient descent method, often gets stuck in local extrema of the cost function. The local stopping points often represent unsatisfactory convergence points. Techniques have been developed to avoid the problem of local extrema, with simulated annealing^[19] being the most common. Simulated annealing incorporates random noise, which acts to dislodge the process from local extremes. It is crucial to the convergence of the process that the random noise be reduced as the system approaches the global optimum. If the noise is too large, the system will never converge and can be mistakenly dislodged from the global solution.

The ALOPEX is another process that incorporates a stochastic element to avoid local extremes in search

of the global optimum of the cost function. The cost function or response is problem dependent and is generally a function of a large number of parameters. The ALOPEX iteratively updates all parameters simultaneously based on the cross-correlation of local changes, ΔX_i , and the global response change ΔR , plus additive noise. The cross-correlation term $\Delta X_i \Delta R$ helps the process move in a direction that improves the response (Eq. 14). Table 1 illustrates how this can be used to find a global maximum of R .

All parameters X_i are changed simultaneously in each iteration according to

$$X_i(n) = X_i(n-1) + \gamma \Delta X_i(n) \Delta R(n) + r_i(n). \quad (14)$$

The basic concept is that this cross-correlation provides a direction of movement for the next iteration. For example, take the case where $X_i \downarrow$ and $R \uparrow$. This means that the parameter X_i decreased in the previous iteration, and the response increased for that iteration. The product $\Delta X_i \Delta R$ is a negative number, and thus X_i would be decreased again in the next iteration. This makes perfect sense because a decrease in X_i produced a higher response; if one is looking for the global maximum then X_i should be decreased again. Once X_i is decreased and R also decreases, then $\Delta X_i \Delta R$ is now positive and X_i increases.

These movements are only tendencies because the process includes a random component $r_i(n)$, which will act to move the weights unpredictably, avoiding local extrema of the response. The stochastic element of the algorithm helps avoid local extrema at the expense of slightly longer convergence or learning periods.

The general ALOPEX updating equation (Eq. 14) is explained as follows. $X_i(n)$ are the parameters to be updated, n is the iteration number, and $R(\cdot)$ is the cost function, of which the best solution in terms of X_i is sought. Gamma, γ , is a scaling constant, $r_i(n)$ is a random number from a gaussian distribution whose mean and standard deviation are varied, and $\Delta X_i(n)$ and $\Delta R(n)$ are found using

$$\Delta X_i(n) = X_i(n-1) - X_i(n-2) \quad (15)$$

$$\Delta R(n) = R(n-1) - R(n-2). \quad (16)$$

Table 1 Combinations of parameter changes and their contribution to the ALOPEX convergence

	ΔX_i		ΔR	$\Delta X_i \Delta R$
$X_i \uparrow$	+	$R \uparrow$	+	+
$X_i \uparrow$	+	$R \downarrow$	-	-
$X_i \downarrow$	-	$R \uparrow$	+	-
$X_i \downarrow$	-	$R \downarrow$	-	+

The calculation of $R(\cdot)$ is problem dependent and can be easily modified to fit many applications. This flexibility was demonstrated in the early studies of Harth and Tzanakou.^[22] In mapping receptive fields, no a priori knowledge or assumptions were made about the calculation of the cost function, instead a response was measured with microelectrodes inserted in the brain. Using action potentials as a measure of the response,^[22,25,26] receptive fields could be determined using the ALOPEX process to iteratively modify the stimulus pattern until it produced the largest response. This response was obtained from the stimulated neurons. (This was the first computer-to-brain interfacing ever used with electrodes inserted in an animal brain to measure neuronal activity and use it as feedback to the computer.)

Because of its stochastic nature, efficient convergence depends on the proper control of both the additive noise and the gain factor γ . Initially, all parameters X_i are random and the additive noise has a gaussian distribution with mean 0 and standard deviation σ , initially large. The standard deviation decreases as the process converges to ensure a stable stopping point (Fig. 5). Conversely, γ increases with iterations (Fig. 6). As the process converges, ΔR

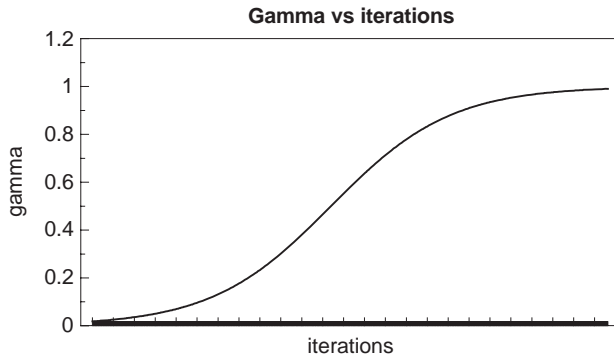


Fig. 5 Gamma vs. iterations.

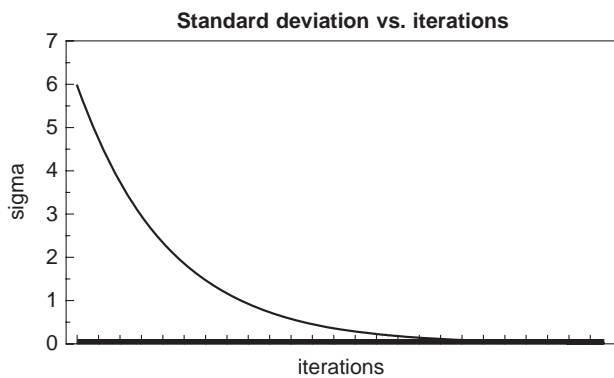


Fig. 6 Gaussian noise standard deviation vs. iterations.

becomes smaller and smaller, requiring an increase in γ to compensate.

Additional constraints include a maximal change permitted for X_i , for a single iteration. This bounded step size prevents the algorithm from drastic changes from one iteration to the next.

These drastic changes often lead to long periods of oscillation, during which the algorithm fails to converge.

Multilayer Perceptron (MLP) Network Training with ALOPEX

An MLP can also be trained for pattern recognition using ALOPEX. A response is calculated for the j th input pattern based on the observed and desired output,

$$R_j(n) = \mathbf{O}_k^{\text{des}} - (\mathbf{O}_k^{\text{obs}}(n) - \mathbf{O}_k^{\text{des}})^2, \quad (17)$$

where $\mathbf{O}_k^{\text{obs}}$ and $\mathbf{O}_k^{\text{des}}$ (the observed and desired outputs, respectively) are vectors corresponding to O_k for all k . The total response for iteration n is the sum of all individual template responses, $R_f(n)$.

$$R(n) = \sum_{j=1}^m R_j(n) \quad (18)$$

In Eq. 18, m is the number of templates used as inputs. The ALOPEX iteratively updates the weights using both the global response information and the local weight histories, according to the following equations,

$$W_{ij}(n) = r_i(n) + \gamma \Delta W_{ij}(n) \Delta R(n) + W_{ij}(n-1) \quad (19)$$

$$W_{jk}(n) = r_j(n) + \gamma \Delta W_{jk}(n) \Delta R(n) + W_{jk}(n-1), \quad (20)$$

where γ is an arbitrary scaling factor, $r_i(n)$ is the additive gaussian noise, ΔW represents the local weight change, and ΔR represents the global response information. These values are calculated by

$$\Delta W_{ij}(n) = W_{ij}(n-1) - W_{ij}(n-2) \quad (21)$$

$$\Delta W_{jk}(n) = W_{jk}(n-1) - W_{jk}(n-2) \quad (22)$$

$$\Delta R(n) = R(n-1) - R(n-2). \quad (23)$$

After the network was trained, it was tested for correct recognition using incomplete or noisy input patterns. The results demonstrate the robustness of the system to noise-corrupted data. Regardless of which training

procedure was used, backpropagation or ALOPEX, the recognition ability of the system was the same. The only difference was in how the response grew with iterations.

The neural network's robustness is derived from its parallel architecture and depends on the network topology rather than the learning scheme used to train. The network used was a three-layer feed-forward network with 35 input nodes, 20 hidden nodes, and 10 output nodes. The network's recognition ability was tested with noisy input patterns. Each 5×7 digit of the training set was subjected to noise of varying gaussian distribution and tested for correct recognition. The single-value response feedback, as opposed to the extensive error propagation schemes of other neural network training algorithms, makes ALOPEX suitable for fast VLSI implementation.

In our laboratory, an analog VLSI chip was designed to implement the ALOPEX algorithm, as described in another section. Making full use of the algorithm's tolerance to noise, an analog design was chosen. Analog designs offer larger and faster implementations than digital designs.

Issues in Neural Networks

Universal approximation

More than 30 years ago, Minsky and Papert^[7] reported that simple two-layer (no hidden layers) networks cannot approximate the non-linearly separating functions (e.g., XOR problems), but a multilayer neural network could do the job. Many results on the capability of multilayer perceptron have been reported. Some theoretical analyses for the network capability of the universal approximator are listed below and are extensively discussed by Hornik et al.^[29]

Kolmogorov^[30] tried to answer the question of Hilbert's 13th problem, i.e., the multivariate function approximation by a superposition of the functions of one variable. The superposition theory sets the upper limit of the number of hidden units to $(2n+1)$ units, where n is the dimensionality of the multivariate function to be approximated. However, the functional units in the network are different for the different functions to be approximated, while one would like to find an adaptive method to approximate the function from the given training data at hand. Thus, Kolmogorov's superposition theory says nothing about the capability of a multilayer network or which method is best.

More general views have been reported. Le Cun^[31] and Lapedes and Farber^[32] demonstrated that monotone squashing functions can be used in the two hidden layers to approximate the functions. Fourier series expansion of a function was realized by a single-layer

network by Gallant and White^[33] with cosine functions in the units. Further related results using the sigmoidal or logistic units are demonstrated by Hecht-Nielsen.^[34] Hornik et al.^[29] presented a general approximation theory of one hidden layer network using arbitrary squashing functions such as cosine, logistic, hyperbolic tangent, etc., provided that sufficiently many hidden units are available. However, the number of hidden units is not considered to attain any given degree of approximation.^[29]

The number of hidden units obviously depends on the characteristics of the training data set, i.e., the underlying function to be estimated. It is intuitive to say that the more complicated functions to be trained, the more hidden units are required.

For the number of the hidden units Baum and Haussler^[35] limit the size of general networks (not necessarily the feed-forward multilayer perceptrons) by relating it to the size of the training sample. The authors analytically demonstrated that if the size of the sample is N , and we want to correctly classify future observations with at least a fraction $(1-\epsilon/2)$, then the size of the sample has a lower bound given by

$$N \geq 0 \left(\frac{W}{\epsilon} \log \frac{N}{\epsilon} \right),$$

where W is the number of the weights and N is the number of the nodes in a network. This, however, does not apply to the interesting feed-forward neural networks and the given bound is not useful for most applications.

There seems to be no rule of thumb for the number of hidden units.^[5] Finding the size of the hidden units can usually be performed by cross-validation or any other resampling methods. The usual starting value for the size is suggested to be about the average of the number of the input and output nodes.^[15] Failure in learning can be attributed to three main reasons^[29]:

- Inadequate learning;
- Inadequate number of hidden units; or
- The presence of a stochastic rather than a deterministic relation between input and target in the training data, i.e., noisy training data.

THE VLSI IMPLEMENTATION OF ALOPEX

Artificial neural networks have existed for many years, yet because of recent advances in technology they are again receiving much attention. Major obstacles in ANN, such as a lack of effective learning algorithms, have been overcome in recent years. Training algorithms have advanced considerably, and now VLSI

technology may provide the means for building superior networks. In hardware, the networks have much greater speed, allowing for much larger architectures.

The tremendous advancement in technology during the past decades, particularly in VLSI technology, renewed interest in ANNs. Hardware implementations of neural networks are motivated by a dramatic increase in speed over software models. The emergence of VLSI technology has and will continue to lead neural network research in new directions. The past few years saw a considerable advancement in VLSI. Chips are now smaller, faster, contain larger memories, and are becoming cheaper and more reliable to fabricate.

Neural network architectures are varied, with over 50 different types explored in the research.^[34] Hardware implementations can be electronic, optical, or electro-optical in design. A major problem in hardware realizations is often not a result of the network architecture, but the physical realities of the hardware design. Optical computers, while they may eventually become commercially available, suffer far greater problems than do VLSI circuits. Thus, for the immediate and near future, neural network hardware designs will be dominated by VLSI.

Much debate exists as to whether digital or analog VLSI design is better suited for neural network applications. In general, digital designs are easier to implement and better understood. In digital designs, computational accuracy is only limited by the chosen word length. Although analog VLSI circuits are less accurate, they are smaller, faster, and consume less power than digital circuits.^[36] For these reasons, applications that do not require great computational accuracy are dominated by analog designs.

Learning algorithms, especially backpropagation, require high precision and accuracy in modifying the weights of the network, which led some to believe that analog circuits are not well suited for implementing learning algorithms.^[37] Analog circuits can achieve high precision at the cost of increasing the circuit size. Analog circuits with high precision (8 bits) tend to be equally large compared with their digital counterparts.^[38] Thus, high-precision analog circuits lose their size advantage over digital circuits. Analog circuits are of greater interest in applications requiring only moderate precision.

Early studies demonstrate that analog circuits can realize learning algorithms, provided that the algorithm is tolerant to hardware imperfections such as low precision and inherent noise. Macq et al. present a fully analog implementation of a Kohonen map, one type of neural network, with on-chip learning.^[39] Because analog circuits have been proven capable of the computational accuracy necessary for weight modification, they should continue to be the choice of neural network research.

Analog circuits are far superior to digital circuits in size, speed, and power consumption, and these areas

constrain most neural network applications. To achieve greater network performance, the size of the network must be increased. The ability to implement larger, faster networks is the major motivation for hardware implementation, and analog circuits are superior in these areas. Power consumption is also a major concern as networks become larger.^[40] As the number of transistors per chip increases, power consumption becomes a major limitation. Analog circuits dissipate less power than digital circuits, thus permitting larger implementations.

Beside its universality to a wide variety of optimization procedures, the nature of the ALOPEX algorithm makes it suitable for VLSI implementation. The ALOPEX is a biologically influenced optimization procedure that uses a single-value global response feedback to guide weight movements toward their optimum. This single-value feedback, as opposed to the extensive error propagation schemes of other neural network training algorithms, makes ALOPEX suitable for fast VLSI implementation.

Recently, a digital VLSI approach to implementing the ALOPEX algorithm was undertaken by Pandya et al.^[41] Results of their study indicated that ALOPEX could be implemented using a single-instruction multiple-data (SIMD) architecture. A simulation of the design was carried out in software and good convergence for a 4×4 processor array was demonstrated.

The importance of VLSI to neural networks has been demonstrated. For neural networks to achieve greater abilities, larger and faster networks must be built. In addition to size and speed advantages, other factors, including cost and reliability, make VLSI implementations the current trend in neural network research. The design of a fast analog optimization algorithm, ALOPEX, is discussed below.

Let us assume that we have an array of 64 pixels that we call $I_i(n)$, where n represents the iteration. The additive gaussian white noise is denoted by $r_i(n)$ and $R_j(n)$ is the response (or cost function) of the j th template at iteration n .

Let us assume that there are 16 templates to choose from, each with 64 pixels. The ALOPEX process is run on each of them with the objective being to recognize (converge to) an input pattern. As a result of the iterative behavior, if allowed to run long enough ALOPEX will eventually converge to each of the templates. However a "match" can be found by choosing the template that took the least amount of time to converge.

By convergence we mean finding either the global maximum or the minimum of the response function. This response function can be calculated many different ways, depending on the application. To allow this chip to be general enough to handle many applications, the response is computed off the chip. A Programmable Read Only Memory (PROM) is used to compute

the response based on the error between the input, $I_i(n)$, and the template. The PROM enables the response function to be changed to meet the needs of the application.

Although the chip design is limited to only 64 ALOPEX subunits, the parallel nature of ALOPEX enables many chips to be wired together for larger applications. Parallel implementations are made easy because each subunit receives a single global response feedback, which governs its behavior. Backpropagation, on the other hand, requires dense interconnections and communication between each node. This flexibility is a tremendous advantage when it comes to hardwired implementations.

Originally, the ALOPEX chip was designed using digital VLSI techniques. Digital circuitry was chosen over analog because it is easier to test and design. Floating point arithmetic was used to insure a high degree of accuracy. The digital design consisted of shift registers, floating point adders, and floating point multipliers. However, after having done much work toward the digital design it was abandoned in favor of an analog design. The performance of the digital design was estimated and was much slower than that of an analog design. The chip area of the digital design was much larger than that of an analog design would be. Also, the ALOPEX algorithm would be tolerant of analog imperfections because of its stochastic nature. For these reasons, it seemed clear that a larger, faster network could be designed with analog circuitry.

The analog design required components similar to the digital design to implement the algorithm. The main requirements were an adder, a multiplier, a difference amplifier, a sample and hold mechanism, and a multiplexing scheme. These cells each perform a specific function and are wired together in a way that implements the ALOPEX process.

The chip is organized into 64 ALOPEX subunits, one for each pixel in the input image. They are stacked vertically, wiring by abutment. Each subunit is composed of smaller components that are wired together horizontally and contains the following cells: a group selector, demultiplexor, follower aggregator, multiplier, transconductance amplifier, multiplexor, and another group selector.

The gaussian white noise required for the ALOPEX process is added to the input before it reaches the chip. This will allow precise control of the noise, which is very important in controlling the stability of the algorithm. If there is too much noise, the system will not converge. If there is too little noise, the system will get stuck in local minima of the cost function. By controlling the noise during execution, using a method similar to simulated annealing^[19] where the noise decays with time, the convergence time can be improved.^[28] By having direct control of the added noise, the component and

functional testing can be performed with no noise added, greatly simplifying the testing.

The addition, multiplication, and subtraction required by the ALOPEX algorithm are performed by the follower aggregator, the Gilbert multiplier,^[36] and the transconductance amplifier, respectively. The follower aggregation circuit computes the weighted average of its inputs. By weighing the inputs equally, the circuit computes the average of the two inputs. The average was chosen instead of the sum because the circuit is more robust, in that the output never has to exceed the supply voltage. A straight summer is more difficult to design because voltages greater than the supply voltage could be required. The output of the follower aggregator is sent to the multiplier, where a C-switch acts as a sample and hold, to store the value of the previous iteration. The difference between these signals is one input to the multiplier. The previous two responses, calculated off chip, are the other two inputs. The output of the multiplier is controlled by the control signal γ . The error signal is computed by a transconductance amplifier.

In designing the chip, much effort was made in making it controllable and testable, while making the chip general enough that it could be used in a wide variety of applications. This is why the gaussian white noise is added off chip and also why the error signal is taken off chip for the computation of the response. This not only allows the response function to be changed to meet the requirements of the specific application, but also provides the operator with accessible test points.

Although backpropagation is the most widely used software tool for training neural networks, it is less suitable for VLSI hardware implementation than ALOPEX for many reasons. Whereas backpropagation converges quickly as a result of its gradient descent method, it can often get stuck in local extrema. The ALOPEX tends to avoid local extrema by incorporating the random noise component at the expense of slightly longer convergence times.

The major differences arise when hardware implementation is discussed. Backpropagation is computationally taxing because of the error computation required for each node in the network. Each error is a function of many parameters (i.e., all the weights of the following layer). In hardware, very complex interconnections between all nodes are required to compute this error.

The ALOPEX is ideal for VLSI implementation for a couple of reasons. First, the algorithm is tolerant to small amounts of noise; in fact, noise is incorporated to help convergence. Second, all parameters change based on their local history and a single-value global response feedback. This single-value feedback is much simpler to implement than the error propagation used in backpropagation.

NEUROMORPHIC MODELS

Since the era of the vacuum tube, a multitude of neuronal models composed of discrete components and off-the-shelf Integrated Circuits (ICs) have been published. Similar efforts in custom VLSI, however, are far fewer in number. A good introduction to a number of neuronal attributes was presented by Linares-Barranco et al.^[42] Complementary Metal Oxide Semiconductor (CMOS)-compatible circuits for approximating a number of mathematical models of cell behavior are described. In its simplest form, this model represents the cell membrane potential in the axon hillock as nothing more than a linear combination of an arbitrary number, n , of dendritic inputs, X , each of which is weighted by a unique multiplier, W , summed, and processed by a non-linear range-limiting operator, f . The mathematical equation for this relationship is

$$Y_k = f \left\{ \sum_{i=1}^n W_i X_i \right\} = f\{S_k\}, \quad (24)$$

This relationship is realized in a CMOS circuit. This circuit is totally static and makes no provision for time courses of changes in input or output signals or intracellular relationships. An operational transconductance amplifier (OTA), as described by Wolpert and Micheli-Tzanakou,^[43] is used in lieu of operational amplifiers for this and most other VLSI neural network applications. Highly compatible with CMOS circuit technology, it is structurally simple and compact, realizable with only nine transistors, and provides reasonable performance. The only consideration it warrants is that its transfer function is a transconductance. As such, operations performed on its output signals must be oriented to its current, rather than its voltage. When driving high load impedances, as is usually the case with CMOS circuitry, this is only a minor inconvenience, necessitating buffering for lower load impedances. In fact, under some circumstances, such as when algebraic summation is being performed, a current output may actually be an advantage, allowing output nodes to be simply tied together.

The non-linear range-limiting operator, f , mentioned earlier, is necessitated by the observation that, for a given biological neuron, there are limits on the strength of the electrochemical gradients the cell's ionic pumps can generate. This imposes limits on how positive and negative cell membrane potential may go. Because a neuron may receive inputs from many other neurons, there is no such limit on the aggregate input voltage applied. As a result, an *activation function*, a non-linearity of the relationship between aggregate input potential and output potential of a neuron, must be imposed. This is typically done in one of three ways: the binary hard-limiter, which assumes one of only two

possible states, active or inactive; the linear-graded threshold, which assumes a linear continuum of active states between its minimal and maximal values; and the sigmoid, which assumes a sigmoidal distribution of values between its negative minimal and positive maximal output values.

Which type of activation function is employed depends on the type of artificial neuron and network in which it is implemented. In networks where cell outputs are all-or-none, such as the McCulloch and Pitts models,^[44] the binary threshold model is used. In networks where neurons are theorized to have variable output levels applied to distinctly designated excitatory and inhibitory inputs, such as Hopfield networks,^[24] the linear threshold model is used. In networks where a synaptic connection must be both excitatory and inhibitory, depending on the level of activity, the sigmoid threshold is used. In either of the latter two activation functions, the slope of the overall characteristic can be varied to suit the sensitivity of the cell in question.

Another well-developed implementation of individual artificial nerve cells is that by Wolpert and Micheli-Tzanakou.^[43,45] Whereas most neuromorphic models are based on the Hodgkin–Huxley equations, this one uses a sequencer to synthesize the action potential in three distinct phases. It also employs a different formulation for cell membrane and threshold potentials known as an integrate-and-fire model, presented and implemented in discrete components by French and Stein in 1970.^[46] It makes use of the aforementioned leaky integrator and provides off-chip control over the response and persistence of stimuli assimilated into the membrane potential. The model affords similar controls over the resting level and time constant of the cell threshold potential and allows refraction to be recreated. This organization also affords control over the shape, resting level, and duration of the action potential and produces a Transistor-Transistor Logic (TTL)-compatible pulse in parallel with the action potential. These controls, all of which are continuously and precisely adjustable, make this model ideal for replicating the behavior of a wide variety of individual nerve cells, and it has been successfully applied as such.

The Wolpert and Micheli-Tzanakou model is organized around three critical nodes, the somatic potential, the axonal potential, and the threshold potential. Each of these nodes is biased off-chip with an R-C network so that its resting level and time constant are independently and continuously controllable. Stimuli to the cell are buffered and standardized by truncation into 10- μ sec impulses. Synaptic weight inputs on the excitatory and inhibitory pathways allow this value to be increased or decreased from off-chip. The impulses are then applied to a somatic potential by a push–pull Metal-Oxide-Semiconductor Field-Effect Transistor (MOSFET) stage and compared with threshold potential by an OTA

acting as a conventional voltage comparator. When threshold is exceeded, an action potential is synthesized and outputted. This waveform is then binarized and buffered to form a binary-compatible output pulse. At the same time, threshold is elevated to form the refractory period. The circuit consists of approximately 130 transistors plus a few on-chip and discrete resistors and capacitors and was implemented in a conventional CMOS technology, requiring a single-ended DC supply of 4–10 V DC and occupying 0.6 mm² of chip area.

With its critical nodes bonded out off-chip, Wolpert and Micheli-Tzanakou's neuromime rate of operation may be accelerated from a biologically compatible time frame over several orders of magnitude. This model was first implemented in 1986^[43] and was intended as a flexible and accurate esthetic, rather than a mathematical model of cell behavior. Since then, it has been used to successfully recreate a number of networks from well-documented biological sources. Waveforms obtained in these recreations have demonstrated a striking similarity to intracellular recordings taken from their biological counterparts. It has also been successfully applied to problems in robotics and rehabilitation.

Neurological Process Modeling

Lateral inhibition is the process in which a cell containing some level of information encoded as its output level acts to inhibit and is inhibited by a similar adjoining cell. For many years, this process has been observed with striking regularity in both 1- and 2-D arrays of sensory receptors in a variety of systems and a variety of organisms. In numerous morphological, mathematical, and circuit studies, it has been identified as a key image preprocessing step that optimizes a sensory image to facilitate fast and accurate recognition in subsequent operations. Lateral inhibition accomplishes this by amplifying differences, enhancing image contrast, lending definition to its outward shape, and isolating the image from its background. Whereas a digital computer would accomplish this process one pixel at a time, biological systems manage it in a manner that is both immediate and simultaneous.

Laterally inhibited behavior has been observed in pairs of cells implemented in hardware and software models by many researchers, but in dedicated VLSI by only a few. Among them, Nabet and Pinter of Drexel University and Darling of the University of Washington have extensively studied the stability and effectiveness of both pairs and linear strings of mutually inhibiting cells in CMOS VLSI and obtained results that correlate well with biological data.^[47] This line of work has been explored in two dimensions in another

series of VLSI-based models by Wolpert and Micheli-Tzanakou.^[48] Arrays of mutually inhibiting cells that inhibit via continuously active connections and cells that inhibit by dynamic or strobed controls both offered stable and variable control over the degree of inhibition. Arrays of hexagonally interconnected cells were more stable than the square array, which tended to “checkerboard” when significant levels of inhibition were attempted. Feedback inhibition, where one array is used to store both the initial and the inhibited images, was as effective, but less convenient to access than feed-forward inhibition, where separate input and inhibited images are maintained.

Characterization of lateral inhibition in the context of a more specific biological model has been pursued in another noteworthy effort by Andreou et al.^[49] Multiple facets of cell–cell interactions, including both mutual inhibition and leakage of information between adjoining cells, were implemented in VLSI as a model of early visual processing in the mammalian retina.^[49] There, adjacent cells on the photoreceptor layer intercommunicate through gap junctions, where their cell membrane potentials couple through a resistive path. Simultaneously, optical information from the photoreceptor cells is downloaded to corresponding cells of the horizontal layer, which have mutually inhibitory connections. One-dimensional arrays, and, subsequently, 2-D models of these relationships were implemented in analog VLSI and tested. Although little numerical data were published from these arrays, the 2-D array produced a number of optical effects associated with the human visual system, including Mach bands, simultaneous contrast enhancement, and the Herman–Herring illusion, all of which are indicative of the real-time image processing known to occur in the mammalian retina.

Finally, the definitive VLSI implementation of a 2-D array is the well-known silicon retina devised by Mead and described in his text, “Analog VLSI and Neural Systems.”^[50] Mead presents a comprehensive treasury of analog VLSI circuits for a variety of mathematical operations necessary to implement neural networks in VLSI. The book then goes on to discuss several applications of analog ANNs, culminating in an auditory model of the cochlea and a visual model of the retina.

ARTICLES OF FURTHER INTEREST

Biopotential Amplifiers; Biosensors; Compression of Digital Biomedical Signals; Control of Movement; Digital Biomedical Signal Acquisition and Processing: Basic Topics; Digital Biomedical Signal Processing: Advanced Topics; Excitable Tissue, Electrical Stimulation Of; Modeling, Biomedical

REFERENCES

1. Lippmann, R.P. An introduction to computing with neural nets. *IEEE ASSP Magazine* **1987**, April, 4–22.
2. Hush, D.; Horne, B. Progress in supervised neural networks. *IEEE Signal Processing Magazine* **1993**, January, 8–39.
3. Kohonen, T. The self-organizing map. *Proc. IEEE* **1990**, 78 (9), 1464–1480.
4. Carpenter, G.A.; Grossberg, S. Art2 Self-organization of stable category recognition codes for analog input patterns. *Appl. Optics* **1987**, 26, 4919–4930.
5. Ripley, B.D. Statistical Aspects of Neural Networks. In *Chaos and Networks: Statistical and Probabilistic Aspects*; Barndorff-Neilsen, O.E., Cox, D.R., Jensen, J.L., Kendall, S.S., Eds.; Chapman & Hall: London, 1993.
6. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning Internal Representations by Error Back Propagation. In *Parallel Distributed Processing Explorations in the Microstructure of Cognition, I: Foundations*; Rumelhart, D.E., McClelland, J.L., the PDP Research Group, Eds.; MIT Press: Cambridge, MA, 1986; Chap. 8.
7. Minsky, M.; Papert, S. *Perceptrons: An Introduction to Computational Geometry*; MIT Press: Cambridge, MA, 1969.
8. Rosenblatt, F. *Principles of Neurodynamics*; Spartan Books: New York, 1959.
9. Bryson, A.E.; Ho, Y.C. *Applied Optimal Control*; Bleisdell: New York, 1969.
10. Werbos, P.J. *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences*, Ph.D. Thesis Ed.; Harvard University: Cambridge, MA, 1974.
11. Parker, D.B. *Learning Logic*, Technical Report; Center for Computational Research in Economics and Management Science, MIT: Cambridge, MA, 1985.
12. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. *Nature* **1986**, 323, 533–536.
13. Widrow, B.; Lehr, M. 30 years of adaptive neural networks: Perceptron, madaline and backpropagation. *Proc. IEEE* **1990**, 78 (9), 1415–1442.
14. Widrow, B. Generalization and Information Storage in Networks of Adaline “Neurons.” In *Self-Organizing Systems*; Yovitz, M., Jacobi, G., Goldstein, G., Eds.; Spartan Books: Washington, DC, 1962; 435–461.
15. Zahner, D.; Micheli-Tzanakou, E. Alopex and back-propagation. In *Supervised and Unsupervised Pattern Recognition, Feature Extraction in Computational Intelligence*; Micheli-Tzanakou, E., Ed.; CRC Press: Boca Raton, Florida, 2000; Chap. 2, 61–78.
16. Micheli-Tzanakou, E.; Uyeda, E.; Sharma, A.; Ramujan, K.S.; Dong, J. Face recognition: Comparison of neural networks algorithms. *Simulation* **1995**, 64, 37–51.
17. Micheli-Tzanakou, E. Neural networks in biomedical signal processing. In *The Biomedical Eng. Handbook*, 2nd Ed.; Bronzino, J., Ed.; CRC Press: Boca Raton, Florida, 1995; Chap. 60, 917–931.
18. Peressini, A.L.; Sullivan, F.E.; Uhl, J.J., Jr. *The Mathematics of Nonlinear Programming*; Springer-Verlag: New York, 1988.
19. Kirkpatrick, S.; Gelatt, C.D., Jr.; Vecchi, M.P. Optimization by simulated annealing. *Science* **1983**, 220 (4598), 671–680.
20. Hinton, G.E.; Sejnowski, T.J. Learning and relearning in Boltzmann machines. In *Parallel Distributed Processing: Explorations in the Microstructure of Cognition I: Foundations*; Rumelhart, D.E., McClelland, J.L., and the PDP Research Group, Eds.; MIT Press: Cambridge, MA, 1986; Chap. 7.
21. Tzanakou, E.; Harth, E. Determination of visual receptive fields by stochastic methods. *Biophys. J.* **1973**, 15, 42a.
22. Harth, E.; Tzanakou, E. Alopex: A stochastic method for determining visual receptive fields. *Vis. Res.* **1974**, 14, 1475–1482.
23. Engel, J. Teaching feed-forward neural networks by simulated annealing. *Complex Syst.* **1988**, 2, 641–648.
24. Hopfield, J.J.; Tank, D.W. Neural computation of decisions in optimization problems. *Biol. Cybernet.* **1985**, 52, 141–152.
25. Tzanakou, E.; Michalak, R.; Harth, E. The ALOPEX process: Visual receptive fields by response feedback. *Biol. Cybernet.* **1979**, 35, 161–174.
26. Micheli-Tzanakou, E. Non-linear characteristics in the frog’s visual system. *Biol. Cybernet.* **1984**, 51, 53–63.
27. Ciaccio, E.; Tzanakou, E. *The ALOPEX Process: Application to Real-Time Reduction of Motion Artifact*. Annual International Conference of IEEE-EMBS, Philadelphia, **1990**, 12 (3), 1417–1418.
28. Dasey, T.J.; Micheli-Tzanakou, E. A Pattern Recognition Application of the Alopex Process with Hexagonal Arrays. International Joint Conference on Neural Networks, Washington, DC, **1989**, II, 119–125.
29. Hornik, K.; Stichcombe, M.; White, H. Multilayer feedforward networks are universal approximators. *Neural Netw.* **1989**, 2, 359–366.
30. Kolmogorov, A.N. On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition. *Doklady Akad. Nauk SSR* **1957**, 114, 953–956.
31. Le Cun, Y. *Modeles Connexionistes de L’apprentissage*, Ph.D. thesis; Université Pierre et Marie Curie: Paris, France, 1987.
32. Lapedes, A.; Farber, R. *How Neural Networks Work*, Technical report; Los Alamos National Laboratory: Los Alamos, NM, 1988.
33. Gallant, A.R.; White, J. *There Exists A Neural Network That Does Not Make Avoidable Mistakes*, IEEE Second International Conference on Neural Networks; San Diego, CA, 1988; I, 657–664.
34. Hecht-Nielsen, R. *Theory of the Back Propagation Neural Network*, Proceedings of the International Joint Conference on Neural Networks; San Diego, CA, 1989; I, 593–608.
35. Baum, E.; Haussler, D. What size net gives valid generalization? *Neural Comput.* **1989**, 1, 151–160.
36. Mead, C.; Ismail, M.; Eds. *Analog VLSI Implementation of Neural Systems*; Kluwer Academic: Boston, 1989.
37. Ramacher, U.; Ruckert, U.; Eds. *VLSI Design of Neural Networks*; Kluwer Academic: Boston, 1991.

38. Graf, H.P.; Jackel, L.D. *Analog Electronic Neural Network Circuits*. IEEE Circuits and Devices Magazine **1989**, July, 44–55.
39. Macq, D.; Verleysen, M.; Jespers, P.; Legat, J. Analog implementation of a Kohonen map with on-chip learning. IEEE Trans. Neural Netw. **1993**, 4 (3), 456–461.
40. Andreou, A.; et al. *VLSI Neural Systems*. IEEE Trans. Neural Netw. **1991**, 2 (2), 205–213.
41. Pandya, A.S.; Shandar, R.; Freytag, L. An SIMD Architecture for the Alopex Neural Network. In *Parallel Architectures for Image Processing*; SPIE: Portland, Oregon, 1990; 1246, 275–287.
42. Linares-Barranco, B.; Sanchez-Sinencio, E.; Rodriguez-Vazquez, A.; Huertas, J.L. A CMOS implementation of Fitzhugh–Nagumo model IEEE. J. Solid-State Circuits **1991**, 26, 956–965.
43. Wolpert, S.; Micheli-Tzanakou, E. *An Integrated Circuit Realization of a Neuronal Model*, Proceedings of the IEEE Northeast Bioengineering Conference; New Haven, CT, 1986; March 13–14.
44. McCulloch, W.S.; Pitts, W. A logical calculus of the ideas imminet in nervous activity. Bull. Math. Biophys. **1943**, 5, 115–133.
45. Wolpert, S.; Micheli-Tzanakou, E. A neuromime in VLSI. IEEE Trans. Neural Netw. **1995**, 6 (6), 1560–1561.
46. French, A.S.; Stein, R.B. A flexible neuronal analog using integrated circuits. IEEE Trans. Biomed. Eng. **1970**, 17, 248–253.
47. Nabet, B.; Pinter, R.B. *Sensory Neural Networks: Lateral Inhibition*; CRC Press: Boca Raton, Florida, 1991.
48. Wolpert, S.; Micheli-Tzanakou, E. Silicon models of lateral inhibition. IEEE Trans. Neural Netw. **1993**, 4 (6), 955–961.
49. Andreou, A.G.; Boahen, K.A.; Pouliquen, P.O.; Pava-sovic, A.; Jenkins, R.E.; Strohbahn, K. Current-mode subthreshold MOS circuits for analog VLSI neural systems. IEEE Trans. Neural Syst. **1991**, 2 (2).
50. Mead, C.A. *Analog VLSI and Neural Systems*; Addison-Wesley, Boston, 1989.

Bioactive Glass

Marivalda M. Pereira

Department of Metallurgical and Materials Eng., Federal University of Minas Gerais, Belo Horizonte, MG, Brazil

Larry L. Hench

Department of Materials, Imperial College London, London, United Kingdom

B

INTRODUCTION

Bioactive glasses are glasses designed to induce a specific biological activity; in most cases the desired biological activity is one that will result in a strong bonding to bone. The concept of a bioactive material capable of bone bonding was originally suggested by Larry Hench in the early 1970s. Since then, several glass compositions have been studied, tissue bonding mechanisms proposed, and clinical applications developed.^[1–3] The bone bonding mechanism includes a series of complex physicochemical reactions on the surface of the implanted material and the formation of a biologically active calcium phosphate layer. The surface reactions lead to the biochemical adsorption of growth factors and to a sequence of cellular events that lead to formation of new bone. The carbonated hydroxyapatite layer that forms is structurally and chemically equivalent to the mineral phase of bone, and the interfacial bonding is generally attributed to its formation.^[2,3] The resulting bonding exhibits high strength, equivalent to or even surpassing that of either bone or the implanted glass.

The bioactivity of glasses was initially observed for silicate glasses based on the system $\text{SiO}_2\text{--CaO--Na}_2\text{O--P}_2\text{O}_5$.^[4] Bioglass[®] is a trade name given to a series of such glass compositions. Other bioactive glass compositions, as well as other bioactive materials such as glass ceramics and calcium phosphate-based ceramics (referred to generally as bioactive ceramics), have also been developed.^[1–3] Among the bioactive ceramics, bioactive glasses have the highest levels of bioactivity based on their rates of reaction and bone bonding.

The following sections describe the structure and processing of bioactive glasses. The bioactivity mechanism and the effect of composition and structure on the bioactivity are discussed, and the mechanical performance of these materials and their clinical applications are briefly treated. Finally, the recent development of bioactive glass scaffolds for tissue engineering applications is described.

SYNTHESIS AND STRUCTURE OF BIOACTIVE GLASSES

At the atomic level, the absence of periodicity in the material structure is what distinguishes glasses from crystalline ceramics. The nonorganized tridimensional network of silicate glasses is based on the silica structure, i.e., silica tetrahedrons (silicon atoms surrounded by four oxygen atoms) bonded by two oxygen atoms. Alkaline and alkaline-earth atoms such as Na, K, Ca, and Mg occupy random sites in the silica network, modifying that structure and creating nonbridging oxygen atoms. Other atoms, such as P and Al, may be present as part of the network structure.

Glasses are normally prepared by cooling a mixture of raw materials from the liquid state.^[2] The cooling rate is important, because slow cooling rates may allow atomic arrangement that leads to crystallization of the material. However, silicate-based melted mixtures form a vitreous phase at normal cooling rates. The raw materials normally used are high-purity silica, calcium and sodium carbonate, and phosphorus oxide. For introduction of other elements such as Mg, Al, Zr, and others, into the glass composition, other oxides or salts are used. The melting temperature of silicate glasses is in the range of 1300–1400°C. Monolithic forms are produced by casting the melt in appropriate molds, whereas powder or granules are produced by rapid-cooling the melt in a liquid, such as acetone, to produce a coarse powder called a frit, followed by grinding and size-screening to a specific particle size range.

An alternative process used for bioactive glass production is the sol-gel method.^[2,5] This method leads to the formation of a highly microporous material with high surface area. The sol-gel route is based on the controlled hydrolysis and condensation of metal alkoxides to form a suspension of colloidal particles (sol), which upon polycondensation forms an interconnected network structure, the gel. Heat treatment stages (aging, drying, and stabilization) are then conducted, leading to the final glass structure. Precursors used to form the silica network of the gel-glasses are Si(OR)_4 ,

where R is normally CH_3 or C_2H_5 . Other metal atoms, such as Ca, P, Ti, and others, can be introduced into the structure as metal alkoxides or salts.

BIOACTIVITY

The search for materials adequate for use in the human body (biomaterials) was initially based on the concept of bioinertness, its principle being the fact that the material would not cause any reaction of the surrounding tissues. This concept has been changed, however, due to evidences that every material causes a certain tissue response, even a minimum one. The different tissue responses induced by a biocompatible material may be divided in three ways.^[2,3]

1. The material is biologically inert (or almost inert). It does not present significant chemical changes in the physiological environment, and a fibrous capsule of variable thickness forms around the implant. Examples of bioinert materials are aluminum oxide, surgical metals and alloys such as titanium and stainless steel, polymers such as poly(methyl methacrylate), polyethylene, and silicone, among others.
2. The material is biologically active. It presents a controlled reactivity and induces a specific biological response that leads to the formation of a continuous interface between the implanted material and tissue. Tissue bonding occurs. Bioactive glasses and calcium phosphate ceramics are examples.
3. The material is biodegradable or bioresorbable. It dissolves or is decomposed, and is substituted by surrounding tissue. The degradation mechanism is gradual and can be mediated by specific biological activity. Examples are polymers such as poly(lactic acid), tricalcium phosphate, and as we will see, some compositions of bioactive glasses.

Factors such as chemical composition, phase composition, pore structure, particle size, implantation site, animal species, and others affect the tissue response.

Based on the classification above, bioactivity can be defined as *the ability of the material to induce the formation of an interfacial bonding between the implant and living tissues, without the formation of a fibrous capsule separating the biomaterial and the tissue*. In the case of the bioactive glasses, various in vitro and in vivo studies show that a series of interfacial reactions occur that lead to the formation of a carbonated hydroxyapatite layer (HCA) on the glass surface responsible for bone bonding. Five reaction stages are identified:^[3]

1. Ion exchange: $\text{Si-O-Na}^+ + \text{H}^+ + \text{OH}^- \rightarrow \text{Si-OH} + \text{Na}^+ + \text{OH}^-$.

2. Loss of soluble $\text{Si}(\text{OH})_4$ and formation of Si-OH (silanols): $\text{Si-O-Si} + \text{H}_2\text{O} \rightarrow \text{Si-OH} + \text{OH-Si}$.
3. Condensation and repolymerization of a SiO_2 -rich layer on the surface.
4. Migration of Ca^{2+} and PO_4 groups to the surface, forming an amorphous $\text{CaO-P}_2\text{O}_5$ film.
5. Crystallization of the amorphous $\text{CaO-P}_2\text{O}_5$ film by incorporation of OH^- , CO_3 .

The surface reactions lead to the biochemical adsorption of growth factors (stage 6) and to the synchronized sequence of cellular events that leads to formation of new bone (stages 7–12).

6. Adsorption of biological moieties and growth factors on the HCA layer.
7. Action of macrophages.
8. Attachment of stem cells.
9. Differentiation of stem cells.
10. Generation of matrix.
11. Crystallization of matrix.
12. Proliferation and growth of bone.

The first five reaction stages occur very rapidly on the surface of most bioactive glasses because of rapid surface reactions. The formation of the silica-rich gel layer provides a favorable substrate for the rapid nucleation and growth of the biologically equivalent HCA layer. Surface compositional profiles resulting from stage 1 through stage 5 reactions have been measured for 45S5 bioactive glasses (a glass with a composition of 45% SiO_2 , 24.5% Na_2O , 24.5% CaO , and 6% P_2O_5) using Auger electron spectroscopy (AES) combined with argon-ion milling.^[6] The results (Fig. 1) show that a calcium and phosphorus-rich reaction layer, heterogeneously nucleated and crystallized from the silica-gel layer formed on the glass

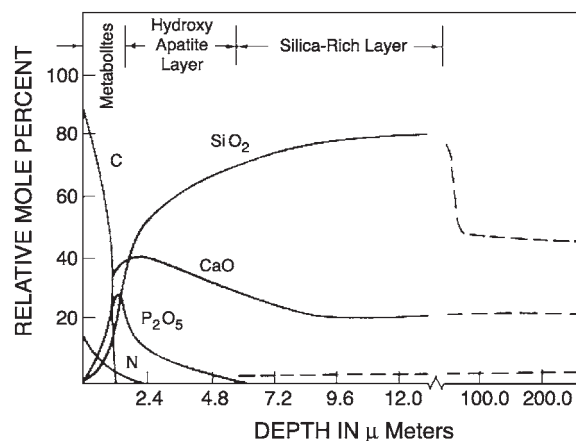


Fig. 1 Chemical composition of the bilayer film formed on 45S5 bioactive glass after 1 hour of implantation in rat bone.

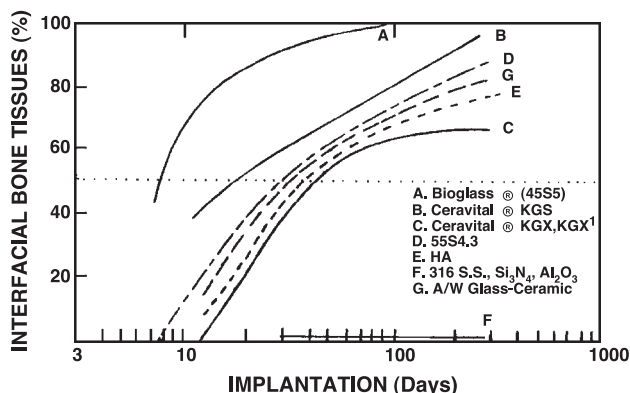


Fig. 2 Rates of bond formation with bone for implant materials with different bioactivity.

surface, has developed after only one hour in vivo. With time, the interfacial bonding zone between living and nonliving material continues to grow by incorporating calcium and phosphorus ions from body fluids. The final result is a gradient in composition that extends over several hundred micrometers, similar in dimension to the bonding zone that occurs naturally at bone interfaces.^[4,7]

Differences observed in the behavior of glasses with different compositions are related to the differences in the rate of formation of this HCA layer, as will be discussed in the next section. In the case of other bioactive ceramics, such as glass-ceramics or calcium phosphate ceramics, an HCA layer also forms in the physiological environment, both in vitro and in vivo.^[1,8,9] However, the rate of formation of the HCA layer is slower than the rate for bioactive glasses. The reactivity of the material and the rate of HCA formation are closely related to the formation of an interface between the material and bone tissue in vivo.^[11] The interfacial bone tissue formation time for bioceramics with different rates of reactivity is shown in Fig. 2. The level of bioactivity of the material can be related to the time for more than 50% of the interface to be bonded ($t_{0.5bb}$), and an index of bioactivity can be defined as $I_B = 100/t_{0.5bb}$. The bioactivity varies from the most bioactive material—the bioactive glass 45S5 with a bioactivity index of $I_B = 12$ —to almost inert ceramics with an index of $I_B = 0$.

EFFECT OF COMPOSITION ON BIOACTIVITY

The basic components of most bioactive glasses are SiO_2 , Na_2O , CaO , and P_2O_5 in specific proportions. Bonding to bone has been demonstrated for a certain compositional range of these glasses.^[1,4,10] Three important compositional features distinguish them from traditional soda–lime–silica glasses: <60 mol% SiO_2 , high Na_2O and CaO content, and a high $\text{CaO}:\text{P}_2\text{O}_5$

ratio. These compositional features make the surface highly reactive when exposed to an aqueous medium, such as body fluids. Figure 3 illustrates the compositional dependence (in weight percent) of bone bonding and soft tissue bonding for the $\text{Na}_2\text{O}-\text{CaO}-\text{P}_2\text{O}_5-\text{SiO}_2$ glasses.^[11] All glasses in Fig. 3 contain a constant 6 wt% of P_2O_5 . Compositions in the middle of the diagram (region A) form a bond with bone. Consequently, region A defines the bioactive bone-bonding boundary. When the concentration of SiO_2 in the glass network exceeds 55%, the rates of reaction decrease, and bonding to bone is very slow. At a concentration of 60% SiO_2 or higher, the rates of reaction are sufficiently slow that the material is biologically inert. Silicate glasses within region B behave as almost inert materials and elicit formation of a fibrous capsule at the implant-tissue interface. Glasses within region C are resorbable and disappear within 10–30 days of implantation. Glasses within region D are not technically practical and, therefore, have not been tested as implants.

Additional compositions of bioactive glasses based on the system $\text{Na}_2\text{O}-\text{K}_2\text{O}-\text{MgO}-\text{CaO}-\text{B}_2\text{O}_3-\text{P}_2\text{O}_5-\text{SiO}_2$ have been developed at Abo Akademi and the University of Turku, Finland, and the compositional dependence for bioactivity has been largely studied.^[11–13] Bioactivity occurs only within certain compositional limits and very specific ratios of oxides in this system. For the glasses containing <59 mol % SiO_2 , the silica-rich layer and the calcium phosphate layer were also observed in this system. In-vivo experiments showed that glasses in the investigated system are bioactive when they contain 14–30 mol % alkali oxides, 14–30 mol % alkaline earth oxides, and <59 mol % SiO_2 . Glasses containing potassium and magnesium bonded to bone in a way similar to that of bioactive glasses of the original system. Substitutions in the 45S5 formula of 5–15 wt% B_2O_3 for SiO_2 , or 12.5 wt% CaF_2 for CaO , or crystallizing the glass compositions to form glass-ceramics have little effect on the ability of

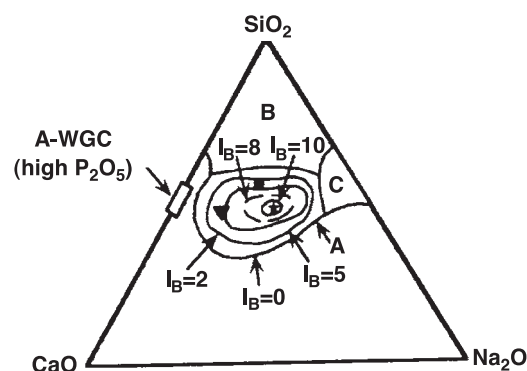


Fig. 3 Composition dependence (in wt%) of bone bonding and soft tissue bonding of bioactive $\text{Na}_2\text{O}-\text{CaO}-\text{SiO}_2$ glasses with a constant 6% P_2O_5 .

the material to form a bone bond. However, addition of as little as 3 wt% Al_2O_3 to the 45S5 formula prevents bonding. Glasses with substantially larger amounts of P_2O_5 do not bond to bone.

A layer of biologically active HCA must form for a bond with tissues to occur. This is the common characteristic of all the known bioactive implant materials. It is the rate of HCA formation (stage 4) and the time for onset of crystallization (stage 5) that vary greatly. When the rate becomes excessively slow, no bond forms, and the material is no longer bioactive. Although composition has a strong influence on bioactivity, the rate of HCA layer formation and bone formation depends not only on the glass composition but also on other aspects, such as particle size, medium composition and volume, adsorption of proteins, and the implantation site.^[14,15]

The aforementioned compositional dependence was defined for dense glasses produced by the conventional melt process. When the sol-gel method is used, the compositional bioactivity range is increased above 60%.^[6,16] In fact, bioactivity is observed even for pure SiO_2 gel-glasses as well as for other sol-gel-derived oxides.^[17] The high bioactivity of the sol-gel glasses is attributed to their high surface areas (up to $600\text{ m}^2/\text{g}$) and the large concentration of M-OH groups on the highly porous glass surface—both features associated with the sol-gel technique. These surface characteristics increase the rate of nucleation and growth of the HCA layer on the surface of the material. The bioactive gel-glasses also have another very important characteristic, which is their higher degradation rate compared to dense glasses.^[18] Therefore, depending on composition and physical characteristics such as particle size and pore structure, bioactive glasses may be considered biodegradable materials.

TISSUE ATTACHMENT AND INTERFACIAL STRENGTH

The formation of a stable interface between an implant material and the living host tissue is crucial to the success of an implant device. The nature and thickness of the tissue-implant interface and the consequent mechanism of tissue attachment are directly related to the type of tissue response at the implant interface.^[3,19] The different means of achieving attachment of prostheses to the musculoskeletal system are categorized as: 1) morphological fixation, achieved by mechanical interlocking on the surface irregularities, cementing the device into the tissue, or press-fitting into a defect; 2) biological fixation, achieved by tissue ingrowth into porous implants or surfaces; and 3) bioactive fixation, in which the material attaches directly by bonding with the tissue.

The relative level of reactivity of an implant influences the thickness of the interfacial zone, or layer, between the material and the tissue. Analysis of the failure of implant materials during the past 20 years generally shows failure originating from the bio-material-tissue interface.^[3] When biomaterials are almost inert and the interface is not chemically or biologically bonded, there is relative movement, causing progressive development of a nonadherent fibrous capsule in both soft and hard tissues. Movement at the biomaterial-tissue interface eventually leads to deterioration in function of the implant or of the tissue at the interface, or both. The thickness of the nonadherent capsule varies greatly, depending upon both the material and the extent of relative motion. The fibrous capsule in almost-inert materials such as pure titanium can be very thin when the implant achieves a tight mechanical fit. If relative movement occurs at the tissue-implant interface, however, the fibrous capsule can become several hundred micrometers thick, loosening the implant and generally leading to clinical failure.

The use of bioactive materials provides a solution to the problem of interfacial attachment. Bioactive materials form an interfacial bond with adjacent tissue, but the time dependence of bonding and the mechanism of bonding differ for the various materials, as discussed earlier. The thickness of the bonding zone and the strength of the bond depend on the material reactivity, and they appear to be correlated.^[3]

The failure strength of a bioactively fixed bond appears to be a nonlinear function of the bonding rate of the implant.^[1,3] Interfacial strength appears to be inversely dependent on the thickness of the bonding zone. For example, 45S5 Bioglass[®] with a very high bonding rate develops a gel-bonding layer of $100\ \mu\text{m}$ thickness that has a relatively low shear strength. In contrast, A/W glass-ceramic, with an intermediate rate of bonding, has an interface of $10\text{--}20\ \mu\text{m}$ and a very high resistance to shear.^[1,20] Therefore, interfacial strength is time-dependent and is a function of morphological factors, such as the change in interfacial area with time, the progressive mineralization of the interfacial tissues, and the resulting increase in elastic modulus of the interfacial bond. It is also a function of shear strength per unit of bonded area and the density and quality of bone bonded to the interface.

SOFT TISSUE BONDING

Although bioactivity is defined as the bonding of tissue to an implant, most of the data and evidence regarding bioactivity have come from bone implants, as shown in the previous discussion. In fact, early studies showed

that bioactive glasses did not bond to soft tissue. It was shown later, however, that a fibrous capsule that formed around bioactive glass implants in muscle was an artifact of the experiment caused by continuous micromotion at the implant-tissue interface.^[1,21] When the implant-tissue interface was immobilized, Wilson et al.^[21] proved that collagen fibers of the soft tissue became embedded and bonded within the growing silica-rich and HCA layers on the 45S5 bioactive glass. Soft tissue bonding was also observed in subcutaneous implants when they were sufficiently immobile.

There are situations in which both soft and hard tissue bonding are desirable, such as in maxillofacial implants and the treatment of periodontal disease. The use of 45S5 bioactive glass powder, with a large I_B value, showed attachment to both bone and soft connective tissue in the treatment of periodontal defects.^[1] Use of the highly bioactive material led to a more rapid and complete repair than was observed when the lower-bioactivity material hydroxyapatite was used, and it also resulted in reattachment of the periodontal ligament with little or no epithelial downgrowth. Bioactive implants with intermediate I_B value do not develop a stable soft tissue bond. Instead, a fibrous interface progressively mineralizes to form bone. Consequently, there appears to be a critical $isoI_B$ boundary beyond which bioactivity is restricted to stable bone bonding. Bioactivity inside the critical $isoI_B$ boundary includes both stable bone bonding and soft tissue bonding, depending on the progenitor stem cells in contact with the implant. Materials within the critical $isoI_B$ boundary have been designated to exhibit class A bioactivity.^[3] An important feature of class A bioactive materials is that they are osteoproliferative as well as osteoconductive. In contrast, class B bioactive materials exhibit only osteoconductivity, defined as the characteristic of bone growth and bonding along a surface. Osteoproliferation occurs when bone proliferates on the particulate surfaces of a mass as a result of enhanced osteoblast activity. Enhanced proliferation and differentiation of osteoprogenitor cells, stimulated by slow resorption of the class A bioactive particles, are responsible for osteoproliferation.

MECHANICAL PROPERTIES

Selected mechanical properties of bioactive glass 45S5 and, for comparison, the properties of bone and other bioactive ceramics are presented in Table 1.^[3] In general, bioactive ceramics present a large modulus of elasticity that determines their use as substitutes for hard tissue replacement. The elastic modulus value of most metals and inert bioceramics is much larger than that of bone. This mismatch of elastic modulus may affect the stress distribution and transfer to the host tissue. The elastic modulus of bioactive glasses is a close match to that of bone, but a low fracture toughness limits the use of bioactive glass implants to low-load-bearing applications or to compression loads. One of the primary restrictions is the uncertain lifetime under the complex stress states, slow crack growth, and cyclic fatigue that occur in many clinical applications. Therefore, current clinical applications of bioactive glasses are restricted to a few implant types or to particulate formulations used to fill spaces and improve the tissue repair processes, as will be presented in the next section.

One way to circumvent the limitation imposed by the mechanical properties of bioactive glasses is to combine them with other materials, improving mechanical behavior while still maintaining the bioactive characteristics of the surface. Two main approaches have been studied. The first approach is the use of bioactive ceramic coatings on metallic substrates, in which the bone-bonding ability of the coating is associated with the mechanical behavior of the substrate, particularly the higher strength and fracture toughness. This approach has been investigated mainly for calcium phosphate coatings.^[2,3] HA coatings can be obtained by different techniques, with plasma spray coating generally preferred. HA has been used clinically in load-bearing dental and orthopedic prostheses, and such coatings are used as an alternative to PMMA "bone cement" fixation of orthopedic devices. Bioactive glass coatings have also been investigated.^[2]

The other creative approach is to use bioactive particles or fibers as the reinforcing phase in polymer



Table 1 Mechanical properties of bioactive glasses and other bioactive ceramics

Property	Cortical bone	Cancellous bone	Bioactive glass 45S5	Glass ceramic A/W	Dense HA (> 99.2%)
Density (g/cm ³)	1.6–2.1		2.66	3.07	3.16
Hardness (HV)	100–230		460	680	600
Compressive strength (MPa)	50–150	2–12		1,080	500–1,000
Flexural strength (MPa)	7–30		42	215	115–200
Young modulus (GPa)	2–12	0.05–0.5	35	218	80–110
Fracture toughness, K_{IC} (MPa.m ^{1/2})				2.0	1.0

matrix composites. Properties equivalent to bone can be achieved by stiffening a compliant biocompatible synthetic polymer, such as polyethylene, with a higher-modulus ceramic second phase, such as bioactive glass or a calcium phosphate powder.^[3,22] The effect is to increase the Young's modulus, to decrease the strain to failure relative to that of the polymer phase, and to increase the fracture toughness relative to that of the ceramic phase. The development of bioactive composites is a research area with large potential. Biomaterials with engineered properties can be obtained through control of structure variables such as composition, type, size, and form of the components, and interface characteristics and properties.^[22,23]

CLINICAL APPLICATIONS

Bioactive glasses may be produced in various forms depending on the repair function they will serve. One of the most important clinical applications of bioactive glasses involves the particulate form that is implanted in an injured bone site. The particulate bioactive glass material has been used in a wide variety of clinical applications where bone grafting is needed to fill spaces or augment the natural repair process.^[1-3] One example is the treatment of periodontal (gum) disease, a clinical problem that affects tens of millions of people. Use of the bioactive glass material rapidly leads to new bone formation around the bioactive glass particles, by means of the reaction mechanisms summarized earlier. Because of the speed of formation of the new bone, the epithelial tissues are stopped from migrating down the tooth, a common problem if nothing is used to fill the space between the tooth and the repairing bone. The junction between the tooth and the periodontal membrane is stabilized by the use of bioactive glass particulate, and the tooth is saved. Similar procedures can be used to augment the jawbone when necessary to improve the stability of dental implants. Other orthopedic and maxillofacial applications that use bioactive glass particulate include repair of total hip and knee replacements that have required revision surgery, spinal fusion, and frontal sinus and calvarial bone defect obliteration, among others. Bioactive glass granules with varying composition and particle size range are currently available for clinical use.^[3,24,25]

Bioactive glasses can also be used as implants or prostheses to substitute for bone tissue. One difficulty surrounding this approach is that it is limited to applications involving no (or low) load bearing, due to the higher elastic modulus, low strength, and low fracture toughness of the material compared to the surrounding tissue. One such application is the use of bioactive glass implants to replace the small bones of the middle ear (ossicles) damaged by chronic infection,

a clinical problem called conductive hearing loss.^[26] Bioinert implants do not bond to the eardrum, and therefore they gradually erode through the tissue and are extruded through the eardrum within two to three years. In contrast, highly bioactive glasses form a bond with the collagen of the eardrum and also bond firmly to the remaining bone of the stapes footplate. They are thereby anchored on both ends, which prevents extrusion. Sound conduction is excellent, and there is no fibrous tissue growth to impair sound transmission. Moreover, there is no micromotion at the implant-tissue interface, so the implant remains stable and functions for the lifetime of the patient. This type of implant has been used successfully for more than 15 years.

Another application of bioactive glass implants, based upon bonding to both soft tissues and bone, preserves the jawbone of patients who have had their teeth extracted prior to being fitted with dentures.^[27] This type of implant, called an endosseous ridge maintenance implant (ERMI), is injection-molded in the shape of a cone, with several sizes available to match the sizes of roots of extracted teeth. X rays show that the implants are stable after many years in bone and have prevented resorption (wearing away) of bone under the dentures. Bioactive glass is also used as a bulk implant to repair damaged orbital floors and for other maxillofacial reconstruction surgeries.

BIOACTIVE GLASS SCAFFOLDS FOR TISSUE ENGINEERING

In the 1980s, emphasis in the field of biomaterials has shifted from bioinert to bioactive materials. However, the improved interfacial stability achieved with bioactive fixation did not improve implant survivability. Other problems related to the biomechanical mismatch of elastic moduli, production of wear debris, and maintenance of a stable blood supply still occur. In addition, all present-day implants lack two of the most critical characteristics of living tissues: the ability to self-repair and the ability to modify their structure and properties in response to environmental factors, such as mechanical load.

These observations and the growing need for long-surviving implants have led to a new shift in biomaterials research, in which the emphasis is on regeneration of tissues rather than on tissue replacement. There are now two biologically oriented approaches:^[28] 1) tissue regeneration, which involves the use of a scaffold that can be implanted into a defect to guide and stimulate tissue regrowth in situ (the scaffold should resorb as the tissue grows, leaving no trace of damage); and 2) tissue engineering, in which the scaffolds are seeded with cells in vitro to produce the basis

of a tissue before implantation. Cells could be extracted from a patient, seeded on a scaffold of the desired architecture, and the replacement tissue grown in the laboratory, ready for implantation.

Recent findings at Imperial College London, England, have established that there is genetic control of the cellular response to class A bioactive glass materials.^[29] Seven families of genes are upregulated when primary human osteoblasts are exposed to the ionic dissolution products of bioactive glasses. The enhanced gene expression occurs very rapidly, within 48 hours. The results indicate that the ionic dissolution products of Bioglass 45S5 may increase IGF-II availability in osteoblasts by inducing the transcription of the growth factor and its carrier protein, and also by regulating the dissociation of this factor from its binding protein. The unbound IGF-II is likely to be responsible for the increase in cell proliferation observed in the primary osteoblast cell cultures, and also in human clinical cases. Similar bioactive induction of the transcription of extracellular matrix components and their secretion and self-organization into a mineralized matrix is responsible for the rapid formation and growth of bone nodules and the differentiation of the mature osteocyte phenotype.

The findings cited in the preceding paragraph indicate that bioactive glass materials are a very interesting option for bone tissue regeneration and tissue engineering. Research has been conducted in this direction to develop scaffolds based on the bioactive glasses.^[30–33] One of the main criteria for a scaffold is that it comprises an interconnected network of pores with diameters in excess of 100 μm (macroporous materials) that allow tissue ingrowth, vascularization, and nutrient delivery. These characteristics have been attained in sol-gel-derived bioactive glass scaffolds developed by a foaming procedure.^[30,31] Foaming was directly achieved by vigorous agitation of the sol, with the aid of surfactants. The scaffolds can be produced with interconnected macropores that have porosities in the range of 60–90% and interconnected pore diameters of up to 500 μm . Figure 4 shows a scanned electron microscopy (SEM) micrograph of a bioactive glass foam

with a composition of 60% SiO_2 , 36% CaO , and 4% P_2O_5 .

Another scaffold material based on bioactive glass with the 45S5 composition was produced using a different processing route.^[32,33] Bioactive glass particles were mixed with a polymeric material, dry-pressed, and heat-treated to eliminate the polymer, leading to scaffolds with varying pore size and porosity depending on synthesis parameters. The porous bioactive glass substrates were capable of supporting the *in vitro* proliferation and maturation of osteoblastlike cells. *In vivo* studies using two tissue-engineered constructs were conducted: 1) osteoprogenitor cells that were seeded onto the scaffold prior to implantation (primary); or 2) those that were culture-expanded to form bonelike tissue on the scaffold prior to implantation (hybrid). The porous scaffolds were conditioned by immersion in simulated body fluid before cell culture. The amount of bone significantly increased for both groups over time as the bioactive ceramic gradually resorbed by 40% at 12 weeks. The authors concluded that porous, surface-modified bioactive glass is a promising scaffold material for tissue-engineered bone repair.

CONCLUSION

Bioactive glasses present a controlled reactivity and induce a specific biological response that leads to the formation of a continuous interface between the implanted material and tissue. The complex physicochemical reactions on the material surface lead to the formation of a biologically active carbonated hydroxylapatite layer, structurally and chemically equivalent to the mineral phase of bone. The reactivity of the material and the rate of HCA formation are closely related to the formation of an interface between the material and bone tissue *in vivo*. Bioactivity occurs only within certain compositional limits for glasses in the $\text{Na}_2\text{O}-\text{K}_2\text{O}-\text{CaO}-\text{MgO}-\text{P}_2\text{O}_5-\text{SiO}_2$ systems. These glasses may be produced by an alternative sol-gel method, in which case the compositional range of

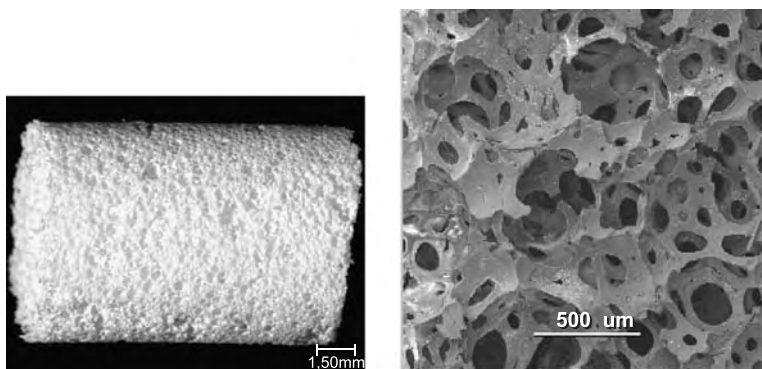


Fig. 4 Optical and SEM micrographs of bioactive gel-glass foam.

bioactivity and the resorbability of the material increase. The sol-gel method and other methods can also be used to generate bioactive glass scaffolds, with high porosity and interconnected pore size above 100 μm . The more bioactive glasses, considered to be class A bioactive materials, form a continuous bonded interface with soft tissues as well as bone. Class A bioactive glasses are both osteopductive and osteoconductive. There is a genetic control of the cellular response to class A bioactive glass materials resulting from controlled release of critical concentrations of Si and Ca ions. All of these features indicate that bioactive glasses are an interesting option for bone tissue engineering. Bioactive glass scaffolds contribute high bioactivity, controlled resorbability, and the potential for the ionic dissolution products (Si and Ca) to control the cell cycle of the osteoblast progenitor cells, stimulating genes in the bone cells to differentiate and enhance bone regeneration.

ARTICLES OF FURTHER INTEREST

Bioactive Materials and Scaffolds for Tissue Engineering; Biologic and Synthetic Apatites; Bone-Implant Interface; Host Reactions; Microporous Materials; Osteoinductive Substances and Materials

REFERENCES

1. Yamamuro, T.; Hench, L.; Wilson, J. *Handbook of Bioactive Ceramics*; CRC Press: USA, 1990; Vol. I.
2. Hench, L.L.; Wilson, J. *An Introduction to Bioceramics*; World Scientific: London, 1993.
3. Hench, L.L. Bioceramics. *J. Am. Ceram. Soc.* **1998**, *81*, 1705–1728.
4. Hench, L.L.; Splinter, R.J.; Allen, W.C.; Greenlee, T.K., Jr. Bonding mechanisms at the interface of ceramic prosthetic materials. *J. Biomed. Mater. Res.* **1971**, *2* (1), 117–141.
5. Li, R.; Clark, A.E.; Hench, L. An investigation of bioactive glass powders by the sol-gel process. *J. Appl. Biomater.* **1991**, *2*, 231–239.
6. Clark, A.E.; Pantano, C.G.; Hench, L.L. Auger spectroscopic analysis of bioglass corrosion films. *J. Am. Ceram. Soc.* **1976**, *59* (1–2), 37–39.
7. Hench, L.L.; Clark, A.E. Adhesion to Bone. In *Biocompatibility of Orthopedic Implants*; Williams, D.F., Ed.; CRC Press: U.S.A., 1982; Vol. 2.
8. Kokubo, T. Surface chemistry of bioactive glass-ceramics. *J. Non-Cryst. Solids* **1990**, *120*, 138–151.
9. Williams, D.F. The Biocompatibility and Clinical Uses of Calcium Phosphate Ceramics. In *Biocompatibility of Tissue Analogs*; Williams, D.F., Ed.; CRC Press: U.S.A., 1985; Vol. II, 43–66.
10. Ogino, M.; Ohuchi, F.; Hench, L.L. Compositional dependence of the formation of calcium phosphate films on bioglass. *J. Biomed. Mater. Res.* **1980**, *14*, 55–64.
11. Andersson, O.H.; Liu, G.; Karlsson, K.H.; Niemi, L.; Miettinen, J.; Juhanoja, J. In vivo behavior of glasses in the $\text{SiO}_2\text{-Na}_2\text{O-CaO-P}_2\text{O}_5\text{-Al}_2\text{O}_3\text{-B}_2\text{O}_3$. *Syst. J. Mater. Sci. Mater. Med.* **1990**, *1*, 219–227.
12. Brink, M.; Turunen, T.; Happonen, R.P.; Yli-Urpo, A. Compositional dependence of bioactivity of glasses in the system $\text{Na}_2\text{O-K}_2\text{O-MgO-CaO-B}_2\text{O}_3\text{-P}_2\text{O}_5\text{-SiO}_2$. *J. Biomed. Mater. Res.* **1997**, *37*, 114–121.
13. Andersson, O.H.; Rosenqvist, J.; Karlsson, K.H. Dissolution, leaching, and Al_2O_3 enrichment at the surface of bioactive glasses studied by solution analysis. *J. Biomed. Mater. Res.* **1993**, *27*, 941–948.
14. Schepers, E.J.; Ducheyne, P. Bioactive glass particles of narrow size range for the treatment of oral bone defects: A 1–24 month experiment with several materials and particle sizes and size ranges. *J. Oral Rehabil.* **1997**, *24*, 171–181.
15. Kaufmann, E.A.; Ducheyne, P.; Radin, S.; Bonnell, D.A.; Composto, R. Initial events at the bioactive glass surface in contact with protein-containing solutions. *J. Biomed. Mater. Res.* **2000**, *52*, 825–830.
16. Pereira, M.M.; Hench, L.L. Mechanisms of hydroxyapatite formation on porous gel-silica substrates. *J. Sol-Gel Sci. Technol.* **1996**, *7*, 59–68.
17. Kokubo, T.; Kim, H.M.; Kawashita, M. Novel bioactive materials with different mechanical properties. *Biomaterials* **2003**, *24*, 2161–2175.
18. Greenspan, D.C.; Zhong, J.P.; Wheller, D.L. Bioactivity and Biodegradability: Melt vs. Sol-Gel Derived Bioglass[®] In Vitro and In Vivo, Proceedings of the 11th International Symposium on Ceramics in Medicine, New York, Nov. 5–8, 1998; LeGeros, R.Z., LeGeros, J.P., Eds.; World Scientific: New York, 1998; 345–348.
19. Gross, U.; Kinne, R.; Schmitz, H.J.; Strunz, V. The response of bone to surface active glass/glass-ceramics. *CRC Crit. Rev. Biocompat.* **1988**, *4*, 2.
20. Yoshii, S.; Kakutani, Y.; Yamamuro, T.; Nakamura, T.; Kitsugi, T.; Oka, M.; Kokubo, T.; Takagi, M. Strength of bonding between A-W glass ceramic and the surface of bone cortex. *J. Biomed. Mater. Res.* **1988**, *22* (A), 327.
21. Wilson, J.; Pigott, G.H.; Schoen, F.J.; Hench, L.L. Toxicology and biocompatibility of bioglass. *J. Biomed. Mater. Res.* **1981**, *15*, 805.
22. Wang, M.; Joseph, R.; Bonfield, W. Hydroxyapatite-polyethylene composites for bone substitution: Effects of ceramic particle size and morphology. *Biomaterials* **1998**, *19*, 2357–2366.
23. Wang, M.; Bonfield, W. Chemically coupled hydroxyapatite-polyethylene composites: Structure and properties. *Biomaterials* **2001**, *22*, 1311–1320.
24. Peltola, M.J.; Suonpää, J.T.; Määttänen, H.S.; Varpula, M.J.; Aitasalo, K.M.; Yli-Urpo, A.; Laippala, P.J. Clinical follow-up method for frontal sinus obliteration with bioactive glass S53P4. *J. Biomed. Mater. Res.* **2001**, *58*, 54–60.
25. Schepers, E.J.; Ducheyne, P.; Barbier, L.; Schepers, S. Bioactive glass particles of narrow size range: A new material for the repair of bone defects. *Implant Dent.* **1993**, *2*, 151–156.

26. Wilson, J.; Douek, E.; Rust, K. Bioglass Middle Ear Devices: Ten Year Clinical Results. In *Bioceramics*; Wilson, J., Hench, L.L., Greenspan, D., Eds.; Pergamon/Elsevier: Oxford, UK, 1995; Vol. 8, 239–246.
27. Stanley, H.R.; Hall, M.B.; Clark, A.E.; King, J.C.; Hench, L.L.; Berte, J.J. Using 45S5 bioglass cones as endosseous ridge maintenance implants to prevent alveolar ridge resorptions—A 5 year evaluation. *Int. J. Oral Maxillofac. Implants* **1997**, *12*, 95–105.
28. Jones, J.R.; Hench, L.L. Biomedical materials for the new millennium: A perspective on the future. *J. Mater. Sci. Technol.* **2001**, *17*, 891–900.
29. Hench, L.L.; Polak, J.M.; Xynos, I.D.; Buttery, L.D.K. Bioactive materials to control cell cycle. *Mater. Res. Innov.* **2000**, *3*, 313–323.
30. Sepulveda, P.; Jones, J.R.; Hench, L.L. Bioactive sol-gel foams for tissue repair. *J. Biomed. Mater. Res.* **2002**, *59*, 340–348.
31. Coelho, M.B.; Soares, I.R.; Mansur, H.S.; Pereira, M.M. Effect of the type of surfactant on bioactive glasses foam formation. *Key Eng. Mater.* **2003**, *240–242*, 257–260.
32. Kaufmann, E.A.B.E.; Ducheyne, P.; Shapiro, I.M. Effect of varying physical properties of porous, surface modified bioactive glass 45S5 on osteoblast proliferation and maturation. *J. Biomed. Mater. Res.* **2000**, *52*, 783–796.
33. Livingston, T.; Ducheyne, P.; Garino, J. In vivo evaluation of a bioactive scaffold for bone tissue engineering. *J. Biomed. Mater. Res.* **2002**, *62*, 1–13.

Bioactive Materials and Scaffolds for Tissue Engineering

Qi-Zhi Chen

Aldo R. Boccaccini

Department of Materials, Imperial College London, London, United Kingdom

INTRODUCTION

Bioactive materials react with physiological fluids and form tenacious bonds to hard tissue through biological interdigitation of collagen fibers with the material surface layer. Thus, these biomaterials can be used to transfer loads to and from living bone. Prominent bioactive materials are inorganic, such as bioceramics, including selected compositions of silicate glasses and their derived glass–ceramics, as well as hydroxyapatite (HA) and related calcium phosphates. Like most ceramic materials, the major disadvantage of bioactive ceramics is their low fracture toughness (i.e., brittleness). They are thus often used in composites combined with polymers, both stable, e.g., poly(methyl methacrylate), and biodegradable, e.g., aliphatic polyesters. For tissue engineering applications, bioactive ceramics and the related bioactive composites are produced in highly porous 3-D structures known as tissue “scaffolds.” For bone tissue engineering, porosity of $\sim 90\%$ and pores of the order of $\sim 400\ \mu\text{m}$ are desirable, as well as high pore interconnectivity, to facilitate the attachment and proliferation of cells and the ingrowth of new tissues into the network, as well to enable mass transport of oxygen, nutrition, and waste products. Numerous methods to produce 3-D tissue scaffolds have been developed.^[1,2]

In this chapter, we describe first the key design factors of hard tissue engineering scaffolds, followed by an up-to-date review on bioactive materials. Then we focus on the state of the art in the field of fabrication technologies for bioactive and biodegradable scaffolds based on bioactive ceramics and related ceramic–polymer composites designed for bone tissue regeneration. The goal of this chapter is to allow readers to fully understand the advantages and disadvantages of various bioactive materials and scaffold processing technologies.

KEY DESIGN FEATURES OF TISSUE ENGINEERING SCAFFOLDS

An ideal scaffold for bone tissue regeneration must satisfy the following essential criteria determined by

both the chemistry of the material and the required morphology (pore structure):^[1,2]

1. Ability to deliver cells. The material should not only be biocompatible (i.e., non-toxic), but also foster cell attachment, differentiation, and proliferation.
2. Osteoconductivity. The material should encourage osteoconduction with the host bone. Osteoconductivity not only eliminates the formation of encapsulating tissue but also brings about a strong bond between the scaffold and host bone.
3. Biodegradability. The composition of the material, combined with the porous structure of the scaffold, should lead biodegradation in vivo at rates appropriate to tissue regeneration.
4. Mechanical properties. The mechanical strength of the scaffold, which is determined by both the properties of the biomaterial and the porous structure, should be sufficient to provide mechanical stability to constructs in load-bearing sites prior to the synthesis of new extracellular matrix by cells.

It is also equally important to meet requirements related to the fabrication technologies, as follows:

5. Porous structure. The scaffold should have an interconnected 3-D porous structure with porosity $>90\%$ and pore diameters in the range $300\text{--}500\ \mu\text{m}$ for cell penetration, tissue ingrowth and vascularization, and nutrient delivery.
6. Processability. The material should possess desired fabrication properties, e.g., the possibility of producing irregular shapes of 3-D scaffolds that match specific bone defects in individual patients.
7. Commercialization. The synthesis of the material and fabrication of the scaffold should be suitable to commercialization.

BIOACTIVE MATERIALS

In this section, the biocompatibility, bioactivity, biodegradability, and mechanical properties of various bioactive materials are reviewed. These properties

should be considered in the design of a tissue regeneration scaffold.

Calcium Phosphates

Because almost two thirds of the weight of bone is HA, $\text{Ca}_{10}(\text{PO}_4)_6(\text{OH})_2$, it is logical to consider this bio-ceramic and related calcium phosphates (e.g., β - or tricalcium phosphate) the major components of scaffold materials for bone tissue engineering. Calcium phosphates have excellent biocompatibility as a result of their close chemical and crystal resemblance to bone mineral. They possess osteoconductive properties, as well as a remarkable ability to bind to bone through biological interdigitation of collagen fibers with calcium phosphate surface layers.^[3] Numerous *in vivo* and *in vitro* assessments have reported that calcium phosphates, regardless of their morphology (bulk, coating, powder, or porous structures) and which phases they are in (crystalline or amorphous), always support the attachment, growth, differentiation, and proliferation of relevant cells (such as osteoblasts and mesenchymal cells).

Typically, crystalline calcium phosphates have long degradation time *in vivo*, often of the order of years. Moreover, the dissolution rates of synthetic HA depend on the type and concentration of the buffered or unbuffered solutions, pH of the solution, degree of the saturation of the solution, solid/solution ratio, and composition and crystallinity of the HA. In the case of highly crystalline HA, the degree of micro- and macroporosity, defect structure, and amount and type of other phases present also have significant influence. Crystalline HA exhibits the slowest degradation rate compared with other calcium phosphates. The dissolution rate decreases in the following order: amorphous tricalcium phosphate > amorphous HA \gg all other crystalline calcium phosphates \gg crystalline HA.

The mechanical properties of synthetic calcium phosphates also vary significantly with their crystalli-

nity, grain size, porosity, and composition (e.g., calcium deficiency). In general, the mechanical strength of synthetic calcium phosphates decreases significantly with increasing content of amorphous phase, microporosity, and grain size. High crystallinity, low porosity, and small grain size tend to give higher stiffness, higher compressive and tensile strength, and greater fracture toughness. A comparison of the properties of HA and related calcium phosphates with those of cortical bone (Table 1) reveals that cortical bone has a reasonably good compressive strength, although lower than that of HA, and better tensile strength and significantly better fracture toughness than HA. The apatite crystals in bone tissue make it strong enough to tolerate compressive loading. The high tensile strength and fracture toughness of bone are attributed to the tough and flexible collagen fibers. Hence, calcium phosphates alone cannot be used for load-bearing scaffolds despite their good biocompatibility and osteoconductivity.

In summary, although the excellent biological performance of HA and related calcium phosphates has been well documented, the slow biodegradation of their crystalline phases and the low mechanical strength in their amorphous state limit their application in engineering of new bone tissue, especially at load-bearing sites.

Bioactive Silicate Glasses

As early as 1969, Hench and colleagues discovered that certain silicate glass compositions had excellent biocompatibility, as well as the ability of bone bonding.^[3] Through interfacial and cell-mediated reactions, bioactive glass develops a carbonated calcium phosphate surface layer that allows it to bond to host bone. This bone-bonding behavior (i.e., bioactivity) has been associated with the formation of a carbonated HA layer on the glass surface when implanted or in contact with biological fluids. Although many details of the

Table 1 Mechanical properties of solid bioactive materials and human cortical bone^[4]

Ceramics	Compressive strength (MPa)	Tensile (T) or bending (B) strength (MPa)	Elastic modulus (GPa)	Fracture toughness (MPa $\sqrt{\text{m}}$)
Hydroxyapatite	>400	~40 (T)	~100	~1.0
Other calcium phosphates	20–900	30–200 (T)	30–103	<1.0
45S5 Bioglass [®]	~500	42 (T)	35	0.5–1
Apatite–wollastonite (A–W)	1080	215 (B)	118	2.0
Parent glass of A–W	NA	72 (B)	NA	0.8
Bioverit [®] I	500	140–180 (B)	70–90	1.2–2.1
Cortical bone	130–180	50–151 (T)	12–18	6–8

bone-bonding mechanisms remain unknown at present, it is clearly recognized that for bonding with bone tissue to occur a layer of biologically active carbonated HA must form.^[3] This is in agreement with the fact that HA has a high capacity to bind the collagens of natural bone. Bioactivity is not an exclusive property of bioactive glasses. Hydroxyapatite and related calcium phosphates also demonstrate an excellent ability to bond to bone, as described above. The capability of a material to form a biological interface with the surrounding tissue is critical in eliminating scaffold loosening upon implantation (similar to orthopedic implants).

Bioactive glasses also support enzyme activity,^[3] enhance vascularization,^[5] foster osteoblast adhesion, growth, and differentiation, and induce the differentiation of mesenchymal cells into osteoblasts.^[6] An important finding is that the dissolution products from bioactive glasses exert a genetic control over the osteoblast cell cycle, affecting the rapid expression of genes that regulate osteogenesis and the production of growth factors.^[7] The above-mentioned advantages make bioactive glasses, in particular the composition 45S5 Bioglass,^{®[3]} the material of choice for several clinical applications, e.g., for treatment of periodontal diseases and as a bone filler material. Bioglass implants have also been used to replace damaged middle ear bone, restoring hearing to patients. More recently, bioactive glasses have gained attention as promising scaffold materials for tissue engineering.^[8]

The key advantage that makes bioactive glasses promising scaffold materials is the possibility of controlling their chemical composition and, thus, the rate of biodegradation. The structure and chemistry of glasses, in particular sol-gel derived glasses, can be tailored at a molecular level by varying either composition or thermal or environmental processing history. It is, in principle, possible to design glasses with degradation properties specific to a particular application of bone tissue engineering.

The primary disadvantages of bioactive glasses are their low mechanical strength and fracture toughness (Table 1). Hence, bioactive glasses on their own have limited applications in load-bearing situations. However, these materials can be used in combination with degradable biopolymers to form composite materials with tailored resorbability and bone repair capability, offering great potential for tissue engineering scaffold development.^[4]

Bioactive Glass–Ceramics

Glasses can be strengthened by the formation of a crystalline phase upon controlled heat treatment of the amorphous glass, forming glass–ceramics. Many

bioactive silicate glass–ceramics were developed over the years for the repair of damaged bone. Among them, apatite–wollastonite (A–W), Ceravital[®], and Bioverit[®] glass–ceramics have long been investigated, as reviewed by Hench and Wilson.^[9] However, these silicate glass–ceramics have not been considered extensively for tissue engineering scaffolds. Recently, a 45S5 Bioglass-derived glass–ceramic was developed specifically for the fabrication of bioactive and biodegradable scaffolds.^[8]

Crystallization of bioactive glasses tends to decrease the level of bioactivity and can even turn a bioactive glass into an inert material. This is one of the disadvantages that has limited the application of melt-derived bioactive glasses as scaffold materials, because full crystallization usually occurs prior to significant densification during heat treatment (i.e., sintering) of glass powders to form the scaffold structure. Extensive sintering is necessary to densify the struts of a (foam-like) scaffold, which would otherwise be made up of loosely packed particles and thus would be too fragile to handle. It is apparent that mechanical stability and biodegradability are difficult to achieve simultaneously in a single scaffold. Mechanically stable, usually crystalline, materials are generally bioinert, whereas degradable materials, which are usually amorphous or weakly crystalline, tend to be weak. This issue is not only relevant in bioactive silicate glasses, but also in HA and associated calcium phosphates, as discussed in the previous section.

A recent investigation at Imperial College London^[8] addressed the problem, focusing on sintered 45S5 Bioglass foam-like scaffolds. It was demonstrated that the mechanically strong crystalline phase $\text{Na}_2\text{Ca}_2\text{Si}_3\text{O}_9$, which forms during sintering of 45S5 Bioglass at 1000°C, can transform into a biodegradable amorphous calcium phosphate upon immersion in simulated body fluid at body temperature. This discovery led to the successful development of highly porous glass–ceramic scaffolds from melt-derived 45S5 Bioglass. Fig. 1 illustrates a typical scaffold microstructure. These novel scaffolds are superior to their HA counterparts in terms of both mechanical strength and biodegradability. Furthermore, a recent quantitative study revealed that the proliferation of osteoblast cells on the glass–ceramic materials is comparable to that on the non-crystallized 45S5 Bioglass.

Bioactive Composites

From a biological perspective, it is a natural strategy to combine polymers and ceramics to fabricate scaffolds for bone tissue engineering because native bone is the combination of a naturally occurring polymer and biological apatite. Moreover, from the materials

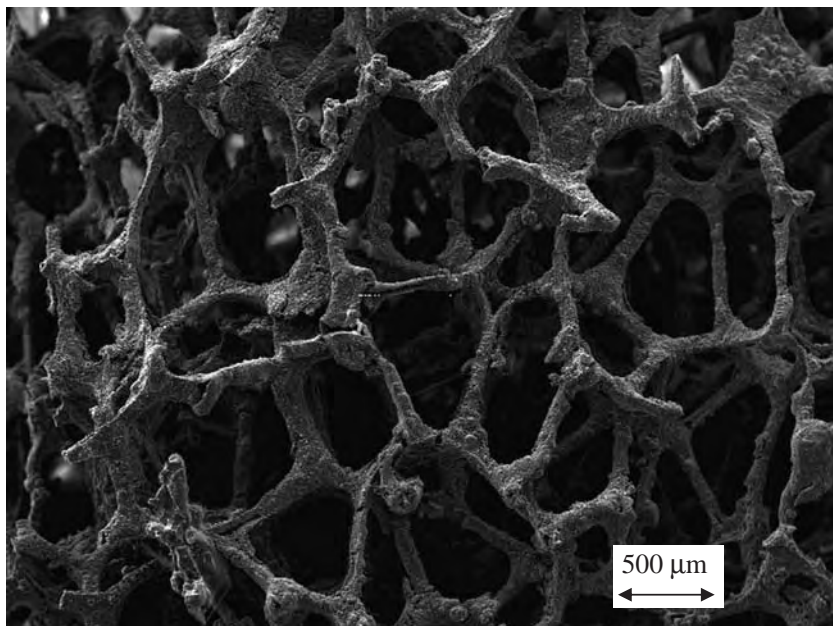


Fig. 1 Scanning electron microscopy image of a Bioglass[®]-derived glass-ceramic scaffold developed using the foam replica technique.^[8]

B

science point of view, a single material type does not always provide the necessary mechanical or chemical properties desired for a particular application. In these instances, composite materials designed to combine the advantages of both materials may be most appropriate.^[4] Polymers and ceramics that degrade *in vivo* should be chosen for designing biocomposites for tissue engineering scaffolds. Although massive release of acidic degradation from biopolymers, e.g., poly(lactic acid), causes inflammatory reactions,^[10] the basic degradation of calcium phosphate or bioactive silicate glasses would buffer the acidic by-products of the polymers and may eventually help to avoid the formation of an unfavorable environment for cells because of a decreased pH. Mechanically, bioceramics are much stronger than polymers and play a critical role in providing mechanical stability and structural integrity to scaffolds prior to the synthesis of new extracellular matrix by the cells upon implantation. However, ceramics and glasses are very fragile as a result of their intrinsic brittleness and flaw sensitivity, whereas polymers are flexible. To capitalize on their advantages and minimize their shortcomings, ceramics and glasses are combined with biopolymers to form bioactive composite biomaterials for fabrication of scaffolds.^[4]

Table 2 lists selected ceramic/glass-polymer composites, which have been developed as biomedical devices or scaffold materials for bone tissue engineering, and their mechanical properties. In general, all these synthetic composites exhibit outstanding biocompatibility and bioactivity. The research team led by

Laurencin, for instance, synthesized porous scaffolds based on poly(lactid-co-glycolic-acid) (PLGA) and HA, which were reported to combine the degradability of PLGA with the bioactivity of HA, fostering cell proliferation and differentiation, as well as bone mineral formation.^[11] Related composites of bioactive glass and polylactic acid (PLA) formed calcium phosphate layers on their surfaces and support rapid and abundant growth of human osteoblasts and osteoblast-like cells when cultured *in vitro*.^[4] Highly porous composites have also been produced from combinations of a poly(D,L-lactic acid) (PDLLA) matrix and Bioglass particle inclusions.^[12] These scaffolds have a well-defined and highly interconnected porous structure. At the same time, their mechanical properties are close to (but lower than) that of cancellous bone. Recently, PDLLA/Bioglass composite scaffolds containing up to 40 wt% Bioglass enhanced bone formation *in vivo* (animal model).^[12]

Table 3 summarizes the advantages and disadvantages of the bioactive materials discussed in the preceding sections.

FABRICATION OF BIOACTIVE SCAFFOLDS FOR TISSUE ENGINEERING

Bioactive Ceramic Scaffolds

Porous ceramics can be produced by a variety of different processes, which may be classified into two categories: 1) manual-based processing techniques and

Table 2 Bioactive composites designed for bone tissue engineering^[4]

Bioactive composite		Polymer	Percentage of ceramic (%)	Porosity (%)	Pore size (μm)	Compressive (C), tensile (T), flexural (F) strength (MPa)			Toughness (kJ/m^2)
Ceramic	Modulus (MPa)					Ultimate strain (%)	Technique		
1. Dense composites									
HA fiber									
		PDLLA	2–10.5 (vol)			45 (F)	$1.75\text{--}2.47 \times 10^3$		
		PLLA	10–70 (wt)			50–60 (F)	$6.4\text{--}12.8 \times 10^3$	0.7–2.3	
HA		PLGA	40–85 (vol)			22 (F)	1.1×10^3		5.29
		Chitosan	40–85 (vol)			12 (F)	2.15×10^3		0.092
		Chitosan + PLGA	40–85 (vol)	NA		43 (F)	2.6×10^3		9.77
		PPhos	85–95 (wt)						
		Collagen	50–72 (wt)						
β -TCP		PLLA-co-PEH	75 (wt)			51 (F)	5.18×10^3		
		PPF	25 (wt)			7.5–7.7 (C)	191–134		
A/W		PE	10–50 (vol)			18–28 (B)	$0.9\text{--}5.7 \times 10^3$		
$\text{Ca}_3(\text{CO}_3)_2$		PLLA	30 (wt)			50	$3.5\text{--}6 \times 10^3$		
Human cortical bone						50–150 (T)	$12\text{--}18 \times 10^3$		
						130–180 (C)			
2. Porous composites									
Amorphous CaP									
		PLGA	28–75 (wt)	75	> 100		65		Microsphere
β -TCP		Chitosan-gelatin	10–70 (wt)		322–355	0.32–0.88 (C)	3.94–10.88		TIPS
HA		PLLA	50 (wt)	85–96	100 \times 300	0.39 (C)	10–14		TIPS
		PLGA	60–75 (wt)	81–91	800–1800	0.07–0.22 (C)	2–7.5		SC/PL
		PLGA	75 (wt)	30–40	110–150		337–1459		
Bioglass [®]		PLGA	20–50 (wt)	43	89	0.42 (C)	51		Microsphere
		PLLA	20–50 (wt)	77–80	\sim 100 (macro)	1.5–3.9 (T)	137–260	1.1–13.7	SC/PL
					\sim 10 (micro)				
		PLGA	0.1–1 (wt)		50–300				
		PDLLA	5–29 (wt)	94	\sim 100 (macro)	0.07–0.08 (C)	0.65–1.2	7.21–13.3	SC/PL
					10–50 (micro)				
Phosphate		PLA-PDLLA	20	90	400–600	0.2–0.7(C)	NA	NA	Foam-coating
Glass A-W		PDLLA	20–40 (wt)	85.5–95.2	98–154	0.017–0.020 (C)	0.075–0.12		SC/PL

Table 3 Advantages and disadvantages of bioactive biomaterials used in bone tissue engineering

Biomaterial	Positive	Negative
Calcium phosphates (e.g., HA, tricalcium phosphate)	(1) Excellent biocompatibility (2) Supporting cell activity (3) Good osteoconductivity	(1) Brittle (2) Generally, they biodegrade too slowly in their crystalline phase, and they are too weak in their amorphous phase
Bioactive glasses	(1) Excellent biocompatibility (2) Support cell activity (3) Good osteoconductivity (4) Enhance vascularization (5) Control of osteoblast gene expression (6) Tailorable degradation rate	(1) Mechanically brittle and weak
Bioactive glass-ceramic	(1) Excellent biocompatibility. (2) Supporting cell activity (3) Good osteoconductivity (4) 45S5 Bioglass [®] -derived glass-ceramics are mechanically comparable to hydroxyapatite	(1) Brittle (2) Generally, they degrade too slowly, (except for crystallized 45S5 Bioglass [®])
Bioactive composite	(1) Excellent biocompatibility (2) Supporting cell activity (3) Good osteoconductivity (4) Tailorable degradation rate (5) Improved mechanical properties	(1) Processing techniques for 3-D porous scaffolds can be complex

B

2) computer-controlled fabrication processes, i.e., solid free-form (SFF) technologies. Manual-based processing techniques can further be divided into conventional powder-forming processes and sol-gel techniques.

Powder-based methods

The powder-forming process starts with the preparation of a slurry, which is a suspension of ceramic powders in a suitable liquid (e.g., water and ethanol) that will be used to prepare green bodies. A green body is porous, and its structure largely determines that of the sintered product. After a programmed heat treatment of the green body, a porous ceramic is produced.

The methods of obtaining green bodies for 3-D porous ceramics include loose packing, compaction, slip casting, injection molding, phase separation, freeze-drying, polymer replication, and gel casting. These processes lead to different porous structures within green bodies. Here we focus on one of the most frequently used techniques for the production of highly porous 3-D bioactive scaffolds: the replication technique.

The foam replication technique is also called the “polymer-sponge” method. In this process, the green

bodies of ceramic foams are prepared by coating a polymer (e.g., polyurethane) foam with a ceramic slurry. The polymer foam, having already the desired pore macrostructure, simply serves as a sacrificial template for the ceramic coating. The polymer template is immersed in the slurry, which subsequently infiltrates the structure and ceramic particles coat the surface of the polymer substrate. Excess slurry is squeezed out, leaving a uniform ceramic coating on the foam struts. After being dried, the polymer is slowly burned out to minimize damage to the porous coating. Once the polymer has been removed, the ceramic is sintered to the desired density. The process replicates the macroporous structure of the polymer foam and results in a rather distinctive microstructure within the struts. This method has been applied for the preparation of highly porous bone tissue engineering scaffolds from a variety of materials including hydroxyapatite,^[13] SiO₂-CaO-CaF₂-Na₂O-K₂O-P₂O₅-MgO glass-ceramic,^[14,15] calcium silicate (CaSiO₃),^[16] calcium phosphates,^[17] and 45S5 Bioglass (Fig. 1).^[8]

Sol-gel methods

The sol-gel process can produce ceramic or glass materials in a wide variety of forms: ultrafine or

spherical-shape powders, thin film coatings, ceramic fibers, microporous inorganic membranes, monolithic ceramics and glasses, or extremely porous aerogel materials. In the processing, alkoxide precursors, such as tetraethyl orthosilicate (TEOS) and triethoxyl orthophosphate (TEP), undergo hydrolysis and condensation reactions to form a sol. Polymerization of -Si-OH groups continue after the completion of hydrolysis, beginning the formation of the silicate (-Si-O-Si-) network. The network connectivity increases until it spans the solvent medium. Eventually, a wet gel forms. The wet gel is then subjected to controlled thermal processes of aging to strengthen the gel, drying to remove the liquid by-product of the polycondensation reaction, and thermal stabilization (or sintering) to remove organic species from the surface of the material and to further densify the structure; as a result, a relatively strong, porous aerogel is produced.

Highly porous bioactive glass foams have been developed by mechanically stirring the sol with the use of a surfactant and catalysts.^[18] The precursors of the glass foams are $\text{Ca}(\text{NO}_3)_2$ and two alkoxides: TEOS and TEP. Sol-gel-derived bioactive glass foams demonstrated favorable results in both in vitro and in vivo tests for bone regeneration.^[18] However, the sol-gel process for cellular bioactive glasses involves hazardous chemicals, such as hydrofluoric acid (HF), which is used as a catalyst. A typical structure of a scaffold fabricated from the silica-calcia system by the sol-gel method is presented in Fig. 2.

Solid free-form techniques

Solid free-form techniques, also known as rapid prototyping (RP), are computer-controlled fabrication processes. They can rapidly produce highly complex 3-D objects using dates generated by computer-aided design (CAD) systems. In a typical application, an image of a bone defect in a patient can be taken and used to develop a 3-D CAD computer model. The computer can then reduce the model to slices or layers. The 3-D objects are constructed layer by layer using rapid prototyping techniques such as fused deposition modeling (FDM), selective laser sintering (SLS), 3-D printing (3D-P), or stereolithography. Calcium phosphate scaffolds have been produced using FDM, SLS, 3D-P,^[19] and stereolithography, as well as RP combined with the replication technique. However, SFF is not a cost-effective method compared with other techniques described above.

Generally, highly porous bioceramic scaffolds are very brittle. An approach to improve the mechanical properties of ceramic scaffolds is to coat the porous network with a polymeric layer.^[20] By comparing the porous structures of ceramic scaffolds made by different techniques with that of trabecular bone, it was possible to conclude that the porous structure produced by the replication technique^[8] is most similar to that of cancellous bone and contains completely interconnecting pores and solid material, forming only the foam struts. The ceramic foams synthesized by gel-casting and sol-gel techniques come next in terms of

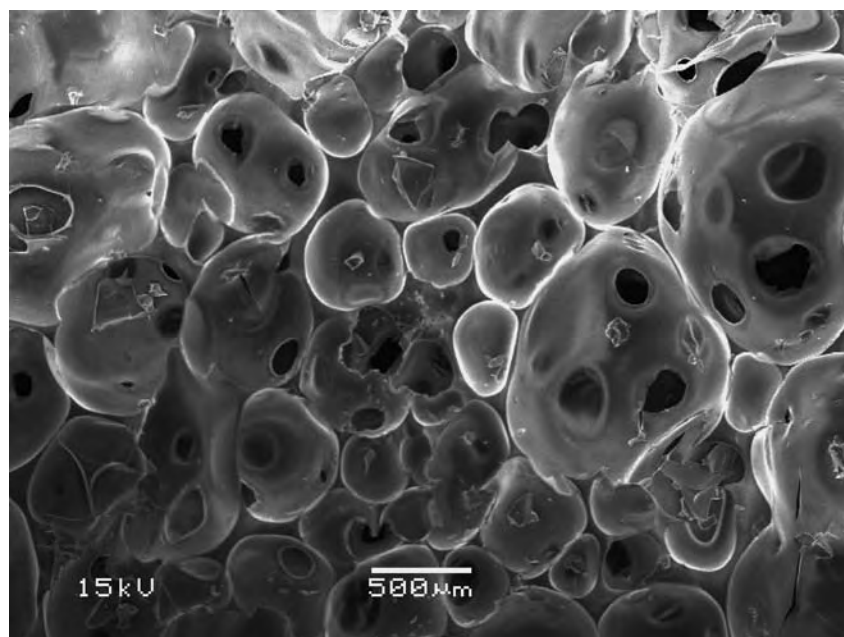


Fig. 2 Scanning electron microscopy image of a bioactive glass (CaO-SiO_2) scaffold fabricated using the sol-gel method.

Table 4 Advantages of the replication technique for production of bioactive ceramic scaffolds

1.	Cancellous bone-like macroporous structure The porous structure produced by the replication technique is very similar to cancellous bone: highly porous and exhibiting open pores of size $\sim 400 \mu\text{m}$.
2.	High commercialization potential This technique is the simplest and most cost-effective method and thus most suitable for commercialization. Rapid-prototyping methods are expensive processes, but may, however, be suitable for producing specific and complex scaffold architectures.
3.	Safety The technique does not involve any toxic chemicals compared with sol-gel and gel-casting techniques, which use HF to accelerate polymerization.
4.	Irregular or complex shape production ability The technique can produce scaffolds of irregular or complex shapes compared with standard dry powder processing.

structural similarity to cancellous bone. Table 4 summarizes the advantages of the replication technique for production of foam-like scaffolds for bone tissue engineering.

Bioactive Composite Scaffolds

The development of fabrication techniques for ceramic-polymer composite scaffolds followed previous achievements in the area of polymeric scaffold production. Numerous techniques have been developed to process porous polymer scaffolds for use in tissue engineering and several reviews exist in the literature.^[1,2,4,21,22] Two of the polymer-foaming processes, namely solvent casting combined with particle leaching and thermally induced phase separation (TIPS) combined with freeze-drying, have been applied to the fabrication of polymer-ceramic composite scaffolds.^[4] In addition, two other methods of significance exist: microsphere sintering and foam coating were developed for the combination of ceramic and polymeric materials with the aim of producing composite scaffolds.

Solvent casting/particle leaching (SC/PL)

In this process, polymer microspheres are first formed from traditional oil/water emulsions. Solvent-aggregated polymer-ceramic scaffolds can then be constructed by

mixing solvent, salt or sugar particles, ceramic granules, and prehardened microspheres. A 3-D structure of controlled porosity is formed based on this method, combined with particle leaching and microsphere packing. The main advantage of this processing technique is the ease of fabrication without the need for specialized equipment. The primary disadvantages of solvent casting are: 1) the limitation in the shapes that can be produced (typically flat sheets and tubes are the only possible shapes); 2) the possible retention of toxic solvent within the polymer; and 3) the denaturation of the proteins and other molecules incorporated into the polymer by the use of solvents. The use of organic solvents to cast the polymer may decrease the activity of bioinductive molecules (e.g., proteins). Another potential drawback of the method is the limited interconnectivity of the pore structures that can be achieved.

Thermally induced phase separation

Porous structures can also be achieved through phase separation and evaporation. One approach to induce phase separation in polymers is to lower the temperature of the suspension of the polymer containing the ceramic additions, e.g., HA or bioactive glass particles. The solvent is solidified first, forcing the polymer and ceramic mixture into the interstitial spaces. The frozen mixture is then lyophilized using a freeze-dryer, in which the ice solvent evaporates. The typical microstructure of a PDLA/Bioglass (5 wt% Bioglass) composite foam scaffold obtained by TIPS is illustrated in Fig. 3, which indicates the high porosity (>90%) and pore interconnectivity achieved.^[12]

Microsphere sintering

In this process, composite microspheres composed of a polymer matrix and ceramic inclusions are synthesized first, using the emulsion/solvent evaporation technique. Sintering the composite microspheres together yields a 3-D porous scaffold.^[23]

Foam coating techniques

An alternative approach to combine a polymeric scaffold and a bioactive ceramic is to coat bioactive ceramic particles onto polymeric foams. This process was successfully demonstrated for the first time on bioactive glass-coated PDLA foams.^[24] Alternatively, a calcium phosphate layer can be grown on the surface of polymeric scaffolds in situ by a biomimetic method, e.g., by immersion of the scaffold in a simulated body fluid.^[25] The inverse method, namely polymer-coated

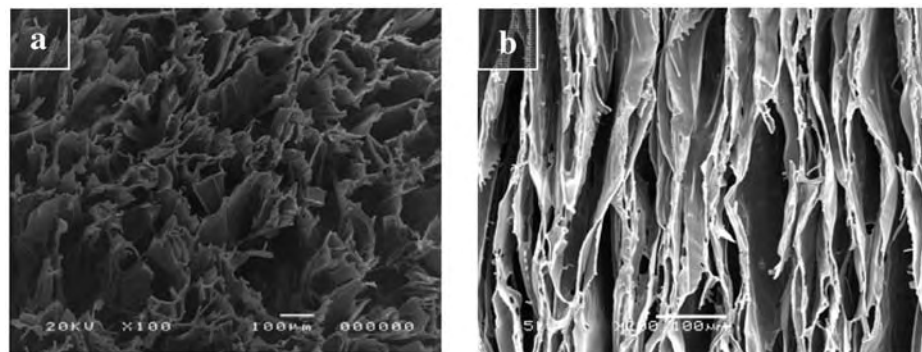


Fig. 3 Scanning electron microscopy images of a PDLLA/Bioglass composite scaffold fabricated by the TIPS technique (5 wt% Bioglass): (a) orthogonal to the pore direction and (b) parallel to the pore direction.^[4,12]

ceramic scaffolds, has also been investigated, for example, by applying a PDLLA coating layer on 45S5 Bioglass-derived glass-ceramic foams.^[20] The bioactivity of Bioglass-based foams was maintained after application of PDLLA coatings, whereas the mechanical stability of the scaffolds was considerably enhanced by the coating. The work of fracture of the PDLLA-coated foams is, on average, approximately 20 times greater than that of the uncoated foams. More importantly, in simulated body fluid the mechanical stability of the scaffolds was greatly improved by the PDLLA coating, which is the result of the in situ formation of a nanocomposite film of PDLLA and hydroxyapatite crystals on the strut surface of the foams.^[20] A similar approach has been investigated for calcium phosphate scaffolds with interconnected macropores ($> 200 \mu\text{m}$) and microporosity ($\sim 5 \mu\text{m}$), which were infiltrated with poly(lactic-co-glycolic acid) (PLGA) to achieve an interpenetrating bioactive ceramic/biodegradable polymer composite structure.^[17] The PLGA filled struts were further coated with a bioactive glass (33 wt%)–PLGA composite coating. The resulting scaffolds proved to be bioactive and exhibited relatively high compressive strength values (up to 7.7 MPa) and compressive moduli (up to 3 GPa), these values being comparable to those of natural spongy bone.

CONCLUSIONS

Numerous synthetic materials have been demonstrated to be cell-supportive, bioactive, and biodegradable. Various material processing techniques exist that can produce highly porous structures for convenient use as tissue engineering scaffolds. Porosity ($> 90\%$), pore size ($> 300 \mu\text{m}$), and high pore interconnectivity are critical design features for enabling 3-D tissue growth and sufficient vascularization. One of the major challenges encountered in hard tissue engineering is the design and fabrication of highly porous, bioactive, and biodegradable scaffolds suitable for load-bearing sites.

Although it might not be necessary to have a starting scaffold with a mechanical strength equal to that of bone because cells and extracellular matrix secreted by cells in vitro or in vivo can increase the scaffold strength, a scaffold should have an appropriate fracture strength and structural integrity for enabling its handling, as well as possessing the required biodegradability. With respect to this design concept, a suitable processing method (such as the replication technique) combined with biomimetic or coating techniques to tailor the scaffold surfaces (chemistry and topography) can be the technologies of choice for future developments in the field of bone tissue engineering.

ARTICLES OF FURTHER INTEREST

Bioactive Glass; Biocompatibility Testing; Biocomposites; Biomimetic Materials; Tissue Engineering

REFERENCES

1. Lanza, R.P.; Langer, R.; Vacanti, J. *Principles of Tissue Engineering*, 2nd Ed.; Academic Press: San Diego, CA, 2000.
2. Karageorgiou, V.; Kaplan, D. Porosity of 3D biomaterial scaffolds and osteogenesis. *Biomaterials* **2005**, *26*, 5474–5491.
3. Hench, L.L.; Wilson, J. Surface-active biomaterials. *Science* **1984**, *226*, 630–636.
4. Rezwan, K.; Chen, Q.Z.; Blaker, J.J.; Boccaccini, A.R. Biodegradable and bioactive porous poly/inorganic composite scaffolds for bone tissue engineering. *Biomaterials* **2006**, *27*, 3413–3431.
5. Day, R.M.; Boccaccini, A.R.; Shurey, S.; Roether, J.A.; Forbes, A.; Hench, L.L.; Gabe, S.M. Assessment of polyglycolic acid mesh and bioactive glass for soft-tissue engineering scaffolds. *Biomaterials* **2005**, *25*, 5857–5866.
6. Gatti, A.M.; Valdre, G.; Andersson, O.H. Analysis of the in vivo reactions of a bioactive glass in soft and hard tissue. *Biomaterials* **1994**, *15*, 208–212.

7. Xynos, I.D.; Edgar, A.J.; Buttery, L.D.K.; Hench, L.L.; Polak, J.M. Ionic dissolution products of bioactive glass increase proliferation of human osteoblasts and induce insulin-like growth factor II mRNA expression and protein synthesis. *Biochem. Biophys. Res. Commun.* **2000**, *276*, 461–465.
8. Chen, Q.Z.; Thompson, I.D.; Boccaccini, A.R. 45S5 Bioglass[®]-derived glass-ceramic scaffolds for bone tissue engineering. *Biomaterials* **2006**, *27*, 2414–2425.
9. Hench, L.L.; Wilson, J. *An Introduction to Bioceramics*; World Scientific: Singapore, 1993.
10. Temenoff, J.S.; Lu, L.; Mikos, A.G. Bone tissue engineering using synthetic biodegradable polymer scaffolds. In *Bone Engineering*; Davies, J.E., Ed.; EM Squared: Toronto, 2000; 455–462.
11. Devin, J.E.; Attawia, M.A.; Laurencin, C.T. Three-dimensional degradable porous polymer–ceramic matrices for use in bone repair. *J. Biomater. Sci. Polym. Ed.* **1996**, *7*, 661–669.
12. Yang, X.B.B.; Webb, D.; Blaker, J.; Boccaccini, A.R.; Maquet, V.; Cooper, C.; Oreffo, R.O. C. Evaluation of human bone marrow stromal cell growth on biodegradable polymer/Bioglass[®] composites. *Biochem. Biophys. Res. Commun.* **2006**, *342* (4), 1098–1107.
13. Callcut, S.; Knowles, J.C. Correlation between structure and compressive strength in a reticulate glass-reinforced hydroxyapatite foam. *J. Mater. Sci. Mater. Med.* **2002**, *13*, 485–489.
14. Vitale-Brovarone, C.; Miola, M.; Balagna, C.; Verne, E. 3D Glass-ceramic scaffolds with antibacterial properties for bone grafting. *Chemical Eng. Journal* **2008**, *137*, 129–136.
15. Vitale-Brovarone, C.; Verné, E.; Robiglio, L.; Appendino, P.; Bassi, F.; Martinasso, G.; Muzio, G.; Canuto, R. Development of glass–ceramic scaffolds for bone tissue engineering: Characterisation, proliferation of human osteoblasts and nodule formation. *Acta Biomaterialia* **2007**, *3*, 199–208.
16. Wu, C.; Ramaswamy, Y.; Boughton, P.; Zreiqat, H. Improvement of mechanical and biological properties of porous CaSiO₃ scaffolds by poly(D,L-lactic acid) modification. *Acta Biomaterialia* **2008**, *4*, 343–353.
17. Miao, X.; Tan, L.P.; Tan, L.S.; Huang, X. Porous calcium phosphate ceramics modified with PLGA-Bioactive glass. *Mater Sci and Eng C* **2007**, *27*, 274–280.
18. Jones, J.R.; Ahir, S.; Hench, L.L. Large-scale production of 3D bioactive glass macroporous scaffolds for tissue engineering. *J. Sol-Gel Sci. Tech.* **2004**, *29*, 179–188.
19. Leong, K.F.; Cheah, C.M.; Chua, C.K. Solid freeform fabrication of three-dimensional scaffolds for engineering replacement tissues and organs. *Biomaterials* **2003**, *24*, 2363–2378.
20. Chen, Q.Z.; Boccaccini, A.R. Poly(D,L-lactic acid) coated 45S5 Bioglass[®]-based scaffolds: processing and characterisation. *J. Biomed. Mater. Res. A* **2006**, *77A*, 445–457.
21. Hutmacher, D.W. Scaffolds in tissue engineering bone and cartilage. *Biomaterials* **2000**, *21*, 2529–2543.
22. Guarino, V.; Causa, F.; Ambrosio, L. Bioactive scaffolds for bone and ligament tissue. *Expert Rev. Medical Devices* **2007**, *4*, 405–418.
23. Khan, Y.M.; Katti, D.S.; Laurencin, C.T. Novel polymer-synthesised ceramic composite-based system for bone repair: an in vitro evaluation. *J. Biomed. Mater. Res.* **2004**, *69A*, 728–737.
24. Roether, J.A.; Boccaccini, A.R.; Hench, L.L.; Maquet, V.; Gautier, S.; Jerome, R. Development and in vitro characterisation of novel bioresorbable and bioactive composite materials based on polylactide foams and Bioglass[®] for tissue engineering applications. *Biomaterials* **2002**, *23*, 3871–3878.
25. Oliveira, A.L.; Mano, J.F.; Reis, R.L. Nature-inspired calcium phosphate coatings: present status and novel advances in the science of mimicry. *Current Opin. Solid State Mater. Sci.* **2003**, *7*, 309–318.

Bioadhesion

John D. Smart

*Biomaterials and Drug Delivery Group, University of Portsmouth, Portsmouth, and
The School of Pharmacy and Biomolecular Sciences, University of Brighton, Brighton,
United Kingdom*

INTRODUCTION

Bioadhesion may be defined as the state in which two materials, at least one of which is biological in nature, are held together for extended periods of time by interfacial forces. In the context of their medical and pharmaceutical use, the term bioadhesion refers to the adhesion of synthetic and biological macromolecules to a biological tissue. The biological substrate may be cells, bone, dentine, or the mucus coating the surface of a tissue. If adhesive attachment is to a mucus coating, the phenomenon is sometimes referred to as mucoadhesion.

Many examples of bioadhesion exist in nature, including such diverse events as cell-to-cell adhesion within a living tissue, barnacles binding to rocks, and bacteria binding to tooth enamel. In health care, bioadhesives were first used as wound dressings, skin adhesives, and denture fixatives. Over the last two decades, bioadhesives have been of interest within the pharmaceutical sciences for their potential to optimize drug delivery. Such drug delivery may be optimized at the site of action (e.g., on the cornea or within the oral cavity) or at the absorption site (e.g., in the small intestine or nasal cavity). Bioadhesives may also be used as therapeutic agents in their own right, to coat and protect damaged tissues (gastric ulcers or lesions of the oral mucosa) or to act as lubricating agents (in the oral cavity, eye, and vagina). Skin adhesives, tissue sealants, and dental and bone adhesives and cements are also defined as bioadhesives.

This article first focuses on the types of muco/bioadhesives currently used in the pharmaceutical sciences, from first-generation hydrophilic polymers to second-generation polymers and lectins. The nature of bioadhesive interactions, types of bioadhesive formulations developed, and regions of the human body to which they may be administered are also considered.

Other types of medical bioadhesives, such as those used in wound management, surgery, and dentistry, are also discussed.

PHARMACEUTICAL BIOADHESIVES

First-Generation Bioadhesives

These are the off-the-shelf and most widely investigated bioadhesives. They are hydrophilic macromolecules containing numerous hydrogen-bond-forming groups.^[1–5] Their first use as bioadhesives was in denture fixative powders or pastes.^[2] The presence of carboxyl or amine groups in the molecules appears to particularly favor adhesion. First-generation bioadhesives are called wet adhesives because they are activated by moistening and will adhere nonspecifically to most surfaces. Once activated, they show stronger adhesion to dry inert surfaces than to those covered with mucus. Unless water uptake is restricted, they may overhydrate to form a slippery mucilage. Like typical hydrocolloid glues, they can form very strong adhesive bonds if the formed adhesive joint is allowed to dry.

The poly(acrylic acid)s/carbomers are among the most important bioadhesive molecules. These are cross-linked acrylic acid-based polymers such as Carbopol 934 P, polycarbophil, and Carbopol 974 (Fig. 1a). Their molecular weight has been estimated to be in the range of $1\text{--}4 \times 10^6$ Da, but may be much higher. These materials form gels at relatively low concentrations (0.3–1%) in neutral pH solutions, and may be better described as water dispersible than as water soluble. They are also used as gelling agents, thickeners, tablet excipients, and emulsifiers.

Chitosan, another bioadhesive of great interest, is obtained by deacetylation of the natural product chitin, a polysaccharide extracted from the exoskeleton of Crustacea (e.g., crabs and shrimps). It is a linear polymer containing glucosamine and N-acetyl glucosamine units, linked by glycosidic bonds (Fig. 1b). Chitosan is cationic, and is water soluble only at lower pHs. It is available in a range of molecular weights (typically $50 \times 10^3\text{--}2000 \times 10^3$ Da) and degrees of acetylation. Although proposed for use as tablet excipient, drug carrier and controlled-release material, chitosin's

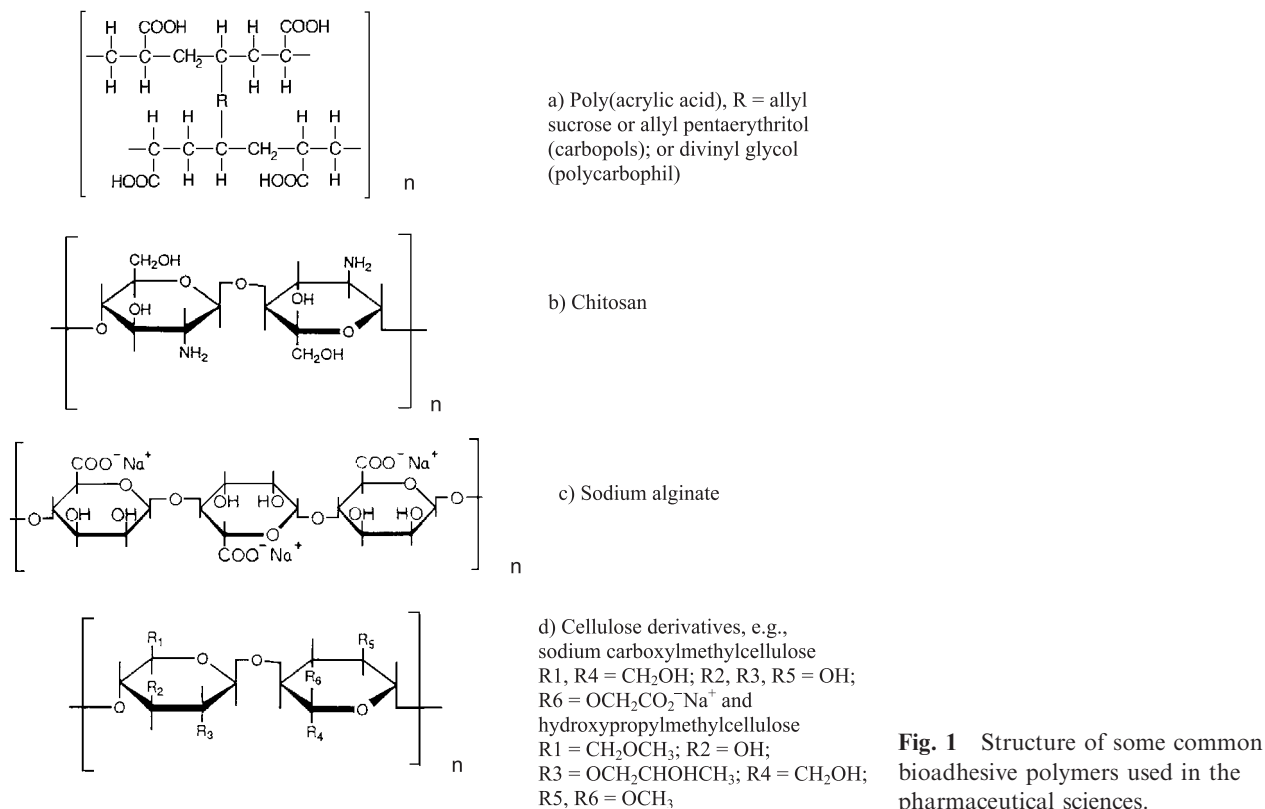


Fig. 1 Structure of some common bioadhesive polymers used in the pharmaceutical sciences.

use in pharmaceutical products has been limited. It has been widely marketed as a putative antiobesity product.

Alginates are polysaccharides extracted from marine algae or produced by bacteria, and consist of units of manuronic and guluronic acid (Fig. 1c). Three distinct regions are possible within an alginate molecule: polymannuronic acid segments (M blocks), poly(guluronic acid) segments (G blocks), and segments of alternating or randomly distributed mannuronic acid and guluronic acid units (MG blocks). Alginates' various pharmaceutical grades contain differing proportions of these blocks, and/or different molecular weight distributions. Alginates are used as thickeners, suspending agents, and tablet excipients. The specific gelling property of alginates with calcium ions is utilized in the production of a viscous gel (a raft) in the stomach to treat gastroesophageal reflux disease.

Cellulose derivatives are another important class of polysaccharide bioadhesives (Fig. 1d).

Some other bioadhesive polymers are listed in Table 1.

The Bioadhesive Bond

Mucous membranes line the walls of various body cavities (e.g., the gastrointestinal and respiratory tracts and the external surface of the eye). These are either simple single-layered epithelia (e.g., the stomach, small and large intestine, and bronchi) or multilayered

stratified epithelia (e.g., in the esophagus, vagina, and cornea). The former contain goblet cells that secrete mucus directly onto the epithelial surfaces; the latter contain or are adjacent to tissues containing specialized glands such as salivary glands that secrete mucus onto the epithelial surface. Mucus is present either as a gel layer adherent to the mucosal surface or in a luminal soluble or suspended form. The thickness of the attached mucus layer varies from 50–450 μm in the stomach to 0.7 μm in the oral cavity. The major functions of mucus are to protect and lubricate; mucus can be said to act as an antiadherent.

The mechanism of bioadhesion to mucosae is still not fully understood, despite many studies. Bioadhesion may arise with both dry (or partially hydrated) dosage forms (tablets or particles) and with hydrated dosage forms (i.e., gels and solutions, or fully swollen formulations such as particles administered into the gastrointestinal tract).^[6] The process by which the bioadhesive bond forms is likely to differ in each case.

For dry or partially hydrated dosage forms, two basic steps have been identified in the process of bioadhesion. Step one is the contact stage, in which intimate contact is formed between the mucoadhesive and mucous membrane; step two is the consolidation stage in which various physicochemical interactions occur to consolidate and strengthen the adhesive joint, leading to prolonged adhesion. Mucus gel dehydration by the bioadhesive polymer and macromolecular interpenetration have been

Table 1 Some bioadhesive polymers

Material	Composition	Solubility	Other uses
<i>Polysaccharides</i>			
Acacia	Calcium salt of arabic acid	Slowly water soluble	Suspending/emulsifying agent, tablet excipient
Carboxymethyl-cellulose (carmellose)	Carboxylated polyglucose ^a	Water dispersible	Suspending/emulsifying agent, tablet excipient, denture fixative
Dextran	A polyglucose	Water soluble	Plasma expander
Hyaluronic acid	N-acetyl glucosamine and glucuronic acid	Water soluble	Eye drops
Hydroxypropyl-cellulose	A polyglucose containing hydroxypropyl-groups	Soluble in cold water some alcohols	Tablet coating, tablet excipient, viscosity enhancer
Hydroxypropyl-methylcellulose (hypromellose)	Polyglucose containing hydroxypropyl- and methoxy-groups ^a	Soluble in cold water	Viscosity enhancer, tablet excipient (sustained release), emulsifier, stabilizer
Karaya gum (sterculia, Indian tragacanth)	Mixture of sugar residues	Sparingly water soluble	Denture fixative powders, ostomy adhesive, suspending agent
Methylcellulose	Polyglucose containing methoxy-groups	Soluble in cold water some alcohols	Viscosity enhancer, tablet excipient (sustained release), emulsifier, stabilizer
Pectin	Polygalacturonic acid containing methoxy-groups	Soluble in water	Viscosity enhancing/suspending agent
Starch	A polyglucose	Hot water soluble	Absorbant (dry) and thickener
Tragacanth	Mixture of sugar residues	Partially water soluble	Suspending/emulsifying agent
<i>Proteins</i>			
Gelatin	Contains hydroxyproline and other amino acids	Soluble in hot water	Capsule shells, gels, stoma and lesion products
<i>Synthetic polymers</i>			
Maleic anhydride co-polymers (Gantrez [®])	Synthetic polymers of methylvinylether and maleic anhydride	Water soluble	Toothpastes, ostomy products, denture fixatives
Polyethylene oxide	Poly oxyethylene	Miscible with water, soluble in some polar organic solvents	Hydrogel wound dressings, viscosity enhancer, tablet excipient, denture fixatives
Polyvinylalcohol	Vinyl alcohol polymer	Water soluble	Film-forming gel, viscosity enhancer, lubricant
Polyvinyl-pyrrolidone	Polymer of vinyl pyrrolidone	Water and alcohol soluble	Tablet excipient, suspending and dispersing agent

^aSee Fig. 1(d).

proposed as two ways in which strengthening of the mucus layer occurs.^[6] Adhesive joint failure will normally occur within the weakest region of the bioadhesive joint (Fig. 2). For weak adhesives this would be at the adhesive interface; for weakly gelling materials this would be at the hydrating layer. For stronger adhesives the weakest region would initially be the mucus layer, but later would be in the hydrating adhesive material.^[7]

The strength of the adhesive joint therefore depends on the cohesive nature of the weakest region.

Several factors have been identified as affecting the strength of the solid bioadhesive joint.^[4,5] Many studies have indicated an optimum molecular weight for bioadhesion, ranging from circa 10^4 Da to circa 4×10^6 Da, although characterizing the molecular weight of large hydrophilic polymers is very difficult. The flexibility of

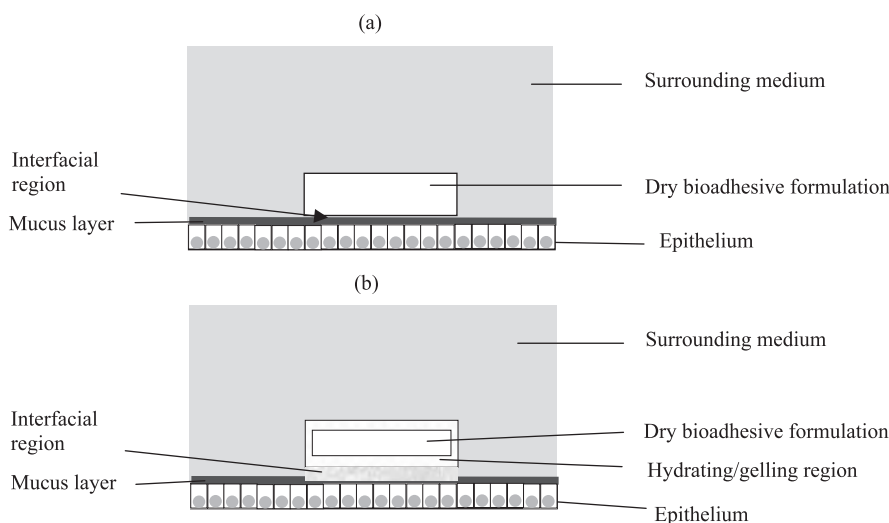


Fig. 2 Regions within a bioadhesive joint upon application of dry formulation to moist mucous membrane. (a) Upon application; (b) after a period of contact.

polymer chains is believed to be important for interpenetration and entanglement. As the cross-linking of water-soluble polymers increases, the mobility of the polymer chains decreases. Bioadhesive polymers that are placed in aqueous media hydrate, swell, and form a gel. The strength of adhesion of such polymers has been shown to reach a maximum at an optimum degree of hydration. Studies have shown that the bioadhesive properties of polymers containing ionizable groups are affected by the pH of the surrounding media. For example, bioadhesion of poly(acrylic acids) is favored when the majority of the carboxylate groups are in the unionized form, which occurs at pHs below the pKa.

The strength of adhesion has been found to change with the initial consolidation force applied to the joint, and with the length of contact time prior to testing. Other factors that may affect bioadhesion include mucin turnover, temperature, formulation, metal ions, and polymer concentration.

Liquid and semisolid bioadhesion typically consist of a viscous liquid or gel clinging to and being retained on a mucous membrane such as the esophagus or vagina. The factors influencing bioadhesion and retention are:

- The thermodynamic affinity of the semisolid for the mucous membrane—in terms of surface energies, contact angles, and spreading coefficient
- Rheology—where a liquid with low viscosity would readily run off the surface, whereas a rigid gel would be less able to spread on and interact with the mucous membrane
- The stability of the gel in the biological system

For a semisolid formulation, characterizing bioadhesion is problematic because contact angle measurement (liquids) and adhesive strength measurements (solids) are inappropriate, resulting in most studies' consideration of the retentive properties of these systems in in-use conditions.

The term bioadhesion can also be used to describe the adsorption of polymers from solution onto mucous membranes, mimicking biopolymers like salivary and ocular mucins. Bioadhesive polymers (such as carbomers and chitosan) have been shown to adsorb onto buccal mucosal cells (and the stomach mucosa) from solution.^[8] The resulting reduction of interfacial free energies and formation of cooperative bonds will facilitate this polymer adsorption process.

Multifunctional Bioadhesives

In many formulations, bioadhesive polymers perform more than one role, acting not only as an adhesive, but also providing a gel matrix that regulates drug release. A more sophisticated development utilizes polymer phase change (liquid–solid) as a response to external conditions such as pH, ionic strength, and temperature as means to closely regulate drug release.^[4] Carbomers have been shown to inhibit proteases and facilitate absorption,^[9] whereas chitosan also has antimicrobial and cell proliferation activity.

Second-Generation Bioadhesives

To overcome the limitations of first-generation off-the-shelf bioadhesive materials, new types of materials have been investigated that allow specificity or prolong and strengthen the bioadhesion process.

One approach to producing improved bioadhesives has been to modify existing materials. For example, thiol groups (by coupling cysteine) have been placed into a range of bioadhesive polymers including the carbomers.^[10] The concept is that in situ they will form disulfide links between not only the polymers themselves, but also with the mucin layer, thus leading to improved adhesive performance. The incorporation of poly(ethylene glycol) into polyacrylic/polymethacrylic acid copolymers

Table 2 Some lectins studied in bioadhesive drug delivery

Source lectin	Type	Sugar specificity
<i>Lycopersicon esculentum</i> (tomato)	Glycoprotein	N-acetyl glucosamine
<i>Phaseolus vulgaris</i> (red kidney bean)	Glycoprotein	Galactose, N-acetyl-galactosamine
<i>Solanum tuberosum</i> (potato)	Glycoprotein	N-acetyl-glucosamine
<i>Triticum vulgaris</i> (wheat germ)	Glycoprotein	N-acetyl-glucosamine, N-acetylneuraminic acid
<i>Ulex europaeus 1</i> (gorse)	Glycoprotein	L-fucose
<i>Helix pomatia</i> (edible snail)	Protein	N-acetyl galactosamine
Bacterial adhesins	Proteins	Mannose
Bacterial invasins	Proteins	β_1 -integrin

has also been investigated.^[11] These comb copolymer films were shown to have favorable adhesion relative to poly(acrylic acid) alone. EDTA has been added to chitosan to enhance the enzyme inhibitory function.^[12]

Specific bioadhesives allow some targeting in the bioadhesion process by binding selectively to the required ligand. The most important members of this group are the lectins, which have the ability to target a dosage form and facilitate absorption.^[13–16] Lectins are defined as proteins or glycoproteins of natural origin capable of specific recognition of, and reversible binding to, carbohydrate moieties of complex glycoconjugates without altering the covalent structure of any of the recognized glycosyl ligands. Lectins of interest usually contain two or more sugar-binding sites; some of the more important in terms of bioadhesion potential are shown in Table 2.

Another interesting approach is to produce polymers containing the appropriate hapten sugars that bind to endogenous lectins on the mucosal surface.^[17]

Formulation Factors

First-generation bioadhesive polymers—as typified by the carbomers—present many challenges to the formulation scientist, often because of the properties that make them adhesive. They are hydrophilic, showing limited solubility in other solvents, and in solution they typically form highly viscous, pH-sensitive solutions at low concentrations. Dry polymers become adhesive on exposure to moisture, and so readily cohere, or adhere to manufacturing equipment or delivery devices.

Tablet formulations usually consist of a matrix containing a bioadhesive polymer into which the drug is distributed. In order to separate the adhesion and controlled-release functions, multilayer preparations have been produced (e.g., with an adhesive layer, a controlled-release layer, and a backing layer to facilitate application or direct drug delivery) (Fig. 3).^[18]

Bioadhesive polymers can be cast from solutions for use as films or patches. Although most of these materials will form good films, building up a substantial

layer may be limited by solution viscosity. Because these films are usually relatively thin, they will rapidly overhydrate and lose their adhesive properties unless access of water is restricted. A bioadhesive patch may be multilayered for the reasons described for tablet preparation.^[19] However, given that the bioadhesive layer will usually swell in contact with moisture, curling of the patch can occur unless it is carefully formulated.

Bioadhesive microparticles or capsules have been produced by a spray drying or oil-in-water emulsification procedure, or during the polymerization process. Lectins—especially plant lectins—are relatively robust protein/glycoproteins, and can be purchased already coupled to a range of markers including colloidal gold, while still maintaining their activity. Using processes like the carbodiimide coupling reaction, they can therefore be readily placed onto the surface of a formed microparticle while maintaining their lectin-binding activity.^[13,16]

Bioadhesive ointments and pastes consist of powdered bioadhesive polymers incorporated into a hydrophobic base. Because the base is anhydrous, the adhesive properties of the bioadhesive polymers are not activated, and it is a relatively simple process to manufacture.

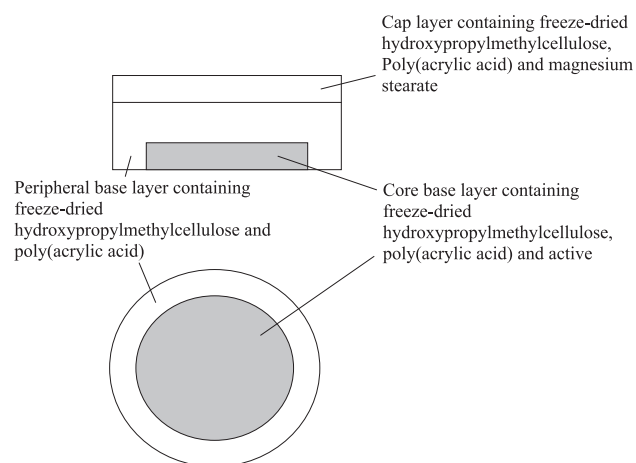


Fig. 3 Example of multilayered tablet formulation. (From Ref. [18].)

In use, the base allows restricted access of water from the saliva, and can therefore prolong the bioadhesive properties of the polymer. An oral paste (e.g., Orabase[®]) was one of the earliest bioadhesive formulations to be described.^[2] This consists of a mix of finely ground pectin, gelatin, and sodium carboxymethylcellulose dispersed in a polyethylene and mineral oil base, which has been used for local application in the oral cavity of steroids for the treatment of mucosal ulceration.

Although materials like the carbomers and cellulose derivatives have been used in aqueous gels in the pharmaceutical and cosmetic industries, these have not normally been considered to be bioadhesive (except perhaps for a polycarbophil-containing vaginal gel formulation).^[4] The content of bioadhesive polymer is normally relatively low (e.g., 0.5–1% for carbomers).

Solutions tend to be viscous and raise issues of polymer dissolution and stability; other proposed bioadhesive formulations include vaginal rods, pessaries, and suppositories.^[5]

Routes of Administration

The oral cavity is a site for local and systemic drug delivery. Local therapy treats conditions such as gingivitis and oral lesions; systemic therapy avoids first-pass metabolism and hostile conditions elsewhere in the gastrointestinal tract.^[20] Considerations associated with bioadhesive drug therapy within the oral cavity include rapid elimination of drugs through the flushing action of saliva, nonuniform distribution of drugs within saliva, organoleptic properties (taste), the relative impermeability of buccal mucosa, and the challenges of eating and drinking.

Suscard Buccal[®] (containing glyceryl trinitrate in a hydroxypropylmethylcellulose matrix) and Buccastem[®] (containing prochlorperazine in a ceratonia/xantham gum base) are two commercial examples of bioadhesive tablets for application to the oral cavity.^[20] Aphtach[®] (a thin double-layered tablet with steroid drug incorporated into a hydroxypropylmethylcellulose and poly(acrylic acid) base, whose upper layer is coated with lactose facilitate application via a nonstick surface) is an example of a multilayered tablet formulation. Adhesive patches have been developed as an alternative to tablets, and typically contain a film of bioadhesive polymer. Patches administered to the buccal mucosa measure up to 12 cm² in size (but typically are 1–3 cm²), and are ellipsoid in shape. Good flexibility of patches is required to ensure good adhesion and to prevent local discomfort.

Oral pastes (e.g., Orabase[®]) and ointments consisting of a bioadhesive polymer (such as carbomer or carmellose) in a mineral oil base have been used for local application to the oral cavity (e.g., to treat mucosal ulceration). After toothpastes, mouthwashes are one of the most widely used preparations for the

oral cavity. Bioadhesive polymers (chitosan, polycarbophil, and Carbopol 934) in dilute solutions have been shown to bind to buccal cells *in vitro*, and to bind and be retained *in vivo* for over 2 hours.^[21] Lectins in solution have also been shown to bind to the cells lining the oral cavity.

The peroral route (swallowing the dosage form) is generally the most favored for drug delivery, due to its acceptability to patients. However, it provides a range of challenges, and to date only limited success in developing bioadhesive formulations has been achieved.^[3–5,16] Challenges to the oral route of delivery include the mucus barrier (and its high turnover rate), the variable environment (particularly in terms of pH and luminal contents), the hostility of the environment (enzymes, pH, colonic microflora), variable and mainly uncontrollable transit rates, and intersubject variability. In addition, the bioadhesive must be delivered to its site of action, because it is rarely possible to locate it directly on its desired attachment site.

The most clinically significant example of bioadhesion in the gastrointestinal tract remains that of dry tablet formulations to the esophagus. Poor techniques of administration can lead to delayed onset of action and esophageal ulceration.^[22] Tablets and capsules are generally considered to be too large, with transit rates that are too variable to be used as gastrointestinal bioadhesives. A more promising approach would employ particulates, believed to distribute more evenly throughout the gastrointestinal tract. Particulates can be delivered in a suspension, or be released in the stomach from a hard gelatin capsule formulation. Much development activity is ongoing in this area, and several claims to enhanced drug delivery by bioadhesive polymer have been made (even on flimsy evidence). It has been shown that polyacrylates, when administered in solution, will bind to gastric mucosa *in vitro*.^[23] Particulate systems containing lectins or other locating agents on their surface have also been studied in cell culture and in animals, and represent a potential new means for enhancing drug delivery.^[13,14,16]

An eye drop is typically retained on the ocular surface for minutes only, thus limiting the time available for drug absorption; ointments can blur the vision and be uncomfortable for the patient. Ocular delivery systems must be sterile before application into the eye, and the rigors of the sterilization process can therefore limit which materials can be used in these formulations. The addition of bioadhesive polymers (carbomer, hyaluronic acid) has been seen to increase the bioavailability of drugs from eye drops. This is due in part to the effect of viscosity in reducing the rate of precorneal clearance. Bioadhesive polymers have also been formulated into adhesive ocular inserts, self-gelling systems, and micro-particles, which have been shown to enhance retention and bioavailability.^[4,5]

However, it has the advantages of avoiding first-pass metabolism, being readily accessible, encouraging particle deposition, and having good vasculature and relatively permeable mucosa (to smaller molecules). The nasal route has the disadvantages of possible local toxicity, rapid mucociliary clearance ($4\text{--}6\text{ mm min}^{-1}$), and being subject to pathological conditions that can adversely affect therapy (such as rhinitis). Bioadhesive solutions and particulate systems have been delivered nasally. The use of bioadhesive polymers would not intuitively be expected to confer a significant advantage, as a major physiological function of the nose is to capture and remove particulates from the inhaled air, clearing them using mucociliary clearance mechanisms. However, bioadhesive formulations (e.g., those containing chitosan or modified starch) have been reported to significantly extend retention in the nasal cavity, facilitating enhanced drug absorption.^[24]

Given that physicochemical properties (aerodynamic diameter) are critical for deposition in the lungs, a particle containing a typical hygroscopic, first-generation bioadhesive polymer would present formidable challenges to a formulation scientist (probably explaining the limited work in this area).

The vaginal route would appear to be highly appropriate for bioadhesive drug delivery systems, to retain systems for treating largely local (although some systemic) conditions, or for use in contraception. Traditional formulations have a relatively short retention period. Apart from menstruation or coitus, the major challenge to a bioadhesive device would arise from body movement and gravity effects. Along with rod, patch, and pessary formulations, a bioadhesive gel containing polycarbophil has been used commercially as a vaginal lubricant, and to deliver drugs such as progesterone.^[5]

MEDICAL BIOADHESIVES

Skin Adhesives

Skin is a squamous stratified (layered) epithelium, with a normally relatively dry, hydrophobic surface covered with lipids. Pressure-sensitive adhesives are used to attach a wide range of appliances, dressings, and devices.^[25] These materials differ greatly from the bioadhesives described previously in that they do not require water to become activated and therefore must possess inherent stickiness (tack). Skin adhesives must be sufficiently plastic and hydrophobic to interact with and adhere to the skin surface on application. They must also be nonirritating, remain adhesive for prolonged periods (typically days), and be removable without damaging the skin or leaving excessive residue. Rubber (polyisoprene) was the first commercial skin adhesive. Today the three main classes of adhesives are based on

polyisobutylenes (or other synthetic rubbers), acrylate polymers (such as 2-ethylhexyl-, butyl-, and octyl-acrylate copolymers), or silicone polymers (such as polydimethyl or poly(dimethyldiphenyl silicone)).

Hydrocolloid pressure-sensitive adhesives contain a combination of the water soluble polymers similar to those used for mucoadhesive drug delivery and the pressure-sensitive adhesives previously described.^[26] They are used in wound dressings because they have both adhesive and absorbent properties. A typical early formulation would contain a continuous phase of a low molecular weight polyisobutylene (circa 40%) into which would be dispersed combinations of hydrocolloids (such as pectin, sodium carboxymethyl cellulose, and gelatine). Later versions use a continuous phase of thermoplastic copolymers (such as styrene-isoprene-styrene) into which the hydrocolloid is dispersed, and may also contain a range of other materials such as tackifiers and antioxidants, cross-linking agents, and mineral oils (e.g., Duoderm[®] or Sigma Dress[®]).

Tissue Sealants

In recent years, adhesives for tissue repair (used in surgery as an alternative to staples, sutures, or surgical tapes) have become popular.^[27] Tissue sealants have the advantages of being relatively easily applied; acceptable, biodegradable, and relatively inexpensive; sealing more surely; reducing infection risk; and virtually eliminating scarring. In some cases, however, they may lack mechanical strength.

Fibrin sealants have been compounded for many years by individual surgeons, but are now available commercially (e.g., Tisseel[®]). They are composed primarily of the proteins fibrinogen and thrombin. These are the building blocks of a blood clot, and upon mixing, work by mimicking the final stage of the body's natural clotting mechanism. The formed fibrin clot is both adhesive and hemostatic, so is therefore particularly applicable to sealing bleeding wounds. The application procedure can be a little complex. Some variants are only activated by an external source (such as light), thus restricting the sealant to being activated only at the site of application.

Cyanoacrylates have been used as tissue adhesives in surgical procedures (e.g., when repairing corneal wounds). They essentially consist of cyanoacrylate monomer that, upon contact with trace amounts of water or bases on a biological surface, initiates polymerization and binds the surfaces together (Fig. 4). The original butyl cyanoacrylate superglues have now been replaced by octyl cyanoacrylates (e.g., Dermabond[®]) that are mechanically stronger and more flexible, and allow slower degradation and release of the potentially inflammatory by-products.

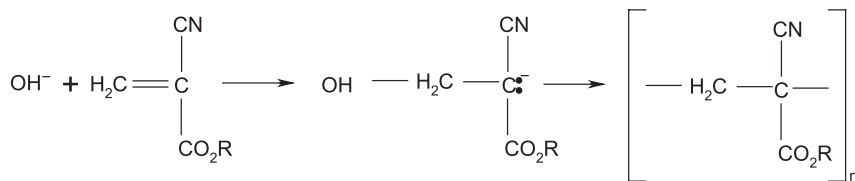


Fig. 4 Polymerization reaction of cyanoacrylate superglues (where R = butyl or octyl groups).

The protein albumin, when cross-linked with gluteraldehyde, forms a strong adhesive for internal surgery that is capable of withstanding high pressures (e.g., BioGlue[®]). The protein collagen can also be used as the matrix to which other materials such as fibrin, thrombin, or polyethylene glycols can attach. Derivatives of polyethylene oxide can be used as swellable hydrogels that rapidly arrest bleeding (e.g., Co-seal[®], Co-stasis[®]).

Dental Adhesives and Cements

Dental adhesives and cements adhere to the dental enamel and dentine of teeth in order to secure orthodontic devices. A range of materials is available, whose choice depends on the properties required for the procedure to be undertaken.^[28]

Zinc phosphate cement (e.g., Tenacin[®]) is based on the neutralizing reaction of zinc oxide with phosphoric acid in an aqueous environment to form solid zinc phosphate, a cement with high compressive strength. Zinc polycarboxylate cement (e.g., Durelon[®]) is the reaction product of zinc oxide with aqueous polyacrylic acid, which forms a cement that adheres to the calcium present in dentine and enamel. Zinc polycarboxylate is considered biocompatible but its mechanical properties limit its use. Glass ionomer cements (e.g., Fuji[®]) consist of polyacrylic or polymaleic acid mixed with an ion-leachable aluminosilicate glass that react together in water to form a cement that will adhere to enamel or dentine. These can be modified to allow a slow release of fluoride and also (via hydrogel domains) allow the release of calcium and other ions associated with tooth remineralization. Drying and low fracture resistance, however, restrict their use. Resin-modified ionomers (e.g., Avance[®]) are similar to the glass ionomers but also contain 10–20% resin monomers that can be polymerized by the application of light or chemical activators to give a more stable and physically stronger cement.

Resin adhesives (e.g., Panavia[®]) consist of resin monomers (such as the acrylate phosphate esters) that can be activated by light, initiators, or a combination of both. Resin adhesives often contain inert fillers. They cure purely through a polymerization process, so contain little water and are less brittle and more fracture resistant than cements. They can be chemically modified to contain polyacids, while aluminosilicate glass particles can also be incorporated to allow the release of remineralizing ions.

Bone Cements and Bone Graft Substitutes

Bone-filling cements have been used in surgery for the fixation of such prostheses as total hip replacements for over 40 years.^[29] The polymethyl methacrylates (PMMA) (e.g., Simplex P[®]) were the first bone cements used, but these can suffer from the mechanical and biological disadvantages of becoming loosened with time and being neither osteoconductive nor bioresorbable. They are prepared by the polymerization of methylmethacrylate monomer. A typical cement contains monomer, an initiator for polymerization (e.g., benzoyl peroxide), a radio-opaque substance (e.g., barium sulphate), and oligomers of PMMA. The initiator is added to the monomer solution prior to application and the properties of the resulting material can be modified by the addition of materials such as dimethyl toluidine or hydroxyapatite.

Although PMMA is still used to anchor hip and knee joint implants, a new class of biodegradable and biologically compatible cements has emerged, made with calcium phosphate (e.g., Norion SRS[®]), which in the form hydroxyapatite is a key component of human bone. These cements are easily accepted by the body and are sometimes used instead of bone graft surgery. They consist of a powder to which water is added to form a paste that hardens to form a material structurally similar to bone, which material can then be integrated into the new growth.

Other materials with adhesive properties that have been used as bone graft substitutes are similar to the materials used as dental cements.^[30] Bioactive glasses (e.g., NovaBone[®]) consist of sodium oxide, calcium oxide, phosphorous pentoxide, and silicon dioxide in varying proportions, and are solid and nonporous. They form a strong bond with bone as a result of a gel layer containing calcium and phosphate ions generated at the surface on exposure to physiological aqueous solutions forming hydroxyapatite. Glass ionomers (e.g., Fugix IX[®]), calcium sulphate (plaster of Paris), aluminium oxide (e.g., Alumina ceramic), and synthetic hydroxyapatite (Calcite[®]) have also been used as synthetic bone grafts.

Preventing infection is important in orthopedic surgery, so sterilisation and the incorporation of antibiotics are key issues in the use of these materials.

Other Adhesives

Muscle glue protein has also been investigated as a bioadhesive material.

As a result of their high surface-to-volume ratio, most nanoparticle systems would be expected to readily adsorb to surfaces, and have been reported as putative bioadhesives.^[16] A range of materials, including polyalkylcyanoacrylates and polymethacrylate copolymers, has been used to produce these nanoparticles. However, the low payload (dose of active agent that can be delivered) and the ease with which these may be flocculated by endogenous biopolymers can be factors limiting the use of these other adhesives.

CONCLUSION

For pharmaceutical bioadhesives, a range of new therapeutically active materials has resulted from advances in biotechnology. Despite the limitations of the first-generation, off-the-shelf bioadhesive polymers highlighted in this review, they have been used in several commercial preparations that are applied directly to the target surface. It should be a priority to design the second generation of polymers that are targetable, have predictable adhesive performance, and can be more readily accommodated in pharmaceutical formulations. To facilitate this, a clear understanding of the bioadhesive process is required.

Significant issues relative to second-generation, site-specific bioadhesives (lectins) remain to be resolved, such as access to their target sites and their toxicity profiles. The significant advances in bioadhesives with a range of medical applications (particularly those in the tissue sealants area) will inevitably result in their significantly expanded use.

ARTICLES OF FURTHER INTEREST

Biological Adhesives from Nature; Bone Cement; Chitosan; Fibrin Sealants; Glues; Surface Coatings

REFERENCES

- Smart, J.D.; Kellaway, I.W.; Worthington, H.E.C. An in-vitro investigation of mucosa-adhesive materials for use in controlled drug delivery. *J. Pharm. Pharmacol.* **1984**, *36*, 295–299.
- Chen, J.L.; Cyr, G.N. Compositions Producing Adhesion Through Hydration. In *Adhesion in Biological Systems*; Manly, R.S., Ed.; Academic Press: New York, 1970; 163–181.
- Harding, S.E.; Davis, S.S.; Deacon, M.P.; Fiebrig, I. Biopolymer mucoadhesives. *Biotechnol. Genet. Eng. Rev.* **1999**, *16*, 41–85.
- Lee, J.W.; Park, J.H.; Robinson, J.R. Bioadhesive-based dosage forms: The next generation. *J. Pharm. Sci.* **2000**, *89*, 850–866.
- Ahuja, A.; Khar, R.P.; Ali, J. Mucoadhesives drug delivery systems. *Drug Dev. Ind. Pharm.* **1997**, *23* (5), 489–515.
- Smart, J.D. The Role of Water Movement and Polymer Hydration in Mucoadhesion. In *Bioadhesive Drug Delivery Systems: Fundamentals, Novel Approaches and Development*; Mathiowitz, E., Chickering, D.E., Lehr, C.-M., Eds.; Marcel Decker: New York, 1999; 11–23.
- Mortazavi, S.A.; Smart, J.D. An in-vitro method for assessing the duration of mucoadhesion. *J. Control. Release* **1994**, *31*, 207–212.
- Patel, D.; Smith, A.W.; Grist, N.W.; Barnett, P.; Smart, J.D. In-vitro mucosal model predictive of bioadhesive agents in the oral cavity. *J. Control. Release* **1999**, *61*, 175–183.
- Lueßen, H.L.; Rental, C.-O.; Kotze, A.F.; Lehr, C.-M.; de Boer, A.G.; Verhoef, J.C.; Junginger, H.E. Mucoadhesive polymers in peroral peptide delivery. IV. Polycarboxophil and chitosan are potent enhancers of peptide transport across intestinal mucosae in vitro. *J. Control. Release* **1997**, *45*, 15–23.
- Bernkop-Schnurch, A.; Scholler, S.; Biebel, R.G. Development of controlled release systems based on thiolated polymers. *J. Control. Release* **2000**, *66*, 39–48.
- Shojaei, A.H.; Li, X. Novel PEG Containing Acrylate Copolymers with Improved Mucoadhesive Properties. In *Bioadhesive Drug Delivery Systems: Fundamentals, Novel Approaches and Development*; Mathiowitz, E., Chickering, D.E., Lehr, C.-M., Eds.; Marcel Decker: New York, 1999; 433–458.
- Bernkop-Schnurch, A.; Krajicek, M.E. Mucoadhesive polymers as platforms for peroral peptide delivery and absorption: Synthesis and evaluation of different chitosan-EDTA conjugates. *J. Control. Release* **1998**, *50*, 215–223.
- Lehr, C.-M. Lectin-mediated drug delivery: The second generation of bioadhesives. *J. Control. Release* **2000**, *65*, 19–29.
- Woodley, J. Bioadhesion. New Possibilities for drug administration. *Clin. Pharmacokinet.* **2001**, *40* (2), 77–84.
- Easson, J.H.; Haltner, E.; Lehr, C.-M. Bacterial Invasion Factors and Lectins as Second Generation Bioadhesives. In *Bioadhesive Drug Delivery Systems: Fundamentals, Novel Approaches and Development*; Mathiowitz, E., Chickering, D.E., Lehr, C.-M., Eds.; Marcel Decker: New York, 1999; 409–432.
- Ponchel, G.; Irache, J.-M. Specific and non-specific bioadhesive particulate systems for oral delivery to the gastrointestinal tract. *Adv. Drug Deliv. Rev.* **1998**, *34*, 191–219.
- Rihova, B.; Rathi, R.C.; Kopeckova, P.; Kopecek, J. In-vitro bioadhesion of carbohydrate containing N-(2-hydroxypropyl)methacrylamide copolymers to the GI tract of guinea pigs. *Int. J. Pharm.* **1992**, *87*, 105–116.
- Nagai, T.; Konishi, R. Buccal/gingival drug delivery systems. *J. Control. Release* **1987**, *6*, 353–360.
- Veillard, M.M.; Longer, M.A.; Martens, T.W.; Robinson, J.R. Preliminary studies of oral mucosal delivery of peptide drugs. *J. Control. Release* **1987**, *6*, 123–131.

20. Smart, J.D. Drug delivery using buccal adhesive systems. *Adv. Drug Deliv. Rev.* **1993**, *11*, 253–270.
21. Kockisch, S.; Rees, G.D.; Young, S.A.; Tsibouklis, J.; Smart, J.D. A direct-staining method to evaluate the mucoadhesion of polymers from aqueous dispersion. *J. Control. Release* **2001**, *77*, 1–6.
22. Wilson, C.G. In Vivo Testing of Bioadhesion. In *Bioadhesion—Possibilities and Future Trends*; Gurny, R., Junginger, H.E., Eds.; Wissenschaftliche Verlagsgesellschaft: Stuttgart, 1990; 93–108.
23. Riley, R.G.; Smart, J.D.; Tsibouklis, J.T.; Young, S.A.; Hampson, F.; Davis, J.A.; Kelly, G.; Dettmar, P.W.; Wilber, W.R. An in-vitro model for investigating the gastric mucosal retention of ¹⁴C-labelled poly (acrylic acid) dispersions. *Int. J. Pharm.* **2002**, *236* (1–2), 87–96.
24. Illum, L.; Farraj, N.F.; Davis, S.S.; Johansen, B.R.; O'Hagan, D.T. Investigation of the nasal absorption of biosynthetic human growth hormone in sheep—Use of a bioadhesive microsphere delivery system. *Int. J. Pharm.* **1990**, *63*, 207–211.
25. Horstmann, M.; Muller, W.; Asmussen, B. Principles of Skin Adhesion and Methods for Measuring Adhesion of Transdermal Systems. In *Bioadhesive Drug Delivery Systems: Fundamentals, Novel Approaches and Development*; Mathiowitz, E., Chickering, D.E., Lehr, C.-M., Eds.; Marcel Decker: New York, 1999; 175–196.
26. Lipman, R. Hydrocolloid PSAs: New formulation strategies. *Med. Device Diagn. Ind. Mag.* **1999**, 132–150.
27. Sharma, A.; Kaur, R.; Kumar, S.; Gupta, P.; Pandov, S.; Patnaik, B.; Gupta, A. Fibrin glue versus N-butyl-2-cyanoacrylate in corneal perforations. *Ophthalmology* **2003**, *110*, 291–298.
28. Ewoldsen, N.; Demke, R.S. A review of orthodontic cements and adhesives. *Am. J. Orthod. Dentofac. Orthoped.* **2001**, *120*, 45–48.
29. DiMaio, F.R. The science of bone cement, a historical review. *Orthopedics* **2002**, *25*, 1399–1407.
30. Moore, W.R.; Graves, S.E.; Bain, G.I. Synthetic bone graft substitutes. *Aust. N.Z. J. Surg.* **2001**, *71*, 354–361.

Biocatalysis

Claire F. Komives

Chemical and Materials Engineering, San Jose State University, San Jose, California, U.S.A.

Ruizhen Rachel Chen

Chemical Engineering Department, Virginia Commonwealth University, Richmond, Virginia, U.S.A.

INTRODUCTION

The discovery that all life forms—from single cells to complex, multicellular organisms—require thousands of controlled chemical reactions to function has enabled the harnessing of naturally occurring catalysts for the fine chemicals industry. Biocatalysis is the use of natural or genetically modified enzymes or whole cells for chemical catalysis of organic molecules.^[1–3]

Enzymes are protein catalysts that accelerate chemical reactions inside a living organism by reducing the activation barrier of the transformation. For every step—from nutrient assimilation to the complex synthesis of bioactive natural molecules—an enzyme is involved. The use of enzymes for organic synthesis of chemicals, from industrial solvents to therapeutics, is now common in industry. This article covers the history of biocatalysis technology, and the use of enzymes and whole cells for the production of chemicals and materials, with a particular focus on biomaterials and biomedical applications.

Enzymes possess three types of selectivities that are unparalleled by traditionally employed chemical catalysts. Usually, enzymes can act on only one type of chemical functional group. This capability of enzymes is referred to as chemoselectivity. A substrate molecule that has more than one functional group from the same chemical family can be differentiated by enzymes possessing regioselectivity. Finally, enzymes can display stereoselectivity or enantioselectivity. Given a racemic mixture of a compound, these enzymes will react on only one type of enantiomer, leaving the other unreacted. A prochiral substrate may be transformed into an optically active product through the action of enzymes. The chiral selectivity of an enzyme is particularly important for drug development, because current regulations mandate either enantiomeric purity or comprehensive toxicity testing to demonstrate the lack of any adverse effects from the enantiomeric impurities. Another hallmark of enzyme biocatalysts is the phenomenal rate of accelerations achieved, the highest to date being 10^{23} with orotidine decarboxylase in the pyrimidine biosynthesis pathway.^[4]

Unlike most of the synthetic processes that necessitate high pressure, high temperature, and organic solvents, enzymes are remarkable in that they achieve this exquisite selectivity and high efficiency at very mild conditions, typically in aqueous solution and close to neutral pH, room temperature, and atmospheric pressure. Enzymes are also environmentally friendly catalysts. They are completely biodegradable and pose no risk to the environment. Using mild conditions in enzyme-catalyzed reactions also means lower energy utilization and less emission of greenhouse gases to the environment.^[5]

In recognizing the potential of biocatalysis, many companies (including large pharmaceutical and other chemical companies) have incorporated research groups with expertise in biocatalysis, because it has been shown in many cases that biocatalysis offers low-cost solutions for chemical synthesis.

HISTORY OF BIOCATALYSIS

The path to progress in the area of biocatalysis was not a linear one. Biocatalysis began with the use of whole cells for the production of ethanol, which is still a major product of biocatalytic reactions and dates back to uncharted historical territory. Because yeast produce ethanol under anaerobic conditions, these fermentations could be successful without an engineering understanding of oxygen mass transfer, and the necessary analytical techniques were those of the nose and eyes of a skilled brewer. Ethanol as a solvent was also produced by fermentation.^[6]

A significant milestone for biocatalysis, related to medicinal applications, was the production of penicillin.^[7] The large-scale fermentation process was initiated during WWII by a number of U.S. universities and pharmaceutical companies under the coordination of Albert Elder. In 1943, typical carbon content yields were 0.0001% and titers were about 0.1 mg/ml, presenting formidable challenges to purification, production rates, and to providing the drug at a reasonable cost for the war effort. Strains were improved through mutation, by X-ray and UV light

exposure, and by selection. By 1967 the titers increased 100-fold to 10 mg/ml. Since then, continuous strain improvements and process optimization have led to current titers reported in the public literature of approximately 40–50 mg/ml.^[8] Likewise, the current production host has been optimized from a strain of *Penicillium chrysogenum* that was screened from an overripe cantaloupe. Current strains show both an increase in copy number of the structural genes and a higher level of mRNA transcribed from the genes. Production was also improved by feeding the penicillin side chain, phenylacetate. However this is also one of the principle contributions to the cost of production. In the current semisynthesis of related β -lactam compounds such as ampicillin, penicillin acylase from *E. coli* or other microbial sources is used. In fact, penicillin acylase represents the best-known enzyme used in pharmaceutical industry with annual worldwide production of 6-aminopenicillanic acid (6-APA) reaching 7500 tons.

The use of microbial species to replace a traditional chemical catalyst was reported in a recent review^[9] to have begun at the UpJohn Company with a step in the production of cortisone, used for the treatment of rheumatoid arthritis.^[10] The oxidation of progesterone, fed in an acetone phase to α -Hydroxyprogesterone, was performed in a *Rhizopus nigricans* 240-liter fermentation with a conversion of 90–97%. This reaction is the first step in the bioconversion of progesterone to cortisone.

With the discovery that microbes and enzymes could be used to catalyze chemical conversions, considerable efforts were put into place for screening of microbes from sources all over the world to seek out enzymes with activities for the conversion of interest. Screening is still in use today by many companies that often possess stocks of microbes for such purposes. Indeed, Diversa Corporation has built its enzyme technology business around the screening of enzymes from unusual environments.^[11] Robots were developed to enable an increase in the rate of processing such experiments.

The infusion of recombinant DNA technology has made an important impact on progress in the biocatalysis field. Recombinant DNA technology has enabled a large number of enzymes to be overexpressed in a convenient host, and provided the enzymes in quantities required for large-scale synthesis with drastically reduced price. In addition, recombinant DNA technology allowed enzyme redesign through modification of the encoding gene. Based on available structural information, site-directed mutagenesis studies (or rational enzyme design), by modifying enzymes at positions important for substrate bindings and/or active sites,^[12] led to some success in improving enzymes for stability, turnover, lower Michaelis Constant

(Km), and substrate specificity, etc. More recently, directed evolution has emerged as a powerful tool in the redesign of enzymes for biocatalysis applications. It is expected that directed evolution will have much greater impact in the field than the rational design approach, because it does not require protein structural information, which is not only largely not available but also poorly predicts structural–functional relationships. In applications involving whole-cell catalyzed reactions, metabolic engineering (modification of cellular genetics to redirect nutrient throughput toward the production of a desired product) has continued to be used for the improvement of production host strains.^[13]

Many scientific disciplines have contributed to the current level of application of biocatalysis, and the history of each field is a story in itself. Considerable contributions of biochemical engineering have included reactor improvements for better mass transfer of oxygen for aerobic processes, methods for low-cost downstream processing, and other significant aspects enabling modern biocatalysis. Bioinformatics has brought about the rapid screening of sequence databases for genes with a high degree of homology with the gene of interest. The knowledge of highly conserved regions on a gene identifies portions of the gene and the protein sequences that are essential for a particular biocatalytic function. Were it not for the advent of high-speed computers, fields such as bioinformatics and many aspects of modeling and control would still be in their infancy.

WHOLE-CELL BIOCATALYSIS

Many industries employ whole-cell biocatalysts (mostly microbial cells) for carrying out chemical reactions, including pharmaceutical and biotech companies, bulk chemical manufacture, and wastewater treatment. A recent review discusses progress in the use of whole cells for oxygenation reactions leading to alcohols and lactones, nitrile and epoxide hydrolysis, and other reactions significant for organic synthesis.^[14] There may be many motivations for using whole cells instead of purified enzymes. Because purifying enzymes can require several steps to achieve adequate purity, using the cells in their native form can save time and effort, provided they do not also contain enzymes with undesirable or competing activities. If the enzyme-catalyzed transformation requires expensive cofactors, whole-cell systems are typically chosen because many cells already produce the necessary compounds. Whole-cell systems are generally more stable because the cell walls and membranes lend some protection to the intracellular enzymes.

A plethora of microbial species are used in whole-cell biocatalysis for products such as antibiotics, immunosuppressors, enzyme inhibitors, and many others. Many of these processes utilize wild-type cells that were isolated from their native habitat. The recent development of recombinant whole-cell biocatalysis has opened new horizons in the field, fueled by the prospect that novel, more potent products can be produced by the recombination of genes from different sources. In addition, it is possible that valuable products can be made more cost efficiently by reengineering the metabolic network of the microbial cells. In those applications, an important issue is host selection.

The reaction product is a primary factor in host selection in recombinant whole-cell biocatalysis. Fast-growing microbes with known genetic tools are the species of choice, especially *E. coli*, yeast, and certain species of fungi. The rapid growth rate of these organisms makes them easier to grow and maintain in a sterile state, and reduces total reaction time, not to mention their robustness in large reactors. Beyond personal preference dictating the selection of a host organism, the general rule for the choice of cells is to use *E. coli* whenever possible because their genetics are among the best known, they grow easily at a rapid rate (maximum doubling time of approximately 20 min) on minimal media, and they can be grown at high density by controlling the rate of carbon source fed.^[15] *E. coli* have an optimal growth temperature of 37°C, but can be adapted for growth at lower or higher temperatures when necessary. There are many commercial plasmids with various properties and markers for selection and expression that include tags for purifying gene inserts and protein products, including enzymes. The overexpression of a protein can be accomplished in *E. coli* in a relatively short time (a few weeks or less is possible). *E. coli* are facultative anaerobes; when given sufficient oxygen, and with carefully controlled feeding of carbon sources, they produce minimal quantities of metabolic by-products.

An example of the use of *E. coli* for the whole-cell, nicotinamide adenine dinucleotide (NADH)-dependent reduction of ketones and the enantioselective oxidation of racemic secondary alcohols at a Japanese pharmaceutical company was reported recently.^[16] The enzymes for the conversions were originally encountered from a yeast strain and expressed in *E. coli* at high levels. The reactions did not require the addition of cofactors and were performed with unpurified enzyme from *E. coli*.

Yeast genetics are also well known but because they are higher organisms, their genetics are much more complex than *E. coli*. and other prokaryotic microorganisms. Yeast are capable of utilizing a number of sugars, as well as highly reduced species such as alkanes or alcohols for production of organic acids. One

of the challenges of working with yeast is that even in the presence of sufficient oxygen, if the sugar carbon source reaches too high a concentration, the yeast metabolism behaves as though it were lacking oxygen, traditionally referred to as fermentation. This phenomenon, known as the Crabtree effect, can be avoided with proper feeding of the substrate. Yeast can have growth rates similar to *E. coli* and generally grow at a slightly lower temperature (25–30°C) than *E. coli*. *Candida* yeasts have been used for alkane oxidation, cellulose degradation, citric acid production, and many other applications.

Method of Use of Whole Cells

When used for a single step in a synthetic route, the cells can be grown, removed from the growth medium, and resuspended in buffer. Nonliving cells can be used for this purpose for some time. The cells can also be encapsulated in some porous material such as alginate that can be formed into beads. This procedure may stabilize the cells' activity and also facilitates the removal and reuse of the cells from the reaction broth. A recent example of this procedure is the production of 4-cyanopentanoic acid by *E. coli* encapsulated in carrageenan beads that were transformed for the overexpression of a nitrilase enzyme from *A. facilis*.^[17] A volumetric productivity of 79 g/l/hr was obtained using the immobilized biocatalyst.

For multistep reactions involving cofactors, living cells are generally used. Products of primary metabolism are produced most actively during the growth phase of the process. These compounds may not be of interest to the medical community. Secondary metabolites can be produced during the resting, or stationary, phase. This is a considerable advantage for process economics, because the duration of cell growth is short in the overall process, but the stationary phase can be extended through the proper feeding of nutrients and cell recycle. In addition, the concentration of cells in the reactor is the highest after the growth phase is completed. The net result is that many more reactions can be catalyzed per cell present in the reactor by resting cells than by growing cells. Once the cell-growth phase is completed, substrate for the reaction is fed to the reactor. Normally during this phase, a small amount of carbon source is fed for cell maintenance, assuming that the reaction substrate cannot be metabolized. It is not desirable that expensive substrates be burned for cell energy, so for efficient chemistry, enzymes resulting in the conversion of substrate to by-products should be knocked out of the organism, provided these steps are not critical for the growth of the organism. One caveat for the large-scale production of secondary metabolites is that the low native

cellular concentration of these compounds is often accompanied by some feedback regulation that prevents the concentration from going above some level. Enzyme engineering, removal of the feedback regulation, or enzyme replacement may be used to obviate the cells' drive to limit production of the compound of interest. Again, this is not as difficult to accomplish in *E. coli* as in other organisms whose genetics are not completely understood.

Metabolic Engineering

Metabolic engineering (or pathway engineering) is a tool for improving the catalytic properties of whole-cell biocatalysts used for carrying out multistep reactions starting from a substrate that the organism can use as a carbon and energy source. In general, the objective of microorganisms in consuming a carbon source is energy generation for growth and reproduction. This does not match the objectives of the pharmaceutical company whose purpose is production of chemicals that may be relatively useless for the microorganism, or even toxic at high concentrations. Effective pathway engineering involves the partial redirecting of carbon flux through the metabolic pathways leading to the product of interest. A number of reviews have covered this subject in detail.^[18,19] Because many organisms can be found to have similar metabolic routes, the enzymes for carrying out similar reactions may be available from a variety of microorganisms, animals, or plants. At times, an entire pathway may be taken from a single organism; at other times, combinations of genes may be used from various sources.

A recent industrial success story is the production of 1,3 propanediol in *E. coli*.^[20,21] The pathway to produce glycerol from glucose was taken from *Saccharomyces cerevisiae*, and the genes for the enzymatic path from glycerol to 1,3 propanediol was taken from *Klebsiella pneumoniae*. The process was dramatically improved by the knockout of the klebsiella dehydratase and the enzymatic activity was replaced with the nonspecific dehydratase from *E. coli*. This switch enabled a significant increase in the final titers. Many examples of successful process development are the result of metabolic engineering.

A technique used in directed evolution, DNA shuffling, has been applied to pathway engineering.^[22] Improvements in yield and titer of processes for the production of carotenoids, aromatics, secondary metabolites, alcohols, and many others have been achieved through DNA shuffling. In addition, novel molecules have been produced that can lead to the development of new drugs.

ISOLATED ENZYME BIOCATALYSIS

Microorganisms such as bacteria, yeast, and fungi are the most convenient sources of enzymes. Microbial cells can be produced in a bioreactor, or fermentor, that provides a controlled environment favorable to cell growth. Methods to facilitate recovery of enzymes excreted to the growth media have been developed. Cells can be removed by centrifugation or ultrafiltration and the enzyme can be removed from the broth. When enzymes are not secreted from the cells, the cells must be lysed and the cell debris must be removed. These crude preparations are adequate for use as catalysts in many industrial applications, in either liquid or dry powder form. For a higher degree of purity, recombinant enzymes expressed with a purification tag may be concentrated and purified in one or two process steps. Native enzymes can also be purified to yield the desired purity by a scheme of salt or solvent precipitation followed by one or more total chromatography steps.

Many enzymatic reactions of synthetic importance require expensive cofactors. For example, oxidoreductases typically require nicotinamide, and glycotransferases require sugar nucleotide as cofactors. Because these cofactors are prohibitively expensive to be used in stoichiometric ratio, an in situ cofactor regeneration system must often be in place. Typically this can be done by coupling another enzymatic reaction that regenerates the cofactors. Formate dehydrogenase is the most commonly used enzyme for regeneration of NADH using formate as the reductant. A variation of this method is to macromolecularize the cofactor by chemically attaching poly(ethylene glycol) of molecular weight over 10,000 to NAD⁺ so that it could be retained in a membrane reactor. This method was successfully used by Degussa in the commercial production of L-tert-leucine.^[5] A number of other approaches have been explored with varying degrees of success. A recent review describes some of the methods.^[23] The necessary reducing power for monooxygenase enzymes has been performed with the help of an electrode and metal reducing agents, and recently with the use of hydrogen peroxide.^[24]

Immobilized Enzymes

In order to reduce the cost of enzymes used in organic synthesis, the immobilization of enzymes on solid supports through covalent or noncovalent attachment is commonly used. The immobilized biocatalysts can be recovered from the reaction medium by settling or filtration and can then be reused. This approach can dramatically increase the stability of the enzyme, but may result in reduced activity. Much research has

improved immobilization of enzymes in recent years. Recent reviews discuss the most recent advances in enzyme immobilization.^[25,26] Recent research has involved the incorporation of functional enzymes into polymer matrices. A review describes research involving both dehydrated enzyme powders and soluble enzymes.^[27] Considerable success was achieved with the encapsulation of soluble protein—including subtilisin, *Candida Rugosa* lipase, and several others—in hydrophilic resins with 47–61% retention of activity. The cured polymers were clear and hard, and the relative proportion of enzyme to polymer was 0.5% to 10% by weight. Polymer properties such as hydrophilicity, hardness, and aqueous swelling capacity could be manipulated by the addition of fillers and by varying the precursor molecules for polymerization. Protein–polyurethane composites are also described, that have been prepared for a number of applications, including the degradation of nerve agents. Much of this research work was carried out by LeJeune and Russell at the University of Pittsburgh in collaboration with a small company, Agentase.^[28] Silicone-based biocatalytic composites have also been developed incorporating an enzyme as a dehydrated powder or attached to a solid polymer as filler. One important goal of biocatalytic polymer research is the development of biocompatible polymers that can resist biofilm formation. The immobilization of lumbrokinase and hirudin in polyurethane composites was recently shown to resist thrombus formation on an artificial heart valve.^[29,30]

Nonaqueous Enzymology

Recent research has demonstrated the capacity of enzymes to function in nonaqueous media. Original work in this area focused on the use of lyophilized enzymes.^[31] It was shown that the pH of the buffer in which the protein resided during the lyophilization process was the pH the enzyme “remembered” when suspended in organic solvent. Recent work in this area has shown a dramatic improvement in the catalytic rate of enzymes lyophilized with high concentrations of certain salts that enable the original conformation of the enzyme to be preserved throughout the dehydration.^[32] Initial efforts to solubilize the enzyme in the organic solvent included the use of enzymes in reversed micelles. Here the enzyme was soluble in the water core of the micelles that were separated from the organic solvent by only the surfactant interface. Finally, the direct solubilization of the enzymes in the organic solvents was achieved by linking the reactive moieties at the surface of the enzyme with poly(ethylene glycol) (PEG) (referred to as the PEGylation of the enzyme). Similar results were achieved by

attachment of the enzyme to certain surfactants. A summary of this work can be found in a recent review.^[33]

Extremozymes

In recent years, the discovery of organisms that grow under extreme conditions has introduced natural enzymes that also have the capacity to function at high temperature, high pressure, high pH, and high salts.^[34] These enzymes have a number of useful applications for organic synthesis. Some of the initial studies with enzymes with such unusual stabilities involved their comparison with enzymes capable of only functioning under mild conditions. These studies indicate that thermally stable enzymes usually possess only small amino acid differences when compared to their mesophilic counterparts. However, these subtle changes often result in increased noncovalent, intramolecular interaction and compaction sufficient to confer thermostability. Important to the practical use of extremozymes is that thermostable enzymes often exhibit enhanced stability to other normally disruptive factors such as denaturation by detergents and organic solvents.^[35]

A prime example of the use of extremozymes is the DNA polymerase from *Thermos aquaticus* (Taq) in the polymerase chain reaction (PCR), which revolutionized the way nucleic acids are being manipulated. Runnion et al. documented the interesting discovery of thermal stable urease from the screening of various sites in Yellowstone National Park.^[35]

Directed Evolution

The need to redesign enzymes for biocatalysis applications is obvious because the enzymes are required to work in drastically different environments in vitro than in their natural physiological conditions. Enhanced stability and activity in nonaqueous environments, and expanded substrate spectra, are a few desirable traits (among others) that are imposed by the biocatalysis conditions.

Using directed evolution to improve a biocatalyst involves three elements. The first is the method to generate enzyme variants. Error-prone PCR and gene shuffling are commonly used to generate a library of 10,000 or more enzyme mutants. The second element is a selection or screening process that allows the beneficial mutation to be selected. This often involves the ingenious design of selection methods that allow only the mutants with the desired trait to survive. When a phenotype is not amenable to such selection, then screening must be applied to all members in the library. In this case, the need to screen hundreds of thousands

of enzyme mutants entails a high-throughput screening mechanism. Some degree of automation in the analysis step is necessary.^[36] The third element of the use of directed evolution to improve a biocatalyst is the method to combine the beneficial mutation.

Many impressive examples of successful directed evolution of enzymes are reported in the literature. One important application is improving thermal stability using directed evolution. In one such study, subtilisin protease was subjected to directed evolution and a thermostable enzyme was generated. It was shown that eight amino acid changes were responsible for enabling the thermostability of the resulting protein.^[37] Simultaneous improvement in activity and stability in organic solvent were achieved by error-prone PCR and DNA shuffling coupled with a simple visual screening.^[38] Directed evolution was also successfully used to engineer enzymes to accept non-natural substrates. In an effort to extend the range of biotransformation of toluence dioxygenase, activity 5.6 times higher toward 4-picoline was achieved by a combination of random mutagenesis and saturation mutagenesis. A more stunning change in substrate specificity was reported by Matsumura and Ellington. The *E. coli* beta-glucuronidase evolved by in-vitro directed evolution into a beta-galactosidase with 500 times improved catalytic efficiency and 52 million-fold inversion in specificity.^[39] Other examples of successful use of directed evolution can be found in recent reviews.^[18,19,23]

CONCLUSION

In recent years, important progress has been made in the application of biocatalysis technology in the production of biomedically important products and materials. However, the enormous potential of biocatalysis and diversity in nature is far from exhausted. So far, about 3000 enzymes have been characterized. This represents only about 10% of all the enzymes believed to exist in nature; only about 300 are commercially available. An overwhelming majority of the catalytic power in nature remains untapped. Many scientific and technological advances have set the stage for the rapid progress of biocatalysis over the next decade.

REFERENCES

1. Faber, K. *Biotransformations in Organic Chemistry*, 4th Ed.; Springer-Verlag: Berlin, 2000.
2. Schmid, A.; Dordick, J.S.; Hauser, B.; Kiener, A.; Wubbolts, M.; Witholt, B. Industrial biocatalysis today and tomorrow. *Nature* **2001**, *409*, 258–267.
3. Panke, S.; Wubbolts, M.G. Enzyme technology and bioprocess engineering. *Curr. Opin. Biotechnol.* **2002**, *13*, 111–116.
4. Walsh, C. Enabling the chemistry of life. *Nature* **2001**, *409*, 226–231.
5. Rozzell, J.D. Commercial scale biocatalysis: Myths and realities. *Bioorg. Med. Chem.* **1999**, *7*, 2253–2261.
6. Turner, M.K. Biocatalysis in organic chemistry (Part I): Past and present. *Trends Biotechnol.* **1995**, *13* (5), 173–177.
7. *The History of Penicillin Production*; Elder, A.L., Ed.; Chem. Eng. Prog. Symp. Ser., American Institute of Chemical Engineers: New York, 1970; Vol. 66 (11).
8. Peñalva, M.A.; Rowlands, R.T.; Turner, G. The optimization of penicillin biosynthesis in fungi. *Trends Biotechnol.* **1998**, *16*, 483–489.
9. Woodley, J.M. Advances in Enzyme Technology—UK Contributions. In *Advances in Biochemical Engineering/Biotechnology*; Springer-Verlag: Berlin, 2000; Vol. 70, 93–108.
10. Peterson, D.H.; Murray, H.C.; Eppstein, S.H.; Reineke, L.M.; Weintraub, A.; Meister, P.D.; Leigh, H.M. Microbiological transformations of steroids. I. Introduction of oxygen at carbon-11 of progesterone. *J. Am. Chem. Soc.* **1952**, *74*, 5933–5936.
11. Keller, M. Microbial Diversity from Extreme Environments Converted into Novel Bioactive Molecules, Society of Industrial Microbiology Annual Meeting, Philadelphia, PA, August 11–15, 2002.
12. Johnson, A.C.; Reitz, M. Site-Directed Mutagenesis. In *Recombinant DNA Principles and Methodologies*; Greene, J.J., Rao, V.B., Eds.; Dekker: New York, NY, 1998; 699–719.
13. Bailey, J.E. Toward a science of metabolic engineering. *Science* **1991**, *5013*, 1668–1675.
14. Holland, H.L. Microbial transformations. *Curr. Opin. Chem. Biol.* **1998**, *2*, 77–84.
15. Reisenberg, D.; Schulz, V.; Knorre, W.A.; Pohl, H.-D.; Korz, D.; Sanders, E.A.; Ross, A.; Deckwer, W.-D. High cell density cultivation of *Escherichia coli* at controlled specific growth rate. *J. Biotechnol.* **1991**, *20*, 17–28.
16. Matsuyama, A.; Yamamoto, H.; Kobayashi, Y. Practical application of recombinant whole-cell biocatalysts for the manufacturing of pharmaceutical intermediates such as chiral alcohols. *Org. Process Res. Dev.* **2002**, *6*, 558–561.
17. Hann, E.C.; Sigmund, A.E.; Hennessey, S.M.; Gavan, J.E.; Short, D.R.; Ben-Bassat, A.; Chauhan, S.; Fallon, R.D.; Payne, M.S.; DiCosimo, R. Optimization of an immobilized-cell biocatalyst for production of 4-cyanopentanoic acid. *Org. Process Res. Dev.* **2002**, *6*, 492–496.
18. Chartrain, M.; Salmon, P.; Robinson, D.; Buckland, B. Metabolic engineering and directed evolution for the production of pharmaceuticals. *Curr. Opin. Biotechnol.* **2000**, *11*, 209–214.
19. Zhao, H.; Chockalingam, K.; Chen, Z. Directed evolution of enzymes and pathways for industrial biocatalysis. *Curr. Opin. Biotechnol.* **2002**, *13*, 104–110.

20. Emptage, M.; Haynie, S.; Laffend, L.; Pucci, J.; Whited, G. *Genetically Engineered Escherichia coli Containing Nonspecific Dehydratase yghD and DHA Regulon for the Biological Production of 1,3-Propanediol with High Titer*; 2001. (E.I. du Pont de Nemours and Company: USA; Genencor International, Inc.); PCT Int. Appl. 109 pp.
21. Zhu, M.M.; Lawman, P.D.; Cameron, D.C. Improving 1,3-propanediol production from glycerol in a metabolically engineered *Escherichia coli* by reducing accumulation of sn-glycerol-3-phosphate. *Biotechnol. Prog.* **2002**, *18* (4), 694–699.
22. Patten, P.; Howard, R.; Stemmer, R. Applications of DNA shuffling to pharmaceuticals and vaccines. *Curr. Opin. Biotechnol.* **1991**, *8*, 724–733.
23. Huisman, G.W.; Gray, D. Towards novel processes for the fine-chemical and pharmaceutical industries. *Curr. Opin. Biotechnol.* **2002**, *13*, 1–7.
24. Joo, H.; Lin, Z.; Arnold, F.H. Laboratory evolution of peroxide-mediated P450 hydroxylation. *Nature* **1999**, *399*, 670–673.
25. Boller, T.; Meier, C.; Menzler, S. EUPERGIT Oxirane acrylic beads: How to make enzymes fit for biocatalysis. *Org. Process Res. Dev.* **2002**, *6*, 509–519.
26. Tischer, W.; Wedekind, F. Immobilized enzymes: Methods and applications. *Top. Curr. Chem.* **1999**, *200*, 95–126.
27. Gill, I.; Ballesteros, A. Bioencapsulation within synthetic polymers (Part 2): Non-sol-gel protein-polymer biocomposites. *Trends Biotechnol.* **2000**, *18*, 469–479.
28. LeJeune, K.E.; Russell, A.J. Covalent binding of a nerve agent hydrolyzing enzyme within polyurethane foams. *Biotechnol. Bioeng.* **1996**, *51*, 450–457.
29. Park; et al. Characterization of antithrombotic activity of lumbrokinase-immobilized polyurethane valves in the total artificial heart. *Artif. Organs* **1999**, *23*, 210–214.
30. Phaneuf, M.D.; et al. Covalent linkage of recombinant hirudin to a novel ionic poly(carbonate) urethane polymer with protein binding sites: Determination of surface antithrombin activity. *Artif. Organs* **1998**, *22*, 657–665.
31. *Enzymic Reactions in Organic Media*; Koskinen, A.M.P., Klibanov, A.M., Eds.; Blackie: Glasgow, UK, 1996.
32. Ru, M.T.; Wu, K.C.; Lindsay, J.P.; Dordick, J.S.; Reimer, J.A.; Clark, D.S. Towards more active biocatalysts in organic media: Increasing the activity of salt-activated enzymes. *Biotechnol. Bioeng.* **2001**, *75*, 187–196.
33. DeSantis, G.; Jones, J.B. Chemical modification of enzymes for enhanced functionality. *Curr. Opin. Biotechnol.* **1999**, *10*, 324–330.
34. Adams, M.W.; Perler, F.B.; Kelly, R.M. Extremozymes: Expanding the limits of biocatalysis. *Bio/Technology* **1995**, *13*, 662–668.
35. Runnion, K.; Combie, J.; Williamson, M. Thermally Stable Urease from Thermophilic Bacteria. In *Biocatalysis at Extreme Temperatures*; Adams, M.W.W., Kelly, R.M., Eds.; ACS Symposium Series, American Chemical Society: Washington, DC, 1992; Vol. 498.
36. Olsen, M.J.; Stephens, D.L.; Georgiou, G.; Iverson, B.L. A Method for Biocatalyst Evolution Using Bacterial Surface Display. In *Book of Abstracts*, 214th ACS National Meeting, Las Vegas, NV, September 7–11, 1997; American Chemical Society: Washington, DC.
37. Zhao, H.; Arnold, F.H. Directed evolution converts subtilisin E into a functional equivalent of thermitase. *Prot. Eng.* **1999**, *12*, 47–53.
38. Song, J.K.; Rhee, J.S. Enhancement of stability and activity of phospholipase A1 in organic solvents by directed evolution. *Biochim. Biophys. Acta* **2001**, *1547*, 370–378.
39. Matsumura, I.; Ellington, A. In vitro evolution of beta-glucuronidase into a beta-galactosidase proceeds through non-specific intermediates. *J. Mol. Biol.* **2001**, *305*, 331–339.

Biocompatibility Testing

Joel D. Bumgardner
Marcia Vasquez-Lee
Keertik S. Fulzele
Daniel H. Smith
Kellye D. Branch
Seth I. Christian
David L. Williams

Biomedical Engineering Program, Mississippi State University, Mississippi State, Mississippi, U.S.A.

B

INTRODUCTION

A biomaterial may be defined as any material or combination of materials used to replace, augment, or restore function of damaged or diseased tissues.^[1] Implicit in this definition is the idea that the materials used in devices must be safe in addition to effective. This characteristic or property of a material is termed biocompatibility. Biocompatibility may generally be regarded as the ability of a material to interact with living cells/tissues or a living system by not being toxic, injurious, or causing immunological reactions while performing or functioning appropriately. The spectrum of biomaterials used in medical applications is quite broad, ranging from cotton pads that stop bleeding of minor cuts within a few minutes; to catheters, dialysis tubes, and contact lenses designed to interact with tissues for a few hours to weeks; to sutures designed to slowly resorb over a few weeks to months; to implant devices like total replacement hips and heart valves intended to last 10–15⁺ years or the lifetime of the patient. To encompass this diversity of materials and applications, a consensus panel of experts has defined biocompatibility as the ability of a material to perform with an appropriate host response in a specific application.^[1] This definition thus accounts for not only the variety of applications of materials in medical devices, but also for the responses of both host tissues to material and material to host tissues. For example, cotton, which may be useful to initiate blood clotting in minor cuts, may not be suitable for use as an artificial blood vessel, because clotting would interfere with or block blood flow through the vessel. The physiological environment causes the release of low levels of metallic ions from biomedical alloys that may induce host metal hypersensitivity reactions such as Ni-allergy. For heart valves derived from animal tissues, the deposition and buildup of calcium from host fluids eventually increases material stiffness and interferes with valve opening and

closing, resulting in failure. It is hence important to realize that 1) no one material will be appropriate for all medical device applications; 2) the material, its composition, and degradation products may affect host cells and tissues; and 3) the host environment may also affect material properties and device performance.

The determination of the biocompatibility of materials and implant devices involves detailed characterization of the material (e.g., bulk and surface chemical composition; density; porosity; and mechanical, electrical, and degradation properties) and extensive testing, first at the protein/cell/tissue or in-vitro level, and then in in-vivo animal models and ultimately in human clinical trials. In-vitro tests are used to screen materials, their components, and or leachable/soluble/degradation products for cytotoxic, genotoxic, immunological, and hemolytic effects. Animal models are used to evaluate material–host tissue interactions and to predict how the device or prototype may perform in humans. Ultimately, the safety and effectiveness of the device must be evaluated in humans prior to widespread use by physicians and their patients. At each stage, biocompatibility data must be correlated with material properties and with manufacturing, sterilization, packaging, storage, and other handling procedures that also may influence test outcomes.

The design and use of biocompatibility testing protocols is provided by a variety of professional and regulatory organizations, including the American Society for Testing and Materials International (ASTM), the International Standards Organization (ISO), the American Dental Association (ADA), the National Institutes of Health (NIH), and the Food and Drug Administration (FDA). The use and documentation of biocompatibility tests are required by law in the United States and other countries, and are used to ensure that biomedical devices and their constituent materials are safe and effective under intended use conditions.^[2,3] This article provides a general introduction to the types of

biocompatibility tests. It is assumed that the physical, mechanical, electrical, and other properties of materials and device designs have been adequately evaluated and determined suitable prior to biocompatibility testing.

IN-VITRO TESTING

As compared with in-vivo and clinical studies, in-vitro environments such as cell or bacterial cultures provide a quick and relatively cheap way of estimating a material's biocompatibility. In-vitro tests allow a great deal of control over the test environment, in contrast to in-vivo tests, whose animal or human subjects may be influenced by variables such as sex, age, activity, diet, etc. If a material does not perform well in-vitro it most likely is not a good candidate for implantation. However, it must be noted that a material's in-vitro characteristics may not necessarily reflect its in-vivo performance, because dynamic cell/tissue interactions and hormonal and other physiological processes are not reproduced. The case of zinc-oxide eugenol dental cement is a classic example of the contradiction between in-vitro and in-vivo performance.^[4] In in-vitro tests, zinc-oxide eugenol cement exhibited severe toxicity, but when used in properly prepared tooth cavities, there is little if any dental pulp tissue reaction. This is because the dentin in the tooth structure acts as a barrier to reduce concentrations of molecules released from the cement to low levels. Therefore, in-vitro cytotoxicity testing should be used only to supplement in-vivo experiments.

There are many options for the type of in-vitro studies that may be performed, and the number is increasing as new technologies (e.g., quantitative polymerase chain reaction, DNA on a chip, cocultures) are developed. Common aspects of host response evaluated in vitro are cell or organelle survival, cell proliferation, metabolic and catabolic activity, cell morphology and function (alteration of shape, size, locomotive ability, phagocytic behavior, expression/response to cytokines/hormones, etc.), and cell damage, including chromosomal damage. Many of these tests are used to investigate and understand basic mechanisms of cell-material interaction that are not possible by other means, as well as to estimate the biocompatibility of materials. This section focuses on several common methods (Table 1) used primarily for estimating biocompatibility in vitro.

Cytotoxicity Tests

Cytotoxicity is the ability to cause death or damage at the cellular level by direct cell lysis or by fatally altering cellular metabolism. Inhibition of enzyme activity, changes in cell membrane permeability, and other sublethal effects may also be included in cytotoxic effects. Cytotoxicity assays are used as a screening mechanism to identify candidate materials for further biocompatibility tests. These tests measure the initial reaction of cells to the material and its components, and correlate fairly well with early tissue inflammatory, necrosis, and toxicity reactions at implant sites. Therefore, materials that do not perform well in cytotoxicity studies generally are not considered for human implantation. The three main types of cytotoxicity tests are elution, agar overlay, and direct contact. Specific test protocols may be found in ISO, ASTM, NIH, and related documents.^[5-8]

Elution tests

In elution tests, an extract or elute of a test material is taken and added to cells in culture. Serial dilutions of the elute are usually tested to determine toxic doses, and changes in cell growth or proliferation are measured and compared to nontreated cells over 24–78-hour periods. Cell proliferation or growth may be determined through direct cell counting using a hemocytometer or electronic cell counter, incorporation of radiolabeled thymidine (because thymidine is used only in DNA replication), or quantification of DNA material (because the amount of DNA is proportional to cell number). Changes in cell membrane permeability are used to estimate cell damage through the uptake of viability dyes (e.g., trypan blue or neutral red) or the release of intracellular enzymes such as lactate dehydrogenase, because large intracellular enzymes are only released from cells with severely damaged or disrupted cellular membranes. The conversion of tetrazolium salts by intracellular enzymes (primarily mitochondrial enzymes) to colored compounds that are read spectrophotometrically may also be used to estimate viability and proliferation because enzyme activity is retained only in viable cells, and because as cells grow, more dye is converted. Generally, a 50% or greater decrease in the number of cells, in the incorporation of

Table 1 Types of in-vitro tests for estimating biocompatibility

Cytotoxicity	Hemocompatibility	Mutagenecity	Hypersensitivity
Elution or extract test	Hemolysis assay	Ames test	Lymphocyte transformation test
Agar or agarose overlay test	Clotting and complement activation		Leukocyte migration inhibition test
Direct contact test			

radiolabeled compounds, or in the absorbance values of colored reagents compared to nontreated controls is associated with a severe cytotoxic response. This test is good for evaluating the toxicity of the leachable, soluble components from test materials, although the concentration of test substances must be carefully considered (because almost any material or chemical in high enough concentration will be toxic).

Agar overlay

In agar overlay tests, materials are placed on an agar or agarose layer covering the cells for 24 hours. Components from the test material are allowed to diffuse through the agar or agarose to the cells. Cytotoxicity of the diffusible components is determined by staining the cells with a viability dye and then measuring the zone of dead cells surrounding the test material. Materials that have no or only a limited zone of dead cells (limited to just under the sample) are given low cytotoxic scores, whereas materials with larger zones extending beyond the test sample are given higher cytotoxic scores. This method is good for low-density materials, powders, liquids, and high-density materials that may crush or damage cells if placed in direct contact. Care must be taken in comparing toxicity of different materials because different chemical compounds may have different diffusion rates and solubilities in the agar or agarose layer.

Direct contact tests

In direct contact tests, test materials are placed in direct contact with the cells. This may include simply growing cells on top of the test material or placing the test material on top of the cells in culture. Changes in cell growth or proliferation are measured in similar manner as in the elution tests, and changes in other cellular parameters (including morphology, intracellular ATP, intracellular signaling, chemokine/cytokine release, and total protein production/secretion) may also be measured. Zones of cell death or inhibited cell growth may also be measured in manner similar to the agar overlay tests, although care must be taken to not disturb the sample from its original position. This test is flexible with respect to testing a variety of materials and assessing of toxicity, although care must be taken with bioresorbable/degradable materials because degradation products may build up to nonrelevant toxic levels.

The degree of cytotoxicity of materials and their released components in these tests is determined in comparison to controls. Positive controls such as pure copper or latex rubber are materials that will cause an effect (e.g., cell death). Negative controls such as

Teflon[®] or titanium are materials that cause no or minimal effect.

Other factors affecting in-vitro tests are the types of cells (e.g., fibroblasts, osteoblasts, macrophage, endothelial, stem cell, etc.), species (e.g., human, rat, pig, etc.), source (normal tissues, tumor tissues, or tissues transformed by viral agents), period of exposure (hours to days), and culture parameters (type of medium, serum additions, number of cells per unit area).^[9] These factors must be considered in comparing results and in establishing experimental designs. Generally, cells derived from tumors or transformed cell lines are recommended for cytotoxicity tests because their characteristics are well-defined, they are less variable in genetic makeup and growth, and they are more readily available than normal (noncancerous, non-transformed) cells. Normal cells are more appropriate for investigating questions of specific cell-material interactions (such as cell attachment and spreading, elaboration of extracellular matrix, attachment and growth of osteoblasts, or activation of platelet or macrophage cells). Also, the selection of cell type should be appropriate for the intended application of the implant material.

Hemocompatibility Tests

Hemocompatibility tests are used to evaluate the effect of a material or released compounds on blood coagulation processes, thrombus formation, and hemolysis (the destruction of red blood cells). An upset in blood chemistry or damage to one of its components may be fatal, and therefore most medical devices that come into contact with blood are required to undergo hemocompatibility tests. Materials such as hypodermic needles that make only brief contact with blood need not be tested.

In hemolysis tests, test and control materials or their extracts are incubated with red blood cells, isolated from rabbits, mice, or rats for three hours with intermittent shaking to keep samples mixed and in contact with blood. The amount of hemoglobin released into the supernatant from the cells is determined spectrophotometrically and reported as percent hemolysis with respect to negative controls (materials that do not cause red blood cell lysis). The concentration of extracts and surface-to-volume ratio of solid or particulate materials to volume of red blood cells must be closely controlled to obtain reproducible results.

To evaluate the effects of materials and their surfaces on blood clotting, materials are exposed to whole blood serum. The time for the development of a clot on the material is compared to that of a reference material. Additionally, the serum may be tested for changes in different blood clotting factors after exposure to the

test material. However, because changes in blood flow may occur due to device design (e.g., turbulent flow may increase hemolysis and or clotting) the evaluation of test materials under flow conditions should be undertaken. For example, the number of adherent platelets may be determined per unit area after exposure to whole blood under controlled flow conditions.^[10] Because of their role in coagulation and thrombus formation, increased adhesion of platelets to the surface may indicate increased clotting rates. Ex-vivo tests are also popular for evaluating hemocompatibility under more in-vivo-like blood flow conditions. In ex-vivo tests, blood is temporarily shunted from an artery and vein of an experimental animal and allowed to flow over or through the test material inserted between inlet and outlet segments.

In addition to the limits of reproducing blood flow dynamics in vitro, the variability in blood from donors due to age, diet, health, etc. makes standardization of responses difficult. Hence, repeated tests with the same donor blood and use of appropriate reference materials are important considerations.

Mutagenicity and Genotoxicity

Mutagens are those materials that modify the genome of a host. Materials that have mutagenic effects on their host may therefore be classified as genotoxic. It is widely accepted that carcinogenic behavior proceeds via a mutation in the genome. It is thus expected that all carcinogens will be mutagens; however, the opposite is not necessarily true (i.e., a mutagen is not always a carcinogen).

One common method for studying mutagenicity is the Ames test. Briefly, the Ames test uses a mutant bacterial cell line (*Salmonella typhimurium* or *Escherichia coli*) that must be supplied with histidine to grow. The cells are cultured in a histidine-free environment, and only those materials that mutate the cells back to a state of histidine independence will allow the cells to grow. The Ames test is often used as an early screen for

mutagenicity, although it will not detect the full range of mutagens.

Hypersensitivity Tests

Tests such as the leukocyte migration inhibition and lymphocyte transformation tests have been used as in-vitro models to estimate delayed hypersensitivity reactions to implant materials and their released components.^[11–13] The use of these tests in screening the sensitization potential of materials is not well justified because genetic factors control individual sensitivity; however, the tests may be useful in assessing the sensitivity of individuals to particular materials and compounds.

IN-VIVO TESTING FOR BIOCOMPATIBILITY

Many systemic physiological processes are complex and cannot be simulated in vitro. Therefore in-vivo animal testing is necessary prior to human clinical testing. In-vivo tests (Table 2) are used to evaluate the local and systemic interactions of host tissues with implant materials or devices and their leachable/soluble/degradation components. These include physiological effects of the implant and its released components on local and systemic tissues, tissue response to the implant materials and designs, changes in extracellular matrices and regulatory biomolecules, and changes in the material and device due to host physiology.

Animal Welfare Issues

Use of animals for scientific purposes may cause them physical or psychological pain or harm. Thus, it is imperative that scientists use the least painful humane methods accepted in modern veterinary or laboratory practice. In-vivo animal testing should be done only when absolutely necessary. Whenever possible, a critical survey of the scientific literature and in-vitro testing

Table 2 Types of in-vitro tests for estimating biocompatibility

Short-term implantation tests	Long-term functional tests	Sensitization	Irritation	Other
Subcutaneous, intramuscular, and intraperitoneal implantation tests to evaluate general tissue necrosis, fibrosis, and inflammation	Devices or compositionally identical prototypes are implanted in appropriate animal models to replicate/simulate intended end-use in humans. Functionality of device and histopathological evaluations of tissues/organs are performed.	Guinea pig maximization test Occluded patch test Open epicutaneous test	Skin Ocular Mucosal	Genotoxic Carcinogenic Reproductive Cerebrospinal Hemocompatible

should be carried out before in-vivo tests are undertaken. In the United States, legislative, regulatory, and professional organizations have developed thorough guidelines for using animals in research.^[14–18] Similar laws and regulations exist in most other countries that provide protection for lab animals and strongly encourage

- Experimental designs based on the relevance to human or animal health, advancement of knowledge, or the good of society.
- Use of appropriate species, quality, and number of animals.
- Avoidance or minimization of discomfort, distress, and pain in concert with sound science.
- Use of appropriate sedation, analgesia, or anesthesia.
- Establishment of experimental end points.
- Conduct of experimentation on living animals only by or under the close supervision of qualified and experienced persons.

Animal and Implant Site Selection

No single species represents an ideal general model for the human species. A variety of sources are devoted to the selection of appropriate animal models.^[5–8,18–21] Anatomical, biochemical, physiological, pathological, and/or psychological characteristics must be considered when choosing an animal model.

Animal Tests

Guidelines to carry out animal biocompatibility tests are described by standards organizations and government regulatory agencies such as ASTM, ISO, NIH, FDA, ADA, U.S. and European Pharmacopeia, and others.^[5–8] The tests can be divided into functional and nonfunctional tests.

Nonfunctional tests

Nonfunctional tests are generally conducted first to study the direct interactions of the implant with the physiological environment. In these tests, it is assumed that acute toxic or inflammatory reactions are non-specific and thus may be evaluated in soft tissues (e.g., subcutaneous, intramuscular, intraperitoneal) because implantation in soft tissue requires minor surgery. Specialized sites such as the cornea and cerebral cortex are used for materials intended for those specific applications. After predetermined periods of time, the animals are euthanized and tissue reactions to the test materials are evaluated histologically and compared to

controls in the form of blanks or currently accepted implant materials that have a well-characterized known tissue response. Histological analyses may be scored or graded based on degree of tissue necrosis/degeneration, fibrosis, and types and amounts of inflammatory (polymorpho-nuclear leukocytes, macrophages, lymphocytes, etc.) and foreign body giant cells present. Because many factors of the implant device and design (such as functional loading and multicomponent interactions for a total hip) are not taken into account, such tests do not fully measure biocompatibility. Thus, they tend to be of short to intermediate duration, usually a few weeks to 24 months (e.g., ASTM F 763 Standard Practice for Short-Term Screening of Implant Materials or ASTM F 981 Standard Practice for Assessment of Compatibility of Biomaterials for Surgical Implants with Respect to Effect of Materials on Muscle and Bone). Fluid extracts from device materials may be used to determine the acute biological reactivity of possible leachable/soluble/degradation compounds when injected into soft tissues.

Functional tests

After evaluation in soft tissues, the next step is the selection of an animal model and implant site for the device or prototype to simulate the site that ultimately will be utilized in humans (i.e., functional testing). Functional tests require that the implant be placed, at least in some degree, in the functional mode that it would experience in human implant service. Animal model functional tests are obviously of much greater complexity because the design, fabrication, surface treatment, packaging, mechanical testing, and implantation of these device may be different in the device finally produced for human use. Variation in testing procedures with different animal models and variation in species histology can create difficulties in comparing, assessing, and interpreting results. Changes in animal behavior (for example, in preferentially not loading a limb with an implant) may also be important.

There are also specific in-vivo physiological assessment tests for genotoxicity, carcinogenicity, reproductive toxicity, delayed-type hypersensitivity, and systemic toxicity. Detailed protocols for these tests may be found in standards documents.^[5–8]

Genotoxicity testing

Genotoxic agents or genotoxins cause alterations in DNA or chromosomal structure or other DNA or gene

damage that result in permanent inheritable changes in cell function. The major genotoxic effects tested for are gene mutations, chromosomal aberrations, and DNA effects. Gene mutation and chromosomal aberration tests detect actual lesions in the DNA molecule, whereas DNA effects tests detect events that may lead to cell damage. Although in-vitro tests for genotoxicity/mutagenicity may be conducted using microorganisms or mammalian cells, an in-vivo animal model is also recommended when the scientific literature or in-vitro tests indicate potential genotoxicity of the implant or its components.^[22]

Gene mutation tests detect base-pair mutations, frameshift mutations, and small deletions. Chromosomal aberration tests detect chromosomal damage induced after one cellular division; structural changes in the chromosomes are evaluated while cells are in the metaphase stage of division. DNA effects tests (generally, mouse bone marrow micronucleus test) detect damage to the chromosomes or the mitotic apparatus of immature red blood cells found in bone marrow.

Carcinogenicity tests

Carcinogenicity testing is done to determine the tumorigenic potential of devices, materials, or extracts (after a single exposure or multiple exposures) over the total life span of the test animal. Specifically, such testing should be considered for a device that will have permanent contact (longer than 30 days) with tissues, either as an implant or as an externally communicating device.

Rodents are invariably chosen as the test species because their relatively short life spans make it practical to carry out lifetime studies. Although most investigators and regulatory agencies agree on undertaking carcinogenicity testing, many available standards lack specific information regarding the number of test animals to be used, the kinds of observations needed, the extent of histopathological evaluations, the number of survivors required at the end of the study, and the type of statistical evaluations to be used. The best scientific practices for carcinogenicity testing have been summarized.^[23]

Sensitization tests

Sensitization testing is done to determine whether chemicals or compounds that may be released from specific biomaterials and devices elicit sensitization reactions. Sensitization or hypersensitivity reactions are a result of immunologically mediated reactions resulting in redness (erythema) and swelling (edema). Because most such reactions are of the dermal cell-mediated type rather than the humoral or antigen-antibody type, the skin of laboratory animals (usually

guinea pigs because they exhibit dermal sensitivity similar to humans) is used in sensitivity testing.^[24] Three common tests are the guinea pig maximization test, the occluded patch test, and the open epicutaneous tests. Typically, the material or material extract is injected subcutaneously. Then one to two weeks after exposure, the test material or solution is reapplied to the skin and the degree of erythema and edema is graded. The use of adjuvants that enhance the immunological response (Freund's adjuvant, sodium lauryl sulphate) may be used during initial exposure to help identify weakly sensitizing materials. These common tests do not detect compounds released from the material that may also act as an adjuvant, or responses to antigens such as the plant proteins found in natural latex. Kimber et al. and Merritt have reviewed sensitization testing.^[12,25]

Irritation tests

Irritation is a localized inflammatory response to single, repeated, or continuous exposure to a material without involvement of an immunological mechanism. The intracutaneous, primary skin, and ocular irritation tests are three in-vivo, nonclinical tests commonly used to evaluate materials for possible contact irritation. The intracutaneous test is similar to the sensitization test. Fresh extract is injected intracutaneously at multiple sites on the shaved backs of albino rabbits and then 24, 48, and 72 hours after injection, the test and control sites are observed and scored for the severity of erythema or edema. Extracts that produce a significantly greater response than controls are considered irritants. In the primary skin irritation test the portions of the test material itself are simply placed on the shaved backs of rabbits with the aid of an occlusive dressing and the response after 24 hours is evaluated by methods similar to those used in the intracutaneous test.

The ocular irritation test is usually performed for ophthalmological materials. A small sample of material (powder or fluid extract) is placed directly into the pocket formed by withdrawing the lower eyelid of a rabbit. The rabbit's other eye is left untreated as a control. After 72 hours, the eye is evaluated for redness and swelling of the conjunctiva, response of the iris to light, corneal opacity, and presence of discharge. Scores are compared with standards to determine the degree of eye irritation. In addition to the intracutaneous, primary skin, and ocular irritation tests, mucosal (oral, rectal, penile, and vaginal) irritation tests may be carried out as necessary.^[26]

Systemic effects

Systemic toxicity may occur when chemicals and compounds released from implant materials are distributed

by the blood and lymphatic system and damage organs and tissues remote from the implant site. The various routes used for sample administration are topical or dermal, inhalation, intravenous, intraperitoneal, and oral. The material may be tested in powder, fluid extract, or solid forms and evaluations may be based on a single or multiple doses. Systemic effects are categorized on the basis of time to initiate adverse effects: acute (within 24 hr), subacute (in 14–28 days), subchronic (10% of an animal's life span), and chronic (longer than 10% of an animal's life span). Test methods used in longer-term systemic toxicity studies are much the same as those used in acute toxicity tests, except that larger groups of animals are used.

Reproduction toxicity tests evaluate the potential effects of devices, materials, or extracts on reproductive function, embryonic development, and prenatal and early postnatal development. Reproductive/development toxicity tests should only be conducted when the device has potential impact on the reproductive system.^[26]

The selection of evaluation in in-vivo protocols depend on the animal used (mice, rabbit, dog, etc.) and the implant material/device. Tests performed during the in-life phase often include measurements of body weight and food consumption, blood and urine analyses, and eye examinations. Recommended hematology test parameters are hematocrit, hemoglobin percentage, erythrocyte counts, and total and differential leukocyte counts. Platelet counts and measurements of prothrombin and thromboplastin concentrations and clotting time are usually performed for blood contacting materials and devices. Clinical pathology protocol should include assays that test proper electrolyte balance, carbohydrate metabolism, and liver and kidney function. Postmortem analyses include gross observations at necropsy, organ weighing, and microscopic examination of selected tissues. Considerable care must be taken in the design of subchronic and chronic studies to make certain they provide assurance of the device's safety.

CLINICAL TRIALS OF BIOMEDICAL IMPLANTS

Clinical trials are designed to test the safety of a new device in humans and are conducted only after they have been extensively evaluated by in-vitro and in-vivo animal models. Although clinical trials do provide vital information on the effectiveness of a biomedical implant device, they do not specifically test the biomaterial; rather, they test the device composed of the biomaterial(s) in a specific application. Unlike in-vivo animal studies, the implants in clinical trials remain exposed to the experimental subject even after the period of observation ends. As time passes and more subjects are exposed to the device without adverse

results, confidence in acceptable biological performance increases.

It is standard practice in new drug trials to employ a double-blind study in which a placebo is randomly administered to a portion of the defined group of patients to ensure that neither doctor nor patient know whether the active drug is being administered. In implant testing, however, it is not possible to pair the implanted patient with a placebo-treated patient, because it is not possible to conceal the implantation site from surgeon or patient. For clinical trials of implant materials, comparisons must be made between the condition of the patient before and after implant surgery. An implanted patient must also be compared with a patient with similar implants made of different materials, as well as a nondiseased (control) individual of the same age, sex, and similar home/workplace environment.

Because comparisons must be made between patients, certain standards exist for the selection and treatment of patients. Inclusion/exclusion criteria of the target patient population clarify the intended use of the device. The exclusion criteria serve to eliminate factors that increase operative risk or that may confound the outcome of the study. Medical care administered in a clinical trial must be under the care of a medical professional, and the patients must give informed consent to any experimental procedure.^[27–29] Patients must be informed of any possible benefits or risks of the procedure, and the identity of the patients must be protected and confidentiality of their medical records preserved. Also, there must be a reasonable possibility of benefit combined with reasonable assurance of usual risk.

Clinical trials are divided into three phases. Phase I (early trial) involves simply selecting a new treatment from several options for further study. The biomaterial is tested on a small group of people (~60–80). Phase II studies a larger group of subjects (~100–300) to test the effectiveness and further evaluate the safety of the biomaterial. It is divided into phase IIA and phase IIB. If the new treatment selected in the early trial is not effective, phase IIA (preliminary trial) examines whether further studies should be performed or the treatment abandoned. Otherwise, phase IIB (follow-up trial) estimates the effectiveness of the new treatment if it appears promising. Phase III is a comparison of the effectiveness of the new treatment with a standard of management or some other treatment, and is also divided into phase IIIA and phase IIIB. A large study group of approximately 1000–3000 people is used. If the new implant improves the clinical outcome (phase IIIA), the treatment is refined and examined further in phase IIIB, usually by multiple investigators and institutions.

Clinical trials must be performed under the control of a defined prospective protocol that describes the

implant device and outlines the indications of the surgical procedure, the uniform surgical procedure used, and the postoperative treatment and follow-up schedule. The number of patients included and needed to complete the trial to demonstrate benefit or improvement of the device is also critical to the study design, because too few patients will reduce significance, and too many may waste time and resources. Inclusion of and consultation with biostatisticians in developing clinical trial designs will help ensure that the device is adequately evaluated for potential use. Other considerations important in the design of a clinical trial include the study duration, examination schedule, and reporting of adverse events. Adverse events are noted and serious intraoperative and postoperative adverse events reported. A list of possible adverse events should also be included in the protocol of the study. To ensure that a clinical trial is conducted ethically and that the rights of the study participants are protected, an institutional review board or independent committee of physicians, statisticians, and community advocates is selected to help prepare the protocols and review the procedures and safeguards in any experimental program involving human subjects before the clinical trial begins.

The reports of clinical trials should discuss the accuracy and precision of all measurements as well as define a minimum confidence level for all statistical measures of data (usually $p < 0.05$). The reports must also include confidence intervals or other measures of significance associated with all derived parameters, and must indicate the significance of any conclusion arrived at by analysis of the trial.

The complication incidence rate for many biomedical materials/devices is extremely small. A study of the costs and benefits associated with research to improve then current orthopedic prosthetic devices concluded that failure/complication rates, even in the 1970s, were acceptably small and that the investment needed to significantly reduce these rates would be unreasonable with respect to the resulting benefits.^[30] In many cases in which the present technology is extremely successful, newer developments cannot demonstrate statistically superior outcomes because of limitations of size and duration of clinical trials. In these cases, the decision to pursue further improvements in the material or device depends on the argument that the human and financial costs to individuals as a result of device malfunction and failure is greater than the financial costs to the manufacturers and research institutions. Because the test of adequate performance described in the legal system is that the device be "safe and effective" and pose no "unreasonable" risk or hazard to the patient, decisions on what may be considered acceptable failure rates remain subjective, and depend on public opinion, expert advice, and administrative action.

Retrieval of implants and associated tissues (due to failure or other complications or at patient death) provides valuable information about the long-term biocompatibility of implants and their materials. Thorough analysis of failed implants provides insights into material design, selection, and manufacture and handling; implant design; surgical protocols; and the interaction between implant and host tissues as they relate to biocompatibility of the device. This information may be used to improve material design, selection, and manufacture, and to revise surgical protocols. Retrieval of implants and tissues is also important because the data provide a true assessment of the safety and efficacy of the device and document long-term host-material interactions. Evaluation of both device failure and success is necessary for improving understanding of the biocompatibility of implant materials and devices.

CONCLUSION

Evaluating the biocompatibility of biomedical materials and devices involves extensive testing at the in-vitro (protein/cellular/tissue), in-vivo animal, and human clinical use levels to determine how the materials and devices will perform in their intended applications. The success of a material and device depend on how the materials and device design affect the host and how the host affects device function and material properties. In-vitro tests have evolved to provide quick and relatively fast biocompatibility screening of candidate biomedical materials, but they are not yet capable of accurately predicting the clinical success of implant devices. However, with societal pressures to develop alternatives to animal testing, and with advances in cell and molecular biology techniques, advances in in-vitro methods for the prediction of clinical success are continuing. The use of in-vivo animal models is necessary to evaluate the interaction of an entire physiological system with the device and its components, and to establish safety and performance capabilities prior to use in humans. It is imperative that great care and planning are used in planning in-vivo animal tests to ensure that relevant and significant biocompatibility data are obtained. Only in clinical studies can the biocompatibility of a material or device and their ability to replace, restore, or augment the function of damaged or diseased human tissues be evaluated. This requires close evaluation of not only patients during the study period, but also of patient health records and of retrieved devices (failed devices and successful devices retrieved at death), even after the planned clinical study time frame to establish epidemiologically successful biomedical materials and devices. Finally, it is noted that methods and evaluation criteria for determining biocompatibility are routinely reviewed

and amended as additional information is collected. New methods and standards are being developed as new hybrid devices are developed and refined that involve both biological and synthetic materials (e.g., a patient's stem cells grown on resorbable substrate).

ACKNOWLEDGMENTS

This manuscript was supported by the Agricultural and Biological Engineering Department and the Mississippi Agriculture and Forestry Experiment Station (MAFES, Ms No. BC 10316) at Mississippi State University.

ARTICLES OF FURTHER INTEREST

Biologic and Synthetic Apatites; Biological Effects of Radiofrequency Electromagnetic Field; Biomaterials Immune Response; Blood-Material Interactions; Cell Culture Assays; Diamond and Diamond-Like Carbons; Host Reactions; In Vivo Cardiovascular Models; Thrombosis

REFERENCES

- Williams, D.F. Definitions in Biomaterials. Proceedings of a Consensus Conference of the European Society for Biomaterials, Chester, England, March 3–5, 1986; Volume 4, Elsevier: New York.
- Medical Device Amendment of the U.S. Food and Drug Act of 1976.
- Safe Medical Devices Act of 1990.
- Mjör, I.A.; Hensten-Pettersen, A.; Skogedal, O. Biological evaluation of filling materials: A comparison of results using cell culture techniques, implantation tests and pulp studies. *Int. Dent. J.* **1977**, *27*, 124–132.
- International Standard ISO 10993—Biological Evaluation of Medical Devices, Part 1: Guidance on Selection of Tests*; International Organization for Standardization: Geneva, Switzerland, 1997.
- ASTM F 748–98 Selecting Generic Biological Test Methods for Materials and Devices. In *Annual Book of ASTM Standards Vol. 13.01 Medical Devices*; Emergency Medical Services, ASTM International: West Conshohocken, PA, 2002.
- Guidelines for Physicochemical Characterization of Biomaterials, Publication 80–2186*; National Institutes of Health, U.S. Department of Health and Human Services: Washington, DC, 1980.
- AAMI Standards and Recommended Practices, Vol. 4 Biological Evaluation of Medical Devices*; Association for the Advancement of Medical Instrumentation: Arlington, VA, 1994.
- Bumgardner, J.D.; Gerard, P.D.; Geurtsen, W.; Leyhausen, G. Cytotoxicity of precious and nonprecious alloys—Experimental comparison of in vitro data from two laboratories. *J. Biomed. Mater. Res. (Appl. Biomater.)* **2002**, *63*, 214–219.
- Skarja, G.A.; Kinlough-Rathbone, K.L.; Perry, F.D.; Rubens, F.D.; Brash, J.L. A cone-and-plate device for the investigation of platelet biomaterial interactions. *J. Biomed. Mater. Res.* **1997**, *34*, 427–438.
- Rae, T. Cell Biochemistry in Relation to the Inflammatory Response to Foreign Materials. In *Fundamental Aspects of Biocompatibility*; Williams, D.F., Ed.; CRC Press, Inc.: Boca Raton, FL, 1981; Vol. I, 159–181.
- Merritt, K. Immunological Testing of Biomaterials. In *Techniques of Biocompatibility Testing*; Williams, D.F., Ed.; CRC Press, Inc.: Boca Raton, FL, 1986; Vol. II, 123–136.
- ASTM F1906–98 Evaluation of Immune Responses in Biocompatibility Testing Using ELISA Tests, Lymphocyte Proliferation and Cell Migration. In *Annual Book of ASTM Standards Vol. 13.01 Medical Devices*; Emergency Medical Services, ASTM International: West Conshohocken, PA, 2002.
- Animal Welfare Act of 1996 as amended. United States Code, Title 7, sections 2131 to 2156.
- Guide for the Care and Use of Laboratory Animals, Publication 86–23*; National Institutes of Health, U.S. Department of Health and Human Services: Washington, DC, 1985.
- Public Health Service Policy on Human Care and Use of Laboratory Animals*; Office of Protection from Research Risks, National Institutes of Health: Bethesda, MD, 1986.
- American Veterinary Medical Associate (AVMA). Report of the AVMA panel on euthanasia. *J. Am. Vet. Med. Assoc.* **1993**, *202*, 229–249.
- International Standard ISO 10993—Biological Evaluation of Medical Devices, Part 2: Animal Welfare Requirements*; International Organization for Standardization: Geneva, Switzerland, 1992.
- Mendenhall, H.V. Animal Selection. In *Handbook of Biomaterials Evaluation: Scientific, Technical and Clinical Testing of Implant Materials*, 2nd Ed.; von Recum, A.F., Ed.; Macmillan Publishing Company: New York, 1999; 475–480.
- Schmidt-Nielsen, K. *Scaling: Why is Animal Size so Important?*; Cambridge University Press: Cambridge, UK, 1985.
- Vale, B.H.; Willson, J.E.; Niemi, S.M. Animal Models. In *Biomaterials Science: An Introduction to Materials in Medicine*; Ratner, B.D., Hoffman, A.S., Schoen, F.J., Lemons, J.E., Eds.; Academic Press: San Diego, CA, 1996; 238–242.
- Dearfield, K.L.; Cimino, M.C.; McCarroll, N.E.; Mauer, I.; Valcovic, L.R. Genotoxicity risk assessment: A proposed classification strategy. *Mutat. Res.* **2002**, *521*, 121–135.
- Combes, R.; Schechtman, L.; Stokes, W.S.; Blakey, D. The international symposium on regulatory testing and animal welfare: Recommendations on best scientific practices for subchronic/chronic toxicity and carcinogenicity testing. *ILAR J.* **2002**, *43*, S112–S117. (suppl).
- ASTM F720–81 Testing Guinea Pigs for Contact Allergens: Guinea Pig Maximization Test. In *Annual*

- Book of ASTM Standards vol. 13.01 Medical Devices; Emergency Medical Services* ASTM International: West Conshohocken, PA, 2002.
25. Kimber, I.; Basketter, D.A.; Berthold, K.; Butler, M.; Garrigue, J.L.; Lea, L.; Newsome, C.; Roggeband, R.; Steiling, W.; Stropp, G.; Waterman, S.; Wieman, C. Skin sensitization testing in potency and risk assessment. *Toxicol. Sci.* **2001**, *59*, 198–208.
 26. Thomas, J.A. In-Use Testing of Biomaterials in Biomedical Devices. In *Handbook of Biomaterials Evaluation: Scientific, Technical and Clinical Testing of Implant Materials*, 2nd Ed.; von Recum, A.F., Ed.; Macmillan Publishing Company: New York, 1999; 313–320.
 27. *Human Subject Research: A Handbook for Institutional Review Boards*; Greenwald, A., Ryan, M., Milvihill, J., Eds.; Plenum Press: New York, 1982.
 28. Federal policy for the protection of human subjects, Department of Health and Human Services regulations, Title 45 Code of Federal Regulations, part 46 (45 CFR46).
 29. World Medical Association Recommendations guiding doctors in clinical research. *Br. Med. J.* **1964**, *2*, 1119–1129.
 30. Piehler, H.R. Risk–benefits of orthopaedic prosthetic devices. *Orthop. Rev.* **1978**, *7*, 75–88.
- Guidelines for Blood–Material Interactions, Publication 85–2185*; National Institutes of Health: Bethesda, MD, 1985.
- Handbook of Biomaterials Evaluation—Scientific, Technical, and Clinical Testing of Implant Materials*, 2nd Ed.; von Recum, A.F., Ed.; Macmillan Publishing Co.: New York, 1999.
- Hanson, S.; Lalor, P.A.; Niemi, S.M.; Northrup, S.J.; Ratner, B.D.; Spector, M.; Vale, B.H.; Willson, J.E. Testing Biomaterials. In *Biomaterials Science—An Introduction to Materials in Medicine*; Ratner, B.D., Hoffman, A.S., Schoen, F.J., Lemons, J.E., Eds.; Academic Press: San Diego, CA, 1996; 215–242.
- International Standard ISO 7405—Dentistry—Preclinical Evaluation of Biocompatibility of Medical Devices Used in Dentistry—Test Methods for Dental Materials*; International Standards Organization: Geneva, Switzerland, 1997.
- Kirkpatrick, C.J.; Bittinger, F.; Wagner, M.; Kohler, H.; van Kooten, T.G.; Klein, C.L.; Otto, M. Current trends in biocompatibility testing. *Proc. Inst. Mech. Eng., H* **1998**, *212*, 75–84.
- Mohan, K.; Sargent, H.E. Clinical trials, an introduction. *Med. Dev. Diag. Ind.* **1996**, *18*, 114–119.
- Mollnes, T.E. Complement and biocompatibility. *Vox Sang.* **1998**, *74* (Suppl. 2), 303–307.
- Technical Committee ISO/TC 94. *International Standard ISO 10993—Biological Evaluation of Medical Devices, Part 1–17*; International Standards Organization: Geneva, Switzerland, 1997.
- Techniques of Biocompatibility Testing*; Williams, D.F., Ed.; CRC Press: Boca Raton, FL, 1986; Vols. I and II.
- Thull, R. Physicochemical principles of tissue material interactions. *Biomol. Eng.* **2002**, *19*, 43–50.
- USP XXIII: The Pharmacopeia of the United States of America, 23rd Revision Incorporating The National Formulary, 18th Revision*; The United States Pharmacopeial Convention, Inc.: Washington, DC, 1995.
- Walum, E.; Stenberg, K.; Jenssen, D. *Understanding Cell Toxicology—Principles and Practice*; Ellis Horwood Ltd.: New York, 1990.
- Wataha, J.C. Biocompatibility of Dental Materials. In *Restorative Dental Materials*, 11th Ed.; Craig, R.G., Powers, J.M., Eds.; Mosby: St. Louis, 2002; 125–163.

SUGGESTED READING/ADDITIONAL RESOURCES

- American National Standard/American Dental Association (ANSI/ADA) Specification No. 41—Biological Evaluation of Dental Materials*; American Dental Association: Chicago, IL, 2001.
- Annual Book of ASTM Standards Vol. 13.01 Medical Devices; Emergency Medical Services*, ASTM International: West Conshohocken, PA, 2002.
- Black, J. *Biological Performance of Materials—Fundamentals of Biocompatibility*; Marcel Dekker, Inc.: New York, 1999.
- Dee, K.C.; Puleo, D.A.; Bizios, R. *An Introduction to Tissue—Biomaterial Interactions*; John Wiley & Sons, Inc.: Hoboken, NJ, 2002; 173–184.
- Fundamental Aspects of Biocompatibility*; Williams, D.F., Ed.; CRC Press: Boca Raton, FL, 1981; Vols. I and II.

Biocomposites

Huinan Liu

*Department of Materials Engineering, Purdue University, West Lafayette, Indiana, U.S.A.,
and Division of Engineering, Brown University, Providence, Rhode Island, U.S.A.*

Grace E. Park

*Weldon School of Biomedical Engineering, Purdue University, West Lafayette,
Indiana, U.S.A.*

Thomas J. Webster

Department of Biomedical Engineering, Purdue University, West Lafayette, Indiana, U.S.A.

B

INTRODUCTION

Currently, there is a high demand to develop new materials that can be used to repair or replace damaged tissue owing to injuries, diseases, and/or genetic malformations. The term “biocomposite” means a material that consists of two or more distinct constituents to obtain complex physicochemical, mechanical, and biological properties, which are required by biomedical applications and cannot be satisfied by any individual components. Owing to the widespread use of biocomposites in regenerating bone, orthopedic applications will be emphasized here. In the U.S.A. in 1995, there were 216,000 total knee replacements, 134,000 total hip replacements, and close to 100,000 bone grafting procedures. In all, musculoskeletal conditions cost the U.S.A. \$214.9 billion in 1995.^[1] Traditionally, autografts, allografts, xenografts, and metal implants have been used to repair bone fractures and other skeletal defects. However, these substitutes are far from ideal as each has its own problems and limitations. For example, autografts are associated with donor shortage and donor site morbidity, whereas allografts and xenografts have the risk of disease transmission and detrimental immune responses. Most common metal implants are made of stainless steel, cobalt–chromium alloys, and titanium-based alloys. Major problems of those metals for implant applications are their mismatched mechanical properties compared to host tissue, which causes necrosis of the surrounding tissue and subsequent implant loosening owing to the well-documented “stress shielding” effect.^[2] It has also been well documented that metals do not possess surface properties that allow for sufficient quick new bone growth.^[2] Collectively, the above-mentioned problems (as well as others) with orthopedic implants have led to an average 10–15 yr current functional implant lifetime before clinical failure necessitates a replacement. To overcome such limitations of these traditional materials, over the past

25 yr, researchers have been interested in developing biocomposites that provide for a wide diversity of properties necessary for successful orthopedic implants.^[1–5] This desire comes from considering that living tissues are composites themselves with a number of levels of hierarchy. In almost all biological systems, a range of tissue properties (such as physicochemical, mechanical, biological, etc.) is present. All properties are of great importance and are key to the success of a tissue. Because of this, often a clinical need can only be fulfilled by a material that exhibits a similar complex combination of properties. The development of biocomposites offers great promise to improve the efficacy of current tissue substitutes that are too frequently single-phase materials that lack such property diversification.

Owing to similarities with the chemical composition of bone, polymer/ceramic composites were one of the first biocomposites considered for orthopedic implant applications.^[3] Specifically, for polymer/ceramic biocomposites, it is possible to obtain a wide range of mechanical and biological properties by modifying the type and distribution of the ceramic phase in the polymer matrix. Importantly, increased interactions between biocomposites and host tissue can be achieved by optimizing polymer/ceramic composites. For this reason, a wide variety of biocomposites have been synthesized and fabricated for various biomedical applications during these years.^[4] This article systematically addresses biocomposites applied to repair or replace damaged tissue (in particular bone) in a comprehensive manner emphasizing polymer/ceramic composites and their applications in tissue engineering.

For biocomposites to be successful, advanced manufacturing techniques must be exploited to take full advantage of their potentially tailorable properties. In this light, this article focuses on three main areas. First, the requirements of biocomposites for bone tissue-engineering applications are described. Second, the article introduces natural tissues as composites to

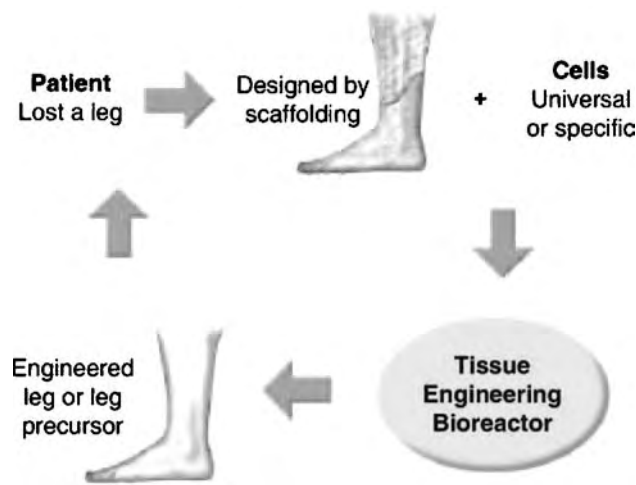


Fig. 1 Schematic diagram of the bone tissue-engineering concept using a hypothetical example of leg regeneration. (From Ref. [15].)

emulate in synthetic biocomposites stressing their mechanical properties. The third area concerns the description of various manufacturing techniques that can be used to control desirable biocomposite properties, including examples of the design and fabrication of three-dimensional composite scaffolds for tissue-engineering applications.

BIOCOMPOSITES USED AS BONE TISSUE-ENGINEERING SCAFFOLDS

A material that is designed to interact with biological systems to solve problems in repair and reconstruction of lost, damaged, or deceased tissue can be regarded as

a biomaterial.^[4] Biomaterials are widely used in implants and medical devices such as artificial hip and knee joints, dental implants, and internal as well as external bone fracture fixators. Of course, there are many applications of biomaterials in orthopedic applications. Most of the research is based either on those polymer/ceramic biocomposites that involve nondegradable materials for use as coatings to promote bone growth or on those that involve degradable materials to fully replace bone while dissolving.^[4] Owing to its current popularity, this section will introduce the application of biomaterials in bone tissue engineering.

Tissue engineering offers a promising opportunity for tissues to regenerate in a natural way. Some would argue that this would be the best approach to healing a bone fracture rather than inserting a permanent material intended to be in the patient for the remainder of their lifetime. Figs. 1 and 2 illustrate the bone tissue-engineering concept using a hypothetical example of leg regeneration.^[15] In this context, scaffolds are designed as temporary three-dimensional mirror matrices onto which cells grow and regenerate the desired tissue. Because the scaffold materials are biodegradable, they will fully resorb after completing the template functions. Importantly, both ceramics and polymers may degrade over time, and thus, have been used extensively in bone tissue-engineering applications.

However, the optimal combination of polymer and ceramic has not been found to date and with good reason.^[4] Currently, the scientific challenges in bone tissue engineering are: 1) developing suitable three-dimensional scaffold fabrication techniques that allow for the creation of appropriate templates for cell adhesion, growth, and proliferation in favorable three-dimensional orientations; 2) understanding cell

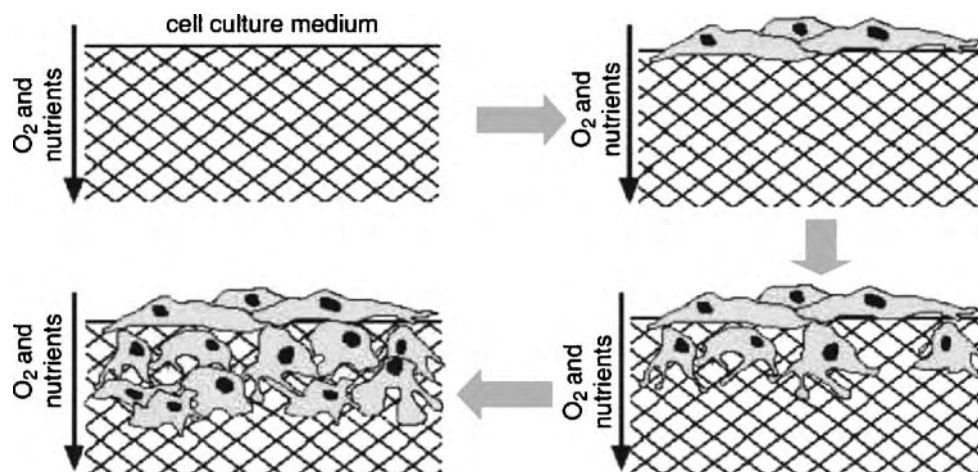


Fig. 2 Schematic diagram of basic functions of scaffolds to promote cellular infiltration over time.

functions on these scaffolds; and 3) combining these two concepts for the construction of an optimal bone regenerative material. This problem is further complicated owing to the multitude of material choices available. In terms of these material solutions, there are several crucial requirements to optimize tissue-engineering scaffolds. The scaffolds should be: 1) cytocompatible; 2) biodegradable with a controllable degradation and resorption rate that matches new cell/tissue ingrowth; 3) favorable to cellular attachment, migration, proliferation, and differentiation leading to the formation of a desirable extracellular matrix; 4) favorable to tissue integration and vascularization owing to proper interconnecting pores (or three-dimensional structure); 5) able to adequately support physiological loads through appropriate mechanical properties that match those of the intended site of implantation; and 6) fabricated through inexpensive but reliable high resolution techniques to match the tissue morphology.^[16] Bearing all of these requirements in mind, several types of polymers and ceramics (as well as fabrication techniques) have been developed for synthesizing biocomposite scaffolds with some success; this is the focus of the sections that follow.

NATURAL AND SYNTHETIC BIOCOMPOSITES

As mentioned, the term “biocomposite” means a material that consists of two or more distinct constituents to obtain complex physicochemical, mechanical, and biological properties that are required by biomedical applications and cannot be satisfied by any individual components.^[4] Thus, biocomposites combine various properties of two or more components to meet the multitude of requirements for select applications. In this respect, there are two main types of biocomposites: natural biocomposites and synthetic biocomposites, which are described in the following sections.

Natural Biocomposites

At the molecular and microstructural level, tissues are natural composite materials with a complex collection of entities (such as organic/inorganic phases) assembled in a controlled manner that allows for a tissue to function. In general, tissues of our body are grouped into hard and soft tissue classifications. Soft tissues (such as cartilage, ligament, tendon, blood vessel, skin, etc.) are formed from elastic (elastin) and nonelastic (collagen) fibers with a cellular matrix actively being remodeled between these fibers. Biological structures such as tendons (which link muscles to bones) are low in elastin content, thus, allowing muscle

Table 1 The relative mass percentages of the four basic components of typical cortical bones

Component	Mass %
Organic—Type I collagen fibers	16%
Inorganic—HA ($\text{Ca}_{10}(\text{PO}_4)_6(\text{OH})_2$)	60%
Ground substance (proteins, polysaccharides, and mucopolysaccharides)	2%
Water	22%

(From Ref. [3].)

movement to be translated to bone.^[3] However, ligaments (which links tissue to bone) are high in elastin content allowing movement between bones but retaining sufficient support to stop joints from dislocating.^[3]

Hard tissues (such as bones and teeth) are formed from four basic components as shown in Table 1.^[3] Of course, hard tissues generally have higher elastic moduli and higher strengths than soft tissues. The relative fractions of each component vary between bone type (i.e., cancellous or porous bone and cortical or dense bone) and condition of the patient (specifically, age, health, disease, etc.). The two most important naturally occurring forms of bone are cancellous and cortical (Fig. 3). Cancellous bone (also called trabecular or spongy bone) is a sponge-like structure, which closely resembles an isotropic material when not fully mature. Although immature, cancellous bone might structurally resemble an isotropic material, its chemical composition is quite diverse and, thus, better suited for composite replacement. In contrast, cortical bone is structurally highly anisotropic with reinforcing fibers along its loading axis. It has a highly organized blood supply called a Haversian system. The relative percentages of mass for typical cortical bones are included in Table 1.

The collagen fibers provide the framework and architecture of bone with the hydroxyapatite (HA) particles located in between the fibers. The ground substance is formed from proteins, polysaccharides, and mucopolysaccharides, which act as a bioactive cement filling the spaces between collagen fibers and HA mineral. These components are important and form the bulk of the bone. However, the living parts of bone are present in its voids that are fluid filled and contain cells that create (osteoblasts) and resorb (osteoclasts) the surrounding bone. It is this part of the bone that governs its structure, state of health, and is difficult to recreate using single-phase materials. Bone is just one simple example of a complex composite material naturally found in our body that researchers are trying to mimic using synthetic materials.

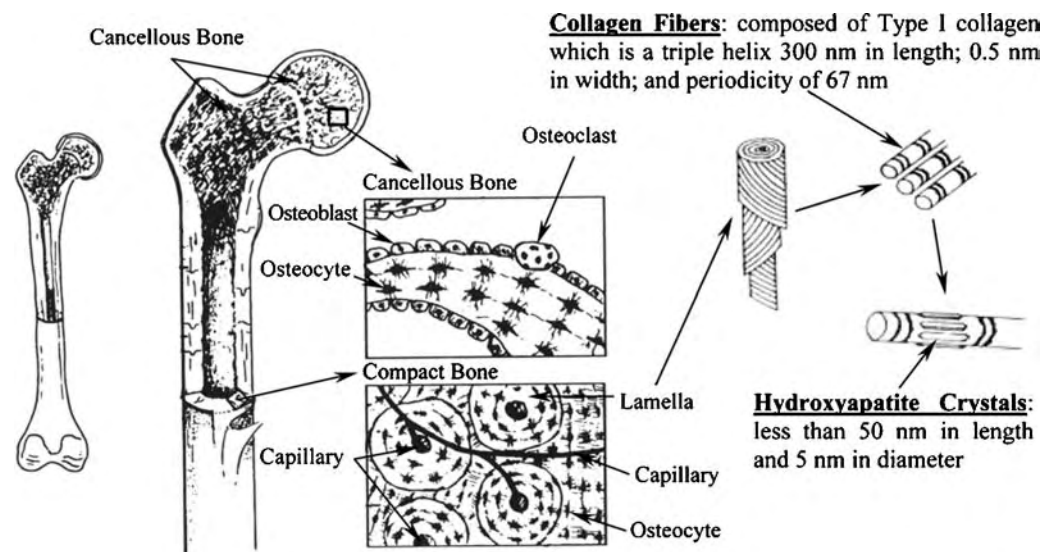


Fig. 3 Schematic composite architecture of the human femur. (From Ref. [17].)

Synthetic Biocomposites

To emulate this complex structure of bone, a wide variety of polymers and ceramics have been studied. Table 2 lists some of these biocomposite formulations and their various implant applications.^[4] These composite formulations can be further broken into resorbable compared to nonresorbable polymer or ceramic components; these are further described in the following sections.

Resorbable compared to nonresorbable polymer components

Polymer/ceramic biocomposites can be classified into resorbable and nonresorbable components

according to their extent of biodegradability. Resorbable biocomposites are composed of a reinforcing phase and a matrix in which at least one component is fully degradable in body fluids; such composites are mainly proposed to bridge small bone fractures. Clearly, the advantage of these resorbable materials is that they do not need to be removed later when the bone has healed, thus reducing the additional risk and pain associated with a subsequent surgery. Moreover, degradable materials are quite useful to provide for a prolonged, controlled release of pharmaceutical agents [like bone morphogenetic proteins (BMPs)] for orthopedic applications. However, most polymers utilized today in orthopedics are non-resorbable (such as those used as polyethylene acetabular cups). These nonresorbable biocomposites

Table 2 Some applications of synthetic biocomposites

Applications	Biocomposites
External fixators	Epoxy resin/carbon fibers (CF)
Spinal surgery	Polyurethane (PU)/bioglass, Polyetheretherketone (PEEK)/CF, epoxy/CF, polysulfone (PS)/CF, and polyethyleneterephthalate (PET)/hydrogel
Bone cements	Polymethymethacrylate (PMMA)/HA, PMMA/ultrahigh molecular weight polyethylene (UHMWPE) fibers
Hip prosthetic stems	Polyether-imide (PEI)/CF-Glass fiber (GF), epoxy/CF, PEEK/CF, PS/CF, UHMWPE/CF, and polytetrafluoroethylene (PTFE)/CF
Bone filling, regeneration	Polypropylene fumarate (PPF)/TCP, PLGA/HA fibers, polyethylene glycol (PEG)-Polybutyleneterephthalate (PBT)/HA, Starch/HA
Tendons and ligaments	PET/Hydrogels, PU/CF, PU/GF
Vascular grafts	PU/polyethylene-oxide-co-lactic acid (PELA), PET/gelatin
Abdominal wall prosthetics	PET/PU, PET/collagen
Dental bridges	PMMA/UHMWPE, PMMA/CF, PMMA/Kevlar (KF)

(From Ref. [4].)

are useful as they provide specific long-lasting mechanical properties necessary for the repair of very large bone defects.

It is important to note that most resorbable biocomposites are so based on the bioresorbable characteristics of the synthetic polymer matrix. The most commonly used polymers in synthetic biocomposites are biodegradable aliphatic polyesters. The degradation products of these polymers can be removed by natural metabolic pathways. Poly(glycolic acid) (PGA), poly(lactic acid) (PLA), and their copolymers poly(lactic-co-glycolic acid) (PLGA) as a family of aliphatic polyesters are some of the most popular degradable scaffolds.^[18,19] Poly(glycolic acid) degrades rapidly in aqueous solutions *in vivo* because of its relatively hydrophilic nature, which loses mechanical integrity between two and four weeks.^[5] For a PLA scaffold, months or even years are required before mechanical integrity is lost *in vivo*.^[6] The extra methyl group in the PLA repeating unit (compared with PGA) provides for more hydrophobic properties, thus, reducing its molecular affinity to water, which leads to a slower hydrolysis rate. Overall, these polymers degrade through nonenzymatic hydrolysis of the ester bonds. The degradation products of PGA, PLA, and PLGA are nontoxic, natural metabolites, and are eventually eliminated from the body in the form of carbon dioxide and water. The degradation rates of these polymers can even be tailored to satisfy the biological requirements from several weeks to several years by altering chemical composition (e.g., the ratio of lactic to glycolic acid in the PLGA copolymers), crystallinity, molecular weight, and molecular weight distribution.^[7] Importantly, these polymers (PLA, PGA, and PLGA) are approved by the US Food and Drug Administration (FDA) for certain human clinical applications.

There are other less common aliphatic polyesters used in bone tissue engineering, such as poly(ϵ -caprolactone) (PCL). The PCL degrades at a significantly slower rate than PLA, PGA, and PLGA.^[8] A slow degradation rate makes PCL less attractive for general tissue-engineering applications, but more attractive for long-term fixation devices. Poly(ϵ -caprolactone)-based copolymers have recently been synthesized to improve degradation properties of other polymer/ceramic biocomposites.^[9] In addition, poly(propylene fumarate) (PPF) is an important synthetic biodegradable polymer gaining in popularity, as it can degrade through hydrolysis of the ester bonds similar to glycolide and lactide polymers. As the mechanical properties of PPF can vary greatly according to the synthesis method and the cross-linking agents used, the versatility of PPF is being explored in numerous orthopedic applications from small to large bone fractures.^[10]

Resorbable compared to nonresorbable protein components

Naturally derived polymers such as collagen have also been used in biocomposite scaffold fabrication. Collagen is a fibrous protein and a major natural extracellular matrix component of many tissues. Collagen, as the most popular natural polymer for tissue regeneration by far, has very useful biological properties desirable for bone tissue engineering, such as degradability, cell adhesiveness, and biocompatibility.^[18] However, there are concerns with regard to the use of collagen because of inconsistent sample preparation, poor mechanical properties as a stand alone orthopedic material, and a less controllable biodegradability compared to the synthetic polymers mentioned above.^[18] Denatured collagen (also called gelatin) has also been processed into porous materials for tissue repair. Importantly, these naturally derived proteins have been incorporated with synthetic polymers and ceramics and have shown promise for orthopedic applications.^[18]

Resorbable compared to nonresorbable ceramic components

For the inorganic component of resorbable polymer/ceramic biocomposites, calcium phosphates [such as calcium tetraphosphate ($\text{Ca}_4\text{P}_2\text{O}_9$), tricalcium phosphate (TCP, $\text{Ca}_3(\text{PO}_4)_2$), HA, $(\text{Ca}_{10}(\text{PO}_4)_6(\text{OH})_2$), and its derivatives, as well as their combinations] are frequently used in orthopedic applications as they resemble the natural inorganic component of bone.^[18] However, calcium phosphate derivatives possess a wide range of degradation properties based mostly on their degree of crystallinity. Specifically, less crystalline materials (like TCP) have higher dissolution rates than more crystalline materials (like HA). These are the most common types of ceramics used to fabricate scaffolds to accommodate bone tissue regeneration. These ceramics are widely considered to be osteoconductive because their surface properties support osteoblast (bone-forming cells) adhesion, growth, and differentiation. They are also reported to be osteoinductive (capable of inducing bone formation) resulting from their capacity to bind and concentrate BMPs *in vivo*.^[11] Moreover, selected ceramics such as HA and TCP can react with physiological fluids and form tenacious bonds to hard and soft tissues through cellular activity; thus, they have been termed “bioactive.” The atomic ratio of Ca/P in calcium phosphates can be varied from 1.5 to 2 and the degradability of calcium phosphates generally varies with Ca/P ratio. Further studies have shown that HA and TCP have different characteristics *in vivo*,

although both forms have Ca/P ratios (1.50~1.67) within the range known to promote bone ingrowth. In general, HA has been found by many to be more osteogenic, while TCP has advantages as it degrades much faster, thus, possibly providing for improved bone integration.^[12] In addition, bone growth factors can be easily embedded in TCP and owing to its degradability it can be released in a controllable fashion to increase bone growth. Although these bioceramics are biocompatible, can promote bone ingrowth, and have a long history in orthopedics, they are inherently brittle, difficult to process into complex shapes, and cannot by themselves match the mechanical properties of true bone.^[18]

Design of optimal biocomposite properties

The design of polymer/ceramic composites offers an exceptional approach to combine the advantages of biodegradable polymers and bioactive ceramics to optimize physical and biological properties of scaffolds for bone regeneration. First, in polymer/ceramic composites, osteoblast functions (such as adhesion, proliferation, and deposition of calcium containing mineral) can be enhanced owing to improved osteoconductivity properties provided by the bioactive ceramic phase. Second, ceramic particles (such as HA and TCP) used as inclusions in biodegradable polyesters can provide a pH buffering effect at the polymer surface to tailor the desired degradation and resorption kinetics of the polymer matrix. This

prevents acceleration of polymer degradation, avoiding the formation of an unfavorable environment for the cells, and reducing the harmful side-effects (such as inflammation) that arise from acidic polymer degradation by-products.^[13]

Third, the mechanical properties may be improved by adding a stiffer particulate ceramic phase in the polymeric matrix. Typical mechanical properties for soft tissues, hard tissues (bone), synthetic bioactive ceramics, degradable polymers, and their composites are shown in Table 3.^[14] As mentioned, one mode of failure for orthopedic implants is stress and strain imbalances at the bone-implant interface owing to a mismatch in mechanical properties. From Table 3, it can be seen that the addition of HA increases the elastic modulus of PLGA closer to bone. This is because pure PLGA has an elastic modulus of 1400–2080 MPa, while pure HA has an elastic modulus of 95,000 MPa. As expected, when combining PLGA and HA, the elastic modulus increases to 4200 MPa, which is closer to the elastic modulus of cortical bone (as measured in either the longitudinal or the transverse directions) (Table 3).^[14] Thus, as can be expected, composites can have mechanical properties that match that of a host tissue, whereas the individual components do not. In this manner, polymer and ceramic individual components do not have optimal properties for bone tissue-engineering applications when used individually, but may (as the next section will further emphasize) when combined. Clearly, however, such individual components must be assembled properly for these advantages to be realized.

Table 3 Mechanical properties of selected biocomposites compared to natural tissue

Materials	Strength (MPa)	Elastic modulus (MPa)	Elongation (%)
Tendon (low elastin content)	46.5–53.0	402	9.4
Skin (high elastin content)	7.6	0.1–0.2	78
Elastin	1.0	—	100
Ligament	29.5	303	—
Collagen	50–100	—	10
Cortical bone (longitudinal)	130–133	12000–17700	3
Cortical bone (transverse)	52–60	12800–13400	1
Cancellous bone	2–7.4	390–400	2.5
Bioglass [®]	42	35	—
HA	50	95000	—
PLLA	55.2–82.7	2400–4200	5–10
PGA	68.9	6900	15–20
PLGA	41.4–55.2	1400–2080	3–10
PCL	20.7	34500	300–500
PLGA/HA	24.9	4200	—

(From Ref. [14].)

BIOCOMPOSITE FABRICATION TECHNIQUES

Biocomposites can be fabricated using a wide range of technologies. The selection of the most appropriate manufacturing technology is also influenced by the production volume capability and allowable cost. But most importantly, fabrication systems must be developed, which can match the complex three-dimensional organization of tissues; this was often

ignored in the first generation of biocomposite synthesis methods. To engineer functional tissues and organs successfully, the scaffolds have to be designed to facilitate cell distribution and guide tissue regeneration in three dimensions. Understandably so, this has been the focus of numerous more sophisticated fabrication techniques^[17,20–50] that are summarized in Table 4^[28] and are elaborated below.



Table 4 Comparison of biocomposite fabrication methods

	Resolution (μm)	Advantages	Disadvantages
<i>Acellular 3-D scaffolds</i>			
Fabrication using heat micro molding ^[28,29]	20–30	Use of well-established fabrication methods, usually automated Simple, reusable molds	Must seed cells postprocessing, less control in cell placement and distribution Limited to thin membranes, each layer must be a contiguous structure, manual alignment required
Selective laser sintering ^[30–32]	400	High porosity, automated	High temperatures during process, powder may be trapped
FDM ^[30,33,34]	250–700	No trapped particles or solvents, automated	High temperatures during processing
3-D plotting ^[45]	1000	Use of hydrogel materials, automated	Limited resolution
Fabrication using light stereolithography ^[31,35]	70–250	Ease of use, easy to achieve small features, automated	Limited choice of materials (must be photosensitive and biocompatible), exposure of material to laser
Fabrication using chemical 3-D printing ^[36–38]	200–500	Versatile, high porosity, automated	Limited choice of materials (e.g., organic solvents as binders), difficult to reduce resolution below polymer particle size
Pressure-assisted microsyringe ^[29]	10	High resolution, not subjected to heat, automated	Viscosity dependent, no inclusion of particles
Fabrication by matrix molding ^[39]	200	Use of biological matrix materials (collagen), mold fabrication can use automated methods	Features must be interconnected, weaker mechanical properties
<i>Cell-laden 3-D scaffolds</i>			
Cellular assembly organ printing ^[40–42]	100	Precise placement of cells throughout construct, ability to place multiple cell types Incorporation of cell aggregates or tissue explants, precise cell placement, automated	Limited fabrication conditions (sterility, temperature, pH), still in earlier phases of development Lack of structural support, dependence on self assembly
Laser-guided deposition ^[43]	<1	Precise single cell placement, automated	Has yet to be extended to 3-D structures, lack of structural support
Cell/biopolymer hybrids hydrogel photopatterning ^[44]	100	Incorporation of living cells within scaffold, leverages existing hydrogel chemistry (incorporation of peptides, degradation domains), versatile	Not yet automated, exposure of cells to ultraviolet light, diffusion of large molecules limited by hydrogel pore size

(From Ref. [27].)

Solvent-Casting Particulate-Leaching

One of the oldest biocomposite techniques developed, which significantly advanced the field is solvent-casting particulate-leaching.^[19] Although particulate-leaching has been widely used to fabricate scaffolds for tissue-engineering applications, to date they have not proved to be sophisticated enough to accurately match the complex structures found in bone (Fig. 3). Briefly, this technique involves producing a suspension of polymer/ceramic composites in a solvent. Salts are ground into small particles, which are then transferred into a mold. The polymer/ceramic suspension is then cast into the salt-filled mold. After the evaporation of the solvent, the salt crystals are leached out (using water) to form a porous structure. The main advantage of this process is its simplicity and control over pore properties. The pore size can easily be controlled by the size of the salt crystals and the porosity by the salt/composite ratio. However, certain critical variables (such as pore shape and interpore openings) are not well controlled using solvent-casting particulate-leaching techniques. Other porogens instead of salt (like gelatin) can allow the manufacturer to control pore structures, such as the creation of spherical as opposed to fractal pores. Fig. 4 illustrates a PLGA/HA scaffold that was fabricated by solvent-casting particulate-leaching techniques.^[47] Sodium hydroxide chemical treatment was further used to increase pore wall roughness that promoted bone ingrowth compared to those with traditional smooth pore walls.^[47] Authors attributed increased bone cell function to pore wall roughness that approximated those found in natural bone. Thus, these early biocomposite fabrication techniques can be modified to create a next-generation, more successful scaffold.

Gas Foaming

Owing to the importance of porosity properties in controlling bone formation in biocomposites, a number of additional techniques have been developed. For example, a polymer/ceramic suspension can be saturated with carbon dioxide (CO₂) at high pressures to control porosity properties through a technique called gas foaming.^[28] The solubility of the gas in the polymer then decreases rapidly by allowing the CO₂ pressure to come back to atmospheric level. This results in nucleation and growth of gas bubbles in the polymer/ceramic matrix as pores with sizes ranging between 100 and 500 μm are created. But the porosity of composites produced by this technique is limited by the solubility of gas in specific solvents. In addition, limited control of pore interconnectivity exists through the use of gas foaming. Nonetheless, gas foaming provides another easy process to control pore sizes in polymer/ceramic biocomposites.

Melt Molding

Several techniques also exist in biocomposite fabrication to control pore interconnectivity, which is important for the exchange of nutrients/wastes from pore to pore.^[28-30] One such process involves filling a Teflon mold with PLGA powder and gelatin microspheres of known dimensions and then heating the mold to above the glass transition temperature of PLGA while applying pressure to the mixture. This treatment causes the PLGA particles to bond together in a controllable fashion, thus, providing for control over pore interconnectivity. Once the mold is removed, the gelatin microspheres are dissolved by immersing the mixture in water and then the scaffolds

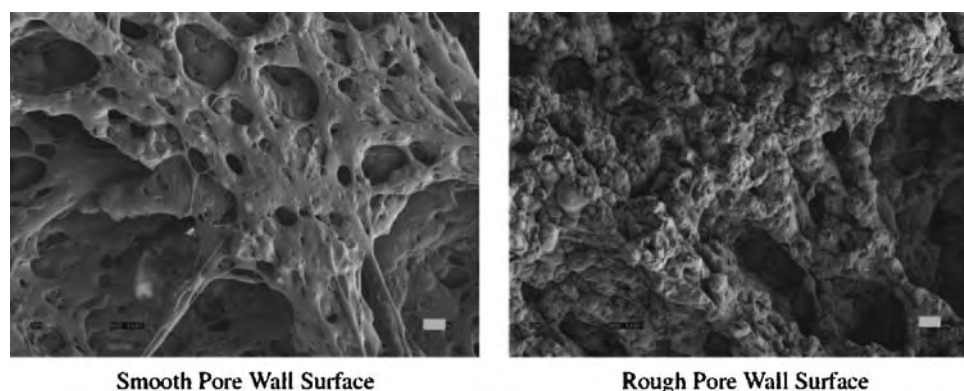


Fig. 4 Increased PLGA/HA pore wall roughness by solvent-casting particulate-leaching techniques enhances bone cell function. Increased pore wall roughness was prepared by chemically etching PLGA/HA scaffolds by solvent-casting particulate-leaching techniques. Bar = 10 μm. (From Ref. [47].)

are dried. Scaffolds produced by this method assume the shape of the mold. Fig. 5 highlights the transfer of unique pore wall surface roughness values in PLGA/titania biocomposites using such molding techniques.^[46] Specifically, the study pictured in Fig. 5 attempted to determine the influence of osteoblast responses on pore walls with nanometer scale features.^[46] These nanoscale pore wall features (both in the original and cast formulations) mimic those found naturally in bone. Results of this study demonstrated that compared to respective smooth pore wall biocomposites, both the molds and PLGA/titania biocomposites prepared from those molds with higher wall nanometer surface roughness values increased bone formation.^[49]

Interestingly, the melt molding process has been easily modified to incorporate other ceramics, specifically short fibers of HA.^[13] The uniform distribution of HA fibers throughout a PLGA scaffold was accomplished by using a solvent-casting technique to prepare a composite material of HA fibers, PLGA matrix, and gelatin (or other salt porogens), followed by a melt molding process. Thus,

similar to the above techniques, various manufacturing techniques can be combined to provide for unique biocomposite properties that can direct bone formation.

Freeze Drying

The primary advantage of adding a ceramic to a polymer phase is to enhance mechanical properties.^[25] For example, increased mechanical properties have been reported for HA and gelatin composites fabricated into foams via novel freeze-drying and cross-linking techniques.^[28] Specifically, in this study, HA powder was added at up to 30 wt.% into a gelatin solution, and the mixtures were freeze-dried and further cross-linked. The pure gelatin foam had a well-developed pore configuration with porosity and pore size diameters of $\sim 90\%$ and 400–500 μm , respectively. With the addition of HA, as expected, the porosity decreased and the pore shape became more irregular. However, the hydrophilicity and degree of water adsorption decreased when HA was added to the

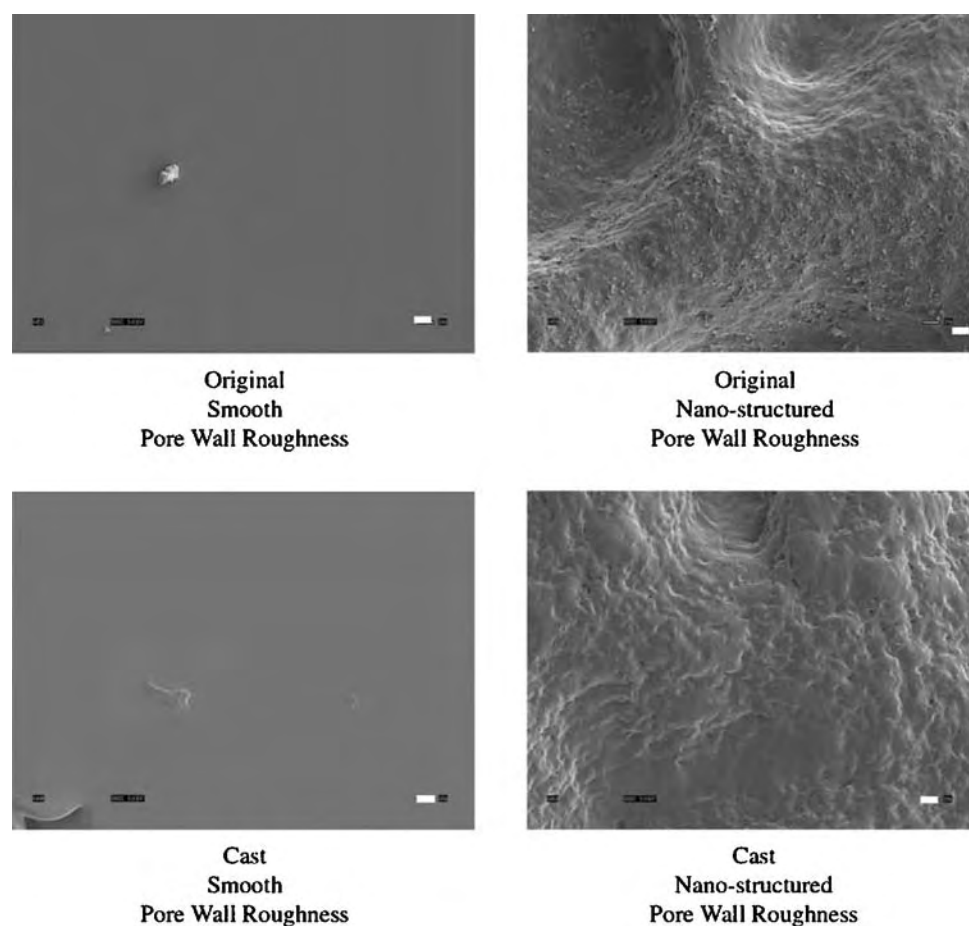


Fig. 5 Increased pore wall roughness in PLGA/titania biocomposites prepared by melt molding techniques enhances bone cell function. Bars (lower right in white) = 1 μm . (From Ref. [46].)

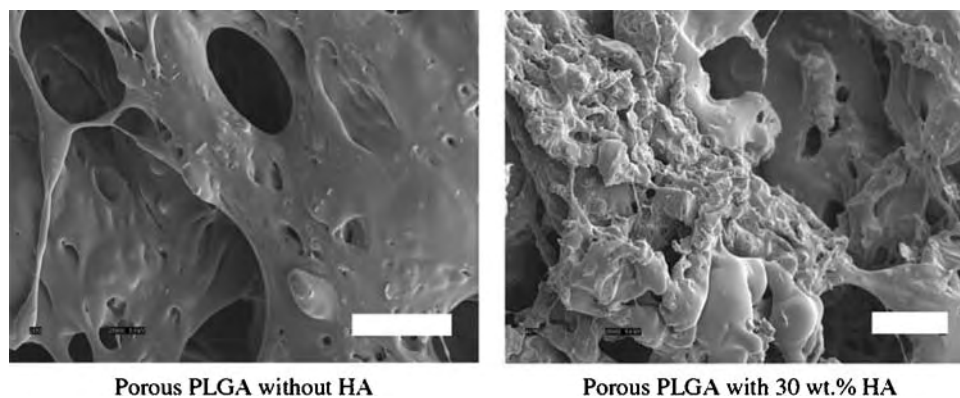


Fig. 6 Altered porosity properties owing to the addition of HA to PLGA. Note the decreased pore size and change in pore wall roughness owing to the presence of HA in PLGA. Increased bone formation was observed on porous PLGA with 30 wt.% HA compared to PLGA without HA. Bar = 10 μm . (From Ref. [49].)

gelatin foams. Clearly, the addition of HA increased the strength of the foam that allowed these composites to be more applicable to the orthopedic tissue-engineering community (i.e., increasing HA amount in the foams allowed for higher compressive stresses and provided for higher elastic moduli in both dry and wet states). The influence of changes in pore properties in biocomposites on bone growth, though, is only beginning to be investigated. An example of this change in pore structure and size owing to ceramic addition to a polymer is presented in Fig. 6. In this study,^[49] the HA particulates, in sizes of about 2–5 μm , were distributed within the PLGA network homogeneously to increase surface roughness, which promoted bone growth. Thus, although more studies are needed, two potential advantages can be obtained simultaneously when adding a ceramic component to polymers: 1) increased mechanical and 2) cytocompatibility properties.

Phase Separation

A number of disadvantages are present when combining polymer and ceramic phases.^[27] Namely, under certain conditions, multiphase materials become thermodynamically unstable and tend to separate into more than one phase to lower the free energy of the system. While investigators have overcome some of these limitations (by using sonication or chemical functionalization)^[27] to increase dispersion of the ceramic into the polymer phase, agglomeration is still an important issue for biocomposite synthesis. Fig. 7 illustrates the reduction of titania agglomeration in PLGA biocomposites using sonication.^[27] Increased bone formation has been documented on well-sonicated PLGA/titania biocomposites.^[27] Authors attribute greater bone cell responses on well-sonicated PLGA/titania biocomposites to matching the dispersion of organic and inorganic phases in natural bone.

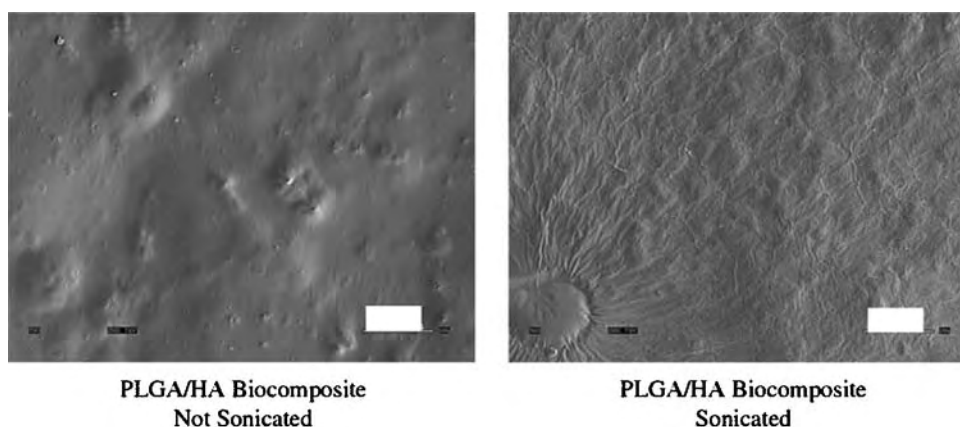


Fig. 7 Altered dispersion of HA in PLGA. Note the increased dispersion of HA in PLGA owing to sonication, which increased bone formation. Bars (lower right in white) = 10 μm . (From Ref. [27].)

On the other hand, phase-separation can be quite useful in the fabrication of biocomposites. For example, phase-separation techniques have been used to fabricate porous membranes for filtration and separation.^[28] Phase separation can also be achieved by lowering the temperature to induce solvent crystallization from a polymer/ceramic suspension. This process is often defined as a solid–liquid phase separation (solid phase formation in a liquid phase). After the removal of the solvent crystals (sublimation or solvent exchange), the space originally occupied by the solvent crystals becomes porous. This technique is quite versatile and can be used to fabricate scaffolds from many types of polymers and ceramic composite materials.^[28]

By manipulating the phase separation conditions, various pore structures can be achieved. For example, many tissues (such as nerve, muscle, tendon, ligament, and dentin) have oriented tubular or fibrous bundle architectures.^[3] To facilitate the organization and regeneration of such tissue types, a scaffold with a high porosity and an oriented array of open microtubules may be desirable. To achieve this goal, a novel phase separation technique has been developed to orient rod-shaped crystals from a polymer solution.^[29] After the removal of these rods, a parallel array of microtubules that can guide tissue growth is formed. This oriented tubular scaffold has anisotropic mechanical properties similar to fibrillar and tubular tissues, and has been shown to facilitate cell organization into oriented tissues.

Moreover, in a novel bone tissue-engineering application, Liu et al.^[20] prepared collagen/HA composite scaffolds by phase-separation techniques. Hydroxyapatite powder was added into a collagen solution and homogenized by a speed stirrer.^[20] Mixtures with different compositions were poured onto 100 mm culture dishes and then rapidly transferred into a refrigerator at -30°C to solidify the mixture and induce solid–liquid phase separation. The solidified mixture was kept at that temperature for two hours and then lyophilized for two days. Resulting collagen/HA scaffolds were porous with three-dimensional interconnected fiber microstructures with pore sizes of 50–150 μm . However, the pores formed using this technique usually have diameters on the order of a few to tens of microns and are often not uniformly distributed. Such small pore sizes not well distributed in biocomposites limit tissue ingrowth and result in poor mechanical properties problematic for many tissue-engineering applications, particularly for those in orthopedics. Controlled phase separation processes, primarily thermally induced phase separation, have recently been explored for scaffold fabrication to overcome these problems.^[28]

Closed-Mold Processes

To mimic the natural orientation of HA and collagen in bone in biocomposites, closed-mold processes are often used.^[40] These techniques gained popularity for producing ceramic fiber-reinforced plastic materials. Closed-mold processes involve the use of a molding apparatus into which mechanical force is applied. While ceramic fiber-reinforced polymer composites possess numerous mechanical property advantages (such as matching the anisotropic mechanical properties of long bones), recent cytocompatibility advantages leading to greater bone growth have been reported.^[50] In this study, carbon nanofibers (CNF) of similar dimensions to HA and collagen in bones were aligned into micropatterns on a model polymer [polycarbonate urethane (PCU)] surface using a closed-mold process. This was accomplished by pressing a mixture of CNF and soluble PCU into micron-patterned gaps (20 μm in width) of a mold placed on the surface of PCU. After allowing the polymer to cure, aligned micron regions of CNF were formed, which directed calcium phosphate mineral deposition by bone cells (Fig. 8). In this manner, substrates were created which promoted osteoblasts to deposit aligned regions of HA and collagen similar to what is observed in long bones (such as the femur).^[50] Current materials used as orthopedic implants for long bones (specifically titanium) do not promote osteoblasts to deposit aligned calcium phosphate minerals.

In addition to the example given above, another example of a process that can be used to align fibers in a biocomposite to mimic the microstructure of long bones is continuous pultrusion (Fig. 9).^[18] The first step in this pultrusion process is the production of a reinforcement supply that usually consists of ceramic fibers impregnated with polymers. These composite fibers are then fed into a heated die, which maintains their orientation while applying compression. Moreover, by altering dimensions of the die (or mold) used in this step, final geometries of the oriented fiber composite can be controlled. Geometries in the final product can be further controlled through the use of pull rollers. Lastly, the finished part is removed after cooling. Pultrusion is an extremely fast and inexpensive technique that has applications from chopped fiber-reinforced thermoplastic composites to bone plates and screws.^[18] Furthermore, continuous pultrusion is a process used for the manufacturing of ceramic fiber-reinforced plastics of constant cross-section, such as structural shapes, beams, channels, pipe, and tubing. Highly oriented parts cut from this stock can then be used in other structures or they can be used alone in such applications as intramedullary rodding or pin fixation of

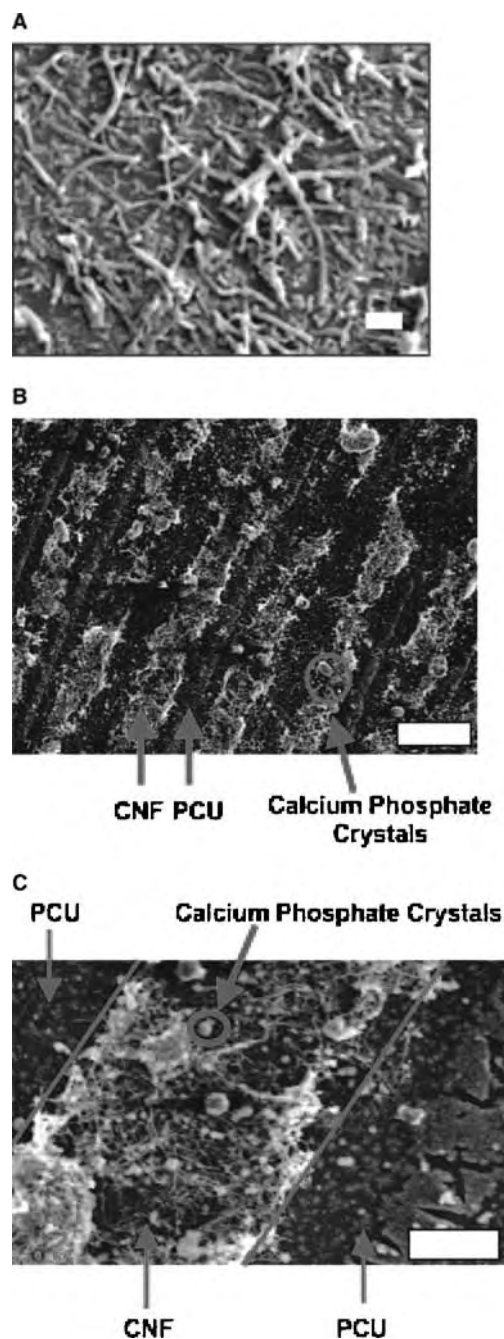


Fig. 8 Micron alignment of CNF for controlling calcium phosphate deposition by bone cells. Directed calcium phosphate deposition by osteoblasts can be achieved by micro-aligning CNF in polymers. Such alignment of calcium phosphate crystals matches the alignment of HA in long bones of the body. Bars = (A) 1 μm ; (B) 20 μm ; and (C) 10 μm .

bone fragments. Clearly, highly aligned ceramic fibers in polymer matrices through continuous pultrusion are attractive to match the anisotropic properties of HA and collagen in bone.

Fiber Meshes/Fiber Bonding

Owing to their similarity to biological entities (like collagen and HA), fibers produced by textile technology have been used to make nonwoven scaffolds of PGA and PLLA.^[16,19] The lack of structural stability of these nonwoven scaffolds often resulted in significant deformation owing to contractile forces from cells seeded on the scaffold. This has led to the development of a fiber-bonding technique to increase the mechanical properties of the scaffolds. Specifically, this is achieved by dissolving PLLA in methylene chloride and casting the polymer over a PGA mesh. The solvent is allowed to evaporate and the construct is then heated above the melting point of PGA. Once the PGA–PLLA construct has cooled, the PLLA is removed by dissolving in methylene chloride again. This treatment results in a mesh of PGA fibers joined at multiple cross-points, which increases mechanical stability. Again, while mostly used for pure polymer matrices, this technique can be easily adapted to polymer/ceramics biocomposites.^[19]

Rapid Prototyping Techniques

Lastly, one of the more exciting biocomposite fabrication techniques is rapid prototyping.^[37–39,41–45] As a common shortcoming of the fabrication technologies discussed above is the lack of automation and resolution when creating three-dimensional porous scaffold architectures, a number of investigators have been actively seeking methods to achieve a high degree of spatial control. To tackle this problem, solid free form fabrication (also known as rapid prototyping) techniques are being developed. The main advantage of these techniques is their ability to produce complex shapes rapidly from a computer-aided design (CAD) model. Thus, the goal here is to have a known bone fractal geometry, which could easily be duplicated in a biocomposite through CAD for use immediately at the surgical table.

Of these rapid prototyping techniques, three-dimensional printing was the first to process biodegradable polymer scaffolds for tissue-engineering applications.^[33] This process generates components by ink-jet printing a binder onto sequential powder layers. The operation parameters such as the speed, flow rate, and drop position can be computer controlled to produce complex three-dimensional polymer scaffolds. Biological agents such as bone growth factors or those that reduce the foreign body response can also be incorporated into the scaffolds in the printing process. However, rapid prototyping techniques have inherent shortcomings such as limited material selection and inadequate resolution. The resolution is determined by

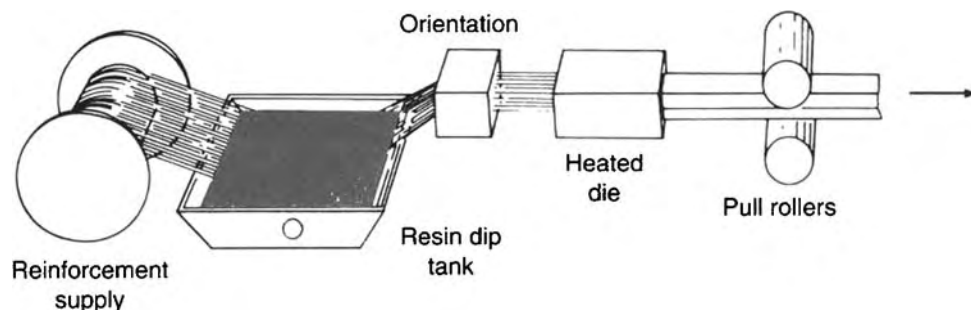


Fig. 9 The pultrusion process for producing aligned fiber-reinforced polymer composites. Fibers impregnated with polymer are fed into a heated die and are then slowly drawn as a cured composite with a constant cross-sectional shape. Compression can be applied in the heated die. (From Ref. [18].)

the size of binder drops as well as the jet and powder particles, which makes it difficult to design and fabricate scaffolds with fine microstructures. The porosity of the scaffolds fabricated with this method is low, and the mechanical properties of the scaffolds have to be significantly improved. Thus, techniques that could easily incorporate ceramic particles would be most welcome to increase mechanical integrity necessary for bone applications. Similarly, other rapid prototyping techniques such as fused deposition modeling (FDM) and stereolithography have been explored for composite fabrication.^[35,36]

Three-dimensional ink-writing techniques can be divided into two approaches: 1) droplet-based or 2) continuous (filamentary) inks. As the name implies, droplet-based ink-writing uses distinct material drops, while continuous (filamentary) ink-writing uses a continuous material stream. For either approach, three-dimensional periodic structures offer the greatest challenge for designing inks because they contain self-supporting (or spanning) features. Inks are typically formulated from colloidal, polymeric, or polyelectrolyte building blocks suspended or dissolved in a liquid or heated to create a stable, homogeneous ink with the desired and reproducible rheological (or flow) behavior. The important rheological parameters for a given ink design include its apparent viscosity, yield stress under shear and compression, and viscoelastic properties, which are tailored for the specific direct-write technique of interest.

Three-dimensional printing (3DP), direct ink-jet printing, and related approaches such as hot-melt printing involve patterning materials using an ink-jet print head, similar to those used in desktop document printing.^[37] These approaches require either low viscosity fluids that must be removed by absorption and evaporation or wax-based inks that are heated during droplet formation and then solidify upon impact cooling. Cima and Sachlos pioneered the concept of using ink-jet printing (3DP) to assemble materials.^[38,39] In 3DP, low viscosity binder droplets are printed onto a powder bed to locally “fuse” material

together in a desired pattern. After defining a given two-dimensional layer, an additional powder layer is spread across the bed surface and subsequently patterned. In other ink-jet approaches, three-dimensional structures (such as high-aspect ratio walls of concentrated nanofiber inks) can be made; however, clogging issues still arise when $D/2a \sim 150$, where D is the finest nozzle used (30 μm in diameter) and a is the radius of largest ink particles. These three-dimensional meso-scale structures may find potential applications as tissue-engineering scaffolds if constructed from a bioactive ceramic material (e.g., HA) or as structural composites if the pore space is filled with a second phase. But clearly, the use of a ceramic phase is still problematic owing to nozzle clogging.

In addition, the resulting constructs have structural heterogeneity because of the “pixel assembly” nature of the rapid prototyping fabrication process. To overcome this shortcoming, a reverse fabrication technique has been developed to form a negative replica of the scaffold.^[42] A polymer/ceramic suspension is cast into such a mold and solidified after the removal of the solvent. The mold is then dissolved away to form the scaffold with the designed three-dimensional pore network. The scaffold is more homogeneous, but the feature resolution is not improved. As will be discussed in the next section, resolutions down to the nanoscale can increase bone regeneration, so this may be problematic and should be addressed in rapid prototyping techniques.

To achieve higher resolution for scaffolds with well-controlled interconnected spherical pores, paraffin spheres have been fabricated by dispersion methods.^[28,46] These paraffin spheres (on the tens of micron size scale) are then transferred into a three-dimensional mold of the desired shape. The spheres are bonded together through a heat treatment process. A composite suspension is then cast into the paraffin assembly in the mold. After removal of the solvent, the paraffin sphere assembly is dissolved away. In this way, an interconnected spherical pore structure is created. Importantly, the features generated (on the hundreds of

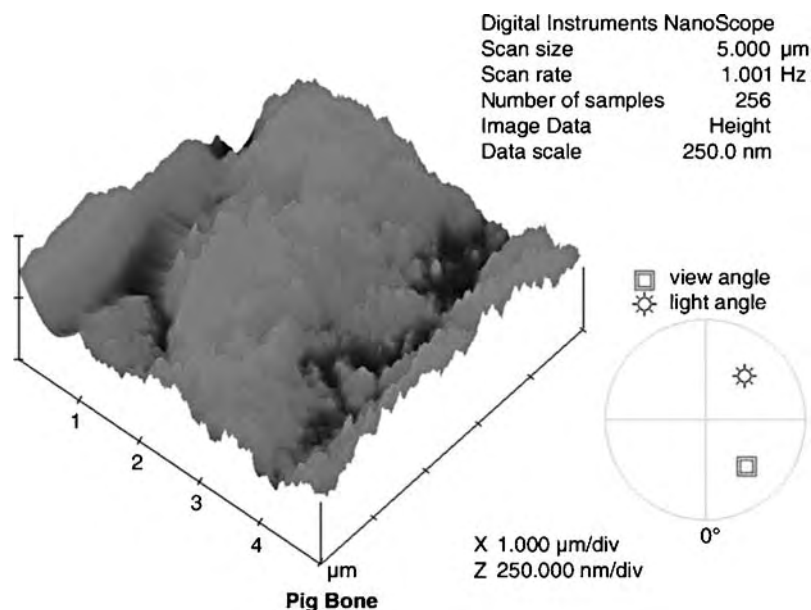


Fig. 10 Atomic force microscopy (AFM) scan illustrating the nanometer roughness of bone. Scan size = 100 nm by 5 μm by 5 μm (Z by X by Y). RMS (root-mean-square surface roughness value) = 33 nm.

nanometer scale) have significantly better resolution than those achieved with current rapid prototyping techniques (tens of microns). In addition, investment in expensive equipment is not required, which allows this technology to be easily adapted to an industrial setting.

Clearly, many adaptations can be made to rapid prototyping techniques for specific outcomes.^[31–46] But, as will be discussed in the next section, those that enable the construction of nanoscale surface features are highly desirable to promote bone growth.

NANOTECHNOLOGY AND BIOCOSITES

One of the biggest advantages to use biocomposites is the ability to match the properties (specifically, mechanical and surface) of natural tissue. That is, if desirable, polymer/ceramic composite scaffolds can be formulated to mimic the constituent components of natural bone (Fig. 10) (naturally present in/on bone). Natural bone is a composite material composed of a polymer matrix (organic compounds, mainly collagen) reinforced with nanometer ceramic particles (inorganic compound, mainly carbonated HA). Moreover, the collagen fibers and HA crystals of natural bone are within the nanometer scale proving for a unique heterogeneous nanosurface roughness to bone cells. Therefore, bone itself is a nanostructured composite composed of nanometer size HA well dispersed in a mostly collagen matrix.

Nanotechnology is the creation and use of materials with dimensions less than 100 nm in at least one direction. For this reason, through the use of nanotechnology, it may be possible to create biocomposites that more accurately mimic the nanoarchitecture of

natural tissues. Indeed, recent research in the biocomposite field has suggested better osteoconductivity on synthetic materials that incorporate nanoscale particles.^[21,23–26,47–50] Compared to respective micron-sized particles, this has been demonstrated for polymers that combine nanoparticles of ceramics (alumina, titania, HA, zinc oxide, etc.), metals (CoCrMo, c.p. Ti, Ti6Al4V, etc.), or CNF/nanotubes.^[23,24,47–50] For example, five times more bone growth was measured on PLGA, which contained nanometer compared to micron-sized titania particles.^[49] In all material cases, increased bone synthesis was attributed to the presence of nanoscale particles that provided for subsequent biologically inspired nanometer roughness to bone cells.

Specifically, when determining the mechanism of increased bone synthesis on nanoceramic biocomposites, studies have demonstrated significantly increased select protein adsorption (fibronectin and vitronectin) that led to promoted osteoblast function on the nano biocomposites compared to traditional micron-sized ceramics in polymers.^[23–25] It was further determined that the reason for altered protein interactions was the unique surface energetics of the biocomposites provided by the nanoparticles. It has been well established that nanomaterials have increased percentages of atoms at their surface, greater defects at their surface, and altered electron distributions;^[23–25] all of which contribute to a greater surface reactivity of nanomaterials compared to conventional materials.

Clearly, scaffold materials with surface properties similar to physiological bone (characterized by surface grain sizes in the nanometer regime) would aid in the formation of new bone at the tissue/biomaterial interface.^[26] Mechanical properties (such as flexural rigidity) are also much closer to natural bones when

nanoceramics as opposed to conventional ceramics are added to polymers.^[25] Although further studies are needed, there could be a significant advantage gained in biocomposite efficacy if nanotechnology is employed.

CONCLUSIONS

While numerous advances have been made toward the use of biocomposites in tissue engineering, much more research is still needed. This article highlighted some advances that have been made toward material selection and fabrication techniques to formulate biocomposites more similar to properties of natural tissue. However, before the field can take advantage of the ability of biocomposites to mimic natural tissues, the following research areas need to be expanded:

1. Improved fabrication techniques for polymer/ceramic composites, emphasizing resolution and preciseness of three-dimensional structures.
2. Improved mechanical properties closer to natural tissue, such as strength and elastic modulus of polymer/ceramic composites.
3. Improved cytocompatibility, emphasizing nano-scale surface properties that favor cell growth and consequently improves the composite-tissue interface.
4. Improved understanding of the degradation behavior of biodegradable polymer and bioactive ceramic composites, especially the effect of the ceramic phase on degradation kinetics and degradation products of the polymer.
5. Understanding of how to most effectively incorporate drugs and growth factors into biocomposite scaffolds for proper diffusion to further promote tissue growth.

In conclusion, we will not be able to fully comprehend the advantages biocomposites can make to the field of tissue engineering until we understand all of the intricate complexities involved when two or more different types of materials are combined. It is clear, though, that as our tissues are composites themselves arranged in a well-controlled manner, we need to develop materials that match this diversity. The perfect biocomposite materials will not be obtained until this is accomplished.

ARTICLES OF FURTHER INTEREST

Bioactive Glass; Biofunctional Polymers; Biomimetic Materials

REFERENCES

1. Laurencin, C.T.; El-Amin, S.F.; Ibim, S.E.; Willoughby, D.A.; Attawia, M.; Allcock, H.R.; Ambrosio, A.A. A highly porous 3-dimensional polyphosphazene polymer matrix for skeletal tissue regeneration. *J. Biomed. Mater. Res.* **1996**, *30*, 133–138.
2. McManus, A.J.; Doremus, R.H.; Siegel, R.W.; Bizios, R. Evaluation of cytocompatibility and bending modulus of nanoceramic/polymer composites. *J. Biomed. Mater. Res.* **2005**, *72A*, 98–106.
3. Marieb, E.N. *Human Anatomy and Physiology*; Benjamin/Cummings Pub Co.: Redwood City, CA, 1989.
4. Ramakrishna, S.; Huang, Z.; Kumar, G.V.; Batchelow, A.W.; Huang, S. *An Introduction to Biocomposites*; World Scientific Publishers Co., Inc.: Singapore, 2004.
5. Ma, P.X.; Langer, R. Degradation, structure and properties of fibrous nonwoven poly(glycolic acid) scaffolds for tissue engineering. *Mat. Res. Soc. Symposium—Proceedings. Polym. Med. Pharm.* **1995**, *394*, 99–104.
6. Zhang, R.; Ma, P.X. Degradation behavior of porous poly(α -hydroxy acids)/hydroxyapatite composite scaffolds. In *Polymer Preprints*; Division of Polymer Chemistry Inc. and American Chemical Society, 2000; Vol. 41 (2), 1618–1619.
7. Taddei, P.; Monti, P.; Simoni, R. Vibrational and thermal study on the in vitro and in vivo degradation of a poly(lactic acid)-based bioabsorbable periodontal membrane. *J. Mat. Sci. Mat. Med.* **2002**, *13* (5), 469–475.
8. Pitt, C.G.; Gratzl, M.M.; Kimmel, G.L.; Surles, J.; Sohndler, A. Aliphatic polyesters II. The degradation of poly(DL-lactide), poly(ϵ -caprolactone), and their copolymers in vivo. *Biomaterials* **1981**, *2*, 215–220.
9. Lin, W. Comparison of thermal characteristics and degradation properties of ϵ -caprolactone copolymers. *J. Biomed. Mater. Res.* **1999**, *47*, 420–423.
10. Temenoff, J.S.; Mikos, A.G. Injectable biodegradable materials for orthopedic tissue engineering. *Biomaterials* **2000**, *21*, 2405–2412.
11. LeGeros, R.Z.; Lin, S.; Rohanzadeh, R.; Mijares, D.; LeGeros, J.P. Biphasic calcium phosphate bioceramics: preparation, properties and applications. *J. Mat. Sci. Mat. Med.* **2003**, *14*, 201–209.
12. Klein, C.P.A.T.; Driessen, A.A.; de Groot, K.; van den Hooff, A. Biodegradation behavior of various calcium phosphate materials in bone tissue. *J. Biomed. Mater. Res.* **1983**, *17*, 769–784.
13. Boccaccini, A.R.; Maquet, V. Bioresorbable and bioactive polymer/bioglass[®] composites with tailored pore structure for tissue engineering applications. *Composites Sci. Technol.* **2003**, *63*, 2417–2429.
14. Black, J.; Hastings, G. *Handbook of Biomaterial Properties*; Chapman & Hall: London, New York, 1998.
15. Ma, P.X. Scaffolds for tissue fabrication. *Mat. Today* **2004**, *7*, 30–40.
16. Hutmacher, D.W. Scaffolds in tissue engineering bone and cartilage. *Biomaterials* **2000**, *21*, 2529–2543.
17. Fung, Y.C. *Biomechanics: Mechanical Properties of Living Tissues*; Springer-Verlag: New York, 1993.
18. Alexander, H. Classes of materials used in medicine. In *Biomaterials Science: An Introduction to Materials in*

- Medicine*; Ratner, B.D., Hoffman, A., Schoen, F., Lemmons, J., Eds.; Academic Press: New York, 18–24.
19. Mikos, A.G. *Biomaterials for Drug and Cell Delivery*; Materials Research Society: Pittsburgh, PA, 1994.
 20. Liu, L.; Zhang, L.; Ren, B.; Wang, F.; Zhang, Q. Preparation and Characterization of collagen-hydroxyapatite composite used for bone tissue engineering scaffolds. *Artif. Cells Blood Substitutes Biotechnol.* **2003**, *31* (4), 435–448.
 21. Du, C.; Cui, F.Z.; Zhu, X.D.; de Groot, K. Three-dimensional Nano-HAP/collagen matrix loading with osteogenic cells in organ culture. *J. Biomed. Mater. Res.* **1999**, *44* (4), 407–415.
 22. Rho, J.Y.; Kuhn-Spearing, L.; Zioupos, P. Mechanical properties and the hierarchical structure of bone. *Med. Eng. Phys.* **1998**, *20*, 92–102.
 23. Webster, T.J.; Ergun, C.; Doremus, R.H.; Siegel, R.W.; Bizios, R. Specific proteins mediate enhanced osteoblast adhesion on nanophase ceramics. *J. Biomed. Mater. Res.* **2000**, *51* (3), 475–483.
 24. Webster, T.J.; Siegel, R.W.; Bizios, R. Osteoblast adhesion on nanophase ceramics. *Biomaterials* **1999**, *20* (13), 1221–1227.
 25. Dulgar Tulloch, A.J.; Bizios, R.; Siegel, R.W. Nanophase alumina/poly(L-lactic acid) composite scaffolds for biomedical applications. *Mat. Res. Soc. Symp. Proc.* **2003**, *740*, 161–166.
 26. Webster, T.J.; Siegel, R.W.; Bizios, R. Nanoceramic surface roughness enhances osteoblast and osteoclast functions for improved orthopaedic/dental implant efficacy. *Scr. Materialia* **2001**, *44*, 1639–1642.
 27. Liu, H.; Slamovich, E.B.; Webster, T.J. Increased osteoblast functions on poly-lactic-co-glycolic-acid with highly dispersed nanophase titania. *J. Biomed. Nanotechnol.* **2005**, *1*, 83–89.
 28. Tsang, V.L.; Bhatia, S.N. Three-dimensional tissue fabrication. *Adv. Drug Deliv. Rev.* **2004**, *56*, 1635–1647.
 29. Borenstein, J.T.; Terai, H.; King, K.R.; Weinberg, E.J.; Kaazempur-Mofrad, M.R.; Vacanti, J.P. Micro-fabrication technology for vascularized tissue engineering. *Biomed. Microdevices* **2002**, *4*, 167–175.
 30. Vozzi, G.; Flaim, C.; Ahluwalia, A.; Bhatia, S.N. Fabrication of PLGA scaffolds using soft lithography and microsyringe decomposition. *Biomaterials* **2003**, *24*, 2533–2540.
 31. Leong, K.F.; Cheah, C.M.; Chua, C.K. Solid freeform fabrication of three-dimensional scaffolds for engineering replacement tissues and organs. *Biomaterials* **2003**, *24*, 2363–2378.
 32. Yang, S.; Leong, K.F.; Du, Z.; Chua, C.K. The design of scaffolds for use in tissue engineering: part II. Rapid prototyping techniques. *Tissue Eng.* **2002**, *8*, 1–11.
 33. Lee, G.; Barlow, J.; Fox, W.; Aufdermorte, T. Biocompatibility of SLS-formed calcium phosphate implants. *Proceedings of Solid Freeform Fabrication Symposium, University of Texas: Austin, TX, 1996*; 15–22.
 34. Hutmacher, D.W. Scaffold design and fabrication technologies for engineering tissues—state of the art and future perspectives. *J. Biomater. Sci. Polym. Ed.* **2001**, *12*, 107–124.
 35. Zein, I.; Hutmacher, D.W.; Tan, K.C.; Teoh, S.H. Fused deposition modeling of novel scaffold architectures for tissue engineering applications. *Biomaterials* **2002**, *23*, 1169–1185.
 36. Cooke, M.N.; Fisher, J.P.; Dean, D.; Rinnac, C.; Mikos, A.G. Use of stereolithography to manufacture critical-sized 3D Bio-degradable scaffolds for bone ingrowth. *J. Biomed. Mater. Res.* **2003**, *64B*, 65–69.
 37. Park, A.; Wu, B.; Griffith, L.G. Integration of surface modification and 3D fabrication techniques to prepare patterned poly(L-lactide) substrates allowing regionally selective cell adhesion. *J. Biomater. Sci. Polym. Ed.* **1998**, *9*, 89–110.
 38. Kim, S.S.; Utsunomiya, H.; Koski, J.A.; Wu, B.M.; Cima, M.J.; Sohn, J.; Mukai, K.; Griffith, L.G.; Vacanti, J.P. Survival and function of hepatocytes on a novel three-dimensional synthetic biodegradable polymer scaffold with an intrinsic network of channels. *Ann. Surg.* **1998**, *228*, 8–13.
 39. Zeltinger, J.; Sherwood, J.K.; Graham, D.A.; Mueller, R.; Griffith, L.G. Effect of pore size and void fraction on cellular adhesion, proliferation, and matrix deposition. *Tissue Eng.* **2001**, *7*, 557–572.
 40. Sachlos, E.; Reis, N.; Ainsley, C.; Derby, B.; Czernuszka, J.T. Novel collagen scaffolds with predefined internal morphology made by solid freeform fabrication. *Biomaterials* **2003**, *24*, 1487–1497.
 41. . EnvisionTec. www.envisionstec.de.
 42. Mironov, V.; Boland, T.; Trusk, T.; Forgacs, G.; Markwald, R.R. Organ printing: computer-aided jet-based 3D tissue engineering. *Trends Biotechnol.* **2003**, *21*, 157–161.
 43. Sciperio, I. www.sciperio.com.
 44. Odde, D.J.; Renn, M.J. Laser-guided direct writing of living cells. *Biotechnol. Bioeng.* **2000**, *67*, 312–318.
 45. Albrecht, D.R.; Sah, R.L.; Bhatia, S.N. Dielectrophoretic cell patterning within tissue engineering scaffolds. In *Second Joint EMBS-BMES Conference. 24th Annual International Conference of the Engineering in Medicine and Biology Society. Annual Fall Meeting of the Biomedical Engineering Society, IEEE, Houston, TX, 2002, Cat. No. 02CH37392, Vol. 2, 1708–1709.*
 46. Landers, R.; Hubner, U.; Schmelzeisen, R.; Mulhaupt, R. Rapid prototyping of scaffolds derived from thermoreversible hydrogels and tailored for applications in tissue engineering. *Biomaterials* **2002**, *23*, 4437–4447.
 47. Webster, T.J.; Smith, T.A. Increased osteoblast function on PLGA composites containing nanophase titania. *J. Biomed. Mater. Res.* **2005**, *74*, 677–686.
 48. Palin, E.; Liu, H.; Webster, T.J. Mimicking the nanofeatures of bone increases bone-forming cell adhesion and proliferation. *Nanotechnology* **2005**, *16*, 1828–1835.
 49. Swaim, J.; Smith, L.; Webster, T.J. Increased osteoblast functions on PLGA containing HA treated with NaOH. *Biomaterials*, . *in press*.
 50. Khang, D.; Webster, T.J. Aligned calcium phosphate in carbon nanofibers aligned in polycrystal. *International Journal of Nanomedicine* **2006**, *1* (1), 65–72.

Biodegradable Polymers: An Overview

Chih-Chang Chu

Fiber & Polymer Science Program, Department of Fiber Science & Apparel & Biomedical Engineering Program, Cornell University, Ithaca, New York, U.S.A.

B

INTRODUCTION

Interests in biodegradable polymers increased dramatically during the past decades as a result of their successful clinical applications in a wide range of biomedical areas. This success is because of the fact that this class of biomaterials has two major advantages that non-biodegradable biomaterials do not have. First, they do not elicit permanent chronic foreign-body reaction because they would be gradually absorbed by the human body and do not permanently retain traces of residual in the implantation sites. Second, some were recently found to be able to regenerate tissues, so-called tissue engineering, through the interaction of their biodegradation with immunologic cells such as macrophages. Hence, surgical implants made from biodegradable biomaterials could be used as a temporary scaffold for tissue regeneration. The approach of using biodegradable polymers as a temporary scaffold either to grow cells/tissues in vitro for tissue engineering application or to regenerate tissues in vivo has very recently become an important aspect of research and development that broadens biodegradable polymers beyond their traditional use in wound closure and drug control/release biomaterials. The scaffolds used in either tissue engineering or regeneration are to provide support for cellular attachment and subsequent controlled proliferation into predefined shape or form. A biodegradable scaffold is preferred because of the elimination of chronic foreign-body reaction and the generation of additional volume for regenerated tissues.

The earliest and most commercially significant biodegradable polymeric biomaterials originated from linear aliphatic polyesters such as polyglycolide and polylactide from poly(α -hydroxyacetic acids). Recent introductions of several new synthetic and natural biodegradable polymeric biomaterials extend the domain beyond this family of simple polyesters. These relatively newer, commercially significant biodegradable polymeric biomaterials include poly(orthoesters), polyanhydrides, polysaccharides, poly(ester amides), tyrosine-based polyarylates or polyiminocarbonates or polycarbonates, poly(D,L-lactide-urethane), poly(β -hydroxybutyrate), poly(ϵ -caprolactone), poly[bis-(carboxylatophenoxy)phosphazene], poly(amino acids), pseudo-poly(amino acids), and copolymers derived

from amino acids and non-amino acids. These commercially important biodegradable polymeric biomaterials are frequently used in one of the following physical forms: fibers, fabrics, gels, microspheres, or films.

The earliest and most successful and frequent biomedical application of biodegradable polymeric biomaterials was in wound closure.^[1] All biodegradable wound closure biomaterials are based on the glycolide and lactide family. This family of biodegradable polymeric biomaterials is also the most studied in chemical, physical, mechanical, morphological, and biological properties. Some of the above materials have been commercially used as surgical meshes for hernia and body-wall repair. Other important biomedical applications of biodegradable polymeric biomaterials that are commercially satisfactory are drug control/release devices and orthopedic fasten devices or implants and dental reconstruction. Biodegradable polymeric biomaterials, particularly totally resorbable composites, have also been used in the field of orthopedics, mainly as components for internal bone fracture fixation devices like bone screws, plates, and pins. Beside the commercial uses described above, biodegradable polymeric biomaterials have been tested as vascular grafts, vascular stents, vascular couplers for vessel anastomosis, nerve growth conduits, augmentations of defective bone, ligament/tendon prostheses, intramedullary plugs during total hip replacement, anastomosis rings for intestinal surgery, and stents in ureteroureterostomies for accurate suture placement.

This chapter provides a brief overview of the commercially most significant and successful biomedical biodegradable polymers based on linear aliphatic polyesters, the effects of some new extrinsic factors on the degradation of the commercially most significant biodegradable polymers, a new theoretical approach to model the hydrolytic degradation of glycolide/lactide-based biodegradable polymers, some of the latest research and development in important classes of synthetic biodegradable polymers outside the conventional aliphatic polyester family. In this chapter, the term “biodegradable” is loosely associated with absorbable, erodable, and resorbable, and refers to materials that could be broken down by nature either through pure hydrolytic or enzyme-catalyzed

mechanisms. The details of the applications of this family and other classes of biodegradable polymeric biomaterials and their chemical, physical, mechanical, biological, and biodegradation properties exist in many other reviews.^[2–13]

Commercially Important Biodegradable Polymers from Linear Aliphatic Polyesters

Biodegradable polymers from linear aliphatic polyesters are the most successful, important, and commercially widely used biodegradable biomaterials in both biomedicine and surgery and were the most extensively studied in terms of degradation mechanisms and structure–property relationship. Among them, polyglycolide or polyglycolic acid (PGA), poly-L-lactide (PLLA), and poly- ϵ -caprolactone (PCL) are the most important building blocks because most other biodegradable polymers are based on these three types of homo- and copolymers, e.g., poly(glycolide-L-lactide) copolymer, poly(glycolide-L-lactide- ϵ -caprolactone) copolymer (Caprosyn™), poly(glycolide-L-lactide-trimethylene carbonate) copolymer (Gore Resolut®), poly(glycolide- ϵ -caprolactone-trimethylene carbonate) copolymers, or modified glycolide monomer, e.g., poly-*p*-dioxanone. Poly(hydroxyalkanoate) (PHA) is an unusual family member of aliphatic polyesters and is mainly derived from genetically engineered microorganisms such as *Escherichia coli*. PHA will be described later. Table 1 illustrates important properties of a few selected commercially important synthetic biodegradable polymers used as biomaterials.

Glycolide, Lactide, and ϵ -Caprolactone-Based Polymers

PGA is the first synthetic biodegradable polymer synthesized for biomaterial application and also the simplest linear aliphatic polyester. For biomedical applications, stannous chloride dihydrate or trialkyl-aluminum is the preferred catalyst for ring-opening polymerization. PGA exhibited an orthorhombic unit cell with dimensions $a = 5.22 \text{ \AA}$, $b = 6.19 \text{ \AA}$, and c (fiber axis) = 7.02 \AA . The planar zigzag-chain molecules form a sheet structure parallel to the ac plane and do not have the polyethylene type arrangement, and the tight molecular packing and the close approach of the ester groups might stabilize the crystal lattice and contribute to the high melting point, T_m ($224\text{--}230^\circ\text{C}$). The glass transition temperature, T_g , ranges from 36 to 40°C . The specific gravities of PGA are 1.707 for a perfect crystal and 1.50 in a completely amorphous state. The heat of fusion of 100% crystallized PGA is 12 KJ/mol (45.7 cal/g). The four Fourier transform infrared spectroscopy bands at 850 , 753 , 713 , and 560 cm^{-1} are associated with the amorphous regions of the PGA and could be used to assess the extent of hydrolysis. Peaks associated with the crystalline phase include those at 972 , 901 , 806 , 627 , and 590 cm^{-1} .

Other commercially successful glycolide-based biodegradable polymeric biomaterials are the copolymers of glycolide with other monomers within linear aliphatic polyesters such as lactides, carbonates, and ϵ -caprolactone. The glycolide–lactide random copolymers are the most studied and have a wide range of properties and applications, depending on the composition ratio of glycolide to lactide. For wound

Table 1 Properties of commercially important synthetic biodegradable polymers

Polymer	Crystallinity	T_m ($^\circ\text{C}$)	T_g ($^\circ\text{C}$)	T_{dec} ($^\circ\text{C}$)	Fiber		
					Strength (MPa)	Modulus (GPa)	Elongation (%)
PGA	High	230	36	260	890	8.4	30
PLLA	High	170	56	240	900	8.5	25
PDLLA	None	—	57	—	—	—	—
Polyglactin 910 ^a	High ^c	200	40	250	850	8.6	24
Polydioxanone	High	106	<20	190	490	2.1	35
Polyglyconate ^b	High ^c	213	<20	260	550	2.4	45
Poliglecaprone 25 ^d	—	<220	$-36 \sim 15$	—	91,100 ^e	113,000 ^e	39

^aGlycolide per lactide = 9/1.

^bGlycolide per trimethylene carbonate = 9/1.

^cDepending on the copolymer composition.

^d2/O size Monocryl (glycolide- ϵ -caprolactone copolymer).

^ePSI unit.

Source: Kimura, Y., 1993. "Biodegradable Polymers," In: *Biomedical Applications of Polymeric Materials*, Ed. T. Tsuruta, T. Hayashi, K. Kataoka, K. Ishihara, & Y. Kimura, CRC Press, Boca Raton, Florida, and Chu, C.C., von Fraunhofer, J.A., and Greisler, H.P., 1996. *Wound Closure Biomaterials and Devices*. CRC Press, Boca Raton, Florida.

closure application, a high concentration of glycolide monomer is required for achieving proper mechanical and degradation properties. Vicryl[®] sutures, sometime called polyglactin 910, contain a 90/10 molar ratio of glycolic to L-lactide. This molar ratio is important for the Vicryl suture to retain crystalline characteristics. Contrary to the conventional high PGA content in the glycolide–lactide copolymers, Ethicon introduced a copolymer of PGA and PLLA with 5-to-95 molar ratio (Panacryl[®]). Because of the high L-lactide component in the copolymer, Panacryl biodegrades much more slowly than Vicryl. Panacryl was recently pulled from the biomedical market. If DL- instead of L-lactide is used as the comonomer, the U-shape relationship between the level of crystallinity and glycolide composition disappears. This is because polylactide from 100% DL-lactide composition is totally amorphous. Infrared (IR) bands associated with Vicryl molecules in the amorphous domains are 560, 710, 850, and 888 cm⁻¹, whereas 590, 626, 808, 900, and 972 cm⁻¹ are associated with the crystalline domains. Like PGA, these IR bands could be used to assess the extent of hydrolysis.

A block copolymer of glycolide and carbonates, such as trimethylene carbonate, has been commercialized as Maxon[®] suture and is made from a block copolymer of glycolide and 1,3-dioxan-2-one (trimethylene carbonate GTMC) and consists of 32.5% by weight (or 36 mol %) trimethylene carbonate. The weight ratio of glycolide to trimethylene carbonate in the middle block is 15:85, and the two ends of this middle block are capped by additional glycolide monomers to form the final triblock copolymer. Maxon is a poly(ester carbonate). The latest glycolide-based copolymer to become commercially successful is Monocryl[®] suture. Monocryl is a segmented block copolymer consisting of both soft and hard segments. The purpose of having soft segments in the copolymer is to provide good handling properties such as pliability, whereas the hard segments provide adequate strength. The soft segment consists of a prepolymer of glycolide and ϵ -caprolactone; like Maxon, the two ends of the prepolymer are capped by polyglycolide, which act as hard segments. Monocryl has a composition of 75% glycolide and 25% ϵ -caprolactone for adequate mechanical property required by sutures. The most unique aspect of Monocryl monofilament suture is its pliability, as claimed by Ethicon, and the force required to bend a 2/O suture is only about 2.8×10^4 lb-in² for Monocryl, whereas the same size PDSII and Maxon monofilament sutures require about 3.9 and 11.6×10^4 lb in² force, respectively. This inherent pliability of Monocryl is because of the presence of soft segments and T_g resulting from the ϵ -caprolactone comonomer unit. The T_g of Monocryl is expected to be between 15 and -36°C .

Poly(*p*-dioxanone) (PDS[®]) is derived from the glycolide family for better flexibility. It is polymerized from ether-containing lactones, 1,4-dioxane-2,5-dione (i.e., *p*-dioxanone) monomers with a hydroxylic initiator, and tin catalyst. The resulting polymer is semicrystalline, with T_m about 106 – 115°C and T_g -10 to 0°C . The improved flexibility of PDS relative to PGA as evidenced in its lower T_g is a result of the incorporation of the ether segment in the repeating unit, which reduces the density of ester linkages for intermolecular hydrogen bonds. Because of the less dense ester linkages in PDS compared with PGA or glycolide-L-lactide copolymers, PDS is expected and has been shown to degrade at a slower rate in vitro and in vivo. Recently, an advanced version of PDS, PDSII, was introduced. PDSII was achieved by subjecting the melt-spun fibers to a high temperature (128°C) for a short period of time. This additional treatment partially melts the outermost surface layer of PDS fibers and leads to a distinctive skin-core morphology. The heat employed also results in larger crystallites in the core of the fiber than in the untreated PDS fiber. The tensile strength loss profile of PDSII sutures is better than that of PDS sutures. A variety of copolymers with high molar ratios of PDS to other monomers within the same linear aliphatic polyester family were reported for the purpose of improving mechanical and biodegradation properties. For example, the copolymer of PDS (80%) and PGA (up to 20%) has an absorption profile similar to that of Dexon[®] and Vicryl sutures, but it has compliance similar to that of PDS. The copolymer of PDS (85%) and PLLA (up to 15%) results in a more compliant (low modulus) suture than homopolymer PDS but with similar absorption profiles as PDS.

Copolymer fibers made from PDS and monomers other than linear aliphatic polyester, such as morpholine-2,5-dione (MD), exhibit interesting biodegradation properties. The copolymer fibers were absorbed 10–25% earlier than PDS. However, the copolymer retained a tensile breaking strength profile similar to that of PDS, with a slightly faster strength loss during the earlier stage, i.e., the first 14 days. This ability to break the inherent fiber structure–property relationship through copolymerization is a major improvement in the biodegradation property of absorbable sutures. It is interesting to recognize that a small percentage (3%) of MD in the copolymer suture is sufficient to result in a faster mass loss profile without the expense of its tensile strength loss profile. The ability to achieve this ideal biodegradation property might be attributed to both an increasing hydrophilicity of the copolymer and the disruption of crystalline domains as a result of MD moieties. As described later, the loss of suture mass is mainly a result of the destruction of crystalline domains, whereas the loss of tensile breaking strength

is chiefly a result of the scission of tie-chain segments located in the amorphous domains. The question is, Why does MD-PDS copolymeric suture retain its strength loss similar to PDS? The possible explanation is that the amide functional groups in MD could form stronger intermolecular hydrogen bonds than ester functional groups. This stronger hydrogen bond contributes to the strength retention of the copolymer of PDS and MD during *in vivo* biodegradation.

To improve γ -irradiation stability of linear aliphatic polyesters, radiostabilizers like poly(ethylene 1,4-phenylene-bis-oxyacetate) (PEPBO) have been copolymerized with glycolide and PDS to form segmented copolymers. Because of the well-known adverse effect of γ -irradiation on the mechanical properties of aliphatic polyester-based synthetic biodegradable polymers, they are generally sterilized by the tedious ethylene oxide process. A great need exists to develop γ -irradiation sterilizable synthetic biodegradable polymers to take advantage of the highly convenient and reliable method of sterilization. The incorporation of about 10 mol % of a polymeric radiostabilizer like PEPBO into PGA backbone chains would make the copolymer sterilizable by γ -irradiation without a significant accelerated loss of mechanical properties upon hydrolysis compared with the unirradiated copolymer control (morpholine-2,5-dione glycolide, or MPG). The changes in tensile breaking force of both MPG and PGA sutures implanted intramuscularly and subcutaneously in rats for various periods demonstrate the great advantage of such copolymers. MPG fibers γ -irradiated at 2.89 Mrad did not demonstrate any loss in tensile breaking force during the first 14 days post-implantation compared with unimplanted samples. On the contrary, PGA sutures γ -irradiated at 2.75 Mrad lost 62% of their tensile breaking force of their unimplanted samples. No tensile breaking force remained for the irradiated PGA at the end of 21 days, whereas both 2.89 and 5 Mrad-irradiated MPG retained 72% and 55% of their corresponding 0 day controls, respectively. The inherent, more hydrolytic resistance of MPG must be attributed to the presence of an aromatic group in the backbone chains, which is also responsible for the observed γ -irradiation stability. The incorporation of 5–10% of such stabilizer in PDS not only improved γ -irradiation resistance considerably, but also increased the compliance of the material. For example, PEPBO-PDS copolymer retained 79%, 72%, and 57% of its original tensile breaking strength at 2, 3, and 4 weeks *in vivo* implantation, whereas PDS homopolymer retained only 43%, 30%, and 25% at the corresponding periods. It appeared that an increasing (CH₂) group between the two ester functional groups of the radiation stabilizers improved copolymer resistance toward γ -irradiation. PEPBO was also incorporated

into glycolide via copolymerization to achieve γ -irradiation capability.

Poly lactides, particularly PLLA, and copolymers having > 50% L- or DL- lactide have been explored for medical use without much success, mainly because of their much slower absorption and difficulty in melt processing. PLLA is a semicrystalline polymer with $T_m = 170^\circ\text{C}$ and $T_g = 56^\circ\text{C}$. This high T_g is mainly responsible for the observed extremely slow biodegradation rate at body temperature. The molecular weight of lactide-based biodegradable polymers suitable for medical use ranges from 1.5 to 5.0 dl/g inherent viscosity in chloroform. Ultrahigh-molecular-weight polylactide has been reported and an intrinsic viscosity as high as 13 dl/g was reported. High-strength PLLA fibers from this ultrahigh-molecular-weight polylactide was made by hot-drawing fibers from solutions of good solvents. The resulting fibers had tensile breaking strength close to 1.2 GPa. Because of the dissymmetric nature of lactic acid, however, the polymer made from optically inactive racemic mixture of D and L enantiomers, poly-DL-lactide is an amorphous polymer.

Lactide-based copolymers with a high percentage of lactide have recently been reported, particularly those copolymerized with aliphatic polycarbonates such as trimethylene carbonate (TMC) or 3,3-dimethyltrimethylene carbonate (DMTMC). The major advantage of incorporating TMC or DMTMC units into lactide is that the degradation products from TMC or DMTMC are largely neutral pH and, hence, considered advantageous. Glycolic- or lactic-acid-rich degradation products have the potential to significantly lower the local pH in closed and less body-fluid-buffered regions surrounded by bone. This is particularly true if the degradation process proceeds with a burst mode (i.e., a sudden and rapid release of degradation products). This acidity tends to cause abnormal bone resorption or demineralization. The resulting environment may be cytotoxic. The increase in local acidity as a result of a faster accumulation of the highly acidic degradation products also leads to an accelerated acid-catalyzed hydrolysis in the immediate vicinity of the biodegradable device. This acceleration in hydrolysis could lead to a faster loss of mechanical property of the device than expected. Copolymers of composition ratio of 10 DMTMC/90 LLA (L-lactide) or 10 TMC/90 LLA appear to be a promising absorbable orthopedic device. Other applications of this type of copolymer include nerve growth conduits, tendon prostheses, and coating materials for biodegradable devices.

Although the copolymer of L-lactide with glycolide (e.g., Vicryl) resulted in commercial success, the copolymer of L-lactide with other aliphatic polyesters such as ϵ -caprolactone is not as well known. Recently,

a copolymer of L-lactide- ϵ -caprolactone and acrylic acid was reported.^[14] The most unique aspect of this new L-lactide- ϵ -caprolactone-acrylic acid copolymer is that the free pendant carboxylic acid group along the acrylic acid backbone was used to attach biologically active species, such as tempamine nitroxyl radicals.

Another unique example of L-lactide copolymer is the copolymer of L-lactide and 3-(S)[(alkyloxy-carbonyl) methyl]-1,4-dioxane-2,5-dione, a cyclic diester. The most unique aspect of this new biodegradable copolymer is the carboxyl acid pendant group which, makes the new polymer not only more hydrophilic, resulting in faster biodegradation, but also more reactive toward future chemical modification through the pendant carboxyl group. These carboxyl-reactive pendant sites could be used to chemically bond antimicrobial agents or other biochemicals such as growth factors, creating future wound closure biomaterials with new and important biological functions. Unfortunately, no reported data evaluate the performance of these new absorbable polymers for biomedical engineering use up to the present time.

Block copolymers of PLLA with poly(amino acids) were also reported as a potential controlled drug delivery system. This new class of copolymers consists of two distinctive blocks in the backbone, the block with only amide linkage and the block with only ester linkages, and is sometimes referred as poly(depsipeptides).^[4] As described later, a new family of poly(ester amide)s (PEA) has both ester and amide linkages within the same repeating unit, instead of being located at two different blocks as above.

The introduction of poly(ethylene oxide) (PEO) into PLLA in order to modulate the hydrophilicity and degradability of PLLA for drug control/release biomaterials has been reported. An example is the triblock copolymer of PLLA/PEO/PLLA. Biomaterials with an appropriate PLLA and PEO block length had hydrogel properties that could deliver hydrophilic drugs as well as hydrophobic ones such as steroids and hormones. Another unique biodegradable biomaterial consisting of a star-block copolymer of PLLA, PGA, and PEO was also reported for protein drug delivery devices. This star-shape copolymer has four or eight arms made of PEO, PLLA, and PGA. The T_g and the crystallinity of this star-shape block copolymer were significantly lower than that of the corresponding linear PLLA and PGA.

Because of the slow biodegradation rate of PLLA and because the copolymers have a high composition ratio of PLLA, their biomedical applications have been limited to mainly orthopedic and dental surgery, drug control/release devices, coating materials for suture, vascular grafts, and surgical meshes to facilitate wound healing after dental extraction.

In addition to glycolide- and lactide-based biodegradable polymers as biomaterials, ϵ -caprolactone-

based biodegradable polymers, particularly pure PCL homopolymer, have not enjoyed the same level of commercial success as their siblings, probably because of their slow biodegradation rate and relatively lower thermal and mechanical properties. Because of its low T_m of 57°C, it is difficult to thermally process poly- ϵ -caprolactone, and its low T_g , near -60°C, would make it difficult for tissue repairs that require high mechanical strength and rigidity. Because of the relative disadvantages of pure PCL homopolymer as biomaterials, research and development efforts have been geared toward the formation of either copolymers as the case with L-lactide or multiarms and gel network structure. The two most recent examples of the latter approach are the synthesis of three-arm star polymers having PCL arms with functionalized arm ends^[15] and the four-arm network structure with gel characteristics.^[16]

Non-glycolide-, lactide-, and ϵ -caprolactone-based linear aliphatic polyesters

The most unique family of biodegradable polymers outside glycolide, lactide, and ϵ -caprolactone is from linear aliphatic polyesters based on poly(ω -hydroxy acids), and the most well-known are poly(3-hydroxybutyrate) (P3HB), poly(3-hydroxyvalerate) (PHV), and the copolymers of P3HB/PHV.^[4] PHB and PHV belong to the family of poly(hydroxyalkanoates) and are mainly produced by a prokaryotic type of microorganism such as *Pseudomonas oleovorans* or *Alcaligenes eutrophus* through biotechnology. PHB and PHV are the principal energy and carbon storage compounds for these microorganisms and are produced when there are excessive nutrients in the environment. These naturally produced PHB and PHV are stereochemically pure and isotactic. They could also be synthesized in laboratories, but the characteristic of stereoregularity would be lost.

This family of biodegradable polyesters is considered environmentally friendly because they are produced from propionic acid and glucose and could be completely degraded to water, biogas, biomass, and humic materials. Their biodegradation requires enzymes. Hence, PHB, PHV, and their copolymers are probably the most important biodegradable polymers for environmental use. However, the biodegradability of this class of linear aliphatic polyesters in human or animal tissues has been questionable. For example, high-molecular-weight PHB or PHB/PHV fibers do not degrade in tissues or simulated environment over periods of up to 6 months. The degradability of PHB could be accelerated by γ -irradiation or copolymerization with PHV.

Because of its brittleness, P3HB has not been commercially successful as a biomaterial. Recently, a

close relative of P3HB was introduced by Tephra. This new biodegradable polymer, poly-4-hydroxybutyrate (P4HB), exhibits unusual elastomeric property with elongation at break as high as 1000% and tensile modulus of 70 MPa. P4HB has a low T_g (around -51°C) and low T_m (about 60°C). Because of its unusual elastomeric characteristic, P4HB may have potential in soft tissue repair.

An interesting derivative of P3HB, poly(β -malic acid) (PMA), was synthesized from β -benzyl malolactonate followed by catalytic hydrogenolysis. PMA differs from P3HB in that the 3-(CH_3) substituent is replaced by $-\text{COOH}$.^[3] The introduction of the pendant carboxylic acid group would make PMA more hydrophilic and easier to absorb than P3HB.

Poly(ester amide)s

The rationale for designing PEAs is to combine the predictable hydrolytic-induced degradability and biocompatibility of linear aliphatic polyesters with the high performance and enzymatically catalyzed biodegradability of polyamides and the potential chemical reactive sites of amide in polyamides into one single entity. Instead of using the block copolymer approach to provide both ester and amide linkages in a polymer backbone, PEAs have both ester and amide linkages within the same repeating unit, and different versions of PEAs have been reported.

Two basic types of PEAs exist: non-amino-acid based PEAs synthesized from polyesterification of amidediol monomers (which contain preformed amide linkages from aliphatic diamines)^[13] and amino-acid-based PEAs synthesized from solution polycondensation of amino acids, diols, and dicarboxylic acids.^[17-19] PEAs obtained from amidediol (i.e., non-amino-acid based) cannot be considered bioassimilative polymers because the water-soluble amidediols (first product of biodegradation) were extensively excreted without change (i.e., without the liberation of toxic diamine) in the urine. The approach of using amino acid, diols, and dicarboxylic acid building blocks to form PEAs appears to be a better one because of the biocompatibility of the amino acids used to provide amide linkages.

These amino-acid-based PEAs consisted of naturally occurring and non-toxic building blocks and had excellent film- and fiber-forming properties. PEAs of M_w from 24,000 to 167,000 with narrow polydispersity ($M_w/M_n = 1.20-1.81$) were successfully synthesized via solution polycondensation of di-*p*-toluenesulfonic acid salts of bis(α -amino acid) α,ω -alkylene diesters and di-*p*-nitrophenyl esters of diacids.^[18] These PEA polymers are initially semicrystalline polymers from solution polycondensation, but some amino-acid-based PEAs, such as L-phenylalanine and L-leucine,

become amorphous after first melting and cooling. This lack of crystallization capability from melt may be attributed to either the extensive inter- and intramolecular hydrogen bond capability from both the amide and the ester linkages in the PEA backbones that retards the chain mobility required for crystallization or the bulky side group that makes regular chain packing difficult. The T_g of these PEA ranged from 11 to 109°C .

The first generation PEAs are of saturated nature (SPEA): saturated carbon-to-carbon linkages on the backbone, i.e., the lack of $>\text{C}=\text{C}<$ double bond functionality in the PEA backbone. Recently, a new generation of PEAs, unsaturated PEAs (UPEA), was reported.^[19] The most unique aspect of these UPEAs is that there are $>\text{C}=\text{C}<$ double bonds built into either the diamide or the diester segments of the polymer backbones. As a result, UPEAs have a much higher T_g and T_m than saturated PEAs because of the rigidity of the $>\text{C}=\text{C}<$ bonds. These $>\text{C}=\text{C}<$ bonds also provide potential reactive sites for either synthesizing additional derivatives or attaching biologically active agents to render biological activity to UPEAs. An example of synthesizing additional derivatives from UPEA is the reported study of UPEA-based hydrogels via photocrosslinking with poly(ethylene glycol) diacrylate precursor,^[20] and these unique hydrogels were used as the paclitaxel carrier for its controlled release. Fig. 1 illustrates a scanning electron microscopic image of a UPEA-based hydrogel.

Beside the UPEA approach, another effort could also provide chemical functionality to SPEAs, the copolymer approach. Jokhadze et al.^[21] recently reported that, via L-lysine comonomer, the resulting co-PEAs would have pendant-free carboxylic acid over a wide range of desirable concentrations. These free carboxylic acids provide the reactive sites for the attachment of biologically active agents, such as nitric oxide derivative, and the resulting co-PEAs would have biological activity and intelligence similar to that of nitric oxide.

An effort to integrate saturated PEA with UPEA into one single entity was recently conducted^[22] and the major advantage of such an integration is to combine the merits of both saturated and unsaturated PEAs into one single entity via chemical linkages so that a wide range of physical, chemical, thermal, and biological properties could be obtained by simply changing the composition ratio of saturated to unsaturated PEAs.

Both saturated PEAs and UPEAs were easily biodegraded by enzymes such as lipase or α -chymotrypsin; however, their biodegradability in pure saline buffer is slow. It appears that an increase in the hydrophobicity of PEAs (via longer methylene groups in diols and dicarboxylic acid segments) leads to a faster enzyme-catalyzed biodegradation.^[19] In addition to the hydrophobicity factor to interpret this relationship

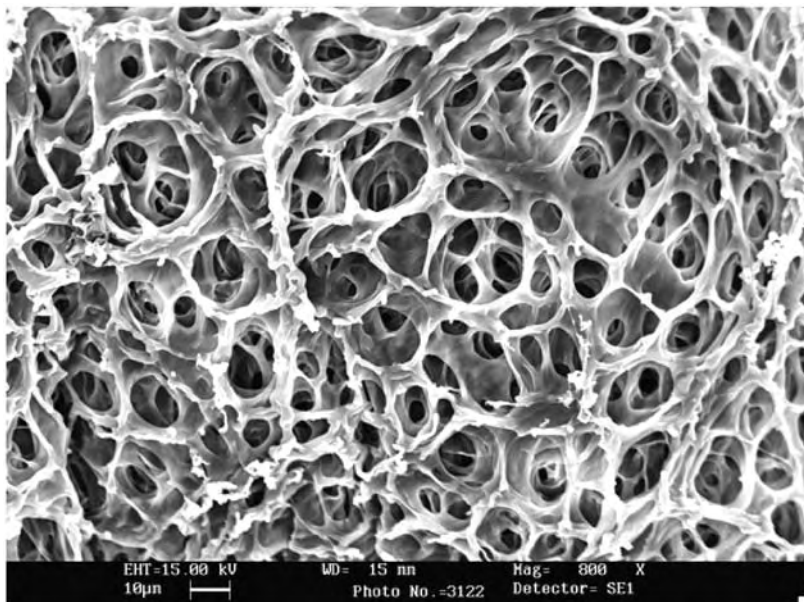


Fig. 1 The scanning electron microscopic image of unsaturated poly(ester amide)-based hydrogel after 1 day of swelling in a buffer solution.

Source: K. Guo and C. C. Chu, "Synthesis, characterization and swelling behaviors of novel biodegradable unsaturated poly(ester amide)s/poly(ethylene glycol) diacrylate hydrogels," *J. Polym. Sci. Polym. Chem. Ed.*, 43: 3932–3944 (2005).

between the length of methylene groups and enzyme-catalyzed biodegradation rate, the lack of intermolecular hydrogen bonds resulting from misalignment of the adjacent PEA macromolecules was also suggested as a possible cause behind such a relationship.

The saturated PEAs were clinically tried as a biodegradable coating on drug-eluting stents. Their biocompatibility data in an *in vivo* porcine model suggest that these new biodegradable polymers are as biocompatible as bare stents under a dynamic vessel condition.^[23] In a reported *in vitro* study of blood and cellular response to saturated PEAs, the proliferation of endothelial cells adhered to SPEA substrates was much higher than on non-biodegradable polymers like poly(*n*-butyl methacrylate) or polyethylene vinyl acetate copolymer.^[24] The amount of ATP release from human platelets on SPEAs (an indication of hemocompatibility) was twofold lower than on the above non-degradable polymers. In addition, monocytes adherent onto saturated PEA induced far less secretion of proinflammatory cytokines such as interleukin (IL)-6 and IL-1 β when comparing tissue culture-treated polystyrene control. Therefore, saturated PEAs appear to support a more natural wound healing process by promoting re-endothelialization and lowering inflammatory response.

Biodegradation Property of Synthetic Biodegradable Polymers

The reported *in vitro* and *in vivo* biodegradation studies of a variety of biodegradable polymeric biomaterials led to the following conclusions:

- Mechanical strength loss always proceeds before mass loss.
- Absorption delay, defined as the difference in time between complete mass loss and mechanical strength loss, is a good indicator for determining the useful life of biodegradable polymers.
- Molecular weights also decrease with duration of biodegradation, and the rate of molecular weight reduction correlates with the profiles of mechanical strength loss, e.g., faster molecular weight loss leads to faster mechanical strength loss. The reduction in molecular weight in an *in vitro* condition is similar to that in an *in vivo* environment.
- Oriented polymers like fibers show little change in birefringence and crystal lattice parameters during the period of mechanical strength loss, an indication of little disruption of overall fiber structure, particularly the crystalline domains.
- Biodegradation property depends on many intrinsic and extrinsic factors, such as pH, enzymes, bacteria, lipids, temperature, γ -irradiation, electrolytes, medium, free radicals, annealing treatment, plasma surface treatment, external stress, orientation, polymer morphology, and type and location of tissues.^[1]
- During the course of biodegradation, the level of crystallinity of biodegradable polymers increases first, reaches a peak, and then decreases thereafter. Thus, the time for the appearance of the crystallinity peak could be used as an indicator of degradation rate.
- Except under the influence of superoxide, the morphological features of biodegradable polymers/fibers upon either *in vitro* or *in vivo* biodegradation exhibit the formation of surface cracks; and the level of surface cracks and their orientation depend on the duration of biodegradation, the severity of the treatment, and the morphological nature of the polymers (e.g., orientation, level of crystallinity). The

onset of apparent surface morphological changes is frequently associated with the onset of mass loss.

- The role of enzymes on the biodegradation of linear aliphatic polyesters remains controversial. It appears, however, that reducing the molecular weight of these aliphatic polyesters by an external means such as γ -irradiation could facilitate subsequent enzyme-catalyzed hydrolysis. Enzymes, however, demonstrate a significant catalytic effect on the biodegradation of biodegradable polymers made from amino acids such as PEAs, which have both ester and amide linkages in their repeating units.
- Superoxide has a profound adverse effect on the mechanical property of synthetic biodegradable polymers and could shorten the duration of complete strength loss from several weeks/months to within 48 hr. The surface morphology from the superoxide-induced biodegradation of linear aliphatic polyester suture fibers reveals a moon-crater-type isotropic impression that defies the anisotropic nature of fibers.
- The degradation products of aliphatic polyesters are of acidic nature, and this acidity build-up, if inside a bulky implant, could lead to acid-catalyzed premature loss of mechanical property. Such acid-catalyzed degradation would become profound if the acid generated could not be neutralized easily as a result of poor circulation, avascular tissue, or size of the biodegradable implants.
- The level of tissue response to biodegradation depends on the biodegradation stage. In addition to the surgical trauma-induced tissue response, the bio-material-induced response would reach a peak during biodegradation and then subside after the polymer mass becomes smaller and eventually disappears.

Table 2 summarizes the structural factors that can be used to control polymer degradability.

Theoretical modeling of degradation property

Computational chemistry was used as a tool to conduct a systematic theoretical modeling study of the degradation properties of biodegradable biomaterials^[1] and the effects of a variety of hypothetical substituents that could exert either steric or inductive effects on the degradation properties of glycolide/lactide-based biodegradable polymers. This new approach could provide scientists with a better understanding of the relationship between the chemical structure of biodegradable polymers and their degradation behavior at a molecular level. It may also help the future research and development of biodegradable polymers through the intelligent prediction of structure–property relationships. Linear aliphatic polyester

Table 2 Structural factors to control polymer degradability

Factors	Methods of control
• Chemical structure of main chain and side groups	Selection of chemical bonds and functional groups
• Aggregation state	Processing, copolymerization
• Crystalline state	Polymer blend
• Hydrophilic/hydrophobic balance	Copolymerization, introduction of functional groups
• Surface area	Micropores
• Shape and morphology	Fiber, film, composite

Source: Kimura, Y., 1993, "Biodegradable Polymers," In: *Biomedical Applications of Polymeric Materials*, Ed. T. Tsuruta, T. Hayashi, K. Kataoka, K. Ishihara, and Y. Kimura, pp.164–190. CRC Press, Inc., Boca Raton, Florida. With permission.

(PGA) was used as a model compound for such a theoretical modeling.

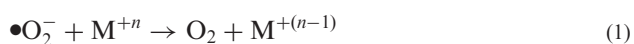
Chu et al.^[1] demonstrated a decrease in the rate of hydrolysis by about a factor of 10^6 with isopropyl α -substituents, but nearly a sixfold increase with *t*-butyl α -substituents. The role of electron-donating and electron-withdrawing groups on the rate of hydrolytic degradation of linear aliphatic polyesters was also theoretically modeled. Electron-withdrawing substituents to the carbonyl group would be expected to stabilize the tetrahedral intermediate resulting from hydroxide attack, i.e., favoring hydroxide attack but disfavoring alkoxide elimination. Electron-releasing groups would be expected to demonstrate the opposite effect. Similarly, electronegative substituents on the alkyl portion of the ester would stabilize the forming alkoxide ion and favor the elimination step. The rate of ester hydrolysis was greatly affected by halogen substituents, primarily because of charge delocalization. Chu et al.^[1] suggest that the magnitude of the inductive effect on the hydrolysis of glycolic esters decreases significantly as the location of the substituent moves further away from the α -carbon because the inductive effect is distance sensitive. In all three locations of substitutions (α , β , and γ), Cl and Br substituents exhibited the largest inductive effect compared with other halogen elements.

Therefore, the rate of ester hydrolysis was greatly affected by both alkyl and halogen substituents, primarily because of either steric hindrance or charge delocalization. In the steric effect, alkyl substituents on the glycolic esters cause an increase in activation enthalpies and a corresponding decrease in reaction rate, up to about three carbon sizes, whereas bulkier alkyl substituents than isopropyl make the

rate-determining elimination step more facile. It appears that aliphatic polyesters containing α -isopropyl groups or slightly larger linear alkyl groups, such as *n*-butyl, *n*-pentyl, etc., would be expected to show longer strength retention given the same fiber morphology. In the inductive effect, α -substituents on the acyl portion of the ester favor the formation of the tetrahedral intermediate through charge delocalization, with the largest effect seen with Cl substitution, but retard the rate-determining alkoxide elimination step by stabilizing the tetrahedral intermediate. The largest degree of stabilization is caused by the very electro-negative F substituent.

The role of free radicals in degradation property

Literature data demonstrated that the biodegradation of synthetic biodegradable polymers is closely related to macrophage activity through the close adhesion of macrophage onto the surface of the polymer. Inflammatory cells, particularly leukocytes and macrophages, are able to produce highly reactive oxygen species such as superoxide (O_2^-) and hydrogen peroxide during inflammatory reactions toward foreign materials. These highly reactive oxygen species participate in the biochemical reaction, frequently referred to as respiratory burst, which is characterized by the one electron reduction of O_2 into superoxide via either NADPH or NADH oxidase. The resulting superoxide radicals are then neutralized to H_2O_2 via cytoplasmic enzyme superoxide dismutase. These reactive oxygen species may be partially responsible for the biodegradation of polymeric implants through the production of highly reactive, potent, and harmful hydroxyl radicals OH in the presence of metals such as iron, as illustrated in the following series of redox reactions.



The net reaction will be



and is often referred to as the metal-catalyzed Haber–Weiss reaction.

The level of H_2O_2 production has been linked to the presence of foreign biomaterials including biodegradable polymers.^[1] Over a period of 14 days in rats, the amounts of H_2O_2 production by leukocytes adherent onto biodegradable PGA mesh consistently reached the highest level and was close to 200 times greater

than non-biodegradable silicone elastomer. Based on Eq. 3 above and Eq. 4 below, a higher H_2O_2 production would lead to more $\bullet OH$ and may be partially responsible for the biodegradation phenomenon of biodegradable polymers in vivo. Although the role of free radicals in the hydrolytic degradation of synthetic biodegradable polymers is largely unknown, a recent study of using absorbable sutures such as Vicryl in the presence of aqueous free radical solution prepared from H_2O_2 and ferrous sulfate, $FeSO_4$, raised the possibility of the role of free radicals in the biodegradation of synthetic biodegradable polymers.^[1] As illustrated below, both $\bullet OH$ radicals and OH^- are formed in the process of oxidation of Fe^{+2} by H_2O_2 and could exert some influence on the subsequent hydrolytic degradation of Vicryl sutures.



Scanning electron microscopy (SEM) results indicated that Vicryl sutures in the presence of free radical solutions exhibited many irregular surface cracks at both 7 and 14 days in vitro, whereas the same sutures in the two controls (H_2O_2 or $FeSO_4$ solutions) did not have these surface cracks. Surprisingly, the presence of surface cracks in Vicryl suture treated in the free radical solutions did not accelerate the tensile breaking strength loss, as would be expected. Thermal properties of Vicryl sutures under the free radical and 3% H_2O_2 media demonstrated the classic well-known maximum pattern of change in the level of crystallinity with hydrolysis time. The level of crystallinity of Vicryl sutures peaked at 7 days in both media (free radical and 3% H_2O_2). Based on the time for the appearance of the crystallinity peak as an indicator of degradation rate, it appears that the two media accelerated the degradation of Vicryl sutures compared with regular physiological buffer solution. It is unknown whether the OH^- species could be more potent than $\bullet OH$ toward hydrolytic degradation of synthetic biodegradable polymers. This is because hydroxyl anions are the sole species that can attack carbonyl carbon of the ester linkages during alkaline hydrolysis. Because an equal amount of $\bullet OH$ and OH^- is generated in Fenton reagents, the observed changes in morphological, mechanical, and thermal properties could be partially attributed to OH^- ions as well as $\bullet OH$ radicals.

Beside hydroxyl radicals, the production of superoxide ions and singlet oxygen during phagocytosis has been well documented. Although the role of superoxide in simple organic ester hydrolysis has been known since the 1970s, its role in the hydrolytic degradation of synthetic biodegradable polyester-based biomaterials has barely been addressed.^[25–27]

Because of the extreme reactivity of superoxide ion, the effect of superoxide ion-induced hydrolytic

degradation of PDLLA and PLLA was significant in terms of their changes in molecular weight and thermal properties. The superoxide ion-induced fragmentation of PDLLA results in a mixture of various species with different chain lengths. A combined Gel Permeation Chromatography (GPC) method with a chemical tagging method revealed that the structure of oligomer species formed during the superoxide-induced degradation of PDLLA and PLLA was linear. The significant reduction in molecular weight of PDLLA by superoxide ion was also evident in the change of thermal property such as T_g . The effect of the superoxide ion-induced hydrolytic degradation on the molecular weight of PLLA was similar to that of PDLLA but with a much smaller magnitude. The mechanism of simple hydrolysis of ester by superoxide ion proposed by Forrester et al. was subsequently modified to interpret the data obtained from the synthetic biodegradable polymers.

In addition to PDLLA and PLLA, superoxide ions have a significant adverse effect on the hydrolytic degradation of synthetic biodegradable sutures.^[25] There was a significant reduction in molecular weight and mechanical and thermal properties of these sutures over a wide range of superoxide ion concentration, particularly during the first few hours of contact with superoxide ions. For example, PGA suture lost almost all of its mass at the end of 24-hr contact with superoxide ions at 25°C, whereas the same suture would take at least 50 days in in vitro buffer for a complete mass loss. The surface morphology of these sutures was also altered drastically. As illustrated in Fig. 2, spherical moon-crater impression of different sizes appeared on the surface of biodegradable suture fibers biodegraded in superoxide medium. This

morphological feature is very different from the longitudinal and circumferential surface cracks observed in PBS degradation medium. The exact mechanism, however, is not fully known yet; the possibility of simultaneous occurrence of several main-chain scissions by three different nucleophilic species was suggested.

The addition of Fenton agent or hydrogen peroxide to the degradation medium retards the well-known adverse effect of the conventional γ -irradiation sterilization of synthetic absorbable sutures.^[27] Lee et al. reported that these γ -irradiated sutures retained better tensile breaking strength in the Fenton medium than in the regular buffer medium. They postulated that the γ -irradiation-induced α -carbon radicals in these sutures react with the hydroxyl radicals from the Fenton agent medium and hence neutralize the adverse effect of α -carbon radicals on backbone chain scission. This mechanism is supported by the observed gradual loss of Electron Spin Resonance (ESR) signal of the sutures in the presence of Fenton agent.

In contrast to the adverse effect of superoxide and hydroxyl free radicals on the degradation properties of synthetic biodegradable polyesters, an innovative approach of covalent bonding nitroxyl radicals onto aliphatic polyesters such as PGA and PLLA so that the nitroxyl radical attached polymers would have biological functions similar to nitric oxide has been proposed.^[28] A preliminary in vitro cell culture study of these new biologically active biodegradable polymers based on PGA and PLLA indicated that they could retard the proliferation of human smooth muscle cells as native nitric oxide does. The same approach was recently extended to functionalized co-PEAs so that these co-PEAs would have biological functionality similar to nitric oxide.^[21]

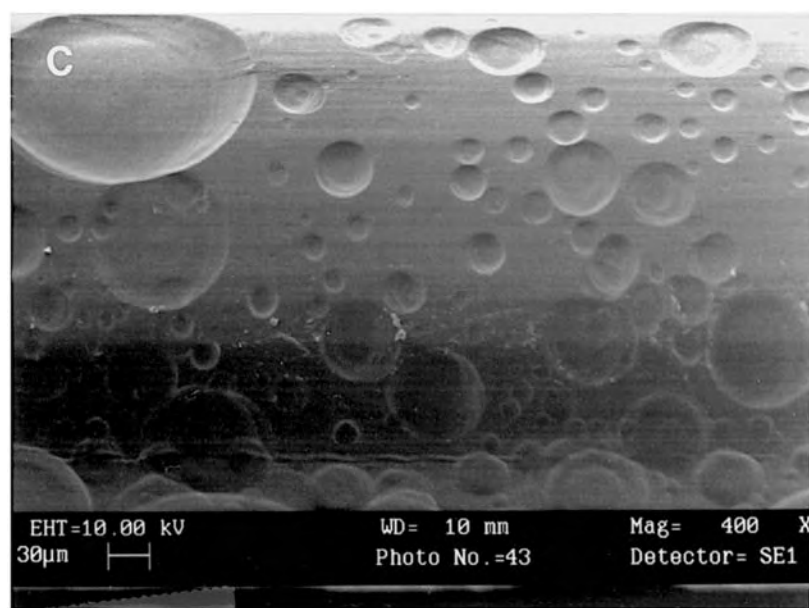


Fig. 2 The surface morphology of 2/0 size Monocryl suture upon superoxide ion-induced in vitro degradation at 25°C for 24 hr. Superoxide concentration was 0.0025 M.

Source: K. H. Lee and C. C. Chu, "The effect of superoxide ions in the degradation of five synthetic absorbable suture materials," *J. Biomed. Mater. Res.*, 49(1): 25–35 (2000).

CONCLUSION

Synthetic biodegradable polymers have become one of the most promising biomaterials for human body repair/reconstruction because of the lack of undesirable permanent chronic foreign-body reactions as well as the potential to serve as temporary scaffolds for tissue regeneration and engineering. The commercially most important and successful synthetic biodegradable polymeric biomaterials are based upon linear aliphatic polyesters. Among the family of aliphatic polyesters, the three most important building blocks used by the medical industry to design synthetic biodegradable polymers and copolymers are polyglycolic acid, polylactic acid, and poly- ϵ -caprolactone. The biodegradation of these aliphatic polyester-based biomaterials is of hydrolytic nature, but the effect of superoxide on their biodegradation was more unusual than normal hydrolytic degradation in terms of the changes in surface morphological and mechanical strength. Recently, biologically intelligent synthetic biodegradable polymers have been designed with designated biological functionality, depending on the type of biological active species incorporated. This new family of synthetic biodegradable polymers derives from PEAs. By grafting nitric oxide derivatives onto the polymer backbones, PEAs were able to retard the proliferation of human smooth muscle cells, as nitric oxide does. This new family of PEAs has been clinically tested as the coating of drug-eluting stents to reduce or eliminate the restenosis problem.

ARTICLES OF FURTHER INTEREST

Biocompatibility Testing; Biomimetic Materials; Tissue Engineering

REFERENCES

1. Chu, C.C.; von Fraunhofer, J.A.; Greisler, H.P. *Wound Closure Biomaterials and Devices*; CRC Press: Boca Raton, FL, 1996.
2. Barrows, T.H. Degradable implant materials: A review of synthetic absorbable polymers and their applications. *Clin. Mater.* **1986**, *1*, 233–257.
3. Kimura, Y. Biodegradable polymers. In *Biomedical Applications of Polymeric Materials*; Tsuruta, T., Hayashi, T., Kataoka, K., Ishihara, K., Kimura, Y., Eds.; CRC Press: Boca Raton, FL, 1993; 164–190.
4. Shalaby, S.W. *Biomedical Polymers: Designed-to-Degrade Systems*; Hanser: New York, 1994.
5. Vert, M.; Feijen, J.; Albertsson, A.; Scott, G.; Chiellini, E. *Biodegradable Polymers and Plastics*; Royal Society of Chemistry: Cambridge, England, 1992.
6. Hollinger, J.O. *Biomedical Applications of Synthetic Biodegradable Polymers*; CRC Press: Boca Raton, FL, 1995.
7. Park, K.; Shalaby, W.S.W.; Park, H. *Biodegradable Hydrogels for Drug Delivery*; Technomic Publishing: Lancaster, PA, 1993.
8. Domb, A.J. *Handbook of Biodegradable Polymers*; CRC Press: Boca Raton, FL, 1998.
9. Smith, R. *Biodegradable Polymers for Industrial Applications*; Woodhead: Cambridge, U.K., 2005.
10. Bastioli, C. *Handbook of Biodegradable Polymers*; Woodhead: Cambridge, U.K., 2005.
11. Chiellini, E.; Solaro, R. *Recent Advances in Biodegradable Polymers and Plastics*; Wiley: Hoboken, NJ, 2003.
12. Platt, D.K. Biodegradable polymers. In *Rapra Market Report*; ChemTec Publishing: Toronto, Ontario, Canada, 2006.
13. Barrows, T.H. Bioabsorbable poly(ester-amides). In *Biomedical Polymers: Designed-to-Degrade Systems*; Shalaby, S.W., Ed.; Hanser: New York, 1994.
14. Lang, M.D.; Chu, C.C. Synthesis and chemical structural analysis of nitroxyl radical incorporated poly(acrylic acid/lactide- ϵ -caprolactone) copolymers. *J. Polym. Sci. A Polym. Chem. Ed.* **2001**, *39* (24), 4214–4226.
15. Lang, M.D.; Wong, R.P.; Chu, C.C. Synthesis and structural analysis of functionalized three-arm star poly(ϵ -caprolactone)s. *J. Polym. Sci. A Polym. Chem. Ed.* **2002**, *40* (8), 1127–1141.
16. Lang, M.D.; Chu, C.C. Functionalized multi-arm poly(ϵ -caprolactone)s: Synthesis, structure analysis and network formation. *J. Appl. Polym. Sci.* **2002**, *86* (9), 2296–2306.
17. Katsarava, R.D.; Kharadze, D.P.; Japaridze, N.; Avalishvili, L.M.; Omiadze, T.N.; Zaalishvili, M.M. Hetero-chain Polymers Based on Natural Amino-acids. Synthesis of Polyamides from N-alpha, N-epsilon-bis(trimethylsilyl)lysine Alkyl Esters *Makromol. Chem.* **1985**, *186*, 939–954.
18. Katsarava, R.; Beridze, V.; Arabuli, N.; Kharadze, D.; Chu, C.C.; Won, C.Y. Amino acid based bioanalogous polymers. Synthesis and study of regular poly(ester amide)s based on bis(alpha-amino acid) α , ω -alkylene diesters and aliphatic dicarboxylic acids. *J. Polym. Sci. Polym. Chem. Ed.* **1999**, *37*, 391–407.
19. Guo, K.; Chu, C.C.; Chkhaidze, E.; Katsarava, R. Synthesis and characterization of novel biodegradable unsaturated poly(ester-amide)s. *J. Polym. Sci. Polym. Chem. Ed.* **2005**, *43*, 1463–1477.
20. Guo, K.; Chu, C.C. Synthesis, characterization and swelling behaviors of novel biodegradable unsaturated poly(ester-amide)s/poly(ethylene glycol) diacrylate hydrogels. *J. Polym. Sci. Polym. Chem. Ed.* **2005**, *43*, 3932–3944.
21. Jokhadze, G.; Machaidze, M.; Panosyan, H.; Chu, C.C.; Katsarava, R. Synthesis and characterization of functional elastomeric biodegradable poly(ester amide)s copolymers. *J. Biomater. Sci. Polym. Ed.* **2007**, *18* (4), 411–438.
22. Guo, K.; Chu, C.C. Synthesis, characterization and biodegradation of copolymers of unsaturated and

- saturated poly(ester amide)s. *J. Polym. Sci. Polym. Chem. Ed.* **2007**, *45*, 1595–1606.
23. Lee, S.H.; Szinai, I.; Carpenter, K.; Katsarava, R.; Jokhadze, G.; Chu, C.C.; Scheerder, I.D.; Hong, M.K. *In vivo* biocompatibility evaluation of stents coated by a new biodegradable elastomeric and functional polymer. *Coron. Artery Dis.* **2002**, *13* (4), 237–241.
 24. DeFife, K.; Gopalan, S.; Price, S.; Cruz-Aranda, G.; Chantung, R.; Turnell, W. *Novel Poly(ester amide) Copolymers Are Blood and Tissue Compatible*. 30th Annual Meeting of the Society for Biomaterials, Memphis, TN, April 27–30, 2005; 159.
 25. Lee, K.H.; Chu, C.C. The effect of superoxide ions in the degradation of five synthetic absorbable suture materials. *J. Biomed. Mater. Res.* **2000**, *49* (1), 25–35.
 26. Lee, K.H.; Won, C.Y.; Chu, C.C.; Gitov, I. The role of superoxide in the biodegradation of synthetic biodegradable polymers. *J. Polym. Sci. A Polym. Chem. Ed.* **1999**, *37*, 3558–3567.
 27. Lee, K.H.; Chu, C.C. *The Role of Free Radicals in Hydrolytic Degradation of Absorbable Polymeric Biomaterials*. 5th World Biomaterials Congress, Toronto, Canada, May 29–June 2, 1996.
 28. Lee, K.H.; Chu, C.C. Molecular design of biologically active biodegradable polymers for biomedical applications. *Macromol. Symp.* **1998**, *130*, 71–80.

Biofilms

Clive Edwards

Division of Microbiology and Genomics, School of Biological Sciences, The University of Liverpool, Liverpool, U.K.

B

INTRODUCTION

The ability of microorganisms to grow attached to surfaces as biofilms has become a topic of immense interest and importance for understanding multicellular behavior, cell–cell interactions, and the induction of altered physiological states. These phenomena are important for understanding the behavior of pathogens and their role in disease as well as in the wider context of microbial lifestyles in natural environments.

This entry describes the current knowledge about how microorganisms commit to a biofilm lifestyle and describes some of the features that are important for cell attachment, colonization, and maturation. Mature microbial biofilms have been likened to primitive tissues in which there appears to be a clear definition of cell function. The role of quorum sensing (QS) mechanisms for successful development of bacterial biofilms is described, particularly for pathogenic species. Because we now have sophisticated genomic technologies, it is possible to identify which genes are expressed during biofilm development, particularly those involved in secretion of extrapolymeric matrices that help anchor and provide stability to biofilm cells. Finally, the importance of the biofilm lifestyle for pathogenic species is discussed and the importance of the altered physiological states that result from cell attachment highlighted. The most important of these appears to be the development of tolerance to inimical processes, especially to antibiotics, which makes control and inhibition of biofilm development a difficult problem in the clinical environment.

BACKGROUND

In natural environments bacteria preferentially exist as biofilms, a growth mode that has been defined as: “Communities of microorganisms attached to a surface that can comprise single or multiple microbial species that attach to a range of biotic and abiotic surfaces and that are encased in an extracellular matrix.”^[1,2] Mixed species biofilms are the norm in natural ecosystems and the ways in which these are formed and respond to environmental fluctuations is complex. Most information concerning biofilm formation stems from

laboratory-based single species studies and this entry will concentrate on this aspect while acknowledging the importance of mixed species biofilms.^[3] Valuable insights into biofilm formation and growth have improved our understanding of the behavior of many microbial pathogens, e.g., *Pseudomonas aeruginosa*, *Staphylococcus aureus*, and *Candida albicans*, which are often problematic in the clinical environment.

The switch from planktonic to biofilm growth is a complex process divided into the stages of initiation, early attachment, maturation, and dissolution. Environmental signals that trigger initiation of biofilm formation appear to differ from one bacterial species to another and enable different species to colonize their preferred environment efficiently. In contrast, environmental cues that regulate the development and structure of mature biofilms, such as nutrient availability and QS, are not species specific.^[4] Nutrient availability controls the depth of the biofilm resulting in large numbers of cells that are exposed to suboptimal concentrations of nutrients. At the extremes of nutrient availability (nutrient rich or limited) there are more cells in the planktonic phase, where they either have more access to nutrients (in excess conditions) or can migrate and access more favorable habitats (nutrient limited). Similarly, QS ensures the development of optimal biofilm architecture (e.g., channels and pillar-like structures) for more efficient liquid flow for movement and availability of nutrients.^[4] Bacteria in biofilms have elevated tolerance to a range of antimicrobial agents, biocides, physical insults, and nutrient limitation but not seemingly to metal cations and oxyanions.^[5–8] Biofilms display a high level of differentiation by means of multicellularity within microcolonies, localized degradation of the biofilm matrix for dispersal of cells, and death of subpopulations within microcolonies.^[9–12]

BIOFILM FORMATION

Overview

A number of reasons have been proposed to explain why bacteria preferentially switch from a planktonic growth lifestyle to sessile, highly differentiated communities.^[13] These include protection against inimical

processes; colonization by establishment and growth in nutrient rich areas; altruism to derive the cooperative benefits of living within a community; and attached growth as the default growth mode of bacteria. It is well known that bacteria display genome plasticity enabling them to express only those genes required for growth and/or survival in a particular environment. Biofilm formation is an example of bacterial adaptation and differentiation that involves cooperative gene expression and is emerging as an important aspect of hospital-acquired infections despite extensive cleaning procedures. Many nosocomial infections are now considered to be due to the surface contamination of indwelling medical devices such as urinary catheters, stents, and central and peripheral venous catheters.^[14] Other medical devices with significant infection problems include peritoneal dialysis catheters, endotracheal tubes, cardiac pacemakers, heart valves, and orthopedic implants.^[15] Intravascular and urinary tract catheters are the most commonly used indwelling medical devices and are also the most commonly infected.^[16] Nosocomial urinary tract infections account for almost 40% of all hospital-acquired infection. Most significantly, the formation of a biofilm is a major virulence factor in device-associated infections that include osteomyelitis, heart valve endocarditis, dental caries, middle ear infections, and chronic lung infections, e.g., cystic fibrosis patients.

Biofilms form as a result of surface attachment of “pioneer” cells and it is thought that bacteria are able to sense contact with a surface leading to gene expression that in turn results in the promotion of attachment. Attachment requires that the surface is preconditioned by adsorbed molecules binding to the surface. At this

early stage, attachment is reversible and bacteria are able to detach from surfaces. Once triggered to form a firm attachment, cells grow and divide to produce microcolonies that require cell–cell interactions, sometimes production of extracellular polysaccharides (EPS), and cell–cell signaling. Signaling often involves QS, a process that has been proposed as a target for prevention or disruption of biofilm formation. Biofilms quickly adopt an antibiotic tolerant phenotype and the failure to control and reduce bacterial infections arising as a result of biofilm formation is well documented. Bacterial biofilms, therefore, seem to be regulated by surface attachment, microcolony formation, cell–cell signaling, depth, and architecture of the biofilm.^[4] The later events are regulated by more conserved processes, largely mediated by events such as QS. The stages of biofilm formation are summarized in Fig. 1.

Early-Attachment Events

Biofilm formation depends on nutrient status—oligotrophic environments produce only a sparse covering of cells with little structural complexity.^[17] Initial contact with surfaces can result in reversible attachment of cells (Fig. 1) and there has been much effort devoted to identifying the surface molecules involved. In *Pseudomonas*, a role for lipopolysaccharide (LPS) synthetic genes has been proposed in attachment but LPSs are more important for regulating surface protein integrity. Treatment of *Pseudomonas fluorescens* cells with low amounts of protease resulted in loss of the ability to attach to a surface, implicating a role for surface proteins.^[18] In *Pseudomonas aeruginosa*, flagella are

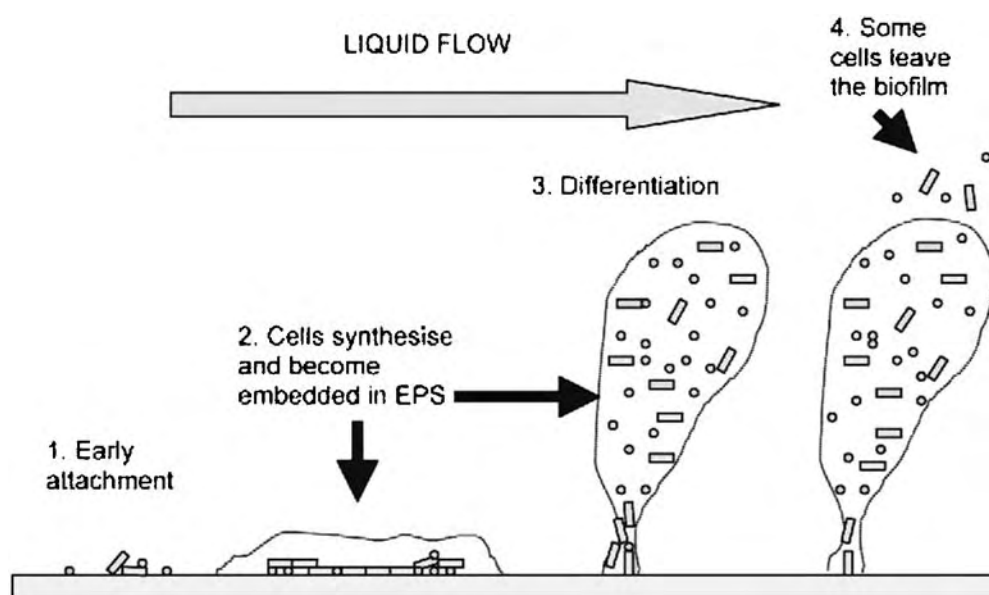


Fig. 1 Formation and developmental stages of biofilm.

important for movement of cells on a surface and after attachment twitching motility, which depends on type IV pili and on cell–cell contact, provides a means for cells to aggregate into a microcolony.^[19]

In the early stages of biofilm formation, microcolonies can move as one across the surface. Mutants defective in type IV pili form only small aggregates and/or dense monolayers of cells compared with wild-type cells that form characteristic mounds. A flagellar mutant of *P. aeruginosa* was unable to colonize a glass flow cell as efficiently as a parental strain.^[17] The *crc* gene (catabolite repression control) has been proposed as part of a signal transduction pathway that relays input signals about carbon availability to the cell regulating the switch from planktonic to biofilm growth. The *crc* gene is also involved in twitching motility and helps regulate the genes needed for type IV pili synthesis.^[20] Alginate synthesis is upregulated in attached *P. aeruginosa* cells presumably by, some form of surface sensing. This polymer is a major constituent of the EPS matrix that helps preserve biofilm integrity. Large numbers of *sad* (surface attachment defective) mutants of *P. aeruginosa* have been described, defective in such properties as flagellum-mediated motility and/or biogenesis of type IV pili. These mutants often attach as a monolayer but fail to form the microcolonies typical of early biofilm development.^[3] *Vibrio cholerae* requires flagella-mediated motility, type IV pilus, and synthesis of the major EPS to form a biofilm. Motility appeared to aid surface colonization because flagellar mutants had larger voids between the microcolonies.^[17] Flagella-mediated motility may be necessary for bacteria on the surface to spread out across the surface and in flagellum minus mutants, for those cells that do attach, the resulting biofilm is identical to that of a wild type implying that flagellar attachment is not a prerequisite for downstream development.^[18] In *Escherichia coli* flagella, type 1 pili and curli fimbriae have been implicated in biofilm formation. Flagellar mutants can still attach but cannot move out across the surface. The importance of LPS is suggested from evidence that signals from the outer membrane or periplasm are important in regulating biofilm development.^[21,22]

Several nonmotile Gram positives such as streptococci, staphylococci, and mycobacteria readily form biofilms. *Streptococcus mutans* is implicated in dental caries as well as, occasionally, infective endocarditis.^[23] It grows and survives as a biofilm in its natural habitat of dental plaque where it is subjected to a variety of stresses. These include periods of starvation and acid stress generated by sugar utilization. Comparisons of acid tolerance of planktonic and biofilm cells has shown that starved bacteria exhibit higher acid tolerance than nonstarved cells and that biofilm cells showed greater resistance to acid shock promoted by

lactic acid than did planktonic ones.^[24,25] In *S. aureus* or *S. epidermidis* attachment is initially via cell–surface interactions mediated via MSCRAMMs (microbial surface components recognizing adhesive matrix molecules) as well as *altE*, which encodes an autolysin that may act as an adhesion factor. Subsequent to cell–surface interactions there is an accumulative phase of biofilm formation involving cell–cell interactions and the formation of cell aggregates on the surface. A polysaccharide intercellular adhesion (ICA) factor has been implicated in this process encoded by the *ica* genes. Once a biofilm is established, expression of a number of adhesins is repressed suggesting that the main role of adhesions, pili and flagella is in the early attachment phase and that they are no longer needed during differentiation into a mature biofilm. The cyclic peptide dependent accessory gene regulator (Agr) QS system of *S. aureus* represses several surface adhesions that mediate contact with the host matrix. These include fibronectin/fibrinogen-binding proteins. Under certain conditions *agr* mutants adhere more efficiently than wild types.

Colonization and Maturation—The Role of QS

In biofilms, higher cell densities and close proximity means that high concentrations of metabolites and secreted molecules exists. Close proximity aids in acquisition of new genes via recombination and horizontal gene transfer. Cell–cell interactions are extremely important and for most species QS is important for biofilm development. The study of this interactive lifestyle has been termed sociomicrobiology to describe the group behavior of microorganisms.^[26] Quorum sensing has been extensively investigated in planktonic cultures, which provide a tractable system for assessing the effects of known concentrations of signaling molecules on different cell types. This is more difficult for biofilms because of the different chemical and physical gradients that exist. The two best understood QS systems are the acyl-homoserine lactone (acyl-HSL) systems of gram-negative bacteria and the secreted peptide signals of many gram-positive species.^[26] A third type of system, found in both gram-positive and -negative species, involves a furanosyl borate diester as the signaling molecule that is synthesized by the *luxS* gene product.

In gram-negative bacteria, acyl-HSL molecules vary in hydrophobicity depending on the length of the acyl side chain. Biofilms are usually encased in EPS that also contains proteins, nucleic acids, and dead cells and acyl-HSLs are assumed to diffuse freely through this matrix, although depending on the relative hydrophobicity of EPS components and the acyl-HSLs themselves, EPS can act as a sink sequestering the signaling molecules. Natural environments pose additional problems of

species that degrade acyl-HSLs as a carbon and energy source, e.g., *Bacillus*. The physiological status of a cell will determine the concentrations of cell metabolites destined for signal production and cells deep in the biofilm may have reduced metabolic activity and lowered concentrations of available acyl-ACP (acyl, acyl carrier protein) and SAM (shoot apical meristem)—the precursors for synthesis of acyl-HSL. Different signal generation rates explain why mature biofilms display cell heterogeneity. Recently, simultaneous treatment of a *P. aeruginosa* biofilm with an antibiotic and a brominated furanone (an HSL analog that blocks the QS response) resulted in killing of the cells, suggesting that inhibition of HSL-mediated QS converted antibiotic tolerant cells to sensitive ones.^[27]

The architecture of mature biofilms can be flat and homogenous or highly structured characterized by liquid flow through channels and pillars of cells encased in EPS. The three-dimensional structure produces physicochemical gradients, which will affect the resistance of cells to inimical processes, e.g., antibiotic resistance. In a tightly packed arrangement of cells, QS must also have a potentially important role in governing the properties of the biofilm. For example, acyl-HSL QS affected the maturation of *Serratia liquefaciens* biofilms, probably by influencing swarming/motility behavior. Wild-type biofilms were heterogenous with cell aggregates of long filamentous cells but in a *swrI* mutant thin biofilms lacking aggregates of cells and filaments were formed.^[26] *CepI/R* QS genes of *Burkholderia cepacia* also control biofilm maturation and mutants in either *cepI* or *cepR* formed biofilms arrested at the microcolony stage of growth.^[28] *AhyR/I* of *Aeromonas hydrophila* were required for biofilm maturation shown by an *ahyI* mutant that formed structurally less well-developed biofilms compared with the wild type. Acyl-HSL controls cell aggregation in *Rhodobacter sphaeroides* and a *cer* mutation (community escape response) resulted in hyperaggregation of cells in liquid culture. *Yersinia pseudotuberculosis*, an enteric pathogen, shows a similar phenotype for QS mutants in the *YpsR/I* system because mutation in *ypsR* caused cells to aggregate in liquid culture. The plant pathogen *Xanthomonas campestris* produces a signal termed DSF (diffusible signal factor) encoded by *rpfF*. A two-component system senses the signal (*rpfC* and *rpfG*) and mutations in the system result in increased aggregation in liquid culture and increased biofilm formation. *V. cholerae* uses QS to regulate production of secreted polysaccharide encoded by the *vps* operon. Repression of *Vps* results in overproduction of EPS and a smaller wrinkled colony phenotype on solid medium. *Vps* overproduction via QS stimulates biofilm formation and aggregation and may protect cells during passage through the acid pH of the human stomach.^[26] Because QS acts as a global regulator it is probably not

surprising that it affects the biofilm formation in many different ways and is not confined to a single biofilm associated event. The primary QS system for *P. aeruginosa* is *lasR/I*, which controls the expression of the second system *rhlR/I*.^[29] Nutrient availability affects the expression of *lasR/I* and *las I* mutants that formed flat, densely packed biofilms compared with the wild type that formed well-differentiated mature biofilms.^[30]

Hitherto, we have discussed the role of QS as a cell density dependent chemically mediated response. Intriguingly, there is now emerging evidence that environmental factors can also influence the expression of QS genes. Recently, the expression of *lasR* was studied at different concentrations of iron and pO₂ and evidence was obtained to show that iron concentration influenced *lasR* expression, which was highest at low iron and triggered by oxidative stress. Under oxygen limitation, *lasR* expression was strongly dependent on iron concentration. A proteomic analysis demonstrated that many proteins upregulated as a result of iron limitation were QS regulated proteins showing that for *P. aeruginosa* there is a strong link between iron/oxygen concentration and QS.^[31] Furanones from the red alga *Delisea pulchra* interfere with biofilm formation by interfering with the QS system by acting as molecular mimics to acyl-HSLs that inactivate LasR. Rhamnolipid is required for maintenance of *P. aeruginosa* biofilm architecture and swarming motility is also QS dependent. Mutants in the *las* or *rhl* systems are deficient in swarming motility although the link between swarming and biofilm development has not been shown for *P. aeruginosa*. Quorum sensing also regulates iron acquisition systems such as biosynthesis of the siderophore pyoverdine—iron limitation inhibits biofilm formation.

Gram-positive organisms like *Streptococcus intermedius* secretes competence stimulating peptide (CSP) as a cell-cell signaling molecule that increases competence of cells for transformation by exogenous DNA and enhances acid tolerance during biofilm formation.^[32] LuxS type QS in *S. mutans* has also been shown to be important for the formation of an ordered and structured biofilm whereby a mutant in *luxS* resulted in mature biofilms of lower biomass and abnormal architecture. Biofilm cells are 10-fold more competent than the corresponding planktonic ones. *S. pneumoniae*, *S. gordonii*, and *S. mutans* also synthesize CSP and the enhanced competence they display is thought to enable cells to acquire useful genes via DNA uptake, e.g., antibiotic resistance. In *S. pneumoniae* CSP interacts with its receptor ComD to initiate a cascade of four temporally distinct transcription profiles—early, late, delayed gene induction, or delayed gene repression. Early competence genes encode proteins for cell-cell signaling including CSP precursor ComC, its histidine kinase receptor comD, a cognate response regulator

ComE, and an alternate sigma factor ComX. The latter induces later waves of transcription and many of the genes expressed control DNA uptake and its integration (late competence genes), stress related functions (delayed class), and protein synthesis (repressed class).^[33] Other roles for the CSP signal in streptococci include biofilm formation, acid tolerance, bacteriocin production, and virulence.^[34] The CSP mediated QS in biofilm formation is illustrated in *S. gordonii*, whereby a mutant in *comD* was defective in biofilm formation. Also in *S. mutans* inactivation of any component of its *comCDE* pathway resulted in a competence and biofilm defective phenotype that had diminished acid tolerance.^[34]

The Role of EPS

Extracellular polysaccharides have been identified in many bacteria. Colanic acid in *E. coli*, alginate, glucose- and mannose-rich Pel, and PLS matrix components in *P. aeruginosa*, cellulose and β -1,6-GlcNAc polymer in *Salmonella* and *E. coli* have all been reported to play important roles in biofilm formation.^[35] These EPS are key elements that shape, provide structural support, and hold the biofilm microcolony cells together in a slime matrix.^[36] Extracellular polysaccharides synthesis is stimulated during nutrient excess and when nutrients become depleted, bacteria detach from the biofilm and become planktonic, presumably for dissemination to another, more favorable site. Structural complexity of *P. aeruginosa* mushroom/mound structures, which develop in mucoid biofilms of wild-type cells isolated from cystic fibrosis (CF) lung, was suppressed in a nonmucoid mutant that formed patchy flat biofilms. Alginate is assumed to be important for tolerance to antibiotics in the CF lung but it is the altered phenotype of biofilm cells that makes them apparently antibiotic resistant, not the alginate barrier. Quorum sensing appears to be involved in EPS formation via *lasI/lasR* as evidenced by the observation that a *lasI* mutant attaches but fails to develop the biofilm architecture of the wild-type cells.

Other EPS have a similar role in other species. Glucose- and galactose-rich EPS in *V. cholerae* was required for complex structural biofilm formation. Increased production of EPS by *V. cholerae* correlated with reduced intestinal colonization in mouse intestine. A strain defective in EPS production was defective in the early stages of biofilm development. In *E. coli* mature biofilms had a similar architecture to that of *P. aeruginosa*. Colanic acid—the major EPS in *E. coli*—is necessary for formation of this characteristic architecture but not in early attachment. Upregulated genes in developing biofilms of *E. coli* include *ompC* porin and the *wca* locus (needed for colanic acid synthesis), whereas *fliC* for flagellar synth was downregulated.^[35]

Some gram-positive pathogens also produce EPS, for example, polysaccharide intercellular adhesin (PIA) and poly-*N*-acetyl galactosamine (PIA-NAG) are encoded by the *icaABCD* locus in *S. aureus* and *S. epidermidis*.^[35] In *S. epidermidis* nutrient limitation can be a trigger for slime (EPS) production. Production of capsular PS/A is closely associated with EPS production. *S. aureus* also makes EPS in a phase-variable manner and σ^B may play a role in cell-cell interactions in the biofilm. After attachment, bacteria adapt further and two major properties include increased synthesis of EPS and development of antibiotic tolerance. These help create a protective environment so that biofilms become a serious problem in the clinical environment because of the problems associated with eliminating them.

B

Gene Expression During Biofilm Development

During biofilm development, a program of gene expression is seen whereby some genes are induced during the early-attachment phases while others are expressed later on during the development and later maturation phases only. Diverse genetic factors are therefore involved in biofilm development. *Shewanella oneidensis* is a facultative Fe(III)- and Mn(IV)-reducing bacterium involved in dissolution of Fe and Mn oxide minerals in soils and sediments where it forms biofilms on mineral surfaces. A total of 173 insertion mutants, of the 15,000 screened, after transposon mutagenesis had defects in initial attachment and/or early stages of biofilm formation and 71 mutants had a nonswimming phenotype suggesting that motility was important in biofilm formation. Disruption mutants in motility genes *flhB*, *fliK*, and *pomA* did not affect initial attachment but stopped biofilm development and differentiation. Mutants defective in mannose-sensitive hemagglutinin type IV pilus biosynthesis and in pilus retraction (*pilT*) showed severe defects in adhesion to abiotic surfaces and biofilm formation.^[37]

Sauer et al examined the changes in protein and gene expression in *P. aeruginosa* over 12 days to identify five distinct stages of biofilm development.^[10] The 2-DE monitored phenotypic differences during each stage to identify 57 biofilm specific proteins that differed from the planktonic profile. Of these 23 were further analyzed by MALDI-ToF (matrix-associated laser desorption ionization-time of flight mass) and found to be associated with metabolic processes such as carbon and lipid metabolism. Evidence was also obtained for the production of multiple phenotypes with time.

A DNA array study of gene expression in a 5-day *P. aeruginosa* biofilm showed relatively few genes (ca. 1%) that were differentially expressed in planktonic and biofilm cells. Specifically, *rpoS*, flagella, and pili genes were downregulated in biofilm cells. Flagella are

not thought to be necessary for later maturation stages of biofilm development because flagellum-encoding genes are repressed soon after bacteria arrive at a surface.^[36] DNA microarrays were also used to look at gene expression in biofilms and planktonic cells by means of a chemostat in which cells grew near to μ_{\max} and into which granite pebbles were placed as surfaces for biofilm growth so that both planktonic and attached cells were exposed to similar cultural conditions.^[38] Only 1% of genes showed differential expression in the two growth modes—about 0.5% of genes were activated and 0.5% repressed in the biofilm cells. Some of the regulated genes are known to affect antibiotic sensitivity of *P. aeruginosa* in free-living cells. Exposure of biofilms to tobramycin caused differential expression of 20 genes. The most highly activated genes in *P. aeruginosa* biofilm cells (not exposed to tobramycin) were those of a temperate phage closely related to the filamentous bacteriophage Pf1. The genome of *P. aeruginosa* PAO1 contains 11 out of 14 Pf1 genes. Phage induction might therefore be important for gene transfer within the biofilm. Genes for pili and flagella were repressed in biofilm cells and therefore these appendages, although important for early attachment, may not have a role in mature biofilms.^[38]

An *rpos* minus mutant showed enhanced biofilm formation and increased tolerance to tobramycin. *Crc* (global carbon metabolism regulator) minus mutants of *P. aeruginosa* produced an attenuated biofilm ascribed to type IV pili expression/function. *CsrA* (global carbon regulator in *E. coli*) disruption led to increased biofilm formation and overexpression of *csr* was inhibitory. Most of the biofilm cells not in contact with nutrient fluids will be subject to limitations of oxygen and nutrients. In *E. coli* a significant part of the biofilm response involves stationary phase induced genes. In WT *B. subtilis*, out of 121 biofilm induced genes 60% are activated during sporulation. However, in many species induction of the stationary phase transcriptional activator *rpoS* has been shown to be repressed twofold or slightly activated in *P. aeruginosa* and its role in biofilms of *E. coli* is debatable.^[36]

The ability to form biofilms by some species is associated with virulence, e.g., *P. aeruginosa*, *S. aureus*, *S. mutans*, and *Enterococcus faecalis*. The majority of clinical isolates of *E. faecalis* have the ability to form biofilms and a secreted zinc metalloprotease controlled by a signal transduction pathway plays a key role. The enzyme is gelatinase, a thermolysin-like M4 protease similar to those found in other pathogens.^[39] It cleaves to a wide range of substrates such as Azocoll, casein, gelatine, hemoglobin, plasmid conjugation factors, collagen, fibrin, and an autolysin. It is encoded by *gelE*, the expression of which is dependent on *fsr* genes that encode a two-component transduction system. Mutants defective in *fsrA* are also impaired in biofilm formation.

FsrC (*FsrA*'s cognate kinase) disruption also impaired biofilm formation as did that for *gelE*. In the disrupted mutants, biofilm formation could be rescued through insertion of a plasmid harboring *gelE*.^[40]

Stress responsive genes have been identified in *S. mutans* and include AP-185 (acid stress protein) and GSP-781 (general stress protein), which is upregulated in response to high osmolarity, temperature, and low pH.^[23] Chaperonin proteins have also been identified as possible candidates for the protection of *S. mutans* cells against numerous stresses and include a degP-like protease that probably degrades damaged proteins resulting from heat shock and oxidative damage and DnaK and GroE, which help protect cells against acid stress and starvation in oral biofilms.^[41,42] A homolog of the HtrA family of stress-response proteases was identified from a partial genome sequence of *S. mutans* and insertional inactivation of the gene resulted in bacterial clumping in liquid media, altered colony morphology, and a reduced ability to withstand high temperature, extremes of pH, and oxidative stress.^[43]

It is clear that bacteria growing in biofilm communities handle stresses and environmental perturbations quite differently from planktonic cells and understanding the nature of protein expression and global gene regulation within biofilm should improve our ability to understand biofilm lifestyles. For example, a recent report highlights the potential disparity between protein synthesis in planktonic and biofilm cells of *S. mutans*. Of the 694 proteins examined, 57 were upregulated and 78 were downregulated in biofilm cells compared with planktonic ones.^[24] No further analysis of these proteins was undertaken. Some of the stress-induced genes that have been identified in bacterial biofilms are listed in Table 1.

Detachment and Dissolution

Cells can actively leave biofilms as dispersal mechanisms for colonization of new surfaces. Active detachment of cells from bacterial biofilms is a critical but poorly understood stage. Detachment of *Shewanella oneidensis* cells from biofilms could be induced by stopping the flow of medium in a hydrodynamic biofilm system. Detachment occurred predominantly in cells most exposed to medium flow while in thin biofilms up to 80% of biomass could be induced to detach. Swimming motility was not required for detachment and the trigger appeared to be the decrease in oxygen tension resulting from cessation of flow. *ArcA* and *crp* were implicated as regulatory genes for detecting changes in oxygen and the induction of detachment.^[44] Biofilms grown under no or low flow tend to dislodge clumps of cells more easily than those developed under high flow. The authors proposed

Table 1 Some stress responses that have been identified in bacterial biofilms

Function	Gene or protein	Organism
Prophages	Pf1	<i>P. aeruginosa</i>
	PBSX	<i>B. subtilis</i>
Proteases	Clp protease	<i>L. monocytogenes</i>
DNA repair	RecO	<i>L. monocytogenes</i>
SOS response	RecA, DinI, SulA	<i>E. coli</i>
Chaperones	DnaK, DnaJ	<i>E. coli</i>
Heat shock	Htpx, HtpG	<i>E. coli</i>
Oxidation stress	Sod proteins, CysK	<i>L. monocytogenes</i>
	SodB	<i>P. aeruginosa</i>
	SoxS	<i>E. coli</i>
Envelope stress	<i>Cpx</i> and <i>rpoE</i>	<i>E. coli</i>
	pathways	<i>E. coli</i>
	<i>Psp</i> pathway	
Sigma factor	σ^w -mediated response	<i>B. subtilis</i>

(From Ref. [36].)

three dispersal strategies: 1) swarming dispersal (e.g., *P. aeruginosa*); 2) clumping dispersal (e.g., *S. aureus*); and 3) surface associated sliding motility due to liquid flow followed by dispersal.^[44] A possible signal for detachment may be starvation. Overexpression of alginate lyase can promote detachment in *P. aeruginosa* and production of proteases that degrade adhesions may also be another mechanism that promotes detachment. Biofilms consist of large numbers of cells and detachment and dissolution have implications for pathogen transmission and disease. For example, *V. cholerae* has an infective dose of approximately 10^4 – 10^6 cells and this is most likely to be attained by detachment of cells as clumps from biofilms rather than in the planktonic mode of growth.

DOES ALTRUISM ENSURE STABILITY OF BIOFILMS?

Biofilms have been likened to multicellular organisms in which bacterial species display cooperation and unselfish behavior for the good of the community.^[45] In the complex architecture of mature biofilms and the prevailing physical and chemical gradients that exist it is not surprising that there is a great deal of heterogeneity that results in phenotypic diversity manifest even in single-species biofilms. The situation in multi-species biofilms is even more complex and difficult to study. Phenotypic diversity can also be viewed as a form of division of labor and therefore altruistic. Bacteria in biofilms communicate using QS leading to

concerted programs of gene expression. There is evidence that bacteria in biofilms undergo programmed cell death, again suggesting multicellularity and altruistic behavior. However, the role of autoinducing signals in biofilms is ambiguous. A role for the *luxS* system in *S. mutans* has been suggested but in the other work no role could be identified. In *lasR/lasI* in *P. aeruginosa* similar contradictory work has been published. Biofilms may also be hot spots for gene transfer because of the close proximity of cells and the induction of prophage genes that can stimulate genetic exchange mechanisms, which in turn may also help spread antibiotic-resistant genes. For example *P. aeruginosa* was grown in flow glass reactor cells from which small colony variants (SCVs) developed. Their emergence correlated with that of plaque-forming Pf1-like filamentous phage, designated Pf4, from the biofilm. It was observed that only SCV cells had high densities of Pf4 filaments on the surface and not the parental cells. These phage filaments were often interwoven into complex lattices surrounding the cells. Infection of planktonic *P. aeruginosa* cells with Pf4 caused emergence of SCVs in culture that exhibited enhanced attachment and accelerated biofilm development, suggesting that Pf4 can mediate phenotypic variation in *P. aeruginosa* biofilms.^[46]

BIOFILMS AND PATHOGENS

Bacteria

Biofilms can act as a reservoir for dispersal of large numbers of cells that subsequently cause disease, a serious problem in the clinical environment when dissolution of cells growing on indwelling medical devices such as catheters occurs. Because biofilms are heterogeneous in nature and promote phenotypic heterogeneity, it is possible that such a milieu also influences virulence genes of pathogens. Quorum sensing in *P. aeruginosa* biofilms induces large numbers of genes that encode virulence determinants. Biofilm development by *V. cholerae* aids in its persistence in the environment, thought to occur in stages consisting of attachment of flagellated cells to surfaces such as exoskeletons of crustaceans; formation of a monolayer of cells that degrades surface polysaccharide polymers generating a nutrient-rich environment; generation of signals for decreased flagellar gene expression, and increased transcription of genes for ICA and EPS production. *V. cholerae* cells growing as biofilms would, during sloughing, be expected to generate significantly high concentrations of cells.^[17] Materials commonly used in clinical environments include silicone elastomer, polyurethanes, and polyvinylchloride in catheter applications, while metal alloys and other polymers are used

in orthopedics. There have been significant developments in the past 20 yr in understanding the interactions that take place between tissues and these materials, often referred to as biomaterials, and how such interactions may be manipulated to improve biocompatibility and biofunctionality. Prevention of device-related infection and biofilm development has been less successful and efforts have mainly been directed toward chemically modifying the biomaterial substrate to provide an antimicrobial coating.^[47,48] In recent years researchers have investigated the effects of substrate topography.^[49–51] Primarily, such studies are intended to design substrates that provide a more acceptable biomaterial/tissue interaction, but recent studies have indicated the potential for nanotopography to inhibit mammalian cellular attachment.^[52] Bacterial attachment to the surface of a medical device is mediated by surface charge attraction, hydrophilic/hydrophobic interaction, specific protein interaction, and surface topography. Many of these factors have been exhaustively investigated; however, relatively little information is available with respect to the influence of the biomaterial substrate.

A number of microbial species have been implicated in biofilm development associated with medical devices.^[53] These include gram-negative bacteria (*E. coli*, *Proteus mirabilis*, and *Pseudomonas aeruginosa*), biofilms of which have been implicated in urinary tract infections and CF; gram-positive bacteria (*Staphylococcus epidermidis* and *S. aureus*), important as skin commensals and implicated in many hospital acquired infections and acquisition and transfer of broad-spectrum antibiotic resistance, i.e., MRSA (methicillin-resistant *S. aureus*); and fungi (*C. albicans* and *Candida tropicalis*) that are opportunistic pathogens, grow as biofilms, and cause infections that range in severity from irritation (e.g., thrush) to life threatening as in the case of immunocompromised patients.^[15,54–63] A list of some human infections that involve biofilms are shown in Table 2.

Candida Biofilms

Candida spp. are associated with infections in humans, which range from irritation (thrush) to life-threatening infections usually associated with immunocompromised patients. They are also major agents of hospital-acquired infections and rank fourth behind coagulase-negative staphylococci, *S. aureus*, and streptococci as the cause of bloodstream infections.^[63] *C. albicans* is the best studied and it has now emerged as one of the most important causes of nosocomial infections. A major reason appears to be the dimorphic nature of this yeast whereby it grows as budding yeast cells or as a hyphal form. Transition from yeast cell to mycelium can be

Table 2 List of some human infections involving biofilms

Infection or disease	Common bacterial species involved
Dental caries	Acidogenic streptococci
Cystic fibrosis pneumonia	<i>P. aeruginosa</i> <i>Burkholderia cepacia</i>
Endocarditis	Viridans group streptococci, staphylococci
Musculoskeletal infections	Gram-positive cocci
Biliary tract infections	Enteric bacteria
Infectious kidney stones	Gram-negative rods
Contact lenses	<i>P. aeruginosa</i> , gram-positive cocci
Sutures	Staphylococci
Arteriovenous shunts	Staphylococci
Urinary catheter infections	<i>E. coli</i> , gram-negative rods
In-dwelling catheters	<i>Candida albicans</i>
Human epithelial cells	<i>Candida</i> spp.

(From Ref. [63,64].)

triggered in the presence of serum and there is evidence that the hyphal form may be the mechanism by which the yeast invades host tissues and establishes itself as a biofilm growing on epithelial cells.^[65] Recently, the mechanisms thought to be important for adherence and pathogenesis have been probed using DNA array analysis of the transcriptome of *C. albicans* growing via germ tubes on polystyrene petri dishes.^[65] This work demonstrated that 77 genes were overexpressed in adherent germ tubes and some of these encoded enzymes were involved in lipid, protein, and carbohydrate metabolism. Some of these were also known to be linked to pathogenesis. Interestingly, nearly 50% of the overexpressed genes were of unknown function.

In recent work, we have been investigating whether fabricated nanoparticulate surfaces that inhibit attachment of mammalian cells to surfaces can be used to control or inhibit attachment of microbial cells to surfaces.^[52] Fig. 2 shows interesting preliminary data whereby *C. albicans* yeast cells were grown in the presence of horse serum on a glass surface (Fig. 2A) and on the same surface coated with silica beads of 14 nm diameter (Fig. 2B). The yeast cells grew and began to produce germ tubes after 4 hr incubation, largely stimulated by the presence of horse serum. On the nanoparticulate surface there was little growth and a failure to differentiate into germ tubes. After 24 hr the glass surface was completely covered by a mixture of yeast cells and hyphae (Fig. 2C) whereas the nanoparticulate-coated surface was also more covered than at 4 hr but

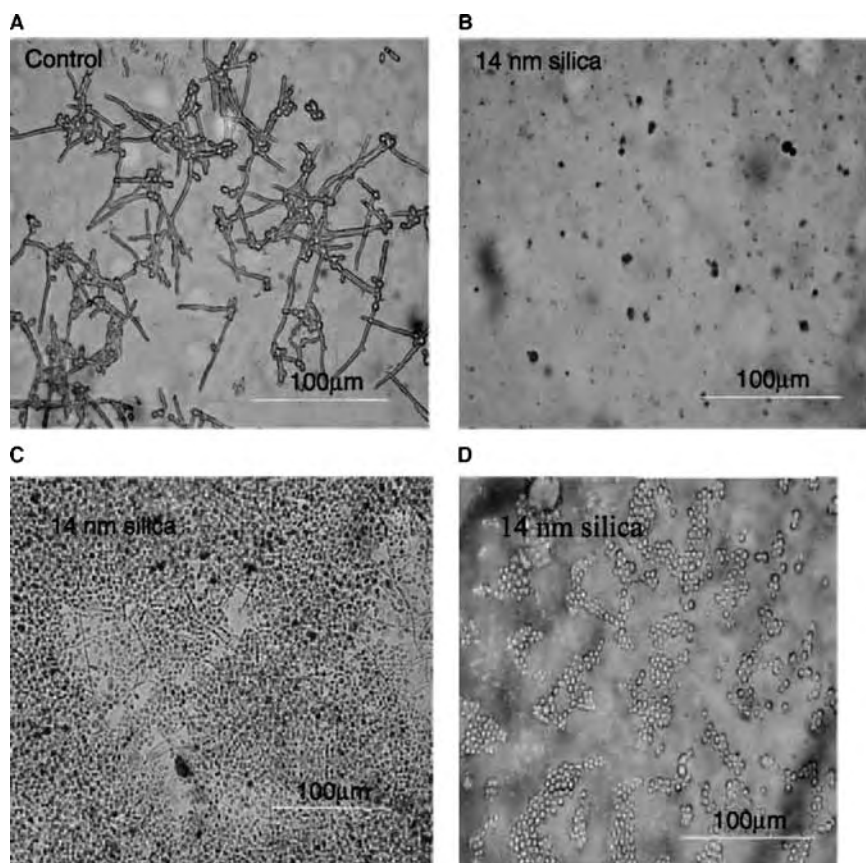


Fig. 2 Response of *C. albicans* grown for 4 hr in the presence of horse serum on a glass surface (A); on the same surface coated with 14 nm-diameter nanoparticles (B); and after 24 hr incubation on glass (C) and 14 nm silica (D).

with a film of yeast cells only. Therefore, not only had the silica nanoparticles reduced the coverage of the yeast to about 15% of that of the untreated glass, but it had also apparently prevented the differentiation of yeast cells to a hyphal mycelium. This work is continuing to assess why this occurs, whether other nanoparticulate surfaces have the same effect, and how the nanoparticulate surface can be manipulated to provide total prevention of yeast cell attachment and proliferation (Edwards and others, unpublished data).

CONCLUSIONS

Biofilm formation has now been recognized as a developmental and possibly the natural mode of growth for microbial populations. Common responses identified during biofilm maturation include repression of flagellum synthesis and hyperexpression of genes for adhesion and ribosomal protein.^[2] Further, work using high-throughput technologies of DNA arrays and proteomics will no doubt begin to identify the molecular basis for biofilm formation. These new techniques should also address the heterogeneity that exists within biofilm populations, even for single-species cultures, and help to unravel the multifaceted roles of these heterogeneous subpopulations. Recently, the concept of “persister cells” helps explain why biofilm cells have

altered properties, particularly antibiotic tolerance.^[66] In *E. coli* and some other bacteria, persister cells arise in a cell density dependent manner to reach ca. 1% of the population in stationary phase liquid cultures or biofilms. These cells neither grow nor die in the presence of bactericidal agents and exhibit multidrug tolerance and appear to be responsible for the robustness and tolerance of biofilms to inimical processes.^[66] Nanotechnology may have a role in modifying cell behavior at surfaces by either altering surface topography (“roughness”) or anchoring biocidal molecules to surface-bound carrier nanoparticles. An area yet to be addressed is that of the ways in which biofilms operate in natural ecosystems and what contribution they make to the global biogeochemical cycling of key elements.

ARTICLES OF FURTHER INTEREST

Bioadhesion; Biological Adhesives from Nature

REFERENCES

1. O’Toole, G.; Kaplan, H.B.; Kolter, R. Biofilm formation as microbial development. *Ann. Rev. Microbiol.* **2000**, *54*, 49–79.

2. Lazazzera, B.A. Lessons from DNA microarray analysis: the gene expression profiles of biofilms. *Curr. Opin. Microbiol.* **2005**, *8*, 222–227.
3. Davey, M.E.; O'Toole, G.A. Microbial biofilms: from ecology to molecular genetics. *Microbiol. Mol. Biol. Rev.* **2000**, *64*, 847–867.
4. Stanley, N.R.; Lazazzera, B.A. Environmental signals and regulatory pathways that influence biofilm formation. *Mol. Microbiol.* **2004**, *52*, 917–924.
5. Costerton, J.W.; Stewart, P.S.; Greenberg, E.P. Bacterial biofilms: a common cause of persistent infections. *Science* **1999**, *284*, 1318–1322.
6. Hogan, D.; Kolter, R. Why are bacteria refractory to antimicrobials? *Curr. Opin. Microbiol.* **2003**, *5*, 472–477.
7. Harrison, D.; Ceri, H.; Stremick, C.A.; Turner, R.J. Biofilm susceptibility to metal toxicity. *Environ. Microbiol.* **2004**, *6*, 1220–1227.
8. Harrison, J.J.; Ceri, H.; Stremick, C.; Turner, R.J. Differences in biofilm and planktonic cell mediated reduction of metalloid oxyanions. *FEMS Microbiol. Lett.* **2004**, *235*, 357–362.
9. Anderson, G.G.; Palermo, J.J.; Schilling, J.D.; Roth, R.; Heuser, J.; Hultgren, S.J. Intracellular bacterial-like pods in urinary tract infections. *Science* **2003**, *301*, 105–107.
10. Sauer, K.; Camper, A.K.; Ehrlich, G.D.; Costerton, J.W.; Davies, D.G. *Pseudomonas aeruginosa* displays multiple phenotypes during development as a biofilm. *J. Bacteriol.* **2002**, *184*, 1140–1154.
11. Kaplan, J.B.; Meyenhofer, M.M.; Fine, D.H. Biofilm growth and detachment of *Actinobacillus actinomycescomitans*. *J. Bacteriol.* **2003**, *185*, 1399–1404.
12. Ausschill, T.M.; Artweiler, N.B.; Netuschil, L.; Brex, M.; Reich, E.; Sculean, A. Spatial distribution of vital and dead microorganisms in dental biofilms. *Arch. Oral Microbiol.* **2001**, *46*, 471–476.
13. Jefferson, K.K. What drives bacteria to produce a biofilm? *FEMS Microbiol. Lett.* **2004**, *236*, 163–173.
14. Hammer, B.K.; Bassler, B.L. Quorum sensing controls biofilm formation in *Vibrio cholerae*. *Mol. Microbiol.* **2003**, *50*, 101–114.
15. Dasgupta, M.K. Biofilms and infection in dialysis patients. *Semin. Dial.* **2002**, *15*, 338–346.
16. Trautner, B.W.; Darouiche, R.O. Catheter associated infections: pathogenesis affects prevention. *Arch. Intern. Med.* **2004**, *164*, 842–850.
17. Hall-Stoodley, L.; Stoodley, P. Biofilm formation and dispersal and the transmission of human pathogens. *Trends Microbiol.* **2005**, *13*, 7–10.
18. Davey, M.E.; O'Toole, G.A. 2000 microbial biofilms: from ecology to molecular genetics. *Microbiol. Mol. Biol. Rev.* **2000**, *64*, 847–867.
19. O'Toole, G.A.; Kolter, R. Flagellar and twitching motility are necessary for *Pseudomonas aeruginosa* biofilm development. *Mol. Microbiol.* **1998**, *30*, 295–304.
20. O'Toole, G.A.; Gibbs, K.A.; Hager, P.W.; Phibbs, P.V.; Kolter, R. The global carbon metabolism regulator Crc is a component of a signal transduction pathway required for biofilm development by *Pseudomonas aeruginosa*. *J. Bacteriol.* **2000**, *182*, 425–431.
21. Danese, P.N.; Pratt, L.A.; Dove, S.; Kolter, R. The outer membrane protein ag43 mediates cell-cell interactions in *Escherichia coli* biofilms. *Mol. Microbiol.* **2000**, *37*, 424–432.
22. Genevaux, P.; Muller, S.; Bauda, P. A rapid screening procedure to identify mini Tn10 insertion mutants of *Escherichia coli* K-12 with altered adhesion properties. *FEMS Microbiol. Lett.* **1996**, *142*, 27–30.
23. Chia, J.-S.; Lee, Y.-Y.; Huang, P.-T.; Chen, J.Y. Identification of stress responsive genes in *Streptococcus mutans* by differential display reverse transcriptional-PCR. *Infect. Immun.* **2001**, *69*, 2493–2501.
24. Svensater, G.; Welin, J.C.; Wilkins, D.; Beighton, D.; Hamilton, I.R. Protein expression by planktonic and biofilm cells of *Streptococcus mutans*. *FEMS Microbiol. Lett.* **2001**, *205*, 139–146.
25. Zhu, M.; Takenaka, S.; Sato, M.; Hoshino, E. Influence of starvation and biofilm formation on acid resistance of *Streptococcus mutans*. *Oral Microbiol. Immunol.* **2001**, *16*, 24–27.
26. Parsek, M.; Greenberg, E.P. Sociobiology: the connections between quorum sensing and biofilms. *Trends Microbiol.* **2005**, *13*, 27–33.
27. Hentzer, M.; Wu, H.; Andersen, J.B.; Riedel, K.; Rasmussen, T.B.; Bagge, N.; Kumar, N.; Schembri, M.A.; Song, Z.; Kristofferson, P.; Manefield, M.; Costerton, J.W.; Molin, S.; Eberl, L.; Steinberg, P.; Kjellenberg, S.; Hoiby, N.; Givskov, M. Attenuation of *Pseudomonas aeruginosa* virulence by quorum sensing inhibitors. *EMBO J.* **2003**, *22*, 3803–3815.
28. Eberl, L.; Tummeler, B. *Pseudomonas aeruginosa* and *Burkholderia cepacia* in cystic fibrosis: genome evaluation, interactions and adaptation. *Int. J. Med. Microbiol.* **2004**, *294*, 123–131.
29. Juhas, M.; Eberl, L.; Tummeler, B. Quorum sensing: the power of cooperation in the world of *Pseudomonas*. *Environ. Microbiol.* **2005**, *7*, 459–471.
30. Withers, H.; Swift, S.; Williams, P. Quorum sensing as an integral component of gene regulatory networks in gram-negative bacteria. *Curr. Opin. Microbiol.* **2001**, *4*, 186–193.
31. Kim, E.-J.; Wang, W.; Deckwer, W.-D.; Zeng, A.-P. Expression of the quorum sensing regulatory protein LasR is strongly affected by iron and oxygen concentrations in cultures of *Pseudomonas aeruginosa* irrespective of cell density. *Microbiol.* **2005**, *151*, 1127–1138.
32. Peterson, F.C.; Pecharki, D.; Scheie, A.A. Biofilm mode of growth of *Streptococcus intermedius* favored by competence-stimulating signalling peptide. *J. Bacteriol.* **2004**, *186*, 6327–6331.
33. Aspiras, M.B.; Ellen, R.P.; Cvitkovitch, D.G. ComX activity of *Streptococcus mutans* growing in biofilms. *FEMS Microbiol. Lett.* **2004**, *238*, 167–174.
34. Suntharalingam, P.; Cvitkovitch, D.G. Quorum sensing in streptococcal biofilms. *Trends Microbiol.* **2005**, *13*, 3–6.
35. Branda, S.S.; Vik, A.; Friedman, L.; Kolter, R. Biofilms: the matrix revisited. *Trends Microbiol.* **2005**, *13*, 20–26.
36. Beloin, C.; Ghigo, J.-M. Finding gene expression patterns in bacterial biofilms. *Trends Microbiol.* **2005**, *13*, 16–19.
37. Thormann, K.M.; Saville, R.M.; Shukla, S.; Pelletier, D.A.; Spormann, A.M. Initial phases of biofilm formation in *Shewanella oneidensis* MR-1. *J. Bacteriol.* **2004**, *186*, 8096–8104.

38. Whiteley, M.; Bangera, M.G.; Bumgarner, R.E.; Parsek, M.R.; Teitzel, G.M.; Lory, S.; Greenberg, E.P. Gene expression in *Pseudomonas aeruginosa* biofilms. *Nature* **2001**, *413*, 860–864.
39. Hancock, L.E.; Perego, M. The *Enterococcus faecalis* fsr two-component system controls biofilm development through production of gelatinase. *J. Bacteriol.* **2004**, *186*, 5629–5639.
40. Carniol, K.; Gilmore, M. Signal transduction, quorum sensing, extracellular protease activity in *Enterococcus faecalis* biofilm formation. *J. Bacteriol.* **2004**, *186*, 8161–8163.
41. Jones, C.H.; Tove, C.; Bolken, K.F.; Jones, G.O.Z.; Dennis, E.H. Conserved DegP protease in gram-positive bacteria is essential for thermal and oxidative tolerance and full virulence in *Streptococcus pyogenes*. *Infect. Immun.* **2001**, *69*, 5538–5545.
42. Diaz-Torres, M.L.; Russell, R.R. HtrA protease and processing of extracellular proteins of *Streptococcus mutans*. *FEMS Microbiol. Lett.* **2001**, *204*, 23–28.
43. Lemos, A.C.; Yi-Ywan, M.C.; Burne, R.A. Genetic and physiologic analysis of the *groE* operon and role of the HrcA repressor in stress gene regulation and acid tolerance in *Streptococcus mutans*. *J. Bacteriol.* **2001**, *183*, 6074–6084.
44. Hormann, K.M.; Saville, R.M.; Shukla, S.; Spormann, A.M. Induction of rapid detachment in *Shewanella oneidensis* MR-1 biofilms. *J. Bacteriol.* **2005**, *187*, 1014–1021.
45. Kreft, J.-U. Biofilms promote altruism. *Microbiology* **2004**, *150*, 2751–2760.
46. Webb, J.S.; Lau, A.U.; Kjellerberg, S. Bacteriophage and phenotypic variation in *Pseudomonas aeruginosa* biofilm development. *J. Bacteriol.* **2004**, *186*, 8066–8073.
47. Baveja, J.K.; Willcox, M.D.P.; Hume, E.B.H.; Kumar, N.; Odell, R.; Poole-Warren, A. Furanones as potential antibacterial coatings on biomaterials. *Biomaterials* **2004**, *25*, 5003–5012.
48. Cicalini, S.; Palmieri, F.; Petrosillo, N. New technologies for prevention of intravascular catheter-related infections. *Crit. Care* **2004**, *8*, 157–162.
49. Flemming, R.G.; Murphy, C.J.; Abrams, G.A.; Goodman, S.L.; Nealey, P.F. Effects of synthetic micro- and nano-structured surfaces on cell behaviour. *Biomaterials* **1999**, *20*, 573–588.
50. Dalby, M.J.; Childs, S.; Riehle, M.O.; Johnstone, H.J.H.; Affrossman, S.; Curtis, A.S.G. Fibroblast reaction to island topography: changes in cytoskeleton and morphology with time. *Biomaterials* **2003**, *24*, 927–935.
51. Curtis, A.; Wilkinson, C. Nanotechniques and approaches in biotechnology. *Trends Biotechnol.* **2001**, *19*, 97–101.
52. Cousins, B.G.; Doherty, P.J.; Williams, R.L.; Fink, J.; Garvey, M.J. The effect of silica nanoparticulate coatings on cellular response. *J. Mater. Sci. Mater. Med.* **2004**, *15*, 353–357.
53. Kojic, E.M.; Darouiche, R.O. Candida infections on medical devices. *Clin. Microbiol. Rev.* **2004**, *17*, 255–267.
54. Daniels, R.; Vanderleyden, J.; Michiels, J. Quorum sensing and swarming migration in bacteria. *FEMS Microbiol. Rev.* **2004**, *28*, 261–289.
55. Friedman, L.; Kolter, R. Genes involved in matrix formation in *Pseudomonas aeruginosa* PA-14 biofilms. *Mol. Microbiol.* **2004**, *51*, 675–690.
56. Stoodley, P.; Sauer, K.; Davies, D.G.; Costerton, J.W. Biofilms as complex differentiated communities. *Ann. Rev. Microbiol.* **2002**, *56*, 187–209.
57. Beloin, C.; Valle, J.; Latour-Lambert, P.; Faure, P.; Kzreminski, M.; Balestrino, D.; Haagensen, J.A.J.; Molin, S.; Prensier, G.; Arbeille, B.; Ghigo, J.M. Global impact of mature biofilm lifestyle on *Escherichia coli* K-12 gene expression. *Mol. Microbiol.* **2004**, *51*, 659–674.
58. HauBler, S. Biofilm formation by the small colony variant phenotype of *Pseudomonas aeruginosa*. *Environ. Microbiol.* **2004**, *6*, 546–551.
59. Yarwood, J.M.; Bartels, D.J.; Volper, E.M.; Greenberg, E.P. Quorum sensing in *Staphylococcus aureus* biofilms. *J. Bacteriol.* **2004**, *186*, 1838–1850.
60. Lyte, M.; Freeston, P.P.E.; Neal, C.P.; Olson, B.A.; Haigh, R.D.; Brayston, R.; Williams, P.H. Stimulation of *Staphylococcus epidermidis* growth and biofilm formation by catecholamine inotropes. *Lancet* **2003**, *361*, 130–135.
61. Carmen, J.C.; Roeder, B.L.; Nelson, J.L.; Beckstead, B.L.; Runyan, C.M.; Schaalje, G.B.; Robison, R.A.; Pitt, W.G. Ultrasonically enhanced vancomycin activity against *Staphylococcus epidermidis*. *J. Biomater. Appl.* **2004**, *18*, 237–245.
62. Lerebour, G.; Cupferman, S.; Bellon-Fontaine, M.N. Adhesion of *Staphylococcus aureus* and *Staphylococcus epidermidis* to the Episkin reconstructed epidermis model and to an inert 304 stainless steel substrate. *J. Appl. Microbiol.* **2004**, *7*, 7–16.
63. Douglas, L.J. Candida biofilm and their role in infection. *Trends Microbiol.* **2003**, *11*, 30–36.
64. Fux, C.A.; Costerton, J.W.; Stewart, P.S.; Stoodley, P. Survival strategies of infectious biofilms. *Trends Microbiol.* **2005**, *13*, 34–40.
65. Marchais, V.; Kempf, M.; Licznar, P.; Lefrancois, C.; Bouchara, J-P; Robert, R.; Cottin, J. DNA array analysis of *Candida albicans* gene expression in response to adherence on polystyrene. *FEMS Microbiol. Lett.* **2005**, *245*, 25–32.
66. Keren, I.; Shah, D.; Spoering, A.; Kaldalu N. Lewis, K. Specialised persister cells and the mechanism of multi-drug tolerance in *Escherichia coli*. *J. Bacteriol.* **2004**, *186*, 8172–8180.

Biofunctional Polymers

Jennifer L. West

Department of Bioengineering, Rice University, Houston, Texas, U.S.A.

INTRODUCTION

In many applications utilizing polymeric biomaterials, such as in scaffolds for tissue engineering, investigators can choose from natural materials, such as collagen or fibrin, or synthetic polymers such as poly(lactic-co-glycolic acid) (PLGA). Each choice has its unique advantages and disadvantages. For example, the extracellular matrix (ECM) proteins have evolved over billions of years to have specific cellular interactions and activities, to be remodeled by cellular activity, and even to guide tissue formation processes. However, the isolation, purification, and processing of these materials can be difficult, mechanical properties are often poor, and there are some concerns about the possibility for disease transmission. Synthetic polymers, on the other hand, offer easier handling and processing as well as a greater degree of control of material properties. Unfortunately, when using synthetic materials such as PLGA, cellular interactions with the material are mediated by nonspecifically adsorbed proteins and are generally difficult to control. In order to obtain some of the advantages of natural bioactive compounds, such as ECM proteins, while gaining the ease of processing and safety of synthetic materials, a number of researchers have developed hybrid materials that are primarily composed of standard synthetic polymers but are also grafted, tethered to, or copolymerized with biologically functional segments such as peptides. Examples discussed in this article include the development of ECM-mimetic materials for use in tissue engineering applications, as well as the development of pharmacologically active biomaterials.

ECM-MIMETIC MATERIALS

In natural tissue, cells grow within a scaffold consisting of a variety of proteins and polysaccharides—the extracellular matrix (ECM). The ECM does far more than just provide mechanical support for the cells, which is all that many of the tissue engineering scaffolds in current use are capable of doing. Rather, the ECM has direct biological interactions with the entrapped cells, influencing cell growth, migration, morphology, and differentiation. Using natural polymers

such as collagen or fibrin gels allows one to reconstitute some of these interactions, but controlling these interactions and other properties of the materials is difficult. Modifying synthetic polymers that have the appropriate mechanical properties and processing conditions with biofunctional moieties that provide some of the ECM interactions may allow one to get the benefits of a synthetic material with many of the advantages normally associated with natural materials.

Cell-Adhesive Materials

In tissues, cell-adhesive interactions occur between cell surface receptors, such as the integrins, and specific ligands on ECM proteins. A number of short peptide sequences derived from adhesive ECM proteins have been identified that are able to bind to cell surface receptors and mediate cell adhesion with affinity and specificity similar to that obtained with intact proteins. Oligopeptides are preferable to intact proteins because they are not subject to denaturation and may be less susceptible to proteolysis. Some of these, along with their biological sources, are shown in Table 1. A number of methods for incorporating peptides into biomaterials have been investigated, including physicochemical adsorption, chemisorption, and covalent attachment. Only covalent attachment will be addressed in this article. Covalent attachment may be in the form of surface grafting, grafting to polymer chains, or incorporation into polymer backbones. Important considerations in designing these types of biohybrid materials include the hydrophobicity of the synthetic substrate, steric hindrance, peptide orientation, and whether interactions need to be two- or three-dimensional. When modifying materials with adhesive peptides, however, one must be aware that activation of these cell-signaling pathways can alter many other cellular activities, including proliferation, migration, and synthesis of new ECM.^[1,2]

Perhaps the most extensively studied cell adhesion peptide is the sequence Arg–Gly–Asp (RGD). It is found in many cell adhesion proteins and binds to integrin receptors on a wide variety of cell types.^[3] Modification with RGD peptides may enhance cell adhesion to a biomaterial and activate integrin

Table 1 Peptide sequences that can mediate cell adhesion

Peptide sequence	ECM protein
RGD	Fibronectin, vitronectin, laminin, collagen
YIGSR	Laminin
IKVAV	Laminin
REDV	Fibronectin
DGEA	Collagen
VGVPAG	Elastin

signaling pathways. A number of studies have been performed with RGD-containing peptides grafted to polymers. In initial studies, RGD peptides were immobilized on a polymer-modified glass substrate, and cell adhesion parameters were correlated with peptide density.^[4] A peptide density of 10 fmol/cm² was sufficient to support adhesion and spreading of fibroblast cells, clustering of integrin receptors, and organization of actin stress fibers. At 1 fmol/cm² the cells were fully spread, but the morphology of the stress fibers was abnormal, and the cells did not form focal contacts.

In addition to grafting RGD peptides to polymeric surfaces, peptides can be incorporated into polymer matrices to allow adhesion throughout the material. For example, a biodegradable copolymer of polylysine and poly(lactic acid) has been used for the attachment of RGD peptides.^[5] These types of materials can be processed into highly porous sponges to allow cells to permeate into the matrix and adhere throughout.^[6,7] When these copolymers were modified with RGD peptides, significant enhancement of cell adhesion was observed.

When designing peptide-modified materials, it is often desirable to start with intrinsically cell-non-adhesive polymers, effectively providing a blank slate upon which one can build desired biological interactions. Polyethylene glycol (PEG)-based materials, such as photocross-linked PEG hydrogels, have been particularly interesting for this approach.^[2,8,9] Figure 1 shows the interaction of fibroblasts with an unmodified PEG hydrogel and with a RGD-containing PEG hydrogel, demonstrating adhesion and spreading only with inclusion of the adhesion ligand. In addition to providing controlled and specific biological interactions, modification of cell-nonadhesive materials can allow one to generate cell-specific materials, i.e., materials that are adhesive only to certain cell types, if adhesion ligands with appropriate selectivity are utilized. For example, the peptide REDV has been shown to be adhesive to endothelial cells but not to smooth muscle cells, fibroblasts, or platelets,^[10] and thus it may be useful in the development of vascular grafts.

Proteolytically Degradable Polymers

The majority of biomaterials currently used in tissue engineering applications are designed to degrade via hydrolysis. With many of these materials, such as poly(lactide-co-glycolide), the degradation rate can be tailored to some degree by altering factors such as the copolymer composition, degree of crystallinity, or initial molecular weight. However, in these types of applications, it can be very difficult to predict a priori what the rate of tissue formation will be and, thus, what the rate of polymer degradation should be. In the ECM, during processes such as wound healing, the rate of scaffold (ECM) degradation is precisely matched to the rate of tissue formation. ECM is degraded by proteolytic enzymes, such as the matrix metalloproteases and plasmin, produced by cells during cell migration and tissue remodeling.^[11] Many of the natural polymer scaffolds, such as collagen and fibrin, are degraded by the same cellular mechanisms. Recently, synthetic biomaterials have been designed that are similarly degraded by cellular proteolytic activity.^[2,12] These materials are block copolymers of synthetic polymers, such as polyethylene glycol (PEG), and peptides that are substrates for proteolytic enzymes. When exposed to solutions containing the targeted protease, PEG-peptide copolymer hydrogels degraded, whereas they remained stable in the absence of protease or in the presence of a nontargeted protease.

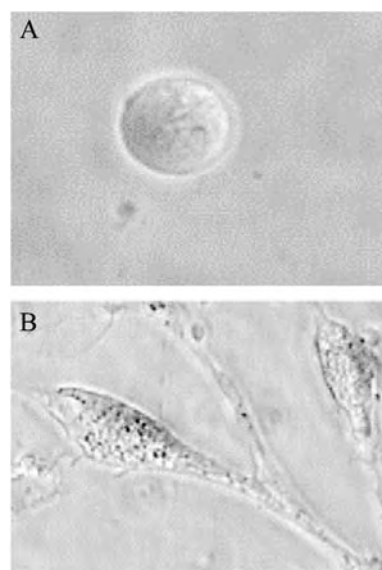


Fig. 1 Fibroblasts were seeded on the surfaces of either PEG hydrogels (A) or PEG hydrogels with tethered RGD peptide (B). Cells are able to adhere to and spread on these materials only in the presence of an immobilized adhesion ligand. If cell-selective adhesion ligands were used, one could generate a biomaterial that would allow adhesion of only certain cell types.

Furthermore, cells can migrate through PEG-peptide hydrogels that contain both proteolytically degradable peptide sequences and cell-adhesive sequences.^[13] In this case, degradable sequences are required so that the cells can create pathways through the solid structure for migration, and the cell adhesion ligands are required so that cells can develop the traction force required for migration.

PHARMACOLOGICALLY ACTIVE BIOMATERIALS

Pharmacologically active polymers are another class of biofunctional materials, with the goal of being used for prolonged, site-specific pharmacotherapy. Covalent modification of materials by means of pharmacological agents has been investigated for a number of applications. Most commonly, the objective is simply to develop a more biocompatible biomaterial, for instance, to improve blood compatibility. In tissue engineering applications, the goal may be to stimulate or inhibit specific cellular processes, such as proliferation or ECM synthesis. The majority of applications of pharmacologically active biomaterials have been performed by grafting the pharmacological agent to the synthetic polymer substrate, though other material designs are possible. A wide range of pharmacological agents have been employed, including polysaccharides, polypeptide growth factors, enzymes, anti-inflammatory agents, and nitric oxide.

Polysaccharides

Heparin has been the most extensively studied polysaccharide for biomaterials modification. Heparin is a polymer of heterogeneously sulfated L-iduronic acid and D-glucosamine, with molecular weights ranging from 2–50 kDa. Heparin has been commonly used as an anticoagulant and antiplatelet agent due to its binding and potent activation of antithrombin III.^[14] Covalent immobilization of heparin onto blood-contacting materials should provide thromboresistance for prolonged periods of time. Numerous synthetic polymers have been covalently heparinized, including polyurethanes,^[15,16] polyethylene terephthalate,^[17] polyvinyl alcohol,^[18,19] poly(dimethylsiloxane),^[20] polypyrrole,^[21] and polylactide.^[22]

Synthesis of covalently heparinized materials often involves the use of a hydrophilic spacer molecule in order to increase the bioavailability of the bound heparin. PEG is often used as a spacer, usually being coupled to the biomaterial and the heparin through diisocyanate groups. PEG has several advantages for this application. PEG is widely known to be highly

resistant to platelet adhesion, so the PEG spacers themselves help to create a more hemocompatible surface. Additionally, when heparin is immobilized through a PEG spacer, it retains its ability to bind both thrombin and antithrombin III, though to a slightly lesser extent than unmodified heparin.^[23] When directly immobilized, however, the heparin is able to bind only to thrombin.

Hyaluronic acid, a linear copolymer of D-glucuronic acid and N-acetylglucosamine, is also generating interest for biomaterials applications. It has been shown to be both antithrombotic and angiogenic.^[24–27]

Hirudin

Hirudin is a 65-amino acid polypeptide, originally derived from the salivary glands of the leech, that acts as an anticoagulant through its ability to directly inhibit thrombin. Both recombinant hirudin and synthetic hirudin are now available. Hirudin has been covalently attached to the surfaces of PET and polyurethane vascular grafts, using bovine serum albumin as a spacer, in order to improve their blood compatibility.^[28,29] Hirudin immobilized in this manner retained its ability to bind and inhibit thrombin. Hirudin has also been directly coupled to PLGA using glutaraldehyde, demonstrating improvements in several aspects of blood compatibility.^[30] Unfortunately, the binding between hirudin and thrombin is irreversible, leading to decreased bioactivity over time.^[30]

Plasminogen Activators

During fibrinolysis, circulating plasminogen is converted to the active enzyme plasmin by plasminogen activators, as depicted in Fig. 2. Plasmin is then able to degrade fibrin to lyse the clot structure (Fig. 2). Plasminogen activators include tissue plasminogen activator (tPA), urokinase, streptokinase, and lumbrakinase. Immobilization of plasminogen activators on blood-contacting materials may be a mechanism to impart thromboresistance. Additionally, immobilization of these enzymes may stabilize them, imparting

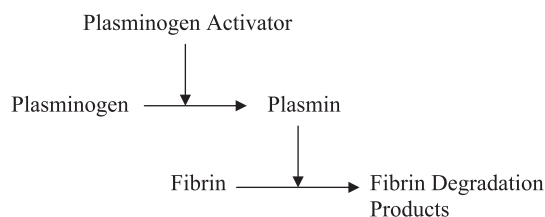


Fig. 2 Plasminogen activators convert circulating plasminogen to the active enzyme, plasmin, which then lyses fibrin clots.

greater resistance to temperature, pH, and proteolytic activity.^[31]

Urokinase, a plasminogen activator expressed in many tissue types, has been immobilized to a wide variety of polymeric materials including polyurethanes,^[32] poly(tetrafluoroethylene),^[33] poly(vinylchloride),^[34] and poly(hydroxyethyl methacrylate).^[35] Thrombus formation on these materials has been shown to be greatly reduced with enzyme immobilization. Similarly, lumbrinase, derived from an earthworm, and streptokinase, derived from streptococcal bacteria, have been immobilized to polymeric materials and shown to improve blood compatibility.^[36,37] However, there are significant concerns regarding adverse immunological interactions with materials modified with these foreign proteins.

Growth Factors

Growth factors are polypeptides involved in the regulation of a variety of cellular activities such as growth and differentiation, as well as many metabolic processes. Many different growth factors have been described in the literature, with potential applications ranging from accelerated wound healing to the induction of bone growth into implanted biomaterials. Modification of biomaterials with growth factors may target activity to the desired tissue and may maintain biological function for a prolonged period of time. Growth factors that have been covalently immobilized to polymeric biomaterials include epidermal growth factor (EGF), basic fibroblast growth factor (bFGF), and transforming growth factor-beta (TGF-B).

EGF is a 6 kDa polypeptide that is mitogenic and chemotactic for many of the cell types involved in wound healing. EGF was also the first growth factor that was shown to remain active after covalent immobilization. In that first study, EGF was immobilized to aminated glass via a star PEG spacer molecule and was shown to retain mitogenic activity, whereas physically adsorbed EGF showed no activity.^[38] Prior to this work, it was believed that receptor dimerization and internalization were required for growth factor activity, and thus that immobilization would not be effective. However, that does not appear to be true. More recently, EGF has also been immobilized on poly(methyl methacrylate)^[39] and polystyrene,^[40] with good retention of bioactivity.

bFGF is a 16 kDa polypeptide that is mitogenic for many cell types and that appears to play an important role in angiogenesis. bFGF binds strongly to heparin, so heparinized materials have been widely used for the complexation and presentation of bFGF.^[41–43] For example, polyurethane vascular grafts with

coimmobilized heparin and bFGF have been shown to support enhanced endothelialization.^[42]

Although in many cell types it does not stimulate growth, TGF-B plays a very important role in regulating ECM metabolism. For example, synthesis of collagen and elastin is greatly upregulated in vascular smooth muscle cells upon exposure to TGF-B. In tissue engineered vascular grafts, a potential strategy to improve mechanical properties is to increase ECM deposition. Thus, PEG-based hydrogel scaffolds have been synthesized with TGF-B tethered to the hydrogel network.^[44] Tethering TGF-B to the polymeric biomaterial actually enhanced its ability, relative to soluble TGF-B, to stimulate ECM protein production by encapsulated vascular smooth muscle cells. This effect may be due to the ability of the cells to interact with TGF-B via the appropriate receptors, while not being able to internalize it. In addition, engineered tissues formed with tethered TGF-B in the scaffold formulation were shown to have improved tensile strength.^[44]

Anti-Inflammatory Biomaterials

Salicylic acid, the active component of aspirin, is commonly used as an anti-inflammatory agent in a broad spectrum of medical applications. Polymeric materials have been synthesized with salicylic acid incorporated into the backbone of the polymer.^[45–47] These materials are polyanhydride esters composed of alkyl chains linked to salicylic acid by ester bonds. Hydrolytic degradation occurs as the ester bonds are cleaved, producing salicylic acid. These materials may be useful in wound closure and wound healing applications, where they could not only release an anti-inflammatory agent but could also serve as a biodegradable scaffold for tissue regeneration. This material could also have use in the treatment of Crohn's disease, a chronic inflammation of the intestine. In this application, the release of salicylic acid following oral ingestion of the polymer would be enhanced in the intestinal tract, due to enhanced polymer degradation at basic pH.

Nitric Oxide-Generating Materials

Nitric oxide (NO) acts on most tissues in the body through its ability to upregulate cGMP.^[48] For example, in the vascular system, NO inhibits platelet adhesion and aggregation, increases endothelial cell proliferation, and suppresses smooth muscle cell proliferation. Thus, NO is attractive for use as a therapeutic agent for the prevention of restenosis or for the improvement of vascular graft or stent performance. A number of small-molecule drugs, referred to as NO

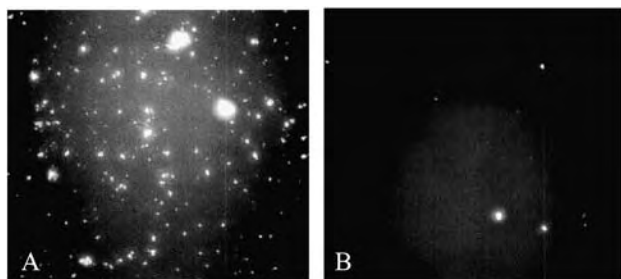


Fig. 3 Pretreatment of mepacrine-labeled whole blood with NO-producing hydrogel materials significantly reduced platelet adhesion to a model thrombogenic surface, collagen-coated glass. A, control hydrogel; B, NO-generating hydrogel.

donors, have been developed that undergo hydrolysis to generate NO in the body.^[49] However, systemic NO therapy for applications such as restenosis has generally failed, due to the difficulties in achieving therapeutic dosages at the appropriate site without incurring substantial systemic side effects.^[50] Bioactive polymers containing NO donor groups may allow the creation of local delivery vehicles for NO therapy, or even NO-generating medical devices.

Diazeniumdiolates have been attached to pendant groups of a polymer chain to provide NO release for up to five weeks.^[51] Diazeniumdiolates attached to crosslinked polyethyleneimine (PEI) inhibited the proliferation of smooth muscle cells in vitro. PTFE vascular grafts were also coated with the PEI-NO polymer and implanted in a baboon arteriovenous shunt model. Platelet adhesion to the NO-releasing grafts was reduced almost five-fold relative to control biomaterials. Microspheres can also be formed from the PEI-NO material.^[52] Diazeniumdiolate NO donors have also been covalently bound to polyurethane and polyvinyl chloride films.^[53] NO release from these materials ranged from 10 to 72 hr, and the NO-modified polymers successfully reduced platelet adhesion in vitro.

Hydrogel materials have been formed by photocrosslinking copolymers of PEG with NO donors, such as diazeniumdiolates and S-nitrosothiols, to achieve NO delivery for periods ranging from hours to months.^[54] Exposure of blood to NO-generating hydrogels dramatically decreased platelet adhesion to collagen-coated surfaces, as shown in Fig. 3. Furthermore, these materials dramatically reduced smooth muscle cell proliferation. These types of hydrogel materials can be polymerized in situ to form thin coatings on the luminal surface of a blood vessel following angioplasty or stent deployment,^[55] which provides highly localized treatment for the prevention of thrombosis and restenosis.

CONCLUSION

The development of bioactive biomaterials is expanding the capabilities and performance of materials used in a variety of biomedical applications. In particular, the development of ECM-mimetic polymers may significantly improve the ability to manipulate cell differentiation and behaviors while growing engineered tissues. As advances are made in cellular and molecular biology, the functionality that can be built into these hybrid materials should greatly expand, hopefully providing solutions to many currently unmet medical needs, such as synthetic small-diameter vascular grafts.

ARTICLES OF FURTHER INTEREST

Biologic and Synthetic Apatites; Biomaterials: Protein-Surface Interactions; Biomimetic Materials; Biorubber/Poly(Glycerol Sebacate); Drug Delivery, Controlled; Elastomers, Biodegradable; Extracellular Matrix Scaffolds; Heparin-Binding to Improve Biocompatibility; Nitric Oxide; Photopolymerization; Surface Coatings; Thrombosis

REFERENCES

1. Mann, B.K.; Tsai, A.T.; Scott-Burden, T.; West, J.L. Modification of surfaces with cell adhesion peptides alters extracellular matrix deposition. *Biomaterials* **1999**, *20*, 2281–2286.
2. Mann, B.K.; Gobin, A.S.; Tsai, A.T.; Schmedlen, R.H.; West, J.L. Smooth muscle cell growth in photopolymerized hydrogels with cell adhesive and proteolytically degradable domains: Synthetic ECM analogs for tissue engineering. *Biomaterials* **2001**, *22*, 3045–3051.
3. Humphries, M.J. The molecular basis and specificity of integrin-ligand interactions. *J. Cell Sci.* **1990**, *97*, 585–592.
4. Massia, S.P.; Hubbell, J.A. An RGD spacing of 440 nm is sufficient for integrin mediated fibroblast spreading and 140 nm for focal contact and stress fiber formation. *J. Cell Biol.* **1991**, *114*, 1089–1100.
5. Barrera, D.A.; Zylestra, E.; Lansbury, P.T.; Langer, R. Synthesis and RGD peptide modification of a new biodegradable polymer—Poly(lactic acid-co-lysine). *J. Am. Chem. Soc.* **1993**, *115*, 11010–11011.
6. Mikos, A.G.; Thorsen, A.J.; Czerwonka, L.A.; Bao, Y.; Langer, R.; Winslow, D.N.; Vacanti, J.P. Preparation and characterization of poly(L-lactide) foams. *Polymer* **1994**, *35*, 1068–1077.
7. Mooney, D.J.; Park, S.; Kaufmann, P.M.; Sano, K.; McNamara, K.; Vacanti, J.P.; Langer, R. Biodegradable sponges for hepatocyte transplantation. *J. Biomed. Mater. Res.* **1995**, *29*, 959–965.

8. Hern, D.L.; Hubbell, J.A. Incorporation of adhesion peptides into nonadhesive hydrogels useful in tissue resurfacing. *J. Biomed. Mater. Res.* **1998**, *39*, 266–277.
9. Mann, B.K.; West, J.L. Cell adhesion peptides alter smooth muscle cell adhesion, proliferation, and matrix protein synthesis on modified surfaces and in polymer scaffolds. *J. Biomed. Mater. Res.* **2002**, *60*, 86–93.
10. Hubbell, J.A.; Massia, S.P.; Drumheller, P.D. Endothelial cell-selective tissue engineering in the vascular graft via a new receptor. *Bio/Technology* **1991**, *9*, 568–572.
11. Rabbani, S.A. Metalloproteases and urokinase in angiogenesis and tumor progression. *In Vivo* **1998**, *12*, 135–142.
12. West, J.L.; Hubbell, J.A. Polymeric biomaterials with degradation sites for proteases involved in cell migration. *Macromolecules* **1999**, *32*, 241–244.
13. Gobin, A.S.; West, J.L. Cell migration through defined synthetic ECM analogs. *FASEB J.* **2002**, *16*, 751–753.
14. Barritt, D.W.; Jordan, S.C. Anticoagulant drugs in the treatment of pulmonary embolism: A controlled clinical trial. *Lancet* **1960**, *1*, 1309–1312.
15. Han, D.K.; Jeong, S.Y.; Kim, Y.H. Evaluation of blood compatibility of PEO grafted and heparin immobilized polyurethanes. *J. Biomed. Mater. Res.* **1989**, *23*, 211–228.
16. Han, D.K.; Park, K.D.; Ahn, K.D.; Kim, Y.H. Preparation and surface characterization of PEO-grafted and heparin immobilized polyurethanes. *J. Biomed. Mater. Res.* **1989**, *23*, 87–104.
17. Kim, Y.J.; Kang, I.K.; Muh, M.W.; Yoon, S.C. Surface characterization and in vitro blood compatibility of poly(ethylene terephthalate) immobilized with insulin and/or heparin using plasma glow discharge. *Biomaterials* **2000**, *21*, 121–130.
18. Cholakakis, C.H.; Sefton, M.V. In vitro platelet interactions with a heparin-polyvinyl alcohol hydrogel. *J. Biomed. Mater. Res.* **1989**, *23*, 399–415.
19. Piao, A.Z.; Jacobs, H.A.; Park, K.D.; Kim, S.W. Heparin immobilization by surface amplification. *ASAIO J.* **1992**, *38*, M638–M643.
20. Grainger, D.; Feijen, J.; Kim, S.W. Poly(dimethyl siloxane)-poly(ethylene oxide)-heparin block copolymers. I: synthesis and characterization. *J. Biomed. Mater. Res.* **1988**, *22*, 231–242.
21. Garner, B.; Georgevich, B.; Hodgson, A.J.; Liu, L.; Wallace, G.G. Polypyrrole-heparin composites as stimulus-responsive substrates for endothelial cell growth. *J. Biomed. Mater. Res.* **1999**, *44*, 121–129.
22. Seifert, B.; Groth, T.; Hermann, K.; Romaniuk, P. Immobilization of heparin on polylactide for application to degradable biomaterials in contact with blood. *J. Biomater. Sci., Polym. Ed.* **1995**, *7*, 277–287.
23. Byun, Y.; Jacobs, H.A.; Kim, S.W. Heparin surface immobilization through hydrophilic spacers: thrombin and antithrombin III binding kinetics. *J. Biomater. Sci. Polym. Ed.* **1994**, *6*, 1–13.
24. West, D.C.; Hampson, I.N.; Arnold, F.; Kuman, S. Angiogenesis induced by degradation products of hyaluronic acid. *Science* **1985**, *228*, 1324–1326.
25. Vercauteren, K.P.; Marecak, D.M.; Marecak, J.F.; Prestwich, G.D. Synthesis and in vitro degradation of new polyvalent hydrazide crosslinked hydrogels of hyaluronic acid. *Bioconjug. Chem.* **1997**, *8*, 686–694.
26. Verheye, S.; Markou, C.P.; Salame, M.Y.; Wan, B.; King, S.B.; Robinson, K.A.; Chronos, N.A.; Hanson, S.R. Reduced thrombus formation by hyaluronic acid coating of endovascular devices. *Arterioscler. Thromb. Vasc. Biol.* **2000**, *20*, 1168–1172.
27. Barbucci, R.; Magnani, A.; Rappuoli, R.; Lamponi, S.; Consumi, M. Immobilization of sulfated hyaluronan for improved biocompatibility. *J. Inorg. Biochem.* **2000**, *79*, 119–125.
28. Ito, R.K.; Phaneuf, M.D.; LoGerfo, F.W. Thrombin inhibition by covalently bound hirudin. *Blood Coagul. Fibrinolysis.* **1991**, *2*, 77–81.
29. Phaneuf, M.D.; Berceci, S.A.; Bide, M.J.; Quist, W.C.; LoGerfo, F.W. Covalent linkage of recombinant hirudin to poly(ethylene terephthalate) (Dacron): Creation of a novel antithrombin surface. *Biomaterials* **1997**, *18*, 755–765.
30. Seifert, B.; Romaniuk, P.; Groth, T. Covalent immobilization of hirudin improves the haemocompatibility of polylactide-polyglycolide in vitro. *Biomaterials* **1997**, *18*, 1495–1502.
31. Torchilin, V.P.; Makismenko, A.V.; Mazaev, A.V. Immobilized enzymes for thrombolytic therapy. *Meth. Enzym.* **1988**, *137*, 552–566.
32. Kitamoto, Y.; Tomita, M.; Kiyama, S.; Inoue, T.; Yabushita, Y.; Sato, T.; Ryoda, H. Antithrombotic mechanisms of urokinase immobilized polyurethane. *Thrombosis* **1991**, *65*, 73–76.
33. Forster, R.I.; Bernath, F. Analysis of urokinase immobilization on the polytetrafluoroethylene vascular prosthesis. *Am. J. Surg.* **1988**, *156*, 130–132.
34. Sugitachi, A.; Tanaka, M.; Kawahara, T.; Kitamura, N.; Takagi, K. A new type of drain tube. *Artif. Organs* **1981**, *5*, 69.
35. Liu, L.S.; Ito, Y.; Imanishi, Y. Biological activity of urokinase immobilized to cross linked poly(2-hydroxyethyl methacrylate). *Biomaterials* **1991**, *12*, 545–549.
36. Drummond, R.K.; Peppas, N.A. Fibrinolytic behavior of streptokinase-immobilized poly(methacrylic acid-g-ethylene oxide). *Biomaterials* **1991**, *12*, 356–360.
37. Ryu, G.H.; Park, S.; Kim, M.; Han, D.K.; Kim, Y.H.; Min, B. Antithrombogenicity of lumbrokinase-immobilized polyurethane. *J. Biomed. Mater. Res.* **1994**, *28*, 1069–1077.
38. Kuhl, P.R.; Griffith-Cima, L.G. Tethered epidermal growth factor as a paradigm for growth factor-induced stimulation from the solid phase. *Nat. Med.* **1996**, *2*, 1022–1027.
39. Ito, Y.; Li, J.S.; Takahashi, T.; Imanishi, Y.; Okabayashi, Y.; Kido, Y.; Kasuga, M. Enhancement of the mitogenic effect by artificial juxtacrine stimulation using immobilized EGF. *J. Biochem.* **1997**, *121*, 514–520.
40. Ito, Y.; Chen, G.; Imanishi, Y. Micropatterned immobilization of epidermal growth factor to regulate cell function. *Bioconjug. Chem.* **1998**, *9*, 277–282.
41. Bos, G.W.; Scharenborg, N.M.; Poot, A.A.; Engbers, G.H.M.; Beugling, T.; van Aken, W.G.; Feijen, J.

- Proliferation of endothelial cells on surface-immobilized albumin-heparin conjugate loaded with basic fibroblast growth factor. *J. Biomed. Mater. Res.* **1999**, *44*, 330–340.
42. Doi, K.; Matsuda, T. Enhanced vascularization in a microporous polyurethane graft impregnated with basic fibroblast growth factor and heparin. *J. Biomed. Mater. Res.* **1997**, *34*, 361–370.
 43. Wissink, M.J.B.; Beernink, R.; Poot, A.A.; Engbers, G.H.M.; Beugling, T.; van Aken, W.G.; Feijen, J. Improved endothelialization of vascular grafts by local release of growth factor from heparinized collagen matrices. *J. Control. Release* **2000**, *64*, 103–114.
 44. Mann, B.K.; Schmedlen, R.H.; West, J.L. TGF-beta increases extracellular matrix production of vascular smooth muscle cells. *Biomaterials* **2001**, *22*, 439–444.
 45. Erdmann, L.; Campe, C.; Palms, D.; Uhrich, K. Polymer prodrugs with pharmaceutically active degradation products. *Polym. Prepr.* **1997**, *41*, 1048–1049.
 46. Erdmann, L.; Campe, C.; Bedell, C.; Uhrich, K. Polymeric prodrugs: Novel polymers with bioactive components. *ACS Symp. Ser.* **1998**, *709*, 83–91.
 47. Krogh-Jespersen, E.; Anastasiou, T.; Uhrich, K. Synthesis of a novel aromatic polyanhydride containing aminosalicyclic acid. *Polym. Prepr.* **2000**, *41*, 1048–1049.
 48. Feldman, P.L.; Griffith, O.W.; Stuehr, D.J. The surprising life of nitric oxide. *Chem. Eng. News* **1993**, *71*, 26–38.
 49. Maragos, C.M.; Morley, D.; Wink, D.A.; Dunams, T.M.; Saavedra, J.E.; Hoffman, A.; Bove, A.A.; Isaac, L.; Hrabie, J.A.; Keefer, L.K. Complexes of NO with nucleophiles as agents for the controlled biological release of nitric oxide—Vasorelaxant effects. *J. Med. Chem.* **1991**, *34*, 3242–3247.
 50. Loscalzo, J. Nitric oxide and restenosis. *Clin. Appl. Thromb./Hemost.* **1996**, *2*, 7–10.
 51. Smith, D.J.; Chakravarthy, D.; Pulfer, S.; Simmons, M.L.; Hrabie, J.A.; Citro, M.E.; Saavedra, J.E.; Davies, K.M.; Hutsell, T.C.; Mooradian, D.L.; Hanson, S.R.; Keefer, L.K. Nitric oxide-releasing polymers containing the [N(O)NO]-group. *J. Med. Chem.* **1996**, *39*, 1148–1156.
 52. Pulfer, S.K.; Ott, D.; Smith, D.J. Incorporation of nitric oxide-releasing crosslinked polyethyleneimine microspheres into vascular grafts. *J. Biomed. Mater. Res.* **1997**, *37*, 182–189.
 53. Mowery, K.A.; Schoenfisch, M.H.; Saavedra, J.E.; Keefer, L.K.; Meyerhoff, M.E. Preparation and characterization of hydrophobic polymeric films that are thromboresistant via nitric oxide release. *Biomaterials* **2000**, *21*, 9–21.
 54. Bohl, K.S.; West, J.L. Nitric oxide-generating polymers reduce platelet adhesion and smooth muscle cell proliferation. *Biomaterials* **2000**, *21*, 2273–2278.
 55. Hill-West, J.L.; Chowdhury, S.M.; Slepian, M.J.; Hubbell, J.A. Inhibition of thrombosis and intimal thickening by in situ photopolymerization of thin hydrogel barriers. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 5967–5971.

Biologic and Synthetic Apatites

Besim Ben-Nissan

Department of Chemistry Materials and Forensic Science, University of Technology, Sydney, NSW, Australia

Racquel LeGeros

Department of Biomaterials and Biomimetics, L. Linkow Professor of Implant Dentistry, Director, Calcium Phosphate Laboratory, New York University College of Dentistry, New York, New York, U.S.A

B

INTRODUCTION

In the early seventies, bioceramics were employed to perform singular, biologically inert roles, such as to provide parts for bone replacement. The realization that cells and tissues perform many other vital regulatory and metabolic roles has highlighted the limitations of synthetic materials as tissue substitutes. Demands of bioceramics have changed from maintaining an essentially physical function without eliciting a host response, to providing a more integrated interaction with the host. This has been accompanied by increasing demands from medical devices to improve the quality of life, as well as extend its duration. Bioceramics, especially hydroxyapatite, incorporating biological additives can be used as body-interactive materials, helping the body to heal, or promoting regeneration of tissues, thus restoring physiological functions.

The crystallography and characterization of biologic and synthetic apatites are very complex. This review attempts to cover over four decades of research on one of the most intriguing and fascinating fields of research.

BACKGROUND

Apatite (Gr, to deceive) was the name given first by Werner in 1788 to describe a group of mineral crystals appearing with various tints (yellow, green, pink, etc.) that were often confused with more precious minerals or gems, such as aquamarine, amethyst, topaz, etc. These minerals have the general formula $M_{10}(RO_4)_6X_2$, where M could be one of several metals (usually calcium, Ca), R is most commonly phosphorus (P), and X is commonly hydroxide (OH), or a halogen such as fluorine (F) or chlorine (Cl). In later years and currently, the name “apatite” describes a family of compounds having similar structure (hexagonal system, space group, $P6_3/m$) in spite of a wide range of composition.^[1,2] The unit cell (smallest building unit) of calcium hydroxyapatite, HA, contains ten calcium (Ca), six PO_4 , and two OH groups, arranged as shown

Fig. 1.^[3] The OH groups located in the corners of the unit cell are surrounded by two sets of Ca (II) atoms arranged in a triangular patterns at levels $z=0$ and $z=1/2$; by two sets of PO_4 tetrahedra, also arranged in triangular patterns at levels $z=1/4$ and $z=3/4$, and by a hexagonal array of Ca (I) atoms at a distance.^[3] Critical to the apatite structure is the network of PO_4 groups in tightly packed arrangements.^[2] Substitutions in the apatite structure affect lattice parameters (a - and c -axis dimensions), infrared absorption characteristics, morphology, dissolution properties, and thermal stabilities.

Biologic apatites are the inorganic phases of calcified tissues (teeth and bones). Similarity in composition of calcined bone to the apatite mineral was proposed by Proust and Klaproth in 1788^[4] and similarity in the x-ray diffraction patterns of bone and mineral apatites (hydroxyapatite, HA and fluorapatite, FA) and similarity in composition (principally, calcium and phosphate ions) led to the conclusion that the inorganic phases of bones and teeth are basically calcium hydroxyapatite.^[5,7] Detection of carbonate associated with biologic apatites led to the speculation that these mineral phases are carbonate-containing apatites similar to the minerals dahllite (carbonate-containing apatite) or staffellite (carbonate and fluoride-containing apatite).^[8] Studies on synthetic carbonate-substituted apatites demonstrated that carbonate substitution in the apatite structure can proceed in two ways: CO_3 -for-OH, or Type A,^[9,10] and CO_3 -for- PO_4 or Type B, coupled with Na-for-Ca^[11,12] and combined analyses of synthetic and biologic apatites using x-ray diffraction, infrared spectroscopy and chemical analyses demonstrated that biologic apatites are carbonate apatites approximated by the formula, $(Ca,Mg,Na)_{10}(PO_4,CO_3, HPO_4)_6(CO_3, OH)_2$.^[12-14]

Early studies on synthetic apatites, and related calcium phosphates, were made to gain a better understanding of the structure, composition and properties of biologic apatites, especially of human enamel apatites. However, studies on synthetic apatites in the last 25 years had focused on their preparation, their applications in medicine and dentistry, and their use as

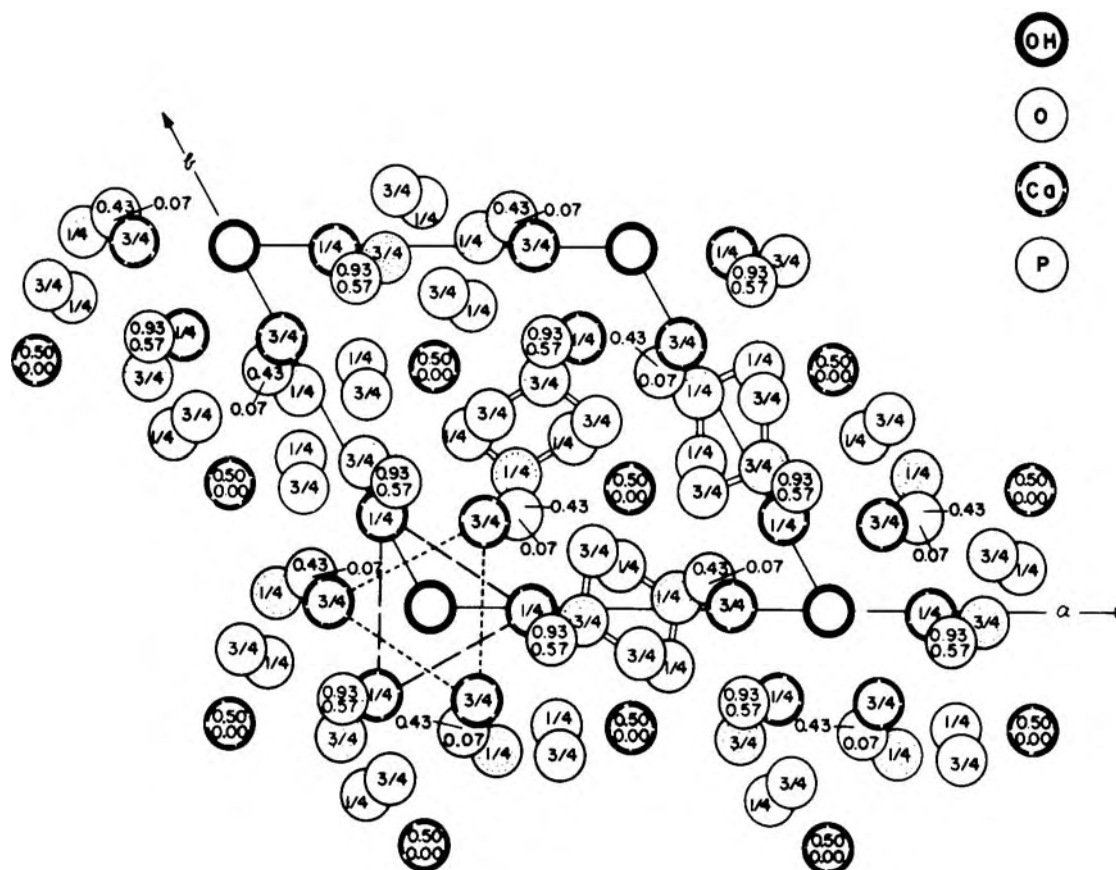


Fig. 1 The unit-cell (smallest building block) of hydroxyapatite, $\text{Ca}_{10}(\text{PO}_4)_6(\text{OH})_2$.

scaffolds for bone and teeth regeneration. Current commercial synthetic calcium phosphate biomaterials classified on the basis of composition include: hydroxyapatite (HA), $\text{Ca}_{10}(\text{PO}_4)_6(\text{OH})_2$; tricalcium phosphate (β -TCP), $\text{Ca}_3(\text{PO}_4)_2$; and biphasic calcium phosphate, BCP, an intimate mixture of HA and β -TCP with varying HA/ β -TCP ratios.^[15–19] Other commercial HA biomaterials are derived from biologic materials, (e.g., processed human bone, bovine bone-derived, hydrothermally converted coral, or derived from marine algae).^[20–22]

This paper is a brief review of biologic and synthetic apatites used as biomaterials, updating reviews published earlier by the authors.^[13,20,23,24]

BIOLOGIC APATITES

Normal and pathologic calcified tissues are composites of organic and inorganic phases. For bone, dentin, and cementum, the organic phase is principally collagen (about 25% by weight) with smaller amounts of non-collagenous proteins.^[25] On the other hand, the main organic phase in enamel is a noncollagenous protein (amelogenin) comprising about 1%, by weight, of the

enamel. In normal calcified tissues, such as in teeth and bones, in fish enameloids (teeth or calcified scales of some species), and in some species of shells, only carbonate- (carbonate hydroxyapatite, CHA) or carbonate- and fluoride-containing apatite (CFA) occurs as the principal inorganic phase.^[13,24,26,27] In pathologic calcifications (dental calculus, urinary stones, vascular calcification, and other soft-tissue calcifications), biologic apatite may occur as one of the mineral phases that include other calcium phosphates, e.g., amorphous calcium phosphate (ACP), Ca_xPO_4 ; dicalcium phosphate dehydrate (DCPD), $\text{CaHPO}_4 \cdot 2\text{H}_2\text{O}$; octacalcium phosphate (OCP), $\text{Ca}_8\text{H}_2(\text{PO}_4)_6 \cdot 5\text{H}_2\text{O}$; magnesium-substituted tricalcium phosphate (β -TCMP), $(\text{Ca},\text{Mg})_3(\text{PO}_4)_2$, calcium pyrophosphate dihydrate (CPPD), $\text{Ca}_2\text{P}_2\text{O}_7 \cdot 2\text{H}_2\text{O}$.^[24,28]

Enamel, Dentin, and Bone Apatite

Enamel, dentin, and bone apatite differ in crystallinity, reflecting crystal size (Fig. 2) and in concentrations of minor constituents, mainly, Mg and CO_3 (Fig. 3A) (Table 1).^[13,24,29] Enamel apatite contains the lowest concentrations of these ions and the highest crystallinity

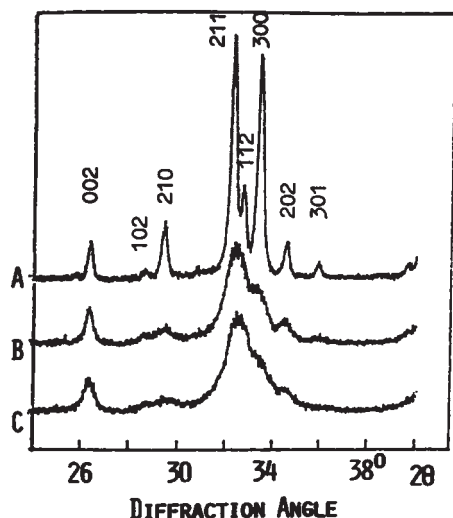


Fig. 2 X-ray diffraction profiles of biological apatites from adult human enamel (A), dentin (B), and bone (C).

(larger crystals), compared to either dentin or bone apatite, which shows much lower crystallinity (smaller crystals) and greater concentrations of Mg and CO_3 (Table 1). These apatites also differ in solubility, decreasing in the order: bone > dentin >> enamel. These differences in crystallinity (crystal size) and solubility may be attributed to the differences in the concentrations of the minor constituents (e.g., Mg, Na, CO_3 , and HPO_4). Studies on the effect of Mg and CO_3 ions on the properties of synthetic apatites demonstrated that incorporation of these ions independently and synergistically cause the growth of smaller and more soluble apatite crystals.^[13,24,29–32] The effects of CO_3 incorporation on apatite crystal size and morphology and on dissolution properties are much more pronounced than that of Mg^[13,24,32] Proteins^[25] and/or other ions (e.g., pyrophosphate, citrate)^[12,24] may also inhibit the crystal growth of biologic apatites.

Biologic apatites are usually calcium-deficient (i.e., with Ca/P molar ratio less than the stoichiometric value of 1.67 obtained for pure HA, $\text{Ca}_{10}(\text{PO}_4)_6(\text{OH})_2$). Calcining above 900°C of human enamel and dentin apatite results in the loss of the CO_3 constituent and formation of HA and Mg-substituted tricalcium phosphate, β -TCMP^[13,24] Calcining of bone (human or bovine) above 900°C results in the loss of CO_3 and formation of mostly HA with small amounts of calcium oxide, CaO.^[13,24]

Partial dissolution of biologic apatites and precipitation of other calcium phosphates (DCPD, OCP, β -TCMP) is believed to occur in human enamel and dentin caries, characterized by the dissolution of the tooth mineral (carbonate apatite) by acids produced by oral bacteria.^[33,34] The nonapatitic calcium phosphates (e.g., DCPD, OCP) may, in turn transform to apatites

by hydrolysis or by dissolution and reprecipitation processes.^[13,24,34]

Biologic apatites have been idealized as calcium hydroxyapatite (HA), $\text{Ca}_{10}(\text{PO}_4)_6(\text{OH})_2$.^[5–8] However, the difference in lattice parameters: a-axis for enamel, 0.9442 nm vs. 0.9422 nm for HA, Ca/P stoichiometry (e.g., about 1.63 for enamel or dentin apatite vs. 1.67 for HA), and the association of other ions, notably, magnesium (Mg) and carbonate (CO_3) with biologic apatites (Table 1) have caused many years of disagreements and research on the structure and composition of biologic apatites.^[8,9,13,35]

The nature of CO_3 incorporation in biologic apatites, especially in human enamel apatite had been a preoccupation of several researchers. The larger a-axis dimension of human enamel apatite, compared to pure or mineral HA, was first attributed to the CO_3 -for-OH substitution (Type A) in these apatites.^[9] Such type of substitution was observed in synthetic carbonate apatites prepared by diffusing CO_2 into HA at 1000°C under very dry conditions, resulting in an expanded a-axis and contracted c-axis dimensions compared to pure HA.^[9,10] However, studies on synthetic apatites prepared at much lower temperatures (60 – 95°C) by precipitation or hydrolysis methods showed a partial CO_3 -for- PO_4 substitution (Type B) coupled with a partial Na-for-Ca substitution, resulting in a contracted a-axis and expanded c-axis dimension compared to CO_3 -free apatites.^[11–13] The observed expanded a-axis of enamel apatite may be attributed to partial Cl-for-OH substitution^[36] and HPO_4 -for- PO_4 substitution,^[37] rather than the CO_3 -for-OH substitution in the apatite. A possible H_3O -for-OH was also offered as a possible cause for the expanded a-axis dimension.^[8] Additional evidence for the dominant partial CO_3 -for- PO_4 substitution in enamel, and all biologic apatites, is the similarity of the characteristic CO_3 absorption bands between the IR spectra of enamel apatite with those of synthetic apatites exhibiting CO_3 -for- PO_4 substitution as shown in Fig. 3B.^[3,24,38]

Studies in synthetic systems also showed that F incorporation (F-for-OH substitution) have the following effects: contraction in the a-axis dimension and no significant effect on the c-axis dimension compared to F-free apatites^[13,24] growth of larger and less soluble apatite crystals,^[13,24,39] and greater structural stability.^[3] Such studies elucidated the nature of fluoride (F) incorporation in some biologic apatites (from modern and fossil teeth and bones). For example, the smaller a-axis dimension of shark enameloid compared to HA or human enamel apatite is due to the high fluoride (F) concentration in shark enameloid (Table 1).^[13,27,40] The therapeutic use of fluoride in dentistry (sealants, topical gels, tablets) was based on the observation of low caries in areas of fluoridated water (1 ppm F).^[41] Fluoride treatment (topically or by the use of fluoridated dentifrices) of enamel and dentin leads to the formation

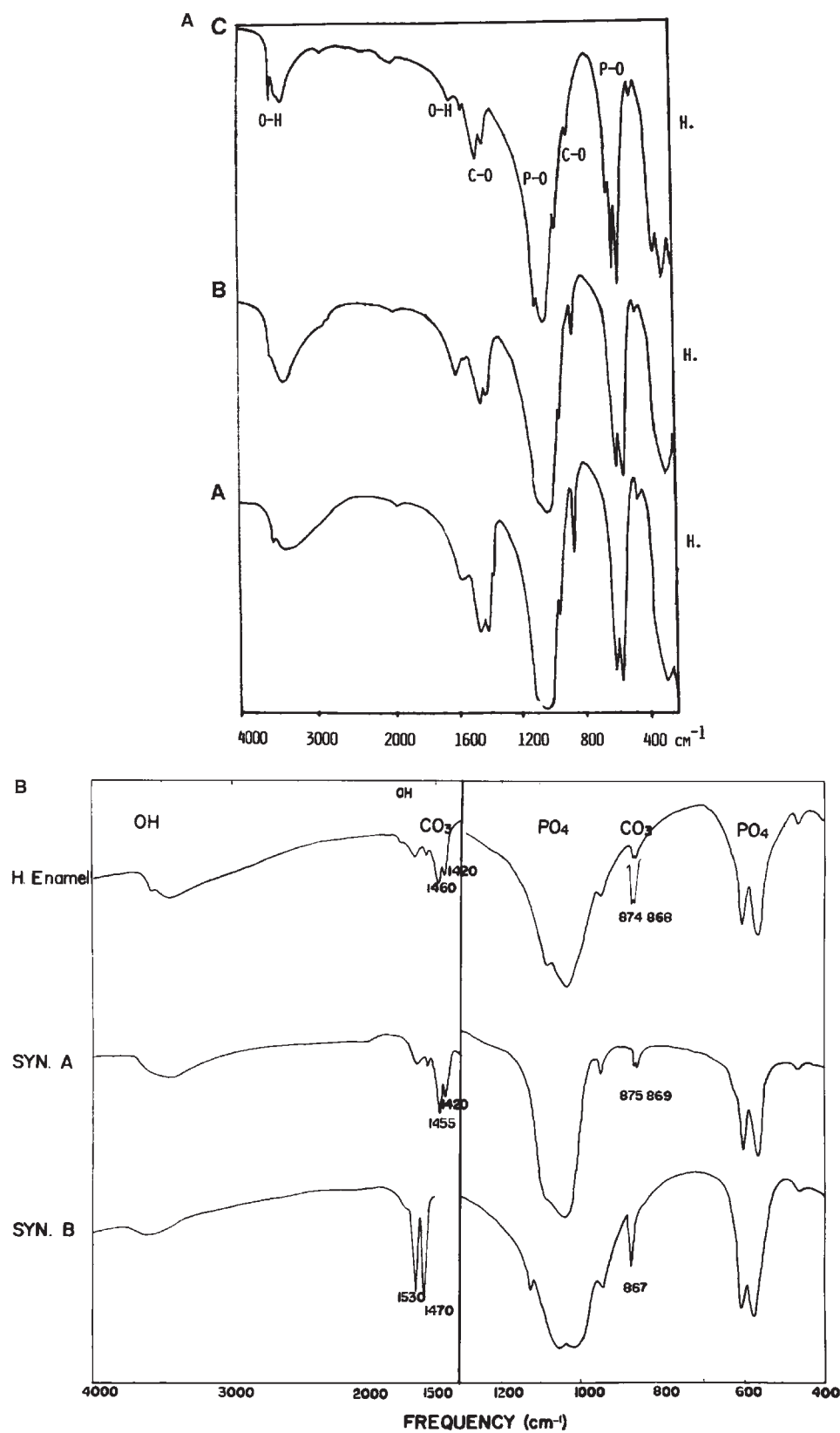


Fig. 3 (A): IR spectra of biological apatites from adult human enamel, dentin, and bone. Note the higher resolution of the P-O (for PO_4 groups) absorption spectra of enamel (C), compared to those of dentin (B) and bone (A) apatites. B: IR spectra comparing the characteristic C-O (for CO_3 groups) absorption bands in human enamel apatite compared to those in synthetic apatite with type A (CO_3 -for-OH) substitution (syn A) and in type B (CO_3 -for- PO_4) substitution (syn B).

of partially-substituted (F,OH) apatite that is more resistant to acid dissolution, (i.e. dental caries).^[29,33,34,39] Administration of F in the drinking water was shown to increase the crystal size and decrease the extent of

dissolution of rat bones^[42] and increase crystal thickness of enamel apatite.^[43] Fluoride therapy (as NaF) has been recommended for the management of osteoporosis to increase bone density.^[44,45]

Table 1 Composition and physical properties of apatites in adult human enamel, dentin, and bone

	Enamel	Dentine	Bone
<i>Composition</i>			
Calcium, Ca ²⁺	36.50	35.10	34.80
Phosphorus, P	17.70	16.20	15.60
Ca/P (molar)	1.63	1.63	1.71
Sodium, Na ⁺	0.50	0.60	0.90
Magnesium, Mg ²⁺	0.34	1.23	0.72
Potassium, K ⁺	0.06	0.05	0.03
Carbonate, CO ₃ ²⁻	3.50	5.60	7.40
Fluoride, F ⁻	0.01	0.06	0.03
Chloride, Cl	0.30	0.01	0.13
Pyrophosphate	0.02	0.10	0.07
Total inorganic (mineral) %	97.0	70.0	65.0
Total organic %	1.5	20.0	25.0
Absorbed H ₂ O %	1.5	10.0	10.0
Trace elements: Zn, Cu, Fe, Sr, etc.			
<i>Crystallographic properties</i>			
<i>Lattice parameters</i>			
a-axis (+0.0003 nm)	0.9441		
c-axis (+0.0003 nm)	0.6880		
Crystallite size (nm, ave)	33 × 3	2 × 0.4	2.5 × 0.3
Crystallinity index b	70–75	33–37	33–37
Ignition products (800C)	HA + β-TCMP	HA + β -TCMP	HA
HA, Ca ₁₀ (PO ₄) ₆ (OH) ₂			
a-axis = 0.9422 nm, c-axis = 0.6882 nm			
Crystallinity index = 100			
Composition: Ca, P, OH			

B

SYNTHETIC APATITES

Earlier extensive studies on synthetic apatites were made to gain a better understanding of biologic apatites and their properties. Because biologic apatite was idealized as hydroxyapatite HA, most of the studies centered on HA preparation and evaluation of HA properties. Studies on synthetic apatites in the last 30 years were motivated by development of calcium phosphate-based biomaterials (principally, HA) for bone repair, substitution, augmentation, and as scaffolds for tissue engineering in bone and teeth regeneration. The rationale for developing HA biomaterials was their similarity in composition to the bone mineral.

Synthetic HA can be made by solid-state reactions or by precipitation or hydrolysis methods and subsequent sintering at high temperatures, usually 1000°C and above. Synthetic apatites can also be prepared using hydrothermal,^[46,47] microwave,^[48,49] or sol-gel^[50,51] method. Apatite nanocrystals are obtained when

prepared by precipitation or hydrolysis at lower temperatures (25°C–60°C). Synthetic apatite crystals approximating the size of human enamel apatite may be obtained by precipitation or hydrolysis methods with reaction temperature, 80°C–95°C (Fig. 4). Apatites may also be prepared in gel systems,^[52] by electrodeposition,^[53,54] or biomimetic precipitation on metallic or polymeric^[55,56] substrates.

Apatites obtained by precipitation involve the reaction of calcium salts (e.g., CaNO₃, Ca(OH)₂, CaCl₂, Ca(Ac)₂) and phosphate salts (Na-, NH₄- or K- phosphates).^[24,57] Hydrolysis of nonapatitic calcium phosphates (e.g., ACP, DCPD, DCPA, OCP, β-TCP, α-TCP) or calcium compounds (e.g., CaCO₃, CaMg(CO₃)₂, CaF₂), in solutions containing OH, CO₃, or F, results in the formation of apatite, CO₃- or F-containing apatites.^[24,32,36,58] Apatites prepared by precipitation or hydrolysis methods when prepared at pH between 5 and 9 are calcium deficient apatites (CDA) and subsequent sintering results in the formation of biphasic calcium

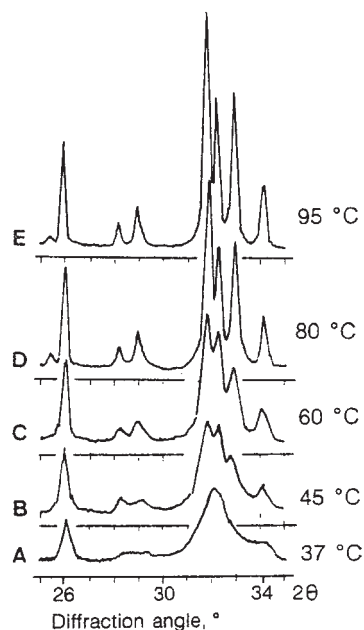


Fig. 4 X-ray diffraction profiles of precipitated apatites obtained at different reaction temperatures. The nanoapatite crystals similar to bone apatite are prepared at 37°C (A) or at room temperature. The narrowing of diffraction peaks reflects increased crystallinity; the broader the diffraction peak the smaller the crystals.

phosphate (BCP), an intimate mixture of HA and β -TCP.^[59] Sintering, or firing, of synthetic apatites results in an increase in crystal size and decrease in microporosity. Sintering of HA at temperatures above 1200°C results in decomposition of apatite-forming other calcium phosphates, such as β -TCP, α -TCP, and possibly even mixed with ACP. Combination of calcium phosphates (e.g., ACP, DCPA, DCPD, CDA, α -TCP, β -TCP) with other calcium compounds ($\text{Ca}(\text{OH})_2$, CaCO_3 , mixed with phosphate solutions or organic acids, results in the formation of apatitic calcium phosphate cements.^[60,61]

Studies on synthetic apatites showed that substitutions for Ca, PO_4 , or OH ions in the apatite structure results in changes in lattice parameters (Table 2) and crystallinity (reflecting crystal size and/or strain) and dissolution properties. For example, CO_3 -for- PO_4 coupled with Na-for-Ca substitution has the following effects on apatite properties: (a) smaller a- and larger c-axis dimensions compared to CO_3 -free apatites, (b) change from needle-like to rod-like to plate-like with increasing CO_3 incorporation, (c) lower resolution of the P-O (for PO_4) absorption bands in the IR spectra, (d) higher solubility, and (e) lower thermal stability.^[11–13,24,31,32,38] Such studies helped explain the contributions of some ions associated with biologic apatites, as well led to the development of some therapies and manufacture of calcium phosphate-based biomaterials.

Synthetic Apatites as Bone-Substitute Materials

The first successful use of a calcium phosphate reagent in bone repair was reported by Albee in 1920,^[62] followed more than 50 years later by the first clinical study by Nery et al.^[63] on periodontal bony defects using porous calcium phosphate identified by the authors as “tricalcium phosphate (TCP)”. X-ray diffraction analysis of Nery’s material years later revealed that the “tricalcium phosphate” consisted of a mixture of hydroxyapatite (HA) and beta-tricalcium phosphate (β -TCP) with a β -TCP/HA ratio of 20/80^[20] and was thus renamed, “biphasic calcium phosphate” or BCP.^[64] Similarity in composition of the synthetic apatite to biologic apatite was the rationale for the development of calcium-phosphate based biomaterials for bone repair, substitution, augmentation, and as scaffolds for bone and tooth regeneration. These calcium phosphate bioceramics include: hydroxyapatite (HA), tricalcium phosphate (β -TCP), biphasic calcium phosphate, BCP (consisting of an intimate mixture of HA and β -TCP of varying HA/ β -TCP ratios), bovine-bone derived apatites (unsintered and sintered), and coral-transformed apatite (Fig. 6). Commercialization of HA as bone graft materials was largely due to the independent efforts of Jarcho,^[15] deGroot,^[16] and Aoki.^[17] Basic studies on BCP led to its commercialization and popularity as bone graft materials and as scaffolds for tissue engineering.^[18,59,64–69]

Commercial HA biomaterials are usually prepared by precipitation at high pH and subsequent sintering at about 1000–1100°C.^[15–17] Coral-derived HA, or coralline HA, is prepared by the hydrothermal reaction of coral (CaCO_3 , in aragonite form) with ammonium phosphate (Fig. 5).^[70] Bovine-bone-derived HA is prepared by removing the organic phase (resulting in bone apatite) or removing the organic phase and sintering at high temperatures. These different preparations and origin (synthetic vs. biologic) are reflected in the difference in their initial crystallinity reflecting crystal size (Fig. 6) and their dissolution rates,^[71] increasing in the order:

HA \ll coralline HA
 \ll bovine bone apatite (sintered)
 $\ll\ll$ bovine bone apatite (unsintered).
 HA \ll BCP \ll β – TCP

Synthetic HA in Implant Surfaces and Coatings

HA and related calcium phosphates, in spite of their many desirable properties, are not strong enough to be used in load-bearing areas.^[15,16] The rationale for the

Table 2 Lattice parameters of mineral and synthetic apatites compared to biologic apatite

Apatite	Substituent	Lattice parameters (+0.0003 nm)	
		a-axis	c-axis
<i>Mineral</i>			
OH-Apatite (Holly Springs)	—	0.9422	0.6880
F-Apatite (Durango, Mex)	F-for-OH	0.9375	0.6880
Dahllite (Wyoming)	CO ₃ -for-PO ₄	0.9380	0.6885
Staffelite (Staffel, Germany)	CO ₃ -for-PO ₄ & F-for-OH	0.9345	0.6880
Marine phosphorite (U.S.A.)	CO ₃ -for-PO ₄ & F-for-OH	0.9322	0.6882
<i>Synthetic (non-aqueous)^a</i>			
OH-Apatite	—	0.9422	0.6882
F-Apatite	F-for-OH	0.9375	0.6880
Cl-Apatite	Cl-for-OH	0.9646	0.6771
CO ₃ -Apatite	CO ₃ -for-OH	0.9544	0.6859
<i>Synthetic (aqueous)^b</i>			
OH-Apatite	—	0.9438	0.6882
OH-Apatite	HPO ₄ -for-PO ₄	0.9462	0.6879
F-Apatite	F-for-OH	0.9382	0.6880
(Cl,OH)-Apatite	*Cl-for-OH	0.9515	0.6858
CO ₃ -OH-Apatite	*CO ₃ -for-PO ₄	0.9298	0.6924
CO ₃ -F-Apatite	*CO ₃ -for-PO ₄ & F-for-OH	0.9268	0.6924
Sr-Apatite	Sr-for-Ca	0.9739	0.6913
Pb-Apatite	Pb-for-Ca	0.9894	0.7422
Ba-Apatite	Ba-for-Ca	1.0161	0.7722
<i>Biologic apatite</i>			
Human enamel	(CO ₃ ,HPO ₄)-for-PO ₄ , (Na,Mg)-for-Ca, & Cl-for-OH	0.9441	0.6882
Shark enameloid	F-for-OH, Mg-for-Ca & (CO ₃ ,HPO ₄)-for-PO ₄	0.9382	0.6880

^aPrepared at a high temperature (1000°C) by solid state reaction or diffusion^[3,9] [Dykes and Elliott 1971].

^bPrepared at 95°C either by precipitation or by hydrolysis of CaHPO₄ in solutions containing the desired substituent.^[12,13,24,58]

development of H-coated orthopedic and dental implants is to combine the strength of the metal (usually titanium or titanium alloy) and the bioactive properties of HA and other calcium phosphates. Dense HA particles are used as the source material for depositing implant coating by the plasma-spray technique.^[15,16,72] The high temperatures (about 30,000°C) and other variable parameters involved in the plasma spray process (e.g., velocity of feeding the HA powder, distance of the gun from the metal substrate) result in the partial transformation of the original HA into ACP and minor amounts of α -TCP, β -TCP, tetracalcium phosphate (TTCP, Ca₄P₂O₉).^[73] Plasma-sprayed HA coatings have nonhomogenous composition (principally ACP/HA ratio), varying from the layer closest to the metal substrate to the outermost layer, and varying from one manufacturer to another.^[73] Alternatives to the plasma-spray method are: electrochemical

deposition,^[53,54] precipitation, or chemical deposition;^[55,56] the latter method being also applicable to nonmetallic substrates.^[56] These other methods provide homogenous implant coating of the desired composition, (e.g., HA, FA, CHA, OC)^[53–56,74] and allows coating deposition at much lower temperatures, thus permitting the incorporation of bioactive molecules and growth factors.

In recent years, HA or BCP has been used as an abrasive material for grit-blasting to roughen the surface and provide a more bioactive surface (compared to alumina or silica abrasive), thus enhancing osseointegration of the implant.^[75]

Synthetic HA in Composites

HA and related calcium phosphates are used as the inorganic component in composites with natural (e.g.,

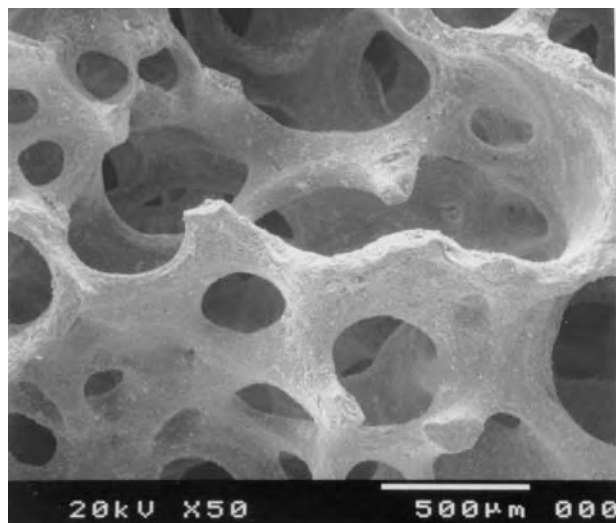


Fig. 5 Scanning electron micrograph of coralline apatite.

collagen, chitosan) or synthetic (polylactic acid or poly(lactide-co-glycolide), PLA or PLGA, high-molecular-weight polyethylene) polymers.^[76,77] The rationale for developing composite biomaterials is the fact that bone is a composite of a biological polymer (collagen) and inorganic phase (carbonate apatite).

Synthetic HA and BCP as Scaffolds for Tissue Engineering

Several investigators have reported that mesenchymal stem cells (MSC) from bone marrow can be cultured in

porous calcium phosphate biomaterials (ceramic HA, coralline HA, BCP ceramic) in vitro and implanted as a tissue engineered material for bone regeneration.^[78,79]

Critical Properties of Synthetic HA and Related Calcium Phosphates: Porosity, Biocompatibility, Bioactivity, Osteoconductivity, and Osteoinductivity

Porosity (interconnecting macro-porosity) is an important property of biomaterials that allows bony ingrowth and vascularization.^[80,81] Macro-porosity is introduced in HA and other calcium phosphates by the incorporation of porogens, such as naphthalene,^[82] H₂O₂, or sugar molecules. Biocompatibility of a material is determined in vitro from the cell response (proliferation, attachment, phenotypic expression) to the material. Material surface composition, surface roughness, or topography influences cell response.^[83] Bioactivity is defined as the property of the material to develop a direct adherent, and strong bonding and interface with the bone tissue.^[84,85] Bioactivity is demonstrated in vitro and in vivo by the ability of the material to form carbonate apatite on the surface from the simulated body fluid in vitro^[56] or biological fluid in vivo in osseous or nonosseous sites.^[65,66,86,87] Osteoconductivity is the property of the material that allows attachment, proliferation, migration, and phenotypic expression of bone cells, leading to formation of new bone in direct apposition to the biomaterial.^[84] Osteoinductivity is the property of the material that allows osteoprogenitor cell growth and development for bone formation to

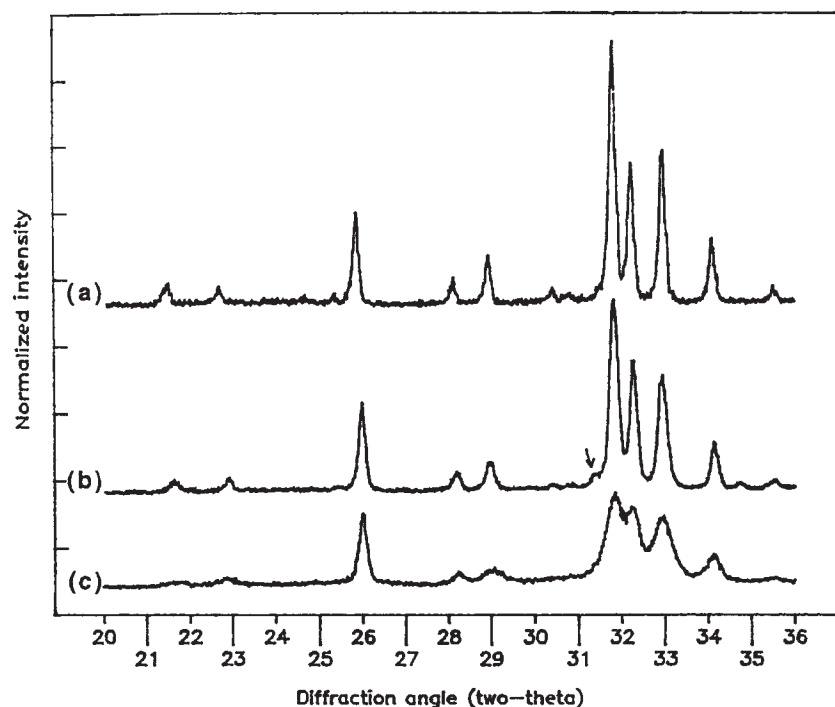


Fig. 6 X-ray diffraction profiles of apatite biomaterials (A,B,C). (A) Ceramic hydroxyapatite (Calciite™), (B) Coralline Hydroxyapatite (Interpore™), (C) Unsintered apatite, calcium-deficient (Osteogen™). The arrow on (b) indicates presence of Mg-substituted β -TCP.

occur^[88] and is usually determined by the formation of bone in non-bone-forming sites (e.g., under the skin or in muscle). HA and related calcium phosphates are generally considered to have all of the above properties, except osteoinductivity. However, it has been reported that with the appropriate composition, geometry and architecture, osteoinductive properties can be promoted.^[89,90]

CONCLUSION

The total analysis of the results obtained over four decades of research in this fascinating field of science suggests that, with newer scientific tools and with further refinement, we will succeed to address problems relating to the structure, morphology, and analysis of the biologic materials and produce synthetic apatites or composites that will emulate the structure and characteristics of natural soft and hard tissues.

ARTICLES OF FURTHER INTEREST

Bioactive Glass; Bioceramics: Chemical, Physical, and Biological Properties; Biologic and Synthetic Apatites; Biphasic Calcium Phosphate (BCP)

REFERENCES

- Naray-Szabo, S. The structure of apatite $(\text{CaF})\text{Ca}_4(\text{PO}_4)_3$. *Z. Kristallogr.* **1930**, *75*, 387–398.
- Beevers, C.A.; McIntyre, D.B. The atomic structure of fluorapatite and its relation to that of tooth and bone mineral. *Miner. Mater.* **1956**, *27*, 254–259.
- Young, R.A.; Elliott, J.C. Atomic scale bases for several properties of apatites. *Arch. Oral Biol.* **1966**, *11*, 699–707.
- Deer, W.A.; Howie, R.A.; Zussman, J. *An Introduction to the Rock Forming Mineral*; Longman: Hong Kong, 1985; 504–509.
- DeJong, W.F. La substance minérale dans les os. *Rec. Trav. Chim.* **1926**, *45*, 415–448.
- Kay, M.; Young, R.A.; Posner, A.S. The crystal structure of hydroxyapatite. *Nature* **1964**, *204*, 1050–1052.
- Carlstroem, D. X-ray crystallographic studies on apatite and calcified tissues. *Acta Radiol. Suppl.* **1955**, 121.
- McConnell, D. The crystal chemistry of carbonate apatites and their relationship to the composition of calcified tissue. *J. Dent. Res.* **1952**, *31*, 53–63.
- Elliott, J.C. The Crystallographic Structure of Dental Enamel and Related Apatites; PhD Thesis, University of London, 1964.
- Bonel, G.; Montel, G. Sur une nouvelle apatite carbonatée synthétique. *CR Séance Acad. Sci. Paris* **1964**, *258*, 923–926.
- Zapanta-LeGeros, R. Effect of carbonate on the lattice parameters of apatite. *Nature* **1965**, *206*, 403.
- LeGeros, R.Z.; Trautz, O.R.; LeGeros, J.P. Apatite crystallites: effect of carbonate on morphology. *Science* **1967**, *155*, 1409–1411.
- LeGeros, R.Z. Apatites in biological systems. *Prog. Crystal Growth Character* **1981**, *4*, 1–45.
- Rey, C.; Renugopalakrishnan, V.; Collins, B. Fourier transform infrared spectroscopic study of the carbonate ions in bone mineral during aging. *Calcif. Tissue Int.* **1991**, *49*, 251–258.
- Jarcho, M. Calcium phosphate ceramics as hard tissue prosthetics. *Clin. Orthop. Relat. Res.* **1981**, *157*, 259–278.
- deGroot, K. *Bioceramics of Calcium Phosphates*; CRC Press: Boca Raton, FL, 1983.
- Aoki, H.; Kato, K.; Ogiso, M. Studies on the application of apatite to dental materials. *J. Dent. Eng.* **1977**, *18*, 86–89.
- Daculsi, G.; Passuti, N.; Martin, S.; LeGeros, R.Z. Macroporous biphasic calcium phosphate ceramic for long bone surgery in human and dogs.: clinical and histological study. *J. Biomed. Mater. Res.* **1990**, *24*, 379–396.
- Metzger, D.S.; Driskell, T.D.; Paulsrud, J.R. Tricalcium phosphate ceramic- a resorbable bone implant: review and current status. *J. Am. Dent. Assoc.* **1982**, *105*, 1035–1040.
- LeGeros, R.Z. Calcium phosphate materials in restorative dentistry: a review. *Adv. Dent. Res.* **1988**, *2*, 164–180.
- Schopper, C.; Moser, D.; Sabbas, A.; Lagogiannis, G.; Spassova, E.; Konig, F.; Donath, K.; Ewers, R. The fluorohydroxyapatite (FHA) FRIOS^R Algipore^R is a suitable biomaterial for the reconstruction of severely atrophic human maxillae. *Clin. Oral Implants Res.* **2003**, *14*, 741–749.
- Holmes, R.E. Bone regeneration within a coralline hydroxyapatite implant. *Plast. Reconstr. Surg.* **1979**, *63*, 626–633.
- Ben-Nissan, Besim.; Chai, Cameron; Evans, Louise. Crystallographic and spectroscopic characterization and morphology of biogenic and synthetic apatites. In *Encyclopedic Handbook of Biomaterials & Bioengineering. Part B: Applications*; Wise, D.L., Trantolo, D.J., Alobelli, D.E., Yanzewski, M.J., Gresser, J.D., Schwartz, E.R., Eds.; Marcel Dekker, Inc.: New York, 1995; Vol. 1, 191–221.
- LeGeros, R.Z. In *Calcium Phosphates in Oral Biology and Medicine*; Myers, H., Karger, S., Eds.; Monographs in Oral Sciences; Basel: Switzerland, 1991; Vol. 15.
- Boskey, A.L. Overview of cellular elements and macromolecules implicated in the initiation of mineralization. In *The Chemistry and Biology of Mineralized Tissues*; Butler, W.T., Ed.; UAB: Birmingham, Alabama, 1984; 334–343.
- LeGeros, R.Z.; Pan, C.M.; Suga, S.; Watabe, N. Crystallo-chemical properties of apatites in atremate brachiopod shells. *Calcif. Tissue Int.* **1985**, *37*, 651–658.
- LeGeros, R.Z.; Bautista, C.; Wong, J.L.; LeGeros, A.; LeGeros, J.P. Biological apatites in modern and fossil shark teeth and in calcified fish scales. In *Fundamentals of Biomineralization*. Allemand, D., Cuif, J.P., Eds.; Bull.; Inst. Oceanographie: Monaco, 1994; Vol. 14, 229–235.

28. LeGeros, R.Z. Formation and transformation of calcium phosphates in biological and synthetic systems—relevance to vascular calcification. *Kardiologie* 90, Suppl 3: Vascular Calcifications: Biophysics and Biology **2001**, 203–209.
29. LeGeros, R.Z.; Sakae, T.; Bautista, C. Magnesium and carbonate in enamel and synthetic apatites. *Adv. Dent. Res.* **1996**, *10*, 225–231.
30. Okazaki, M.; LeGeros, R.Z. Crystallographic and chemical properties of Mg-containing apatites before and after suspension in solutions. *Magnes. Res.* **1992**, *5*, 103–108.
31. LeGeros, R.Z.; Tung, M.S. Chemical stability of carbonate and fluoride containing apatites. *Caries Res.* **1983**, *17*, 419–429.
32. LeGeros, R.Z.; LeGeros, J.P.; Trautz, O.R.; Klein, E. Conversion of monetite, CaHPO₄, to carbonate apatite: effect on crystallinity. *Adv. X-ray Appl.* **1971**, *14*, 57–66.
33. Margolis, H.C.; Moreno, E.C. Kinetic and thermodynamic aspects of enamel demineralization. *Caries Res.* **1985**, *19*, 22–35.
34. LeGeros, R.Z. Chemical and crystallographic events in caries. *J. Dent. Res.* **1990**, *69*, 567–574. Special Issue.
35. Driessens, F.C.M. In *Mineral Aspects of Dentistry*; Myers, H.M., Ed.; Monographs in Oral Sciences; Basel: Karger, 1982; Vol. 10.
36. LeGeros, R.Z. The unit-cell dimensions of human enamel apatite: effect of chloride incorporation. *Arch. Oral Biol.* **1974**, *20*, 63–68.
37. LeGeros, R.Z.; LeGeros, J.P.; Bonel, G. Types of 'H₂O' in human enamel and in precipitated apatites. *Calcif. Tissue Res.* **1979**, *26*, 111–116.
38. LeGeros, R.Z.; LeGeros, J.P.; Trautz, O.R.; Klein, E. Spectral properties of carbonate in carbonate-containing apatites. *Dev. Appl. Spectrosc.* **1970**, *7*, 3–12.
39. Moreno, E.G.; Kresak, M.; Zahradnik, R.T. Physico-chemical aspects of fluoride-apatite systems relevant to the study of dental caries. *Caries Res.* **1977**, *11* (Suppl. 1), 142–171.
40. LeGeros, R.Z.; Silverstone, L.M.; Daculsi, G.; Kerebel, L.M. In vitro caries-like lesion formation in F-containing tooth enamel. *J. Dent. Res.* **1983**, *62*, 138–144.
41. Dean, H.T.; Arnold, F.A., Jr; Elvove, E. Domestic water and dental caries. V. Additional studies of the relations of fluoride domestic waters to dental caries experience in 4,425 white children aged 12 to 14 years, of 13 cities in 4 states. *Public Health Rep.* **1942**, *57*, 115–125.
42. LeGeros, R.Z.; Singer, L.; Ophaug, R.; Quirolgico, G.; LeGeros, J.P. The effect of fluoride on the stability of synthetic and biological (bone mineral) apatites. In *Osteoporosis*; Menczel, J., Robin, G.C., Makin, M., Eds.; Wiley: New York, 1982; 327–341.
43. Robinson, C.; Kirkham, J. The effect of fluoride on the developing mineralized tissues. *J. Dent. Res.* **1990**, *69*, 184–185.
44. Pak, C.Y.C.; Sakhaee, K.; Bell, N.K.; Licata, A.; Johnson, C.; Rubine, B. Comparison of non-randomized trials with slow-release sodium fluoride with a randomized placebo-controlled trial in postmenopausal osteoporosis. *J. Bone Miner. Res.* **1996**, *11*, 160–168.
45. Kleerkoper, M. Fluoride and the skeleton. In *Principles of Bone Biology*; Bilezikian, J.P., Raisz, L.G., Rodan, F.A., Eds.; Academic Press: San Diego, 1996; Chapter 75, 1053–1062.
46. Ito, A.; Nakamura, S.; Aoki, H.; Akao, M.; Traoka, K.; Tsutsumi, S.; Onuma, K.; Tateishi, T. Hydrothermal growth of carbonate-containing hydroxyapatite single crystals. *J. Crystal Growth* **1996**, *163*, 311–317.
47. Hattori, T.; Iwadate, Y. Hydrothermal preparation of calcium hydroxyapatite powders. *J. Am. Ceram. Soc.* **1990**, *73*, 1803–1805.
48. Pena, J.; LeGeros, R.Z.; Rohanizadeh, R.; LeGeros, J.P. CaCO₃–CaP biphasic materials prepared by microwave processing of natural aragonite and calcite. *Key Eng. Mater.* **2001**, *192–195*, 267–270.
49. Lerner, E.; Sarig, S.; Azoury, R. Enhance maturation of hydroxyapatite from aqueous solutions using microwave irradiation. *J. Mater. Sci.: Mater. Med.* **1991**, *2*, 138–141.
50. Milev, A.S. Chemistry, Synthesis and Morphological Stability of sol–gel derived Carbonate Substituted Plate-Like Hydroxyapatite. PhD Thesis, University of Technology: Sydney, 2002.
51. Deptula, A.; Lada, W.; Olczak, T.; Sarowska, S.; LeGeros, R.Z.; LeGeros, J.P. Complex Sol–Gel Process (CSGP) Preparation of Calcium Phosphate Biomaterials (Powders, Monoliths, Fibers) In *Bioceramics 11*, World Scientific Publishing: Singapore, 1998; 743–746.
52. LeGeros, R.Z.; Morales, P. Renal stone crystals grown in gel systems. *Invest. Urol.* **1973**, *11*, 12–20.
53. Vijayraghavan, T.V.; Bensalem, A. Electrodeposition of apatite coating on pure titanium and titanium alloys. *J. Mater. Sci. Lett.* **1994**, *13* (24), 1782–1785.
54. LeGeros, J.P.; Lin, S.; LeGeros, R.Z. Electrochemically deposited CaP coating on titanium alloy. *J. Dent. Res.* **2000**, *79*, 560. 3332 Sp.Iss,SI.
55. Rohanizadeh, R.; LeGeros, R.Z.; Harsono, M.; Ben-David, A. Adherent apatite coating on titanium substrate using chemical deposition. *J. Biomed. Mater. Res.* **2004**, *72A*, 428–438.
56. Kokubo, T. Formation of biologically active bone-like apatite on metals and polymers by a bioamimetic process. *Thermochim. Acta* **1996**, *280*, 479–490.
57. Hayek, E.; Newesely, H. Pentacalcium monohydroxy-orthophosphate. *Inorg. Synth.* **1963**, *7*, 63–69.
58. LeGeros, R.Z.; Go, P.; Vandemaele, K.H.; LeGeros, D.J. Transformation of calcium carbonates and calcium phosphate to carbonate apatites: possible mechanism for phosphorite formation. Proceedings of 2nd International Congress on Phosphate Compounds, Boston, 1980, 41–53.
59. LeGeros, R.Z.; Lin, S.; Rohanizadeh, R.; Mijares, D.; LeGeros, J.P. Biphasic calcium phosphate bioceramics: preparation, properties and applications. *J. Mater. Sci.: Mater. Med.* **2003**, *14*, 201–210.
60. Brown, W.E.; Chow, L.C. A new calcium phosphate water setting cement. In *Cements Research Progress*; Brown, P.W., Ed.; American Ceramic Society: Westerville, OH, 1987.
61. Niwa, S.; LeGeros, R.Z. Injectable calcium phosphate cements for repair of bone defects. In *Tissue Engineering and Biodegradable Equivalents: Scientific and Clinical*

- Applications*; Lewandrowski, K.-U., Wise, D.L., Trantolo, D.J., Gresser, J.D., Eds.; Marcel Dekker Inc.: New York, 2002; 385–500.
62. Albee, F.H. Studies in bone growth: triple calcium phosphate as a stimulus to osteogenesis. *Ann. Surg.* **1920**, *71*, 32–36.
 63. Nery, E.B.; Lynch, K.L.; Hirthe, W.M.; Mueller, K.H. Preliminary clinical studies of bioceramics in periodontal osseous defects. *J. Periodontol.* **1978**, *49*, 523–527.
 64. Ellinger, R.F.; Mery, E.G.; Lynch, K.L. Histological assessment of periodontal osseous defects following implantation of hydroxyapatite and biphasic calcium phosphate ceramics: a case report. *Int. J. Periodontics Restorative Dent.* **1986**, *3*, 223–233.
 65. LeGeros, R.Z.; Daculsi, G.; Nery, E.; Lynch, K.; Kerebel, B. In vivo transformation of biphasic calcium phosphates of varying β -TCP/HA ratios: ultrastructural characterization. Third World Biomaterials Congress, Kyoto, Japan, 1988.
 66. LeGeros, R.Z.; Daculsi, G. In vivo transformation of biphasic calcium phosphate ceramics: histological, ultrastructural and physico-chemical characterization. In Yamamuro, T., Hench, L., Wilson-Hench, J., Eds.; Handbook in Bioactive Ceramics; CRC Press: Boca Raton, 1990; Vol. 2, 17–28.
 67. Daculsi, G.; Passuti, N. Bioactive ceramics fundamental properties and clinical applications: the osseo-coalescence process. In *Bioceramics 2*; Heimcke, G., Oonishi, H., Eds.; Cologne: Butterworth—Heinemann, Germany, 1990; 3–10.
 68. Wykrota, L.L.; Garrido, C.A.; Wykrota, F.H.I. Clinical evaluation of biphasic calcium phosphate ceramic used in orthopaedic lesions. In *Bioceramics 11*; LeGeros, R.Z., LeGeros, J.P., Eds.; World Scientific Publishing: Singapore, 1998; 641–644.
 69. Nery, E.; LeGeros, R.Z.; Lynch, K.L.; Lee, K. Tissue response to biphasic calcium phosphate ceramic with different ratios of HA/ β -TCP in periodontal osseous defects. *J. Periodontol.* **1992**, *63*, 729–735.
 70. Roy, D.L.; Linnehan, S.K. Hydroxylapatite formed from coral skeletal carbonate by hydrothermal exchange. *Nature* **1974**, *247*, 220–222.
 71. LeGeros, R.Z.; Bautista, C.; LeGeros, J.P.; Vijaraghavan, T.V.; Retino, M. Comparative properties of bioactive bone graft materials. In *Bioceramics 8*; Sedel, L., Ed.; Pergamon Press: London, 1995; 1–7.
 72. Berndt, C.C.; Haddad, G.N.; Farmer, A.J.D.; Gross, K.A. Review: thermal spraying for bioceramic applications. *Mater. Forum* **1990**, *14*, 161–173.
 73. LeGeros, R.Z.; LeGeros, J.P.; Kim, Y.; Kijkowska, R.; Zheng, R.; Bautista, C.; Wong, J.L. Calcium phosphates in plasma-sprayed HA coatings. *Ceram. Trans.* **1995**, *48*, 173–189.
 74. Lin, S.; LeGeros, R.Z.; LeGeros, J.P. Adherent octacalcium phosphate coating on titanium alloy using a modulated electrochemical deposition method. *J. Biomed. Mater. Res.* **2003**, *66A*, 820–828.
 75. Salgado, T.; LeGeros, J.P.; Wang, J. Effect of alumina and apatitic abrasives on Ti alloy substrates. In *Bioceramics 11*; Ohgushi, H., Ed.; World Scientific Publishing: Singapore, 1998; 683–686.
 76. Doyle, C. Bioactive composites in orthopedics. In Yamamuro, T., Hench, L., Wilson-Hench, J., Eds.; Handbook of Bioactive Ceramics; CRC Press: Boca Raton 1990; Vol. 2, 195–207.
 77. Okazaki, M.; Ohmae, H.; Hino, T. Insolubilization of apatite-collagen composites by UV irradiation. *Biomaterials* **1989**, *10*, 564–568.
 78. Ohgushi, H.; Caplan, A.I. Stem cell technology and bioceramics: from cell to gene engineering. *J. Biomed. Mater. Res. Appl. Biomater.* **1999**, *48*, 913–927.
 79. Toquet, J.; Rohanizadeh, R.; Guicheux, J.; Daculsi, G. Osteogenic potential in vitro of human bone marrow cells cultured on macroporous biphasic calcium phosphate ceramic. *J. Biomed. Mater. Res.* **1999**, *44*, 98–109.
 80. LeGeros, R.Z. Properties of osteoconductive biomaterials: calcium phosphates. *Clin. Orthopaed Related Res.* **2002**, *395*, 81–98.
 81. Klawitter, J.J. A Basic Investigation of Bone Growth in Porous Materials. PhD Thesis, Clemson University: Clemson, 1979.
 82. Hubbard, W. Physiological calcium phosphate as orthopedic implant material. PhD Thesis. Marquette University: Milwaukee, 1974.
 83. Harada, Y.; Want, J.-T.; Doppalppudi, V.A.; Willis, A.A.; Goldring, S.R. Differential effects of different forms of hydroxyapatite and hydroxyapatite/tricalcium phosphate particles on human monocyte/macrophages in vitro. *J. Biomed. Mater. Res.* **1996**, *31*, 19–26.
 84. Hench, L.L. Bioceramics: from concept to clinic. *J. Am. Ceram. Soc.* **1994**, *74*, 1487–1510.
 85. Osborn, J.F.; Newesely, H. The material science of calcium phosphate ceramic. *Biomaterials* **1980**, *1*, 108–111.
 86. Heughebaert, M.; LeGeros, R.Z.; Gineste, M.; Guilhem, M.; Bonel, G. Physico-chemical characterization of deposits associated with HA ceramics implanted in non-osseous sites. *J. Biomed. Mater. Res.* **1988**, *22*, 257–268.
 87. LeGeros, R.Z.; Daculsi, G.; Orly, I. Substrate surface dissolution and interfacial biological mineralization. In *The Bone Biomaterials Interface*; Davies, J.E., Ed.; University of Toronto Press: Toronto, 1991; 76–88.
 88. Urist, M.R. Bone formation by autoinduction. *Science* **1965**, *150*, 893–898.
 89. Ripamonti, U.; Ma, S.; Reddi, A.H. The critical role of geometry of porous hydroxyapatite delivery system induction of bone by osteogenin, a bone morphogenetic protein. *Matrix* **1992**, *12* (3), 202–212.
 90. Le Nihouannen, D.; Daculsi, G.; Saffarzadeh, A.; Gauthier, O.; Delplace, S.; Pilet, P.; Layrolle, P. Ectopic bone formation by microporous calcium phosphate ceramic particles in sheep muscles. *Bone* **2005**, *36* (6), 1086–1093.

Biological Adhesives from Nature

Lloyd D. Graham

CSIRO Molecular and Health Technologies, Sydney Laboratory, Riverside Life Sciences Centre, New South Wales, Australia

INTRODUCTION

This entry reviews some of the adhesive secretions used by organisms for attachment, construction, obstruction, defense, and predation. These glues often have to work on surfaces that are wet, dirty, or otherwise refractory—conditions that confound even the best synthetic adhesives. The entry reviews what is known of the natural glues produced in marine and other wet environments (by fish, holothurians, molluscs, arthropods, worms, bacteria, algae, and fungi) and those produced on land (by amphibians, spiders, insects, molluscs, etc.). The entry also considers whether some of these glues may prove useful in medical contexts, such as wound healing, surgical repair, or dentistry. There is a significant unmet clinical need for strong, elastic, and biocompatible adhesives that work in moist environments, and this entry reviews the steps that have been taken toward meeting this need with materials based on the design of high-performance glues found in nature.

An adhesive must engage in bonding interactions both with the adherends (adhesion) and within the adhesive layer (cohesion). The set glue should have a high elastic modulus (stiffness) and/or a high loss modulus (viscosity), as it must resist flow and be able to dissipate energy by deformation.^[1] Most biological adhesives rely on polymers—typically, proteins or polysaccharides—to form adhesive interactions with the substrata and cohesive interactions with each other. The concentration of the glue components is an important consideration.^[2] The “interpenetrating network” required for cohesion may be achieved either by providing preformed chains of high molecular mass or by linking smaller units in situ, either noncovalently or covalently. Many natural glues are of the second type; they probably owe their cohesive strength to a mixture of intermolecular forces acting in concert, where the possibilities for interaction include polymer chain entanglement, noncovalent association, joint coordination of metal ions, and covalent bonding.^[2] Glues of the second type also have the advantage of relatively low initial viscosities, which facilitates their mechanical engagement with the surface irregularities of the adherends and enhances their adhesive strength when set. Of the types of biological glues that have

been examined biochemically, 18 are thought to be predominantly protein, 7 involve proteins and carbohydrates in comparable amounts, and 3 consist mainly of carbohydrate (Table 1). Of course, the apparent bias toward protein-based adhesives may simply reflect the fact that it is generally easier to characterize and manipulate proteins than polysaccharides.

Sequences are known for some 37 protein families found in biological glues (Table 1). These often have similarities with the sequences of structural and elastomeric proteins: many are Gly-rich and contain repeated motifs and posttranslational modifications (PTMs). Common PTMs include glycosylations, which give rise to N- or O-linked glycoproteins, and hydroxylations, which give rise to 3,4-dihydroxyphenyl-L-alanine (Dopa), 4-hydroxyproline (Hyp), 5-hydroxylysine (Hyl), or 4-hydroxyarginine. Both types of PTM increase the number of hydroxyl groups present, and therefore enhance the potential for hydrogen bonding, metal coordination, and so on. Dopa has very specific properties in this regard, and its contribution to bioadhesion has been explored in detail. Despite their prominence in the literature, adhesive mechanisms that depend on Dopa are actually likely to be quite rare.^[3] Dopa-based reactions (e.g., quinone tanning) do, however, play important roles in other biological processes, such as the sclerotization of insect cuticles. Intermolecular cross-linking via covalent bonds or high-affinity metal coordination might be expected to contribute to the high cohesive strength of many irreversible biological adhesives (e.g., those intended for permanent attachment), and such bonds are known to occur *inter alia* in mantid ootheca and in the adhesives used to anchor marine algae and mussels. Glues designed to produce temporary adhesion (e.g., for short-term anchorage, prey capture, or predator evasion) are more likely to rely on noncovalent interactions and chain entanglement. Table 2 indicates the range of bond strengths associated with natural adhesives.

Many adhesive proteins are rich in Gly, Pro, or Ser, residues that have high β -turn potential.^[4] Indeed, tandem repeats of Pro- and Gly-rich motifs can form a range of specialized poly-Pro/ β -turn conformations called pro- β helices.^[5] The dipeptide Pro-Gly is particularly likely to constitute the center of a type II β -turn, and tandem turns of this type constitute a

Table 1 Key features of natural adhesives mentioned in the text

Organisms/material	Type of adhesion	Main component	Protein sequences	Covalent-like cross-links
<i>Amphibia</i>				
Salamander exudate	R	P		No
Frog exudate	R	P		No (Disulfides)
<i>Fish</i>				
Stickleback nest	I	P	1	?Disulfides
<i>Molluscs</i>				
Mussel plaque	I	P	5	Fe:Dopa ₃ , quinones
Limpet mucus	R	P		(Disulfides)
Periwinkle mucus	R	PC		(Disulfides)
Land snail mucus	R	PC		(Disulfides)
Slug mucus	R	PC		(Disulfides)
Abalone shell	I	P	1	
<i>Crustaceans</i>				
Barnacle cement	I	P	8	(Disulfides)
<i>Echinoderms</i>				
Sea star footprint	R	P		
Sea cucumber tubules	R	PC		(Disulfides)
<i>Marine worms</i>				
Monogenean adhesive	R	P		
Polychaete cement	I	P	2	Quinones
<i>Insects and spiders</i>				
Tick cement	I	P	9	
Blackfly larval glue	I	P		No
Fruit fly larval glue	I	PC	6	No
Fruit fly mating plug	R	P	1	No
Bumblebee mating plug	R	F		No
Silkworm sericin	I	P	2	No
Mantid ootheca	I	P		NADA-quinones
Spumaline/egg glues	I	C		?No
Rove beetle secretion	R	PF		?No
Spiderweb glue	R	P		?No
<i>Other animals</i>				
Velvetworm ejectate	R	P		?No
<i>Fungi, algae, and bacteria</i>				
Fungal hydrophobins	I	P	2	No
Fungal mannoproteins	I	PC		?No
Algal attachment	I	PC, C		Polyphenolics
Microbial EPS	I	C		No

For brevity, specialized cell-surface interactions between microbial pathogens and their host tissues are not included. The type of adhesion is presumed from the purpose of the attachment to be I, irreversible or long-term, or R, reversible or short-term. The main type of organic component in the adhesive is indicated by P, protein; C, carbohydrate; F, lipid or fatty acid; two letters denote two components in comparable amounts. Items in parentheses may be of secondary importance or only present in some circumstances or species; items preceded by “?” are uncertain or speculative. Protein sequences are usually deduced from cloned genes, and homologs are generally counted as a single sequence. The heading “Covalent-like cross-links” is mainly directed toward intermolecular tethering via covalent bonds, high-affinity metal coordination, etc., but the entry “(Disulfides)” may refer to intramolecular bridges that are important to adhesive function.

Table 2 Bond strengths of natural and reference adhesives

Material	Environment	Bond time	Surfaces	Bond strength	Reference ^a
				(kPa)	
<i>Tensile strength^b</i>					
Amphibian					
<i>Notaden</i> exudate	WD	1–25 hr	Plastic	57–78	[63]
Tomato frog glues	WD	3 min	Plastic	92–104	[60]
Other frog/salamander	WD	3 min	Plastic	20–63	[60]
Mussel					
MAP (extract)	[D]	24 hr	Skin	330	[97]
MAP + fillers, XL	D	~1 hr	Meniscus, bone	[<8]	[96]
Mussel plaque	[W]		Hydrophobic	15	[117]
Mussel plaque	[W]		Polar	320–850	[117,118]
MAP (Cell-Tak TM)	W	22 hr	Skin	0.3	[95]
MAP (extract)	W	24 hr	Skin	930	[97]
Barnacle					
Cyprid (larva), mobile	W		Diverse rigid	60–227	[117,119]
Cyprid (larva), fixed	W		Diverse rigid	160–2,140	[117,119]
Adult barnacle	W	28 weeks	Diverse rigid	80–2,080	[117,119]
Other marine organisms					
Sea cucumber tubules	W	[<1 hr]	Glass	30–135	[120]
Sea star—live podia	W	>10 sec	Plastic, glass	180–198	[31,121]
Sea anemone—live			Plastic	460	[120]
Sea urchin—live	W	>10 sec	Plastic	140–340	[31]
Limpet—live	WD		Diverse rigid	77–227	[122]
Limpet—live ^c				230–518	[33]
Tubeworm—live	W		Slate	2,190–2,475	[123]
Bacterial EPS					
Montana C-902	[D]		[Aluminum]	3,448–10,342	[38]
Analogues and reference materials					
Fibrin	W	1 hr	Cartilage	5–19	[95]
Fibrin (Tisseel [®])	W	48 hr	Skin	1,040	[97]
Fibrin (Tisseel [®])	D	48 hr	Skin	430	[97]
GRF	D		Aorta	24–48	[95]
GRF	W	1 hr	Skin/cart/bone	70–200	[95]
GRG	W	1 hr	Cartilage	21	[95]
(AKPSYPPPT _{8–12} YK) _{8–12}	M	3 days	Iron	2,740	[118]
(Lys) _{460–3260}	M	3 days	Iron	4,510–12,050	[118]
Poly(DL-Ala), poly(YK)	M	3 days	Iron	3,630–3,720	[118]
Gelatin (20%)	M	3 days	Iron	7,740	[118]
Gelatin (5%)	D	2–3 hr	Plexiglass	18	[2]
Pectin (15%)	[D]			34	[2]
Natural gums (65–75%)	[D]			45–88	[2]
Rubber cement	D	3 min	Plastic	35	[60]
<i>n</i> -Bu-2-cyanoacrylate ^c			ST, cart, tendon	100–1,000	[95]
<i>n</i> -Bu-2-cyanoacrylate	W	1 hr	Skin/cart/bone	1,000–1,200	[95]
<i>i</i> -Bu-2-cyanoacrylate ^c		2 hr	Tooth-plastic	2,000–3,000	[95]

(Continued)

Table 2 Bond strengths of natural and reference adhesives (*Continued*)

Material	Environment	Bond time	Surfaces	Bond strength	Reference ^a
Et-2-cyanoacrylate	M	3–24 hr	Skin	990–2,010	[97]
Et-2-cyanoacrylate	W	12–48 hr	Skin	960–1,270	[97]
<i>Shear strength^d</i>				(kPa)	
Amphibian					
<i>Notaden</i> exudate/solid	D	1 day–1 week	Wood	1,700–2,800	[62,63]
Marine					
Sea cucumber tubules	[W]		Paraffin wax	15	[120]
MAP (Cell-Tak)			Cornea	<26	[95]
MAP + fillers, XL	W + D	10–20 min	Skin, cornea	[14–15]	[96]
Limpet—mobile	D		Glass	2.3	[2]
Barnacle—adult	W	<40 days	Plastic	480	[124]
Limpet—adhered	D		Glass	125	[2]
Periwinkle—adhered	[M]		Glass, plastic	73	[34]
Brown macroalga—live	W		Slate	>333, >422	[123]
Bacterial EPS					
Montana C-902 EPS			Al, epoxy glass	1,655–5,516	[38]
Montana EPS (31%)	M	1 week	Wood	12,200–14,500	[39]
<i>Alteromonas</i> LST			Steel, Al, wood	276–545	[39]
Analogues and reference materials					
Fibrin	W	1 hr	Skin	8–36	[95]
Fibrin	W	2 hr	Skin	≤0.7–27	[125]
BioGlue™ (alb/gluteral)	D		Pericardium	40	[126]
Gelatin (20%)	M	3 days	Aluminum	2,060	[118]
Gelatin, XL	W	2 hr	Skin	12–23	[125]
(XYK) _{19–42} XL (20%)	D	30 min	Skin	8–12	[106]
(YK) _{19–42} XL (20%)	D	30 min	Skin	9	[106]
(GK) _{19–42} XL (20%)	D	30 min	Skin	5	[106]
(AKPSYPPTYK) ₁₁ XL	D	30 min	Skin	10	[106]
(GPKTYPPTYK) ₁₁ XL	D	30 min	Skin	11	[106]
(AKPSYPPTYK) _{8–12}	M	3 days	Aluminum	294	[118]
(Lys) _{460–980}	M	3 days	Aluminum	200–300	[118]
Poly(DL-Ala)	M	3 days	Aluminum	780	[118]
(Lys) ₃₂₆₀	M	3 days	Aluminum	690–1,670	[118]
Clear gum	D	1 days	Wood	2,900	[62]
Dextrin (50%)	D	1 week	Wood	4,100	[39]
Na alginate (17%)	D	1 week	Wood	9,000	[39]
Pullulan (33%)	D	1 week	Wood	13,600	[39]
Chitosan, XL ^c			[Biol tissue]	3	[125]
PVA	D	1 day, 1 week	Wood	1,300–4,900	[62,63]
PVA (45%)	M	1 week	Wood	13,000–19,700	[39]
Glue Stick	D	1 day, 1 week	Wood	900–1,200	[62,63]
Et-2-cyanoacrylate	D	1 day, 1 week	Wood	1,700–3,200	[62,63]
<i>n</i> -Bu-2-cyanoacrylate	W	1 hr	Cartilage	700	[95]

(Continued)

Table 2 Bond strengths of natural and reference adhesives (*Continued*)

Material	Environment	Bond time	Surfaces	Bond strength	Reference ^a
<i>Peel strength^c</i>				(N/m)	
Ascidian—live	W		Shell, plant	4.7–6.9	[127]
<i>Notaden</i> exudate	W	1 day	Meniscus	97	[62,115,116]
Fibrin	W	1 day	Meniscus	20	[62,115,116]
GRF	W	1 day	Meniscus	39	[62,115,116]
<i>n</i> -Bu-2-cyanoacrylate	W	1 day	Meniscus	149	[62,115,116]

^aFor brevity, some references are not the original sources of the data but do cite the originals.

^bNormal tenacity, i.e., force exerted at right-angles to adhesive layer (butt joint).

^cMixture of tensile and shear data, or uncertain which format was used.

^dForce exerted parallel to the adhesive layer so that the surfaces slide over one another (lap joint).

^eJoint undergoes progressive separation from one end to the other in response to a force applied (usually at right-angles to adhesive layer) to one end of a flexible adherend.

Abbreviations: W, wet or humid environment (relative humidity of $\geq 80\%$); D, dry environment (relative humidity of $\leq 40\%$); M, midrange environment (relative humidity 41–79%); WD, moist bonding environment followed by rapid test in dry; W + D, bonding or testing in a variety of environments; %, grams per 100 ml solution; MAP, mussel adhesive proteins; XL, covalently cross-linked; ST, soft tissue; cart, cartilage; GRG, gelatin/resorcinol/glyoxal; GRF, gelatin/resorcinol/formaldehyde; MAP, mussel adhesive proteins, without covalent cross-linking unless otherwise indicated; alb/gluteral, albumin cross-linked with gluteraldehyde; P, hydroxyproline; Y, Dopa; PVA, polyvinylacetate; [] indicates an assumed condition or interpolated value. Where a material was tested with different surface textures, the result for the smoothest one is cited.

β -turn spiral (nanospring) that is thought to be highly elastic.^[4,6] There is, however, no uniformity in the secondary structures encountered in protein-based adhesives. Some (e.g., mantid oothecal protein and keratin-like tick proteins) are believed to form α -helical coiled coils, others (e.g., silk sericins, *Drosophila* mating plug protein) probably contain β -sheets or β -turn spirals, while yet others (e.g., collagen-like tick proteins, mussel glue repeat motifs) are likely to adopt poly-Pro type II helices. In terms of tertiary structure, the need to form an interpenetrating polymer network suggests that proteins will be more useful in long extended conformations than in compact globular folds.

AQUATIC ADHESIVES

Mussel Glue

The adhesive plaque material of *Mytilus* sp. has been under detailed scrutiny for several decades and has been reviewed extensively (e.g., Refs. [3,7–10]). The components of *Mytilus* adhesive plaque are mostly basic polyphenolic proteins that are soluble in dilute acids. The genes encoding many of the *Mytilus edulis* foot proteins (Mefp) have been cloned and sequenced; the gene products span the size range 5–120 kDa and typically contain repetitive sequences. Selected Tyr, Pro, and Arg residues in these proteins undergo PTMs (hydroxylations) so that they are present as Dopa, Hyp, and hydroxyarginine residues, respectively, in the mature proteins. The Dopa content of the various Mefps varies from 3 to 30 mol%, and the proteins with

the highest Dopa contents (Mefp-3 and -5) tend to predominate at the interface with the substratum.

Tyrosinase (also known as catechol oxidase or diphenol oxidase) can convert peptidyl Tyr residues into Dopa residues, and can further convert Dopa into *o*-quinone or its tautomer, α,β -dehydroDopa.^[8,9] Recently, Jonathan Wilker's group at Purdue has confirmed (for mussel plaque, Mefp-1 and -2, and model peptides) that the Dopa residues chelate iron, mainly as Fe(Dopa)₃ (Fig. 1).^[11] This complex alone can account for much protein–protein cross-linking (cohesion). However, the iron also facilitates the oxidation of Dopa by molecular oxygen to generate quinone-like radicals; these radicals may possibly go on to react with each other or with Dopa to form C–C linked rings (cohesion), or alternatively they may react with the substratum (adhesion).^[8] While the generation of Dopa in mussel plaque is probably enzymatic, further conversions associated with curing need not be. Indeed, Monahan and Wilker found that optimal curing (i.e., hardening and darkening) of Dopa-containing Mefp-1 and -2 was achieved by metal ion oxidizing agents (Fe³⁺, MnO₄[−]), and that enzymes (tyrosinase, laccase) were completely ineffective.^[12]

Other interactions may also serve to enhance the adhesive and cohesive strength of mussel plaque. During adhesion, the numerous hydroxyl groups of the Dopa/Ser/Thr/Hyp/Hyl residues in the mussel adhesive proteins may compete with and displace surface-bound water molecules, leading to zipper-like chemisorption of extended polypeptides.^[7] The Dopa side chain is particularly effective at hydrogen bonding, and it also chelates metals/metal oxides, allowing binding to mineral surfaces. Dopa and other residues

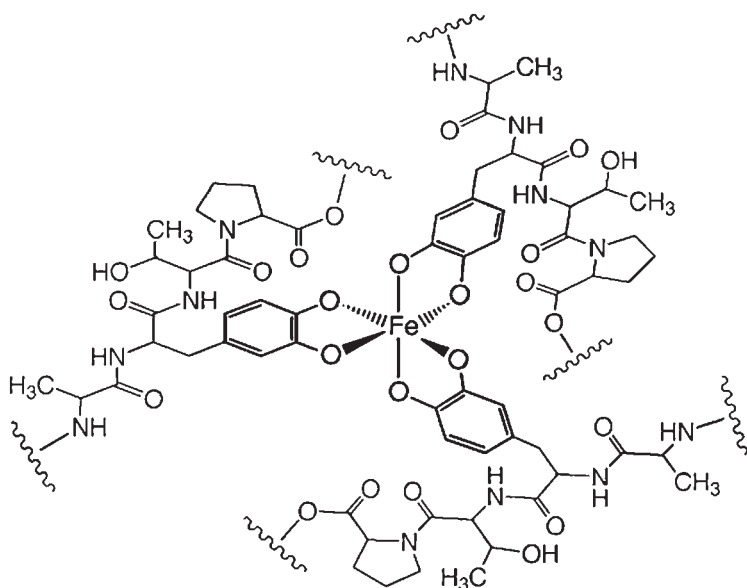


Fig. 1 Metal-protein cross-link proposed for mussel adhesive proteins. (From Ref. [11].)

(e.g., hydroxyarginine) can engage in pi-based interactions that allow binding to aromatic surfaces.^[3,9] The numerous basic residues in the mussel adhesive proteins promote binding to natural surfaces, which typically carry a net negative charge. Although the reaction of Dopa/quinones with Lys amino groups, Cys thiols, or any of the numerous hydroxyl groups is no longer thought to occur as a primary mechanism of cross-linking, Lys amino groups are now believed to form Schiff bases with completed diquinone cross-links.^[3] It is thought that the covalent cross-linking of Dopa residues probably enhances cohesion at the expense of adhesion.^[3] The interactions between surface-bound Mefp-1 and Mefp-2, including effects due to cross-linking, have recently been investigated.^[13]

The properties of the key proteins in the adhesive plaque of *M. edulis* may be summarized as follows:^[3,7-9]

- Mefp-1 (110 kDa, calculated $pI > 10$, 13 mol% Dopa) contains up to 80 repeats of the decapeptide Ala-Lys-Pro-Ser-Tyr-Hyp-Hyp-Thr-Dopa-Lys, and probably adopts a poly-Pro type II helix. Mefp-1 is found as a protective lacquer at byssus-seawater interfaces. Each molecule binds 50–75 Fe^{3+} with huge affinity, and for this reason Mefp-1 has been likened to armor plating. It is strongly adhesive in vitro and is considered to be the main ingredient of Cell-Tak™, the cell/tissue adhesive sold by BD Biosciences.^[3] An NMR study of the isolated Mefp-1 decapeptide concluded that it was a bent helix based on a type III β -turn involving the Hyp residues, possibly stabilized by hydrogen bonding and Tyr/Dopa ring stacking.^[14] More recent NMR of a synthetic decapeptide (with Tyr in place of Dopa) robustly concluded that it adopted a standard poly-Pro type II helix, which

leaves all of the hydroxyls free for adhesive or cohesive interactions.^[15]

- Mefp-2 (40 kDa, calculated $pI \sim 9$, 3 mol% Dopa) is the dominant bulk adhesive (to 40% w/w of plaque) and contains EGF domains interspersed with acidic Dopa-rich regions.
- Mefp-3 (6 kDa, calculated $pI > 11$, 20 mol% Dopa) is the surface primer; it is a family of >10 proteins rich in hydroxyarginine and Dopa.
- Mefp-4 (80 kDa, 4 mol% Dopa) is the other bulk adhesive; it is elevated in Gly, Arg, and His, contains Dopa, and may contain Tyr-rich octapeptide repeats.
- Mefp-5 (9 kDa, calculated $pI \sim 9$, 30 mol% Dopa) has >75 mol% of its Dopa residues adjacent to basic residues (e.g., Dopa-Lys) and contains phosphoserine (Ser-P) in putative mineral-binding motifs.^[16]

Barnacle Cement

Unlike mussel adhesive proteins, barnacle cement proteins do not contain Dopa; rather, they are rich in Gly, Ser, Thr, and Ala residues.^[7,17] The adhesive strength of barnacle cement is comparable to that of the mussel glue, and may even exceed it (Table 2).^[17,18] Barnacle cement proteins are not solubilized by 8 M urea or 6 M guanidinium hydrochloride, but they generally do dissolve when heated in the presence of sodium dodecyl sulfate (SDS) and reducing agents.^[17] Interestingly, adult barnacles of different species appear to use different adhesive proteins, as described below, and at least eight gene sequences have been patented.^[18]

Balanus eburneus

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) revealed constituent proteins

with molecular masses of 7, 22, 36, 52, and 58 kDa. Oddly, an 80 kDa species (apparently with the same N-terminus as the 36 kDa species) was also detected when samples had been reduced with 2-mercaptoethanol.^[17] Amino acid compositions have been determined and many peptide sequences have been obtained. The 36 kDa protein is the most abundant one, and is elevated in Tyr (10 mol%), Pro (10 mol%), and Leu.

Semibalanus balanoides

Like the cement of *B. eburneus*, the base plate and hemolymph of *S. balanoides* contains a prominent protein around 35 kDa in size. This has been cloned and sequenced; the encoded protein has homology to filensin, an intermediate filament protein found in eye lenses, and to pernin, a self-aggregating mussel hemolymph protein.^[18,19] Toward the C-terminus, there are repeats of a His-Asp-Asp-His motif which may bind metals, perhaps allowing adhesion to metal surfaces or metal-mediated protein cross-links.

Balanus crenatus

Initially, major proteins were seen at 84, 50, 47, 39, and 15–19 kDa in SDS-PAGE of reduced samples.^[20] In later work by the same group, more than six defined bands (3–50 kDa) were observed in SDS-PAGE of nonreduced samples, but reduced samples gave only a smear which was dominated by material of <5 kDa.^[17] Accordingly, it was suggested that the upper smear comprised multimers of a <5 kDa monomer corresponding to the *B. eburneus* 7 kDa protein, and that the proteins homologous to the other *B. eburneus* ones seen earlier must have remained insoluble.^[20] The cured cement is a trabecular mesh.^[20] One patent reports a 35 kDa protein for *B. crenatus* that has a similar N-terminal sequence to the corresponding protein from *B. improvisus* and *S. balanoides*.^[19]

Balanus perforatus

This cement seems to be a mixture of very hydrophobic proteins cross-linked through disulfide-linked Cys residues. Sodium dodecyl sulfate-PAGE revealed bands of 65, 62, 39, 24 and 15 kDa; the N-terminal sequence of the 39 kDa protein has been determined.^[20]

Megabalanus rosa

This cement can be completely solubilized by 7 M guanidinium hydrochloride containing 0.5 M

dithiothreitol at 60°C. Three proteins (Mr_{cp}-100k, Mr_{cp}-68k, and Mr_{cp}-52k) account for >90% (w/w) of the cement, while a fourth (Mr_{cp}-20k) accounts for >2% (w/w).^[21,22]

Mr_{cp}-100k was cloned and sequenced; it contains no repeat sequences, and has a high content of hydrophobic residues dominated by Leu (11 mol%) and Gly (10 mol%).^[21,23] Interestingly, the calculated pI value rises along the polypeptide chain, ranging from pI~5 (near the N-terminus) to pI~11 (near the C-terminus). The low Cys content for this protein (1.4 mol%) is at odds with the need for reducing agents to solubilize it. The investigators speculate that this protein may form β-strands with alternating hydrophobic and hydrophilic residues. A cyanogen bromide fragment from *B. eburneus* cement has homology to Mr_{cp}-100k.

Mr_{cp}-68k, which was originally called the “60 kDa protein,” is rich in Ser, Thr, Ala, and Gly.^[24] In the absence of 2-mercaptoethanol a band of ~180 kDa is the only one seen in SDS-PAGE, suggesting that the protein forms disulfide-mediated trimers. A patent specification provides sequence information for a 47 kDa protein that seems to be a truncation product of Mr_{cp}-68k.^[25] Another patented sequence (546 residues) is also likely to relate to Mr_{cp}-68k.^[26]

Mr_{cp}-20k accounts for >2% (w/w) of *M. rosa* cement.^[22,27] It readily dissolved without reduction and was isolated by reversed-phase high-performance liquid chromatography. The DNA-derived amino acid sequence indicates a protein rich in Cys (17 mol%) and charged residues (Asp, Glu, and His all feature at ~10 mol%) with a calculated pI~4.7. Somewhat surprisingly, all of the Cys residues are involved in intramolecular (rather than intermolecular) disulfides. The protein's sequence comprises six repeats based loosely on the motif, [Xaa]₄₋₁₄CysAsn[Xaa]₂[His/Ala]ProCys-Tyr[His/Arg]Cys[Xaa]₅₋₁₀CysXaaCysXaaCys. Interestingly, the Cys-Xaa-Cys-Xaa-Cys motif is shared with silk protein sp185 of *Chironomus* sp., which is used in making the larval tubes of these aquatic midges. There are also similarities with the mussel protein Mefp-2, which has clusters of acidic residues and contains multiple EGF domains featuring six conserved Cys residues. The presence of Mg²⁺, Mn²⁺, Ca²⁺, or Zn²⁺ ions did not alter the electrophoretic mobility of Mr_{cp}-20k. Above 0.5 M NaCl, synthetic peptides based on the Mr_{cp}-20 repeats showed self-assembly into fibrils that were hundreds of nanometers long and did not appear to involve β-elements.^[28]

Sea Star and Sea Cucumber Adhesives

The podial adhesive secretion of the sea star *Asterias rubens* is an insoluble matrix that when dry comprises

40% (w/w) inorganic material (presumably sea salts), and contains 2.6 times more protein than carbohydrate. The material has somewhat higher levels of Gly (9.7 mol%), Pro, Ile, and Cys than the average eukaryotic protein.^[29] Immunohistochemistry indicates that a similar adhesive is used by many species of sea star.^[30] In terms of gross composition, the *Asterias rubens* adhesive resembles pedal mucus from the limpet *Patella vulgata*.^[29] The podial discs of sea stars and sea urchins adapt their surface topography to mirror substratum roughness, so the adhesive layer is always just a thin (nm range) film.^[31]

Many holothurians (sea cucumbers) eject tacky tubules as a defense against potential predators. The adhesive is mainly organic and comprises 60% protein and 40% carbohydrate. It is also very insoluble, with urea, SDS, and reducing agents required to achieve partial dissolution. Extracts contain around 10 main proteins (17–220 kDa); these have broadly similar amino acid compositions and are all Gly-rich (16–22 mol%). The proteins visible in isoelectric focusing gels had *pI* values of 3.0–5.5 (denatured).^[32]

Limpet and Periwinkle Mucus

The limpet *Lottia limatula* secretes a mucus that is 93% (w/w) water and contains six times as much protein as carbohydrate.^[33] The mucus of adhered limpets is twice as concentrated as that of mobile ones, and is harder to dissolve. Urea, SDS, and 2-mercaptoethanol cooperate to solubilize the mucus, which contains ~10 proteins ranging from 20 to 220 kDa that are elevated in polar and acidic residues (denatured *pI* typically 4.7–5.3). A 45 kDa protein is Gly-rich (13 mol%), while the only glycoprotein present (140 kDa) has a high Pro content (11 mol%). The mucus proteins have a strong tendency toward noncovalent aggregation. A 118 kDa protein is characteristic of adhesive mucus while a 68 kDa protein is distinctive for the nonadhesive form, but their functions are unclear.^[33]

The mucus of the marsh periwinkle *Littorina irrorata* is also mainly water, with inorganic material accounting for around half of its dry weight. The organic fraction of the trail mucus from mobile periwinkles consists more of large polysaccharides and less of proteins (those present are 59–65 kDa). In contrast, the mucus of tightly adhered snails contains equal amounts of protein and carbohydrate because of the additional (and abundant) presence of acidic proteins (denatured *pI* ~4.8) of 59 and 65 kDa (unreduced) or 36 and 41 kDa (reduced).^[34] Both proteins are lightly glycosylated and have similar amino acid compositions.^[34] It is noteworthy that the periwinkle and the limpet secretions lack the megadalton-sized

glycoproteins typical of mammalian mucus, while in both cases the adhesiveness of the secretion correlates with the presence of particular proteins.^[2]

In the absence of reducing agents, fractions containing the *L. irrorata* 36 and 41 kDa proteins triggered the stiffening or gelling of commercial polysaccharide solutions, especially polyanionic ones, and so did preparations of the proteins elevated in the adhesive forms of mucus from the slug *Arion subfuscus* (61 and 15 kDa) and the land snail *Helix aspersa* (175, 97, and 82 kDa).^[1] It is likely that the glue proteins [which comprise 25–50% (w/w) of the organic fraction of slug, land snail, and periwinkle adhesive mucus] are carbohydrate-binding agents rather than enzymes.^[1] One such protein might bridge several polysaccharide chains in a noncovalent manner.

Stickleback Spiggin

Spiggin is a proteinaceous material that is secreted by the breeding male stickleback and is used as an adhesive thread in the construction of an underwater nest. Although spiggin is encoded by a single gene, it is found as three distinct subunits, which are in fact splice variants.^[35,36] The α , β , and γ subunits are proteins of 103, 70, and 53 kDa, respectively. In kidney tissue, which is the seat of spiggin biosynthesis, the main forms of spiggin appear in Western blots as bands of 130 and 51 kDa; the larger band is thought to represent a glycosylated form of the α -subunit. In contrast, the main form of spiggin detected in bladder urine is a 203 kDa glycoprotein, which is thought to be a dimer of the 130 kDa species. Unfortunately, it is unclear whether the samples used in these gels were reduced or not. Sequence analysis reveals that the three subunits contain different numbers of domains that have high homology to the D-domains of mammalian von Willebrand factor. These are interspersed by Cys-rich regions of comparable size to the D-domains. von Willebrand factor (250 kDa) can multimerize to species >20 MDa, and does so by way of intermolecular disulfide bridges that depend on the presence of two sets of the motif Cys-Gly-Leu-Cys-Gly.^[37] As the same motifs (and/or truncated variants thereof) are present in spiggin, it is possible that the latter multimerizes in a similar way.^[35,36]

Bacterial Adhesives

The extracellular polymeric substances (EPS) secreted by bacteria as an aid to biofilm and floc formation are typically polysaccharides of high molecular mass. They can readily be produced on a commercial scale by fermentation. Xanthan gum, an EPS secreted by the bacterium *Xanthomonas campestris*, is one such

product, and Montana Biotech's C-902 Polymer is another. The latter is noncytotoxic, biodegradable, and, unlike the highly branched Xanthan polysaccharide, does not expand in water.^[38] Although produced in an aquatic environment, bacterial EPS seem to rely on noncovalent interactions and/or polymer chain entanglement for adhesion, and provide the highest bond strengths when dry. Additives and chemical modification have been used to enhance the moisture resistance of C-902 Polymer.^[39]

Pathogenic bacteria use many different cell-surface adhesive proteins (adhesins) to bind specifically to host glycan and protein determinants. Their importance to human health means that many of these molecular recognition systems have been characterized in great detail, but their discussion lies outside the scope of this entry.^[40]

Algal Adhesives

In brown macroalgae (*Fucus distichus*) and green macroalgae (*Caulerpa prolifera*) the elongating rhizoid tips of zygotes, embryos, or cut sections secrete glycoproteins (60–70 kDa) that are recognized by antibodies to vitronectin (65–75 kDa), one of the major cell adhesion proteins of mammalian tissues.^[41,42] These glycoproteins are implicated in rhizoid-substratum adhesion, which for example was inhibited by antivitronection antibodies.^[41] Rhizoid tip adhesion in *Fucus* sp. also requires a sulfated fucan called fucoidan or F2, inviting speculation that the vitronectin-like proteins bind fucoidan in the same way that vitronectin binds heparan sulfate.^[41]

Spores of the marine green alga *Ulva* sp. (formerly *Enteromorpha* sp.) attach to underwater substrata via a glycoprotein adhesive. Western blots reveal a monomer of ~110 kDa, which is thought to become progressively cross-linked after zoospore attachment. The hydrophilic swollen gel cures within 1 hr, becoming less adhesive and more cohesive, but the nature of the cross-links remains unknown.^[43,44]

Like bacteria, microalgae can secrete EPS as an aid to biofilm formation and aggregation. Moreover, the cohesive strength of many algal adhesives may be enhanced by the covalent cross-linking of their polysaccharide chains. Indeed, taxonomically diverse marine algae may share a common mechanism of adhesion in which zygotes (and sometimes eggs and/or embryos) secrete polyphenolic compounds that are activated by extracellular peroxidase enzymes to form covalent cross-links between carbohydrate fibers. For example, marine brown macroalgae use vanadium haloperoxidase to cross-link acidic polysaccharides with polymers of phloroglucinol.^[45,46]

Fungal Adhesives

Like bacteria, fungi can secrete EPS as an aid to biofilm and aggregate formation. Fungal hyphae also secrete small proteins (~100 amino acid residues) called hydrophobins, and to date six or more genes from the hydrophobin gene family have been identified.^[47] At hydrophilic/hydrophobic interfaces, class I hydrophobins self-assemble into thin (5–12 nm) films that present a hydrophobic surface to the nonpolar milieu (oil, plastic, air, etc.) and a hydrophilic surface to the polar milieu (water, hyphal cell wall, etc.). For example, the SC3 hydrophobin allows hyphae of *Schizophyllum commune* to attach firmly to Teflon[®] surfaces when grown in air, while MPG1 of *Magnaporthe grisea* and ABH3 of *Agaricus bisporus* also enable the attachment of submerged hyphae.^[48] Some hydrophobins are decorated with mannose sugars, while others are unglycosylated.^[47]

Macroconidia of the plant pathogen *Nectria haematococca* produce a hydrophobic 90 kDa glycoprotein from the spore tip.^[49] The polypeptide (whose amino acid composition is known) carries N-linked glycans, including mannose residues which seem to be essential to adhesion.^[49] The conidia of another phytopathogen, *Collectrichum graminicola*, adhere to hydrophobic substrata using a glycoprotein (apparent molecular mass ~200 kDa) whose core protein is ~116 kDa and whose dominant sugar is mannose.^[50] Thus, the *N. haematococca* and *C. graminicola* glycoproteins, which are not considered hydrophobins, may exemplify another class of fungal adhesive.^[49,50]

The oral pathogen *Candida albicans* also uses mannoprotein adhesin(s) (25% carbohydrate by weight) to attach to plastic surfaces.^[51] It can attach to host epithelia by a variety of mechanisms, including lectin-like adhesins and hyphal wall proteins that mimic the mammalian substrates of host transglutaminase enzymes.^[52] As with pathogenic bacteria, the cell-surface molecules used by fungal pathogens to recognize and colonize host tissues lie outside the scope of this entry.

Other Aquatic Adhesives

Some parasitic marine monogeneans (flatworms) produce insoluble protein secretions for temporary attachment; this material is Dopa-free and rich in Gly (12–19 mol%) and Ala.^[53] An oyster adhesive features a protein that contains little Cys, some Pro (but no Pro-OH), and a large complement of Gly (25 mol%).^[54] In abalone shell, the insoluble organic adhesive between layers of nacre contains lustrin A, a modular protein that comprises a single domain with a huge Ser content and 10 Cys-rich domains interspersed

by eight Pro-rich domains.^[55] The reef-building polychaete *Phragmatopoma californica* secretes a protein mortar rich in Ser-P (25 mol%), Gly, and Lys. It contains a highly acidic Ser-P protein and two proteins of 10–20 kDa (Pc-1 and Pc-2; calculated $pI \geq 9.7$) that feature repeat motifs rich in Dopa, Gly, and Lys.^[56,57] The aquatic larvae of the blackfly *Simulium ornatum* are anchored by a proteinaceous adhesive secreted from their labial glands. The main proteins are 70, 40, and 20 kDa.^[58] N-terminal sequences have been determined for the two larger proteins, and these contain two- and three-residue repeats of particular residues, such as Tyr.^[58] The blackfly larval adhesive is considered to be a form of silk. Most anuran (amphibian) larvae have a transient cement gland that secretes a mucous adhesive for underwater attachment.^[59]

TERRESTRIAL ADHESIVES

Salamander and Frog Glue

Some adult amphibia secrete tacky exudates that probably form a defense against potential predators.^[60] Salamander skin exudates are proteinaceous but fail to congeal when secreted directly into water.^[61] The adhesive strength of skin secretions from adult salamanders, frogs, and other amphibians has been measured (Table 2).^[60]

When provoked, Australian frogs of the genus *Notaden* secrete a sticky nontoxic exudate.^[62] This sets rapidly into an elastic solid (Fig. 2) that is actually a

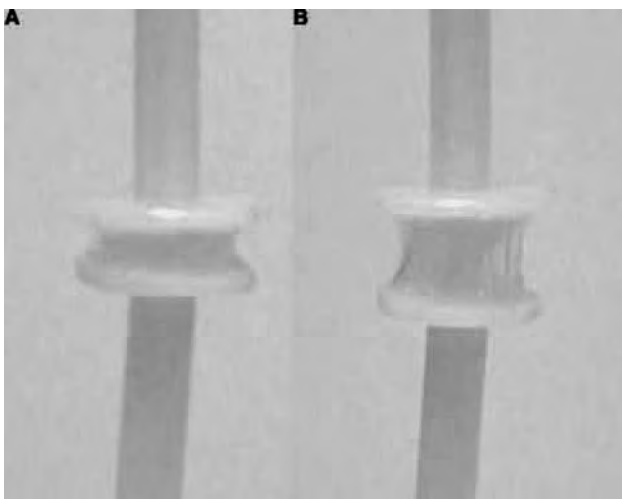


Fig. 2 Elasticity of frog glue. A sample of moist glue from *Notaden bennetti* was adhered to plastic test discs (1 cm diameter) and the joint photographed without stress (A) and under tension (B), as described in Ref. [63]. [Photograph courtesy of CSIRO Molecular Science (Australia) and Adelaide Research and Innovation Pty. Ltd. (Australia).]

hydrogel (85–90% water by weight) and functions as a pressure-sensitive adhesive.^[63] It adheres tightly, even in moist conditions, to a wide range of materials including glass, plastic, metal, wood, skin, and Teflon.^[62,64] Dry glue contains little carbohydrate and consists mainly of proteins. The protein complement is rich in Gly (16 mol%), Pro (9 mol%), and Glx (14 mol%), and contains some Hyp (5 mol%) but no Dopa.^[63] Sodium dodecyl sulfate-PAGE reveals a characteristic pattern of proteins spanning 13–400 kDa. The largest protein (apparent molecular mass 350–500 kDa) is also the most abundant, and this protein appears to be the key structural component.^[63] The solid glue can be dissolved in 5% (v/v) acetic acid; raising the ionic strength causes the glue components to self-assemble spontaneously into a solid that resembles the starting material.^[63] The solid glue was shown by scanning electron microscopy to consist of an open, microporous meshwork that—on hydration—would hold a large volume of water. The material's high extensibility may derive from β -turn spiral nanosprings and reflect the relative lack of covalent cross-linking within the material.^[63] The glue may have elements in common with self-assembling amphiphilic peptides, silks, elastin, and other biological adhesives.^[63]

Tick Salivary Cement

Some blood-feeding ectoparasites secure themselves to their hosts using nonimmunogenic adhesives that bind mammalian tissues with great strength. For example, the cement cone of the ixodid tick *Rhipicephalus appendiculatus* consists of two components, a rigid quick-hardening core and a flexible cortex that serves as a hinge or gasket. Some seven structural proteins from the cement cone have been cloned. Generally, the proteins are rich in Gly and Pro or Tyr, and many have repeated motifs.^[65] They have similarities with keratins, collagens, and glutenins; the resemblance to mammalian structural proteins may help the tick evade immune surveillance. Curiously, one protein (*R. appendiculatus* clone 1) is homologous to a reef cement component from the polychaete *Phragmatopoma californica*, seemingly the Pc-1 protein.^[56,65] Cross-linking enzymes may play a role, but this has not been demonstrated. Various regions from a collagen- and keratin-like cement protein named 64P were expressed recombinantly in bacteria, purified, and used to raise antibodies.^[65,66] These and other recombinant cement proteins may be useful for immunizing livestock against tick infestations.

Other researchers have also cloned, sequenced, and expressed tick cement proteins with a view to

producing antitick vaccines.^[67,68] These proteins again resemble collagens and other vertebrate structural proteins, and one of them (RIM36) contains very obvious repeat motifs. A Gly-rich protein (Is9) expressed by feeding *Ixodes scapularis* ticks has homology with both the 64P and RIM36 proteins from *R. appendiculatis*.^[69]

Fruit Fly Larval Glue

The salivary secretions used by the fruit fly *Drosophila melanogaster* to affix its puparium to the substratum contain some five to seven major proteins with apparent sizes of 8–360 kDa, whose polypeptides are encoded by the *sgs* and *ng* gene families.^[70,71] *Drosophila* glue proteins usually have a high Thr content, and many are heavily glycosylated. For example, the proteins derived from the *sgs3* and *sgs4* genes of *D. melanogaster* and the *lgp1* and *lgp3* genes of *D. virilis* comprise tandem repeats of (typically 5- to 12-mer) motifs that are hugely Thr-rich and sometimes also contain Pro.^[72,73] Moreover, while the *sgs4* gene encodes a polypeptide of just 32 kDa, the secreted glycoprotein has an apparent molecular mass of ~100 kDa.^[71] For *lgp1*, the core protein (~138 kDa in SDS-PAGE) is decorated on Thr residues with O-linked hexose-rich glycans to give a glycoprotein (~230 kDa).^[74] The water-binding capacity of the large carbohydrate component in these glycoproteins is thought to account for their stickiness.^[73] *D. nasuta* larvae produce particularly abundant glue glycoproteins that have been characterized to some extent.^[75] However, most work on *Drosophila* larval/salivary glue proteins has focused on the regulation of their expression during development rather than on the proteins themselves.

Insect Mating Plugs

PEB-me is the major protein in the mating plug of *D. melanogaster*.^[76] It is a basic 377-residue protein (calculated $pI \sim 9.7$, 38 kDa) that is rich in Gly (22 mol%) and Pro (17 mol%), and its central region contains tandem Pro-Ser-Pro-Gly-(Gly/Glu) repeats that are somewhat reminiscent of spider flagelliform silk fibroin. PEB-me is autofluorescent.^[76] The excitation/emission filters used to image the protein in vivo (340–360/450–460 nm) can also be used to monitor the formation of dityrosine cross-links (328/410 nm) in proteins, so the presence of a 28-residue segment containing six Tyr residues suggests that the autofluorescence of monomeric PEB-me in SDS-PAGE gels might be caused by intramolecular dityrosine linkages.^[77] However, PEB-me does not seem to engage in covalent intermolecular cross-linking.^[76]

Nonprotein mating plugs are used by some other insects; for example, bumblebees use a sticky mixture of fatty acids.^[78]

Silk Sericins

Silk sericins are polar silkworm proteins with a high Ser content (average 38 mol%) that act as a glue for silk threads in the cocoon. Circular dichroism indicates that renatured sericins have a β -sheet structure.^[79] Transcripts from two sericin genes undergo differential splicing to generate a range of protein sizes up to 400 kDa.^[80] One gene (*Ser 1*) consists of 70 repeats of a 38-residue motif. Synthetic genes expressing up to 20 repeats of the motif were expressed in *E. coli*, but the largest product contained only 12 repeats—a situation reminiscent of bacterial silk production.^[81] All of the recombinant proteins migrated anomalously slowly in denaturing PAGE. Removing urea from the recombinants caused the self-assembly of amyloid fibrils. Natural sericins, which are less repetitive in sequence, are thought to be rich in β -sheet but nonamyloid. A propensity for β -style hydrogen bonding is suggested as the basis for the silk-gluing ability of sericins.

Mantid Oothecal Protein

The fibrous protein found in the ootheca of mantid insects is a coiled-coil protein.^[82] The mantid's left colleterial (i.e., female accessory) gland stores this structural protein in membrane-bound globules as stacked sheets of parallel α -helical fibrils.^[83] On extrusion, the globule membranes coalesce and the contents become a viscous mass. Over 30 min the fibrils change into long flat ribbons, and the final form of the ootheca consists mainly of this ribbon material.^[84] Globule contents can be induced to form ribbons in vitro by adding CaCl_2 .^[83] The ribbons consist of close-packed two-stranded ropes (α -helical coiled coils).^[82,84] Alongside this structural conversion, some covalent cross-linking takes place, as follows. The left colleterial gland secretes a phenoloxidase enzyme along with the precursors of sclerotizing agents, such as 3-*O*- β -glucosides of *N*-acyl-Dopamines (NADAs). The right colleterial gland secretes a β -glucosidase enzyme, which liberates NADAs that are then oxidized by the phenoloxidase to NADA-quinones, which may covalently cross-link proteins directly. In addition, another enzyme isomerizes the NADA-quinones to highly reactive intermediates (possibly quinone methides) that react with nucleophilic groups and thereby covalently cross-link the structural protein in other ways.^[83,85]

Other Terrestrial Adhesives

Silk dope is often used by insects for permanent attachment, and a possible aquatic example has been mentioned above, but a discussion of silk lies outside the scope of this entry.^[58] In several orders of insects, the accessory glands (or sometimes the oviducts) of gravid females produce a glue that attaches the eggs to each other or to the substratum. In some cases, this material takes the form of a protective jelly (if moist) or varnish (if dry) and is called spumaline.^[86] Trichopteran spumaline contains polysaccharide but almost no protein, and the same seems to apply to at least some lepidopteran egg attachment glues.^[86,87] The adhesive mucus of terrestrial slugs and snails has been dealt with above, alongside limpet and periwinkle mucus.^[1] Other organisms use glues to capture prey. Thus, cribellate spiders catch insects using silk threads coated with a sticky aqueous solution that contains glycoproteins (including a major one of ≥ 65 kDa), which appear rich in Gly (22 mol%) and Pro (16 mol%).^[88–90] Likewise, onychophorans (velvet-worms) eject a sticky fluid to immobilize their prey. This material is mainly protein; large collagen-like species (> 600 kDa) were tentatively identified with elastic fibers in the secretion, while small proteins (which, when reduced, were dominated by a 15 kDa species) were presumed to be components of the adhesive fluid.^[91] Last, rove beetles (*Stenus* sp.) capture prey using paraglossae coated with an adhesive secretion that is probably a mixture of protein and lipid.^[92]

APPLICATIONS IN WOUND HEALING AND TISSUE REPAIR

Currently, most of the biological glues used in medicine are based on fibrin, albumin, or gelatin and exhibit low bond strength. Synthetic glues, such as cyanoacrylates, are very strong, but they are somewhat toxic and form rigid nonporous films that can hinder wound healing. In contrast, many of the natural adhesives discussed above are strong and potentially biocompatible. Permanent adhesives from aquatic organisms (and some from terrestrial endo- and ectoparasites) should be able to provide high bond strengths in physiological environments. Bioadhesives with lower bond strengths may still prove useful as sealants and hemostats, while temporary/reversible adhesives from aquatic or terrestrial sources may find niche applications in areas such as wound dressings. In addition to biocompatibility, many natural adhesives possess elasticity and microscopic porosity. Indeed, electron micrographs of mussel plaque, barnacle cement, polychaete cement, monogenean adhesive, sea

star footprints, and solidified frog exudate all reveal a sponge-like ultrastructure that, in a biomedical context, should permit gas and nutrient exchange, and may support the cellular infiltration needed for eventual resorption.^[18,20,29,53,57,64] Some of the steps taken toward adapting natural adhesives for biomedical use are described below.

Mussel Glue

Polyphenolic adhesive proteins are widely reported to be both nontoxic and biodegradable. Moreover, difficulties in raising antibodies against Dopa-containing mussel adhesive proteins suggest that they have low immunogenicity, another desirable property from a biomedical perspective.^[3,7] In some circumstances, however, mussel adhesive proteins seem to induce apoptosis.^[93]

Air-dried films of *M. edulis* adhesive protein have been tested for mucoadhesion and found to be as good or better than polycarbophil, one of the best-known mucoadhesive polymers.^[94] Purified extracts containing mussel adhesive proteins have also been used directly as tissue bonding agents. Chivers and Wolowacz tested the mussel adhesive Cell-Tak in submerged pigskin butt joints without a cross-linking enzyme, and reported a tensile strength of only 0.0003 MPa after 22 hr (Table 2).^[95] Under similar conditions they obtained general tensile strengths of cyanoacrylate (1 MPa) > gelatin/resorcinol/formaldehyde (10^{-1} MPa) > fibrin (10^{-2} MPa). Benedict and Picciano describe formulations in which purified mussel adhesive proteins are mixed with fillers (e.g., collagen), cross-linked using tyrosinase or chemical agents, and tested (over 10–60 min) for the ability to bond meniscal cartilage, vertebrae, lap-jointed skin, and lap-jointed corneal tissue.^[96] The adhesive strength range observed by these investigators (0.002–0.015 MPa) is slightly higher than that reported by Chivers and Wolowacz. In contrast, Ninan et al. claim that the adhesive strength of a mussel extract in butt joints of pigskin reached a maximum of ~ 0.4 MPa (at 40% humidity) or ~ 1 MPa (at 80% humidity) in the absence of cross-linking enzymes, the latter result being similar to the strength of control joints formed using a cyanoacrylate glue.^[97] The bond strengths in this study were generally much higher than those reported by Benedict and Picciano or Chivers and Wolowacz. At the lower humidity, Tisseel[®] fibrin glue was usually twice as strong as the mussel extract, and both glues disimproved beyond 24 hr because of dehydration of the skin. At higher humidity, the mussel extract reached its maximum strength faster (24 hr), and the two glues were at full strength at 48 hr. Inventors continue to file patent specifications directed toward the use of polyphenolic protein

adhesives in medical and other applications, but to date no such product has been approved for use in surgical procedures.

Recent strategies to generate polypeptides based on mussel adhesive protein motifs have included polycondensation of the active esters of synthetic peptides followed by tyrosinase-mediated cross-linking.^[98] Other researchers have focused on more biotechnological approaches. Some have simply opted for the recombinant expression of natural genes that encode mussel adhesive proteins, treating the resulting proteins with tyrosinase to introduce Dopa and promote cross-linking.^[99,100] In this regard, it is worth remembering that ascorbate inhibits the ability of tyrosinase to oxidize Dopa to quinones.^[101] Others have constructed synthetic genes that express artificial proteins containing tandem repeats of the canonical decapeptide repeat of Mefp-1; tyrosinase treatment of such gene products results in cross-linking and insolubilization.^[102,103] Some investigators have suggested the use of plants as expression hosts to encourage posttranslational hydroxylation.^[104] While this may enhance the Hyp content of recombinant Mefp-1 analogs, it is unlikely to introduce Dopa residues as these are not known to occur in plant proteins. Nevertheless, it is interesting to realize that the repetitive proline-rich proteins and nodulins of plants are based on high-order repeats of Pro-Hyp-Val-Tyr-Lys, a sequence that resembles the C-terminal half of the Mefp-1 decapeptide motif.^[105]

Some researchers have focused on multimerizing and cross-linking synthetic peptides that are rich in Tyr and/or Dopa without attempting to mimic the sequences actually found in the mussel proteins. For example, Tatehata et al. synthesized (Xaa-Tyr-Lys)₁₉₋₄₂ and (Tyr-Lys)₁₉₋₄₂ polypeptides, cross-linked them enzymatically using tyrosinase, and measured the resulting adhesive strength in pigskin lap joints. These polydipeptides and polytripeptides provided similar adhesive strength to synthetic polydecapeptides based on the repeat sequences found in mussel adhesive proteins (Table 2).^[106] Subsequently, the same workers found that a formulation of poly(Gly-Tyr-Lys) with tyrosinase was less immunoreactive than Tisseel at closing an incision wound in a live pig.^[107]

Other researchers have avoided the use of polypeptides and focused instead on incorporating Dopa moieties into synthetic polymers. Huang et al. included Dopa residues as end groups in synthetic pluronic acids and found that this led to enhanced mucoadhesion.^[108] The same group has developed a method whereby *N*-methacrylated Dopa monomers mixed with poly(ethylene glycol) (PEG) diacrylate can be copolymerized to form Dopa-containing PEG hydrogels.^[109]

Bacterial, Algal, and Fungal Adhesives

Although most applications of bacterial EPS have been as commodity adhesives, a bacterial EPS from the marine bacterium *Vibrio diabolicus* has proven effective as a coating and void filler in promoting bone repair. Thus, HE 800, a linear exopolysaccharide resembling hyaluronic acid, is reported to restore large bone defects and to be biocompatible, non-immunogenic, and resorbable.^[110,111]

Protein extracts from macroalgae and microalgae have been characterized in terms of amino acid composition and subjected to functional tests, including skin bonding and T-cell adhesion tests.^[112] The resulting patent claims protein-based adhesive preparations from macroalgal holdfasts that fulfill certain criteria, and which potentially include the vitronectin-like proteins discussed previously.^[41,42,112] The European Union FP5 Algal Bioadhesives Consortium (2001–2005) aims to develop materials based on algal bioadhesives for use in wound closure and drug delivery.^[113] Their multidisciplinary approach includes the purification and characterization of algal bioadhesives, as well as gene cloning and recombinant expression. Organisms of interest include *Ulva*, *Caulerpa*, *Fucus*, and *Laminaria*. The consortium also aims to develop new models of human skin and mucosal membranes, which will be used in assessing the adhesive properties of alga-derived materials.^[113] One of the industrial partners in the consortium is Biota Ltd., an Israeli company that was established to develop and commercialize surgical adhesives made from proprietary alga-derived materials.^[114]

Fungal hydrophobins do not elicit cytotoxic or immunogenic responses. These small-protein surfactants can be used to reverse surface wettability or—after chemical or genetic modification—impart specific functionalities to surfaces, such as determinants for mammalian cell attachment.^[47]

Frog Glue

The highly porous nature of the solid glue from *Notaden* sp. seems likely to allow the migration of nutrients, gases, and, perhaps, cells in clinical applications. Initial tests on the compatibility of *N. melanoscaphus* glue with animal tissue preparations showed no adverse effects.^[62,64] In cell culture tests, bovine corneal epithelial cells adhered well to tissue culture plastic that had been coated with dissolved *N. bennetti* glue, and proliferated thereon as if on a hydrophobic surface.^[64] Some initial tests of tissue bonding and wound healing efficacy have also been conducted. In organ culture experiments, dissolved *N. bennetti* glue adhered plastic lenticules to debrided bovine

corneas and permitted epithelial regrowth.^[64] The *N. melanoscaphus* glue was compared to various surgical glues by a tear propagation test on longitudinal bonded tears in excised sheep menisci.^[64,115,116] Meniscal fragments were pressed firmly together for 90 sec with adhesive, and then incubated in moist conditions at room temperature for 24 hr. These data indicated that the frog adhesive was much stronger than the commonly used fibrin or gelatin/resorcinol/formaldehyde adhesives, but not as strong as cyanoacrylate. However, the cyanoacrylate formed a hard, brittle film, while the frog glue remained elastic. A preliminary evaluation of the *in vivo* effectiveness of freshly secreted glue from *N. melanoscaphus* in repairing an artificial meniscal tear was undertaken in a trial involving 16 sheep (10 test, 6 control). After 10 weeks, it was found that the adhesive had successfully held the cut meniscal fragments together. Histology showed new, mainly fibrous, tissue with well-oriented collagen filling the gap.^[62,64]

CONCLUSIONS

There is a significant unmet clinical need for strong, elastic, and biocompatible adhesives that work in moist environments. Many natural glues provide adhesive strengths in or above the 200–700 kPa range considered desirable by the U.S. Department of Defense for an underwater adhesive (Table 2), and can in addition be biologically innocuous. Indeed, parasite adhesives (such as those from worms, ticks, and microbial pathogens) that have evolved to bond tightly to host tissues without causing immune or inflammatory responses seem particularly appropriate for development into surgical glues. The softer hydrogel bioadhesives may prove useful as sealants, void fillers, or reversible glues. In addition, the highly porous nature of many biological adhesives may allow them to function either internally or topically as depots for therapeutic agents. The challenge with all of the natural adhesives is to understand their modes of action sufficiently well that their essential features can be mimicked by simplified systems, such as those consisting of synthetic and/or recombinant components. Early tests of mussel adhesive proteins and their analogs have given mixed results, but—given recent discoveries—the performance of Dopa-based adhesives might improve greatly if Fe³⁺ is provided as a cross-linking agent. The genes encoding the structural proteins of many other types of biological adhesive have been cloned, and it is only a matter of time before their recombinant products are tested as tissue bonding agents. The results of these experiments are eagerly awaited.

ACKNOWLEDGMENTS

The author gratefully acknowledges support from the Crop Biofactories Initiative, a joint venture of CSIRO and the Grains Research & Development Corporation, and CSIRO's National Research Flagship Program.



ARTICLE OF FURTHER INTEREST

Bioadhesion

REFERENCES

1. Pawlicki, J.M.; Pease, L.B.; Pierce, C.M.; Startz, T.P.; Zhang, Y.; Smith, A.M. The effect of molluscan glue proteins on gel mechanics. *J. Exp. Biol.* **2004**, *207* (7), 1127–1135.
2. Smith, A.M. The structure and function of adhesive gels from invertebrates. *Integr. Comp. Biol.* **2002**, *42* (6), 1164–1171.
3. Waite, J.H. Adhesion a la moule. *Integr. Comp. Biol.* **2002**, *42* (6), 1172–1180.
4. Hutchinson, E.G.; Thornton, J.M. A revised set of potentials for beta-turn formation in proteins. *Prot. Sci.* **1994**, *3* (12), 2207–2216.
5. Matsushima, N.; Creutz, C.E.; Kretsinger, R.H. Polyproline, beta-turn helices. *Proteins* **1990**, *7* (2), 125–155.
6. Hayashi, C.Y.; Lewis, R.V. Evidence from flagelliform silk cDNA for the structural basis of elasticity and modular nature of spider silks. *J. Mol. Biol.* **1998**, *275* (5), 773–784.
7. Burzio, L.O.; Burzio, V.A.; Silva, T.; Burzio, L.A.; Pardo, J. Environmental bioadhesion: themes and applications. *Curr. Opin. Biotechnol.* **1997**, *8* (3), 309–312.
8. Deming, T.J. Mussel byssus and biomolecular materials. *Curr. Opin. Chem. Biol.* **1999**, *3* (1), 100–105.
9. Waite, J.H. Reverse engineering of bioadhesion in marine mussels. *Ann. N.Y. Acad. Sci.* **1999**, *875* (1), 301–309.
10. Waite, J.H.; Andersen, N.H.; Jewhurst, S.; Sun, C.J. Mussel adhesion: finding the tricks worth mimicking. *J. Adhes.* **2005**, *81* (3–4), 297–317.
11. Sever, M.J.; Weisser, J.T.; Monahan, J.; Srinivasan, S.; Wilker, J.J. Metal-mediated cross-linking in the generation of a marine-mussel adhesive. *Angew. Chem. Int. Ed.* **2004**, *43* (4), 448–450.
12. Monahan, J.; Wilker, J.J. Cross-linking the protein precursor of marine mussel adhesives: bulk measurements and reagents for curing. *Langmuir* **2004**, *20* (9), 3724–3729.
13. Fant, C.; Elwing, H.; Hook, F. The influence of cross-linking on protein–protein interactions in a marine adhesive: the case of two byssus plaque proteins from the blue mussel. *Biomacromolecules* **2002**, *3* (4), 732–741.

14. Olivieri, M.P.; Wollman, R.M.; Alderfer, J.L. Nuclear magnetic resonance spectroscopy of mussel adhesive protein repeating peptide segment. *J. Pept. Res.* **1997**, *50* (6), 436–442.
15. Kanyalkar, M.; Srivastava, S.; Coutinho, E. Conformation of a model peptide of the tandem repeat decapeptide in mussel adhesive protein by NMR and MD simulations. *Biomaterials* **2002**, *23* (2), 389–396.
16. Waite, J.H.; Qin, X. Polyphosphoprotein from the adhesive pads of *Mytilus edulis*. *Biochemistry* **2001**, *40* (9), 2887–2893.
17. Naldrett, M.J.; Kaplan, D.L. Characterization of barnacle (*Balanus eburneus* and *B. crenatus*) adhesive proteins. *Mar. Biol.* **1997**, *127* (4), 629–635.
18. Wiegemann, M. Adhesion in blue mussels (*Mytilus edulis*) and barnacles (genus *Balanus*): mechanisms and technical applications. *Aquat. Sci.* **2005**, *67* (2), 166–176.
19. Kaplan, D.L.; Gatenholm, P.; Berglin, K.M.; Platko, J.D.; Pepper, L.R.; Ngangan, A.V. Barnacle adhesion proteins. International Patent WO 03/093413, Apr 24, 2003.
20. Naldrett, M.J. The importance of sulphur cross-links and hydrophobic interactions in the polymerization of barnacle cement. *J. Mar. Biol. Ass. U.K.* **1993**, *73* (3), 689–702.
21. Kamino, K.; Inoue, K.; Maruyama, T.; Takamatsu, N.; Harayama, S.; Shizuri, Y. Barnacle cement proteins. Importance of disulfide bonds in their insolubility. *J. Biol. Chem.* **2000**, *275* (35), 27360–27365.
22. Kamino, K. Novel barnacle underwater adhesive protein is a charged amino acid-rich protein constituted by a Cys-rich repetitive sequence. *Biochem. J.* **2001**, *356* (2), 503–507.
23. Barnacle Adhesion Protein Gene—For Production of Adhesive Protein. Japanese Patent JP7,265,081, Mar 31, 1994 (Kaiyo Biotechnology Kenkyusho KK).
24. Kamino, K.; Odo, S.; Maruyama, T. Cement proteins of the acorn barnacle, *Megabalanus rosa*. *Biol. Bull.* **1996**, *190* (3), 403–409.
25. Barnacle Second Adhesive Protein Gene—Providing Protein Useful as Starting Material for Adhesive or as Substrate in Cell Culture. Japanese Patent JP9,047,288, Aug 9, 1995 (Kaiyo Biotechnology Kenkyusho KK).
26. New Barnacle Gene That Encodes the Fourth Adhesive Protein—Which May Be Used as a Bioadhesive. Japanese Patent JP10,327,867, May 30, 1997 (Kaiyo Biotechnology Kenkyusho KK).
27. Barnacle Third Adhesion Protein Gene—Useful for Producing Protein Recombinantly Used as Raw Material for Adhesives. Japanese Patent JP9,299,089, May 17, 1996 (Kaiyo Biotechnology Kenkyusho KK).
28. Nakano, M.; Shen, J.R.; Urushida, Y.; Mori, Y.; Kamino, K. Peptide-based material design from barnacle underwater adhesive proteins. In *Transactions; 7th World Biomaterials Congress*, Sydney, Australia, May 17–21, 2004; Australian Society for Biomaterials Inc.: Melbourne, Victoria, 2004; 1649.
29. Flammang, P.; Miche, A.; Cauwenberge, A.V.; Alexandre, H.; Jangoux, M. A study of the temporary adhesion of the podia in the sea star *Asterias rubens* (Echinodermata, Asteroidea) through their footprints. *J. Exp. Biol.* **1998**, *201* (16), 2383–2395.
30. Santos, R.; Haesaerts, D.; Jangoux, M.; Flammang, P. Comparative histological and immunohistochemical study of sea star tube feet (Echinodermata, Asteroidea). *J. Morphol.* **2005**, *263* (3), 259–269.
31. Santos, R.; Gorb, S.; Jamar, V.; Flammang, P. Adhesion of echinoderm tube feet to rough surfaces. *J. Exp. Biol.* **2005**, *208* (13), 2555–2567.
32. DeMoor, S.; Waite, J.H.; Jangoux, M.; Flammang, P. Characterization of the adhesive from Cuvierian tubules of the sea cucumber *Holothuria forskali* (Echinodermata, Holothuroidea). *Mar. Biotechnol.* **2003**, *5* (1), 45–57.
33. Smith, A.M.; Quick, T.J.; St. Peter, R.L. Differences in the composition of adhesive and non-adhesive mucus from the limpet *Lottia limatula*. *Biol. Bull.* **1999**, *196* (1), 34–44.
34. Smith, A.M.; Morin, M.C. Biochemical differences between trail mucus and adhesive mucus from marsh periwinkle snails. *Biol. Bull.* **2002**, *203* (3), 338–346.
35. Jones, I.; Lindberg, C.; Jakobsson, S.; Hellqvist, A.; Hellman, U.; Borg, B.; Olsson, P.E. Molecular cloning and characterization of spiggin. *J. Biol. Chem.* **2001**, *276* (21), 17857–17863.
36. Olsson, P-E.; Jaas, J.; Jones, I.; Jakobsson, S.; Borg, B. Adhesive Protein and Uses Thereof. U.S. Patent 2004/0072183, Apr 15, 2004.
37. Sadler, J.E. Biochemistry and genetics of von Willebrand factor. *Annu. Rev. Biochem.* **1998**, *67*, 395–424.
38. Combie, J. Natural polymer with adhesive properties produced by bacteria. Adhesives & Sealants Industry Website, Cover Story, Feb 2003; <http://www.adhesivesmag.com/CDA/ArticleInformation/coverstory/BNPCoverStoryItem/0,2103,101154,00.html> (accessed Jan 2005).
39. Haag, A.P.; Maier, R.M.; Combie, J.; Geesey, G.G. *Int. J. Adhes. Adhes.* **2004**, *24* (6), 495–502.
40. Remaut, H.; Waksman, G. Structural biology of bacterial pathogenesis. *Curr. Opin. Struct. Biol.* **2004**, *14* (2), 161–170.
41. Wagner, V.T.; Brian, L.; Quatrano, R.S. Role of a vitronectin-like molecule in embryo adhesion of the brown alga *Fucus*. *Proc. Natl. Acad. Sci. USA* **1992**, *89* (8), 3644–3648.
42. Levi, B.; Friedlander, M. Identification of two putative adhesive polypeptides in *Caulerpa prolifera* rhizoids using an adhesion model system. *J. Appl. Phycol.* **2004**, *16* (1), 1–9.
43. Stanley, M.S.; Callow, M.E.; Callow, J.A. Monoclonal antibodies to adhesive cell coat glycoproteins secreted by zoospores of the green alga *Enteromorpha*. *Planta* **1999**, *210* (1), 61–71.
44. Callow, J.A.; Crawford, S.A.; Higgins, M.J.; Mulvaney, P.; Wetherbee, R. The application of atomic force microscopy to topographical studies and force measurements on the secreted adhesive of the green alga *Enteromorpha*. *Planta* **2000**, *211* (5), 641–647.
45. Vreeland, V.; Waite, J.H.; Epstein, L. Polyphenols and oxidases in substratum adhesion by marine algae and mussels. *J. Phycol.* **1998**, *34* (1), 1–8.

46. Berglin, M.; Delage, L.; Potin, P.; Vilter, H.; Elwing, H. Enzymatic cross-linking of a phenolic polymer extracted from the marine alga *Fucus serratus*. *Biomacromolecules* **2004**, *5* (6), 2376–2383.
47. Scholtmeijer, K.; Wessels, J.G.; Wösten, H.A. Fungal hydrophobins in medical and technical applications. *Appl. Microbiol. Biotechnol.* **2001**, *56* (1–2), 1–8.
48. Wessels, J.G.H. Fungi in their own right. *Fungal Gen. Biol.* **1999**, *27* (2–3), 134–145.
49. Kwon, Y.H.; Epstein, L. Involvement of the 90 kDa glycoprotein in adhesion of *Nectria haematococca* macroconidia. *Physiol. Mol. Plant Pathol.* **1997**, *51* (5), 287–303.
50. Sugui, J.A.; Leite, B.; Nicholson, R.L. Partial characterization of the extracellular matrix released onto hydrophobic surfaces by conidia and conidial germ-lings of *Colletotrichum graminicola*. *Physiol. Mol. Plant Pathol.* **1998**, *52* (6), 411–425.
51. Vardar-Ünlü, G. Mannoprotein adhesin of *Candida albicans* germ tubes. *Turk. J. Med. Sci.* **1998**, *28* (5), 469–474.
52. Staab, J.F.; Bahn, Y.S.; Tai, C.H.; Cook, P.F.; Sundstrom, P. Expression of transglutaminase substrate activity on *Candida albicans* germ tubes through a coiled, disulfide-bonded N-terminal domain of Hwp1 requires C-terminal glycosylphosphatidylinositol modification. *J. Biol. Chem.* **2004**, *279* (39), 40737–40747.
53. Hamwood, T.E.; Cribb, B.W.; Halliday, J.A.; Kearns, G.C.; Whittington, I.D. Preliminary characterisation and extraction of anterior adhesive secretion in monogenean (platyhelminth) parasites. *Folia Parasitol.* **2002**, *49* (1), 39–49.
54. Yamamoto, H. Adhesive Composition with Potential use as Bioadhesive in Humans. Japanese Patent JP7,188,640, Jul 19, 1991.
55. Shen, X.; Belcher, A.M.; Hansma, P.K.; Stucky, G.D.; Morse, D.E. Molecular cloning and characterization of lustrin A, a matrix protein from shell and pearl nacre of *Haliotis rufescens*. *J. Biol. Chem.* **1997**, *272* (51), 32472–32481.
56. Waite, J.H.; Jensen, R.A.; Morse, D.E. Cement precursor proteins of the reef-building polychaete *Phragmatopoma californica* (Fewkes). *Biochemistry* **1992**, *31* (25), 5733–5738.
57. Stewart, R.J.; Weaver, J.C.; Morse, D.E.; Waite, J.H. The tube cement of *Phragmatopoma californica*: a solid foam. *J. Exp. Biol.* **2004**, *207* (26), 4727–4734.
58. Kiel, E.; Röder, T. Geoelectrophoretic studies on labial gland secretions of immature blackflies (Simuliidae, Diptera). *Limnologica* **2002**, *32* (3), 201–205.
59. Nokhbatolfighahai, M.; Downie, J.R. Larval cement gland of frogs: comparative development and morphology. *J. Morphol.* **2005**, *263* (3), 270–283.
60. Evans, C.M.; Brodie, E.D. Adhesive strength of amphibian skin secretions. *J. Herpetol.* **1994**, *28* (4), 499–502.
61. Williams, T.A.; Anthony, C.D. Technique to isolate salamander granular gland products with a comment on the evolution of adhesiveness. *Copeia* **1994**, (2), 540–541.
62. Tyler, M.J.; Ramshaw, J.A.M. An Adhesive Derived From Amphibian Skin Secretions. International Patent WO 02/22756, Sep 18, 2001.
63. Graham, L.D.; Glattauer, V.; Huson, M.G.; Maxwell, J.M.; Knott, R.B.; White, J.W.; Vaughan, P.R.; Peng, Y.; Tyler, M.J.; Werkmeister, J.A.; Ramshaw, J.A. Characterization of a protein-based adhesive elastomer secreted by the Australian frog *Notaden bennetti*. *Biomacromolecules*. Available online at <http://dx.doi.org/10.1021/bm050335e>.
64. Graham, L.D.; Glattauer, V.; Peng, Y.Y.; Vaughan, P.R.; Werkmeister, J.A.; Tyler, M.J.; Ramshaw, J.A. An adhesive secreted by Australian frogs of the genus *Notaden*. In *Biological Adhesives*; Smith, A.M., Callow, J.A., Eds.; Springer-Verlag: Berlin, *in press*.
65. Nuttall, P.A.; Paesen, G.C. Tissue Cement Proteins Produced by Blood-Feeding Ectoparasites and Related Polynucleotides. International Patent WO 99/24567, Nov 12, 1998.
66. Trimnell, A.R.; Paesen, G.C.; Nuttall, P.A. Vaccine Comprising a Tick Cement Protein. International Patent WO 01/80881, Apr 25, 2001.
67. Mulenga, A.; Sugimoto, C.; Sako, Y.; Ohashi, K.; Musoke, A.; Shubash, M.; Onuma, M. Molecular characterization of a *Haemaphysalis longicornis* tick salivary gland-associated 29-kilodalton protein and its effect as a vaccine against tick infestation in rabbits. *Infect. Immun.* **1999**, *67* (4), 1652–1658.
68. Bishop, R.; Lambson, B.; Wells, C.; Pandit, P.; Osaso, J.; Nkonge, C.; Morzaria, S.; Musoke, A.; Nene, V. A cement protein of the tick *Rhipicephalus appendiculatus*, located in the secretory *e* cell granules of the type III salivary gland acini, induces strong antibody responses in cattle. *Int. J. Parasitol.* **2002**, *32* (7), 833–842.
69. Guilfoile, P.G.; Packila, M. Identification of four genes expressed by feeding female *Ixodes scapularis*, including three with sequence similarity to previously recognized genes. *Exp. Appl. Acarol.* **2004**, *32* (1–2), 103–110.
70. Beckendorf, S.K.; Kafatos, F.C. Differentiation in the salivary glands of *Drosophila melanogaster*: characterization of the glue proteins and their developmental appearance. *Cell* **1976**, *9* (3), 365–373.
71. Korge, G. Larval saliva in *Drosophila melanogaster*: production, composition, and relationship to chromosome puffs. *Dev. Biol.* **1977**, *58* (2), 339–355.
72. Swida, U.; Lucka, L.; Kress, H. Glue protein genes in *Drosophila virilis*: their organization, developmental control of transcription and specific mRNA degradation. *Development* **1990**, *108* (2), 269–280.
73. Lanio, W.; Swida, U.; Kress, H. Molecular cloning of the *Drosophila virilis* larval glue protein gene Lgp-3 and its comparative analysis with other *Drosophila* glue protein genes. *Biochem. Biophys. Acta* **1994**, *1219* (2), 576–580.
74. Kress, H. Biochemical and ontogenetic aspects of glycoprotein synthesis in *Drosophila virilis* salivary glands. *Dev. Biol.* **1982**, *93* (1), 231–239.
75. Ramesh, S.R.; Kalisch, W.E. Glue proteins in *Drosophila nasuta*. *Biochem. Genet.* **1988**, *26* (7–8), 527–541.

76. Lung, O.; Wolfner, M.F. Identification and characterization of the major *Drosophila melanogaster* mating plug protein. *Insect Biochem. Mol. Biol.* **2001**, *31* (6–7), 543–551.
77. Graham, L.D. Unpublished data.
78. Baer, B.; Maile, R.; Schmid-Hempel, P.; Morgan, E.D.; Jones, G.R. Chemistry of a mating plug in bumblebees. *J. Chem. Ecol.* **2000**, *26* (8), 1869–1875.
79. Zhu, L.J.; Arai, M.; Hirabayashi, K. Sol-gel transition of sericin. *Nippon Sanshigaku Zasshi* **1996**, *65* (4), 270–274.
80. Takasu, Y.; Yamada, H.; Tsubouchi, K. Isolation of three main sericin components from the cocoon of the silkworm, *Bombyx mori*. *Biosci. Biotechnol. Biochem.* **2002**, *66* (12), 2715–2718.
81. Huang, J.; Valluzzi, R.; Bini, E.; Vernaglia, B.; Kaplan, D.L. Cloning, expression, and assembly of sericin-like protein. *J. Biol. Chem.* **2003**, *278* (46), 46117–46123.
82. Bullough, P.A.; Tulloch, P.A. High-resolution spot-scan electron microscopy of microcrystals of an alpha-helical coiled-coil protein. *J. Mol. Biol.* **1990**, *215* (1), 161–173.
83. Kenchington, W.; Flower, N.E. Studies on insect fibrous proteins: the structural protein of the ootheca in the praying mantis, *Sphodromantis centralis* (Rehn). *J. Microsc.* **1969**, *89* (2), 263–281.
84. Rudall, K.M. Protein ribbons and sheets. In *Lectures on the Scientific Basis of Medicine*; Athlone Press: London, 1959; Vol. 5, 217–230.
85. Yago, M.; Sato, H.; Oshima, S.; Kawasaki, H. Enzymatic activities involved in the oothecal sclerotization of the praying mantid, *Tenodera aridifolia Sinensis* (Saussure). *Insect Biochem.* **1990**, *20* (7), 745–750.
86. Hinton, H.E. *Biology of Insect Eggs*; Pergamon Press: Oxford, U.K., 1981; Vol. I, 193–196, 200, and 634.
87. Li, D.; Graham, L.D. Unpublished data.
88. Tillinghast, E.K. Selective removal of glycoproteins from the adhesive spiral of the spiders orb web. *Naturwissenschaften* **1981**, *68*, 526–527.
89. Dreesbach, K.; Uhlenbruck, G.; Tillinghast, E.K. Carbohydrates of the trypsin soluble fraction of the orb web of *Argiope trifasciata*. *Insect Biochem.* **1983**, *13* (6), 627–631.
90. Vollrath, F.; Tillinghast, E.K. Glycoprotein glue beneath a spider web's aqueous coat. *Naturwissenschaften* **1991**, *78* (12), 557–559.
91. Benkendorff, K.; Beardmore, K.; Gooley, A.A.; Packer, N.H.; Tait, N.N. Characterisation of the slime gland secretion from the peripatus, *Euperipatoides kanangrensis* (Onychophora: Peripatopsidae). *Comp. Biochem. Physiol. B.* **1999**, *124* (4), 457–465.
92. Kolsch, G. The ultrastructure of glands and the production and function of the secretion in the adhesive capture apparatus of *Stenus* species (Coleoptera: Staphylinidae). *Can. J. Zool.* **2000**, *78* (3), 465–475.
93. Benthien, J.P.; Russlies, M.; Behrens, P. Investigating the effects of bone cement, cyanoacrylate glue and marine mussel adhesive protein from *Mytilus edulis* on human osteoblasts and fibroblasts *in vitro*. *Ann. Anat.* **2004**, *186* (5–6), 561–566.
94. Schnurrer, J.; Lehr, C.M. Mucoadhesive properties of the mussel adhesive protein. *Int. J. Pharm.* **1996**, *141* (1–2), 251–256.
95. Chivers, R.A.; Wolowacz, R.G. The strength of adhesive-bonded tissue joints. *Int. J. Adhes. Adhes.* **1997**, *17* (2), 127–132.
96. Benedict, C.V.; Picciano, P.T. Adhesives Derived from Polyphenolic Proteins. U.S. Patent 5,015,677, Jun 27, 1988.
97. Ninan, L.; Monahan, J.; Stroshine, R.L.; Wilker, J.J.; Shi, R. Adhesive strength of marine mussel extracts on porcine skin. *Biomaterials* **2003**, *24* (22), 4091–4099.
98. Yamamoto, H.; Sakai, Y.; Ohkawa, K. Synthesis and wettability characteristics of model adhesive protein sequences inspired by a marine mussel. *Biomacromolecules* **2000**, *1* (4), 543–551.
99. Hwang, D.S.; Yoo, H.J.; Jun, J.H.; Moon, W.K.; Cha, H.J. Expression of functional recombinant mussel adhesive protein Mgf-5 in *Escherichia coli*. *Appl. Environ. Microbiol.* **2004**, *70* (6), 3352–3359.
100. Hwang, D.S.; Gim, Y.; Cha, H.J. Expression of functional recombinant mussel adhesive protein type 3A in *Escherichia coli*. *Biotechnol. Prog.* **2005**, *21* (3), 965–970.
101. Benedict, C.V.; Picciano, P.T. Method for Making Dopa-Containing Bioadhesive Proteins from Tyrosine-Containing Proteins. European Patent EP242,656, Apr 2, 1987.
102. Salerno, A.J.; Goldberg, I. Cloning, expression, and characterization of a synthetic analog to the bioadhesive precursor protein of the sea mussel *Mytilus edulis*. *Appl. Microbiol. Biotechnol.* **1993**, *39* (2), 221–226.
103. Kitamura, M.; Kawakami, K.; Nakamura, N.; Tsutomoto, K.; Uchiyama, H.; Ueda, Y.; Kumagai, I.; Nakaya, T. Expression of a model peptide of a marine mussel adhesive protein in *Escherichia coli* and characterization of its structural and functional properties. *J. Polym. Sci. A.* **1999**, *37* (6), 729–736.
104. McQueen Mason, S.; Filatov, V. Transgenic Cells Expressing Adhesive Proteins. International Patent WO 03/082912, Mar 24, 2003.
105. Kieliszewski, M.J.; Lamport, G.T.A. Extensin—repetitive motifs, functional sites, posttranslational codes, and phylogeny. *Plant J.* **1994**, *5* (2), 157–172.
106. Tatehata, H.; Mochizuki, A.; Kawashima, T.; Yamashita, S.; Yamamoto, H. Model polypeptide of mussel adhesive protein. I. Synthesis and adhesive studies of sequential polypeptides (X-Tyr-Lys)_n and (Y-Lys)_n. *J. Appl. Polym. Sci.* **2000**, *76* (6), 929–937.
107. Tatehata, H.; Mochizuki, A.; Ohkawa, K.; Yamada, M.; Yamamoto, H. Tissue adhesive using synthetic model adhesive proteins inspired by the marine mussel. *J. Adhes. Sci. Technol.* **2001**, *15* (9), 1003–1013.
108. Huang, K.; Lee, B.P.; Ingram, D.R.; Messersmith, P.B. Synthesis and characterization of self-assembling block copolymers containing bioadhesive end groups. *Biomacromolecules* **2002**, *3* (2), 397–406.
109. Lee, B.P.; Huang, K.; Nunalee, F.N.; Shull, K.R.; Messersmith, P.B. Synthesis of 3,4-dihydroxyphenylalanine (Dopa) containing monomers and their co-polymerization with PEG-diacrylate to form

- hydrogels. *J. Biomater. Sci. Polym. Ed.* **2004**, *15* (4), 449–464.
110. Zanchetta, P.; Lagarde, N.; Guezennec, J. A new bone-healing material: a hyaluronic acid-like bacterial exopolysaccharide. *Calcif. Tissue Int.* **2003**, *72* (1), 74–79.
111. Guezennec, J.; Zanchetta, P.; Durand, P. Use of a Polysaccharide Excreted by the *Vibrio diabolus* Species in Bone Repair. International Patent WO 02/02051, Jul 4, 2001.
112. Haber, M. Plant Proteins. U.S. Patent 5,859,198, Aug 31, 1995.
113. <http://algal-adhesives.ocean.org.il/index.html> (accessed Mar 2005).
114. <http://www.algawish.com/index.html> (accessed Mar 2005).
115. Szomor, Z.L.; Appleyard, R.; Tyler, M.J.; Murrell, G.A.C. Meniscal repair with a frog glue. In Abstracts, Pre-Olympic Congress on Sports Medicine and Physical Education and International Congress on Sport Science, Brisbane, Australia, Sep 7–13, 2000; National Sport Information Centre: Canberra, Australia, 2000; 297. <http://www.ausport.gov.au/fulltext/2000/preoly/abs297.htm> (accessed Nov 2004).
116. Szomor, Z.L.; Appleyard, R.; Tyler, M.J.; Murrell, G.A.C. Meniscal repair with a new biologic glue. Abstracts, 47th Annual Meeting, Orthopaedic Research Society, San Francisco, CA, Feb 25–28, 2001; Orthopaedic Research Society: Illinois, 2001; Session 8, 0047.
117. Crisp, D.J.; Walker, G.; Young, G.A.; Yule, A.B. Adhesion and substrate choice in mussels and barnacles. *J. Colloid Interface Sci.* **1985**, *104* (1), 40–50.
118. Yamamoto, H. Synthesis and adhesive studies of marine polypeptides. *J. Chem. Soc. Perkin Trans.* **1987**, *1* (3), 613–618.
119. Yule, A.B.; Walker, G. Adhesion in barnacles. In *Crustacean Issues, Biology of Barnacles*; Southward, A.J., Ed.; Balkema: Rotterdam, 1987; 389–402.
120. Flammang, P.; Ribesse, J.; Jangoux, M. Biomechanics of adhesion in sea cucumber Cuvierian tubules (Echinodermata, Holothuroidea). *Integr. Comp. Biol.* **2002**, *42* (6), 1107–1115.
121. Flammang, P.; Walker, G. Measurement of the adhesion of the podia in the asteroid *Asterias rubens* (Echinodermata). *J. Mar. Biol. Assoc. U.K.* **1997**, *77* (4), 1251–1254.
122. Grenon, J-F.; Walker, G. The tenacity of the limpet *Patella vulgata* L.: an experimental approach. *J. Exp. Mar. Biol. Ecol.* **1981**, *54*, 277–308.
123. Roscoe, D.T.; Walker, G. Observations on the adhesion of the calcareous tube of *Pomatoceros lamarckii* and the holdfast of *Laminaria digitata*. *Biofouling* **1995**, *9* (1), 39–50.
124. Berglin, M.; Gatenholm, P. The nature of bioadhesive bonding between barnacles and fouling-release silicone coatings. *J. Adhes. Sci. Technol.* **1999**, *13* (6), 713–727.
125. McDermott, M.K.; Chen, T.; Williams, C.M.; Markley, K.M.; Payne, G.F. Mechanical properties of biomimetic tissue adhesive based on the microbial transglutaminase-catalyzed crosslinking of gelatin. *Bio-macromolecules* **2004**, *5* (4), 1270–1279.
126. Páez, J.M.G.; Herrero, E.J.; Rocha, A.; Maestro, M.; Castillo-Olivares, J.L.; Millan, I.; Sanmartin, A.C.; Cordon, A. Comparative study of the mechanical behaviour of a cyanoacrylate and a bioadhesive. *J. Mater. Sci. Mater. Med.* **2004**, *15* (2), 109–115.
127. Edlund, A.F.; Koehl, M.A.R. Adhesion and reattachment of compound ascidians to various substrata: weak glue can prevent tissue damage. *J. Exp. Biol.* **1998**, *201* (16), 2397–2402.

Biological Effects of Radiofrequency Electromagnetic Field

Henry Lai

Department of Bioengineering, University of Washington, Seattle, Washington, U.S.A.

INTRODUCTION

Radiofrequency electromagnetic field/radiation (RFR) covers a large segment of the electromagnetic spectrum and falls within the nonionizing bands. Its frequency ranges between 3 KHz and 300 GHz. Cellular phones, wireless transmission towers for radio, TV, and telecommunication, radar and many other applications emit RFR. Different frequencies of RFR are used in different applications. For example, the frequency range of 540–1600 KHz is used in AM radio transmission, while 76–108 MHz is used for FM radio. Cell phone technology uses frequencies mainly between 800 MHz and 3 GHz, and RFR of 2450 MHz is used in microwave cooking.

Because of the proliferation of wireless communication in recent years, a large population of people is exposed to RFR constantly. There are two major concerns on the possible biological/health effects of RFR: the effect of RFR absorbed during cell phone use and the exposure to RFR emitted from transmission towers. These two situations represent very different exposure conditions. The close proximity of a cell telephone antenna to the user's head leads to the deposition of a relatively large amount of radiofrequency energy in the head. The relatively fixed position of the antenna relative to the head causes a repeated irradiation of a more or less fixed amount of body tissue. Exposure to RFR from cell phones is of a short-term, repeated nature at a relatively high intensity, whereas exposure to RFR emitted from transmission towers is of long duration but at a very low intensity and, in general, the whole body of a person is exposed.

Biological effects of RFR depend on how energy is deposited in the exposed organism. There are three major physical parameters: frequency, intensity, and duration of exposure. To understand the possible health effects of exposure to RFR, one needs first to understand the effects of these different parameters and how they interact with each other. These are discussed in this entry.

FREQUENCY, INTENSITY, AND PATTERN OF ENERGY ABSORPTION

The frequency of RFR is analogous to the color of a light bulb, and intensity is its wattage. There is a

question of whether the effects of RFR of one frequency are different from those of another frequency. In this case, one is basically asking the question, "Are the effects of red light different from those of green light?" The answer to this is that it depends on the situation. They are different: if one is looking at a traffic light, "red" means "stop" and "green" means "go." But, if one is going to send some information by Morse Code using a light (on and off, etc), it will not matter whether one uses a red or green light, as long as the receiver can see and decode it. We do not know which of these two cases applies to the biological effects of RFR regarding frequency.

It must be pointed out that data are sparse showing either different frequencies producing different effects or an effect observed at one frequency but not at another. An example is the study by Sanders, Joines, and Allis, who observed that RFR at frequencies of 200 and 591 MHz, but not at 2450 MHz, produced effects on energy metabolism in neural tissues.^[1] There are also several studies that showed that different frequencies of RFR produced different effects.^[2–5] However, it is not certain whether these differences were actually due to differences in the pattern of energy absorption in the body of the exposed animal at the various frequencies. In addition, some studies showed frequency-window effects, i.e., effect is only observed at a certain range of frequencies and not at higher or lower ranges.^[6–17] These results may suggest that the frequency of an RFR can be a factor in determining the biological outcome of exposure.

On the other hand, there are more studies showing that different frequencies can produce the same effect. For example, changes in blood-brain barrier have been reported after exposure to RFRs of 915, 1200, 1300, 2450, and 2800 MHz, and effects on calcium metabolism have been reported at 50, 147, 450, and 915 MHz.^[6,8,9,12,13,16–20] If there is any difference in effects among different frequencies, it is a difference in quantity and not in quality.

The intensity of RFR in the environment is the power density measured in units such as milliwatts per square centimeter. However, power density provides little information on the biological consequence unless the amount of energy absorbed by the irradiated object is known. This is generally given as the specific

absorption rate (SAR), which is the rate of energy absorbed by a unit mass of tissue of the object, and usually expressed as watts per kilogram. Thus, to understand the biological effect of RFR, one should know the SAR. It is a more reliable determinant and index of RFR biological effects than power density. Specific absorption rate is used in the setting of exposure standards for RFR.

Biological effects can occur after exposure to high intensity of RFR (high SAR) that can cause general or local heating. In some RFR exposure guidelines, the limits of 0.4 W/kg for occupational exposure and 0.08 W/kg for general public exposure are used based on the experimental results that “disruption of behavior” in animals occurs at 4 W/kg. However, there are many studies that show biological effects at SARs less than 4 W/kg after short-term exposure to RFR. For example, behavioral effects have been observed at SARs less than 4 W/kg: D’Andrea et al., 0.14–0.7 W/kg; DeWitt et al., 0.14 W/kg; Gage, 3 W/kg; King, Justesen, and Clarke, 2.4 W/kg; Lai et al., 0.6 W/kg; Mitchell, Switzer, and Bronaugh, 2.3 W/kg; Navakatikian and Tomashkevskaya, 0.027 W/kg; Schrot, Thomas, and Banvard, 0.7 W/kg; Thomas et al., 1.5–2.7 W/kg; Wang and Lai, 1.2 W/kg.^[21–31] There are also many reports of other biological functions affected by RFR at a SAR less than 4 W/kg. Most of the energy from a cellular telephone antenna is deposited in the skin and the outer portion of the brain. The peak energy output of cell telephones can range from 0.3 to 1 W, although the average output could be much smaller. Relatively high SARs have been determined in various dosimetry studies: Dimbylow and Mann, 2.3 and 4.8 W/kg per watt output at 900 MHz and 1.8 GHz; Anderson and Joyner, 0.12–0.83 W/kg; Gandhi et al., 0.13–5.41 W/kg at 0.6 W output (835 and 1900 MHz); and Van de Kamer and Lagendijk, 1.72–2.55 W/kg at 0.25 W output (915 MHz).^[32–35]

Surprisingly, effects have also been reported in cells and animals after exposure to very low-intensity RFR that apparently cannot cause a physiologically significant change in temperature. Some 40 studies can be listed as low-intensity effects. The following are examples of some of these studies (some studies only give the power density, in mW/cm², of the radiation, whereas others give the SAR, in W/kg, in the exposed objects). Lebedeva et al. showed changes in brain wave activation in human subjects exposed to cellular phone RFR at 0.06 mW/cm².^[36] Magras and Xenos reported a decrease in reproductive function in mice exposed to RFR at power densities of 0.000168–0.001053 mW/cm².^[37] Phillips et al. reported DNA damage in cells exposed to RFR at SAR of 0.0024–0.024 W/kg.^[38] Salford et al. reported nerve cell damage in brain of rats exposed for 2 hr to cell phone signal at 0.02 W/kg.^[39] Tattersall

et al. showed that low-intensity RFR (0.0016–0.0044 W/kg) modulated the function of a part of the brain called the hippocampus, in the absence of gross thermal effects.^[40]

In addition, there are some indications that biological effects may also depend on how energy is deposited in the body. The rate of absorption and the distribution of RFR energy in an organism depend on many factors. These include the dielectric composition of the irradiated tissue, e.g., bones with a lower water content absorb less of the energy than muscles; the size of the object relative to the wavelength of the RFR (thus, the frequency); the shape, geometry, and orientation of the object; and the configuration of the radiation, e.g., how close is the object from the RFR source? These factors make the distribution of energy absorbed in an irradiated organism extremely complex and nonuniform, and also lead to the formation of so called “hot spots” of concentrated energy in the tissue. For example, an experiment reported by Chou et al., measuring local energy absorption rates (SARs) in different areas of the brain in a rat exposed to RFR, has shown that two brain regions less than a millimeter apart can have more than a twofold difference in SAR.^[41] The rat was stationary when it was exposed. The situation is more complicated if an animal is moving in an RF field. Depending on the amount of movement of the animal, the energy absorption pattern in its body could become either more complex and unpredictable or more uniform. Thus, the pattern of energy absorption inside an irradiated body is nonuniform, and biological responses are dependent on distribution of energy and the body part that is affected.^[42,43] Related to this is that we have found that different areas of the brain of the rat have different sensitivities to RFR.^[44] This further indicates that the pattern of energy absorption could be an important determining factor of the nature of the response.

Different propagation characteristics such as “modulation,” or different waveforms and shapes may have different effects on a living system. For example, the same amount of energy can be delivered to tissue “continuously” or “in short pulses.”

Another interesting observation of the research is that modulated and pulsed RFR seem to be more effective in producing an effect. They can also elicit a different effect when compared with continuous-wave radiation of the same frequency.^[16,19,43,45–47] This observation is important because cell phone radiation is modulated at low frequencies. This also raises the question of how much do low-frequency electric and magnetic field components contribute to the biological effects. Biological effects of extremely low-frequency (<100 Hz) electric and magnetic fields are quite well established.

REPEATED EXPOSURE AND DURATION OF EXPOSURE

The majority of the biological studies on RFR have been conducted with short-term exposures, i.e., a few minutes to several hours. Little is known about the effects of long-term exposure. However, in actual human exposure situations, RFR exposure tends to be repetitive and long term. What are the effects of long-term exposure? Does long-term exposure produce different effects from short-term exposure? Do effects accumulate over time?

An important question regarding the biological effects of RFR is whether the effects are cumulative, i.e., after repeated exposure, will a biological system adapt to the perturbation and, with continued exposure, when will homeostasis break down leading to irreparable damage? The question of whether an effect will cumulate over time with repeated exposure is particularly important in considering the possible health effects of cell phone usage, because it involves repeated exposure of short duration over a long period (years) of time. Existing results indicate changes in the response characteristics of a biological system with repeated exposure, suggesting that the effects are not “forgotten” after each episode of exposure. Various biological outcomes have been reported after long-term/repeated exposure to RFR:

1. Effects were observed after prolonged, repeated exposure but not after short-term exposure.^[48,49]
2. Effects that were observed after short-term exposure, disappeared after prolonged, repeated exposure (i.e., habituation) (e.g., Refs. [50–52]).
3. Different effects were observed after different durations of exposure (e.g., Refs. [26,53,54]).
4. There is also an indication that an animal becomes more sensitive to the radiation after long-term exposure (e.g., Refs. [4,21,22,23]). For example, the conclusion from a series of experiments on disruption of behavior in animals after one-time exposure to RFR is that “disruption of behavior occurred when an animal was exposed at a SAR of approximately 4 W/kg.”^[55] However, after long-term exposure (7 hr/day, 7 days/week for 90 days to 14 weeks), the threshold for behavioral and physiological effects of RFR was found to occur between 0.14 and 0.7 W/kg.^[21,22] Thus, RFR can produce an effect at much lower intensities after an animal is repeatedly exposed. This can have very significant implications for people exposed to RFR in the environment.

The conclusion from this body of research is that effects of long-term exposure can be quite different from those of short-term exposure. There is also some

evidence that effects of RFR accumulate over time. Here are some examples: Phillips et al. reported DNA damage in cells after 24 hr of exposure to low-intensity RFR. DNA damage can lead to gene mutation, which accumulates over time.^[38] Magras and Xenos reported that mice exposed to low-intensity RFR became less able to reproduce.^[37] After five generations of exposure, the mice were not able to produce offspring. This shows that the effect of RFR can pass from one generation to another. Persson, Salford, and Brun reported an increase in permeability of the blood–brain barrier in mice when the total energy deposited in the body exceeded 1.5 J/kg, i.e., the effect depends on the amount of energy deposited and not the rate of deposition.^[56] This suggests that, under similar exposure conditions, a short-term/high-intensity exposure can produce the same effect as a long-term/low-intensity exposure. This is an indication that RFR effects can accumulate over time.

Related to this is that various lines of evidence suggest that responses to RFR could be a stress response.^[57,58] Stress effects are well known to cumulate over time and involve first adaptation and then an eventual breakdown of homeostatic processes when the stress persists.

The possibility that effects cumulate over time and that acute effects change with repeated exposure have important implications on the setting of standards of RFR exposure. This suggests that the total amount of energy absorbed, the specific absorption (SA) ($SA = SAR \times \text{time}$) rather than SAR, should be used as the index.

THERMAL AND NONTHERMAL EFFECTS

When RFR energy is absorbed, it is converted into heat. A readily understandable mechanism of effect of RFR is tissue heating (thermal effect). Biological systems alter their functions as a result of change in temperature. However, there is also a question as to whether “nonthermal” effects can occur from RF exposure. There can be two meanings to the term nonthermal effect. It could mean that an effect occurs under the condition of no apparent change in temperature in the exposed animal or tissue, suggesting that physiological or exogenous mechanisms maintain the exposed object at a constant temperature. The second meaning is that somehow RFR can cause biological effects without the involvement of heat energy (i.e., temperature independent). For practical reasons, it may be futile to make these distinctions simply because it is very difficult to rule out thermal effects in biological responses to RFR, and because heat energy is inevitably dissipated when RFR is absorbed.

In some experiments, thermal controls (i.e., samples subjected to direct heating) have been studied. Indeed, there are reports showing that “heating controls” do not produce the same effect of RFR.^[59–62] These were taken as an indication of nonthermal effects. However, as we have discussed earlier, it is difficult to reproduce the same pattern of internal heating of RFR by external heating, as we know that a conventional oven cooks food differently from a microwave oven. And the pattern of energy distribution in the body is important in determining the effect of RFR (e.g., Refs. [19,42,43]). Thus, “heating controls do not produce the same effect of RFR,” does not really support the existence of nonthermal effects.

Furthermore, even though no apparent change in body temperature during RFR exposure occurs, it cannot really rule out a “thermal effect.” In one of our experiments, we have shown that animals exposed to a low SAR of 0.6 W/kg are actively dissipating the energy absorbed, thus, no significant increase in body temperature was observed in these animals.^[42] This suggests that the nervous system involved in body temperature regulation is activated. The physiology of body temperature regulation is complicated and can involve many organ systems. Thus, changes in thermoregulatory activity can indirectly affect biological responses to RFR.

Another difficulty in eliminating the contribution of thermal effects is that it can be “microthermal.” An example of this is the auditory effect of pulsed RFR. We can hear RFR delivered in pulses. One of the explanations for this “hearing” effect is that it is caused by thermoelastic expansion of the head of the “listener.” In a paper by Chou et al., it was stated that “... one hears sound because a miniscule wave of pressure is set up within the head and is detected at the cochlea when the absorbed microwave pulse is converted to thermal energy.”^[63] The threshold of hearing was determined to be approximately 10 μ J/g/pulse, which causes an increment of temperature in the head of one millionth of a degree centigrade. Lebovitz gives another example of the microthermal effect of RFR on the vestibulocochlear apparatus, an organ in the inner ear responsible for keeping the body balance and sense of movement.^[64] He proposed that an uneven distribution of RFR absorption in the head can set up a temperature gradient in the semicircular canals of the vestibulocochlear apparatus, which in turn affects the function of the vestibular system. The semicircular canals are very minute organs in our body.

What about in vitro experiments in which isolated organs or cells are exposed to RFR? Generally, these experiments are conducted with the temperature controlled by various regulatory mechanisms. However, it turns out that the energy distribution in culture disks, test tubes, and flasks used in these studies are often

very uneven.^[65] Hotspots are formed. There is a question of whether the temperature within the exposed samples can be efficiently controlled.

However, the existence of intensity windows, reports of modulated fields producing stronger or different effects than continuous-wave fields, and the presence of effects when exposed to RFR at very low intensity described in the sections above could be indications of nonthermal effects. My argument is that, in the setting of exposure standards, it may not be practical to differentiate these effects owing to the difficulty of eliminating the thermal effects.

OTHER CONSIDERATIONS

It has been repeatedly pointed out that the results reported in RFR research are difficult to replicate or reproduce. Problems in data reproducibility are not uncommon in science. In RFR research, it is further complicated by the complexity of interaction among the various exposure parameters as discussed above. This may make the response more sensitive to experimental procedures. Moderate variations in procedures could lead to different results. Examples are recent attempts to study RFR-induced DNA damage and spatial learning deficit.^[38,66–71] Different results were observed when different experimental procedures were used. However, by comparing them carefully, these differences in results may actually help to reveal the mechanisms of interaction between RFR and biological systems.

An area of research that requires more study is the role played by the physiological conditions of an organism on its response to RFR. For example, The British Stewart Report on “Mobile Phone and Health” recommends caution in the use of cellular phones by young children because of their “developing nervous system.”^[72] Yet, little is known on whether developing biological systems are actually more vulnerable to the effects of RFR. On the other hand, people under certain drug therapies may be more susceptible to RFR. This is suggested by a study by Kues et al. in the early 1990s showing that the ophthalmic drug timolol significantly enhanced corneal endothelial lesion induced by RFR in the monkey.^[73] Little research has been carried out to further investigate this type of interactions.

Another consideration is that the genetics of an organism may affect its response to electromagnetic fields. For example, two groups of researchers reported different effects of magnetic field on 7,12-dimethylbenz[α]anthracene (DMBA)-induced breast tumors in Sprague–Dawley rats.^[74] One team, led by Wolfgang Löscher, later found that two substrains of Sprague–Dawley rats responded differently to the carcinogen

and the magnetic field and that this could account for the different results found by the two research groups.^[75]

CONCLUSIONS

1. It is quite certain that RFR exposure can cause biological effects even at low intensity. However, the potential hazardous health effects of such exposure to humans are not clear. We do not know if these effects occur in humans exposed to RFR, or whether the reported effects are health hazards. Biological effects do not automatically mean adverse health effects. Many biological effects are reversible. However, it is very clear that low-intensity RFR is not biologically inert. Much more needs to be learned, however, before a presumption of safety can be made.
2. Frequency, intensity, exposure duration, and the number of exposure episodes can affect the response to RFR, and these factors can interact with each other and produce different effects. To understand the biological consequence of RFR exposure, one must know whether the effect is cumulative, whether compensatory responses result, and when homeostasis will break down. The response is not likely to be linear with respect to the intensity of the radiation. Other parameters of RFR exposure, such as waveform, frequency and amplitude modulation, etc., are also important determinants of biological responses and affect the shape of the dose–response relationship.
3. Not much is known about the biological effects of long-term exposure. The effects of long-term exposure can be quite different from those of short-term exposure. The effects may accumulate. In that case, the total energy absorbed (i.e., specific absorption) is a more relevant index of biological effect than the rate of energy absorption (SAR).
4. In many RFR exposure guidelines, a limit of 0.4 W/kg is used based on the experimental results that disruption of behavior in animals occurs at 4 W/kg. However, there are many studies that show biological effects at SARs less than 4 W/kg after short-term exposure to RFR. Therefore, the rationale of 4 W/kg should be reconsidered. The guidelines also only consider short-term exposure effect. Effects of long-term exposure, modulation, and other propagation characteristics are not considered. Another omission is that the physiological conditions of the exposed organism are not taken into consideration. Therefore, the present guidelines are questionable in protecting the public from possible harmful effects of RFR exposure.

5. Owing to the uncertainty in science, exposure of the general population to RFR should be kept to a minimum and should follow the ALAR-principle (As Low as Reasonably Achievable).

ARTICLE OF FURTHER INTEREST

Biocompatibility Testing

REFERENCES

1. Sanders, A.P.; Joines, W.T.; Allis, J.W. The differential effect of 200, 591, and 2450 MHz radiation on rat brain energy metabolism. *Bioelectromagnetics* **1984**, *5*, 419–433.
2. D'Andrea, J.A.; Gandhi, O.P.; Lords, J.L.; Durney, C.H.; Johnson, C.C.; Astle, L. Physiological and behavioral effects of chronic exposure to 2450-MHz microwaves. *J. Microw. Power* **1979**, *14*, 351–362.
3. D'Andrea, J.A.; Gandhi, O.P.; Lords, J.L.; Durney, C.H.; Astle, L.; Stensaas, L.J.; Schoenberg, A.A. Physiological and behavioral effects of prolonged exposure to 915 MHz microwaves. *J. Microw. Power* **1980**, *15*, 123–135.
4. de Lorge, J.; Ezell, C.S. Observing-responses of rats exposed to 1.28- and 5.62-GHz microwaves. *Bioelectromagnetics* **1980**, *1*, 183–198.
5. Thomas, J.R.; Finch, E.D.; Fulk, D.W.; Burch, L.S. Effects of low level microwave radiation on behavioral baselines. *Ann. N.Y. Acad. Sci.* **1975**, *247*, 425–432.
6. Bawin, S.M.; Kaczmarek, L.K.; Adey, W.R. Effects of modulated VHF fields on the central nervous system. *Ann. N.Y. Acad. Sci.* **1975**, *247*, 74–81.
7. Blackman, C.F.; Elder, J.A.; Weil, C.M.; Benane, S.G.; Eichinger, D.C.; House, D.E. Induction of calcium-ion efflux from brain tissue by radiofrequency radiation: effects of modulation frequency and field strength. *Radio Sci.* **1979**, *14*, 93–98.
8. Blackman, C.F.; Benane, S.G.; Elder, J.A.; House, D.E.; Lampe, J.A.; Faulk, J.M. Induction of calcium ion efflux from brain tissue by radiofrequency radiation: effect of sample number and modulation frequency on the power-density window. *Bioelectromagnetics* **1980**, *1*, 35–43.
9. Blackman, C.F.; Benane, S.G.; Joines, W.T.; Hollis, M.A.; House, D.E. Calcium ion efflux from brain tissue: power density versus internal field-intensity dependencies at 50-MHz RF radiation. *Bioelectromagnetics* **1980**, *1*, 277–283.
10. Blackman, C.F.; Kinney, L.S.; House, D.E.; Joines, W.T. Multiple power density windows and their possible origin. *Bioelectromagnetics* **1989**, *10*, 115–128.
11. Chang, B.K.; Huang, A.T.; Joines, W.T.; Kramer, R.S. The effect of microwave radiation (1.0 GHz) on the blood-brain-barrier. *Radio Sci.* **1982**, *17*, 165–168.
12. Dutta, S.K.; Subramoniam, A.; Ghosh, B.; Parshad, R. Microwave radiation-induced calcium ion efflux

- from human neuroblastoma cells in culture. *Bioelectromagnetics* **1984**, *5*, 71–78.
13. Dutta, S.K.; Ghosh, B.; Blackman, C.F. Radiofrequency radiation-induced calcium ion efflux enhancement from human and other neuroblastoma cells in culture. *Bioelectromagnetics* **1989**, *10*, 197–202.
 14. Dutta, S.K.; Das, K.; Ghosh, B.; Blackman, C.F. Dose dependence of acetylcholinesterase activity in neuroblastoma cells exposed to modulated radiofrequency electromagnetic radiation. *Bioelectromagnetics* **1992**, *13*, 317–322.
 15. Lin-Liu, S.; Adey, W.R. Low frequency amplitude modulated microwave fields change calcium efflux rate from synaptosomes. *Bioelectromagnetics* **1982**, *3*, 309–322.
 16. Oscar, K.J.; Hawkins, T.D. Microwave alteration of the blood-brain-barrier system of rats. *Brain Res.* **1977**, *126*, 281–293.
 17. Sheppard, A.R.; Bawin, S.M.; Adey, W.R. Models of long-range order in cerebral macro-molecules: effect of sub-ELF and of modulated VHF and UHF fields. *Radio Sci.* **1979**, *14*, 141–145.
 18. Salford, L.G.; Brun, A.; Stureson, K.; Eberhardt, J.L.; Persson, B.R. Permeability of the blood-brain barrier by 915 MHz electromagnetic radiation, continuous wave and modulated at 8, 16, 50, and 200 Hz. *Microsc. Res. Tech.* **1994**, *27*, 535–542.
 19. Frey, A.H.; Feld, S.R.; Frey, B. Neural function and behavior: defining the relationship. *Ann. N.Y. Acad. Sci.* **1975**, *247*, 433–439.
 20. Albert, E.N. Light and electron microscopic observations on the blood-brain-barrier after microwave irradiation. In *Symposium on Biological Effects and Measurement of Radio Frequency Microwaves*; Hazzard, D.G., Ed.; HEW Publication (FDA): Rockville, MD, 1977.
 21. D'Andrea, J.A.; DeWitt, J.R.; Emmerson, R.Y.; Bailey, C.; Stensaas, S.; Gandhi, O.P. Intermittent exposure of rat to 2450-MHz microwaves at 2.5 mW/cm²: behavioral and physiological effects. *Bioelectromagnetics* **1986**, *7*, 315–328.
 22. D'Andrea, J.A.; DeWitt, J.R.; Gandhi, O.P.; Stensaas, S.; Lords, J.L.; Nielson, H.C. Behavioral and physiological effects of chronic 2450-MHz microwave irradiation of the rat at 0.5 mW/cm². *Bioelectromagnetics* **1986**, *7*, 45–56.
 23. DeWitt, J.R.; D'Andrea, J.A.; Emmerson, R.Y.; Gandhi, O.P. Behavioral effects of chronic exposure to 0.5 mW/cm² of 2450-MHz microwaves. *Bioelectromagnetics* **1987**, *8*, 149–157.
 24. Gage, M.I. Behavior in rats after exposure to various power densities of 2450 MHz microwaves. *Neurobehav. Toxicol.* **1979**, *1*, 137–143.
 25. King, N.W.; Justesen, D.R.; Clarke, R.L. Behavioral sensitivity to microwave irradiation. *Science* **1971**, *172*, 398–401.
 26. Lai, H.; Carino, M.A.; Horita, A.; Guy, A.W. Low-level microwave irradiation and central cholinergic systems. *Pharmacol. Biochem. Behav.* **1989**, *33*, 131–138.
 27. Mitchell, D.S.; Switzer, W.G.; Bronaugh, E.L. Hyperactivity and disruption of operant behavior in rats after multiple exposure to microwave radiation. *Radio Sci.* **1977**, *12*, 263–271.
 28. Navakatikian, M.A.; Tomashevskaya, L.A. Phasic behavioral and endocrine effects of microwaves of non-thermal intensity. In *Biological Effects of Electric and Magnetic Fields*; Carpenter, D.O., Ed.; Academic Press: San Diego, CA, 1994; Vol. 1, 333–342.
 29. Schrot, J.; Thomas, J.R.; Banvard, R.A. Modification of the repeated acquisition of response sequences in rats by low-level microwave exposure. *Bioelectromagnetics* **1980**, *1*, 89–99.
 30. Thomas, J.R.; Finch, E.D.; Fulk, D.W.; Burch, L.S. Effects of low level microwave radiation on behavioral baselines. *Ann. N.Y. Acad. Sci.* **1975**, *247*, 425–432.
 31. Wang, B.M.; Lai, H. Acute exposure to pulsed 2450-MHz microwaves affects water-maze performance of rats. *Bioelectromagnetics* **2000**, *21*, 52–56.
 32. Dimbylow, P.J.; Mann, J.M. SAR calculations in an anatomically realistic model of the head for mobile communication transceivers at 900 MHz and 1.8 GHz. *Phys. Med. Biol.* **1994**, *39*, 1527–1553.
 33. Anderson, V.; Joyner, K.H. Specific absorption rate levels measured in a phantom head exposed to radio frequency transmissions from analog hand-held mobile phones. *Bioelectromagnetics* **1995**, *16*, 60–69.
 34. Gandhi, O.P.; Lazzi, G.; Tinniswood, A.; Yu, Q.S. Comparison of numerical and experimental methods for determination of SAR and radiation patterns of handheld wireless telephones. *Bioelectromagnetics* **1999**, *4* (suppl.), 93–101.
 35. Van de Kamer, J.B.; Legendijk, J.J.W. Computation of high-resolution SAR distributions in a head due to a radiating dipole antenna representing a hand-held mobile phone. *Phys. Med. Biol.* **2002**, *47*, 1827–1835.
 36. Lebedeva, N.N.; Sulimov, A.V.; Sulimova, O.P.; Kotrovskaya, T.I.; Gailus, T. Cellular phone electromagnetic field effects on bioelectric activity of human brain. *Crit. Rev. Biomed. Eng.* **2000**, *28*, 323–337.
 37. Magras, I.N.; Xenos, T.D. RF radiation-induced changes in the prenatal development of mice. *Bioelectromagnetics* **1997**, *18*, 455–461.
 38. Phillips, J.L.; Ivaschuk, O.; Ishida-Jones, T.; Jones, R.A.; Campbell-Beachler, M.; Haggren, W. DNA damage in Molt-4 T-lymphoblastoid cells exposed to cellular telephone radiofrequency fields in vitro. *Bioelectrochem. Bioenerg.* **1998**, *45*, 103–110.
 39. Salford, L.G.; Brun, A.R.; Eberhardt, J.L.; Malmgren, L.; Persson, B.R.R. Nerve cell damage in mammalian brain after exposure to microwaves from GSM mobile phones. *Environ. Health Perspect.* **2003**, *111*, 881–883.
 40. Tattersall, J.E.; Scott, I.R.; Wood, S.J.; Nettell, J.J.; Bevir, M.K.; Wang, Z.; Somasiri, N.P.; Chen, X. Effects of low intensity radiofrequency electromagnetic fields on electrical activity in rat hippocampal slices. *Brain Res.* **2001**, *904*, 43–53.
 41. Chou, C.K.; Guy, A.W.; McDougall, J.; Lai, H. Specific absorption rate in rats exposed to 2450-MHz microwaves under seven exposure conditions. *Bioelectromagnetics* **1985**, *6*, 73–88.
 42. Lai, H.; Horita, A.; Chou, C.K.; Guy, A.W. Acute low-level microwave irradiation and the actions

- of pentobarbital: effects of exposure orientation. *Bioelectromagnetics* **1984**, *5*, 203–212.
43. Lai, H.; Horita, A.; Guy, A.W. Acute low-level microwave exposure and central cholinergic activity: studies on irradiation parameters. *Bioelectromagnetics* **1988**, *9*, 355–362.
 44. Lai, H.; Carino, M.A.; Horita, A.; Guy, A.W. Acute low-level microwave exposure and central cholinergic activity: a dose-response study. *Bioelectromagnetics* **1989**, *10*, 203–209.
 45. Arber, S.L.; Lin, J.C. Microwave-induced changes in nerve cells: effects of modulation and temperature. *Bioelectromagnetics* **1985**, *6*, 257–270.
 46. Frey, A.H.; Feld, S.R. Avoidance by rats of illumination with low power nonionizing electromagnetic radiation. *J. Comp. Physiol. Psychol.* **1975**, *89*, 183–188.
 47. Sanders, A.P.; Joines, W.T.; Allis, J.W. Effect of continuous-wave, pulsed, and sinusoidal-amplitude-modulated microwaves on brain energy metabolism. *Bioelectromagnetics* **1985**, *6*, 89–97.
 48. Baranski, S. Histological and histochemical effects of microwave irradiation on the central nervous system of rabbits and guinea pigs. *Am. J. Physiol. Med.* **1972**, *51*, 182–190.
 49. Takashima, S.; Onaral, B.; Schwan, H.P. Effects of modulated RF energy on the EEG of mammalian brain. *Radiat. Environ. Biophys.* **1979**, *16*, 15–27.
 50. Johnson, R.B.; Spackman, D.; Crowley, J.; Thompson, D.; Chou, C.K.; Kunz, L.L.; Guy, A.W. *Effects of Long-Term Low-Level Radiofrequency Radiation Exposure on Rats, Vol. 4. Open field Behavior and Corticosterone*; Brooks AFB: San Antonio, TX, 1983; USAF SAM-TR83-42; Report of USAF School of Aerospace Medicine.
 51. Lai, H.; Horita, A.; Chou, C.K.; Guy, A.W. Effects of low-level microwave irradiation on hippocampal and frontal cortical choline uptake are classically conditionable. *Pharmacol. Biochem. Behav.* **1987**, *27*, 635–639.
 52. Lai, H.; Carino, M.A.; Horita, A.; Guy, A.W. Single vs. repeated microwave exposure: effects on benzodiazepine receptors in the brain of the rat. *Bioelectromagnetics* **1992**, *13*, 57–66.
 53. Di Carlo, A.; White, N.; Guo, F.; Garrett, P.; Litovitz, T. Chronic electromagnetic field exposure decreases HSP70 levels and lowers cytoprotection. *J. Cell. Biochem.* **2002**, *84*, 447–454.
 54. Dumansky, J.D.; Shandala, M.G. The biologic action and hygienic significance of electromagnetic fields of super high and ultra high frequencies in densely populated areas, *Biologic Effects and Health Hazard of Microwave Radiation: Proceedings of an International Symposium*; Czerski, P., et al., Ed.; Polish Medical Publishers: Warsaw, 1974.
 55. de Lorge, J.O. Operant behavior and colonic temperature of *Macaca mulatta* exposed to radiofrequency fields at and above resonant frequencies. *Bioelectromagnetics* **1984**, *5*, 233–246.
 56. Persson, B.R.R.; Salford, L.G.; Brun, A. Blood-brain barrier permeability in rats exposed to electromagnetic fields used in wireless communication. *Wireless Network* **1997**, *3*, 455–461.
 57. Lai, H.; Carino, M.A.; Horita, A.; Guy, A.W. Single vs. repeated microwave exposure: effects on benzodiazepine receptors in the brain of the rat. *Bioelectromagnetics* **1992**, *13*, 57–66.
 58. Lai, H.; Horita, A.; Chou, C.K.; Guy, A.W. A review of microwave irradiation and actions of psychoactive drugs. *IEEE Eng. Med. Biol.* **1987**, *6*, 31–36.
 59. D’Inzeo, G.; Bernardi, P.; Eusebi, F.; Grassi, F.; Tamburello, C.; Zani, B.M. Microwave effects on acetylcholine-induced channels in cultured chick myotubes. *Bioelectromagnetics* **1988**, *9*, 363–372.
 60. Johnson, C.C.; Guy, A.W. Nonionizing electromagnetic wave effect in biological materials and systems. *Proc. IEEE* **1971**, *60*, 692–718.
 61. Seaman, R.L.; Wachtel, H. Slow and rapid responses to CW and pulsed microwave radiation by individual *Aplysia* pacemakers. *J. Microw. Power* **1978**, *13*, 77–86.
 62. Wachtel, H.; Seaman, R.; Joines, W. Effects of low-intensity microwaves on isolated neurons. *Ann. N.Y. Acad. Sci.* **1975**, *247*, 46–62.
 63. Chou, C.K.; Guy, A.W.; Galambos, R. Auditory perception of radiofrequency electromagnetic fields. *J. Acoust. Soc. Am.* **1982**, *71*, 1321–1334.
 64. Lebovitz, R.M. Detection of weak electromagnetic radiation by the mammalian vestibulocochlear apparatus. *Ann. N.Y. Acad. Sci.* **1975**, *247*, 182–193.
 65. Guy, A.W.; Chou, C.K.; McDougall, J.A. A quarter century of in vitro research: a new look at exposure methods. *Bioelectromagnetics* **1999**, (suppl. 4), 21–39.
 66. Diem, E.; Schwarz, C.; Adlkofer, F.; Jahn, O.; Rudiger, H. Non-thermal DNA breakage by mobile-phone radiation (1800 MHz) in human fibroblasts and in transformed GFSH-R17 rat granulosa cells in vitro. *Mutat. Res.* **2005**, *583*, 178–183.
 67. Lai, H.; Singh, N.P. Single- and double-strand DNA breaks in rat brain cells after acute exposure to radiofrequency electromagnetic radiation. *Int. J. Radiat. Biol.* **1996**, *69*, 513–521.
 68. Malyapa, R.S.; Ahern, E.W.; Straube, W.L.; Moros, E.G.; Pickard, W.F.; Roti Roti, J.L. Measurement of DNA damage after exposure to 2450 MHz electromagnetic radiation. *Radiat. Res.* **1997**, *148*, 608–617.
 69. Cassel, J.C.; Cosquer, B.; Galani, R.; Kuster, N. Whole-body exposure to 2.45 GHz electromagnetic fields does not alter radial-maze performance in rats. *Behav. Brain Res.* **2004**, *155*, 37–43.
 70. Cobb, B.L.; Jauchem, J.R.; Adair, E.R. Radial arm maze performance of rats following repeated low level microwave radiation exposure. *Bioelectromagnetics* **2004**, *25*, 49–57.
 71. Lai, H.; Horita, A.; Guy, A.W. Microwave irradiation affects radial-arm maze performance in the rat. *Bioelectromagnetics* **1994**, *15*, 95–104.
 72. National Radiological Protection Board (NRPB). Mobile phones and health 2004. *Doc. NRPB* **2004**, *15* (5), 1–114.

73. Kues, H.A.; Monahan, J.C.; D'Anna, S.A.; McLeod, D.S.; Luty, G.A.; Koslov, S. Increased sensitivity of the non-human primate eye to microwave radiation following ophthalmic drug pretreatment. *Bioelectromagnetics* **1992**, *13*, 379–393.
74. Anderson, L.E.; Morris, J.E.; Sasser, L.B.; Loscher, W. Effects of 50- or 60-hertz, 100 microT magnetic field exposure in the DMBA mammary cancer model in Sprague-Dawley rats: possible explanations for different results from two laboratories. *Environ. Health Perspect.* **2002**, *108*, 797–802.
75. Fedrowitz, M.; Kamino, K.; Loscher, W. Significant differences in the effects of magnetic field exposure on 7,12-dimethylbenz(a)anthracene-induced mammary carcinogenesis in two substrains of Sprague-Dawley rats. *Cancer Res.* **2004**, *64*, 243–251.

B

Biomaterials Immune Response

Sandra C.P. Cachinho

John A. Hunt

UKCTE, Clinical Engineering, UKBioTEC, School of Clinical Sciences,
University of Liverpool, United Kingdom

INTRODUCTION

Materials used in medical applications that necessarily come into intimate contact with the tissues of the body may initiate several and complex defense and healing biological reactions: cell and material contact evokes the release of chemotactic mediators and growth factors that may elicit and sustain inflammatory responses at the implant site.^[1,2] Because biocompatibility has been defined in terms of “the ability of a material to perform with an appropriate host response in a specific application,” it is important to determine the influence of these reactions on biocompatibility because this could influence the length of time in which a biomedical device can remain in the host, as well as the biofunctionality of such a device.^[1,3,4] Because all implanted materials are foreign to the host, the assessment of biocompatibility is an important step to prevent implant failure and includes the determination of the potential host responses to the material.^[4] The implantation of a material (without cells) initiates a sequence of events akin to a foreign body reaction, starting with an acute inflammatory response and leading, in some cases, to a chronic inflammatory response and granulation tissue development, a definitive foreign body reaction, and fibrous capsule formation. The duration and intensity of each of these is dependent upon the extent of injury created in the implantation, biomaterial chemical composition, surface free energy, surface charge, porosity, roughness, and implant size and shape. The extent and duration of the deviation from optimal wound healing conditions and homeostasis determine the biocompatibility of the material.^[5,6]

Vertebrates are provided with an immune system, composed of several lymphoid organs, highly specialized cells, and a circulatory system, that serves to protect the body from damage and disease. These defense mechanisms may be divided into two categories—acquired, also commonly referred as specific, and innate or non-specific.^[7,8] The processes in these categories are quite different, although they may be influenced and mediated by the same agents. Cells such as macrophages play an important role in both specific and non-specific responses.^[9]

The innate immune system is present in all individuals, requires little or no induction period, and does not respond more quickly or more vigorously upon the second encounter with a particular microbe (i.e., it does not exhibit a memory response). It is mediated mostly by neutrophils, eosinophils, basophils, macrophages, and natural killer cells (natural killer cells, also called large granular lymphocytes, are the major population of null lymphocytes, i.e., lymphocytes that are neither B cells nor T cells since they do not carry the markers to be B or T cells, and spontaneously attack and kill infected cells).^[9–11] These cells work by direct antimicrobial action, by killing virus-infected cells or tumor cells, and by secreting cytokines, which orchestrate inflammation and other defense mechanisms.^[9] The acquired immune response requires stimulation or immunization in order to be activated and is mediated by lymphocytes. Each lymphocyte has receptors on its surface that will recognize only a single conformation (antigenic determinant). Upon recognition of a specific antigen, a lymphocyte will proliferate, forming clones. Thus, upon reintroduction of the antigen, the host response is greater and accelerated.^[3] The immune system includes effector mechanisms that are capable of destroying a broad range of microbial cells and particles. A critical aspect of an effective immune response is the ability to avoid unleashing these destructive mechanisms against the host’s own tissues. Avoidance of destruction of self-tissues is referred to as self-tolerance and mechanisms to avoid reactions against self-antigens are expressed in both the innate and the adaptive immune responses.^[8]

MACROPHAGES

Macrophages are quintessentially the central cell in directing host inflammatory and immune processes. As such, determining the macrophage response to materials is extremely important in understanding the material-mediated host response. Several characteristic macrophage functions have been identified as critical events in the material host interaction.^[12] Macrophages are referred to as mononuclear phagocytes, reflecting their important ability to phagocytose

particulate material.^[10] They provide a first line of defense as part of the innate immune response against microorganisms and are therefore essential in the control of some bacterial infections. However, they cannot always eliminate infectious organisms, and there are some pathogens they cannot recognize. The lymphocytes of the adaptive immune system have evolved to provide a more versatile means of defense, which, in addition, provides increased protection against subsequent reinfection by the same pathogen.

Microorganisms such as bacteria that penetrate the epithelial surfaces of the body for the first time are met immediately by cells and molecules that can mount a vigorous, innate immune response. Phagocytic macrophages conduct the defense against bacteria by means of surface receptors that are able to recognize and bind common constituents of many bacterial surfaces. Bacterial molecules binding to these receptors trigger and activate the macrophage to engulf and surround the bacterium and also induce the secretion of biologically active molecules. Activated macrophages secrete cytokines, which are defined as proteins released by cells that affect the behavior of other cells that bear receptors for them. They also release proteins known as chemokines that mobilize cells with chemokine receptors such as neutrophils from the bloodstream along a chemotactic gradient toward a site of infection or injury.

Macrophages recognize pathogens by means of cell surface receptors that can discriminate between the surface molecules displayed by pathogens and those of the host. These receptors include the macrophage mannose receptor, scavenger receptors, which bind many charged ligands, and CD14, a receptor for bacterial lipopolysaccharide.

Upon phagocytosis, macrophages also produce a variety of other toxic products that destroy the engulfed microorganism. The most important of these are hydrogen peroxide (H_2O_2), the superoxide anion (O_2^-), and nitric oxide, which are directly toxic to bacteria.^[13]

In addition to phagocytosis, macrophages play a key role in acquired immunity acting as antigen-presenting cells (APCs). In the initial step of an acquired immune response, antigen processing and presentation by APCs is crucial.^[14] A phagocytic cell may translocate fragments of phagocytosed material to the cell surface in association with major histocompatibility complex (MHC) molecules and engage in binding to a T lymphocyte.^[15,16] It is important to recognize in this context that there are APCs with the capacity to initiate a primary immune response by the presentation of antigen to naive T lymphocytes and also APCs that will stimulate a secondary response by the presentation of antigen to T lymphocytes that have already been exposed to the antigen.^[16,17] Macrophages play an

essential role not only as APCs, but also as a source of cytokines that can manipulate the emerging T-lymphocyte response.^[15] Antigen processing is divided into two pathways that lead to peptide fragments of foreign antigens being presented by MHC class I or class II molecules:

The first pathway is the endogenous pathway: proteins that are derived within the cell (endogenously), such as those produced during viral or intracellular bacterial infections, result in the production of foreign proteins in the cytoplasm of the infected cell. Some of these proteins are cleaved into peptides in the cytosol by a large proteolytic enzyme complex called the proteasome.^[18] These peptides are translocated to the endoplasmic reticulum (ER) and bind to a protein complex called MHC class I. From the ER, the MHC class I–antigen complex is transferred to the cell surface, where it stimulates antigen-specific cytotoxic T lymphocytes (CTLs)—killer cells that seek out and eliminate infected cells.^[19]

The second pathway is the exogenous pathway: proteins that are derived from outside the cell (exogenously) are usually processed differently from endogenous antigens and produce different results.^[18,19] Extracellular pathogens are engulfed by APCs and degraded in phagosomes. This produces peptide fragments but they are not transported to the ER. Instead, they bind directly to a different set of MHC molecules, MHC class II, inside the phagosome and are then transported to the cell surface, where they stimulate a subset of helper T cells. Unlike the CTLs, helper T cells do not kill other cells, but instead produce chemical signals. These trigger the initiation of the humoral response and the production of antigen-specific antibodies by mature B cells (plasma cells), and they elicit an inflammatory response that eliminates the microbial pathogens.^[19]

The rules governing antigen presentation in principle are quite simple—endogenous antigens are presented by MHC class I molecules to stimulate a cell-mediated response, and exogenous antigens are presented by MHC class II molecules to stimulate a helper response. In reality, the situation is not so straightforward. Professional APCs can break these rules and present exogenous antigens on MHC class I molecules in a process known as cross-presentation. The cross-presentation pathway seems to be required for CTL responses to certain exogenous antigens. For example, some viruses do not infect professional APCs, but in order to generate a CTL response that will eliminate virus-infected cells, the APC must present exogenous viral proteins on MHC class I molecules. Cross-presentation might also be important for triggering CTL responses to some cancer cells.^[19]

How do exogenous antigens cross over to the MHC class I pathway? Rodriguez and colleagues^[20] indicated

that exogenous proteins could escape from phagosomes and enter the cytosol, suggesting that cross-presentation might simply involve the transport of exogenous proteins from the phagosome into the cytosol, where they could join the pathway used to process endogenous antigens. This view was supported by data demonstrating that cross-presentation depends on two components of the MHC class I pathway—the proteasome complex and the “transporter associated with antigen presentation” (TAP). The proteasome complex degrades proteins in the cytosol and TAP shuttles the resulting peptide antigens into the ER for loading onto MHC class I molecules.^[19]

LYMPHOCYTES

Lymphocytes are the cells responsible for the acquired immune response. T lymphocytes are responsible for the cell-mediated immune response and B cells are responsible for the humoral immune response, which is mediated by antibodies.^[21] T lymphocytes are derived from pluripotent stem cells in the bone marrow. They undergo differentiation in the thymus, which leads to the expression of their antigen receptor (i.e., T-lymphocyte antigen receptor or TCR).^[18] Each T lymphocyte has a unique receptor generated by recombination of the variable (V: V α and V β), diversity (D: D β), and joining (J: J α and J β) gene segments during thymic maturation.^[18,22,23] This random process generates many non-functional receptors. T lymphocytes bearing these non-functional receptors die by a process called positive selection. Large numbers of the T lymphocytes whose receptors are specific for normal proteins produced by host cells (self-proteins) are also generated. Self-reactive T lymphocytes are eliminated by a process called negative selection. This is one of the mechanisms by which the immune system is made tolerant to a body's own self-cells.^[18]

The immune system has evolved to protect the body from substances and organisms that threaten life.^[7,24] These may be present in extracellular body fluids (e.g., toxins, bacteria) or harbored in the individuals' own cells (e.g., viruses, oncogene products). This distinction is critically important.^[24] Immunity comprises cellular and extracellular components that reflect the diversity of the immune system. B cells produce immunoglobulins that function as extracellular, antigen-specific mediators in the recognition and clearance of foreign or domestic invaders.^[25] T cells carry membrane-associated TCRs that participate in the activation of T cells and recognize antigen that has been “processed” by APCs and presented in association with MHC glycoproteins.^[26–28]

Graft rejection is initiated by host T-cell recognition of alloantigen and activation of major T-cell

subpopulations, such as CD4⁺ and CD8⁺ T cells. Most CD4⁺ T cells function as helper T cells (Th), acting primarily to regulate the cellular and humoral immune response, and most CD8⁺ T cells are cytotoxic, acting primarily to kill cells infected with intracellular microbes.^[7,29,30] In the thymus developing T cells follow a developmental program in which, in the cortex, they at first express neither CD4 nor CD8 (double negative) and then express both CD4 and CD8 (double positive). Double-positive cells are tested by positive selection in the thymic cortex; those that are selected on class I MHC molecules become CD4⁻CD8⁺, and those that are selected on class II MHC molecules become CD4⁺CD8⁻. The cells then move to the thymic medulla for negative selection and export to the periphery. In the blood and secondary lymphoid organs, 60% to 70% of T cells are CD4⁺CD8⁻ (CD4⁺) and 30% to 40% are CD4⁻CD8⁺ (CD8⁺).^[7,31] The role of CD4⁺ and CD8⁺ T cells is well defined in early acute rejection. The chief contribution of CD8⁺ T cells to graft rejection is through direct lysis of target cells, whereas activated CD4⁺ cells contribute to graft rejection by releasing cytokines and recruiting a variety of other effector cells.^[29]

CD4 and CD8 are important in determining which MHC molecule is recognized by the TCR. CD4 binds to MHC class II molecules, helping the TCR to recognize antigens presented by MHC class II molecules. In contrast, CD8 binds to MHC class I molecules and CD8⁺ cells recognize antigens presented by MHC class I molecules.^[10]

CD4⁺ T LYMPHOCYTE SUBSETS

During immune responses, CD4⁺ T lymphocytes differentiate into different functional subsets of Th cells. Initially Th cells were subclassified into two subsets, Th1 and Th2, according to the cytokines secreted. For example, Th1 secrete interleukin 2 (IL-2), interferon- γ (IFN- γ), tumor necrosis factor- α (TNF- α) and TNF- β , whereas Th2 cells produce IL-4, IL-5, IL6, IL-10, and IL-13,^[32–34] among others. The balance between Th1 and Th2 cytokines is the key event that influences the maturation of naive CD4⁺ T cells to differentiate into Th1 and Th2 effectors and is important in the development of inflammatory diseases. Excessive Th1-type cytokines have been associated with the tissue destruction found in autoimmune diseases, whereas overabundant Th2-type cytokines have been implicated in atrophy and allergic asthma.^[33,35] Recently, a third subset was discovered, Th17, that has been associated with major human autoimmune diseases: rheumatoid arthritis, pancreatitis, inflammatory bowel disease, transplant rejection, allergic asthma, or

multiple sclerosis.^[36] The Th17 subset is characterized by the production of IL-17 cytokines, which is a set of six molecules named IL17A, IL17B, IL17C, IL17D, IL17E, and IL17F, which induce the production of proinflammatory cytokines such as TNF- α , IL-1 β , and IL-6 as well as chemokines CXCL1, -2, and -8.^[37-40] IL-17 is produced by activated T cells and mediates its proinflammatory effects via its receptor, IL-17R, which is expressed on all cell types. The IL-17 system appears to be a separate, unique, and potent signaling system involved in the control of the immune responses.^[41]

Different cytokines are involved in the development and proliferation of the Th17 subset from naive CD4⁺ T cells. High levels of transforming growth factor- β and IL-6 are crucial for the differentiation. IL-23, a cytokine produced by APCs, plays an important role in maintaining Th17 effector function.^[40,42]

MAJOR HISTOCOMPATIBILITY COMPLEX

The MHC is a genetic region that contains many genes involved in immune responses.^[43,44] These molecules carry the major histocompatibility antigens, which are tissue cell surface antigens capable of inducing an immune response in a genetically dissimilar (allogeneic) recipient, resulting in the rejection of the tissue or cells bearing those antigens.^[45,46] The immune response to a foreign protein is determined by the expression of specific MHC molecules, designated class I or class II, that bind and present peptide fragments of the protein to T cells. Without these specific recognition elements the immune system would be incapable of responding to a given antigen challenge.^[47-49] The cytosol is predominantly, but not exclusively the source of precursor proteins of peptides bound to MHC class I molecules, whereas peptides bound to MHC class II molecules are mainly derived from the extracellular compartment.^[50]

The complex is divided into three regions on the basis of tissue distribution, structure, and function—class I (HLA-A, -B, -C, -E, -F, and -G genes), class II (HLA-DR, -DP, and -DQ genes), and class III (includes genes encoding complement and TNF). These genes control immune responses to pathogens, graft rejection, and tumor surveillance. The functions of HLA-E, HLA-F, and HLA-G are not completely understood.^[46,51]

The constitutive expression of class I and II molecules is different. MHC class I molecules are cell-surface glycoproteins expressed on most cells of the body, although expression varies widely and is typically highest on hematopoietic cells.^[50,52] Class II molecules are generally restricted to a subset of APCs, such as macrophages, dendritic cells, and B cells. Expression can be induced on other cell types by stimulation with cytokines such as IFN- γ .^[49,53,54] Class I

molecules bind peptides derived from proteins synthesized endogenously by a cell and then display these peptides at the cell surface for recognition by CD8⁺ T cells. Class II molecules are responsible for presenting peptides derived from extracellular pathogens to T cells bearing the CD4⁺ marker.^[54-56]

Determination of the roles played by MHC antigens has focused on the role played by these proteins in the development of tumors, bacterial infections, articular diseases, and organ/tissue transplants. This research led to the development of treatments and vaccinations to combat many of these medical issues. Although there is still a seemingly exponential growth in the widespread use of materials as implanted devices, very little is understood about the immunological response of the body to materials. There is now evidence that some implanted materials may cause acquired immune reactions in a small percentage of the population.^[57] Implanted materials in general have been considered biologically inert; however, there is increasing evidence that no material implanted in living tissues is truly inert and will stimulate an immune response. The wear debris disseminated from these materials is not biologically inert^[58,59] as macrophages phagocytose debris particles. Much research has been focused on their role in the innate immune system; however, studies of periprosthetic tissue suggest that the acquired immune system may also be implicated by these cells and play an important role in the complications of the host response.^[63,64] As APCs, macrophages have the potential to present a biomaterial as an antigen. The material combines with self-proteins becoming antigenic and processed by macrophages for presentation to T cells in combination with MHC.^[62] Understanding the interactions between macrophages and T lymphocytes and the release of proinflammatory cytokines is critical in the development of strategies against current issues like the aseptic loosening of implants.

The immune system modulates many key biological processes in humans; however, the exact role played by the immune system in periprosthetic osteolysis is complex and not completely understood. Human tissue retrieval studies, and in vivo and in vitro experiments suggest that the immune response to polymeric particles is non-specific and macrophage mediated and direct lymphocyte activation by polymer particle protein complexes seems unlikely. On the other hand, metallic products may complex with serum proteins and lead to a T lymphocyte-mediated immune reaction. In predisposed individuals, this reaction may rarely lead to persistent painful joint effusions, necessitating the removal of the bearing surfaces of the prosthesis.^[61]

Hercus et al.^[63] observed T cell infiltration in retrieved periprosthetic tissues from patients with aseptic loosening of prosthetic joints. PCR tests did not show a predominance of Th1/Th2 cytokines. The major

cytokines expressed included IL-2, IL-4, IL-13, and IFN- γ .

Lalor et al.^[64] performed immunohistochemical studies of tissues from patients undergoing revision hip replacement from a previous failed cobalt chrome alloy or titanium alloy implant. All the tissues contained particulate metal debris and T lymphocytes in abundance with no evidence of B lymphocytes. Macrophages surrounding the T lymphocytes expressed HLA-DR and CD11c, indicating that they were antigen presenting cells. These findings suggested that in cases of excessive metal particle production, a T lymphocyte-mediated immune reaction may occur.

Wooley et al.^[65] studied the cellular responses in vitro to particles of polymethylmethacrylate, ultrahigh molecular weight polyethylene, and alloys of cobalt chromium and titanium from patients with underlying osteoarthritis, rheumatoid arthritis, and avascular necrosis. All patients were radiographically diagnosed as having aseptically loosened prostheses. The study confirmed the presence of a positive proliferative response to polymethylmethacrylate, a finding that was not in agreement with the findings by Santavirta et al.^[66] that reported polymethylmethacrylate failed to stimulate proliferation of peripheral blood mononuclear cells, although there was an increase in the expression of the two major markers of immune T cell activation: MCH-II and IL-2R on the cells surface. These different findings can be associated with particle size and culture conditions. The proliferative response to polyethylene was undetectable in vitro. Related to metallic particles neither alloys provoked a significant response.

In the study Arora et al.^[60] the specific roles of T lymphocyte subsets in osteolysis and aseptic loosening of total joint replacements was investigated. Tissue from periprosthetic regions from patients with loose, cemented acetabular components was used to determine the TH1 and TH2 cytokine profile. Immunohistochemistry revealed a significant number of T cells and TH1 (IL-2 and IFN- γ) and TH2 (IL-4 and IL-10) immune cytokines in the periprosthetic tissues. However, no significant differences in the percentage of positive cells expressing cytokines characteristic of the TH1 pathway or TH2 pathway were found when comparing the osteolytic and nonosteolytic groups. It has been suggested that the distinction between TH1 and TH2 cytokines in humans might be obscure.

The less intense reaction around metal-on-metal total hip replacement (THR) compared to metal-on-polyethylene has been reported and might be associated with the differences in the characteristics of metal wear particles. Doorn et al.^[67] characterized metal wear particles extracted from tissues from patients undergoing revision of metal-on-metal total hip replacements with cobalt-chromium-molybdenum bearing couples. It was demonstrated that metal

particles did not cause a more severe local reaction when compared with polymer particles, possibly due to the metal particles being smaller than the polyethylene particles and continuing to degrade further in soft tissues eventually.

Studies of the cement-bone interface of THRs have shown that a synovium-like membrane forms which has the capacity to stimulate the production of prostaglandin E2, collagenase and other osteolytic molecules such as interleukin-1 and tumor necrosis factor. Haddad et al.^[68] investigated a group of patients suffering from early loosening in whom polyethylene debris is unlikely to be responsible. The results indicated that hypersensitivity to DMT (N,N-dimethylparatoluidine, a constituent to bone cement) is the cause of rapid aseptic loosening.

Methe et al.^[69] examined how endothelial immunogenicity was controlled by the interaction with 3-D matrices. Human aortic endothelial cells (HAE) were grown to MHC class II, and integrin, IFN- γ receptor expression, and signaling were analyzed. The matrix architecture is critical for the modulation of endothelial immunogenicity. Embedding HAE within a physiologic 3-D environment affects the activity of intracellular signalling pathways, MHC II expression, and subsequent activation of immune cells. These findings offer insights into understanding endothelial-mediated diseases and may improve the potential for cell-based therapies.

Granchi et al.^[70] evaluated the sensitization status in patients with Co-Cr hip prosthesis in terms of the expression of the CD69 antigens in peripheral blood mononuclear cells collected from and cultured with the metal ions derived from the materials used to produce implanted medical devices. The activation antigens, HLA-DR, were expressed later and required a long incubation time. These data supported the hypothesis that an acquired immune response could contribute to the pathogenesis of hip prosthesis loosening.

Implant localized soft tissue from patients with aseptic loosening of cemented total joint arthroplasties contained more activated T lymphocytes than that of patients without aseptic loosening.^[71] The potential interactions of lymphocytes from peripheral blood and periprosthetic membranes with particulate orthopedic materials was examined in vitro, with the potential interactions between macrophages and lymphocytes in relation to IL-6 and TNF- α release examined further. The results demonstrated that the exposure of macrophages or macrophages cocultured with lymphocytes to particles of all materials tested led to a dose-dependent release of IL-6 and TNF- α . Baldwin et al.^[72] combined histological and molecular biological techniques to build a detailed picture of the cell types present in tissue surrounding failed Accord Knee prostheses and determined the occurrence of

inflammatory cells and cytokines, which may have contributed to implant failure. Tissue samples from patients undergoing revision surgery for aseptic loosening were analyzed to determine the presence and possible significance of inflammatory cells, cytokines, and chemokines present in these prosthesis failures. The research indicated that macrophages produced both the inflammatory cytokines and the chemokines, which may recruit further inflammatory cells to the implant site to perpetuate the cycle. The macrophages sequester a proportion of wear particles but inadequate removal or continued production of wear particles sustained the macrophage-driven response. The results from this immunohistochemical study confirmed the infiltration of macrophages, T cells, granulocytes, and neutrophils to the implant site.

Voggenreiter et al.^[73] analyzed the immunoinflammatory response to titanium and stainless-steel plates in the surrounding soft tissue. The results clearly demonstrated the presence of a marked inflammatory response and tissue reaction in the soft tissue covering stainless-steel and titanium plates used for the internal fixation of the fractures of long bones. Cells localized with metal particles expressed HLA-DR as an indicator of their activation and expression of HLA-DR was observed in the majority of cases, indicating that macrophages that phagocytosed steel or titanium particles were activated as APCs.

In vitro studies by Vallés et al.^[74] indicated that titanium particles in the phagocytosable range induced macrophages to secrete cytokines and initiate an inflammatory response. Rutile particles stimulated the release of TNF- α , IL-6, and IL-1 β to a lesser extent than titanium particles in these cellular models. Luyn et al.^[75] investigated whether repeated subcutaneous application of the same type of material could alter the foreign-body reaction. Two biodegradable (cross-linked collagens) and two non-biodegradable biomaterials (polyurethane and silicone) were tested in a repetitive implantation model. Enhancement of the foreign body reaction was observed by increased ingrowth of cells, increased vascularization, upregulation of MHC class II-positive cells, and the infiltration of plasma cells in the surrounding tissues. Except for the cellular ingrowth, changes occurred with both the biodegradable and the non-biodegradable materials, indicative that alterations were not specifically related to the biodegradation of biomaterials.

CONCLUSION

The immune mechanisms that are triggered when a material is introduced into a body play an important role in determining its biocompatibility. Immediately, cells of the immune system accumulate at the site of

implantation and a sequence of events targeted against the material can take place. Among all the cells of the immune system, macrophages and T lymphocytes are the most abundant or persistent in the material tissue interface or the surrounding tissue. The interaction between these cells results in the secretion of proinflammatory cytokines, soluble molecules involved in the initiation of the inflammatory process that may contribute to material rejection. The implantation of materials has been associated with innate immune inflammatory responses. However, experimental evidence indicates that materials could induce varied and significant expression of MHC molecules that are implicated in acquired immune responses. As cell-based therapies develop and become widespread, the intricacies of the detail of the acquired immunity must be understood. The conventional boundaries between innate and acquired immune responses will be breached, but should ultimately be predictable and therefore controllable. A further characterization of the immunoreactions to implant materials will permit the development of appropriate strategies that could expand the useful life of an implanted material.



ARTICLES OF FURTHER INTEREST

Biocompatibility Testing; Thrombosis

REFERENCES

1. Bosetti, M.; Ottani, V.; Kozel, D.; Raspanti, M.; Pasquale, V.D.; Ruggeri, A.; Cannas, M. Structural and functional macrophages alterations by ceramics of different composition. *Biomaterials* **1999**, *20* (4), 363–370.
2. Sharma, C.P. Biomaterials and artificial organs: Few challenging areas. *Trends Biomater. Artif. Organs* **2005**, *18* (2), 148–157.
3. Remes, A.; Williams, D.F. Immune response in biocompatibility. *Biomaterials* **1992**, *13* (11), 731–743.
4. Hunt, J.A.; Shoichet, M. Biomaterials: Surface interactions. *Curr. Opin. Solid State Mater. Sci.* **2001**, *5* (2–3), 161–162.
5. Babensee, J.E.; Anderson, J.M.; McIntire, L.V.; Mikos, A.G. Host response to tissue engineered devices. *Adv. Drug Deliv. Rev.* **1998**, *33* (1–2), 111–139.
6. Li, Y.; Schutte, R.J.; Abu-Shakra, A.; Reichert, W.M. Protein array method for assessing in vitro biomaterials-induced cytokine expression. *Biomaterials* **2005**, *26* (10), 1081–1085.
7. Bakker, T.C.M.; Zbinden, M. Counting on immunity. *Nature* **2001**, *414* (15), 262–263.
8. Chaplin, D.D. Overview of the immune response. *J. Allerg. Clin. Immunol.* **2003**, *111* (2), S442–S459.
9. Pruett, S.B. Stress and the immune system. *Pathophysiology* **2003**, *9* (3), 133–153.

10. Lee, K.A.; Klasing, K.C. A role for immunology in invasion biology. *Trends Ecol. Evol.* **2004**, *19* (10), 523–529.
11. Pier, G.B.; Lyczak, J.B.; Wetzler, L.M. *Immunology, Infection, and Immunity*; ASM Press: Washington, 2004.
12. Kao, W.J. Evaluation of protein-modulated macrophage behavior on biomaterials: Designing biomimetic materials for cellular engineering. *Biomaterials* **1999**, *20* (23–24), 2213–2221.
13. Parham, P. *The Immune System*; Garland Science: New York, USA, 2004.
14. Muno, D.; Kominami, E.; Mizuochi, T. Generation of both MHC class I- and class II-restricted antigenic peptides from exogenously added ovalbumin in murine phagosomes. *FEBS Lett.* **2000**, *478* (1–2), 178–182.
15. Doherty, T.M. T-cell regulation of macrophage function. *Curr. Opin. Immunol.* **1995**, *7* (3), 400–404.
16. Perry, V.H. A revised view of the central nervous system microenvironment and major histocompatibility complex class II antigen presentation. *J. Neuroimmunol.* **1998**, *90* (2), 113–121.
17. Boes, M.; Ploegh, H.L. Translating cell biology in vitro immunity in vivo. *Nature* **2004**, *430* (8), 264–271.
18. King, C.A.; Wills, M.R. Immunology II: Acquired immunity. *Surgery* **2005**, *23* (9), 319–323.
19. Roy, C.R. Professional secrets. *Nature* **2003**, *425* (25), 351–352.
20. Rodriguez, A.; Regnault, A.; Kleijmeer, M.; Ricciardi-Castagnoli, P.; Amigorena, S. Selective transport of internalized antigens to the cytosol for MHC class I presentation in dendritic cells. *Nat. Cell Biol.* **1999**, *1* (6), 362–368.
21. Sirica, A.E. *Cellular and Molecular Pathogenesis*; Lippincott-Raven: Philadelphia, 1996.
22. Ignatowicz, L.; Kappler, J.; Marrack, P. The repertoire of T cells shaped by a single MHC/peptide ligand. *Cell* **1996**, *84* (4), 521–529.
23. Carmichael, A.; Wills, M. The immunology of infection. *Medicine* **2005**, *33* (3), 3–9.
24. Yewdell, J.W.; Bennink, J.R. The binary logic of antigen processing and presentation to T cells. *Cell* **1990**, *62* (2), 203–206.
25. Cone, R.E.; Malley, A. Soluble, antigen-specific T-cell proteins: T-cell-based humoral immunity? *Immunol. Today* **1996**, *17* (7), 318–322.
26. Bachmann, M.F.; Ohashi, P.S. The role of T-cell receptor dimerization in T-cell activation. *Rev. Immunol. Today* **1999**, *20* (12), 568–575.
27. Patel, D.M.; Mannie, M.D. Intercellular exchange of Class II major histocompatibility complex/peptide complexes is a conserved process that requires activation of T cells but is constitutive in other types of antigen presenting cells. *Cell. Immunol.* **2001**, *214* (2), 165–172.
28. Choudhuri, K.; Kearney, A.; Bakker, T.R.; Merwe, P.A.V.D. Immunology: How do T cells recognize antigen? *Curr. Biol.* **2005**, *15* (10), R382–R385.
29. Guo, W.H.; Chan, K.L.; Dallman, M.; Tam, P.K.H. Role of CD4+ and CD8+ T cells in early and late acute rejection of small bowel allograft. *J. Pediatr. Surg.* **2001**, *36* (2), 352–356.
30. Moalem, G.; Tracey, D.J. Immune and inflammatory mechanisms in neuropathic pain. *Brain Res. Rev.* **2006**, *51* (2), 240–264.
31. Robey, E.A. Immunology guide for a cell-fate decision. *Nature* **2005**, *433* (24), 813–814.
32. Romagnani, S. The Th1/Th2 paradigm. *Immunol. Today* **1997**, *18* (6), 263–266.
33. Frossi, B.; Carli, M.D.; Piemonte, M.; Pucillo, C. Oxidative microenvironment exerts an opposite regulatory effect on cytokine production by Th1 and Th2 cells. *Mol. Immunol.* **2008**, *45* (1), 58–64.
34. Furuzawa-Carbadella, J.; Vargas-Rojas, M.I.; Cabral, A.R. Autoimmune inflammation from the Th17 perspective. *Autoimmun. Rev.* **2007**, *6* (3), 169–175.
35. Horiuchi, Z.; Kawano, Y.; Yamasaki, K.; Minohara, M.; Furue, M.; Taniwaki, T.; Miyazaki, T.; Kira, J.-I. Th1 dominance in HAM/TSP and the optico-spinal form of multiple sclerosis versus Th2 dominance in mite antigen-specific IgE myelitis. *J. Neurol. Sci.* **2000**, *172* (1), 17–24.
36. Hofstetter, H.H.; Ibrahim, S.M.; Koczan, D.; Kruse, N.; Weishaupt, A.; Toyka, K.V.; Gold, R. Therapeutic efficacy of IL-17 neutralization in murine experimental autoimmune encephalomyelitis. *Cell. Immunol.* **2005**, *237* (2), 123–130.
37. Schmidt-Weber, C.B.; Akdis, M.; Akdis, C.A. TH17 cells in the big picture of immunology. *J. Allergy Clin. Immunol.* **2007**, *120* (2), 247–254.
38. Bettelli, E.; Oukka, M.; Kuchroo, V.K. TH-17 cells in the circle of immunity and autoimmunity. *Nat. Immunol.* **2007**, *8* (4), 345–350.
39. Wilson, N.J.; Bonidace, K.; Chan, J.R.; McKenzie, B.S.; Blumenschein, W.M.; Mattson, J.D.; Basham, B.; Smith, K.; Chen, T.; Morel, F.; Lecron, J.-C.; Kastelein, R.A.; Cua, D.J.; McClanahan, T.K.; Bowman, E.P.; Malefyt, R.D.W. Development, cytokine profile and function of human interleukin 17-producing helper T cells. *Nat. Immunol.* **2007**, *8* (9), 950–957.
40. Bettelli, E.; Korn, T.; Kuchroo, V.K. Th17: The third member of the effector T cell trilogy. *Curr. Opin. Immunol.* **2007**, *19* (6), 652–657.
41. Aggarwal, S.; Gurney, A.L. IL-17: Prototype member of an emerging cytokine family. *J. Leukocyte Biol.* **2002**, *71* (1), 1–8.
42. Stockinger, B.; Veldhoen, M. Differentiation and function of Th17 T cells. *Curr. Opin. Immunol.* **2007**, *19* (3), 281–286.
43. Belich, M.P.; Trowsdale, J. Proteasome and class I antigen processing and presentation. *Mol. Biol. Rep.* **1995**, *21* (1), 53–56.
44. Modo, M.; Mellodew, K.; Rezaie, P. In vitro expression of major histocompatibility class I and class II antigens by conditionally immortalized murine neural stem cells. *Neurosci. Lett.* **2003**, *337* (2), 85–88.
45. Brown, R.E.; Roser, B.; Singh, P.B. Class I and Class II regions of the major histocompatibility complex both contribute to individual odors in congenic inbred strains of rats. *Behav. Genet.* **1989**, *19* (5), 659–674.

46. Buckley, R.H. Transplantation immunology: Organ and bone marrow. *J. Allergy Clin. Immunol.* **2003**, *11* (2), S733–S744.
47. Maiorana, A.; Cesinaro, A.M.; Fano, R.A.; Collina, G. Expression of MHC class I and class II antigens in primary breast carcinomas and synchronous nodal metastases. *Clin. Exp. Metast.* **1995**, *12* (1), 43–48.
48. Peters, V.B.; Sperber, K.E. The effect of viruses on the ability to present antigens via the major histocompatibility complex. *Microbes Infect.* **1999**, *1* (4), 335–345.
49. Schäffer, M.; Bongartz, M.; Hoffmann, W.; Viebahn, R. MHC-Class-II-deficiency impairs wound healing. *J. Surg. Res.* **2007**, *138* (1), 100–105.
50. Maffei, A.; Harris, P.E. Peptides Bound to major histocompatibility complex molecules. *Peptides* **1998**, *19* (1), 179–198.
51. Moss, D.J.; Khanna, R. Major histocompatibility complex: From genes to function. *Immunol. Today* **1999**, *20* (4), 165–167.
52. Germain, R.N. MHC-dependent antigen processing and peptide presentation: Providing ligands for T lymphocyte activation. *Cell* **1994**, *76* (2), 287–299.
53. Kittur, D.S.; Wilasrusmee, C.; Han, W.-F.; Xu, R.; Burdick, J.F.; Adler, W. Locally derived cytokines and upregulation of MHC Class II genes in allografts. *J. Heart Lung Transplant.* **2002**, *21* (8), 882–889.
54. Ting, J.P.-Y.; Trowsdale, J. Genetic control of MHC Class II expression. *Cell* **2002**, *109* (2), S21–S33.
55. Williams, D.B.; Vassilakos, A.; Suh, W.-K. Peptide presentation by MHC class I molecules. *Trends in Cell Biol.* **1996**, *6* (7), 267–273.
56. Faulkner, L.; Borysiewicz, L.K.; Man, S. The use of human leucocyte antigen class I transgenic mice to investigate human immune function. *J. Immunol. Methods* **1998**, *221* (1–2), 1–16.
57. Toumbis, C.A.; Kronick, J.L.; Wooley, P.H.; Nasser, S. Total joint arthroplasty and the immune response. *Semin. Arthritis Rheum.* **1997**, *127* (1), 44–47.
58. Case, C.P.; Langkamer, V.G.; James, C.; Palmer, M.R.; Kemp, A.J.; Heap, P.F.; Solomon, L. Widespread dissemination of metal debris from implants. *Journal of Bone and Joint Surgery [Br]* **1994**, *76B* (5), 701–712.
59. DiCarlo, E.F.; Bullough, P.G. The biologic responses to orthopaedic implants and their wear debris. *Clinical Materials* **1992**, *9* (3–4), 235–260.
60. Arora, A.; Song, Y.; Chun, L.; Huie, P.; Trindade, M.; Smith, R.L.; Goodman, S. The role of the TH1 and TH2 immune responses in loosening and osteolysis of cemented total hip replacements. *Journal of Biomedical Materials Research* **2003**, *64A* (4), 693–697.
61. Goodman, S.B. Wear particles, periprosthetic osteolysis and the immune system. *Biomaterials* **2007**, *28* (34), 5044–5048.
62. Hallab, N.J.; Mikecz, K.; Vermes, C.; Skipor, A.; Jacobs, J.J. Differential lymphocyte reactivity to serum-derived metal-protein complexes produced from cobalt-based and titanium-based implant alloy degradation. *Journal of Biomedical Materials Research* **2001**, *56* (3), 427–436.
63. Hercus, B.; Revell, P. Phenotypic characteristics of T lymphocytes in the interfacial tissue of aseptically loosened prosthetic joints. *Journal of Materials Science: Materials in Medicine* **2001**, *12* (10–12), 1063–1067.
64. Lalor, P.; Revell, P. T lymphocytes and titanium aluminium vanadium (TiAlV) alloy: evidence for immunological events associated with debris deposition. *Clinical Materials* **1993**, *12* (1), 57–62.
65. Wooley, P.H.; Nasser, S.; Fitzgerald, R.H. The Immune Response to Implant Materials in Humans. *Clinical Orthopaedics and Related Research* **1996**, *326* (3), 63–70.
66. Santavirta, S.; Kontinen, Y.; Bergroth, V.; Grönblad, M. Lack of immune response to methyl methacrylate in lymphocyte cultures. *Acta Orthopaedica Scandinavica* **1991**, *62* (1), 29–32.
67. Doorn, P.F.; Campbell, P.A.; Worrall, J.; Benya, P.D.; McKellop, H.A.; Amstutz, H.C. Metal wear particle characterization from metal on metal total hip replacements: transmission electron microscopy study of periprosthetic tissues and isolated particles. *Journal of Biomedical Materials Research* **1998**, *42* (1), 1998.
68. Haddad, F.S.; Cobb, A.G.; Bentley, G.; Levell, N.J.; Dowd, P.M. Hypersensitivity in aseptic loosening of total hip replacements. The role of constituents of bone cement. *The Journal of Bone and Joint Surgery [BR]* **1996**, *78B* (4), 546–549.
69. Methé, H.; Edelman, E.R. Tissue Engineering of Endothelial Cells and the Immune Response. *Transplantation Proceedings* **2006**, *38* (10), 3293–3299.
70. Granchi, D.; Ciapetti, G.; Savarino, L.; Stea, S.; Filipini, F.; Sudanese, A.; Rotini, R.; Giunti, A. Expression of the CD69 activation antigen on lymphocytes of patients with hip prosthesis. *Biomaterials* **2000**, *21* (20), 2059–2065.
71. Trindade, M.C.D.; Lind, M.; Sun, D.; Schurman, D.J.; Goodman, S.B.; Smith, R.L. In vitro reaction to orthopaedic biomaterials by macrophages and lymphocytes isolated from patients undergoing revision surgery. *Biomaterials* **2001**, *22* (3), 253–259.
72. Baldwin, L.; Flanagan, B.F.; McLaughlin, P.J.; Parkinson, R.W.; Hunt, J.A.; Williams, D.F. A study of tissue interface membranes from revision accord knee arthroplasty: the role of T lymphocytes. *Biomaterials* **2002**, *23* (14), 3007–3014.
73. Voggenreiter, G.; Leiting, S.; Brauer, H.; Leiting, P.; Majetschak, M.; Bardenheuer, M.; Obertacke, U. Immuno-inflammatory tissue reaction to stainless-steel and titanium plates used for internal fixation of long bones. *Biomaterials* **2003**, *24* (2), 247–254.
74. Vallés, G.; González-Melendi, P.; González-Carrasco, J.L.; Saldaña, L.; Sánchez-Sabaté, E.; Munuera, L.; Vilaboa, N. Differential inflammatory macrophage response to rutile and titanium particles. *Biomaterials* **2006**, *27* (30), 5199–5211.
75. Luyn, M.J.A.v.; Plantinga, J.A.; Brouwer, L.A.; Khouw, I.M.S.L.; Leij, L.F.M.H.d.; Wachem, P.B.v. Repetitive subcutaneous implantation of different types of (biodegradable) biomaterials alters the foreign body reaction. *Biomaterials* **2001**, *22* (11), 1385–1391.

Biomaterials: Protein–Surface Interactions

Robert A. Latour, Jr.

Department of Bioengineering, Clemson University, Clemson, South Carolina, U.S.A.

INTRODUCTION

The study of protein–surface interactions represents one of the most important topics in the field of biomaterials, and as such, it has been a focus of intensive study for several decades. The reason for the great interest in this topic is the realization that protein–surface interactions are fundamentally responsible for the biocompatibility of medical devices, or the lack thereof.

When a solid material (e.g., a catheter, stent, hip joint replacement, or tissue engineering substrate) comes in contact with a fluid that contains soluble proteins (e.g., blood, interstitial fluid, cell culture media), proteins rapidly adsorb onto the surface of the material, saturating the surface within a time frame of seconds to minutes. Therefore, when living cells (which are much larger than proteins and thus much more slowly moving) approach the biomaterial surface, they do not actually contact the molecular structure of the material surface itself, but rather they contact and interact with the molecular structure of the adsorbed protein layer. Cells, of course, cannot “see” the adsorbed protein layer, but rather they interrogate their surroundings by way of membrane-bound receptors that can bind to specific bioactive features presented by the adsorbed proteins. Then, through a series of orchestrated molecular mechanisms, these receptor–protein binding events are transduced through the cell membrane in a manner that stimulates specific intracellular processes that then determine a cell’s response. Accordingly, at the most fundamental level the key to controlling cellular response is to control the type of bioactive sites that are presented by the adsorbed layer of proteins. This, in turn, can be controlled by controlling the amounts and the types of proteins that are adsorbed and their orientation, conformation, and packing arrangement on the biomaterial surface. While this is conceptually simple to understand, the numerous types of soluble proteins contained in physiological fluids combined with their structural complexity has made, and continues to make, this an extremely challenging problem.

Many excellent reviews and entire books have been written over the past couple of decades that address the multitude of issues related to the interaction of proteins with surfaces. Several of these reviews and books are referenced in the “Further Readings” section at the

end of this article and the readers interested in this topic are encouraged to seek out these additional sources of information. Given the vast number of studies that have been conducted on this topic, it is of course impossible to provide a complete review of the literature in one article. Instead, the goal here is to provide an overview of the basic understanding that has been achieved over the years regarding how proteins interact with surfaces. This will be accomplished by first providing an overview of protein adsorption processes and then by addressing the special case of the design of surfaces to prevent protein adsorption. Following these topics, attention will be focused on highlighting several of the most interesting relatively recent techniques that have been developed and applied to further our understanding of the sub-molecular-level mechanisms involved in how surface chemistry influences the orientation, conformation, and organization of adsorbed proteins. The continued development of our understanding of these processes is critical if we are to get beyond the current era of surface design largely by trial and error, and move into an era where surfaces are proactively designed to directly control adsorbed protein bioactivity, and thereby control cellular response. This article will then close with a conclusion section that addresses future directions toward the continuing goal of the development of truly biocompatible materials for the design of medical devices for improved patient care.

OVERVIEW OF PROTEIN–SURFACE INTERACTIONS

As introduced in the prior section, the surfaces of synthetic biomaterials (e.g., polymers, metals, and ceramics) are generally not bioactive themselves. Rather, surface bioactivity is provided by the proteins that adsorb to the biomaterial surface following exposure of the surface to biological fluids. The types and the amounts of proteins adsorbed determine the types and surface density of the bioactive sites that may be available for cell interactions, and the orientation, conformation, and packing density of the adsorbed proteins determine whether the available bioactive sites are presented in a manner such that they can be recognized by the membrane-bound receptors of cells as they interrogate the adsorbed protein layer.

Accordingly, to control cellular response, it is important to first understand how surface chemistry and surface topology influence the formation of the adsorbed protein layer and the bioactive sites presented by this layer.

Although the understanding of protein adsorption to biomaterial surfaces is still far from complete, research over the past several decades has led to a general understanding of the complex and multifaceted processes involved in the interactions between proteins and surfaces. In this section, an overview of protein structure is first presented to establish a framework from which protein adsorption behavior can begin to be understood. Following this, an overview of the current understanding of protein adsorption behavior to biomaterial surfaces is presented, initially for the relatively simple case involving protein adsorption from single-component protein solutions, and then for the more complex situation involving competitive adsorption from multicomponent protein solutions. This section will then be completed by addressing the special case of the development of nonfouling surfaces that are highly resistant to protein adsorption.

Protein Structure

Before protein adsorption behavior can be understood, it is first important to understand the basic makeup of protein structure. Proteins are complex copolymers that are made up of four levels of structure, designated as the primary, secondary, tertiary, and quaternary structures.^[1] The primary structure involves the specific sequence of the 20 L-amino acids coded for by the DNA of a cell. As part of a protein, a given amino acid is referred to as a peptide residue. Each amino acid, or residue, has the general backbone structure of $\{-\text{NH}-\text{C}_\alpha\text{HR}-\text{CO}-\}$, with R designating a specific side-group structure that gives the residue its specific functional characteristics. Accordingly, the amino acids are subcategorized into three primary types: nonpolar (i.e., hydrophobic), \pm -charged, and polar. Examples of each type of amino acid are shown in Fig. 1A. The polypeptide chain formed by the primary sequence is then organized into three basic types of secondary structure: α -helices, β -sheets, and loops that connect helix and sheet elements (Fig. 1B). These secondary elements are then organized together

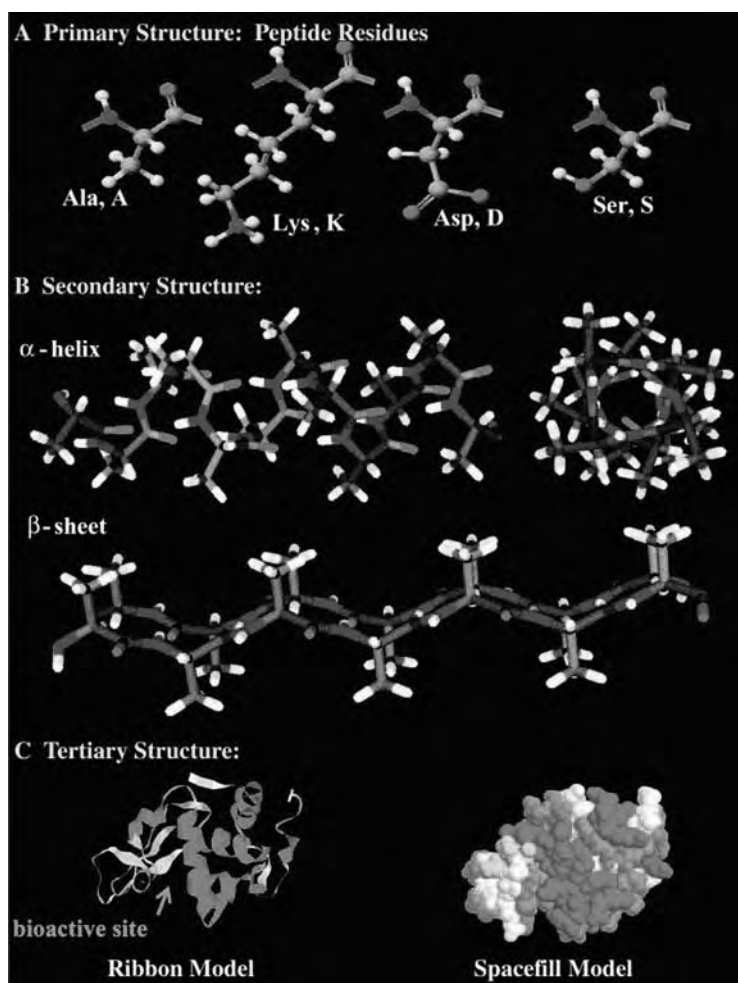


Fig. 1 Protein structure. (A) Primary structure showing examples of four types of residues with designated three-letter and one-letter codes: nonpolar (alanine, Ala, A), positively charged (lysine, Lys, K), negatively charged (aspartic acid, Asp, D), polar (serine, Ser, S). Color code: gray = C, white = H, blue = N, red = O. (B) Secondary structures composed of an alanine sequence: α -helix (views from side and end), β -sheet. Same color code as in (A). (C) Tertiary structure (lysozyme, PDB ID 7LYZ). Color code: yellow = β -sheet, pink = α -helix, blue and white = connecting loops.

to form the tertiary structure (Fig. 1C). Finally, more than one polypeptide chain can be organized together to form quaternary structure, with each individual chain having a separate beginning (N-terminus) and ending (C-terminus), which are typically positively and negatively charged, respectively. The specific functional group character of the peptide residues combined with their spatial organization creates specific bioactive domains in a protein's structure that enable it to perform its specific biological function, whether that is as an enzyme, growth factor, intracellular structure, or extracellular matrix (ECM) assembly, or cell adhesion molecule for cell–cell and cell–ECM interactions.

While the subject of how a polypeptide chain goes about folding into its native structure is still a very active area of research, it is well understood that one of the primary driving forces that causes a protein to fold into its native state, and to be maintained in that state in aqueous solution, is the reduction in free energy due to a decrease in the solvent accessible surface area of the nonpolar residues contained in the protein's structure, i.e., hydrophobic effects.^[1] Accordingly, proteins are generally structured with their hydrophobic residues buried within the core of the protein and their hydrophilic residues (charged and polar) lining the protein's solvent accessible surface. While this generally describes a protein's structural makeup, it is common to have some hydrophobic residues on the surface and hydrophilic residues buried within the protein's core. Thus, a protein's surface is highly amphiphilic, meaning that it displays a myriad of different types of functional groups (nonpolar, \pm -charged, polar) on its surface. As a further feature that can greatly influence protein adsorption behavior, each charged residue on a protein's surface has a designated pK_a value, which defines the residue's protonation state (i.e., whether it is in its charged or neutral state). The overall charged state of a protein is thus sensitive to the pH of the surrounding solution, with the pH value that results in the protein having a net zero value of charge being designated as the isoelectric point (or pI) of the protein. Further details of protein structure are provided in standard biochemistry texts and readers interested in protein adsorption behavior are encouraged to study this material to provide a foundation from which protein adsorption behavior can then be better understood.^[2]

Protein Adsorption Behavior from Single-Component Protein Solutions

The general model used to describe protein adsorption to a surface is illustrated in Fig. 2.^[3] This process

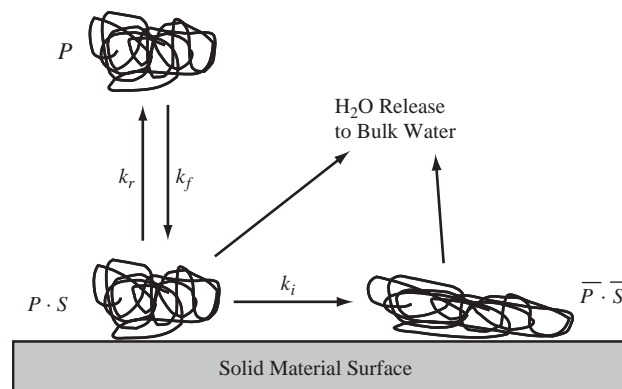
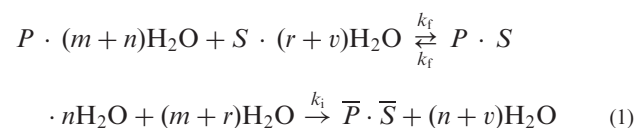


Fig. 2 Illustration of the protein adsorption process. A protein molecule in solution (P) adsorbs to the material surface ($P \cdot S$) in a reversible manner with forward and reverse reaction rate constants, k_f and k_r , respectively. The adsorbed protein then may transition to an irreversibly adsorbed state ($\bar{P} \cdot \bar{S}$) as a result of orientational or conformational changes that further reduce the free energy of the system, which can be described with a reaction rate constant k_i . The release of structured water from the protein and surface back to bulk water represents an important factor that greatly influences these processes. (From Ref. [3].)

can be described by the following chemical reaction equation:



where P is the protein concentration in solution, S is the concentration of available surface area sites for adsorption, $P \cdot S$ is the concentration of reversibly adsorbed protein on the surface with k_f and k_r designating reaction rate constants for the forward and reverse reactions, respectively, and $\bar{P} \cdot \bar{S}$ is the concentration of irreversibly adsorbed protein on the surface with k_i designating the reaction rate constant for the transition of the adsorbed protein from a reversibly bound to an irreversibly bound state. Water is shown complexed to the protein and surface, which is then released to bulk solution during the various stages of adsorption. This is included to emphasize its importance as a molecular constituent in the adsorption process rather than simply being an inert media that the protein and surface happen to be immersed in. Because of the presence of the irreversible step shown at the right-hand side of this equation, a surface that adsorbs a protein in a manner that includes this part of the process will inevitably become saturated with the adsorbed protein if sufficient interaction time is provided. The state of the final adsorbed protein layer will

be determined by the manner in which the proteins are able to organize themselves on the surface, which is influenced by both protein–surface and protein–protein interactions in the presence of the surrounding aqueous solution. The proteins that adsorb to the surface in the latter stages of this process will tend to be increasingly inhibited from the irreversible step of this process by steric restrictions caused by the previously adsorbed proteins. Because of this, the final fraction of the adsorbed protein layer is more likely to be maintained in its native and possibly reversibly adsorbed state.^[4] This behavior, which is illustrated in Fig. 3, explains the frequently observed phenomenon that a fraction of protein is found to desorb from an adsorbed protein layer following exposure of the surface to pure buffer solution while another fraction of the protein layer remains irreversibly adsorbed to the surface. This type of behavior can also be caused by a nonhomogenous surface (e.g., owing to defects in the surface chemistry or phase separation), thus presenting one surface phase that irreversibly adsorbs the protein and another more weakly interacting phase that reversibly adsorbs the protein.

Assuming the conditions of adsorption from a dilute solution and unity activity coefficients of each species, and defining the standard states of the products and reactants of Eq. (1) as being 1.0 M (protein solution) and 1.0 mole fraction (adsorbed protein and surface), the kinetics of protein adsorption for Eq. (1) can be expressed as:^[2,5]

$$\begin{aligned} \frac{d[P \cdot S]}{dt} &= k_f[P][S] - k_r[P \cdot S] - k_i[P \cdot S], \\ \frac{d[\bar{P} \cdot \bar{S}]}{dt} &= k_i[P \cdot S] \end{aligned} \quad (2)$$

where $[S]$, $[P \cdot S]$, and $[\bar{P} \cdot \bar{S}]$ are the mole fraction of surface sites occupied by bare surface and the reversibly and irreversibly adsorbed protein, respectively, and $[P]$ is the molar concentration of the protein in solution. The change in the free energy for the reversible part of the adsorption process can then be expressed as:

$$\Delta G_{\text{ads}} = \Delta G_{\text{ads}}^{\circ} + RT \ln \left(\frac{[P \cdot S]}{[P][S]} \right)$$

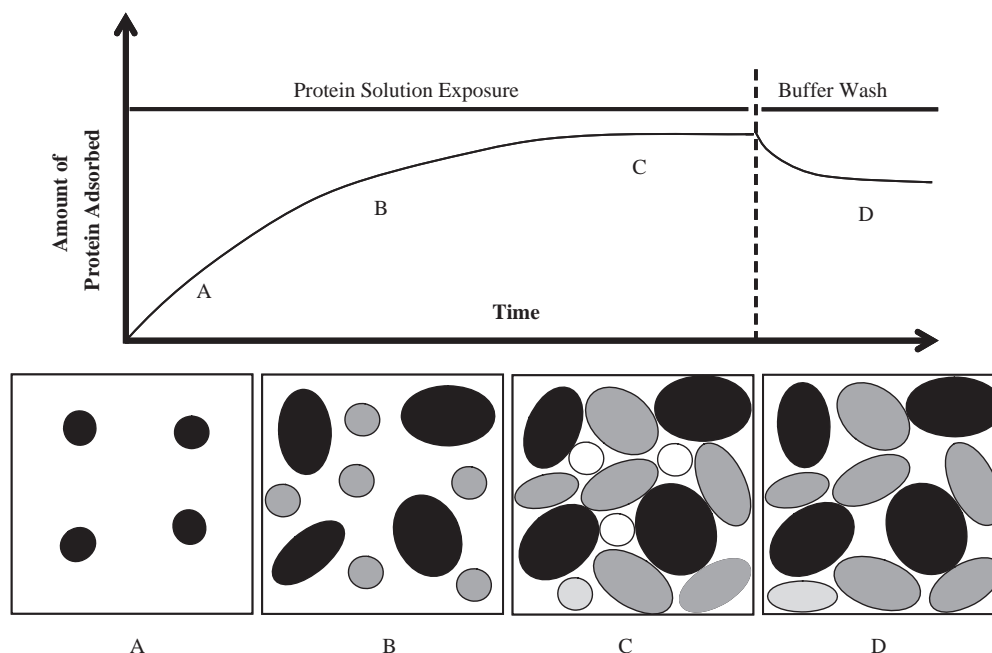


Fig. 3 Illustration of the development of an adsorbed protein layer formed when a surface is exposed to a single-component protein solution of constant concentration followed by a rinsing step of pure buffer solution. (A) Proteins initially adsorb to random areas of the surface, possibly in a reversible fashion. (B) The previously adsorbed proteins begin reorienting and spreading out on the surface to become irreversibly adsorbed (black ellipses) while new protein molecules (gray circles) continue to adsorb to open surface areas. (C) This process continues until most of the available surface area is covered with irreversibly adsorbed proteins. The last proteins to adsorb may be sterically prevented from reorienting or spreading on the surface by the previously adsorbed proteins, thus being maintained in a more weakly held, reversibly adsorbed state by the surface (white circles). (D) When the surface is exposed to pure buffer solution, the reversibly adsorbed fraction is able to desorb from the surface with the irreversibly adsorbed fraction remaining.

and at equilibrium:

$$\begin{aligned}\Delta G_{\text{ads}}^{\circ} &= -RT \ln \left(\frac{[P \cdot S]}{[P][S]} \right)_{\text{eq}} = -RT \ln \left(\frac{q}{C_e(Q - q)} \right) \\ &= -RT \ln \left(\frac{\Gamma}{C_e(1 - \Gamma)} \right) = -RT \ln(K) \quad (3)\end{aligned}$$

where ΔG_{ads} and $\Delta G_{\text{ads}}^{\circ}$ are the change in adsorption free energy and change in standard state adsorption free energy, R is the ideal gas constant, T is the absolute temperature, q is the amount of protein per unit area adsorbed in equilibrium with the molar concentration of protein in solution (C_e), Q is the amount of protein per unit area adsorbed at monolayer saturation, Γ is the fraction of surface saturation (or q/Q), and K is the equilibrium constant, which also equals (k_f/k_r).

Eqs. (2) and (3) provide a means to quantitatively describe the reversible part of the protein adsorption process for a given set of conditions. For example, Table 1 provides values for $\Delta G_{\text{ads}}^{\circ}$ and KC_e that correspond to various values of Γ for two different proteins, lysozyme (14.6 kDa) or fibrinogen (340 kDa), reversibly adsorbing to a surface in equilibrium with their respective protein solution concentrations of 1.0 mg/mL (commonly used solution concentration for experimental studies). Here, KC_e , which is derived from Eq. (2) for the reversible part of the adsorption process, expresses the ratio of the initial adsorption rate of protein ($d[P \cdot S]/dt$) onto a completely bare surface compared to the initial desorption rate from a saturated surface into a pure buffer solution over the surface. Taking the energy of a single hydrogen bond to be on the order of -3.5 kcal/mol, which is also on the order of the energy of adsorption of a single hydrophobic peptide residue onto a hydrophobic

surface, the values shown in Table 1 indicate that if the interaction of either of these proteins with a surface involves a decrease in $\Delta G_{\text{ads}}^{\circ}$ of the system no more than that of the formation of about one additional hydrogen bond, or one hydrophobic contact, then the adsorbed amount can be kept to about 0.1% of a monolayer of adsorbed protein.^[6] In contrast to this, a $\Delta G_{\text{ads}}^{\circ}$ value of only -15 kcal/mol (i.e., equivalent to approximately four hydrogen bonds or hydrophobic interactions) will result in essentially 100% monolayer coverage of the surface. Furthermore, considering the ratio of the rates of adsorption to desorption, as indicated by KC_e , an adsorption energy of only -15 kcal/mol results in an adsorption rate that is about one million times faster than desorption. In other words, the number of protein macromolecules adsorbed in 1 sec would take over 10 days to desorb, thus effectively representing an irreversible condition. Therefore, considering the size of many proteins (such as fibrinogen) and the resulting large number of functional group interactions involved when such a protein contacts a surface, it is not surprising that the adsorption of proteins to biomaterial surfaces generally occurs in an effectively irreversible manner, even without considering the actual irreversible stage of protein adsorption indicated in Eq. (1).

Based on the adsorption processes indicated by Eq. (1), when a solid surface is placed in a solution containing a given protein, the protein will generally tend to rapidly adsorb until it saturates the surface. If the surface is hydrophobic, the protein will tend to adsorb by the various hydrophobic patches of residues present on the protein's amphiphilic surface, with the protein then tending to unfold and spread its hydrophobic core over the surface owing to the thermodynamic driving force to reduce the net hydrophobic surface area of the system exposed to the solvent.^[7]

Table 1 Corresponding values of the standard state free energy of adsorption ($\Delta G_{\text{ads}}^{\circ}$) and relative adsorption/desorption rates (KC_e) for a given percent surface saturation

Percent saturation ($\Gamma \times 100$)	Lysozyme (14,600 Da, 1 mg/mL solution)		Fibrinogen (340,000 Da, 1 mg/mL solution)	
	$\Delta G_{\text{ads}}^{\circ}$ (kcal/mol)	KC_e (unitless)	$\Delta G_{\text{ads}}^{\circ}$ (kcal/mol)	KC_e (unitless)
0.10	-1.65	0.001	-3.59	0.001
1.00	-3.08	0.010	-5.02	0.010
10.00	-4.55	0.111	-6.49	0.111
30.00	-5.38	0.429	-7.32	0.429
60.00	-6.16	1.500	-8.10	1.500
90.00	-7.26	9.000	-9.20	9.000
99.00	-8.74	99.00	-10.68	99.00
99.99	-11.58	10,000	-13.52	10,000
99.9999	-14.42	1,000,000	-16.36	1,000,000

Hydrophilic surfaces, on the other hand, tend to interact with the charged and polar functional groups of the protein's surface, thus influencing adsorbed protein orientation, but with a lower tendency to cause the protein to unfold and spread over the surface.^[7] Because proteins generally do not tend to adsorb nonspecifically to themselves (which is why they are soluble in solution in the first place), protein adsorption tends to stop once complete monolayer coverage is achieved. The final organization of the adsorbed protein layer (i.e., the orientation, conformation, and packing arrangement of the adsorbed proteins) depends on the chemical and physical structure of the protein, the surface, and the aqueous solution, and the thermodynamics of the interactions between these system components. It is also strongly influenced by the kinetics of mass transport of the protein to the surface from solution relative to the kinetics of the protein's movement on the surface (i.e., orientational and translational motions), the kinetics of protein spreading on the surface (i.e., conformational changes), and the physical constraints imposed by protein–protein interactions on the surface.^[8]

Because protein adsorption is influenced by so many factors, many of which result in irreversible processes, the final state of an adsorbed protein layer is history dependent, meaning that very different results can be obtained for the same protein–surface system depending on how the adsorption process is carried out. For example, it is generally true that a hydrophobic surface will more strongly adsorb proteins than a neutrally charged hydrophilic surface. Because of this, it is usually believed that a hydrophobic surface will then adsorb a greater amount of protein.^[9] This will tend to be true, however, only if the protein is adsorbed from a solution with high protein concentration such that mass transport to the surface is much faster than the rate of protein spreading/reorientation on the surface. Under this condition, the hydrophobic surface will tend to adsorb more protein than the hydrophilic surface because its stronger attraction for the protein will result in a more densely packed layer of protein at surface saturation. This situation is illustrated in Fig. 4 for adsorption to a hydrophobic surface “a” compared to a hydrophilic surface “b” from a relatively concentrated protein solution. The opposite result can occur, however, if the same protein is adsorbed from a very dilute solution, which will greatly slow down the rate of mass transport of the protein to the surface. If the rate of mass transfer is slow compared to the rate of protein spreading/reorientation on the surface once it adsorbs, then the hydrophilic surface may actually adsorb more protein than the hydrophobic surface. This occurs because spreading tends to occur to a much greater extent on the more hydrophobic surface, thus causing an adsorbed protein

molecule to occupy a greater area of surface, leading to a lower amount of protein that can be adsorbed before the surface is saturated (illustrated in Fig. 4; conditions “c” and “d”). As a further consequence of this concentration effect, a plot of the amount of adsorbed protein vs. solution concentration can erroneously give the appearance of a Langmuir isotherm adsorption plot (see Fig. 4III), which has often mistakenly been taken to imply a reversibly adsorbing system.^[10] In fact, many investigators have wrongly applied Eq. (3) to such isotherms and calculated $\Delta G_{\text{ads}}^{\circ}$ for the adsorption process.^[11,12] As illustrated in Fig. 4, however, a Langmuir-looking isotherm will often occur from an irreversible protein adsorption process simply owing to protein spreading effects. In such cases, the calculation of $\Delta G_{\text{ads}}^{\circ}$ is meaningless; complete reversibility must be established as a necessary (but still not sufficient) requirement for the application of the Langmuir adsorption model for the meaningful calculation of $\Delta G_{\text{ads}}^{\circ}$.

The types of processes described in the preceding paragraph have been clearly demonstrated in experimental studies by Wertz and Santore.^[7] These investigators compared the kinetics and adsorption behavior of albumin and fibrinogen on hydrophobic and hydrophilic surfaces (CH_3 , CF_3 , OH silane-functionalized glass) using total internal reflectance fluorescence spectroscopy. In these studies, the rate of mass transfer of the protein to the surface was controlled by varying both the solution concentration and flow rate over each surface, with the adsorption process being continued until the condition of surface saturation was achieved. The total amount of protein adsorbed to each surface was found to directly correlate with the rate of mass transport provided, with the surface area per adsorbed protein (referred to as the protein's footprint) thus increasing as the mass transport rate was decreased. Their results also indicated that the increase in the size of the adsorbed protein's footprint as the adsorption process was slowed down was associated with the reorientation of the adsorbed protein on the hydrophilic surface and conformational spreading on the hydrophobic surface. Similar events have been demonstrated by Agnihotri and Siedlecki, who used tapping-mode atomic force microscopy (AFM) to observe time-dependent structural differences in fibrinogen adsorbed to a hydrophilic (mica) vs. a hydrophobic (graphite) surface.^[13] In this study, four different orientational states were observed on each surface, with the height of the adsorbed protein on the hydrophobic surface then undergoing a significant reduction over a time period of 2 hr following initial adsorption, while a height increase was actually observed over time on the hydrophilic surface. It was also observed that the AFM tip tended to move the protein around on the hydrophilic surface, but not on

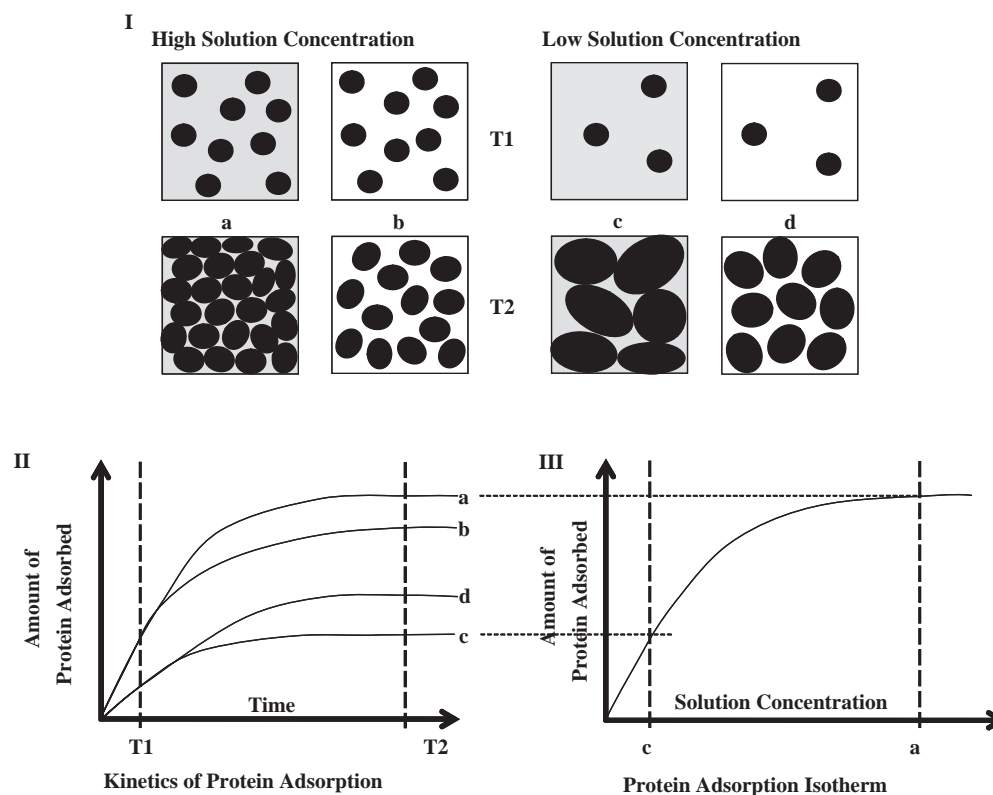


Fig. 4 Illustration of the history dependence of the final state of a protein layer adsorbed from a single-component protein solution. (I) Comparison of the resulting adsorbed protein layers when the same protein is adsorbed from a high vs. low protein concentration solution on hydrophobic (a, c) vs. hydrophilic (b, d) surfaces at time points T1 and T2 ($T1 < T2$), with T2 representing the final condition at surface saturation. (II) The amount of adsorbed protein vs. time for conditions (a)–(d) showing that mass transport controls the kinetics of the initial adsorption process as opposed to surface chemistry, with the final amount adsorbed at surface saturation dependent on both mass transport and surface chemistry. (III) Resulting irreversible adsorption isotherm with reversible Langmuir isotherm-like appearance. The concentration dependence is due to a change in the molecular footprint of the adsorbed protein as opposed to an equilibrium exchange with the protein in solution. Each point on the isotherm plot represents a fully saturated irreversibly adsorbed protein layer. (From Refs. [7,10].)

the hydrophobic surface, thus suggesting a more tightly bound interaction of the protein with the hydrophobic surface. These results are again consistent with the understanding that hydrophobic surfaces tend to more strongly adsorb proteins and induce conformational spreading, while hydrophilic surfaces tend to more weakly adsorb protein and induce orientational changes in the adsorbed protein as a function of time.

Protein Adsorption Behavior from Multicomponent Protein Solutions

When more than one type of protein is present in solution, a competitive process occurs between the different proteins for adsorption to the surface. The first important factor that influences the adsorption of proteins from a multicomponent system is that the mass transfer rate of a given solute molecule to a surface is directly related to its solution concentration

and inversely related to its molecular weight.^[14] Accordingly, when a material is exposed to a solution containing several different soluble proteins, such as blood plasma, the more concentrated and smaller proteins tend to adsorb to the surface first, and then be displaced by larger, more strongly interacting proteins that may arrive at the surface at a later point in time.^[15] This exchange process, known as the Vroman effect, was first recognized by Vroman and Adams in the late 1960s for the case of fibrinogen exchange from various types of surfaces.^[16,17] Although initially believed to be unique for fibrinogen exchange, numerous subsequent studies have indicated that this is actually a general phenomenon that occurs for many other types of proteins, including albumin, IgG, and fibronectin.^[16] Interestingly, this process is observed to occur even for proteins that are otherwise irreversibly adsorbed to a surface. Thus, while an adsorbed protein may not tend to desorb from the surface under the flow of pure buffer solution, it may be displaced from the surface by another more strongly interacting protein

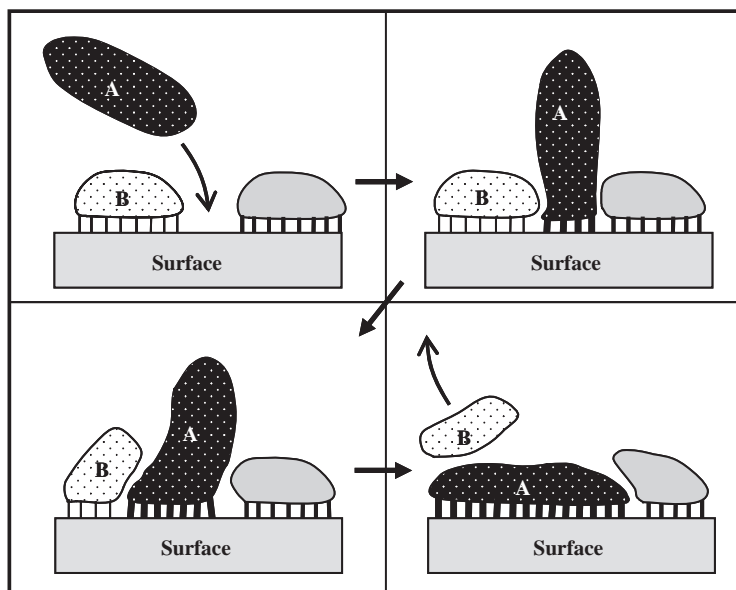


Fig. 5 Illustration of the Vroman effect. Protein “B,” which is initially adsorbed to the surface, is displaced from the surface in an exchange reaction by protein “A” owing to the ability of the residue functional groups presented by “A” to form more stable bonds with the available binding sites on the surface than the residue functional groups of “B.”

that can more successfully compete for the same surface functional groups, thus exchanging itself for the previously adsorbed protein. This process is illustrated in Fig. 5. Many interesting studies have been conducted over the past several decades to investigate the competitive nature of proteins; only a few of which will be highlighted here. Readers are therefore referred to the “Further Readings” section for further references related to this subject.

As an example of the intriguing types of behavior described by Vroman effects, fibrinogen preadsorbed on a negatively charged, hydrophilic glass surface has been observed to be readily displaced following exposure of the surface to various dilutions of blood plasma.^[16] Immunochemistry studies have revealed that fibrinogen is primarily displaced by high molecular weight kininogen (HMWK) under these conditions. In other studies involving the preadsorption of fibrinogen to polyethylene, the behavior is quite different.^[16,18] While fibrinogen is also displaced from a polyethylene surface following exposure to blood plasma solutions, it has been shown that HMWK is not involved in the exchange process. Instead, other proteins, especially high-density lipoproteins, have been shown to be responsible for fibrinogen exchange. Warkentin et al. have conducted studies with multi-component protein solutions that shed some light on why this difference occurs.^[19] Using hydrophobic-gradient surfaces, they found that HMWK exhibited distinctly higher affinity for the hydrophilic end of the gradient-surface compared to other proteins, thus demonstrating that HMWK’s ability to displace fibrinogen from hydrophilic but not hydrophobic surfaces is simply due to HMWK’s inherent ability to bind most strongly to hydrophilic surfaces rather than

due to a specific interaction between HMWK and fibrinogen. More recent studies by Cremer and coworkers have provided insights into some of the actual molecular-level mechanisms that influence Vroman effects with regard to fibrinogen displacement from surfaces.^[20] Their studies showed that at a pH of 8.0, fibrinogen adsorbs to a negatively charged silica surface via the relatively small positively charged α C domains of the protein, while the larger D and E domains of fibrinogen, which are negatively charged, are not adsorbed to the surface because of electrostatic repulsion. This results in the protein being only weakly bound and enables it to be readily displaced from the surface by other, more strongly adsorbing proteins like HMWK. However, at a pH of 3.2, the negative charges of both the protein and this surface are largely neutralized, thus eliminating the electrostatic repulsive effects that previously prevented the D and E domains from adsorbing to the surface. Under these conditions, the D and E domains of fibrinogen are able to participate in the adsorption process via hydrogen bonding and van der Waals interactions with functional groups on the silica, which enables fibrinogen to bind much more strongly to the silica surface. This was observed to reduce the rate of fibrinogen displacement from the silica surface by other plasma proteins by over two orders of magnitude. Other studies have shown that Vroman effects are not only influenced by surface chemistry and solution pH, but also, not surprisingly, by many other variables including solution composition, temperature, and wall shear strain rate.^[16]

Other interesting protein adsorption effects due to protein–protein interactions have been shown by Wertz and Santore. As part of the same study referred

to in the previous section, Wertz and Santore studied the competitive adsorption behavior between fibrinogen and albumin.^[7] In these studies, the surfaces were first saturated with albumin and then aged under an albumin containing solution, following which the surfaces were exposed at various time points to a fibrinogen solution. The amount of fibrinogen adsorption and albumin desorption was then monitored as a function of time by separate radiolabeling studies. The results from these studies indicated that albumin saturated each surface very quickly and then maintained a constant surface density irrespective of aging time. Following exposure to the fibrinogen solution, they found that about the same amount of albumin was displaced by the fibrinogen irrespective of prior albumin exposure time. However, the amount of fibrinogen adsorbed was found to be strongly dependent on the aging time of the adsorbed albumin. This was interpreted to indicate that although the amount of albumin adsorbed did not change with time during the aging phase of the experiment, the pre-adsorbed albumin was able to undergo substantial reorganization during this time period, more effectively covering the surface area, and perhaps strengthen its interaction by interprotein associations, such that less surface area was then available for subsequent fibrinogen adsorption. This effect was observed to occur most strongly on the hydrophobic surfaces, with the implication that surface reorganization on the hydrophilic surface mainly occurred as a result of reorientational effects, while the hydrophobic surfaces induced additional effects due to conformational spreading.

A final study that will be addressed here, reported by Lee et al., compared the competitive binding capabilities on glass for a wild-type T4 lysozyme and two single-residue mutants, one designed to be more conformationally stable and one less conformationally stable than the wild-type lysozyme.^[21] Using binary mixtures of each combination of the proteins, the authors were able to show that in each case, the more stable protein was more easily displaced from the surface by the less stable protein, thus nicely demonstrating that conformation stability plays an important role in the competitive adsorption behavior of proteins.

Surfaces that are Resistant to Protein Adsorption

Surfaces that are resistant to protein adsorption, also referred to as nonfouling surfaces, represent a special case in the study of protein–surface interactions. The development of nonfouling surfaces is of substantial interest for marine, food processing, bioMEMs, and

biosensor applications, in addition to their potential application for the development of medical devices with improved biocompatibility. Of the various types of surface chemistries that have been studied, poly(ethylene glycol) (PEG; $[-(\text{O}-\text{CH}_2-\text{CH}_2)_n-\text{OH}]$) has received the greatest amount of attention because it has been widely demonstrated to be very resistant to the adsorption of proteins from aqueous solution. Many of the earlier attempts to understand and explain the resistance to protein adsorption imparted by PEG-coated surfaces centered on the Alexander-De Gennes theory of polymer interfaces.^[22] This theory essentially presents a “physical view” that supposes that protein resistance is provided by steric repulsion of the protein due to the conformational entropy of the very flexible PEG chains on the surface, which acts as a physical barrier that prevents a protein from approaching the underlying substrate. The development of this theory, however, does not address why the protein does not favorably interact with the PEG molecule itself. In fact, its development is based on the a priori assumption of equal attraction between the PEG chains, water, and protein, thus inherently neglecting any enthalpic contribution to this process. In a relatively recent paper by Morra (22), this deficiency is highlighted and the viewpoint expressed that a more sophisticated theory is needed to explain the non-adsorptive behavior of PEG.^[22] Morra suggests that the physical view be combined with a “chemical view” that includes the strongly favorable enthalpic interactions between the PEG chains and water. This position is supported by a report by Kjellander that addresses the thermodynamics of PEG–water interactions and indicates that PEG apparently interacts with water structure in a very energetically favorable manner that is rather unique among the class of polyether polymers, which are otherwise largely insoluble in water.^[23] While providing an improvement over the relatively simple steric repulsion theory of PEG, this approach still neglects the other side of the issue, which is equally important, namely, why are PEG–protein interactions thermodynamically less favorable than the interactions of PEG with water?

Although a definitive theory regarding how to design molecular structures to provide resistance against protein adsorption has not yet been agreed upon, several attempts have been made to characterize the general qualities necessary for the design of such surfaces. Several studies using oligo(ethylene glycol) (OEG) functionalized alkanethiol self-assembled monolayer (SAM) surfaces have demonstrated that protein adsorption resistance of OEG with either hydroxyl or methoxy end groups ($[-(\text{O}-\text{CH}_2-\text{CH}_2)_n-\text{X}]$; X: OH, O–CH₃) is dependent on the degree of polymerization and chain packing density.^[24–27] Herrwerth et al. have shown that protein resistance requires that

the degree of OEG polymerization be 2 or more (i.e., $n \geq 2$) and suggested that the chains must be sufficiently spaced to enable water to penetrate the OEG layer, especially for the methoxy-terminated surfaces. These findings were summarized by stating that protein resistance requires both internal and external hydrophilicity. Similar findings for hydroxyl- and methoxy-terminated OEG, and other functional groups, have been reported by Whitesides and coworkers.^[27,28] Of particular interest, Ostuni et al. have reported values for the protein resistance provided by 48 different surface functional types covalently linked to SAM surfaces, many of which were very different from OEG, although still providing a high level of resistance to protein adsorption.^[27] In particular, several compounds with dimethyl amine ($-\text{N}(\text{CH}_3)_2$) and methoxy ($-\text{O}-\text{CH}_3$) end groups were shown to adsorb very low levels of protein compared to the other functionalities tested. Based on these findings, these authors stated that four common elements were present in each of the protein-resistance functionalities found in their studies: the functional groups 1) were hydrophilic; 2) included hydrogen bond acceptors; 3) did not include hydrogen bond donors; and 4) were neutrally charged. One immediate problem with this set of four necessary characteristics, however, is that condition (3) is violated by the most protein-resistant molecular structure tested, namely hydroxyl-terminated OEG. Thus, while these studies provide very interesting results, the requirement that protein-resistant functional groups not be hydrogen bond donors obviously cannot be universally applied. Another limitation to the above-described studies is that they again focus on functional group–water interactions and do not address the issue concerning why these interactions are thermodynamically preferred over interactions between the functional groups and the protein.

As indicated from their adsorption/desorption plots, the protein-resistant surfaces reported by Ostuni et al. represent systems exhibiting reversible adsorption behavior with low surface coverage, and thus the results from this study can be considered from a thermodynamic perspective with the approximate determination of the standard state free energy of adsorption ($\Delta G_{\text{ads}}^{\circ}$) calculated from Eq. (3) above.^[27] The data provided by Ostuni et al. are reported as the percent monolayer adsorption of protein (lysozyme and fibrinogen) from 1.0 mg/mL solutions, which, fortuitously, are the same example conditions provided in Table 1 of this entry. Accordingly, the $\Delta G_{\text{ads}}^{\circ}$ values in Table 1 indicate that under equilibrium conditions, to achieve 1% monolayer coverage or less requires that $\Delta G_{\text{ads}}^{\circ} \geq -3.0$ kcal/mol and $\Delta G_{\text{ads}}^{\circ} \geq -5.0$ kcal/mol for the adsorption of lysozyme and fibrinogen, respectively. Considering fibrinogen, which has a side-on molecular footprint of 235 nm^2 ($5.0 \text{ nm} \times 47 \text{ nm}$)

compared to 0.21 nm^2 per OEG surface group on a homogeneously functionalized SAM surface, an adsorbed fibrinogen molecule would potentially interact with over 1000 OEG chains.^[7,29] This thus averages out to $\Delta G_{\text{ads}}^{\circ} \geq -0.005$ kcal/mol per OEG chain, which is approximately stating that $\Delta G_{\text{ads}}^{\circ} \geq 0$ kcal/mol for this interaction. Considering the thermodynamic relationship of $\Delta G_{\text{ads}}^{\circ} = \Delta H_{\text{ads}}^{\circ} - T\Delta S_{\text{ads}}^{\circ}$, it is clear that to provide protein resistance, the interaction of a surface with a protein should thus ideally result in $\Delta H_{\text{ads}}^{\circ} \geq 0$ and $\Delta S_{\text{ads}}^{\circ} \leq 0$; or in other words, the unbound hydrated state of the surface functional group should be lower in enthalpy and higher in entropy than when bound to a protein. Considering enthalpy, this will be provided if the surface functional group is able to bond with water more favorably than with functional groups of the protein. Considering entropy, the functional group must be able to maintain a higher state of system entropy when bonded to water than when bonded to the protein. Of course, the water structure itself must also be considered, with the surface being able to maintain the surrounding water structure in a free-energy state that is equal to or lower than that of bulk water.

With this thermodynamic perspective, it is possible to interpret the results shown by Hernwerth et al. and Ostuni et al. somewhat differently, with a key feature being that water is much smaller and more mobile than the functional groups presented by a protein.^[26,27] Accordingly, if hydrogen bonding groups are presented by the surface molecules in a manner such that water can access the groups much more readily than can the functional groups of a protein, then a lower state of enthalpy will be provided for the hydrated, nonadsorbed state. Similarly, because of the mobility of water, water is able to maintain hydrogen bonding with a flexible surface chain without substantially restricting either the chain's or the water's configurational space, whereas the bonding of the protein to a flexible surface molecule will substantially restrict the motion of the surface chain, thus decreasing the entropy of the system relative to the nonbonded state. Accordingly, the longer the surface chain, the greater the entropic penalty will be for restricting its motion. These arguments are consistent with experimental evidence regarding the types of molecular structure that provide resistance to protein adsorption. Longer OEG chains that are not restricted by too high a packing density are more resistant to protein adsorption (entropic contribution), and chains that have well-hydrated hydrogen bonding groups that are relatively inaccessible to the protein, such as ($-\text{O}-\text{CH}_2-\text{CH}_2-\text{X}$) and ($-\text{N}(\text{CH}_3)_2$), are resistant to protein adsorption (enthalpic contribution). It is interesting to note that similar molecular features are found in phosphatidylcholine molecules presented by cell

membranes and 2-methacryloyloxyethyl phosphor-ylcholine molecules for surface functionalization that have been designed as a biomimetic of the cell surface, both of which provide a very high level of resistance to protein adsorption.^[2,30,31]

PROBING SUBMOLECULAR MECHANISMS OF PROTEIN ADSORPTION BEHAVIOR

While a great deal has been learned from studies focusing on the amount and type of protein that adsorbs to various types of surfaces, if surfaces are to be designed to actually control protein adsorption behavior, it will be necessary to develop a much more detailed understanding of the submolecular mechanisms involved in protein–surface interactions, and how these interactions influence the orientation, conformation, packing, and subsequent bioactivity of adsorbed proteins. Fortunately, several very exciting relatively new technologies are being developed and applied that should greatly advance the understanding of protein–surface interactions. Several of these technologies are highlighted in this section.

Drobny, Stayton, and coworkers are developing solid-state nuclear magnetic resonance (ssNMR) techniques for the study of peptide–surface interactions to enable the distance between isotopically labeled atom types to be measured with subangstrom resolution.^[32,33] By the careful selection of which atoms are labeled in their systems, they have been able to quantitatively determine the secondary structure of adsorbed peptides and to identify the types of peptide residues that specifically interact with functional groups of the surface for several different types of surfaces. This method thus enables specific interactions governing the structure of adsorbed peptides and their associated surface functional groups to be targeted and directly probed.

Circular dichroism spectropolarimetry (CD) has also been demonstrated by several groups to be able to quantitatively determine the secondary structure of adsorbed peptides and proteins, as well as provide information regarding the adsorbed tertiary structure for proteins containing tryptophan residues.^[34–37] In particular, the high sensitivity of this method has enabled its use to be extended to proteins adsorbed onto flat surfaces in addition to particles in solution, thus expanding its versatility to a much broader range of materials while also avoiding issues related to signal noise problems due to light scattering by dispersed solid particles. Recent studies by Hylton et al. have demonstrated that surface chemistry strongly influences the amount of adsorption-induced loss in the α -helix structure of both fibrinogen and albumin and, more importantly, platelet adhesion following protein

adsorption was found to directly correlate to the degree of structural change in the adsorbed layer of each protein.^[36] In another very interesting study, Engel et al. combined CD with stop-flow fluorescence spectroscopy and demonstrated that up to 70% of adsorption-induced unfolding of bovine α -lactalbumin (BLA) on polystyrene nanospheres in aqueous solution occurs within 15 msec of nanoparticle exposure, thus providing direct experimental measurement of the kinetics of the adsorption-induced unfolding process.^[35]

In a very interesting follow-up experimental study by Engel and coworkers, submolecular events related to adsorption-induced BLA unfolding were probed using hydrogen/deuterium (H/D) amide proton exchange and 2D NMR.^[38] This method provides an excellent technique to probe the effect of adsorption on protein structure because H/D exchange is directly related to the degree of solvent accessibility of a given peptide residue. When the H/D exchange behavior for the native BLA in solution was compared to adsorbed BLA, the investigators found distinct differences in H/D exchange before and after adsorption, thus providing direct insights into how BLA unfolds during adsorption by identifying the specific residues that transitioned from a buried to a solvent-exposed state following adsorption. The results of this study support the presence of distinctly different populations of adsorbed protein—one that exhibited substantial unfolding and another, smaller population with near-native-state H/D exchange behavior. This finding was believed to occur because the last fraction of adsorbed proteins was prevented from unfolding because of the physical restrictions provided by the previously adsorbed proteins, which is in agreement with the process illustrated in Fig. 3.

Castner and coworkers are developing a different experimental technique to probe adsorbed protein orientation and conformation using static time-of-flight secondary ion mass spectrometry (ToF-SIMS), which uses an ion beam to release molecular fragments from the top 1.0–1.5 nm of an adsorbed protein layer.^[39,40] Using multivariate analysis techniques, the released fragments are able to be associated with their parent peptide residue types.^[41] Thus, if the adsorbed protein layer is thicker than 1.5 nm (which is smaller than the smallest dimension of most clinically relevant proteins), and if the protein structure has an anisotropic distribution of peptide residues, the specific distribution of peptide residues present in the ToF-SIMS spectrum can be used to provide information relating to the orientation and conformation of the adsorbed protein layer. Castner and coworkers have already demonstrated that differences in both the type and conformational state of adsorbed proteins can be determined with this method, with continued research

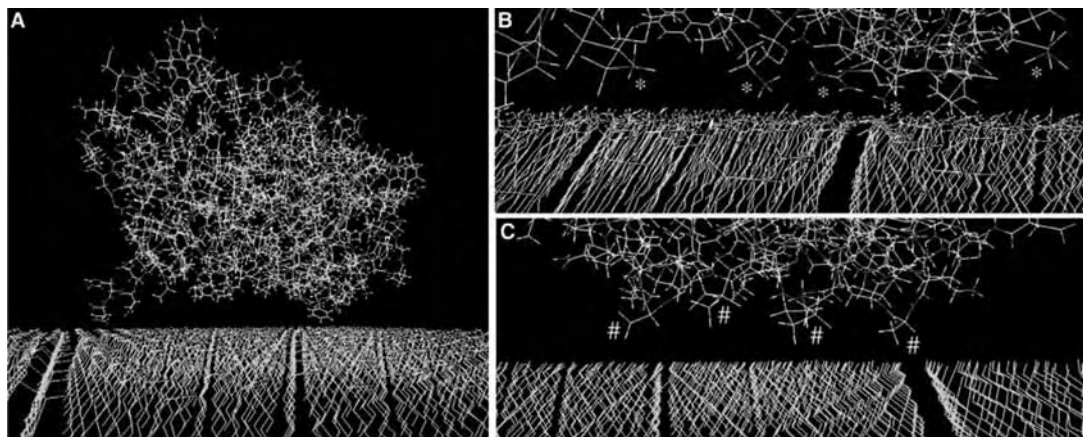


Fig. 6 (A) Molecular model of a 30 kDa segment from the C-terminus end of the γ -chain of fibrinogen (γ Fg, PDB ID 1FID) over a SAM surface in 150 mM saline. Water molecules and salt ions surrounding the protein and surface, which were present during the simulation, are not shown for clarity. (B, C) Close-up views of the resulting peptide–surface interactions observed during the 5.0 nsec molecular dynamics simulations for adsorption of γ Fg to a negatively charged COOH-SAM surface (B) and a hydrophobic CH₃-SAM surface (C). Note that positively charged peptide residues (e.g., lysine, arginine) preferentially interacted with the COOH-SAM surface (indicated by * in B) while hydrophobic peptide residues (e.g., alanine, leucine) preferentially interacted with the CH₃-SAM surface (indicated by # in C).

being conducted to further develop and evaluate its potential as a tool for the characterization of adsorbed protein structure.

Advances in biochemistry techniques have also greatly assisted in the study of protein adsorption behavior. As previously noted, extensive peptide residue mutation studies have been conducted with T4 lysozyme.^[42] McGuire and coworkers have used this versatile model protein system to elucidate relationships between protein stability, adsorption behavior, and adsorbed-state bioactivity.^[21,43] García and coworkers have conducted extensive studies with recombinant fragments of fibronectin, which is a clinically relevant ECM cell-adhesion protein with cell-binding domains recognized by various types of integrins displayed by fibroblasts.^[44,45] These studies have clearly demonstrated that surface chemistry substantially influences the adsorption behavior of the cell-binding domains of this protein, with subsequent changes in the integrin expression and adhesion strength of fibroblast cells. In another set of recombinant protein studies, Tang and coworkers have demonstrated that a specific bioactive site located in the C-terminus segment of the D-fragment of fibrinogen, which is not accessible in the soluble state of this protein, becomes exposed when fibrinogen adsorbs to a biomaterial surface, with this bioactive site serving as an important mediator of the inflammatory response.^[46]

In addition to experimental methods, substantial advances in computational hardware, software, and analysis methods have been made over the past few decades. These advances, coupled with the structural determination of thousands of proteins that are now

available via the Protein Data Bank (PDB), have now provided the capability to conduct all-atom molecular simulations to investigate the atomic level behavior of proteins in solution.^[47] While these capabilities have been widely developed and exploited for application to the study of protein folding and ligand–receptor interactions for rational drug design, they represent a largely untapped resource for assisting in the understanding of protein adsorption behavior on biomaterial surfaces.^[48] Latour and colleagues have been working over the past several years to develop the necessary methods to enable molecular modeling to be accurately applied to simulate protein–surface interactions.^[49,50] As an example from this work, Fig. 6 shows the interactions of a 30 kDa segment of fibrinogen with two different SAM surface chemistries.^[49] As illustrated, distinctly different types of peptide residue–surface interactions were predicted for each type of SAM surface, with subsequent hydrophobic, electrostatic, and hydrogen-bonding interactions being observed. In other studies, Jiang and coworkers and Kreuzer and coworkers have applied molecular simulation methods to theoretically investigate protein–surface and OEG–water interactions to gain further insight into the possible submolecular interactions responsible for these types of adsorption systems.^[51–54]

CONCLUSIONS

With the advent of tissue engineering and regenerative medicine, the goals of biomaterial surface design have

shifted from a focus on attempting to minimally disturb the surrounding biological processes to the design of surfaces that are able to direct them to attain a desired response. Although the desired type of response varies widely depending on the specific application, be it the incorporation of an implant without fibrous capsule formation or thrombosis, or the regeneration of a targeted tissue or organ, at the fundamental level the biological response will still largely be governed by the bioactive signals that are provided by the biomaterial surface.

It has now been widely accepted that protein adsorption behavior must be controlled to control the bioactivity of a biomaterial surface.^[55,56] For many applications, the complete prevention of protein adsorption may represent the optimal condition, thus eliminating the presentation of bioactive signals from the surface altogether. Proteins, however, are still the body's natural means of directing biological processes and thus, if properly harnessed, still possess the greatest potential to directly interact with natural cellular processes to direct cellular responses. The central issue still remains: how to control protein adsorption behavior in a manner to realize this potential. It is obvious that there is much more to be discovered and understood. The continued development of experimental methods to probe the submolecular mechanisms that control protein adsorption behavior, combined with the capabilities that can be provided by molecular simulation to theoretically assess and visualize protein–surface interactions at the atomistic level, should provide the means to continue the advancement of our understanding of protein adsorption behavior. Given the widespread recognition of its importance, the study of protein–surface interactions should continue to be of central importance for biomaterials development well into the foreseeable future as the biomaterials community works toward the goal of proactively designing surfaces that are able to directly control biological response.

ARTICLES OF FURTHER INTEREST

Biocompatibility Testing; Biomaterials Immune Response; Thrombosis

REFERENCES

- Brandon, C.; Tooze, J. The building blocks. In *Introduction to Protein Structure*, 2nd Ed.; Garland Publishing: New York, 1999; 3–12.
- Voet, D.; Voet, J.G.; Pratt, C.W. *Fundamentals of Biochemistry*; John Wiley & Sons Inc.: New York, 2002.
- Norde, W.; Haynes, C.A. Reversibility and the mechanism of protein adsorption. In *Proteins at Interfaces II. Fundamentals and Applications*; Horbett, T.A., Brash, J.L., Eds.; American Chemical Society: Washington, DC, 1995; 26–40.
- Wertz, C.F.; Santore, M.M. Fibrinogen adsorption on hydrophilic and hydrophobic surfaces: geometrical and energetic aspects of interfacial relaxations. *Langmuir* **2002**, *18*, 706–715.
- Adamson, A.W. The solid–liquid interface—adsorption from solution. In *Physical Chemistry of Surfaces*, 5th Ed.; John Wiley & Sons: New York, 1990; Chap. XI, 421–435.
- Latour, R.A.; Rini, C.J. Theoretical analysis of adsorption thermodynamics for hydrophobic peptide residues on SAM surfaces of varying functionality. *J. Biomed. Mater. Res.* **2002**, *60*, 564–577.
- Wertz, C.F.; Santore, M.M. Effect of surface hydrophobicity on adsorption and relaxation kinetics of albumin and fibrinogen: single-species and competitive behavior. *Langmuir* **2001**, *17*, 3006–3016.
- Talbot, J.; Tarjus, G.; Van Tassel, P.R.; Viot, P. From car parking to protein adsorption: an overview of sequential adsorption processes. *Colloids Surf. A.* **2000**, *165*, 287–324.
- Malmsten, M. Formation of adsorbed protein layers. *J. Colloid Interface Sci.* **1998**, *207*, 186–199.
- Wertz, C.F.; Santore, M.M. Adsorption and relaxation kinetics of albumin and fibrinogen on hydrophobic surface: single-species and competitive behavior. *Langmuir* **1999**, *15*, 8884–8894.
- Haynes, C.A.; Norde, W. Globular proteins at solid/liquid interfaces. *Colloids Surf. B* **1994**, *2*, 517–566.
- Nakanishi, K.; Sakiyama, T.; Imamura, K. On the adsorption of proteins on solid surfaces, a common but very complicated phenomenon. *J. Biosci. Bioeng.* **2001**, *91*, 233–244.
- Agnihotri, A.; Siedlecki, C.A. Time-dependent conformational changes in fibrinogen measured by atomic force microscopy. *Langmuir* **2004**, *20*, 8846–8852.
- Fournier, R.L. Solute transport in biological systems. In *Basic Transport Phenomena in Biomedical Engineering*; Taylor & Francis: Philadelphia, PA, 1999; 27–28.
- Dee, K.C.; Puleo, D.A.; Bizios, R. Protein–surface interactions. In *Tissue-Biomaterials Interactions*; John Wiley & Sons Inc.: Hoboken, NJ, 2002; Chap. 3, 45–49.
- Slack, S.M.; Horbett, T.A. The Vroman effect: a critical review. In *Proteins at Interfaces II. Fundamentals and Applications*; Horbett, T.A., Brash, J.L., Eds.; American Chemical Society: Washington, DC, 1995; 112–128.
- Vroman, L.; Adams, A.L. Findings with the recording ellipsometer suggesting rapid exchange of specific plasma proteins at liquid/solid interfaces. *Surf. Sci.* **1969**, *16*, 438–446.
- Turbill, P.; Beugeling, T.; Poot, A.A. Competitive adsorption of proteins during exposure of human blood plasma to glass and polyethylene. In *Proteins at Interfaces II. Fundamentals and Applications*; Horbett, T.A., Brash, J.L., Eds.; American Chemical Society: Washington, DC, 1995; 150–162.
- Warkentin, P.H.; Lundstrom, I.; Tengvall, P. Protein–protein interactions affecting proteins at surfaces. In

- Proteins at Interfaces II. Fundamentals and Applications*; Horbett, T.A., Brash, J.L., Eds.; American Chemical Society: Washington, DC, 1995; 163–180.
20. Jung, S.-Y.; Lim, S.-M.; Albertorio, F.; Kim, G.; Gurau, M.C.; Yang, R.D.; Holden, M.A.; Cremer, P.S. The Vroman effect: a molecular level description of fibrinogen displacement. *J. Am. Chem. Soc.* **2003**, *125*, 12782–12786.
 21. Lee, W.-K.; McGuire, J.; Bothwell, M.K. Competitive adsorption of bacteriophage T4 lysozyme stability variants at hydrophilic glass surfaces. *J. Colloid Interface Sci.* **2004**, *269*, 251–254.
 22. Morra, M. On the molecular basis of fouling resistance. *J. Biomater. Sci. Polym. Ed.* **2000**, *11*, 547–569.
 23. Kjellander, R.; Florin, E. Water structure and changes in thermal stability of the system poly(ethylene oxide)-water. *J. Chem. Soc. Far. Trans.* **1981**, *1*, 2053–2077.
 24. Prime, K.L.; Whitesides, G.M. Adsorption of proteins onto surfaces containing end-attached oligo(ethylene oxide): a model system using self-assembled monolayers. *J. Am. Chem. Soc.* **1993**, *115*, 10714–10721.
 25. Li, L.; Chen, S.; Zheng, J.; Ratner, B.D.; Jiang, S. Protein adsorption on oligo(ethylene glycol)-terminated alkanethiolate self-assembled monolayers: the molecular basis of nonfouling behavior. *J. Phys. Chem. B.* **2005**, *109*, 2934–2941.
 26. Herrwerth, S.; Eck, W.; Reinhardt, S.; Grunze, M. Factors that determine the protein resistance of oligoether self-assembled monolayers—internal hydrophilicity, terminal hydrophilicity, and lateral packing density. *J. Am. Chem. Soc.* **2003**, *125*, 9359–9366.
 27. Ostuni, E.; Chapman, R.G.; Holmlin, R.E.; Takayama, S.; Whitesides, G.M. A survey of structure–property relationships of surfaces that resist the adsorption of protein. *Langmuir* **2001**, *17*, 5605–5620.
 28. Chapman, R.G.; Ostuni, E.; Yan, L.; Whitesides, G.M. Preparation of mixed self-assembled monolayers (SAMs) that resist adsorption of proteins using the reaction of amines with a SAM that presents inter-chain carboxylic anhydride groups. *Langmuir* **2000**, *16*, 6927–6936.
 29. Ulman, A.; Eilers, J.E.; Tillman, N. Packing and molecular orientation of alkanethiol monolayers on gold surfaces. *Langmuir* **1989**, *5*, 1147–1152.
 30. Ishihara, K.; Iwasaki, Y. Reduced protein adsorption on novel phospholipid polymers. *J. Biomater. Appl.* **1998**, *13*, 111–127.
 31. Feng, W.; Brash, J.; Zhu, S. Atom-transfer radical grafting polymerization of 2-methacryloyloxyethyl phosphorylcholine from silicon wafer surfaces. *J. Polym. Sci. Pt. A: Polym. Chem.* **2004**, *42*, 2931–2942.
 32. Shaw, W.J.; Long, J.R.; Campbell, A.A.; Stayton, P.S.; Drobny, G.P. A solid-state NMR study of dynamics in a hydrated salivary peptide adsorbed to hydroxyapatite. *J. Am. Chem. Soc.* **2000**, *122*, 7118–7119.
 33. Bower, P.V.; Louie, E.A.; Long, J.R.; Stayton, P.S.; Drobny, G.P. Solid-state NMR structural studies of peptides immobilized on gold nanoparticles. *Langmuir* **2005**, *21*, 3002–3007.
 34. Vermeer, A.W.P.; Norde, W. CD spectroscopy of proteins adsorbed at flat hydrophilic quartz and hydrophobic Teflon surfaces. *J. Colloid Interface Sci.* **2000**, *225*, 394–397.
 35. Engel, M.F.M.; van Mierlo, C.P.M.; Visser, A.J.W.G. Kinetic and structural characterization of adsorption-induced unfolding of bovine α -lactalbumin. *J. Biol. Chem.* **2002**, *277*, 10,922–10,930.
 36. Hylton, D.M.; Shalaby, S.W.; Latour, R.A. Direct correlation between adsorption-induced changes in protein structure and platelet adhesion. *J. Biomed. Mater. Res.* **2005**, *73A*, 349–358.
 37. Lundqvist, M.; Sethson, I.; Jonsson, B.-H. Protein adsorption onto silica nanoparticles: conformational changes depend on the particles' curvature and the protein stability. *Langmuir* **2004**, *20*, 10639–10647.
 38. Engel, M.F.M.; Visser, A.J.W.G.; van Mierlo, C.P.M. Conformation and orientation of a protein folding intermediate trapped by adsorption. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 11,316–11,321.
 39. Tidwell, C.D.; Castner, D.G.; Golledge, S.L.; Ratner, B.D.; Meyer, K.; Hagenhoff, B.; Benninghoven, A. Static time-of-flight secondary ion mass spectrometry and x-ray photoelectron spectroscopy characterization of adsorbed albumin and fibronectin films. *Surf. Interface Anal.* **2001**, *31*, 724–733.
 40. Wagner, M.S.; Castner, D.G. Analysis of adsorbed proteins by static time-of-flight secondary ion mass spectrometry. *Appl. Surf. Sci.* **2004**, *231–232*, 366–376.
 41. Wagner, M.S.; Graham, D.J.; Ratner, B.D.; Castner, D.G. Maximizing information obtained from secondary ion mass spectra of organic thin films using multivariate analysis. *Surf. Sci.* **2004**, *570*, 78–97.
 42. Matthews, B.W. Studies on protein stability with T4 lysozyme. *Adv. Prot. Chem.* **1995**, *46*, 249–278.
 43. Bower, C.K.; Xu, Q.; McGuire, J. Activity losses among T4 lysozyme variants after adsorption to colloidal silica. *Biotechnol. Bioeng.* **1998**, *58*, 658–662.
 44. Michael, K.E.; Keselowsky, B.G.; Vernekar, V.N.; Meredith, C.A.; Latour, R.A.; Garcia, A.J. Adsorption-induced conformational changes in fibronectin due to interactions with well-defined surface chemistries. *Langmuir* **2003**, *19*, 8033–8040.
 45. Keselowsky, B.G.; Collard, D.M.; Garcia, A.J. Surface chemistry modulates fibronectin conformation and directs integrin binding and specificity to control cell adhesion. *J. Biomed. Mater. Res.* **2003**, *66A*, 247–259.
 46. Hu, W.J.; Eaton, J.W.; Tang, L. Molecular basis of biomaterial-mediated foreign body reactions. *Blood* **2001**, *98*, 1231–1238.
 47. <http://www.pdb.org/>.
 48. Latour, R.A. Molecular modeling of biomaterial surfaces. *Curr. Opin. Solid State Mater. Sci.* **1999**, *4*, 413–417.
 49. Agashe, M.; Raut, V.; Stuart, S.J.; Latour, R.A. Molecular simulation to characterize the adsorption behavior of a fibrinogen gamma-chain fragment. *Langmuir* **2005**, *21*, 1103–1117.
 50. Sun, Y.; Latour, R.A. Prediction of the orientations of adsorbed protein using an empirical energy function with implicit solvation. *Langmuir* **2005**, *21*, 5616–5626.
 51. Zheng, J.; Li, L.; Chen, S.; Jiang, S. Molecular simulation study of water interactions with oligo(ethylene

- glycol)-terminated alkanethiol self-assembled monolayers. *Langmuir* **2004**, *20*, 8931–8938.
52. Zhou, J.; Zheng, J.; Jiang, S.Y. Molecular simulation studies of the orientation and conformation of cytochrome c adsorbed on self-assembled monolayers. *J. Phys. Chem. B* **2004**, *108*, 17,418–17,424.
 53. Kreuzer, H.J.; Wang, R.L.C.; Grunze, M. Effect of stretching on the molecular conformation of oligo(ethylene oxide): a theoretical study. *New J. Phys.* **1999**, *1*, 21.1–21.16.
 54. Wang, R.L.C.; Kreuzer, H.J.; Grunze, M. The interaction of oligo(ethylene oxide) with water: a quantum mechanical study. *Phys. Chem. Chem. Phys.* **2000**, *2*, 3613–3622.
 55. Castner, D.G.; Ratner, B.D. Biomedical surface science: foundations to frontiers. *Surf. Sci.* **2002**, *500*, 28–60.
 56. Ratner, B.D.; Bryant, S.J. Biomaterials: where we have been and where we are going. *Annu. Rev. Biomed. Eng.* **2004**, *6*, 41–75.
- FURTHER READINGS**
- Brash, J.L.; Horbett, T.A., Eds. *Proteins at Interfaces. Physicochemical and Biochemical Studies*; American Chemical Society: Washington, DC, 1987.
- Brash, J.L.; Wojciechowski, P.W., Eds. *Interfacial Phenomena and Bioproducts*; Marcel Dekker, Inc.: New York, 1996.
- Dee, K.C.; Puleo, D.A.; Bizios, R. Protein–surface interactions. In *Tissue–Biomaterial Interactions*; John Wiley & Sons: Hoboken, NJ, 1999; 37–52.
- Gray, J.J. The interaction of proteins with solid surfaces. *Curr. Opin. Struct. Biol.* **2004**, *14*, 110–115.
- Hlady, V.; Buijs, J. Protein adsorption on solid surfaces. *Curr. Opin. Biotechnol.* **1996**, *7*, 72–77.
- Hoffman, A.S. Non-fouling surface technologies. *J. Biomater. Sci. Polym. Ed.* **1999**, *10*, 1011–1014.
- Horbett, T.A. The role of adsorbed proteins in tissue response to biomaterials. In *Biomaterials Science—An Introduction to Materials in Medicine*, 2nd Ed.; Ratner, B.D., Hoffman, A.S., Schoen, F.J., Lemons, J.E., Eds.; Elsevier Academic Press: New York, 2004; 237–246.
- Horbett, T.A.; Brash, J.L., Eds. *Proteins at Interfaces II. Fundamentals and Applications*; American Chemical Society: Washington, DC, 1995.
- Malmsten, M. *Biopolymers at Interfaces*; Marcel Dekker, Inc.: New York, 1998.
- Ramsden, J.J. Puzzles and paradoxes in protein adsorption. *Chem. Soc. Rev.* **1995**, *24*, 73–78.

Biomedical Ethics

Daniel E. Wueste

Robert J. Rutland Center for Ethics, Clemson University, Clemson, South Carolina, U.S.A.

B

INTRODUCTION

The purpose of this article is to survey and discuss the salient features of biomedical ethics in a way that makes its value apparent and its methodology accessible to persons whose interest in it probably arises from involvement in biomedical practice rather than familiarity with ethical theory.

Although biomedical ethics is a relative newcomer to the field of applied or practical ethics, an enormous amount of ink has been put to paper in discussions of issues in the field, alternative ways of approaching them, and even whether any of the alternative approaches is actually viable. The literature is, in a word, vast. Readers are most cordially invited, if they choose, to dive into that literature. Here, however, our focus has to be narrow. Accordingly, after a brief discussion of what brought the newcomer onto the field, we will focus on the principlist approach in biomedical ethics. As it happens, this approach emerged at the start and, although it now has competitors, it remains the standard approach among practitioners and teachers of future healthcare professionals.

The key elements of the principlist approach are the principles of autonomy, nonmaleficence, beneficence, and justice. We will examine each of these principles in some detail to reveal the character of each, how they differ from one another, and how they can be applied in practice. Along the way, we will clarify the notion of applied or practical ethics; distinguish two sorts of ethical standards and explain how they are related to one another; and explain the distinction between absolute and prima facie principles, which is enormously important to the principlist program. We will conclude with an example that brings us back to a point made at the beginning and helps to underscore the necessity for judgment that emerges from systematic, reflective, and responsible engagement with the pressing issues of biomedical ethics.

IN THE BEGINNING

Biomedical ethics is a branch of applied or practical ethics. Like its cousins, engineering ethics, business ethics, and environmental ethics, it focuses on and provides tools for the systematic resolution of ethical

problems that arise in practice. By most accounts, it was born in the 1970s in response to significant and rapid changes in medicine, particularly those tied to advances in research or emerging technologies. Changes in the relationship between physician and patient and the contexts of medical practice and research were also influential. (At that time, the essentially paternalistic character of the doctor/patient relationship came into question and the autonomy of medical professionals began to diminish as the practice of medicine became institutionally entrenched.) However, the most important factor was a medical and scientific scandal revealed in the *Washington Evening Star* on July 25, 1972 under this headline: “Syphilis Patients Died Untreated.” The story, written by Jean Heller of the Associated Press, reported on a US Public Health Service (PHS) study, begun in 1932, “in which human guinea pigs, not given proper treatment, have died of syphilis and its side effects.” Heller reported that the aim of the study, which at that time had been going on for 40 years, was “to determine from autopsies what the disease does to the human body.”^[1] The next day, the story ran on the front page of the *New York Times* with this headline: “Syphilis Victims in US Study Went Untreated for 40 Years,” and the world learned that this study, which has come to be known as The Tuskegee Syphilis Study, was “the longest nontherapeutic experiment on human beings in medical history.”^[2] The PHS study, conducted in Macon County Alabama, involved 600 African-American males, 399 who were syphilitic and 201 who were free of the disease (as control). From the very beginning, the subjects of the experiment were not aware that they were part of a PHS study, let alone what its aim was or what alternatives were open to them.^[3] Like a bacterium subjected to intense heat, the study could not survive for long in the heat and light generated by publicity. It ended on November 16, 1972.

The official name of the study, “Tuskegee Study of Untreated Syphilis in the Negro Male,” reveals a good bit of what was scandalous in it, for although there was no effective treatment for syphilis when the study began, one became available in 1943—penicillin—and it was deliberately withheld from the subjects. In fact, the PHS worked hard to make sure that syphilitic study subjects did not receive treatment. For example, during World War II when these men were subject to

the draft, because draftees with syphilis were required to undergo treatment, the PHS provided the Draft Board of Macon County Alabama with the names of 256 study participants under the age of 45, and formally requested that they be excluded from the required treatment. The Board complied with the request. Similarly, from 1952 until the study ended, local health departments helped the PHS both keep track of study participants who left Macon County and prevent them from receiving treatment.^[3] According to the timeline of the study at the web site of the Centers for Disease Control (CDC), questions about the ethics of the study were raised in 1968.^[4] However, the CDC confirmed the need for the study in 1969; the local chapters of the American Medical Association and the National Medical Association supported its continuation as well. Thus, for 29 years, study participants, *none* of whom had been told that they were infected with syphilis—they had been told that they had “bad blood”—were denied treatment for a disease that could have been effectively treated. Moreover, as a result of the deception and denial of treatment, their wives and lovers were exposed to the disease; when the study ended, 40 wives had been infected and 19 children had congenital syphilis.^[5]

There are many troubling aspects to this sad story. For instance, Alabama law (1927) required reporting and treatment of syphilitic individuals and the Henderson Act (1943) required testing and treatment for venereal disease. In addition, according to the first principle of the Nuremberg Code (1946), voluntary informed consent “is absolutely essential” in experiments with human beings; the Helsinki Declaration of the World Health Organization (1964) makes the same claim. Yet, so far as this study is concerned, such requirements came to nothing. Published articles reporting on the study betray nothing in the way of ethical misgivings. For example, a report in the December 1964 issue of the *Archives of Internal Medicine* marked the 30th year of the study noting 1) that because many of the younger participants had received some treatment, “14 young, untreated syphilitics were added to the study to compensate” and 2) that this “group continues to have higher mortality.”^[6] Indeed, it appears that John Heller, director of the Venereal Diseases Unit of the PHS from 1943 to 1948, summed up the thinking of those involved with the study when, in an interview in 1976, he said that the “men’s status did not warrant ethical debate. They were subjects, not patients; clinical material, not sick people.”^[7]

THE NATIONAL RESEARCH ACT OF 1974

The literature about the Tuskegee study is vast. There is much to ponder and explain in this sad chapter in the

history of biomedicine; there are lessons to be learned and *relearned* as well.^[8] What is most important for the present purpose, however, is the federal legislation enacted in response to the revelation of what, in the name of science, the PHS had been doing in Macon County Alabama for 40 years. The National Research Act of 1974 created the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. Prominent among the tasks undertaken by the Commission was the identification of “basic ethical principles that should underlie the conduct of biomedical and behavioral research involving human subjects” and the development of “guidelines which should be followed to assure that such research is conducted in accordance with those principles.”^[9]

The Commission started its work in December 1974. Its first report, *Research on the Fetus: Report and Recommendations*, was submitted to the Department of Health Education and Welfare (DHEW), Congress and the President in July 1975. The Commission submitted nine more reports. These reports dealt with a variety of issues: 1) research on prisoners (1976); 2) psychosurgery (1977); 3) disclosure of research information under the Freedom of Information Act (1977); and 4) research on children (1977). Five of the reports were submitted in 1978. The topics were: 1) research on those institutionalized as mentally infirm; 2) institutional review boards; 3) ethical guidelines for the delivery of health services by DHEW; 4) the Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research; and 5) implications of advances in biomedical and behavioral research.^[10]

The best known in this series of formal reports is the *Belmont Report*. This report differs from the others in not making specific recommendations for administrative action. In this case, the Commission recommended adoption of the report in its entirety, as a statement of policy.^[9] As this report articulates the basic ethical principles and guidelines the Commission had been charged with identifying, its recommendation was, in effect, that these principles and guidelines be an official policy of the federal agencies that fund and oversee research involving human subjects, which is, of course, a rather large portion of the biomedical research undertaken in the U.S.A. What the Commission proposed is, in effect, the case:

Title 45 Code of Federal Regulations, Part 46, Protection of Human Subjects (45 CFR Part 46), embodies the ethical principles of *The Belmont Report*. These regulations apply to all research involving human subjects conducted or supported by the Intramural Research Program (IRP) of the NIH ... The NIH IRP operates under an OHRP [Office for Human

Research Protections] approved Federal Wide Assurance (FWA). The FWA obligates the NIH to conduct all its research activities involving human subjects consistent with the ethical principles of *The Belmont Report*.^[11]

While the Commission was doing its work, Tom L. Beauchamp and James F. Childress were writing a groundbreaking book, *Principles of Biomedical Ethics*. First published in 1979, the book was written “almost coextensively” with the Commission’s work. Indeed, Beauchamp “served as staff philosopher”^[12] for the Commission. Like the Commission in the *Belmont Report*, Beauchamp and Childress present a few basic principles (they offer four principles, the Commission offers three) that can be used by healthcare professionals to work their way through ethical difficulties. It is fairly clear, then, why this is called applied ethics. The approach is also called “principlism;” it has proved to be very popular among healthcare professionals. It is, in fact, the best-known approach in biomedical ethics constituting what can be called, without hesitation, the received or standard approach among practitioners and the teachers of future healthcare professionals. To be sure, as even a cursory survey of the vast literature in biomedical ethics reveals, there are other approaches. For the most part, however, they have emerged in response to perceived weaknesses in the principlist approach and have not (yet) succeeded in usurping the throne of principlism. We will take a look at the principles put forward by the Commission as well the principles championed by Beauchamp and Childress after a few preliminaries about applied ethics.

APPLIED OR PRACTICAL ETHICS

As noted above, biomedical ethics is a branch of applied or practical ethics. Here, as with other varieties of applied ethics, the task is *normative* rather than descriptive; the focus of attention is on how things *ought* to be. In this, we find a sharp contrast with the scientific enterprise as it is commonly understood. For on this conception of science, the scientist’s task is to discover and report the way the world is. Indeed, on some accounts, each scientist is contributing to what will one day be a *complete description* of the way the world is. In any case, the distinction between normative and descriptive is fundamental. When, for example, a social scientist reports that in fact people make promises planning not to keep them, or that, say, studies reveal that 75% of undergraduates cheat on exams or make up data for lab reports, the ethicist is likely to remark that these are *sad* facts; it *ought* to be otherwise. Her response is evaluative and prescriptive.

Of course, in making such judgments, she is applying a standard. Her criterion may be more or less general. It may take the form of a principle or a rule, for example. But either way, and necessarily, a standard is being applied. If it is an ethical standard, then 1) it will be based on impartial considerations favoring neither the interests of an agent/actor nor those of an individual who is making an ethical judgment; 2) it will be general in scope, or put another way, it will be universal in the sense of being applicable in all relevantly similar situations; 3) it will be practicable, which is another way of saying that ought implies can; 4) it will be addressed to serious rather than frivolous matters; and 5) it will be peremptory or overriding, i.e., poised to carry the day in a conflict with another normative standard, for example, a law.

Consider the recent scandals in business—Enron, WorldCom, or Adelphia, for instance. Recognizing the ethical wrongdoing in these cases involves the application of ethical standards. The standards we need are not arcane; what was done in these cases violated basic and familiar ethical standards readily available for application not only in business but also in the professions, government, and the church, for example. It may be that the actors in these cases believed sincerely that they were not doing anything wrong. Indeed, they may have believed, again sincerely, that the relevant standards did not apply to them in their work. Claims of this sort are in fact common in such cases. However, their exculpatory force is illusory. This is so for three reasons. First, concerted efforts to make one’s dealings opaque suggest collusion rather than innocence. (This is true generally.) Second, it is far from clear why the occupations of the actors in these cases would carry a moral exemption; indeed, the idea is so mysterious that such a belief seems explicable only in terms of its convenience. Third, and most important, so far as judging actions is concerned, the sincerity of an actor’s beliefs about their rightness or wrongness cannot itself be decisive, for surely, thinking or believing something is so, however sincerely, does not make it so. A rapist may sincerely believe that what he is doing is perfectly okay; nevertheless, what he believes is false. What he is doing is wrong.

TWO TYPES OF ETHICAL STANDARDS: PRINCIPLES AND RULES

At this juncture we should pause to consider how, if we were challenged to do so, we could make the case for saying that rape is wrong. To be sure, it is illegal, and if the rapist is caught and convicted, he will be punished. But rape is not wrong because it is illegal and carries a heavy sanction. That is the wrong way around—if we say this, we’ve put the cart before the horse. Similarly,

while it is true that customary morality condemns rape, it is not the case that rape is wrong because customary morality condemns it. Again, law and customary morality condemn rape; each includes a rule prohibiting it. However, the rule itself acknowledges or recognizes but does not explain the wrongness of what it prohibits. The challenge we started with is not answered by an appeal to law or customary morality. Rather, it is recast as a question about the rules: Why should we have a rule prohibiting rape? More precisely, what's called for is a principled response to the challenge, which is precisely what will emerge as we articulate the moral considerations that justify this rule (and others as well).

A principle resides at a higher level of abstraction than rules. Rules characteristically require or prohibit actions. Consequently, they have more specificity than principles, which have a different normative function. The relationship between principle and rules is rather like that between a constitution and laws made in the exercise of the powers it bestows. It is, then, a relationship of authorization/justification. Questions about the validity of rules—for instance, about their moral necessity or warrant—are answered by reference to principles. Thus, as we reflect and respond to the question above about the reason(s) for the rule, we will be articulating a relevant principle.

1. Interactions with others should be beneficial, not detrimental. It is obvious that rape entails harm, physical and/or psychological, and this is surely the first thing that comes to mind, if one is asked to defend the judgment that rape is wrong. It is also clearly relevant to the justification of a rule prohibiting rape: Such a rule has the aim of preventing harm. The principle that is in play here is something like this: *It is wrong to willfully harm another person without justification; we ought to make things better for others, not worse; at the very least we have a duty not to harm others.*
2. Another way of making the case for the wrongness of rape, as well as a case in favor of a rule against it, points in a different direction. The rapist treats his victim as a thing, as a mere means to his end of sexual gratification (a rapist may have other ends, for instance, to establish, at least in his own mind, his dominance over women). In treating her this way, he denies her human dignity. He fails to treat her as she *deserves* to be treated; she has a right to be self-governing rather than controlled by another; she has a right to control over her body. Each of us thinks this is true in our case and there is no good reason to suppose that it is any less true of others. The rapist violates this basic right. The principle in play here is something like this: *Persons must be treated as self-governing*

individuals: Their capacity for self-direction must be respected; unlike things, persons are never properly treated as a mere means to an end.

The principles that emerge here speak in terms of consequences and deserved treatment, respectively. The first principle focuses on consequences. It approves of good consequences (i.e., those that are beneficial) and disapproves of bad consequences (i.e., those that are detrimental). Indeed, it speaks of a duty not to harm others, which appears to be where we start. Its concept of positive duty is tied to consequences as well: We ought to help others; the aim is to increase the amount of good in the world. The second principle embraces the idea of human dignity, which is understood in terms of self-governance, and demands respect for persons, in particular, that persons never be treated as if they were things. These principles differ from one another in a fundamental way. One suggests that questions of duty, right and wrong, are to be answered by careful consideration of results, the other focuses on something altogether different. Demanding respect for persons in the name of human dignity, this principle declines the invitation to consider results in judging what is right or wrong. Its claim is not that one should respect persons if or because doing so will have good consequences. Respecting persons is a duty regardless of the consequences. It may be true that respecting persons has good consequences and failing to respect persons has bad consequences. Indeed, the correlation may be very strong. However, an act is not right only because it has good consequences (nor wrong only because it has bad consequences). One acts rightly by respecting persons even when doing so has unhappy consequences. For example, because we have a duty to respect persons, it is necessary to obtain informed consent from research subjects. When research is conducted on human beings without their consent, the research subjects are being treated as a mere means to an end (the successful completion of the study). They are being treated like things rather than persons; their capacity for self-governance is not being respected. This is wrong *even if* the results of the study would be of tremendous value. In fact, so far as this principle is concerned, the results—good or bad—are morally irrelevant; ethically they make no difference at all.

It should be clear that the Tuskegee study violated this principle. It may be less obvious, but it is equally true, that the first principle was violated as well. Because the syphilitic study participants were not told that they had syphilis and because they were denied treatment (and were kept from receiving treatment as inductees during World War II), it was not only foreseeable, it was a virtual certainty that some of the wives and children of the syphilitic study participants would

contract the disease. And as noted earlier, this did in fact happen. Moreover, it strains credulity to the limit to suppose that no one could foresee that if the truth about the study were to become known in the African-American community, the trust essential to effective medical care and public health programs would be seriously eroded. These foreseeable consequences are very bad. Perhaps they are not weighty enough by themselves to justify the claim that the first principle was violated. However, combined with the quite substantial harm to the study participants themselves, the claim that the principle was violated can scarcely be doubted.

PRINCIPLIST PRINCIPLES

Earlier we noted that *The Belmont Report* advances three principles and that in their book, *Principles of Biomedical Ethics*, which was written virtually coextensively with the work of the Commission, Beauchamp and Childress put forward four principles. The first two principles have been the subject of the preceding discussion, though we have not used the names given to them in these two publications and did not arrive at them by the same means. In *The Belmont Report*, the second principle above is called the principle of respect for persons and the first one is called the principle of beneficence. Beauchamp and Childress call the second principle the principle of autonomy and, seeing two distinct aspects in the first principle, they speak of two principles: the principle of beneficence and the principle of nonmaleficence. The remaining principle, discussed below, is called the principle of justice in both places.

CONSEQUENTIALIST PRINCIPLES

Let us take a closer look at the first principle discussed above. The key feature of this principle is that it directs attention to the consequences of actions and indicates that actions are to be morally evaluated on the basis of the overall character of those consequences, that is, the key question is whether the balance of good and bad (benefit and cost) weighs heavier for good or ill. We are to approve of the former and disapprove of the latter. Because it focuses on consequences this principle is aptly categorized as a consequentialist principle. The best-known consequentialist principle is the principle of utility championed by the English utilitarians Jeremy Bentham and John Stuart Mill. The utilitarians taught that one acts rightly when one acts so as to produce the greatest good for the greatest number. Put another way, the criterion for choosing among available courses of action dictates the choice that is likely

to produce the largest benefit with the least in the way of cost for all those who have a stake in the matter (the stakeholders). The utilitarians were realistic. They recognized and in fact emphasized that the choices we make involve a mix of benefit and detriment (i.e., disadvantage or harm). Thus, they were frank in admitting the need to weigh and balance the likely benefits and detriments. However, when one thinks about rape, for example, it seems reasonable to say that the “benefit” the rapist enjoys should not figure in our judgment of his act (so too, *a fortiori*, any benefits that might accrue to his friends or fellow travelers). With this move, one rejects as ludicrous the suggestion that we must weigh all the benefits and all of the costs leaving open the possibility that it could turn out, on balance, that having weighed all of the benefits and costs, the rapist’s action was right. It may be true that we have a duty to maximize benefits—a *duty of beneficence*.^[9] However, we also have a duty not to harm others and this duty takes precedence; it is the first order of business: first, no harm (*primum non nocere*).

The primacy of a duty not to harm seems straightforward to many. For example, H.L.A. Hart observes that “the most common requirements of law and morality” are negative in form, and the most important of these prohibitions “restrict the use of violence in killing or inflicting bodily harm.” Hart suggests that the sense in which a duty not to harm is fundamental emerges with a question: without it, he asks, “what point could there be for beings such as ourselves in having [duties] of *any* other kind?”^[13] Like Hart, Bernard Gert holds that rules that prohibit harming others are quite basic. Indeed, Gert maintains that such rules, which prohibit killing, causing pain, and disabling, as well as depriving others of freedom, opportunity, or pleasure, constitute the core of morality.^[14]

In any case, thinking along lines such as these, one may, as Beauchamp and Childress do, distinguish two consequentialist principles, introducing, in addition to the principle of beneficence, the principle that entails our duty not to harm, namely, the principle of non-maleficence. With this move, it is very clear that consequentialist considerations entail two sorts of moral requirements. One variety is negative: We are duty-bound *not to act* in ways that inflict harm (non-maleficence). The other is positive: We are duty-bound *to act* beneficently; at the end of the day, as it were, as a result of our actions, there ought to be more good in the world. This result might be accomplished by creating value—doing or promoting good. It might also be accomplished by preventing or removing evil or harm (what is bad). As it happens, it is much easier to accept the latter, preventing or removing what is bad, as *duties*. Indeed, it might be said that because

promoting good is not morally incumbent on any and every moral agent independent of social/institutional role, it is not a “duty in the strict sense.”^[15] In any case, should these duties conflict, it is clear that the duty not to harm takes precedence over a duty to benefit others.

TWO POINTS OF SPECIAL INTEREST

Two points of special interest emerge here. First, prominent among the things that do not count as “duties in the strict sense” are the responsibilities individuals have in virtue of roles or relations, for instance, the responsibilities of a physician or a father. While it is probably true there is no duty in the strict sense to benefit others, it is clear that some roles require beneficent acts. For example, in taking the Hippocratic Oath a physician vows that “Into as many houses as I may enter, I will go for the benefit of the ill,” and again, that “I will use regimens for the benefit of the ill in accordance with my ability and my judgment, but from [what is] to their harm or injustice I will keep [them].” (This last, interestingly, captures the thrust of both the principle of beneficence and the principle of nonmaleficence and in the same voice.)^[16]

The second point is that these principles give rise to “prima facie duties.” This is a very important point, for among other things, it holds with respect to all four principles in the principlist approach championed by Beauchamp and Childress. Consider the duty not to lie and the duty not to kill. On one interpretation, the moral point of the claim that we have these duties is that lying and killing are wrong always and everywhere for anyone. Or again, the point is that one must never lie or kill, where “never” means never; under no circumstances is it permissible to lie or kill. Understood in this way, the duties not to lie and not to kill are *absolute* duties. That is to say, they cannot be overridden. Thus, for example, it would be wrong for a doctor to answer a young patient’s question, “Am I going to die?” with anything but the truth, even if the child’s parents had requested that the child not be told or there were good reason to expect that a truthful answer would be injurious. Similarly, it would be morally wrong for a police sniper to kill a terrorist to protect the hostages he was holding. For, as in the lying example, the good achieved makes no difference morally. In either case, thinking otherwise involves the (alleged) mistake of thinking that there are some circumstances in which the duty can be overridden—to achieve some good or prevent some evil, for example. Understanding duties in this way guarantees that they are taken seriously. That is a good thing. But a good thing can be taken too far. The appeal of the notion of prima facie duty is that it allows us to take the notion

of duty seriously without the embarrassment of having to say something that smacks of absurdity, for instance, that the good that could be achieved by overriding a duty makes no difference morally. Where the absolutist conception of duty announces in advance, independent of experience in the situation, that this is so, the prima facie conception of duty says that we cannot know this a priori. To say that one has a prima facie duty not to lie, for example, is to say that there is a presumption in favor of this moral prohibition. That presumption can be overcome—the duty can be overridden—but the justificatory burden for an override is heavy, rather like the burden imposed on a prosecutor by the presumption of innocence in criminal law. That presumption is overcome when, using evidence and argument, a prosecutor eliminates any reasonable doubt on the part of jurors that the defendant is guilty. Similarly here, the case one makes for overriding a prima facie duty has to be strong enough to withstand critical scrutiny by reasonable men and women (e.g., one’s professional peers). Convincing oneself is not enough; the conflict of interest is patent: One ought not be judge in one’s own case.

A DEONTOLOGICAL PRINCIPLE

In *The Belmont Report*, the second principle above is called the principle of respect for persons; Beauchamp and Childress call it the principle of autonomy. The latter is the more common name for the principle. As noted earlier, the key difference between this principle and the first one is that the principle of autonomy does not invite and in fact eschews consideration of consequences. Its demand that human dignity be respected is not based on a forecast of good results, if it occurs. That respecting persons has good consequences is a happy coincidence rather than a reason for judging it right. Thus, in the interest of human dignity, those whose capacity for autonomous action is limited (e.g., children, the infirm) are *entitled* to protection.

This mode of ethical thinking, traditionally called deontological, makes a claim similar to the one made in the American *Declaration of Independence*, which speaks of rights that persons have independent of government, that are, in fact, the reason that governments are established (governments are established to secure and protect them). Rights created or given by government could be as easily destroyed or taken away, presumably on the basis of some calculations about advantage (for the king or the people, for example). The “must” associated with the right/duty relationship here is conditional, i.e., it depends on the advantage generated by its being respected. Deontological thinking denies this contingency; its “must” is

unconditional. Thus, the “must” of respect for persons has nothing to do with efficiency in the pursuit of a goal and the gain it promises. An idea found in the moral philosophy of Immanuel Kant may be helpful. Kant taught that a person has an intrinsic rather than an instrumental value (dignity instead of a price). Accordingly, a person must be treated as an end and never as a mere means to an end. When a person is treated as a mere means to an end, human dignity is denied; the person is treated as a thing. This is a very powerful and appealing idea. Indeed, in the form of a doctrine of natural rights it has inspired men to revolution.

In much the same spirit that the duty not to harm was said to have primacy in consequentialist thinking, it has been argued that the principle of autonomy must carry the day in any conflict with a consequentialist principle (i.e., be regarded as absolute). However, in biomedical ethics, no principle has pride of place; all four are prima facie principles, each one establishes a strong presumption in favor of the duties it entails. In each case, the considerations that argue for the principle are weighty. Yet, in each case, particularly when the principle is brought to bear in context, there are things that make one pause. If these principles were treated as absolute, these difficulties would likely derail any attempt to apply them. For example, in the case of nonmaleficence, does an amputation necessary to save a life constitute prohibited harm? If not, why not? If the answer is that it does not count as prohibited harm because a greater good is achieved by means of it, surely we would be better off to frankly admit that what we are doing is weighing competing principle-based considerations. Practitioners do not have time to engage in a semantic charade aimed at preserving the absolute character of a principle. Similarly, with the principle of autonomy, under what conditions does an individual possess or cease to possess autonomy in the requisite form? It will be granted readily that infants, small children, and adults seriously impaired by Alzheimer’s, for example, do not have it. But what, precisely, are the conditions for recognition or denial? A moment of reflection on the ethical questions about suicide and passive voluntary euthanasia reveal the problems: Does a would-be suicide or a person requesting euthanasia satisfy the conditions of autonomy and thus have a right to deference by healthcare professionals to his/her autonomous decision? Can/should this question be answered without carefully considering the merits of individual cases?

One of the attractions of an absolute principle is that wherever it is applied, it would generate answers in the fashion of an algorithm thereby freeing us from the burdens and responsibilities of judgment. Trouble is, so far, no purportedly absolute moral principle has

emerged from theoretical interrogation intact. Moreover, in light of the stakes and the nuances of human interaction and interdependence, reliance on a set of prima facie principles applied in good faith is clearly to be preferred to “moral algorithms” that are not quite what they purport to be after all.

JUSTICE

The need to rely on judgment exercised in good faith is especially clear when we turn to the fourth principle, justice. Another thing becomes clear as well. The scope of biomedical ethics is such that it is nearly impossible to maintain sharp boundaries between it and other normative disciplines such as social and political philosophy, jurisprudence (legal philosophy), law, and politics. When we ask about the distribution of benefits and burdens in society, for example, when we query whether it is just or fair that the rich and powerful benefit from medical advances that emerge from clinical trials or studies conducted on the weak and the poor, we are in an area where ethical concerns overlap mightily with social, legal, and political concerns. So too when the problem we confront is the allocation of a scarce resource such as organs for transplantation.

What is justice? One answer is that justice consists in treating similar cases similarly and dissimilar cases dissimilarly, the dissimilarity in treatment being proportional to the difference between the cases. This essentially Aristotelian answer is likely to meet with nodding agreement. This is, however, a formal principle. Clearly, we need to know how, that is on what basis, similarity is measured. The relevant similarity might be in need, effort, contribution, or merit, for example. Or we might think that everyone should be treated equally regardless of differences in need and so on. Fig. 1 presents these competing (available) bases for justice in distribution; each gives rise to what is called a material principle of justice. Each of these material principles has some claim on our attention. However, in the context of biomedicine, many find need especially relevant and thus the most compelling of the competing alternatives. It is true, however, as Beauchamp and Childress observe, that there is no obvious obstacle to an embrace of all five as principles that entail prima facie duties.^[12] The key point, as before, is that the moral force of these principles (“their actual weight”) cannot be determined a priori; such an assessment is possible only on the ground, so to speak, in light of the particulars of a situation. Adjudicating among these material principles is a challenge of the first order to be undertaken seriously and in good faith.

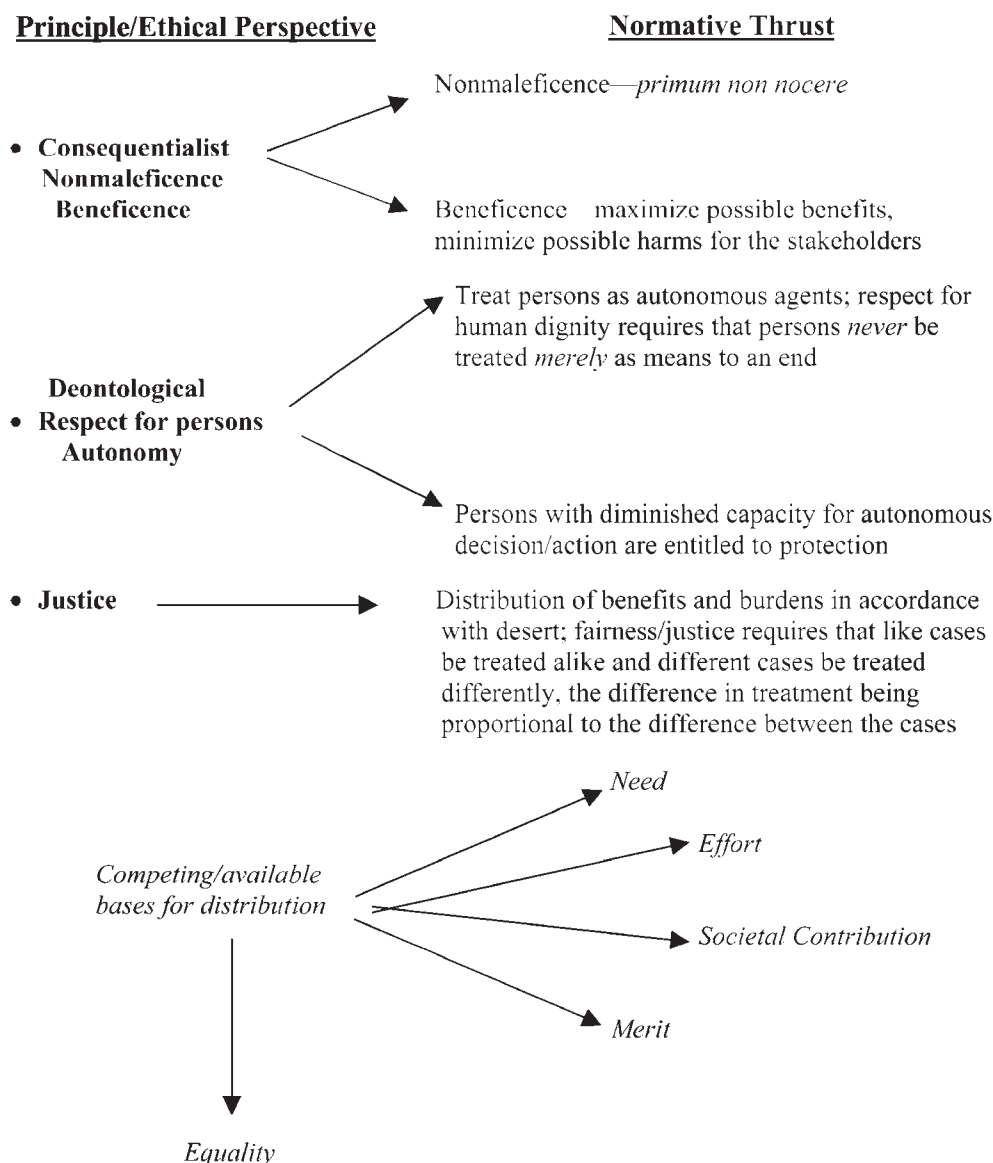


Fig. 1 The prima facie principles of biomedical ethics are tools/instruments of practical ethics. They are not absolute; they do not function algorithmically. Their application in concrete situations requires the exercise of judgement in good faith.

CONCLUSIONS

In closing, it will be well to return to a point about the birth of biomedical ethics. One of the key factors in the beginning was the need to respond to significant and rapid changes in medicine, especially advances resulting from research or emerging technologies. Thus, for example, practitioners faced ethical questions not encountered previously at the end of life (e.g., how to ensure a patient a meaningful measure of control over treatments he or she will receive as life ends; how, in the absence of advance directives, to determine whether further medical intervention is appropriate); at the start of life too, new questions emerged (e.g., whether it is permissible to conduct stem cell research that requires the destruction of human embryos left

over from and slated for destruction after successful in vitro fertilization, which was itself a source of moral controversy when the first “test tube baby” was born).

The need to respond to rapid changes in medicine continues and in fact seems destined to grow. Consider, for example, organ transplantation. There is a shortage of transplantable organs. The shortage is the result of an increase in demand, which, in turn, is the upshot of improvements in the methods for preserving and transporting organs and, most important, the discovery of cyclosporine, which suppresses the immune system so that the recipient’s body does not reject the transplanted organ. Consequently, questions that seemed merely academic 20 yrs ago are quite real and pressing. For example, should there be a market in transplantable organs? Or, alternatively, (a) should

requests to allow transplantation be legally required or (b) should there be a legal presumption that potential donors would agree to their organs being taken for transplant, unless they have opted out? It is essential that these questions and others like them are addressed in a systematic, reflective, and a responsible way. The principles of biomedical ethics discussed above will not provide algorithmic answers to these questions, but they are clearly germane. Moreover, seen as tools/instruments of practical ethics and applied in good faith, they appear well suited to meet this very real need.

ARTICLE OF FURTHER INTEREST

Telemedicine

REFERENCES

1. <http://www.npr.org/programs/morning/features/2002/jul/tuskegee/index.html> (accessed June 2005).
2. New York Times. 26 July 1972, 1, 8.
3. <http://www.tuskegee.edu/Global/Story.asp?s=1207598> (accessed June 2005).
4. <http://www.cdc.gov/nchstp/od/tuskegee/time.htm> (accessed June 2005).
5. http://webk.ask.com/redirect?u=http%3a%2f%2fwww.kn.pacbell.com%2fwired%2fBHM%2ftuskegee_quest.html (accessed June 2005).
6. Wysiyg//24<http://www.britannica.com/magazine/print?/ebSCO-id=11382> (accessed April 2001).
7. Jones, J. *Bad Blood: The Tuskegee Syphilis Experiment: A Tragedy of Race and Medicine*; The Free Press: New York, 1981.
8. Pellegrino, E. The Nazi doctors and Nuremberg: some moral lessons revisited. *Ann. Intern. Med.* 1997, 127 (4), 307–308.
9. <http://www.cdc.gov/od/ads/ethcodes/belm-eng.pdf> (accessed June 5, 2005).
10. <http://www.georgetown.edu/research/nrcbl/nrc/archives/ncphsguide.pdf> (accessed June 2005).
11. *Guidelines for the Conduct of Research Involving Human Subjects at the National Institutes of Health*; 00-4783; United States Department of Health and Human Services, Public Health Services, National Institutes of Health: Washington, DC, 2004; 1–24.
12. Beauchamp, T.L.; Childress, J.F. *Principles of Bio-medical Ethics*, 2nd Ed.; Oxford University Press: New York, 1983.
13. Hart, H.L.A. *The Concept of Law*; Oxford University Press: New York, 1961.
14. Gert, B. *The Moral Rules*; Harper and Row: New York, 1973.
15. Frankena, W.K. *Ethics*, 2nd Ed.; Prentice Hall: Englewood Cliffs, New Jersey, 1973.
16. <http://www.indiana.edu/~anmed/oath.htm> (accessed June 2005).

Biomimetic Materials

Jun-ichi Kikuchi
Atsushi Ikeda
Mineo Hashizume

Graduate School of Materials Science, Nara Institute of Science and Technology, Nara,
Japan

INTRODUCTION

Although the subdivision of science has given us numerous discoveries and a wealth of knowledge to understand biological systems, the importance of the boundaries between biology and other research fields was pointed out in the middle of the last century. The concept of biomimetic materials was born during the course of studies aimed at mimicking excellent and fascinating biomaterials. Much effort has been devoted to developing biomimetic materials from various points of view. The significance of biomimetic approaches is not only simple mimicry of biomaterials, but also the development of new materials exhibiting unique biomimetic functions. This article discusses biomimetic materials research and its future.

MIMICS OF CELL COMPONENTS

In biochemistry, biomaterials are classified into four main categories depending on each characteristic molecular structure and its consequence in the biochemical function. The categories are proteins, lipid membranes, nucleic acids, and polysaccharides. These biomaterials are macromolecules composed of covalently or noncovalently assembled small molecules, such as amino acids, lipids, nucleotides, and sugars. There are thus two approaches to the design of biomimetic materials: structural and functional mimicking. Major biomimetic materials and their biomaterial building blocks are shown in Fig. 1.

Protein Mimetic Materials

The focus of the early stages of biomimetic materials research was on the mimicking of functional proteins and on catalysis (or enzymes), because studies on enzyme models were useful in unveiling the reaction mechanism of enzymes.^[1–3] In general, enzymatic reactions proceed through formation of an enzyme-substrate (ES) complex. The process of forming this complex is intrinsic to the catalytic performance of enzymes, including high reactivity and strict substrate

specificity. Three types of enzyme mimicry capable of forming the ES complex—macrocyclic compounds, micelles, and polymers—have been proposed.^[4]

Macrocyclic compounds having an inner cavity as the substrate-binding site are simplified structural models of the enzyme active site (Fig. 2). Cyclodextrins (cyclic oligomers of glucose) are typical examples of the macrocyclic enzyme models.^[5] They are called α -, β -, and γ -cyclodextrins depending on whether the number of the glucopyranose units is 6, 7, or 8, respectively. The cyclodextrins bind hydrophobic substrates into their cavities in aqueous solution, and have two rims of hydroxyl groups that can either react with substrates themselves or be used to attach to other catalytic and functional groups. The introduction of multiple functional groups into cyclodextrins is essential to the design of more sophisticated artificial enzymes. For example, a β -cyclodextrin bearing two imidazolyl groups acts as an excellent artificial ribonuclease showing turnover behavior (possibly through the bifunctional catalytic assistance of the imidazolyl groups). On the other hand, a compound carrying both a pyridoxamine and an ethylenediamine unit attached to a β -cyclodextrin on the neighboring primary methylene groups has behaved as a potent artificial aminotransferase. Aromatic L- α -amino acids such as phenylalanine, tryptophan, and phenylglycine have been generated with high enantiomeric excess through transamination reaction.

Cyclophanes—synthetic macrocycles with aromatic ring(s)—have also proved to be useful as attractive components of artificial enzymes.^[6] Whereas a relatively limited range of structural modifications is possible with cyclodextrins, a wide synthetic variation of cyclophanes can be achieved so that, with regard to size, shape, and microenvironment, an appropriate recognition site can be provided for a target substrate molecule. Crown ethers, discovered by Pedersen in 1967 (for which discovery Pedersen, Cram, and Lehn were awarded the Nobel Prize in Chemistry in 1987), can act not only as ionophores but also as key components of artificial enzymes in organic solutions.^[7] Calixarenes, developed by Gutsche, belong

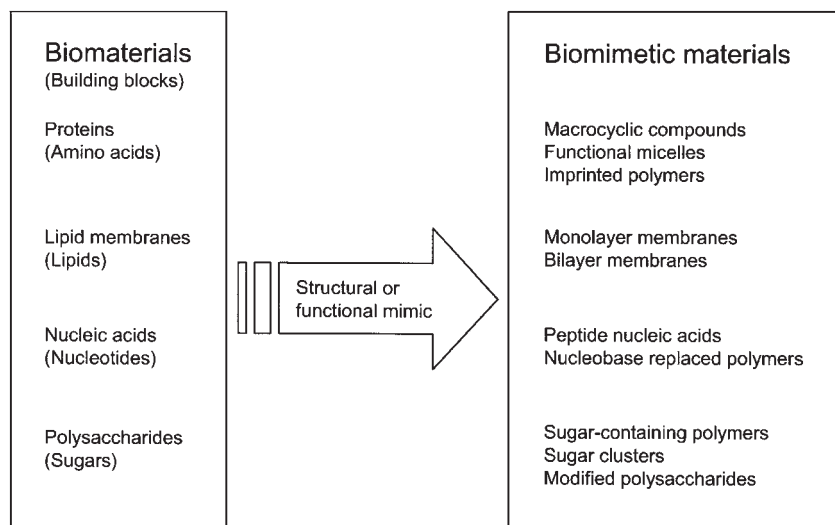


Fig. 1 Classification of biomimetic materials.

by definition to a family of cyclophanes; much attention has been focused on their catalytic functions as artificial enzymes based on their versatile molecular recognition abilities.^[8] Other macrocyclic compounds such as cryptands, spherands, and cucurbiturils seem to be feasible as equally promising enzyme mimics.

Another type of enzyme mimic provides a substrate-binding site as a result of the self-aggregation of functional elements through noncovalent intermolecular interactions in solutions.^[4] These are the surfactant

micelles, which, in aqueous media, have been widely employed as enzyme models. Reverse micelles formed with various surfactants in apolar solvents in the presence of small amounts of water have also been extensively studied for the characteristic features of their cores in enzyme-mimetic reactions.

On the basis of their structural resemblance to enzymes, various kinds of biomimetic polymers have been developed as artificial enzymes. These polymers are mainly classified into modified biopolymers and

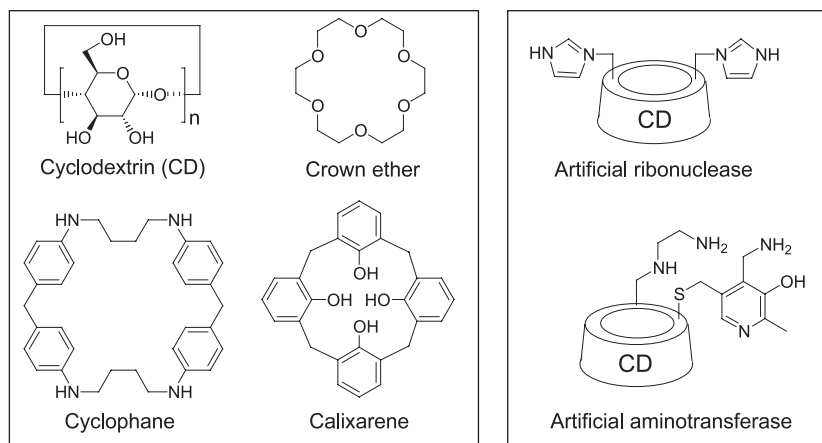
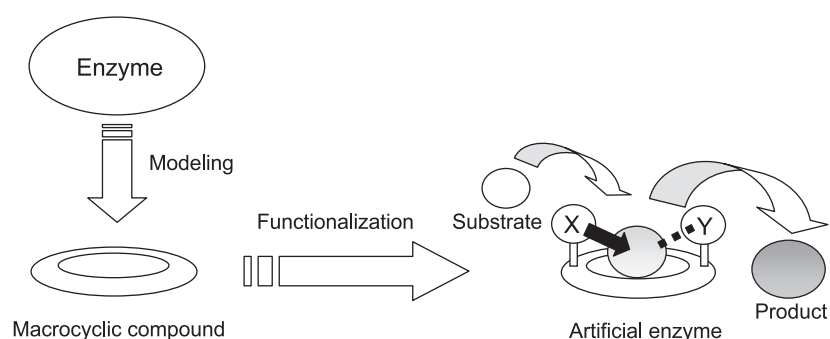


Fig. 2 Examples of biomimetic materials employing macrocyclic compounds.

totally synthetic functional polymers. The design of artificial enzymes through site-directed mutagenesis (an example of the former classification) has been well established.^[9] By single or multiple replacement of amino acid residue placed in the enzyme protein, enzymatic functions such as reactivity, substrate specificity, and reaction selectivity can be altered. Incorporation of unnatural amino acids and nonamino acids into proteins has also been made possible by employing site-directed mutagenesis, expanding the genetic code.^[10] Development of catalytic antibodies or abzymes is another fascinating approach to artificial enzymes.^[11] The concept of catalytic antibodies is that, if a putative transition state analogue of a reaction for which a selective and efficient catalyst is desired is used to elicit antibodies, they are likely to catalyze the reaction of a parent substrate. Although the site-directed mutagenesis and the catalytic antibody technique are useful genetic methods, the chemical modification of enzymes remains a valuable tool for protein engineering. The latter approaches are divided into direct modification of an enzyme protein and modification of a cofactor bound to the enzyme.^[12,13]

As regards enzyme mimics formed with totally synthetic polymers, relatively simple water-soluble polymers with catalytic functionalities were employed in early studies.^[4] Although the simulation of enzymatic behavior was successful to a degree by employing such prototype models, recent interest has focused on more intelligent synthetic polymers, such as imprinted polymers^[14] and dendrimers.^[15] When a polymer is prepared in the presence of print molecules as the transition state or intermediate analogue (extracted after polymerization), the remaining polymer may contain catalytic cavities capable of recognizing the print molecules. On the other hand, because dendrimers are structurally regulated polymers with an inner core like an active enzyme site, the characteristics that make them enzyme mimics have also been attracting attention recently.

The basic concepts of artificial enzymes can be applied to the design of artificial receptors, which recognize a signaling ligand to respond.

Membrane Mimetic Materials

Lipids are major constituents of biomembranes. As regards membrane-forming lipids such as phospholipids and glycolipids, their important functions generally derive from their formation of molecular self-assemblies that provide internal phases with defined sizes, shapes, and microenvironments for the incorporation of various biomolecules into their fluid matrices. Biomembranes exercise recognition and response to signal molecules, material separation and

transport, fusion, endocytosis, and exocytosis as cellular functions. Spontaneous formation of the lipid bilayer membranes was first reported by Bangham and Horne in 1964 for aqueous dispersions of naturally occurring phospholipids. In 1977 Kunitake and Okahata demonstrated that a totally synthetic bilayer membrane can be formed by a simple lipid, didodecyl-dimethylammonium bromide. These topics led to extensive research on lipid assemblies as biomembrane mimics.^[16]

There are several types of lipid assemblies as biomembrane models. The spreading of a lipid at an air-water interface leads to formation of monolayer membranes. Because they present one-half of the biomembrane's lipid bilayer, the lipid monolayer often utilized to understand the physicochemical characteristics of biomembranes.^[17] Investigations of the monolayer membranes supply precise information about the molecular packing and orientation of lipids that cannot be obtained from the bilayer membranes. Highly ordered, ultrathin Langmuir-Blodgett (LB) films, which are obtained by transferring lipid monolayers onto solid substrates,^[4] and self-assembled monolayers (SAMs) of thiolate lipids covalently bound on metal surfaces^[18] are expected to be useful as functional elements for biomimetic materials, especially for molecular devices. Liposomes are spherically closed lipid bilayer vesicles with an aqueous inner compartment formed in aqueous media. Vesicles prepared by various methods differ in diameter as well as in number of bilayers. For example, the sonication of lipid suspensions in water leads to small unilamellar vesicles whose diameter range is 20–100 nm. The swelling of lipid films with water under appropriate conditions, however, affords to uni- or multilamellar giant vesicles with the size of a cell. Planar lipid bilayers are another type of biomembrane model. They can be prepared either as freestanding black lipid membranes^[4] or as supported planar bilayer membranes.^[19] Various kinds of biomembrane processes have been simulated by employing these membrane mimetic materials.

In focusing on lipid bilayer membranes, relationships between the molecular structures of lipids and their consequence on assembling behavior have been systematically investigated up to the present. Thus it is relatively easy to design synthetic lipids showing expected aggregate morphology and physicochemical characteristics. As one such example, synthetic peptide lipids bearing amino acid residue, interposed between a polar head moiety and a hydrophobic double-chain segment through peptide bonds, have been developed.^[20] The peptide lipids that have an ionic head generally form stable bilayer membranes in aqueous media. Morphological stability of the unilamellar vesicles is greatly enhanced through formation

of the hydrogen-belt domain in the membrane. Owing to such strong hydrogen bonding interactions among amino acid residues of the lipid, phase transition parameters for the membrane from gel to liquid-crystalline phase are predictable from the molecular structure. In addition, the nonlamellar aggregates adopting inverted cubic or inverted hexagonal phase (suggested to participate in the membrane fusion process) can be systematically created by adjusting the critical packing parameters of the peptide lipids in the aggregate. Whereas the peptide lipids are effective biomembrane models, the bilayer vesicles act as artificial enzymes upon incorporation of the hydrophobized coenzyme moiety such as vitamin B₆ and B₁₂.^[6]

Morphological stabilization of lipid membranes is an important requirement for biomimetic materials. Numerous methods to stabilize the lipid assemblies have been developed, mainly by using polymeric systems.^[21] The polymerization of lipids before or after their orientation in membranes has been studied most intensively. Other possibilities to stabilize lipid assemblies are based on the noncovalent binding of polymers to the membrane surface. Morphological stabilization of the bilayer vesicles has been achieved by introducing the concept of the organic-inorganic hybrid into the liposome chemistry. Thus bilayer-forming lipids with a trialkoxysilyl head moiety form the cerasome, a liposomal membrane with ceramic surface, in aqueous media.^[22] Hierarchical assembly formed with the cerasome is expected to apply to the biomimetic multicellular systems or artificial organs.

Nucleic Acid Mimetic Materials

Nucleotides serve as building blocks in the construction of nucleic acids, in which nucleotide subunits are covalently linked by the formation of a phosphodiester bond between the phosphate group attached to the sugar of one nucleotide and a hydroxyl group attached to the sugar of the next nucleotide. The fundamental role of nucleic acids in the cell is in the storage and retrieval of biological information. To elucidate and mimic the versatile functions of deoxyribonucleic acid (DNA) as genetic material and ribonucleic acid (RNA) as its transcribed messenger molecule, a wide variety of nucleic acid model compounds have been designed, which show sequence-specific recognition of messenger RNA and double-stranded DNA.

Peptide nucleic acids (PNAs), defined as oligomers of nucleobase amino acids, were originally designed as mimics of DNA-recognizing, major-groove-binding, triplex-forming oligonucleotides.^[23] The PNA is constructed of three basic parts: an α -amino acid, a nucleobase attached via an amide linker, and an

ethylene spacer. The pseudopeptide backbone of a PNA-containing glycine unit has proven to be a good structural mimic of the ribose-phosphate backbone of nucleic acids. Thus, PNA has attracted wide attention in medicinal chemistry for development of therapeutic gene (antisense and antigene) drugs, and in genetic diagnostics.

DNA is not only a carrier of genetic information, but is also a versatile supramolecular scaffold, arranging smaller organic structures into predesigned geometries. Replacement of the nucleobases in DNA with designer molecules (such as fluorophores, ligands for metals, helix stabilizers, and DNA shape mimics) has thus been reported.^[24]

Sugar (Polysaccharide) Mimetic Materials

Polysaccharides and glycoproteins participate in a number of biochemical reactions and cell-to-cell recognition processes. Synthetic carbohydrate-based polymers are increasingly being explored as biodegradable, biocompatible, and biorenewable materials for use as water absorbents, chromatographic supports, and medical devices. Moreover, synthetic polymers bearing sugar residues also offer a good surface for cell attachment and might thus be applied to cell recognition events in antimicrobial/viral and tissue engineering.

Various types of synthetic polymers bearing sugar residues have been developed.^[25] These polymers are classified into four major categories: linear polymers, comb polymers, dendrimers, and hydrogels. Linear polymers are usually linked through the anomeric position with other hydroxyl groups in the adjacent sugar residue. Comb polymers are prepared from polymerizable sugar derivatives. Dendrimers are synthesized in stepwise fashion from simple branched monomer units, and are considered one of the prime nanometer-scale building blocks in the construction of nanoscale objects, molecular devices, and advanced drug delivery systems. To design intelligent sugar mimetic materials, the introduction of a cluster glycoside effect (or the enhancement of valence-corrected binding activities of saccharide ligands) are of great importance.^[26] Cross-linked polymers and hydrophobized polysaccharides have been developed as sugar-based hydrogels. Cholesterol-bearing pullulan forms self-assembled hydrogel, which assembly acts as an artificial molecular chaperone that selectively traps heat-denatured proteins and intermediates in refolding to prevent irreversible aggregation.^[27]

HOW TO MIMIC BIOMATERIALS

Although the foregoing presents a wide variety of approaches to the mimicking of biomaterials, one

guidepost in the creation of intelligent biomimetic materials is the hybridization of the functional building blocks of biological and artificial molecules. Typical examples of biomimetic materials prepared through the hybridization of functional building blocks are listed in Table 1.

As regards amino acids, their hybrids with sugars through covalent bonds have been present as biomaterials (glycoproteins). Hybrids of a polypeptide with a naturally occurring functional small molecule (such as coenzyme and cofactor) act as biocatalysts, (holoenzymes) for the reactions that cannot be catalyzed by the apoprotein alone. Thus, hybrid materials of amino acids with lipids, nucleotides, and artificial building blocks are candidates for biomimetic materials. Peptide lipids^[20] and peptide nucleic acids^[23] that have been designed according to this concept are excellent biomimetic materials, combining characteristics of each building block. Functional fusion of the catalytic group of amino acids with various artificial building blocks (such as cyclodextrins, cyclophanes, and crown ethers) produces catalytic macrocycles as artificial enzymes.^[5,6]

Several kinds of biomimetic hybrid materials that use lipids as building blocks have been developed. In addition to peptide lipids, hydrophobized polysaccharides, and cerasomes, nucleolipids^[28] form a hybrid with nucleotides, and octopus cyclophanes^[6] form a hybrid with cyclophanes. Nucleolipids (with a nucleobase as a polar head moiety of the lipid molecule) have been developed to clarify complementary interactions at the interface of the lipid assemblies. Octopus cyclophanes are unique amphiphilic macrocycles, capable of performing induced-fit molecular recognition and utilized as artificial enzymes and receptors.

Because of the diversity of molecular design in artificial building blocks and connection modes with

other counterparts, a large number of biomimetic hybrid materials (in addition to the examples in Table 1) have been reported.

MIMICS OF BIOLOGICAL SYSTEMS

An examination of biological systems highlights the existence of various functionally connected supramolecular systems, including metabolic and energy conversion systems and information networks. The second stage of biomimetic materials science should thus aim at construction of the supramolecular systems or molecular devices in which the functional connection of biomimetic materials with each other or with biomaterials can be achieved. A molecular platform capable of arranging individual functional elements is required for this purpose. The layer-by-layer assembling technique through multiple electrostatic interactions is a powerful tool to construct such nanoarchitectures on solid substrate.^[29] In the biomimetic sense, however, lipid membranes as biomembrane mimic are a fascinating platform for molecular organization; much effort has been devoted to constructing membrane mimetic supramolecular systems.^[6,22] Typical examples of such systems constructed in lipid bilayer membranes are shown in Fig. 3.

Setting up multienzyme systems in combination with natural and artificial enzymes is fascinating. An artificial multienzyme system has been developed that combines a catalytic bilayer membrane as artificial aminotransferase with a natural lactate dehydrogenase (Fig. 3a). The artificial aminotransferase is composed of a peptide lipid bilayer as apoenzyme model, a hydrophobized pyridoxal derivative as potent coenzyme derivative, and metal ions as cofactor. The multienzyme complex is formed with the lactate

Table 1 Strategy to create biomimetic materials through hybridization of functional building blocks

Biological and artificial building blocks	Hybridization of building blocks	Examples of biomimetic hybrid materials	Ref.
(1) Amino acids	(1) × (2)	Peptide lipids	[20]
(2) Lipids	(1) × (3)	Peptide nucleic acids	[23]
(3) Nucleotides	(1) × (6.1)	Catalytic cyclodextrins	[5]
(4) Sugars	(1) × (6.2)	Catalytic cyclophanes	[6]
(5) Coenzymes or cofactors	(2) × (3)	Nucleolipids	[28]
(6) Artificial building blocks	(2) × (4)	Hydrophobized polysaccharides	[27]
(6.1) Cyclodextrins	(2) × (6.2)	Octopus cyclophanes	[6]
(6.2) Cyclophanes	(2) × (6.5)	Cerasomes	[22]
(6.3) Crown ethers	(4) × (6.4)	Sugar clusters	[26]
(6.4) Dendrimers	(5) × (6.1)	Artificial holoenzymes	[5]
(6.5) Inorganic components			

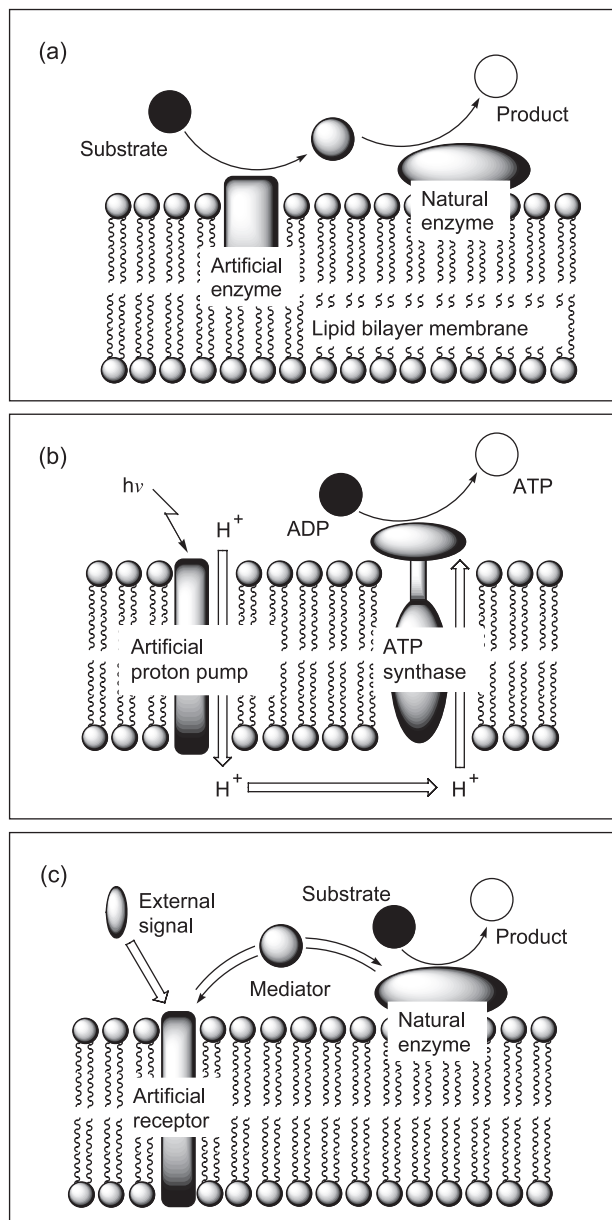


Fig. 3 Supramolecular systems exhibiting the functional connection of a biomaterial with a biomimetic material in lipid bilayer membranes: (a) multienzyme system; (b) energy conversion system; (c) signal transduction system.

dehydrogenase and the artificial enzyme, primarily through multiple electrostatic interactions. This hybrid system enhances catalytic activity if the artificial enzyme promotes the sequential transformation of α -amino acids to the corresponding α -hydroxy acids via formation of α -keto acids, showing marked substrate specificity. A multicomponent vesicular system has also been developed in which electrons derived from the enzymatic decarboxylation of pyruvate serves to reduce a synthetic, membrane-spanning steroidal

Mn(III) porphyrin and to mediate subsequent oxygen activation and transfer.

As an energy conversion system, an artificial photosynthetic membrane can be constituted that combines a liposomal bilayer, an artificial proton-pumping photocycle, and F-type ATP synthase (Fig. 3b). The proton pump is driven by vectorial photo-induced electron transfer in a synthetic carotene-porphyrin-naphthoquinone molecular triad. On accumulation of efficient proton motive force, proton flow through the coupling factor in thermodynamically allowed, with the concomitant formation of ATP from ADP and inorganic phosphate.

Detection of an external signal on the cell surface is accomplished by a specific receptor, and the signal recognition of the receptor molecule then stimulates the required intracellular response. Such signal transduction behavior mediated by cell-surface receptors presents various examples of supramolecular systems. In relation to the ion channel-linked receptor, a bio-sensing system has been developed, in which the conductance of a population of molecular ion channels is switched by the molecular recognition event. This approach mimics biological sensory functions and can be used with many types of receptors, including antibodies and nucleotides. An artificial signaling system inspired by biological signal transduction and mediated by the G-protein-linked receptor is another example. Because the G-protein is regarded as a signal transmitter between a cell-surface receptor and an enzyme, a strategy can be adopted that replace the G-protein in an artificial system by a simple mediator species capable of affecting both receptor and enzyme functions. Thus, a supramolecular system has been designed, based on an artificial receptor, a signal mediator, and a natural enzyme, placed on the lipid bilayer membrane (Fig. 3c). In this system, information of an external signal binding to the receptor molecule is transferred to the effector via a signal transmitter, followed by amplified signal output as enzymatic reaction.

CONCLUSION

Extensive studies of biomimetic materials have unveiled structural and functional mysteries in biological systems and have advanced synthetic and analytical procedures in materials science. At present, we have a wide variety of biomimetic materials exist that are not simple mimics of biomaterials, but are new materials and molecular devices that exceed biological functions. Further conceptual fusion of the biological and materials sciences may lead to a new field of science—bioinspired materials science.

ARTICLES OF FURTHER INTEREST

Biofunctional Polymers; Biologic and Synthetic Apatites; Extracellular Matrix Scaffolds; Surface Coatings; Tissue Engineering Scaffolds

REFERENCES

- Bruice, T.C.; Benkovic, S.J. *Bioorganic Mechanisms*; Breslow, R., Karplus, M., Eds.; W.A. Benjamin Inc.: New York, 1966; Vol. 2.
- Jencks, W.P. *Catalysis in Chemistry and Enzymology*; McGraw-Hill: New York, 1969.
- Bender, M.L. *Mechanisms of Homogeneous Catalysis from Protons to Proteins*; Wiley-Interscience: New York, 1971.
- Fendler, J.H.; Fendler, E.J. *Catalysis in Micellar and Macromolecular Systems*; Academic Press: New York, 1975.
- Breslow, R.; Dong, S.D. Biomimetic reactions catalyzed by cyclodextrins and their derivatives. *Chem. Rev.* **1998**, *98* (5), 1997–2011.
- Murakami, Y.; Kikuchi, J.; Hisaeda, Y.; Hayashida, O. Artificial enzymes. *Chem. Rev.* **1996**, *96* (2), 721–758.
- Lehn, J.-M. Supramolecular chemistry—Scope and perspectives. Molecules, supramolecules, and molecular devices. *Angew. Chem., Int. Ed. Engl.* **1988**, *27* (1), 89–112.
- Calixarenes: A Versatile Class of Macrocyclic Compounds*; Vecens, J., Böhmer, V., Eds.; Kluwer Academic Publishers: Dordrecht, 1991.
- Kondo, H. Towards Design of Artificial Enzymes and Receptors Through Site-Directed Mutagenesis of Proteins. In *Comprehensive Supramolecular Chemistry*; Atwood, J.L., Davies, J.E.D., MacNicol, D.D., Vögtle, F., Lehn, J.-M., Murakami, Y., Eds.; Pergamon: Oxford, 1996; Vol. 4, 528–547.
- Wang, L.; Schultz, P.G. Expanding the genetic code. *Chem. Commun.* **2002**, (1), 1–11.
- Feiters, M.C. Supramolecular Catalysis. In *Comprehensive Supramolecular Chemistry*; Atwood, J.L., Davies, J.E.D., MacNicol, D.D., Vögtle, F., Lehn, J.-M., Reinhoudt, D.N., Eds.; Pergamon: Oxford, 1996; Vol. 10, 268–360.
- Qi, D.; Tann, C.-M.; Haring, D.; Distefano, M.D. Generation of new enzymes via covalent modification of existing proteins. *Chem. Rev.* **2001**, *101* (10), 3081–3111.
- Hayashi, T.; Hisaeda, Y. New functionalization of myoglobin by chemical modification of heme-propionates. *Acc. Chem. Res.* **2002**, *35* (1), 35–43.
- Wulff, G. Enzyme-like catalysis by molecularly imprinted polymers. *Chem. Rev.* **2002**, *102* (1), 1–27.
- Astruc, D.; Chardac, F. Dendritic catalysts and dendrimers in catalysis. *Chem. Rev.* **2001**, *101* (9), 2991–3023.
- Kunitake, T. Synthetic Bilayer Membranes: Molecular Design and Molecular Organization. In *Comprehensive Supramolecular Chemistry*; Atwood, J.L., Davies, J.E.D., MacNicol, D.D., Vögtle, F., Lehn, J.-M., Sauvage, J.-P., Hosseini, M.W., Eds.; Pergamon: Oxford, 1996; Vol. 9, 351–406.
- Ariga, K.; Kunitake, T. Molecular recognition at air-water and related interfaces: Complementary hydrogen bonding and multisite interaction. *Acc. Chem. Res.* **1998**, *31* (6), 371–378.
- Ulman, A. Formation and structure of self-assembled monolayers. *Chem. Rev.* **1996**, *96* (4), 1533–1554.
- Sackmann, E. Supported membranes: Scientific and practical applications. *Science* **1996**, *271* (5245), 43–48.
- Murakami, Y.; Kikuchi, J. Supramolecular Assemblies Formed with Synthetic Peptide Lipids. Functional Models of Biomembranes and Enzymes. In *Bioorganic Chemistry Frontiers*; Dugas, H., Ed.; Springer-Verlag: Berlin, 1991; Vol. 2, 73–113.
- Mueller, A.; O'Brein, D.F. Supramolecular materials via polymerization of mesophases of hydrated amphiphiles. *Chem. Rev.* **2002**, *102* (3), 727–757.
- Kikuchi, J.; Ariga, K.; Sasaki, Y. Molecular Recognition and Functional Connection in Lipid Membranes. In *Advances in Supramolecular Chemistry*; Gokel, G.W., Ed.; Cerberus Press: South Miami, 2001; 131–173.
- Nielsen, P.E. Peptide nucleic acid. A molecule with two identities. *Acc. Chem. Res.* **1999**, *32* (7), 624–630.
- Kool, E.T. Replacing the nucleobases in DNA with designer molecules. *Acc. Chem. Res.* **2002**, *35* (11), 936–943.
- Wang, Q.; Dordick, J.S.; Linhardt, R.J. Synthesis and application of carbohydrate-containing polymers. *Chem. Mater.* **2002**, *14* (8), 3232–3244.
- Lundquist, J.J.; Toone, E.J. The cluster glycoside effect. *Chem. Rev.* **2002**, *102* (2), 555–578.
- Akiyoshi, K.; Sasaki, Y.; Kuroda, K.; Ueminami, A.; Sunamoto, J. Controlled Association of Macromolecules in Water: Hydrophobized Polymers and Proteins. In *Molecular Interactions and Time-Space Organization in Macromolecular Systems*; Morishima, Y., Norisue, T., Tashiro, K., Eds.; Springer-Verlag: Berlin, 1999; 151–158.
- Ringsdorf, H.; Schlarb, B.; Venzmer, J. Molecular architecture and function of polymeric oriented systems: Models for the study of organization, surface recognition, and dynamics of biomembranes. *Angew. Chem., Int. Ed. Engl.* **1988**, *27* (1), 113–158.
- Decher, G. Layered Nanoarchitectures Via Directed Assembly of Anionic and Cationic Molecules. In *Comprehensive Supramolecular Chemistry*; Atwood, J.L., Davies, J.E.D., MacNicol, D.D., Vögtle, F., Lehn, J.-M., Sauvage, J.-P., Hosseini, M.W., Eds.; Pergamon: Oxford, 1996; Vol. 9, 507–528.

Biopotential Amplifiers

Albert Lozano-Nieto

Engineering Technology and Commonwealth Engineering, The Pennsylvania State University, Lehman, Pennsylvania, U.S.A.

B

INTRODUCTION

The vast majority of biomedical amplifiers have been specifically designed to detect, measure, and record biopotentials, i.e., voltages and electrical fields generated by nerves and muscles. The measurement of these biopotentials is not a trivial task as their level is low (around 1 mV although they can be as low as 1 μ V) and is also immersed in noisy environments including internal and external interference.

This article starts with a short description of the different types of biopotential signals, followed by the requirements for amplifiers to detect biopotential signals and their design. After discussing the different interference sources for biopotential amplifiers, the entry analyzes the methods to mitigate their effects. The final section in this article describes how the generic biopotential amplifier described earlier is used for the measurement and detection of the different types of biopotential signals described at the beginning of this article.

BIOPOTENTIALS—ORIGIN AND NATURE

Biopotentials are produced as the result of some electrochemical activities in excitable cells. The voltage that can be measured across the membrane of these cells presents two very distinct periods: *action potential* when they are properly stimulated and *resting potential* without stimuli. During the resting potential phase, the voltage across the membrane of the cell is kept constant at a value that can range from 50 to 400 mV, with the exterior of the cell being more positive than the interior of the cell. This is also referred as the cell membrane being *polarized*. In the presence of an adequate excitation, the cell becomes *depolarized* acquiring an opposite voltage up to 60 mV that is independent of the amount of stimuli applied to the cell. This period of change in voltage polarity is called the action potential. After the action potential has been triggered, the ability of the cell to respond to another stimulus is extremely restricted. During the initial period of the action potential, the cell is not able to respond to any stimulus no matter how intense. This period is called *absolute refractory period*. It is followed

by the *relative refractory period* in which another action potential can be triggered by a stimulus of much higher value.^[1,2]

Biopotentials can then be viewed as the additive contribution of all the several action potentials propagating through the conductive tissue. Therefore, biopotentials can be classified depending on the type of cells and tissue, thus yielding the traditional biopotential signals. Furthermore, it is also important to note that because of the travel and propagation of these signals, the shape, amplitude, and other characteristics of biopotentials detected on the surface of the body will be very different from those detected inside the specific tissue that one wants to measure.

TYPES OF BIOPOTENTIALS

The following are some of the most common types of biopotentials. This is not an exhaustive list however. A more complete and detailed list of biopotentials can be found for example in Ref. [3].

Electrocardiogram

The electrocardiogram (ECG or EKG) is probably the most and best characterized biopotential and therefore is the one used mostly in the clinical practice. The cells that form the different tissues in the heart (sinoatrial and atrioventricular nodal tissues, atrial, Purkinje, and ventricular tissues) have very differentiated action potentials that in turn originate the composite signal known as ECG. The detailed explanation of the sequence of events for the electrical activation of the heart chambers as well as the interpretation of the resulting signal is beyond the scope of this section and can be found elsewhere.

The electrical activity of the heart as recorded on the surface of the body is the signal known as ECG. However, when it is necessary to record signals of higher magnitude to evaluate the significance of very small fluctuations, or when it is necessary to acquire the high-frequency components of the action potentials in the heart, the traditional ECG has several problems. These are mainly owing to the attenuation that the

electrical signals suffer as they travel through the different tissues of the body surface, the low-pass filtering associated to the same condition and the problems associated with the electrode-to-skin interface.^[3] In these situations, there are several alternatives for measuring the biopotentials of the heart with a higher degree of accuracy: In the *esophageal* ECG, the detection electrodes are placed inside a pill that is ingested by the patient and is stopped when it reaches the esophagus.^[4] Because of the closer proximity of the electrodes to heart and the absence of the electrode-to-skin interface and of motion artifacts, the signal obtained is much clearer and with more details, being especially useful to analyze the lower voltage amplitudes of the P-wave. By means of a cardiac catheterization—most of times done for other medical reasons—it is then possible to detect and record the action potentials inside the heart.^[5] This technique allows recording these signals with the highest bandwidth possible.

Electroencephalogram

The term “electroencephalogram” (EEG) is referred to the diverse techniques that are used to record the fluctuations in potential owing to the activity of the brain. Different types of electrodes are used depending on their location, giving rise to different types of signals.^[6]

1. Electrodes on the scalp surface that are small silver pads or cups for the standard EEG.
2. Sphenoidal EEG, using alternate insulated silver and bare wire and chloride tip inserted through muscle tissue by a needle.
3. Nasopharyngeal EEG, by means of silver rods with silver balls at their tips inserted through the nostrils.
4. Electroencephalographic signal, using cotton wicks soaked in saline solution that rests on the brain surface.
5. Intracerebral signal using sheaves of Teflon-coated gold wires, used to electrically stimulate the brain.

However, the most common type of EEG is performed with electrodes placed on the scalp, after the area has been cleaned with alcohol to remove skin oils. In this case, 20 small surface electrodes are placed on the patient’s scalp according to a normalized location map. The resulting signal for all these three cases is the superposition of the signals generated by a wide variety of neuronal current generators. In contrast to the ECG in which each part of the signal is clearly differentiated and the point of the heart where it originated is known, the EEG signal is made of a series of irregular waves,

without a general pattern. The clinical analysis of the EEG signal is based on the frequency classification of the waves and their amplitude.^[7]

Electromyogram

The electromyogram (EMG) is the combined result of action potentials from skeletal muscle tissue called the *motor unit*. The EMG can be recorded using surface electrodes, although they can only be used with superficial muscles. The EMG signal recorded this way is sensitive over a wide area, thus originating from an unpredictable signal in which only its amplitude may be relevant. To avoid these problems, the EMG can also be recorded using several types of insertion-type electrodes that are useful to detect signals from deep muscles or from small regions within a muscle.^[8]

Electroretinogram

The electroretinogram (ERG) uses two electrodes: one electrode known as the exploring electrode attached to the cornea in a manner similar to a contact lens, and a reference electrode placed elsewhere on the body surface. When the retina is stimulated with a brief flash of light, it is possible to record a very specific sequence of action potential changes known as ERG.

Electrooculogram

The electrooculogram (EOG) is a technique used to record the movement of the eyes, normally for sleep and dream research, although it has also been used to measure levels of alertness. In an EOG, two electrodes are placed to the left and right of the eye, in the nose and temple. When the eye moves in one direction, the voltage increases in the electrode closer to the eye movement. By placing the electrodes above and beyond the eye, it is also possible to measure vertical eye movements.

REQUIREMENTS FOR BIOPOTENTIAL AMPLIFIERS

Given the characteristics of bioelectrical signals discussed previously, the general requirements for biopotential amplifiers can be summarized as:

1. Have high input impedance to not distort the biopotential signal that needs to be measured.
2. Provide electrical isolation to not influence or modify the physiological parameter to be measured.
3. Have a bandwidth adequate to the biopotential signal to measure, normally including DC-rejection

and limited enough to not increase the noise level.

4. Provide enough gain within its bandwidth to have an output level compatible with the rest of the measurement system.
5. Reject, as much as possible, the noise in the system as well as the other interfering signals.
6. Have low output impedance and be able to supply the amount of current necessary to the load.
7. Have the means or provisions for calibration to guarantee the accuracy of the processed signals.

In addition to these generic requirements, each application will demand additional requirements to meet the specific requirements demanded by a specific biopotential signal.

DESIGN OF BIOPOTENTIAL AMPLIFIERS

A typical configuration of an amplifier for biopotentials is shown in Fig. 1. It consists of two input electrodes that are used to measure the biopotential (*measuring electrodes*) and a third electrode used as reference (*reference electrode*), normally connected away from the other two. The two measuring electrodes are connected to the input stage of a differential amplifier. The mission of the differential amplifier is to amplify the signal between the measuring electrodes while rejecting any unwanted signals, either other biopotentials or external sources of interference.

In Fig. 1, V_d represents biopotential to be measured that with these configurations appear as a differential voltage, while V_{cm} represents the voltage that appears in both inputs, known as *common-mode voltage*. This voltage is normally produced as interference

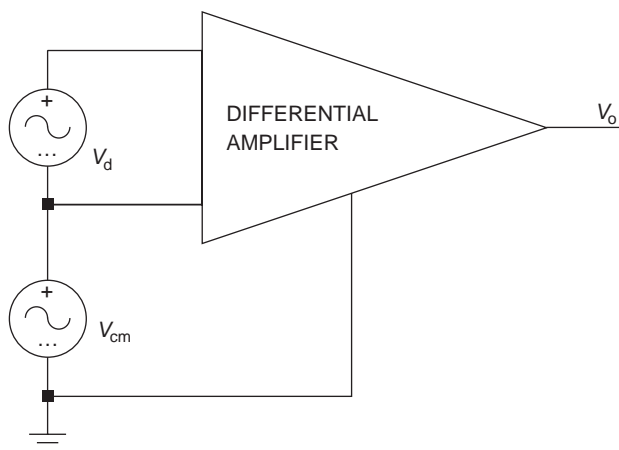


Fig. 1 Typical configuration of an amplifier for biopotentials. For simplification purposes, this figure does not include electrode impedance.

by the power lines surrounding the measurement system as well as interference from nearby instruments and even interference from other biopotential signals.

Ideally, the output of the biopotential amplifier could be represented as

$$V_o = A_d V_d \quad (1)$$

where V_o is the output of the amplifier, V_d is the differential voltage, that is, the biopotential to be measured and A_d is the differential gain of the biopotential amplifier.

From Eq. (1) it can be seen that for an ideal biopotential (differential) amplifier, the output is totally independent of the interfering voltage and only responds to the biopotential to be measured as differential input. It is necessary to mention that strictly speaking, not any differential amplifier can be used as a biopotential amplifier. Only those differential amplifiers that meet the requirements specified in the previous section—and in particular the requirement of high input impedance—can be used as biopotential amplifiers. These differential amplifiers are commonly known as *Instrumentation Amplifiers*.

With real electronic devices, however, this ideal situation is never achieved.^[9] Eq. (2) shows more realistic representation of the voltage at the output of the differential biopotential amplifier

$$V_o = A_d V_d + A_{cm} V_{cm} \quad (2)$$

where V_{cm} is the common-mode voltage at the input of the differential amplifier, and A_{cm} is the common-mode gain for the amplifier.

An additional problem arises when considering the relative magnitude values of the biopotentials to detect the interference signals. While biopotentials—differential signal—do not normally exceed 1 mV, the common-mode interfering signals can very easily have values around 30 V. With these circumstances it is then imperative that $A_{cm} \ll A_d$ in order for the output voltage to be meaningful by having a low value of error.

To give a figure of merit or quality to a differential amplifier, the parameter *common-mode rejection ratio* (CMRR) is generally used. The CMRR of an amplifier is defined as the ratio of differential gain to common-mode gain:^[10]

$$\text{CMRR} = A_d / A_{cm} \quad (3)$$

and is normally expressed in logarithmic units (decibels).

$$\text{CMRR}(\text{dB}) = 20 \log\{A_d / A_{cm}\} \quad (4)$$

We can then find the error term as follows

$$V_o = A_d\{V_d + V_{cm}/CMRR\} \quad (5)$$

Eq. (5) clearly shows that the error term ($V_{cm}/CMRR$) can be minimized by increasing the value of CMRR as much as possible. However, there are some factors that limit the increase of CMRR from a design and construction point of view. The previous calculations have been derived from an overidealistic case in which we have assumed that the input impedances to the differential amplifier, both in common-mode and in differential modes were infinity. In a realistic situation, the finite values of these impedances as well as their unbalances will contribute to degrade the value of CMRR, thus increasing the error in the output measurement. The CMRR decreases as the frequency of the signal increases. The study of how CMRR is affected by design parameters, design topologies, and type of semiconductors used is beyond the scope of this entry. However, it is generally accepted that the CMRR for a biopotential amplifier should be higher than 100 dB, preferably in the range of 120–140 dB.^[11] It is also accepted that the minimum values of the input impedances to be at $10^9 \Omega$.

The bandwidth requirements for the biopotential amplifier are carried out with a low-pass filter (LPF) and a high-pass filter (HPF). The HPF is placed as close to the front-end of the system as possible. The goal of this filter is to eliminate the DC voltages (*offset voltage*) that appear as a consequence of the electrode-to-skin interface. This offset voltage does not show a constant value, making it necessary to eliminate its drift. The LPF is designed with a cut off frequency high enough so it does not distort the biopotential signal, but low enough so it does not add unnecessary noise to the system. Given the small magnitude of

biopotential signals, it is essential to preserve the signal-to-noise ratio in the measurement system. As mentioned earlier, a biopotential amplifier needs to incorporate galvanic isolation to prevent the possibility of inducing a macroshock or a microshock on the organism that is measuring. Although there are several methods to create electrical isolation, one of the most common methods is to use an isolation amplifier. This isolation amplifier will create galvanic isolation as well as will introduce additional gain in the measurement gain. In any case, it is necessary to keep in mind that if there are active components in the patient side of the system, these need to be powered by an isolated power supply. Some isolation amplifiers incorporate an isolated power supply to solve this problem, while in using other isolation amplifiers or other methods for isolation, it is necessary to specifically introduce this isolated power supply. Moreover, international standards specify the maximum amount of leakage that is permissible on the patient side for different applications. Fig. 2 shows a typical disposition of the different systems and components for a biopotential amplifier, although as it has been mentioned earlier, some of the systems or components can be placed at different locations.

INTERFERENCE IN BIOPOTENTIAL AMPLIFIERS

Any biopotential amplifier has two main requirements that are in principle opposite: Provide electrical safety to the organism under study and have the highest degree of protection against interferences, therefore having to reach a compromise between them. This happens because the most obvious approach and solution to eliminate interferences in a measurement system such as grounding the object under

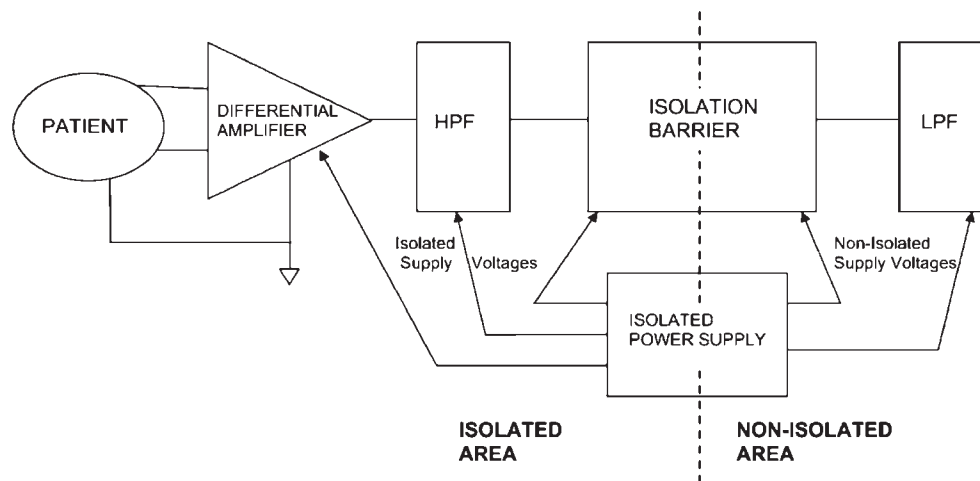


Fig. 2 Typical configuration for a biopotential amplifier. Depending on the specific type of biopotentials to detect, the HPF and LPF may be interchanged.

measurement cannot be applied here because it would result in a high risk of electrical shock. In fact, the degree of electrical isolation required for using biopotential amplifiers in human beings is a requirement that forces the need for a very careful design of the whole system to minimize the effect of interferences.

Interferences in biopotential amplifiers can be classified attending to their origin as

1. Other biopotential signals that occur simultaneously to the signal to be measured, such as the EMG signal when interested in measuring the EEG for example.
2. The voltage known as contact potential that appears in the electrode-to-skin interface owing to the chemical reactions. This signal can be eliminated by using a HPF, although the biopotential to detect may have energy components at a very low frequency, the design of this filter is not a trivial task.
3. Motion artifacts owing to electrode movements. The elimination of these interfering signals that can saturate the amplifiers requires the use of electrolytic gel to lower the contact impedance of the electrode, the selection of a front-end with a very high input impedance (above $1\text{ G}\Omega$) as well as choosing the front-end active devices with input offset currents as low as possible, ideally below 50 pA .
4. Respiratory movements that change the electrical characteristics of the thorax.
5. Triboelectric charges that may result in an increase of the patient-to-ground voltage owing to the accumulation of charges.
6. Interference from nearby high-power electrical signals such as electrosurgery units or even nearby high power radio, television, or radar facilities. In these situations, the patient leads act like an antenna. Other sources of similarly related electromagnetic interference are x-ray machines or switches and relays on heavy-duty electrical equipment.
7. Interference from the mains electrical distribution network.

From the enumerative list above, it is easy to see how the last item, the interference resulting from the main line distribution network can be the most harmful one, because it is the most difficult to eliminate and will always be present in a measurement system. Furthermore, the frequency of the mains line usually falls within the bandwidth of the biopotential amplifier, thus increasing the difficulties for its elimination. There are three paths for the energy from the mains electrical distribution network to enter the biopotential amplifier system and therefore become an interference signal.

These are capacitive coupling, inductive coupling, and resistive coupling.^[12]

Fig. 3 shows the basis for capacitive coupling: as the patient—that is acting as a conductor—is not grounded owing to safety reasons, it is possible to distinguish two stray capacitances, C_1 and C_2 , that are charged by the stray current I_s . I_s is generated by the voltage difference created by the mains generator and has typical values from 0.01 to $10\text{ }\mu\text{A}$. This originates a patient-to-ground voltage in the range from 100 mV to 1 V at 60 Hz depending on the values of C_1 and C_2 . It is important to note that the current I_s does not flow through the amplifier because of its high input impedance. Instead, it flows through the much lower skin-electrode resistance and stray capacitance to ground, thus closing the path. Although the patient voltage may seem low, it is indeed several orders of magnitude higher than the biopotentials to be measured. In addition to the patient voltage generated by the stray patient capacitances C_1 and C_2 , additional stray capacitances appear owing to the chassis of the amplifier not being grounded, thus increasing the amount of stray current and therefore the patient voltage with respect to ground.

Inductive coupling is owing to the current induction on wires and conductors generated by a variable magnetic field. These magnetic fields are generated by the power lines as well as by the transformers and ballasts in fluorescent lights. The main entry for this type of interference is through the cables that connect the amplifier to the electrodes. Inductive coupling can be greatly reduced by twisting the cables as much as possible and by avoiding cable loops of larger areas, as the value of the interfering voltage is proportional to the area of the effective loop.

Resistive coupling arises as a consequence of the resistance in the measuring wires, in particular the wires that connect the distant and proximal points in a measurement system. Because the return path is used by leakage currents to be drained to ground, they

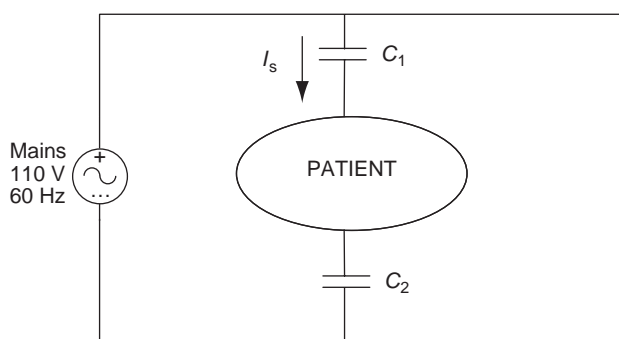


Fig. 3 Principle of capacitive coupling in which stray current flows through stray capacitances C_1 and C_2 .

generate an additional voltage that is not negligible and is present as a common-mode voltage superposed to the biopotential signals to be measured.

INTERFERENCE REDUCTION TECHNIQUES FOR BIOPOTENTIAL AMPLIFIERS

As seen earlier, much of the interfering signals are present as common-mode voltages. This is another reason why a differential amplifier with a high CMRR is absolutely necessary in the measurement of biopotential signals. A complementary approach for interference reduction is the elimination of the cause that originates the interfering signal when possible.

Electrostatic shielding is a technique useful to eliminate interference generated by electrical field, mostly through capacitive coupling.^[13] Shielding is accomplished by placing a grounded conducting plane between the source of electrical current and the measurement system as shown in Fig. 4. In this case, the interference current I_s is drained to ground as the impedance of the shield is much lower than the impedance presented by the stray capacitances inside the shield. Many hospitals have shielded rooms to measure EEG and other low-level signals. It is necessary to note, however, that the traditional electrostatic shielding is ineffective against magnetic fields. For shielding to be effective against magnetic fields, the shielding has to be made from a material that is a good electric and magnetic conductor, making its cost to increase for several orders of magnitude.^[14]

The elimination of ground loops is a technique useful to eliminate both inductive and resistive interference. As the magnetic interference is proportional to the area of the effective loop, by eliminating all ground loops, the number of areas that can pick up magnetic fields is drastically reduced. Furthermore, by designing grounding circuits so that there is a single unique path between each component and ground as well as a single

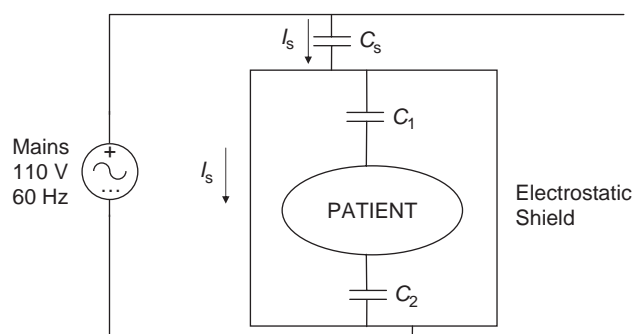


Fig. 4 Principle for electrostatic shielding, making the stray current I_s not to modify the voltage of the patient referred to ground.

unique path between the ground circuit of one component and that of any other component of the system, the interfering voltages generated by resistive coupling are eliminated.

Finally, a very common technique for interference reduction is the so-called *driven-right-leg* approach. In this technique, the right leg of a patient is connected to the output of an auxiliary amplifier whose voltage follows the inverted common-mode voltage in the patient. This approach keeps the patient ideally at 0 V, becoming a virtual ground without the risks associated to directly connecting the patient to ground. In reality, the characteristics of the amplifier used to reinject the common-mode voltage into the patient will determine the patient voltage, although in any case it will be greatly reduced.

TYPES OF BIOPOTENTIAL AMPLIFIERS

Amplifiers for Electrocardiographic Signals

A single electrocardiographic signal requires three electrodes as described earlier: two measuring electrodes and a reference electrode. The combination of how the measuring electrodes are placed on the body is called a *lead*. In the clinical practice, it is necessary to record more than one lead to fully characterize the electrical activity of the heart. There are a total of three basic leads (differential measurements between measuring electrodes), three augmented leads (unipolar measurements between one measuring electrode and a reference electrode denominated *Wilson's central terminal*), and six precordial leads. The detailed characteristics of all these leads can be found elsewhere in this volume.

An amplifier for electrocardiographic signals needs to accept and process the electrical signals originated by each one of these possible leads that have very different characteristics in terms of their amplitude and frequency components. For this reason, the first stage on an electrocardiograph is a lead selector that will choose the appropriate combination of the biopotentials on the patient for the lead or leads that need to be recorded.

Commercial amplifiers for ECG signals also incorporate a calibration signal for the operator or technician to verify the accuracy of the gain. The calibration signal is a 1 mV signal that produces a specific displacement of the needle on paper or on the output screen. The overall frequency response for an amplifier for ECG signals that has a flat bandwidth ranges from 0.14 to 25 Hz. The traditional method of recording ECG signals has been on a paper chart by using thermal-sensitive paper to avoid the friction of a writing instrument with the paper. However, with the advent of computers and networks of computers, ECG signals are increasingly displayed on a computer screen

and recorded on magnetic or solid-state media. However, when ECG is displayed on a computer screen, they still follow the traditional protocol of using a speed of 25 mm/sec as the clinical community is used to observing them.

Amplifiers for ECG signals also need to be protected against large signals generated by systems other than the amplifier itself. This happens, for example, in patients undergoing electrosurgery or being defibrillated while having an ECG amplifier attached to them. Ideally, an amplifier for ECG signals should be designed to be completely unaffected by these signals and continue functioning normally even after encountering these large signals. However, in practice this is not possible and designers need to settle for limiting the effects of these transient signals on the ECG signal to measure as well as to protect the equipment. This is achieved by connecting a voltage-protection device between ground and each one of the inputs from an electrode. These voltage-protection devices do not respond to the small-level signals generated by the cardiac biopotentials. However, when a certain threshold is exceeded—this threshold can be chosen by selecting the type of voltage-protection device to be used, normally in the range of 300–600 mV—they become active and drain all the signal to ground, thus protecting the ECG amplifier. It is also common to use different types of voltage-protection devices to ensure the protection of the biopotential amplifier against a wide range of harmful voltages. In addition to protecting the ECG amplifier, these voltage-protection devices also need to have a short recovery period because while they are active, blocking all the signals at the input of the amplifier, there is no recording of the patient's ECG. Another source of blocking the ECG signal in the presence of an over-voltage is the transient time required to discharge the HPF once it is saturated. Because the low-frequency response of the filter is at 0.25 Hz, the ECG signal may be blocked for several seconds after the offending signal has been eliminated.

Amplifiers for Electroencephalographic Signals

The clinical use of EEG signals requires the recording of more than one EEG channel, normally 8, 16, or 32 different channels chosen from the different combinations possible with the 20-scalp electrodes that are used following the standardized electrode placement. Therefore, the first stage is a matrix-like set of switches that routes and combines the appropriate signals that have been selected. The output of the switching system will contain 8, 16, or 32 signals—depending on the type of EEG system used—that are connected to the input of a set of amplifiers acting as preamplifiers.

These preamplifiers have differential inputs and have a constant gain of 1000. Given the very small amplitude of the EEG signals recorded on the scalp (approximately $5\ \mu\text{V}$), all the front-end is extremely susceptible to interference from the power-supply lines as well as noise from other physiological signals, mostly from muscle contractions. For this reason, it is very important that the front-end preamplifiers are powered with a power-supply as clear as possible. Other requirements for these preamplifiers are: low internal noise, CMMR in excess of 100 dB at the mains frequency, a low-pass cut-off frequency of 0.1 Hz or below, very low DC drift, and input impedance higher than $10\ \text{M}\Omega$. The outputs of the array of differential preamplifiers are now unipolar signals that are sent to an array of adjustable gain amplifiers. These amplifiers are used to condition the signals to the gain selected by the operator to a level that allows them to be digitized for computer storing and display, or drive power amplifiers for chart recording. We need to note once again, that the very low cut-off frequency of the amplifiers will increase the desaturation time if the amplifiers become saturated, normally owing to the presence of noise artifacts. In addition to the preamplifiers and amplifiers, it is necessary to introduce one adjustment for the sensitivity of the whole system and a set of adjustments for the sensitivity of each one of the unipolar amplifiers. These adjustments need the presence of a known and stable calibration signal that ranges from 5 to $1000\ \mu\text{V}$, injected into the differential preamplifiers.

The technique known as *averaging repetitive EEG* allows for detection and separation of the very small amplitude-evoked potentials from the background EEG. These evoked potentials or *cortical potentials* can be evoked from sensory stimuli, although their amplitude on the scalp is too small to be recorded by typical EEG amplifiers, thus requiring the averaging repetitive approach. The averaging technique has two effects: it enhances the evoked potentials signal while at the same time it reduces the ongoing EEG signals because of its random nature. The number of averages needs to be selected carefully as some information is lost in the process. A typical number of averages are approximately 100. The approach to detect evoked potentials is similar to the EEG, although in this case there is only one differential channel. After the signal is amplified, it is fed to a system for averaging. This system also synchronizes the stimuli that produce the evoked potential, either visual or auditory.

Amplifiers of other Bioelectrical Signals

Amplifiers used to detect and record the other bioelectric signals described earlier follow a similar

structure, mostly differing in the frequency bandwidth and the amplification level, thus placing different constraints in the design of the amplifiers. Additionally, the electrodes that are used to detect the signals may also have a strong influence. While the impedance of the interface between skin and ECG or EEG electrodes is generally low, other type of measurements may be affected by higher impedance at the interface as well as by the effects of more abundant drifts and presence of noise artifacts.

As mentioned earlier, the EMG signal has energy in the frequency range from 25 Hz to several kHz, with amplitudes that can range from 100 to 100 mV. This makes an EMG to have a much wider bandwidth than ECG or EEG amplifiers, but liberates them from the strong constraints of a low frequency as low as these amplifiers. This originates, in turn, that EMG amplifiers will have faster response times when saturated owing to motion artifacts. Furthermore, these low-frequency signals created by noise and electrode artifacts can be filtered to eliminate them. When using surface electrodes for the EMG detection, the amplifier will have to accommodate a bioelectrical signal in the range from 0.1 to 1 mV, with an electrode impedance generally low, below 1 kHz. These requirements are not very different from the specifications for an ECG amplifier. However, when using intramuscular needle electrodes for EMG detection, the amplitude of the bioelectrical signal increases in about one order of magnitude, thus requiring less gain. The electrode impedance also increases considerably owing to the reduction of the contact area, therefore forcing to use amplifiers with much higher input impedance values than those used for ECG or EEG.

The use of very small electrodes is necessary when we want to measure the potential across the membrane of a cell that is in the range from 50 to 100 mV. This is achieved with the use of intracellular electrodes or microelectrodes that can penetrate a single cell. The problem associated to the use of these electrodes resides in the fact that the small size and small effective contact area of these electrodes give them a very high source impedance, needing an amplifier with a value of input impedance higher than several G Ω . Moreover, the high shunting capacitance of the electrode itself affects the frequency characteristics of the system, reducing the effective bandwidth to values lower than desired. The cell has information of physiological value as low as DC and as high as 10 kHz. To compensate for this high capacitance, the detecting amplifier must use positive-feedback approaches, even with the risks of instability and oscillation associated to positive feedback. Another problem linked to the use of positive feedback is the increase in the noise level, although

this problem is diminished by the high value of cellular potentials. All this makes the design of amplifiers for microelectrodes to be a nontrivial task, more complicated than the design of amplifiers for any other biopotential signal.

CONCLUSIONS

Biopotential amplifiers are electrical voltage amplifiers that have been specifically designed to detect the electrical voltage signals generated in the human body due to several physiological processes. The detection of biopotential signals is difficult due to biopotential having voltages several orders of magnitude below the other signals present in the measurement system that become sources of interference. Biopotential amplifiers are therefore required to have very critical requirements in order to be able to provide an accurate representation of the biopotential signal they have been designed to measure.

REFERENCES

1. Barr, R.C. Basic electrophysiology. In *The Biomedical Engineering Handbook*; Bronzino, J.D., Ed.; CRC Press: Boca Raton, Florida, 1995.
2. Thakor, N.V. Biopotentials and electrophysiology measurement. In *The Measurement, Instrumentation and Sensors Handbook*; Webster, J.G., Ed.; CRC Press: Boca Raton, FL, 1999.
3. Clark, J.W. The origin of biopotentials. In *Medical Instrumentation Application and Design*, 3rd Ed.; Webster, J.G., Ed.; Wiley: New York, 1997.
4. Aston, R. *Principles of Biomedical Instrumentation and Measurement*; Merrill: New York, 1990; 195 pp.
5. Beams, D.M. Pacemaker sense amplifiers. In *Design of Cardiac Pacemakers*; Webster, J.G., Ed.; IEEE Press: New York, 1995.
6. Neuman, M.R. Biopotential electrodes. In *The Biomedical Engineering Handbook*; Bronzino, J.D., Ed.; CRC Press: Boca Raton, FL, 1995.
7. Bronzino, J.D. Principles of electroencephalography. In *The Biomedical Engineering Handbook*; Bronzino, J.D., Ed.; CRC Press: Boca Raton, FL, 1995.
8. Henneberg, K.A. Principles of electromyography. In *The Biomedical Engineering Handbook*; Bronzino, J.D., Ed.; CRC Press: Boca Raton, FL, 1995.
9. Franco, S. *Design with Operational Amplifiers and Analog Integrated Circuits*, 3rd Ed.; Mc Graw Hill: New York, 2002.
10. Pallas-Areny, R.; Webster, J.G. Common-mode rejection ratio in differential amplifiers. *IEEE Trans. Instrum. Meas.* **1991**, *40*, 669–676.

11. Nagel, J.H. Biopotential amplifiers. In *The Biomedical Engineering Handbook*; Bronzino, J.D., Ed.; CRC Press: Boca Raton, FL, 1995.
12. Ott, H.W. *Noise Reduction Techniques in Electronic Systems*; Wiley: New York, 1988.
13. Pallas-Areny, R.; Webster, J.G. *Sensors and Signal Conditioning*; John Wiley & Sons: New York, 1991.
14. Kimmel, W.D.; Gerke, D.D. *Electromagnetic Compatibility in Medical Equipment*; IEEE Press: New York, 1995.

A black square containing a white, bold, serif capital letter 'B'.

Bioreactors

Lito C. Mejia

EnduraTEC Systems Corporation, Minnetonka, Minnesota, U.S.A.

Kent S. Vilendrer

Tissue Growth Technologies, Inc., Minnetonka, Minnesota, U.S.A.

INTRODUCTION

Bioreactors play a significant role in the ever-expanding world of functional tissue engineering. A new generation of bioreactors that are increasingly comprehensive in capability is emerging. These bioreactors do more than just culture cells; they apply multi-axial physiological pressure/loading conditions to scaffolds and the growth environment, they analyze mechanical properties during the growth stages, and they can optimize the stimulation protocol to intelligently achieve the desired properties of functional tissue-engineered medical products.

As of this writing, the primary role of the bioreactor has been basic tissue research. These devices have often been home built, lacking the design and experimental discipline to be readily comparable to other home built systems and their respective results. The authors of this paper have proposed a technology platform for bioreactor design that can provide flexibility for both cardiovascular and orthopedic applications, and thus advance the uniformity of inter-laboratory comparisons, leading to improved development of functional tissue engineering processes. Commonality of design also provides a scalability factor and allows multiple constructs to be evaluated under a varying array of parameters. Bioreactor technologies that have scalability with faster evaluation and established testing protocols will become extremely important as functional tissue engineering moves from the research world to the production world.

TISSUE ENGINEERING

Tissue engineering is a promising field that offers a potentially unlimited source for transplantable tissues to replace damaged human ligaments, repair cartilage, or be an outright replacement of bone. It is not limited to tissues in the musculoskeletal system; there is also significant research and commercial activity in the cardiovascular arena for such tissues as pericardium, heart valves, arteries, and supporting blood vessels and organs.

Tissue Engineered Medical Products (TEMP) are expected to play a major role in the future of heart valve replacement, cardiovascular bypass surgery, venous

valve repair, and other intravascular surgeries. Research in this area has been focused on creating a three-dimensional construct from porous matrices (collagen, elastin, poly(glycolic acid), etc.) and seeding it with cells. For example, arterial construct development requires the appropriate biochemical and mechanical environments so that endothelial cells, smooth muscle cells, and fibroblasts proliferate to create an extracellular matrix. Mechanical stimulation of a seeded scaffold plays a significant role in inducing cell differentiation. Applying pulsatile pressure and media flow has shown that the applied stresses help promote faster cell growth, orientation, and mechanical strength.^[1-8]

Functional tissue engineering aims to develop biological substitutes for the repair, reconstruction, regeneration, or replacement of tissues. Its long-term goal is to construct biomaterials that are biocompatible, biodegradable, and capable of integrating molecules (e.g., growth factors) or cells.^[1,8] Currently, many different materials such as ceramics, polymers of lactic and glycolic acid, collagen gels, and other polymers have been tested in vitro and in vivo.^[1]

An important aspect of functional tissue engineering is that the discipline provides the basis for quantitative in-vitro studies of tissue development by culturing cells on 3-D constructs that are then exposed to physical and biochemical factors. It is here that bioreactors play their part. The bioreactor can be designed to provide cells capable of differentiating into desired lineages, seeding those cells onto a scaffold or other structural matrix, exposing the cultured cells to specific stimuli, and providing the necessary biochemical and growth factors that direct the cells to differentiate and develop. Bone and cartilage are the tissues in which most tissue engineering techniques have been applied, and the need for more research is substantial.

BACKGROUND

Cells communicate with each other chemically and respond to cues from their environment. The cues important to tissue engineers are those that tell cells to divide, migrate, and differentiate into different types of tissue. Mesenchymal stem cells, for example, have the

ability to become bone, cartilage, tendon, teeth, fat, or skin. Within the bioreactor there must exist a sterile environment of natural substances, such as growth factors, hormones, and other regulatory molecules, that guide these cells to differentiate along specific lineages (e.g., ligament lineage). It has been demonstrated that mechanical stimulation of the environment or mechanical stress (particularly multidimensional strain) of the seeded scaffold plays a significant role in inducing cell differentiation.^[9] Altman et al. documented that mechanical stimulation of ligaments based on human or bone marrow-derived cells induced elongated, ligamentlike cell morphology and cell alignment in the direction of loading. This was contrasted to round and randomly distributed cells with nonstimulated controls.

Why mechanical stimulation enhances cell differentiation is not clearly understood. It has been speculated that the mechanical forces signal or trigger cells to respond in some biological manner to produce ligament extracellular matrix components.^[9,10] Natural chemical environments also assist growth, making for challenging research in order to determine the optimal factors. Also, it is not known if the influence of mechanical stimuli on the differentiation process is transient or permanent. Ratcliffe and Niklason^[8] have noted that, under conditions of periodic compressive strain, the properties of engineered cartilage are significantly improved. Mol et al.^[11] note that straining resulted in more pronounced and organized tissue formation. Overall tissue properties improved with increasing levels of strain.

CLINICAL SIGNIFICANCE

According to the American Heart Association (2002 Heart and Stroke Statistical Update), approximately 1 million people suffered a myocardial infarction in 1999. Forty-five percent (45%) of these cases were fatal. Other statistics for 1999 include

- Cause of death was cardiovascular related in 41% of all deaths.
- 571,000 cardiac bypass surgeries were performed.
- Valvular heart disease claimed 19,612 lives.
- 96,000 valve replacement surgeries were performed.
- About 40,000 babies are born each year with cardiovascular defects.

The replacement or repair of diseased vessels with natural synthetic vascular grafts has become a routine treatment for certain types of intravascular disease. For coronary bypass surgery, the autologous saphenous vein remains the graft of choice for its nonthrombogenic flow surface, and its ability to be healed by the host as well for its strength and elasticity. However, efforts to

create a suitable synthetic small-diameter vascular graft have been largely unsuccessful.^[1]

Likewise, man-made medical devices have been used to replace heart valves and repair other intravascular complications. These devices are typically made from metals (stainless steel, nickel titanium alloys, carbon, fiber) and fabric (e-PTFE, Dacron, carbon fiber) that are foreign to the body. Their use often requires special blood-thinning medication that can lead to further health complications. Additionally, the devices are fixed form and do not conform to the body as the patient grows from childhood to adulthood, thus necessitating multiple surgeries. There are some areas (i.e., venous valves) where medical device designs offer low patency and the only means of repair is through reconstructive surgery or transplantation.

Knee ligament damage, in particular the anterior cruciate (ACL), is a frequently sustained knee injury. More than 100,000 ruptures of the ACL are estimated to occur every year in the United States,^[12,13] and, with the low healing capacity of the ACL, surgical reconstruction of some sort is necessary in order to restore normal knee function. A tissue-engineered ligament can address the shortcomings of traditional ACL therapies.

Knee meniscal tears due to twisting or compression forces are also common sports injuries. Nearly a million are estimated candidates for tissue-engineered meniscus. Traumatic meniscus tears accompany half of all anterior cruciate ligaments tears.

Damage to knee articular cartilage is a common problem following sports injuries. It leads to premature arthritis, causes a considerable decrease in quality of life, and has enormous long-term healthcare costs.^[12,14] Regeneration of damaged articular cartilage is very limited because adult cartilage lacks blood supply, lymphatic drainage, and innervation. Furthermore, chondrocytes are sheltered from synovial fluid nourishment and reparative recognition by their large extracellular matrix.^[15]

Arthritis, which can occur as either an inflammatory or degenerative joint disorder, affects tens of millions of people worldwide. Osteoarthritis is a form of arthritis that involves degeneration of the joint, with the most common being osteoarthritis of the knee. With the aging population, and the expectations for more active lifestyles, it is expected these numbers will increase dramatically in the coming years. With increasing obesity in younger, more inactive people, osteoarthritis is expected to become more prevalent in the young and middle age groups as well.^[16,17]

THE ROLE OF INTELLIGENT BIOREACTORS

Using intelligent bioreactors, a researcher can seed the construct and then apply pressure/flow waveforms of

varying shapes and frequencies to stimulate cell proliferation and matrix deposition leading to constructs with properties suitable for use in humans. The waveforms are generated using a computer-controlled dynamic pumping system that allows experimentation with different pressure/flow conditions. Godbey and Atala^[18] have reported that pulsatile fluid flow has been used to develop tissue-engineered heart valves, with the assumption that mimicking in-vivo conditions will yield a stronger tissue. Hoerstrup et al.^[19] and Sodian et al.^[20] describe bioreactors that utilize a respirator for driving pulsatile flow through heart valves and blood vessels.

Hoerstrup et al. as well as Dumont and Demarteau^[19-27] reported an increase in material properties for tissue-engineered heart valves, pulmonary artery conduits, and small-caliber vascular grafts stimulated for several weeks in a bioreactor that had pulsed flow. Jockenhoovel et al.^[28] describes a laminar flow chamber for in-vitro improvement of mechanical tissue properties. Ratcliffe and Niklason^[8] have noted that, under conditions of periodic compressive strain, the properties of engineered cartilage are significantly improved.

The intelligent bioreactor can also provide a non-destructive means for evaluating the strength of the developing tissue construct during its growth cycle. The loading system is capable of pausing the normal stimulation and introducing dynamic pressure or mechanical loading to the tissue construct to measure the construct's response. With special software, the complex modulus of the tissue can be estimated. This information may be stored for comparison of different constructs and for monitoring tissue construct development.

A bioreactor can culture in-vitro ligament tissues by seeding anterior cruciate or medial collateral ligament cells (fibroblasts) on porous scaffolds (synthetic biodegradable polymer or collagen-based matrices). Similarly, a bioreactor can culture in-vitro cartilage tissues by seeding cartilage cells (chondrocytes) on porous scaffolds (also of synthetic biodegradable polymer or collagen-based matrices). In time, the new tissue will develop its own supportive matrix and no longer need the scaffolding; therefore, biodegradable scaffolds are essential. Another approach involves the direct use of autologous tissues, which can differentiate within the host tissues (bone, meniscus, ligament, muscle) and can then be cultured in a bioreactor.

A six-station intelligent bioreactor represents an advancement in bioreactor technology in that several constructs can be developed simultaneously, allowing for more thorough and quantitative analysis of how a bioreactor environment enhances functional tissue growth. In addition, the ability to culture multiple constructs simultaneously allows the researcher to address important bioprocess engineering issues such

as scale-up and tissue uniformity. These criteria will become important as tissue engineering moves from the research laboratory to a production environment.

With the commercially available ITEMS (Intelligent Tissue Engineering via Mechanical Stimulation) BioReactor, the researcher can evaluate various combinations of culture techniques, scaling parameters, monitoring technologies, and control strategies to produce differentiated tissues with suitable biomolecular and mechanical properties. Until now, the production of one or a small number of constructs has been the norm, with little attention given to inter- or intrabatch variability for key tissue markers. Clearly, process development for production of tissue-engineered medical products will require knowledge of the variability of tissue characteristics, and definition of acceptable ranges for key cellular and molecular components and processing technologies.

CONSIDERATIONS FOR DEFINING CELLULAR CHARACTERISTICS

Experiments to test the effect of the appropriate conditions (hydrodynamic loading, oxygen tension, growth factors, etc.) on construct development can be performed. Constructs can be harvested periodically to provide kinetic data, and morphological analysis can be performed to include immunohistochemical staining and fluorescence microscopy to characterize cell and matrix distribution in growing constructs. Morphometric techniques to quantify construct matrix architecture for different bioreactor conditions will need to be developed. Quantitative morphometric analysis can include the use of confocal laser scanning microscopy (CLSM), scanning electron microscopy (SEM), and transmission electron microscopy (TEM).

CLSM can be used to provide spatial relationship information between cells, matrix proteins, undegraded polymer, and voids, and to allow for three-dimensional reconstruction of the construct. For example, CLSM can quantify tissue heterogeneity,^[29,30] including identification of secretory cells and degenerating cells,^[31] and structural organization of native tissue.^[31-33] These techniques have been proposed for the development of the quantitative ultrastructural detail of tissue-engineered medical products. SEM can provide high-detail matrix composition and allow visualization of collagen fibers^[34] in the matrix.^[29] Observations can then be confirmed using TEM. Relating tissue mechanical properties as a function of matrix organization, especially collagen and elastin, to the quantification of collagen and elastin fiber diameter, length, and branching in tissue sections will allow comparison of tissue structure to function (mechanical properties) and bioreactor growth conditions.

BIOREACTORS: A LITERATURE AND PATENT BRIEF

Various bioreactor designs exist, and most are single-purpose home-built designs. These include simple spinner flasks and rotating-wall vessels such as slow-turn-lateral and high-aspect-ratio vessels. Rotating-wall perfused vessels (such as hollow fiber bioreactors), along with perfused column and perfused chambers, complete the mix of bioreactor designs.^[34–37] These systems provide basic medium and gas exchange with mixing mechanisms combined with fluid flow to provide mass transfer in the bulk medium. Although they are initially used to attain uniform seeding of three-dimensional scaffolds, rotating bioreactors have been used to study effects of turbulent and laminar flow (mass transfer rates or direct hydrodynamic effects) on tissue development via continuous recirculation. Effects of microgravity have also been studied in these bioreactors and have included the use of magnetic resonance imaging (MRI) as a nondestructive means to monitor the development of engineered cartilage. In certain cases, local measurements of the velocity field made using particle-image velocimetry (PIV) permit the determination of local shear stresses (cartilage applications).^[38]

Biodynamic bioreactors attempt to maintain the spatial uniformity of cells during the seeding process as well as expose the (vascular) cells to pulsatile physical forces during vasculogenesis and throughout its life.^[4,5,39,40] Although chemical regulators mediate tissue morphogenesis, tissue-engineering researchers critically emphasize that the signals that are responsible for tissue pattern are often mechanical in nature.^[41] Examples include the pattern-generation effects of compression on bone, shear on blood vessels, and tension on muscles. Mechanical stresses also play a role at the cellular level (cell-generated forces) during tissue morphogenesis.^[41]

Mechanical testing and in-vivo implantation studies ultimately assess the functioning engineered tissue's (vascular graft) ability to withstand long-term exposure to systemic pressures and hemodynamic forces. Stress-strain analysis to provide elastic modulus and compliance data can help predict the graft's resistance to dilation and aneurysm formation in vivo.^[39] While in-vivo implantation is the definitive means for evaluating the performance of the engineered tissue, a literature search, reveals that current bioreactors lack the capability of multi-axis pressure/physiological loading, as well as the means to provide real-time mechanical properties assessment to allow for more intelligent functional tissue engineering.

There is a critical need for more advanced bioreactors that can provide 1) complete mechanical stimulation to enhance tissue growth with simultaneous

mechanical properties analysis, 2) real dynamic mechanical analysis (DMA) for viscoelastic characterization of the tissue, and 3) elastic modulus and compliance information attained during and after the growth process.

A new generation of more sophisticated bioreactors is emerging that can provide larger three-dimensional tissue-engineered constructs. Altman et al.'s^[8] tissue engineering of ligaments uses multiple bioreactor vessels housed in a mechanical device that can stimulate translationally and rotationally the collagen gel suspended between two bone anchors (see Fig. 1). Sodian et al.^[20] have designed a new pulsatile bioreactor for the fabrication of tissue-engineered patches. The closed-loop perfused bioreactor design combines continuous pulsatile perfusion and mechanical stimulation by periodically stretching the autologous tissue-engineered patch construct. This approach is aimed at providing optimal biomechanical and biodynamic stimuli for controlled tissue development. Dumont et al.^[21] describe their bioreactor pulsatile capability for aortic heart valves as able to perform hydrodynamic studies over a wide physiological range with controllable resistance, compliance, stroke volume, and frequency. Demartean et al.^[22] have designed multiple-chamber bioreactors for cartilage applications using computer-controlled stepper motors to provide compressive loading while the bioreactor chamber provides different mixing velocities and different polymer scaffolds.

Vorp et al. at the University of Pittsburgh have designed a single-specimen apparatus that can apply cyclic twist and extension of perfused vascular segments.^[42] Fig. 2 is a detailed sketch of the specimen-housing chamber. Perfusion tees enter into the chamber while allowing desired stretching and twisting motion. The chamber allows access for a video camera for

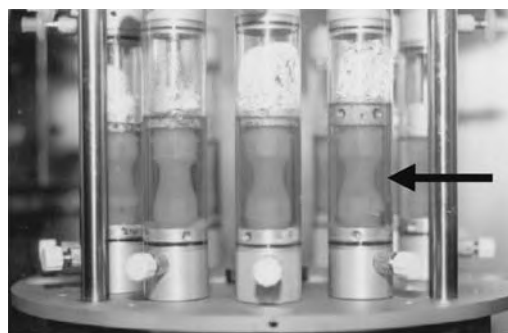


Fig. 1 This bioreactor design by Altman et al. (Tufts University) shows multiple chambers in a mechanical assembly that provides both translation and torsional stimulation. Collagen gel supported by bone anchors is shown.

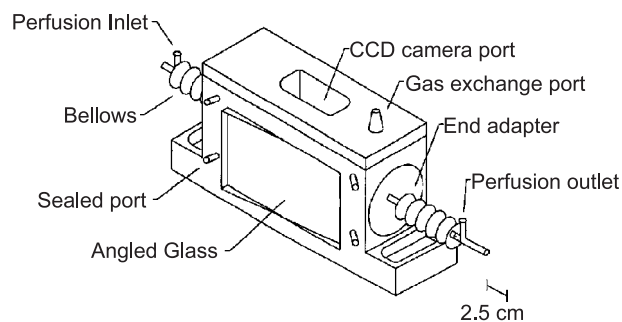


Fig. 2 This schematic details the specimen-housing chamber of a single-station bioreactor. Stretching and twisting of the vascular scaffold is accomplished inside the assembly. (Courtesy of D. A. Vorp: University of Pittsburgh.)

strain measurement. Ports allow for physiological sampling and oxygenation of the bath medium. Bioreactors employing nonperfusion, biaxial stretching techniques (e.g., Banes et al.), via a rubber membrane (via vacuum) are also available.^[43] This biaxial stretching technique, however, is more difficult for programming a specific physiological pressure or loading and strain profile to the gel matrix.

Enhanced techniques were used by Kasra et al.^[44] to determine the effect of hydrostatic pressure on rabbit intervertebral disc cells. Short-term application of high loading amplitudes and frequencies (to 20 Hz) was found to be beneficial in the stimulation of protein synthesis and the reduction of protein degradation. Bioreactor designs of the future can provide these higher levels of performance. Latterman at the University of Pittsburgh Hospital has designed a complete ligament bioreactor using a highly responsive dynamic linear motor; the complete system can be placed inside an incubator, as shown in Fig. 3.

Notable patents have surfaced within the past several years. Many are in the cardiovascular area. Vilendrer et al., in “Bioreactor with Plurality of Chambers for Conditioning Intravascular Tissue Engineered Medical Products,” provides a complete scalable bioreactor platform that incorporates multiple chambers and attempts to combine pressure, axial loading, torsional or lateral shearing, and bending of a seeded scaffold (see Fig. 4).^[45] Others are more specific to culturing tissues as a primary goal without any further conditioning of the construct.

Macallister et al. focus on a tissue-engineered blood vessel that is made from a cultured fibroblast sheet rolled into a multilayer vessel (via a rolling mandrel) to achieve sufficient burst strength to withstand physiological blood pressure without the use of a synthetic scaffold.^[46] One often finds variations of rotary bioreactors with enhanced features such as a two-chamber perfusion bioreactor (Dimitrijevič et al.) to counteract

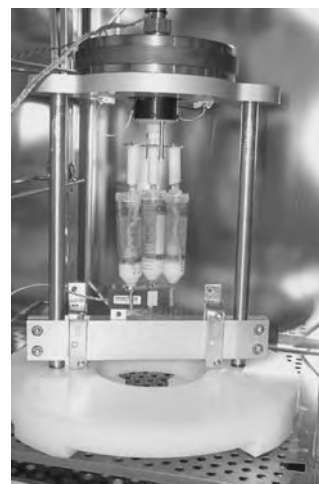


Fig. 3 This first-generation bioreactor shown inside an incubator uses a voice coil actuator as the actuation method and provides only axial loading. The three ligament chambers share a common load cell. The next design will feature independent load measurement and incorporate a moving magnet motor instead of the voice coil. (Photo courtesy of C. Latterman of University of Pittsburgh Hospital.)

significant shear stresses while allowing high nutrient transfer and removal.^[47]

INTELLIGENT BIOREACTOR DESIGN CONSIDERATIONS FOR CARDIOVASCULAR APPLICATIONS

The following sections describe a scalable biodynamic environmental system that provides a complete closed-loop, nutrient-flow subsystem capable of introducing mechanical stresses and strains to a tissue-engineered construct during its growth cycle. In addition to mechanical stimulation, the researcher can query the tissue construct to evaluate its mechanical condition. This procedure allows adjustment of the stimulation profile to best meet the desired growth plan for the chosen tissue-engineered construct.

The ITEMS (Intelligent Tissue Engineering via Mechanical Stimulation) BioReactor is based on the design of the EnduraTEC 9100 Series of Stent Graft Testers, and it includes several important design features for the tissue-engineering researcher. Additionally, experience with the single-specimen ITEMS BioReactor configuration in collaboration with a major research institute has provided insight into its functionality and usability. Several of the significant capabilities of this second-generation, commercial system are summarized below.

Servocontrolled Bioreactor

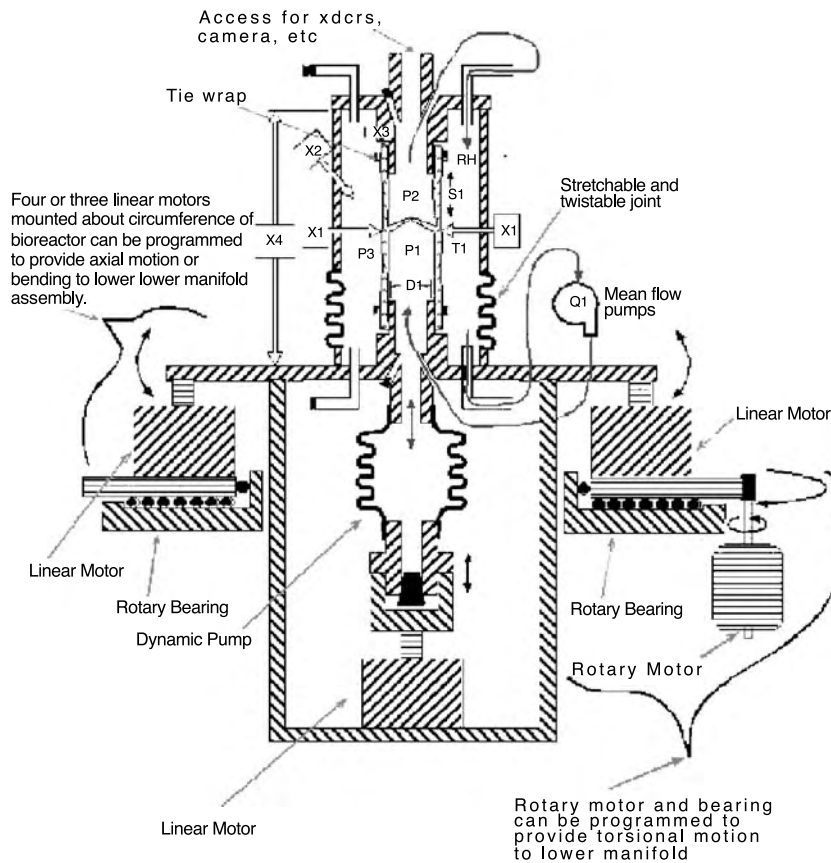


Fig. 4 Patents such as those described by Vilendrer et al. attempt to incorporate almost every facet of mechanical motion to apply to the construct/manifold. These include pressure, axial loading (stretching), torsion (twisting), lateral shearing, and bending.

Moving Magnet Motor

A central element of the ITEMS BioReactor is the moving magnet linear motor, which provides the cyclic motion that in turn provides cyclic pressure or cyclic loading (see Fig. 5). The ITEMS unit (with the linear motor) fits into a tissue culture incubator and is not affected by moisture.

The moving magnet motor is composed of three basic elements: the magnet, the coils, and the core. The magnet moves up and down while the coils and core remain stationary. The core and coil combination may be viewed as an electromagnet producing a north and south pole as a function of current. When the current is applied, the appropriate poles of the magnet are either attracted or repelled, producing the force and thus moving the actuator up or down. The stronger the current, the stronger is the force.

Magnetically, the magnet is charged, or magnetized, with two polarities. The function of the coil and core assembly is to carry and modulate the flux across the gap where the magnet resides. The flux generated by the coils interacts with the flux generated by the magnet, causing the magnet to react with a force either up or down depending on the polarity of the coil flux. The

implemented structure will produce a force that is directly proportional to the current in the coils. Maximum force is limited only by the thermal performance of the overall system design.

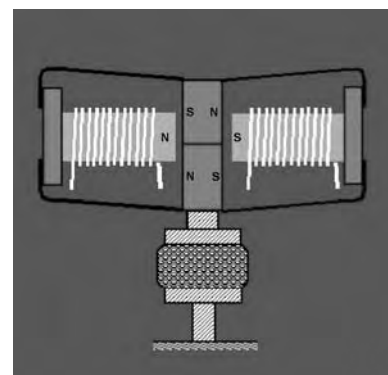


Fig. 5 This schematic details the moving magnet as part of a simple electromagnet principle. Its simplicity and clean design make it ideal for bioreactor applications. The magnet/flexure assembly offers low magnet mass and a frictionless-flexure-suspension design and can be controlled to provide both high-frequency performance and step-loading capability that is highly suited for characterizing soft tissue.

The moving magnet linear motor system is inherently a high-fidelity device due to its low magnet-mass and frictionless-flexure-suspension design. The capability of high-acceleration and high-frequency performance for cyclic/step load and displacement profiles makes it ideal for tissue stimulation and characterization. Naturally quiet in operation, and with very few moving components (only the magnet/flexure assembly is moving), the super-clean, efficient design (no seals or pneumatics/hydraulic residue) is highly desirable for use within an incubator system. Also noteworthy is the infinite fatigue life of the frictionless-flexure-suspension design, which provides a durability factor into tens of billions of cycles with literally no maintenance.

Ergonomic Features for a Laboratory Setting

The ITEMS BioReactor is a compact design; it can be completely placed within medium-sized incubation systems that control temperature, humidity, and the environment (see Figs. 6–9). A typical single-station bioreactor configuration is 305 mm (12 in) wide by 381 mm (15 in) deep and 560 mm (22 in) high, with the multistation only slightly wider and taller. The weight ranges from 14 to 20.4 kg (31–45 lb). The BioReactor



Fig. 6 Single-station BioReactor showing the moving magnet motor to provide dynamic pressure pulsing.

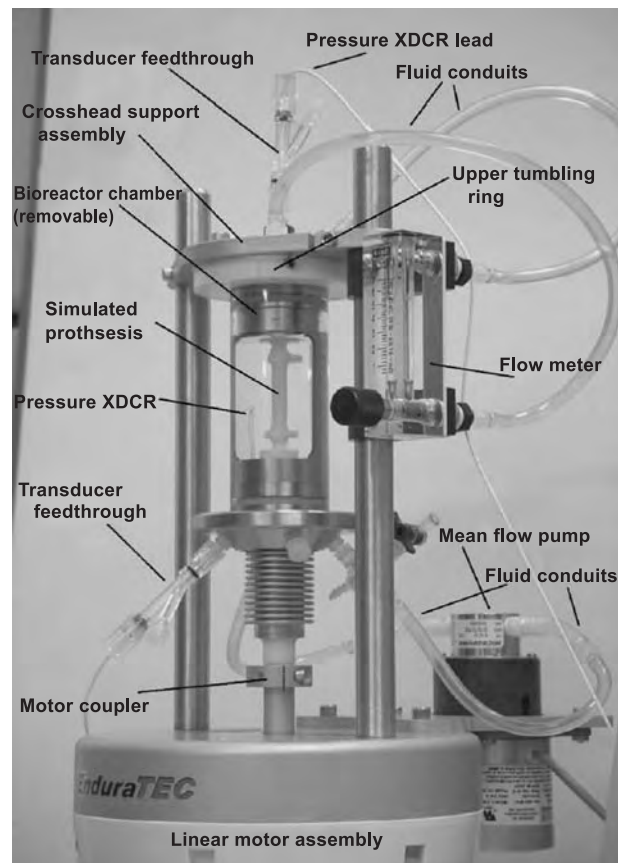


Fig. 7 Close-up of BioReactor showing components, transducers, and feed-throughs. The typical bioreactor chamber utilizes stainless (or titanium), plastic, and glass components. Each chamber can be readily disassembled for ETO or autoclave sterilization.

chambers and dynamic pumps are fabricated from titanium, stainless steel, delrin, and glass for long life and durability. The stand-alone device is designed to minimize contamination and allow experimentation in a nutrient environment. The biochamber is a modular design that allows rotational cellular seeding of the scaffold independent of the stimulation system. In addition, the chamber is a high-visibility design that allows optical monitoring of the construct, and it is ported to provide independent internal and external flow of nutrient solution to the intravascular construct (Fig. 10).

Pulsatile Pressure Control

Mean flow and superimposed dynamic flow can be varied independently to provide a full range of hydrodynamic loading conditions. The combination of dynamic and mean flow pumps allows the researcher to create almost any flow waveform, and an example of the achievable waveforms can be found in Fig. 11a.



Fig. 8 Six-station BioReactor configuration showing individual chambers and mean pressure pumps.

For growing tissue-engineered medical products such as arteries, combined dynamic and mean flows can be utilized to deliver the wall shear stress that duplicates the average wall shear stress deemed suitable for stimulation. The mean flow can be adjusted from 5 to 25 L/min while pulse rate and dynamic bellows pump displacements can be varied to characterize the effect of varying flow conditions on construct morphology and function. Typical dynamic displacement and pressure range are 6.4 mL/pulse and 0–300 mm Hg respectively. The integrated mean flow pump is a commercial gear variety, and it features a DC brushless motor and reusable gear head pump.

Integrated Control Electronics

The ITEMS BioReactor system uses a controller that is implemented on a Windows-based personal computer and provides great flexibility for exploring different stimulating conditions, transducer measurement, and algorithm development. The mechanical stimulation parameters that can be controlled open up new possibilities to understand their effects on the formation of tissue with improved quality.

Measurement of Material Properties

In addition to reduced cell-growth time periods, a major desired outcome of mechanical stimulation is to create enhanced material properties. These might include resistance-to-burst characteristics and a more

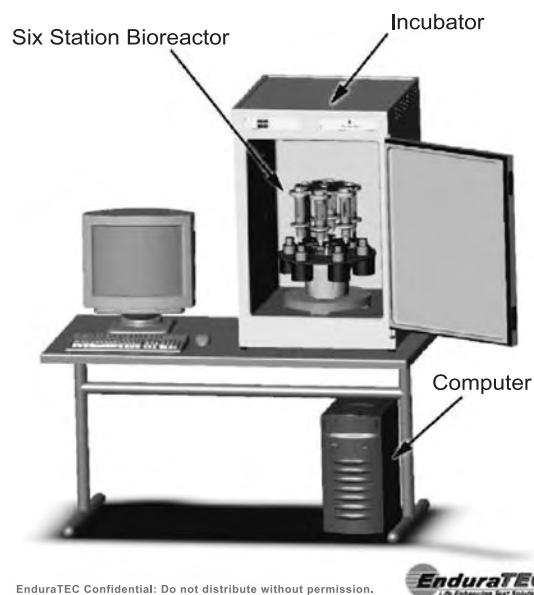


Fig. 9 Overview of the six-station BioReactor within an incubator showing the compact design of the BioReactor.

suitable modulus of elasticity. The ITEMS BioReactor has the capability to measure the differential pressure between the inside and outside of the construct using catheter pressure transducers. The resistance to burst is easily determined using these two readings. In addition, combining this pressure information with the applied flow pulse information allows a pseudo determination of the relative stiffness (modulus) can be made. This provides some feedback to the researcher about how the stimulation may be affecting the modulus; see Fig. 11b. The BioReactor can display these measurements automatically, which will allow rapid progress in understanding which stimulation parameters have the greatest impact on tissue strength.

Use of Material Properties As Input to Algorithms That Can Automatically Adjust Applied Conditions

Once the pseudo determination of relative stiffness is understood, the researcher can develop control algorithms for optimizing the applied conditions. The computerized monitoring and control system enables the researcher to program these algorithms into the BioReactor's stimulating system so that the computer can make changes to the conditioning sequence based on the pseudo stiffness measurements on a real-time basis. The ability to adjust parameters and conditioning sequences is a dramatic improvement to existing practices. With this innovation, it is possible to optimize the stimulation input to the condition of the

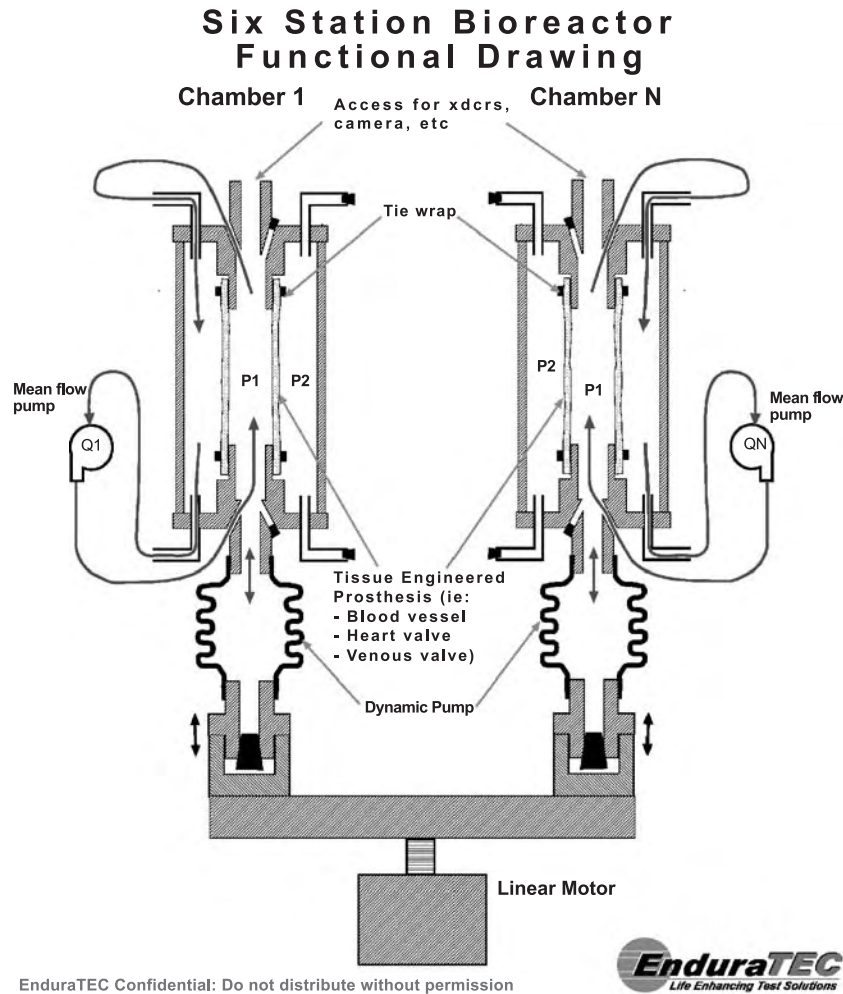


Fig. 10 Functional schematic of the multiple-chamber BioReactor. The movement of the bellows provides the pulsatile pressures. Each chamber has its own, independently controlled mean flow pumps.

construct, leading to significantly improved tissue. A potential added benefit is that the tissue may also be produced in a shorter amount of time because of the optimization algorithms. The adjustment of more

than one material measurement may also accelerate the tissue development, saving additional time. The ITEMS BioReactor has the capability to measure and control multiple factors.

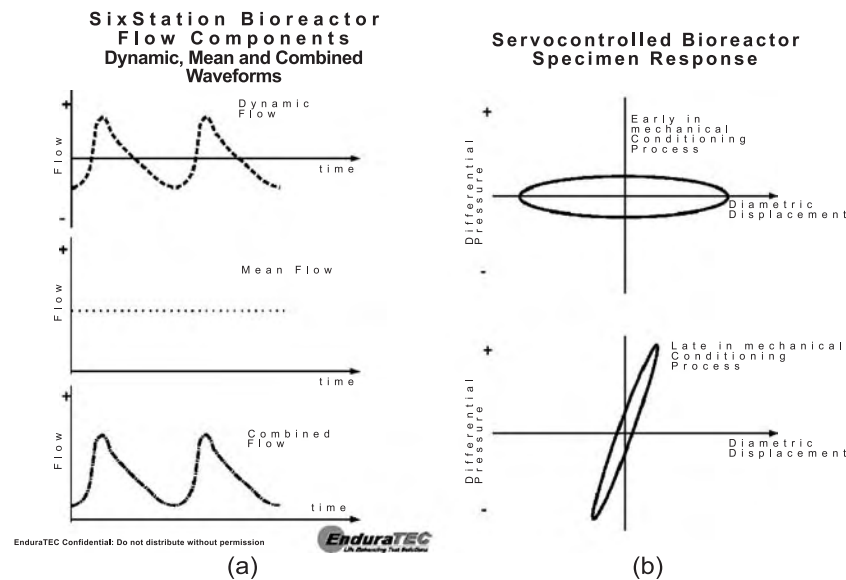


Fig. 11 (a) The dynamic, mean, and combined flow profiles that are used for stimulating the vascular construct. (b) The early and late mechanical conditioning process. The hysteresis loops for the late conditioning process will exhibit one to several orders of magnitude in stiffness as compared to the original construct properties.

Pressure/Flow Measurement Instrumentation

The ITEMS BioReactor system features catheter-based pressure and flow measurement transducers. This means the researcher is able to accurately measure the mechanical stimulating conditions that are applied to the construct, and with this capability quantitative studies can be performed to understand the impact of these factors. The measurement of internal conduit conditions is critical to understanding the tissue growth process, and as a result, more detailed discussion of the measurement approaches are presented.

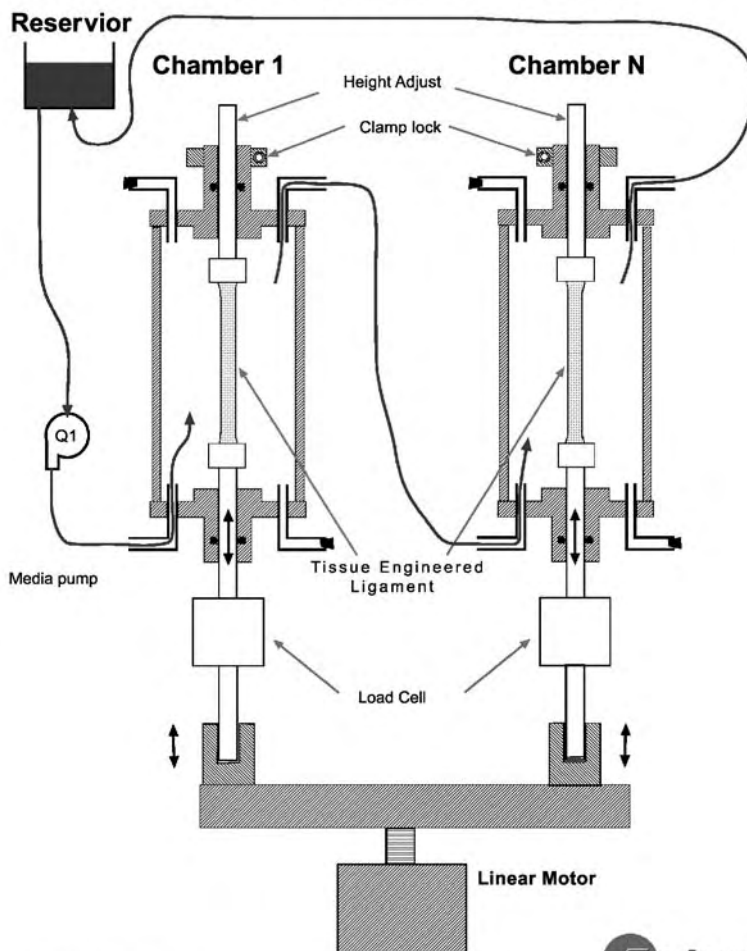
The ITEMS BioReactor utilizes internal and external pressure transducers to measure differential pressure conditions across a vascular conduit. These are shown in Figs. 7 and 10. The fluid flow velocities are measured using the mean flow pump input voltage, then differentiating the dynamic pump motion multiplied by the pump area, and then summing these two components.

Measurement of the flow conditions at the input and output of the construct's inner lumen will provide better flow information, and these two readings could be used to measure pseudo material properties. For example, if the construct has a high radial compliance, there will be a substantial phase and amplitude difference between the dynamic velocity readings at these two locations. This is equivalent to a fluid bulge working its way down a soft hose.

Methods for measuring the internal pressure conditions from outside the BioReactor chamber using a laser- or ultrasonic-based method are also being investigated. Transducers that could measure the diameter change of the construct with each pulse would provide a direct measurement of the radial compliance of the construct. These measurements would be useful in more accurately ascertaining the construct's material properties.

The approach of measuring radial strains directly is used in the EnduraTEC 9100 Series Stent Graft Test Instruments. Although these Stent Graft systems use a

Multi Station Ligament Bioreactor Functional Drawing



EnduraTEC Confidential: Do not distribute without permission

EnduraTEC
Life Enhancing Test Solutions

Fig. 12 This figure details the functional design for ligaments. Single or independent mean flow pumps can be provided. Manual length adjustment is required to properly insert each ligament. Axial motion and torsion motion are accomplished via the lower unit assembly, which utilizes either two linear motors or a linear motor and rotary motor.

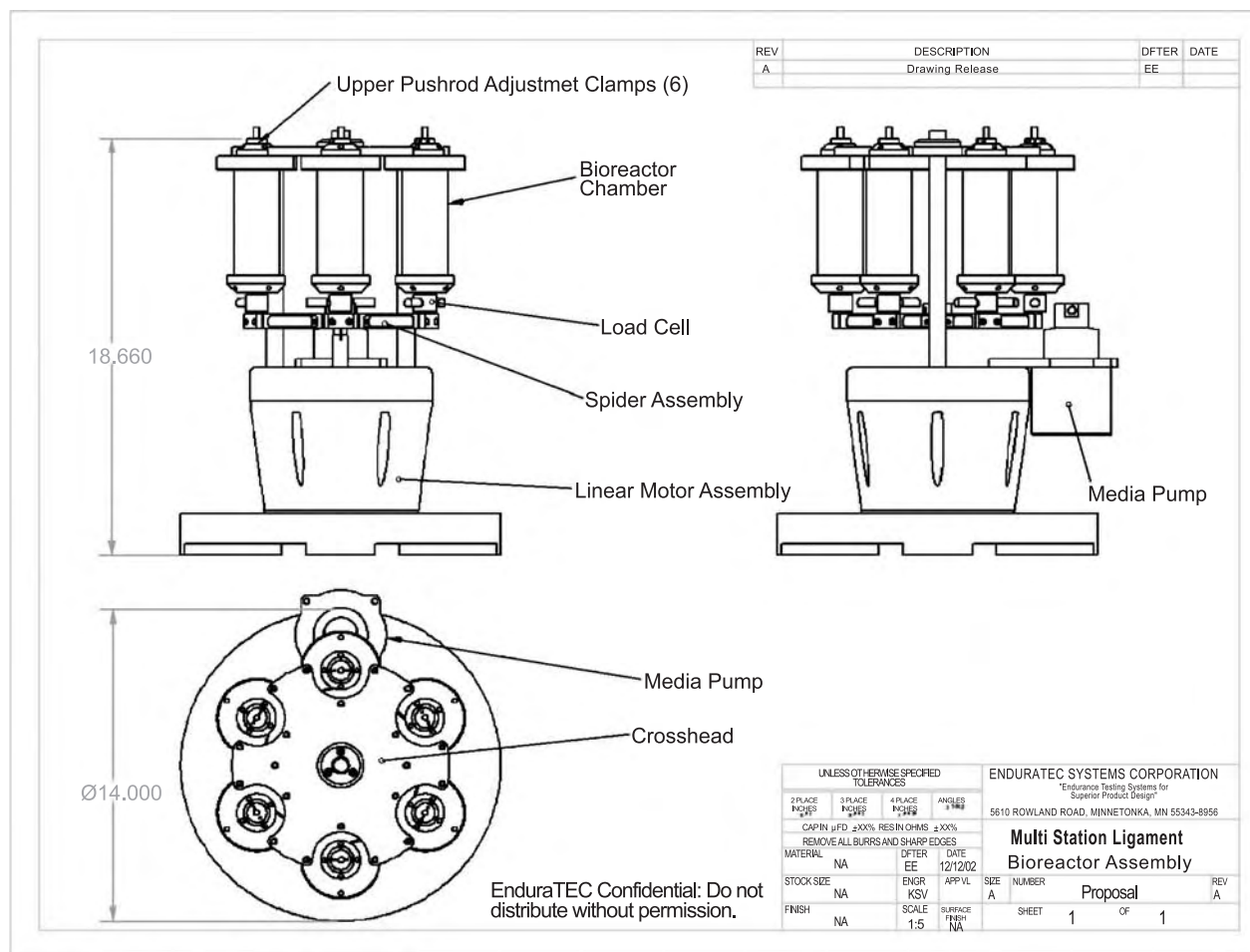


Fig. 13 This drawing describes the individual components along with overall dimensions of the ligament BioReactor.

laser micrometer to make the diametral measurements, and the same micrometer could be used in the ITEMS BioReactor, the use of ultrasonic sensors as the means for this measurement is a more attractive possibility. EnduraTEC believes that ultrasonic transducers can be made more robust and could be fabricated for hundreds of dollars versus thousands of dollars for a laser system. These measurement options are important considerations for researchers depending on their specific requirements.

Multistation Configuration—Multiple BioReactor Chambers

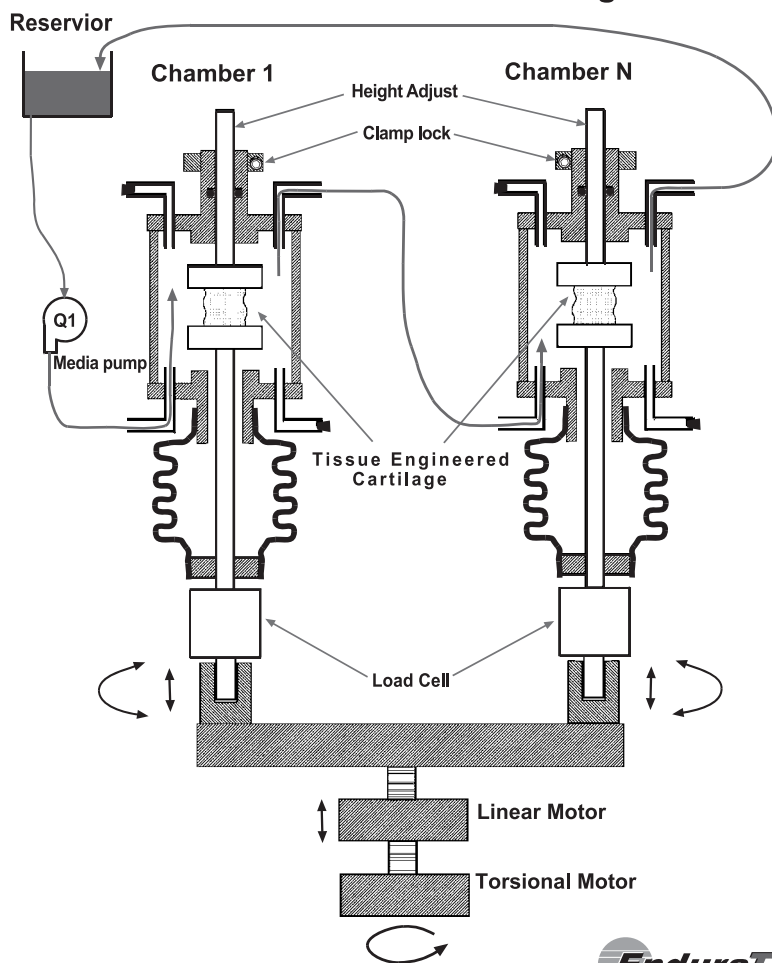
A multiple-chamber BioReactor configuration is available so that multiple experiments can be performed simultaneously. In an effort to minimize the expense associated with the linear pump motor and controls, the multistation BioReactor has six chambers with dynamic pumps that are driven from a common linear motor. Each chamber has its own mean flow pump,

which enables the researcher to vary that parameter if desired.

BIOREACTOR DESIGN CONSIDERATIONS FOR ORTHOPEDIC (LIGAMENT AND CARTILAGE) APPLICATIONS

The ITEMS BioReactor can be configured for ligament and cartilage applications and has several important features for these types of tissues. Concept drawings of the ligament accessories are shown in Figs. 12–15. The moving magnet linear motor design remains as the heart of the system for both ligament and cartilage applications. However, higher forces for the multispecimen configuration requires dual linear motors. Additionally, cartilage stimulation requires shearing motion (either simple lateral or pure torsional shearing) as part of the design. A third motor is needed to provide this shearing motion. For vascular tissue and ligaments, torsional loading is an option that is easily incorporated.

Multi Station Cartilage Bioreactor Functional Drawing



EnduraTEC Confidential: Do not distribute without permission



Fig. 14 This drawing details chamber modifications to accommodate small cartilage samples. The chamber is shorter in length, and the bellows arrangement provides for both axial loading and lateral motion (simple shear). As in the ligament, a common reservoir with a single or independent mean flow pumps can be provided.

Mechanical Stimulation Subassembly

The axial-loading subassembly utilizes dual linear motors to provide a maximum of 450 newtons of force. The total load is distributed to the six stations through a spider load distribution subassembly so that each specimen can be subjected to as much as 75 newtons. Each specimen station will have its own dedicated load cell for measurement of the exact applied force (see Figs. 12 and 13).

For cartilage, shear loading is provided by another linear motor that rotates the spider assembly to provide an effective linear displacement of ± 3 mm of the bottom compression platen in each of the chambers. This motion provides shear forces of up to 58 newtons per sample (see Figs. 14 and 15).

The flexible bellows design allows both vertical and lateral displacement. There is a chamber through-hole that has sufficient clearance to allow lateral displacement (shear) of the bottom piston when the spider assembly is programmed for rotational motion.

The piston assembly is designed with the appropriate stiffness to minimize any lateral displacement error.

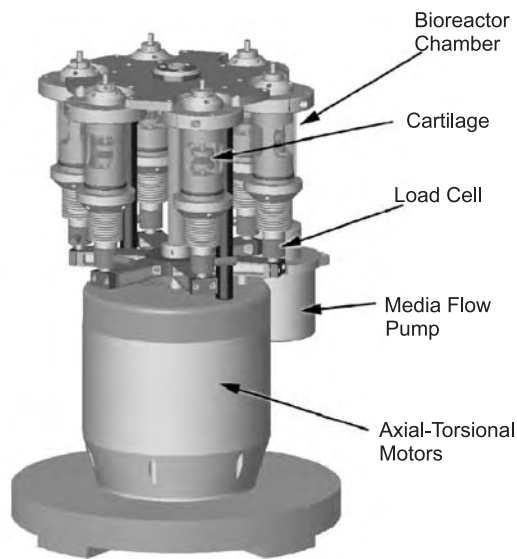
Modified BioReactor Chambers

Each transparent chamber will house gripping and piston assemblies for dynamic loading of the ligament or cartilage scaffold as shown in Figs. 12–15. The chambers will incorporate commercially available seals that are best rated for low pressure, short dynamic strokes, low friction, and biocompatibility to the nutrient matrix.

For ligaments, each pull/push rod is attached to a load cell below the chamber to measure load. The top push rod will have a manual slack adjustment to accommodate the length of the scaffold. Manual locking mechanisms for the push/pull rods are provided to facilitate moving/transfer of the BioReactor chamber after assembly, seeding (if seeding is done remotely),

B

Multi Station Cartilage Bioreactor Assembly



EnduraTEC Confidential: Do not distribute without permission.



Fig. 15 This Solidworks™ drawing details the main components of the cartilage BioReactor configuration. The lower unit assembly that houses the linear and rotary motors makes the unit taller but maintains overall functionality to be used within an incubator as the cardiovascular and ligament configuration.

and placement into the loading spider carousel of the BioReactor. These locking mechanisms are similar for both ligament and cartilage applications.

It is intended that the whole chamber assembly can be easily removed so that a chamber and scaffold that has been seeded at a separate processing station can be

quickly inserted into the ITEMS BioReactor. In the case where the chambers are routinely moved between incubator and into the BioReactor, the locking mechanisms can act as “zero” start positions and prevent looking for the zero electronically.

Each chamber is serially connected to a mean flow pump (for nutrient/environment media), which enables the researcher to vary that parameter if desired. The chambers could also be connected in parallel to the mean flow pump. For those applications where independent pumps are required, this capability can be provided as an option. This could be desirable to minimize any contamination between chambers, or if different fluid media are required for each chamber.

It is intended that the multiple-chamber configuration may be removed as a unit. The intent, in this case, is that mechanical stimulation could be set up as a scheduled periodic activity. The six-chamber subassembly could reside in another incubator and then be placed on the ITEMS BioReactor for a stimulation routine. This feature allows the use of multiple six-chamber assemblies, thus providing cost-effective/productive tissue-engineering practices.

Gripping and Attachment Considerations

The attachment of the ligament scaffold or construct to the stimulation mechanism is an area that requires significant collaboration with the researcher. EnduraTEC has created several mechanical designs that have been successful; however, each application may require some variation. Some situations may entail allowing the tissue to grow directly onto the gripping mechanism; in other cases, either attaching to or mechanically gripping the ends of the scaffold may be appropriate. For cartilage, simple compression platens (10–15 mm in diameter) with detent for scaffold positioning are provided.

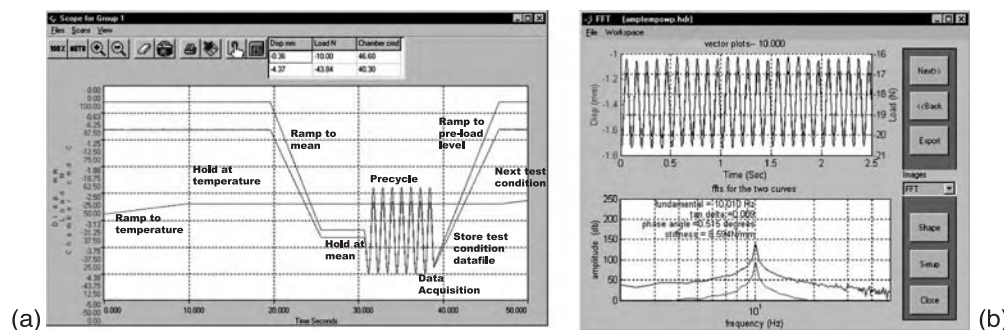


Fig. 16 (a) Graph showing how DMA, or dynamic mechanical analysis, frequency, dynamic amplitude, mean level, and the control mode can be defined. (b) A typical DMA plot showing results at a particular fundamental frequency.

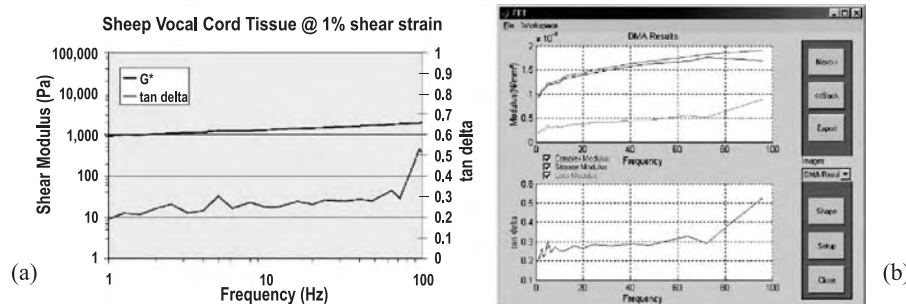


Fig. 17 These figures show DMA results for sheep vocal cord tissue. (a) A 50-gram cell was used to measure forces in the shear direction. This testing was performed on the same linear motor technology described by the BioReactor so that it could be compared to potential tissue-engineered cord tissue. (b) Compression DMA testing on porcine fold was also performed during this exercise. Compression modulus is higher than existing data in shear or tension due primarily to the pre-compressive stress required to maintain contact during testing. A 45-N cell was used for these compression tests.

B

Measurement Instrumentation

The ITEMS BioReactor for ligaments and cartilage utilizes a mean flow internal pressure transducer to measure pressure conditions of the nutrient environment. The fluid flow velocity is measured using the mean flow pump input voltage. In addition, the loading mechanism for each chamber will have a load cell for measuring dynamic loading imposed to the ligament specimen. For overall operation, the six sensors will be averaged to get the control command desired. Lateral or torsional shear measurements are possible.

Material/Construct Properties

Unlike cardiovascular applications, Dynamic Characterization Analysis (DMA) applications software is available. It uses load, displacement, or strain information to evaluate how the scaffold and scaffold/tissue construct is performing during its growth/stimulation stages. By performing this material characterization exercise, the constructs' viscoelastic properties such as dynamic modulus, storage modulus, and tan delta can be readily calculated. With DMA software, the researcher can define frequency, dynamic amplitude, mean level, and the control modes that are most

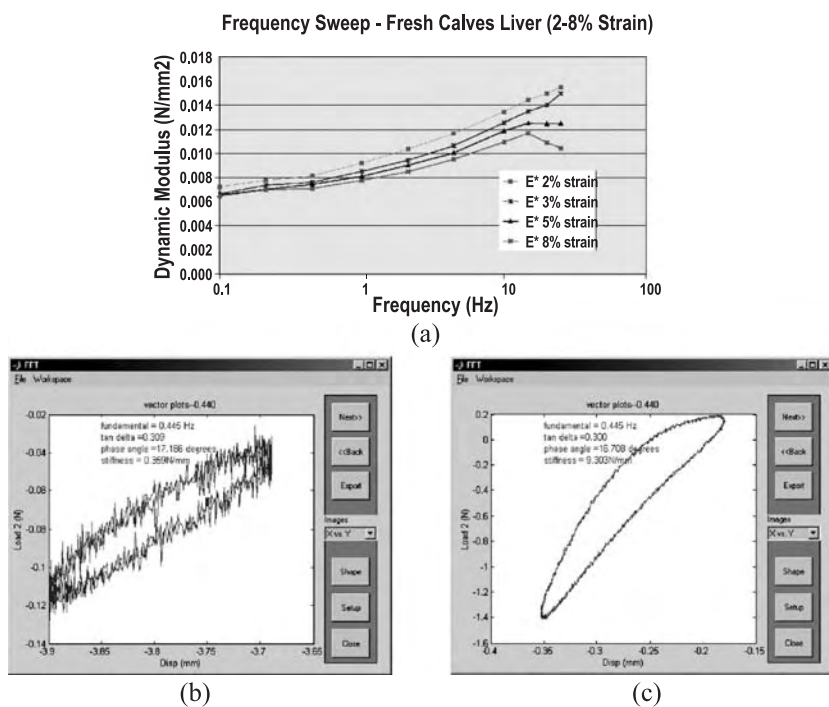


Fig. 18 These graphs show another example of soft tissue (liver) DMA results. Programming both frequency and strain sweeps is relatively easy to perform once the test protocol is established. Typical loads at 1% strain were about 0.2 N (200 g). Figures (b) and (c) show the hysteresis curves for the same sample as (a). In (c), the liver was ablated (radio frequency) to correlate viscoelastic properties derived from DMA to sonoelastography measurements. There are two order-of-magnitude differences in the stiffness between the two curves. (b) shows a relatively flat hysteresis curve due to its very low stiffness and shows more “noise” in the measured feedback to extremely low loads, as would be seen in the initial stages of tissue growth. As the tissue/scaffold gets stronger/stiffer, hysteresis curves as dramatic as the ablated example may be measured.

applicable to characterize the tissue. The graphs in Fig. 16 describe typical data presented after DMA analysis (via FFT) is performed. Figs. 17 and 18 show viscoelastic data on sheep vocal cord material and calf liver. The linear motor is ideal for characterizing such soft tissue. This data would be the baseline information (reference curves) that one would compare the properties of tissue-engineered liver and tissue-engineered vocal cord. Ideally, by using these reference curves and querying the tissue-engineered medical product during the growth cycle, and then making comparisons, optimized growth via manual or automatic adjustments to the stimulation process can be realized.

SUMMARY

A flexible tool such as the ITEMS BioReactor provides well-defined mechanical forces during construct development to enhance the biomolecular composition and mechanical function of tissue-engineered medical products such as arteries, ligaments, and cartilage. It can be utilized in fundamental studies to develop precise quantitative relationships between chemical and mechanical environments and construct growth. In addition, with this second-generation BioReactor, researchers can begin to address some of the concerns regarding the production of large numbers of uniform constructs suitable for human implantation.

A bioreactor with the ability to grow multiple constructs independently under well-defined hydrodynamic conditions represents an important advancement for functional tissue engineering. Since each chamber is independent, cells, scaffold, nutrients, and mechanical loading can be varied between constructs, providing the development of quantitative relationships between hydrodynamic or mechanical loading on cell proliferation and matrix deposition.

The use of a scalable, intelligent bioreactor that is commercially produced and supported will allow tissue-engineering researchers to devote more time and resources to rapidly move forward with their functional tissue engineering research programs. With a highly functional platform that can be easily configured for either vascular or orthopedic constructs, the uniform design attributes of the ITEMS BioReactor will allow researchers to readily compare their tissue engineering results in the hope that functional tissue engineering milestones can be greatly expedited.

The major design elements of the ITEMS BioReactor are 1) a simple geometry to allow for accurate modeling and visualization of the internal (luminal) and external flow fields and mass transfer, 2) compact and easy-to-use biochamber designs for different applications, and 3) advanced control and measurement capabilities to evaluate material properties and adjust

the conditioning parameters in order to optimize the stimulation input to the construct, leading to significantly improved tissue.

ARTICLES OF FURTHER INTEREST

Blood Purification; Cartilage Regeneration; Extracellular Matrix Scaffolds; Rotary Cell Culture Systems; Tissue Engineering; Tissue Engineering of Blood Vessel; Tissue Engineering of Ligament; Tissue Engineering of Rotator Cuff Tendons

REFERENCES

1. L'Heureux, N.; Paquet, S.; Labbe, R.; Germain, L.; Auger, F.A. A completely biological tissue-engineered human blood vessel. *FASEB J.* **1998**, *12*, 47–56.
2. Shinoka, T.; Shum-Tim, D.; Ma, P.X.; Tanel, R.E.; Isogai, N.; Langer, R.; Vacanti, J.P.; Mayer, J.E., Jr. Creation of viable pulmonary artery autografts through tissue engineering. *J. Thorac. Cardiovasc. Surg.* **1998**, *115*, 536–545.
3. Shum-Tim, D.; Stock, U.; Hrkach, J.; Shinoka, T.; Lien, J.; Moses, M.A.; Stamp, A.; Taylor, G.; Moran, A.M.; Landis, W.; Langer, R.; Vacanti, J.P.; Mayer, J.E., Jr. Tissue engineering of autologous aorta using a new biodegradable polymer. *Ann. Thorac. Surg.* **1999**, *68*, 2298–2304.
4. Niklason, L.E.; Gao, J.; Abbott, W.M.; Hirschi, K.K.; Houser, S.; Marini, R.; Langer, R. Functional arteries grown in vitro. *Science* **1999**, *284*, 489–493.
5. Niklason, L.E.; Abbott, W.; Gao, J.; Klagges, B.; Hirschi, K.K.; Ulubayram, K.; Conroy, N.; Jones, R.; Vasanawala, A.; Sanzgiri, S.; Langer, R. Morphologic and mechanical characteristics of engineered bovine arteries. *J. Vasc. Surg.* **2001**, *33*, 628–638.
6. Kobashi, T.; Matsuda, T. Fabrication of compliant hybrid grafts supported with elastomeric meshes. *Cell Transplant.* **1999**, *8*, 477–488.
7. Weinberg, C.B.; Bell, E. A blood vessel model constructed from collagen and cultured vascular cells. *Science* **1986**, *231*, 397–400.
8. Ratcliffe, A.; Niklason, L.E. Bioreactors and bioprocessing for tissue engineering. *Ann. NY Acad. Sci.* **2002**, *961*, 210–215.
9. Altman, G.H.; Horan, R.L.; Martin, I. Cell differentiation by mechanical stress. *FASEB J. Vol. 16*. Express article 10.1096/fj.01-0656fje. Published online December 28, 2001.
10. Altman, G.H.; Stark, P.; Martin, I.; Richmond, J.C.; Vunjak-Novakovic, G. Advanced Bioreactor with Controlled Application of Multi-Dimensional Strain for Tissue Engineering. In *Transactions of ASME; American Society for Mechanical Engineers*, Dec. 2003; Vol. **124**, 742–749.
11. Mol, A.; Vouten, C.V.C.; Zunde, G.; Gunter, C.I.; Visjager, J.F.; Turina, M.I.; Baaijens, F.P.T.; Hoerstrup, S.P. The relevance of large strains in functional tissue engineering of heart valves. *Thorac.*

- Cardiovasc. Surg. **2003**, *51*, 78–83. © Georg Thieme Verlag Stuttgart, New York ISSN 0171-6425.
12. Martinek, V.; Fuh, F.H.; Huard, J. *Gene Therapy and Tissue Engineering in Sports Medicine: The Physician and Sports Medicine*; McGraw Hill, February 2000; Vol. 28, No. 2.
 13. Miyasaka, K.C.; Daniel, D.M.; Hirschman, P. The incidence of knee ligament injuries in the general population. *Am. J. Knee Surg.* **1991**, *4* (1), 3–8.
 14. O'Driscoll, S.W. The healing and regeneration of articular cartilage. *J. Bone Jt. Surg., (Am)* **1998**, *80* (12), 1795–1812.
 15. Shapiro, F.; Koide, S.; Glimcher, M.J. Cell origin and differentiation in the repair of full-thickness defects of articular cartilage. *J. Bone Jt. Surg., Am* **1993**, *75* (4), 532–553.
 16. Solchaga, L.A.; Dennis, J.E.; Goldberg, V.M. Hyaluronic acid-based polymers as cell carriers for tissue-engineered repair of bone and cartilage. *J. Orthop. Res.* **1999**, *17* (2), 205–213.
 17. Picciolo, G.L. Enabling biomaterials technology for tissue engineering: Introduction. *Tissue Eng.* **1997**, *3* (1), 67–70.
 18. Godbey, W.T.; Atala, A. In vitro systems for tissue engineering. *Ann. N.Y. Acad. Sci.* **2002**, *961*, 10–26.
 19. Hoerstrup, S.P.; Sodian, R.; Sperling, J.S.; Vacanti, J.P.; Mayer, J.E., Jr. New pulsatile bioreactor for in vitro formation of tissue engineered heart valves. *Tissue Eng.* **Feb. 2000**, *6* (1), 75–79.
 20. Sodian, R.; Lemke, T.; Loebe, M.; Hoerstrup, S.P.; Potapov, E.V.; Husmann, H.; Meyer, R.; Hetzer, R. New pulsatile Bioreactor for fabrication of tissue-engineered patches. *J. Biomed. Mater. Res.* **2001**, *58* (4), 401–405.
 21. Dumont, K.; Yperman, J.; Verbeken, E.; Segers, P.; Meuris, B.; Vandenberghe, S.; Flameng, W.; Verdonck, P.R. Design of a new pulsatile bioreactor for tissue engineered aortic heart valve formation. *Artif. Organs* **2002**, *26* (No. 8), 710–714.
 22. Demarteau, O.; Jakob, M.; Schafer, D.; Heberer, M.; Martin, I. *Development and Validation of a Bioreactor for Physical Stimulation of Engineered Cartilage*; Biorheology, IOS Press, 2003; Vol. 40, 331–336.
 23. Hoerstrup, S.P.; Sodian, R.; Daebritz, S.; Wang, J.; Bacha, E.A.; Martin, D.P.; Moran, A.M.; Guleserian, K.J.; Sperling, J.S.; Kaushal, S.; Vacanti, J.P.; Schoen, S.J.; Mayer, J.E. Functional living trileaflet heart valves grown in vitro. *Circulation* **2000**, *102* (suppl III), III 44–III 48. American Heart Association, Inc.
 24. Sodian, R.; Hoerstrup, S.P.; Sperling, J.S.; Daebritz, S.H.; Martin, D.P.; Shoen, F.J.; Vacanti, J.P.; Mayer, J.E., Jr. Tissue engineering of heart valves: In vitro experiences. *Ann. Thorac. Surg.* **2000**, *70* (1), 140–144.
 25. Hoerstrup, S.P.; Kadner, A.; Melnitchouk, S.; Trojan, A.; Eid, K.; Tracy, J.; Sodian, R.; Visjager, J.F.; Kolb, S.A.; Grunenfelder, J.; Zund, G.; Turina, M.I. Tissue engineering of functional trileaflet heart valves from human marrow stromal cells. *Circulation* **2002**, *106* (Suppl 1), I143–I150.
 26. Hoerstrup, S.P.; Kadner, A.; Breyman, C.; Maurus, C.F.; Guenter, C.I.; Sodian, R.; Visjager, J.F.; Zund, G.; Turina, M.I. Living, autologous pulmonary artery conduits tissue engineered from human umbilical cord cells. *Ann. Thorac. Surg.* **2002**, *74*, 46–52.
 27. Hoerstrup, S.P.; Zund, G.; Sodian, R.; Schnell, A.M.; Grunenfelder, J. Tissue engineering of small caliber vascular grafts. *Eur. J. Cardio-Thorac. Surg.* **2001**, *20* (1), 164–169.
 28. Jockenhoovel, S.; Zund, G.; Hoerstrup, S.P.; Schnell, A.; Turina, M. Cardiovascular tissue engineering: A new laminar flow chamber for in vitro improvement of mechanical tissue properties. *ASAIO J. Jan–Feb.* **2002**, *48* (1), 8–11.
 29. Kouri, J.B.; Arguello, C.; Luna, J.; Mena, R. Use of microscopical techniques in the study of human chondrocytes from osteoarthritic cartilage: An overview. *Microsc. Res Tech.* **1998**, *40*, 22–36.
 30. Aydelotte, M.B.; Schumacher, M.A.; Kuettner, K.E. Heterogeneity of Articular Cartilage. In *Articular Cartilage and Osteocartilage*; Kuettner, K.E., Schleyerbach, R., Peyron, J.G., Hascall, V.C., Eds.; Raven Press: New York, 1991; 237–249.
 31. Hunziker, E.B.; Quinn, T.M.; Hauselmann, H.J. Quantitative structural organization of normal adult human articular cartilage. *Osteoarthr. Cartil.* **2002**, *10*, 564–572.
 32. Mori, R.; Ochi, M.; Sakai, Y.; Maniwa, S.; Kawasaki, K. Quantitative cell and matrix changes in human humeral head osteoarthritic cartilage. *APMIS* **2000**, *108*, 642–648.
 33. Wantanabe, M.; Leng, C.-G.; Toriumi, H.; Hamada, Y.; Akamatsu, N.O.S. Ultrastructural study of upper surface layer in rat articular cartilage by “in vivo cryotechnique” combined with various treatments. *Med. Electr. Microsc.* **2000**, *33*, 16–24.
 34. Freed, L.E.; Vunjak-Novakovic, G. Tissue Engineering Bioreactors. In *Principles of Tissue Engineering*; Lanza, R., Langer, R., Vacanti, J., Eds.; Academic Press: San Diego, 2000; 143–156.
 35. Freed, L.E.; Vunjak-Novakovic, G. Culture Environments: Cell-Polymer-Bioreactor Systems. In *Methods of Tissue Engineering*; Atala, A., Lanza, R.P., Eds.; Academic Press: San Diego, 2002; 97–111. Chapter 26. .
 36. Lelkes, P.I.; Unsworth, B.R. *Neuroectodermal Cell Culture: Endocrine Cells. Methods of Tissue Engineering*; Atala, A., Lanza, R.P., Eds.; Academic Press: San Diego, 2002; 371–382. Chapter 29.
 37. Vastag, B. Cell biology update: A decade of simulating space on earth. *JAMA* **May 2, 2001**, *285* (No 17), 2181–2182.
 38. Moreau, D.; Neitzel, G.P.; Pennecot, J.; Smith, M.K.; Sucusky, P. *Fluid Dynamics of Bioreactors for Tissue-engineered Cartilage*; Proceedings of the 2002 IEEE Engineering in Medicine and Biology and 24th Annual Conference and the 2002 Fall Meeting of the Biomedical Engineering Society, 2002; Vol. 1, 875–876.
 39. Niklason, L.E.; Seruya, M. Small-Diameter Vascular Grafts. In *Methods of Tissue Engineering*; Atala, A., Lanza, R.P., Eds.; Academic Press: San Diego, 2002; 905–913. Chapter 80.
 40. Risau, W.; Flamme, I. Vasculogenesis. *Annu. Rev. Cell Dev. Biol.* **1995**, *11*, 73–79.

41. Ingber, D.E. Mechanical and Chemical Determinants of Tissue Development. In *Principles of Tissue Engineering*; Lanza, R., Langer, R., Vacanti, J., Eds.; Academic Press: San Diego, 2002; 101–109. Chapter 9.
42. Vorp, D.A.; Severyn, D.A.; Steed, D.L.; Webster, M.W. A device for the application of cyclic twist and extension on perfused vascular segments. *Am. J. Physiol.* **1996**, *270*, H 787–H795. (Heart Circ. Physiol. 39).
43. Banes, A.J.; Link, G.W., Jr; Gilbert, J.W.; Tran Son Tay, R.; Monbureau, O. Culturing cells in a mechanically active environment: The Flexercell Strain Unit can apply cyclic or static tension or compression to cells in culture. *Am. Biotechnol. Lab.* **May 1990**, *8* (7), 12–22.
44. Kasra, M.; Goel, V.; Martin, J.; Wang, S.-T.; Choi, W.; Buckwalter, J. Effect of Dynamic Hydrostatic Pressure on Rabbit Intervertebral Disc Cells. In *Journal of Orthopedic Research*; Elsevier, 8 February 2003. DTD 4.3.1 8 pp.
45. Vilender, K. Bioreactor with Plurality of Chambers for Conditioning Intravascular Tissue Engineered Medical Products. Patent Pending: US Patent Application Serial Number 10/371175, November 27, 2002.
46. McAllister; Todd, N.; L'Heurex, N. Cytograph Tissue Engineering, Inc., La Jolla CA (US). *Improved Tissue Engineered Blood Vessels and Methods and Apparatus for their Manufacture*; World Intellectual Property Organization. WO 01/37884 A2 published May 31, 2001.
47. Dimitrijevič, D.S.; Dodd, C.W.; Anderson, W.J.; Schwarz, R.P. *Two Chamber Cell Culture Vessel*; World Intellectual Property Organization. WO 02/38735 A2, published May 16, 2002.

FURTHER READING

- Huard, J.; Li, Y.; Peng, H.R.; Fu, F.H. Gene therapy and tissue engineering for sports medicine. *J. Gene Med.* **Feb 2003**, *5* (2), 93–108.
- Kuriwaka, M.; Ochi, M.; Uchio, Y.; Maniwa, S.; Adachi, N.; Mori, R.; Kawasaki, K.; Kataoka, H. Optimum combination of monolayer and three-dimensional cultures for cartilage-like tissue engineering. *Tissue Eng.* **Feb 2003**, *9* (1), 41–49.
- Sah, R.; Kim, Y.; Doong, JY.; Grodzinsky, A.J.; Plass, A.H.; Sandy, J.D. BioSynthetic response of cartilage explants to dynamic compression. *J. Orthop. Res.* **1989**, *7*, 619–636.
- Toyoda, T.; Matsumoto, H.; Fujikwas, K.; Saito, S.; Inoue, K. Tensile load and the metabolism of anterior cruciate ligament cells. *Clin. Orthop. Relat. Res.* **1998**, *353*, 247–255.
- Woo, SL-Y.; Young, E.P.; Kwan, M.K. Fundamental Studies in Knee Ligament Mechanics. In *Knee Ligaments: Structure, Function, Injury, and Repair*; Daniel, D.M., Akeson, W.H., O'Connor, Eds.; Raven Press: New York, NY, 1990; 115–134.

Biorubber/Poly(Glycerol Sebacate)

Yadong Wang

Wallace H. Coulter School of Biomedical Engineering and School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, Georgia, U.S.A.

B

INTRODUCTION

Biodegradable polymers have significant potential in biotechnology and bioengineering. However, for some applications, they are limited by their mechanical properties' mismatch and unsatisfactory compatibility with cells and tissues. A strong, biodegradable, biocompatible elastomer could be useful in fields such as tissue engineering, drug delivery, and medical devices. Poly(glycerol sebacate) (PGS) is a tough biodegradable elastomer designed and synthesized from biocompatible monomers. PGS forms a covalently cross-linked three-dimensional network of random coils with hydroxyl groups attached to its backbone. Both cross-linking and the hydrogen bonding interactions between the hydroxyl groups likely contribute to the unique properties of the elastomer. *In vitro* and *in vivo* studies show the polymer has good biocompatibility. *In vivo* degradation shows a near linear loss of mass with time, and a gradual decrease of mechanical strength with very limited water uptake. Polymer implants under animal skin are absorbed completely within 60 days, followed by restoration of the implantation sites to their normal architectures. PGS supports the growth of many cell types, and may be useful in tissue engineering, medical devices, and other fields in biomedical engineering.

BACKGROUND

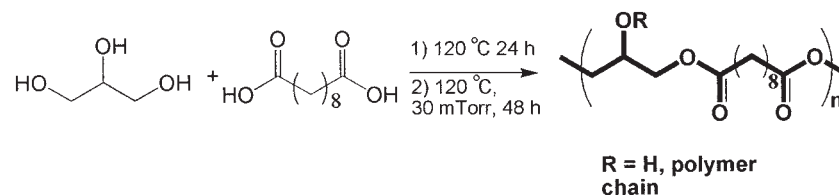
Advances in biomaterials have significantly impacted many fields in life sciences and technology, and improved the quality of life of tens of millions of individuals.^[1–3] Annual sale of medical systems using biomaterials in the United States alone exceeds \$100 billion.^[4] An increasingly important category of biomaterials is biodegradable polymers, which have been used widely in fields such as tissue engineering, drug delivery, and medical devices. Many biomedical devices are placed in a mechanically dynamic environment in the body, which requires the implants to sustain and recover from various deformations without mechanical irritations to the surrounding tissues. In many cases, the matrices and scaffolds of these implants would ideally be made of biodegradable

polymers whose properties resemble those of extracellular matrices (ECM)—soft, tough, and elastomeric in order to provide mechanical stability and structural integrity to tissues and organs. Hence a soft biodegradable elastomer that readily recovers from relatively large deformations is advantageous for maintaining the implant's proper function without mechanical irritation to the host. One such material is a series of biodegradable elastomers based on glycerol and sebacic acid, hence the name poly(glycerol sebacate).

These elastomers are biodegradable, biocompatible, tough, and inexpensive. Upon degradation *in vivo*, PGS loses mass almost linearly with time, and its mechanical strength decreases gradually with very limited water uptake. Unmodified PGS supports the growth of many different cell types, including endothelial and smooth muscle cells, Schwann cells, hepatocytes, cardiomyocytes, and chondrocytes. PGS has been fabricated into different physical forms such as flat sheets, foams, and tubes. Microfabricated PGS has also been used as a scaffolding material to make patent endothelialized branching capillary networks that may serve as artificial microvasculature for engineering vital organs. As one of the few biodegradable and biocompatible elastomers, PGS may be useful in soft tissue engineering, medical devices, and other fields in biomedical engineering.

DESIGN OF PGS

PGS is analogous to vulcanized rubber in that it forms a cross-linked three-dimensional network of random coils. Such a strategy to achieve tough and elastomeric materials is also found in nature; collagen and elastin, the major fibrous protein components of ECM, are both cross-linked. In addition to covalent cross-linking, hydrogen bonding interactions through hydroxyproline hydroxyl groups also contribute to the mechanical strength of collagen. The design of our polymer is based on two hypotheses: 1) good mechanical properties can be obtained through covalent cross-linking and hydrogen bonding interactions; and 2) rubberlike elasticity can be obtained by building a three-dimensional network of random coils through copolymerization whereby at least one monomer is



Scheme 1

trifunctional. To realize this design, the following criteria were considered: 1) degradation mechanism—hydrolysis was preferred to minimize individual differences in degradation characteristics caused by enzymes;^[2] 2) hydrolyzable chemical bond—ester was chosen for its established and versatile synthetic methods; 3) cross-link density—low density was preferred, as a high degree of cross-linking usually leads to rigid and brittle polymers; 4) cross-link chemical bonds were chosen to be hydrolyzable and identical to those in the backbone to minimize the possibility of heterogeneous degradation; and 5) specific monomers—they should be nontoxic, at least one should be trifunctional, and at least one should provide hydroxyl groups for hydrogen bonding. Glycerol [CH₂(OH)-CH(OH)CH₂OH], the basic building block for lipids, satisfies all three requirements and was chosen as the alcohol monomer. From the same toxicological and polymer chemistry standpoints, sebacic acid [HOOC(CH₂)₈COOH] was chosen as the acid monomer. Sebacic acid is the natural metabolic intermediate in ω -oxidation of medium- to long-chain fatty acids,^[5,6] and has been shown to be safe in vivo.^[7] Both monomers are inexpensive, an advantage for large-scale applications. This approach was likely to yield biodegradable polymers with improved mechanical properties and biocompatibility.

SYNTHESIS OF PGS

The polymer was synthesized by melt polycondensation of an equimolar amount of glycerol and sebacic acid at 120°C under argon for 24 h (Scheme 1). The pressure was then reduced from 1 Torr to 40 mTorr over 5 h. The polymer at this stage is a highly viscous liquid above 80°C, and is soluble in organic solvents such as tetrahydrofuran (THF), *N,N*-dimethylformamide, and ethanol. The polymer was cured at 40 mTorr and 120°C for 48 h. The resultant polymer is a transparent, almost colorless elastomer.^a The resulting polymer features a small amount of cross-links and

^aRigid, totally crosslinked polymer has been synthesized from glycerol and sebacic acid (glycerol/sebacic acid molar ratio: 2/3) under different conditions. See Ref. [8].

hydroxyl groups directly attached to the backbone. PGS is insoluble in water and swells only $2.1 \pm 0.33\%$ after soaking in water for 24 h. As is the case for other hydrolyzable polymers, PGS should be kept in an anhydrous environment for long-term storage.

CHARACTERIZATION OF PGS

The chemical and physical properties of PGS were characterized by various instrumental methods. Fourier transformed infrared spectroscopy (FTIR) on a newly prepared KBr pellet of PGS showed an intense C=O stretch at 1740 cm^{-1} , which confirmed the formation of ester bonds. The absence of C=O stretches corresponding to carboxylic acids suggested a high conversion rate to the polymer. FTIR also showed a broad intense OH stretch at 3448 cm^{-1} , indicating the hydroxyl groups were hydrogen-bonded (Fig. 1).

Differential scanning calorimetry (DSC) measurement showed two crystallization temperatures, -52.14°C and -18.50°C , and two melting temperatures, 5.23°C and 37.62°C , respectively. No glass transition temperature was observed above -80°C , the lower detection limit of the instrument. The DSC results indicate that the polymer was totally amorphous at 37°C .

Elemental analysis on vacuum-dried PGS samples confirmed the composition of the polymer as approximately 1 glycerol:1 sebacic acid. Water-in-air

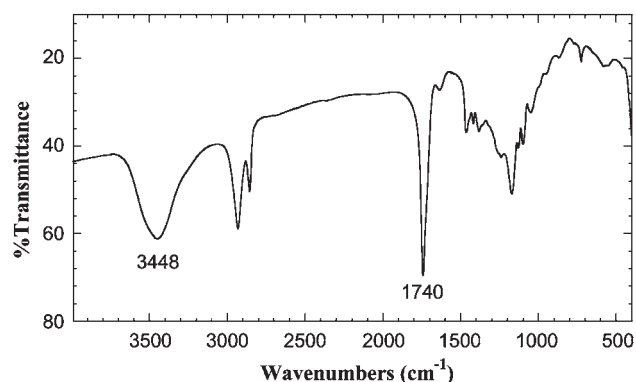


Fig. 1 FTIR spectrum of PGS. Note the intense C=O, and O-H stretches at 1740 cm^{-1} and 3440 cm^{-1} , respectively.

contact angle was measured at room temperature using the sessile drop method and an image analysis of the drop profile on slabs of polymer fixed on glass slides. The polymer surface is very hydrophilic due to the hydroxyl groups attached to its backbone. Its water-in-air contact angle was 32.0° , almost identical to that of a flat 2.7 nm thick type I collagen film (31.9°).^[9] The cross-linking density is expressed by n , moles of active network chains per unit volume, which was $38.3 \pm 3.40 \text{ mol/m}^3$, and M_c , the molecular weight between cross-links, which was $18,300 \pm 1620$, calculated from equation:^[10]

$$n = E_0/3RT = \rho/M_c$$

where E_0 is Young's modulus, R is the universal gas constant, T is the temperature, and ρ is the density. The low M_c value confirms that the cross-linking density is low in PGS. Depending on the monomer ratio and process parameters, the modulus of PGS can be controlled within at least an order of magnitude. On the low modulus (soft) limit, tensile tests on thin strips of PGS revealed a stress-strain curve characteristic of an elastomeric and tough material (Fig. 2). The nonlinear shape of the tensile stress-strain curve is typical for elastomers, and resembled those of ligament^[11-13] and vulcanized rubber.^[14] PGS can be elongated repeatedly to at least three times its original length without rupture. The total elongation is unknown, as grip breaks occurred at $267 \pm 59.4\%$ strain. The tensile Young's modulus of the polymer was $0.282 \pm 0.0250 \text{ MPa}$ with an ultimate tensile strength greater than 0.5 MPa (grip break). The value of the Young's modulus of PGS is

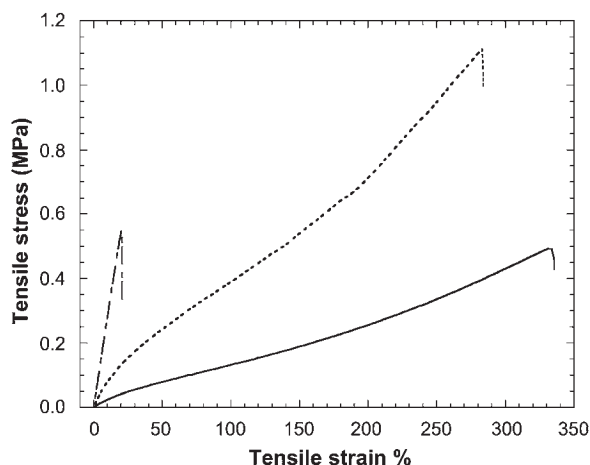


Fig. 2 Stress strain curves of two extremes of PGS, the soft (-----) and the stiffer one (.....), and vulcanized rubber (—). Both soft PGS and vulcanized rubber are marked by low modulus and large elongation ratio, indicating elastomeric and tough materials. The modulus of the stiffer PGS is 10 times higher than that of the soft one. Grip break for both PGS.

between those of ligaments (kPa scale),^[11-13] which contain a large amount of elastin in addition to collagen, and tendon (GPa scale),^[15-17] which is mainly made of collagen. The strain to failure of PGS is similar to that of arteries and veins (up to 260%),^[18] and much larger than that of tendons (up to 18%).^[19] On the high modulus limit, PGS can be stretched to about 20%, still much higher than stiff biodegradable polymers such as poly(lactide-co-glycolide) (PLG). Once cross-linked, PGS is insoluble in water or organic solvents. After soaking 24 h in water, the weight of PGS barely changed and the mechanical properties were virtually the same as for dry polymer.

DEGRADATION OF PGS

The degradation characteristics of PGS were examined both in vitro and in vivo. Agitation for 60 days in phosphate buffered saline solution (PBS) at 37°C caused the polymer to degrade $17 \pm 6\%$ as measured by change of dry sample weight. In contrast, subcutaneous (SC) implants in rats were totally absorbed in 60 days. For the in vivo experiment, enzymes, and perhaps macrophages as well, may have caused differences in degradation rate. In vivo degradation thinned the polymer implants, with the explants maintaining their square shape and relatively sharp edges up to at least 35 days. Both the mechanical strength and the mass decreased almost linearly as the polymer degraded. Both results suggest PGS predominantly degrades through surface erosion. The preservation of integrity during the degradation process can be important for certain types of tissue-engineered implants, drug delivery devices, and medical devices.

In Vivo Degradation

Appropriate degradation properties of biodegradable polymers are critical for the success of their clinical applications. The in vivo degradation characteristics of PGS were evaluated using a subcutaneous model in Sprague-Dawley (SD) rats with PLG as the control material. Carboxyl-terminated PLG (50:50, carboxyl-ended, MW 15,000) powder was pressed into round disks. The surface area/volume ratio was the same for PGS and PLG samples. PGS samples were autoclaved, and PLG samples were sterilized by ethylene oxide (4 h sterilization, 48 h ventilation) before implantation. All samples were implanted subcutaneously in the backs of SD rats. The animals were randomly divided into five groups. At days 7, 14, 21, 28, and 35, implants were explanted from one group of animals under general anesthesia. The implants were carefully retrieved and rinsed sequentially with PBS solution and deionized (DI) water. The geometry and surface

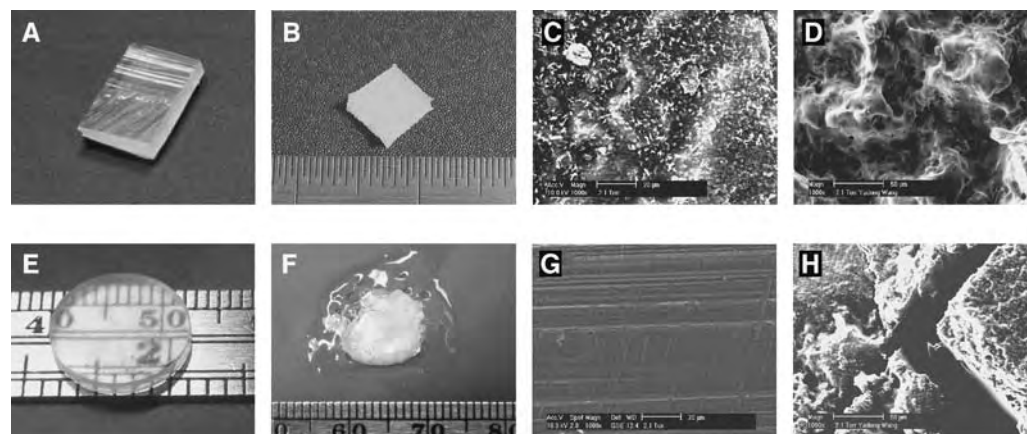


Fig. 3 Photographs and scanning electron micrographs ($1000\times$) of PGS (top) and PLG (bottom) explants at various time points of degradation. PGS implants preserve their geometry and surface integrity throughout the degradation period, whereas PLG implants deform severely. PGS, A–D; PLG, E–H. A: new PGS slab; B: 21 days implantation; C: surface of autoclavated PGS slab—new; D: surface of PGS slab implanted for 21 days. E: new PLG disk implantation; G: surface of ethylene oxide sterilized PLG disk—new; H: surface of PLG disk implanted for 21 days.

properties of the explants were examined, and the change of mass, water content, and mechanical strength with time were also monitored. All these parameters are important indications for the mechanism of degradation, and are crucial for the integrity and proper function of an implant. Macroscopically, the PGS explants maintained their geometry throughout the time period tested (Fig. 3). In contrast, the geometry of PLG explants was distorted significantly within 14 days, most likely because of bulk degradation and swelling. They were deformed considerably from transparent disks to white opaque irregular lumps (Fig. 3). No PLG implants were retrieved successfully beyond three weeks due to excessive swelling and the fragile nature of the implants.

The surface integrity of the explants was observed by scanning electron microscopy (SEM). Both pristine sterilized samples and cleaned and dried explants were mounted on aluminum stubs, and their surface morphologies were observed. SEM observation showed that the PGS explant surface maintained its integrity well. Contoured features developed on the surface after autoclaving. Such features remained throughout the course of the experiment; however, no crack formation was observed (Fig. 3). In case of PLG implants (Fig. 3), $20\ \mu\text{m}$ holes developed on the surface 7 days postimplantation; cracks about $20\ \mu\text{m}$ wide formed within 14 days, and by 21 days, networks of both larger cracks wider than $40\ \mu\text{m}$ and smaller cracks could be seen throughout the PLG surface.

The degree of swelling of degradable polymers in vivo is a key parameter for proper implant materials. Excessive swelling is usually undesirable for an implant, as it distorts the shape of the implant and softens the polymer.^[20,21] The degree of swelling was estimated by measuring the explant weight differences

before and after drying. The explants were cleaned and the surface water removed with a Kimwipe before weighing. Each explant was thoroughly dried at 40°C under vacuum ($85\ \text{mTorr}$) for 48 h. Each explant was weighed again before any subsequent testing. The swelling ratio was calculated from explant weight difference before and after drying: $(W_w - W_d)/W_d$, where W_w is the wet sample weight and W_d is the dry sample weight. The water content of PGS implants rose linearly and reached 15% in 35 days, at which time the polymer had degraded $> 70\%$ (Fig. 4A). In contrast, water uptake of PLG implants showed a time lag followed by a surge of water content, in a pattern similar to its mass loss (Fig. 4B). The water content of PLG implants increased gradually to 11% within 14 days, then increased abruptly, and reached 49% within the next 7 days. One of the key functions of a degradable polymer in an implant is to provide mechanical support. Hence it's critical to know how the mechanical strength changes with degradation. The PGS implants lost mechanical strength gradually and slowly after implantation, about 8% each week. At day 35, when $< 30\%$ of the PGS implant's mass was left, the modulus was $> 50\%$ (Fig. 4A). In contrast, PLG implants lost their mechanical strength shortly after implantation ($> 98\%$ within 7 days). At 14 days, PLG implants' moduli were reduced to 0.25%. At 21 days, with 42% of the mass left, their moduli were reduced to 0.023% (Fig. 4B). This demonstrates that PGS implants maintained their mechanical strength much better than PLG implants. The explants were tested according to ASTM standard D575-91. Briefly, the explants were compressed at a fixed ramp speed of $2\ \text{mm/min}$. PGS explants were compressed to 50% strain, whereas PLG explants were compressed to failure. Pristine samples and PGS explants had regular

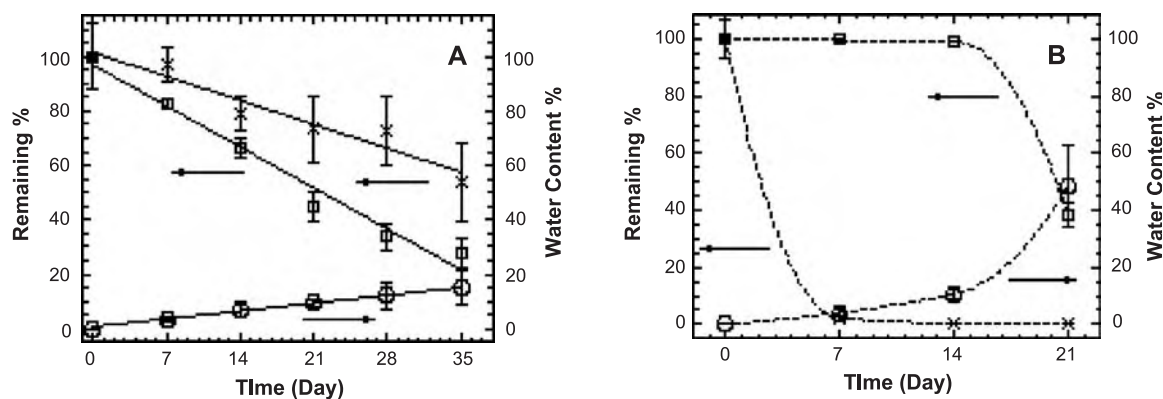


Fig. 4 Comparison of the changes in mass (\square), mechanical strength (\times), and water content (\circ) of PGS (solid line) and PLG (dashed line) implants upon degradation. A) Steady, almost linear changes of PGS implants' properties upon degradation. B) Abrupt, nonlinear changes of PLG implants' properties upon degradation. Data were plotted as mean with standard deviation. (Adapted from Ref. [22].)

geometry and were measured with a digital caliper. PLG explants deformed upon degradation, and their dimensions were measured to the best approximation. The rate of mass loss is of fundamental importance in studying the degradation characteristics of a biodegradable polymer. PGS implants lost their weight steadily and linearly over the test period of 35 days, at which time they had lost $> 70\%$ of their mass (Fig. 4A).^[22] The mass loss of PLG was negligible ($< 1\%$) within 14 days, then surged abruptly, and reached 61% in the next 7 days (Fig. 4B). The dramatic mass loss after the initial lag for PLG upon degradation was similar to what has been reported in the literature.^[23-25] The differences in degradation characteristics between PGS and PLG under identical conditions indicate they probably degrade by different mechanisms. Unlike PLG, which degrades mostly by bulk degradation, in vivo degradation of PGS is dominated by surface erosion, as indicated by linear mass loss with time, preservation of implant geometry, better retention of mechanical strength, absence of surface cracks, and minimal water uptake. Upon degradation, PGS implants keep their integrity better than PLG implants, and may prove useful in biomedical applications where other such polymers are unsuccessful.

BIOCOMPATIBILITY OF PGS

PGS appears to be biocompatible both in vitro and in vivo. Under identical conditions, cultured fibroblasts grew faster on PGS surfaces than on PLG surfaces. Subcutaneous implantation experiments showed that PGS induced only acute inflammation, whereas PLG implantation sites presented signs of chronic inflammation such as fibrous capsule formation and foreign body giant cells.

In Vitro Biocompatibility

The in vitro biocompatibility of PGS was evaluated in comparison with PLG, both in morphology and in cell growth rate. Glass Petri dishes were coated with PGS and PLG, respectively. The coated dishes were sterilized by UV radiation for 15 min. Each dish was soaked in growth medium for 4 h; the medium was replaced with fresh medium and the dishes soaked for another 4 h to remove any unreacted monomers or residual solvents before cell seeding. NIH 3T3 fibroblast cells were seeded homogeneously on these coated dishes. The cells in PGS sample wells were viable and showed normal morphology with higher growth rate than the control, as tested by methyl thiazol tetrazolium (MTT) assay.^[26] Cells in PLG wells tended to form clusters, and the number of floating cells was higher; furthermore, most of the attached cells adopted a long thin threadlike morphology. These experiments suggested that PGS is at least as biocompatible as PLG in vitro.

In Vivo Biocompatibility

Subcutaneous implantation in SD rats was used to compare the in vivo biocompatibility of PGS and PLG. The surface area/volume ratio was kept the same for both PGS and PLG implants. Autoclaved PGS slabs of approximately $6 \times 6 \times 3$ mm, and ethylene oxide-sterilized PLG (50:50, carboxyl-ended, MW 15,000, Boehringer Ingelheim, Inc., Germany) disks were implanted SC in female SD rats by blunt dissection under general anesthesia. Two implants each of PGS and PLG were implanted symmetrically on the upper and lower back of the same animal. Every implantation site was marked by tattoo marks. The animals were randomly divided into five groups. At each predetermined time point (7, 14, 21, 28, 35 days),

one group of rats was sacrificed, and tissue samples surrounding the implants were harvested. The samples were fixed and embedded in paraffin after dehydration. The slides were stained with hematoxylin and eosin (H&E) and Masson's trichrome stain (MTS). The thickness of the inflammatory zone (H&E) and collagen deposition (MTS) for each polymer implant is expressed as the average value of three readings per slide for six slides at each time point.

The inflammatory responses subsided with time for both PGS and PLG implants. In the first three weeks, the inflammatory response of PLG implantation sites was about 16% thinner than that of PGS (Fig. 5). The thickness of the inflammatory zone in both implantation sites was approximately the same at weeks 4 and 5. Fibrous capsules surrounding PLG implants

developed within 14 days, and their thickness hovered around 140 μm . Collagen deposition did not appear around PGS implants until 35 days. The collagen layer was highly vascularized and was only about 45 μm thick. Foreign body giant cells were present in the PLG implantation sites 35 days postimplantation. However, no foreign body giant cells were observed around PGS implants. The inflammatory response and fibrous capsule formation observed for PLG is similar to those reported in the literature.^[27,28] Thick fibrous capsules may block mass transfer between the implants and surrounding tissues, which can impair implant functions. In an *in vivo* study with PGS alone, the SC implantation sites were undetectable despite repeated sectioning of the specimens at multiple levels in 60 days (two implantation sites each in three animals).

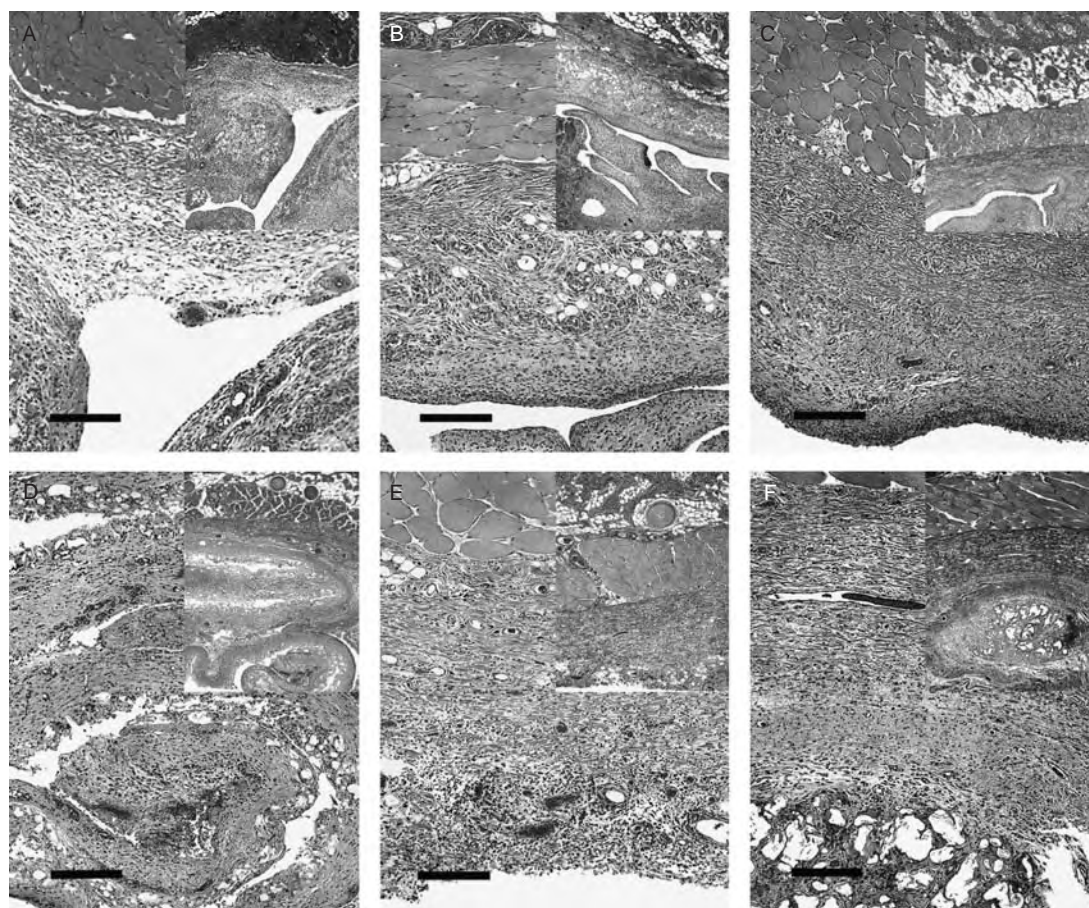


Fig. 5 Photomicrographs of rat skin; comparisons of lumen wall characteristics (H&E, 10 \times) and fibrous capsule thickness (insets, MTS, 5 \times) at implantation sites across time. A, C, E: PGS; B, D, F: PLGA. A, B: 7 days postimplantation (pi)—lumenal wall was markedly thickened by a zone of dense vascular proliferation and mild inflammation without detectable collagen deposition. C, D: 21 days pi—lumenal wall was significantly thinner with a modest degenerative inflammatory infiltrate immediately adjacent to the polymer. PLGA implantation site (D) was marked by a significant collagen fibrous capsule, which was absent in PGS. E, F: 35 days pi—lumenal wall was reduced to a thin zone of cell debris with no vascular proliferation. Collagen deposition in PGS implantation site was much thinner than that surrounding the fragmented PLGA implant. Top: skin, implantation site: *, scale bar = 200 μ . (Adapted from Ref. [26].)

The implants were completely absorbed without granulation or scar tissues, and the implantation site was restored to its normal histological architecture. Overall, the acute inflammatory response of PGS is similar to that of PLG. However, unlike PLG, PGS barely induces any significant fibrous capsule formation, and the implantation site is free of foreign body giant cells, suggesting the absence of chronic inflammation.

APPLICATIONS OF PGS

Similar to vulcanized rubber, this PGS elastomer is a thermoset polymer. However, the uncross-linked prepolymer can be processed into various shapes, because it can be melted into a liquid and is soluble in common organic solvents. In a typical procedure, a mixture of NaCl particles of appropriate size and an anhydrous THF solution of the prepolymer is poured into a Teflon mold. The polymer is cured in the mold in a vacuum oven at 120°C and 100 mTorr. A porous scaffold that conforms to the shape of the mold is obtained after salt leaching in DI water. So far, the applications of PGS are focused on soft-tissue engineering. The material supports the growth of a wide range of cell types such as endothelial and smooth muscle cells, Schwann cells, cardiomyocytes, hepatocytes, chondrocytes, and fibroblasts.

Artificial Microvasculature

Vital organs maintain dense microvasculatures to sustain the proper function of their cells. For tissue-engineered organs to function properly, artificial capillary networks have to be developed. Due to its elastomeric properties, PGS might be a good candidate for fabricating microvasculature. Standard MicroElectroMechanical Systems (MEMS) techniques were used to etch capillary network patterns onto silicon wafers. The resultant silicon wafers served as micromolds for the devices. The patterned PGS film was bonded with a flat one to create capillary networks. The devices were perfused with a syringe pump at a physiological flow rate, and became endothelialized to near confluence within 7 days. These endothelialized capillary networks could be used in tissue engineering of vital organs.

Cardiac Patch

To evaluate the performance of PGS as a scaffolding material for cardiac tissue engineering, neonatal rat cardiomyocytes were seeded in PGS, collagen, and PLG scaffolds. The PGS constructs beat spontaneously for 3 days after seeding, and they beat

synchronously with a DC pulse of 1 Hz until at least 14 days. The collagen constructs behaved similarly up to 6 days. After 6 days, the constructs collapsed, because the collagen matrix was degraded or absorbed so much that it had little mechanical strength to hold the construct in proper shape. PLG scaffolds were also used to build cardiac patches. The construct failed to show any beating at macroscopic level, most likely because of the stiffness of the PLG scaffolds. Compared with collagen foams and PLG, PGS appears to be as good, if not better suited, for tissue engineering of cardiac muscles.

Liver Tissue Engineering

One challenge in liver tissue engineering is to keep the liver-specific functions of the hepatocytes in the constructs. PGS and PLG scaffolds of the same porosity and pore sizes were treated under identical conditions, and seeded with hepatocytes. Hepatocytes on PGS scaffolds showed approximately four times higher ammonia metabolism rate and 40% higher albumin secretion rate than those on PLG.

CONCLUSION

Compared with conventional biodegradable polymers, PGS appears to be tougher, less expensive, and more flexible. In the models tested, the material is biocompatible both in vitro and in vivo. The degradation of PGS is mostly through surface erosion, which gives a more linear decrease in mass and gradual change in mechanical properties. In addition, the polymer doesn't swell in aqueous media conditions when fresh, and has limited swelling upon degradation. Thus, the geometry and surface integrity of PGS implants are maintained very well. These properties of PGS may be desirable in many applications in biomedical engineering beyond those discussed previously. The polymer's properties—such as hydrophilicity, degradation rate, and pattern—can potentially be tailored by grafting hydrophobic moieties onto the hydroxyl groups. To further control or regulate polymer interaction with cells, biomolecules could be coupled to the hydroxyl groups or integrated into the polymer backbone.

ACKNOWLEDGMENTS

The author would like to thank Drs. Robert Langer, Hiroyuki Ijima, Gordana Vunjak-Novakovic, Milica Radisic, Hyounghsin Park, and Mrs. Christina Fidkowschi for their contributions to research in PGS.

ARTICLES OF FURTHER INTEREST

Biocompatibility Testing; Biodegradable Polymers, An Overview; Biofunctional Polymers; Bioreactors; Diamond and Diamond-Like Carbons; Extracellular Matrix Scaffolds; Tissue Engineering of Liver; Tissue Engineering of Microvascular Networks

REFERENCES

1. Peppas, N.A.; Langer, R. New challenges in biomaterials. *Science* **1994**, *263* (5154), 1715–1720.
2. Langer, R. Biomaterials: Status, Challenges, and Perspectives. *AIChE J.* **2000**, *46* (7), 1286–1289.
3. Hench, L.L.; Polak, J.M. Third-generation biomedical materials. *Science* **2002**, *295*, 1014–1007.
4. Langer, R. Drug delivery and targeting. *Nature* **1998**, *392* (6679 Suppl.), 5–10.
5. Liu, G.; Hinch, B.; Beavis, A.D. Mechanisms for the transport of α,ω -dicarboxylates through the mitochondrial inner membrane. *J. Biol. Chem.* **1996**, *271* (41), 25338–25344.
6. Grego, A.V.; Mingrone, G. Dicarboxylic acids, an alternate fuel substrate in parenteral nutrition: An update. *Clin. Nutr.* **1995**, *14* (3), 143–148.
7. Tamada, J.; Langer, R. The development of poly-anhydrides for drug delivery applications. *J. Biomater. Sci., Polym. Ed.* **1992**, *3* (4), 315–353.
8. Nagata, M.; Machida, T.; Sakai, W.; Tsutsumi, N. Synthesis, characterization, and enzymatic degradation of network aliphatic copolyesters. *J. Polym. Sci., A, Polym. Chem.* **1999**, *37*, 2005–2011.
9. Dupont-Gillain, C.C.; Nysten, B.; Rouxhet, P.G. Collagen adsorption on poly(methyl methacrylate): Net-like structure formation upon drying. *Polym. Int.* **1999**, *48* (4), 271–276.
10. Sperling, L.H. *Introduction to Physical Polymer Science*; John Wiley & Sons: New York, 1992.
11. Yamaguchi, S. Analysis of stress-strain curves at fast and slow velocities of loading in vitro in the transverse section of the rat incisor periodontal ligament following the administration of beta-aminopropionitrile. *Arch. Oral Biol.* **1992**, *37* (6), 439–444.
12. Komatsu, K.; Chiba, M. The effect of velocity of loading on the biomechanical responses of the periodontal ligament in transverse sections of the rat molar in vitro. *Arch. Oral Biol.* **1993**, *38* (5), 369–375.
13. Chiba, M.; Komatsu, K. Mechanical responses of the periodontal ligament in the transverse section of the rat mandibular incisor at various velocities of loading in vitro. *J. Biomech.* **1993**, *26* (4–5), 561–570.
14. Nagdi, K. *Rubber as an Engineering Material: Guideline for Users*; Hanser: Munich, 1993.
15. Fratzl, P.; Misof, K.; Zizak, I.; Rapp, G.; Amenitsch, H.; Bernstorff, S. Fibrillar structure and mechanical properties of collagen. *J. Struct. Biol.* **1998**, *122* (1–2), 119–122.
16. Misof, K.; Rapp, G.; Fratzl, P. A new molecular model for collagen elasticity based on synchrotron x-ray scattering evidence. *Biophys. J.* **1997**, *72* (3), 1376–1381.
17. Wang, J.L.; Parnianpour, M.; Shirazi-Adl, A.; Engin, A.E. Failure criterion of collagen fiber: Viscoelastic behavior simulated by using load control data. *Theor. Appl. Fract. Mech.* **1997**, *27* (1), 1–12.
18. Lee, M.C.; Haut, R.C. Strain rate effects on tensile failure properties of the common carotid artery and jugular veins of ferrets. *J. Biomech.* **1992**, *25* (8), 925–927.
19. Haut, R.C. The effect of a lathyritic diet on the sensitivity of tendon to strain rate. *J. Biomech. Eng.* **1985**, *107* (2), 166–174.
20. Yoon, J.J.; Park, T.G. Degradation behaviors of biodegradable macroporous scaffolds prepared by gas foaming of effervescent salts. *J. Biomed. Mater. Res.* **2001**, *55* (3), 401–408.
21. Kranz, H.; Ubrich, N.; Maincent, P.; Bodmeier, R. Physicomechanical properties of biodegradable poly(D,L-lactide) and poly(D,L-lactide-co-glycolide) films in the dry and wet states. *J. Pharm. Sci.* **2000**, *89* (12), 1558–1566.
22. Wang, Y.; Kim Yu, M.; Langer, R. In vivo degradation characteristics of poly(glycerol sebacate). *J. Biomed. Mater. Res.* **2003**, *66A* (1), 192–197.
23. Lu, L.; Peter, S.J.; Lyman, M.D.; Lai, H.L.; Leite, S.M.; Tamada, J.A.; Uyama, S.; Vacanti, J.P.; Langer, R.; Mikos, A.G. In vitro and in vivo degradation of porous poly(dl-lactic-co-glycolic acid) foams. *Biomaterials* **2000**, *21* (18), 1837–1845.
24. Vert, M.; Li, S.; Garreau, H. More about the degradation of LA/GA-derived matrixes in aqueous media. *J. Control. Release* **1991**, *16* (1–2), 15–26.
25. Kenley, R.A.; Lee, M.O.; Mahoney, T.R., II; Sanders, L.M. Poly(lactide-co-glycolide) decomposition kinetics in vivo and in vitro. *Macromolecules* **1987**, *20* (10), 2398–2403.
26. Wang, Y.; Ameer, G.; Sheppard, B.; Langer, R. A tough biodegradable elastomer. *Nat. Biotechnol.* **2002**, *20*, 602–606.
27. Cadee, J.A.; Brouwer, L.A.; den Otter, W.; Hennink, W.E.; Van Luyn, M.J.A. A comparative biocompatibility study of microspheres based on crosslinked dextran or poly(lactic-co-glycolic)acid after subcutaneous injection in rats. *J. Biomed. Mater. Res.* **2001**, *56* (4), 600–609.
28. van der Elst, M.; Klein, C.P.A.T.; de Bleeck-Hogervorst, J.M.; Patka, P.; Haarman, H.J.T.M. Bone tissue response to biodegradable polymers used for intra medullary fracture fixation: A long-term in vivo study in sheep femora. *Biomaterials* **1999**, *20* (2), 121–128.

Biosensors

Jerome S. Schultz

Center for Bioengineering Research, University of California, Riverside, California, U.S.A.

B

INTRODUCTION

Biosensors are small devices that utilize biological components such as enzymes, membranes, or antibodies for detecting specific analytes. These devices intimately couple a biological recognition element that interacts with an analyte with a physical transducer that converts the biorecognition event into a useful electric signal. Common transducing elements include optical, electrochemical, mass-sensitive, or thermal devices that generate light, current, frequency, or temperature signals, respectively.

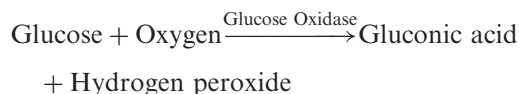
OVERVIEW

The earliest chemical sensor, which was in the form of a probe that could be placed in a solution, was the glass pH electrode for the measurement of hydrogen ion concentration. Because of the relative inertness of glass to most chemicals and ions, the pH electrode—which was developed in the 1920s—was one of the first devices that could be used directly in biological media such as blood. One can distinguish two categories of biosensors by the nature of the recognition event. Bioaffinity devices rely on selective and sometimes reversible binding of the target analyte to a surface-bound ligand. No net consumption of analyte or other reagent takes place. In contrast, biosensors that use biocatalytic components such as an immobilized enzyme to recognize target substrate incur loss of some component(s) of the assay mixture. These two types of biosensor systems have different operational characteristics. In the latter case, because of the consumption of reagents, the diffusion of the analyte to the sensor can become a limiting factor in sensor response or sensitivity. For example, if the biosensor is placed in a sample that is not well stirred, then variations in the rate at which the analyte can reach the sensor can affect the magnitude of the signal generated and slow down the response of the system. For the former (bioaffinity) type, the rate at which the sensor reaches a steady reading may be affected by sample mixing, but the calibration will not be affected.

With the increasing interest in miniaturized sensing systems, the distinction between a biosensor and a portable bioassay is becoming blurred. Portable

bioassays were introduced in 1965 with the paper strip assay system for glucose called Dextrostix, which was manufactured by Miles Laboratories (now Bayer). This “dry chemistry” technology consisted of a glucose oxidase enzyme impregnated into a paper strip with associated chemicals that produced a color change when a drop containing glucose was applied to the strip. The method was used for monitoring blood glucose.

In that case, the chemical reactions were



Hydrogen peroxide + Orthotodine

→ Oxidized orthotodine (colored)

The amount of color produced was quantitated either by comparing to a standard color chart or by measurement in a reflectometer device called the Ames Reflectance Meter.

These types of assays have expanded greatly with the introduction of immunoassay cartridge units that have been devised for a variety of purposes from pregnancy tests to tests for drugs of abuse. When the readout of these bioassays has been coupled to an instrument such as a colorimeter or an electrode system, the device is sometimes called a biosensor. While the original use of the term “biosensor” implied a handheld probe of some type that could be placed in a sample to obtain a continuous measurement of a specific analyte, today the term biosensor is used so broadly that the only remaining feature of the original meaning relates to a miniature bioanalytic system that incorporates a numerical readout device.

The history of biosensors goes back to 1956 when Leland Clark described an oxygen electrode^[1] and 1962 when he coupled an enzyme to the oxygen electrode to make a glucose electrode.^[2] The problem that prompted this invention was the need to continuously measure metabolites in biological tissues. As is still true, most methods required sampling of tissue, blood, or other fluids and processing the material separately through various laboratory procedures in order to obtain accurate estimates of various biochemicals. The lack of suitable sensor systems limited

the understanding of the dynamics of biological systems, and in certain applications such as diabetes and chemotherapy, the clinical benefit of invoking control algorithms by feedback techniques was not possible.

The critical new concept introduced by Clark in 1956 was the use of permeable membranes to form separate but intercommunicating regions within the device. In a polarographic cell, he interposed a membrane between the platinum working electrode and the external media to protect the electrode from deleterious components in the sample solution. Clark's oxygen electrode, shown as the upper component in Fig. 1, is an amperometric polarographic device. A constant voltage is applied to a platinum electrode versus the reference silver/silver chloride that places the electrode in the diffusion-limited regime, i.e., the oxygen concentration at the electrode surface is reduced to zero. Thus, the amount of current developed at the electrode is directly proportional to the rate of oxygen diffusion to the electrode. Clark's invention was to place a plastic membrane that provided a constant diffusion resistance to oxygen between the electrode and the test solution, ensuring that the current was proportional to the oxygen. The other function of the

plastic membrane was to prevent other electroactive substances in the test solution from reaching the platinum electrode. (Gases can diffuse through plastic but ions cannot.) Because of the success of Clark's glucose electrode (Fig. 1), much of the early work on biosensors also utilized electrochemical detection devices.

The emergence of alternative measurement methods that utilized spectrophotometric techniques was initiated in 1976 by Lubbers and Optiz,^[3] who coined the term "optodes" as a methodology coupling optical systems to miniaturized analytical methods for sensors. (The word "optode" was meant to convey a similarity of this technique to electrodes, where the interaction occurs at the surface of the device.) This opened doors to transforming many of the analytical colorimetric, spectrophotometric, and fluorometric methods used for many years in analytical chemistry to sensor formats. Lubbers showed that these approaches could be miniaturized by placing appropriate chemistries at the end of an optic fiber, thereby producing probes that could be used in a variety of environments and solutions. One of the very first probes of this sort was produced by Peterson and colleagues at NIH,^[4] who developed a fiber-optic pH probe. This device was

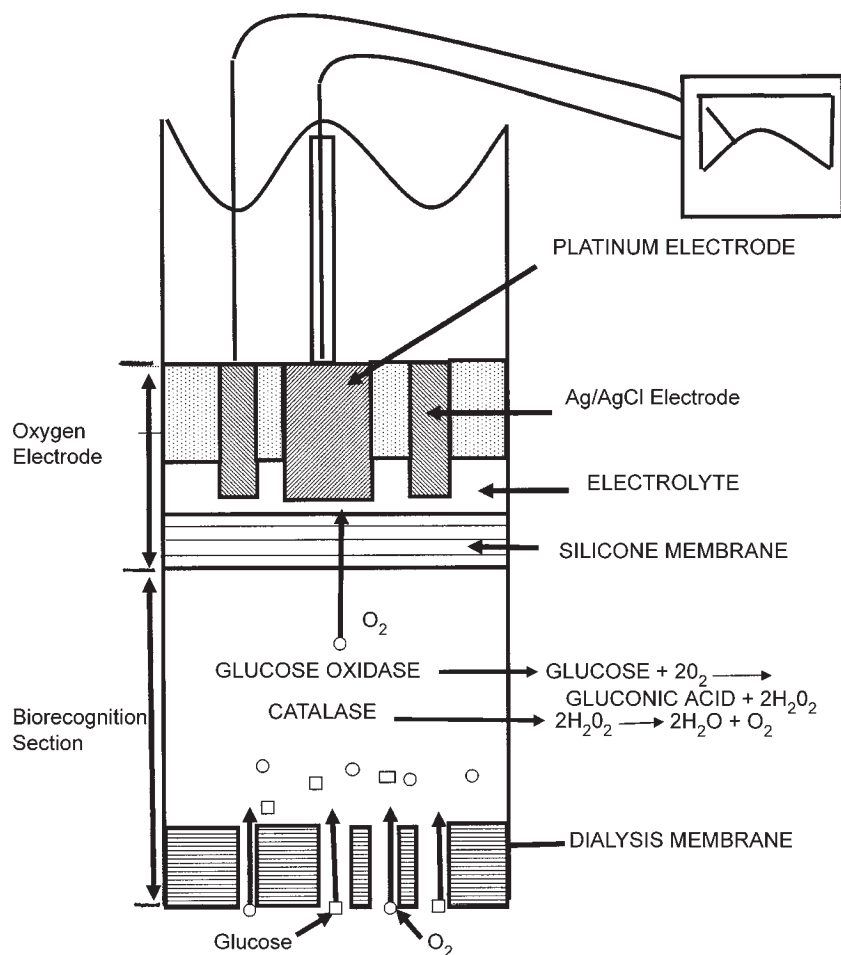


Fig. 1 Structure of Clark sensors. The upper section of this composite sensor is the "oxygen electrode." This portion consists of an amperometric polarographic cell that is separated from the environment by a silicone rubber membrane. The lower "biorecognition section" contains the enzyme glucose oxidase that converts glucose to gluconic acid. Together, the two sections compose Clark's "glucose biosensor."

based on the changes in the absorption spectra of pH-sensitive dyes placed at the end of a fiber-optic catheter. Peterson also developed a fluorescent probe for oxygen by using a dye that exhibits quenching in the presence of oxygen (first described and utilized in this fashion by Lubbers).

This concept was elaborated further by David Walt and colleagues,^[5] who made enzyme optodes by placing enzymes at the end of optical fibers and using chemical reagents that produced an optical change (either absorption or fluorescence) when presented with the analyte.

A key advantage of most biosensors is the ability these probes can provide for reliably measuring specific analytes when inserted into diverse samples consisting of complex mixtures of chemical and biochemical entities. This capability makes biosensors particularly attractive for research and consumer use. Since Clark's invention, hundreds, if not thousands, of biosensors have been described in the literature. A few sensor devices have reached commercialization with applications still primarily in the biomedical area, such as for blood glucose. However, as these devices are proven to be reliable, robust, and portable, their applications to other fields, such as the food industry (for freshness) and environmental/toxic monitoring, is expanding rapidly.

There are two major components of biosensors, as shown in Fig. 1. These consist of the biochemical component that provides the recognition of the analyte of interest, and a detector that responds to a physical/chemical change in the local environment that is the result of the interaction of the analyte with the biochemical component. Often membranes are used to separate regions of the sensor for various reasons—e.g., to prevent the leakage reagents from touching the

sensor, to protect the detector from deleterious interactions with materials in the sample, to allow only certain types of chemical species to enter the sensor.

Although there are many sensors that could be used as a detector in a biosensor, the choice is somewhat limited to the type of chemical interaction that occurs between the analyte and recognition element. Fig. 2 illustrates some of the schemes that have been tried for a glucose biosensor. The sensitivity of the final biosensor system is strongly related to the characteristics of the physical sensor used as the detector element.

The biological recognition element is the component that gives biosensors their unique qualities of selectivity. These molecules are a gift of millions of years of biochemical evolution. During the eons that living systems evolved, several unique structures were evolving at the molecular level as well. For example, antibodies (Fig. 3) are an essential component of the immune system that only came into existence when mammals made their appearance on earth. The immune system is the most powerful molecular identification system yet discovered. A typical animal has the ability to generate about 100,000 different antibodies that have binding selectivities for an equal number of analytes. In addition, a typical cell has hundreds of different enzymes, each of which could serve as the recognition element in a biosensor. Also, the enormous diversity of receptors for signaling species such as G proteins just becoming documented are another potential resource for biosensors.

Some appreciation of the discriminatory capability of biosensors can be obtained from the data in Table 1 (from YSI) that gives the relative response of their glucose biosensor to various compounds that might be found in blood or other biological samples. As can be seen, the potential interference of a glucose measurement

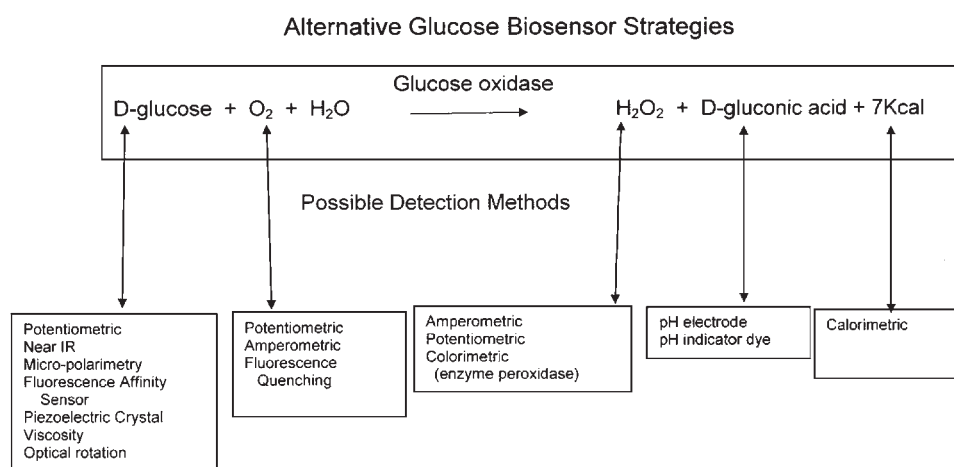


Fig. 2 Examples of the wide variety of strategies for the measurement of glucose primarily based on the biorecognition afforded by glucose oxidase. Biosensors corresponding to each of these strategies have been described in the literature. (Adapted from Ref. [6].)

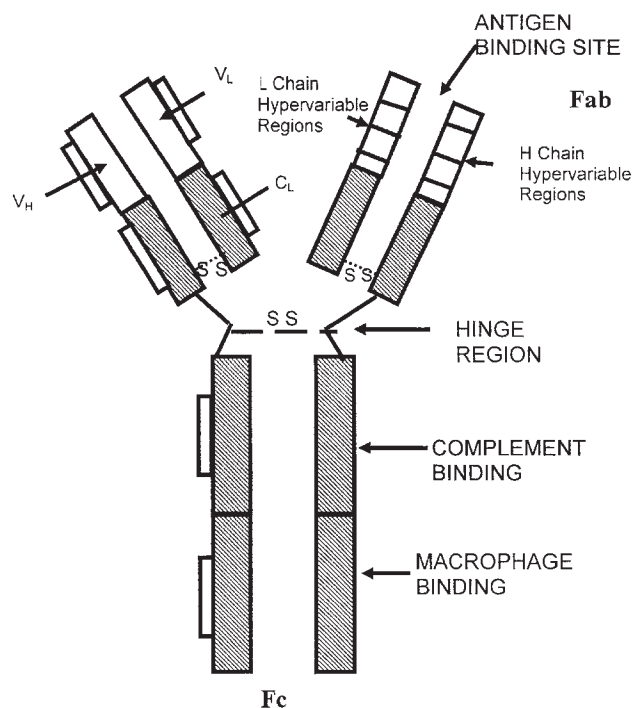


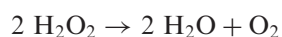
Fig. 3 General structural features of IgG antibodies, molecular weight about 180,000. The binding sites are located at the terminal ends of the arms of the molecule. The molecule can be cleaved near the hinged region to result in active mono-binding fragments known as Fabs.

by other compounds in these samples can be quite high (e.g., acetaminophen). This illustrates the need to evaluate the response of biosensors in the type of samples of interest, since there might be unknown components that may affect the calibration of the biosensor.

GENERAL FEATURES OF A BIOSENSOR

The components of a commercial model of a Clark-type glucose sensor produced by YSI are shown in a more expanded view in Fig. 4. The primary component is the enzyme glucose oxidase that provides the selectivity to the sensor. As mentioned earlier, glucose oxidase had been used in the Miles Dextrostix glucose assay strip.

The reaction steps for the glucose electrode are shown below:



The net reaction is



For every mole of glucose that reaches the enzyme and is consumed, $1/2$ mole of oxygen is also consumed. Since the detector, an oxygen electrode, measures

oxygen concentration, the net result of the presence of glucose is a reduced oxygen reading from the oxygen electrode. Thus for the sensor to work satisfactorily, the oxygen concentration in the solution presented to the glucose electrode should be the same from sample to sample. This condition is easy to achieve if the solutions presented are saturated with air, but difficult to achieve if the sensor is placed in tissue where the oxygen concentration can vary widely.

This illustrates that there are special design considerations and compromises that are required in most sensor designs. For example, one solution to deal with the variation in oxygen concentration in vivo was developed by David Gough, who fashioned a dual-electrode system consisting of a glucose electrode and a separate oxygen electrode in the same catheter. Thus variations in blood oxygen could be accounted for. In one series of animal tests, Gough's glucose sensor gave reliable measurements for more than six months when implanted in a blood vessel in a dog.^[7]

Although there is ongoing very active research on new biosensor technologies, translating these new concepts into practical products for commercial use has been difficult. Issues to be resolved are stability (such as storage life), reliability, and accuracy. Since many of these sensors contain complex biomolecules such as antibodies and enzymes, long-term stability is an important issue. One solution is to manufacture the biosensor in a manner so that various elements can be replaced easily. YSI markets a biosensor platform that can be disassembled to replace sensitive components while preserving the configuration of the various elements as shown in Fig. 4.

Membrane matrices are available with a variety of different immobilized oxidase enzymes to allow the measurement of a variety of analytes such as choline, D-glucose, dextrose, ethanol, galactose, hydrogen peroxide, lactose, L-glutamate, L-glutamine, L-lactic acid, methanol,

Table 1 Inhibitors of glucose oxidase-based biosensor

Analyte	Response relative to glucose ^a (percent)
Fructose	0.1
Sorbitol	0.3
Xylose	1.0
Galactose	1.5
Ascorbic acid	2.0
Mannose	3.0
Acetaminophen	300

^aResponse relative to glucose is calculated as follows: If we assume that the instrument reading for glucose at some concentration (A) is 50 and the instrument reading for another analyte at the same concentration (A) is x, then the response relative to glucose is $(x/50) \times 100$.

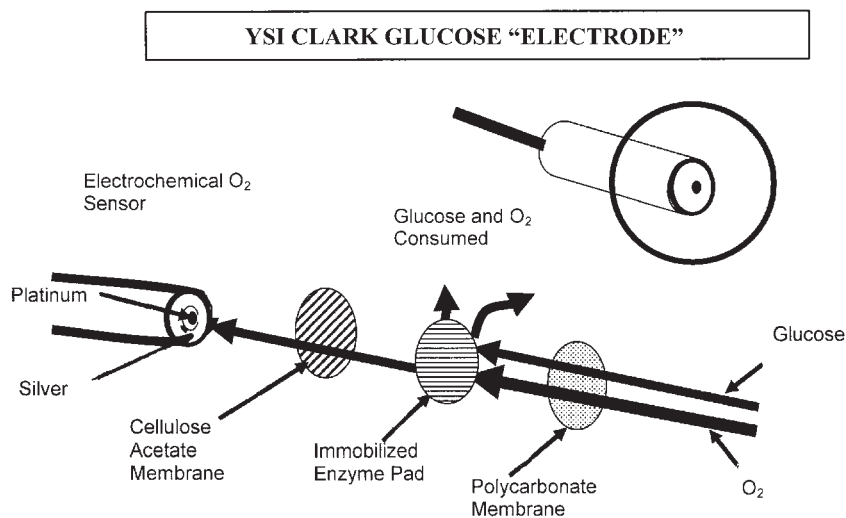


Fig. 4 Exploded view of a commercial biosensor (manufactured by YSI). The sensor is designed to provide a platform for several analytes. The section titled “Enzyme Pad” can be substituted with membranes containing different enzymes so that the biosensor can be used to measure several analytes.

and sucrose. Typical sample volume for these commercial biosensors is on the order of 10–25 microliters.

Another example of a commercial biosensor is the handheld glucose analyzer that is sold to diabetics for daily monitoring of blood glucose (Fig. 5). In this device, a drop of blood is placed on a disposable strip that contains the reagents, including glucose oxidase. The monitoring device may have a data storage unit that allows the review of blood glucose values over a period of time.

BIOLOGICAL RECOGNITION ELEMENTS

Enzymes have been the most frequently selected biorecognition element in biosensors for a number of

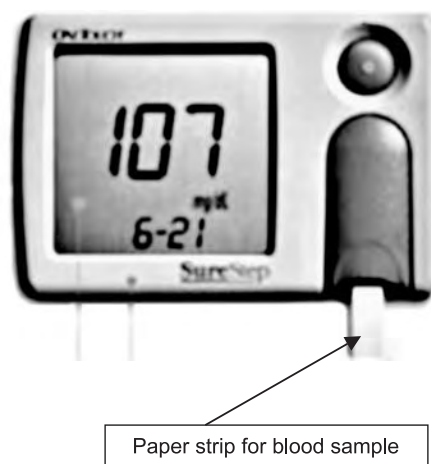


Fig. 5 Photo of a commercial biosensor used by diabetics. First, one draws a drop of blood by pricking a finger. The drop is placed on a treated paper strip that is then placed inside the sensor to obtain a measure of the glucose concentration in the blood. Most of the commercial glucose sensors to date use glucose oxidase as the active reagent.

reasons. The technology for producing enzymes for commercial purposes is very well developed, and a number of enzymes are commonly used in conventional bioassays such as ELISAs (enzyme-linked immunoassays). Technologies to isolate, purify, and immobilize these reagents makes them a very practical resource for biosensor applications. In addition, the reaction kinetics and substrate specificities of many enzymes have been well documented and reported in compendiums of enzyme literature. Some of the enzymes that have been used in biosensors are listed in Table 2. In some cases, the role of the enzyme is not for biorecognition but rather as a biochemical amplifier to enhance the measurement of a binding event with the selective bioreceptor. This strategy was first used in immunoassays—for example, using horseradish peroxidase with the substrate *o*-phenylenediamine to produce a highly colored product—resulting in a well-established technology that is useful for developing biosensors.

Another principle that has been evaluated for the construction of optically based sensors is the use of chemiluminescence. In these cases, an enzyme system specific for the analyte is coupled to reactions that produce light through chemiluminescence. In principle, systems of this type could be very sensitive, first of all due to the amplification factor of enzyme reactions, and secondarily because fluorescence measurements are among the most sensitive of optical techniques.

At nearly every level of biological behavior, there are recognition events that direct biological function. Doubled-stranded DNA holds together due to the complementary structure of the two strands, proteins are synthesized through a series of steps involving transfer RNA and ribosomes, some proteins become enzymes and have specific binding characteristics for certain substrates, and so on, including membrane

Table 2 Enzymes that have been used in commercial biosensors

Enzyme	K_M (mol l ⁻¹)	Substrate
Alcohol dehydrogenase	1×10^{-2}	Ethanol
Catalase	1	Sucrose, glucose
Choline dehydrogenase	7×10^{-3}	Phospholipids
Lactate dehydrogenase	2×10^{-3}	Lactate
Galactose oxidase	2×10^{-1}	Galactose
Glycerol kinase	1×10^{-6}	Glycerol
Glucose oxidase	1×10^{-2}	Glucose
Lactate dehydrogenase	2×10^{-2}	Lactate
Lactate oxidase	2×10^{-2}	Lactate
Mutarotase	3×10^{-2}	Glucose
Sulfite oxidase	3×10^{-5}	Sulfite
Urate oxidase	2×10^{-5}	Uric acid

(From Ref. [8].)

proteins that have specific transport properties. Each of these biological elements can in principle be used as a recognition element in a biosensor. Recently there has been a great deal of research to understand and predict the binding characteristics of various proteins that control essential physiological functions. These studies use computer programs to simulate the binding (or docking) of reagents to the “active site” of proteins. Interest in these programs (available from companies such as Tripos and Accelrys) derives primarily from the pharmaceutical industry to obtain some educated guesses for potential drug candidates. But these programs can also benefit groups working on sensor development. One may wonder why such complex molecules evolved to carrying our signal transduction in biological systems. The likely explanation is that the macrostructure is needed to stabilize the three-dimensional configuration of the binding site.

One class of bioreceptors that has made a considerable impact on bioassays is antibodies. The use of antibodies as an analytical reagent was pioneered by Berson and Yalow,^[9] who devised the type of assay known as immunoassays. Their technique is illustrated in Fig. 6. It depends first on finding an antibody to the analyte of interest and synthesizing an analogue to the analyte, denoted by the symbol (*), with some type of signaling agent as part of its structure (such as a radioisotope or a fluorochrome). (In the immunology literature, the analyte is called the antigen.) In one mode of operation, the antibody is immobilized on the surface of the test vessel. Then, the test solution with a known amount of the analogue analyte is introduced into the test vessel. After a period of incubation, the test solution is removed, and the amount of analogue analyte remaining in the tube is measured. This

amount will be inversely related to the amount of analyte in the test solution. The sensitivity of the immunoassay is related to the affinity of the analyte to the antibody. As shown in Fig. 7, depending on the method of preparation, antibodies can be prepared with affinities that range over six orders of magnitude. However, as also shown in this figure, the higher the affinity the slower the rate of dissociation of the antibody–analyte complex.

Today hundreds of immunoassay kits are produced commercially for use in hospitals and clinical laboratories, and thousands of these tests are performed daily. This assay is widely used because of the high specificity of antibody–analyte binding in the presence of many other biochemicals. For example, the test sample can be from blood plasma and can be used in an immunoassay without further purification. This

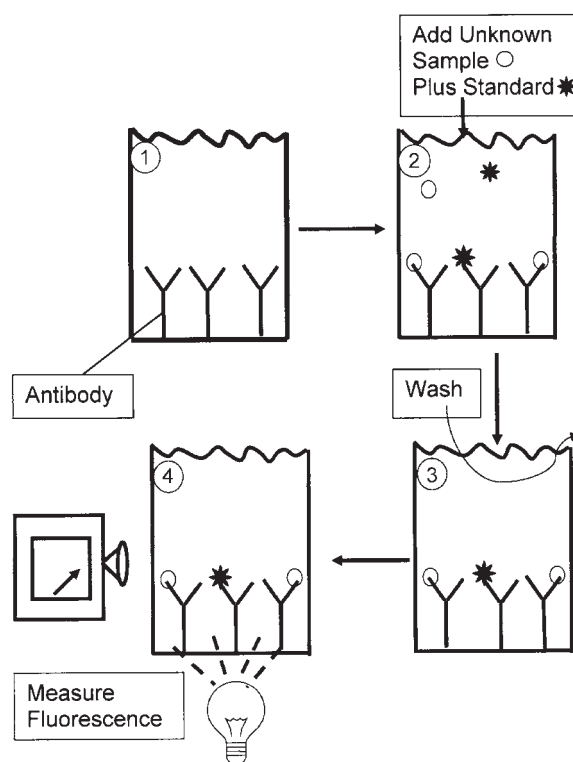


Fig. 6 Illustration of the immunoassay method developed by Berson and Yalow. An antibody (denoted by the symbol Y) for the analyte of interest (denoted by the symbol O) is immobilized on the surface of a well of a multiwell plastic analysis plate (1). A sample containing an unknown amount of the analyte (O) and a known amount of an analogue analyte (*) is pipetted into the well (2). Both the analyte and analogue compete for antibody binding sites. The more analyte in the unknown sample, the less of the analogue will bind. The well is rinsed with buffer to remove unbound analyte and analogue (3). The remaining amount of analogue analyte is measured by a fluorescent photonic device (4).

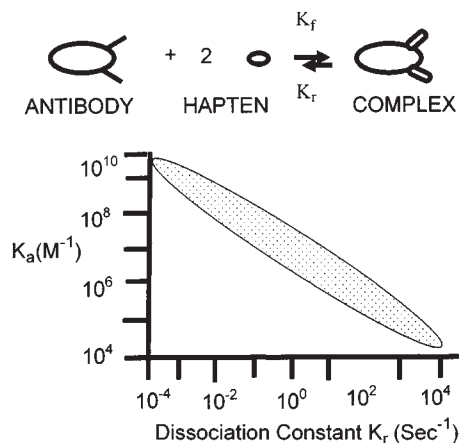


Fig. 7 The range in binding constants that can be prepared from animals. The greater the affinity of an antibody for the analyte, the more sensitive the immunoassay (or biosensor) will be for the analyte. (Data reported in Ph.D. thesis of Ref. [10].)

selective characteristic of antibodies has found many applications in the production of biosensors.

Schultz and Mansouri^[11] were one of the first to apply the immunoassay technique to make a biosensor. First, they placed the reagents in a porous microdialysis “test tube” at the end of an optical fiber. Since an antibody for glucose was not known at the time, they used the protein Concanavalin-A, a lectin well known for its affinity for glucose, as a surrogate for the antibody. They thus demonstrated that any protein that shows specific binding behavior could be used in a biosensor, and termed the technique “affinity biosensors.” A typical calibration curve, Fig. 8, shows the saturation behavior at high analyte concentrations for this type of biosensor.

More recently, Ballerstadt and Schultz^[12] have developed a minimally invasive biosensor that is designed to be placed just under the skin. The implanted capsule is illuminated by an external light source, and fluorescence emitted by the biosensor is measured with an external photodetector. The structure of the sensor is illustrated in Fig. 9a. The hollow dialysis microcapsule is loaded with Sephadex microbeads and fluorescein-labeled Concanavalin-A. The dialysis membrane allows glucose to freely exchange between tissue fluid and the inside of the capsule. The polymer composing the microbeads has many pendant glucose residues, and when the glucose concentration in the microporous capsule is low, most of the Concanavalin-A will reside inside the microbeads because of its affinity for the pendant glucose residues inside the bead. The microbeads are treated with dyes so that light cannot enter the beads and does not elicit emission from the fluorescein/Concanavalin-A within the beads. Whenever the glucose concentration within the capsule

increases, glucose competes for Con-A binding sites, releasing some Con-A from within the beads. The fluorescein/Concanavalin-A now is exposed to the external light, and emission occurs. Thus the magnitude of emission is directly related to the concentration of glucose in the tissue outside of the microdialysis tube. This sensor is designed to be placed subdermally and operated with an external light source and detector as shown in Fig. 9b.

With the recent sequencing of entire genes, the measurement of DNA fragments has become an essential clinical tool for tracking diseases. Biosensor approaches for DNA detection use the hybridization property of complementary DNA sequences as a molecular recognition event. One example is the synthesis of molecular beacons (Tyagi and colleagues^[13]) that are small nucleotide strands designed to self-hybridize by incorporating complementary DNA sequences at each end of the linear strand. At the two terminals of the strand, fluorescent dyes are attached. These dyes are selected to show fluorescence quenching when in close proximity. This is shown diagrammatically in Fig. 10a, the so-called stem-and-loop configuration. There are 5–8 bases at each of the two ends of the beacon that are complementary to each other. When the molecular beacon is exposed to a DNA fragment that is complementary to its sequence, it will bind and the structure will open up, separating the terminal dyes (Fig. 10b). When separated the dyes will fluoresce, and this signal can be used to identify the unknown DNA fragment in the sample.

S. Fodor^[14] introduced the concept of developing array technologies for massive analysis in biotechnology research in 1991. The group at Affymax

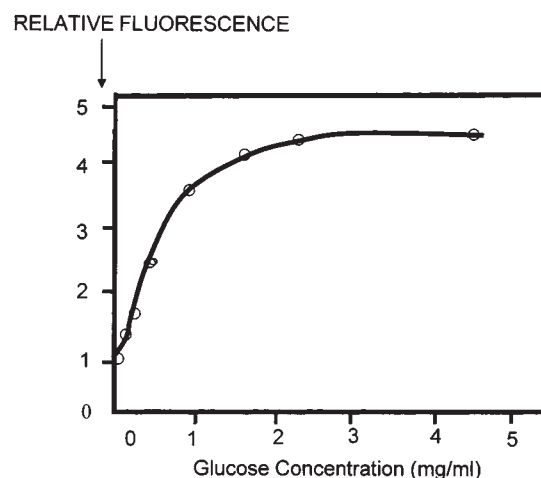


Fig. 8 Typical calibration curve for an affinity biosensor. This glucose biosensor data from the work of Schultz and Mansouri shows saturation behavior at high glucose levels corresponding to complete saturation of Concanavalin-A with glucose.

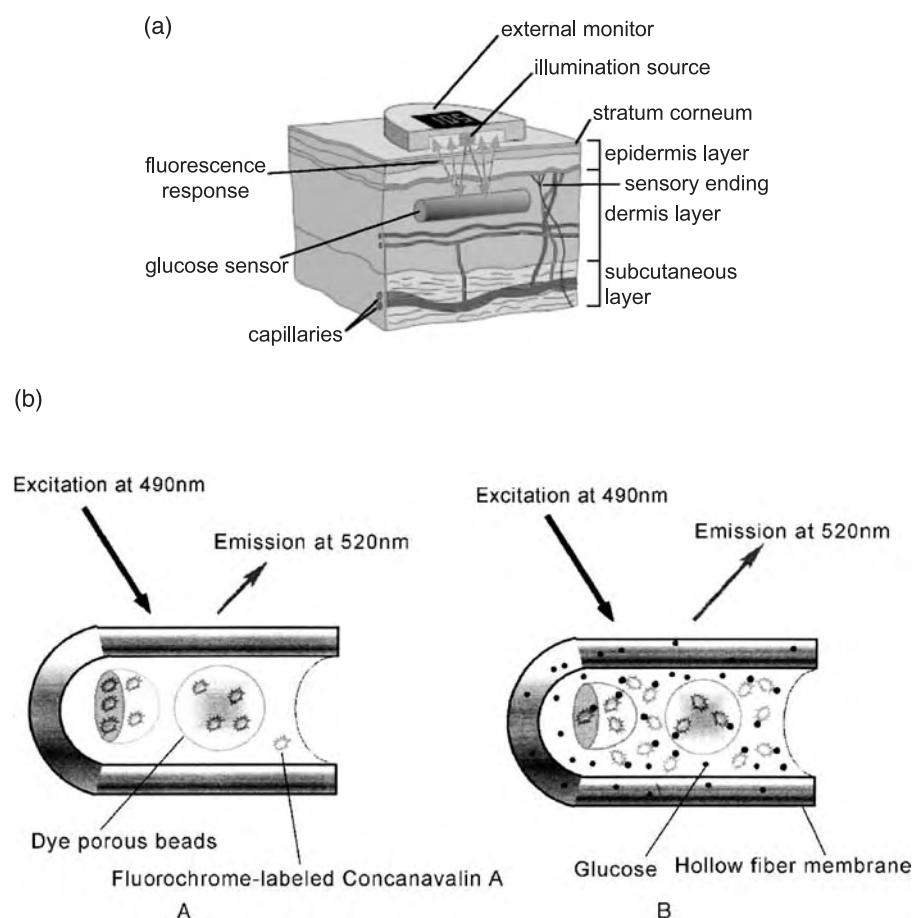


Fig. 9 Concept for a subdermal optical biosensor for various metabolites or drugs. A) The biosensor “capsule” consists of a microporous tube containing the sensor reagents. B) The glucose biosensor components consist of porous Sephadex beads coated by appropriate dyes to prevent fluorescence from fluorescein/Con-A adsorbed inside the beads. When the biosensor is exposed to glucose, glucose enters the capsule and competes with Sephadex for Con-A binding sites, resulting in the displacement of the fluorescent Con-A from the beads into the void space where they are excited by light and emit light at a longer wavelength. The emitted light is measured by a photonic device placed on the surface of the skin.

showed that gene sequencing and gene discovery could be accomplished by immobilizing thousands of polynucleotide fragments of known characteristics as a matrix of micron-sized spots on microscope-like slides. The array chip is then exposed to fluorescent-labeled gene fragments from a sample under study. Hybridization is allowed to occur, and then by the pattern of fluorescent spots on the chip, the DNA sequence of the unknown sample can be deduced. Array technology has been expanded to be useful in identifying genetic defects and in drug discovery.

For DNA chip technologies, the DNA fragments are bound to the surface. A variety of methods have been used to attach the DNA probes, but they usually involve a carboxyimide coupling using the five prime amino-modified nucleotides placed directly on the surface with a mechanical micropipette. Current commercial versions of these chips hold 20,000–40,000 probes in approximately a one-square-centimeter chip area. Another approach for putting the DNA on the chip is by using microjet dispensing systems. The protocol for using a DNA microarray chip usually is as follows. DNA is extracted from the sample of interest, and the DNA is amplified by a PCR technique. At some point during this process, fluorescent tags are incorporated into the DNA strands. The DNA is

separated into individual strands and then exposed or put on the microarray plate. DNA fragments will associate with immobilized DNA on the microarray plate wherever there is a match between the unknown DNA in the sample and a known fragment on the plate, resulting in a spot at that location. These spots where there is a combination between the DNA sample and the DNA on the plate will give rise to a fluorescent spot, and this can be measured with a reader. Since the DNA on the chip is known, the corresponding sequence on the DNA in the unknown is determined.

Another type of biosensor that uses artificial ion channels has been reported by Cornell et al.^[15] (Fig. 11). The basic mechanism for this type of device depends on the use of artificial bilayer membranes that have pore-forming materials in the surface, such as the antibiotic gramicidin, that are cross-linked to an affinity ligand. Gramicidin has a doughnut-type structure and randomly floats in each surface of the bilayer. Normally the different floating doughnuts do not coincide and there isn't an aqueous channel across the bilayer membrane through them. However, if due to binding events with the affinity ligand these doughnut-like molecules are aligned, aqueous pores are formed through the membrane, allowing a high flux of ions to pass through. Along with this ion flux is a large change

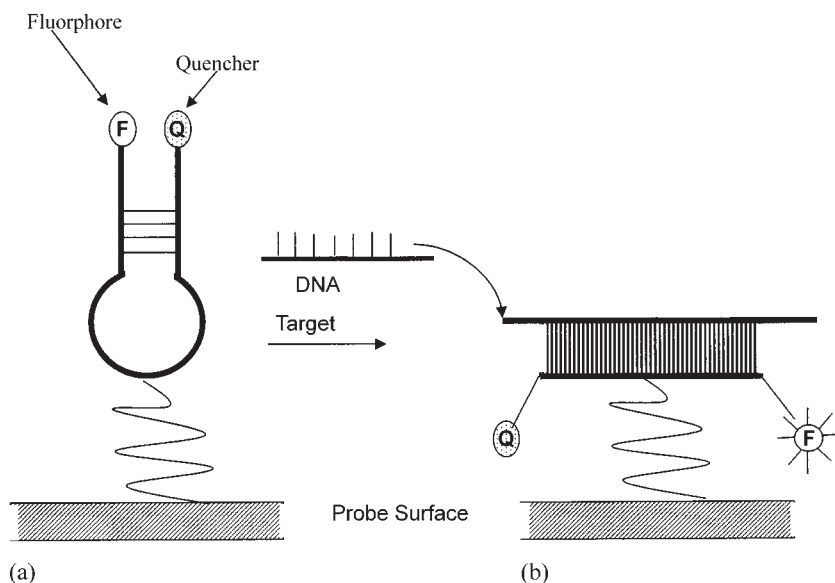


Fig. 10 “Molecular beacon” sensor mechanism. A DNA loop structure is synthesized containing the mating sequence to a DNA fragment to be identified. In the loop configuration (a), the terminal ends of the strand are in close proximity, causing fluorescence from the pendant dye molecule to be quenched. When the sensor with the molecular beacon on its surface is exposed to a sample containing the corresponding DNA fragment, the molecule opens up, separating the ends and resulting in an increase in emitted fluorescence (b).

in conductivity that can be quantified by impedance measurements. Thus the change in impedance due to the binding of the analyte to the affinity ligand composes this biosensor.

CELL-BASED SENSORS

One of the first applications of cells as biosensors has been in the field of pollutants (reviewed by Pancrazio et al.^[16]). Some pollutants may affect the metabolism of bacteria or microorganisms, and the change in the metabolism of these microorganisms to produce more

acid or decrease oxygen consumption is an indication of the concentrations of the pollutants. Typically cells are placed near a sensor that measures a change in the local environment—for example, either the pH or the oxygen concentration. In the presence of the pollutant, if the cells’ metabolism is changed, there will be a change in the concentration of one of these environmental components. By measuring the change in pH, for example, one can estimate the amount of the pollutant. A commercial device called a Cytosensor, once manufactured by Molecular Devices, used whole cells as biosensors. Essentially this instrument was a microphysiometer that consisted of very small wells in

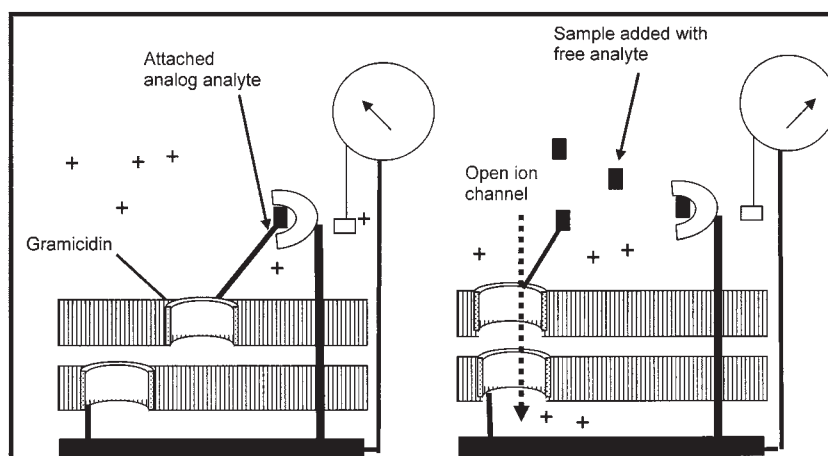


Fig. 11 Concept for a biosensor based on the conductivity through a bilayer (Cornell and coworkers). A membranelike bilayer is deposited on a conductive surface (e.g., gold). The bilayer is seeded with a molecule that has an aqueous core (e.g., gramicidin). In the lower bilayer, the gramicidin is anchored to the metal surface. The upper bilayer is seeded with gramicidin that is tethered to an analogue analyte whose motion is hindered because it is bound to a bioreceptor for the analyte. When the biosensor is exposed to free analyte, the upper gramicidin is released and moves freely in the upper bilayer. Occasionally (right panel) it intersects with a gramicidin molecule in the lower bilayer, creating an aqueous channel that allows the passage of ions and resulting in a measurable current.

a culture dish where the organisms were placed. An optical system to measure pH changes was used to measure changes in the metabolism of cells. In addition to toxicology studies, this type of device is useful for pharmacokinetic studies and drug screening and has been used with a variety of cellular types including neurons, hepatocytes, T-cells, etc.

Alternatively, instead of measuring the metabolic response of cells to various analytes, another sensitive indicator that reflects cell "well being" is the impedance of cells to electric fields. Cells have dielectric properties that are a function of the status of the cell membrane. Since intact living cells are very good insulators at low frequencies, slight changes in membrane properties can be measured accurately. By culturing cells over electrode arrays, changes in the electrode impedance allows a noninvasive method of following cell behavior such as adhesion, spreading, and motility. Impedance measurements have been used to monitor behavior of nonexcitable cell types including macrophages, endothelial cells, and fibroblasts.

Excitable cells such as neurons and myocardiocytes provide a very sensitive resource for monitoring certain analytes, particularly toxic agents such as drugs of abuse and nerve agents. Typically the cells are cultured on chips consisting of microarrays of electrodes. This type of biosensor is especially useful when one wants to establish whether the sample contains substances that may be harmful to cells, rather than to identify a particular material or its concentration. Also with these devices, signal propagation through synaptic transmission or gap junction connectivity can give a measure of cell coupling which is inaccessible under other techniques. Unfortunately, excitable cells are typically quite unstable, making this technique difficult to use routinely.

Another approach using microorganisms has been to use a genetically engineered protein that causes bioluminescence in the presence of the pollutant. Typically a hybrid plasmid is introduced into the organism with dual functionality. One section is activated by a particular analyte (e.g., a pollutant such as mercury), and the second section codes for an enzyme that produces an easily detected product. An example is the construct reported by Selifonova et al.^[17] The pollutant turns on the gene; in the same gene, coding for bioluminescent proteins is activated to give a visible signal.

DETECTOR ELEMENTS

Optical biosensors often rely on changes in color or fluorescence resulting from the biorecognition event and commonly use flexible fiber-optic wave guides to illuminate the sensor constituents that are placed at the distal end of the optical fiber.

The use of optical methods became popularized after the work of Lubbers and Opitz, who developed technology that they termed "optrodes." These are fiber-optic devices in which a chemically sensitive material such as pyrene butric acid that shows fluorescence quenching by oxygen is placed at the terminal end of a fiber-optic system. Light is transmitted along the optical fiber, and when it interacts with the material at the terminal end, a change in the characteristics of the light occurs, e.g., absorption, fluorescence, scattering, and polarization. Some of this light energy is captured and retransmitted back through the same or an alternative fiber to a photodetector device. The change in optical characteristics can then be related to the concentration of material at the end of the optical fiber.

Optical fiber detectors are typically utilized as either miniature spectrophotometers or fluorimeters and employ indicator dyes that can be detected by absorbance or fluorescence methods. For example, by measuring the absorption of solutions of phenol red at two wavelengths, the isobestic point at 490 millimicrons where absorption is virtually independent of pH and at 600 millimicrons where the intensity of absorption is very pH sensitive, one can estimate the pH of a solution or region optically. Thus, in the glucose biosensor systems shown in Fig. 2, such a fiber-optic pH sensor could be used to estimate the amount of gluconic acid formed.

Many different chromophores have been developed for use in bioassays. These compounds either change color as a result of some change in the environment such as a pH shift or change their fluorescent properties as a result of some chemical environmental change. Fluorescence methods tend to be more sensitive for analytical purposes if in the absence of the analyte light emission is minimal, resulting in a desirable high signal-to-noise ratio.

Another method uses a phenomenon known as surface plasmon resonance. This effect occurs at a surface that is coated with a thin layer of metal (e.g., gold). The critical angle of total internal reflection is then a function of the thickness of an adsorbed film on the metal surface. Lundstrom and colleagues^[18] showed that if the metal surface is first coated with a biorecognition component, such as an antibody, the subsequent binding of an analyte will cause a change in the angle of maximum reflection as shown in Fig. 12. Instruments using this technology are manufactured by Biacore and Thermo Lab Systems.

In situ label-free (i.e., without the use of chromophores) optical detection can be achieved through changes in other optical properties designed into the biorecognition event. For example, an array of nanoparticles changes its color as a function of spacing or aggregation. Asher^[19] showed that these particles can be coated with antibodies for a specific analyte and

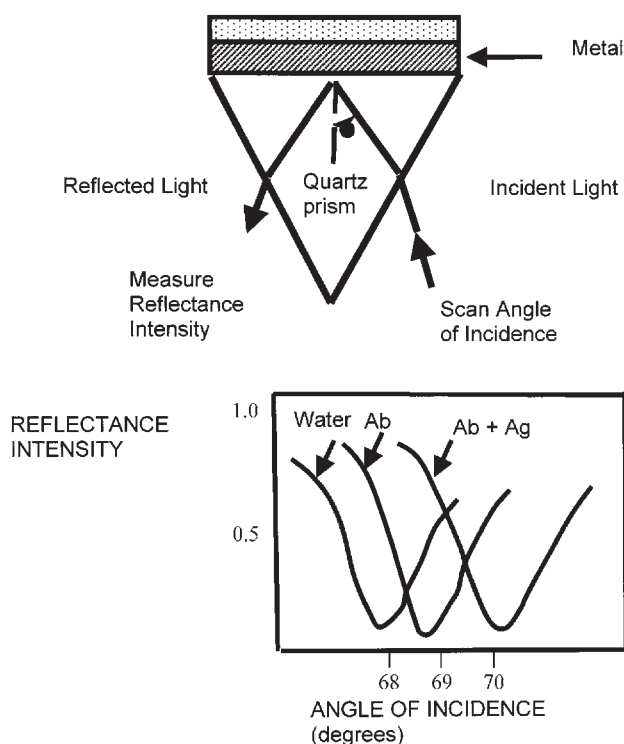


Fig. 12 An optical biosensor technique based on surface plasmon resonance. When the gold film is on the order of 50 nanometers in thickness, the angle of total internal reflection at a gold-quartz interface is directly proportional to the thickness of protein adsorbed on that surface. To make a biosensor, the surface is coated with a biorecognition species such as an antibody (Ab). When the surface is exposed to a solution containing the corresponding antigen (Ag), a thicker protein layer (Ab + Ag) forms, and the angle of maximum reflection changes as shown in the lower panel.

fabricated in a membrane form. When exposed to a sample containing the analyte, the analyte binds to the antibody, causing a change in the electric field around the particles. The particles will disperse and a change in color of the film will result.

Some other techniques include quartz crystal microbalance transducers. These are very sensitive mass-measuring devices that allow the dynamic monitoring of binding events. These devices consist of an oscillating crystal where the material is immobilized on the surface. When the analyte binds, the increase in mass associated with the binding events causes a decrease in oscillation frequency.

MICROFABRICATION

Miniaturization of devices and mass-production manufacturing techniques are some of the key reasons for commercial interest in biosensors at this time. Manufacturing technologies developed for completely

different applications, such as micromachining for integrated circuits and fiber optics components for telecommunications, have allowed rather novel designs to be contemplated and developed for biosensors. Another key feature of these technologies is the miniaturization of devices that one is able to achieve. Usually a microsized analytical system has components for the following functions: sample injection, preparation, separation, and detection. While current techniques for computer devices can make features at the nanometer level, for microfabricated analytical devices the usual feature dimensions are more likely at the micron level. This capability leads to the potential for very small, sensitive, and very stable devices that allow the development of portable and perhaps disposable biosensors. This will permit bringing the analysis system very close to the source of the analyte rather than the current mode of bringing samples of the analyte to centralized analysis laboratories. For example, if the application is environmental, one could put the sensor in remote locations to monitor many sites simultaneously. For medical uses, one can put the sensor in a catheter to be inserted into a blood vessel, or an individual could have a portable analyzer and periodically place a drop of blood in contact with the sensor.

An example of a commercial microanalysis system that utilizes microfabrication and incorporated fluidic circuits and biosensors is the product produced by i-STAT for point-of-care monitoring of blood chemistries. An exploded view of their device is shown in Fig. 13 with an explanation for the operation of this device.

SENSOR DYNAMICS

The response time of sensors to changes in concentration of an analyte depends on the response rate of the various components of a sensor. Usually, diffusional processes are the limiting factor; these include diffusion of the analyte to the surface of the sensor (dependent on external mixing), diffusion through membranes, and diffusion through the various regions of the sensor structure. Since diffusion lag times increase with the square of distance, it is imperative to maintain the active layers of the sensor to dimensions on the order of tenths of a millimeter or less. Enzyme reactions can usually be made nonlimiting if necessary by increasing the amount of enzyme in the system. In general terms, this behavior can be understood by considering the two processes involved that influence the behavior of biosensors. Diffusion of the analyte, such as glucose, into the sensor usually involves transport through a permeable membrane.

The i-STAT sensor arrays insure the integrity of the overall analytical process. The solid-state chips contain biosensors configured to perform specific tests with chemically sensitive membranes and films containing reagent chemicals. Sensors perform other functions such as monitoring the quality of the sample being tested. Silicon-type microfabrication utilizing high quality materials that exhibit exceptional stability allows consistent reproducibility in a high-volume manufacturing environment. This well accepted technology ensures that each cartridge offers a high level of accuracy and reliability.

When blood samples contact the sensors, they are measured electrochemically as follows:

Sodium, Potassium, Chloride, Ionized Calcium, pH and PCO₂ are measured by ion-selective electrode potentiometry. Concentrations are calculated from the measured potential through the Nernst equation.

Urea is first hydrolyzed to ammonium ions in a reaction catalyzed by the enzyme urease. The ammonium ions are measured by an ion-selective electrode and concentration is calculated from the measured potential through the Nernst equation.

Glucose is measured amperometrically. Oxidation of glucose, catalyzed by the enzyme glucose oxidase, produces hydrogen peroxide. The liberated hydrogen peroxide is oxidized at an electrode to produce an electric current which is proportional to the glucose concentration.

PO₂ is measured amperometrically. The oxygen sensor is similar to a conventional Clark electrode. Oxygen permeates through a gas permeable membrane from the blood sample into an internal electrolyte solution where it is reduced at the cathode. The oxygen reduction current is proportional to the dissolved oxygen concentration.

Hematocrit is determined conductometrically. The measured conductivity, after correction for electrolyte concentration, is related to the hematocrit.

A variety of calculated results are available that include HCO₃, TCO₂, BE, sO₂, Anion Gap and Hemoglobin.

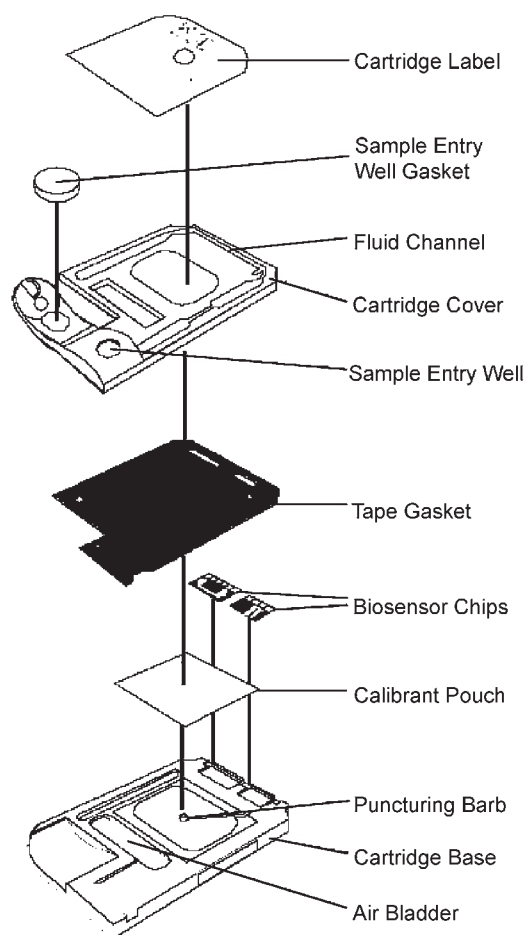


Fig. 13 Exploded view of a commercial biosensor for six different analytes manufactured by i-STAT.

$$N = PA (S_e - S_i) \quad (1)$$

where

N = rate of substrate transfer (moles/s)

P = permeability of membrane (cm/s)

A = area of membrane (cm²)

S_e = external concentration of substrate (mole/cm³)

S_i = internal concentration of substrate (mole/cm³)

The rate of consumption (e.g., conversion of glucose to gluconic acid) is given by the following equation:

$$r_c = \frac{kVES_i}{K_M + S_i} \quad (2)$$

where

r_c = rate of substrate consumption (mole/s)

k, K_M = kinetic constants

V = internal volume of biosensor (cm³)

E = enzyme concentration (mole/cm³)

At steady state

$$N = r_c$$

Or

$$PA (S_e - S_i) = \frac{kVES_i}{K_M + S_i}$$

From this relationship, it can be shown that increasing the enzyme concentration (E) results in a lowering of the internal substrate concentration (S_i), eventually

resulting in the condition $S_e \gg S_i$. Under these conditions, by Eq. 1, the rate of substrate diffusion into the biosensor (and thus the magnitude of response of the biosensor) will be nearly proportional to the external concentration of the analyte, S_e .

Some of the considerations that are important in determining the structure of sensors for use in analytical assays are discussed in the following sections.

Sensitivity

Sensitivity relates to the lowest concentration that the system can reliably detect. The biochemicals in blood can have average concentration ranges from milligrams per cubic centimeter to nanograms per cubic centimeter. The wide range of concentration of different biochemicals that are present in blood presents a technical challenge to the measurement of several analytes simultaneously in the same detection system. For example, colorimetric methods can provide reliable estimates of concentrations in the millimeter range, but fluorescence techniques often are required if the analyte is in the micromolar range of concentration.

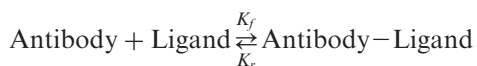
However, the rate of dissociation of antibody–analyte complexes decreases directly with increasing binding affinity. For high levels of sensitivity, high binding constants are needed. For example, to measure an analyte at a concentration of nanomoles, a binding constant on the order of nanomoles is required. The dissociation rate for such antibodies is on the order of tens of minutes.

Dynamic Range

In addition to sensitivity of an assay system, the range of sensitivity of the device is another important consideration. The dynamic range is usually defined as the ratio of highest concentration to the lowest concentration that a particular technique can reliably measure. For many analytical methods based on biosensors, the dynamic range is rarely more than a factor of about 10. Thus for any particular application, the system needs to be engineered to achieve a dynamic range that covers the normal variation in concentration for that analyte. For example, for the measurement of blood glucose, the normal level is about 100 milligrams/deciliter and the dynamic range should be between 50 and 200 milligrams/deciliter and somewhat higher for diabetics. For systems that use biological materials such as enzymes or antibodies, this means that one has to select the appropriate protein with the correct binding characteristics or reactive characteristics for the molecule that is to be measured. Fortunately, with the advent of modern molecular biology, one can either select or modify the protein's

active site to change its selectivity to fulfill the analytical requirements.

An equation that depicts the interaction of an antibody with a ligand is



The affinity constant for these antibodies is the concentration at which the protein is half saturated with the ligand.

$$K_a = K_f / K_r$$

where

K_f = forward rate constant

K_r = reverse rate constant

K_a = equilibrium constant

As can be seen by Eq. 2, enzymes have kinetic characteristics that show saturation behavior. That is, when the numerical value of the substrate concentration (S_i) is much larger than the kinetic constant K_M , the right-hand side increases to a constant asymptotic value of kVE .

The parameter for enzymes that characterizes the sensitivity level for sensors based on them is the kinetic parameter known as the Michaelis Menton constant (K_M). Again looking at the behavior of Eq. 2, when the substrate concentration is numerically equal to K_M (i.e., $S_i = K_M$) or smaller, the reaction rate is approximately linearly proportional to substrate concentrations. Enzymes can be found in nature or readily genetically modified to have K_M that vary over five orders of magnitude. Table 2 lists K_M values for some enzymes that have been used in biosensors.

Response Time

The response behavior of sensor-based analytical devices is affected by a number of parameters. With respect to the sensor component, response characteristics are usually determined by a combination of diffusional and reaction kinetic processes. The usual configuration for a biosensor includes a membrane that protects the sensor element. Thus, in the Clark glucose sensor shown in Fig. 7, the diffusion processes include the passage of glucose and oxygen from the external fluid through the dialysis membrane into the enzyme compartment and the subsequent passage of oxygen through the silicone rubber membrane into the electrode compartment.

The approximate lag time introduced by diffusion layers (either membranes or liquid layers) around or

within biosensors is related to the thickness (L) and analyte diffusivity (D) through the layer.

$$\text{Lag time (s)} = 0.2L^2/D$$

where

D = diffusivity of solute in the layer (cm^2/s)

L = thickness of layer (cm)

If the biological element is an enzyme then the kinetics of the reaction of analyte will be governed by an equation similar to that given in Eq. 2 and the overall rate of reaction will be determined by the enzyme concentration in the preparation (E) and the inherent substrate turnover of the enzyme characterized by V_m .

If the biological element is an antibody or a bioreceptor, the kinetics will be related to the magnitude of the affinity constant K_a . The affinity constant is essentially the equilibrium constant for the reversible association and dissociation of the ligand for the antibody. Kinetic studies of these reactions have shown that an increase in binding constants is directly correlated with a reduction in the inherent dissociation rates of these reactions. The implication of this behavior is that the response rate of sensors based on antibodies becomes slower with higher affinities.

An approximate value for the association rate constant (K_f) for ligands with antibodies is $10^8 \text{ M}^{-1} \text{ s}^{-1}$. Thus, an approximation for the lag time for dissociation of antibody–ligand complexes is

$$\text{Lag time (s)} = 10^{-8}/K_a$$

where

K_a = ligand–antibody association equilibrium constant (M^{-1})

ARTICLES OF FURTHER INTEREST

Biopotential Amplifiers; Glucose Sensors; Microelectromechanical Systems (MEMS) Manufacturing

REFERENCES

- Clark, L.C., Jr. Monitoring and control of blood tissue PO_2 tensions. *Trans. Am. Soc. Artif. Intern. Organs* **1956**, 2, 41–48.
- Clark, L.C., Jr.; Lyons, C. Electrode systems for continuous monitoring in cardiovascular surgery. *Ann. N.Y. Acad. Sci.* **1962**, 102, 29–45.
- Lubbers, D.W.; Opitz, D. Quantitative fluorescence photometry with biological fluids and gases. *Arch. Exp. Med. Biol.* **1976**, 75, 6511.
- Peterson, J.I.; Goldstein, S.R.; Fitzgerald, R.J.; Ruckold, D.R. Fiberoptic pH probe for physiological use. *Anal. Chem.* **1980**, 52, 864–869.
- Kulp, T.J.; Camins, I.; Angel, S.M.; Munkholm, C.; Walt, D.R. Polymer immobilized enzyme optodes for the detection of penicillin. *Anal. Chem.* **Dec. 15, 1987**, 59 (24), 2849–2853.
- Fischer, U.; Rebrin, K.; von Woedtke, T.; Abel, P. Clinical usefulness of the glucose concentration in the subcutaneous tissue—Properties and pitfalls of electrochemical biosensors. *Horm. Metab. Res.* **1994**, 26, 515–522.
- Armour, J.C.; Lucisano, J.Y.; McKean, B.D.; Gough, D.A. Application of chronic intravascular blood glucose sensor in dogs. *Diabetes* **1990**, 39, 1519–1526.
- Cass, A. *Biosensors. A Practical Approach*; IRL Press at Oxford University Press: Oxford, 1990; 271 pp.
- Berson, S.A.; Yalow, R.S. General principles of radioimmunoassay. *Clin. Chim. Acta* **1968**, 22, 51–69.
- Liu, B. Characterization of Immunosensor Binding. Ph.D. Thesis; University of Michigan; 1986.
- Schultz, J.S.; Mansouri, S.; Goldstein, I.J. Affinity sensor: A new technique for developing implantable sensors for glucose and other metabolites. *Diabetes Care* **1982**, 5, 245–253.
- Ballerstadt, R.; Schultz, J.S. A fluorescence affinity hollow fiber sensor for continuous transdermal glucose monitoring. *Anal. Chem.* **1999**, 72, 4185–4192.
- Tyagi, S.; Kramer, F.R. Molecular beacons: Probes that fluoresce upon hybridization. *Nat. Biotechnol.* **1996**, 14, 303–308.
- Fodor, S.P.A.; Need, J.L.; Pirrung, M.C.; Stryer, L.; Lu, A.T.; Solas, D. Light directed, spatially addressable parallel chemical synthesis. *Science* **1991**, 251, 767–773.
- Cornell, B.A.; Braach-Maksvytis, V.L.B.; King, L.G.; Osman, P.D.J.; Raguse, B.; Wiczorek, L.; Pace, R.J. A biosensor that uses ion-channel switches. *Nature* **1997**, 387, 580–583.
- Pancrazio, J.J.; Whelan, D.A.; Borkholder, W.M.A.; Stenger, D.A. Development and application of cell-based biosensors. *Ann. Biomed. Eng.* **1999**, 27, 697–711.
- Selifonova, O.; Burlage, R.S.; Barkay, T. Bioluminescent sensors for detection of bioavailable Hg(II) in the environment. *Appl. Environ. Microbiol.* **1993**, 59, 3083–3090.
- Liedberg, B.; Nylander, C.; Lundstrom, I. Biosensing with surface plasmon resonance—how it all started. *Biosens. Bioelectron.* **1995 Fall**, 10 (8), i–ix.
- Holtz, J.H.; Asher, S.A. Polymerized colloidal crystal hydrogel films as intelligent chemical sensing materials. *Nature* **1997**, 389, 829–832.

FURTHER READING

- Buck, R.P.; et al. *Biosensor Technology. Fundamentals and Applications*; Harold Dekker: New York, 1990; 419 pp.

Hall, A.H. *Biosensors*; Open University Press: Buckingham, 1990; 351 pp.

Janata, I. *Principles of Chemical Sensors*; Plenum Press: New York, 1989; 317 pp.

Kress-Rogers, E. *Handbook of Biosensors and Electronic Noses*; CRC Press: New York, 1996; 695 pp.

Taylor, R.F.; Schultz, J.S. *Handbook of Chemical and Biological Sensors*; Institute of Physics Press: Philadelphia, 1996; 604 pp.

Turner, A.P.F.; Karube, I.; Wilson, G.S. *Biosensors. Fundamentals and Applications*; Oxford Science Publications: Oxford, 1987; 770 pp.

B

Biosynthesis and Applications of Alginates

Bernd H. A. Rehm

Institute of Molecular BioSciences, Massey University, Palmerston North, New Zealand

INTRODUCTION

Alginate is a nonrepeating copolymer of β -D-mannuronic acid (M) and α -L-guluronic acid (G), linked by 1–4 linkages. Alginates adopt in nature various material properties ranging from slimy and viscous solutions to pseudoplastic materials when cross-linked with divalent cations. The comonomers are arranged in blocks of continuous M-residues (M-blocks), G-residues (G-blocks), or alternating residues (MG-blocks) (Fig. 1). Alginates are produced by bacteria and brown seaweeds. The M-residues of only the bacterial alginates are acetylated to a variable extent at positions O-2 and/or O-3. The variability in comonomer blocks and the acetylation degree strongly affect the material properties of the polymer. To understand how biological systems control this variability is of scientific and applied importance. Moreover, the analysis and understanding of the structure–function relationships of alginates would provide the basis for the production of tailor-made polymers with particular material properties for a certain application.

Since the early 20th century, alginate is an industrial product obtained commercially by harvesting of brown seaweeds from coastal regions. Only bacteria belonging to the genera *Pseudomonas* and *Azotobacter* are capable of producing significant quantities of this polymer as an exopolysaccharide. In both genera, the genetics, gene regulation, and biochemistry of alginate biosynthesis have been well characterized and only differences with respect to the alginate modifying enzymes called epimerases were obtained. In contrast to *Pseudomonas*, *Azotobacter vinelandii* produces various extracellular epimerases that enable enhanced control of epimerization of nascent alginate. Nevertheless, the polymerization process and the export of the alginate are still not known. The direct precursor of polymerization is GDP-mannuronic acid, which is polymerized to polymannuronate. This homopolymer can be further modified by acetylation at positions O-2 and/or O-3 and by epimerization, leading to a variable content of acetyl groups and G-residues, respectively. The G-residue content and distribution strongly affect the gel-forming capacity of alginates. Thus, the epimerization process is of great interest from an applied point of view. Alginate is currently used for a variety of industrial purposes, and the production is hitherto exclusively based on brown seaweeds. This entry provides

an overview of the state of art in alginate biosynthesis and the biotechnological potential of the production of tailor-made alginates focusing on medical application. Alginates have been considered as valuable biomaterials because of their interesting material property range as well as their biocompatibility. The use of highly purified alginates in medical application is already targeted by strong research and development activities as well as by steadily increasing market shares. Highly purified alginates can serve for immunoisolation of allogenic and xenogenic cells as well as tissues enabling long-term and functional implantation into the human body, as shown for allogenic parathyroid tissue implanted into patients with severe hypoparathyroidism. Further examples comprise treatment of Parkinson's disease with neuroblastic cells and diabetes mellitus using encapsulated porcine Langerhans islets.

Biomaterials play already an important role in medicine. Natural and synthetic polymers are used as biomaterials implementing a multidisciplinary approach including wound dressings, dental implants, tissue engineering, and microencapsulation of drugs, peptides, and cells as controlled drug delivery systems. The current crucial problems in purity, biocompatibility, and batch-to-batch reproducibility, especially in the case of naturally occurring polymers, must be addressed in future to enable application in medicine. The main commercial sources are species of *Ascophyllum* and *Laminaria* (Europe), *Lessonia* (South America), *Ecklonia* (South Africa), *Durvillaea* (Australia and Chile), and *Macrocystis* (California). Polyphenolic compounds and proteins, which are difficult to remove, contaminate alginates derived from harvested seaweeds and result in bioincompatibility. Hence, further thorough purification of alginates from seaweeds is required for implantation and other medical applications. These circumstances do certainly attract more interest in biotechnological production of alginates using bacteria in defined production processes.

ALGINATE BIOSYNTHESIS

Biosynthesis of the Alginate Precursor, GDP-Mannuronic Acid

The GDP-mannuronic acid is the activated precursor for alginate polymerization. Early biochemical

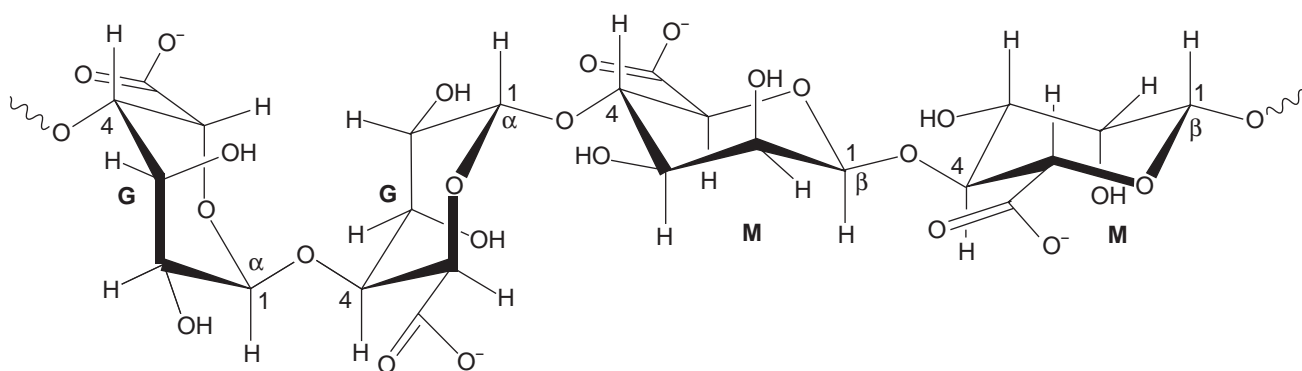


Fig. 1 Secondary structures of polymannuronate (M-blocks) and polyguluronate (G-blocks).

analysis of brown algae indicated enzyme activities related to alginate biosynthesis. The phosphomannose isomerase–guanosine diphosphomannose pyrophosphorylase (PMI–GMP) is a bifunctional protein catalyzing the initial and third steps of alginate synthesis. The PMI catalyzed reaction channels fructose-6-phosphate into alginate synthesis via formation of mannose-6-phosphate. Mannose-6-phosphate is converted to mannose-1-phosphate catalyzed by phosphomannomutase. The GMP activity of the PMI–GMP catalyzes conversion of mannose-1-phosphate to GDP-mannose while hydrolyzing GTP. Although the enzyme favors the reverse reaction, the efficient subsequent removal of GDP-mannose in the next step enables efficient channeling toward alginate synthesis. GDP-mannose is oxidized to GDP-mannuronic acid catalyzed by the enzyme guanosine diphosphomannose dehydrogenase (GMD). These first biosynthesis steps were confirmed in bacteria and brown algae (Fig. 2).^[1,2] Because the intracellular enzyme activities are rather low, it has been assumed that the enzymes PMI–GMP and PMM form an enzyme complex (“metabolon”).

Genes Involved in Alginate Biosynthesis

The alginate biosynthesis genes involved in the synthesis of the precursor GDP-mannuronic acid have all been identified and characterized in bacteria, and they have been designated *algA* (encoding PMI–GMP), *algC* (encoding phosphomannomutase), and *algD* (encoding GMD). With the exception of *algC* (encoding phosphomannomutase), the other two genes and all other known structural genes involved in alginate biosynthesis in *bacteria* are clustered (Table 1). So far, the respective genes were not identified in eukaryotic alginate producers such as brown algae. Several bacterial gene products putatively involved in

the polymerization process are functionally not understood. This is due to a lack of biochemical characterization of the respective proteins and to a lack of an *in vitro* alginate synthesis assay. The gene products of *alg8*, *alg44*, *algK*, and *algX* are candidates for being subunits of the alginate polymerase. The deduced amino acid sequences of these proteins contain hydrophobic regions, indicating a localization in the cytoplasmic membrane.^[4,5] Because the *algK* deletion led to the formation of uronic acids, a role in polymerization had been suggested.^[6] Recent biochemical characterization of AlgX showed that this protein is localized in the periplasm and required for alginate formation.^[7,8] Because no undecaprenol-linked intermediate has been identified in *Pseudomonas*

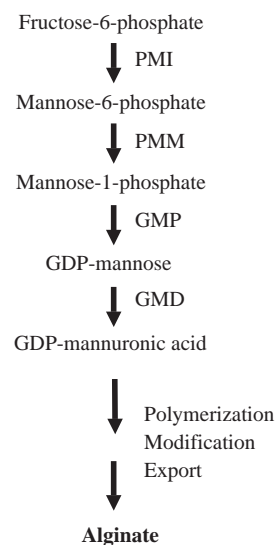


Fig. 2 Alginate biosynthesis pathway. PMI, phosphomannose isomerase; PMM, phosphomannose mutase; GMP, GDP-mannose pyrophosphorylase; GMD, GDP-mannose dehydrogenase.

Table 1 Bacterial alginate biosynthesis genes

Gene	Gene product/function
<i>AlgD</i>	GDP-mannose dehydrogenase
<i>alg8</i>	Polymerase/export function
<i>alg44</i>	Polymerase/export function
<i>AlgK</i>	Polymerase/export function
<i>AlgE</i>	Outer-membrane porin/export
<i>AlgG</i>	Mannuronan C-5-epimerase/biosynthesis
<i>AlgX</i>	Putative polymerase/export function and high-sequence similarity to <i>algJ</i>
<i>AlgL</i>	Alginate lyase/biosynthesis
<i>algI</i>	O-Acetylation
<i>AlgJ</i>	O-Acetylation
<i>AlgF</i>	O-Acetylation
<i>Alga</i>	PMI/GDP-mannose yrophosphorylase
<i>AlgB</i>	Member of <i>ntrC</i> subclass of two-component transcriptional regulators
<i>AlgC</i>	Phosphomannomutase
<i>AlgH</i>	Unknown function
<i>algR1</i>	Regulatory component of two-component signal transduction system
<i>algR2 (algQ)</i>	Protein kinase or kinase regulator
<i>algR3 (algP)</i>	Histone-like transcription regulator
<i>AlgZ</i>	AlgR cognate sensor
<i>AlgU (algT)</i>	Homologous to <i>E. coli</i> σ^E global stress response factor
<i>algS (mucA)</i>	Anti- σ factor
<i>algN (mucB)</i>	Anti- σ factor
<i>AlgM (mucC)</i>	Regulator
<i>algW (mucD)</i>	Homologous to serine protease (HtrA)

(Adapted from Ref. [3].)

aeruginosa or *A. vinelandii*, the polymerase, presumably localized as a protein complex in the cytoplasmic membrane, could catalyze alginate formation via an undecaprenol-independent mechanism.^[2] Thus, alginate synthesis might be similar to bacterial cellulose synthesis. In *Acetobacter xylinum* the cellulose synthase, a protein complex of 420 kDa, is localized in the cytoplasmic membrane, and catalyzes the processive polymerization of glucose residues.^[9] Accordingly, the alginate polymerase in the cytoplasmic membrane might bind the GDP-mannuronic acid at the cytosol site leading to polymerization and simultaneous translocation of nascent alginate.^[3]

In the alginate biosynthesis cluster of bacteria a gene encoding an outer membrane protein has been identified.^[10–13] Biochemical and biophysical characterization of the outer membrane protein indicated an anion-selective pore function.^[12] Topological models of these outer membrane proteins were developed showing these proteins as β -barrels consisting of 18 β -strands. Since these channel proteins could be partially blocked by GDP-mannuronic acid, these proteins

might form alginate-specific pores, which enable export of the nascent alginate chain through the outer membrane. Genes encoding the three alginate modifying enzymes (lyase, transacetylase, and C5-mannuronan epimerase) were identified and characterized in bacteria.^[3] Recently, a family of C5-epimerase genes has also been identified in the brown algae *Laminaria digitata*.^[14] Modifications of bacterial alginate are introduced after polymerization and presumably in the periplasm. The bacterial proteins involved in transacetylation (AlgI, AlgJ, AlgF), the epimerase (AlgG), and the lyase (AlgL) have been characterized.^[15–17] O-2 and/or O-3 positions of the mannuronic acid residue are targeted by the transacetylase. Acetylated residues are protected against epimerization to guluronic residues by the epimerase AlgG. Thus, the degree of epimerization is controlled by the transacetylase activity. Increasing acetylation degree results in changes of the physicochemical properties of alginate impacting on its role in biofilm formation.^[18] Thus, acetylation might be particularly relevant under dehydrating conditions, e.g., during infection and

colonization of the lungs from cystic fibrosis patients. The alginate lyase might function by controlling the molecular weight of the alginate and/or provides primer for the polymerase. However, this biosynthesis lyase does not play a role in utilizing alginate as a carbon source.

In the bacterium *A. vinelandii*, it has been found that the genome encodes a seven-member family (AlgE1-7) of Ca²⁺-dependent mannuronan C-5 epimerases, which are secreted to the surface and extracellular environment.^[19-21] Interestingly, after production in *Escherichia coli*, each epimerase has been found to introduce its characteristic sequence distribution of G residues in the substrate, explaining the extensive variability in the monomer sequences of the alginates produced by *A. vinelandii*.^[20] It is probable that this ability is biologically significant for the formation of a particular resting stage, called cyst.^[22,23] These enzymes, therefore, potentially explain the monomer sequence differences observed in *P. aeruginosa* and *A. vinelandii* alginates. Further aspects of these discoveries are discussed in the paragraph below dealing with the potential applications of bacterial alginates.

Surprisingly, the *A. vinelandii* genome also encodes a biosynthesis mannuronan C5-epimerase (AlgG) that belongs to another class than AlgE1-E7, and this epimerase seems to represent the equivalent of AlgG in *P. aeruginosa*.^[24] In contrast to the extracellular epimerases (AlgE1-E7), the biosynthesis epimerases were found to be required for alginate production.^[25,26]

Regulation of Alginate Synthesis

All the information about regulation of alginate biosynthesis has been derived from bacteria. The regulatory genes, genotypic switch genes, and biosynthesis genes are not colocalized in bacteria. The alginate structural genes are controlled via the positively regulated *algD* promoter.^[27] In *A. vinelandii*, on the other hand, there seems to exist an additional *algD*-independent promoter, from which at least *algL* (the lyase gene) and *algA* (encoding PMI-GMP) are transcribed.

Regulation of alginate biosynthesis is becoming increasingly complex and involves various specific and globally acting gene products (Figs. 3 and 4).^[28] Alginate biosynthesis is regulated based on environmental stimuli and/or by a genotypic switch comprising a mutation inactivating the antisigma factor MucA.^[29]

Alginate production is regulated by a number of transcriptional factors.^[30] The genotypic switch region comprising *algU*, *mucA*, *mucB*, *mucC*, and *mucD* is found in bacteria, and the corresponding gene sequences are highly homologous (see Ref. [3] for a

review). Regulation of alginate biosynthesis by the sigma factor AlgU (AlgT) via the Muc proteins has been extensively investigated.^[29,31-37] All the known *alg* genes and their corresponding proteins are listed in Table 1.

Mutations in *mucA* prevent MucA, an anti-sigma factor, from interaction with AlgU, leading to AlgU-mediated transcriptional activation of the alginate biosynthesis operon. Additionally, the periplasmic protein MucB is required for alginate regulation because mutations in *mucB* also cause the conversion of *P. aeruginosa* from a nonmucoid to a mucoid phenotype. Mutations in MucD were also found to enhance transcription of the alginate biosynthesis operon.^[38] For MucC a negative regulatory role in alginate production had been suggested, too.^[39]

The two-component transcriptional regulator AlgR is required for alginate production in response to environmental stimuli.^[40-42] Environmental conditions such as high osmolarity, N- and P-starvation, dehydration, and the presence of phosphorylcholine activates a cascade of regulatory proteins involved in activation of *algD* promoter-dependent transcription. The regulator AlgR binds to three sites within the *algD* promoter and thereby activates transcription of the alginate operon. Furthermore, AlgR regulates alginate production through *algC* by binding to its promoter.^[43] AlgR binds to three positions within the *algC* promoter region, but the orientation and positioning of the AlgR binding sites differ in the *algC* and *algD* promoters. AlgR exerts a more global function by regulating additional genes besides those required for alginate production. AlgR has been shown to be required for twitching motility.^[43-46] AlgR overexpression was recently used in *P. aeruginosa* to identify genes regulated by AlgR using DNA-microarray analysis. These data suggested an even more global role of AlgR.^[44,46,47]

The entire alginate biosynthesis gene cluster of *P. aeruginosa* is transcribed under control of the *algD* promoter and the region (*algD*-*algA*) comprises an operon.^[48] In *A. vinelandii* two additional *algD*-independent promoters were identified as being involved in the transcription of *alg8*-*algJ* and *algG*-*algA*.^[49,50]

MEDICAL APPLICATIONS OF ALGINATES

Besides traditional applications of algal alginates as food additives, alginates have been increasingly applied in the medical field particularly as dressings in surgery and wound management.^[51] Large quantities of alginate dressings are used each year to treat exuding wounds, such as leg ulcers, pressure sores, as well as infected surgical wounds. These dressings are

made of Ca^{2+} -alginate fibers to form a product with increased fabric strength when it is soaked with exudate or blood. Some products contain a proportion of sodium alginate to enhance the gelling properties of the applied dressing. Other dressings have been produced from freeze-dried alginate. Upon contact with an exuding wound, the calcium ions in the dressing are exchanged by the sodium ions in serum or wound fluid. Replacement by sodium causes the fibers to swell and partially dissolve, resulting in an amorphous gel. This also makes it easy to remove these dressings from large open wounds or burns because they do not adhere to the wound. Removal can be assisted by applying saline solutions to the dressing to ensure its conversion to soluble sodium alginate. The degree of swelling is strongly dependent on the chemical composition of the alginate, which varies between algal species. However, some evidence was provided that alginates may influence wound healing in an unknown way. Recently, the consumer division of a multinational pharmaceutical company launched a new line of adhesive bandages and gauze pads based on calcium alginate fibers. They are being promoted as helping blood to clot faster—twice as fast as their older, well-established product.

The alginic acid powder swells when in contact with water. This property was used for the application as disintegrant in certain tablets. The swelling property was also used in dietary foods to provide a “full” feeling, and in tablets taken to relieve heartburn and acid indigestion. The swollen alginic acid reduces reflux irritating the lining of the esophagus.

Alginate is also used in controlled release of medical drugs and other chemicals. The active ingredient is entrapped in calcium alginate beads and slowly released upon exposure to certain environmental conditions. Recently, alginate microspheres, sometimes coated with chitosan to enhance the mechanical strength, have been developed and tested as oral controlled-release systems for drug delivery.

NONMEDICAL APPLICATIONS

Paper: Alginate is used in the paper industry for surface sizing and high-performance printability.

Welding rods: Alginate is applied in coatings to welding rods or electrodes. The quantities of alginates used are very dependent on the type of welding rod being coated and the extrusion equipment being used.

Binders for fish feed: Crude alginate is used as a binder in fish feeds. Bound alginate can lower the consumption by up to 40% and culture ponds are less polluted.

Release agents: Alginates are used as mold release agents for the formation of fiberglass plastics.

POTENTIAL APPLICATIONS OF BACTERIAL ALGINATES

Because the prices of algal alginates are low, the establishment of a competitive bacterial production process appears to be rather difficult unless high-value applications are targeted. The following two strategies are currently pursued: 1) optimization of the bacterial production process to reduce prices and 2) bioengineering of alginates with high-value applications justifying high production costs. However, it is increasingly more likely that bacterial alginates may become commercial products because environmental concerns related to the traditional seaweed harvesting and processing are increasing as well as the demand for higher-quality alginates. Because the past decades provided extensive insight into the molecular processes of alginate biosynthesis in bacteria, this knowledge is an important foundation for manipulation and exploitation of the microbial production process. Strategies for the production of high-quality and/or tailor-made bacterial alginates with improved qualities will be discussed on the basis of recent advances in alginate research.

Bioengineering of Bacterial Alginates

The material properties of alginates are determined by their molecular weight, degree of acetylation, and monomer composition and sequence. Based on the current knowledge it seems likely that all these three parameters can be controlled in the bacterial alginates. In contrast, seaweeds harvested from the ocean will not enable this degree of control and therefore the product range will be limited by the potential of fractionation of polymer mixtures and upgrading using alginate modifying enzymes.

The degree of polymerization impacts on viscosity properties of alginates. It might be possible to control this parameter in bacteria by controlling both the fermentation conditions and the alginate lyase activity. Bacterial production could provide alginates with a given viscosity by means of reduced molecular weight. However, very high viscosities will impair oxygen transfer in fermentors and sophisticated aeration techniques will have to be developed and applied. Whether this will be feasible is currently unclear.

The most interesting potential provided by bacterial alginates is based on the possibility of controlling their monomer composition as well as sequential structures. The mannuronan C-5-epimerases catalyze the conversion of mannuronic acid residues into guluronic residues in the polymer chain without breaking the glycosidic bonds. This reaction is similar to the conversion of β -D-GlcA into α -L-IdoA in the biosynthesis

of heparin, heparan sulfate, and dermatan sulfate.^[52] According to the current knowledge it appears that *Pseudomonas* species have only one mannuronan C-5-epimerase, encoded by the *algG* gene, and this enzyme is able to introduce only single G residues in the polymannuronate chain. These alginates are strongly impaired in their gel-forming capacity. Divalent cation-dependent (typically Ca^{2+}) gels are formed by alginates containing G-blocks. Thus, the in vivo manipulations of monomer composition in *Pseudomonas* alginates is rather limited. However, inactivation of the *algG* gene should enable production of pure mannuronan, and supporting evidence for this has been obtained by Franklin et al.^[15] Polymannuronate is known to be a strong immunostimulant and could have a commercial potential in applications where this property is of interest.

In *A. vinelandii* the epimerization system is much more complicated. This bacterium encodes seven secreted Ca^{2+} -dependent epimerases, which have been sequenced, cloned, and produced recombinantly in *E. coli*.^[53,54] Each of these recombinant enzymes generates specific nonrandom epimerization patterns when acting upon mannuronan or alginate as substrate. The AlgE4 epimerase catalyzes predominantly an alternating residue sequence, whereas the action of AlgE1, AlgE2, and AlgE6 generates long G blocks.^[24,53–57] It therefore seems clear that the monomer sequence variability observed in alginates isolated from natural sources may originate from the action of a set of different epimerases, each representing a particular specificity with respect to the final reaction product.

Insight into enzymes involved in sequence variability provides a potential for applications. Tailor-made alginates could be made by initially producing deacetylated poly-mannuronic acid using *Pseudomonas* mutants, which are deficient in acetylation and thus enable epimerization.^[16,18] Alternatively, the acetyl groups could be removed from the purified polymer by standard chemical deacetylation procedures. Such polymannuronates could be epimerized by one particular recombinantly produced epimerase, an engineered enzyme or a combination of such enzymes. It is also conceivable to fractionate and purify commercially produced seaweed alginates employing these recombinant epimerases. It is now likely that almost any alginate structure could be produced by such processes aiming at particular material properties of applied interest.^[19,58–60]

Besides the bioengineering of alginates, an enormous design space is provided by the combination of other biopolymers with alginate. Recent studies indicate a promising potential of composites for future applications, particularly in the medical field.^[61,62]

Bacterial alginates might be first applied as immunostimulants or for immobilization of cells. Cells

immobilized in gel-forming alginates might be used for various biotechnological production processes or in medical transplantation technologies.^[63–66] The reversal of Type I diabetes by immobilizing insulin-producing cells in alginate capsules has been already achieved. These capsules have been implanted into the body of whole animals and even humans, and long-term biological efficacy was found.^[67,68]

CONCLUSIONS

Currently, pharmaceutical and medical uses of alginates are about 20% of the market value with 2–4% annual growth rates, which are maintained by continuous developments in controlled release technologies and the use of alginates in wound care applications. Food applications are worth about 20% of the market value and the annual growth of 1–2% is rather low. The best profitability is related to high-value applications, such as pharmaceutical and medical applications. This has already justified research and development activities implementing bacterial alginates using bioengineering approaches, grafting, and blending to obtain high-performance materials for medical applications.

ARTICLE OF FURTHER INTEREST

Biocatalysis

REFERENCES

1. Lin, T.Y.; Hassid, W.Z. Pathway of alginic acid synthesis in the marine brown alga, *Fucus gardneri* Silva. *J. Biol. Chem.* **1966**, *241*, 5284–5297.
2. Rehm, B.H.; Valla, S. Bacterial alginates: biosynthesis and applications. *Appl. Microbiol. Biotechnol.* **1997**, *48*, 281–288.
3. Rehm, B.H.A. *Alginates from Bacteria*; Wiley-VCH, 2002.
4. Maharaj, R.; May, T.B.; Wang, S.K.; Chakrabarty, A.M. Sequence of the *alg8* and *alg44* genes involved in the synthesis of alginate by *Pseudomonas aeruginosa*. *Gene* **1993**, *136*, 267–269.
5. Wang, S.K.; Sa'-Correia, I.; Darzins, A.; Chakrabarty, A.M. Characterization of the *Pseudomonas aeruginosa* alginate (*alg*) gene region II. *J. Gen. Microbiol.* **1987**, *133*, 2303–2314.
6. Jain, S.; Ohman, D.E. Deletion of *algK* in mucoid *Pseudomonas aeruginosa* blocks alginate polymer formation and results in uronic acid secretion. *J. Bacteriol.* **1998**, *180*, 634–641.
7. Monday, S.R.; Schiller, N.L. Alginate synthesis in *Pseudomonas aeruginosa*: the role of AlgL (alginate lyase) and AlgX. *J. Bacteriol.* **1996**, *178*, 625–632.
8. Robles-Price, A.; Wong, T.Y.; Sletta, H.; Valla, S.; Schiller, N.L. AlgX is a periplasmic protein required

- for alginate biosynthesis in *Pseudomonas aeruginosa*. *J. Bacteriol.* **2004**, *186*, 7369–7377.
9. Ross, P.; Mayer, R.; Benziman, M. Cellulose biosynthesis and function in bacteria. *Microbiol. Rev.* **1991**, *55*, 35–58.
 10. Chu, L.; May, T.B.; Chakrabarty, A.M.; Misra, T.K. Nucleotide sequence and expression of the algE gene involved in alginate biosynthesis by *Pseudomonas aeruginosa*. *Gene* **1991**, *107*, 1–10.
 11. Rehm, B.H.; Boheim, G.; Tommassen, J.; Winkler, U.K. Overexpression of algE in *Escherichia coli*: sub-cellular localization, purification, and ion channel properties. *J. Bacteriol.* **1994**, *176*, 5639–5647.
 12. Rehm, B.H.; Grabert, E.; Hein, J.; Winkler, U.K. Antibody response of rabbits and cystic fibrosis patients to an alginate-specific outer membrane protein of a mucoid strain of *Pseudomonas aeruginosa*. *Microb. Pathog.* **1994**, *16*, 43–51.
 13. Rehm, B.H. The *Azotobacter vinelandii* gene algJ encodes an outer-membrane protein presumably involved in export of alginate. *Microbiology* **1996**, *142* (pt 4), 873–880.
 14. Nyvall, P.; Corre, E.; Boisset, C.; Barbeyron, T.; Rousvoal, S.; Scornet, D.; Kloareg, B.; Boyen, C. Characterization of mannuronan C-5-epimerase genes from the brown alga *Laminaria digitata*. *Plant Physiol.* **2003**, *133*, 726–735.
 15. Franklin, M.J.; Chitnis, C.E.; Gacesa, P.; Sonesson, A.; White, D.C.; Ohman, D.E. *Pseudomonas aeruginosa* AlgG is a polymer level alginate C5-mannuronan epimerase. *J. Bacteriol.* **1994**, *176*, 1821–1830.
 16. Franklin, M.J.; Ohman, D.E. Identification of algI and algJ in the *Pseudomonas aeruginosa* alginate biosynthetic gene cluster which are required for alginate O acetylation. *J. Bacteriol.* **1996**, *178*, 2186–2195.
 17. Schiller, N.L.; Monday, S.R.; Boyd, C.M.; Keen, N.T.; Ohman, D.E. Characterization of the *Pseudomonas aeruginosa* alginate lyase gene (algL): cloning, sequencing, and expression in *Escherichia coli*. *J. Bacteriol.* **1993**, *175*, 4780–4789.
 18. Nivens, D.E.; Ohman, D.E.; Williams, J.; Franklin, M.J. Role of alginate and its O acetylation in formation of *Pseudomonas aeruginosa* microcolonies and biofilms. *J. Bacteriol.* **2001**, *183*, 1047–1057.
 19. Bjerkan, T.M.; Lillehov, B.E.; Strand, W.I.; Skjak-Braek, G.; Valla, S.; Ertesvag, H. Construction and analyses of hybrid *Azotobacter vinelandii* mannuronan C-5 epimerases with new epimerization pattern characteristics. *Biochem. J.* **2004**, *381*, 813–821.
 20. Ertesvag, H.; Hoidal, H.K.; Schjerven, H.; Svanem, B.I.; Valla, S. Mannuronan C-5-epimerases and their application for in vitro and in vivo design of new alginates useful in biotechnology. *Metab. Eng.* **1999**, *1*, 262–269.
 21. Valla, S.; Li, J.; Ertesvag, H.; Barbeyron, T.; Lindahl, U. Hexuronyl C5-epimerases in alginate and glycosaminoglycan biosynthesis. *Biochimie* **2001**, *83*, 819–830.
 22. Campos, M.; Martinez-Salazar, J.M.; Lloret, L.; Moreno, S.; Nunez, C.; Espin, G.; Soberon-Chavez, G. Characterization of the gene coding for GDP-mannose dehydrogenase (algD) from *Azotobacter vinelandii*. *J. Bacteriol.* **1996**, *178*, 1793–1799.
 23. Hoidal, H.K.; Glaerum Svanem, B.I.; Gimmestad, M.; Valla, S. Mannuronan C-5 epimerases and cellular differentiation of *Azotobacter vinelandii*. *Environ. Microbiol.* **2000**, *2*, 27–38.
 24. Rehm, B.H.; Ertesvag, H.; Valla, S. A new *Azotobacter vinelandii* mannuronan C-5-epimerase gene (algG) is part of an alg gene cluster physically organized in a manner similar to that in *Pseudomonas aeruginosa*. *J. Bacteriol.* **1996**, *178*, 5884–5889.
 25. Gimmestad, M.; Sletta, H.; Ertesvag, H.; Bakkevig, K.; Jain, S.; Suh, S.J.; Skjak-Braek, G.; Ellingsen, T.E.; Ohman, D.E.; Valla, S. The *Pseudomonas fluorescens* AlgG protein, but not its mannuronan C-5-epimerase activity, is needed for alginate polymer formation. *J. Bacteriol.* **2003**, *185*, 3515–3523.
 26. Jain, S.; Franklin, M.J.; Ertesvag, H.; Valla, S.; Ohman, D.E. The dual roles of AlgG in C-5-epimerization and secretion of alginate polymers in *Pseudomonas aeruginosa*. *Mol. Microbiol.* **2003**, *47*, 1123–1133.
 27. May, T.B.; Shinabarger, D.; Maharaj, R.; Kato, J.; Chu, L.; DeVault, J.D.; Roychoudhury, S.; Zielinski, N.A.; Berry, A.; Rothmel, R.K.; Misra, K.T.; Chakrabarty, A.M. Alginate synthesis by *Pseudomonas aeruginosa*: a key pathogenic factor in chronic pulmonary infections of cystic fibrosis patients. *Clin. Microbiol. Rev.* **1991**, *4*, 191–206.
 28. Schmitt-Andrieu, L.; Hulen, C. Alginates of *Pseudomonas aeruginosa*: a complex regulation of the pathway of biosynthesis. *C.R. Acad Sci III* **1996**, *319*, 153–160.
 29. Schurr, M.J.; Yu, H.; Martinez-Salazar, J.M.; Boucher, J.C.; Deretic, V. Control of AlgU, a member of the sigma E-like family of stress sigma factors, by the negative regulators MucA and MucB and *Pseudomonas aeruginosa* conversion to mucoidy in cystic fibrosis. *J. Bacteriol.* **1996**, *178*, 4997–5004.
 30. Schurr, M.J.; Martin, D.W.; Mudd, M.H.; Hibler, N.S.; Boucher, J.C.; Deretic, V. The algD promoter: regulation of alginate production by *Pseudomonas aeruginosa* in cystic fibrosis. *Cell. Mol. Biol. Res.* **1993**, *39*, 371–376.
 31. Baynham, P.J.; Wozniak, D.J. Identification and characterization of AlgZ, an AlgT-dependent DNA-binding protein required for *Pseudomonas aeruginosa* algD transcription. *Mol. Microbiol.* **1996**, *22*, 97–108.
 32. Cochran, W.L.; Suh, S.J.; McFeters, G.A.; Stewart, P.S. Role of RpoS and AlgT in *Pseudomonas aeruginosa* biofilm resistance to hydrogen peroxide and monochloramine. *J. Appl. Microbiol.* **2000**, *88*, 546–553.
 33. Firoved, A.M.; Boucher, J.C.; Deretic, V. Global genomic analysis of AlgU (sigma(E))-dependent promoters (sigmulon) in *Pseudomonas aeruginosa* and implications for inflammatory processes in cystic fibrosis. *J. Bacteriol.* **2002**, *184*, 1057–1064.
 34. Firoved, A.M.; Deretic, V. Microarray analysis of global gene expression in mucoid *Pseudomonas aeruginosa*. *J. Bacteriol.* **2003**, *185*, 1071–1081.
 35. Keith, L.M.; Bender, C.L. AlgT (sigma22) controls alginate production and tolerance to environmental stress in *Pseudomonas syringae*. *J. Bacteriol.* **1999**, *181*, 7176–7184.

36. Mathee, K.; McPherson, C.J.; Ohman, D.E. Post-translational control of the algT (algU)-encoded sigma22 for expression of the alginate regulon in *Pseudomonas aeruginosa* and localization of its antagonist proteins MucA and MucB (AlgN). *J. Bacteriol.* **1997**, *179*, 3711–3720.
37. Wu, W.; Badrane, H.; Arora, S.; Baker, H.V.; Jin, S. MucA-mediated coordination of type III secretion and alginate synthesis in *Pseudomonas aeruginosa*. *J. Bacteriol.* **2004**, *186*, 7575–7585.
38. Yorgey, P.; Rahme, L.G.; Tan, M.W.; Ausubel, F.M. The roles of mucD and alginate in the virulence of *Pseudomonas aeruginosa* in plants, nematodes and mice. *Mol. Microbiol.* **2001**, *41*, 1063–1076.
39. Boucher, J.C.; Schurr, M.J.; Yu, H.; Rowen, D.W.; Deretic, V. *Pseudomonas aeruginosa* in cystic fibrosis: role of mucC in the regulation of alginate production and stress sensitivity. *Microbiology* **1997**, *143* (pt 11), 3473–3480.
40. Deretic, V.; Dikshit, R.; Konyecsni, W.M.; Chakrabarty, A.M.; Misra, T.K. The algR gene, which regulates mucoidy in *Pseudomonas aeruginosa*, belongs to a class of environmentally responsive genes. *J. Bacteriol.* **1989**, *171*, 1278–1283.
41. Deretic, V.; Konyecsni, W.M. Control of mucoidy in *Pseudomonas aeruginosa*: transcriptional regulation of algR and identification of the second regulatory gene, algQ. *J. Bacteriol.* **1989**, *171*, 3680–3688.
42. Mohr, C.D.; Hibler, N.S.; Deretic, V. AlgR, a response regulator controlling mucoidy in *Pseudomonas aeruginosa*, binds to the FUS sites of the algD promoter located unusually far upstream from the mRNA start site. *J. Bacteriol.* **1991**, *173*, 5136–5143.
43. Zielinski, N.A.; Chakrabarty, A.M.; Berry, A. Characterization and regulation of the *Pseudomonas aeruginosa* algC gene encoding phosphomannomutase. *J. Biol. Chem.* **1991**, *266*, 9754–9763.
44. Carterson, A.J.; Morici, L.A.; Jackson, D.W.; Frisk, A.; Lizewski, S.E.; Jupiter, R.; Simpson, K.; Kunz, D.A.; Davis, S.H.; Schurr, J.R.; Hassett, D.J.; Schurr, M.J. The transcriptional regulator AlgR controls cyanide production in *Pseudomonas aeruginosa*. *J. Bacteriol.* **2004**, *186*, 6837–6844.
45. Lizewski, S.E.; Lundberg, D.S.; Schurr, M.J. The transcriptional regulator AlgR is essential for *Pseudomonas aeruginosa* pathogenesis. *Infect. Immun.* **2002**, *70*, 6083–6093.
46. Lizewski, S.E.; Schurr, J.R.; Jackson, D.W.; Frisk, A.; Carterson, A.J.; Schurr, M.J. Identification of AlgR-regulated genes in *Pseudomonas aeruginosa* by use of microarray analysis. *J. Bacteriol.* **2004**, *186*, 5672–5684.
47. Penalzoza-Vazquez, A.; Fakhr, M.K.; Bailey, A.M.; Bender, C.L. AlgR functions in algC expression and virulence in *Pseudomonas syringae* pv. *syringae*. *Microbiology* **2004**, *150*, 2727–2737.
48. Chitnis, C.E.; Ohman, D.E. Genetic analysis of the alginate biosynthetic gene cluster of *Pseudomonas aeruginosa* shows evidence of an operonic structure. *Mol. Microbiol.* **1993**, *8*, 583–593.
49. Lloret, L.; Barreto, R.; Leon, R.; Moreno, S.; Martinez-Salazar, J.; Espin, G.; Soberon-Chavez, G. Genetic analysis of the transcriptional arrangement of *Azotobacter vinelandii* alginate biosynthetic genes: identification of two independent promoters. *Mol. Microbiol.* **1996**, *21*, 449–457.
50. Vazquez, A.; Moreno, S.; Guzman, J.; Alvarado, A.; Espin, G. Transcriptional organization of the *Azotobacter vinelandii* algXLVIFA genes: characterization of algF mutants. *Gene* **1999**, *232*, 217–222.
51. Thomas, S. Alginate dressings in surgery and wound management—part 1. *J. Wound Care* **2000**, *9*, 56–60.
52. Malmstrom, A.; Fransson, L.A. Biosynthesis of dermatan sulfate. I. Formation of L-iduronic acid residues. *J. Biol. Chem.* **1975**, *250*, 3419–3425.
53. Ertesvag, H.; Hoidal, H.K.; Hals, I.K.; Rian, A.; Doseth, B.; Valla, S. A family of modular type mannuronan C-5-epimerase genes controls alginate structure in *Azotobacter vinelandii*. *Mol. Microbiol.* **1995**, *16*, 719–731.
54. Svanem, B.I.; Skjak-Braek, G.; Ertesvag, H.; Valla, S. Cloning and expression of three new *Azotobacter vinelandii* genes closely related to a previously described gene family encoding mannuronan C-5-epimerases. *J. Bacteriol.* **1999**, *181*, 68–77.
55. Ertesvag, H.; Doseth, B.; Larsen, B.; Skjak-Braek, G.; Valla, S. Cloning and expression of an *Azotobacter vinelandii* mannuronan C-5-epimerase gene. *J. Bacteriol.* **1994**, *176*, 2846–2853.
56. Ertesvag, H.; Hoidal, H.K.; Skjak-Braek, G.; Valla, S. The *Azotobacter vinelandii* mannuronan C-5-epimerase AlgE1 consists of two separate catalytic domains. *J. Biol. Chem.* **1998**, *273*, 30927–30932.
57. Svanem, B.I.; Strand, W.I.; Ertesvag, H.; Skjak-Braek, G.; Hartmann, M.; Barbeyron, T.; Valla, S. The catalytic activities of the bifunctional *Azotobacter vinelandii* mannuronan C-5-epimerase and alginate lyase AlgE7 probably originate from the same active site in the enzyme. *J. Biol. Chem.* **2001**, *276*, 31542–31550.
58. Campa, C.; Holtan, S.; Nilsen, N.; Bjerkan, T.M.; Stokke, B.T.; Skjak-Braek, G. Biochemical analysis of the processive mechanism for epimerization of alginate by mannuronan C-5 epimerase AlgE4. *Biochem. J.* **2004**, *381*, 155–164.
59. Donati, I.; Draget, K.I.; Borgogna, M.; Paoletti, S.; Skjak-Braek, G. Tailor-made alginate bearing galactose moieties on mannuronic residues: selective modification achieved by a chemoenzymatic strategy. *Biomacromolecules* **2005**, *6*, 88–98.
60. Donati, I.; Holtan, S.; Morch, Y.A.; Borgogna, M.; Dentini, M.; Skjak-Braek, G. New hypothesis on the role of alternating sequences in calcium-alginate gels. *Biomacromolecules* **2005**, *6*, 1031–1040.
61. King, A.; Strand, B.; Rokstad, A.M.; Kulseng, B.; Andersson, A.; Skjak-Braek, G.; Sandler, S. Improvement of the biocompatibility of alginate/poly-L-lysine/alginate microcapsules by the use of epimerized alginate as a coating. *J. Biomed. Mater. Res. A* **2003**, *64*, 533–539.
62. Pandey, R.; Khuller, G.K. Chemotherapeutic potential of alginate-chitosan microspheres as anti-tubercular drug carriers. *J. Antimicrob. Chemother.* **2004**, *53*, 635–640.
63. Kulseng, B.; Skjak-Braek, G.; Ryan, L.; Andersson, A.; King, A.; Faxvaag, A.; Espevik, T. Transplantation of alginate microcapsules: generation of antibodies against

- alginates and encapsulated porcine islet-like cell clusters. *Transplantation* **1999**, *67*, 978–984.
64. Rokstad, A.M.; Kulseng, B.; Strand, B.L.; Skjak-Braek, G.; Espevik, T. Transplantation of alginate microcapsules with proliferating cells in mice: capsular overgrowth and survival of encapsulated cells of mice and human origin. *Ann. N.Y. Acad. Sci.* **2001**, *944*, 216–225.
 65. Rokstad, A.M.; Strand, B.; Rian, K.; Steinkjer, B.; Kulseng, B.; Skjak-Braek, G.; Espevik, T. Evaluation of different types of alginate microcapsules as bioreactors for producing endostatin. *Cell Transplant.* **2003**, *12*, 351–364.
 66. Strand, B.L.; Morch, Y.A.; Syvertsen, K.R.; Espevik, T.; Skjak-Braek, G. Microcapsules made by enzymatically tailored alginate. *J. Biomed. Mater. Res. A* **2003**, *64*, 540–550.
 67. Schneider, S.; Feilen, P.J.; Brunnenmeier, F.; Minnemann, T.; Zimmermann, H.; Zimmermann, U.; Weber, M.M. Long-term graft function of adult rat and human islets encapsulated in novel alginate-based microcapsules after transplantation in immunocompetent diabetic mice. *Diabetes* **2005**, *54*, 687–693.
 68. Soon-Shiong, P.; Heintz, R.E.; Merideth, N.; Yao, Q.X.; Yao, Z.; Zheng, T.; Murphy, M.; Moloney, M.K.; Schmehl, M.; Harris M., et al. Insulin independence in a type 1 diabetic patient after encapsulated islet transplantation. *Lancet* **1994**, *343*, 950–951.

Biphasic Calcium Phosphate (BCP) Bioceramics: Chemical, Physical, and Biological Properties

B

Guy Daculsi

Nantes University, and INSERM EM 9903, INSERM, Nantes, France, France

Racquel LeGeros

Department of Biomaterials and Biomimetics, New York University College of Dentistry, New York, New York, U.S.A.

INTRODUCTION

Albee, in 1920, reported the first successful application of a calcium phosphate (CaP) reagent for the repair of bone defect in human.^[1] More than 50 yr later, Nery et al.^[2] reported successful clinical use of a porous “tricalcium phosphate (TCP)” in surgically created periodontal defects in dogs and Denissen^[3] reported the use of dense hydroxyapatite (HA) as immediate tooth root replacements. In the early 1980s, synthetic HA and β -TCP became commercially available as bone-substitute materials for dental and medical applications largely through the efforts of Jarcho, de Groot, and Aoki.^[4–8]

The term “biphasic calcium phosphate (BCP)” was first used by Nery et al.^[9] to describe the bioceramic that consisted of a mixture of HA and β -TCP, after the x-ray diffraction analysis by LeGeros^[8] in 1986 showed that the porous TCP preparation material used in their early publications^[2] was, in fact, a mixture of 20% HA and 80% β -TCP. The first studies on BCP with varying HA/ β -TCP reported by LeGeros, Nery, Daculsi et al.^[10–13] demonstrated that the bioactivity of these ceramics may be controlled by manipulating the HA/ β -TCP ratios. Subsequently, focussed and intensive studies on macroporous BCP (MBCP) by Daculsi and coworkers^[14–18] led to the development and commercialization of MBCP as bone-substitute biomaterial for orthopedic and dental applications.^[19–26] BCP bioceramics consist of a mixture of HA, $\text{Ca}_{10}(\text{PO}_4)_6(\text{OH})_2$ and β -TCP, $\text{Ca}_3(\text{PO}_4)_2$ of varying HA/ β -TCP ratio. The BCP ceramics concept is based on the combination of the more stable phase of HA and more soluble phase, β -TCP. The more soluble β -TCP gradually dissolves in vivo, seeding new bone formation as it releases calcium and phosphate ions into the biological medium.^[10–12,26]

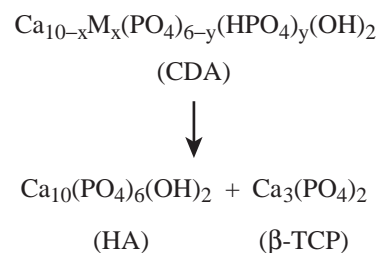
At present, BCP is commercially available in Europe, U.S.A., Brazil, Japan, Korea, Taiwan, and China as a bone-graft or bone-substitute materials for orthopedic and dental applications under various trade marks (e.g., BCP[®], MBCP[®], Triosite[®], Zimmer,

Europe, Hatric[®], Tribone 80[®], TricOs[®], Eurocer[®], and Biceram[®]). Macroporous BCP is available in blocks, particulates, and customized design (Fig. 1) and as an injectable material in a polymer carrier (MBCP Gel[®]). The main attractive feature of bioactive bone-graft materials such as CaP (HA, β -TCP, and BCP), ceramics, and bioactive glasses is their ability to form a strong direct bond with the host bone resulting in a strong interface compared to bioinert or biotolerant materials that form a fibrous interface.^[3–5,15,27–31] The formation of this dynamic interface is believed to result from a sequence of events involving interaction with cells and formation of carbonate HA (CHA) (similar to bone mineral) by dissolution/precipitation processes.^[10–12,26,32–34] This review briefly describes preparation of BCP bioceramics, their in vitro properties, cellular events after implantation, and subsequent biodegradation and biodissolution including the significance of the biological apatite precipitation.

PREPARATION OF BCP BIOCERAMICS

Calcium-Deficient Apatites

The BCP or mixtures of HA and β -TCP are obtained when biologic or synthetic calcium-deficient apatites (CDAs) are sintered above 700°C^[26,28,35] according to the following reaction



The apatite is considered calcium deficient when the Ca/P ratio is lower than the stoichiometric value



Fig. 1 MBCP[®] block, granules, cylinders, wedges, and customized design available for bone reconstruction.

of 1.67 for pure calcium HA, $\text{Ca}_{10}(\text{PO}_4)_6(\text{OH})_2$. The CDAs may be represented by the formula, $\text{Ca}_{10-x}\text{M}_x(\text{PO}_4)_{6-y}(\text{HPO}_4)_y(\text{OH})_2$, where M represents other ions (e.g., sodium and magnesium) that can substitute for calcium (Ca) ions.

The Ca deficiency of the apatites depends on the synthesis conditions (precipitation or hydrolysis methods), reaction pH, and/or temperature.^[26,28,35,36] The CDAs may be obtained by precipitation method under different conditions of pH and temperature; the lower the pH, the higher the temperature required for the precipitation of apatite.^[26,28] For example, CDA is obtained by precipitation at 80°C to 100°C even at low pH (pH 4–6). At lower temperatures and low pH, nonapatitic CaPs, e.g., dicalcium phosphate dihydrate (DCPD), $\text{CaHPO}_4 \cdot 2\text{H}_2\text{O}$ or octacalcium phosphate (OCP), $\text{Ca}_8\text{H}_2(\text{PO}_4)_6 \cdot 5\text{H}_2\text{O}$, is obtained.^[28]

Another method of CDA preparation is by hydrolysis of nonapatitic CaPs including amorphous CaP, $\text{Ca}_x(\text{PO}_4, \text{HPO}_4)_y$; DCPD; dicalcium phosphate anhydrous, CaHPO_4 ; OCP; or β -TCP, $\text{Ca}_3(\text{PO}_4)_2$.^[26,28,35,36] In the hydrolysis of DCPD in NaOH solutions, the Ca deficiency of the unsintered apatite and, subsequently, the HA/ β -TCP ratio of the BCP obtained after sintering can be controlled by two variables: the concentration of the NaOH solution and the ratio of the weight of the DCPD to the volume of the NaOH solution.^[36]

Variation in HA/ β -TCP Ratio in BCP

The BCP composition (HA/ β -TCP ratio) obtained after sintering depends on the Ca deficiency of the unsintered biologic or synthetic apatite and on the sintering temperature.^[26,28]

Sintering commercial CaP reagents (labelled as “HA” or “calcium phosphate, tribasic” or “TCP” (e.g., Baker or Fisher or Mallinckrodt, U.S.A.; or Merck, Darmstadt) above 900°C was shown to result in pure HA, pure β -TCP, or BCP.^[28,37] The BCP

ceramic may also be prepared by mechanically mixing two types of synthetic apatites or commercial CaP reagents.^[26,38]

Introduction of Macroporosity and Microporosity

Two physical properties of bioceramics are believed to be very important for optimum biological performance, including bioceramic–cell interaction, bioceramic resorption, bioceramic–tissue interface, and new bone formation. These fundamental physical properties are interconnecting macroporosity and appropriate microporosity.

Macroporosity in the BCP ceramic is introduced by incorporating volatile materials (e.g., naphthalene, hydrogen peroxide, or other porogens), heating at temperature below 200°C, and subsequent sintering at higher temperatures.^[39–41] Macroporosity results from the release of the volatile materials (Fig. 2).

Microporosity is a consequence of the temperature and duration of sintering.^[26] The higher the temperature, the lower the microporosity content and the lower the specific surface area (Figs. 3A and B).

Presently, commercial BCP products of different or similar HA/ β -TCP ratios are manufactured in many parts of the world and their successful use in medicine and dentistry has been reported.^[14,16,21–24] The total porosity (macroporosity plus microporosity) of these products is reported to be about 70% of the bioceramic volume. The current BCP products used by the surgeons (with 60/40–75/25 HA/TCP ratio) present similar percent macroporosity (50–60%), but the additional microporosity was very different among them, varying from 3% to 25%. Low percent microporosity and low surface area can result in lower

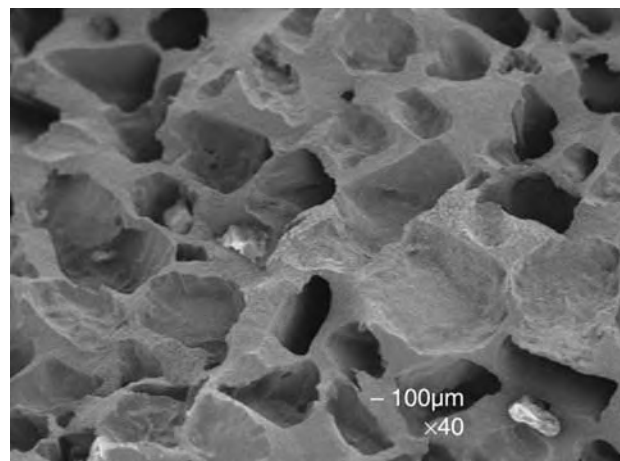


Fig. 2 Scanning electron microscopy blocs showing macroporous structure.

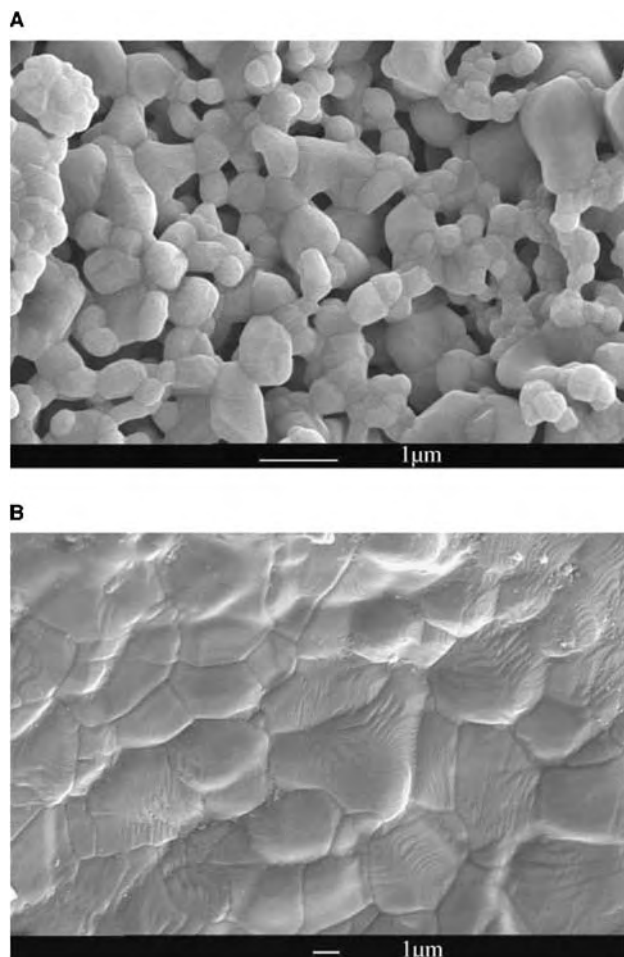


Fig. 3 The MBCP sintered at (A) 1050°C and (B) at 1200°C.

bioactivity and lower dissolution property. A microporosity of at least 20% with specific surface area higher than $2 \text{ m}^2/\text{g}$ is suggested for optimal BCP efficacy.

The ideal pore size for a bioceramic approximates that of bone. It has been demonstrated that microporosity (diameter $< 10 \mu\text{m}$) allows body fluid circulation whereas macroporosity (diameter $> 100 \mu\text{m}$) provides a scaffold for bone-cell colonization. It was reported that BCP ceramic with an average pore size diameter of $565 \mu\text{m}$ (compared to those with average pore size diameter of $300 \mu\text{m}$) and 40% macroporosity (compared to 50% macroporosity) had greater bone ingrowth.^[17]

IN VITRO PROPERTIES OF BCP CERAMICS

Crystal Properties

Hydroxyapatite crystals are not distinguishable from the β -TCP crystals in the BCP sintered above 900°C

using scanning electron microscopy (SEM). However, using transmission electron microscopy (TEM), electron diffraction is able to distinguish between HA and β -TCP crystals in the BCP.^[42] The crystallinity (reflecting crystal size and perfection) of the HA and β -TCP phases in BCP depends on the sintering temperature of CDA: The higher the sintering temperature, the higher the crystallinity.^[26]

Mechanical Properties

It is expected that the pore size and percent macroporosity of the BCP ceramic will affect the mechanical properties.^[17,38] The preparation method also was found to have a significant influence on the compressive strength. The BCP ceramic prepared from a single CDA phase was reported to exhibit higher compressive strength compared to BCP ceramic prepared by mixing two unsintered CaP preparations, one which, after sintering at 1200°C , resulted in only HA or only β -TCP.^[43] The initial mechanical property is not the best criterion for assessing the efficacy of bone ingrowth. For example, BCP with high mechanical property owing to low microporosity (resulting from high sintering temperature) may exhibit reduced bioresorption and bioactivity. On the contrary, it was demonstrated that the initial mechanical property of BCP increased two or three times in a few weeks after implantation owing to physicochemical events of dissolution and biological precipitation into the micropores.^[17]

Dissolution Properties

The extent of dissolution in acidic buffer in vitro is much higher for the β -TCP ceramic compared to that for the HA ceramic.^[26,28,44] Thus, the extent of dissolution of BCP ceramic of comparable macroporosity and particle size will depend on the HA/ β -TCP ratio: The higher the ratio, the lower the extent of dissolution.^[10,27,44] The dissolution properties are also affected by the methods of obtaining BCP: whether from a single calcium-deficient apatite phase (BCP1) or from a mechanical mixture of two unsintered CaP preparations (BCP2): BCP2 exhibited higher extent of dissolution compared to BCP1.^[38] In some cases, BCP ceramic with similar HA/ β -TCP ratios could present different dissolution rates.^[26] This phenomenon may be caused by processing variables (sintering time and temperature) that could affect the total macroporosity and microporosity: The greater the macroporosity and microporosity, the greater the extent of dissolution. In vivo, dissolution of BCP ceramics is manifested by a decrease in crystal size and increase in macroporosity and microporosity.^[10,12,15]

B

In Vitro Cell–BCP Ceramic Interaction

BCP ceramics, such as HA or β -TCP ceramics, present hospitable surfaces for spreading of monocytes^[45,46] or osteoclasts.^[29,47] In vitro studies using commercial BCP ceramic with an HA/ β -TCP ratio of 60/40 (Triosite) showed less resorption pits on the BCP discs compared to those in dentin.^[46] Yamada et al.^[47] reported that osteoclastic resorption was highest for the BCP ceramic with HA/ β -TCP ratio of 25/75 compared to those with higher ratios. This observation seemed to indicate that osteoclastic resorption is a function of the solubility of the BCP, which in turn, is a function of the HA/ β -TCP ratio: The lower the ratio, the higher the solubility. However, the same study also demonstrated that the BCP (HA/ β -TCP ratio, 25/75) had a higher osteoclastic resorption than pure β -TCP in spite of the higher solubility of the latter. This may be related to the observation that the amount of calcium and phosphate ions released in the environment is critical to cellular activities.^[48]

CELLULAR EVENTS AND TISSUE RESPONSE AFTER IMPLANTATION OF BCP

The BCP materials elicit responses from bone cells and related cells in vitro and in vivo, which are similar to those elicited by bone.^[49] These materials allow cell attachment, proliferation, and phenotypic expression. The first biological events after BCP ceramics implantation are biological fluid diffusion, followed by cell colonization into the macropores of the implanted BCP (Fig. 4).



Fig. 4 Newly formed bone into the macropores of MBCP implanted in rabbit femoral epiphysis after 14 days of implantation showing osteoclasts and osteoblasts. Decalcified section stained with Masson's trichrome staining.

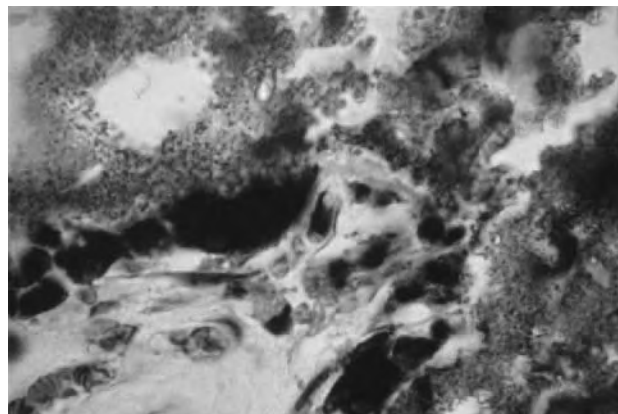


Fig. 5 Tartrate-resistant acid phosphatase staining of osteoclast associated with MBCP implanted in rabbit femoral epiphysis after 14 days of implantation.

These cells are macrophages in early steps, followed by mesenchymal stem cells, osteoblasts, and osteoclasts. The resorbing cells forming both on the surface of the newly formed bone and on the bioceramic surface look like osteoclast and are tartrate-resistant acid phosphatase (TRAP)-positive^[25] as shown in Fig. 5.

In human spine arthrodesis, we observed that after two months, bone remodelling occurs with secondary osteoclast resorption of the BCP and bone ingrowth at the expense of the implanted BCP (Fig. 6).

Generally, when granules are used in osteoarticular surgery, some grains are released in the cartilage or nonosseous sites. No foreign body reaction or material rejection was observed. Resorption or tissue incorporation was demonstrated. For example, in human spine arthrodesis after 3.5 months of implantation, granules

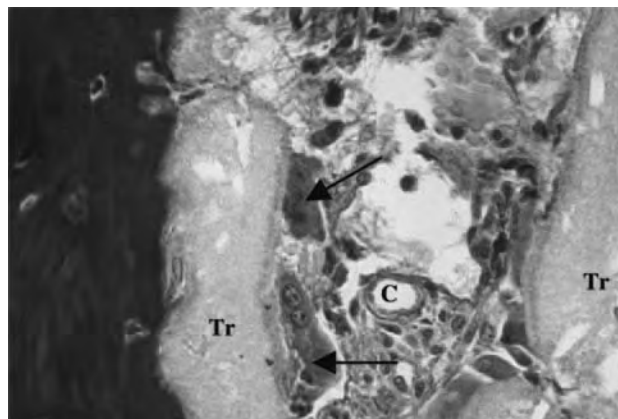


Fig. 6 Human spine arthrodesis using Triosite blocks after 3.5 months of implantation showing bone ingrowth at the expense of the Triosite (Tr) with osteoclast (arrow) near vascular channel (V).

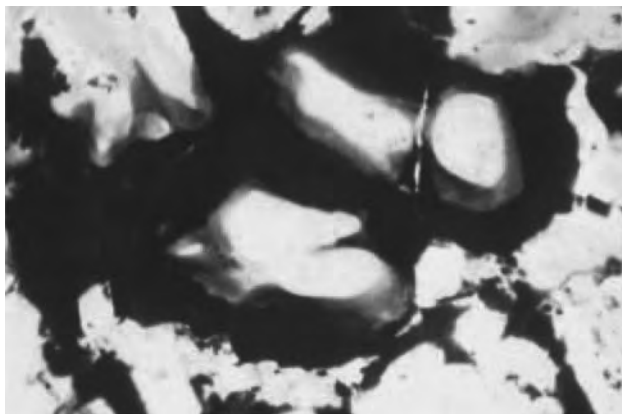


Fig. 7 MBCP[®] granules implanted for two weeks in muscular area of rabbit. Undecalcified section with Movat staining.

of Tr (BCP with a 60/40 HA/ β -TCP ratio) appeared surrounded by newly formed cartilage without fibrous encapsulation.^[30] Moreover, after implantation in nonosseous site (in subcutaneous area), osteoid formation into some MBCP macropores was observed (Fig. 7).^[24]

These observations suggest that BCP with appropriate interconnecting macroporosity present suitable chemical environment associated with efficient architecture that allows the capture of mesenchymal STEM cells and induce their phenotype to osteogenic cell lines. These observations have been also described by other groups.^[50–53] This property can be used for artificial bone in irradiated implantation site. Irradiation produces irreversible effects on normal tissues, involving damages on their repair properties. Nevertheless quality of life of patients who undergo radiotherapy could be improved by bone reconstructions. A preclinical study performed in irradiated dogs demonstrated bone ingrowth at the expense of structured implants of micro MBCP filled with autologous bone marrow after implantation in irradiated soft and bone tissue,^[24] shown in Fig. 8.

We have conducted a comparative study^[54] on cell interaction with three biomaterials: 0.5 to 1 mm granules of unsintered bovine bone mineral (BioOss[™], Geistlich, Switzerland, 60% porosity) sintered bovine bone mineral (“BonAp,” 60% porosity, prepared by LeGeros, New York University), and synthetic BCP (MBCP[™], 70% total porosity and 60/40 HA/ β -TCP ratio) implanted in epiphysis of rat femur. The three biomaterials allowed cell attachment, proliferation, and osteogenic expression. Light microscopy and SEM examinations showed bone ingrowth with osteoblasts, and osteoid or bony formation between and on the surfaces of the residual granules. However, the amount of resorption was very limited for the bovine bone-derived biomaterials, sintered and unsintered, in spite

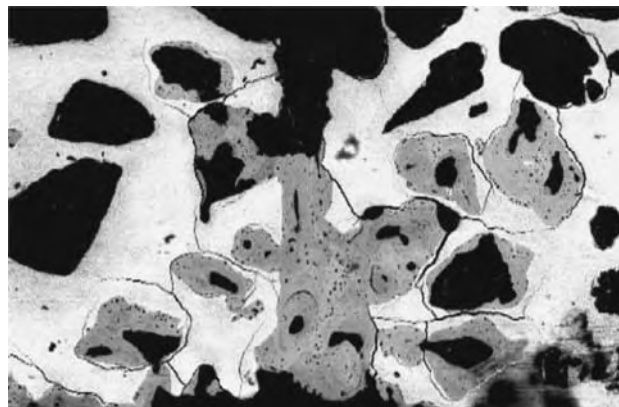


Fig. 8 Bone reconstruction into MBCP implant mixed with autologous bone marrow and implanted in irradiated (65 grays) femoral canine bone defects.

of an evidence of higher amount of macrophagous cells associated with BioOss. On the contrary, the biodegradation of the MBCP[™] was associated with bone ingrowth with numerous bone trabeculae between the granules and on the surface of the residual materials. Microscanner revealed bone architecture of physiological bone ingrowth on MBCP samples. However, it was more difficult to distinguish the host bone from the bone implant (BioOss). These results suggested that all the bone-substitute material tested had osseointegrative properties; however, the kinetics of bone ingrowth was higher for synthetic micro and macroporous MBCP. The sintered bovine bone (“BonAP”) showed low amount of bone formation perhaps owing to the larger apatite crystal size caused by high sintering temperature. For the unsintered bovine bone (BioOss), resorption was limited and bone ingrowth appeared delayed compared to synthetic MBCP, in spite of the microcrystalline nature of BioOss.

BIODEGRADATION, BIODISSOLUTION, AND SIGNIFICANCE OF BIOLOGICAL APATITE PRECIPITATION

The processes of biodegradation and bioresorption of bone-substitute biomaterials are the main factors for promoting bioactivity and for providing the space for bone ingrowth.

The biodegradation of BCP include the dissolution of the individual HA or β -TCP crystals.^[41] The HA/ β -TCP ratio in the BCP appeared greater after implantation,^[42] confirming the greater solubility of the β -TCP compared to the HA phase in the BCP.^[44]

Formation of microcrystals with Ca/P ratios similar to those of bone apatite crystals was also observed after implantation of the MBCP. The abundance of

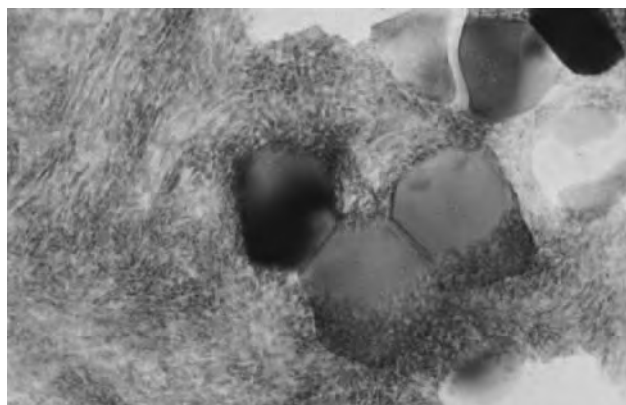


Fig. 9 Biological apatite precipitation at the surface of residual crystals in BCP, observed in high-resolution TEM.

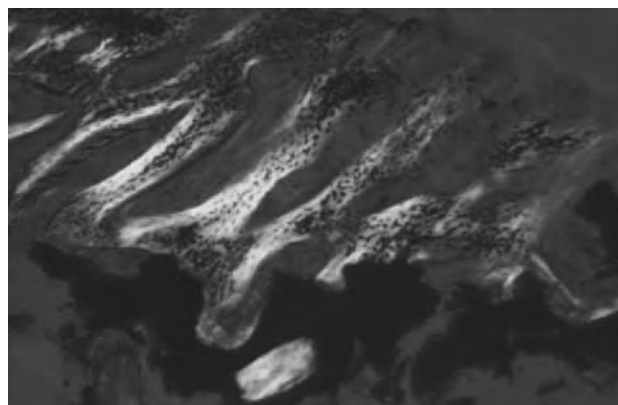


Fig. 10 Spongy bone formed at the expense of Tr blocks in human spine arthrodesis.

these crystals was directly related to the initial β -TCP/HA ratio in the BCP: The higher the ratio the greater the abundance of the microcrystals associated with the BCP crystals. According to these data it is possible to control the kinetics of dissolution and precipitation, and subsequently the bioactivity.^[10–12] Using high-resolution TEM, we have demonstrated for the first time that the formation of these bone apatite-like microcrystals (Fig. 9) after implantation of CaPs (HA, BCP) were nonspecific, i.e., not related to implantation site (osseous or nonosseous sites), subjects of implantation, and types of CaP ceramics.^[10–12,32–34]

The coalescing interfacial zone of biological apatite and residual BCP ceramic crystals (mostly HA), provide a scaffold for further bone-cell adhesion and further bone ingrowth. The bone repair or bone regeneration process involves dissolution of CaP crystals and then a precipitation of CHA needle-like crystallites in micropores close to the dissolving crystals (Fig. 9). The coalescing zone constitutes the new biomaterial–bone interface, which includes the participation of proteins and CHA crystals originating from the BCP ceramic crystals, but does not include the biomaterial surface. The ensuing events of bone ingrowth and formation of new bone progressively replace the initially formed CHA from the CaP biomaterials.^[11,12,34]

The processes of cell colonization, adhesion, phagocytosis and osteoclastic resorption, extracellular matrix elaboration and mineralization, bone ingrowth, and bone remodelling associated with the biological apatite precipitation during CaP ceramics dissolution are continuously in progress. Consequently, the interface is not static but dynamic and in constant evolution, taking into account bone physiopathology, biomechanical factors, and bone maturation. The processes involved a well-organized and mineralized bone ingrowth at the expense of the bone-substitute biomaterial (Fig. 10).

BIOACTIVE, OSTEOCONDUCTIVE, AND OSTEOINDUCTIVE PROPERTIES

The bioactive concept of BCP has been applied to a new composite consisting of water-soluble polymer and MBCP granules in an injectable form.^[55–57] We have elaborated on such an injectable bone substitute (IBS) (MBCP Gel™ Biomatlante, France) that is ready to use, biocompatible, able to conform to the shape of any bone defect, and bioresorbable. The IBS (MBCP Gel) cannot have mechanical properties like the self-setting hydraulic CaP bone cements.^[58,59] However, bone cells are able to invade the spaces released by the disappearance of the polymer and bone ingrowth take place at the expense of the resorption of the BCP grains (Fig. 11).

In time, mechanical property increases owing to increasing new bone formation.

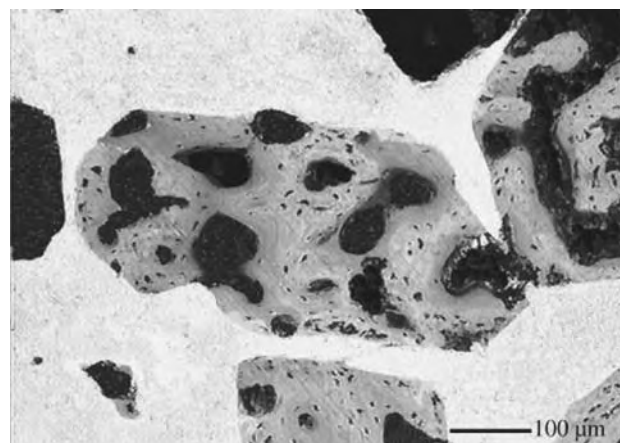


Fig. 11 The SEM using backscattered electron imaging of polished section of IBS after three weeks of implantation in rabbit showing bone growth closely associated with the residual grain of CaP.

Osteoconductive and Osteoinductive Properties

The CaP biomaterials, including BCP, have been described as osteoconductive but not osteoinductive.^[60] Osteoconductive materials serve as scaffold or template, allowing “vascular ingress, cellular infiltration and attachment, cartilage formation, and calcified tissue deposition.” Osteoinductive materials [(e.g., bone morphogenetic proteins (BMPs))] “stimulate uncommitted cells (e.g., mesenchymal stem cells) to convert phenotypically to chondroprogenitor and osteoprogenitor cells.”^[61] Ripamonti, Crooks, and Kirbride.^[51–53] and Kuboki, Takita, and Kobayashi^[50] demonstrated that some CaP materials with appropriate geometry can have osteoinductive properties. Osteoinductive properties have been reported for other CaP biomaterials including BCP.^[62–68]

Reddi^[61] attributed the phenomenon of osteoinductive properties associated with specific (but undefined) biomaterial geometry to the binding of an optimum amount of endogenous BMPs in circulation to the CaP biomaterial allowing the biomaterial to be osteoinductive. Ripamonti, Crooks, and Kirbride^[53] and Kuboki, Takita, and Kobayashi^[50] have postulated that the geometry of the material is a critical parameter in bone induction. Kuboki, Takita, and Kobayashi^[50] concluded that the geometry of the BMP carrier is one of the important factors controlling the efficacy of the phenotype induction. Daculsi et al. have shown that micropores combined with the high solubility of β -TCP and stability of HA induce the precipitation of biological apatite crystals on to BCP ceramics.^[10–12,25,66] The osteoprogenitor cells might in turn recognize the bone-like apatite layer formed by dissolution–reprecipitation on the material^[10–12,34] and produce bone. Others have thought that low oxygen tension in the central region of the implants might provoke a dedifferentiation of pericytes from blood microvessels into osteoblasts.^[68,69]

CONCLUSIONS

The BCP bioceramic is an intimate mixture of HA and β -TCP, obtained by sintering CDA. The BCP has the advantage of controlled bioactivity by manipulation of the HA/ β -TCP ratio owing to the preferential dissolution of the more soluble β -TCP component. The BCPs like other bioactive materials have the ability to form carbonateapatite on the surface by the process of dissolution and precipitation. Biphasic calcium phosphates, like other CaP biomaterials, are osteoconductive but can be made osteoinductive by appropriate geometry, specifically, macroporosity and microporosity. Besides its current application in orthopedics and dentistry, BCP can also be used as scaffolds for

tissue engineering using stem cells, delivery system for drugs, antibiotics, and hormones, and carriers for growth factors.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the collaboration of other researchers in the studies performed at Nantes University, Nantes Hospital, and Nantes Microscopy SC3M department, and at New York University College of Dentistry.

The collaborative studies cited in this review were supported by research grants from the INSERM U225, CJF 93-05 and E99-03, and CNRS EP59 (Dr. G. Daculsi, Director) and from the National Institute for Dental Research of the National Institutes of Health Nos. DE04123 and DE07223 (Dr. R. Z. LeGeros, Principal Investigator).

Parts of this review were reprinted from a previous publication^[25,26] with the permission of the publisher (Journal of Materials Science: Materials in Medicine).

ARTICLES OF FURTHER INTEREST

Bioactive Glass; Biologic and Synthetic Apatites; Zirconia Ceramics

REFERENCES

1. Albee, F.H. *Ann. Surg.* **1920**, *71*, 32–36.
2. Nery, E.B.; Lynch, K.L.; Hirthe, W.M.; Mueller, K.H. *J. Periodontol.* **1975**, *46*, 328–339.
3. Denissen, H.W. Ph.D. Thesis. Vrije Universiteit: Amsterdam, 1979.
4. Jarcho, M. *Clin. Orthop.* **1981**, *157*, 259–278.
5. De Groot, K. *Bioceramics of Calcium Phosphates*; CRC Press: Boca Raton, 1983; 100–114.
6. Metsger, S.D.; Driskell, T.D.; Paulsrud, J.R. *J. Am. Dent. Assoc.* **1982**, *105*, 1035.
7. Aoki, H.; Kato, K. *Jpn. Ceram. Soc.* **1975**, *10*, 469.
8. LeGeros, R.Z. *Adv. Dent. Res.* **1988**, *2*, 164–183.
9. Ellinger, R.F.; Nery, E.B.; Lynch, K.L. *J. Periodont. Restor. Dent.* **1986**, *3*, 223.
10. LeGeros, R.Z.; Nery, E.; Daculsi, G.; Lynch, K.B. *Third World Biomaterials Congress*; Abstract no. 35; 1988.
11. LeGeros, R.Z.; Daculsi, G. *Handbook of Bioactive Ceramics*; Yamamuro, N., Hench, L., Wilson, J., Eds.; Boca Raton, 1990; Vol. 2, 1728 pp.
12. Daculsi, G.; LeGeros, R.Z.; Nery, E.; Lynch, K.; Kerebel, B. *J. Biomed. Mat. Res.* **1989**, *23*, 883–894.
13. Nery, E.B.; LeGeros, R.Z.; Lynch, K.L.; Kalbfleisch, J. *J. Periodontol.* **1992**, *63*, 729.
14. Daculsi, G.; Passuti, N. *Biomaterials* **1990**, *11*, 86.
15. Daculsi, G.; Passuti, N.; Martin, S.; Deudon, C.; LeGeros, R.Z.; *J. Biomed. Mater. Res.* **1990**, *24*, 379.

16. Daculsi, G.; Bagot D'arc, M.; Corlieu, P.; Gersdorff, M. *Ann. Orol. Rhinol. Laryngol.* **1992**, *101*, 669.
17. Trecant, M.; Delecrin, J.; Royer, J.; Goyenvallé, E.; Daculsi, G. *Clin. Mater.* **1994**, *15*, 233.
18. Gouin, F.; Delecrin, J.; Passuti, N.; Touchais, S.; Poirier, P.; Bainvel, J.V. *Chir. Orthop.* **1995**, *81*, 59.
19. Ransford, P.; Passuti, N.; Chopin, D.; Morin, C. J. *Bone Joint Surg. Br.* **1998**, *80*, 13.
20. Cavagna, R.; Daculsi, G.; Bouler, J.-M. *Longterm Effects Med. Impl.* **1999**, *9*, 403.
21. Soares, E.J.C.; Franca, V.P.; Wykrota, L.; Stumpf, S. *Bioceramics 11*; LeGeros, R.Z., LeGeros, J.P., Eds.; World Scientific: Singapore, 1998; 633 pp.
22. Wykrota, L.L.; Garrido, C.A.; Wykrota, F.H.I. *Bioceramics 11*; LeGeros, R.Z., LeGeros, J.P., Eds.; World Scientific: Singapore, 1998; 641 pp.
23. Malard, O.; Guicheux, J.; Bouler, J.M.; Gauthier, O.; Beauvillain de Montreuil, C.; Aguado, E.; Pilet, P.; LeGeros, R.; Daculsi, G. *Bone* **2005**, *36*, 323–330.
24. Daculsi, G. *Biomaterials* **1998**, *19*, 1473–1478.
25. Daculsi, G.; Laboux, O.; Malard, O.; Weiss, P. *J. Mater. Sci. Mater. Med.* **2003**, *14*, 195–200.
26. LeGeros, R.Z.; Lin, S.; Rohanizadeh, R.; Mijares, D.; LeGeros, J.P. *J. Mater. Sci. Mat. Med.* **2003**, *14*, 201–210.
27. Hench, L.L.; Splinter, R.J.; Allen, W.C.; Greelee, T.K. *J. Biomed. Mater. Res.* **1971**, *2*, 117–141.
28. LeGeros, R.Z. Calcium phosphates in oral biology and medicine. In *Monographs in Oral Sciences*; Myers, H., Ed.; S. Karger: Basel, 1991; Vol. 15.
29. Basle, M.; Chappard, D.; Grizon, F.; Filmon, R.; Daculsi, D.; Rebel, A. *Calcif. Tissue Int.* **1993**, *53*, 348.
30. Gauthier, O.; Bouler, J.M.; Aguado, E.; Pilet, P.; Daculsi, G. *Biomaterials* **1998**, *11*, 133.
31. Osborne, J.; Newesely, H. *Biomaterials* **1980**, *1*, 108.
32. Heughebaert, M.; LeGeros, R.Z.; Gineste, M.; Guilhem, A. *J. Biomed. Mat. Res.* **1988**, *22*, 257–268.
33. Daculsi, G.; LeGeros, R.Z.; Heughebaert, M.; Barbieux, I. *Calcif. Tissue Int.* **1990**, *46*, 20–27.
34. LeGeros, R.Z.; Daculsi, G.; Orly, I.; LeGeros, J.P. *The Bone Biomaterial Interface*; Davies, J.E.D., Ed.; University of Toronto Press, 1991; 76–88.
35. LeGeros, R.Z. *Prog. Cryst. Growth Charact.* **1981**, *4*, 1.
36. Bouler, J.M.; LeGeros, R.Z.; Daculsi, G. *J. Biomed. Mater. Res.* **2000**, *51*, 680.
37. LeGeros, R.Z.; Zheng, R.; Kijkowska, R.; Fan, D.; LeGeros, J.P. *Characterization, Performance of Calcium Phosphate Coatings for Implants*; Horowitz, E., Parr, J.E., Eds.; American Society For Testing Materials Stp: Philadelphia, 1994; Vol. 1, 198,43 pp.
38. Bouler, J.M.; Trecant, M.; Delecrin, J.; Royer, J.; Passuti, N.; Daculsi, G. *J. Biomed. Mater. Res.* **1996**, *32*, 603.
39. Hubbard, W.G. Ph.D. Thesis. Marquette University: Milwaukee, 1974.
40. Schmitt, M. Ph.D. Thesis. Université De Nantes: Nantes, 2000.
41. Driessen, A.A.; Klein, C.P.A.T.; De Groot, K. *Biomaterials* **1982**, *3*, 113.
42. Rohanizadeh, R.; Padrines, M.; Bouler, J.M.; Couchourel, G.; Fortun, Y.; Daculsi, G. *J. Biomed. Mater. Res.* **1998**, *42*, 530.
43. Gauthier, O.; Bouler, J.M.; Aguado, E.; LeGeros, R.Z.; Pilet, P.; Daculsi, G. *J. Mat. Sci. Mat. Med.* **1999**, *10*, 99.
44. LeGeros, R.Z. *Clin. Mater.* **1993**, *14*, 65.
45. Soueidan, A.; Gan, O.I.; Bouler, J.M.; Gouin, F.; Daculsi, G. *Cells Mater.* **1995**, *5*, 31.
46. Benahmed, M.; Bouler, J.M.; Heymann, D.; Gan, O.; Daculsi, G. *Biomaterials* **1996**, *17*, 2173.
47. Yamada, S.; Heymann, D.; Bouler, J.M.; Daculsi, G. *Biomaterials* **1997**, *18*, 1037.
48. Chang, Y.L.; Stanford, C.M.; Keller, J.C. *J. Biomed. Mater. Res.* **2000**, *52*, 270.
49. Yuan, H.; Kurashina, K.; Joost de Bruijn, D.; Li, Y.; de Groot, K.; Zhang, X. *Biomaterials* **1999**, *20*, 1799–1806.
50. Kuboki, Y.; Takita, H.; Kobayashi, D. *J. Biomed. Mater. Res.* **1998**, *3*, 190.
51. Ripamonti, U. *J. Bone Joint Surg. Am.* **1991**, *73*, 692–703.
52. Ripamonti, U. *Biomaterials* **1996**, *17*, 31–35.
53. Ripamonti, U.; Crooks, J.; Kirbride, A. *South Africa J. Sci.* **1999**, *95*, 335–343.
54. Daculsi, G.; Corre, P.; Malard, O.; Legeros, R.; Goyenvallé, E. *Bioceramics 18*, key engineering, 2005.
55. Daculsi, G.; Weiss, P.; Delecrin, J.; Grimandi, G.; Passuti, N.; Guerin, F. *Compositions pour Biomateriaux*. Patent 94-01-414, 235 1994.
56. Daculsi, G.; Weiss, P.; Bouler, J.M.; Gauthier, O.; Aguado, E. *Bone* **1999**, *25*, 59–61.
57. Millot, F.; Grimandi, G.; Weiss, P.; Daculsi, G. *Cells Mat.* **1999**, *9*, 21–30.
58. Constanz, B.R.; Ison, I.C.; Fulmer, M.T.; Poser, R.D.; Smith, S.T.; VanWagoner, M.; Ross, J.; Goldstein, S.A. *Science* **1995**, *267*, 1796–1799.
59. Niwa, S.; LeGeros, R.Z. *Tissue Engineering and Biodegradable Equivalents. Scientific and Clinical Applications*; Lewandrowski, K.-U., Wise, D.L., Trantolo, D.J., Gresser, J.D., Eds.; Marcel Dekker: New York, 2002; 395 pp.
60. LeGeros, R.Z. *Clin. Orthopaed. Rel. Res.* **2002**, *395*, 81.
61. Reddi, A.H. *Tissue Eng.* **2000**, *6*, 351.
62. Barrere, F.; van der Valk, C.M.; Dalmeijer, R.A.; Meijer, G.; van Blitterswijk, C.A.; de Groot, K.; Layrolle, P. *J. Biomed. Mater. Res.* **2003**, *66A*, 779.
63. Habibovic, P.; van der Valk, C.M.; van Blitterswijk, C.A.; De Groot, K.; Meijer, G. *J. Mater. Sci. Mater. Med.* **2004**, *15*, 373–380.
64. Habibovic, P.; Li, J.; van der Valk, C.M.; Meijer, G.; Layrolle, P.; van Blitterswijk, C.A.; de Groot, K. *Biomaterials* **2005**, *26*, 23.
65. Le Nihouhannen, D.; Daculsi, G.; Gauthier, O.; Saffarzadeh, A.; Delplace, S.; Pilet, P.; Layrolle, P. *Bone* **2005**, *36*, 1086–1096.
66. Daculsi, G.; Layrolle, P. *Key Eng. Mat.* **2004**, *254–256*, 1005.
67. De Groot, J. *Tissue Eng.* **1998**, *4*, 337.
68. Diaz-Flores, L.; Gutierrez, R.; Lopez-Alonso, A.; Gonzalez, R.; Varela, H. *Clin. Orthop.* **1992**, *6*, 280.
69. Steitz, S.A.; Speer, M.Y.; Curinga, G.; Yang, H.Y.; Haynes, P.; Aebbersold, R.; Schinke, T.; Karsenty, G.; Giachelli, C.M. *Circ. Res.* **2001**, *89*, 1147–1154.

Blood-Material Interactions

Stephen R. Hanson

Dept. of Biomedical Engineering, Emory University, Atlanta, Georgia, and Department of Biomedical Engineering, Oregon Health & Science University, Beaverton, Oregon, U.S.A.

B

INTRODUCTION

In general terms, blood-material interactions are the interactions occurring reversibly and irreversibly between surfaces and blood components, including small solute molecules, proteins, and cells. The types of interactions include adsorption, absorption, adhesion, denaturation, activation, cellular spreading, and cell-cell interactions, all of which occur dynamically under variable conditions of exposure time, blood composition, and blood flow. Following contact with blood, most—if not all—artificial surfaces quickly acquire a layer of adsorbed blood proteins whose composition may vary with time in a complex manner, which depends in part on the physicochemical properties of the substrate surface. Plasma contains more than 100 identified proteins with specific functions and varying biologic properties. Unfortunately, the relationships between material surface properties, the initially deposited protein layer, cell-surface interactions, and later events, which could be of clinical consequence for the evaluation of biocompatibility and the design of blood-contacting cardiovascular devices, are not well understood. This is because the dynamics of protein-cell-surface interactions under flow are extraordinary complex and therefore difficult to study. Accordingly, the following discussion will consider primarily those blood proteins, cellular elements, biochemical pathways, and physical factors that contribute most directly to adverse blood-material interactions, e.g., thrombosis and thromboembolism, which currently limit the usefulness and more widespread application of devices used in contact with blood.

HEMOSTASIS

Mechanisms of physiological hemostasis are designed to maintain the integrity of the vascular system by quickly causing the arrest of bleeding from injured blood vessels, which thereby facilitates later vessel repair through normal healing mechanisms. These usually advantageous processes may also produce adverse reactions when the synthetic artificial surfaces comprising most prosthetic cardiovascular devices are placed in contact with blood. Primary hemostasis

involves a complex set of interdependent biochemical and cellular reactions between: 1) the material or device surface; 2) circulating blood platelets, with less direct or later interactions involving erythrocytes and leukocytes; 3) blood proteins, especially coagulation enzymes that may trigger the formation of a blood clot or thrombus that may obstruct blood flow or detach (embolize) to occlude distal microcirculatory beds; and 4) the fibrinolytic mechanism, which ultimately lyses the residual thrombus mass. The process is localized to the surface of a device or injured blood vessel by a balance between activation and inhibition mechanisms, which thereby maintains the fluidity of blood in the circulation. Although a great deal is known regarding blood responses to injured arteries and blood-contacting devices, the interrelationships are complex and not fully defined in many instances and particularly regarding the physicochemical properties of synthetic materials that are exposed to blood. Accordingly, thrombosis and thromboembolism complicate to some extent the use of all blood-contacting devices, whether intended for short-term or long-term applications and despite the general use of anticoagulants. These topics will be reviewed briefly. A more detailed discussion has been given in other reports.^[1–3]

Blood Coagulation

At least 12 plasma proteins interact in a series of enzymatic and cofactor reactions leading to blood clotting. The coagulation factors are identified by Roman numerals and were named in order of their discovery while the clotting scheme was being elucidated experimentally. The biochemical properties of the coagulation factors are summarized in Table 1. Initiation of clotting occurs either intrinsically through surface-mediated reactions or extrinsically through activation of factor X by the tissue factor:factor VIIa complex. The two systems converge upon a final common pathway, which leads to the formation of the key procoagulant enzyme thrombin. Thrombin acts enzymatically on fibrinogen, which then polymerizes to form an insoluble fibrin gel that stabilizes the thrombus mass.

Table 1 Properties of human clotting factors

Clotting factor	Apparent molecular weight (number of chains)	Approximate normal plasma concentration ($\mu\text{g/ml}$)	Active form
Intrinsic clotting system			
Prekallikrein	100,000 (1)	50	Serine protease
High molecular weight kininogen	120,000 (1)	80	Cofactor
Factor XII	80,000 (1)	30	Serine protease
Factor XI	160,000 (2)	3–6	Serine protease
Factor IX	57,000 (1)	3–5	Serine protease
Factor VIII ^a	330,000 (1)	0.2	Cofactor
Von Willebrand factor ^a	250,000 (1)	10	Cofactor for platelet adhesion
Extrinsic clotting system			
Tissue factor	44,000 (1)	0 ^b	Cofactor
Factor VII	50,000 (1)	1	Serine protease
Common pathway			
Factor X	59,000 (2)	5	Serine protease
Factor V	330,000 (1)	5–12	Cofactor
Prothrombin	72,000 (1)	140	Serine protease
Fibrinogen	340,000 (6)	2500	Fibrin polymer
Factor XIII	320,000 (4)	10	Transglutaminase

^aIn plasma, Factor VIII is complexed with von Willebrand factor, which circulates as a series of multimers ranging in molecular weight from about 600,000 to 2×10^6 .

^bThe tissue factor concentration in cell-free plasma is minimal because tissue factor is an integral cell membrane-associated protein expressed by vascular and inflammatory cells, although a soluble form of tissue factor has recently been described.^[4]

Coagulation proceeds through a cascade of reactions whereby normally inactive factors, e.g., proenzymes, become activated after surface contact or proteolytic cleavage by precursor enzymes, e.g., factor XIIa activates factor XI to factor XIa. These enzymes, in turn, activate other precursor molecules, e.g., factor XIa converts factor IX to factor IXa. Because this sequence involves a series of steps and one enzyme's molecule can act upon many substrate molecules, the reactions are quickly amplified, which result in the generation of significant amounts of thrombin. Thrombin potently activates circulating platelets, which further amplify coagulation reactions by localizing clotting factors to platelet membrane phospholipids that act as a surface catalyst, i.e., several clotting reaction steps proceed at an effective rate only on the surface of activated platelets. The accumulation of platelets and the formation of fibrin at sites of vessel injury result in the arrest of bleeding. The process is further localized, i.e., widespread or systemic clotting does not occur, due to the dilution of activated factors by blood flow and the actions of coagulation enzyme

inhibitors that are present in circulating blood or are generated within clotting blood.

Fig. 1 presents a diagram of the classical clotting factor cascade for both the intrinsic and extrinsic systems and the common pathway that begins with the activation of factor X. Except for the contact phase, calcium (Ca^{++}) is required for most clotting reactions. Accordingly, chelators of calcium such as sodium citrate and ethylenediamine tetra-acetic acid (EDTA) are commonly used as anticoagulants for in-vitro blood studies. It should be recognized that the in-vitro interactions of clotting factors, clotting, is not identical with coagulation in living systems due to the presence in vivo of complex blood flow phenomena, the proximity of vascular tissues that also modulate coagulation reactions, and normal physiological mechanisms of blood recirculation, inactivation, and dilution. There are also interrelationships between the intrinsic and extrinsic systems such that crossover or reciprocal activation reactions may be important, such as the activation of factor IX by either extrinsic pathway factor VIIa and tissue factor or intrinsic pathway

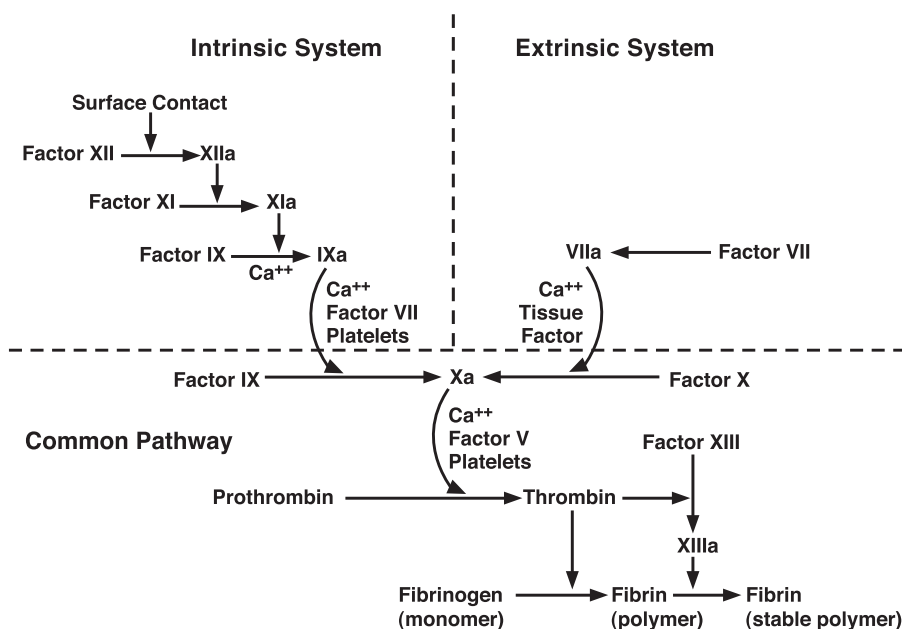


Fig. 1 Blood coagulation. Clotting factors (proenzymes), identified by Roman numerals, interact in a sequential series of enzymatic activation reactions (coagulation cascade) leading to the amplified production of the enzyme thrombin, which in turn activates fibrogen to form a fibrin polymer that stabilizes the clot or thrombus.

factor XIa.^[1-5] Due to the complexity of the hemostatic system and the difficulty of studying hemostasis *in vivo*, many aspects of coagulation, thrombus formation, and adverse blood-material interactions remain to be elucidated.

In the intrinsic system, contact factor activation refers to reactions following adsorption to surfaces of a group of proteins known as contact factors. The most potent artificial surfaces for activating blood are negatively charged, e.g., glass. Although the contact reactions have been well studied *in vitro*, their pathologic significance remains uncertain. In hereditary disorders, only very low levels of factor XI are associated with abnormal bleeding although normal levels of factor XI may contribute to the growth and propagation of thrombus. Contact factors include factors XII, XI, prekallikrein, and high molecular weight kininogen (HMWK). All contact reactions take place in the absence of calcium. Kallikrein is also involved in the fibrinolytic system and inflammation.^[5]

A middle phase of intrinsic clotting begins with the first calcium-dependent step, the activation of factor IX by factor XIa. Factor IXa subsequently activates factor X. Factor VIII is an essential cofactor for the intrinsic pathway activation of factor X. To exhibit cofactor activity, factor VIII first requires modification by an enzyme such as thrombin. In the presence of calcium, factors IXa and VIIIa form a complex on phospholipid surfaces, e.g., on the surface of activated platelets, that then activates factor X. This reaction proceeds slowly in the absence of an appropriate phospholipid surface that serves to localize these biochemical events to the surface (versus bulk fluid) phase of blood.

The extrinsic system of blood clotting is initiated by the activation of factor VII. When factor VIIa interacts with tissue factor, which is expressed by many cell types, the tissue factor:factor VIIa complex becomes an active enzyme that is the extrinsic pathway activator of factor X. Tissue factor becomes available when underlying vascular structures (the subendothelium) are exposed to flowing blood upon vessel injury, or tissue factor may be expressed by activated white cells that are involved in inflammatory reactions.

The common pathway begins when factor X is activated by either factor VIIa:tissue factor or the factor IXa-VIIIa complex. After the formation of factor Xa the next step involves factor V, a cofactor, e.g., factor VIII, which has activity only after modification by another enzyme such as thrombin. Factor Xa-Va, in the presence of calcium and platelet phospholipids, is then capable of converting prothrombin (factor II) to thrombin. Like the conversion of factor X by phospholipid-VIIIa-IXa (tenase complex), prothrombin activation is similarly catalyzed by phospholipid-Va-Xa (prothrombinase complex). The higher plasma concentration of prothrombin (Table 1), as well as the biologic amplification capacity of the clotting cascade, allows for a few molecules of an activated initiator to generate a large amount of thrombin activity. Thrombin, in addition to its ability to modify factors V, factor VIII, factor XI, and activate platelets, acts enzymatically upon two other important substrates: fibrinogen and factor XIII. The action of thrombin on fibrinogen releases small peptides from fibrinogen, e.g., fibrinopeptide A, that can be assayed in plasma as evidence of thrombin activity. After thrombin cleavage, the fibrin monomers then polymerize to form an

insoluble fibrin gel. Factor XIII is either trapped within the clot or provided by platelets and is activated directly by thrombin. A tough, cross-linked fibrin polymer is formed by interaction of the fibrin polymer with factor XIIIa.

Control Mechanisms

Obviously, the body has mechanisms for avoiding explosive coagulation and thrombus formation once coagulation is initiated. At least four types of mechanisms may be considered. First, blood flow may reduce the localized concentration of precursors and removes activated materials by hemodilution into the circulating blood pool. Filtration and rapid removal of inactivated factors is accomplished by passage through the liver. Second, the rate of several clotting reactions is fast only when the reaction is catalyzed by a surface. These reactions include the contact reactions, the activation of factor X by factor VII-tissue factor at sites of tissue injury, and reactions that are accelerated by deposited platelet phospholipids (activation of factor X and prothrombin). Third, there are naturally occurring inhibitors of coagulation enzymes, including serine protease inhibitors such as antithrombin III, which potently inactivate thrombin and other coagulation enzymes. Consequently, plasma levels of thrombin:antithrombin III complex can also be assayed as a measure of thrombin production *in vivo*. Fourth, during the process of coagulation, enzymes are generated that not only activate coagulation factors but also degrade cofactors. For example, the fibrinolytic enzyme plasmin degrades fibrinogen and fibrin monomers and can inactivate cofactors V and VIII. Of particular importance is the thrombin:thrombomodulin pathway. Thrombin is ultimately removed from the circulation after it binds to thrombomodulin, a protein found on the surface of blood vessel endothelial cells. Thrombin interaction with endothelial cell thrombomodulin also results in a complex that activates circulating protein C to form activated protein C. Activated protein C, a key physiologic anticoagulant protein, in turn inactivates coagulation factors V and VIII, which thereby down regulates thrombin production. While thrombin produced locally upon vessel injury or in association with the use of cardiovascular devices is a potent trigger for thrombosis, circulating thrombin, paradoxically, may produce a systematic anticoagulant state through activation of protein C.^[6] A number of other physiologic inhibitory systems have also been described.^[1] In summary, the platelet, coagulation, and endothelial systems interact in a number of ways that serve to promote localized hemostasis while preventing generalized blood clotting and thrombosis.

PLATELETS

Introduction

Platelets are nonnucleated, disk-shaped cells having a diameter of 3–4 μm and an average volume of $10 \times 10^{-9} \text{ mm}^3$. Platelets are produced by the fragmentation of bone marrow megakaryocytes and circulate at an average concentration of about 250,000 cells per μl of blood. Platelets occupy approximately only about 0.23% of the total blood volume. In contrast, red cells typically circulate at 5×10^6 cells per μl and may compose 40–45% of the total blood volume. A comparison of the types, concentrations, and sizes of primary blood cellular elements is given in Table 2. Platelets are of particular interest since they markedly accelerate coagulation reactions and play a key role in thrombus formation in arteries—processes that can lead to heart attacks and strokes. As described subsequently, platelet functions are designed to initially arrest bleeding through formation of platelet plugs and stabilize platelet plugs by catalyzing coagulation reactions leading to the formation of fibrin.

Platelet structure provides a basis for understanding platelet function. In the normal (nonstimulated) state, the platelet discoid shape is maintained by a circumferential bundle (cytoskeleton) of microtubules. The external surface coat of the platelet exposes membrane bound receptors, e.g., glycoproteins Ib and IIb/IIIa, that mediate the platelet functions of adhesion (platelet-surface) and platelet aggregation (platelet-platelet). The membrane of activated platelets also contains negatively charged phospholipids, e.g., phosphatidyl serine, upon which coagulation reactions efficiently proceed. The platelet membrane forms a spongy, canal-like (canalicular) open network that represents an expanded reactive surface to which plasma factors are selectively adsorbed and activated. Thus, when platelets aggregate, platelet coagulant activity is produced, including expression of membrane phospholipids that accelerate two critical steps of the blood coagulation sequence, factor X activation and the conversion of prothrombin to thrombin. Platelets also promote the proteolytic activation of factors XII and XI. Thus, the surface of the aggregated platelets serves as a site

Table 2 Blood cell parameters

	Concentration (per μl)	Average size (fl)	Volume (% of blood)
Red cells	5,000,000	87	43
Platelets	250,000	9	0.23
White cells	7,000	383	0.27

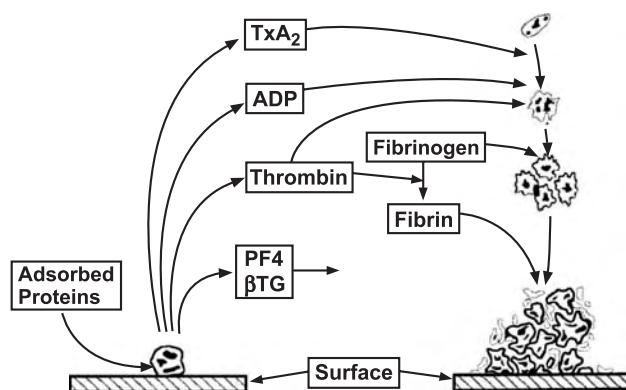


Fig. 2 Platelet reaction at artificial surfaces. Following protein adsorption, platelets adhere and release their α -granule contents, including ADP. Thrombin production is catalyzed locally by platelet membrane phospholipids. Thromboxane A₂ (TxA₂) is synthesized. ADP, TxA₂ and thrombin then recruit additional circulating platelets into an enlarging platelet mass that is stabilized by fibrin.

where thrombin can form rapidly and locally in excess of the capacity of the anticoagulant mechanisms of blood.

Platelets also contain substantial quantities of muscle proteins, e.g., actin and myosin, which allow for internal contraction when platelets are activated. Platelets also contain three types of cytoplasmic storage granules: 1) α -granules, which are numerous and contain the platelet-specific proteins platelet factor 4 (PF-4) and β -thromboglobulin (β -TG) as well as other proteins that are found in plasma (fibrinogen, albumin, fibronectin, and coagulation factors V and VIII); 2) dense granules that contain adenosine diphosphate (ADP), calcium ions (Ca^{++}), and serotonin; and 3) lysosomal granules containing enzymes (acid hydrolases). Platelets are extremely sensitive cells that may respond to stimulation by a variety of agonists, such as thrombin and epinephrine, as well as other agonists which are released by platelets themselves, e.g., ADP and serotonin. Activation causes platelets to become adhesive and change in shape to irregular spheres with multiple spiny pseudopods, which facilitate platelet adhesive encounters, accompanied by internal contraction and extrusion of storage granule contents into the extracellular environment. These secreted platelet products stimulate ambient circulating platelets and cause irreversible platelet aggregation, which leads to the formation of a fused platelet thrombus (Fig. 2).

Platelet Adhesion

In the process of hemostasis, platelets initially adhere to artificial surfaces and injured blood vessels. At sites of vessel injury, platelet adhesion involves the

interaction of platelet glycoprotein Ib (GP Ib) and connective tissue elements that become exposed, e.g., collagen, and requires plasma von Willebrand factor (vWF) as an essential cofactor. GP Ib (about 15,000 molecules per platelet) acts as the surface receptor for vWF. The hereditary absence of GP Ib or vWF results in defective platelet adhesion and a serious bleeding tendency. Platelet adhesion to artificial surfaces may be mediated by other plasma adhesive glycoproteins, such as fibrinogen, which bind to the integrin receptor platelet glycoprotein IIb/IIIa. GP IIb/IIIa (about 40,000 copies per platelet) is the platelet receptor for adhesive plasma proteins that support cell attachment and spreading, including fibrinogen, vWF, fibronectin, and vitronectin. Resting platelets do not bind these proteins, which normally occurs only after platelet activation causes a conformational change in GP IIb/IIIa that renders it competent to bind ligands. Platelets activated near surfaces, e.g., by exposure to factors released from thrombus such as thrombin and ADP, could adhere through this mechanism, e.g., to surface-adsorbed fibrinogen. Also, normally unactivated GP IIb/IIIa receptors could react with surface proteins that have undergone conformational changes or denaturation as a result of their adsorption to surfaces. Supporting this view is the enhanced adhesiveness of platelets towards surfaces preadsorbed with fibrinogen that has been commonly observed in in-vitro studies. Following adhesion, activation, and release reactions, the expression of functionally competent GP IIb/IIIa receptors may also support tight binding and platelet spreading through multiple focal contacts with fibrinogen and other surface-adsorbed adhesive proteins.

Platelet Aggregation

Following platelet adhesion, a complex series of reactions is initiated involving: 1) the release of dense granule contents such as ADP; 2) the formation of small amounts of thrombin at the platelet surface; and 3) the activation of platelet biochemical processes to generate other platelet-activating species such as thromboxane A₂. The release of ADP, thrombin formation, and thromboxane A₂ generation act in concert to recruit platelets into an enlarging platelet aggregate (Fig. 2). Platelet stimulation by these agonists causes the increased expression on the platelet surface of GP IIb/IIIa, mobilized from platelet granular stores, which then binds plasma proteins that support platelet aggregation. Under normal circumstances, fibrinogen, due to its relatively high concentration in blood (Table 1), is probably the most important protein-supporting platelet aggregation. The platelet-platelet interaction is calcium-dependent and involves bridging of adjacent

platelets by fibrinogen molecules, e.g., platelets aggregate poorly or not at all in the absence of fibrinogen, GP IIb/IIIa, and calcium. Thrombin binds directly to platelet-associated thrombin receptors (PARs) which, following cleavage by thrombin, cause platelet activation.^[7] Thus, thrombin plays a key role in platelet aggregation and thrombus formation by activating platelets that then catalyze the production of more thrombin; stimulating the release of ADP, thromboxane A₂ and other platelet agonists; and enzymatically cleaving fibrinogen to form fibrin that stabilizes the growing platelet mass.

Platelet Release Reaction

The release reaction is the secretory process whereby substances stored in platelet granules are extruded from the platelet into its external environment. ADP, collagen, epinephrine, and thrombin are physiologically important release-inducing agents and interact with the platelet through specific receptors for these agonists on the platelet surface membrane. The alpha-granule contents (PF-4, β -TG, and other proteins) are readily released by relatively weak agonists such as ADP. Release of the dense granule contents (ADP, calcium, and serotonin) requires platelet stimulation by a stronger agonist such as thrombin. Agonist binding to platelets also initiates the formation of intermediates that cause activation of the contractile-secretory apparatus, production of thromboxane A₂, and mobilization of calcium from intracellular storage sites. Elevated cytoplasmic calcium is probably the final mediator of platelet aggregation and release. As noted previously, substances that are released (ADP), synthesized (TxA₂), and generated (thrombin) by platelets affect other platelets and actively promote their recruitment into growing platelet aggregates and thrombi. Measurements of plasma levels of platelet-specific proteins, such as PF-4 and β -TG, have been widely used as indirect markers of platelet activation and release that may occur in various disease in vivo.

Platelet Consumption

The turnover or rate of destruction of platelets has also been used as a diagnostic test for thrombosis. In man, platelets are cleared from circulating blood in an approximately linear fashion over time with an apparent lifespan of approximately 10 days. With ongoing thrombosis that may be produced by cardiovascular devices, platelets may be removed from circulating blood at a more rapid rate. Thus, steady-state elevations in the rate of platelet destruction, as reflected in a shortening of platelet lifespan, have been used as a

measure of the thrombogenicity of artificial surfaces and prosthetic devices.^[8,9]

Fibrinolysis

The fibrinolytic system removes unwanted fibrin deposits and thrombus in order to improve blood flow through injured vessels after the arrest of bleeding and facilitate the healing process after injury and inflammation. The fibrinolytic mechanism is a multi-component system composed of precursors, activators, cofactors, and inhibitors and has been studied extensively.^[1-3] In brief, the fibrinolytic system interacts with the coagulation system at the level of contact activation.^[5] The most well-studied fibrinolytic enzyme is plasmin, which circulates in blood as the precursor plasminogen. Plasminogen adheres to a fibrin clot and is incorporated into the fibrin mesh during polymerization. Plasminogen is activated to plasmin by the actions of plasminogen activators that may be present in blood or released from tissues or may be administered therapeutically. Important plasminogen activators occurring naturally in man include tissue plasminogen activator (tPA) and urokinase, which are also given therapeutically in patients having heart attacks, strokes, and other thromboembolic disorders. Following activation, plasmin digests the fibrin clot by releasing soluble fibrin/fibrinogen digestion products (FDP) into circulating blood that may be assayed as markers of in-vivo fibrinolysis, e.g., D-D dimer fragment of fibrin. In a manner analogous to the regulation of coagulation, the activity of the fibrinolytic system is modulated by naturally circulating inhibitors of plasmin, e.g., antiplasmins such as α -2-antiplasmin, or plasminogen activator inhibitors, e.g., plasminogen activator inhibitor 1 or PAI-1.

COMPLEMENT

The complement system is primarily designed to effect a biologic response to antigen-antibody reactions. In a manner similar to coagulation and fibrinolytic reactions, complement proteins are activated enzymatically through a complex series of reaction steps.^[5] Several proteins in the complement cascade exhibit activity as inflammatory mediators. Ultimately, activation of the complement system results in the generation of an enzymatic complex that causes irreversible damage by lytic mechanisms to the membranes of antigen-presenting cells, e.g., bacteria.

There are multiple interactions between the complement, coagulation, and fibrinolytic systems that have been reviewed elsewhere.^[5] In recent years, there has been considerable interest in the problem of

complement activation by artificial surface, which has been prompted by observation that devices having large surface areas, e.g., hemodialyzers, may cause reciprocal activation reactions between complement enzymes and white cells and complement activation that may mediate both white cell and platelet adhesion to artificial surfaces.^[3,5]

ERYTHROCYTES

Red cells (erythrocytes) are usually considered as passive participants in processes of hemostasis and thrombosis. Although, under some conditions, e.g., low shear or venous flows, red cells may comprise a large proportion of total thrombus mass. The concentration and motions of red cells under flow have important mechanical effects on the diffusive transport of blood elements.^[10] Specifically, blood flow phenomena control the rate of transport by diffusion and convection of cells and proteins in the vicinity of artificial surfaces and thrombi. While physiologically encountered blood shear forces probably do not damage or activate platelets or red cells directly, such forces can dislodge platelet aggregates and thrombi that may attach farther downstream or are carried away (embolize) to distal circulatory beds. Platelet diffusion in flowing blood and early platelet attachment to surfaces may be increased 100-fold by the presence of red blood cells that greatly enhance movement of platelets across parallel streamlines and hence increase the rate at which platelets can participate in thrombus development. As a consequence of red cell motions under laminar flow, platelets may become more concentrated at the periphery of the blood vessel while red cells may have a reduced concentration in the near-wall region. Under some conditions, red cells may also contribute chemical factors that influence platelet reactivity.^[11] The process of direct attachment of red cells to artificial surfaces has been considered to be of minor importance and has therefore received little attention in studies of blood-material interactions.

LEUKOCYTES

The various classes of white cells perform many functions in inflammation, infection, wound healing, and blood response to foreign materials. The reactions of white cells to artificial surfaces may depend upon both the properties of the biomaterial surface as well as local flow conditions, i.e., these reactions may be sensitive to shear stress.^[12,13] White cell adhesive interactions with artificial surfaces may proceed through mechanisms related to activation of the complement, coagulation, fibrinolytic, and other enzyme systems, which result in

the expression by white cells of procoagulant, fibrinolytic, and inflammatory activities. For example, stimulated monocytes express tissue factor that can initiate extrinsic coagulation. Neutrophils may contribute to clot dissolution by releasing potent fibrinolytic enzymes, e.g., neutrophil elastase. White cell interactions with devices having large surface areas may be extensive, especially with surfaces that activate complement, which result in their marked depletion from circulating blood. Activated white cells, through their enzymatic and other activities, may produce organ dysfunction in other parts of the body. The relative importance of many of these reactions for outcomes following the contact of biomaterials with blood remains to be defined.

ARTIFICIAL SURFACES AND CARDIOVASCULAR DEVICES

Introduction

Thousands of devices made from synthetic materials are commonly used in settings that involve their exposure to blood. Since cardiovascular devices as well as the biological environment are complex, even though the materials comprising a device may be considered to be blood compatible, i.e., no adverse events are expected that could compromise device function or host well-being, an implanted device that is fabricated from combinations of those materials will not necessarily be blood compatible. Such devices may be used short-term or implanted permanently and include the extracorporeal pump-oxygenator (heart-lung machine) used in many surgical procedures; catheters for blood access and blood vessel manipulation, e.g., angioplasty and stenting; heart ventricular assist devices; hollow fiber and other types of hemodialyzers for treatment of renal failure; and devices for replacement or repair of diseased heart valves, e.g., prosthetic heart valves, and arteries, e.g., vascular grafts and endovascular stents. Since these and many other blood-contacting devices have been successfully used in patients for many years, they are judged to be therapeutically beneficial. Nonetheless, the performance of many existing devices is also less than optimal.^[14] For example, prolonged cardiopulmonary bypass and membrane oxygenation can produce platelet dysfunction and a severe bleeding tendency. Mechanical heart valves can accumulate thrombus and occasionally shed emboli to the brain, which produce neurological dysfunction. Synthetic vascular grafts perform less well than grafts derived from natural arteries or veins. Graft failure due to thrombosis can lead to ischemia, i.e., lack of oxygen, and injury in downstream tissue beds. At the present time, there are no acceptable synthetic substitutes for

the replacement of small diameter arteries, e.g., less than 6 mm in diameter, which might be used to replace diseased coronary arteries and other blood vessels. Due to the pressing need for blood vessel substitutes, this area remains one of intensive investigation. Thus, while the performance characteristics of many devices is acceptable, i.e., the benefit/risk ratio is high, many existing devices could be improved to extend their period of safe operation, e.g., oxygenators; reduce adverse long-term events, e.g., heart valves; and eliminate or reduce the need for anticoagulating drugs that are required in many instances, e.g., oxygenators, heart valves, and hemodialyzers, and pose a significant risk of increasing bleeding. Device improvements in these areas could result in enormous medical and economic benefits.

Evaluation of Blood-Material Interactions

Despite intensive efforts, the blood responses to specific materials or devices for particular applications are not well-established or predictable because the devices used may be complex in terms of geometry and material composition; the blood flow phenomena associated with device use may be poorly understood regarding the mass transport phenomena that control the rate of blood-material interactions; the possible blood responses are numerous, complex, and dynamic; and it is difficult to measure device thrombogenicity in a manner that is clinically relevant (i.e., predictive of outcomes in man) in either experimental animals or man. Thus, the assessment and interpretation of blood compatibility testing remains uncertain. Most tests purported to measure blood compatibility really evaluate only certain blood-material interactions that are specific to the experimental configuration used and cannot be extrapolated to other conditions involving changes in blood exposure time; blood chemistry, e.g., native versus anticoagulated blood; flow geometry; or species, e.g., animals versus humans. In general, materials and devices that perform well across a wide range of test conditions and, in several animal species, are more likely to perform acceptably in humans as well although this conclusion is not assured until clinical testing has actually been performed. Nonetheless, such tests have provided insights into the mechanisms of thrombus formation and the role of device properties and have permitted some general guidelines for device development. Further improvement of approaches for the rational design of materials and devices that contact blood will require a greater understanding of the physical and biological mechanisms of blood-material interactions.

BLOOD-CONTACTING MATERIALS

Introduction

As a result of the above considerations, it is still not possible to simply rank or classify materials with respect to their suitability for use in particular blood-contacting devices. Nor is it possible to predict in any general way, based on the properties of devices and their blood-contacting surfaces, the behavior of blood in contact with materials or the propensity of devices to produce clinically adverse events. Despite many attempts to correlate biologic responses with device physicochemical property measurements, our success has been largely empirical. These findings suggest that no material may be simply blood-compatible or nonthrombogenic since this assessment will strongly depend on details of the test system or usage configuration. Indeed, under conditions of sluggish (low-shear) blood flow or stasis, most—if not all—synthetic materials may become associated with localized blood clotting and thus be considered thrombogenic. This is because synthetic materials, unlike the endothelium that lines all blood vessels, cannot actively inhibit thrombosis and clotting by directly producing and releasing inhibitors or inactivating procoagulant substances. The development of vascular tissue constructs that possess such favorable biologic activities is one goal of tissue engineering.

Conventional Polymers

Conventional polymers such as poly(ethylene), poly(vinyl chloride), poly(tetrafluoroethylene), and poly(dimethyl siloxane), and certain poly(urethanes) have been used for many decades in short-term blood contact applications, including catheters, cannulas, arteriovenous shunts for hemodialysis, and tubing components of extracorporeal circuits.^[8,14] When used in patients with systemic anticoagulation, the performance of such materials has usually been acceptable. For example, although thrombus on angiographic catheters can be demonstrated in about half of all cases, most thromboembolic or occlusive events are clinically silent and produce significant complications in less than about 1% of procedures. With the use of catheters for venous sampling or drug administration, even the total occlusion of peripheral veins is usually inconsequential. However, longer-term indwelling catheters in a variety of configurations and polymer compositions, particularly in infants and children, are now recognized to involve a significant risk of thrombosis that can ultimately lead to organ or limb damage and even death. Poly(urethanes), due in part to their flexibility and toughness, are perhaps the polymer of

choice for ventricular assist devices, blood pumps, and the next generation of prosthetic heart valves. While all conventional polymers are at risk of eventual degradation and failure due to hydrolysis, calcification, and absorption of lipids, considerable effort is being directed towards overcoming these challenges.

Hydrophilic Polymers

These materials, which preferentially adsorb or absorb water (hydrogels), were initially postulated to be blood-compatible based on the view that many naturally occurring phospholipids comprising the cell membranes of blood contacting tissues are also hydrophilic. Thus, water, as a biomaterial, was expected to have minimal interaction with blood proteins and cells. Interestingly, in animal studies, highly hydrophilic polymers based on acrylic and methacrylic polymers and copolymers as well as poly(vinyl alcohol) were all found to consume platelets, i.e., shorten the circulating platelet lifespan, when implanted for periods up to several weeks although these polymers accumulated little visible thrombus.^[8] While the rates of platelet destruction were variable between materials, the hydrophilic polymers had little capacity to retain adherent thrombus, i.e., they shed deposited blood elements as microemboli. Thus, while hydrogels, which are mechanically weak, are currently used to improve catheter lubricity and serve as reservoirs for drug delivery, they have not received widespread application for improving blood-compatibility. Another hydrophilic polymer that has received considerable attention is poly(ethylene oxide). While poly(ethylene oxide) surfaces have been shown, e.g., hydrogels, to reduce interactions with blood proteins and cells *in vitro*, their use for clinical device applications has not been established.

Metals

Metals tend to be thrombogenic in experimental test systems and are most commonly applied in situations requiring considerable mechanical strength, e.g., the struts of mechanical heart valves and as endovascular stents. Stents are metallic mesh devices mechanically placed within blood vessels to expand the vessel lumen diameter, e.g., after balloon angioplasty, in order to preserve vessel patency and improve blood flow. Metals most commonly employed are stainless steel and tantalum. Guidewire thrombogenicity, although readily documented, has been less of a problem clinically because of the usually short period of blood exposure. Theoretically, reduced thrombogenicity of metallic stents and heart valve components might be achieved by thin film polymer coatings although the

clinical effectiveness of this strategy has not been demonstrated. Thus, systemic anticoagulation or long-term platelet inhibitor therapy, e.g., aspirin, is generally employed in patients with prosthetic heart valves and stents.

Carbons

Cardiac valves fabricated from low temperature isotropic carbons, e.g., pyrolytic carbon, are successfully used clinically. These materials are appropriate for such applications as mechanical valves that require long-term chemical inertness, smoothness, and wear-resistance. The reasons for the marked improvement in the performance, e.g., reduced thrombosis and thromboembolic stroke rates, of these newer versus older style heart valves are not entirely understood, but they are undoubtedly multifactorial and relate to improved patient management and valve design with respect to flow phenomena as well as to the nature of the carbon surface. Clearly, these carbons confer specific benefits with respect to blood cell and protein interactions, which result in a very low frequency of clinical complications, although their development has been largely empirical and dictated by the mechanical requirements of prolonged implantation. The use of carbon coatings has been proposed for other devices, e.g., vascular grafts, and such devices are gaining acceptance clinically.

Biological Surfaces

The use of biological and bioactive molecules as device surface coatings dates back many years to the use of coatings containing albumin, phospholipids, and the anticoagulant heparin. These coatings may confer significant thromboresistance. Phospholipids such as phosphorylcholine, a normal constituent of cell membranes, may orient polar head groups towards the aqueous phase and locally organize water molecules—much like hydrogel surfaces. These surfaces may minimize protein and complement interactions.^[15] In preliminary animal studies, phosphorylcholine-coated stents, guidewires, and vascular grafts have shown improved thromboresistance. This approach is being actively developed for clinical applications. Heparin, a naturally occurring anticoagulant glycosaminoglycan, has been considered an attractive surface coating based on its ability to directly catalyze the inactivation of procoagulant enzymes. Recently, the reduced thrombogenicity and improved biocompatibility of heparinized stents and vascular grafts has been demonstrated in animals. This method has also been widely used clinically for the heparin coating of catheters and other devices, and it is clear that the improved

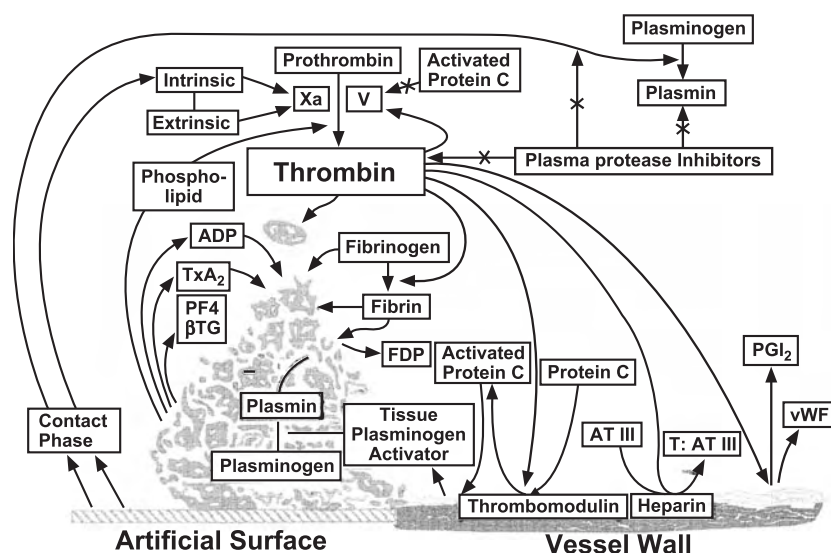


Fig. 3 Integrated hemostatic reactions between a foreign surfaces and platelets, coagulation factors, the vessel endothelium, and the fibrinolytic system.

biocompatibility is a function of heparin's anticoagulant activity as opposed to heparin's nonspecific physicochemical properties. Finally, coatings based on hyaluronic acid have also shown promise for reducing the thrombogenicity of metallic endovascular stents.^[16] In general, biomolecule coatings have shown considerable promise as blood-contacting materials.

CONCLUSION

Interrelated blood systems respond to tissue injury in order to quickly minimize blood loss and remove unneeded deposits after healing has occurred. For purposes of illustration, Fig. 3 depicts several of these relationships and inhibitory pathways that apply to blood reactions following contact with both natural and artificial surfaces. When artificial surfaces are exposed to blood, an imbalance between processes of activation and inhibition of these systems can lead to excessive thrombus formation and an exaggerated inflammatory response. While certainly the key blood cells, proteins, and reaction steps have been identified, many aspects of their biological interrelationships remain poorly understood. Because the variables affecting cardiovascular device responses are sufficiently numerous and complex, including the effects of flow phenomena, those properties of blood-contacting surfaces that would be desirable to minimize adverse reactions cannot, in most instances, be predicted with confidence. The choice of material is often constrained by mechanical property considerations and variable requirements for material durability and chemical stability. Newer materials, including biomolecule coatings, and newer devices, including tissue-engineered constructs having beneficial biological properties

that are found in normal vascular tissue, hold considerable promise for improving the performance of blood-contacting materials that are used in the diagnosis and treatment of vascular disease.

ACKNOWLEDGMENTS

This work was supported by Research Grant HL 31469 from the Heart, Lung, and Blood Institute, the National Institutes of Health, and a grant from the National Science Foundation (EEC-9731643) to the Georgia Institute of Technology.

ARTICLES OF FURTHER INTEREST

Biocompatibility Testing; Foreign Body Response; Heparin-Binding to Improve Biocompatibility; Host Reactions; Hydrogels; In Vivo Cardiovascular Models; Poly(carbonate)urethanes; Protein Adsorption Modeling; Thrombosis

REFERENCES

1. *Hemostasis and Thrombosis*, 3rd Ed.; Colman, R.W., Hirsh, J., Marder, V.J., Salzman, E.W., Eds.; J.B. Lippincott: New York, 1994; 1-1713.
2. Forbes, C.D.; Courtney, J.M. Thrombosis and Artificial Surfaces. In *Haemostasis and Thrombosis*, 2nd Ed.; Bloom, A.L., Thomas, D.P., Eds.; Churchill Livingstone: New York, 1987; 902-921.
3. Cardiovascular biomaterials and biocompatibility. *Cardiovasc. Pathol.* **1993**, *2* (3), 1S-224S.
4. Bogdanov, V.Y.; Balasubramanian, V.; Hathcock, J.; Vele, O.; Lieb, M.; Nemerson, Y. Alternatively spliced

- human tissue factor: A circulating, soluble, thrombogenic protein. *Nat. Med.* **2003**, *9* (4), 458–462.
5. Bennett, B.; Booth, N.A.; Ogston, D. Potential Interactions Between Complement, Coagulation, Fibrinolysis, Kinin-Forming, and Other Enzyme Systems. In *Haemostasis and Thrombosis*, 2nd Ed.; Bloom, A.L., Thomas, D.P., Eds.; Churchill Livingstone: New York, 1987; 267–282.
 6. Hanson, S.R.; Griffin, J.H.; Harker, L.A.; Kelly, A.B.; Esmon, C.T.; Gruber, A. Antithrombotic effects of thrombin-induced activation of endogenous protein C in primates. *J. Clin. Invest.* **1993**, *92*, 2003–2012.
 7. Coughlin, S.R. Protease-activated receptors in vascular biology. *Thromb. Haemost.* **2001**, *86* (1), 298–307.
 8. Hanson, S.R.; Harker, L.A.; Ratner, B.D.; Hoffmann, A.S. In vivo evaluation of artificial surfaces using a nonhuman primate model of arterial thrombosis. *J. Lab. Clin. Med.* **1980**, *95*, 289–304.
 9. *Platelet Kinetics*; Paulus, J.M., Ed.; American Elsevier: New York, 1971; 1–360.
 10. Slack, S.M.; Cui, Y.; Turitto, V.T. The effects of flow on blood coagulation and thrombosis. *Thromb. Haemost.* **1993**, *70* (1), 129–134.
 11. Turitto, V.T.; Weiss, H.J. Red cells: Their dual role in thrombus formation. *Science* **1980**, *207*, 541–544.
 12. Brodbeck, W.G.; Nakayama, Y.; Matsuda, T.; Colton, E.; Ziats, N.P.; Anderson, J.M. Biomaterial surface chemistry dictates adherent monocyte/macrophage cytokine expression in vivo. *Cytokine* **2002**, *18* (6), 311–319.
 13. Shive, M.S.; Brodbeck, W.G.; Colton, E.; Anderson, J.M. Shear stress and material surface effects on adherent human monocyte apoptosis. *J. Biomed. Mater. Res.* **2002**, *60* (1), 148–158.
 14. Salzman, E.W.; Merrill, E.D. Interaction of Blood with Artificial Surfaces. In *Haemostasis and Thrombosis*, 2nd Ed.; Colman, R.W., Hirsh, J., Marder, V.J., Salzman, E.W., Eds.; J.B. Lippincott: Philadelphia, 1987; 1335–1347.
 15. Yu, J.; Lamba, N.M.; Courtney, J.M. Polymeric biomaterials: Influence of phosphorylcholine polar groups on protein adsorption and complement activation. *Int. J. Artif. Organs* **1994**, *17* (9), 499–504.
 16. Verheye, S.; Markou, C.P.; Salame, M.Y.; Wan, B.; King, S.B.; Robinson, K.A.; Chronos, N.A.; Hanson, S.R. Reduced thrombus formation by hyaluronic acid coating of endovascular devices. *Arterioscler. Thromb. Vasc. Biol.* **2000**, *20*, 1168–1172.

Blood Purification

Yukihiko Nosé
Hisashi Okubo

Michael E. DeBakey Department of Surgery, Baylor College of Medicine,
Houston, Texas, U.S.A.

INTRODUCTION

Currently, approximately one million patients are kept alive with the help of hemodialysis. If there were no dialyzers today, all these patients would have died within two weeks of kidney failure. Some hemodialysis patients have lived for longer than 30 years. However, in order to avoid chronic hemodialysis syndrome, which is produced by the insufficient removal of renal toxins, hemodiafiltration or plasmapheresis should be effectively employed. These new technologies should be useful in the treatment of not only metabolic diseases but also immunologically induced diseases.

Plasmapheresis involves the nonspecific or specific removal of plasma factors and cellular removal or modulation, reducing their circulating concentrations and thereby improving the clinical state of the patient.

It is estimated that approximately half a million therapeutic plasma treatment procedures are currently carried out in the world per year, and that this number will increase yearly. Growth is related to the cost-effectiveness of the procedure compared with conventional treatments. Plasma treatment has been employed in a wide variety of disease states that may be generally categorized as metabolic or immunologic. The original procedure of plasma exchange has been expanded to plasma fractionation by membrane, adsorbents, or cytapheresis. The physical characteristics of membrane filters in plasmapheresis will be summarized. While many studies have been anecdotal in nature, plasmapheresis is highly recommended in certain diseases. Through further investigation the role of plasma treatment will be more scientifically understood.

These blood purification technologies are described in detail in *Manual on Artificial Organs: The Artificial Kidney*,^[1] and *Blood Purification—Artificial Kidney and Plasmapheresis*,^[2] together with two issues of *Therapeutic Apheresis*.^[3,4]

HEMODIALYSIS: THE DIALYSIS OF BLOOD

Blood returned to the body from the artificial kidney has a lower concentration of chemicals than the surrounding body tissues. These chemicals will diffuse out

of the tissues into the blood. Blood is returned to the dialyzer many times. This repeated return of the dialyzed blood to the body provides for a continuous diffusion of excess body chemicals from the body cells (Fig. 1).

Routine dialysis procedures take three to five hours and are performed three times a week in chronic patients. Hemodialysis is routinely done in outpatient dialysis centers or in the home. Twelve liters of blood are processed per hour. Blood access is normally made via an internal fistula. An anticoagulant such as heparin is required to prevent blood clotting in the circuit during treatment. Hemodialysis is performed to remove small molecules in blood, physiologically removable by the kidney, based on the principle of osmosis.

Principles of Hemodialysis

The hemodialyzer, or artificial kidney, can perform some but not all of the functions of a natural kidney. The artificial kidney removes waste metabolites and water and maintains chemical equilibrium sufficiently to return the patient to a nearly normal condition for several days. Dialyzers operate outside the patient's body to cleanse the blood. One important issue is that hemodialysis does not fully reproduce the natural kidney function. Thus, after several years of hemodialysis treatment, new diseases emerge due to the insufficient restoration of kidney function. However, the current blood purification technologies will help hemodialysis patients maintain a healthy condition. The blood is usually drawn from a tube in an artery and, after passing through the dialyzer, is returned through a tube inserted into a vein.

What makes it possible to use only the blood to restore the patient to a nearly normal condition? As it flows through the body, the blood collects waste materials and excess water that are discharged when the blood reaches the kidney. The kidneys act as a regulating mechanism by determining how much of the various waste materials, chemicals, and water to remove. By effectively removing these various materials from the blood, the dialyzer temporarily performs the function of the natural kidney and is able to keep the patient in

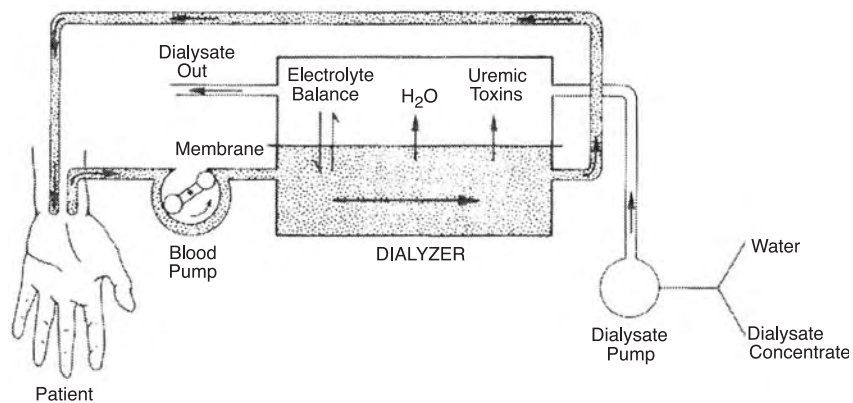


Fig. 1 Hemodialysis circuitry. (From Ref. [2].)

close to a normal condition. Deficiencies in the removal of these waste products can be compensated for by hemodiafiltration or apheresis (described in a later section) (Fig. 2).

Just how the dialyzer performs each of its separate functions is described in the following paragraphs.

Waste Product Removal

Waste products are defined as the unusable end products of metabolism. Metabolism is the process by which food is converted to energy and structural components for the body. The waste products from metabolism accumulate in the bloodstream, and during normal renal function are removed when the blood passes through the kidneys. As a partial substitute for disabled kidneys, these waste products are removed by the process of dialysis.

Dialysis Membrane

Dialysis involves the use of a thin, porous membrane that keeps the blood separated from another fluid, the dialysate. The membrane can be compared to a metal netting. The membrane has perforations, but they are so small that they are invisible to the naked eye. Waste products in the blood are able to pass through these minute perforations into the dialysate fluid, where they are immediately washed away. The perforations in the

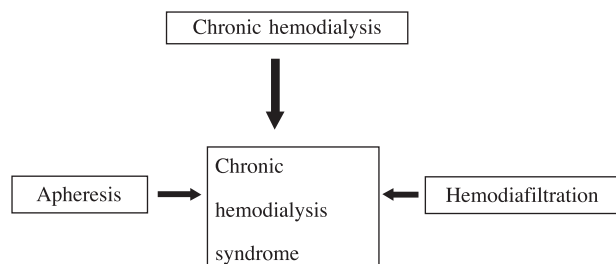


Fig. 2 Hemodialysis and blood purification.

dialysis membrane have an average diameter of 50 Å, with an estimated range of 30 to 90 Å. The original cellulose membrane utilized for hemodialysis was manufactured primarily for sausage casing.

Concentration Gradient

The waste products (as molecules) pass through the membrane by action of a concentration gradient (Fig. 3). Molecules move in a random manner when they occupy a given volume, and in so doing they tend to distribute themselves evenly throughout the entire

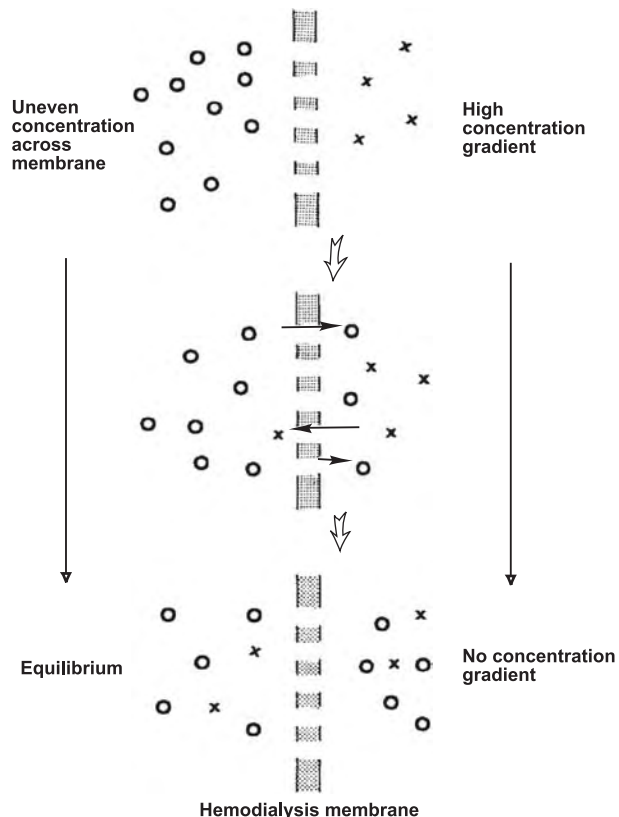


Fig. 3 Concentration gradient. (From Ref. [1].)

volume. The number of molecules in a specific volume represents the concentration. In the artificial kidney, the dialysis fluid is free of waste product molecules; consequently, those in the blood are able to pass through the membrane and tend to distribute themselves evenly throughout the blood and the dialysate. Thus, there is a net movement of waste product molecules from a region of high concentration (the blood) to a region of low concentration (the dialysate). This difference in concentration between the blood and the dialysate, which is the cause of the movement of molecules, is known as the concentration gradient.

Selective Permeability of Molecules in the Blood

There are many different types, as well as sizes, of particles (molecules) in the blood (Table 1). These molecules include waste products such as urea, creatinine, and uric acid; ions (small particles that have an electrical charge), and many other, larger molecules. In addition to the proteins, there are even larger particles (such as erythrocytes and leukocytes) that are the formed elements of the blood. These blood cells, along with the majority of proteins, do not pass through the membrane during dialysis; they are just too large for the perforations (Fig. 4). This selective quality of the membrane—allowing small molecules to pass through

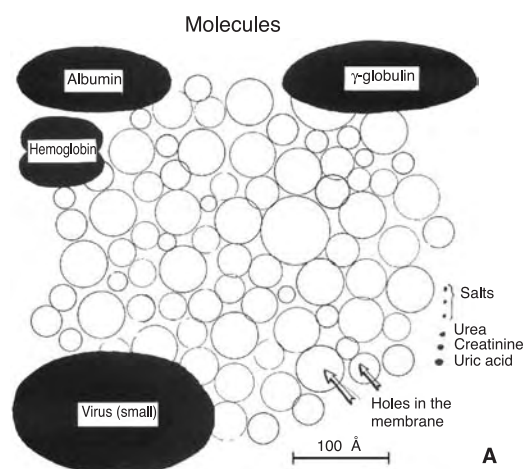


Fig. 4 Permeable smaller molecules and non-permeable larger molecules.

but retarding the passage of progressively larger molecules and particles—represents the real advantage of dialysis. Proteins, erythrocytes, and leukocytes are invaluable to the body, and their loss would have fatal consequences for the patient. Because the membrane can effectively retain these larger substances and at the same time eliminate unwanted waste materials from the blood, dialysis has become an effective agent in the treatment of renal diseases. Furthermore, no bacteria or viruses can be introduced into the blood compartment because they are too large to pass through the perforations in the dialysis membrane.

Maintenance of Balance of Body Chemistry

Many chemicals are present in all parts of the human body. These chemicals appear in several forms—compounds, ions, and aggregates. The relative amounts of these chemicals are critically important to normal body function. For example, potassium is present both in the body cells and in extracellular fluid such as blood. There is a definite relationship between the amount of potassium in the blood and the corresponding amount in the body cells. In blood, the ionic form of potassium is minute and can easily pass through the dialysis membrane into the dialysate fluid. If dialysis were performed with dialysate that contained no potassium, the amount of potassium in the blood would steadily decrease. The ratio between the amount of potassium in the body cells and in the blood would consequently be upset, and the effects could be dangerous. In actual practice, potassium-free baths are used for hemodialysis on uremic patients without any danger. Imbalances are avoided by the addition of chemicals to the dialysate fluid to make their concentration the same as the concentration in the blood. With no concentration

Table 1 Molecules in the blood

Molecules	Molecular weight
Electrolytes	
Sodium	22.989
Chloride	35.453
Potassium	39.102
Calcium	40.08
Magnesium	24.312
Urea	60.06
Creatine	113.14
Uric acid	168.11
Dextrose	180.16
Bilirubin	584.67
Amino acids	
Glycine (smallest)	75.07
Tryptophane (largest)	204.22
Proteins	
Albumin	68,000
γ -globulin	150,000
Hemoglobin	68,000
Fibrinogen	400,000

Source: Ref. [1].

gradient, there is no net movement of these critically important chemicals through the pores of the dialysis membrane. Thus, valuable chemicals small enough to pass through the pores of the dialysis membrane are retained in the blood. Conversely, when the body balance of some vital chemical is upset as a result of disease, it may be restored by properly adjusting the concentration of the chemical in the dialysate fluid. When it is desirable to remove an excess of a chemical from the blood, that chemical is not added to the dialysate fluid. This will result in a large concentration gradient, favoring net movement of the chemical from the blood to the dialysate. To add a given chemical to the blood, the procedure is reversed.

Water Removal

Water molecules are extremely small and pass easily through the pores of the dialysis membrane. The water

concentration is the percentage of water molecules in a specific volume containing many different molecules. All other molecules and particles dispersed among the water molecules are known as solutes, and the water is called the solvent. The two together are referred to as a solution. If the number of solutes is significantly increased, then the water concentration is decreased. In practice this principle is one basis for removing excess water from the blood during dialysis. There are two methods of water removal: the addition of a substance to the dialysate fluid and ultrafiltration.

A large amount of some harmless substance such as dextrose is added to the dialysate fluid to decrease the water concentration. Because the water concentration in the blood is higher than that in the dialysate fluid, a concentration gradient is set up that results in a net movement of water out of the blood into the dialysate fluid. This phenomenon is commonly referred to as osmosis (Fig. 5).

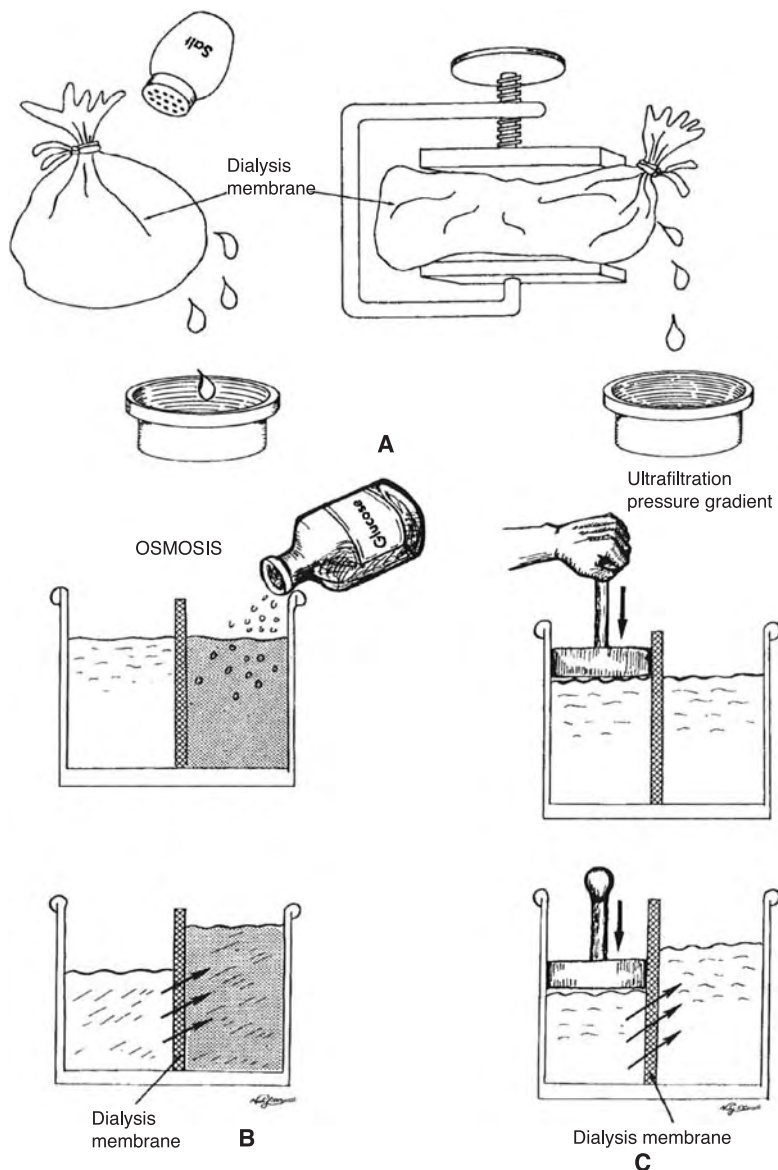


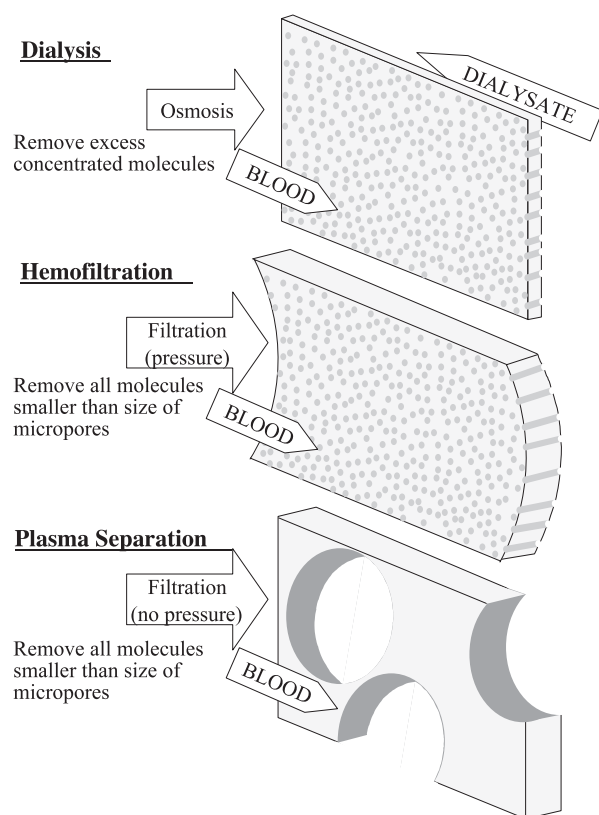
Fig. 5 Water removal. (From Ref. [1].)

Table 2 Conditions for an effective dialyzer (optimal metabolic clearance)

Large surface area of membrane
Good dialysis membrane
Small, constant priming volume
Thin channel for blood
Even distribution of blood
Short channel of blood
Effective mixing of blood
High blood pressure (if a blood pump is not used)
High blood flow
Fresh dialysate and effective dialysate flow
Countercurrent flow of dialysate and blood
High flow rate of dialysate
Effective exposure of the dialysis membrane to the dialysate
High temperature of dialysate

Source: Ref. [1].

Excess water is usually removed from the blood by another process known as ultrafiltration. This process is more effective than the addition of dextrose to the dialysate. Ultrafiltration of water through a dialysis membrane works on the principle of pressure difference. In order to make water flow through a pipe or other

**Fig. 6** Differences of dialysis, hemofiltration, and plasma separation.

conduit, the pressure applied to the fluid at the upstream end of the pipe must be greater than the pressure at the downstream end. One can imagine an analogous situation in the artificial kidney during dialysis. Blood and dialysate flow through compartments separated from each other by the dialysis membrane. Since the water molecules in both the blood and the dialysate fluid can easily pass through the membrane, the direction of flow will be primarily determined by the pressure difference (Fig. 5). At points along the membrane length where the pressure on the blood is greater than the pressure on the dialysate fluid, the water flows from the blood into the dialysate. When the situation is reversed, water flows from the dialysate into the blood.

Since ultrafiltration is a much more effective means of removing excess water than osmosis, it is usually favored if practical. The pressure difference in the dialyzer can be adjusted to remove variable amounts of excess water. Hemofiltration and hemodiafiltration (the combination of dialysis and filtration) are very efficient for removing smaller molecules (smaller than albumin) and low-concentration molecules. On the other hand, hemodialysis is very efficient for removing smaller (than albumin) molecules with a higher concentration.

Dialyzer Effectiveness

Several points should be made in regard to the optimal performance of various types of dialyzers.

Metabolite clearance (MC) is a measure of the removal rate of waste products. The mathematical equation

$$MC = \frac{C_1 - C_2}{C_1} Q_B$$

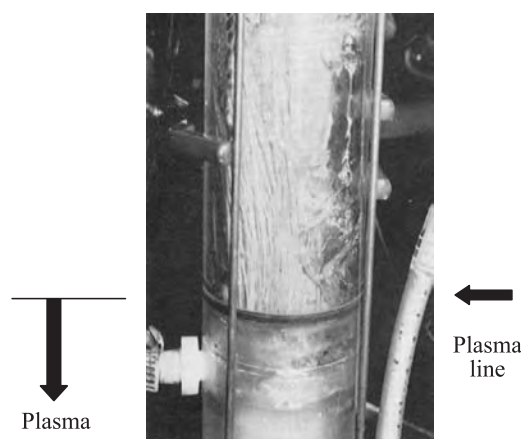
**Fig. 7** Monsanto hollow-fiber device evaluated as an experimental plasma separator in 1972. Separated plasma is seen at the bottom of the hollow-fiber column. (From Nosé, Y.; Malchesky, P.S. Therapeutic Membrane Plasmapheresis. In *Therapeutic Plasmapheresis*; Oda, T., Ed.; F.K. Schattauer Verlag: Stuttgart, 1981; 3.)

Table 3 Procedures in therapeutic apheresis

Procedure	Method
I. Plasma exchange (plasmapheresis for plasma separation)	Centrifugal method (manual intermittent, automated continuous) Membrane method (TMP of plasma flow controlled, rotating filter system)
II. On-line plasma fractionation	Membrane method
1. On-line membrane plasma fractionation	
a. Cascade filtration (double-filtration plasmapheresis)	
b. Cryofiltration	
c. Thermofiltration	
2. On-line plasma fractionation and plasma sorption by sorbents	Biological sorbent (protein A column, immunoabsorbent IM-TR, etc.) Nonbiological sorbent (resin, activated carbon, etc.)
3. Other precipitation methods	
III. Cytapheresis	Methods include centrifugal cell removal, cell removal filter, and photopheresis
a. Leukocytapheresis	
b. Plateletapheresis	
c. Erthrocytapheresis	
IV. Leukoplasmapheresis	Centrifugal method

Source: Ref. [2].

where C_1 = blood inlet concentration of waste product to the dialyzer, C_2 = blood outlet concentration of waste product from the dialyzer, and Q_B = blood flow rate (ml/min), shows that waste product removal depends on, or is a function of, the blood flow and the fractional change in waste product concentration as the blood passes through the dialyzer. The magnitude of these quantities depends on such physiological factors as blood pressure and blood vessel patency, as well as on physical factors such as blood film thickness and resistance, dialysate flow rate, and membrane form and permeability. Optimizing the metabolite clearance of any dialyzer requires manipulation of these factors

until a satisfactory waste removal rate is established (Table 2).

Another commonly used indicator of the efficiency of dialysis is dialysance. This is the molecular exchange ratio between the blood and dialysate. Dialysance increases when blood flow increases:

$$D = \frac{C_1 - C_2}{C_1 - C_D} Q_B$$

where D = dialysance (ml/min) and C_D = concentration in the dialysate. The difference between clearance and dialysance is evident when the dialysate is recirculated.

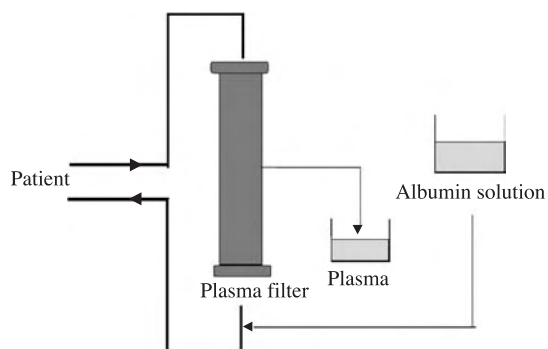


Fig. 8 Membrane plasma exchange. (From Ref. [2].)

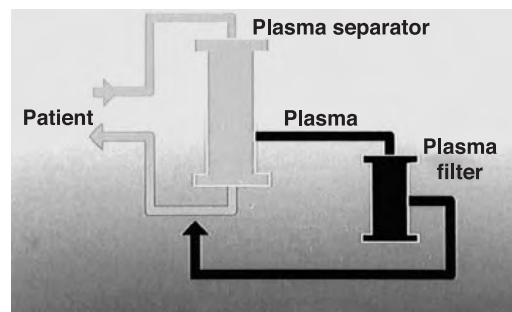


Fig. 9 Plasma filtration. (From Ref. [2].)



However, when the dialysate is used in only a single pass, these two values are the same.

In many cases, the variables themselves depend on other variables, and often compromises have to be made between opposing factors to achieve optimal efficiency of the dialyzer. Ultrafiltration (ml/hr) is just as important in many patients as waste product removal. Often, however, conditions for good ultrafiltration are in opposition to those for adequate waste product removal. For example, in a continuous-flow dialyzer, when a patient requires little or no water removal, the dialysate is generally pumped through the dialyzer. If more water removal is necessary, the dialysate is sucked through the dialyzer. Sucking creates a negative pressure in the dialysate ("negative pressure" refers to a pressure that is less than normal atmospheric pressure). The effect of this negative pressure is to increase the pressure difference between the blood and the dialysate, thereby increasing the rate of water removal. However, the same increase in pressure difference that increases ultrafiltration also increases the priming volume and blood film thickness and consequently reduces the waste metabolite removal. It is important to recognize that many of the physiological and physical variables affecting dialysis (waste removal) also affect ultrafiltration and vice versa. A dialyzer capable of a urea clearance of over 80 or 90 ml/min at a flow rate of 150 ml/min and ultrafiltration rate of over 150 ml/hr is clinically acceptable.

HEMOFILTRATION

Hemofiltration is a process for removing water and small molecules in the blood by the ultrafiltration principle (Fig. 5). The sizes of molecules removable by hemofiltration are dictated by the pore size of the membrane. Typically, hemofiltration is utilized to remove molecules smaller than albumin (and occasionally the same size as albumin) whose concentration in the blood is very small. Osmosis is not an effective removal method when the concentration of the molecules is very low. Typically, approximately 20 liters of filtrate are removed from the patient's blood while 17 liters of pyrogen-free sterile dialysate are infused at one session of the treatment. Then the low-concentration toxic substances are effectively removed. Hemodiafiltration is the combination of hemodialysis and hemofiltration. This combination therapy can take advantage of osmosis (dialysis) for the removal of molecules of higher concentration and filtration for the removal of molecules of lower concentration.

PLASMAPHERESIS

Separation of plasma from the cellular components of the blood is achieved by either the plasma separator

membrane or the centrifugal method. In this chapter, only membrane separation is described. A membrane with much larger pore size than is used for hemofiltration should be used for this purpose. For the effective removal of plasma, the transmembrane pressure should be kept as low as possible. The differences between dialysis, hemofiltration, and plasma separation are summarized in Fig. 6.

Membrane Plasma Separation and Plasma Treatments

By 1972, the development of membrane technology was such that hydrophobic membranes with a porosity of 0.2 to 0.5 μm were available and used initially for membrane oxygenators. I. O. Slayer together with Y. Nosé had developed the Monsanto hollow-fiber membrane oxygenator (Fig. 7). Nosé discovered that the polymeric hollow-fiber device showed reasonable gas exchange characteristics but also produced a large amount of plasma by leakage through the membrane walls.^[3] The membrane was not hydrophilic enough to be practical for oxygenators.

But Nosé thought that this type of hydrophilic membrane would be effective for plasma separation. Thus, the Cleveland Clinic group lead by Nosé and P. S. Malchesky devoted their efforts to the development of a membrane for plasma separator and plasma fractionator systems.^[3] Initially a plate-type plasma separator was utilized. In 1977, a hollow-fiber plasma separator became available from Asahi Medical Inc. in Japan. Z. Yamazaki and N. Inoue in Tokyo performed the first clinical application of this filter for hepatic failure patients. The CCF group initiated its clinical application using this filter the next year.^[3]

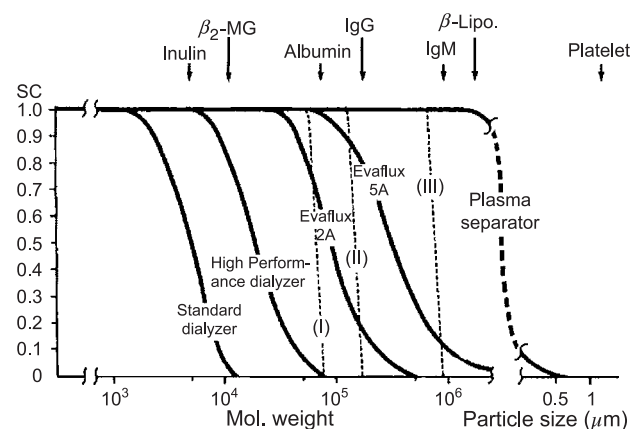


Fig. 10 Plasma fractionator for the removal of different sizes of macromolecules (Kuraray plasma fractionator). (From Ref. [2].)

Table 4 Physical characteristics of membrane fibers in plasmapheresis

Supplier or manufacturer	Filters	Membrane material: hollow fiber (F) or plate or pile (P)	Surface area (m ²)	Maximum pore size (μm)
<i>Plasma separator</i>				
Travenol Lab., USA	CPS-10	Polypropylene (F)	0.17	0.55
AB Gambro, USA	Fiber plasmafilter PP	Polypropylene (F)	0.38	0.5
	Fiber plasmafilter PC-1000	Polycarbonate (F)	0.27	NA
Cobe Lab, USA	TPE plasma separator	Polyvinylchloride (P)	0.13	0.6
Fresenius MTS, FRG	Plasmaflux P2	Polypropylene (F)	0.50	0.5
Dideco S.p.A., Italy	Hemaplex BT 900	Polypropylene (F)	0.20	0.5
Asahi Medical Japan	Plasmaflo AP-06-M	Cellulose diacetate (F)	0.65	0.2
	Plasmaflo AP-03-H	Cellulose diacetate (F)	0.30	0.2
	Plasmaflo AP-05-H	Cellulose diacetate (F)	0.50	0.2
	Plasmaflo AP-08-H	Cellulose diacetate (F)	0.80	0.2
Kuraray Co., Japan (Kawasumi Lab. Inc.)	Plasmacure type M	Polyvinylalcohol F	0.30	0.2
	Plasmacure type L	Polyvinylalcohol (F)	0.50	0.2
Toray Ind., Japan	Plasmax PS-05	Polymethylmethacrylate (F)	0.50	0.5
	Plasmax PS-02	Polymethylmethacrylate (F)	0.15	NA
Teijin, Japan	Plasma separator TP-50	Polymer alloy (F)	0.40	0.2
Mitsubishi, Japan	Plasma separator MPS 0250	Polyethylene (F)	0.50	NA
Terumo, Japan	PS-4000	Cellulose acetate (P)	0.50	0.45
Organon Teknika, USA	Curesis	Polypropylene (F)	0.12	0.65
Cordis Dow Inc.	Plasmafilter	Polyethylene (F)	0.70	NA
Nipro, Japan	Plasma separator PEX-50	Cellulose triacetate (F)	0.50	0.4
Kawasumi Lab. Inc., Japan	Plasma separator KPS02	Polypropylene (F)	0.20	0.6
	Plasma separator KPS03	Polypropylene (F)	0.30	0.6
	Plasma separator KPS04	Polypropylene (F)	0.40	0.6
<i>Plasma fractionator</i>				
Dideco, S.p.A., Italy	Albusave BT 902	Cellulose diacetate (F)	1.00	0.02
Enka, AG	PF prototype	Cellulose diacetate (F)	1.00	NA
Asahi Medical, Japan	Cascadeflo AC-1730	Cellulose diacetate (F)	1.70	NA
	Cascadeflo AC-1760	Cellulose diacetate (F)	1.70	NA
	Cascadeflow AC-1770	Cellulose diacetate (F)	1.70	NA
	Plasmaflo AP-06-M	Cellulose diacetate (F)	0.65	0.2
Kuraray Co., Japan	Eval filter type 2A	Ethylene vinyl alcohol (F)	1.00	0.01
	Eval filter type 3A	Ethylene vinyl alcohol (F)	NA	0.02
	Eval filter type 4A	Ethylene vinyl alcohol (F, T)	2.00	0.03
Toray Ind., Japan	Plasmax AS-08	Polymethylmethacrylate (F)	0.80	NA
	Plasmax ASC-08	Polymethylmethacrylate (F)	0.80	NA
Teijin, Japan	TA-100	Cellulose diacetate (F)	0.80	NA
	TA-200	Cellulose diacetate (F)	0.80	NA
Terumo, Japan	CF-01	Cellulose acetate (F)	1.0	NA

NA, data not available from company.

Source: Ref. [2].

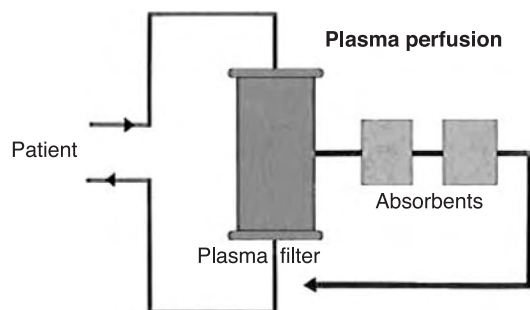


Fig. 11 Plasma adsorption. (From Asanuma, Y.; Malchesky, P.S.; Zawicki, I.; Smith, J.W.; Carey, W.D.; Ferguson, D.R.; Hermann, R.E.; Nosé, Y. Clinical hepatic support by on-line plasma treatment with multiple sorbents—Evaluation of system performance. *Trans. Am. Soc. Artif. Intern. Organs* 1980, 26, 400.)

The following plasma treatment methods were developed using these microporous membrane filters (Table 3).

Plasma exchange

Utilizing a plasma separator (either membrane or centrifuge) (Fig. 8), the diseased patient's plasma is separated on-line. At the same time, the healthy patient's plasma or albumin solution replaces the lost plasma volume.

Plasma filtration

After plasma is separated with a plasma separator, the diseased plasma is filtered through a second filter called

a plasma filter or plasma fractionator (Fig. 9). The molecules of unwanted size are removed by the secondary plasma fractionator module (Fig. 10). LDL cholesterol is a much larger molecule than HDL cholesterol; it is very easily removed by the proper plasma fractionator (Kuraray 5A) while HDL cholesterol is retained. To enhance the efficiency of plasma filtration, cryofiltration and thermofiltration (described in a later section) were introduced.

The plasma filter or plasma fractionator separates unwanted larger-size molecules inside the second filter and retains these larger molecules (Table 4). Filtered plasma without the unwanted molecules is returned to the patient.

Plasma adsorption

Specific adsorption columns can remove unwanted macromolecules in plasma (Fig. 11). This procedure is called plasma adsorption (Table 5). To enhance its efficiency, multiple adsorption columns are often necessary.

On-line membrane filtration with plasma treatment

Membrane plasma separators are very similar to capillary hemodialyzers in terms of results. Their basic operational constraints, however, are completely different. Membrane plasma separators are designed to separate the cellular components of the blood from the non-cellular components (Fig. 12). After separation, plasma can either be discarded (Fig. 8), or processed by sorbents (Fig. 11), secondary membranes (Fig. 9), or

Table 5 Adsorbents for therapeutic apheresis

Material	Target substance or disease
<i>Adsorbents for direct hemoperfusion</i>	
Activated charcoal	Drug intoxication
Amberlite XAD-4 resin	Drug intoxication
Polymixin B bound polystyrene fiber	Endotoxin
Polyacrylamide gel	LDL, LP(aq)
Porous cellulosic beads	β_2 -microglobulin amyloidosis
<i>Adsorbents for plasma perfusion</i>	
Styrene-divinylbenzene copolymer	Icterus, acute hepatic failure
Tryptophan fixed polyvinyl alcohol gel	MG, Guillain-Barré syndrome
Phenylalanine fixed cellulose fiber	SLE, antiphospholipid antibody syndrome
Acetylcholine receptor peptide immobilized resin	MG
Protein A column	Idiopathic thrombocytopenia, malignancy
Blood typing antigen fixed silica	A,B,O mismatch organ transplantation
<i>Adsorbents for blood cell adsorption</i>	
Polyester nonwoven fabric	IBD, RA, skin disease
Anti-CD4 + monoclonal antibody immobilized fabric	Autoimmune disease, multiple sclerosis

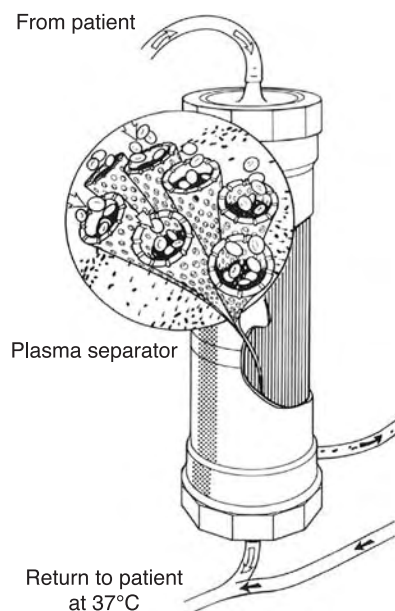


Fig. 12 Membrane plasma separators for plasma treatment. Blood is perfused through the lumens of the hollow fiber. Under a low hydraulic transmembrane pressure gradient, plasma is freed through the membrane wall, collected, and thrown away (in the discarded plasma treatment scheme). Blood cells are retained in the lumen and are not allowed to pass through the wall before being returned to the patient. (From Nosé, Y.; Horiuchi, T.; Malchesky, P.S.; Smith, J.W.; Matsubara, S.; Abe, Y. Therapeutic Cryogel Removal in Autoimmune Disease: What is Cryogel? In *Therapeutic Plasmapheresis (II)*; Oda, T., Ed.; F.K. Schattauer Verlag: New York, 1982; 15.)

cryofiltration. With on-line plasma treatment systems, we not only discard a large variety of macromolecules, which has metabolic or immunological consequences for the body, but also eliminate the need for infusion of foreign macromolecules from plasma products, which are potentially incompatible with the recipient.

On-line plasmapheresis is a treatment modality consisting of

1. Removal of whole blood from the patient
2. Separation of plasma from the cellular blood components via a porous membrane or a centrifugal process
3. Subjecting the separated plasma to specific procedures, such as cooling, filtration, and perfusion through columns
4. Reunion of the treated plasma with the cellular blood components and its return to the patient

Cascade membrane technology

If it is known that certain macromolecules should be removed, cascade membrane technology can be used

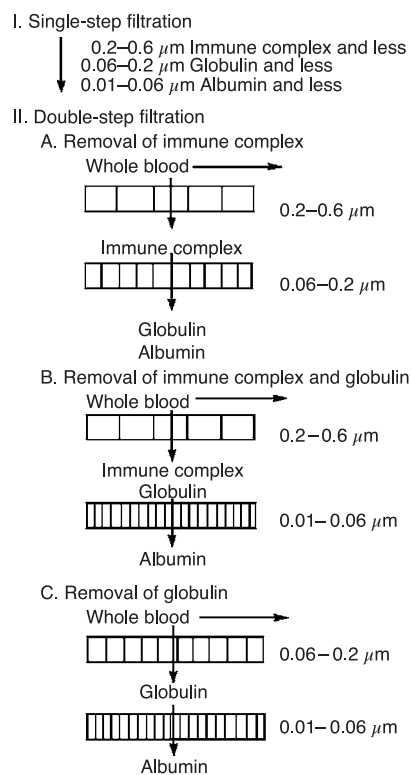


Fig. 13 Cascade membrane technology for selective removal of macromolecules. The general objective of this technology is selection of filter membranes of various porosities suitable for selective removal of macromolecules from blood. (From Nosé, Y.; Malchesky, P.S.; Smith, J.W. Hybrid artificial organs: Are they really necessary? *Artif. Organs* **1980**, 4(4), 285.)

(Fig. 13). Using this principle we can selectively separate molecules of a few million daltons molecular weight (e.g., immune complexes), of a few hundred thousand daltons molecular weight (e.g., gamma globulin fractions), or below 100,000 daltons molecular weight (e.g., albumin fractions). Aiming for the separation of these various types of macromolecules, several industrial firms have developed macromolecular filters of various porosities.

Cryofiltration

Selective separation of globulin fractions or larger molecules from albumin or smaller molecules is theoretically possible, but the molecular sizes of albumin and globulin are quite similar, and it is rather difficult to separate them effectively using simple membrane filtration systems. Because cryoprotein is very often present in the autoimmune disease state treated by plasmapheresis, an attempt was made to utilize temperature as a variable in the membrane separation of pathological macromolecules in the plasma (Fig. 14A). Cryoproteins

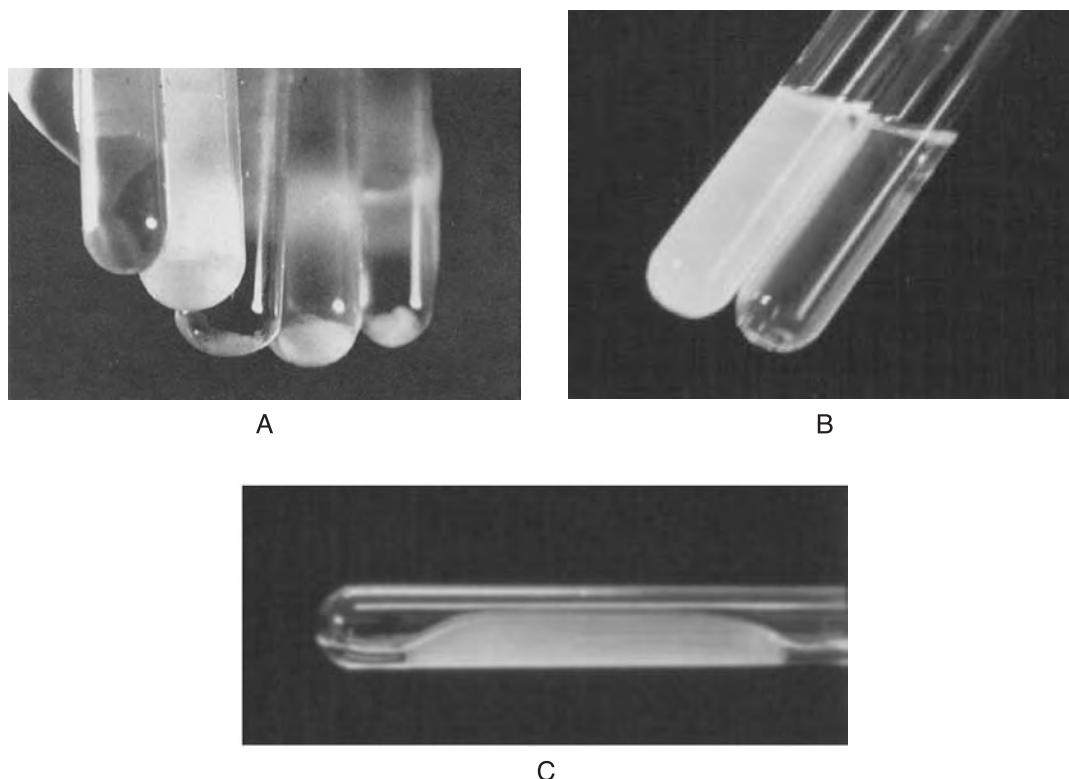


Fig. 14 (A) Cryoprecipitates formed from the serum of patients by cooling at 4°C for 72 hours. From the left: control, rapidly progressive glomerulonephritis, rheumatoid arthritis, myasthenia gravis, and Sezary syndrome. These cryoprecipitates are different from cryogel, which is formed in heparinized plasma. (B) Heparinized plasma (right) is cooled to below 35°C (typically near 0 degrees). Then the plasma changes to a milkylike solution, indicating the formation of cryogel (left). (C) After the cooled heparinized plasma is cryofiltered, the cryogel is formed inside the filter and removed from the plasma. When the cryogel is heated above 35°C, this gel changes its appearance to a liquid form. (From Nosé, Y.; Malchesky, P.S.; Smith, J.W. Problems of Current Technology for the Selective Removal of Pathological Macromolecules. In *Therapeutic Plasmapheresis III*, Oda, T., Ed.; Schattauer Verlag: Stuttgart, 1984; 3.)

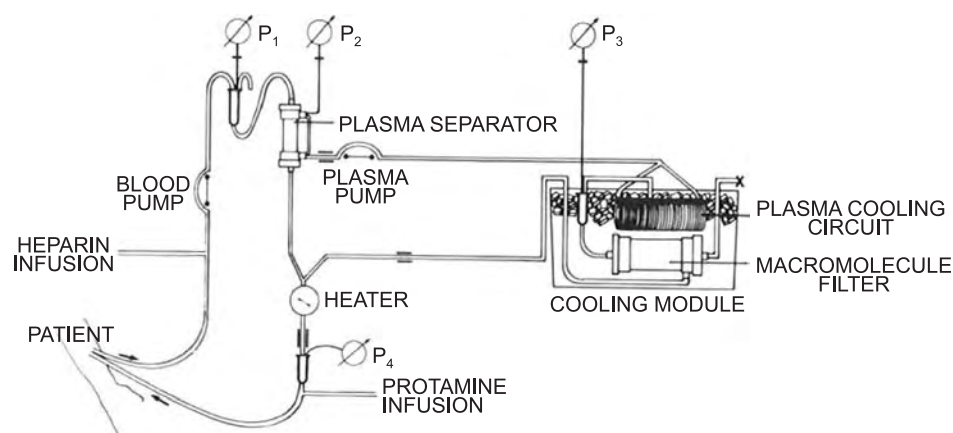


Fig. 15 The cryofiltration circuit. Cryogel is primarily formed as fibrinogen-heparin-fibronectin complexes. Antibodies, cryoprotein, immunocomplexes, pathological albumin, endotoxin, and other forms are also included. (From Malchesky, P.S.; Asanuma, Y.; Zawicki, I.; Blumenstein, M.; Calabrese, L.; Kyo, A.; Krakauer, R.; Nosé, Y. On-Line Separation of Macromolecules by Membrane Filtration with Cryogelation. In *Plasma Exchange*; Sieberth, H.G., Ed.; F.K. Schattauer-Verlag: Stuttgart, 1980; 133.)

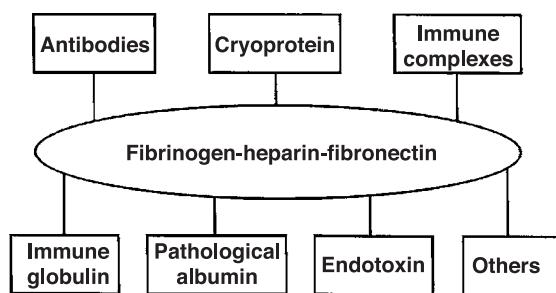


Fig. 16 Composition of cryogel. (From Nosé, Y. Congress presidential address: 5th WAA congress, therapeutic artificial organs, 10 years after. *Artif. Organs* **1995**, *19*(3), 204.)

are serum proteins that precipitate and gel spontaneously at a low temperature (near 4°C for three to five days) and return to a soluble form on elevation of the temperature. Cryoprecipitates also occur in heparinized plasma more dominantly and represent a homogeneous mixture of proteins, immunoglobulins, immune complexes, certain complements, lipids, and fibrinogen fractions. Particularly at a temperature below 35°C, heparin used as an anticoagulant for the extracorporeal circulation establishes fibrinogen–heparin–fibronectin complexes and attracts antibodies, immunocomplexes, and other forms inside of these complexes. This gel-type residual inside the secondary filter was coined “cryogel” by Y. Nosé^[3] (Fig. 14B,C). While normal values exist in nondisease states, concentrations of cryogel over 10 times higher are not uncommon in the autoimmune-diseased patient.

Cryofiltration was introduced by P. S. Malchesky and Y. Nosé, taking advantage of cryogel formation. Cryofiltration is the technique of cooling (possibly to near freezing) and filtering plasma on-line. Cryogel is formed and retained in the membrane filter. While distinctly different from cryoglobulins as measured by

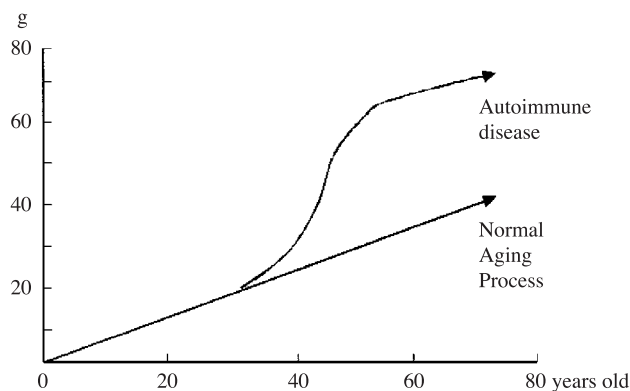


Fig. 17 Kinetics of cryogel in aged and diseased individuals. (From Nosé, Y. Congress presidential address: 5th WAA congress, therapeutic artificial organs, 10 years after. *Artif. Organs* **1995**, *19*(3), 204.)

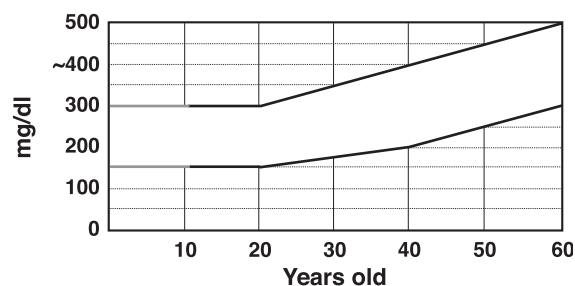


Fig. 18 Aging and increased level of fibrinogen in plasma. (From Nosé, Y. Congress presidential address: 5th WAA congress, therapeutic artificial organs, 10 years after. *Artif. Organs* **1995**, *19*(3), 204.)

B

classical assays, cryogel removal has been shown to be beneficial in the treatment of rheumatoid arthritis and other autoimmune diseases. A cryofiltration circuit is shown in Fig. 15. Blood is pumped from the patient into a plasma separator, which separates the plasma from the cellular components of the blood. The plasma is then pumped into a cooling module, where it is cooled and filtered on-line by means of a very porous macromolecular filter. The filtered plasma is reunited with the whole blood, warmed to physiological temperature, and returned to the patient. Cryogel removal is achieved with the bulk of the plasma returned to the patient accompanied by the passage of albumin and

Table 6 Commercialized leukocyte removal filters for blood transfusion

Manufacturer	Filter (material)
Asahi Medical (Tokyo, Japan)	Sepacell (polyester fiber)
Dideco (Mirandole, Italy)	Leukoseize (cellulose acetate fiber)
Fenwall (Glendale, CA, U.S.A.)	4C2423 (polyester wool)
Miramed (Mirandole, Italy)	Miopore (cellulose acetate fiber), Leukostop (polyester fiber)
NPBI (Emmercompascum, The Netherlands)	Cellselect (cellulose acetate fiber), Optima (polyester fiber)
Organon (Boxtel, The Netherlands)	Erypur (cellulose acetate fiber)
Pall (East Hill, NY, U.S.A.)	RC100 (polyester fiber), Ultipor SQ40S (polyester mesh)
Terumo (Tokyo, Japan)	Immugard IG500 (cotton wool)
Travenol (Thetford, U.K.)	Leukopak (nylon fibers)

Source: Ref. [1].

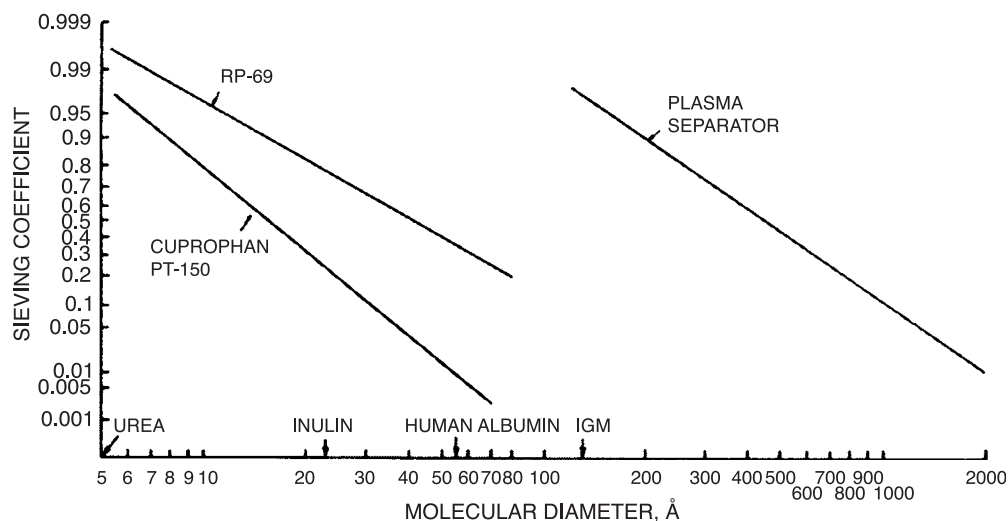


Fig. 19 Comparison of sieving coefficients as a function of molecular diameter for dialysis (Cuprophane PT-150), hemofiltration (RP-69), and plasma separator membrane modules. The sieving coefficient is the ratio of the concentration of the solute in the filtrate (C_F) to that in whole plasma (C_P) and is represented as $SC = C_F/C_P$. (From Ref. [2].)

low-molecular-weight plasma. The composition of cryogel is shown in Fig. 16.

Cryogel exists even in a healthy individual, and the level of cryogel increases with the aging process (Fig. 17). So removal of this cryogel might contribute to the maintenance of physical youth. This is explainable on a physiological level by the increase of fibrinogen during the aging process (Fig. 18).

Cryofiltration has been used for the treatment of malignant rheumatoid arthritis. The initial intensive treatment was approximately 30 sessions in six months. After one year, treatment from once a month to once every three months kept the syndrome under control.

Thermofiltration

For the removal of lipids such as LDL cholesterol, the formation of cryogel is not desirable. Thus, for the removal of LDL cholesterol, thermofiltration was developed by Y. Nosé and P. S. Malchesky, with the second filter's environmental temperature kept at around 37–40°C.^[3] Cryogel is formed when the heparinized plasma is cooled down below 35°C, so cryogel formation is prevented by this method, and effective removal of LDL cholesterol was established for the treatment of arteriosclerosis. A once-a-month session of thermofiltration should be sufficient to maintain the LDL cholesterol level within a reasonable range.

Sorbent systems including leukocyte filters

The utilization of chemical sorbents such as charcoal and resins has been investigated since 1963 for use in the

artificial kidney and for drug overdose. For biomedical applications of sorbents in the extracorporeal circulation, the system must be free from particle release and must be biocompatible. Direct perfusion of blood over sorbents (hemoperfusion) in clinical use has been hampered due to damage to blood cellular components and embolization. The microencapsulation of sorbents, the direct perfusion of sorbents with plasma, and the application of sorbent membranes have made the chronic application of sorbent devices clinically safe.

Recently, many specific adsorption columns have been developed (Table 5). The mechanisms include

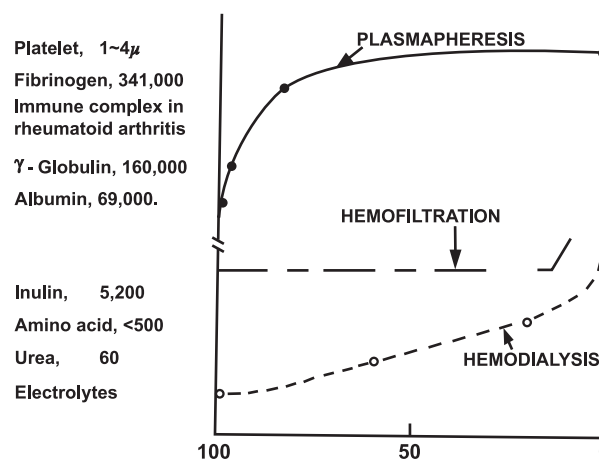


Fig. 20 Comparison of solute extraction for hemodialysis, hemofiltration, and membrane plasmapheresis, illustrating the effect of increasing molecular weight on solute removal. Only plasmapheresis provides the transfer of protein-bound toxins. For effective hepatic assist, it is necessary to remove molecules up to the size of albumin. (From Ref. [2].)

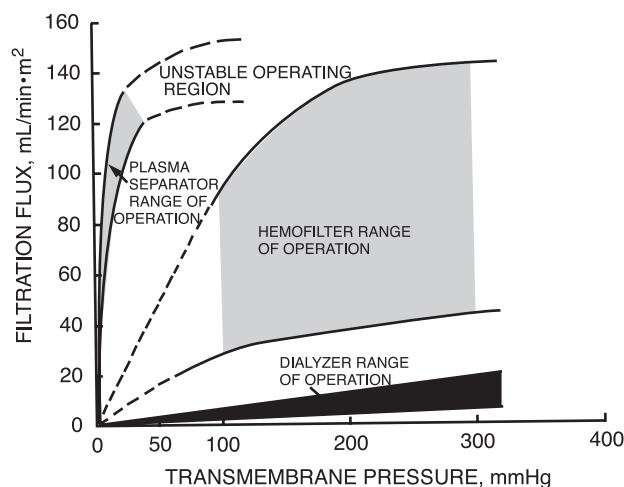


Fig. 21 Filtrate flux is represented as a function of the transmembrane pressure for dialysis, hemofiltration, and plasma separation. The shaded regions represent normal operating ranges. Filtration flux is equal to the number of milliliters per minute transferred per meter squared (ml/min/m²). (From Ref. [2].)

electrostatic, hydrophobic, immunosorption, and others. The FDA has approved some of them, but most are not allowed for use in the United States.

Similar to Polymixin B bound polystyrene fiber, removal of endotoxins by a specially designed adsorption column was developed for the treatment of septic patients and multi-organ failure patients. Various types of adsorbents are packed in the form of beads or granules inside the column. Also, adsorbents are adhered to the polymeric fibers and packed in the column.

The centrifugal separation method effectively separates not only plasma but also leukocytes and erythrocytes. Leukocyte removal is achieved, and thus leukocytapheresis results in not only plasma modulation but also cellular modulation.

Recent advances in apheresis technologies have introduced various types of cellular adsorption columns and cellular filters (Table 6). The effective removal of leukocytes is possible with such columns. The selective removal of granulocytes and lymphocytes is also becoming possible. Filter material used for leukocyte removal includes polymeric thin fibers, netted or woven fibers, and polymeric sponges.

Recent blood purification technologies are expanding from classical hemodialysis, hemofiltration, and plasma separation to selective removal of pathological plasma macromolecules and pathological cellular factors by specifically designed adsorption columns or adsorption filters. Theoretically, for some disease states to be treated, it is essential to provide not only macromolecular modulation but also cellular modulation. Recently, immunologically active molecules such as cytokines (molecular weight below

30,000 daltons) should be considered for removal in the treatment of some immunologically-induced diseases, even though this hypothesis still needs to be proven.

For the treatment of autoimmune diseases, it is necessary to produce an immunosuppressive state through the removal of immunologically active macromolecules and cells. Along with the treatment of immunologically deficient states, such as a malignant tumor or an infectious prone state, it is necessary to induce immunostimulation in the patient.

Immunostimulation can be established with bio-compatible columns or filters. Thus, plasma therapy could induce both immunosuppressive and immunostimulative states.

CONCLUSION

Dialysis is effective in treating renal failure because of its ability to remove solutes of low molecular weight (generally less than 5000 daltons), including water molecules. The dialysis procedure does, however, have limitations in the removal of solutes whose molecular weight is larger or that are protein bound. Such abnormal concentrations of solutes exist in many disease states. Hemofiltration, sorption, and plasma exchange have been applied as solute removal techniques in the removal of plasma solutes that cannot be removed by conventional dialysis (Figs. 19 and 20). Pore sizes in the membranes dictate molecular removal. Because pore sizes are different, transmembrane pressures required for molecular removal by hemodialysis, hemofiltration, and plasma separation are quite different (Fig. 21).

ARTICLE OF FURTHER INTEREST

Bioreactors

REFERENCES

1. Nosé, Y. *Manual on Artificial Organs. Volume 1: The Artificial Kidney*; CV Mosby: Saint Louis, MO, 1969; pp. 13, 19, 182.
2. Nosé, Y.; Kambic, H.E.; Ichikawa, S. *The ICMT Publication on Artificial Organs. Volume 1: Blood Purification—Artificial Kidney and Plasmapheresis—Past, Present and Future*, 2nd Ed.; ICAOT/ICMT Press: Houston, TX, 2001; . pp. 128, 133, 134, 126, 127, 187, 192.
3. Nosé, Y. Millennium memorabilia issue on apheresis: Part 1, membrane therapeutic apheresis. *Therap. Apher.* **2000**, *4* (1), 1–68.
4. Murabayashi, S. Millennium memorabilia issue on apheresis: Part 2, centrifugal devices, adsorbents, leukocyte filters, and extracorporeal immunomodulations. *Therap. Apher.* **2000**, *4* (2), 71–182.

Blood Vessel Mechanics

Ruth J. Okamoto

Departments of Mechanical Engineering and Biomedical Engineering, Washington University, St. Louis, Missouri, U.S.A.

INTRODUCTION

The primary role of the cardiovascular system is to distribute oxygenated blood to organs and tissues. The vascular system (vasculature) is the distribution system and includes arteries, arterioles, capillaries, venules, and veins. Oxygen exchange occurs in the capillaries, while arteries and veins serve as conduits and to regulate blood flow. While blood vessels share common structural features, there is a wide variation in the structural details of vessels located in different parts of the vascular tree, because of differences in blood pressure and blood flow. Thus, while common governing principles can be delineated for blood vessels, the structure and mechanical properties of specific blood vessels must be investigated individually.

This article begins with a description of the structural features of blood vessels. Experimental methods for studying blood vessel mechanics are described and their strengths and limitations identified. Finite elasticity is presented as a theoretical framework for interpreting experimental observations. Differences between blood vessel mechanics are examined in light of differences in the mechanical environment during development and after maturity. The ability of blood vessels to respond to their mechanical environment is discussed, and examples of normal and pathologic responses to changes in mechanical loads are provided. The article concludes with comments on current and future directions in understanding blood vessel mechanics, in particular, developing a more detailed understanding of the relationships between the biochemical and mechanical environments. Because of the brevity of this article, it cannot provide a complete literature review, but instead highlights a few of the many papers written on this topic, and the reader is referred to lengthier review articles where appropriate.

BLOOD VESSEL SIZE, STRUCTURE, AND FUNCTION

This section briefly summarizes the relationship between blood vessel size, location, and structure; the

work of Rhodin^[1] provides an excellent in-depth description. Two principles govern blood vessel dimensions: (1) the inner (lumen) diameter of vessels decreases as their flow decreases and (2) the ratio of wall thickness to lumen diameter in large and medium vessels is proportional to the transmural (internal minus external) pressure. The drop from arterial to venous pressure is primarily due to the resistance in the small arteries and arterioles and many tiny capillaries.^[2] Venules and veins, which return de-oxygenated blood to the heart, are under lower pressure than arteries and have a much smaller wall thickness to diameter ratio, as shown in Fig. 1.

All blood vessels are lined with a layer of endothelial cells, have a medial (middle) layer of smooth muscle cells in an extracellular matrix of collagen and elastin, and have an adventitial (outer layer) of fibroblasts in a primarily collagenous extracellular matrix. The endothelial cells serve an important function as mechano-sensors and regulators of wall shear stress because of blood flow.^[3] In large arteries, the media is structured into alternating layers of elastic fibers and smooth muscle cells called lamellar units.^[4] The smooth muscle cells are elongated and appear to be oriented predominantly in circumferential direction, although the orientation may change through the wall thickness in larger mammals. In smaller muscular arteries, the media is predominantly composed of well-organized layers of smooth muscle cells with few elastic fibers interspersed between them, and the lamellar units are no longer evident.

Capillaries are composed of a single layer of endothelial cells that is integrated into the surrounding tissue. The media of venules and veins are also composed primarily of smooth muscle cells; however, the orientation of the smooth muscle cells can be either circumferential or axial. Fig. 1 graphically shows the relative proportions of endothelial cells, elastic fibers, and collagen and smooth muscle cells under each type of blood vessel. The consequences of these structural differences on blood vessel mechanics have been illustrated by a variety of experimental methods, described in the following section.

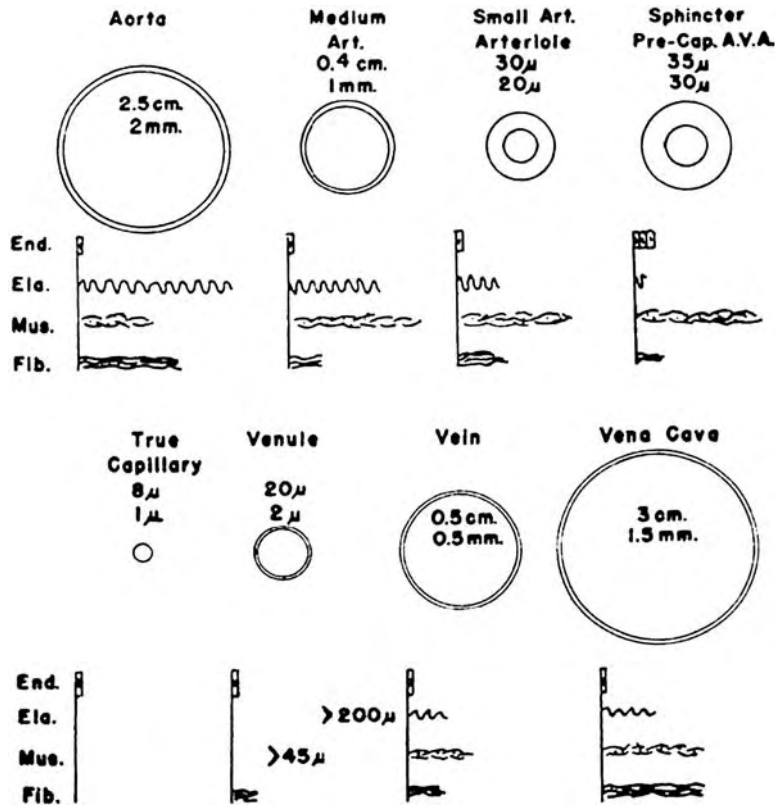


Fig. 1 The relative sizes of human blood vessels. The numbers above each vessel type show their lumen diameter and wall thickness. The bars below each vessel type show the relative proportion of endothelium (End.), elastic fibers (Ela.), smooth muscle cells (Mus.), and collagen (Fib.). The true capillary contains only endothelial cells and is not discussed in the text. Source: Reproduced with permission from The American Physiological Society (see Ref. [52]).

EXPERIMENTAL METHODS

In Vivo Studies

In vivo studies give insight into the mechanics of blood vessels undergoing physiologic loading and deformation and generally use non-invasive imaging modalities. McVeigh provides a useful review.^[5] Ultrasound is used for real-time measurement of vessel dimension changes during the cardiac cycle. Magnetic resonance imaging currently has more limited temporal resolution but is useful for reconstruction of dynamic vessel geometry. Computed tomography (CT) scans provide high-quality spatial resolution, but currently, they do not yet have the temporal resolution to measure deformation during the cardiac cycle.^[6] Pulse wave velocity (PWV) is used to infer vessel stiffness within a finite length of the vessel. The modified Moens-Korteweg equation^[7] relates the PWV to the elastic modulus in the circumferential direction (E_c):

$$PWV = \sqrt{\frac{E_c h}{2R_i \rho}} \quad (1)$$

where h is the vessel wall thickness, R_i is the vessel lumen radius, and ρ is the density of blood (1.05 g/cm³). This equation assumes that the vessel is thin walled (i.e., $h \ll R_i$), straight, and linearly elastic

within the range of physiologic pressures. Vessel lumen radius and wall thickness must be measured non-invasively or estimated from post-mortem measurements of excised tissue. The velocity is usually determined by calculating the time it takes for a flow or pressure wave to pass two locations. A finite distance separates these locations, so E_c calculated from this equation is an average modulus.

A second widely calculated index is vessel compliance (distensibility). Several equations are given in the literature, but all rely on the basic definition of compliance as a change in vessel volume during the cardiac cycle divided by the pulse pressure (the difference between the systolic and diastolic pressure). In most vessels, with the exception of the ascending aorta, vessel length does not change appreciably within a cardiac cycle, so compliance is defined with respect to either an area change or diameter change,^[8] e.g.:

$$C = \frac{d_{sys} - d_{dias}}{d_{dias}(P_{sys} - P_{dias})} \quad (2)$$

where d_{sys} , P_{sys} , d_{dias} , and P_{dias} are the measured luminal diameters and pressures at systole and diastole, respectively. Invasive pressure measurements (e.g., with a catheter) should be made in the same location as diameter measurements, but compliance in central arteries is often calculated from the more easily measured

brachial artery pressures. Compliance does not require the measurement of vessel wall thickness, but it should be emphasized that it includes both structural and material properties that cannot be decoupled. For instance, two rubber tubes of the same material and inner diameter, but different wall thicknesses, will have different values of compliance.

The circumferential wall tension, T_c , can also be estimated from *in vivo* measurements of transmural pressure, P , and the corresponding lumen radius R_i . If the wall thickness, h , is also known, the average circumferential wall stress, $\bar{\sigma}_c$, can also be estimated. For a thin-walled artery (i.e., $h \ll R_i$), the relationships are:

$$T_c = PR_i; \quad \bar{\sigma}_c = \frac{T_c}{h} = \frac{PR_i}{h} \quad (3)$$

These equations, both known as the Law of Laplace, are used to compare wall tension and average circumferential wall stress between different arteries or species. Unlike PWV or compliance, T_c and $\bar{\sigma}_c$ do not depend on vessel stiffness, but still require accurate measurements or estimates of P , R_i , and h at each pressure of interest.

In Vitro Studies

When a blood vessel is dissected free of its surrounding tissue, it usually retracts (shortens) in the axial direction. The ratio of the *in situ*, or *in vivo*, length to the retracted or *in vitro* length of a vessel segment is called the *in vivo* axial stretch ratio. Excised tissue used for *in vitro* tests may include tubular vessel segments, vessel rings, vessel strips, or rectangular segments.

During *in vitro* studies, the excised tissue is usually bathed in a physiologic salt solution (PSS). The exact composition of the PSS varies from laboratory to laboratory. Humphrey gives a concise summary.^[8] The composition of the PSS can be modified to activate the smooth muscle cells, or conversely, to render the vessel segment passive (i.e., no contractile force is generated by smooth muscle cells). To obtain active properties, the bathing solution is modified by increasing the calcium concentration and adding substances that activate smooth muscle, such as potassium chloride or norepinephrine. If active properties are being studied, it is important to keep the bathing solution near physiologic temperatures, as smooth muscle activation is temperature dependent. To obtain passive properties, the PSS is modified to omit calcium and add relaxing agents, such as papaverine or sodium nitroprusside. The difference between the response of the activated vessel segment and the passive response is considered the active contribution of the smooth muscle cells.

Uniaxial Tests

Rings and strips of isolated blood vessel tissue are used for uniaxial tests in which only one direction is subject to mechanical loads. Uniaxial tests are the simplest to set up, but provide limited information about possible differences in the mechanical response in the axial and circumferential directions and may overestimate or underestimate passive properties.^[9] Furthermore, because only one direction has an applied load, they do not measure any coupling between the responses in the axial and circumferential directions. Nonetheless, they have been useful in delineating the active response of vascular smooth muscle. The active response can be measured isometrically, that is, the specimen is incrementally stretched to specified lengths and the active force is measured after an equilibration period. Alternatively, specimens can be tested isotonicity, where the specimen is stretched to specified force levels and the length is adjusted to maintain that force. Rings and strips are also useful for studying viscoelastic properties^[10] where the simple setup facilitates studies over a wide range of frequencies.

Biaxial and Shear Tests

Mechanical tests on excised vessel segments using computer control provide a more complex loading environment.^[11] An excised blood vessel is placed in a bath of PSS, and the ends of the vessels are secured over hollow cannulae. The vessel is stretched in the axial direction to the *in vivo* axial stretch ratio. The vessel is then inflated over a range of pressures while measuring the increase in outer diameter and axial force. With appropriate measurements of vessel dimensions, this data can be used to calculate compliance or intrinsic mechanical properties of blood vessels. In addition, by changing the axial stretch ratio of the vessel before or during inflation, a wide range of loading conditions can be investigated. These vessel segment test systems are thus biaxial (control loading in both the circumferential and axial directions). In some systems, a twist can be applied to one cannula and the applied torque measured to obtain estimates of the response to axial shearing.^[12,13] While more expensive and complex than uniaxial test systems, these systems have been extremely useful in delineating the intrinsic mechanical properties of a number of blood vessels.^[14,15] Note that not all tests of excised tubular segments of blood vessels are biaxial. Often, the axial stretch ratio and axial force are not reported, and only a pressure–diameter curve is obtained.^[16,17]

Tests of planar biaxial specimens (rectangular or square) are less common than tests of vessel segments, primarily because of vessel geometry. In some vessels,

notably the ascending aorta and pulmonary artery, the vessel segments are quite short compared to the vessel diameter, and planar biaxial testing has been used to delineate mechanical properties^[18,19] of these vessels.

Testing of Vessel Components and Layers

Early experiments on vessel segments delineated the contributions of elastin and collagen to blood vessel elasticity by the addition of elastase or collagenase to selectively remove the elastin or collagen from the vessel wall. These studies^[20,21] established that elastase treatment resulted in an immediate and irreversible dilation of the vessel from its intact, unloaded state, but the vessel was still able to withstand large transmural pressures. Collagenase had little effect on the vessel's deformation at low transmural pressures, but resulted in vessel rupture at much lower transmural pressures than normal, suggesting collagen is important in maintaining the structural integrity of blood vessels.

Experimental methods have also been developed to test the medial and adventitial layers of blood vessels separately. Usually, the vessel is tested in its intact configuration, followed by removal of the adventitial layer by dissection. If the vessel is inverted, or if strips are used, it is also possible to remove the media and test the adventitial layer of some vessels.^[22,23]

Opening Angle Measurements

Intact, but unloaded, blood vessels are in a state of circumferential residual strain.^[24] Rings isolated from vessel segments often spring open when cut radially (Figs. 2A and B). This phenomenon is characterized by an opening angle α (Fig. 2B) and can be used to quantify the circumferential residual strain. It is possible, although technically difficult, to separate the layers of the blood vessel wall before measuring the opening angle in some vessels. The opening angle of each layer may be different, suggesting that even a single radial cut does not relieve residual strain throughout the vessel wall. The implications of circumferential residual strain are discussed in the framework of finite elasticity below.

Morphologic Measurements

After testing, many investigators preserve the test specimen for morphological measurements, either by freezing or formalin fixation. Preserved specimens can be sectioned and stained to identify histologic features, including the thickness of the intimal, medial, and adventitial layers. If preserved with formalin, blood

vessels may be fixed with a transmural pressure equal to the physiologic pressure or at zero pressure and at the *in vivo* length or the unloaded length. These conditions will affect the measured wall thickness and should be noted when using morphologic data. If vessels are fixed under physiologic pressure, incompressibility is used to determine the wall thickness in the unloaded configuration. When computerized measurements are made from stained tissue sections, the luminal area and wall cross-sectional area may be reported. Fixation and embedding can cause tissue shrinkage, which may underestimate tissue dimensions, particularly wall thickness, although the effect can be minimized by a careful choice of fixative and embedding agents.^[25]

In Situ Studies

In situ studies allow the investigator to manipulate mechanical loads over a wider range than *in vivo* studies, as well as alter the biochemical environment by changing the fluid both in and surrounding the blood vessel. In these tests,^[26,27] the vessel is isolated from the circulation. This allows the vessel to be inflated from zero pressure to supraphysiologic pressures under controlled loading conditions. The advantage of these tests is that external structures supporting the vessel remain intact, while the disadvantage is that the axial stretch ratio cannot be independently controlled.

THEORETICAL FRAMEWORK

In vivo, *in vitro*, and *in situ* measurements of blood vessel mechanics are interpreted within a theoretical framework. In many cases, that framework has been linear, isotropic elasticity, a branch of continuum mechanics. The continuum hypothesis assumes that the macroscopic response of a body can be described by locally averaged quantities. When applied to solid mechanics, the quantities of interest are the components of stress and strain.^[24] For a linearly elastic material, the relationship between stress and strain is linear and without time dependence. When the material has no preferred direction, it is isotropic, and the relationship between stress and strain components is described by a single elastic modulus. However, the definitions of stress and strain used in linear elasticity assume that the displacements of the body are small relative to its size, an assumption that is generally not true for blood vessels. An alternative formulation is incremental elasticity,^[28] in which non-linear behavior is linearized around an operating point, such as the mean arterial blood pressure and *in vivo* axial stretch ratio.^[29] The simplicity of incremental linear elasticity

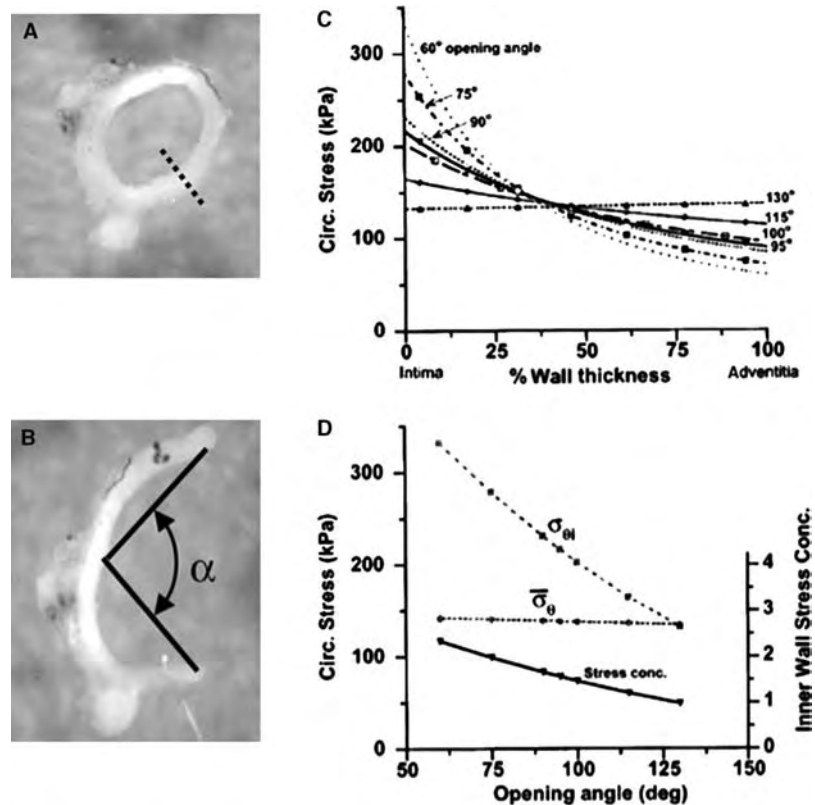


Fig. 2 Opening angle and effect on transmural stress distribution (A) A closed ring of mouse descending thoracic aorta. A radial cut is made at the dotted line. (B) The open ring, showing measurement of the opening angle, α , as defined by Fung. (C) The effect of opening angle on the transmural distribution of circumferential stress in the rabbit carotid artery. Larger opening angles lead to more even distribution of circumferential stress at systolic pressure (120 mmHg). (D) Circumferential stress concentration, defined as the circumferential stress at the inner wall divided by the average circumferential stress as a function of opening angle. Panels (C) and (D) Source: Reproduced with permission from American Society of Mechanical Engineers (see Ref. [35]).

is appealing, but its application is limited to small changes in deformations and loads near the linearization point.

Non-linear continuum mechanics, in particular finite elasticity, provides a more general framework for interpreting mechanical test data for blood vessels, as described in texts by Fung^[24] and Humphrey.^[8] While this framework comes with the cost of added experimental and analytical complexity, the availability of computational methods for solving finite elasticity problems^[30,31] has made the framework more accessible.

Finite elasticity assumes that the blood vessel undergoes substantial deformations from a stress free to a loaded state, such that the dimensions of the blood vessel in the loaded state have changed substantially and thus introduce non-linear geometric relationships. The stresses in the elastic material are related by a non-linear function to the strains. This function is known as the constitutive relation.

Pseudo-Elasticity

Fung and co-workers observed that the response of blood vessels to mechanical load is viscoelastic, but has a weak dependence upon strain rate over several orders of magnitude, and introduced the concept of pseudo-elasticity to describe the viscoelastic behavior of blood vessels in a more tractable form.^[32] A pseudo-elastic

material is an elastic material in which the constants in the constitutive relation are different in loading and unloading. This concept has been widely employed in blood vessel mechanics within the framework of finite elasticity.

The constitutive relation contains parameters that are determined from experimental data. In the case of in vitro tests of vessel segments, the measured pressure, force, diameter, and applied axial stretch can be converted to stress and strain and the experimental data used to estimate parameters in constitutive relations. These parameters can then be used in computational models. Biaxial data has been used to fit parameters in several types of constitutive relationships, most commonly using a hyperelastic strain energy function, in which stresses are obtained from the derivatives of the strain energy with respect to strain or other measures of deformation. While pseudo-elasticity has been broadly adopted in the biomechanics community, some models also include viscoelastic effects.^[33] A recent review by Vito and Dixon^[34] describes many of the constitutive models proposed for blood vessels in the past decade.

Residual Strain and Layered Models

Including the circumferential residual strain in wall stress calculations provides a more uniform distribution

of wall stresses through the wall thickness (Fig. 2C). As the opening angle increases, the circumferential stress concentration (i.e., inner wall stress divided by mean wall stress) may disappear (Fig. 2D). Models that individually describe properties of each layer of the blood vessel wall also affect the distribution of wall stresses.^[8,35,36] These models require experimental measurements to determine the residual strains and constitutive relations for each layer, as well as morphologic measurements of their relative thickness.

Models of Blood Vessel Growth and Remodeling

In the past two decades, interest has focused on how blood vessels adapt to changes in mechanical loads during both development and adulthood. Two principles of adaptation (remodeling) have been identified: normalization of wall shear stress and normalization of circumferential and axial wall stress. Shear stress on the inner surface of the blood vessel wall, τ_w , can be expressed as^[3]:

$$\tau_w = \frac{32\mu Q}{\pi D_i^3} \quad (4)$$

where μ is the blood viscosity, Q is the volume flow rate, and D_i is the inner (lumen) diameter of the artery. Since flow in large arteries is pulsatile, the shear stress varies within each cardiac cycle. Typical (homeostatic) values of time-averaged τ_w are approximately 15 dyne/cm² and are similar between arteries within an animal and between different animal species.^[3] Average circumferential stress ($\bar{\sigma}_c$) is given by Eq. 3, while estimates of local values of circumferential and axial wall stress require a constitutive relationship, as described above. Homeostatic values of $\bar{\sigma}_c$ are approximately 200–300 kPa in elastic arteries.^[36]

Liu and Fung^[37] propose indicial response functions to describe remodeling, where the time course of tissue growth or degradation after a sudden (step) change in a quantity, such as pressure, is measured and then used to describe changes in that quantity after an arbitrary change in the input quantity. Taber and co-workers^[36] developed a theoretical framework to explain blood vessel growth as a response to altered local stresses. It was recognized that growth and remodeling involved cellular response to altered stress or strain, and that this could trigger local biochemical processes, such as protein synthesis or degradation. To explain the consequences of this mechano-transduction, Humphrey and co-workers^[38] developed constrained mixture models that describe remodeling of individual constituents (i.e., collagen, elastin, and smooth muscle cells), where the rates of

growth and degradation can vary between components. An important assumption is that the tissue stress is the sum of the component stresses multiplied by their mass fractions, but the zero-stress state of each constituent can be different, and this zero-stress configuration can change as the constituent grows or degrades.

ELASTIC ARTERIES

The large elastic arteries include the aorta, carotid, larger branches of the aorta, and pulmonary arteries. The distinction between elastic and muscular arteries is based on the histologic observation of well-organized medial lamellae,^[39] rather than on size, although the larger conduit arteries tend to be elastic and the smaller ones muscular. Of the large elastic arteries, the mechanics of aorta and carotid have been the most intensively studied. The ascending aorta, which carries oxygenated blood from the left ventricle, originates at the aortic valve. Its elasticity helps reduce the pulsatility of downstream flow. Approximately 25% of the aortic flow leaves through blood vessels, branching from the aortic arch, that provide blood flow to the head and upper limbs,^[40] and a further 50% of the blood flow leaves through side branches in the descending thoracic and abdominal aortas. The remaining 25% flows from the distal end of the aorta into the left and right femoral arteries. Thus, the diameter of the aorta decreases along its length, with concurrent changes in the number of lamellar units and organization of elastic fibers. From a mechanics point of view, the aorta is not a single structure, and thus each region has been studied separately.

Passive Properties

When a segment of an elastic artery is inflated while held at its *in vivo* length, the diameter change with pressure is non-linear and may appear as a reverse S-shape (Fig. 3). The slope of the pressure–diameter curve appears slightly steeper at low transmural pressures, becomes more compliant, and then becomes stiffer again at or above the physiologic blood pressure range. The non-linearity of the pressure–diameter relationship is often overlooked, and a single value compliance or stiffness is calculated for a vessel over the entire physiologic pressure range,^[41] or an incremental circumferential elastic modulus is calculated as shown in Eq. 3. When biaxial data are fit to anisotropic constitutive relationships, the parameter fits indicate that arteries are nonlinear and anisotropic, although the direction of anisotropy is not consistent for all arteries. Although the pulmonary arteries are

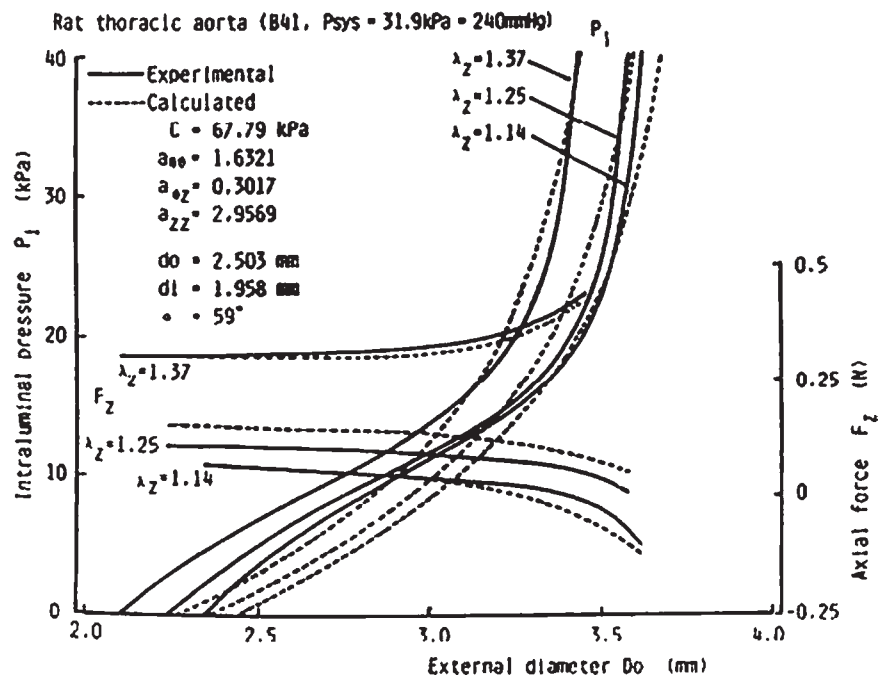


Fig. 3 Internal pressure and axial force for rat carotid artery as a function of the deformed outer radius for different axial stretch ratios ($\lambda_z = 1.14, 1.25, 1.37$). The artery was obtained from an 18-week old Wistar rat eight weeks after inducing hypertension with a renal clip. Source: Reproduced with permission from The American Physiological Society (see Ref. [14]).

considered elastic arteries, the pressure range in the pulmonary arterial circulation (10–20 mmHg) is much lower than in the systemic circulation (80–120 mmHg) in mammals. Hence, pulmonary arteries have a lower wall thickness to diameter ratio and are more compliant than the aorta, although their radii are similar.^[2] Biaxial tests of passive canine pulmonary arteries show their mechanical response is nearly linear under physiologic loads.^[18]

Active Properties

Elastic arteries held at a constant pressure contract in response to smooth muscle activation (Fig. 4). The amount of contraction is a function of the concentration of the activating agent, as well as the composition and arrangement of the smooth muscle cells in the artery studied. The active response of the carotid artery has been widely studied.^[17,42] When an activated segment contracts at a constant length, the active smooth muscle contribution increases with circumferential strain. When an activated artery is slowly inflated and deflated, a similar shift in the pressure–diameter curve is seen. The maximum active response appears to occur near the physiologic range of blood pressures. If the contractile response (CR) is defined as $CR = (D_a - D_p) / D_p$, where D_a and D_p are the measured outer diameters of the carotid under activated and passive conditions at a given transmural pressure, the maximum CR for the carotid artery in Fig. 4 is approximately 20% at 100 mmHg.

MUSCULAR ARTERIES AND ARTERIOLES

Segments of muscular arteries (e.g., brachial, renal, mesenteric, iliac, or coronary arteries) also exhibit a non-linear pressure diameter relation when inflated at their in vivo length under passive conditions (Fig. 4). Most muscular arteries contract circumferentially to a larger degree than elastic arteries when the smooth muscle cells are stimulated with potassium chloride^[17] or in response to other vasoactive substances, although the coronary artery may be an exception. While fewer studies have been conducted on the multiaxial properties of passive muscular arteries, there has been a recent interest in the passive properties of the individual layers of coronary arteries. In strips from aged human coronary arteries, the media appears less stiff than the adventitia, and the thickened intima contributed to the mechanical response.^[23] In contrast, inflation of tubular segments of porcine coronary arteries at their in vivo length indicated that the medial layer was stiffer than the adventitial layer, and the thin intima could not be separated from the media.^[22] These studies illustrate the diversity of mechanical responses in similar arteries, likely due to age and species differences.

Arterioles are small arteries having nerves and at least one layer of smooth muscle cells and ranging in diameter from 10 to 125 μm .^[7] As regulators of regional blood flow, their active properties are of primary interest to physiologists. Under basal (non-stimulated) conditions, isolated segments of arterioles exhibit non-linear pressure–diameter relations similar to larger arteries.^[24] When stimulated with potassium

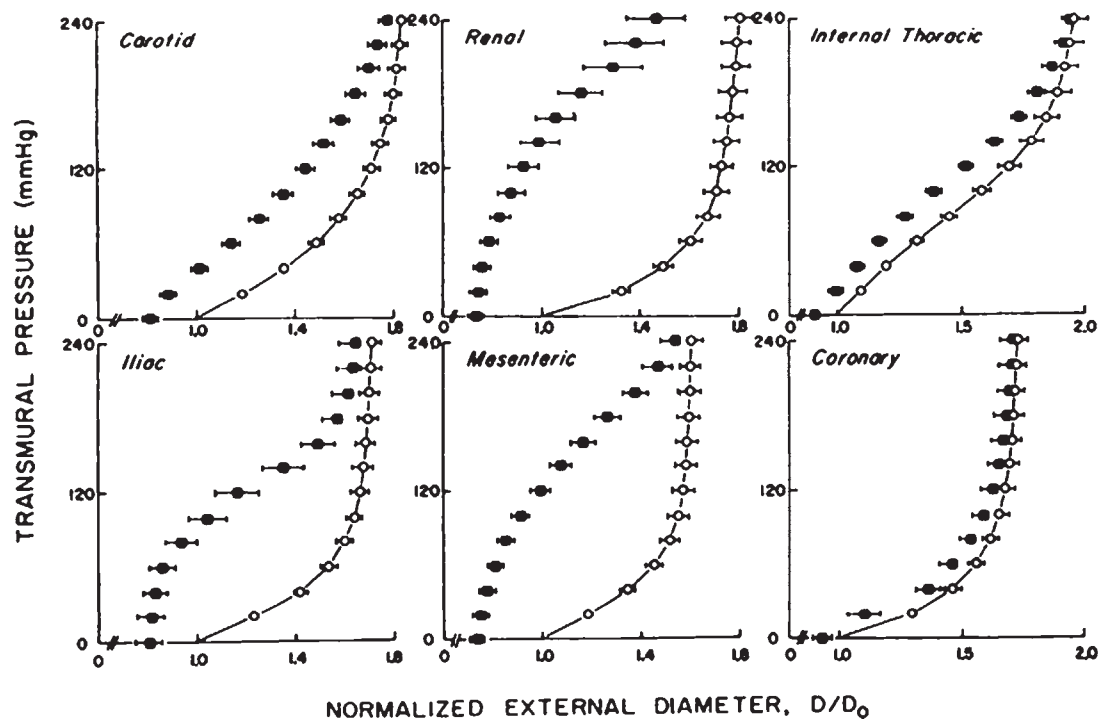


Fig. 4 Pressure vs normalized diameter curves from elastic (carotid and internal thoracic) and muscular (iliac, renal, mesenteric, and coronary) arteries. Open symbols: vessels tested under passive conditions; Closed symbols: vessels tested after stimulation with 120 mM KCl. Source: Reproduced with permission from The American Physiological Society (see Ref. [17]).

chloride and norepinephrine, arterioles contract substantially.^[43]

VEINS

Venous pressure is much lower than arterial pressure, and the ratio of vessel diameter to wall thickness is larger in veins, as illustrated in Fig. 1. A consequence of this structural difference is that veins may collapse when the external pressure is lower than the internal pressure in a vein. The major component of the venous wall is the smooth muscle cells, which may be oriented either circumferentially or axially, depending upon the vein. Venous volume can be changed substantially by smooth muscle activation or relaxation.^[24] A second structural difference is the presence of one-way valves in many veins, which serve to prevent blood from pooling in the lower limbs when standing. Mechanical properties of veins have been studied far less frequently than arteries, and little biaxial test data is available. Canine jugular veins and human saphenous veins exhibit both non-linear diameter–pressure and force–pressure relations, but this non-linearity is evident at much lower pressures in the canine jugular vein.^[44]

BLOOD VESSEL DEVELOPMENT, ADAPTATION, AND AGING

In mammals, the arrangement of the major blood vessels occurs during fetal development. However, the fetal circulation differs greatly from the post-natal period.^[2] At birth, when the pulmonary system begins to function and umbilical circulation ceases, blood pressure rises rapidly in the systemic circulation (i.e., aorta) and drops in the pulmonary circulation. This is reflected in a rapid accumulation of extra cellular matrix proteins in the aorta after birth and a decrease in those proteins in the pulmonary artery. Mean arterial blood pressure continues to rise slowly until adulthood. As body mass increases, blood vessel size and wall thickness increase to accommodate increased flow and pressure. The production of elastin by vascular smooth muscle cells appears to end at maturity, except under wound healing conditions.^[8]

The lumen diameters of arteries during development and maturity have been shown to normalize shear stress. As blood flow leaves the aorta through side branches. For example, the diameter of the aorta decreases, keeping τ_w at the homeostatic value.

As Eq. 4 implies, an increase in blood flow through the vessel lumen, without an increase in lumen diameter,

also increases τ_w . Elastic arteries respond to a sudden increase in blood flow with endothelial cells producing nitric oxide to induce vasorelaxation and increasing lumen diameter. This is followed by a further slow increase in lumen diameter if the higher flow rate is maintained.^[45] The vasorelaxation and increase in lumen diameter return τ_w to or near the homeostatic value.

Conversely, an increase in systemic blood pressure induces wall thickening without an increase in lumen diameter,^[14,42] but with an increase in basal smooth muscle tone.^[42] In this case, the remodeling appears to return both τ_w and $\bar{\sigma}_c$ to homeostatic values. A sudden increase in pressure initially increases D_i , increasing $\bar{\sigma}_c$ and reducing τ_w . This causes the endothelial cells to increase smooth muscle activation, reducing D_i and returning τ_w to the homeostatic value. If the increased pressure is maintained, the wall thickness gradually increases, thus reducing $\bar{\sigma}_c$ to near homeostatic values.^[14]

Many studies of blood vessel mechanics and adaptation have focused on animals that are just reaching maturity; however, the mechanics of aging human blood vessels are substantially different. The changes include an increase in structural stiffness, as measured by higher PWV,^[41] thickening of the intimal layer,^[23] and increase in opening angle.^[46] The stiffening is thought to be caused by increased synthesis of collagen and increased crosslinking of collagen molecules.^[47] These studies suggest that studies of human blood vessels, as well as those using animal models, will be needed to characterize adaptation during aging.

VASCULAR DISEASE AND ALTERED BLOOD VESSEL MECHANICS

Vascular diseases such as hypertension also lead to altered blood vessel mechanics. Nichols et al. provide a useful overview.^[40] Atherosclerosis, or hardening of the arteries, is a common occurrence during aging. Atherosclerotic plaques are focal regions where accumulations of lipids and fibrous caps cause local changes in arterial mechanics. Plaques are most common in regions of low or oscillating wall shear stress, such as vessel bifurcations.^[3] Recent work has focused on the mechanics of atherosclerotic plaque rupture, where the fibrous cap detaches from the plaque. These studies require computational models that allow the local heterogeneity of the tissue to be modeled.^[48]

Vascular aneurysms are local enlargements of a blood vessel lumen, accompanied by changes in vessel wall morphology and microstructure, which may lead to potentially fatal rupture. Aneurysms most commonly occur in the ascending aorta, abdominal aorta, cerebral arteries, and lower limb veins. Ascending aortic aneurysms are associated with congenital

connective tissue disorders, such as Marfan Syndrome, and usually consist of enlargement (dilation) of the vessel around its entire circumference.^[40] Abdominal aortic aneurysms, which occur predominantly in older patients with other cardiovascular risk factors, are characterized by the disappearance of elastic fibers from the aortic wall and increased stiffness with an asymmetric bulge on the anterior side. In both of these conditions, a large aortic diameter is associated with an increase in the risk of aneurysm rupture.^[40] In vitro biaxial tests suggest that tissue in both ascending and abdominal aortic aneurysms is stiffer than non-aneurysmal tissue and is somewhat anisotropic.^[19,49] Intracranial aneurysms are characterized by a loss of elastic fibers and local thinning and are typically saccular, i.e., almost spherical with a small connecting neck to the main vessel. They have been modeled as thin-walled membranes subject to pulsatile loading, but a very limited amount of mechanical test data is available, perhaps due to difficulty locating and isolating the small segments.^[8]

Diabetes is generally associated with stiffening of the blood vessel walls through the formation of sugar-mediated collagen crosslinks.^[50] Most measurements are of regional properties, such as PWV, with little in vitro mechanical testing of human tissue to date; however, animal models, such as streptozocin-induced diabetes, also show an increase in aortic stiffness.^[51]

CONCLUSIONS

The vascular system is a complex network of blood vessels that serve multiple functions in the distribution of blood to organs and tissues. As summarized in this article, blood vessel mechanics have been studied at the system, regional, and tissue levels. Understanding the complex relationships between blood vessel mechanics and biological and biochemical processes is the future challenge for this field. Modeling the alterations in blood vessel mechanics in transgenic animal models for human disease is just one example. Engineering of vascular tissue grafts is another. Integrative models that explain mechanical responses at the vascular tissue and cellular levels are needed to allow biomechanics researchers, physiologists, and biologists to link the ever-growing bodies of knowledge in each of these areas.

ARTICLES OF FURTHER INTEREST

Artificial Heart Fluid Dynamics; Bioreactors; Positive Displacement Pumps; Tissue Engineering of Blood Vessel

REFERENCES

1. Rhodin, J. Architecture of the vessel wall. In *Handbook of Physiology, Section 2, the Cardiovascular System*; Berne R.M., Ed.; Vascular Smooth Muscle; American Physiological Society: Bethesda, MD, 1979; Vol. 2, 1–31.
2. Milnor, W.R. *Cardiovascular Physiology*; Oxford University Press: New York, 1990; xiii.
3. Ku, D.N. Blood flow in arteries. *Annu. Rev. Fluid Mech.* **1997**, *29*, 399–434.
4. Wolinsky, H.; Glagov, S. A lamellar unit of aortic medial structure and function in mammals. *Circ. Res.* **1967**, *20* (1), 99–111.
5. McVeigh, G.E.; Hamilton, P.K.; Morgan, D.R. Evaluation of mechanical arterial properties: clinical, experimental, and therapeutic aspects. *Clin. Sci. (Lond.)* **2002**, *102* (1), 51–67.
6. Stanford, W. Advances in cardiovascular CT imaging: CT clinical imaging. *Int. J. Cardiovasc. Imaging* **2005**, *21* (1), 29–37.
7. Fung, Y.C. *Biomechanics: Circulation*; Springer: New York, 1997; xvii.
8. Humphrey, J.D. *Cardiovascular Solid Mechanics*; Springer: New York, 2002.
9. Cox, R.H. Comparison of arterial wall mechanics using ring and cylindrical segments. *Am. J. Physiol.* **1983**, *244* (2), H298–H303.
10. Yin, F.C.; Spurgeon, H.A.; Kallman, C.H. Age-associated alterations in viscoelastic properties of canine aortic strips. *Circ. Res.* **1983**, *53* (4), 464–472.
11. Humphrey, J.D.; Kang, T.; Sakarda, P.; et al. Computer-aided vascular experimentation: a new electro-mechanical test system. *Ann. Biomed. Eng.* **1993**, *21* (1), 33–43.
12. Deng, S.X.; Tomioka, J.; Debes, J.C.; et al. New experiments on shear modulus of elasticity of arteries. *Am. J. Physiol.* **1994**, *266* (1 Pt 2), H1–H10.
13. Lu, X.; Yang, J.; Zhao, J.B.; et al. Shear modulus of porcine coronary artery: contributions of media and adventitia. *Am. J. Physiol. Heart Circ. Physiol.* **2003**, *285* (5), H1966–H1975.
14. Matsumoto, T.; Hayashi, K. Stress and strain distribution in hypertensive and normotensive rat aorta considering residual strain. *J. Biomech. Eng.* **1996**, *118* (1), 62–73.
15. Zulliger, M.A.; Fridez, P.; Hayashi, K.; et al. A strain energy function for arteries accounting for wall composition and structure. *J. Biomech.* **2004**, *37* (7), 989–1000.
16. Cox, R.H. Effects of norepinephrine on mechanics of arteries in vitro. *Am. J. Physiol.* **1976**, *231* (2), 420–425.
17. Cox, R.H. Regional variation of series elasticity in canine arterial smooth muscles. *Am. J. Physiol.* **1978**, *234* (5), H542–H551.
18. Debes, J.C.; Fung, Y.C. Biaxial mechanics of excised canine pulmonary arteries. *Am. J. Physiol.* **1995**, *269* (2 Pt 2), H433–H442.
19. Okamoto, R.J.; Wagenseil, J.E.; DeLong, W.D.; et al. Mechanical properties of dilated human ascending aorta. *Ann. Biomed. Eng.* **2002**, *30* (5), 624–635.
20. Dobrin, P.B.; Baker, W.H.; Gley, W.C. Elastolytic and collagenolytic studies of arteries. Implications for the mechanical properties of aneurysms. *Arch. Surg.* **1984**, *119* (4), 405–409.
21. Roach, M.R.; Burton, A.C. The reason for the shape of the distensibility curves of arteries. *Can. J. Biochem. Physiol.* **1957**, *35*, 681–690.
22. Pandit, A.; Lu, X.; Wang, C.; et al. Biaxial elastic material properties of porcine coronary media and adventitia. *Am. J. Physiol. Heart Circ. Physiol.* **2005**, *288* (6), H2581–H2587.
23. Holzapfel, G.A.; Sommer, G.; Gasser, C.T.; et al. Determination of layer-specific mechanical properties of human coronary arteries with nonatherosclerotic intimal thickening and related constitutive modeling. *Am. J. Physiol. Heart Circ. Physiol.* **2005**, *289* (5), H2048–H2058.
24. Fung, Y.C. *Biomechanics: Mechanical Properties of Living Tissues*, 2nd Ed.; Springer: New York, 1993.
25. Choy, J.S.; Mathieu-Costello, O.; Kassab, G.S. The effect of fixation and histological preparation on coronary artery dimensions. *Ann. Biomed. Eng.* **2005**, *33* (8), 1027–1033.
26. Guo, X.; Kassab, G.S. Variation of mechanical properties along the length of the aorta in C57bl/6 mice. *Am. J. Physiol. Heart Circ. Physiol.* **2003**, *285* (6), H2614–H2622.
27. Hamza, L.H.; Dang, Q.; Lu, X.; et al. Effect of passive myocardium on the compliance of porcine coronary arteries. *Am. J. Physiol. Heart Circ. Physiol.* **2003**, *285* (2), H653–H660.
28. Vaishnav, R.N.; Vossoughi, J. Incremental formulations in vascular mechanics. *J. Biomech. Eng.* **1984**, *106* (2), 105–111.
29. Lu, X.; Pandit, A.; Kassab, G.S. Biaxial incremental homeostatic elastic moduli of coronary artery: two-layer model. *Am. J. Physiol. Heart Circ. Physiol.* **2004**, *287* (4), H1663–H1669.
30. Holzapfel, G.A.; Stadler, M.; Schulze-Bauer, C.A. A layer-specific three-dimensional model for the simulation of balloon angioplasty using magnetic resonance imaging and mechanical testing. *Ann. Biomed. Eng.* **2002**, *30* (6), 753–767.
31. Wang, D.H.; Makaroun, M.S.; Webster, M.W.; et al. Effect of intraluminal thrombus on wall stress in patient-specific models of abdominal aortic aneurysm. *J. Vasc. Surg.* **2002**, *36* (3), 598–604.
32. Fung, Y.C.; Fronek, K.; Patitucci, P. Pseudoelasticity of arteries and the choice of its mathematical expression. *Am. J. Physiol.* **1979**, *237*, H620–H631.
33. Holzapfel, G.A.; Gasser, T.C.; Stadler, M. A structural model for the viscoelastic behavior of arterial walls: continuum formulation and finite element analysis. *Eur. J. Mech. Solids* **2002**, *21* (3), 441–463.
34. Vito, R.P.; Dixon, S.A. Blood vessel constitutive models-1995–2002. *Annu. Rev. Biomed. Eng.* **2003**, *5*, 413–439.
35. Peterson, S.J.; Okamoto, R.J. Effect of residual stress and heterogeneity on the stress distribution in the arterial wall. *J. Biomech. Eng.* **2000**, *122* (4), 454–456.

36. Taber, L.A.; Humphrey, J.D. Stress-modulated growth, residual stress, and vascular heterogeneity. *J. Biomech. Eng.* **2001**, *123* (6), 528–535.
37. Liu, S.Q.; Fung, Y.C. Indicial functions of arterial remodeling in response to locally altered blood pressure. *Am. J. Physiol.* **1996**, *270* (4 Pt 2), H1323–H1333.
38. Gleason, R.L.; Taber, L.A.; Humphrey, J.D. A 2-D model of flow-induced alterations in the geometry, structure, and properties of carotid arteries. *J. Biomech. Eng.* **2004**, *126* (3), 371–381.
39. Bunce, D.F.M. *Atlas of Arterial Histology*; Warren H. Green: St. Louis, MO, 1974.
40. Nichols, W.W.; O'Rourke, M.F.; Hartley, C. *McDonald's Blood Flow in Arteries: Theoretical, Experimental, and Clinical Principles*, 4th Ed.; Oxford University Press: New York, 1998.
41. O'Rourke, M.F.; Staessen, J.A.; Vlachopoulos, C.; et al. Clinical applications of arterial stiffness; definitions and reference values. *Am. J. Hypertens.* **2002**, *15* (5), 426–444.
42. Fridez, P.; Makino, A.; Miyazaki, H.; et al. Short-term biomechanical adaptation of the rat carotid to acute hypertension: contribution of smooth muscle. *Ann. Biomed. Eng.* **2001**, *29* (1), 26–34.
43. Davis, M.J.; Gore, R.W. Length–tension relationship of vascular smooth muscle in single arterioles. *Am. J. Physiol.* **1989**, *256* (3 Pt 2), H630–H640.
44. Wesly, R.L.; Vaishnav, R.N.; Fuchs, J.C.; et al. Static linear and nonlinear elastic properties of normal and arterIALIZED venous tissue in dog and man. *Circ. Res.* **1975**, *37* (4), 509–520.
45. Ben Driss, A.; Benessiano, J.; Poitevin, P.; et al. Arterial expansive remodeling induced by high flow rates. *Am. J. Physiol.* **1997**, *272* (2 Pt 2), H851–H858.
46. Saini, A.; Berry, C.; Greenwald, S. Effect of age and sex on residual stress in the aorta. *J. Vasc. Res.* **1995**, *32* (6), 398–405.
47. Bruel, A.; Oxlund, H. Changes in biomechanical properties, composition of collagen and elastin, and advanced glycation endproducts of the rat aorta in relation to age. *Atherosclerosis* **1996**, *127* (2), 155–165.
48. Chau, A.H.; Chan, R.C.; Shishkov, M.; et al. Mechanical analysis of atherosclerotic plaques based on optical coherence tomography. *Ann. Biomed. Eng.* **2004**, *32* (11), 1494–1503.
49. Vande Geest, J.P.; Sacks, M.S.; Vorp, D.A. The effects of aneurysm on the biaxial mechanical behavior of human abdominal aorta. *J. Biomech.* **2006**, *39* (7), 1324–1334.
50. Sims, T.J.; Rasmussen, L.M.; Oxlund, H.; et al. The role of glycation cross-links in diabetic vascular stiffening. *Diabetologia* **1996**, *39* (8), 946–951.
51. Liu, S.Q.; Fung, Y.C. Changes in the rheological properties of blood vessel tissue remodeling in the course of development of diabetes. *Biorheology* **1992**, *29* (5–6), 443–457.
52. Burton, A.C. Relation of structure to function of the tissues of the wall of blood vessels. *Physiol. Rev.* **1954**, *34* (4), 619–642.

Bone Cement

Julie Hasenwinkel

Department of Bioengineering and Neuroscience, Syracuse University, Syracuse, New York, U.S.A.

B

INTRODUCTION

Acrylic bone cement has been used extensively for the fixation of total joint replacements since its introduction by Sir John Charnley over forty years ago. To date, all commercial bone cements have been based on poly(methyl methacrylate) (PMMA), a polymer that was developed just prior to and during World War II. PMMA was first introduced for medical applications in 1937 when it was used as a denture-base material, and the self-curing resin was subsequently used for cranioplasties and for filling other skeletal defects. In recent years, the application of bone cement in orthopedic surgery has been expanded to vertebroplasty, a procedure in which cement is injected percutaneously into the vertebral body in order to stabilize fractures that occur primarily as a result of osteoporosis. Another variation of this procedure is kyphoplasty, during which a balloon is inserted percutaneously into the vertebral body, inflated to restore the height of the compressed vertebrae, and subsequently filled with injected bone cement to stabilize the fracture. These newer applications of bone cement have been successful in relieving pain and restoring vertebral strength and function to patients.

Although the variety of applications for PMMA in medicine and dentistry illustrate the importance and utility of this polymer as a biomaterial, the focus of this chapter is the application of bone cement for total joint fixation. Approximately 50% of all total joint replacements utilize bone cement to achieve implant fixation, and currently the survival probabilities of cemented total joint arthroplasties average at least 90% after 15 years, especially for hip and knee implants in patients over age 50. In this application, the primary functions of bone cement are to stabilize the prosthesis and transfer mechanical loads between the implant and the bone. PMMA does not chemically bond with either interface, but rather serves as a grouting or space-filling material. Commercial bone cements are supplied as two component systems, consisting of a powder and a liquid, which are mixed in the operating room during the surgical procedure and delivered to the implant site prior to introduction of the prosthesis. As the polymerization reaction proceeds in situ, the powder/liquid admixture experiences a transition from the liquid state

to the solid state, forming a rigid amorphous polymer. This polymerization reaction along with the important physical properties of bone cement, including rheological and setting characteristics, mechanical properties, porosity, and residual monomer, are critical to its success in orthopedic surgery. These properties, which have been extensively investigated, along with the chemistry, mixing, and delivery of bone cement will be discussed.

CHEMISTRY AND POLYMERIZATION REACTION

Chemical Components

Commercial bone cements are supplied prepackaged with the powdered portion contained in a polymeric pouch and the liquid portion contained in a glass ampule. The powder and liquid components contain a variety of chemical ingredients that play an important role in the polymerization process. The powdered portion of the cement contains approximately 90% (by weight) prepolymerized beads of PMMA or PMMA-based copolymers.^[1] The remainder of the powder is typically composed of benzoyl peroxide (BPO), an initiator for the polymerization reaction, and barium sulfate (BaSO_4) or zirconium dioxide (ZrO_2), both of which are radiopacifiers. The liquid portion of acrylic bone cement contains three basic components: methyl methacrylate (MMA) monomer, *N,N*-dimethyl-*p*-toluidine (DMPT), and hydroquinone (HQ).^[1] DMPT is a tertiary aromatic amine that acts as an activator in the polymerization reaction. Hydroquinone is added to the liquid to limit spontaneous polymerization of the monomer due to exposure to heat or light during storage.^[1] It acts as a free radical scavenger and is termed an inhibitor. The structures of these chemical components are shown in Fig. 1.

There is a wide variety of commercial bone cement formulations used worldwide. These different products vary in their exact chemical composition, specifically the relative proportions of initiator, activator, and monomer, along with the types of copolymers used in the powder portion of the cement. Variations in chemical composition affect many of the important mixing, handling, and rheological characteristics of the material during its application in the operating theater

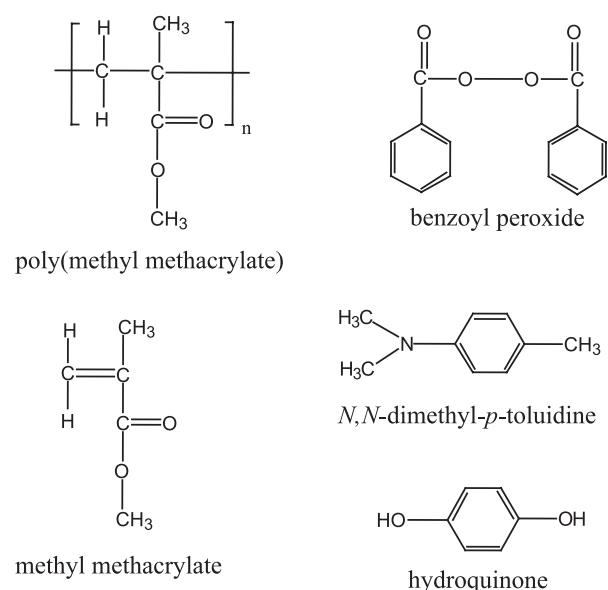


Fig. 1 Chemical structures of the powder and liquid components of acrylic bone cements.

along with the physical properties of the polymerized cement. The specific chemical compositions of seven different commercial cements are shown in Table 1. All of these commercial cements are mixed in a 2:1 (wt/vol) powder-to-liquid ratio, as recommended by their manufacturers.

Polymerization Reaction

The polymerization process by which PMMA and other vinyl polymers are formed is known as free radical or addition polymerization. In bone cement, this polymerization reaction has three major kinetic processes: initiation, propagation or growth, and termination. The process of initiation in free radical polymerization encompasses the reaction or set of reactions that generate the primary radicals that give rise to the active center for polymer chain growth. Radicals can be formed by several different methods including the collision of two monomer molecules of sufficient energy or the decomposition of an initiator molecule by means of energy generated from heat, light, or chemical reaction. In the case of cold-cured PMMA, primary radicals are generated by the chemical reaction of BPO and DMPT. The scheme shown in Fig. 2 outlines a possible mechanism for the reaction of BPO and DMPT, proposed by Horner and Schwenk.^[2] As this figure illustrates, radical formation is not the primary result in amine-peroxide reactions. The toluidine and benzoyloxy free radicals arise from the breakdown of the unstable intermediate products of the initial polar reaction.^[3] Horner and Schwenk^[2] have concluded that the benzoyloxy radical is primarily responsible for initiating polymerization of methyl methacrylate.

Propagation is the process by which PMMA polymer chains grow by the addition of MMA monomer units to

Table 1 Chemical compositions of six commercial formulations of bone cement

Constituent ^a	CMW TM -1	CMW TM -3	Palacos [®] R	Simplex P [®]	Zimmer regular [®]	Zimmer LVC [®]
Powder						
BPO	2.60	2.20	0.5–1.6	1.19	0.75	0.75
BaSO ₄	9.10	10.00	–	10.0	10.00	10.00
ZrO ₂	–	–	14.85	–	–	–
Chlorophyll	–	–	200 ppm ^b	–	–	–
PMMA	88.30	87.80	–	16.55	89.25	89.25
P(MMA/MA) ^c	–	–	83.55–84.65	–	–	–
P(MMA/ST) ^d	–	–	–	82.26	–	–
Liquid						
<i>N,N</i> -DMPT	0.40	0.99	2.13	2.48	2.73	2.75
Hydroquinone	15–20 ppm	15–20 ppm	64 ppm	75 ppm	75 ppm	75 ppm
MMA	98.66	98.07	97.87	97.51	97.27	97.25
Ethanol	0.92	0.92	–	–	–	–
Ascorbic acid	0.02	0.02	–	–	–	–
Chlorophyll	–	–	267 ppm	–	–	–

^aConstituents are in weight percent unless stated otherwise.

^bppm = parts per million.

^cP(MMA/MA) = poly(methyl methacrylate)/poly(methacrylic acid) copolymer.

^dP(MMA/ST) = poly(methyl methacrylate)/poly(styrene) copolymer.

(From Ref. [21].)

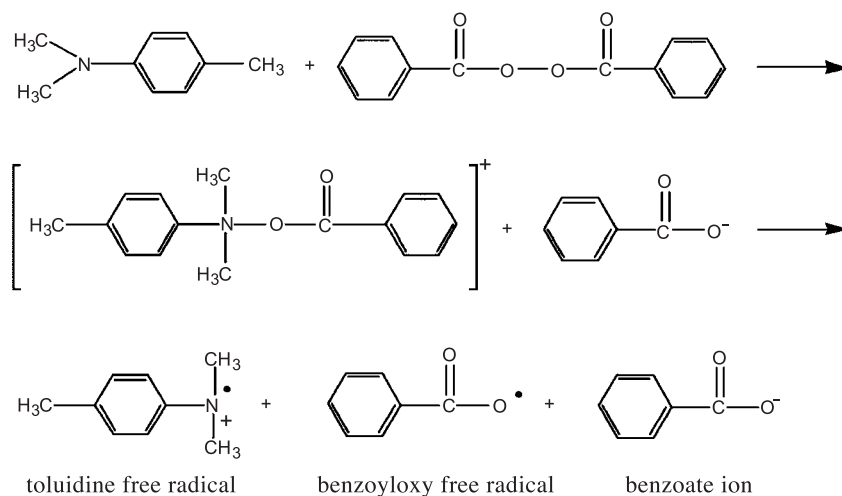


Fig. 2 Redox reaction between benzoyl peroxide and *N,N*-dimethyl-*p*-toluidine. (Figure courtesy of Ref. [2].)

the active center of a free radical. Polymer chain growth is influenced by thermodynamic, chemical, and kinetic conditions. In order for the process to be thermodynamically favorable, it must result in an overall decrease in the free energy of the reacting system. Monomer molecules must have sufficient chemical reactivity to participate in the propagation reaction. And from a kinetic standpoint, the rate of propagation must be on the order of one thousand times higher than the rate of termination in order to grow high polymer chains with sufficient molecular weight. The propagation phase of polymerization is a period of rapid chain growth in which a portion of the available monomer molecules in the system is consumed. A significant consequence of the propagation phase is an increase in the viscosity of the system. This change is due to increases in both the polymer concentration in the system and the molecular weight of the growing PMMA chains.

Termination is the final step in the polymerization process, whereby propagation ceases with the annihilation of radicals in the system. Termination and the formation of “dead” polymer chains occurs via two main mechanisms: combination or disproportionation. Combination is the term used to describe the simple connection of two macroradical chains through mutual reaction. In this case, two radical chains of length m and n combine to form a single polymer chain of length $m+n$. Disproportionation involves the transfer of a hydrogen atom from one radical chain to another and results in two polymer chains.

Physical Aspects of Mixing

When the powder and liquid components of bone cement are mixed together, there are several physical events that occur during the polymerization process. These events include: 1) solvation of polymeric materials in the powder, as well as the BPO into the liquid;

2) diffusion (swelling) of liquid into the organic matrix of the powder; 3) polymer–polymer diffusion from the liquid to the solid phase; and 4) monomer evaporation from the mixture.^[1] In this physical transition, the mixture goes through different states that are often referred to as sandy, stringy, doughy, and finally rigid or set.^[10] During this process, the BPO from the powder and the DMPT from the liquid participate in a redox reaction which produces free radicals that initiate polymerization of the MMA monomer. Physically speaking, the polymerization process of bone cement can be divided into four different phases: 1) the mixing phase; 2) the waiting phase; 3) the working phase; and 4) the hardening or setting phase.^[4]

The mixing phase is simply the initial combination of powder and liquid, typically in a commercial mixing system. The second phase, or waiting phase, occurs during the mixing process while the cement slurry is gradually transformed into a more homogeneous mixture. The waiting phase is characterized by a cement viscosity that is too low for the material to be handled and manipulated by the surgeon. Once the viscosity has increased to an appropriate level, due to the solvation and swelling of the polymer beads, the working phase commences and the cement is transferred from the mixing system into a delivery system or cement gun for insertion into the implant site. During this phase, the cement has a doughy consistency that allows it to flow through the cement gun into the interstices of the bone and around the metallic implant upon its insertion. It is critical that the surgeon delivers the cement, inserts the implant, and manipulates the prosthesis into its final position during the working phase. The final phase is the setting or hardening phase, which is characterized by a sharp increase in cement viscosity associated with the rapid polymerization of the methyl methacrylate monomer in the mixture. This entire process occurs over the course of six to twelve minutes, depending on the particular cement composition, the cement mass, and

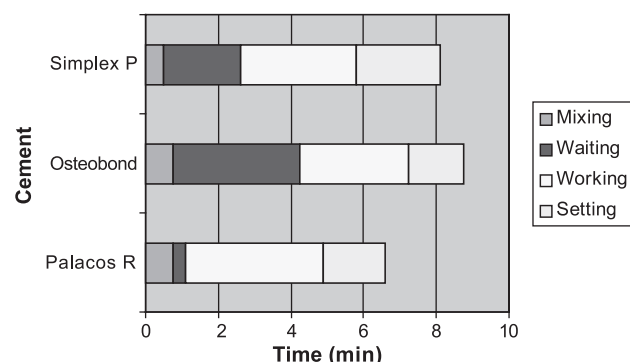


Fig. 3 Time course of the mixing, waiting, working, and setting phases for three different commercial bone cements at 23°C. (Figure courtesy of Ref. [4].)

the ambient temperature. Figure 3 shows the time course of these four phases for Palacos R, Osteobond, and Simplex P.

The final cement mass is actually an inhomogeneous multiphase material with a matrix of PMMA, derived from the polymerization of the methyl methacrylate monomer, surrounding the prepolymerized polymer beads and radiopacifier particulate phase.^[5,6] This is illustrated in Fig. 4. In this sense, bone cement is actually a polymeric composite material. It is interesting to note that the molecular weight of the polymerized cement is commonly higher than the molecular weight of the prepolymerized polymer beads. For Simplex P, the average molecular weight of the powder is between 179,000 and 194,000 g/mol, whereas the molecular weight of the set cement is 242,000 g/mol.^[1,7] From these data, it is hypothesized that the PMMA matrix derived from the liquid monomer yields a higher molecular weight during the polymerization process than the preexisting polymer beads embedded in the matrix.^[7-9]

PHYSICAL PROPERTIES OF BONE CEMENT

Exotherm

During the polymerization reaction, the carbon double bond in the methyl methacrylate monomer is cleaved

and replaced by a single bond. This process results in the release of 130 cal/g of heat for each bond that is broken.^[1] Therefore, the reaction is highly exothermic which leads to a large rise in the temperature of the material during the setting process. The degree of this temperature rise is dependent on the mass and thickness of the cement, the ambient temperature, and the efficiency of heat dissipation to the surroundings. In a comparison of nine different bone cements, Hansen et al.^[10] found a temperature range of 66–82.5°C at the center of 6 mm samples. Meyer et al.^[11] demonstrated a peak exotherm of 60°C for Simplex P when measured in a 3 mm thick specimen and an exotherm of 107°C for a 10 mm thick specimen. Wang et al.^[12] reported temperatures between 67 and 124°C at the center of the cement mantle in vivo, depending on the cement composition.

Although the exothermic character of the polymerization reaction and its potential adverse effects, including thermal necrosis and implant loosening, have been a longstanding concern in the literature,^[13-15] in vitro studies have shown that the temperatures at the bone-cement interface average 40–43°C.^[16,17] This decrease in temperature at the bone-cement interface in the in vivo scenario is due to the relatively thin cement mantle, the wet environment, and the dissipation of heat by the metallic implant.^[18] Eriksson found that an exotherm in excess of 47°C is sufficient to cause thermal necrosis of bone; protein denaturation occurs at 56°C.^[19] Therefore, it is important to limit the temperature increase at the bone-cement interface.

Viscosity and Rheology

The viscosity and rheological properties of bone cement play an important role in the handling characteristics of the material. These properties are often the most apparent difference between the various commercial bone cements from the surgeon's perspective. The viscosity of the powder/liquid admixture increases over time as the polymerization process proceeds. The initial increase in viscosity is due to swelling and solvation of the polymer beads as the monomer ingresses during the wetting process. A more significant increase in viscosity with

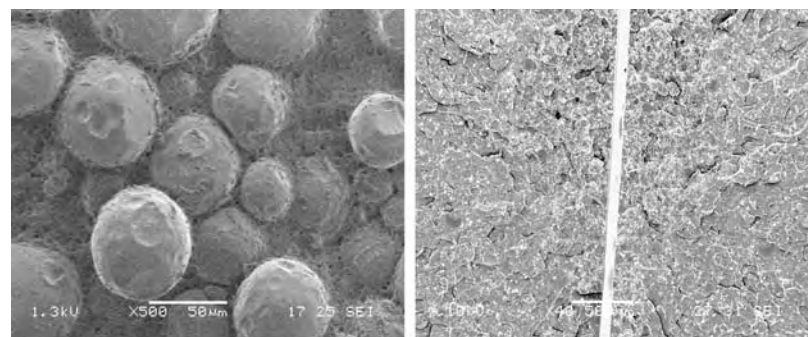


Fig. 4 Scanning electron micrographs of Palacos-R bone cement. The sample on the left was etched with methyl isobutyrate to better reveal the PMMA bead microstructure within the polymerized matrix. The image on the right shows the two halves of a fracture surface after three-point bend testing. The bright white spots distributed throughout the matrix are zirconium dioxide radiopacifier.

Table 2 Mean false apparent dynamic viscosity (Pa · s) of selected bone cements^a

Shear rate (s ⁻¹)	Formulation		
	Simplex P	Zimmer regular	Zimmer LVC
1.00	800	500	100
10.00	500	200	30
97.00	220	80	6

^aMeasurements were made three minutes after commencement of mixing.

(From Ref. [20].)

time corresponds to the increase in weight fraction of polymer in the material as the monomer is polymerized. At a constant shear rate of 97sec^{-1} , the apparent viscosity of Simplex-P bone cement increases threefold, from $200\text{Pa}\cdot\text{s}$ to $600\text{Pa}\cdot\text{s}$, between 3 and 7 min after mixing.^[20]

In addition to the dynamic viscosity changes over the course of the polymerization reaction, the viscosity of PMMA-based cements is also dependent on the shear rate, as shown Table 2. Most polymer solutions and melts are pseudoplastic or shear-thinning, meaning their viscosity decreases with increasing shear rates. This phenomenon is attributed to an alignment of polymer chains along the direction of the applied force, decreasing the resistance to flow. Therefore, the viscosity of bone cement is dependent on the shear force applied in mixing the powder and liquid components. The rheological properties of bone cement, which vary across commercial brands, influence both the ability of the surgeon to handle the material and the ability of the cement to penetrate into the porous bone interface and obtain an adequate mechanical interlock.

Residual Monomer

During the polymerization of bone cement, the conversion of methyl methacrylate is always less than 100%. This is largely due to the rather uncontrolled in situ free radical polymerization process, which lacks the application of additional heat or significant pressures typically used in commercial polymerization processes. One result of this incomplete conversion is the presence of residual monomer (MMA) within the bulk of the cement mantle in vivo. Residual monomer has been implicated in chemical necrosis of the bone,^[21] and it can also affect the material properties of the cement. During the early clinical use of acrylic bone cements for total joint fixation, there was a great deal of concern about the potential toxic effects of unreacted methyl methacrylate on patients.^[13–15,22,23] Hypotension, local tissue irritation, and alveolar lesions have all been

attributed to the release of monomer from polymerizing bone cement.^[24] Although these are valid issues of concern, the clinical success of acrylic bone cements has overshadowed the issue of monomer toxicity. According to Wixson and Lautenschlager,^[18] the larger issues with regard to monomer are the extent of polymerization and its effect on the material properties.

There are several factors that effect the monomer phase of bone cement both during and after the polymerization process. First, methyl methacrylate is highly volatile, leading to significant evaporation upon exposure to air. Monomer loss through evaporation can approach 5%,^[24] and this is dependent on the rate of mixing and the amount of material exposed to air.^[25,26] Residual monomer due to incomplete polymerization has been measured at 3.3% in samples of Simplex P stored in air 30 to 60 minutes after mixing.^[8] This amount declined over time when samples were stored in both air (2.4% after 215 days) and water at 37°C (1.4% after 4.5 months).^[8] The change in residual monomer content over time is due to a combination of continued polymerization and leaching of the monomer into water. Unreacted MMA can influence the mechanical properties of the cement by affecting the degree of polymerization and molecular weight of the polymer, and by acting as a plasticizer within the set polymer. As residual monomer content increases, the strength of the polymer is expected to drop due to a decrease in the length of polymer chains in the bulk material.

Porosity and Volumetric Shrinkage

Porosity and volumetric shrinkage, due to the conversion of monomer into polymer, are two factors that adversely affect the performance of commercial cements. According to Wixson et al.,^[27] there are at least five potential sources of porosity in the bone cement system: 1) air which initially surrounds the powdered constituents; 2) air trapped during the wetting of the powder by the liquid monomer; 3) air stirred into the admixture during spatulation; 4) air trapped during the transfer of the cement mixture to the delivery system; and 5) monomer volatilization during the curing process. There are two different types of pores in polymerized bone cement: macropores ($d > 1\text{mm}$) and micropores ($d \approx 0.1\text{--}1\text{mm}$).^[28] In general, pores act as stress risers and initiating sites for crack formation. Topoleski et al.^[9] found that pores in bone cement act as nucleation sites for microcracks. The stress concentration of the pores is particularly important in fatigue loading, because the relatively low stress intensity allows for the development of a microcracking damage zone over time.^[9] This analysis predicts that a reduction in porosity may enhance the fatigue life of the material.^[9] Although Topoleski et al. also suggested that

Table 3 Mechanical properties of human bone, PMMA, and prosthesis material

	Ultimate stress (MPa)			Modulus (GPa)
	Tensile	Compressive	Shear	
Commercial PMMA	68.9	103		2.1
Simplex P (hand mixed)	25	77	41	2.0
Wet cortical bone	123	166	84	17.2
Wet lumbar vertebrae	3.9	6.3	2.2	0.34
316L stainless steel	480			200

(From Ref. [30] and Ref. [31].)

porosity-enhanced microcrack nucleation may form a larger damage zone and effectively blunt cracks,^[9] the general consensus is that all efforts should be made to reduce both micro- and macropores in bone cement.

Many monomers undergo shrinkage during polymerization due to the difference in density between the monomer and its polymer. The density of methyl methacrylate is 0.937 g/cm³ whereas the density of PMMA is 1.18 g/cm³. There is approximately a 21% decrease in volume with the polymerization of pure monomer to polymer. In the case of commercial bone cements with a 2:1 powder-to-liquid ratio, one-third of the original volume is composed of monomer; therefore, the theoretical shrinkage is reduced to about 7%. This is a maximum shrinkage value, and smaller values are usually observed experimentally due to monomer evaporation, incomplete polymerization, and residual porosity.^[8] Gilbert et al. found volumetric shrinkage values of 6.67±0.40 and 6.50±0.24% for vacuum-mixed Simplex-P and vacuum-mixed Endurance bone cements, respectively.^[29] Shrinkage of the setting cement can have several different detrimental effects on the material. Volumetric contraction can induce residual stresses in the bulk of the cement mantle. The shrinkage process is also capable of inducing porosity along the stem of the implant and retracting the cement from the porous bony interface.^[29] In a constrained material, volumetric contraction can result in the growth of pores that are already present due to entrapped air. High residual stresses and large pores will increase the likelihood of cement fracture and failure.

Mechanical Properties

The mechanical properties of bone cement are very relevant to the clinical performance of the material. The cement mantle must be able to transfer loads from the implant to the surrounding bone, and in serving this function, it experiences a combination of tensile, compressive, and shear forces. The three components of a joint prosthesis construct, bone, cement, and metal, have very different mechanical properties, as shown in

Table 3.^[30,31] The cement is viewed as the weak link in the construct,^[32] due to its low tensile strength and modulus of elasticity with respect to cortical bone and metallic implant materials.

Table 4 shows the range of values, obtained from the literature, for eleven different mechanical properties of commercial cements. The mechanical properties of bone cement are dependent on a number of different variables, including cement formulation, mixing method, curing and aging conditions, and test conditions.^[21] Lewis^[21] has discerned several different trends in an extensive review of cement mechanical properties. First, for a given formulation, the mixing method has a significant influence on the tensile properties of the material. For example, vacuum mixing yields a mean increase of 44% in ultimate tensile strength versus hand mixing (attributed to reduced porosity).^[27] Lewis found that the cement formulation and mixing method have a strong effect on the ultimate compressive strength and compressive modulus. He proposes that the chemical composition, powder size and distribution, polymer morphology, rheological properties, and porosity are

Table 4 Summary of relevant mechanical properties for acrylic bone cements

Property	Value
Ultimate tensile stress (MPa)	23.6–49.2
Compressive strength (MPa)	72.6–117
Three-point flexural strength (MPa)	49.9–84.5
Four-point flexural strength (MPa)	52.7–90.0
Tensile modulus (GPa)	1.58–3.08
Compressive modulus (GPa)	1.94–3.10
Three-point flexural modulus (GPa)	1.29–2.91
Four-point flexural modulus (GPa)	1.95–3.12
Tensile strain to failure (%)	0.86–2.05
Fracture toughness (K _{Ic} , Mpa*m ^{1/2})	1.03–2.02
Fatigue endurance limit (MPa)	6.9–17.0

(From Ref. [21].)

the primary factors that influence the compressive strength of bone cement.

Although static mechanical tests give useful information for comparisons among different types of cements, fatigue tests are a more accurate predictor of the clinical performance of bone cement. Stresses *in vivo* are believed to be much less than the quasistatic strength of the cement; however, hip and knee implants are subjected to cyclical loading in their service environment which can lead to fracture and failure of the cement through fatigue mechanisms. Fatigue failure and fatigue crack propagation have been implicated as the “primary cement failure mechanisms that contribute to aseptic loosening of the prosthesis.”^[21] In a retrieval analysis of cemented femoral components from postmortem asymptomatic patients, Jasty et al.^[33] found evidence of debonding at the cement-prosthesis interface and fatigue crack initiation at this interface and at voids in the cement. The clinical success of cemented implants is highly dependent on maintaining a bone-cement-prosthesis construct free of aseptic loosening. The loosening phenomenon involves interfacial failure between the cement and the bone and/or the implant, bond failure, bone remodeling, and cement failure.^[21] The large modulus mismatch between the cement and surrounding bone and the propensity for debonding at the cement-implant interface contribute to these mechanisms. Cracks that typically develop due to flaws in the bulk of the cement or stress overloads at the cement-implant interface^[34] contribute to cement fracture and failure. Also, cement particles produced by fracture or loosening can evoke an osteolytic response, increasing loosening and hastening failure of the implant.

MIXING AND DELIVERY TECHNIQUES

Vacuum Mixing

Many improvements have been made to bone cement since its introduction by Charnley. The most dramatic area of improvement has likely come in the form of better mixing and delivery methods, such as vacuum mixing,^[27] which have significantly reduced the amount of porosity and improved the mechanical properties of commercial cements. Prior to the advent of vacuum mixing technology, manual or hand mixing was the primary mixing method for combining the powder and liquid components of bone cement. In this process, generally the powder component is added to the liquid in a polymeric bowl and the mixture is stirred with a spatula at a frequency of 1 to 2 Hz for 45 to 120 seconds. This type of spatulation process can introduce a significant amount of air into the mixture during the blending phase and result in a relatively high degree of trapped porosity in the set cement, as high as 10% or

greater. Demarest et al.^[35] first introduced the concept of mixing the powder and liquid components of bone cement under partial vacuum as a means of reducing this porosity. Wixson et al.^[27] subsequently demonstrated that an effective vacuum mixing unit can reduce the porosity in bone cement to 0.1%.

The reduction in porosity of acrylic bone cements associated with vacuum mixing has translated into improved mechanical properties, thus elevating vacuum mixing to the primary mixing method for clinical use. Vacuum mixing results in significantly higher compressive, diametral tensile, and uniaxial tensile strengths as compared to hand mixing.^[27] Perhaps more clinically relevant is the significant improvement in fatigue resistance that accompanies vacuum mixing.^[27] There are currently a wide variety of commercially available vacuum mixing systems, including the Advanced Cement Mixing System (Stryker Instruments, Kalamazoo, MI), the MixOR System and Vortex™ Vacuum Mixer (Smith & Nephew, Memphis, TN), the Optivac® (Biomet Merck, Sjobo, Sweden), and the HiVac™ Bowl and Syringe systems (Summit Medical, Gloucestershire, UK). Although each system has its own unique features, the general operation of each includes mixing the powder and liquid components of the bone cement under a partial vacuum until the desired viscosity is achieved, followed by delivery of the cement through a nozzle-fitted syringe.

Other Mixing Methods

Other bone cement mixing methods that have received considerable attention include the use of centrifugation to decrease porosity and prechilling of the liquid component of the cement prior to mixing in order to improve the workability and decrease the porosity of the cement. Centrifugation typically involves hand mixing the cement and then placing it in a centrifuge for 30 to 180 seconds at maximum speeds ranging from 2300 to 4000 rpm.^[21] This process is designed to induce the flow of bubbles from the bulk of the mixture out to the surface of the mixture, thus significantly reducing trapped porosity prior to the complete polymerization of the cement. Burke et al.^[36] demonstrated that centrifugation not only significantly reduces the porosity of bone cement, but also increases the fatigue life by 136%, compared to hand mixing. The effectiveness of centrifugation in reducing porosity and improving mechanical properties is dependent on the relative viscosity of the cement formulation. The cement viscosity has to be low enough to allow bubbles to flow through it and to the surface. Davies et al.^[37] demonstrated that centrifugation was most effective with Simplex P bone cement, reducing the porosity from 9.4 to 2.9% and increasing the mean fatigue life from 15,000 to 70,000 cycles.

Prechilling of the liquid component of bone cement has also been investigated as a means of improving mixing and reducing porosity. In 1990, Hansen and Jensen^[10] studied nine different bone cements, focusing on their handling characteristics, doughing time, setting time, and exothermic temperature. They compared the effects of manual and vacuum mixing, along with the effects of mixing the components at room temperature versus 5°C. At room temperature, they were unable to adequately vacuum mix high-viscosity cement formulations due to the reduction in setting and working time associated with vacuum mixing. Prechilling the cement components sufficiently prolonged the setting time and preserved a lower viscosity during the handling period, which enabled successful vacuum mixing and the expected reduction in porosity.

FUTURE DIRECTIONS IN BONE CEMENT RESEARCH

Much of the recent and current bone cement research is focused on improving cement properties and interactions with the biological environment through changes in cement chemistry and formulations. One example of this is the development of bioactive bone cements, which are designed to improve interfacial bonding with bone for enhanced mechanical fixation and biological performance. There are two different types of bioactive cements, ones in which bioactive agents are added to existing PMMA-based bone cements, and others that rely on the design of alternative matrices for these bioactive agents.^[38] Other chemical compositions are also being investigated as alternatives to the current acrylic cement formulations. Vazquez, Deb, Liso, and others^[39–43] have studied a number of different tertiary aromatic amines as alternatives to DMPT in the redox initiation system, with the goals of decreasing the polymerization exotherm, increasing the setting time, and reducing the degree of leaching of toxic residual chemicals after polymerization. They have demonstrated in vitro biocompatibility, reduced exotherms and increased setting times, and comparable mechanical properties and residual monomer content as compared to current commercial cements; however, it remains to be seen whether these alternate cement chemistries will have an impact on clinically used formulations.

Finally, another approach to improving the properties of powder–liquid bone cements has been the development of two-solution bone cement with the same chemical constituents as current commercial formulations.^[44,45] In this design, PMMA powder is predissolved in methyl methacrylate monomer to form two separate high viscosity, porosity-free solutions with BPO added to one and DMPT added to the other. Because of their high initial viscosity, similar to the

viscosity of powder–liquid cements in the dough stage, these solutions can be simultaneously mixed and delivered to the surgical site from a single closed system. This approach eliminates the issues of technique sensitivity due to cement mixing, transfer, and delivery. Hasenwinkel et al.^[44,45] have also shown that relevant material properties such as the polymerization exotherm, setting time, mechanical properties, and residual monomer content of two-solution cements can be controlled by the concentration and ratio of initiator chemicals along with other factors. These cements compare favorably in these properties to current powder–liquid cements and hold promise for the future.

CONCLUSIONS

Acrylic bone cements have had a significant impact on medicine and specifically orthopedic surgery since their introduction by Charnley in the 1960s. While significant research efforts have led to improvements in material properties and clinical outcomes over the past forty years, it is clear that there is still potential for alternate materials and formulations for these applications. However, until further progress is made, powder–liquid based acrylic bone cements remain the gold standard for fixation of total joint prostheses.

ARTICLES OF FURTHER INTEREST

Antimicrobial/Antibiotic (Infection Resistance) Materials; Bioadhesion; Bone Fracture Fixation; Bone–Implant Interface; Dental Implants; Implant, Total Hip; Knee Joint Replacement; Orthopedic Biomaterials; Orthopedic Fixation Devices

REFERENCES

1. Lautenschlager, E.P.; Stupp, S.I.; Keller, J.C. Structure and Properties of Acrylic Bone Cement. In *Functional Behavior of Orthopedic Biomaterials*; Ducheyene, P., Hastings, G.W., Eds.; CRC Press: Boca Raton, FL, 1984; 87–119.
2. Horner, L.; Schwenk, E. Über die polymerisation von vinylverbindungen durch das system: Diacylperoxyol-tertiäres amin. *Ann. Chem.* **1950**, *566*, 69–84.
3. Walling, C. *Free Radicals in Solution*; John Wiley & Sons: New York, 1957; 590–594.
4. Kuhn, K.D. *Bone Cements: Up-to-Date Comparison of Physical and Chemical Properties of Commercial Materials*; Springer-Verlag: Berlin, 2000; 272.
5. Kusy, R.P.; Turner, D.T. Characterization of the microstructure of a cold-cured acrylic resin. *J. Dent. Res.* **1974**, *53*, 948.

6. Cameron, H.U.; Mills, R.H.; Jackson, R.W.; Macnab, I. The structure of polymethylmethacrylate cement. *Clin. Orthop.* **1974**, *100*, 287–291.
7. Brauer, G.M.; Termini, D.J.; Dickson, G. Analysis of the ingredients and determination of the residual components of acrylic bone cements. *J. Biomed. Mater. Res.* **1977**, *11*, 577–607.
8. Haas, S.S.; Brauer, G.M.; Dickson, G. A characterization of polymethylmethacrylate bone cement. *J. Bone Jt. Surg.* **1975**, *57A* (3), 380–391.
9. Topoleski, L.D.T.; Ducheyne, P.; Cuckler, J.M. Microstructural pathway of fracture in poly(methyl methacrylate) bone cement. *Biomaterials* **1993**, *14* (15), 1165–1172.
10. Hansen, D.; Jensen, J.S. Prechilling and vacuum mixing not suitable for all bone cements. Handling characteristics and exotherms of bone cements. *J. Arthroplast.* **1990**, *5*, 287–290.
11. Meyer, P.R.; Lautenschlager, E.P.; Moore, B.K. On the setting properties of acrylic bone cement. *J. Bone Jt. Surg.* **1973**, *55A*, 149–156.
12. Wang, J.-S.; Franzen, H.; Toksvig-Larsen, S.; Lidgren, L. Does vacuum mixing of bone cement affect heat generation? Analyses of four cement brands. *J. Appl. Biomater.* **1995**, *6*, 105–108.
13. Charnley, J. *Acrylic Cement in Orthopaedic Surgery*; Williams and Wilkins: Baltimore, 1970.
14. Feith, R. Side-effects of acrylic cement implanted into bone. A histological, (micro)angiographic, fluorescence-microscopic and autoradiographic study in rabbit femur. *Acta Orthop. Scand. Suppl.* **1975**, *161*, 3–136.
15. Mjoberg, B. Loosening of the cemented hip prosthesis. The importance of heat injury. *Acta Orthopaed. Scand. Suppl.* **1986**, *221*, 1–40.
16. Reckling, F.W.; Dillon, W.L. The bone-cement interface temperature during total joint replacement. *J. Bone Jt. Surg.* **1977**, *59A*, 80–82.
17. Toksvig-Larsen, S.; Franzen, H.; Ryd, L. Cement interface temperature in hip arthroplasty. *Acta Orthoped. Scand.* **1991**, *62*, 102–105.
18. Wixson, R.L.; Lautenschlager, E.P. Methyl Methacrylate. In *The Adult Hip*; Callaghan, J.J., Rosenberg, A.G., Rubash, H.E., Eds.; Lippincott-Raven Publishers: Philadelphia, 1998; 135–157.
19. Eriksson, R.A.; Albrektsson, T. The effect of heat on bone regeneration. An experimental study in the rabbit using the bone growth chamber. *J. Oral Maxillofac. Surg.* **1984**, *42*, 707–711.
20. Krause, W.R.; Miller, J.; Ng, P. The viscosity of acrylic bone cements. *J. Biomed. Mater. Res.* **1982**, *16*, 219–243.
21. Lewis, G. Properties of acrylic bone cement: State of the art review. *J. Biomed. Mater. Res. (Applied Biomaterials)* **1997**, *38*, 155–182.
22. Linder, L.G.; Harthorn, L.; Kullberg, L. Monomer leakage from polymerizing acrylic bone cement. An in vitro study on the influence of speed and duration of mixing, cement volume and surface area. *Clin. Orthop.* **1976**, *119*, 242–249.
23. Willert, H.G.; Ludwig, J.; Semlitsch, M. Reaction of bone to methacrylate after hip arthroplasty. A long-term gross, light microscopic, and scanning electron microscopic study. *J. Bone Jt. Surg.* **1974**, *56A*, 1368–1382.
24. Bayne, S.C.; Lautenschlager, E.P.; Greener, E.H.; Meyer, P.R. Clinical influences on bone cement monomer release. *J. Biomed. Mater. Res.* **1977**, *11*, 859–869.
25. Lee, A.J.; Wrighton, J.D. Some properties of polymethylmethacrylate with reference to its use in orthopedic surgery. *Clin. Orthop.* **1973**, *95*, 281–287.
26. Lee, A.J.; Ling, R.S. Further studies of monomer loss by evaporation during the preparation of acrylic cement for use in orthopaedic surgery. *Clin. Orthop.* **1975**, *106*, 122–125.
27. Wixson, R.L.; Lautenschlager, E.P.; Novak, M.A. Vacuum mixing of acrylic bone cement. *J. Arthroplast.* **1987**, *2* (2), 141–149.
28. Wang, J.-S.; Franzen, H.; Jonsson, E.; Lidgren, L. Porosity of bone cement reduced by mixing and collecting under vacuum. *Acta Orthop. Scand.* **1993**, *64*, 143–146.
29. Gilbert, J.L.; Hasenwinkel, J.M.; Wixson, R.L.; Lautenschlager, E.P. A theoretical and experimental analysis of polymerization shrinkage of bone cement: A potential major source of porosity. *J. Biomed. Mater. Res.* **2000**, *52*, 210–218.
30. Black, J. *Orthopaedic Biomaterials in Research and Practice*; Churchill Livingstone, Inc.: New York, 1988.
31. Saha, S.; Pal, S. Mechanical properties of bone cement: A review. *J. Biomed. Mater. Res.* **1984**, *18*, 435–462.
32. Collis, D.K. Femoral stem failure in total hip replacement. *J. Bone Jt. Surg.* **1977**, *59A*, 1033–1041.
33. Jasty, M.; Maloney, W.J.; Bragdon, C.R.; O'Connor, D.O.; Haire, T.; Harris, W.H. The initiation of failure in cemented femoral components of hip arthroplasties. *J. Bone Jt. Surg.* **1991**, *73B*, 551–558.
34. Huiskes, R. Mechanical failure in total hip arthroplasty with cement. *Curr. Orthop.* **1993**, *7*, 239–247.
35. Demarest, V.A.; Lautenschlager, E.P.; Wixson, R.L. Vacuum mixing of acrylic bone cement. *Trans. Soc. Biomater.* **1983**, *9*, 37.
36. Burke, D.; Gates, E.; Harris, W. Centrifugation as a method of improving tensile and fatigue properties of acrylic bone cement. *J. Bone Jt. Surg.* **1984**, *66A*, 1265–1273.
37. Davies, J.P.; Jasty, M.; O'Connor, D.O.; Burke, D.W.; Harrigan, T.P.; Harris, W.H. The effect of centrifuging bone cement. *J. Bone Jt. Surg.* **1989**, *71*, 39–42.
38. Harper, E.J. Bioactive bone cements. *Proc. Inst. Mech. Eng., H J. Eng. Med.* **1998**, *212* (2), 113–120.
39. Liso, P.A.; Vazquez, B.; Rebuelta, M.; Hernaez, M.L.; Rotger, R.; Roman, J.S. Analysis of the leaching and toxicity of new amine activators for the curing of acrylic bone cements and composites. *Biomaterials* **1997**, *18* (1), 15–20.
40. Deb, S.; Silvio, L.D.; Vazquez, B.; Ramon, J.S. Water absorption characteristics and cytotoxic and biological evaluation of bone cements formulated with a novel activator. *J. Biomed. Mater. Res.* **1999**, *48* (5), 719–725.
41. Vazquez, B.; Elvira, C.; Levenfeld, B.; Pascual, B.; Goni, I.; Gurruchaga, M.; Ginebra, M.P.; Gil, F.X.; Planell, J.A.; Liso, P.A.; Rebuelta, M.; Roman, J.S. Application of tertiary amines with reduced toxicity to the curing

- process of acrylic bone cements. *J. Biomed. Mater. Res.* **1997**, *34*, 129–136.
42. Vazquez, B.; Roman, J.S.; Deb, S.; Bonfield, W. Application of long chain amine activator in conventional acrylic bone cement. *J. Biomed. Mater. Res.* **1998**, *43* (2), 131–139.
 43. Vazquez, B.; Deb, S.; Bonfield, W.; Roman, J.S. Characterization of new acrylic bone cements prepared with oleic acid derivatives. *J. Biomed. Mater. Res.* **2002**, *63* (2), 88–97.
 44. Hasenwinkel, J.M.; Lautenschlager, E.P.; Wixson, R.L.; Gilbert, J.L. Effect of initiation chemistry on the fracture toughness, fatigue strength, and residual monomer content of a novel high-viscosity, two-solution acrylic bone cement. *J. Biomed. Mater. Res.* **2002**, *59*, 411–421.
 45. Hasenwinkel, J.M.; Lautenschlager, E.P.; Wixson, R.L.; Gilbert, J.L. A novel high viscosity, two solution acrylic bone cement: Effect of chemical composition on properties. *J. Biomed. Mater. Res.* **1999**, *47* (1), 36–45.

Bone Fracture Fixation

Eleftherios Tsiridis

The Whittington Hospital, University College London, London, Institute of Orthopaedics, Royal National Orthopaedic Hospital, Stanmore, United Kingdom, and Boston Medical University, Boston University School of Medicine, Boston, MA, U.S.A.

Constantin Schizas

The Whittington Hospital, University College London, London, and Institute of Orthopaedics, Royal National Orthopaedic Hospital, Stanmore, United Kingdom

B

INTRODUCTION

The recent advances in orthopedic trauma research have brought up a new era in understanding and treating skeletal injuries. Since the beginning of the last century, metallurgic advances and mechanical principles have been introduced in orthopedic science with particular reference to trauma surgery. The further understanding of bone healing biology and bone biomechanics has led orthopedic surgery to the evolution of bone fracture fixation. An overview of the main fracture fixation techniques in modern orthopedic surgery, with emphasis on the biological and biomechanical principles of each, is presented in this article. Indications, complications, surgical techniques, and controversies surrounding each method of treatment are noted, but not discussed in depth. Internal and external fixation devices are described, namely wires, screws, plates, intramedullary nails, and several types of external fixators.

BONE BIOLOGY

Bone Structure

Bone is a living tissue organized so as to support the human body and serve as an attachment for muscles, allowing locomotion as well as protecting vital structures. The basic unit of bone is the osteon, consisting of concentric lamellae around a central canal of 20 microns, occupied by blood vessels. These lamellae are formed by cells (osteocytes) and matrix composed of organic (40%) and inorganic components (60%). Organic components include collagen, proteoglycans, and growth factors. Inorganic components include calcium hydroxy-apatite and osteocalcium phosphate. Bone is either cortical or cancellous. Cortical, or compact, bone (80% of the skeleton) is formed of tightly assembled osteons and has a relatively high Young's modulus. It is mainly found in the outer aspect of long bones. Cancellous bone is formed by less

densely assembled osteons and has a smaller Young's modulus. It is mainly found toward the ends (epiphysis) of long bones.

Bone is three times lighter than iron but ten times more flexible, although bone and iron have the same tensile strength. Bone is, nevertheless, anisotropic and has a modulus of around 18 giganewtons/m². Fractures can be caused by direct or indirect trauma. The latter type can be traction, angulation, rotational, or compression fractures, or fractures caused by a combination of these forces.^[1]

Biology of Bone Healing

The immediate reaction to fracture is an inflammatory phase that lasts up to 72 hours. The ends of the fractured bone do not actively participate in this process, but are in fact dead. They play only a passive role in the bridging process between the more distant living bone edges. Injured tissues and platelets release vasoactive mediators (which induce vasodilatation) and cytokines. Cytokines influence migration and proliferation of mesenchymal stem cells and their differentiation to osteoprogenitor cells, as well as matrix synthesis (i.e., members of the Tissue Growth Factor- β superfamily, such as Bone Morphogenetic Proteins 2 and 7, Platelet-Derived Growth Factor, Fibroblast Growth Factor, Insulin Growth Factor, and Interleukins 1 and 7). During this process macrophages remove the dead tissue.

Following the inflammatory phase, the reparative response begins and can last from two days to two weeks. Development of the repair tissue includes organization of the fracture hematoma, the appearance of fibroblasts and chondroblasts between the bone ends, and the formation of cartilage (type-II collagen). The clot is then replaced by provisional bone called callus. Three types of calluses are recognized: the external (or bridging) callus, which ossifies by endochondral ossification to form woven bone; the internal (or medullary) callus, which forms at a

slower rate, and the periosteal callus, which ossifies by intramembranous ossification. When enough callus is formed to bridge the ends, a remodeling phase will take place. This phase may last up to seven years and consists of remodeling of the woven bone to form mature cortical and cancellous bone, affected by mechanical function according to Wolff's law.

The initial bridging process can be achieved either by callus formation (secondary bone healing) or no callus formation (primary bone healing). The amount of callus formed depends on the amount of movement of the fractured bone ends. Rigid internal fixation can abolish this micromotion and lead to primary bone healing with little or no callus formation. In this situation the bone ends do not undergo resorption but are recanalized by Haversian systems (osteons) directly.^[1]

Bone Healing and Fracture Fixation

Development of external callus is the most rapid of all processes. It predominates in fractures treated with external fixation devices. Primary bone union, by contrast, is the slowest method of healing, and artificial stability must be maintained for many months. The advantage of rigid fracture fixation is that it allows early patient or limb mobilization. An example of this type of bone healing is a fracture rigidly and anatomically fixed with a compression plate. Intra-articular fractures require accurate or anatomical reduction and stable fixation in order to minimize the late onset of posttraumatic arthritis. Fractures at a distance from the joint can be allowed to heal without necessarily achieving full apposition of the fracture ends, provided that there is no axial or rotational deformity. A new concept known as biological fixation is aimed at reducing surgical trauma produced while exposing the fracture ends. It also involves less-rigid internal fixation. A more flexible fixation is thought to induce formation of callus and is achieved by bridging the fracture area with an implant, without fixing all intermediate fragments. New implants have been designed in order to achieve this type of fixation.

FRACTURE MECHANISM

Five different mechanisms of injury—tension, compression, shear, bending, and torsion—can be described in association with the most common skeletal injuries. Comminution is the expression of the amount of energy transmitted to the bone during the injury.

MATERIALS

The most commonly used materials for fracture fixation devices are cold-worked stainless steel (ISO 5832-1), unalloyed titanium (ISO 5832-2), and titanium alloy (Ti-6Al-7Nb, ISO 5832-11). Titanium has the advantage of producing fewer artifacts during magnetic resonance imaging (MRI) and is therefore very useful in spinal surgery, where imaging of the neural structures might be required after surgery. For certain fractures, biodegradable materials composed mainly of polyglycolic or polylactic acid have been developed, and they are currently under evaluation. Theoretically, they offer the advantage of gradual transfer of load to the bone and elimination of the need for removal. However, their tensile strength and rate of degradation are not ideal, and reported high incidence of inflammatory reactions has made them unpopular.

INTERNAL FIXATION

Wire Fixation

The simplest method to hold two fragments together has been simple wire fixation. This is still commonly used in wrist fractures. The wires can be introduced percutaneously and either buried under the skin or left protruding through the skin. In the latter case, it is necessary to remove them when satisfactory bony healing has been achieved, in order to minimize risk of infection. Other anatomical areas where wires can be used include the medial malleolus, the foot, and the hand. Wires are also being used in a circular manner, usually in order to supplement fixation of a long-bone fracture or in fractures around a previously implanted hip prosthesis. This is known as the cerclage wiring technique.

A particular method of wire fixation is known as tension band wiring. It consists of using straight, smooth wires and a figure-of-eight wire. This technique is applied in anatomical areas, where forces displacing the fracture fragments are of tensile nature. Typically, this situation is encountered in fractures of the olecranon at the elbow, where the triceps tendon tends to displace the bony fragments (Fig. 1). Another application of tension band wiring is in fractures of the patella where the extensor mechanism of the knee displaces the fracture fragments.

Screw Fixation

Screws can be used in isolation or in combination with other fixation devices, such as plates and intramedullary nails. Screw fixation alone is rarely used in



Fig. 1 Tension band wiring of an olecranon injury.

long-bone fractures, because it can fail. In such cases they are used as an adjunct in order to compress the fracture fragments. Screws are also used in the proximity of joints and in intra-articular fractures, and different diameters and types of screws are available.

In the 1950s, a group of surgeons from Switzerland founded the Association for the Study of Internal Fixation (AO) and popularized the use of standardized screw designs. Most manufacturers now use the same types and diameters as well as the same screw heads, so it is quite possible to continue treatment of a patient originally treated in another country.

Biomechanical Characteristics

A screw is a complex device with a four-part construction: head, shaft, thread, and tip. The main use of the screw in orthopedics is to fix two fragments together by compressing them against each other. The head serves as an attachment for the screwdriver, which can be hexagonal, cruciate, or slotted in design. It also serves as a counterforce against compression generated by the screw during insertion into the bone.

The shaft is the smooth portion of the screw between the head and the threaded portion. There are three different screw diameters to consider: the core diameter, the thread diameter (threaded part), and the shaft diameter (smooth part).

The thread has four characteristics: shape, depth, pitch and lead. The shape is asymmetrical—flat on the upper surface and rounded underneath. The depth of the thread is half the difference between the thread diameter and the core diameter. The thread is deeper in cancellous screws and shallower in cortical screws. Pitch is the distance between adjacent threads, and lead is the distance that the screw advances into the bone with a turn of 360°. The tip of the screw is either blunt-rounded (requires pretapping), or fluted or trocar (self-tapping).

Screw Types

Screws are mainly two types: cancellous and cortical. Either type can be fully or partially threaded.

Cancellous screws are designed to provide a better hold in the softer cancellous bone, so they have a deeper thread and coarse pitch. Bone screws usually require predrilling and tapping, but some screws are self-tapping. Screws are produced in diameters of 6.5 mm (large cancellous screws), 4.5 mm (large cortical screws), 3.5 mm (small-fragment cancellous and cortical screws), or smaller screws usually employed to fix hand fractures (2.7 mm, 2 mm and 1.5 mm). Partially threaded cancellous screws come in two thread lengths: 16 mm or 32 mm.

The Lag Screw Principle

Compression of two fragments can be achieved with the lag screw technique, in two ways. The screw can be designed as a lag screw, which means being partially threaded with the thread near to the tip. On the other hand a plain, fully threaded screw can be inserted in such a way as to act as a lag screw. The latter can be achieved by drilling a large-diameter proximal hole (sliding hole) in which the screw has no perches and a smaller (diameter equal to the screw core diameter) distal hole (tread hole) in which the screw has got strong perches. By inserting the screw, compression of the distal fragment to the proximal one can be achieved. The lag screw is best positioned at a right angle to the fracture plane.

There are four principal applications of screws:

1. In the cancellous epiphyseal or metaphyseal bone, as lag screws
2. In cortical bone with plates
3. With intramedullary nails to increase construct rigidity
4. To permit displacement in an axial direction (dynamic hip or dynamic condylar screw)

Plate Fixation

The AO group has also influenced plate design, and several manufacturers offer similar sizes. Large plates accommodate 4.5 mm screws and tend to be used in lower-limb fractures. Smaller plates accommodating 3.5 mm screws are available for fractures around joints or for forearm fractures (Fig. 2). Mini-plates for hand and foot fractures can accommodate 2.7 mm and 1.5 mm screws. Plates are available in different shapes to accommodate the anatomy of the bone that needs internal fixation (e.g., condylar plates for fractures of the distal femur and tibial head buttress plates for fractures of the proximal tibia). With dynamic compression plates, an additional improvement in design of the screw holes has allowed axial compression to be applied as the screws are fully tightened. More



Fig. 2 Plating of a forearm fracture.

recently, plates have been designed to hardly touch the bone surface, the screws being locked to the plate holes. These plates act as internal splints and are named LC-DCP (low-contact dynamic compression plate). One of the most commonly used plate designs is the so-called dynamic hip screw (DHS) used in the fixation of extracapsular fractures of the femoral neck. This design consists of a plate with a fixed-angle barrel, inside which slides a large partially threaded screw. This results in axial loading being translated to compression force at the fracture side, enhancing callus formation and healing.^[2]

Biomaterial and Biocompatibility

The most commonly used material for plate manufacturing is stainless steel, but titanium is more biocompatible and offers desirable plasticity. Plates and screws must be of the same material to prevent galvanic corrosion.

Principle of dynamic compression

The dynamic compression plate has specially designed screw holes, which have an inclined plane at one end so the two ends are at different levels. When the screw is inserted at the higher-level end and tightened down, it slides along this plane, pulling the plate and causing compression of the bone fragments. If, on the other hand, the screw is placed at the center of such a hole, the screw does not move and is said to be a neutral position.

Positioning and type of structural support

Plates can be used as:

1. Buttress devices, usually around joints.
2. Dynamic compression devices, usually in bones with uneven loading (i.e., femur) where there is a

tension and a compression side. The plate should always be placed on the tension side to achieve fracture healing.

3. Static compression devices: The forearm bones are the best example.
4. Neutralization devices, used to protect single screw fixation (lateral malleolar fractures).
5. Bridging devices, used for length restoration without fracture fixation. These can be augmented occasionally with bone graft.

Plates and screws in different bones

Every bone requires a certain number of screws when a single plate is applied. The humerus requires three to four screws on each side of the fracture, whereas the femur requires at least four, the radius and ulna six, and the tibia six to four.

INTRAMEDULLARY NAILING

History and Development of the Technique

The first reported intramedullary device is found in Simson's *Textbook of Surgery* (1883), where the author describes the use of ivory pegs introduced into the medullary canal of a bone. Hey-Grooves (1921) inserted the first metallic femoral intramedullary device during the First World War, and Smith-Petersen treated hip fractures with percutaneous medullary devices. Professor Gerhard Kuntscher (1939) focused his research and clinical work on intramedullary devices for fracture fixation and since, he has been considered the pioneer of the modern era of intramedullary nail fixation.^[3] See Figs. 3 and 4.



Fig. 3 Tibial nailing (anteroposterior view).



Fig. 4 Tibial nailing (lateral view).

Basic Biomechanical Definitions

Inertia (Newton's first law)	If a zero net external force acts on a body, the body will remain at rest or move uniformly.
Force	A push or pull causing external (acceleration) and internal (strain) effect
Moment (M)	Rotational effect of a force on a body about a point
Deformation	Change of shape of the material to which the force is applied. Deformation may be linear (i.e., compression or tension) or bending and torsion.
Rigidity or stiffness	Measure of amount of load required to produce unit deformation under given load configuration
Resistance to deformation	In bending, defined as bending rigidity; in torsion, defined as torsional rigidity.

Biomechanical Characteristics of Intramedullary Fixation

The most important mechanical factors in a biologically compatible intramedullary fixation system are strength, rigidity, and gripping strength.^[1] The amount of load that a structure can withstand before failure occurs is its strength. In characterizing an intramedullary nail, two modes of rigidity need to be considered: torsional rigidity and bending rigidity.

Torsional rigidity is the amount of torque needed to produce (or resistance to) a unit amount of torsional (rotational) deformation. It depends on both material properties (shear modulus, G) and structural properties (polar moment of inertia, J). Bending rigidity is the amount of force required to produce (or resistance to) a unit amount of deflection. It depends on both

material properties (elastic modulus) and structural properties (area moment of inertia, I). In both cases the resistance to deformation is very strongly influenced by the distribution of material about the axis (neutral axis) of bending or torsion. Along the neutral axis, no stress is present. Maximum stress occurs along the outermost fibers of the material in both bending and torsion. How the material is distributed away from the neutral axis is represented by the moment of inertia. In cases of bending, this is called the area moment of inertia, I , whereas polar moment of inertia, J , is used for cases of torsion. In other words, the more material can be distributed away from the neutral axis, the higher the moment of inertia and the greater the capability of the implant to bear load.

For example, two cylindrical nails containing the same amount of metal in the cross section (one being hollow with a greater diameter, and the other being solid with a smaller diameter) will have different values for I and J . Although both nails have the same cross-sectional area of metal, the larger nail is stiffer in bending and torsion because both I and J are larger, due to the fact that the material is distributed farther away from the bending or the torsional axis.^[4] Reaming increases the implant–bone interface contact area, allows larger nails to be inserted, and increases torsional and bending rigidity, respectively.

Gripping strength is the resistance to slipping at the implant–bone interface and is essential for the transmission of torque between fracture fragments. Gripping strength can be augmented by cortical reaming to increase the length of cortical contact. Gripping is optimized by the interlocking screw placement.

Nails are made of stainless steel 316L (L designates low interstitial carbon content, which is less susceptible to corrosion in vivo), 22:13:5 steel, titanium, titanium–aluminum–vanadium (Ti–6Al–4V), and titanium–aluminum–niobium (Ti–6Al–7Nb). The smaller-diameter nails should be made of alloys having superior fatigue strength.

Technical Characteristics

The working length—not true nail length—is the important biomechanical factor. Working length is defined as the length of a nail between the most proximal point of fixation in the distal fragment and the most distal point of fixation in the proximal fragment, and is directly related to the implant stiffness. Bending stiffness is inversely proportional to the square of the working length. Torsional stiffness is inversely proportional to the working length. In other

words, a short working length increases nail rigidity in both bending and torsion. In any case, a fracture should be bypassed by at least 5 cm by the intramedullary nail, because the peak stress around the distal screw holes may exceed the endurance limit of the metal.

The diameter of the nail is the principal factor that alters bending stiffness, and is related to the fourth power (of the radius) in a cylindrical nail.

A wide range of nail widths is available, to match different bone configurations. Nails may have a small diameter (i.e., Rush rods and Ender pins) relative to the medullary canal or a large diameter (Kuntscher, Kempf and Crosse, Russell and Taylor, AO/ASIF) to fit the canal following reaming.

The curvature of the nail can simulate bone anatomy (i.e., a femoral nail is usually curved in a sagittal plane) or facilitate nail insertion (i.e., humeral or tibial nail proximal to sagittal curvature to accommodate insertion portals offset from the center of the medullary canal).

Cross-Sectional Geometry

Nails can be hollow (can be inserted over a guide rod following reaming) or solid (can be inserted reamed or unreamed) (i.e., AO/ASIF). Slotted nails may be open-sectioned or close-sectioned. Open-sectioned nails may be slotted anteriorly (i.e., AO/ASIF) or posteriorly (i.e., Kempf and Cross). The presence of a slot increases the torsional flexibility of the implant and reduces hoop stresses during insertion. Nails may be cylindrical, cloverleaf, rectangular, or square in shape. The cross-sectional shape affects bending and torsional stiffness.

Locking Capacity

Most intramedullary nails are designed to allow interlocking fixation with screws at both ends of the bone. The great variety in the number, location, and angle of screw holes offers advantages over specific indications. Two screws proximally and two screws distally are usually applied, allowing control in rotation and bending and allowing fixation of comminuted fractures. Oblique screws (usually applied proximally for femoral or humeral nails) reduce the rotational deformation around the screw and alleviate the need for a second screw. Proximal locking capacity in the upper femur has been extensively modified to treat fractures proximally (e.g., initially the Y Kuntscher nail was used; more recently the Russell and Taylor reconstruction nail has been used) and from the subtrochanteric region.

Dynamic interlocking fixation

This technique involves the use of proximal interlocking screws only. In this type of fixation there is control in bending and rotation of the bone around the fracture, but the bone transfers the main axial load. Dynamic fixation is used in axially stable fractures and in some nonunions, and potentially increases the fatigue life of the nail.

Static interlocking fixation

This technique involves the use of proximal and distal interlocking screws. It controls rotation, bending, and axial loading, and may reduce the fatigue life of the implant as the nail becomes the load-bearing device. Static fixation is used in comminuted and metaphyseal fractures.

Dynamization of static interlocking fixation

This technique involves conversion of static to a dynamic fixation by removing the interlocking screws from the longest fragment. This is necessary in nearly 2% of femoral fractures. Dynamization is indicated in nonunions because it potentially increases the axial compression forces at the fracture site and augments fracture healing. Dynamization increases the fatigue life of the implant as its load-bearing capacity is decreased.

Systemic Effects of Intramedullary Nailing

Effect of reaming and nail insertion on blood flow, fracture union, and bone porosity

Normal bone blood flow is centrifugal from the endosteum to periosteum. Cortical reaming and nail insertion damages the medullary endosteal blood flow, resulting in avascularity of significant portions of the diaphyseal cortex. On the other hand, reaming increases the periosteal blood flow, which is not dependent on the nutrient artery.

There is a marked increase in blood flow during the first two weeks following reaming. After six weeks, flow in cortical bone recovers to normal when periosteal blood flow remains high, because the normal centrifugal pattern of flow is reversed.^[5] The disturbance in vascular supply is the basis of osteoporosis of cortical bone seen after intramedullary nailing rather than the stress shielding.^[6] Reaming also weakens the bone and the recommended bone removal (being no more than 4 cm and no more than half the original cortical thickness).

Fracture healing proceeds mainly by the formation of periosteal callus, which is largely independent of endosteal blood supply. Reaming products (fibroblasts, bone, and mesenchymal stem cells) induce bone on subperiosteal surface and augment the periosteal callus.^[7]

Effects of reaming and nail insertion in intramedullary pressures; fat intravasation and pulmonary complications

In experimental animal models, pressure can increase as much as 300 mmHg (and in some cases has increased up to 1000 mmHg). In patients with femoral fractures, elevations of intramedullary pressure between 140–830 mmHg have been observed. Fat embolization is a well-known complication of intramedullary nailing of the lower limbs. There is some controversy in the literature with regard to the extent of fat embolization that results from the reamed vs. unreamed technique. It seems logical, however, to consider that reaming may cause the accumulated effect of fat emboli, compared with the single-nail passage of the unreamed technique. Bone marrow fat embolization in the lung may cause mechanical obstruction of pulmonary vessels. Several pathogenetic pathways have been suggested (e.g., that the high thromboplastin content of bone marrow may cause increased coagulation or activation of the alternative pathway of disrupted fat cells' coagulation by-products. These mechanisms however, do not seem strong enough to explain all of the observed changes. Some authors therefore have recently postulated the adverse effect of fat in the lung endothelium. Direct or cytokine-mediated toxicity has been postulated, but at present there is insufficient evidence to support one hypothesis or the other. It is known, however, that the use of the unreamed nailing technique evokes a reduced systemic inflammatory response and fewer embolization effects.^[8]

EXTERNAL SKELETAL FIXATION

History and Development of the Technique

Two and a half thousand years ago, Hippocrates first described a method of external skeletal fixation for a tibial fracture (Fig. 5). Proximal and distal frame-type wooden rings were interconnected with long double bars on each side of the leg. This configuration permitted wound inspection and allowed fracture healing. Malgaigne in 1853 described a claw device for patellar fractures. Clayton Parkhill, innovator of the modern version of an external fixator, in 1897 presented to the

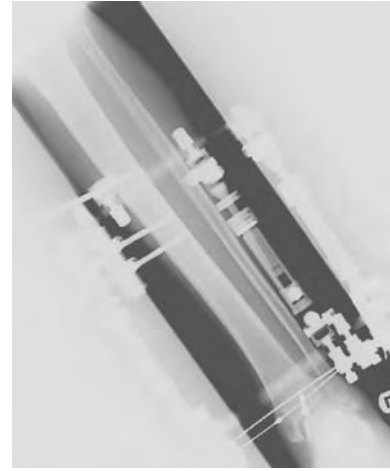


Fig. 5 Hybrid external fixation of a distal tibial fracture.

American College of Surgeons a device designed with two pins above a fracture line and two below, joined externally by a clamp. Lambotte in 1909 and Humphry in 1917 were the first advocates of the use of threaded pins. During the Second World War, external fixation devices became popular among army surgeons, although they fell into disrepute in North America because of their lack of rigid fixation and due to pin track infection. Charnley in 1948 popularized an external compression device for joint arthrodesis, that has remained attractive to this day. Hoffmann from Switzerland in 1954 presented his successful external fixation device with multiplanar frames that increase stability; Hoffman's external fixation technique has become a standard in modern orthopedic surgery. In 1970, Dr. Giovanni De Bastiani from Verona, Italy, introduced the dynamic axial fixator. To facilitate union, the telescopic portion of De Bastiani's unilateral fixator can be unlocked to provide dynamic compression. In 1980 Professor Ilizarov from Kurgan, Siberia, treated Italian explorer Carlo Mauri for the infected nonunion of his tibia. Since then, the Ilizarov method of external fixation and the distraction osteogenesis technique first popularized by Professor Ilizarov became significant parts of the orthopedic armamentarium. See Fig. 5.^[9]

Biology of Fracture Healing Following External Fixation

Fracture healing and construct rigidity

Knowledge of the basic biology of fracture healing and of the biomechanical principles of the external fixation device are necessary to obtain the full benefit of the technique. Fracture healing is divided into two patterns: primary and secondary healing. Primary

healing is achieved by callusless union, following rigid stabilization of the fracture to prevent interfragmentary motion. In contrast, secondary healing is achieved by periosteal and endosteal callus formation as a result of interfragmentary strain or motion following less rigid fixation. Remodelling of the bone cortex, however, takes time, and rigid fixation devices (i.e., plates) may need to remain in situ for up to 18 months before removal. Thus, rigid fixation not only precludes the development of callus, but also results in the protracted biomechanical dependency of fracture healing on the fixation system. Early external fixators employed rigid frames in multiplanar configurations to achieve primary healing. Although they did occasionally achieve anatomical restoration, it is now known that these techniques may have delayed or even prevented union. Current external fixation systems, while maintaining alignment and length, allow micromotion at the fracture site and promote callus formation (secondary healing).^[9]

Dynamization of an external fixator

Dynamization is the conversion of an external fixator from a statically locked device to a more load-sharing construct to promote micromotion at the fracture site. In current practice, dynamization is usually recommended after the initial stages of fracture healing. Progressive closure-type dynamization is achieved by sliding a telescopic component of the fixator body and approximating the bone ends. This controlled load-sharing eliminates fracture gaps; enhances callus formation; and accelerates maturation, union, and remodelling.

Elements and Design of External Fixators

External fixators are composed of a bone anchorage system in the form of wires or pins, articulations, and longitudinal supports.^[10] The two types of external fixator are pin and ring. Pin fixators are further subdivided into simple fixators, which provide independent application of single pins, and clamp fixators, which permit spatially constrained pin clusters.

Pin fixators can be used in four basic configurations: 1) a unilateral frame with a support device and half-pins on one plane constitutes a unilateral one-plane configuration; 2) the addition of a second plane of half-pins forms a unilateral two-plane configuration; 3) trans-fixation pins connected to support devices at either side are used to build the bilateral one-plane configuration; and the addition of a second plane of half-pins (or possibly transfixion pins) forms a bilateral two-plane configuration.

Ring fixators consist of complete or partial rings connected by rods or articulated devices. Rings are attached to bone with wires or half pins. Pinless fixators have been developed, designed to be attached to the bone by clamps rather than pins or wires. These devices have proven strong enough for temporary fixation in cases of emergency stabilization. They are relatively easy to apply and prevent pin loosening, pin track infection, and neurovascular injury.

Hybrid fixators combine wire and half-pin fixation. These devices are used for intraarticular fractures (wires) with diaphyseal extension (pins). Ilizarov external fixators employ complete modular rings with tensioned plain or olive-tipped wires. Although they are only 25% as stiff axially as unilateral fixators, Ilizarov external fixators have similar stiffness in bending and shear. This satisfies the main biomechanical principle of secondary fracture healing by enhancing micromotion and callus formation. Wire diameter and tension are the most important factors affecting frame stability.

Ilizarov external fixators are particularly useful in high-energy injuries because they create minimal surgical trauma, in malunions and nonunions as well as in osteomyelitis. Limb deformities and discrepancies in length can be corrected by the technique of distraction osteogenesis as well as by salvage arthrodesis of the knee, ankle, and hind foot.

Biomechanical Principles of External Fixation

Half-threaded pins

The most significant parameter affecting the stability of an external fixator is the radius of the pin.^[9] The bending stiffness of the pin is proportional to the fourth power of the radius of the implant. However, the diameter of any pin should not exceed 30% of the diameter of bone diaphysis because this increases the risk of fracture. Pins vary in diameter from 2.5 to 6 mm. Optimal pin design strives for greater core diameter, to increase rigidity. Bicortical fixation with the threaded part is preferred, and radial preload techniques can improve screw fixation and prevent loosening. Radial preload can be achieved by first drilling a pilot hole slightly smaller than the root diameter of the pin or by using tapered-root diameter screw design. This results in lower bending stresses at the entry-site cortex and reduced rate of osteolysis and loosening.

Wires

Kirschner wires under tension are used for Ilizarov frames. These wires are 1.5–1.8 mm in diameter.

Increasing wire tension and diameter, and decreasing the diameter of the ring enhance fixation stability. The yield point for a Kirschner wire has been measured at 120 kg/mm², which equals 210 kg for a 1.5 mm wire and 305 kg for a 1.8 mm wire. The optimal wire tension should not be greater than 50 % of the yield strength (or 105 kg for 1.5 mm wire and 150 kg for 1.8 mm wire). The average tension for most trauma applications recommended for Ilizarov frames is 80–90 kg.

Optimal frame configuration

1. Ilizarov frames with tensioned wires positioned at 90° angles to each other increase overall stability, compared with wires applied at smaller angles.
2. Minimizing the ring diameter as much as clinically possible (no less than 2 cm around the entire ring) increases stability.
3. Increasing the number of fixation devices (screws or wires) to the bone, spreading them evenly over the bone, and using large-framed fixators increase stability.
4. When multiplanar frames are used, then delta or triangular configuration offers greater stability.
5. The closer the frame is to the bone, the more stable the construct will be.
6. Fracture reduction reduces stresses at the pin–bone interface, off-loads the implant, and decreases the rates of pin loosening.

Advantages of an External Fixator

1. Provides a method of fixation in Gustilo type II and III open fractures in which other methods of treatment (i.e., internal fixation) would devitalize and contaminate larger areas of soft tissue, increasing the risk of infection and limb loss.^[2]
2. Compression, neutralization, and distraction of fracture fragments are possible.
3. Direct surveillance of the wound is possible, including neurovascular status, skin viability, and muscle compartmental pressure.
4. Additional treatment is possible by means of dressings, skin grafting, debridement, and bone grafting without distracting the bone fixation.
5. Immediate motion of the proximal and distal joint is allowed.
6. Extremity elevation is allowed by suspension from ropes from overhead frames, aiding edema resolution and relief of the posterior soft tissue.
7. Early patient mobilization is allowed.
8. Application can be performed under local anesthesia if necessary.
9. Can be applied in infected or noninfected nonunions.

10. Can be used for joint arthrodesis.
11. Could correct malunions following appropriate osteotomy.
12. Can be used for limb lengthening procedure or for deformity correction.



Disadvantages of an External Fixator

1. Pin track infection may complicate fixation; requires meticulous care.
2. Requires training on how to mechanically assemble the frame.
3. The frame can be cumbersome, and the patient may reject it for aesthetic reasons.
4. Fracture through pin track may occur.
5. Unless the limb is adequately protected, refracture after frame removal may occur until the underlying bone remodels the disrupted cortex and redevelops original strength.
6. May be expensive.
7. If the fixator needs to bridge a joint, joint stiffness may occur.

Absolute Indications for External Fixation

1. Gustilo type II and III.
2. Fracture associated with burns.
3. Fractures requiring additional plastic or vascular surgical procedures.
4. Fractures requiring destruction (i.e., associated with significant bone loss).
5. Limb lengthening.
6. Arthrodesis.
7. Osteomyelitis and nonunions.

Relative Indications for External Fixation

1. Certain pelvic fractures and dislocations.
2. Open, infected pelvic nonunions.
3. Reconstructive pelvic osteotomy.
4. Following tumor excision with autograft or allograft replacement.
5. Femoral osteotomies in children.
6. Limb reimplantation.
7. Polytraumatized patients.
8. Correction of congenital joint contractures.
9. Supplement to inadequate internal fixation.
10. Ligamentotaxis. Reduction of comminuted intra-articular fractures by traction of the ligaments and capsule around the joint.
11. Fixation of fractures in patients with head injuries having frequent seizures or spasms.
12. Fixation of floating knee fractures (ipsilateral fracture of distal femur and proximal tibia).

CONCLUSION

Fracture fixation remains currently the treatment of choice for bony injuries. Internal or external fixation can be alternatively or in combination employed to stabilize skeletal trauma. However, orthopedic surgery has evolved over the last 25 years. Modern technologies like tissue engineering have come to the front line of research and clinical practice, and more biological methods of treating fracture disease are emerging.

ARTICLES OF FURTHER INTEREST

Bone Cement; Bone–Implant Interface; Bone Plates and Screws, Bioabsorbable; Bone Remodeling; Orthopedic Biomaterials; Orthopedic Fixation Devices; Titanium and Its Alloys

REFERENCES

1. Brinker, M. Basic Sciences. In *Review of Orthopaedics*, 3rd Ed.; Miller, M., Brinker, M., Eds.; W.B. Saunders: Philadelphia, 2000; pp. 1–22, 119–127.
2. Claiborne Christian General Principles of Fracture Treatment. In *Operative Orthopaedics*, 9th Ed.; Canale, T., Ed.; Mosby: St. Louis, 1998; Vol. 3, 2011–2015, 2018–2021.
3. Harkess, J.; Ramsey, W.; Harkess, J. Principles of Fracture and Dislocations. In *Rockwood and Green's Fractures in Adults*, 3rd Ed.; Rockwood, C., Jr., Green, P., Bucholz, R., Eds.; J.B. Lippincott Company: Philadelphia, 1991; Vol. 1, 119–121.
4. Allen, C.W.; Piotrowski, G.; Burstein, A.H.; Frankel, V.H. Biomechanical principles of intramedullary fixation. *Clin. Ortop.* **1968**, *60*, 13–20.
5. Rhinelander, F.W. Effects of medullary nailing on the normal blood supply of diaphyseal cortex. *Clin. Ortop.* **1998**, *350*, 5–17.
6. Perren, S.M. The biomechanics and biology of internal fixation using plates and nails. *Orthopedics* **1989**, *12* (1), 21–34.
7. Tydings, J.D.; Martino, L.J.; Kircher, M.; Alfred, R.; Lozman, J. The osteoinductive potential of intramedullary canal bone reamings. *Curr. Surg.* **1986**, *43* (2), 121–124.
8. Giannoudis, P.V.; Pape, H.C.; Cohen, A.P.; Krettek, C.; Smith, R.M. Systemic effects of femoral nailing. *Clin. Ortop.* **2002**, *404*, 378–386.
9. Nepola, J.V. External Fixation. In *Rockwood and Green's Fractures in Adults*, 4th Ed.; Rockwood, C., Jr., Green, P., Bucholz, R., Heckman, J.D., Eds.; J.B. Lippincott-Raven Publishers: Philadelphia, 1996; Vol. 1, 229–231, 235–240.
10. Behrems, F. A primer of fixator devices and configurations. *Clin. Ortop.* **1989**, *241*, 5–15.

Bone–Implant Interface

D. A. Puleo

Center for Biomedical Engineering, University of Kentucky, Lexington, Kentucky, U.S.A.

B

INTRODUCTION

Placement of endosseous implants has improved the quality of life for millions of people. It is estimated that over 500,000 total joint replacements, primarily hips and knees, and between 100,000 and 300,000 dental implants are used each year in the United States alone. Total joint arthroplasty relieves pain and restores mobility to people such as those afflicted with osteoarthritis, and dental implants provide psychological and aesthetic benefits in addition to improving masticatory function for edentulous patients.

Past achievements in the use of endosseous implants have led to the use of these devices in expanded patient populations. Whereas implants once were placed exclusively in elderly patients who are relatively sedentary as they approach the end of their lifespan, prostheses now are being placed in younger, more active patients. This trend has led to a steady increase in the number of devices implanted. Consider, for example, hip and knee replacements. Knee procedures are growing at an average rate of 14% per year, and hip replacements at 3% per year. In addition, with the “graying” of the U.S. population (meaning that the proportion of the population 55 years or older is steadily increasing), more and more people will be candidates for orthopedic and dental implants.

The success of endosseous implants depends on acquiring and retaining stable fixation of the device in the bony site. In turn, stable fixation largely depends on obtaining intimate apposition of bone to the implant. Several factors—including the particular implant, the surgical procedure (and site of implantation), and the patient—can compromise periprosthetic bone formation. Subsequent to the immediate postoperative phase, in which infection can be a cause of failure, the primary mode of failure is loosening, or loss of fixation between bone and implant. In this case, a soft, fibrous membrane is present at the interface and thereby decreases stability as the implant is mechanically loaded. Therefore, much effort is being devoted to enhancing stable initial fixation at the bone–implant interface.

DEVELOPMENT OF THE BONE–IMPLANT INTERFACE

Following implantation of a biomaterial in bone, events analogous to those that occur during fracture healing will occur. The nature of the bone–implant interface will depend on factors that affect the body’s normal response to injury, which involves inflammation, repair/modeling, and remodeling/maturation. Fig. 1 schematically depicts events at the bone–implant interface.

Insertion of an implant into a surgically prepared site will result in the adsorption of proteins from blood and other tissue fluids. Although still incompletely understood, these adsorbed biomolecules play a significant role in the unfolding of subsequent events at the interface, such as the adhesion and function of cells. Initially, a hematoma will be present between the implant and bone. Fibrin within the coagulated blood serves as a scaffold for the infiltration of inflammatory cells. Cytokines and growth factors secreted by macrophages can stimulate invasion by mesenchymal cells, including osteoprogenitor cells. With adequate chemical, biochemical, and mechanical factors, the osteoprogenitor cells will differentiate into osteoblasts that begin synthesizing woven bone. With time, the woven bone will be remodeled into mature, lamellar bone, which further stabilizes the implant.

As shown in Fig. 1, bone formation in the periprosthetic region can occur in two directions. Not only does the healing bone approach the biomaterial, but bone formation can also begin on the implant surface and extend outward. The terms distance osteogenesis and contact osteogenesis have been used to describe these phenomena.^[1] Distance osteogenesis is similar to normal appositional bone growth, in which the surrounding bone is a source of osteogenic cells and provides a substrate for these cells to produce new bone matrix. The distance between implant and bone decreases as distance osteogenesis proceeds and the newly formed bone approaches the implant. In this case, cells and their matrix will be present between implant and bone.

In contact osteogenesis, bone formation occurs directly on the implant surface and moves outward toward the surrounding bone. For this to happen,

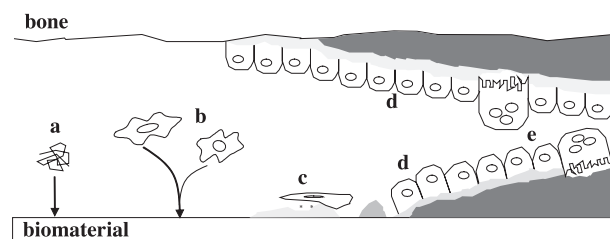


Fig. 1 Schematic representation of the events at the bone–implant interface. **a:** Protein adsorption from blood and tissue fluids. **b:** Inflammatory and connective tissue cells approach the implant. **c:** Formation of an afibrillar mineralized layer and adhesion of osteogenic cells. **d:** Bone deposition on both the bone and the implant surfaces. **e:** Remodeling of newly formed bone.

osteogenic cells must be recruited to the implant surface and subsequently must be induced to differentiate into osteoblasts that will secrete bone matrix. As indicated previously, fibrin provides a pathway for migration of osteogenic cells toward the implant. Once at the surface, these cells initially deposit a collagen-free organic matrix that contains important non-collagenous bone proteins. The matrix provides sites for the nucleation of calcium phosphate crystals and subsequent crystal growth. Finally, a collagenous matrix that mineralizes is secreted adjacent to the collagen-free layer. This process results in the mineralized, collagen-free layer separating periprosthetic bone from the implant, with the interfacial layer being similar to the cement lines found at bone interfaces, such as between osteons.

Extrapolating from the study of dental implants placed in cortical bone of rabbits, Roberts has described the expected time course for development of the bone–implant interface in humans.^[2] The process is divided into four stages: 1) woven callus, 2) lamellar compaction, 3) interface remodeling, and 4) compacta maturation. In stage 1, over a period of about six weeks, woven bone forms to bridge gaps between implant and bone. In stage 2, lamellar bone forms within the woven bone scaffold to provide additional support for the implant by 18 weeks. In stage 3, remodeling of the interface also occurs by 18 weeks to repair damage to periprosthetic bone resulting from the implantation procedure. In stage 4, further maturation and long-term maintenance of the supporting bone bed occur through 54 weeks after implantation.

Osseointegration

The term osseointegration is used frequently when discussing the bone–implant interface. Unfortunately, various investigators use different definitions of the term, which can be confusing when interpreting

and comparing different studies. The concept of osseointegration was developed by Brånemark's group after noticing what appeared to be direct contact between bone and titanium (Ti) bone chambers.^[3] This observation led to nonalloyed (commercially pure) Ti being used to construct root form dental implants. Even Brånemark's use of the term osseointegration has undergone an evolution. In 1977, the term was used to describe intimate apposition of bone to Ti implants without intervening tissue.^[4] Several years later, osseointegration was defined as a "direct structural and functional connection between ordered, living bone and the surface of a load-carrying implant."^[5] Although bone cells and mineralized tissue approach the surface of the implant, regions without direct contact also are observed at higher magnification (Fig. 2). Osseointegrated implants have regions in which they are separated from the surrounding bone by at least a 20–40 nm layer of amorphous, collagen-free material. Considering the many factors affecting bone–implant interactions, osseointegration occurs only under precisely controlled conditions.

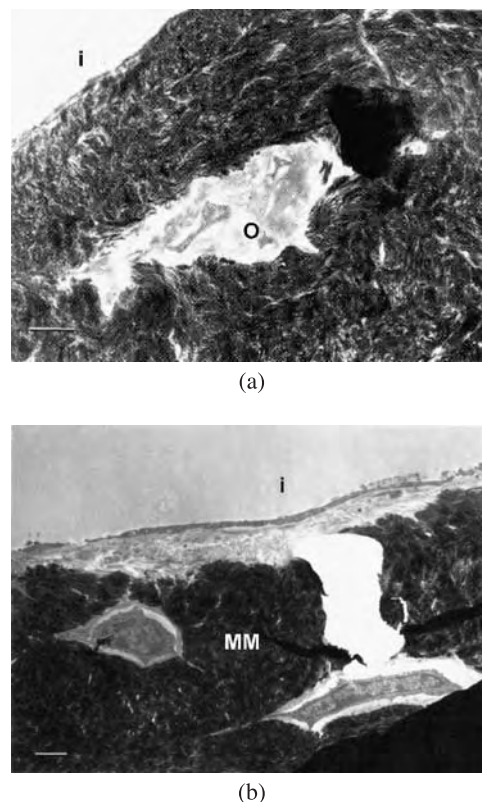


Fig. 2 Transmission electron micrographs showing the varied interface between bone and titanium dental implants. (a) Intimate apposition of mineralized matrix to implant surface. (b) Separation of implant from bone by an unmineralized zone. **i** = implant space, **o** = osteocyte, **MM** = mineralized matrix. Bar = 2 μ m. (From Ref. [5]. Reprinted by permission of John Wiley & Sons, Inc.)

FACTORS INFLUENCING THE BONE-IMPLANT INTERFACE

As indicated in the preceding section, the host response to implants placed in bone involves a series of cell and matrix events that ideally culminate in intimate apposition of bone to biomaterial. For this intimate contact to occur, bone damaged during surgical preparation of the implant site must be repaired, and gaps between bone and implant must be filled. Additionally, unfavorable conditions that disrupt the newly forming tissue and lead to formation of a fibrous capsule must be minimized.

Implant Materials

In an attempt to enhance the stability of endosseous implants, a large number of implant materials and designs have been used. Although a treatment of the various designs intended to alter interfacial mechanics is beyond the scope of this chapter, it is nonetheless important to note that changes in implant design impact biomechanical attributes and thus affect the bone-implant interface.

Bioinert materials

Historically, so-called bioinert materials have been used principally for endosseous implants. These materials include oxide ceramics, such as alumina (Al_2O_3), and metals that form stable surface oxides, such as Co-Cr alloys, commercially pure Ti, and Ti-6Al-4V. Because these surface oxides are chemically stable, it was thought that the implant materials would only minimally affect host tissues and that host responses would minimally affect the material properties.

Although oxide ceramics were important for early research on porous implants (discussed further in the section on cementless implants), their brittleness makes them generally unsuitable for use in highly loaded sites. Consequently, metallic biomaterials are predominantly used for load-bearing prostheses such as hip, knee, and dental implants.

Bioinert materials are not exactly inert, however. Metallic ions and particles are released into surrounding tissues even from Ti and its alloys, which form a stable TiO_2 surface film. For example, metal levels in tissues encapsulating implants have been measured at up to 21 ppm Ti, 10.5 ppm Al, and 1 ppm V around Ti-6Al-4V, and up to 2 ppm Co, 12.5 ppm Cr, and 1.5 ppm Mo around Co-Cr-Mo.^[6] Depending on the amount of metal released, grey to black tissue discoloration may be observed, a phenomenon called metallosis. Because of the blood vessel network adjacent to the implant, metallic species can be

distributed systemically. Even in patients with well-functioning implants, the level of Ti in the serum of recipients of Ti alloy devices may be elevated threefold over levels in control groups. The level of Cr in recipients of Co-Cr alloy implants may be five- and eightfold higher in serum and urine, respectively.^[7] In light of the potential for metallic species to affect the differentiation and activity of bone cells,^[8] they consequently may affect the integration of implants in bone.

Another consequence of the lack of reactivity of bioinert materials is that even though bone will approach, and may even directly contact, the implant surface, bone will not bond to the material. In the case of an osseointegrated implant, as shown in Fig. 2, regions with apparently direct apposition of bone to implant will exist, but they are unlikely to provide significant resistance to shear and tensile loading. Furthermore, there will also be regions where an even weaker electron-lucent zone separates the implant from bone. High resolution immunocytochemical studies have demonstrated that this interfacial layer is rich in noncollagenous proteins such as osteopontin (OPN) and bone sialoprotein (BSP), as well as certain plasma proteins such as $\alpha_2\text{HS-glycoprotein}$.^[9] The absence or relative paucity of serum proteins, such as albumin, in this layer indicates selective accumulation/deposition of molecules at the bone-implant interface.^[9] Because they contain Arg-Gly-Asp and polyacidic sequences, OPN and BSP are believed to play roles in cell adhesion and binding of mineral. With materials other than commercially pure Ti, the interfacial zone becomes thicker. In more extreme cases, a fibrous capsule with a thickness in the microns scale will separate bone from implant. The search for a stronger bone-implant interface has led to the development of bioactive materials.

Bioactive materials

Hench et al. conducted pioneering research in the development of bioactive materials.^[10] Rather than have a material sit passively in tissues, these researchers hypothesized that materials should be designed to have controlled surface reactivity under aqueous conditions. The original Bioglass[®] formulation (45S5) contained 45 wt% SiO_2 , 24.5 wt% CaO , 24.5 wt% Na_2O , and 6 wt% P_2O_5 . By inducing ion exchange reactions, release of calcium and phosphate, alteration of interfacial pH, and nucleation and crystal growth, chemical bonding between bone and biomaterial can be achieved (Fig. 3a). During formation of a bonelike carbonated apatite on the surface of bioactive glasses, bone bonding is enhanced by the interdigitation of collagen fibrils (produced by adjacent osteoblasts) with

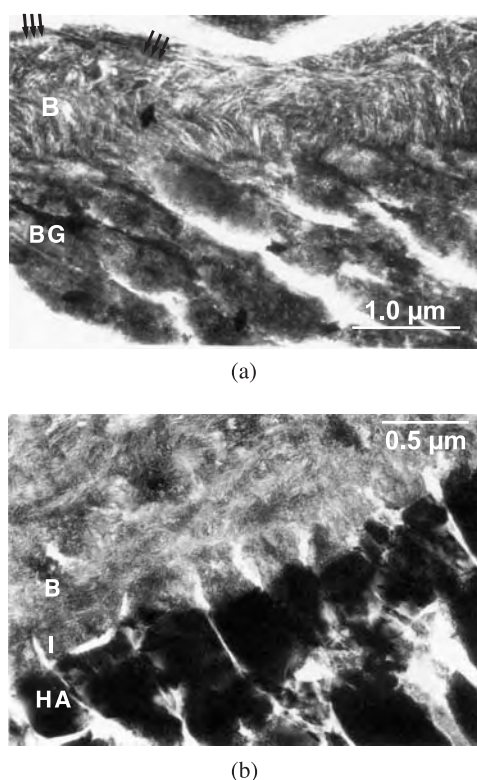


Fig. 3 Transmission electron micrographs showing the interface between bone (B) and bioactive materials. (a) Bioglass[®] (BG). (b) Hydroxyapatite (HA). I = intervening apatite layer. The arrows in frame (a) indicate the banding pattern of collagen fibers in bone. (From Ref. [11]. Reprinted by permission of John Wiley & Sons, Inc.)

the mineralizing layer. By changing the formulation, such as by altering the silica content, the reactivity of the material can be controlled. Furthermore, it is speculated that chemical changes in the interfacial microenvironment stimulate bone cell activity and thus enhance bone bonding.^[12]

The chemical similarity of calcium phosphate (Ca-P) ceramics to the inorganic matrix of bone has made them a subject of intense investigation. The two most commonly studied Ca-P materials are hydroxyapatite [HA; $\text{Ca}_{10}(\text{PO}_4)_6(\text{OH})_2$] and tricalcium phosphate [TCP; $\text{Ca}_3(\text{PO}_4)_2$]. Because these materials are composed primarily of calcium and phosphate ions, which are normally present in the bone microenvironment, they participate in solid-solution equilibria at the interface. Partial dissolution of calcium and phosphate from the implant causes supersaturation of the tissue–implant interface microenvironment with respect to apatite, and thereby results in reprecipitation of bone-like carbonated apatite.^[13] Matrix proteins and glycosaminoglycans secreted by adjacent cells adsorb on, and become incorporated into, the mineralizing surface layer, resulting in bonding with bone (Fig. 3b). A collagen-free bonding zone of up to a micron in thickness

separates the implant from surrounding bone. Because alterations in the microstructure, such as phase composition and crystallinity, affect dissolution and reprecipitation of Ca-P materials, such changes also modify the bioactivity of these materials.

A frequent observation reflecting the ability of bioactive materials to form strong chemical bonds with bone is that attempting to remove them results in fracture of either implant or bone. The fracture lines run across, rather than along, the interface. At weaker interfaces, such as between bioinert materials and bone, failure generally results in separation of the implant from the surrounding tissue.

Modes of Fixation

The lifetime of endosseous implants is dependent on stability of the bone–implant interface. Furthermore, this lifespan can be divided into two parts: the initial period following implantation, during which stable fixation must be obtained; and the remaining time, during which stable fixation must be retained. Several studies have shown that early subsidence of hip implants correlates with, and can be used to predict, eventual failure.^[14] Overall, stable initial fixation is a necessary, although not sufficient, requirement for long-term implant stability. To this end, two approaches have been used to achieve fixation of endosseous implants.

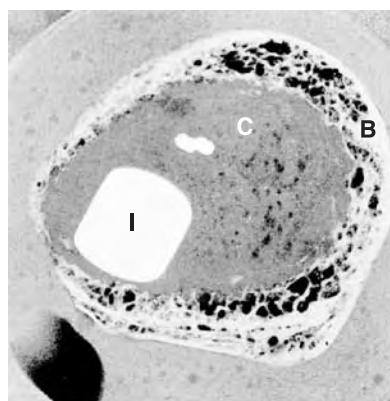
Cemented implants

Cemented implantation involves the use of polymethyl methacrylate (PMMA) as a bone cement to anchor the implant in place. Partially polymerized PMMA has a doughy consistency. It is inserted into the prepared implantation site during surgery, and the implant is subsequently pressed into the still-soft PMMA. As the PMMA continues to cure, a solid layer of polymer is interposed between the implant and surrounding bone. The cement interdigitates with the walls of the surgically prepared bone cavity. As the implant is loaded, forces are transmitted from the implant through the cement and to the surrounding bone. Because cemented implants have two interfaces (one between metal and polymer, and one between polymer and bone), the long-term stability of an implant depends on retaining the integrity of both interfaces.

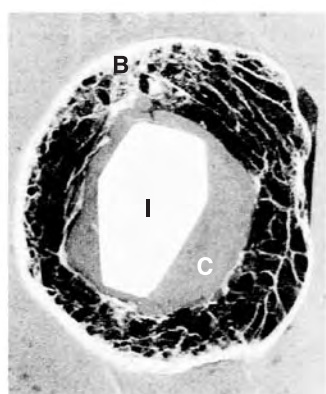
With respect to the cement–implant interface, approaches to both minimize and maximize bonding have been investigated. The chemical nature of the polymer and metallic implant precludes chemical bonding, but mechanical interlocking can be obtained. To minimize interaction, polished surfaces have long been used successfully. In light of reports of debonding

at the interface between PMMA and smooth implants,^[15] prostheses with matte finish surfaces were investigated to allow for mechanical interdigitation. Implants precoated with PMMA to facilitate adhesion between implant and cement also have been tested and used. Rougher implants have shown excellent results, but when debonding does occur, micromotion between the rough surface and polymer can cause abrasion and result in the generation of particulate matter, leading to bone resorption and loosening of the implant.

Analysis of the cement-bone interface of some implants retrieved from humans has shown a thin shell of trabecular bone around the polymer, whereas other samples have a layer of fibrous tissue or fibrocartilage interposed between bone and cement (Fig. 4). In addition to interfacial biomechanics, which will be discussed further in a later section (Additional Factors Affecting Bone Ongrowth or Ingrowth), the cement-bone interface is affected by the mechanical, thermal, and chemical trauma generated during the



(a)



(b)

Fig. 4 Contact radiographs of sections of retrieved cemented femoral stems. (a) Significant contact of bone with the cement. (b) Less bone-cement contact. I = implant; C = cement; B = bone. (From Ref. [16]. Adapted by permission of The British Editorial Society of Bone and Joint Surgery.)

implantation procedure. The surface of the surrounding bone bed is inevitably damaged during the broaching and/or reaming procedures used to prepare the cavity for the implant. Also, the polymerization reaction for PMMA is exothermic, which may expose periprosthetic bone to elevated temperatures. Exposure of bone to a temperature of between 44 and 47°C for 1 min. significantly decreases the bone formation rate and often is considered to be a threshold for thermal necrosis.^[17] Intraoperative measurements have shown a median maximum temperature of 48°C (range 41–67°C) at the interface between PMMA and acetabular bone.^[18] Additionally, the methyl methacrylate monomer and dimethyl-*p*-toluidine initiator can have toxic effects on the surrounding bone cells.

Furthermore, integrity of the cement mantle surrounding the implant is essential for stable support of the prosthesis. Fragmentation of bone cement, resulting from repeated loading during the implant service life, stimulates an inflammatory reaction involving macrophages and cytokines, eventually leading to bone resorption. This phenomenon has been termed osteolysis. The destruction of periprosthetic bone, which is reflected in radiolucent spaces surrounding the implants, allows formation of a soft tissue membrane interposed between the cement and bone. Believing that this type of reaction was unique to bone cement, early reports classified this phenomenon as “cement disease.” Subsequent studies have shown that many types of particulate matter, particularly the ultrahigh-molecular-weight polyethylene commonly used as a bearing surface in joint replacements, can induce osteolysis. For bone cement, mantle thickness is an important factor; thicker layers produce better clinical results.

Cementless implants

Although cemented implants have benefited millions of patients, the loosening that often occurs 10 to 20 years after implantation requires complicated and less successful revision surgery. To expand the use of implants in younger patient populations, noncemented implants were developed. Because the afibrillar mineralized layer at the bone-implant interface would not be expected to have substantial tensile or shear strength, mechanical interdigitation is needed to stabilize the implant. With cementless devices, the original concept was that fixation could be obtained by direct mechanical interlocking between bone and implant; the implant is press-fitted directly into an undersized bony cavity. Undersizing of the cavity is intended to decrease gaps and force initial mechanical interdigitation between the implant and surrounding bone.

Several designs of cementless implants have been explored. The goal of each, however, has been to

encourage mechanical interlocking by promoting bone ongrowth or ingrowth. Macrointerlock surfaces range from those having a matte or textured finish to those that are grit-blasted. Grit-blasted surfaces have fared well in both hip and dental implants. For example, the sandblasted and acid-etched titanium SLA surface used for root form dental implants exhibits good interfacial bone contact, even under loaded conditions.^[19] An alternative approach is the use of microinterlock implants, which have porous surfaces to allow ingrowth of bone. Porous coatings have been produced by sintering or diffusion-bonding metallic beads or fibers to select regions of the implant surface. Plasma-spraying techniques also have been used to deposit rough and porous coatings.

For a cementless implant to become stably fixed, the surface must not simply allow tissue to infiltrate the coating; it must also allow formation of bone within the pore volume (Fig. 5). The optimal pore size for bone ingrowth is still a subject of debate. Early work by Hulbert et al. using oxide ceramics showed that a minimum interconnected pore diameter of approximately 100 μm was needed for adequate bone ingrowth.^[21] Smaller pore sizes allowed incomplete mineralization or even prevented mineralization of the infiltrating fibrous tissue. Later work by Bobyn et al. using metallic implants showed good bone ingrowth with pore sizes between 50 and 400 μm .^[22] More recent discussions have presented the possibility that much smaller pores may allow bone ingrowth when bioactive, rather than bioinert, substrates are used.

A newer cementless approach that may be an adjunct to textured or porous surfaces is to use osteoconductive or osteoinductive coatings, which enhance apposition of bone from existing bone surfaces or which stimulate de novo bone formation, respectively. The most studied of these methods is the use of bioactive calcium phosphate coatings. The ability of Ca-P coatings to enhance bone formation at the tissue-implant interface depends on

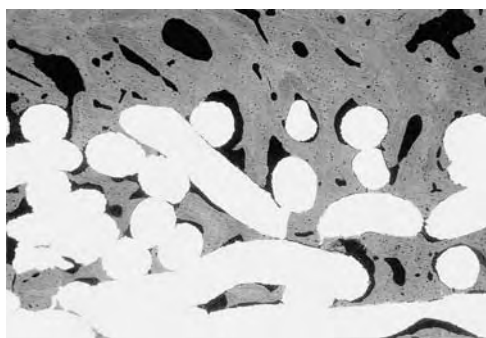


Fig. 5 Backscattered electron micrograph showing good ingrowth of bone into a porous coating. (Ref. [20]. Reprinted by permission of The Journal of Bone and Joint Surgery.)

several factors, including the type of calcium phosphate and its crystallinity, thickness, and adhesion to the implant. The first two factors play a large role in determining the degradability of these Ca-P coatings. Although some dissolution is needed for reprecipitation of carbonated apatite, too-rapid degradation could lead to disintegration of the Ca-P coating. To date, most clinical research has focused on high-density, highly crystalline HA. Examination of implants retrieved from animals and humans reveals apposition of bone within two or three weeks post surgery.^[23] Even though short-term clinical results have been encouraging,^[24,25] the dissolution of coatings as well as cracking and separation from metallic substrates remain concerns.^[26,27] Although still in a research stage, the use of osteoinductive biomolecules, such as the bone morphogenetic proteins, to induce periprosthetic osteogenesis is intriguing.

ADDITIONAL FACTORS AFFECTING BONE ONGROWTH OR INGROWTH

During periprosthetic osteogenesis, anything that affects or interferes with the recruitment, migration, or function of osteogenic cells will affect implant integration. This section will discuss patient-related factors and the surgical procedure, which also play significant roles in determining the nature of the bone-implant interface.

Preparation of the implantation site by reaming, broaching, or drilling, for example, can cause bone damage via mechanical and/or thermal means. Healthy bone cells and their supporting vasculature are needed not only for repair of the damaged bone itself but also for bone formation to proceed toward the implant. Although dental implants are placed in sites that are prepared by drilling but subsequently heal well (become osseointegrated), meticulous surgical care is needed. As mentioned in the section on cemented implants, a temperature of between 44 and 47°C is considered the threshold for thermal damage.^[17] Thus, copious irrigation and controlled drill speed are needed to prevent overheating of the bone bed during preparation of the implantation site.

Gaps between bone and the implant reduce the likelihood of good bone apposition. Dental implants fit well into precisely prepared sites, but joint replacements do not fit nearly as well in the medullary cavity. For example, even though a range of sizes of hip implants is available, geometric differences between the implant and femoral canal permit only sporadic bone-implant contact to be achieved initially.^[28] Furthermore, because bone has limited ability to bridge gaps, excessive space makes good fixation difficult to obtain. For example, although 1.5–2-mm gaps can be spanned

in animal models, the process may take up to three months, even under unloaded conditions.^[29] Loading of the implant (further discussed later in this section) will likely delay bridging, or may even prevent it. Because of the closer proximity of bone to the implant, better ingrowth is observed on the anterior and medial aspects of femoral components. Osteoconductive and osteoinductive coatings are being explored to enhance bridging.

A related factor that affects the fate of implants is the type of bone—compact or cancellous—adjacent to the implant. Implantation into compact bone will provide better primary support of the prosthesis than will placement in cancellous (spongy) bone. However, woven bone, which is rapidly synthesized (up to 50 μm per day), can form between the trabeculae of the cancellous bone to bridge the gap.^[2] Subsequent formation of stronger, but slower growing, lamellar bone within the woven bone framework (lamellar compaction) can provide good secondary fixation.

In the dental implant literature, the quantity and quality of bone are major considerations in surgical planning. Because the clinical success of an implant will depend, in part, on how much of its surface is in intimate contact with bone, quantitative factors, such as the thickness and height of bone present at the site of implantation, are important. Bone quality is less easily determined, but it relates to the degree of bone density present. Type 1 bone provides good anchorage in compact bone, but type 4 bone provides a soft texture that will poorly support an implant. The significance of these differences becomes apparent when the success of implants placed in the mandible, which has higher bone density, is compared to the poorer results of implants placed in the less dense maxilla.^[30]

Excessive motion at the bone–implant interface has been long recognized as a factor that interferes with periprosthetic osteogenesis and thereby results in fibrous encapsulation. Some investigators speculate that relative motion between bone and implant damages the fibrin network and the new vasculature that are part of the early bone healing process, consequently rerouting the healing response into repair by scar tissue. Considering the complexity of the tissue–implant interface and reports of mechanical loading altering cell responses *in vitro* and *in vivo*,^[31] this explanation may be too simplistic. Regardless of the precise mechanisms of failure, a two-stage procedure for placing dental implants has been commonly used, and limited weight-bearing following joint arthroplasty is prescribed. Two-stage dental implantation involves inserting the root form implant and then allowing three to six months for bone healing before the abutment and crown are attached. Similarly, protected weight-bearing of joint implants for up to six weeks after surgery is intended to allow initial bone repair. The

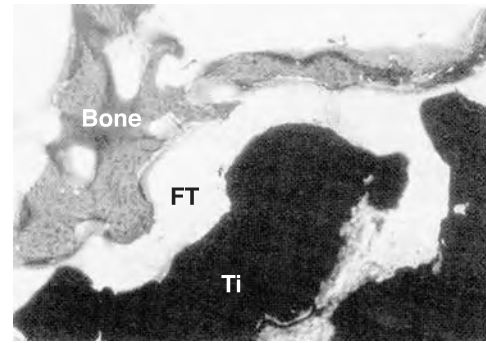


Fig. 6 Photomicrograph of a loaded implant showing fibrous tissue (FT) between titanium (Ti) and bone. (Ref. [33]. Reprinted by permission of John Wiley & Sons, Inc.)

critical amount of relative motion between implant and bone has not been definitely established, but minimization of loading during the early postoperative period is generally accepted as necessary to obtain good initial fixation. Results from animal studies suggest that the threshold for excessive micromotion is between 50 and 150 μm .^[32] Below this level, movement at the interface is tolerated, and bone ingrowth/ongrowth can occur. Motion above this threshold interferes with bone repair and results in fibrous encapsulation of the implant (Fig. 6).

Patient factors may further affect development of the bone–implant interface. Age, health condition, diseases, and medications or treatments that severely compromise healing can be contraindications for receiving an implant. These conditions may include metabolic bone diseases, such as osteoporosis or Paget's disease, and radiation therapy following tumor resection. Uncontrolled diabetes, chronic smoking, and systemic administration of glucocorticoids also can interfere with periprosthetic bone repair.

CONCLUSION

Long-term success of endosseous implants depends on obtaining and maintaining a stable bone–biomaterial interface. Significant advances have been made in understanding events at the interface between bone and implants. Advances in biomaterials and their processing, implant design, and surgical technique have improved the success of endosseous implants. Under certain circumstances, these implants can become intimately apposed by bone. However, numerous factors can shunt the bone repair process to a situation in which the implant becomes encapsulated by fibrous tissue. Further improvements in the performance of endosseous implants will require modifications in implant materials and designs, further manipulating biological responses to stimulate early periprosthetic

bone formation and to subsequently retain the supporting bone tissue.

ARTICLES OF FURTHER INTEREST

Alumina; Bioactive Glass; Bioadhesion; Blood-Material Interactions; Bone Fracture Fixation; Bone Plates and Screws, Bioabsorbable; Bone Remodeling; Corrosion; Dental Implants; Orthopedic Biomaterials; Orthopedic Fixation Devices; Osteogenic Progenitor Cell; Tissue Engineering of Bone; Titanium and Its Alloys

REFERENCES

- Osborn, J.F.; Newsley, H. Dynamic Aspects of the Implant-Bone Interface. In *Dental Implants—Materials and Systems*; Heimke, G., Ed.; Carl Hanser: Munich, 1980; 111–123.
- Roberts, W.E. Bone tissue interface. *J. Dent. Educ.* **1988**, *52*, 804–809.
- Brånemark, P.-I. Introduction to Osseointegration. In *Tissue-Integrated Prostheses*; Brånemark, P.-I., Zarb, G.A., Albrektsson, T., Eds.; Quintessence: Chicago, 1985; 11–76.
- Brånemark, P.I.; Hansson, B.O.; Adell, R.; Breine, U.; Lindstrom, J.; Hallen, O.; Ohman, A. Osseointegrated implants in the treatment of the edentulous jaw. Experience from a 10-year period. *Scand. J. Plast. Reconstr. Surg., Suppl.* **1977**, *16*, 1–132.
- Steffik, D.E.; Sisk, A.L.; Parr, G.R.; Lake, F.T.; Hanes, P.J.; Berkerey, D.J.; Brewer, P. Transmission electron and high-voltage electron microscopy of osteocyte cellular processes extending to the dental implant surface. *J. Biomed. Mater. Res.* **1994**, *28*, 1095–1107.
- Dorr, L.D.; Bloebaum, R.; Emmanuel, J.; Meldrum, R. Histologic, biochemical, and ion analysis of tissue and fluids retrieved during total hip arthroplasty. *Clin. Orthop.* **1990**, *261*, 82–95.
- Jacobs, J.J.; Skipor, A.K.; Patterson, L.M.; Hallab, N.J.; Paprosky, W.G.; Black, J.; Galante, J.O. Metal release in patients who have had a primary total hip arthroplasty. A prospective, controlled, longitudinal study. *J. Bone Jt. Surg., Am.* **1998**, *80*, 1447–1458.
- Thompson, G.J.; Puleo, D.A. Ti-6Al-4V ion solution inhibition of osteogenic cell phenotype as a function of differentiation timecourse in vitro. *Biomaterials* **1996**, *17*, 1949–1954.
- Nanci, A.; McKee, M.D.; Zalzal, S.; Sakkal, S. Ultrastructural and Immunocytochemical Analysis of the Tissue Response to Metal Implants in the Rat Tibia. In *Biological Mechanisms of Tooth Eruption, Resorption and Replacement by Implants*; Davidovitch, Z., Mah, J., Eds.; Harvard Society for the Advancement of Orthodontics: Boston, 1998; 487–500.
- Hench, L.L.; Splinter, R.J.; Allen, W.C.; Greenlee, T.K., Jr. Bonding mechanisms at the interface of ceramic prosthetic materials. *Biomed. Mater. Symp.* **1972**, *2*, 117–141.
- Neo, M.; Kotani, S.; Nakamura, T.; Yamamuro, T.; Ohtsuki, C.; Kokubo, T.; Bando, Y. A comparative study of ultrastructures of the interfaces between four kinds of surface-active ceramic and bone. *J. Biomed. Mater. Res.* **1992**, *26*, 1419–1432.
- Xynos, I.D.; Hukkanen, M.V.; Batten, J.J.; Buttery, L.D.; Hench, L.L.; Polak, J.M. Bioglass 45S5 stimulates osteoblast turnover and enhances bone formation in vitro: Implications and applications for bone tissue engineering. *Calcif. Tissue Int.* **2000**, *67*, 321–329.
- Jarcho, M. Calcium phosphate ceramics as hard tissue prosthetics. *Clin. Orthop.* **1981**, *157*, 259–278.
- Kim, Y.H.; Kim, V.E. Early migration of uncemented porous coated anatomic femoral component related to aseptic loosening. *Clin. Orthop.* **1993**, *295*, 146–155.
- Jasty, M.; Maloney, W.J.; Bragdon, C.R.; O'Connor, D.O.; Haire, T.; Harris, W.H. The initiation of failure in cemented femoral components of hip arthroplasties. *J. Bone Jt. Surg., Br.* **1991**, *73*, 551–558.
- Kwong, L.M.; Jasty, M.; Mulroy, R.D.; Maloney, W.J.; Bragdon, C.; Harris, W.H. The histology of the radiolucent line. *J. Bone Joint Surg. Br.* **1992**, *74*, 67–73.
- Eriksson, R.A.; Albrektsson, T. The effect of heat on bone regeneration: An experimental study in the rabbit using the bone growth chamber. *J. Oral Maxillofac. Surg.* **1984**, *42*, 705–711.
- Wykman, A.G.M. Acetabular cement temperature in arthroplasty. *Acta Orthop. Scand.* **1992**, *63*, 543–544.
- Cochran, D.L.; Schenk, R.K.; Lussi, A.; Higginbottom, F.L.; Buser, D. Bone response to unloaded and loaded titanium implants with a sandblasted and acid-etched surface: A histometric study in the canine mandible. *J. Biomed. Mater. Res.* **1998**, *40*, 1–11.
- Urban, R.M.; Jacobs, J.J.; Sumner, D.R.; Peters, C.L.; Voss, F.R.; Galante, J.O. The bone-implant interface of femoral stems with non-circumferential porous coating. *J. Bone Joint Surg. Am.* **1996**, *78*, 1068–1081.
- Klawitter, J.J.; Hulbert, S.F. Application of porous ceramics for the attachment of load bearing internal orthopedic applications. *Biomed. Mater. Symp.* **1972**, *2*, 161–229.
- Boby, J.D.; Pilliar, R.M.; Cameron, H.U.; Weatherly, G.C. The optimum pore size for the fixation of porous-surfaced metal implants by the ingrowth of bone. *Clin. Orthop.* **1980**, *150*, 263–270.
- Lee, T.M.; Wang, B.C.; Yang, Y.C.; Chang, E.; Yang, C.Y. Comparison of plasma-sprayed hydroxyapatite coatings and hydroxyapatite/tricalcium phosphate composite coatings: In vivo study. *J. Biomed. Mater. Res.* **2001**, *55*, 360–367.
- Dhert, W.J. Retrieval studies on calcium phosphate-coated implants. *Med. Prog. Technol.* **1994**, *20*, 143–154.
- Geesink, R.G.; Hoefnagels, N.H. Six-year results of hydroxyapatite-coated total hip replacement. *J. Bone Jt. Surg., Br.* **1995**, *77*, 534–547.

26. Bauer, T.W.; Taylor, S.K.; Jiang, M.; Medendorp, S.V. An indirect comparison of third-body wear in retrieved hydroxyapatite-coated, porous, and cemented femoral components. *Clin. Orthop.* **1994**, *298*, 11–18.
27. Bloebaum, R.D.; Beeks, D.; Dorr, L.D.; Savory, C.G.; DuPont, J.A.; Hofmann, A.A. Complications with hydroxyapatite particulate separation in total hip arthroplasty. *Clin. Orthop.* **1994**, *298*, 19–26.
28. Noble, P.C.; Alexander, J.W.; Lindahl, L.J.; Yew, D.T.; Granberry, W.M.; Tullos, H.S. The anatomic basis of femoral component design. *Clin. Orthop.* **1988**, *235*, 148–165.
29. Boby, J.D.; Pilliar, R.M.; Cameron, H.U.; Weatherly, G.C. Osteogenic phenomena across endosteal bone-implant spaces with porous surfaced intramedullary implants. *Acta Orthop. Scand.* **1981**, *52*, 145–153.
30. Adell, R. Long-Term Treatment Results. In *Tissue-Integrated Prostheses*; Brånemark, P.-I., Zarb, G.A., Albrektsson, T., Eds.; Quintessence: Chicago, 1985; 175–186.
31. Raab-Cullen, D.M.; Thiede, M.A.; Petersen, D.N.; Kimmel, D.B.; Recker, R.R. Mechanical loading stimulates rapid changes in periosteal gene expression. *Calcif. Tissue Int.* **1994**, *55*, 473–478.
32. Szmukler-Moncler, S.; Salama, H.; Reingewirtz, Y.; Dubruille, J.H. Timing of loading and effect of micro-motion on bone-dental implant interface: Review of experimental literature. *J. Biomed. Mater. Res.* **1998**, *43*, 192–203.
33. Mouzin, O.; Soballe, K.; Bechtold, J.E. Loading improves anchorage of hydroxyapatite implants more than titanium implants. *J. Biomed. Mater. Res.* **2001**, *58*, 61–68.

Bone Plates and Screws, Bioabsorbable

Riitta Suuronen

Institute for Regenerative Medicine and Medical School, Tampere University and University Hospital, Tampere, Finland

Christian Lindqvist

Department of Oral and Maxillofacial Surgery, Helsinki University and University Hospital, Helsinki, Finland

INTRODUCTION

The possibility of using biodegradable devices in bone surgery has been studied intensively for more than 40 years. The first clinical trials of use of such materials in fracture surgery took place more than 30 years ago, with disappointing results. The stability of the devices was not adequate, and osteosynthesis did not result in rigid, function-stable fixation. Research since the early 1970s has resulted in major improvements, however. Advances in polymer chemistry and material processing procedures has improved the quality of biodegradable materials to the extent that clinical use has become safe: Stability has been achieved and maintained for a sufficiently long time to allow normal bone healing to occur.

In the human skeleton, the facial bones seem to be optimally suited for the use of biodegradable devices. An osteotomy or sagittal mandibular fracture can be repaired using biodegradable screws that allow free postoperative mandibular movement. Some other fractures can be treated and certain reconstructions undertaken with resorbable plates and screws, requiring no maxillo-mandibular fixation (MMF). Considering their value in achieving the goal of avoiding MMF in surgery of the facial bones, the question remains, why biomaterials are not used more extensively in maxillofacial surgery today.

One explanation relates to the fact that the devices available are fairly large, and some are inconvenient to use. Another explanation for the limited use of biomaterials in surgery of the facial bones is that many surgeons consider that there remains significant need for improved development of absorbable fixation devices, although some say that bioresorbable screws and plates are excellent solutions for a problem that does not exist. By this is meant that the metallic devices commercially available today are so reliable, stable, and easy to use that there is no need for absorbable materials. For example, it is believed that small titanium plates and screws can be safely left in the body,

even after the implant has ceased to serve any function (i.e., after fracture or osteotomy consolidation). Another factor limiting the use of biodegradable devices may be the adverse reactions noted after use of poly-L-lactic acid (PLLA) or polyglycolic acid (PGA) plates and screws. Sterile skin fistulas and swellings have been described in the vicinity of absorbable plates and screws. The clinical resorption time seems to be unknown, and whether the products of degradation are harmful to patients is also uncertain.

OVERVIEW

Modern titanium plates and screws are very different from devices used at the beginning of the era of rigid fixation. In the beginning, stainless steel plates meant for forearm fractures were used in treating mandibular fractures.^[1] Significant progress has since been made, improving both the quality of materials (i.e., enhancing their biocompatibility), and the design and surfaces of implants. The titanium microplates used in midfacial fracture and osteotomy surgery are excellent examples. Plates and screws as small as these are almost always left in place, with no known need for their removal. It is not known whether it is justifiable to leave plates and screws in place, nor what are the safe limits of plate and screw sizes. Although many manufacturers claim that their implants are made of pure titanium, other metals (e.g., aluminium and nickel) have been found in soft and hard tissues around titanium implants.^[2] Titanium, thought to be biologically inert, has been found in fairly large quantities in lymph nodes and distant organs after the placement of dental implants.^[3] There have also been reports of tumors in the vicinity of metallic prostheses.^[4-7] Postoperative radiation may cause problems in cancer surgery. Metallic devices used in the fixation of access osteotomies have disturbed radiologic examinations. Bioresorbable material can be considered tissue equivalent and does not cause scattering in neighboring structures. Furthermore, the

material does not disturb postoperative imaging, which is extremely important in oncologic patients, for whom early detection of possible recurrence or metastasis is of utmost importance.

In the growing skeleton, restriction of growth caused by metallic fixation devices may lead to consistent asymmetry.^[8] There are also several reports of passive intracranial translocation of metallic hardware, causing great concern and discussion.^[9] Plates, screws, and wires have been found embedded in the dura.^[10]

The foregoing discussion suggests reasons that nonfunctional implants should be removed, at least in growing individuals. Hazards related to removal operations, inconvenience to patients, and costs to patients and society must be considered in determining whether biomaterials can be considered safe to use.

Biomaterials used today all degrade to CO₂ and water, but absorption times differ significantly and depend on various factors. If degradation is rapid, tissues may not be able to tolerate some products of degradation. Swelling, sterile skin, or mucosal fistulas may develop. Traces of biomaterials have been found in tissues several years after implantation, a period of time that is excessive in relation to fracture healing. Currently used copolymers seem to disappear totally within a period of two years, although some investigators claim that the materials never totally degrade. Despite this surgeons universally regularly use bioresorbable suture materials, which they leave in place without hesitation.

It should be a goal that bioabsorbable screws and plates disappear soon after fracture consolidation. Degradation should occur in a way that makes it possible for tissues to tolerate end products. Various factors, including the material's means of sterilization, affect degradation rate. Research proceeding in laboratories and clinics worldwide will overcome these problems associated with use of biodegradable devices. One such problem relates to the size of the devices. Recent advances in materials research have made it possible for screws, rods, pins, and tacks to be significantly smaller than before. Bioabsorbable screws for mandibular osteotomy fixation are nearly as small as metallic screws in clinical use. Miniplates and miniscrews are likely to be of size and strength similar to those made of titanium.

Various chemicals can be incorporated into implants, increasing indications for implant's use. It would be useful, especially when working in contaminated areas, if antibiotics could be incorporated into screws and membranes. Growth factors and enzymes could also be incorporated into implants, to accelerate healing and bone growth.

This article describes the basic research and recent progress in the use of bioresorbable materials in the craniofacial region.

MATERIALS USED

Materials in use in the beginning of the era bioabsorbable materials were mainly homopolymers of poly-L-lactide (PLLA). Unfortunately, the first promising reports on the use of poly-L-lactide devices were soon followed by reports of long-term adverse reactions during resorption.^[11] These reports created doubt in the international craniomaxillofacial surgery community, which led to further research on both biomaterials engineering and the behavior of these materials in vivo.

The self-reinforcing (SR) method developed by Törmälä and Rokkanen represents a cornerstone in circumventing some of the problems associated with use of the early polymers.^[12] See Fig. 1. This internationally patented SR technique entails sintering the polymer material to yield a composite in which matrix and reinforcing elements are of the same polymer. With this technique, it is possible to avoid using large implants, while producing strong implants that provide adequate mechanical stability to the healing bone. This technique also solves the problem of the increased relative crystallinity of polymers during their in-vivo degradation, which some have speculated as causing the complications seen in highly crystalline polymers such as pure poly-L-lactides (PLLA). The advent of the SR technique has also made it possible to obtain strong, malleable devices from the amorphous poly (L/DL) lactide (PLDLA). Without the SR technique, plates must be heated before bending to correspond to the underlying bony structure. Plates manufactured in the self-reinforcing technique can be bent with pliers at room temperature, making them easy to use. The use of these plates and screws has gradually become accepted worldwide, even though several surgeons have

B



Fig. 1 Self-reinforcing structure is seen in an electron microscopy image. The long fibers are protruding out of the fractured surface of an SR-PLLA rod.

questioned the mechanical strength of the bioabsorbable devices, especially in the load-bearing bone of the craniofacial skeleton, the mandible. In 2002, Maurer et al.^[13] published a finite element study on the strength of several different bioabsorbable screws in the fixation of sagittal osteotomies. They concluded that, as concerns chewing forces, all materials studied were sufficiently stable at the osteotomy gap.

Different materials are available today, both self-reinforced and non-self-reinforced. However, most are copolymers of D- and L-lactic acid (PLDLA) or PLA and PGA (PLGA). Currently, most devices are manufactured by combining two polymers (mainly poly-L-lactide with D-lactide or polyglycolide). This combining of two polymers reduces the crystallinity of the material (considered to be a cause of the clinically reported adverse reactions), and avoids the long degradation time of PLLA homopolymers.

EXPERIMENTAL STUDIES

The use of bioabsorbable plates and screws was first reported in the 1970s.^[14] The plates and screws were used in the fixation of mandibular osteotomies in sheep^[15,16] and in the fixation of frontal bone craniotomies in lambs.^[17]

Suuronen and coworkers undertook a series of studies on the use of SR-PLLA plates and screws for fixation of mandibular osteotomies in 92 sheep.^[15,16] The SR-PLLA plates and screws were compared with similar metallic devices for fixation. SR-PLLA screws, SR-PLLA plates and screws, and SR-PLLA plates and metallic screws were used in both the condylar neck and mandibular body areas. (See Fig. 2.) The SR-PLLA plates were composed of four plate layers. SR-PLLA screws or SR-PLLA plates and metallic screws provided sufficient fixation of mandibular bone osteotomies. The body osteotomies healed with heavy callus formation in both groups. The plate and screw

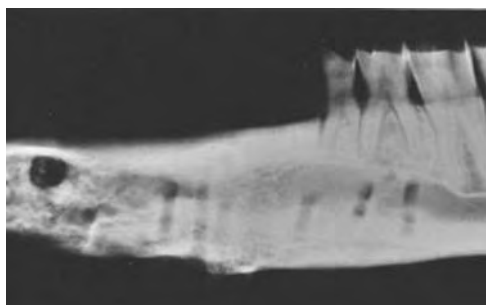


Fig. 2 X ray showing osteotomy on the mandible of a sheep, fixed with SR-PLLA plate and screws. Only the screw holes can be seen in the X ray.

prototypes used in these studies were too large for clinical applications.

Peltoniemi et al. used SR-PGA miniscrews and PLDLA plates or titanium miniplates and screws to fix frontal bone craniotomies in lambs.^[17] During the follow-up period (4 to 104 weeks), no clinically manifested foreign body reactions or differences in consolidation were observed. Fixation with SR-PGA miniscrews resulted in thicker bone and wider connective tissue-filled bone defects than fixation with titanium devices on the contralateral (control) side. The SR-PGA screws were hydrolyzed and fragmented at 4 to 6 weeks, and were resorbed by 12 weeks. Histologically, PGA degradation was accompanied by a nonspecific foreign-body reaction and initial transient osteolysis with decreased osteoid formation around PGA screw channels, but intense compensatory osteoid formation followed after PGA resorption.

CLINICAL STUDIES

The first clinical reports on the use of resorbable plates and screws were published by Bos et al.^[18] Although the biocompatibility of bioabsorbable sutures is well-documented in the literature, implants used for osteofixation have been associated with certain complications.^[11] However, as a result of research that continued in spite of difficulties, it was possible to resolve some of the complications and regain interest and confidence in these biodegradable devices.

Orthognathic Surgery

Orthognathic surgery of the cranio-maxillofacial (CMF) skeleton is needed to rehabilitate patients with congenital or acquired abnormalities of the facial skeleton. The most common orthognathic procedures are mandibular bilateral sagittal split osteotomy (BSSO) and maxillary Le Fort I osteotomy. Fuente del Campo et al. performed maxillary osteotomies with thin SR-PLLA plates and titanium microscrews with successful outcomes.^[19] The longest follow-up time was 17 months. Apart from two cases, the fixation was stable. No complications were reported except for the exposure of one plate in one patient through the mucosa of the mouth one month postoperatively, which was treated by plate removal.

Suuronen and coworkers reported successful fixation of sagittal mandibular osteotomies with SR-PLLA screws without maxillomandibular fixation (with follow-up exceeding 10 years).^[20] Widening of screw holes during healing has been detected in nearly 25% of screw holes. This phenomenon, however, has not caused any subjective or objective signs or

symptoms. The long-term stability of these orthognathic procedures has been similar to the stability of traditional metallic devices.

Similar results were reported by Turvey and coworkers,^[21] who used SR-PLDLA plates and screws in 70 patients. The patients underwent 194 osteotomies of the maxilla and/or mandible with stabilization of each osteotomy achieved by use of self-reinforced polylactide bone plates and/or screws of size and configuration similar to those of titanium systems. They concluded that patients' acceptance of the material was good (70/74) and that stabilization was accomplished as planned in all patients. Three patients experienced problems that resulted in immediate loosening of bone screws. The remaining 67 experienced no short-term problems (6 to 24 months), and healing progressed uneventfully. In each case, acceptable occlusion and favorable aesthetic changes were noted.

Laine and coworkers recently analyzed all complications in orthognathic procedures carried out in the Department of Oral and Maxillofacial Surgery at Helsinki University Central Hospital during the past 10 years (between November 1991 and November 2001).^[22] During this period we operated on 163 patients who underwent 329 orthognathic osteotomies fixed with bioresorbable (both SR-PLLA and SR-PLDLA) devices. (See Fig. 3.) No postoperative intermaxillary fixation was used. Patients' acceptance was generally excellent and very few complications occurred during the 10-year follow-up. The complications were minor and did not affect the end results of the operations. Minor complications occurred in 14 patients (8.6%). Only one patient (0.6%) experienced a postoperative infection. The other minor complications consisted mainly of dehiscence of the wound and plate exposure, together with granulation tissue in the operation field. The remainder of the complications occurred early in the study, when large screw heads on top of bone irritated the patient and had to be removed. The newly developed devices used in the beginning of the trial led to insufficient fixation and resulted in open bite in three patients (1.8%). (The early devices are no longer used.) Based on the study by Laine et al., it can be concluded that bioresorbable devices are safe to use in orthognathic procedures, although there is a learning curve, as in the introduction of all new methods.

Trauma Surgery

In CMF trauma surgery, selection of patients for bioabsorbable devices is very important. If the patient's cooperation is not adequate (e.g., if he or she is an alcoholic or drug addict), these devices should not be used. The key is that bioabsorbable systems not be

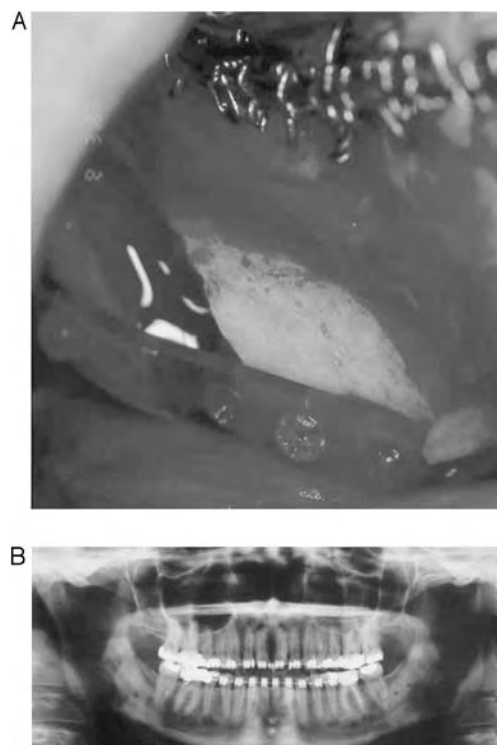


Fig. 3 Orthognathic procedures. (A) Intraoperative view of sagittal osteotomy of the mandible fixed with SR-PLDLA miniplates and screws. The mandible has been advanced 5 mm. (B) Panoramic tomogram showing split on both sides. Because the plate and screws are not radiolucent, only the screw holes can be seen.

used in load-bearing situations; they should be limited to use in load-sharing situations, only.

Self-reinforcing PLLA lag screws have been used in the fixation of anterior mandibular fractures in 11 patients.^[23] Healing in all cases has been uneventful without postoperative maxillomandibular fixation (follow-up exceeds 4 years). Today, SR-PLDLA miniplates are mainly used for mandibular trauma surgery at Helsinki University Central Hospital. (See Fig. 4.) Yerit et al. (2002) followed 22 patients with a variety of patterns of fracture of the mandible, who underwent management with a biodegradable fixation system.^[24] After surgery, maxillomandibular fixation was applied in three cases. The follow-up period averaged 49.1 weeks. The complication rate was low in this study. Mucosal dehiscence on top of the resorbable device was present in two patients. In one of these cases, the material had to be replaced with titanium plates. Mucosal healing and consolidation of the fracture was normal in all other patients. Yerit et al. were able to conclude that self-reinforced biodegradable osteosynthesis materials provide a reliable alternative to conventional titanium plate systems.

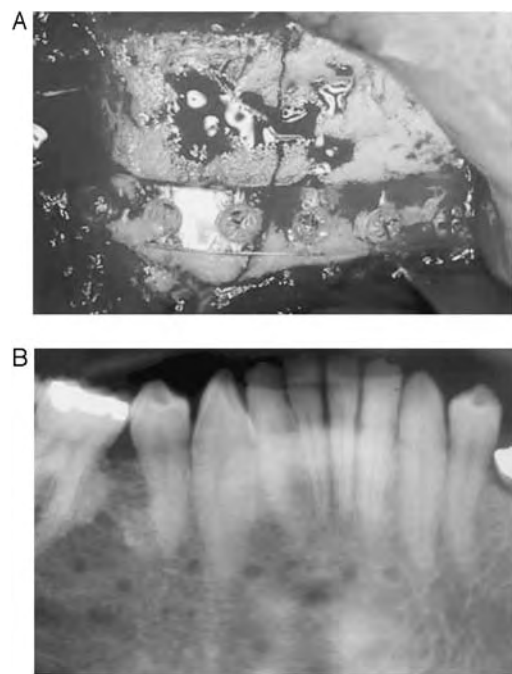


Fig. 4 Trauma surgery. (A) Intraoperative view of anterior mandibular fracture fixed with SR-PLDLA miniplates and screws. (B) Close-up of panoramic tomogram shows good reduction of fragments (visibility not impaired by plates and screws).

Cancer Surgery

To make it possible to remove a malignant tumor from the base of the mouth or tongue, access osteotomies frequently need to be made in the mandible. For such patients, the use of bioabsorbable devices offers many benefits over conventional metallic plates and screws. Bioabsorbable devices are radiolucent and tissue equivalent; hence they do not disturb postoperative imaging or radiotherapy and do not cause scattering of radiation. Dental implants are often needed as post-surgical rehabilitation. Bioabsorbable screws can be used to fix the access osteotomy and, simultaneously, to place the dental implant. (See Fig. 5.)

Suuronen and coworkers reported on 16 access osteotomies fixed with SR-PLDLA miniplates and screws.^[25] With the exception of one access osteotomy carried out at Le Fort I level, all were mandibular swing osteotomies and all patients received postoperative radiotherapy. Follow-up for the 16 patients now exceeds one year. No problems were encountered during the operation, although postoperatively three patients developed problems with fixation and two patients experienced non-union after six weeks, one requiring splinting and one healing spontaneously. Only one patient required another operation due to the failure of fixation.

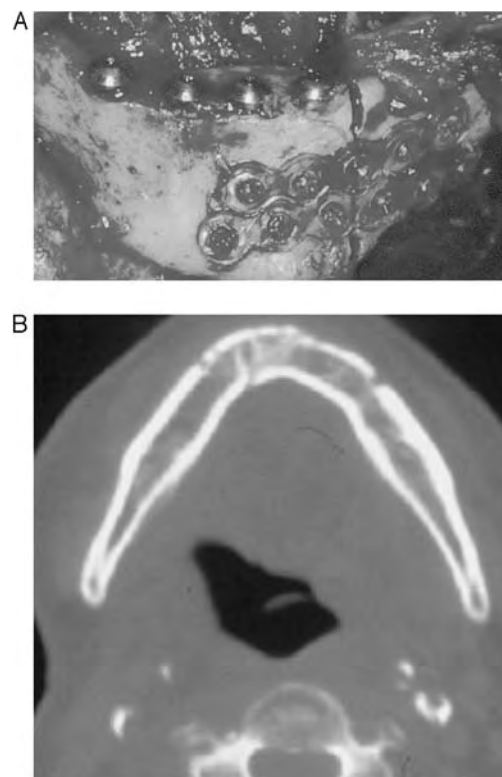


Fig. 5 Oncologic surgery. (A) Intraoperative view of anterior mandibular access osteotomy fixed with SR-PLDLA miniplates and screws. Titanium dental implants were placed simultaneously to facilitate proper rehabilitation of the masticatory system. (B) Bioresorbable plates and screws do not inhibit imaging, enabling early detection of metastasis or recurrence in the region. Only screw holes can be seen in outer cortex.

Fixation of Bone Grafts

Bone grafts are often needed in the mandible and maxilla, either to enable placement of dental implants for dental rehabilitation, or to replace lost bone. Bioabsorbable bone grafts are ideal for bone grafting. (Grafts always resorb to some extent, and metallic screws must be removed because they sit on top of the graft and disturb the patient.) Suuronen et al.^[20] used SR-PLDLA and SR-PLLA lag screws to fix conchondral (rib) grafts to replace a lost mandibular condyle in 10 patients. All recoveries were uneventful without graft displacement.

Pediatric Surgery

The second SR device to be commercially available is SR-PLGA (80/20). Ninkovic et al.^[26] reported on early experience with SR-PLGA miniplates and screws in four patients for osteofixation in cranial remodelling

operations. None of the plates broke, and sufficient fixation stability was achieved without additional support. In a recent review, Suuronen et al.^[20] concluded that the use of SR-PLGA (80/20) should be limited to pediatric use because of its PGA component and because children seem to tolerate rapid resorption better than adults. (In the growing skeleton, rapid resorption is essential to avoid growth impairment.)

CONCLUSION

The use of bioabsorbable self-reinforced plates and screws in CMF surgery is an appealing alternative to titanium devices. They do not disturb growth in the growing CMF skeleton, and hence do not require a second (removal) operation. They do not interrupt postoperative imaging, which is extremely important for early detection of possible metastasis or recurrence. The fact that SR plates and screws gradually lose their strength may even enhance ossification of the osteotomy; however, the use of these devices should be restricted to those operations wherein miniplate fixation is stable enough.

According to reports in the international literature, the plate should not be placed immediately under the incision (especially in the mandible), but should be placed low, under the nonattached gingival tissue in the mandible to enable good soft tissue coverage and blood supply for removal of resorbed polymer. This reduces the risk of dehiscence and makes the plate less palpable and less disturbing to the patient. If the plate is exposed immediately after the operation, revision and closure are necessary. The plate and screws should be carefully examined and any loose material removed to prevent future problems. Good closure of the wound is essential because the plate will not degrade properly if it is not covered by soft tissue.

A new bioresorbable fixating method has recently been introduced that uses tacks and a tackshooter instead of screws.^[27] This new device was used in 15 children who underwent various cranioplasty procedures to demonstrate its reliable application in cranial bone osteofixation. Stable, secure fixation was obtained intraoperatively, with operating time reduced 10–15% compared with earlier use of plates and screws. The use of tacks and a tackshooter reduces operating time and risk of infection and blood loss. This consequently reduces costs, and makes the tackshooter very useful in selected craniofacial cases.

Based on reports in the international literature, it can be concluded that the use of the newest biomaterials presents a learning curve, as occurs with any innovation in methods and materials. Similar to the case when titanium devices were first used, problems with biomaterials are rare and complications can be

regarded as minor. The only difference is the rare occurrence of granulation tissue, which does not require surgical treatment.

ACKNOWLEDGMENTS

The authors wish to thank the State Subsidiary of Finland (TYH1331), the Paulo Foundation, and the European Union Project QLRT-2000-00487 for funding part of this study.

ARTICLES OF FURTHER INTEREST

Bone Cement; Bone Fracture Fixation; Bone-Implant Interface; Dental Implants; Orthopedic Biomaterials; Orthopedic Fixation Devices; Poly(Glycolic Acid); Poly(lactic acid)s

REFERENCES

1. Spiessl, B. *Internal Fixation of the Mandible. A Manual of AO/ASIF Principles*; Springer-Verlag: Berlin, 1989.
2. Moberg, L.E.; Nordenram, A.; Kjellman, O. Metal release from plates used in jaw fracture treatment. A pilot study. *Int. J. Oral. Maxillofac. Surg.* **1989**, *18*, 311–314.
3. Schliephake, H.; Reiss, G.; Urban, R.; Neukam, F.W.; Guckel, S. Metal release from titanium fixtures during placement in the mandible: An experimental study. *Int. J. Oral Maxillofac. Implants* **1993**, *8*, 502–511.
4. Bagó-Granell, J.; Aguirre-Canyadell, M.; Nardi, J.; Tallada, N. Malignant fibrous histiocytoma of bone at the site of a total hip arthroplasty. *J. Bone Jt. Surg.* **1984**, *66-B*, 38–40.
5. Penmann, H.G.; Ring, P.A. Osteosarcoma in association with total hip replacement. *J. Bone Jt. Surg.* **1984**, *66-B*, 632–634.
6. Swann, M. Malignant soft-tissue tumour at the site of a total hip replacement. *J. Bone Jt. Surg.* **1984**, *66-B*, 629–631.
7. Weber, P.C. Epithelioid carcinoma in association with total knee replacement. A case report. *J. Bone Jt. Surg.* **1986**, *68-B*, 824–826.
8. Wong, L.; Richtsmeier, J.T.; Manson, P. Craniofacial growth following rigid fixation: Suture excision, miniplating, and microplating. *J. Craniofac. Surg.* **1993**, *4*, 234–244.
9. Posnick, J.C.; Yaremchuk, M.J. The effects of non-resorbable internal fixation devices placed on and within a child's cranial vault: Brain function, morbidity, and growth restriction [editorial]. *Plast. Reconstr. Surg.* **1995**, *96*, 966–968.
10. Papay, F.A.; Hardy, S.; Morales, L., Jr.; Walker, M.; Enlow, D. "False" migration of rigid fixation appliances in pediatric craniofacial surgery. *J. Craniofac. Surg.* **1995**, *6*, 309–313.

11. Bergsma, E.J.; Rozema, F.R.; Bos, R.R.M.; De Bruijn, W.C. Foreign body reactions to resorbable poly (L-lactide) bone plates and screws used for the fixation on unstable zygomatic fractures. *J. Oral Maxillofac. Surg.* **1993**, *51*, 666–670.
12. Törmälä, P.; Rokkanen, P.; Laiho, J.; Tamminmäki, M.; Vainionpää, S. Material for Osteosynthetic Devices. U.S. Pat. 4,743,257, May 10, 1988.
13. Maurer, P.; Holweg, S.; Knoll, W.D.; Schubert, J. Study by finite element method of the mechanical stress of selected biodegradable osteosynthesis screws in sagittal ramus osteotomy. *Br. J. Oral Maxillofac. Surg.* **2002**, *40*, 76–83.
14. Cutright, D.E.; Hunsuck, E.E.; Beasley, J.D. Fracture reduction using a biodegradable material, polylactic acid. *J. Oral Surg.* **1971**, *29*, 393–397.
15. Suuronen, R. Comparison of absorbable self-reinforced poly-L-lactide screws and metallic screws in the fixation of mandibular condyle osteotomies: An experimental study in sheep. *J. Oral Maxillofac. Surg.* **1991**, *49*, 989–995.
16. Suuronen, R.; Pohjonen, T.; Vasenius, J.; Vainionpää, S. Comparison of absorbable self-reinforced multilayer poly-L-lactide and metallic plates for the fixation of mandibular body osteotomies: An experimental study in sheep. *J. Oral Maxillofac. Surg.* **1992**, *50*, 255–262.
17. Peltoniemi, H.H.; Tulamo, R.M.; Toivonen, T.; Hallikainen, D.; Törmälä, P.; Waris, T. Biodegradable semirigid plate and miniscrew fixation in experimental calvarial osteotomies: A comparative study with rigid titanium fixation. *J. Neurosurg.* **1999**, *90*, 910–917.
18. Bos, R.R.M.; Boering, G.; Rozema, F.R.; Leenslag, J.W.; Pennings, A.J.; Verwey, A.B. Resorbable poly(L-lactide) plates and screws for the fixation of zygomatic fractures. *J. Maxillofac. Surg.* **1987**, *45*, 751–753.
19. Fuente del Campo, A.; Pohjonen, T.; Törmälä, P.; Waris, T. Fixation of horizontal maxillary osteotomies with biodegradable self-reinforced absorbable polylactide plates: Preliminary results. *Eur. J. Plast. Surg.* **1996**, *19*, 7–9.
20. Suuronen, R.; Haers, P.; Lindqvist, C.; Sailer, H. An update on bioresorbable plates in maxillofacial surgery: Invited review. *Facial Plast. Surg.* **1999**, *15*, 61–72.
21. Turvey, T.A.; Bell, R.B.; Tejera, T.J.; Proffitt, W.R. The use of self-reinforced biodegradable bone plates and screws in orthognathic surgery. *J. Oral Maxillofac. Surg.* **2002**, *60*, 59–65.
22. Laine, P.; Kontio, R.; Lindqvist, C.; Suuronen, R. Are there any complications with bioabsorbable fixation devices? 10 years of experience in orthognathic surgery. *Int. J. Oral Maxillofac. Surg.* accepted 2003.
23. Kallela, I.; Iizuka, T.; Salo, A.; Lindqvist, C. Lag screw fixations of anterior mandibular fractures using biodegradable polylactide screws: Preliminary report. *J. Oral Maxillofac. Surg.* **1999**, *57*, 113–118.
24. Yerit, K.C.; Enislidis, G.; Schopper, C.; Turhani, D.; Wanschitz, F.; Wagner, A.; Watzinger, F.; Ewers, R. Fixation of mandibular fractures with biodegradable plates and screws. *Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endo.* **2002**, *94*, 294–300.
25. Suuronen, R.; Laine, P.; Salo, A.; Kontio, R.; Lindqvist, C. Bioabsorbable miniplates in the fixation of access osteotomies. *Int. J. Oral Maxillofac. Surg.* **2003**, *32*, S74.
26. Ninkovic, M.; Rumer, A.; Kofler, A. Preliminary Clinical Experience with Biodegradable Plate and Screw Fixation: The European Union Demonstration Project BMH4-98-3892. In *Treatment of Craniofacial Syndromes with Biodegradable Plates, Screws and Wire [Abstract 6]*, Proceedings of the 2nd Symposium on Biodegradable Materials in Craniofacial Surgery, Saariselkä, Finland, April 8–11, 1999.
27. Spanio, S.; Ashammakhi, N.; Ilomäki, J.; Välimaa, T.; Rainer, C.; Waris, T.; Törmälä, P.; Nincovic, M. Use of new bioabsorbable tacks and a tackshooter in cranial bone osteofixation saves operative time. *J. Craniofac. Surg.* **2002**, *13*, 693–696.
28. Bos, R.R.M.; Rozema, F.R.; Boering, G.; Nijenhuis, A.J.; Pennings, A.J.; Verwey, A.B. Bio-absorbable plates and screws for internal fixation of mandibular fractures. A study in six dogs. *Int. J. Oral Maxillofac. Surg.* **1989**, *18*, 356–369.
29. Gerlach, K.L.; Krause, H.R.; Eitenmüller, J. Use of Absorbable Osteosynthesis Material for Mandibular Fracture Treatment of Dogs. In *Biomaterials and Clinical Applications*; Pizzoferrato, A., Marchetti, P.G., Ravagliori, A., Lee, A.J.C., Eds.; Elsevier: Amsterdam, 1987; 459–464.
30. Törmälä, P.; Vasenius, J.; Vainionpää, S.; Laiho, J.; Pohjonen, T.; Rokkanen, P. Ultra-high-strength absorbable self-reinforced polyglycolide (SR-PGA) rods for internal fixation of bone fractures. In vitro and in vivo study. *J. Biomed. Mater. Res.* **1991**, *25*, 1–22.

Bone Remodeling

Elisabeth H. Burger
Jenneke Klein-Nulend

Department of Oral Cell Biology, Vrije Universiteit, Amsterdam, The Netherlands

Theo H. Smit

Department of Physics and Biomedical Engineering, Vrije Universiteit Medical Center, Amsterdam, The Netherlands

B

INTRODUCTION

Bone remodelling is the biological process whereby living bone tissue renews itself in the course of life. This article explains the details of this process, including the separate activities of osteoclasts (the bone-resorbing cells) and osteoblasts (the bone-forming cells); and discusses recent findings that explain how remodeling leads to mechanically adapted bone tissue, and the role of the mature bone cells or osteocytes in this process. A hypothesis is presented that relates osteocyte apoptosis resulting from disuse, to the loading-aligned tunnelling activity of osteoclasts. Last, the article discusses how adaptive remodelling leads to unwanted bone loss, e.g., around arthroplastic implants and after tooth extraction.

BONE TISSUE

Bone is a living connective tissue that consists of bone cells surrounded by their extracellular matrix (ECM). In all connective tissues, the ECM contains collagen, but in bone, the collagen is mineralized and is therefore hard. The combination of long collagen fibers that are pliable but resistant to stretch, and calcium-phosphate crystals that are brittle but hard, provides the unique material properties of bone ECM: stiff as well as tough, and therefore very strong. A second factor contributes to the strength of bone: the strain-adapted inner structure of bone tissue. In both trabecular bone that has a spongelike structure and in compact bone that is more or less solid, the tissue elements have a tendency to run in one of the major directions of habitual loading. This phenomenon was described by the end of the nineteenth century by Wolff,^[1] and is sometimes called Wolff's Law, or the Law of Bone Remodeling. It makes a skeletal element (a bone) much stronger for its normal tasks than when the tissue material had been evenly distributed, and is therefore an important contribution to skeletal strength. A third factor contributing to skeletal performance is the continuous renewal of tissue during life. Because bone is a living tissue, it is

able to renew itself by cellular activity by removing old, fatigued tissue and replacing it with new material. This process continues throughout the life of the individual, and is called bone remodeling.

BONE REMODELING BY BASIC MULTICELLULAR UNITS (BMUs)

Since the pioneering work of Frost,^[2] it has become generally accepted that bone tissue renews itself by means of basic multicellular units (BMUs). These are groups of bone-resorbing cells (or osteoclasts), and bone-forming cells (or osteoblasts), that act in coordinated fashion to first resorb a small area of existing bone tissue and subsequently refill the gap with new bone. This process is called bone remodeling.^[3] In trabecular bone, consisting of struts and plates, remodeling is a surface phenomenon. The osteoclasts dig a trench along the surface of the strut or plate, which is subsequently filled in by osteoblasts that produce new bone matrix. During the process of filling, some osteoblasts are entombed in the new ECM and become osteocytes. (See the article titled *Histogenesis* in this encyclopedia.) In compact bone, the osteoclasts literally eat themselves deep into the old bone. They dig a tunnel whose tip is called the cutting cone (Fig. 1). The gap made by the osteoclasts is subsequently filled in by osteoblasts, but a small channel is left open in the center of the tunnel where one or more blood vessels grow during the process of refilling (see Fig. 1). These blood vessels ensure that the osteocytes of the new bone that fills the tunnel are close enough to the circulating blood to receive nutrients and discard their waste products. The length of tunnel that is in the process of being filled by osteoblasts is called the closing cone (Fig. 1). The filled-in tunnel is called an osteon or Haversian system, and the central canal is the Haversian channel. In trabecular bone, the filled-in trench is called a (hemi-)osteon, because it resembles an osteon cut in half along its length (Fig. 2). The large pores of the spongelike trabecular bone are filled with

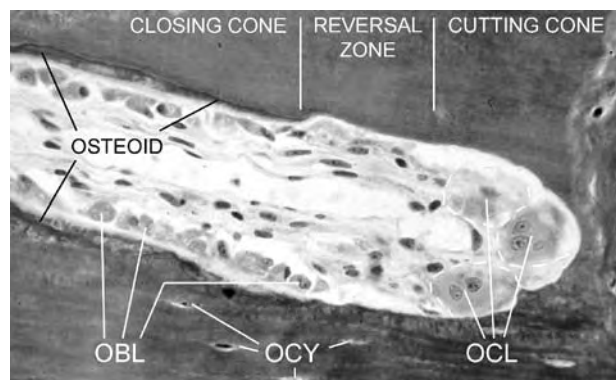


Fig. 1 Longitudinal section of cortical bone illustrating the typical histology of remodeling by a BMU. Osteoclasts (multinucleated giant cells; OCL) erode bone tissue and are followed by osteoblasts (small mononucleated cells; OBL) that refill the gap with new bone tissue. Osteocytes (OCY) are former osteoblasts embedded within the bone matrix. The spherical end of the tunnel, where osteoclasts are active, is called the cutting cone, and the zone where osteoblasts refill the tunnel is called the closing cone. In the reversal zone the bone surface is bare. In the closing cone a narrow channel—the Haversian channel—will be left open to allow space for one or two blood vessels. The filled-in resorption tunnel is the new osteon. (Figure reproduced with permission from Ref. [8].)

loose connective tissue containing many blood vessels; the osteocytes of the new bone are therefore always close to circulating blood. However, it is quite likely that new vessels will also sprout along the remodeling (hemi-)osteon, although the abundance of existing vessels makes their formation less evident than is the case in compact bone.

Because osteons and (hemi-)osteons are replacing preexisting bone tissue, they are called secondary (to distinguish them from the first bone tissue in the developing skeleton). Secondary osteons are thick-walled tubes packed together to make a thick layer of compact bone tissue, whereas (hemi-)osteons are half-tubes packed to form a trabecula, while the other half of the tube consists of loose connective tissue with blood vessels.

THE ADAPTED NATURE OF SECONDARY BONE

As already mentioned, remodeled bone is adapted to the prevailing loads of daily usage. Adaptation means two things: First, the density of the bone tissue relates to the magnitude of habitual loads; second, the osteons and (hemi-)osteons run in the direction of habitual loading. Compact or osteonal bone is found in areas where loads are concentrated; for instance, in the diaphyses of tubular bones. Because mechanical loads are transmitted mostly along the length of the diaphysial

tube, the osteons also run in parallel along the length of the tube.^[4] Trabecular bone is found in areas where the orientation of applied mechanical loads resulting from normal usage is diverse (e.g., in the spine or near the joints of long bones). Because movements involving joints are often complex, it follows that the morphology of trabecular bone around joints is also complex, and that its load-oriented structure is less obvious than in diaphysial compact bone. Recent computer simulations, however, have confirmed that the tissue elements (the (hemi-)osteons of the trabeculae) in trabecular bone are aligned in the dominant loading directions, the same as the osteons in compact bone.^[5] Bone tissue is thus able to renew itself in accordance with the magnitude and orientation of the daily mechanical loads placed upon it, as originally postulated by Wolff.^[1] The question remains how the bone cells perform this task.

STRAIN-RELATED REMODELING

BMUs literally move through preexisting bone tissue during remodeling. In order for the new (hemi-)osteon to align in the direction of strain, the resorption tunnel or trench must be dug along the direction of strain. Through their resorbing activity, therefore, osteoclasts determine not only the width of the future (hemi-)osteon, but also its orientation. The activity of the osteoblasts

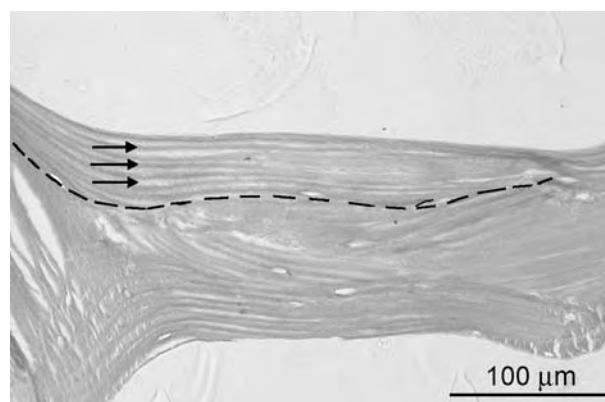


Fig. 2 Histological picture of a remodeled trabecula, showing the longitudinal orientation of secondary (hemi-)osteons. Bone ECM stains pink (or gray in black-and-white reproduction). The dashed line indicates the border of a secondary (hemi-)osteon on the upper side of the trabecula. Its lamellae (indicated by arrows) are oriented along the length of the trabecula, demonstrating that resorption occurred along the longitudinal surface. The resorption trench was subsequently refilled with lamellar bone by osteoblasts. Note the osteocyte lacunae, which show up as flattened white ovals dispersed through the ECM. (Figure reproduced with permission from Ref. [6].)

that follow the group of osteoclasts at the bone resorption site determines the extent to which the defect is refilled. This osteoblastic activity is important because, in the long run, it determines the amount of bone mass. Inadequate refilling during remodelling is the cause of osteoporosis, an important threat to skeletal health in old age. Orientation, however, is determined by the osteoclastic activity, more precisely by the direction in which the team of resorbing osteoclasts moves. So how do osteoclasts know their way; how are they steered by the loading process? A recently developed theory links the very local tissue strains around a progressing BMU to the activation of osteoclasts and osteoblasts, respectively.^[6-9] This theory is based on the concept that the cells involved in mechanosensation are neither osteoclasts nor osteoblasts, but osteocytes.

OSTEOCYTES AND MECHANOSENSING

Osteocytes are the prominent cells of the adult skeleton. They are the only bone cells that actually lie *within* the hard bone ECM. On the surface of adult bone a thin sheath of lining cells is found that covers all bony surfaces, i.e., the periosteum, the trabecular surfaces, and the Haversian channels. Lining cells are only absent from those surfaces where osteoclasts and osteoblasts are active. Both osteocytes and lining cells represent the end stage of differentiation of osteoblasts, the cuboidal cells that produce the bone ECM (see Fig. 3). Osteocytes are left in the mineralizing bone matrix during the process of bone formation, and are therefore the only cells that are truly embedded in the bone ECM. Lining cells are formed when bone formation has finished at a certain surface and the last remaining osteoblasts flatten out to become lining cells (Fig. 3). Osteocytes keep contact with other osteocytes and with lining cells via long, thin, fingerlike processes that run through the ECM and connect to other osteocyte processes by means of gap junctions. Calcification of the ECM always stops before the osteocyte cell body or cell processes are reached. As a result, osteocytes are completely surrounded by a thin sheath of nonmineralized ECM. In deproteinized preparations of bone, the osteocyte cell bodies and their surrounding sheath show up as small holes called lacunae, while the cell processes plus their sheaths show up as tiny canals, the canaliculi of bone. Bone canaliculi are approximately 0.2 micrometer in diameter, about 250 times smaller than the Haversian channels (which are about 50 micrometer in diameter and contain blood vessels). Osteocytes may have up to 80 cell processes (and therefore 80 canaliculi); via these processes they keep contact with neighboring osteocytes as well as with bone lining cells. Therefore, compact as well as trabecular bone may be considered as a three-

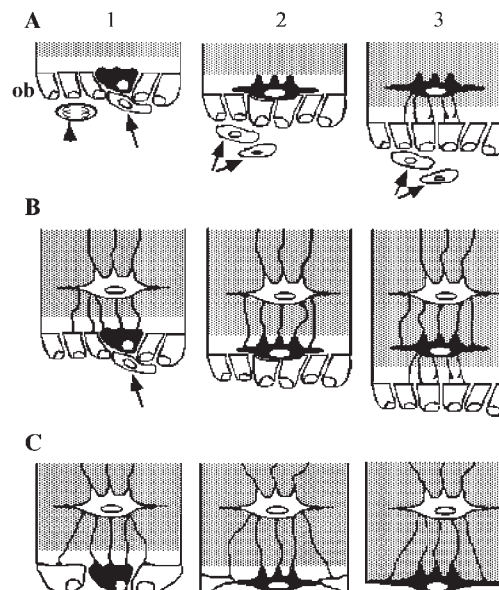


Fig. 3 Schematic presentation of the growth of bone tissue. In A and B, a quiescing osteoblast (black cell) turns into an osteocyte because its neighbors continue to produce bone ECM. Note that during rapid growth (A), proliferation of osteoblast progenitor cells (arrows) ensures that new osteoblasts can be recruited. In B, bone growth slows down as cell proliferation stops. In C, bone growth has stopped completely. The last echelon of osteoblasts stops producing ECM and turns into a layer of lining cells. No new osteoblasts are recruited at this site. Note that when the ECM mineralizes (indicated by gray stippling), the osteocytes remain in contact with the bone surface by means of thin cell processes in canaliculi (thin black lines). (Figure reproduced with permission from Ref. [12].)

dimensional network of osteocytes with bone ECM in the interstices and lining cells on the surface. The function of this cellular network throughout a piece of bone was until recently unclear; recent studies, however, suggest that it functions as the mechanosensory system that regulates and orchestrates adapted bone remodeling. First, the anatomical location of the osteocytes is perfect for sensing strains throughout the tissue. Animal studies have shown that mechanical loading activates several cellular processes in osteocytes, including energy metabolism, gene activation, growth factor production, and matrix synthesis.^[10-12] In addition, comparative studies of the responsiveness of isolated osteocytes, osteoblasts, and periosteal preosteoblasts to mechanical strain in-vitro, showed that osteocytes were most responsive. Of the other two cell types, osteoblasts were more responsive than preosteoblasts.^[13-15] Thus, in the course of differentiation from immature preosteoblast via osteoblast to osteocyte, bone cells increase their sensitivity to mechanical strain; this development is suggestive of a specific role for osteocytes in mechanosensing. Even so, the manner

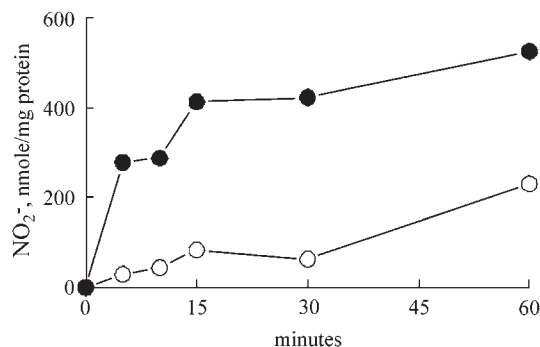


Fig. 4 Effect of pulsating fluid flow (5 Hz sigmoidal pulses, 0.6 ± 0.03 Pa shear stress) on NO production by cultured human bone cells obtained from a transiliac bone biopsy. After 1 hr with or without flow, the cell culture medium was assayed for NO (measured as NO₂). Black dots are stressed cells; white dots are nonstressed controls. Note that after 15 minutes the initial rapid stimulation of stressed cells stopped. (Figure reproduced with permission from Ref. [25].)

whereby osteocytes sense the strains of the mineralized matrix remained obscure because of the very small strains in bone during daily loading (as compared to, e.g., muscle tissue). In healthy bone, the strains resulting from physiological loading do not exceed 0.3%, and are usually of the order of 0.05%,^[16] whereas cultures of isolated bone cells need strains of the order of 1 to 10% to evoke a response.^[17] This apparent contradiction was overcome by the concept of strain-derived canalicular fluid flow as the physical mediator of mechanosensing by osteocytes in bone tissue.^[18,19] According to this view, the osteocytes of loaded bone are activated by the flow of extracellular fluid through the network of canaliculi and lacunae. When bone is loaded, extracellular fluid is squeezed through the thin layer of nonmineralized matrix surrounding the osteocytes toward the bone surface. This flow, rather than the small bulk strain of the loaded bone, produces fluid shear stress at the osteocyte membrane that activates the osteocyte. Several biological studies support this concept. Animal studies show that bone formation in response to loading is sensitive to strain rate, and that large strains alone are not sufficient to activate bone cells, suggesting that interstitial fluid flow is the physical stimulus involved.^[20,21] In cell culture studies, osteoblasts are more responsive to fluid forces than to mechanical strain.^[22] Isolated osteocytes show strong responses to fluid flow of the order of 0.5 pascals (Pa), the magnitude predicted by the studies of Weinbaum and colleagues,^[18,23] responses include rapid release of nitric oxide (NO) (Fig. 4) and prostaglandin E₂ (PGE₂).^[13-15,24-26] Together, these findings suggest that the osteocyte network, with its accompanying lacuno-canalicular porosity, is the site of mechanosensing in bone tissue. Mechanotransduction then

includes translation of canalicular fluid shear stress into signals in osteocytes that can recruit osteoclasts and osteoblasts. The following sections discuss how this concept fits into a cellular mechanism of strain-adapted remodeling. (See the article titled *Histogenesis* in this encyclopedia.)

STRAIN-INDUCED CANALICULAR FLUID FLOW AROUND A BMU

Assuming that osteocytes regulate adaptive bone remodeling, the very local strain fields around a progressing BMU become important, because they could activate the local osteocytes to recruit first osteoclasts, and subsequently osteoblasts. Finite Element (FE) analysis of remodeling (hemi-)osteons, performed on a microscopic, supracellular level, found opposite strain fields around the tip and the tunnel of a progressing BMU.^[6] Equivalent strain was used as a measure for the deformation of the tissue. In the FE model of a tunneling osteon, a region of decreased strain appeared in front of the tunnel tip or cutting cone (Fig. 1) when the model was loaded naturally (i.e., in the longitudinal direction) (Fig. 5). Behind the tunnel tip, however, elevated strains appeared in the tunnel wall in the area known as the closing cone where new bone is formed by osteoblasts (Fig. 5). The bone tissue around the site where osteoclasts are active (the cutting cone) is thus understrained, while the tissue around the site where osteoblasts are active (the closing cone) is overstrained. A similar relation was found for (hemi-)osteons of

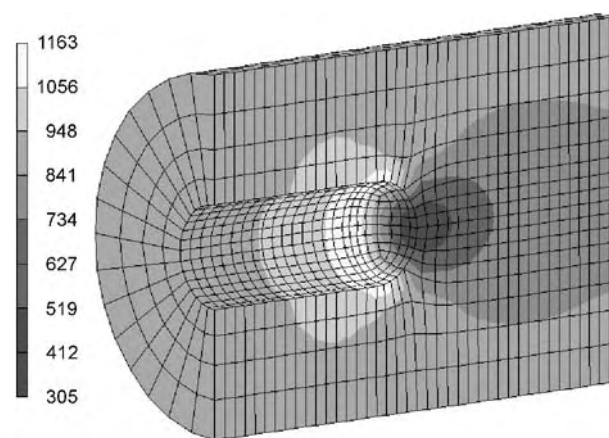


Fig. 5 Equivalent strain field in compact bone around the resorption tunnel of a tunneling BMU (half model). Loading occurred in the longitudinal direction. Reduced strains are present in front of the cutting cone tip where osteoclasts are active (blue or dark grey). Elevated strain levels are present directly after the cutting cone, where osteoblasts are recruited to refill the tunnel (yellow or white band). (Figure reproduced with permission from Ref. [6].)

remodeling trabecular bone (Fig. 6). Here the initial stage of remodeling was modeled as a resorption lacuna in a trabecula. Loading of the trabecula in the longitudinal direction produced increased strain at the bottom and transverse border of the lacuna, the sites where further osteoclast activity is inhibited while osteoblasts are actively filling up the gap. Reduced strain levels appeared, however, along the longitudinal border and surface of the trabecula just where resorption continues (Fig. 6). This study therefore concluded that subsequent activation of osteoclasts and osteoblasts during remodeling may be a strain-regulated phenomenon.^[6]

The opposite strain fields around cutting and closing cones were subsequently linked to fluid flow patterns in the lacuno-canalicular network around the cutting and closing cones.^[7,8] Biot's theory of poroelasticity was used to calculate the interaction between bone deformation and extracellular fluid flow in a finite element model of a tunneling osteon.^[8] Remodeling was considered to occur within the diaphyseal cortex of a human femur; the force pattern applied was a typical loading curve recorded from a walking person. As expected, the pattern and magnitude of flow were very different around the cutting and closing cones. At maximum load during the heel strike phase, a high fluid outflow was found at the bone surface of the closing cone that dampened out at a depth of some 100 micrometers within the bone tissue (Fig. 7). Along the cutting cone, however, a much lower fluid inflow at the bone surface reversed to a very low outflow at a

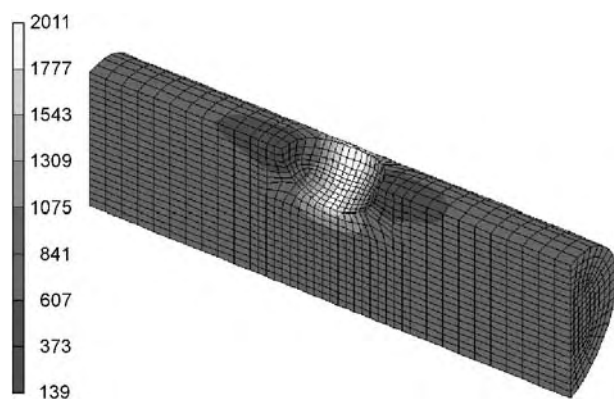


Fig. 6 Equivalent strain field around the initiation of a resorption trench along a remodeling trabecula (half model). Elevated strains are found at the bottom and transverse border of the resorption pit, where further osteoclast activity is inhibited and osteoblasts are recruited (yellow or white area). Reduced strain levels appear along the longitudinal rim and the trabecular surface (blue or dark grey), just where resorption will continue. (Figure reproduced with permission from Ref. [6].)

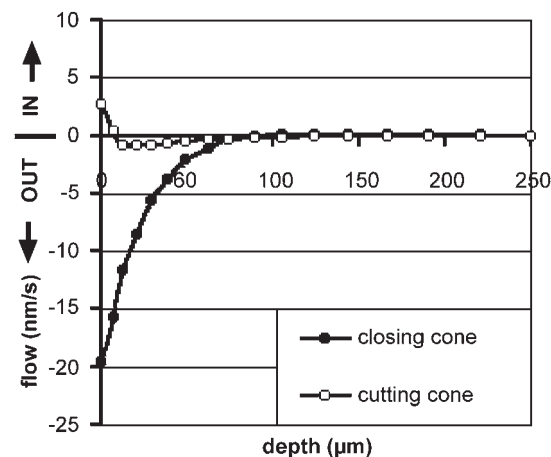


Fig. 7 Fluid flow within the lacuno-canalicular porosity around the cutting and closing cones of a remodeling resorption tunnel, in the human femoral diaphysis at maximum load during the walking cycle. Along the closing cone, fluid outflow (indicated by negative values) is high at the surface and dampens out at a depth of some 100 micrometers. Along the cutting cone, low fluid inflow (indicated by positive values) is reversed to low fluid outflow (negative values) at a depth of some 10 micrometers. (Figure reproduced with permission from Ref. [8].)

depth of some 10 micrometers within the bone tissue (Fig. 7). This led to a situation of almost complete fluid stasis in front of the cutting cone during the entire walking cycle (Fig. 8). The bone tissue layer immediately in front of the cutting cone is therefore an area that lacks fluid flow, whereas the bone tissue layer around the closing cone is an area of enhanced fluid flow.^[8] Assuming that unloaded osteocytes attract osteoclasts while overloaded osteocytes attract osteoblasts, this study is the first to link the progress of a BMU, and the coupling of bone formation to bone resorption, to the mechanical strain of osteocytes around a loaded remodeling site. Speculation is now in order about the cellular mechanisms by which unloaded osteocytes attract osteoclasts and overloaded osteocytes attract osteoblasts.

OSTEOCYTE APOPTOSIS AND REMODELING IN THE DIRECTION OF LOADING

The studies in the foregoing discussion^[6-8] provide evidence for mechanical regulation of osteoclast and osteoblast recruitment during bone remodeling. Osteoblast recruitment to loaded bone is generally believed to involve growth factors such as insulin-like growth factor^[11] as well as prostaglandins,^[27] but the pathway by which overloaded osteocytes recruit osteoblasts to the bone surface is largely unknown.^[28]

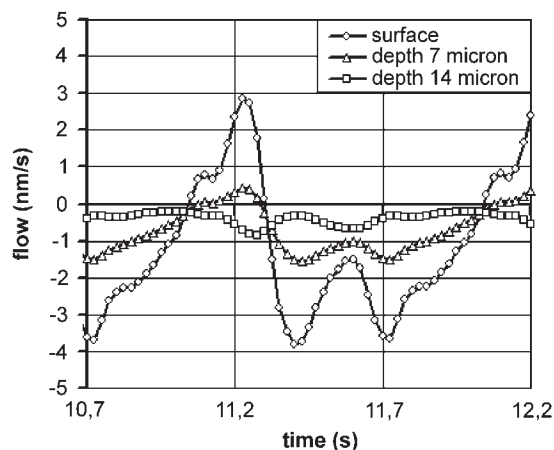


Fig. 8 Fluid flow at three levels of depth around the cutting cone tip of the remodeling resorption tunnel of Fig. 7. At a depth of about 10 micrometers, a situation of almost complete fluid stasis occurs during the whole walking cycle, leading to absence of fluid flow in the canaliculi during normal loading. (Figure reproduced with permission from Ref. [8].)

A cellular mechanism by which osteoclasts may be recruited by underloaded osteocytes has recently been proposed.^[9] This cellular mechanism is based on osteocyte apoptosis (or programmed cell death) as the steering mechanism for attracting osteoclasts.^[29] In the proposal by Burger, Klein-Nulend, and Smit, apoptosis of underloaded osteocytes results from lack of NO production due to lack of canalicular fluid flow. Production of low levels of endogenous NO is important for preventing apoptosis in many cell types, including the endothelial cells that line blood vessels.^[30] The example of endothelial cells is particularly important because these cells need a certain amount of fluid shear stress (provided by the blood stream) to produce enough NO to prevent the cells' apoptosis.^[31] The shear stress-dependent NO is produced by a specific enzyme, endothelial cell-NO-synthase (ecNOS), that is found in osteocytes as well as in endothelial cells.^[25,32] Because osteocytes rapidly produce low levels of NO by this enzyme in response to shear stress,^[24] it is therefore likely that NO plays a role in preventing their apoptosis similar to the role it plays with endothelial cells.^[9] Fluid shear stress has been reported to inhibit experimentally induced apoptosis in osteoblasts.^[33] The Burger/Klein-Nulend/Smit hypothesis proposes that, under normal mechanical loading conditions, osteocytes are protected against apoptosis by a basal amount of NO and probably by PGE₂ production under normal canalicular shear stress. At the tip of the cutting cone of a BMU, however, osteocytes enter apoptosis as a result of insufficient fluid flow and, therefore, insufficient NO (and PGE₂) production. These apoptotic

osteocytes attract osteoclasts (Fig. 9) that behave as macrophages by phagocytosing the dying osteocytes.^[29,34] By contrast, at the base of the cutting cone where the osteocytes become overloaded, NO production is higher than normal. Because released NO can cause retraction of osteoclasts from a resorbing bone surface,^[35] it is proposed that the osteoclasts of a BMU move in a treadmill: attaching at the tip of the cutting cone, detaching at the base, and reattaching again at the tip (Fig. 10).^[9] While osteoclasts further excavate the cutting cone, the front of low fluid flow and osteocyte apoptosis also moves along, steered by the strain fields in the correct (i.e., strain-oriented direction). This process will lead to secondary osteons and (hemi-)osteons that run in the direction of mechanical loading.

OSTEOCYTE APOPTOSIS, ACTIVATION OF REMODELING, AND UNWANTED BONE LOSS

Osteoclast recruitment by apoptotic osteocytes has been demonstrated experimentally in adult rat long bones after fatigue loading^[36] and after damage applied by inserting an orthopedic screw.^[37] Both studies linked local osteocyte apoptosis to rapid recruitment of osteoclasts and resorption of damaged bone surfaces.

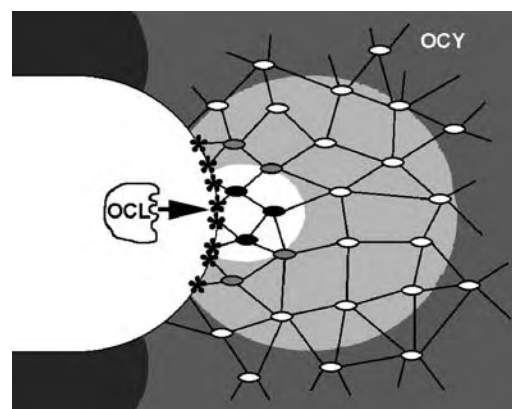


Fig. 9 Schematic illustration of the proposed relation between canalicular fluid flow stasis, osteocyte apoptosis, and osteoclast recruitment. White ovals are vital osteocytes (OCY); black ovals are apoptotic osteocytes. Black interconnecting lines represent canaliculi. OCL is an osteoclast in the cutting cone (white void) that is attracted to the cutting cone surface by apoptotic osteocytes. Longitudinal loading causes fluid stasis (indicated by white background) in front of the cutting cone, which in turn causes osteocyte apoptosis. Apoptotic osteocytes (black ovals) attract osteoclasts, possibly by expressing apoptosis-specific markers on the surface of their cell processes (indicated by asterisks). (Figure reproduced with permission from Ref. [9].)

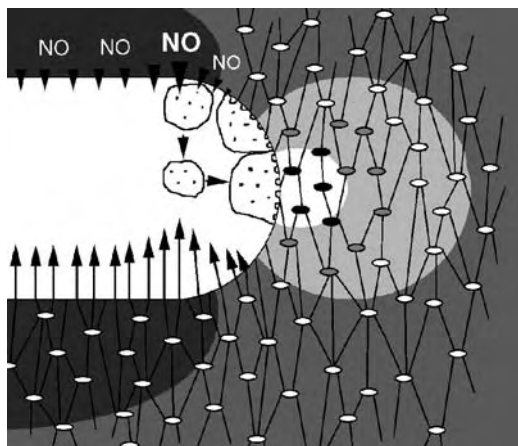


Fig. 10 Schematic illustration of how overloaded osteocytes around the base of the cutting cone may suppress further osteoclastic resorption by producing enhanced amounts of NO. White ovals are vital osteocytes, black ovals are apoptotic osteocytes, and interconnecting lines are canaliculi, similar to Fig. 9. Osteoclasts (large cells with several nuclei, indicated by black dots) are attracted by apoptotic osteocytes to attach to the cutting cone surface and resorb bone. Behind the cutting cone, where the wall of the tunnel is overloaded (indicated by dark gray background), the vital osteocytes are mechanically overstimulated (indicated by long arrows) to produce enhanced amounts of NO. High levels of NO cause retraction of the osteoclasts from the bone surface, so that osteoclasts now appear as round cells. A retracted osteoclast can only reattach at the cutting cone tip where no NO is produced, which leads to a treadmill of osteoclasts that continue to dig a tunnel along the direction of mechanical loading. (Figure reproduced with permission from Ref. [9].)

Osteocyte apoptosis thus served to start remodeling—a process called activation—rather than continuing it, as discussed in the previous section.^[9] However, many authors assume that in physiological bone remodeling, microdamage due to tissue fatigue is also a signal for activation of remodeling.^[38,39] Microdamage due to fatigue of the ECM can easily lead to osteocyte apoptosis by disrupting canaliculi and severing osteocyte processes.^[12] Osteoclast recruitment by apoptotic osteocytes will therefore also explain remodeling activation. A single mechanism—programmed cell death of local osteocytes—may thus cause first activation and subsequently adaptation during bone remodeling. The only difference is the cause of osteocyte apoptosis: ECM fatigue and damage for activation; strain-oriented disuse and lack of canalicular flow for (hemi-)osteon formation and alignment. Interestingly such a mechanism also explains the clinically observed unwanted loss of bone that occurs around orthopedic joint-replacing implants and in the edentulous jaw. Both cases may be seen as the outcome of the biological response

to bone disuse. In the edentulous jaw, the processus alveolaris of the jaw that carried the teeth completely loses its function after tooth extraction. In the case of the replaced hip joint, the femoral diaphysis is shielded against stress by the pin of the prosthesis. In both cases the resulting disuse of bone tissue may lead to lack of strain-derived fluid shear stress, and therefore to osteocyte apoptosis. These events will lead to untimely activation of remodeling, and unlimited progress of the osteoclastic attack. In addition, the space created by the resorptive action of osteoclasts is not refilled, because resorption does not lead to overstraining of the surrounding osteocytes, again due to generalized disuse of the surrounding bone tissue. Trabecularization of the femoral diaphysis and resorption of the processus of the jaw are the final result of this process.

CONCLUSION

Recent experimental and theoretical studies have for the first time provided an explanation for the strain-oriented remodeling of bone tissue. The histological structure of bone, with its three-dimensional network of osteocytes and lacuno-canalicular porosity, seems to provide the intrinsic mechanosensory system whereby bone tissue can renew itself in relation to the direction and magnitude of its daily mechanical loads. Although many details of this process remain to be unraveled, a cellular explanation of Wolff's Law now exists.

ARTICLES OF FURTHER INTEREST

Biologic and Synthetic Apatites; Bone Fracture Fixation; Bone-Implant Interface; Finite Element Analysis; Nitric Oxide; Osteogenic Progenitor Cells of Bone; Osteoinductive Substances and Materials; Tissue Engineering of Bone; Tissue Engineering of Bone Marrow

REFERENCES

1. Wolff, J. *Das Gesetz der Transformation der Knochen*; Springer Verlag: Berlin, Germany, 1892; 1986. Translated as: *The Law of Bone Remodelling*.
2. Frost, H.M. *Bone Remodeling Dynamics*; C.C. Thomas: Springfield, IL, USA, 1963.
3. Parfitt, A.M. Osteonal and hemi-osteonal remodeling: The spatial and temporal framework for signal traffic in adult human bone. *J. Cell. Biochem.* **1994**, *55*, 273–286.
4. Petrtyl, M.; Hert, J.; Fiala, P. Spatial organisation of the Haversian bone in man. *J. Biomech.* **1996**, *29*, 161–169.
5. Huiskes, R.; Ruimerman, R.; Van Lenthe, G.H.; Janssen, J.D. Effects of mechanical forces on maintenance

- and adaptation of form in trabecular bone. *Nature* **2000**, *405*, 704–706.
6. Smit, T.H.; Burger, E.H. Is BMU-coupling a strain-regulated phenomenon? A finite element analysis. *J. Bone Miner. Res.* **2000**, *15*, 301–307.
 7. Smit, T.H.; Huyghe, J.M.; Cowin, S.C. Estimation of the poro-elastic parameters of cortical bone. *J. Biomech.* **2002**, *35*, 829–835.
 8. Smit, Th.H.; Burger, E.H.; Huyghe, J.M. A case for strain-induced fluid flow as regulator of BMU-coupling and osteonal alignment. *J. Bone Miner. Res.* **2002**, *17*, 2021–2029.
 9. Burger, E.H.; Klein-Nulend, J.; Smit, T.H. Strain-derived canalicular fluid flow regulates osteoclast activity in a remodelling osteon—A proposal. *J. Biomech.* **2003**, *36*, 1453–1459.
 10. Skerry, T.M.; Bitensky, L.; Chayen, J.; Lanyon, L.E. Early strain-related changes in enzyme activity in osteocytes following bone loading in vivo. *J. Bone Miner. Res.* **1989**, *4*, 783–788.
 11. Lean, J.M.; Jagger, C.J.; Chambers, T.J.; Chow, J.W. Increased insulin-like growth factor I mRNA expression in rat osteocytes in response to mechanical stimulation. *Am. J. Physiol.* **1995**, *268*, E318–E327.
 12. Burger, E.H.; Klein-Nulend, J. Mechanotransduction in bone—Role of the lacunocanalicular network. *FASEB J.* **1999**, *13*, S101–S112.
 13. Ajubi, N.E.; Klein-Nulend, J.; Nijweide, P.J.; Vrijheid-Lammers, T.; Alblas, M.J.; Burger, E.H. Pulsating fluid flow increases prostaglandin production by cultured chicken osteocytes—A cytoskeleton-dependent process. *Biochem. Biophys. Res. Commun.* **1996**, *225*, 62–68.
 14. Klein-Nulend, J.; Van der Plas, A.; Semeins, C.M.; Ajubi, N.E.; Frangos, J.A.; Nijweide, P.J.; Burger, E.H. Sensitivity of osteocytes to biomechanical stress in vitro. *FASEB J.* **1995**, *9*, 441–445.
 15. Pitsillides, A.A.; Rawlinson, S.C.F.; Suswillo, R.F.L.; Bourrin, S.; Zaman, G.; Lanyon, L.E. Mechanical strain-induced NO production by bone cells—A possible role in adaptive bone (re)modeling. *FASEB J.* **1995**, *9*, 1614–1622.
 16. Burr, D.B.; Milgrom, C.; Fyhrie, D.; Forwood, M.; Nyska, M.; Finestone, A.; Hoshaw, S.; Saiag, E.; Simkin, A. In vivo measurement of human tibial strains during vigorous activity. *Bone* **1996**, *18*, 405–410.
 17. Burger, E.H.; Veldhuijzen, J.P. Influence of Mechanical Factors on Bone Formation, Resorption, and Growth in vitro. In *Bone*; Hall, B.K., Ed.; CRC Press: Boca Raton, FL, USA, 1993; Vol. 7, 37–56.
 18. Weinbaum, S.; Cowin, S.C.; Zeng, Y. A model for the excitation of osteocytes by mechanical loading-induced bone fluid shear stresses. *J. Biomech.* **1994**, *27*, 339–360.
 19. Cowin, S.C.; Weinbaum, S.; Zeng, Y. A case for bone canaliculi as the anatomical site of strain generated potentials. *J. Biomech.* **1995**, *28*, 1281–1296.
 20. Turner, C.H.; Forwood, M.R.; Otter, M.W. Mechanotransduction in bone: Do bone cells act as sensors of fluid flow? *FASEB J.* **1994**, *8*, 875–878.
 21. Turner, C.H.; Owan, I.; Takano, Y. Mechanotransduction in bone: Role of strain rate. *Am. J. Physiol.* **1995**, *269*, E438–E442.
 22. Owan, I.; Burr, D.B.; Turner, C.H.; Qui, J.; Tu, Y.; Onyia, J.E.; Duncan, R.L. Mechanotransduction in bone: Osteoblasts are more responsive to fluid forces than mechanical strain. *Am. J. Physiol.* **1997**, *273*, C810–C815.
 23. Weinbaum, S.; Guo, P.; You, L. A new view of mechanotransduction and strain amplification in cells with microvilli and cell processes. *Biorheology* **2001**, *38*, 119–142.
 24. Klein-Nulend, J.; Semeins, C.M.; Ajubi, N.E.; Nijweide, P.J.; Burger, E.H. Pulsating fluid flow increases nitric oxide (NO) synthesis by osteocytes but not periosteal fibroblasts—Correlation with prostaglandin upregulation. *Biochem. Biophys. Res. Commun.* **1995**, *217*, 640–648.
 25. Klein-Nulend, J.; Helfrich, M.H.; Sterck, J.G.H.; MacPherson, H.; Joldersma, M.; Ralston, S.H.; Semeins, C.M.; Burger, E.H. Nitric oxide response to shear stress by human bone cell cultures is endothelial nitric oxide synthase dependent. *Biochem. Biophys. Res. Commun.* **1998**, *250*, 108–114.
 26. McAllister, T.N.; Frangos, J.A. Steady and transient fluid shear stress stimulate NO release in osteoblasts through distinct biochemical pathways. *J. Bone Miner. Res.* **1999**, *14*, 930–936.
 27. Westbroek, I.; Ajubi, N.E.; Alblas, M.J.; Semeins, C.M.; Klein-Nulend, J.; Burger, E.H.; Nijweide, P.J. Differential stimulation of prostaglandin G/H synthase-2 in osteocytes and other osteogenic cells by pulsating fluid flow. *Biochem. Biophys. Res. Commun.* **2000**, *268*, 414–419.
 28. Imai, S.; Kaksonen, M.; Raulo, E.; Kinnunen, T.; Fages, C.; Meng, X.; Lakso, M.; Rauvala, H. Osteoblast recruitment and bone formation enhanced by cell-matrix associated heparin-binding growth-associated molecule (HB-GAM). *J. Cell Biol.* **1998**, *143*, 1113–1128.
 29. Bronckers, A.L.J.J.; Goei, S.W.; Luo, G.; Karsenty, G.; D'Souza, R.N.; Lyaruu, D.M.; Burger, E.H. DNA fragmentation during bone formation in neonatal rodents assessed by transferase-mediated end labeling. *J. Bone Miner. Res.* **1996**, *11*, 1281–1291.
 30. Dimmeler, S.; Zeiher, A.M. Nitric oxide and apoptosis: Another paradigm for the double-edged role of nitric oxide. *Nitric Oxide* **1997**, *1*, 275–281.
 31. Rossig, L.; Haendeler, J.; Hermann, C.; Malchow, P.; Urbich, C.; Zeiher, A.M.; Dimmeler, S. Nitric oxide down-regulates MKP-3 mRNA levels: Involvement in endothelial cell protection from apoptosis. *J. Biol. Chem.* **2000**, *275*, 2552–2557.
 32. Zaman, G.; Pitsillides, A.A.; Rawlinson, S.C.; Suswillo, R.F.; Mosley, J.R.; Cheng, M.Z. Mechanical strain stimulates nitric oxide production by rapid activation of endothelial nitric oxide synthase in osteocytes. *J. Bone Miner. Res.* **1999**, *14*, 1123–1131.
 33. Pavalko, F.M.; Gerard, R.L.; Ponik, S.M.; Gallagher, P.J.; Jin, Y.; Norvell, S.M. Fluid shear stress inhibits TNF-alpha-induced apoptosis in osteoblasts: A role for fluid shear stress-induced activation of caspase-3. *J. Cell. Physiol.* **2003**, *194*, 194–205.
 34. Elmardi, A.S.; Katchburian, M.V.; Katchburian, E. Electron microscopy of developing calvaria reveals

- images that suggest that osteoclasts engulf and destroy osteocytes during bone resorption. *Calcif. Tissue Int.* **1990**, *46*, 239–245.
35. MacIntyre, I.; Zaidi, M.; Alam, A.S.M.T.; Datta, H.K.; Moonga, B.S.; Lidbury, P.S.; Hecker, M.; Vane, J.R. Osteoclast inhibition: An action of nitric oxide not mediated by cyclic GMP. *Proc. Natl. Acad. Sci. U. S. A.* **1991**, *88*, 2936–2940.
36. Bentolilla, V.; Boyce, T.M.; Fyhrrie, D.B.; Drumb, R.; Skerry, T.M.; Schaffler, M.B. Intracortical remodeling in adult rat long bones after fatigue loading. *Bone* **1998**, *23*, 275–281.
37. Verborgt, O.; Gibson, G.J.; Schaffler, M.B. Loss of osteocyte integrity in association with microdamage and bone remodeling after fatigue in vivo. *J. Bone Miner. Res.* **2000**, *15*, 60–67.
38. Burr, D.B.; Martin, R.B. Calculating the probability that microcracks initiate resorption spaces. *J. Biomech.* **1993**, *26*, 613–616.
39. Burr, D.B.; Forwood, M.R.; Fyhrrie, D.P.; Martin, R.B.; Schaffler, M.B.; Turner, C.H. Bone microdamage and skeletal fragility in osteoporotic and stress fractures. *J. Bone Miner. Res.* **1997**, *12*, 6–15.

Burn Dressing

Andrew Burd

Department of Surgery, The Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, Hong Kong

INTRODUCTION

The concept of burn dressings is undergoing radical change. Traditionally, the goal was to protect the wound and promote natural healing. Now we are revisiting the concept of dermal healing, a process of repair resulting in unwanted scars. The focus has shifted toward preventing repair and promoting regeneration. A burn is a specific type of wound caused by a pathological change in energy within a tissue, resulting in the loss of functional and structural integrity. All of this energy—thermal (hot), chemical, electrical, or radiation—destroys cells and denatures proteins. In thermal (cold), proteins will not be denatured and a different outcome will result. Healing is the process whereby wounded tissues respond and attempt to restore functional and structural integrity. In the skin there are two major mechanisms of healing; regeneration and repair.

This article begins with a brief overview of the structure and function of skin and the effects of burn injury. Burn injury represents a wide spectrum of trauma, and treatment strategies are varied. Burn dressings provide an essential part of treatment. They may be all that is required in a superficial wound, or they may be used as adjuncts to surgical treatment. In this latter context there has been a surge in the development of tissue-engineered products incorporating biological interactions.

STRUCTURE AND FUNCTION OF SKIN

Skin is the largest immunologically competent organ in the body.^[1,2] It extends to over 1.6 square meters in the adult and weighs approximately 3000 gm. The skin has two layers: epidermis and dermis (Fig. 1). The epidermis is rich in cells and is of ectodermal origin. The cells are specialized in the formation of keratin and are called keratinocytes. The basal keratinocytes adhere to the basement membrane, which forms part of the zone between epidermis and dermis: the dermo-epidermal junction. Basal keratinocytes are unique in that they are capable of proliferating to form either new basal cells or terminally differentiating keratinocytes. The terminally differentiating keratinocyte moves up

through several layers and undergoes certain changes: the cytoplasm becomes increasingly packed with keratin and the nucleus shrinks. The outermost layer—the stratum corneum—contains dead keratinocytes packed with keratin, forming a barrier between the living tissues and the external environment.

The area of the basement membrane far exceeds that of the surface of the stratum corneum, because multiple dermal papillae project from the surface of the dermis. These papillae are formed by loose approximations of collagen bundles referred to as the papillary dermis, which lies on the relatively much thicker reticular dermis. Collagen bundles in the reticular dermis are thicker and more condensed.

Between the organized structure of elastin and collagen fibers is a thick, viscous fluid made up of glycosaminoglycans and hyaluronan. This arrangement of the dermis gives the skin its major biomechanical properties, allowing stretching and recoil and deformation without destruction.

The dermis contains a complex vascular arrangement of capillary and venous plexuses. Within the dermis are adnexal structures of ectodermal origin—hair follicles, sweat glands, and sebaceous glands—which are lined with keratinocytes. This arrangement becomes particularly important when considering the mechanisms of burn wound healing.

Multifunctional Langerhan's cells found in the epidermis have an important role as antigen-presenting cells in cutaneous immune reactions. Skin color is determined principally by the pigment melanin, produced by melanocytes. These cells package the melanin into melanosomes, which are then transferred to keratinocytes.

The skin is richly endowed with sensory nerve endings that enable the skin to play its vital role as a tactile interface between the body and the environment. Other sensory functions provide important stimuli for behavioral modification, e.g., the withdrawal reflex associated with pain or excessive temperature.

THE BURN INJURY

The damage caused by a burn is a function of the amount and the speed of the energy change within the

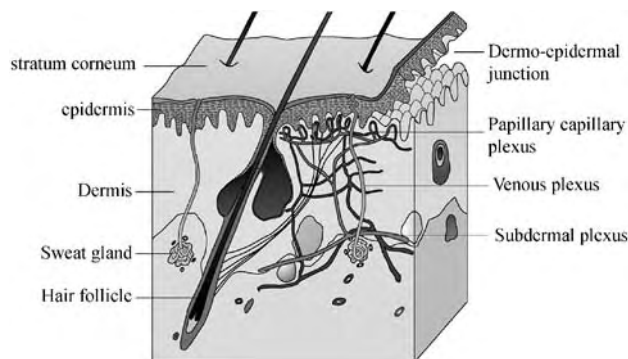


Fig. 1 A diagrammatic section of skin showing bilaminar structure, vascularity, and appendageal elements.

system.^[3] In a burn from hot water, for example, two major factors that influence the potential damage are the temperature of the water and the duration of contact with the skin. Prompt cooling can reduce this damage. The actual damage will be influenced by other factors, however, including the anatomical location of the burn, the local blood supply, and the treatment.

Following the burn the immediate damage will be to the keratinocytes. The integrity of the cell membrane is lost and the cell contents are released. This results in an uncontrolled series of devastating chemical interactions that produce a range of biologically active substances with wide-reaching systemic effects. Capillary permeability increases, resulting in a fluid shift from the vascular to the extravascular compartment. Edema develops in the tissues and the patient suffers from increasing hypovolemia that can lead to renal failure and death, if not treated. For the uncomplicated burn injury, the immediate concern is to treat the potential fluid imbalance by a prospective infusion of intravenous fluids. Once the acute resuscitative needs of the patient have been addressed, attention is then focused on the burn wound.

A number of factors influence the management of the burn wound:

- the percentage of body surface area (BSA) involved, affecting logistics
- the depth of the burn, affecting surgical strategy
- the resources available
- the ability to utilize the resources
- the preburn condition of the patient

The depth of injury is classified by descriptive terminology. The skin may be either partially or completely destroyed; partial destruction may be superficial, intermediate, or deep. Most burn injuries will contain areas of varying depth of tissue damage, and the depth of burn determines the capacity of the wound to heal.

BURN WOUND HEALING, REGENERATION, AND REPAIR

Regeneration is the capacity of a tissue to renew itself so that the end result is indistinguishable from the preinjured tissue.^[4] Regeneration is a feature seen in superficial partial-thickness burns, in which the injury involves the loss of epidermis and basement membrane and the papillary dermis. There may be a highly exudative and painful wound.

The exudative phase persists for several days, and as it decreases the nature of the exudates changes. The viscosity and relative protein content increase, and eventually a fibrin layer seals the wound. In the meantime, the basal keratinocytes at the margin of the wound begin to undergo mitosis. In the normal resting state only approximately 12% of basal keratinocytes are proliferating at any time, giving the skin a tremendous reserve capacity. Re-epithelialization begins not just at the wound margin, but also from the appendageal structures. The rate of keratinocyte proliferation and migration is extremely high, and when the exposed dermal surface is recovered, contact inhibition abruptly stops migration and redirects the cells to stratification. A superficial partial-thickness burn will heal with a stratified squamous epithelial in a matter of days. Disturbance in normal pigment expression can occur, even with no scarring. In people with darker pigmentation, areas of absent pigment can have major social and psychological sequelae.

As the burn becomes deeper, the nature of the wound changes. With an intermediate partial-thickness burn the epidermis is destroyed (as with the superficial partial-thickness burn) and there is more dermal damage. The damage to the dermis involves an irreversible denaturation of the collagen. The inert collagen has to be removed for re-epithelialization to take place. Removal involves an autolytic process with enzymatic degradation and phagocytosis, augmented by an inflammatory response. Healing takes longer. As the depth of burn increases through the thickness of the dermis the phenomenon of inflammation plays an even more important role in the healing process. Typically, inflammation initiates a cascade of events, with polymorphonuclear leucocytes being attracted to the wound site. Their principal role is proteolysis and phagocytosis of debris. These leucocytes signal to macrophages which, when activated, enter the wound site to undertake a more detailed assessment of the damage. Through a cytokine-mediated signaling process, they recruit fibroblasts to begin the process of replacing the damaged collagen. Fibroblasts involved in wound healing have the capacity to produce abundant amounts of collagen, but they have lost the capacity to place and organize it in a highly structured way. The end result of dermal repair is the deposition

of disorganized collagen, which is physically apparent as scar tissue. Scarring represents a very complex biological phenomenon. Hypertrophic and keloid scars represent clinical descriptions of a heterogeneous group of disorders with various etiologies and pathological mechanisms that result in the production of a disorganized connective tissue. The duration of the phase of wound closure (i.e., re-epithelialization) and the incidence of adverse scarring increase as the depth of burn increases.

Scarring is the major cause of long-term morbidity after a burn. It can result in physical disability when function is impaired and psychological and social isolation when deformity distorts features, particularly on the face and other regions of high aesthetic impact.

Burn wounds and their treatment vary widely, and there are few universal standards. The most widely accepted surgical principles are that superficial partial-thickness wounds *do not* require surgery and that full thickness wounds *do* require surgery. There is the qualification that a full-thickness wound of no more than 2–3 cm in diameter in a nonvital part of the body can close by the natural process of wound contraction. Debate revolves around the role of surgery for the intermediate and partial-thickness burns.

The goals of burn treatment have evolved, and the major outcomes are no longer described in simple actuarial terms of death, survival, or time-to-healing but in more qualitative terms that relate to the degree, effects, and influence of scarring.

THE IDEAL BURN DRESSING

Properties of the ideal dressing for burns are listed in Table 1. In addition, the dressing should be easy to apply and readily available.

Protects

The skin provides a number of protective functions. The stratum corneum acts as a barrier that prevents invasive infection, and sebum secreted by sebaceous glands has an antibacterial action. The burn wound is vulnerable to bacterial colonization and invasive

bacterial contamination and sepsis. As burn depth increases the potential reservoir for bacterial invasion increases before viable tissue and host defenses are reached. Burns dressing should have either an antibacterial barrier function or a bactericidal/bacteriostatic function or, ideally, both. The dressing also should protect the regenerating epidermis from further mechanical injury.

Proteolytic

Except for the most superficial injuries, the wound is going to produce a layer of dead tissue. The natural process of removing this tissue involves proteolysis and phagocytosis. The deeper the wound the more tissue has to be removed, and the longer it will take for healing to occur. The ideal dressing would either support or augment the natural autolytic process or, even better, have an intrinsic enzymatic action.

Promotes Healing

The ideal dressing would promote healing through regeneration. Regeneration occurs when the epidermis is damaged, but repair is the mechanism for dermal wound healing. Repair is a macrophage-driven fibroproliferative process in which collagen is rapidly but randomly deposited. The fibroblasts involved are recruited by the macrophages, but their origin remains a subject of controversy. They are mesodermal cells, and one possibility is that they are derived from a differentiated endothelial cell at the front end of sprouting capillary buds. Whatever their origin, they seem to have lost the blueprint for the original dermal architecture.

Apart from the selective promotion of specific cell behavior, the ideal wound dressing would also provide an environment that would be more conducive to healing. In this respect the healing wound performs better in a moist healing environment, but not a macerated healing environment. Accordingly, the dressing should have the physical ability to deal with the exudates produced by the wound, but at the same time not allow the wound to dry out and the tissues to desiccate.

Pain Relieving

Partial-thickness burns can be extremely painful, and one aspect of pain control is the nature of the dressing applied to the wound. In addition, when the dressing needs to be changed it should be possible to remove it easily from the wound prior to reapplication.

Table 1 Ideal burns dressing

Properties
<ul style="list-style-type: none"> • Protects • Proteolytic • Promotes healing • Pain relieving

BURNS DRESSINGS—DEFINITIONS AND CLASSIFICATION

On a global scale a vast range of nonmedical materials are applied to burn wounds (Table 2).^[5]

Banana leaves could be regarded as a dressing, as well indeed could river clay, but what about curry powder or, for that matter, urine? It would appear that a distinction must lie somewhere between treatments that could be regarded as dressings and treatments that are not dressings.

The Chambers Dictionary^[6] refers to a dressing as being “material applied to a wound.” Material, in turn, is described as relating to matter, and matter is defined as “that which occupies space and with which we become acquainted by our bodily senses”! On the face of it, then, chili powder is a dressing, as indeed is urine. Using such simplistic logic, even air can be described as a dressing. Indeed, facial burns are routinely left exposed, and more extensive burns also can be

Table 2 List of nonmedical treatments used for burns in the Hermans Survey

Butter
Eggs and egg white
Aloe vera
Honey
Milk and condensed milk
Banana leaves
Bicarbonate powder
Coffee powder
Cowdung
Flour
Herb mixtures
Lettuce leaves
River clay
Toothpaste
Yoghurt
Oil
Curry powder
Gentian violet
Ink
Metal-containing powder
Nivea cream
Saffron
Talcum powder
Turmeric
Urine
Zambuk

(From Ref. [5].)

treated by exposure to filtered, warm laminar-flow air in individualized cubicles. The wound exudate dries within a few days to form a thin eschar, which acts as a protective dressing for the underlying process of re-epithelialization (Fig. 2).

How then to classify burns dressings? We could consider the type of material—solid, liquid, or cream—or the complexity of the dressing with single or multiple components. There is a tendency toward developing more complex dressings that serve a number of functions, e.g., antibacterial and exudate control and/or wound protection. Another classification is to consider the range of available treatment options and develop groupings of similar types of dressing. Table 3 lists a range of treatments for partial-thickness burns, and Table 4 lists the treatments for full-thickness burns. We thus arrive at the classification outlined in Table 5. This classification differs from that proposed by Quinn et al. in their seminal 1985 paper,^[7] where she proposed that burn dressings fall into three categories: conventional, biological, and synthetic. The conventional dressing was cotton wool, gauze and lint, or Gamgee (gauze and cotton tissues), which are still widely used for dressing many types of wound including burns, principally because of their absorptive capacity (Fig. 3). The biological dressings in Quinn’s classification comprise autograft, allograft, xenograft, and tissue derivatives. In turn, the allograft can be skin or amnion, and the tissue can be used fresh, frozen, or lyophilized. The third category in Quinn’s classification—synthetic—is divided into films, gel, foams, and composites. There are a number of limitations in Quinn’s classification, notwithstanding the fact that it was published at a time when there was an exponential growth in the concept of tissue engineering. That growth resulted in a new category of biological skin substitutes that also have synthetic elements and are known as biological composites.

Skin Substitutes: Biological and Synthetic

There have been a number of more recent reviews of skin substitutes, both biological and synthetic. These act as a temporary dressing and may incorporate elements that modulate the healing response (Table 6). Pruitt gives an excellent review of the evolutionary development of biologic dressings and skin substitutes, ranging from rubbing frogs dipped in warm oil onto a burn wound (as described in the Ebers Papyrus dated 1500 B.C.) to the use of chimeric allogenic tissue proposed over three thousand years later.^[8] An increasing number of new products on the market could be grouped together under the heading of biological skin substitutes,^[9,10] and the majority of these (except for cadaveric allograft) are engineered (Table 7). The

B



Fig. 2 A burn patient treated by exposure. Note that the hands are being treated in a closed environment, covered with polythene bags containing silver sulphadiazine cream. Exudate collects in the bags and they are changed daily. The hands are free to move and the physiotherapist will assist in active exercises to maintain joint function and decrease edema.

largest not-for-profit skin bank is the Euroskin bank, based in Holland, which provides glycerolized cadaver skin.^[11]

Skin substitutes vary widely in concept and application, but Integra sets the standard in addressing the problem of dermal repair. Integra is a bilaminar material; the outer layer—silicone—is synthetic (Fig. 4). The inner layer is derived from bovine type I collagen and shark-derived chondroitin sulphate, combined in a tissue-engineering process to produce a biodegradable matrix. The design of this matrix is such that the process of degradation results in the formation of a highly organized autocollagenous replacement. With

Table 3 List of treatments for partial-thickness burns in the Hermans Survey

Silver sulfadiazine cream 1% (different brands)
Cerium silver sulfadiazine cream (Flammacerium)
Betadine cream
Sulfamylon (different concentrations possible)
Silver nitrate solutions (different concentrations possible)
Nonbiological occlusive dressings (such as hydrogels, hydrocolloids, polyurethane films)
Xenograft (usually pig skin)
Human cadaver skin, fresh
Human cadaver skin, frozen
Human cadaver skin, lyophilized
Human cadaver skin, glycerolized
Amnion membrane
Tanning techniques
Tulle, vaseline gauze, medicated vaseline gauze
Tangential excision
Other materials and techniques

(From Ref. [5].)

concurrent vascularization the new dermis can be covered with autologous keratinocytes after the silicone outer layer is removed. The concept and practice come close to being realized in terms of permanent skin replacement. The process is influenced by many factors, however. Whereas the best results compare extremely favorably with autologous full and thick-split thickness grafts, there is still a considerable range in outcomes that suggests more work needs to be done in refining the product and defining the clinical procedures and applications. Another tissue-engineered

Table 4 List of treatments for full-thickness burns in the Hermans Survey

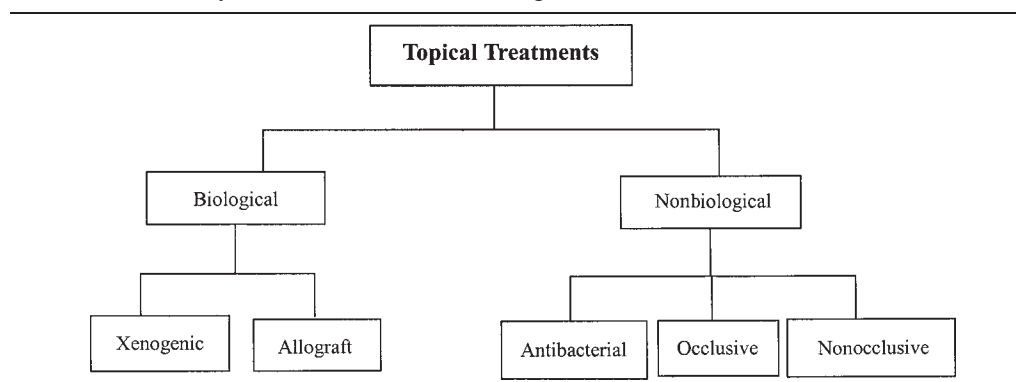
Silver sulfadiazine cream 1% (different brands)
Cerium silver sulfadiazine cream (Flammacerium)
Betadine
Sulfamylon or silver nitrate solutions (different concentrations possible)
Tanning techniques
Tangential excision and grafting
Avulsion (excision to fat or fascia) and grafting
Other

If grafting, preferred technique:

Meshed autografts
Meshed autografts
Sandwich technique
Chinese intermingled techniques
Cultured keratinocytes (nonspecific on preparation of wound bed, with or without dermal component such as allografts with shaved off epidermis, or Integra)

(From Ref. [5].)

Table 5 Preliminary classification of burn dressings



B

product, Apligraf, adopts an alternative approach. Apligraf has been described as “the nearest thing to living skin” that is produced in a laboratory. It has a bilaminar structure with a collagen matrix populated with allogenic neonatal fibroblasts, over which there is a stratified layer of neonatal keratinocytes. This living dressing is applied to the wound and begins a complex interaction. The cytokine profile of the wound bed is detected and modulated by cytokine expression by the neonatal cells. The nature of the wound changes, and the process and outcome of the healing are modified.

In a review of the clinical application of tissue-engineered skin substitutes, Kearney provides a useful categorization:^[12]

	<i>Synthetic</i>	<i>Biological</i>
Cell-free	e.g., Integra	e.g., Alloderm
Cell-containing (or cell-conditioned)	e.g., Dermagraft	e.g., Apligraf

Kearney also describes the therapeutic strategies of the various products:

1. Materials that are applied for short intervals and then removed or replaced at intervals. Function: to stimulate autologous healing.
2. Cell-free materials that encourage colonization by autologous cells and/or biodegradation and remodeling. Function: to stimulate neo-skin function.
3. Cell-containing skin substitutes. Function: to provide an immediate functional skin replacement.

A major problem with these new concepts in biointeractive, tissue-engineered burn wound dressings is cost. Other factors that affect the availability of biologically derived dressings are regulatory approval and legislation. One of the major reasons for the considerable interest in tissue-engineered skin substitutes is the lack of availability of what many consider to be the gold standard for temporary biological burn wound dressing: viable cutaneous allograft.^[13] The beneficial properties of viable cutaneous allograft are



Fig. 3 This patient has been treated with the closed method. The extensive burns are covered with antibacterial dressings adjacent to the skin and with absorbent Gamgee and retaining dressings. The dressings require daily changing, and the process is extremely time consuming.

Table 6 Pruitt's list of biological dressings

- Naturally occurring tissues
 - Cutaneous allografts
 - Cutaneous xenografts
 - Amniotic membranes
- Skin substitutes
 - Synthetic bilaminate
 - Collagen-based composites
 - Biobrane
 - Integra
 - Dermagraft-TC
 - Collagen-based dermal analogs
 - De-epithelialized allograft
 - Alloderm
- Culture-derived tissue
 - Cultured autologous keratinocytes
 - Fibroblast-seeded dermal analogs

(From Ref. [8].)

listed in Table 8, but there are several problems associated with its use that are independent of its biological effects. The principal concerns are donor availability, effective processing, and safety in regard to disease transmission.^[14,15] Allograft skin can be either viable or nonviable. Viable skin, which can be fresh or frozen in liquid nitrogen, has advantages in providing cytokine release into the wound bed and also providing temporary vascular connection. With appropriate immunosuppression, viable allograft can be considered a transplant, but such clinical application is rarely indicated or justified. Nonviable cadaver skin has certain advantages in terms of increased safety, and it also has a longer shelf life. The most common ways to render skin nonviable are either gamma irradiation, glycerolization, or lyophilization.

There are three principal uses for cadaver skin:

1. Both viable and nonviable cadaver skin are used as temporary skin substitutes after the excision of full-thickness cutaneous burns. Rejection will occur in both, but the period of take can be extended by the use of immunosuppression.

Table 7 Tissue engineered skin substitutes

Trade name	Source	Layers	Purpose
Alloderm	Lifecell corporation	1) Acellular de-epithelialized cadaver dermis	Permanent dermal replacement.
Apligraf	Organogenesis corporation	1) Neonatal keratinocytes 2) Collagen seeded with neonatal fibroblasts	An 'intelligent' bio-interactive wound healing modulator.
Biobrane	Dow Hickman/Bertek Pharmaceuticals	1) Silicone 2) Nylon mesh 3) Collagen	A temporary dressing for partial-thickness burns and a temporary skin substitute for excised burn wounds.
Comp Cult Skin	Ortec International, Inc.	1) Allogenic keratinocytes 2) Allogenic fibroblasts and collagen	A temporary bio-interactive wound dressing.
Dermagraft	Advanced Tissue Sciences	1) Polyglycolic acid (Dexon) or polyglactin-910 (Vicryl) seeded with neonatal fibroblasts.	When covered with autologous keratinocytes, this can be a permanent skin substitute.
Epicel	Genzyme Biosurgery	Cultured autologous keratinocytes	Permanent epidermal replacement.
Integra	Integra Life Sciences	1) Silicone 2) Collagen and glycosaminoglycans	Temporary skin substitute for excised wounds. Permanent dermal regeneration template.
Laserskin	Fidia Biopolymers (Italy)	1) Culture autologous keratinocytes 2) Hyaluronic acid membrane with laser drilled perforations.	Keratinocyte delivery system for permanent epidermal replacement.
Polyactive	HC Implants (Netherlands)	1) Polyethylene oxide and Polybutylene terephelate with autogenous human fibroblasts.	Covered with autologous human keratinocytes to give a permanent skin replacement.
Transcyte	Advanced Tissue Sciences	1) Silicone 2) Nylon mesh 3) Collagen seeded with neonatal fibroblasts.	Temporary, bio-interactive, skin substitute.



Fig. 4 Integra dermal regeneration template is being stapled to the excised burn wound of this young patient.

B

2. A more common use of nonviable allogenic skin is to act as a temporary cover to widely meshed autograft, which is placed on a bed of viable tissue after the excision of a full-thickness or very deep partial-thickness burn. The autograft will take and the interstices of the mesh will close, but this process takes time. The allogenic skin prevents desiccation from occurring in the viable exposed graft bed in the interstices of the meshed autograft. Once re-epithelialization has occurred, the allograft (the nonviable allogenic skin) falls off. It has acted as a temporary dressing.
3. The third use for nonviable allograft is as a dressing for intermediate-depth partial-thickness burns. The burn is not excised, but it is cleaned and dressed with the allogenic skin. The cadaver skin adheres to the wound and protects it while healing occurs underneath. When the underlying skin heals the allogenic skin desiccates and falls off.

Allogenic skin can also run into problems with regulatory authorities, and the classification of allogenic skin as a transplant or a dressing does have medical/legal implications. Figure 5 shows a scheme for determining the nature of allogenic skin.^[16]

The other allogenic material used as a burn wound dressing is amniotic membrane. Where alternatives do not exist and risks of use are otherwise minimized and justified, it can provide an effective, temporary, biologically interactive dressing.

The other principal category of biological dressing includes xenogenic dressings. Pruitt describes the evolution of the use of xenograft tissue, identifying Canaday as being among the pioneers when he reported the use of water lizard skin in 1692.^[8] It is interesting to note that Dr. Nelson Piccolo in Brazil regularly uses frog skin dressings to good effect on burn wounds today. The skins of guinea pigs, rabbits, chickens, and dogs have all been tried, but it is porcine xenograft that finds the widest clinical application in contemporary burn care. Like human allograft, porcine xenograft can be used fresh, frozen, irradiated, or lyophilized. Availability depends on local resources, although a lyophilized preparation is commercially available. In most cases, pig skin is used as a temporary dressing for superficial burns. Typically applied after the exudative phase has passed, it provides a protective barrier dressing that relieves pain, allows movement, and promotes healing. Porcine skin can also be used as a biological dressing for a debrided partial-thickness

Table 8 Properties of viable cutaneous allograft

- | |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <ol style="list-style-type: none"> 1) The intact stratum corneum provides a barrier function which <ol style="list-style-type: none"> a. limits desiccation of the wound. b. limits evaporative water loss. c. reduces bacterial contamination. d. protects underlying viable tissue. e. limits exudative tissue fluid loss. 2) The biomechanical properties of the dermal component allows <ol style="list-style-type: none"> a. effective draping of the wound. b. movement of joints. 3) The biochemical properties of the skin cause <ol style="list-style-type: none"> a. In the partial-thickness burn <ol style="list-style-type: none"> i reduced pain ii enhanced healing iii decreased scar b. In full-thickness excised wounds <ol style="list-style-type: none"> i angiogenesis on the wound bed, which aids subsequent autograft take. |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

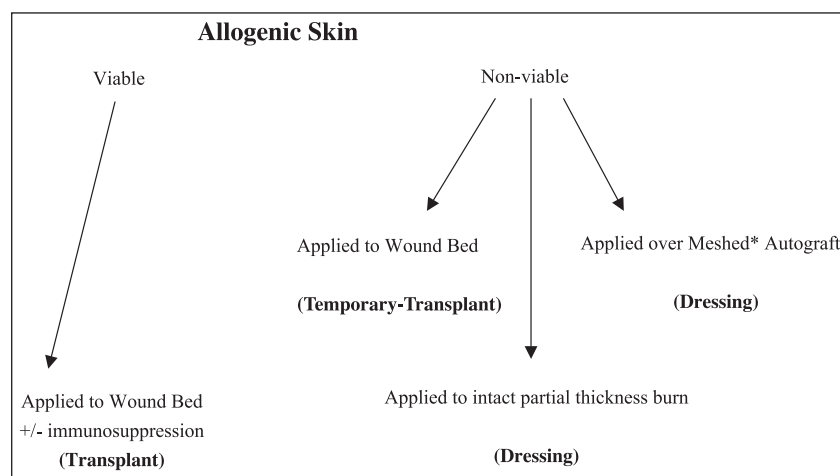


Fig. 5 A schema for determining the nature of allogenic skin. *If mesh ratio small enough for interstices to re-epithelialize before temporary 'take' of allograft i.e. <1:3.

burn. There are cultural limitations to the use of certain xenografts; porcine sources are not acceptable among Muslim patients, for example.

Autologous tissue has another role in dermal replacement. This involves the strategy to reduce scar formation by reducing the need for repair. The source may be acellular, highly processed dermis such as Alloderm, or it could be the incorporated dermis from either viable or nonviable cadaver skin. Cuono has described a strategy of using this dermal bed and shaving off the allogenic epidermis, replacing it with autologous keratinocytes.^[17] The nature of autologous applications is changing, however, particularly with regard to the keratinocytes. Autologous keratinocytes are the key cells for survival of the patient because of the essential nature of the protective function of the stratum corneum.

In 1975, when Rheinwald and Green first described the technique of producing keratinocyte cultures in vitro, it appeared that a new era of burns management was about to commence.^[18] Indeed it has, but the expectation of rapid wound closure with laboratory-produced cultured epithelial cell autografts (CEAs) has not been fulfilled in clinical practice. The process of preparation is expensive. Although it takes only a few weeks to create enough CEAs from a few square centimeters of original biopsy to cover the entire body, the actual take and survival of the CEAs on the body has been a disappointment. Keratinocyte delivery vehicles have been developed to facilitate the process and improve the rate of successful engraftment, but this has further increased the costs of this strategy while not significantly increasing the success.^[19]

There has been a tremendous amount of research and development in keratinocyte preparation and application, and there has been a parallel development in the area of dermal replacement and/or regeneration. One major problem that has faced workers in this field is how to combine the two layers of the skin. The dermo-epidermal junction (DEJ) is extremely complex.

It is a combined unit with structural components produced by both keratinocytes and by dermal fibroblasts. In nature, a new DEJ is formed when proliferating keratinocytes migrate across the dermal collagen. This occurs in all superficial wounds and in healing donor sites. There is no problem in the function of a DEJ in these situations, and this has focused attention on the role of the undifferentiated, proliferating keratinocyte cell suspensions in burn wound coverage. Fiona Wood, a pioneer in this field, has used cell suspensions of keratinocytes to treat burns in an increasing number of clinical applications, including dermal collagen remodeling, pigment expression, and keratin expression.^[20] Thus, the interaction between the products of the keratinocytes and the dermal fibroblasts becomes as important as the speedy regeneration of the lost epidermis.

Antimicrobial Dressings

Biological dressings will not be effective in the presence of infection. The treatment and prevention of infection encompasses another vast range in the field of burn dressing. The most common topical antimicrobial agents currently used in burns care are silver based. The history of the use of silver in burn wound care has been extensively reviewed by Klasen.^[21,22]

The contemporary use of silver really dates from 1965, when Moyer described his observations on silver nitrate solution. Through clinical observation, Moyer determined that a 0.5% solution topically applied to burn wounds did not impede the process of re-epithelialization, but this concentration was quite toxic to bacteria when evaluated in in vitro studies. Silver nitrate solution had to be reapplied at frequent intervals to maintain an effective local concentration of silver; it was thus labor-intensive and also very messy. Nevertheless, it was associated with a major reduction in lethal infection.

A significant development in topical silver preparation was the introduction of silver sulphadiazine cream. This combination of silver with a sulphonamide, in a cream formulation, heralded a dramatic reduction in the once-feared *Pseudomonas septicemia*. The cream was liberally applied to the wound and then covered with a paraffin gauze layer, several layers of absorbant gauze or Gamgee, and a retaining crepe bandage. This became and remains the standard burn dressing in many centers world-wide. Silver sulphadiazine is an excellent dressing where surgical expertise is limited and facilities are poor. It promotes desloughing of the burn wound, initially by maceration. As this process becomes established in the burn wound there is considerable exudate, and the process of dressing change becomes more demanding on nursing staff. Breakthrough requires repadding of dressings or twice daily changes. The mechanism of action involves stimulating the autolytic process through an enhanced inflammatory response. The viable tissue below the eschar demonstrates increased leucocytosis. As observed earlier, when healing is associated with inflammation there is liable to be more scarring.

A further concern about silver sulphadiazine cream is that when surgery is performed the macerated, hyperemic tissue is more hostile for the surgeon. As the trend toward a more aggressive surgical management of the burn wound developed, surgeons became more critical of silver sulphadiazine cream. At the same time, a new generation of nurses who had not experienced the devastating pseudomonas septicemia began to express dissatisfaction with the “messy” application of silver sulphadiazine. A new topical preparation of silver—flammacerium—appeared to address both of these concerns.^[23]

Flammacerium is a combination of flammazine, the silver sulphadiazine combination, and cerium nitrate. The effect of this formulation was to provide the antibacterial properties of the flammazine along with improved handling properties due to the tannalizing effect of the cerium nitrate. Cerium acts on nonviable collagen, and instead of creating a macerated eschar, the burn wound develops a crust like eggshell. It is dry and favorable for the surgeons, and once formed the dressing requirement is reduced. A further advantage of flammacerium is that there is less inflammatory response in the subeschar tissue and thus less bleeding at surgery—and more importantly from the patient’s perspective, less hypertrophic scarring.

The development of silver-based dressings continues, and new products are emerging in which the silver is incorporated into a material. One such product is Acticoat, developed by Burrell, which heralds a new concept of dressings in the era of nanotechnology.^[24] Acticoat is a nonadherent, trilaminar dressing

comprising a core of polyester and absorbent rayon with polyethylene mesh on either side. This mesh is coated with nanocrystalline silver, using a physical vapor deposition. When applied to a wound and activated by moistening with water, Acticoat delivers a controlled, sustained, antimicrobial dose of 50–100 mg/ml of silver ions (0.005–0.01% silver). The dressing can remain intact for three to four days, and newer preparations can remain unchanged for even longer (Fig. 6). This makes Acticoat an extremely attractive dressing from the perspective of the nursing management of burns patients.

Another approach to delivering silver to the wound has been pioneered by Dr. A. B. Flick, the Silverlon range of products manufactured by Argentum Medical is based on silver-plated fabric. This product range includes elastic burn wraps as well as absorbent burn dressings. Silver-impregnated dressings have another interesting aspect regarding wound healing: The silver is highly conductive. Using a current density of $0.6 \mu\text{A}/\text{cm}^2$, Chu and McManus at the Army Surgical Research Center at Fort Sam Houston studied the effect of using silver-coated fabrics as anodes. The biological effects observed were increased antibacterial effects, modified vascular permeability, enhanced dermal blood flow, and enhanced rate of donor site wound healing.^[25]

Yet another silver-impregnated product to be promoted for burn dressing is Aquacel Silver, produced by Convatec. Aquacel is a hydrofiber dressing promoted for its absorbent properties. The incorporation of silver gives an antibacterial dimension to the product.

There are many other wound dressings that have antibacterial properties by virtue of medical additives. Such products include chlorexidine-impregnated paraffin gauze, and povidone-iodine impregnated gauzes.

Nonbiological, Nonmedicated Dressings

The final category of dressings to be discussed includes nonbiological, nonmedicated materials developed from the classic concept of wound dressings. Natural fabric derived from wool contains lanolin and is thus poorly absorbent. Lawrence describes wound dressings prior to the latter half of the nineteenth century that were often based on well washed old rags and spent cloth that were more absorbent than new cloth.^[26] Cotton wool in its natural state is also nonabsorbent, because of natural greases.

In 1880 the surgeon Joseph Sampson Gamgee discovered that bleaching cotton wool rendered it highly absorbent. Gamgee tried to contain the wool in an open-weave gauzelike material. The resultant dressing, known epinonymously as Gamgee, is still in widespread use today. One of the main disadvantages of



Fig. 6 This child is a septic burn patient being nursed in the ICU. The dressing is Acticoat, which is applied to the wound with no absorbent or retention dressings. The Acticoat needs to be kept moist to activate the silver it contains.

using the cotton wool is its adherence to the burn wound. During the First World War, Lumiere devised the wide-meshed cotton-weave gauze impregnated with soft, medical-grade paraffin. This dressing alone and its medicated versions referred to earlier are commonly applied to small and superficial burns.

In 1962 Winter published the classic paper in the journal *Nature* describing the process of healing in wounds dressed with nonocclusive dressings and in those dressed with occlusive dressings.^[27] This observation in a porcine model, subsequently verified in human experiments, led to the development of new dressing materials based on film technology.

Films represent the most extensive group of synthetic material dressings, and the therapeutic strategy is to cover the wound to prevent bacterial invasion while allowing for some water loss. Clingfilm is possibly the most common film dressing used in burns care. A polyvinylchloride film widely used in food preparation and storage, its principal application is in the emergency treatment of burns to cover and protect burn wounds and later, in the burns unit, to cover wounds between dressing changes to allow for clinical inspection. Clingfilm and other commercially available polyvinylchloride films have excellent flexibility and mechanical properties, but their low rate of water vapor transmission leads to pooling of exudates under the dressing with prolonged use.

The discovery of a methyl methacrylate film called Hydron was a major advance in film technology. It had a high permeability to water vapor and yet was occlusive to bacteria. It was thus capable of creating the optimum environment for re-epithelialization, but it avoided the buildup of bacteria associated with earlier film dressings. This initial material was subsequently

modified to a vapor-permeable polyurethane, Opsite, that had similar properties but was more flexible. Trials indicated it considerably enhanced the rate of re-epithelialization of burns. Another interesting feature concerned the nature of the exudate that collected under Opsite dressings. It was rich in leucocytes, and the dressing appeared to exert an *in vitro* antibacterial effect.

Another polyurethane dressing that has widespread clinical application is Tegaderm. Nonadhesive coated films are also available. Omiderm is a highly vapor-permeable polyurethane that adheres to superficial wounds without the need for additional retention.

Film dressings have been medicated are also available. It should be pointed out that films can also be used in conjunction with other medicated dressings, as well as topical antibacterial preparations.

Gel dressings have an established place in the spectrum of burn dressings. The advantage of gels is that they are highly conformable and impermeable to bacteria, but are permeable to water and oxygen. They can also act as carriers for antimicrobial agents. They lack durability, however, but are useful for short-term application. One example is Geliperm, a hydrophilic gel composed of agar and polyacrylamide that is an excellent dressing to protect eyelid skin in the ventilated burn patient. Further advances in developing gels for dressings, using nanotechnology, are anticipated.

Foams and sponges have particular advantages as dressing materials. They are absorbent, their thickness provides protection to the wound, and they are relatively inexpensive and easy to apply. Lyofoam is a polyurethane foam manufactured from a soft polyether foam subjected to further processes of heat and pressure to form a smoothly hydrophilic, porous

surface membrane. It has been targeted at the burns market. Another foam product is Allevyn, produced by Smith and Nephew and marketed as a donor site dressing. This product has been shown to be clinically effective and is widely used.

Hydrocolloid dressings have an established place in burns treatment. One example of a hydrocolloid is Granuflex/Duoderm, marketed by Convatec. This dressing has an outer layer of polyurethane foam bonded to an inner adhesive layer of pectins, methyl cellulose, and other materials. The hydrocolloid liquefies on the wound, and the enclosed environment promotes re-epithelialization while suppressing bacterial colonization.

The final dressings to be described are the alginates, derived from seaweed. The common preparation is calcium alginate, which forms a gel when moistened. The calcium does have a hemostatic effect, and the gel-like alginate does appear to promote epithelialization. For this reason, calcium alginate is a popular dressing for skin graft donor sites.

This overview of burn dressings does not attempt to be an exhaustive list; the range of commercially available products is vast. Examples have been given of different types of dressing, but this is not meant as a recommendation. New dressings will continue to be developed and older methods rediscovered.

THE FUTURE

The spectrum of burn dressings is vast, but this is in part due to the spectrum of burn injury. The principles remain, though:

1. Promote healing
2. Prevent infection
3. Protect from further injury
4. Pain relief

As we move ever further into the era of regenerative medicine, the nature of burn dressings is going to change to reflect an increasing focus on the ultimate goal—skin regeneration.^[28] The previous philosophy of burn dressings has been to optimize the natural process of repair and prevent infection. The natural process of repair is not good enough, however. The result of repair is scar, and scar is the enemy. The burn dressing of the future must modulate healing. It must, in fact, prevent repair and promote regeneration. As our understanding of the biology of wound healing increases, we may well have to confront some paradoxes. Our present strategy is to speed up healing and to promote re-epithelialization. It may be that we should be paying more attention to the dermis, stimulating the mesodermal stem cells and keeping our

window of opportunity—the defect in the epidermis—open for a little longer.

ARTICLES OF FURTHER INTEREST

Allografts; Antimicrobial/Antibiotic (Infection Resistance) Materials; Biofilms; Fibrin; Fibrin Sealants; Sili-cones; Tissue Engineering of Skin; Xenografts

REFERENCES

1. Eady, R.A.J.; Leigh, I.M.; Pope, F.M. Anatomy and Organization of Human Skin. In *Textbook of Dermatology*, 6th Ed.; Champion, R.H., Burton, J.L., Burns, D.A., Breathnach, S.M., Eds.; Blackwell Science: Oxford, 1998; Vol. 1, 37–111.
2. Archer, C.B. Functions of the Skin. In *Textbook of Dermatology*, 6th Ed.; Champion, R.H., Burton, J.L., Burns, D.A., Breathnach, S.M., Eds.; Blackwell Science: Oxford, 1998; Vol. 1, 113–122.
3. Kramer, G.C.; Lund, T.; Herndon, D.N. Pathophysiology of Burn Shock and Burn Edema. In *Total Burn Care*, 2nd Ed.; Herndon, D.N., Ed.; W.B. Saunders: New York, 2002; 78–87.
4. Mitchell, R.N.; Cotran, R.S. Tissue Repair: Cell Regeneration and Fibrosis. In *Basic Pathology*, 7th Ed.; Kumar, V., Cotran, R.S., Robbins, S.L., Eds.; Saunders: Philadelphia, 2003; 61–78.
5. Hermans, M.H.E. Results of a survey on the use of different treatment options for partial and full thickness burns. *Burns* **1998**, *24*, 539–551.
6. *The Chambers Dictionary*; Schwarz, C.M., Ed.; Chambers: Edinburgh, 1993.
7. Quinn, K.J.; Courtney, J.M.; Evans, J.H.; Gaylor, J.D.S. Principles of burn dressings. *Biomaterials* **1985**, *6*, 369–377.
8. Pruitt, B.A. The evolutionary development of biologic dressings and skin substitutes. *J. Burn Care Rehabil.* **1997**, *18*, S2–S5.
9. Jones, I.; Currie, L.; Martin, R. A guide to biological skin substitutes. *Br. J. Plast. Surg.* **2002**, *55*, 185–193.
10. Balasubramani, M.; Kumar, T.R.; Babu, M. Skin substitutes: A review. *Burns* **2001**, *27*, 534–544.
11. De Backere, A.C.J. Euro skin bank: Large scale skin-banking in Europe based on glycerol-preservation of donor skin. *Burns* **1994**, *20*, S4–S9.
12. Kearney, J.N. Clinical evaluation of skin substitutes. *Burns* **2001**, *27*, 545–551.
13. Herndon, D.N. Perspectives in the use of allograft. *J. Burn Care Rehabil.* **1997**, *18*, S6.
14. Kealey, G.P. Disease transmission by means of allograft. *J. Burn Care Rehabil.* **1997**, *18*, S10–S11.
15. Heck, E. Operational standards and regulations for tissue banks. *J. Burn Care Rehabil.* **1997**, *18*, S11–S12.
16. Burd, A. Glycerolised allogenic skin: Transplant or dressing? A medico-legal question. *Burns* **2002**, *28*, S34–S39.

17. Cuono, C.B.; Langdon, R.; Birchall, N.; Barttelbort, S.; McGuire, J. Composite autologous-allogeneic skin replacement: Development and clinical application. *Plast. Reconstr. Surg.* **1987**, *80*, 626–637.
18. Rheinwald, J.G.; Green, H. Serial cultivation of strains of human epidermal keratinocytes: The formation of keratinizing colonies from single cells. *Cell* **1975**, *6*, 331.
19. Boyce, S.T. Design principles for composition and performance of cultured skin substitutes. *Burns* **2001**, *27*, 523–533.
20. Navarro, F.A.; Stoner, M.L.; Lee, H.B.; Park, C.S.; Wood, F.M.; Orgill, D.P. Melanocyte repopulation in full-thickness wounds using a cell spray apparatus. *J. Burn Care Rehabil.* **2001**, *22*, 41–46.
21. Klasen, H.J. A historical review of the use of silver in the treatment of burns. II. Renewed interest for silver. *Burns* **2000**, *26*, 131–138.
22. Klasen, H.J. Historical review of the use of silver in the treatment of burns. I. Early uses. *Burns* **2000**, *26*, 117–130.
23. Boeckx, W.; Focquet, M.; Cornelissen, M.; Nuttin, B. Bacteriological effect of cerium-flamazone cream in major burns. *Burns Incl. Therm. Inj.* **1985**, *11*, 337–342.
24. Yin, H.Q.; Langford, R.; Burrell, R.E. Comparative evaluation of the antimicrobial activity of ACTICOAT antimicrobial barrier dressing. *J. Burn Care Rehabil.* **1999**, *20*, 195–200.
25. Chu, C.S.; McManus, A.T.; Matylevich, N.P.; Mason, A.D., Jr.; Pruitt, B.A. Enhanced survival of auto-epidermal-allodermal composite grafts in allosensitized animals by use of silver-nylon dressings and direct current. *J. Trauma* **1995**, *39*, 273–278.
26. Lawrence, F.C. Dressings for Burns. In *Principles and Practice of Burns Management*; Settle, J.A.D., Ed.; Churchill Livingstone: London, 1996; 259–269.
27. Winter, G.D. Formation of the scab and rate of epithelialization of superficial wounds in the domestic pig. *Nature* **1962**, *193*, 293–294.
28. Martin, P. Wound healing—Aiming for perfect skin regeneration. *Science* **1997**, *276*, 75–81.

Calcium Phosphate Ceramics: New Generation Produced in Japan

Atsuo Ito

*Advanced Biomaterials Group, Institute for Human Science and Biomedical Engineering,
National Institute of Advanced Industrial Science and Technology, Tsukuba, Ibaraki, Japan*

Hajime Ohgushi

*Tissue Engineering Research Group, Research Institute for Cell Engineering (RICE), National
Institute of Advanced Industrial Science and Technology, Amagasaki City, Hyogo, Japan*



INTRODUCTION

Many types of calcium phosphate biomaterials are clinically used in Japan, including calcium phosphate ceramics, calcium phosphate cements (CPCs), and composites. The first calcium phosphate biomaterial to be officially approved for clinical use was dense hydroxyapatite [HA: $\text{Ca}_5(\text{PO}_4)_3\text{OH}$] in the form of a dental implant in 1983. Although the application of dense HA to a dental implant was unsuccessful, a variety of dense and porous calcium phosphate ceramics have been developed as bone grafting materials since then: five HA, one tricalcium phosphate, and one biphasic calcium phosphate products have been officially approved and are commercially available. They vary widely in porosity, porous structure, mechanical strength, surface chemistry, and biological reaction. Calcium phosphate cement is an injectable paste-like material that hardens in the body. Two different classes of CPC have been developed; one has been officially approved and the other is pending approval. The calcium phosphate composites described in this entry are the recently developed, calcium phosphate/collagen, calcium phosphate/gelatin, and calcium phosphate/poly L-lactic acid composites, all of which have distinctive features. Two of the composites are pending approval and the rest have been officially approved. All these calcium phosphate biomaterials are described in this article with a brief introduction to the calcium phosphate family to help understanding of calcium phosphate biomaterials.

The addition of an osteogenic ability or osteoinductive property to calcium phosphate biomaterials holds great promise for treating patients with diminished bone forming ability. The calcium phosphate biomaterials are osteoconductive materials that only allow the formation of bone on their surface by serving as a scaffold. In patients with diminished bone forming ability, bone tissues do not necessarily bond to calcium phosphate biomaterials at a clinically satisfactory level. If the calcium phosphate biomaterials possess osteoinductive properties, the expansion of clinical applications for the

ceramics is certain. In this context, the addition of bone marrow cells to calcium phosphate biomaterials is promising. The present article also describes this tissue engineering approach in which bone marrow cells are added to calcium phosphate biomaterials.

NEW GENERATION OF CALCIUM PHOSPHATE CERAMICS PRODUCED IN JAPAN

Calcium Orthophosphates

There are many types of calcium orthophosphates (Table 1). Calcium orthophosphate is simply termed calcium phosphate in this article. Calcium phosphates are conventionally classified according to their Ca/P molar ratio, which varies from 0.5 to 2.0. Monocalcium phosphate anhydrous [MCPA: $\text{Ca}(\text{HPO}_4)_2$] and monocalcium phosphate monohydrate [MCPM: $\text{Ca}(\text{HPO}_4)_2 \cdot \text{H}_2\text{O}$] are soluble calcium phosphates, unlike all the rest of the calcium phosphates. Dicalcium phosphate dehydrate (DCPD: $\text{CaHPO}_4 \cdot 2\text{H}_2\text{O}$) and dicalcium phosphate anhydrous (DCPA: CaHPO_4) are used as powder components of CPCs. Octacalcium phosphate [OCP: $\text{Ca}_8\text{H}_2(\text{PO}_4)_6 \cdot 5\text{H}_2\text{O}$] is a transient intermediate phase formed in the precipitation of the thermodynamically more stable HA or calcium-deficient HA. Amorphous calcium phosphate (ACP) is an amorphous phase with a typical Ca/P molar ratio of 1.50. Amorphous calcium phosphate is precipitated in moderately or highly supersaturated calcium phosphate solutions at or above neutral pH. It tends to convert into thermodynamically more stable OCP or calcium-deficient HA. Tricalcium phosphate [TCP: $\text{Ca}_3(\text{PO}_4)_2$] has three polymorphs β , α , and super α , with transition temperatures $\beta \rightarrow \alpha$ at 1125°C and $\alpha \rightarrow \text{super } \alpha$ at 1430°C . The β phase (β -TCP) powder is converted into a dense solid, thus sintered into a ceramic body, when the powder is compacted followed by heating at a temperature in the range of $700\text{--}1125^\circ\text{C}$. The α phase (α -TCP) powder is

Table 1 Calcium orthophosphate family

Ca/P	Abbreviation	Composition	Feature or application
0.50	MCPA	Ca(HPO ₄) ₂	Soluble in water
0.50	MCPM	Ca(HPO ₄) ₂ · H ₂ O	Soluble in water; cement powder
1.00	DCPD	CaHPO ₄ · 2H ₂ O	Cement powder
1.00	DCPA	CaHPO ₄	Cement powder
1.33	OCP	Ca ₈ H ₂ (PO ₄) ₆ · 5H ₂ O	Transient intermediate phase; reaction product in cement setting
1.50	ACP	—	Amorphous; transient intermediate phase
1.50	β-TCP	Ca ₃ (PO ₄) ₂	Resorbable; sintered into ceramics
1.50	α-TCP	Ca ₃ (PO ₄) ₂	Cement powder
1.50	α'-TCP	Ca ₃ (PO ₄) ₂	Appears only at above 1430°C
1.50–1.67		Calcium deficient HA	Low or moderately crystalline; decomposes above approx. 700°C; low crystalline material is resorbable
1.67	HA	Ca ₅ (PO ₄) ₃ OH	Low to highly crystalline; sintered into ceramics; low crystalline material is resorbable; highly crystalline material is nonresorbable and osteoconductive
2.00	TeCP	Ca ₄ (PO ₄) ₂ O	Cement powder

MCPA, monocalcium phosphate anhydrous; MCPM, monocalcium phosphate monohydrate; DCPD, dicalcium phosphate dehydrate; DCPA, dicalcium phosphate anhydrous; OCP, octacalcium phosphate; ACP, amorphous calcium phosphate; TCP, tricalcium phosphate; HA, hydroxyapatite; TeCP, tetracalcium phosphate.

hardly sintered. The α-TCP powder is used for CPCs. The super-α phase is unable to survive quenching to room temperature. Hydroxyapatite is the most stable, and hence the most insoluble, among the calcium phosphates in neutral and alkaline solutions. Hydroxyapatite powder compact is sintered into a ceramic body by heating at a temperature in the range of 700–1200°C. Hydroxyapatite decomposes at temperatures higher than 1200°C. Hydroxyapatite with a Ca/P molar ratio lower than 1.67 is termed calcium-deficient HA. Calcium-deficient HA is formed without difficulty by aqueous reactions between calcium ions and phosphate ions. Calcium-deficient HA powders can be sintered into a biphasic ceramic body consisting of β-TCP and HA. Tetracalcium phosphate [TeCP: Ca₄(PO₄)₂O] is used for CPCs.

Biological reactions to HA and calcium-deficient HA depend on their crystallinity. Hydroxyapatite and calcium-deficient HA have low crystallinity when they are precipitated in aqueous solutions. Such low-crystallinity HA or calcium-deficient HA has an extremely small particle size (usually some tens of nanometers), and thus a large surface area, high reactivity, high dissolution rate, and high adsorption property. Low-crystallinity HA or calcium-deficient HA can be resorbed in bone tissue. When powders or powder compacts of these low-crystallinity materials are heated, there is an increase in crystallinity, and in particle or grain size. Moreover, calcium-deficient HA decomposes into β-TCP and highly crystalline HA upon being heated.

The highly crystalline HA, typically a ceramic HA, is a nonresorbable and osteoconductive material.

CALCIUM PHOSPHATE BIOMATERIALS PRODUCED IN JAPAN

Hydroxyapatite

The crystalline phase of natural bone is basically HA and thus HA ceramics have extensively been used as a bone graft substitute. Dense or porous ceramic bodies as well as granules are available. There is diversity in pore size, porosity, interconnectivity of pores, and mechanical strength depending on manufacturers.

Bonfil[®] (Pentax Co.)

Porous blocks and granules have been officially approved for orthopedic use. Bonfil[®] is sintered at 900°C, which is the lowest sintering temperature among the commercially available HA ceramic bodies in Japan. The sintering temperature is selected so that the porous ceramic has the maximum bone ingrowth. Bonfil[®] has compressive and bending strengths of 15 MPa and 2–3 MPa, respectively. All the pores are of irregular shape with a size of 200–300 μm, and have interconnectivity. The porosity is 60–70%. The previous manufacturer of Bonfil[®] was Mitsubishi Materials Co. The

Biomaterial Business Center of Mitsubishi Materials Co. has merged with Pentax Co. since October 2004.

Bonetite[®] (Pentax Co.)

Porous blocks and dense granules have been officially approved for dental use. Bonetite[®] is sintered at 1200°C. The pores of the porous block are of irregular shape with a size of 200 μm and a porosity of 70%. The previous manufacturer of Bonetite[®] was Mitsubishi Materials Co. The Biomaterial Business Center of Mitsubishi Materials Co. has merged with Pentax Co. since October 2004.

Bonoceram[®] (Sumitomo Osaka Cement Co., Ltd.)

Porous blocks, porous granules, and dense bodies, all sintered at 1150°C, have been officially approved for orthopedic use. The porous bodies are termed Bonoceram-P[®] while the dense bodies are termed Bonoceram-K[®]. The porous body has a compressive strength of 44.1–68.6 MPa and a bending strength higher than 8.33 MPa. The pores are spherical in shape with a size of 50–300 μm. The porosity is 35–48%. The dense body has a bending strength higher than 58.8 MPa.

Apaceram[®] (Pentax Co.)

Porous blocks, porous granules, and dense bodies have been officially approved for dental, orthopedic, and brain surgery uses. The porous blocks are applied to dental and orthopedic cases as well as to filling the bone defects formed in brain surgery. The granules are applied to dental and orthopedic cases, and the dense bodies are used as plugs to fill the hole formed in brain surgery.

The porous blocks and granules, sintered at 1200°C, have a porosity of 15–60%. The compressive and bending strengths vary from 16 to 250 MPa and from 8 to 47 MPa, respectively, depending on the porosity (Table 2). The porous bodies have an interconnected bimodal pore configuration consisting of spherical macropores about 300 μm in size and micropores about 2 μm in size with irregular shapes. The dense body, sintered at 1200°C, has a porosity of less than 0.8%. The compressive and bending strengths of the dense body are 750 and 210 MPa, respectively.

Neobone[®] (Toshiba Ceramics Co., Ltd.)

Porous blocks and granules have been officially approved for orthopedic use. Neobone has fully interconnected pores. The interconnected pores are produced by sintering foamed and solidified hydroxyapatite slurry.

Table 2 Mechanical strength of Apaceram[®]

Porosity (%)	Three-point bending strength (MPa)	Compressive strength (MPa)
0	210	750
15	47	250
30	25	180
40	16	66
50	11	30
55	9.5	21
60	8.0	16

(Courtesy of Mr. T. Ogawa, Pentax Co.)

The slurry contains hydroxyapatite and a polyethyleneimine aqueous solution, and is foamed with the addition of polyoxyethylenelaurylether. For solidification, a polyfunctional epoxy compound is used. The solid and foamy body is dried and sintered at 1200°C.

Neobone has a porosity of 72–78% and a compressive strength of 8–10 MPa. All the pores are spherical in shape and approximately 100 μm in diameter with uniform interconnections among adjacent pores. The interconnecting pores are 10–80 μm in diameter. The pore walls are densely sintered, which leads to the relatively high compressive strength.

Bone ingrowth throughout the pores has been observed in Neobone: new bone with vessels and abundant bone marrow entered the pores located at the central part of the cylindrical body 6 mm in diameter and 15 mm in length implanted in the femoral condyles of rabbits for 6 weeks. Owing to the dominant bone ingrowth, the compressive strength of Neobone gradually increased from its original value (10 MPa) with increasing implantation period, and finally reached about 30 MPa 9 weeks after the implantation.^[1]

β-Tricalcium Phosphate

In contrast to HA ceramics, a β-TCP ceramic is gradually degraded with time in bone tissue. A β-TCP ceramic is called “resorbable ceramic,” which means that the β-TCP ceramic has the ability to be replaced by new bone tissue during the process of degradation.

Osferion[®] (Olympus Optical Co., Ltd.)

Porous blocks and granules have been officially approved for orthopedic use. Osferion has a porosity of 75% with a sponge-like interconnected bimodal pore configuration. The macro- and micropores, both irregular in shape, are 100–400 μm and 1 μm in diameter, respectively. The sintering temperature is 1050°C. The



compressive strength is very low (2 MPa). Owing to the low strength, additional grafting of cancellous bone together with cortical bone with internal fixation may be required when the ceramics are used in high load-bearing areas including those just beneath the knee joint. Furthermore, a long period is required before the start of load bearing compared with the use of other types of ceramic. Osferion is completely resorbed in a period ranging from 6 mo to a few years, with replacement by regenerated new bone tissue in many cases.

Osferion has a distinctive feature in its powder preparation method.^[2] Unlike the conventional aqueous precipitation methods, β -TCP powder is prepared by mechanochemical reaction between DCPD and calcium carbonate powder in water. In this method, DCPD and calcium carbonate are mixed with a Ca/P molar ratio of 1.50 with the addition of water followed by grinding. During the wet grinding, a chemical reaction takes place between DCPD and calcium carbonate. The reaction product is calcium-deficient HA with a Ca/P molar ratio of 1.50. Pure β -TCP is obtained by heating the reaction product. This method is superior to the conventional precipitation methods in its simplicity of powder separation from water and in the control of Ca/P molar ratio.

Biphasic Ceramic

Ceratite[®] (NGK Spark Plug Co., Ltd.)

Porous blocks, porous granules, and dense bodies have been officially approved for orthopedic and brain surgery uses. Ceratite consists of 70% HA and 30% β -TCP. Ceratite is a nonresorbable material. Ceratite supports the osteogenic action of host bone tissue near the implantation site, resulting in new bone formation around the implant. Ceratite bonds directly to the newly formed bone. The porous Ceratite has two different classes: one is a microporous ceramic body with a pore diameter of 1–10 μm and a porosity of 35–40 %, and the other is a macroporous ceramic body with a pore diameter of 150–250 μm and a porosity of 45–55%. The microporous ceramic body, sintered at 1100°C, has compressive and bending strengths of 120–160 MPa and 30–40 MPa, respectively. The macroporous ceramic body, sintered at 1300°C, has compressive and bending strengths of 20–30 MPa and 7–10 MPa, respectively. The macroporous ceramic body has a bimodal pore size distribution: many micropores are present in the matrix surrounding the macropores. The macropores are fairly well connected three dimensionally to adjacent pores with interconnecting pores with a size of 60–80 μm . A dense Ceratite ceramic body, sintered at 1300°C, has

compressive and bending strengths of 600–700 MPa and 140–170 MPa, respectively.

Two clinical cases treated with granular and dense Ceratite are detailed below:

Case 1: Twenty-five-year-old male case (benign bone tumor at tibia side) Ceratite granules were mixed with autogenous bone chips and implanted in the excised bone tumor area. Six months after the operation, a radiograph showed the homogenous appearance of the implants.

Case 2: Fifty-three-year-old male case (cerebral aneurysm) Dense Ceratite buttons were implanted for a burr hole skull defect. The postoperative x-ray showed bone integration around the ceramic buttons.

Calcium Phosphate Cement

A CPC consists of aqueous liquid and finely pulverized calcium phosphate powders. Calcium phosphate cements date from the discovery of hardening of α -TCP in water reported by Monma and Kanazawa.^[3] Calcium phosphate cements are of considerable interest in orthopedics owing to their ability to be shaped easily during surgical operations, their ability to be injected with low operative invasion, their ability to harden in the body, and their good osteoconductivity. Calcium phosphate cements are used to fill bony defects and to augment bone tissue. The calcium phosphates used in the cement powder are precursors of calcium-deficient HA. The precursors convert into calcium-deficient HA or OCP by hydrolysis reactions in aqueous solutions. Such precursors include α -TCP, DCPD, DCPA, and TeCP. During the hydrolysis reaction, the calcium-deficient HA and OCP become tangled and join together, which leads to hardening of the cement (Fig. 1).

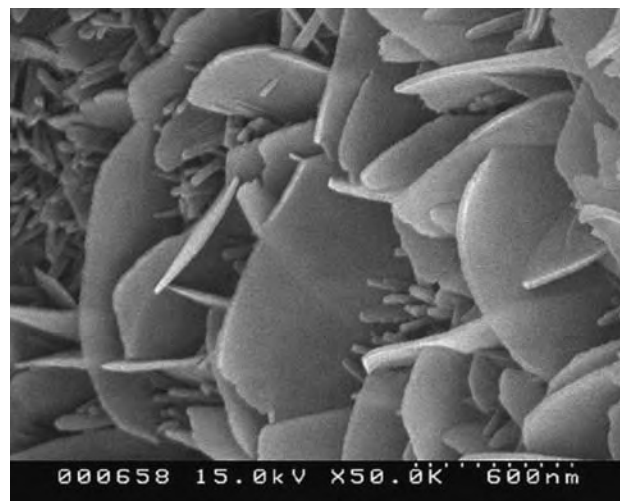
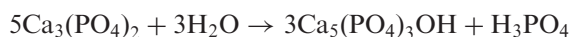
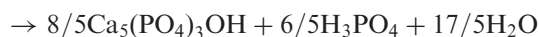
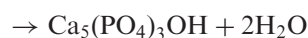
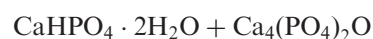
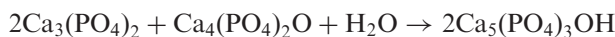


Fig. 1 Microstructure of hardened Biopex-R. (Courtesy of Dr. M. Hirano, Pentax Co.)

Biopex-R[®] (Pentax Co.)

Biopex-R has been officially approved for: 1) repairing bone defects; 2) repairing and strengthening/supporting broken bones; 3) reinforcing orthopedic screws and implants; and 4) filling gaps between cement-less artificial joints and bones. The previous manufacturer of Biopex-R was Mitsubishi Materials Co. The Biomaterial Business Center of Mitsubishi Materials Co. has merged with Pentax Co. since October 2004. Biopex-R contains α -TCP as the major powder component. Other powder components include TeCP, DCPD, HA, and magnesium phosphate. The aqueous liquid contains sodium chondroitin sulfate, sodium succinate, and sodium hydrogensulfite. The cement hardens in 6–10 min at 37°C when the powder and liquid are mixed at a powder/liquid (P/L) ratio in the range from 3.3 to 4. The ideal hardening reactions can be described as follows:



However, in practice the HA finally formed is either calcium-deficient HA or HA containing sodium and carbonate ions instead of stoichiometric HA.

The hardening speed and final mechanical strength depends on the temperature and P/L ratio.^[4] The higher the temperature and P/L ratio, the higher the hardening speed and strength. A hardened body, which was set at 37°C for 1 hr with the P/L ratios of 2.5, 3.0, and 3.5, was subjected to four-point bending and compression tests after immersion in a phosphate buffer solution for 7 days. The maximum bending and compressive strength were 10.5 ± 1.1 MPa and 50.2 ± 4.3 MPa, respectively, which were attained at a P/L ratio of 3.5. The compressive strength is higher than that of human tibial cancellous bone (1.9 ± 4.3 MPa), morselized bone graft (1–2 MPa) and porous hydroxyapatite ceramic with a porosity of 60% (15 MPa) although the bending strength is low. The cement paste was inserted and hardened in a hole 5.2 mm in diameter and 15 mm in depth at femoral condyle of rabbit. After 1 day, 3 weeks, and 6 weeks, the push-out test was performed. The push-out strength was 1.4 MPa after 1 day and

2 MPa after 6 weeks at a P/L ratio of 3.5. All these mechanical strength data indicate that the cement is applicable only to non-load-bearing sites.

Biopex-R[®] is used to maintain normal bone architecture after reduction of fractures such as radius (metaphyseal areas), tibia plateau, calcaneus, and vertebral body (compressioned types) fractures.^[5–8] Biopex-R is also used for unstable femoral intertrochanteric fractures with hip screw fixation devices, and for fixation of cementless artificial hip joints.^[9,10]

Using the cement, percutaneous transpedicular vertebroplasty procedures were performed in patients having thoracic or lumbar osteoporotic vertebral fractures to obtain early pain relief and improve the integrity of the osteoporotic vertebral body (VB). Following repositioning and curettage of the pathological soft tissues, cement-assisted vertebroplasty was percutaneously performed in four patients with osteoporotic burst fractures and pseudarthrosis (Procedure A). In situ cement-assisted vertebroplasty was performed in 12 patients with fresh vertebral compression fractures due to osteoporosis (Procedure B). Back pain and low-back pain were evaluated using a visual analog scale (VAS). The deformity index of the VB was measured on a lateral radiograph as the ratio of the VB's height (sum of measurements at anterior, middle, and posterior regions) to its longitudinal diameter. On the basis of VAS scores, pain was decreased in all patients immediately after surgery, and pain relief was maintained at the last follow-up. The mean preoperative deformity index score of the VB was 1.43 in Procedure A and 1.67 in Procedure B; postoperative scores improved to 1.59 and 1.93, respectively. At the 6 mo follow-up examination, the mean deformity index score rebounded to 1.52 in Procedure A and 1.79 in Procedure B. Bone union was documented in all patients. Complications, such as a temporary respiratory insufficiency and a small amount of cement leakage into the spinal canal, were observed in patients. It is concluded that percutaneous transpedicular cement-assisted vertebroplasty is a minimally invasive procedure that provides early relief of pain and prevents vertebral collapse and pseudarthrosis in patients with osteoporotic vertebral fractures.^[11]

TeCP-DCPA cement

A new TeCP-DCPA cement is pending approval. The cement uses an equimolar mixture of TeCP and DCPA as the powder. The liquid is a 50% solution of dextran sodium sulfate that effectively prevents the cement from absorbing excess water, and hence from disintegrating during the hardening reaction. The ideal hardening reaction is described as:



However, the HA finally formed contains sodium and carbonate ions in practice.

The cement reached the maximum compressive strength of 48 MPa in 8 hr; it was hardened at 37°C and 95% relative humidity after mixing the powder and the liquid. The cement was implanted in a rabbit femoral condyle for up to 52 weeks. A greater amount of new bone formation was observed around the cement than around sintered HA. The cement was considerably resorbed without inflammation: the cross-sectional area of the implanted cement decreased to as low as two-thirds of the initial area at 24 weeks as demonstrated by microcomputed tomography.^[12]

Composite of HA and Poly-L-Lactic Acid

SuperFIXSORB30[®] (Takiron Co., Ltd.)

SuperFIXSORB30 is the screw fixation device consisting of 30% HA and 70% poly-L-lactic acid (Fig. 2). SuperFIXSORB30 has been officially approved and commercialized since August 2003. SuperFIXSORB30 has the bending strength of 270 MPa and Young's modulus of 7.6 GPa.

SuperFIXSORB30 is osteoconductive and bioresorbable material. The HA used in SuperFIXSORB30 is low-crystallinity HA that is precipitated by a wet method without being subjected to any heat treatment. The HA, therefore, can be well resorbed. SuperFIXSORB30 bonds directly to the surrounding bone tissue without intervening fibrous tissue, and is gradually resorbed and totally replaced by newly formed bone tissue.

HA-Collagen Composite

Bonject[®] (Koken Co., Ltd.)

Bonject is a viscous and injectable bone filling material that has been officially approved for use in dentistry.

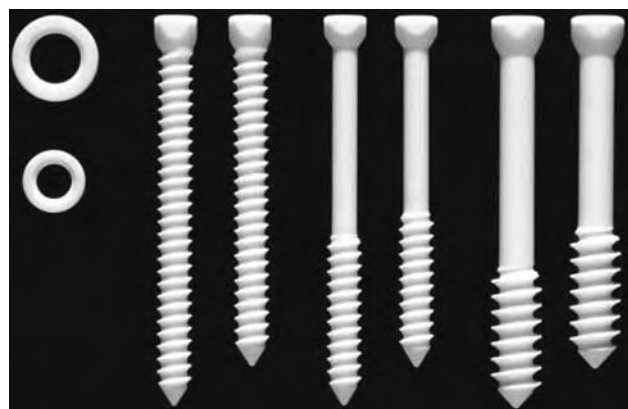


Fig. 2 SuperFIXSORB30[®] screw fixation devices. (Courtesy of Mr. R. Kashiwadani, Takiron Co., Ltd.)

Bonject is a mixture of HA granules and atelocollagen solution. The HA granules are synthesized by heating cattle bone at 1100°C. Two classes of particle size, 300–600 μm or 600–1000 μm, are available for the HA granules depending on clinical cases. The atelocollagen is derived from the dermis of cattle by enzymatic treatments and is highly purified. The HA to atelocollagen ratio is 3:2. The atelocollagen solution is sufficiently viscous to prevent the HA granules from leaking from the implanted site. Bonject is applicable to the bony defects caused by marginal periodontitis, radicular cysts, and resection of benign bone tumors.

Composite of HA and collagen type I

A novel composite of HA and collagen type I has been recently developed. The HA-collagen composite is synthesized by reaction between a collagen-containing phosphoric acid solution and a calcium hydroxide aqueous suspension followed by removal of excessive water. The collagen-containing phosphoric acid and the calcium hydroxide suspension are mixed at pH 7–9 and 25–40°C at a Ca/P molar ratio of 1.67 that corresponds to that of HA. The excessive water is removed by either high-pressure squeezing (200 MPa) or freeze-drying. The high-pressure squeezing is used to make solid HA-collagen composite while the freeze-drying is used to make highly porous and sponge-like composites. The solid HA-collagen composites contain different amounts of collagen depending on the reaction pH: 7–14 wt% at pH 7–8 and 17–19 wt% at pH 9. Water content is 10–20 wt%. The solid HA-collagen composite shows mechanical properties that are similar to those of cancellous bone: a three-point bending strength of 39.5 ± 0.9 MPa and Young's modulus of 2.54 ± 0.38 GPa.

The HA-collagen composite is an intertwined assembly of fibers 20 μm in length or longer. Each fiber consists of HA nanocrystals 50 nm in diameter bound to collagen fibrils 300 nm in length. The fibril axis and the crystallographic *c*-axis of HA aligned at an angle within 30°. This alignment of HA and collagen fibrils is very similar to that in bone tissue. The HA *c*-axis is best aligned when the composite is prepared at pH 9 and 40°C.

The solid HA-collagen composite is a resorbable material and can be used as the carrier of bone morphogenetic protein-2 (BMP-2). A composite 5 × 5 × 10 mm³ in size adsorbing 40 and 80 μg recombinant human BMP-2 (rhBMP-2) implanted in holes 6 mm in diameter in a beagle radius or ulna led to significant callus formation on the grafted composite as early as 2 weeks after the implantation. The average time needed until bone union was 9.2 ± 2.3 weeks for the 0 rhBMP-2 μg group, 7.6 ± 0.9 weeks for 40 rhBMP-2 μg, and 5.6 ± 0.9 weeks

for the 80 rhBMP-2 μg group. The time until bone union for the 80 rhBMP-2 μg group was significantly shorter than that for the 0 rhBMP-2 μg group ($p < 0.05$). The solid apatite–collagen composite $5 \times 5 \times 10 \text{ mm}^3$ in size adsorbing 80 μg rhBMP-2 used in anterior interbody fusion of the cervical spine of a beagle dog led to the enhancement of callus formation and bone bridging as well as a reduction in the intervertebral distance within 13 weeks, indicating the efficacy of this material in anterior interbody fusion of the cervical spine.^[13]

The sponge-like HA–collagen composite has a structure similar to that of cancellous bone. Owing to its highly porous nature, the sponge-like HA–collagen composite has sufficient internal spaces for osteoblasts and osteoclasts to enter. The sponge-like HA–collagen composite easily absorbs liquids and swells when it is immersed in a liquid, which is a clinically useful characteristic for the incorporation of any medicine and growth factor. The sponge-like HA–collagen composite shows distinctive elasticity when the composite absorbs water: the sponge-like HA–collagen composite deforms in response to the applied compressive stress, and it returns to its original shape after the stress is removed (Fig. 3). The sponge-like HA–collagen composite is scheduled to be commercialized in 2006 in Japan.

APPLICATION OF CALCIUM PHOSPHATE CERAMICS TO TISSUE ENGINEERING

Adding an osteogenic ability or osteoinductive property to calcium phosphate biomaterials holds great promise for treating patients with diminished bone forming ability. The calcium phosphate biomaterials are osteoconductive materials that only allow the formation of bone on their surface by serving as a scaffold. In patients with diminished bone forming ability, bone tissues do not necessarily bond to calcium phosphate biomaterials at a clinically satisfactory level. If the calcium phosphate biomaterials possess osteoinductive properties, the expansion of clinical applications for the ceramics is certain. Bone morphogenetic protein is a protein that demonstrates osteogenesis when implanted at subcutaneous and intramuscular sites. Therefore, BMP is osteoinductive.^[14] The composite of BMP and calcium phosphate ceramic enhances the inherent osteogenic ability of BMP.^[15,16] In contrast to the utilization of the osteoinductive activity of BMP, the composites of marrow cells and porous calcium phosphates such as HA show an osteogenic response when implanted in vivo.^[17] The osteogenic capability of the marrow cells resides in not only fresh marrow cells but also in in vitro culture-expanded mesenchymal stem cells (MSCs) derived from marrow cells.^[18]

Mesenchymal stem cells residing in marrow cell population can differentiate into osteoblasts and

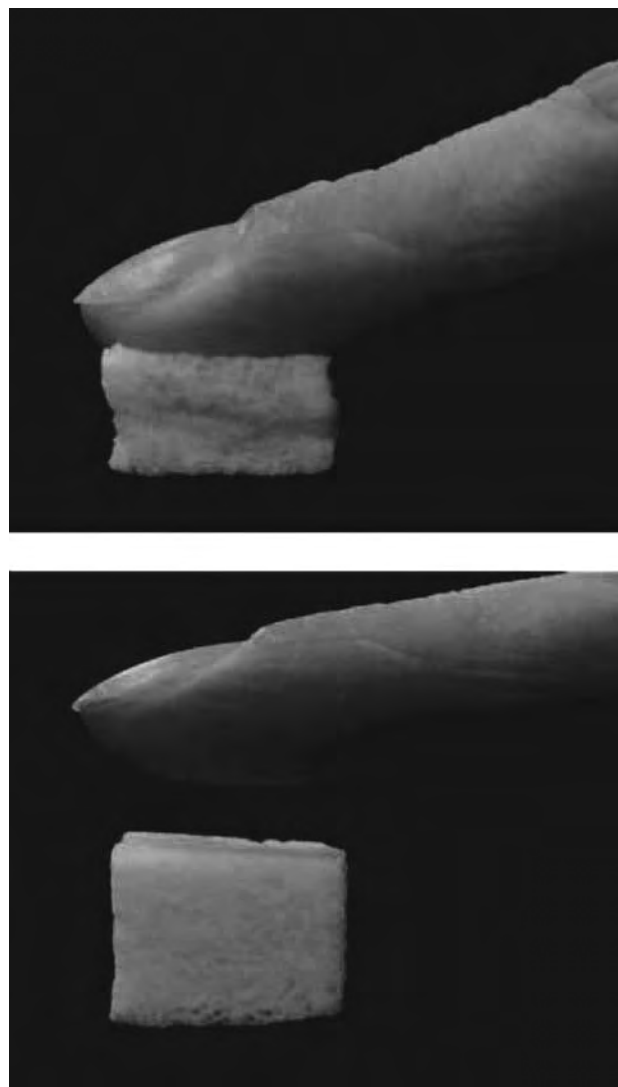


Fig. 3 The sponge-like HA–collagen composite. The sponge-like HA–collagen composite containing water deforms in response to the applied compressive stress (upper), and returns to its original shape after the stress is removed (lower). (Courtesy of Mr. D. Shoji, Pentax Co.)

produce mineralized bone matrix in in vitro culture conditions in the presence of dexamethasone and β -glycerophosphate.^[19,20] Although the number of MSCs in fresh marrow is extremely small, their number can be expanded without loss of their osteogenic potential by using a culture technique.^[21] When MSCs are cultured in a porous framework of HA under conditions where the stem cells can differentiate into osteoblasts and fabricate bone matrix, the porous areas of the HA are covered with a thin layer of the “in vitro bone” containing active osteoblasts.^[22] The construct can show the ability of immediate new bone formation after in vivo implantation (Fig. 4). On the basis of the fact that the construct fabricated with human cultured MSCs can

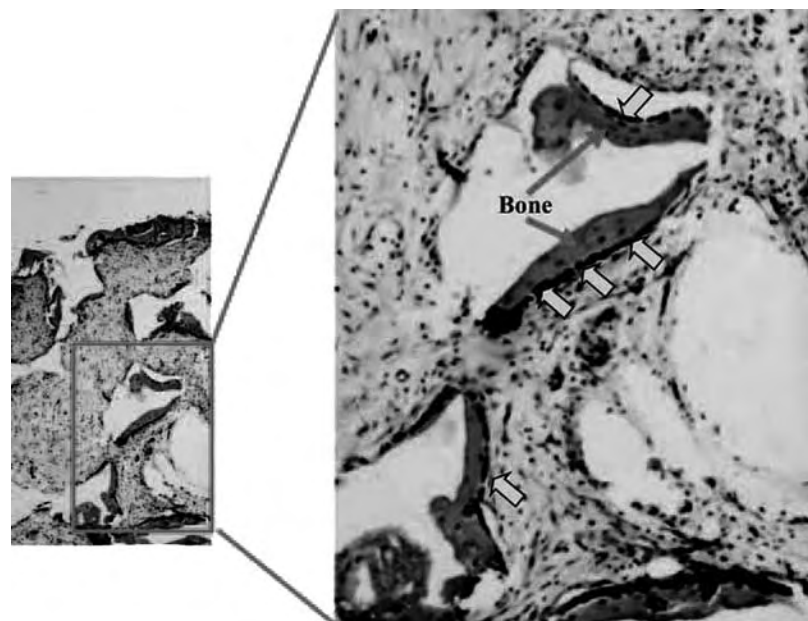


Fig. 4 Two weeks after subcutaneous implantation of HA covered with bone matrix/active osteoblasts (rat-cultured bone). Newly formed bone can be seen in the pore region of the HA ceramics. Arrows indicate the bone forming osteoblasts.

show promising osteogenic ability, we initiated clinical applications of various kinds of ceramic loaded with in vitro bone tissue/osteoblasts derived from patient's marrow cells.^[23] These techniques using patient's cultured MSCs convert the nonosteogenic materials into osteogenic materials; i.e., the osteoconductive ceramics turn to osteoinductive ceramics. These new tissue engineering approaches, therefore, can be applied to various patients with massive skeletal defects to regenerate skeletal tissues.^[24]

CONCLUSIONS

We described the current status of calcium phosphate biomaterials produced in Japan. The calcium phosphate biomaterials detailed are HA, TCP, biphasic HA/TCP, CPC, the HA–collagen composite, the HA–collagen composite, and the HA–PLLA composite. They have been officially approved or are pending approval for clinical use. All the materials show biocompatible and osteoconductive nature with a variety of characteristics in porosity, pore size, mechanical strength, and ability to be resorbed. We also focused on tissue engineering approaches using mesenchymal stem cells/ceramic composites. These tissue engineering approaches can offer new cell-based therapies to various patients who need skeletal reconstruction.

ACKNOWLEDGMENTS

We wish to express our appreciation to Mr. Ogawa, Mr. Shoji, Dr. Hirano, Mr. Takeuchi, Mr. Asaoka (Pentax Co.), Dr. Hakamatsuka, Mr. Irie (Olympus

Optical Co., Ltd.), Mr. Okura, Mr. Hattori (NGK Spark Plug Co., Ltd.), Mr. Oomamiuda (Sumitomo Osaka Cement Co., Ltd.), Mr. Kashiwadani (Takiron Co., Ltd.) for their advice and recommendations in the preparation of this manuscript.

ARTICLES OF FURTHER INTEREST

Bioactive Glass; Biphasic Calcium Phosphate (BCP) Bioceramics: Chemical, Physical, and Biological Properties; Tissue Engineering of Bone; Zirconia Ceramics

REFERENCES

1. Tamai, N.; Myoui, A.; Tomita, T.; Nakase, T.; Tanaka, J.; Ochi, T.; Yoshikawa, H. Novel hydroxyapatite ceramics with an interconnective porous structure exhibit superior osteoconduction in vivo. *J. Biomed. Mater. Res.* **2002**, *59* (1), 110–117.
2. Toriyama, M.; Kawamura, S. Synthesis of β -tricalcium phosphate by use of wet milling (in Japanese). *Yogyo-Kyokai-Shi (J. Ceram. Soc. Jpn.)* **1986**, *94* (9), 1004–1008.
3. Monma, H.; Kanazawa, T. The hydration of α -tricalcium phosphate. *Yogyo-Kyokai-Shi (J. Ceram. Soc. Jpn.)* **1976**, *84* (4), 209–213.
4. Ohashi, H.; Kitanno, T.; Kakutani, Y.; Yamano, Y.; Morita, M.; Tanabe, H. mechanical properties of calcium phosphate bone paste (in Japanese). *Orthop. Surg. Traumatol.* **2002**, *45* (10), 959–967.
5. Arai, T.; Aoki, H. Treatment with calcium phosphate bone paste for distal radius fracture (in Japanese). *Orthop. Surg. Traumatol.* **2002**, *45* (10), 969–974.
6. Sogo, T.; Uchida, S.; Saito, S.; Ikeuchi, M.; Bando, K. Treatment with Biopex[®] (a calcium phosphate cement)

- for tibia plateau fracture (in Japanese). In *Transactions of the 3rd Symposium on Bioactive Paste*, The 3rd Symposium on Bioactive Paste, Tokyo, Nov 30, 2002; Organizing Committee of the 3rd Symposium on Bioactive Paste: Tokyo, 2003; 6-III-3, 116–118.
7. Sogo, T.; Uchida, S.; Saito, S.; Ikeuchi, M.; Bando, K. Treatment with Biopex[®] for calcaneus fracture (in Japanese). In *Transactions of the 3rd Symposium on Bioactive Paste*, The 3rd Symposium on Bioactive Paste, Tokyo, Nov 30, 2002; Organizing Committee of the 3rd Symposium on Bioactive Paste: Tokyo, 2003; 6-II-1, 95–97.
 8. Takemasa, R.; Yamamoto, H. Clinical application of calcium phosphate cement for repair of vertebral fractures in the osteoporotic spine (in Japanese). *Orthop. Surg. Traumatol.* **2002**, *45* (10), 989–1001.
 9. Hashimoto, S.; Kawashima, M.; Sakuragi, T.; Sato, K. Clinical results of augmentation of femoral neck fractures and intertrochanteric fractures with bioactive calcium-phosphate bone cement (in Japanese). *Orthop. Surg. Traumatol.* **2002**, *45* (10), 975–981.
 10. Moriwaki, T.; Tahara, Y.; Takayama, T.; Kodera, M.; Satomi, K.; Ishi, Y.; Bando, K. Clinical applications of calcium phosphate paste to revision surgery of hip arthroplasty (in Japanese). In *Transactions of the 3rd Symposium on Bioactive Paste*, The 3rd Symposium on Bioactive Paste, Tokyo, Nov 30, 2002; Organizing Committee of the 3rd Symposium on Bioactive Paste: Tokyo, 2003; 4–4, 75–77.
 11. Nakano, M.; Hirano, N.; Matsuura, K.; Watanabe, H.; Kitagawa, H.; Ishihara, H.; Kawaguchi, Y. Percutaneous transpedicular vertebroplasty with calcium phosphate cement in the treatment of osteoporotic vertebral compression and burst fractures. *J. Neurosurg. Spine* **2002**, *97* (3), 287–293.
 12. Hoshikawa, A.; Fukui, N.; Fukuda, A.; Sawamura, T.; Hattori, M.; Nakamura, K.; Oda, H. Quantitative analysis of the resorption and osteoconduction process of a calcium phosphate cement and its mechanical effect for screw fixation. *Biomaterials* **2003**, *24* (27), 4967–4975.
 13. Itoh, S.; Kikuchi, M.; Koyama, Y.; Matsumoto, H.N.; Takakuda, K.; Shinomiya, K.; Tanaka, J. Development of a novel biomaterial, hydroxyapatite/collagen (HAp/Col) composite for medical use. *Biomed. Mater. Eng.* **2005**, *15* (1–2), 29–41.
 14. Urist, M.R. Bone: formation by autoinduction. *Science* **1965**, *150* (698), 893–899.
 15. Noshi, T.; Yoshikawa, T.; Dohi, Y.; Ikeuchi, M.; Horiuchi, K.; Ichijima, K.; Sugimura, M.; Yonemasu, K.; Ohgushi, H. Recombinant human bone morphogenetic protein-2 potentiates the in vivo osteogenic ability of marrow/hydroxyapatite composites. *Artif. Organs* **2001**, *25* (3), 201–208.
 16. Takaoka, K.; Nakamura, H.; Yoshikawa, H.; Masuhara, K.; Tsuda, T.; Ono, K. Ectopic bone induction on and in porous hydroxyapatite combined with collagen and bone morphogenetic protein. *Clin. Orthop. Relat. Res.* **1988**, *234*, 250–254.
 17. Ohgushi, H.; Okumura, M.; Tamai, S.; Shors, E.C.; Caplan, A.I. Marrow cell induced osteogenesis in porous hydroxyapatite and tricalcium phosphate: a comparative histomorphometric study of ectopic bone formation. *J. Biomed. Mater. Res.* **1990**, *24* (12), 1563–1570.
 18. Ohgushi, H.; Dohi, Y.; Yoshikawa, T.; Tamai, S.; Tabata, S.; Okunaga, K.; Shibuya, T. Osteogenic differentiation of cultured marrow stromal stem cells on the surface of bioactive glass ceramics. *J. Biomed. Mater. Res.* **1996**, *32* (3), 341–348.
 19. Caplan, A.I. Mesenchymal stem cells. *J. Orthop. Res.* **1991**, *9* (5), 641–650.
 20. Maniopoulos, C.; Sodek, J.; Melcher, A.H. Bone formation in vitro by stromal cells obtained from marrow of young adult rats. *Cell. Tissue Res.* **1988**, *254* (2), 317–330.
 21. Yoshikawa, T.; Ohgushi, H.; Tamai, S.; Ichijima, K. Self setting hydroxyapatite cement as a carrier for bone-forming cells. *Biomed. Mater. Eng.* **1996**, *6* (5), 345–351.
 22. Yoshikawa, T.; Ohgushi, H.; Akahane, M.; Tamai, S.; Ichijima, K. Analysis of gene expression in osteogenic cultured marrow/hydroxyapatite construct implanted at ectopic sites: a comparison with the osteogenic ability of cancellous bone. *J. Biomed. Mater. Res.* **1998**, *41* (4), 568–573.
 23. Yoshikawa, T.; Ohgushi, H.; Uemura, T.; Nakajima, H.; Ichijima, K.; Tamai, S.; Tateishi, T. Human marrow cells-derived cultured bone in porous ceramics. *Biomed. Mater. Eng.* **1998**, *8* (5–6), 311–320.
 24. Ohgushi, H.; Caplan, A.I. Stem cell technology and bioceramics: from cell to gene engineering. *J. Biomed. Mater. Res.* **1999**, *48* (6), 913–927.



Cardiac Assist Devices

Keefe B. Manning

Department of Bioengineering, The Pennsylvania State University, University Park, Pennsylvania, U.S.A.

Conrad M. Zapanta

Department of Surgery, H151, Division of Artificial Organs, Penn State College of Medicine, Hershey, Pennsylvania, U.S.A.

John M. Tarbell

Department of Bioengineering, The Pennsylvania State University, University Park, Pennsylvania, and Department of Biomedical Engineering, The City College of New York, New York, New York, U.S.A.

INTRODUCTION

Cardiac assist devices are designed to assist or (in some cases) replace the natural heart. These devices can be used as short-term therapy, bridge to transplantation, and as permanent cardiac support. The different types of cardiac assist devices described in this article are the roller pump, intraaortic balloon pump, ventricular assist device (pulsatile and continuous flow), and total artificial heart. In-vitro and in-vivo methods that are used to test and evaluate cardiac assist devices will also be detailed. Finally, the complications associated with cardiac assist devices will be presented.

NATIVE HEART ANATOMY, FUNCTION, AND CLINICAL NEED

Anatomy

The anatomy of the native heart is shown in Fig. 1. The heart may be thought of as two pumps (right and left side) working together. The right pump delivers blood to the pulmonary circulation, while the left pump delivers blood to the systemic circulation. Each pump consists of an atrium and ventricle making up the heart's four distinct chambers: right atrium, right ventricle, left atrium, and left ventricle. The atria act as priming chambers for the ventricles. The ventricles pump blood out of the heart to either the pulmonary or systemic circulation. Heart valves located between each atrium and ventricle and at the outlet of each ventricle maintain flow direction during pulsatile flow.

Function

Blood from the systemic circulation (the body) enters the right atrium through the superior vena cava (from

the head and upper extremities) and inferior vena cava (from the trunk and lower extremities). The blood is then pumped to the right ventricle, which then pumps blood to the pulmonary circulation (the lungs) via the pulmonary arteries. Oxygenated blood returns to the heart from the lungs via the pulmonary vein to the left atrium and is then pumped to the left ventricle. The left ventricle pumps blood to the systemic circulation via the aorta.

Table 1 lists the nominal pressures and flows in the native heart.^[1] Any cardiac pumping system must generate these pressures and flows in order to meet the needs of recipients.

Clinical Need for Cardiac Assist Devices

Cardiovascular disease was the leading cause of death for almost 1.4 million people in the United States in 1999, representing over 40% of the reported deaths. Coronary heart disease (55%) represented the majority of these deaths, followed by stroke (17%), and congestive heart failure (6%). Almost 4.8 million Americans suffer from congestive heart failure, with about 550,000 new cases diagnosed each year. According to the Framingham Heart Study, 80% of men and 70% of women with congestive heart failure under the age of 65 will die within 8 years.^[2]

One treatment for congestive heart failure is heart transplantation. It is estimated that 40,000 Americans could benefit from a heart transplant each year. However, only 2300 donor hearts are available each year (1998–2001), with over 4000 patients on the waiting list during this period. From 1988 to 2002, approximately 640 patients died each year waiting for transplant.^[2,3] These numbers clearly demonstrate the need for blood pumps that either sustain the patient until transplant or permanently assist or replace the natural heart.

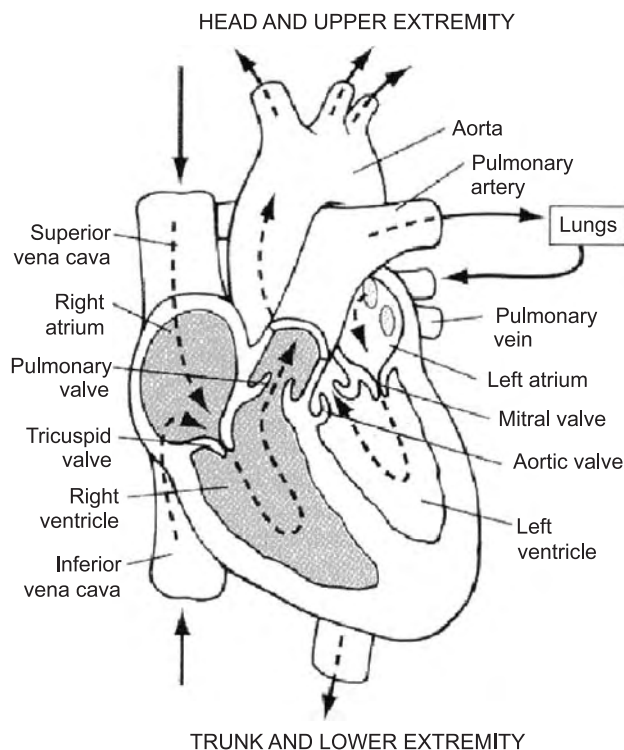


Fig. 1 Natural heart. (Permission pending.)

Different types of cardiac assist devices have been developed to aid the failing heart. These devices can be divided into roller pumps, balloon pumps, ventricular assist devices, and total artificial hearts. Roller pumps provide temporary support during or immediately following cardiac surgery. Balloon pumps aid the heart during cardiopulmonary bypass and cardiogenic shock. Ventricular assist devices assist the function of either the left or right ventricle. Total artificial hearts entirely replace the human heart.

Table 1 Nominal pressures and flows in the natural heart

Pressures:

1. Left ventricle
 - a. 120 mmHg peak systolic normal (into aorta)
 - b. 10 mmHg mean diastolic (from left atrium)
2. Right ventricle
 - a. 25 mmHg peak systolic (into pulmonary artery)
 - b. 5 mmHg mean diastolic (from right atrium)

Flows:

Normal healthy adult at rest: 5 liters/minute.
Maximum flow: 25 liters/minute.

ROLLER PUMPS

Roller pumps were the first generation of cardiac assist devices. These devices are designed to allow the natural heart to rest and recuperate after the surgical procedures have been completed. Roller pumps were first used in humans in 1953.^[4] These pumps are non-pulsatile and are used during cardiopulmonary bypass (CPB). They remain an integral part of most modern extracorporeal membrane oxygenators (ECMO), supporting patients for an average of 10 days. During open heart surgery, blood is routed extracorporeally (outside the body) through arterial and venous cannulation via elastic tubing to membranes that facilitate oxygen and carbon dioxide exchange with the blood.

Roller pumps used in CPB and ECMO have resulted in extremely high levels of damage to red blood cells (hemolysis) and other blood components with increasing damage with extended support time. Patients are then required to take high dosage levels of anticoagulants that may lead to excessive bleeding. Additional complications include blood cavitation due to tubing fatigue resulting in the tubing becoming less compliant.^[5] Because of these complications, roller pumps are typically limited to pediatric patients, and those adult patients who do not require large flow rates.^[5]

INTRAAORTIC BALLOON PUMPS

Intraaortic balloon pumps (IABP), illustrated in Fig. 2, are commonly used to assist in weaning patients from cardiopulmonary bypass or when patients experience cardiogenic shock secondary to an acute myocardial infarction.^[6] Initial success in human implants occurred in 1968.^[7] The IABP is inserted through the femoral or axillary artery and placed within the descending aorta. The patient's electrocardiogram (ECG) is used to trigger the IABP in order to adequately support the native heart. A polyethylene balloon inflates with gas (helium) via a percutaneous tube during diastole and deflates during systole. This counterpulsation reduces the amount of work that the heart has to perform by supplementing cardiac output by approximately 15%.^[8]

The complications associated with IABP are predominantly due to vascular injury that occurs during insertion into the aorta. Additional complications include excessive bleeding and limb ischemia.^[6] The IABP has been used for temporary support (mean of 53 hours),^[6] for short term ventricular assistance, and bridge to transplantation (up to 70 days).^[9] The IABP is currently being adapted for permanent ventricular assistance.^[10] The IABP has been used successfully in



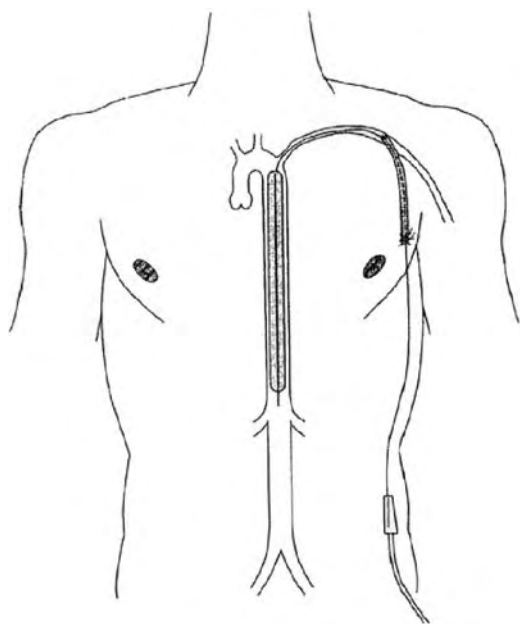


Fig. 2 Intraaortic balloon pump. (Permission pending.)

adult patients (regardless of size or age) and in pediatric patients.^[11]

VENTRICULAR ASSIST DEVICES

Ventricular assist devices (VADs) are designed to assist or replace the function of either the left or right ventricle. These devices are intended to provide temporary support until a donor heart has been located, myocardial recovery occurs, or as a permanent device. Ventricular assist devices can be classified as pulsatile or continuous flow devices.

Pulsatile Flow

Pulsatile flow ventricular assist devices are composed of a single ventricle that mimics the native ventricle. The ventricle is placed either within the abdominal cavity or outside the patient's body. There are two types of pulsatile flow ventricular assist devices: pneumatic and electric.

Pneumatic ventricular assist device

Figure 3 shows a pneumatic ventricular assist device that was originally developed by Penn State. The ventricle contains a flexible blood sac (of segmented polyurethane) that is housed within a rigid (polysulfone or titanium) case. The blood sac is extremely smooth to prevent the formation of clots (or thrombi). Prosthetic heart valves are placed in the inlet and outlet positions

of the ventricle to regulate flow direction. Air pulses that are generated by an external drive unit are used to compress the flexible blood sac periodically. An automatic control system varies the cardiac output by adjusting the heart rate and the time for ventricular filling in response to an increase in filling pressure.^[12]

The pump is placed extracorporeally on the patient's abdomen. The inlet cannula is inserted into the apex of the left ventricle and connected to the inlet port of the ventricular assist device. The outflow cannula is attached between the outflow port of the ventricular assist device and the ascending aorta, as an end-to-side anastomosis. Both the inflow and outflow cannulae pass through the skin below the rib cage. The first human implant of this type of assist device occurred in 1976. The Penn State pneumatic ventricular assist device was later purchased by Thoratec Corporation (Pleasanton, CA) and implanted in over 1700 patients worldwide. The pneumatic pump can also be used to assist the right ventricle by itself, or simultaneously with an additional pump that assists the left ventricle.^[13]

Other types of pneumatic devices include the ABIOMED (Danvers, MA) BVS (biventricular system) 5000 and the Heartmate (Thoratec) IP (intra-peritoneal) LVAS (left ventricular assist system). The ABIOMED BVS 5000 is an external dual-chamber device that can provide support to one or both ventricles. The chambers utilize polyurethane valves to regulate the flow direction.^[14] More than 3000 implants have been performed worldwide.^[15] The

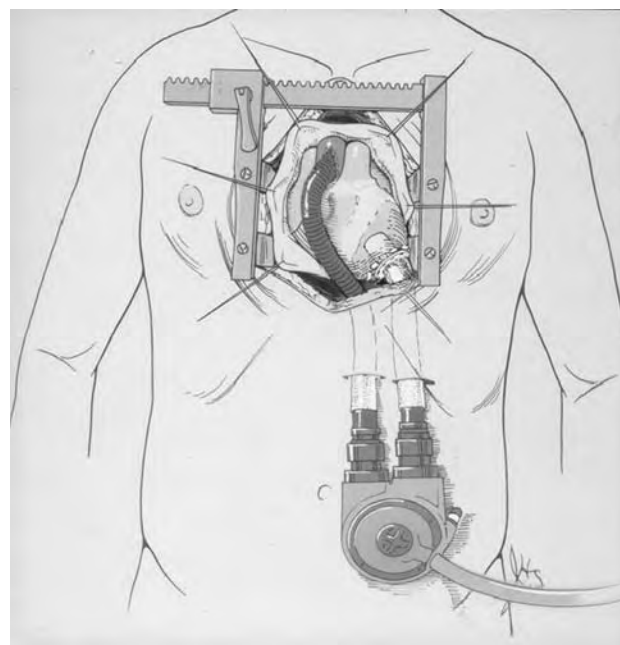


Fig. 3 Penn State/Thoratec pneumatic ventricular assist device. (Permission pending.)

Heartmate IP LVAS features an implantable blood pump that is connected to an external drive unit via a percutaneous air drive line. A unique feature of this device is the use of a textured blood surface in the ventricle that promotes the formation of a cell layer. This layer is believed to decrease thrombus formation because the layer mimics the blood contacting surface of a blood vessel. The Heartmate IP has been implanted in over 1200 patients worldwide.^[16]

Electric ventricular assist devices

Figure 4 illustrates an electric ventricular assist device developed by Penn State. Electric ventricular assist devices are typically placed within the abdominal cavity. The inlet cannula is inserted into the apex of the native left ventricle and connected to the inlet port of the device. The outlet cannula is attached between the outflow port of the device and the ascending aorta via an end-to-side anastomosis. The blood pump assembly utilizes a rollerscrew energy converter with a pusher plate. The motion of the pusher plate compresses the blood sac and ejects blood from the ventricle. The controller and internal battery supply are also implanted in the abdomen. This allows tether-free operation for approximately 30 minutes. Energy passes from the external power coil to the subcutaneous coil by inductive coupling via the transcutaneous energy transmission system (TETS). Air displaced by the blood pump housing enters the compliance chamber.

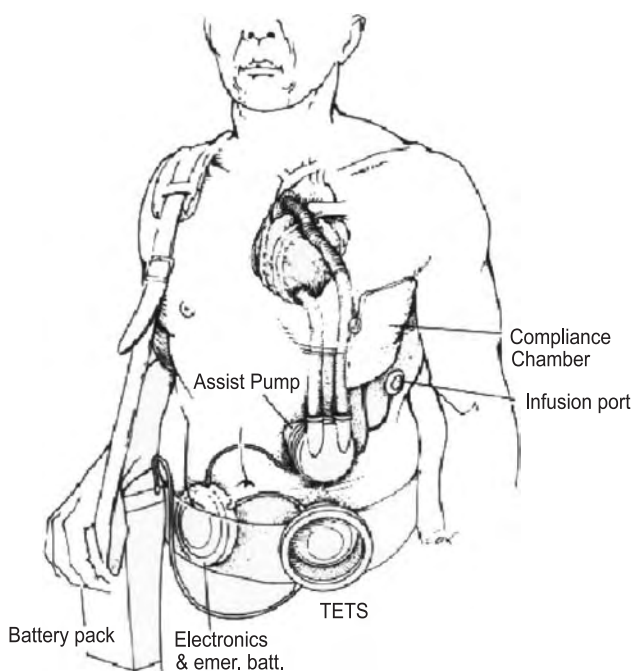


Fig. 4 Penn State/Arrow International electric ventricular assist device. (Permission pending.)

Air that diffuses out of the compliance chamber is periodically replenished by medical personnel via the subcutaneous access port. The Penn State system shown in Fig. 4 is being developed with Arrow International (Reading, PA) and is currently in clinical trials in the United States and Europe.^[17]

Other types of electric VADs include the Novacor (WorldHeart, Oakland, CA) LVAS and Heartmate VE (Vented Electric) LVAS. The Novacor LVAS contains a polyurethane blood sac that is compressed between dual pusher plates. A percutaneous power line connects the pump to an external battery pack and controller. The Novacor LVAS has been implanted in over 1350 patients worldwide.^[18] The Heartmate VE LVAS is similar to the Heartmate IP LVAS except for the use of an electric motor and pusher plate system. A percutaneous power line is also used to connect the pump to an external battery pack and controller. The Heartmate VE has been implanted in over 1600 patients worldwide.^[19]

A recent study, Randomized Evaluation of Mechanical Assistance for the Treatment of Congestive Heart Failure (REMATCH), examined the utility of ventricular assist devices. Patients with end-stage heart failure who were ineligible for cardiac transplantation were split into two groups. The first group received the Heartmate VE LVAS ($n=68$) while the second group received optimal medical management ($n=61$) consisting of angiotensin-converting-enzyme (ACE) inhibitors. The results showed a reduction of 48% in the risk of death from any cause in the LVAD group versus the medical-therapy group ($p=0.001$). The one-year survival was 52% for the VAD group and 25% for the medical-therapy group ($p=0.002$). The two-year survival was 23% for the VAD group and 8% for the medical-therapy group ($p=0.09$). Finally, the median survival was 408 days for the VAD group and 150 days for the medical-therapy group. This study clearly showed the clinical utility of ventricular assist devices.^[20]

Continuous Flow

Continuous flow ventricular assist devices deliver nonpulsatile flow and are classified as rotary pumps. There are two types of continuous flow ventricular assist devices: centrifugal or axial.

Centrifugal pumps

Figure 5 shows a centrifugal pump configuration. The inlet and outlet ports of the device are orthogonal to one another and are connected to the native ventricle and the aorta, respectively, in a similar fashion as described previously for electric ventricular assist devices. Blood entering the pump through the inlet port is





Fig. 5 Centrifugal flow ventricular assist device. (Permission pending.)

propelled towards the outer edge of the impeller due to centrifugal and lifting forces into a collecting volute. This volute is a channel carved into the housing where fluid is collected and stabilized before it exits the pump. The impeller can consist of blades, cones, or disks with the number and geometric configuration of the impeller being optimized for maximal efficiency. Typically, the impeller and housing are constructed of the same material (titanium or acrylic/plastic). Acrylic is used for some pump housings to provide optical access and examination of the impeller while in clinical use. Current centrifugal pumps serve as extracorporeal ventricular assist devices. However, newer generation pumps are being designed for implantable use with TETS.^[21]

Centrifugal pumps have several advantages over pulsatile ventricular assist devices. For example, centrifugal pumps require less energy than pusher plate designs because there is not an active load to move (piston or diaphragm). The impeller transfers its energy directly to the fluid to assist in flow. In addition, centrifugal pumps do not require valves for operation. Centrifugal pumps, however, are outlet pressure sensitive and may not produce flow should the outflow pressure be greater than the outlet pressure. In this case, the impeller will continue to spin without producing flow. To remedy this, either the outflow pressure must be reduced or the impeller speed must be increased, thus allowing the pump to operate regardless of the pressure at the inlet.

Typical motor speeds (or rotation rates) for centrifugal pumps range from 1500 to 5000 RPM (rotations per minute). This results in flow rates of 2 to 10 liters/minute. Although centrifugal pumps can be operated under steady or pulsatile flow conditions, there is debate regarding the superiority of either type

of flow.^[22] The long term effects of continuous flow on circulatory physiology and overall homeostasis remain unknown.

The Medtronic (Minneapolis, MN) Bio-Pump (Fig. 5) is the most commercially successful of all centrifugal blood pumps.^[23,24] This pump is placed extracorporeally and used during cardiac surgery for cardiopulmonary bypass, surgical recovery, and bridge to transplantation. The pump uses stacked acrylic cones as its impeller within a conical acrylic housing. The fluid enters the center of the cones, which creates swirling flow that is similar to a tornado. Flow is forced beyond the outer edge of the impeller and then exits at the bottom of the pump. A single magnetic drive unit is coupled with a magnet in the impeller. The Bio-Pump also comes in a pediatric version that can be used for assisted circulation. However, the Bio-Pump has limited use for long-term assistance because of excessive thrombus formation that has been mitigated through the use of heparin-coated circuits reducing the frequency for pump head exchange.^[24] Despite thrombus formation, the Bio-Pump has been intentionally designed so that the pump head (impeller) may be quickly replaced once thrombus has formed without detrimental effects on the patient. Clinical use of the Bio-Pump is said to be up to 1.5 million times. One clinical report from Baylor College of Medicine reported its use on 141 patients resulting in only 54% being weaned and only 22% being discharged from the hospital.^[25] An additional report from St. Louis University and Wake Forest University reported the use of the Bio-Pump on 52 patients resulting in only an 11% survival rate.^[26]

Other centrifugal pumps in use or under development include the Sarns pump (Terumo, Ann Arbor, MI),^[27] the Baylor College of Medicine's series of Gyro pumps,^[21] Ventracor's VentrAssist pump (Sydney, Australia),^[28] and MedQuest's HeartQuest pump (Salt Lake City, UT).^[29] Clinical postcardiotomy use of the Sarns pump in 91 patients resulted in 46% of them being weaned from the device but only a 21% survival rate.^[30]

Axial pumps

Figure 6 illustrates an axial flow pump, which is another type of continuous flow ventricular assist device. The flow through these pumps moves in the axial direction assisted by tangential forces and is accelerated through the device by a rotating impeller (rotor) that transfers energy from the impeller to the fluid. Stators (stationary flow straighteners) stabilize the blood as it enters and exits the rotor/impeller to prevent additional blood damage. Magnets are embedded within the rotor and are coupled with a rotating magnetic

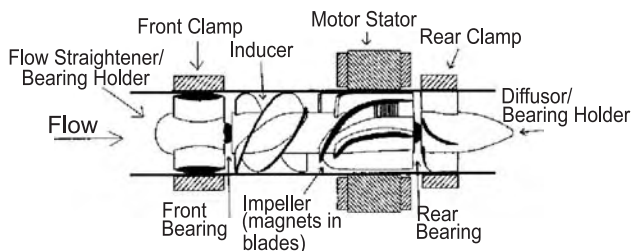


Fig. 6 Axial flow ventricular assist device. (Permission pending.)

field on the housing to act as a brushless DC motor. The pumps are constructed of titanium or a similar metal alloy that is biocompatible.

Axial pumps are either implanted directly within the left ventricle using a sewing cuff or are attached outside of the heart near the apex like most devices. Apical cannulae and outlet grafts have been evaluated to optimize each pump's efficiency. The outlet grafts are attached to the ascending or thoracic aorta.

Axial flow pumps run at speeds of 10,000–20,000 RPM, generating flow rates of up to 10 liters/minute. Such high motor speeds do not cause excessive hemolysis because of the short residence time of blood within the axial flow pump.^[31] Like centrifugal pumps, these pumps are also outlet pressure sensitive and may not produce flow should the outflow pressure exceed the outlet pressure.

Two types of axial flow pumps are currently in clinical trials as ventricular assist devices in the United States. They are the Jarvik2000 (New York, NY) and the Micromed DeBakey VAD (Houston, TX). The Jarvik2000 pump was developed through a collaborative effort between Jarvik Heart, Inc. and the Texas Heart Institute (Houston, TX). It employs a titanium stator and rotor that uses immersed bearings in the blood.^[32] The pump is placed within the left ventricle via a sewing cuff sutured to the ventricle, eliminating the inflow cannula. The outflow cannula is attached to the thoracic aorta as an end-to-side anastomosis. In-vivo studies with calves and humans illustrate the pump's effectiveness in assisting cardiac output and myocardial recovery.^[33,34] The initial 10 cases of human implantation yielded a 70% bridge-to-transplantation success rate.^[34] Pump speeds range from 8000 to 12,000 RPM with a maximum flow rate of 8 liters/min. In order to accommodate a wide range of patients, adult and pediatric versions have been developed.

The MicroMed DeBakey VAD (Fig. 6) is a collaborative effort between MicroMed Technology, Inc. and the Baylor College of Medicine. The rotor and stator were optimized through the extensive use of computational fluid dynamics prior to the manufacturing of the first prototype.^[31,35] Studies in calves

and humans have suggested great potential as an assist device.^[36] The DeBakey VAD operates from 7500 to 12,000 RPM and is attached to the left ventricle via an inflow cannula near the apex. The outflow tract attaches to the ascending aorta via an end-to-side anastomosis.^[37] Fifty-one patients had received the MicroMed DeBakey VAD as of September 2000 with an estimated probability of survival at 30 days to be 81%.^[38]

C

TOTAL ARTIFICIAL HEARTS

The total artificial heart (TAH) is designed to support both the pulmonary and systemic circulatory systems by replacing the natural heart. Two types of artificial hearts have been developed: pneumatic and electric.

Pneumatic Total Artificial Heart

A pneumatic total artificial heart is composed of two ventricles that replace the native left and right ventricle. Each ventricle is of similar design to the Penn State/Thoratec pneumatic ventricular assist device. Both ventricles are implanted within the chest. The air pulses are delivered to the ventricles via percutaneous drivelines. An automatic control system varies cardiac output by adjusting the heart rate and the time for ventricular filling in response to an increase in filling pressure.^[12]

Pneumatic artificial hearts can be used as a mechanical bridge to transplant (until a donor heart is implanted) and as a permanent replacement. The only pneumatic TAH approved as a bridge to transplant in the United States is the CardioWest/Jarvik (Tucson, AZ) Total Artificial Heart. Since 1993, the CardioWest has been implanted in over 100 patients. Sixty-seven patients survived to receive a cardiac transplant.^[39]

Pneumatic artificial hearts have also been used as a permanent replacement device. The Jarvik-7 pneumatic TAH was implanted in four patients.^[40] Although the longest survivor lived for 620 days, all four patients had hematologic, thromboembolic, and infectious complications. This type of artificial heart is no longer considered for permanent use because of infections associated with the percutaneous pneumatic drive lines and quality-of-life issues related to the bulky external pneumatic drive units.

Electric Total Artificial Heart

The electric TAH is completely implantable and is designed for permanent use. A schematic of the Penn State Electric TAH is shown in Fig. 7. The artificial heart is composed of two blood pumps that are of similar design to the Penn State electric ventricular

assist device. However, the electric TAH uses a single implantable energy converter that drives both ventricles asynchronously. The implantable controller adjusts the heart rate in response to ventricular filling and maintains left-right balance. The Penn State electric TAH has been successfully implanted in animals for over 1 year without thromboembolic complications.^[41]

The ABIOMED AbioCor TAH uses an electrohydraulic energy converter^[42] to compress the blood sacs. The AbioCor is currently undergoing clinical trials in the United States. Seven patients have been implanted. Two patients were discharged from the hospital, with one patient surviving for over one year. No significant device malfunctions were noted.^[43]

TESTING AND EVALUATION

Cardiac assist devices are subjected to a series of in-vitro and in-vivo tests prior to implantation in humans.

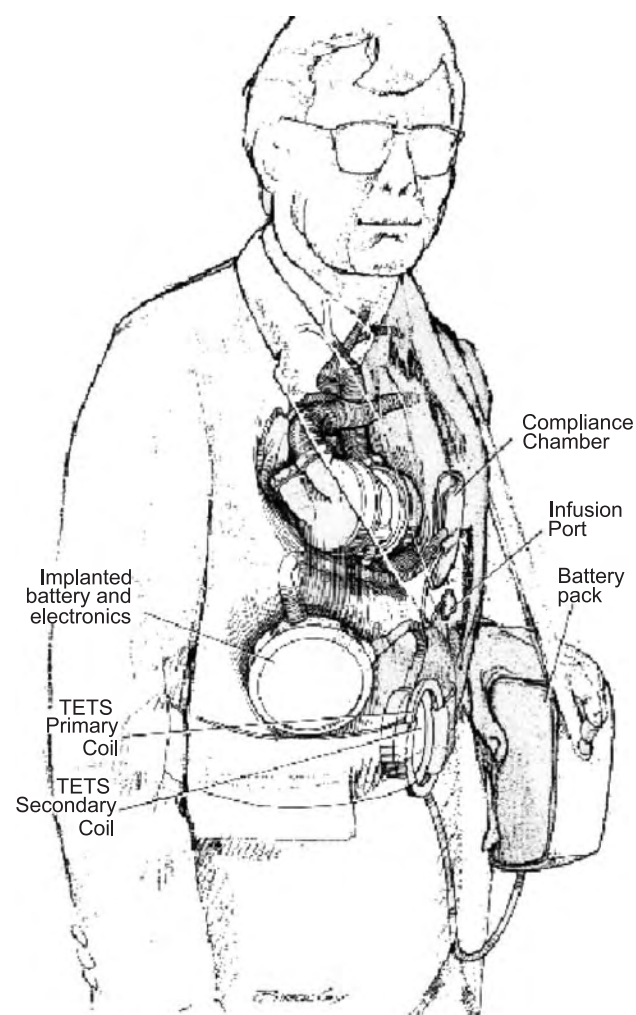


Fig. 7 Penn State electric total artificial heart. (Permission pending.)

These tests are designed to demonstrate the safety and efficacy of the devices.

In-Vitro Testing

In-vitro testing permits the evaluation of a cardiac assist device under controlled conditions that simulate in-vivo conditions. A mock circulatory loop is typically used to mimic the human circulatory system by replicating the systemic arterial capacitance and resistance.^[44] The fundamental components that compose a mock circulatory loop are blood pump, aortic compliance chamber, peripheral resistance element (constricted tubing), and venous reservoir. Tests are conducted to assess the performance characteristics (pressure-flow characteristics and power consumption) of the cardiac assist device under a variety of operating conditions simulating physiologic flows and pressures.^[45]

Flow visualization studies in transparent pump models are often conducted to elucidate any potential areas of risk (high shear or flow stasis) that could lead to hemolysis or thrombus formation. Laser Doppler velocimetry and particle image velocimetry systems are two noninvasive experimental techniques that can be used to evaluate flow fields within cardiac assist devices for fluid shear stress and turbulence levels.^[46,47] These techniques can also be used to assess the inlet and outlet flow regions quantitatively.^[48]

Durability testing (bearings, seals, diaphragms, heart valves, etc.) of the cardiac assist device is also performed to determine the lifetime and failure modes of the device through individual assessment in mock circulatory loops and accelerated wear testers.^[41,49] Hemolysis testing may also be performed to determine the degree of blood damage caused by the device relative to other systems.^[50]

In-Vivo Testing

In-vivo testing is employed to demonstrate the viability of the pump design in an animal model that simulates the human patient. Typical animal models include pigs, goats, dogs, and calves. The cardiac assist pump is implanted using surgical procedures and anatomical placements that will be used in human implants. After implantation, the animal is typically subjected to an anticoagulant therapy that includes heparin, coumadin, and/or aspirin. Plasma hemoglobin levels, erythrocyte, platelet, and leukocyte counts, and additional blood parameters are monitored to determine the existence and extent of blood trauma. After sacrifice, the cardiac assist device is explanted and carefully evaluated for thrombus deposition. Histopathological

studies are also conducted on the animals' organs (liver, kidney, lung, brain) to evaluate damage.^[51,52]

COMPLICATIONS

Although cardiac assist devices fulfill a major clinical need, many device-related complications exist. Bleeding has been reported in 79% of the patients with a Biomedicus centrifugal pump.^[51] Additional complications reported in this study included high levels of hemolysis (35%), renal failure (33%), thrombus formation in the pump (23%), and infection (15%). For patients with the Thoratec pneumatic LVAD, infection was the most frequent complication (46%), followed by bleeding (44%), thrombus formation in the device (15%), renal failure (14%), and hemolysis (9%). Results from the Registry Review for 1983–1993 indicated similar complication rates. Additional complications include the response to the implanted materials and device failure.^[53] These complications are also associated with other types of cardiac assist devices.^[5,43]

CONCLUSIONS

Current cardiac assist devices are able to provide viable cardiac support successfully by either assisting or replacing the native heart. Future designs will utilize novel battery technologies to permit longer time between battery changes, new biomaterials to enhance biocompatibility, improved control schemes that enhance the response of the cardiac assist device to meet physiologic demands, and smaller electronics to decrease the size of the pumps. These designs will also incorporate features that minimize blood trauma, reduce thromboembolic complications, and eliminate the need for anticoagulants. Future pumps will also be designed in a range of sizes to fit easily in a variety of patient populations including women and infants.

ACKNOWLEDGMENTS

The authors would like to acknowledge the support of William S. Pierce, M.D., Gus Rosenberg, Ph.D., David B. Geselowitz, Ph.D., and the past and present faculty and graduate students at the Division of Artificial Organs at the Penn State College of Medicine and the Department of Bioengineering at the Pennsylvania State University. The financial support from the National Institutes of Health and The Whitaker Foundation is also recognized.

ARTICLES OF FURTHER INTEREST

Artificial Heart Fluid Dynamics: Positive Displacement Pumps; Centrifugal Blood Pumps; Implanted Medical Devices, Power Sources For; Thrombosis; Tissue Engineering of Blood Vessel

REFERENCES

1. Guyton, A.C. *Textbook of Medical Physiology*; Saunders, W.B., Ed.; Company: Philadelphia, 2000.
2. American Heart Association. *2002 Heart and Stroke Statistical Update*; American Heart Association: Dallas, 2001.
3. UNOS 2002 *Annual Report of the U.S. Scientific Registry of Transplant Recipients and the Organ Procurement and Transplantation Network: Transplant Data 1989–2002. (2002, November 12)*; HHS/HRSA/OSP/DOT and UNOS: Rockville, MD, 2002.
4. Gibbon, J.H. Extracorporeal maintenance of cardiopulmonary functions. *Harvey Lect.* **1959**, *53*, 186–224.
5. Scherr, K.; Jensen, L.; Koshal, A. Mechanical circulatory support as a bridge to cardiac transplantation: Toward the 21st century. *Am. J. Crit. Care* **1999**, *8*, 324–339.
6. Ferguson, J.J., III; Cohen, M.; Freedman, R.J., Jr.; Stone, G.W.; Miller, M.F.; Joseph, D.L.; Ohman, E.M. The current practice of intra-aortic balloon counterpulsation: Results from the benchmark registry. *J. Am. Coll. Cardiol.* **2001**, *38*, 1456–1462.
7. Kantrowitz, A.; Tjonneland, S.; Freed, P.S.; Phillips, S.J.; Butner, A.N.; Sherman, J.L. Initial clinical experience with intraaortic balloon pumping in cardiogenic shock. *JAMA* **1968**, *203*, 113–118.
8. Nanas, J.N.; Mouloupoulos, S.D. Counterpulsation: Historical background, technical improvements, hemodynamic and metabolic effects. *Cardiology* **1994**, *84*, 156–167.
9. Cochran, R.P.; Starkey, T.D.; Panos, A.L.; Kunzelman, K.S. Ambulatory intraaortic balloon pump use as bridge to heart transplant. *Ann. Thorac. Surg.* **2002**, *74*, 746–752.
10. Jeevanandam, V.; Jayakar, D.; Anderson, A.S.; Martin, S.; Piccione, W., Jr.; Heroux, A.L.; Wynne, J.; Stephenson, L.W.; Hsu, J.; Freed, P.S.; Kantrowitz, A.D. Circulatory assistance with a permanent implantable IABP: Initial human experience. *Circulation* **2002**, *106* (suppl. 1), I183–I188.
11. Pinkney, K.A.; Minich, L.; Tani, L.Y.; Di Russo, G.B.; Veasy, G.; McGough, E.D.; Hawkins, J.A. Current results with intraaortic balloon pumping in infants and children. *Ann. Thorac. Surg.* **2002**, *73*, 887–891.
12. Richenbacher, W.E.; Pierce, W.S. Assisted Circulation and the Mechanical Heart. In *Heart Disease: A Textbook of Cardiovascular Medicine*; Braunwald, E., Ed.; W.B. Saunders Company: Philadelphia, 2001; Vol. 1, 534–547.
13. McBride, L.R.; Naumheim, K.S.; Fiore, A.C.; Moroney, D.A.; Swartz, M.T. Clinical experience with



- 111 Thoratec ventricular assist devices. *Ann. Thorac. Surg.* **1999**, *67*, 1233–1239.
14. Samuels, L.E.; Holmes, E.C.; Thomas, M.P.; Entwistle, J.C.; Morris, R.J.; Narula, J.; Wechsler, A.S. Management of acute cardiac failure with mechanical assist: Experience with the ABIOMED BVS 5000. *Ann. Thorac. Surg.* **2001**, *71* (suppl. 3), S67–S72.
 15. Lavee, J.; Paz, Y. Mechanical alternatives to the human heart: paracorporeal assist systems. *Isr. Med. Assoc. J.* **2002**, *4*, 125–130.
 16. Poirier, V.L. Worldwide experience with the TCI HeartMate system: Issues and future perspective. *Thorac. Cardiovasc. Surg.* **1999**, *47* (suppl. 2), 316–320.
 17. Mehta, S.M.; Pae, W.E., Jr.; Rosenberg, G.; Snyder, A.J.; Weiss, W.J.; Lewis, J.P.; Frank, D.J.; Thompson, J.J.; Pierce, W.S. The LionHeart LVD-2000: A completely implanted left ventricular assist device for chronic circulatory support. *Ann. Thorac. Surg.* **2001**, *71* (suppl. 3), S156–S161.
 18. Wheeldon, D.R.; LaForge, D.H.; Lee, J.; Jansen, P.G.; Jassawalla, J.S.; Portner, P.M. Novacor left ventricular assist system long-term performance: Comparison of clinical experience with demonstrated in vitro reliability. *ASAIO J.* **2002**, *48*, 546–551.
 19. Frazier, O.H.; Rose, E.A.; Oz, M.C.; Dembitsky, W.; McCarthy, P.; Radovancevic, B.; Poirier, V.L.; Dasse, K.A. Multicenter clinical evaluation of the HeartMate vented electric left ventricular assist system in patients awaiting heart transplantation. *J. Thorac. Cardiovasc. Surg.* **2001**, *122*, 1186–1195.
 20. Rose, E.A.; Gelijns, A.C.; Moskowitz, A.J.; Heitjan, D.F.; Stevenson, L.W.; Dembitsky, W.; Long, J.W.; Ascheim, D.D.; Tierney, A.R.; Levitan, R.G.; Watson, J.T.; Meier, P.; Ronan, N.S.; Shapiro, P.A.; Lazar, R.M.; Miller, L.W.; Gupta, L.; Frazier, O.H.; Desvigne-Nickens, P.; Oz, M.C.; Poirier, V.L. Randomized evaluation of mechanical assistance for the treatment of congestive heart failure (REMATCH) study group. Long-term mechanical left ventricular assistance for end-stage heart failure. *N. Engl. J. Med.* **2001**, *345*, 1435–1443.
 21. Nose, Y.; Yoshikawa, M.; Murabayashi, S.; Takano, T. Development of rotary blood pump technology: Past, present, and future. *Artif. Organs* **2000**, *24*, 412–420.
 22. Nose, Y.; Kawahito, K.; Nakazawa, T. Can we develop a nonpulsatile permanent rotary blood pump? Yes, we can. *Artif. Organs* **1996**, *20*, 467–474.
 23. Lynch, M.F.; Peterson, D.; Baker, V. Centrifugal blood pumping for open heart surgery. *Minn. Med.* **1978**, *61*, 536–538.
 24. Noon, G.P.; Lafuente, J.A.; Irwin, S. Acute and temporary ventricular support with BioMedicus centrifugal pump. *Ann. Thorac. Surg.* **1999**, *68*, 650–654.
 25. Noon, G.P.; Lafuente, J.A.; Irwin, S. Acute and temporary ventricular support with BioMedicus centrifugal pump. *Ann. Thorac. Surg.* **1999**, *68*, S133–S138.
 26. Pennington, D.G.; Swartz, M.T.; Lohmann, D.P.; McBride, L.R. Cardiac assist devices. *Surg. Clin. North Am.* **1998**, *78*, 691–704.
 27. Curtis, J.J.; Walls, J.T.; Demmy, T.L.; Boley, T.M.; Schmaltz, R.A.; Goss, C.F.; Wagner-Mann, C.C. Clinical experience with the Sarns centrifugal pump. *Artif. Organs* **1993**, *17*, 630–633.
 28. Tansley, G.; Vidakovic, S.; Reizes, J. Fluid dynamic characteristics of the VentrAssist rotary blood pump. *Artif. Organs* **2000**, *24*, 483–487.
 29. Allaire, P.E.; Wood, H.G.; Awad, R.S.; Olsen, D.B. Blood flow in a continuous flow ventricular assist device. *Artif. Organs* **1999**, *23*, 769–773.
 30. Curtis, J.J.; Wagner-Mann, C. Extracorporeal Support: Centrifugal Pumps. In *Cardiac Assist Devices*; Goldstein, D.J., Oz, M.C., Eds.; Futura Publishing Company: Armonk, NY, 2000; 215–233.
 31. Mizuguchi, K.; Damm, G.; Benkowsky, R.; Aber, G.; Bacak, J.; Svikovsky, P.; Glueck, J.; Takatani, S.; Nose, Y.; Noon, G.P.; DeBakey, M.E. Development of an axial flow ventricular assist device: In vitro and in vivo evaluation. *Artif. Organs* **1995**, *19*, 653–659.
 32. Jarvik, R.K. System considerations favoring rotary artificial hearts with blood immersed bearings. *Artif. Organs* **1995**, *19*, 565–570.
 33. Macris, M.P.; Parnis, S.M.; Frazier, O.H.; Fuqua, J.M., Jr.; Jarvik, R.K. Development of an implantable ventricular assist system. *Ann. Thorac. Surg.* **1997**, *63*, 367–370.
 34. Frazier, O.H.; Myers, T.J.; Gregoric, I.D.; Khan, T.; Delgado, R.; Croitoru, M.; Miller, K.; Jarvik, R.; Westaby, S. Initial clinical experience with the Jarvik 2000 implantable axial-flow flow ventricular assist system. *Circulation* **2002**, *105*, 2855–2860.
 35. Wernicke, J.T.; Meier, D.; Mizuguchi, K.; Damm, G.; Aber, G.; Benkowski, R.; Nose, Y.; Noon, G.P.; DeBakey, M.E. A fluid dynamic analysis using flow visualization of the Baylor/NASA implantable axial flow pump for design improvement. *Artif. Organs* **1995**, *19*, 161–177.
 36. Wieselthaler, G.M.; Schima, H.; Hiesmayr, M.; Pacher, R.; Laufer, G.; Noon, G.P.; DeBakey, M.E.; Wolner, E. First clinical experience with the DeBakey VAD continuous-axial-flow pump for bridge to transplantation. *Circulation* **2000**, *101*, 356–359.
 37. Noon, G.P.; Morley, D.; Irwin, S.; Benkowski, R. Development and clinical application of the MicroMed DeBakey VAD. *Curr. Opin. Cardiol.* **2000**, *15*, 166–171.
 38. Noon, G.P.; Morley, D.N.; Irwin, S.; Abdelsayed, S.V.; Benkowski, R.J.; Lynch, B.E. Clinical experience with the MicroMed DeBakey ventricular assist device. *Ann. Thorac. Surg.* **2001**, *71* (suppl. 3), S133–S138.
 39. Arabia, F.A.; Copeland, J.G.; Smith, R.G.; Banchy, M.; Foy, B.; Kormos, R.; Tector, A.; Long, J.; Dembitsky, W.; Carrier, M.; Keon, W.; Pavie, A.; Duveau, D. CardioWest total artificial heart: A retrospective controlled study. *Artif. Organs* **1999**, *23*, 204–207.
 40. DeVries, W.C. The permanent artificial heart. four case reports. *JAMA* **1988**, *259*, 849–859.
 41. Weiss, W.J.; Rosenberg, G.; Snyder, A.J.; Pierce, W.S.; Pae, W.E.; Kuroda, H.; Rawhouser, M.A.; Felder, G.; Reibson, J.D.; Cleary, T.J.; Ford, S.K.; Marlotte, J.A.; Nazarian, R.A.; Hicks, D.L. Steady state hemodynamic and energetic characterization of the Penn State/3M

- health care total artificial heart. *ASAIO J.* **1999**, *45*, 189–193.
42. Kung, R.T.; Yu, L.S.; Ochs, B.D.; Parnis, S.M.; Macris, M.P.; Frazier, O.H. Progress in the development of the ABIOMED total artificial heart. *ASAIO J.* **1995**, *41*, M245–M248.
43. Dowling, R.D.; Gray, L.A.; Etoch, S.W.; Laks, H.; Marelli, D.; Samuels, L.; Entwistle, J.; Couper, G.; Vlahakes, G.J.; Frazier, O.H. The AbioCor implantable replacement heart. *Ann. Thorac. Surg.* **2003**, *75*, S93–S99.
44. Rosenberg, G.; Phillips, W.M.; Landis, D.L.; Pierce, W.S. Design and evaluation of the Pennsylvania State University mock circulatory system. *ASAIO J.* **1981**, *4*, 41–49.
45. Rosenberg, G. Artificial Heart and Circulatory Assist Devices. In *The Biomedical Engineering Handbook*; Bronzino, J.D., Ed.; CRC Press, Inc.: Boca Raton, FL, 1995; 1839–1870.
46. Baldwin, J.T.; Deutsch, S.; Geselowitz, D.B.; Tarbell, J.M. LDA measurements of mean velocity and reynolds stress fields within an artificial heart ventricle. *J. Biomech. Eng.* **1994**, *116*, 190–200.
47. Hochareon, P.; Manning, K.B.; Fontaine, A.A.; Deutsch, S.; Tarbell, J.M. In *Development of High Resolution Particle Image Velocimetry for Use in Artificial Heart Research*, Second Joint Meeting of the IEEE-EMBS and BMES, Houston, TX, Oct. 2002.
48. Manning, K.B.; Miller, G.E. Flow through an outlet cannula of a rotary ventricular assist device. *Artif. Organs* **2002**, *26*, 714–723.
49. Lee, J.; Miller, P.J.; Chen, H.; Conley, M.G.; Carpenter, J.L.; Wihera, J.C.; Jassawalla, J.S.; Portner, P.M. Reliability model from the in vitro durability tests of a left ventricular assist system. *ASAIO J.* **1999**, *45*, 595–601.
50. Ducko, C.T.; McGregor, M.L.; Rosenberg, G.; Pierce, W.S. The effect of valve type and drive line dP/dt on hemolysis in the pneumatic ventricular assist device. *Artif. Organs* **1994**, *8*, 454–460.
51. Wampler, R.K.; Moise, J.C.; Frazier, O.H.; Olsen, D.B. In vivo evaluation of a peripheral vascular access axial flow blood pump. *ASAIO Trans.* **1988**, *34*, 450–454.
52. Weiss, W.J.; Rosenberg, G.; Snyder, A.J.; Donachy, J.; Reibson, J.; Kawaguchi, O.; Sapirstein, J.S.; Pae, W.E.; Pierce, W.S. A completely implanted left ventricular assist device: Chronic in vivo testing. *ASAIO J.* **1993**, *39*, M427–M432.
53. Mehta, S.M.; Aufiero, T.X.; Pae, W.E., Jr.; Miller, C.A.; Pierce, W.S. Results of mechanical ventricular assistance for the treatment of post cardiectomy cardiogenic shock. *ASAIO J.* **1996**, *42*, 211–218.



Cardiac Bioelectricity

Clive M. Baumgarten

*Department of Physiology, Biomedical Engineering, and Internal Medicine (Cardiology),
Virginia Commonwealth University, Richmond, Virginia, U.S.A.*

INTRODUCTION

An electrical event—the action potential—is the signal that initiates cardiac contraction and the pumping of blood. Action potentials reflect the intrinsic bioelectricity of the heart. Electrical activity of individual cardiac myocytes (muscle cells) arises both from the characteristics of the ion channels found in the plasma membrane (sarcolemma) and the ionic gradients that exist between the cytoplasm and extracellular spaces. Individual myocytes are electrically coupled by low-resistance gap junctions, located at the intercalated disk at the end of cells and at lateral points of contact. This coupling creates an electrical syncytium and permits the electrical impulse to spread to every cell in the heart. Normal function requires a precise coordination of cardiac bioelectric activity, and the action potential propagates over the heart in a specific sequence and with specific timing. The precise coordination of cardiac electrical activity arises from both the electrical anatomy of the heart and the specialized characteristics of myocytes in each region. The distribution of ion channels and, to a lesser extent, ionic gradients is not uniform. Regional differences exquisitely tune the electrophysiological properties of myocytes for their functional requirements.

This article reviews several facets of cardiac bioelectricity, including the electrical anatomy of the heart, the ionic basis for the cardiac action potential, pacemaker activity, and excitability. These concepts form the physiological basis for understanding the electrocardiogram, stimulation of the heart by pacemakers, and defibrillation, but clinical aspects of electrophysiological function will not be covered.

CARDIAC ELECTRICAL ANATOMY

The cardiac action potential propagates to every myocyte in the heart in a specific sequence and with specific timing. This is called normal sinus rhythm (NSR). NSR requires excitation to originate at the sino-atrial node (SAN) at a regular rate of 60–100 min⁻¹ and to follow the sequence and timing described in the following section.^[1] The resulting coordination of electrical activity reflects, in part, both

the presence of myocytes with specialized bioelectric properties and the electrical anatomy of the heart, which is illustrated in Fig. 1.

Sino-Atrial Node

The SAN is the physiologic pacemaker, the site where the heartbeat originates.^[1] It is approximately 8 mm long and approximately 2 mm thick in humans and is located at the roof of the right atrium (RA) along the posterior margin of the superior vena cava, where it joins the RA free wall at the crista terminalis.^[2,3] SAN myocytes spontaneously and repetitively give rise to action potentials that propagate to the rest of the heart, a property called normal automaticity. Cardiac electrical activity does not require neural input. The SAN beats spontaneously when excised from the heart, even after blocking both the sympathetic and parasympathetic neurons that richly innervate this structure.

Autonomic neurons modulate the frequency of SAN firing and thereby modulate heart rate.^[1] Sympathetic neurons innervating the SAN arise in the superior cervical ganglion and release norepinephrine, which binds to β_1 adrenoreceptors on the myocytes and accelerates SAN automaticity. Parasympathetic neurons reach the heart via the vagus nerve and release acetylcholine (ACh), which binds to muscarinic (M₂) receptors and slows the heart. At rest, the heart is under parasympathetic control, meaning heart rate accelerates if both sympathetic and parasympathetic influences are eliminated. The two autonomic branches are regulated in a coordinated but antagonistic manner by the vasomotor center in the medulla. To increase heart rate, for example, the vasomotor center depresses the firing of parasympathetics (parasympathetic withdrawal) while simultaneously augmenting sympathetic activity. Slowing of the heart is accomplished by the converse process. These actions are termed positive and negative chronotropic effects, respectively.

Atria

Because SAN is located within the RA, RA becomes excited before the left atrium (LA). The excitation

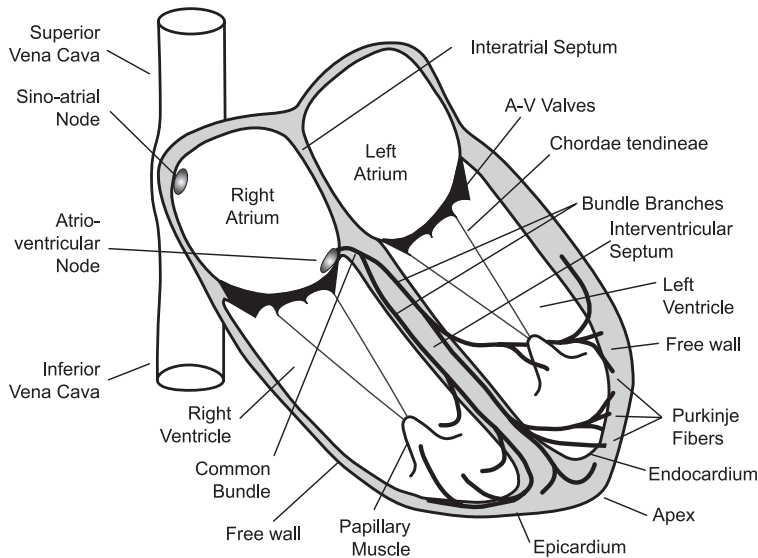


Fig. 1 Diagram of the electrical anatomy of the heart. Tissue specialized for electrical properties includes: sino-atrial node, at the roof of right atrium; atrio-ventricular node, at floor of right atrium adjacent to interatrial septum; and the His-Purkinje system, comprising the common bundle, bundle branches, and Purkinje fiber network.

wavefront conducts in all directions in atria with a conduction velocity of approximately $0.5\text{--}1\text{ m}\cdot\text{s}^{-1}$.^[1] Preferential conduction pathways—the atrial internodal tracts—have been noted. Bachmann's bundle (anterior tract) directs excitation from RA to LA, and there is evidence that its interruption slightly delays LA excitation. Wenkebach's and Thorel's bundles (middle and posterior tracts) direct the impulse from SAN to the atrio-ventricular node (AVN). These tracts are not of major functional significance, however. Preferential conduction largely reflects local fiber orientation and variations in atrial wall thickness (i.e., the pectinate muscles).

Atrio-Ventricular Node

Atria and ventricles are physically separated by connective tissue that serves to anchor the valves and electrically isolate the upper and lower chambers. Normally, the only muscular connection is the AVN, and to pass from atria to ventricles the electrical impulse must propagate through this 5 mm-long structure.^[1–3] The AVN is adjacent to the coronary sinus at the RA floor and posterior along the atrial septum (Fig. 1). AVN conduction velocity is extremely slow ($0.02\text{--}0.05\text{ m}\cdot\text{s}^{-1}$), which gives rise to the AV delay. AV delay allows time for completion of atrial contraction and ventricular filling before the initiation of ventricular contraction.

Classically, AVN is divided into AN (atrio-nodal), N (nodal), and NH (nodal-His) regions based on anatomic and electrophysiologic criteria.^[1,2] The AN and NH regions contain transitional cells. Conduction velocity is slowest in the central N region, although the majority of AV delay occurs in the AN region because

of its greater length. A more recent approach better reflecting the complex 3-dimensional AVN anatomy divides AVN into transitional cells, compact node, posterior node extensions, and the lower nodal cell bundle.

The AVN has two additional functions. First, it acts as an electrical filter. Excitability recovers slowly after the action potential in AVN, and therefore the AVN cannot follow accelerated atrial rates that occur in certain dysrhythmias, such as atrial tachycardias and atrial fibrillation. If the atrial rate is too fast, an impulse may reach AVN during the effective refractory period (ERP), when the myocytes are inexcitable.^[1,3] This can give rise to 2:1 or 3:1 AV block, where only every second or third impulse conducts through. The basis for the ERP will be discussed later. The second function of the AVN is to serve as a backup pacemaker. AVN exhibits normal automaticity, although its spontaneous rate (typically $40\text{--}50\text{ min}^{-1}$) is slower than that of SAN. Therefore, the AVN is said to be a latent pacemaker, and its automaticity is not normally expressed in the intact heart.

Occasionally, an accessory pathway (bundle of Kent) directly links the atria and ventricle.^[1,2] When functional, bypass tracts allow rapid impulse conduction from atria to ventricle and premature ventricular excitation known as Wolff–Parkinson–White syndrome, or preexcitation. Additional accessory pathways—bypass fibers of James that connect RA to the distal end of AVN and Mahaim fibers that link AVN to ventricle—sometimes are noted. Such pathways can serve as the anatomical substrate for reentrant arrhythmias and preclude normal AV delay. Recent evidence indicates that the posterior extension of the AVN also may serve as a slowly conducting pathway in reentrant arrhythmias.

AVN is highly innervated by both sympathetic and parasympathetic fibers, and each aspect of AVN function is regulated by autonomic neurons.^[1,3] Sympathetic stimulation speeds conduction velocity (positive dromotropic effect); quickens recovery of excitability (decreased ERP), which allows AVN to pass impulses at a higher frequency; and augments its spontaneous firing rate (normal automaticity). On the other hand, parasympathetic (vagal) stimulation slows conduction velocity (negative dromotropic effect), delays recovery of excitability (increased ERP), and suppresses AVN automaticity.

His-Purkinje System

The location of the aortic and pulmonic valves at the base of the heart (plane dividing atria and ventricles) presents an engineering dilemma. The electrical signal to contract reaches the ventricle via AVN at its base, whereas efficient ejection of blood requires excitation from apex to base, creating a wave of contraction that propels blood toward the valves. This dilemma is solved by the His-Purkinje system, also called the ventricular specialized conducting system. Conduction velocity in the His-Purkinje system ($2\text{--}4\text{ m}\cdot\text{s}^{-1}$) is the fastest in the heart, and it rapidly distributes the impulse to the ventricles.^[1,3] The His (common) bundle runs in the ventricular septum and subdivides into right and left bundle branches, anterior (superior) and posterior (inferior) fascicles of the left bundle, and the Purkinje fiber network, which terminates under the endocardial surface near the apex and along the free walls of the ventricles (Fig. 1).^[1,2] The detailed anatomical pattern of Purkinje network branching is highly variable. Portions of the network are free-running,

meaning they cross part of the ventricular cavity rather than adhering to the walls of the chamber. The His-Purkinje system also expresses automaticity and serves as a latent pacemaker. Intrinsic rate decreases from 40 to 20 s^{-1} from the AVN junction to terminal Purkinje fibers.

Ventricle

Within the ventricular wall, the action potential spreads in all directions with a conduction velocity of $0.5\text{--}1\text{ m}\cdot\text{s}^{-1}$,^[1] depending on fiber orientation (conduction is fastest parallel to the myocyte's long axis and slowest transversely). The first region to be excited is the ventricular septum adjacent to the left ventricular chamber. From this point, depolarization spreads across the septum toward the right ventricle and down the septum toward the apex. Then, distributed via the Purkinje network, excitation begins at multiple subendocardial sites near the apex. The excitation coalesces and spreads over the free walls from endocardium to epicardium and from apex to base.

CARDIAC ACTION POTENTIAL

Action potentials from various regions and the phases of the action potential are shown in Fig. 2. Recordings can be made either by impaling a cell in a multicellular preparation with a 3 M KCl -filled microelectrode or from isolated single myocytes using a patch electrode filled with a physiological intracellular solution. By convention, the potential difference across the membrane (membrane potential, E_m) is taken with the outside as ground.

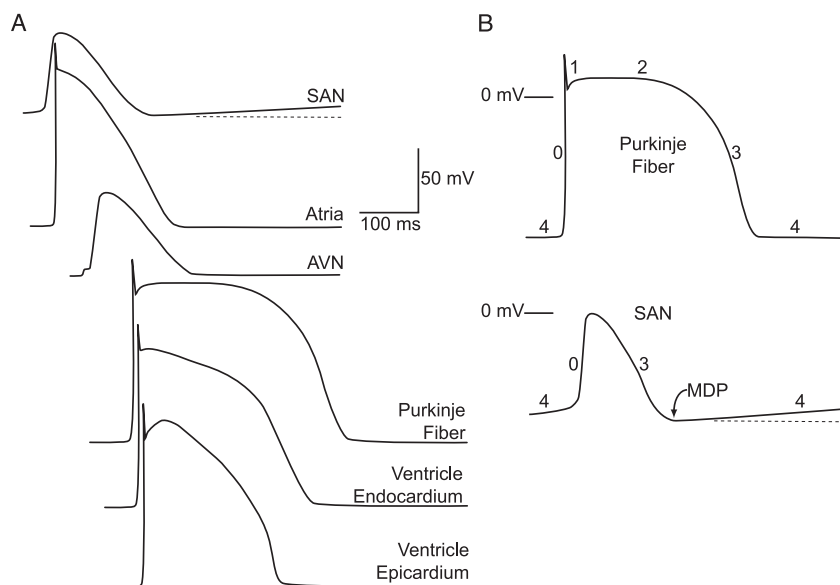


Fig. 2 (A) Action potentials from regions of the heart. Upstrokes are offset, representing the normal timing of activation. SAN undergoes diastolic depolarization (dashed line at MDP). The step in AVN upstroke is due to asynchronous arrival of action potentials via different AVN pathways. (B) Action potential phases in Purkinje fiber and SAN: 0, upstroke; 1, initial repolarization; 2, plateau; 3, repolarization; and 4, electrical diastole (or diastole). SAN action potentials do not have a distinct phase 2 (plateau) or phase 1 (initial repolarization), and they undergo phase 4 depolarization. Time and voltage calibrations apply to both panels.

Action potentials have distinct features in each region (Fig. 2A).^[1,4,5] Phase 0, the upstroke, is rapid in atrial, ventricular, and His-Purkinje myocytes (maximum $dV/dt = 150\text{--}600 \text{ V} \cdot \text{s}^{-1}$), but is much slower ($dV/dt = 2\text{--}10 \text{ V} \cdot \text{s}^{-1}$) in SAN and AVN. Action potential duration (APD) ranges from nearly 400 ms in the terminal Purkinje fibers to about 150 ms in SAN and AVN. The sequence of APD is:

terminal Purkinje fibers > His-Purkinje system
 > ventricle (midmyocardium > endocardium
 > epicardium) > atria \approx SAN \approx AVN

APD also is inversely related to heart rate.^[1,4,5] This effect is prominent in ventricle, which explains the heart-rate dependence of the Q-T interval in the electrocardiogram, as well as in the His-Purkinje system. Ventricular and Purkinje action potentials have a distinct plateau (phase 2), whereas atrial action potentials are more triangular, and phases 2 and 3 (repolarization) blend together in SAN and AVN. Finally, phase 4 (diastolic) depolarization is responsible for normal automaticity observed in SAN, whereas E_m during phase 4 (electrical diastole or diastole) typically is constant in other regions. Nevertheless, the AVN and the His-Purkinje system exhibit phase 4 depolarization under certain conditions.

Ionic Basis for Cardiac Action Potential

Ultimately, the electrical activity in the heart reflects the time- and voltage-dependent behavior of ion channels and the transmembrane ionic gradients. The heart is particularly rich in ion channels; Na^+ channels and several types of K^+ , Ca^{2+} , and Cl^- channels play key roles.^[4-6] The main types of cardiac ion channels and their localization and molecular correlates are summarized in Table 1. To understand the basis for the action potential and its distinct characteristics across the heart, we will first review the equations for ionic current and E_m , then consider a prototypic ventricular action potential, and finally discuss how heterogeneous ion channel expression is responsible for distinct functional characteristics.

Equations for ionic current and E_m

Ionic current (I) crossing the membrane is determined by conductance (g) and driving force, the difference between E_m and the Nernst equilibrium potential for the ion carrying the current. For example, we

write for K^+ current

$$I_K = g_K(E_m - E_K)$$

where the K^+ equilibrium potential, E_K , is given by

$$E_K = -RT/zF \ln[K^+]_i/[K^+]_o$$

where R is the gas constant, T is temperature (K), z is the valence, F is Faraday's constant, and the bracketed terms refer to intracellular, i , and extracellular, o , concentrations, respectively. For a monovalent ion such as K^+ , at 37°C this is equivalent to

$$E_K = -61 \log[K^+]_i/[K^+]_o$$

Measurements of intracellular ion concentrations in various parts of the heart and the resulting equilibrium potentials are reviewed elsewhere.^[7] By convention, an outward current is taken as cation efflux from the cell (or the electrical equivalent, anion influx) and will result in repolarization or hyperpolarization of E_m , whereas inward current is cation influx (or anion efflux), which causes depolarization. In terms of the sign of I , an outward current is positive and an inward current is negative, regardless of whether a cation or an anion is carrying current.

When E_m is constant ($dV/dt = 0$), the sum of the ionic currents, I_{ionic} , must be 0, and E_m is given by the chord conductance equation. For a membrane with K^+ , Na^+ , Ca^{2+} , and Cl^- channels,

$$I_{\text{ionic}} = g_K(E_m - E_K) + g_{\text{Na}}(E_m - E_{\text{Na}}) \\ + g_{\text{Ca}}(E_m - E_{\text{Ca}}) + g_{\text{Cl}}(E_m - E_{\text{Cl}}) = 0$$

and solving for E_m gives

$$E_m = \frac{g_K E_K + g_{\text{Na}} E_{\text{Na}} + g_{\text{Ca}} E_{\text{Ca}} + g_{\text{Cl}} E_{\text{Cl}}}{g_K + g_{\text{Na}} + g_{\text{Ca}} + g_{\text{Cl}}}$$

where the subscript j in g_j and E_j denotes the species, and the chord conductance for each species is taken as the chord of the current-voltage relationship between E_m and the equilibrium potential, E_j . (Note that the chord conductance, $\Delta I/(E_m - E_j)$, is different from the slope conductance, which is equal to $\delta I/\delta E$ at a selected voltage.) E_m is constant at the point at which inward and outward currents balance, and E_m approaches the equilibrium potential for the ion that has the greatest conductance. Nevertheless, the chord conductance equation as written is a simplification. It lumps all of the conductance for a particular ion into a single term, even though multiple distinct channel types may contribute, and it ignores the small contributions of electrogenic pumps and exchangers that produce a



Table 1 Major ionic currents in heart

Current	Molecular basis		Role in action potential	Localization of current
	α subunit	β subunit		
<i>Inward cation</i>				
I_{Na}	$Na_v1.5$ (SCN5A)	β_1 (SCN1B) β_2 (SCN2B)	Upstroke in A, V, PF	PF > A ~ V
I_{Ca-L}	α_{1C} ; $Ca_v1.2$ (CACNA1C; CACNL1A1)	β_1 (CACNB1) β_2 (CACNB2) $\alpha_2\delta$ (CACNA2D1)	Plateau in A, V, PF; upstroke and automaticity in SAN, AVN	SAN ~ AVN
I_{Ca-T}	α_{1H} ; $Ca_v3.1$ (CACNA1H)		Upstroke in SAN; automaticity in SAN?	SAN \gg A \gg AVN, V, PF
I_f	(HCN4, HCN1, HCN2)	MiRP1 (KCNE2)	Phase 4 (diastolic) depolarization	SAN: HCN4 \gg HCN1 \gg HCN2 PF: HCN4 ~ HCN1 > HCN2 A, V: HCN2 \gg HCN4
<i>Outward K^+</i>				
I_{K1}	$K_{ir}2.1$ (KCNJ2) $K_{ir}2.2$ (KCNJ12)		Resting potential; terminal phase 3 repolarization	V > A \gg SAN ~ AVN
I_{to}	$K_v4.3$ (KCND3) $K_v1.4$ (KCNA4)	KChIP2 (KCNIP2) MiRP1 (KCNE2)	Phase 1 repolarization	A, PF, V (mid, epi > endo) $K_v4.3$ (man) $K_v1.4$ (rabbit, V endo in man)
I_{Kr}	hERG (KCNH2)	minK (KCNE1) MiRP1 (KCNE2)	Termination of plateau, phase 3 repolarization	Ubiquitous, LA > RA, V > PF
I_{Ks}	K_vLQT1 (KCNQ1)	minK (KCNE1)	Termination of plateau, phase 3 repolarization	A ~ V > PF
I_{Kur}	$K_v1.5$ (KCNA5) $K_v3.1$ (KCNC1)		Phase 1 and 2	A only $K_v1.5$ (man), $K_v3.1$ (dog)
I_{K-ACh}	KCNJ3 ($K_{ir}3.1$; GIRK1) KCNJ5 ($K_{ir}3.4$; GIRK4)		Response to ACh	SAN ~ AVN > A \gg V
I_{K-ATP}	$K_{ir}6.2$ (KCNJ11)	SUR2 (ABCC9)	Shortens APD during ischemia (low $[ATP]_i$)	A, V, PF
<i>Anion</i>				
$I_{Cl,cAMP}$	CFTR (ABCC7)		Alters APD	V, PF; not human A
$I_{Cl,CFTR}$				
$I_{Cl,swell}$	(ClC3) ?		Alters APD	A, SAN > V
$I_{Cl,vol}$				
I_{to2}	(CLCAx) ?		Phase 1 repolarization	A, V, not identified in man

Molecular basis given as: common name for protein (gene). SAN, sino-atrial node; AVN, atrio-ventricular node; A, atria; V, ventricle; PF, Purkinje fibers; endo, endocardium; epi, epicardium; ?, uncertain. For additional details, see Refs. [4-6,11,17].

net ionic current, such as the $\text{Na}^+ - \text{K}^+$ pump and $\text{Na}^+ - \text{Ca}^{2+}$ exchanger.

Ionic basis for a prototypic ventricular action potential

The relationship between the action potential and the flow of time-dependent ionic currents is shown diagrammatically in Fig. 3. Starting with phase 4 (diastole), each phase of the action potential will be reviewed in turn. E_m in ventricle remains constant during phase 4 and is about -85 mV , which approaches the value of E_K , -95 mV , with a physiologic serum K^+ concentration of $3.5\text{--}4.5\text{ mM}$. Resting E_m approaches E_K because K^+ is the ion with the greatest conductance during phase 4. The specific K^+ channel responsible is the I_{K1} channel,^[7,8] and its conductance is termed g_{K1} (Table 1). It is important to note, however, that the constancy of E_m during phase 4 requires that inward and outward currents are exactly in balance. The outward I_{K1} is balanced by inward currents carried primarily by Na^+ . The resting Na^+ conductance, g_{Na} , is much smaller than g_{K1} because almost all of the Na^+ channels are closed at -85 mV , but E_{Na} is $+60\text{ mV}$, and the driving force for Na^+ entry, $E_m - E_{\text{Na}}$, is nearly 15-fold greater than that for K^+ efflux. As expected from this analysis, replacement of extracellular Na^+ with an impermeant cation causes E_m to hyperpolarize and approach E_K . The remaining discrepancy between

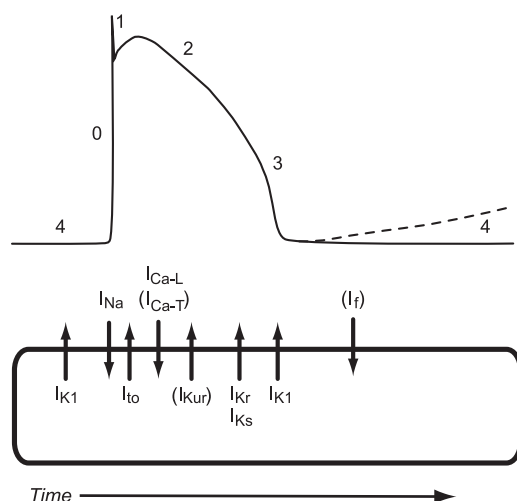


Fig. 3 Ionic currents during generic action potential. Main inward and outward currents are denoted, respectively, by arrows entering or leaving the cell and are placed at time of their activation. Currents found throughout the heart are included, although the distribution of ionic currents is heterogeneous (Table 1). For the illustrated ventricular epicardial action potential, currents in parentheses normally are not significant.

E_m and E_K is largely explained by Cl^- conductance, which carries inward current during phase 4.

Because the K^+ conductance dominates, resting E_m is sensitive to serum K^+ and will follow E_K . For example, increasing serum K^+ from 4 to 6 mM : 1) shifts E_K positively by approximately 10 mV ; 2) reduces K^+ driving force; 3) reduces outward I_K ; and thereby 4) causes a depolarization of approximately 10 mV . Elevation of serum K^+ causes the myocyte to depolarize until the inward Na^+ and outward K^+ currents are again in balance. The final value of E_m primarily reflects the restoration of the driving force for K^+ and thereby K^+ efflux, although voltage-dependent changes in conductance and the effect of E_m on the Na^+ driving force also make small contributions.

Phase 0, the upstroke, begins when a ventricular myocyte starts to depolarize. Normally, depolarization is initiated by the flow of current through gap junctions from nearby myocytes that already are depolarized, or it may be initiated by an electrical stimulus such as that provided by an implantable pacemaker. Depolarization opens Na^+ channels, giving rise to an inward I_{Na} .^[9,10] If the depolarization is sufficient to reach the threshold potential, the inward current produced is larger than the outward current, and as a result, the myocyte will continue to depolarize without a further stimulus. As it continues to depolarize, additional Na^+ channels open, producing an even larger inward I_{Na} , further depolarization, and the opening of more Na^+ channels in a positive feedback loop. Because g_{Na} is much greater than any other conductance during the upstroke, the activation of Na^+ channels elicits a large inward I_{Na} that rapidly drives E_m to about $+50\text{ mV}$ in $<2\text{ ms}$. I_{Na} is transient, however, for two reasons. First, depolarization not only activates (opens) Na^+ channels, but it also causes them to inactivate (close). Therefore, both the Na^+ channel opening and the increase in g_{Na} will be transient. Individual Na^+ channels remain open for $<1\text{ ms}$. Second, as the upstroke is inscribed, the driving force for Na^+ entry decreases. At a resting E_m of -85 mV , the driving force for Na^+ ($E_m - E_{\text{Na}}$) is 145 mV , whereas at the peak of the action potential, $+50\text{ mV}$, it is only approximately 10 mV . Thus, I_{Na} is much reduced near the peak of the upstroke.

Phase 1, initial repolarization, begins when outward currents again become larger than inward currents as I_{Na} decays. Because the driving force on K^+ at the peak of the action potential is large, one might expect an enormous outward I_{K1} . This is not the case, however. I_{K1} exhibits strong inward-going rectification, also termed anomalous rectification, and its conductance is markedly reduced almost instantaneously on depolarization.^[7,8] This decrease in g_{K1} minimizes the K^+ efflux that otherwise would occur and permits development of the remainder of the cardiac action potential. The rectification of I_{K1} is not due to an



intrinsic voltage-dependence of the channels, however. Outward current through the I_{K1} is suppressed because the pore is blocked by millimolar concentrations of intracellular Mg^{2+} and micromolar concentrations of intracellular polyamines, such as spermine and spermidine.^[8]

Other currents contribute to phase 1 repolarization. The transient outward current, I_{to} , is a K^+ current that rapidly activates upon depolarization and inactivates over roughly 100 ms, so it contributes to phase 1 and phase 2.^[4,5,11] As will be noted later, the expression of I_{to} varies in different parts of the heart. A second transient outward current, I_{to2} , also has been identified. I_{to2} is a Ca^{2+} -activated Cl^- repolarizing current due to a conductance that tracks changes in intracellular Ca^{2+} .^[6]

Ventricular action potentials have a pronounced plateau, phase 2, that constitutes most of APD. During the plateau, E_m is near 0 mV and is changing slowly, representing a near balance of inward and outward currents. The inward current during the plateau is primary I_{Ca-L} .^[4,5,10] Like the Na^+ channel, the L-type Ca^{2+} channels open on depolarization. I_{Ca-L} activates more slowly than I_{Na} , reaching a peak in about 10 ms, and inactivates slowly during the plateau. Its amplitude is about 1/50 that of I_{Na} . The inward current is balanced primarily by the outward K^+ currents, I_{Kr} and I_{Ks} .^[10-12] These currents are termed delayed rectifiers because they activate over time upon depolarization. I_{Kr} is the rapidly activating component, and I_{Ks} activates slowly. Overall, the amplitude of I_{Ca-L} and the kinetics of its inactivation and the amplitude of I_{Kr} and I_{Ks} and the kinetics of their activation are the primary factors that set the duration of the plateau, and thereby APD. During the plateau, the inward Ca^{2+} current is decaying and the outward K^+ currents are increasing. Eventually this leads to a more rapid repolarization and the onset of phase 3. The critical role of I_{Kr} and I_{Ks} in setting APD is emphasized by analysis of certain types of dysrhythmias. Forms of genetic, acquired, and drug-induced long-QT syndrome have been traced to modification of these K^+ conductances and a resulting prolongation of ventricular APD.^[13]

The transition between phase 2 and phase 3 (repolarization) is characterized by an acceleration of the rate of repolarization rather than by a distinct transition. As the membrane repolarizes, g_{K1} and I_{K1} begin to increase again. In addition, a fraction of I_{Kr} channels produce extra outward current while transiently passing through the open state as they move from the inactivated state back to the closed state.^[4,5,12] Phase 3 also has an important effect on Na^+ and Ca^{2+} channels. Repolarization is required for both Na^+ and Ca^{2+} channels to recover from the inactivated (closed) state, to the resting (closed) state, from which they can

again open.^[9,10] Recovery of Na^+ channels occurs rapidly during phase 3, and the recovery of Na^+ channels restores the excitability of the cell. Because of the high density and rapid recovery kinetics of Na^+ channels, ERP in ventricular as well as atrial and His-Purkinje myocytes ends during phase 3.

Membrane conductance is much lower (i.e., resistance is higher) during the plateau than during diastole. Therefore, although I_{Ca-L} , I_{Kr} , and I_{Ks} are most important, the potential attained during the plateau and the timing of repolarization are sensitive to several smaller currents. One is due to the Na^+-Ca^{2+} exchanger, which produces a current by coupling the efflux of one Ca^{2+} ion to the influx of three Na^+ ions. At resting E_m , Ca^{2+} is extruded and an inward current is produced, but depolarization causes reversal of the Na^+-Ca^{2+} exchanger, generating an outward current. The Na^+-Ca^{2+} exchanger current again flows inward later in the action potential because of the transient increase in $[Ca^{2+}]_i$ and, eventually, phase 3 repolarization. The exact voltage and, therefore, the timing of the reversal of the Na^+-Ca^{2+} exchanger depends on $[Na^+]_i$. A second current is due to the Na^+-K^+ pump, which extrudes 3 Na^+ while taking up 2 K^+ and contributes a small time-independent outward current to the plateau. A third is I_{to} , which contributes a decaying outward current, especially early in the plateau. Finally, at least two Cl^- currents are important.^[6] There is a background Cl^- current that has been attributed to the same channels that produce a volume-sensitive current, and a cAMP-sensitive Cl^- current that is due to the cardiac variant of the cystic fibrosis transmembrane regulator (CFTR). Both of these Cl^- currents are largely time- and voltage-independent at the plateau potential. Nevertheless, they have significant effects on APD and smaller effects on resting membrane potential.

Because of the interaction of multiple currents and their time- and voltage-dependence, control of the action potential plateau and APD is more complex than may be apparent. Ion channel kinetics are highly voltage-dependent, and in some situations, this can lead to counterintuitive outcomes. For example, an intervention that shifts the plateau positive (e.g., increased inward or decreased outward current) can, in certain situations, result in a more rapid activation of I_{Kr} and I_{Ks} and a paradoxical decrease in APD. Because of these complex interactions, computer models are useful and sometimes necessary for predicting changes in APD.^[14,15]

Recovery of excitability

The prolonged depolarization that is characteristic of cardiac action potentials has an important

consequence. It sets the maximum frequency at which action potentials can be elicited. The current required to initiate an action potential is described as excitability, and a reduction in excitability means that a greater current is required to excite the cell and give rise to another upstroke.

The recovery of excitability is time-dependent, and three periods can be defined, as illustrated in Fig. 4. During the absolute refractory period (ARP), which begins at the upstroke, myocytes are inexcitable and no stimulus can elicit another action potential. Following the ARP is the relative refractory period (RRP). An action potential can be evoked during the RRP, but a larger current is required to excite the cell than when the cell is at rest (phase 4), and therefore excitability is said to be reduced. Moreover, threshold potential during the RRP is more positive than at rest. Excitability and threshold potential return to their resting levels as phase 4 begins.

As a practical matter, the effective refractory period (ERP) is both more important for the heart and more easily measured. The ERP extends from the upstroke to the middle of the RRP and is defined as the period during which one cannot elicit a conducted action potential. Thus, the end of the ERP marks the time when a stimulus can give rise to an action potential that is robust enough to depolarize neighboring cells and spread to the rest of the heart. This definition of the ERP implies that an action potential can be elicited early during the RRP, but it will fail to conduct.

The basis for the time-dependent recovery of excitability is the recovery of the ability to generate an inward current. Depolarization during the upstroke

causes Na^+ channels to open and then inactivate.^[9,10] The Na^+ channels remain in the inactivated state until repolarization occurs during phase 3. Repolarization leads to Na^+ channels returning to their closed state, from which they can open again. This process is rapid and voltage-dependent, and more and more Na^+ channels recover to the closed state as phase 3 repolarization progresses. As a result, sufficient Na^+ channels recover to generate a net inward current and elicit an action potential by the middle of phase 3. Excitability returns as more and more Na^+ channels recover over time. The opening of Na^+ channels is voltage-dependent, so when a greater number of Na^+ channels are in the closed state, a smaller current and less depolarization are required to open a sufficient number of channels to generate a net inward current.

The recovery of Na^+ channels is responsible for the recovery of excitability in atria, ventricle, and the His-Purkinje system, and the relationship of the refractory periods and repolarization is similar in each. Certain pharmacologic agents that are useful in treating ventricular arrhythmias, such as lidocaine, act in part by trapping Na^+ channels in the inactivated state, which delays recovery of excitability. As discussed in a subsequent section, recovery of excitability also is delayed in the AVN because Ca^{2+} channels rather than Na^+ channels are responsible for the upstroke.

REGIONAL VARIATIONS IN CARDIAC ELECTROPHYSIOLOGY

Sino-Atrial Node

The most negative potential attained by SAN myocytes—maximum diastolic potential (MDP)—is about -50 mV; the upstroke in SAN is slow, $2\text{--}5 \text{ V} \cdot \text{s}^{-1}$; and spontaneous depolarization during phase 4 is responsible for pacemaker activity (Fig. 2). Earliest excitation is in the center of SAN. The precise location is not fixed, but moves in response to altered autonomic input and other interventions.^[16] The SAN is heterogeneous.^[2,3,16] Microscopic observation reveals both stellate and elongated spindle cells. Electrophysiologically, the stellate myocytes have a faster intrinsic rate, a more positive MDP, and a longer APD, and they are the primary pacemaker cells.

The expression of ion channels in SAN is different from that in ventricle, atria, and the His-Purkinje system, and these differences can explain the distinct electrophysiological properties of SAN.^[4,5,17] I_{Na} is absent in the primary pacemaker cells. L- and T-type Ca^{2+} channels ($I_{\text{Ca-L}}$ and $I_{\text{Ca-T}}$) rather than I_{Na} are responsible for the upstroke (Table 1). $I_{\text{Ca-T}}$ is particularly prominent in SAN, and it activates and inactivates more rapidly and at more negative potentials

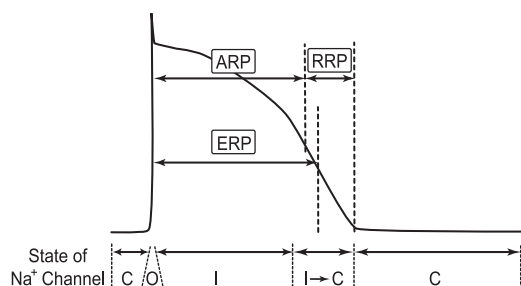


Fig. 4 Recovery of excitability during action potential. Absolute refractory period (ARP): no stimulus can elicit action potential; relative refractory period (RRP): action potential can be elicited, but current required is greater than during phase 4, and threshold potential is more positive; effective refractory period (ERP): a conducted action potential cannot be elicited. Time-dependence of excitability reflects the state of Na^+ channels. Depolarization causes opening and then inactivation. Channels recover from the inactivated to the closed (resting) state during phase 3. C, closed (resting); O, open; I, inactivated; I \rightarrow C, inactivated to closed transition.



than I_{Ca-L} . The upstroke in SAN is very slow because I_{Ca-L} and I_{Ca-T} are much smaller than I_{Na} , and their activation upon depolarization is slower. The depolarized level of MDP in SAN is due to the absence of I_{K1} , the current responsible for the high resting K^+ conductance in other parts of the heart. SAN cells do possess I_{Kr} , I_{Ks} , and also I_{K-ACh} , a K^+ current activated by vagal stimulation. Another critical feature in SAN is a non-selective cation current called I_f , or the pacemaker current.^[17] The density of I_f is much greater in the stellate cells that serve as the primary pacemaker than it is in spindle cells. Activation of I_f upon repolarization drives phase 4 depolarization. In addition, deactivation of I_{Kr} and I_{Ks} following repolarization and activation of Ca^{2+} channels as the cell nears threshold potential contribute to SAN pacemaker activity.

Atria

E_m in atria is about -80 mV, and the upstroke is rapid (150 – 250 $V \cdot s^{-1}$). Atrial action potentials are approximately 150 ms in duration and take on a triangular configuration, because the transition from phase 2 to phase 3 is less distinct and phase 3 repolarization is slower than in ventricle. Regional differences within the atria may contribute to atrial arrhythmias.^[4,5,11] APD and ERP are longer in the crista terminalis than in the pectinate muscles in RA, and longer in RA than in LA. Atrial APD and ERP are both significantly abbreviated by parasympathetic stimulation, although the magnitude of the effect is not uniform throughout the atria. These regional differences in atrial APD, ERP, and parasympathetic responsiveness are thought to play an important role in reentry and atrial fibrillation. Recent studies have recognized that atrial myocytes grow into and form a collar in the pulmonary veins. In some cases, these atrial cells become spontaneously active and serve as ectopic foci that initiate atrial tachyarrhythmias.

The ionic mechanisms underlying atrial and ventricular activity are similar, but some differences have been noted.^[4,5,11] The density of I_{K1} is lower in atria, explaining the more positive E_m and the greater difference between E_m and E_K . The shorter APD in atria is explained in part by expression of I_{Kur} , an ultra-rapidly activating component of delayed rectifier K^+ current in atria, that is found along with I_{Kr} , I_{Ks} , and I_{to} . In addition, atria contain I_{K-ACh} , which plays a critical role in the parasympathetic-induced shortening of atrial APD but is sparse in ventricle. As in SAN, both I_{Ca-L} and I_{Ca-T} are found in atria, whereas I_{Ca-T} is observed in ventricle only after hypertrophy. Interestingly, I_f also is expressed in atrial myocytes,^[17] even though working atrial muscle does not normally

exhibit automaticity. It has been suggested that I_f may play a role in atrial ectopic beats in diseased myocytes. At normal atrial diastolic potentials, however, there is little change in the activation of I_f .

AVN

Overall, the cellular electrophysiology and ionic mechanisms in AVN resemble those of the SAN.^[3-5] The action potential shapes and durations are similar, but MDP in AVN myocytes is slightly more negative, about -60 mV; the upstroke is slightly faster, about 10 – 20 $V \cdot s^{-1}$ and the intrinsic automaticity arising in the N and NH regions is slower, 40 – 50 min^{-1} . Importantly, ERP is significantly longer than APD in AVN, because recovery of I_{Ca-L} is slow.^[10] Slow recovery of AVN excitability prevents rapid atrial rhythms from propagating to the ventricle at an excessive frequency. Ovoid-shaped cells isolated from AVN have the characteristics of N and NH regions, and rod-shaped cells appear to arise from the AN region.

I_{Ca-L} is responsible for the upstroke in the N and NH regions,^[4,5] and its small amplitude and slow activation explain the slow upstroke and slow conduction in the AVN. The presence of I_{Ca-L} without I_{Ca-T} gives rise to useful pharmacologic distinctions. L-type Ca^{2+} channel blockers, such as dihydropyridines and verapamil, can impede AVN conduction and prolong ERP with minimal effects on SAN. In contrast to N and NH cells, I_{Na} is found in transitional AN cells. I_{Kr} , I_{to} , and I_{K-ACh} are prominent in AVN, whereas I_{Ks} is less important in AVN than in SAN. As expected from the depolarized MDP, only low levels of I_{K1} are found. The N and NH regions also exhibit strong expression of I_f ,^[17] which is consistent with the pacemaker activity arising in these portions of the AVN. Nevertheless, activation of I_{Ca-L} and the deactivation of I_{Kr} also modulate pacemaker activity.

His-Purkinje System

Action potentials in the His-Purkinje system are noted for the fastest upstroke velocity in the heart— 500 – 600 $V \cdot s^{-1}$ —corresponding to the fastest conduction velocity.^[3-5] E_m is typically -90 mV, slightly more negative than in ventricle. APD is greater in the Purkinje system than elsewhere, especially at slow heart rates, and a distinct notch is observed between phases 1 and 2. Within the His-Purkinje system, the longest APD and ERP are in the terminal Purkinje fiber branches. The prolonged ERP in these myocytes is thought to reduce the risk of reentrant rhythms involving Purkinje fibers and ventricle. On the other hand, prolonged APD favors drug-induced early

afterdepolarizations, which propagate to the ventricle and may induce ventricular tachyarrhythmias such as torsade de pointes.^[13]

Fast conduction velocity in Purkinje fibers is due in part to a high density of I_{Na} . Both I_{Ca-L} and a substantial I_{Ca-T} are found, although the density of I_{Ca-L} is less than in ventricle.^[4,5] Several factors contribute to the long APD. The density of I_{K1} may be less than in ventricle, but this is controversial. A slowly inactivating component of I_{Na} also has been suggested. On the other hand, the density of I_{Kr} and I_{Ks} is similar to that in ventricle. Another difference is in I_{to} . In human heart, the recovery of I_{to} is much slower in Purkinje cells than in ventricle, and it is more sensitive to blockers such as 9-aminopyridine and tetraethyl ammonium, suggesting that different molecular species are responsible for I_{to} in different regions of the heart.^[4,5,11]

Ventricle

Regional variations in ventricular APD are responsible for the upright T-wave in the electrocardiogram, and they contribute to arrhythmogenesis. Epicardial action potentials have a much more prominent notch separating phases 1 and 2 and a shorter APD than those of endocardial myocytes. In addition, a population of midmyocardial cells (M cells) exhibit a pronounced prolongation of APD at slow heart rates, as found also in Purkinje cells.^[18] Left ventricular M cells have a shorter duration and more pronounced notch than those in right ventricle. M cells were first identified in isolated myocytes. In the presence of electrical coupling in intact tissue, the prolongation of APD in M cells at slow heart rates should be minimized by the influence of surrounding myocytes with shorter APD. Nevertheless, the preponderance of evidence argues that midmyocardial cells have a prolonged APD in vivo. The heart rate-dependent transmural gradient of APD and ERP is thought to be an important factor in reentrant arrhythmias, especially when APD is prolonged by slow heart rates or drugs.

CONCLUSION

Thus, the heart is anything but uniform in its bioelectric behavior. Heterogeneity of the action potential, conduction, and automaticity in the atria, ventricles, His-Purkinje system, and SA and AV nodes—and within each of these regions—can now be explained by differences in the expression of ion channels. Detailed knowledge of the diversity of cardiac bioelectricity is essential for understanding both normal electrophysiologic function and the consequences of disease processes.

ACKNOWLEDGMENTS

Preparation of this article was supported by grants from the National Heart, Lung and Blood Institute (HL-46764) and the American Heart Association.



ARTICLES OF FURTHER INTEREST

Biopotential Amplifiers; Cardiac Pacemaker; Defibrillators; Electrocardiography; Optical Mapping

REFERENCES

1. Katz, A.M. *Physiology of the Heart*, 2nd Ed.; Raven: New York, 1992; 2–36, 466–481.
2. James, T.N. Structure and function of the sinus node, AV node and His bundle of the human heart: Part I—Structure. *Prog. Cardiovasc. Dis.* **2002**, *45*, 235–267.
3. James, T.N. Structure and function of the sinus node, AV node and His bundle of the human heart: Part II—Function. *Prog. Cardiovasc. Dis.* **2003**, *45*, 327–360.
4. Roden, D.M.; Balsler, J.R.; George, A.L., Jr.; Anderson, M.E. Cardiac ion channels. *Annu. Rev. Physiol.* **2002**, *64*, 431–475.
5. Schram, G.; Pourrier, M.; Melnyk, P.; Nattel, S. Differential distribution of cardiac ion channel expression as a basis for regional specialization in electrical function. *Circ. Res.* **2002**, *90*, 939–950.
6. Hume, J.R.; Duan, D.; Collier, M.L.; Yamazaki, J.; Horowitz, B. Anion transport in heart. *Physiol. Rev.* **2000**, *80*, 31–81.
7. Baumgarten, C.M.; Fozzard, H.A. Cardiac Resting and Pacemaker Potentials. In *The Heart and Cardiovascular System*, 2nd Ed.; Fozzard, H.A., Haber, E., Jennings, R.B., Katz, A.M., Morgan, H.E., Eds.; Raven: New York, 1991; 963–1001.
8. Lopatin, A.N.; Nichols, C.G. Inward rectifiers in the heart: an update on I_{K1} . *J. Mol. Cell. Cardiol.* **2001**, *33*, 625–638.
9. Hanck, D.A. Biophysics of Sodium Channels. In *Cardiac Electrophysiology: From Cell to Bedside*, 2nd Ed.; Zipes, D.P., Jalife, J., Eds.; Saunders: Philadelphia, 1995; 65–73.
10. Heath, B.; Gingrich, K.; Kass, R.S. Ion Channels in the Heart: Cellular and Molecular Properties of Cardiac Na, Ca, and K Channels. In *Handbook of Physiology, Section 2: The Cardiovascular System, Vol. 1: The Heart*; Page, E., Fozzard, H.A., Solaro, J., Eds.; Oxford: New York, 2001; 548–567.
11. Nerbonne, J.M. Molecular basis of functional voltage-gated K^+ channel diversity in the mammalian myocardium. *J. Physiol. (Lond.)* **2000**, *525*, 285–298.
12. Tristani-Firouzu, M.; Sanguinetti, M.C. Structural determinants and biophysical properties of HERG and KCNQ1 channel gating. *J. Mol. Cell. Cardiol.* **2003**, *35*, 27–35.

13. Makielski, J.; Fozzard, H.A. Ion Channels and Cardiac Arrhythmias in Heart Disease. In *Handbook of Physiology, Section 2: The Cardiovascular System, Vol. 1: The Heart*; Page, E., Fozzard, H.A., Solaro, J., Eds.; Oxford: New York, 2001; 709–740.
14. Noble, D. Ionic Mechanisms of Cardiac Electrical Activity. In *Cardiac Electrophysiology: From Cell to Bedside*, 2nd Ed.; Zipes, D.P., Jalife, J., Eds.; Saunders: Philadelphia, 1995; 305–313.
15. Rudy, Y. The Cardiac Ventricular Action Potential. In *Handbook of Physiology, Section 2: The Cardiovascular System, Vol. 1: The Heart*; Page, E., Fozzard, H.A., Solaro, J., Eds.; Oxford: New York, 2001; 531–547.
16. Boyett, M.R.; Hono, H.; Kodama, I. The sinoatrial node, a heterogeneous pacemaker structure. *Cardiovasc. Res.* **2000**, *47*, 658–687.
17. Accili, E.A.; Proenza, C.; Baruscotti, M.; DiFrancesco, D. From funny current to HCN channels: 20 years of excitement. *News Physiol. Sci.* **2002**, *17*, 32–37.
18. Antzelevitch, C.; Dumaine, R. Electrical Heterogeneity in the Heart: Physiological, Pharmacological, and Clinical Implications. In *Handbook of Physiology, Section 2: The Cardiovascular System, Vol. 1: The Heart*; Page, E., Fozzard, H.A., Solaro, J., Eds.; Oxford: New York, 2001; 654–692.

Cardiac Catheters

Tommy L. Miller

Division of Cardiology, Medical College of Virginia, Virginia Commonwealth University, Richmond, Virginia, U.S.A.

Juana María Alfaro

Cordis de Mexico, Juarez, Mexico

Deepak Banerjee

Benjamin Lee

Division of Cardiology, Medical College of Virginia, Virginia Commonwealth University, Richmond, Virginia, U.S.A.

Anthony J. Minisi

Cardiac Cath Lab—111J1, McGuire VA Medical Center, Richmond, Virginia, U.S.A.



INTRODUCTION

From the perspective of the operating physician, cardiac catheters represent technologic marvels and are a triumph of bioengineering. The design and ease of use of these devices have enabled physicians around the world to perform important diagnostic and therapeutic procedures to the benefit of countless numbers of patients.

HISTORICAL PERSPECTIVE

Physicians in the Greco–Roman and early Egyptian period had access to a variety of medical instruments that were designed to aid them in their treatment of the sick. Many instruments—including scalpels, hooks, forceps, drills, and catheters—were available to Hippocrates and are used today in a modified form of what was used in antiquity.^[1] As early as 3000 B.C., catheters were employed by physicians in order to cannulate obstructed urinary tracts to allow urine to pass freely from the body. The design of these early catheters was a simple, crude hollow tube made of steel or bronze. The S-curve catheter was specifically designed for men and the straighter catheter for women (Fig. 1).

In the early 1700s, Hale performed the first examination of the heart of a horse. He used an archaic catheter contraption made of brass pipes, a glass tube, and the trachea of a goose (Fig. 2).^[2] In 1844, Claude Bernard, using the horse as a model, inserted a mercury thermometer into the carotid artery and advanced it through the aortic valve and into the left ventricle to measure blood temperature. Over the next 40 years, his work with catheters and the measurement of intracardiac pressure in various animals became the

standard method for the study of cardiovascular hemodynamics.^[3]

The extension of these techniques to humans is credited to Werner Forssmann, a German surgical resident.^[4] In 1929 he experimented with a human cadaver and found that it was relatively easy to guide a urological catheter from an arm vein into the right atrium.^[4] The purpose of his investigation was to find a safer approach for intracardiac drug delivery. He took this approach one step further and passed a catheter through one of his own dissected left antecubital veins. He guided the catheter with the aid of fluoroscopy until it entered his right atrium. He confirmed the position of the catheter by a chest roentgenogram.^[4]

Forssmann made multiple attempts at self-catheterization during his training and was heavily criticized by his colleagues because of the perceived danger of his experimentation, as well as his unorthodox approach to scientific investigation. However, he was the first to document the feasibility of right heart catheterization in humans using radiographic techniques. Because of the constant harsh criticism, his interest in cardiology waned and he pursued a career in urology.

Despite the negative response to Forssmann's innovation in Germany, Andre Cournand in New York began utilizing cardiac catheters as a diagnostic tool in the early 1940s. He used these catheters to investigate cardiac function in normal and diseased patients.^[5,6] In 1956, Forssmann and Cournand shared the Nobel Prize in Physiology and Medicine for their foresight and contribution to the advancement of cardiac catheterization.

Over the ensuing years, catheter use for the diagnosis of heart disease advanced rapidly. The diagnosis of congenital and acquired (rheumatic) heart diseases, once so elusive to physicians, was facilitated greatly by this new catheter technology.^[7–10]



Fig. 1 Photograph of first known catheters to be used to access the urinary bladder in humans. (From website: www.Indiana.edu/~anccmed/instr1/html.)

Cardiac catheters were used in an indirect fashion to visualize the coronary arteries by flooding the aortic root with contrast dye, which then flowed into the coronary arteries. In the late 1950s, Mason Sones, a pediatric cardiologist at The Cleveland Clinic, conducted an imaging procedure in which dye was injected into the aortic root of a patient. During the procedure, the catheter accidentally entered the patient's right coronary artery and before it could be removed, a small amount of dye was released.^[11,12] The coronary artery tolerated the dye load and remained intact, and—most significantly—the patient suffered no adverse complications.^[12] Sones further modified the catheter and developed new designs that could enter the coronary arteries selectively. Combining selective coronary catheterization with high-quality imaging techniques, coronary angiography became a significant milestone in the diagnosis of coronary disease. Judkins and Adams further modified the technique of cardiac catheterization, resulting in a better understanding of the pathophysiologic mechanism of heart disease.^[13]

Catheter development continued to improve in the ensuing years but a full discussion of these improvements is beyond the scope of this article. However, a brief chronology of several advancements warrants mentioning. They include:

1. In 1964, Dotter introduced transluminal angioplasty to open blocked peripheral arteries with increasing catheter diameters.^[14]
2. The development of the balloon flotation catheter by Swan and Ganz in 1970.^[15]

3. In 1977, Grüntzig expanded on Dotter's techniques and introduced percutaneous transluminal coronary angioplasty (PTCA) to open blocked coronary arteries.^[16,17]
4. Catheter-based therapy combining PTCA and stenting of the coronary arteries provided significant improvement in the treatment of acute coronary thrombosis.^[18,19]

DIAGNOSTIC CARDIAC CATHETERIZATION IN ADULTS

The development of cardiac catheters and the many advances in imaging techniques have made coronary angiography a major tool in the diagnosis and treatment of coronary artery disease. Many noninvasive techniques such as magnetic-resonance imaging, fast computed tomography, and intravascular ultrasound have aided in the diagnosis of coronary artery disease. However, selective coronary angiography remains the gold standard.^[20,21]

There are numerous indications for the use of catheter-based coronary angiography, and as catheter technology improves, the indications may broaden. However, the most common indication for cardiac catheterization is the evaluation of coronary atherosclerosis.

There are several techniques used to access and view the coronary arteries using coronary catheters. In most patients, coronary angiography can be performed by brachial cut-down or by the more common



Fig. 2 Artist drawing depicting the first known attempt to measure intravascular pressures in a live horse. (From website: www.PTCA.org/history_timeline.html. This also appears in *Cardiac Angiography*; edited by S.W. Miller, published by Little, Brown, and Company, New York, 1984; citation for this figure in this book is Medical Times, Nov. 1944, p. 314.)

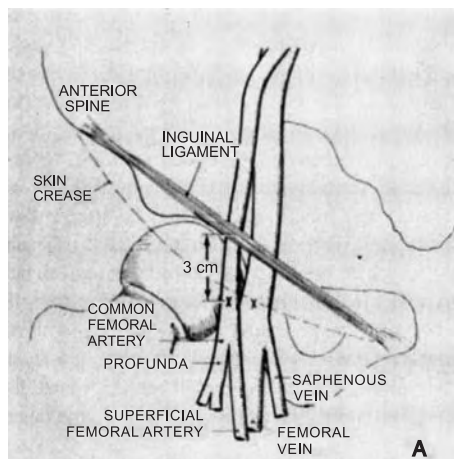


Fig. 3 Anatomy of the femoral artery and femoral vein at the site of puncture during cardiac catheterization performed from the groin. (From *Grossman's Cardiac Catheterization, Angiography, and Intervention*, 6th ed.; edited by D. S. Baim and W. Grossman; published by Lippincott, Williams, and Wilkins, Baltimore, 2000; p. 71; Fig. 4.1, panel A.)

percutaneous approach from the femoral, brachial, or radial artery.

FEMORAL ARTERY APPROACH

Femoral, brachial, and radial arterial entry gives access to the aorta and the left side of the heart. The femoral artery approach is most often used during coronary artery angiography.^[22] Trauma can occur to the vessel during access and may result in vascular compromise during or following the cardiac catheterization procedure. Assessment of the peripheral pulses is mandatory prior to accessing the arterial circulation. It is most important to assess the femoral artery and its

distal branches, including the dorsalis pedis (top of the foot below the ankle) and the posterior tibial (inside and behind the ankle) prior to cardiac catheterization. Once adequate pulses are determined, assessment of any changes can easily be made postprocedurally.

The femoral artery is a tube approximately 1 cm in diameter that crosses the inguinal ligament, which extends from the iliac crest to the pubis. The actual arterial entry is 1–2 cm below the inguinal ligament (Fig. 3).

Local anesthesia with 1% lidocaine is delivered to the superficial and deep tissue planes. Adequate anesthesia is provided to cover the complete depth of the skin-to-artery path to decrease the level of discomfort for the patient.

Single, anterior percutaneous arterial wall entry technique with a hollow tip-beveled needle is preferred to the original Seldinger double-wall puncture technique. The latter was associated with increased hemorrhagic complications.

The needle is introduced through the skin and advanced toward the artery at a 30–45° angle (Fig. 4). Once the artery is punctured, a flexible J-tipped guidewire is introduced through the hollow needle and advanced slowly and gently into the artery. Fluoroscopy is used frequently throughout the procedure to check for guidewire placement.

The guidewire is advanced to the appropriate position in the ascending aorta and the puncture needle is removed. Many laboratories utilize a catheter sheath inserted over the guidewire for increased comfort for the patient and limitation of the number of catheter exchanges through the artery. Various femoral catheters are then inserted with a J-tipped guidewire for left and right coronary artery evaluation.

If there is significant peripheral vascular disease and the femoral artery cannot be used to access the aorta and the left side of the heart, then the brachial and/or radial artery can be used. Care again must be taken to

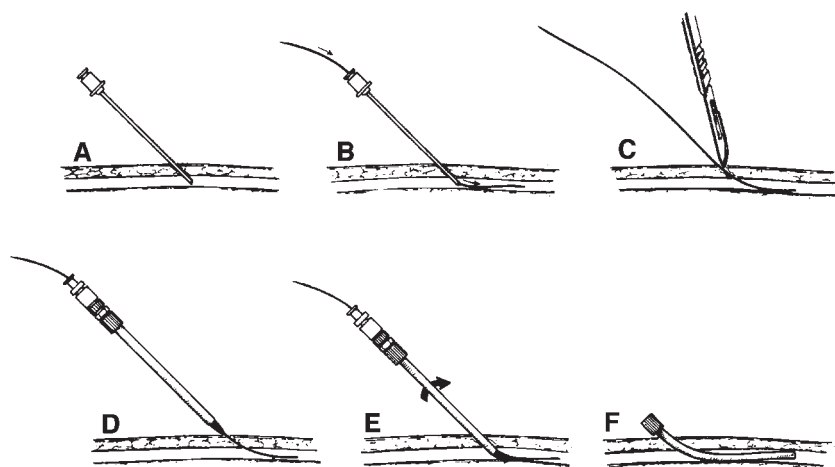


Fig. 4 Schematic depicting the various steps in obtaining vascular access using the Modified Seldinger technique. See text for details. (From *Diagnostic and Therapeutic Cardiac Catheterization*, 2nd ed.; edited by C. J. Pepine, J. A. Hill, and C. R. Lambert; published by Williams and Wilkins, Baltimore, 1994; p. 91; Fig. 8.1.)

properly assess for adequate pulses prior to the procedure.

CORONARY CATHETER TECHNOLOGY

Of all the cardiac catheters, the development and manufacture of coronary catheters presents the most unique bioengineering challenge. Early angiographers were severely limited with primitive, nonradiopaque polyethylene catheters that were hand-fabricated and had to be filled with contrast agent to be visible under rudimentary fluoroscopes that provided poor imaging quality.^[23]

Charles Dotter, M.D., an interventional radiologist, made significant contributions to the field of cardiac catheterization and catheter design.^[24-26] Dotter was very innovative, and initially designed and invented his own interventional catheters. He used unconventional materials such as guitar strings, Volkswagen automobile speedometer cables, and vinyl insulation stripped from pieces of intercom cable. Dotter refined his inventions and finally developed catheters and guidewires made from Teflon. Since then, advances in catheter technology have contributed to improved diagnosis and treatment of atherosclerotic heart disease.

Key to the successful engineering of intravascular and intracardiac catheters was the explosive development of a class of materials known as plastics. Plastics include numerous organic, synthetic, or processed materials that revolutionized the manufacturing industry. A plastic is made up principally of a binder together with plasticizers, fillers, pigments, and other additives. The binder gives a plastic its main characteristics and often its name. Binders consist of very long-chainlike molecules called polymers. Plasticizers are added to a binder to alter the material properties of the compound such as flexibility and toughness. Fillers are added to further manipulate properties such as hardness and resistance to shock. Pigments are used to impart optical characteristics such as radiopacity.

There are two basic types of plastic useful to cardiac catheter manufacturing: thermosetting plastics, which cannot be resoftened after initially being subjected to

heat and pressure; and thermoplastics, which can be repeatedly softened and remolded by heat and pressure. When heat and pressure are initially applied to a thermosetting binder, the molecular chains become cross-linked, thus preventing any slippage if heat and pressure are reapplied. When heat and pressure are applied to a thermoplastic binder, the chainlike polymers slide past each other, giving the material plasticity, or the ability to be molded to a predetermined shape, and later capable of being deformed to some degree while retaining much of its previous shape.

Diagnostic cardiology catheters have evolved over the past few decades to include such modifications as smaller outer diameters, larger inner diameters, soft atraumatic tips, enhanced torque control, and a variety of shape designs. All this has been done in an attempt to make catheterizations easier for both physician and patient.

The construction of these multilayered plastic tubes involves the union of several elements, which determine the performance of the product (Fig. 5):

- The hub connects the catheter to syringe or power ejector.
- The strain relief acts as a reinforcement near the hub to prevent kinking of the catheter's main tubular section.
- The main tubular section is commonly referred as the body of the catheter.
- The proximal tip function is to support the shape transition between both braided and nonbraided segments.
- The distal tip of the catheter is designed to be soft and atraumatic in order to prevent vascular damage.

The market offers several catheter brands that basically share the same manufacturing steps:

Compounding

In the compounding process, plastic pellets are blended into a compound. Thermoplastic materials such as polyurethane or nylon are heated and blended with lubricants and radiopaque materials into a molten mix

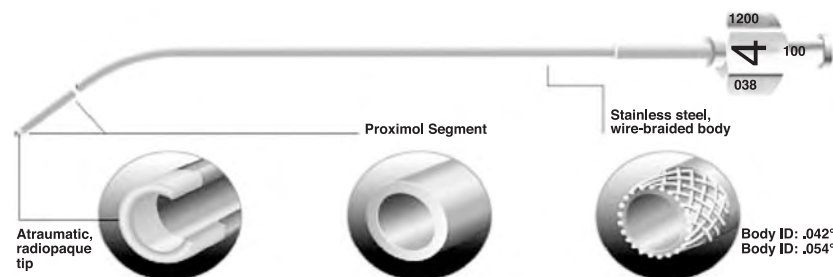


Fig. 5 Schematic representation of the elements of a diagnostic cardiac catheter. (Courtesy of Cordis, Corporation, Miami, FL, a Johnson & Johnson company.)

that has the consistency of bread dough. The material is mixed further on rollers that spread the material to a consistent thickness. It is pulled off the rollers, cut into strips, and allowed to cool. The strips are diced into pellets that will be used in the extrusion process.

Extrusion

Extrusion refers to forcing, pressing, or pushing a material through a tool or devices for imparting a desired shape. The pelletized lot of compound is taken to the extrusion floor. The extrusion process heats the thermoplastic material and forces it into a tubular shape. This is accomplished in one of two ways, either by air extrusion or mandrel extrusion.

In the air extrusion process, the molten thermoplastic material is blown into its tubular shape by a volume of air, then passed through distilled water and allowed to cool.

In the mandrel extrusion process, the hot thermoplastic material is spread in a layer over a mandrel wire to form its tubular shape. This becomes the base coat layer of the catheter. It is passed through a column of water to cool. Sophisticated computers precisely control the thickness of the base coat.

Braiding

Stainless steel braiding makes up the middle of the catheter wall. A large spool of extruded, hollow tubular thermoplastic is placed into the braiding machine, which weaves a knitted pattern of braid onto the base coat using a certain number of strands of very fine stainless steel wire.

Following the braiding process, the large spool of braided base coat is taken back to the extrusion floor, where another layer of plastic is extruded over the braiding. This is the topcoat of the catheter.

Catheter Assembly

At the end of the topcoat extrusion line, the catheter bodies are cut into their approximate final length. The mandrel wire is removed. Each end of the catheter body is prepared for assembly by grinding and in some cases flaring the end.

At one end, a nonbraided tip segment is fused to the catheter body. At the other end of the catheter body, a strain relief is fused on, and then a polycarbonate hub is bonded to the catheter. Finally, the catheter is cut to its final length, fully inspected, and tested using hydrostatic pressure test.

Shape Forming

This is a key step with exacting standards. For instance, the success of the Judkins line of catheters in minimizing the time for coronary artery catheterization is due in large part to the recognition that the ostia of the left and right coronary arteries lie in relatively set planes with respect to the catheters' long axis. The catheters are taken to a station to form the tip into its specified shape. Forming rods preshaped in the intended curve are inserted into the catheter tip segments. The tip segment is heated for a specified time and then immediately cooled. The forming rods are removed and the catheter tips retain their desired shapes because they have been heat-set. This is the key characteristic of the thermoplastic materials used in the manufacture of coronary catheters.

The functional specifications of coronary catheters are influenced by the materials used in the manufacturing process and by engineering specifications such as catheter length, diameter size, wall thickness, and maximal injection pressure and flow rates.^[27] Important and key aspects of the physician-device and patient-device interface include the following:

- Torque control.
- Kink resistance.
- Flexibility.
- Pushability, related to friction.
- Stability during injection.
- Radiopacity.
- Shape retention.
- Thrombogenicity.

Relative to these key parameters, the ideal coronary catheter should be soft enough to avoid vascular trauma. This is particularly true for the tip of the catheter that ultimately engages the ostium of the coronary artery. However, the catheter must also be firm enough to maintain its preformed shape, especially as it warms from room temperature to body temperature. In addition, these catheters must also be advanced, retracted, and maneuvered by the application of rotational torque. During difficult cases in which there is extensive tortuosity of the aorta or iliofemoral arteries, extensive catheter manipulation may be required. The composition of the catheter shaft must allow for external forces that are applied to the distal end of the catheter outside the body to be transmitted reliably to the tip of the catheter in the ascending aorta.

Superimposed on these important parameters is the issue of catheter and lumen diameter size. Ideally, a coronary catheter would be the smallest size possible to minimize trauma to the arterial entry site (femoral, brachial, or radial). However, the need to maneuver



the catheter limits the extent of miniaturization that is possible. In addition, there are limits as to how small the catheter lumen can be. There are two major classes of coronary catheters that have two distinct purposes. Angiographic catheters are designed to inject contrast material into the coronary artery at a rate and volume that is sufficient to create a diagnostic quality angiogram. Achieving this goal is more difficult when using catheters with smaller lumen diameters are used. Angioplasty guiding catheters must not only allow for injection of contrast material, but also have lumens of sufficient size to allow the passage of the various devices used in contemporary interventional cardiology procedures.

Thus, there are important tradeoffs between the desire for the smallest possible coronary catheter and the ability of that catheter to accomplish its intended purpose. For instance, manufacturing a smaller catheter while keeping the lumen diameter constant requires that the shaft of the catheter be composed of thinner material. This increases the susceptibility of this catheter to kinking during catheter manipulation, particularly rotational torsion.

Historically, catheters have been manufactured from various biomaterials. Synthetic polymers have made a significant impact on catheter development, virtually replacing older biomaterials in cardiovascular procedures. This transition began with the shift from metal catheters to those made of polyethylene. Nylon became available in the 1930s, polyurethane was introduced in 1937 as a surface coating, silicones were introduced during World War II, and other biocompatible polymers were developed in the 1950s. A summary of the relevant polymers used in coronary catheter construction follows.

Teflon

This biomaterial has a low frictional coefficient and high memory, which allow for excellent gliding performance with decreased resistance. Because it has good memory, it cannot be shaped at temperatures below 100°C.

Nylon

This is the common name of a thermoplastic linear polyamide resin that has outstanding properties. This polymer has excellent chemical resistance, good stress crack resistance, low moisture absorption (0.9%), low coefficient of friction and high-pressure resistance with dimensional stability. Based on these properties, it is possible to design smaller catheters with a reduced inner diameter (ID) without compromising performance (the largest ID could be 0.057" and the smallest

0.042"). Nylon catheter construction is highly responsive to torque transmission. In addition, nylon catheters can be designed to provide a uniform ID from the hub of the catheter to its tip. This results in excellent flow dynamics and optimal angiographic results.

Polyurethane

This biomaterial holds an important role in cardiac applications, and the benefits of this material lie in its physical properties. This polymer has been described to bridge the gap between rubber and plastic and also has a load-bearing capacity comparable to cast steel. Outstanding properties of polyurethane include exceptionally smooth surfaces, resistance to colonization by fungal organisms, and excellent hydrolytic stability. Braided polyurethane catheters are built for maximum pushability and precise torque control. Nonbraided proximal segments and distal tips are designed for flexibility and shape retention.

Polyethylene

This thermoplastic biomaterial exhibits moderate-to-high tensile strength with moderate elongation. Polyethylene has high memory and can easily be reshaped during the manufacturing process.

Silicone

This is an unusual polymer in that it is partly organic and partly inorganic. Silicones are based on polymers in which the main polymer chain consists of alternating silicon and oxygen atoms. They are derived from sand and methyl chloride. The notable properties of this polymer are a high degree of lubricity, excellent water resistance, chemical inertness, retention of most physical properties at elevated temperatures, and compatibility with the human body. Silicone is often prepared with barium sulfate—a substance required to make it radioopaque—but this addition can also make it less elastic and flexible and more fragile.

PVC

Polyvinylchloride is a thermoplastic material; during its compounding into commercial resin, ingredients are added, such as stabilizers to prevent heat degradation, plasticizers to impart flexibility, lubricants as processing aids, fillers for cost-reduction purposes, and impact modifiers to improve toughness. The hubs of most catheters are composed of PVC and the production process is done by injection-molding. This polymer

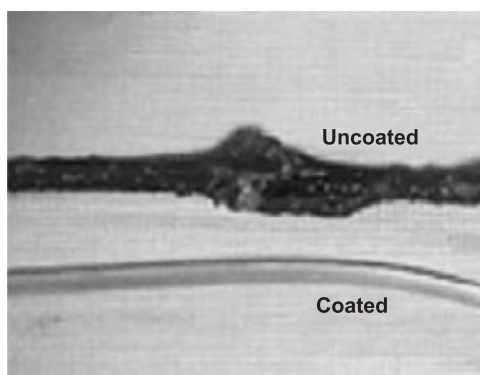


Fig. 6 Photograph illustrating the differing thrombogenic potential between a coated and a noncoated catheter. (From website: www.surmodic.com.)

exhibits high friction and has the potential to grow hard over time.

A very important property for medical devices that come into contact with blood after insertion in a human body is thrombogenicity. Thrombogenicity is the ability of the device to induce formation of blood clots. For blood to clot, it is believed that thrombin must be generated in the blood. Several potent anticoagulants can prevent the formation of thrombin. The most widely used anticoagulant is heparin. When heparin is incorporated onto the surface of a medical device, the ability of the surface to induce formation of clots decreases, or in other words, the surface becomes antithrombogenic (Fig. 6). However, many heparinized surfaces tend to have poor stability *in vivo* and the antithrombogenicity of a heparinized surface tends to decrease as the surface heparin loses its potency to retard clot formation. New heparin compositions and coatings are needed to improve the antithrombogenic stability of heparinized surfaces. Examples of these coatings used in the diagnostic catheters industry are the MDX4-4159 fluid 50% medical-grade dispersion, trichloroethene, carnauba wax, and other mixtures of non-heparin synthetic hemocompatible materials such as hydrophilic coatings.

Heparin is a pharmaceutical that has been used clinically for decades as an intravenous anticoagulant to treat inherent clotting disorders and to prevent blood clot formation during surgery and interventional procedures. Heparin molecules are polysaccharides with a unique chemical structure that gives them specific biological activity. When heparin is immobilized onto the surface of a medical device material, it improves the performance of the material when it comes in contact with blood in several ways:

1. Provides local catalytic activity to inhibit several enzymes critical to the formation of fibrin (which holds thrombi together).

2. Reduces the absorption of blood proteins, many of which lead to undesirable reactions on the device surface.
3. Reduces the adhesion and activation of platelets, which are the primary component of thrombus.

Nevertheless, heparin compounds and methods to prepare them present many bioengineering challenges such as: 1) heparin's poor adhesion to surfaces, which requires special surface and thermal treatments that may result in degradation of surface properties; and 2) they must be applied in thick coating layers using toxic chemical compounds.

The MDX4-4159 fluid 50% medical-grade dispersion material provides an alternative solution that circumvents many of the problems related to preparation of heparin compounds. The MDX coating is a dispersion containing 50% active silicone ingredients in mixed aliphatic and isopropanol solvents. The significant features of this dispersion are:

1. It has excellent lubricity, providing a smooth surface to the catheters.
2. It has high adhesion to some plastics such as polyurethane used to build diagnostic catheters.
3. It offers an easy application from dilute solvent solutions.
4. It cures at room temperature upon exposure to atmospheric moisture.

The coating of the catheter surface may be applied by dipping or wiping the catheter body with the MDX4-4159 fluid (Fig. 7). Spray application can be used but is not recommended because of the high risk of inhalation toxicity. The coating thickness can be varied to fit individual needs. The coating should be thin enough to be invisible to the unaided eye. At 25°C (77°F) and 55% relative humidity, the coating should cure enough in 24 hours to allow handling and use of the coated catheters.

Trichloroethene (TCE) is a colorless liquid at room temperature with a somewhat sweet odor similar to ether or chloroform. It is used primarily as a solvent to remove grease from metal surfaces. In the diagnostic catheter industry, TCE is used to create a repellent or nonstick catheter surface. This prevents the build-up of cells and proteins that causes blood clot formation during the catheterization procedure. Carnauba wax can be added to TCE to create a solution that provides additional lubricity to the catheter surface. The manufacturing process to apply this solution involves dipping and wiping the surface of the catheter, and is similar to the process used with the MDX4-4159 coating (Fig. 7).

Because friction is a very important factor related to catheter performance, modern hydrophilic coatings





Fig. 7 Photograph illustrating the process of wiping the catheter surface following dipping in MDX4-4159 fluid. (Courtesy of Cordis de Mexico S.A. de C.V., Chihuahua, Mexico.)

(wet lubricity) have been applied to the external portion of the catheter in order to smoothen the otherwise rough surface. In many cases, these hydrophilic coatings were found to be antithrombogenic, causing many to employ them as an alternative to heparinization. Hydrophilic coatings can also limit bacterial colonization and reduce protein and platelet adhesion.

In comparison to known heparin compounds and the challenges of making and using them, hydrophilic solutions provide one or more of the following advantages:

- Spontaneous wetting and improved flow.
- Reduced procedure time.
- Reduced insertion forces.
- Increased patient comfort.
- Enhanced ability to maneuver through tortuous vascular anatomy.
- Reduced tissue irritation and damage.

The manufacturing process to coat the catheters with hydrophilic solutions is versatile and simple. The reagents are dissolved in water or a water/alcohol mixture, applied to a cleaned catheter shaft, and finally activated by an ultraviolet light lamp. The entire activation stage takes only seconds to minutes of exposure to ultraviolet light to accomplish the chemical bonding reaction at the catheter surface. This is a distinct advantage over other coating processes that require lengthy curing times. This technology uses no harsh chemicals and produces no hazardous by-products, and there is no need for surface-expanding pretreatment that can often affect the catheter's physical properties or performance. In addition, no expensive equipment is required to comply with environmental regulations.

Finally, this kind of technology adapts well to either batch processing or incorporation into continuous-flow production lines allowing high volume and lean manufacturing systems. Inspections can be done online for early feedback, and as a result, fewer catheters are rejected for failing to meet specifications. In general, this process maintains consistent performance with predictable outcomes, eliminating batch-to-batch variability.

COMMONLY USED CARDIAC CATHETERS

As stated earlier, initial attempts to perform coronary angiography used nonselective injections of contrast medium into the aorta to simultaneously opacify both the left and right coronaries. The development of new catheter technology has made nonselective coronary angiography obsolete except in cases where coronary anatomy is distorted or anomalous or the origins of saphenous vein bypass grafts must be visualized to guide subsequent selective cannulation. Today, cardiac catheterization involves selectively cannulating the coronary vasculature and the chambers of the heart. Several catheters are available that are preshaped and vary in design and luminal diameter. This is due to the anatomical differences in the right and left coronary origins as well as variations and pathology in the ascending aorta and arch. There are left and right catheters for engaging the corresponding coronary arteries respectively. The outer diameter of a catheter is determined in French units (1 French unit = 0.33 mm).^[28] The catheters are available in a variety of sizes ranging from 3.5 F to 8 F.

Smaller-diameter catheters allow for smaller arterial punctures, greater patient comfort, and earlier patient ambulation. On the other hand, catheters with a larger lumen offer better opacification of coronary vessels, thus improving the image resolution. They give the operator greater maneuverability and stability during the procedure and minimize dislodgement during contrast injection. The selection of the proper catheter size necessitates a balance between these properties, consideration for the size of the target vessels, and physician preferences based on individual experience and comfort level.

The Judkins-type catheters are the preferred catheters for coronary arteriography via the femoral approach in most cases. Other catheters commonly used in this route include the Amplatz-type catheters and the Multipurpose catheters. Preshaped bypass graft catheters are available for the saphenous venous grafts and the internal mammary grafts. Each of these various catheters has a unique method for engaging and disengaging the coronary ostia. Pigtail catheters are used for instrumentation of the left ventricle as well as

for any angiographic study that requires injection of larger volumes of contrast material.

JUDKINS CORONARY CATHETERS AND JUDKINS TECHNIQUE

The left and right Judkins catheters are preshaped to selectively engage the left and right coronary ostia. These catheters have tapered end holes for contrast injection. Both catheter types have a primary and a secondary curve to negotiate the aortic root and facilitate engagement of the left and right coronary, respectively. The secondary bend represents the major difference between the left and right Judkins catheters. In the case of the left, the bend approaches 180°, whereas it is about 30° with the right.^[29] Both catheters are available in various sizes ranging from 3.5 to 8 cm. Larger catheters can be custom ordered. In the left Judkins catheter, these sizes represent the distance between the primary and secondary curves (Figs. 8 and 9).

Because the right Judkins catheter has a less prominent primary curve, it is sized by the length of its secondary curve (Fig. 10). A 4.0-cm catheter will fit most adults. However, in patients with dilated or enlarged aortic roots, a larger-sized catheter may be required. The total length of a catheter from tip to hub is 100 cm. Custom-designed catheters of longer length are required in very tall patients or patients with extreme tortuosity of the aorta.

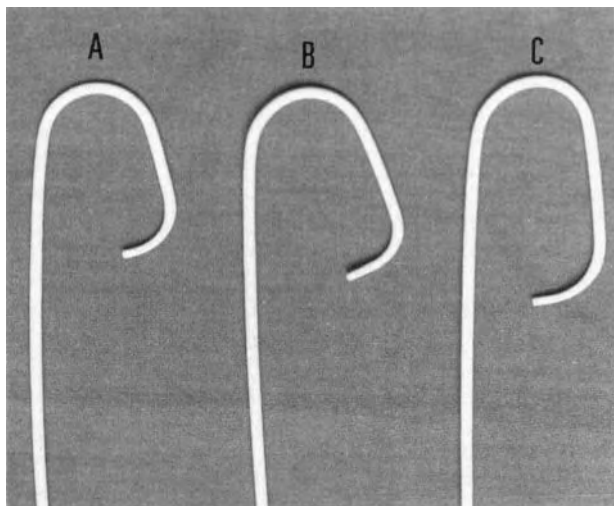


Fig. 8 Photograph of Judkins left coronary catheters. A:JL3.5; B:JL4; C:JL5. These numbers designate the length in centimeters between the primary and secondary curves. (From *Diagnostic and Therapeutic Cardiac Catheterization*, 2nd ed.; edited by C. J. Pepine, J. A. Hill, and C. R. Lambert; published by Williams and Wilkins, Baltimore, 1994; p. 136; Fig. 9.32.)

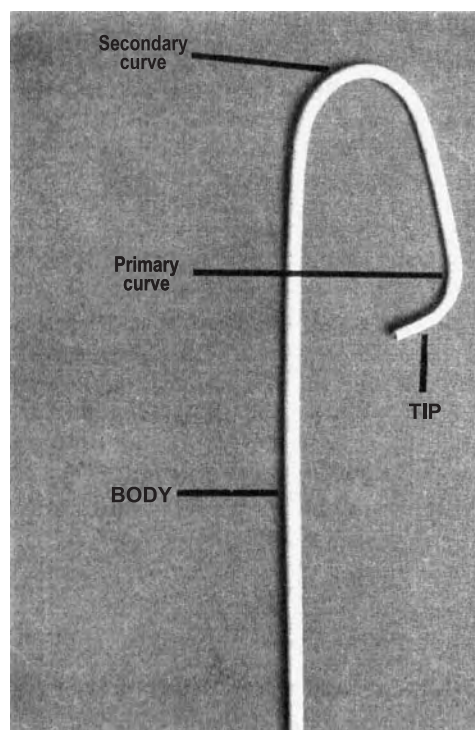


Fig. 9 Photograph of a JL4 coronary catheter demonstrating the different sections of the catheter. (From *Diagnostic and Therapeutic Cardiac Catheterization*, 2nd ed.; edited by C. J. Pepine, J. A. Hill, and C. R. Lambert; published by Williams and Wilkins, Baltimore 1994; p. 126; Fig. 9.18.)

The left Judkins catheter is easily directed into the left coronary artery with very little manipulation. The catheter tip is slowly advanced under fluoroscopic guidance as it follows the ascending aorta border. The unique double-curve design allows the catheter tip to fall into the opening of the left coronary artery, often with an abrupt forward motion. Sometimes, a minimal amount of clockwise rotation improves alignment of the tip with the left main trunk (Fig. 11).

The technique for cannulation of the right coronary artery is different than that of the left and requires slightly more catheter manipulation.^[28,30] With the fluoroscopy unit placed in the left anterior oblique position, the catheter is advanced into the ascending aorta with the tip directed caudally toward the feet. The catheter is advanced into the right coronary cusp of the aortic valve. The catheter is then pulled back about 2–3 cm above the valve and rotated 45–90° clockwise simultaneously. This combined maneuver transmits the torque to the catheter tip and the right coronary ostium is slowly engaged (Fig. 12). Another technique is to engage the right coronary artery from above its opening. The catheter is advanced 2–4 cm above the aortic valve. While the catheter is rotated

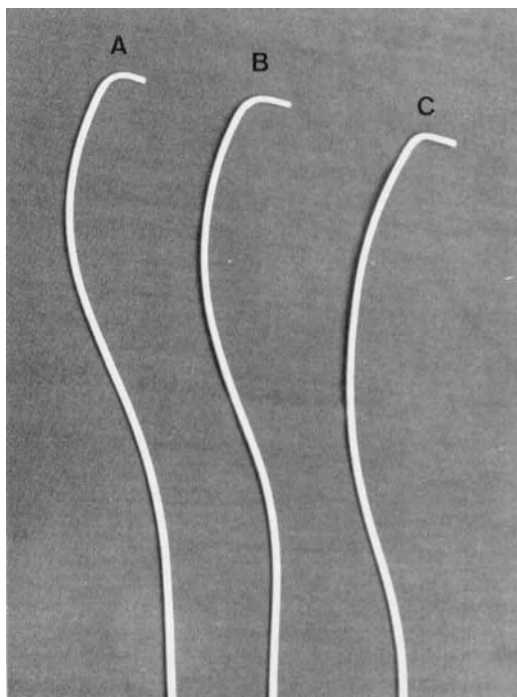


Fig. 10 Photograph of Judkins right coronary catheters. A:JR4; B:JR5; C:JR6. These letters designate the length of the secondary curve in centimeters. (From *Diagnostic and Therapeutic Cardiac Catheterization*, 2nd ed.; edited by C. J. Pepine, J. A. Hill, and C. R. Lambert; published by Williams and Wilkins, Baltimore, 1994; p. 136; Fig. 9.33.)

45–90°, it is advanced downward approximately 1–2 cm and the right coronary can be slowly engaged.

All catheters used during cardiac catheterization are attached to fluid-filled manometer systems for the measurement of arterial pressures. Close monitoring of the arterial pressure waveform during selective engagement of the left and right coronary artery is crucial because the first indication of important atherosclerotic disease involving the origin of these vessels could



Fig. 11 Diagram illustrating the selective cannulation of the left coronary artery using the Judkins left catheter. (From *Cardiac Catheterization and Angiography*, 2nd ed.; edited by W. Grossman; published by Lea and Febiger, Philadelphia, 1980; p. 162; Fig. 13-9.)



Fig. 12 Illustration of selective cannulation of the right coronary artery using the Judkins right catheter. (From *Cardiac Catheterization and Angiography*, 2nd ed.; edited by W. Grossman; published by Lea and Febiger, Philadelphia, 1980; p. 163; Figs. 13–10.)

be changes in the pressure waveforms as the catheter enters the vessel.

AMPLATZ CORONARY CATHETERS

Occasionally the Judkins catheters are unsuitable for engaging the coronary ostia. This is due to variations in aortic root configurations, aberrations in the left and right coronary artery origins that can displace them superoposteriorly, or anomalous origins of the coronary arteries. In these select instances, an operator may preferably change to Amplatz-type catheters that allow more rotational maneuvering and controlled probing of the ascending aorta than do Judkins catheters.

The distal end of Amplatz catheters has a preshaped half-circular curve with a tapered tip angled perpendicular to this curve. The left Amplatz catheter is available in various sizes ranging from AL0.75 to AL4 based on progressively larger curve diameters (Fig. 13).^[31]

The right Amplatz catheter has a similar—although smaller—distal end, and is available in two sizes, AR1 and AR2. As with the left Amplatz catheters, these numbers refer to progressively larger diameters of the half-circular curves. An AL1 or 2 and AR1 will fit most adults.

The left Amplatz catheter has a unique method of engaging and disengaging the left coronary artery. The catheter is advanced around the arch with its tip directed toward the left coronary ostium and pointing downward. The tip usually comes to rest below the ostium in the sinus of Valsalva. Further advancement causes the catheter tip to ride up the aortic wall and engage the left coronary ostium. At this point, slight retraction causes deeper engagement of the left main coronary artery (Fig. 14). To disengage the catheter, the operator has to paradoxically advance the catheter farther instead of withdrawing it. This backs the tip upward and out of the left main ostium. When

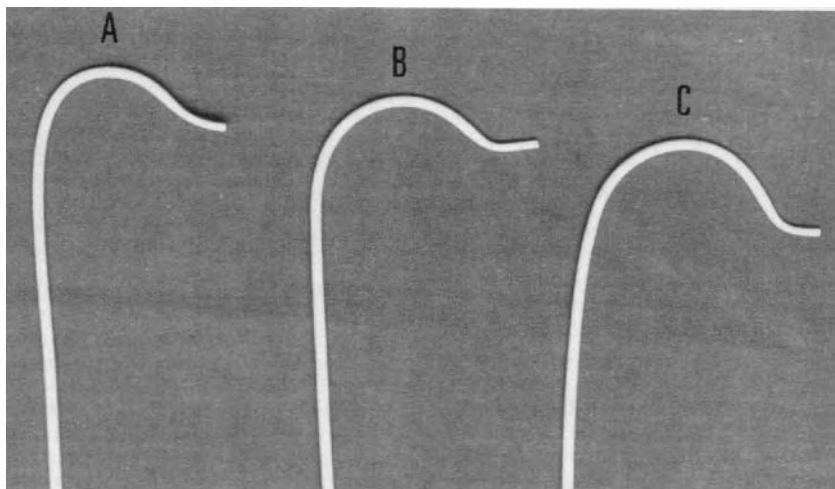


Fig. 13 Photograph of different sized Amplatz left coronary catheters. A:AL1; B:AL2; C:AL3. The letters refer to progressively larger diameters of the half circular curve. (From *Diagnostic and Therapeutic Cardiac Catheterization*, 2nd ed.; edited by C. J. Pepine, J. A. Hill, and C. R. Lambert; published by Williams and Wilkins, Baltimore, 1994; p. 137; Fig. 9.34.)



positioned properly, withdrawal (i.e., pulling) of the left Amplatz catheter during attempts to disengage the left main can actually cause the catheter to advance deeper into the vessel. The left Amplatz catheter is particularly useful in situations where the left anterior descending artery and the left circumflex artery have separate ostia, and subselective engagement of each is required to acquire diagnostic quality angiograms. This catheter can also provide greater stability and support for the passage of various interventional tools during the treatment of complex coronary stenoses. The left Amplatz catheter is more technically challenging to use and has higher incidence of coronary artery dissections compared to the Judkins catheter.

The right Amplatz catheter is engaged similarly to the right Judkins catheter by rotating clockwise 45° to 90° and simultaneously withdrawing it 1 to 2 cm (Fig. 15). In order to disengage, it may be pulled, advanced, or rotated.

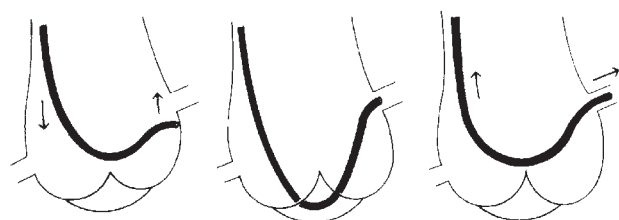


Fig. 14 Illustration demonstrating the catheter maneuvers necessary to engage the left coronary with a left Amplatz catheter. (From *Grossman's Cardiac Catheterization, Angiography, and Intervention*, 6th ed.; edited by D. S. Baim and W. Grossman; published by Lippincott, Williams, and Wilkins, Baltimore, 2000; p. 219; Fig. 11.5.)

BYPASS GRAFT CATHETERIZATION

During prior coronary artery bypass surgery, reversed saphenous veins and the left and right internal mammary arteries may have been used as bypass conduits in a patient. It is helpful to know their presence and location during subsequent cardiac catheterizations. Vein bypass grafts are anastomosed to the anterior wall of the ascending aorta. The right coronary artery bypass graft is usually slightly anterosuperior to the right coronary artery origin. Left anterior descending artery and diagonal grafts are anastomosed farther to the left and higher, whereas the obtuse marginal grafts

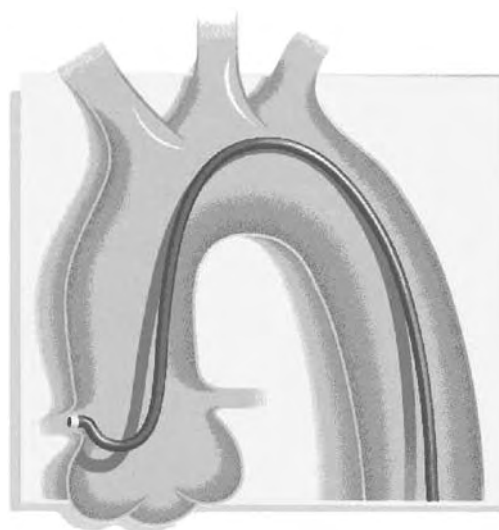


Fig. 15 Graphic demonstrating engagement of the right coronary artery with a right Amplatz catheter. (From "The Guiding Catheter Shape Selector"; A Cordis Corporation (Miami, FL) Educational Resource.)



Fig. 16 Illustration showing the distal shape of the right coronary bypass catheter. (From “The Guiding Catheter Shape Selector”; A Cordis Corporation (Miami, FL) Educational Resource.)

are the highest and farthest left.^[28] Occasionally, an aortic root injection is necessary to locate the origin or to document occlusion of the aortic anastomosis of a vein graft. The left internal mammary artery originates anteriorly from the caudal wall of the subclavian artery distal to the vertebral artery origin. The right internal mammary artery originates from a similar location in the right brachiocephalic trunk.

The right Judkins catheter is versatile and is the catheter of choice for initial attempts at catheterization of the saphenous vein grafts. It can also be used to cannulate internal mammary bypass grafts, but, optimal engagement of the mammary artery is frequently not possible due to the shape of this catheter. For angiography of saphenous vein grafts with the right Judkins catheter, counterclockwise rotation usually works best to initially engage grafts anastomosed to either the right or the left coronary. After initial engagement, further gentle torque in either the clockwise or counterclockwise direction is needed to obtain a fully coaxial position of the catheter with the graft.

For internal mammary artery angiography, the right Judkins catheter must first be manipulated into either the subclavian artery on the left or the innominate artery on the right. Typically, the catheter is initially placed beyond the origin of the mammary artery and then slowly withdrawn. Once the catheter reaches the level of the thyrocervical trunk, slight counterclockwise rotation will often turn the catheter tip inferiorly into the origin of the internal mammary artery.

The Amplatz catheters are also suited for saphenous vein graft angiography due to the rotational maneuverability that allows controlled probing of the ascending aorta as described earlier.

In cases where cannulation of coronary vein bypass grafts is more challenging, the right and left coronary bypass catheters may have to be used. Similarly, use of an internal mammary graft catheter is often necessary in patients with a very vertical origin of the mammary artery.

The right and left bypass catheters are shaped similar to the right Judkins catheter with minor modifications to facilitate engagement of the vein graft ostia. The right coronary bypass catheter has a wider, more open primary curve that points the catheter tip inferiorly and facilitates the cannulation of the vertically oriented coronary artery vein graft (Figs. 16 and 17). As with the right Judkins catheter, counterclockwise rotation of the right coronary bypass catheter is required to engage the graft. Multipurpose catheters can also frequently be used successfully in situations where the right coronary graft is vertically oriented.

The left coronary bypass catheter has a smaller and sharper secondary curve when compared to the right Judkins coronary catheter. This facilitates the cannulation of the left coronary vein grafts that originate with a horizontal and upward takeoff from the aorta (Figs. 18 and 19). When a left coronary bypass catheter must be used to cannulate bypass grafts to the anterior descending or circumflex arteries, clockwise rotation of

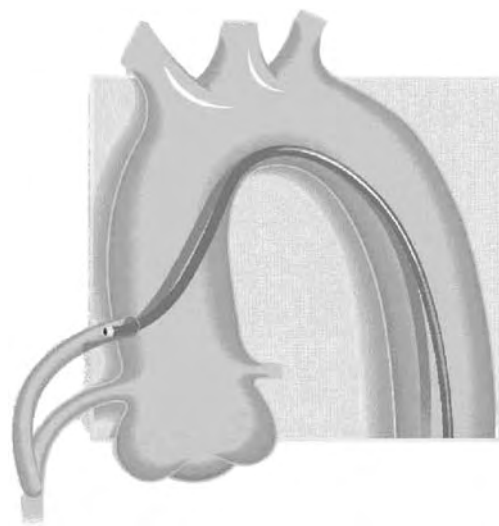


Fig. 17 Graphic representation of the engagement of a saphenous vein bypass graft to the right coronary artery using a right coronary bypass catheter. (From “The Guiding Catheter Shape Selector”; A Cordis Corporation (Miami, FL) Educational Resource.)



Fig. 18 Illustration showing the distal shape of the left coronary bypass catheter. (From “The Guiding Catheter Shape Selector”; A Cordis Corporation (Miami, FL) Educational Resource.)

the catheter is usually required for successful engagement.

The internal mammary artery graft catheters have a hook-shaped tip that makes the engagement of the mammary arteries comparatively easier. This shape allows the catheter to be manipulated into a more vertical position with caudal orientation than a standard right Judkins catheter can be (Fig. 20).

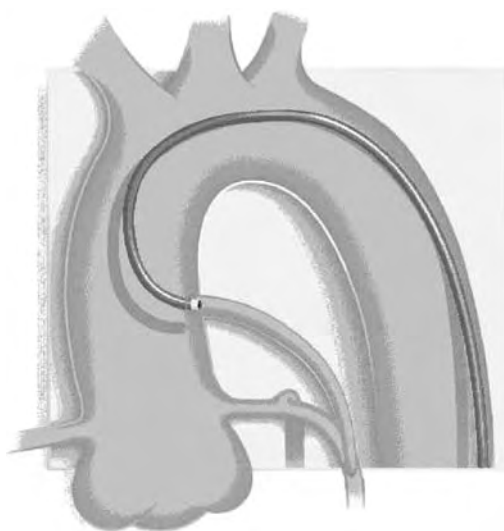


Fig. 19 Graphic representation of the engagement of a saphenous vein bypass graft to the left coronary artery using a left coronary bypass catheter. (From “The Guiding Catheter Shape Selector”; A Cordis Corporation (Miami, FL) Educational Resource.)

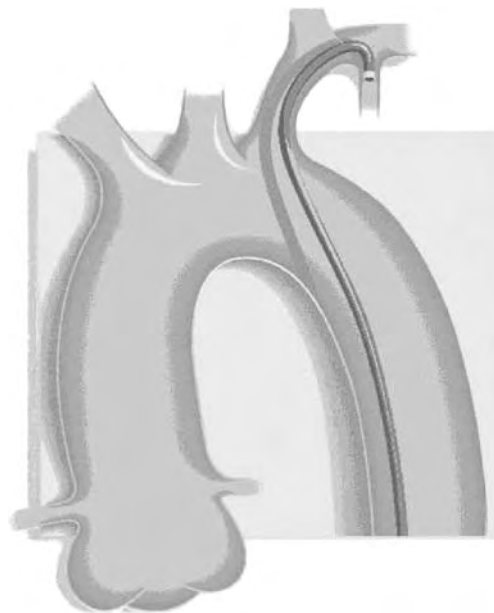


Fig. 20 Graphic representation of the engagement of the left internal mammary artery using an internal mammary artery catheter. (From “The Guiding Catheter Shape Selector”; A Cordis Corporation (Miami, FL) Educational Resource.)

PIGTAIL CATHETERS

For left ventriculography, a specialized pigtail catheter developed by Judkins is used. It has a preshaped tapered tip that makes a circle 1 cm in diameter with five to 12 side holes located on the straight portion of the catheter just proximal to the curve. Side holes in this section of the catheter help to create a cloud of contrast medium and improve stability during injection. Several methods exist to cross the aortic valve and enter the left ventricle. Most frequently, the pigtail catheter is advanced onto the aortic valve forming a gentle loop (Fig. 21). The catheter is then slowly withdrawn and frequently, as the loop uncoils, the catheter will prolapse across the valve and into the ventricle. When the aortic root is dilated, this maneuver is facilitated by using a guidewire to stiffen the pigtail catheter. In cases of aortic valve disease such as aortic stenosis, the pigtail catheter alone will not cross the stenotic valve. In these cases, the valve must be crossed with a guidewire (Fig. 21) and then the pigtail catheter is tracked over the guidewire and into the left ventricle. When using a guidewire to cross the aortic valve, great care must be taken to avoid inadvertent trauma to the left main coronary artery.

Upon entry into the left ventricle, the catheter is usually manipulated to find a position that does not cause frequent premature ventricular contractions. For most patients, a straight pigtail catheter sits most

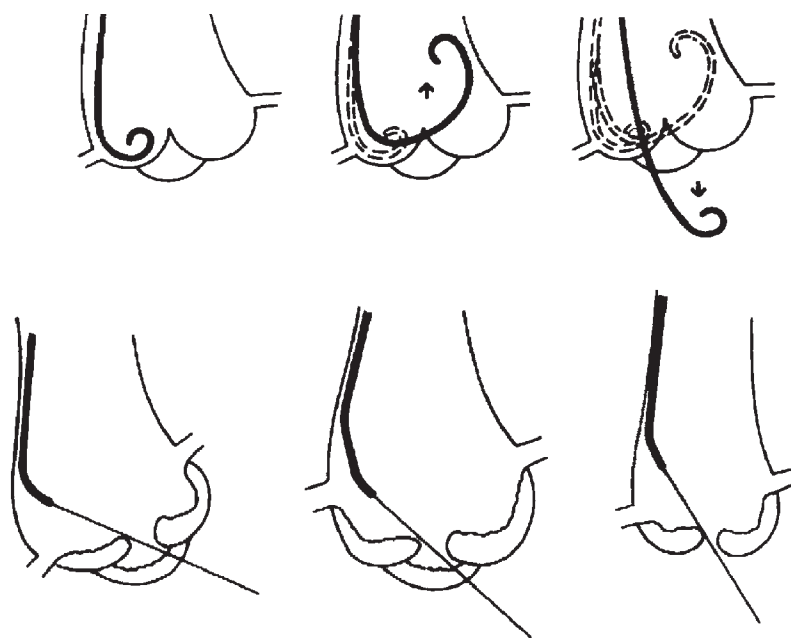


Fig. 21 Graphic representation of the techniques used to cross the aortic valve with a pigtail catheter. This can be done either with (bottom panel) or without (top panel) a guidewire. See text for details. (From *Grossman's Cardiac Catheterization, Angiography, and Intervention*, 6th ed.; edited by D. S. Baim and W. Grossman; published by Lippincott, Williams, and Wilkins, Baltimore, 2000; p. 86; Fig. 4.11.)

quietly in the left ventricle. However, for horizontally oriented hearts, a 145° angled pigtail catheter may be preferable. The multiple side holes provide simultaneous exit routes for contrast media, thereby reducing the magnitude of catheter recoil and stabilizing the catheter. This design also eliminates endocardial staining that can be seen with side-hole catheters and further reduces the incidence of ectopy.

RIGHT HEART CATHETERIZATION

Right heart catheterization is often performed simultaneously at the time of coronary angiography. Right heart catheterization allows for hemodynamic measurements in the right atrium, right ventricle, and the pulmonary arteries. Blood oxygen saturations from these locations can also be assessed. The pulmonary capillary wedge pressure is an indirect measurement of left ventricular filling pressure. Measurements can also be utilized to obtain cardiac output to calculate valve orifice areas. In addition to the femoral vein approach, the internal jugular and subclavian veins can also be used for right heart catheterization, especially in an intensive care unit setting.

The balloon-tipped flotation catheter, originally designed by Drs. H. J. C. Swan and W. Ganz, is the most commonly used catheter for right heart pressure measurements.^[15] The balloon tip is inflated with room air or with carbon dioxide if entry into the arterial circulation is a possibility or concern. It allows the catheter to float through the right side of the heart with ease and safety in the majority of cases. The proximal end usually has four ports. Two of these open at the

level of the right atrium and pulmonary artery and are used for hemodynamic measurements, instillation of fluids, and aspiration. The third port inflates the balloon at the catheter tip, and the fourth port is used for cardiac output measurements using the thermodilution technique.

CONCLUSIONS

We have attempted to provide the reader with a general overview of techniques and equipment most commonly used during the procedure of cardiac catheterization in adult patients. Our discussion here is not comprehensive, particularly regarding the many other catheters that are available for special situations. The interested reader is referred to one of several excellent textbooks devoted entirely to the field of cardiac catheterization and vascular intervention.^[28-31]

ARTICLES OF FURTHER INTEREST

Distal Protection Devices; Heparin-Binding to Improve Biocompatibility; Polymers; Surface Coatings

REFERENCES

1. Milne, J.S. *Surgical Instruments in Greek and Roman Times*; Clarendon Press: Oxford, 1907.
2. Mueller, R.L.; Sanborn, T.A. The history of interventional cardiology: Cardiac catheterization, angioplasty, and related interventions. *Am. Heart J.* **1995**, *129*, 146.

3. Cournand, A. Cardiac catheterization: Development of the technique, its contributions to experimental medicine, and its initial application in man. *Acta Med. Scand., Suppl.* **1975**, 579, 1.
4. Forssmann, W. *Experiments on Myself: Memoirs of a Surgeon in Germany*; St. Martin's Press: New York, 1974.
5. Cournand, A.F.; Ranges, H.S. Catheterization of the right auricle in man. *Proc. Soc. Exp. Biol. Med.* **1941**, 45, 462.
6. Cournand, A.F.; Riley, R.L.; Breed, E.S.; Baldwin, E.F.; Richards, D.W. Measurement of cardiac output in man using the technique of catheterization of the right auricle. *J. Clin. Invest.* **1945**, 24, 106.
7. Brannon, E.S.; Weens, H.S.; Warren, J.V. Atrial septal defect: Study of hemodynamics by the technique of right heart catheterization. *Am. J. Med. Sci.* **1945**, 210, 480.
8. Hellem, H.K.; Haynes, F.W.; Dexter, L. Pulmonary "capillary" pressure in man. *J. Appl. Physiol.* **1949**, 2, 24.
9. Zimmerman, H.A.; Scot, R.W.; Becker, N.D. Catheterization of the left side of the heart in man. *Circulation* **1950**, 1, 357.
10. Seldinger, S.J. Catheter replacement of the needle in percutaneous arteriography. *Acta Radiol.* **1953**, 39, 368.
11. Sones, F.M., Jr.; Shirey, E.K.; Proudfit, W.L.; Wescott, R.N. Cine coronary arteriography. *Circulation* **1959**, 20, 773.
12. Sones, F.M., Jr.; Shirey, E.K. Cine coronary arteriography. *Mod. Concepts Cardiovasc. Dis.* **1962**, 31, 735.
13. Kohchi, K.; Takebayashi, S.; Block, P.C.; Hiroki, T.; Nobuyoshi, M. Arterial changes after percutaneous transluminal coronary angioplasty: Results at autopsy. *J. Am. Coll. Cardiol.* **1987**, 10, 592-599.
14. Dotter, C.T.; Judkins, M.P. Transluminal treatment of arteriosclerotic obstruction: Description of a new technique and preliminary report of its application. *Circulation* **1964**, 30, 654.
15. Swan, H.J.C.; Ganz, W.; Forrester, J.; et al. Catheterization of the heart in man with use of a flow directed balloon-tipped catheter. *N. Engl. J. Med.* **1970**, 283, 447.
16. Grüntzig, A.; et al. Coronary transluminal angioplasty. *Circulation* **1977**, 56 (Suppl II), 319. (abst).
17. Grüntzig, A.; Senning, A.; Siegenthaler, W.E. Non-operative dilatation of coronary artery stenosis; percutaneous transluminal coronary angioplasty. *N. Engl. J. Med.* **1979**, 301, 61.
18. Grines, C.L.; Browne, K.F.; Marco, J.; et al. A comparison of immediate angioplasty with thrombolytic therapy for acute myocardial infarction. *N. Engl. J. Med.* **1998**, 328, 673.
19. Tiefenbrunn, A.J.; Chandra, N.C.; French, W.J.; et al. Clinical experience with primary percutaneous transluminal coronary angioplasty compared with alteplase (recombinant tissue-type plasminogen activator) in patients with acute myocardial infarction: A report from the Second National Registry of Myocardial Infarction (NORMI-2). *J. Am. Coll. Cardiol.* **1998**, 31, 1240.
20. Botnar, R.M.; Stuber, M.; Danias, P.G.; et al. Improved coronary artery definition with T2-weighted, free-breathing, three dimensional coronary MRA. *Circulation* **1999**, 99, 3139.
21. Achenbach, S.; Moshage, W.; Ropers, D.; et al. Value of electron-beam computed tomography for the noninvasive detection of high-grade coronary-artery stenosis and occlusions. *N. Engl. J. Med.* **1998**, 339, 1964.
22. Noto, T.J.; Johnson, L.W.; Krone, R.; et al. Cardiac catheterization 1990: A report of the registry of the Society for Cardiac Angiography and Interventions. *Cathet. Cardiovasc. Diagn.* **1991**, 24, 75.
23. Katzen, B.T. Interventional radiology: Past, present and future. *Med. Mundi* **2002**, 40 (3), 151-155.
24. Payne, M. Charles Theodore Dotter: The father of intervention. *Texas Heart Inst. J.* **2001**, 28, 28-38.
25. Anonymous Portraits in radiology: T. Charles, Dotter, MD. *Appl. Radiol.* **1981**, 10 (28), 116. (Jan-Feb).
26. Rosch, J.; Abrams, H.L.; Cook, W. Memorials: Charles Theodore Dotter, 1920-1985. *Am. J. Roentgenol.* **1985**, 144, 1321-1323.
27. Physics, Techniques and Procedures. In *The Encyclopedia of Medical Imaging*; Pettersson, H., Ed.; Isis Medical Medica, 1998; Vol 1. s.v. "Catheter".
28. *The Cardiac Catheterization Handbook*, 3rd Ed.; Kern, M.J., Ed.; Mosby, Inc.: St. Louis, 1999.
29. *Invasive Cardiology: Principles and Techniques*; Bashore, T.M., Ed.; B.C. Decker: Toronto, 1990.
30. *Grossman's Cardiac Catheterization, Angiography, and Intervention*, 6th Ed.; Baim, D.S., Grossman, W., Eds.; Lippincott, Williams, and Wilkins: Philadelphia, 2000.
31. *Diagnostic and Therapeutic Cardiac Catheterization*, 3rd Ed.; Pepine, C.J., Ed.; Williams and Wilkins: Baltimore, 1998.



Cardiac Elastography, Full-Field Development

Tomy Varghese

Departments of Medical Physics and Biomedical Engineering, The University of Wisconsin–Madison, Madison, Wisconsin, U.S.A.

Q. Chen

Department of Medical Physics, The University of Wisconsin–Madison, Madison, Wisconsin, U.S.A.

P. Rahko

Section of Cardiovascular Medicine, Department of Medicine, The University of Wisconsin–Madison, Madison, Wisconsin, U.S.A.

James A. Zagzebski

Department of Medical Physics, The University of Wisconsin–Madison, Madison, Wisconsin, U.S.A.

INTRODUCTION

Echocardiography is used routinely for the assessment of regional myocardial function, with stress echocardiography used to identify ischemia in patients with coronary artery disease (CAD). However, clinical diagnosis based on visually assessed wall motion scoring from echocardiography images is semi-quantitative, operator dependent, and heavily weighted by operator experience. Tissue Doppler imaging (TDI), a newer technique used to assess myocardial muscle velocity, provides quantitative evaluation of systolic and diastolic motion velocities and intervals and derived parameters, such as the strain and strain-rate. As these methods rely on narrow-band Doppler phase-shift analysis, they inherit the disadvantages associated with Doppler, such as poor axial resolution, aliasing, and angle dependence of velocity estimates. Estimation of tissue displacements using cross-correlation, as in elastography, has been repeatedly shown to overcome many of the limitations associated with Doppler phase-shift methods.

In this entry we illustrate the production of two-dimensional strain images for evaluating cardiac function. Loops of in-phase/quadrature (IQ) data from a clinical scanner (GE Vingmed Vivid 5) were recorded during echocardiography examinations. Data were transferred to a computer for off-line processing. The IQ data were converted to radiofrequency (RF) echo signals, and elastograms depicting myocardial strain were obtained from successive RF echo signal frames. Full-field and scan-converted cardiac strain images were obtained along all the standard echocardiographic views. Cardiac strain image loops for normal subjects were compared with those obtained from two patients with CAD. Initial observations suggest differences in temporal and

spatial strain distribution between normal and diseased ventricles.

Cardiac elastography may provide quantitative and translation-independent measures of myocardial strain during the cardiac cycle, thereby providing a two-dimensional method for characterizing regional myocardial function.

BACKGROUND

Coronary artery disease is the leading cause of morbidity and mortality in the United States. Despite advances in prevention and treatment of this disorder, there remains a large patient population who are difficult to diagnose noninvasively, yet require percutaneous or surgical revascularization.^[1] Myocardial ischemia is generally associated with impaired regional myocardial function. Current clinical assessment uses analysis of myocardial wall motion abnormalities using echocardiography, nuclear imaging, or magnetic resonance imaging (MRI).^[1–4]

Echocardiography has been routinely used for the assessment of regional myocardial function as it provides real-time information, is portable, and is readily available. During systole short-axis echo, images of the left ventricle (LV) show wall thickening in the radial direction and shortening in the circumferential direction, while in the long-axis view shortening is observed as the base moves toward the apex. Thickening and shortening of the wall muscle are very useful indicators of myocardial performance.^[5] This provides strong evidence that the measurement of the regional strain will provide direct information on the regional myocardial function. In conventional echocardiography, global LV function is estimated by the shortening or ejection fraction, while regional LV

function is evaluated using segmental endocardial excursion or wall thickening. Although endocardial excursion is the most commonly used parameter, it does not reliably differentiate actively contracting from passively drawn segments (scar tissue). Wall thickening also may be difficult to evaluate, as the epicardium is often not clearly seen.^[6]

Use of myocardial strain for cardiac diagnosis was first described by Mirsky and Parmley.^[7] Regional strain has been measured invasively in experimental animal models using optical markers, sonomicrometry, and radio-opaque markers.^[8–10] Invasive measurements have shown that strain is inhomogeneous even in healthy animals, and that experimental ischemia induces increasing regional inhomogeneity.^[8,10,11] Changes in myocardial stiffness with ischemic changes and infarction have also been reported in the literature.^[12]

Noninvasive measurement of myocardial strain in human subjects was initially performed using MRI.^[13] Magnetic resonance tagging and phase-contrast MRI were used to obtain noninvasive and accurate measurements of myocardial motion.^[13,14] Estimation of the strain information required the acquisition of long-axis images during MR tagging, or velocity encoding in the slice direction for phase-contrast MRI.^[3,4] The main limitations of strain measurement using MRI are the low sampling rates and long examination times and fundamental limitations of the tag spacing of around 5–7 mm.

Tissue Doppler imaging, also called tissue velocity imaging, estimates local tissue velocities and tracks heart wall motion using Doppler techniques.^[15,16] One way of displaying TDI information is by color-coding and superimposing TDI estimates on a B-scan image, similar to color-flow imaging. The original method used pulsed Doppler and indeed that is now used on most routine exams at the lateral or septal base. However, the examination time required to survey 16 segments for TDI, one segment at a time, is not practical. Tissue Doppler imaging has been tried and found to be impractical or unworkable for the analysis of wall motion; however, it is used extensively for diastolic and systolic global function assessment. Tissue Doppler imaging also does not differentiate between active contraction and simple rotation or translation of the heart wall, nor does it differentiate passively following tissue from active contraction. Fleming et al. used the spatial gradient of the TDI velocity to measure relative changes in wall thickness, or strain-rate, to overcome this problem.^[17] However, as velocities decrease toward the apex, the signal-to-noise ratio decreases, limiting the usefulness in the apical part of the ventricle. Strain-rate plots using cross-correlation also have been reported, but only on M-mode data.^[18]

Recent studies evaluating myocardial and sub-endocardial function using strain-rate imaging compared to an invasive global evaluation of LV function conclude that despite the angle dependence of strain-rate imaging it is superior to MRI with respect to spatial and temporal resolution.^[19,20] The high spatial resolution with strain-rate imaging enabled the evaluation of segmental myocardial layers, differentiating subendocardium, midmyocardium, and epicardium.^[19] Hashimoto et al. obtained a significant difference in strain-rate values between the three layers identified above in the baseline condition and under hemodynamic alterations.^[19] The authors mention that the results of their study agree with recent tissue tagging studies in normal subjects showing that endocardial deformation is greatest in the longitudinal direction and verifying the endoepicardial gradient in normal LVs.^[21]

Strain-rate imaging enables the clinical evaluation of the LV layers that were previously possible only using invasive techniques, such as sonomicrometry.^[20] Noninvasive discrimination of subendocardium and midwall, however, is highly desirable as this analysis yields clinically significant findings.^[22] However, tethering effects of adjacent myocardium and translational heart motion can cause misinterpretation of regional wall motion with strain-rate imaging.^[20]

We are developing full-field cardiac elastography and evaluating its role in displaying myocardial strain and for detecting ischemia and infarction. In cardiac elastography, strain is imaged during the contraction and relaxation of the heart muscle. This differs from conventional elastography or stiffness-imaging techniques, which measure strain from externally applied compressions.^[23–25] However, virtually all the knowledge derived from the practice of external compression elastography can be applied for cardiac elastography. Recent improvements in the performance of cardiac ultrasound systems provide an opportunity for the application of elastography for estimations of myocardial strain.^[26–28] Fast acquisition of RF data frames is required of the ultrasound system to significantly minimize motion artifacts accompanying cardiac translation.

It has been estimated that 50% or more of all echocardiography studies performed are directed toward assessment of LV function. Our work addresses an important issue related to interpretation of these studies. Although scoring of wall motion is useful, particularly in skilled hands, clinical personnel seek more objective assessments of LV function than those provided in conventional echocardiography studies. The proposed cardiac elastography modality could be a powerful tool to extend the capabilities of echocardiography enabling differentiation between actively and passively moving muscle tissues.



MATERIALS AND METHOD

Method

A GE Vingmed Vivid FiVe ultrasound system (GE-Vingmed, Horten, Norway) was used to obtain two-dimensional gray scale and raw echo data frames for cardiac elastography. Data acquisition was performed in the University of Wisconsin–Madison Hospital echocardiography laboratory. A 2.5 MHz phased array transducer was used to collect images from the standard apical and parasternal long-axis views. Loops of IQ data collected over one heart cycle were transferred via a magneto-optical disk to a personal computer for off-line processing.

The IQ data were first converted back to RF echo signals sampled at 20 MHz using Echomat, a custom software program implemented in Matlab (Mathworks, Inc., Natick, Massachusetts, U.S.A.), and provided by GE-Vingmed. Cardiac elastograms were then generated by comparing corresponding beam-line data from consecutive RF frames using cross-correlation to detect tissue displacements.^[23] The localized myocardial strain was estimated from the axial gradient of the cardiac muscle displacement. A window length of 3 mm with a 75% overlap between data segments was used to generate the cardiac elastograms. The frame rate of 50 frames/sec for the GE-Vingmed system FiVe enabled tracking of small compressions of the heart muscle and limited motion artifacts. Batch programs developed using Matlab software were used to generate both “directional” and “absolute” cardiac elastograms. Directional elastograms are obtained by incorporating both the sign and the magnitude of the estimated strain in the elastogram, while the absolute elastograms depict only the strain magnitude.

Data were acquired on patients and normal healthy male volunteers. The patient results depicted in this paper show results from one patient diagnosed with CAD, and the other with dilated cardiomyopathy. Cardiac elastograms for both patients and normal subjects were evaluated by two clinical observers using qualitative visual analysis (QVA) of the strain images.

RESULTS

Full-field and scan-converted cardiac elastograms obtained in all standard echocardiographic views are presented in this paper. The cardiac elastograms shown present a one-to-one correspondence with the gray-scale images, allowing for easy comparison of structures of interest. Cardiac elastograms in the parasternal long-axis view; short-axis view; and two-, three-, and four-chamber apical views are shown.

Fig. 1 presents a gray-scale image and cardiac elastograms in the parasternal long-axis view at end-systole (ES). The ultrasound gray-scale image is illustrated in Fig. 1(A); a vector or directional strain elastogram, i.e., an elastogram that incorporates both positive (red) and negative (blue) strain information, as shown in Fig. 1(B); and an absolute strain elastogram, as shown in Fig. 1(C). Tissue strains in the anterioseptal (AS) wall, the posterior medial papillary muscle (PM), and the posterior wall (PW) are depicted in Fig. 1. Near ES, the PM is actively contracting, which is clearly observed in Fig. 1(B) as a red region, with muscles surrounding the PM undergoing relaxation (blue regions). The brightest spot is at the PM, which also yields the highest percent strain (see color bar at the right-hand side). Figs. 1(D) and 1(E) present one-dimensional axial strain profiles along the fifth beam line (along the PM) used to form the cardiac elastogram. Note the variation in this axial-strain profile for both absolute strain (D) and positive and negative strains (E). Peak strains are measured at the level of the PM, where the 4 cm depth corresponds to the AS, the 5–8 cm depth the PM, and the 10–12 cm depth the PW.

We use a frame rate of 50 frames/sec and capture strain information over 20 msec snapshots of the cardiac cycle. As we provide a snapshot of the cardiac deformation over 20 msec in each cardiac elastogram, it is extremely unlikely that the images would contain many artifacts. If artifacts were present, we would observe large noise errors (abrupt strain jumps in the elastogram similar to salt-and-pepper type noise artifacts) in the elastograms.^[23,25] Such artifacts clearly are not seen in the images presented, and were not seen during the image sequences processed.

Elastograms for the same patient at ED are illustrated in Fig. 2. From the elastogram in Fig. 2(B) the PM is now undergoing relaxation, as shown by the blue regions at the level of the PM at ED. One can also observe an increase in the volume at ED, along with the relaxation of the AS and PW. The peak strain occurred in both patients at ES at the PM (see Figs. 1 and 2).

Qualitative visual analysis (QVA) by the two clinical observers concluded that peak strain in normals occurred at the PM and the PW in ES and was more evenly distributed in ED. The patients, however, exhibited peak strain at the PM in ES. One of the primary differences observed from ES to ED is that the strains near ED are smaller than at the ES. Typically, peak strains at ED are much lower than at ES. This is consistent with the fact that during systole the heart is actively contracting, while during diastole the heart chamber is in the passive-refilling phase. This pattern is reflected on the values of the color bar that shows the range of strains depicted on the elastogram.

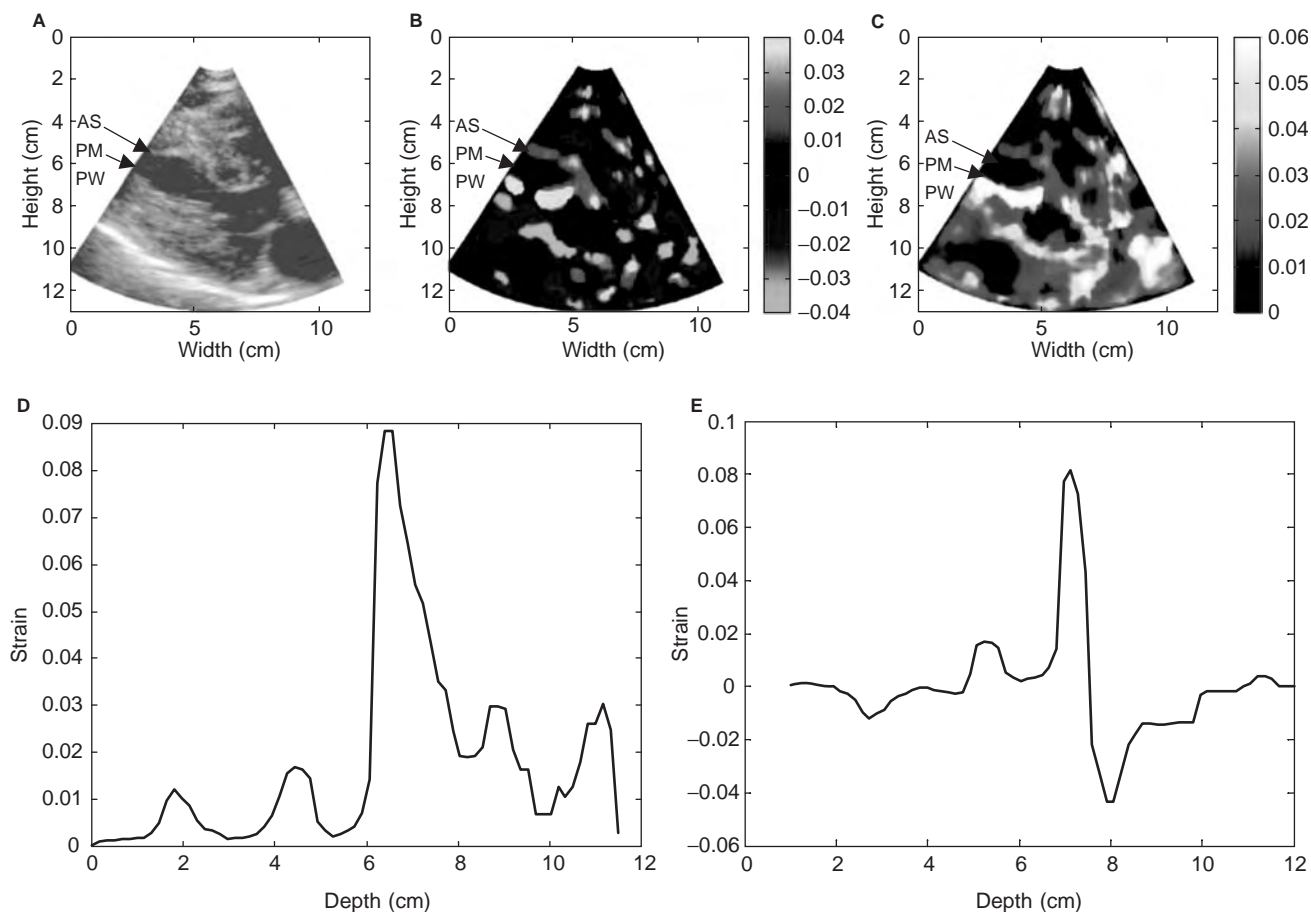


Fig. 1 Parasternal long-axis view, depicting (A) two-dimensional B-mode echocardiogram at end systole along with two-dimensional cardiac (B) vector or directional strain (incorporating both positive and negative strain information) and (C) absolute strain elastograms. The brightest spot is at the papillary muscle (PM), which also yields the highest percent strain (see colorbar at the right hand side). The spatial distribution of the strain is demonstrated by plotting the 5th A-line along the papillary muscle in (D) for absolute strain and (E) for both positive and negative strains. Peak strains measured at the level of the papillary muscle, where the 4 cm depth corresponds to the anterior septum, the 5-8 cm depth the papillary muscle, and the 10-12 cm depth to the posterior wall.

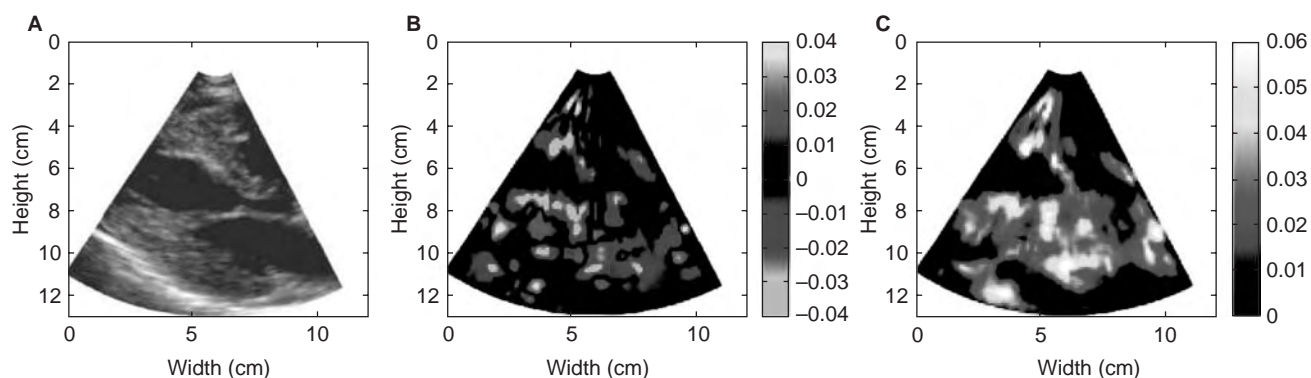


Fig. 2 (A) A two-dimensional gray-scale echocardiogram at ED is shown along with the respective elastograms. (B) Comparison of elastograms obtained from a parasternal long-axis view that show the absolute values of the strain. (C) Both positive and negative strains at ED.

Spatial and temporal interrogation along individual beam acquisition lines allowed quantitative assessment of anatomical and dynamic strain throughout the cardiac cycle. In the parasternal long-axis view of tissue strain in the AS wall, the posterior medial PM and the PW were measured at the ES and the ED. The peak strain occurred in the patients at ES at the PM (see Figs. 1 and 2). QVA by the two clinical observers revealed that peak strain in normals occurred at the PM and PW in ES and was more evenly distributed in ED. The patients, however, exhibited peak strain at the PM in ES, while it was more uniformly distributed in the normal volunteers.

As the heart is actively contracting near ES, and the local strains observed in the myocardium are significantly larger than at other times during the cardiac cycle, we examined the ability of cardiac elastography to depict myocardial strains near ES for all the standard echocardiographic imaging views. Examples are shown in Fig. 3. Short-axis echo images and elastograms of the patient with CAD are shown at ES (Fig. 3i). At ES, the chamber is at its smallest volume, which is clearly seen in the gray-scale echo images and in the cardiac elastograms. Elastograms also exhibit compressive strains in the walls of the ventricle at this time in the cardiac cycle.

Imaging along apical views provides the most challenging test for cardiac elastography, as motion of the heart from apex to base also occurs along with contraction and relaxation of heart muscle. However, this situation did not preclude the production of cardiac elastograms. Evidently, the 50 Hz frame rate resulted in only a small incremental motion between successive frames, while still providing adequate frame-to-frame heart muscle contraction. The motion seen in successive frames is roughly the total movement of the heart from apex to base divided by the number of frames over which this motion occurs. In our studies, this incremental motion was not large enough to decorrelate RF signals between successive frames, which would necessitate more robust and time-consuming algorithms for strain imaging. Fig. 3(ii), presents a two-chamber apical view of the heart in one of the patients with CAD. Maximal strain is seen in the anterior basal and midsegments.

In a similar manner, cardiac elastograms obtained using a three-chamber apical view of a normal subject are illustrated in Fig. 3(iii). Maximal strain is seen in the AS basal and midsegment. Figs. 3(iv) and 3(v) present the four-chamber apical view of the heart for the patient diagnosed with dilated cardiomyopathy and for a normal subject, respectively. Inhomogeneous strain distributions are noted in the cardiomyopathy patient and the patient with CAD, while the strain was more homogeneously distributed in the normal subject.

DISCUSSION

Cardiac elastography may provide improvements over observer-dependent judgments of myocardial contraction and relaxation estimated from conventional echocardiography for assessment of myocardial function. This new modality may also provide quantitative indices of regional myocardial function. We have presented cardiac elastograms obtained in the standard echocardiographic imaging views in this paper. Two-dimensional cardiac elastograms obtained using acquisition frame rates of 50 frames/sec appear to provide qualitative and quantitative strain imaging without translation artifacts.^[26,27] This two-dimensional strain imaging technique could be incorporated into future clinical scanners. However, much work needs to be performed to develop and evaluate this imaging modality.

Using an image acquisition rate of 50 frames/sec, we obtain unsaturated measured strains between frames. In addition to obtaining absolute estimates of strain, cardiac elastography can provide the direction of strain, i.e., contraction (positive values of the strain) and relaxation (negative values of the strain) of the heart muscle. The cardiac elastogram mode that provides the most diagnostic information has yet to be determined. Possible display choices, for example, include absolute strains, directional elastograms that display positive and negative strains, either positive or negative strains, or other modalities. The cardiologist may prefer one of these modes to evaluate diastolic function but may prefer another to evaluate systolic function. Accurate registration between the echocardiograms and the elastograms (similar to the display for color flow or TDI) may also facilitate the interpretation of these results.

Cardiac elastography addresses many of the limitations of tissue Doppler-based strain-rate imaging that have been described above.^[26,27] First, cardiac elastography could be performed to provide a two-dimensional strain information. Strains in the axial and lateral direction, along with Poisson's ratio have been estimated in elastography without using an incompressibility assumption.^[29] Radiofrequency echo signal data acquired along multiple angular insonification directions of the ultrasound beam have also been utilized to obtain axial and lateral strain tensors.^[30] Cardiac elastography provides strain information with excellent spatial resolution, dependent on signal processing parameters, i.e., the window length and overlap, along with the same temporal resolution provided by strain-rate imaging. As the strain information in elastography is obtained from successive frames of RF data, motion artifacts due to the translation and rotation of the heart may be minimized if high frame rates are available.

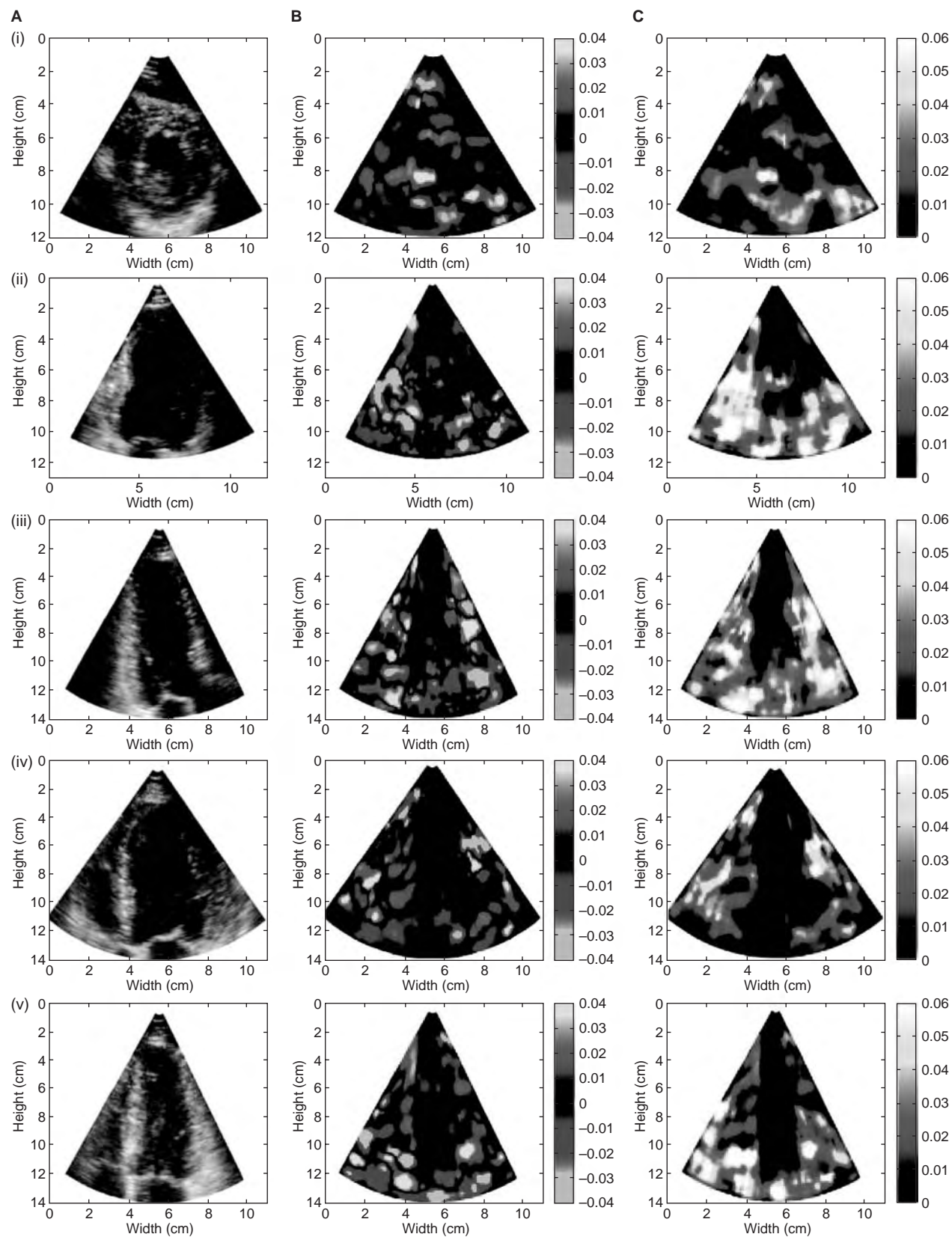


Fig. 3 (A) Two-dimensional gray-scale echocardiograms, (B) two-dimensional vector cardiac elastograms, and (C) two-dimensional absolute cardiac elastograms obtained near ES. (i) Short-axis view for a patient with CAD, (ii) two-chamber apical view for this patient, (iii) three-chamber apical view for a normal subject, (iv) four-chamber apical view for a patient with cardiomyopathy, and (v) four-chamber apical view for a normal subject.

These are currently implemented on the state-of-the-art scanners.

Cardiac elastography may be able to clearly characterize areas of hypokinesia and akinesia. Likely, it will also be used in infarction diagnosis, for displaying the infarcted areas, in addition to the ability of detecting residual contractility in infarcted areas. New physiologic information on systolic and diastolic events should be accessible. This may lead to the use of cardiac elastography in the assessment of conduction disturbances and regional systolic and diastolic function. In addition, models of stress-strain or force velocity relationships in the heart may be used to estimate regional wall stresses by combining regional strain with pressure values measured in the ventricle.

The raw capabilities to acquire full-field strain images are beginning to emerge on modern ultrasound scanners, such as the Vivid 7, with availability of digital echo data for both real-time and off-line processing. The algorithms and techniques developed in this project still must be assessed for both accuracy and diagnostic capability to determine the capabilities of this new form of echo data processing. Nevertheless, the analysis methods can be made available to other researchers who have access to RF echo data, either from the VingMed systems or from the other echocardiography apparatus supporting RF signals. We anticipate that future work will be directed toward real-time, on-line solutions to successful algorithms developed in this research. Cardiac elastography may provide quantitative information on myocardial strain and may lead to development of operator-independent indices for characterization of cardiac dysfunction and regional myocardial function.

CONCLUSIONS

Full-field cardiac elastograms can be obtained in the standard echocardiographic imaging views. Cardiac elastography may provide quantitative and translation-independent measures of myocardial strain during the cardiac cycle, thereby providing a two-dimensional method for characterizing regional myocardial function. Further work is required to evaluate the role of cardiac elastography in the detection of ischemia and infarction. Cardiac elastography may also provide valuable, noninvasive information of resting global and regional dysfunction and of induced ischemic regional dysfunction. The potential of cardiac elastography to successfully and noninvasively characterize ischemic myocardial tissue is also supported by results reported in the literature that demonstrate changes in myocardial stiffness with ischemia.^[12]

ACKNOWLEDGMENTS

This work was supported in part by start-up grant funds from the Department of Medical Physics, Medical School and Graduate School at the University of Wisconsin-Madison, to Dr. Varghese.

REFERENCES

1. Marwick, T.H. Current status of stress echocardiography for diagnosis and prognostic assessment of coronary artery disease. *Coron. Artery Dis.* **1998**, *9* (7), 411-426.
2. Sinusas, A.J.; Beller, G.A.; Watson, D.D. Cardiac imaging with technetium 99m-labeled isonitriles. *J. Thorac. Imaging* **1990**, *5* (3), 20-30.
3. Pelc, N.J.; Herfkens, R.J.; Shimakawa, A.; Enzmann, D.R. Phase contrast cine magnetic resonance imaging. *Magn. Reson.* **1991**, *7* (4), 229-254Q.
4. Denney, T.S., Jr.; Prince, J.L. Reconstruction of 3-D left ventricular motion from planar tagged cardiac MR images: an estimation theoretic approach. *IEEE Trans. Med. Imaging* **1995**, *14* (4), 625-635.
5. McVeigh, E.R.; Zerhouni, E.A. Noninvasive measurement of transmural gradients in myocardial strain with MR imaging. *Radiology* **1991**, *180* (3), 677-683.
6. Heimdal, A.; Stoylen, A.; Torp, H.; Skjaerpe, T. Real-time strain rate imaging of the left ventricle by ultrasound. *J. Am. Soc. Echocardiogr.* **1998**, *11* (11), 1013-1019.
7. Mirsky, I.; Parmley, W.W. Assessment of passive elastic stiffness for isolated heart muscle and the intact heart. *Circ. Res.* **1973**, *33* (2), 233-243.
8. Van Leuven, S.L.; Waldman, L.K.; McCulloch, A.D.; Covell, J.W. Gradients of epicardial strain across the perfusion boundary during acute myocardial ischemia. *Am. J. Physiol.* **1994**, *267* (6 pt 2), H2348-H2362.
9. Piene, H.; Myhre, E.S. Position of interventricular septum during heart cycle in anesthetized dogs. *Am. J. Physiol.* **1991**, *260* (1 pt 2), H158-H164.
10. Hashima, A.R.; Young, A.A.; McCulloch, A.D.; Waldman, L.K. Nonhomogeneous analysis of epicardial strain distributions during acute myocardial ischemia in the dog. *J. Biomech.* **1993**, *26* (1), 19-35.
11. Azhari, H.; Weiss, J.L.; Rogers, W.J.; Siu, C.O.; Zerhouni, E.A.; Shapiro, E.P. Noninvasive quantification of principal strains in normal canine hearts using tagged MRI images in 3-D. *Am. J. Physiol.* **1993**, *264* (1 pt 2), H205-H216.
12. Miyaji, K.; Sugiura, S.; Inaba, H.; Takamoto, S.; Omata, S. Myocardial tactile stiffness during acute reduction of coronary blood flow. *Ann. Thorac. Surg.* **2000**, *69* (1), 151-155.
13. Young, A.A.; Axel, L. Three-dimensional motion and deformation of the heart wall: estimation with spatial modulation of magnetization—a model-based approach. *Radiology* **1992**, *185* (1), 241-247.
14. O'Dell, W.G.; Moore, C.C.; Hunter, W.C.; Zerhouni, E.A.; McVeigh, E.R. Three-dimensional myocardial

- deformations: calculation with displacement field fitting to tagged MR images. *Radiology* **1995**, *195* (3), 829–835.
15. McDicken, W.N.; Sutherland, G.R.; Moran, C.M.; Gordon, L.N. Colour Doppler velocity imaging of the myocardium. *Ultrasound Med. Biol.* **1992**, *18* (6–7), 651–654.
 16. Sutherland, G.R.; Stewart, M.J.; Groundstroem, K.W.; Moran, C.M.; Fleming, A.; Guell-Peris, F.J.; Riemersma, R.A.; Fenn, L.N.; Fox, K.A.; McDicken, W.N. Color Doppler myocardial imaging: a new technique for the assessment of myocardial function. *J. Am. Soc. Echocardiogr.* **1994**, *7* (5), 441–458.
 17. Fleming, A.D.; Xia, A.; McDicken, W.N.; Sutherland, G.R.; Fenn, L. Myocardial velocity gradients detected by Doppler imaging. *Br. J. Radiol.* **1994**, *67* (799), 679–688.
 18. Kanai, H.; Hasegawa, H.; Chubachi, N.; Koiwa, Y.; Tanaka, M. Noninvasive evaluation of local myocardial thickening and its color-coded imaging. *IEEE Trans. Ultrason. Ferroelectric. Freq. Control* **1997**, *44*, 752–768.
 19. Hashimoto, I.; Li, X.; Hejmadi Bhat, A.; Jones, M.; Zetts, A.D.; Sahn, D.J. Myocardial strain rate is a superior method for evaluation of left ventricular sub-endocardial function compared with tissue Doppler imaging. *J. Am. Coll. Cardiol.* **2003**, *42* (9), 1574–1583.
 20. Urheim, S.; Edvardsen, T.; Torp, H.; Angelsen, B.; Smiseth, O.A. Myocardial strain by Doppler echocardiography. Validation of a new method to quantify regional myocardial function. *Circulation* **2000**, *102* (10), 1158–1164.
 21. Prinzen, F.W.; Hunter, W.C.; Wyman, B.T.; McVeigh, E.R. Mapping of regional myocardial strain and work during ventricular pacing: experimental study using magnetic resonance imaging tagging. *J. Am. Coll. Cardiol.* **1999**, *33* (6), 1735–1742.
 22. Zabalgoitia, M.; Rahman, S.N.; Haley, W.E.; Yarows, S.; Krause, L.; Anderson, L.C.; Oraby, M.A.; Amarena, J. Effect of regression of left ventricular hypertrophy from systemic hypertension on systolic function assessed by midwall shortening (HOT echocardiographic study). *Am. J. Cardiol.* **2001**, *88* (5), 521–525.
 23. Ophir, J.; Cespedes, I.; Ponnekanti, H.; Yazdi, Y.; Li, X. Elastography: a quantitative method for imaging the elasticity of biological tissues. *Ultrason. Imaging* **1991**, *13* (2), 111–134.
 24. O'Donnell, M.; Skovoroda, A.R.; Shapo, B.M.; Emelianov, S.Y. Internal displacement and strain imaging using ultrasonic speckle tracking. *IEEE Trans. Ultrason. Ferroelec. Freq. Control* **1994**, *41* (3), 314–325.
 25. Varghese, T.; Ophir, J.; Konofagou, E.; Kallel, F.; Righetti, R. Tradeoffs in elastographic imaging. *Ultrason. Imaging* **2001**, *23* (4), 216–248.
 26. Konofagou, E.E.; D'Hooge, J.; Ophir, J. Myocardial elastography—a feasibility study in vivo. *Ultrasound Med. Biol.* **2002**, *4*, 475–482.
 27. Varghese, T.; Zagzebski, J.A.; Rahko, P.; Breburda, C.S. Ultrasonic imaging of myocardial strain using cardiac elastography. *Ultrason. Imaging* **2003**, *25* (1), 1–16.
 28. Varghese, T.; Breburda, C.S.; Zagzebski, J.A.; Rahko, P. Method and Apparatus for Cardiac Elastography. U.S. Patent 6,749,571, United States: Wisconsin Alumni Research Foundation, University of Wisconsin–Madison, 2004.
 29. Konofagou, E.; Ophir, J. A new elastographic method for estimation and imaging of lateral displacements, lateral strains, corrected axial strains, and Poisson's ratios in tissues. *Ultrasound Med. Biol.* **1998**, *24* (8), 1183–1199.
 30. Techavipoo, U.; Chen, Q.; Varghese, T.; Zagzebski, J.A. Estimation of displacement vectors and strain tensors in elastography using angular insonifications. *IEEE Trans. Med. Imaging* **2004**, *23* (12), 1479–1489.

FURTHER READING

- Konofagou, E.E.; D'hooge, J.; Ophir, J. Myocardial elastography—a feasibility study in vivo. *Ultrasound Med. Biol.* **2002**, *28* (4), 475–482.
- Varghese, T.; Zagzebski, J.A.; Rahko, P.; Breburda, C.S. Ultrasonic imaging of myocardial strain using cardiac elastography. *Ultrason Imaging* **2003**, *25* (1), 1–16.



Cardiac Output, Basic Measurement Principles

Leslie A. Geddes

Professor Emeritus, School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana, U.S.A.

CARDIAC OUTPUT

Introduction

The heart is a two-sided stroke pump in a double circulatory system. The right heart pumps blood through the lungs; the left heart pumps blood through the rest of the body. The output of the right heart is equal to the output of the left heart. In fact, the output of the right heart, after flowing through the lungs, is the input to the left heart. Similarly, the output of the left heart, after flowing through the arteries to the capillaries and collecting veins, is the input to the right heart. Fig. 1 illustrates these points. Cardiac output is, therefore, the output of either the right or left heart. Cardiac output (CO) is equal to the volume of blood ejected per beat, which is called stroke volume (SV), multiplied by heart rate (HR). Thus,

$$CO = SV \times HR.$$

It is easy to measure heart rate, but there are very few methods for measuring stroke volume. It is for this reason that recourse is taken to methods that measure the average cardiac output, which is usually expressed in liters per minute. To compare cardiac output in large and small subjects and among different species, two methods of normalization are used. One method expresses cardiac output per unit of body weight; the other divides it by the body-surface area.

Thus, the units are mL/min/kg of body weight and L/min/m² of body surface area. The latter unit is more frequently used and is given the special name “cardiac index.” Table 1 presents typical basal values for both quantities for a variety of species. Note the similarity of cardiac index values.

It is of some interest to note that the cardiac output for a typical 70-kg adult male subject at rest is about 5 L/min. If the heart rate is 72 beats/min, the stroke volume is $5000/72 = 69.4$ mL. Such a subject would have a body surface area of about 1.85 m²; therefore, the cardiac index is $5/1.85 = 2.7$ L/min/m². The cardiac output per kilogram of body weight is $5000/70 = 71.4$ mL/min/kg. In such a subject, if cardiac output dropped to 3 L/min, consciousness would be impaired.

Muscular exercise increases cardiac output dramatically. Table 2 presents typical values. Highly

trained long-distance runners, when in condition, can increase their cardiac output from a resting value of 5–25, or 30 L/min with maximal exercise. It is interesting to note that trained athletes have a resting heart rate that is much lower than untrained subjects; the heart rate of a trained athlete is 40–50 beats/min at rest. The resting cardiac output is about the same as in other subjects of similar size; therefore, the stroke volume of an athlete is much larger than that of an untrained subject. In general, there are two major cardiovascular differences in athletes: (1) the resting heart rate is low and (2) following exercise, the heart rate returns to a resting level very rapidly. In fact, the ability of the heart to increase its output and to return quickly to a resting value after exercise are cardiovascular parameters that identify physical fitness.

The cardiac output is shared by the various organs and tissues. A typical resting distribution is shown in Table 3 and Fig. 1B. It is interesting to note that the heart takes about one-twentieth of its output (via the coronary arteries) for its own operation. The brain requires about one-seventh of the cardiac output. These two important areas have a high priority for circulation, and there is little control over their flow rates. However, the amount of blood going to the other vascular beds is under considerable control, depending on the circumstances. Kidney, hepatic portal, muscle, and skin blood flows can and do vary considerably depending on circumstances. Fig. 1B illustrates these controls.

Table 3 shows the extraction of oxygen from the blood for the various circulatory regions. Note that the heart extracts the highest proportion of oxygen in comparison to other organs. It is noteworthy that venous blood contains a wide range of oxygen saturations. It is for this reason that the most representative sample of venous blood is found on the pulmonary artery.

BLOOD VOLUME

Before discussing the methods for measuring cardiac output, it is of value to identify the amount of blood in the body. Table 4 presents typical values for blood volume for different species in mL/kg of body weight. In humans, the blood volume is typically about 7% of

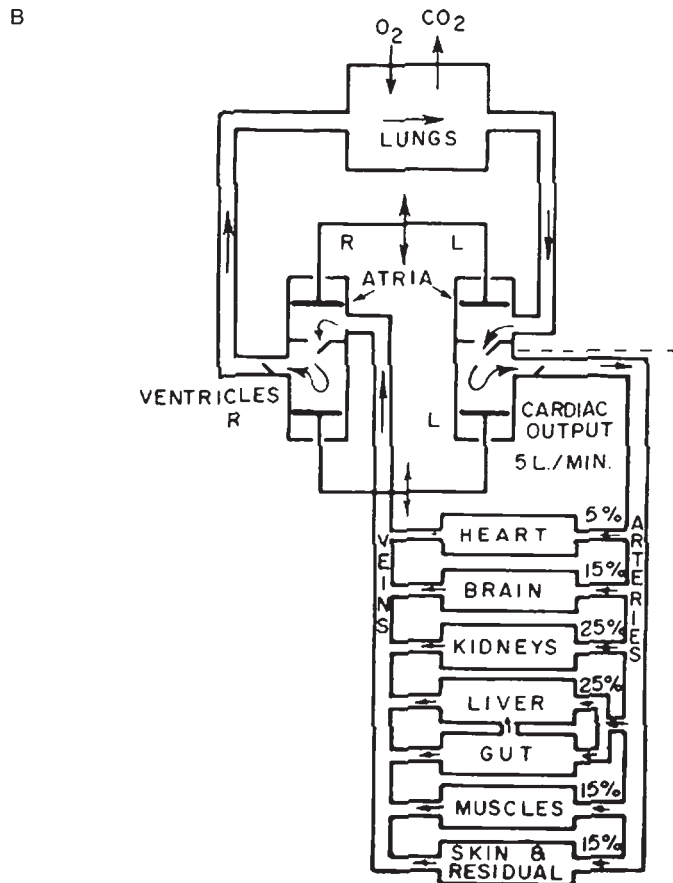
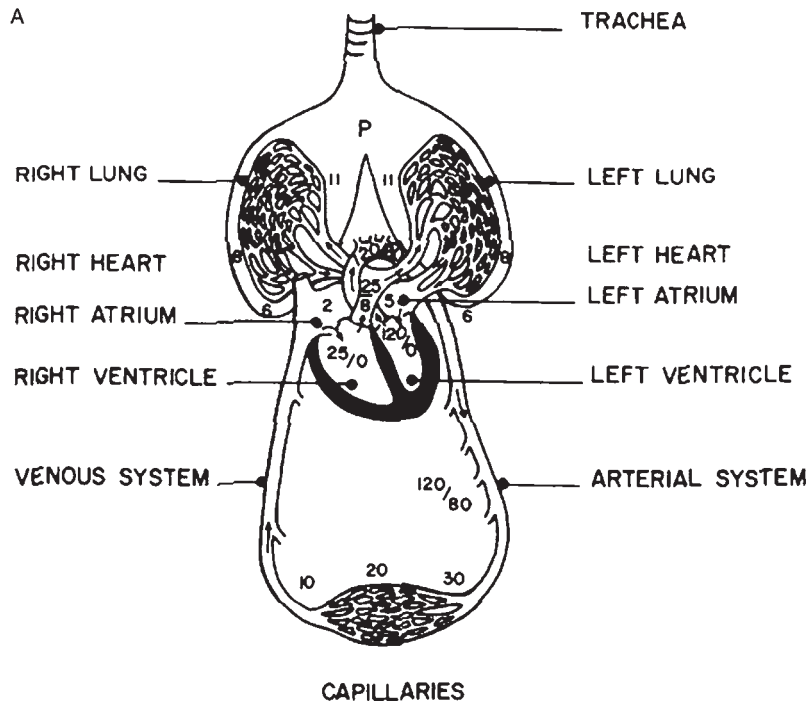


Fig. 1 (A) The circulatory system.
(B) Mechanical analog of the circulatory system.



Table 1 Cardiovascular data

Species	Weight (kg)	Area (m ²)	State	Cardiac output	Cardiac index (L/min/m ²)	CO (mL/min/kg)
Human (5 yr)		0.82	Basal	2.28	2.78	122
Human (25 yr)	68.5	1.85	Basal	5.10	2.76	74.7
Human (60 yr)	69.8	1.79	Basal	4.55	2.54	65.2
Dog		0.39	Basal	1.12	2.9	175
Goat		0.71		2.2	3.1	137
Goat		0.91	Basal	3.1	3.4	131
Horse	283	4.30	Standing	18.8	4.4	66.4

Source: Data from Saunders (see Ref. [1]).

the body weight. Often the blood volume is stated as 70 mL/kg of body weight.

The blood volume of a subject is remarkably constant; however, its distribution is of some interest. For example, 60% is found in the veins and 13% is in the arteries. The lungs accommodate 12%, the heart 8%, and, surprisingly enough, only about 7% is in the capillaries, where all of the metabolic exchange takes place. It should be clear that a very small change in the size of the venous reservoir would allow accommodation of a large volume of blood. Venous pooling of blood is a serious problem in cardiovascular shock. The spleen is a reservoir for red cells that can be delivered to the circulation following blood loss. Parenthetically, when blood is donated by human volunteers, about 500 mL are withdrawn. Such a

donation amounts to about 10% of the blood volume and is relatively unnoticed by the average donor. The volume of blood in a subject is measured with the dilution technique, which is used to measure the volume of any irregularly shaped or inaccessible container. A measured amount (m) of detectable substance (indicator) is introduced to the fluid-filled container and, after complete mixing, a sample is withdrawn. The concentration (c) of the indicator in the sample is then measured; therefore, the volume of the container $V = m/c$.

In applying the dilution principle to measure blood volume, the indicator must have special properties. For example, there must be a minimum loss of indicator through the kidneys or by trapping in an organ. For this reason, two types of indicator are used: one combines with the plasma proteins and the other becomes incorporated in the red cells. Thus what is measured is either plasma volume or red-cell volume. Blood volume is calculated by measuring the fraction of cells (by centrifugation) and scaling the result. The fraction of cells is called the packed-cell volume (PCV) and is often expressed as a percentage. Because there are so many more red cells, than platelets or leukocytes, the percent red-cell volume (or hematocrit) is often used as a descriptor of the cell fraction in blood.

Evan's blue, a dye that attaches to plasma albumin, and radioiodine labeled albumin are the most frequently used indicators. The concentration of Evan's blue is measured with a spectrophotometer (at 640 nm) and the concentration of radioiodine-labeled albumin is measured with a radioactivity detector. The technique employed consists of drawing a blood sample and centrifuging it to separate the plasma from the cells. This procedure produces a plasma sample (blank) without the indicator and allows calculation of the packed-cell volume (cell fraction).

The indicator is injected intravenously and time is allowed for mixing of the indicator. At 10, 20, and

Table 2 Range of cardiac output in humans

Activity	Duration (hr)	Output (L/min)	Power to blood (W)
Dinner	1	6.37	1.41
Relaxation	3	7.17	1.59
Sleep	8	6	1.33
Dressing, etc.	1	9.03	2.00
Driving	0.5	8.32	1.85
Moderate work	4	8.6	1.91
Walking	0.25	10.5	2.33
Lunch	0.5	6.37	1.41
Walking	0.25	10.5	2.33
Moderate work	4.0	8.6	1.91
Driving	0.5	8.32	1.85
Moderate exercise	1.0	11.5	2.55
Average		7.6	1.68

Table 3 Distribution of cardiac output

Distribution (Human) region	CO (%)	Approximate fraction	O ₂ cm ³ /100 cm ³ A-V O ₂ difference	Venous ^a saturation (%)
Hepatic/portal	25	1/4	3.4	83.3
Kidneys	25	1/4	1.4	93
Brain	15	1/7	6.2	69
Heart (coronaries)	5	1/20	11.4	43
Skeletal muscle	15.5	1/6	6	70
Skin and residual	14.5	1/7		

^aNote that venous system contains blood of varying saturations.

30 min, venous blood samples are withdrawn, centrifuged, and the concentration of indicator is measured in the plasma sample. With Evan's blue, the concentration is proportional to the optical density [log e (1/transmission)]. Similarly, the count-per minute rate is proportional to the concentration of the radioiodine albumin. These concentrations are plotted on a graph, which shows the slight loss of indicator. Fig. 2 illustrates a typical concentration-time curve which is extrapolated to zero time to identify the concentration of the indicator corresponding to zero loss.

This zero-time concentration is divided into the amount of indicator injected to obtain plasma volume. As a practical matter, it is more convenient to plot the optical density or counts/minute versus time and extrapolate to obtain the zero-time value. This Fig. 2 can then be converted to concentration by adding a known amount of indicator to the plasma bank and measuring it. The blood volume is the plasma volume divided by (1-PCV), where PCV is the packed-cell fraction.

Table 4 Blood volumes

Species	mL/kg
Human	
Male	74-85
Female	63-74
Dog	66-102
Monkey	60 to -5
Guinea pig	72-75
Rabbit	55-72
Rat	45-63
Cattle	52-62
Goat	57-82
Swine	61-69
Horse	71-109

Source: From National Academy of Science (see Ref. [2]).

DILUTION METHODS FOR CARDIAC OUTPUT

Although there are numerous methods of measuring cardiac output, the two that are frequently used are the Fick and the indicator-dilution methods. Both employ the dilution principle and will now be described. An indirect method which employs recording of the electrical impedance changes that appear between thoracic electrodes is also used in some clinics; this method will also be described.

Fick Method

Venous blood flowing through the lungs from the right heart to the left heart picks up oxygen and releases carbon dioxide (Fig. 3A). In 1870 Fick^[3] stated that if the amount of oxygen taken up (or the amount of carbon dioxide liberated) per minute is known and the oxygen (or carbon dioxide) concentration in the blood entering and leaving the lungs is known, it is possible to calculate blood flow through the lungs (cardiac output). A convenient way of understanding Fick's proposal is shown in Fig. 3B.

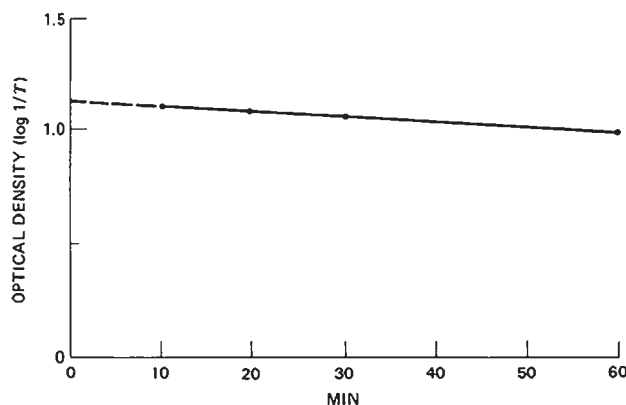


Fig. 2 Calculation of blood volume by determination of the zero-time concentration of an injected indicator.



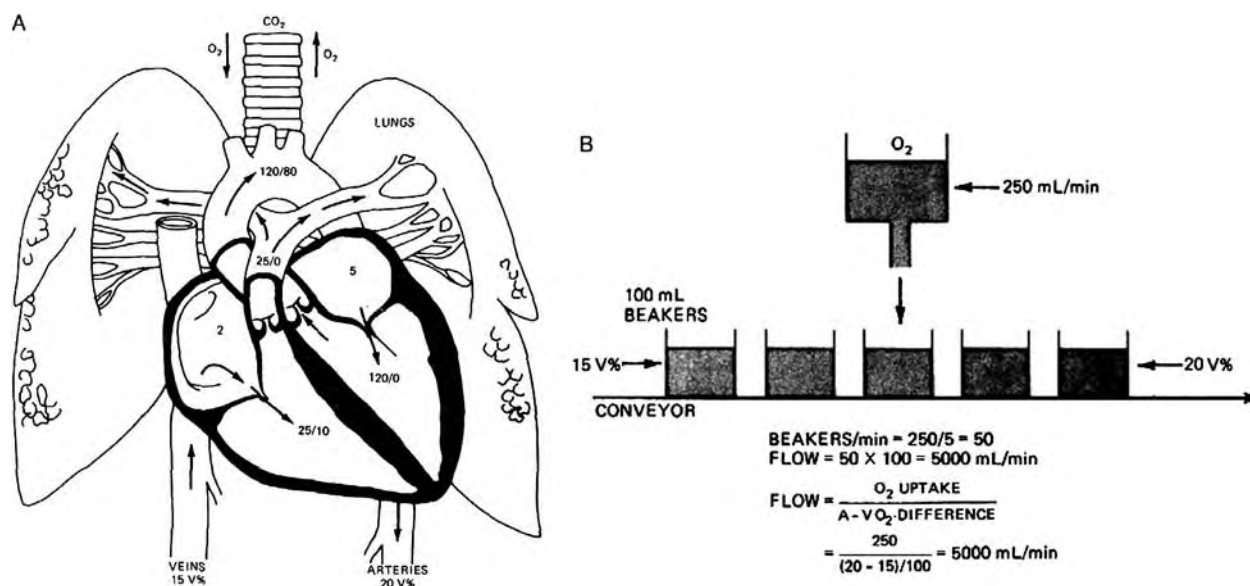


Fig. 3 (A) The pulmonary circuit. (B) Principle employed in the Fick method using the conveyor-belt analog.

Consider delivery of oxygen to the blood to be equivalent to a conveyor belt in which venous blood from the right heart comes along in 100 mL beakers, which pass a dispensing reservoir that delivers oxygen. Thus the (arterial) beakers that have passed the oxygen dispenser contain more oxygen. Assume that the concentration of oxygen in the venous blood is 15 mL/100 mL blood (i.e., 15 %) and that the concentration of oxygen in the arterial beakers is 20 mL/100 mL blood. In this steady-state-conveyor-belt system it was found that 250 mL of oxygen disappeared from the dispensing reservoir each minute. Each 100 mL beaker of blood took up $20 - 15 = 5$ mL of oxygen; therefore, $250/5 = 50$ beakers must have passed each minute. Cardiac output must therefore be $50 \text{ beakers} \times 100 \text{ mL/beaker} = 5000 \text{ mL/min}$. These facts are usually expressed by the following familiar expression:

Cardiac Output

$$= \frac{\text{oxygen uptake/min}}{\text{arterial} - \text{venous oxygen concentration}}$$

Using the numbers given in the conveyor-belt analog:

$$\begin{aligned} \text{CO} &= \frac{250 \text{ (mL/min)}}{(20/100) - (15/100)} = \frac{250 \times 100}{5} \\ &= 5000 \text{ mL/min} \end{aligned}$$

The oxygen uptake is measured using a spirometer equipped with a carbon dioxide absorber. A typical record of oxygen uptake is shown in Fig. 4. Note that the oxygen uptake, which is the slope of the recording, must be corrected to body temperature and saturated

with water vapor by using the correction factor shown in Table 5.

$V_s = FV_c$ where V_s is the standardized condition and V_c is the collected condition.

$$V_s = \frac{(1 + (37/273)) (P_B - P_{H_2O})}{(1 + (t^\circ C/273)) (P_B - 47)} V_c = FV_c$$

The concentration of oxygen in the arterial and venous blood is less easily measured. Two instruments, the Van Slyke manometric apparatus or the LexO₂con analyzer, are used. The former is a gasometric instrument in which the oxygen content is measured by removal of the oxygen from the blood sample. The LexO₂con is an electrochemical instrument that counts the oxygen molecules and indicates the O₂ content in the sample of blood introduced. The indication represents volume percent at 0°C and 760 mm Hg dry air. Therefore, this Fig. 4 must be corrected to body temperature and ambient pressure, saturated.

The Fick method is reliable but requires steady-state conditions. It can be applied using carbon dioxide and, with slightly more difficulty, foreign gases. The highest accuracy is obtained when the mixed venous sample is obtained from the pulmonary artery; right ventricular blood is next best. The arterial sample can be obtained from any convenient artery.

Indicator-Dilution Method

Stewart^[5] showed that if a known concentration of indicator is introduced into a flowing stream and the temporal concentration of the indicator is measured at

Subject: 36-Year-old male
 Height: 5'10½"
 Weight: 175 lb*
 Body surface area: $A = W^{0.425} \times H^{0.725} \times 71.84$
 where W = weight in kg
 H = height in cm
 A = area in cm²
 $A = 1.98 \text{ m}^2$
 O₂ Uptake at 26°C: 400 mL/min
 Barometer: 750 mm Hg P_{H_2O} at 26°C: 25.2 mm Hg
 O₂ Uptake correction factor: F

$$F = \frac{273 + 37}{273 + 26} \times \frac{750 - 25.2}{750 - 47}$$

 $F = 1.069$
 F from Table 4.5 = 1.069
 O₂ Uptake BTPS: $400 \times 1.069 = 427.6$
 Blood O₂:
 Arterial = 20 V%
 Mixed venous = 15 V%
 Cardiac output:

$$CO = \frac{427.6}{(20-15)/100} = 8552 \text{ mL/min}$$

 Cardiac index:

$$CI = \frac{8.552}{1.98} = 4.32 \text{ L/min/m}^2$$

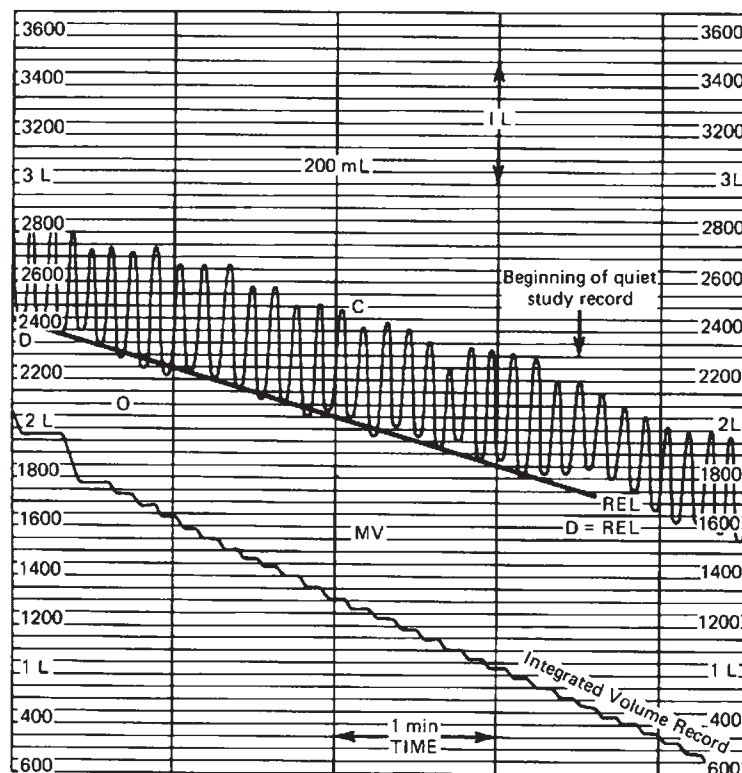


Fig. 4 Measurement of oxygen uptake with a spirogram, and sample calculation of cardiac output using the Fick method.

a downstream point, it is possible to calculate the volume flow. To illustrate the salient points of the indicator-dilution method, consider a fluid flowing at a constant rate in a tube of cross-sectional area a , as shown in Fig. 5. Assume that m grams of a soluble and detectable indicator are rapidly introduced to form a uniform cylinder of indicator in the fluid. If a concentration detector is placed slightly downstream and its output is recorded, an idealized rectangular concentration-time curve of amplitude C will be recorded as the cylinder passes. The practical use of this idealized concentration-time curve will now be illustrated.

In a tube containing a flowing fluid, flow is equal to the velocity of flow (v) multiplied by the cross-sectional area (a); flow $\phi = va$. Velocity is distance divided by time, and in the model the cylinder of indicator moves L units of distance in t s, the value for t being identified on the concentration-time curve; therefore, $v = L/t$. Substituting L/t for velocity, $\phi = La/t$. However, La is the volume of the cylinder (V) and, therefore, $\phi = Vt$. It is not possible to measure the volume of the cylinder, but it is possible to measure concentration (C), which is equal to m/V . Substituting this quantity gives:

$$\phi = \frac{m}{Ct} = \frac{\text{mass of indicator injected}}{\text{concentration} \times \text{time}}$$

The expression just derived illustrates that in the idealized case, if m grams of indicator are injected and

the calibrated downstream concentration-time curve is recorded, it is possible to calculate flow. In a practical situation, the cylinder of indicator would be spread out into a teardrop shape owing to the velocity profile in the tube. Therefore, the concentration-time curve will rise rapidly and fall slowly as shown in Fig. 5B. However, a rectangular concentration-time curve can be obtained from this typical dilution curve by measuring the area under it and dividing by the base. Performing this operation provides a mean concentration C for the time t . Thus,

$$\phi = \frac{m}{Ct}$$

There is an extremely important subtlety about this flow expression. Note that it is only necessary to know the amount (m) of indicator injected and the downstream concentration-time curve. The concentration-time curve can be recorded from any downstream branch. The important requirement is that the indicator is injected so that it mixes with all of the flow.

In the measurement of cardiac output the indicator is injected into the right atrium, right ventricle, or pulmonary artery and the dilution curve is obtained by withdrawing arterial blood (from any large artery) into the detector at a constant rate. With the thermal-dilution method, injection is into the right atrium and detection is in the pulmonary artery.



Table 5 Correction factor (F) for standardization of collected volume

Temperature (°C)/P _B	640	650	660	670	680	690	700	710	720	730	740	750	760	770	780
15	1.1388	1.1377	1.1367	1.1358	1.1348	1.1339	1.1330	1.1322	1.1314	1.1306	1.1298	1.1290	1.1283	1.1276	1.1269
16	1.1333	1.1323	1.1313	1.1304	1.1295	1.1286	1.1277	1.1269	1.1260	1.1253	1.1245	1.1238	1.1231	1.1224	1.1217
17	1.1277	1.1268	1.1266	1.1249	1.1240	1.1232	1.1224	1.1216	1.1208	1.1200	1.1193	1.1186	1.1179	1.1172	1.1165
18	1.1222	1.1212	1.1203	1.1194	1.1186	1.1178	1.1170	1.1162	1.1154	1.1147	1.1140	1.1133	1.1126	1.1120	1.1113
19	1.1165	1.1156	1.1147	1.1139	1.1131	1.1123	1.1115	1.1107	1.1100	1.1093	1.1086	1.1080	1.1073	1.1067	1.1061
20	1.1108	1.1099	1.1091	1.1083	1.1075	1.1067	1.1060	1.1052	1.1045	1.1039	1.1032	1.1026	1.1019	1.1014	1.1008
21	1.1056	1.1042	1.1034	1.1027	1.1019	1.1011	1.1004	1.0997	1.0990	1.0984	1.0978	1.0971	1.0965	1.0960	1.0954
22	1.0992	1.0984	1.0976	1.0969	1.0962	1.0964	1.0948	1.0941	1.0935	1.0929	1.0923	1.0917	1.0911	1.0905	1.0900
23	1.0932	1.0925	1.0918	1.0911	1.0904	1.0897	1.0891	1.0884	1.0878	1.0872	1.0867	1.0861	1.0856	1.0850	1.0845
24	1.0873	1.0866	1.0859	1.0852	1.0846	1.0839	1.0833	1.0827	1.0822	1.0816	1.0810	1.0805	1.0800	1.0795	1.0790
25	1.0812	1.0806	1.0799	1.0793	1.0787	1.0781	1.0775	1.0769	1.0764	1.0758	1.0753	1.0748	1.0744	1.0739	1.0734
26	1.0751	1.0710	1.0738	1.0732	1.0727	1.0721	1.0716	1.0710	1.0705	1.0700	1.0696	1.0691	1.0686	1.0682	1.0678
27	1.0688	1.0682	1.0677	1.0671	1.0666	1.0661	1.0656	1.0651	1.0640	1.0641	1.0637	1.0633	1.0629	1.0624	1.0621
28	1.0625	1.0619	1.0614	1.0609	1.0604	1.0599	1.0595	1.0591	1.0586	1.0582	1.0578	1.0574	1.0570	1.0566	1.0563
29	1.0560	1.0555	1.0550	1.0546	1.0548	1.0537	1.0533	1.0529	1.0525	1.0521	1.0518	1.0514	1.0519	1.0507	1.0504
30	1.0494	1.0496	1.0486	1.0482	1.0478	1.0474	1.0470	1.0467	1.0463	1.0460	1.0450	1.0453	1.0450	1.0447	1.0444

Source: From Annual Reviews (see Ref. [4]).

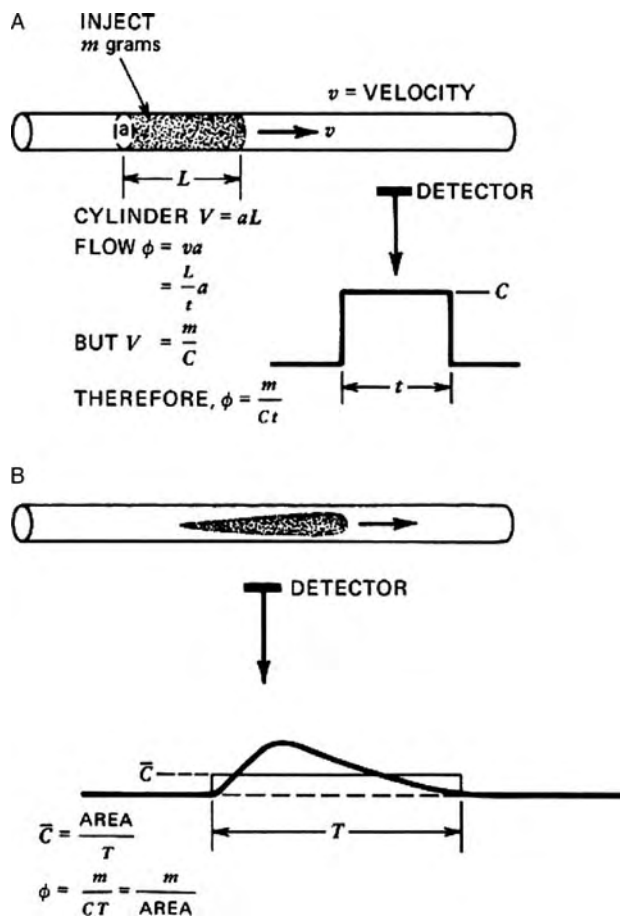


Fig. 5 Genesis of the dilution curve.

Before describing practical application of the indicator-dilution method, several important facts should be recognized, the most important of which is that cardiac output must remain constant for the duration of measurement. The characteristics of the indicator are also very important. Obviously, it must be detectable and not be a stimulant or depressant to the circulatory system. It must be retained for a period long

enough to permit inscription of a dilution curve. Indicators that are lost from the circulation between the injection and measuring sites provide a falsely high value for cardiac output because C is too small and is in the denominator of the flow expression. Indicators that are rapidly cleared from the circulation allow repeated determinations without indicator buildup. Indicators that are retained in the circulation can build up to excessively high concentrations if repeated determinations are made. If the indicators are colored dyes, there is the possibility of altering skin color until the indicator is cleared.

A wide variety of indicators are used to measure cardiac output; Table 6 presents a typical list, along with the type of detector employed for each. However, in clinical medicine, the dye and thermal-dilution methods are perhaps the most popular.

The method of obtaining a dilution curve from which cardiac output can be calculated is shown in Fig. 6. The site of injection for the indicator is the right atrium, ventricle, or pulmonary artery.

In the illustration, a catheter connected to the indicator-filled syringe has been advanced down the right jugular vein. In the human the basilic vein in the arm is used as the entry point.

To inscribe the dilution curve, an arterial catheter is connected to the detector for the indicator, which is in turn connected to a motor-driven withdrawal syringe as shown in Fig. 6. The catheter is introduced into a major artery, such as the femoral, and advanced into the aorta. The indicator is then injected rapidly and blood is drawn through the detector, the output of which is recorded to obtain the dilution curve as shown. Note that the dilution curve does not return to zero indicator concentration because of recirculation of the indicator. To process the dilution curve, this recirculation must be removed and the dilution curve for a single pass must be synthesized. The method of performing this correction will now be described.

Table 6 Indicators

Material	Detector	Retention data
Evans blue (T1824)	Photoelectric, 640 μ m	50% loss in 5 days ^a
Cardiogreen	Photoelectric, 800 μ m	50% loss in 10 min ^b
Coomassie blue	Photoelectric, 585–600 μ m	50% loss in 15–20 min ^c
Saline (5%)	Conductivity cell	Diffusible ^d
Albumin ¹³¹ I	Radioactive	
²⁴ Na, ⁴² K, D ₂ O, DH	Radioactive	Diffusible
Hot-cold solutions	Thermo-detector	Diffusible

^aConnolly et al.^[6]

^bFox,^[7] Wheeler^[8]

^cTaylor^[9]

^dIt is estimated that there is less than 15% loss of diffusible indicators during the first pass.



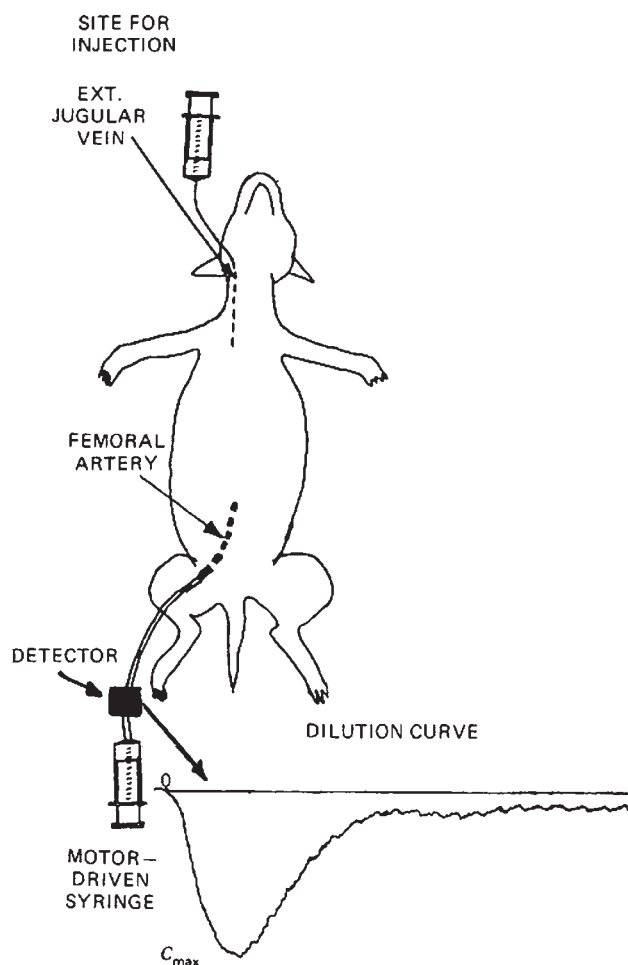


Fig. 6 The indicator-dilution technique.

Because the descending limb of the dilution curve is exponential, semilogarithmic plotting of this part of the curve provides a means of identifying the end of the first pass, which, for practical purposes, is taken as the time for the dilution curve to fall to 1% of its maximum concentration. Semilogarithmic plotting is usually performed first on the uncalibrated dilution curve so that a true single-pass curve can be synthesized and measured for area to determine the mean height for the time of one pass. Fig. 7A illustrates a typical dilution curve showing: recirculation on the downslope; Fig. 7B presents a plot of the downslope of this curve. Note that the part of the curve reflecting recirculation departs from the linear portion of the downslope. Extrapolation of the linear portion allows identification of the time for the practical end of the dilution curve, which is taken as 1% of the maximum height (C_{\max}) of the dilution curve, that is, the time corresponding to $C_{\max}/100$. The dilution curve can be completed by plotting the heights for times earlier than the extrapolated end time as shown by the dashed line in Fig. 7A. Then the area under the corrected dilution curve is measured and divided by its length; this gives a

mean height which is converted to a mean concentration (C) after calibration of the detector in terms of amplitude on the record.

Calibration of the detector is accomplished easily by obtaining a blood sample and adding a known amount of indicator to it. Then this mixed sample is drawn through the detector and the deflection is recorded. In this example, the indicator calibration amplitude corresponds to 0.43 g/L of indicator in blood. The mean concentration (area of the dilution curve divided by its base) corresponds to a mean concentration of 0.173 g/L of indicator in blood. Since the corrected dilution curve is 18.5 s in duration, the cardiac output per minute is as follows:

$$CO = \frac{60 \times 0.15}{0.173 \times 18.5} = 2.81 \text{ L/min}$$

As stated previously, this value for cardiac output would be expressed as cardiac index, that is, cardiac output per square meter of body surface area. In this case the body weight of the dog was 12 kg. The body surface area was 0.59 m² (area = 11.2 $W^{0.667}$, where W is body weight in grams and the area is in square centimeters). Therefore, the cardiac index is 4.76 L/min/m²; obviously, the animal was not in a basal state.

Thermal-Dilution Method

The thermal-dilution method in which the indicator is a cold liquid is a variant of the indicator-dilution method. A thermal detector is used to record the downstream decrease in temperature which, when recorded, constitutes the thermal-dilution curve. The usual site for injection is the right atrium, and the detecting site is the pulmonary artery. A special balloon-tipped catheter,^[10] with a thermistor at the tip, is used to detect the temperature change. The same catheter has a side port which permits injection of the cold solution into the right atrium. Another lumen is used for pressure monitoring. Fig. 8 illustrates the principle, and a typical thermal-dilution curve is shown in Fig. 9.

The advantages of the thermal-dilution method are many and, in part, relate to the way it is used. Any suitable cold biocompatible solution can be used as the indicator; isotonic dextrose (5%) at 0°C is the most popular. Because the detecting site is in the pulmonary artery and the indicator is injected into the right atrium, recirculation minimally obscures the downslope of the dilution curve. Therefore, a simple integrator can be used to determine the area under the dilution curve. Since the dilution curve is inscribed with a detector calibrated in degrees centigrade, it is not necessary to prepare calibrating solutions. Moreover, there is no indicator buildup, and dilution curves can be

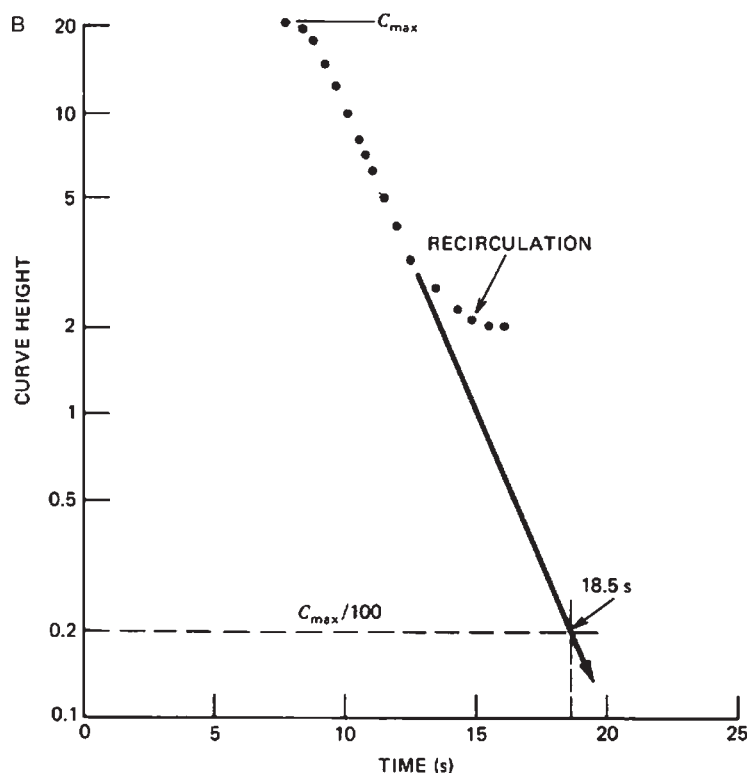
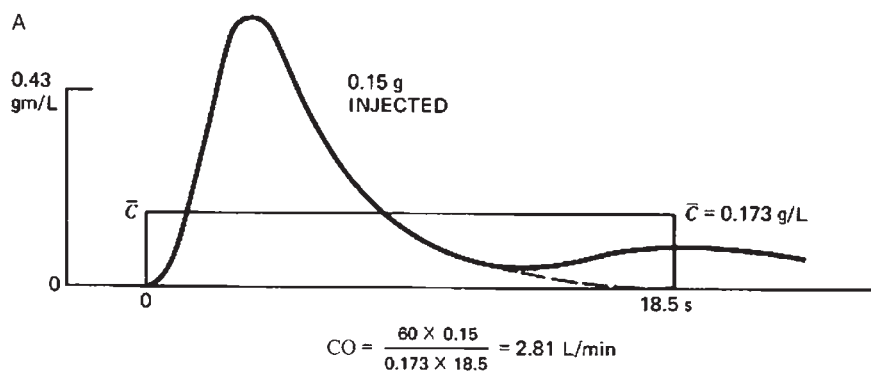


Fig. 7 (A) Calculation of cardiac output from the corrected dilution curve. (B) Semilog plot of the downslope of the dilution curve shown in Fig. 7A. Note that the portion of the downslope of the dilution curve that is exponential is linear on the semilog plot. Extrapolation of the line joining the points forming the straight line allows the acquisition of values for the dilution curve during recirculation. In practice, extrapolation is carried out to find a time corresponding to 1% of the maximum height of the dilution curve. In this example the end of the dilution curve was found to occur at 18.5 s and the amplitudes between 12 s (where recirculation became apparent) and 18.5 s (the end of the curve) were used to complete the dilution curve shown in Fig. 7A.)

obtained as often as every few minutes. Diminishing these considerable advantages only slightly is the need to apply a correction factor to account for the heat transfer through the catheter and loss to the vessel walls during injection of the indicator. There is also the need to obtain the temperature of the pulmonary artery blood prior to injection of the thermoinicator. Fig. 9 illustrates small variations in temperature prior to appearance of the dilution curve; averaging takes care of this.

The fact that a cold solution can be used as an indicator was reported by Fegler^[11] who validated the method in a model and then in dogs in which thermocouples were placed in the right ventricle and aorta. The indicator was 3–5 mL of Ringer's solution at 18°C, which was injected into the inferior vena cava.

The model and the dog studies provided conclusive evidence for the accuracy of the method and the expression he developed for cardiac output. Additional corroborative data in dogs were provided by Goodyer et al.^[12] who used the dye and Fick methods as references. Two important improvements were reported in this study: (1) a thermistor was employed as the detector and (2) a correction factor was applied to correct for heat transfer through the catheter. The method was then used in man by Ganz et al.^[13] with 5% dextrose in water at 0°C as the indicator. To reduce the need for correcting for heat loss in the catheter, a thermistor at the injection site was used to measure the temperature of the indicator as it entered the right atrium. The cardiac output values were compared with those obtained using cardiogreen dye; excellent agreement was



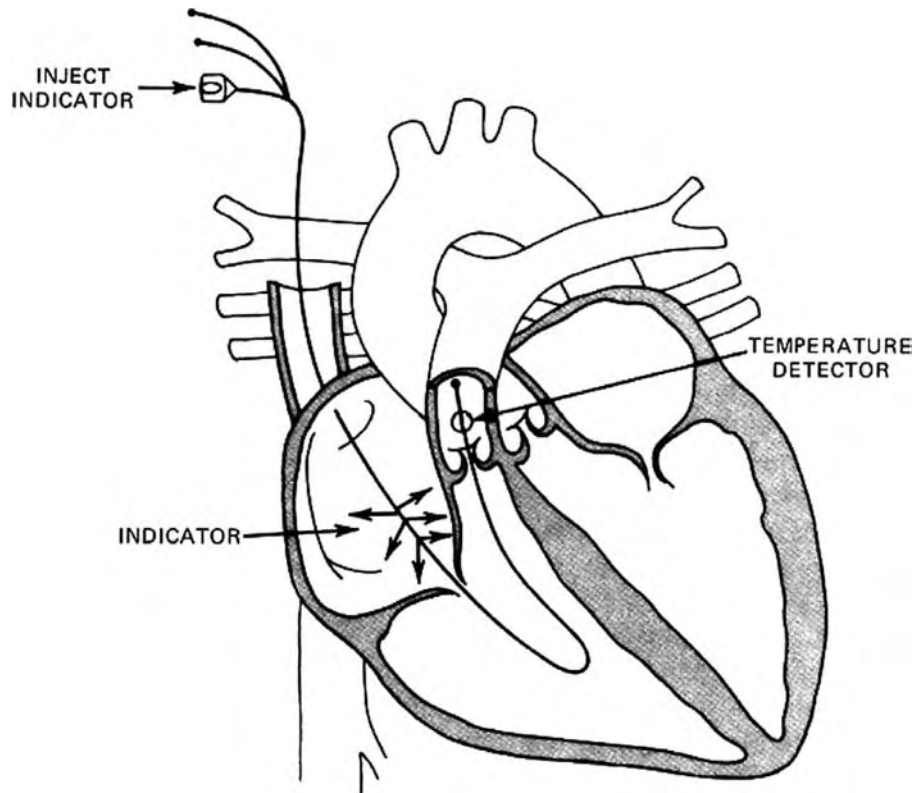


Fig. 8 Method of injecting the thermal indicator into the heart.

obtained. Finally, the technique was made practical enough for bedside use by Forrester et al.^[14] who developed a balloon-tipped catheter that incorporated a thermistor at the tip. A second lumen in the catheter allowed recording pressure from the catheter tip. A third lumen communicated with an injection port 30 cm proximal from the tip of the catheter in the pulmonary artery. The catheter was introduced into an arm vein, advanced into the right atrium, where the balloon was inflated. The blood flow guided the tip into the right ventricle, and its presence there was confirmed by recording the pressure. Then the catheter

was advanced into the pulmonary artery, its presence there being confirmed by the pressure recording. At this time the balloon was deflated and thermodilution curves were obtained by injection of the cold dextrose solution into the atrium via the third lumen which was in communication with the side port.

Verification Studies. After applying the correction factor for heat transfer through the #7F multilumen catheter, Forrester et al. obtained cardiac outputs on 40 critically ill patients. They validated the method with in vitro pump studies, and the average difference in flow amounted to about 3%. A later evaluation of the

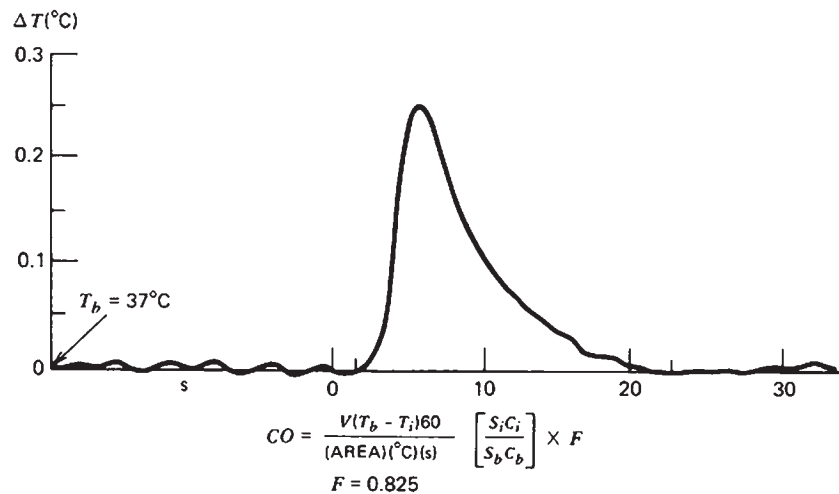


Fig. 9 The thermal indicator-dilution curve.

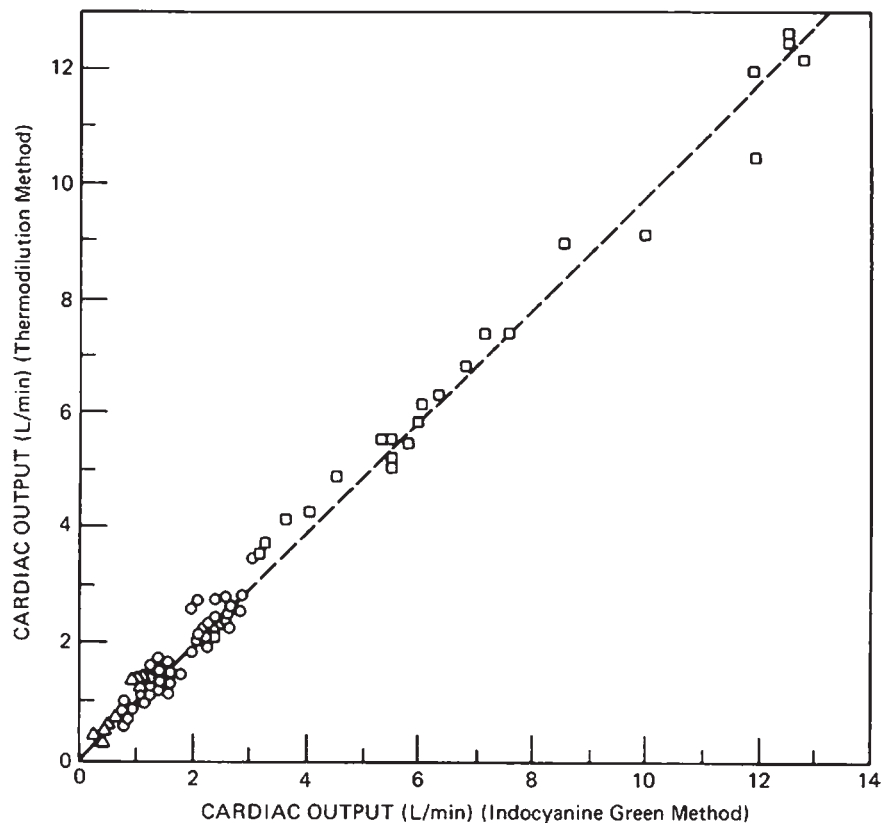


Fig. 10 Eighty-three simultaneous measurements of cardiac output using thermodilution (CO_{TD}) and indocyanine green (CO_{IG}) showing close agreement [$r = 0.994 \pm 2.79$ (SEM)]. Source: With permission from Elsevier (see Ref. [15]).

thermodilution method in man was presented by Weissel et al.^[15] who reported excellent agreement with values obtained with cardiogreen when a thermal correction factor was applied; Fig. 10 presents the data.

Calculation of Cardiac Output. The expression for calculating cardiac output with the thermal-dilution method contains several terms not found in other indicator-dilution formulas because of the physical properties of the indicator and blood. The indicator is a cold solution, which subtracts calories from the blood that it contacts; thus, the indicator is negative calories. Heat-exchange calculations depend on the mass of material, its specific heat, and the temperature difference through which the material is raised or lowered. In fact, the product of these three quantities is the number of calories exchanged. For this reason the thermodilution equation must include specific heat terms. In addition, with the thermal-dilution method, a known volume (not mass) of cold liquid is injected. Volume can be converted to mass by multiplying by specific gravity. Therefore, since the thermodilution method is based on heat exchange measured in calories, the flow equation must contain terms for the specific heat and the specific gravity of the indicator and blood.

The expression usually employed when a #7F thermistor-tipped catheter is employed and 5% dextrose is injected into the right atrium is as follows:

$$CO = \left(\frac{V(T_B - T_i)60}{A} \right) \left(\frac{S_i C_i}{S_b C_b} \right) F$$

where V = Volume of indicator injected in mL, T_B = Temperature (average) of pulmonary artery blood in degrees centigrade before injection of the indicator, T_i = Temperature of the indicator ($^{\circ}C$), 60 = Multiplier required to convert mL/s into mL/min, A = Area under the dilution curve in (s) ($^{\circ}C$), S = Specific gravity of indicator (i) and blood (b), C = Specific heat of indicator (i) and blood (b). The ratio $S_i C_i / S_b C_b = 1.08$ for 5% dextrose and blood of 40% packed-cell volume, F = Empiric factor employed to correct for heat transfer through the injection catheter wall. For a #7F catheter, $F = 0.825$ (Weissel et al.^[15]).

Entering these factors into the expression gives:

$$CO = \frac{V(T_B - T_i)53.46}{A}$$

where CO = cardiac output in mL/min, $53.46 = 60 \times 1.08 \times 0.825$.

To illustrate how a thermodilution curve is processed, cardiac output will be calculated using the dilution curve shown in Fig. 9:

$V = 5$ mL of 5% dextrose in water $T_B = 37^{\circ}C$ $T_i = 0^{\circ}C$
 $A = 1.59^{\circ}C$ s

$$CO = \frac{5(37 - 0)53.46}{1.59} = 6220 \text{ mL/min}$$

Despite the potential source of error due to heat loss through the injection catheter from which the empiric



factor F is derived, the thermal dilution method is remarkably reliable. Correlation values for cardiac output with cardiogreen dye in patients were reported by Weissel et al.^[15]; these data are shown in Fig. 10. The high degree of agreement using the empiric factor ($F=0.825$) is impressive. Perhaps the most attractive feature of the thermal-dilution method is the minimal recirculation, which makes it easy to process the dilution curve by hand or a simple computer. In low-output situations, conventional dilution curves are obscured by recirculation and the endpoint is not easily determined. The lack of recirculation artifact with the thermal-dilution method makes it attractive in low-output situations. However, in very-low-output situations, there may be significant indicator heat loss to vessel walls, which will result in an overestimate of cardiac output.

Diffusible Indicators. In many instances, diffusible indicators, which leave the vascular bed via the capillaries, are used to measure cardiac output. Saline and dextrose 5% in water (D5W) are examples, and whether cardiac output is overestimated depends on the way the indicator is used. The determining factor is the presence or absence of a capillary bed between the point of indicator injection and measurement. If there is an intervening capillary bed, the indicator loss therein will result in an overestimate of cardiac output. However, if there is none, and good mixing is achieved, cardiac output will not be overestimated.

The foregoing explains why D5W, a diffusible indicator, can be used without error because injection is in the right atrium and measurement occurs in the pulmonary artery. Grubbs et al.^[16] has used saline, another diffusible indicator, to measure cardiac output without error by injecting it into the right atrium and detecting it in the pulmonary artery. Likewise, we (Geddes et al.^[17]) measured left-sided output without error by injecting the saline in the left ventricle and measuring the dilution curve in the descending aorta.

If a diffusible indicator is injected into the venous system and the dilution curve recorded on the arterial side, an overestimate of 15% is typical. Chinard et al.^[18] predicted this situation, which was quantified by Grubbs et al.^[16] using saline.

IMPEDANCE CARDIOGRAPHY

The impedance cardiograph is an instrument that employs four circumferential thoracic electrodes to measure the pulsatile changes in resistance (impedance) offered to the flow of high-frequency alternating current to obtain a signal related to stroke volume. The pulsatile change in impedance is usually displayed on a graphic record, which can be processed to provide a value for stroke volume. At present, impedance

cardiography is one of the very few noninvasive methods that provides a measure of stroke volume, which when multiplied by heart rate is cardiac output.

Principle

Fig. 11 illustrates the principle employed in impedance cardiography. Low-intensity (e.g., a few milliamperes), high-frequency (e.g., 100 kHz) sinusoidal alternating current from a constant-current generator (I_k) is applied to metal-ribbon electrodes placed around the neck and abdomen (1,4). With this arrangement, the same current flows through the subject irrespective of the subject size and electrode impedance.

The flow of current through the thoracic tissues produces a potential difference detected by the two inner electrodes (2, 3), which are also metal ribbons placed around the thorax so that the heart is between them. Any variation in conducting properties of the thoracic contents will alter the potential difference between electrodes 2 and 3. Inspiration increases, and ejection of the stroke volume from the right and left ventricles decreases, the impedance between the potential-measuring electrodes. The respiratory induced change is much larger than the cardiac-induced change. Therefore, to obtain a clear cardiac signal from electrodes 2 and 3, the subject must be breathing slowly or hold his breath for a few heart beats.

Although band electrodes were used in early experimental studies, present systems for clinical measurements use standard ECG spot electrodes.

Fig. 12 illustrates the relative magnitude of the respiratory and cardiac impedance changes and illustrates typical cardiac impedance signals obtained during breath-holding.

High-frequency alternating current is used for impedance cardiography for reasons of safety. The

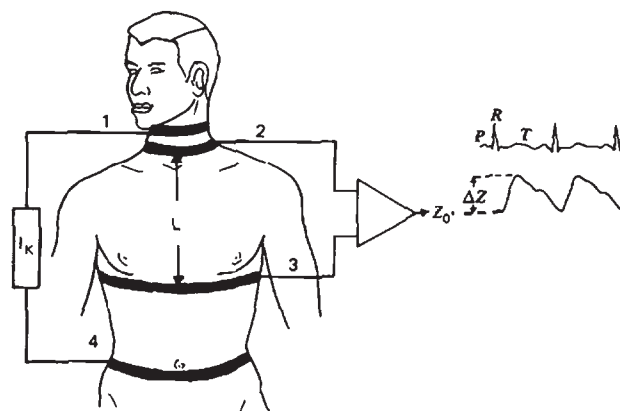


Fig. 11 The tetrapolar method of recording the thoracic impedance cardiogram (Z_0 and ΔZ).

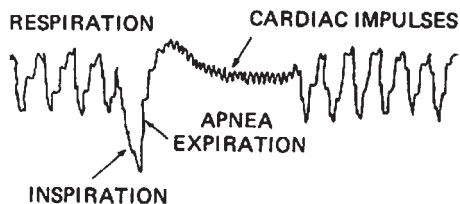


Fig. 12 Respiratory and cardiac impedance changes.

change in thoracic impedance provides a high-frequency signal of varying amplitude appearing between electrodes 2 and 3. The varying alternating voltage is amplified and demodulated to provide an analog signal that can be recorded graphically; Fig. 13 is an example.

Calculation of Stroke Volume

Various methods have been used to process the impedance cardiogram to obtain a value for stroke volume. The two that are presently used will be described. It is important to recognize that the peak-to-peak value for ΔZ illustrated in Fig. 13 is not the quantity used to calculate stroke volume, although it is related. The cardiac-induced impedance change reflects a change in impedance that is the difference between the inflow of blood to the thoracic vessels (from the heart) and the outflow of blood from the thoracic vessels to all other regions of the body. Therefore, to estimate the true volume change that would occur in the absence of thoracic outflow, the maximum rate of impedance

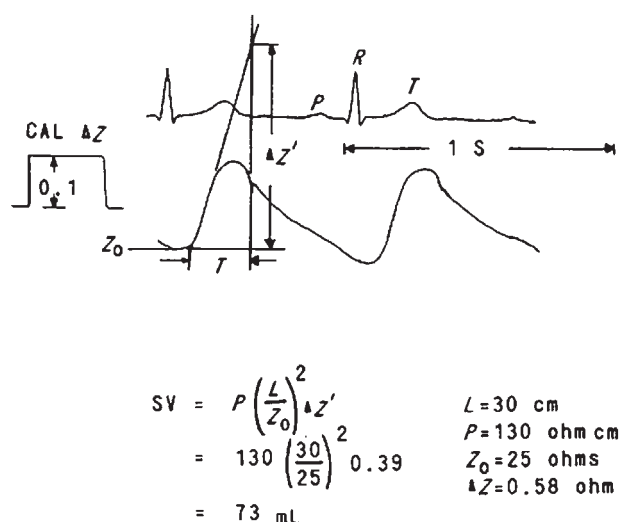


Fig. 13 The impedance cardiogram pulse and its relation to the R and T waves of the ECG, along with a sample calculation of stroke volume.

change during the early ejection period is determined and used to obtain the impedance change that would have occurred with no outflow from the thoracic vessels. The rationale underlying this technique is that outflow from the thoracic vessels is minimum at the beginning of cardiac ejection. Fig. 13 illustrates the cardiac impedance pulse and the graphical extrapolation technique used to obtain the corrected value ($\Delta Z'$) for entry in the equation developed by Kubicek et al.^[19]:

$$SV = \rho \left(\frac{PL^2}{Z_0^2} \right) (\Delta Z')$$

In this expression ρ is the resistivity of blood in ohm-cm at body temperature, L is the distance between the voltage-measuring electrodes (2, 3) in cm, and Z_0 is the basal impedance of the thorax in ohms, which is the baseline on which the pulsatile impedance rides, and $\Delta Z'$ is the extrapolated impedance change in ohms, as shown in Fig. 13.

Because extrapolation of the rising phase of the cardiac-induced impedance change is open to subjective error, it is often more convenient to include an electrical differentiation circuit in the impedance cardiograph. Thus, two output signals are provided: one is the cardiac-induced impedance change (ΔZ) and the other is the time derivative or differential (dZ/dt) as shown in Fig. 14. By knowing the maximum rate of change of impedance ($(dZ/dt)_{\max}$) and the period over which it occurs, namely, the ejection period (T), the following expression can be used to calculate

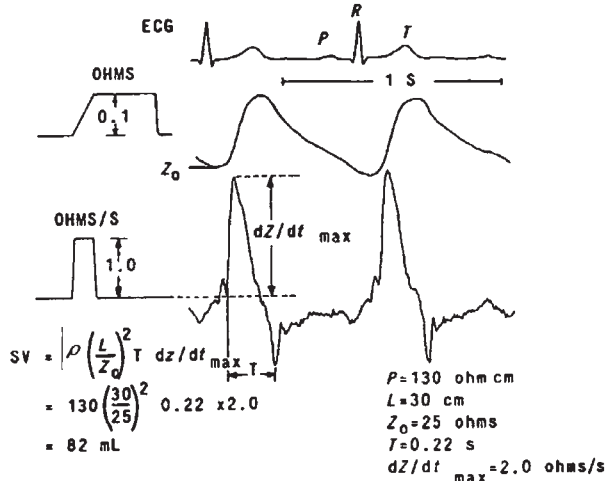


Fig. 14 The impedance cardiogram and its time derivative (dZ/dt), along with a sample calculation of stroke volume.



Table 7 Resistivity (ρ) of blood at 37°C

Species	Resistivity (ohm-cm)	Correlation coefficient
Dog	$55.5e^{0.024H}$	0.98
Human	$51.6e^{0.024H}$	0.97
Baboon	$57.5e^{0.019H}$	0.99
Cow	$59.9e^{0.016H}$	0.97
Goat	$58.3e^{0.015H}$	0.99
Cat	$53.9e^{0.023H}$	0.99
Monkey	$53.7e^{0.019H}$	0.99
Sheep	$56.9e^{0.019H}$	0.97
Camel	$56.4e^{0.016H}$	0.98
Average for nine species	$55.2e^{0.02H}$	0.96

Data obtained for packed-cell volumes (H) ranging from 0 to 50%.

stroke volume:

$$SV = \rho \left(\frac{L^2}{Z_0^2} \right) (T) \left(\frac{dZ}{dt} \right)_{\max}$$

Calibration signals in the impedance cardiograph allow quantification of the record to obtain Z_0 and $(dZ/dt)_{\max}$. Fig. 14 illustrates a typical record. The value for the resistivity (ρ) of blood depends on the percent packed-cell volume (H). For humans, the resistivity in ohm-cm was given by Geddes and Baker (20) as:

$$P = 51.6 e^{0.024H}$$

Often it is undesirable to draw a blood sample to measure the packed-cell volume (H) or resistivity (ρ). In such a case it is often assumed that a normal packed-cell volume of 40% exists. Therefore, the resistivity at 37°C is 130 ohm-cm. Table 7 presents values for a variety of species.

Verification Studies

Of particular importance is the relationship between cardiac output measured by impedance cardiography and by conventional techniques such as the indicator-dilution and Fick methods. Table 8 presents a summary of the validation studies to 1974. In some of these studies, the correct value for blood resistivity was used; in others an assumed value or a room-temperature value was entered into the equation. Nonetheless, it would appear that, in general, impedance cardiography provides a value for cardiac output that agrees with the true value (as obtained by an accepted method) or overestimates cardiac output by as much as

about 25%. A comparative study by Denniston et al.^[21] is shown in Fig. 15.

A large number of studies have shown that impedance cardiography tracks changes in cardiac output with good accuracy; this point was made by Denniston et al.^[21] in a study of normal subjects exercising on a bicycle ergometer. Fig. 16 illustrates the ability of the impedance cardiogram to track cardiac output as determined by the indicator-dilution method using indocyanine-green dye.

Mohapatra et al.^[22] pointed out the important fact that the degree of overestimation of cardiac output by the impedance method is dependent on the packed-cell volume. From a study on infants and adults they developed a correction factor (F), which when multiplied by the impedance cardiac output gives true cardiac output. Thus the Kubicek stroke volume expression becomes

$$SV = \frac{\rho L^2 T F}{Z_0^2} \left(\frac{dZ}{dt} \right)_{\max}$$

The value for F is shown in Fig. 17 and the mathematical expression for it is

$$F = 2.524 - 1.121 \log_{10} H$$

in which H is the packed-cell volume in percent. Note that for a packed-cell volume of 40%, the value for F is 0.83.

A practical consideration worthy of note is the value of L , the separation between the two (inner) potential-measuring electrodes (2, 3). Usually this dimension measured along the sternum is different than the magnitude when measured along midaxillary lines and on the back. In this situation it is of value to determine the mean value for L by making multiple measurements (chest, back, and along midaxillary lines). This technique was advocated by Denniston et al.^[20] Because stroke volume is usually overestimated, some advocate using the minimum value for L .

In conclusion, although cardiac-output values obtained by impedance cardiography usually overestimate true cardiac output, the technique is of considerable value in a variety of situations when invasive measurements cannot be made. It must be emphasized that the comparison studies made in humans were based on techniques (e.g., Fick and indicator-dilution) that averaged cardiac output over a minute or more, whereas the impedance method provided a value for stroke volume that was multiplied by average heart rate to obtain cardiac output. The fact that it is sometimes necessary for the subject to stop breathing for a few heart beats irritates some investigators. However, its ability to provide a value for stroke volume on a beat-by-beat basis is a highly attractive feature. Many newer instruments incorporate processing circuitry so that graphic recording is not needed and cardiac output is indicated on a digital display.

Table 8 Verification of cardiac output values determined by impedance cardiography

	Correlation coefficient	Standard error of estimate	Subjects	Reference method for A(CO)+B	ΔZ processing method	Principal investigator and year	Remarks
Cardiac output relationship (L/min)							
$Z(\text{CO}) = A(\text{CO}) + B$							
$0.942(\text{CO}) + 0.007$	0.959	0.026	Dog	Fick	Backslope	Allison (1966)	14–42 kg body weight
$1.124(\text{CO}) - 0.097$	0.976	0.107	Dog	Cardio-green	Backslope	Allison (1966)	14–42 kg body weight
$1.18(\text{CO})$	—	—	Human	Cardio-green	Max derivative	Kubicek (1966)	Resting and postexercise
$0.86(\text{CO}) + 2.93$	0.680	2.33	Human	Dye	Forward slope	Harley (1967)	24 cardiac patients
$0.73(\text{CO}) + 4.91$	0.579	2.34	Human	Radioisotope	Max derivative	Judy (1969)	
$0.71(\text{CO}) + 2.06$	0.83	—	Human	Cardio-green	Max derivative	Smith (1970)	Adult-tilt table
$0.90(\text{CY}) + 0.62$	0.94	—	Human	Cardio-green	Max derivative	Smith (1970)	Adult-tilt table
$0.8(\text{CO})$ to $1.2(\text{CO})$	0.95	—	Human	Fick	Double slope	Kinmen (1970)	25 patients
$0.8(\text{CO})$ to $1.2(\text{CO})$	0.84	—	Human	Fick	Double slope	Kinmen (1970)	67 patients
$0.8(\text{CO})$ to $4.3(\text{CO})$	0.58	—	Human	Radioisotope	Max derivative	Baker (1971)	17 normal adults
$1.0(\text{CO}) + 0.52$	0.68	1.63	Human	Cardio-green	Max derivative	Baker (1971)	10 normal adults
$0.96(\text{CO}) + 0.56$	0.66	1.54	Human	Cardio-green	Max derivative	Baker (1971)	21 normal adults
$1.43(\text{CO}) + 1.08$	0.923	0.593	Human	Cardio-green	Max derivative	Van De Water (1971)	One subject
Ca. $1.055(\text{CO})$	—	—	Human	Cardio-green	Max derivative	Lababidi (1971)	Children (cardiac patients)
$0.91(\text{CO}) + 0.218$	0.92	—	Dog	EM flowmeter	Max derivative	Baker (1971)	Aortic blood flow
Stroke-volume relationship (mL)							
$Z(\text{SV}) = P(\text{SV}) + Q$							
$1.42(\text{SV}) + 3.5$	0.94	—	Human	Pressure gradient	Max derivative	Bache (1969)	Cardiac patient
$1.42(\text{SV}) - 5.8$	0.96	—	Human	Pressure gradient	Max derivative	Bache (1969)	Cardiac patient
$0.93(\text{SV}) + 2.0$	0.91	—	Human	Pressure gradient	Max derivative	Bache (1969)	Cardiac patient
$0.44(\text{SV}) + 5.3$	0.58	—	Human	Pressure gradient	Max derivative	Bache (1969)	Cardiac patient
$1.37(\text{SV}) + 12.4$	0.62	—	Human	Pressure gradient	Max derivative	Bache (1969)	Cardiac patient
$0.65(\text{SV}) - 1.0$	0.78	—	Human	Pressure gradient	Max derivative	Bache (1969)	Cardiac patient
$0.77(\text{SV}) + 0.4$	0.68	—	Human	Pressure gradient	Max derivative	Bache (1969)	Cardiac patient
$1.53(\text{SV}) - 5.5$	0.87	—	Human	Pressure gradient	Max derivative	Bache (1969)	Cardiac patient
$0.92(\text{SV}) + 2.3$	0.94	1.9	Dog	EM flowmeter	Max derivative	Kubicek (1970)	Aortic blood flow
$0.90(\text{SV}) + 15.2$	0.87	—	Human	Cardio-green	Max derivative	Smith (1970)	Head up (8 subjects)
$1.01(\text{SV}) + 0.82$	0.96	—	Human	Cardio-green	Max derivative	Smith (1970)	Head up (8 subjects)

Source: From Wiley (see Ref. [20]).



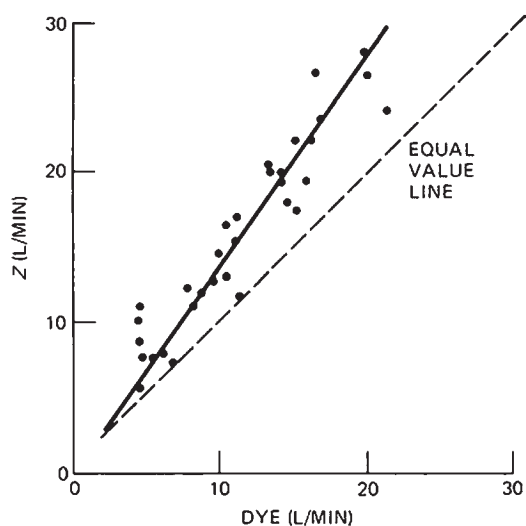


Fig. 15 The relationship between impedance cardiac output (solid line) and dye-cardiac output in exercising subjects. Source: Redrawn by permission from American Physiological Society (see Ref. [21]).

Contraindication

Impedance cardiography should not be applied to patients with implanted pacemakers. Using impedance cardiography to monitor a patient with an implanted pacemaker during surgery, Aldrete et al.^[23] reported that the patient's heart rate increased to 120/min. Drug therapy failed to reduce it and the ECG showed

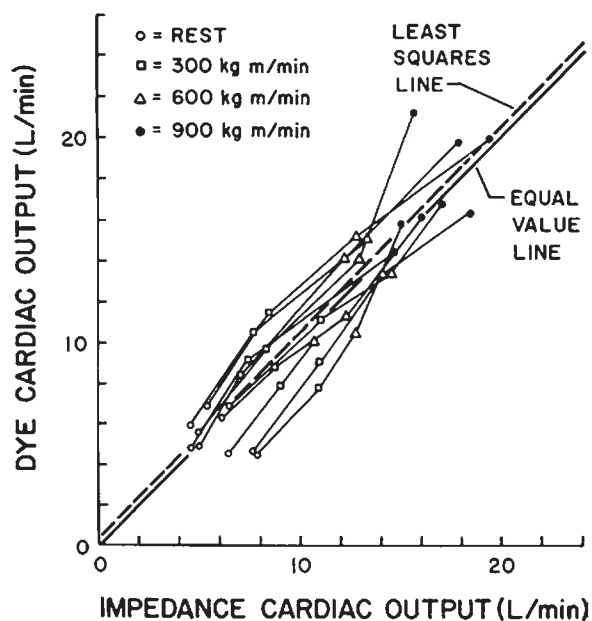


Fig. 16 The relationship between dye-cardiac output and impedance cardiac output (corrected) in subjects exercising on a bicycle ergometer. Source: Redrawn by permission from American Physiological Society (see Ref. [21]).

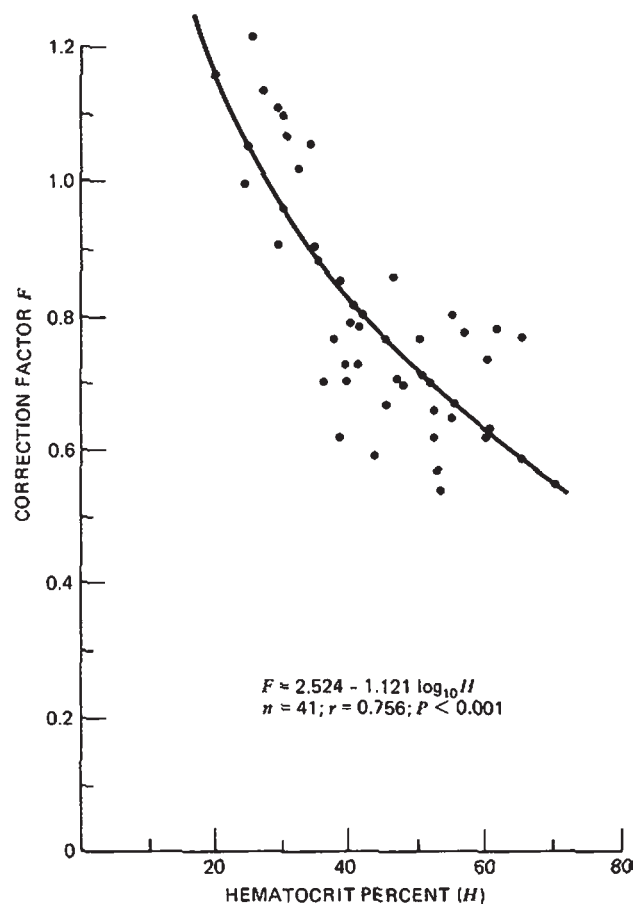


Fig. 17 The correction factor (F) reported by Mohapatra et al.^[22] for correcting impedance cardiac output values.

pacemaker spikes at 120/min. When the impedance cardiograph was disconnected, the heart rate returned to 72/min. The event occurred a second time.

Worthy of note is the fact that the 2.5 mA, 70-kHz current applied for impedance cardiography gained access to the pacemaker via the pacing electrodes, which are virtually in line with the 70-kHz current path. The pacemaker was of the programmable type and had been operating properly for 11 months prior to the event. It is clear that the 70 kHz current caused the pacemaker rate to increase from 72 to 120 beats/min. Neither the instruction manual for the pacemaker, nor that for the impedance cardiograph identified this hazard.

ARTICLES OF FURTHER INTEREST

Artificial Heart Fluid Dynamics: Positive Displacement Pumps; Blood Vessel Mechanics

REFERENCES

1. Spector, W.S. Ed.; *Handbook of Biological Data*; Saunders: Philadelphia, 1956.

2. Altman, P.L.; et al. Handbook of Circulation, NRC. National Academy of Science, Contract AF 33 (616)-972, 1960.
3. Fick, A. Uber die messung des blutstroms in den hertzventrikle. Verhandl. d. phys. Med. Ges zu Wurzburg **1870**, 2, XVI. [See Hoff, 1965].
4. Kovach, J.C.; Paulos, P.; Arabadjis, C. J. Thorac. Surg. **1955**, 29, 552–554.
5. Stewart, G.N. The output of the heart in dogs. Am. J. Physiol. **1921**, 57, 27–50.
6. Connolly, D.C.; Wood, E.H. Simultaneous Measurement of the appearance and disappearance of T-1824 (Evans Blue) in blood and tissue after intravenous injection in man. J. Appl. Physiol. **1954**, 7, 73–83.
7. Fox, I.J. Indicators and detectors for circulatory dilution studies and their application to organ or regional blood-flow determination. Circ. Res. **1962**, X (3), 447–471.
8. Wheeler, H.O.; Cranston, W.I.; Meltzer, J.I. Hepatic uptake and biliary excretion of indocyanine green in the dog. Proc. Soc. Exp. Biol. Med. **1958**, 99, 11–14.
9. Taylor, S.H.; Thorp, J.M. Properties and biological behavior of coomassie blue. Brit. Heart J. **1959**, 21, 492–496.
10. Swan, H.; Ganz, W.; Forrester, X. Catherization of the heart in man with use of a flow-through balloon-tipped catheter. N. Engl. J. Med. **1970**, 283, 447–451.
11. Fegler, G. Measurement of cardiac output in anesthetized animals by a thermodilution method. Q. J. Exp. Physiol. **1954**, 39, 153–164.
12. Goodyer, V.N.; Huvos, A.; Eckhardt, W.F.; Ostberg, R.H. Thermal dilution curves in the intact animal. Circ. Res. **1959**, 7, 432–441.
13. Ganz, W.; Donoso, R.; Marcus, H.; Forester, J.S.; Swan, H.J.C. A new technique for measurement of cardiac output by thermodilution in man. Am. J. Cardiol. **1971**, 27, 392–396.
14. Forrester, J.S.; Ganz, W.; Diamond, G.; McHugh, T.; Chanette, D.W.; Swan, H.J.C. Thermodilution cardiac output determination with a single flow-directed catheter. Am. Heart J. **1972**, 83, 306–311.
15. Weissel, R.D.; Berger, R.L.; Hechtman, H.B. Measurement of cardiac output by thermodilution. N. Engl. J. Med. **1975**, 292, 682–684.
16. Grubbs, D.S.; Geddes, L.A.; Voorhees, W.D. Right-side cardiac output determined with a newly developed catheter-tip resistivity probe using saline indicator. Jpn. Heart J. **1984**, 25 (1), 105–111.
17. Geddes, L.A.; Babbs, C.F. A new technique for repeated measurements of cardiac output during cardiopulmonary resuscitation. Crit. Care Med. **1980**, 8, 131–133.
18. Chinard, F.P.; Enns, T.; Nolan, M.F. Indicator-dilution studies with diffusible indicators. Circ. Res. **1962**, 10, 473–490.
19. Kubicek, W.G.; Kanegris, J.N.; Patterson, R.P.; Witsoe, D.A.; Mattson, R.H. Development and evaluation of an impedance cardiac output system. Aerosp. Med. **1966**, 37, 1208–1212.
20. Geddes, L.A.; Baker, L.E. *Principles of Applied Biomedical Instrumentation*; Wiley: New York, 1975.
21. Denniston, J.; Maher, J.T.; Reeves, F.T.; Cruz, J.C.; Gymerman, A.; Grover, R.F. Measurement of cardiac output by electrical impedance at rest and during exercise. J. Appl. Physiol. **1976**, 40, 91–95.
22. Mohaptra, S.N.; Costeloe, K.L.; Hill, D.W. Blood resistivity and its implication for calculation of cardiac output by the thoracic electrical impedance technique. Intens. Care Med. **1977**, 3, 1–5.
23. Aldrete, A.; Brown, C.; Daily, J.; et al. Pacemaker malfunction due to microcurrent injection from a bioimpedance noninvasive cardiac output monitor. J. Clin. Monit. **1995**, 11 (2), 131–133.



Cardiac Pacemaker

Rakesh K. Pai

Department of Internal Medicine, University of New Mexico School of Medicine, Albuquerque, New Mexico, U.S.A.

Fred M. Kusumoto

Electrophysiology and Pacing Service, Lovelace Medical Center, Albuquerque, New Mexico, U.S.A.

INTRODUCTION

The first asynchronous single-chamber permanent pacemaker was implanted during the late 1950s.^[1] During the past 50 years, permanent pacemakers have developed after many years of cooperation among physicians and biomedical engineers and fierce competition among device companies into an important and commonly used technology for treating slow heart rates. While the initial pacing systems were the size of a breadbox and used an electrical outlet as a power source, current devices weigh approximately 25 grams, are completely self-contained, and provide complex programmable options that mimic physiologic activity of the heart. In the year 2000, over 225,000 pacemakers were implanted in the United States and more than 600,000 worldwide.^[2] This article will review the anatomy and physiology of the cardiac conduction system, the components of a pacing system, pacemaker function, and the clinical indications for permanent pacing.

PHYSIOLOGY OF THE CARDIAC CONDUCTION SYSTEM

Cellular Activation

Cardiac cells (myocytes) undergo a series of characteristic changes in surface voltage when exposed to electrical stimulation.^[3] Atrial and ventricular myocytes are surrounded by a cell membrane that, among other things, provides an electrical insulation between the intracellular ionic concentrations (mostly K^+) and extracellular ones (mostly Na^+ and Ca^{2+}). The cell membrane consists of bilayer structures composed of molecules with a polar (charged) side and a nonpolar fatty side. Because ions are not very soluble in hydrophobic environments like the cell membrane, they cannot normally cross the membrane, thus maintaining a concentration gradient and a net membrane potential of about -85 mV. Cells can synthesize linear polymers of amino acids (membrane proteins) that can interact with the membrane. Some of these can span all the way through the membrane,

thereby forming aqueous pores (ion channels) that connect the intra- and extracellular and allowing the flow of ions. If an electrical stimulus is provided to resting cells, ion channels for Na^+ and Ca^{2+} open and the cell becomes quickly depolarized ($0-10$ mV) due to the inward flow of Na^+ and Ca^{2+} ions. Cellular depolarization initiates the release of intracellular stores of Ca^{2+} ions, which in turn leads to contraction of the myocyte. An outward flow of K^+ returns the cell to the polarized rest state. The whole excitable cycle for a human cardiac cell usually lasts for a few hundred milliseconds.

Specialized myocytes in the sinoatrial (SA) node and atrioventricular (AV) node (see below) have a special additional property of spontaneous depolarization. Gradual relative accumulation of cations on the inside of these cells leads to continuous regular depolarization of the cells. Depolarization of these “pacemaker” cells causes depolarization of atrial and ventricular cells and leads to repetitive electrical activation of the heart.^[4]

Cardiac Conduction System

The normal electrical impulse of the heart is initiated at the SA node, which was originally identified by Keith and Flack in 1907. The SA node is located high in the right atrium at the junction between the superior vena cava and the atrium; this structure serves as the natural “pacemaker” for the heart. The SA node is richly innervated by nerve fibers from the autonomic nervous system, which regulate the heart rate under a variety of physiologic conditions (Fig. 1).

When relaxed, the right atrium receives blood that has circulated through the body after delivering oxygen and nutrients, while the left fills with newly oxygenated blood returning from the lungs. After an electrical impulse leaves the SA node, it spreads throughout the atria, causing contraction that pushes the blood through valves into the relaxed ventricles. The electrical wavefronts from both atria converge at the atrioventricular (AV) node, a small subendocardial structure located within the interatrial septum. The AV node, like the SA node, is connected to the autonomic nervous

system, so that if SA node disease results in loss of pacemaker activity, the AV node can serve as a “backup” pacemaker, although at a slower heart rate than that of the SA node. The AV node produces a small delay in the electrical activation that allows the ventricles to be filled with the blood coming from the atria. After this delay, the activation is quickly spread to the ventricles through a specialized conduction system (Bundle of His, the right bundle, the left bundle, and the Purkinje fibers). This architecture allows rapid simultaneous activation and contraction of both the right ventricle, which pumps blood to the lungs to be oxygenated, and the left ventricle, which pumps oxygenated blood to the body; ventricular activation is usually complete within 100 ms.

The electrical impulse just described can be altered or disrupted in many different ways. In this article, our primary concern is slow heart rates, also known as bradycardia. When the heart rate is too slow, tissues throughout the body may not receive sufficient oxygen and other nutrients. Bradycardia can be due to two physiologic causes: 1) decreased automaticity of normal pacemaker tissue (usually the SA node) or 2) blocked conduction of the electrical impulse (usually at the level of the AV node or Bundle of His).

Electrocardiogram

The electrocardiogram (ECG) represents the sum of all the cardiac electrical activity measured from the surface (Fig. 1). The instantaneous wave of depolarization can be followed throughout the myocardium by looking at the ECG. The electrical signals generated by the SA node are not large enough to be measured by the ECG. However, as the atria are activated, an electrical signal of 1–3 mV can be recorded on the ECG. This first deflection is called the P wave. The relatively small mass of the AV node, the His bundle, and the right and left bundles are also not of sufficient amplitude to be recorded on the ECG, but ventricular activation results in a QRS complex. The QRS complex is relatively large (7–12 mV), because of the larger number of ventricular myocytes. A final wave called the T wave is due to ventricular repolarization. The duration of ventricular contraction can be estimated from the QT interval, which is measured from the beginning of the QRS complex to the end of the T wave. If the SA node fails to fire normally, a P wave will not be seen (since the atria are not depolarized) and a pause is observed (Fig. 2). If AV node conduction is completely blocked, P waves will be seen (since the atria are depolarized by the SA node), but QRS complexes will not be seen after the P waves. Instead, QRS complexes will occur at a slower rate due to ventricular depolarization from “pacemaker” cells distal to the site of block.

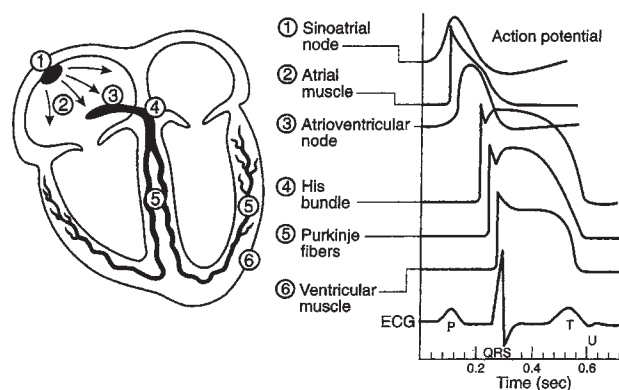


Fig. 1 Conducting system of the action potentials of various tissue types and their correlation with the surface electrocardiogram. The cardiac impulse starts in the SA node and spreads throughout the atria, and the P wave is inscribed. The AV node and His bundle form the normal axis of atrioventricular conduction. The QRS complex is inscribed as the ventricles are depolarized. The T wave is due to ventricular repolarization. The duration of ventricular contraction or systole correlates with the QT interval (the interval from the beginning of the QRS complex to the end of the T wave). Diastole, or ventricular filling, occurs from the T wave to the beginning of the next QRS complex.

CARDIAC PACEMAKERS

Present-day permanent pacemakers are composed of two main components: the pulse generator that contains the battery, specialized circuitry, and the connector or header block; and one or two leads that transmit the electrical impulse from the pulse generator to atrial or ventricular tissue (Fig. 2).

Pacemakers function by repetitively delivering electrical impulses to atrial and ventricular tissue. This small amount of energy (1–2 mA) is sufficient to depolarize tissues and initiate the wave of depolarization and contraction. The stimulation or pacing threshold is defined as the minimal amount of energy or stimulus to consistently depolarize the heart muscle.^[5] The pacing threshold is measured in terms of both amplitude in volts (pulse strength) and the duration of time in milliseconds for which it is applied to the cardiac tissue. There is a predictable hyperbolic relationship between strength duration and pulse amplitude, resulting in the strength-duration curve. At short pulse durations, a small change in the pulse duration is associated with a significant change in the pulse amplitude required to achieve depolarization.

In addition to providing electrical stimulation to the myocardium, pacemakers can also monitor intrinsic cardiac activity. Sensing is the ability of the pacemaker to detect or sense native cardiac depolarization. Intrinsic cardiac signals (known as electrograms) are produced by the wave of electrical current through the myocardium and can be recorded by the pacemaker



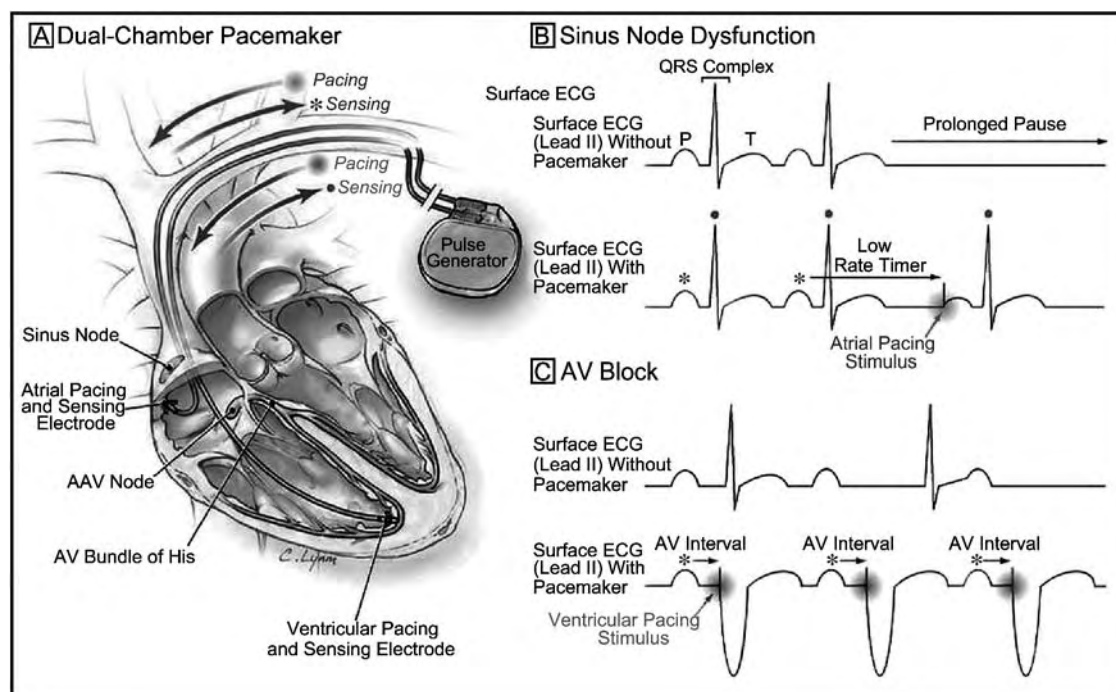


Fig. 2 Dual chamber pacing. A) Schematic of a dual-chamber pacemaker. Leads are usually placed in the right atrium and right ventricle. The leads are connected to a pulse generator usually located in the shoulder area. The leads are used to transmit current from the pulse generator (pacing) and also to transmit intrinsic electrical activity from the cardiac tissue (sensing). B) In sinus node dysfunction, loss of atrial activity leads to a prolonged pause. When a pacemaker is implanted, atrial activity (pink asterisk) and ventricular activity (blue circle) are sensed. If the low rate timer expires without sensed atrial activity, an atrial pacing stimulus is provided to maintain the heart rate. C) In complete AV block, there is no relation between atrial and ventricular activity. With dual-chamber pacing, atrial activity is sensed and the AV interval is initiated. When the interval expires, a ventricular pacing stimulus is provided. In patients with sinus node dysfunction, sensed ventricular activity due to normal AV conduction inhibits the pacemaker. Using this set of timers (low-rate and AV interval), dual-chamber pacemakers maintain AV synchrony regardless of the cause of bradycardia. (Used with permission from Ref. [2].)

electrodes and transmitted back to the pulse generator via the pacing leads. Reliable sensing depends mostly on the amplitude of the electrograms; pacemakers function optimally if the atrial electrogram is greater than 2 mV and the ventricular electrogram is greater than 7 mV. Ventricular electrograms are considerably larger than atrial electrograms due to increased ventricular mass.

Pacing Leads

Cardiac pacing leads are the weakest link in permanent pacing; most pacing failures are due to some malfunction of a lead component.^[6] These components include the electrode, the conductor, the insulation, and the connector or header pin. In order to function reliably, pacing leads must be able to survive in the harsh endocardial environment *in vivo*. Ideally the leads must be constructed of durable materials, which provide stability, flexibility, adequate conductive and resistive properties, and most importantly endocardial contact

within the heart. Not surprisingly, the “perfect” lead has yet to be developed and is a work in progress.

Most current lead designs utilize a tip electrode and a ring electrode 1–2 cm proximal to the tip. In almost all cases, the tip electrode is used as the cathode and the ring electrode as the anode. Optimal pacing and sensing characteristics favor pacing electrodes with a small radius and large surface area. By decreasing the lead radius, current flow can be minimized by providing increased electrode resistance and greater current density, which results in extended battery life and reduced pacing thresholds.

The composition of electrodes has changed in both shape and material composition over the years; to increase the electrode surface area, manufacturers have used pores, filament mesh, and various coatings with microspheres. The composition of the electrode is essential to its long-term performance; unfortunately, relatively few materials have proven satisfactory for pacing electrodes. Electrode degradation and corrosion can happen over time, and the “ideal” electrode must be biologically inert, resist degradation, and produce little

in the way of tissue inflammation and fibrosis at the electrode–endocardial interface. A number of metals, such as copper, zinc, mercury, nickel, lead, and silver, cause toxic reactions with the myocardium.^[7] Current materials in use are platinum–iridium, Elgiloy (a composite alloy of nickel, chromium, cobalt, manganese, molybdenum, and iron), and carbon. Typically after lead implantation, pacing thresholds increase due to the intense inflammatory response at the electrode–endocardial interface. Complex electrode designs that gradually release steroids have been developed that markedly attenuate the inflammatory response and are associated with reduced pacing thresholds.

Leads can be fixed to the endocardium via several methods. In passive fixation leads, small tines or fins positioned immediately behind the electrode are used to secure the lead within the trabeculae of the right atrium and/or ventricle. Active fixation leads use various types of barbs or screws that are imbedded within the myocardium to provide lead stability. Unlike passive fixation leads, actively fixed leads invade the endocardium and can cause transient elevation of the pacing thresholds at implantation. Both types of lead fixation, when implanted properly, have a very low incidence of dislodgment. Passive fixation leads can be placed in a relatively limited number of areas that have well-defined trabeculae (right atrial appendage and right ventricular apex), whereas active fixation leads can be placed anywhere within the right atrium or right ventricle.

The pacing lead conductor is commonly made of a multifilament design to facilitate tensile strength and reduce resistance to metal fatigue. Conductor materials are made of relatively noncorrosive alloys such as MP35N, which is a composite of nickel, chromium, cobalt, and molybdenum.^[7] The lead conductor is made of wire that conducts current from the pulse generator to the pacing electrode and sensed cardiac signals from the electrode(s) to the sensing amplifier of the pulse generator. The current generation of leads most commonly uses a coaxial design, with an inner coil extending to the distal electrode and the outer coil terminating at the proximal electrode. A layer of inner insulation separates the conductor coils.

Silicone and polyurethane are the two materials that have predominated lead insulation, each having its advantages and disadvantages.^[6] Silicone rubber has a performance record of more than 30 years; it is easily fabricated, molded, and repairable. Silicone's disadvantages are that it tears and cuts easily, is thrombogenic, has a higher friction coefficient in the blood, and is subject to cold failure. In contrast, polyurethane insulation exhibits very high tear and tensile strengths, high cut resistance, low friction in blood, high abrasion resistance, less thrombogenicity, and superior compressive properties. However, the disadvantages of

polyurethanes include relative stiffness, lack of repairability, sensitivity to manufacturing processes, their potential for environmental stress cracking (ESC), and metal ion oxidation (MIO). The two most common causes of polyurethane lead insulation failures are ESC and MIO, particularly in leads manufactured with Pellethane 80A and 55D. In ESC, if cracking propagates through the full thickness of the insulator to the conductor, lead failure can occur. In MIO, oxidative degradation of the polyurethane insulation occurs because of a reaction between the polymer and oxygen released by hydrogen peroxide produced from inflammatory cells at the electrode–endocardial interface.

The lead connector unites the lead to the pulse generator; contemporary leads are standardized to conform to international guidelines, which mandate that leads have a 3.2 mm in-line diameter. This mandate was established because some leads and connector (header) blocks were incompatible, requiring the development of multiple adaptors. Some patients who have functioning leads of the older variety (5.0 or 6.0 mm diameter designs) still require lead adaptors when the pulse generator is replaced. The development of an international lead connector standard has enabled physicians to combine various lead designs with different pulse generators.

Pulse Generator

The cardiac pulse generator is a machine having a power source and circuitry that produces the output stimuli. Energy stored within the device's battery travels through the connecting pathway to pace the myocardium, and energy then flows back to the pacemaker to complete the circuit. Specifically, all pulse generators include a power source, an output circuit, a sensing circuit, a timing circuit, and a header block with connector(s) to attach lead(s) to the pulse generator. Today, most pacemakers also include some type of rate-adaptive sensor, the ability to communicate with a programmer (telemetry), and the ability to record and store clinically relevant electrograms. Recent advances in pacemaker technology have exponentially increased the complexity of modern-day devices, while at the same time device size continues to decrease and the longevity continues to increase.

Over the last three decades pacemakers have utilized a number of battery technologies.^[8] There have been many different power sources ranging from nuclear, which generated toxic and radioactive elements; photoelectric cells; rechargeable nickel–cadmium cells that had to be charged on a weekly basis; and biogalvanic sources. Lithium iodine is currently the only power source being used in pulse generators today. Lithium is an anodal element and provides the supply of electrons;



iodine is the cathodal element and accepts electrons. The cathodal and anodal elements are separated by electrolyte, which serves as a conductor of ionic movement but as a resistive barrier to the transfer of electrons (lithium iodide). The pacing leads and the endocardium complete the circuit. In contrast to first-generation power sources, such as mercury-zinc cells, a lithium iodine cell has a more predictable and gradual decrease in battery voltage as it approaches depletion.

At the beginning of life for a lithium iodine battery, the cell generates approximately 2.8 V, which decreases to 2.4 V when approximately 90% of the useable battery life has been attained. The voltage exponentially degrades to 1.8 V when the battery reaches end of life (EOL). At the EOL, most devices lose programmability, frequently reverting to a fixed high output mode to maintain pacing and ensure patient safety. Lithium iodine cells have a relatively low but measurable self-discharge rate. Thus, even when a pulse generator is sitting in packaging on a shelf, not connected to a lead, its battery capacity is slowly being depleted. This has prompted manufacturers to provide device-specific shelf-life dates. Manufacturers also designate a device-specific voltage that triggers the elective replacement indicator (ERI), which alerts the physician that replacement of the pulse generator will be required within several months.

Two things determine the pacemaker longevity: battery capacity and device drain, which is mostly determined by usage. More specifically, longevity of any battery is determined by a number of factors, including chemical composition of the battery, size of the battery, external load factors, amount of internal discharge, and voltage decay characteristics of the cell. The external load factors that influence device longevity are pulse duration and amplitude, stimulation frequency, pacing lead impedance, and the amount of current required to operate the device and store diagnostic information. As a general rule, pacemaker longevity is primarily determined by how much actual pacing is required. Pacemaker longevity will be extended if a patient has intrinsic cardiac activity that is sensed by the pacemaker and inhibits delivery of pacemaker stimuli.

For pacing, the current drain is a function of output voltage, pulse width, pacing rate, and lead impedance. The pacing pulse is generated by charging of the output capacitor, which is then discharged to the cathode tip of the pacing lead. Since the voltage of a lithium iodine cell is 2.8 V, specialized circuitry is used to allow the physician to vary the voltage delivered to the endocardium. By programming the pacemaker to provide the minimal voltage required to reliably pace the heart, the physician can safely extend pacemaker longevity. Pulse generators are constant voltage, implying a constant voltage pulse throughout the pulse duration.

Actually, there is some decay of voltage from the leading and trailing edges of the stimulus impulse; the magnitude of the drop depends on the pacing impedance and pulse duration. The lower the impedance, the greater the current flow from the fixed quantity of charge on the capacitor and the greater the voltage drop throughout the pulse duration.

The intracardiac electrogram is amplified and filtered via the sensing circuitry of the pacemaker. The input impedance must be larger than the sensing impedance to avoid attenuation of the intracardiac electrograms. The signal is filtered and compared with a reference voltage, known as the sensitivity setting. Signals with an amplitude of this reference voltage or greater are sensed as true cardiac events and are then sent to the timing circuitry within the pulse generator (Fig. 3). Signals with amplitude below the reference amplitude do not reset the timing circuitry of the pacemaker. Sensing circuits also incorporate noise reversion circuitry that causes the pacemaker to revert to a noise reversion mode, such as asynchronous pacing whenever the rate of signal received by the sensing circuitry exceeds the noise reversion rate. This feature is key to preventing the inhibition of pacing when the device is exposed to environmental electromagnetic interference.

The timing circuitry of the permanent pacemaker uses a crystal oscillator that regulates the pacing functions with great accuracy. The output from the oscillator and signals from the sensing circuit are sent to a timing and logic control board that operates the internal clocks, which in turn regulates several different timing cycles of the pulse generator. The logic functions of present-day devices are controlled by microprocessors that form a specialized integrated circuit. The microprocessor serves as the central process unit of the device and utilizes two kinds of memory, RAM (random access memory) and ROM (read-only memory). The trend in pacemaker technology has been to use increasing amounts of RAM, which provides greater ease in storing data. Static RAM is gaining popularity among device engineers because critical data is not lost when the power is turned off.

CARDIAC PACEMAKER FUNCTION (PACING MODES)

In order to understand how permanent pacemakers function, it is essential to comprehend the coding system that has been developed by the NASPE/BPEG alliance (North American Society of Pacing and Electrophysiology/British Pacing Electrophysiology Group). For pacemakers, the NASPE/BPEG code consists of a four-position system using a letter in each position to describe a specific function of the pacing system. The four

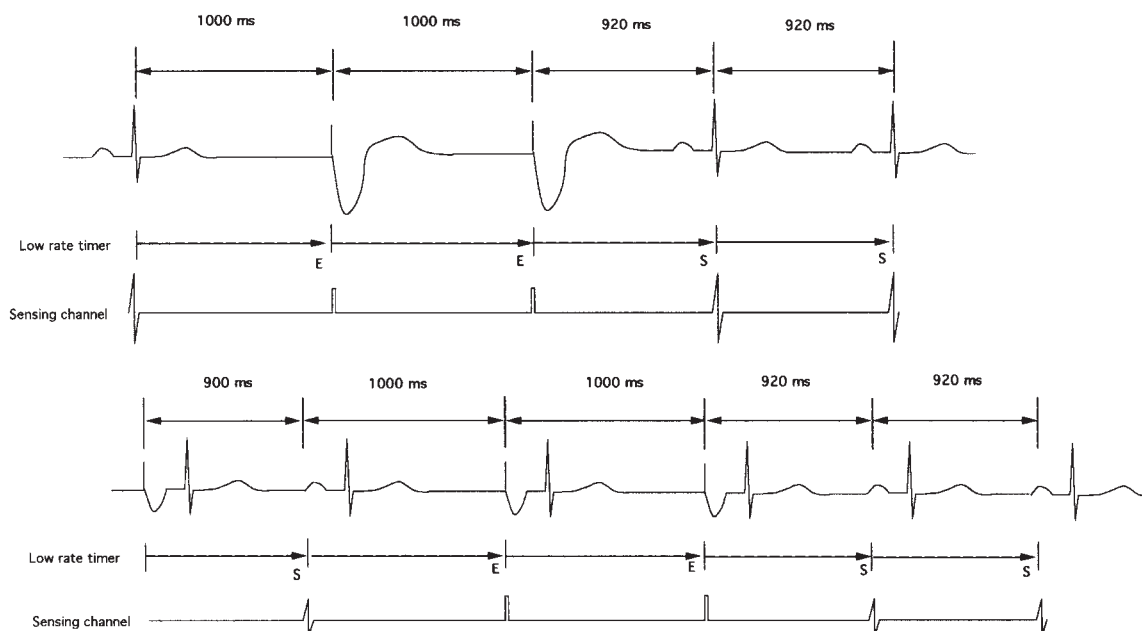


Fig. 3 Single-chamber pacing in a patient with sinus node dysfunction: differences between VVI and AAI pacing modes. *Top:* In the VVI pacing mode, intrinsic ventricular activation is sensed (S) by the pacemaker (first QRS complex), which inhibits the pacemaker and resets the low rate timer. The pacemaker is set at a base rate of 60 bpm so that after 1000 ms, when the low rate timer expires (E), and since the pacemaker did not sense any cardiac electrical activity, a pacing stimulus is delivered that leads to ventricular activation (second QRS complex) and the low rate timer is again reset. The low rate timer again expires without sensed cardiac activity and another pacing stimulus is delivered (third QRS complex) and the low rate timer is reset. Before the low rate timer expires, intrinsic cardiac activity occurs and is sensed by the pacing system that inhibits the pacemaker output and resets the low rate timer again. This continuous resetting of the low rate timer after paced or sensed events prevents the heart rate from dropping below a programmable level. *Bottom:* In the AAI pacing mode, cardiac activity in the atrium is sensed and the response to a sensed signal (S) is inhibition of pacing output. Since the lead is in the right atrium, when the low rate timer expires (E), atrial pacing occurs. The AAI pacing mode is effective for treating bradycardia due to sinus node dysfunction. However, if the patient develops atrioventricular conduction block either spontaneously or as a result of necessary medications, significant bradycardia may occur. (Used with permission from Phillips, R., Kusumoto, F.M., Goldschlager, N. Pacing in the elderly. *Am. J. Ger. Card.* 2002, in press.)

letters together describe the pacing mode of a specific pacemaker. The first position designates the chamber(s) paced. The letters utilized are **V**, **A**, **D**, and **O** to designate **V**entricle, **A**trium, **D**ual chamber, or **O**ff. The first generation of pacemakers could only pace; they had no ability to respond to or “sense” a patient’s intrinsic cardiac rhythm. Shortly after the introduction of the asynchronous pacemaker, it was discovered that the pacemaker could induce life-threatening tachyarrhythmias. Tachyarrhythmias occurred when a pacing stimulus was delivered during a vulnerable period of the cardiac cycle, usually at the end of ventricular repolarization (R on T phenomenon). The ability to sense intrinsic cardiac activity was an important advance in pacemaker function. The second letter of the NASPE/BPEG code denotes what chamber(s), if any, are being monitored for intrinsic cardiac activity (**V**, **A**, **D**, or **O**).

The third letter corresponds to the response of a pacemaker to a sensed event. The letters used in this position are **I**, **T**, **D**, and **O** to designate **I**nhibited,

Triggered (or **T**racking), **D**ual response, or **O**ff (no response). In the inhibited response, if the pacemaker senses an event, it will inhibit the pacemaker and reset the device to start a timing cycle called the lower rate timer. Delivery of a stimulus (pacing) occurs only if the lower rate timer expires without any sensed activity. For example, a pacemaker programmed to the VVI pacing mode would pace and sense the ventricle, and if a sensed QRS occurred before the pacemaker stimulus was due, it would withhold the stimulus (be inhibited) and reset for another cycle (Fig. 3). An AAI pacemaker is identical in all ways to a VVI pacemaker except for the fact that the pacing lead is in the right atrium rather than the right ventricle.

The triggered mode is often a source of confusion since it is not normally used in the current generation of pacemakers. Instead of inhibiting the output when the pacemaker senses an intrinsic event, a pacing stimulus is delivered when the sensed event occurs. In a VVT programmed device, the pacemaker will pace at the



lower rate timer unless a sensed QRS occurs. If a QRS is detected before the lower rate timer expires, the device will immediately deliver (or trigger) a pacing stimulus and reset the lower rate timer. On the surface electrocardiogram, pacing stimuli will be observed within the native QRS complexes.

Most pacemakers implanted today use the DDD pacing mode. While the DDD pacing mode is more complex since it requires the interaction between the atrial and ventricular signals and two separate leads, it provides the most physiologic responses in a variety of conditions. In the DDD pacing mode, if an atrial event is sensed, the atrial output is inhibited and a timer called the atrioventricular interval (AVI) is started in the ventricular channel. If the AVI expires without a sensed ventricular event, the pacemaker will assume AV conduction is blocked and deliver a ventricular pacing stimulus. If a spontaneous ventricular event occurs before the AVI expires, the pacemaker will assume intrinsic AV conduction is present and will not deliver a ventricular pacing stimulus. The lower rate timer is most commonly "atrially based" and will deliver an atrial stimulus if the lower rate timer expires without a sensed event. The DDD pacing mode provides physiologic depolarization of the atria and ventricles regardless of whether bradycardia is due to SA node dysfunction or AV block (Fig. 4).

The three most commonly used pacing modes are AAI, VVI, and DDD (Table 1, Figs. 2, 3, and 4). The AAI pacing mode, which uses a single lead placed in the

right atrium, prevents bradycardia in patients with sinus node dysfunction and maintains AV synchrony. However, if the patient develops AV conduction block, bradycardia is not prevented. The VVI pacing mode, which uses a single lead placed in the right ventricle, prevents bradycardia in both abnormal automaticity and AV block, but AV synchrony is not maintained. The DDD pacing mode, which usually requires implantation of two separate leads in the right atrium and right ventricle, prevents bradycardia and maintains AV synchrony in both sinus node dysfunction and AV block.

The fourth position in the NASPE/BPEG code designates the presence of certain programmable features of the device. The designations are presumed to be hierarchical, so if a device has a more advanced function, it is presumed to have all lower functions as well. The letter codes for the fourth position are **O**, **P**, **M**, **C**, and **R**, for **nO**-programmable features, **P**rogrammable for three or fewer parameters, **M**ulti-programmable (more than three parameters), **C**ommunicating with telemetry, and **R**ate adaptive (sensor driven). Since the current generation of pacemakers are all multiprogrammable and allow telemetry, the fourth letter is usually reserved to denote the presence of a rate modulation sensor.

Thirty years ago, a series of studies demonstrated that approximately 50% of patients that required permanent pacemakers for bradycardia at rest also had blunted heart rate responses to exercise.^[5,9] Manufacturers have developed a variety of sensors that

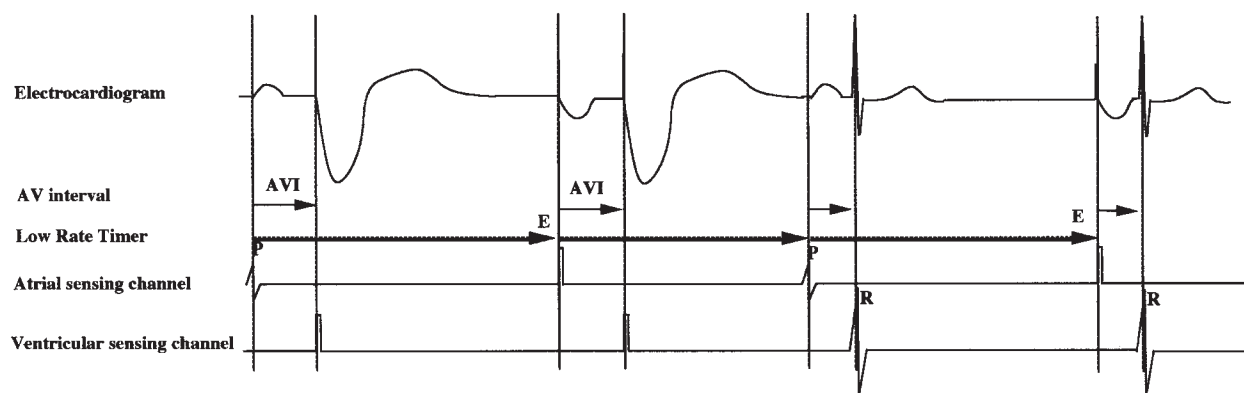


Fig. 4 Timing cycles and function of the DDD pacing mode. The first P wave is sensed by the pacing system, which inhibits the delivery of an atrial stimulus and begins the initiation of the AV interval (AVI). The AVI is designed to mimic the PR interval and allow efficient atrial filling of the ventricles. No intrinsic ventricular activity occurs so that when the AVI expires a ventricular stimulus is delivered (first QRS complex). The ventricular stimulus initiates the low rate escape interval. When the low rate timer expires (E), an atrial output stimulus is delivered since no sensed atrial activity has occurred during this interval. After the atrial stimulus the AVI begins, and when the AVI expires a ventricular stimulus is delivered (second QRS complex) and the low rate timer is reset. Before the low rate timer expires, atrial activity is sensed (P in the atrial sensing channel), which inhibits the delivery of an atrial stimulus and begins the AVI. However, before the AVI expires, intrinsic atrioventricular conduction occurs, which results in a spontaneous QRS (R in the ventricular sensing channel) that is sensed by the pacing system and causes inhibition of the ventricular output and again resets the low rate timer. The low rate timer expires again and an atrial output stimulus is delivered. Intrinsic atrioventricular conduction results in a spontaneous QRS (fourth QRS complex), which causes inhibition of the ventricular output and again resets the low rate timer. (Used with permission from Phillips, R., Kusumoto, F.M., Goldschlager, N. Pacing in the elderly. *Am. J. Ger. Card.* 2002, in press.)

Table 1 Commonly encountered pacing modes

Mode	Advantages	Disadvantages	Usual clinical uses
AAI (R)	<ul style="list-style-type: none"> • Requires only a single lead • Simple 	<ul style="list-style-type: none"> • Slow ventricular rates may develop if AV block occurs 	<ul style="list-style-type: none"> • Sinus node dysfunction without AV node dysfunction
VVI (R)	<ul style="list-style-type: none"> • Requires only a single lead • Simple 	<ul style="list-style-type: none"> • During pacing, if the atrial rhythm is sinus, atrioventricular synchrony is not preserved 	<ul style="list-style-type: none"> • AV block in patients with atrial fibrillation
DDD (R)	<ul style="list-style-type: none"> • AV synchrony is maintained for patients with sinus node and AV node disease 	<ul style="list-style-type: none"> • Requires two leads • More complex 	<ul style="list-style-type: none"> • Bradycardia due to sinus node disease or AV node disease

AV: atrioventricular.

Source: Used with permission from Phillips, R., Kusumoto, F.M., Goldschlager, N. Pacing in the elderly. *Am. J. Ger. Card.* 2002, in press.

respond to physical, electrical, or chemical stimuli that estimate metabolic need and vary the pacing rate depending on input from the sensor. In the United States, motion is the most commonly used sensor. In general, motion correlates with activity; motion is sensed by a piezoelectric crystal or by an integrated circuit silicon accelerometer located within the pulse generator. In a person with a pacemaker programmed to the VVIR pacing mode that uses a motion sensor, the pacing rate will be set to a lower rate at rest (60 beats per minute is commonly programmed), but with activity the pacemaker will gradually increase the pacing rate. The amount of activity required to increase the pacing rate and the rate of increase in the pacing rate can be separately programmed by the physician so that rate-adaptive pacing systems can be adjusted for use in both sedentary and active patients.

CLINICAL INDICATIONS FOR PERMANENT PACING

A detailed discussion regarding the clinical indications for permanent pacing is beyond the scope of this article. The decision to implant a permanent pacemaker requires interaction between the physician and patient. In an effort to standardize the indications for permanent pacing, the American College of Cardiology (ACC) and the American Heart Association (AHA) formed a joint task force to review the clinical data and recommend indications for permanent pacemaker implantation based on the clinical evidence.^[10]

Traditionally, pacemakers have been used for the treatment of bradycardia due to sinus node dysfunction and atrioventricular block. Other potential indications for permanent pacemaker implantation include autonomic nervous system disorders, cardiomyopathy, and the prevention of atrial fibrillation.

Sinus Node Dysfunction

Sinus node dysfunction (sick sinus syndrome and tachycardia-bradycardia syndrome) includes a broad spectrum of cardiac arrhythmias. Electrocardiographic manifestations of sinus node dysfunction include sinus bradycardia, sinus arrest, sinoatrial block, and paroxysmal supraventricular tachycardias alternating with bradycardia or even asystole. Sinus bradycardia can be a normal physiologic finding in the trained athlete, who commonly has awake resting heart rates in the 40-beats-per-minute range. These athletes may also have sleeping heart rates in the 30s with sinus pauses as long as 2.8 seconds.^[11] As emphasized in the following sections, the presence of symptoms, rather than a specific heart rate or duration of a pause, is the most important factor for deciding whether a permanent pacemaker should be implanted. Sinus node dysfunction disorders account for more than half of permanent pacemaker implants.^[12] Prior to pacemaker implantation it is important to correlate symptoms with the rhythm disturbance via a 24-hour ambulatory ECG monitor or a patient-activated event recorder; however, the correlation of symptoms with rhythm disturbance often proves difficult given the intermittent nature of these episodes. It should also be emphasized that a precise correlation of the symptoms of cerebral hypoperfusion with bradycardia may not be discovered. Decreased perfusion pressure to the brain itself may cause confusion which may allow the event to be unrecognized by the patient, and the sedentary lifestyle of the patient may lead to under-recognition of sinus node dysfunction by patients, family, and physicians.^[13] Currently, patients with documented symptomatic bradycardia as a result of sinus node dysfunction should receive a permanent pacemaker. In addition, patients in whom sinus node dysfunction is thought to be the cause of symptoms should be considered for permanent pacemaker implantation, even in the absence of direct correlation between symptoms and rhythm disturbances.



Atrioventricular Block

Atrioventricular (AV) block is electrocardiographically classified as first-, second-, and third-degree or complete heart block; anatomically, AV block is classified as supra-, intra-, or infra-His. In first-degree block, there is a 1:1 relationship between atrial and ventricular contraction, but the time between these events is abnormally long. Electrocardiographically, in first-degree AV block, the PR interval (measured from the beginning of the P wave to the beginning of the QRS complex) is greater than 220 ms. First-degree AV block is generally not considered an indication for pacing. However, some patients with severely prolonged PR interval (usually >300 ms) may require permanent pacing, as the result of nonphysiologic timing of atrial and ventricular filling and contraction; pacing may alleviate symptoms by restoring normal physiologic timing. In second-degree AV block, some but not all P waves are conducted to the ventricles. Second-degree AV block is subclassified as Type I when the PR interval prior to a dropped QRS complex is longer than the PR interval after the dropped QRS complex and Type II when the PR intervals before and after the dropped QRS complex are the same. The distinction is important, since Type I second-degree AV block is associated with supra-Hisian block and Type II second-degree AV block is associated with intra- or infra-His block. Intra-Hisian and infra-Hisian block are associated with a poorer prognosis due to the unreliability of subsidiary pacemakers below the His bundle. In practical terms, all symptomatic patients with second-degree AV block and asymptomatic patients with Type II second-degree AV block should receive permanent pacemakers. In third-degree AV block, no P waves are conducted to the ventricles. In general, all patients with third-degree AV block receive permanent pacemakers.

Other Indications for Pacing

Autonomic nervous system disorders are the causative mechanism in 10% to 40% of patients that experience transient loss of consciousness and postural tone (syncope). In these conditions, triggering of the autonomic nervous system results in a transient decrease in blood pressure due to bradycardia and peripheral vasodilation. The two most common examples of neurally mediated syncope are carotid sinus hypersensitivity syndrome and vasovagal syncope. The pathophysiology of vasovagal syncope is not completely understood but appears in part to be initiated by activation of myocardial mechanoreceptors, which leads to inhibition of efferent sympathetic tone and increased efferent parasympathetic tone, resulting in peripheral vasodilation and inappropriate bradycardia.

Pacing does not treat the peripheral vasodilation but may attenuate symptoms by preventing bradycardia. Several recent multicenter studies have suggested that pacemaker implantation may be beneficial in selected patients with severe symptoms.^[14,15] In older patients, carotid sinus hypersensitivity may be associated with syncope and unexplained falls. In carotid sinus hypersensitivity, stimulation of the carotid artery leads to bradycardia (pauses of more than three seconds) and hypotension (>50 mm Hg drop in systolic pressure). A recent multicenter study found that pacing therapy was associated with a fourfold decrease in recurrent falls in patients with carotid sinus hypersensitivity.^[16]

Pacing therapy may be useful for patients with advanced heart failure due to cardiomyopathy. Some patients with cardiomyopathy have dysynchronous left ventricular activation where the lateral wall of the left ventricle is activated very late relative to the inter-ventricular septum. Since left ventricular contraction is occurring sequentially rather than simultaneously, the efficiency of ventricular contraction is significantly reduced. In these patients, the duration of the QRS complex is often >150 ms, which reflects the slow rate of ventricular depolarization. Specially designed leads that are placed in the left ventricular venous system are used to restore a more normal left ventricular activation pattern. Pacemakers utilizing these leads have been associated with improved hemodynamics and quality-of-life in patients with severe congestive heart failure.^[17,18] As more information becomes available, the authors believe that pacing will become more commonly used for this condition.

Finally, pacing may have an adjunctive role for the treatment of atrial fibrillation. Preliminary data from the Dual Site Atrial Pacing for Prevention of Atrial Fibrillation (DAPPAF) trial suggests that atrial pacing reduces the frequency of atrial fibrillation episodes.^[19] However, additional studies will be required before pacing can be widely recommended in this situation.

CONCLUSION

It is clear that permanent pacemakers are truly an extraordinary technology, perhaps among the most important medical innovations of the twentieth century. As the number of elderly Americans with SA node and conduction system disease increases, it is inevitable that the number of permanent pacemaker implants will continue to rise. In addition, the indications for pacing will continue to expand beyond the treatment of bradycardia. Device and lead technologies will no doubt continue to advance, becoming even more reliable and durable. The evolution of the pacemaker provides an important example of the interplay

between medicine and engineering to develop an important tool for the treatment of cardiovascular diseases.

ARTICLES OF FURTHER INTEREST

Cardiac Bioelectricity; Defibrillators; Electrocardiography; Biopotential Amplifiers

REFERENCES

1. Elmqvist, R.; Senning, A. In *Implantable Pacemaker for the Heart*, Medical Electronics: Proceedings of the Second International Conference on Medical Electronics, Paris, June 24–27, 1959; Smyth, C.M., Ed.
2. Kusumoto, F.M.; Goldschlager, N. Device therapy for cardiac arrhythmias. *JAMA* **2002**, *287*, 1848–1852.
3. Kusumoto, F.M. Arrhythmias. In *Cardiovascular Pathophysiology*, 1st Ed.; Kusumoto, F.M., Ed.; Blackwell Science: Malden, MA, 1999; 54.
4. Keith, A.; Flack, M.W. The form and nature of the muscular connections between the primary divisions of the vertebrate heart. *J. Anat. Physiol. Lond.* **1907**, *41*, 172–189.
5. Holden, W.; McAnulty, J.H.; Rahimtoola, S.H. Characterisation of heart rate response to exercise in the sick sinus syndrome. *Br. Heart J.* **1978**, *40*, 923–930.
6. Mond, H.G. Pacing Leads. In *Cardiac Pacing for the Clinician*, 1st Ed.; Kusumoto, F.M., Goldschlager, N.F., Eds.; Lippincott Williams and Wilkins: Philadelphia, PA, 2001; 20–24.
7. Lloyd, M.A.; Hayes, D.L.; Friedman, P.A. Clinically Relevant Basics of Pacing and Defibrillation. In *Cardiac Pacing and Defibrillation: A Clinical Approach*, 1st Ed.; Lloyd, M.A., Hayes, D.L., Friedman, P.A., Eds.; Futuraco Publishing: Armonk, NY, 2000; 11–12.
8. Tyers, G.F.; Brownlee, R.R. Power pulse generators, electrodes, and longevity. *Prog. Cardiovasc. Dis.* **1981**, *23*, 421–434.
9. Wiens, R.; Lafia, P.; Marder, C.M.; Evans, R.G.; Kennedy, H.L. Chronotropic incompetence in clinical exercise testing. *Am. J. Cardiol.* **1984**, *54*, 74–82.
10. Gregaratos, G.; Abrams, J.; Epstein, A.E.; Freedman, R.A.; Hayes, D.L.; Hlatky, M.A.; Kerber, R.E.; Naccarelli, G.V.; Schoenfeld, M.H.; Silka, M.J.; Winters, S.L.; Gibbons, R.J.; Antman, E.M.; Alpert, J.S.; Hiratzka, L.F.; Faxon, D.P.; Jacobs, A.K.; Fuster, V.; Smith, S.C. ACC/AHA/NASPE 2002 guidelines for implantation of cardiac pacemakers and antiarrhythmia devices: Summary article. *Circulation* **2002**, *106*, 2145–2161.
11. Dreifus, L.S.; Michelson, E.L.; Kaplinsky, E. Bradycardia: Clinical significance and management. *J. Am. Coll. Cardiol.* **1983**, *1*, 327–338.
12. Lamas, G.A.; Pashos, C.L.; Normand, S.L.T.; McNeil, B. Permanent pacemaker selection and subsequent survival in elderly medicare pacemaker recipients. *Circulation* **1995**, *91*, 1063–1069.
13. Kusumoto, F.M.; Goldschlager, N.F. Cardiac pacing. *N. Engl. J. Med.* **1996**, *334*, 89–97.
14. Connolly, S.J.; Sheldon, R.; Roberts, R.; Gent, M. The North American vasovagal pacemaker study: A randomized trial of permanent cardiac pacing for the prevention of vasovagal syncope. *J. Am. Coll. Cardiol.* **1999**, *33*, 16–20.
15. Sutton, R.; Brignole, M.; Menozzi, C.; Raviele, A.; Alboni, P.; Giani, P.; Moya, A. Dual-chamber pacing in the treatment of neurally mediated tilt-positive cardioinhibitory syncope. *Circulation* **2000**, *102*, 294–299.
16. Kenny, R.A.; Richardson, D.A.; Steen, N.; Bexton, R.S.; Shaw, F.E.; Bond, J. Carotid Sinus Syndrome: A modifiable risk factor for nonaccidental falls in older adults (SAFE PACE). *J. Am. Coll. Cardiol.* **2001**, *38*, 1491–1496.
17. Cazeau, S.; Leclercq, C.; Lavergne, T.; Walker, S.; Varma, C.; Linde, C.; Garrigue, S.; Kappenberger, L.; Haywood, G.A.; Santini, M.; Bailleul, C.; Daubert, J.C. Effects of multisite biventricular pacing in patients with heart failure and intraventricular conduction delay. *New Engl. J. Med.* **2001**, *344*, 873–880.
18. Abraham, W.T.; Fisher, W.G.; Smith, A.L.; Delurgio, D.B.; Leon, A.R.; Loh, E.; Kocovic, D.Z.; Packer, M.; Clavell, A.L.; Hayes, D.L.; Ellestad, M.; Trupp, R.J.; Underwood, J.; Pickering, F.; Truex, C.; McAtee, P.; Messenger, J. Cardiac resynchronization in chronic heart failure. *New Engl. J. Med.* **2002**, *346*, 1845–1852.
19. Saksena, S. In *DAPPAF Trial Data*, Oral Presentation at the American College of Cardiology Meeting, Orlando, FL, March 18, 2001.



Cardiac Patch Engineering

Michael J. Yost

Robert L. Price

University of South Carolina, Columbia, South Carolina, U.S.A.

David G. Simpson

Virginia Commonwealth University, Richmond, Virginia, U.S.A.

Wentao Yan

Louis Terracio

New York University, New York, New York, U.S.A.

INTRODUCTION

The field of cardiac tissue engineering is advancing at a rapid pace. Cardiac function is intimately linked to the three-dimensional organization of the heart. Cell-based therapies designed to reconstruct dysfunctional or damaged domains of the heart must successfully recapitulate the structure and function of the intact tissue. To fabricate a three-dimensional prosthesis that exhibits the structural and functional properties of intact cardiac tissue presents many challenges. In this article we provide an overview of the research in this field with emphasis on collagen scaffold-based engineering of a cardiac patch.

OVERVIEW

Normal Heart

The normal adult heart weighs about 300 grams and is an incredibly efficient and durable pump lasting up to 120 years. On average, the heart pumps roughly 6000 liters of blood per day. This blood supply ensures that all of the tissues and vital organs are sufficiently perfused, oxygenated, and supported nutritionally and that CO₂ and other waste products are effectively eliminated.^[1] To tissue-engineer a replacement heart, or even replacement components, is therefore a formidable challenge.

Heart Disease

Cardiac dysfunction currently accounts for 750,000 deaths annually in the United States, which is nearly twice the rate of all cancer deaths combined. The annual cost of ischemic heart disease is approximately \$100 billion dollars. Malfunction of the cardiac system has profound negative physiological consequences. Cardiac dysfunction can be broken down into five main categories: 1) Pump failure, which most

frequently consists of damaged muscle that contracts too weakly or inadequately, causing ineffective emptying of the chambers. Less frequently the muscle may develop the inability to relax fully resulting in the ineffective filling of the chambers. 2) Obstruction of flow is usually caused by the formation of a lesion that interferes with the proper functioning of one or more of the valves. These obstructions cause an excessive pressure drop reducing the available pressure to perfuse the tissues. They can also impair valve closure leading to regurgitation within a chamber. 3) Regurgitation of flow is typically a partial failure of one or more of the valves in the heart causing flow reflux and increased workload on the emptying chamber. 4) Conduction system malfunctions cause incomplete or asynchronous contraction of the cardiac myocytes, usually detected as an atrial or a ventricular fibrillation. 5) Breach of the circulatory system is a more acute condition than the others and is usually an externally inflicted wound or aneurysm of the aorta or ventricle. The patient loses sufficient blood volume to allow the heart to maintain blood pressure and the victim can eventually exsanguinate.^[1]

Any one of the above conditions is sufficient to transport the patient along the path toward congestive heart failure, regardless of the cause; the only long-term treatment for this end-stage condition is transplantation. However, this treatment option is available only on a limited basis and there are presently long lists of patients waiting for available hearts. The material available for the transplantation and correction of cardiac abnormalities in children is even more limited. In a variety of conditions, it is theoretically possible to correct the damaged myocardium by surgically excising the dysfunctional area and reconstructing that domain of the heart by implanting healthy cardiac tissue. In order to successfully carry out this type of treatment, two major hurdles must be overcome. The first is to identify a source of donor tissue that will not be rejected. The tremendous advances in stem-cell technology indicate that it will be feasible in the near

future to have a source of autologous cells for transplantation. The second hurdle is the focus of this chapter, the development of the technology necessary to grow in vitro a segment of myocardium composed of cardiac myocytes, extracellular matrix constituents, and interstitial cells that closely mimic the intact tissue.

BACKGROUND ON CARDIAC TISSUE ENGINEERING

In recent years there has been significant engineering and scientific effort expended in an attempt to shift the treatment paradigm from one of replacement and substitution via cardiac transplantation to one of regeneration via either in-vivo or in-vitro tissue engineering.^[2] Much of the effort has been focused on regeneration of the non-functional segments of the damaged myocardial wall, repair of damaged valves, and/or ventricular assist devices for both the right and left ventricles.

The two main areas of regenerative cardiac research are: 1) injection of cells in an effort to repopulate damaged areas of muscle, and 2) in-vitro tissue engineering of cardiac valves and patches of cardiac muscle, and engineered ventricular assist devices to restore the pressure differential necessary for normal cardiac function. The development of the biomaterials, fabrication techniques, and tissue/cell delivery systems and technology to engineer mechanical devices must proceed in parallel with stem-cell technology in order to realize a tissue-engineered treatment for heart disease.

Injection of Cells

A variety of cell sources to repair injured myocardium have been used including skeletal myoblasts,^[3,4] fetal and neonatal cardiomyocytes,^[5,6] and stem cells.^[7] The basic premise is that live cells can be deposited by injection into and around a developing cardiac scar. The cells will then engraft into the existing tissue and be guided by the microenvironment and the extracellular matrix to adopt the appropriate architectural arrangement. Signaling mechanisms provided by the microenvironment will then stimulate the cells to develop into functioning myocytes that will proceed to integrate and connect electrically to the surrounding functional tissue. Injection studies have demonstrated healthy and functioning transplanted cells within heterotopic areas as well as the areas of myocardial infarction. Additionally, some studies have noted minor improvements in regional systolic performance, compliance, peri-infarct perfusion, and global ventricular function.^[8-10] On the cellular level, grafted cardiomyocytes have the capacity to form gap-junctions with host myocardium.^[11] These results indicate that this is a promising area that needs further investigation.

Tissue Engineering of Constructs

The most popular in-vitro tissue-engineering approach is to fabricate a scaffold consisting of a biocompatible material in the desired geometry. The scaffold is then seeded with the desired cell types and the composite is allowed to form the requisite cellular phenotype, function, signaling, and intercellular connections necessary for function as an intact tissue. Much of the early work in tissue engineering used two-dimensional culture techniques. These techniques, while providing significant and useful information, failed to produce viable engineered constructs that could be used clinically in heart repair. The major drawback of these constructs was that they were too thin and lacked sufficient mechanical integrity to be clinically significant. These two-dimensional systems did demonstrate the significance of the extracellular matrix (ECM) composition and geometry to the fabrication of in-vitro tissues. The aligned collagen system of Simpson et al. shows the importance of collagen and that the collagen must be arranged in a parallel, fibrillar form to elicit the elongated rod-shaped phenotype typical of in-vivo cardiomyocytes.^[12] McDermitt et al. further demonstrated the importance of the ECM in micro printing laminin streaks on tissue culture plastic. This work showed that cardiomyocytes cultured on micro-patterned laminin streaks of appropriate line width and spacing can achieve an elongated rod-shaped phenotype and intercalated disks.^[13] These aligned cultures exhibit polarized electrical properties, a characteristic necessary for normal excitation and the generation of organized contractions.

The current thinking in tissue engineering of cardiac systems is that repair of areas affected by ischemic heart disease can be accomplished by the fabrication and implantation of either a muscle patch or, for more advanced forms of heart failure, a tubular ventricular-assist device. Both of these approaches require a three-dimensional tissue scaffold. The scaffold could provide temporary biomechanical support for the cells in a more clinically appropriate geometry and result in more mature cultures with respect to cellular differentiation and electrical properties.^[14]

The primary difficulties encountered with the early three-dimensional scaffolds were inefficient methods for cell seeding, poor cell distribution, poor cell-cell and cell-scaffold interactions, and standard culture conditions in dishes and flasks that were not able to support the cultures. Thus, they failed to elicit biologically appropriate phenotypes and tissue function or failed to survive. Several investigators overcame these problems with cell seeding in ground-breaking studies by attempting to culture cells in the microgravity environment of space.^[15] These studies showed that in an environment with reduced gravitational force, cell-cell



and cell-scaffold interactions permitted much higher density cultures with more in-vivo-like cellular phenotypes.^[15-18]

Bioreactors

The National Aeronautical and Space Administration (NASA) has developed a randomized gravity-vector rotating-wall bioreactor that can be used to transport sensitive biological experiments into space. This reactor randomizes the acceleration force generated by the launch of the space shuttle and minimizes its impact on these experiments. These NASA bioreactors are also being used in ground-based studies to minimize the influence of gravity on cell culture by mimicking the microgravity conditions found on the space shuttle in orbit.^[18] When combining a three-dimensional scaffold with cardiac myocytes in the bioreactor, a cardiac construct with a high cellular density and mature, elongated cells that contract spontaneously and synchronously develops.^[19] Further advantages to these systems have been an improved rate of cellular seeding efficiency, most likely due to an improvement in aerobic metabolism, and reduced cellular stress within the dynamic environment due to high bulk flow conditions with minimal shear stress.^[15] In addition, in-vitro cultures grown within a dynamic environment more closely resemble native tissue with respect to protein characteristics and ultrastructural features.^[16,17] Overall, the utilization of bioreactors has become a powerful research tool that improves the cultivation of cardiac myocytes.

Approach to Tissue-Engineering a Cardiac Patch

One approach to achieving a successful cardiac implant is to pattern the development of the cultured extracellular and cellular components after the intact heart. The logic justifying this approach is that structure and function are intimately associated; therefore, if we are able to recreate the structure, the necessary function should follow. So what can we learn from intact tissue? The intact heart is composed of a series of overlapping cellular layers.^[3,18] Within any given cell layer the rod-like myocytes of the myocardium are arrayed in parallel with one another along a common axis. This parallel alignment is essential for the efficient distribution of electrical signals throughout the myocardium and the ability of the heart to eject blood effectively into the vascular tree. As the individual layers of parallel myocytes develop within the heart during embryogenesis and begin to contract, the increased force must be contained and directed to the ventricular chambers.^[19] To direct these mechanical forces, a three-dimensional, stress-tolerant, connective tissue network forms early in

development.^[4] The coordinate elaboration of the ECM during the development of the multilayered myocardium indicates that the ECM plays a role in directing the formation of the complex fiber patterns of the heart by providing positional information to the differentiating myocytes. The lesson to be learned is that if the tissue engineering process develops a patterned ECM, the cellular components will develop the same pattern. Thus, the composition and arrangement of this connective tissue network is essential.

The cardiac ECM is composed primarily of interstitial collagens. The fibrils of this collagen network surround the myocytes and tether adjacent cell layers together (Fig. 1). It is now evident that the changes in the ECM that occur in development and disease represent a dynamic adaptation to physiological change. Any perturbation that alters the normal arrangement of the myocytes within the myocardium, as the result of either a congenital defect or a pathological condition, alters the efficiency of electrical conduction and the ability of the heart to generate contractile force. To correct these defects, it is necessary to replace the tissue with an engineered construct that mimics the electrical and mechanical properties of the normal tissue. As a consequence of the functional characteristics and architecture of the intact heart, any cell-based procedure that is used to repair a segment of damaged myocardium must consist of myocytes that exhibit the functional and structural properties of the surrounding myocardium.

In addition to its structural role in the heart, the ECM also functions as a repository for signal molecules such as growth factors, cytokines, and extracellular proteases. The complex interplay of these biochemical factors and the mechanical events associated with cardiac function direct the formation of the heart and biochemically

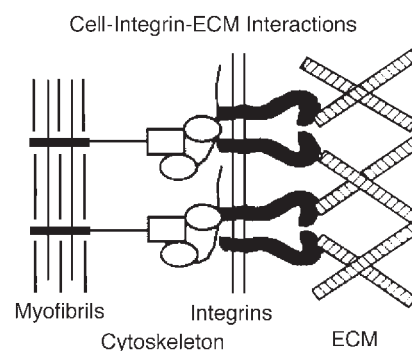


Fig. 1 A schematic drawing representing the physical communication that exists between the intercellular and extracellular environments in a cardiac myocyte. The connection can be traced from the myofibrils, across the cytoskeleton-integrin complex, and out into the ECM. This connection is essential for the formation and maintenance of normal cardiac myocyte phenotype.

link the myocytes to the other organ systems, enabling the heart to adapt to a variety of physiological conditions.^[6-8] Signals that originate in the ECM play an important role in development as well as in the adaptation to physiological signals throughout life.^[9]

Structural and Functional Arrangement of the ECM in the Heart

The ECM of the heart consists of a complex mixture of collagens, proteoglycans, noncollagenous glycoproteins, and growth factors that are produced and secreted into the extracellular environment by the fibroblasts. However, the principal structural component of the ECM is collagen. The collagen-based network of the heart appears to direct the expression of the rod-like cell shape by cardiac myocytes and the formation of the three-dimensional structure of the heart. The chemical identity and arrangement of this network is subject to change during development and in response to different physiological conditions.^[20] During neonatal development the collagen-based matrix is rapidly elaborated and develops into a stress-tolerant network that assists in distributing the mechanical forces associated with cardiac function.^[21] The collagen network of the heart has been demonstrated to control the elastic properties of the ventricular wall.^[12,22-24] In the adult, several factors, including nutritional factors, mechanical stimulation and biochemical feedback, have been shown to regulate the expression and accumulation of collagen.^[25]

Regulation of Myocyte Phenotype by the ECM

Most in-vitro culture methods for isolated cardiac myocytes have shown that it is necessary to use purified components of the ECM to get isolated cells to attach and adapt to the culture environment.^[26] While it is known that these ECM components are essential to promote the attachment and survival of myocytes in vitro, it has only recently been documented that the arrangement of a particular ECM component can dictate the shape that a cardiac myocyte will express in vitro (discussed in preceding text).^[2]

Interaction of Collagen and Myocytes; Integrins

In primary cell culture the specific interaction of collagen with receptors of the integrin family controls many aspects of the processes that regulate the shape that a cardiac myocyte expresses under a given set of culture conditions. Many of these phenotypic signals appear to be communicated across the plasma membrane by members of the integrin family of ECM receptors. The

functional integrin receptor consists of an α and a β chain that are non-covalently linked to one another. Different α and β chains interact with one another to form receptors that bind to components of the ECM. Within the intracellular space, the cytoplasmic tail of the integrin complex interacts with components of the cytoskeleton, the elaborate internal scaffolding that stabilizes myofibrillar structure and cell shape. In the heart during neonatal life, the α -1, β -1 integrin is the predominate isoform. However, α_2 , α_3 , α_5 , α_7 , and β_1 isomers have been identified and are present at varying concentrations under different physiological conditions and states of development.^[27]

In myocytes, collagen associates with the cell surface in two distinct ways: 1) perpendicular at a site near the Z band (Fig. 1), and 2) lateral associations parallel to the surface of the myocyte. Members of the integrin family mediate the first type of attachment to the surface. These specialized receptors transduce biochemical and mechanical signals from the ECM to the internal milieu of the cell.^[28] The integrins are coordinately synthesized with ECM components,^[29] associated with regulation of the cardiac myocyte phenotype,^[12] and intimately associated with the delivery of contractile force.^[30,31] Collagen is also associated and linked to the surface of the cardiac fibroblast through specific interactions mediated by the integrins; both high- and low-affinity interconnections between the interstitial cells of the heart and collagen have been identified.^[32] These different binding affinities appear to enable the fibroblasts to migrate on collagen (low-affinity) as well as to contract and place mechanical tension across the collagen network (high-affinity).^[33]

Mechanical Forces and Myocyte Biology

Mechanical signals play an important role in directing myocytes to assume the distinctive cytoarchitectural features of the differentiated myocyte as well as the overall structural arrangement of intact muscle.^[12] Mechanical forces associated with cardiac function during various stages of development have been shown to regulate the expression^[34] and degradation^[30,35] of a variety of contractile and regulatory proteins. Mechanical stimulation is of particular importance in the cardiovascular system in a variety of developmental and disease conditions. During development and in the adult, the cells and tissues are continually subjected to physical forces. A primary developmental force in the heart is mechanical stimulation resulting from contractile force as well as an increase in pressure and volume.^[36] The role of mechanical stimulation in cell function has been difficult to investigate due to methodological problems preventing a logical and systematic approach to the isolation of the mechanical



stimulation. To study the role of mechanical stimulation in a controlled environment, several types of *in vitro* mechanical cell stimulators have been developed. A variety of biological responses such as cell size, regulation, expression,^[34] synthesis, and degradation^[35] of a variety of contractile and regulatory proteins are measured and shown to be influenced by the application of mechanical tension. It will be essential to pre-adapt cardiac tissue constructs to mechanical stresses prior to transplantation.

Characteristics of a Biomaterial as a Tissue-Engineering Substrate

Based on the structure, function, and endogenous composition of the ECM of cardiac muscle, the logical choice for a scaffold material to construct a tissue-engineered construct would be collagen. This would be true for the majority of tissues in the body. So why have many tissue engineers chosen to avoid collagen? The answer is that they have not been able to obtain the correct physical properties for the collagen when used in other systems, such as artificial blood vessels, and thus the artificial vessels have failed when transplanted. So, what are the necessary properties for a scaffold for artificial muscle? From a biological standpoint, the material must be noncytotoxic and biodegradable, the byproducts of biodegradation must be nontoxic, and the scaffold must promote or at least support the maintenance and synthesis of new tissue. Collagen meets these requirements as the endogenous ECM and when used in culture. From a mechanical standpoint, the scaffold must have an appropriate modulus, possess flexural rigidity, have a surface energy that is not too hydrophobic or too hydrophilic, and thus provide a mean pore size and pore volume fraction to support the cells and fluid necessary to form a tissue.

In nature, collagen, as the dominant ECM component of cardiac muscle, possesses all of these features. The challenge is to recreate them *in vitro*. The scaffold must have a biodegradation rate that is fast enough to allow for the normal homeostatic remodeling of the tissue, but not so high as to have a tissue engineered construct fail before it integrates into the surrounding tissue. In other systems such as blood vessels, this is where conventional preparations using collagen gels and permutations of this type of material have fallen short. One can crosslink collagen with aldehydes to the point that they will last for years, but this collagen will not support cell growth or tissue remodeling. The key is to develop alternative processing and/or crosslinking methods to achieve the desired properties. There are also other practical/clinical properties necessary for an ideal scaffold. It must be able to “hold a stitch” so that

it can be sutured into the surgical site. It must not undergo contracture or scarring and should not promote a chronic inflammatory response. These combined characteristics are a tall order for any scaffold and most tissue engineers have turned to synthetic materials to achieve many but not all of the necessary characteristics. We propose that it is worth the effort to alter the “manufacturing” of collagen substrates since in nature it is the ideal scaffold.

DATA IN SUPPORT OF USING THE ROTATING-WALL BIOREACTOR

The data presented below were obtained from experiments conducted in the microgravity of space in collaboration with NASA. The goal of the experiments was to determine if rounded cardiac myocytes added as a suspension would: 1) interact with a template monolayer of cells on a collagen substrate to form a multilayered culture and, 2) if the rounded cells added to the aligned cells on the collagen substrate would attach to the template of aligned cells and adopt an aligned cell shape even though they were not in direct contact with the collagen-coated substrate. If both conditions were met, the experiments would provide important information showing that a pre-existing template of cells could communicate the phenotypic information stored in an underlying collagen gel to a second population of myocytes not in direct contact with the collagen fibrils. These conditions are essential if a multilayered culture of cardiac myocytes that has a phenotype similar to that of an *in vivo* myocardial wall is to be developed for subsequent implantation into a damaged heart.

To this end, we used a modification of the aligned collagen gel system^[12] to engineer a three-dimensional prosthesis of cardiac muscle cells. In the aligned collagen gel system, cells are plated onto a thin, polarized matrix of type I collagen fibrils prepared on a solid substrate. This matrix is produced by applying a solution of ice-cold, neutral type I collagen to a specially treated silicone rubber film or other suitable surface, using a sterile nylon brush, at a concentration of 0.62 $\mu\text{g}/\text{mm}^2$. The substrate is then tipped and the collagen solution is allowed to drain across the surface. Once this application process is completed, the collagen is allowed to polymerize at 37°C for 1 hour. This procedure results in a thin layer of collagen fibrils that are arrayed in parallel with one another along the direction that the solution of collagen was originally drained (Fig. 2A and B).

Under conventional culture conditions, fetal,^[37] neonatal cardiac myocytes,^[12] and neonatal cardiac fibroblasts^[38] will attach to these collagen gels and express an elongated *in vivo*-like, cell shape. The

rod-like myocytes of cultures plated onto aligned collagen are distributed in parallel with one another (Fig. 2C) and as a population; the cells exist as a monolayer in a tissue-like pattern of organization. The aligned phenotype of muscle cells cultured on this substrate is very different from the pattern of organization that is observed when these cells are cultured on native plastic, nonaligned collagen, or laminin (Fig. 2D). The geometry and biochemical nature of this interaction appear to be critical in promoting the expression of the aligned phenotype. Disrupting the aligned structure of the collagen or altering its chemical composition suppresses the expression of the aligned phenotype.^[12] Blocking the function of either the α_1 or β_1 integrin (the predominate integrin isoforms expressed by these cells) with antibodies directed against the extracellular domains of these molecules inhibits cell adhesion and disrupts the expression of the aligned phenotype. Similar results were obtained when we used antisense oligonucleotides (ODNs) to inhibit the accumulation of the α_1 subunit of the integrin complex.^[39] Blocking the accumulation of vinculin, a cytoskeletal protein of the cardiac costamere and Z-disk with antisense ODNs also suppresses the expression of the aligned

phenotype.^[37] Cells treated with antisense ODNs against vinculin did not align and were unable to assemble normal myofibrils.

The aligned phenotype in these cultures resembles the organization of muscle cells within a single cellular layer of intact heart. Data from our experiments indicate that the phenotypic information that directs the expression of this pattern of organization is stored in the chemical identity and tertiary structure of the ECM. This information is translated into signals that control the muscle cell shape. Unfortunately, the monolayer nature of these cultures makes them unsuitable as a tissue-engineered construct capable of surviving transplantation. In order to be useful for this type of clinical application, the cultures must more closely mimic the three-dimensional structure and mechanical integrity of intact tissue. This requires the prosthesis to be composed of multiple layers of muscle cells arrayed along a common axis. Our attempts to induce the formation of a multi-layered culture under conventional conditions have been unsuccessful. However, large multicellular aggregates can be induced to form in the microgravity of space and in devices designed to provide high nutrient delivery in a buoyant, low-shear environment on the ground^[40,41] mimicking the high buoyancy and low-shear conditions of space. While these aggregates can be quite large, they are not distributed in a clearly defined, three-dimensional pattern of organization, a hallmark of intact cardiac tissue. For this reason, they are also unsuitable for use in reconstructive surgery.

On four separate space shuttle flights, we have successfully used a confluent monolayer of aligned myocytes as a template for the fabrication of a multilayered aligned culture. In these experiments, we first established a confluent layer of aligned cardiac myocytes on a solid substrate for several days on the ground, and then placed them on the shuttle in one chamber of a two-chambered device designed to mix different cell populations in the microgravity of space. In the second chamber we placed a suspension culture of freshly isolated myocytes. When the shuttle reached orbit the suspension of rounded myocytes in chamber two was mixed with the template layer of aligned myocytes in chamber one. The cultures were fixed and preserved for microscopic examination 24–96 hrs after mixing.

To distinguish the two populations of cells, prior to mixing them we labeled the suspended cell population in the second chamber with the membrane-intercalating agent, DiI, prior to loading them into the chamber. Myocytes in the template layer were preloaded with DiO. Both dyes are incorporated into the membranes of cells and when excited at the appropriate wavelengths of light, DiI fluoresces green and DiO fluoresces red. By the use of confocal microscopy, this allowed us to identify unambiguously the origin of each cell

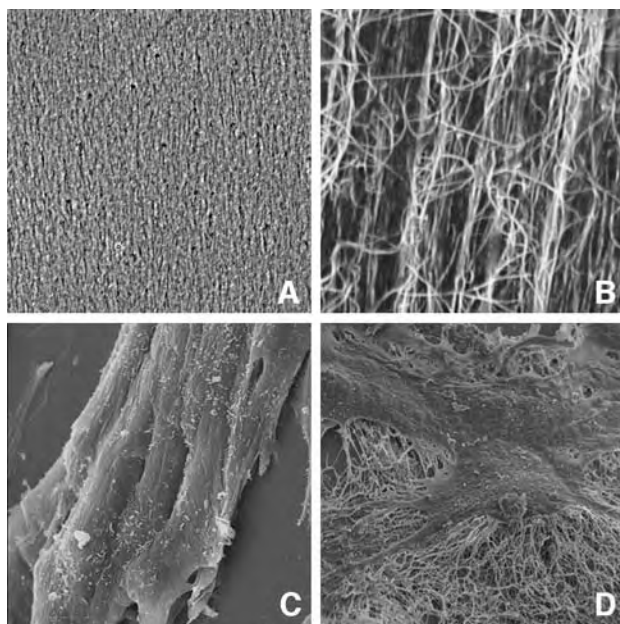


Fig. 2 A) Phase contrast micrograph of an aligned collagen gel painted onto a culture dish. B) SEM micrograph of a similar gel showing the aligned grooves in the polymerized collagen and the fibrils. C) SEM micrograph of myocytes plated onto aligned collagen fibrils. They express an elongated rod-like phenotype similar to that seen in vivo. D) Neonatal cardiac myocytes plated onto a thin gel of nonaligned collagen fibrils. These myocytes express a stellate cell shape. Magnification, A ~ 100 ×, B ~ 2500 ×, C ~ 2100 ×, D ~ 1800 ×.

C

population at the conclusion of the experiment (Fig. 3). In mixing experiments conducted at microgravity aboard the space shuttle, the rounded cells added in suspension attached uniformly over the surface of the template layer (Fig. 3A) and spread out in parallel with the long axis of the underlying cells to form multilayered cultures of aligned myocytes. Unlike the cultures of myocytes combined aboard the space shuttle in orbit, myocytes mixed in ground-based control experiments and exposed to normal gravity became unevenly distributed and formed aggregates on top of the template culture (Fig. 3B). The cells of these aggregates did not spread on the template layer even after 72 hours of

culture and the myocytes of these aggregates failed to assemble myofibrils (Fig. 3B, inset).

The results of these experiments suggest that phenotypic information in the aligned collagen is communicated by the template layer of myocytes, which are in direct contact with the collagen, to the overlaying myocytes, which are not in direct contact with the collagen. To characterize how this information might be conveyed by the template layer, we used anti-N cadherin antibodies to block cell–cell communication after mixing. As before, we prepared a template layer of aligned DiO-labeled cardiac myocytes in a culture vessel. In a second culture vessel, freshly isolated DiI-labeled cardiac myocytes were mixed in suspension with anti-N cadherin antibodies. The two culture vessels were then loaded onto the space shuttle; upon reaching orbit the two cell populations were mixed as described previously. After 48 and 96 hours, replicate cultures were fixed in 2% paraformaldehyde. Cells treated with irrelevant antibodies served as controls. Microscopic analysis indicated that cells mixed in ground-based experiments in the presence or absence of anti-N cadherin antibodies failed to form multilayered cultures (Fig. 3C). As before, cells added in suspension in ground-based experiments formed a series of aggregates on the surface of the template layer and failed to spread in parallel with the underlying myocytes. There were no obvious differences in the appearance of these two treatment groups. Control cultures mixed in the microgravity of space formed multilayered sheets of myocytes that were arrayed along a common axis and in parallel with the underlying collagen fibrils, as previously reported (Fig. 3E). In contrast, cell cultures mixed in microgravity and in the presence of anti-N cadherin antibodies failed to form multilayers (Fig. 3D). Myocyte organization in these cultures resembled those prepared under conventional ground-based conditions. Rounded cells added under these conditions formed a discontinuous layer of cell aggregates consisting of rounded myocytes on the surface of the template layer. These data indicate that cell–cell interactions mediated by the N cadherin molecule play a role in transducing the phenotypic information necessary to induce the formation of multilayered constructs.

To mimic the low-shear force environment of microgravity we have used the Synthecon Rotary Cell Culture System (RCCS). The RCCS provides a unique environment with low-shear force, simulated microgravity, and high mass transfer, all taking place in a conventional tissue culture incubator. Similar to the flight studies, a template layer of aligned myocytes on solid substrate (silicone rubber in this case, tissue culture plastic on the shuttle) were placed in the bioreactor and cultured for 3 days. Figure 4 shows the multilayer mixed culture of myocytes and fibroblasts

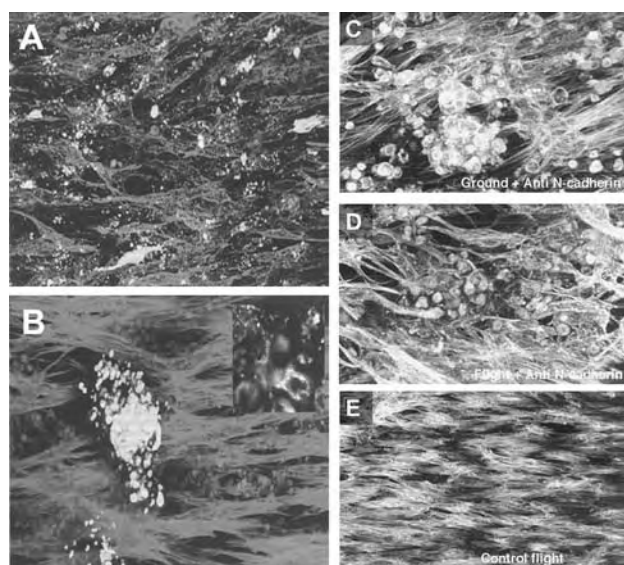


Fig. 3 A) Myocyte cultures 96 hours after mixing in microgravity. Myofibrils have been stained with rhodamine phalloidin and are depicted in the red channel. The distribution of Dye I is illustrated in the green channel (identifies cells originally added in suspension to the template layer of myocytes). Rounded cells appear uniformly distributed over the template layer when mixed in microgravity. This micrograph illustrates the uniform distribution of cells added in suspension to a template layer of aligned myocytes. B) Ground-based control experiments where cells in suspension were added to a template culture consisting of a dense monolayer of aligned myocytes (red). Cells added in suspension formed aggregates on the surface of the template culture and did not spread out (green). C) Cultures prepared under conventional conditions (ground-based, normal gravity) in the presence or absence (not shown) of N-cadherin antibodies fail to form multilayers. D) Blocking cell–cell communication by the addition of N-cadherin antibodies disrupts the formation of multilayered cultures of aligned myocytes. E) Control cells mixed in microgravity (flight) form multilayered cultures of aligned myocytes. Magnification, A 130 \times , B 200 \times , C,D,E 125 \times .

grown in the bioreactor from an aligned monolayer and additional cells added in suspension.

These experiments demonstrate the feasibility of developing multilayered cultures of myocytes. In actual use, cardiac implants would be prepared in a conventional tissue culture facility in a rotating-wall bioreactor. These devices are designed to promote cell-to-cell contacts while providing high nutrient delivery in a buoyant, low-shear environment as previously described. In order to add stability and biocompatibility to the cardiac muscle constructs for transplantation, we have chosen to use sheets of extruded collagen as the initial base substrate upon which the multilayered construct of aligned myocytes is assembled. Fig. 5A is a scanning electron micrograph of representative sheets of collagen used in the bioreactor to support the aligned cardiac cells. These sheets have considerable material integrity and an aligned orientation of fibrils that facilitates the alignment of the myocytes. We have taken these collagen substrates and coated them with our thin, aligned collagen gel templates using a soft nylon brush as described above and plated neonatal rat cardiac myocytes in a culture dish and maintained for 3–5 days *in vitro*. These substrates were then placed into the RCCS bioreactor and seeded with freshly isolated myocytes (7 million cells). The substrates were allowed to free-float or were attached to the outer wall of the RCCS. The excess cells in the reactor form aggregates, as previously described by others.^[40] We have added more cells to the construct either in the RCCS or in a culture dish. These cells add to the construct and align yielding a thicker more robust construct. To date, after adding cells to the

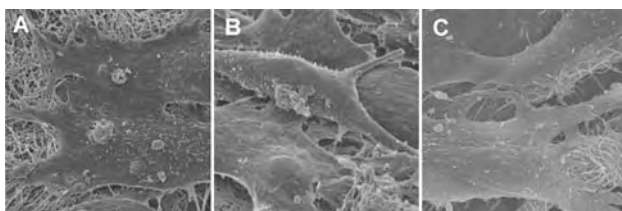


Fig. 4 A) SEM of template layer in cultures mixed in ground-based experiments. In regions that lack the large multicellular aggregates the template layer consists of a single layer of cells. The monolayer nature of the template layer is evident. The elongated cells of the template layer rest on a matrix of collagen fibrils. B) Scanning electron micrograph of the cultures prepared in microgravity. Myocytes exhibit an elongated rod-like shape and are distributed in multiple layers. The image depicts a cell resting on a bed of myocytes. Multiple layers of cells are visible; the underlying collagen fibrils are completely covered by myocytes. C) Scanning electron micrograph of the cultures prepared in the RCCS. The myocytes exhibit an elongated rod-like shape and are distributed in multiple layers.

construct two times, we have achieved spontaneously beating constructs as large as $18 \times 95 \times 1$ mm (Fig. 5B–D). Although cells add to the construct if it is floating, the rotation needs to be higher than when attached to the outer wall. We have experienced problems with the constructs dislodging from the wall and we are working on ways to attach the substrates more effectively. Maintaining the cultures on the outer wall and running the reactor at relatively slow rates (15 rpm) yields the best results. However, if the construct dislodges and then floats in the space between the walls, then the rotation rate necessary to allow cells to interact with the construct is much higher (30 rpm). The best results to date (Fig. 5) have been obtained with the construct on the outer wall and at a rotation rate of 15 rpm and adding cells in a culture dish for the second seeding. The whole construct appears to be contracting in unison indicating that the seeded cells are electrically connected. The construct in Fig. 5 is the largest we have made to date, but the RCCS comes with larger chambers that could accommodate larger pieces of substrate. These data suggest that bioreactors may provide the environment necessary to fabricate large-scale constructs of clinically relevant mass.

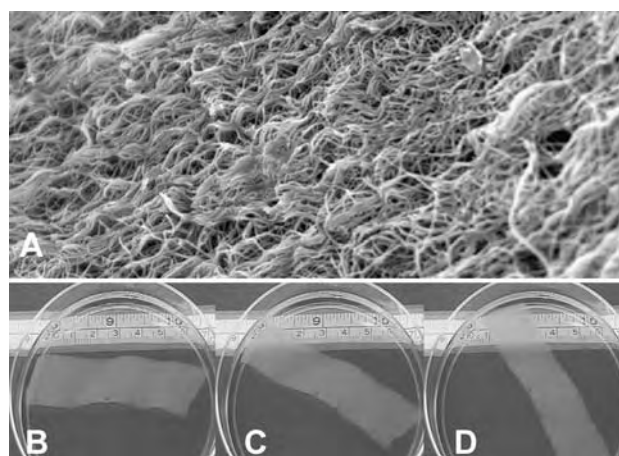


Fig. 5 A) Scanning electron micrograph of an extruded type I collagen substrate. The collagen fibrils are shown with their characteristic wavy structure. This is an example of the type of substrate used in the RCCS bioreactor. The cardiac construct seen in B–D was produced using this type of collagen substrate. B, C, and D are three still frames from a real-time video of a cardiac construct spontaneously contracting in a culture dish. The three frames are at 0, 12, and 32 seconds into the video, respectively. The construct is four weeks old and is constructed using the collagen substrate described above in A and seeded with neonatal rat cardiac myocytes in a RCCS bioreactor. The construct is about $18 \times 95 \times 1$ mm and rotates in the culture dish as it contracts. The video can be seen at: www.nyu.edu/dental/research/faculty/terracio.html. A magnification $\sim 500 \times$.

C

CONCLUSIONS

Cardiac function is intimately linked to the three-dimensional organization of the heart. Cell-based therapies designed to reconstruct dysfunctional or damaged domains of the heart, regardless of the source of the donor material, must successfully recapitulate the structure and function of the intact tissue. To fabricate a three-dimensional prosthesis that exhibits the structural and functional properties of intact cardiac tissue presents many challenges, some of which have been addressed in this article. In addition to those discussed, the tissue engineer will also be challenged with pre-adapting the construct to the mechanical environment of the heart and designing each construct to the individual patient. The field of cardiac tissue engineering is advancing at a rapid pace and these are exciting times. It is conceivable that in the near future the pathology associated with a severe myocardial infarct may be alleviated by an implanted tissue-engineered heart patch derived from the patient's own stem cells or donor myocytes grown on an artificial scaffold such as collagen.

ACKNOWLEDGMENTS

This work was funded in part by NIH grants HL072096 and HL42249 to LT, HL67097 to MJY, and DAMD17-00-1-0512 to DGS. The authors would like to thank Drs. Marian Lewis and Ed Piepmeier and NASA for their help and support of the proof of concept experiments performed on the space shuttle. The authors would also like to thank Dr. Vladimir Mironov for helpful discussions.

ARTICLES OF FURTHER INTEREST

Bioreactors; Cardiac Bioelectricity; Collagen; Rotary Cell Culture Systems; Tissue Engineering; Tissue Engineering of Cardiac Muscle

REFERENCES

- Cotran, R.; Kumar, V.; Collins, T. The Heart. In *Pathological Basis of Disease*, 6th Ed.; W.B. Saunders Co.: Philadelphia, PA, 1999; 543–599.
- Chiu, R.C.-J. Therapeutic cardiac angiogenesis and myogenesis: The promises and challenges on a new frontier. *J. Thorac. Cardiovasc. Surg.* **2001**, *122* (5), 851–852.
- Marelli, D.; Desrosiers, C.; El-Alfy, M.; Kao, R.L.; Chiu, R.C. Cell transplantation for myocardial repair: An experimental approach. *Cell Transplant.* **1992**, (1), 383–390.
- Taylor, D.A.; Atkins, B.Z.; Hungspreugs, P.; Jones, T.R.; Reedy, M.C.; Hutchenson, K.A.; Glower, D.D.; Kraus, W.E. Delivery of primary autologous skeletal myoblasts into rabbit heart by coronary infusion: A potential approach to myocardial repair. *Proc. Assoc. Am. Physicians* **1997**, *109*, 245–253.
- Soonpaa, M.H.; Koh, G.Y.; Klug, M.G.; Field, L.G. Formation of nascent intercalated disks between grafted cardiomyocytes and host myocardium. *Science* **1994**, *264*, 98–101.
- Koh, G.Y.; Soonpaa, M.H.; Klug, M.G.; Field, L. Long-term survival of AT-1 cardiomyocyte grafts in syngenic myocardium. *Am. J. Physiol.* **1993**, *264*, H1727–H1733. (*Heart Circ Physiol* 33).
- Klung, M.; Soonpaa, M.; Koh, G.; Field, L. Genetically selected cardiomyocytes from differentiating embryonic stem cells form stable intracardiac grafts. *J. Clin. Invest.* **1996**, *98*, 216–224.
- Taylor, D.A.; Atkins, B.Z.; Hungspreugs, P.; Jones, T.R.; Reddy, M.C.; Hutchenson, K.A.; Glower, D.D.; Kraus, W.E. Regenerating functional myocardium: Improved performance after skeletal myoblast transplantation. *Nat. Med.* **1998**, *4* (8), 929–933.
- Li, R.-K.; Weisel, R.D.; Mickle, D.A.; Jia, Z.-Q.; Kim, E.J.; Sakai, T.; Tomita, S.; Schwartz, L.; Iwanocko, M.; Husain, M.; Cusimaro, R.J.; Burns, R.J.; Yau, T.M. Autologous porcine heart cell transplantation improved heart function after a myocardial infarction. *J. Thorac. Cardiovasc. Surg.* **2000**, *119*, 62–68.
- Sakai, T.; Li, R.-K.; Weisel, R.D.; Mickle, D.A.; Kim, E.J.; Tomita, S.; Jia, Z.-Q.; Yau, T.M. Autologous heart cell transplantation improves function after myocardial injury. *Ann. Thorac. Surg.* **1999**, *68*, 2074–2081.
- Reincke, H.; Zhang, M.; Bartosek, T.; Murry, C.E. Survival, integration, and differentiation of cardiomyocyte grafts. A study in normal and injured rat hearts. *Circulation* **1999**, *100*, 193–202.
- Simpson, D.G.; Terracio, L.; Terracio, M.; Price, R.L.; Turner, D.C.; Borg, T.K. Modulation of cardiac myocyte phenotype in vitro by the composition and orientation of the extracellular matrix. *J. Cell. Physiol.* **1994**, *161*, 89–105.
- McDevitt, T.C.; Angello, J.C.; Whitney, M.L.; Reinecke, H.; Hauschka, S.D.; Murry, C.E.; Stayton, P.S. In vitro generation of differentiated cardiac myofibers on micro-patterned laminin surfaces. *J. Biomed. Mater. Res.* **2002**, *60*, 472–479.
- Bursac, N.; Papadaki, M.; Cohen, R.J.; Schoen, F.J.; Eisenberg, S.R.; Carrier, R.; Vunjak-Novakovic, G.; Freed, L. Cardiac muscle tissue engineering: Toward an in vitro model for electrophysiological studies. *Am. J. Physiol.* **1999**, *277*, 433–444. (*Heart Circ Physiol* 4e).
- Carrier, R.L.; Papadaki, M.; Rupnick, M.; Schoen, F.J.; Bursac, N.; Langer, R.; Freed, L.E.; Vunjak-Novakovic, G. Cardiac tissue engineering: Cell seeding, cultivation parameters, and tissue construct characterization. *Bio-technol. Bioeng.* **1999**, *64*, 580–589.
- Freed, L.E.; Vunjak-Novakovic, G. Tissue Culture Bioreactors: Chondrogenesis as a Model System. In

- Principles of Tissue Engineering*; RG Landes: Austin, TX, 1997; 151–165.
17. Freed, L.E.; Vunjak-Novakovic, G. Cultivation of cell-polymer tissue constructs in simulated microgravity. *Biotech. Bioengineering*. **1995**, *46*, 306–315.
 18. Freed, L.E.; Vunjak-Novakovic, G. Microgravity tissue engineering. *In Vitro Cell. Dev. Biol., Anim.* **1997**, *33*, 381–385.
 19. Jankowski, R.J.; Wagner, W.R. Directions in cardiovascular tissue engineering. *Clin. Plast. Surg.* **1999**, *26* (4), 605–616.
 20. Rubin, K.; Gullberg, D.; Tomasini-Johansson, B.; Reed, R.K.; Ryden, C.; Borg, T.K.; Part, H. Molecular Components. In *Structure and Function of the Extracellular Matrix of Connective Tissues*; Comer, W.D., Ed.; Harwood Academic Press: New York, 1995.
 21. Borg, T.K.; Lill, P.H.; Johnson, L.D. Specific attachment of collagen to cardiac myocytes: In vivo and in vitro. *Dev. Biol.* **1983**, *97*, 417–423.
 22. Borg, T.K.; Ranson, W.F.; Moslehy, F.A. Structural basis of ventricular stiffness. *Lab. Invest.* **1981**, *44*, 49–54.
 23. Weber, K.T.; Kalil, J.E.; Janicki, J.S. Myocardial collagen remodeling in pressure overload hypertrophy: A case for interstitial heart disease. *Am. J. Hypertens.* **1989**, *2*, 93.
 24. Weber, K.T.; Pick, R.; Janicki, J.S. Inadequate Collagen tethers in dilated cardiopathy. *Am. J. Heart.* **1988**, *116*, 1641.
 25. Weber, K.T.; Brilla, C.G. Pathological hypertrophy and cardiac interstitium. *Circulation* **1991**, *86* (6), 1849–1865.
 26. Borg, T.K.; Terracio, L. Attachment Substrates for Heart Muscle Cells. In *Cell Culture Techniques in Cardiovascular Research*; Piper, M., Ed.; Springer-Verlag: Berlin, 1989.
 27. Ross, R.S.; Borg, T.K. Integrins and the myocardium. *Circ. Res.* **2001 Jun 8**, *88* (11), 1112–1119. [Review] [102 refs] [Journal Article. Review. Review, Tutorial].
 28. Hynes, R.O. Integrins: Versatility, modulation, and signaling in cell adhesion. *Cell* **1992**, *69* (1), 11–25.
 29. Borg, T.K.; Nakagawa, M.; Carver, W.; Terracio, L. Extracellular Matrix, Receptors and Heart Development. In *Developmental Mechanisms of Heart Disease*; Clark, E.B., Markwald, R.R., Takao, A., Eds.; Futura Publ. Co.: Armonk, NY, 1995.
 30. Simpson, D.G.; Sharp, W.W.; Borg, T.K.; Price, R.L.; Terracio, L.; Samarel, A.M. Mechanical regulation of cardiac myocyte protein turnover of a myofibrillar structure. *Am. J. Physiol.* **1996**, *270* (4 pt 1), C1075–C1087.
 31. Sharp, W.W.; Simpson, D.G.; Borg, T.K.; Samarel, A.M.; Terracio, L. Contractile activity and external mechanical tension affect focal adhesion formation by cultured neonatal rat cardiac myocytes. *Am. J. Physiol.* **1997**, *273* (2 Pt 2), H546–H556.
 32. Tingstrom, A.; Heldin, C.H.; Rubin, K. Regulation of fibroblast mediated collagen gel contraction by PDGF, IL-1, and TGFB I. *J. Cell. Sci.* **1992**, *102* (pt 2), 315–322.
 33. Lundgren, E.; Gullberg, D.; Rubin, K.; Borg, T.K.; Terracio, M.J.; Terracio, L. In vitro studies on adult cardiac myocytes: Attachment and biosynthesis of collagen type IV and laminin. *J. Cell. Physiol.* **1988**, *136*, 43–53.
 34. Izumo, S.; Lompre, A.M.; Matsuoka, R.; Koren, G.; Schwartz, K.; Ginard-Nadal, B.; Mahdavi, V. Myosin heavy chain messenger RNA and Protein isoform transitions during Cardiac Hypertrophy. Interaction between hemodynamic and thyroid-induced signals. *J. Clin. Invest.* **1987**, *79*, 977–979.
 35. Samarel, A.M.; Spragia, M.L.; Maloney, V.; Kamal, S.A.; Englmann, G.L. Contractile arrest accelerates myosin heavy chain degradation in neonatal rat heart cells. *Am. J. Physiol.* **1992**, *263* (3 Pt 1), C642–C652.
 36. Terracio, L.; Peters, W.; Durig, B.; Miller, B.; Borg, K.; Borg, T.K. Cellular Hypertrophy Can be Induced by Cyclical Stretch In Vitro. In *UCLA Symposium on Molecular and Cellular Biology*; Skalak, R., Fox, C.F., Eds.; Alan R Liss, Inc.: New York, NY, 104, 51–56.
 37. Shiraishi, I.; Simpson, D.G.; Fu, L.; Zhao, Z.; Majewski, M.; Borg, T.K.; Terracio, L. Inhibition of vinculin with antisense ODNs in fetal mouse cardiac myocytes. *I. Mol. Cell. Cardiol.* **1996**, *29*, 2041–2052.
 38. Yost, M.J.; Simpson, D.G.; Wrona, K.; Ridley, S.; Ploehn, H.; Borg, T.; Terracio, L. Design and construction of a uniaxial cell stretcher. *Am. J. Physiol. Heart Circ. Physiol.* **2000**, *279*, H3124–H3130.
 39. Simpson, D.G.; Reaves, T.A.; Shih, D.; Burgess, W.; Borg, T.K.; Terracio, L. Cardiac integrins: The ties that bind. *Cardiovasc. Pathol.* **1998**, *7*, 135–143.
 40. Akins, R.E.; Schroedi, N.A.; Gonda, S.R.; Hartzell, C.R. Neonatal rat heart cells cultured in simulated microgravity. *In Vitro* **1997**, *33*, 337–343.
 41. Molnar, G.; Schroedl, N.A.; Gonda, S.R.; Hartzell, C.R. Skeletal muscle satellite cells cultured in simulated microgravity. *In vitro, Animal* **1997**, *33* (5), 386.



Cartilage Regeneration

Nic D. Leipzig
Kyriacos A. Athanasiou

Department of Bioengineering, Rice University, Houston, Texas, U.S.A.

INTRODUCTION

Joint pain is one of the leading causes of disability for an aging population. Joint pain is normally associated with the degeneration of hyaline cartilage, a form of cartilage that lines the articulating surfaces of joints. Cartilage degeneration normally occurs because of injury or primary osteoarthritis.

Cartilage is considered a simple tissue due to its avascular, aneural, alymphatic, and nearly acellular nature; however, clinical and experimental attempts to regenerate functional hyaline cartilage with mechanobiological properties matching those of native tissue have proved to be extremely challenging. Implantation of other tissues and devices has proved inadequate to match hyaline cartilage's ability to absorb and redistribute physiological loads. Tissue engineering has recently emerged as an exciting and possible solution for providing functional restoration of hyaline cartilage. The strategy of tissue engineering aims for the regeneration of cartilage through the utilization of cells, scaffolds, chemical factors, and mechanical forces. Current research following various tissue engineering paradigms has had a variety of success levels. This article will provide an overview of articular cartilage structure-function characteristics, and will focus on current research and results for the regeneration of this vital and complex tissue.

ARTICULAR CARTILAGE ANATOMY AND PHYSIOLOGY

Articular cartilage is a specialized form of hyaline cartilage that is found in diarthrodial joints. A variety of other cartilages exist in the body, including fibrocartilage and elastic cartilage. Tissue engineering has mainly focused on regeneration of articular cartilage, which at present could stand to solve the largest number of orthopedic problems including the treatment of osteoarthritis. It is estimated that osteoarthritis alone affects 20.7 million Americans. Furthermore, osteoarthritis and other musculoskeletal disorders cost the U.S. economy nearly \$65 billion dollars per year in direct expenses and lost wages and

productivity.^[1] A full understanding of the tissue's anatomy and physiology must occur to foster the accomplishment of articular cartilage regeneration.

Articular cartilage is essential for normal diarthrodial joint function.^[2] Existing as the contact surface of joints, articular cartilage functions to transfer and effectively lessen loads from one subchondral bone to another. Articular cartilage and synovial fluid combine to form a wear-resistant and friction-reduced surface that is only a few millimeters thick. This solid-fluid nature of cartilage creates a mechanoprotective environment that is favorable for the tissue's exceptional durability.^[3]

Unlike most tissues, cartilage is avascular, aneural, and alymphatic and contains a sparse number of cells (chondrocytes) that are encased in an extracellular matrix (ECM). In articular cartilage chondrocytes constitute approximately 10% of the total volume.^[4] Articular cartilage is often considered a fiber-reinforced gel containing a fluid phase of primarily water that makes up 75–80% of the wet weight of cartilage.^[4] The remaining solid phase contains approximately 50% collagen, 30–35% proteoglycans and 15–20% glycoproteins by dry weight of tissue.^[5] These assembled macromolecules are responsible for the matrix architecture and give articular cartilage its inherent structure and properties.

Collagens constitute the major portion of the dry weight of cartilage. Type II collagen is the most abundant, accounting for 80–85% of the collagens and is a specific marker of chondrocyte differentiation.^[3,6] Type II collagen is largely responsible for the tensile strength of cartilage. Small amounts of collagen types V, VI, IX, X, and XI are also found in hyaline cartilage. The associated collagens, types IX and XI account for 3–10% of total collagen and are important for the organization and structure of fibrils.^[3]

Proteoglycans are a specific class of glycoproteins consisting of glycosaminoglycan (GAG) chains covalently attached to a central core protein. Aggrecan is the predominant proteoglycan found in articular cartilage. Aggrecan interacts with hyaluronan, an unbranched polysaccharide with a molecular weight in the millions, to form large multi-molecular aggregates that are entrapped within the collagen network.^[3,6]

Mature articular cartilage is an anisotropic material that is often classified into four distinct zones in succeeding vertical sections. Starting from the articular surface, cartilage is separated into the superficial, middle (transitional), deep (radial) and calcified zones. These zones vary greatly in structure and composition (Fig. 1).^[7]

The superficial zone composes 10–20% of the full thickness of cartilage.^[7] This zone is characterized by densely packed collagen II fibrils that are oriented parallel to the surface and chondrocytes that appear to be flattened. A high tensile strength and stiffness is imparted to this zone from the combination of tangentially oriented collagen fibrils and a high water content.^[8] The middle zone serves as a transition from the tangential orientation of the superficial zone to the radial orientation of the deep zone and contains randomly oriented collagen and more rounded cells.^[7] The deep or radial zone is often the thickest zone in adult articular cartilage. This zone has the lowest water and total collagen content, but collagen fibrils are the largest in diameter. Deep zone chondrocytes are surrounded by a pericellular matrix and are arranged in columns perpendicular to the joint surface. From the radial zone, sheets of collagen fibrils extend into the calcified zone to strengthen the bond of cartilage to bone. The calcified zone is noticeably separated from the deep zone by a plane called the tidemark and contains fewer metabolically active chondrocytes that are encased in calcified matrix.^[3,8]

The chondrocyte is the sole metabolic unit responsible for the production and maintenance of cartilage. Chondrocytes originate from mesenchymal stem cells (MSCs), which have been shown to be able to differentiate *in vitro* into a number of phenotypes.^[9] Mature chondrocytes are long-lived cells that show no detectable cell division in healthy adult cartilage. Chondrocytes exist in a low-oxygen-tension environment and preferentially maintain anaerobic glycolysis.^[4,6] Since cartilage is avascular, chondrocytes must obtain oxygen and nutrients from synovial fluid. Transport is facilitated during joint movement when compressive forces change the synovial fluid composition in cartilage.

Healthy cartilage exists in a constant state of matrix remodeling. Collagens are turned over slowly, while proteoglycans have a half-life of a couple of weeks.^[10] Chondrocytes regulate the process of matrix remodeling through the production of catabolic and anabolic growth factors and cytokines. In articular cartilage, anabolic pathways generate ECM components such as collagen II and proteoglycans. These molecules are often used as positive indicators of articular cartilage production. Catabolic pathways lead to degenerative disorders such as osteoarthritis. Production of collagen type I, interleukin-6, and nitric oxide are often indicators of osteoarthritic cartilage.^[11]

THE BIOMECHANICAL ENVIRONMENT

The interaction between the collagen matrix and large multimolecular aggregates gives articular cartilage its important physical properties. Aggrecan trapped within the network is highly sulfated and creates a fixed negative charge. This charge imbalance attracts cations into the matrix creating a high osmotic swelling pressure due to the Donnan effect. Resistance in the collagen network acts to balance the swelling property of cartilage. This swelling property allows articular cartilage to dissipate transient loads through changes in osmotic and hydrostatic pressure when fluid is transferred from loaded to unloaded areas.^[3,6]

When cartilage is placed under compressive loads, fluid flows through the porous matrix, as modeled by the biphasic theory.^[12] The application of the biphasic model yields three intrinsic properties from a creep indentation test: the compressive aggregate modulus, which indicates ability to respond to compressive loads; Poisson's ratio, which provides a measure of apparent compressibility; and permeability which quantifies the resistance to flow in the porous solid.^[2,12] The mechanical properties of articular cartilage depend on biochemical composition and anatomical location. Human articular cartilage has demonstrated a compressive aggregate modulus ranging from 0.530 MPa to

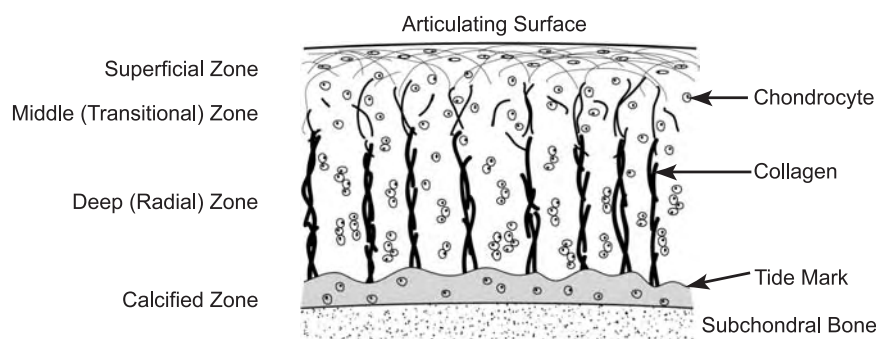


Fig. 1 Articular cartilage structure. Schematic representation of adult articular cartilage organization of matrix and cells into zones.

1.816 MPa, a Poisson's ratio ranging from 0.00 to 0.098, and a permeability ranging from $0.710 \times 10^{15} \text{ m}^4/\text{Ns}$ to $2.173 \times 10^{15} \text{ m}^4/\text{Ns}$.^[13,14] In diarthrodial joints, articular cartilage is exposed to a wide range of static and dynamic loads. During vigorous activities, stress amplitudes can rise to 10–20 MPa and cartilage compression can be up to 45%.^[15]

ARTICULAR CARTILAGE REPAIR

It is well known that articular cartilage has a limited capacity for self-repair. Numerous experimental and clinical procedures have attempted to induce or enhance the healing of focal cartilage lesions created by injury or disease. Generally, articular cartilage lesions do not heal and result in symptoms such as joint pain, locking, and disability.^[16]

The repair response to nonpenetrating cartilage defects is often minimal since no access is gained to the vascular supply. As a result, the healing response is not initiated and there is no source of progenitor cells to fill the site. Chondrocytes in the adjacent matrix are responsible for any repair that occurs. Often these cells will divide but have trouble migrating into the defect site due to entrapment in the cartilage matrix. As a result, an open defect is left in the cartilage surface, which can worsen with time and use.^[10]

Spontaneous cartilage healing is associated with defects that penetrate the bone and bone marrow spaces to gain access to the vascular supply. This allows the healing response to initiate the formation of a fibrin clot over the defect. Eventually the site will fill with a tissue that resembles hyaline cartilage in appearance. However, the cartilage is often primarily composed of fibrocartilage, which is mechanically inferior to healthy articular cartilage. This tissue is unable to function properly and degrades in a relatively short period of time.

A variety of clinical procedures for the treatment of cartilage defects have been developed including lavage and arthroscopy, shaving, debridement, Pridie drilling, microfracture, and tissue grafting.^[16] Most of these treatments have seen short-term success in the restoration of joint function and in the reduction of pain, but nothing has produced long-term, functional hyaline cartilage tissue. Of these procedures, autologous tissue grafting techniques have had the most success. Autografts can be obtained from low-weight-bearing areas in the knee and transplanted to cartilage defects. These cartilage grafts show short-term survival, however, and donor site morbidity is of concern. Similar results have been observed with autologous periosteal flaps that are sutured over penetrating defects created by drilling techniques. Excitement has been generated over a new technique named autologous chondrocyte cell

transplantation, which is employed frequently in clinical practice. This technique involves the injection of cultured chondrocytes underneath a periosteal flap that is sutured over a nonpenetrating cartilage lesion. Chondrocytes are harvested from a donor site in an earlier procedure. In the initial clinical studies, good to excellent results were seen in most patients. Upon two-year examination the repair cartilage resembled hyaline cartilage in gross morphological appearance. Longer-term evaluation is still needed to determine the function and durability of the new tissue. Autologous chondrocyte cell transplantation should not be considered a panacea since encouraging results have been observed only for the healing of focal chondral defects, not for the treatment of osteoarthritic joints.^[17]

CARTILAGE REGENERATION: APPLIED TISSUE ENGINEERING

The ultimate goal for the treatment of cartilage injuries is articular cartilage regeneration, where the neocartilage is completely indistinguishable from host cartilage. This differs from repair cartilage, where the result is a functional tissue that resembles hyaline cartilage in appearance but differs in biochemical composition, morphology, and biomechanical properties.^[2,18]

Historically the assessment of repair cartilage has been based on gross morphology and histology. More recently, the importance of determining the tissue's functional abilities has become evident. Biochemical and biomechanical properties can serve as indirect measures of neocartilage's functional ability. Quantitative measurement of collagen and glycosaminoglycan content serve as adequate means for biochemical evaluation of neocartilage. Common measurements for both use radioactive isotope incorporation techniques. Measurement of biomechanical properties is the most important evaluation of an engineered tissue's ability to function as native cartilage. An indentation test is one of the most common means of obtaining the mechanical properties of a neotissue *in vitro*.^[2] More recently techniques have been developed for *in vivo* mechanical assessment of cartilage. An arthroscopic indentation probe named Actaeon^[19] is an example that currently enjoys clinical use.

The functional restoration of damaged and diseased articular cartilage has proved to be extremely challenging. Current clinical procedures meant to repair cartilage defects often have unpredictable results. As a consequence, tissue engineering of articular cartilage has evolved as an important and possible solution to the enigma of cartilage regeneration. Cartilage tissue engineering aims to create regenerated articular cartilage through the manipulation of four main variables:

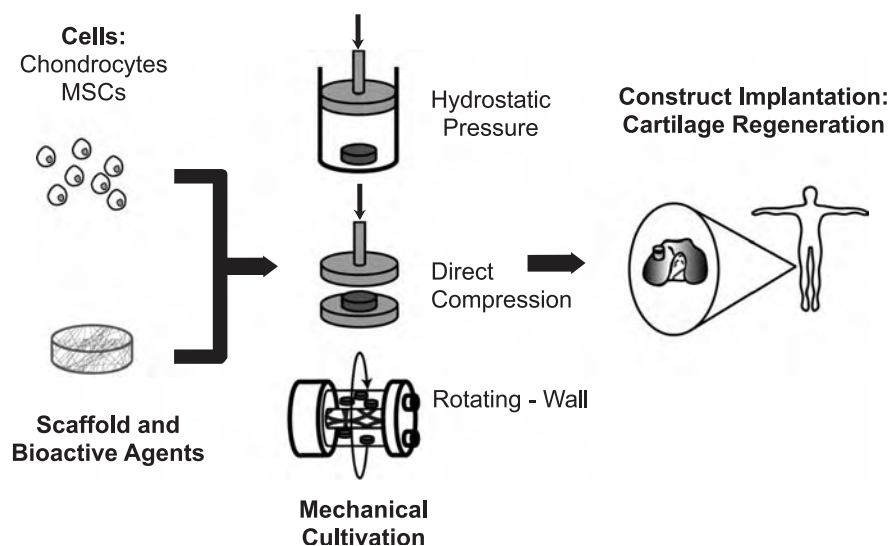


Fig. 2 Tissue-engineered cartilage. Illustration of application of the tissue engineering methodology to cartilage regeneration.

scaffold materials and structure, cell type, chemical factors (peptides and cytokines/growth factors), and mechanical forces (Fig. 2).^[7] The complex interactions of these variables make the engineering task difficult, but great strides have been made in recent years.

SCAFFOLD SELECTION

Current methods for regeneration of cartilage have concentrated on the use of biodegradable polymer scaffolds of natural and synthetic origin. The choice of material is an important consideration when designing a scaffold for *in vitro* or *in vivo* use. Scaffolds selected for cartilage regeneration *in vivo* should have several important properties. First, the scaffold should be biodegradable and bioresorbable, with degradation kinetics matching the rate of new tissue formation. The material should be nontoxic and should be non-stimulatory to inflammatory cells. Degradation products of the material should follow the same criteria. The scaffold should contain a porous network, to allow for tissue growth and diffusive transport of nutrients. Finally, the surface should facilitate cell attachment, proliferation, and differentiation of seeded cells.^[20] The *in vitro* requirements of a scaffold often are not as strict concerning biocompatibility, but are important since in the end *in vitro* studies should lead to *in vivo* work.

The mechanical requirements of an ideal scaffold often depend on whether application will be *in vitro* or *in vivo*. For both instances, the scaffold must provide the proper support structure to protect and facilitate the growth of seeded cells. For *in vivo* use, the scaffold should have good fixation to the defect site. Mechanical properties of the *in vivo* scaffold should be similar to that of native hyaline cartilage and should be able to

function under the native loading environment. A scaffold that is seeded and cultured *in vitro* does not require the same mechanical rigidity to a similar *in vivo* construct. *In vitro* constructs can be cultured for implantation for a sufficient amount of time for a mechanically acceptable tissue to form.^[20,21]

Because new materials are constantly being discovered, the mechanical and chemical characterization of biomaterials is an important field of study for cartilage scaffold materials. Many biomaterials have seen application to other cell types; in this article only the materials that have seen positive results for chondrocytes and cartilage regeneration will be discussed. For a more extensive review of scaffold materials and design see Refs. [16] and [20].

Natural Polymers

Naturally derived materials often provide superior strength and exhibit high molecular and macroscopic organization. Contrary to the popular belief that natural materials are nonimmunogenic, many natural polymers are antigenic and produce a chronic inflammatory response. Due to natural derivation, the possibility of disease transmission is also of concern. Another common problem with natural polymers is processing them to consistent grade and properties.^[22]

One of the most extensively studied natural polymers for articular cartilage engineering is collagen, specifically type I collagen. Collagens are found in many connective tissues throughout the body. Scaffolds made of collagen can be degraded and remodeled by chondrocytes to allow for tissue growth. It has been demonstrated that proliferation of chondrocytes is encouraged in collagen I gels, but chondrocytes are only able to maintain their phenotype for a few weeks before dedifferentiation into a more fibroblastic



phenotype.^[2] Success has been seen using collagen I scaffolds in conjunction with cross-linked proteoglycans.^[23] Unlike in collagen I gels, chondrocytes cultured in collagen type II scaffolds do not show signs of dedifferentiation,^[24] but unfortunately collagen II scaffolds are hard to fabricate.^[20]

Agarose is a natural derivative of seaweed that is used to fabricate gels. Cartilage constructs can be made from agarose so that chondrocytes are fully encapsulated within. Agarose gels provide a three-dimensional environment that helps maintain the chondrocytic phenotype. These constructs have been used as a vehicle for mechanical stimulation of chondrocytes and have shown positive results for matrix production. However, tissue formation may be hindered in the gel since agarose degrades very slowly both *in vitro* and *in vivo*. Unlike synthetic polymers, the degradation rates of agarose gels cannot be tailored to match that of proliferating cartilage tissue.^[20]

SYNTHETIC POLYMERS

Most successful attempts to achieve engineered cartilage constructs have utilized synthetic materials. With synthetic polymers physical and chemical properties can be modified. As a result, mechanical strength and degradation rates can be tailored to specific applications. Unlike natural polymers, there is no fear of pathogen transmission and synthetic polymers can be mass-produced with relatively reproducible properties and characteristics.

The most widely used synthetics in engineered cartilage constructs are the poly(α -hydroxy esters): poly(glycolic acid) (PGA), poly(L-lactic acid) (L-PLA) and their copolymer poly(D,L-lactic glycolic acid) (PLGA). These materials are all FDA approved for certain human implantations. In addition, these alpha polyesters degrade by hydrolysis into bioresorbable by-products.

PGA has seen wide use in medicine for some time as a suture material. *In vivo*, PGA shows a degradation time of 2–4 weeks. PGA is often extruded into small polymer strands that are molded into a nonwoven felt for use as a scaffold. Porous PGA scaffolds can also be fabricated using salt leaching techniques. Molded felt PGA scaffolds provide seeded chondrocytes with a very porous matrix (~90%) for attachment and proliferation. Unseeded PGA meshes are mechanically inferior to native tissue and require tissue formation for mechanical strength. Freed and associates have used PGA scaffolds extensively for cartilage regeneration research.^[25] Their studies have shown good production of matrix compared to controls, as characterized by measurements of total collagen and GAGs. During

tissue formation the PGA scaffolds also demonstrate suitable degradation characteristics.^[25]

Lactic acid exists in two stereo-isoforms, which give rise to four distinct types of PLA: poly(D-lactide), poly(L-lactide), poly(D,L-lactide) and poly(meso-lactide).^[26] Generally, only the semi crystalline L-PLA and the amorphous D,L-PLA are used in cartilage scaffolds. L-PLA usually shows degradation rates ranging from a few months to many years. Similar to PGA, L-PLA is often extruded into small fibers and molded into a nonwoven mesh for cartilage scaffolds. Porous L-PLA scaffolds have also been utilized for bone regeneration scaffolds.

PLGA is a copolymer of lactic acid and glycolic acid. The material properties of PLGA can be varied by the ratios of the two monomers used during polymerization. Ratios higher in monomeric lactic acid show longer degradation times. For example, PLGA that has a 75/25 ratio of D,L-PLA/PGA has a degradation time of 4–5 months, while the 50/50 ratio shows a degradation time of 1–2 months.^[27] Similar to L-PLA and PGA, the copolymer can be fabricated into felt, or as a porous, solid scaffold for use in cartilage regeneration. Athanasiou and associates have experimented extensively with 50/50 ratios of D,L-PLA/PGA showing biocompatibility and ability to carry mechanical loads *in vivo*.^[28,29]

CELLS FOR CARTILAGE REGENERATION

Cell types for engineered cartilage constructs have come from both fully differentiated chondrocytes and progenitor cells. Chondrocytes are easily obtained from freshly excised articular cartilage. Tissue is harvested from specimen joints and then enzymatically digested with collagenase to isolate chondrocytes from their extracellular matrix. Often these primary chondrocytes must be expanded in two-dimensional culture to reach the cell densities needed for seeding three-dimensional scaffolds. Interestingly, chondrocytes that propagate in monolayer do synthesize cartilage matrix. However, the tissue that forms is fibrous in nature and has poor organization. For cell expansion, often several passages are needed to reach the required cell densities. A passage is defined as enzymatic (e.g., via trypsin) release of chondrocytes from monolayer followed by subsequent reseeding at lower cell number. Two to four passages can be performed without dedifferentiation to a fibroblastic phenotype.^[18]

Mesenchymal stem cells (MSCs) are highly proliferative, multipotent cells that have shown increased use in bone and cartilage engineering. MSCs can be isolated from bone marrow and expanded in culture without loss of osteochondrogenic potential.^[9] Differentiation can proceed *in vitro* or *in vivo* and can be further

guided by the addition of growth and differentiation factors.

The importance of initial cell seeding density of constructs is important for tissue formation. Few systematic studies have been implemented to decipher optimal initial cell densities to induce cartilage formation. It has been previously observed that scaffolds seeded with densities of less than 10 million cells per mL will result in little or no cartilage tissue formation.^[18]

CHEMICAL FACTORS

Chemical modifications to tissue-engineered scaffolds have many benefits for tissue formation. Depending on the modification, such additions can aid attachment, proliferation, differentiation, and matrix production of seeded cells. Surface modifications such as protein coating and peptide attachment have been used to control cell attachment to scaffold surfaces. Growth factors and cytokines are often included to guide the differentiation, proliferation, and matrix production of cells in scaffolds.

Surface Modification

Integrin-receptor reactions mediate the attachment, spreading, and transmembrane signaling of chondrocytes on substrates. Chondrocytes express a variety of integrin subunits including α_2 , α_5 , α_6 , α_{10} , α_v , and β_1 .^[30] These integrins form heterodimers that bond to a variety of extracellular molecules. The choice of chemical modification to the surface of a scaffold requires consideration of the integrin and ligand involved.

Protein adsorption is a common method of enhancing surface attachment to a scaffold. The choice of protein coating used can greatly affect scaffold biocompatibility and can induce an immune response upon implantation. Protein adsorption is easily accomplished by choosing a hydrophobic scaffold material. ECM proteins such as collagen or fibronectin are often used as scaffold coatings. Collagen I and fibronectin have been shown previously to enhance chondrocyte attachment to surfaces.^[20]

Peptides are short amino acid sequences that are derived from larger adhesion proteins. Sequences from the active binding regions of proteins can be isolated so that only a few amino acids are required for receptor-mediated attachment instead of the entire protein. Often, peptide sequences are securely tied to scaffold surfaces using covalent bonds. A linker chain can be added to avoid steric hindrance and to provide enough clearance from the surface to allow receptor-ligand binding.^[20] The density of peptides on the scaffold surface is an important consideration because high

concentrations of peptides can decrease cell motility and therefore hinder spreading and proliferation of chondrocytes. The Arginine-Glycine-Aspartic acid (RGD) sequence is the most common peptide sequence used for cell attachment because chondrocytes have multiple integrin dimers that recognize and bind to the RGD sequence. Longer peptides that contain RGD sequences are also utilized to facilitate attachment to certain materials. Examples include sequences such as **GRGD**, **GRGDSP**, and **CGNGEPRGDTYRAY**.^[20]



Growth Factors and Cytokines

The addition of growth factors to scaffolds or areas of cartilage degeneration offer a means of enhancing cell differentiation and proliferation. Proteins such as transforming growth factor beta (TGF- β), insulinlike growth factor (IGF), platelet-derived growth factor (PDGF), hepatocyte growth factor (HGF), basic fibroblastic growth factor (bFGF), bone morphogenetic proteins (BMPs) and interleukin-4 (IL-4) alone and in combination have been experimentally shown to affect chondrocytes. The majority of growth factor–cartilage scaffold studies have been conducted in vitro where factors can be delivered and distributed through media. For in vivo application, special consideration must be made to incorporate growth factors into the scaffolds to achieve even distribution throughout the construct. Often growth factors do not have the same effects in vivo as seen in vitro due to their soluble nature as well as unforeseen interactions with the body.

Transforming growth factor- β 1 (TGF- β 1) and insulinlike growth factor-I (IGF-I) are two of the most widely researched regulatory factors for cartilage regeneration. Various studies using TGF- β 1 have found conflicting results, showing the inhibition of chondrocyte proliferation or the production of proteoglycans and collagen. IGF-I has been consistently shown to stimulate chondrocyte proliferation and synthesis of ECM molecules. A recent systematic study exposed non-woven PGA scaffolds cultured in spinner flasks to varying concentrations and combinations of IGF-I, TGF- β 1, IL-4, and PDGF.^[31] It was found that IGF-I substantially increased construct wet weights and total amounts of collagen and glycosaminoglycans (GAGs), but did not significantly change wet weight fractions of either compared to controls. TGF- β 1 was found to increase total collagen and the wet weight fraction of collagen while decreasing the amount of GAGs. The combination of IL-4 and TGF- β 1 increased construct wet weight as compared to TGF- β 1 or IGF-I alone while inhibiting GAG depletion. These results confirm previous findings that demonstrated IL-4's ability to inhibit proteoglycan degradation in cartilage.^[31]

IN VITRO CULTURE—MECHANICAL ENVIRONMENT

Chondrocytes in loaded articular joints experience a multitude of forces, including hydrostatic, compressive, tensile, and shear forces (Fig. 3). The application of these forces has been shown to be important for maintenance of the chondrocyte phenotype and for the production of cartilaginous tissue. Through the process of mechanotransduction, cells regulate transcriptional activities based on mechanical signals received at their surface. Mechanical signals can induce cytoskeletal remodeling and can stimulate anabolic or catabolic pathways in cartilage. Focal adhesions allow chondrocytes to “sense” their mechanical environment; they also provide anchorage to ECM molecules. Integrins and other constituents (e.g., cytoskeletal related proteins) allow focal adhesions to transmit mechanical forces across the cell membrane to coordinate cytoskeletal and transcriptional changes.^[32]

The biochemical and mechanical characterization of individual chondrocytes is vital for elucidation of the complex pathways influencing cartilage formation or destruction. The single-cell approach,^[33] a recently formulated methodology, offers a means of simplifying this characterization. Techniques such as cytoindentation, cytodetachment, and reverse transcriptase-polymerase chain reaction (RT-PCR) are tools to

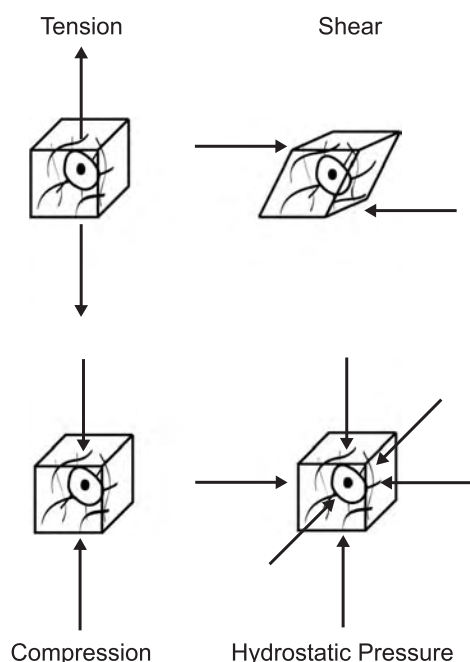


Fig. 3 Forces experienced by chondrocytes. Illustration of forces borne onto chondrocytes through matrix and fluid interactions in articular cartilage.

decipher the effects of forces on single chondrocytes. More information on cell mechanics of chondrocytes and the single-cell approach can be found elsewhere.^[33]

In vitro culturing conditions for cartilage should mimic the mechanical environment of native hyaline cartilage to induce proper tissue formation. In general, constructs exposed to extraneous forces show increased proliferation and production of matrix. Several methods have been employed to offer mechanical stimulation to chondrocytes in vitro. These methods can be grouped into three categories: those that use hydrostatic pressure, direct compression, or shear as the biomechanical stimulus. For a more complete review of mechanical culturing of cartilage see Ref. [11].

Hydrostatic Pressure

During joint articulation, synovial fluid inside the joint capsule generates hydrostatic pressure as a result of resistance to flow in the cartilage matrix. As a result chondrocytes experience hydrostatic pressure during periods of loading. During regular activity, the interstitial pressure within a normal diarthrotic joint cycles between 7 and 10 MPa.^[34] As part of an in vitro tissue-engineering approach, cultures are placed in chambers to expose cell-seeded constructs to hydrostatic pressure. Often the pressure is cycled at a prescribed frequency and at physiological loads to mimic the host environment of cartilage. Loads between 0.1 and 15 MPa and frequencies between 0.05 and 1 Hz have generated positive results in chondrocyte cultures compared to controls.^[11]

Direct Compression

Compressive forces are generated within the articular joint due to surface-to-surface contact. During joint loading, cartilage is compressed as forces are distributed throughout the matrix. Direct compression bioreactors have been utilized to expose cartilage constructs to compressive forces during culturing. A variety of designs exist, but all use a similar mechanism that involves the translation of a flat surface onto the top of a construct, under load or displacement control. Dynamic compression, unlike static compression, has proved to be effective for stimulating cartilage formation in constructs. Studies have produced positive results for cartilage regeneration, but were often limited by the machinery used. Experiments have examined frequencies ranging from 0.001 to 3 Hz, loads from 0.1 to 24 MPa, and strains from 0.2 to 25%.^[11]

Shear Forces: Bioreactors

The use of bioreactors is the most common way of culturing cartilage constructs in a shear environment. Beyond the mechanical environment, bioreactors provide mixing of media for more efficient nutrient transport and waste removal. Some of the bioreactors used for the *in vitro* culture of chondrocyte-seeded scaffolds include spinner flasks, perfusion bioreactors, and rotating-wall bioreactors.

Spinner flasks are the simplest type of bioreactor used for culturing cartilage constructs. Usually scaffolds are attached to needles that hang from a stopper. Media circulation is facilitated through gentle mixing by a magnetic stir bar. Cells are normally added to the media and are allowed to attach on their own to the scaffolds. Spinner flasks are often used as a means of seeding scaffolds with chondrocytes before transferring the constructs to more complex bioreactors.^[25]

Direct perfusion reactors offer a means of reducing shear gradients in a culture environment and eliminating mechanical mixing. Mass transfer is achieved by a simple design that forces media through fixed scaffolds. The bioreactor usually consists of a hollow vessel where media enters one end and exits out the other. Fresh media can be constantly supplied, but research has shown that recycling of media can encourage tissue growth by retaining the extracellular proteins (e.g., growth factors and cytokines) produced by growing chondrocytes.^[11] The cells in three-dimensional perfused cartilage scaffolds experience shear forces, which have been shown to up-regulate matrix production. Perfusion studies have demonstrated mixed results for culturing of cartilage constructs. Using direct perfusion, Pazzano and colleagues demonstrated an increase of 118% in DNA content, an increase of 184% in GAG content, and an increase of 155% in hydroxyproline incorporation (measure of collagen synthesis) compared to static culture.^[35] Alignment of chondrocytes to the direction of flow was also observed. Perfusion bioreactors are also able to align collagen fibers and thus offer a method of influencing cartilage matrix architecture.

The most promising overall results for culturing cell-seeded cartilage scaffolds have come from rotating-wall reactors. These reactors are a modification of the perfusion design that instead offers a low-shear environment. The rotating-wall reactor was originally developed by NASA to simulate the effects of microgravity on cell culture but has found application for the *in vitro* culture of cartilage constructs.^[11] The design consists of concentric cylinders with adequate space for media and scaffolds to be placed in between the walls. Normally the outer wall of the reactor rotates, thus providing shear force and facilitating fluid flow inside. Indirect perfusion systems are often added as a means

of replacing the media. Results from rotating-wall cultures have shown higher fractions of collagen and GAGs compared to mixed flasks and static culture. Freed and associates found that bovine chondrocyte-seeded PGA scaffolds cultured in rotating-wall reactors produced 68% as much GAGs and 33% as much collagen II per gram wet weight as compared to natural calf cartilage.^[25] The neotissue generated in the constructs produced morphologies that approached that of native articular tissue, which has not been seen in other bioreactors or culturing conditions.

CONCLUSION

Full regeneration of cartilage is the best solution to the problem of cartilage damage caused by disease and injury. Tissue engineering is poised to offer the best approach of achieving cartilage regeneration through the selection and manipulation of four main variables: cells, scaffold, chemical factors, and mechanical environment.

Successes of current research have demonstrated the feasibility of the tissue engineering paradigm for the achievement of cartilage regeneration. There is an important need for cartilage regeneration, which currently cannot be accomplished with clinical procedures in a consistently successful manner. This solution is still not attainable, because there is still much knowledge and experience to be gained before fully functional cartilage can be created.

ARTICLES OF FURTHER INTEREST

Allografts; Bioreactors; Collagen; Finite Element Analysis; Poly(Glycolic Acid); Poly(lactic acid)s; Rotary Cell Culture Systems; Tissue Engineering of Cartilage

REFERENCES

1. Arthritis Foundation. <http://www.arthritis.org/conditions/DiseaseCenter/oa.asp> (accessed July 2002).
2. Athanasiou, K.A.; Shah, A.R.; Hernandez, R.J.; LeBaron, R.G. Basic science of articular cartilage repair. *Clin. Sports Med.* **2001**, *20* (2), 223–247.
3. Hasler, E.M.; Herzog, W.; Wu, J.Z.; Muller, W.; Wyss, U. Articular cartilage biomechanics: Theoretical models, material properties, and biosynthetic response. *Crit. Rev. Biomed. Eng.* **1999**, *27* (6), 415–488.
4. Stockwell, R.A. Biology of Cartilage Cells. In *Biological Structure and Function*; Harison, R.J., McMinn, R.M.H., Eds.; Cambridge University Press: Cambridge, 1979; 7.
5. Buckwalter, J.A.; Hunziker, E.B.; Rosenberg, L.C.; Coutts, R.; Adams, M.; Eyre, D. Articular Cartilage: Composition and Structure. In *Injury and Repair of the Musculoskeletal Soft Tissues*, 2nd Ed.; Woo, S.L.,



- Buckwalter, J.A., Eds.; American Academy of Orthopedic Surgeons: Park Ridge, 1991; 405–425.
6. Muir, H. The chondrocyte, architect of cartilage. Biomechanics, structure, function and molecular biology of cartilage matrix macromolecules. *Bioessays* **1995**, *17* (12), 1039–1048.
 7. Hu, J.C.Y.; Athanasiou, K.A. The Role of Mechanical Forces in Tissue Engineering of Articular Cartilage. In *Functional Tissue Engineering*; Guilak, F., Butler, D.L., Mooney, D., Eds.; Springer Verlag: New York, 2003; 227–243.
 8. Buckwalter, J.A.; Mankin, H.J. Articular cartilage: Tissue design and chondrocyte-matrix interactions. *Instr. Course Lect.* **1998**, *47*, 477–486.
 9. Kadiyala, S.; Young, R.G.; Thiede, M.A.; Bruder, S.P. Culture expanded canine mesenchymal stem cells possess osteochondrogenic potential in vivo and in vitro. *Cell Transplant.* **1997**, *6* (2), 125–134.
 10. Coutts, R.D.; Sah, R.L.; Amiel, D. Effects of growth factors on cartilage repair. *Instr. Course Lect.* **1997**, *46*, 487–494.
 11. Darling, E.M.; Athanasiou, K.A. Articular cartilage bioreactors and bioprocesses. *Tissue Eng.* **2003**, *9* (1), 9–26.
 12. Mow, V.C.; Kuei, S.C.; Lai, W.M.; Armstrong, C.G. Biphasic creep and stress relaxation of articular cartilage in compression: Theory and experiments. *J. Biomech. Eng.* **1980**, *102* (1), 73–84.
 13. Athanasiou, K.A.; Rosenwasser, M.P.; Buckwalter, J.A.; Malinin, T.I.; Mow, V.C. Interspecies comparisons of in situ intrinsic mechanical properties of distal femoral cartilage. *J. Orthop. Res.* **1991**, *9* (3), 330–340.
 14. Athanasiou, K.A.; Agarwal, A.; Dzida, F.J. Comparative study of the intrinsic mechanical properties of the human acetabular and femoral head cartilage. *J. Orthop. Res.* **1994**, *12* (3), 340–349.
 15. Grodzinsky, A.J.; Levenston, M.E.; Jin, M.; Frank, E.H. Cartilage tissue remodeling in response to mechanical forces. *Annu. Rev. Biomed. Eng.* **2000**, *2*, 691–713.
 16. Hunziker, E.B. Articular cartilage repair: Basic science and clinical progress. A review of the current status and prospects. *Osteoarthr. Cartil.* **2002**, *10* (6), 432–463.
 17. Buckwalter, J.A. Articular cartilage: Injuries and potential for healing. *J. Orthop. Sports Phys. Ther.* **1998**, *28* (4), 192–202.
 18. LeBaron, R.G.; Athanasiou, K.A. Ex vivo synthesis of articular cartilage. *Biomaterials* **2000**, *21* (24), 2575–2587.
 19. Niederauer, M.Q.; Wilkes, R.P.; Niederauer, G.M.; Cristante, S.; Cline, S.F.; Tynmann, R.A. Hand-Held Materials Tester. US Patent 5,904,658, 1997.
 20. Darling, E.M.; Athanasiou, K.A. Bioactive Scaffold Design for Articular Cartilage. In *Biomedical Technology and Devices Handbook*; Moore, J., Zouridakis, G., Eds.; CRC Press: New York, 2003; Ch21 (1–16).
 21. Coutts, R.D.; Healey, R.M.; Ostrander, R.; Sah, R.L.; Goomer, R.; Amiel, D. Matrices for cartilage repair. *Clin. Orthop.* **2001**, *391* (Suppl), S271–279.
 22. Lu, L.; Valenzuela, R.G.; Yaszemski, M.J. Articular cartilage engineering. *J. Regenerative Med.* **2000**, *2*, 99–114.
 23. van Susante, J.L.C.; Pieper, J.; Buma, P.; van Kuppevelt, T.H.; van Beuningen, H.; van Der Kraan, P.M.; Veerkamp, J.H.; van den Berg, W.B.; Veth, R.P.H. Linkage of chondroitin-sulfate to type I collagen scaffolds stimulates the bioactivity of seeded chondrocytes in vitro. *Biomaterials* **2001**, *22* (17), 2359–2369.
 24. Nehrer, S.; Breinan, H.A.; Ramappa, A.; Shortkroff, S.; Young, G.; Minas, T.; Sledge, C.B.; Yannas, I.V.; Spector, M. Canine chondrocytes seeded in type I and type II collagen implants investigated in vitro. *J. Biomed. Mater. Res.* **1997**, *38* (2), 95–104.
 25. Freed, L.E.; Hollander, A.P.; Martin, I.; Barry, J.R.; Langer, R.; Vunjak-Novakovic, G. Chondrogenesis in a cell-polymer-bioreactor system. *Exp. Cell. Res.* **1998**, *240* (1), 58–65.
 26. Kohn, J.; Langer, R. Bioresorbable and bioerodable materials. In *Biomaterials Science*; Ratner, B.D., Hoffman, A.S., Shoen, F.J., Lemons, J.E., Eds.; Academic Press: New York, 1996; 64–84.
 27. Middleton, J.C.; Tipton, A.J. Synthetic biodegradable polymers as orthopedic devices. *Biomaterials* **2000**, *21* (23), 2335–2346.
 28. Athanasiou, K.A.; Korvick, D.; Schenck, R.C., Jr. Biodegradable implants for the reatment of osteochondral defects in a goat model. *Tissue Eng.* **1997**, *3* (4), 363–373.
 29. Athanasiou, K.A.; Singhal, A.R.; Agrawal, C.M.; Boyan, B.D. In vitro degradation and release characteristics of biodegradable implants containing trypsin inhibitor. *Clin. Orthop.* **1995**, *315*, 272–281.
 30. LeBaron, R.G.; Athanasiou, K.A. Extracellular matrix cell adhesion peptides: functional applications in orthopedic materials. *Tissue Eng.* **2000**, *6* (2), 85–103.
 31. Blunk, T.; Sieminski, A.L.; Gooch, K.J.; Courter, D.L.; Hollander, A.P.; Nahir, A.M.; Langer, R.; Vunjak-Novakovic, G.; Freed, L.E. Differential effects of growth factors on tissue-engineered cartilage. *Tissue Eng.* **2002**, *8* (1), 73–84.
 32. Zhu, C.; Bao, G.; Wang, N. Cell mechanics: Mechanical response, cell adhesion, and molecular deformation. *Annu. Rev. Biomed. Eng.* **2000**, *2*, 189–226.
 33. Shieh, A.C.; Athanasiou, K.A. Biomechanics of single chondrocytes and osteoarthritis. *Crit. Rev. Biomed. Eng.* **2002**, *30* (4–6), 307–343.
 34. Hall, A.C.; Horwitz, E.R.; Wilkins, R.J. The cellular physiology of articular cartilage. *Exp. Physiol.* **1996**, *81* (3), 535–545.
 35. Pazzano, D.; Mercier, K.A.; Moran, J.M.; Fong, S.S.; DiBiasio, D.D.; Rulfs, J.X.; Kohles, S.S.; Bonassar, L.J. Comparison of chondrogenesis in static and perfused bioreactor culture. *Biotechnol. Prog.* **2000**, *16* (5), 893–896.

Cell Culture Assays

P. J. Doherty

Department of Clinical Engineering, University of Liverpool, Liverpool, United Kingdom



INTRODUCTION

Biomaterials science has developed very rapidly in recent years and has now achieved a high level of sophistication in terms of research, manufacture, and testing. Every new biomaterial, whether selected for implantation or to be used externally in contact with the human body, undergoes an extensive array of testing using a diversity of techniques. Increasingly, in-vitro test methods are playing a significant role. Appropriate selection of testing methods can fundamentally assist in the research and development of biomaterials. Although animal-testing procedures (in-vivo testing) remain valuable and are still the most widely employed and dependable means of biological assessment before clinical trials, the quantity of animal testing has been restricted in recent years. This is partly because the huge increase in the use of artificial materials in the body has been accompanied not only by an increase in the required legislation, but also by an increase in public concerns about the use of such materials and about the appropriate use of animals.

In-vitro cell culture techniques have been proposed for many years as an alternative to animal testing. Although such a suggestion remains controversial, it is certainly true that the effective use and sophistication of the cell culture test has improved. The application of cell culture assays in the field of biomaterial evaluation has developed alongside the significant use of cell culture in drug toxicology studies, and the potential for replacing animal studies is ever closer. New techniques in cell biology and cellular genetics will continue to provide in-vitro assays that are more relevant to in-vivo and clinical situations. Nevertheless, it is also true that much of the cell culture testing carried out and reported today is not useful and may indeed be misleading.

The primary purpose of the ISO standard 10993 “Biological Evaluation of Medical Devices” is the protection of humans. Part five of the standard deals with ‘Tests for cytotoxicity—in vitro methods’ and presents guidelines for the evaluation of toxicity. Although toxicity testing is not the sole function of the cell culture test, it is the increase in toxicity tests for biomaterial evaluation that has led to the plethora of in-vitro cell culture testing employed in the biomaterials field today. These include testing regimes that attempt to

measure cytotoxicity, cytocompatibility, cell adhesion, genotoxicity, haemocompatibility, cell proliferation, and a host of other measurable cell parameters.

BACKGROUND

The use of cell culture testing for biomaterial evaluation has relatively recent origins, despite the fact that we have had the ability to culture living cells outside of the body for a much longer period of time. The development of cell culture for fundamental biological investigations can be divided into three phases.^[1] The first, up to 1838, was mainly concerned with the definition and hypothesis of cell origin and form. The second, from 1838 to the beginning of the 20th century, was relatively inactive; in-vitro techniques had been demonstrated but proved difficult to master. The work of Harrison in 1907,^[2] and Carrell in 1912^[3] confirmed these researchers as the initiators of the cell culture techniques that are employed today. Although tissue and cell culture methods developed from that time, problems of infection and contamination from bacteria and fungi meant that the potential benefits offered by these new techniques were limited to specialist workers.

The discovery of antibiotics, the introduction of the laminar airflow cabinet, and the development of synthetic culture media during the 1950s and 1960s has transformed the use of this valuable biological tool and has allowed researchers from a diversity of backgrounds to employ the techniques of cell culture successfully.

The terms tissue culture and cell culture are often used to describe the same processes. The original studies of Harrison and Carrell employed small biopsies of tissue or embryo from which discrete cells propagated. Most in-vitro testing today employs cloned cells and is properly named cell culture. In 1961, Hayflick and Moorhead^[4] carried out studies on the life span of cells in culture and stimulated subsequent interest in the development of continuous cell lines. These continuous or established cell lines became commercially available and led to biological test methods that introduced a degree of uniformity and standardization (Figs. 1 and 2).

The advantages of using cell culture tests in biomaterial evaluation are numerous. A significant saving of time and money is possible when compared with



Fig. 1 Optical photomicrograph showing normal murine L929 fibroblast cells, following methylene blue staining. The plasma membrane and nuclear membrane are clearly visible and the nuclear contents can be observed. Cell mitosis is also in evidence. This is an example of normal fibroblast cell line culture (original magnification $\times 200$).

the same material evaluation carried out *in vivo*. The numbers of animals required can be greatly reduced. A huge range of cell types can be employed *in vitro*. These include the use of human cells and other animal primary cell lines. The test environment can be carefully controlled and monitored. The exact composition of the test sample is known (or should be known) and can be altered at any stage in the procedure. The duration of

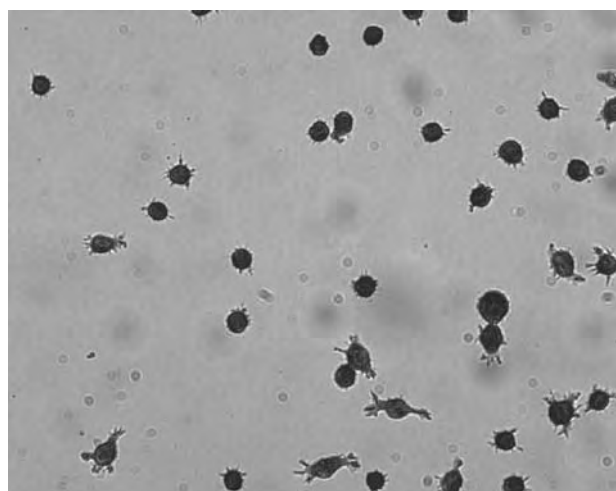


Fig. 2 Optical photomicrograph showing rounded murine L929 fibroblast cells, following methylene blue staining. Cell spreading has been inhibited following contact with a biomaterial substrate. This is an example of a reversible cellular effect (original magnification $\times 200$).

contact with cells is also easily controlled. The effect on the cells, induced by the test material, can be observed by microscope or quantified by cell counting or biochemical measurement of another cell parameter. A rapid result can be obtained if required. The potential for transfer of cell lines between laboratories allows for greater reliability in testing and the opportunity for standard testing.

It is easy to identify many differences between the *in-vitro* environment and the behavior of a cell *in situ*, and this provides the argument for the many skeptics who view cell culture as at best an inadequate alternative to animal testing and at worst an ambiguous and misused biological tool. However, many cell functions can be demonstrated and challenged in culture, and, as long as the limitations of the techniques are recognized, *in-vitro* testing will continue to prove a valuable resource in biomaterial evaluation. The methodology of culturing cell lines and deriving and culturing primary source cells is well established and many protocols are available.^[5]

A Confusion of Terms

The term biocompatibility is often associated with cell culture testing of biomaterials. In many cases simple *in-vitro* test methods have been employed with a potential biomaterial as the test sample. If a positive result is obtained, *i.e.*, the cells remain viable, the material is labeled biocompatible. If the cells die or are significantly adversely affected, the material is labeled nonbiocompatible. This is a serious misuse of the term biocompatibility and a misuse of the *in-vitro* test.

Biocompatibility is not the result of a single event or phenomenon. It refers to a collection of processes involving different but independent mechanisms of reaction between a material and its host tissue. Biocompatibility refers to the ability of a material to perform function. The ability to perform and to continue to perform this function depends not only on its interaction with tissue but also on the intrinsic mechanical and physical properties of the material. The definition of biocompatibility^[6] refers to appropriate host response. It does not stipulate that there should be no response. This may be a minimal response or indeed a more aggressive response. The definition refers to specific application. Biocompatibility should always be described with reference to the situation in which a material or device will be used. No material should be described as biocompatible without further qualification. Biocompatibility, therefore, is not an intrinsic property of any material. The majority of *in-vitro* tests employed in biomaterial evaluation measure cytotoxicity. However, lack of toxicity cannot be equated with biocompatibility. Cell culture assays are not biocompatibility assays.

BIOMATERIAL TESTING

Using cell culture to evaluate the effects of drugs was already a well-established procedure in the 1960s when such testing was first used to examine biomaterials. The studies of Rosenbluth^[7] and Guess^[8] detail the development of the agar overlay assay. A refined and standardized format of this test is presented by Autian^[9] as part of a structured program of biomaterial evaluation. The agar overlay test remains widely employed and reported today despite a number of significant limitations. These include the presence of the agar as a barrier between cell and test sample leading to a subsequent reduction in sensitivity. The test model itself provides an essentially nongrowing culture of cells that is incapable of revealing selective effects on the cell cycle. There is uncertainty in ascertaining an end point, although this is true of many of the cell culture tests in use today. No account is taken of lesser injury indiscernible by vital or nonvital staining. Nevertheless, it should be noted that the agar overlay test formed part of the earliest attempts to introduce a degree of uniformity and standardization to in-vitro biomaterial evaluation. This was true not only for the test procedure itself, but also in attempts to introduce an effective means of subsequent evaluation. The agar overlay test formed the foundation for the variety of in-vitro screening tests that are employed today in biomaterial evaluation and is one of the recommended test methods described in the relevant ISO standard.^[10]

Standards

The safety evaluation of medical devices is generally based on in-vivo testing and evaluates acute and chronic toxicity and inflammation, irritation and sensitivity, carcinogenicity and mutagenicity, allergic responses, haemocompatibility, and systemic toxicity. It is not possible to assess all of these factors using in-vitro assays; however, test models have now become established that permit the in-vitro evaluation of acute toxicity, genotoxicity,^[11] haemocompatibility,^[12] necrotic cell death, and apoptosis.^[13]

The introduction and development of in-vitro testing has raised many concerns with respect to reproducibility and proper regulatory control. The establishment of new in-vitro assays requires three steps: development, validation, and acceptance. Acceptance of in-vitro assays involves satisfying the regulatory bodies, the manufacturers, and the users. At present there are few, if any, properly validated in-vitro tests.^[14]

International standard ISO10993, part 5 was developed as the result of a harmonization of various existing standard procedures and the outcome reports of working groups. The standard acknowledges the

widespread use of in-vitro cytotoxicity testing and defines a scheme for testing that allows the user to identify the most appropriate test procedures. These tests specify the incubation of cultures of mammalian cells with the medical device or component part of the device (i.e., the direct test method) or with an extract derived from the device (i.e., the indirect test method). Tests using agar diffusion and filter overlay are also detailed here. Information is provided on appropriate test sample preparation for direct testing and the most appropriate means of producing extracts for indirect testing is presented in some detail. The standard provides details on a number of recommended established cell lines, although these are not exclusive. The standard further provides details on how cytotoxicity may be most appropriately determined. This evaluation is categorized as follows:

1. Assessment of cell damage by morphological means.
2. Measurement of cell damage.
3. Measurement of cell growth.
4. Measurement of specific aspects of cell metabolism.

This standard notes that evaluation may be by qualitative or quantitative means. It should also be noted that the standard does not present a single in-vitro method for cytotoxicity evaluation but allows the user to select an appropriate range of tests. Furthermore, at no point does the standard use the term biocompatibility in connection with this type of test. Another important aspect of the standard is the reference made to negative and positive control materials that should be employed in all test systems.

Replacement of animal testing should not be the objective of in-vitro assays, and an appropriate use of both in-vitro and in-vivo testing is the most sensible approach to biomaterial evaluation. Simple in-vitro assays can provide screening information and play a significant role in the early phase of investigation. Subsequent studies should employ more specific in-vitro models in parallel with animal studies. This level of testing should include the use of human cell lines.

Screening Tests and Cytotoxicity Testing

The test procedures identified in ISO 10993, part 5 are designed to evaluate cytotoxicity. The primary goal of such tests is to determine safety. As such, these can be considered screening tests for use in the development and research of novel materials and for the quality control of medical devices (both component parts and final products). Even at this simple level of testing relevant information can be garnered with respect to the final use of the test material. However, the user must be careful to appreciate the simplicity of such



testing. For example, it is prudent that any cytotoxicity evaluation should employ more than one cell line and where appropriate and possible make use of human primary cell lines. Even simple screening tests can benefit from some level of quantification where this can be sensibly applied. Although ISO 10993 does make reference to the use of subjective, semiquantitative grading systems based on microscopy observations, these will rarely produce any meaningful data and do not easily allow for the comparison of data between different laboratories or indeed different users. Qualitative, descriptive assessment of cell morphological changes by an appropriately trained individual remains a more meaningful technique. However, quantification using measurements of cell metabolic functions are readily employed and can provide accurate and reproducible data easily transferred between researchers. These include cell-counting procedures, both manual and automatic, assessment of DNA levels,^[15] or the determination of cell activity by the 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazoliumbromide (MTT) assay.^[16]

Specific Testing

Cytotoxins may have reversible or irreversible effects, which may be immediate or delayed. Toxic effects do not always result in cell death. A cytostatic cell response is often reversible and may be the result of deprivation of a nutrient or growth factor or the presence of a negative growth regulator.^[17] Establishing an appropriate end point for cytotoxicity studies is difficult as there is often no clear boundary between cytotoxic effects that will lead to cell death and other discernable but potentially irreversible effects. There is no single in-vitro test available that will supply all of the information required to assess a potential biomaterial.

The earliest in-vitro test was the Salmonella test for mutagenicity developed by Ames.^[18] This remains the established in-vitro procedure for evaluation of the potential mutagenicity of materials. In recent years a number of more sophisticated and hopefully more reliable test procedures have been introduced. Most cytotoxicity screening tests employ established cell lines. However, it is clearly possible to employ primary cell lines, in many cases cells derived from human source, as the basis for in-vitro tests. The introduction of primary cells and the development of methods of evaluation that provide information on aspects other than cytotoxicity may be considered a step beyond the simple screening tests detailed and recommended in ISO10993. Primary cell lines can provide cultures that are close to the biological nature of the tissue of origin. However, subculture of primary cells can be achieved for only a limited time period before the artificial culture conditions stimulate a

change in cell phenotype. Finite cell lines are generally restricted to fibroblasts and have, as the name suggests, a limited life span. The use of continuous culture i.e., a cell line that has the capacity for infinite survival, overcomes these deficiencies. Yet, continuous cultures are often derived from tumors and may not express the characteristics of the original tissue.

Although aspects of cytotoxicity remain relevant, it is appropriate to consider other cell parameters that in-vitro testing can challenge, e.g., cell adhesion, cell spreading and orientation, cell proliferation and metabolic stimuli, cell activation, and cell and tissue engineering.

Cell adhesion is an extremely important factor to assess when considering biomaterial performance. In the majority of cases cell-material adhesion is a positive and required property, and many studies on implant materials are designed to improve the degree of cellular adhesion. However, cell adhesion may be an undesirable property for a number of medical device applications. It may also be possible to design biomaterials such that they are selective in allowing cells to adhere, and in-vitro assays are ideal research tools for such studies.^[19]

Cell motility and spreading, which involves cell cytoskeletal reorganization, the establishment of focal contact points, and the response of cells to physical features on the biomaterial substrate are other areas of considerable interest to the biomaterial scientist^[20,21] with respect to cell guidance and tissue engineering applications.

A measurement of cell proliferation is widely used to determine the response of cells to biomaterial contact. Proliferation can be quantified by manual cell-counting techniques (haemocytometer) or electronic counting. Electronic methods are rapid and have lower inherent error compared to manual methods. Also, counting can be prone to misinterpretation as cell clumps and cell debris may be counted indiscriminately. Fluorescent or antibody labeling are now routinely used to determine cellular constituents and these methods allow for accurate quantification by flow cytometry.^[22]

The MTT test is now widely in use for biomaterial evaluation. It is a quantitative measure of the metabolic activity of cells that measures the activity of the mitochondrial enzyme succinate dehydrogenase. The test relies upon the reduction of the tetrazolium salt component of the MTT to formazan, an aqueous insoluble compound. Using an appropriate solvent, the quantity of formazan can be determined by spectrophotometry. It is important to note that the MTT assay cannot distinguish between cell loss and metabolic inhibition. Furthermore, stringent controls must be employed when examining short-term cultures.

Most viability assays rely on damage to the integrity of the cell membrane. This is readily examined by

measuring the ability of the cell to uptake or exclude dyes such as methylene blue or neutral red.

It should be noted that there are a large number of test models available. Users should take care that they select methods of testing that are most likely to provide the information desired. In-vitro testing must be employed in such a way as to provide a rigorous examination of the material under test.

Meaningful evaluation of the results obtained is undoubtedly the major challenge with any cell culture test system. Ideally the test should have an evaluation program that is easily and reliably repeatable. The difficulty lies in producing an accurate method of quantitatively evaluating the effect a material may have on the cells, such that the test produces a meaningful result. Cell culture testing should of course be only one stage in any biocompatibility evaluation and the results obtained from such testing should not be used alone to make a judgement about a material. However, there are a number of factors that can be taken into account in order to maximize the usefulness of these tests.

Measurements of pH are easy to make but may not provide a great deal of useful information. Cells are initially incubated in a buffered environment. In a static system a normal rise in cell population will inevitably lead to a fall in pH levels. Some test materials will cause a rapid and deleterious change in pH levels leading to cell death. However, in the static and highly sensitive test system this may not be a relevant measure of cytocompatibility.

Cultured cells are prone to genetic instability. This is clearly demonstrated as primary cell cultures undergo dedifferentiation in culture. Continuous cell lines may

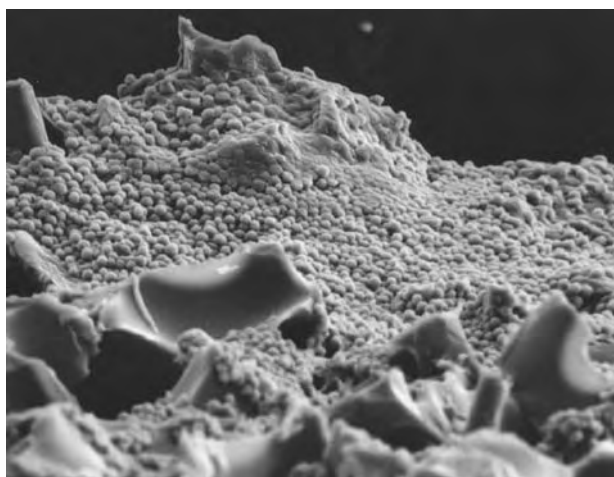


Fig. 3 Scanning electron micrograph showing murine L929 fibroblast cells cultured within a three-dimensional polyether (urethane) biomaterial scaffold. The cells are demonstrating an altered morphology; however, viable cells proliferate within this structure for time periods in excess of 20 weeks (original magnification $\times 200$).

produce a cell population that is significantly different from the source tissue. Cell cultures also lack the cell–cell and cell–extra-cellular matrix interactions of tissue, which are required to define phenotype. However, the significant developments made in tissue engineering in recent years have seen the creation of three-dimensional cell constructs, usually with biomaterial scaffolds in which the cell–cell and extra cellular matrix contacts are developed and maintained (Fig. 3).^[23]

In recent years an examination of cell adhesion molecules (CAMs) and cytokines, regulators of immune and inflammatory responses, have routinely been used as markers of cell activation and potential cytocompatibility.^[24]

REQUIREMENTS OF THE IN-VITRO TEST

Three types of in-vitro assay are generally available. Short-term assays of viability, long-term assays measuring cell survival, and intermediate assays that measure cell proliferation or a cellular metabolic parameter as a measure of viability or toxicity.

The disadvantage of viability assays is that they often overestimate cell survival and are often subjective. Growth curves can be used to compare cytotoxic effects or growth inhibition. It should be noted that cell numbers at any specific point of time are dependent on many factors including the lag phase of the cell population, the initial cell inoculum, the effect any toxic insult has on an initial inoculum, and the doubling time of the cell line.

There are two basic approaches to in-vitro testing when employed in the evaluation of biomaterials. One is to determine the potential for a negative effect on the cells. This is the way most in-vitro testing is employed. This is particularly true of the identification of cytotoxic or mutagenic effects. However, it is increasingly true that biologically active medical devices are designed such that they have a positive and beneficial effect on cell growth or cell repair. In-vitro testing can readily be employed to monitor this.

If the in-vitro test is to be presented as a viable supplement as well as an alternative to animal experimentation, there are a number of requirements it must satisfy. The most obvious and widely quoted is a reduction in cost and time. However, there are other requirements that must be taken into consideration. Dividing and metabolizing cells must be in contact with the material or an appropriate extract of the material under test. This is not the case with the often-quoted agar overlay test. The test method must be easily and reliably repeated, and there should be a simple and rapid means of evaluation. Such testing can be readily carried out; however, tests should not be used to simply indicate pass or fail. Where possible an



accurate quantitative means of recording results should be employed.

LIMITATIONS OF THE IN-VITRO TEST

Throughout cytotoxicity testing procedures for bio-material evaluation, the major disadvantage in the use of cell cultures is the difficulty in interpreting the results obtained in a clinically significant manner. The objective is to provide meaningful results. Although test samples can be examined in contact with human primary cells, the response of such cells cannot be assumed to be the same as the response produced by cells that are part of the human body. Complex interactions and responses are absent during these tests. An often cited criticism of cell culture testing is that it is too sensitive. This is the result of the relatively static conditions that normally exist in the Petri dish. These conditions exclude the detoxification pathways provided in the body by the kidney and liver. When operating cell culture assays, a number of practical problems may be encountered that will affect the way the test procedure is evaluated. In many cases, cytotoxicity is encountered as a result of substances leaching from a bulk/solid material. Test samples may not be homogeneous in nature, and substances may leach at different rates from the same sample. It is usually necessary that a device is broken down into its constituent parts for cytotoxicity testing. These individual component parts should be examined individually and in combination. Occasionally, a nontoxic or inert material is produced through the combination of two or more toxic or volatile constituents. In such a case it must be established that, under the conditions in which the device will be used, there is no likelihood of the device breaking down into its constituent parts. To satisfy this requirement the device should be examined under a variety of extreme conditions. Normally, however, materials can readily be examined as components of a final product.

When examining a number of test materials it is important to ensure that, as far as possible, the test conditions are the same for all materials. It is not difficult to ensure that there is no significant difference in the number of cells added to each test well or dish; however, the manner in which cells are added can have a significance that is often overlooked. Cells may be added to the tissue culture dish before the addition of the test material. A suspension of cells may be introduced to the dish and test materials at the same time. Cells may be added exclusively to the surface of the test material. The consequences of each technique should be considered when evaluating cellular response.

One of the most significant disadvantages of the agar overlay test method is the presence of agar in the

test environment. However, many direct contact test methods employ some form of adhesive or anchoring material. The presence of a third party in the test system will have a considerable significance on the response of cells. This can be true even with the use of commercially available tissue culture polystyrene dishes.

Contamination by bacteria and fungi is the most commonly encountered hazard in the cell culture laboratory. Despite the use of laminar flow cabinets that provide a sterile working air flow and the presence of antibiotics in the culture medium, proper aseptic technique is essential. It is appropriate that cell laboratory workers be appropriately trained and follow strict operating procedures. The media used for the cultivation of mammalian cells in culture provide perfect growth conditions for the majority of bacteria and fungi. Most of these organisms will grow more rapidly than the animal cells, whose environment they have invaded, and will eventually create conditions toxic to the cells. Contamination in a laboratory can rapidly spread and wipe out entire stocks of cells. The most obvious means of avoiding contamination is to ensure that there are no unwanted organisms present at any stage of the procedure. This is achieved by good aseptic technique and by following proper sterilization procedures.

Test materials from devices, which are normally supplied nonsterile but are sterilized before use, should be sterilized by the method recommended by the manufacturer and subsequently handled aseptically throughout any extraction or direct contact testing. The effects of sterilization methods on the material or device must be considered. Any material introduced into a cell culture environment should be sterile, via heat treatment (steam/dry), gamma irradiation, or gaseous ethylene oxide.

Safety

The fact that primary cell cultures are capable of propagating viral contaminants means they should always be considered potentially infectious.^[25] It is important that researchers involved in the culture of cells in vitro, particularly primary cell cultures, be able to adequately assess the risks involved and take the steps required to provide appropriate protection. A guide to the relevant risk assessment is provided by Jones.^[26]

CONCLUSION

In-vitro testing is now widespread and becoming ever more complex and sophisticated. Developments in cell biology and genetics and the significant deployment of

cell culture testing in toxicology studies are advancing such test methods ever closer to the day when they may effectively replace animal testing. At present, however, this is not the case, and it is not likely that *in vitro* will replace *in vivo* in medical device evaluation in the foreseeable future. Indeed, it is apposite that in the majority of cases *in vitro* will not replace but supplement animal testing.

The simplest form of *in-vitro* models in biomaterial evaluation is the commonly used screening tests that, by an assortment of means, measure cytotoxicity. More specific test methods are available. These may make use of human cell lines, primary cell lines, and testing regimes that challenge the medical device or biomaterial in an environment that simulates much more closely the intended clinical application. Great care must be taken in the evaluation of all *in-vitro* test methods, and the significant limitations must be considered during any means of evaluation. *In-vitro* testing should be only one stage in a program of biomaterial evaluation. Cytotoxicity tests do not assess biocompatibility and should not be used to do so.

ARTICLES OF FURTHER INTEREST

Biocompatibility Testing; Biological Effects of Radio-frequency Electromagnetic Field; Electric Cell-Substrate Impedance Sensing; In Vivo Cardiovascular Models

REFERENCES

1. Paul, J. *Cell and Tissue Culture*, 3rd Ed.; Churchill Livingstone Press: Edinburgh, 1975.
2. Harrison, R.G. Observations on the living developing nerve fibre. *Proc. Soc. Exp. Biol. N.Y.* **1907**, *4*, 140–143.
3. Carrell, A. On the permanent life of tissues outside the organism. *J. Exp. Med.* **1912**, *15*, 516–528.
4. Hayflick, L.; Moorhead, P.S. The serial cultivation of human diploid cell strains. *Exp. Cell Res.* **1961**, *25*, 585–621.
5. Freshney, R.I. *Culture of Animal Cells, a Manual of Basic Technique*, 4th Ed.; Wiley-Liss: New York, 2000.
6. Williams, D.F. *Progress in Biomedical Engineering, Definitions in Biomaterials*; Elsevier: Amsterdam, 1987; Vol. 4, 54.
7. Rosenbluth, S.A.; Weddington, G.R.; Guess, W.L.; Autian, J. Tissue culture method for screening toxicity of plastic materials to be used in medical practice. *J. Pharm. Sci.* **1965**, *54*, 156–159.
8. Guess, W.L.; Rosenbluth, S.A.; Schmidt, B.; Autian, J. Agar diffusion method for toxicity screening of plastics on cultured cell monolayers. *J. Pharm. Sci.* **1965**, *54*, 1545–1547.
9. Autian, J. Toxicological evaluation of biomaterials: Primary acute toxicity screening program. *Artif. Organs* **1977**, *1*, 53–60.
10. *ISO 10993 The Biological Evaluation of Medical Devices, Part 5, Tests for Toxicity—In vitro Methods*; International Standards Organisation 1995. ISO 10993-5.
11. Chauvel-Lebret, D.J.; Auroy, P.; Tricot-Doleux, S.; Bonnaure-Mallet, M. Evaluation of the SCGE assay to assess the genotoxicity of biomaterials. *Biomaterials* **2001**, *22*, 1795–1807.
12. *ISO 10993 The Biological Evaluation of Medical Devices, Part 4, Selection of Tests for Interactions with Blood*; International Standards Organisation 1995. 10993-4.
13. Renvoize, C.; Biola, A.; Pallardy, M.; Breard, J. Apoptosis: Identification of dying cells. *Cell Biol. Toxicol.* **1998**, *14*, 111–120.
14. Ferro, M.; Doyle, A. Standardisation for *in vitro* testing. *Cell Biol. Toxicol.* **2001**, *17*, 205–212.
15. Kirkpatrick, C.J.; Mittermayer, C. Theoretical and practical aspects of testing potential biomaterials *in vitro*. *J. Mater. Sci., Mater. Med.* **1990**, *1*, 9–13.
16. Ciapetti, G.; Cenni, E.; Pratelli, L.; Pizzoferrato, A. *In vitro* evaluation of cell/biomaterial interaction by MTT assay. *Biomaterials* **1993**, *14*, 359–364.
17. Freshney, I. Application of cell cultures to toxicology. *Cell Biol. Toxicol.* **2001**, *17*, 213–230.
18. Ames, B.N. The Detection of Chemical Mutagens with Enteric Bacteria. In *Chemical Mutagens: Principles and Methods for Their Detection*; Hollaender, Ed.; Plenum: New York, 1971; 267–282.
19. Kirkpatrick, C.J.; Bittinger, F.; Wagner, M.; Kohler, H.; van Kooten, T.G.; Klein, C.L.; Otto, M. Current trends in biocompatibility testing. *Proc. Inst. Mech. Eng.* **1998**, *212 Part H*, 75–84.
20. Meyle, J.; Gultig, K.; Brich, M.; Hammerle, H.; Nisch, W. Contact guidance of fibroblasts on biomaterials surfaces. *J. Mater. Sci., Mater. Med.* **1994**, *5*, 463–466.
21. Flemming, R.G.; Murphy, C.J.; Abrams, G.A.; Goodman, S.L.; Nealy, P.F. Effects of synthetic micro- and nano-structured surfaces on cell behaviour. *Biomaterials* **1999**, *20*, 573–588.
22. Pizzoferrato, A.; Ciapetti, G.; Stea, S.; Cenni, E.; Arciola, C.R.; Grachi, D.; Savarino, L. Cell culture methods for testing biocompatibility. *Clin. Mater.* **1994**, *15*, 173–190.
23. Hoffman, R.M. The 3-dimensional question—Can clinically relevant tumour drug resistance be measured *in vitro*? *Cancer Metastasis Rev.* **1994**, *13*, 169–173.
24. Cenni, E.; Granchi, D.; Ciapetti, G.; Verri, E.; Cavada, D.; Gamberini, S.; DiLeo, A.; Pizzoferrato, A. Cytokine production and adhesive protein expression by endothelial cells after contact with polyethylene terephthalate. *Biomaterials* **1996**, *17*, 2071–2076.
25. Stacey, G.; MacDonald, C. Immortalisation of primary cells. *Cell Biol. Toxicol.* **2001**, *17*, 231–246.
26. Jones, B.; Stacey, G. Safety considerations for *in vitro* toxicology testing. *Cell Biol. Toxicol.* **2001**, *17*, 247–270.



Cell-Material Interaction

Aditya Chaubey

Karen J.L. Burg

Department of Bioengineering, Clemson University, Clemson, South Carolina, U.S.A.

INTRODUCTION

The fate of any biomedical device depends on its initial contact with the surrounding tissue, because proteins start adsorbing on its surface within seconds. The intrinsic property of the material, including its chemistry, governs the quantity and orientation of the protein that is adsorbed on its surface, which, in turn, regulates the performance of the device. Hence, to successfully create a clinically viable cell-based device, the initial interaction between the material and the cells must be understood. Although much progress has been made toward better understanding this process, there are still many issues to resolve before whole organs can be developed.

Biomaterials have evolved from inert materials solely providing mechanical support and function to materials that actively elicit a response from the surrounding tissue. As such, polymeric biomaterials are key components in cell-based devices, providing support and direction for proliferating cells. Both the bulk and the surface properties differently affect the nature of a polymer; hence, surface modification is used to modulate the surface properties to control the proliferation and differentiation of cells.

BIOMATERIALS: A BASIC INTRODUCTION

Organ transplantation is currently a standard clinical treatment for patients suffering from permanent organ failure. The practice of transplanting organs has existed for many decades and, with the considerable improvement in surgical techniques and immunosuppressive drugs, has become a viable clinical option. However, the biggest problem associated with organ transplantation is that the demand for organ replacement far exceeds the available supply; the difference will only increase with time. Tissue engineering in the past decade has emerged as a key technology in the potential regeneration of whole organs.

One form of tissue engineering involves harvesting healthy cells from a patient's body, culturing those cells in vitro, seeding them on an absorbable scaffold, and then implanting the cellular construct into the patient's body. The scaffold directs the cells to form tissue, while slowly absorbing. A more formal

definition of tissue engineering is “application of the principles and methods of engineering and the life sciences towards the fundamental understanding of structure–function relationships in normal and pathological mammalian tissues and the development of biological substitutes to restore, maintain or improve tissue function.”^[1] The most crucial aspect for the success of a tissue-engineered device is the interaction between the cells and the scaffold material onto which the cells are seeded. In these unique devices, biomaterials serve to provide three-dimensional, structural environments for the cells. Tissue engineering scaffolds should encourage proliferation and differentiation of the particular cell type and should be designed with the local implantation site in mind.

Biomaterials were traditionally defined as materials that could be implanted in the body to perform a specific function and interact in a defined, positive manner, ideally causing minimal tissue reaction. The earliest hip implants, for example, were made with smooth surfaces from chemically nonreactive materials like titanium or stainless steel. These materials fulfilled their desired function and were termed “inert” due to the low level of associated foreign body response.^[2,3] The lack of normal tissue in growth, due to implant smooth surface textures, resulted in high rates of component loosening and bone absorption, highlighting the need for improved implant fixation.^[4,5] These early failures led to the use of implants with microporous surfaces, which offered better stabilization at the bone–implant interface by allowing bone ingrowth. The observation that a material could actively encourage the surrounding tissue to act in a specific and desired way led to a change in the direction of the biomaterials industry and also a new definition of “biomaterial.” Hence, a biomaterial can be defined as any synthetic or natural material that is implanted in the body and elicits a specific reaction from the body at either a local or a systemic level. This revised definition marked an increased effort to understand the interaction between an implant material and surrounding tissue.

BIOMATERIALS CLASSIFICATION

The majority of biomaterials can broadly be separated into three classes: metals, ceramics, and polymers.

Polymers can be subdivided into synthetics or naturally derived plastics, which can be further subclassified as absorbable/degradable or nonabsorbable. Common metals used as implants include titanium, stainless steel, and other alloys. Metals are most commonly used for load-bearing applications, although they are also sometimes used in other devices requiring high mechanical integrity, like stents. Issues of note with metal implants include wear, corrosion in the body environment, and the release of metal ions in the body. Bioactive ceramics like calcium phosphate ceramics and bioactive glasses avoid these problems. The surface of a hip implant, for example, may be coated with bioactive glass to form a thin layer of calcium phosphate. This layer of calcium phosphate stimulates the adjacent bone cells to form new bone,^[6] and thus, in accordance with the contemporary understanding of biomaterials, plays an active role in inducing a favorable reaction from the body. The third class of biomaterials is polymers; polymers are increasingly finding more use in a wide variety of applications, including scaffolds in tissue engineering applications. Some of the most common synthetic polymers used are polyurethanes (used to form pacemaker leads), poly(2-hydroxyethyl methacrylate) (pHEMA; used to form contact lenses), poly(L-lactide) (PLL), polyglycolide (PG), and poly(lactide-coglycolide) (PLG; used to manufacture sutures and drug delivery vehicles). PLL, PG, and PLG are absorbable, which makes them very attractive for tissue engineering applications. PLL, PG, and PLG are among the best characterized absorbable polymers and are used extensively in a variety of clinical applications. Because many variables govern the interaction between a cell type and a biomaterial, a complete characterization of the bulk and surface properties of a biomaterial is absolutely necessary for a thorough understanding of the interaction between the two.

EFFECT OF MATERIAL PROPERTIES ON CELLS

The interaction between a material and the surrounding tissue depends on a number of material properties, such as the material wettability, surface chemistry, rigidity, and topography. This interaction can be visualized in terms of cell functions like attachment, locomotion, proliferation, and differentiation. Wettability of a surface can be measured as a function of contact angle and is an indicator of surface energy of the material. Materials with a high surface energy (for example, ceramics and metal oxides) have high affinity for serum proteins, the presence of which in turn increases cell attachment.^[7] In the absence of serum proteins, cell attachment is enhanced on positively charged surfaces. Materials having an

intermediate wettability are most often best suited for cell attachment. The rigidity of a material surface also exerts an influence on the attachment of cells. Generally speaking, so called “anchorage-dependent” cells require a solid substratum for growth and maintenance of function. Many anchorage-dependent cells do not proliferate and, instead, retain a rounded shape on materials with very low modulus, such as hydrogels. The texture of an implant material is also an important parameter in determining its success and is designed according to the intended application of the implant. A rough surface increases the overall surface area of the implant surface and hence provides more potential for cell attachment; however, the effect of texture is dependent on the cell type. For example, it has been demonstrated that whereas fibroblasts prefer smooth surfaces for attachment, macrophages prefer relatively rougher surfaces, a phenomenon termed rugophilia.^[8] The roughness of an implant surface can also alter the affinity of cell types to hydrophobic and hydrophilic surfaces. The geometrical configuration of an implant influences its biocompatibility by determining the level of development and orientation of the capsule formed as a result of the normal foreign body response. An implant with sharp corners, for example, induces higher inflammatory activity around the corners, with the eventual preferential formation of a thick, dense capsule at these regions.^[2,3,9] Excellent review articles in the literature overview various surface properties and their effect on the cellular response. Properties/responses reviewed include surface energy,^[10] surface charge, and cell spreading.^[11]

Three-dimensional scaffolds are meant to provide support to the cells or tissues by providing an appropriate three-dimensional architecture. Additionally, these scaffolds can be tailored to provide proper cell adhesion and proliferation, which in turn will promote better signaling pathways, consequently leading to better incorporation of the cell-material construct in the body. Surface modification has emerged as an important tool to modulate surface properties and render materials more attractive to any particular application. For example, it has been demonstrated that when tissue-culture polystyrene (TCPS) is coated with pHEMA, cell spreading decreases with increasing amounts of pHEMA^[12]; however, when pHEMA is treated with sulfuric acid, cell spreading is greatly enhanced. Cell adhesion and spreading are similarly increased when TCPS is treated with glow discharge or radiofrequency plasma deposition. It is believed that surface modification primarily increases the adsorption of serum proteins and thereby leads to an increased cell attachment, rather than directly affecting the cell attachment to the surface. PL and PG, for example, can be easily processed into complex 3-D shapes, are well characterized, are approved by the U.S. Food and



Drug Administration for specific *in vivo* applications, and are particularly well suited to tissue engineering applications because they are biodegradable. Despite all these advantages, their surface properties, which are not ideal for cell adhesion and proliferation, make them unsuitable for many applications. These absorbable polyesters lack functional groups on the polymer chain; the presence of methyl groups in PLL makes it highly hydrophobic in nature. Because cell growth and proliferation are better suited to hydrophilic surfaces, PLL in unmodified state is not ideal for cell attachment. However, cell affinity to PLL surfaces may be increased by surface modification. Some routinely used methods for this purpose are plasma treatment^[13], entrapment of a functional group within the polymer chain^[14], and partial hydrolysis by acid and base^[15] treatments. Surface modification can also be accomplished by incorporating extracellular matrix (ECM) molecules on the surface of the implant material; these molecules enhance cell attachment because cells recognize them through their adhesion ligands and thus are more likely to adhere to an ECM-coated surface. Liu and coworkers, for example^[16], reported that surface modification of poly(α -hydroxy acids) (PHAs) could be accomplished by immobilizing gelatin on the PHA surface, thus facilitating increased cell attachment and proliferation and augmenting the deposition of collagen fibers and other cell secretions. Although modifying a 2-D surface (e.g., a thin polymer film) is relatively straightforward, the same is not true for a 3-D scaffold, because other parameters such as bulk porosity and surface features increase the scaffold complexity; modification treatments are limited by the rate of diffusion.

NATURAL BIOMATERIALS

Synthetic biomaterials that are biodegradable in nature are well suited to tissue engineering applications. Synthetics are, for example, easily processed into 3-D shapes, reproducibly manufactured on a large scale, and have controlled mechanical and degradation properties. Also, because these materials absorb with time, they do not lead to lingering, long-term deleterious effects. However, a critical shortcoming of many synthetic biomaterials is their lack of cell-recognition sites. The success of any tissue engineered device is based on the attachment of cells to the substrate; in general, cell adhesion involves the cell surface receptor recognition of small molecules called ligands that are located in the extracellular space. The presence of specific ligands on a biomaterial surface, commonly true of naturally derived biomaterials, greatly enhances and facilitates the process of cell adhesion to any surface. Natural biomaterials include polysaccharides,

polypeptides, amino acids, and their combinations. Some specific examples of natural biomaterials are collagen, fibronectin, fibrin, chitosan, and gelatin. The advantages offered by naturally derived biomaterials can often include increased cellular affinity and similarity of structure to the surrounding tissue. However, these biomaterials also have disadvantages, such as increased susceptibility to infection, antigenicity, limited production scale-up potential, and unpredictable degradation properties over an extended period of time. The process of designing a 3-D scaffold from natural materials is also challenging because of the difficulty associated with controlling design parameters like architecture and controlled porosity.

Collagen is one of the most commonly used natural materials as a scaffold. It is composed of three chains wound into a triple helix, is the most abundant protein in the body, and is responsible for structural integrity. More than 20 different kinds of collagen types are found in the body, affording different properties. Type I collagen is the most common among all the types of collagen and is the chief component of tendons, cartilage, and bones. Collagen is mainly synthesized by cells responsible for laying down the ECM, such as fibroblasts, myofibroblasts, osteoblasts, and chondrocytes. Because collagen has high mechanical strength, encourages cell attachment, offers good resistance to degradation, and has low antigenicity, it can be used in a wide variety of applications. The mechanical stability and the rate of degradation of collagen can be enhanced by chemical crosslinking and, hence, many different agents such as glutaraldehyde, tannic acid, and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) have been used to crosslink collagen. An example of a tissue engineered product containing collagen is Apligraf[®] (manufactured by Organogenesis and marketed in the United States by Novartis), a bioengineered skin composed of a bovine collagen scaffold containing human fibroblasts and overlaid with a differentiated epithelium. However, this product is allogenic in nature (i.e., the cells do not belong to the patient on whom the product is intended to be used) and, hence, although the cells provide growth factors that enhance wound healing, complete integration of these cells in the patient's body likely does not occur. DuraGen[™] is another clinical product that uses collagen; this product is an absorbable collagen matrix graft that is used in spinal and cranial dural repair. DuraGen has had limited success in many patients requiring dural repair following anterior and posterior spinal surgery.^[17] Arginine-glycine-aspartic acid (RGD) peptides form a sequence of amino acids, found routinely in the ECM, that promote cell adhesion. Modification of collagen scaffolds by incorporation of RGD peptides leads to a significant increase in the attachment of cells on the

scaffold without any significant increase in the proliferation of cells.^[18,19]

Fibronectin is a common glycoprotein that is involved in many cellular processes including tissue repair, blood clotting, and cell migration/adhesion. Fibronectin sometimes serves as a general cell adhesion molecule by binding cellular integrins to collagen or proteoglycan substrates. Fibronectin matrix formation is initiated by fibronectin binding to cell-surface receptors, followed by assembly and reorganization of the cell-surface-associated fibronectin into fibrils. Three-dimensional fibrin gel scaffolds have demonstrated promise for vascular engineering applications.^[20] Furthermore, when the initial cell-biomaterial interaction of silicone is modified by covalently immobilizing fibronectin on its surface,^[21] silicone samples appear to have higher cell affinity. The selection of a particular natural material as a material to modulate cell attachment or proliferation depends on the specific application or the cell type under consideration. Hyaluronic acid (HA) is a naturally occurring biopolymer that has desirable properties and potential as scaffolding for tissue engineering. HA is a major component of connective tissue and of the synovial fluid, and it functions to cushion and lubricate various joints, such as the knee joint. A promising use of HA is in alleviating the pain caused by osteoarthritis. Supplements of hyaluronic acid help restore the health of the joint by increasing the thickness of the synovial fluid. This fluid lubricates the joint, which can allow a greater range of motion with less pain. Hyaluronan binding proteins on cell surfaces are HA receptors and mediate the effects of HA on the cell (especially cell adhesion and migration). Researchers have demonstrated that hyaluronic acid induces chondrogenesis of mesenchymal stem cells.^[22] The structure of HA can be chemically modified to generate a wide range of physical properties to better aid in the differentiation of chondrocytes. Natural biomaterials, because of the intrinsic nature that promotes their interaction with different cell types, hold considerable promise in the field of tissue engineering.

SYNTHETIC MATERIALS WITH RGD PEPTIDES

Synthetic materials may be surface modified by covalently binding proteins or peptides that are involved in various adhesion and proliferation signal pathways. The result affords the predictability of a synthetic biomaterial as well as the superior interaction that natural materials offer with the cells or the surrounding tissue. The most common material covalently attached to a synthetic biomaterial surface and used to enhance cell adhesion to a biomaterial is the RGD peptide sequence. Integrins are membrane proteins

found in the cellular plasma membrane that serve to bind to ligands found in the extracellular space. Integrins link the intracellular cytoskeleton of cells with the extracellular matrix by recognizing and interacting with ligands, such as RGD peptides. Thus, ligands such as RGD may cause the relay of information into the cell that triggers processes associated with cell attachment, growth, division, survival, differentiation, and apoptosis. RGD peptides found on biomaterials interact with the cell integrin receptor sites, which can, in turn, initiate cell signaling processes and influence the cellular adhesion rate of the cells on a material.^[23]

Ceramics are extremely biocompatible and also display osteoconductive properties; hence, a large number of calcium phosphate biomaterials have been proposed for use as bone filler materials. However, in addition to inadequate bulk load-bearing properties, these materials also lack any significant osteoinductive properties. The covalent attachment of defined densities of RGD peptides onto the surface of hydroxyapatite can increase the attachment and differentiation of osteoblasts,^[24] specifically causing increased DNA content, alkaline phosphatase activity, matrix protein expression, and mineralized nodule formation. Other research demonstrated that incorporation of RGD peptide on polymer surfaces for small-diameter vascular graft applications allows enhanced cell adhesion and spreading.^[25]

MICROTOPOGRAPHY AND ITS INFLUENCE ON CELL BEHAVIOR

Tissues and the cells that comprise the tissues in the body are arranged in precise spatial patterns, where the type of patterning observed depends on the tissue type and function. Microscopic evaluation of the structure of bone, for example, reveals a definite and characteristic pattern of arrangement. The ground substance of bone is arranged in concentrated layers (lamellae) around the small canals that run parallel to the long axis (shaft) of the bone. These canals, called haversian canals, are interconnected with one another via Volkmann's canals and contain a blood vessel, a nerve, and a lymph vessel. Each haversian canal is surrounded by concentric layers of bone matrix (called lamellae) and concentric rings of bone-forming cells (osteoblasts). This compact and highly defined structure maximizes the load-bearing ability of the bone. The configuration demonstrates the complex and definitive architecture of tissue in the body and the role of the structure in achieving its functionality. Chemical and physical cues are provided to the proliferating cells to arrange themselves in the requisite patterns. Logic suggests that for a tissue engineered device to be



successful in vivo it must not only provide the biochemical environment of the native tissue, but also provide a spatial arrangement that facilitates normal tissue function. Over the years many efforts have been made by researchers to study the interaction of cells with materials of a defined surface microtopography.^[16] Cells respond to surface features and react to them, leading to what is termed “contact guidance.” Efforts include study of the behavior of nerve cells cultured on spiderwebs and coverslips, as well as the evaluation of cell orientation on randomly scratched (grooved) mica substrates. These types of studies contributed to our basic understanding of cell interaction with microtopography. However, the studies fail to address, among other things, the structural dimensions of the tissues and precise repeatability over 3-D areas, comparable to that in vivo.

Advances in microfabrication technology opened new avenues for fabrication of substrates that have dimensional and chemical features similar to that found in tissues. Use of microfabrication technology enables the control of cell–surface interactions at the micro- or nanometer scale. Many reports in the literature suggest that surfaces with microgrooves affect the orientation and migration of cells.^[9,27] However, the mechanism and even the exact role of the microgrooves is poorly understood because it is difficult to compare reported studies as a result of methodology differences. For example, cell alignment is significantly affected by differences in groove depth, but differences in groove width are not very critical in influencing cell alignment.^[28] Several reports in the literature claim that microtexture is not correlated to cell proliferation; alternately, many reports claim exactly the opposite.^[29] Clearly, there is much work to be done in this area in order to provide viable tissue-engineered devices, because they are dependent on the adhesion and differentiation mechanisms/behaviors of cells on scaffolds. Surfaces with rough microtopography, for example, enhance osteoblast differentiation by encouraging osteoblasts to secrete factors that enhance osteogenesis, while decreasing osteoclast formation and activity.^[30] Researchers have demonstrated that, in many instances, implant surface topography effects can override implant chemistry, and by modulating the surface topography, one can radically alter the implant/tissue interface and overall biocompatibility.^[9] Similarly, micropatterning of natural materials has been investigated to enhance cell–surface interactions and the tissue–implant interface by spatially directing the growth of cells. Cellular differentiation and proliferation are affected by the interaction among hormones, growth factors, surface chemistry, and surface microtopography.^[31] To control cell-material interactions, a huge array of factors must be considered, including the form of the

material, the bulk and surface chemistry, the surface texture, the surface charge, and the biochemical and mechanical attributes of the implant site. Understanding of cell-material interaction has increased markedly over the past years, but there remain many important questions to answer and issues to address in order to realize the full potential of biomedical device design.

CONCLUSIONS

Biomaterials have evolved from providing simple mechanical support to providing support while purposefully influencing a desired response from the surrounding tissue. At the heart of this advance is the increase in our understanding of the cell-material interaction at both the micro and the macro level. As interest in developing cell-based devices grows, so will the necessity for extensive and well-controlled, repeatable studies of cell-material interactions.

ARTICLES OF FURTHER INTEREST

Biocompatibility Testing; Biological Effects of Radio-frequency Electromagnetic Field; Biomaterials Immune Response; Biomaterials: Protein–Surface Interactions; Thrombosis

REFERENCES

1. Skalak, R.; Fox, C.F. *Tissue Engineering*; A.R. Liss: New York, 1988.
2. Anderson, J.M. Inflammatory response to implants. *Trans. Am. Soc. Artif. Intern. Organs* **1988**, *XXXIV*, 101–107.
3. Woodward, S.C. Evaluation by light microscopy. In *Handbook of Biomaterials Evaluation*, 2nd Ed.; von Recum, A.F., Ed.; CRC Press: Boca Raton, FL, 1999; 599–612.
4. Adam, F.; Hammer, D.S.; Pfautsch, S.; Westermann, K.J. Early failure of a press-fit carbon fiber hip prosthesis with a smooth surface. *Arthroplasty* **2002**, *17* (2), 217–223.
5. Otani, T.; Whiteside, L.A. Failure of cementless fixation of the femoral component in total hip arthroplasty. *Orthop. Clin. North Am.* **1992**, *23* (2), 335–346.
6. Hench, L.L.; Paschall, H.A. Direct chemical bond of bioactive glass–ceramic materials to bone and muscle. *J. Biomed. Mater. Res.* **1973**, *7* (3), 25–42.
7. Baier, R.E. Implant dentistry forefront ‘85. Surface preparation. *J. Oral Implantol.* **1986**, *12* (3), 389–395.
8. Rich, A.; Harris, A.K. Anomalous preferences of cultured macrophages for hydrophobic and roughened substrata. *J. Cell Sci.* **1981**, *50*, 1–7.

9. von Recum, A.F.; Shannon, C.E.; Cannon, C.E.; Long, K.J.; van Kooten, T.G.; Meyle, J. Surface roughness, porosity, and texture as modifiers of cellular adhesion. *Tissue Eng.* **1996**, *2* (4), 241–253.
10. Kaelble, D.H.; Moacanin, J. A surface energy analysis of bioadhesion. *Polymer* **1977**, *18* (5), 475–482.
11. Sugimoto, Y. Effect of adhesion and locomotion of mouse fibroblasts by their interacting with differently charged substrates. *Exp. Cell Res.* **1981**, *135* (1), 39–45.
12. Folkman, J.; Moscona, A. Role of cell shape in growth control. *Nature* **1978**, *273* (5661), 345–349.
13. Gugala, Z.; Gogolewski, S. Attachment, growth, and activity of rat osteoblasts on polylactide membranes treated with various low temperature radiofrequency plasmas. *J. Biomed. Mater. Res. Part A.* **2006**, *76A* (2), 288–299.
14. Janorkar, A.V.; Fritz, E.W.; Burg, K.J.L.; Metters, A.T.; Hirt, D.E. Grafting amine-terminated branched architectures from poly(L-lactide) film surfaces for improved cell attachment. *J. Biomed. Mater. Res. Part B. Appl. Biomater.* **2007**, *81B*, 142–152.
15. Atthoff, B.; Hilborn, J. Protein adsorption onto polyester surfaces: Is there a need for surface activation? *J. Biomed. Mater. Res. Part B. Appl. Biomater.* **2007**, *80* (1), 121–130.
16. Liu, X.; Won, Y.; Ma, P.X. Surface modification of interconnected porous scaffolds. *J. Biomed. Mater. Res. A* **2005**, *74* (1), 84–91.
17. Narotam, P.K.; Jose, S.; Nathoo, N.; Taylon, C.; Vora, Y. Collagen matrix (DuraGen) in dural repair: Analysis of a new modified technique. *Spine* **2004**, *29* (24), 2868–2669.
18. Hsu, S.-H.; Chang, S.-H.; Yen, H.J.; Whu, S.W.; Tsai, C.-L.; Chen, D.C. Evaluation of biodegradable polyesters modified by type II collagen and arg-gly-asp as tissue engineering scaffolding materials for cartilage regeneration. *Artif. Organs* **2006**, *30* (1), 42–55.
19. Zhang, L.; Hum, M.; Wang, M.; Li, Y.; Chen, H.; Chu, C.; Jiang, H. Evaluation of modifying collagen matrix with RGD peptide through periodate oxidation. *J. Biomed. Mater. Res. Part A* **2005**, *73A* (4), 468–475.
20. Ye, Q.; Zund, G.; Benedikt, P.; Jockenhoevel, S.; Hoerstrup, S.P.; Sakyama, S.; Hubbell, J.A.; Turina, M. Fibrin gel as a three dimensional matrix in cardiovascular tissue engineering. *Eur. J. Cardiothorac. Surg.* **2000**, *17* (5), 587–591.
21. Siggelkow, W.; Gescher, D.M.; Siggelkow, A.; Klee, D.; Malik, E.; Rath, W.; Faridi, A. In vitro analysis of modified surfaces of silicone breast implants. *Int. J. Artif. Organs* **2004**, *27* (12), 1100–1108.
22. Hegewald, A.A.; Ringe, J.; Bartel, J.; Kruger, I.; Notter, M.; Barnewitz, D.; Kaps, C.; Sittinger, M. Hyaluronic acid and autologous synovial fluid induce chondrogenic differentiation of equine mesenchymal stem cells: A preliminary study. *Tissue Cell.* **2004**, *36* (6), 431–438.
23. Rowley, J.A.; Mooney, D.J. Alginate type and RGD density control myoblast phenotype. *J. Biomed. Mater. Res.* **2002**, *60* (2), 217–223.
24. Itoh, D.; Yoneda, S.; Kuroda, S.; Kondo, H.; Umezawa, A.; Ohya, K.; Ohyama, T.; Kasugai, S. Enhancement of osteogenesis on hydroxyapatite surface coated with synthetic peptide (EEEEEEPRGDT) in vitro. *J. Biomed. Mater. Res.* **2002**, *62* (2), 292–298.
25. Fussell, G.W.; Cooper, S.L. Endothelial cell adhesion on RGD-containing methacrylate terpolymers. *J. Biomed. Mater. Res. A* **2004**, *70* (2), 265–273.
26. Folch, A.; Toner, M. Microengineering of cellular interactions. *Annu. Rev. Biomed. Eng.* **2000**, *2* (1), 227–256.
27. Campbell, C.E.; von Recum, A.F. Microtopography and soft tissue response. *J. Invest. Surg.* **1989**, *2* (1), 51–74.
28. Clark, P.; Connolly, P.; Curtis, A.S.; Dow, J.A.; Wilkinson, C.D. Topographical control of cell behavior: II. Multiple grooved substrata. *Development* **1990**, *108* (4), 635–644.
29. Matsuzaka, K.; Walboomers, F.; de Ruijter, A.; Jansen, J.A. Effect of microgrooved poly-L-lactide (PLA) surfaces on proliferation, cytoskeletal organization, and mineralized matrix formation of rat bone marrow cells. *Clin. Oral Implants Res.* **2000**, *11* (4), 325–333.
30. Lossdorfer, S.; Schwartz, Z.; Wang, L.; Lohmann, C.H.; Turner, J.D.; Wieland, M.; Cochran, D.L.; Boyan, B.D. Microrough implant surface topographies increase osteogenesis by reducing osteoclast formation and activity. *J. Biomed. Mater. Res. A* **2004**, *70* (3), 3361–3369.
31. Boyan, B.D.; Bonewald, L.F.; Paschalis, E.P.; Lohmann, C.H.; Rosser, J.; Cochran, D.L.; Dean, D.D.; Schwartz, Z.; Boskey, A.L. Osteoblast-mediated mineral deposition in culture is dependent on surface microtopography. *Calif. Tissue Int.* **2002**, *71* (6), 519–529.

Centrifugal Blood Pumps

Gerald E. Miller

The Artificial Heart Laboratory, Virginia Commonwealth University, Richmond, Virginia, U.S.A.

INTRODUCTION

The use of centrifugal blood pumps for cardiovascular support and heart assist have resulted from the marriage of a very successful class of standard fluid pumps with an urgent need to stem one of the largest causes of death in the United States—cardiovascular and heart pathologies. Centrifugal pumps have long been used as water pumps, oil pumps, air fans, and a host of other industrial, automotive, space, and home applications. They are one class of rotary pumps, the others being axial pumps and mixed-flow pumps. The use of such a pumping technology to address a growing need for cardiac support is a natural extension of these devices, and has resulted in the development of dozens of such pumps for use to propel blood. Over the last two decades, the use of rotary blood pumps has steadily grown to the point that they are now employed clinically as often as the more historically known positive displacement pumps, such as the older Jarvik pumps, and the latest pumps such as the HeartMate pump, the LionHeart pump, the Abiomed pump, and so on.

Centrifugal pumps offer advantages over other pumping modalities in that they can produce large pressure increases and large flow rates at relatively low rotation speeds. As the history of centrifugal pumps for nonmedical applications is long, the developments and advances in such a technology have been significant. Many centrifugal pumps have been used for specialized fluids, including those that are highly volatile (such as rocket fuel) as well as those that have suspensions of particles (such as motor oil). Development of centrifugal pumps for such applications has resulted in versions that are designed for sensitive fluids and can thus produce stable flow regimes. The use of such pumps to propel blood is thus a natural extension, as this is another sensitive fluid that has deformable particles and that requires stable flow. It is thus not surprising that many centrifugal pumps are now employed for cardiovascular support and are in various stages of clinical and bench-top trials.

CENTRIFUGAL PUMPS—GENERAL INFORMATION

Centrifugal pumps are one class of rotary pumps or rotodynamic pumps that are sometimes called

turbomachinery. Fluid motion is directed within a housing by means of blades, vanes, or impellers (propellers that are within a housing). The housing has no valves and thus is not closed, contrary to positive displacement pumps (such as piston pumps and diaphragm pumps used to propel blood), all of which have valves and closed volumes. The use of a housing for a centrifugal pump is to collect fluid along or from the rotating elements and direct it towards the outlet. Fluid normally enters the pump housing within the “eye” of the pump near the center (above) the rotor and leaves at the periphery. A centrifugal pump increases the pressure of the fluid from a small value at the inlet to a larger value at the outlet. This is much the same way that air flows within a tornado or hurricane. The center or eye of the tornado has little flow and is a source of low pressure. However, at the periphery, due to the spinning action of the tornado, the flow rate is high, as is the pressure.

For centrifugal pumps, flow and pressure increases are affected by several factors including size of the rotating element (blades, rotor, impeller), number of rotating elements, their configuration (slanted, straight, curved), size and shape of the housing, and rotation rate of the rotating elements. A typical centrifugal pump with housing and rotating element is shown in Fig. 1. Standard measures of performance and design analysis include the specific speed N_s , which relates the pressure head (pressure rise) and volume flow rate when operating at optimal efficiency. $N_s = N \times Q^{1/2} / H^{3/4}$, where N = rotation rate of the impeller, Q = volume flow rate, and H = pressure head (rise).^[1] The net positive suction head (NPSH), which evaluates the pump inlet pressure with respect to vapor (dissolved gas) pressure and is a measure of the tendency toward cavitation (bubbling of gases) in the fluid; and pump energy can be obtained from Bernoulli's equation. $N \times T = \{(P_{out} - P_{in}) / \rho\} + \{(V_{out}^2 - V_{in}^2) / 2\} + g (Z_{out} - Z_{in})$, where T = torque, ρ = fluid density, V = fluid velocity, g = coefficient of gravity, and Z = elevation/height.^[1]

Centrifugal pumps, unlike their positive displacement counterparts, are not affected by inlet pressure. In effect, they draw their own inlet vacuum and thus are not affected by preload, the venous pressure that supplies blood for heart assist devices. This is contrary to the natural heart and to all positive displacement

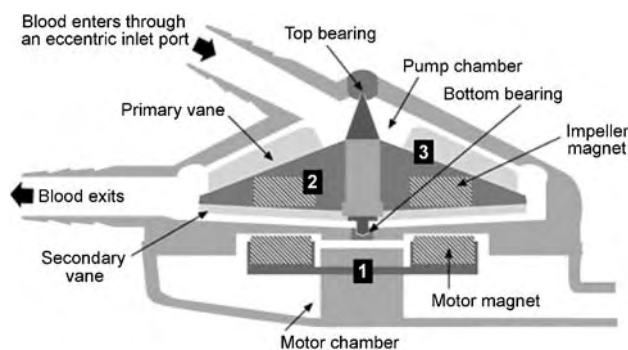


Fig. 1 Typical design of a centrifugal blood pump with flow entering above the rotor and exiting peripherally. The rotor can be magnetically rotated or shafted. The central inlet flow above the rotor is at a low pressure and the exit flow is at a higher pressure. The resulting pressure rise is called the pressure head.

blood pumps (such as the HeartMate and other diaphragm/piston type blood pumps), all of which are affected by inlet pressure and flow conditions. If venous return to the natural heart is reduced, cardiac output is reduced. Centrifugal pumps are not so affected. This is a great advantage over traditional blood pumping systems.

However, the outlet flow rate of centrifugal pumps is affected by the outlet load pressure that the fluid faces upon leaving the pump. If the outlet (load) pressure is high, then the pump must produce flow against a greater load. Because centrifugal pumps are designed to produce a pressure rise as opposed to a guaranteed outlet flow, then a large afterload to the device will adversely affect the outlet flow rate. It is possible in such pumps to produce a “dead head” effect—the pump spins, but no flow occurs. This is due to the pressure at the outlet (tubing/vessel) being larger than the pressure at the outlet port of the pump. The solution for such situations is to increase the rotation rate of the centrifugal pump, which supplies more energy to the fluid, increases the outlet pressure, and results in increased outlet flow of the pump. This adverse effect of increased afterload on outlet flow of centrifugal pumps represents a disadvantage over positive displacement pumps that produce an outlet flow regardless of afterload pressure, thus, their name—positive displacement, meaning they displace a certain volume of blood (stroke volume) regardless of the load against which they pump. The cardiovascular analogy is that aortic hypertension adversely affects centrifugal pumps while not affecting positive displacement pumps. However, decreased venous return does not affect centrifugal blood pumps but it does adversely affect positive displacement pumps. A description of the effects of afterload on the performance of centrifugal pumps at various rotation rates is shown in Fig. 2.

Centrifugal pumps offer other advantages over the more classical positive displacement blood pumps. They do not require valves and thus avoid calcification, pannus formation, and thrombus formation that have often plagued positive displacement blood pumps.^[2–4] They usually require less energy to operate, as there is not a large piston or diaphragm to push against. They can operate in either steady or pulsatile mode and are quite responsive to changing drive conditions in that the rotation rate can be adjusted “on the fly” if there are feedback mechanisms to be employed. As there is no fixed and constrained volume, there is no need for a compliance chamber, as has been required for many positive displacement pumps.^[1,5] Thus, the use of centrifugal pumps to propel blood and serve as cardiac assist offers significant promise. It is no wonder that numerous versions are in various stages of clinical trials for cardiac assist and replacement.

CENTRIFUGAL BLOOD PUMPS

Centrifugal pumps have been investigated to propel blood for several decades, producing versions that have different rotating elements and different housing configurations. Rotating elements, known as rotors, consisting of bladed impellers, stacked cones, parallel discs, and ridge-type rotors, have all been employed in various designs of centrifugal blood pumps. Housings that have straight inlets centered above the rotor, or canted inlets with the opening above the rotor, or offset inlets have all been employed. The only constant among all versions of such pumps has been that the fluid enters above the rotor in a region of low pressure and radial flow, and exits at the periphery in a region of high pressure and radial flow. Centrifugal pumps thus convert axial flow into radial flow through the

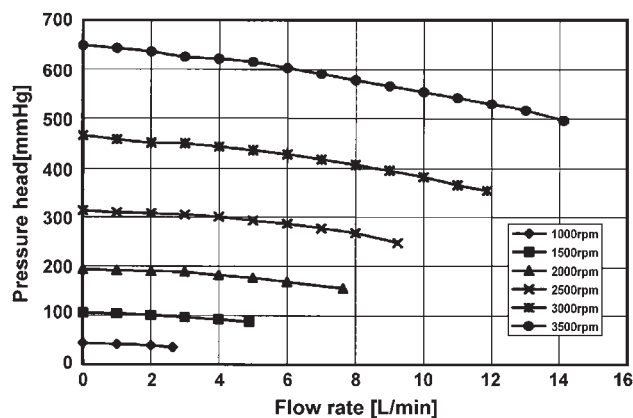


Fig. 2 Typical centrifugal blood pump relationship between pressure head and flow rate as a function of pump rpm. Pressure head drops as a function of outlet flow rate for a constant rotation rate.

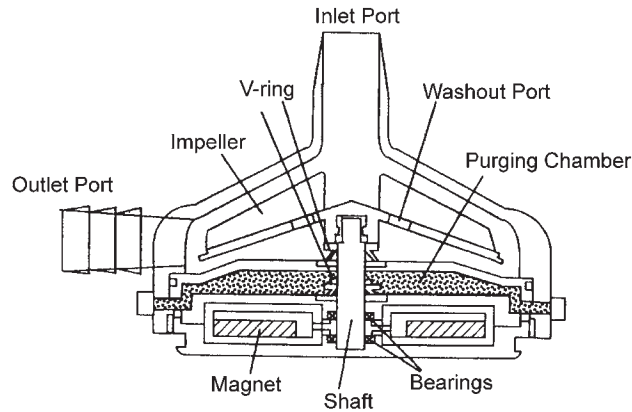


Fig. 3 Baylor-Nikkiso centrifugal blood pump with purging chamber below the rotor to reduce thrombosis as a result of blood stagnation. (From Ref. [6].)

rotating element at the base of the housing. The inlet and outlet ports of the pump are typically perpendicular to each other. Most centrifugal pumps have conical housings that conform to the tornadolike fluid flow fields that are generated by the impeller/rotor assembly. This housing naturally increases in diameter from the inlet (at the top center of the device) toward the periphery where the radial fluid flow collects to exit the pump. A typical centrifugal blood pump and housing (using a centrally placed rotor/impeller) are 50–75 mm in diameter and 30–40 mm high, weigh 200–300 g, and use a brushless DC motor.

Many centrifugal pumps have employed a shaft through the housing (with a shaft seal), which connects the rotor to the drive motor. Other versions have employed some form of magnetic suspension of the rotor, which might achieve stability by means of pivot bearings that ride within a cavity at the base of the housing. Still other pumps produce complete magnetic suspension including both levitation and radial stabilization of the rotor within the housing. Many of the developed centrifugal blood pumps have resulted from

collaboration between companies and universities/medical centers, including several collaborations between Japanese firms and U.S. medical centers.

One such example is the Baylor-Nikkiso centrifugal pump which has been under investigation for many years and represents a collaboration between the Japanese firm Nikkiso and the Baylor College of Medicine in Houston.^[6–10] This pump has a shaft and shaft seal as well as magnetic coupling. Together, these provide radial and axial stability as well as a mechanism to spin the impeller. One unique aspect of this design is the use of a fluid purging system to avoid stagnation near the shaft seal by providing a washout flow rate via a purging fluid. This pump is shown with the purging system in Figs. 3 and 4, indicating the effects of purging and nonpurging versions on the development of thrombus formation. The pressure/flow relationship for this pump is presented in Fig. 5 and is typical of centrifugal blood pumps. At higher rotation rates of the impeller, the flow rates increase, as do the outlet pressures at the periphery/outlet of the pump, in part to counterbalance any tendency for a flow decrease as a result of increased afterload.

The Baylor College of Medicine has also collaborated on the Baylor Gyro series of pumps, which utilize magnetic coupling to spin the rotor, but also utilize a pivot bearing to allow the rotor some degree of radial and axial stability. This is a collaboration with Kyocera Corporation. The spinning rotor rides above a bearing with a conical projection from the rotor that sits above the bearing, acting like a gyroscope. This pump is shown in Fig. 6. One interesting aspect of the Baylor Gyro series of pumps is that various inlet configurations have been tested, including canted inlets and offset inlets in an attempt to optimize flow patterns within the housing.^[11–19] An eccentric inlet port design is presented in Fig. 7 and the pressure/flow relationship of this pump can be seen in Fig. 8. Again, the family of pressure/flow relationships for various rotor speeds indicates higher rotation rates resulting in

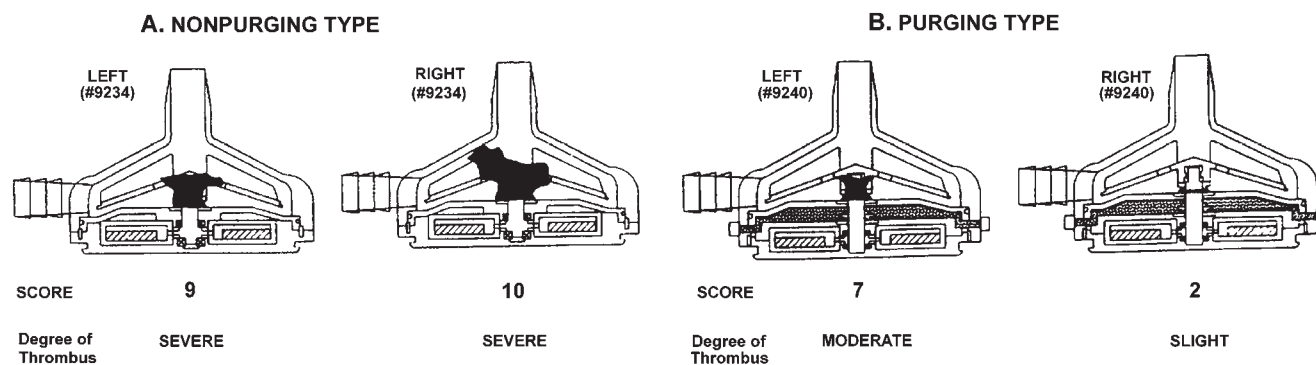


Fig. 4 Baylor-Nikkiso centrifugal blood pump: the effect of purging (A) vs. nonpurging (B) of the spacing below the rotor. More thrombosis occurs in the nonpurged system, indicating areas of blood stagnation. (From Ref. [6].)

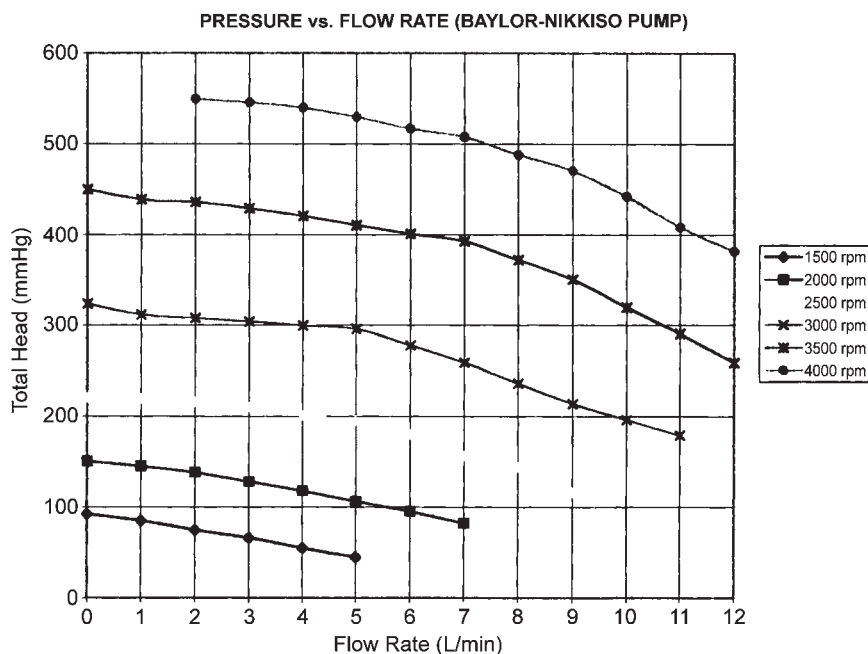


Fig. 5 Baylor-Nikkiso centrifugal blood pump performance curves depicting pressure head as a function of outlet flow rate for varying pump rpm. Pressure head drops as a function of flow rate for a constant rpm, typical of centrifugal pumping systems. (From Ref. [7].)

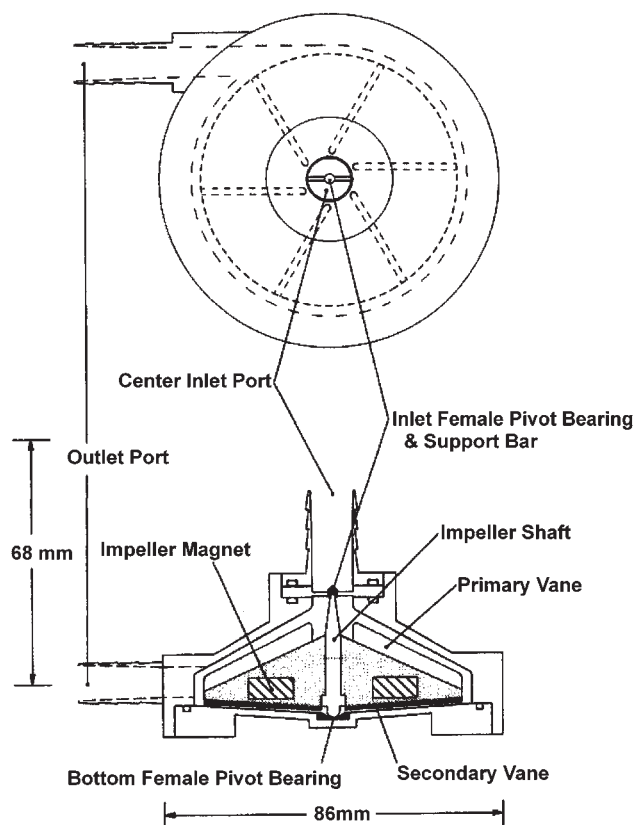


Fig. 6 Baylor Gyro centrifugal blood pump with a pivot bearing and magnetically driven rotor. Hydrodynamic forces lift the rotor within the housing with the pivot bearing serving as a guide. (From Ref. [13].)

higher outlet pressures and flow rates in order to counterbalance any tendency for a flow decrease as a result of increased afterload.

Another mechanism to propel blood by means of a centrifugal design utilizes spinning disks within a spiral volute housing. This work was initiated at Texas A&M University in collaboration with the Texas Heart Institute and was later continued at Virginia Commonwealth University. This system employs multiple, donut-shaped, rotating disks with fluid entering axially into the spacing at the center of the disks. The fluid is then propelled via shearing and centrifugal forces within the spacing between the disks to provide radial flow.^[20-25] The spiral

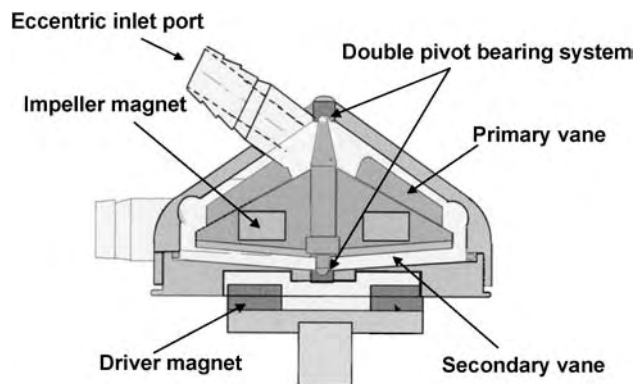


Fig. 7 Baylor Gyro pump with diagonal input. Output is at the periphery as is common with other centrifugal pump designs. (From Ref. [13].)



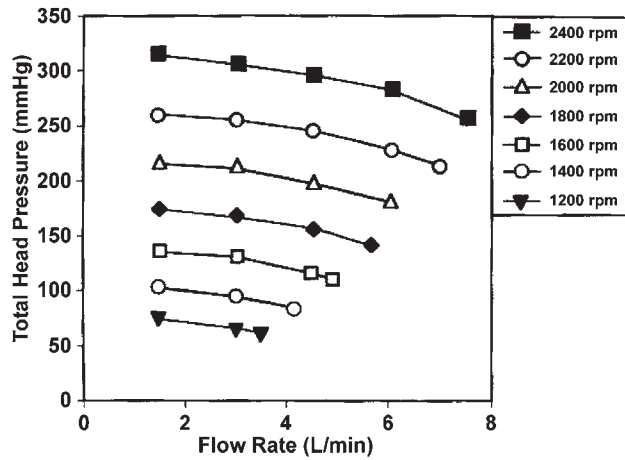


Fig. 8 Baylor Gyro centrifugal blood pump performance curves indicating pressure head as a function of outlet flow rate for varying rpm. There is a slight decrease in pressure head for higher flow rates, which is typical of centrifugal pumping systems. (From Ref. [13].)

housing collects the radially spun fluid and redirects the fluid toward the outlet at the periphery to reproduce axial flow, albeit at right angles to the inlet flow. This pump is based upon the Tesla Turbine, a multiple-disk pump that has previously been investigated for use to propel liquid rocket fuel for over five decades.^[26–29] The multiple-disk pump is shown in Fig. 9 and its pressure/flow relationship is presented in Fig. 10. Unlike other centrifugal

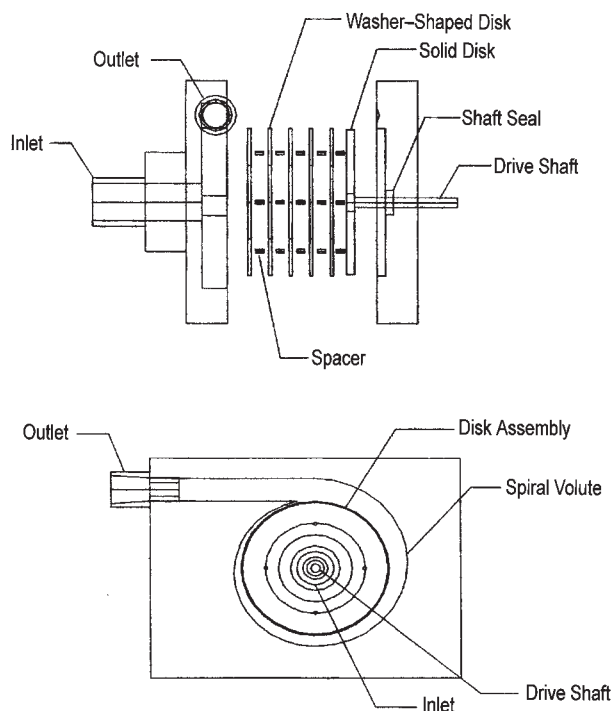


Fig. 9 Multiple-disk pump with parallel rotating disks and a spiral volute housing. (From Ref. [21].)

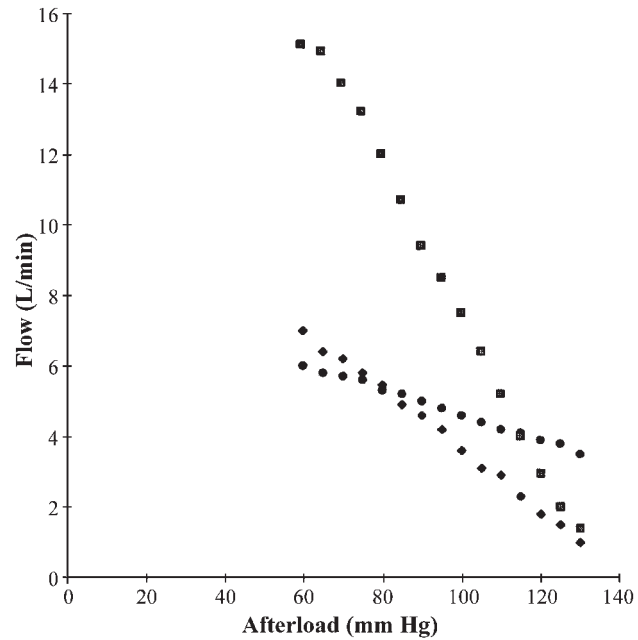


Fig. 10 Multiple-disk blood pump performance chart depicting flow rate as a function of afterload, the load against which the pump must produce flow. (From Ref. [24].)

blood pumps, the housing of the multiple-disk pump is circular, not conical, as it surrounds the circular-shaped disk assembly.

One of the more clinically successful pumps in terms of overall usage is the BioMedicus BP-80 pump and its successors.^[30–32] This pump utilizes magnetic forces to not only suspend, but also rotate the rotor. The rotor consists of a series of stacked cones with small blade-like projections on the cones to assist in propelling the fluid. As with most centrifugal pumps that utilize rotating elements located at the base of the housing, the flow field within the housing resembles a vortex with the fluid entering above the rotor and spinning outward from the center of the housing toward the periphery, where it builds momentum, pressure, and velocity. In fact, the BioMedicus pump is advertised to produce a “constrained forced vortex.” One advantage of this design is that the rotor has a magnetic disk at its base, which is married to another magnetic disk in the drive unit that exists outside the housing. This allows the entire pump head to be twisted off the drive unit and replaced with a clean pump head, should the old head be affected by thrombosis. This pump is depicted in Fig. 11 and, like most impeller-type pumps, has a conical housing.

The Cleveland Clinic has developed the CorAide centrifugal blood pump, which has been bench-tested on a mock circulatory system^[33] and tested in calves.^[34,35] The pump is designed to be permanently implantable and consists of a rotating assembly that is



Fig. 11 BioMedicus BP 80 pump with rotating cones and a full-bottom magnetic disk connected to a driving magnet. (From Ref. [30].)

completely suspended and noncontacting (shaftless). This pump operates in steady flow mode and has produced low hemolysis, plasma-free hemoglobin, and thrombosis levels that require less anticoagulant therapy. This pump is shown in Fig. 12.

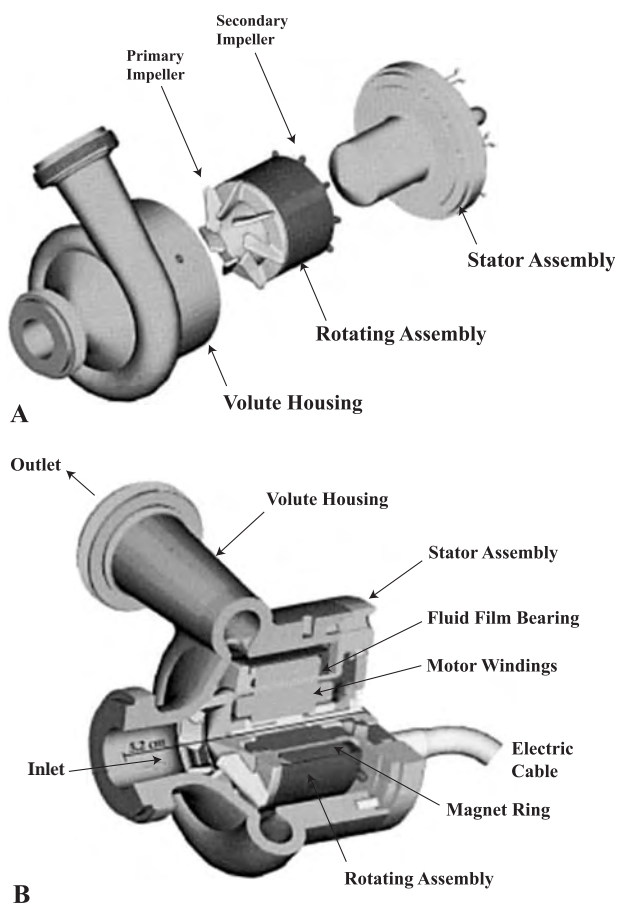


Fig. 12 Cleveland Clinic CorAide centrifugal blood pump. (From Ref. [33].)



Fig. 13 HeartMate III centrifugal blood pump. (From Ref. [3].)

The HeartMate III centrifugal blood pump^[2-4] has been used as a left ventricular assist device and was developed by ThoratecCardiosystems, the same company that designed the positive displacement, pusher plate Heartmate pump.^[5] This pump operates with a magnetically suspended, levitated, and rotated rotor and has produced physiological pressures and flow rates in experimental studies in calves. The pump design is presented in Fig. 13 and its performance (pressure vs. flow vs. rpm) can be seen in Fig. 14.

The aforementioned pumps are but a sampling of numerous designs from Japan, China, Austria, France, Korea, Singapore, Australia, and the United States, among others. The majority of devices under development have originated from either Japan or the United States and often involve the collaboration of both countries. Many centrifugal pumps under development have similar designs with a centrally placed impeller/rotor, a conical housing, and pressure/flow characteristics that indicate decreasing flow at higher load pressures and a series of performance curves with higher flow rates at higher rotation rates of the impeller assembly. All of these devices have been tested on the bench (in vitro) using flow visualization techniques to evaluate the flow field and predict any regions of flow irregularities, as well as pressure/flow measurements to evaluate pump performance. Hemolysis testing has been accomplished both in vitro and in vivo to determine the efficacy of these pumps as viable blood pumping devices. Thrombosis has also been evaluated in these pumps by examining the excised pumps after a significant period of pumping to visually determine any regions where flow stagnation has resulted in thrombus formation.

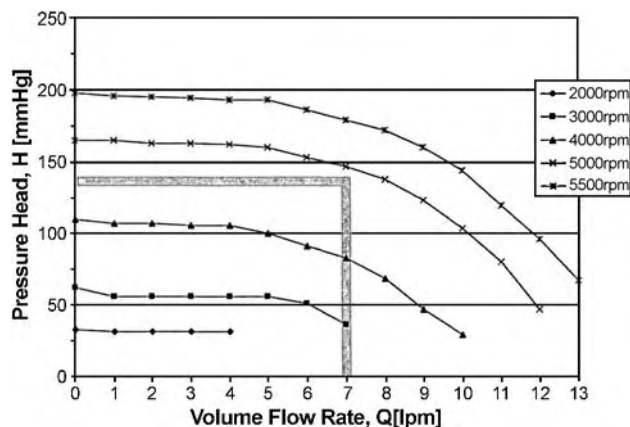


Fig. 14 HeartMate III centrifugal pump performance curves depicting pressure head vs. flow rate for varying pump rpm. Pressure head decreases at higher flow rates typical of centrifugal pumps. (From Ref. [4].)

EVALUATION OF CENTRIFUGAL BLOOD PUMPS

In vitro evaluation of centrifugal blood pumps employs a mock circulatory system that produces a physiologic load on the device, which is similar to that which may be experienced in the body during clinical or animal trials. A typical mock circulatory system consists of a fluid capacitor, fluid resistor, and venous reservoir as shown in Fig. 15. The capacitor is a fluid chamber with an elastic back-resistance element that consists of either a leaf spring or a pressurized air chamber, separated from the fluid chamber by an elastic diaphragm. By varying the leaf spring position or the back pressure of the air chamber, one can vary the compliance of the capacitance element and thus simulate a variety of physiological and/or pathological conditions. The capacitance chamber moves up and down either rapidly or slowly depending on the back-resistance setting. During steady flow, the compliance chamber is inactive, as there are no flow pulsations to cause any movement of the fluid chamber. The fluid resistor element is merely tubing placed between two flat plates with wing nuts (or a drive motor) controlling the compression of the tubing between the two plates. The tubing cross-sectional area is thus manipulated, which produces a controllable fluid resistance by creating a larger or smaller pressure drop resulting from the area reduction increment. The inlet pressure to the pump is affected by the height of fluid in the venous reservoir.

Although centrifugal pumps are not affected by preload conditions (inlet/venous pressure), the use of a venous reservoir is designed to mimic the reduction in fluid perturbation aspects of the human venous system. A blood analogue, typically a glycerin/water mixture (40/60%), is employed in the mock circulatory system

to avoid the need to use real blood (and the resulting thrombosis that may occur) and to allow various fluid visualization techniques to be employed. The blood analogue produces a viscosity (3.5–4 centipoise) and density (1.05 g/cc) similar to those of whole blood, and is a commonly employed technique for in vitro testing.

Flow visualization techniques are often employed to evaluate the flow field within the pumping chamber of centrifugal blood pumps as well as within the inlet and/or outlet cannula of such pumps. Particle image velocimetry (PIV) is a commonly used technique to evaluate the flow fields, and consists of lasers that illuminate neutrally buoyant particles embedded within the blood analogue, a digitizing camera to monitor the tracks produced by the illuminated particles, and controlling software to analyze the flow field data. A typical particle image velocimetry system is shown in Fig. 16. The system produces snapshots of velocity vectors (movement of numerous particles along pathlines) such as that depicted in Fig. 17. By examining the particle tracings, one can determine regions where stagnation occurs (which may lead to thrombosis,) or regions where large shear rates occur (which may lead to hemolysis). By placing various portions of the housing or cannula within the camera viewpoint, one can zoom in on regions of particular interest. Numerous snapshots can be summed and averaged to produce trend graphs, such as that shown in Fig. 18. Other flow visualization systems employ laser Doppler anemometer (LDA) techniques whereby the laser energy illuminates particles or the fluid itself, and the scattering or reflection is monitored as an indication of fluid velocity—hence the Doppler aspect to this system. The output of such a system produces snapshots or trend graphs similar to that generated by the PIV

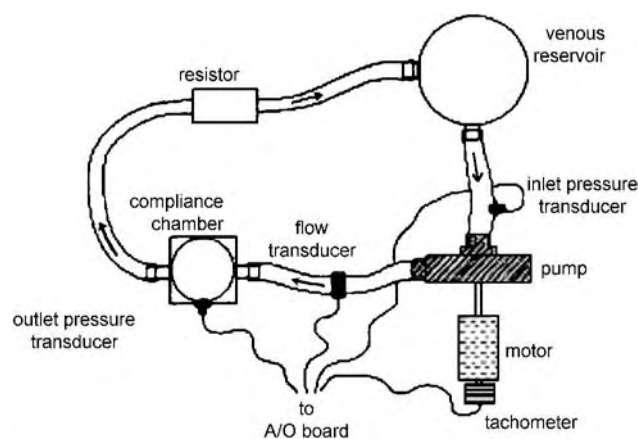


Fig. 15 Typical mock circulatory system for in vitro testing of artificial hearts and ventricular assist devices. The system mimics the circulatory load that an artificial heart must face.

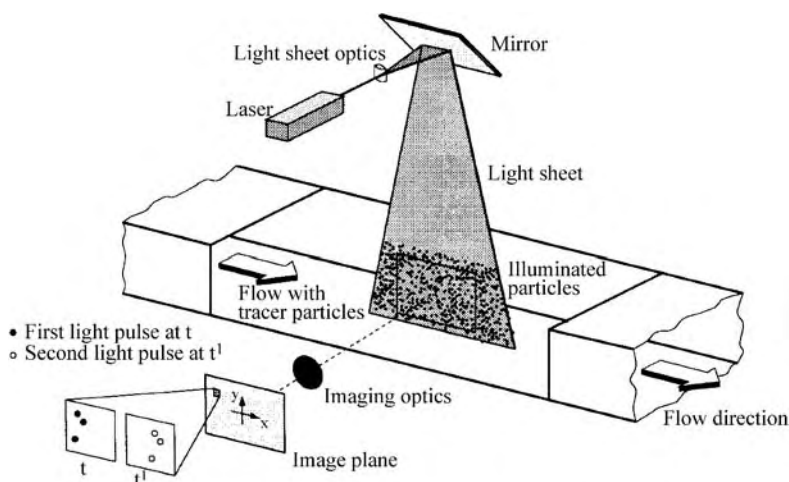


Fig. 16 Particle image velocimetry (PIV) system for visualizing the velocity patterns in and around ventricular-assist devices and artificial hearts. The lasers illuminate fluorescent particles imaged by a charged coupled device (CCD) camera. (From Ref. [25].)

system, albeit by different means. A typical LDA system is shown in Fig. 19.

Hemolysis testing, both *in vitro* and *in vivo*, is commonly utilized to evaluate the efficacy of centrifugal blood pumps to serve as long-term cardiac assist devices. The pumps operate for an extended period using whole blood, with the blood sampled periodically for plasma-free hemoglobin, an indication of excessive lysis of red cells. For those studies conducted *in vivo*, there are other tests that are conducted to evaluate the performance of the pump in a physiological setting including platelet levels and activation, renal function, and cardiac perfusion. Often, these pumps are compared to each other and to the roller pump (the commonly used pump for cardiopulmonary bypass) in terms of hemolysis. A typical hemolysis curve for centrifugal pumps is shown in Fig. 20 and indicates a slow rise in plasma-free hemoglobin toward a plateau. Despite the nature of the vortexlike flow field within the typical conical housing of a centrifugal blood pump, the level of hemolysis generated by such pumps is minimal. This is likely due to the short duration that red cells are in contact with the neighboring zone of the

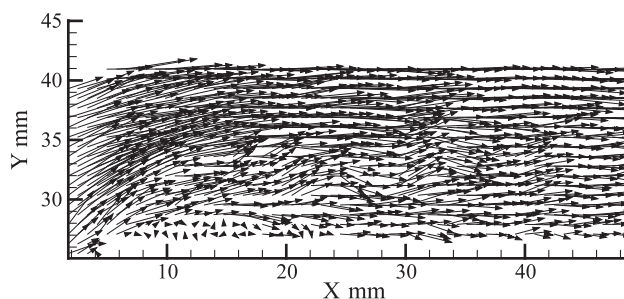


Fig. 17 Vector mapping of the velocity field from a PIV flow visualization system. This is a snapshot image of the neutrally buoyant particles in the flow field. The vectors (arrows) depict the direction of each particle and the arrow length indicates the velocity magnitude. (From Ref. [25].)

rotor/impeller. Although the rotation speeds of most centrifugal blood pumps range from 1200–4000 rpm, each red cell is in contact or near the impeller for only an extremely brief time period. Although the vortex generated by many centrifugal blood pumps can produce somewhat unpredictable velocity patterns within the housing, this does not appear to significantly affect blood pumped by such a device.

ISSUES RELATED TO CLINICAL DEVELOPMENT OF CENTRIFUGAL BLOOD PUMPS

The use of centrifugal blood pumps in a clinical setting has resulted in limited success. There are three major uses of artificial heart technology: an assist pump for temporary assistance of a recovering heart following open heart surgery, a bridge-to-transplant pump that is implanted for longer periods as a long-term assist until a donor heart is obtained for a transplant, and as a total heart replacement for a severely failing heart. As an assist pump, both centrifugal pumps and other blood pumping modalities have been quite successful, because their use is temporary (a few days) and the patient is typically hospitalized in a cardiac intensive care unit. The total heart replacement has been a difficult hurdle for any pumping system to produce, as the long-term issues of hemolysis, thrombosis, power availability, feedback and control, and other patient health issues have been difficult to successfully address. It is the middle ground, the bridge-to-transplant, where many blood pumping systems have been employed as a staging system to provide adequate cardiac support to the patient until a donor heart is available. Because this use may result in long-term application of the blood pump and usually requires implantation, the issues related to the total heart replacement are similar to those for the bridge-to-transplant versions.

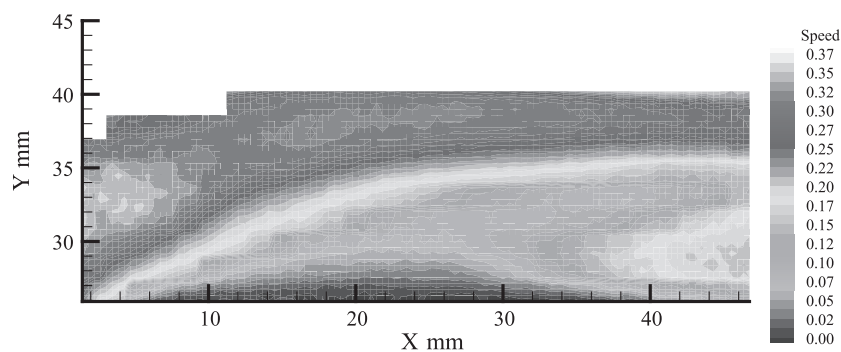


Fig. 18 Contour mapping of the flow field from a PIV flow visualization system, representing the average of 100 snapshot vector plots. The color scheme is shown at right with velocity values. The red colors represent higher velocities and blue colors represent lower velocities. (From Ref. [25].)

Thrombosis is still a significant issue in the long-term use of centrifugal blood pumps—in fact, with virtually all types of blood pumps. Once the thrombus is released due to vibrations or shear and becomes an embolus, then the chances for stroke and other vascular complications arise. As the generation of a thrombus is associated with stagnation zones within the blood velocity field, it is imperative to avoid stagnant zones within both the rotor and housing and the inlet and outlet cannula of the device. Feedback and control of the device is another area of considerable concern and interest. The development of a reliable long-lasting series of sensors to monitor blood pressure, flow rate, plasma-free hemoglobin, blood oxygen levels, and other associated physiological measures is of paramount importance in order to determine whether the device is functioning properly or to predict potentially dangerous situations before there are irreversible adverse effects on the patient.

Power consumption and the ability to renew and recharge battery sources transcutaneously is an issue that must be successfully addressed for centrifugal blood pumps as well as other pumping modalities. Thus, it is common to bench-test centrifugal blood pumps in order to monitor voltage and current needs in the pump as well as pump efficiency. It is in this arena where centrifugal pumps may have an advantage over other blood pumping modalities, as the long and

successful history of nonmedical applications of such pumps has resulted in devices that are highly efficient, utilize small motors, require small driving voltages, and draw small currents. In addition, these pumps do not need to push a large volume of fluid in contact with a piston or diaphragm, do not need to build pressure within a closed housing (with valves) such as occurs in isovolumetric contraction in the natural heart, and operate at rotation rates that are smaller than axial flow blood pumps. Nonetheless, the long-term viability of these or any blood pumps is dependent on the ability to continually power the pump *in vivo/in situ* and to continually recharge the power system in a reliable manner.

CONCLUSION

Centrifugal blood pumps have been widely utilized in numerous clinical settings, and various versions have been examined by numerous research groups worldwide. The development of centrifugal blood pumps represents an important marriage between industry, academia, and medical centers. It is perhaps the field that produces the greatest such collaboration. It also represents a strong collaboration among research

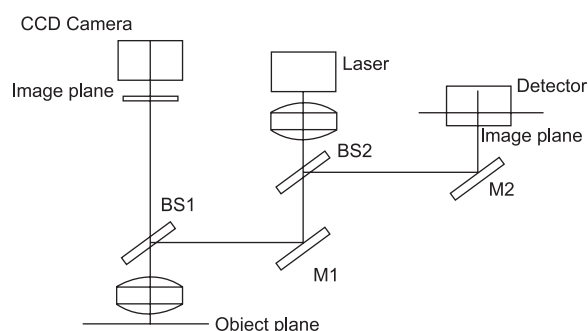


Fig. 19 Typical laser Doppler anemometer flow visualization system that uses backscattered laser light imaged by a CCD camera to view the velocity field.

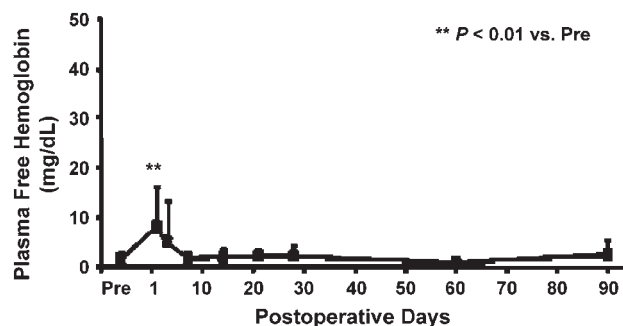


Fig. 20 Typical hemolysis curve of a centrifugal blood pump indicating levels of plasma-free hemoglobin (from blood samples) postimplantation. Hemolysis levels for centrifugal blood pumps are typically low and remain so for extended periods of use.

teams from many nations where the pumps are designed and bench-tested in one location and then evaluated clinically elsewhere. The International Society for Rotary Blood Pumps and its host journal *Artificial Organs* are ample evidence of such widespread collaboration. A multinational group dealing with rotary blood pumps meets annually to present their latest findings and demonstrate the latest pump designs.

The successful long-term use of centrifugal pumps in nonmedical settings has given rise to a number of advances that have been in evidence in the blood pumping versions developed over the last three decades. The advantage of centrifugal pumps is the large pressure gradient and the potential flow rates that can be generated and thus produce more than adequate support for cardiac assist or replacement over a wide range of physiological and pathological conditions. The very rapid response time between altered pump controls and the resulting changes in pressure and flow are another large advantage over other cardiovascular pump designs. The future is indeed promising for the further development and clinical testing of centrifugal blood pumps. However, as with any long-term implantable device to be used in the human body, there are issues to overcome before the term “artificial heart” becomes a reality.

ARTICLES OF FURTHER INTEREST

Artificial Heart Fluid Dynamics: Positive Displacement Pumps; Blood-Material Interactions; Blood Vessel Mechanics; Cardiac Assist Devices; Thrombosis

REFERENCES

1. Turton, R. *Rotodynamic Pump Design*; Cambridge University Press, 1994; 1–9.
2. Loree, H.M.; Bourque, K.; Gernes, D.B.; Richardson, J.S.; Poirier, V.L.; Barletta, N.; Fleischli, A.; Foiera, G.; Gempp, T.M.; Schoeb, R.; Litwak, K.N.; Akimoto, T.; Kameneva, M.; Watach, M.J.; Litwak, P. The HeartMate III: Design and in vivo studies of a maglev centrifugal left ventricular assist device. *Artif. Organs* **May 2001**, *25* (5), 386–391.
3. Maher, T.R.; Butlet, K.C.; Poirier, V.L.; Gernes, D.B. HeartMate left ventricular assist devices: A multi-generation of implanted blood pumps. *Artif. Organs* **May 2001**, *25* (5), 422–426.
4. Bourque, K.; Gernes, D.B.; Loree, H.M., II; Richardson, J.S.; Poirier, V.L.; Barletta, N.; Fleischli, A.; Foiera, G.; Gempp, T.M.; Schoeb, R.; Litwak, K.N.; Akimoto, T.; Watach, M.J.; Litwak, P. HeartMate III: Pump design for a centrifugal LVAD with a magnetically levitated rotor. *ASAIO J.* **Jul–Aug 2001**, *47* (4), 401–405.
5. Piacentino, V., III; Jones, J.; Fisher, C.A.; Singhal, A.K.; Macha, M.; McClurken, J.B.; Furukawa, S. Off-pump technique for insertion of a HeartMate Vented Electric left ventricular assist device. *J. Thorac. Cardiovasc. Surg.* **Jan 2004**, *127* (1), 262–264.
6. Naito, K.; Miyazoe, Y.; Aizawa, T.; Mizuguchi, K.; Tasai, K.; Ohara, Y.; Orime, Y.; Glueck, J.; Takatani, S.; Noon, G.P. Development of Baylor-Nikkiso centrifugal pump with a purging system for circulatory support. *Artif. Organs* **Jul 1993**, *17* (7), 614–618.
7. Ohtsubo, S.; Naito, K.; Matsuura, M.; Kawahito, K.; Shimono, T.; Makinouchi, K.; Tasai, K.; Ohara, Y.; Damm, G.; Glueck, J. Initial clinical experience with the Baylor-Nikkiso centrifugal pump. *Artif. Organs* **Jul 1995**, *19* (7), 769–773.
8. Ohtsubo, S.; Tayama, E.; Short, D.; Noon, G.P.; Nose, Y. Clinical comparative study of cardiopulmonary bypass with Nikkiso and BioMedicus centrifugal pumps. *Artif. Organs* **Jun 1996**, *20* (6), 715–720.
9. Coselli, J.S.; LeMaire, S.A.; Ledesma, D.F.; Ohtsubo, S.; Tayama, E.; Nose, Y. Initial experience with the Nikkiso centrifugal pump during thoracoabdominal aortic aneurysm repair. *J. Vasc. Surg.* **Feb 1998**, *27* (2), 378–383.
10. Ichikawa, S.; Nose, Y. Centrifugal blood pumps for various clinical needs. *Artif. Organs* **Nov 2002**, *26* (11), 916–918.
11. Ohara, Y.; Sakuma, I.; Makinouchi, K.; Damm, G.; Glueck, J.; Mizuguchi, K.; Naito, K.; Tasai, K.; Orime, Y.; Takatani, S.; et al. Baylor Gyro Pump: A completely seal-less centrifugal pump aiming for long-term circulatory support. *Artif. Organs* **Jul 1993**, *17* (7), 599–604.
12. Ohara, Y.; Makinouchi, K.; Orime, Y.; Tasai, K.; Naito, K.; Mizuguchi, K.; Shimono, T.; Damm, G.; Glueck, J.; Takatani, S.; et al. An ultimate, compact, seal-less centrifugal ventricular assist device: Baylor C-Gyro pump. *Artif. Organs* **Jan 1994**, *18* (1), 17–24.
13. Ohara, Y.; Nose, Y. The next generation Baylor C-Gyro Pump: Antithrombogenic “free impeller” design for long-term centrifugal VAD. *Artif. Organs* **Mar 1994**, *18* (3), 238–243.
14. Nakazawa, T.; Benkowski, R.; Makinouchi, K.; Takami, Y.; Ohtsubo, S.; Glueck, J.; Kawahito, K.; Sueoka, A.; Schmalleger, H.; Schima, H.; Wolner, E.; Nose, Y. Recent advances in the gyro centrifugal ventricular assist device. *ASAIO J.* **Jan–Feb 1998**, *44* (1), 94–97.
15. Yoshikawa, M.; Nakata, K.; Ohtsuka, G.; Takano, T.; Glueck, J.; Fujisawa, A.; Makinouchi, K.; Yokokawa, M.; Nose, Y. Feasibility of a tiny Gyro centrifugal pump as an implantable ventricular assist device. *Artif. Organs* **Aug 1999**, *23* (8), 774–779.
16. Takano, T.; Nakata, K.; Yoshikawa, M.; Maeda, T.; Glueck, J.; Fujisawa, A.; Makinouchi, K.; Yokokawa, M.; Murabayashi, S.; Nose, Y. Development of an antithrombogenic and antitraumatic blood pump: The Gyro C1E3. *ASAIO J.* **Jan–Feb 2000**, *46* (1), 123–127.
17. Nonaka, K.; Linneweber, J.; Ichikawa, S.; Yoshikawa, M.; Kawahito, S.; Mikami, M.; Motomura, T.; Ishitoya, H.; Nishimura, I.; Oestmann, D.; Glueck, J.;



- Schima, H.; Wolner, E.; Shinohara, T.; Nose, Y. Development of the Baylor Gyro permanently implantable centrifugal blood pump as a biventricular assist device. *Artif. Organs* **Sep 2001**, 25 (9), 675–682.
18. Takano, T.; Schulte-Eistrup, S.; Kawahito, S.; Maeda, T.; Nonaka, K.; Linneweber, J.; Glueck, J.; Fujisawa, A.; Makinouchi, K.; Yokokawa, M.; Nose, Y. Inlet port positioning for a miniaturized centrifugal blood pump. *Artif. Organs* **Jan 2002**, 26 (1), 45–48.
 19. Nishimura, I.; Ichikawa, S.; Mikami, M.; Ishitoya, H.; Motomura, T.; Kawamura, M.; Linneweber, J.; Glueck, J.; Shinohara, T.; Nose, Y.; Nishimura, I. Evaluation of floating impeller phenomena in a gyro centrifugal pump. *ASAIO J.* **Nov–Dec 2003**, 49 (6), 744–747.
 20. Miller, G.E.; Etter, B.D.; Dorsi, J.M. A multiple disk centrifugal pump as a blood flow device. *IEEE Trans. Biomed. Eng.* **Feb 1990**, 37 (2), 157–163.
 21. Miller, G.E.; Sidhu, A.; Fink, R.; Etter, B.D. Evaluation of a multiple disk centrifugal pump as an artificial ventricle. *Artif. Organs* **Jul 1993**, 17 (7), 590–592.
 22. Miller, G.E.; Madigan, M.; Fink, R. A preliminary flow visualization study in a multiple disk centrifugal artificial ventricle. *Artif. Organs* **Jul 1995**, 19 (7), 680–684.
 23. Manning, K.B.; Miller, G.E. Shaft/shaft-seal interface characteristics of a multiple disk centrifugal blood pump. *Artif. Organs* **Jun 1999**, 23 (6), 552–558.
 24. Miller, G.E.; Fink, R. Analysis of optimal design configurations for a multiple disk centrifugal blood pump. *Artif. Organs* **Jun 1999**, 23 (6), 559–565.
 25. Manning, K.B.; Miller, G.E. Flow through an outlet cannula of a rotary ventricular assist device. *Artif. Organs* **Aug 2002**, 26 (8), 714–723.
 26. Rice, W. An analytical and experimental investigation of multiple disk pumps and compressors. *J. Eng. Power* **1963**, 85, 191–200.
 27. Brieter, M.; Pohlhausen, K. Laminar Flow Between Two Parallel Rotating Disks. In *Aeronautical Research Laboratory Report, WPAFB*; U.S. Air Force, 1962. Project No. 7071, Task No. 70437.
 28. Hassinger, S.H.; Kehrt, L.G. Investigation of a shear-force pump. *J. Eng. Power* **July 1963**, 85, 201–207.
 29. Tesla, N. Turbine. US Patent No. 1,061,142, May, 1913.
 30. Noon, G.P.; Ball, J.W., Jr.; Papaconstantinou, H.T. Clinical experience with BioMedicus centrifugal ventricular support in 172 patients. *Artif. Organs* **Jul 1995**, 19 (7), 756–760.
 31. Noon, G.P.; Lafuente, J.A.; Irwin, S. Acute and temporary ventricular support with BioMedicus centrifugal pump. *Ann. Thorac. Surg.* **Aug 1999**, 68 (2), 650–654.
 32. Reiss, N.; El-Banayosy, A.; Mirow, N.; Minami, K.; Korfer, R. Implantation of the Biomedicus centrifugal pump in post-transplant right heart failure. *J. Cardiovasc. Surg.* **Oct 2000**, 41 (5), 691–694.
 33. Ochiai, Y.; Golding, L.A.; Massiello, A.L.; Medvedev, A.L.; Gerhart, R.L.; Chen, J.F.; Takagai, M.; Fukamachi, K. In vivo hemodynamic performance of the Cleveland Clinic CorAide blood pump in calves. *Ann. Thorac. Surg.* **Sep 2001**, 72 (3), 747–752.
 34. Ochiai, Y.; Golding, L.A.; Massiello, A.L.; Medvedev, A.L.; Horvath, D.J.; Gerhart, R.L.; Chen, J.F.; Krogulecki, A.Y.; Takagaki, M.; Doi, K.; Howard, M.W.; Fukamachi, K. Cleveland Clinic CorAide blood pump circulatory support without anticoagulation. *ASAIO J.* **May–Jun 2002**, 48 (3), 249–252.
 35. Fukamachi, K.; Ochiai, Y.; Doi, K.; Massiello, A.L.; Medvedev, A.L.; Horvath, D.J.; Gerhart, R.L.; Chen, J.F.; Krogulecki, A.Y.; Takagaki, M.; Howard, M.W.; Kopcak, M.W., Jr.; Golding, L.A. Chronic evaluation of the Cleveland Clinic CorAide left ventricular assist system in calves. *Artif. Organs* **Jun 2002**, 26 (6), 529–533.

Ceramics

Robert H. Doremus

Department of Material Science and Engineering, Rensselaer Polytechnic Institute, Troy, New York, U.S.A.



INTRODUCTION

Ceramic materials have been used in human activities for many millennia. In recent years they have found application as replacement or repair materials in the human body, starting with the use of porcelain in restorative dentistry.^[1] Compared with other materials used in the body, such as polymers and metals, ceramics can have excellent biocompatibility, but suffer from lack of impact resistance in applications where high stress is applied to the implant, such as in hip replacements. Nevertheless, ceramics are used in prosthetic and biomedical applications where impact resistance is not so important, as listed in Table 1. Materials used in the body must be compatible with physiological materials and functions. In Table 2 some desirable properties of materials in the body are listed.

Processing of crystalline ceramics and glass ceramics is briefly described in the next section. Then, there is a section on fracture and fatigue of ceramics. Mechanical stability is often essential in the use of ceramics as implant materials. Ceramics have been used as implants without a thorough understanding of their mechanical properties, and the result has often been failure of the material and its rejection as an implant material. The following section is devoted to a discussion of hydroxylapatite (HA) as an implant material. It is an attractive material for implants, because it is completely biocompatible and bonds strongly to bones. Bones and teeth are composed of HA with some carbonate, so the bone cells recognize synthetic HA as a compatible material. Other ceramic materials with potential as implants are described in the section on other ceramics.

Composite materials can enhance the properties of the matrix or embedded material, and have attractive potential as implants. To take advantage of the chemical compatibility of ceramics with body fluids, and the impact resistance of metals, ceramics are coated onto metal substrates; these combined materials are described in a section on composites and one on coatings. Many natural materials are combinations of crystalline ceramic particles in a polymeric matrix. An example is the shell of a clam, which contains calcium carbonate particles in a protein matrix, and is very strong and durable. Synthetic materials are made that try to copy some of these structures. This approach is discussed in

the biomimetics section. Some of the present review is taken directly from Ref. [2]. There are also other reviews in Refs. [3–5].

PROCESSING OF CERAMICS

Crystalline ceramics are usually made in three steps: forming of a powder to the desired shape, partial drying, and firing at high temperatures to produce dense material.^[6,7] The firing step can include a hold at intermediate temperature to burn out organic binders. In slip casting the powder in a slurry is poured into a mold that absorbs excess water; the resulting ware can be fired directly without further drying.

In traditional ceramics such as porcelain, powdered raw materials such as clay (hydrated aluminosilicates), feldspar (alkali aluminosilicates), and silica sand are used directly as mined. For specialty uses, such as in electronics and most bioceramics, properties are strongly influenced by impurities and particle size of the starting materials, which must therefore be carefully controlled. Precipitation from solution and pyrolysis of organic precursors, including polymers or gels, are methods of producing pure powders.

Powders are usually pressed damp in metal dies or dry in lubricated dies at high enough pressure to form material strong enough to hold together until it is sintered. An organic binder such as polyvinyl alcohol helps to bind the powder together. Drying at about 100°C is a critical step in preparing damp-formed pieces for firing. Too much or too little water in pressed powder can lead to blowing apart the ware on heating, or to crumbling, respectively. Binder is removed by heating to a higher temperature in air to oxidize the organic binder to carbon dioxide and water.

When a piece is fired, the powder particles join together to form a solid, porous mass; as firing continues, most of the pores are removed, forming a compact, dense, polycrystalline solid. Sintering together of the particles can occur entirely in the solid state or from a mixture of solid particle and liquid. A liquid phase aids sintering; a small amount of liquid coats grain boundaries in the dense solid, and a larger amount can provide a glassy matrix for solid particles, as in porcelain, and is called vitrification.

Table 1 Current and proposed applications of bioceramics

Dental restorations: crown, caps, dentures, fillings
Tooth root replacements
Rebuilding the alveolar ridge
Whole-tooth replacements
Reconstructive middle ear surgery
Fracture healing
Spine surgery
Rebuilding of facial and cranial bones
Filling mastoid defects
Filling bony defects, for example, in tibial plateau fractures
Female sterilization

Materials that are difficult to sinter to dense bodies can sometimes be sintered by a combination of heat and pressure called hot pressing. The main problem in hot pressing is to find materials for dies that will withstand pressure of 30 MPa or higher above 1000°C. Graphite is widely used, but it is quite expensive and has a short life as a die, so hot pressing is used only for pieces of high value. Hot pressing is also slow because of the time required to heat and cool the dies for processing each piece.

FRACTURE AND FATIGUE OF CERAMICS

The mechanical failure of brittle materials such as ceramics results from propagation of cracks from flaws in the material, either on its surface or in the bulk volume. When a tensile stress, S , is applied to a thin, deep flaw of depth c with a flaw tip radius of p , the stress S_t at the crack tip is:^[8]

$$S_t = 2(c/p)^{1/2}S \quad (1)$$

Since the tip radius is usually several orders of magnitude smaller than the crack depth, the applied stress is greatly multiplied at the crack tip. Thus, brittle materials fail at applied tensile stresses much lower

Table 2 Desirable properties of implants

No toxicity
Mechanically stable: do not fracture or deform excessively
No foreign body reaction
No stimulation of undesirable growths or chemical reactions in body fluids and tissues
No decomposition
No resorption, or at a desired rate

than the theoretical stress required to break the chemical bonds in the solid material.

The strengths of ceramics are usually measured in bend (flexure) tests of rod or bar samples, with three or four contact points on the specimen. Fibers are usually tested in tension. A convenient method for measuring strengths of ceramics is the diametral test.^[9,10] In this test a right circular cylinder, a few millimeters in diameter and in thickness, is compressed along the diameter of the cylinder between flat plates. Pads of ductile material (e.g., copper) are inserted between the sample and the plates to prevent crushing of the sample at the diameter ends. A tensile stress is developed normal to the loading diameter along both surfaces of the sample and also across a plane through the sample from one surface to another. Thus, both surface and volume flaws can contribute to sample failure. Samples for the diametral test can be cut from a long cylindrical sample, and are therefore easy to make in large quantities.

Vardar and Finnie^[9] tested rock samples and showed that the failure strengths from the diametral test were a factor of about 1.2 smaller than failure strengths from bend tests. Thus, strengths from diametral tests should be multiplied by 1.2 to compare them with bend strengths.

When the failure strengths of a number of identical brittle samples are measured under the same conditions of sample preparation and loading, the strengths vary over a wide range, for example, by a factor of 2. This distribution of strengths can be expressed statistically. A distribution function often used for strengths is the Weibull function.^[11]

$$F = 1 - \exp[-S/S_0]^m \quad (2)$$

in which F is the fraction of samples that breaks below stress S , and m and S_0 are constants. The Weibull modulus, m , is a measure of the spread and the skewness of the distribution. It is desirable to have a method of measuring the bonding between bone and an implant. A widely used way for semiquantitatively measuring the bone-implant bonding is the pushout test. In this test a cylindrical sample is implanted, for example, into the cortex of a femur of an animal, and after the bone plus the implant is recovered the force needed to push the implant out of the bone is measured.

A more quantitative test for measuring the strength of bone-implant adhesion has been described by Nakamura, Yamamuro, Higushi, Kukubi, Ito.^[12] A plate ($2 \times 10 \times 15 \text{ mm}^3$) was implanted into a hole in a rabbit tibia such that the plate extended right through the tibia from the medial to the lateral cortex. After sacrifice of the animal, at the end of the experiment, the bone was cut so that two semicircular bone segments were each bonded to the plate on either side. Each bone segment was held by a hook connected to an Instron test

machine. In this way the bonding strength of the bone to the implant could be measured directly. A segment of bone without an implant was tested in the same way for comparison.

When oxides are held under a tensile stress in aqueous solutions or water vapor, they become weaker. This fatigue can be tested in several ways.^[13] In a static fatigue test the samples are held under constant load until they fail. In metals, fatigue tests are usually conducted with a variable (cyclical) applied stress, because this mode leads to more rapid failures than a static load. However, in ceramics the integrated load time function gives the same fatigue times as a cyclical stress, so there is no special deterioration of the strength for cyclical loading.

To make predictions of the time to failure for a brittle material, one must choose a suitable stress–failure time function. The function most often used is a power law:

$$\log t/\tau = n \log(S/S_N) \quad (3)$$

where t is the time to failure, S is the applied stress, and τ , n , and S_N are constants. S_N is the failure strength under conditions where there is no fatigue. The parameter n is a measure of the fatigue susceptibility of the material. A large (>50) value of n means a material has little fatigue sensitivity (large times to failure); a low (>10) value of n means the material is highly sensitive to fatigue (short time to failure). Tau (τ) is a scaling parameter for the time of failure.

In silicate glasses a function that fits the distribution of $\log t/\tau$ values better than Eq. (3) is:^[13]

$$\log(t/\tau) = g(S_N/S) \quad (4)$$

in which g is a constant. This function can be extrapolated to long fatigue times more reliably than Eq. (3); neither Eq. (3) nor Eq. (4) shows a fatigue limit, which is an applied stress below which no failure occurs.

A brittle material such as a ceramic or glass deforms elastically until it fractures; thus, if the applied stress is plotted as a function of the relative deformation or strain, a straight line results. A ductile material such as aluminum deforms linearly at very low stresses, but at higher stresses plastic deformation results and the stress–strain curve becomes nonlinear; more and more stress is required to increase the strain. Toughness can be defined as the ability of a material to absorb energy during plastic deformation, and an indication of toughness can be obtained from the area under the stress–strain curve for a static test, which represents the work required to fracture the material.

A quantity called the critical stress intensity factor K_c can be measured for a solid material. If there is a crack

or notch of depth c in a material which is subjected to a tensile stress (mode I loading), K_c is defined by

$$K_c = S(\pi c)^{1/2} \quad (5)$$

in which S is the stress required to cause the crack or notch to propagate to failure. K_c is sometimes considered as a material or intrinsic property of a material, but many experiments show that K_c depends on extraneous factors such as sample microstructure and treatment history. Thus, K_c is not an intrinsic property of a material. The critical stress factor, K_c , is sometimes called the toughness of a brittle material, a term which is highly misleading because brittle materials are not “tough” (have good impact resistance). The term toughness should be limited to descriptions of ductile materials, as described in the preceding paragraph.



HYDROXYLAPATITE

The mineral phases in mammalian bones and teeth are HA, $\text{Ca}_{10}(\text{PO}_4)_6(\text{OH})_2$, containing some additional carbonate. Growing bone attaches directly to HA implants by chemical bonding; there is no foreign body reaction, fibrous layers, or other chemical changes at the bone–HA interface.^[14–16] Thus, HA is an ideal implant material from a chemical and bone-bonding perspective. HA, of course is brittle, so it is not suitable for application which requires impact resistance, such as hip socket replacements.

Dense HA containing no pores or second phases was made by pressureless sintering of fine, precipitated crystals of HA in a damp cake at 1100°C for 1 hr.^[17] To make samples without pores or second phases the stoichiometry must be exactly that of HA, that is (Ca:P ratio of 5:3). Without pores or second phases the dense HA is translucent to transparent; the pores and second phases scatter light to make HA samples opaque.

Translucent HA with no pores or second phases has high strength (up to 200 MPa in diametral tests) and excellent fatigue resistance.^[10,17] HA samples of lower density, containing pores and second phases, have lower strength and fatigue resistance. A detailed examination of some previous results on inferior quality samples of HA shows that their reduced strength and fatigue resistance results because of pores, second phases, and impurities in these materials.^[2] When high-purity HA is properly made so it has no pores or second phases, it has strength and fatigue properties that permit many biomedical implant applications.

Coatings of HA on metals and composites of HA with other ceramics are being explored to overcome the lack of impact resistance of HA and are discussed in the composites and coatings section. One can conclude that HA is the most attractive ceramic material for implants,

especially because of its excellent bonding to bones and teeth.

OTHER CERAMIC MATERIALS AS IMPLANTS

Dense, pore-free β -tri-calcium phosphate $\text{Ca}_3(\text{PO}_4)_2$ can be made by sintering precipitated fine crystals in much the same way as HA was made.^[18] The resulting material is strong, and bone bonds well to it without any foreign body reactions. A possible advantage of tri-calcium phosphate over HA as an implant material is that it is slowly reabsorbed by body fluids, because it is more soluble in aqueous solutions than apatite. There is more background information on calcium phosphates in Ref. [19].

Jade (base composition $\text{Na}_2\text{O} \cdot \text{Al}_2\text{O}_3 \cdot 4\text{SiO}_2$) was found implanted in teeth of Mayan women, of the period 600–800 A.D., perhaps for ornamental and restorative purposes.^[20] Natural jade is a quite hard, durable crystalline mineral of ceramic.

Porcelain has been used for restorative dentistry since the late 18th century. (See Ref. [1] for a review of materials in dentistry.) Porcelain inlays and crowns were made by Charles H. Land in the late 19th century by fusing the porcelain on to platinum foil in the form of a model of the tooth. Porcelain can be made with optical properties that match teeth with additions of either silica or alumina particles; such particles also increase the strength of the porcelain. Porcelain crowns can also be made strong by fusing the porcelain to a metal coping; however, the metal changes the appearance of the crown. Porcelain restorations can also be cast in a mold.

Porcelains contain various mixtures of clay (e.g., kaolin, $\text{Al}_2\text{O}_3 \cdot 2\text{SiO}_2 \cdot 2\text{H}_2\text{O}$), feldspar ($\text{K}_2\text{O} \cdot \text{Al}_2\text{O}_3 \cdot 6\text{SiO}_2$), and quartz sand. A typical hard porcelain for whiteware contains 50% clay, 25% feldspar, and 25% sand. Dental porcelains have more feldspar and little or no sand. A typical composition is 95% feldspar, 5% clay. The result is a highly glassy material that is translucent and melts at a lower temperature because of the high feldspar content. During firing of porcelain the feldspar melts first to provide a molten matrix that speeds consolidation of the powders and gives a glassy matrix on cooling. The clay dehydrates and transforms to fine, elongated crystals of mullite, $3\text{Al}_2\text{O}_3 \cdot 2\text{SiO}_2$, that provide additional strength to the resulting material. Any quartz or alumina particles partially dissolve, and the residual crystals increase strength but reduce translucency.

Material properties of porcelains are not sensitive to small changes in composition, forming methods and firing temperatures. As the firing temperature or amount of liquid (feldspar) increases, more silica and alumina dissolve in the feldspar melt, increasing the

viscosity and lowering the reactivity of the molten feldspar. Porcelains, therefore, have a glassy matrix, and chemical properties that resemble those of a potassium aluminosilicate glass; they are strong, chemically durable, and can be fired at relatively low temperatures and short times. These properties have made porcelains a satisfactory restorative material in dentistry, with the particular advantage of optical matching teeth, but the disadvantage of brittleness.

Aluminum oxide is very hard (9 on the Mohr scale), stable to high temperatures, and chemically durable. Polycrystalline sintered alumina is very strong compared to most ceramics, although it is brittle. Its strength, abrasion resistance, and chemical inertness have made it attractive for dental and bone implants.

Porous calcium aluminate was made by sintering fine alumina powder with coarse calcium carbonate at 1450°C for 24 hr.^[21] Pore sizes from 20 to 100 μm resulted by using different sizes of the calcium carbonate particles; the pore size was about the same as the particle size. The carbonate decomposed to calcium oxide, which then reacted with alumina to form two calcium aluminate phases, $\text{CaO} \cdot \text{Al}_2\text{O}_3$ and $3\text{CaO} \cdot 5\text{Al}_2\text{O}_3$. With a starting composition of about 27% CaO and 73% Al_2O_3 (by weight), the amounts of the two phases were about those expected at equilibrium (20% $\text{CaO} \cdot \text{Al}_2\text{O}_3$, 80% $\text{CaO} \cdot 5\text{Al}_2\text{O}_3$). Porous $\text{CaO} \cdot \text{TiO}_2$ and $\text{CaO} \cdot \text{ZrO}_2$ were also made by reaction of calcium carbonate with TiO_2 or ZrO_2 . One reason for porous implants is to provide a place for bone to grow into, to hold the implant mechanically. However, porosity weakens the solid implant, and the spaces in the pores not filled by bone invite bacterial infiltration and infection.

Glassy carbon is produced by pyrolysing cross-linked organic polymers such as cellulose, phenylformaldehyde, and polyfurfuryl alcohol under carefully controlled conditions.^[22] The microstructure consists of random crystallites about 5 nm in size, in contrast to the ordered layer structure of graphite. The density of glassy carbon (1.4–1.5 gm/cm^3) is lower than that of graphite, but glassy carbon has low porosity and permeability. The mechanical properties are similar to those of silicate glasses; the strength and modulus are higher than those of graphite. Glassy carbon is chemically inert compared to graphite, probably because of its low porosity. Glassy carbon is quite expensive to manufacture.

There was a flurry of interest in glassy carbon implants in dentistry in the mid-1970s.^[5] Bone does not bond to glassy carbon, so the implants and tooth replacements were made in various shapes, such as blades, with holes for bone to grow into to anchor the implant. Some of these implants were initially successful, but fractured or loosened after a few months to 1 or 2 yr after implantation. The fatigue resistance of glassy carbon should be high.

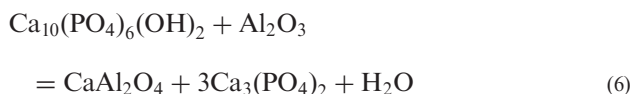
Pyrolytic carbon coatings have been used in a variety of implant applications, such as heart valves and artificial arteries.^[5] Artificial joints and parts of heart valves have been made of solid carbon. Pyrolytic carbon has good fatigue resistance.

COMPOSITES

Combinations of ceramics with either other ceramics or other materials can lead to improved properties and thus to more applications. Most work has been done on composites with HA because of its excellent biocompatibility, so this review is devoted to composites with HA, either as a matrix or as a filler.

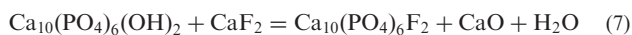
Reinforcing materials have been incorporated into HA to improve its mechanical properties. Metal and ceramic wires have been hot-pressed in an HA matrix at 1000°C and 1 kbar for 15 min;^[23] in such processes, it is necessary to match the coefficients of thermal expansion (CTE) closely to make dense and crack-free composites. The most successful alloys for wires were Hastalloy (51% Ni, 22% Cr, 9% Mo, 18% Fe) of 8 μm diameter and FeCrally (79% Fe, 16% Cr, and 5% Al) of 20 μm diameter; the oxide reaction layers on the wire surfaces gave good bonding to HA. FeCrally composites containing 20% and 30% wire volume were stronger than HA.

Alumina is hard and strong, so Al₂O₃ powder has been sintered with HA at temperatures of 1100–1400°C to form dense composites.^[24–29] One problem with such constructs was that the alumina catalyzed the decomposition of HA by the formation of calcium aluminates (CaAl₂O₄, CaAl₄O₇), for example, in the reaction



The water formed in this reaction can lead to porosity in the sintered composite and thus weaken it.

Addition of 5 wt% calcium fluoride to the sintering mixture reduces the porosity of the alumina–HA composites and increases their strengths.^[24–29] Partial substitution of hydroxyl groups in apatite by fluoride at temperatures of about 800°C reduces decomposition of apatite by Eq. (6) and leads to lower porosity in the apatite because less water is produced. In this case, the pertinent reaction is



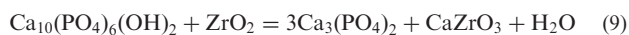
In the alumina–HA composites the CaO reacts with Al₂O₃ to form CaAl₂O₄:



Strengths of alumina–HA composites above the 200 Mpa strength of pure apatite were reported for these composites with low porosity.^[24,27,28] The fracture toughness of alumina–HA composites were also reported to be higher than the value of 1 Mpa√m for pure, dense HA.^[24,26,27]

Because of its high mechanical strength and chemical inertness in the human body, zirconia (ZrO₂) is also an attractive additive for zirconia–HA composites. These composites can be produced in two different ways. In the first, apatite and zirconia powders are mixed, cold-pressed into a desired shape, and sintered at temperatures from 1100°C to 1300°C.^[24,30–33] In the second method, the apatite and zirconia are precipitated from aqueous solution by mixing solutions of Ca(NO₃)₂, (NH₄)₃PO₄, and ZrOCl₂^[24,34,35] to form mixtures of HA and ZrO₂ crystals, which are then sintered.

Zirconia also increases the tendency for decomposition of the HA, as does alumina. In this case, a possible reaction is:^[24,32,33]



Again, the formation of water increases the porosity of the sintered composite and reduces its strength. Addition of 5 wt% CaF₂ to the mixture substitutes fluorine for OH at about 800°C by the reaction of Eq. (7) and reduces the formation of water and the resultant porosity.^[24,33] Diametral tensile strengths of greater than 300 MPa are possible for HA–zirconia composites containing 40 wt% ZrO₂.^[24] Higher fracture toughness was also reported for HA–zirconia composites as compared to pure HA.^[24,33,35]

Mixtures of HA granules in a variety of matrices have also been used in various implant applications. HA in plaster of Paris (calcium sulfate or gypsum, CaSO₄·2H₂O) has given good results in alveolar ridge augmentation and in the treatment of periodontal defects.^[4] This mixture can be shaped to fit a variety of defects and sterilized without changing its form or properties. After implantation, the plaster of Paris is gradually replaced by fibro-vascular tissue, and the HA particles remain firmly in place. A more complex mixture containing hydroxylapatite, corticocancellous cranial bone, Bacitracan powder, and thrombin was implanted into a variety of bony defects.^[4] The autogenous bone was presumed to stimulate subsequent bone growth; however, the osteocytes originally in this bone did not survive the transplantation surgical procedure, so new cells must grow to provide new bone growth.

COATINGS

To take advantage of the excellent biocompatibility and bonding of HA to bone, the HA can be coated on



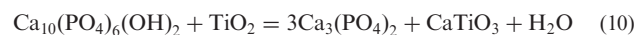
a metallic substrate to make a composite material with good impact resistance.^[36,37] The HA coating also prevents the metal from reacting with body fluids, and migration of metal reaction products into neighboring tissues.

The most commonly used alloys for implants are based on titanium, cobalt–chromium, and stainless steel (iron, nickel, and chromium). They have good resistance to body fluids and reasonable strength, but they all induce fibrous encapsulation and some degree of foreign body reactions.

Many different methods of coating HA onto metals have been used including plasma spraying,^[36–39] sputtering,^[40,41] polymer-assisted^[42] dipping,^[43] sol–gel processing,^[44] electrophoretic deposition,^[45] and hot isostatic pressing.^[46] However, none of these methods has led to a satisfactory coating because such coatings contained cracks, pores, second phases, and residual stresses that reduced their durability and led to partial or complete disintegration of the coating in body fluids.^[2,47]

In plasma spraying, the most used coating method, a fine (<50 μm) ceramic powder is blown through a plasma torch flame onto a metallic substrate. The plasma is very hot, but if conditions of flow are regulated, the powder can have a temperature much lower than the plasma. In ambient air, HA begins to decompose above about 1200°C; this reaction can be suppressed by adding water to the powder stream. Some plasma-sprayed deposits of HA show evidence of melting before deposition, which means that the HA powder reached a temperature above its melting temperature of about 1500°C.

In a detailed study of bonding of HA to titanium in several different kinds of experiments, it was shown that the HA bonds strongly and chemically to titanium^[48] to form perovskite, CaTiO₃, at the HA–titanium interface by the following reaction:



The TiO₂ is formed on the titanium surface either by oxidating in air or by residual oxygen or water. Since the HA and titanium react and bond strongly, it was concluded that cracks and weakness at the HA–titanium interfaces result from stresses generated by the mismatch between the CTE of these materials.^[48] In a study of bonding of HA composites to different alloys,^[24] it was found that when the CTE of the HA composite and the alloy were close (within about ±5%), the HA and alloy bonded tightly without cracks at the HA–alloy interface. For example, composites consisting of 25–40 wt% Al₂O₃, HA, and 5 wt% CaF₂ had CTE close to that of an alloy of titanium containing 6% Al and 4% V; the HA composite formed thick coatings on this alloy with good interfacial bonding.

Thus, the key to making successful coatings of HA on metals and alloys is to match their CTE closely.

BIOMIMETICS

Many biological structures have remarkable mechanical properties, in addition to the chemical properties required in physiology. Examples are bone, many sea-shells, and sponge spicule. Some of these structures have high strength and remarkable impact resistance. They are made up of combinations of an organic matrix and inorganic crystals. For example, bone consists of fine plate-like HA crystals (with dimensions of about 20 nm) bonded in a matrix of the fibrous protein collagen. Abalone has aragonite (CaCO₃) platelets in an organic film that grows to become nacre (mother of pearl).

The unusual properties of these structures have led to the idea of making similar structures synthetically. Since this process copies nature, it has been given the name biomimetics. There is a review of many aspects of biomimetics in the new journal *Nature Materials* with 97 references.^[49] For the present discussion, I will select a few examples from this review; the reader is referred to the review article for more details and references.

The goal of biomimetics research is to construct structures of various interesting inorganic materials in a protein matrix. The protein bonds to the inorganic solids; the main difficulty is to assemble these micro-composites into larger-scale structures that mimic natural structures (such as bone and nacre) that have a complex hierarchy of architectures from the initial scale (a few tens of nanometers) to macroscopic structures (millimeters). The first step in this process is to find proteins and polypeptides that bind to the nanoinorganic crystals and then to assemble them into macrostructures.

Proteins are large complex organic “molecules” vital to life. They come in a wide variety of sizes, shapes, and chemical constituents. Examples are the fibrous collagen that forms the bone matrix, and globular albumin in blood. Proteins are made up of amino acids, which are molecules containing carboxyl (COOH) and amine (NH₂) groups. Amino acids form chains of molecules called peptides or polypeptides, which are then folded and bonded into the complex structures of proteins. Proteins bond to solid surfaces (and other molecules) as determined by the specific organization of the carboxylate and amine groups in them; proteins also have recognition mechanisms related to their size and shape. Thus, the challenge of biomimetics is to find or synthesize polypeptides and proteins that bond as specifically as possible to inorganic materials.

Some particular inorganic materials that are being studied in biomimetic architectures are fine gold

particles,^[50] “quantum dots” (small particles) of GaInAs,^[51] and carbon nanotubes.^[52] Bonding of proteins to a variety of other inorganic solids^[49] such as silicates, ZnO, CaCO₃, Fe₂O₃, and ZnS is also possible.

Some other biomimetic applications involve proteins that copy unusual behavior in living systems. One possibility is proteins that serve as actuators in micro- and nanodevices.^[53] Another application is a contractile protein from broad bean as a driver for a micro-motor.^[54] Much more work is required to copy the architectures and functions of proteins and protein-inorganic combinations in nature.

PROSPECTS

Ceramics are being developed for a variety of implant applications, with HA, especially in composites with other ceramics and as coatings on metals, being the most promising. Architectures that copy natural combinations of proteins and inorganic materials (biomimetics) also are highly attractive for applications.

ARTICLES OF FURTHER INTEREST

Biologic and Synthetic Apatites; Biphasic Calcium Phosphate (BCP) Bioceramics: Chemical, Physical, and Biological Properties; Zirconia Ceramics

REFERENCES

1. Brackett, S.E.; Duncanson, M.G. *Biomedical Materials and Devices*; Materials Research Society: Pittsburgh, PA, 1989; 579–588.
2. Doremus, R.H. Bioceramics. *J. Mater. Sci.* **1992**, *27*, 285–297.
3. Hench, L.L. Bioceramics: from concept to clinic. *J. Am. Ceramic Soc.* **1991**, *74*, 1487–1510.
4. Hanker, J.S.; Giannakopoulos, B.L.; Eds. *Biomedical Materials and Devices*; Materials Res. Soc.: Pittsburgh, PA, 1989.
5. Ducheyne, P.; Hastings, G.W.; Eds. *Metal and Ceramic Biomaterials*, Vols. I and II. CRC Press: Boca Raton, FL, 1984.
6. Kingery, W.D. *Introduction to Ceramics*; CRC Press: Wiley, New York, 1960.
7. Rahaman, M.N. *Ceramic Processing and Sintering*; Marcel Dekker: New York, 2003.
8. Inglis, C.E. Stresses in a plate due to the presence of cracks and sharp corners. *Trans. Inst. Naval Arch.* **1913**, *55*, 219–230.
9. Vardar, O.; Finnie, J. An analysis of the brazilian disk fracture test using weibull probabilistic treatment of brittle strength. *Int. J. Fracture* **1974**, *11*, 495–508.
10. Thomas, M.B.; Doremus, R.H.; Jarcho, J.; Salsbury, R.L. Dense hydroxylapatite: fatigue and fracture strength after various treatments from diametral tests. *J. Mater. Sci.* **1980**, *16*, 891–896.
11. Doremus, R.H. Fracture statistics: A comparison of the normal, Weibull, and type I extreme value distributions. *J. Appl. Phys.* **1983**, *54*, 193–200.
12. Nakamura, T.; Yamamuro, T.; Higushi, S.; Kukubi, T.; Ito, J. A new glass-ceramic, apatite and wollastonite containing, for bone replacement. *J. Biomed. Mater.* **1985**, *19*, 685–698.
13. Doremus, R.H. *Glass Science*; Wiley, New York, 1994, Chpt. 10.
14. Jarcho, M.; Kay, J.F.; Gumaer, K.J.; Doremus, R.H.; Drobeck, H.P. Tissue, cellular and subcellular events at a bone-ceramic hydroxylapatite interface. *J. Bioengineering* **1977**, *1*, 79–92.
15. Tracy, B.M.; Doremus, R.H. Direct electron microscopy studies of the bone hydroxyapatite interface. *J. Biomed. Matls. Res.* **1984**, *18*, 719–726.
16. DeLange, G.L.; EdPutter, C.; DeWijns, F.L.J.A. Histological and ultrastructural appearance of the hydroxyapatite-bone interface. *J. Biomed. Matls. Res.* **1990**, *24*, 829–845.
17. Jarcho, M.; Bolen, C.H.; Thomas, M.B.; Bobick, J.; Kay, J.F.; Doremus, R.H. Hydroxylapatite synthesis and characterization in dense polycrystalline form. *J. Mats. Sci.* **1976**, *11*, 2027–2035.
18. Jarcho, M.; Salsbury, R.L.; Thomas, M.B.; Doremus, R.H. Synthesis and fabrication of β -tricalcium phosphate ceramics for potential prosthetic applications. *J. Matls. Sci.* **1979**, *14*, 142–151.
19. DeGroot, K., Ed. *Bioceramics of Calcium Phosphate*; CRC press: Boca Raton, FL, 1983.
20. Epstein, N. Smithsonian, December 1989, 98–102.
21. Hulbert, S.F.; Klawitter, J.J.; Leonard, R.B. *Ceramics in Severe Environments*; Kriegel, W.W., Palmour, N., Eds.; Plenum Press, New York, 1971; 417–421.
22. *Encyclopedia of Chemical Technology*; Vol. 4, Wiley, N.Y., 1978; 417.
23. DeWith, G.; Corbijon, A.J. Metal fibre reinforced hydroxyapatite ceramics. *J. Matls. Sci.* **1989**, *24*, 3411–3415.
24. Evis, Z. Ph.D. Thesis. Rensselaer Polytechnic Institute: Troy, N.Y., 2003.
25. Ji, H.; Marquis, P.M. Sintering behavior of hydroxyapatite reinforced with 20wt% Al₂O₃. *J. Matls. Sci.* **1993**, *28*, 1941–1945.
26. Gautier, S.; Champion, E.; Bernache-Assollant, D. Toughening in alumina-platelet-hydroxyapatite composites. *J. Matls. Sci.: Matls in Medicine* **1999**, *10*, 533–540.
27. Kim, S.; Keng, U.; Lee, J.; Kim, H. Effect of calcination of starting powder on mechanical properties of hydroxylapatite-alumina bioceramic composites. *J. Matls. Sci.: Matls. In Medicine* **2002**, *13*, 307–310.
28. Li, J.; Fartach, B.; Hermansson, L. Hydroxyapatite-alumina composites and bone bonding. *Biomaterials* **1995**, *16*, 417–422.
29. Adonfsson, E.; Nygren, N.; Hermansson, L. Decomposition mechanisms in aluminum oxide—apatite systems. *J. Am. Ceramic Soc.* **1999**, *82*, 2909–2912.
30. Li, J.; Liao, H.; Hermansson, L. Sintering of partially stabilized zirconia–hydroxyapatite composites by hot



- isostatic pressing and pressureless sintering. *Biomaterials* **1996**, *17*, 1787–1790.
31. Heimann, R.B.; Vu, T.A. Effect of CaO on thermal decomposition during sintering of composite hydroxyapatite zirconia mixtures for monolithic bioceramic implants. *J. Matls. Sci.: Letters* **1997**, *16*, 437–439.
 32. Rao, R.R.; Kannan, T.S. Synthesis and sintering of hydroxyapatite zirconia composites. *Matls. Sci. and Eng. C* **2002**, *20*, 187–193.
 33. Kim, H.W.; Noh, Y.J.; Koh, Y.H.; Kim, H.E.; Kim, H.M. Effect of CaF₂ on densification and properties of hydroxyapatite-zirconia composites for biomedical applications. *Biomaterials* **2002**, *23*, 4113–4121.
 34. Wu, J.; Yeh, F. Sintering of hydroxyapatite-zirconia composite materials. *J. Matls. Sci.* **1988**, *23*, 3771–3777.
 35. Ahn, E.S.; Gleason, N.J.; Nakahira, A.; Ying, J.Y. Nanostructure processing of hydroxyapatite based bioceramics. *Nano Letters* **2001**, *1–3*, 149–153.
 36. Cook, S.D.; Kay, J.F.; Thomas, K.A.; Jarcho, M. Interface mechanics and histology of titanium and hydroxylapatite coated metal implants. *J. Biomed. Mater. Res.* **1984**, *18*, 719–726.
 37. Ducheyne, P.; Healy, K.E. The effect of plasma sprayed calcium phosphate ceramic coatings on the metal ion release from porous titanium and cobalt-chromium alloys. *J. Biomed. Mater. Res.* **1988**, *22*, 1137–1163.
 38. Ji, H.; Ponton, C.B.; Marquise, P.M. Microstructural characterization of hydroxylapatite coating on titanium. *J. Mater. Sci.: Mater. Med.* **1992**, *3*, 283–287.
 39. Park, E.; Condrate, R.A.; Hoelzer, D.T.; Fischman, G.S. Interfacial characterization of plasma-spray coated calcium phosphate on Ti-6Al-4V. *J. Mater. Sci.: Mater. Med.* **1998**, *9*, 643–649.
 40. VanDijk, K.; Schaeken, H.G.; Wolke, J.G.C.; Jansen, J.A. Influence of annealing temperature on RF magnetron sputtered calcium phosphate coatings. *Biomaterials* **1996**, *17*, 405–410.
 41. Liu, X.; Weng, J.; Tong, W.; Zuo, C.; Zhang, X.; Wang, P.; Liu, Z. Characterization of a hydroxylapatite film with mixed interface by Ar⁺ ion beam enhanced deposition. *Biomaterials* **1997**, *18*, 1487–1493.
 42. Brendel, T.; Engel, A.; Russel, C. Hydroxylapatite coating by polymeric route. *J. Mater. Sci.: Mater. Med.* **1992**, *3*, 175–179.
 43. Li, T.; Lee, J.; Kobayashi, T.; Aoki, H. Hydroxylapatite coating by dipping method. *J. Mat. Sci.: Mat. Med.* **1996**, *7*, 355–357.
 44. Russell, S.W.; Luptak, K.A.; Suchicital, C.T.A.; Alford, T.L.; Pizzicano, V.B. Chemical and structural evaluation of sol-gel derived hydroxylapatite thin films under rapid thermal processing. *J. Am. Ceram. Soc.* **1996**, *79*, 837–842.
 45. Kim, C.S.; Ducheyne, P. Compositional variations in the surface and interface of calcium phosphate ceramic coatings on Ti and Ti-6Al-4V due to sintering and immersion. *Biomaterials* **1991**, *12*, 461–469.
 46. Hero, H.; Wie, H.; Jorgensen, R.B.; Ruyter, I.E. Hydroxylapatite coating on Ti by hot isostatic pressing. *J. Biomed. Mater. Res.* **1994**, *28*, 343–348.
 47. Garcia, R.; Doremus, R.H. Electron microscopy of the bone-hydroxylapatite interface from a human dental implant. *J. Mater. Sci.: Mater. Med.* **1992**, *3*, 154–156.
 48. Ergun, C.; Doremus, R.H.; Lamford, W.A. Hydroxylapatite and titanium: interfacial reactions. *J. Biomed. Matls. Res.* **2003**, 336–343.
 49. Sarikaya, M.; Tamerler, C.; Jen, A.K.Y.; Schulten, K.; Baneyx, F. Molecular biomimetics: nanotechnology through biology. *Nature Materials*, **2003**, *2*, 577–585.
 50. Brown, S.; Sarikaya, M.; Johnson, E. Genetic analysis of crystal growth. *J. Mol. Biol.* **2000**, *299*, 725–732.
 51. Beverly, K.C. Quantum dot artificial solids: understanding the static and dynamic role of size and packing disorder. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 6456–6459.
 52. Kempa, K. Photonic crystals based on periodic arrays of aligned carbon nanotubes. *Nano Letters* **2003**, *3*, 13–18.
 53. Knoblauch, M. ATP-independent contractile proteins from plants. *Nature Materials*, **2003**, *2*, 800–803.
 54. Mavroidis, C.; Dubey, A. Biomimetics: from pulses to motors. *Nature Materials* **2003**, *2*, 573–574.

Ceramics in Dentistry

Nicolas M. Jedyakiewicz

School of Dentistry, The University of Liverpool, Liverpool, U.K.



INTRODUCTION

Teeth are complex organs of the human body and consist of several component tissues, both hard and soft (Fig. 1). If lost, a tooth cannot be regenerated. If damaged, a tooth cannot be repaired by the body. The tooth is subject to many damaging influences. Dental caries (tooth decay) damages teeth by bacterial attack, demineralising and softening the enamel and the underlying dentine. Acidic fluids such as canned drinks and salad dressings cause chemical dissolution and demineralization of the tooth surface. Trauma may chip, crack, or fracture teeth. Those who are fortunate enough to retain their teeth into later life discover the problems brought about by a lifelong developing payload of wear, fatigue, and dietary corrosion within the substance of teeth. The body has a limited capacity for remineralization, where early decay or chemical dissolution has damaged a tooth surface; but if the integrity of the tooth is further threatened, then the best that the tooth can do is to form a calcific mass within the dental pulp. This creates a physical barrier against the intrusion of bacteria and oral fluids, but in no way replaces damaged or missing tooth tissue.

Such a limited response to damage means that few people in the developed world are able to avoid encounters with restorative dentistry. Restorative dentistry concerns itself with repairing damaged teeth and their supporting structures. Patients are increasingly demanding that such repairs should not only function in the manner of an undamaged tooth, but also look like an undamaged tooth. Aesthetics are today of paramount concern, and the only medical material that in any way provide a durable and satisfactory solution to the aesthetic repair of teeth is ceramic. This article examines the manner in which dental porcelains attempt to meet the challenges set by the need to achieve an appropriate visual appearance, whilst providing sufficient strength to accept the loads placed upon teeth during function.

THE STRUCTURE OF TEETH AND THE DERIVATION OF THE AESTHETIC FORM

Teeth have a complex structure with a layer of enamel firmly attached to a supporting framework of dentine.

Enamel forms a crystalline shell of hydroxyapatite. It has an anisotropic structure that is aligned to the direction of maximum loading. The dentine is a strongly anisotropic framework of fluid-filled microtubules that contain cellular components of the underlying dental pulp. The dentine has viscoelastic properties and functions with the overlying enamel in accepting the mechanical load of the bite. The union between the enamel and dentine is profoundly strong and enables the synergistic mechanical behavior of these two rather different materials. It is the creamy blue–white translucency of the enamel sitting upon the rather more opaque and yellow dentine that gives the distinctive appearance to the tooth. The main physical characteristics of a tooth, which contribute to its aesthetic form, are the color (in terms of hue, chroma, and saturation), translucency, and surface reflectance. Opalescence, fluorescence, and metamerism also challenge the dental ceramist in matching the appearance of a tooth. The complexity of such a color can be seen in the handmade ceramic crown, shown in Fig. 2. Analyzing and matching the aesthetic characteristics of a patient's tooth is one of the most challenging elements of restorative dentistry. Furthermore, this description must be communicated to a technician in a remote laboratory where ceramic restorations are made. The conventional approach is to use a series of color-matching tabs that are held adjacent to the tooth surface, and an evaluation and match is made by the dentist. This is fraught with problems such as the color balance of the illuminating light and even the incidence of minor color blindness in the operator. Recently, several automated color analysis devices have been produced, the most sophisticated using hand-held spectrophotometers to analyze the color.^[1]

One problem in matching the optical characteristics of the existing teeth is that the natural appearance is not constant throughout life. Surface biofilms attach differentially to teeth and restorations and may absorb stain, although these may be removed by professional cleaning. A greater problem is that the color and translucency of teeth alter as age progresses, and a ceramic restoration that matches the tooth of a teenager precisely, will not be an effective match when the patient reaches middle age. Cosmetic bleaching of teeth can complicate matters, because unnatural levels of whiteness can be achieved. The bleaching process must

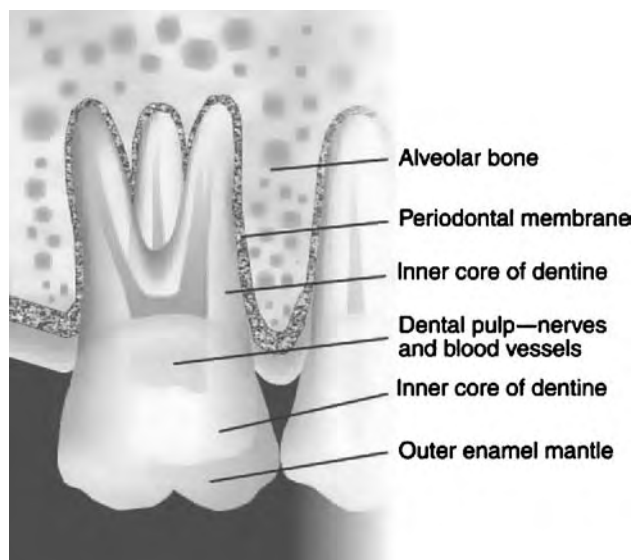


Fig. 1 Diagrammatic representation of the structure of a human tooth.

be refreshed regularly to maintain the brightness. Whilst a gradual color change in the dentition as a whole is not so noticeable, the presence of a fixed reference color in the form of a ceramic restoration can be very distracting.

Overall, the replacement of enamel and dentine with prosthetic materials is extremely challenging. Repair is equally difficult. Any repair needs to have sufficient physical properties to resist repetitive biting forces, and must be fully sealed to the remaining tooth tissue to avoid microleakage at the interface. Furthermore, such a repair may need to function and the seal may need to the tooth remain intact for several decades.

REQUIREMENTS OF CERAMICS IN DENTISTRY

The Oral Environment

The human mouth provides one of the most inhospitable environments for structural materials. It is warm, wet, and both chemically and biologically active. It is capable of generating quite massive force concentrations, and the variance in such characteristics between individuals is enormous. The challenges in restoring damaged tooth tissues are therefore considerable in terms of the mechanical behavior of structural repair materials. The challenges are compounded by the need for the repairs to blend visually into the existing tooth structures. A range of materials are used commonly in repairing teeth and replacing missing teeth, and these include cast metals, dental amalgam, glass polyalkenoate cements, and resin composites. Metals offer the greatest strength and durability. The most successful aesthetically is dental

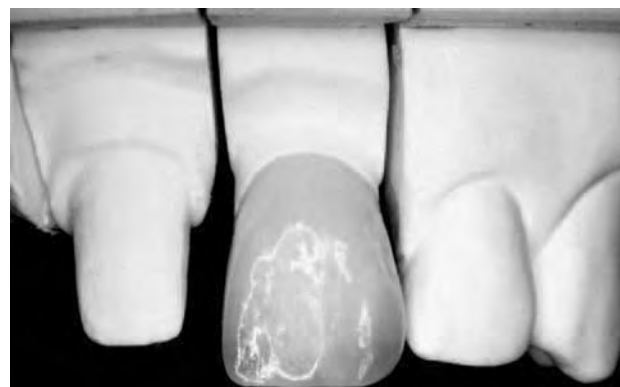


Fig. 2 A completed ceramic crown showing the complexity of color and translucency that is needed to match adjacent teeth.

porcelain, but the mechanical properties are often barely adequate. The challenge to dental science is to endow dental porcelains with sufficient attributes of strength, corrosion resistance, fatigue resistance, and other such factors to enable them to restore damaged teeth for the remaining lifetime of the patient.

Teeth

Physical properties

Teeth, and in consequence dental restorations, are subjected to considerable but varying and repetitive loads. Loading occurs not only during normal biting and chewing, but also in aberrant processes such as the form of nocturnal grinding known as “bruxism.” A small adult may apply as little as 80 N load upon maximal bite. A large and muscular individual may raise 1000 N. Typical values for normal chewing are in the order of 50–250 N, rising to 550–880 N during parafunction.^[2] Therefore, the clinician has to analyze each situation very carefully to make an appropriate choice of dental restoration.

Generally speaking, even when selected correctly, dental ceramics are only marginally able to withstand the loading that may be expected to be applied to them during function and parafunction. The greatest challenge with dental ceramics is to achieve an adequate level of strength in any given clinical situation. The required physical properties and testing methods for dental ceramics are specified in ISO standard 6872.^[3]

Biological properties

Artificial repair materials, when used in the human body, must be biocompatible—that is, they must

invoke an appropriate response from the host organism. This may range from no response at all, to repair and regeneration. When teeth are repaired, there are three main concerns regarding biocompatibility. The first concern is the development of an appropriate biofilm on the surface of the restoration. All surfaces of normal teeth are coated with a biofilm, and it is via the biofilm that remineralization of teeth can occur. It is also the cause of demineralization, when local areas where microorganisms are attached are able to metabolically alter the local pH and commence the processes of dental caries or tooth decay. It is important that any artificial surface does not attract colonization with bacteria that may be pathogenic to the adjacent host tissues. This is essential in cases where a dental restoration comes into contact with the soft tissues of the gums and the periodontal ligament. The periodontal ligament is the attachment of the gum and bone to the tooth. It is a unique structure in that it is the only point in the entire body where the envelope of epithelium that protects us from the outside world is broken. The ligament is complex in form, and the attachment to the tooth surface is mediated via hemidesmosomes. This attachment is the structure that becomes damaged in periodontal disease, and this process may be exacerbated directly by irritant materials or indirectly by the change in profile of the bacterial flora brought about by the nature of the surface of the restoration.

Secondly, there must be some form of seal or adhesive fusion between the restoration and the adjacent tooth tissue. This is essential to prevent the ingress of bacteria and their metabolic products past the restoration and into deep dentine, where irritation of and damage to the underlying dental pulp may occur.

Thirdly, the material itself must not in any way be toxic or leach out any toxic or mutagenic components into the dentine of the tooth or into the mouth.

Ceramics are generally excellent in all these respects. They are usually relatively inert. When polished or glazed correctly, they do not facilitate the development of an inappropriate biofilm, and they do not sustain an overgrowth of pathogenic bacteria. They can be firmly attached to the tooth with dental adhesives and mostly do not leach out any toxic components.

LIMITATIONS OF CERAMICS IN DENTAL APPLICATIONS

Applications of Ceramics in Dentistry

When teeth are damaged, three main types of restoration are used to effect a repair: block replacement (fillings and inlays), shell coverage (crowns), and

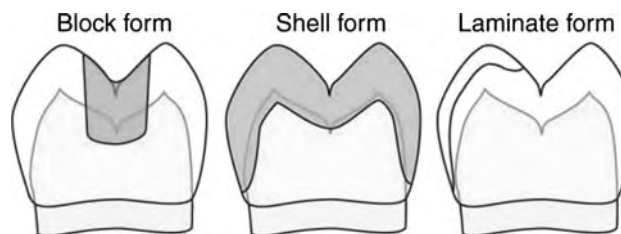


Fig. 3 Repair of teeth (from left): block form (fillings and inlays), laminate form (dental veneers), and shell form (crowns).

surface lamination (dental veneers) (Fig. 3). In the case of block replacement, there is a large volume of material that is usually supported on several surfaces by the containing tooth. The significant thickness of the restoration helps in resisting loads, and adequate strength is relatively easily achieved. Shell form coverage represents the typical dental crown. The tooth is reduced to a generally conical form upon which a crown can be cemented. Traditionally, the cement layer is merely a lute or watertight seal composed of an inorganic cement. In such a configuration, the ceramic crown needs great strength to resist biting loads.

Surface lamination is a more recent innovation. An all-ceramic restoration—either a full crown veneer or a part veneer—is fabricated. The fitting surface of the ceramic restoration is then etched with hydrofluoric acid to create microscopic convolutions into which a resin may be run (Fig. 4). Similarly, the tooth surface is prepared by etching with phosphoric acid. This creates a surface on the enamel, similar to that of the etched ceramic. The surface of the dentine behaves rather differently, but the demineralizing effect of the acid etching leaves a surface that can be diffused with polymer resins to gain an adhesive bond. In this manner, the ceramic restoration can be firmly attached to the underlying tooth substance, and this compound form achieves a very high working strength. Clearly, such a process requires a disciplined and careful clinical stage, and contamination at any point in the attachment procedure may compromise the mechanical strength of the restoration enormously. From an aesthetic viewpoint, however, a superlative result can be achieved, and the advantages have been demonstrated in clinical practice.^[4,5]

DENTAL CERAMICS

Feldspathic Porcelain

The original form of dental porcelain is feldspathic porcelain. This type of porcelain presents as a powder, which is mixed with water to form a slurry, and this is



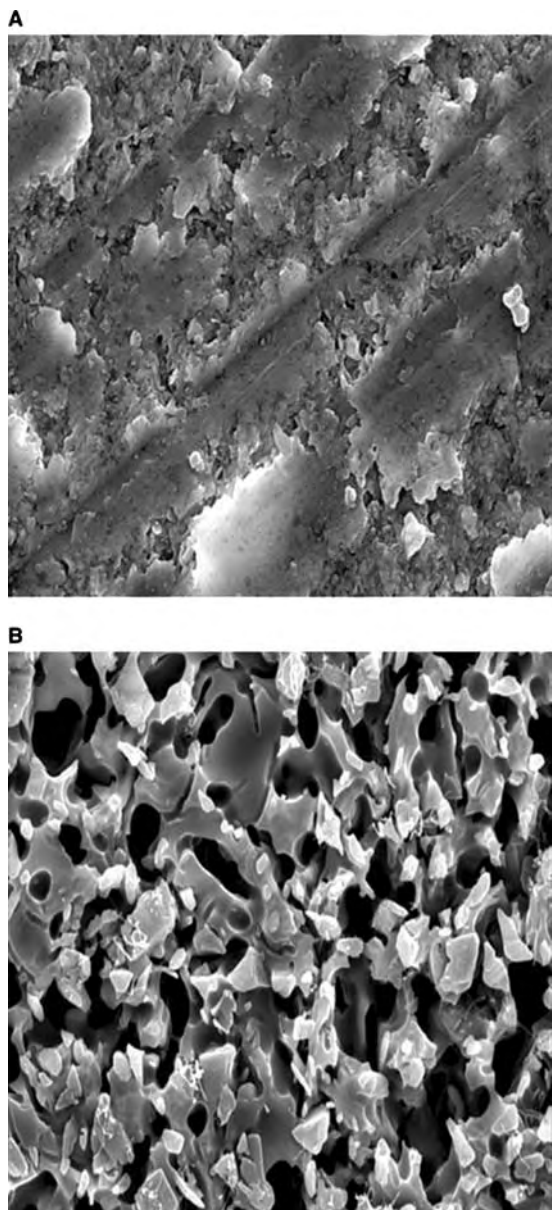


Fig. 4 Ceramic surface. (A) After machining to form and (B) after etching with hydrofluoric acid. Field of view: $60 \times 60 \mu\text{m}$.

stacked up onto a die form by hand and then sintered in a furnace. The die is cast in dental stone from a dental impression, and it is covered with a foil of platinum upon which the restoration is built up (Fig. 5). The foil is removable from the die and is used to carry the stacked porcelain powder to the furnace where it is sintered, maintaining the form of the fitting surface of the crown. The process of sintering causes the porcelain mass to shrink considerably, and this shrinkage must be accommodated by overbuilding before firing. Sintering leaves a dense but microscopically rough and porous outer surface to the restoration, and this must be filled



Fig. 5 Stacking feldspathic dental porcelain powder upon a die to form a dental crown.

in and sealed by the application of a glaze layer of a lower melting point ceramic.

Whilst the shape of the fitting surface is maintained by the foil matrix, shrinkage of the ceramic still occurs. This leads to small crack defects on the fitting surface of the restoration. These blemishes are known as Griffith's flaws, and they contribute greatly to a reduction in the overall strength of the restoration, because it is from these defects that cracks propagate when the crown is loaded (Fig. 6). The whole process of building up a ceramic crown is immensely labor intensive, and very high levels of skill are needed. Nevertheless, by building up crowns with various matching colors and translucencies of porcelain powder, a result that is indistinguishable from the natural teeth can be achieved. Glazed ceramics also provide a surface that is very resistant to the buildup of dental plaque and is very acceptable, biologically, to soft

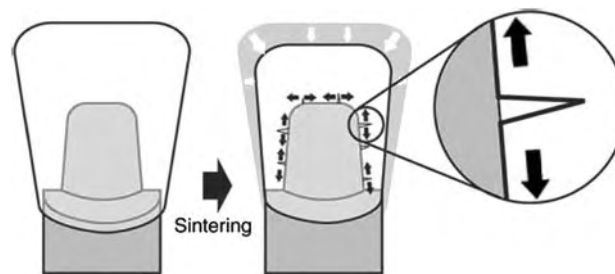


Fig. 6 Crack defects occur on the fitting surfaces of ceramic crowns owing to shrinkage upon firing. The carrier resists deformation as the porcelain shrinks and therefore crack defects open up. These are known as Griffith's flaws, and their presence or absence determines the strength of the crown.

tissues that contact it. The greatest problem is that the strength of such a restoration is very low, limiting the application to situations where loading forces are minimal.

Porcelain Bonded to Metal

The problem of Griffith's flaws on the fitting surface of feldspathic porcelain crowns has been noted above. By bonding the porcelain to a metal carrier, the inner surface of the porcelain mantle is left under tension, and crack flaws do not form at the metal–ceramic interface (Fig. 7). The metal carrier then becomes part of the crown. This increases the strength of restoration by eliminating the initiation points of potentially propagating cracks. The strength of this combination enables porcelain to be used on load-bearing surfaces, but the downside is that the metal must be hidden under a layer of opaque ceramic before the translucent layers are added.

The challenge in achieving the combination of ceramic bonded to metal is that the metal must have a higher melting point than the ceramic, so that the ceramic powder can be fused over it. The second problem is that normally ceramics have a much lower coefficient of thermal expansion than metals. This would mean that as the metal carrier cooled and shrank, the overlying ceramic would fracture off.

In 1962, Weinstein formulated a porcelain powder that had a high coefficient of thermal expansion and low fusion temperature.^[6] A gold alloy with a relatively low thermal expansion and a high fusion temperature was achieved by alloying gold with platinum and palladium. Finally, by adding a small amount of tin to the metal, a tin oxide layer can be created on the metal carrier that can fuse with the metal oxides of the applied porcelain. The glassy phase of the porcelain dissolves the oxide at the metal surface to form a bond with the metal. In this manner, porcelain could be coated and bonded onto a metal base giving an

enormous gain in strength. A second advantage is in the precision of fit, because this is now determined by metal casting rather than by the sintered porcelain powder.

A metal substructure allows an extension of the restoration where necessary. For example, a tooth may be replaced by a crown that is attached to one or more crowns on teeth adjacent to the gap in the form of a bridge. Fig. 8 shows metal–ceramic crowns that have been precision milled to enable the attachment of a removal denture to replace missing teeth.

In recent years, several lower cost alloys have been developed to substitute for the original precious metal system. These include palladium–silver alloys and nickel–chromium alloys.

Porcelain bonded to metal is of the highest strength, and this is the most ubiquitous method of using ceramics in dentistry.

High-Alumina Porcelain

To the clinical dentist there seems to be an inverse relationship between aesthetics and strength in dental ceramics. As the strength increases there is a significant loss of aesthetics. One of the earliest attempts to achieve greater strength in ceramic crowns was the use of high alumina feldspathic porcelain—typically 40% Al_2O_3 . The alumina increased the strength by blocking crack propagation, but this is at the expense of an almost complete loss of translucency. Therefore porcelain crowns with high-strength core structures were developed, whereby an aesthetic porcelain overlay was fired against a high-strength but opaque core (Fig. 7B). The downside of this technique is that the minimum thickness of the crown increases, requiring the dentist to cut back the tooth further. If insufficient space is provided to the technician, then a crown is returned to the dentist with a bright alumina core showing through the surface.

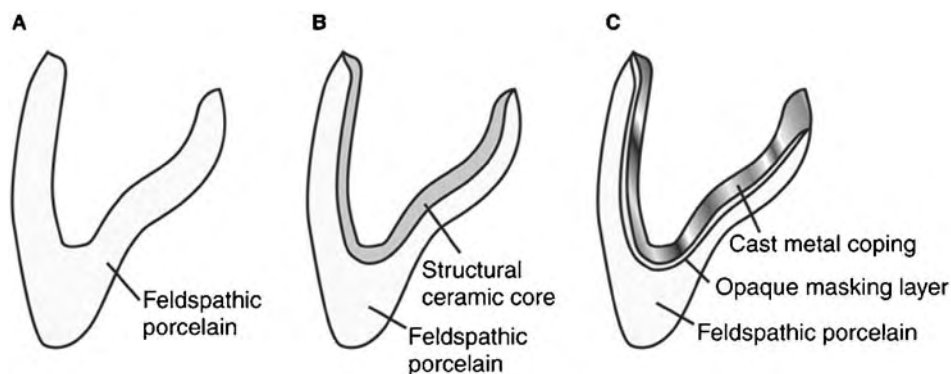


Fig. 7 Structural form of dental crowns. (A) A conventional porcelain jacket crown, (B) a porcelain crown reinforced with a structural ceramic core and (C) a porcelain-fused-to-metal crown.



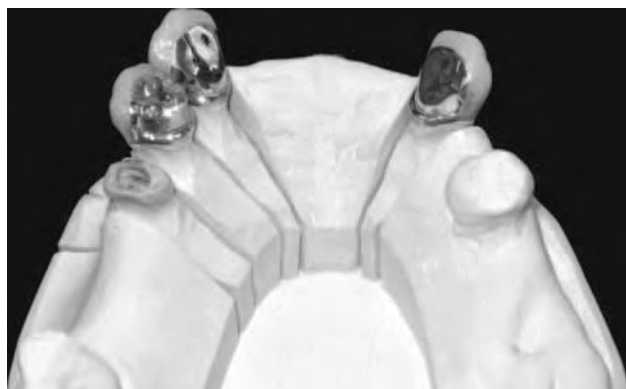


Fig. 8 Three porcelain-bonded-to-metal crowns on a laboratory model. The metal substructure is left open on the hidden side and has been precision milled to enable a denture to be attached onto the crowns. The retention of these teeth increases the stability of the denture very considerably.

Glass-Infiltrated Cores

A further development of the high-strength ceramic core approach uses a sintered core of alumina or zirconia that is subsequently infused with molten lanthanum glass. The system is marketed as “In-Ceram” (Vita-Zahnfabrik, H. Rauter GmbH & Co, Bad Sackingen, Germany).

The ceramic core is formed by applying a slurry of alumina powder to a die in a process that is known as slip casting. This core is dried and then fired very slowly in a furnace, typically at over 1120°C for 10–12 hr. This temperature is insufficient to melt the alumina, but does permit solid-phase sintering to occur between adjacent particles. This results in a highly porous structure with little shrinkage compared with conventional dental porcelain. This porous shell form is placed in a dish and sprinkled with lanthanum glass powder. In a further firing at 1100°C for around 5–6 hr, the glass melts and infiltrates into the porous alumina core. At the end of the process, excess glass is attached to the surface of the core and is ground off with rotary diamond grinders. The resultant core is used as a base upon which conventional feldspathic porcelain powders are applied to develop the final structural and aesthetic form. Whilst the strength is increased significantly, the fitting surface of the restoration does not lend itself easily to adhesive cementation techniques because it resists etching.

Glass Ceramics

Glass ceramics were developed for dental applications by the Corning Glass company in the late 1980s,^[7] and were marketed under the brand name “Dicor” and

“Dicor MGC” (machineable glass ceramic). Glass ceramics offer some novel advantages for dental applications. Firstly, they are relatively, dimensionally constant when cast or hot pressed to shape. Secondly, they can be machined relatively easily, allowing them to be fabricated by computer-aided design–computer-aided manufacturing (CAD–CAM). The surface that results from machining can be polished successfully to give an excellent external surface to the restoration. The machined surface can also be etched to create a microscopically porous surface that is appropriate for adhesive attachment on the fitting surface.^[8] Glass ceramics are also very aesthetic, as they can be formulated to replicate both the color and the translucency of human enamel.

Glass ceramics have several possible formulations. Typically, the microstructure consists of a highly interlocked array of tetrasilicic fluormica crystals at approximately 70% by volume in a nonporous glass matrix.^[9] Oxide colorants are added. The glass is melted at temperatures in excess of 1400°C, and then cooled and annealed. At this stage the glass is clear and may be shaped by casting or rolling. The material becomes a glass ceramic at the next stage, where it is, in its final shaped form, reheated at between 650°C and 1150°C for several hours, a process of tempering that is sometimes referred to as “ceramming.” This produces a uniform dispersion of randomly oriented tetrasilicic mica crystals. The tempering process is critical in both time and temperature to ensure appropriate crystal growth. If the crystals are either too small or too large, then the physical properties of the final material will be compromised. The mica crystal plates are around 2 μm in size, and their formation causes the transparent material to become translucent and to develop the final color and aesthetic form, as well as the final strength.

If the glass ceramic is to be cast or pressure formed, it is supplied as clear glass ingots that require post-processing in a furnace. If the material is to be machined, it is supplied fully processed as a glass ceramic block, with the final optical characteristics fully established. Glass ceramic crowns are pressed or machined in the full, required shape, and additional ceramic material is not normally added. Therefore, the final aesthetic form may require the application of surface stains to achieve a fully characterized color match to the underlying tooth.

Leucite-Reinforced Ceramics

The glass ceramic “Empress” (Ivoclar-Vivadent, Schaan, Liechtenstein) is a glass ceramic with leucite crystal precipitation within the material.^[10] Like the Dicor material, it can be either machined or hot pressed.

Alumina-Reinforced Ceramics

High-purity alumina powder can be isostatically compacted into ingots that can then be machined into a substructure for dental ceramic restorations. Following machining, the shaped mass is sintered at temperatures in the range 1600–1800°C, a process that transforms it into a dense polycrystalline solid with substantial physical properties. Dimensional changes occur during the sintering process, and therefore the pattern shaping of the ingot is usually performed by a CAD–CAM system that allows an overbuild proportional to the final shrinkage. Conventional dental porcelains can then be built up over this core to form an aesthetic restoration. This principle underlies the Procera AllCeram system (Nobel Biocare AB, Gothenburg, Sweden), which uses a digitizer to scan a conventional model die. The data are then sent to a central processing factory, where the construction of the alumina coping occurs. The core is then returned to the participating dental laboratory for the completion of the process by the addition of aesthetic feldspathic porcelains.

The fitting surface of an alumina-cored crown cannot be etched to gain micromechanical attachment of adhesive luting cements. Therefore, many of the advantages of adhesive luting are unavailable in this method.

Zirconia-Reinforced Ceramics

Zirconia offers the highest strength material for the ceramic reconstruction of teeth. Although impressive by comparison with other forms of dental ceramics, when compared with metals, there is still a massive discrepancy. A zirconia-based restoration may have a flexural strength one order of magnitude greater than that of a feldspathic porcelain crown. Metals may achieve more than two orders of magnitude. Therefore cast metal restorations remain the most durable method of repairing in high-stress situations such as multiple unit bridgework.^[11] The vulnerable zone in bridgework is the connector, which is the component that attaches the replacement tooth to the crowns on the adjacent teeth. Ceramic connectors need to be substantial to resist flexural loading as biting occurs. Zirconia connectors over 7 mm in diameter have been shown to be unable to resist maximum biting loads.^[12]

Zirconia has a dense white color. This is unacceptable for the final appearance of a dental restoration, but it does give a suitable backdrop for the buildup of other aesthetic ceramics as a surface veneer. It is in this manner that zirconia is used, mainly to form a substructure around which aesthetic but lower strength veneer porcelains can be applied. Coloration of the

zirconia core is now becoming available, and this means that the balance between the thickness of the high-strength core and the overlying aesthetic porcelain can be altered in favor of strength.^[13]

The great strength of zirconia is attributed to its resistance to crack propagation by a mechanism of transformation toughening. Zirconia undergoes a volume change related to a tetragonal phase to monoclinic phase transformation. After initiation at the surface, a propagating crack causes a transformation of adjacent zirconia grains into a monoclinic phase. The monoclinic phase occupies a larger volume, and therefore compressive stresses are induced in the material at the crack front. More energy is required for the crack to propagate further, and therefore the progress of the crack may be halted.

Zirconia is a difficult material to fabricate. During machining, the same phase transformation that provides it with strength causes a dimensional change. Hence, it resists machining aggressively because of its high fracture toughness.

As a work around, many systems use a block of consolidated zirconia particles that is machined to an appropriate form. The machined block is then fired slowly in a furnace to sinter the zirconia into the high-strength structure. Sintering results in a considerable shrinkage of the material that may be in the range of 25%. This must be compensated for, by shaping the material into an appropriately oversize block, before firing. This cannot be done by a technician, and therefore CAD–CAM systems are used (Fig. 9). Like the alumina cores, zirconia cores cannot be etched to gain micromechanical attachment of luting cements to the fitting surface.



Fig. 9 Zirconia frameworks for porcelain crowns machined oversize by CAD–CAM (CERCON), prior to trimming and firing.



Ceramic Resin Composites

The commonest method of using ceramic materials to restore teeth is to insert a composite paste of powdered glass or ceramic bound in a polymerizable plastic resin. The resin enables adhesives to be used, to bind the restoration to the underlying tooth. The disadvantage of this system is that the resin systems that are used currently shrink upon setting, and this polymerization shrinkage must be accounted for in the placement process. Deterioration of the resin in the hostile environment of the mouth is a limiting factor in the life of such restorations. Nevertheless, it is such resin composite materials that provide the best interface between ceramic restorations such as crowns and veneers and the underlying tooth substrate.

Bulk Fillers

Methods of fabrication

Some aspects of fabrication methods for dental ceramics have been touched upon earlier, but their importance requires further examination.

Accuracy and fit

The dental ceramic restoration must be made to very high degrees of accuracy. It must fit precisely to the prepared tooth surface, with just enough space to permit an interposed adhesive film. It must have an occlusal (biting surface) shape that provides a precise series of contacts with the opposing teeth, and a shape that precludes interfering contacts when the jaw moves in a grinding or chewing pattern. The restoration must also fit between adjacent teeth with enough contact pressure to prevent the interposition of food debris, but without so much pressure that the restoration could not be fitted without the contacting teeth being forced aside.

Whilst certain adjustments can be made by the dentist at the time of fitting the restoration, notably to the occlusal surface and the points contacting adjacent teeth, such adjustments can damage the surface finish. Some ceramics can be quickly reglazed if a glazing furnace is to hand, but such a stage is time consuming and expensive in surgery time. An unglazed rough ceramic surface can cause significant wear of the opposing tooth surfaces and should be avoided.^[14] One advantage of machineable ceramics is that they are usually both internally dense and free of porosity, and are polishable to a high glaze. This can be achieved easily at the chair-side.

Many aspects of the subjects discussed herewith are involved with accuracy of fit and the dimensional

changes that occur within the fabrication processes. Therefore, this raises the question of how accurate a fit is needed. All laboratory made restorations need to be cemented into place, and the cement must occupy an evenly space layer between the crown and the tooth. When a crown is being fitted, it is filled with a low-viscosity cement and then pressed onto the prepared tooth. The cement has to flow, and the excess material has to be expressed out at the marginal gap. The limit of this flow and the resultant marginal fit will be a function of both the rheological behavior of the cement and fit of the crown. To allow for the thickness of the cement layer, the die upon which the crown is fabricated is often coated with several layers of a varnish to build up the potential cement space. The varnish is omitted from the margins of the die to allow tight closure of the crown.

Overall, a marginal fit of 50 μm is generally accepted as good. In actual practice, the marginal fit is frequently in excess of 100 μm .

To achieve such a precision of fit, the technician must be provided with an accurate high-resolution impression of the prepared teeth. Vinyl polysiloxane impression materials provide both accuracy and resolution to contemporary dental impressions. Such materials are also acceptably stable in transit, and enable an effective model to be cast in the dental laboratory in a high-strength dental stone. Such models are visible in Figs. 5 and 8.

Mapping the biting surfaces of dental restorations to the contacts and movements of the opposing jaw is more complicated. Various methods and devices are available to record this information. This is used to set up and adjust a model of the working jaws that is referred to as a dental articulator.

Sintering

Sintering is a process whereby a paste of particulate form is heated to the point where surface melting of the particles just begins to occur. This allows fusion of the surfaces of adjacent particles without full melting, and conglomeration into a single mass. Sintering allows a structure to retain its shape. One step further in melting and flow would occur, and the shape would be either compromised or lost altogether. The process is accompanied by a considerable degree of shrinkage, and this must be accommodated in the fabrication process. The methods of compensation for this shrinkage are effective, and sintering produces accurately fitting dental restorations.^[15] Contemporary furnaces raise and lower the temperature of the workpiece according to programmes that are specific for the particular porcelain product. A slow preheat outside the furnace drives off water from the porcelain, to prevent the generation of steam later in the firing cycle. The workpiece is introduced into the furnace,

and the binders are burnt off to allow intimate contact between the porcelain powder particles. The temperature is increased subsequently until sintering occurs. Exclusion of air by the establishment of a vacuum during the firing cycle is an important element of achieving the correct porcelain density during sintering. The cooling phase is equally important to prevent internal stresses and subsequent cracking.

Casting

Casting is the process whereby a molten mass is injected into a void within a mould in the manner of the lost wax process. The mass is allowed to cool, and then the mould is broken away to release the cast object. Again, dimensional changes occur for many materials that are processed in this way. The exceptions are the glass ceramics. These materials exhibit very little change in dimension upon casting.

Hot pressing

Hot pressing is where a glass-like mass of ceramic is heated until softening occurs. The softened mass is then squeezed under pressure into a mould. The advantage of this method is that the ceramic is not molten, and therefore dimensional changes during the forming process are reduced.

Copy milling

The celay machine is a sophisticated copy-milling machine that can be used to follow a pattern and copy the shape into a block of ceramic. The operation requires the preparation of a die and the manufacture of a pattern from the die. The pattern is then placed on one side of the machine, with a ceramic blank being inserted into a chuck on the other side. A diamond-grinding cutter follows the motion of a tracing stylus, which is moved by the operator over the pattern until contact is achieved. The process is repeated over the entire surface of the pattern until the shape has been fully translated over to the recipient ceramic block. A range of matching cutters and styli allow fine detail to be copied.

CAD-CAM

Computer-controlled machining of dental ceramics offers several advantages over manual systems. Firstly, it allows the use of materials that could not be fabricated by conventional methods. Secondly, it may reduce the need for skilled dental technicians for basic restorations.

The CAD-CAM systems consist of several key components. There must be a shape acquisition device to determine the shape of the substrate tissue against

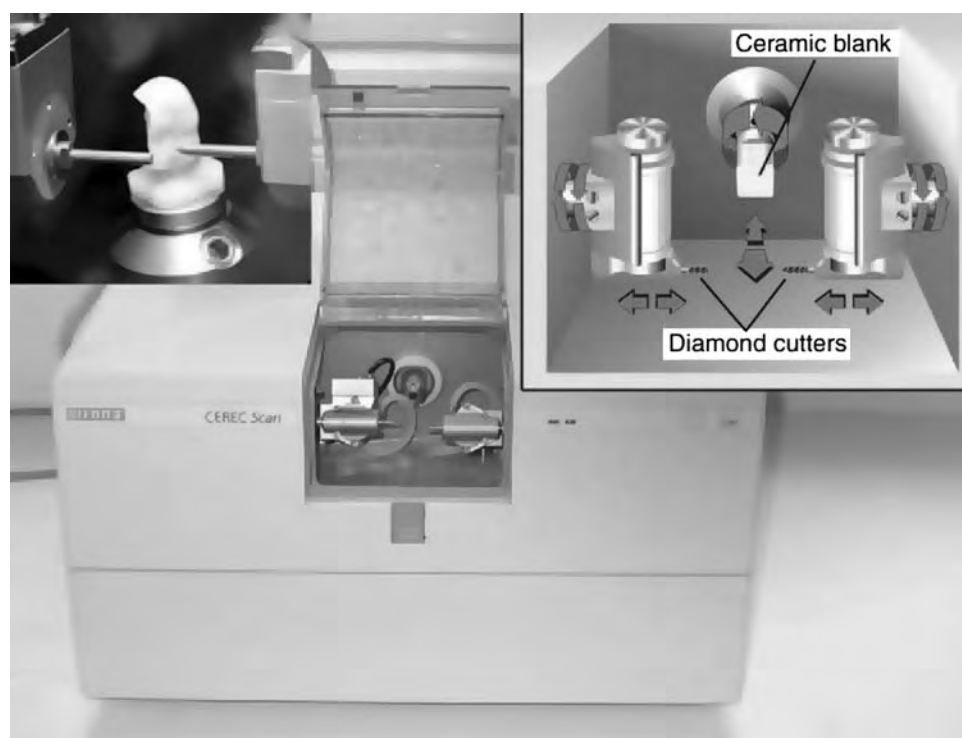


Fig. 10 A Cerec CAD-CAM milling unit. Inset left: machining an onlay restoration. Inset right: detail of the movements possible with the milling cutters.

which the restoration is to be made. Secondly, there must be a computer system to generate the spatial form of the restoration. Thirdly, there must be some form of machining or milling device to fabricate the restoration.

One of the first CAD-CAM systems in dentistry was the laboratory-based Procera system (Nobel Biocare, Göteborg, Sweden). This uses a touch probe scanner to map the surface shape of model die. High-strength alumina or zirconia copings are machined, and then aesthetic dental porcelain is added and fired conventionally. High levels of success have been claimed for this system.^[16]

The most established CAD-CAM system in dentistry is the Cerec system (Cerec, Sirona, Bensheim, Germany).^[17] This started off as a chair-side system that could scan a prepared tooth surface and generate an all ceramic restoration to fit it, whilst the patient was in the chair for a single appointment. The device combined a shape acquisition unit, a milling unit, and a computer system into a small mobile trolley that was used in the dental operatory. In the latest version of the system, the shape acquisition unit and the milling machine have been separated. The machine is multi-functional and can be used either at the chair-side or in the dental laboratory.

A feature unique to the Cerec system is the hand-held three-dimensional scanner that is used to acquire the morphologic profile of the prepared tooth. A striped reference pattern is projected onto the tooth surface in infrared light through telecentric lens. The distortion in the stripes brought about by the surface contour is observed through the same lens by a digital camera. The recorded pattern is used to build up a shape model of the tooth surface. The whole scanning and recording process takes a fraction of a second, so the scanner can be held by the dentist without a rigid support.

The machining system comprises two electrically driven and water-cooled machining heads fitted with diamond cutters (Fig. 10). These cutters act simultaneously upon opposing the sides of a centrally mounted and rotating ceramic blank. A full range of the shapes required for dental restorations can be cut. Restorations produced by the system have been demonstrated in clinical trials to offer high levels of clinical performance.^[18]

The CAD-CAM also offers an alternative to slip casting in building up restorations using glass infiltration techniques (see above). Blocks of alumina, zirconia, or spinel are manufactured by hot pressing. These porous blocks machine quite freely into a core form. The cores are then infiltrated with an appropriate glass and processed conventionally, with the technician building up the shape, form, and color with overlays of feldspathic porcelain.

Ceramic Preforms

Ceramic inlays are normally constructed to be tight fitting to the surrounding tooth structure, and to be retained with a composite resin cement. One method of achieving this is to cut a cavity into the tooth using a truncated cone-shaped diamond burr and then to insert a size-matched truncated cone of silanated glass ceramic. The Cerana system (Cerana, Nordiska Dental AB, Helsingborg, Sweden) offers such a range of matched cutters and ceramic cones. Clearly, it will not always be possible for a precise truncated cone form to match the carious lesion in the tooth, and therefore any deviation of the cavity form from a circular shape can be accommodated by an infill of composite resin. Such composites are often used as restorative materials by themselves, and, therefore, the extensions of the cavity form that are filled with a composite do not compromise the restoration to any great degree.

Adhesion to Ceramics

Ceramic restorations gain significant strength from being firmly bonded to the underlying tooth. Whilst this



Fig. 11 Trauma in the form of being kicked by a horse has fractured off the entire enamel shell from the outer surface of this patient's first premolar tooth. A veneer of ceramic was machined by CAD-CAM and bonded on in a single visit.

is a desirable approach for all ceramic restorations, some types such as veneers are entirely dependent upon adhesive placement (Fig. 11). The conventional method of preparing a ceramic for adhesion is to etch it with hydrofluoric acid, typically at a 5% concentration. In the case of sintered ceramics, this opens up microscopic involutions into the surface and allows interpenetration of polymeric resins into the ceramic surface, giving a micromechanical bond (Fig. 4). This attachment is enhanced by the use of coupling agents—methacrylate functional organosilanes—that are applied to the etched surface. The silanes attach to the silica in the ceramic and link to the polymers in the cement system.^[19]

Some ceramics, such as zirconia, that are used for coping construction are highly resistant to hydrofluoric acid etching, and other forms of surface modification such as tribochemical silica coatings may be necessary.^[20] Bonding may be improved by grit blasting the fitting surface. While this does not create the micro-involutions of etching, it roughens the surface micro-profile very considerably, and therefore increases the surface area that is available for bonding.

CONCLUSIONS

Ceramics are used widely in dentistry in many different forms. No single system yet offers the desirable combination of high strength, excellent aesthetics, and simplicity of fabrication. Metals form an essential component of the dentist's reconstructive technology, and the gap between the strength of metallic restorations and the strength of ceramics remains significant. Whilst zirconia cores with porcelain veneer layers offer much promise, long-term clinical trials are not yet available. The CAD–CAM is a maturing technology that reduces many of the variables in fabrication; but on a global basis, the cost remains quite high, limiting its use to populations that enjoy a strong economic structure.

ARTICLES OF FURTHER INTEREST

Biologic and Synthetic Apatites; Biphasic Calcium Phosphate (BCP) Bioceramics: Chemical, Physical, and Biological Properties; Zirconia Ceramics

REFERENCES

1. Paul, S.J.; Peter, A.; Rodoni, L.; Pietrobon, N. Conventional visual vs spectrophotometric shade taking for porcelain-fused-to-metal crowns, a clinical comparison. *Int. J. Periodontics Restorative Dent.* **2004**, *24* (3), 222–231.
2. Lundgren, D.; Laurell, L. Occlusal forces in prosthetically restored dentitions, a methodological study. *J. Oral Rehabil.* **1984**, *11* (1), 29–37.
3. International Standards Organisation, ISO Standard 6872, Dental Ceramics, 1997.
4. Burke, F.J.; Fleming, G.J.; Nathanson, D.; Marquis, P.M. Are adhesive technologies needed to support ceramics? An assessment of the current evidence (1988–2002). *J. Adhes. Dent.* **2002**, *4*, 7–22.
5. Fradeani, M.; Redemagni, M.; Corrado, M. Porcelain laminate veneers, 6- to 12-year clinical evaluation—a retrospective study. *Int. J. Periodontics Restorative Dent.* **2005**, *25* (1), 9–17.
6. Weinstein, M.; Katz, S.; Weinstein, A.B. Fused Porcelain-To-Metal Teeth. US Patent 3,052,982, Sept. 11, 1962; Porcelain-Covered Metal-Reinforced Teeth. US Patent 3,052,983, Sept. 11, 1962.
7. Grossman, D.G.; Johnson, J.L. Glass Ceramic Compositions for Dental Constructs. U.S. Patent 4,652,312, March 24, 1987.
8. Martin, N. Interfaces in Dental CAD-CAM Restorations. Ph.D. thesis, The University of Liverpool, 1996.
9. Grossman, D.G. Structure and physical properties of Dicolor/MGC glass-ceramic. In *International Symposium on Computer Restorations*. Moermann, W.H., Ed.; Quintessence: Berlin, 1991; Vol. 103, 115.
10. Hoeland, W.; Frank, M. Material Science of Empress Glass-Ceramics, Ivoclar-Vivadent Report No. 10; July 1994, 3–8.
11. Raigrodski, A.J.; Chiche, G.J. The safety and efficacy of anterior ceramic fixed partial dentures, a review of the literature. *J. Prosthet. Dent.* **2001**, *86* (5), 520–525.
12. Luthy, H.; Filser, F.; Loeffel, O.; Schumacher, M.; Gauckler, L.J.; Hammerle, C.H. Strength and reliability of four-unit all-ceramic posterior bridges. *Dent. Mater.* **2005**, *21* (5), 476–482.
13. Devigus, A.; Lombardi, G. Shading Vita YZ substructures, influence on value and chroma, Part I. *Int. J. Comput. Dent.* **2004**, *7* (3), 293–301.
14. Schuh, C.; Kinast, E.J.; Mezzomo, E.; Kapczynski, M.P. Effect of glazed and polished surface finishes on the friction coefficient of two low-fusing ceramics. *J. Prosthet. Dent.* **2005**, *93* (3), 245–252.
15. Balkaya, M.C.; Cinar, A.; Pamuk, S. Influence of firing cycles on the margin distortion of 3 all-ceramic crown systems. *J. Prosthet. Dent.* **2005**, *93* (4), 346–355.
16. Fradeani, M.; D'Amelio, M.; Redemagni, M.; Corrado, M. Five-year follow-up with Procera all-ceramic crowns. *Quintessence Int.* **2005**, *36* (2), 105–113.
17. Jedyakiewicz, N.M.; Martin, N. The Cerec method, science, research and clinical application. *Compend. Contin. Educ. Dent.* **2001**, *22* (6), 7–13.
18. Martin, N.; Jedyakiewicz, N.M. Clinical performance of Cerec ceramic inlays, a systematic review. *Dent. Mater.* **1999**, *15*, 54–61.
19. Guler, A.U.; Yilmaz, F.; Ural, C.; Guler, E. Evaluation of 24-hour shear bond strength of resin composite to porcelain according to surface treatment. *Int. J. Prosthodont.* **2005**, *18* (2), 156–160.
20. Bottino, M.A.; Valandro, L.F.; Scotti, R.; Buso, L. Effect of surface treatments on the resin bond to zirconium-based ceramic. *Int. J. Prosthodont.* **2005**, *18* (1), 60–65.



Chitosan

M.N.V. Ravi Kumar

Department of Pharmaceutics, National Institute of Pharmaceutical Education and Research (NIPER), Mohali, India

S. M. Hudson

Fiber and Polymer Science Program, North Carolina State University, Raleigh, North Carolina, U.S.A.

INTRODUCTION

Chitin is the second most abundant biopolymer on earth (after cellulose). Its annual production is estimated to be almost as high as the production of cellulose. The chemical structure of chitin is similar to that of cellulose, with the hydroxyl group at position C-2 replaced by an acetamide group. Chitosan is obtained by alkaline deacetylation of chitin, present in fungal cell walls and exoskeletons of arthropods such as insects, crabs, shrimps, lobsters, and other vertebrates.^[1] The length of the chitin molecule ranges from 5000 to 8000 N-acetyl-D-glucosamine residues in crabs and may be only 100 residues in the case of some yeasts. However, these values often decrease as a consequence of chitin degradation during its deproteinization in hot alkali solution. Chitosan is a nontoxic biocompatible polymer used unambiguously in all sectors of medicine and pharmacology.^[2] This article discusses biomedical applications of chitosan, with emphasis on developments in drug and gene delivery, and a brief discussion of tissue engineering and food technology.

USES OF CHITOSAN

Chitosan (Fig. 1) may be the only biomaterial used in every sector of the medical and pharmaceutical industries, as well as in other applications.^[3–10] One characteristic of chitosan—its low aqueous solubility at a physiological pH of 7.4—limits its use as an absorption enhancer in some pharmaceutical applications. To overcome this limitation, researchers have taken advantage of the reactive functional groups of chitosan at the C-2 (amino group), C-3 (primary hydroxyl), and C-6 (secondary hydroxyl) positions to make various chitosan derivatives.^[2,3,11] Chitosan was discovered to accelerate wound healing in 1970, leading to applications in wound-healing bandages.^[12] Chitosan has been investigated as an excipient in the pharmaceutical industry in direct tablet compression,^[13] as a tablet disintegrant,^[14] for production of controlled-release solid dosage forms,^[2,4,7,8,15] and

for the improvement of drug dissolution.^[16,17] Consumption of chitosan has been clinically approved and is being marketed around the world as a trapper of fat. Chitosan is usually a high-molecular weight polymer possessing positive charge. It has film-forming ability as well as gelation characteristics and therefore finds applications in the formulation of nonviral gene delivery systems.^[18,19]

DRUG DELIVERY

Literature suggests that chitosan has been extensively investigated for its potential in the development of controlled-release systems.^[20,21] Controlled drug delivery technology is based on the systematic release of a pharmaceutical agent to maintain a therapeutic level of the drug in the body for a sustained period of time (Fig. 2). This may be achieved by incorporating the therapeutic agent into a degradable polymer vehicle, releasing the agent continuously as the matrix erodes. This concept has been demonstrated using chitosan as an excipient in various forms (viz., in hydrogels, microspheres, nanoparticles, films, etc.) (Fig. 3).^[2,22–26] Numerous such systems have been reported in the literature.

Risbud and Bhonde^[27] studied the sustained release of antibiotics in polyacrylamide–chitosan hydrogels, in an attempt to explore the favorable properties of both polyacrylamide and chitosan. Hydrogels were prepared by cross-linking an acrylamide–chitosan mixture (8:2 v/v) with N,N' methylene bisacrylamide. The gels were further characterized by SEM and octane contact angle measurements. Swelling, cytotoxicity, and release studies of the hydrogels were performed. Amoxicillin was used as a model drug. The researchers found the hydrogels to be porous with a pore size of approximately $104 \pm 7.61 \mu\text{m}$, and observed matrix degradation after the drug was released from the hydrogels at pH 7.4.^[27] They observed no pH-dependent swelling; however, a swelling ratio of about 16 was found in all buffers studied at the end of 24 hours. About 56% of the drug was released after 24 hours and 77% was

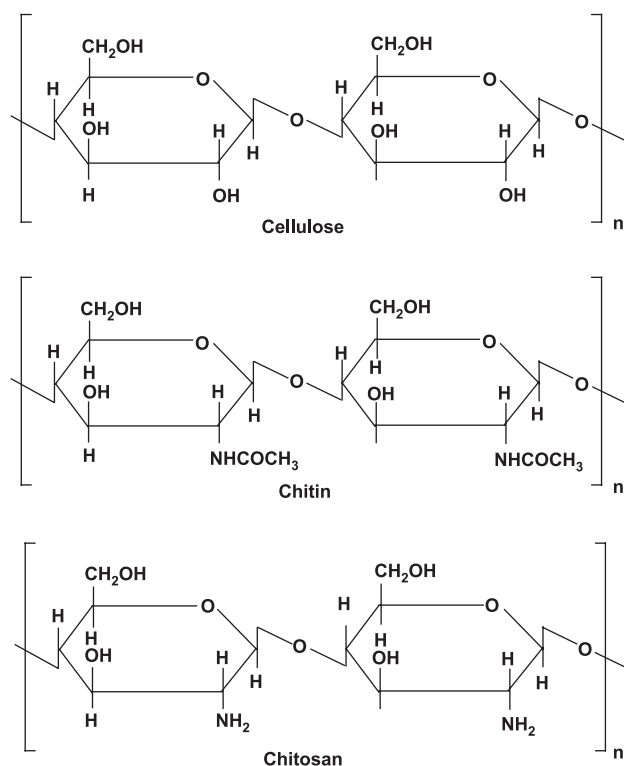


Fig. 1 Structural difference between cellulose, chitin, and chitosan.

released after 75 hours, showing the sustained-release nature of the matrix.^[27]

Chenite and coworkers^[28–31] reported novel thermosensitive chitosan gels possessing a physiological pH that can be held liquid below room temperature for encapsulating living cells and therapeutic proteins; they form monolithic gels at body temperature. The first report discusses the transformation of pH-gelling cationic polysaccharide solutions into thermally sensitive pH-dependent gel-forming aqueous solutions, without

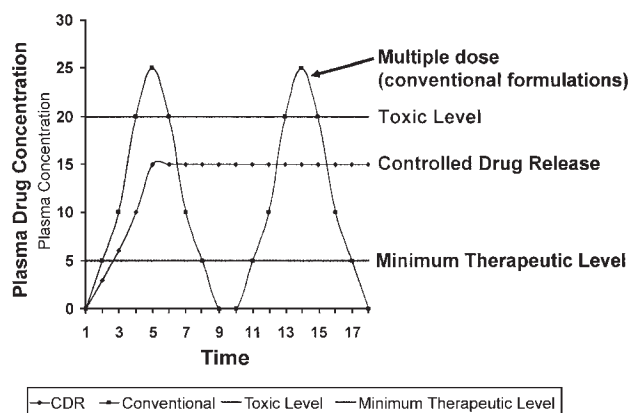


Fig. 2 Schematic representation of conventional versus controlled drug release formulations.

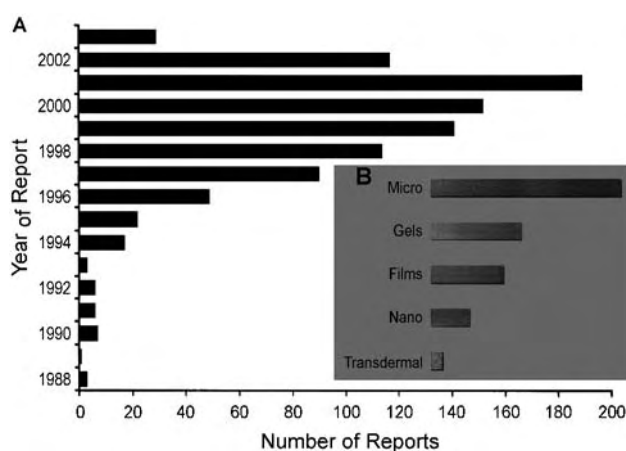


Fig. 3 (A) Chitosan drug delivery; (B) various chitosan formulations in drug delivery. (From Science Finder, April, 2003.)

any chemical modification or cross-link.^[28] The gel formation resulted from the addition of polysalts bearing a single anionic head, such as glycerol-, sorbitol-, fructose-, or glucose-phosphate salts (polyol- or sugar-phosphate).^[28] Chenite et al. found the chitosan/glycerol phosphate (C/GP) gels well suited to the delivery of sensitive biological materials such as proteins and living cells. Subcutaneous injections of gels formulated with a bone-inducing growth factor (BP) demonstrate that C/GP gels can deliver active BP, leading to de novo cartilage and bone formation in an ectopic site.^[28] Subcutaneous gel implants had beanlike shape and remained morphologically intact when excised at 3, 6, 12, 24, 48, and 72 hours and at 7 days.^[28] Studies were also performed by encapsulating various cell lines, in an attempt to demonstrate the cytocompatibility of C/GP gels. These studies found cell viability higher than 80% over extended periods of time when cultured in vitro.^[28] Subsequent studies analyzed the gels for the sustained delivery of drugs;^[29] for rheologic concerns,^[30] and for the effect of steam sterilization on thermo-gelling.^[31] Molinaro et al.^[32] evaluated the biocompatibility of the four chitosan/GP solutions in vivo. From these studies it was concluded that a higher degree of deacetylation of the chitin chain is desirable for superior biocompatibility.^[32] Ruel-Garièpy et al.^[33] reported on thermosensitive chitosan-based hydrogels containing liposomes for the delivery of hydrophilic molecules. The system thus developed is capable of releasing macromolecules over a period of several hours to few days in a controlled manner; however, with low-molecular weight lipophilic compounds, the release is completed within 24 hours.^[33] Chenite et al. have proposed a mathematical model to characterize the kinetics of release.^[30] Martin et al.^[34,35] investigated palmitoyl glycol chitosan hydrogels for the delivery of macromolecules.

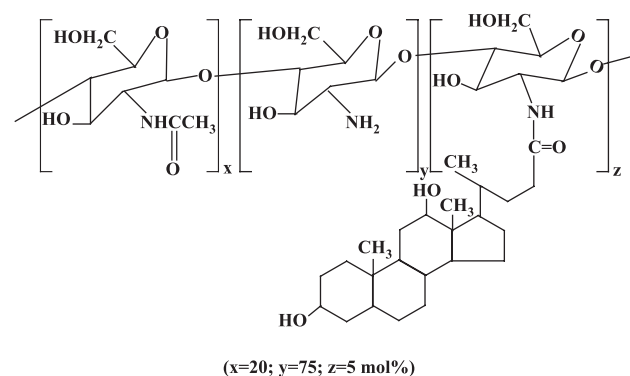


Fig. 4 Chemical structure of deoxycholic acid-modified chitosan.

Lee et al.^[36] reported self-aggregates of deoxycholic acid-modified chitosan as a novel carrier of adriamycin, where deoxycholic acid was covalently conjugated to chitosan via EDC-mediated reaction to generate self-aggregated chitosan nanoparticles (Fig. 4). About 49.6 wt% of adriamycin (ADR) was physically entrapped inside the self-aggregates, and the characteristics of the self-aggregates were analyzed by dynamic light scattering, fluorescence spectroscopy, and atomic force microscopy (AFM). Drug release studies were performed in PBS solution pH 7.2.^[36]

Alonso and coworkers^[37,38] reported chitosan nanoparticles for controlled release of doxorubicin and cyclosporin A. The entrapment of a cationic hydrophilic molecule (doxorubicin) was achieved by masking the positive charge of doxorubicin by complexing it with the polyanion-dextran sulfate. This modification resulted in efficient encapsulation of doxorubicin in comparison to controls.^[37] They have also investigated the possibility of forming complexes between chitosan and doxorubicin prior to the formation of nanoparticles, which resulted in low complexation efficiency and indicated no dissociation of the complex upon formation of the nanoparticles.^[37] Cell culture studies showed that those containing dextran sulfate were able to maintain cytostatic activity relative to free doxorubicin-loaded nanoparticles, whereas doxorubicin complexed to chitosan before nanoparticle formation showed slightly decreased activity. The foregoing studies suggest that the drug-bound nanoparticles entered the cells before release occurred in the cells.^[37] Another study reports ocular drug delivery.^[38] The preparation technique (based on an ionic gelation process) is extremely mild and involves the mixture of two aqueous phases at room temperature. The formation of the nanoparticles resulted from the interaction between negative groups of the sodium tripolyphosphate (TPP) and the positively charged amino groups of chitosan.^[38] These nanoparticles had a mean size of 293 nm, a zeta potential of +37 mV,

and high cyclosporin-association efficiency and loading (73 and 9%, respectively). The in-vitro studies revealed a fast release during the first hour followed by a constant release over a period of 24 hours.^[38] In-vivo experiments in rabbits following topical instillation of cyclosporin-loaded nanoparticles led to therapeutic concentrations in external ocular tissues for 48 hours, while maintaining negligible or undetectable cyclosporin levels in inner ocular structures, blood, and plasma. The drug levels were found to be significantly higher than those achieved by cyclosporin suspension or in chitosan solution containing cyclosporin.^[38] Felt et al. also studied the delivery of antibiotics to the eye using chitosan.^[39] The influence of the molecular weight, degree of deacetylation, and concentration of 4 different samples of chitosan on pharmacokinetic parameters of antibiotics in tears was investigated over time. Felt et al. observed that the chitosan products of high molecular weight and low deacetylation degree significantly increased the antibiotic availability.^[39] Ko et al.^[40] investigated preparation and characterization of chitosan microparticles for controlled drug delivery following a procedure similar to the foregoing. Ionic cross-linking of chitosan with TPP led them to microparticles within the range 500–710 μm with encapsulation efficiencies higher than 90%. Other methods for preparation of chitosan particles through ionic gelation have been investigated as well for slow release of drugs.^[41] Gadolinium-loaded chitosan nanoparticles (Gd-nanoCPs) were prepared by the novel emulsion-droplet coalescence technique.^[42] In-vitro cellular accumulation of gadolinium incorporated into chitosan nanoparticles was studied in L929 fibroblast cells.

Banerjee et al.^[43] reported a procedure for preparation of chitosan nanoparticles cross-linked with glutaraldehyde and also studied the biodistribution of nanoparticles thus prepared in rabbits. They found that the nanoparticles measure 30 nm in length when 10% of the amine groups are cross-linked, and 110 nm in length when 100% are cross-linked. When intravenously injected in mice, these cross-linked particles appeared in blood for a considerable period of time. The administration of ^{99m}Techneium-tagged nanoparticles in rabbits also confirms the presence of the administered particles in blood, and distributed in the heart, liver, kidneys, bladder, and vertebral column. Banerjee et al. also found these particles distributed in the bone marrow, raising the possibility that these particles may be used for bone imaging and targeting.^[43]

Junginger and coworkers^[44,45] extensively investigated chitosan and its derivatives as intestinal absorption enhancers. They synthesized trimethyl chitosan chloride (TMC) at different degrees of quaternization and demonstrated its activity as absorption enhancer at neutral pH values similar to those found in the intestinal tract.^[46–49] This quaternized polymer

forms complexes with anionic macromolecules and gels or solutions with cationic or neutral compounds in aqueous environments and neutral pH values. Trimethyl chitosan chloride has been shown to considerably increase the permeation of neutral and cationic peptide analogues across Caco-2 intestinal epithelia. The mechanism by which TMC enhances intestinal permeability is similar to that of protonated chitosan. It reversibly interacts with components of the tight junctions, leading to widening of the paracellular routes. This chitosan derivative does not provoke damage of the cell membrane, and does not alter the viability of intestinal epithelial cells. Coadministration of TMC with peptide drugs was found to substantially increase the bioavailability of the peptide in both rats and juvenile pigs compared with administration without the polymer.^[46–49]

Vandenberg et al.^[50] carried out a series of experiments to evaluate the influence of a number of physicochemical factors on the diffusion of a model protein—bovine serum albumin (BSA)—from dried chitosan-coated alginate microcapsules. The factors tested include alginate and chitosan concentration, calcium chloride concentration in the gelation medium, loading rate, chitosan molecular mass, and pH of the gelation medium. Alginate and chitosan concentrations were found to have influence on BSA retention during microcapsule manufacture and acid incubation, whereas calcium chloride had influence during the gelation process. Bovine serum albumin retention during manufacture was not significantly altered by the protein-loading rate or pH of the encapsulation medium. However, protein retention during acid incubation decreased significantly with increasing protein-loading rate and encapsulation medium pH. Vandenberg et al. also investigated the influence of the post-manufacture washing step, and found that microcapsules washed with acetone significantly increased protein retention during acid incubation.^[50] Continuing along a similar line of research, Huang et al.^[51] investigated formulation factors in preparing betamethasone (BTM)-chitosan microspheres with a spray-drying method using gelatin-A and ethylene oxide-propylene oxide block copolymer as modifiers. They found that BTM-loaded microspheres demonstrated good drug stability, high entrapment efficiency, and a positive surface charge. They also observed that in-vitro drug release is dependent on gelatin content.^[51] Gentamicin-loaded microparticles were prepared using chitosan hydroglutamate (CH), hyaluronic acid (HA), and a combination of both polymers by a solvent evaporation technique.^[52] In-vivo evaluation of these newly developed particles for nasal delivery of gentamicin was carried out in rabbits. The bioavailability of the drug was found to increase significantly when administered as microparticles rather than free drug

administered as 1.1% solution of 2.1% dry powder.^[52] Mi et al.^[53] investigated cross-linked chitosan microspheres as injectable implants. The chitosan microspheres were prepared by spray-drying and were further cross-linked by glutaraldehyde and genipin, where glutaraldehyde cross-linked microspheres served as a control. When the in-vivo biocompatibility and degradability of these microsphere implants was tested, intramuscular injections into the skeletal muscle of a rat showed less inflammatory reaction in the case of genipin cross-linked than did its control glutaraldehyde cross-linked microspheres. Mi et al. also found the glutaraldehyde cross-linked microspheres degraded within 12 weeks, whereas the genipin microspheres stayed intact after 20 weeks, as observed by SEM experiments.^[53] The structural differences between glutaraldehyde cross-linked and genipin chitosan microspheres are shown in Fig. 5.

Physical mixtures of sodium diclofenac (SD)-chitosan salts have been investigated.^[54] Chitosan salts were obtained by dissolving chitosan in solutions containing aspartic, glutamic, hydrochloric, lactic, and citric acids. The salts were further collected as dry powders by simple spray-drying and were then mixed with SD. How different types of acid influence the release of SD from the physical mixture was studied during gastrointestinal transit. Glutamic and aspartic salts were found to provide the best release.^[54]

GENE THERAPY

Applications that use various lipids and polymers to deliver DNA have been the subject of intense examination for more than a decade. Polymers display striking advantages as gene delivery carriers. Furthermore, they can be specifically tailored for a proposed application by choosing appropriate molecular weights, by cell coupling, by tissue-specific targeting moiety, or by performing other modifications that confer specific physiological or physicochemical properties. In this respect, chitosan enjoys a unique place in the pharmaceutical arena due to its versatile properties. It is being investigated as a potential nonviral carrier for gene delivery (Fig. 6). Depending on its molecular weight and degree of deacetylation, chitosan is capable of forming small and stable complexes with plasmids. Subsequent paragraphs discuss recent reviews of chitosan research^[18,19] and chitosan's role in gene delivery.

MacLaughlin et al.^[55] studied chitosan and depolymerized chitosan oligomers as condensing carriers for in-vivo plasmid delivery. To form complexes with plasmid, each chitosan oligomer or polymer (2% w/v) was dissolved in acetic acid by sonication. The final concentration of the solution was brought to 0.4% w/v



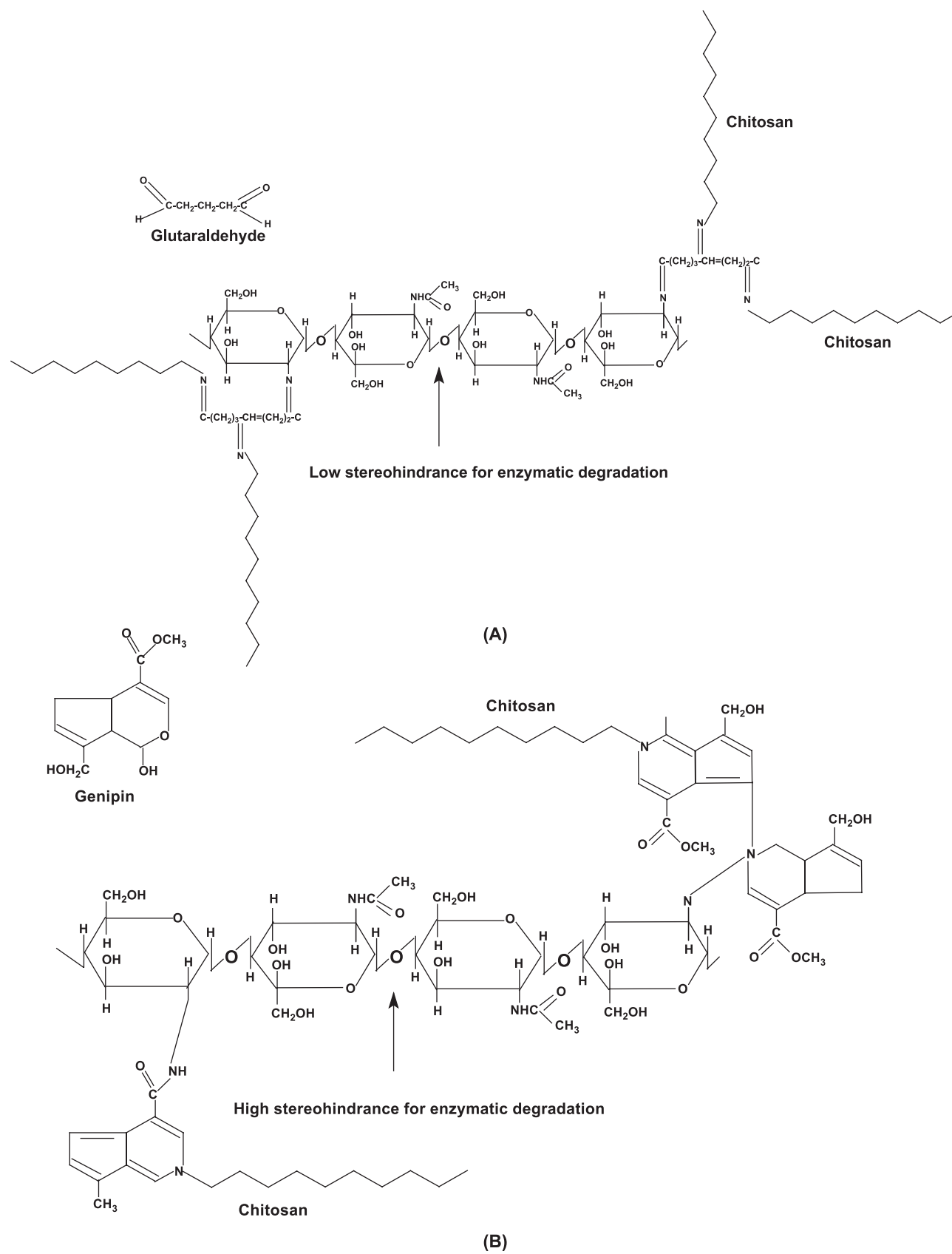


Fig. 5 Structural differences between glutaraldehyde and genipin cross-linked chitosan microspheres.

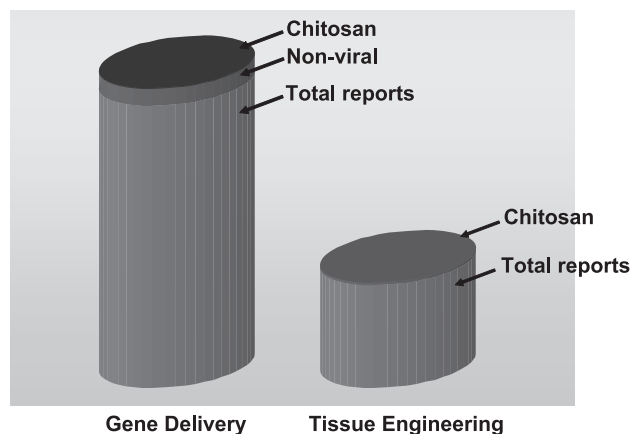


Fig. 6 Chitosan in gene delivery and tissue engineering (From Science Finder, April, 2003.)

by adding water and the solution was sterile-filtered for removal of particles greater than 200 nm. This chitosan solution was then added to an aqueous suspension containing 25–400 mg of plasmid in a total volume of 1 ml by gentle pipetting to form nanocomplexes of a selected charge ratio.^[55] The solution was vortexed rapidly for 3–5 min and left for 30 min at room temperature to ensure complete complexation.^[55] The in-vitro transfection was tested in Cos-1 cells and the in-vivo expression was tested in the intestinal tissues. In vivo, higher levels of expression were measured with the plasmid/chitosan/GM225.1 formulation than with naked plasmid in the upper small intestine. The overall expression levels achieved using the DOTMA:DOPE-based formulation were lower than those achieved with plasmid/chitosan/GM225.1 complexes, and there were fewer tissue extracts showing positive chloramphenicol acetyltransferase (CAT) expression (Table 1). It is interesting to note that the in-vitro results were reversed. The parameters influencing particle size and stability include chitosan molecular weight, plasmid concentration, and charge ratio. Plasmid/chitosan complexes made of higher-molecular weight chitosan were more stable to salt and serum challenge. Complexes of a 1:2 (–/+) charge ratio were shown to be most stable. In vitro, the highest level of expression in the absence of serum was obtained using a 1:2 (–/+) complex made with 102 KDa chitosan, which was approximately 250 times less than the level of expression observed with Lipofectamine™, a positive control. Particle size was found to not influence expression, although inclusion of the pH-sensitive endosomolytic peptide GM227.3 in the formulation enhanced levels of expression with the 1:2:0.25 (–/+ /–) complex (200 µg/mL), with very low expression levels (100-fold less) compared to Lipofectamine™. Cui and Mumper^[56] used chitosan-based

nanoparticles for topical genetic immunization. They investigated two different formulations (viz., pDNA-condensed chitosan nanoparticles and pDNA coated on preformed cationic chitosan/carboxymethylcellulose (CMC) nanoparticles). From these studies it is evident that both chitosan and chitosan oligomer can complex CMC to form stable cationic nanoparticles and thereby attain pDNA coating. Topical application of the various chitosan-based formulations under investigation resulted in both detectable and quantifiable levels of luciferase expression in mouse skin after 24 hours.^[56]

Thorough investigation of chitosan as an efficient gene transfection agent in vitro and in vivo by Leong and coworkers^[57–61] resulted in a series of reports describing various modifications. They investigated important parameters for the preparation of nanoparticles and characterized the physicochemical properties of the system. Protection for encapsulated DNA by chitosan particles was confirmed.^[60] Leong et al. investigated the effects of coencapsulating a lysosomolytic agent—chloroquine—on transfection efficiency.^[60] They also proposed plausible schemes (shown in Fig. 7) for transferrin and KNOB protein conjugation in an attempt to improve the surface property and result in improved transfection efficiency.^[60] Furthermore, the chitosan–DNA particles were PEGylated (shown in Fig. 8) to improve their storage stability and yield a formulation that can be lyophilized without loss of transfection ability.^[60] Tissue distribution of the chitosan–DNA nanoparticles and PEGylated nanoparticles following intravenous administration was investigated.^[60] It was found that clearance of PEGylated nanoparticles was slower than clearance of unmodified particles at 15 mins. Higher depositions in the kidney and liver, however, did not produce any difference after 1 hour.^[60] It was observed that transfection efficiency was cell-type dependent and 3–4 times higher than the background level in HEK293 cells, and 2–10 times lower than achieved by Lipofectamine™.^[60] Mohapatra and coworkers^[61] used these chitosan nanoparticles^[60] to demonstrate their protection factor in acute respiratory syncytial virus infection in BALB/c mice following intranasal gene transfer. Illum et al.^[62] reviewed chitosan as a novel nasal delivery system for vaccines.

Hoggard et al.^[63] used chitosan nanoparticles of different molecular-weight chitosan (Table 2) for intratracheal administration in mice and compared the results with polyethyleneimine in vitro. After intratracheal administration, they observed both polyplexes distributed to the mid-airways, where transgene expression was observed in virtually every epithelial cell using a sensitive pLacZ reporter containing a translational enhancer element.^[63] Hoggard et al. observed a similar result, where PEI polyplexes induced more



Table 1 CAT expression in tissue extracts of rabbits dosed in the upper small intestine or colon with plasmid/chitosan complexes

Formulation	Region	Rabbit number				CAT expression [mean \pm S.D. (pg/mg)]
		#1	#2	#3	#4	
Dosed in upper small intestine	PP1	0	0	0	0	0 \pm 0
	PP2	0	0	0	0	0 \pm 0
Plasmid/CT/GM225.1 (<i>n</i> = 4)	PP3	6.12	5.01	0	5.83	4.24 \pm 2.87
	ENT1	1.45	0	8.10	6.62	4.04 \pm 3.92
	ENT2	0	4.50	0	9.60	3.52 \pm 4.57
	ENT3	6.16	0	11.41	4.32	5.47 \pm 4.72
	Col	0	0	0	0	–
	MLN	0	0	3.68	10.96	3.66 \pm 5.17
Plasmid in 10% lactose (<i>n</i> = 2)	PP3	0	0	–	–	0 \pm 0
	ENT1	0	0	–	–	0 \pm 0
	MLN	0	0	–	–	0 \pm 0
Plasmid/DOTMA:DOPE (<i>n</i> = 4)	PP1	0	0	0	6.75	1.69 \pm 3.37
	PP2	0	0	0	0	0 \pm 0
	PP3	0	8.85	0	6.83	3.9 \pm 2.46
	ENT1	0	0	0	10.13	2.53 \pm 5.07
	ENT2	0	0	0	0	0 \pm 0
	ENT3	0	0	0	0	0 \pm 0
	Col	0	0	0	4.74	1.19 \pm 2.37
	MLN	0	0	0	4.60	1.15 \pm 2.30
Dosed in the colon						
Plasmid/CT/GM225.1 (<i>n</i> = 4)	Col1	0	0	0	6.30	1.58 \pm 3.20
	Col2	0	0	7.47	6.30	3.57 \pm 4.13
	Col3	0	0	0	0	0 \pm 0
	MLN	5.97	10.11	12.14	0	7.06 \pm 5.36
Plasmid in 10% lactose (<i>n</i> = 2)	PP3	0	0	–	–	0 \pm 0
	ENT1	0	0	–	–	0 \pm 0
	MLN	0	0	–	–	0 \pm 0
Plasmid/DOTMA:DOPE (<i>n</i> = 4)	Col1	0	0	0	0	0 \pm 0
	Col2	0	0	0	0	0 \pm 0
	Col3	0	0	0	0	0 \pm 0
	MLN	0	0	0	0	0 \pm 0

rapid onset of gene expression than did chitosan.^[63] Corsi et al.^[60] evaluated the potential of chitosan–DNA nanoparticles for gene transfer on human mesenchymal stem cells (MSCs), human osteosarcoma cells (MG63), and human embryonic kidney cells (HEK293). HEK293 cells were found to be the most transfected cells, compared to MSCs and MG63.^[64]

In their studies, Sato and coworkers^[65] optimized transfection conditions for chitosan-mediated gene delivery. The transfection studies were carried out using a tumor cell line (Human-lung Carcinoma A549

cells, HeLa cells, and B16 melanoma cells 1×10^5 cells/well), and indicated that the transfection efficiency of chitosan complexes is dependent on the pH of culture medium, the stoichiometry of pGL3:chitosan, and serum and molecular mass of chitosan.^[65] Transfection efficiency at pH 6.9 was higher than that at pH 7.6. The optimum charge ratio of the pGL3:chitosan was 1:5. Chitosan polymer of 15 and 52 kDa largely promoted luciferase activities. Transfection efficiency mediated by chitosan of >100 kDa was less than that by chitosan of 15 and 52 kDa.^[65]

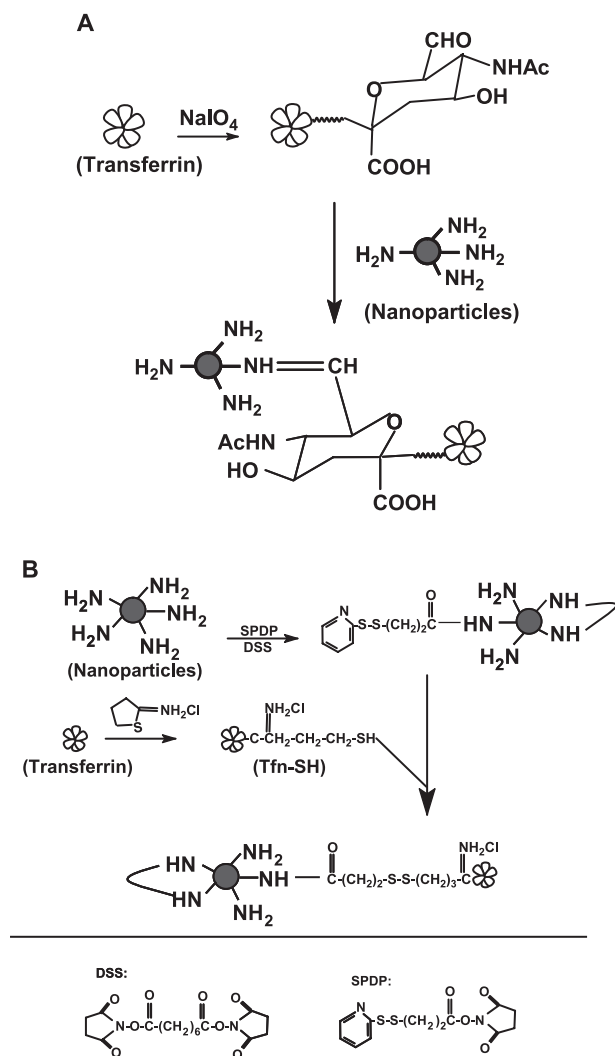


Fig. 7 (A) Conjugation of transferrin through periodate oxidation; (B) conjugation of transferrin through a reversible disulfide linkage.

Kim et al.^[66] reported self-aggregates of deoxycholic acid-modified chitosan as DNA carriers. In order to control the size of self-aggregates, chitosan was depolymerized with various amounts of sodium nitrite, and hydrophobically modified with deoxycholic acid to form self-aggregates in aqueous media, shown in Fig. 4. They observed that the size of the self-aggregates varied in the range of 130–300 nm in diameter, where the chitosan molecular mass has influence on the sizes.

Thanou et al.^[67] reported quaternized chitosan oligomers as novel gene delivery vectors in epithelial cell lines. They synthesized trimethylated chitosan (TMO) using oligomeric chitosan (<20 monomer units). They observed spontaneous complex formation of TMOs with RSV- α 3 luciferase plasmid DNA ranging in size from 200 to 500 nm.^[67] They found that, when tested in Cos-1, the transfection efficiencies of

chitoplexes are lower than those of DOTAP (N-(1-(2, 3-dioleoyloxy)propyl)-N,N,N-trimethylammonium sulphate)-DNA lipoplexes, although both these complexes showed lower transfection in Caco2 cells, with TMOs taking the lead over DOTAP.^[67]

Galactosylated chitosan-graft-dextran (GCD) was prepared and characterized for Hepatocyte-targeting and complexed with DNA by Park and coworkers.^[68] Lactobionic acid-bearing galactose group was coupled with chitosan for liver specificity, and dextran was grafted to galactosylated chitosan (GC) for stability in water. The structures of GCD and GC are as shown in Fig. 9. Park et al. found the GCD/DNA complex to be more stable compared to the GC/DNA complex, and found that the particle size of the GCD/DNA complex becomes smaller as the charge ratio of GCD to DNA increased. They observed transfection of GCD/DNA complexes only in Chang Liver cells. They also found the transfection of Hep G2 having asialoglycoprotein receptors (ASGR), indicating the specific interaction of cells and galactose ligands on chitosan.^[68]

Okamoto et al.^[69] adopted a supercritical carbon dioxide technique to make chitosan-plasmid complex powders, which they applied as a pulmonary gene delivery system in vivo. The luciferase activity in mouse lung was evaluated after pulmonary administration of the powders or pCMV-Luc solutions. The chitosan-pDNA powders increased the luciferase activity in mouse lung compared with pCMV-Luc powders without chitosan or pCMV-Luc solutions with or without chitosan. Okamoto et al. found a drastic increase of luciferase activity (up to 2700%) when chitosan-pDNA powder with an N/P ratio = 5 in comparison to that of the pCMV-Luc solution.

TISSUE ENGINEERING

The tissue-engineering approach to repair and regeneration is based on the use of polymer scaffolds that serve to support, reinforce, and in some cases organize the regenerating tissue.^[70–72] The scaffold may be required to release bioactive substances at a controlled rate or to directly influence the behavior of incorporated or ingrowing cells. A number of natural and synthetic polymers are currently employed as tissue scaffolds.^[73–75] The microstructures of these systems range from hydrogels, to open-pore structures, to fibrous matrices. Because the range of potential tissue-engineered systems is broad, there is a continuous ongoing search for materials which either possess particularly desirable tissue-specific properties, or which may have broad applicability and can be tailored to several tissue systems. Chitosan finds application in tissue engineering because of its biocompatibility and capacity to deliver drugs at a



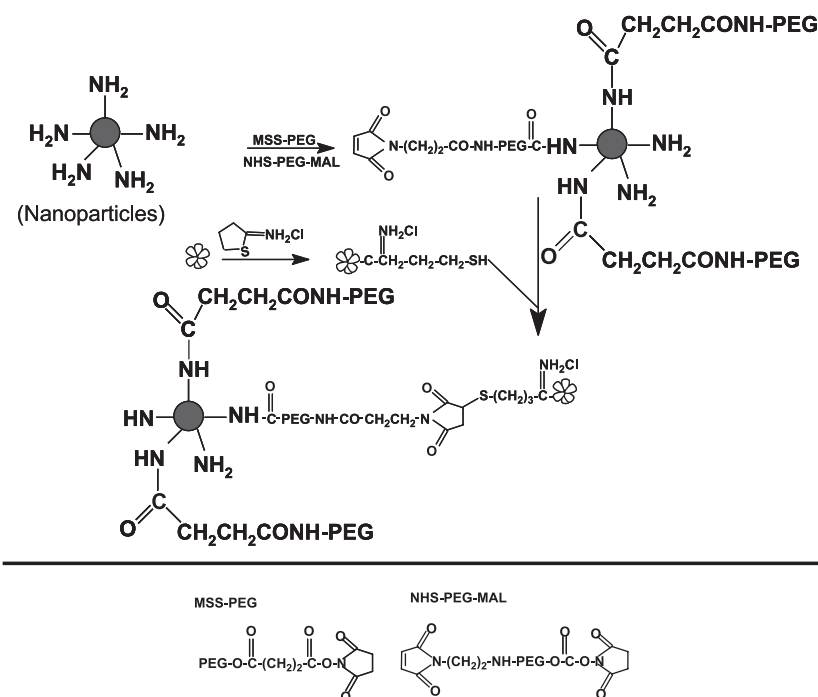


Fig. 8 PEGylation of chitosan-DNA nanoparticles and conjugation of transferring through a PEG spacer.

controlled rate. From a survey of the literature, it is evident that chitosan forms approximately one of every 50 reports on tissue engineering. (See Fig. 6.) One of the most promising features of chitosan is its excellent ability to be processed into porous structures for use in cell transplantation and tissue regeneration. Francis Suh and Matthew^[76] recently reviewed the application

of chitosan-based polysaccharide biomaterials in cartilage tissue engineering.

Chitosan has been thoroughly investigated in tissue engineering; several researchers have examined the response of tissue to various chitosan-based implants.^[77,78] These materials have generally been found to evoke minimal foreign body reaction. In most cases, no major fibrous encapsulation has been observed.^[79] Porous chitosan structures can easily be formed by freezing and lyophilizing chitosan-acetic acid solutions in suitable molds.^[80] The mechanical properties of chitosan scaffolds depend mainly on pore sizes and orientation.

Zhang and Zhang^[81] synthesized and characterized macroporous chitosan scaffolds reinforced by β -tricalcium phosphate (β -TCP) and calcium phosphate invert glass by a freeze-drying technique via thermally induced phase separation for tissue-engineering applications. They observed the composite scaffolds to be porous and the pore structures of the scaffolds with β -TCP and the glass to be very different. The bioactivity tests showed a continuous decrease in both Ca and P concentrations of a simulated body fluid (SBF) after the scaffolds with β -TCP were immersed in the SBF for more than 20 hours, suggesting the formation of an apatite layer on the scaffolds (which was not the case with pure chitosan scaffolds or scaffolds incorporated with glass due to the absence of nucleation sites on scaffold surfaces).^[81] There are, however, positive reports of chitosan scaffolds used alone without any

Table 2 Transfection efficiency in 292 cells and particle sizes of different chitosan/pDNA complexes at their optimal charge ratios

Chitosan ^c	Charge ratio ^a (+/-)	Gene expression ^b (pg/ μ g protein)	Particle size (nm)
C(1;31)	3.6:1	6.0 \pm 2.8	131 \pm 9
C(1;170)	3.6:1	5.0 \pm 2.0	174 \pm 23
C(15;190)	3.0:1	7.2 \pm 1.5	144 \pm 12
C(35;170)	3.0:1	0.2 \pm 0.04	195 \pm 15
C(49;98)	3.6:1	0.1 \pm 0.04	229 \pm 2

^aCharge ratios covering the range 0.6:1–4.2:1 (+/-) were investigated.

^bCells were analyzed for CAT gene expression 48 h after transfection. Data are expressed as mean values \pm s.d. from one representative experiment ($n=4$) of three performed.

^cChitosan represented by previously published nomenclature. The numbers in parentheses are degree of deacetylation expressed in %, followed by the molecular weight in kDa.

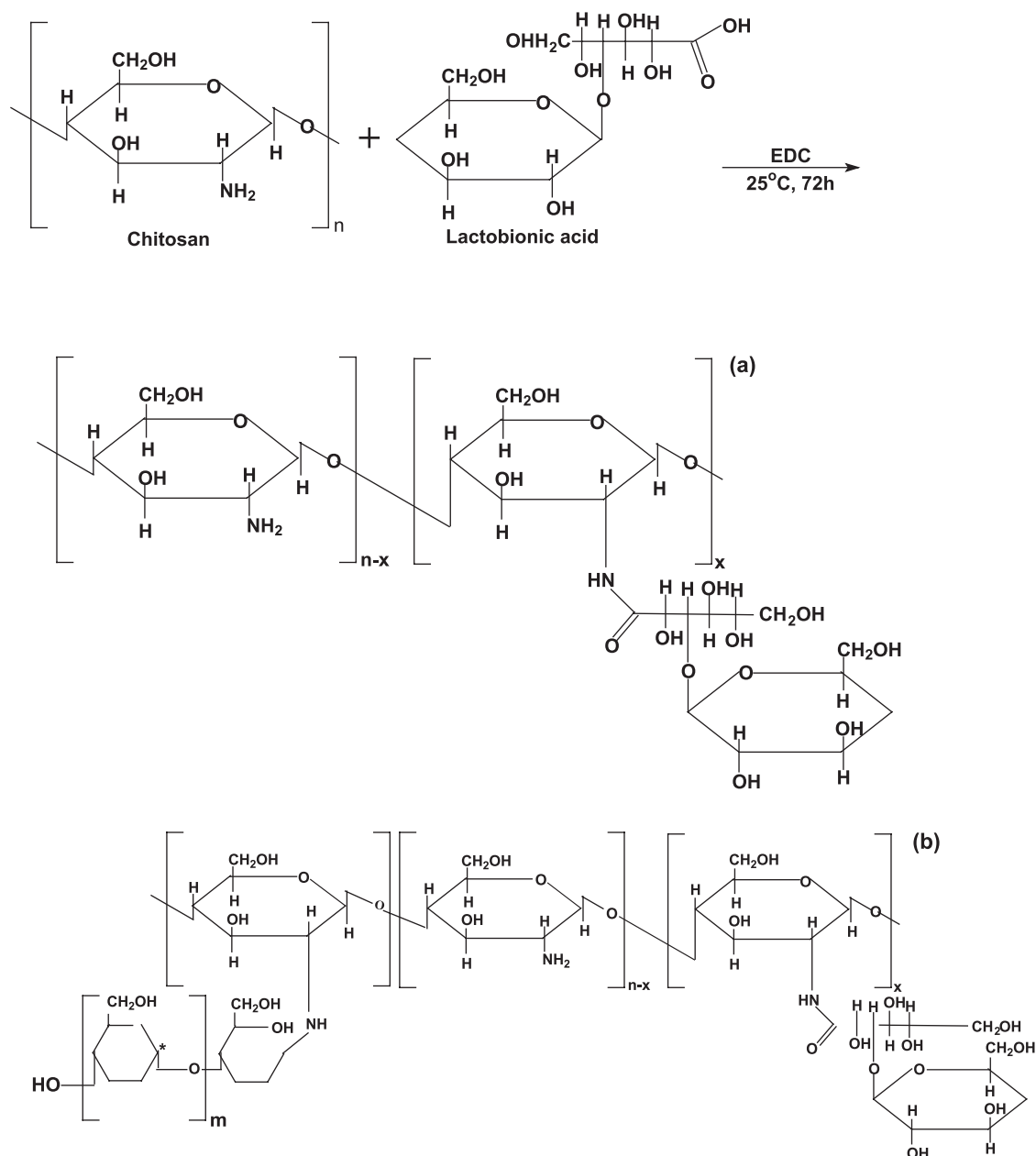


Fig. 9 Chemical structures of (a) galactosylated chitosan (GC); and (b) galactosylated chitosan-graft-chitosan (GCD).

modification.^[82] Lee and coworkers^[83] studied bone formation by controlled growth factor delivery from chitosan, chitosan-poly(L-lactide) (PLLA) matrices, and chitosan coated on PLLA matrices.

FOOD TECHNOLOGY

Chitin and chitosan have been materials of interest over the past decades due to the broad range of their potential applications, as discussed in the foregoing. However, only limited attention has been paid to the food application of these versatile biopolymers. The

conversion of processing discards into edible by-products and alternative specialties is a timely challenge for food research and development. Chitosan finds application in various sectors of the food industry, as shown in Table 3. Shahidi et al.^[84] reviewed the food application of chitin and chitosan.

A most intently investigated food industry application of chitosan stems from its antimicrobial activity. The growing consumer demand for foods without chemical preservatives has led to the discovery of chitosan's antimicrobial properties.^[85] The cationic nature of chitosan is expected to influence the mechanism of the antimicrobial activity through binding



Table 3 Applications of chitinous materials in food technology

Area of application	Examples
Antimicrobial agent	Bactericidal Fungicidal Measure of mold contamination in agricultural commodities
Edible films	Controlled moisture transfer between food and surrounding environment Controlled release of antimicrobial substances Controlled release of nutrients, flavors, and drugs Reduction of oxygen partial pressure Controlled rate of respiration Temperature control Controlled enzymatic browning in fruits Reverse osmosis membranes
Additive	Clarification and deacidification of fruits and beverages Natural flavor extender Texture controlling agent Emulsifying agent Food mimetic Thickening and stabilizing agent Color stabilization
Nutritional quality	Dietary fiber Hypocholesterolemic effect Livestock and fish food additive Reduction of lipid absorption Production of single cell protein Antigastritis agent Infant feed ingredient
Recovery of solid materials from food processing waste	Affinity flocculation Fractionation of agar
Purification of water	Recovery of metal ions, pesticides, phenols, and PCBs Removal of dyes

sialic acid in negatively-charged phospholipids, thus restraining the movement of microbiological substances.^[84] Oligomeric chitosan penetrates into the cells of microorganisms and prevents the growth of cells by prohibiting the transformation of DNA into RNA.^[86] Jung et al.^[87] synthesized amphiphilic chitosan copolymers by grafting two anionic soluble monomers,

mono(2-methacryloyl oxyethyl) acid phosphate and vinylsulfonic acid sodium salt. The antimicrobial properties of these copolymers were tested and indicated improved activity. The antimicrobial activity was found to depend largely on the amount and type of grafted chains and on changes of pH against *Candida albicans*, *Trichophyton rubrum*, and *Trichophyton violaceum*. Enhanced activity was determined at pH 5.75 with strain selectivity against *Candida albicans* and *Trichophyton violaceum* due to a difference in affinity between the cell walls of fungi and chitosan or its derivatives.^[87]

Jumaa et al.^[88] investigated the antimicrobial activity of chitosan in lipid emulsion as well as in aqueous solutions using different deacetylated chitosan (92% and 73%). They studied the influence of pH on antimicrobial activity by incubating the microorganisms at different pH in lactic acid solution (1% w/v) containing different concentrations of chitosan up to 1.5% (w/v). Both types of chitosan were pH-dependent; optimal activity was found at pH 5 for *P. aeruginosa* and pH 5.3 for *S. aureus*. This study revealed that lipid emulsions containing 0.5% (w/v) chitosan displayed antimycotic efficacy, conforming to the requirements of the preservation efficacy test for topical formulations.^[88]

CONCLUSION

From the foregoing sections it is evident that chitosan is a material of choice for versatile applications. The problem of chitosan's ineffectiveness at neutral pH can be dealt with by derivatization to render the polymer soluble. Although much research and development have been accomplished during the past few decades, untiring efforts are still in progress to develop increasingly versatile applications.

ARTICLES OF FURTHER INTEREST

Bioadhesion; Burn Dressing; Drug Delivery, Controlled; Gene Therapy; Tissue Engineering Scaffolds

REFERENCES

1. Muzzarelli, R.A. *Chitin*; Pergamon Press: Oxford, UK, 1977.
2. Ravi Kumar, M.N.V. A review of chitin and chitosan applications. *React. Funct. Polym.* **2000**, *46*, 1–27.
3. Dutkiewicz, J.K. Superabsorbent materials from shellfish waste—A review. *J. Biomed. Mater. Res.* **2002**, *63* (3), 373–381.

4. *Applications of Chitin and Chitosan*; Goosen, M.F.A., Ed.; Technomic Publishing Co. Inc.: Lancaster, 1997; 1–336.
5. Skaugrud, O.; Hagen, A.; Borgersen, B.; Dornish, M. Biomedical applications of alginate and chitosan. *Biotechnol. Genet. Eng. Rev.* **1999**, *16*, 23–40.
6. Hirano, S. Chitin biotechnology applications. *Biotechnol. Annu. Rev.* **1996**, *2*, 237–258.
7. Singla, A.K.; Chawla, M. Chitosan: Some pharmaceutical and biological aspects—An update. *J. Pharm. Pharmacol.* **2001**, *53* (8), 1047–1067.
8. Genta, I.; Perugini, P.; Pavanetto, F.; Modena, T.; Muzzarelli, R.A.A.; Conti, B. Chitosan Crosslinking for Pharmaceutical Purposes: An Overview. In *Chitosan in Pharmacy and Chemistry*; Muzzarelli, R.A.A., Muzzarelli, C., Eds.; Atec, Italy, 2002; 11–20.
9. Muzzarelli, R.; Biagini, G.; Pugnali, A.; Filippini, O.; Baldassarre, V.; Castaldini, C.; Rizzoli, C. Reconstruction of parodontal tissue with chitosan. *Biomaterials* **1989**, *53* (2), 131–140.
10. Rathke, T.D.; Hudson, S.M. Review of chitin and chitosan as fibre and film formers. *JMS Rev. Macromol. Chem.* **1994**, *C34*, 375.
11. Kurita, K. Chemical Modifications of Chitin and Chitosan. In *Chitin in Nature and Technology*; Muzzarelli, R.A.A., Jeuniaux, C., Gooday, G.W., Eds.; Plenum Press: New York, 1986; 287–293.
12. Dodane, V.; Vilivalam, V.D. Pharmaceutical applications of chitosan. *Pharm. Sci. Technol. Today* **1998**, *6*, 246–253.
13. Sawayanagi, Y.; Nambu, N.; Nagai, T. Directly compressed tablets containing chitin or chitosan in addition to lactose or potato starch. *Chem. Pharm. Bull.* **1982**, *30*, 2935–2940.
14. Ritthidej, G.C.; Chomto, P.; Pummangura, S.; Menasveta, P. Chitin and chitosan as disintegrants in paracetamol tablets. *Drug Dev. Ind. Pharm.* **1994**, *20*, 2109–2134.
15. Illum, L. Chitosan and its use as a pharmaceutical excipient. *Pharm. Res.* **1998**, *15* (9), 1326–1331.
16. Sawayanagi, Y.; Nambu, N.; Nagai, T. Enhancement of dissolution properties of griseofulvin from ground mixtures with chitin and chitosan. *Chem. Pharm. Bull.* **1982**, *30*, 4464–4467.
17. Dutta, P.D.; Ravi Kumar, M.N.K.; Dutta, J. Chitin and chitosan for versatile applications. *J. Macromol. Sci., Part C, Polym. Rev.* **2002**, *C42* (3), 307–354.
18. Liu, W.G.; Yao, K.D. Chitosan and its derivatives—A promising non-viral vector for gene transfection. *J. Control. Release* **2002**, *83*, 1–11.
19. Borchard, G. Chitosans for gene delivery. *Adv. Drug Deliv. Rev.* **2001**, *52*, 145–150.
20. Paul, W.; Sharma, C.P. Chitosan, a drug carrier for the 21st century: A review. *STP Pharma Sciences* **2000**, *10* (1), 5–22.
21. Felt, O.; Buri, P.; Gurny, R. Chitosan: A unique polysaccharide for drug delivery. *Drug Dev. Ind. Pharm.* **1998**, *24* (11), 979–993.
22. Bernkop-Schnurch, A. Chitosan and its derivatives: Potential excipients for peroral peptide delivery systems. *Int. J. Pharm.* **2000**, *194*, 1–13.
23. Cascone, M.G.; Maltinti, S. Hydrogels based on chitosan and dextran as potential drug delivery systems. *J. Mater. Sci., Mater. Med.* **1999**, *10* (5), 301–307.
24. Zhang, M.; Kohr, E.; Hirano, S. Hydrogels of Chitin and Chitosan. In *Food Hydrocolloids: Proc. Int. Conf. Ind. Exhib.*; Plenum: New York, 1993; 65–70.
25. Kas, H.K. Chitosan: Properties, preparations and application to microparticulate systems. *J. Microencapsul* **1997**, *14* (6), 689–711.
26. Janes, K.A.; Calvo, P.; Alonso, M.J. Polysaccharide colloidal particles as delivery systems for macromolecules. *Adv. Drug Deliv. Rev.* **2001**, *47*, 83–97.
27. Risbud, M.V.; Bhone, R.R. Polyacrylamide–chitosan hydrogels: In vitro biocompatibility and sustained antibiotic release studies. *Drug Deliv.* **2000**, *7*, 69–75.
28. Chenite, A.; Chaput, C.; Wang, D.; Combes, C.; Buschmann, M.D.; Hoemann, C.D.; Leroux, J.C.; Atkinson, B.L.; Binette, F.; Selmani, A. Novel injectable neutral solutions of chitosan form biodegradable gels in situ. *Biomaterials* **2000**, *21*, 2155–2161.
29. Ruel-Gariépy, E.; Chenite, A.; Chaput, C.; Guirguis, S.; Leroux, J.-C. Characterization of thermosensitive chitosan gels for the sustained delivery of drugs. *Int. J. Pharm.* **2000**, *203*, 89–98.
30. Chenite, A.; Buschmann, M.; Wang, D.; Chaput, C.; Kandani, N. Rheological characterization of thermogelling chitosan/glycerol–phosphate solutions. *Carbohydr. Polym.* **2001**, *46*, 39–47.
31. Jarry, C.; Chaput, C.; Chenite, A.; Renaud, M.-A.; Buschmann, M.; Leroux, J.-C. Effects of steam sterilization on thermogelling chitosan-based gels. *J. Biomed. Mater. Res.* **2001**, *58*, 127–135.
32. Molinaro, G.; Leroux, J.-C.; Damas, J.; Adam, A. Biocompatibility of thermosensitive chitosan-based hydrogels: An in vivo experimental approach to injectable biomaterials. *Biomaterials* **2002**, *23* (13), 2717–2722.
33. Ruel-Gariépy, E.; Leclair, G.; Hildgen, P.; Gupta, A.; Leroux, J.C. Thermosensitive chitosan-based hydrogel containing liposomes for the delivery of hydrophilic molecules. *J. Control. Release* **2002**, *82*, 373–383.
34. Martin, L.; Wilson, C.G.; Koosha, F.; Tetley, L.; Gray, A.I.; Senel, S.; Uchegbu, I.F. The release of model macromolecules may be controlled by the hydrophobicity of palmitoyl glycol chitosan hydrogels. *J. Control. Release* **2002**, *80*, 87–100.
35. Martin, L.; Wilson, C.G.; Koosha, F.; Uchegbu, I.F. Sustained buccal delivery of the hydrophobic drug denbufylline using physically cross-linked palmitoyl glycol chitosan hydrogels. *Eur. J. Pharm. Biopharm.* **2003**, *55*, 35–45.
36. Lee, K.Y.; Kim, J.-H.; Kwon, I.C.; Jeong, S.Y. Self-aggregates of deoxycholic acid-modified chitosan as a novel carrier of adriamycin. *Colloid Polym. Sci.* **2000**, *278*, 1216–1219.
37. Janes, K.A.; Fresneau, M.P.; Marazuela, A.; Fabra, A.; Alonso, M.J. Chitosan nanoparticles as delivery systems for doxorubicin. *J. Control. Release* **2001**, *73*, 255–267.
38. De Campos, A.M.; Sánchez, A.; Alonso, M.J. Chitosan nanoparticles: A new vehicle for the improvement of the



- delivery of drugs to the ocular surfaces application to cyclosporin A. *Int. J. Pharm.* **2001**, *224*, 159–168.
39. Felt, O.; Gurny, R.; Buri, P.; Baeyens, V. Delivery of antibiotics to the eye using a positively charged polysaccharide as vehicle. *AAPS Pharm. Sci.* **2001**, *3* (4), 1–7.
 40. Ko, J.A.; Park, H.J.; Hwang, S.J.; Park, J.B.; Lee, J.S. Preparation and characterization of chitosan microparticles intended for controlled drug delivery. *Int. J. Pharm.* **2002**, *249*, 165–174.
 41. González-Rodríguez, M.L.; Holgado, M.A.; Sánchez-Lafuente, C.; Rabasco, A.M.; Fini, A. Alginate/chitosan particulate systems for sodium diclofenac release. *Int. J. Pharm.* **2002**, *232*, 225–234.
 42. Shikata, F.; Tokumitsu, H.; Ichikawa, H.; Fukumori, Y. In vitro cellular accumulation of gadolinium incorporated into chitosan nanoparticles designed for neutron-capture therapy of cancer. *Eur. J. Pharm. Biopharm.* **2002**, *53*, 57–63.
 43. Banerjee, T.; Mitra, S.; Singh, A.K.; Sharma, R.K.; Maitra, A. Preparation, characterization and biodistribution of ultrafine chitosan nanoparticles. *Int. J. Pharm.* **2002**, *243*, 93–105.
 44. Thanou, M.; Verhoef, J.C.; Junginger, H.E. Chitosan and its derivatives as intestinal absorption enhancers. *Adv. Drug Deliv. Rev.* **2001**, *50*, S91–S101.
 45. van der Lubben, I.M.; Verhoef, J.C.; Borchard, G.; Junginger, H.E. Chitosan and its derivatives in mucosal drug and vaccine delivery. *Eur. J. Pharm. Sci.* **2001**, *14*, 201–207.
 46. Thanou, M.; Kotzé, A.F.; Scharringhausen, T.; Lueßen, H.L.; de Boer, A.G.; Verhoef, J.C.; Junginger, H.E. Effect of degree of quaternization of *N*-trimethyl chitosan chloride for enhanced transport of hydrophilic compounds across intestinal Caco-2 cell monolayers. *J. Control. Release* **2000**, *64*, 15–25.
 47. Thanou, M.; Florea, B.I.; Langemeijer, M.W.E.; Verhoef, J.C.; Junginger, H.E. *N*-trimethylated chitosan chloride (TMC) improves the intestinal permeation of the peptide drug busserelin in vitro (Caco-2 cells) and in vivo (rats). *Pharm. Res.* **2000**, *17*, 27–31.
 48. Thanou, M.; Verhoef, J.C.; Marbach, P.; Junginger, H.E. Intestinal absorption of octreotide: *N*-trimethyl chitosan (TMC) ameliorates the permeability and absorption properties of the somatostatin analogue in vitro and in vivo. *J. Pharm. Sci.* **2000**, *89*, 951–957.
 49. Thanou, M.; Verhoef, J.C.; Verheijden, J.H.M.; Junginger, H.E. Intestinal absorption of octreotide using *N*-trimethyl chitosan (TMC): Studies in pigs. *Pharm. Res.* **2001**, *18*, 823–828.
 50. Vandenberg, G.W.; Drolet, C.; Scott, S.L.; de la Noüe, J. Factors affecting protein release from alginate-chitosan coacervate microcapsules during production and gastric/intestinal simulation. *J. Control. Release* **2001**, *77*, 297–307.
 51. Huang, Y.C.; Yeh, M.K.; Chiang, C.H. Formulation factors in preparing BTM-chitosan microspheres by spray drying method. *Int. J. Pharm.* **2002**, *242*, 239–242.
 52. Lim, S.T.; Forbes, B.; Berry, D.J.; Martin, G.P.; Brown, M.B. In vivo evaluation of novel hyaluronan/chitosan microparticulate delivery systems for the nasal delivery of gentamicin in rabbits. *Int. J. Pharm.* **2002**, *231*, 73–82.
 53. Mi, F.-L.; Tan, Y.-C.; Liang, H.-F.; Sung, H.-W. In vivo biocompatibility and degradability of a novel injectable-chitosan-based implant. *Biomaterials* **2002**, *23*, 181–191.
 54. Orienti, I.; Cerchiara, T.; Luppi, B.; Bigucci, F.; Zuccari, G.; Zecchi, V. Influence of different chitosan salts on the release of sodium diclofenac in colon-specific delivery. *Int. J. Pharm.* **2002**, *238*, 51–59.
 55. MacLaughlin, F.C.; Mumper, R.J.; Wang, J.; Tagliaferri, J.M.; Gill, I.; Hinchcliffe, M.; Rolland, A.P. Chitosan and depolymerized chitosan oligomers as condensing carriers for in vivo plasmid delivery. *J. Control. Release* **1998**, *56* (1–3), 259–272.
 56. Cui, Z.; Mumper, R.J. Chitosan-based nanoparticles for topical genetic immunization. *J. Control. Release* **2001**, *75*, 409–419.
 57. Mao, H.-Q.; Roy, K.; Truong-Le, V.L.; August, J.T.; Leong, K.W. DNA-Chitosan Nanospheres: Derivatization and Storage Ability. *Proceedings Int. Symp. Control. Rel. Bioact. Mater.*, Stockholm, Sweden 1997; 671–672.
 58. Leong, K.W.; Mao, H.-Q.; Truong-Le, V.L.; Roy, K.; Walsh, S.M.; August, J.T. DNA-polycation nanospheres as non-viral gene delivery vehicles. *J. Control. Release* **1998**, *53*, 183–193.
 59. Roy, K.; Mao, H.-Q.; Huang, S.K.; Leong, K.W. Oral gene delivery with chitosan-pDNA nanoparticles generate immunologic protection in a murine model of peanut allergy. *Nat. Med.* **1999**, *5*, 387–391.
 60. Mao, H.-Q.; Roy, K.; Janes, K.A.; Truong-Le, V.L.; Lin, K.Y.; August, J.T.; Leong, K.W. Chitosan-DNA nanoparticles as gene carriers: Synthesis, characterization and transfection efficacy. *J. Control. Release* **2001**, *70*, 399–421.
 61. Kumar, M.; Behera, A.K.; Lockey, A.F.; Zhang, Z.; Bhullar, G.; De La Cruz, C.P.; Chen, L.-C.; Leong, K.W.; Huang, S.K.; Mohapatra, S.S. Intranasal gene transfer by chitosan-DNA Nanospheres protects BALB/c mice against acute respiratory syncytial virus infection. *Hum. Gene Ther.* **2002**, *13*, 1415–1425.
 62. Illum, L.; Jabbal-Gill, I.; Hinchcliffe, M.; Fisher, A.N.; Davis, S.S. Chitosan as a novel nasal delivery system for vaccines. *Adv. Drug. Deliv. Rev.* **2001**, *51*, 81–96.
 63. Koping-Hoggard, M.; Guan, I.T.; Edwards, K.; Nilsson, M.; Varum, K.M.; Artursson, P. Chitosan as a nonviral gene delivery system. Structure-property relationships and characteristics compared with polyethylenimine in vitro and after lung administration in vivo. *Gene Ther.* **2001**, *8*, 1108–1121.
 64. Corsi, K.; Chellat, F.; Yahia, L.H.; Fernandes, J.C. Mesenchymal stem cells, MG63 and HEK293 transfection using chitosan-DNA nanoparticles. *Biomaterials* **2003**, *24*, 1255–1264.
 65. Sato, T.; Ishii, T.; Okahata, Y. In vitro gene delivery mediated by chitosan. Effect of pH, serum, and molecular mass of chitosan on the transfection efficiency. *Biomaterials* **2001**, *22*, 2075–2080.
 66. Kim, Y.H.; Gihm, S.H.; Park, C.R. Structural characteristics of size-controlled self-aggregates of deoxycholic acid-modified chitosan and their application as a

- DNA delivery carrier. *Bioconjug. Chem.* **2001**, *12*, 932–938.
67. Thanou, M.; Florea, B.I.; Geldof, M.; Junginger, H.E.; Borchard, G. Quarternized chitosan oligomers as novel gene delivery vectors in epithelial cell lines. *Biomaterials* **2002**, *23*, 153–159.
68. Park, Y.K.; Park, H.K.; Shin, B.A.; Choi, E.S.; Park, Y.R.; Akaike, T.; Cho, C.S. Galactosylated chitosan-graft-dextran as hepatocyte-targeting DNA carrier. *J. Control. Release* **2000**, *69*, 97–108.
69. Okamoto, H.; Nishida, S.; Todo, H.; Sakakura, Y.; Iida, K.; Danjo, K. Pulmonary gene delivery by chitosan-pDNA complex powder prepared by a supercritical carbon dioxide process. *J. Pharm. Sci.* **2003**, *92* (2), 371–380.
70. Langer, R.; Vacanti, P. Tissue engineering. *Science* **1993**, *260*, 920–926.
71. Niklason, L.E.; Langer, R.S. Advances in tissue engineering of blood vessels and other tissues. *Transplant. Immunol.* **1997**, *5*, 303–306.
72. Kim, B.S.; Mooney, D.J. Development of biocompatible synthetic extracellular matrices for tissue engineering. *Trends Biotechnol.* **1998**, *16*, 224–230.
73. Hayashi, T. Biodegradable polymers for biomedical uses. *Prog. Polym. Sci.* **1994**, *9*, 663–702.
74. Han, D.K.; Park, K.D.; Hubbell, J.A.; Kim, Y.H. Surface characteristics and biocompatibility of lactide-based poly(ethylene glycol) scaffolds for tissue engineering. *J. Biomater. Sci., Polym. Ed.* **1998**, *9*, 667–6680.
75. Yannas, I.V. Applications of ECM analogs in surgery. *J. Cell. Biochem.* **1994**, *56*, 188–191.
76. Francis Suh, J.-K.; Matthew, H.W.T. Application of chitosan-based polysaccharide biomaterials in cartilage tissue engineering: A review. *Biomaterials* **2000**, *21*, 2589–2598.
77. Chandy, T.; Sharma, C.P. Chitosan—As a biomaterial. *Biomater. Artif. Cell Artif. Organs* **1990**, *18* (1), 1–24.
78. Eser Elcin, A.; Elcin, Y.M.; Pappas, G.D. Neural tissue engineering: Adrenal chromaffin cell attachment and viability on chitosan scaffolds. *Neurol. Res.* **1998**, *20* (7), 648–654.
79. Nishimura, K.; Nishimura, S.; Nishi, N.; Saiki, I.; Tokura, S.; Azuma, I. Immunological activity of chitin and its derivatives. *Vaccine* **1984**, *2* (1), 93–99.
80. Madhivaly, S.V.; Matthew, H.W.T. Porous chitosan scaffolds for tissue engineering. *Biomaterials* **1999**, *20*, 1133–1142.
81. Zhang, Y.; Zhang, M. Synthesis and characterization of macroporous chitosan/calcium phosphate composite scaffolds for tissue engineering. *J. Biomed. Mater. Res.* **2001**, *55*, 304–312.
82. Vande Vord, P.J.; Matthew, H.W.T.; DeSilva, S.P.; Mayton, L.; Wu, B.; Wooly, P.H. Evaluation of the biocompatibility of chitosan scaffold in mice. *J. Biomed. Mater. Res.* **2001**, *59*, 585–590.
83. Lee, J.-Y.; Nam, S.-H.; Im, S.-Y.; Park, Y.-J.; Lee, Y.-M.; Seol, Y.-J.; Chung, C.-P.; Lee, S.-J. Enhanced bone formation by controlled growth factor delivery from chitosan-based biomaterials. *J. Control. Release* **2002**, *78*, 187–197.
84. Shahidi, F.; Vidana Arachchi, J.K.; Jeon, Y.-J. Food applications of chitin and chitosans. *Trends Food Sci. Technol.* **1999**, *10*, 37–51. References therein.
85. Yalpani, M.; Johnson, F.; Robinson, L.E. Antimicrobial Activity of Some Chitosan Derivatives. In *Advances in Chitin and Chitosan*; Brine, C.J., Sandford, P.A., Zikakis, J.P., Eds.; Elsevier Applied Science: London, 1992; 543–555.
86. Sudharshan, N.R.; Hoover, D.G.; Knorr, D. Antibacterial action of chitosan. *Food Biotechnol.* **1992**, *6*, 257–272.
87. Jung, B.-O.; Kim, C.-H.; Choi, K.-S.; Lee, Y.-M.; Kim, J.-J. Preparation of amphiphilic chitosan and their antimicrobial activity. *J. Appl. Polym. Sci.* **1999**, *72*, 1713–1719.
88. Jumaa, M.; Furkert, F.H.; Müller, B.W. A new lipid emulsion formulation with high antimicrobial efficacy using chitosan. *Eur. J. Pharm. Biopharm.* **2002**, *53*, 115–123.



Chondroitin Sulfates

Nicola Volpi

Department of Biologia Animale, Biological Chemistry Section, University of Modena and Reggio Emilia, Modena, Modena, Italy

INTRODUCTION

The study of chondroitin sulfate (CS) and related proteoglycan (PG) structure and function has significantly influenced our understanding of a wide variety of biological processes and pharmacological applications. Historically, these cellular and extracellular macromolecules were generically designated as “ground substances” or “mucopolysaccharides” and, owing to their carbohydrate content and heterogeneity, were extremely difficult to study in relation to proteins. The advent of the field of glycobiology with tools and techniques for studying glycosaminoglycans (GAGs) has done much toward furthering appreciation of how carbohydrate moieties can affect cellular responses to a variety of stimuli.

Despite their wide distribution and common synthetic pathway to heparan sulfate, CS and dermatan sulfate (DS) have attracted less attention than heparan sulfate and heparin. Growing evidence suggests, however, that CS and DS chains have intriguing functions in central nervous system (CNS) development, wound repair, infection, growth factor signaling, morphogenesis, and cell division, in addition to their conventional structural roles. Furthermore, from a pharmacological point of view, CS produces a slow but gradual reduction of the clinical symptoms of osteoarthritis (OA) and these benefits last for a long period after the end of treatment. In fact, many animal studies and clinical trials have proved the efficacy of CS as a structure-modifying OA drug agent able to reverse, retard, or stabilize the pathology of OA, thereby providing symptomatic relief in the long-term treatment.

Chondroitin sulfate and DS are synthesized as GAG side chains of PGs and, as a consequence, many of their biological functions are related to these macromolecules. Furthermore, the recent molecular cloning of the various genes encoding the glycosyltransferases and sulfotransferases required for the synthesis of GAG chains has outlined major biosynthetic pathways that are common to vertebrates and invertebrates.

The aim of this entry is to focus on the structure and activity of CSs also as components of more complex PGs and to discuss recent advances in their pharmacological applications and their possible use as biomaterials. The

main arguments focused on in this entry will be the structure of CSs, their role in the structure and functions of PGs, their biological and pharmacological activities, and their applications as biomaterials.

STRUCTURE OF CS/DS

Glycosaminoglycans are complex natural heteropolysaccharides distributed among all organisms.^[1,2] These biological molecules are major structural components of PGs, the main macromolecular complex of the extracellular matrix, but also localized at the cellular level as intracellular components and distributed on the cellular membrane.^[3] Evidence derived from recent glycobiology studies in the fields of biochemistry, cell and developmental biology suggests that PGs are not only structural components, but they participate in and regulate many cellular events and physiological processes, such as cell proliferation and differentiation, cell–cell and cell–matrix interactions.^[4] The strong capacity of GAGs to be interactive at the molecular and cellular levels is due to their great structural diversity, such as the GAG type, size, composition and charge density, as well as the degree of substitution and domain arrangement.

Chondroitin sulfate chains are polymers constructed from 40 to more than 100 repeating disaccharide units comprising GlcA and GalNAc, whereas DS, known as CS type B, chains contain varying proportions of IdoA in place of GlcA. In addition, the two chains are often found as CS/DS copolymeric structures. Chondroitin sulfate and DS have the potential to display an enormous structural diversity, comparable to that of heparan sulfate, by embedding multiple overlapping sequences constructed with distinct disaccharide units modified by different patterns of sulfation. This diversity is the basis for their wide variety of functions.

Chondroitin sulfate and DS chains are formed by the repeating unit $[-\rightarrow 4\text{GlcA}\beta 1 \rightarrow 3\text{GalNAc}\beta 1 \rightarrow]$, which is then variously sulfated in different positions of the hexosamine unit and/or uronic acid. Some, but very few, of these positions remain unsulfated. The regular disaccharide sequence of CS type A, chondroitin 4-sulfate, is constituted by the repeating unit sulfated in position 4 of the GalNAc unit, while CS type C, chondroitin 6-sulfate, is mainly composed of a disaccharide unit

sulfated in position 6 (Fig. 1). Disaccharides with a different number and position of sulfate groups can be located, in different percentages, within the polysaccharide chains, such as the disulfated disaccharides in which two sulfate groups are O-linked in position 2 of GlcA and position 6 of GalNAc (disaccharide D) or in positions 4 and 6 of GalNAc (disaccharide E) (Fig. 1). In the case of DS, further enzymatic modifications complete the final structure, such as C-5 epimerization of GlcA to IdoA, and O-sulfation at C-2 of IdoA. As a consequence, polysaccharide chains of DS consist of a prevailing disaccharide unit [$\rightarrow 4$ IdoA $\beta 1 \rightarrow 3$ GalNAc $\beta 1 \rightarrow$] sulfated in position 4 of the hexosamine unit with a minor concentration of disulfated disaccharides, in particular sulfated in position 4 of GalNAc and position 2 of the IdoA unit (Fig. 1).

These heterogeneous structures are responsible for different and more specialized functions of these GAGs (see below). L-Iduronic acid imparts conformational flexibility to the DS chain, altering the shape and spatial

orientation of sulfate residues, endowing the chain with a higher negative charge content than the GlcA.^[5] Although the principles of the biosynthetic process have not yet been fully elucidated, it is well known that this process results in the generation of highly modified oligosaccharide domains within the polymer chain, which are separated by regions of relatively low-degree structural modifications. Thus, the DS chain has a hybrid copolymeric structure consisting of low modified (CS) and highly modified (DS) domains. The IdoA-containing units are often sulfated at C-4 of the GalNAc residue, while sulfation at C-6 is frequently associated with GlcA-containing disaccharides.^[6]

CHONDROITIN SULFATE/DERMATAN SULFATE PGs

Proteoglycans are macromolecules that consist of a protein backbone to which GAG chain(s) and

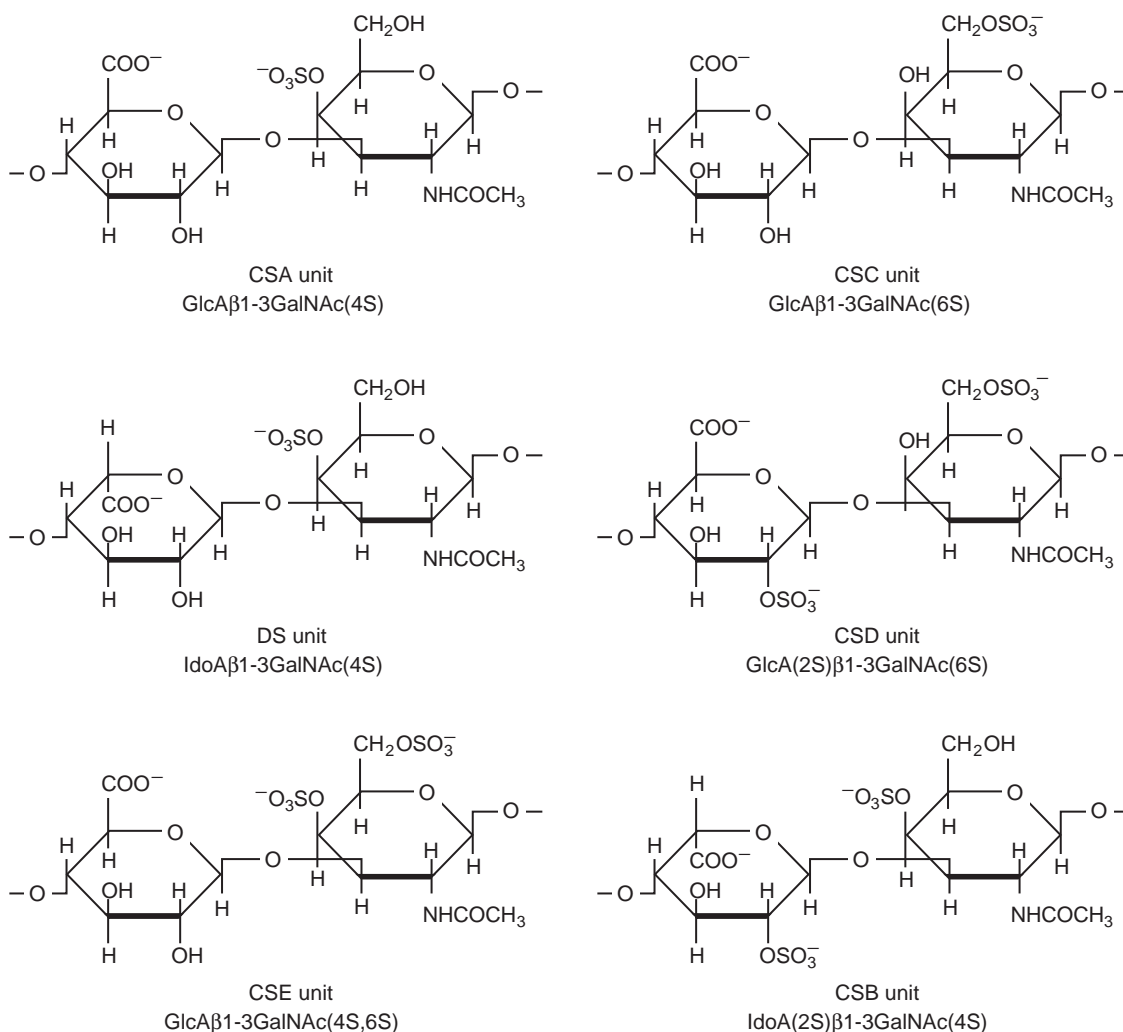


Fig. 1 Six typical disaccharide units found in CS and DS chains.



oligosaccharides are covalently attached. In view of the considerable complexity of PG structure and composition, we need to know the identities of the constituent GAGs but also the core-protein sequences to organize a logical classification scheme for PGs.

Chondroitin Sulfate PGs that Aggregate with Hyaluronic Acid

The large aggregating PGs in cartilage represent the most studied of all PGs (Table 1). They are characterized by their very large size. A hyaluronic acid-binding region (HABR) is located at the N-terminus and constitutes about one-quarter to one-third of the total core protein.^[7,8] This region can bind specifically to hyaluronic acid with the formation of macromolecular aggregates. Interaction of the PG with hyaluronic acid is stabilized by a link protein that binds to the HABR in a molar ratio of 1:1.

Some molecules contain a region where keratan sulfate chains are concentrated and that is adjacent to the HABR. To the remainder of the core protein are attached most of the CS chains.^[7] The very high charge density reached by aggregated cartilage PGs permits these molecules to imbibe reversibly large quantities of water. They can occupy a solution volume of 30–50 times their dry weight but in cartilage this swelling is limited by the network of collagen fibrils. The resulting pressure after swelling endows cartilage with its compressive rigidity and its ability to sustain repeated loading during articulation.

When cartilage develops, the structures of these PGs change. The molecule synthesized in the fetus initially lacks keratan sulfate. In man, this is added mainly after

birth and increases in amount with age. This is accompanied by a shortening of CS chains, with an increase in the proportion of 6-sulfate, a reduced molecular size, and an increase in the proportion of core protein. The significance of these GAG changes remains to be established.

Aging of human cartilage is accompanied by an accumulation of the HABR and partially fragmented link proteins. In rabbits there is a decrease in link-stabilized PG aggregates with age. These observations indicate that there is progressive degradation in cartilage with aging, which would favor the development of osteoarthritic changes where a net loss of PG is associated with macroscopic degenerative changes. In inflammatory arthritis, PGs can act as autoantigens and have been implicated in the immunopathology of the disease.^[7]

The aorta also contains large CS-PGs that aggregate with hyaluronic acid (Table 1). A link protein similar in size to the larger of the two major species found in cartilage is also present in the aorta. As in cartilage, aggregates of these PGs probably play a key role in maintaining shape and sustaining the repeated loading due to blood flow. A large aggregating CS-PG is also present in tendons and may perform an important role in dispersing load because it is more abundant in those parts of the tendon subjected to increased pressure.^[7]

Large Nonaggregating CS PGs

Skin contains a large nonaggregating CS-PG (Table 1). Its amino acid composition is similar to that of the large aggregating cartilage PG but it lacks cystine and methionine. This is consistent with the lack of aggregation, because both of these amino acids are present

Table 1 Chondroitin sulfate proteoglycans

Type	M_r	Glycosaminoglycans		Core protein		Tissue
		Chain number	M_r	M_r	Binds to HA	
Aggregating	2.5×10^6	100 CS	20,000	200,000	Yes	Cartilage
		30–60 KS	6,000			
	$1-2 \times 10^5$	12	43,000	200,000	Yes	Aorta
	$1-2 \times 10^6$	50	17,000	>200,000	Yes	Tendon
	nd	nd	nd	250,000	Yes	Melanoma
Large nonaggregating	150,000	1–2	30,000	70,000	Yes	Brain
	1×10^6	50	20,000	50,000	No	Skin
	1×10^6	20	40,000	200,000	No	Metaphysis
Small nonaggregating	$>2.5 \times 10^6$	nd	70,000	nd	No	Skeletal muscle
	70,000	1–2	35,000	35,000	No	Bone
	76,000–120,000	2–3	35,000	42,000	No	Cartilage

HA, hyaluronic acid; KS, keratan sulfate; nd, not determined.
(From Ref. [7].)

in the HABR of the cartilage molecule where cystine disulfide linkages are essential for the functional integrity of this region.

Another large PG is present in the soft tissue mesenchyme of the metaphysis lying between the trabeculae of forming bone (Table 1). It has a protein core similar in size to that of the aggregating cartilage PGs and cross-reacts with antibodies raised against these cartilage molecules.

Small Nonaggregating CS PGs

A major PG of bone is a molecule of small size and is similar to the small widely distributed DS-PG. It is present in osteoblasts and bone, but not in cartilage (Table 1). The major difference between the small PGs is probably whether DS or CS chains are synthesized.

Another similar small CS-PG is found in cartilage. Immunochemical analysis and peptide mapping have revealed that it is different from the small DS-PGs of cornea, sclera, and tendon, but similar to a small PG isolated from the aorta.

Dermatan Sulfate PGs

A number of PGs bearing DS have been described (Table 2). They vary in size and iduronic acid content. A small PG with an M_r of 70,000–100,000 and a

protein core probably of M_r 38,000 is widely distributed (Table 2).^[7] The core protein can constitute from 30% to 60% of the total weight. This molecule may have one or possibly two DS chains which, within a tissue, can vary in iduronate content from 45% to 85% of the total hexuronate. The length of this chain also shows tissue variations from M_r 15,000 to 45,000. In both skin, cartilage, and placenta, position 4 of the core protein is a serine residue and in each case there is probably either a DS chain or an O-linked oligosaccharide attached to it. Some PGs may have the same or very similar core proteins to which either CS or DS chains are attached, depending on whether glucuronic acid is epimerized to iduronic acid during synthesis. Several studies have indicated that this small DS-PG is bound to type I collagen. Because the PG can inhibit fibrillogenesis of type I and type II collagens, it may play a key role in the regulation of collagen fibril assembly and in the biomechanical strength of collagen fibrils.

Dermatan sulfate chains with a high iduronic acid content, e.g., those found in the small DS-PG, can self-associate to form multimeric PG complexes.^[7] This is a property of the arrangement of the glucuronic acid and iduronic acid components and may indicate that DS-PGs on adjacent collagen fibrils can reversibly associate through their GAGs, assuming that they are bound to collagen by their core proteins, as indicated. Such DS chains also show high-affinity binding to low-density lipoproteins and this may be of importance in

Table 2 Dermatan sulfate proteoglycans

M_r	Glycosaminoglycans			Core protein		
	Chain number	M_r	Iduronic acid (%)	M_r	Bind to HA	Cell/tissue
70–100 × 10 ³	1–2	15–24 × 10 ³	35–85	40,000	No	Skin Sclera Cervix Cartilage Tendon Placenta
100–150 × 10 ³	1	55,000	10	40,000	No	Cornea
1 × 10 ⁶	nd	nd	25	400,000	Yes	Fibroblast
160–410 × 10 ³	nd	nd	20	100,000	Yes	Sclera
2.5 × 10 ⁶	50	40,000	8	200,000	Yes	Glial, glioma cells
70–100 × 10 ³	1–2	15–24 × 10 ³	23–36	40,000	No	Cartilage (DS-PG I)
400–800 × 10 ³	nd	40,000	10	30,000	No	York sarc tumor
500,000	8	52,000	7	43,000	No	Embryonic cartilage
420,000	1–2	40,000	12	1 × 100,000	No	Embryonic cartilage
420,000	1–2	40,000	12	2 × 120,000	No	Embryonic cartilage
2–3 × 10 ⁵	20	56,000	3	500,000	No	Foallicular fluid

HA, hyaluronic acid; nd, not determined.
(From Ref. [7].)



atherosclerosis if these molecules are synthesized in excess in atheromas. A small glucuronic acid-rich DS-PG is also present in the cornea (Table 2). The core protein is immunologically related to the small DS-PG of cartilage and sclera and has the same size. The molecular size of this corneal PG enables it to fit between the collagen fibrils, permitting corneal transparency.

Fibroblasts secrete the small DS-rich PG into culture medium but retain a larger molecule that is cell associated and can bind to hyaluronic acid (Table 2). A large glucuronic acid-rich DS-PG is also found in the sclera (Table 2). Like the aggregating PGs of cartilage, aorta, and tendon, this molecule may play a key role in maintaining shape (of the eye) and sustaining intraocular pressure.

The other DS-PG of cartilage (DS-PGI) has an overall size and core protein similar to the smaller molecule (Table 2). Although it has a similar iduronic acid content to the smaller molecule, only the larger species can self-associate, probably through its DS chains.

BIOLOGICAL PROPERTIES OF CS

Chondroitin sulfate exhibits a wide variety of biological functions reviewed in Ref. [9,10], mainly owing to the presence of rare oversulfated structural building units forming domain structures interacting specifically with other molecules, inducing the modulation of several biological functions, such as the regulation of neuronal patterning in the retina, interactions with fibronectin, neurite outgrowth promoting activity, modulation of the adhesive function of $\alpha_4\beta_1$ integrin, monocyte and B-cell activation, plasminogen activation, binding of *Plasmodium falciparum*-infected erythrocytes, and upregulation of the antigen-specific Th1 immune response of murine splenocytes sensitized with ovalbumin in vitro and to suppress the antigen-specific IgE responses.

Highly sulfated PG is correlated with axon boundaries in the developing CNS, which suggests that these molecules affect neural pattern formation. In the developing mammalian retina, gradual regression of CS may help control the onset of ganglion cell differentiation and the initial direction of their axons. Changes induced by the removal of CS from intact retinas in culture confirm the function of CS in retinal histogenesis.

Interactions between fibronectin and GAGs are essential for extracellular matrix morphology and cell adhesion. Recombinant fibronectins (deminectins) containing the carboxyl-terminal cell, heparin, and fibrin domains bind specifically to CS in affinity chromatography assays.

A mouse brain PG-CS epitope was shown to possess neurite outgrowth promoting properties. Several

hexasaccharides isolated from commercial shark cartilage CSD, which contains a higher proportion of characteristic D units (GlcA2S β 1 \rightarrow 3GalNAc6S), has a tetrasaccharide sequence composed of an A disaccharide unit (GlcA β 1 \rightarrow 3GalNAc4S) and a D disaccharide unit. The CSD promoted neurite outgrowth of embryonic day 18 hippocampal neurons. Eight octasaccharide fractions were isolated from CS type D after partial digestion with bacterial chondroitin ABC lyase to investigate the frequency and the arrangement of the tetrasaccharide unit in the polymer sequence. Structural analysis demonstrated that the isolated octasaccharides shared the common core structure Δ hexA (Glc/Ido) α 1 \rightarrow 3GalNAc β 1 \rightarrow 4(GlcA β 1 \rightarrow 3GalNAc)₃, with four, five, and six sulfate esters at various hydroxyl groups in different combinations.

Melanoma CS-PG plays a role in modulating the adhesive function of $\alpha_4\beta_1$ integrin. A peptide from within α_4 integrin termed SGI binding to cell surface melanoma CS-PG was identified. Soluble SGI inhibits $\alpha_4\beta_1$ integrin-mediated human melanoma cell adhesion to CS. These results demonstrate that $\alpha_4\beta_1$ integrin interacts directly with CS through the SGI site, and that this site can affect the ligand binding properties of the integrin.

At inflammatory sites, PG are both secreted by activated mononuclear leukocytes and released as a consequence of extracellular matrix degradation. Chondroitin 4-sulfate PGs constitute the predominant ones produced by activated human monocytes/macrophages. Two chondroitin 4-sulfate forms, CSA and CSB, can activate distinct peripheral blood mononuclear cell types. Whereas CSA activates monocytes (to secrete monokines), CSB activates B-cells (to proliferate). In contrast, the chondroitin 6-sulfate CSC and heparin do not exert these functional effects. These findings point to GAGs, and specifically chondroitin 4-sulfates, as a novel class of immunological mediators at inflammatory sites. Furthermore, the data link CD44 to B-cell activation, paralleling the established roles of CD44 in T-cell and monocyte activation.

CS type E markedly enhanced plasminogen activation by tissue plasminogen activators (t-PAs) and urinary plasminogen activator (u-PA) in vitro. Though the potentiation activity of CSE gradually decreased when it was depolymerized by chondroitinase ABC, the specific disaccharide from CSE still showed significant activity. These results suggested that the minimal structural requirement in CSE to enhance plasminogen activation by plasminogen activators is GlcA β 1-3GalNAc(4S,6S) and that additional branching sugars abolish the activity.

Plasmodium falciparum infection in pregnant women results in chondroitin 4-sulfate-mediated adherence of the parasite-infected red blood cells (IRBCs) in the placenta, adversely affecting the health

of the fetus and the mother. The unusually low sulfated CS-PGs in the intervillous spaces of the placenta are the receptors for IRBC adhesion, which involves a chondroitin 4-sulfate motif consisting of six disaccharide moieties with approximately 30% 4-sulfated residues. The placental CS-PGs are a mixture of two major populations, which are similar by all criteria except for their differing sulfate contents. The majority of the sulfate groups in the CS-PGs are clustered in CS chain domains consisting of 6–14 repeating disaccharide units. While the sulfate-rich regions of the CS chains contain 20–28% 4-sulfated disaccharides, the other regions have little or no sulfate. The oligosaccharides corresponding to the sulfate-rich domains of the CS chains efficiently inhibited IRBC adhesion. Finally, the unique distribution of sulfate groups in the CS chains of placental CS-PGs and the sulfate-clustered domains have the necessary structural elements for the efficient adhesion of IRBCs.

Chondroitin 4-sulfate was shown to upregulate the antigen-specific Th1 immune response of murine splenocytes sensitized with ovalbumin in vitro, and CS was able to suppress the antigen-specific IgE responses. In a further recent study, a specific sulfation pattern of the CS was demonstrated to be necessary for the Th1-promoted activity, suggesting that the GlcA(β -3)GalNAc(4,6-*O*-disulfo) sequence in CSE is important for Th1-promoted activity.^[11] These findings describe a new mechanism for the anti-inflammatory and chondroprotective properties of CS.

BIOLOGICAL ACTIVITIES OF DS

Dermatan sulfate proteoglycans bind a diverse range of molecules including matrix molecules, growth factors, protease inhibitors, cytokines, chemokines, and pathogen virulence factors.^[12] The precise binding site has neither been identified for many of these proteins, nor has the binding affinity for different structural variants of DS been thoroughly investigated.

Table 3 lists a number of DS-binding proteins. Although many of these interactions were first identified using other GAGs, it has been appreciated that many proteins can interact with more than one type of GAG chain albeit with differing affinities. The characterization of GAG/PG-protein interactions is complicated by the fact that many interactions are between the PG protein core and the binding protein rather than the GAG. It is conceivable that certain GAG/PG-protein binding interactions involve both the GAG side chain(s) and the PG protein core.

One particularly well-investigated DS binding interaction, reviewed by Trowbridge and Gallo, occurs with heparin cofactor II (HCII).^[12] This serpin homolog of antithrombin III acts by inhibiting the procoagulative effect of thrombin. This effect is enhanced 1000-fold in the presence of DS or heparin. It is postulated that the GAG forms a stable ternary complex between the serpin and the protease. Dermatan sulfate fragments reported to bind HCII include octasaccharides and hexasaccharides with the



Table 3 Binding interactions of dermatan sulfate and dermatan sulfate proteoglycans

Binding protein	Binding sequence	Physiologic effect
Heparin cofactor II	IdoA-2S-GalNAc-4S Hexasaccharide	Enzymatic inactivation of thrombin
Thrombin	nd	Anticoagulation
Activated protein C	nd	Anticoagulation
Protein C inhibitor	nd	Stimulates serpin activity
Platelet factor 4	nd	nd
Collagen	Binds protein core	Extracellular matrix stability
Fibronectin	Binds protein core	Extracellular matrix stability
Tenascin-X	Binds GAGs core	Collagen matrix stability
<i>Borrelia burgdorferi</i> adhesions	nd	Increased infectivity
α -Defensin	nd	Increased infectivity
RANTES	nd	Modulation of inflammatory response
Interferon gamma (INF- γ)	nd	Receptor for INF- γ
Transforming growth factor-beta	Binds protein core	Growth regulation
Fibroblast growth factors 1 and 2	nd	Cellular proliferation
Hepatocyte growth factor/scatter factor	IdoA-GalNAc-4S Octasaccharide	Cellular proliferation, organogenesis
Low-density lipoprotein	nd	Atherosclerotic plaque stabilization

nd, not determined.
(From Ref. [12].)

highest-affinity fraction bearing a sulfate on IdoA C2 and on the C4 of GalNAc. Interestingly, this species represents only about 5% of porcine skin DS. Despite its high binding affinity, the purified fraction was estimated to be 20 times less active than unfractionated DS in a thrombin inhibition assay. With a different source of DS (bovine mucosa), the total charge of the DS disaccharide subunits appears to matter more than the IdoA content per se. Two disaccharide species (IdoA 2S-GalNAc-4S and GlcA-GalNAc-4,6diS) were primarily identified in DS that activated HCII.

Heparin has been shown to have a fivefold greater affinity for HCII than DS, but the complex formed between heparin, fibrin, and thrombin is different from that formed by DS. It was postulated that heparin binds both thrombin and fibrin, thereby forming a ternary complex and causing a change in the HCII active site of thrombin such that it is no longer sensitive to inhibition. On the contrary, DS binds directly to thrombin only, not to fibrin, and the bound thrombin retains sensitivity toward the serpin. This suggests that DS might be a useful clinical tool for specifically treating or preventing thrombosis.

Dermatan sulfate may also influence coagulation by enhancing the effects of activated protein C (APC), an endogenous inhibitor of the clotting cascade. Other CS samples (CSA and CSC) also enhance APC activity. At a physiologically relevant concentration, DS enhanced APC activity more than the other GAGs tested. Moreover, when fractions of DS separated on the basis of charge density were tested, the fraction containing the highest charge density enhanced APC activity more than unfractionated DS. These findings further suggest that DS oligosaccharide size and sulfation determine binding affinity and lead to different physiological consequences.

PHARMACOLOGICAL EFFECTS OF CS

Chondroitin sulfate has anti-inflammatory activity and affects cartilage metabolism (reviewed in Ref. [9,10]). Oral administration of CS was shown to significantly decrease the granuloma formation secondary to cotton, pellet, or sponge implants, the inflammatory response in adjuvant arthritis, and lysosomal enzyme release in carrageenan pleurisy. Furthermore, CS and its fractions inhibit the directional chemotaxis induced by zymosan-activated serum, are able to decrease the phagocytosis and the release of lysozyme induced by zymosan and to protect the plasma membrane from oxygen reactive species. When the effects of CS on chymopapain-induced articular cartilage injury in rabbits were investigated, it was found that rabbits that were given intramuscular or oral doses of CS before

injection of chymopapain demonstrated less PG loss from the articular cartilage than the controls.

The antioxidant activity of chondroitin 4-sulfate was tested on a rat model of collagen-induced arthritis (CIA). Arthritis was induced in Lewis rats by multiple intradermal injections of emulsion containing bovine type II collagen in complete Freund's adjuvant at the base of the tail and into three to five other sites on the back. Treatment with CS, starting at the onset of arthritis for 10 days, limited the erosive action of the disease in the articular joints of knee and paw, reduced lipid peroxidation, restored the endogenous antioxidants, reduced glutathione (GSH) and superoxide dismutase, decreased plasma TNF-alpha levels, and limited synovial neutrophil infiltration. These data confirm that erosive destruction of the joint cartilage in CIA is due at least in part to free radicals released by activated neutrophils and produced by other biochemical pathways. Furthermore, in another study, CS treatment was found to reduce the damage in carbon tetrachloride-induced acute rat liver injury.

Chondroitin sulfate efficacy and tolerance have been evaluated and demonstrated in many clinical studies (Table 4) reviewed in Ref. [9,10]. Oliviero et al. conducted a clinical trial, lasting 6 mo, performed on 200 patients. The results showed a considerable improvement both in pain and in mobility. No relevant side effects were found; only 3% of patients, with oral administration, noticed slight nausea and found it necessary to interrupt treatment.

Mazieres et al. treated 120 patients with CS in a placebo-controlled study. Significant improvement of joint pain and physician global assessments was evident in the CS-treated group. Morreale and coworkers evaluated 146 patients with knee OA. Seventy-four patients received 400 mg CS three times a day for 3 mo, followed by 3 mo without treatment. These patients were compared with 72 patients who were treated with 50 mg of diclofenac three times a day for 1 mo, followed by 5 mo of treatment. Those treated with diclofenac had a decrease in joint pain at 10 days, which disappeared shortly after discontinuation of treatment; those treated with CS had a significant response at 30 days, which lasted for 3 mo after discontinuation of the drug. Thus, the effects of CS treatment persisted longer after discontinuation than those of traditional therapy.

In a study by Bucsi and Poor, 40 patients treated with 800 mg of CS were compared with 40 who received placebo. The CS group showed significant improvement in the Lequesne index, visual analog pain scale, 20 m walk time, and patient and physician efficacy ratings. All these results strongly suggest that CS acts as a symptomatic slow-acting drug in knee OA.

Bourgeois et al. evaluated 40 patients with knee OA who were treated with CS in the form of a 1200 mg

Table 4 Summary of results in chondroitin sulfate trials

Author(s)	Year	No. of patients	Follow-up period	Complications	Results
Rovetta	1991	40	25 wk	Well tolerated	Higher therapeutic effect on all symptoms of OA
Oliviero et al.	1991	200	6 mo	3% Mild adverse effects	Considerable improvement in both pain and mobility
Mazieres et al.	1992	120	3 mo	Well tolerated	Slowly effective against symptoms of OA and reduction of the need for NSAIDs
Morreale et al.	1996	146	3 mo	Minor	Benefits of CS appeared later but lasted for up to 3 mo after end of treatment
Fleish et al.	1997	56	1 yr	Well tolerated	Improvement in mobility, joint effusion, and swelling
Bucsi and Poor	1998	80	6 mo	Minor	43% reduction in joint pain
Bourgeois et al.	1998	127	3 mo	No adverse events	Significant reduction in joint pain
Uebelhart et al.	1998	42	1 yr	None	Decreased joint pain and improved mobility
Verbruggen et al.	1998	119	3 yr	Not documented	Radiographic demonstration of decrease in number of patients with "new" erosive finger-joint OA
Uebelhart et al.	2004	120	1 yr	Minor adverse events	Decreased pain and improved knee function. Reduction of the progression of medial femoro-tibial joint space
Rovetta et al.	2004	24	2 yr	Not documented	Efficacy of oral CS in improving some aspects of erosive OA of the hands

(From Ref. [10].)

dose oral gel. The Lequesne index and pain score (visual analog scale) decreased significantly in both CS groups, compared with the placebo group, and patient and physician global assessments significantly favored the CS-treated group.

A computerized technique was used to measure medial tibiofemoral joint space in patients treated with 800 mg CS per day or with placebo. After 1 yr, tibiofemoral joint space width had decreased significantly in placebo-treated patients but had not changed from the baseline value in the CS treatment group. The aim of this study was to assess the clinical, radiological, and biological efficacy, and tolerability of the chondroitin 4- and 6-sulfate, in patients suffering from knee OA.

Verbruggen and colleagues evaluated hand radiographs of 119 patients with OA (Table 4). Thirty-four patients who received CS 400 mg three times a day were compared with 85 patients who received placebo. Radiographs were performed annually over a 3 yr period. A total of 119 patients were included in a randomized, double-blind, placebo-controlled trial to assess the properties in OA of CS. In the CS group, a significant decrease in the number of patients with new erosive OA finger joints was observed.

In a very recent randomized, double-blind, multicenter study, the efficacy and tolerability of oral CS in knee OA patients in a 3 mo duration, twice yearly,

intermittent treatment was investigated (Table 4). Radiological progression at month 12 showed significantly decreased joint space width in the placebo group with no change in the CS group. Tolerability was good with only minor adverse events identically observed in both groups. This study provides evidence that oral CS decreased pain and improved knee function.

PHARMACOLOGICAL USE OF DS

As previously reported, although the anticoagulant activity of DS is significantly lower than that of heparin, its venous antithrombotic activity appears to be significantly higher. Thus, the hemorrhagic properties of DS are greatly reduced when compared to heparin. This has made DS an interesting target in developing new therapeutic agents for the prevention of thrombosis.^[13] Dermatan sulfate has two major pharmacokinetic problems, a relatively short half-life and low bioavailability, when compared to heparin delivered by subcutaneous or intramuscular routes. However, DS does not bind to the endothelium and thus gives simpler pharmacokinetics than those observed for heparin.

The use of DS in clinical trials has been extremely limited. A pharmaceutical preparation of DS has been used in healthy volunteers and patients under



treatment for renal failure by dialysis, venous thromboembolism, disseminated intravascular coagulation syndrome, and arterial atherothrombotic diseases. In all cases (more than 300 patients) no serious bleeding complications have occurred. Furthermore, DS has been used to prevent deep vein thrombosis in hip fracture patients. In another study, the efficacy of moderate intravenous doses of DS was compared to heparin for the treatment of disseminated intravascular coagulation in acute leukemia patients. Dermatan sulfate was as effective as heparin in preventing deep vein thrombosis. Moreover, new bleeding complications were less frequent with DS than heparin and no patient experienced thromboembolic complications.

CHONDROITIN SULFATE APPLICATION IN BIOMATERIALS

Several materials are used for regeneration therapies in orthopedics and dentistry, such as heterologous graft of frozen or lyophilized bone, both demineralized or not, use of biomaterials, reabsorbable or not, or combinations of the above-mentioned materials with organic macromolecules composing the extracellular matrix such as proteins and GAGs. Apatites and bioglass are excellent biomaterials with good biocompatibility and associated with a lack of toxicity or proinflammatory activity.

As previously reported, mineralized connective tissues predominantly contain small CS-PG species. For these reasons, GAGs and in particular CS, have received major attention as molecules implicated in the absorption process onto hydroxyapatite. This biomaterial has been used for the construction of artificial bone and as bone-substitute biomaterial used to fill bone defects around immediate implants in humans.^[14] Furthermore, material composed of gelatin, CS, and hyaluronic acid was utilized for cartilage tissue engineering, and GAG hydrogel films, formed of hyaluronic acid and CS, were applied as biointeractive dressings for wound healing.^[15]

CONCLUSIONS

The rapidly accumulating evidence obtained in these last decades suggests that CS/DS have crucial biological functions in growth factor signaling, axonal growth and regeneration, neuritogenesis during CNS development, morphogenesis, and cell division. In addition, the involvement of cell-surface CS/DS chains in infection has been recently reported. Although the evidence is fragmentary and immature, future analysis will help reveal the molecular mechanisms underlying these functions.

The study of CS will certainly extend well into this new century with so many questions still left unanswered. Some focal points in the near future include: 1) improved preparation and synthesis of CS/DS; 2) new DS-based anticoagulants with improved properties; 3) new therapeutic applications for CS/DS; 4) new biomaterials; and 5) development of an improved understanding of physiology and pathophysiology through glycomics.

The preparation of CS from mammalian tissues creates concern particularly after the recent appearance of bovine spongiform encephalopathy in Europe. As a consequence, there are growing concerns about CS prepared by alternative routes, defined, recombinant mammalian cell lines capable of being cultured in large-scale fermentations, or non-mammalian tissues, offer an exciting alternative to the present preparations.

New anticoagulants based on DS structure might provide enhanced specificity, targeting one or selected groups of coagulation proteases and avoiding undesired interactions with other proteins, thus decreasing the side effects associated with the use of heparins. These new potential applications of DS/CS may be strictly associated with molecules possessing a peculiar and unusual structure, such as those isolated from invertebrates.

Finally, one of the most important future directions of CS research is driven by the recent sequencing of the human genome and the field of genomics. While much attention is currently focused on the proteome encoded by the genome and the rapidly developing field of proteomics, the glycome has attracted little attention. Glycomics is the study of the structure and function of the glycome, the most important and complex of the posttranslation modifications that proteins undergo. An improved understanding of the glycome should be beneficial in better understanding genetic diseases, offering new therapeutic approaches to treat these very serious pathologies. Moreover, improved knowledge of glycomics, in vertebrate and invertebrate species, should lead to a better understanding of physiology and pathophysiology, offering new approaches to drug development.

ABBREVIATIONS

APC	Activated protein C
CNS	Central nervous system
CS	Chondroitin sulfate
CIA	Collagen-induced arthritis
DS	Dermatan sulfate
GalNAc	N-acetyl-D-galactosamine
GAGs	Glycosaminoglycans
GlcA	D-glucuronic acid
HABR	Hyaluronic acid-binding region

HCII	Heparin cofactor II
IdoA	L-iduronic acid
IRBCs	Infected red blood cells
OA	Osteoarthritis
PG	Proteoglycan
TNF	Tumor necrosis factor

ARTICLE OF FURTHER INTEREST

Tissue Engineering of Cartilage

REFERENCES

- Volpi, N., Ed.; Analytical techniques to evaluate the structure and function of natural polysaccharides, glycosaminoglycans. *Res. Signpost, India* **2002**.
- Sugahara, K.; Mikami, T.; Uyama, T.; Mizuguchi, S.; Nomura, K.; Kitagawa, H. Recent advances in the structural biology of chondroitin sulfate and dermatan sulfate. *Curr. Opin. Struct. Biol.* **2003**, *13*, 612–620.
- Hardingham, T.E.; Fosang, A.J. Proteoglycans: many forms and many functions. *FASEB J.* **1992**, *6*, 861–870.
- Jackson, R.L.; Busch, S.J.; Cardin, A.D. Glycosaminoglycans: molecular properties, protein interactions, and role in physiological processes. *Physiol. Rev.* **1991**, *71*, 481–539.
- Casu, B.; Choay, J.; Ferro, D.R.; Gatti, G.; Jacquinet, J.C.; Petitou, M.; Provasoli, A.; Ragazzi, M.; Sinay, P.; Torn, G. Controversial glycosaminoglycan conformations. *Nature* **1986**, *322*, 215–216.
- Karamanos, N.K.; Syrokou, A.; Vanky, P.; Nurminen, M.; Hjerpe, A. Determination of 24 variously sulfated galactosaminoglycan- and hyaluronan-derived disaccharides by high-performance liquid chromatography. *Anal. Biochem.* **1994**, *221*, 189–199.
- Poole, A.R. Proteoglycans in health and disease: structures and functions. *Biochem. J.* **1986**, *236*, 1–14.
- Silbert, J.E. Metabolism and structure-function relationship of connective tissue glycosaminoglycans and proteoglycans. *Curr. Organic Chem.* **2004**, *8*, 395–411.
- Volpi, N. The pathobiology of osteoarthritis and the rationale for using the chondroitin sulfate for its treatment. *Curr. Drug. Targets-Immune Endocr. Metab. Disord.* **2004**, *4*, 119–127.
- Volpi, N. Chondroitin sulfate for the treatment of osteoarthritis. *Curr. Med. Chem. Anti-inflammatory Anti-allergy Agents.* **2005**, *4*, 221–234.
- Akiyama, H.; Sakai, S.; Linhardt, R.J.; Goda, Y.; Toida, T.; Maitani, T. Chondroitin sulfate structure affects its immunological activities on murine splenocytes sensitized with ovalbumin. *Biochem. J.* **2004**, *382*, 269–278.
- Trowbridge, J.M.; Gallo, R.L. Dermatan sulfate: new functions from an old glycosaminoglycan. *Glycobiology* **2002**, *12*, 117R–125R.
- Linhardt, R.J.; Hileman, R.E. Dermatan sulfate as a potential therapeutic agent. *Gen. Pharmacol.* **1995**, *26*, 443–451.
- Rebaudi, A.; Silvestrini, P.; Trisi, P. Use of a resorbable hydroxyapatite-collagen chondroitin sulfate material on immediate postextraction sites: a clinical and histologic study. *Int. J. Periodontics Restorative Dent.* **2003**, *23*, 371–379.
- Kirker, K.R.; Luo, Y.; Nielson, J.H.; Shelby, I.; Prestwich, G.D. Glycosaminoglycan hydrogel films as bio-interactive dressings for wound healing. *Biomaterials* **2002**, *23*, 3661–3671.



Collagen

Matthew J. Beckman

*Department of Biochemistry Orthopaedic Research Laboratory,
Virginia Commonwealth University, Richmond, Virginia, U.S.A.*

Kelly J. Shields

*Department of Biomedical Engineering Orthopaedic Research Laboratory, Virginia
Commonwealth University, Richmond, Virginia, U.S.A.*

Robert F. Diegelmann

*Department of Biochemistry, Anatomy & Emergency Medicine,
Virginia Commonwealth University, Richmond, Virginia, U.S.A.*

INTRODUCTION

Collagen is the most abundant protein in the human body, accounting for 30% of the total protein. In normal tissues, collagen provides tensile strength and structural integrity. Collagen has a structure of splendid hierarchical order. The structural complexity involves numerous types of collagen associated with tissue-specific expression, organization into heteropolymeric or homopolymeric fibrils of uniform size, unique posttranslational processing and interaction with proteoglycans, adhesion glycoproteins, cells, and other collagens. When tissues are disrupted following injury, collagen is needed to repair the defect and restore structure and function. Too much collagen deposited in a wound site results in loss of normal anatomical structure, function, and fibrosis. Conversely, if insufficient amounts of collagen are deposited, the wound is weak and may dehisce. Therefore, to fully understand wound healing, it is essential to understand the basic biochemistry of collagen.

Collagen forms the extracellular matrix that defines the ground substance for all connective tissues and is a major functional component of skin, bone, cartilage, and other tissues to be discussed. There are over 1000 commercial products on the market today that contain collagen and collagen enhancers. These products are represented by body and hand lotions, nail treatments, firming gels, wrinkle injections, eye pads, and even anti-cancer treatments, to name but a few. In recent years, new high tech wound dressing materials and skin substitutes have become available for the treatment of partial and full thickness dermal injuries and chronic dermal ulcers. Future applications of collagen products hold promise in the fields of drug delivery and tissue engineering.

COLLAGEN STRUCTURE AND FUNCTION

Collagen provides structural integrity for connective tissues such as dermis, bone, cartilage, tendon, ligament,

and internal organs.^[1,2] Due to the wide distribution throughout various tissues, collagen represents one of the most abundant naturally occurring proteins on earth.^[3] The collagen super-family includes over 30 different gene products that assemble into at least 20 genetically distinct collagen types which may be homotrimeric or heterotrimeric in nature.^[4,5] The ubiquitous nature of collagen is such that a single cell may co-express several collagen types. For example, skin fibroblasts are able to synthesize six highly homologous but genetically distinct procollagen chains necessary for the assembly of fibrillar collagen. The ability of individual procollagen chains to discriminate between each other is a central feature of the assembly process by ensuring that each collagen molecule has the correct complement of pro chains. Collagens I–V are the five major types, and those which are the main focus of this review (Table 1). The predominant form is type I collagen, found in bone, and is by far the major single protein. Type I collagen is also found in other tissues such as skin, dentin, cornea, blood vessels, fibrocartilage, and tendon. The fibrillar form of collagen represents over 90% of our total collagen and is composed of three very long protein chains. In type I collagen, each protein chain is referred to as an alpha chain (α -chain). Two of the α -chains are identical and are called α 1-chains whereas the third chain is slightly different and is labeled the α 2-chain. The three chains are twisted with each other to form a triple helical structure called a collagen monomer or a gamma structure (γ -structure) (Fig. 1a). The equivalent configuration imparts tremendous strength to the protein. The overall structure of the collagen molecule can be compared to the reinforcement rods (called rebar) used in concrete construction. The scaled molecular dimensions of a collagen molecule would be one inch in diameter and 17 feet in length. Collagen is indeed nature's rebar due to its incredible strength and provision of structural integrity for connective tissues and organ structures.

Table 1 Collagens

Type ^a	Composition	Tissue distribution
I	($\alpha 1(I)$) ₂ $\alpha 2(I)$, trimer ($\alpha 1(I)$) ₃	Dermis, bone cornea, tendon, fibrocartilage, vessels, intestine, dentin, skin, uterus
II	($\alpha 1(II)$) ₃	Cartilage, notochord, vitreous humor, nucleus pulposus
III	($\alpha 1(III)$) ₃	Dermis, intestine, gingiva, heart valve, uterine wall, vessels
IV	($\alpha 1(IV)$) ₂ $\alpha 2(IV)$	Basement membranes
V	$\alpha 1(V)\alpha 2(V)(\alpha 3(V))$ or ($\alpha 1(V)$) ₂ $\alpha 2(V)$ or ($\alpha 1(V)$) ₃	Cornea, placental membranes, bone, vessels, cartilage, gingiva
XI	1 $\alpha 2\alpha 3\alpha 1$ or $\alpha 1(XI)\alpha 2(XI)\alpha 3(XI)$	Cartilage, intervertebral disc, vitreous humor

^aI-V are the most abundant; I, II, III, V, and XI are capable of fibril formation.

All of the collagens share the triple-helical molecular structure. However, the various other types of collagens have slightly different amino acid compositions and provide other specific functions in our bodies. Type II collagen is the form found exclusively in cartilaginous tissues. It is usually associated with water, proteoglycans, glycoproteins, and noncollagenous proteins. The collagen provides for the tensile integrity of the tissue whereas the water and proteoglycans are responsible for the swelling. Type II collagen acts as a shock absorber in our joints and vertebrae. Type III collagen is found in our skin as well as in ligaments, blood vessels, and internal organs. In the adult, the skin contains about 80% type I and 20% type III collagen. In newborns, the type III content is greater than that found in the adult and may contribute to the supple nature of newborn skin. Type III also contributes to the flexibility of blood vessels and is expressed during the initial period of wound healing.^[6]

Type IV collagen is found in basement membranes and basal lamina structures and functions as a filtration system. The complex interactions between type IV collagen and the noncollagenous components of the basement membrane form a meshwork that filters cells, molecules, and light. For example, the basement membrane in the lens capsule of the eye plays a role in light filtration. The glomerulus basement membrane of the kidney is responsible for filtration of the blood to remove waste products. The basement membrane form in the walls of blood vessels controls the movement of oxygen and nutrients out of circulation and into the tissues. Likewise, the basal lamina in the skin delineates the dermis from the epidermis and controls the movement of materials in and out of the dermis.

Type V collagen is found in essentially all tissues and is associated with types I and III. In addition it is often found around the perimeter of many cells and functions as a cytoskeleton. It is of interest to note that there appears to be a particular abundance of type V collagen in the intestine compared to other tissues.^[7] The remaining types of collagen play minor but

important roles participating as extracellular signaling molecules and eliciting intracellular responses.

Primary Structure

The collagens most involved in forming fibrillar structures include types I, II, III, V, and XI.^[8] The structures are assembled in precise combinations utilizing 10 varying polypeptide chains that include, $\alpha 1(I)$, $\alpha 2(I)$, $\alpha 1(II)$, $\alpha 1(III)$, $\alpha 1(V)$, $\alpha 2(V)$, $\alpha 3(V)$, $\alpha 1(XI)$, $\alpha 2(XI)$, and $\alpha 3(XI)$ (Table 1). The chains derive from a common ancestral gene containing a nucleotide segment that encodes a peptide representing strings of five or six amino acid triplets comprised of Glycine-X-Y, where the residues for X and Y are frequently the amino acids proline (Pro) and lysine (Lys). At the nucleotide level this corresponds to a 45- or 54-nucleotide segment that is repeated several times uninterrupted and provides a common feature within the highly conserved collagen family (Fig. 1b). At the amino acid level, there is extensive posttranslational modification that takes place, which includes hydroxylation of Lys and Pro residues into hydroxylysine and hydroxyproline. More variable is the glycosylation of hydroxylysine residues into galactosyl- or glucosyl-galactosyl-hydroxylysine, and oxidation of some Lys and hydroxylysine residues into an aldehyde group. In diabetic patients sometimes the collagen in tendons and basement membranes becomes glycosylated by nonenzymatic mechanisms and this abnormal collagen is responsible for the many pathologic characteristics associated with the disease such as blindness, kidney failure, neuropathy, and atherosclerosis.^[9]

Secondary Structure

The α -chains form tight left-handed α -helices. The peptide bond backbone forms with the side chains of the amino acids pushing to the outside of the helix. Approximately three residues form a turn in the helix



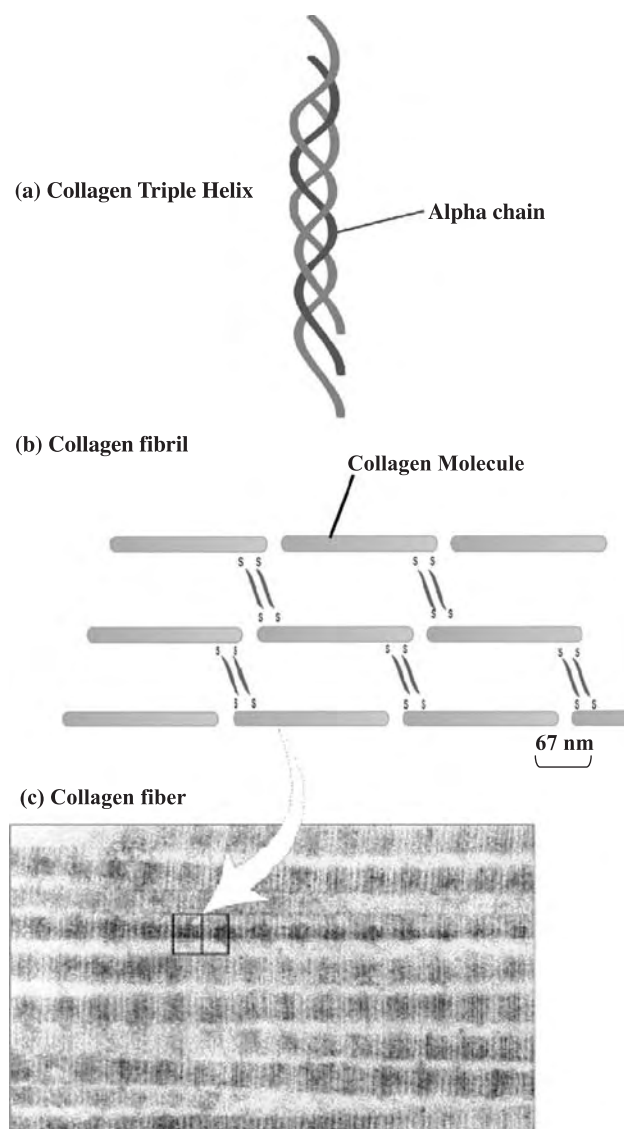


Fig. 1 The structure of collagen: (a) The basic structural unit of collagen is a triple-stranded helical molecule, and each basic unit is approximately 300 nm in length. (b) Collagen fibrils are formed with collagen units oriented side by side and adjacent strands positioned to overlap and form covalent cross-linking bonds (in dark) between the N-terminus of one strand and the C-terminus of the adjacent strand. The periodic pattern of collagen formation in fibril packing creates 67 nm end-to-end separations. (c) The periodic pattern gives an appearance of striations as seen in a scanning electron micrograph (SEM) of embryonic chick tendon. (SEM scan reproduced from *Molecular Cell Biology*, Fourth Ed. Lodish, Berk, Zipursky, Matsudaira, Baltimore, and Darnell, eds. New York: W. H. Freeman and Company, 2000, with permission of the publisher.)

with glycine positioned at every third residue in the (Gly-X-Y) motif. The Gly-residues align and form what is known as the clear line that winds in a left-handed fashion on the surface of the molecule.

Tertiary Structure

The Gly-clear line is a left-handed formation that helps stabilize the reciprocal formation of right-handed threads and vice versa.^[10] The triple helix between three collagen fibers occurs in the C- to N-terminal direction and is facilitated by a lack of steric hindrance due to the glycine rows forming a central axis. Point mutations in the X or Y positions can be tolerated, whereas mutations that replace Gly impair the formation of the triple helix, and most often result in one of many forms of connective tissue disorders collectively referred to as osteogenesis imperfecta.^[11] The type of triple helix that forms is dictated by the collagen and tissue type and can be either a homotrimer or a heterotrimer triple helix formation (Table 1).^[12] As collagen forms into its triple helix it resembles a rodlike structure that is both stiff and flexible (Fig. 1c), and can be over 300 nm long (~1.5 nm wide), with globular domains capping both N- and C-terminal ends. This form of secreted collagen is termed procollagen.

COLLAGEN SYNTHESIS

The biosynthetic pathway responsible for collagen production is complex.^[4,13] Each specific collagen type is encoded by a specific gene; the genes for all of the collagen types are found on a variety of chromosomes. As the messenger RNA (mRNA) for each collagen type is transcribed from the gene, or DNA blueprint, it undergoes many processing steps to produce a final code for that specific collagen type, called mRNA processing. Translation occurs once the final pro- α -chain mRNA is produced and attaches to the site of actual protein synthesis. The site of pro- α -chain mRNA translation is found on the membrane-bound ribosomes called the rough endoplasmic reticulum (rER) for function in the extracellular environment (Fig. 2).

Collagen, like most proteins that are destined for transport to the extracellular spaces for their function or activity, is produced initially as the larger precursor molecule called procollagen.^[14] Procollagen contains extension proteins on each end called amino and carboxy propetides that are unlike collagen.^[14] These nonhelical portions of the molecule increase solubility for ease of movement within the cell as it undergoes production. While inside the cell and when the procollagen peptides are intact, the molecule is about 1000 times more soluble than it is at a later stage when the extension peptides are removed.^[15] This high degree of solubility allows the procollagen molecule to be transported easily within the cell. It is moved by means of specialized structures called microtubules to the cell

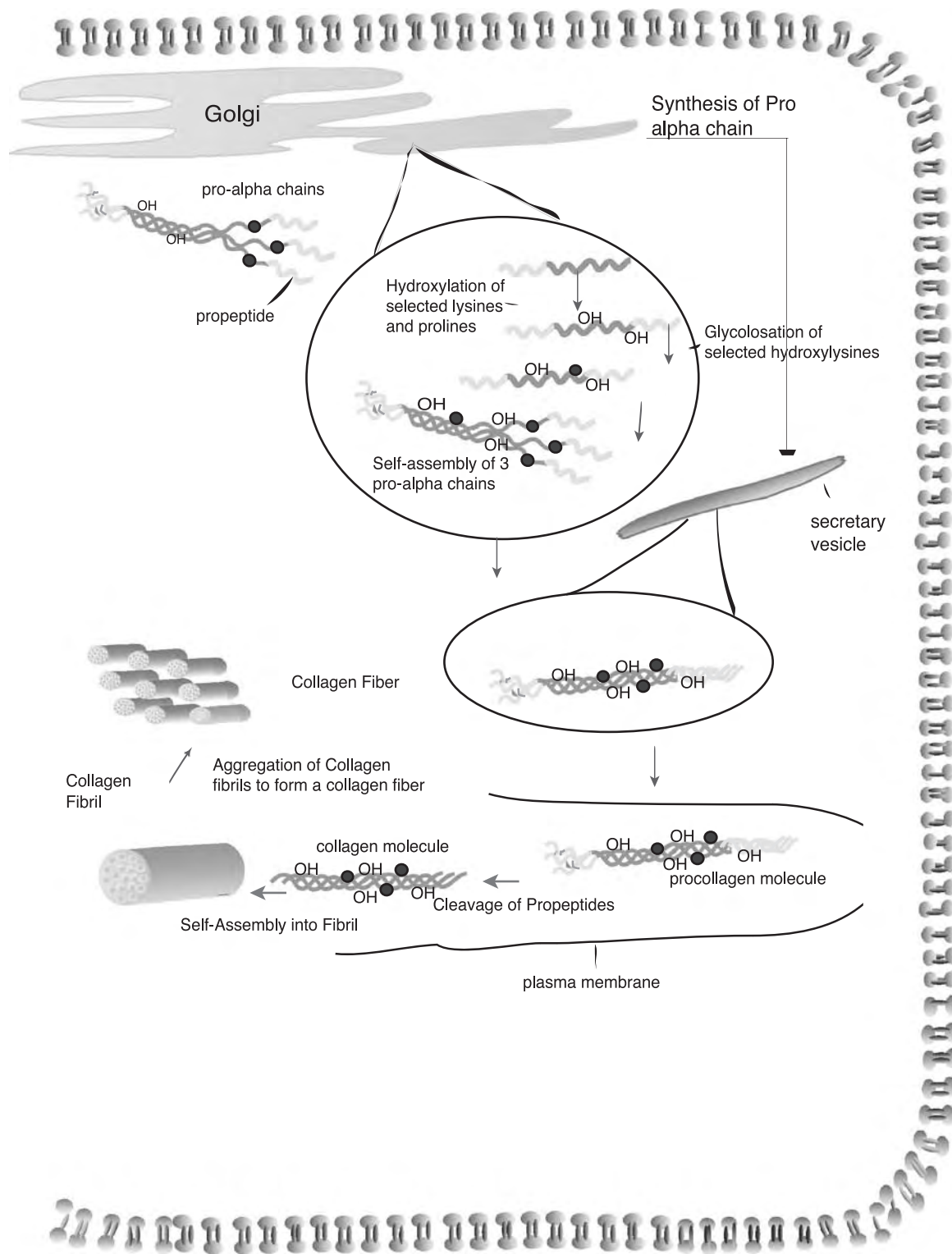


Fig. 2 The intracellular and extracellular events involved in the formation of a collagen fibril.

surface where it is secreted into the extracellular spaces.^[16] As the procollagen is secreted from the cell, it is acted upon by specialized enzymes called procollagen proteinases that remove the extension

peptides from the ends of the molecule.^[17] Portions of these digested end pieces are thought to reenter the cell and regulate the amount of collagen synthesis through a feedback type mechanism.^[18,19] The newly processed

molecule is referred to as collagen and now begins the important process of fiber formation.

On the amino- or N-terminal end, special disulfide bonds are formed among three procollagen chains and insure proper alignment of the chains. This step is called registration. Once registration occurs, then the three chains wrap around each other forming a stringlike structure. As collagen molecules are produced, they undergo many posttranslational modifications that take place in the Golgi compartment of the ER.^[4,13]

One of the first modifications to take place is the very critical step of hydroxylation of selected Pro and Lys amino acids in the newly synthesized procollagen protein (Fig. 2). Specific enzymes called hydroxylases are responsible for these important reactions needed to form hydroxyproline and hydroxylysine. The hydroxylase enzymes require vitamin C and iron as cofactors and molecular oxygen as a cosubstrate.^[20] If a patient is vitamin C deficient, the hydroxylation reaction is inhibited. In the absence of hydroxyproline, the collagen chains cannot initiate hydrogen bond formation to form a proper helical structure and the resultant molecule is weak and quickly destroyed.^[21] The end result is poor wound healing and the clinical condition is called scurvy.^[22] The current Recommended Dietary Allowance for vitamin C is 60 mg per day; however, research has shown 200 mg may be optimal.^[23,24]

Some of the newly formed hydroxylysine amino acids are glycosylated by the addition of sugars such as galactose and glucose.^[25] The enzymes that catalyze the glycosylation step, galactosyl and glucosyl transferases, require the trace metal manganese (Mn^{+2}). The glycosylation step imparts unique chemical and structural characteristics to the newly formed collagen molecule and may influence fibril size.^[26] It is of interest to note that the glycosylation enzymes are found with the highest activities in the very young and decrease as we age.^[27]

Fibrillogenesis

In the extracellular spaces, another posttranslational modification takes place as the triple helical collagen molecules (see Fig. 1) line up and form fibrils and then

fibers. The cross-link formation is promoted by another specialized enzyme called lysyl oxidase.^[28] The reactions place stable cross-links within (intramolecular) and between (intermolecular) the molecules (Fig. 3). Cross-link formation is the critical step giving the collagen fibers, on a per weight basis, tensile strength approaching that of steel.

The structural hierarchy of collagen can be visualized as resembling a piece of sewing thread for the individual collagen molecules. Many of these threads are wrapped around one another to form a string or fibril. The strings form cords, which associate to form ropes, and then interact to form cables or fibers. The structure can be compared to the steel rope cables on the Golden Gate bridge. The highly organized structure is responsible for the strength of tendon, ligament, bone, and dermis.

When the normal collagen in our tissues is injured and replaced by scar tissue, the connective tissue does not regain the original highly organized structure and is weaker. The maximum tensile strength of scar tissue is about 70 to 80% of the original.^[29] Collagen synthesis and remodeling continue at the wound site long after the injury. The body is constantly trying to remodel the scar tissue to achieve the original collagen ultrastructure. This remodeling involves ongoing collagen synthesis and collagen degradation. Interference with protein synthesis can cause the synthesis/degradation equilibrium to shift allowing increased collagen degradation. For example, patients who are malnourished or patients receiving chemotherapy may experience wound dehiscence because the wound site will become weak due to a shift in the balance toward collagen degradation. It is of interest to note that when wounds in the fetus heal, they do so in such a manner that the original collagen ultrastructure is achieved.^[30] If only we understood more about the biology and mechanisms responsible for the rapid and optimal wound healing response seen in the fetus, we would have greater insight into the management of adult wounds.^[31]

Spontaneous aggregation of the collagen molecules begins with the formation of microfibrils which aggregate to form fibrils. Fibrillogenesis is dependent on the glycoprotein content and the distribution of the glycosylated residues along the molecule. Any effects

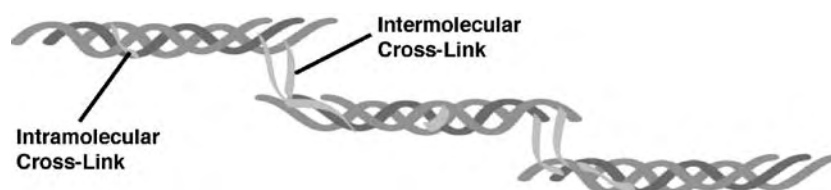


Fig. 3 The intramolecular and intermolecular cross-links formed within a collagen fibril.

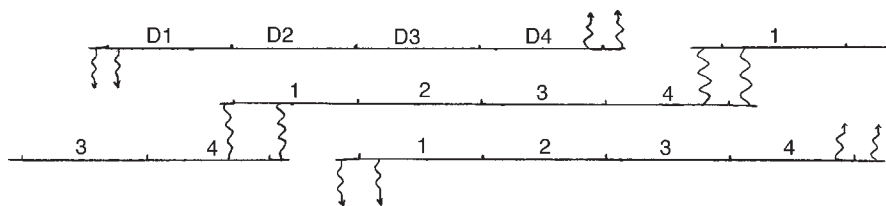


Fig. 4 D-Period banding of collagen molecules. (From Ref. [32].)

of copolymerization with different collagen Types within the fibril also contribute to differences in fibrillogenesis.^[32,33] The aggregation is also controlled by periodic distribution of polar charged amino acid residues and amino acids with hydrophobic side chains along the triple helix.^[33–36]

The differing types of amino acid R-groups include:

- nonpolar, i.e., glycine, alanine, proline.
- cationic, i.e., lysine, arginine.
- anionic, i.e., free carboxyls of aspartic and glutamic acid.
- polar ionic, i.e., residues of serine and hydroxyproline.^[35]

The collagen fibrils are the major source of strength for all connective tissue and range from 16 to 200 nm in diameter.^[34] The nonanastomosing fibrils form a $\frac{1}{4}$ stagger arrangement with a 67 nm periodic banding pattern which is termed D-period banding. D is equivalent to $\frac{1}{4}$ the molecular length of the collagen molecule, approximately 234 residues.^[35,37,38] There is a 0.4 D overlap which contributes to the D-period banding pattern and also contains the functional groups for cross-linking (see Figs. 3 and 4). The D4 period contains amino acid sequences critical for cell motility.^[39] The potential nonpolar and ionic interaction between the collagen molecules is optimized at multiples of D.^[32,35,36,40]

The kinetics of fibril formation is also dependent on pH, ionic strength, temperature, cross-links, aldehydes which form cross-links, the concentration of phosphate, and the telopeptide region.^[32] In cartilage, Type II collagen forms fibrils approximately 20 to 30 nm in diameter but does not form fibers as found in tendons. The tissues composed of collagen Type II are comprised of a network of thin fibrils.^[32,37] The $\alpha 1(\text{II})$ chain contains 6% hydroxylysine-linked carbohydrates

which is a comparatively large amount of hydrophilic carbohydrate groups when compared with other collagen molecules. The increase in the lateral spacing produced by the carbohydrate groups allows twice as much water to be dispersed among the molecules and the fibrils than does a typical collagen Type I molecule.^[32,35] The lateral spacing is achieved by a layering effect of collagen molecules in a microfibril. The most popular model for microfibril arrangement is the Smith model, in which five molecules are staggered by a D period in a layer (see Fig. 5).^[32,34] Recent advancements upon this theory have been demonstrated using atomic force microscopy.^[41] With this technology one can study fully hydrated tendon specimens to simultaneously demonstrate fibril-bound material and surface cross-binding, at the same time excluding artifacts typically associated with fixation and dehydration. This work predicts a compressed microfibril model forming a fibril of five-stranded microfibrils. The molecular segments follow an individual path within the gap zone and pack tightly in the overlap zone.

Cross-Links

The inherent tensile strength of the collagen fibrils is primarily dependent on the covalent nondisulfide cross-linking. Disulfide bonds are found in the propeptide regions of the procollagen molecule and play no role in intermolecular cross-linking; however, the aldehyde-derived cross-links are multifunctional.^[34] Although the cross-links are not visible by scanning electron microscopy (SEM), the cross-links can be thought of as welds used to bind and direct the building blocks of collagen molecules into a rope-like unit.^[35,40] Inhibition of cross-linking significantly increases the diameter and the volume density of collagen fibrils. However, the mechanical integrity of the

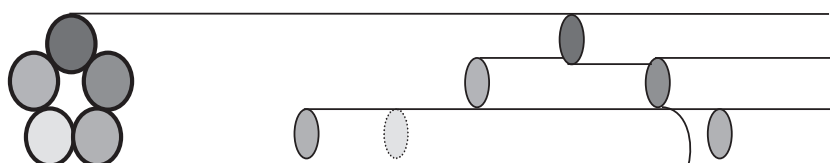


Fig. 5 Smith Model—Collagen Microfibril Arrangement. (From Ref. [33].)



thicker under-cross-linked fibrils is significantly less than that of normal collagen fibrils.^[42] A decrease in the percentage of cross-linking also contributes to an increased susceptibility to proteolytic degradation and subsequent solubilization.^[35,37] Cofactors are necessary to induce cross-linking of the fibrils by the conversion of lysine and hydroxylysine to allysine and hydroxyallysine. The cofactors include pyridoxal phosphate, Cu^{++} , O_2 , and lysyl oxidase.^[34-36] The spontaneity of the condensation reactions is mediated by the physical location of the molecules with respect to each other and their correct alignment. The aldimine product is a combination of lysine and hydroxylysine in which the lysine has been altered to allysine by the cofactors and subsequently combines with the hydroxylysine to form a terminal group.^[34] The aldehydes are restricted to the N-telopeptide and the helical region.^[36,37] The initial condensation reactions produce difunctional cross-links capable of linking two separate polypeptide chains.

COLLAGEN DEGRADATION

Of equal importance in the total picture of collagen metabolism is the complex process of collagen degradation. Normally the collagen in our connective tissues turns over at a very slow and controlled rate. However, during rapid growth and in disease states such as arthritis, cancer, and chronic non-healing ulcers, the extent of collagen degradation can be quite extensive. In normal healthy tissues where the collagen is fully hydroxylated and in a triple helical structure, the molecule is resistant to attack by most proteases. Under these normal healthy conditions only specialized enzymes called collagenases can attack the collagen molecule.^[43] The group of collagenases belongs to a family of enzymes called matrix metalloproteinases (MMP).

Many cells in our bodies can synthesize and release collagenases including fibroblasts, macrophages, neutrophils, osteoclasts, and tumor cells. One of the reasons that some neoplastic cells can be so invasive is because they release potent collagenases and can break down the collagen around them. They degrade the basement membranes of blood vessels and spread throughout the body. In chronic pressure ulcers, there is a massive invasion of neutrophils releasing a very potent collagenase called MMP-8 which is responsible for connective tissue breakdown and delayed healing.^[44] Some exciting new research suggests that members of the tetracycline family of antibiotics, such as doxycycline, when given systemically, and at low doses, may be useful to treat pressure ulcers by inhibiting MMP-8.^[45]

Collagen is particularly resistant to attacks by neutral proteases due to its function as the primary structural protein in the body. At neutral pH only specific collagenases cleave the native helix at a position about three-quarters of the way from the N-terminus.^[46] They are zinc-containing MMP of approximately 40–50 kDa and generally require calcium as a cofactor for full activity.^[47] Fibrils as aggregates of collagen molecules are degraded starting from the exterior. Collagenase binds tightly to triple helices at or near the surface, whereas molecules in the interior become accessible to enzymes in the course of the progressive degradation from the outside.^[48] The various types of collagen show different susceptibility to collagenolytic degradation.^[49] After the triple helix is disrupted, further degradation of the collagen molecules is facilitated by enzymes such as gelatinases and non specific proteinases which cleave the primary fragments into small peptides and amino acids. Elastases and cathepsins assist in proteolysis of collagen fibers with a higher degree of natural cross-linking based on their ability to digest nonhelical telopeptide regions.^[50]

Synthetic Cross-Linking

Naturally occurring intramolecular cross-links are disrupted during collagen extraction purification procedures. Upon subsequent processing, collagen does not molecularly re-cross-link; therefore, a synthetic cross-linking process must be performed.

Synthetic cross-linking prevents material degradation, increases resistance to cell-mediated contraction, and in some cases decreases immune reaction responses.^[51-55] Physical cross-linking methods involve processes allowing the collagen to polymerize without cytotoxic products, but may degrade or fragment the collagen. Chemical cross-linking methods involve polymerization with reagents potentially forming cytotoxic products.

Different methods of cross-linking affect the collagen material characteristics regarding cell adhesion qualities, degeneration, and mechanical properties which must be considered depending on the application.^[56]

Physical cross-linking methods included dehydrothermal treatments (DHT) and ultraviolet light (UV). DHT forms intermolecular cross-links by removing water from the collagen through condensation reactions.^[51,54,57] The high temperature required for cross-linking may partially denature the collagen.^[51,58] UV exposure produces radicals of unpaired electrons located in the nuclei of aromatic amino acid residues located on the α -chains, which then band to form cross-links.^[55,57] The process is limited by the number of aromatic amino acids and the attenuation of the UV

light by the collagen fibers.^[57] Overexposure to the UV light may partially denature the collagen fibers.^[55]

Chemical cross-linking methods include glutaraldehyde (GTA), 1-ethyl-3-(3-dimethylamino-propyl) carbodiimide (EDC or EDAC), or diphenylphosphorylazide (DPPA). GTA reacts with amino groups to form methylene bridges capable of spanning large intermolecular spaces. Disadvantages of this process include its reactivity during foreign body reactions in which the cross-links are dissolved and absorbed, unsealing the collagen material.^[54] The inflammatory response is associated with continued hydrolysis of free monomeric GTA, thereby increasing cytotoxicity.^[56,59] EDC involves many functional groups by forming amide bonds in the presence of lysine or hydroxylysine.^[51,55] In vivo testing has shown biocompatibility with the unbound carbodiimide eliminated from the collagen through simple washes and the conversion of the bound EDC into the form of nontoxic urea.^[56,59,60] DPPA synthesizes acryl azide on lateral carboxyl groups of aspartic and glutamic acid which then react with amino groups of lysine and hydroxy lysine to produce cross-links.^[61]

Extraction

Although collagen is ubiquitous in the mammalian body, those tissues rich in fibrous collagen such as skin and tendon are generally used as starting materials to generate collagen for use in implants, wound dressings, or drug delivery systems. The major impediment to dissolution of collagen type I from tissue is the presence of covalent cross-links between molecules. Collagen is an extremely stable protein and resists action of nonspecific proteases.^[62] In order to extract the protein from tissue it is necessary to solubilize and appropriate the covalent cross-links. One way to aid in extraction is to inhibit the cross-links from forming. Inhibitors of cross-links include a copper deficient diet, carbonyl reagents such as phenyl hydrazine and hydroxyl amine, and lathyrogens such as β -aminopropionitrile.^[34] In general, however, three different solvent reactions exist to extract collagen. The first employs a salt solvent (0.15–2.0 M NaCl, pH 7.5), the second uses a dilute organic acid (0.5 M acetic acid), and the third uses a dilute organic acid in conjunction with an acid protease (pepsin).^[34]

Neutral Salt-Soluble Collagen

The most commonly used solvents are neutral salt solution (0.15–2.0 M NaCl) or dilute acetic acid.^[63] Neutral salt solutions will extract freshly synthesized and negligibly cross-linked collagen molecules present in the tissue. The extracted material is purified by

dialysis, precipitation, and centrifugation. Most tissues have little or no salt-extractable collagen.

Acid-Soluble Collagen

Dilute acidic solvents, e.g., 0.5 M acetic acid at 4°C, citrate buffer, or hydrochloric acid at pH 2–3, are more efficient than neutral salt solutions. The intermolecular cross-links of the aldimine Type of collagen are dissociated by the dilute acids, and the repelling charges on the triple helices lead to swelling of fibrillar structures.^[64] Dilute acids will not disassociate less labile cross-links such as keto-imine bonds. Therefore, collagen from tissues containing higher percentages of keto-imine bonds, such as bone, cartilage, or material from older animals, has a lower solubility in dilute acid solvents.

Alkali- and Enzyme-Treated Collagen

Additional collagen material can be solubilized by treating connective tissue with an aqueous solution comprised of alkali hydroxide and alkali sulfate, for example, approximately 10% sodium hydroxide and 10% sodium sulfate, for approximately 48 h.^[65] Much higher yields than with acidic extraction can be achieved by taking advantage of the fact that the collagen triple helix is relatively resistant to proteases, that is, pronase[®], ficin, pepsin, or chymotrypsin, below approximately 20°C.^[66] The pepsin cleaves at carboxylic groups of aromatic amino acids such as phenylalanine and tyrosine. It will not cleave at bonds with valine, alanine, or glycine, which allows for the maintenance of the triple helix.^[67] Therefore, the selective cleavage of pepsin occurs at the N- and C-telopeptide nonhelical regions breaking the peptide bonds near cross-links and forming an atelopeptide collagen molecule capable of dissolving in aqueous solution.^[34,68] Polyacrylamide gel electrophoresis in sodium dodecyl sulfate is used extensively to identify collagens. Performed with relative ease, speed, and a high degree of sensitivity, it is capable of identifying microgram quantities instead of the milligram quantities necessary for other analyses.^[34]

Denaturation

Denaturation of a protein is a result of damage to the internal structure of the molecule. The rearrangement of peptide chains is due to rupture of weak interchain bonds by heat, urea, collagenases, or acids.^[62] The melting temperature (T_M) of collagen is proportional to the body temperature of the particular species. The T_M is stable enough to remain native at common



temperatures but not so stable as to prevent collagen degradation and turnover. The T_M is regulated by the proline and hydroxyproline content.^[34,62] Mammalian collagenases separate the α -chain at specific and limited sites (i.e., Gly-Ile) to denature the molecule, and nonspecific proteolytic enzymes take over and digest the collagen fragments.^[32,34] Bacterial collagenases sever the α -chain at several sites.^[32] As stated, collagen is slightly susceptible to proteolytic enzymes under physiological conditions, and resorption of collagenous tissues does occur normally. However, resorption is dependent on the ability of the collagenase to degrade the collagen fibril and upon other proteases to depolymerize the collagen fibrils. Highly cross-linked collagens are resistant to digestion, but can be degraded with an increase in temperature.^[34,40] Due to the amino acid sequence, the degree of glycosylation, and the number of cross-links, type II collagen degrades five to six times more slowly than either type I or type III collagen.^[34]

CONCLUSION

Collagen metabolism is indeed one of the most complex and highly regulated processes in our bodies. As we move forward in the future to design new strategies and technologies to treat the many challenging clinical problems associated with wound healing, we need to keep in mind how our connective tissues are assembled and how they are remodeled. Collagen is an attractive biomaterial because it is a natural material of low immunogenicity and is therefore seen by the body as a normal constituent rather than as foreign matter.^[69] The versatility of the collagen matrix allows for it to be processed into a number of forms such as sheets, tubes, sponges, electrospun matrices, powders, fleeces, injectable solutions, and dispersions, all of which have extensive potential in medical procedures.^[70-72] Future work with collagen-based products will be at the forefront of a variety of applications in systems for drug delivery, ophthalmology, wound and burn dressing, tumor treatment, and cartilage and musculoskeletal tissue engineering.

ACKNOWLEDGMENT

The authors graciously thank Elisabeth A. Beckman for her assistance with artwork in Figs. 1–3.

ARTICLES OF FURTHER INTEREST

Biologic and Synthetic Apatites; Biomimetic Materials; Collagen Fixation; Collagen Processing; Elastin;

Fabrics; Matrix Metalloproteinases; Polyepoxy Compound Fixation; Small-Angle Light Scattering Methods for Soft Connective Tissue Structural Analysis

REFERENCES

- Burgeson, R.E. The collagens of skin. *Curr. Probl. Dermatol.* **1987**, *17*, 61–75.
- Prockop, D.J. What holds us together? Why do some of us fall apart? What can we do about it? *Matrix. Biol.* **1998**, *16*, 19–28.
- van der Rest, M.; Garrone, R. Collagen family of proteins. *FASEB J.* **1991**, *5*, 2814–2823.
- Prockop, D.J.; Kivirikko, K.I. Collagens: Molecular biology, diseases, and potentials for therapy. *Annu. Rev. Biochem.* **1995**, *64*, 403–434.
- Miller, E.J.; Gay, S. The collagens: An overview and update. *Methods Enzymol.* **1987**, *144*, 3–41.
- Clore, J.N.; Cohen, I.K.; Diegelmann, R.F. Quantitation of collagen types I and III during wound healing in rat skin. *Proc. Soc. Exp. Biol. Med.* **1979**, *161*, 337–340.
- Graham, M.F.; Diegelmann, R.F.; Elson, C.O. Collagen content and types in the intestinal strictures of Crohn's disease. *Gastroenterology* **1988**, *94*, 257–265.
- Ottani, V.; Martini, D.; Franchi, M.; Ruggeri, A.; Raspanti, M. Hierarchical structures in fibrillar collagens. *Micron* **2002**, *33*, 587–596.
- Reiser, K.M.; Amigable, M.A.; Last, J.A. Non-enzymatic glycation of type I collagen. *J. Biol. Chem.* **1992**, *267*, 24207–24216.
- Jones, E.Y.; Miller, A. Analysis of structural design features in collagen. *J. Mol. Biol.* **1991**, *218*, 209–219.
- Royce, P.M.; Steinmann, B. *Connective Tissue and its Heritable Disorders: Molecular, Genetic, and Medical Aspects*; Wiley: New York, 2001.
- Ayad, S.; Boot-Handford, R.P.; Humphries, M.J.; Kadler, K.E.; Shuttleworth, C.A. *The Extracellular Matrix Facts Book*, 2nd Ed.; Academic Press: San Diego, 1998.
- Kivirikko, K.I.; Risteli, L. Biosynthesis of collagen and its alterations in pathological states. *Med. Biol.* **1976**, *54*, 159–186.
- Bellamy, G.; Bornstein, P. Evidence for procollagen, a biosynthetic precursors of collagen. *Proc. Natl. Acad. Sci. U. S. A.* **1971**, *68*, 1138–1142.
- Prockop, D.J.; Kivirikko, K.I.; Tuderman, L.; Guzman, N.A. The biosynthesis of collagen and its disorders (first of two parts). *N. Engl. J. Med.* **1979**, *301*, 13–23.
- Diegelmann, R.F.; Peterkofsky, B. Inhibition of collagen secretion from bone and Referencescultured fibroblasts by microtubular disruptive drugs. *Proc. Natl. Acad. Sci. U. S. A.* **1972**, *69*, 892–896.
- Lapiere, C.M.; Lenaers, A.; Kohn, L.D. Procollagen peptidase: An enzyme excising the coordination peptides of procollagen. *Proc. Natl. Acad. Sci. U. S. A.* **1971**, *68*, 3054–3058.
- Lichtenstein, J.R.; Martin, G.R.; Kohn, L.D.; Byers, P.H.; McKusick, V.A. Defect in conversion of

- procollagen to collagen in a form of Ehlers-Danlos syndrome. *Science* **1973**, *182*, 298–300.
19. Wiestner, M.; Krieg, T.; Horlein, D.; Glanville, R.W.; Fietzek, P.; Muller, P.K. Inhibiting effect of procollagen peptides on collagen biosynthesis in fibroblast cultures. *J. Biol. Chem.* **1979**, *254*, 7016–7023.
 20. Mussini, E.; Hutton, J.J.; Udenfriend, S. Collagen proline hydroxylase in wound healing, granuloma formation, scurvy, and growth. *Science* **1967**, *157*, 927–929.
 21. Bienkowski, R.S.; Baum, B.J.; Crystal, R.G. Fibroblasts degrade newly synthesised collagen within the cell before secretion. *Nature* **1978**, *276*, 413–416.
 22. Hirschmann, J.V.; Raugi, G.J. Adult scurvy. *J. Am. Acad. Dermatol.* **1999**, *41*, 895–910.
 23. Gross, R.L. The effect of ascorbate on wound healing. *Int. Ophthalmol. Clin.* **2000**, *40*, 51–57.
 24. Ayello, E.A.; Thomas, D.R.; Litchford, M.A. Nutritional aspects of wound healing. *Home Healthc Nurse* **1999**, *17*, 719–730.
 25. Anttinen, H.; Myllyla, R.; Kivirikko, K.I. Further characterization of galactosylhydroxylsyl glucosyltransferase from chick embryos. Amino acid composition and acceptor specificity. *Biochem. J.* **1978**, *175*, 737–742.
 26. Kivirikko, K.I.; Myllyla, R. Collagen glycosyltransferases. *Int. Rev. Connect. Tissue Res.* **1979**, *8*, 23–72.
 27. Anttinen, H.; Oikarinen, A.; Kivirikko, K.I. Age-related changes in human skin collagen galactosyltransferase and collagen glucosyltransferase activities. *Clin. Chim. Acta* **1977**, *76*, 95–101.
 28. Bailey, A.J.; Robins, S.P.; Balian, G. Biological significance of the intermolecular crosslinks of collagen. *Nature* **1974**, *251*, 105–109.
 29. Schilling, J.A. Wound healing. *Physiol. Rev.* **1968**, *48*, 374–423.
 30. Mast, B.A.; Flood, L.C.; Haynes, J.H.; Depalma, R.L.; Cohen, I.K.; Diegelmann, R.F.; Krummel, T.K. Hyaluronic acid is a major component of the matrix of fetal rabbit skin and wounds: Implications for healing by regeneration. *Matrix* **1991**, *11*, 63–68.
 31. Mast, B.A.; Diegelmann, R.F.; Krummel, T.M.; Cohen, I.K. Scarless wound healing in the mammalian fetus. *Surg. Gynecol. Obstet.* **1992**, *174*, 441–451.
 32. Nimni, M.E.; Olsen, B.R. *Biochemistry and Biomechanics*; Collagen Ed.; CRC Press: Boca Raton, FL, 1988; Vol. 2.
 33. Kuhn, K. Structure and biochemistry of collagen. *Aesthet. Plast. Surg.* **1985**, *9*, 141–144.
 34. Piez, K.A.; Reddi, A.H. *Extracellular Matrix Biochemistry*; Elsevier: New York, 1984.
 35. Kanke, Y.; Bashey, R.I.; Mori, Y. Biochemistry of collagen. Present knowledge. *N. Y. State J. Med.* **1981**, *81*, 1045–1052.
 36. Veis, A. The biochemistry of collagen. *Ann. Clin. Lab. Sci.* **1975**, *5*, 123–131.
 37. Riesle, J.; Hollander, A.P.; Langer, R.; Freed, L.E.; Vunjak-Novakovic, G. Collagen in tissue-engineered cartilage: Types structure, and crosslinks. *J. Cell. Biochem.* **1998**, *71*, 313–327.
 38. Olsen, B.R. New Insights into the function of collagens from genetic analysis. *Curr. Biol.* **1995**, *7*, 720–727.
 39. Fertala, A.; Han, W.B.; Ko, F.K. Mapping critical sites in collagen II for rational design of gene-engineered proteins for cell-supporting materials. *J. Biomed. Mater. Res.* **2001**, *57*, 48–58.
 40. Bornstein, P.; Byers, P.H. *Collagen Metabolism*; Upjohn Co.: Kalamazoo, MI, 1980.
 41. Friess, W. Collagen–biomaterial for drug delivery. *Eur. J. Pharm. Biopharm.* **1998**, *45*, 113–136.
 42. Wong, M.; Siegrist, M.; Gaschen, V.; Park, Y.; Graber, W.; Studer, D. Collagen fibrillogenesis by chondrocytes in alginate. *Tissue Eng.* **2002**, *8*, 979–987.
 43. Gross, J.; Harper, E.; Harris, E.D. Animal collagenases: Specificity of action, and structures of the substrate cleavage site. *Biochem. Biophys. Res. Commun.* **1974**, *61*, 605–612.
 44. Yager, D.R.; Zhang, L.Y.; Liang, H.X.; Diegelmann, R.F.; Cohen, I.K. Wound fluids from human pressure ulcers contain elevated matrix metalloproteinase levels and activity compared to surgical wound fluids. *J. Invest. Dermatol.* **1996**, *107*, 743–748.
 45. Golub, L.M.; Lee, H.M.; Ryan, M.E.; Giannobile, W.V.; Payne, J.; Sorsa, T. Tetracyclines inhibit connective tissue breakdown by multiple non-antimicrobial mechanisms. *Adv. Dent. Res.* **1998**, *12*, 12–26.
 46. Cawston, T.E.; Murphy, G. Mammalian Collagenases: Methods in Enzymology. In *Proteolytic Enzymes*; Lorand, L., Ed.; Academic Press: London, 1981; 80, 711–722.
 47. Stricklin, G.P.; Hibbs, M.S. Biochemistry and Physiology of Mammalian Collagenases. In *Collagen–Biochemistry*; Nimni, M.E., Ed.; CRC Press: Boca Raton, FL, 1988; 187–206.
 48. Welgus, H.G.; Jeffrey, J.J.; Stricklin, G.P.; Roswit, W.T.; Eisen, A.Z. Characteristics of the action of human skin fibroblast collagenase on fibrillar collagen. *J. Biol. Chem.* **1980**, *255*, 6806–6813.
 49. Woolley, D.E. Mammalian Collagenases. In *Extracellular Matrix Biochemistry*; Piez, K.A., Reddi, A.H., Eds.; Elsevier: New York, 1984; 119–158.
 50. Kucharz, E.J. *The Collagens: Biochemistry and Pathophysiology*; Springer-Verlag: Berlin, 1992; 55–67.
 51. Pieper, J.S.; Oosterhof, A.; Dijkstra, P.H.; Veerkamp, J.H.; van Kuppevelt, T.H. Preparation and characterization of porous crosslinked collagenous matrices containing bioavailable chondroitin sulphate. *Biomaterials* **1999**, *20*, 847–858.
 52. Nehrer, S.; Breinan, H.A.; Ramappa, A.; Young, G.; Minas, T.; Shortkroff, S.; Louie, L.K.; Sledge, C.B.; Yannas, I.V.; Spector, M. Matrix collagen type and pore size influence behaviour of seeded canine chondrocytes. *Biomaterials* **1997**, *18*, 769–776.
 53. Lee, C.R.; Grodzinsky, A.J.; Spector, M. The effects of crosslinking of collagen-glycosaminoglycan scaffolds on compressive stiffness, chondrocyte-mediated contraction, proliferation and biosynthesis. *Biomaterials* **2001**, *22*, 3145–3154.
 54. Ma, X.H.; Noishiki, Y.; Yamane, Y.; Iwai, Y.; Marato, D.; Matsumoto, A. Thermal cross-linking for biologically degradable materials. Preliminary report. *ASAIO J.* **1996**, *42*, 866–871.



55. Weadock, K.; Olson, R.M.; Silver, F.H. Evaluation of collagen crosslinking techniques. *Biomater. Med. Dev. Artif. Organs* **1984**, *11*, 293–318.
56. Kato, Y.; Silver, F. Formation of continuous collagen fibres: Evaluation of biocompatibility and mechanical properties. *Biomaterials* **1990**, *11*, 169–175.
57. Weadock, K.S.; Miller, E.J.; Bellincampi, L.D.; Zawadsky, J.P.; Dunn, M.G. Physical crosslinking of collagen fibers: Comparison of ultraviolet irradiation and dehydrothermal treatment. *J. Biomed. Mater. Res.* **1995**, *29*, 1373–1379.
58. Nehrer, S.; Breinan, H.A.; Ramappa, A.; Hsu, H.P.; Minas, T.; Shortkroff, S.; Sledge, C.B.; Yannas, I.V.; Spector, M. Chondrocyte-seeded collagen matrices implanted in a chondral defect in a canine model. *Biomaterials* **1998**, *19*, 2313–2328.
59. Kato, Y.; Christiansen, D.; Hahn, R.; Shieh, S.; Goldstein, J.; Silver, F. Mechanical properties of collagen fibres: A comparison of reconstituted and rat tail tendon fibres. *Biomaterials* **1989**, *10*, 38–42.
60. Cavallaro, J.; Kemp, P.; Kraus, K. Collagen fabrics as biomaterials. *Biotechnol. Bioeng.* **1993**, *43*, 781–791.
61. Petite, H.; Frei, V.; Huc, A.; Herbage, D. Use of diphenylphosphorylazide for cross-linking collagen-based biomaterials. *J. Biomed. Mater. Res.* **1994**, *28*, 159–165.
62. Gustavson, K.H. *The Chemistry and Reactivity of Collagen*; Academic Press: New York, 1956.
63. Fielding, A.M. Preparation of Neutral Salt Soluble Collagen. In *The Methodology of Connective Tissue Research*; Hall, D.A., Ed.; Joynson–Bruvvers: Oxford, 1976; 9–12.
64. Trelstad, R.L. Immunology of Collagens. In *Immunochemistry of the Extracellular Matrix*; Furthmayer, H., Ed.; CRC Press: Boca Raton, FL, 1982; Vol. 1, 32–39.
65. Chandrakasan, G.; Torchia, D.A.; Piez, K.A. Preparation of intact monomeric collagen from rat tail tendon and skin and the structure of the nonhelical ends in solution. *J. Biol. Chem.* **1976**, *251*, 6062–6067.
66. Piez, K.A. Molecular and Aggregate Structures of the Collagens. In *Extracellular Matrix Biochemistry*; Piez, K.A., Reddi, A.H., Eds.; Elsevier: New York, 1984; 1–40.
67. Hall, B.K. *Cartilage*; Academic Press: New York, 1983.
68. Lee, C.H.; Singla, A.; Lee, Y. Biomedical applications of collagen. *Int. J. Pharm.* **2001**, *221*, 1–22.
69. Nimni, M.E.; Harkness, R.D. Molecular Structures and Functions of Collagen. In *Collagen–Biochemistry*; Nimni, M.E., Ed.; CRC Press: Boca Raton, FL, 1988; 1–79.
70. Chvapil, M.; Kronentahl, R.L.; van Winkle, W., Jr. Medical and Surgical Applications of Collagen. In *International Review of Connective Tissue Research*; Hall, D.A., Jackson, D.S., Eds.; Academic Press: New York, 1973; 1–61.
71. Gorham, S.D. Collagen. In *Biomaterials*; Byrom, V., Ed.; Stockton Press: New York, 1991; 55–122.
72. Thies, C. Microcapsules as drug delivery devices. *Crit. Rev. Biomed. Eng.* **1982**, *8*, 335–383.

Collagen Fixation

Thomas J. Koob

Skeletal Biology Section, The Centre for Research in Skeletal Development and Pediatric Orthopaedics, Shriners Hospitals for Children, Tampa, Florida, U.S.A.



INTRODUCTION

Collagenous tissue grafts and artificial constructs produced from reconstituted collagen fibers have been examined for potential use as bioprostheses for over forty years. Tissues, both allografts and xenografts, have warranted systematic investigation because the inherent collagen fiber system retains its tissue-specific shape and organization coupled with functional biomechanical properties. Grafts generally require processing to eliminate undesirable and potentially problematic compounds such as cell-associated molecules, enzymes, nucleotides, fats, etc. Methods to render grafts more biocompatible and long lasting have been developed over the past several decades and remain an active area of research. In contrast to grafts, reassembled fibers can be fabricated to satisfy mechanical requisites of the intended tissue replacement with the added advantage of allowing incorporation of additional macromolecules, therapeutic agents, and autologous or stem cells. Numerous innovative methods using diverse forms of collagen have been developed and optimized recently for fabricating functional collagen-based structures.

Both forms of bioprosthesis, whether a tissue or engineered construct, must be treated to reduce antigenicity and foreign body response, to attenuate degradation by host cells, or to enhance biomechanical properties. The basic strategy of most treatment protocols developed to date centers on the introduction of stable, covalent, intermolecular cross-links between collagen fibrils. Two related criteria are essential for the implementation of this strategy. The collagen fibrils must contain the reactive group, typically an amino acid side chain, targeted by the chemistry of the cross-linking agent and coupling mechanism. Second, the conditions, including temperature, pH, and nature of the solvent, must not adversely affect the inherent collagen molecules, fibrils, fiber systems, or properties of the bioprosthesis. Additional considerations might also be necessary depending on the cross-linking agent employed, e.g., residual reactants, unreacted intermediates, and byproducts of the reaction, all of which might be detrimental to the performance of the bioprosthesis in vivo.

What has proven most demanding is designing treatment strategies that are entirely biocompatible,

particularly with respect to cytotoxicity, antigenicity, and induction of a foreign body response. Chemical cross-linking agents such as formaldehyde, glutaraldehyde, and epoxy compounds, while effective at cross-linking collagen matrices, are themselves cytotoxic, and their reaction products in the cross-linked bioprostheses can be problematic in vivo. Nevertheless, advanced research on these strategies may in the future provide novel materials for biomedical engineering. New strategies incorporating natural biological cross-linking compounds or inducing cross-link formation without assimilation of foreign agents are emerging that may be as biomechanically effective as conventional cross-linking agents. Plant-derived compounds such as genipin and nordihydroguaiaretic acid are two examples of biological compounds that not only effectively improve the functional properties of collagenous materials but also appear to be biocompatible. An alternative method is to catalyze cross-link formation between reactive groups within the collagen fibrillar system. Such treatments include carbodiimide induction of isopeptide bond formation, ultraviolet light, and dehydrothermal treatments, all of which improve the properties of collagenous constructs crucial for use as bioprostheses.

Research and development of fixation strategies for collagenous materials have escalated enormously concomitant with the burgeoning fields of biomedical and tissue engineering. New strategies employing natural chemistries and biomimetic approaches will emerge that may supersede traditional methodologies. The present article reviews the current status of traditional approaches and describes potentially applicable novel fixation strategies.

COLLAGENOUS TISSUES AND MATERIALS

Fibrous assemblies of collagen molecules provide the major structural support of all vertebrate tissues. The molecular type of collagen and organization of these assemblies, however, are tissue-specific. In tendon and ligament, the predominantly type I collagen fibrils form fibers and fiber bundles that are aligned nearly parallel to the direction of load. In dermis, fibrils of type I collagen form loose networks. In bone, type I collagen forms the fibrillar phase of the mineralized

cortical and trabecular composites. Type II collagen in cartilage forms a dense network array of individual thin fibrils that constrain the proteoglycan-rich, inter-fibrillar matrix. In cornea, the thin type I collagen fibrils (co-polymerized with type V collagen) are arrayed parallel in each lamella with alternating layers arranged orthogonally with respect to the orientation of the fibrils. These examples illustrate the diverse arrangements of natural collagenous scaffolds to attain specific physicochemical properties and serve as archetypes for bioengineering designs.

Natural Collagen Fiber Systems

In order to devise strategies for using collagen in the development of advanced biomaterials for biomedical engineering, it is first necessary to understand the natural mechanisms underlying collagen fibril assembly and cross-linking pathways, the physicochemical properties of assembled fibril systems, and, thereby, appropriate means of obtaining diverse and potentially exploitable collagenous materials. Of principal concern here are the fibril-forming collagens that make up the extracellular matrices in collagen-dominated tissues and that can be obtained and purified in relatively large amounts (the interstitial collagen types I, II, and III). Other fibril-forming collagens,

types V and XI, are beyond the scope of the present article, as are the other twenty or more genetic types of collagen that are present in humans.^[1,2]

Interstitial collagen molecules within fibrils consist of a triple helix of three polypeptide chains (α -chains) with short, flanking, nonhelical telopeptides (Fig. 1). These molecules spontaneously assemble into fibrils with the characteristic D-period cross-striated repeat pattern measuring approximately 67 nm. Following incorporation into the fibril, covalent, intermolecular, lysine-derived cross-links form between the telopeptides and helical domains of neighboring molecules. The cross-linked fibril is the fundamental unit of mature collagen fiber systems within all connective tissues.

Collagen fibrils are organized into tissue-specific suprafibrillar architectures.^[3] Individual fibrils form a network in tissues such as skin, cartilage, and cornea. Fibrils are aligned and bundled into fibers and these in turn are aligned and bundled into fascicles in tissues such as tendon and ligament. Intact, native collagen fibrils have never been isolated from normal vertebrate tissues, indicating that their hierarchical organization is crucial to the tissues' structural properties. In addition, noncovalent as well as covalent interactions may provide additional stabilizing bridges linking fibrils where they are appropriately aligned. The nature of functionally relevant interfibrillar interactions remains an active area of connective tissue research and may

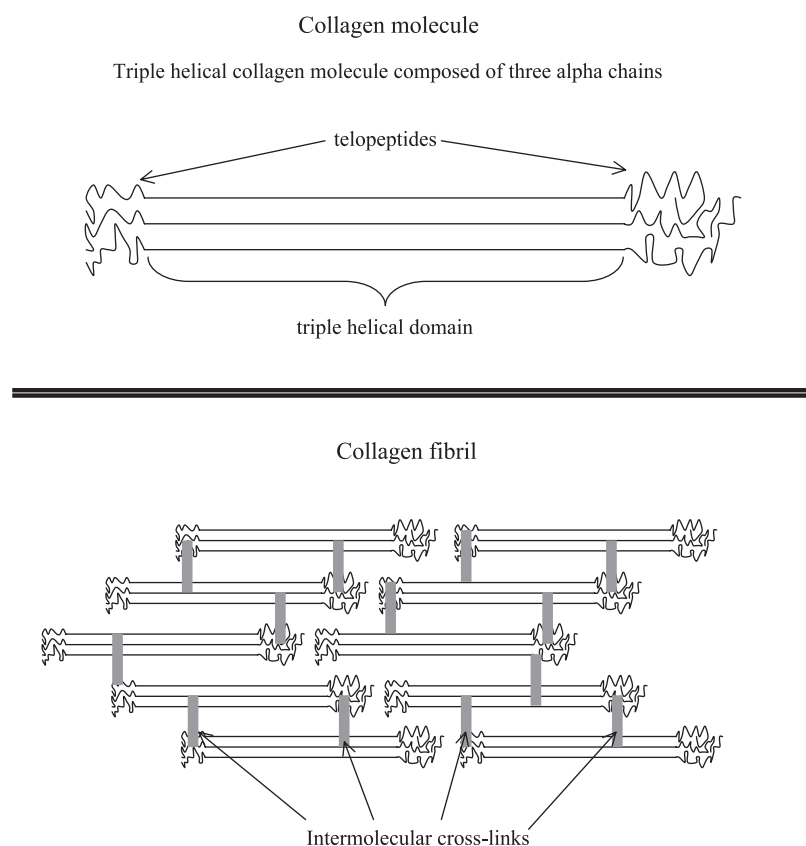


Fig. 1 Representative structure of fibril-forming, interstitial collagens and the organization of the fibrils they form.

eventually have important ramifications for collagen fixation strategies.

Collagen Scaffolds

Tissue grafts

Various forms of collagen have been examined for potential use as bioprosthetic devices and scaffolds for biomedical engineering. In the absence of synthetic mechanisms or molecular techniques to produce biomimetic collagen in any of its forms, the principal means of taking advantage of collagen is to use the native structure. Three forms are readily applicable for biomedical applications: collagenous fibril systems as they exist in tissue grafts, processed insoluble collagen, and purified collagen molecules that can be reassembled to satisfy specific geometrical and biomechanical objectives. Techniques to use each of these natural biomaterials have been optimized to various degrees. Before describing the strategy underlying implementation of these techniques, the distinct properties of each class of material must be understood.

Tissue grafts for use in humans can be obtained from several sources and each of these has advantages as well as drawbacks or limitations. Host-derived autogenous grafts function quite well in many instances and eliminate complications associated with immune reactions, foreign body response, and eventual rejection. They retain the native biomechanical properties of the extracellular matrix. Therefore, autogenous grafts do not require chemical fixation. However, autografting of collagenous tissues is limited by the paucity of sources with the added consideration that the source of the grafted tissue must not be compromised by resection of the graft itself. Examples of autogenous grafts used with success include skin grafts for treating severely burned patients, tendons for anterior cruciate ligament repair, and blood vessels for replacement of atherosclerotic arteries.

Tissue grafts derived from human cadaveric donors, i.e., allografts, constitute a potentially plentiful source of structural materials for use in the surgical repair of the musculoskeletal system. In contrast to vital organs for transplantation that are harvested alive and sterile, and utilized without modification, structural tissues such as bone and soft connective tissues—tendons, ligaments, and knee menisci—must be processed to remove blood, cells, soluble proteins, and enzymes before they can be used *in vivo*. Nevertheless, if treated properly, allografts retain much of their intrinsic structural and biomechanical properties. However, processed allografts must also be treated to eliminate immunogenic reactions and to reduce host-mediated breakdown. The only FDA-approved allografts that

are in general use today are bone-derived materials including bone chips and demineralized bone powder. No soft skeletal tissue has attained acceptable performance, due primarily to inflammatory responses and premature catabolism by invading cells.

Grafts originating from nonhuman species, termed xenografts, offer yet another source of collagenous scaffolds for use in tissue augmentation and repair. In contrast to allografts, there is potentially an unlimited supply of xenograft materials from a variety of animals, but they present a greater challenge *in vivo* with respect to antigenicity and induction of a foreign body response, including rapid degradation and removal. Xenografts must be processed even more stringently than allografts to eliminate potentially antigenic compounds and fixed to mask antigenic sites and to reduce host-mediated biodegradation. Fixation strategies to render xenografts more biocompatible have focused on methods for removal of nonstructural elements and chemical fixation of the collagen scaffold.

Although there has been a great deal of clinical and laboratory work exploring the potential use of fixed xenografts for musculoskeletal and vascular system applications, glutaraldehyde-fixed pig heart valves are the only xenograft bioprostheses in general use in humans at this time. Development and verification of the protocols that have enabled use of pig heart valves have a long and successful history. The most relevant advances will be described under specific fixation strategies to follow.

Collagen fibers

Naturally cross-linked collagen fibers purified to remove noncollagenous compounds provide effective starting materials for fabrication of collagen-based bioprostheses. This material is generally referred to as insoluble collagen and is commercially available. As long as the collagen is not denatured by the extraction procedure, the purified fibers retain the biomechanical properties that give tissues their particular functional properties and are thus like allografts. Concerns about biocompatibility are minimized due to the exhaustive extraction that leaves only the terminally cross-linked collagen. Unlike allografts, however, the tissues from which the fibers are obtained are mechanically fragmented, producing small segments of the fiber system. Fabrication of bioprostheses from these segments requires reassembly and fixation. Tissues that are used to produce insoluble collagen include tendons (Achilles) and corium, and these are predominantly from nonhuman species. The possibility of using human allograft tendons or other collagen-rich tissues as a source of insoluble collagen has not been systematically examined.



Because biomaterials created from insoluble collagen are essentially xenografts, they must be treated or fixed to reduce antigenicity and retard degradation *in vivo*. The same methods that have been applied to allografts and xenografts have been examined for their effects on insoluble collagen and its derivatives. In addition, following assembly of the fiber segments, the construct must be fixed to join the segments and thereby augment its biomechanical properties. One significant advantage of using insoluble collagen as opposed to grafts is that the fibers can be assembled into variably sized structures with geometries that satisfy specific physical or mechanical requirements. Another advantage is that other structural macromolecules or therapeutic agents can be incorporated into the construct during the fabrication process. Insoluble collagen fibers have been used to produce biomaterials for ligament and tendon repair, ocular implants and dressings, drug delivery vehicles, and as scaffolds for tissue engineering.

Reconstituted fibrils

Intact native collagen molecules containing the telopeptides can be obtained from tissues of fetal and growing vertebrates by acid extraction. Extraction of the newly synthesized collagen is possible because the mature intermolecular covalent cross-links have not yet formed or have progressed on the allysine pathway to a stage where they remain acid-labile. Since acid extraction is limited to the newly incorporated molecules, only a small proportion of the collagen is extractable from the tissues. More collagen is extracted from growing animals; only small amounts of collagen can be extracted from mature animals.

Substantially more collagen can be obtained from fetal and neonatal animals by enzyme extraction. The triple helical domain of the native collagen molecule is obtained from connective tissues by pepsin treatment in acid. Pepsin cleaves in the telopeptide region of the molecule inside the cross-linking domain but outside the helical domain, thereby liberating the molecule.

Acid-extracted and pepsin-solubilized collagen is purified by salt fractionation in acetic acid. The extract containing acid-soluble or pepsin-extracted collagen is brought to 0.7 M NaCl by addition of solid or dissolved NaCl in acetic acid. The major fibrillar collagens precipitate and are collected by centrifugation. The collagen is redissolved in acetic acid and precipitated a second time. The precipitate is redissolved in acetic acid and residual salt is removed by dialysis. The purity of the collagen preparation is assessed by polyacrylamide gel electrophoresis and amino acid analysis. Of note here is that bovine collagen prepared

in this way does not elicit significant antigenic response and has been FDA-approved for use in humans.

Both intact and pepsin-solubilized collagen molecules spontaneously form fibrils *in vitro* at neutral pH at 37°C. These fibrils are ultrastructurally identical to native fibrils *in vivo*. The purified collagen molecules dissolved in acid are neutralized, generally by dialysis, into phosphate-buffered saline. The molecules are soluble at neutral pH at 4°C. To form fibrils, the temperature of the collagen preparation is increased to 37°C. Because the initial collagen solution is a viscous liquid, it can be easily loaded into molds, vessels, or chambers for specific geometries. Following fibrillogenesis, the viscous liquid is transformed into a loose gel. However, gels formed this way are relatively dilute and fragile. The subsequent treatment of reconstituted collagen gels depends on the intended use of the biomaterial. However, in virtually all cases, fixation is required at some stage to stabilize the collagen phase of the construct and improve its biomechanical properties. Reconstituted collagen fibrils have been used to fabricate fibers for tendon and ligament repair and for collagen scaffolds for tissue engineering.

COLLAGEN FIXATION STRATEGIES

There are four fundamental strategies for fixing collagenous tissue grafts and biomaterials constructed of processed fibers or purified collagen. The traditional approach involves the introduction of exogenous chemical cross-linking agents that covalently couple neighboring collagen fibrils using targeted reactive moieties in the collagen fibrillar system and the cross-linking molecules themselves. Such agents include glutaraldehyde, epoxy compounds, and isocyanates. The exogenous agents themselves are incorporated as part of the cross-link (Fig. 2). These treatments result in formation of chemical entities that are not normally present in biological tissues and, therefore, care must be taken to ensure that the cross-linked materials are not cytotoxic, are not inflammatory, do not induce a foreign body response, and, if biodegradable, the mobilized products do not cause immunogenic reactions and do not impact organs involved in their metabolism or excretion.

Another approach is based on catalyzing cross-link formation by physicochemical means that covalently join collagen fibrils via the naturally occurring reactive amino acid side chains. Included in this class of fixation methods are photo-oxidation, microwave irradiation, dehydration, and dehydrothermal treatment. Physicochemical methods eliminate concerns centering on nonbiological cross-linking agents, residual cross-linking precursors and intermediates, and the products derived from reactions of these agents

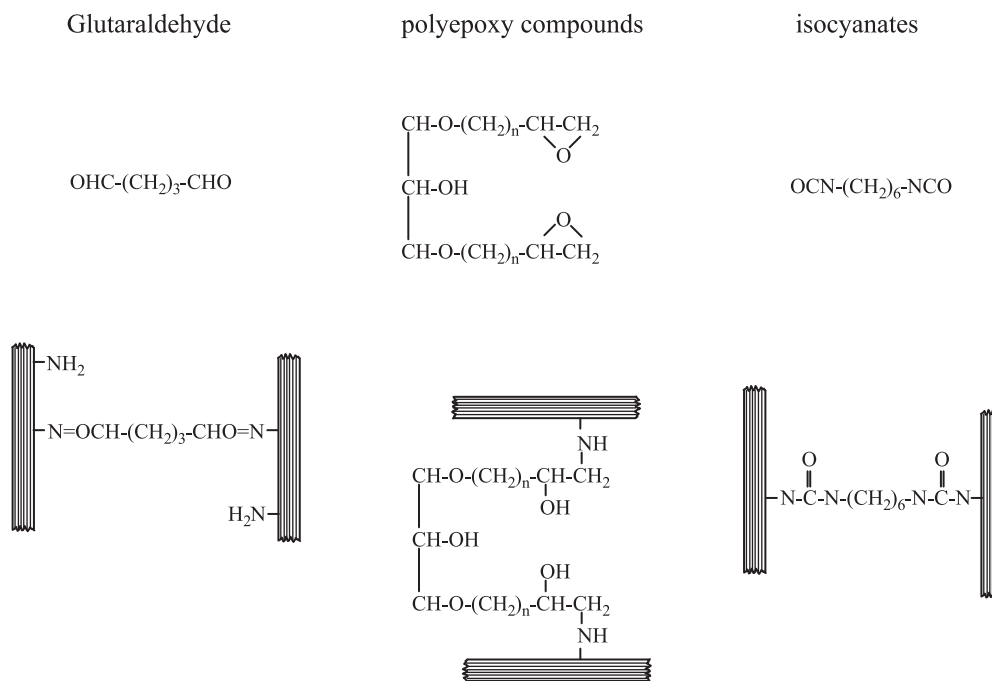


Fig. 2 Structure of chemical cross-linking agents that are themselves incorporated into the cross-link and the structure of the cross-links that form between collagen fibrils. It should be noted that the scheme for glutaraldehyde fixation is an extremely simplified representation of only the primary reaction product generated by glutaraldehyde. The extremely complex nature of glutaraldehyde cross-linking is described in Refs. [4] and [5] and these should be consulted for a full understanding of glutaraldehyde fixation and for further reading.

with collagen. It is assumed that the cross-links formed with these protocols are similar to naturally occurring biological adducts thereby reducing potential adverse effects *in vivo*. Physicochemical methods such as UV light irradiation may adversely affect the integrity of the collagen fibrillar system, so care must be taken to ensure that the design of the treatment protocol reduces detrimental reactions.

The third strategy is to chemically catalyze intermolecular cross-links between reactive amino acid side chains on the collagen fibrils. The advantage of this strategy is that the catalytic agent does not become incorporated into the cross-link and can be completely washed out of the biomaterial. Carbodiimide and the azyl azide method are two treatments that fix collagenous tissue grafts in this way.

The fourth strategy is based on polymerizing compounds that are mixed with collagenous assemblies and form polymeric interactions that do not chemically react with the collagen fibrils. These interactions can be either noncovalent or covalent. Materials produced with this approach can be considered polymer-reinforced composites. While the collagen is not chemically altered, the polymer formed may or may not be problematic *in vivo*.

Implementing one or another of the collagen fixation strategies described above involves distinct considerations based on the type of collagen scaffold to be

fixed, its inherent properties, and intended use *in vivo*. Allografts and xenografts require fixation to limit adverse reactions *in vivo*, including a foreign body response, an inflammatory reaction, and immunogenicity. A fixation strategy that retards biodegradation and absorption during the repair phase is also necessary in most cases. Premature debilitation of the biomaterial's functional properties must be avoided during the repair phase. The objective of most strategies is to design a fixation strategy that produces a biodegradable biomaterial, the rate of degradation of which is closely tuned to the repair process and the development of normal properties in the repair tissue. Since these grafts retain their natural biomechanical properties, less emphasis need be placed on fixation methods that are designed to improve physical properties. However, fixation will inevitably alter the graft's biomechanical properties, and this must be taken into account when designing graft fixation strategies.

Biomaterials produced from insoluble and reconstituted collagen also require fixation for biocompatibility and attenuation of biodegradation. However, of equal importance for these bioprotheses is fixation to achieve appropriate biomechanical performance. In practice, the principal goal in the design and development of collagen-based biomaterials is optimization of the fixation strategy for the desired functional properties. Nonetheless, only fixation methods that



result in biocompatible materials need be applied. In the best case scenario, the fixation method not only produces a mechanically competent biomaterial, but also renders the construct biocompatible. Some of the methods described below are capable of this dual effect.

Standard methods are used to measure the effectiveness of the fixation strategy, and these relate directly to the intended use of the biomaterial. The temperature at which collagenous structures shrink, the thermal shrinkage temperature, is generally taken as a measure of the density of intermolecular cross-links. The technique is simple to set up and reliably measures differences in effective cross-link density. If the fixation method creates new cross-links, the shrinkage temperature should increase over that of the untreated biomaterial. Differential scanning calorimetry is a more sensitive and informative tool to quantify thermal stability of the cross-linked collagen, but it requires costly equipment.

Biodegradability can be assessed initially in the laboratory by assessing the susceptibility of the cross-linked material to bacterial collagenase. The treated biomaterial is incubated with commercially available collagenase and the amount of degraded collagen released over time is compared to that of the untreated control. This method is employed primarily as an assessment of the effectiveness of the fixation method, although decreased susceptibility to collagenase *in vitro* indicates that the biomaterial will be less degradable *in vivo*. In the final analysis, the fixed biomaterial must be tested *in vivo* for biodegradability by implanting it, usually subcutaneously, and following its fate histologically.

Biocompatibility can be tested *in vitro* with isolated cells to determine whether the fixation agents or the cross-links they produce are cytotoxic. In addition, the response of specific cell types to the fixed biomaterial can be assessed. As with biodegradability, biocompatibility must be determined *in vivo*. Following implantation, the local tissue is analyzed histologically over time for inflammatory reactions, including the presence of T-cells, monocytes and macrophages, and multinucleated giant cells. Antigenicity can be evaluated by biochemical analysis for antibodies to the fixed biomaterial.

The functional properties of the fixed biomaterial must also be evaluated. The type of test will depend on the intended use of the fixed biomaterial. For example, cross-linked collagen fibers for tendon and ligament repair are tested with uniaxial tensile tests to failure. These tests yield tensile strength, stiffness, and strain at failure, which can be compared to the natural tensile properties of tendon and ligaments. The stiffness of fixed pig heart valves and pericardium is measured to determine whether they will perform *in vivo* similarly

to the natural valves. Beyond structural and mechanical properties, biochemical analyses are generally required to determine the nature of the reaction products formed by fixation protocols and to monitor the efficacy of the cross-linking conditions. Many of the fixation strategies described below involve free amines. Relatively straightforward colorimetric analyses of free amines, e.g. with picrylsulfonic acid (TNBS), can be employed to evaluate the cross-linked biomaterials.

Aldehyde-Based Fixation

Glutaraldehyde is an effective cross-linking agent that is used primarily to fix biological tissues for electron microscopy. Fixation is based on the introduction of covalent cross-links between neighboring proteins; however, the mechanism of bond formation and exact nature of the introduced bonds are not entirely clear. Glutaraldehyde contains two reactive aldehyde groups flanking a short 3-carbon aliphatic chain ($\text{OH}-\text{C}-(\text{CH}_2)_3-\text{C}-\text{HO}$). An aqueous solution of glutaraldehyde contains a mixture of the free aldehydes, monomeric and polymeric cyclic derivatives, and unsaturated polymers. At neutral pH, glutaraldehyde and its derivatives react with primary amines ($\text{R}-\text{NH}_2$) and, if the amines are constituent side chains of separate proteins, then glutaraldehyde cross-links these proteins. The number of cross-links formed depends on the number of available primary amines coupled with their intermolecular distance. It should also be noted that glutaraldehyde reacts with other functional groups in proteins, including carboxy and amido groups, though adducts formed from this reaction appear to play only a minor role in glutaraldehyde fixation.

Though simple in theory, glutaraldehyde fixation is encumbered with potential difficulties. First, glutaraldehyde in aqueous solutions polymerizes with time, forming polymers that contain fewer free aldehydes available for cross-linking. This problem can be minimized by using fresh solutions prepared from concentrated (>25%) electron microscopy-grade glutaraldehyde that has been stored frozen until dilution to the appropriate concentration. Second, the effectiveness of fixation varies with glutaraldehyde concentration and treatment duration. There is no straightforward general formula to apply to all tissue. The optimal concentration for any particular application must be empirically determined. There is good evidence that lower concentrations or short durations are more effective for some tissues, whereas high concentrations produce better outcomes for others. Third, coupled with concentration and duration is the size of the tissue or biomaterial to be fixed.

Complete and effective fixation relies on thorough penetration into the tissue, which can be problematic in large structures. Despite these potential problems, glutaraldehyde is one of the most investigated of the collagen fixation schemes and remains the standard to which new fixation schemes are systematically compared.^[4,5]

Glutaraldehyde reacts with ϵ -amino groups of lysine residues in collagen, thereby forming cross-links between collagen molecules, presumably at all hierarchical levels.^[5] Greatest biomechanical stability would result when cross-links occur between neighboring collagen fibrils. Interfibrillar cross-links can occur only where fibrils are within the cross-linking distance of glutaraldehyde and where lysine side chains are juxtaposed. Thus, glutaraldehyde is most effective in achieving the highest cross-link density in structures with high concentrations of collagen or aligned collagen fibrils.

Glutaraldehyde has been used extensively for treating allografts and xenografts, in which the primary objective is to make the graft more biocompatible and less biodegradable. Bioprosthetic bovine pericardium and porcine heart valves fixed with glutaraldehyde have been used in humans for over twenty-five years with success.^[4] The principal shortcoming of glutaraldehyde as applied to heart valves is that the cross-link products facilitate calcium deposition. Calcification is the major cause of failure as it results in stiffening the valves to an extent that they no longer function properly. Nevertheless, bioprosthetic heart valves fixed with glutaraldehyde remain the only such bioprosthesis in use today, largely because a suitable alternative, differently fixed heart valves, valve biomaterials, or synthetic substitutes, have yet to achieve adequate performance.

Glutaraldehyde fixation has been applied to a variety of other collagenous allografts and xenografts, including vascular grafts, pericardium, knee menisci, tendons, and ligaments.^[4] Research on other potential applications of glutaraldehyde fixation has been aimed at cross-linking reassembled collagenous constructs for fabricated bioprostheses, tissue repair scaffolds, and drug delivery vehicles.^[4-7] Despite the long history and large body of empirical work on glutaraldehyde-fixed collagenous biomaterials, problems associated with the cytotoxic, inflammatory, and antigenic effects resulting from this fixation strategy have yet to be resolved adequately to allow general use of these biomedical devices.

Polyepoxy Compounds

Polyepoxy compounds have been examined as potential collagen cross-linking agents for fixation of grafts

with the idea that they could serve as alternatives to glutaraldehyde and ameliorate the adverse reactions, e.g., cytotoxicity and calcification, associated with glutaraldehyde-fixed tissues. These compounds consist of short branched polymers that terminate in reactive epoxy functionalities. Both the number of epoxy groups, and therefore potential cross-linking sites, and the molecular distance separating them can vary significantly. The most commonly employed group of multifunctional epoxy reagents, the glycerol ethers, is based on glycerol, the hydroxyl side chains of which having been modified by the addition of alkanes of varying length. Other polyepoxy compounds include glycol and glycerol polyglycidyl ethers.^[5,8]

Cross-linking is based on the terminal epoxy functionality, a three-atom cyclic ether, the high strain energy of which favors ring opening, especially in the presence of nucleophiles. Polyepoxy compounds react primarily with the free amines of lysine side chains on neighboring proteins. Fixation is relatively straightforward in that the reaction takes place in aqueous solutions at neutral pH at room temperature. However, rates and degrees of cross-linking vary depending on the chemistry and size of the branched polymer. Experimental studies on specific tissues must be carried out in order to optimize the effectiveness of fixation. Epoxy compounds have been used to fix bovine pericardium, arteries, aortic valve cusps, and tendons.^[8] Cytotoxicity of the fixation strategy is equal to or less than that of glutaraldehyde.^[5]

Isocyanates

Isocyanates (R-NCO) react with primary amines to form the urea bond (R-NH-CO-NH-R). Difunctional isocyanates thus have the ability to cross-link collagen via its lysine side chains. One such compound, hexamethylene diisocyanate, in which the two reactive isocyanate groups are separated by a linear ten atom chain (OCN-(CH₂)₈-NCO), has been used to cross-link bovine pericardium,^[9] sheep dermal collagen,^[10] and processed collagen fibers.^[11]

In principle, isocyanates can be the terminal groups on a variety of polymers with distinct chemistries. For example, a polyurethane polymer with terminal isocyanate groups has been examined as a fixation strategy for porcine pericardium.^[12] Though effective at cross-linking the surface, it failed to penetrate the deep layers of the pericardium, likely due to the relatively large size of the polyurethane polymer, which prevented diffusion. An advantage of using polyurethane polymers with isocyanates is that the chain size separating the cross-linking moieties can be varied and thereby potentially result in differential



cross-linking effects. The potential use of difunctional isocyanates has not been systematically examined.

The major concern in designing fixation strategies using isocyanates is the solvent in which the reaction takes place. Aqueous solvents largely favor destruction of the isocyanate over reaction with collagen. This can be eliminated by using nonaqueous solvents. In the case of hexamethylene diisocyanate, iso-propanol was used effectively as the solvent. Dimethyl sulfoxide was used as the solvent for polyurethane diisocyanate treatment. Use of these non-aqueous reagents, however, requires additional considerations, as they will affect the collagen scaffold itself.

Studies to date, though few, have clearly established that diisocyanate treatment effectively cross-links collagenous materials. However, whether isocyanate and its derivatives are more or less biocompatible than other covalent cross-linking agents remains an open question. In-vitro experiments indicated that hexamethylene diisocyanate-treated sheep collagen was less cytotoxic than that treated with glutaraldehyde.^[13] In-vivo experiments with hexamethylene diisocyanate-fixed synthetic collagen fibers as anterior cruciate ligament replacements in sheep showed that they were less cytotoxic than glutaraldehyde-fixed fibers, and the degradation products were minimally cytotoxic.^[11] More work is necessary to determine the full range of biocompatibility with isocyanate-fixed materials.

Carbodiimide cross-linking

The recognition that exogenous cross-linking agents that become incorporated into the cross-link can cause serious complications in vivo has led to the development of strategies to catalyze cross-link formation with agents that act as catalysts only and are subsequently removed from the treated tissue following the reaction. The cross-links formed mimic chemical structures normally present in vivo and therefore could potentially eliminate concerns with biocompatibility. One such strategy that has recently received considerable attention is carbodiimide-generated iso-peptide bond formation.

Carbodiimide, 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) being the most widely used, reacts with the carboxyl groups of aspartic and glutamic acid side chains in collagen to form the *o*-isoacylurea derivative.^[5] This then reacts by nucleophilic attack with primary amines of lysine side chains on collagen, generating the amide cross-link and liberating the urea derivative of the carbodiimide, which can be readily washed out of the biomaterial. With the addition of N-hydroxysuccinimide (NHS), internal rearrangement of the activated isoacylurea

derivative is minimized resulting in little remaining reaction intermediates and more efficient cross-linking. No residual cross-linking agent remains in the biomaterial, and the resulting iso-peptide bond mimics the natural peptide bond in proteins, except that it bridges neighboring peptides. The reaction takes place in aqueous solution at pH 5. The extent of cross-linking can be controlled by adjusting the ratio of carbodiimide to collagen since the first step proceeds to completion.

Both tissues and collagenous biomaterials have been fixed with carbodiimide, although the predominant use of carbodiimide is to cross-link biomaterials created from insoluble and reconstituted collagen fibers. Its effect on tissues include stabilization of the collagen fibrillar network, enhancement of biomechanical properties, and a decrease in biodegradability.^[14] Carbodiimide cross-linking strengthens fibers produced from insoluble collagen. [15, 16] Recently, carbodiimide has been the cross-linking agent of choice for production of collagen scaffolds for tissue engineering, either alone or in combination with other macromolecules, primarily because properly washed carbodiimide-fixed scaffolds are not cytotoxic and they are entirely permissive to incorporation of cells (see, e.g., Refs. [17, 18] and [19]). Biomaterials cross-linked with carbodiimide are in general significantly more biocompatible than those cross-linked with glutaraldehyde; however, carbodiimide fixation would not necessarily reduce antigenicity.

Acyl azide-based cross-linking

As with carbodiimide treatment, the acyl azide method produces covalent bond formation between the carboxylic acid side chains of aspartic and glutamic acids and the ϵ -amino groups of the lysines of collagen (Fig. 3).^[20] Following esterification of the carboxyl groups in which a methyl group is added to the acid, the biomaterial is treated with hydrazine to form the corresponding hydrazide. Sodium nitride is added to react with the hydrazide forming the acyl azide. These reactive acyl azides then spontaneously couple with the primary amines of lysine to form the cross-link. The cross-linked biomaterial is washed thoroughly to remove all of the reactants, leaving the collagen with the newly formed iso-peptide bonds. A somewhat simpler one-step acyl azide method that uses diphenylphosphoryl azide has been reported, but this method requires a nonaqueous environment. While the acyl azide method has been shown to effectively stabilize collagenous biomaterials that are biocompatible,^[21] it is more difficult to carry out and has been superseded by the much simpler carbodiimide method.

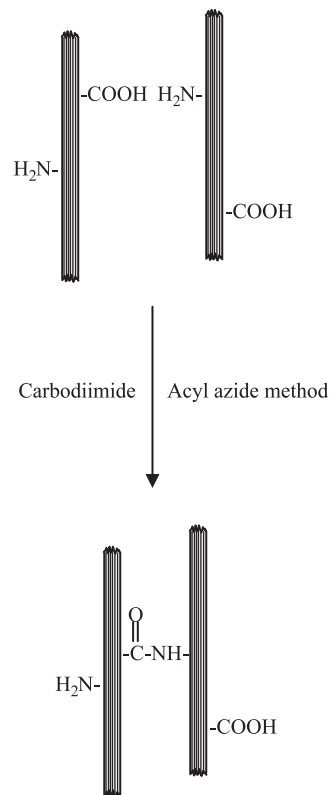


Fig. 3 Structure of the isopeptide cross-links formed between collagen fibrils by carbodiimide and acyl azide method.

Other Strategies

Dehydration

The aforementioned cross-linking strategies require that the molecular distance between the targeted functionalities in the collagen fiber system have the proper spatial dimension in order for cross-links to form. Both the side-to-side and lateral spacing must be in the proper register, and optimal spacing will depend on the fixation method employed. The three-dimensional configuration of the collagen fibers will thus dictate the cross-link distance. Fixation agents that become incorporated into the cross-link vary in length and therefore occupy a portion of the intermolecular distance. Glutaraldehyde is a straight chain molecule five atoms long; multifunctional epoxy compounds can vary between three and 46 carbons; hexamethylene diisocyanate spans twelve atoms and polyurethane diisocyanates can vary immensely in length. Isopeptide bond formation catalyzed by carbodiimide treatment or the acyl azide method requires that the carboxylic acid group be directly adjacent to the primary amine.

In order to facilitate optimal cross-link number and density, methods to concentrate and align collagen

fibers have been examined. These methods are generally applied to bioprostheses produced from purified insoluble collagen fibers and reconstituted collagen fibrils. The simplest method is to air dry the construct at room temperature. Not only do the collagen fibrils become closely apposed, it is likely that both covalent and noncovalent cross-links form among juxtaposed fibers during drying. This has proven most effective in the production of cross-linked fibers with advanced tensile properties for tendon and ligament repair (see, e.g., Refs. [15] and [16]). The addition of heat (110°C) and vacuum, dehydrothermal shrinkage, has been reported to produce concentrated cross-linked fibers with improved mechanical properties.^[22] However, dehydrothermal treatment must be used with caution because of progressive denaturation of the collagen triple helix and degradation.

UV irradiation

Another physicochemical cross-linking method that avoids introduction of exogenous agents into collagen scaffolds is ultraviolet light irradiation.^[22,23] Dry collagenous materials are exposed to UV light at 254 nm wavelength. Both the distance from the source and the duration of exposure can be adjusted to vary cross-link density and optimize the intended biomechanical properties. The extent of irradiation must be well controlled because UV light has two competing effects, cross-linking and fragmentation of the collagen triple helix. At the appropriate intensity, cross-linking proceeds rapidly and achieves its maximum effect before significant fragmentation occurs. Intense and prolonged UV treatment will not only break up the collagen but will also denature it, thereby eliminating the effect of cross-linking. UV treatment has been shown to increase the mechanical strength of collagenous materials, but it is not as effective as other cross-linking schemes at reducing degradation of materials *in vivo*.

Glycation and UV irradiation

Glucose-mediated cross-linking of collagen, termed glycation, is a well recognized process that occurs in connective tissues with aging and is accelerated in diabetic humans and laboratory animals. Experimental studies have established that glycation of collagen alters the mechanical properties of connective tissues such as tendon and skin and it reduces susceptibility of these tissues' collagen to enzymatic degradation. Both of these properties are beneficial for collagen-based bioprostheses. Initial attempts to take advantage of collagen glycation met with mixed success, primarily because of the relative inefficiency of cross-linking by



conventional in-vitro incubations with glucose alone.^[24] Of note, however, was that glycated collagen materials implanted subcutaneously did not cause an inflammatory response and allowed integration of host cells. Glycated collagen scaffolds have also been shown to provide a permissive environment in vitro for tissue engineering.^[25] Increased levels of glycation cross-links have been produced by UV treatment of glucose-saturated collagen films that resulted in improved biomechanical properties.^[26]

Biological cross-linking compounds

Genipin and its related iridoid compounds are natural products isolated from the fruits of the gardenia plant (*Gardenia jasminoides*). It is used in herbal medicines as an antiphlogistic and cholagogue, for dyeing food products, in the preparation of gelatin capsules, and immobilizing enzymes. Its capacity to cross-link proteins plus being a naturally occurring biological

compound make it potentially applicable and particularly interesting for collagen fixation. The protocol for cross-linking of collagen-based biomaterials with genipin is straightforward and has no adverse effects on the collagen matrix. Cross-linking is carried out in aqueous buffers such as phosphate at room temperature and neutral pH. The putative mechanism by which genipin cross-links collagen involves aldehydes and primary amines, similar to glutaraldehyde cross-linking, except that in aqueous solutions genipin is a dialdehyde.^[27] The aldehydes react with ϵ -amino groups on lysine side chains of neighboring collagen molecules. The genipin molecule becomes incorporated into the cross-link. Pericardium treated with genipin demonstrated increased stability to enzymatic degradation, improved biomechanical properties, and overall was comparable or superior to pericardium fixed with glutaraldehyde.^[27,28] In addition, genipin-fixed tissues are significantly more biocompatible and less cytotoxic than those fixed with glutaraldehyde and epoxy compounds.^[29]

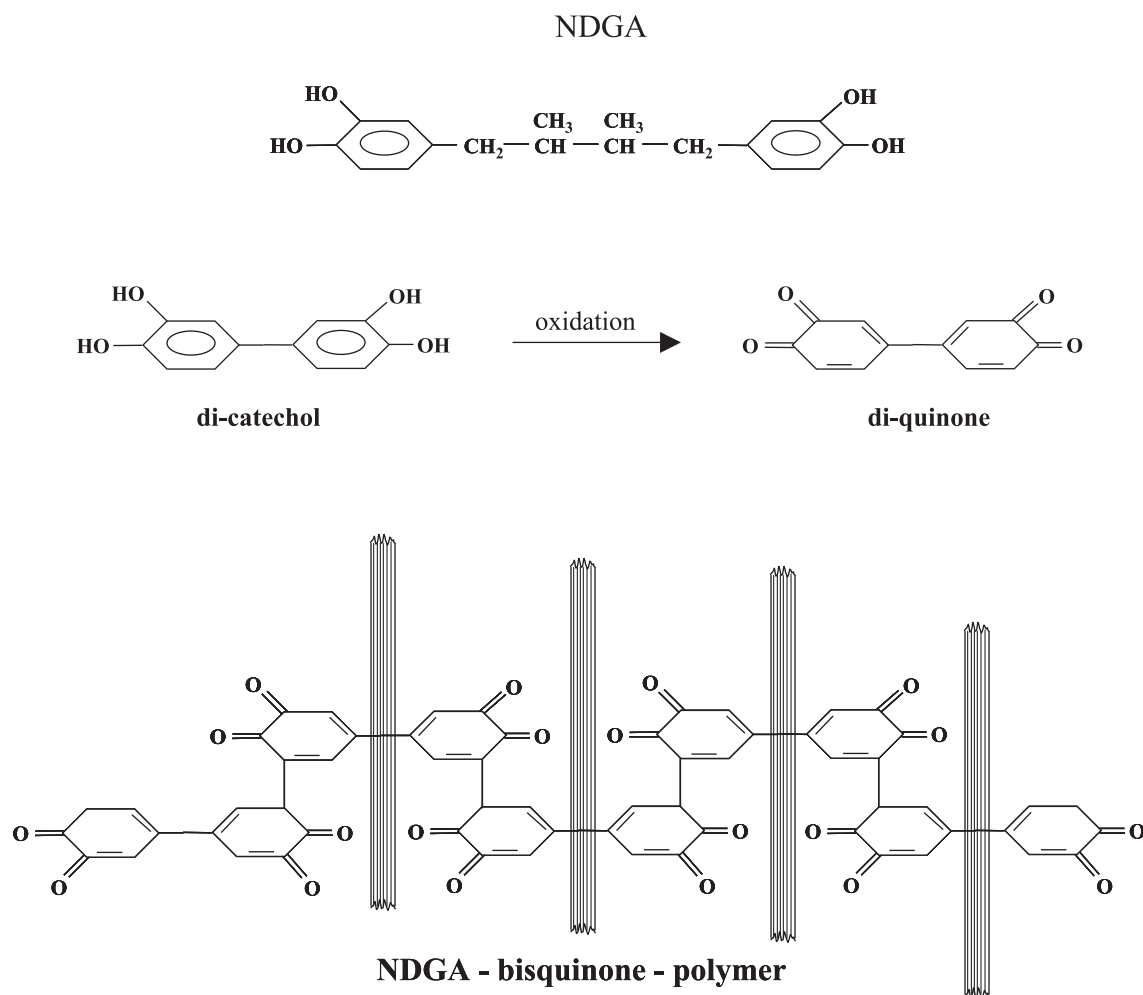


Fig. 4 Structure of NDGA, the mechanism of autoxidation and the bisquinone polymer that polymerizes with collagen fibrils.

Table 1 Annotated list of references for the principal collagen fixation strategies

Fixation strategy	Reference	Focus
Glutaraldehyde	[4]	Cross-link chemistry and applications to bioprostheses and drug delivery vehicles
	[5]	Cross-link chemistry and applications
	[7]	Applications for drug delivery vehicles
	[35]	Cross-linking methods for collagen scaffolds
Polyepoxy compounds	[8]	Chemistry of epoxy compounds, fixation methods, and properties of fixed tissues
	[5]	Chemistry of epoxy fixation and applications
Isocyanates	[5,7]	Chemistry of isocyanate fixation and its applications
Carbodiimide	[5,7]	Chemistry of carbodiimide-catalyzed isopeptide bond formation
Acyl azide	[5,7]	Chemistry of acyl azide cross-link formation and its applications
Glycation	[26]	UV-catalyzed collagen cross-linking and its effects on collagen scaffolds
Genepin	[27,28]	Chemistry of genepin and its effects on tissue properties
NDGA	[30]	Chemistry of NDGA fixation as applied to tendon bioprostheses

Catechol-quinone tanning systems are prevalent in a wide variety of animals wherein they strengthen otherwise vulnerable extracellular matrices. In all cases the catechol is derived from tyrosine. In insect cuticle, a chitin-based material, circulating tyramine is hydroxylated forming 3,4-dihydroxytyramine, which enters the newly secreted and assembled cuticle where it polymerizes with the chitin to form the hard exoskeleton. Mussel byssus threads are collagenous structures by which the animal secures a firm holdfast to the substrate. The foot gland produces an amalgam of collagenous material and catechol containing proteins. The catechol in this case is 3,4-dihydroxyphenylalanine produced intracellularly by hydroxylation of tyrosine in the polymerizing proteins. Skate egg capsules are polymerized by a quinone tanning mechanism that involves conversion of tyrosine in the structural proteins to 3,4-dihydroxyphenylalanine in the assembled biomaterial. In all these structures, polymerization is induced by enzymatic oxidation of the catechol moieties to the corresponding quinones, which subsequently form covalent adducts.

Nordihydroguaiaretic acid (NDGA), isolated from the creosote bush, is a small molecular weight dicatechol containing two *ortho*-catechols that are located

at either end of a short alkane chain, 3,4 dimethylbutane (2,3-bis(3,4-dihydroxyphenylmethyl)butane) (Fig. 4). The two catechols on NDGA will slowly autooxidize in aqueous solutions at neutral pH forming reactive quinones. Oxidized NDGA thus has the potential to fix collagenous biomaterials. However, the fixation mechanism differs from those described above. The NDGA-quinone does not form cross-links with side chains of collagen, but rather polymerizes to form an NDGA bisquinone polymer interpolated within the collagen network.^[30] The NDGA treatment protocol is suitable for collagenous biomaterials in that it occurs in aqueous buffers at neutral to slightly basic pH.^[31]

Fibers produced from reconstituted collagen treated with NDGA exhibit mechanical properties superior to collagen fibers fixed with glutaraldehyde, carbodiimide, epoxy compounds, and the acyl azide method.^[31] NDGA-collagen fibers were reported to be biocompatible *in vitro* and *in vivo*.^[30,32,33] They created little foreign body response or inflammatory reaction when placed subcutaneously in an animal model.^[30] NDGA fixation establishes that strategies other than those based on forming covalent cross-links between collagen molecules are potentially viable for producing collagen-based biomaterials.

Table 2 Annotated list of references for applications of collagen-based biomaterials

Biomaterial	Reference	Focus
Bioprostheses	[4]	Heart valves, vascular prostheses, wound dressings, hemostat
Collagen scaffolds	[4,7]	Drug delivery vehicles
	[36]	Soft tissue repair
	[37]	Shields, sponges, gels, skin, bone, and engineered tissues
Biocompatibility	[38]	Methods of assessment <i>in vivo</i> , immunogenicity
	[39]	Fate of collagen implants



CONCLUSION

The burgeoning research on collagen fixation strategies is both a testament to the need to devise safe and effective methods for the biomedical use of biologically based, collagenous materials, be they grafts or engineered constructs, and a clear indication that a wholly biocompatible and biomechanically competent fixation method has yet to be developed. Collagen fixation strategies continue to evolve. Traditional cross-linking methods such as glutaraldehyde could be improved by modifying the problematic properties that result from fixation. For example, much research has focused on efforts to mitigate calcification in glutaraldehyde-fixed pig heart valves.^[5] Synthetic mechanisms for producing novel cross-linkers may derive from developments in the industrial and chemical engineering sectors, similar to polyurethanes and isocyanates. New fixation chemistries will no doubt be coopted from other disciplines, in the same way carbodiimide cross-linking was transferred from analytical biochemistry to biomedical applications. As more is discovered about natural cross-linking mechanisms in collagen, such as glycation, these reactions might be mimicked to develop an entirely biomimetic collagen fiber. Though still far from realization, molecular biology could devise cellular techniques to produce natural collagens with cross-linking precursors built in. Perhaps the most enticing is the potential to capture and adapt stabilization chemistries from biological materials with unique and exploitable physical properties.

Finally, it should be pointed out that several excellent reviews on the use and treatment of collagen-based biomaterials have been published recently (Tables 1 and 2). These reviews provide more in-depth treatment of specific aspects covered in this chapter and should be consulted by those who wish to learn more about particular cross-linking strategies and their applications. Moreover, an extensive introduction to the literature on collagen fixation strategies as well as biomaterials development can be found in these papers.

ACKNOWLEDGMENTS

This entry was prepared under the auspices of the Shriners Hospitals for Children and Grant #8610 awarded to the author.

ARTICLES OF FURTHER INTEREST

Allografts; Collagen; Collagen Processing; Polyepoxy Compound Fixation; Tissue Engineering Scaffolds; Xenografts

REFERENCES

1. Kadler, K.E. Extracellular matrix 1: Fibril-forming collagens. *Protein Profile* **1995**, *2*, 491–619.
2. Myllyharju, J.; Kivirikko, K.I. Collagens and collagen-related diseases. *Ann. Med.* **2001**, *33*, 7–21.
3. Hulmes, D.J.S. Building collagen molecules, fibrils, and suprafibrillar structures. *J. Struct. Biol.* **2002**, *137*, 2–10.
4. Jayakrishnan, A.; Jameela, S.R. Glutaraldehyde as a fixative in bioprotheses and drug delivery matrices. *Biomaterials* **1996**, *17*, 471–484.
5. Khor, E. Methods for the treatment of collagenous tissues for bioprotheses. *Biomaterials* **1997**, *18*, 95–105.
6. Milthorpe, B.K. Xenografts for tendon and ligament repair. *Biomaterials* **1995**, *15*, 745–752.
7. Friess, W. Collagen—Biomaterial for drug delivery. *Eur. J. Pharm. Biopharm.* **1998**, *45*, 113–136.
8. Sung, H.-W.; Hsu, H.-L.; Shih, C.-C.; Lin, D.-S. Cross-linking characteristics of biological tissues fixed with monofunctional or multifunctional epoxy compounds. *Biomaterials* **1996**, *17*, 1405–1410.
9. Naimark, W.A.; Pereira, C.A.; Tsang, K.; Lee, J.M. HMDC crosslinking of bovine pericardial tissue: A potential role of the solvent environment in the design of bioprosthetic materials. *J. Mater. Sci., Mater. Med.* **1995**, *6*, 235–241.
10. Van Wachem, P.B.; Van Luyn, M.J.A.; Olde Damink, L.H.H.; Dijkstra, P.J.; Feijen, J.; Niewenhuis, P. Biocompatibility and tissue regenerating capacity of cross-linked dermal sheep collagen. *J. Biomed. Mater. Res.* **1994**, *28*, 353–363.
11. Chvapil, M.; Speer, D.P.; Holubec, H.; Chvapil, T.A.; King, D.H. Collagen fibers as a temporary scaffold for replacement of ACL in goats. *J. Biomed. Mater. Res.* **1993**, *27*, 313–325.
12. Loke, W.K.; Khor, E.; Wee, A.; Teoh, S.H.; Chain, K.S. Hybrid biomaterials based on the interaction of polyurethane with porcine pericardium. *Biomaterials* **1996**, *17*, 2162–2172.
13. Van Luyn, M.J.A.; Van Wachem, P.B.; Olde Damink, L.; Dijkstra, P.J.; Feijen, J.; Niewenhuis, P. Relations between in vitro cytotoxicity and cross-linked dermal sheep collagens. *J. Biomed. Mater. Res.* **1992**, *26*, 1091–1110.
14. Sung, H.W.; Chang, W.H.; Ma, C.Y.; Lee, M.H. Crosslinking of biological tissues using genipin and/or carbodiimide. *J. Biomed. Mater. Res.* **2003**, *64*, 427–438.
15. Kato, Y.P.; Christiansen, D.L.; Hahn, R.A.; Shieh, S.-J.; Goldstein, J.D.; Silver, F.H. Mechanical properties of collagen fibers: A comparison of reconstituted and rat tail tendon fibers. *Biomaterials* **1989**, *10*, 38–42.
16. Dunn, M.G.; Prashanti, N.A.; Zawadsky, J.P. Optimization of extruded collagen fibers for ACL reconstruction. *J. Biomed. Mater. Res.* **1993**, *27*, 1545–1552.
17. Pieper, J.S.; van der Kraan, P.M.; Hafmans, T.; Kamp, J.; Buma, P.; van Susante, J.L.; van den Berg, W.B.;

- Veerkamp, J.H.; van Kuppevelt, T.H. Crosslinked type II collagen matrices: Preparation, characterization and potential for cartilage engineering. *Biomaterials* **2002**, *23*, 3183–3192.
18. Park, S.N.; Lee, H.J.; Lee, K.H.; Suh, H. Biological characterization of EDC-cross-linked collagen hyaluronic acid matrix in dermal tissue restoration. *Biomaterials* **2003**, *24*, 1631–1641.
 19. Daamen, W.F.; van Moerkerk, H.T.; Hafmans, T.; Buttafoco, L.; Poot, A.A.; Veerkamp, J.H.; van Kuppevelt, T.H. Preparation and evaluation of molecularly-defined collagen-elastin-glycosaminoglycan scaffolds for tissue engineering. *Biomaterials* **2003**, *24*, 4001–4009.
 20. Petit, H.; Rault, I.; Huc, A.; Menasche, P.; Herbage, D. Use of acyl azide method for cross-linking collagen-rich tissues such as pericardium. *J. Biomed. Mater. Res.* **1990**, *24*, 179–187.
 21. Petit, D.; Duval, J.-L.; Frei, V.; Abdul-Malik, N.; Sigot-Luizard, M.-F.; Herbage, D. Cytocompatibility of calf pericardium treated by glutaraldehyde and by the acyl azide methods in an organotypic culture model. *Biomaterials* **1995**, *16*, 1003–1008.
 22. Weadcock, K.S.; Miller, E.J.; Bellincampi, L.D.; Zawadsky, J.P.; Dunn, M.G. Physical crosslinking of collagen fibers: Comparison of ultraviolet irradiation and dehydrothermal treatment. *J. Biomed. Mater. Res.* **1995**, *29*, 1373–1379.
 23. Wang, M.-C.; Pins, G.D.; Silver, F.H. Collagen fibers with improved strength for the repair of soft tissue injuries. *Biomaterials* **1994**, *15*, 507–512.
 24. Girton, T.S.; Oegema, T.R.; Tranquillo, R.T. Exploiting glycation to stiffen and strengthen tissue equivalents for tissue engineering. *J. Biomed. Mater. Res.* **1999**, *46*, 87–92.
 25. Girton, T.S.; Oegema, T.R.; Grassl, E.D.; Isenberg, B.C.; Tranquillo, R.T. Mechanisms of stiffening and strengthening in media-equivalents fabricated using glycation. *J. Biomech. Eng.* **2000**, *122*, 216–223.
 26. Ohan, M.P.; Weadcock, K.S.; Dunn, M.G. Synergistic effects of glucose and ultraviolet irradiation on the physical properties of collagen. *J. Biomed. Mater. Res.* **2002**, *60*, 384–391.
 27. Sung, H.-W.; Huang, R.-N.; Huang, L.L.H.; Tsai, C.-C.; Chiu, C.-T. Feasibility study of a natural cross-linking reagent for biological tissue fixation. *J. Biomed. Mater. Res.* **1998**, *42*, 560–567.
 28. Sung, H.-W.; Chang, Y.; Chiu, C.-T.; Chen, C.-N.; Liang, H.-C. Crosslinking characteristics and mechanical properties of a bovine pericardium fixed with a naturally occurring crosslinking agent. *J. Biomed. Mater. Res.* **1999**, *47*, 116–126.
 29. Huang, L.L.H.; Sung, H.-W.; Tsai, C.-C.; Huang, D.-M. Biocompatibility study of a biological tissue fixed with a naturally occurring crosslinking reagent. *J. Biomed. Mater. Res.* **1998**, *42*, 568–576.
 30. Koob, T.J. Biomimetic approaches to tendon repair. *Comp. Biochem. Physiol., Part A* **2002**, *133*, 1171–1192.
 31. Koob, T.J.; Hernandez, D.J. Material properties of polymerized NDGA-collagen composite fibers: Development of biologically based tendon constructs. *Biomaterials* **2002**, *23*, 203–212.
 32. Koob, T.J.; Willis, T.A.; Hernandez, D.H. Biocompatibility of NDGA polymerized fibers. I. In vitro evaluation of cytotoxicity with tendon fibroblasts. *J. Biomed. Mater. Res.* **2001**, *56*, 31–39.
 33. Koob, T.J.; Willis, T.A.; Qiu, Y.S.; Hernandez, D.H. Biocompatibility of NDGA polymerized fibers. II. Attachment, replication and migration of tendon fibroblasts. *J. Biomed. Mater. Res.* **2001**, *56*, 40–48.
 34. Weadcock, K.; Olson, R.M.; Silver, F.H. Evaluation of collagen crosslinking techniques. *Biomater. Med. Dev. Artif. Organs* **1983–1984**, *11*, 293–318.
 35. Rault, I.; Herbage, D.; Abdul-Malak, N.; Huc, A. Evaluation of different chemical methods for cross-linking collagen gel, films and sponges. *J. Mater. Sci., Mater. Med.* **1996**, *7*, 215–221.
 36. Pachence, J.M. Collagen-based devices for soft tissue repair. *J. Biomed. Mater. Res.* **1996**, *33*, 35–40.
 37. Lee, C.H.; Singla, A.; Lee, Y. Biomedical applications of collagen. *Int. J. Pharm.* **2001**, *221*, 1–22.
 38. Ramshaw; Werkmeister, J.A.; Glattauer, V. Collagen-based biomaterials. *Biotechnol. Genet. Eng. Rev.* **1995**, *13*, 335–382.
 39. Bailey, A.J. The fate of collagen implants in tissue defects. *Wound Repair Regen.* **2000**, *8*, 5–12.



Collagen Processing

Michael J. Yost

University of South Carolina, Columbia, South Carolina, U.S.A.

Louis Terracio

New York University, New York, New York, U.S.A.

Robert L. Price

University of South Carolina, Columbia, South Carolina, U.S.A.

INTRODUCTION

The transformation of animal dermis into useful products predates recorded history. Primitive people who lived during the Ice Age approximately 500,000 years ago were likely the first to use animal skins to protect their bodies from the elements. Just as leather today is a by-product, our ancient ancestors hunted animals primarily for food. But, once they had eaten the meat, they would clean the skin by scraping off the flesh and then slinging it over their shoulders as a crude coat. They also made footwear by using smaller pieces of animal skin that were made to fit loosely over the foot and tied it at the ankle with thin strips of skin or vines to protect their bare feet from rocks and thorns. The main problem that primitive man encountered was that, after a relatively short time, the skins decayed and rotted away. With his limited knowledge and experience, primitive man had no idea how to preserve these hides. As centuries passed, it was noticed that several processes could slow down the decay of the skins. If the skins were stretched out and allowed to dry in the sun, even though they became stiff and hard, they lasted much longer. Various oily substances were then rubbed into the skins to soften them. As time passed, it was eventually discovered that the bark of certain trees contained tannin or tannic acid, which could be used to convert raw skins into what we recognize today as leather. It is quite hard to substantiate chronologically at exactly what time this tanning method materialized, but the famous Iceman, which dates from at least 5,000 BC that was discovered in the Italian Alps, was clothed in very durable leather.

Techniques used by the American Indians were very similar to those used in this early period. Indians soaked ashes from their campfires in water and then soaked the skins in this solution. In a few weeks, the hair and bits of flesh came off and only the raw hide was left. The hide was then tanned in a solution made from a mixture of hemlock and oak bark. This took about three months to complete. After which, the leather was worked by hand to make the hide soft and pliable.

THE MAKING OF LEATHER GOODS

The tanning of leather was used by mankind in numerous geographical areas during early periods of human civilization. As certain leather characteristics began to emerge, man realized that leather could be used for many purposes besides footwear and clothing. The uses and importance of leather increased greatly. For example, it was discovered that a leather bag acts as an insulator that keeps water fresh and cool. It was also found suitable for items such as tents, beds, rugs, carpets, armor, and harnesses. Ancient Egypt, one of the most developed civilizations in this early period, valued leather as an important item of trade. The Egyptians made leather sandals, belts, bags, shields, harnesses, cushions, and chair seats from tanned skins. Many of these items are still made from leather today.

The Greeks and Romans also used leather to make many different styles of sandals, boots, and shoes. When the Roman legions marched in conquest across Europe, they were well-attired in leather armor and capes. In fact, right up until the early 18th century, the shield carried by the ordinary soldier was more likely to be made of leather than metal.^[1]

Today, many items of clothing are still made from leather. However, other uses have been developed for collagen, the primary component of leather. These include food substances in addition to medical devices and therapies. In this article, we will cover various aspects of collagen processing, recovery, and fibrillogenesis that allows collagen to be used as an industrial raw material and in biomedical applications.

COLLAGEN RECOVERY

The goal of collagen processing is to isolate collagen from cellular and extracellular material for use as a unique polymeric material. All procedures used to isolate collagen rely heavily on several key processing variables, including temperature–time, pH, and ionic strength. The ionic strength on a concentration basis is defined as $I_c = 1/2 \sum c_B z_B^2$. The sum goes over all the

ions B . z_B is the charge number of ion B . Time itself is not a process variable, but it becomes important in diffusion of ions through the collagen articles and the temperature–time history of the collagen.

COLLAGEN PRODUCTION FOR INDUSTRIAL APPLICATIONS

Initial Processing

Freshly killed bovine steers which are up to 18 months old, are hooked onto a chain conveyor and transported into the processing area (5°C winter and 35°C summer, ambient conditions). The hide is removed in the skinning area and allowed to fall into a water bath (10°C winter and 30°C summer, ambient conditions). It is then cooled from a normal body temperature of 39°C. The hides are removed from the water bath, trimmed to remove extraneous flesh, and boxed (5°C winter and 35°C summer, ambient conditions) for shipment to the tannery by dedicated refrigerated truck (4°C).

Beamhouse

Once received at the tannery, the hides are washed for 30–120 minutes in potable water (30°C) that contains a small amount (0.2 wt%) of NaHCO_3 ($\text{pH}=9.0$, $\text{IS}=0.01$). The water is drained and refilled for a second wash or soak. Soaking is accomplished by adding $\text{Na}(\text{CO}_3)_2$ at 0.2 wt% which is based on wet hide weight, ($T=27^\circ\text{C}$, $\text{pH}=9.0$, $\text{IS}=0.01$) for 3.5 hours. Surfactant may be added at this time to help remove the fat from the hides. The soak water is drained, and the hides are washed for a third time in 27°C water for 30–60 minutes to rinse off the $\text{Na}(\text{CO}_3)_2$.

Hair Removal

Hair is removed from the hides in the hair burn step. Water is added to equal 50 wt% of the wet hide weight. Then, 0.6 wt% $\text{Na}(\text{CO}_3)_2$, 2.0 wt% $\text{Ca}(\text{OH})_2$, and 4.3 wt% NaHS is added and mixed for two hours to remove all of the hair from the exterior of the hide ($T=30^\circ\text{C}$, $\text{pH}=13.0$, $\text{IS}=0.09$). After 30 minutes in the hair burn, more water can be added if the hides swell and absorb all of the available water. The hair burn treatment is drained, and the hides are limed. Water is added at 1.5 to 2.0 times the wet hide weight and then 2 wt% $\text{Ca}(\text{OH})_2$ is added and mixed for seven hours. ($T=27^\circ\text{C}$, $\text{pH}=12.5$, $\text{IS}=0.4$). The liming solution is drained, and the hides are washed in water for 90 minutes. The final treatment in the beamhouse is with 0.1 wt% sodium metabisulfite solution for

20 minutes ($T=27^\circ\text{C}$, $\text{pH}=12.5$, $\text{IS}=0.1$). From the beamhouse, the hides are transported to the tannery.

Tannery

In the tannery, the hides are stacked on pallets and then split. Splitting is accomplished by feeding the hide pieces through a continuous knife that separates the top grain from the corium and flesh layers of the hide. The top grain is then further processed into leather goods for products such as shoes, belts, briefcases, car upholstery, and other industrial uses. Leather making is beyond the scope of this text. The top grain layer of the hide is unsuitable for use in food and medical collagen applications because of highly developed nonacid labile crosslinking. After splitting, the corium layer is stripped of any remaining flesh or subdermal tissue. The corium layer is then transferred to the white room for further processing.

White Room

In this step, the hide amide–N concentration is established. This is a critical determinant in the production of collagen. Amide–N is an indication of the extent of treatment with base. The hides are again treated with 3 wt% $\text{Ca}(\text{OH})_2$ for 90 minutes ($T=15^\circ\text{C}$, $\text{pH}=12.5$, $\text{IS}=0.6$). The goal is for the hide to absorb and retain 1.0–2.0 wt% lime based on the total amount of solids. Typical solids content at this point is 24.0 wt%. The hides are returned to a 3.0 wt% $\text{Ca}(\text{OH})_2$ solution in the temper step, removed from the lime, stacked in shipping containers, and sent to cold storage for use in industrial applications. The hides are good for up to 12 days in storage before the collagen begins to significantly degrade into gelatin. Again, amide–N content will be a good indicator of the state of the material (Fig. 1).

Treatment with base, usually lime, drastically changes the characteristics of the collagen. The treated collagen will associate a great deal more water than untreated collagen (Fig. 2). This process is commonly referred to as opening up of the hide, which means that treatment with lime allows much faster and more consistent absorption of tanning chemicals than before treatment.

There are a number of changes that take place during treatment with base or liming of the collagen tissues. The addition of salts, sulfur compounds, and amines controls the degree of swelling of the collagen. The swelling state influences the mass transport of noncollagenous cellular and extracellular materials from the hide. Any remaining cells are destroyed. The content of proteoglycans and carbohydrates is greatly reduced. Ammonia is liberated from glutamine



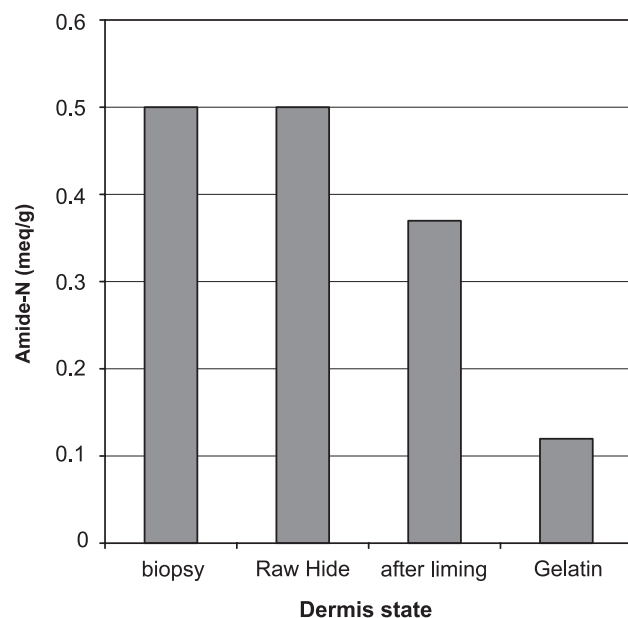


Fig. 1 Bovine hide amide concentration at various stages of processing the hide for leather or other applications. As the hide becomes more chemically treated, the lower the hide amide content. (From Ref. [2].)

and asparagine—forming additional carboxyl groups. This is indicated by a drop in the amide nitrogen content.^[2]

Deliming

$\text{Ca}(\text{OH})_2$ is effectively removed from the hides by acid treatment. Deliming begins with a water wash at 28°C

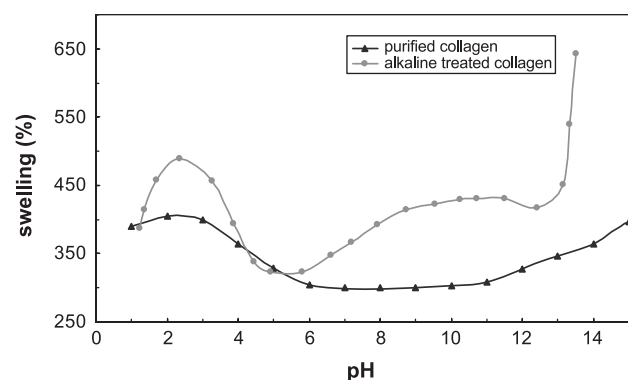


Fig. 2 Collagen swelling characteristics before and after treatment with lime. After treatment with lime, the collagen swells much more strongly with changes in pH. Swelling is maximized around pH 2.5 and above pH 13. This behavior is useful for further treatment by tanning chemicals or prior to dissociation of the collagen fibrils. Adapted from de la Burde (1961).

and a pH of 12.5, which results from residual lime, for 95 minutes. The wash water is drained, and the vessel is refilled with water at 50.0 wt% of the total hide weight. Salt can be added at this point—up to 1 M—to facilitate the removal of $\text{Ca}(\text{OH})_2$ from the hides. Salt increases the ionic strength of the solution and causes a partial shrinking of the hides and opening of the pore structure. Different salts have been used. Two of the most effective are NaCl and NH_4Cl . During deliming, acid, and which is usually concentrated HCL, is added to slowly and consistently drop the pH from 12.5 to roughly 5.5. It is critical that the acid be added stepwise and the pH of the solution monitored. The hides should be allowed to equilibrate with the aqueous solution. If the acid is added too quickly, as indicated by the solution's pH being below 5.5, the hide pieces will swell on the outside, which essentially traps calcium on the interior and results in a high ash content and poor subsequent processing performance. Deliming can be accomplished in roughly 600 minutes ($T=28\text{ C}$, $\text{pH}=12.5\text{--}5.5$, $\text{IS}=0.36$). The deliming solution is drained, and the hides are washed in ice water to reduce their temperature. At this point, hide pieces may be split, and the internal pH checked is by using a surface pH probe. Targets for sufficient lime removal are an internal pH between 6.0 and 6.8. The delimed hides may now be stored at 4°C for up to 72 hours or frozen for up to one year.

Size Reduction

Bovine hides have dimensions of about 1.5 meters \times 1.2 meters and must be reduced in size for further processing to extract the collagen. The delimed hide is a tough, fibrous material that can be a very challenging task for the size-reduction equipment. Several types of mills have been tried, including rotary knife, hammer, rotating discs, and ball. Rotary knife mills work best because they can shred the collagen hide without overheating it. Size-reduction processing can cause intense, localized heating of the material being reduced and, thus, can be detrimental to the collagen. This can be minimized by adding ice during the size-reduction process.

Mixing

Ground hide pieces have a solids content of 20.0 to 25.0 wt%. For most applications, the collagen solids content is set between 3.0 and 9.0 wt%. At these concentrations, the collagen's consistency is between that of thick batter (3.0 wt%) and stiff dough (9.0 wt%). For industrial purposes, the collagen hide particles can now be acidified to dissociate the collagen molecules. It is unlikely that the collagen will dissociate

completely into individual triple helix molecules. It is much more likely that the collagen will dissociate into sub- and microfibrils rather than completely to individual collagen molecules. Mixing, in this case, is usually a two-step process. First, there is usually a bulk mixing of acidified water into the comminuted hide particles. This can be accomplished in planetary or sigma blade type mixers that are designed for high-bulk flow of viscous materials. Upon acidification, the collagen swells associates water with the protein molecules, and the collagen fibrils will then dissociate from each other. The viscosity will greatly increase at this stage. The second mixing step is usually a high-shear local mixing. This step is necessary to help break apart clumps of collagen that are stuck together and improve the overall distribution of the acid within the mixture. Localized heating can be a significant problem for this step; therefore, cooling is required. ($T=11^{\circ}\text{C}$, $\text{pH}=2.5\text{--}3.0$, $\text{IS}=\text{low}$). The ionic strength should be kept to a minimum during the swelling and dissociation processes because excess ionic strength will shrink the collagen and make it difficult to create a smooth aqueous dispersion.

Homogenization

The collagen dispersion requires 24–48 hours for the acid to diffuse into the hide particles and complete the swelling and dissociation. Further high-shear intensity mixing is required to complete the processing of the dispersion. A two-stage homogenizer operating at pressures of 4000–6000 psi is used for this mixing. After homogenization, the dispersion is filtered to remove any collagen aggregates that did not disintegrate during the processing (40–60 micron slotted stainless steel type filter). Homogenization is a two-step storage, homogenization, and filter process. The dispersion is now ready for shaping.

Shaping

Shaping the collagen dispersion is the same as for any polymer system. It can be molded, extruded in tubes and sheets, spun into fibers, etc. Treatment after shaping can be a dry or wet process, depending on the type and properties of the desired final product. Post shaping and drying in the acidic state followed by neutralization and cross-linking is considered the dry process and neutralization followed by cross-linking and drying is considered the wet process. The dry process produces a tough film that has a very small pore structure and low water permeability. The wet process produces a collagen film that has a much more open pore structure that behaves more like a fabric than a polymer film.

Dry Process

The collagen dispersion is shaped and then dried at a maximum temperature of 50°C . Higher temperatures will cause irreversible denaturation of the collagen molecules and turn it into gelatin. Drying is completed when the moisture content is very low at approximately 7.0 wt% water. This extreme drying develops some dehydrothermally induced cross-links. After drying, the collagen is neutralized in a water bath containing 0.3 wt% NaHCO_3 . After neutralization, the collagen is washed in water to remove the bicarbonate salt. It then can be treated with cross-linking agents, such as glutaraldehyde 200–1000 ppm, pH approximately 7.2–7.5, water-washed, and dried again. For thin films under 3 millimeters thick, 10-second exposure time is sufficient to produce measurable changes in the modulus and ultimate tensile properties of collagen films. Longer times may be required for thicker films or higher cross-link densities.

Wet Process

The collagen dispersion is shaped at temperatures between 4°C and 10°C and then quickly heated to 37°C while simultaneously neutralizing the pH in either an anhydrous ammonia chamber or ammonium sulfate bath that is saturated with ammonia to a pH between 8.0 and 12.0. Further processing includes water washes that lower the dispersion pH to between 5.5 and 7.0 to remove residual salt. The collagen can then be air-dried or treated with cross-linking agents such as aldehydes. Air-drying can be accomplished at temperatures up to 135°C . These are much higher temperatures than can be used in the dry process because the collagen is at a much more neutral pH. The collagen is dried to moisture contents as low as 7.0 wt% to induce chemical and dehydrothermal cross-links. Collagen can then be rehydrated to the desired final moisture content of 20.0–25.0 wt% water.

Further postprocessing cross-linking can be accomplished by treatment with UV radiation or drying to a very low moisture, such as 7.0 wt%, and then holding at an elevated temperature of 80°C for up to 48 hours.

Collagen Extraction for Laboratory Uses

Research in the medical field and other industries frequently requires that collagen be extracted from noncollagenous tissues, cells, and extracellular matrix components.^[3] During this procedure, intact bovine dermis is obtained from slaughterhouses, the hair removed, and all subdermal tissue is scraped off. The top grain is then split from the hide with a sharp knife



to reveal the corium layer. At this point, the hide can be cut into smaller pieces and stored at -20°C for up to one year.

Unless otherwise noted, all subsequent processing is performed at 4°C . Once ready for processing, the hides are partially thawed and ground into small particles of approximately 5 millimeters diameter to generate sufficient surface area to maximize subsequent chemical processing while minimizing mechanical degradation of the collagen fibrils. The most frequently used processing method involves a 24-hour treatment of the particles with 0.5 N acetic acid (10:1 based on solids) and pepsin (1:100 based on wet hide weight) while mixing with a high-shear mixer. Acetic acid (0.5 M to glacial) is used in this step because it is a very effective solvent for collagen. Other solvents that may be used include mineral acids such as HCl, which is less expensive than acetic acid but is a significantly poorer solvent. The addition of pepsin to the reaction mixture is for two important reasons. First, pepsin is active against noncollagenous peptides, but it does not attack collagen. Second, pepsin is active at a low pH (2.5–3.5) in which the collagen molecule has unfolded and associated significant amounts of water. While the collagen is unfolded and at a low pH, pepsin will cleave any noncollagenous telopeptides from the collagen triple helix and leave the two $\alpha 1$ and the $\alpha 2$ chains intact.^[4] This significantly removes noncollagenous components from the mix. It is important to prevent localized heating of the collagen dispersion during this step, and ice may be incorporated into the mixture as necessary to maintain a temperature of 4°C .

Following the 24 hours in the acetic acid: pepsin mixture, the collagen dispersion should be centrifuged at 20,000 g for 50 minutes, which will result in the separation of the solubilized collagen and the nonacid-labile cross-linked collagen in the pellet that can be discarded. The solubilized collagen should be mixed in a high-flow, low-shear mixer while solid NaCl is slowly added to a final concentration of 2 M. Once all of the salt is added, subsequent mixing is not required. If the NaCl is added too rapidly, the collagen fibrils will clump around the salt and significantly slow dissolution. If done properly, the collagen will form a white precipitate in the mixture that should be complete in 24 hours. The collagen can then be collected by either centrifuging (20,000 g for 50 minutes) or filtration through a 50–100 μm filter.

Residual salts can be removed from the collagen by dissolving the pellet or filtrate in 0.5 N acetic acid (10:1 based on wet solids weight) and dialyzing in 0.5 N acetic acid (30:1 based on wet solids weight). The dialysis solution should be changed a minimum of three times over a 24-hour period. Once dialysis is complete, the pH of the solution should be adjusted to

6.0 with 1 M NaOH while maintaining the temperature of the solution between 0 and 4°C . The collagen can then be recovered by centrifugation or filtration as described above.

At this point, it is necessary to separate the neutral from the nonneutral salt soluble collagens. Neutral salt soluble collagen can be collected by dialyzing the neutralized collagen dispersion versus 0.05 M tris buffer containing 1 M NaCl at a pH of 7.5 (30:1 buffer based on sample weight). As above, the buffer should be changed a minimum of three times in a 24-hour period. The solution should then be centrifuged, which will result in the precipitation of the nonneutral salt soluble collagen that should be discarded.

To further separate types I and III collagen, the solution is dialyzed versus 0.05 M tris buffer containing 1.8 M NaCl at pH 7.5 for 48 hours at 4°C with a minimum of at least three changes of buffer. Type III collagen can be removed by centrifugation at 35,000 g for 50 minutes. To precipitate type I collagen, the collagen solution is dialyzed vs 0.05 M tris containing 3.5 M NaCl at pH 7.5 for 48 hours at 4°C while changing the buffer at least three times. Type I collagen can then be collected by centrifugation at 35,000 g for 50 minutes.

Degradation

There are three main methods that cause degradation of the intact collagen triple helix: proteolysis, thermal, and chemical. Proteolysis occurs when bacteria, mold, or yeast contaminates the collagen material during processing and secrete collagenase enzymes that degrade the collagen. Most bacteria are killed during the pH shifts from very basic to very acidic levels. Molds and especially yeasts can be very tenacious by surviving and adapting to pH extremes. Preservatives can be added to inhibit microbial growth. Sodium benzoate (500–1000 ppm) can prevent growth of almost all microorganisms, including yeast, bacteria, and fungi. It operates best in an acidic environment at a pH below 3.5 and works synergistically with other preservatives especially potassium sorbate (500–1000 ppm).

The second common cause of degradation of collagen is when it is exposed to high temperatures. In the extreme case of thermal degradation, the collagen triple helix will lose its conformation and become a random coil. The collagen will become gelatin with significantly different physical properties than intact collagen. The most common thermal degradation case causes hydrolysis of the peptide bonds along the backbone of the collagen triple helix. The triple helix is damaged and, thus, causes changes in the physical characteristics of the collagen and its performance in

the final product. It is important to note that thermal degradation is usually manifested as low molecular weight oligomers breaking away from the collagen triple helix rather than complete denaturation of the triple helix into gelatin. Finally, strong concentrations of mineral acids such as HCL and H₂SO₄ can cause irreversible chemical degradation that is similar to thermal degradation.

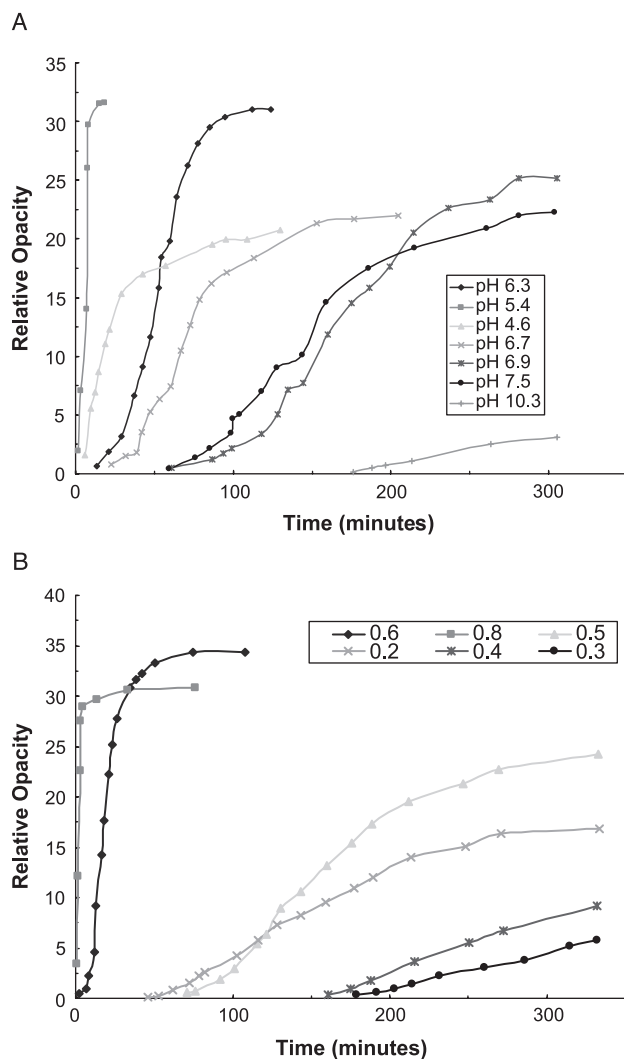


Fig. 3 A) The effect of pH on the rate of fibrillogenesis. These data are for a 0.1% solution at constant temperature of 23°C and IS of 0.4. Note that the rate of fibrillogenesis is fastest near the isoelectric pH, 5.5, of the collagen and not at physiological pH conditions. B) The effect of ionic strength on collagen fibrillogenesis. These data are for 0.1% collagen solutions at constant temperature of 23°C and pH of 7.5. Increasing the ionic strength of the buffer has the effect of neutralizing the ionic repulsions present within the triple helix allowing hydrophobic/hydrophilic interactions to predominate. In both cases, the buffer used was sodium phosphate buffered saline. (From Ref. [6].)

Fibrillogenesis

The self-assembly of collagen fibrils has been studied extensively in vitro for many years. Studies have measured the effect of temperature, pH, ionic strength, and the types of ions on the rate of fibrillogenesis and the size and type of fibrils obtained under various experimental conditions.^[1,4-18] These investigations have demonstrated that proteoglycans, glycosaminoglycans, and carbohydrates alter the rate of fibrillogenesis and the size and type of fibrils obtained.^[19-21] The general procedure for these investigations is to mix the collagen solution with an equal part of double strength buffer, heat it to reaction temperature, and monitor the increase in opacity or absorbance with respect to time. The absorbance curve has three distinct phases: a lag phase with no increase in absorbance; a growth phase; and a plateau phase in which the reaction has come to completion. These investigations show that the rate of collagen fibrillogenesis increases with increasing temperature occurring between 16°C and 37°C, reacts readily in the pH range of 6.0–8.0, and increases with increasing ionic strength, as can be seen in Figs. 3A and 3B. Some studies have been done showing the effect of specific ion type on the length of the lag phase of the reaction. In general, near physiological conditions, increasing the ionic strength or lowering the pH lengthens the lag phase. The lengthening effect of the salt is both anion and cation specific with anions having the greater effect. The anion effect is in the order $\text{SCN}^- < \text{HCO}_3^- < \text{I}^- < \text{Br}^- < \text{SO}_4^{2-} < \text{S}_2\text{O}_3^{2-}$ and cations in the order $(\text{K}^+, \text{Na}^+, \text{Li}^+) < (\text{Ba}^{2+}, \text{Mg}^{2+})$.^[17]

CONCLUSION

Humankind has benefited from the use of collagen for thousands of years. From crude protection from the elements to scaffold materials for tissue engineering applications, collagen has proven itself to be a remarkably versatile and useful material. There are several means to isolate collagen, which depends on the desired end product though, most methods are based on either removing noncollagenous materials from the intact collagen matrix (leather and food applications) or extracting the collagen from noncollagenous material and then recovering the collagen from this extract. Both methods have benefits and drawbacks. Most recently, collagen has seen a revival in the field of tissue engineering in its use as a scaffold material. Tissue engineering is advancing at a rapid pace, which thus results in some exciting findings. It is conceivable that, in the near future, numerous pathological conditions will be treated by implantation of biologically engineered devices. Some of which will likely have collagen as a major component.

C

ARTICLES OF FURTHER INTEREST

Collagen; Collagen Fixation; Polyepoxy Compound Fixation; Tissue Engineering Scaffolds

REFERENCES

1. The Olde Hide House. *History of Leather*; 2003. www.leathertown.com. Ontario, Canada.
2. Heidemann, E. The Chemistry of Tanning. In *Collagen*; Nimni, M.E., Ed.; CRC Press: Boca Raton, FL, 1988; Vol. 3, 39–62.
3. Miller, E.J.; Rhodes, R.K. Preparation and Characterization of the Different Types of Collagen. In *Methods in Enzymology*; Cunningham, L.W., Frederiksen, D.W., Eds.; Academic Press: New York, 1982; Vol. 82, 33–64.
4. Wallace, D.G.; Thompson, A. Description of collagen fibril formation by a theory of polymer crystallization. *Biopolymers* **1983**, *22*, 1793–1811.
5. Wood, G.C. The precipitation of collagen fibers from solution. *Biochem. J.* **1960**, *75*, 598–605.
6. Cassel, J.M.; Mandelkern, L.; Roberts, D.F. The kinetics of the heat precipitation of collagen. *J. Am. Leather Chem. Assoc.* **1962**, *57*, 556–575.
7. Cassel, J.M. Collagen aggregation phenomena. *Biopolymers* **1966**, *4*, 989–997.
8. Yannas, I.V. Collagen and gelatin in the solid state. *J. Macromol. Sci., Rev. Macromol. Chem.* **1972**, *C7* (1), 49–104.
9. Comper, W.D.; Vies, A. The mechanism of nucleation for in vitro collagen fibril formation. *Biopolymers* **1977**, *16*, 2113–2131.
10. Trelstad, R.L.; Birk, D.E.; Silver, F.H. Collagen fibrillogenesis in tissues, in solution and from modeling: A synthesis. *J. Invest. Dermatol.* **1982**, *79*, 109s–112s.
11. Silver, F.H.; Birk, D.E. Kinetic analysis of collagen fibrillogenesis: I. Use of turbidity–time data. *Collagen Relat. Res.* **1983**, *3*, 393–405.
12. Birk, D.E.; Silver, F.H. Collagen fibrillogenesis in vitro: Comparison of types I, II, and III. *Arch. Biochem. Biophys.* **1984**, *235*, 178–185.
13. Birk, D.E.; Silver, F.H. Kinetic analysis of collagen fibrillogenesis: II. Corneal and scleral type I collagen. *Collagen Relat. Res.* **1984**, *4*, 265–277.
14. Holmes, D.F.; Capaldi, M.J.; Chapman, J.A. Reconstitution of collagen fibrils in vitro; the assembly process depends on the initiating procedure. *Int. J. Macromol.* **1986**, *8*, 161–166.
15. Yannas, I.V.; Lee, E.; Orgill, D.P.; Skrabut, E.M.; Murphy, G.F. Synthesis and characterization of a model extracellular matrix that induces partial regeneration of adult mammalian skin. *Proc. Natl. Acad. Sci.* **1989**, *86*, 933–937.
16. Birk, D.E.; Fitch, J.M.; Babiarz, J.P.; Doane, K.J. Collagen fibrillogenesis in vitro: Interaction of types I and V collagen regulates fibril diameter. *J. Cell. Sci.* **1990**, *95*, 649–657.
17. Veis, A.; Payne, K. Collagen Fibrillogenesis. In *Collagen*; Nimni, M.E., Ed.; CRC Press: Boca Raton, FL, 1988; Vol. 1, 113–137.
18. Birk, D.E.; Nurminskaya, M.V.; Zycband, E.I. Collagen fibrillogenesis in situ: Fibril segments undergo post-depositional modifications resulting in linear and lateral growth during matrix development. *Dev. Dyn.* **1995**, *202*, 229–243.
19. Obrink, B.; Wasteson, A. Nature of the interaction of chondroitin 4-sulfate and chondroitin sulfate-proteoglycan with collagen. *Biochem. J.* **1971**, *121*, 227–233.
20. Obrink, B.A. Study of the interactions between monomeric tropocollagen and glycosaminoglycans. *Eur. J. Biochem.* **1973**, *33*, 387–400.
21. Vogel, K.G.; Paulsson, M.; Heinegard, D. Specific inhibition of type I and type II collagen fibrillogenesis by the small proteoglycan of tendon. *Biochem. J.* **1984**, *223*, 587–597.
22. Reed, R.K.; Aukland, K. Transcapillary fluid balance in immature rats. Interstitial fluid pressure, serum and interstitial protein concentration, and colloid osmotic pressure. *Microvasc. Res.* **1977**, *14*, 37–43.
23. Aukland, K.; Reed, R.K. Interstitial-lymphatic mechanisms in the control of extracellular fluid volume. *Physiol. Rev.* **1993**, *73* (1), 1–78.
24. Zurovsky, Y.; Mitchell, G.; Hattingh, J. Composition and viscosity of interstitial fluid of rabbits. *Exp. Physiol.* **1995**, *80*, 203–207.
25. Borg, T.K.; Rubin, K.; Carver, W.; Samuel, A.; Terracio, L. The cell biology of the cardiac interstitium. *TCM* **1996**, *6* (2), 65–70.

Composites

Chia-Wei Wang

Department of Mechanical Engineering, University of Michigan, Ann Arbor, Michigan, U.S.A.

Ann Marie Sastry

Department of Mechanical Engineering, University of Michigan, and Department of Biomedical Engineering, University of Michigan, Ann Arbor, Michigan, U.S.A.



INTRODUCTION

Composite materials can be generally defined as those materials having two or more distinct material phases. With the advent of advanced polymeric composite materials, the term composite became somewhat synonymous with engineered carbon-epoxy, Kevlar-epoxy or ceramic- or metal-matrix composites, though this term later came to refer to a broader set of materials; more recently, the descriptor heterogeneous has been used to characterize study of such materials. Porous materials may also be considered composite materials, with one phase composed of void or air spaces. Examples of composites include the familiar carbon-epoxy airframe skins, and glass/epoxy or glass/polyester structural materials, e.g., helicopter rotor blades, or even furniture. Sporting goods, e.g., golf clubs, tennis rackets, and skis are also often constructed of advanced composites. Even wood, which contains reinforcing cellulose fibers, bone, which may be considered a porous reinforcement, at a smaller scale, extracellular matrices, reinforced by structural proteins such as collagen, that are surrounded by ground substance (Fig. 1) constitute composites. There is a large body of literature available on both properties^[1-5] and manufacturing^[6] of many types of engineered composite materials, and much of this work has found, and will continue to find application in improved understanding of heterogeneous biomaterials and design of biocompatible materials.^[7] As an example, of the 20–30% of the human body that is composed of proteins, up to 50% is collagen;^[8] collagen's precursors have been on the planet nearly as long as multicellular life.^[9] Undoubtedly, the need for improved micro and nanoscale models for the behavior of such critically important fibrous biomaterials will continue, and will support new insights into biochemistry and evolutionary science.

HISTORICAL BACKGROUND

General modeling of the properties of heterogeneous materials is of great importance to almost all engineering and scientific disciplines, and can be traced to the

mid-19th century (Table 1), in work on properties of gases.^[10-53] Closed-form solutions for effective properties in gases led to similar analyses for conductivity and stiffness in composite solids. Determination of engineering properties, from conductivity to stiffness, was classically accomplished via solution of Laplace's equation, whose linearizing field assumption allows simultaneous solution of a number of important problems using the same partial differential equations, rescaled using appropriate material constants (Tables 2 and 3, following the description in Ref. [54]). Other techniques that have been widely used to determine, or bound, properties of heterogeneous materials, include solutions of stress fields in representative volume elements (RVEs), models of anisotropic sheets, and models of continuum anisotropic phases. In this article, we do not attempt to survey each of these areas thoroughly; we aim, however, to give an overview of approaches in modeling composite materials, with specific results of classic models, and an eye toward the modeling of biological materials. Thus, we omit discussion of manufacture of composite materials (see, for example, Refs. [6,55,56]). Instead, we emphasize analysis of elastic and transport properties, both for their common roots in the literature, and also their importance in study of biomaterials. We begin with general formulations of anisotropic elasticity, and discuss simplifications for layered structures, which are abundant both in engineered and biological materials. Bounds on elastic properties are also discussed, since they allow estimation of material response (important in analysis of damage and growth modeling of bone, skin, and other tissues). And, a discussion of the role of phase geometry and percolation, relative to transport properties, is presented for its usefulness in estimation of both mechanical response and permeability of specific phases (e.g., structural proteins) in biomaterials.

CONTINUUM AND MICROMECHANICS OF COMPOSITE MATERIALS

Composite materials may be isotropic or anisotropic, depending on the shapes, locations, and relative sizes

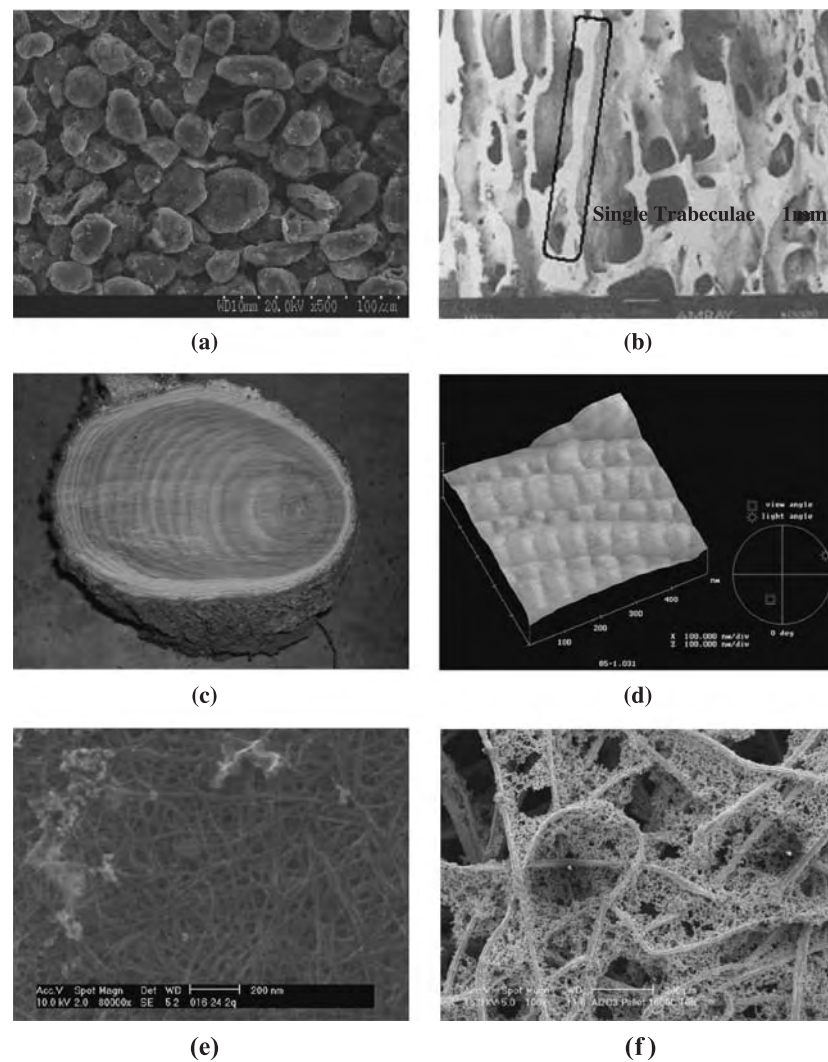


Fig. 1 Examples of composite materials, including (a) an anode of a Li-ion battery, containing carbon particles and polymeric binder, (b) trabecular bone (courtesy Dr. Scott Hollister, University of Michigan), (c) wood of an elm tree, and (d) collagen fibrils of the rat sciatic nerve perineurium, (e) carbon nanotube sheet, and (f) substrate of NiMH electrode.

of the material phases. Many particulate, porous, or short-fiber systems, for example, exhibit isotropic stiffness and conductivity; long-fiber and fabric-reinforced systems generally exhibit some degree of anisotropy.

Broadly, analysis of composite materials can be divided into two categories. Work that incorporates anisotropy into material models without detailed modeling of each phase is termed continuum mechanics, and can be used to predict effective properties of a heterogeneous material (see Ref. [57] for example). Work that directly models the shapes, locations, and relative sizes of model phases in a material is termed micromechanics, and more recently, nanomechanics, though the latter properly includes atomistic or molecular dynamics modeling and often is applied only to very small volumes due to its inherent computational intensiveness. Micromechanics is generally used to determine the details of stress, current, or other distributions

(Table 2) within a heterogeneous material, along with effective properties (e.g., Refs. [1,58]). The units of the common parameters used in the derivation of effective properties and their governing equations are listed in Table 3 and Table 4. Though theories for failure of material have been developed using continuum mechanics,^[59,60] understanding of specific failure mechanisms often requires analysis of the load sharing among the constituent materials, especially for brittle reinforcements. Statistical approaches have been shown to be useful in developing scaling rules for failure (e.g., Refs. [61–65]). The subject of failure of heterogeneous materials is quite broad and spans modeling of ductile failure, fracture, fatigue, and creep, to name a few key phenomena. Here, we introduce modeling of constitutive properties of elastic, heterogeneous materials, from the continuum to microscale, and comment on applications for both constitutive and failure modeling.

Table 1 Contributions to heterogeneous mechanics, including work in disperse gases

Disperse gases		Dielectrics		Disperse media	
1850	<i>O.F. Mossotti</i>	1955	<i>W.F. Brown Jr.</i>	1972	W.B. Russel and A. Acrivos
1879	<i>R. Clausius</i>	1956	J.D. Eshelby	1973	W.B. Russel
1880	<i>L.V. Lorenz</i>	1956	<i>E.H. Kerner</i>	1975	<i>K.S. Mendelson</i>
1880	<i>H.A. Lorentz</i>	1957	J.D. Eshelby	1978	W.T. Doyle
1891	J.C. Maxwell	1958	<i>E. Kroner</i>	1978	R.C. McPhedran and D.R. McKenzie
1892	J.W. Rayleigh	1958	C. van der Pol	1978	R. Landauer
1906	<i>A. Einstein</i>	1960	R.E. Meredith and C.W. Tobias	1979	W.T. Perrins, McPhedran, and McKenzie
1912	<i>O. Wiener</i>	1962	Z. Hashin	1979	R.M. Christensen and K.H. Lo
1924	<i>H. Fricke</i>	1962	Z. Hashin and S. Shtrikman	1988	<i>D.S. McLachlan</i>
1924	<i>K. Lichtenecker</i>	1964	<i>J.B. Keller</i>	1990	<i>K.D. Bao, J. Axell, and G. Grimvall</i>
1925	I.Z. Runge	1964	Z. Hashin and B.W. Rosen	1990	J.M. Gudes and N. Kikuchi
1933	J.N. Goodier	1964	R. Hill	1991	G.Q. Gu and Z.R. Liu
1935	D.A.G. Bruggeman	1965	R. Hill	1993	G.Q. Gu
1947	<i>J.M. Dewey</i>	1965	B. Budiansky	1993	R.M. Christensen
1952	R. Landauer	1966	Z. Hashin	1995	S.Y. Lu
1954	A.V. Hershey				

Gases are in italics, conductivity of solids are in normal type, and mechanics of solids are in bold. (Adapted from Ref. [54].)

Table 2 Effective medium theories and solution using laplace's equation

Linear problem of interest for a two-phase material	Quantity represented by q	Quantity represented by $EE_j = U_j$	Transport coefficient K	Local differential equation satisfied in each phase (in steady state)
Thermal conduction	Heat flux	Temperature gradient	Thermal conductivity	$q_i = K_{ij}E_j = K_{ij}U_j$
Electrical conduction	Electric current	Electric field intensity	Electrical conductivity	$\nabla \cdot \vec{q} = 0$
Electrical insulation	Electric displacement	Electric field intensity	Dielectric constant	
Permeation of a porous medium consisting of a fixed array of small rigid particles with an incompressible Newtonian fluid	Force on particles in unit volume of mixture (= pressure gradient calculated from pressure drop between distant parallel planes)	Flux of fluid volume relative to particles	Permeability (Darcy constant divided by μ)	$\nabla p = \mu \nabla^2 \vec{u}$ $\nabla \cdot \vec{u} = 0$ where \vec{u} = velocity; p = pressure; μ = viscosity
Elasticity of a medium containing elastic inclusions embedded in an elastic matrix	Stress	Strain	Lame constants (or rigidity and bulk moduli)	$\vec{q} = 2\mu \vec{E} + \lambda \vec{E}l$ $\nabla \cdot \vec{q} = 0$ μ, λ = local Lame constants

(Adapted from Ref. [54].)

Table 3 Common units for conversion of parameters in Table 2

	q	U	E	K
Thermal conduction	J/m^2	K (temperature)	K/m	$J/m \cdot K$
Electrical conduction	Amp/m^2	V	V/m	S/m
Electrical insulation	C/m^2	V	V/m	f/m
Porous medium	m/s	m	1	m/s
Elasticity	F/m^2	m	1	F/m^2

CONTINUUM, ANISOTROPIC STIFFNESS

The number of stiffnesses required to fully characterize a material’s response depends upon the degree of its anisotropy. Counterintuitively, tensorial stiffnesses are denoted C_{ijkl} , and tensorial compliances are denoted S_{ijkl} . An elastic constitutive relation (using the notation of Fig. 2) can be expressed as either

$$\sigma_{ij} = C_{ijkl}\epsilon_{kl} \tag{1}$$

or

$$\epsilon_{ij} = S_{ijkl}\sigma_{kl} \tag{2}$$

respectively, where ϵ_{kl} is the infinitesimal strain tensor and σ_{ij} is the stress tensor. We note that 2-D sections of the internal structure of a composite material depend on the plane examined, but models using 3-D ellipsoidal or cylindrical inclusions can be used to represent a wide range of reinforcement shapes (Fig. 3), from particles (2-D circles or ellipses, or 3-D spheres or ellipsoids) to fibers (1-D lines, 2-D ellipses or circles, or 3-D cylinders or ellipsoids).

The number of nonzero components of the stiffness tensor and the relationships among its components can be determined using material symmetry and equilibrium considerations. As a first cut, the 81 coefficients in the C_{ijkl} tensor can be seen immediately to have only 36 independent coefficients due to the symmetry of σ_{ij} and ϵ_{kl}

$$\sigma_{ij} = \sigma_{ji} \text{ and } \epsilon_{kl} = \epsilon_{lk}$$

We can then write the constitutive rule in matrix form (e.g., Refs. [66,67] for notation and general methods

Table 4 Common governing equations for modeling physical phenomena

Equation	Formula	Phenomena	Solution
Wave	$\nabla^2 V = \kappa \frac{\partial V}{\partial t}$	1. Wave	1. Analytical solution by various transformations, e.g. Bäcklund transformation, Green’s function, integral transform, Lax Pair, separation of variables 2. Numerical solution, e.g. finite element method
Diffusion	$\frac{\partial V}{\partial t} = \kappa \nabla^2 V$	1. Heat conduction 2. Mass diffusion	1. Analytical solution by separation of variables, Laplace transform, Fourier transform, Green’s function 2. Numerical solution, e.g. finite element method
Poisson	$\nabla^2 V = -4\pi\rho$	1. Electrostatics with constant source or sink 2. Thermal field with constant source or sink	1. Analytical solution by separation of variables, 2. Numerical solution, e.g. finite element method
Laplace	$\nabla^2 V = 0$	1. Thermal conduction 2. Electrostatics 3. Incompressible fluid flow 4. Membrane mechanics 5. Elasticity	1. Analytical solution by separation of variables. 2. Numerical solution, e.g. finite element method

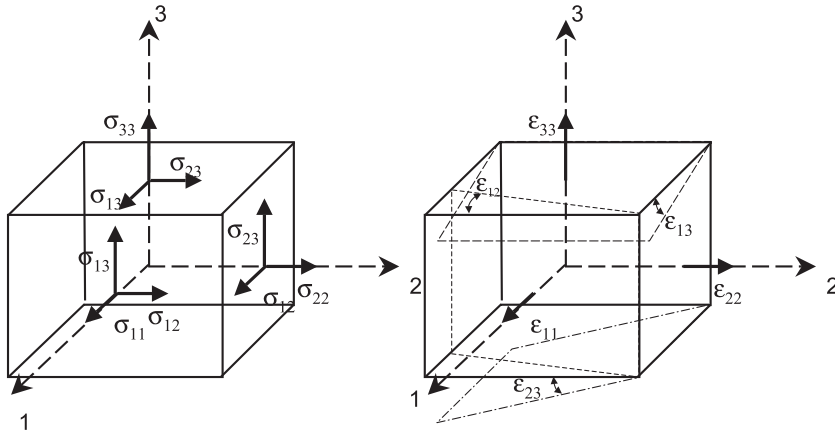


Fig. 2 Coordinate axes and components for the stress and strain tensors in 3D.

that follow) as

$$\begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \\ \sigma_6 \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{21} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ C_{31} & C_{32} & C_{33} & C_{34} & C_{35} & C_{36} \\ C_{41} & C_{42} & C_{43} & C_{44} & C_{45} & C_{46} \\ C_{51} & C_{52} & C_{53} & C_{54} & C_{55} & C_{56} \\ C_{61} & C_{62} & C_{63} & C_{64} & C_{65} & C_{66} \end{bmatrix} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \end{bmatrix} \quad (3)$$

where C_{ij} are the elastic stiffness coefficients, and

$$\begin{aligned} \sigma_1 &= \sigma_{11}; \sigma_2 = \sigma_{22}; \sigma_3 = \sigma_{33}; \sigma_4 = \sigma_{23}; \sigma_5 \\ &= \sigma_{13}; \sigma_6 = \sigma_{12} \\ \varepsilon_1 &= \varepsilon_{11}; \varepsilon_2 = \varepsilon_{22}; \varepsilon_3 = \varepsilon_{33}; \varepsilon_4 = 2\varepsilon_{23}; \varepsilon_5 \\ &= 2\varepsilon_{13}; \varepsilon_6 = 2\varepsilon_{12} \end{aligned} \quad (4)$$

The requirement of symmetry in C_{ij} further reduces the number of independent coefficients to 21.

Further, nontrivial reductions of the stiffness tensor (i.e., for a number of independent coefficients greater than 2) can be obtained for aligned fibrous materials, wherein the transverse arrangement of fibers determines the degree of anisotropy (Fig. 4). Materials containing square-packed fibers (Fig. 4a) are termed orthotropic, since they contain at least two mutually orthogonal planes of symmetry; in this case, the number of elastic coefficients is reduced to 9, as

$$\begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \\ \sigma_6 \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & 0 \\ C_{12} & C_{22} & C_{23} & 0 & 0 & 0 \\ C_{13} & C_{23} & C_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{66} \end{bmatrix} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \end{bmatrix} \quad (5)$$

Materials containing aligned, hexagonally packed fibers or randomly arranged fibers (Figs. 4b and 4c) are termed transversely isotropic, since the elastic properties are invariant with respect to an arbitrary rotation about an axis parallel to the fibers' axis; in this case, the

number of independent stiffness coefficients is reduced to 5, as

$$\begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \\ \sigma_6 \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & 0 \\ C_{12} & C_{11} & C_{13} & 0 & 0 & 0 \\ C_{13} & C_{13} & C_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{C_{11} - C_{12}}{2} \end{bmatrix} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \end{bmatrix} \quad (6)$$

Composites composed of layers of fibrous sheets, or laminae, are termed laminates or laminated composites. Analysis of the elastic properties of a stack of such layers (Fig. 5) requires modeling of both the in-plane properties of each layer and an assumed for an out-of-plane displacement function for the stack. Classically, this is accomplished using the well known laminate theory, in which an elastic stiffness matrix, the ABBD matrix, is assembled for the stack. A standard notation is shown in Fig. 5; the laminate code is enclosed in brackets, with sequential plies designated by the in-plane orientation of their fibers. Subscripts s and t are used to denote symmetric or total, respectively. Symmetry in the orientations of the layers about the midplane prevents bend-twist coupling in the laminate, discussed later. Also, many laminates are designed to have in-plane properties that are independent of rotation (quasi-isotropy), as in the example of Fig. 5.

Plane stress is assumed in each layer. Normal stress resultants in the x direction, N_x , in the y direction, N_y , and in shear, N_{xy} , are defined as

$$\begin{aligned} N_x &= \int_{-H/2}^{H/2} \sigma_x dz \\ N_y &= \int_{-H/2}^{H/2} \sigma_y dz \\ N_{xy} &= \int_{-H/2}^{H/2} \sigma_{xy} dz \end{aligned} \quad (7)$$

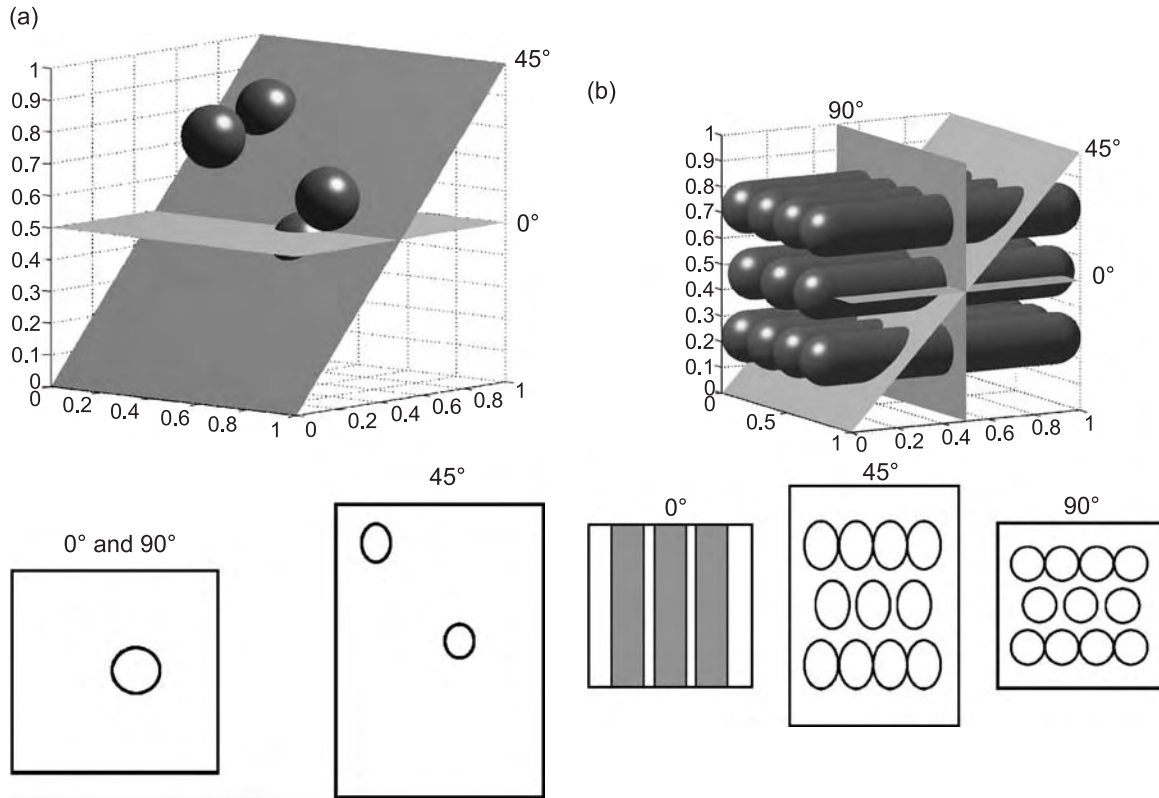


Fig. 3 Cross-sections of (a) particulate and (b) fibrous composite materials.

and moment resultants, M_x , M_y , and M_{xy} , are defined as

$$M_x = \int_{-H/2}^{H/2} \sigma_x z dz$$

$$M_y = \int_{-H/2}^{H/2} \sigma_y z dz$$

$$M_{xy} = \int_{-H/2}^{H/2} \tau_{xy} z dz$$

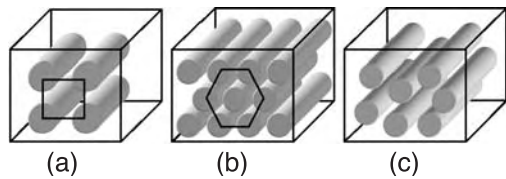


Fig. 4 Several classical transverse arrangements assumed for aligned-fiber composites, including (a) square packing, (b) hexagonal close-packing and (c) random packing.

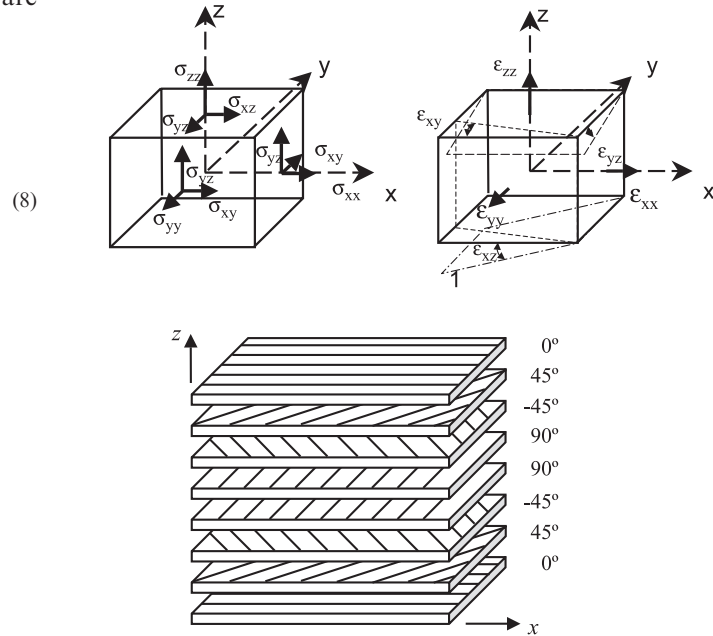


Fig. 5 Notation for laminated composites, with in-plane orientation of fibers in each layer designated in degrees. The example shown is a symmetric, quasi-isotropic laminate, $[0/\pm 45/90]_s$.

The plane stress assumption (σ_3 , τ_{23} , and $\tau_{13}=0$) results in the 2-D reduction of Eq. 3 to

$$\begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \tau_{12} \end{bmatrix} = \begin{bmatrix} Q_{11} & Q_{12} & 0 \\ Q_{12} & Q_{22} & 0 \\ 0 & 0 & Q_{66} \end{bmatrix} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \gamma_{12} \end{bmatrix} \quad (9)$$

where

$$\begin{aligned} Q_{11} &= C_{11} - \frac{C_{13}^2}{C_{33}} \\ Q_{12} &= C_{12} - \frac{C_{13}C_{23}}{C_{33}} \\ Q_{22} &= C_{22} - \frac{C_{23}^2}{C_{33}} \\ Q_{66} &= Q_{66} \end{aligned} \quad (10)$$

We note that for the 2-D case, these relations apply to both transversely isotropic or orthotropic laminae. Transformation of stiffnesses (Eq. 9) to a global coordinate system is required for each layer, where the transformed stiffness are denoted \bar{Q}_{ij} , as in

$$\begin{bmatrix} \sigma_x \\ \sigma_y \\ \tau_{xy} \end{bmatrix} = \begin{bmatrix} \bar{Q}_{11} & \bar{Q}_{12} & \bar{Q}_{16} \\ \bar{Q}_{12} & \bar{Q}_{22} & \bar{Q}_{26} \\ \bar{Q}_{16} & \bar{Q}_{26} & \bar{Q}_{66} \end{bmatrix} \begin{bmatrix} \varepsilon_x \\ \varepsilon_y \\ \gamma_{xy} \end{bmatrix} \quad (11)$$

and are obtained from the Q_{kl} by

$$\begin{aligned} \bar{Q}_{11} &= Q_{11} \cos^4 \theta + 2(Q_{12} + 2Q_{66}) \sin^2 \theta \cos^2 \theta \\ &\quad + Q_{22} \sin^4 \theta \\ \bar{Q}_{12} &= (Q_{11} + Q_{22} - 4Q_{66}) \sin^2 \theta \cos^2 \theta \\ &\quad + Q_{12}(\sin^4 \theta + \cos^4 \theta) \\ \bar{Q}_{16} &= (Q_{11} - Q_{12} - 2Q_{66}) \sin \theta \cos^3 \theta \\ &\quad + (Q_{12} - Q_{22} - 2Q_{66}) \cos \theta \sin^3 \theta \\ \bar{Q}_{22} &= Q_{11} \sin^4 \theta + 2(Q_{12} + 2Q_{66}) \sin^2 \theta \cos^2 \theta \\ &\quad + Q_{22} \cos^4 \theta \\ \bar{Q}_{26} &= (Q_{11} - Q_{22} - 2Q_{66}) \sin^3 \theta \cos \theta \\ &\quad + (Q_{12} - Q_{22} + 2Q_{66}) \sin \theta \cos^3 \theta \\ \bar{Q}_{66} &= (Q_{11} + Q_{22} - 2Q_{12} - 2Q_{66}) \sin^2 \theta \cos^2 \theta \\ &\quad + Q_{66}(\sin^4 \theta + \cos^4 \theta) \end{aligned} \quad (12)$$

Use of the Kirchhoff–Love hypothesis, that plane transverse sections remain plane during loading,

results in

$$\begin{aligned} u &= u_o - z \frac{\partial w_o}{\partial x} \\ v &= v_o - z \frac{\partial w_o}{\partial y} \\ w &= w_o \end{aligned} \quad (13)$$

for the displacement (u , v , w) of a point at (x, y, z) , given midplane displacements (u_o , v_o , w_o). The strains can then be calculated via

$$\begin{aligned} \varepsilon_x &= \frac{\partial u}{\partial x} = \frac{\partial u_o}{\partial x} - z \frac{\partial^2 w_o}{\partial x^2} = \varepsilon_x^o - z \frac{\partial^2 w_o}{\partial x^2} \\ \varepsilon_y &= \frac{\partial v}{\partial y} = \frac{\partial v_o}{\partial y} - z \frac{\partial^2 w_o}{\partial x^2} = \varepsilon_y^o - z \frac{\partial^2 w_o}{\partial x^2} \\ \gamma_{xy} &= \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} = \frac{\partial u_o}{\partial y} + \frac{\partial v_o}{\partial x} - 2z \frac{\partial^2 w_o}{\partial x \partial y} \\ &= \gamma_{xy}^o - 2z \frac{\partial^2 w_o}{\partial x \partial y} \end{aligned} \quad (14)$$

Assuming small rotations, displacements u and v in Eq. 13, the strains, can expressed as

$$\begin{aligned} \varepsilon_x &= \frac{\partial u}{\partial x} = \frac{\partial u_o}{\partial x} - z \frac{\partial^2 w_o}{\partial x^2} = \varepsilon_x^o + z\kappa_x^o \\ \varepsilon_y &= \frac{\partial v}{\partial y} = \frac{\partial v_o}{\partial y} - z \frac{\partial^2 w_o}{\partial x^2} = \varepsilon_y^o + z\kappa_y^o \\ \gamma_{xy} &= \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} = \frac{\partial u_o}{\partial y} + \frac{\partial v_o}{\partial x} - 2z \frac{\partial^2 w_o}{\partial x \partial y} \\ &= \gamma_{xy}^o + z\kappa_{xy}^o \end{aligned} \quad (15)$$

and Eq. 11 can be rewritten as

$$\begin{bmatrix} \sigma_x \\ \sigma_y \\ \tau_{xy} \end{bmatrix} = \begin{bmatrix} \bar{Q}_{11} & \bar{Q}_{12} & \bar{Q}_{16} \\ \bar{Q}_{12} & \bar{Q}_{22} & \bar{Q}_{26} \\ \bar{Q}_{16} & \bar{Q}_{26} & \bar{Q}_{66} \end{bmatrix} \begin{bmatrix} \varepsilon_x^o + z\kappa_x^o \\ \varepsilon_y^o + z\kappa_y^o \\ \gamma_{xy}^o + z\kappa_{xy}^o \end{bmatrix} \quad (16)$$

Finally, a global stiffness matrix for both in-plane displacements and out-of-plane moments can be expressed in terms of the ABBD matrix as

$$\begin{aligned} \begin{bmatrix} N_x \\ N_y \\ N_{xy} \end{bmatrix} &= \begin{bmatrix} A_{11} & A_{12} & A_{16} \\ A_{12} & A_{22} & A_{26} \\ A_{16} & A_{26} & A_{66} \end{bmatrix} \begin{bmatrix} \varepsilon_x^o \\ \varepsilon_y^o \\ \gamma_{xy}^o \end{bmatrix} \\ &\quad + \begin{bmatrix} B_{11} & B_{12} & B_{16} \\ B_{12} & B_{22} & B_{26} \\ B_{16} & B_{26} & B_{66} \end{bmatrix} \begin{bmatrix} \kappa_x^o \\ \kappa_y^o \\ \kappa_{xy}^o \end{bmatrix} \end{aligned} \quad (17)$$



$$\begin{bmatrix} M_x \\ M_y \\ M_{xy} \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} & B_{16} \\ B_{12} & B_{22} & B_{26} \\ B_{16} & B_{26} & B_{66} \end{bmatrix} \begin{bmatrix} \epsilon_x^o \\ \epsilon_y^o \\ \gamma_{xy}^o \end{bmatrix} + \begin{bmatrix} D_{11} & D_{12} & D_{16} \\ D_{12} & D_{22} & D_{26} \\ D_{16} & D_{26} & D_{66} \end{bmatrix} \begin{bmatrix} \kappa_x^o \\ \kappa_y^o \\ \kappa_{xy}^o \end{bmatrix} \quad (18)$$

where

$$\begin{aligned} A_{ij} &= \int_{-H/2}^{H/2} \bar{Q}_{ij} dz \approx \sum_{k=1}^N \bar{Q}_{ijk} (z_k - z_{k-1}) \\ B_{ij} &= \int_{-H/2}^{H/2} \bar{Q}_{ij} z dz \approx \sum_{k=1}^N \bar{Q}_{ijk} (z_k^2 - z_{k-1}^2) \\ D_{ij} &= \int_{-H/2}^{H/2} \bar{Q}_{ij} z^2 dz \approx \frac{1}{3} \sum_{k=1}^N \bar{Q}_{ijk} (z_k^3 - z_{k-1}^3) \end{aligned} \quad (19)$$

The matrices represent extensional stiffness (A), bending (D), and bend-twist coupling (B). If each layer of laminate is thin, the terms A_{ij} , B_{ij} , and D_{ij} can be approximated from the sums shown in Eq. 19, and we can write the laminate constitutive law as

$$\begin{bmatrix} N_x \\ N_y \\ N_{xy} \\ M_x \\ M_y \\ M_{xy} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{16} & B_{11} & B_{12} & B_{16} \\ A_{12} & A_{22} & A_{26} & B_{12} & B_{22} & B_{26} \\ A_{16} & A_{26} & A_{66} & B_{16} & B_{26} & B_{66} \\ B_{11} & B_{12} & B_{16} & D_{11} & D_{12} & D_{16} \\ B_{12} & B_{22} & B_{26} & D_{12} & D_{22} & D_{26} \\ B_{16} & B_{26} & B_{66} & D_{16} & D_{26} & D_{66} \end{bmatrix} \begin{bmatrix} \epsilon_x^o \\ \epsilon_y^o \\ \gamma_{xy}^o \\ \kappa_x^o \\ \kappa_y^o \\ \kappa_{xy}^o \end{bmatrix} \quad (20)$$

We note that the dimensions of these parameters are scaled per unit length of the laminate, i.e., N_i and A_{ij} are expressed in force per unit length, M_i and B_{ij} , in force, and D_{ij} as force \times length. As usual, ϵ is dimensionless; curvature κ is dimensioned as inverse length.

These relations have been widely implemented in free and commercially available codes, and are a cornerstone of laminate design. However, though laminate theory can satisfactorily model the behavior of the composite at the lamina or laminate scale, the details of load sharing among the constituent materials and many aspects of failure require more detailed modeling of the phases.

MICROMECHANICS OF ORDERED AND DISORDERED COMPOSITES

Micromechanical approaches take specific account of each phase in modeling. Classical approaches (Fig. 6a) span simple strength-of-materials series^[68] and parallel^[69] models, to early 2-D elastic field solutions for stresses around elliptical particles (e.g., Ref. [70]). All of these stemmed, along with much other work in the field, from the classical work of Eshelby.^[27] To model polymeric systems having viscoelastic behavior, elements in each model can be replaced with Voigt–Kelvin (parallel spring-dashpot) or Maxwell (series spring-dashpot) elements, as illustrated in Fig. 6b. In present applications, modeling of the details of even elastic load transfer in constituent phases (with engineering fibers of O(10–100 μ m)) in a simulation of structural properties (of components of O(>1 cm)) is still somewhat beyond computational capability, though

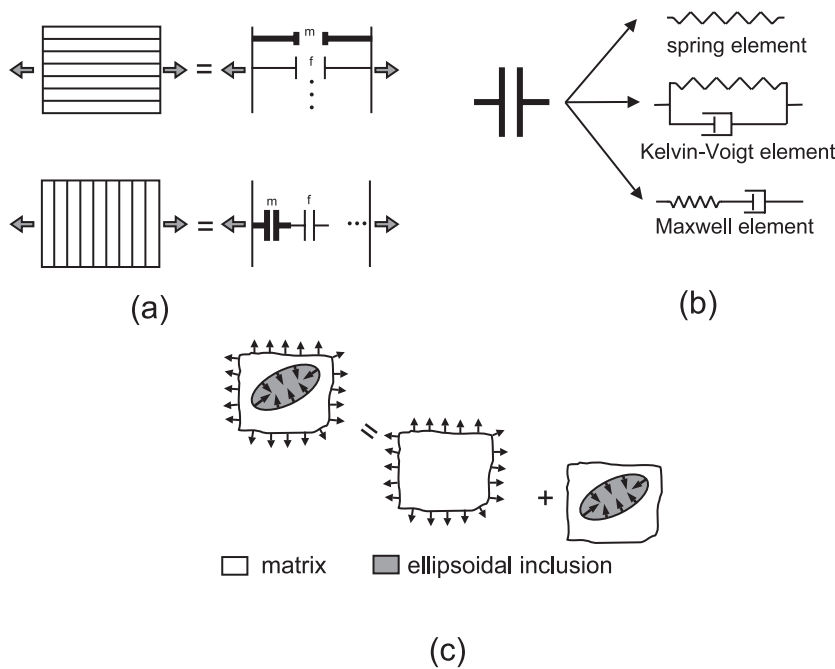


Fig. 6 Possible levels of detail in micromechanical models, from (a) simple rule of mixtures in series and parallel for stiffness modeling, (b) the element in (a) can be represent for spring, Kelvin–Voigt (spring-dashpot in parallel) element, or Maxwell (spring-dashpot in series) element, to (c) derivation of the stress fields around elliptical particles.

detailed finite element analyses of load transfer have found great utility in informing less-detailed models of structural response and failure (e.g., Ref. [67]). Here, we emphasize the classical work in analytical elasticity (Fig. 7), employing simplifying assumptions for the shapes of the phases in composite media in order to directly perform energy and mechanics of materials analyses to derive field solutions for stresses, strains, and stiffnesses.

Though many fibers are anisotropic, especially with regard to thermal expansion (e.g., many carbon fibers have a slightly negative axial coefficient of thermal expansion and a positive transverse CTE), the classical micromechanics models view each phase as isotropic. Thus, we complete the reduction of the stiffness tensor described earlier for single, isotropic phases to (see below)

$$\begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \\ \sigma_6 \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{12} & 0 & 0 & 0 \\ C_{12} & C_{11} & C_{12} & 0 & 0 & 0 \\ C_{12} & C_{12} & C_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & (C_{11} - C_{12}) & 0 & 0 \\ 0 & 0 & 0 & 0 & (C_{11} - C_{12}) & 0 \\ 0 & 0 & 0 & 0 & 0 & (C_{11} - C_{12}) \end{bmatrix} \times \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \end{bmatrix} \quad (21)$$

for which there are only two independent components of C_{ij} . Engineering constants, such as Young's modulus E , shear modulus μ , and bulk modulus K can be readily obtained from these tensorial stiffnesses. For example, if we specify

$$\varepsilon_1 = \varepsilon_2 = \varepsilon_3 = \varepsilon \quad (22)$$

and

$$\sigma_1 = \sigma_2 = \sigma_3 = \sigma \quad (23)$$

we can find from the definition of bulk modulus K ,

$$\sigma = 3K\varepsilon \quad (24)$$

that

$$K = \frac{1}{3}(C_{11} + 2C_{12}) \quad (25)$$

The shear moduli are defined as

$$\mu = \mu_{12} = \mu_{31} = \mu_{23} = \frac{1}{2}(C_{11} - C_{12}) \quad (26)$$

and Young's modulus and Poisson's ratio can be then determined by bulk modulus K and shear modulus μ as

$$E = \frac{9K\mu}{3K + \mu} \quad (27)$$

and

$$\nu = \frac{3K - 2\mu}{2(3K + \mu)} \quad (28)$$

Both somewhat realistic and purely theoretical constructs have been used to derive effective properties (both conductive and mechanical) using micromechanical analyses, and both classes of RVEs have contributed to literature on the bounding of effective properties. In the first category, wherein fibers or particles are somewhat literally represented in 2-D as circles or ellipsoids (see Fig. 3), the derivations of Maxwell^[14] and Rayleigh^[15] were among the first to allow calculation of effective conductive properties based on relative fraction of materials packed in a regular fashion. Later, Bruggeman analyzed both a "symmetric effective medium" and an "asymmetrical effective medium" by assuming a wide distribution for the sizes of inclusions.^[22] In 1962, Hashin introduced a composite sphere model (see Fig. 7c) using Eshelby's energy approach (Fig. 6b) to develop a closed-form solution for effective stiffness of a continuous matrix phase infused with a variable-diameter sphere.^[31] The ratio of radii a/b for the phases was taken as a constant, proportional to the volumetric ratio of each composite sphere and independent of their absolute size.

Schemes involving theoretical material constructs have also allowed for solution of the field equations to estimate effective properties; these so-called self-consistent domains are analyzed by matching average stress and strain in the inclusion phases and the uniform stress and strain in the surrounding, infinite, isotropic medium. The first self-consistent formulations were developed by Hershey^[71] and Kröner^[72] in modeling polycrystalline media (Fig. 7d). Later, Budansky extended their work in order to determine



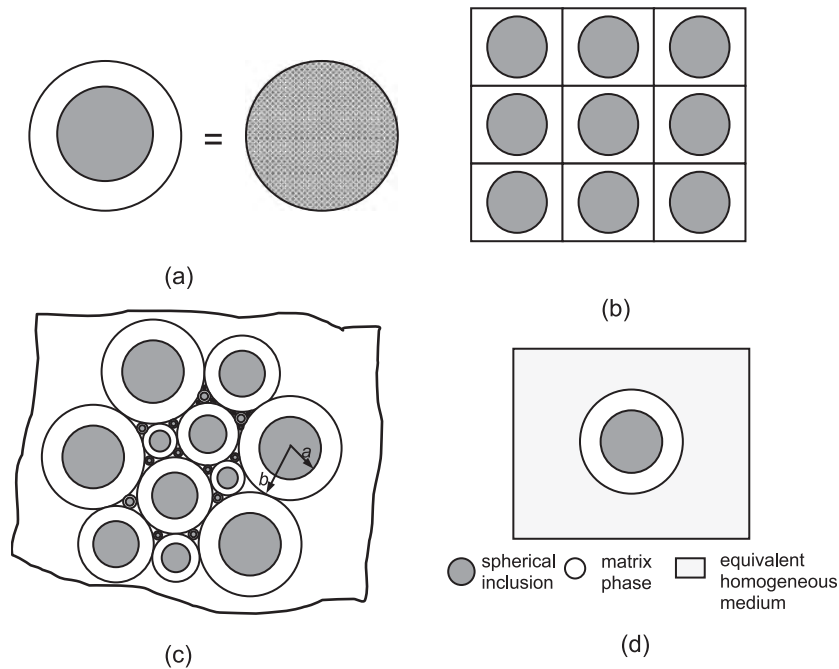


Fig. 7 Micromechanical models of ordered composite materials, including representative volume elements for the effective medium theories of (a) Maxwell's model (see Ref. [14]), (b) Rayleigh's rectangular array (see Ref. [15]), (c) Hashin's composite sphere model (see Ref. [31]) and (d) Christiansen and Lo's three phases model (see Ref. [46]).

bounds on shear and bulk moduli of multiphase materials.^[37] Hashin and Shtrikman used a variational method to minimize potential energy in a model domain and provide rigorous bounds on effective magnetic permeability, shear, and bulk moduli of multiphase materials.^[32,73-75] Christensen and Lo later devised a three-phase self-consistent approach to calculate effective shear moduli of materials containing spherical and cylindrical inclusions.^[46]

Importantly, many classical results for effective conductivity and effective bulk modulus coincide for both physical and self-consistent RVEs. Effective elastic properties can be found for these arrangements via minimization of potential energy, or minimization or work principles, expressing stored elastic energy as either

$$U_C = \frac{1}{2} \sum_{i=1}^6 \sum_{j=1}^6 \varepsilon_i C_{ij} \varepsilon_j \quad (29)$$

or

$$U_S = \frac{1}{2} \sum_{i=1}^6 \sum_{j=1}^6 \sigma_i (S_{ij}/m_i m_j) \sigma_j \quad (30)$$

These two expressions result in different magnitudes for the total stored elastic energy in heterogeneous domains due to simplifying assumptions regarding stress and strain fields. Together, they allow calculation of bounds on properties of anisotropic fields of fibrous (Table 5) and particulate (Table 6) materials, with assumed geometry and volume fraction of phases (summarized in compact form, for example, by McCullough,^[76] following work by Hashin and Shtrikman^[74,75]). These

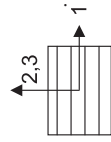
estimates are useful for design for stiffness, in both fibrous and particulate materials, and can be readily used to model a wide range of materials, and compare well with elastic finite element simulations in prediction of stiffnesses^[67] we note that these models assume transverse isotropy in the composite, with isotropic fiber and matrix.

Though effective properties can be estimated and bounded using these simplified elastic estimates, however, the importance of disorder in material geometry at the scale at which the material is used must frequently be considered. Such disorder can create internal non-uniformities in material response, and thus variability among devices or structures created from the material. Also, variability in location of reinforcement phases can result in higher stress concentrations in the material, producing earlier-than-expected failure (Fig. 8^[77-79]).

Indeed, real, disordered, or stochastic materials exhibit a range of values in material response, and attempts have been made for various geometries (e.g., randomly laid fibers or particles, as in Fig. 9) to bound these properties. Work of this type models disorder in a material specifically, rather than through use of an energy approach, as described in the previous section, to simply bound properties of an ordered, representative arrangement. In early work, Bergman devised an analytical method to determine bounds for dielectric conductivity^[80] Lurie and Cherkaev^[81] and Milton and Kohn^[82] presented bounds for macroscopically anisotropic media by extending the work of Hashin and Shtrikman.^[32,73,74] Gibiansky and Torquato used a "translation method" to determine the rigorous bounds for the relation between conductivity and elastic moduli of two-dimensional, globally isotropic

Table 5 Summary of the rigorous bounds of physical properties for transverse isotropic composites materials

Property	Lower bound	Upper bound
\mathbf{E}_1	$E_1^L = \mathbf{v}^f E_1^f + \mathbf{v}^m E_1^m + \frac{4 \left(\mathbf{v}_{12}^m - \mathbf{v}_{12}^f \right)^2 k^f k^m G_{12}^m}{\left[k^m k^f + G_{12}^m (\mathbf{v}^f k^f + \mathbf{v}^m k^m) \right]}$	$E_1^U = \mathbf{v}^f E_1^f + \mathbf{v}^m E_1^m + \frac{4 \left(\mathbf{v}_{12}^f - \mathbf{v}_{12}^m \right)^2 k^f k^m G_{12}^f}{\left[k^m k^f + G_{12}^f (\mathbf{v}^f k^f + \mathbf{v}^m k^m) \right]}$
$\nu_{12} = \nu_{13}$	$\nu_{12}^L = \nu_{13}^L = \mathbf{v}^m \nu_{12}^m + \mathbf{v}^f \nu_{12}^f + \frac{\left(\mathbf{v}_{12}^m - \mathbf{v}_{12}^f \right) (k^f - k^m)}{\left[k^m k^f + G_{12}^m (\mathbf{v}^m k^m + \mathbf{v}^f k^f) \right]}$	$\nu_{12}^U = \nu_{13}^U = \mathbf{v}^m \nu_{12}^m + \mathbf{v}^f \nu_{12}^f + \frac{\left(\mathbf{v}_{12}^f - \mathbf{v}_{12}^m \right) (k^m - k^f)}{\left[k^m k^f + G_{12}^f (\mathbf{v}^m k^m + \mathbf{v}^f k^f) \right]}$
\mathbf{G}_{12}	$G_{12}^L = \frac{G_{12}^m \left[\left(G_{12}^f + G_{12}^m \right) + \mathbf{v}^f \left(G_{12}^f - G_{12}^m \right) \right]}{\left[\left(G_{12}^f + G_{12}^m \right) - \mathbf{v}^f \left(G_{12}^f - G_{12}^m \right) \right]}$	$G_{12}^U = \frac{G_{12}^f \left[\left(G_{12}^f + G_{12}^m \right) + \mathbf{v}^m \left(G_{12}^m - G_{12}^f \right) \right]}{\left[\left(G_{12}^f + G_{12}^m \right) - \mathbf{v}^m \left(G_{12}^m - G_{12}^f \right) \right]}$
\mathbf{G}_{23}	$G_{23}^L = \frac{G_{23}^m \left[k^f \left(G_{23}^f + G_{23}^m \right) + 2G_{23}^f G_{23}^m + \mathbf{v}^f k^m \left(G_{23}^f - G_{23}^m \right) \right]}{\left[k^m \left(G_{23}^f + G_{23}^m \right) + 2G_{23}^f G_{23}^m - \mathbf{v}^f \left(G_{23}^f - G_{23}^m \right) \right] \left(k^m + 2G_{23}^m \right)}$	$G_{23}^U = \frac{G_{23}^f \left[k^m \left(G_{23}^f + G_{23}^m \right) + 2G_{23}^f G_{23}^m + \mathbf{v}^m k^f \left(G_{23}^f - G_{23}^m \right) \right]}{\left[k^f \left(G_{23}^f + G_{23}^m \right) + 2G_{23}^f G_{23}^m - \mathbf{v}^m \left(G_{23}^f - G_{23}^m \right) \right] \left(k^f + 2G_{23}^f \right)}$
\mathbf{k}	$k^L = \frac{k^m \left(k^f + G_{23}^m \right) + \mathbf{v}^f G_{23}^m \left(k^f - k^m \right)}{\left[\left(k^f + G_{23}^m \right) - \mathbf{v}^f \left(k^f - k^m \right) \right]}$	$k^U = \frac{k^f \left(k^m + G_{23}^f \right) + \mathbf{v}^m G_{23}^f \left(k^m - k^f \right)}{\left[\left(k^m + G_{23}^f \right) - \mathbf{v}^m \left(k^m - k^f \right) \right]}$
$\mathbf{E}_2 = \mathbf{E}_3$	$E_2^L = E_3^L = \frac{1}{\frac{1}{4G_{23}^L} + 4k^L + \frac{1}{E_1^L}} \frac{1}{\left(\nu_{12}^U \right)^2}$	$E_2^U = E_3^U = \frac{1}{\frac{1}{4G_{23}^U} + 4k^U + \frac{1}{E_1^U}} \frac{1}{\left(\nu_{12}^L \right)^2}$
ν_{23}	$\nu_{23}^L = \frac{2E_1^L k^L - E_1^L E_2^L - 4 \left(\nu_{12}^U \right)^2 k^L E_2^L}{2E_1^L k^L}$	$\nu_{23}^U = \frac{2E_1^U k^U - E_1^U E_2^U - 4 \left(\nu_{12}^L \right)^2 k^U E_2^U}{2E_1^U k^U}$



Superscripts “L” and “U” indicate lower and upper bounds, respectively, for corresponding physical properties. Superscripts “f” and “m” indicate fiber and matrix phases, respectively. All formulae adapted from Ref. [76].



Table 6 Summary of rigorous bounds physical properties for particulate-filled composite materials

Property	Lower bound	Upper bound
K	$k^L = \frac{k^m(k^f(1 + (v^f)^2(2v^f - 2v^m) + v^m) + 2k_m((v^f)^2 - 1)(v^m - v^f))}{k^f(v^m)(1 + v^m) + k^m(2 + v^f + (v^f - 4)v^m)}$	$k^U = \frac{k^f \{ 2k^f (v^f)^2 (v^f - v^m) + k^m [(2v^f + v^m)(3v^m - 3v^f + 4(v^f)^2)] \}}{-k^m v^f (v^m + 2v^f) + k^f [v^f (5v^f + 4v^m) - 3]}$
G	$G^L = -\frac{G^m [G^m (v^m)(7v^f + 2v^m) + G^f (8 + 7v^f - 5(2 + v_f)v_m)]}{2G^f (v^m)(v^m - 4v^f) + G^m (-7v^f - 2v^m + 2v^f (v^m - 4v^f))}$	$G^U = \frac{G^f [G^f (v^f) (7v^m - 2v^f) + G^m (15 + 7v^f (5v^f - 22))]}{2G^m v^f (4v^m - v^f) + G^f (15 + v^f (10v^f - 23))}$
E	$E^L = \frac{9k^L G^L}{3k^L + G^L}$	$E^U = \frac{9k^U G^U}{3k^U + G^U}$
v	$v^L = \frac{3k^L - 2G^L}{2(3k^L + G^L)}$	$v^U = \frac{3k^U - 2G^U}{2(3k^U + G^U)}$

Superscripts “L” and “U” indicate lower and upper bounds, respectively, for corresponding physical properties. Superscripts “f” and “m” indicate filler and matrix phases, respectively. Physical properties of filler phase are assumed to be superior to those of the matrix phase. All formulae adapted from Ref. [76].

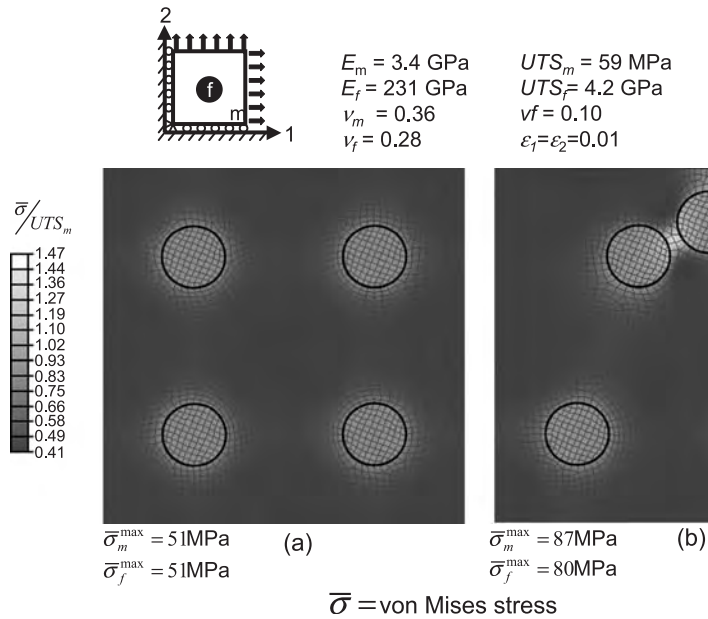


Fig. 8 2D models of transverse stress distributions in an elastic, carbon-epoxy, fiber-reinforced composites. For the simulations shown, $E_f = 231 \text{ GPa}$, $E_m = 3.4 \text{ GPa}$, $\nu = 0.28$, $\nu_m = 0.36$ and $f = 0.1$; both simulations are subjected to a biaxial tension of 1%, resulting in far field von Mises stresses of 26 MPa. In the regular array (a), the highest stress is 51 MPa at both matrix and fiber phases, whereas in the disordered array (b), the highest stresses are 87 MPa in matrix phase and 80 MPa in fiber phases between the closely-spaced fibers. Material constants for AS4C Hexel[®] Carbon fibers were obtained from Refs. [77,78]. Material constants for Hercules 3501-6 epoxy were obtained from Refs. [6,79].

composite materials.^[83] Later, Torquato and coworkers used a discrete network and homogenization theory to determine the effective mechanical and transport properties of cellular solids.^[84] Studying fibrous materials, Lu, Carlsson, and Andersson used a micromechanical approach to obtain rigorous bounds on elastic properties.^[85–87] And Ostoja-Starzewski et al. developed network techniques to simulate effective properties of generalized composites by changing the

spring constants of individual springs within the network to model multiphase material properties.^[88] Sastry and coworkers studied the stochastic fibrous networks to determine both conductivity and variance in conductivity in battery materials.^[89,90] This work was later extended to model fiber type bonding conditions, and failure mechanisms in greater detail^[90,91] and was also extended to include the effects of fiber waviness on material mechanical properties^[92–94] to model variance in electrical conductivity of porous networks with elliptical particles.^[95]

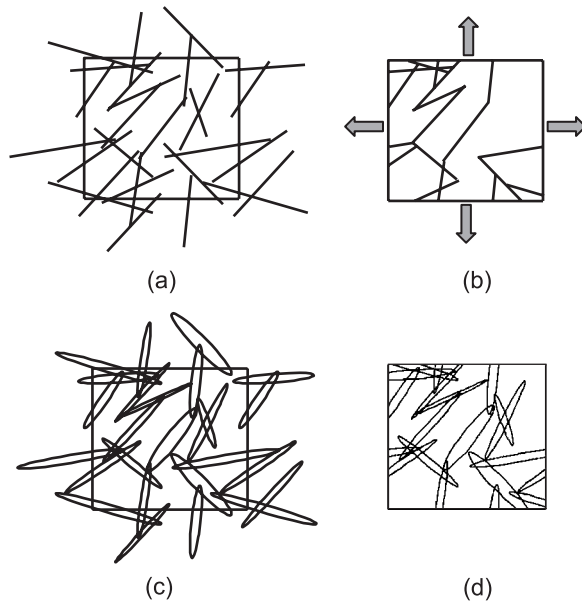


Fig. 9 Micromechanical models of disordered composite materials, including disordered 1D fibrous inclusions, with (a) randomly-laid, and (b) reduced, and periodic geometry, and 2D, fibrous or particulate inclusions, with (c) randomly-laid and (b) reduced, and periodic geometry.

PHASE CONTINUITY AND PERCOLATION IN DISORDERED COMPOSITES

All of these efforts further underscored the importance of variance in properties in real materials. Indeed, a key requirement in producing dramatic changes in material transport properties (e.g., thermal and electrical conduction) is percolation of the additive phase. We can define percolation as the formation of at least one, continuous, domain-spanning path of the percolating phase in the material. Dramatic improvements in computing speeds from the 1980s to the present have allowed direct, stochastic simulation of transport in disordered arrays.

Key features of such models include physically realistic simulation domains and use of many statistically equivalent realizations, i.e., domains in which the statistical parameters describing particles, sizes, etc. are the same for all, but the locations, sizes, etc. are different in each. As a result, the minimum amount of a phase required to produce dramatic improvements in composite properties can be determined, along with

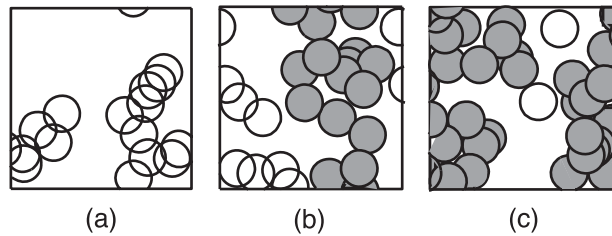


Fig. 10 Percolating and non-percolating arrays of circles, with volume fractions (a) 32.7%, (b) 51.2%, and (c) 64.0%.

discrete and averaged predictions of material properties. Percolation concepts have been used not only in design of materials for mechanical^[96–98] properties, but also in filtration^[99–101] and conductive^[89–103] properties.

Classic work by Kirkpatrick, who studied site and bond percolation using a resistor network, pointed out the importance of the percolation threshold, or threshold volume fraction above which a percolating phase would form at least one single, domain-spanning path (Fig. 10).^[104] In 1988, McLachlan introduced a more general effective-media equation for binary conductivity media.^[47] The fact that higher aspect ratio phases percolate at lower volume or area fractions than lower aspect ratio phases has been well documented. In early work, Kirkpatrick's simulations showed that the percolation threshold ρ_c , i.e., the density or volume fraction of the fiber phases at percolation onset, exhibited a power law dependence upon bond fraction v .^[104] Pike and Seager also examined conduction and percolation phenomena in stick networks

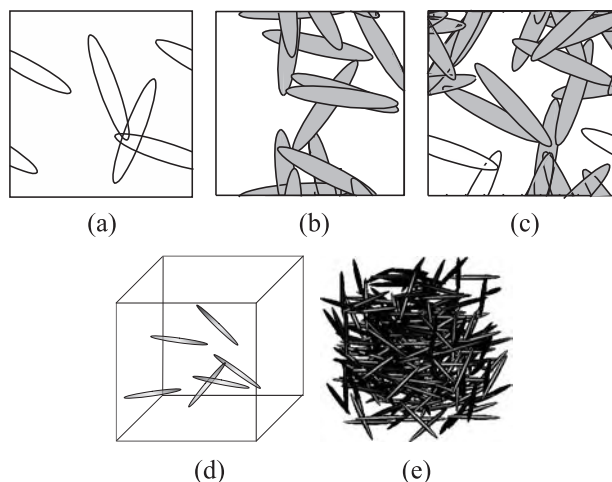


Fig. 11 Two-dimensional percolating and nonpercolating arrays of ellipses of aspect ratio 6 and volume fractions (a) 17.9%, (b) 31.6%, and (c) 50.7%. Three-dimensional percolating and nonpercolating arrays of ellipsoids are also shown, of aspect ratio 10, and volume fraction (d) 6%, and (e) 20%.

(among others), using 2-D and 3-D Monte Carlo simulations.^[105]

Recently, analytical approximations of the percolation points have been developed for both 2-D and 3-D arrays of generalized ellipses and ellipsoids of uniform shape and size,^[106,107] which verified and extended earlier results on simulation of simpler 2-D networks of 1-D fiber percolation,^[97,98] 3-D networks of 2-D ellipsoids, and other analytical approaches for determination of percolation of circular arrays;^[108–114] combined results^[91,106] are shown in Fig. 11. These illustrate that the percolation point in realistic materials is probabilistic (i.e., only a statistical estimation of percolation point can be made, for any given volume fraction of particles), and also that 2-D fiber models in models are quite satisfactory for determining percolation properties for aspect ratios greater than 100, as shown in Fig. 12. Similarly, Fig. 13 shows that percolation probability of ellipsoid models in 3-D model is also strongly dependent upon of aspect ratio.

APPLICATION OF COMPOSITE THEORIES TO BIOMATERIALS

Modeling of the shapes and effects of various phases in materials is tremendously important in understanding the combined mechanical and physiological role of biomaterials. Natural materials exhibit a high degree of variability, and use of statistical theories touched upon in this chapter can be helpful in anticipating differences in clinical results of both in-vitro and in-vivo tissue response. Generalized domains, particularly fields of ellipses and ellipsoidal particles, can be used to describe a wide variety of materials and bound a number of important effective engineering properties. Percolative properties have important implications

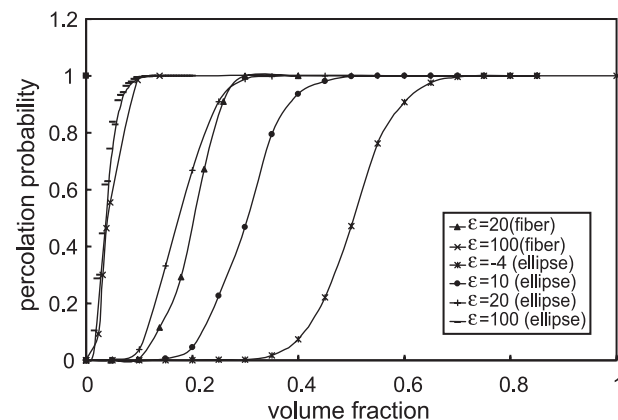


Fig. 12 Probability of percolation in 2D arrays, for various particle-fiber geometries. (Adapted from Ref. [106].)

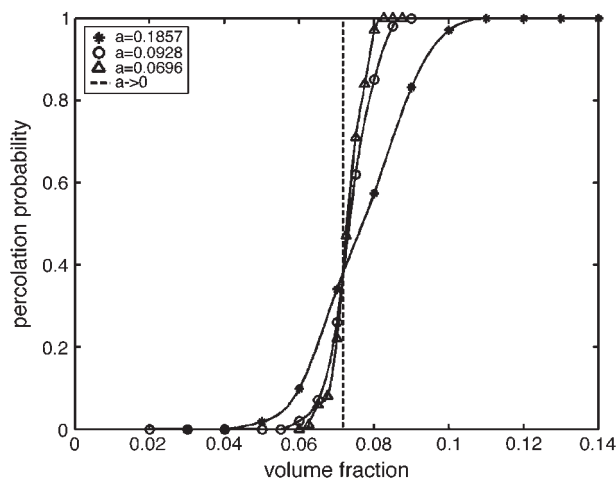


Fig. 13 Simulation results for percolation probability versus particle size, for arrays of overlapping ellipsoids of revolution. In this figure, 'a' represents the semi-axis length, and the aspect ratio of the ellipsoids is 10. (Adapted from Ref. [107].)

for understanding the role of particulates (either voids or material phases) and fibers, both adaptation and selection, in biomaterials (Fig. 12).

Thus, in conclusion, the modeling of heterogeneous domains performed by the composites community has broad application in modeling biomaterials. Elasticity provides an initial estimate of properties, and percolation theories provide a method of determining the connectivity of phases, and thus their importance in transport.

ACKNOWLEDGMENTS

The authors wish to acknowledge the invaluable assistance of Dr. Yun-Bo Yi, Dr. Bradley Layton, Dr. Hui Wang, and Dr. Scott Hollister, who provided simulation data and/or images for this chapter, and of Mr. Taeyong Kim, in helping to collect materials.

ARTICLES OF FURTHER INTEREST

Alumina; Biocomposites; Degradable Polymer Composites; Finite Element Analysis

REFERENCES

- Christensen, R.M. *Mechanics of Composite Materials*; John Wiley and Sons: New York, 1979.
- Carlsson, L.A.; Pipes, R.B. *Experimental Characterization of Advanced Composite Materials*; Prentice-Hall: Englewood Cliffs, NJ, 1987.
- Vinson, J.R. *The Behavior of Sandwich Structures of Isotropic and Composite Materials*; Technomic Pub. Co.: Lancaster, PA, 1999.
- Carlsson, L.A.; Gillespie, J.W.; Zweben, C.H. *Delaware Composites Design Encyclopedia*; Technomic Pub. Co.: Lancaster, PA, 1989.
- Chou, T.-W. *Microstructural Design of Fiber Composites*; Cambridge University Press: Cambridge, 1992.
- Gutowski, T.G.P. *Advanced Composites Manufacturing*; John Wiley & Sons: New York, 1997.
- Taboas, J.M.; Maddox, R.D.; Krebsbach, P.H.; Hollister, S.J. Indirect solid free form fabrication of local and global porous, biomimetic and composite 3d polymer-ceramic scaffolds. *Biomaterials* **2003**, *24* (1), 181–194.
- Kimball, J.W. *Biology*, 2nd Ed.; Addison-Wesley Pub. Co.: Reading, MA, 1968.
- Rubin, G.M.; Yandell, M.D.; Wortman, J.R.; Miklos, G.L.G.; Nelson, C.R.; Hariharan, I.K.; Fortini, M.E.; Li, P.W.; Apweiler, R.; Fleischmann, W.; Cherry, J.M.; Henikoff, S.; Skupski, M.P.; Misra, S.; Ashburner, M.; Birney, E.; Boguski, M.S.; Brody, T.; Brokstein, P.; Celniker, S.E.; Chervitz, S.A.; Coates, D.; Cravchik, A.; Gabrielian, A.; Galle, R.F.; Gelbart, W.M.; George, R.A.; Goldstein, L.S.B.; Gong, F.C.; Guan, P.; Harris, N.L.; Hay, B.A.; Hoskins, R.A.; Li, J.Y.; Li, Z.Y.; Hynes, R.O.; Jones, S.J.M.; Kuehl, P.M.; Lemaitre, B.; Littleton, J.T.; Morrison, D.K.; Mungall, C.; O'farrell, P.H.; Pickeral, O.K.; Shue, C.; Vosshall, L.B.; Zhang, J.; Zhao, Q.; Zheng, X.Q.H.; Zhong, F.; Zhong, W.Y.; Gibbs, R.; Venter, J.C.; Adams, M.D.; Lewis, S. Comparative genomics of the eukaryotes. *Science* **2000**, *287* (5461), 2204–2215.
- Mossotti, O.F. *Memorie di Matematica e Fisica della Societa Italiana delle Scienze*. **1850**, *24*, 49.
- Clausius, R. *Die mechanische warmetheorie II*. Braunschweig **1879**, 62.
- Lorenz, L.V. *Ann. Phys.* **1880**, *11*, 70.
- Lorentz, H.A. *Ann. Phys.* **1880**, *9*, 641.
- Maxwell, J.C.A. *Treatise on Electricity and Magnetism*; Dover: New York, 1891.
- Rayleigh, L.J.W. On the influence of obstacles arranged in rectangular order upon the properties of a medium. *Philos. Mag. Ser.* **1892**, *5* (34), 481–502.
- Einstein, A. Eine neue bestimmung der molekuldimensionen. *Ann. Phys.* **1906**, *19*, 289.
- Wiener, O. Die theorie des mischkörpers für das feld des stationären strömung. *Abh. Math.-Phys. Kl. Königl. Sächs. Gesel. Wissen.* **1912**, *32*, 509–604.
- Fricke, H. *Phys. Rev.* **1924**, *24*, 575.
- Lichtenecker, K. *Phys. Z.* **1926**, *27*, 115.
- Runge, I.Z. *Teck. Phys.* **1925**, *6*, 61.
- Goodier, J.N. Concentration of stress around spherical and cylindrical inclusion and flaws. *J. Appl. Mech.* **1933**, *55*, 39–44.
- Bruggeman, D.A.G. Berechnung verschiedener physikalischer konstanten von heterogenen substanzen: I. Dielektrizitätskonstanten und leifähigkeiten der mischkörper aus isotropen substanzen. *Ann. Phys.* **1935**, *24* (5), 636–679.



23. Dewey, J.M. The elastic constants of materials loaded with non-rigid fillers. *J. Appl. Phys.* **1947**, *18* (6), 578–581.
24. Landauer, R. The electrical resistance of binary metallic mixtures. *J. Appl. Phys.* **1952**, *23* (7), 779–784.
25. Hershey, A.V. The elasticity of an isotropic aggregate of anisotropic cubic crystals. *J. Appl. Mech., Trans. ASME* **1954**, *21* (3), 236–240.
26. Brown, W.F. Solid mixture permittivities. *J. Chem. Phys.* **1955**, *23* (8), 1514–1517.
27. Eshelby, J.D. The determination of the elastic field of an ellipsoidal inclusion, and related problems. *Proc. R. Soc. Lond., A Math. Phys. Sci.* **1957**, *241* (1226), 376–396.
28. Kerner, E.H. The electrical conductivity of composite media. *Proc. Phys. Soc. Lond., B* **1956**, *69* (8), 802–807.
29. Van Der Pol, C. On the rheology of concentrated dispersions. *Rheol. Acta* **1958**, *1*, 198.
30. Meredith, R.E.; Tobias, C.W. Resistance to potential flow through a cubical array of spheres. *J. Appl. Phys.* **1960**, *31* (7), 1270–1273.
31. Hashin, Z. The elastic moduli of heterogeneous materials. *J. Appl. Mech., Trans. ASME* **1962**, *29* (3), 143–150.
32. Hashin, Z.; Shtrikman, S. A variational approach to the theory of the elastic behaviour of polycrystals. *J. Mech. Phys. Solids* **1962**, *10* (4), 343–352.
33. Keller, J.B. Theorem on conductivity of composite medium. *J. Math. Phys.* **1964**, *5* (4), 548.
34. Hashin, Z.; Rosen, B.W. The elastic moduli of fiber-reinforced materials. *J. Appl. Mech., Trans. ASME* **1964**, *31*, 233.
35. Hill, R. Theory of mechanical properties of fibre-strengthened materials. 1. Elastic behaviour. *J. Mech. Phys. Solids* **1964**, *12* (4), 199–212.
36. Hill, R. A self-consistent mechanics of composite materials. *J. Mech. Phys. Solids* **1965**, *13* (4), 213.
37. Budianski, B. On elastic moduli of some heterogeneous materials. *J. Mech. Phys. Solids* **1965**, *13* (4), 223.
38. Hashin, Z. Viscoelastic fiber reinforced materials. *AIAA J.* **1966**, *4* (8), 1411.
39. Russel, W.B.; Acrivos, A. Effective moduli of composite-materials—Slender rigid inclusions at dilute concentrations. *Zeit. Angew. Math. Phys.* **1972**, *23* (3), 434–463.
40. Russel, W.B. Effective moduli of composite-materials—Effect of fiber length and geometry at dilute concentrations. *Zeit. Angew. Math. Phys.* **1973**, *24* (4), 581–600.
41. Mendelson, K.S. Theorem on effective conductivity of a two-dimensional heterogeneous medium. *J. Appl. Phys.* **1975**, *46* (11), 4740–4741.
42. Doyle, W.T. Clausius-mossotti problem for cubic arrays of spheres. *J. Appl. Phys.* **1978**, *49* (2), 795–797.
43. Mcphedran, R.C.; McKenzie, D.R. Conductivity of lattices of spheres. 1. Simple cubic lattice. *Proc. R. Soc. Lond., A Math. Phys. Eng. Sci.* **1978**, *359* (1696), 45–63.
44. Landauer, R. Stability in dissipative steady-state. *Phys. Today* **1978**, *31* (11), 23–29.
45. Perrins, W.T.; McKenzie, D.R.; Mcphedran, R.C. Transport-properties of regular arrays of cylinders. *Proc. R. Soc. Lond., A Math. Phys. Eng. Sci.* **1979**, *369* (1737), 207–225.
46. Christensen, R.M.; Lo, K.H. Solutions for effective shear properties in 3 phase sphere and cylinder models. *J. Mech. Phys. Solids* **1979**, *27* (4), 315–330.
47. McLachlan, D.S. Measurement and analysis of a model dual-conductivity medium using a generalized effective-medium theory. *J. Phys. C. Solid State Phys.* **1988**, *21* (8), 1521–1532.
48. Bao, K.D.; Axell, J.; Grimvall, G. Electrical-conduction in checkerboard geometries. *Phys. Rev., B* **1990**, *41* (7), 4330–4333.
49. Guedes, J.M.; Kikuchi, N. Preprocessing and post-processing for materials based on the homogenization method with adaptive finite-element methods. *Comput. Methods Appl. Mech. Eng.* **1990**, *83* (2), 143–198.
50. Gu, G.Q.; Liu, Z.R. Thermal-conductivity of periodic composite media with spherical inclusions. *Commun. Theor. Phys.* **1991**, *15* (2), 141–148.
51. Gu, G.Q. Calculation methods for effective constants of periodic composite media. *J. Phys., D. Appl. Phys.* **1993**, *26* (9), 1371–1377.
52. Christensen, R.M. Effective properties of composite-materials containing voids. *Proc. R. Soc. Lond., A Math. Phys. Eng. Sci.* **1993**, *440* (1909), 461–473.
53. Lu, S.Y. The effective thermal-conductivities of composites with 2-d arrays of circular and square cylinders. *J. Compos. Mater.* **1995**, *29* (4), 483–506.
54. Batchelor, G.K. Transport properties of 2-phase materials with random structure. *Annu. Rev. Fluid Mech.* **1974**, *6*, 227–255.
55. Whitney, J.M.; McCullough, R.L. Micro-Models for Predicting Viscoelastic Behavior. In *Delaware Composites Design Encyclopedia*; Whitney, J.M., McCullough, R.L., Eds.; Technomic Pub. Co.: Lancaster, PA, USA, 1990; Vol. 2, 189–201.
56. Sastry, A.M. 2.17 Impregnation and Consolidation Phenomena. In *Comprehensive Composite Materials*; Kelly, A., Zweben, C.H., Eds.; Elsevier: Amsterdam, 2000; Vol. II, 609–622.
57. Fung, Y.C. *Introduction to Bioengineering*; World Scientific: Singapore, 2001.
58. Torquato, S. *Random Heterogeneous Materials: Microstructure and Macroscopic Properties*; Springer: New York, 2002.
59. Tsai, S.W.; Wu, E.M. General theory of strength for anisotropic materials. *J. Compos. Mater.* **1971**, *5*, 58. (JAN).
60. Tsai, S.W.; Hahn, H.T. *Introduction to Composite Materials*; Technomic Pub. Co.: Westport, CT, 1980.
61. Sastry, A.M.; Phoenix, S.L. Load redistribution near nonaligned fiber breaks in a 2-dimensional unidirectional composite using break-influence superposition. *J. Mater. Sci. Lett.* **1993**, *12* (20), 1596–1599.
62. Sastry, A.M.; Phoenix, S.L.; Schwartz, P. Analysis of interfacial failure in a composite microbundle pull-out experiment. *Compos. Sci. Technol.* **1993**, *48* (1–4), 237–251.
63. Sastry, A.M.; Phoenix, S.L. Shielding and magnification of loads in elastic, unidirectional composites. *SAMPE J.* **1994**, *30* (4), 61–67.

64. Beyerlein, I.J.; Phoenix, S.L.; Sastry, A.M. Comparison of shear-lag theory and continuum fracture mechanics for modeling fiber and matrix stresses in an elastic cracked composite lamina. *Int. J. Solids Struct.* **1996**, *33* (18), 2543–2574.
65. Phoenix, S.L. Statistical Strength Theory for Fibrous Composite Materials. In *Comprehensive Composite Materials*; Kelly, A., Zweben, C.H., Eds.; Elsevier: Amsterdam, 2000; Vol. I, 559.
66. Reddy, J.N. *Mechanics of Laminated Composite Plates: Theory and Analysis*; CRC Press: Boca Raton, 1997.
67. Hyer, M.W. *Stress Analysis of Fiber-Reinforced Materials*; WCB McGraw-Hill: Boston, 1998. Especially Chapter 3.
68. Reuss, A. Berechnung der fließgrenze kristallen von mischkristallen auf grund der plastizitätsbedingung für einkristalle. *Math. Mech. Z. Angew.* **1929**, *9*, 49.
69. Voigt, W. *Lehrbuch der Kristallphysik*; B.G. Teubner: Leipzig, 1910.
70. Boucher, S. On the effective moduli of isotropic two-phase elastic composites. *J. Compos. Mater.* **1974**, *8* (82).
71. Hershey, A.V. The plasticity of an isotropic aggregate of anisotropic face-centered cubic crystals. *J. Appl. Mech.*, *Trans. ASME* **1954**, *21* (3), 241–249.
72. Kröner, E. Berechnung der elastischen konstanten des vielkristalls aus den konstanten des einkristalls. *Zeit. Phys.* **1958**, *151* (4), 504–518.
73. Hashin, Z.; Shtrikman, S. A variational approach to theory of effective magnetic permeability of multiphase materials. *J. Appl. Phys.* **1962**, *33* (10), 3125.
74. Hashin, Z.; Shtrikman, S. On some variational principles in anisotropic and nonhomogeneous elasticity. *J. Mech. Phys. Solids* **1962**, *10* (4), 335–342.
75. Hashin, Z.; Shtrikman, S. A variational approach to the theory of the elastic behaviour of multiphase materials. *J. Mech. Phys. Solids* **1963**, *11* (2), 127–140.
76. McCullough, R.L. Micro-Models for Composite Materials. In *Delaware Composites Design Encyclopedia*; Whitney, J.M., McCullough, R.L., Eds.; Technomic Pub. Co.: Lancaster, PA, USA, 1990; 2, 49–142.
77. <http://www.matweb.com> (2003).
78. Kriz, R.D.; Stinchcomb, W.W. Elastic-moduli of transversely isotropic graphite fibers and their composites. *Exp. Mech.* **1979**, *19* (2), 41–49.
79. <http://structures.ucsd.edu> (2003).
80. Bergman, D.J. Analytical properties of the complex effective dielectric constant of a composite medium with applications to the derivation of rigorous bounds and to percolation problems.
81. Lurie, K.A.; Cherkhev, A.V. Exact estimates of conductivity of composites formed by 2 isotropically conducting media taken in prescribed proportion. *Proc. R. Soc. Edinb., Sect. A, Math.* **1984**, *99*, 71–87.
82. Milton, G.W.; Kohn, R.V. Variational bounds on the effective moduli of anisotropic composites. *J. Mech. Phys. Solids* **1988**, *36* (6), 597–629.
83. Gibiansky, L.V.; Torquato, S. Link between the conductivity and elastic-moduli of composite-materials. *Phys. Rev. Lett.* **1993**, *71* (18), 2927–2930.
84. Torquato, S.; Gibiansky, L.V.; Silva, M.J.; Gibson, L.J. Effective mechanical and transport properties of cellular solids. *Int. J. Mech. Sci.* **1998**, *40* (1), 71–82.
85. Lu, W.T.; Carlsson, L.A.; Andersson, Y. Micro model of paper. 1. Bounds on elastic properties. *Tappi J.* **1995**, *78* (12), 155–164.
86. Lu, W.T.; Carlsson, L.A. Micro-model of paper. 2. Statistical analysis of the paper structure. *Tappi J.* **1996**, *79* (1), 203–210.
87. Lu, W.T.; Carlsson, L.A. Influence of viscoelastic behavior on curl of paper. *Mech. Time-Depend. Mater.* **2001**, *5* (1), 79–100.
88. Ostojca-Starzewski, M.; Sheng, P.Y.; Jasiuk, I. Damage patterns and constitutive response of random matrix-inclusion composites. *Eng. Fract. Mech.* **1997**, *58* (5–6), 581–606.
89. Cheng, X.; Sastry, A.M. On transport in stochastic, heterogeneous fibrous domains. *Mech. Mater.* **1999**, *31* (12), 765–786.
90. Wang, C.W.; Berhan, L.; Sastry, A.M. Structure, mechanics and failure of stochastic fibrous networks: Part I—Microscale considerations. *J. Eng. Mater. Technol. Trans. ASME* **2000**, *122* (4), 450–459.
91. Wang, C.W.; Sastry, A.M. Structure, mechanics and failure of stochastic fibrous networks: Part II—Network simulations and application. *J. Eng. Mater. Technol. Trans. ASME* **2000**, *122* (4), 460–468.
92. Berhan, L.; Yi, Y.B.; Sastry, A.M. Effect of nanorope waviness on the effective moduli of nanotube sheets. *J. Appl. Phys.* **2004**, *95* (8), 4335–4345.
93. Berhan, L.; Yi, Y.B.; Sastry, A.M.; Munoz, E.; Selvidege, M.; Baughman, R. Mechanical properties of nanotube sheets: Alteration in joint morphology, and achievable moduli in manufacturable materials. *J. Appl. Phys.* **2004**, *95* (9), 5027–5034.
94. Yi, Y.B.; Berhan, L.; Sastry, A.M. Statistical geometry of random fibrous networks, revisited: Waviness, dimensionality and percolation. *J. Appl. Phys.* **2004**. Accepted.
95. Wang, C.W.; Cook, K.A.; Sastry, A.M. Conduction in multiphase particulate/fibrous networks simulations and experiments on li-ion anodes. *J. Electrochem. Soc.* **2003**, *150* (3), A385–A397.
96. Sastry, A.M.; Cheng, X.; Wang, C.W. Mechanics of stochastic fibrous networks. *J. Thermoplast. Compos. Mater.* **1998**, *11* (3), 288–296.
97. Cheng, X.; Wang, C.W.; Sastry, A.M.; Choi, S.B. Investigation of failure processes in porous battery substrates: Part II—Simulation results and comparisons. *J. Eng. Mater. Technol. Trans. ASME* **1999**, *121* (4), 514–523.
98. Wang, C.W.; Cheng, X.; Sastry, A.M.; Choi, S.B. Investigation of failure processes in porous battery substrates: Part I—Experimental findings. *J. Eng. Mater. Technol. Trans. ASME* **1999**, *121* (4), 503–513.
99. Nielaba, P.; Privman, V. Multilayer adsorption with increasing layer coverage. *Phys. Rev., A* **1992**, *45* (8), 6099–6102.



100. Deckmyn, P.; Davies, G.A.; Bell, D.J. Properties of percolating clusters on finite lattices applied to model filtration processes. *Appl. Math. Model.* **1995**, *19* (5), 258–269.
101. Datta, S.; Redner, S. Gradient and percolative clogging in depth filtration. *Int. J. Mod. Phys. C* **1998**, *9* (8), 1535–1543.
102. Cheng, X.; Sastry, A.M.; Layton, B.E. Transport in stochastic fibrous networks. *J. Eng. Mater. Technol. Trans. ASME* **2001**, *123* (1), 12–19.
103. Mohanty, S.; Sharma, M.M. A Monte-Carlo RSRG method for the percolation conduction properties of correlated lattices. *Phys. Lett., A* **1991**, *154* (9), 475–481.
104. Kirkpatrick, S. Percolation and conduction. *Rev. Mod. Phys.* **1973**, *45* (4), 574–588.
105. Pike, G.E.; Seager, C.H. Percolation and conductivity—Computer study. 1. *Phys. Rev., B* **1974**, *10* (4), 1421–1434.
106. Yi, Y.B.; Sastry, A.M. Analytical approximation of the two-dimensional percolation threshold for fields of overlapping ellipses. *Phys. Rev., E* **2002**, *66* (6). Art. no.-066130.
107. Yi, Y.B.; Sastry, A.M. Analytical approximation of the percolation threshold for overlapping ellipsoids of revolution. *Proc. R. Soc. Lond., A* **2004**, accepted.
108. Haan, S.W.; Zwanzig, R. Series expansions in a continuum percolation problem. *J. Phys., A, Math. Gen.* **1977**, *10* (9), 1547–1555.
109. Coniglio, A.; Deangelis, U.; Forlani, A.; Lauro, G. Distribution of physical clusters. *J. Phys., A, Math. Gen.* **1977**, *10* (2), 219–228.
110. Coniglio, A.; Deangelis, U.; Forlani, A. Pair connectedness and cluster size. *J. Phys., A, Math. Gen.* **1977**, *10* (7), 1123–1139.
111. Hill, T.L. Molecular clusters in imperfect gases. *J. Chem. Phys.* **1955**, *23* (4), 617–622.
112. Drory, A. Theory of continuum percolation. I. General formalism. *Phys. Rev., E* **1996**, *54* (6), 5992–6002.
113. Quintanilla, J.; Torquato, S. Clustering properties of d-dimensional overlapping spheres. *Phys. Rev., E* **1996**, *54* (5), 5331–5339.
114. Quintanilla, J.; Torquato, S. Clustering in a continuum percolation model. *Adv. Appl. Probab.* **1997**, *29* (2), 327–336.

Compression of Digital Biomedical Signals

Tapio Grönfors

Department of Computer Science, University of Kuopio, Finland

Martti Juhola

Department of Computer Sciences, University of Tampere, Finland



INTRODUCTION

This article introduces compression methods for biosignals routinely used in clinical work, electrocardiography (ECG), electromyography (EMG), electrooculography (EOG), and electroencephalography (EEG), including brain auditory evoked potential (BAEP). The compression of ECG signals is the most studied biosignal compression approach, with hundreds of published scientific papers. The compression of BAEP signals is far less studied, with only a few published papers. This article discusses some less studied biosignals, such as EMG and EOG, and determines fundamental quality questions regarding biosignal compression methods.

The article begins with a brief background introduction to compression methods and quality evaluation. Selected, cited application examples are presented after the Introduction. In every case, remarks are made from the most typical methods and most prominent results. However, the aspect is kept on an abstract level, without technical details of the methods. A concise conclusion is drawn about these methods and their applicability.

MOTIVATION FOR BIOSIGNAL COMPRESSION

Several electrophysiological signals of humans are used routinely in clinical work. Many of these signals fulfill one or more of the following conditions: high sampling rate (≥ 1 kS/sec), high precision (12 bits or more), multichannel recording (four channels or more), or long-term recording (≥ 24 hr). Based on these figures, hundreds of megabytes of data can be easily evoked per recording.

Although large portions of these recordings are disposable, growing portions are archived for scientific purposes and clinical use in electronic patient documents. Notable numbers of recordings change positions in information networks. Novel biosignal recording technologies are made to utilize wireless measuring processes.^[1] Typically, these portable, wearable, or even implantable recording devices use a radio system with limited transmission capabilities.

Both archiving and telecommunication benefit from data compression. However, demands of the task and available approaches are quite specific. Requirements exist to work with novel biosignal recording technologies for low-power computing and real-timing in tandem with computationally reduced hardware.

ASPECT FOR BIOSIGNAL COMPRESSION: LOSSLESS, NEAR LOSSLESS, AND LOSSY

Lossless compression methods are the footing for compression. As the name suggests, these methods are totally reversible and there is no information loss in the compression process. These methods are based on the results of information theory^[2] and compression is attained by setting the length of codes according to their probability. For compressing a text or computer program, the lossless property is an obvious requirement. Lossless methods are also widely used for biosignals for fear of losing critical information.

The main methods of lossless compression are^[2] Huffman coding and its variations, such as Golomb coding, arithmetic coding, and dictionary techniques, including well-known LZ77/LZ78 approaches and predictive coding. Arithmetic coding is computationally more complex than Huffman coding, but can code sequences of data and effect better performance and adaptability on changing data statistics. Dictionary techniques are most useful when frequently occurring structural patterns in data are a small subset of all possible structural patterns, like in text. However, this is not typical with biosignals. Predictive methods, which use past values of the data sequence to predict the actual value of the current data, are endogenously suitable for signal compression.

In the group of near-lossless compression methods, the main requirement is ensuring that the maximum error between the original and the processed data does not exceed a fixed threshold. This kind of a bound on the error magnitude can apply to biosignals on a sample-by-sample basis.

In the group of lossy compression methods, there is an unlimited error between the original and the

processed data. Typically, the error increases while a better compression ratio is attained. The lossy methods are based on models. Via the model, worthless, or at least non-critical, information is lost in the compression process. This approach is acceptable for recorded biosignals, because they contain extra data beside the information.

A variety of lossy compression approaches are available. The main methods are^[2] scalar and vector quantization with multiple variations on the theme, differential encoding, transform methods, subband coding, and wavelet-based methods. Differential encoding does not provide as high compression as quantization methods, but it is very simple and suited to implementation for signal-like data.

In a practical biosignal compression system, few consecutive compression stages can follow each other. In every stage, the results of the preceding stage are compressed with another method, gaining more compression. Some methods are lossless and some are lossy. Compression methods can also be dynamic, adapting to local variations of data.

EVALUATION OF EFFICIENCY OF COMPRESSION METHODS

The most obvious measurement for a compression method is the attained ratio between the original and the compressed data. This is simply the ratio between the original size and the compressed size, expressed, for example, in the form 5:1, meaning that the original data was five times as large as the compressed data. The compression ratio can also be represented by a reduction in the amount of data required as a percentage of the size of the original data, for example, 80%. From a telecommunications aspect, the compression ratio can be referred to as the rate, or the average number of bits per sample in the compressed representation.

Obtaining the desired compression ratio requires computational effort. The algorithmic complexity is the second measure of a compression method. Algorithmic complexity can be represented as the relative structural complexity of an algorithm, required amount of memory, required sophistication of a processor, or fastness of an implementation with a given processor. In novel challenging applications the algorithmic complexity of compression methods is emphasized.

With lossy compression methods, the third measure is how closely the reconstruction resembles the original data. This is the compression error. With lossless methods, this is omitted. With lossy methods, the measure is computed by investigating the influence of compression. First, the signal data is compressed and then decompressed. The decompressed signal data is compared with the original, uncompressed data. The

natural objective is that their difference be as small as possible. On the other hand, the more effective the compression, the larger such a difference tends to be. The compression error is often called the distortion.

The compression errors of biosignals are typically evaluated using the measures^[2] mean squared error (MSE), average of absolute differences, or mean absolute error (MAE), maximum value of the error magnitude (MEM), signal-to-noise ratio [SNR, including its logarithmic representation in decibels, SNR(dB)], peak signal-to-noise ratio (PSNR), and normalized root-mean-square error (NMRS). In biosignal-related studies, the NMRS value is represented as a percentage and is called the percent root-mean-square difference (PRD) for original signal samples o_i and processed samples p_i and their variances, σ_o^2 and σ_p^2 :

$$MSE = \frac{1}{N} \sum_{n=1}^N (o_n - p_n)^2 \quad (1)$$

$$MAE = \frac{1}{N} \sum_{n=1}^N |o_n - p_n| \quad (2)$$

$$MEM = \max_n |o_n - p_n| \quad (3)$$

$$SNR = \frac{\sigma_o^2}{\sigma_p^2} \quad (4)$$

$$SNR(dB) = 10 \log_{10} \frac{\sigma_o^2}{\sigma_p^2} \quad (5)$$

$$NMRS = \sqrt{\frac{\sum_{n=1}^N (o_n - p_n)^2}{\sum_{n=1}^N o_n^2}} \quad (6)$$

All of the above-mentioned measures are used to summarize the error of the sequence of signal samples. Sometimes the differences are also examined on a sample-by-sample basis tracking the values of squared or absolute difference measures. Nevertheless, these measures only express the general changes of an entire signal affected by compression. Thus, they cannot directly reveal the risk of the loss of useful clinical information. Such a risk has to be separately clarified for each medical parameter, as we will describe in the later sections.

FIDELITY OF BIOSIGNALS AND CLINICAL QUALITY

Audio and video compressions are well-known areas of lossy methods for signal-like data. The fidelity and the quality of processed signals are evaluated via human

perception of distortion: if they sound or look agreeable, compression is valid. When the compression is optimized for perception, the data are inaccurate for computational processing.

Visual similarity is sometimes used as a criterion for biosignals. This can be a sufficient condition only if the clinical use is based on visual inspection. A fundamental criterion of fidelity is that the processed biosignal is not allowed to lead a physician to a wrong or insufficient diagnosis.

Most biosignals are interpreted via medical parameters and the time evolution of medical parameters. The medical parameters are typically calculated from a time-domain or frequency-domain representation of biosignals. Although the measured compression errors between the original and processed biosignals correlate with errors of medical parameters, the dependency is not straightforward. A novel and fruitful aspect involves inspecting the medical parameter values of processed signals against original signals over the whole clinically interesting range.

APPLICATION EXAMPLES OF SELECTED BIOSIGNALS

Compression of ECG Signals

Perhaps the most important, or at least most used, biomedical signal is ECG, which is applied to the diagnosis of heart problems and even leisure-time heart pulse recorders. Numerous articles investigated compression of ECG signals since the 1960s. Presumably, the majority of all biosignal compression investigations consider ECG signals. The investigators concentrated on lossy compression, or ECG signals, which are typically explored in time domain. An ECG signal is usually composed of rather clear waves, segments, and QRS complexes, which create a successful basis for effective lossy compression. It suffices to keep the main features of the ECG signal components almost intact or only slightly deformed by compression. Therefore, there has not been much reasonable cause to study lossless compression for ECG signals. Several different methods were introduced in the past. The methods can be divided into two main categories: direct and transformation techniques. Another method, called the parametric method, is performed on the basis of statistical models applying signal distribution. The latter category is seldom applied to ECG signals.

Direct compression methods were introduced first because several such methods are fairly straightforward and, therefore, fast. In fact, investigations mostly focused on compression effectiveness, but their computational effectiveness, which materializes as running time, was only briefly considered. The quality of compression

was usually understood to include (normalized) MSE, PRD, or the measures mentioned above. Instead, investigators rarely dealt with the issue of how a compression method can exactly deform ECG signal components in detail. A statistical value like PRD might look fairly good, but it does not express much, e.g., how amplitudes of QRS complexes could have been decreased by compression. Later research mainly addressed transformation methods. Transformation methods seem promising in order to compress efficiently and, at the same time, handle information content of signals conservatively, i.e., not deform ECG signal components too strongly.

Several direct compression methods are applied to ECG signals. Typically, they are designed on the basis of signal segmentation in a suitable way to change a signal curve into broken lines of best fit corner data-points and to represent the curve as a sequence of such segments. Segments can also be accomplished with nonlinear short, e.g., third-degree, curves. However, short linear segments are more typically applied because their consideration is simpler, faster, and possibly equally useful in the sense of compression.

Perhaps one of the oldest direct compression methods is differential pulse code modulation, which utilizes a polynomial predictor like

$$\hat{u}_i = u_{i-1} + \Delta_1 u_{i-1} + \Delta_2 u_{i-2} + \dots + \Delta_k u_{i-k},$$

where \hat{u}_i is the i th predicted sample, u_{i-1} is the $(i-1)$ st sample, $\Delta_1 u_{i-1}$ is equal to $u_{i-1} - u_{i-2}$, and generally $\Delta_k u_{i-k}$ is equal to $\Delta_{k-1} u_{i-k} - \Delta_{k-1} u_{i-k-1}$. Usually either the zero-order predictor of

$$\hat{u}_i = u_{i-1}$$

or the first-order predictor,

$$\hat{u}_i = u_{i-1} + (u_{i-1} - u_{i-2}) = 2u_{i-1} - u_{i-2},$$

is used. For example, for the former the first-difference signal or amplitude between successive samples, $e_i = u_i - \hat{u}_i$, is replaced for the original signal. Because a magnitude of e_i is clearly smaller than that of u_i , it spares storage counted in bits. To decode the compressed data, the reverse procedure is performed. Several methods are derived from this approach,^[3] where typically the prediction procedure is developed further. The basic differential pulse code modulation (DPCM) is naturally suitable for any type of signal. It is also used as a preprocessing method and then built for some other compression, which obtains successive differences e_i as input. In other words, first DPCM is performed as preprocessing for a signal and then another compression method is executed for the output of DPCM.



Other simple direct methods are often based on various segmentation procedures of ECG signals. Fan and scan-along polygonal approximation (SAPA) are algorithms^[3,4] in which linear slopes are fitted on a signal. In Fan the longest possible line was set between the current start sample and a new end sample so that all intermediate samples were within a given small error limit. In SAPA, this technique was developed slightly further. Another idea was to take advantage of peaks, maxima, and minima in ECG signals. By recognizing maxima and minima, slope changes, and zero-crossings it was possible to approximate ECG signals.^[3] These simple methods can be fast, but can lose more information from ECG signals than more sophisticated segmentation algorithms. Information loss is understood to mean inaccurate decoding with considerable sample (amplitude) differences between the original and the decoded (returned or decompressed) signal. There may also be differences along the time axis after decompression. These features are even visually considerable. In principle, they could slightly affect frequency domain properties, but this has been little explored in compression literature. After the aforementioned early ECG compression methods, more sophisticated segmentation algorithms were introduced. A simple idea starts from the computation of linear regression slopes, which are successively set on the signal to approximate this signal. However, this way is not very effective. The regression procedure implies an optimization task along with a vertical axis, which may cause extra amplitude errors in segmentation, because the shortest distance between a line and a sample is not along the vertical axis, but is equal to the Euclidean distance between them. Such a distance to be minimized is obtained by assigning a normal (perpendicular segment of a line) between the sample and the line. This is computationally harder than linear regression. We constructed such a method, which also used error distributions to predict for the minimization purpose.^[5]

Transformation-based compression methods cover a large area, which can loosely be described at the highest level as a signal that is transformed into another form, from which its information content (frequencies, waveforms, peaks, and phenomena in general) can be extracted and effectively returned via an inverse transformation or corresponding procedure to almost equivalent representation compared with the original signal. Compared with direct methods, transformations are more complicated and probably more time consuming, but capable of yielding high compression ratios with tolerance to the decay of information content. In short, a more complicated and sophisticated method may perform better than a simple method.

Several available transformations are applicable to ECG and other biosignal compression, such as discrete fast Fourier, cosine, Walsh, Haar, and Karhunen–Loeve

(principal components) transforms. The Karhunen–Loeve transform is optimal in a theoretical sense, but requires heavy computation. It results in decorrelated transform coefficients, minimizing total entropy. The computationally hard task of its eigenvalue process supports the use of other theoretically suboptimal techniques, in which the basis vectors of a transform are other than the eigenvectors (derived from eigenvalues) of the covariance matrix. Instead, they are based on simple sines and cosines (Fourier and cosine transforms), squares, or other waves or sequences.

Although articles on ECG compression are numerous, we mainly restrict ourselves to a small collection of recent publications. A comprehensive review of the ECG compression of “traditional” methods is provided by Jalaeddine et al.^[3] Chen et al. introduced a method on the basis of pattern recognition using heartbeat segmentation and normalization, two-stage pattern matching with template updating, and residual beat coding as Huffman or run-length coding or vector quantization.^[6] A Max-Lloyd quantizer was developed for the ECG compression of a telemetry system for dolphins and human divers.^[7] Because of the application context, i.e., allowing low computational complexity, the traditional quantizer was employed to code differences of current and predicted samples given by a first-order differential pulse modulation. Ambulatory ECG compression was studied with average beat subtraction and residual differencing.^[8] Fast Fourier transform was adopted by Kulkarni et al.^[9] Vector quantization was used by Mammen and Ramamurthi^[10] and Cárdenas-Barrera and Lorenzo-Ginori.^[11] Miaou and Larn discussed adaptive vector quantization of ECG.^[12] Discrete cosine transform was applied to model heartbeats.^[13] Sun and Tai explored vector quantization for residuals after QRS detection and segmentation of heartbeats.^[14]

Wavelet transforms have been at the forefront of current research. Alesanco segmented an ECG signal into beats, subtracted a template from them, and applied wavelet compression tuned with thresholds as usual.^[15] A filter bank-based algorithm was introduced,^[16] in which a cosine-modulated filter bank was compared with a wavelet packet technique. Tohumoglu and Sezgin utilized embedded zero-tree wavelets.^[17] Wavelet-based vector quantization was considered by Miaou and Chao.^[18] Wavelet-based compression for the purpose of telecardiology was presented.^[19] Signal-dependent wavelet compression was employed to adapt wavelets to a signal according to signal distortion rate determined by the desired compression ratio.^[20] This method was applied to EMG and EEG as well. We tested and compared discrete cosine transform with enumerative coding, a Daubechies-6 wavelet with γ run-length coding or enumerative coding, and a Daubechies-6 wavelet packet with γ run-length coding

or enumerative coding.^[21] The four wavelet compressions were virtually equally good and slightly but clearly better than the discrete cosine transform. For ECG signals recorded at the sampling frequency of 400 Hz and amplitude resolution of 13 bits, average PRD values were approximately 0.5% for the wavelets and 1% for the cosine transform when the compression ratio was 6.5:1. For a compression ratio equal to 13:1, average PRD values were, respectively, about 1.5% and 5.5%.

Neural networks were used for prediction in ECG compression.^[22] A genetic segmentation algorithm for compression was presented.^[23] Two-dimensional coding techniques for ECG signals were investigated by Alexandre et al.^[24] and Lee and Buckley.^[25] A Markov process was used to consider the coding redundancy and intersample redundancy, which yielded better results than Huffman coding.^[26]

It is not straightforward to liken the efficacy of the compression methods considered, because their data materials vary greatly. Most studies used their own signals, but some utilized the MIT-BIH arrhythmia signal database (Massachusetts Institute of Technology, 1992). Compression effectiveness clearly depends on recording parameters. The parameters are sampling frequency, amplitude resolution, and filtering mode, which are usually 100–1000 Hz, 8–16 bits, and lowpass filtering with a cutoff frequency of 30–100 Hz. There are numerous ways to design suitable filters, which may somewhat affect compression results. The use of a high sampling frequency may naturally cause better compression results than a low sampling frequency. This is assumed to be a roughly direct linear relation: by doubling the sampling frequency, the compression ratio can approximately be doubled if everything else is similarly performed in the compression process. Amplitude resolution may follow this property. However, these ideas are simplified approaches, because in reality there is always more or less noise present in signals; even after lowpass filtering, frequency bands of actual information and noise overlap. Increased sampling frequencies and amplitude resolutions also manifest more noise in addition to actual signal phenomena. We must remember that noise is mostly random, with the exception of some systematic types of noise, such as that caused by surrounding electric devices using a main current of 50 or 60 Hz.

The lossy compression literature regarding ECG signals reports compression ratios between 2:1 and 12:1 depending on the methods and types of ECG signals. Lossless compression has seldom been studied with regard to ECG signals. A compression ratio of 2:1 was obtained for ECG signals at a 400-Hz sampling frequency and 13-bit amplitude resolution.^[27] Thus, this value can be taken as the baseline ratio for lossy compression. Compression quality was then evaluated

according to the measures, e.g., PRD, listed at the beginning of the chapter. Nevertheless, these techniques only compare decompressed and original signals in a general sense, because they deal with parallel samples. Such a general view does not express precisely what may happen to amplitudes of QRS peaks or waves in ECG signals. This conception is strongly supported by Blanco-Velasco et al.,^[28] who pointed out that relying merely on PRD or corresponding measures may result in wrong conclusions.

The quality of ECG signal compression statistical measures other than those given above has been little investigated. The statistical measures provide an overall influence of compression on ECG signals, but more precise scrutiny is also affordable. A subjective evaluation of cardiologists was used by Hilton^[29] for ECG signals of the MIT-BIH database. A diagnostic distortion measure was prepared and tested by Zigel et al.^[30] to assess the impact of compression on several medically interesting ECG attributes such as magnitudes of QRS complex and T and P waves. A wavelet-based measure was suggested for the quality assessment of ECG compression.^[31] Kulkarni et al. clarified how many fast Fourier transform coefficients were necessary for the reliable compression validated with differences of ECG attributes like QRS peaks between original and decompressed signals.^[9] Likewise, we explored ECG attributes regarding amplitudes and durations of QRS peaks and that of T waves. For our four wavelet compression modes named above,^[21] we determined that, for instance, for a compression ratio of 13:1, the two T-wave attributes received tolerable average errors of 5–7% evaluated as PRD values. Correspondingly, QRS amplitude errors were 2% and 3% on average, but QRS duration errors were as large as 10–12%. To diminish the latter to 5%, the compression ratio should be reduced to about 4:1. Taking these in-depth evaluations into account, better compression ratios than 4:1 to 8:1 cannot be recommended for ECG signal compression, depending on the compression methods and other factors described.

Compression of EMG

The recording of muscle activity is widely used in clinical neurophysiology, sports medicine and rehabilitation, sleep studies, and even psychophysiology to measure emotions. EMG is best suited to examining the way in which tension develops within a muscle. EMG can reveal onset and offset of activities too small to produce visible movements. Via EMG, muscular force and fatigue can be assessed. EMG recording can be made intramuscularly with needle electrodes or on the skin with surface electrodes. Non-invasive surface myoelectric signals are relatively simple to record.^[32]



EMG recordings are noisy-like signals generated by the complex non-linear system of skeletal muscles. EMG, like regular patterns, is not typical. The amplitude of the surface EMG signal is typically a few hundred microvolts and most of the activity is between 50 and 150 Hz. Hence, 1000 samples/sec is a suitable sampling frequency. Depending on the application area of the EMG signal, different medical parameters can be computed^[33]: spectral parameters such as the mean frequency (MNF) and the median frequency (MDF) of power spectra and amplitude parameters such as the averaged rectified value (ARV) and the root mean square value (RMS).

Different lossy compression methods were widely tested^[34] and results evaluated via compression ratio versus SNR(dB). Differential pulse code modulation, a multipulse coder, a code excited linear predictive coder, discrete cosine transform discarding non-significant components, and wavelet-based subband coding were used. Both needle-electrode and surface-electrode recordings were used for testing. Compression ratios varied up to 13:1. The wavelet transform outperformed other compression methods in the compression ratio range of 5:1 to 9:1.

Wellig et al.^[35] used two modified wavelet-based methods (single-tree and zero-tree wavelet coder) and evaluated results via both compression ratio versus SNR(dB) and percentual distortion PRD versus bitrate. The modified zero-tree wavelet algorithm gives better results than the single-tree wavelet algorithm. Berger et al.^[36] also used wavelet-based methods. Orthogonal wavelet transformations (DWT) with Daubechies wavelets are compared with discrete cosine transformation. In assessments between methods based on compression ratio and energy-based SNR(dB), DWT methods demonstrate better performance than discrete cosine transform methods and Daubechies-4 demonstrates the best performance within DWT methods with a compression ratio of 10:1.

Chan et al. first published a paper in which evaluation was based on the medical parameters MDF and variance.^[37] They used a variation of differential encoding, the adaptive differential pulse code modulation (ADPCM) technique, in a real-time system and achieved compression ratios up to 4:1.

Carotti et al. used a compression method based on the algebraic code excited linear prediction.^[38] The test procedure utilized a fixed 87.3% compression ratio and the quality was assessed in two ways: MSE provided a global indicator of the quality and medical parameters ARV, RMS, MNF, and MDF and spectral skewness calculated for several force levels provided detailed quality indicators. Both simulated and recorded EMG signals were used.

We studied lossy compression of EMG signals^[39,40] with different variations of vector quantization: basic

vector quantization (VQ), mean residual VQ, gain shape VQ, and classified VQ for signal segments. In all our studies the behavior of MNF and MDF parameters played a central role. The fidelity depended on the size of the codebook and segment length. Longer segments presented attenuating lowpass filter effects and shorter segments added segment junction impulse noise. The gain shape method was most advantageous and was able to reach agreeable quality up to a compression ratio of 10:1.

Compression of EOG Recordings

We shed light on biosignal compression by exploring eye movements measured at otoneurological clinics applied to balance investigations. These signals were electro-oculographically recorded with surface electrodes on the canthi of each subject. We recorded many eye movement signals at a sampling frequency of 400 Hz and amplitude resolution of 13 bits. They were filtered with a median-hybrid lowpass filter to suppress noise above about 70 Hz. To study lossless compression of eye movement signals,^[41] the following algorithms were programmed: γ and exponential Golomb universal codes, LZSS, LZW, and LZ77 dictionary methods, Huffman algorithm, arithmetic coding, and adaptive Rice and exponential Golomb coding. All were implemented with the zero-order predictor. For the unfiltered signals the four latter techniques performed best, yielding average compression ratios of 2.2–2.3:1 when MAE and MSE were computed to output reasonably low values. After filtering, the compression ratios of the same signals were 2.6–2.7:1.

We also studied lossy compression methods.^[42] We implemented compression methods on the basis of wavelets, wavelet packets, and discrete cosine transform, similar to that described above.^[21] In addition, Walsh–Hadamard was used. Results of these transform methods were compared with simple pure scalar quantization (levels up to 64) and differential pulse code modulation with the zero-order predictor. For the unfiltered eye movement signals, the latter two simple methods gave low compression ratios of 2.2–3.2:1, with average PRD values no higher than 5%. Using compression ratios as high as 18:1, the wavelet and wavelet packet methods reached average PRD values of less than 4%, the Walsh–Hadamard transform obtained 8%, and the discrete cosine transform produced the best PRD values, at 3%. We continued this research by analyzing precisely how the best methods (discrete cosine transform and wavelet packet) affected eye movement attributes such as amplitudes and durations and maximum velocities and gains derived from the preceding amplitudes and durations.^[43] Fast eye movement types, i.e., saccades and fast phases of nystagmus,

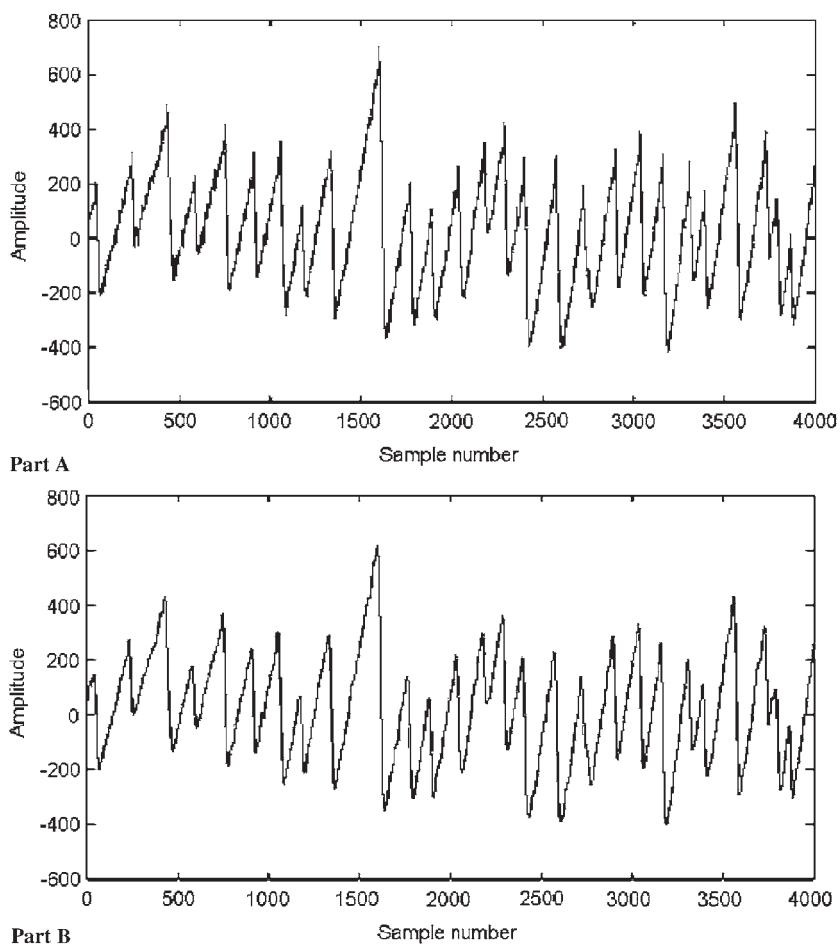


Fig. 1 A 10-sec segment of a fast nystagmus eye movement signal (A) was compressed with the discrete cosine transform using a very high compression ratio of 19:1 and then decompressed (B). Note how, for example, the highest beat at the approximate location of sample 1500 in (B) is lower than that in (A).

were more sensitive to compression than slower eye movement types like tracking eye movements (Fig. 1). In conclusion, we were able to compress and decompress eye movement signals with a compression ratio of 8:1 with reliable analysis results, including average errors of at most 5% for critical attributes of fast eye movements and less for others. Even a compression ratio as high as 18:1 produced average errors below 5% for slow eye movements.

Compression of EEG Including BAEP

EEG is a technique for recording the summation of electrical activity from the scalp as related to cortical activity. It is widely used in neurology, clinical neurophysiology, sleep studies, and even novelty areas like biofeedback and human-computer interfaces. Spontaneous EEG signals are interpreted with rhythms: α activity (about 8–12 Hz), β activity (13 Hz and larger), δ activity (0.5–4 Hz), and θ activity (5–7 Hz), λ waves, K-complex waves, and sleep spindles, which are not defined solely in terms of frequency.^[32]

Counterpoints to spontaneous EEG recordings are evoked responses (ER). ER are time locked to specific

sensory stimulus and response, alternations in cortical or neuronal pathway activity, following that stimulus. BAEP are a sequence of electrical potentials generated in the brainstem and the central auditory pathway after stimulation of the ears with various kinds of sounds. A brief repeated click delivered through earphones is the stimulus most often used. The response can be characterized by a series of peaks within 10 msec after a stimulus onset, e.g., so-called Jewett components (Fig. 2).^[44] BAEP recordings have been used in audio- and otoneurological clinics to evaluate symptoms and dysfunctions related to auditory pathways.

The amplitude of an EEG signal on the scalp is typically only a few microvolts. A common sampling frequency for EEG recordings is between 125 Hz and 200 Hz, although the newest technology allows sampling frequencies up to 1 kHz. A sampling frequency for evoked responses is higher, even 20 kHz for short potentials. Both spontaneous EEG and ER recordings use multichannel setups.

A summary paper^[45] provides a good background to lossless EEG compression techniques. Lossy methods are passed, because clinicians consider exact EEG reconstruction an essential requirement. Widely employed compression algorithms (repetition count and



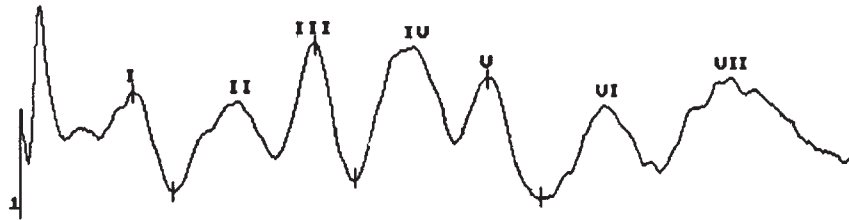


Fig. 2 A typical 10-msec BAEP signal of a healthy subject with Jewett I–VII components

Huffman coding), VQ, transformation compression, and techniques based on signal predictors (linear predictors, maximum likelihood, and neural networks) were evaluated and presented.^[45] Data compression was represented in percentage format, comparing it with two widely used compression programs: *gzip* (38.8%) and *lharc* (38.9%). Huffman coding resulted in the same compression ratio as *gzip* and *lharc*, 39.5%. Also, the authors used Huffman coding on the compression of prediction error of predictor-based techniques and the difference between the original signal and the processed signal in transformation compression was run with discrete cosine transform. The VQ method includes the use of error vectors to ensure losslessness. The best compression ratios achieved were 59.0% with VQ and 59.7% with linear predictive coding.

A substantial number of EEG compression studies use only lossless methods. Magotra et al. used a two-stage technique: a reversible filter or recursive least-squares lattice filter followed by arithmetic coding.^[46] The best results with reversible (high-pass) filters approached a factor of 2. An instantaneous adaptive bit-assignment algorithm for lossless compression of multichannel EEG was used.^[47] This real-time technique approximated a compression ratio of 2:1. Neural network-based techniques were used.^[48] Lossless compression through neural network is achieved using the net of a predictor and coding the prediction error in a lossless manner with Huffman or arithmetic coding. The best results exceeded a compression ratio of 6:1. Lossless wavelet transformation was utilized for multichannel real-time EEG compression.^[49] The system also performs denoising, spike detection, and sorting. The achieved compression ratio could exceed 10:1.

Cinkler et al. and Memon et al. performed a versatile and thorough comparison study on lossless and near-lossless EEG compression methods.^[50,51] The motivation for near-lossless methods is to improve the bit rate and preserve diagnostic accuracy. For lossless compression, methods include a range from simple dictionary-based approaches to more sophisticated context modeling techniques. For near-lossless compression the authors used a context-based trellis-searched DPCM technique that minimizes the entropy of the quantized prediction error sequence. They proved that dictionary-based techniques do not work

well when applied to EEG signals because of the inherent nondeterministic nature.

In a study by Mitra and Sarbadhikari, an interactive function system- and genetic algorithm-based lossy compression method was used.^[52] The fidelity of the compression process was evaluated with SNR and PSNR and by examining the cross-correlation and the power spectra of original and processed signals. The authors concluded the quality was sufficiently good for clinical purposes with 85% compression. Computationally effective ADPCM subband coding was used.^[53] The authors achieved a compression ratio of 70% with very little distortion.

We studied lossy compression of BAEP signals^[42] with scalar quantization (SQ), DPCM combined with SQ, combinations of discrete cosine transformation (DCT) and successive approximation quantization (SAQ) with different bits per sample (bps) rates, and combinations of wavelet transforms and SAQ with different bps rates. The results were evaluated via MAE of latencies (in milliseconds) and amplitudes between original and processed signals for Jewett components I and V and mean NMRS between the original and processed signals. Results demonstrated that both DCT and SAQ with the bps 1 approach and wavelets and SAQ with the bps 1 approach could keep the fidelity at a compression ratio of 12:1.

CONCLUSIONS

If the biosignal is clinically interpreted via well-defined medical parameters, the use of lossy compression method is easier, because the effect of processing is measurable. Confidence in correct diagnosis remains. A good counterexample is EEG signals: they are visually inspected by human experts and automatically analyzed via signal processing algorithms in the time domain.

Provided that signal compression is carefully accomplished, fairly high compression ratios such as 4–8:1, and even up to 10:1, can be applied to biosignals in the case of lossy compression. However, compression ratios depend on the types of signals. For slow changes in signals, i.e., for low frequencies, even greater compression ratios could be used. The most effective

methods to compress biosignals and yet preserve their main features seem to be methods-based wavelet transforms. In the case of lossless compression, compression ratios are essentially smaller, e.g., 2–3:1. Compression ratios also depend on sampling frequencies, amplitude resolutions, and the filters employed. Thus, no direct comparison between reports of various tests is reasonable. A more general measure than the compression ratio is actually the average number of bps used in a compressed signal, but we only used it secondarily here because the great majority of the literature regarding biosignal compression does not apply it.

REFERENCES

1. Schwiebert, L.; Gupta, S.K.S.; Weinmann, J. *Research challenges in wireless networks of biomedical sensors, MobiCom '01*. Proceedings of the 7th Annual International Conference on Mobile Computing and Networking, Rome, Italy; ACM Press: New York, 2001; 151–165.
2. Sayoon, K. *Introduction to Data Compression*; Morgan Kaufmann: New York, 2000.
3. Jalaliddine, S.M.S.; Hutchens, C.G.; Strattan, R.D.; Coberly, W.A. ECG data compression techniques—A unified approach. *IEEE Trans. Biomed. Eng.* **1990**, *37* (4), 329–343.
4. McKee, J.J.; Evans, N.E.; Owens, F.J. Efficient implementation of FAN/SAPA-2 ECG compression methods using fixed point arithmetic. *Automedica* **1994**, *16*, 109–117.
5. Koski, A.; Juhola, M. Segmentation of digital signals based on estimated compression ratio. *IEEE Trans. Biomed. Eng.* **1996**, *43* (9), 928–938.
6. Chen, W.S.; Hsieh, L.; Yuan, S.Y. High performance data compression method with pattern matching for biomedical ECG and arterial pulse waveforms. *Comput. Methods Programs Biomed.* **2004**, *74* (1), 11–27.
7. Rodriguez, M.; Ayala, A.; Rodriguez, S.; Rosa, F.; Diaz-Gonzalez, M. Application of the Max–Lloyd quantizer for ECG compression in diving mammals. *Comput. Methods Programs Biomed.* **2004**, *73* (1), 13–21.
8. Hamilton, P.S.; Tompkins, W.J. Compression of the ambulatory ECG by average beat subtraction and residual differencing. *IEEE Trans. Biomed. Eng.* **1991**, *38* (3), 253–259.
9. Kulkarni, P.K.; Kumar, V.; Verma, H.K. Diagnostic acceptability of FFT-based ECG data compression. *J. Med. Eng. Techn.* **1997**, *21* (5), 185–189.
10. Mammen, C.P.; Ramamurthi, B. Vector quantization for compression of multichannel ECG. *IEEE Trans. Biomed. Eng.* **1990**, *37* (9), 821–825.
11. Cárdenas-Barrera, J.L.; Lorenzo-Ginori, J.V. Mean-shape vector quantizer for ECG signal compression. *IEEE Trans. Biomed. Eng.* **1999**, *46* (1), 62–70.
12. Miaou, S.-G.; Larn, J.-H. Adaptive vector quantization for electrocardiogram signal compression using overlapped and linearly shifted codevectors. *Med. Biol. Eng. Comput.* **2000**, *38*, 547–552.
13. Nunes, J.C.; Nait-Ali, A. ECG compression by modelling the instantaneous module/phase of its DCT. *J. Clin. Monit. Comput.* **2005**, *19* (3), 207–214.
14. Sun, C.C.; Tai, S.C. Beat-based ECG compression using gain-shape vector quantization. *IEEE Trans. Biomed. Eng.* **2005**, *52* (11), 1882–1888.
15. Alesanco, A.; Olmos, S.; Istepanian, R.S.; Garcia, J. Enhanced real-time ECG coder for packetized telecardiology applications. *IEEE Trans. Inf. Technol. Biomed.* **2006**, *10* (2), 229–236.
16. Blanco-Velasco, M.; Cruz-Roldan, F.; Lopez-Ferreras, F.; Bravo-Santos, A.; Matinez-Munoz, D. A low computational complexity algorithm for ECG signal compression. *Med. Eng. Phys.* **2004**, *26* (7), 553–568.
17. Tohumoglu, G.; Sezgin, K.E. ECG signal compression by multi-iteration EZW coding for different wavelets and thresholds. *Comput. Biol. Med.* **2007**, *37*, 173–182.
18. Miaou, S.G.; Chao, S.N. Wavelet-based lossy-to-lossless ECG compression in a unified vector quantization framework. *IEEE Trans. Biomed. Eng.* **2005**, *52* (3), 539–443.
19. Kim, B.S.; Yoo, S.K.; Lee, M.H. Wavelet-based low-delay ECG compression algorithm for continuous ECG transmission. *IEEE Trans. Inf. Technol. Biomed.* **2006**, *10* (1), 77–83.
20. Nielsen, M.; Kamavuako, E.N.; Andersen, M.M.; Lucas, M.F.; Farina, D. Optimal wavelets for biomedical signal compression. *Med. Biol. Eng. Comput.* **2006**, *44* (7), 561–568.
21. Koski, A.; Tossavainen, T.; Juhola, M. On lossy transform compression of ECG signals with reference to deformation of their parameter values. *J. Med. Eng. Technol.* **2004**, *28* (2), 61–66.
22. Chatterjee, A.; Nait-Ali, A.; Siarry, P. An input-delay neural-network-based approach for piecewise ECG signal compression. *IEEE Trans. Biomed. Eng.* **2005**, *52* (5), 945–947.
23. Gacek, A.; Pedrycz, W. A genetic segmentation of ECG signals. *IEEE Trans. Biomed. Eng.* **2003**, *50* (10), 1203–1208.
24. Alexandre, E.; Pena, A.; Sobeira, M. On the use of 2-D coding techniques for ECG signals. *IEEE Trans. Inf. Technol. Biomed.* **2006**, *10* (4), 809–811.
25. Lee, H.; Buckley, K.M. ECG data compression using cut and align beats approach and 2-D transforms. *IEEE Trans. Biomed. Eng.* **1999**, *46* (5), 556–564.
26. Tai, S.C. An extensive Markov system for ECG exact coding. *IEEE Trans. Biomed. Eng.* **1995**, *42* (2), 230–232.
27. Koski, A. Lossless ECG coding. *Comput. Methods Programs Biomed.* **1997**, *52*, 23–33.
28. Blanco-Velasco, M.; Cruz-Roldan, F.; Godino-Llorente, J.I.; Blanco-Velasco, J.; Armiens-Aparicio, C.; Lopez-Ferreras, F. On the use of PRD and CR parameters for ECG signal compression. *Med. Eng. Phys.* **2005**, *27* (9), 798–802.
29. Hilton, M.L. Wavelet and wavelet packet compression of electrocardiograms. *IEEE Trans. Biomed. Eng.* **1997**, *44*, 394–402.
30. Zigel, Y.; Cohen, A.; Katz, A. The weighted diagnostic distortion (WDD) measure for ECG signal



- compression. *IEEE Trans. Biomed. Eng.* **2000**, *47* (11), 1422–1430.
31. Al-Fahoum, A.S. Quality assessment of ECG compression techniques using a wavelet-based diagnostic tool. *IEEE Trans. Inf. Technol. Biomed.* **2006**, *10* (1), 182–191.
 32. Stern, R.; Ray, W.; Quigley, K. *Psychophysiological Recording*, 2nd ed. Ed.; Oxford University Press: New York, 2001.
 33. Farina, D.; Merletti, R. Comparison of algorithms for estimation of EMG variables during voluntary isometric contractions. *J. Electromyogr. Kinesiol.* **2000**, *10* (5), 337–349.
 34. Cuerrero, A.; Mailhes, C. *On the choice of an electromyogram data compression method*, Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society: Chicago, IL, 1997; 1558–1561.
 35. Wellig, P.; Cheng, Z.; Semling, M.; Moschytz, G.S. *Electromyogram data compression using single-tree and modified zero-tree wavelet encoding*, Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Oct. 29–Nov. 1, 1998; 3, 1303–1306.
 36. Berger, P.A.; Nascimento, F.A.O.; Carmo, J.C.; Rocha, A.F.; dos Santos, I. *Algorithm for compression of EMG signals*, Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Sept. 17–21, 2003; 2, 1299–1302.
 37. Chan, A.; Lovely, D.; Hudgins, B. Errors associated with the use of adaptive differential pulse code modulation in the compression of isometric and dynamic myo-electric signals. *Med. Biol. Eng. Comp.* **1997**, *36*, 215–219.
 38. Carotti, E.; De Martin, J.C.; Merletti, R.; Farina, D. Compression of surface EMG signals with algebraic code excited linear prediction. *Med. Eng. Phys.* **2007**, *29* (2), 253–258.
 39. Grönfors, T.; Päivinen, N. *Comparison of vector quantization methods for medical fidelity preserving lossy compression of EMG signals*. International Conference on Computational Intelligence for Modelling Control and Automation CIMCA 2005, Wien: Austria, Nov. 28–30, 2005.
 40. Grönfors, T.; Reinikainen, M.; Sihvonen, T. Vector quantization as a method for integer EMG signal compression. *J. Med. Eng. Techn.* **2006**, *30* (1), 41–52.
 41. Tossavainen, T.; Juhola, M. Lossless compression of otoneurological eye movement signals. *J. Clin. Monit. Comput.* **2002**, *17*, 393–402.
 42. Tossavainen, T.; Juhola, M.; Grönfors, T. Lossy compression of eye movement and auditory brainstem response signals. *Comput. Methods Programs Biomed.* **2003**, *72*, 43–54.
 43. Juhola, M.; Tossavainen, T.; Aalto, H. Influence of lossy compression on eye movement signals. *Comput. Biol. Med.* **2004**, *34*, 221–239.
 44. Jewett, D.; Williston, J. Auditory-evoked far fields averaged from the scalp of humans. *Brain* **1971**, *94*, 681–696.
 45. Antonioli, G.; Tonella, P. EEG data compression techniques. *IEEE Trans. Biomed. Eng.* **1997**, *44* (2), 105–114.
 46. Magotra, N.; Mandyam, G.; Mingui, S.; McCoy, W. *Lossless compression of electroencephalographic (EEG) data*, IEEE International Symposium on Circuits and Systems, “Connecting the World,” 1996; 2, 313–315.
 47. Yi, K.C.; Sun, M.; Li, C.C.; Sclabassi, R.J. *A lossless compression algorithm for multichannel EEG*, Proceedings of the First Joint BMES/EMBS Conference, 1999; 1, 429.
 48. Sriraam, N.; Kannan, R.; Eswaran, C. *Lossless compression of EEG data using neural network predictors*, Proceedings of the 9th International Conference on Neural Information Processing (ICONIP ‘0Z), 2002; 4, 2046–2048.
 49. Folkers, A.; Mosch, F.; Malina, T.; Hofmann, U.G. Realtime bioelectrical data acquisition and processing from 128 channels utilizing the wavelet-transformation. *Neurocomputing* **2003**, *52–54*, 247–254.
 50. Cinkler, J.; Kong, X.; Memon, N. *Lossless and near-lossless compression of EEG signals*. Conference Record of the 31st Asilomar Conference on Signals, Systems, & Computers, 1997; 2.
 51. Memon, N.; Kong, X.; Cinkler, J. Context-based lossless and near-lossless compression of EEG signals. *IEEE Trans. Inf. Technol. Biomed.* **1999**, *3* (3), 231–238.
 52. Mitra, S.K.; Sarbadhikari, S.N. Iterative function system and genetic algorithm based EEG compression. *Med. Eng. Phys.* **1997**, *19*, 605–617.
 53. Sijerčić, Z.; Agarwal, G.C.; Anderson, C.W. *EEG signal compression with ADPCM subband coding*. IEEE 39th Midwest Symposium on Circuits and Systems, 1996; 2, 695–698.

Computer-Assisted Surgery

Thomas L. Husted
Timothy J. Broderick

Department of Surgery, University of Cincinnati College of Medicine, Cincinnati, Ohio,
U.S.A.



INTRODUCTION

Computer-assisted surgery has naturally evolved from surgical applications of improved computer and robotic technology within the minimally invasive surgical revolution. Computer-assisted surgery addresses several shortcomings of present day minimally invasive surgery (MIS) and thereby enables an intuitive return to more traditional surgical practice. Currently, computer-assisted surgery is gaining acceptance as the standard of care in several complex surgical procedures. Specific esophageal, cardiac, and pelvic operations have been improved with the use of surgical robots and, with time, the benefit of surgical robots in the performance of many more operations will become apparent. The brightest star in the rapidly expanding horizon of computer-assisted surgery is the use of the robotic platform to integrate digital technology into surgical training, practice, and quality improvement. As the integration of surgery with computers and robots continues to evolve, this article offers a broad overview of the potential of surgical robotics.

EVOLUTION OF COMPUTER-ASSISTED SURGERY

While computer and robotic technologies have continually increased in sophistication, surgeons have only recently begun to exploit these advancements. The latest innovations in surgical practice represent a complex amalgamation of advanced computer technology, robotic engineering, and paradoxically, a return to traditional surgical techniques. “Robot” is derived from the Czech word *robota* and is defined as slave labor.^[1] Therefore, the use of the term “surgical robot” is a misnomer. The correct description is surgical telemanipulator, as surgical robotic systems lack significant or sufficient autonomy in their subservient role. Currently, anatomic variability, difficulty in registering imaging data of deformable tissues in real time, and complex decision-making processes prohibit automation of most operations. For this reason, in the clinical arena, surgical robots function as slaves under

control of their human masters. Despite master–slave functionality, the surgical and lay communities continue to refer to these systems as surgical robots as opposed to telemanipulators. As such, surgical robotics is used interchangeably with computer-assisted surgery throughout the remainder of this article.

The introduction of video technology into surgical practice in the late 1980s resulted in a MIS revolution. One of the initial minimally invasive surgical procedures performed was the laparoscopic cholecystectomy (use of video equipment within the abdominal cavity to remove the gallbladder). During laparoscopic cholecystectomy, the surgeon operates through several small incisions into which are introduced a small diameter video camera and elongated surgical instruments. The laparoscopic cholecystectomy demonstrated significant advantages over traditional open cholecystectomy which included decreased tissue injury, decreased blood loss, decreased postoperative pain, faster recovery, and improved cosmetic result. Surgeons soon realized the benefit of performing a wide variety of procedures through such small incisions. Minimally invasive surgery has grown quickly and broadly to include such areas as laparoscopy (instruments and viewing optics inserted into the abdomen through small incisions), endoscopy (instruments and viewing optics inserted into the body through existing orifices), and angiography (catheters inserted through the skin to diagnose and treat abnormalities within blood vessels).

Minimally invasive surgery has several limitations, and computer-assisted surgery has been developed in part in an attempt to overcome these restrictions. The minimally invasive surgeon does not directly see or feel the tissues on which he or she operates; impaired visual and haptic perception limits the application and performance of MIS procedures. By separating a surgeon’s hands and eyes from the tissues, MIS imposes a conceptually awkward operative presence and requires a steep learning curve to master the techniques. Other limitations of MIS include a decreased range of motion with the elongated instruments, an abdominal wall fulcrum effect that requires counterintuitive hand movements, reliance on two-dimensional (2D) viewing

Table 1 Description of degree of freedom

-
- DOF is an important concept as it defines the number of independent motions of which an instrument is capable. Robotic instruments allow 7 DOF, similar to the human hand holding a surgical instrument. These seven degrees are motion in any of x, y, or z planes as well as any rotation in pitch, yaw, or roll. The seventh degree is achieved with the ability of the instrument to pinch (Fig. 1).
 - How this translates into the operative environment is that the first 2 DOF are the ability to pivot about the entry port in two planes (similar to the human wrist rotating a fist). The next 2 DOF are of the ability to move the instrument in and out of the entry port as well as rolling the instrument (similar to moving the human forearm forward and back plus pronation and supination of the forearm). The next 2 DOF allow for an arbitrary choice of position and orientation of the jaws of the instrument, a freedom not permitted with traditional endoscopic instruments (similar to the human hand being positioned in any orientation). The final degree of freedom is gained from the grasping ability of the instrument (achieved with the human hand holding a surgical instrument.)
-

with a narrow field of vision and instruments which convey only 4 degrees of freedom (DOF) (Table 1). Additionally, elongated instruments magnify hand tremors and reduce tactile feedback.

Computer-assisted surgery has successfully addressed several of these limitations, while ignoring some limitations and creating several limitations of its own. The current commercially available robotic devices have instruments with the same 7 DOF as the surgeon's hand (Table 1). Moreover, the view is stereoscopic [three-dimensional (3D)], the computer is able to filter natural hand tremor, the motion can be scaled to suit the surgeon and operation, and the fulcrum effect is eliminated. As robotic surgery approximates the traditional surgeon-patient interaction experienced in open surgery, many surgeons perceive robotic surgery as more intuitive than laparoscopic surgery.

CURRENT COMPUTER-ASSISTED SURGERY PLATFORMS

At present there are two robotic surgical systems used clinically, although one of these system is no longer manufactured or supported. The research, development, government approval, and production costs of new surgical robot systems are extremely high. In addition, intellectual property issues have stifled collaboration and development. Unfortunately, financial collapse of robotic surgical companies has been the rule rather than the exception.

Zeus System

The Zeus robotic system is a telemanipulator comprised of three robotic arms and a surgeon's console. The surgeon operates two instrument arms and a camera arm from a remotely located console usually positioned in the same room. The robotic arms individually attach

to the operating room table, and the latest version of a 5-mm-diameter-articulating instrument allows 6 DOF.

The surgeon operates the Zeus system based upon a video image that can be provided by a variety of different video systems. The video system most widely utilized with Zeus provides either high quality 2D or lower quality 3D image. Stereoscopic vision is



Fig. 1 A da Vinci intracorporeal operating instrument, illustrating the similar range of motion and DOF with the surgeon's hand.

provided by viewing polarized images specific for each eye on a monitor while wearing polarized glasses. This video system permits magnification to increase surgical precision. The surgeon uses voice commands to direct the robotic arm holding the camera. The voice-controlled robotic arm that positions the video camera was originally marketed as a stand-alone product for laparoscopic surgery as the automated endoscopy system for optimal placement (AESOP). This Robotic platform permits adjustable motion scaling and tremor filtration. Additionally, the Zeus system affords the option to use instruments as small as 3.5 mm diameter. These smaller diameter instruments are well suited to pediatric surgery.

A significant shortcoming of this robotic system was the lack of haptic feedback. Furthermore, the original design did not include articulated instruments and thus did not increase the DOF beyond traditional laparoscopic surgery. Subsequently, an articulated version was introduced, but the man-machine interface was shown to be more difficult to use than the rival commercial system described later in this entry. The performance characteristics of the Zeus system led surgeons to conclude that this computer-assisted device was more similar to laparoscopic surgery than it was to traditional open surgery.^[2] The lack of demonstrable clinical benefit over laparoscopic surgery combined with high system cost limited the adoption of this technology.

A landmark event for the Zeus system came in September 2001. A successful transoceanic operation was performed, demonstrating the feasibility of remote presence surgery or telesurgery.^[3] From a surgeon's console located in New York, U.S.A., surgeons connected via a broadband telecommunication link to a Zeus robotic system in Strasbourg, France, were able to perform a telerobotic laparoscopic cholecystectomy. This expensive, proof of concept operation was successful and the patient recovered nicely. Significantly, the surgeons demonstrated that they could overcome latency—or communication time delay—and effectively remotely operate on patients. Most importantly, they demonstrated the value of the robotic platform as a means to introduce advanced technology—in this case telecommunication technology—into the surgical process. Despite these successes, intellectual property battles and competition within a limited market forced a merger of the two major robot manufacturers and the Zeus system is no longer produced or supported.

da Vinci System

While being similar to Zeus, the da Vinci system is currently produced and supported for clinical use. The

system is made up of two major components. The first is the control console (Fig. 2). The surgeon operates while seated at the control console with his head tilted forward and eyes peering down; the surgeon's hands are in an ergonomic position and inserted in the system's master interface. The optics deliver a high-resolution 3D image through display of stereoscopic images on separate monitors directed to the right and left eyes (Fig. 3). The position of the camera can be adjusted by the surgeon to achieve the optimal view, with the possibility of magnification (2–10X). The control console is located up to 3 m from the patient-side cart positioned at the operating table. The 3 m distance is required secondary to direct cable connections between the components; although, in theory, this distance may be extended or even wirelessly transmitted.

The second component is the patient-side cart, which, in addition to the robotic arm holding the camera, also has three arms for tissue manipulation (Fig. 4). The surgeon can toggle the console to simultaneously control any two of the three instrument arms. The surgeon uses one static instrument arm to assist with exposure and retraction and the remaining two arms to operate. These complex manipulating



Fig. 2 The da Vinci control console at which the operating surgeon is seated.

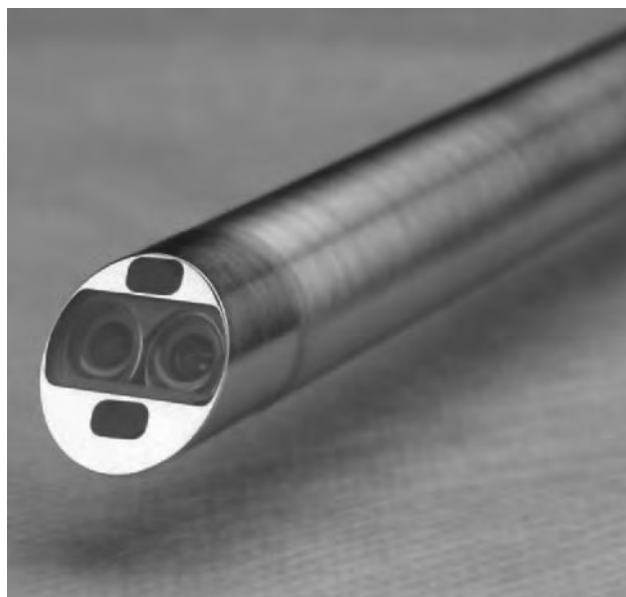


Fig. 3 Dual channeled camera of the da Vinci surgical robot, delivering stereoscopic vision to the operating surgeon.

arms have cable-driven joints and are able to achieve 7 DOF. The surgeon's hand movements at the console are digitized, sampled >1300 times per second, and sent to the slave components of the manipulators. Any nonpurposeful oscillations (tremor) are filtered. Additionally, motion scaling (up to 5X) can translate comfortable and natural hand motions into precise and delicate instrument movements. The stereoscopic imaging, intuitive interface, and articulated instruments of the da Vinci system permit the use of operative techniques very similar to traditional open techniques. The novice MIS surgeon learns MIS techniques more quickly with use of the da Vinci robot.^[2] Ongoing research has demonstrated that the da Vinci system also can serve as a platform to introduce digital technologies that enable functionality such as virtual reality simulation and telesurgery.

However, da Vinci has limitations when compared to traditional open and MIS surgery. The most significant disadvantage is the cost. A four-armed da Vinci Surgical System currently sells for approximately 1.25 million US dollars. Operating costs include a maintenance contract (10% annually) and the semi-reusable instruments (\$2000 for an individual 10-use instrument). The cost for a noncomplex robotic operation has been factored to be an additional \$2000 as compared to a traditional open operation.^[4] Another disadvantage of the da Vinci is the lack of haptic feedback. The surgeon receives limited haptic feedback from gross interactions such as forceful instrument collision, but fine haptic feedback is not obtained from interaction with soft tissues. When using this robot,

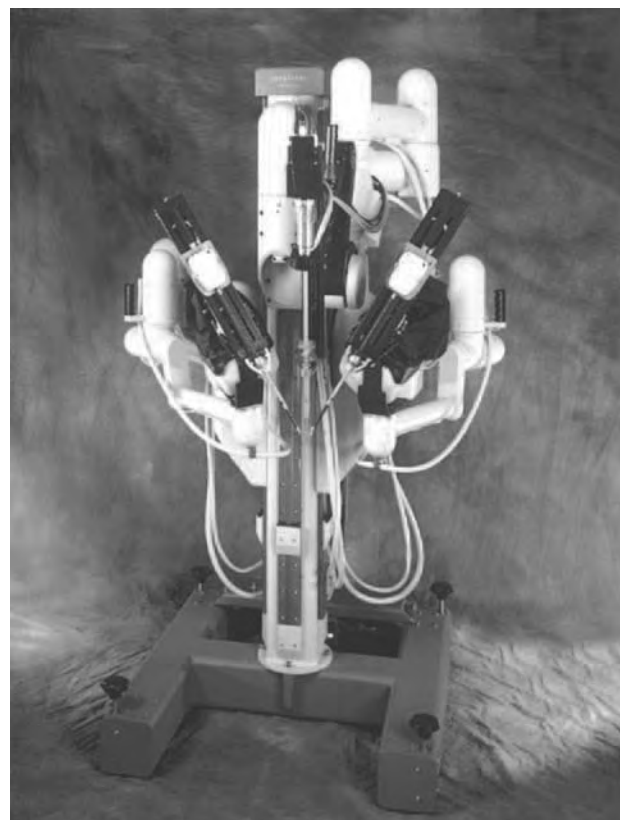


Fig. 4 The da Vinci patient side-cart, illustrating the camera and three manipulation operating arms.

surgeons cannot feel tissues and it can be difficult to diagnose and localize pathology. Proper surgical technique also mandates gentle handling of tissues and sutures. While haptic feedback is currently an area of intense research, the present day system relies on exceptional imaging to provide clues to the motion and strength of the robotic arms as they stress the operative tissues. Fortunately, the possibility of tissue injury and suture breakage from the lack of haptic feedback can be partially compensated for by training. A robotic surgeon quickly learns to use visual cues to compensate for the lack of haptic feedback.

Additional limitations are the complexity of the system. It takes longer to prepare and set-up the da Vinci system than traditional surgical equipment. Use of the da Vinci also requires specialized training for the operating room staff. A result of these limitations is the practical limitation of robotic surgery to large academic medical centers: "Centers of Excellence." In these centers, volume of use allows economy of scale to support computer-assisted surgery with clinical, technical, and financial expertise. Robotic surgery can become more expedient and effective as the surgeons and staff become familiar with the technology and procedures.^[5]

CURRENT USE OF COMPUTER-ASSISTED SURGERY

The historically glacial pace of change in medicine and the conservative nature of surgeons have translated into slow integration of computer-assisted surgery into clinical practice over the last decade. Initial experience has demonstrated safe and effective application of surgical robots in a variety of cardiac, abdominal, and pelvic operations. In addition, recent data suggest that robotic surgery benefits patients who require selective procedures as compared to standard MIS.

Esophageal Surgery

Esophageal surgery has traditionally been technically demanding, owing in part to the location of the esophagus in the posterior thoracic and abdominal cavities. Conventional open esophageal surgery has been performed through the abdomen or the chest and has been associated with significant scarring and prolonged rehabilitation. The development of advanced laparoscopic techniques enabled better visualization and manipulation of the esophagus while decreasing recovery time and improving the outcome of the surgery. However, the application and effectiveness of MIS esophageal surgery is limited by the use of rigid laparoscopic instruments.

The da Vinci system has proven itself useful in esophageal surgery. While a number of robotic esophageal operations have been demonstrated to be safe, one of the most promising applications of the da Vinci system has been the esophageal myotomy (in which the surgeon divides an unusually tight lower esophageal muscle to allow the patient to swallow normally). In a multicenter study, surgeons performed nearly 100 esophageal myotomies with no perforations (the most serious complication) and only minor complications not attributable to the use of the robotic system. The researchers credit this success to the magnified 3D visualization and the enhanced range of motion of the high DOF articulating instruments.^[6-8]

Additionally, robotic application in antireflux surgery has been validated in two controlled clinical trials in which it was found that computer-assisted fundoplication (wrapping the stomach around the lower esophagus to create a valve that will limit heartburn caused by the reflux of stomach contents into the esophagus) to be feasible and safe in a series of 21 and 40 patients, respectively.^[9,10] However, these studies did find that the use of the robot resulted in longer operative time and increased the cost when compared to conventional laparoscopic fundoplication. With more frequent use of the device, robotic operative times approach laparoscopic operative times.

Cardiac Valve Surgery

As early as in 1998, the first truly endoscopic mitral valve repair was performed using a prototype of the da Vinci system.^[11] Evolution of cardiac valve repair has seen the requirement for an assistant's hand eliminated with the totally robotic mitral valve repair using only one-centimeter ports as a routine choice for patients having isolated mitral valve pathology.^[12]

Initial experience with robotic mitral valve repair during Food and Drug Administration (FDA) Phase Two trials showed good success for 112 patients among 10 centers.^[13] Despite a 5.4% reoperation rate, the study demonstrated that multiple surgical teams could perform robotic mitral valve surgery safely early in the development of this technique, thus predicting a flattened early learning curve.^[12]

After receiving FDA approval for mitral valve surgery, the da Vinci system has been used successfully in several series of patients.^[12,14,15] The largest series report 150 robotic mitral valve repairs with one operative death (0.67%) from a medication reaction and no device-related complications. Operative times continued to decrease until 100 operations had been performed, attributed to improved set up and deployment of the robot and closer working relations among the operating room staff. As technology advances, there will be an increased use of surgical robot in such cardiac operations as coronary artery bypass grafting and correction of arrhythmias (irregular heart beat). Innovative robotic and anastomotic technology currently under development will spur the application of MIS in cardiac surgery.

Pelvic Surgery

Removal of the prostate for localized prostate cancer is the most commonly performed robotic operation. A traditional approach to cancer of the prostate requires a large abdominal incision to remove the prostate, but the laparoscopic approach uses several small incisions and provides similar results with regard to tumor resection and side effects. With both of these approaches, however, visualization of and suturing the bladder to the urethra after removal of the prostate remain problematic. The difficulty in carefully dissecting the prostate and avoiding the adjacent neurovascular bundle and sphincter complex has led to life altering side effects: retained tumor, loss of urinary continence, and sexual impotence. Surprisingly, incisional pain and operative discomfort are minor complaints among men undergoing this procedure.^[16]

Computer-assisted surgery has demonstrated significant advantages in removal of the prostate in completeness of tumor resection as well as in decreasing the



side effects of loss of continence and sexual potency. Superior visualization and articulated instruments allow more precise dissection of the prostate from the surrounding nerves and vessels. Improved dissection leads to more complete tumor resection and preservation of the nerves responsible for sexual potency. Additionally, suturing facilitated by the articulated instruments allows a more accurate anastomosis of the urinary bladder and urethral sphincter, preserving urinary continence.^[16]

The largest series to date include over 250 patients who received robotic-assisted radical prostatectomy with less intraoperative blood loss, fewer positive tumor margins, decreased length of stay in hospital, and only a 4% complication rate. Furthermore, sexual potency was preserved in 78% of patients and urinary continence was preserved in 96% patients.^[16,18] Better patient outcomes confirm that surgical robots provide benefit primarily by enabling the surgeon to dissect and suture better in this difficult anatomic region.

Additional Applications

In every facet of surgery where laparoscopy has taken a foothold, the use of surgical robots has been introduced and has been favorably received by some surgeons and patients. Surgeon acceptance originates from excellent 3D visualization, and improved capability to dissect and suture with high DOF instrumentation. A broad range of operations have been successfully performed with the robot including cardiac, thoracic, esophageal, gastric, intestinal, biliary, hepatic, pancreatic, gynecologic, pediatric, and urologic procedures.^[19] Only time will reveal whether robotic systems will evolve sufficiently to permit application in the few remaining specialties such as trauma and transplant surgery.

FUTURE OF COMPUTER-ASSISTED SURGERY

The potential of computer-assisted surgery is nearly limitless. The next step is correction of the current list of shortcomings. Initial improvements will include adaptation of common surgical instrumentation for use with the robot. These instruments include suction-irrigators, multifire clip appliers, and surgical staplers. Improvements will also address the increased length of time required to perform robotic surgery as the result of longer set up and instrument exchange times. Adapting the platform to make it more “plug and play” will decrease the set up and disassembly time. Additionally, multifunction instruments should decrease operative time by decreasing instrument exchange time. Another limitation that hopefully will be

addressed in the near future is the lack of tactile sensation. Providing haptic feedback to the surgeon should improve surgical performance and quality. In the short term, surgeon compensation for the lack of haptics through visual clues is promising.^[17]

Keeping in step with ever quicker and more powerful computer technology, surgical assist devices will decrease their footprint despite increasing capability. Eventually, the ever-increasing portability and capability of modern computers will be translated into new robotic platforms. This development is eagerly awaited by the military and National Aeronautic and Space Administration. Small, portable, deployable robotic systems could serve as a platform to improve surgical treatment in hostile environments. Despite the lack of a local surgeon, remotely guided expert treatment provided at the point of injury could save the life and limbs of soldiers on the battlefield. From a centralized location and out of the direct line-of-fire, a single surgeon could become the battlefield medic for countless combat troops. Networked robotic platforms could become a force multiplier for limited surgical assets. A working model of telesurgery has been attempted using the internet as the communication network. Surgeons in Cincinnati, Ohio have mentored and guided Surgical and Urological Residents in Sunnysvale, California, through robotic procedures using porcine models. While the latency of the internet transmissions was cumbersome, the teams were able to compensate for the delay and the principle was proven possible. As developments in transmission bandwidth are advanced, the latency issue should be overcome.

As we further explore space in the upcoming decades, robust surgical capabilities are required to insure crew health and mission success. While an earth-bound surgeon could successfully operate on a sick astronaut in earth orbit, distance and corresponding communication latency dictate that autonomous surgical care will be required in planetary exploration. Exploration enabling surgical technology is currently being developed. Astronauts on a Martian mission in the upcoming decades could benefit from a robotic platform that enables autonomous surgical care that includes just-in-time virtual reality training, planning, and practice of an operation on patient specific data, and semiautonomous and autonomous function of some surgical tasks.

In the future, robotic technology could facilitate development of intraluminal surgery (operating inside structures such as the intestinal tract or blood vessels). Natural orifice surgery is thought to represent the next frontier of surgery by some surgical thought leaders. Robotics could allow surgeons to perform such “natural orifice” surgery as taking a gallbladder out through the mouth. In the longer term, deployable miniaturized robotics and intraluminal endovascular

therapy could revolutionize minimally invasive therapy of organs such as the heart or brain. Surgeons have begun to explore the use of early versions of such systems.

Future surgical robots will also be used to facilitate surgical training. Similar to the use of virtual reality simulators in the aviation industry, surgical robotic systems will serve as the platform and interface through which surgeons obtain training. Surgical education will move from a time-based apprenticeship to a proficiency-based educational process. Trained surgeons will use such systems to plan and practice complex operations on virtual representations of the patient and thereby optimize surgical performance and patient outcome.

As all the information transmitted via the robot and optical console is digitized, there is incredible potential for integrating patient, surgeon, and procedure related data. Surgeon controlled inputs, endoscopic video, radiologic imaging, and physiologic data are available as digital media. As the digital interface between the surgeon and patient, the robotic system is the information platform that will integrate these data. For example, “augmented reality” that is similar to the onlay of images seen in a fighter pilot’s heads-up display could allow the surgeon to safely remove a tumor within a solid organ by “viewing” the tumor with fused radiological and video images. Aviation’s digital “fly-by-wire” systems are finally reaching medicine as digital “surgery-by-wire” systems.

The ultimate evolution of robotic surgery is a true surgical robot. A database of human anatomy, human physiology, and operative procedures will be combined with improved robotics and artificial intelligence to allow transition of the robot from a slave to a colleague. Its role will broaden from reducing surgical error (establishing and preventing known common errors) to eventually replacing the surgeon as a truly automated robot. Eliminating the surgeon from robotic surgery is considered by many to be science fiction, but the rapid development and application of computers and robotics in the last century suggest this is probably better considered science eventuality.

CONCLUSIONS

The last half century has seen amazing developments in technology and the integration of computers, robotics, and medicine is natural. Current robotic surgical systems deliver two advantages: greatly improved visualization and high DOF instruments that continue to push the application and performance of MIS. Computer-assisted surgery will continue to evolve and further improve surgical quality and safety.

While still suffering from growing pains and limitations, computer-assisted surgery has proven itself useful, safe, and effective in esophageal, cardiac, and urologic surgery. While these operations have little in common anatomically or procedurally, robotic surgery has proven of benefit secondary to improved 3D visualization and dexterity in manipulation of tissues. Increasing collaboration between creative surgeons and engineers will push development and application of robotic surgery and thereby improve surgical practice.

The future of surgery will increasingly incorporate robots, as new devices prove themselves to be safe, easy to operate, capable, and cost effective. Although faced with several issues, robotic surgery will ultimately demonstrate its value with superior safety and performance of a variety of minimally invasive procedures.

ARTICLE OF FURTHER INTEREST

Telemedicine

REFERENCES

1. Capek, K. *Rossum's Universal Robots*; Landes, W.A., Ed.; Doubleday: New York, 1923.
2. Sung, G.; Gill, I. Robotic laparoscopic surgery: a comparison of the da Vinci and Zeus systems. *Urology* **2001**, *58*, 893–898.
3. Gottlieb, S. Surgeons perform transatlantic operation using fiberoptics. *BMJ* **2001**, *323*, 177–181.
4. Costi, R.; Himpens, J.; Bruyns, J.; Cadiere, G.B. Robotic fundoplication: from theoretic advantages to real problems. *J. Am. Coll. Surg.* **2003**, *197*, 500–507.
5. Nifong, L.W.; Chitwood, W.R., Jr. Building a surgical robotics program. *Am. J. Surg.* **2004**, *188* (Suppl. 4A), 16S–18S.
6. Jacobsen, G.; Elli, F.; Horgan, S. Robotic surgery update. *Surg. Endosc.* **2004**, *18*, 1186–1191.
7. Ruurda, J.P.; Gooszen, H.G.; Broeders, I.A. Early experience in robot-assisted Heller myotomy. *Scand. J. Gastroenterol. Suppl.* **2004**, *241*, 4–8.
8. Talamini, M.A.; Chapman, S.; Horgan, S.; Melvin, W.S. A prospective analysis of 211 robotic-assisted surgical procedures. *Surg. Endosc.* **2003**, *17*, 1521–1524.
9. Cadiere, G.B.; Himpens, J.; Vertruyen, M.; Bruyns, J.; Gernay, O.; Leman, G.; Izizaw, R. Evaluation of tele-surgical (robotic) Nissen fundoplication. *Surg. Endosc.* **2001**, *15*, 918–923.
10. Melvin, W.S.; Needleman, B.J.; Krause, K.R.; Schneider, C.; Ellison, E.C. Computer-enhanced vs. standard laparoscopic antireflux surgery. *J. Gastrointest. Surg.* **2002**, *6*, 11–16.
11. Carpentier, A.; Loulmet, D.; Aupecle, B.; Kieffer, J.P.; Tournay, D.; Guibourt, P.; Fiemeyer, A.; Meleard, D.;



- Richomme, P.; Cardon, C. Computer assisted open-heart surgery: first case operated on with success. *C R Acad. Sci. III* **1998**, *321*, 437–442.
12. Kypson, A.P.; Chitwood, W.R., Jr. Robotic mitral valve surgery. *Am. J. Surg.* **2004**, *188* (Suppl. 4A), 83S–88S.
 13. Chitwood, W.R., Jr.; Nifong, L.W. Minimally invasive videoscopic mitral valve surgery: the current role of surgical robotics. *J. Card. Surg.* **2000**, *15*, 61–75.
 14. Mohr, F.W.; Falk, V.; Diegeler, A.; Walther, T.; Gummert, J.F.; Bucarius, J.; Jacobs, S.; Autschbach, R. Computer-enhanced “robotic” cardiac surgery: experience in 148 patients. *J. Thorac. Cardiovasc. Surg.* **2001**, *121*, 842–853.
 15. Kypson, A.P.; Nifong, L.W.; Chitwood, W.R., Jr. Robotic mitral valve surgery. *Semin. Thorac. Cardiovasc. Surg.* **2003**, *15*, 121–129.
 16. Smith, J.A., Jr. Robotically assisted laparoscopic prostatectomy: an assessment of its contemporary role in the surgical management of localized prostate cancer. *Am. J. Surg.* **2004**, *188* (Suppl. 4A), 63S–67S.
 17. Hanly, E.J.; Talamini, M.A. Robotic abdominal surgery. *Am. J. Surg.* **2004**, *188* (Suppl. 4A), 19S–26S.
 18. Tewari, A.; Menon, M. Vattikuti institute prostatectomy: surgical technique and current results. *Curr. Urol. Rep.* **2003**, *4*, 119–123.
 19. Camarillo, D.B.; Krummel, T.M.; Salisbury, J.K., Jr. Robotic technology in surgery: past, present, and future. *Am. J. Surg.* **2004**, *188* (Suppl. 4A), 2S–157S.

Conductive Polymers

Gordon Wallace
Leon Kane-Maguire

Intelligent Polymer Research Institute, University of Wollongong,
Wollongong, New South Wales, Australia



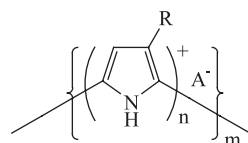
INTRODUCTION

The use of electrical stimuli to control and promote the adhesion of living cells and biomolecular events occurring within such structures is not a new area of study. Previous studies have involved the use of inorganic (metal) or simple organic (carbon) electrodes as conduits for electrical stimulation to facilitate bone regrowth wound healing and nerve regeneration.

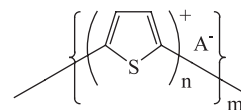
Although the mechanism of operation is not clearly understood, it is believed that changes in surface hydrophobicity, interfacial pH, and ionic strength contribute to the overall effects observed. In summary, it is not only the ability to provide a conduit for electron transfer reactions that may be important in biomolecular events but also the ability to induce appropriate interfacial chemistries. This makes the use of inherently conducting polymers (ICPs) as biomaterials a very exciting prospect. In recent years an understanding of the dynamic interfacial chemistry that accompanies the oxidation/reduction of ICPs has been established and is now being used to advantage in designing revolutionary new biomaterials.

WHAT ARE ICPs?

A large array of organic polymeric structures capable of functioning as electronic conductors are now available.^[4,5] To date most attention has been paid to the easily prepared and relatively stable ICP structures based on polypyrroles, polythiophenes, and polyanilines, shown as Chemical Structures 1, 2, and 3 (where A^- = counteranion, n determines the level of doping (normally between 3–4 for PPy and PTh), and m determines the molecular weight)

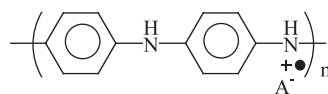


1
Polypyrrole (PPy)



2

Polythiophene (PTh)

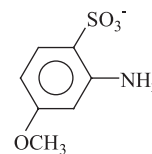


3

Polyaniline (PAN)

These simple ICP materials typically have electrical conductivities in the range of $1\text{--}200\text{ S cm}^{-1}$. For comparison, copper (an excellent conductor) exhibits a conductivity of $5 \times 10^5\text{ S cm}^{-1}$. The simplest ICP materials formed are generally neither solution- nor melt-processable.

However, even within these simple materials the chemistry available is rich and diverse (discussed later). Polypyrrole or polythiophene backbones can be chemically modified to covalently attach groups (R) capable of inducing organic solvent solubility or alkyl sulfonate groups inducing solubility in aqueous media. With polyanilines, attachment of alkyl or alkoxy groups to the aromatic ring can similarly induce organic solvent solubility. Alternatively, the aromatic ring can be directly sulfonated to induce water solubility after polymer synthesis, or sulfonated monomers (shown in Chemical Structure 4) can be used to produce water-soluble sulfonated polyanilines.



4

An alternative, highly flexible, and often facile route to functionalizing ICPs is by varying the nature of the counteranions (A^-) incorporated along the polymer chains. These dopant anions may vary from simple ions

(such as Cl^- or BF_4^-) to complex biochemicals, including proteins, enzymes, antibodies, ATP, DNA, and even a whole living cell (discussed later).

The organic nature of these ICP structures provides a potential molecular-level compatibility with living systems, and the versatile chemistry currently available provides the biomaterials engineer with an array of approaches to the design of appropriate structures for particular bioenvironments. The unique ability of ICPs to conduct charge at charge densities comparable to or greater than conventional electrodes (see Cellular Communications) and the ability to tune the chemistry or biochemistry of these surfaces in situ explains the excitement of materials scientists and engineers working in this field.

SYNTHESIS OF ICPs

ICPs such as polypyrrole, polythiophene, and polyaniline are most commonly prepared via the chemical or electrochemical oxidative polymerization of the respective monomers. However, a range of polymerization techniques has now been developed, including 1) electrochemical polymerization; 2) chemical polymerization; 3) photochemically-initiated polymerization; and 4) enzyme-catalyzed polymerization.

Both aqueous and organic solvents have been widely employed for the synthesis of polypyrroles and polythiophenes, whereas aqueous solutions are typically used for synthesizing polyanilines. A counteranion (A^-) is incorporated during synthesis in order to balance the positive charges that are generated on the developing polymer chains. With pyrrole and thiophene oxidations, the counteranions can be added to the polymerization solutions as neutral salts. However, for polyaniline syntheses, acidic conditions ($\text{pH} \leq 3$) are usually required, and the counteranions are often introduced as the corresponding acids (HA). These acidic conditions are necessary to solubilize the aniline monomers in the aqueous solvent and to ensure the deposition of the electrically conducting emeraldine salt form of the polyaniline.

Depending on the conditions employed, the ICPs can be produced as films, powders, colloids, or water-soluble materials.

Electrochemical Synthesis

Electrochemical syntheses of ICPs are usually carried out in a three-electrode cell, with the polymer products depositing as films on the working electrode (anode), as illustrated in Fig. 1 for polypyrrole syntheses. A variety of working electrodes may be used, including platinum or gold plates, glassy carbon, and reticulated vitreous

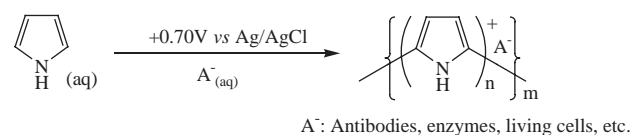


Fig. 1

carbon, as well as glass coated with indium-tin-oxide (ITO). Use of the optically transparent ITO-coated glass working electrodes allows the measurement of the absorption (UV-visible) and circular dichroism spectra of the ICP films.

Constant potential (potentiostatic), potentiodynamic (potential cycling), or galvanostatic (constant current) techniques may be employed. The rate-determining step (requiring the highest positive potential) in these electropolymerizations is the oxidation of the monomers to their respective radical cation species (shown in Fig. 2 for pyrrole ($\text{X}=\text{NH}$) and thiophene ($\text{X}=\text{S}$)). The anodic potential required to achieve monomer oxidation increases along the series, pyrrole, aniline, thiophene. Dimeric and higher oligomeric species formed in subsequent steps are considerably easier to oxidize (e.g., bithiophene has an oxidation potential 0.76 lower than thiophene itself). This has given rise to some problems with overoxidation of the polymer products, especially in the case of polythiophenes.^[6] At the potential required to oxidize the thiophene monomer, the polythiophene product becomes overoxidized, causing a deterioration in its physical and chemical properties. An approach that has been successfully employed to overcome this problem is the addition of small amounts of bi- or ter-thiophenes to reduce the polymerization potential.^[7] Defects arising from overoxidation of polyaniline can also occur, which may be associated with crosslinking or the opening of the chains after the formation of a paraquinone.^[8] Therefore, the use of potentiostatic growth conditions is recommended for polyaniline syntheses in order to minimize the likelihood of overoxidation.

An important requirement for the oxidative formation of ICPs with extended π -conjugation and high conductivity is regioselective coupling of the initially

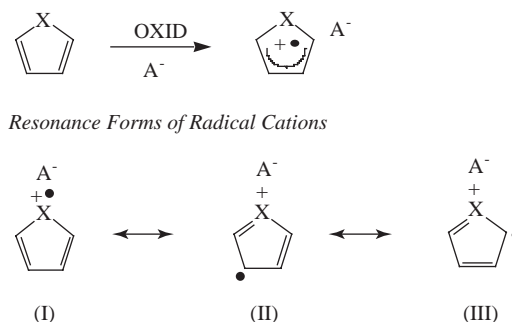


Fig. 2

formed radical cations. For example, this involves coupling of the alpha-sited radicals (III, Fig. 2) for pyrrole and thiophene.

A useful recent advance for the electrochemical synthesis of water-soluble or -dispersible ICPs has been the development of an electrohydrodynamic technique, employing a flow-through electrochemical cell with porous working electrodes such as reticulated vitreous carbon.^[9]

Chemical Polymerizations

The most common oxidants employed in the chemical oxidation approach to ICPs are FeCl_3 and $(\text{NH}_4)_2\text{S}_2\text{O}_8$, which provide Cl^- and $\text{HSO}_4^-/\text{SO}_4^{2-}$, respectively, as the dopant anions (A^-). An attractive feature of chemical polymerization is its experimental simplicity and ready scale-up to produce bulk quantities of ICPs. However, the alternative electrochemical oxidation route has the advantage of greater control over the oxidation potential applied (thereby limiting over-oxidation) and greater flexibility in terms of the anions that can be incorporated from the electrolyte (MA salt or HA acid).

Properties of ICPs such as their morphology, electrical conductivity, molecular weight, and stereoregularity are sensitive to the polymerization conditions employed (as well as to the nature of the monomer and dopant anion).^[10] For example, it has recently been shown from circular dichroism spectral studies that electrochemically and chemically generated PAn.(+)-HCSA polymers (where the CSA^- dopant is the chiral camphorsulfonate anion) possess different conformations/structures for their polyaniline chains.^[11]

Photochemical Polymerizations

Alternative photochemical routes to ICPs have been reported. Using either $[\text{Ru}(\text{bipy})_3]^{2+}$ (bipy = 2,2'-bipyridine) or $[\text{Cu}(\text{dpp})_2]^+$ (dpp = 2,9-diphenyl-1,10-phenanthroline) as the photosensitizer and an appropriate electron acceptor (sacrificial oxidant), pyrrole has been successfully polymerized to polypyrrole by irradiation with visible light.^[12] More recently, Kobayashi et al.^[13] reported the photopolymerization of aniline (in the presence of a small amount of dimer) via the visible irradiation of films containing $[\text{Ru}(\text{bipy})_3]^{2+}$ and methylviologen (MV^{2+}) as sacrificial oxidant. Using appropriate masks, these photochemical routes have been used to generate patterns of ICPs on substrates (with definitions as low as $2\ \mu\text{m}$). We have recently found^[14] that by employing $[\text{CoX}(\text{NH}_3)_5]^{n+}$ ($\text{X} = \text{Cl}$ or H_2O) as the sacrificial oxidant instead of MV^{2+} , aniline can be readily polymerized in the presence of $[\text{Ru}(\text{bipy})_3]^{2+}$ to polyaniline without the need to add

aniline dimer. We have employed the route to produce conducting polyanilines as bulk powders and as polyaniline-impregnated fabrics.

Enzyme-catalyzed Polymerizations

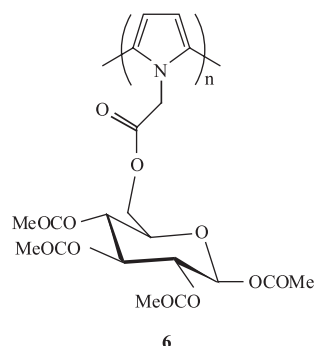
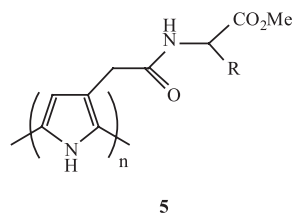
Enzymes such as horseradish peroxidase have seen limited employment as catalysts in the synthesis of polyanilines and polypyrroles.^[15] For example, Samuelson et al.^[16] used polyelectrolytes such as poly(styrenesulfonate) as templates in enzyme-catalyzed polymerizations to induce more structural order during polyaniline growth. Another advantage of this enzyme-catalyzed route is the considerably higher pHs that can be employed, compared to chemical and electrochemical polymerization methods. The mild (pH 4.3) conditions for these HRP-catalyzed syntheses have permitted the use of more delicate biological polyelectrolytes such as DNA as the aligning templates.^[17]

Adding Biological Functionality to ICPs

The presence of substituents such as alkyl, alkoxy, or sulfonate groups on the monomer repeat units of polypyrroles, polythiophenes, and polyanilines can lead to markedly enhanced solubility in organic solvents or water (as discussed earlier). Related modifications of ICPs have improved their biocompatibility and provided them with a range of biological functions. Three approaches have typically been employed to achieve this additional functionality: 1) covalent attachment of specific groups to the CEP backbone (either pre- or postpolymerization); 2) incorporation of specific dopant anions during polymerization; and 3) formation of hybrid ICP/biopolymer composites.

In the first approach, a monomer containing the appropriate biologically active group is synthesized and then oxidized to produce the modified ICP.^[18] For example, the covalent attachment of amino acids to the β -ring position of pyrrole and thiophene monomers prior to their oxidation has provided routes to substituted ICPs, as shown in Chemical Structure 5, ($\text{R} = \text{CH}_2\text{OH}$, CHMe_2 , Ph).^[19] The presence of the chiral substituents is believed to induce a one-handed helical structure on the polymer backbones of the polypyrroles and polythiophenes formed. *N*-substituted polypyrroles possessing main chain chirality, as shown in Chemical Structure 6, have also been generated via the covalent attachment of chiral groups (e.g., sugars) to the N center of the pyrrole monomer prior to polymerization.^[20] In some cases, e.g., the polymers, shown as Chemical Structures 5 and 6, these optically active polypyrroles and polythiophenes have been shown to possess the ability to discriminate between the enantiomers of chiral molecules and ions.





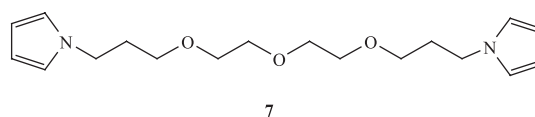
A drawback to this approach to functionalized ICPs is potential difficulties in the synthesis of the initial substituted monomer. Furthermore, subsequent oxidation to the desired polymer may prove difficult due to steric hindrance by the functional group or electronic effects, causing an increase in the oxidation potential of the monomer. A useful development in this regard is the recent report of a route involving the facile modification of preformed polypyrroles containing good leaving groups such as *N*-hydroxysuccinamide.^[21] Using this approach, oligonucleotides have been covalently attached to polypyrrole backbones and their hybridization with appropriate base pairs used for direct hybridization detection and nucleotide sequencing information.^[22] This has led to the development of DNA biochips for biochemical sensing.^[23,24]

Others^[25] have attached 3-*N*-hydroxyphthalimide to polypyrrole or polythiophene backbones to provide facile leaving groups for attachment of antigens. Enzymes (e.g., GOD) have been covalently attached to polypyrroles^[26] or polythiophenes^[27] containing carboxy functional groups. A range of poly(3,4-ethylenedioxythiophene) (PEDOT) derivatives containing hydroxyl groups have been used to covalently attach GOD.^[28] Covalently attached biotin groups have been similarly used to facilitate immobilization of biological entities.^[29,30]

The second approach to adding biological functionality to ICPs avoids the synthetic problems of the first method by entrapping the desired biologically active group as the dopant anion (A^-) or via physical entrapment into the conducting polymer chains during synthesis. These dopants may be relatively small biomolecules such as adenosine triphosphate (ATP). For example, Reynolds et al. have reported the

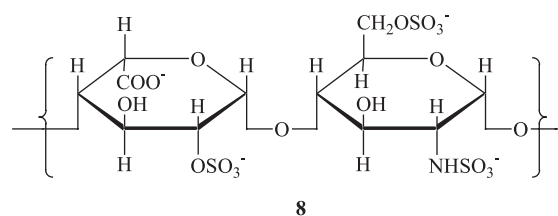
synthesis of PPy.ATP species.^[31] They showed that the ATP could be released via electrochemical stimulation, suggesting that such ICP-based materials may be useful as ATP supply devices.

Enzymes have also been directly entrapped during the electropolymerization process carried out in the presence of additional electrolyte. This approach has been used to incorporate glucose oxidase,^[32,33] sulfatase,^[34] and urease.^[35] The use of specific monomers (see Chemical Structure 7) to induce a chemical environment to increase trapping efficiency has been demonstrated.^[36] Antibodies have also been incorporated directly into conducting polymers at the point of electrosynthesis. Incorporation of antihuman serum albumin,^[37] an antibody to thaumatin,^[38] and anti-isoproturon^[39] are examples of immunopolymers fabricated using this approach.



A range of polyelectrolyte biological dopant anions (A^-) have also been successfully incorporated into ICP chains. In the synthesis of polyaniline-based biomaterials, the polyelectrolyte provides the additional benefit of binding to and preferentially aligning with the aniline monomers prior to polymerization, thereby facilitating the desired head-to-tail coupling of the monomer units, as described by Samuelson et al. in the template-assisted synthesis of PAn.DNA. In related studies, salmon sperm DNA has been incorporated into polypyrrole chains, producing PPy.DNA membranes that could be employed for electrochemically controlled transport of simple ions.^[40]

Similarly, biologically active polyelectrolytes such as the anticoagulant heparin have been incorporated into polypyrrole^[41] (see Chemical Structure 8.) During redox cycling, these large biological dopants are retained in the polymer film and instead cations are reversibly incorporated from solution. Oligonucleotides have also been employed as dopants in ICPs,^[42] and the hybridization of these oligonucleotides with the appropriate base pair gives rise to useful electronic signals that can be used for diagnostic purposes.



The third approach is related to the second, but the polymeric biological entities are not necessarily entrapped physically or as dopants, but may form blends via interactions such as H-bonding or hydrophobic or dispersion forces. For example, hybrid biomaterials have been reported that combine the biologically active polysaccharide chitosan with both polypyrrole and polyaniline.^[43,44] These were prepared via the oxidation of the pyrrole and aniline monomers, respectively, in the presence of 2% (w/w) chitosan in 2% (v/v) acetic acid.

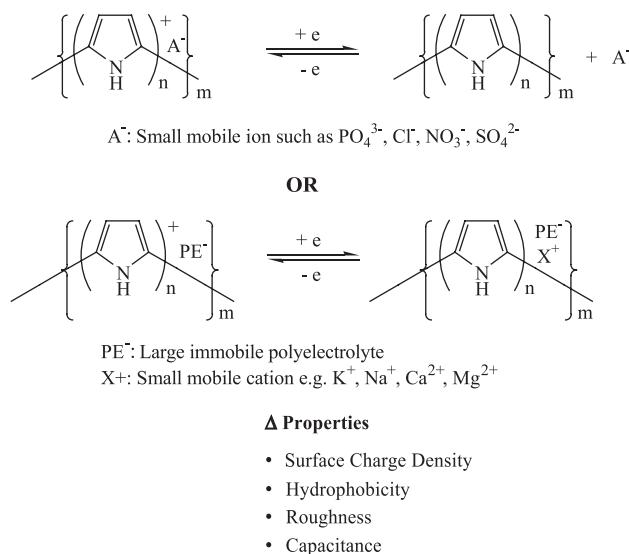
Polypyrrole-collagen composites have been similarly synthesized that contain between 4% and 35% (w/w) of the protein collagen.^[45] The solid products resemble polymer blends and show good electrical conductivity (e.g., 62 and 1.4 S cm⁻¹ for composites containing 4% and 26% collagen, respectively). It has been suggested that these solid hybrid biomaterials may be useful for applications of collagen in areas such as artificial nerves and blood vessels where charge or electrical conductivity are necessary. Using a water-soluble pyrrole derivative (bearing a butanesulfonate substituent) as monomer, related substituted polypyrrole-collagen hybrids have been generated as gels.^[46] As well as thereby achieving improved processability, the collagen was found to be more strongly bound to its ICP partner due to additional possibilities for electrostatic and hydrogen-bonding interactions between the two polymeric components.

ELECTROCHEMICAL PROPERTIES OF ICPs

In addition to their electrical conductivity, a characteristic of ICPs most useful in many bioapplications is their ability to be readily and reversibly switched between different oxidation states, as illustrated in Fig. 3 for polypyrrole. These reversible oxidations/reductions may be achieved by the simple application of an electrical potential. The oxidized and reduced species have markedly different chemical and physical properties. Such redox switching therefore provides an in situ mechanism of control over chemical, electronic, and physical properties (Fig. 3).

Besides providing a route to incorporate specific bioactive dopants (such as antibodies or enzymes) into the ICPs, the choice of dopant is also critical in determining a number of other important factors that govern behavior during electrochemical switching.^[5] These include initial conductivity, morphology, switching potential (i.e., the nature of the electrochemical processes occurring), hydrophobicity/hydrophilicity, water content, and mechanical properties (tensile strength/elongation to break).

It has been shown that the strength of individual molecular interactions occurring on ICPs is determined



(All expected to impact on extracellular protein/cell adhesion)

Fig. 3 Oxidation/reduction of polypyrrole.

by the initial composition (polymer backbone and dopant), and can be altered by electrical stimulation. For example, the ion exchange capacity of polypyrrole is obviously diminished in the reduced state, making it a more hydrophobic material.

Secondary effects (such as reduction of oxygen in solution with a subsequent increase in pH at the polymer solution interface) will also influence the molecular interactions that take place on the polymer surface. With some ICPs (e.g., polyaniline), these secondary (solution) effects are amplified by the fact that the change in oxidation state of the polymer involves release/uptake of protons.

It is presumed that the effect of applied potential (and the subsequent redox state of the polymer) on these primary molecular interactions results in the ability to control the adsorption of larger molecules such as proteins on ICP surfaces.^[47,48]

ICP electrode substrates have also been used to demonstrate the ability to provide in situ control over how some proteins interact with each other. It has been shown^[37,38] that antibody protein molecules can be incorporated directly into ICPs during electrodeposition (shown as the dopant A⁻ in Fig. 1). These antibody-containing materials retain the ability to selectively interact with the target (antigen) molecule. It is particularly interesting that the degree and reversibility of this interaction can be modulated using appropriate electrical stimuli.

CELLULAR COMMUNICATIONS USING ICPs

The earliest report involving cells and inherently conducting polymers was in 1990 by Deshpande and

Hall,^[49] wherein whole banana cells were immobilized into polypyrrole to create a biosensor for detection of dopamine. The enzyme polyphenol oxidase (contained in banana cells) converts dopamine to quinone, involving oxygen consumption that can be monitored to provide the sensing signal. The whole cells were incorporated during electropolymerization (Fig. 1) from a phosphate buffer-containing solution.

Aizawa and coworkers^[50] demonstrated the ability to modulate the acid phosphatase gene expression in *Saccharomyces cerevisiae* (yeast cells) using inherently conducting polymer platforms. Expression of the PHO5 gene, which codes for the major acid phosphatase in *Saccharomyces cerevisiae*, is inversely proportional to the phosphate concentration in the medium, coming out of repression when external phosphate levels fall below ca. 700 μM . The application of appropriate electrical stimuli to conducting polymers in contact with yeast cells was found to modulate the gene's expression via localization of ions. Aizawa et al. also demonstrated the use of the conducting polymer to facilitate electron transfer and hence control enzyme activity (fructose dehydrogenase).^[50]

Yeast cells have also been brought into contact with polypyrrole surfaces by introducing an agarose layer containing the cells.^[51] A change in pH resulting from metabolic activity of the cells caused a change in the electronic properties of the polypyrrole platform, which could then be used to monitor metabolic activity and to detect the presence of metabolic inhibitors (Fig. 4).

A more recent example^[52] of utilizing molecular events within cells as the basis of a biosensor involves the immobilization of red blood cells (RBCs) in polypyrrole (Fig. 5). These studies found that the use of a polyelectrolyte as the dopant (represented by A^- in Fig. 1) resulted in a cell-compatible environment for the red blood cells. The use of a polyelectrolyte essentially produced an electronic hydrogel whose resultant material had water content greater than 90%. The biomolecular event of interest here was the binding of anti-D to the D antigen in the RBC membrane. It was established that in the presence of anti-D, the electronic properties of the polymer containing RBCs changed.

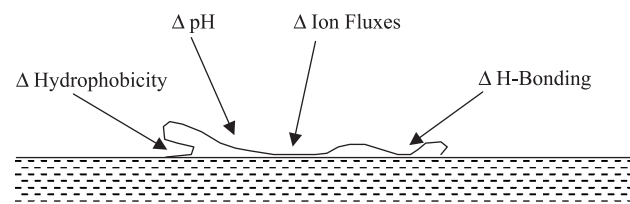


Fig. 4 Factors affecting protein adsorption on ICP surfaces.

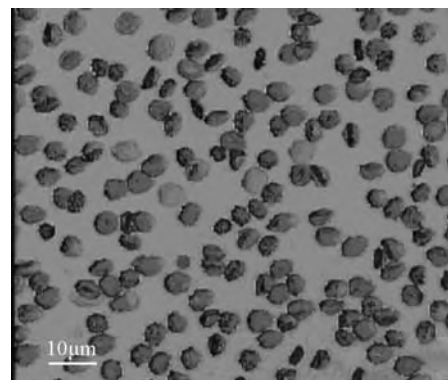


Fig. 5 Polypyrrole-containing red blood cells fixed in glutaraldehyde (0.15%)/phosphate buffer immediately postsynthesis. Scale bar 10 μm . (Courtesy of T. Campbell, Ph.D. thesis, University of Wollongong, 1999.)

Others^[53] have electrodeposited polypyrroles with polystyrene sulfonate as dopant onto the gold or indium electrodes of micromachined neural probes. It was found that the impedance of the coated electrode decreased with polymer thickness up to at least approximately 13 μm . Acute in-vivo recording tests were conducted on the cerebellum of a guinea pig, and high-quality neural signals were obtained.

ICPs have also been employed as platforms for cell adhesion and proliferation. Langer and coworkers^[54] showed that the adsorption of extracellular matrix molecules—namely fibronectin—is a critical step in facilitating cell attachment. When aortic endothelial cells were cultured, an oxidized polypyrrole coated with fibronectin cells spread normally and synthesized DNA as expected. When the polymer was reduced (Fig. 3), cell extension and DNA synthesis were inhibited, but cells remained viable. In other studies^[55] it was shown that the presence of vitronectin was critical for adhesion of endothelial cells to heparin-containing polypyrroles. With this conducting polymer platform the involvement of fibronectin was less critical.

In studies^[56] using polystyrene sulfonate as dopant in polypyrrole, the use of electrical stimulation was shown to enhance fibronectin adsorption on the polymer if it was stimulated immediately upon exposure to the protein-containing solutions. Subsequently, adhesion of nerve cells to the substrate and neurite and outgrowth were improved. However, if the electrical stimulation was delayed (e.g., applied after only a two-hour period of exposure to the Fn-containing solution), no increase in protein adsorption was observed.

The adhesion of PC-12 nerve cells to polypyrroles containing polyelectrolytes as dopants has also been demonstrated,^[57] and the incorporation of nerve growth factor and the electrically controlled release of this to trigger cell growth and differentiation has been demonstrated.^[58]

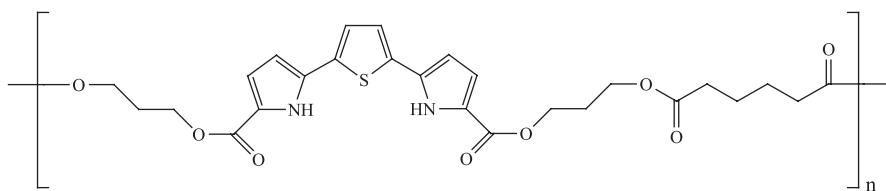
Langer and coworkers^[59] utilized polypyrrole platforms for PC-12 cell adhesion and proliferation. In-vitro studies revealed that neurite excursions from PC-12 cells were more pronounced than on tissue culture polystyrene, and that application of an electrical stimulus resulted in a significant increase in neurite length. In the same work,^[59] polypyrrole discs laminated with PLGA poly(lactic acid-co-glycolic acid) were also implanted into adult male Lewis rats in the subcutaneous and intramuscular regions. It was found that inflammation was not as severe as when PLGA alone was inserted. It was also found that the conducting polymer conduits facilitated sciatic nerve regeneration over a 10 nm gap, with no application of electrical stimulus and no apparent adverse effects observed.

Other workers^[60] have carried out in-vivo studies using polypyrrole tubes to facilitate regrowth of a severed (14 mm gap) sciatic nerve in rats. In other studies^[61] using PC-12 cells, it was shown that the medium nerve outgrowth was 18.1 μm with electrical stimuli applied and 9.5 μm for controls. These latter studies from

200% compared to nonpolymer coated titanium was observed.^[64]

Others^[65] have incorporated the nonapeptide (CDPGYIGSR) during electrodeposition of polypyrrole. This peptide is a laminin fragment proven to mediate in cell attachment and to promote neurite outgrowth. The study showed that neuroblastoma cells grew preferentially on the PPy-polypeptide materials.

Even a biodegradable ICP has recently been devised.^[66] The polymer consists of conducting oligomers of pyrrole and thiophene connected via degradable ester linkages (see Chemical Structure 9). An electrical conductivity on the order of $10^{-4} \text{ S cm}^{-1}$ was obtained. It was verified that the material was biodegradable and was subject to degradation by enzymes normally found in the body. Nerve cells adhered to the new polymer structures, and neurite outgrowth was observed after one day. In-vivo biocompatibility was determined by subcutaneous implantation into rats. Inflammation was mild and similar to that observed with PLGA at 29 days.



9

Langer's group also involved the use of a polyelectrolyte (polystyrene sulfonate) as dopant. Biologically derived polyelectrolytes (shown earlier in Chemical Structure 8, heparin) have also been included as dopants in ICPs used as platforms for cell growth.^[62] The advantage of such dopants is that they not only result in cellularly compatible electronic hydrogels, but also bring additional biofunctionality to the platform. For example, heparin is a powerful anticoagulant and is a component of the extracellular matrix of blood vessels promoting endothelial cell growth. In these studies the growth of endothelial cells was indeed sustained on the conducting polymer composites. The effect of the dopant on cell densities observed during growth was obvious. Accessibility to the incorporated heparin moiety was also shown to be dependent on the electrical stimuli applied to the conducting polymer composite. (Composites remained stable after autoclaving to sterilize.)

Protocols to enable covalent attachment of polypeptides to polypyrrole backbones have been developed.^[63] These techniques were used to attach the polypeptide Cys-Gly-(Arg-Gly-Asp)-Ser-Pro-Lys to polypyrroles coated on titanium or platinum. This polypeptide (RGD) contains the cell adhesive region of fibronectin. Increase in osteoblast cell adhesion on the order of

Zhang and coworkers^[67] recently investigated the biocompatibility of PPy-coated polyester fabrics for possible use as vascular prostheses. The fabrics were coated in two ways: 1) using phosphorylation to bind pyrrole groups to the polyester surface, with subsequent gas-phase polymerization; or 2) through plasma-activated polymerization of pyrrole. No adverse effect on hemolysis or coagulation time was observed. Low cell adhesion for endothelial cells was found for both coated textiles at 24 hours, but no cytotoxicity was detected at 72 or 168 hours. Further confirming biocompatibility the fabrics did not activate polymer phagocytic cells (PMN) (i.e., inflammatory cells).

CONCLUSION

Inherently conducting polymers—polypyrrole in particular—provide an organic environment capable of efficient charge transfer that is compatible with biological systems. The incorporation of enzymes, antibodies, and even whole living cells during polymerization (while retaining molecular biological function) verifies the high degree of compatibility. Further evidence is provided by the ability to culture mammalian cells on polypyrrole

surfaces. The exciting possibilities of controlling cell adhesion and proliferation and the direction of neurite outgrowth have been demonstrated.

The mechanisms underpinning such control are yet to be fully elucidated. Once understood, they will provide a sound basis for the design of new biomaterials based on ICPs. In parallel, the rapid advances in ICP processing and device fabrication are particularly pertinent to the biomaterials engineering field. The ever-expanding ability to generate ICPs in aqueous solution or as aqueous dispersions that can be used to fabricate structures of different sizes, shapes, and configurations is important for the biomaterials engineer.

Such solutions and dispersions can be used to create patterned two-dimensional surfaces using traditional methods such as ink-jet printing, or three-dimensional structures using electrospinning techniques. The development of protocols that result in ordered structures at the nanodomain will also impact the biomaterials field.^[68] This is a rapidly developing field, in which strategies to enable synthesis of nanocomponents of controlled dimensions and the assembly of nanodimensional structures are being developed at an accelerating rate.^[69]

ARTICLES OF FURTHER INTEREST

Artificial Muscles; Electroactive Polymeric Materials; Tissue Engineering of Peripheral Nerve

REFERENCES

- Lavine, L.S.; Grodzinsky, A.J. Current concepts review electrical stimulation of repair of bone. *J. Bone Jt. Surg.* **1987**, *69-A* (4), 626–630.
- Karba, R.; Semrov, D.; Vodovnik, L.; Benko, H.; Savrin, R. DC electrical stimulation for chronic wound healing enhancement. Part 1. Clinical study and determination of electrical field distribution in the numerical wound model. *Bioelectrochem. Bioenerg.* **1997**, *43*, 265–270.
- Kerns, J.M.; Fakhouri, A.J.; Weinrib, H.P.; Freeman, J.A. Electrical stimulation of nerve regeneration in the rat: The early effects evaluated by a vibrating probe and electron microscopy. *Neuroscience* **1991**, *40* (1), 93–107.
- Skotheim, T.A.; Elsenbaumer, R.L.; Reynolds, J.R. *Handbook of Conducting Polymers*, 2nd Ed.; Marcel Dekker: New York, 1998.
- Wallace, G.G.; Teasdale, P.R.; Spinks, G. *Conductive electroactive polymers: Intelligent material systems*; Technomic Publ. Co.: Lancaster, 1997.
- Krische, B.; Zagorska, M. The polythiophene paradox. *Synth. Met.* **1989**, *28* (1–2), C263–C268.
- Wei, Y.; Chan, C.C.; Tian, J.; Jang, G.W.; Hsueh, K.F. Electrochemical polymerization of thiophenes in the presence of bithiophene or terthiophene: Kinetics and mechanism of the polymerization. *Chem. Mater.* **1991**, *3* (5), 888–897.
- Huang, W.S.; Humphrey, B.D.; MacDiarmid, A.G. Polyaniline, a novel conducting polymer. Morphology and chemistry of its oxidation and reduction in aqueous electrolytes. *J. Chem. Soc., Faraday Trans* **1986**, *82* (8), 2385–2400, and references cited therein
- Aboutanos, V.; Barisci, J.N.; Innis, P.C.; Wallace, G.G. Factors affecting the electrochemical formation of polypyrrole-nitrate colloids. *Colloids Surf., A* **1998**, *137* (1–3), 295–300.
- Adams, P.N.; Monkman, A.P. Characterization of high molecular weight polyaniline synthesized at -40°C using a 0.25:1 mole ratio of persulfate oxidant to aniline. *Synth. Met.* **1997**, *87* (2), 165–169, and references cited therein.
- Norris, I.D.; Kane-Maguire, L.A.P.; Wallace, G.G. Thermochromism in optically active polyaniline salts. *Macromolecules* **1998**, *31* (19), 6529–6533.
- Kern, J.-M.; Sauvage, J.-P. Photochemical deposition of electrically conducting polypyrrole. *J. Chem. Soc., Chem. Commun.* **1989**, 657–658.
- Teshima, K.; Uemura, S.; Kobayashi, N.; Hirohashi, R. Effect of pH on photopolymerization reaction of aniline derivatives with the tris(2,2'-bipyridyl)ruthenium complex and the methylviologen system. *Macromolecules* **1998**, *31*, 6783–6788, and references cited therein.
- Kane-Maguire, L.A.P.; Pirkle, B.; Wallace, G.G.; Watson, R.; Dozier, T.; Kane-Maguire, N.A.P. In *Photochemical synthesis of polyaniline and substituted polyanilines*, 11th RACI Convention. Symposium on Functional Polymers and Electrochemistry, Canberra, ACT, Feb. 6–11, 2000.
- Aizawa, M.; Wang, L. *Enzyme-catalyzed oxidative polymerization (aromatic compounds)*; Salamone, J.C., Ed.; Polymer Materials Encyclopedia, CRC Press: Boca Raton, USA, 1996; Vol. 3; 2107–2115.
- Liu, W.; Kumar, J.; Tripathy, S.; Senecal, K.J.; Samuelson, L. Enzymatically synthesized conducting polyaniline. *J. Am. Chem. Soc.* **1999**, *121*, 71–78.
- Nagarajan, R.; Liu, W.; Kumar, J.; Tripathy, S.K.; Bruno, F.F.; Samuelson, L.A. Manipulating DNA conformation using intertwined conducting polymer chains. *Macromolecules* **2001**, *34*, 3921–3927.
- Higgins, S.J. Conjugated polymers incorporating pendant functional groups-synthesis and characterization. *Chem. Soc. Rev.* **1997**, *26* (4), 247–258.
- Delabouglise, D.; Garnier, F. Chiral metals: Amino acid-substituted conducting polypyrroles. *Synth. Met.* **1990**, *39*, 117–120.
- Moutet, J.-C.; Saint-Aman, E.; Tran-Van, F.; Angibeaud, P.; Utille, J.-P. Poly(glucose-pyrrole) modified electrodes: A novel chiral electrode for enantioselective recognition. *Adv. Mat.* **1992**, *4* (7/8), 511–513.
- Korri-Youssoufi, H.; Godillot, P.; Srivastava, P.; El Kassmi, A.; Garnier, F. New method of polypyrrole functionalization toward molecular recognition. *Synth. Met.* **1997**, *84* (1–3), 169–170.
- Livache, T.; Roget, A.; Dejean, E.; Barthet, C.; Bidan, G.; Teoule, R. Preparation of a DNA matrix via an electrochemically directed copolymerization of pyrrole

- and oligonucleotides bearing a pyrrole group. *Nucleic Acids Res.* **1994**, *22* (15), 2915–2921.
23. Bidan, G.; Billon, M.; Galasso, K.; Livache, T.; Mathis, G.; Roget, A.; Torres-Rodriguez, L.M.; Vieil, E. Electropolymerization as a versatile route for immobilizing biological species onto surfaces: Application to DNA biochips. *Appl. Biochem. Biotechnol.* **2000**, *89* (2–3), 183–193.
 24. Lassalle, N.; Mailley, P.; Vieil, E.; Livache, T.; Roget, A.; Correia, J.P.; Abrantes, L.M. Electronically conductive polymer grafted with oligonucleotides as electrochemical sensors of DNA. Preliminary study of real time monitoring by in situ techniques. *Electroanal. Chem.* **2001**, *509* (1), 48–57.
 25. Korri-Youssoufi, H.; Richard, C.; Yassar, A. A new method for the immobilisation of antibodies in conducting polymers. *Mater. Sci. Eng., C* **2001**, *15*, 307–310.
 26. Kojima, K.; Unuma, T.; Yamauchi, T.; Shimomura, M.; Miyauchi, S. Preparation of polypyrrole covalently attached with glucose oxidase and its application to glucose sensing. *Synth. Met.* **1997**, *85*, 1417–1418.
 27. Shimomura, M.; Kojima, N.; Oshima, K.; Yamauchi, T.; Miyauchi, S. Covalent immobilization of glucose oxidase on film prepared by electrochemical copolymerization of thiophene-3-acetic acid and 3-methylthiophene for glucose sensing. *Polym. J.* **2001**, *33* (8), 629–631.
 28. Kros, A.; Nolte, R.J.M.; Sommerdijk, N.A.J.M. Poly(3,4-ethylenedioxythiophene)-based copolymers for biosensor applications. *J. Polym. Sci., Part A, Polym. Chem.* **2002**, *40*, 738–747.
 29. Ouerghi, O.; Touhami, A.; Jaffrezic-Renault, N.; Martelet, C.; Ben Ouada, H.; Cosnier, S. Impedimetric immunosensor using avidin-biotin for antibody immobilization. *Bioelectrochemistry* **2002**, *56*, 131–133.
 30. Kumpumbu-Kalemba, L.; Leclerc, M. Electrochemical characterization of monolayers of a biotinylated polythiophene: Towards the development of polymeric biosensors. *J. Chem. Soc., Chem. Commun.* **2000**, 1847–1848.
 31. Pyo, M.; Reynolds, J.R. Electrochemically stimulated adenosine 5'-triphosphate (ATP) release through redox switching of conducting polypyrrole films and bilayers. *Chem. Mater.* **1996**, *8*, 128–133.
 32. Umana, M.; Waller, J. Protein-modified electrodes. The glucose oxidase/polypyrrole system. *Anal. Chem.* **1986**, *58*, 2979–2983.
 33. Belanger, D.; Nadreau, J.; Fortier, G. Electrochemistry of the polypyrrole glucose oxidase electrode. *J. Electroanal. Chem.* **1989**, *274*, 143–155.
 34. Adeloju, S.B.; Barisci, J.N.; Wallace, G.G. Electroimmobilisation of sulfite oxidase into a polypyrrole film and its utilisation for flow amperometric detection of sulphite. *Anal. Chim. Acta* **1996**, *332*, 145–153.
 35. Adeloju, S.B.; Shaw, S.J.; Wallace, G.G. Pulsed-amperometric detection of urea in blood samples on a conducting polypyrrole urease biosensor. *Anal. Chim. Acta* **1997**, *341*, 155–160.
 36. Mousty, C.; Galland, B.; Cosnier, S. Electrogeneration of a hydrophilic cross-linked polypyrrole film for enzyme electrode fabrication. Application to the amperometric detection of glucose. *Electroanalysis* **2001**, *13* (3), 186–190.
 37. Sadik, O.; Wallace, G.G. Pulsed amperometric detection of proteins using antibody containing conducting polymers. *Anal. Chim. Acta* **1993**, *279*, 209–212.
 38. Barnett, D.; Sadik, O.A.; John, M.J.; Wallace, G.G. Pulsed amperometric detection of thaumatin using antibody-containing poly(pyrrole) electrodes. *Analyst* **1994**, *119*, 1997–2000.
 39. Petit, M.A.; Zante, D.; Colin, C.; Combellas, C.; Hennion, M.-C. Evaluation of a polypyrrole film containing anti-isoproturon antibodies for the detection of isoproturon. *Electroanalysis* **2001**, *13* (10), 856–859.
 40. Misoska, V.; Price, W.E.; Ralph, S.F.; Ogata, N.; Wallace, G.G. Synthesis, characterisation and ion transport studies on polypyrrole/deoxyribonucleic acid conducting polymer membranes. *Synth. Met.* **2001**, *123*, 279–286.
 41. Zhou, D.; Too, C.O.; Wallace, G.G. Synthesis and characterization of polypyrrole/heparin composites. *React. Funct. Polym.* **1999**, *39*, 19–26.
 42. Wang, J.; Jiang, M.; Fortes, A.; Mukherjee, B. New label-free DNA recognition based on doping nucleic acid probes within conducting polymer films. *Anal. Chim. Acta* **1999**, *402* (1–2), 7–12.
 43. Khor, E.; Whey, J.L.H. Interaction of chitosan with polypyrrole in the formation of hybrid biomaterials. *Carbohydr. Polym.* **1995**, *26*, 183–187.
 44. Yang, S.; Tirmizi, S.A.; Burns, A.; Barney, A.A.; Risen, W.M. Chitiline materials: Soluble chitosan-polyaniline copolymers and their conductive doped forms. *Synth. Met.* **1989**, *32*, 191–200.
 45. Li, H.C.; Khor, E. Interaction of collagen with polypyrrole in the production of hybrid materials. *Polym. Int.* **1994**, *35*, 53–59.
 46. Chze, H.; Khor, E. A collagen-polypyrrole hybrid: Influence of 3-butanedisulfonate substitution. *Macromol. Chem. Phys.* **1995**, *196*, 1801–1812.
 47. Smith, A.B.; Knowles, C.J. Potential role of a conducting polymer in biochemistry: Protein-binding properties. *Biotechnol. Appl. Biochem.* **1990**, *12*, 661–669.
 48. Khan, G.F.; Wernet, W. Adsorption of proteins on electro-conductive polymer films. *Thin Solid Films* **1997**, *300*, 265–271.
 49. Deshpande, M.V.; Hall, E.A.H. An electrochemically grown polymer as an immobilisation matrix for whole cells: Application in an amperometric dopamine sensor. *Biosens. Bioelectron.* **1990**, *5*, 431–448.
 50. Aizawa, M.; Haruyama, T.; Khan, G.F.; Kobatake, E.; Ikariyama, Y. Electronically modulated biological functions of molecular interfaced enzymes and living cells. *Biosens. Bioelectron.* **1994**, *9*, 601–610.
 51. Palmqvist, E.; Kriz, C.B.; Khayyami, M.; Danielsson, B.; Larsson, P.-O.; Mosbach, K.; Kriz, D. Development of a simple detector for microbial metabolism, based on a polypyrrole dc resistometric device. *Biosens. Bioelectron.* **1994**, *9*, 551–556.
 52. Campbell, T.E.; Hodgson, A.J.; Wallace, G.G. Incorporation of erythrocytes into polypyrrole to form the basis of a biosensor to screen for rhesus (D) blood groups and rhesus (D) antibodies. *Electroanalysis* **1999**, *11* (4), 215–222.



53. Cui, X.; Hetke, J.F.; Wiler, J.A.; Anderson, D.J.; Martin, D.C. Electrochemical deposition and characterization of conducting polymer polypyrrole/PSS on multichannel neural probes. *Sens. Actuators, A* **2001**, *93*, 8–18.
54. Wong, J.Y.; Langer, R.; Ingber, D.E. Electrically conducting polymers can noninvasively control the shape and growth of mammalian cells. *Proc. Natl. Acad. Sci.* **1994**, *91*, 3201–3204.
55. Garner, B.; Hodgson, A.J.; Wallace, G.G.; Underwood, P.A. Human endothelial cell attachment to and growth on polypyrrole-heparin is vitronectin dependent. *J. Mater. Sci., Mater. Med.* **1999**, *10*, 19–27.
56. Kotwal, A.; Schmidt, C.E. Electrical stimulation alters protein adsorption and nerve cell interactions with electrically conducting biomaterials. *Biomaterials* **2001**, *22*, 1055–1064.
57. Hodgson, A.J.; Gilmore, K.; Small, C.; Wallace, G.G.; MacKenzie, I.L.; Aoki, T.; Ogata, N. Reactive supramolecular assemblies of mucopolysaccharide, polypyrrole and protein as controllable biocomposites for a new generation of 'intelligent biomaterials'. *Supramol. Sci.* **1994**, *1*, 77–83.
58. Hodgson, A.J.; John, M.J.; Campbell, T.; Georgevich, A.; Woodhouse, S.; Aoki, T.; Ogata, N.; Wallace, G.G. Integration of biocomponents with synthetic structures- use of conducting polymer polyelectrolyte composites. *SPIE Int. Soc. Opt. Eng.* **1996**, *2716*, 164–176.
59. Schmidt, C.E.; Shastri, V.R.; Vacanti, J.P.; Langer, R. Stimulation of neurite outgrowth using an electrically conducting polymer. *Proc. Natl. Acad. Sci.* **1997**, *94*, 8948–8953.
60. Chen, S.J.; Wang, D.Y.; Yuan, C.W.; Wang, X.D.; Zhang, P.Y.; Gu, X.S. Template synthesis of the polypyrrole tube and its bridging in vivo sciatic nerve regeneration. *Mater. Sci. Lett.* **2000**, *19*, 2157–2159.
61. Shastri, V.R.; Schmidt, C.E.; Kim, T.H.; Vacanti, J.P.; Langer, R. Polypyrrole—a potential candidate for stimulated nerve regeneration. *Mater. Res. Soc. Symp. Proc.* **1996**, *414*, 113–118.
62. Garner, B.; Georgevich, A.; Liu, L.; Hodgson, A.J.; Wallace, G.G. Polypyrrole-heparin composites as stimulus-responsive substrates for endothelial cell growth. *J. Biomed. Mater. Res.* **1999**, *44*, 121–129.
63. De Giglio, E.; Sabbatini, L.; Zambonin, P.G. Development and analytical characterization of cystein-grafted polypyrrole films electrosynthesized on Pt-and Ti-substrates as precursors of bioactive interfaces. *Biomater. Sci. Polym. Ed.* **1999**, *10* (8), 845–858.
64. De Giglio, E.; Sabbatini, L.; Colucci, S.; Zambonin, G. Synthesis, analytical characterization, and osteoblast adhesion properties on RGD-grafted polypyrrole coatings on titanium substrates. *Biomater. Sci. Polym. Ed.* **2000**, *11* (10), 1073–1083.
65. Cui, X.; Lee, V.A.; Raphael, Y.; Wiler, J.A.; Hetke, J.F.; Anderson, D.J.; Martin, D.C. Surface modification of neural recording electrodes with conducting polymer/biomolecule blends. *Biomed. Mater. Res.* **2001**, *56*, 261–272.
66. Rivers, T.J.; Hudson, T.W.; Schmidt, C.E. Synthesis of a novel, biodegradable electrically conducting polymer for biomedical applications. *Adv. Funct. Mater.* **2002**, *12* (1), 33–37.
67. Zhang, Z.; Roy, R.; Dugré, F.J.; Tessier, D.; Dao, L.H. In vitro biocompatibility study of electrically conductive polypyrrole-coated polyester fabrics. *Biomed. Mater. Res.* **2001**, *57* (1), 63–71.
68. Boeckl, M.S.; Baas, T.; Fujita, A.; Hwang, K.-O.; Bramblett, A.L.; Ratner, B.D.; Rogers, J.W.; Sasaki, T. Template-assisted nano-patterning of solid surfaces. *Biopoly* **1998**, *47*, 185–193.
69. Innis, P.C.; Wallace, G.G. Inherently conducting polymer nanostructures. *J. Nanosci. Nanotechnol. in press.*

Confocal Microscopy

Denis Semwogerere

Eric R. Weeks

Department of Physics, Emory University, Atlanta, Georgia, U.S.A.



INTRODUCTION

A confocal microscope creates sharp images of a specimen that would otherwise appear blurred when viewed with a conventional microscope. This is achieved by excluding most of the light from the specimen that is not from the microscope's focal plane. The image has less haze and better contrast than that of a conventional microscope and represents a thin cross-section of the specimen. Thus, apart from allowing better observation of fine details it is possible to build three-dimensional (3D) reconstructions of a volume of the specimen by assembling a series of thin slices taken along the vertical axis.

BACKGROUND

Confocal microscopy was pioneered by Marvin Minsky in 1955 while he was a Junior Fellow at Harvard University.^[1] Minsky's invention would perform a point-by-point image construction by focusing a point of light sequentially across a specimen and then collecting some of the returning rays. By illuminating a single point at a time Minsky avoided most of the unwanted scattered light that obscures an image when the entire specimen is illuminated at the same time. Additionally, the light returning from the specimen would pass through a second pinhole aperture that would reject rays that were not directly from the focal point. The remaining "desirable" light rays would then be collected by a photomultiplier and the image gradually reconstructed using a long-persistence screen. To build the image, Minsky scanned the specimen by moving the stage rather than the light rays. This was to avoid the challenge of trying to maintain sensitive alignment of moving optics. Using a 60 Hz solenoid to move the platform vertically and a lower-frequency solenoid to move it horizontally, Minsky managed to obtain a frame rate of approximately one image every 10 sec.

MODERN CONFOCAL MICROSCOPY

Modern confocal microscopes have kept the key elements of Minsky's design: the pinhole apertures and

point-by-point illumination of the specimen. Advances in optics and electronics have been incorporated into current designs and provide improvements in speed, image quality, and storage of the generated images. Although there are a number of different confocal microscope designs, this entry will discuss one general type—the other designs are not markedly different.^[2]

The majority of confocal microscopes image either by reflecting light off the specimen or by stimulating fluorescence from dyes (fluorophores) applied to the specimen. The focus of this entry will be on fluorescence confocal microscopy as it is the mode that is most commonly used in biological applications. The difference between the two techniques is small. There are methods that involve transmission of light through the specimen, but these are much less common.^[3]

Fluorescence

If light is incident on a molecule, it may absorb the light and then emit light of a different color, a process known as fluorescence. At ordinary temperatures most molecules are in their lowest energy state, the ground state. However, they may absorb a photon of light (for example, blue light) that increases their energy causing an electron to jump to a discrete singlet excited state.^[4] In Fig. 1, this is represented by the top black line. Typically, the molecule quickly (within 10^{-8} sec) dissipates some of the absorbed energy through collisions with surrounding molecules causing the electron to drop to a lower energy level (the second black line). If the surrounding molecules are not able to accept the larger energy difference needed to further lower the molecule to its ground state, it may undergo spontaneous emission, thereby losing the remaining energy, by emitting light of a longer wavelength (for example, green light).^[5] Fluorescein is a common fluorophore that acts this way, emitting green light when stimulated with blue excitation light. The wavelengths of the excitation light and the color of the emitted light are material dependent.

Microscopy in the fluorescence mode has several advantages over the reflected or transmitted modes. It can be more sensitive. Often, it is possible to attach fluorescent molecules to specific parts of the specimen,

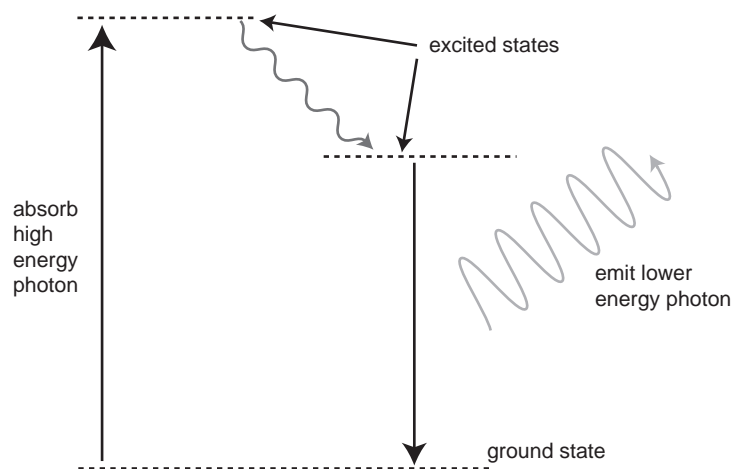


Fig. 1 Mechanism of fluorescence. The horizontal lines indicate quantum energy levels of the molecule. A fluorescent dye molecule is raised to an excited energy state by a high-energy photon. It loses a little energy to other molecules and drops to a lower excited state. It loses the rest of the energy by emitting light of a lower energy.

making them the only visible ones in the microscope and it is also possible to use more than one type of fluorophore.^[6] Thus, by switching the excitation light different parts of the specimen can be distinguished.

FLUORESCENCE MICROSCOPY

In conventional fluorescence microscopy a dyed specimen is illuminated with light of an appropriate wavelength and an image is formed from the resulting fluorescent light. In Fig. 2 the excitation light is blue and the emitted light is green. The microscope uses a dichroic mirror (also called a “dichromatic mirror”) that reflects light shorter than a certain wavelength but transmits light of longer wavelength. Thus, the light from the main source is reflected and passes through the objective to the sample, while the

longer-wavelength light from the fluorescing specimen passes through both the objective and the dichroic mirror. This particular type of fluorescence microscopy, in which the objective used by the illuminating light is also used by the fluorescing light in conjunction with a dichroic mirror, is called epifluorescence. In the case of reflected light microscopy, a beamsplitter is used in place of the dichroic mirror.

CONFOCAL MICROSCOPY

To understand confocal microscopy it is instructive to imagine a pair of lenses that focuses light from the focal point of one lens to the focal point of the other. This is illustrated by the dark blue rays in Fig. 3. The light blue rays represent light from another point in the specimen, which is not at the focal point of the

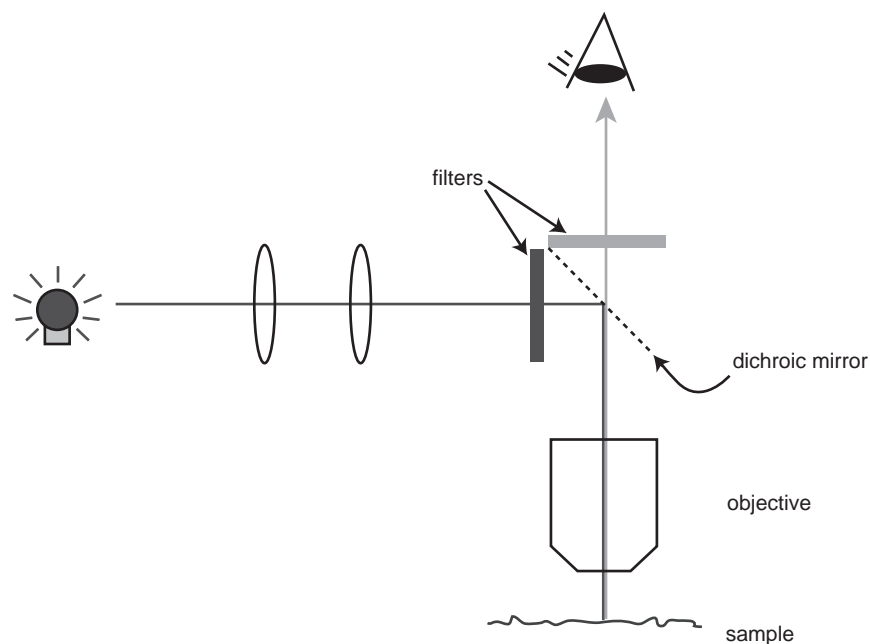


Fig. 2 Basic setup of a fluorescence microscope. Light from the source is reflected off the dichroic mirror toward the specimen. Returning fluorescence of a longer wavelength is allowed to pass through the dichroic mirror to the eyepiece.

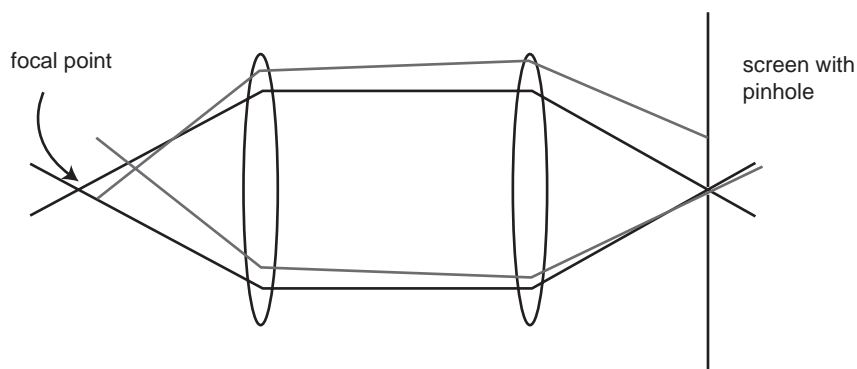


Fig. 3 Rejection of light not incident from the focal plane. All light from the focal point that reaches the screen is allowed through. Light away from the focal point is mostly rejected.



left-hand-side lens. (Note that the colors of the rays are purely for purposes of distinguishing the two sets—they do not represent different wavelengths of light.) Clearly, the image of the light blue point is not at the same location as the image of the dark blue point. (Recall from introductory optics that points do not need to be at the focal point of the lens for the system of lenses to form an image.)

In confocal microscopy, the aim is to see only the image of the dark blue point.^[1] Accordingly, if a screen with a pinhole is placed at the other side of the lens system, then all of the light from the dark point will pass through the pinhole.^a Note that at the location of the screen the light blue point is out of focus. Moreover, most of the light will get blocked by the screen, resulting in an image of the light blue point that is significantly attenuated compared to the image of the dark blue point.

To further reduce the amount of light emanating from “light blue” points, the confocal microscope setup minimizes how much of the specimen is illuminated. Normally, in fluorescence microscopy the entire field of view of the specimen is completely illuminated, making the whole region fluoresce at the same time. Of course, the highest intensity of the excitation light is at the focal point of the lens, but the other parts of the specimen do get some of this light and they do fluoresce. Thus, light at a “dark blue” point may include light that has been scattered from other “light blue” points, thereby obscuring its fluorescence. To reduce this effect the confocal microscope focuses a point of light at the in-focus dark blue point by imaging a pinhole aperture placed in front of the light source.^[1] Thus, the only regions that are illuminated are a cone of light above and below the focal (dark blue) point (Fig. 9A).

Together the confocal microscope’s two pinholes significantly reduce the background haze that is typical

^aActually not all the light from the focal point reaches the pinhole. Some of it is reflected and absorbed by the optics in between.

of a conventional fluorescence image, as shown in Fig. 5. Because the focal point of the objective lens forms an image where the pinhole/screen is, those two points are known as “conjugate points” (or alternatively, the specimen plane and the pinhole/screen are conjugate planes). The pinhole is conjugate to the focal point of the lens, hence the name “confocal” pinhole.

HOW DOES A CONFOCAL MICROSCOPE WORK?

The confocal microscope incorporates the ideas of point-by-point illumination of the specimen and rejection of out-of-focus light.

One drawback with imaging a point onto the specimen is that there are fewer emitted photons to collect at any given instant. Thus, to avoid building a noisy image each point must be illuminated for a long time to collect enough light to make an accurate measurement.^[1] In turn, this increases the length of time needed to create a point-by-point image. The solution is to use a light source of very high intensity, which Minsky did with a zirconium arc lamp. The modern choice is a laser light source, which has the additional benefit of being available in a wide range of wavelengths.

In Fig. 4A the laser provides the intense blue excitation light. The light reflects off a dichroic mirror, which directs it to an assembly of vertically and horizontally scanning mirrors. These motor-driven mirrors scan the laser across the specimen. Recall that Minsky’s invention kept the optics stationary and instead scanned the specimen by moving the stage back and forth in the vertical and horizontal directions. As awkward (and slow) as that method seems it does have among others the following two major advantages:^[7]

- The specimen is everywhere illuminated axially, rather than at different angles as in the case of the scanning mirror configuration, thereby avoiding optical aberrations. Thus, the entire field of view is illuminated uniformly.

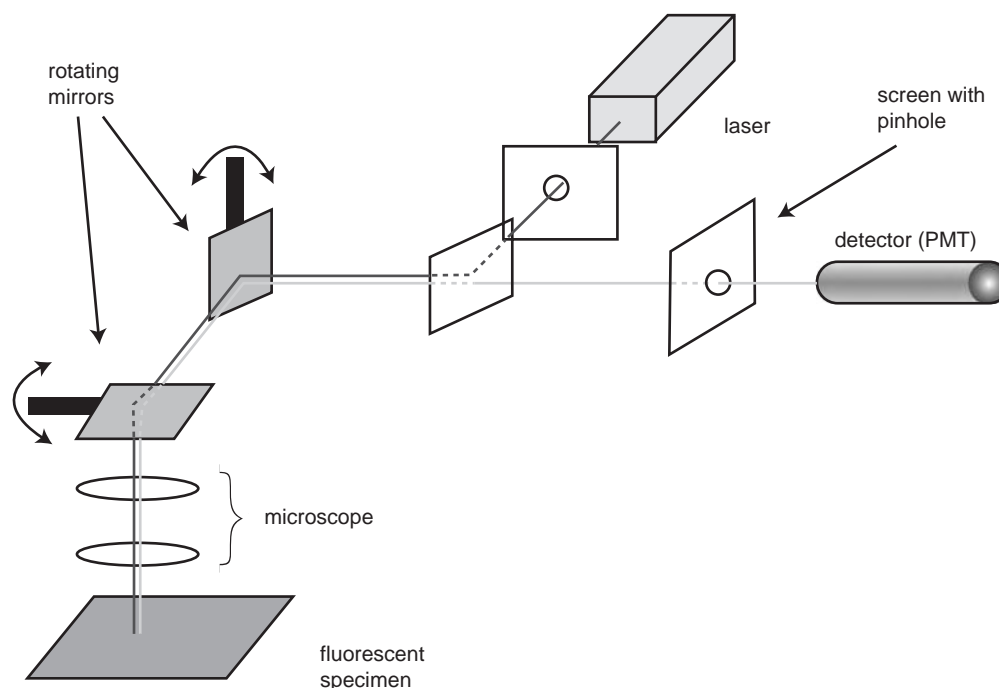


Fig. 4 Basic setup of a confocal microscope. Light from the laser is scanned across the specimen by the scanning mirrors. Optical sectioning occurs as the light passes through a pinhole on its way to the detector.

- The field of view can be made larger than that of the static objective by controlling the amplitude of the stage movements.

In Fig. 4 the dye in the specimen is excited by the laser light and fluoresces. The fluorescent (green) light is descanned by the same mirrors that are used to scan the excitation light (blue) from the laser and then passes through the dichroic mirror. Thereafter, it is focused onto the pinhole. The light that makes it through the pinhole is measured by a detector such as a photomultiplier tube.

In confocal microscopy, there is never a complete image of the specimen because at any instant only one point is observed. Thus, for visualization the detector is attached to a computer, which builds up the image

one pixel at a time. For a 512×512 -pixel image this is typically done at a frame rate of 0.1–30 Hz. The large range in frame rates depends on a number of factors, the most important of which will be discussed below.

The image created by the confocal microscope is of a thin planar region of the specimen—an effect referred to as optical sectioning. Out-of-plane unfocused light has been rejected, resulting in a sharper, better-resolved image. Fig. 5 shows an image created with and without optical sectioning.

THREE-DIMENSIONAL VISUALIZATION

The ability of a confocal microscope to create sharp optical sections makes it possible to build 3D

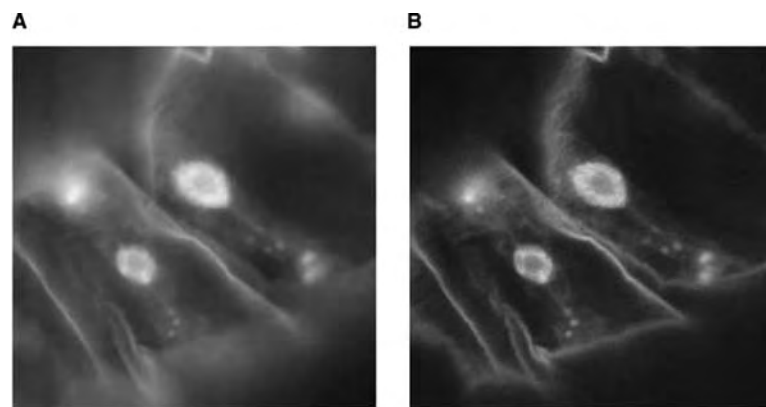


Fig. 5 Images of cells of spirogyra generated with and without optical sectioning. The image in (B) was created using a slit rather than a pinhole for out-of-focus light rejection. Most of the haze associated with the cell walls of the filamentous algae is absent, allowing clearer distinction of the different parts.

renditions of the specimen. Data gathered from a series of optical sections imaged at short and regular intervals along the optical axis are used to create the 3D reconstruction. Software can combine the 2D images to create a 3D rendition. Representing 3D information in a meaningful way out of 2D data is nontrivial, and a number of different schemes have been developed. Fig. 6 shows a 3D reconstruction, from slices of a suspension of 2 μm diameter colloidal particles using “alpha blending”—a technique that combines images by first making each of their individual pixels less or more transparent according to a computed weight called the “alpha” value.^[8] The result is a 3D-like structure.

OTHER CONSIDERATIONS

A confocal microscope, as with every instrument, has some limitations and often compromises must be made to optimize performance. The following is an outline of some of the most important of them.

Resolution

As with conventional microscopy, confocal microscopy has inherent resolution limitations due to diffraction. In the discussion above it is assumed that the point source used produces a point of light on the specimen. In fact it appears in the focal plane as an Airy disk, whose size depends on the wavelength of the light source and the numerical aperture of the objective lens.^[9] (The numerical aperture of a lens is a measure of how well it gathers light.) The graph of Fig. 7B shows a plot of the intensity of light as a function of

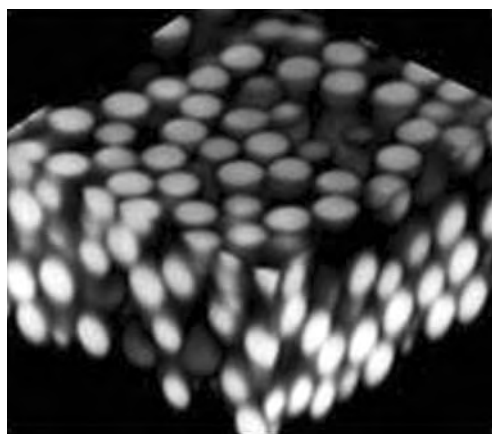


Fig. 6 Three-dimensional reconstruction of a series of 2D images of PMMA spheres suspended in a cyclohexylbromide and decalin solution. The image was created using “alpha blending.”

radius of an Airy disk—the image is circularly symmetric, as shown in Fig. 7A.

The Airy disk limits the maximum resolution that can be attained with the confocal microscope—the best resolution is typically about 200 nm. Ideally, the image of a point would just be a single intense point right at radius = 0. However, the finite size of the Airy disk sets the scale for which details can be resolved. According to the Rayleigh criterion, the minimum separation between two Airy disks for which they are distinguishable is equal to their radius. This corresponds to the maximum of one Airy disk superimposed on the minimum of the other.

Resolution along the optical axis is also limited by diffraction effects. As in the lateral direction there is a periodic, but elliptical distribution of intensity in the shape of an Airy disk.^[10]

Pinhole Size

The optical sectioning capability of a confocal microscope derives from having a pinhole to reject out-of-focus light rays. In turn, the strength of the optical sectioning (the rate at which the detected intensity drops off in the axial direction) depends strongly on the size of the pinhole.^[11]

It is tempting to assume that making the pinhole as small as possible is the best choice. However, as the pinhole size is reduced, so too are the number of photons that arrive at the detector from the specimen. This may lead to a reduced signal-to-noise ratio. To offset the weaker signal more fluorescence is needed from the specimen. This usually can be done, to a limit, by raising the intensity of the excitation light. But high intensities can damage the specimen, and in the case of fluorescence, also degrade the fluorophore. Moreover, it has been shown that optical sectioning does not improve considerably with the pinhole size below a limit that approximates the radius of the first zero of the Airy disk.^[2,11] (Note that the study considered imaging in coherent mode; however, the result still qualitatively applies to inherently incoherent confocal fluorescence microscopy.) Thus, a good approximation is to make the pinhole about the size of the Airy disk.

Intensity of Incident Light

An important component of a confocal microscope is the photodetector that captures light from the specimen. In confocal fluorescence imaging the pinhole along with the optics preceding it significantly reduce the intensity of the emission that reaches the detector. Thus, the detector’s sensitivity and noise behavior are vitally important.^[3] The sensitivity is characterized by the quantum efficiency, which, as in any measurement



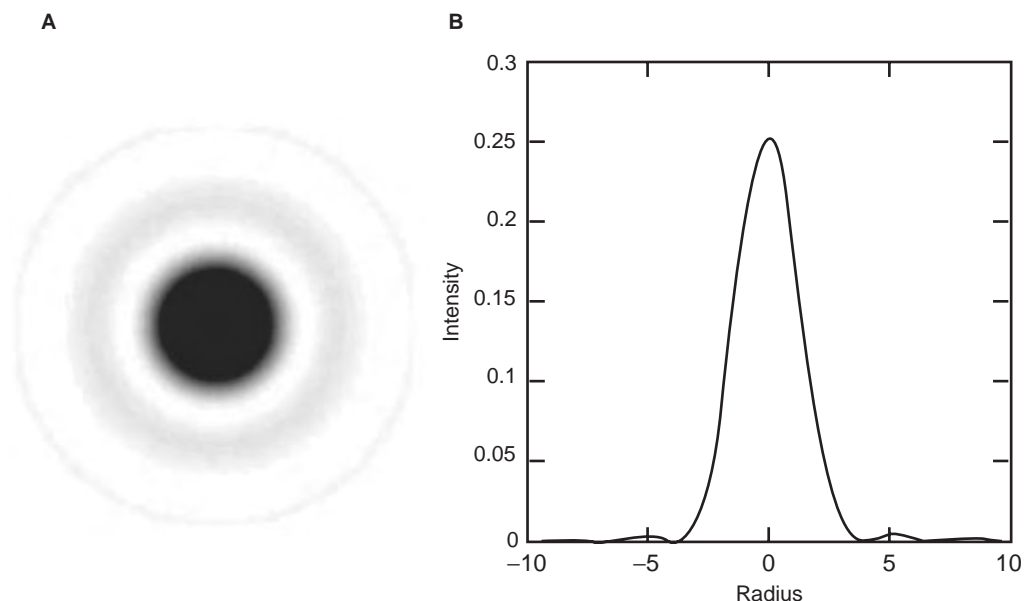


Fig. 7 Airy disk similar to that of an image of a very small particle. (A) The image is “overexposed” and in reverse color to allow distinction of the faint secondary peak. (B) A graph of the intensity change with radius.

involving quantum interactions, is limited by Poisson statistics.^[12] That is, the accuracy of the measurement is improved by increasing the number of photons arriving at the detector. In practical terms this can be done by averaging data from many frames—which has the drawback of slowing down the effective frame rate of the microscope. Or, it can be done by increasing the intensity of the fluorescence signal.

Fluorescence can be increased by dyeing the specimen with a larger concentration of fluorophore molecules or by raising the intensity of the excitation light. However, each of these methods increases excitation up to some limit. For high fluorophore concentrations, the individual molecules can quench each other.^[5] They may also reduce the amount of fluorescence deep inside the specimen—fluorophores nearest the light source can absorb enough light to significantly reduce the portion available to the rest of the specimen.^[3] Increasing fluorescence by increasing the excitation light intensity leads eventually to saturation of the fluorophore. Higher intensities drive a larger fraction of fluorophore molecules into excited states, which in turn leads to a smaller fraction of ground state molecules. Ultimately, the rate at which fluorophores are excited matches their decay rate, causing ground state depopulation. Fluorescence then ceases to increase with excitation intensity.^[13]

Fluorophores

Among the most important aspects of fluorescence confocal microscopy is the choice of fluorophore. It is

typically influenced by several factors. The fluorophore should tag the correct part of the specimen. It must be sensitive enough for the given excitation wavelength. For living specimens it should not significantly alter the dynamics of the organism; and an extra consideration is the effect of the specimen on the fluorophore—its chemical environment can affect the position of the peaks of the excitation and emission spectra.^[14]

Photobleaching

A major problem with fluorophores is that they fade (irreversibly) when exposed to excitation light (Fig. 8). Although this process is not completely understood, it is believed in some instances to occur when fluorophore molecules react with oxygen and/or oxygen radicals and become nonfluorescent.^[13,15] The reaction can take place after a fluorophore molecule transitions from the singlet excited state to the triplet excited state. Although the fraction of fluorophores that transitions to the triplet state is small, its lifetime is typically much longer than that of the singlet state. This can lead to a significant triplet state fluorophore population and thus to significant photobleaching.^[5]

Several strategies have been developed to reduce the rate of photobleaching.^[5,13] One method is to simply reduce the amount of oxygen that would react with the triplet excited states. This can be done by displacing it using a different gas.^[5] Another method is by the use of free-radical scavengers to reduce the oxygen radicals. Shortening the long lifetime of the triplet excited state has also been shown to be effective.^[16] Other ways

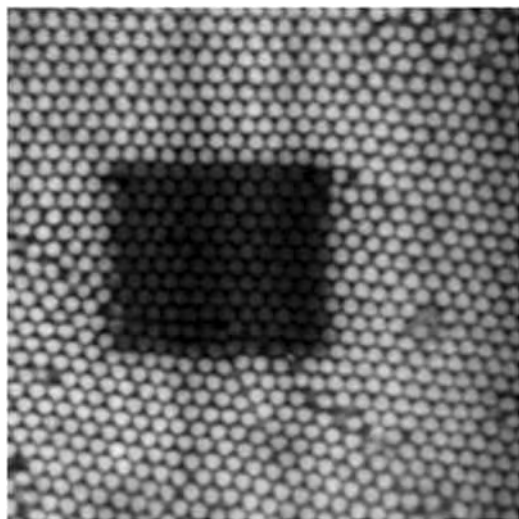


Fig. 8 Dyed suspension of densely packed polymethyl-methacrylate beads with significant photobleaching. The rectangular region near the center faded after about 30 sec of exposure to excitation light.

include using a high numerical aperture lens to collect more fluorescence light and thus use less excitation light.^[17] Also, keeping the magnification as low as is permissible spreads the excitation light over a larger area, thereby reducing the local intensity.

While photobleaching makes fluorescence microscopy more difficult, it is not always undesirable. One technique that takes advantage of it is fluorescence photobleaching recovery (FPR) or fluorescence recovery after photobleaching (FRAP). It involves exposing a small region of the specimen to a short and intense laser beam, which destroys the local fluorescence, and then observing as the fluorescence is recovered by transport of other fluorophore molecules from the surrounding region. Quantities such as the diffusion coefficient of the dyed structures can then be determined.^[18]

LIVING CELLS

Confocal microscopy has been used effectively for the 3D study of dynamics in living cells. However, the imaging of living specimens adds the challenge of maintaining the life and normal function of the organism.^[19] There are of course difficulties involved in preparing the sample for viewing as is the case in conventional microscopy. But on top of that the effect of photodamage on the specimen caused by the focused high-intensity excitation light must be taken into account. This is compounded by the repeated exposure required for tracking the cellular dynamics—a problem that is worsened for 3D data collection.

Fluorescence also introduces the problem of the fluorophore influencing the cell behavior as well as the risk that oxygen molecules reacting with fluorophores in triplet excited states may generate free radicals that damage the cell.^[19]

Despite the challenges, a wide variety of sophisticated fluorophores have been developed to study different aspects of cell biology. They are designed to mark specific parts of the cell interior and often can simply be introduced to the cell wall.^[19] The fluorophore molecules make their way into the cell and attach to the intracellular structures of interest such as the mitochondria and the Golgi apparatus. This is not always the case, however, as some fluorophores must be injected directly into the cell.^[19] Labeling is even applied to the study of “nonphysical” structures of the cell—some fluorophores have been developed for the measurement of dynamic processes such as membrane potentials and ion concentrations.^[5]

Multicolor Fluorescence

To distinguish between small features such as proteins within a cell it is useful to tag them with different fluorophores and image them as separate colors. There are two ways to do this: in one method fluorophores are selected to correspond with the wavelengths of a multiline laser and in the other their response to the same excitation wavelength causes emission at different wavelengths.^[20] In both cases the resulting emission is separated with appropriate filters and directed to different detectors. However, there can be cross talk between channels of the emitted light.^[14] For most of the commonly used fluorophores there is usually some overlap between their emission spectra, making perfect channel separation impossible by filtering alone. To first order this can be corrected by determining the level of overlap of emission of each individual fluorophore into the channels of the other fluorophores and subtracting it out mathematically.

FAST CONFOCAL MICROSCOPY

Most confocal microscopes generate a single image in 0.1–1 sec.^[21] For many dynamic processes this rate may be too slow, particularly if 3D stacks of images are required. Even for a single 2D image, slow frame rates translate into long exposure times of the specimen to intense laser light, which may damage it or cause photobleaching. Two commonly used designs that can capture images at high speed are the Nipkow disk confocal microscope and a confocal microscope that uses an acousto-optic deflector (AOD) for steering the excitation light.



Acousto-Optic Deflector

Speeding up the image acquisition rate can be achieved by making the excitation light beam scan more quickly across the specimen. For most confocal microscopes, the limitation is the galvanometers that move the mirrors back and forth in the characteristic saw-tooth pattern. The usual configuration is a slow vertical scan combined with a rapid scan in the horizontal direction. For 512×512 -pixel images at a frame rate of ~ 30 frames per second the horizontal galvanometer would have to scan at a frequency of approximately $30 \times 512 = 15\text{kHz}$, which is beyond its normal capability of several kilohertz.^[21]

Fast horizontal scans are achieved using an AOD. An AOD is a device that deflects light by creating a diffraction grating out of a crystal using sound waves. The sound waves are high-frequency pressure waves that locally alter the refractive index of the crystal. Thus, when monochromatic light shines through the crystal, it forms sharp fringes at an angle of deflection that depends on the wavelength of the acoustic pressure waves. Rapidly changing the frequency, and hence the wavelength, of the sound waves allows quick and accurate steering.^[22]

The major disadvantage of AODs is that they are wavelength sensitive. That is, different wavelengths experience different degrees of deflection. This presents a problem for fluorescence microscopy because the light from fluorescence has a different wavelength from the excitation light and thus cannot be descanned by the AOD as is done in Fig. 4 with the mirrors. To get around this problem the confocal microscope is designed to descann only along the vertical direction that is controlled by the slow galvanometer and to then collect the light using a slit rather than a pinhole.^[21] The penalty is a reduction in the amount of optical sectioning and a very slight distortion in the image caused by the loss of circular symmetry. Nevertheless, it is possible to obtain high-quality images (Fig. 5B) using slits.^[10] Note that descanning is not a problem for monochromatic reflected light microscopy because the incident and reflected light are of the same wavelength.

Nipkow Disk

An even faster technique is the so-called Nipkow disk microscope. Instead of scanning a single point across the specimen the Nipkow disk microscope builds an image by passing light through a spinning mask of pinholes, thereby simultaneously illuminating many discrete points. In the setup by Xiao, Corle, and Kino the mask is a disk with thousands of pinholes arranged in interleaved spirals.^[23] At any given time only a small

section of the disk with a few hundred pinholes is illuminated. The light travels through the pinholes and onto the specimen and the returning light passes through the same pinholes for optical sectioning. As the disk spins, the entire specimen is covered several times in a single rotation. At a rotation of 40 revolutions per second Xiao, Corle, and Kino were able to generate over 600 frames per second.

The disadvantage of the Nipkow disk microscope is that only a small fraction ($\sim 1\%$) of the illuminating light makes it through the pinholes to the specimen.^[24] While that is not a major problem when operating in reflected light mode, it can lead to a weak signal and poor imaging in fluorescence mode. However, with strong fluorophores an image as good as that of a confocal laser scanning microscope can be obtained.^[24] Increasing transmission would require an increase in pinhole size, which would lead to less effective optical sectioning and xy resolution.

TWO-PHOTON MICROSCOPY

A fast-growing technique that is related to confocal microscopy and also provides excellent optical sectioning is two-photon microscopy. It addresses a fundamental drawback of confocal laser scanning microscopy: that the beam also excites the specimen above and below the focal-plane (Fig. 9A). For each 2D image created using confocal microscopy there is a significant nonimaged portion of the specimen that has been subjected to photobleaching and photodamage. Two-photon microscopy cleverly limits fluorescence to the focal volume while also reducing the average intensity of the excitation light.^[25] The lower amount of photobleaching and photodamage makes it particularly desirable for biological and medical imaging.

Two-photon microscopy is based on a novel method of exciting fluorophore molecules. In confocal microscopy, a single high-energy photon excites a fluorophore molecule while it takes two lower-energy photons absorbed simultaneously to achieve the same result in two-photon microscopy. By itself a single photon cannot be absorbed because it does not have enough energy to excite a fluorophore molecule and thus causes no fluorescence. It must reach the fluorophore molecule within approximately 10^{-16} sec of the arrival of another photon so that together they reach the energy threshold for excitation. Not surprisingly, this simultaneous absorption of light is a very low-probability event at moderate light intensities. To get a sense of how rare it is, a molecule of rhodamine B will absorb a single photon approximately once a second in bright sunlight, but will absorb a two-photon pair once every 10 million years.^[26] Because

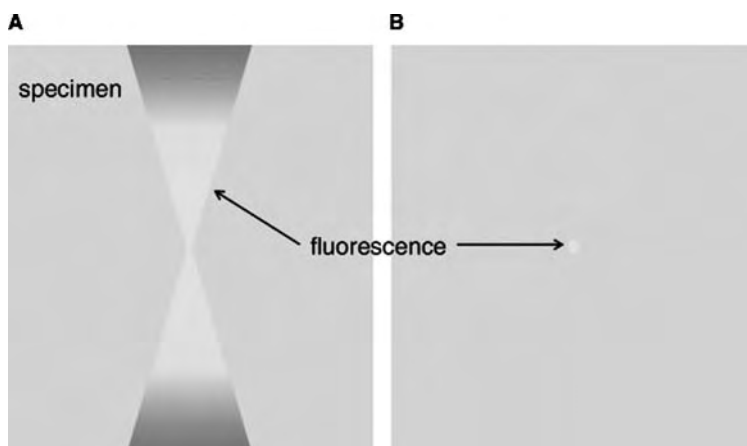


Fig. 9 One-photon vs. two-photon emission from a solution of fluorescein. (A) One-photon emission shows strong fluorescence at the waist, but also fluorescence in a cone-like region above and below the focal volume. (B) In two-photon emission fluorescence is limited to the focal volume. (From Ref. [25].)

C

the probability of absorption is proportional to the square of the instantaneous light intensity, lasers that generate very short ($<10^{-12}$ sec) but intense pulses are used.^[26,27] Such lasers can be operated at low average power—typically, five orders of magnitude lower than for confocal microscopy—thereby causing less photo-damage.^[25]

How does the above help in microscopy? The key is that because two-photon excitation events are unlikely at low laser intensities, fluorescence only occurs within the focal volume of the objective where the intensity is highest (Fig. 9B). The hourglass-shaped region of fluorescence characteristic of one-photon excitation fluorescence confocal microscopy is instead a small volume at the focus, which needs no detector pinhole for optical sectioning.^[25,26]

A major advantage of two-photon microscopy is that it is possible to image deeper into a specimen than with confocal microscopy. The reason is that without the pinhole in front of the detector more of the fluorescence signal is collected. This is important because as the observation depth within the specimen is increased, fluorescent light from the focal volume becomes more likely to be scattered. In confocal microscopy, the pinhole will reject that scattered light if its trajectory makes it appear not to originate from the focal plane. To make matters worse, the pinhole may accept out-of-plane fluorescence light that is scattered and appears to come from the focal plane. Together the two effects significantly reduce the image contrast at depths of $\sim 100\ \mu\text{m}$. In two-photon microscopy all fluorescent light is from the focal volume and is thus acceptable. Plus it is much more likely to reach the detector because it does not have to pass through a pinhole.^[26] This arrangement allows imaging deeper than $\sim 400\ \mu\text{m}$.^[25]

Why not avoid the confocal microscope altogether in favor of a two-photon microscope? The main drawback with two-photon microscopy is the challenges associated with the type of laser it needs. Ultrashort pulse

lasers are still very expensive compared with the continuous wave lasers used in confocal microscopy. They are also typically more demanding to set up.^[27] For example, they have to be on the same optical table as the microscope because of difficulties in maintaining the pulse characteristics through an optical fiber. They need a stronger power supply and much more cooling—usually water pumped through it at several gallons per minute.

CONCLUSIONS

A confocal microscope provides a significant imaging improvement over conventional microscopes. It creates sharper, more detailed 2D images, and allows collection of data in three dimensions. In biological applications it is especially useful for measuring dynamic processes. A number of designs have been developed to achieve video-rate confocal microscopy, which enables the capture of short-timescale dynamics.

FURTHER READING

For detailed discussions about most aspects of confocal microscopy the book cited in Ref. [6] is an excellent source. Ref. [25] cites a book that along with providing insight and detail about confocal microscopy also gives plenty of information on two-photon microscopy.

ACKNOWLEDGMENT

The authors would like to acknowledge the support from the National Science Foundation under Grant DMR-0239109.

REFERENCES

1. Minsky, M. Memoir on inventing the confocal microscope. *Scanning* **1988**, *10*, 128–138.
2. Wilson, T.; Carlini, A.R. Three dimensional imaging in confocal imaging systems with finite-sized detectors. *J. Microsc.* **1988**, *141*, 51–66.
3. Sheppard, C.J.R.; Shotton, D.M. Image formation in the confocal laser scanning microscope. In *Confocal Laser Scanning Microscopy*; Springer-Verlag, New York Inc.: New York, 1997; 15–31.
4. Guilbault, G. General aspects of luminescence spectroscopy. In *Practical Fluorescence*, 2nd Ed.; Guilbault, G., Ed.; Marcel Dekker, Inc.: New York, 1990; 1–40.
5. Atkins, P. Spectroscopy 2: electronic transitions. In *Physical Chemistry*, 5th Ed.; W. H. Freeman and Company: New York, 1994; 590–622.
6. Tsien, R.Y.; Waggoner, A. Fluorophores for confocal microscopy. In *Handbook of Biological Confocal Microscopy*, 2nd Ed.; Pawley, J.B., Ed.; Plenum Press: New York, 1995; 267–280.
7. Sheppard, C.J.R.; Shotton, D.M. Introduction. In *Confocal Laser Scanning Microscopy*; Springer-Verlag New York Inc.: New York, 1997; 1–13.
8. Porter, T.; Duff, T. Compositing digital images. *Comput. Graphics* **1984**, *18* (3), 253–259.
9. Inoué, S.; Spring, K.R. Microscope image formation. In *Video Microscopy the Fundamentals*; Plenum Press: New York, 1997; 13–118.
10. Keller, H.E. Objective lenses for confocal microscopy. In *Handbook of Biological Confocal Microscopy*, 2nd Ed.; Pawley, J.B., Ed.; Plenum Press: New York, 1995; 111–126.
11. Wilson, T. The role of the pinhole in confocal imaging system. In *Handbook of Biological Confocal Microscopy*, 2nd Ed.; Pawley, J.B., Ed.; Plenum Press: New York, 1995; 167–182.
12. Pawley, J. Fundamental limits in confocal microscopy. In *Handbook of Biological Confocal Microscopy*, 2nd Ed.; Pawley, J.B., Ed.; Plenum Press: New York, 1995; 19–38.
13. Becker, P.L. Quantitative fluorescence measurements. In *Fluorescence Imaging Spectroscopy and Microscopy*; Wang, X.F., Herman, B., Eds.; John Wiley & Sons, Inc.: New York, 1996; 1–29.
14. Sheppard, C.J.R.; Shotton, D.M. Confocal fluorescence microscopy. In *Confocal Laser Scanning Microscopy*; Springer-Verlag New York Inc.: New York, 1997; 61–70.
15. Chen, H.; Swedlow, J.R.; Grote, M.; Sedat, J.W.; Agard, D.A. The collection, processing, and display of digital three-dimensional images of biological specimens. In *Handbook of Biological Confocal Microscopy*, 2nd Ed.; Pawley, J.B., Ed.; Plenum Press: New York, 1995; 197–210.
16. Song, L.; Varma, C.A.; Verhoeven, J.W.; Tanke, H.J. Influence of the triplet excited state on the photobleaching kinetics of fluorescein in microscopy. *Biophys. J.* **1996**, *70*, 2959–2968.
17. Lemasters, J.J. Confocal microscopy of single living cells. In *Fluorescence Imaging Spectroscopy and Microscopy*; Wang, X.F., Herman, B., Eds.; John Wiley & Sons, Inc.: New York, 1996; 157–177.
18. Axelrod, D.; Koppel, D.E.; Schlessinger, J.; Webb, W.W. Mobility measurement by analysis of fluorescence photobleaching recovery kinetics. *Biophys. J.* **1976**, *16*, 1055–1069.
19. Teraski, M.; Dailey, M.E. Confocal microscopy of living cells. In *Handbook of Biological Confocal Microscopy*, 2nd Ed.; Pawley, J.B., Ed.; Plenum Press: New York, 1995; 327–346.
20. Stelzer, E.H. The intermediate optical system of laser-scanning confocal microscopes. In *Handbook of Biological Confocal Microscopy*, 2nd Ed.; Pawley, J.B., Ed.; Plenum Press: New York, 1995; 139–154.
21. Tsien, R.Y.; Bacskai, B.J. Video-rate confocal microscopy. In *Handbook of Biological Confocal Microscopy*, 2nd Ed.; Pawley, J.B., Ed.; Plenum Press: New York, 1995; 459–478.
22. Iyer, V.; Losavio, B.E.; Saggau, P. Compensation of spatial and temporal dispersion for acousto-optic multiphoton laser-scanning microscopy. *J. Biom. Opt.* **2003**, *8* (3), 460–471.
23. Xiao, G.Q.; Corle, T.R.; Kino, G.S. Real-time confocal scanning optical microscope. *Appl. Phys. Lett.* **1988**, *53* (8), 716–718.
24. Kino, G.S. Intermediate optics in Nipkow disk microscopes. In *Handbook of Biological Confocal Microscopy*, 2nd Ed.; Pawley, J.B., Ed.; Plenum Press: New York, 1995; 155–165.
25. Diaspro, A.; Sheppard, C. Two-photon microscopy: basic principles and architectures. In *Confocal and Two-Photon Microscopy Foundations, Applications and Advances*; Diaspro, A., Ed.; Wiley-Liss, Inc.: New York, 2002; 39–73.
26. Denk, W.; Svoboda, K. Photon upmanship: why multiphoton imaging is more than a gimmick. *Neuron* **1997**, *18*, 351–357.
27. Denk, W.; Piston, D.W.; Webb, W.W. Two-photon molecular excitation in laser-scanning microscopy. In *Handbook of Biological Confocal Microscopy*, 2nd Ed.; Pawley, J.B., Ed.; Plenum Press: New York, 1995; 445–458.

Contact Lenses: Silicone Hydrogels

Paul C. Nicolson

Atlanta, Georgia, U.S.A.

Jürgen Vogt

CIBA Vision Corp., Duluth, Georgia, U.S.A.



INTRODUCTION

The subject of conventional soft hydrogel contact lenses has been covered previously in this encyclopedia.^[1] Further background can be found in the literature.^[2,3] This entry will discuss the exciting new family of contact lenses and the polymers of which they are composed. This new family of lenses, prepared from silicone hydrogels, offers the wearer a more ophthalmically compatible, more comfortable contact lens having a lower risk of adverse events, especially for those wearing lenses extended wear, which means wearing the contact lens continuously day and night for several days, or even weeks. The designations silicone and siloxane are essentially equivalent and can be used interchangeably, with “silicone hydrogel” used as the main terminology in the contact lens industry.

This entry will discuss the structural features of silicone hydrogels as contact lens polymers, including morphology and how innovators in this field have addressed the well-recognized problems of corneal hypoxia caused by contact lens wear and the notorious hydrophobic and lipophilic characteristics of silicone containing polymers.

BIOMATERIALS IN THE OCULAR ENVIRONMENT

The ocular environment places high demands on the performance of contact lenses as biomaterials. Ophthalmic compatibility of a contact lens requires that the lens maintain a stable, continuous tear film for clear vision, is resistant to deposition of tear film components, sustains normal hydration, is permeable to oxygen to maintain normal corneal metabolism, is permeable to ions to maintain movement, and is nonirritating and comfortable. Therefore, the lens must have excellent bulk and surface characteristics, being neither completely hydrophobic nor lipophilic, and must possess the appropriate polymer composition and morphology to be successful. The polymers must also be optically clear and homogeneous and have low modulus for comfort and sufficient mechanical strength for normal handling.

OXYGEN PERMEABILITY

The use of silicone hydrogel materials has been vigorously pursued in contact lens, as well as intraocular lens applications since the late 1970s. As greater understanding of the needs of human cornea for oxygen and for removal of metabolic products was gained, researchers sought materials that would better meet those needs. The cornea has no blood supply, and thus obtains oxygen predominantly from the air. Contact lenses reduce the amount of oxygen reaching the cornea, and thus can cause edema, microcysts, red eyes, and more serious and sight-threatening complications.

The permeability (Dk) of a specific gas through a given material is the product of the diffusivity D times the solubility of that gas in that material. For the diffusion of oxygen through contact lens materials, values of Dk are usually given in units of barrer = 10^{-11} [(cm³ O₂ STP) cm²/(cm³ s (mm Hg))]. The oxygen transmissibility Dk/t for a material of a given thickness is conveniently given in unit of barrer/(0.1 mm) = 10^{-9} [(cm³ O₂ STP)/(s (mm Hg))], thus the numerical value of the Dk of a material in barrer is equal to the same numerical value of a lens of 0.1 mm thickness, a typical value for soft contact lenses.

In conventional hydrogel contact lenses, the water contained in the lens is providing the pathway for oxygen transport through the lens. Thus, the achievement of higher oxygen permeability with hydrogels was first pursued through the development of high-water content (>50%) hydrogels that contained strongly hydrophilic monomers such as *N*-vinylpyrrolidone (NVP) and methacrylic acid (MAA). However, these endeavors soon reached their limit. Pure water has an oxygen permeability (Dk) of about 80 barrer. The revelation that a minimum oxygen transmissibility (Dk/t) of 87 barrer/(0.1 mm) was required for a lens to be worn safely overnight^[4] meant that thin, and thus more oxygen transmissible, designs were necessary in the high-water hydrogels, if they were to approach the required performance level for corneal health. But the use of extreme high water content (>80%) materials would have distinct disadvantages such as poor mechanical properties, difficult handling, difficult manufacture,

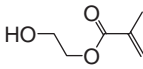
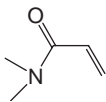
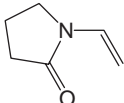
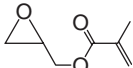


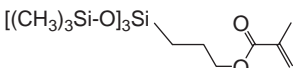
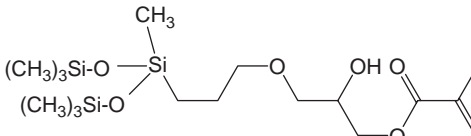
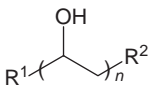
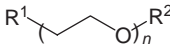
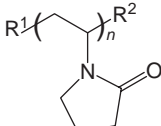
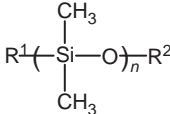
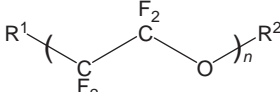
uptake of tear proteins, and care system components. Additionally, the findings of corneal desiccation with those material/design combinations eliminated them from serious consideration for extended wear.

Innovations in contact lens materials designed to better meet the oxygen needs of the eye appeared in 1978 through the efforts of Gaylord,^[5-7] scientists at Dow Corning^[8,9] and others in the rigid gas permeable (RGP) contact lens field. Those innovations included

the use of organosilicon-containing compounds, such as silicone acrylates composed of the tris-(trimethylsiloxy-silyl)-propyl methacrylate (TRIS) (Table1), polydialkyl- and/or diaryl-siloxane, as well as silicone resins. Since the focus of this paper will be soft contact lens materials, nothing further will be said concerning the rigid materials.

First, why is there so much interest in high oxygen permeability (Dk)? The work done by DeCarle^[10] in

Table 1 Typical monomers and units used in contact lens manufacturing

2-Hydroxyethyl methacrylate	HEMA	
<i>N,N</i> -dimethyl acrylamide	DMA	
<i>N</i> -vinyl pyrrolidone	NVP	
Glycidyl methacrylate	GMA	
Methyl methacrylate	MMA	
Methacrylic acid	MAA	
3-Methacryloxypropyl tris(trimethylsiloxy)silane	TRIS	
	SiGMA	
Poly(vinyl alcohol)	PVA	
Poly(ethylene glycol)	PEG	
Poly(vinyl pyrrolidone)	PVP	
Silicone, poly(dimethyl siloxane)	PDMS	
Perfluoro poly(ethylene glycol)		

England in the early 1970s, fitting contact lenses for overnight (extended wear), created a demand for the convenience that such a modality offered. Wearers were demanding “natural vision,” waking up and seeing the alarm clock, etc. without the hassle of daily lens removal and requisite cleaning and disinfection. However, it wasn't long before the negative physiological effects of the corneal oxygen deprivation created by extended wear of the lenses available at that time began to be realized, first in Europe^[11] in the late 1970s and then in the United States in the mid-1980s.^[12–14] As a result, in the United States in 1987, the Food and Drug Administration rescinded all 30 night extended wear approvals they had originally granted starting in 1981, leaving the only approved extended wear modality to be 6 nights' wear.^[15] However, researchers have shown that 6 nights of corneal compromise caused by wearing conventional hydrogels can still cause problems in some wearers.^[16–18]

A basic physical property of organosilicon-containing polymers is their high permeability to oxygen. The reported Dk of pure dimethylsiloxane is about 600 barrer, which is significantly higher than the 8–20 barrer of the average hydrogel. This high permeability to oxygen comes about by virtue of the bulkiness of the siloxane group generating free volume, and thus having high chain mobility that characterizes such materials, and which is responsible for the high diffusivity of oxygen through silicone-containing materials. The first commercial, soft silicone lens, elastofilcon A, was developed in 1980 by Dow Corning.^[19] This lens was a silicone elastomer containing no water and, as such, the lens manifested the typical rubbery behavior of silicone. These properties required the lenses be fabricated in small diameters, much like RGP lenses, to avoid binding or adhering to the cornea, and therefore were never able to deliver the comfort characteristics of hydrogels. Further, the inherent lipophilicity of the material required the lenses be surface treated. Elastofilcon A lenses are still available today, but only in aphakic powers for children and adults.^[20]

SILICONE HYDROGELS: BACKGROUND

The high oxygen permeability of polysiloxanes, offset by their characteristic hydrophobicity and rubbery character, led investigators to the design of a novel family of materials called silicone hydrogels. It was hypothesized that such materials would offer the softness, wettability and on-eye comfort of a conventional hydrogel, while at the same time providing the higher oxygen transmission required by the cornea.

The siloxane group can be represented structurally as follows: $-\text{Si}(\text{R})_2-\text{O}-$. It is a silicon atom bound to two oxygen atoms and is bonded to two R groups.

In contact lens technology the R's are most often CH_3 (methyl) groups or another siloxane group. However, in the general chemical literature, and in some silicone hydrogel contact lens patents, R can be a wide range of groups. n is the number of times the siloxane group is repeated in a “chain” structure. When n is greater than about five or six, then the term “poly(dimethyl siloxane)” (PDMS) is used to define the unit. Such units are usually linear. In some contact lens materials, n can be as large as 25 and even larger. When the R groups are siloxane groups, the entire unit becomes relatively bulky and space filling. A classic example of such a bulky siloxane can be seen in Gaylord's TRIS siloxane monomer. The end result of incorporating these groups into a polymer network is that significant free volume is created, especially when these units aggregate together to form a domain or phase. It is through this free volume that the relatively small oxygen molecules can pass, or diffuse, freely. The outcome is that the polymer network has a high permeability to oxygen and some other gases. The TRIS monomer, or closely related bulky siloxanes, are very frequently included in silicone hydrogel patents as part of mixtures or formulations of monomers and macromers that, when polymerized, form useful contact lens materials.^[21]

Both the linear PDMS and the bulky TRIS-type siloxanes can be incorporated as a part of a monomer. Again, that means the siloxane group is chemically bonded, through various linking groups, to a carbon-carbon double bond. Therefore the siloxane groups are present, pendant to the carbon backbone of the polymer network in the same fashion as hydrophilic groups are pendant in hydrogels.

In the authors' experience, particularly with linear PDMS groups, the longer the chain (i.e., the larger the value of n) the more efficient is the improvement in oxygen permeability. This is believed to be the case because larger PDMS groups more readily aggregate to form domains or phases within the polymeric material. For any polymer however, there is a practical upper limit to the length of the PDMS due to problems of macromer viscosity, ultimate phase size, compatibility with hydrophilic monomers, and optical clarity of the polymer.

There is a third means of incorporating siloxane groups into a polymer network. Macromeric hydrophiles can also contain silicone. Such combination macromers are disclosed in silicone hydrogel patents,^[22] but by far the most frequently disclosed means of incorporating siloxane into silicone hydrogel contact lens materials is through the use of silicone-only macromers that contain two double bonds. Such macromers are generally linear PDMS-containing, have large values of n (20–25 and higher) and serve two purposes. The first is, of course, to build into the polymer, large siloxane units that can more readily aggregate and form phases



for the facile transport of oxygen. Secondly, such bifunctional macromers serve as crosslinking agents that, in combination with carbon backbone polymer formed by the chains of all the monomers and macromers employed, form the polymer network. The siloxane chains from the macromer in this scenario can be considered “main-chain” siloxanes, in contrast to the pendant siloxanes discussed previously.

Bifunctional macromer silicones, because they form crosslinks within the polymer, tend to increase the modulus or stiffness of the resultant silicone hydrogel. In fact, 3 of the 5 current commercial silicone hydrogel contact lenses, demonstrate moduli or stiffness that is somewhat greater than most conventional hydrogel lenses. This is because those materials, balafilcon A and lotrafilcon A and B, contain silicone macromer crosslinkers.

The modulus of a crosslinked polymer generally increases proportionally with the number of crosslinks present and also decreases with increasing distance between crosslink points (the value of n in the bifunctional macromer) i.e., the points where crosslinking monomers or macromers connect to the carbon backbone. Therefore, longer silicone macromers are preferred when preparing silicone hydrogel materials for contact lenses. Structures of monomers and macromers commonly used in contact lens polymers are shown in Tables 1 and 2, respectively.

The higher modulus characteristic of the early silicone hydrogel polymers prompted polymer development to move in the direction of reduced modulus.^[23] In fact, a recent patent discloses a relationship between the modulus of the polymer, the thickness of the lens, and the clinical incidence of SEALs (superior epithelial arcuate lesion) observed with some lenses.^[24]

EARLY INNOVATIONS

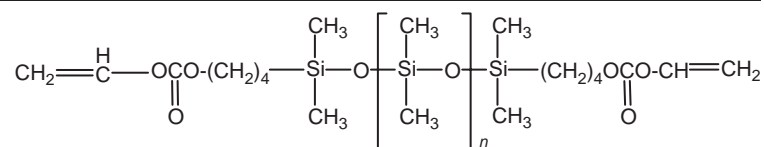
Chang perhaps can be credited with the first patent dealing with polymers composed of siloxane-containing monomers combined with hydrophilic monomers that absorbed small amounts of water for use as contact lens materials.^[25]

Independently, in July and November 1977, and February 1978, CIBA-Geigy (Mueller, Kleiner),^[26] Toyo Contact Lens Company in Japan (Tanaka et al.)^[27] and Bausch and Lomb (Deichert et al.),^[28] respectively, each applied for patents on new families of high Dk polymers for contact lens polymers called silicone hydrogels. All three claimed these hydrophilic siloxane-containing materials as being novel compositions offering high oxygen permeability, thereby overcoming the shortcomings of conventional hydrogels for extended wear. Mueller and Deichert both employed macromers derived from bi- and/or poly-functional, methacrylated polydimethyl siloxane (PDMS) polyols, (main-chain siloxanes) in combination with hydrophilic monomers (NVP, HEMA). Chang^[25] and Tanaka,^[27] on the other hand, used monofunctional bulky siloxane-containing acrylates (pendant siloxanes). More recent patents describe the use of combinations of siloxane macromers with pendant siloxanes (e.g., TRIS). The use of TRIS in combination with linear macromeric PDMS is said to provide silicone hydrogel contact lens polymers with improved mechanical properties, compared to those using macromeric PDMS alone due to reduced crosslink density.^[21]

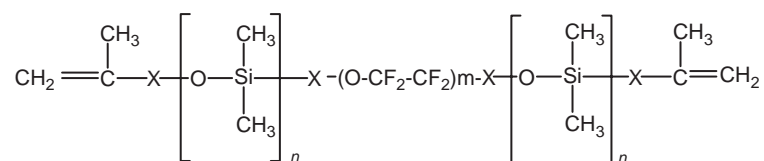
The macromer technologies disclosed by these early researchers differed largely in the linking groups employed. Condensation chemistries, namely urethane, ester, carbonate, thiourethane, urea, and amide group formation were the principle modes by which the final

Table 2 Macromonomers used in commercial silicone hydrogel contact lenses

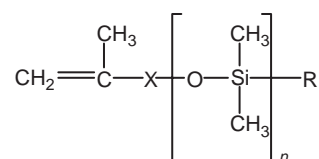
Balafilcon A



Lotrafilcon A, B



Galyfilcon A, senofilcon A



structures were assembled. Urethanes were particularly popular in that no byproducts were formed, although they were criticized in some early patents as being undesirable for use in medical devices, particularly contact lenses.^[29] No reasons were given, however.

OTHER OXYGEN PERMEABILITY-ENHANCING MOIETIES

PDMS was not the only oxygen permeability enhancing component pursued for contact lens polymers. Rice at 3M presents novel macromers derived from perfluoropolyethers (PFPE).^[30,31] PFPE moieties function by increasing the oxygen solubility in polymers (the *k* in the *Dk* expression) while siloxanes increase diffusion (the *D* component) of oxygen. The PFPE technology appeared briefly on the contact lens market as fluorofacon A contact lenses, a terpolymer of a PFPE dimethacrylate, methyl methacrylate and NVP.^[32] Their mechanical properties fell somewhere between soft and rigid materials, thus they were termed flexible lenses. Lacking siloxane, fluorofacon A lenses did not manifest lipophilicity and their high *Dk* made them attractive for extended wear. However, because they were neither hard nor soft and they contained significant amounts of fluorochemicals that are relatively expensive, their unique fitting behavior and high cost contributed to their relatively short life on the market.

Building on the concept of permachor,^[33] the relationship of structure to oxygen permeability developed by the packaging industry, other nonsiloxane structures were studied. Patents on polyethylene/polypropylene oxide-containing soft polymers, with and without water, have issued.^[34,35] Although the *Dk* values reported are not as high as comparable materials derived from siloxanes, the materials did outperform conventional hydrogels with regard to oxygen permeability. There have not been any contact lens polymers derived from polyether technology that have been commercialized to our knowledge. To date, there have been no reports of other polymer materials for contact lens application that have the oxygen permeability characteristics of silicones or fluorosilicones.

SILICONE HYDROGELS: MORE RECENT DEVELOPMENTS

Following the earlier silicone hydrogel patents, a great number of patents have been issued in the field of silicone- and later fluorosilicone- and fluoro-hydrogels for contact lenses. The bulk of these patents, emanated from commercial enterprises, i.e., Bausch and Lomb, CIBA Vision, Johnson & Johnson, Vistakon, and a number of Japanese companies, such as Menicon and

Shin-Etsu, reflecting a commercial interest in developing new, better performing silicone and/or fluoro-silicone hydrogel polymers for contact lenses. For example, recent published applications by Alli et al.^[36] and Vanderlaan et al.^[37] disclose the use of a monofunctional macromer (mPDMS) as a means of incorporating siloxane into contact lens polymers. The benefit noted is that use of such monofunctional entities is the achievement of high oxygen permeability without the concurrent increase in modulus that occurs with the previously disclosed bifunctional macromers.

Historically, the use of TRIS in silicone hydrogel contact lens polymers has been very frequently disclosed in patents. A recent patent by McCabe et al. discloses a new, bulky siloxane acrylate monomer called SiGMA (2-propenoic acid, 2-methyl-2-hydroxy-3-[3-[1,3,3,3-tetramethyl-1-[(trimethylsilyloxy)-disiloxanyl]-propoxy] propyl ester).^[38] This new, bulky siloxane monomer shares a degree of similarity to that described by Tanaka, in that hydrophilicity in the form of hydroxyl and ether groups is incorporated in the linkage between the acrylate group and the siloxane group. This monomer is said to improve compatibility between the hydrophiles and siloxanes present in the polymers formed.

SILICONE MACROMER ARCHITECTURE

Silicone hydrogel materials development, based on a review of issued patents and patent applications, has followed different paths. Materials development strategies have included bifunctional macromers in large part, but the architecture of those macromers has been varied significantly. In the simplest form, silicone macromers were assembled by “capping” one mole of a PDMS diol with two moles of methacryloyl chloride. Alternatively, one mole of the PDMS diol could be reacted with exactly two moles of a diisocyanate, followed by consumption of the resulting diisocyanate with HEMA. Simple bi-functional PDMS derivatives such as those described by Deichert and Mueller have been chain extended^[39] in an effort to increase silicone content as well as to reduce modulus in such materials. Chain extension refers to the concept of linking two or more PDMS moieties together in a linear fashion. The starting PDMS units, generally available as α - ω diols or diamines, are reacted with a diisocyanate in a stoichiometry such that the PDMS units are joined to form an oligo-urethane or oligo-urea chain. The commercial availability of a new isocyanate methacrylate, isocyanatoethylmethacrylate (IEM), simplified the construction of methacrylate macromers for biomaterials.

Methacrylate end-capped, segmented hard–soft–hard polyurethane macromers have also been disclosed, offering improved mechanical strength according to the inventors.^[40] The soft segment of the materials contained



PDMS, as well as other long chain, flexible moieties, while the hard segment was composed of short chain diols linked through di-isocyanates such as isophorone di-isocyanate (IPDI).

Construction of methacrylate functional PDMS macromers by first reacting a di-isocyanate with the PDMS diol, followed by a monomer such as HEMA, inevitably lead to some chain extension even when exact stoichiometry was used. Use of IEM reduced the extent of that phenomenon and made macromer structure more controlled and predictable.

As mentioned earlier, monofunctional macromers have appeared more recently in patents. They allow the incorporation of blocks of PDMS to achieve oxygen permeability but they do not tend to increase modulus.

The hydrophobic and lipophilic character of PDMS led many investigators to explore more hydrophilic polysiloxanes. Patents^[22,41,42] report the preparation of amphiphilic segmented A–B–A-type block macromers, where A represents a hydrophilic segment such as polyethylene oxide or polyoxazoline with molecular weights in excess of 400, and B is a hydrophobic PDMS or PFPE segment with molecular weights of 1000–5000 or more. Since these macromers are inherently hydrophilic, homopolymers as well as co- and terpolymers, with other hydrophilic and hydrophobic monomers, have, in general, good soft contact lens bulk material properties, namely phase morphology, water content, and mechanical properties.

Keough et al. make a point in their patent about the troublesome incompatibility of PDMS-containing components and hydrophilic monomers leading to opaque, phase-separated products. To resolve this, they disclose the preparation of PDMS macromers with polar or hydrophilic side-chains.^[43] Such macromers are claimed to have enhanced compatibility with hydrophilic monomers. The hydrophilic side chains incorporate PEO or glycol moieties. An earlier investigator, LeBoeuf, discloses similar hydrophilic side-chain siloxanes, indicating the polymers derived from them are more wettable and that the siloxanes are compatible with hydrophilic monomers.^[44,45] Interestingly, Künzler reported problems of significant instability when the side chains are PEO, in contrast to when glycols are incorporated.^[46] A limited shelf lifetime was also reported by Nicolson in A-B-A lens polymers, where A is PEO. The materials manifested an inadequate shelf life for commercial purposes, whereas when A is poly-oxazoline, no such problems were observed.^[47]

Other strategies followed to obtain compatibility with hydrophilic components included PDMS macromers containing side chains with polar fluoro-substituents ($-\text{CF}_2\text{H}$). A second benefit of reduced lipophilicity, imparted by the fluoro-substituent, is claimed.^[23]

Bambury and Seelye introduced a new family of silicone hydrogel polymers employing vinyl carbonate- and

vinyl carbamate-functionalized compounds.^[48] They described silicone copolymer hydrogel films that contain NVP. Such vinyl-functional macromers co-polymerize more readily with NVP than do acrylate-functional macromers, improving NVP incorporation into the final polymer. Unique cross-linking agents containing both vinyl and acrylate functions have also been disclosed and have been claimed to improve co-polymerization of vinyl and acrylate-functional monomers and macromers.^[49]

SILICONE HYDROGEL COMPOSITION—THE HYDROPHILIC COMPONENTS

The hydrophilic monomers that can be, and are, used in silicone hydrogels are very much the same as those used in conventional hydrogels. However, a review of patents reveals that by far, the predominant hydrophilic monomers where the pendant group is small and contains a nitrogen atom. The monomers NVP and DMA appear much more frequently than does HEMA. This is, at first sight, strange due to the prevalent use of HEMA monomer in conventional hydrogels. What is essential in silicone hydrogel contact lenses, as in any contact lens, of course, is optical clarity. Clarity in polymers is a result of the achievement of a degree of compatibility amongst the hydrophilic- and silicone-containing components. Predominant citation of nitrogen containing hydrophilic monomers in patents suggests compatibility of significant amounts of HEMA with siloxane components, in a hydrogel polymer, may be more difficult to achieve.

A recent U.S. patent application filed by Menicon discloses a new hydrophilic monomer for silicone hydrogel contact lenses.^[50] It is based upon a methacrylate monomer, and the hydrophilic portion is triethyleneglycol, a member of the ether family, and does not contain nitrogen. Menicon does not yet have a commercial silicone hydrogel contact lens. Time will tell whether this new technology will emerge to compete within this exciting field of high *Dk* silicone hydrogels.

The amounts of hydrophilic components used in silicone hydrogels for contact lenses can vary widely, depending upon the amount of water desired in the final polymer. Since the intent of silicone incorporation into hydrogels is to significantly increase oxygen permeability, generally it is desirable to keep water contents as low as practicable to achieve high oxygen permeability, as well as other properties including low modulus, mechanical strength, comfort, etc. From our review of silicone hydrogel patents, hydrophilic component levels fall in the range from about 30% up to about 45% based on dry polymer weight.

Silicone Hydrogel Composition; Hydrophobic Comonomers

One would wonder why there might be a need for additional hydrophobic monomers in polymer systems already rich in hydrophobic siloxane monomers and macromers. Hydrophobic monomers have been described as being “toughening agents,” which improve mechanical strength, often at the expense of oxygen permeability. They also provide an additional means of fine tuning the final physical properties of polymers, impacting water content, refractive index, and modulus. Monomers containing fluorine offer the opportunity of maintaining oxygen permeability, while reducing the lipophilicity imparted by the silicones. The more common hydrophobic comonomers disclosed, however, are simple aliphatic esters of methacrylic acid, such as methyl methacrylate (MMA).

TRANSPORT PROPERTIES OF SILICONE HYDROGELS

It has been known for a long time that some polymeric materials possess excellent oxygen (and other gas) transport properties, one of the most prominent of such materials being silicone. Silicones are hydrophobic and do not take up significant amounts of water. Due to the bulkiness of the silicone polymer backbone ($-\text{Si}(\text{R})_2-\text{O}-$ chains), the amorphous silicone bulk material has a rather large “free volume,” or space, between the polymeric chains, through which small, nonpolar molecules, such as oxygen and nitrogen, can diffuse quite easily. This characteristic of silicone was exploited some 20 years ago in silicone elastomer lenses. But such lenses had profound disadvantages, including nonwettability and the need to fabricate the lenses in small diameters to avoid binding to the cornea. Even after appropriate surface coating or modification, such lenses were not sufficiently comfortable and biocompatible.

Attempts to improve the hydrophilicity and biocompatibility of silicones were numerous, mainly in chemical modification of the silicone polymer backbone with hydrophilic groups and in “blending” siloxane monomers with hydrophilic monomers. The former approach makes the overall polymer more hydrophilic, thus enabling water uptake. This approach, however, essentially fills up the free volume created by the siloxane groups in the polymer with water, rendering this type of silicone hydrogel a conventional homogeneous (mono-phase) hydrogel material. Oxygen molecules must still traverse the polymer network through the water. As discussed earlier, the oxygen permeability of such hydrogels is too low to be suitable as extended wear contact lens materials. The latter approach of blending silicone with hydrogel creates the water-and-oil

immiscibility dilemma: if forced to mix, the system forms an oil-in-water or a water-in-oil emulsion, turning opaque due to scattering of light by the droplets of the dispersed phase in the continuous phase.

Opacification of the biphasic system can be avoided by carefully designing the architecture of the polymer chains, controlling the domain size of the two immiscible polymer phases in such a way as to stay significantly below the wavelength of visible light (i.e., less than 400 nm).

Intensive research work in the authors' laboratories, with blending siloxane-containing and hydrophilic monomers, resulted in numerous optically clear silicone hydrogel lenses. The silicone hydrogel materials had excellent oxygen permeability, however they initially lacked sufficient comfort and movement on eye. This finding caused great consternation for our researchers until an unanticipated discovery concerning contact lens polymer morphology was made.^[51]

POLYMER PHASES IN SILICONE HYDROGELS

Up until 1995, literature and patents were silent on phases in contact lenses. If phases were mentioned, they were mentioned in the context of something that was to be avoided. Phase separation, and the resultant typical opacity associated with it, lies counter to the optical performance required by such products. The major message was that monomers employed in contact lenses should be as compatible as possible to avoid phase separation.

Silicones and most hydrophilic polymers are typically immiscible and when given the opportunity, they will phase separate. Phase separation occurs when cohesive interactions between chemically identical, or similar molecules, are stronger than the adhesive interactions between chemically different molecules. If mutual disliking groups are part of the same molecule, as exist in silicone hydrogel (amphiphilic) crosslinked networks, covalent bonds limit the maximal separation length between the phases. Separation lengths of about 100 nm or less generally lead to microphase separated systems. Such systems then are optically clear, as the phase size is significantly less than the wavelength of light. A recent publication by Gido describes the application of field emission gun scanning transmission electron microscopy (FEGSTEM) to the study of connectivity of domains in microphase separated polymer materials. Results obtained on CIBA Vision's Focus Night & Day lens (Iotrafalcon A) indicate a morphological length of about 20 nm.^[52]

In distinct contrast to prior conceptions about phase separation, Nicolson et al., in a series of comprehensive patents,^[51,53–56] describe the requisite properties for an ophthalmically compatible silicone hydrogel contact



lens, namely the need for separate, and cocontinuous ion/water permeable (ionoperm) hydrogel and oxygen permeable (oxyperm) pathways. An ophthalmically compatible lens is defined as a lens which has high oxygen transmissibility, moves on the eye, and does not accumulate significant lipid or protein deposits, as some of the necessary and desirable features.

The revelation that co-continuous phases or pathways are necessary to achieve an extended wear contact lens came about from the clinical evaluation of a large number of prototype silicone hydrogel-based contact lenses. Although generally, these test lens compositions had properties suitable for a soft extended-wear contact lens, clinical assessment revealed almost all of them ceased movement immediately or shortly after placement on the eye. It became clear that only those compositions manifesting a certain level of ion permeability (e.g, sodium chloride) would move on the eye. Differential scanning calorimetry (DSC) analysis of these compositions clearly showed the presence of separate silicone and hydrogel phases. The presence of micro-phases alone was clearly not sufficient however. The extent and continuity of both phases is key to on-eye movement and the achievement of high-oxygen permeability. A conceptual model illustrating cocontinuous phases is the cellulose sponge, with continuous sponge material and air "phases." Contrast this model with the Styrofoam cup, where the air "phase" is discontinuous.

ION PERMEABILITY AND THE PREDICTION OF ON-EYE MOVEMENT

Domschke et al. in a 1997 ACS presentation,^[57] described the research and the ion permeability test method developed to further elucidate the presence and nature of the continuity of the hydrogel phase. A percolation model can distinguish lens materials that bind on the cornea from those that will move.

Research is currently underway to determine just how water and ion permeability affect the ability of a lens to move on the eye. Tighe has proposed that the permeation of ions through the lens material allows for the maintenance of a hydrodynamic boundary layer between the lens and the cornea to avoid hydrophobic binding.^[58]

There does seem to be a level of ion permeability above which a lens material will move on the eye. The fact that greater permeability does not seem to promote greater movement supports the hydrodynamic boundary level theory and implies that once the layer is formed, its thickness is not critical. A previously accepted theory in the contact lens field is that surface characteristics of a lens will dictate its movement on the eye. It has been our experience that surface treatment of a lens is not enough to overcome the bulk polymers

physical performance with regard to movement. If a particular lens material did not move on the eye, application of various hydrophilic and wettable surfaces did not change the mobility on eye.

ON PHASES

Phase morphology can take many forms. Silicone hydrogels disclosed in patents are random, statistical polymers that can be best represented as isotropic structures, either as a dispersed system (only one phase is continuous) or as interpenetrating cocontinuous polymer phases (two phases continuous throughout the material) similar to an interpenetrating polymer network (IPN). Anisotropic structures typical of highly ordered systems are unlikely to work in contact lens application. One possible exception may be the polymer systems disclosed by Hirt, which may form mesophases.^[43]

A model analysis evaluation of phase-separated systems, using the parameters of phase volume fraction^[59] (Fig. 1), and the connectivity factor ξ ^[60] (Fig. 2), yields models allowing the prediction of oxygen permeability theoretically achievable (Fig. 3). Based on the analogy^[60] between the integrated Fourier law and Ohm's law, such calculations offer two boundary solutions, one representing the serial, the other the parallel arrangements of the conductors. Both solutions are shown in Fig. 3 as solid lines. The broken line in between represents the geometrical mean of those two solutions. As can be deduced from such calculations, a maximum oxygen permeability of 225 barrer could be obtained at 50% volume fraction of the PDMS phase, assuming the ideal connectivity $\xi = 1$. With the less ideal (but more realistic) $\xi = 0.5$, and the same PDMS volume fraction of 50%, the predicted Dk is still as high as 194 barrer. It becomes obvious from such model calculations that a material with dispersed, thus unconnected, oxygen transport phases would fall in its Dk value, far below the requirement for extended wear contact lenses (Fig. 3, right diagram). Furthermore, one can easily conclude from such simple model considerations that a complete phase separation (least amount of interphase regions) is essential for obtaining high oxygen permeabilities. Also, the continuity of the oxygen transport phase should be high ($\xi = 0.5$ to 1) for maintaining $Dk > 100$ barrer at a 50% volume fraction of PDMS. A similar treatment can be used to predict

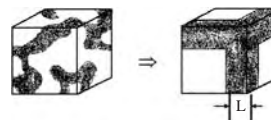


Fig. 1 Finite element concept two phases, one parameter: volume fraction.

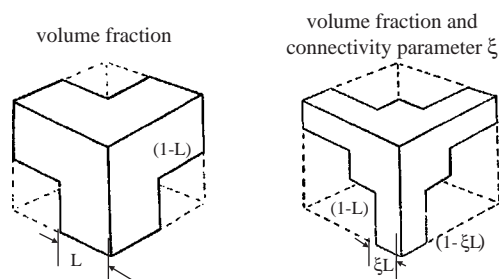


Fig. 2 Refined model showing connectivity.

the ion permeability. From such simple model treatment, one can predict the theoretical maximum oxygen permeability of a perfectly phase-separated (no inter-phase regions) silicone hydrogel contact lens, which will move on the eye. Achievement of such a polymer system is a goal for future polymer research.

The preparation of silicone or fluorosilicone hydrogel polymers with cocontinuous phases, in our experience, can be achieved either by careful design of the macromer architecture or by attention to polymerization kinetics, especially in polymers with relatively low water contents (<30%). Such systems are particularly advantageous for extended-wear contact lens polymers, as the lower water content allows for higher oxygen permeability to be achieved. Amphiphilic segmented macromers or prepolymers,^[43] composed of hydrophilic and hydrophobic segments bonded together, can provide spontaneous cocontinuous products. In polymer systems prepared from all-hydrophobic macromers, copolymerized with hydrophilic monomers such as DMA, inhibition of polymerization by oxygen has been shown to have a dramatic influence on continuous ionophase formation. Using ion permeability as a diagnostic

for the presence of a continuous hydrophilic phase, we have shown that continuity is achieved only when oxygen has been effectively excluded from the polymerizing mixture.^[51] However, differential scanning calorimetric (DSC) analysis shows no difference between two polymers cured in the presence and absence of oxygen, implying that two phases still exist. Apparently, kinetics of the complex competing polymerization reactions are sufficiently different in presence or absence of oxygen to hinder or allow the formation of the cocontinuous phases.^[47]

THE SURFACE OF SILICONE HYDROGEL CONTACT LENS MATERIALS

Neither silicones, nor fluorocarbon compounds, are hydrophilic, and silicones are highly lipophilic, neither of which are advantageous properties for the surface of a contact lens. The need for surfaces with excellent wettability and low deposition potential are well-understood in the contact lens field.^[61]

In soft hydrogel polymers, siloxane and fluorocarbon chain segments have the ability to move in the polymer matrix. Given the opportunity, these groups will migrate toward, and be enriched at, the lens surface in contact with air, which is hydrophobic, because “like-attracts-like,” thus lowering the surface energy. Siloxanes and fluorocarbons, lying at the lens-air interface, are energetically preferred over being immersed in the rather hydrophilic core of the hydrogel lens polymer. Siloxane chain segments in particular, enriched at a lens surface, will have a strong propensity to attract lipids from the tear film during lens wear. Little or no tear film stability, excessive deposits, discomfort, and poor vision would be the result of wearing a soft silicone hydrogel lens that

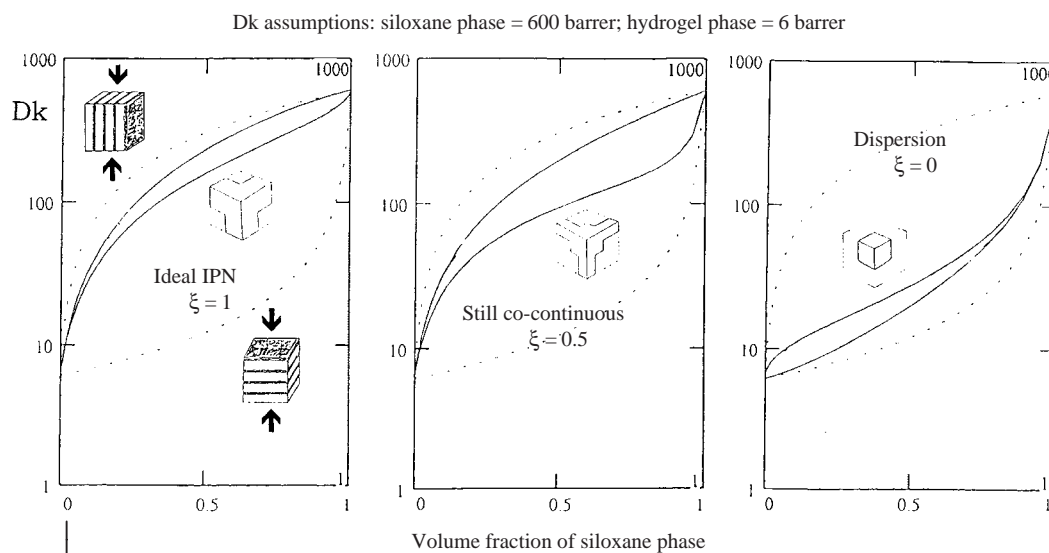


Fig. 3 Oxygen permeability prediction models.

does not employ some means of rendering the surface of the lens more hydrophilic than the core polymer.

This problem can also be considered an opportunity. In the case of conventional hydrogel contact lenses, the surface character of the lens is identical to the character of the core material, which is hydrophilic, and is thus easily wettable by water and tear fluid.

The necessity for some surface masking of the hydrophobic/lipophilic nature of silicone hydrogels and fluorosilicone hydrogels creates the opportunity to separately design and optimize the surface of the lens to achieve maximum ophthalmic compatibility. This is particularly important for silicone hydrogel contact lenses, since they are designed to provide high oxygen throughput to the cornea for overnight and continuous wear.

One of the major challenges facing researchers developing surface treatment technologies is to keep the desired changes at the surface of the lens material and not perturb the bulk properties. That is, one wishes not to alter or compromise the desired properties designed into the core material at the start.

Methods for surface treatment

Numerous methods for achieving and enhancing compatibility of silicone hydrogel contact lenses have been disclosed in patents and reported in literature. All the methods share one primary goal; that is to achieve a lens surface with minimum deposition of lipids and proteins, and one which will support a stable tear film on the anterior surface of the lens during wear.

Before a discussion of various means of surface treatments is offered, it is important to point out that the relative costs of the alternative surface treatment processes must be carefully considered by the researcher and the manufacturer. It is the nature of contact lenses that they are a challenge to handle individually in high volume and high yield. Most surface treatments require additional process steps in the overall manufacture of silicone hydrogel lenses, thus the need for careful cost considerations.

It is interesting to note that a review of early silicone hydrogel contact lens material patents seldom mention the need for surface treatment. It was believed that because the polymers absorbed water, they would be hydrophilic at their surface. During the 1980s, several patents appeared disclosing silicone hydrogel polymer materials that would be wettable. Approaches, including the inclusion of surface-active monomers,^[62] the incorporation of latent hydrophilic monomers, and the use of polar mold materials in the forming of the lens^[63] were disclosed. These very attractive, low cost means of achieving a wettable silicone hydrogel contact lens apparently did not deliver the desired results, as no commercial products ever emerged.

Plasma

Plasmas are considered to be the fourth state of matter. They can be described as a highly reactive gas, capable of reacting with any surface present. The gas is partially ionized, and thus rendered reactive by a strong, high-frequency electric field applied to it. Plasmas most often are employed under reduced pressure. Reactions in the ionized gas, and at the exposed surfaces, depend upon many factors, the most important being the composition of the gas, the electric power applied to the gas plasma reactor, the nature of the surface, the pressure in the reactor, and the frequency of the oscillating electric field. For contact lens silicone containing polymers, many different outcomes can be achieved. They are listed as follows:

1. Coating of the surface by compounds derived from the component(s) of the gas mixture
2. Etching of the surface manifested as a roughening of the surface
3. Ablation of the surface
4. Change of the chemical nature of the elements already present in the surface
5. Creation of chemically active sites on a surface; sites where subsequent reactions can occur

Illustrative examples of each of the above outcomes are:

- a. Coating of the surface with compounds containing carbon, and oxygen when the gas contains oxygen and methane, rendering the surface more hydrophilic.
- b. A loss of surface material and an observed roughening of the surface. The localized rate of etching can vary depending upon the local surface composition. Oxygen gas is often employed to achieve surface etching. A rough surface can be more wettable than a smooth surface. Higher energy electric fields are usually applied.
- c. A loss of surface material much like (b) above. Inert gases such as helium are sometimes used to achieve ablation in combination with higher energy electric fields.
- d. Conversion of chemical groups in the original surface. For example, conversion of hydrocarbon ($-\text{CH}_2-$) groups to $-(\text{C}-\text{OH})-$ or $-(\text{C}=\text{O})-$ containing groups which are more hydrophilic. Oxidizing plasma gases are generally used at higher energy to achieve such chemical conversions.
- e. This is a special case where a plasma reactor can be used to chemically activate the surface of a contact lens material. In this case, the activated lens is retrieved from the reaction chamber and subsequently treated with some hydrophilic chemical entity capable of reacting with that activated

surface. The reaction product is a lens with a more hydrophilic surface.

The authors hasten to point out that in most plasma processes more than one of the above a–e events occurs. For example, some change in the chemical nature of the substrate lens material may occur while a coating is being deposited.

The specific chemical reactions that take place in a plasma reactor are in most cases very complex and are not easily defined or characterized. It is up to the researcher to assure oneself that what has been achieved is safe, non-toxic, stable, durable, and reproducible. One of the particular advantages of plasma-based processes is that they are very reproducible when conditions in the reactor are maintained constant, the reactor is well maintained, and that the lens material being fed into the reactor is itself reproducible and constant.

While plasma systems that are appropriate for the high production volumes in commercial contact lens manufacturing require large initial capital investment, particularly when they are automated, the cost of operating the systems over time, both in dollars and in environmental costs, are competitive with other processes.

Chemical Coatings

In the patent literature, manufacturers and researchers have disclosed many methods of applying coatings to silicone hydrogel surfaces. One such process includes chemically linking a preformed hydrophilic polymer, for example poly-HEMA or PEG, to the surface using a class of chemical reagents known as carbodiimides.^[64] The preformed hydrophilic polymer can be applied to the lens substrate using methods such as lens immersion in a solution of the polymer, spraying, spin-coating etc. In some cases, patents disclose the preparation of unique hydrophilic polymers specifically designed by the researcher to deliver a desired surface quality.^[65]

Another method disclosed involves sequential adsorption onto the lens surface alternating cationic and anionic hydrophilic polymers.^[66] This is achieved by immersion of the lens into a series of solutions of hydrophilic polymers. Each polymer solution in the series bears opposite charges. That is, the first solution would be of an anionic polymer, the second would contain a cationic polymer, and so on. This process is known as the layer-by-layer method.

A third method of applying a preformed hydrophilic polymer is a method generally known as transfer grafting. In this procedure, the preformed polymer is applied to interior surfaces of the plastic molds used to form the lens. This application requires great care that the entire forming surfaces of the molds are uniformly coated.^[67] Then the monomer mixture that is to be

polymerized into the lens core is filled into the coated mold. The mold halves are brought together and polymerization commenced. As the lens polymer is formed in the coated mold, the preformed hydrophilic polymer on the mold surfaces is incorporated onto the surface of the formed lens.

A second general class of chemical coatings involves the initiation and growth of a new hydrophilic polymer directly from the surface of the substrate silicone hydrogel lens. This process, known as surface grafting, is challenging in that penetration of the hydrophilic monomers into the core of the lens substrate is difficult to control. If such penetration occurs, the properties and dimensions of the substrate silicone hydrogel can be significantly altered.

Surface Treatment from within the Core

A recent patent granted to J&J discloses a method of adsorption of a hydrophilic polymer from within the core lens that has been preloaded with the hydrophilic polymer.^[68] In this particular disclosure, an “internal wetting agent,” high molecular weight PVP, is incorporated into the initial blend or formulation of monomers used for the lens fabrication. The formulation, or blend, is polymerized, however the PVP may not be bonded or incorporated into the network silicone hydrogel formed. Rather the PVP may be “entangled” in the network, forming a kind of interpenetrating polymer network (IPN). It is suggested by the inventors that when the lens polymer is slowly cured, hydrated, and packaged in saline, the PVP may be able to migrate in the cross-linked network of the silicone hydrogel and adsorb onto the lens surface and provide wettability.^[68] A commercial conventional hydrogel, vifilcon A, also contains preformed PVP. Those lenses are known for their very slippery surface and that is due to the same phenomenon, that is, PVP migrating to the lens surface.

One can appreciate the economic attractiveness of such a treatment process from within the lens core. No additional processing steps for surface treatment are required.

In summary, as of today, silicone hydrogels require some surface treatment. It remains a major goal, to develop an inherently wettable silicone hydrogel polymer having high oxygen permeability.

CURRENT COMMERCIAL SILICONE HYDROGEL CONTACT LENSES

There are five commercial silicone hydrogel contact lenses available on the international market today. They are lotrafilcon A, as FOCUS Night&Day™ from CIBA Vision, balafilcon A as PureVision™ from Bausch and



Table 3 Commercial siloxane hydrogel contact lenses

Name	Night&Day	PureVision	Acuvue Advance	O2Optix	Acuvue Oasys
Manufacturer	CIBA vision	Bausch&lomb	J&J vistakon	CIBA vision	J&J vistakon
USAN	Lotrafilcon A	Balafilcon A	Galyfilcon A	Lotrafilcon B	Senofilcon A
%Water	24	36	47	33	38
Dk	140	99(92 ^a)	60	110	103
Modulus	1.2 Mpa	1.1 Mpa	0.5 MPa	1.2 MPa	0.7 MPa
Polymer	Acrylate	Vinyl	Acrylate	Acrylate	Acrylate
Surface	Plasma deposition	Plasma oxidation	Entangled adsorbed PVP	Plasma deposition	Entangled adsorbed PVP
Oxygen transport enhancers	Difunctional fluorosiloxane macromer, TRIS	TRIS-like monomer, difunctional siloxane macromer	SiGMA monomer, monofunctional siloxane macromer	Difunctional fluorosiloxane macromer, TRIS	SiGMA monomer, monofunctional siloxane macromer
Hydrophilic components	DMA	NVP amino acid monomer	DMA HEMA PVP	DMA	DMA HEMA PVP

ND, No data.

^aCurrent *Dk* value.

Lomb, galyfilcon A as Acuvue Advance™ from Johnson & Johnson Vision Care, and most recently, lotrafilcon B as O₂Optix™ and senofilcon A as OASYS™ from CIBA Vision and Johnson & Johnson Vision Care, respectively. Some of the properties of these four products are listed in Table 3.

While the exact compositions of the core polymers and surfaces of these lenses are proprietary, general information is available. Patent disclosures offer additional insight.

Lotrafilcon A—the Core Polymer

This material has the highest oxygen permeability of the five materials, due likely, at least in part, to the lowest water content. It is reported to be composed of a fluorosiloxane difunctional macromer, DMA and TRIS. The difunctional macromer serves as a crosslinker in the polymer, thus the relatively high reported modulus. The polymer morphology of lotrafilcon A is reported to be biphasic, having fluoro-silicone rich pathways extending throughout the material, allowing facile transport of oxygen. The second phase is then hydrophilic and is also extending throughout the material.^[69] The hydrophilic phase allows the transport of ions and water from surface to surface, allowing for movement of the lens on the eye.

Balafilcon A—the Core Polymer

The patents claimed to cover this material disclose a different polymer system known as vinyl carbamate and vinyl carbonate polymers. It is sufficient to say here that

the monomers used to form vinyl polymers have carbon-carbon double bonds that are linked differently to the pendant groups than most all other monomers used in contact lens materials. Almost all other contact lens polymers are formed from acrylate carbon-carbon double bond containing monomers. A recent U.S. patent granted to B&L describes balafilcon A as containing a TRIS-like monomer, NVP, a bifunctional siloxane monomer and an amino acid derived hydrophilic monomer.^[65] This last ingredient is responsible for the classification of balafilcon A in FDA group III, namely as a low water, ionic polymer. Note the higher modulus of this material, again due to the siloxane crosslinking macromer. Balafilcon A has been reported to be a homogeneous polymer material.^[70] However, the high reported water permeability (twice the value of poly-HEMA) and the high *Dk* would tend to argue otherwise. Such high permeabilities would not be anticipated for a homogeneous silicone hydrogel. It is oxygen transport through the above described silicone phase which allows the attainment of high permeability.^[3]

Galyfilcon A—the Core Polymer^[71]

The polymer is composed of the monomer identified as SiGMA. Recent patents state SiGMA improves the compatibility of this monomer with hydrophilic monomers and the high molecular weight PVP and provides improved mechanical properties of the final polymer.^[38] The other monomers present include a monofunctional silicone macromer, DMA, HEMA and a crosslinker, ethyleneglycol dimethacrylate (EGDMA). The PVP is stated to be the “internal wetting agent” (termed

HYDRACLEAR™ in promotional material from J&J responsible for the wettability of the lens. Note the silicone macromer is monofunctional. This may be the primary reason why the galyfilcon A lens has a lower modulus than the other materials. No reports or literature have surfaced concerning the phase morphology of galyfilcon A.

Lotrafilcon B—the Core Polymer

Since this polymer shares the same generic name with lotrafilcon A, having only a different suffix, one can assume the same qualitative composition for both materials. The “B” indicates the materials differ in the amounts of the components present. The lower modulus measured on lotrafilcon B is a strong indication the amount of difunctional macromer in this material is less than that in lotrafilcon A. Because the macromer is hydrophobic, the lesser amount present has the predictable effect of higher water content (Table 3). The biphasic morphology is also anticipated to be present.

Senofilcon A—the Core Polymer^[72]

This is the newest silicone hydrogel-based contact lens to come to the market. The qualitative composition is identical to galyfilcon A, with the exception that the crosslinker is changed from EGDMA to the longer chain TEGDMA (tetra(ethyleneglycol) dimethacrylate), likely rendering the polymer slightly softer.

FOCUS Night&Day™ and O₂Optix™—the Surface

The surface applied to lotrafilcon A and lotrafilcon B to make FOCUS Night&Day™ and O₂Optix™, respectively, is reported to be a plasma-derived coating covalently and permanently bound to the lens polymer substrate.^[73] This process and outcome would fit best under subsection (a), under Plasma, above. Published images of the surface from atomic force microscopic (AFM) analysis reveal the surface to be smooth and homogeneous. Only insignificant scratches originating from tool fabrication are visible. The surface is reported to be low depositing.

PureVision™—the Surface

The surface of balafilcon A core material lenses is reported to be the result of partial plasma oxidation of the substrate polymer siloxanes to silicate. Silicate materials are hydrophilic. This process would be best described as falling under subsection (d), under Plasma, above. Published images and a patent clearly show the PureVision™ lens surface to be a surface of fragmented glassy islands.

Wettability is reported to be achieved as water “bridges over” the fissures between the glassy islands.^[73] As mentioned above under the plasma discussion, the plasma processes are usually carried out under partial vacuum; therefore, the lens is dry when coated. Perhaps the fissures develop as a result of the expansion of the substrate hydrogel as it swells with water to equilibrium.

Acuvue Advance™ and OASYS™, the Surface^[38]

The surface is enriched in PVP polymer by migration from within the core and adsorption on the lens surface.

CONCLUSION

There have been many exciting new innovations in the field of contact lens materials. The proliferation of these new silicone hydrogel products are likely only the beginning of a new generation of lenses that may render obsolete conventional hydrogels as technological advances continue to evolve. A major area of research activity, based on patents, involves lenses that diminish or eliminate bacterial growth on lens surfaces. Should these efforts prove successful, a greater margin of safety will be enjoyed by the contact lens wearer.

REFERENCES

1. Watanabe, J.; Kiritoshi, Y.; Woo Nam, K.; Ishihara, K. *Hydrogels. Encyclopedia of Biomaterials and Biomedical Engineering*; Marcel Dekker: New York, 2004; 790–801.
2. Nicolson, P.; Vogt, J. Soft contact lens polymers: an evolution. *Biomaterials* **2001**, *22* (24), 3273–3283.
3. Lai, Y.; Wilson, A.; Zantos, S. Contact Lens. In *Kirk-Othmer Encyclopedia of Chemical Technology*, 4th Ed.; Wiley: New York, 1993; Vol. 7, pp. 191–218.
4. Gaylord, N. US Patent 3,808,178, 1974.
5. Gaylord, N. US Patent 3,808,179, 1974.
6. Gaylord, N. US Patent 4,120,570, 1978.
7. Birdsall, A.; et al. US Patent 4,198,131, 1980.
8. Mitchell, D. US Patent 4,487,905, 1984.
9. DeCarle, J. Developing hydrophilic lenses for continuous wearing. *Aust. J. Optom* **1972**, *55*, 343–346.
10. Reuben, M.; Brown, N.; Lobascher, D.; Chaston, J.; Morris, J. Clinical manifestations secondary to soft lens wear. *Br. J. Ophthalmol.* **1976**, *60*, 529–530.
11. Poggio, E.; Schein, O.; Kenyon, K.; Glynn, R.; Seddon, J. *The Risk of Microbial Keratitis Among Cosmetic Soft Contact Lens Wearers: An Incidence Study and a Case-Control Study*; The Contact lens Institute: Washington, DC, 1998.
12. Poggio, E.; Glynn, R.; Schein, O.; et al. The incidence of ulcerative keratitis among users of daily-wear and extended-wear soft contact lenses. *N. Engl. J. Med.* **1989**, *321*, 779–783.



13. Schein, O.; Glynn, R.; Poggio, E.; Seddon, J.; Kenyon, K. The relative risk of ulcerative keratitis among users of daily-wear and extended-wear soft contact lenses. *N. Engl. J. Med.* **1989**, *321*, 773–778.
14. F.D.A. Practitioner Letter, Ulcerative keratitis and contact lens wear, May 30, 1989.
15. Schein, O.; Buehler, P.; Stamler, J.; Verdier, D.; Katz, J. The impact of overnight wear on the risk of contact lens associated ulcerative keratitis. *Arch. Ophthalmol.* **1994**, *112*, 186–190.
16. Hamano, H.; Watanabe, K.; Hamano, T.; Mitsunaga, S.; Katani, S.; Okada, A. A study of complications induced by conventional and disposable contact lenses. *CLAO J.* **1994**, *20*, 103–108.
17. Boswall, G.; Ehlers, W.; Luistro, A.; Worrall, M.; Donshik, P. A comparison of conventional and disposable extended wear contact lenses. *CLAO J.* **1993**, *19*, 158–165.
18. *U.S.P. Dictionary of USAN and International Drug Names*; U.S. Pharmacopeia: Rockville, MD, 2000; 259.
19. Professional Edition; *Tyler's Quarterly Soft Contact Lens Parameter Guide*; Tyler's Quarterly, Inc.; Little Rock, 2000; Vol. 18, 29.
20. Holden, B.; Mertz, G. Critical oxygen levels to avoid corneal edema for daily and extended wear contact lenses. *Invest. Ophthalmol. Vis. Sci.* **1984**, *25*, 1161–1167.
21. Lai, Y. The role of bulky polysiloxanyl methacrylates in oxygen-permeable hydrogel materials. *J. Appl. Polym. Sci.* **1996**, *56*, 317–324.
22. Robertson, R.; et al. US Patent 5,070,169, 1991.
23. Kunzler, J.; Ozark, R. US Patent 5,321,108, 1994.
24. Vanderlaan, D.; et al. PCT Application WO 01/27174, 2001.
25. Chang, H. US Patent 4,182,822, 1980.
26. Mueller, K.; Kleiner, E. US Patent 4,136,250, 1979.
27. Tanaka, K.; Takahashi, K.; Kanada, M.; Kanome, S.; Nakajima, T. US Patent 4,139,513, 1979.
28. Deichert, W.; Su, K.; van Buren, M. US Patent 4,153,641, 1979.
29. Keogh, P.; Künzler, J.; Niu, G. US Patent 4,259,467, 1981.
30. Rice, D.; Ihlenfeld, J. US Patent 4,440,918, 1984.
31. Rice, D.; Ihlenfeld, J. US Patent 4,818,801, 1989.
32. Isaacson, W. President of 3M Vision Care, Minneapolis, MN, private communication.
33. Salame, M.; Steingeiser, S. Barrier polymers. *Polym. Plast. Technol. Eng.* **1977**, *8*, 155–175.
34. Molock, F. US Patent 5,039,769, 1991.
35. Su, K.; Robertson, R. US Patent 4,740,533, 1988.
36. Alli, A.; et al. PCT Application WO 01/30867, 2001.
37. Vanderlaan, D.; et al. PCT Application WO 01/27174, 2001.
38. McCabe, K.; et al. US Patent Application 2004/6822016, 2004.
39. Mueller, K.; Heiber, S.; Plankl, W. US Patent 4,486,577, 1984.
40. Lai, Y.; Baccei, L. US Patent 5,034,461, 1991.
41. Goldenberg, M. US Patent 4,929,692, 1990.
42. Hirt, T. US Patent 6,039,913, 2000.
43. Keogh, P.; Künzler, J.; Niu, G. US Patent 4,260,725, 1981.
44. LeBoeuf, A. US Patent 4,261,875, 1981.
45. LeBoeuf, A. US Patent 4,294,974, 1981.
46. Künzler, J.; Ozark, R. Hydrogels based on hydrophilic side-chain siloxanes. *J. Appl. Polym. Sci.* **1995**, *55*, 611–619.
47. CIBA Vision; unpublished results.
48. Bambury, R.; Seelye, D. US Patent 5,070,215, 1991.
49. Lai, Y. US Patent 5,310,779, 1994.
50. Ichihara, M. US Patent Application 2004/0014898, 2004.
51. Nicolson, P.; Baron, R.; Chabreck, P.; et al. US Patent 5,760,100, 1998.
52. Gido, S. Connectivity of domains in microphase separated polymer materials. In *New Polymer Materials*; Kourgic-Karasz L.S., MacKnight, W.J., Mortuscelli, E., Eds.; A.C.S. Symposium Series; Oxford University Press: Oxford, London, 2005; 309–325.
53. Nicolson, P.; et al. US Patent 5,776,999, 1998.
54. Nicolson, P.; et al. US Patent 5,789,461, 1998.
55. Nicolson, P.; et al. US Patent 5,848,811, 1998.
56. Nicolson, P.; et al. US Patent 5,965,631, 1999.
57. Domschke, A.; Lohmann, D.; Winterton, L. On-eye mobility of soft oxygen permeable contact lenses. Proceedings of the ACS Spring National Meeting, San Francisco, PMSE; 1997.
58. Tighe, B. Silicone hydrogel materials—how do they work?. In *Silicone Hydrogels: the Rebirth of Continuous Wear Contact Lenses*; Sweeney, D., Ed.; Butterworth, Heinemann: Oxford, 2000; 1–21.
59. Takayanagi, M.; Harima, H.; Iwata, Y. *J. Soc. Mater. Sci. Jpn.* **1963**, *12*, 389.
60. Marcussen, L. In *Thermal Conductivity*; Ashworth, T., Smith, D., Eds.; Plenum: New York, 1985; 585.
61. Nicolson, P. Continuous wear contact lens surface chemistry and wearability. *Eye Contact Lens* **2003**, *29*, S30–S32.
62. Lai, Y.; et al. US PCT Application WO 92/09644, 1992.
63. Lai, Y. PCT Application WO 92/09421, 1992.
64. Vanderlaan, D.; et al. US Patent US 6,087,415, 2000.
65. McGee, J.; et al. US Patent 6,599,559, 2003.
66. Winterton, L.; et al. US Patent 6,451,871, 2002.
67. Turner, D.; et al. US Application 2003/005242, 2003.
68. Maiden, A.; et al. US Patent 6,367,929, 2002.
69. Jones, L.; Tighe, B. Silicone hydrogel contact lens materials update—Part 1, On the Silicone Hydrogels web site at www.siliconehydrogels.org/editorials/index.asp, July 22, 2004.
70. Jones, L.; Dumbleto, K. Silicone hydrogel contact lenses, Part 1: Evolution and current status, *Optometry Today*, 2002, September 20, 26–31.
71. American Medical Association (USAN) USAN Assigned to Hydrophilic Contact Lens Materials, available at http://www.ama-assn.org/ama1/pub/upload/mm/365/galyfilcon_a.doc, 2005.
72. American Medical Association (USAN) USAN Assigned to Hydrophilic Contact lens Materials, available at http://www.ama-assn.org/ama1/pub/upload/mm/365/senofilcon_a.doc, 2005.
73. Jones, L.; Tighe, B. Silicone hydrogel contact lens materials update—Part 2, On the Silicone hydrogels, Web site available at www.siliconehydrogels.org/editorials/index.asp, 2005.

Control of Movement

Dejan B. Popović

Department of Health Science and Technology, Aalborg University, Aalborg, Denmark,
and Faculty of Electrical Engineering, University of Belgrade, Belgrade, Serbia



INTRODUCTION

The most remarkable attribute of movement is how easy it is. It is only when we face a sensory-motor disability that we are reminded of the problems associated with movement control with which our nervous systems copes so uncomplainingly. This becomes even more of a problem when we are faced with the task of designing an assistive system that would allow a disabled human to function like an able-bodied person in everyday activities. The biological control developed through the process of self-organization is impressive. This biological control is capable of encompassing the redundancy and richness of the anatomy. The basic concept presented here is that man-made control has to mimic as well as integrate into the biological control. The man-made control of movement should maximize the utilization of preserved sensory-motor mechanisms. The control has to consider the changes that have occurred to the sensory-motor systems as a result of a disability. The command interface between the biological and man-made control must be simple and intuitive, possibly at the subconscious level. The man-made control that we are describing here is based on the current understanding of life sciences and the technology that is available.

To learn more about the topics in this entry we recommend the book *Control of Movement for the Physically Disabled*.^[1] To learn in depth about neural control we suggest the book *Principles of Neural Science*.^[2] The biomechanics and motor control details can be found in *Biomechanics and Neural Control of Posture and Movement*.^[3]

BIOLOGICAL CONTROL OF MOVEMENT

The motor system in humans comprises three inter-related anatomical systems: the skeletal system; the muscle system, which supplies the power to move the skeleton; and the nervous system, which directs and regulates the activities of muscles. Man uses about 350 pairs of skeletal muscles of many shapes and sizes that are situated across more than 300 joints, being attached to bones via tendons. A shortening of the

muscle brings the muscle ends closer to each other, which changes the angle at the joint across which it acts. A muscle has three characteristic properties: contractility, excitability, and elasticity. Contractility is the capacity of the muscle to produce a force between its ends; relaxation is the opposite of contraction, which is entirely passive. Excitability is the ability of the muscle tissue to respond to stimulation. Elasticity relates to the ability of the muscle to be stretched by an external force and return to its original length. The muscle suffers no harm so long as it is not stretched in excess of its physiological limits.

A skeletal muscle is composed of two types of structural components: active contractile elements and inert compliant materials. The contractile elements are contained within the muscle fibers. Each muscle is composed of many muscle fibers, of which a medium-sized muscle contains approximately one million. The fibers vary in length from a few millimeters to more than 20 cm, and their width is between 50 and 150 μm . About 85% of the mass of a muscle consists of muscle fibers while the remaining 15% is composed largely of connective tissues, which contain variable proportions of collagen, reticular, and elastic fibers.

The muscle activation is linked to the activity of neural cells. The essential function of a nerve cell is to transmit excitation to other cells. This is achieved by releasing a chemical transmitter substance at its synaptic terminal. Many different kinds of stimuli may excite the neurons. The neurons normally carry trains of impulses. The motor neurons usually conduct at frequencies of 5–30 impulses per second although sometimes at higher rates. The upper-limit frequencies for sensory neurons are normally between 100 and 200 impulses per second. The motor neurons conduct at velocities between 50 and 80 m/sec. The nerves enter the muscle near the main arterial branch and divide to distribute both motor and sensory fibers to the muscle bundles. One motor neuron and all of the muscle fibers that are innervated by axon terminals constitute a motor unit. Small muscles and the muscles that are involved in the fine gradations of a contraction necessarily have smaller motor units compared with larger bulky muscles that are involved in the maintenance of a strong contraction.

The excitation of a muscle (contraction) can be described as follows. The nerve impulses arriving at the neuromuscular junction cause the release of a transmitter substance, which diffuses across the junction and chemically excites the muscle fiber. The action potentials travel along the fiber membrane at a speed of 1–5 m/sec and initiate the events that lead to the shortening of the contractile elements of the myofibrils and the consequent production of force in the muscle.

Movement results from the interaction of muscular force, gravity, and other external forces that impinge on skeletal levers. The muscles rarely act singly; rather, groups of muscles interact in several ways for the desired movement to occur. Whenever a muscle causes movement by shortening, it is functioning as a mover or agonist. If an observed muscle constitutes the major contribution to a movement, it is that muscle that is called the prime mover. Other muscles crossing the same joint on the same aspect, but that are less relevant for the movement, are termed secondary or assistant movers or agonists. The muscles that are in opposite action to the prime mover are called antagonists.

The sine qua non of the effective control of movement involves greater stabilization of the more proximal joints in order to provide greater freedom to the distal segments to move effectively. The greater the amount of force to be exerted by the open end of the system (whether it is the peripheral end of an upper or lower extremity), the greater the amount of stabilizing force that is needed at the proximal joints.

The natural control systems for human movements have a hybrid hierarchical structure and organization (Fig. 1). The control systems have a component that is

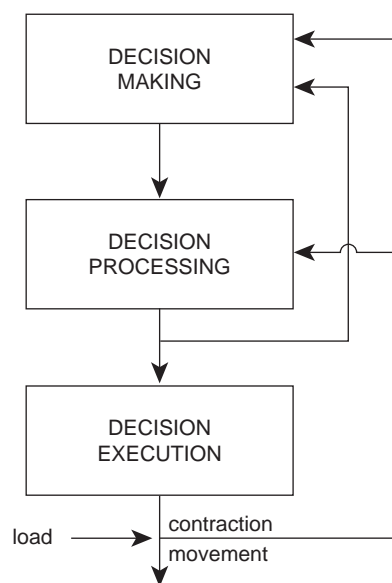


Fig. 1 Hierarchical hybrid model of control of movement in humans.

centrally programmed. For example, the entire act of signing a letter requires only one conscious thought: “Write the name.” The entire sequence of muscle activities and movements that follows occurs entirely automatically although the action can be changed at any time. Thus, it is possible to distinguish the *idea* (decision making in Fig. 1) from the *plan* that is used to carry it out. The movement plans are stored in the cortical parts of the central nervous system. The plans are developed mostly in childhood through a trial and error procedure, but they also develop later in life, especially when facing the loss of some sensory-motor mechanisms. The plans are likely composed of primitives (simple movement plans), which are placed into an integrated structure during each execution that is based on the visual or other information on the relation between the internal and external reference frames. The control systems also use feedback (arrows in Fig. 1).

The exteroceptors provide the control systems with information about the spatial coordinates of the objects. The proprioceptors relay information about the position of the body vs. the vertical, the angles of the joints, the length and tension of the muscles, etc. Via the proprioceptors, the control systems gain access to information about the condition of the muscles and skeletal parts that have to be moved. The control systems need information about the consequences of the motor systems’ actions. The sensory systems provide this information, which is thereafter used to calibrate the next series of motor commands. Thus, motor mechanisms are intimately related to and functionally dependent on central programming and sensory feedback.

Control of posture: Posture is characterized by the equilibrium of gravity and muscle generated forces. The multisegmental system has to be supported in an erect position with a coordinated activity of muscles at the lowest levels possible. Most of the body segments behave as inverted pendulums in this motor task. The control systems have to generate commands that ensure a slow oscillation around a metastable position where body segments are aligned to the direction of gravity. Many sensors are integrated into this task: vestibular system, somatosensory systems, muscle spindles, Golgi tendon organs, joint receptors, etc. The task is to keep the direction between the centers of mass and pressure (point where the resultant ground reaction force acts at the sole) aligned with the vertical. The center of mass is called the center of gravity because the external forces are gravity forces. Therefore, the center of pressure belongs to the central zone of the support surface during stable standing. The small region where the position of the center of gravity is located, to guarantee the stable posture of the body, is called the stability zone.

Control of walking: Walking follows the voluntary decision to progress from one point to another in space as well as to maintain vertical posture. The first element in the planning of locomotion in able-bodied subjects is a visual cue. Walking is feasible without visual cues (e.g., walking in a dark room), but it is then based on the anticipation of no obstacles and no change in the contact surface. Good performance requires knowing the environment, which is rarely the case in normal life. A hallmark of successful and safe walking is the ability to adapt the basic gait patterns to meet environmental demands. The visual system allows a human to make an anticipatory adjustment to the walking pattern necessary for overcoming environmental constraints.

Bipedal walking is a cyclic activity. The prerequisite for bipedal walking is the ability to maintain posture. Walking occurs once the equilibrium required for stable posture ceases to exist owing to the changes in internal forces (muscle activities). The internal forces cause the center of gravity to shift outside the stability zone, in which the body then starts falling. The falling of the body is prevented by moving one leg in the direction of the fall, thereby providing new support and a new stability zone. After the moved leg has safely supported the body, the contralateral leg pushes the body up and forward and generates momentum, resulting in a progression of the center of gravity. The cyclic repetition of the above-described events is defined as walking. In short, the essentials for walking are: 1) provision of support to the body segments against gravity; 2) production of a locomotor rhythm, 3) production of muscle forces that will result in the friction force (between the foot and the ground) required for propelling it to the intended direction; 4) dynamic equilibrium of the moving body; and 5) adaptation of movements to meet the environmental demands and the tasks selected by the individual.

The model that is often used for describing locomotion is based on an assumption of the existence of the central pattern generator (CPG). A CPG is a neuronal network capable of generating a rhythmic pattern of motor activity in the absence of sensory input from peripheral receptors. The CPG has been identified and analyzed in many animals, and has been suggested, with great reservation, in humans.

Control of goal-directed movement: Goal-directed movement is a planned change of the body segments' positions, which ultimately leads to a task. The nervous system solves the following three tasks during goal-directed movements: 1) multimodal sensory integration or sensory intake; 2) transformations between different spatial and intrinsic coordinate systems; and 3) motor commands adequate to attain the objective. Goal-directed movement is reliant upon a balance of initial programming and subsequent correction. Initial

programming is based partly on the visual perception of objects and partly on proprioception (e.g., a visual cue is used to decide whether to pick up the object with one or two hands, and what type of grasp to use). The accuracy of visual perception determines the initial programming. The visual information is used to identify a target and its location in space, and also for corrections of ongoing movement. Goal-directed movement for a seen object benefits from visual feedback.^[4] Here, we will only concentrate on goal-directed movements of the upper extremities.

Biological control uses task dependent constraints to reduce the number of variables that have to be controlled voluntarily to the lowest amount possible by forcing other variables to follow the memorized schemes. This method was called synergistic control, and was first analyzed and described by Bernstein.^[5] Bernstein suggested that the problem of motor redundancy is solved using flexible relations (synergies) among the control variables that are directed at individual elements. Synergy can be found between individual joint movements, thus the dimensionality of the control is greatly reduced.^[6] Recent observations have indicated that humans' reaching movements are characterized by certain stereotypical kinematic features such as straight hand paths and bell-shaped velocity profiles, thus suggesting that the generated trajectories are not selected at random, but in a rather preferred way according to some general criteria or organizing principles.^[7] Humans can learn to make a virtually unlimited set of trajectories with widely varying levels of cocontraction. Thus, the nervous system must be able to control muscles almost individually. However, humans often do not use the full repertoire of possible trajectories, but produce movements that have certain invariant properties, suggesting a tendency to select one trajectory from the many available.

Goal-directed movement can be divided into reaching and grasping. The goal-directed movement control differs for the following tasks:

- *Pointing*—The target position is imprecisely designed, yet there is only the direction of the distal segment of the arm.
- *Point-to-point reaching without grasping*—Both the initial and the target positions are known; however, the orientation of the hand is irrelevant, as well as the trajectory of the hand between the end points.
- *Point-to-point movement with grasping*—Complex movement that requires the orientation of the hand that corresponds to the shape of the object to be grasped, in addition to having precision in reaching the target with the velocity being about the same as that of the object to minimize the impact.
- *Tracking*—Movement along a prescribed trajectory.



Grasping is a part of the prehension, which is a process of orienting the hand, opening it so that the object fits comfortably, coming in contact with the object, and then forming a firm grip. On the kinetic level, prehension entails applying forces during interaction with an object. Stable grasping requires that the forces are applied by hand surfaces in opposition to other hand surfaces or external objects to overcome perturbations. The human hand has a variety of ways to grasp objects firmly. The selection of the grasp depends on the function that has to be achieved and the physical constraints of the object and hand. A fundamental functional constraint is that once the object is grasped, it is not to be dropped. The posture used by the hand during the task must be capable of overcoming perturbations and include anticipated forces that may act at the object.

MAN-MADE CONTROL OF MOVEMENT

In terms of control engineering, the human body that is being controlled is referred to as the *plant*. The configuration of the system at any given time comprises the plant states. The system components that power the system are called actuators. The signals driving the actuators are called controls. The controller is the component that generates controls. The time histories of the plant states in response to the controls are referred to as the system trajectory.

The open-loop control operates based on a pre-programmed set of commands. The open loop control cannot change the command sequence even when the trajectories that are achieved are different from the ones desired. To correct disturbances and errors, a closed-loop controller needs to be utilized. Closed-loop controllers, or so-called error driven controllers, use feedback information based on sensors signaling about the actual plant states. The error based controllers require the following: 1) the complete knowledge of the system parameters; 2) desired trajectories of the system components; and 3) acceptable tracking errors. In parallel, control requires optimization because the systems are highly redundant. The optimization criterion can consider a single mechanical category (e.g., velocity, acceleration, energy, tracking error); yet, often it considers several categories.

Open-loop and closed-loop controllers are based on the simulation of the system performance. The simulation requires the modeling that has to be derived by way of using mechanics and mathematics.^[8] A mathematical model is a quantitative description that connects the states and controls. The models of the real world are never exactly correct. Errors in the models of physiological systems arise from the practical modeling necessity of omitting many factors that are either of

secondary importance to the problem studies or unknown. The second error arises because of the necessity of describing the spatially distributed systems by lumped systems. The first step is always to choose the scope of the model. Although it would be ideal to adopt one general model, it is not practical because the generality increases the mathematical complexity to an inadequately treatable level in real time. In parallel, the general model comprises too many states, controls, and parameters that cannot be determined with sufficient accuracy. The second step in the modeling process is the selection of the model form. One can choose between the black-box model, structurally based model, and reductionist model.^[9,10] In a structural model, the elements of the system are isolated on the basis of a known or assumed system structure. The models of the individual elements are combined into multielement systems to describe the overall behavior. The advantages of a structural model are that the knowledge of the internal structure and operation can be incorporated, and the internal states and parameters can be directly related to the physical properties and experimental estimations. A possible disadvantage of structural models is that the individual element models require knowledge of parameters that are impossible to estimate.

The reductionist models are in essence identical to the structural models. The reductionist models were introduced by engineers who design artificial control and assistive systems, but were accepted by kinesiologists for studying specific behaviors. The reductionist models consider only the isolated components that are chosen to be instrumental for the selected movement.

Fig. 2 shows one possible structural model of movement.

Typically, the structural models of movement place an emphasis only on the extremities. A body segment, when exposed to physiological loading, is modeled as a rigid body although it comprises soft tissues (e.g., muscles, tendons, ligaments, skin, etc). The inertial properties of a rigid body are its mass and the inertia tensor. The inertia tensor comprises information about the inertial properties of the body when rotating, and the mass is a measure of inertial properties when translating. In reality, it is impossible in vivo to determine the body segments' properties with the precision that is required from the instrumentation available today. The second problem is that the body segment parameters change during movement because of the redistribution of the soft tissues (muscle contraction).

An important issue in the modeling relates to the joint structures. The term in the theory of mechanisms for the connection of two rigid bodies is a kinematic pair. The position of the free rigid body is determined by way of six independent parameters (coordinates).

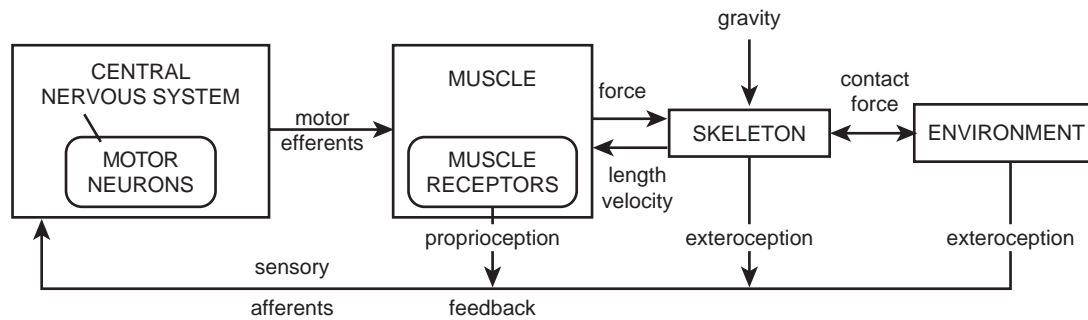


Fig. 2 The organization of sensory-motor systems leading to movement. The model corresponds to the general model shown in Fig. 1.

The independent parameter is termed degree of freedom (DoF). The theory of mechanisms defines the connection of segments (rigid bodies) as kinematic pairs having up to five DoFs. A typical human joint is a rotational kinematic pair having three DoFs (ball joint), or one DOF (pin joint). Many bodily joints cannot be reduced to a pin or ball joint. Some human joints are a structure that comprises two mechanical joints (e.g., the ankle joint consists of one joint for internal/external rotation and inversion/eversion, and the other for the flexion/extension that allows movements in the sagittal plane). In some cases the body segment motion results from complex structures such as two parallel bones connecting two joints (e.g., elbow and wrist). This structure provides a degree of freedom that cannot be associated with a joint (e.g., supination/pronation of the forearm).

The mathematical model allows the determination of the position of body parts in the absolute reference frame when the joint angles are known (direct kinematics), whereas the inverse kinematics compute the joint angles when a position of the body part is known. The direct kinematics analysis has a unique solution. An inverse kinematics solution depends on the redundancy of the structure, as well as the offsets and constraints, and in most cases the solution is not unique. Redundancy is an important feature in biological systems. An understanding of redundancy requires a distinction between a DoF and a permissible motion (e.g., rotation). The term DoF exclusively pertains to the permissible motion being independent from other motions. When analyzing the body as a whole, the total number of permissible rotations is termed system DoFs. It is difficult to determine which permissible rotations should be controlled, and which will be constrained by themselves owing to anatomical and physiological limitations.

Dynamic analysis deals with two problems: determination of forces and torques when desired laws of motion are known (forward dynamics), and determination of trajectories and laws of motion when torques and forces are prescribed (inverse dynamics).

A dynamic model is a set of equations that emulate the behavior of the physical system. Dynamic models specify how the current states of the system (e.g., body trajectories) depend on the past (e.g., the prior body trajectories and muscle forces). The inputs to a dynamic model of a human motor task are the muscle stimulation patterns, and the outputs are the computed body trajectories:

$$X' = f(X, U)$$

$$Y = g(X, U)$$

The differential equations can be written using the so-called Cauchy form. The vector X' comprises the first derivatives of the state vector X , vector Y comprises the outputs, nonlinear functions f and g describe the plant dynamics, and vector U consists of the controls.

There are twice more first-order differential equations compared with the number of DoFs. The first-order Cauchy form can also be represented in the following form:

$$\ddot{\theta} = A^{-1}(\theta)[M_a + M_{in}(\theta, \dot{\theta}) + M_g(\theta)]$$

where θ is the vector of joint angles, M_a , M_{in} , and M_g are the active, inertial, and gravity force based torques, respectively, and A^{-1} is the inverse matrix of the system. The muscle torque contribution is a vector of the torques produced by muscles spanning the joints. These torques are multiplied by A^{-1} , an inverse of the system matrix $A_{n \times n}$. A is the time variable, which depends on the inertial properties of the system, and also on the positions of the system components, which vary during the movement (e.g., trigonometric functions of joint angles). When A^{-1} has only diagonal elements, a muscle would only cause the angular acceleration of the joint it spans. A^{-1} is generally a full matrix with significant off-diagonal terms. Biarticular muscles (muscles crossing two joints) could accelerate a joint they span in a direction opposite to their torque at the corresponding joints.



Using the laws of mechanics in inverse dynamics works only for open chain structures such as an arm when reaching, or a leg during swinging, but it is not applicable for closed chain structures such as legs during the double support phase in walking, or the arm when contacting an immobile object, or in redundant systems. Every closed chained structure is dynamically undetermined, and the only possibility of determining forces and torques is to use the theory of elasticity. The theory of elasticity assumes that the elements of the structure are solid (not rigid), and this theory is inappropriate for control design because of its extreme complexity. The approximate solution of the dynamics of a closed chain mechanism assumes how forces and torques are distributed in the structure.

The complexity of a model depends on the number and types of body segments, the joints connecting the segments, and the interaction among the segments and the environment. The analysis of multijoint structures is always characterized by a reductionist approach, that is, the reduction of the number of DoFs to a manageable level (e.g., walking is often reduced only to the sagittal plane). Furthermore, there are three-dimensional models in the literature; yet, they typically add only a few joints that are out of the sagittal plane. Using a smaller number of links is essential for the simulation of movement. The dynamic analysis requires the decomposition of the system.^[11] The actuators in the body are muscles. The muscle produces a pair of forces of identical intensity but opposite direction, acting on two bodily segments that they connect. Some muscles connect non-neighboring segments, act over more than one joint, and are termed biarticular muscles or multiarticular muscles. When analyzing the body as a whole, the sum of forces generated by muscles is zero, and the muscle forces are called internal forces. The decomposition is virtual “dissection” of the body into segments. The result of decomposition is that now a muscle force acts as an external force to a specific bodily segment. The pair of agonist and antagonist muscles therefore results after decomposition as an external momentum with the net joint torque producing movement.

The modeling of muscles has received a great deal of attention. In, short muscle models can be grouped into three classes: 1) cross-bridge models including a sarcomere—microscopic, conventional cross-bridge, and unconventional cross-bridge models; 2) fiber models; and 3) whole-muscle—macroscopic models (viscoelastic and black box models).^[12] Here, we will introduce only macroscopic viscoelastic models of muscles.^[13] Winters generalized Hill’s model in a simple enhancement of the original, which makes no claim based on the microstructure.^[12,13]

Winters’ model deals with joint moments:

$$M = M_0 - B\dot{\Theta}, M_0 = AM_{0m}f(\Theta)$$

where M is the muscle moment exerted about the joint, M_0 is the isometric moment (active state) with a maximum possible value M_{0m} , Θ is the joint angle corresponding to the current contractile element length, A is a normalized activation variable ($A=0$ for relaxed muscle, $A=1$ for a tetanized muscle), $f(\Theta)$ is the normalized tetanic length–tension relation, and B is a variable damping coefficient. For shortening muscle (SM) and lengthening muscles (LM), the equations are:

$$B = M_0(1 + \alpha)/(\dot{\Theta} + \alpha\dot{\Theta}_0), \quad \dot{\Theta} < 0 \quad (\text{SM})$$

$$B = M_0(1 - \beta)/(\dot{\Theta} + \gamma\dot{\Theta}_0), \quad \dot{\Theta} > 0 \quad (\text{LM})$$

where α , β , and γ are constants.

The black-box model only considers a subset of states of interest and controls without assuming anything about how the system operates internally. The relation in the black-box model is often nonparametric, and nonparametric black-box models were also termed nonanalytical models. The system operation, not the plant, is modeled as a set of mappings between inputs and outputs. Nonanalytic control can be described as a replica of biological control.^[14,15] The basic mechanisms of nonanalytical control are rule-based controllers (RBC).^[16] An RBC is a system implementing “If-Then” rules, where the “If” part describes the sensory and motor states of the system while the “Then” part of a rule defines the corresponding motor activity to follow. The If-Then operation in nature is called reflex; hence, the man-made If-Then operation was termed “artificial reflex control.” Artificial reflex control belongs to discrete control systems, and in essence it decomposes the movement into time windows and defines synergies between the state space variables by the If-Then rules. The execution of the artificial reflex depends on how fine the tuning was, as determined by observing and analyzing biologically controlled movements. This control method comprises the following: 1) adaptivity; 2) modularity; and 3) ease of application and the possibility of integration into a man–machine system. The main problem is how to design the rules for artificial reflex control. Various machine learning techniques were shown to be suitable for acquiring the needed structure of rules for the RBC. Learning in general can be described as the capturing and memorizing of connectivities between facts. Machine learning is a computerized method for the capturing and memorizing of the processes.

HIERARCHICAL HYBRID CONTROL OF MOVEMENT

Hybrid means, in general, heterogenous in nature or composition. The term “hybrid systems” is understood to describe systems with behavior defined by entities or processes of distinct characteristics. The hybrid systems of interest here are dynamic systems where the behavior is determined by interacting continuous and discrete dynamics. These systems typically contain variables or signals that take values from a continuous set (e.g., the set of real numbers) and also variables that take values from a discrete and typically finite set (e.g., the set of symbols {a, b, c}). These continuous or discrete-valued variables or signals depend on independent variables such as time, which also may be continuous or discrete; some of the variables may also be a discrete event that is driven in an asynchronous manner.^[17,18]

There are many examples of hybrid systems. In the control area, a very well-known instance of a hybrid system is when a continuous-time linear time-invariant plant that can be described by linear differential equations (which involve continuous-valued variables that depend on continuous time) is controlled by a discrete-time linear time-invariant plant described by linear difference equations (which involve continuous-valued variables depending on discrete time). These types of systems are termed sampled data systems or digital control systems; digital control systems may of course include more general types of systems such as time-varying and nonlinear plants and controllers. If one were to also consider the quantization of the continuous-valued variables or signals, then the hybrid systems contain not only continuous valued variables that are driven by continuous and discrete times, but also discrete-valued signals. Another familiar example of a hybrid control system is a switching system where the dynamic behavior of interest can be adequately described by a finite (small) number of dynamical models, i.e., typically the sets of differential or difference equations together with a set of rules for switching among these models. These switching rules are described by logic expressions or a discrete event system with a finite automaton or a petri net representation.

There are several reasons for using hybrid models to represent movement. Reducing the complexity was and still is an important reason for dealing with hybrid systems. This is accomplished in hybrid systems by incorporating models of dynamic processes at different levels of abstraction. For another example, to avoid dealing directly with a set of nonlinear equations, one may choose to work with sets of simpler equations (e.g., linear) and switch among these simpler models.

This is a rather common approach in modeling physical phenomena. In control, switching among simple dynamical systems has been used successfully in practice for many decades.

The hybrid control systems typically arise from the interaction of discrete planning algorithms and continuous processes, and, as such, they provide the basic framework and methodology for the analysis and synthesis of autonomous and intelligent systems, i.e., planning to move the hand and grasp an object. The hybrid control systems contain two distinct types of components, subsystems with continuous dynamics and subsystems with discrete event dynamics, that interact with each other.

Another important way in which hybrid systems arise is from the hierarchical organization of complex control systems (Fig. 3). In these systems, a hierarchical organization helps manage the complexity, and higher levels in the hierarchy require less detailed models (discrete abstractions) of the functioning of the lower levels, necessitating the interaction of discrete and continuous components.

There are analogies between certain current approaches to hybrid control and digital control system methodologies. In digital control, one could carry the control design in the continuous-time domain, then approximate or emulate the controller by way of a discrete controller and implement it using an interface consisting of a sample and a hold device (analog to digital, A/D, and digital to analog, D/A, respectively). Alternatively, one could first obtain a discrete model of the plant taken together with the interface and then carry the controller design in the discrete domain. In hybrid systems, in a manner analogous to the latter case, one may obtain a discrete event model of the plant together with the interface using automata or petri nets.^[19]

Hybrid systems inherently combine logical and continuous processes, which are usually coupled with finite automata and differential equations. Thus, the continuous dynamics are modeled by a differential equation:

$$\dot{x}(t) = \xi(t), \quad t \geq 0$$

where $x(t)$ is the continuous component of the state taking values in some subset of a Euclidean space, and $\xi(t)$ is a controlled vector field that depends on $x(t)$ and the aforementioned logical or finite dynamics. Two categories of the switched systems are of interest: the autonomous switching where the vector field $\xi(\cdot)$ changes discontinuously when the state $x(\cdot)$ reaches certain boundaries, and the controlled switching where $\xi(\cdot)$ changes abruptly in response to a control command, and possibly with an associated cost.



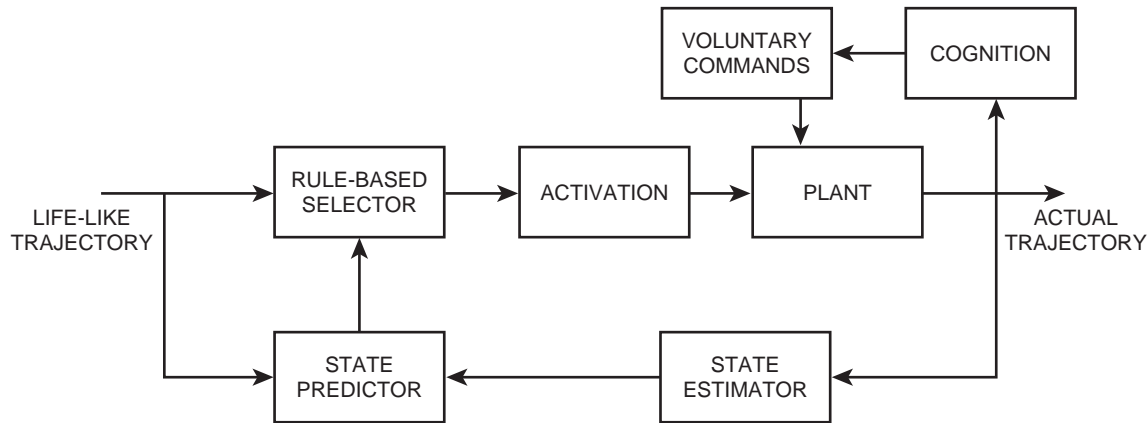


Fig. 3 The organization of a man-made hierarchical hybrid control system for restoring movement in humans.

A (continuous time) autonomous hybrid system may be defined as follows:

$$\begin{aligned}\dot{x}(t) &= f(x(t), q(t)) \\ q(t) &= v(x(t), q(t^-))\end{aligned}$$

where $x(t) \in R^n$, $q(t) \in Q \cong \{1, \dots, N\}$. Here, $f(\cdot, q): R^n \rightarrow R^n$, $q \in Q$, each globally Lipschitz continuous represents the continuous dynamics, and $v: R^n \times Q \rightarrow Q$ is the finite state dynamics. Here, the notation t^- indicates that the finite state is piecewise continuous. Thus, starting at $[x_0, i]$, the continuous-state trajectory $x(\cdot)$ evolves according to $\dot{x}(t) = f(x, i)$. If $x(\cdot)$ reaches $(v(\cdot, i))^{-1}(j)$ at time t_i , then the state becomes $[x(t_i), j]$, from which the process continues.^[20]

A continuous-time controlled hybrid system is:

$$\begin{aligned}\dot{x}(t) &= f(x(t), q(t), u(t)) \\ q(t) &= v(x(t), q(t^-), u(t))\end{aligned}$$

where everything is as above except that $u(t) \in R^m$, with f and v modified appropriately.

The schema of a hierarchical hybrid controller for movement incorporates three levels (Figs. 4 and 5). We have opted to describe the model of walking, but similar models can be easily envisioned for other movements.

The top artificial control structure is the interface between the user and the machine. This interface is the principal command channel, and it allows the user to trigger the operation of their choice volitionally.

The interface initiates the activity of a discrete, RBC. This RBC operates as a discrete, sample data feedback system, and its main role is to distribute the commands to the lowest actuator levels. The RBC implements the finite-state model of movement, and the rules have to be determined with sufficient generality to allow the application over many assistive systems to be applicable to the entire population with a

similar level of disability. The actuator level deals with specific muscle groups responsible for the flexion or the extension of a single joint, or in other cases, the action of several joints when a multiarticular muscle is externally stimulated. The actuator level implements the continuous feedback control and structural modeling.

The RBC system is applicable for the coordination level of control. This level deals with the following: 1) the strategy of how to employ the resources available and 2) the methodology of how to maximize the efficiency of the resources. The RBC in this multilevel control uses the simulation of movement as the pattern that has to be followed. This is to say that the RBC does not clone normal, able-bodied movement,

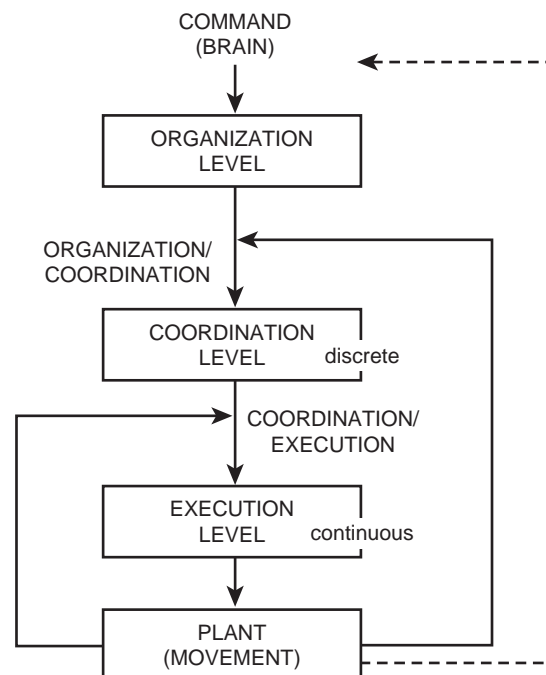


Fig. 4 The model of life-like movement controller based on a nonanalytical model of movement.

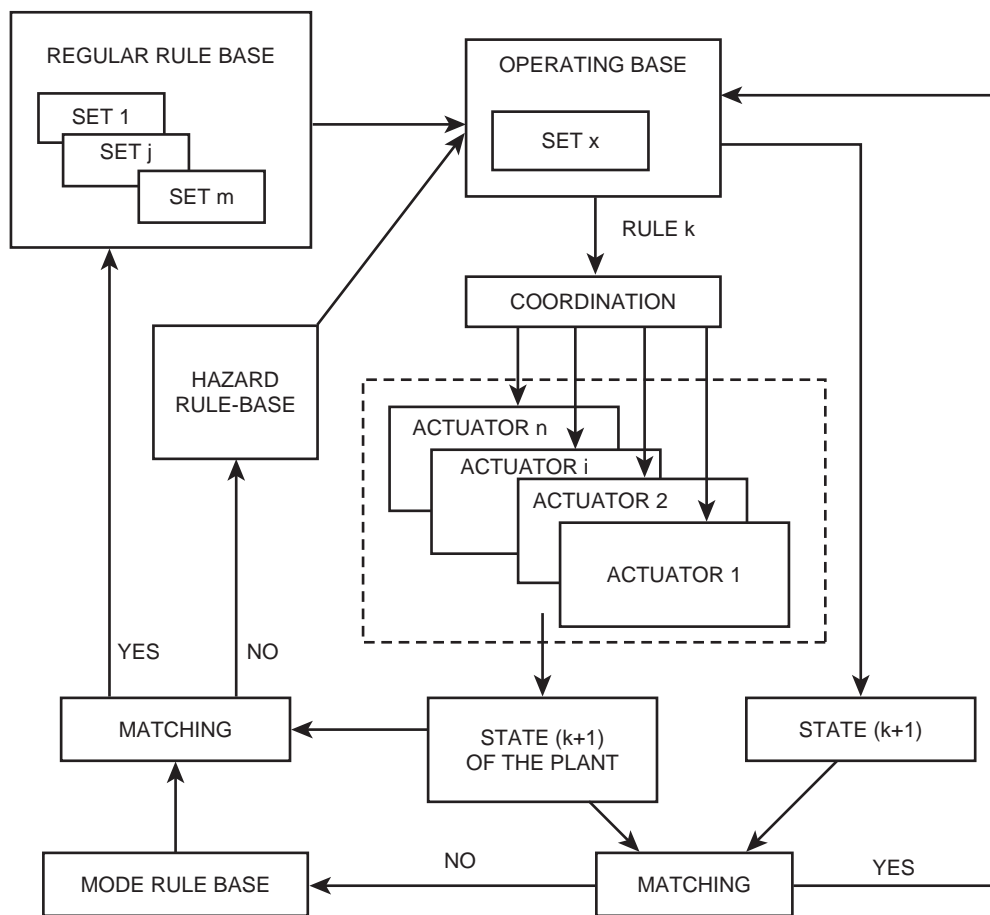


Fig. 5 The schema of a hierarchical hybrid controller for paralyzed extremities in humans with sensory-motor disability.

although the simulated movement along the desired trajectory is achievable. This means that the structural model is of great importance for this off-line analysis. The RBC comprises the following elements: the regular rules that sequentially switch from one to the expected action based on sensory input; mode recognition rules that are responsible for selecting the appropriate set of regular rules; and the hazard states that deal with conflict situations. The conflict situations occur as a result of the uncertainty of the available sensory information and/or hardware limitations, in addition to unexpected gait events. The hazard rules result in the safety “behavior,” which attempts to minimize the eventual catastrophic consequences of the hazard (e.g., falling, obstacles, nonphysiological loading).

The lower, actuator control level is responsible for executing decisions from the coordination level. Executing commands in the sense of artificial reflexes means that the electrical stimulation has to be delivered to a group of muscles that are controlling a joint. Single-joint control is achieved through a coordinated action of several muscles acting at the neighboring segments. Details on the various methods that are implemented for the control of movement by

means of neural prostheses of upper and lower extremities can be found in Chapter 5 of *Control of Movement for the Physically Disabled*, while the specifics on the hierarchical hybrid control regarding the rehabilitation of humans with disabilities can be found in Popović.^[1,21,22]

CONCLUSIONS

Compared with other classes of dynamical systems, the sensory and motor systems supporting the execution of functional motions in humans have three unique features: 1) they are highly redundant; 2) they are organized in a hierarchical structure, yet with many parallel channels; and 3) they are self-organized, relying, among other things, on an extremely complex connectionism. In spite of the above complexity, the movements resulting from the action of the sensory system, as a rule, are deterministic and they follow the preferred way of performing the intended motor task. In this entry some of the above phenomena are briefly discussed, taking a comprehensive approach, i.e., discussing general findings from motor physiology,



biomechanics, and automatic control. The entry offers a plausible explanation as to how the self-organized systems are able to force otherwise uncontrollable system components to perform the desired movement, essentially in a deterministic, nonstochastic way. In external control of movement the aim is typically to activate joints in a controlled way so as to restore as much motor function as possible in individuals with motor disabilities. The control strategies implemented in most of the assistive systems have so far been fairly simple, and have been developed largely in relation to the design of machines rather than to the design of nervous systems. Neural engineering is where the ultimate successes at this stage must come. The development of new implantable devices that interface directly the central and peripheral nervous system allowing wireless communication with the outside world opens the space for integrating the sophisticated artificial control. The nanotechnology is making a dramatic impact, and provides what has been difficult to imagine. This technology is ideal for implementing control that mimics natural regulation, and integration into the preserved sensory-motor systems of individuals with disabilities could be the method for improving the quality of life.

ARTICLE OF FURTHER INTEREST

Artificial Muscles

REFERENCES

1. Popović, D.B.; Sinkjær, T. *Control of Movement for the Physically Disabled*; Springer: London, 2000.
2. Kandel, E.R., Schwartz, J.H., Jessel, T.M., Eds.; *Principles of Neural Science*, 4th Ed.; Appleton and Lange: Norwalk Connecticut, 1991.
3. *Biomechanics and Neural Control of Posture and Movement*; Winters, J.M., Crago, P.E., Eds.; Springer-Verlag: New York, 2000.
4. Georgopoulos, A.P. On reaching. *Annu. Rev. Neurosci.* **1986**, *9*, 147–170.
5. Bernstein, N.A. *The Coordination and Regulation of Movements*; Pergamon Press: London, 1967 (Original work in 1926–1935.).
6. Soechting, J.F. Elements of coordinated arm movements in three-dimensional space. In *Perspectives on the Coordination of Movement*, Wallace, S.A., Ed.; Amsterdam: North-Holland, 1989; 47–83.
7. Hogan, N. An organization principle for a class of voluntary movements. *J. Neurosci.* **1984**, *4*, 2745–2754.
8. Crago, P.E. Creating neuromuscular models. In *Biomechanics and Neural Control of Posture and Movement*; Winters, J.M., Crago, P.E., Eds.; Springer: London, 2000; 119–133.
9. Kearney, R.E.; Kirsch, R.F. System identification and neuromuscular modeling. In *Biomechanics and Neural Control of Posture and Movement*; Springer Verlag: New York, 2000; 134–147.
10. Brown, I.E.; Loeb, G.E. A reductionist approach to creating and using neuromuscular models. In *Biomechanics and Neural Control of Posture and Movement*; Winters, J.M., Crago, P.E., Eds.; Springer: London, 2000; 148–163.
11. Zajac, F.E. Muscle and tendon: properties, models, scaling, and application to biomechanics and motor control. *Crit. Rev. Biomed. Eng.* **1989**, *17* (4), 359–411.
12. Hill, T.L. The heat of shortening and the dynamic constants of muscle. *Proc. R. Soc. Lond. Biol.* **1938**, *126*, 135–195.
13. Winters, J.M. Hill-based muscle models: a systems engineering prospective. In *Multiple Muscle Systems: Biomechanics and Movement Organization*; Winters, J.M., Woo, S.L.-Y., Eds.; Springer Verlag: New York, 1990; 66–93.
14. Popović, D.B. Finite state model of locomotion for functional electrical stimulation systems. *Prog. Brain Res.* **1993**, *97*, 397–407.
15. Popović, D.B.; Stein, R.B.; Oğuztöreli, M.N.; Lebedowska, M.; Jonić, S. Optimal control of walking with functional electrical stimulation: a computer simulation study. *IEEE Trans. Rehabil. Eng.* **1999**, *TRE-7* (1), 69–79.
16. Tomović, R.; Popović, D.B.; Stein, R.B. *Nonanalytical Methods for Motor Control*; World Science: Singapore, 1995.
17. Antsaklis, P.J.; Nerode, A. Hybrid control systems: an introductory discussion to the special issue. *IEEE Autom. Control* **1988**, *AU-43*, 457–459.
18. Lima, P.U.; Saridis, G.N. Intelligent controllers as hierarchical stochastic automata. *IEEE Trans. Syst. Man. Cybern.* **1999**, *SMC (B)-29* (2), 151–163.
19. Antsaklis, P.; Kohn, W.; Nerode, A.; Sastry, S. Eds.; *Hybrid Systems II*, Lecture Notes in Computer Science, No. 999 Ed.; Springer Verlag: New York, 1995.
20. Branicki, M.S. Multiple Liapunov functions and other analysis tool for switched and hybrid systems. *IEEE Trans. Autom. Control.* **1998**, *AU-43*, 475–482.
21. Popović, M.B. Control of neural prostheses for grasping and reaching. *Med. Eng. Phys.* **2003**, *25* (1), 41–50.
22. Popović, D.B.; Popović, M.B. Tuning of a nonanalytic hierarchical control system for reaching with FES. *IEEE Trans. Biomed. Eng.* **1998**, *BME-45*, 203–212.

Controlled Release

Brent Vernon

Michael Wegner

The Harrington Department of Bioengineering, Arizona State University, Tempe, Arizona, U.S.A.



INTRODUCTION

Early drug delivery methods typically involved simple oral administration as liquid, powder, or tablets or topical delivery. As drugs were developed, such as proteins and peptide drugs, which cannot overcome the natural barriers present in oral and topical delivery, they began to be injected or delivered intravenously. Recently developed drugs with narrow therapeutic concentrations, systemic side effects, short half lives, or other barriers to efficient application require advanced methods of delivery. Delivery of these compounds to a specific location (spatial control) or with a specific pattern (temporal control), or both, is being pursued to overcome these barriers. “Controlled release” refers to this type of delivery.

Several system types have been or are under development to provide this spatial and temporal drug release. These different system types are classified as 1) mechanical/electrical systems, 2) polymeric systems, and 3) biological systems. Examples and history of these different types of controlled-release systems will be presented here. In these systems, modeling and analysis of the release profile is required because the rate of release is crucial. Furthermore, controlled-release technology often combines drug and device and sometimes biologics into one product. This combination of drug and device results in regulatory issues unique to these systems.

ADVANTAGES AND PURPOSES OF CONTROLLED RELEASE

Many drugs benefit from their controlled release to overcome a number of difficulties. Controlled release is often used to reduce toxicity associated with bolus oral or injected delivery by delivering the drug at steady lower therapeutic concentrations for an extended time instead of at high initial concentrations.^[1–3] These systems can replace multiple doses with a single dose or fewer doses. Controlled release can also reduce systemic toxicity by providing localized delivery, avoiding high concentrations of the drug throughout the body.

As an example, biodegradable polyanhydride wafers (GLIADEL) have been used to deliver toxic chemotherapy agents to malignant brain tumors in phase-three clinical trials.^[2] Many drugs also require precise timing in delivery; controlled-release technology can provide a solution for this need. For example, insulin delivery to diabetics needs to be correlated to high blood glucose levels,^[4,5] and other drugs require delivery timed with circadian cycles.^[6] In many other drug delivery devices, the goal is to obtain constant drug delivery rates. Controlled-release technology can also be used to bypass various barriers to drug delivery, such as cell membranes^[7] or aggressive digestive environments.^[8] Controlled-release systems provide protection from in-vivo metabolism to many drugs, especially proteins and peptides, allowing higher stability of the drugs and longer periods of efficacy.^[9] Controlled-release systems can also help to enhance delivery of poorly soluble drugs.^[10] Controlled-release systems often enhance the delivery of many cancer agents that are poorly soluble.^[10] Table 1 summarizes the advantages and purposes of controlled release.

MECHANICAL/ELECTRICAL SYSTEMS

There are several controlled-release systems developed or under development that are mechanical and/or electrical in nature. Early systems of this type were generally passive pumps that were intended to provide constant, sustained release over a long period of time using a constant pump rate. Common pump systems include infusion pumps, peristaltic pumps, and osmotic pumps.^[11] The first of these, the infusion pump, is commonly used for insulin release. First attempts at controlled release of insulin using infusion pumps began in the late 1970s.^[12] These systems improve diabetic treatment by providing a sustained continuous delivery of insulin.^[4] Future systems improved on this by adding programmable pumps to provide additional control over the timing of insulin delivery. More recent models combine insulin pumps with glucose sensors.^[5] However, the current unsolved issue is the development of a continuous and reliable sensor that provides

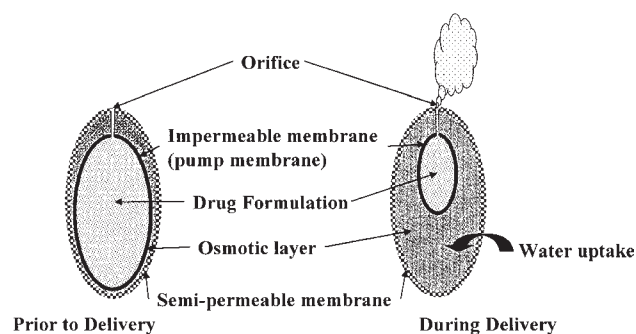
Table 1 Advantages and purposes of controlled-release technology

1. Reduce toxicity associated with bolus oral, injected, or intravenous delivery.
2. Reduce systemic toxicity by providing localized delivery.
3. Provide precise timing in delivery.
4. Bypass various barriers to drug delivery, such as cell membranes or aggressive digestive environments.
5. Protect drugs from in-vivo metabolism allowing higher stability of the drug and longer periods of efficacy.
6. Enhance delivery of poorly soluble drugs.

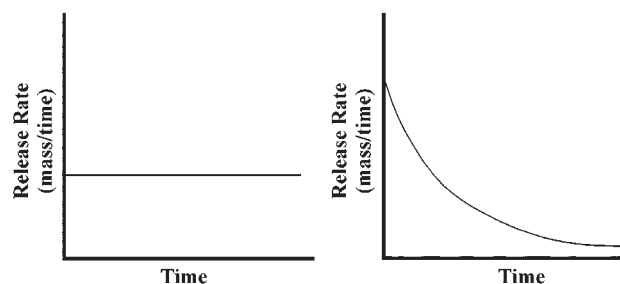
accurate and real-time glucose monitoring. The eventual development of such a “feedback-controlled” artificial pancreas will require a long-term stable glucose sensor; this would allow the controlled release of insulin in response to blood glucose levels.^[4] Although a preponderance of these systems have been investigated for the delivery of insulin, infusion pumps—external and implantable—have been pursued for delivery of numerous other drugs such as chemotherapy agents^[13] and neurological drugs.^[14] The use of an implantable infusion pump to treat cancer metastases in the liver provides a prime example of infusion pump technology used to localize the delivery of chemotherapy agents.^[13] Since the liver obtains a large portion of its blood supply from a single hepatic artery, slow infusion of drugs such as 5-fluorodeoxyuridine provides an increased concentration of the drug locally (in the liver) but avoids high systemic concentrations.^[13] In a 2000 review, Rise summarizes several implantable infusion drug delivery systems designed for delivery of neurological drugs and other chemical agents to the central nervous system (CNS).^[14] Medtronic offers an implantable and programmable drug infusion system called “Synchromed EL.”^[14] This system can be used to provide spatial and temporal control of drugs to the CNS. Such systems have been used to deliver neuro-modulating and pain-control agents.^[14]

Similar to the infusion pumps, peristaltic pumps have been developed for controlled delivery of insulin for diabetics. Early implantable peristaltic pumps for insulin release appeared in 1980.^[15] Early systems of this type showed the advantage of remote control using an external remote.^[15] So far, however, these systems have proven to be too expensive to be used practically.^[11]

The osmotic pump controlled-release system was first described by Theeuwes and Yum in 1976.^[16] Since then, the Alza Corporation has developed osmotic pump delivery systems for various drugs such as hydromorphone,^[3] verapamil hydrochloride,^[8] and pseudoephedrine.^[17] These osmotic pump drug-delivery systems work by surrounding a drug-filled impermeable

**Fig. 1** Representation of the mechanism used for osmotic pump controlled-release drug delivery systems. (Adapted from Ref. [18].)

membrane with a semipermeable membrane and an osmotic layer between the impermeable and semi-permeable membranes. Water is driven through the semipermeable membrane by osmotic pressure compressing the inner impermeable membrane, driving the drug formulation from the device. Figure 1 presents this mechanism for osmotic pumps.^[18] Various different configurations and designs have been developed using the same principle. Osmotic pumps are popular because they can provide zero-order release of the drug and avoid the “initial burst effect” seen with many other types of controlled-release systems. Figure 2 compares zero-order release to first-order release rates. As an example, one of Alza’s pumps released 262 ug/hour of the pain management agent, hydromorphone, over a two-week period while avoiding the burst effect seen in various degradable and nondegradable matrix systems.^[3] The Alza osmotic systems, ALZET osmotic pumps, are available for laboratory and veterinary use. More recently, osmotic pump systems have been shown to deliver pseudoephedrine with zero-order kinetics over 12 hours.^[17] In another osmotic pump device designed by G.D. Searle & Alza Corporation, a hydrophilic cellulosic membrane is used to delay the delivery of verapamil hydrochloride, a hypertension drug, for

**Fig. 2** Comparison between zero-order release and first-order release mass release rate with time showing constant release in zero-order systems compared with decreasing release rate in first-order systems.

4 to 5 hours after oral administration.^[6] Laser-drilled micropores allow release of the drug over 24 hours. This design is used to match delivery of the verapamil hydrochloride to circadian rhythm of hypertension.^[6]

Less common pump systems include positive displacement pumps and micropumps.^[11] Additionally, a pressure-induced permeation pump described by Uchiyama et al. in 2002 utilizes pressure-induced changes in insulin permeation through a permeable membrane to control release. With an applied pressure of 18 kPa, the release of insulin from this device increased 3.4-fold.^[19] Different mechanical and electrical technologies offer achievable devices that can control the release of numerous drugs. Many of these systems, when combined with computer or sensor control, can provide a sophisticated and useful delivery pattern.

POLYMERIC SYSTEMS

In examples of polymeric controlled-release systems, drugs combine with polymers and then release from the polymer. Early pioneer work was performed by Langer et al. to demonstrate the release of large macromolecule drugs such as proteins from polymeric devices.^[20] Since this early work, the use of polymeric materials for controlled release has expanded extensively.

These polymer systems can be either degradable or nondegradable. In both cases, there are two important classes: matrix and reservoir (see Fig. 3). In matrix systems, the drug is dispersed or dissolved within a solid or porous polymer matrix. The drug is generally delivered by one or more of three mechanisms from these polymer matrices: diffusion, degradation, and swelling-enhanced diffusion.^[21]

Degradation-controlled systems come in two classes: erodible systems and drug-polymer conjugates.^[22] In the first class, the drug is entrapped within a polymer matrix and is released as the matrix erodes. In the second class, the polymer is chemically attached to the matrix or polymer and is released as the drug-polymer bond is broken, usually hydrolytically or enzymatically. The degradable matrices in the erodible systems can undergo either bulk degradation or surface erosion. Bulk degradation occurs when water, or another agent that causes the degradation, can penetrate into the matrix and cause uniform degradation throughout the entire volume of the matrix. In bulk degradation systems, the drug is often released by a combination of diffusion and degradation mechanisms, referred to as heterogeneous degradation.^[22] In this case, the release rate of the drug is dependent on the permeability of the polymer matrix, which is changing with time due to bulk degradation of the matrix. For these systems, in a slab configuration, the rate of release is shown in Eq. 1,

$$\frac{dM_t}{dt} = \frac{A}{2} \left[\frac{2P_o e^{k_h t} C_o}{t} \right]^{1/2} \quad (1)$$

where dM_t/dt is the rate of release, A is the surface area, P_o is the permeability of the matrix, C_o is the original concentration of the drug, t is time, and k_h is the hydrolysis rate constant, a constant associated with the rate of change in the matrix permeability.^[22] Solutions are also available for other geometries.^[22] Surface erosion, on the other hand, occurs when the matrix is impermeable to water or other degradation agents and the degradation must occur on the surface. In surface erosion systems, the drug release rate is often equal to the polymer degradation rate, referred

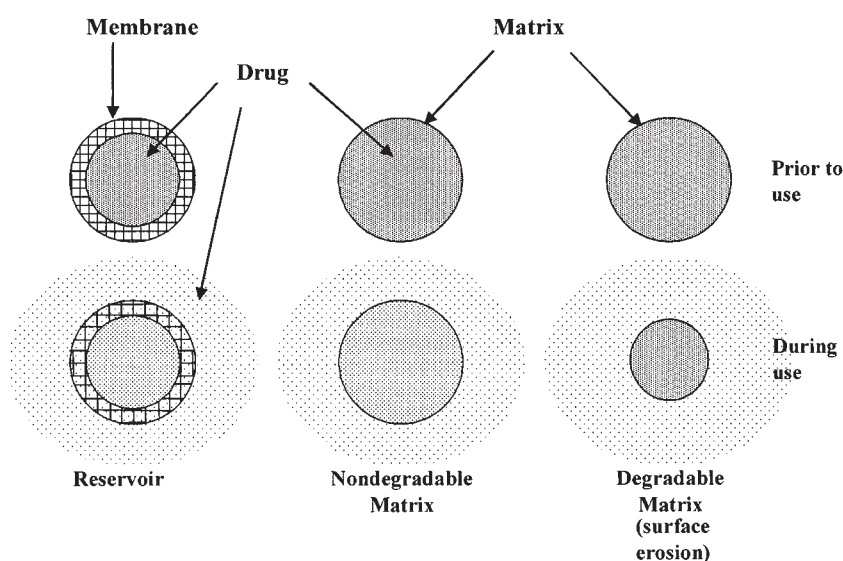


Fig. 3 Depiction of possible reservoir, nondegradable, and degradable matrix controlled-release systems designs.

to as homogeneous degradation.^[22] In this case, the release of the drug is dependent on the surface area of the system and the surface degradation rate. Device designs that provide constant surface area with time, as the material erodes, can provide constant zero-order release when release occurs by this mechanism. In systems with heterogeneous release, the release rate is dependent on the surface area of the system and the surface degradation rate as in Eq. 2,

$$\frac{dM_t}{dt} = k_{sd}C_oA \quad (2)$$

where dM_t/dt is the rate of release, C_o is the original drug concentration, k_{sd} is the surface erosion rate constant, and A is the surface area of the device.^[22] Many polymeric drug delivery devices show attributes of both surface and bulk erosion, to varying degrees.^[22] In these systems, the release profile can become more complicated.

As an example of bulk degradation, poly(lactide:glycolide/75:25) (PGLA) can be prepared as microspheres that degrade by bulk erosion.^[21] In controlled-release applications, these systems exhibit uniform degradation throughout the entire matrix. Generally, in these systems, the drug must diffuse out of the device as the device degrades to avoid drug dumping when the end-stage bulk erosion causes rapid mass loss in the matrix. In contrast, in surface-eroded devices, drug is released as the surface erodes; therefore, the release rate is dependent on the surface area of the drug delivery system.^[21] In an example of surface-erosion controlled delivery, Hanes et al. demonstrated drug delivery systems of polyether anhydrides for pulmonary delivery of drugs.^[23] The surface erosion properties of the ether anhydride provided a continuous drug-release profile over time spanning hours to days.

A current commercially available erodible system, Lupron Depot, is available from TAP Pharmaceuticals for prostate disease, endometriosis, and precocious puberty.^[24] The Lupron Depot system provides constant delivery of leuprolide at a rate of 120 ug/day for at least 12 months. As another example of eroding systems, injectable PGLA systems have been developed, made injectable by delivering in organic solvents, for drug delivery over periods of weeks to months.^[25] In these injectable PGLA systems, the polymer is dissolved in Food and Drug Administration (FDA)-approved water-miscible organic solvents such as dimethyl sulfoxide (DMSO) and then drug is added to form a homogeneous solution or suspension. When this solution or suspension is injected, the polymer precipitates due to insolubility in water as the organic solvent is replaced by water. The drug is then slowly released from the system. A major challenge in these systems is an initial burst of drug released at injection.^[25]

In the drug-polymer conjugate systems, the polymer-drug conjugate is often referred to as a prodrug. The polymer can be the polymer matrix of a drug delivery device, or the polymer may be a water-soluble polymer chain making the prodrug water soluble.^[26] This can be an important method for improving the solubility and delivery of poorly soluble drugs. In these systems, especially if the drug is inactive when attached to the polymer, the drug is released after the attachment between the polymer and the drug is broken. A number of factors can influence the drug delivery rate of polymer-bound drugs from matrix systems. These factors include.^[22]

1. The mode of degradation of the matrix (i.e., surface versus bulk, homogeneous versus heterogeneous).
2. The permeability of the matrix to water.
3. The diffusion coefficient of the released drug or prodrug in the matrix.
4. The solubility of the drug in water and in the matrix.
5. The rate of hydrolytic or enzymatic cleavage of the drug from the polymer.

In reservoir-type systems, different than in the matrix-type systems, the drug is enclosed within the lumen of a permeable membrane.^[27] The thickness and the permeability properties of this membrane control the rate of drug release from these types of devices. Zero-order release can be obtained from reservoir-type devices if the reservoir concentration is constant. This can occur when the drug is loaded within the reservoir above the saturation concentration and the diffusion through the permeable membrane is sufficiently slow to allow equilibrium of the concentration within the reservoir, particularly at the inner surface of the permeable membrane. A disadvantage to the reservoir systems is that the permeable membrane is generally constructed from nondegradable materials. If removal is required for safety or reapplication purposes, surgery would be required. These systems can also “dump” toxic quantities of the drug if the membrane is compromised. An example of a current reservoir system is Norplant, a silicone membrane-encapsulated hormone system for long-term contraception. The Norplant system was first seen in clinical trials in the early 1980s.^[27] Since then, Norplant has become an important controlled-release system available in the clinic.

More advanced systems are also under development using environmentally sensitive polymers to provide additional control mechanisms beyond diffusion and degradation for the delivery of drugs.^[21] Polymers that are sensitive to temperature, pH, electric field, ultrasound, magnetic field, and ionic strength have been used in development of controlled-release systems.^[28] In these systems, this environmental trigger can control the release of the drug. As an example, pH-sensitive

systems for oral delivery of insulin have been investigated to protect the insulin in the stomach but allow release in the intestine.^[8]

RELEASE PROFILES AND MODELS

In the polymeric systems, important release models have been developed to assist in understanding and designing these systems. For many drug delivery systems, zero-order delivery is ideal, while most polymeric systems provide first-order release. A truly zero-order delivery system would be time-independent with respect to release. This can occur when the concentration at the diffusion boundary is constant. On the other hand, in first-order delivery, the rate of release is dependent on the concentration of the drug in the device as it decreases by diffusion. Typically, controlled-release data are reported as mass released and as percent mass released or fraction mass released. In the reservoir and matrix system, diffusion becomes an important mechanism to control the rate of release. Because of this importance, a number of models and analytic methods have been developed to understand and predict the release. Insight into the release mechanism can often be gained by determining the dependence of the release profile on time. The profiles can be analyzed using Eq. 3, where M_t/M_∞ is the fraction of drug released at time t , t is the release time, k is a kinetic constant dependent on the system geometry and diffusion coefficients, and n is the diffusional release exponent.^[29]

$$\frac{M_t}{M_\infty} = k \cdot t^n \quad (3)$$

The diffusional release exponent found in Eq. 1 can provide indications regarding the drug-release mechanism. An exponent, n , of 0.5 suggests normal concentration-controlled Fickian diffusion, while n of 1.0 indicates zero-order (time-independent) release.^[29] Values between 0.5 and 1.0 indicate that there is some anomalous or non-Fickian diffusion occurring in the system.^[29] Exponent values greater than 1 indicate a super-release mechanism such as swelling-or degradation-enhanced release.^[29] For systems with Fickian diffusion, solutions have been derived for Fick's second law (conservation of mass) for a variety of different controlled-release system configurations including spherical, cylindrical, and flat sheet. Many of these solutions have been outlined by Saltzman.^[30]

BIOLOGICAL SYSTEMS

The final system type is biological. Biological systems provide controlled release of a drug by using living

cells or tissue. The primary difference between biological systems and the polymeric systems is the use of the living cells or tissues for the in-situ production of the delivered drug. Typically, the living cells or tissue are encapsulated in a semipermeable membrane (polymer), physically isolating the cells from the host immune system. Several polymers have been used to immunoisolate the cells including poly(hydroxyethyl methacrylate-co-methyl methacrylate)^[31] and alginate/poly-L-lysine.^[32] With this immune protection, different cell sources can be used: autologous, allogenic, xenogenic or genetically engineered. Autologous cells are cells derived from the organism to be implanted, allogenic cells are derived from the same species, and xenogenic cells are derived from another species.

Although not the only examples, two important drugs, for which this type of system is being developed, are L-dopa for the treatment of Parkinson's disease^[31] and insulin for the treatment of diabetes.^[32] These types of systems are often referred to as tissue-engineered or biohybrid systems. In the first example, Sefton et al. have demonstrated the microencapsulation of PC-12 cells in a copolymer of hydroxyethyl methacrylate-methyl methacrylate (HEMA-MMA) for the prolonged delivery of L-dopa.^[31] However, they also address the difficulty of maintaining the PC-12 cells in vivo after implantation and the difficulty of implanting microencapsulated cells. They describe that the biocompatibility and long-term viability of the system is contingent on the method and site of implantation.^[31] In the treatment of diabetes, several clinical trials have been conducted using alginate-polylysine-encapsulated islets of Langerhans, a subset of cells in the pancreas that provide insulin production and secretion.^[32] Clinical trials have been conducted using human and porcine islets.^[32] Some of the challenges faced in the islet encapsulation clinical trials are scarcity of the human tissue for implantation and the need for immunosuppression. However, approval for human trials in the United States was obtained, and VivoRx planned to start trials in the United States in 1998 for nonimmunosuppressed diabetics.^[32]

There are two main advantages to biological systems. First, biological systems provide physiological delivery of the drug. The drug is synthesized and released by the entrapped cells in response to physiological stimuli or in some physiological rate or pattern. As an example, encapsulated islets respond to glucose concentration with modified insulin production and release.^[33] Figure 4 presents rat islets of Langerhans that are immobilized within a 5-wt% poly(N-isopropylacrylamide-co-acrylic acid) hydrogel with 2-mol% acrylic acid as part of work^[33] toward the development of a refillable biohybrid artificial pancreas. Secondly, biological systems also provide in situ drug production. Therefore, these systems do not require



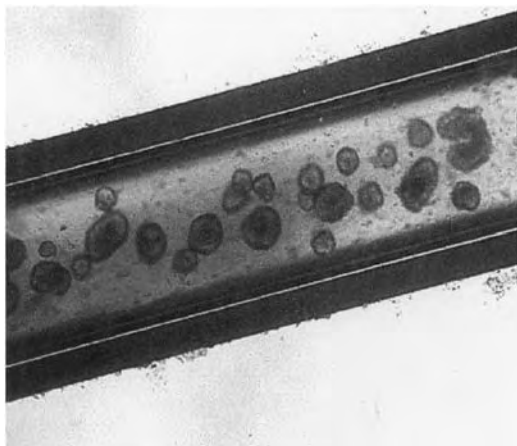


Fig. 4 Rat islets of Langerhans entrapped in 5-wt% hydrogels of poly(N-isopropylacrylamide-co-acrylic acid) with 2-mol% acrylic acid. Encapsulated islets are envisioned as a potential controlled-release system for insulin.

application of toxic quantities of the drug, guaranteeing that even device failure will not result in a bolus drug release. This provides extra protection against potentially dangerous drug overdoses. The in situ drug generation also allows continuous, potentially permanent, delivery of the drug over longer periods with low initial drug load.

Despite these advantages, there are a number of disadvantages that must be overcome prior to widespread clinical use of these systems. The main disadvantage to the biological systems is the need for immune protection. Unless the cells are autologous, these systems require either immune suppression or immune barriers. Although device failure would not result in bolus drug release, it would result in an immune response due to the immunogenicity of the entrapped cells. Also, the implanted cells have a finite lifespan, generally due to immune components or diffusion restrictions, requiring device replacement after a certain time.

REGULATORY ISSUES

Prior to marketing any of these controlled-release systems in the United States, approval must be obtained from the Food and Drug Administration of the U.S. federal government. However, the FDA considers many of these controlled-release systems as combination products. In 1991, the FDA defined [combination products] as (definition from 21 CFR Part 3, Subpart A, Section 3.2 (e)).^[34]

1. A product comprised of two or more regulated components, i.e., drug/device, biologic/device,

drug/biologic, or drug/device/biologic, that are physically, chemically, or otherwise combined or mixed and produced as a single entity;

2. Two or more separate products packaged together in a single package or as a unit and comprised of drug and device products, device and biological products, or biological and drug products;
3. A drug, device, or biological product packaged separately that according to its investigational plan or proposed labeling is intended for use only with an approved individually specified drug, device, or biological product where both are required to achieve the intended use, indication, or effect and where upon approval of the proposed product the labeling of the approved product would need to be changed, e.g., to reflect a change in intended use, dosage form, strength, route of administration, or significant change in dose; or
4. Any investigational drug, device, or biological product packaged separately that according to its proposed labeling is for use only with another individually specified investigational drug, device, or biological product where both are required to achieve the intended use, indication, or effect.

From the preceding definition, it is apparent that most controlled-release systems fit this classification, involving both device and drug and in some cases adding biologics. The combined nature of these devices complicates the regulatory environment, adding challenges to the premarket review, approval, and post-market surveillance processes. The three classes—devices, drugs, and biologics—are reviewed, approved, and monitored by different divisions or centers of the FDA. Device products are under the jurisdiction of the Center for Devices and Radiological Health (CDRH); drug products fall under the jurisdiction of the Center for Drug Evaluation and Research (CDER), while biologic products are controlled by the Center for Biologics Evaluation and Research (CBER). For products with combinations of drugs, devices, and biologics, it is not immediately clear which center should have jurisdiction, and it is necessary that multiple centers participate in the review, approval, and monitoring processes. Therefore, in 1991, the three centers entered into multiple intercenter agreements regarding the handling of combination products; these agreements can be found online at www.fda.gov/oc/combination/intercenter.html. For many controlled-release products, the intercenter agreement between CDRH and CDER is important.^[35] This agreement outlines the primary jurisdiction for many combination device/drug products.

In 2002, the Medical Device User Fee and Modernization Act of 2002^[36] helped to centralize the combination product review, approval, and postmarket

surveillance responsibilities of the FDA. This act amended the Food and Drug Act requiring the establishment of an Office of Combination Products within the FDA. In December of 2002, this office was established with the responsibilities to determine the center with primary jurisdiction and to coordinate multicenter input in the review, approval, and monitoring processes. This office is required to assign responsibility for these products to the appropriate center based on the products' "primary mode of action." Although multicenter participation can occur during the review, the center assigned primary jurisdiction defines what requirements must be met. As an example, systems found to have the drug as the primary mode of action are required to submit an Investigational New Drug (IND) application, while systems found to have the device as the primary mode of action are required to submit an Investigational Device Exemption (IDE) application to perform clinical trials. Additional information regarding regulatory issues for drug/device combination products, especially those where the primary function is "drug," can be found in a review by P. Marroun.^[37]

CONCLUSION

The primary goal of controlled-release technology is to improve the efficacy and safety of drug delivery by controlling the release rate and/or location. Numerous design approaches have been pursued in the development of controlled-release systems. Mechanical/electrical systems of insulin pumps combined with glucose sensors are the hope for improved delivery of insulin for type I and type II diabetics. Osmotic pumps are being used to deliver a variety of different drugs. Polymeric systems often rely on diffusion, swelling, and degradation to control the release of a compound. More advanced polymeric systems add potential environmental control by using materials that are sensitive to pH, temperature, or other stimuli. Biological systems combine living cells with devices to provide physiologically relevant delivery of important biological molecules. These systems have the potential to deliver their compound in response to physiologically specific signals or conditions and provide long-term delivery with low drug load. In all of these systems, the rate and timing of the release are extremely important and must be fully considered during development. Also, these systems have special requirements in the regulatory arena because of their characteristic combination of drug, biologics, and devices. Finally, controlled release provides spatial and/or temporal control that can be used to reduce toxicity, improve efficacy, bypass barriers, enhance solubility, and prolong activity for the delivery of many different drugs.

ACKNOWLEDGMENTS

The authors gratefully acknowledge Michael Caplan, Alyssa Panitch, and Colette Vernon for assistance in preparation of this manuscript.

ARTICLES OF FURTHER INTEREST

Drug Delivery, Controlled; Insulin Delivery; Poly (Glycolic Acid); Poly(lactic acid)s; Transdermal Drug Delivery; Stents

REFERENCES

1. Storm, P.B.; Moriarity, J.L.; Tyler, B.; Burger, P.C.; Brem, H.; Weingart, J. Polymer delivery of camptothecin against 9L gliosarcoma: Release, distribution, and efficacy. *J. Neuro-Oncol.* **2002**, *56* (3), 209–217.
2. Westphal, M.; Hilt, D.C.; Bortey, E.; Delavault, P.; Olivares, R.; Warnke, P.C.; Whittle, I.R.; Jaaskelainen, J.; Ram, Z. A phase 3 trial of local chemotherapy with biodegradable carmustine (BCNU) wafers (Gliadel wafers) in patients with primary malignant glioma. *Neuro-Oncology* **2003**, *5* (2), 79–88.
3. Lesser, G.; Grossman, S.; Leong, K.; Lo, H.; Eller, S. In vitro and in vivo studies of subcutaneous hydro-morphone implants designed for the treatment of cancer pain. *Pain* **1996**, *65*, 265–272.
4. Schiel, R. Continuous subcutaneous insulin infusion in patients with diabetes mellitus. *Therap. Apher.* **2003**, *7* (2), 232–237.
5. Renard, E.; Costalat, G.; Moran, B.; Shah, R.; Zhang, Y.N.; Villegas, D.; Kolopp, M.; Lebel, R.; Bringer, J. First combined implantations of a long-term IV glucose sensor and an intra-peritoneal insulin pump in diabetic patients. *Diabetes* **2001**, *50*, A3.
6. Carter, B.L. Optimizing delivery systems to tailor pharmacotherapy to cardiovascular circadian events. *Am. J. Health-Syst. Pharm.* **1998**, *55*, S17–S23.
7. Koning, G.A.; Kamps, J.; Scherphof, G.L. Efficient intracellular delivery of 5-fluorodeoxyuridine into colon cancer cells by targeted immunoliposomes. *Cancer Detec. Prev.* **2002**, *26* (4), 299–307.
8. Lowman, A.; Peppas, N. Complexation graft copolymers as oral drug delivery systems. *Polym. Preprints* **1997**, *38* (2), 566–567.
9. Blanco, M.D.; Garcia, O.; Gomez, C.; Sastre, R.L.; Teijon, J.M. In-vivo drug delivery of 5-fluorouracil using poly-(2-hydroxyethyl methacrylate-co-acrylamide) hydrogels. *J. Pharm. Pharmacol.* **2000**, *52* (11), 1319–1325.
10. Torchilin, V.P.; Lukyanov, A.N.; Gao, Z.G.; Papahadjopoulos-Sternberg, B. Immunomicelles: Targeted pharmaceutical carriers for poorly soluble drugs. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (10), 6039–6044.



11. Dash, A.K.; Cudworth, G.C. Therapeutic applications of implantable drug delivery systems. *J. Pharmacol. Toxicol. Methods* **1998**, *40* (1), 1–12.
12. Selam, J.L.; Charles, M.A. Devices for insulin administration. *Diabetes Care* **1990**, *13* (9), 955–979.
13. Skitzki, J.; Chang, A. Hepatic artery chemotherapy for colorectal liver metastases: Technical considerations and review of clinical trials. *Surg. Oncol.* **2002**, *11* (3), 123–135.
14. Rise, M. Instrumentation for neuromodulation. *Arch. Med. Res.* **2000**, *31* (3), 237–247.
15. Schade, D.S.; Eaton, R.P.; Edwards, W.S.; Doberneck, R.C.; Spencer, W.J.; Carlson, G.A.; Bair, R.E.; Love, J.T.; Urenda, R.S.; Gaona, J.I. A remotely programmable insulin delivery system- successful short-term implantation in man. *JAMA J. Am. Med. Assoc.* **1982**, *247* (13), 1848–1853.
16. Theeuwes, F.; Yum, S. Principles of the design and operation of generic osmotic pumps for the delivery of semisolid or liquid drug formulations. *Ann. Biomed. Eng.* **1976**, *4*, 343–353.
17. Makhija, S.; Vavia, P. Controlled porosity osmotic pump-based controlled release system of pseudoephedrine- I. Cellulose acetate as a semipermeable membrane. *J. Control. Release* **2003**, *89* (1), 5–18.
18. Wong, P.S.L.; Gupta, S.K.; Stewart, B.E. Osmotically Controlled Tablets. In *Modified-Release Drug Delivery Technology*; Rathbone, M.J., Hadgraft, J., Roberts, M.S., Eds.; Marcel Dekker, Inc.: New York, 2003; 101–114.
19. Uchiyama, T.; Watanabe, J.; Ishihara, K. Pressure-induced change in permeation of insulin through a polymer alloy membrane for an implantable insulin pump. *J. Membr. Sci.* **2002**, *210* (2), 423–431.
20. Langer, R. Controlled release- a new approach to drug delivery. *Technol. Rev.* **1981**, *83* (5), 26–34.
21. Brannon-Peppas, L. Polymers in controlled drug delivery. *Med. Plast. Biomater. Mag.* **1997**, **November**, 34
22. Ron, E.; Langer, R. Erodible Systems. In *Treatise on Controlled Drug Delivery*; Kydonieus, A., Ed.; Marcel Dekker, Inc.: New York, 1992; 199–224.
23. Fu, J.; Fiegel, J.; Krauland, E.; Hanes, J. New polymeric carriers for controlled drug delivery following inhalation or injection. *Biomaterials* **2002**, *23* (22), 4425–4433.
24. Lupron Depot TAP Pharmaceuticals, www.lupron.com Accessed Sept. 9, 2003.
25. Wang, L.; Kleiner, L.; Venkatraman, S. Structure formation in injectable poly(lactide-co-glycolide) depots. *J. Control. Release* **2003**, *90* (3), 345–354.
26. Greenwald, R.B.; Choe, Y.H.; McGuire, J.; Conover, C.D. Effective drug delivery by PEGylated drug conjugates. *Adv. Drug Deliv. Rev.* **2003**, *55* (2), 217–250.
27. Sivin, I.; Diaz, S.; Holma, P.; Alvarezsanchez, F.; Robertson, D.N.; Stern, J. Long-term experience with norplant-Tm in international clinic trials. *Contracept. Deliv. Syst.* **1982**, *3* (3-4), 90.
28. Kim, S. Temperature Sensitive Polymers for Delivery of Macromolecular Drugs. In *Advanced Biomaterials in Biomedical Engineering and Drug Delivery Systems*; Ogata, N., Kim, S., Feijen, J., Eds.; Springer: Tokyo, 1996; 126–133.
29. Korsmeyer, R.W.; Gurny, R.; Doelker, E.; Buri, P.; Peppas, N.A. Mechanisms of solute release from porous hydrophilic polymers. *Int. J. Pharm.* **1983**, *15* (1), 25–35.
30. Saltzman, W.M. *2001. Drug Delivery: Engineering Principles for Drug Therapy*; Oxford University Press: New York, 2001; 235–279.
31. Vallbacka, J.J.; Nobrega, J.N.; Sefton, M.V. Tissue engineering as a platform for controlled release of therapeutic agents: Implantation of microencapsulated dopamine producing cells in the brains of rats. *J. Control. Release* **2001**, *72* (1–3), 93–100.
32. Soon-Shiong, P. Treatment of type I diabetes using encapsulated islets. *Adv. Drug Deliv. Rev.* **1999**, *35* (2–3), 259–270.
33. Vernon, B.; Kim, S.W.; Bae, Y.H. Insulin release from islets of Langerhans entrapped in a poly(N-isopropylacrylamide-co-acrylic acid) polymer gel. *J. Biomater. Sci. Polym. Ed.* **1999**, *10* (2), 183–198.
34. Code of Federal Regulations. *Title 21*, 2003; Vol. 1. 21CFR3.
35. Food and Drug Administration Intercenter agreement between the center for drug evaluation and research and the center for devices and radiological health. HFA-305. *Food Drug Admin.* **1991**.
36. *Medical Device User Fee and Modernization Act of 2002 PL 107-250*; Office of the Federal Register (OFR), National Archives and Records Administration (NARA), 2002; 403.
37. Marroum, P.J. Regulatory Issues Relating to Modified-Release Drug Formulations. In *Modified-Release Drug Delivery Technology*; Rathbone, M.J., Hadgraft, J., Roberts, M.S., Eds.; Marcel Dekker, Inc.: New York, 2003; 943–975.

Corneal Implants

Traian V. Chirila

Celia R. Hicks

Centre for Ophthalmology and Visual Science, Lions Eye Institute, Nedlands,
Western Australia, Australia



INTRODUCTION

Over the last 30 years, many surgical procedures have been devised to modify the refractive power of the eye in order to improve vision.^[1] Being the most accessible layer of the ocular system and providing about 70% of the total dioptric power of the human eye, the cornea is a preferred candidate for these procedures. Refractive change can be achieved by increasing the cornea's refractive power either by altering, in part, its refractive index by means of an insert or changing the radius of curvature of its anterior surface where, at the air-tissue interface, most of the refraction of light occurs. This can be done either by means of an insert, surface sculpting, or precisely locating cuts in the cornea. These surgical procedures are included under the umbrella term of refractive corneal surgery or refractive keratoplasty.^[2] Many refractive procedures require implantation of inserts in which human donor corneal tissue and collagen have been used, but this has been associated with problems such as immune response, proteolytic degradation, and, in the case of donor tissue, limited supply. With the advent of synthetic polymers, artificial inserts became an established alternative. Surgery on the cornea may be also performed when the cornea becomes opaque due to disease or trauma causing poor vision or blindness such that its central part must be excised and replaced. This procedure is known as penetrating keratoplasty (PK) when the replacement is donor tissue. Interest has been also directed at developing synthetic alternatives to donor tissue for the replacement of the cornea. Historical roots of this procedure, known as prosthokeratoplasty, can be traced back to the 18th century.^[3]

This article focuses on the use of polymeric biomaterials as corneal implants for refractive purposes or full-thickness corneal replacement as an artificial cornea. Since the 1960s, numerous research groups have examined the responses of corneal tissue to implanted biomaterials in terms of inflammation, opacification, tissue melting, and refractive changes. Important issues that soon became apparent to researchers were the thickness, diameter, edge profile, refractive index, water content, and porosity of the materials. Much of

the early research involved implantation of biomaterial samples at approximately midstromal depth within animal corneas and assessing clinical and histological outcomes. Pioneers in this work included Krwawicz, Knowles, Choyce, and Dohlman. By the end of the 1970s, it was well established that certain materials could be retained in the medium- to long-term and that both clarity and refractive changes could be achieved. This work led to the development of artificial refractive devices and further progression in the development of artificial corneas (keratoprotheses).

GLOSSARY

Bowman's layer: an acellular layer in the cornea that is located between the epithelium and stroma.

Descemet's membrane: a basement membrane in the cornea that is located between the stroma and endothelium.

Endothelium: a single, ordered layer of hexagonal cells on the posterior surface of the cornea.

Epithelium: the external layer of the cornea consisting of specific cells organized into five to seven layers.

Hyperopia/hypermotropia: the refractive status of the eye that is either too short or in which its focusing apparatus (cornea, lens) is too weak so that light rays come into focus behind the retina, and additional converging power is required through surgical alteration or optical aids.

Intraocular lens (IOL): a synthetic replacement for the natural crystalline lens that may be removed due to cataract, trauma, or refractive reasons.

Keratoprosthesis (KPro): a synthetic, or artificial, cornea, i.e. a full-thickness replacement for a diseased cornea that is intended to maintain ocular integrity and allow transmission and refraction of light.

Melting: a loss of corneal stromal tissue through pathologically induced enzymatic degradation.

Myopia: the refractive status of the eye that is either too long or its focusing apparatus (cornea, lens) is too strong so that light rays come into focus in front of the

retina, and diverging power is required through surgical alteration or optical aids.

Penetrating keratoplasty (PK): a replacement of a diseased cornea with a human donor corneal graft.

Prosthokeratoplasty: a surgical procedure for insertion of a keratoprosthesis.

Pseudophakia: a removal and replacement of the natural crystalline lens with a synthetic IOL. An eye with its natural crystalline lens removed and not replaced is called aphakic.

Stroma: the main layer of the cornea comprising about 90% of the total corneal thickness and consisting of collagen fibrils, proteoglycan ground substance, water, and a small number of specialized fibroblasts (keratocytes).

IMPLANTS FOR REFRACTIVE KERATOPLASTY

Lamellar refractive keratoplasty constitutes the main group of procedures within refractive keratoplasty^[2] and includes all procedures involving addition of material by placing a lenticule or a ring on or within the cornea. Three of these procedures (keratophakia, epikeratoplasty, and intracorneal ring insertion) may make use of synthetic polymers (Fig. 1) and will be further discussed.

Implants for Keratophakia

Keratophakia involves placement of a lens within the cornea (Fig. 1) to alter its refractive power either by altering its anterior radius of curvature or incorporating a lens of higher refractive index. Such lenses, usually termed lenticules or inlays, can be made of donor tissue (homoplastic) or, as will be discussed here, a synthetic material (alloplastic). The technique has been felt most suitable for aphakic patients who are unsuited for intraocular surgery or contact lens wear. Usually, a lamellar disc of cornea is excised, the

lenticule is placed in the lamellar bed, and the disc is sutured into its original position. Early proof-of-principle for the ability of polymer inserts to be retained and cause permanent refractive alterations was provided by experiments with silicone rubber and poly(methyl methacrylate) (PMMA), which caused little tissue response.^[4] Other researchers felt that silicone rubber was more suited to use as a barrier implant, which causes a desirable degree of anterior dehydration in edematous corneas.^[5]

Hydrogels are hydrophilic polymers that are able to swell in water without dissolution due to the existence of water-attracting functional groups in the polymer chains. Attached water molecules act to hold the polymer strands apart, which results in an inherent structural microporosity. The best known representative of this class is poly(2-hydroxyethyl methacrylate) (HEMA), which is used in a large range of biomaterial applications and attracted interest as a material for keratophakia in early stages.^[6] McCarey and co-workers used copolymers of HEMA to improve the solute permeability and aimed for an ideal implant material that was described as transparent, chemically inactive, permeable, nonabsorbable, dimensionally stable, soft, and pliable.^[7] These researchers established^[8] that lenticules made of copolymers of HEMA with methacrylic acid (MAA) and/or 1-vinyl-2-pyrrolidinone (VP) permitted bulk flow of water and nutrients, a refractive change could be achieved, and the endothelium and epithelium remained essentially normal. But the researchers also established that the refractive change could be unpredictable when placed in middepth pockets. It was appreciated that, to achieve a significant alteration in curvature of the anterior corneal surface, the surgical technique must breach the strong Bowman's layer, which lies just deep to the corneal epithelium, and through most of its circumference to allow the tissue to deform. The earliest experiments used convex implants to steepen the cornea and correct aphakia or hyperopia in which greater refractive power was required. But, researchers then found that doughnut-shaped hydrogel implants could cause corneal flattening and correct myopia.^[9] A commercial lens became available, the Keratogel[®] intracorneal lens (Allergan Medical Optics, United States of America), manufactured from Lidofilcon[®] A (an MMA/VP copolymer), which was evaluated in human clinical trials. This insert provided significant refractive corrections—initially in aphakic patients—although best-corrected visual acuity took several months to return to its preoperative level. The outcomes were often disappointing when large hydrogel implants (made from Lidofilcon[®] A or Perfilcon[®] A, a HEMA/VP/MAA terpolymer) were implanted since poor predictability and astigmatism were common features. Successful implantation of these hydrogel lenses depends

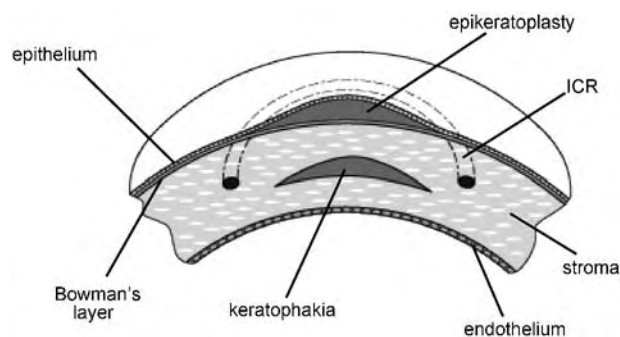


Fig. 1 Diagram of the cornea showing the major structural layers and position of refractive implants.

upon skillful manipulation of microkeratomes and is, thus, a more complex procedure than implantation of a high refractive index lens using freehand stromal dissection.

An alternative approach to implanting hydrogels and relying upon creating a predictable change in anterior surface curvature was proposed by Choyce.^[10] Namely, this approach was to implant intrastromally a lens that was sufficiently thin enough not to affect corneal curvature but with a significantly higher refractive index so that it would provide an additional refractive interface when implanted deeply within the cornea. The ideal diameter of a lenticule implant was proposed to lie within 4–5 millimeters to ensure that the size of the impermeable device does not affect anterior nutrition while the optical aberrations and decentration effects were minimized. A polysulfone produced from Bisphenol A and 4,4'-dichlorophenylsulfone was proposed and assessed for this purpose. Adverse events including cloudiness, which may have been related to instability to UV light, were reported^[11] along with lipid deposits and inflammatory reactions in some cases. Fenestration of the material to improve permeability and tolerance was proposed but, in turn, caused optical problems.

Recent work to address the limitations of hydrogel intracorneal implants for keratophakia has been conducted at Anamed, Inc. (United States of America). Their PermaVision[®] lenticule, which was made from a proprietary hydrogel (Nutrapore), was evaluated in animals^[12] with generally acceptable results in the short- to medium term. Although, it appears that safety problems are likely in the long term and may limit the use of this device. Synthetic intracorneal lenses are yet to have a proven place in refractive surgery.

Implants for Epikeratoplasty

Epikeratoplasty, a technique first suggested by Werblin and Kaufman in 1979,^[13] involves the placement of a lenticule superficially, just beneath the corneal epithelium, by suturing it to a groove in the de-epithelialized surface of the cornea and allowing re-epithelialization over it (Fig. 1). It has an advantage over keratophakia by avoiding incisions through the optical zone. Progress in this field was reviewed by Thompson and coworkers.^[14] The positioning of the lenticule anterior to Bowman's membrane, an important structural layer of the cornea, makes the procedure theoretically completely reversible and avoids adverse reaction to stromal healing. But, it does depend upon achieving adhesion of the lenticule and a stable overlying epithelialization. This has been achieved for donor tissue inserts but not to date for

synthetic inserts in such a position. The lenticule is concave when the procedure is performed for correction of myopia and convex when it is performed in an aphakic or hyperopic eye.

The use of synthetic lenticules for this application was explored during the 1980s and, due to their location just deep to the epithelium, a high degree of permeability was essential. Hydrogels developed for contact lenses (e.g., Lidofilcon[®] A) were assessed for this purpose, and it was further realized that surface modification with fibronectin, epidermal growth factor, or other biological adhesion promoters would be desirable to enhance epithelialization. Collagen-hydrogel combinations were also proposed, but disadvantages include a susceptibility to collagenolytic degradation. Other researchers achieved the epithelialization of poly(vinyl alcohol) (PVA) copolymers after coating them with basement membrane components, but no lenticules for clinical use were produced. Synthetic lenticules can be made with great accuracy, avoid disease transmission, compensate for the lack of donor availability, and are not subject to scarring or remodelling. Additionally, it was realized that the option of postoperative excimer laser ablation of the lenticule could provide for fine-tuning of the refractive outcome.^[15] However, there are significant biomaterial-associated challenges when a synthetic lenticule is used since the polymer needs not only to remain inert but must actively promote recovery and maintenance of the epithelium. Fixation methods attempted included mechanical, i.e., surgical dissection of a cut ring, laser welding, biological nontoxic glues, and suturing. However, suturing is difficult due to the delicacy of the thin layer. In practice, reversibility is often limited by residual scarring. Since the refractive outcomes from pre-lathed donor tissue implants were not as good as had been expected, they are no longer supplied. Clinical interest in the technique, whether with donor tissue or synthetic lenticules, has abated. Research interest continues in developing a synthetic onlay, which is sometimes inappropriately referred to as an artificial cornea.^[16] Some of the current research is focused on polymers that will allow re-epithelialization of the onlay. Promising results in animal experiments have been reported with perfluoropolyether-based polymers produced by the photo-oxidative polymerization of perfluoroalkylenes and surface-modified by covalent bonding of components of extracellular matrix.^[16,17]

Intracorneal Ring

An intracorneal ring, known also as intrastromal corneal ring (ICR), in the mid-peripheral cornea (Fig. 1) could either steepen the central corneal radius of



curvature and increase its power if the ring is tightened in situ or would flatten the central cornea and provide a myopic correction if the ring maintains its position or tends to stretch peripherally. Important variables include the diameter and thickness of the ring, the width of the channels dissected for it to lie within, and the material itself. Generally, the channel is cut at about two-thirds of the stromal depth by using a special circular cutter that makes only a small entry incision. So far, the preferred material for ICRs was exclusively PMMA. Proof-of-principle was first demonstrated in the rabbit model and then in eye bank human corneas.^[18] The bulk of the device itself was found to contribute to central corneal flattening. Results of ICR in humans with low myopia were good with reversibility and lack of significant complications.^[19] Higher degrees of myopia were treated less reliably, and the procedure is not without risks, including infection or perforation of the anterior chamber. Commercially known as Intacs[®] (KeraVision, United States of America), the devices are approved for use in low- to moderate myopia and myopic astigmatism. By the same principle, injection of a synthetic gel has a similar effect.^[20] In this procedure, known as gel injection adjustable keratoplasty (GIAK), crosslinked polymer gels, such as polysiloxanes^[20] or poly(ethylene oxide),^[21] have been used so far in eye bank corneas or animals.

KERATOPROSTHESIS

A true artificial (synthetic) cornea, or keratoprosthesis (KPro), is meant to totally replace the central part of an opaque cornea. Figure 2 shows three basic conceptual forms of KPros and their placement within the cornea. An artificial cornea is not merely a refractive device since its purpose is to restore vision lost due to corneal opacification in cases in which the transplantation of donor corneal tissue has little chance of success. Research on KPro has a long history, which includes substantial progress along with many failures, as shown in some major reviews.^[22-25] The most serious complication of prosthokeratoplasty is the loss, or extrusion, of the KPro, which is usually preceded by tissue melting around the prosthetic rim. Other complications include the formation of retroprosthetic membranes, inflammation, infection, retinal detachment, and cataract. Despite its complications, prosthokeratoplasty is the only hope for the visual recovery of those patients in which the transplantation of donor cornea is impracticable.

Due to a high incidence of extrusion, the early KPros made of glass or quartz were abandoned at the end of the 19th century, and, for an ensuing long period of time, no significant research was reported on KPro. The introduction in the late 1940s of synthetic

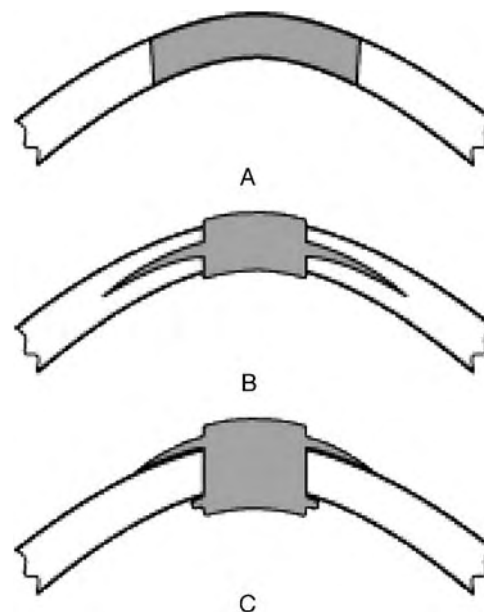


Fig. 2 Basic types of keratoprosthesis. A: full thickness; B: core-and-skirt; C: nut-and-bolt.

polymers as materials marked the beginning of a modern era in the development of KPro. PMMA was the first polymer to be used for developing a variety of one- or two-piece KPros,^[26] and it is still used in some models.^[27] The introduction of lateral-supporting perforated or meshwork plates (flanges) around the PMMA optical cylinder led to the so-called core-and-skirt KPros, which were aimed at promoting integration of the host cornea with the periphery of the KPro (Fig. 2B). However, these devices led to only slightly better clinical outcomes when compared with nonperforated devices. Fixation of the implants on both external and internal aspects of the host cornea, as exemplified in the nut-and-bolt KPro designs (Fig. 2C), led to some improvement in the clinical outcome. But, the presence of a retrocorneal nut remained a source of complications. PMMA is still preferred by many researchers as a material for artificial cornea, and models comprising various fixation means of increasing complexity are still a frequent feature of contemporary KPro research.

The research on KPro was generally conducted along two directions. The first trend was aimed at finding a suitable synthetic polymer with the underlying criterion being ocular biocompatibility. As many materials more or less fulfill this prerequisite, the range of synthetic polymers proposed or used for optical core and prosthetic skirts was rather extensive,^[24] including polysiloxanes, PHEMA, polytetrafluoroethylene (PTFE), poly(ethylene terephthalate) (usually Dacron[®]), polyurethanes, polyalkylenes, poly(vinyl alcohol), polycarbonates, poly(glyceryl monomethacrylate), poly(ethylene oxide), carbon fibers, and carbon fiber composites. A second

research trend was triggered by the perceived need for biointegration at the polymer-tissue interface. As an unavoidable consequence, KPros of high level of complexity have been developed, which in some cases were indeed retained for a long time in the host cornea. However, the sophistication involved in their design rendered these KPros unavailable off-the-shelf to surgeons and patients. A revealing example is perhaps one of the Cardona's through-and-through KPro. It consisted of a central PMMA cylinder (peripherally pigmented to prevent glare) surrounded by a perforated PTFE plate, which was reinforced with both a mesh of Dacron[®] and a layer of autologous tissue (usually periosteum excised from the same patient). As this KPro had to pass through the eyelid, extensive surgery was needed, including the total removal of the crystalline lens and iris, the partial removal of the vitreous, the excision of eyelid muscles and, finally, the permanent closure of the eyelid skin. In other KPro models, the central PMMA cylinder was surrounded by autologous tissue, such as bone and tooth (osteodontokeratoprosthesis), and covered with mucous membrane, which requires multistage surgery on patients. Although providing some chance for visual improvement in the long term, such KPros are acceptable only for the most severe bilaterally blind cases due to their complexity, complicated and lengthy surgery, aesthetically disfiguring result, and visual fields lower than 30 degrees.

By the end of 1980s, it became clear that acceptable integration of the host corneal tissue into the prosthetic peripheral region could not be assured by large holes in the skirts or through autologous tissue. KPro researchers began to understand that this integration should occur through growth within voids of a width in the micron range, which is not much larger than the size required to accommodate the stromal fibroblasts, and these voids should be of unlimited length through the material, i.e., contiguous. In addition, it was realized that the simplest conceptual form of KPro (Fig. 2A) could be employed without any need for mechanical fixation means. The attempts to achieve this goal led to the development of core-and-skirt KPros with a porous skirt.

Keratoprotheses with Porous Skirts

A relatively small number of research groups pioneered independently KPros comprising skirts with a true porous structure. From a tissue-engineering standpoint, the skirts can be regarded as nonbiodegradable scaffolds. These developments were extensively reviewed elsewhere.^[24,28] One of the first KPro models with a porous skirt comprised a polysiloxane core surrounded by carbon fibers,^[29] which was eventually abandoned despite partial success in animal

experiments. Other researchers chose a melt-blown fibrous web of polybutylene/polypropylene blend for the skirt and pol(vinyl alcohol-co-vinyl acetate) copolymers as core materials.^[30] This model was also abandoned after little success in animal trials. Legeais' group in France developed KPros with skirts of porous PTFE. Their latest model, in which the core was made from polysiloxane coated with PVP,^[31] has been approved for distribution and clinical use in Europe as BioKP[®] Sil (FCI, France). The KPro developed by Caldwell in the United States of America, in which the skirt was also made from porous PTFE and the core from transparent polyurethane, had only partial success in human patients.^[32] The concept of using hydrophilic polymers in a KPro has been applied in the development of a model in which the core and the porous skirt were both made from PHEMA.^[33,34] The main features of this KPro include a PHEMA sponge as a skirt, which was created by phase-separation polymerization, and interpenetrating polymer networks as an attachment between core and skirt. In clinical human trials,^[35,36] this model (Fig. 3) led to a lower risk of some of the complications traditionally associated with prosthokeratoplasty. The device has obtained widespread regulatory approval for distribution and use in human patients, including the U.S. FDA approval, and is currently marketed as AlphaCor[™] (Argus Biomedical, Australia). In order to promote better fixation of the device within the host tissue, some researchers included not only a porous skirt but also elements for additional anchoring to the sclera. For instance, the Seoul-type KPro consists of a PMMA optical cylinder, a skirt form porous polyurethane or fibrous polypropylene, and loops made of monofilament polypropylene. This KPro has only been evaluated in a small number of human patients.^[37]

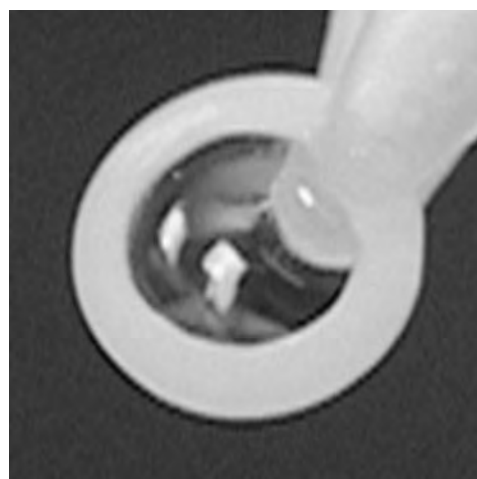


Fig. 3 The AlphaCor[™] keratoprosthesis.

CONCLUSION

Corneal refractive devices for visual correction and artificial corneas for visual restoration are becoming major fields for the application of polymeric biomaterials. However, the implantation of polymers into the cornea remains significantly associated with clinical complications. To solve such problems, researchers have adopted multidisciplinary approaches that enabled them to optimize polymer selection and their structural modification. The range of synthetic polymers proposed and evaluated as refractive devices (corneal onlays, corneal inlays, intracorneal rings) is rather limited and requires more work and diversification. The development of artificial corneas has involved a much greater number of polymers, but devices produced to date are justifiable for implantation only in high-risk cases. An artificial cornea of sufficiently low risk to replace donor grafts in all cases is still some years away.

ARTICLES OF FURTHER INTEREST

Hydrogels; Ocular Implants; Ocular Implants for Drug Delivery; Tissue Engineering of Cornea

REFERENCES

1. Stulting, R.D.; Lahnert, W.J.; Carr, J.D. Advances in refractive surgery: 1975 to the present. *Cornea* **2000**, *19* (5), 741–753.
2. Waring, G.O. Making sense of ‘keratospark’: A classification of refractive corneal surgery. *Arch. Ophthalmol.* **1985**, *103* (10), 1472–1477.
3. Chirila, T.V.; Hicks, C.R. The origins of the artificial cornea: Pellier de Quengsy and his contribution to the modern concept of keratoprosthesis. *Gesnerus* **1999**, *56* (1–2), 96–106.
4. Belau, P.G.; Dyer, J.A.; Ogle, K.N.; Henderson, J.W. Correction of ametropia with intracorneal lenses. *Arch. Ophthalmol.* **1964**, *72* (10), 541–547.
5. Brown, S.I.; Dohlman, C.H. A buried corneal implant serving as a barrier to fluid. *Arch. Ophthalmol.* **1965**, *73* (5), 635–639.
6. Mester, U.; Roth, K.; Dardenne, M.U. Trial with 2-hydroxy-ethyl-methacrylate lenses as keratophakia material [Ger.]. *Ber. - Dtsch. Ophthalmol. Ges.* **1974**, *72*, 326–329.
7. McCarey, B.E. Current status of refractive surgery with synthetic intracorneal lenses. *Refract. Corneal Surg.* **1990**, *6* (1), 40–46.
8. McCarey, B.E.; van Rij, G.; Beekhuis, W.H.; Waring, G.O. Hydrogel keratophakia: A freehand pocket dissection in the monkey model. *Br. J. Ophthalmol.* **1986**, *70* (3), 187–191.
9. Binder, P.S. Hydrogel implants for the correction of myopia. *Curr. Eye Res.* **1983**, *2* (7), 435–441.
10. Choyce, D.P. The correction of refractive errors with polysulfone corneal inlays: A new frontier to be explored? *Trans. Ophthalmol. Soc. U. K.* **1985**, *104* (3), 332–342.
11. Kirkness, C.M.; Steele, A.D.M.; Garner, A. Polysulfone corneal inlays. Adverse reactions: A preliminary report. *Trans. Ophthalmol. Soc. U. K.* **1985**, *104* (3), 343–350.
12. Ismail, M.M. Correction of hyperopia with intracorneal implants. *J. Cataract Refract. Surg.* **2002**, *28* (3), 527–530.
13. Werblin, T.P. Lamellar refractive surgery: Where have we been and where are we going? *Refract. Corneal Surg.* **1989**, *5* (3), 167–176.
14. Thompson, K.P.; Hanna, K.; Waring, G.O.; Gipson, I.; Liu, Y.; Gailitis, R.P.; Johnson-Wint, B.; Green, K. Current status of synthetic epikeratoplasty. *Refract. Corneal Surg.* **1991**, *7* (3), 240–248.
15. Thompson, K.P.; Hanna, K.D.; Waring, G.O. Emerging technologies for refractive surgery: Laser adjustable synthetic epikeratoplasty. *Refract. Corneal Surg.* **1989**, *5* (1), 46–48.
16. Xie, R.Z.; Stretton, S.; Sweeney, D.F. Artificial cornea: Towards a synthetic onlay for correction of refractive error. *Biosci. Rep.* **2001**, *21* (4), 513–536.
17. Evans, M.D.M.; McLean, K.M.; Hughes, T.C.; Sweeney, D.F. A review of the development of a synthetic corneal onlay for refractive correction. *Biomaterials* **2001**, *22* (24), 3319–3328.
18. Burris, T.E.; Ayer, C.T.; Evensen, D.A.; Davenport, J.M. Effects of intrastromal corneal ring size and thickness on corneal flattening in human eyes. *Refract. Corneal Surg.* **1991**, *7* (1), 46–50.
19. Hoh, H.; Rehfeldt, K.; Reiss, G. Correction of myopia with the intracorneal ring [Ger.]. *Ophthalmologie* **1999**, *96* (11), 717–723.
20. Simon, G.; Parel, J.-M.; Lee, W.; Kervick, G.N. Gel injection adjustable keratoplasty. *Graefe’s Arch. Clin. Exp. Ophthalmol.* **1991**, *229*, 418–424.
21. Legeais, J.-M.; Parel, J.-M.; Savoldelli, M.; Villain, F.L.; D’Hermies, F.; Renard, G.; Pouliquen, Y. Reticulated polyethylene oxide for gel injection adjustable keratoplasty. Biocompatibility in critical situation [Fr.]. *J. Franc. Ophthalmol.* **1997**, *20* (1), 31–36.
22. Day, R. Artificial cornea implants. *Trans. Am. Ophthalmol. Soc.* **1957**, *55*, 455–475.
23. Hicks, C.R.; Fitton, J.H.; Chirila, T.V.; Crawford, G.J.; Constable, I.J. Keratoprosthesis: Advancing towards a true artificial cornea. *Surv. Ophthalmol.* **1997**, *42* (2), 175–189.
24. Chirila, T.V.; Hicks, C.R.; Dalton, P.D.; Vijayasekaran, S.; Lou, X.; Hong, Y.; Clayton, A.B.; Ziegelaar, B.W.; Fitton, J.H.; Platten, S.; Crawford, G.J.; Constable, I.J. Artificial cornea. *Prog. Polym. Sci.* **1998**, *23* (3), 447–473.
25. Mannis, M.J.; Dohlman, C.H. The Artificial Cornea: A Brief History. In *Corneal Transplantation: A History in Profiles*; Mannis, M.J., Mannis, A.A., Eds.; Hirschberg History of Ophthalmology—The Monographs, J.P.

- Wayenborgh: Oostende, Belgium, 1999; Vol. 6, 321–335.
26. Cardona, H. Keratoprosthesis: Acrylic optical cylinder with supporting intralamellar plate. *Am. J. Ophthalmol.* **1962**, *54* (2), 284–294.
 27. Dohlman, C.H.; Doane, M.G. Some factors influencing outcome after keratoprosthesis surgery. *Cornea* **1994**, *13* (3), 214–218.
 28. Chirila, T.V. Modern artificial corneas: The use of porous polymers. *Trends Polym. Sci.* **1994**, *2* (9), 296–300.
 29. Kain, H.L. A new concept for keratoprosthesis [Ger.]. *Klin. Mon.bl. Augenheilkd.* **1990**, *197*, 386–392.
 30. Trinkaus-Randall, V.; Wu, X.Y.; Tablante, R.; Tsuk, A. Implantation of a synthetic cornea: Design, development and biological response. *Artif. Organs* **1997**, *21* (11), 1185–1191.
 31. Legeais, J.-M.; Renard, G. A second generation of artificial cornea (Biokpro II). *Biomaterials* **1998**, *19* (16), 1517–1522.
 32. Caldwell, D.R. The soft keratoprosthesis. *Trans. Am. Ophthalmol. Soc.* **1997**, *95*, 751–802.
 33. Chirila, T.V.; Vijayasekaran, S.; Horne, R.; Chen, Y.-C.; Dalton, P.D.; Constable, I.J.; Crawford, G.J. Interpenetrating polymer network (IPN) as a permanent joint between the elements of a new type of artificial cornea. *J. Biomed. Mater. Res.* **1994**, *28* (6), 745–753.
 34. Chirila, T.V. An overview of the development of artificial corneas with porous skirts and the use of PHEMA for such an application. *Biomaterials* **2001**, *22* (24), 3311–3317.
 35. Crawford, G.J.; Hicks, C.R.; Lou, X.; Vijayasekaran, S.; Tan, D.; Mulholland, B.; Chirila, T.V.; Constable, I.J. The Chirila keratoprosthesis: Phase I human clinical trials. *Ophthalmology* **2002**, *109* (5), 883–889.
 36. Hicks, C.R.; Crawford, G.J.; Lou, X.; Tan, D.T.; Snibson, G.R.; Sutton, G.; Downie, N.; Werner, L.; Chirila, T.V.; Constable, I.J. Corneal replacement using a synthetic hydrogel cornea, AlphaCor™: Device, preliminary outcomes and complications. *Eye* **2003**, *17* (3), 385–392.
 37. Kim, M.K.; Lee, J.L.; Wee, W.R.; Lee, J.H. Seoul-type keratoprosthesis: Preliminary results of the first 7 human cases. *Arch. Ophthalmol.* **2002**, *120* (6), 761–766.



Corrosion

Rachel L. Williams

Department of Clinical Engineering, University of Liverpool, Liverpool, United Kingdom

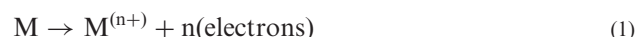
INTRODUCTION

Corrosion of metallic biomaterials occurs as an electrochemical reaction at the surface of the metal with the biological fluids acting as the electrolyte.^[1-3] The consequences of corrosion are changes in the surface chemistry of the metal and the accumulation of metal ions in the electrolyte. Both are important in terms of the behavior of an implanted device. These, along with measurement of the electrochemical reactions, form the basis for evaluation of the corrosion phenomena that occur. The important consequences of corrosion for implanted medical devices are not only degradation of the material properties but also, and often more importantly, the biological response to the degradation products. This article will consider the corrosion of metallic biomaterials under the following three headings:

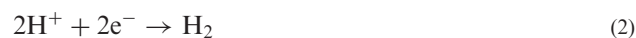
1. The electrochemical reactions
2. Surface analysis of the metal
3. Chemical analysis of the electrolyte

THE ELECTROCHEMICAL REACTIONS

The electrochemical reactions relevant to implantable medical devices take place on a metal surface in an aqueous electrolyte. There are always two reactions that occur: the anodic or oxidation reaction, which yields metallic ions resulting from the oxidation of the metal to its salt,



and the cathodic or reduction reaction, in which the electrons generated are consumed. The precise cathodic reaction will depend on the nature of the electrolyte, but two of the most important in aqueous environments are the reduction of hydrogen,



and the reduction of dissolved oxygen,



in acidic solutions or



in neutral or basic solutions.

In all corrosion processes, the rate of the anodic or oxidation reaction must equal the rate of the cathodic or reduction reaction. This is a basic principle of electrochemically based metallic corrosion. It also explains how variations in the local environment can affect the overall rate of corrosion by influencing either the anodic or cathodic reactions. The whole corrosion process can be controlled by altering either of these reactions.

When a metal is placed in an electrolyte, the electrochemical reactions will occur until an equilibrium is set up across the interface such that the rate of the anodic reaction equals the rate of the cathodic reaction. At this point, the potential across the interface is related to the corrosion current density by the Nernst equation:

$$E = E_0 + (RT/nF) \ln(a_{\text{anod}}/a_{\text{cath}}) \quad (5)$$

where E_0 is the standard electrode potential, RT/F is a constant, n is the number of electrons transferred, and a is the activity of the anodic and cathodic reactants. E_0 is defined as the electrode potential measured against a standard hydrogen electrode in a 1 N solution of its salt at 25°C. At low concentrations, the activity can be approximated to the concentration. In this situation, there is a net dissolution of the metal and a current will flow. At equilibrium, the rate of the metal dissolution is equal to the rate of the cathodic reaction, and the rate of the reaction is directly proportional to the current density by Faraday's law; therefore

$$i_{\text{anodic}} = i_{\text{cathodic}} = i_{\text{corrosion}} \quad (6)$$

and the Nernst equation can be rewritten

$$E - E_0 = \pm \beta \ln(i_{\text{corr}}/i_0) \quad (7)$$

where β is a constant and i_0 is the exchange current density, which is defined as the anodic (or cathodic) current density at the standard electrode potential.

The relationship between electrode potential and the corrosion current can be measured experimentally. The

data are useful in evaluation of the corrosion behavior of a particular metal/electrolyte system.

For an actively corroding metal, if the potential difference relative to a reference electrode is driven by some external process, then the current density will increase exponentially and a plot of the potential against the \ln current is a straight line. Most implantable structural metals and alloys will form a passivating oxide layer on their surface rapidly on exposure to the atmosphere such that when placed in an aqueous environment, active corrosion is inhibited. These materials result in a typical S-shaped polarization curve as shown in Fig. 1. Important information about the corrosion behavior of a metal or alloy can be obtained from these curves. The position of the equilibrium or corrosion potential (E_c) will lie within the passive region of this curve. It is important to notice that in this region there is still a finite current density and therefore a release of metal ions into the environment. As the potential is raised, the breakdown potential (E_b) will be reached, at which point the passive layer will begin to break down allowing active corrosion to occur and thus the current density begins to increase. If the potential is cycled, a potential will be reached at which repassivation occurs. The differences between the equilibrium potential, the breakdown potential, and the repassivation potential provide useful information to characterize the corrosion behavior of a particular metal/electrolyte system.

When a metallic device is implanted, there are also various practical considerations that can modify the basic corrosion of the metal in the electrolyte. These include geometrical considerations (crevice corrosion), relative movement of components (fretting corrosion), incidences where metal types are combined (galvanic corrosion), or situations where the implant is stressed (stress corrosion cracking).

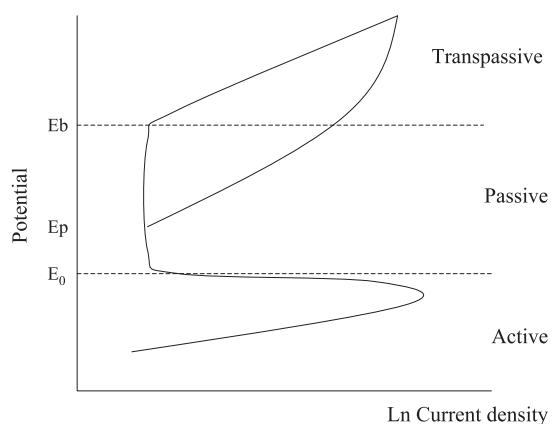


Fig. 1 Typical S-shaped polarization curve of a passivating metal or alloy.

Crevice Corrosion

Crevice corrosion can occur due to the geometry of the device. The area between the head of a bone screw and countersink on the fracture fixation plate can be influenced by the crevice conditions that the geometry creates.^[4] Porous coated implants may also demonstrate crevice corrosion.^[5] The crevice geometry restricts diffusion of oxygen into the crevice. Initially, the anodic and cathodic reactions occur uniformly over the surface, including within the crevice. As the crevice becomes depleted of oxygen, the reaction is limited to metal oxidation balanced by the cathodic reaction on the remainder of the surface, resulting in a small actively corroding area which will not repassivate. In an aqueous sodium chloride solution, the buildup of metal ions within the crevice causes the influx of chloride ions to balance the charge by forming the metal chloride. In the presence of water, the chloride will dissociate to its insoluble hydroxide and acid. This is a rapidly accelerating process since the decrease in pH causes further metal oxidation within the crevice.

Fretting Corrosion

Fretting corrosion results from the removal of the passive layer by a mechanical process.^[6,7] This can be a scratch that does not repassivate, resulting in the formation of a pit, or a continuous cyclic process in which any reformed passive layer is removed. It is suggested that this can contribute to the corrosion observed between a fracture fixation plate and the bone screws attaching the plate to the bone. Fretting can also affect the corrosion rate due to plastic deformation of the contact area; this can subject the area to high strain fatigue and may cause fatigue corrosion. Furthermore, stirring of the electrolyte can increase the limited current density of the cathodic reaction by upsetting the charge equilibrium that is established at the metal/electrolyte interface.

Galvanic Corrosion

If two metals are independently placed within the same solution, each will establish its own electrode potential with respect to the solution. If these two metals are placed in electrical contact, then a potential difference will be established between them, and electrons will pass from the more anodic to the more cathodic metal. Thus equilibrium is upset and a continuous process of dissolution from the more anodic metal will take place. This accelerated corrosion process is galvanic corrosion. It is important if two different alloys are used in an implantable device when the more reactive may corrode freely.



Whenever stainless steel is joined to another implant alloy, it will suffer from galvanic corrosion. If both alloys remain within their passive region when coupled in this way, the additional corrosion may be minimal. Some modular orthopedic systems are made of titanium alloys and cobalt-based alloys on the basis that both should remain passive, but evidence of corrosion has been reported.^[8] Galvanic corrosion may also take place on a microscopic scale in multiphase alloys where phases are of considerably different electronegativity. In dentistry, some amalgams may show extensive corrosion because of this mechanism.

Crevice, fretting, and galvanic corrosion could all occur on modular hip prostheses,^[9] where fretting initiates the breakdown of the passive layer, and once initiated, the crevice conditions cause continued corrosion without mechanical loads. This was shown to occur on both Co-Cr-Mo head/Co-Cr-Mo neck and Co-Cr-Mo head/Ti-6Al-4V neck prostheses.

Stress Corrosion Cracking

Stress corrosion cracking is a form of corrosion where an applied stress and a corrosive environment work together and cause complete failure of a component, when neither the stress nor the environment would be a problem on their own. The stress level may be very low, possibly only residual stress, and the corrosion may be initiated at a microscopic crack tip that does not repassivate rapidly. Incremental crack growth may then occur, resulting in fracture of the implant. Industrial uses of stainless steels in saline environments have shown susceptibility to stress corrosion cracking, and therefore it is a potential source of failure for implanted devices.

Electrochemical techniques can be used to build up an understanding of the corrosion behavior of a metal/electrolyte system. Measurement of the equilibrium corrosion potential, relative to a standard electrode, is used to assess the corrosion behavior. If the equilibrium electrode potential reaches a stable value rapidly, it is concluded that the metal forms a stable and protective passive layer. In some incidences, the equilibrium electrode potential fluctuates with time, leading to the conclusion that the passive layer is continually breaking down and reforming. Conditions that result in a lowering of the equilibrium electrode potential suggest reduced likelihood of device corrosion. Changes in the electrolyte composition—for example, the addition of H₂O₂ to PBS to mimic the in-vivo inflammatory response—caused the equilibrium electrode potential of Ti to increase, suggesting a reduction in the corrosion resistance, although when the H₂O₂ was replaced with PBS alone, the equilibrium electrode potential returned to a value similar to that in PBS

alone.^[10] Changes in the surface properties of the metal—for example, surface treatment of 316L and Nitinol stents, causing the formation of an amorphous oxide—resulted in a lower and more stable equilibrium potential.^[11] In bacterial cultures, the equilibrium electrode potentials of Ni-Cr and precious metal alloys were shown to drop in response to the formation of a biofilm on the surface, thus influencing the electrolyte interaction with the surface.^[12]

Small polarization of the electrode on either side of the equilibrium electrode potential (e.g., ± 20 mV) gives rise to the polarization resistance. This is a measure of the corrosion rate of the material under the defined conditions. This is a nondestructive technique and can be used to compare different materials under similar conditions, demonstrating the greater corrosion resistance of certain metals or the enhanced corrosion resistance of a particular surface modification or environment.^[12,13]

Anodic polarization, in which the metal is polarized over a wide potential range ($\pm 1-2$ V), is often used to compare materials. In particular, the position of the breakdown potential and the magnitude of the passive current are used to compare the corrosion resistance of different alloys, different surface modifications, or the influence of different electrolyte compositions. Three different titanium alloys—Ti6Al7Nb, Ti13Nb13Zr, and Ti15Zr4Nb—were studied using measurement of corrosion current density versus time and anodic polarization, and both studies concluded that the Ti6Al7Nb was the most corrosion resistant and that all three alloys had a high pitting corrosion resistance.^[14,15] Electrochemical processes have been used to evaluate the effect of casting conditions and subsequent sandblasting or polishing processes on titanium and titanium alloys^[13] and demonstrated that all the titanium materials tested were highly corrosion resistant and that the surface treatments had more effect than the casting processes. Surface treatment to form an amorphous oxide on 316L stainless steel and NiTi stents was analyzed using anodic polarization. The original polycrystalline oxide on the 316L had a high susceptibility to pitting corrosion as demonstrated by a relatively low breakdown potential and a large hysteresis loop before repassivation. The treated surfaces were highly pitting corrosion resistant as demonstrated by a high breakdown potential and insignificant hysteresis.^[11] Various surface treatments of NiTi^[16] have also demonstrated an effect on corrosion resistance by evaluation of a reduction in passive current density and an increase in the breakdown potential for all treated surface compared to an untreated control.

The characteristics of the passive layer can also be evaluated using AC impedance measurements. This technique evaluates the resistance of the electrolyte at

the metal/solution interface and therefore can be used to follow the interaction of the components of the electrolyte—for example, proteins, cells, and bacteria—with the surface of the metal and their influence on the corrosion behavior. Using impedance spectroscopy,^[17] it was concluded that the addition of H₂O₂ caused a decrease in the corrosion resistance of Ti owing to its influence on the structure of the oxide. It was suggested that the H₂O₂ thickened the oxide and increased its porosity. Thermal oxidation of Ti6Al4V caused a change in the electrochemical impedance spectroscopy results but did not impair the osteoblast cell response to the material.^[18] AC impedance measurements for commercial pure Ti in culture medium with and without L929 murine fibroblasts demonstrated that a layer of biomolecules formed on the surface and acted as a capacitor, reducing the diffusion of molecules and therefore increasing corrosion resistance.^[19]

Fretting corrosion can be assessed by measuring the corrosion current and anodic polarization behavior.^[7] Comparison of different Ti alloys demonstrated the poor fretting behavior of the conventional Ti alloy Ti-6Al-4V and the much improved behavior of Ti-15Zr-4Nb-4Ta.

There are many studies in the literature that evaluate the electrochemical behavior of implant alloys using a range of electrolytes to model the in-vivo environment. It is important to consider the properties of the environment so that these studies are useful predictors of the in-vivo corrosion. The biological environment to which biomaterials are exposed during prolonged use (i.e., the internal milieu of the body) can be described as an aqueous medium containing various anions, cations, organic substances, and dissolved oxygen. The anions are mainly chloride, phosphate, and bicarbonate ions. The principal cations are Na⁺, K⁺, Ca²⁺, and Mg²⁺, but there are smaller amounts of many others. This represents an environment with a chloride concentration of approximately a third of that of sea water.^[20] The concentration of dissolved oxygen also influences the aggressive nature of the environment and in venous blood is approximately a quarter of that in air. The organic substances include low-molecular-weight species as well as relatively high-molecular-weight proteins and lipids. The protein content of the environment is known to have a significant influence on the corrosive nature of body fluids.^[21,22] The pH in this well-buffered system is around 7.4, although it may change for short periods following surgery due to inflammation to as low as 4 or 5.^[2] The temperature remains constant around 37°C. When metals are placed in this environment, the electrochemical reactions result in the dissolution of metal ions into the electrolyte. Once in solution, the metal ions can interact with components

of the electrolyte, and this can influence the corrosion process and the biological response to the corrosion products.

There is sufficient evidence to support the premise that the presence of proteins, cells, and bacteria can influence the rate of corrosion of some metals.^[6,12,20–23] The stability of the oxide layer depends on the electrode potential and the pH of the solution. Proteins and cells can be electrically active and interact with the charges formed at the interface and thus affect the electrode potential.^[2] Bacteria^[12] and inflammatory cells^[10,20] can alter the pH of the local environment through the generation of acidic metabolic products which can shift the equilibrium. The stability of the oxide layer is also dependent on the availability of oxygen. The adsorption of proteins and cells onto the surface of materials could limit the diffusion of oxygen to certain regions of the surface. This could cause preferential corrosion of the oxygen-deficient regions and lead to the breakdown of the passive layer. Alternatively, the biomolecule adsorption layer could act as a capacitor preventing the diffusion of molecules from the surface.^[19]

The biological molecules could upset the equilibrium of the corrosion reactions by consuming one or more of the products of the anodic or cathodic reaction. For example, proteins can bind to metal ions and transport them away from the implant surface. This will upset the equilibrium across the charged double layer and allow further dissolution of the metal. Clearly if either the accumulating positive metal ions in the surrounding media or the accumulating electrons in the metal are removed, the balance between the dissolution and the replacement of the ions will be disturbed. This will occur in the biological environment surrounding implanted alloys due to the interaction of the proteins with the metal ions. Metal ions can form complexes with proteins,^[3,24,25] and these complexes can be transported away from the immediate vicinity. This removes the metal ions from the charged double layer at the interface allowing further release of metal ions to reestablish the equilibrium. Similarly, relative movement between the implant and the tissue—for example, at a bearing surface or on a cyclically loaded implant—will cause mixing at the interface and will modify the composition of the electrolyte and may modify the surface of the alloy.^[6] The equilibrium is established precisely because of the imbalance of charge, so that if the charge balance is disturbed, further corrosion will occur to attempt to reestablish the balance. The result will be continued dissolution as the system attempts to achieve this equilibrium. An environment that allows the removal of electrons in contact with the metal or stirring of the electrolyte will achieve this.



SURFACE ANALYSIS OF THE METAL

The effect of corrosion on the metal occurs at the surface. It is possible to investigate the corrosion behavior of a metal/electrolyte system by evaluating changes to the surface properties of the metal. This includes changes in the composition of the surface both in terms of a depth profile and as a function of time. Many of the commonly used surgical alloys contain highly reactive metals such as titanium, aluminium, and chromium. Because of this high reactivity, they will react with oxygen upon initial exposure to the atmosphere. This initial oxidation leaves an impervious oxide layer firmly adherent to the metal surface; thus all other forms of corrosion may be significantly reduced because the oxide layer acts as a protective barrier. The manufacturing process for implant alloys may include a passivating step to enhance the oxide layer prior to implantation, for example, nitric acid treatment of 316L stainless steel.^[26]

Of the three elements that are strongly passivated (i.e., aluminium, chromium, and titanium), only Ti is used on its own. Chromium is widely used in alloys, especially in stainless steels and in the cobalt-chromium-based alloys, where it is normally considered that a level of above 12% gives good corrosion resistance and about 18% provides excellent resistance. Aluminium has an important role in several Ti alloys, whereas titanium is used as a pure metal or as the major constituent of alloys.^[27] In alloys, the passivating layer promoting the corrosion resistance is predominantly composed of one of these metal oxides. For example, chromium oxide passivates 316L stainless steel and Co-Cr-based alloys and Ti oxide in Ti alloys. The other alloying elements may be present in the surface oxide, and this can influence the passivity of the layer.^[28] Careful pretreatment of the alloys can be used to control the passivity of these alloys by controlling the structure and chemistry of the oxide layer.^[11,16] In particular, production procedures need to be controlled due to their influence on the surface oxides, for example, the cleaning^[29] and sterilization^[30] procedures.

The stainless steels used in implantable devices are passivated by the chromium oxide that forms on the surface. It has been shown, however, that in a physiological saline environment, the driving force for repassivation of the surface is not high.^[5] Thus, if the passive layer is broken down, it may not repassivate and active corrosion can occur.

Localized corrosion can occur as a result of imperfections in the oxide layer, producing small areas in which the protective surface is removed.^[31] These localized spots will actively corrode and pits will form in the surface of the material. This can result in a large degree of localized damage because the small areas of

active corrosion become the anode and the entire remaining surface becomes the cathode. Since the rate of the anodic and cathodic reactions must be equal, it follows that a relatively large amount of metal dissolution will be initiated by a small area of the surface, and large pits may form.

If some areas of the alloy become depleted in chromium, as can happen if carbides are formed at the grain boundaries, the regions adjacent to the grain boundaries become depleted in chromium. The passivity of the surface in these regions is therefore affected and preferential corrosion can occur. Although this problem can easily be overcome by heat-treating the alloys,^[32] it has been observed on retrieved implants^[33] and can cause severe problems since, once initiated, it will proceed rapidly and may cause fracture of the implant and the release of large quantities of corrosion products into the tissue.

Characterization of the oxide can provide useful information about the corrosion processes that have occurred. As seen in Fig. 2, components of the electrolyte will interact with the surface of the oxide. Metal ions can diffuse through the oxide and be released into the environment, and dissolution of the oxide can occur, also releasing metal ions.^[34] Oxygen can diffuse through the oxide and react with the bulk metal, increasing the thickness of the oxide. Hydroxides can interact with the surface oxide, producing a metal hydroxide surface layer.^[35] Calcium and phosphate ions can become incorporated into the surface oxide. In in-vitro electrolytes containing Ca and phosphate ions, it has been shown that repassivation results in the incorporation of titanium phosphate into and the adsorption of calcium phosphate or calcium titanium phosphate onto the surface of Ti.^[23,36] Macromolecules (for example, proteins and lipoproteins) can adsorb onto the surface.

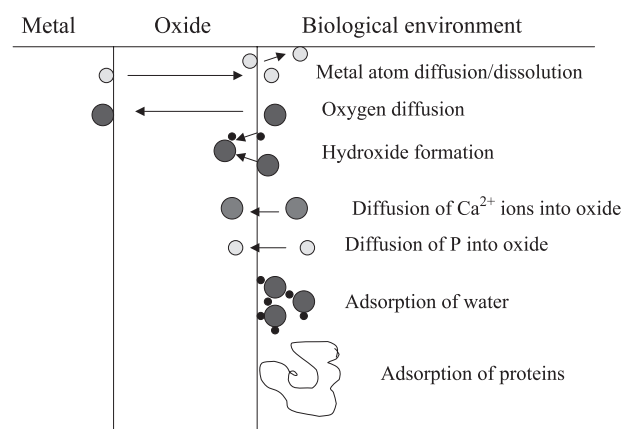


Fig. 2 Schematic representation of the metal/oxide/biological environment interface.

Superimposed on these reactions are the particular conditions under which the metal is being used—for example, the mechanical environment, the geometry of the device, any relative movement between components, and any compositional variations. There is a dynamic process occurring at this interface, and therefore changes in the surface oxide can have longer-term implications in terms of the performance of a device.

The principal techniques available to study the surface chemical composition are X-ray photoelectron spectroscopy (XPS)^[14,36] and auger electron spectroscopy (AES).^[36–38] They provide information on the elements present in the surface and how the composition changes with depth into the oxide. XPS can analyze the bonding environment of the elements present, and AES can give spatial resolution of the elements.

There have been many studies that have investigated the surface properties of implant alloys and how these change on implantation.^[34] It is well known that the composition of the oxide is not representative of the bulk chemistry of the alloy. For commercially pure Ti, the natural occurring oxide was found to be composed of TiO₂, Ti₂O₃, and TiO.^[28] Ti alloys were shown to have additional alloy component oxides (for example, Al₂O₃, Nb₂O₅, and V-oxides), and the concentration of these oxides in the surface layers was dependent on the pretreatment of the materials. Spontaneously formed oxides on Ti-7Nb-6Al were shown to have large surface enrichment of Al oxide, and Ti-13Nb-13Zr and Ti-15Zr-4Nb were found to have significant surface enrichment of Zr oxide and some Nb oxide but to a lesser extent.^[14] Surface treatments of Ni-Ti stents were shown to influence the oxide composition and thickness,^[16] and this in turn influenced the corrosion resistance. It was concluded that a thin uniform oxide provided the most corrosion-resistant surface. The composition of the oxide and its solubility can be particularly important if there is concern about the biological toxicity of some of the alloy components, for example, Ni in Ni-Ti alloys or Al and V in Ti-6Al-4V.

These analyses can also be used to evaluate the relationship of the corrosion of the metal with time. This can be performed on model in-vitro experiments as well as on retrieval studies, either in animal studies or following clinical implantation.^[34]

CHEMICAL ANALYSIS OF THE ELECTROLYTE

Corrosion results in the release of metal ions into the environment. It is important^[3] to investigate the amount of metal ion released, the chemical form of the metal ion, where and how much is transported away

from the local site, and thus the pathological consequences of the corrosion process.

In an in-vitro closed environment, the quantification of the metal ion concentration in the electrolyte is a measure of the amount of corrosion that has occurred. This can be measured using atomic absorption spectroscopy (AAS) or inductively coupled plasma mass spectroscopy (ICP-MS).^[2] Another important consideration is the valence of the metal ions that have been released; neither AAS or ICPMS will provide this information, but polarography can be used to gain insight into this problem.

In vivo, however, the metal ions can bind to proteins and be transported away from the immediate vicinity of the implant. This has various implications. Since the metal ions are removed from the electrolyte surrounding the metal, the equilibrium will be disrupted and thus allow continual replacement of the metal ions from the surface, i.e., continual corrosion. There is a large body of literature on the release and distribution of metal ions from orthopedic devices as reviewed by Jacobs et al.^[3]

In in-vitro studies, metal ion release can be measured under controlled conditions, and therefore the effect of changing the conditions can be followed closely. For example, the influence of fretting,^[20] the composition of the electrolyte,^[2,39] and the surface treatment of the alloy^[40] can be investigated. These studies also demonstrate how comparison of changes in the surface properties of the oxide layer and ion release into the electrolyte can help to characterize the corrosion behavior of a particular implant alloy/environment system.

In in-vivo studies, total metal ion release is more difficult to obtain, but the distribution of metal ions systemically can be assessed by collecting tissue and organs from around the body (either in animal studies^[41] or at autopsy^[3]) and analyzing for metal ion content using AAS or ICPMS. In-vivo release of metal ions can also be correlated with the local or systemic tissue response using histological methods.^[2,3]

The collection of samples of body fluids (for example, serum and urine from patients following joint replacement surgery) and the analysis of metal ions in these samples have demonstrated significant increases in the concentration of various metal ions, and in some studies the increase correlates with implant loosening.^[3] Different implant systems and the influence of different mechanical environments, for example, can be investigated and compared with clinical outcome.

Studies on the systemic reaction to implant corrosion have concentrated on hypersensitivity reactions, although potential carcinogenicity has also been considered.^[2,3] In these studies, the valence of the metal ion is important. A number of studies have investigated the valence state of chromium released from



implants because the biological activity of Cr^{3+} and Cr^{6+} differ. To study this, polarographic techniques are used and have shown that in Ringer's solution and in blood in-vitro 316L stainless steel releases Cr^{6+} rather than Cr^{3+} .

CONCLUDING REMARKS

One of the important criteria to be taken into account when choosing structural implant alloys is their corrosion resistance. The corrosion resistance of most of these materials is due to a naturally occurring passive oxide layer. Even under these conditions, metal ions are released into the surrounding environment on implantation. Any process that modifies, either intentionally or not, the surface oxide may influence the corrosion resistance of the material. The amount of metal ion release may be so low that it is inconsequential to the structural integrity of the material; however, the biological response to the metal ions may lead to severe adverse biological responses either locally or systemically.

ARTICLES OF FURTHER INTEREST

Biocompatibility Testing; Bone-Implant Interface; Host Reactions; Implant, Total Hip

REFERENCES

- Williams, D.F.; Williams, R.L. Degradative Effects of the Biological Environment on Metals and Ceramics. In *Biomaterials Science: An Introduction to Materials in Medicine*; Ratner, B., Hoffman, A., Lemons, J., Schoen, F., Eds.; Academic Press: San Diego, 1996; 260–267.
- Bundy, K.J. Corrosion and other electrochemical aspects of biomaterials. *Crit. Rev. Biomed. Eng.* **1994**, *22* (3/4), 139–251.
- Jacobs, J.J.; Gilbert, J.L.; Urban, R.M. Current concepts review corrosion of metal orthopaedic implants. *J. Bone Jt. Surg.* **1998**, *80-A*, 268–282.
- Cook, S.D.; Tomas, K.A.; Harding, A.F.; Collins, C.L.; Haddad, R.J.; Milicic, M.; Fischer, W.L. The in vivo performance of 250 internal fixation devices; a follow up study. *Biomaterials* **1987**, *8*, 177–184.
- Seah, K.H.W.; Thampuran, R.; Teoh, S.H. The influence of pore morphology on corrosion. *Corr. Sci.* **1998**, *40*, 547–556.
- Khan, M.A.; Williams, R.L.; Williams, D.F. Conjoint corrosion and wear in titanium alloys. *Biomaterials* **1999**, *20* (8), 765–772.
- Okazaki, Y. Effect of friction on anodic polarization properties of metallic biomaterials. *Biomaterials* **2002**, *23*, 2071–2077.
- Gilbert, J.L.; Buckley, C.A.; Jacobs, J.J. In vivo corrosion of modular hip prosthesis components in mixed and similar metal combinations. The effect of crevice, stress, motion and alloy coupling. *J. Biomed. Mater. Res.* **1993**, *27*, 1533–1544.
- Goldberg, J.R.; Gilbert, J.L. In vitro corrosion testing of modular hip tapers. *J. Biomed. Mater. Res. (Appl. Biomater.)* **2003**, *64B*, 73–78.
- Fonseca, C.; Barbosa, M.A. Corrosion behaviour of titanium in biofluids containing H_2O_2 studied by electrochemical impedance spectroscopy. *Corr. Sci.* **2001**, *43*, 547–559.
- Shih, C.-C.; Lin, S.-J.; Chung, K.-H.; Chen, Y.-L.; Su, Y.-Y. Increased corrosion resistance of stent materials by converting current surface film of polycrystalline oxide into amorphous oxide. *J. Biomed. Mater. Res.* **2000**, *52*, 323–332.
- Laurent, F.; Grosgeat, B.; Reclaru, L.; Dalard, F.; Lissac, M. Comparison of corrosion behaviour in presence of oral bacteria. *Biomaterials* **2001**, *22*, 2273–2282.
- Cai, Z.; Shafer, T.; Watanabe, I.; Nunn, M.E.; Okabe, T. Electrochemical characterization of cast titanium alloys. *Biomaterials* **2003**, *24*, 213–218.
- López, M.F.; Gutiérrez, A.; Jiménez, J.A. In vitro corrosion behaviour of titanium alloys without vanadium. *Electrochim. Acta* **2002**, *47*, 1359–1364.
- López, M.F.; Jiménez, J.A.; Gutiérrez, A. Corrosion study of surface-modified vanadium-free titanium alloys. *Electrochim. Acta* **2003**, *48*, 1395–1401.
- Trepanier, C.; Tabrizian, M.; Yahia, L'H.; Bilodeau, L.; Piron, D. Effect of modification of oxide layer on NiTi stent corrosion resistance. *J. Biomed. Mater. Res.* **1998**, *43*, 433–440.
- Pan, J.; Thierry, D.; Leygraf, C. Electrochemical impedance spectroscopy study of the passive oxide film on titanium for implant applications. *Electrochim. Acta* **1996**, *41*, 1143–1153.
- García-Alonso, M.C.; Saldaña, L.; Vallés, G.; González-Carrasco, J.L.; González-Cabrero, J.; Martínez, M.E.; Gil-Garay, E.; Munuera, L. In vitro corrosion behaviour and osteoblast response of thermally oxidized Ti6Al4V alloy. *Biomaterials* **2003**, *24*, 19–26.
- Hiroto, S.; Noda, K.; Hanawa, T. Development of electrolytic cell with cell-culture for metallic biomaterials. *Corr. Sci.* **2002**, *44*, 955–965.
- Hanawa, T. Evaluation techniques of metallic biomaterials in vitro. *Sci. Technol. Adv. Mater.* **2002**, *3*, 289–295.
- Khan, M.A.; Williams, R.L.; Williams, D.F. The corrosion behaviour of Ti-6Al-4V, Ti-6Al-7Nb and Ti-13Nb-13Zr in protein solutions. *Biomaterials* **1999**, *20* (7), 631–637.
- Williams, D.F. Physiological and microbiological corrosion. *Crit. Rev. Biocompat.* **1985**, *1* (1), 1–24.
- Hanawa, T. In vivo metallic biomaterials and surface modification. *Mater. Sci. Eng.* **1999**, *A267*, 260–266.
- Steinemann, S.G. Metal implants and surface reactions. *Injury* **1996**, *27* (S-3), 16–22.
- Büdingner, L.; Hertl, M. Immunological mechanisms in hypersensitivity reactions to metal ions: An overview. *Allergy* **2000**, *55*, 108–115.

26. *Corrosion and Degradation of Implant Materials*; Fraker, A.C., Griffith, C.D., Eds.; ASTM S.T.P. No. 859, American Society for Testing and Materials: Philadelphia, 1985.
27. Long, M.; Rack, H.J. Titanium alloys in total joint replacement—A materials science perspective. *Biomaterials* **1998**, *19*, 1621–1639.
28. Sittig, C.; Textor, M.; Spencer, N.D.; Wieland, M.; Vallotton, P.-H. Surface characterization of implant materials c.p.Ti, Ti-6Al-7Nb and Ti-6Al-4V with different pretreatments. *J. Mater. Sci., Mater. Med.* **1999**, *10*, 5–46.
29. Aronsson, B.-O.; Lausmaa, J.; Kasemo, B. Glow discharge plasma treatment for surface cleaning and modification of metallic biomaterials. *J. Biomed. Mater. Res.* **1997**, *35*, 49–73.
30. Thierry, B.; Tabrizian, M.; Savadogo, O.; Yahia, L'H. Effects of sterilization processes on NiTi alloy: Surface characterisation. *J. Biomed. Mater. Res.* **2000**, *49*, 88–98.
31. Rondelli, G.; Vicentini, B. Localized corrosion behaviour in simulated human body fluids of commercial Ni-Ti orthodontic wires. *Biomaterials* **1999**, *20*, 785–792.
32. Disegi, J.A.; Eschbach, L. Stainless steel in bone surgery. *Injury* **2000**, *31* (Suppl. 4), 2–6.
33. Walczak, J.; Shahgaldi, F.; Heatley, F. In vivo corrosion of 316L stainless-steel hip implants: Morphology and elemental compositions of corrosion products. *Biomaterials* **1998**, *19*, 229–237.
34. Textor, M.; Sittig, C.; Frauchiger, V.; Tosatti, S.; Brunette, D.M. Properties and Biological Significance of Natural Oxide Films on Titanium and Its Alloys. In *Titanium in Medicine*; Brunette, D.M., Tengvall, P., Textor, M., Thomsen, P., Eds.; Springer: Berlin, 2001; 171–230.
35. Tengvall, P.; Lundström, I.; Sjökvist, L.; Elwing, H.; Bjursten, L. Titanium-hydrogen peroxide interactions: Model studies of the influence of the inflammatory response on titanium implants. *Biomaterials* **1989**, *10*, 166–175.
36. Hanawa, T.; Asami, K.; Asaoka, K. Repassivation of titanium and surface oxide film regeneration in simulated bioliquid. *J. Biomed. Mater. Res.* **1998**, *40*, 530–538.
37. Ong, J.L.; Lucas, L.C. Auger electron spectroscopy and its use for the characterization of titanium and hydroxyapatite surfaces. *Biomaterials* **1998**, *19*, 455–464.
38. Esposito, M.; Lausmaa, J.; Hirsch, J.-M.; Thomsen, P. Surface analysis of failed oral titanium implants. *J. Biomed. Mater. Res. (Appl. Biomater.)* **1999**, *48*, 559–568.
39. Ryhänen, J.; Niemi, E.; Serlo, W.; Niemelä, E.; Sandvik, P.; Pernu, H.; Salo, T. Biocompatibility of nickel-titanium shape memory metal and its corrosion behaviour in human cell cultures. *J. Biomed. Mater. Res.* **1997**, *35*, 451–457.
40. Thierry, B.; Tabrizian, M.; Trepanier, C.; Savadogo, O.; Yahia, L'H. Effect of surface treatment and sterilization processes on the corrosion behaviour of NiTi shape memory alloy. *J. Biomed. Mater. Res.* **2000**, *51*, 685–693.
41. Ryhänen, J.; Kallioninen, M.; Serlo, W.; Perämäki, P.; Jumila, J.; Sandvik, P.; Niemelä, E.; Tuukkanen, J. Bone healing and mineralization, implant corrosion, and trace metals after nickel-titanium shape memory metal intramedullary fixation. *J. Biomed. Mater. Res.* **1999**, *47*, 472–480.

Cryopreservation of Living Cells

Dayong Gao

Department of Mechanical Engineering and Center for Biomedical Engineering, University of Kentucky, Lexington, Kentucky, U.S.A.

J. K. Critser

Department of Comparative Medicine Center, University of Missouri, Columbia, Missouri, U.S.A.

INTRODUCTION

Biological metabolism in living cells dramatically diminishes and eventually stops at low temperatures, a fact that permits the long-term cryopreservation (i.e., low-temperature preservation) of the living cells and tissues. With recent and rapid advances made in the clinical diagnosis and treatment of diseases, there is an increasing requirement for using cell and organ transplantation to cure acquired diseases and to correct genetic defects. Coupled with this, more and more types of living cells, tissues, organs, and engineered tissues are urgently required to be cryopreserved and banked at low temperatures. Cryopreservation is not only important for modern medicine/healthcare but also for many other areas, because it is indispensable for: 1) the banking of a large quantity of living cells/tissues for HLA typing and matching between the recipients and donors to meet the greatly increased clinical needs and, sometimes, the urgent needs (e.g., in a war or in events such as the “9-11” terrorist attack, natural disasters, or other accidents); 2) facilitating the transport of cells/tissues between different medical centers; and 3) allowing sufficient time for the screening of transmissible diseases (e.g., HIV) in donated cells/tissues before transplantation. In addition, 4) any engineered tissues (from tissue engineering research) need to be successfully cryopreserved before their practical use in any applications and commercialization; and 5) the cryopreservation of sperm and oocytes/eggs of endangered or transgenic species is urgently needed.

However, there is an apparent contradiction between the concept of cryopreservation and experimental findings that cells/tissues can be killed by the cryopreservation process itself. In fact, the first successful cryopreservation of living mammalian cells became reality in 1949 when Polge et al.^[1] found glycerol as a cryoprotective agent (CPA), reducing cell's injury during cryopreservation. Since then, a number of glycerol-like CPAs have been discovered and used in cryopreservation. The major steps in the cryopreservation process include: 1) adding cryoprotective

agents (CPAs) to cells/tissues before cooling, 2) cooling the cells/tissues to a low temperature (e.g., -196°C , liquid nitrogen temperature at pressure of 1 atm) at which the cells/tissues are stored, 3) warming the cells/tissues, and 4) removing the CPAs from the cells/tissues after thawing. Injury to cells has been shown to be caused by each or a combination of the above steps. Cryobiology is a multidisciplinary science to study the physical and biological behaviors of cells and tissues (including their interactions with environment) at low temperatures (especially at temperatures below the freezing point of water, 0°C). A very important part of the research in fundamental cryobiology is to reveal the underlying physical and biological mechanisms related to the injury of cells at low temperatures (so-called cryoinjury), especially associated with the phase change of water in both extra- and intracellular environments. With the understanding of the mechanisms of cell cryoinjury, one may establish biophysical/mathematical models describing cell responses to environmental change during the cryopreservation process and use these models to develop the optimal procedures/devices for the long-term cryopreservation of living cells, preventing cryoinjury.

CRYOINJURY MECHANISMS DURING FREEZING PROCESSES

Contrary to popular belief, the challenge to cells during cryopreservation is not their ability to endure storage at low temperature; rather, it is the lethality of an intermediate zone of temperature (-15 to -60°C) that cells must traverse twice: once during cooling and once during warming.^[2] As cells are cooled to about -5°C , both the cells and surrounding medium remain unfrozen and supercooled. Between -5 and about -15°C , ice forms (either spontaneously or as a result of artificially introducing ice crystals to the solution) in the external medium but the cells' contents remain unfrozen and supercooled, presumably because the plasma membrane blocks the growth of ice crystals into the cytoplasm. The supercooled water in the cells

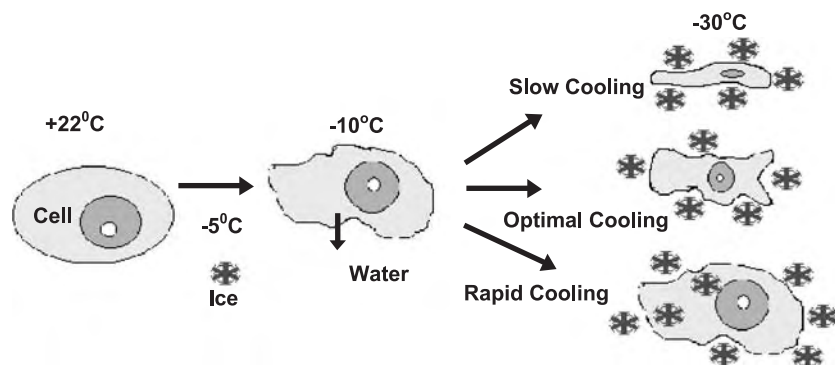


Fig. 1 Schematic of physical events in cells during freezing.

has, by definition, a higher chemical potential than that of water in the partially frozen extracellular solution, and thus water flows out of the cells osmotically and freezes externally (Fig. 1). The subsequent physical events in the cells depend on the cooling rate as shown in Fig. 1. If cells are cooled too rapidly, water is not lost fast enough to maintain equilibrium; the cells become increasingly supercooled, eventually attaining equilibrium by freezing intracellularly.^[3,4] Muldrew and McGann invoked an osmotic rupture hypothesis and model of intracellular freezing injury based on solid experimental data and theoretical derivations,^[4] indicating that the water flux across the cell membrane during the freezing/thawing can induce intracellular ice formation causing injury. Diller and Cravalho developed novel microscopy to observe the cell behavior as well as intracellular ice formation (IIF) temperatures and IIF dynamics during the freezing processes.^[5,6] Recently, based on heterogeneous nucleation theory and thermodynamics, as well as experimental results, Toner et al. developed an IIF model and equations describing the probability of IIF as a function of cooling rate, temperature, and cryobiological characteristics of a specific cell type.^[7,8] The freezing behavior of the cells can be modified by the addition of cryoprotective agents (CPAs), which affect the rates of water transport, nucleation, and crystal growth. Karlsson and Toner further developed a model of diffusion-limited ice growth inside cells in the presence of CPAs.^[9] They successfully predicted the IIF phenomena as a function of cooling rate, temperature, and CPA concentration, and also predicted optimum cooling protocols preventing IIF. In most cases, cells undergoing intracellular ice formation (IIF) are killed.^[4,7,8,10–19]

If cooling is sufficiently slow, the cells will lose water rapidly enough to concentrate the intracellular solutes sufficiently to eliminate supercooling. As a result, the cells dehydrate and do not freeze intracellularly. However, if the cells are cooled too slowly, they will experience a severe volume shrinkage and long-time exposure to high solute (mainly electrolyte)

concentrations before eutectic temperature (the temperature at which all components in a solution are solidified) is achieved, both of which may cause cell injury (Fig. 1). Lovelock proposed that the increased concentration of solutes and cell dehydration have deleterious effects on the lipid-protein complexes of cell membranes, weakening them and increasing lipid and phospholipid losses.^[20] The cell is rendered permeable to electrolytes and swells, eventually rupturing. Levitt proposed that loss of water from the protoplasm brings protein molecules into apposition, presenting the opportunity for the formation of new chemical bonds previously too distant and rigidly structured in hydrated form to permit combination.^[21] Thawing would have a disruptive force on the new combination, permitting unfolding and denaturation. Karow and Webb explained slow freezing injury as a consequence of the extraction of bound water from cellular structures for incorporation into ice crystals, denuding proteins of lattice-arranged bound water essential to cell integrity.^[22] The possible effects that may cause cell injury as a result of concentration of solutes have been characterized collectively by Mazur et al. as “solution effects.”^[23] Mazur et al. suggested that solution effects on cells are greatly enhanced in a slow cooling process during which the time of the exposure of cells to a highly concentrated solution is prolonged. They also indicated that hyperosmotic stress may cause a net leak/influx of nonpermeating solutes (mainly, electrolytes); when cells are returned to isotonic conditions, they swell beyond their normal isotonic volume and lyse. Meryman proposed that the slow freezing results in the decrease in intracellular volume beyond a critical volume.^[24] As the cell reduces in size in response to increasing extracellular osmolarity, the compression of contents enhances the resistance to further shrinking. This results in a hydrostatic pressure difference across the cell membrane, incurring cell membrane damage. Steponkus and Wiest proposed a maximum cell surface area hypothesis: Cell shrinkage induces irreversible membrane fusion, and hence the effective area of cell membrane is reduced; when returned to isotonic



condition, the cells lyse before their normal volume is recovered.^[17] Some cells could be shown to exhibit a reversible inhibition of shrinkage at some point during a progressive increase in extracellular solute concentration, implying the development of membrane stress.^[25] Alternative theories have embraced such mechanisms of slow-freezing injury as macromolecule denaturation from dehydration or phase transition of membrane lipids.^[26] Pegg et al. proposed a packing effect, which is an inverse dependence of cell survival upon the proportion of the initial sample volume that is occupied by the cells.^[27] Mazur and colleagues, however, have proposed that a considerable portion of the damage, at least in human erythrocytes, is due to the reduction in the size of the unfrozen channels, which results in increased cell-ice or cell-cell contacts or interactions.^[28]

In general, either too high or too low a cooling rate can kill cells, although the mechanisms underlying cell damage are different. These cell responses to freezing were first expressed quantitatively by Mazur^[2] and directly linked with cell cryoinjury by Mazur's two factor hypothesis:^[23] 1) at slow cooling rates, the cryoinjury occurs due to the solution effects (i.e., the solute/electrolyte concentration, severe cell dehydration, and the reduction of unfrozen fraction in the extracellular space), and 2) at high cooling rates, cryoinjury occurs due to the lethal intracellular ice formation (IIF). The optimum cooling rate for cell survival should be low enough to avoid IIF but high enough to minimize the solution effects. Based upon this, an optimal cooling rate for cell cryosurvival should exist between the high and the low rates, as shown in Fig. 1. This has been confirmed experimentally. As shown in Fig. 2, cell survival plotted as a function of cooling rate produces a characteristic inverted U-shaped curve. There is an optimal cooling rate for a specific cell type. For a given cell type, low cooling rate (relative to the optimal cooling rate) results in

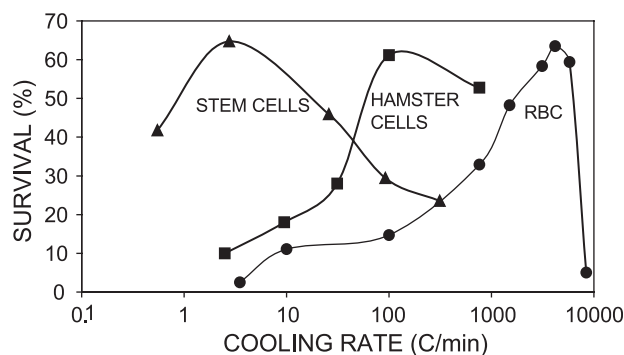


Fig. 2 Relationship (inverted U shape) between percent survival and cooling rate in three types of cells. Stem cells refer to mouse bone marrow stem cells, and hamster cells refer to hamster oocytes. (Redrawn from Ref. [13].)

solution-effects cryoinjury while high cooling rate (relative to the optimal cooling rate) results in lethal IIF.

It is clear from the above cryobiology theory that whether a given cooling rate is too high or too low for a given cell type depends on the ability of water to move across the cell membrane, i.e., water permeability of cell membrane, which is cell-type dependent. Differences in water permeability account largely for the magnitude of difference in optimal cooling rates for different cell types (Fig. 2). Diller et al.,^[6] Levin et al.,^[29] McGrath et al.,^[30,31] and Gao et al.^[32,33] developed devices/modeling to determine cell membrane permeabilities. Devireddy and Bischof et al. developed a novel differential scanning calorimeter (DSC) method to determine cell permeabilities during the freezing processes.^[34] The rate of water transport across cell membrane during the freezing process with different cooling rates can be described by four simultaneous equations.^[35] The first relates the loss of cytoplasmic water to the chemical potential gradient between intracellular supercooled water and external ice, expressed as a vapor pressure ratio,

$$\frac{dV}{dt} = \frac{L_p A R T}{v} \ln \frac{P_e}{P_i} \quad (1)$$

where V is the volume of cell water, t is time, L_p is the permeability coefficient for water (hydraulic conductivity), A is the cell surface area, R is the universal gas constant, T is temperature, v is the molar volume of water, and P_e and P_i are the vapor pressures of extracellular and intracellular water, respectively. The change in this vapor pressure ratio with temperature can be calculated from the Clausius-Clapeyron equation and Raoult's law:

$$\frac{d \ln(P_e/P_i)}{dT} = \frac{L_f}{RT^2} - \frac{N_v}{(V + N_v)V} \frac{dV}{dT} \quad (2)$$

Here, N is osmoles of intracellular solute and L_f is the latent heat of fusion of ice. Time and temperature are related by the cooling rate (B), which, if linear, is given by

$$\frac{dT}{dt} = B \quad (3)$$

Finally, L_p is related to temperature by the Arrhenius relation,

$$L_p = L_{p0} \exp \left[-\frac{E_a}{R} \left(\frac{1}{T} - \frac{1}{T_0} \right) \right] \quad (4)$$

where L_{p0} is the permeability coefficient of water at a known temperature, T_0 ; E_a is the activation energy of L_p ; and R is the universal gas constant.

To avoid intracellular freezing, the water content of the cell given by the above four equations must, before reaching the intracellular ice-nucleation temperature (usually between -5 and -40°C), have approached the equilibrium water content given by

$$V = \frac{vM_i}{\exp\left(\frac{L_f}{RT} - \frac{L_f}{273R}\right) - 1} \quad (5)$$

where M_i is the initial osmolality of the extracellular solution. The above quantitative expressions permit one to calculate the extent of supercooling in cells as a function of the cooling rate provided one knows or can estimate the permeability of the cell to water (L_p), its temperature coefficient or activation energy (E_a), the osmoles of solute initially in the cell, and the ratio of the cell surface area to volume. The results of such calculations are commonly expressed as plots of the water content of a cell as a function of temperature relative to the normal or isotonic water volume. Figure 3, for example, shows computed water loss curves for mouse ova. The calculated extent to which a cell becomes supercooled is the number of degrees that any given curve is displaced to the right of the equilibrium curve at a given subzero temperature. In the example shown, mouse ova cooled at $4^{\circ}\text{C}/\text{min}$ will be supercooled by 15°C as the temperature passes through -20°C . Three biological parameters have special influence on the position and shape of these curves, namely water permeability coefficient (L_p), its activation energy (E_a), and the size of the cell, or more properly, its surface-to-volume ratio (A/V). An increase in L_p produces the same effect as a comparable decrease in cooling rate. The effect of cell size is the opposite. An increase in size reduces the cooling rate required to produce a given probability of intracellular freezing. Changes in these two parameters shift the positions of the curves. Changes in the activation energy, E_a , have a major effect on the shape of the curves. For instance, the dashed line in Fig. 3 shows the effect of changing E_a from 14 to 17 kcal/mol. All the parameters needed to compute the curves can be estimated from permeability measurements. Once computed, these curves can be used to estimate the probability of intracellular freezing as a function of cooling rate. Cells that have dehydrated close to equilibrium, prior to reaching their ice-nucleation temperature, will have a zero probability of undergoing intracellular freezing. Cells that are still extensively supercooled when cooled to their nucleation temperature, and therefore still hydrated, will have a high probability of undergoing intracellular freezing.

It should be mentioned that in the preceding text, Mazur's theory and method have been introduced as an example to show how to predict an optimal cooling

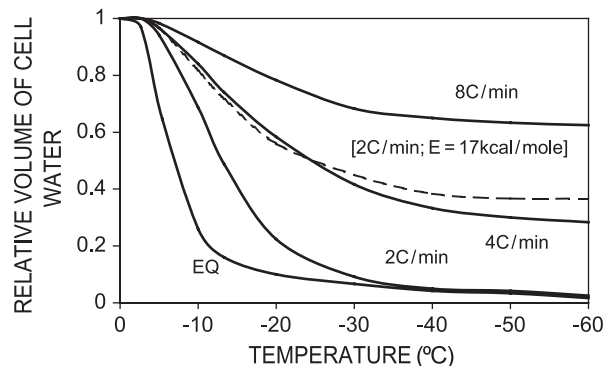


Fig. 3 Computed kinetics of water loss from mouse ova cooled at 2 to $8^{\circ}\text{C}/\text{min}$ in 1 M DMSO. The curve labeled EQ shows the water content that ova have to maintain to remain in equilibrium with extracellular ice. The other solid curves, labeled 2 to $8^{\circ}\text{C}/\text{min}$, were computed assuming the activation energy E of L_p to be 14 kcal/mol. The dashed curve shows the effect of changing E to 17 kcal/mol. (Redrawn from Ref. [15].)

rate to prevent IIF. Different parameters and cryobiological properties of cells will be needed to predict optimal cooling conditions to prevent IIF if different physical models are used, as described by Toner et al.,^[7,8] and Muldrew and McGann.^[4]

PREVENTING CRYOINJURY DURING SLOW FREEZING

It could be inferred from the preceding text that a slow cooling process should be used to avoid IIF. While the avoidance of numerous and large intracellular ice crystals is necessary for cell survival, it is not sufficient. Most cells also require the presence of cryoprotective agents (CPAs). The role of glycerol as an effective CPA for sperm and human erythrocytes was discovered nearly 50 years ago by a research group in Mill Hill, England.^[1,35] Glycerol remains one of the most effective and commonly used CPAs to date, although others have since emerged and are, in some cases, more effective, namely dimethylsulfoxide, ethylene glycol, methanol, propylene glycol, and dimethylacetamide. The most important characteristic of all these solutes is that they are, for the most part, readily able to permeate cells and they are relatively nontoxic to cells in concentrations approaching 1 M or more. The primary basis for their protective action is that they lower the concentration of electrolytes during freezing and decrease the extent of osmotic shrinkage at a given low temperature.^[15,36] The extent of protection depends primarily on the molar ratio of the CPA to endogenous solutes inside and outside the cells, and the general protective mechanism of action is thus

colligative. The effectiveness of a given CPA for a given cell type usually depends on the permeability of that cell to that CPA and that CPA's toxicity. While this class of CPAs must permeate to protect, their permeation before freezing and their removal after thawing generates osmotic volume changes, which themselves can be damaging,^[37-39] a topic that will be further discussed in the following text.

The second class of CPAs is composed of non-permeating solutes. This includes sugars and higher molecular weight compounds such as polyvinylpyrrolidone (PVP), hydroxyethyl starch (HES), polyethylene glycols (PEG), and dextrans. In most cases, these solutes will not protect in the absence of a permeating CPA, but will often substantially augment the effectiveness of a permeating CPA or permit the use of a lower concentration of permeating CPA. These non-permeating CPAs may contribute to enhance vitrification of the solutions, stabilize proteins and membranes, and prevent progressive ice formation.^[40-42]

PREVENTING CRYOINJURY BY VITRIFICATION

All preceding discussions refer to the role of CPAs during slow equilibrium freezing at low enough cooling rates to prevent intracellular ice formation (IIF). A diametrically different approach to cryopreservation is to use either high concentrations of certain CPAs or ultrarapid cooling rates ($>10^6$ °C/min) to induce the cell cytoplasm to form a glass (i.e., to vitrify cells/tissues) rather than to crystallize. However, ultrarapid cooling rates are technically difficult to achieve. Several of the CPAs that are effective in ameliorating slow freeze injury also act to promote glass formation, but the required concentrations are so high, e.g., 4–8 M,^[41] that they can be very toxic to the cells/tissues.^[41,42] Recently, Jackson et al.^[43] demonstrated the significant effect of interaction between the microwave and CPA concentration on vitrification of aqueous solutions and is developing a novel single-mode microwave cavity to enhance the vitrification with relatively low CPA concentrations and relatively low cooling rates.

CRYOINJURY DURING THE STORAGE PERIOD

Usually, the cryopreserved samples are stored at or below -70 °C. The question is often asked as to the length of time that cells can be kept in the frozen state without damage. The question is probably moot if the storage temperature is below -120 °C and is certainly moot at -196 °C (liquid nitrogen). Below -120 °C, chemical reactions cannot occur in human-relevant times. At -196 °C, no thermally driven reactions can

occur in less than geologically relevant times. The reactions that can occur are the slow accumulation of direct damage from ionizing radiation, but this becomes significant only after centuries of storage.^[15]

CRYOINJURY ASSOCIATED WITH WARMING PROCESS

A cell that has survived cooling to low subzero temperatures still faces challenges during warming and thawing that can exert effects on survival comparable to those of cooling.^[15] These effects depend on whether the prior rate of cooling has induced intracellular freezing or cell dehydration. In the former case, rapid thawing can rescue many cells, possibly because it can prevent the growth of small intracellular ice crystals into harmful large ice crystals (i.e., so-called recrystallization). Even when cells are cooled slowly enough to preclude intracellular freezing, the response to warming rate is often highly dependent on the freezing conditions and cell type and is difficult to produce a priori. For human sperm, previous reports used a wide variety of imprecise methods yielding a wide variety of warming rates. These include thawing in ambient air, in water baths with temperatures ranging between 5 and 37°C, in coat pockets, and thawing the samples in the investigator's hands. A more comprehensive study was conducted by Mahadevan in which warming rates between 9.2°C/min and 2140°C/min utilizing coat pocket, hands, air (5, 20, or 37°C), or water baths (5, 20, 37, 55, or 75°C) were examined in combination with a standard freezing rate of 10.0°C/min from +5 to -80 °C.^[43] The results of that study indicated that slower warming in 20 or 35°C air resulted in more optimal cryosurvival of human sperm in terms of both motility and supravital staining, although there were significant differences in post-thaw motilities only between the two fastest rates (1837 and 2140°C/min) and the other 12 rates. Additionally, Mahadevan examined the interaction between cooling and warming rates.^[43] Slow warming was reported optimal for samples frozen slowly; however, there was little effect of warming rate when a fast cooling rate was employed probably because the cells had been killed by the IIF during the fast cooling process before the warming process started.

OSMOTIC AND POTENTIAL TOXIC EFFECTS OF CRYOPROTECTIVE AGENTS (CPAs) ON CELL INJURY

Survival of cells subjected to cryopreservation depends not only on the presence of CPAs but also on the concentration of the CPAs. For example, the

percentage of mouse bone marrow stem cells that survive freezing after cooling at the optimal rate increases dramatically as the concentration of glycerol increases from 0.4 to 1.25 M.^[44] The results are similar for other cells.^[15] Two important procedures related to the use of permeating cryoprotective agents (CPAs) in the cryopreservation of cells are: 1) the addition of a CPA to the cells before freezing and 2) the removal of the CPA from the cells after thawing. Cells transiently shrink upon addition of a CPA and then return to near-normal volume as the CPA permeates. They undergo transient volume expansion during the removal of the CPA, the magnitude of which depends on how the removal is effected and on the inherent permeability of the cell to water and CPA. In several cell types, for example, oocytes, the addition of cryoprotectants (DMSO and 1,2-propanediol) has been shown to cause disruption of the cytoskeleton chemically.^[45,46] Sperm of many species (e.g., bull, boar, and human) are damaged by exposure to glycerol.^[47] Sherman^[48] reported that glycerol exposure caused a reduction in human sperm motility. Critser et al.^[49] have demonstrated that exposure of human sperm to 7.5% glycerol for as little as 15 min causes a dramatic reduction in sperm motility, accounting for approximately 50% of the motility loss observed after 24 h post-thaw. Historically, bull spermatozoa were exposed to medium containing glycerol for an extended period of time, which was referred to as equilibration time. However, Bemdtson and Foote reported that glycerol was able to permeate bull sperm at either 25 or 5°C within 3 to 4 min, and found further that maximum post-thaw motility occurred when the sperm were exposed to glycerol for the shortest time measured (10 sec).^[50,51] These data suggest that bull spermatozoa are either highly permeable to glycerol or, unlike other cells, permeation is not required for protection. What then is the role of equilibration time? Is it, as some have proposed, a matter of allowing for protective membrane alterations?^[47] It is important to note that human sperm differs from the sperm of most domestic species and does not show enhanced cryosurvival with increasing delay between addition of CPA and freezing.^[48,52] Karow et al. found that the survival of human sperm in vitro at 3°C was significantly enhanced by 20 mM potassium (K⁺) in Tyrode's solution relative to Tyrode's solution with less K⁺.^[53] A metabolizable sugar such as glucose was essential to maintain sperm viability in K⁺-free media. The addition of raffinose to media containing glucose improved motility of sperm stored at 3°C for 6 h. The detrimental effect of glycerol to sperm can be alleviated to a certain extent by reducing the temperature at which the cryoprotectant is added in both bull and human sperm cells.^[49,52,54] To what extent the sensitivity of sperm to glycerol reflects the osmotic

consequence of its addition and removal, or to what extent it reflects chemical toxicity, is a current subject of investigation. Recently, it has been shown that for human sperm, the adverse effect of glycerol is mainly attributable to the osmotic volume changes induced during addition and removal of glycerol rather than specific chemical action.^[37] In this study, a systematic investigation was also conducted to determine the human sperm volume excursion (shrinkage or expansion) tolerance limits beyond which the cells will be irreversibly injured. The tolerance limits determined for human sperm in combination with the permeability data of cell membrane to water and CPA have been used to develop the optimal multistep procedures for addition and removal of glycerol in sperm to prevent cell osmotic injury.^[37]

COLD SHOCK INJURY AT LOW TEMPERATURES ABOVE THE WATER FREEZING POINT

Currently, most procedures for cryopreservation of sperm, across species, utilize an initial slow cooling rate between body temperature or the temperature at which the sample is collected (often ambient temperature) and +5°C.^[47,52,55,56] In general, this is due to the fact that some species of mammalian spermatozoa are sensitive to temperature changes in this range, with abrupt cooling (a high cooling rate) causing a high frequency of the cells to become irreversibly damaged. This phenomenon has been termed cold shock or chilling injury.^[47,57] The sensitivity of sperm to cold shock varies among species; bull, ram, boar, and stallion sperm are highly sensitive, dog and cat sperm are moderately sensitive, and rabbit and human sperm are relatively resistant.^[47,49,54,57] The resistance of human sperm to cold shock is of particular importance in the context of developing new approaches to human sperm freezing. It is also an important example of the differences in fundamental cryobiology of cell types among species, highlighting the fact that simple modifications of procedures developed for freezing of domestic animal sperm are unlikely to yield optimal results in other species. Currently, the mechanism of cold shock is unclear.

OTHER POTENTIAL MECHANISMS OF CRYOINJURY

In most previous studies to date, cells have generally been considered simple compartments of cytoplasmic solution enclosed by a semipermeable plasma membrane. The emphasis has been on how cell survival is related to the physical responses of the whole cell to



the physical-chemical events involved in the cryopreservation process. Survival has generally been defined in terms of its viability (e.g., the intactness of cell membrane, etc.) and/or the ability of the cells to undergo subsequent growth and development (e.g., hematopoietic stem cells or embryos). Intactness of the plasma membrane is obviously necessary for functional cell survival. However, in many cases (e.g., granulocytes; Armitage and Mazur^[58]), intactness of the plasma membrane is not sufficient. Findings like the latter emphasized that cells are not simply bags of cytoplasm bound by a plasma membrane. Rather, within the bounds of the plasma membrane are other membrane-bound structures and organelles essential to cell functions. Little is known about how these intracellular structures and organelles respond to freezing, partly because it is difficult to assay their state and function *in situ* in an unambiguous fashion.^[59]

The spermatozoon has been used as an especially good model for investigating the cryobiological role of both intracellular structures and the plasma membrane, because sperm possess two clearly defined and measurable characteristics, motility and the acrosome reaction phenomenon, both of which depend on both functional integrity of organelles and the plasma membrane. An especially attractive feature of the acrosome reaction endpoint in cryobiology is that it involves membrane fusional events.^[60] Investigation of the effects of freezing on such a reaction may provide an experimental model for the growing view that membrane fusion plays a major role in cryobiological injury.^[28,61] To be fully functional, sperm must reach the site of fertilization and the oocyte, capacitate, undergo the acrosome reaction, penetrate the zona pellucida, and fuse with the plasma membrane of the oocyte. Freezing could interfere with or ablate the capacity of the sperm to undergo one or more of these steps.^[62-64] While a motile spermatozoon is not necessarily a fully functional cell, a nonmotile sperm or one exhibiting a damaged plasma membrane is almost certainly nonfunctional (in the context of unassisted fertilization). Consequently, multiple levels of biological function must be considered in assessing the role of the biophysical parameters that have been shown to be critical to cell survival.

SUMMARY

Although there is great diversity in the cryobiological response of different cell types or given cells among different mammalian species, cryosurvival requires that cell freezing and thawing be carried out within certain biophysical and biological limits defined by cryobiology principles. Namely: 1) Cells must be

frozen in such a way that little or none of their water freezes intracellularly. In addition, they must be warmed in such a way that any unfrozen intracellular water remains unfrozen during warming, or that small ice crystals that form during cooling remain small during warming.^[15] 2) Even when these conditions are met, most cells will not survive unless substantial concentrations of CPAs are present. These CPAs must be introduced before freezing and removed after thawing in ways that do not exceed osmotically tolerable limits. Their concentrations also must not be toxic. Although these general limits are necessary, they may not be sufficient, for a number of possible reasons. One is that cells may be injured by factors such as cold shock that have nothing to do with ice formation or CPA damage. Another reason is that cell viability limits are defined primarily in terms of an intact plasma membrane that retains normal, semipermeable properties. It is possible that conditions that allow the plasma membrane to survive may not allow the survival of critical organelles inside cells.

As reviewed in the preceding text, with the progress in understanding mechanisms of injury to cells during cryopreservation, various physical modeling and mathematical formulations have been developed in cryobiology research to simulate a cell's response to environmental change during the cryopreservation process and to predict optimal cryopreservation conditions.^[2-4,7,65] The ability of these predictive models has been confirmed using data collected on cell types ranging from plant protoplasts to mammalian embryos.^[4,7,15,37,38,65] Application of these models to a given cell type requires the determination of a series of cell cryobiological characteristics, including 1) osmotic behavior of the cells, cell volume and surface area, as well as osmotically inactive intracellular water fraction; 2) temperature dependence of cell membrane permeability coefficients to water and the CPA as well as reflection coefficient of cell membrane to the CPA; 3) temperature dependence of safe volume excursion limits of cells; 4) intracellular ice formation temperatures, 5) kinetic and thermodynamic heterogeneous nucleation parameters, etc. These characteristics of many cell types have not been investigated until recently. Determination of these cryobiological characteristics as well as further understanding of the mechanisms of cryoinjury and cryoprotection of CPAs are ongoing research in fundamental cryobiology. Other important on-going research areas in fundamental cryobiology include the investigations of 1) mechanisms and modeling of cryoinjury and cryoprotection to multicellular systems, i.e., tissues and organs^[66,67] and 2) mechanisms of new potential CPAs, such as antifreeze proteins^[68] and inhibitors of cell apoptosis.^[69]

ARTICLES OF FURTHER INTEREST

Stem Cells; Vitrification

REFERENCES

1. Polge, C.; Smith, A.U.; Parkes, A.S. Revival of spermatozoa after vitrification and dehydration at low temperatures. *Nature* **1949**, *164*, 666–676.
2. Mazur, P. Kinetics of water loss from cells at subzero temperatures and the likelihood of intracellular freezing. *J. Gen. Physiol.* **1963**, *47*, 347–369.
3. Mazur, P. Equilibrium quasi-equilibrium, and nonequilibrium freezing of mammalian embryos. *Cell Biophys.* **1990**, *17*, 53–92.
4. Muldrew, K.; McGann, L.E. The osmotic rupture hypothesis of intracellular freezing injury. *Biophysics* **1994**, *66* (1), 532–541.
5. Diller, K.R.; Cravalho, E.G. A cryomicroscope for the study of freezing and thawing process in biological systems. *Cryobiology* **1970**, *7*, 191–199.
6. Diller, K.R. Quantitative low temperature optical microscopy of biological systems. *J. Microsc.* **1972**, *126*, 9–28.
7. Toner, M.; Cravalho, E.G.; Karel, M. Thermodynamics and kinetics of intracellular ice formation during freezing of biological cells. *J. Appl. Phys.* **1990**, *67*, 1582–1593.
8. Toner, M.; Cravalho, E.G.; Stachecki, J.; Fitzgerald, T.; Tompkins, R.G.; Yarmuch, M.L.; Armant, D.R. Non-equilibrium freezing of one-cell mouse embryos. *Biophys. J.* **1993**, *64*, 1921–1980.
9. Karlsson, J.M.; Cravalho, E.G.; Toner, M. A model of diffusion-limited ice growth inside biological cell during freezing. *J. Appl. Phys.* **1994**, *75*, 4442–4452.
10. Farrant, J. Mechanisms of Injury and Protection in Living Cells and Tissues at Low Temperatures. In *Current Trends in Cryobiology*; Smith, A.U., Ed.; Plenum: New York, 1970; 139–152.
11. Farrant, J. Water transport and cell survival in cryobiological procedures. *Philos. Trans. R. Soc. Lond., B* **1977**, *278*, 291–306.
12. Fujikawa, S. Freeze-fracture and etching studies on membrane damage on human erythrocytes caused by formation of intracellular ice. *Cryobiology* **1980**, *17*, 351–362.
13. Mazur, P. Freezing and Low Temperature Storage of Living Cells. In *Proceedings of the Workshop on Basic Aspects of Freeze Preservation of Mouse Strains*; Muhlbock, O., Ed.; , 1976; 1–12. Jackson Laboratory: Bar Harbor, Gustav Fisher Verlag, Stuttgart. .
14. Mazur, P. The role of intracellular freezing in the death of cells cooled at superoptimal rates. *Cryobiology* **1977**, *14*, 251–272.
15. Mazur, P. Freezing of living cells: Mechanisms and implications. *Am. J. Physiol.* **1984**, *M7*, C125–C142. (Cell Physiol. 16).
16. Rapatz, G.; Nath, J.; Luyet, B. Electron microscope study of erythrocytes in rapidly frozen mammalian blood. *Biodynamica* **1963**, *9*, 83–94.
17. Steponkus, P.L.; Wiest, S.C. Freeze-Thaw Induced Lesions in the Plasma Membrane. In *Low Temperature Stress in Crop Plants: The Role of the Membrane*; Lyons, M., Graham, D.G., Raison, J.K., Eds.; Academic Press: New York, 1979; 231–253.
18. Stowell, R.E.; Young, D.E.; Arnold, E.A.; Tromp, B.F. Structural, chemical and functional alterations in mammalian nucleus following different conditions of freezing, storage and thawing. *FASEB Fed. Proc.* **1965**, *24* (Suppl. 15), S115–S141.
19. Trump, B.F.; Young, D.F.; Arnold, E.A.; Stowell, R.E. Effects of freezing and thawing on the structure, chemical constitution and function of cytoplasmic structures. *FASEB Fed. Proc.* **1965**, *24* (Suppl. 15), S144–S167.
20. Lovelock, J.E. The denaturation of lipid-protein complexes as a cause of damage by freezing. *Proc. R. Soc. Lond., B* **1957**, *147*, 427–434.
21. Levitt, J.A. Sulfhydryl disulphide hypotheses of frost injury and resistance in plants. *J. Theor. Biol.* **1962**, *3*, 355.
22. Karow, A.M.; Gilbert, W.B., Jr.; Black, J.B. Effects of temperature, potassium and sugar on human spermatozoa motility: A cell preservation model from reproductive medicine. *Cryobiology* **1992**, *29*, 250–254.
23. Mazur, P.; Leibo, S.P.; Chu, E.H.Y. A two-factor hypothesis of freezing injury. *Exp. Cell Res.* **1972**, *71*, 345–355.
24. Meryman, H.T. The Exceeding of a Minimum Tolerable Cell Volume in Hypertonic Suspension as a Cause of Freezing Injury. In *The Frozen Cell*; Wolstenhoime O.E.W., O'Connor, M., Eds.; Ciba Foundation Symposium, Churchill: London, 1970; 51–64.
25. Williams, R.J. The mechanisms of cryoprotection in the intestinal mollusk *Mytilus*. *Cryobiology* **1979**, *4*, 250.
26. Fishbein, W.M.; Winkert, J.W. Parameters of biological freezing damage in simple solutions: Catalase 11 demonstration of an optimum recovery cooling rate curve in a membraneless system. *Cryobiology* **1978**, *15*, 168.
27. Pegg, D.E.; Diaper, M.P. The mechanism of injury to slowly-frozen erythrocytes. *Biophys. J.* **1988**, *54*, 471–488.
28. Mazur, P.; Cole, K.W. Influence of cell concentration on the contribution of unfrozen fraction and salt concentration to the survival of slowly frozen human erythrocytes. *Cryobiology* **1985**, *22*, 509–536.
29. Levin, R.L.; Cravalho, E.G.; Huffins, C.G. A membrane model describing the effect of temperature on the water conductivity of erythrocyte membranes at subzero temperatures. *Cryobiology* **1976**, *13*, 415–429.
30. McGrath, J.J. Preservation of Biological Material by Freezing and Thawing. In *Heat Transfer in Medicine and Biology*; Shitzer, A., Eberhart, R.C., Eds.; Plenum: New York, 1985.
31. McGrath, J.J. Membrane Transport Properties. In *Low Temperature Biotechnology: Emerging Applications and Engineering Contributions*; McGrath, J.J., Diller, K.R., Eds.; A.S.M.E. 1988; 273–330. BED, HTD- Vol. 10; HTD- Vol. 98.
32. Gao, D.Y.; Liu, C.; Benson, C.; Liu, J.; Lin, S.; Critser, E.S.; Critser, J.K. Theoretical and Experimental



- Analyses on the Optimal Experimental Design for Determination of Hydraulic Conductivity of Cell Membrane. In *Advances in Heat and Mass Transfer in Biological System*; Hayes, L., Roemer, R.B., Eds.; ASME Press, HTD: New York, 1994; Vol. 288, 151–158.
33. Gao, D.Y.; Benson, C.T.; Liu, C.; McGrath, J.J.; Critser, J.S.; Critser, J.K. Development of a novel microperfusion chamber for determination of cell membrane transport properties. *Biophys. J.* **1996**, *71*, 443–450.
 34. Devireddy, R.V.; Raha, D.; Bischof, J.C. Measurement of water transport during freezing in cell suspensions using a differential scanning calorimeter. *Cryobiology* **1998**, *36*, 124–155.
 35. Polge, C.; Lovelock, J.E. Preservation of bull sperm at -70°C . *Vet. Rec.* **1952**, *64*, 296–297.
 36. Woods, E.; Zieger, M.; Gao, D.Y.; Critser, J.K. Equations for obtaining melting points for the ternary system ethylene glycol/sodium chloride/water and their application to cryopreservation. *Cryobiology* **1999**, *38*, 403–407.
 37. Gao, D.Y.; Liu, J.; Liu, C.; McGann, L.E.; Watson, P.F.; Kleinhans, F.W.; Mazur, P.; Critser, E.S.; Critser, J.K. Prevention of osmotic injury to human spermatozoa during addition and removal of glycerol. *Hum. Reprod.* **1995**, *10*, 1109–1122.
 38. Mazur, P.; Schneider, U. Osmotic response of preimplantation mouse and bovine embryos and their cryobiological implications. *Cell Biophys.* **1986**, *8*, 259–285.
 39. Schneider, U.; Mazur, P. Osmotic consequences of cryoprotectant permeability and its relation to the survival of frozen-thawed embryos. *Theriogenology* **1988**, *21*, 68–79.
 40. Takahashi, T.; Inada, S.; Pommier, C.G.; O'Shea, J.J.; Brown, E.J. Osmotic stress and the freeze-thaw cycle cause shedding of Fc and C3b receptors by human polymorphonuclear leukocytes. *J. Immunol.* **1985**, *34*, 4062.
 41. Fahy, G.M.; MacFarlane, D.R.; Angell, C.A.; Meryman, H.T. Vitrification as an approach to cryopreservation. *Cryobiology* **1984**, *21*, 407–426.
 42. Fahy, G.M. The relevance of cryoprotectant "toxicity" to cryobiology. *Cryobiology* **1986**, *23* (1), 1–13.
 43. Mahadevan, M. Cryobiological and Biochemical Studies of Human Semen. In *Ph.D. Thesis*; Monash University: Melbourne, Victoria, Australia, 1980.
 44. Leibo, S.P.; Farrant, J.; Mazur, P.; Hanna, M.G.; Smith, L.H. Effects of freezing on marrow stem cell suspensions: Interactions of cooling and warming rates in the presence of PVP, sucrose or glycerol. *Cryobiology* **1970**, *6*, 315–332.
 45. Johnson, M.H.; Pickering, S.J. The effects of dimethylsulfoxide on the microtubular system of the mouse oocyte. *Development* **1987**, *100*, 313–324.
 46. Joly, C.; Behini, O.; Boulekbache, H.; Testart, J.; Maro, B. Effects of 1,2-propanediol on the cytoskeletal organization of the mouse oocyte. *Hum. Reprod.* **1992**, *7*, 374–378.
 47. Watson, P.F. The Preservation of Semen in Mammals. In *Oxford Reviews of Reproductive Biology*; Finn, C.A., Ed.; Oxford Univ. Press: London, 1979; 283–350.
 48. Sherman, J.K. Improved methods of preservation of human spermatozoa by freezing and freezing-drying. *Fertil. Steril.* **1964**, *14*, 49–64.
 49. Critser, J.K.; Huse-Benda, A.R.; Aaker, D.V.; Arneson, B.W.; Ball, G.D. Cryopreservation of human spermatozoa. III. The effect of cryoprotectants on motility. *Fertil. Steril.* **1998**, *50*, 314–320.
 50. Berndtson, W.E.; Foote, R.H. Bovine sperm cell volume at various intervals after addition of glycerol of 5C. *Cryobiology* **1972**, *9* (1), 29–33.
 51. Berndtson, W.E.; Foote, R.H. The freezability of spermatozoa after minimal prefreezing exposure to glycerol or lactose. *Cryobiology* **1972**, *9* (1), 57–60.
 52. Sherman, J.K. Synopsis of the use of frozen human semen since 1964: State of the art of human semen banking. *Fertil. Steril.* **1973**, *24*, 397–416.
 53. Karow, A.M.; Gilbert, W.B., Jr.; Black, J.B. Effects of temperature, potassium and sugar on human spermatozoa motility: A cell preservation model from reproductive medicine. *Cryobiology* **1992**, *29*, 250–254.
 54. Critser, J.K.; Arneson, B.W.; Aaker, D.V.; Huse-Benda, A.R.; Ball, G.D. Factors affecting the cryosurvival of mouse two-cell embryos. *J. Reprod. Fertil.* **1988**, *82* (1), 27–33.
 55. Foote, R.H. Semen quality from the bull to the freezer: An assessment. *Theriogenology* **1975**, *3*, 219–235.
 56. Graham, E.G. Fundamentals of the Preservation of Spermatozoa. In *The Integrity of Frozen Spermatozoon*; National Academy of Sciences: Washington, DC, 1978; 4–44.
 57. Watson, P.F. Recent developments and concepts in the cryopreservation of spermatozoa and the assessment of their post-thawing function. *Reprod. Fertil. Dev.* **1995**, *7*, 871–891.
 58. Armitage, W.J.; Mazur, P. Osmotic tolerance of human granulocytes. *Am. J. Physiol.* **1984**, *247*, C373–C381.
 59. McGann, L.E.; Yang, H.; Walterson, M. Manifestations of cell damage after freezing and thawing. *Cryobiology* **1988**, *25*, 178–185.
 60. Yudin, A.L.; Gottlieb, W.; Meizel, S. Ultrastructural studies of the early events of the human sperm acrosome reaction as initiated by human follicular fluid. *Gamete Res.* **1988**, *20*, 11–24.
 61. Mazur, P.; Cole, K.W. Roles of unfrozen fraction, salt concentration, and changes in cell volume in the survival of slowly frozen human erythrocytes. *Cryobiology* **1989**, *26*, 1–29.
 62. Critser, J.K.; Huse-Benda, A.R.; Aaker, D.V.; Arneson, B.W.; Ball, G.D. Cryopreservation of human spermatozoa. I. Effects of holding procedure and seeding on motility, fertilizability, and acrosome reaction. *Fertil. Steril.* **1987**, *47*, 656–663.
 63. Critser, J.K.; Arneson, B.W.; Aaker, D.V.; Ball, G.D. Cryopreservation of human spermatozoa. II. Post-thaw chronology of motility and of zona-free hamster ova penetration. *Fertil. Steril.* **1987**, *47*, 980–984.
 64. Wheeler, M.B.; Seidel, G.E. Time course of in vitro capacitation of frozen and unfrozen bovine spermatozoa. *Intl. Embryo Trans. Soc.* **1986**, *216*. (ABSTRACT).

65. Pitt, R.E.; Steponkus, P.L. Quantitative analysis of the probability of intracellular ice formation during freezing of isolated protoplasts. *Cryobiology* **1989**, *24*, 44–63.
66. Rubinsky, B.; Pegg, D. A mathematical model for the freezing process in biological tissue. *Proc. R. Soc.* **1988**, *234*, 343–358.
67. Bischof, J.; Hunt, C.J.; Rubinsky, B.; Burgess, A.; Pegg, D.E. The effect of cooling rate and glycerol concentration on the structure of the frozen kidney: Assessment by cryo-scanning electron microscopy. *Cryobiology* **1990**, *27*, 301–310.
68. Rubinsky, B.; Arav, A.; DeVries, A.L. The cryoprotective effect of antifreeze glycopeptides from antarctic fishes. *Cryobiology* **1992**, *229*, 69–72.
69. Baust, J.M.; Hollister, W.; Van Buskirk, R.; Baust, J.G. Cryopreservation outcome is enhanced by intracellular-type media and inhibition of apoptosis. *Cryobiology* **1998**, *37* (4), 410–411.



Defibrillators

Paul Wang

Stanford University Medical Center, Stanford, California, U.S.A.

INTRODUCTION

Defibrillators are devices that deliver electrical shocks to the heart in order to convert rapid irregular rhythms of the upper and lower heart chambers to normal rhythm. External defibrillators deliver energy to the chest externally, creating a high energy field in the heart. Automatic external defibrillators (AEDs) have automated features that permit their use by nonmedical personnel and lay people. Implantable defibrillators are fully implantable devices that deliver electric energy internally to the heart.

DEFIBRILLATION

Ventricular Fibrillation and Ventricular Tachycardia

Basic definitions

Normal Conduction System and Sinus Rhythm. The normal impulse originates in the sinus node, which lies within the right atrium, one of the two upper chambers of the heart. The impulse from the sinus node spreads through the two atria and activates the auricular-ventricular (A-V) node and His-Purkinje system, which spreads the electrical activation through the ventricles (Fig. 1). The left ventricle is the primary chamber of the heart responsible for pumping blood to the body. This pattern of activation is responsible for the normal rhythm, called sinus rhythm. With activity there is an increase in heart rhythm; the increased heart rate provides an increased amount of blood flow to the body required by exertion.

Ventricular Tachycardia. In some cases, particularly in the setting of severe abnormalities of function of the ventricles, the electrical impulse in the heart originates in the ventricles rather than in the atria. These rhythms, when rapid, are termed ventricular tachycardia or ventricular fibrillation. Ventricular tachycardia is a rapid heart rhythm originating in the ventricles, usually from one location at a time, at a rate over 100 beats per minute and usually between 150 and 250 beats per minute. Ventricular tachycardia most commonly originates from a relatively constant position

within the ventricles. Each beat of such ventricular tachycardias on the surface electrocardiogram (ECG) is nearly identical. These ventricular tachycardias are considered monomorphic. In contrast, ventricular tachycardias may be less commonly polymorphic, meaning that they originate from shifting locations in the ventricles, and each beat's shape changes over time. Ventricular tachycardias frequently occur in individuals who have had prior heart attacks, called myocardial infarctions. The rapid activation of the ventricles during ventricular tachycardia is inappropriate to the metabolic demands of the heart and body and causes low blood pressure and poor output of the heart. The decreased blood flow through the heart, called cardiac output, may manifest as the absence of a pulse.

Ventricular Fibrillation. Ventricular fibrillation is an extremely rapid rhythm originating in the ventricles but without a discrete point of origin. Multiple regions of the ventricles are activated simultaneously by moving wavefronts that are changing over time (Fig. 2). The ventricular rate is difficult to count and exceeds 300 beats per minute. Activation of many regions of the left ventricle results in no significant contraction of the ventricles. Ventricular fibrillation uniformly results in immediate death unless defibrillation is performed.

Atrial Fibrillation. Atrial fibrillation is an extremely rapid rhythm originating in the atria without a discrete point of origin (Fig. 2). There are multiple wavefronts of electrical activation throughout the atria. Atrial fibrillation results in no significant contraction in the atria. However, passive blood flow from the atria fills the left ventricle. Therefore, patients are usually able to function fairly normally during atrial fibrillation, and the ventricles continue to contract. Patients may be asymptomatic or may experience symptoms such as shortness of breath, fatigue, or rapid heart sensation.

Mechanisms of ventricular fibrillation

Ventricular fibrillation is characterized by multiple waves of cardiac excitation that travel from one region to the next adjacent one. The activation pattern changes from beat to beat. Each beat is called a cycle.

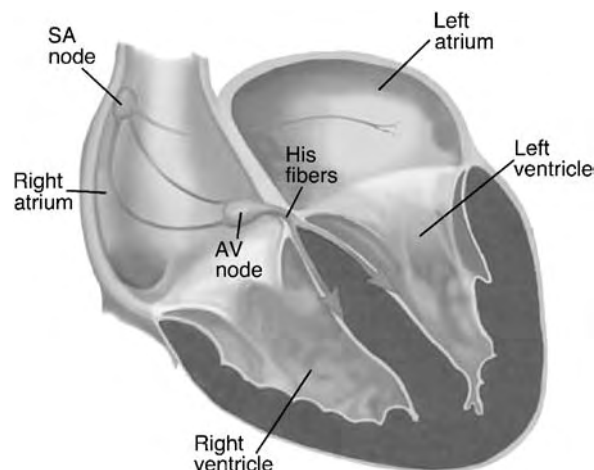


Fig. 1 Normal conduction system of the heart. The normal impulse originates in the sinus node of the heart within the upper chamber called the right atrium. The impulse travels through the right and left atria and reaches the A-V node. The impulse then travels through the His-bundle and Purkinje fibers to the ventricles.

The period of time between cycles is approximately 210 milliseconds (ms).^[1] Activation occurs so rapidly that each region of tissue has just recovered the ability to be electrically activated, a state called excitability. There may be brief periods during which the ventricular tissue is excitable prior the arrival of the next wavefront, consistent with a so-called excitable gap.^[2] However, existence of an excitable gap during human ventricular fibrillation has not been clearly demonstrated.^[1] Ventricular fibrillation is often described as proceeding through four stages.^[3] In the first two stages, comprising the first one to two minutes, activation occurs rapidly at a cycle length of 90 to 120 ms. In the last two stages, activation slows significantly.

Biophysics of Defibrillation

Mechanism of defibrillation

Defibrillation is based on the termination of the multiple activation wavefronts that are responsible for sustaining ventricular fibrillation.^[4] An electrical field is created, altering the membrane potential of the heart muscle cells called myocytes, and disrupting the activation wavefronts present in ventricular fibrillation.^[4,5] Electrical activation of cardiac tissue, referred to as depolarization, must occur over a sufficiently large region or mass of myocardial tissue to terminate the preexisting wavefronts. Depolarization of up to 90% of the myocardial mass may be necessary for successful defibrillation.^[6]

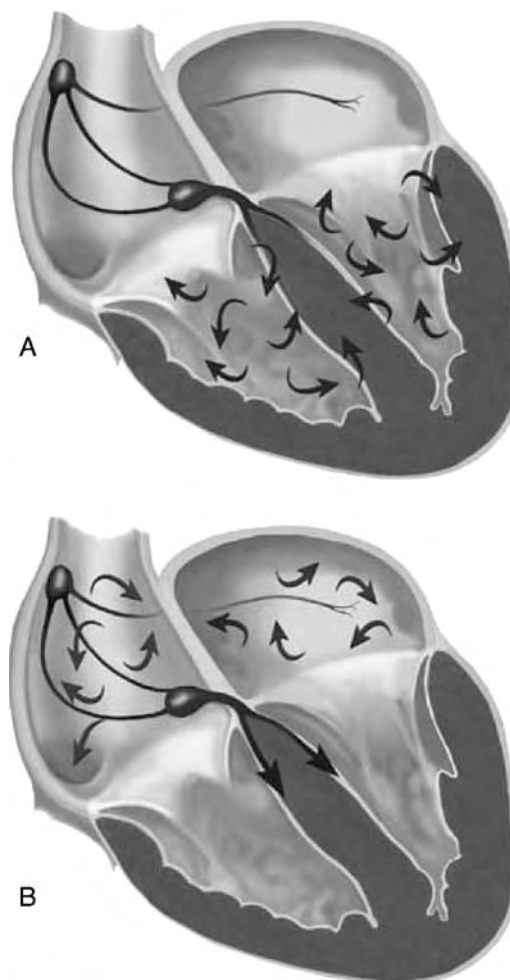


Fig. 2 Schematic diagrams of ventricular fibrillation and atrial fibrillation. (A) Ventricular fibrillation is an extremely rapid rhythm created by multiple wavefronts of electrical activation that are constantly changing. The rate is over 300 beats per minute and results in no significant contraction of the ventricles. Death occurs unless defibrillation is rapidly performed. (B) Atrial fibrillation is an extremely rapid rhythm in the atrium resulting from multiple wavefronts of electrical activation without a discrete point of origin.

Waveforms

The two commonly used clinical waveforms for defibrillation are monophasic and biphasic shocks. In monophasic shock waveforms, the polarity does not change during the shock. In contrast, in biphasic shocks, the polarity of the shock reverses during the shock delivery (Fig. 3). In general, biphasic shocks are more successful in defibrillating the heart than monophasic shocks.^[7]

Implantable defibrillators most commonly use biphasic shock waveforms. Their shocks use a truncated exponential waveform. External defibrillators more recently have begun using biphasic shock waveforms with a damped sinusoidal shape.

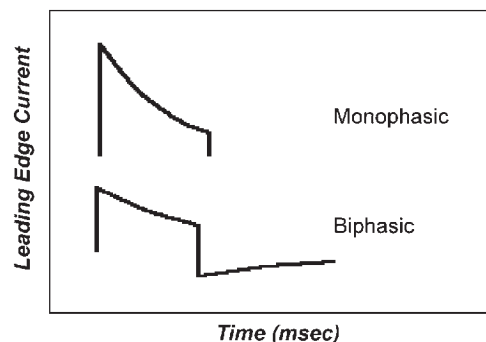


Fig. 3 Monophasic and biphasic waveforms. During monophasic waveforms, the polarity of the energy delivered is constant throughout the duration of the shock. During biphasic waveforms, the polarity reverses during the shock.

EXTERNAL DEFIBRILLATORS

Clinical Use

External defibrillators used by medical personnel

External defibrillators are nonimplantable devices used primarily to convert ventricular fibrillation by delivering a shock to the chest. External defibrillators may also be used to convert ventricular tachycardia and atrial fibrillation. External defibrillators are used by hospital and emergency medical personnel.

Clinical use of AEDs

The survival of patients with cardiac arrest depends critically on the time that elapses before defibrillation is performed. To decrease the time to defibrillation, automatic external defibrillators have been introduced for use by first responders, including lay people and nonmedical personnel such as police and firefighters. Automatic external defibrillators (AEDs) recognize the presence of ventricular fibrillation automatically and prompt a lay person or nonmedical personnel to initiate shock delivery.^[8-12] This strategy of defibrillation using AEDs has been called public access defibrillation (PAD). Studies have demonstrated that the time to defibrillation is decreased by Automatic external defibrillator use by police or firefighter personnel. Automatic external defibrillator use by security personnel in casinos and on airplanes has also been shown to decrease the time to defibrillation.

Design

Electrode system

In external defibrillators, the shock is delivered to the chest using one or two types of electrode systems:

1) a pair of hand-held devices called paddles that have metal plates serving as electrodes, or 2) adhesive electrode pads. A conductive gel or paste is used with the defibrillator paddles to improve conductivity through the skin and minimize skin burns. Increased pressure on the paddles or adhesive electrode pads decreases the impedance through the skin and improves defibrillation success.

Defibrillators

Standard External Defibrillator. A standard external defibrillator consists of a power supply, an electrocardiographic monitor, and a connector to the electrode system. The defibrillator has a battery power supply and an AC power supply. The defibrillator uses a step-up transformer and a capacitor to store the charge. External defibrillators typically deliver from 200 to 360 Joules biphasic energy and up to 360 Joules monophasic energy. For pediatric patients, the size of the paddles or adhesive patches is reduced and the amount of energy may be reduced and adjusted to a target Joule/kg of energy delivered. Most manual external defibrillators are capable of synchronizing the shock to the ECG signal to achieve synchronized cardioversion. Some external defibrillators have the additional ability to stimulate the chest wall via the adhesive patches to pace the heart at, for example, 60 beats per minute as a therapy for slow heart rates.

Automatic External Defibrillator (AED). Significant advances in AED technology have reduced the size, weight, and cost of AEDs, and have improved the accuracy of the ability of AEDs to recognize rhythms automatically. These advances have led to the explosion in the use of AEDs in public settings.

AEDs use self-adhesive electrode pads with diagrams indicating the method of applying them. An automatic algorithm interprets the rhythm. The AEDs provide voice and text prompts to guide the user through the steps of use.

AEDs use biphasic waveforms and large electrode pads to convert the rhythm with a high frequency of success. AEDs use nickel cadmium or nonrechargeable lithium batteries that provide a long shelf life. AEDs test themselves automatically in regular intervals.

IMPLANTABLE CARDIOVERTER-DEFIBRILLATORS (ICDS)

History

Implantable cardioverter-defibrillators (ICDs) are fully implantable devices capable of cardioversion (conversion

of atrial fibrillation and ventricular tachycardia) and defibrillation (conversion of ventricular fibrillation). These devices were first implanted in patients in 1980 and were approved by the Food and Drug Administration in 1985. The implantation of the first ICDs required placing an epicardial patch on the surface of the heart, necessitating open-chest surgery. However, the introduction of transvenous ICD lead systems obviated the need for open-chest surgery.

Design

Implantable defibrillator design elements

Implantable cardioverter-defibrillator (ICD) systems consist of two main components: the ICD itself, called the generator, and the lead(s).

ICD generator

The ICD generator has several critical components: microprocessor, memory units, high-voltage system, pacing circuitry, and sensing amplifiers. Each device is hermetically sealed within a titanium case and is fully implantable (Fig. 4).

Microprocessor and Memory Units. The microprocessor units are responsible for timing and function control in the ICDs. The ICD receives electrical signals from the sensing amplifier system and classifies the rhythm, using a series of recognition algorithms. Once the ICD determines the therapy should be given, charging of the ICD capacitors occurs.

The ICD is capable of digitally storing the electrical signals from the heart during arrhythmia episodes. These stored signals are called electrograms. The ICD



Fig. 4 ICD generator. The ICD generator has a hermetically sealed titanium casing—called a can—that contains all of the components required for defibrillation and pacing other than the lead(s). The ICD lead is inserted into the top (header) portion of the ICD generator.

may now store several minutes of these electrograms. Physicians can retrieve these stored electrograms in order to reconstruct the arrhythmic events leading to device therapy.

High-Energy System. The high-energy system consists of a battery, DC-to-DC converters, and high-energy output switches.^[13] The batteries used in the ICD industry are predominantly based on vanadium oxide. These batteries have a low internal impedance. High current pulsing of the batteries and reforming of the electrolytic capacitors are automatically performed periodically for optimal high-energy system function. Aluminum oxide may accumulate on the anodal plate of the capacitor in aluminum electrolytic capacitors that are used in most ICDs.

The battery of the beginning of device life of an ICD has a voltage output of approximately 3.2 V. However, defibrillation requires delivery of voltages in the range of 600–800 V. DC-to-DC or step-up converters are used to achieve the high voltage needed for ICDs. These converters consist of a low-voltage high-current switch, a fly-back transformer, a rectifying diode, and a capacitor element. The low-voltage switch closes, permitting current flow to the fly-back transformer. The switch opens and the energy stored in the fly-back transformer is transferred to the secondary circuit, resulting in charge storage on the capacitors. This process is repeated in small steps until the required voltage is stored.

Pacing System. The ICD functions not only by shocking the heart but by pacing it, using small-amplitude electrical stimuli. These electrical stimuli, approximately 3 to 5 volts with a 0.4- to 0.6-ms pulse width, are used to depolarize the heart from leads in the atrium and ventricle in dual-chamber ICDs or from the ventricle alone in single-chamber ICDs.

Sensing Amplifiers. Sensing amplifiers are responsible for sensing the cardiac signal. The cardiac signal amplitude typically varies between fractions of a millivolt during ventricular fibrillation to up to 20 millivolts in normal rhythm. The ICD sensing amplifiers must detect signals with this range of amplitudes even when the amplitude is rapidly changing over time. Because the cardiac signal is used mainly to detect the heart rate, it is important that each heart beat be detected only once per cycle. Because repolarization creates an electrical signal in addition to depolarization, sensing amplifiers must distinguish between the primary depolarization signal and the subsequent repolarization signal.

The electrical signal from the sensing lead typically undergoes amplification and band-pass filtering. There are two basic methods of sensing the rapid

signals of varying amplitude that are present in ventricular fibrillation. An automatic gain control system is one sensing method. Using a decay constant, there is a limit to the rate at which the amplifier may respond to variations in amplitude. Each beat reaching a threshold is counted toward the heart rate. In one modification of this system, the threshold is dynamic and varies based on the amplitude of the prior signals.

ICD leads

The primary functions of ICD leads are to provide a path for the high-voltage energy from the ICD generator to the heart and to carry electrical signals from the heart to the ICD generator for rhythm determination. In addition, the ICD leads serve as pacing leads to prevent slow heart rates.

Implantable cardioverter-defibrillators leads use a coil structure to create the high-current density within large regions of the myocardium required for defibrillation (Fig. 5).

Clinical Function and Use

Implantation of ICDs

Implantation of ICDs involves a number of steps: 1) making an incision for insertion of the ICD and leads; 2) accessing the venous system; 3) advancing



Fig. 5 ICD defibrillation lead. An ICD lead typically has two long defibrillation coils along its length. Energy travels from the ICD generator through the length of the ICD lead to the defibrillation coils. The energy delivered from the ICD defibrillation lead coils creates sufficient current density within a large percentage of the ventricular mass of the heart to convert the rhythm to normal. At the tip of the lead is an electrode used for sensing the cardiac signal and for transmitting electrical stimuli from the ICD generator to the heart to maintain an adequate heart rate.

and positioning the ICD leads within the heart; 4) connecting the leads to the ICD generator; 5) placing the ICD generator under the subcutaneous tissue; 6) testing the ICD function; and 7) surgically closing the skin. Implantable cardioverter-defibrillators implantation is performed in a sterile surgical environment in order to minimize the risk of ICD device infection. Implantation of the ICD system requires expertise gained from specific training in this area. The ICD is fully implantable, meaning that there are no external wires or devices required for its function.

Radiofrequency programming

A device called a programmer is used to communicate and send commands to the ICD. Most current programmers are modified portable computers that use radiofrequency energy to communicate with the implanted ICD. Hundreds of parameters or settings in ICD function may be reversibly altered by commands sent from the programmer. The programmer is used to record and measure the electrical signals from the ICD in order to confirm appropriate sensing and proper connections. Device and lead abnormalities are typically detected using this system. Stored electrograms of the arrhythmia episodes necessitating therapy are transferred by radiofrequency energy from the device to the programmer and may be downloaded to a floppy disk for archiving.

Clinical indications and clinical trials

Implantable cardioverter-defibrillators detect and treat life-threatening ventricular arrhythmias. They are primarily indicated in several groups of patients: 1) patients who have survived life-threatening ventricular tachycardia or ventricular fibrillation, requiring resuscitation; 2) patients with short runs of ventricular tachycardia, prior heart attack, and ventricular tachycardia produced during a special test called an electrophysiologic study; and 3) patients with a prior heart attack and severely reduced heart function. Randomized clinical trials have demonstrated improved survival of patients with ICDs in these three patient groups. Implantable cardioverter-defibrillators are also indicated in patients with episodes of loss of consciousness and ventricular tachycardia produced during electrophysiologic study.

Clinical functions

In addition to terminating ventricular arrhythmias using shock therapy, ICDs in some cases may terminate

ventricular tachycardia using rapid pacing, called antitachycardia pacing. By interrupting the activation pathways resulting in ventricular tachycardia, antitachycardia pacing may terminate the rhythm without the need for shock therapy.

Most defibrillators may pace the heart in response to patient activity, a feature called rate-adaptive or rate-responsive pacing. In addition, a new generation of ICDs is capable of pacing not only the atrium and the right ventricle but also the left ventricle, via a lead placed in a vein on the outside of the left ventricle. This vein is accessed via the vein called the coronary sinus, which drains into the right atrium. The activation of both ventricles is called biventricular pacing or resynchronization therapy and has been shown to increase exercise tolerance in patients with severely impaired heart function.

CONCLUSION

External defibrillators deliver energy to the heart via the chest wall. Automatic external defibrillators are being used by nonmedical personnel and lay people in numerous settings. Implantable defibrillators are implantable devices, capable of automatically recognizing life-threatening ventricular arrhythmias and delivering shock therapy directly to the heart. Defibrillators have become critical to improving survival of patients with life-threatening rhythm problems of the heart. Significant technological advances have made these devices more accurate and more successful in converting these rhythms.

ARTICLES OF FURTHER INTEREST

Cardiac Bioelectricity; Cardiac Pacemaker; Electrocardiography

REFERENCES

- Swartz, J.F.; Jones, J.L.; Fletcher, R.D. Characterization of ventricular fibrillation based on monophasic action potential morphology in the human heart. *Circulation* **1993**, *87*, 1907–1914.
- KenKnight, B.H.; Bayly, P.V.; Gerstle, R.J.; Rollins, D.L.; Wolf, P.D.; Smith, W.M.; Ideker, R.E. Regional capture of fibrillating ventricular myocardium: Evidence of an excitable gap. *Circ. Res.* **1995**, *77*, 849–855.
- Worley, S.J.; Swain, J.L.; Colavita, P.G.; Smith, W.M.; Ideker, R.E. Development of an endocardial-epicardial gradient of activation rate during electrically induced, sustained ventricular fibrillation. *Am. J. Cardiol.* **1985**, *55*, 813–820.
- Chen, P.S.; Shibata, N.; Dixon, E.G.; Wolf, P.D.; Daniele, N.D.; Sweeney, M.B.; Smith, W.M.; Ideker, W.M. Activation during ventricular defibrillation. Evidence of complete cessation and regeneration of ventricular fibrillation after unsuccessful shocks. *J. Clin. Invest.* **1986**, *77* (3), 810–883.
- Mower, M.M.; Mirowski, M.; Spear, J.F.; Moore, N.E. Patterns of ventricular activity during catheter defibrillation. *Circulation* **1994**, *48*, 858–861.
- Zipes, D.P.; Fischer, J.; King, R.M.; Nicoll, A.; Jolly, W.W. Termination of ventricular fibrillation by depolarizing a critical amount of myocardium. *Am. J. Cardiol.* **1975**, *36*, 37–44.
- Bardy, G.H.; Marchlinski, F.E.; Sharma, A.D.; Worley, S.J.; Luceri, R.M.; Yee, R.; Halperin, B.D.; Fellows, C.L.; Ahern, T.S.; Chilson, D.A.; Packer, D.L.; Wilber, D.J.; Mattioni, T.A.; Reddy, R.; Kronmal, R.A.; Lazzara, R. Multicenter comparison of truncated biphasic shocks and standard damped sine wave monophasic shocks for transthoracic ventricular defibrillation. *Circulation* **1996**, *94*, 2507–2514.
- Cummins, R.O.; Eisenberg, M.S.; Litwin, P.E.; Graves, J.R.; Hearne, T.R.; Hallstrom, A.P. Automatic external defibrillators used by emergency medical technicians: A controlled clinical trial. *JAMA* **1987**, *27*, 1605–1610.
- Page, R.L.; Joglar, J.A.; Kowal, R.C.; Zagrodzky, J.D.; Nelson, L.L.; Ramaswamy, K.; Barbera, S.J.; Hamdam, M.H.; McKenas, D.K. Use of automated external defibrillators by a U.S. airline. *N. Engl. J. Med.* **2000**, *343*, 1210–1216.
- Valenzuela, T.D.; Roe, D.J.; Nichol, G.; Clark, L.L.; Spaite, D.W.; Hardman, R.G. Outcomes of rapid defibrillation by security officers after cardiac arrest in casinos. *N. Engl. J. Med.* **2000**, *343*, 1206–1209.
- Nichol, G.; Hallstrom, A.P.; Ornato, J.P.; Riegel, B.; Stiell, I.G.; Valenzuela, T.; Wells, G.A.; White, R.D.; Weisfeldt, M.L. Potential cost-effectiveness of public access defibrillation in the United States. *Circulation* **1998**, *97*, 1315–1320.
- Marengo, J.P.; Wang, P.J.; Link, M.S.; Homoud, M.K.; Estes, N.A.M., III. Improving survival from sudden cardiac arrest: The role of the automated external defibrillator. *JAMA* **2001**, *285*, 1193–1200.
- Warren, J.A.; Nelson, J.P. Pacemaker and ICD Pulse Generator Circuitry. In *Clinical Cardiac Pacing and Defibrillation*, 2nd Ed.; Ellenbogen, K.A., Kay, G.N., Wilkoff, B.L., Eds.; W.B. Saunders Co: Philadelphia, 2000; 194–216.

RECOMMENDED READING

The antiarrhythmics versus implantable defibrillators (AVID) Investigators: A comparison of antiarrhythmic-drug therapy with implantable defibrillators in patients resuscitated from near-fatal ventricular arrhythmias. *N. Engl. J. Med.* **1997**, *337*, 1576–1583.

Moss, A.J.; Hall, W.J.; Cannom, D.S.; Daubert, J.P.; Higgins, S.L.; Klein, H.; Levine, J.H.; Saksena, S.; Waldo,



- W.D.; Brown, A.L.; Heo, M.W.M. Improved survival with an implanted defibrillator in patients with coronary disease at high risk for ventricular arrhythmia. *N. Engl. J. Med.* **1996**, *335* (12), 1933–1940.
- Moss, A.J.; Zareba, W.; Hall, W.J.; Klein, H.; Wilber, D.J.; Cannom, D.S.; Daubert, J.P.; Higgins, S.L.; Brown, M.W.; Andrews, M.L. The multicenter automatic defibrillator implantation trial II investigators. Prophylactic implantation of a defibrillator in patients with myocardial infarction and reduced ejection fraction. *N. Engl. J. Med.* **2002**, *346* (12), 877–883.

Degradable Polymer Composites

Debra D. Wright

Department of Biomedical Engineering, Michigan Technological University, Houghton, Michigan, U.S.A.

INTRODUCTION

Degradable polymers have found a number of uses in biomedical engineering. Drug delivery and tissue engineering are but two examples gaining widespread use. In drug delivery, degradable polymers hold the promise of zero order release, or continuous release of a drug that is independent of time. Current devices have not achieved this gold standard, but they still aid in delivering appropriate amounts of drugs with less reliance on patient compliance.^[1] In tissue engineering, a significant amount of research has investigated the use of scaffolds seeded with cells to replace or augment human tissues. Scaffolds have been shown to mediate cell growth in critically sized bone defects^[2] and mediate nerve cell regeneration.^[3] However, degradable polymers by themselves lack the mechanical properties necessary to sustain fixation of many load-bearing bones while they are healing. Therefore, composite materials are being investigated as new options to traditional materials.

OVERVIEW

Current methods of internal fracture fixation include plates, pins, and screws that are fabricated primarily from metallic materials. There are several limitations with metallic implants. Metallic implants lead to stress shielding, i.e., local weakening of the bone, in the area of the implant due to the high mismatch in stiffness between metals and bone. Metallic plates also prevent the natural healing of bone. Natural healing of bone promotes a large callus formation at the fracture site that supports the fracture as it heals and the bone remodels. This was shown by Sarmiento et al.^[4] when he imposed simple transverse fractures of the radius and ulna on two sets of adult dogs. One group had fixation by a metallic plate while the other group had no fixation method. The group with a metallic plate had significantly less bone remodeling activity and no callus formation. This was also seen by Claes^[5] when comparing metallic plates of different stiffnesses and carbon fiber-reinforced polymer plates. Clinically, this is seen after plate removal as the weakened bone is more susceptible to breakage in up to 30% of

patients,^[6-8] which recreated the original problem. Furthermore, approximately 66% of patients require a second surgery to remove the implant.^[9] For these reasons, alternative devices are desired for internal fracture fixation.

Degradable materials have the potential to solve the current limitations with metallic materials. The challenge is to create materials that are strong enough to support the loads seen in bone while degrading at a rate that matches the remodeling of bone. With a degrading material, a second surgery to remove the plate is avoided, which reduces the cost of the procedure and recuperation time for the patient.^[10] Since degradable polymers by themselves do not have sufficient strength to support bone while it heals, composite materials are the next progression.

This article presents an overview of the materials used in degradable composites, the properties and manufacturing of composite materials, clinical uses of these materials, and current limitations of these devices.

DEGRADABLE POLYMERS

Polymer Structure and Properties

Polymers are long chain molecules derived from many repeating units. Typically, a carbon backbone comprised of covalent bonds is along the chain while weak secondary bonds exist between the molecules. When the carbon backbone is made from groups that are hydrolytically unstable, these long chain molecules will degrade when placed in water. Initially, these polymers were thought to be unsuitable for any application, but the need for biomedical devices that degrade over time has led to a variety of applications for these materials.^[11] There are many degradable polymers clinically used in devices today. The most commonly used materials are the class of materials known as poly(α -hydroxy acids), such as poly(lactide) (PLA) and poly(glycolide) (PGA).^[12] These contain an ester linkage that is hydrolytically unstable. Other bonds that are unstable in an aqueous environment include anhydrides, orthoesters, and amides.^[11] PLA and PGA both have U.S. Food and Drug Administration (FDA)

approval and are currently in clinical use for a variety of applications. An excellent review on the properties of many degradable polymers has been written by Middleton and Tipton.^[11] This article focuses on the poly(α -hydroxy acids) and how the structure and properties of these materials affect their mechanical behavior, degradation behavior, and use in polymer composites because they are the most common polymers found in clinical use today. The most important structural characterizations are percent crystallinity and molecular weight.

The percent crystallinity plays an important role in determining the properties of a polymer. With all other factors remaining constant, a polymer with more crystalline regions will be stronger and degrade at a slower rate. Crystalline regions in the polymer have stronger secondary bonds between the polymer chains that make them less susceptible to degradation. PGA is a highly crystalline material (45–55%), and can be made into stiff, high-strength fibers. The crystallinity leads to a reduction of strength of 50% at two weeks, and it completely loses its structural integrity at four weeks.^[11] It can take anywhere from four to six months to degrade completely.

In contrast, PLA can be found as one of two isomers, *D* (PDLA) and *L* (PLLA). These isomers can also be blended into a copolymer poly(DL-lactide) (PDLLA). PDLLA is completely amorphous, degrades too rapidly to be used in orthopedic applications, and has a low tensile strength. The naturally occurring form, poly(L-lactide) (PLLA) is approximately 37% crystalline and requires over two years to be completely resorbed. However, copolymers of PLLA and PGA can speed the degradation time by reducing the crystallinity of the resultant polymer.^[11]

The second factor that has a major impact on the degradation of these polymers is the molecular weight. As the molecular weight of a polymer increases, the polymer is strengthened by the covalent bonds along the chain and the increasing number of entanglements in the structure. The degradation process also slows because more covalent bonds must be broken per chain before the polymer chains can be metabolized. The molecular weight degrades before the strength or mass of the polymer so this can be a method to track the degradation and predict the degradation rate of a new material.^[11]

Degradation Mechanisms

In general, polymers may degrade by a variety of mechanisms, including oxidation, heat, radiation, or chemical exposure. The focus of this article is polymers that degrade in a biological system. When a biodegradable polymer is placed in water, the

hydrolytically weak functional group is susceptible to degradation from the aqueous environment. As the polymer chains become shorter, they become water soluble and can be metabolized. Degradation methods have been well studied for drug delivery applications, and there are two main mechanisms:

- Surface erosion.
- Bulk erosion.

Surface erosion occurs on the surface of the implant and when water is unable to penetrate the device in large concentrations. The device will shrink in size as the water degrades the polymer on the surface. This type of erosion will occur in all devices and particularly devices with a low porosity. It is also more common with the polyanhydrides and polyorthoesters because of their high hydrophobicity.

In bulk erosion, water penetrates the bulk of the device and breaks bonds from within the polymer device. This type of erosion is more common in devices with porosity. Bulk erosion is also more common for polymers based on poly(α -hydroxy acids). As these polymers degrade, lactic or glycolic acid are formed that can then be converted into natural metabolites. If the production of these degradation products is greater than the metabolism of them, the pH of the surrounding area increases. As the local acidic concentration is raised, the degradation process is further accelerated. Clinically, this can result in pain, discomfort, or swelling.^[13] In a porous implant, nutrients and fluid flow may be quite small in a low-porosity implant, which leads to even more acid buildup and more rapidly accelerates the degradation of the polymer.

Biocompatibility

Perhaps the biggest challenge in any biomaterial's application is the design of materials that will integrate with the human body and cause no adverse reactions. The biocompatibility of a material refers to how well the material is accepted in the body.^[1] When a foreign material is implanted in the body, the surrounding tissues are compromised, and the inflammatory response is activated to heal the damage. During this process, proteins and cells adsorb to the surface of the implanted material. The overall response of the body to the material is a complicated reaction to a combination of factors such as material properties, surface topography, and surface chemistry. The vascularity of the implant site also affects the biocompatibility. Well-vascularized sites can transport degradation products more easily than poorly vascularized tissues, which leads to an improved cellular response.^[14] Clinical

responses to degradable implants include irritation at the implant site, a sterile sinus, or osteolysis.^[14] Examples of the biocompatibility of specific devices will be presented subsequently.

COMPOSITE MATERIALS

Materials are generally classified into three categories: polymer, metal, and ceramic. A composite material is a combination of two or more categories or two or more distinct materials from one of the categories. Composite materials are particularly useful in engineering because they typically take advantage of the best properties of each of the individual materials used to create the composite.

The common element in any composite material is the existence of a reinforcing phase surrounded by a matrix. The reinforcing phase could be particulate or fibrous. Fibers possess a high strength and high modulus due to the orientation of the polymer molecules. The matrix can be ductile or brittle and has several primary functions. It acts to transfer the applied load between the fibers, bonds the fibers together, and deflects cracks during fracture. A critical component to the success of a composite material is the bond between the reinforcing phase and the matrix.

This bonding is even more critical in the corrosive environment of the human body. If water can leach into the composite and degrade the bond between fiber and matrix, the results can be disastrous. The first polymer composites intended for orthopedic fracture fixation failed because of this very reason. A partial list of materials under consideration for these plates included carbon fiber-reinforced epoxy resin, polyetheretherketone (PEEK), polysulphone (PS), polybutylene terephthalate (PBT) reinforced with polyacrylonitrile (PAN), carbon fiber-reinforced PS, and chopped carbon fiber-reinforced PEEK.^[5,15–19] The benefits of these materials include less stress shielding in the bone, a modulus that can be tailored from about 1.5 to five times that of bone,^[18] and the elimination of corrosion effects from metallic plates. Despite these improvements, however, a loss of strength and stiffness was noted after in-vitro conditioning or in-vivo use in animal models. Delamination of the fibers from the matrix was noted and a foreign body response to the carbon fibers eliminated these composite materials from in-vivo applications.

Another desirable aspect of composite materials is the ability to tailor the properties of the material more precisely. By adjusting the percentage of reinforcing material, the properties can be varied dramatically. Often, there is an optimum concentration of reinforcing material and deviating from this optimum will yield an inferior material.

Self-Reinforced Composite Materials

A subcategory of composite materials are self-reinforced composites (SRC). These composites consist of a reinforcing fiber and matrix material that are fabricated from the same polymer—hence the term self-reinforced. They are processed by a hot compaction mechanism. They were developed at least in part because of the stringent requirements of the FDA. A common problem in the biomaterials field is the development of new materials to meet new device requirements and determining their biocompatibility. New materials are often developed from existing materials having proven biocompatibility, which eliminates some of the toxicity testing that surrounds a new material.

A schematic of the process commonly used to fabricate SRCs is shown in Fig. 1. Fibers are laid in a metal mold and heated under pressure so that the outer surfaces of the fiber soften and meld together to form the matrix. There is no addition of a matrix material or other bonding agent. Two main events occur during processing:

1. The reinforcing fibers lose their orientation (and strength).
2. The bonding between the fibers, or development of the matrix, increases.

The matrix must be developed so that it is strong enough to bond the fibers together while still maintaining the orientation and integrity of the original fibers. The fiber orientation gives the composite its strength. The optimum balance between the loss of orientation and the development of the matrix yields the best mechanical properties.

Although the focus of this article is on degradable polymer composites, SRCs have been developed for a variety of materials and applications. Ward et al. have used this method with great success with polyethylene (PE),^[20–28] poly(ethylene terephthalate),^[29] and polypropylene fibers.^[30] Wright et al. have used these

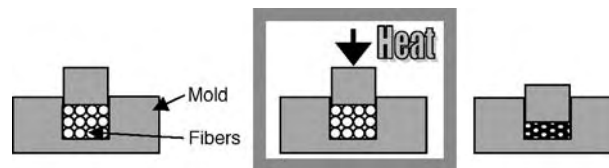


Fig. 1 Schematic of the standard self-reinforced composite processing method. At left, a constant mass of the fibers is placed in a mold. The middle picture illustrates the composite being heated under pressure. At right, the fibers are now joined to each other and considered a composite material.

methods to fabricate SRCs from poly(methyl methacrylate) in an effort to improve the bone cement in total hip replacements.^[31,32] Törmälä and colleagues have performed the most extensive studies with self-reinforced degradable composites,^[10,33–54] and this work will be discussed in detail subsequently.

POLYMER FIBER COMPOSITES

Self-Reinforced Composites

Self-reinforced composites fabricated from PLLA and PGA materials are currently in clinical use and marketed by Linvatec Corporation. They can be fabricated into screws, pins, rods, tacks, wires, and membranes, and can be used for a variety of fractures, especially in the humerus, radius, and craniofacial applications. The manufacturers of these implants have written a number of review articles on the subject of their manufacturing, use, and clinical successes.^[34,35,47,48,52] Table 1 shows a summary of representative mechanical properties for self-reinforced rods and injection-molded controls for PLLA and PGA materials. The self-reinforced composites show a two to three times increase for most mechanical properties when compared to the injection-molded unreinforced material.^[52]

Early work presented the development of the materials and initial in-vitro biocompatibility results.^[40–42,44,46,49,51,54,55] Animal studies have examined their viability in a variety of applications^[41,42,44,49,51] and have shown some success in load-bearing applications for femoral osteotomies in rabbits.^[41] Amorphous composites based on PGA tend to absorb more quickly—about eight weeks—whereas crystalline composites based on PLLA absorb in about 36 weeks.^[52]

Clinical Uses and Experiences

Clinically, these materials have been in use since 1984, and over 300,000 devices have been implanted.^[48] The originators of this device have performed more than 3,200 operations with these devices^[48] so the clinical

results of these devices are well documented in studies by the originators of the device^[36,37,39,45,53,56] and others.^[57,58] Complications have been minimal, but they include the formation of a sterile sinus at the site of the implant. McVicar et al.^[57] reported no adverse effects due to the degradation of a self-reinforced PGA membrane, which may be due to the small quantity of material present in the material. For larger implants, where these effects will be magnified, it is recommended that the smallest implant necessary be used in areas that are well vascularized.^[52] These problems also become more pronounced when particulate debris from the device is present and is more prevalent for PGA implants.^[48,13] Ashammakhi et al.^[35] note that PLA particulate debris is found for 0.1% of orthopedic patients. Since self-reinforced PGA implants degrade more rapidly than self-reinforced PLA implants, Ashammakhi et al.^[35] recommend that these implants be reserved for children because of their increased metabolism in bone tissue.

Limitations and Challenges

Self-reinforced composites have revolutionized the degradable composites available for fracture fixation. Their mechanical properties are improved over unreinforced materials, and their clinical use has shown good biocompatibility with few serious complications. However, these materials are mainly processed using a trial-and-error method. Although optimization is mentioned in some articles, no known attempt has been made to link the properties of the starting fibers with the processing method of the composite. Improving the processing techniques to a greater extent could further improve the strength of these materials. Additionally, in some cases, the formation of a sterile sinus can lead to a second operation to remove the implant, which would negate the benefits of a degradable device. Finally, although these implants are biocompatible in the traditional sense, they are surrounded by a fibrous capsule and do not provide good surfaces to guide bone growth and bone bonding to the implant. Further improvements to the manufacturing, reductions in inflammatory responses, and increased

Table 1 Representative mechanical properties for self-reinforced and injection-molded PLLA and PGA

Polymer	Processing	Bending modulus (GPa)	Bending strength (MPa)	Shear strength (MPa)
PGA	Self-reinforced	13	320	240
PGA	Injection-molded	7	218	95
PLLA	Self-reinforced	10	300	220
PLLA	Injection-molded	3	119	68

From Ref. [52].

bone bonding would further improve degradable self-reinforced composites for future applications.

DEGRADABLE POLYMER-CERAMIC COMPOSITES

A recent advance in the area of degradable polymer composites is the incorporation of ceramic particles into the polymer matrix. This can be achieved for porous scaffolds,^[59–65] nonporous materials,^[66–74] membranes,^[75] and injectable materials.^[76,77] The form that the material takes indicates its intended application. Foams and injectable materials are generally targeted toward bone graft substitutes while solid materials can be fabricated into plates, pins, or screws. Membranes are used to guide tissue growth in maxillofacial applications. Table 2 summarizes the polymers and ceramics used in these studies. Although there is a great variety in the polymers and ceramics chosen, the majority of studies focus on derivatives of poly(α -hydroxy acids).

Ceramic particles yield several benefits for degradable polymer composites. First, they add a considerable amount of strength and stiffness to the polymer matrix. Secondly, degradable ceramics such as hydroxyapatite (HA) and α -tricalcium phosphate (α -TCP) have been shown to be extremely biocompatible with bone. Transmission electron microscopy studies have well documented the interface between ceramic particles and natural bone. In these studies, HA is seen to directly bond to bone whereas TCP bonds to bone through an intermediate apatite layer. In both cases, the new bone formed is indistinguishable from new natural bone.^[78–80] Glass ceramics have also demonstrated these bone bonding properties. Dissolution of the calcium and phosphorus from the glass ceramic mediates the formation of an apatite layer on the material and subsequent bone bonding can occur.^[81,82] Despite these excellent properties, ceramics are too brittle to be used in orthopedic applications. However, polymers lack the necessary strength for these same applications, and the build up



Table 2 Polymers and ceramics used for composite applications

Polymer	Ceramic	Use	Authors
PLA	45S5 bioactive glass	Microspheres	Qui et al. ^[77]
PLLA	α -TCP	Membrane	Ignatius et al. ^[75]
	Calcined or uncalcined HA	Nonporous	Furukawa et al. ^[72]
	HA fibers	Nonporous	Kasuga et al. ^[67]
PLA-co-PGA (85:15)	Carbonated apatite	Scaffold	Murphy et al. ^[64]
PLA-co-PGA (50:50)	Calcium phosphate	Scaffold	Ambrosio et al. ^[59]
PLA-co-PGA (50:50)	HA	Scaffold	Devin et al. ^[60]
Poly(L,DL-lactide)	β -TCP	Nonporous	Ignatius et al. ^[66]
PDLLA	TCP	Nonporous	Lin et al. ^[69]
	α -TCP	Scaffold	Ignatius et al. ^[62]
	Glass ceramic GB14N	Scaffold	Ignatius et al. ^[62]
Poly(ϵ -caprolactone/ L-lactide) (90:10 or 10:90)	Sodium-calcium-aluminum-polyphosphate glass fiber or calcium-sodium-metaphosphate microfiber	Nonporous	Andriano et al. ^[71]
Poly(caprolactone) and PDLLA-co-PGA blends	HA	Scaffold	Marra et al. ^[63]
Poly(ortho ester)	Sodium-calcium-aluminum-polyphosphate glass fiber or calcium-sodium-metaphosphate microfiber	Nonporous	Andriano et al. ^[70,71]
Polyhydroxybutyrate	HA	Nonporous	Knowles et al. ^[74]
	HA and ceramic glass	Nonporous	Knowles et al. ^[74]
Blends of corn starch and ethylene vinyl alcohol	HA	Nonporous	Mendes et al. ^[73]
	HA	Nonporous	Gomes et al. ^[61]
	HA	Nonporous	Vaz et al. ^[68]
Hydroxypropylmethylcellulose (HPMC)	Biphasic calcium phosphate	Injectable	Nguyen et al. ^[76]

of acidic by-products is still a limitation to their use in some applications. The combination of these materials may yield composite materials with properties that match the given need more closely.

Work in this area is relatively new and centers on material development and in-vitro characterization of the resultant materials and apatite formation or cell viability in vitro. Only a few studies have examined the behavior of these materials in an animal model,^[62,69,71–74] and none are in clinical use. However, these materials represent the future of degradable polymer composites, and their development is important to consider in this context.

Polymer-Ceramic Scaffolds

The addition of ceramic particles is highly desirable in three-dimensional scaffold applications. In these applications, the polymer is highly porous, and cell growth is mediated through the interconnected pores. However, these materials lack structural integrity, and, if the degradation products from the polymers cannot be transported out of the scaffold, cell growth will be limited because of the increased acidic environment in the interior of the scaffold. By adding ceramic particles that are also degradable, the strength of the scaffold can be increased, and bone growth can be accelerated or enhanced.

Ignatius et al. has reported use of porous poly(L, DL-lactide) scaffolds with ceramic or glass ceramics in an animal model^[62] as a bone graft substitute. Implants containing either α -TCP or a glass ceramic were implanted in the tibia of sheep, and the bone growth in them was studied at 6, 12, or 24 months. No significant differences were seen between the two ceramic materials, and adverse effects were not reported until 24 months. At 24 months, a significant amount of PLA had degraded, and histological examination found that the bone had osteolytic characteristics. These materials were found to be unsuitable for clinical use primarily because the degradation time of the polymer component did not adequately match the bone healing and ceramic degradation.^[62] The ceramics were completely absorbed at six months whereas the polymer component did not have significant degradation at two years. Their future solutions will seek to better match the degradation time of each component to the healing time of bone.

In another study, Murphy et al. reported the use of a copolymer of PLA and PGA in a 85:15 ratio.^[64] They point to the challenges of bony reconstruction using only a single material. Because the problem is so complex, they have elected to investigate a composite consisting of 85:15 poly(lactide-co-glycolide) with a carbonated apatite layer incorporated in the scaffold. Their hypothesis is that the addition of the apatite

layer prior to cell seeding will enhance bone growth and mineralization within the scaffold. The processing method consisted of one-step solvent casting and particulate leaching. Their work centered on an in-vitro study that characterized the apatite formation in the scaffold and subsequent mechanical properties. Their work showed a continuous apatite formation within the scaffold and a fivefold increase in compressive modulus after 16 days compared to the control samples.^[64] These studies were not verified in an animal model but instead points to the positive early results seen with these scaffolds.

Marra et al.^[63] has approached this challenge with new materials. Blends of two degradable polymers, poly(caprolactone) and poly(D,L-lactic-co-glycolic acid), have been combined with hydroxyapatite (HA). HA has a higher strength than the TCP ceramics considered earlier, but it also has a slower degradation rate. In their work, mechanical studies and in-vitro degradation was studied. Cells were cultured and incubated in physiological conditions. One challenge with all porous scaffolds is promoting growth of cells deep within the scaffold. Their work showed cellular activity and collagen formation 500 μ m within the scaffold, which is an improvement over results seen by others.^[83]

Another novel approach is to fabricate calcium phosphate-PLAGA microspheres—each an individual composite—and sinter the microspheres to obtain a porous disc as done by Ambrosio et al.^[59] Fabrication of the calcium phosphate occurred within the microsphere to promote noncrystalline formation of calcium phosphate that more closely resembles the low crystallinity of natural bone tissue. The size and concentration of the microspheres can be varied to modify the resultant pore size. Composites were fabricated that have a minimum pore size of 100 μ m, and 75% porosity. It is hypothesized that these scaffolds will be able to support cellular growth for bone repair applications.

Nonporous Polymer-Ceramic Materials

Nonporous polymer-ceramic composites are used in applications such as orthopedic pins, screws, or plates. The earliest work in this area was performed by Andriano et al.^[70,71] and examined copolymers of poly(ϵ -caprolactone/L-lactide) and poly(ortho ester) reinforced with three proprietary glass fibers. Two fibers were shown to be nontoxic in preliminary tests, but composites with the sodium-calcium-aluminum-polyphosphate glass fiber caused necrosis when implanted intracutaneously in a rabbit.^[71] This may have been due to the radiation sterilization method employed for these materials.

Positive results have also been seen with PLLA/HA^[72] and PDLA/TCP^[69] composites. Furukawa et al.^[72] compared two composites fabricated with a matrix of PLLA and a reinforcing particle of either calcined or uncalcined HA. These materials were implanted in the tibia of a rabbit, and bone bonding was studied at a period of 8, 16, and 25 weeks after implantation. Bone bonding as measured by a detaching mechanical test was improved by at least an order of magnitude for all composites when compared to controls of PLLA at all time periods. No difference was seen between the experimental composites during the experiment. Direct bone bonding was noted for the PLLA/HA composites while a layer of connective tissue surrounded the PLLA implant. Similar bone bonding results were seen by Lin et al.,^[69] but mechanical tests were not performed on the implants after retrieval. In Lin's study, a two to threefold increase in compressive strength is seen with the addition of TCP, which indicated the mechanical benefit of incorporating ceramic particles in a polymer matrix.

Recent research is also investigating new, starch-based polymers as matrices for HA particles.^[73] Initial biocompatibility assays show good results in intracortical and intramuscular implantation in Dutch milk goats at 12 weeks. Bone contact was seen for both implants, although a fibrous layer was noted in areas where bone was remodeling. No strength data was presented for these materials. New materials may yield an improved match between bone healing rates and polymer degradation rates.

NANOCOMPOSITES

Cutting-edge research is investigating the incorporation of nanometer-sized particles in polymer matrices to control material properties on a much smaller size scale. The primary benefit of these materials is the ability to mimic living tissues in engineered materials in structure and assembly.^[84] Du et al. fabricated a collagen-nanohydroxyapatite that bears resemblance to the organic-inorganic structure of living bone.^[85,86] Their work found that bone remodeling occurred in a similar manner to natural bone in intimate contact with the nanocomposite. Although work is just beginning in this area, the ability to design materials with structures at the nano-scale represents an exciting future for new devices.

CONCLUSION

In 1995, the market for biodegradable medical devices was \$300 billion in the United States alone. Only 5%

of this market was for orthopedic devices whereas 95% was for degradable sutures.^[11] However, the market for orthopedic devices is growing at approximately 6% per year. If new materials, methods, and devices are developed for orthopedic applications utilizing degradable polymers, the potential for market growth in this area is limitless. Degradable polymer composites have the potential to drastically reduce the medical costs associated with these procedures because the need for a second operation to remove the device is eliminated. For example, one British study calculated the cost savings associated with degradable implants to be £608–1011 (approximately US\$ 1000–1700) per patient when compared to metallic implants due to decreased time in the hospital.^[10] The market growth and savings in health care costs are limited by the properties of the polymeric materials and the current technologies used to process them.

It is clear that novel new approaches need to be considered to solve the complex problem of fracture fixation and bone healing. Traditional devices fabricated from metals do not have mechanical properties that match natural bone and require more recuperation time and a higher medical expense. Degradable polymer devices do not have the necessary strength for many applications and are plagued with problems of acidic by-product buildup, which leads to an inflammatory response and loss of bone.

Composite materials, such as the ones presented here, represent the future for these devices. Self-reinforced degradable devices reduce the stress shielding seen in metallic implants and eliminate the need for removal, but they are still not strong enough as fracture fixation plates for load-bearing bones, such as the femur and tibia. These implants also have the complication of a foreign body reaction due to the reaction of the degradation products with the human body and a limited ability to bond directly to bone. Polymer materials reinforced with ceramic particles hold the promise of better mechanical properties and bone bonding ability, but only clinical studies will verify whether these promises become reality.

ACKNOWLEDGMENT

The author gratefully acknowledges the assistance of Ms. Amy Latimer in conducting the literature search for this article.

ARTICLES OF FURTHER INTEREST

Composites; Elastomers, Biodegradable; Host Reactions; Poly(Glycolic Acid); Poly(lactic acid)s; Polymers

REFERENCES

1. Ratner, B.D.; Hoffman, A.S.; Scheon, F.J.; Lemons, J.E. *Biomaterials Science: An Introduction to Materials in Medicine*; Academic Press: San Diego, 1996.
2. Whang, K.; Goldstick, T.K.; Healy, K. A biodegradable polymer scaffold for delivery of osteotropic factors. *Biomaterials* **2000**, *21*, 2545–2551.
3. Evans, G.; Brandt, K.; Niederbichler, A.; Chauvin, P.; Hermann, S.; Bogle, M.; Otta, L.; Wang, B.; Patrick, C.W., Jr. Clinical long-term in vivo evaluation of poly(L-lactide acid) porous conduits for peripheral nerve regeneration. *J. Biomater. Sci., Polym. Ed.* **2000**, *11* (8), 869–878.
4. Sarmiento, A.; Mullis, D.L.; Latta, L.L.; Tarr, R.R.; Alvarez, R. A quantitative comparative analysis of fracture healing under the influence of compression plating vs. closed weight-bearing treatment. *Clin. Orthop.* **1980**, *149*, 232–239.
5. Claes, L. The mechanical and morphological properties of bone beneath internal fixation plates of differing rigidity. *J. Orthop. Res.* **1989**, *7* (2), 170–177.
6. Beaupré, G.S.; Csongradi, J.J. Refracture risk after plate removal in the forearm. *J. Orthop. Trauma* **1996**, *10* (2), 87–92.
7. Böstman, O.M. Refracture after removal of a condylar plate from the distal third of the femur. *J. Bone Jt. Surg.* **1990**, *72* (7), 1013–1018.
8. McAuliffe, J.A. Forearm fixation. *Hand Clin.* **1997**, *13* (4), 689–701.
9. Lovell, M.E.; Galasko, C.S.; Wright, N.B. Removal of orthopedic implants in children: Morbidity and post-operative radiologic changes. *J. Pediatr. Orthop., Part B* **1999**, *8* (2), 144–146.
10. Juutilainen, T.; Päätiälä, H.; Ruuskanen, M.; Rokkanen, P. Comparison of costs in ankle fractures treated with absorbable or metallic fixation devices. *Arch. Orthop. Trauma. Surg.* **1997**, *116* (4), 204–208.
11. Middleton, J.C.; Tipton, A.J. Synthetic biodegradable polymers as orthopedic devices. *Biomaterials* **2000**, *21*, 2335–2346.
12. Hirashima, Y.; Ishiguro, N.; Kondo, S.; Iwata, H. Osteoclast induction from bone marrow cells is due to pro-inflammatory mediators from macrophages exposed to polyethylene particles: A possible mechanism of osteolysis in failed THA. *J. Biomed. Mater. Res., Appl. Biomater.* **2001**, *56*, 177–183.
13. Böstman, O.M. Intense granulomatous inflammatory lesions associated with absorbable internal fixation devices made of polyglycolide in ankle fractures. *Clin. Orthop.* **1992**, *278*, 193–199.
14. Böstman, O.; Pihlajamäki, H.K. Clinical biocompatibility of biodegradable orthopaedic implants for internal fixation: A review. *Biomaterials* **2000**, *21*, 2615–2621.
15. Pemberton, D.J.; McKibbin, B.; Savage, R.; Tayton, K.; Stuart, D. Carbon-fibre reinforced plates for problem fractures. *J. Bone Jt. Surg.* **1992**, *74* (1), 88–92.
16. Tayton, K.; Johnson-Nurse, C.; McKibbin, B.; Bradley, J.; Hastings, G. The use of semi-rigid carbon-fibre-reinforced plastic plates for fixation of human fractures. Results of preliminary trials. *J. Bone Jt. Surg.* **1982**, *64* (1), 105–111.
17. Ali, M.S.; French, T.A.; Hastings, G.W.; Rae, T.; Rushton, N.; Ross, E.R.; Wynn-Jones, C.H. Carbon fibre composite bone plates. Development, evaluation and early clinical experience. *J. Bone Jt. Surg.* **1990**, *72* (4), 586–591.
18. Brown, S.A.; Hastings, R.S.; Mason, J.J.; Moet, A. Characterization of short-fibre reinforced thermoplastics for fracture fixation devices. *Biomaterials* **1990**, *11* (8), 541–547.
19. Jockisch, K.A.; Brown, S.A.; Bauer, T.W.; Merritt, K. Biological response to chopped-carbon-fiber-reinforced peek. *J. Biomed. Mater. Res.* **1992**, *26* (2), 133–146.
20. Bonner, M.J.; Hine, P.J.; Ward, I.M. Hot compaction of crosslinked high modulus polyethylene fibers and fabrics. *Plast. Rubber Compos. Process. Technol.* **1998**, *27* (2), 58–64.
21. Hine, P.J.; Ward, I.M.; Olley, R.H.; Bassett, D.C. The hot compaction of high modulus melt-spun polyethylene fibres. *J. Mater. Sci.* **1993**, *28*, 316–324.
22. Kabeel, M.A.; Bassett, D.C.; Olley, R.H.; Hine, P.J.; Ward, I.M. Compaction of high-modulus melt-spun polyethylene fibres at temperatures above and below the optimum. *J. Mater. Sci.* **1994**, *29*, 4694–4699.
23. Kabeel, M.A.; Bassett, D.C.; Olley, R.H.; Hine, P.J.; Ward, I.M. Differential melting in compacted high-modulus melt-spun polyethylene fibres. *J. Mater. Sci.* **1995**, *30*, 601–606.
24. Morye, S.S.; Hine, P.J.; Duckett, R.A.; Carr, D.J.; Ward, I.M. Comparison of the properties of hot compacted gel-spun polyethylene fibre composites with conventional gel-spun polyethylene fibre composites. *Compos., Part A Appl. Sci. Manuf.* **1999**, *30* (5), 649–660.
25. Olley, R.H.; Bassett, D.C.; Hine, P.J.; Ward, I.M. Morphology of compacted polyethylene fibres. *J. Mater. Sci.* **1993**, *28*, 1107–1112.
26. Tissington, B.; Pollard, G.; Ward, I.M. A study of the influence of fibre/resin adhesion on the mechanical behaviour of ultra-high-modulus polyethylene fibre composites. *J. Mater. Sci.* **1991**, *26*, 82–92.
27. Ward, I.M. A review of recent developments in the processing and properties of oriented polyethylene. *Plast. Rubber Compos. Process. Appl.* **1993**, *19*, 7–13.
28. Woods, D.W.; Ward, I.M. Study of the interlaminar shear strength of unidirectional high-modulus polyethylene fibre composites. *Polymer* **1994**, *29*, 2572–2578.
29. Rasburn, J.; Hine, P.J.; Ward, I.M.; Olley, R.H.; Bassett, D.C.; Kabeel, M.A. The hot compaction of polyethylene terephthalate. *J. Mater. Sci.* **1995**, *30*, 615–622.
30. Abo El-Maaty, M.I.; Bassett, D.C.; Olley, R.H.; Hine, P.J.; Ward, I.M. The hot compaction of polypropylene fibres. *J. Mater. Sci.* **1996**, *31*, 1157–1163.
31. Wright, D.D.; Lautenschlager, E.P.; Gilbert, J.L. Interfacial properties of self-reinforced composite poly(methyl methacrylate). *J. Biomed. Mater. Res.* **1998**, *43* (2), 153–161.
32. Wright, D.D.; Lautenschlager, E.P.; Gilbert, J.L. Bending and fracture toughness of woven self-reinforced

- composite poly(methyl methacrylate). *J. Biomed. Mater. Res.* **1997**, *36* (4), 441–453.
33. Andriano, K.P.; Pohjonen, T.; Törmälä, P. Processing and characterization of absorbable polylactide polymers for use in surgical implants. *J. Appl. Biomater.* **1994**, *5*, 133–140.
 34. Ashammakhi, N.; Rokkanen, P. Absorbable polyglycolide devices in trauma and bone surgery. *Biomaterials* **1997**, *18* (1), 3–9.
 35. Ashammakhi, N.; Peltoniemi, H.; Waris, E.; Suuronen, R.; Serlo, W.; Kellomäki, M.; Törmälä, P.; Waris, T. Developments in craniomaxillofacial surgery: Use of self-reinforced bioabsorbable osteofixation devices. *Plast. Reconstr. Surg.* **2001**, *108* (1), 167–180.
 36. Hirvensalo, E.; Böstman, O.; Törmälä, P.; Vainionpää, S.; Rokkanen, P. Chevron osteotomy fixed with absorbable polyglycolide pins. *Foot Ankle* **1991**, *11* (4), 212–218.
 37. Jukkala-Partio, K.; Partio, E.K.; Hirvensalo, E.; Rokkanen, P. Absorbable fixation of femoral head fractures. A prospective study of six cases. *Ann. Chir. Gynaecol.* **1998**, *87* (1), 44–48.
 38. Jukkala-Partio, K.; Laitinen, O.; Partio, E.K.; Vasenius, J.; Vainionpää, S.; Pohjonen, T.; Törmälä, P.; Rokkanen, P. Comparison of the fixation of subcapital femoral neck osteotomies with absorbable self-reinforced poly-L-lactide lag-screws or metallic screws in sheep. *J. Orthop. Res.* **1997**, *15* (1), 124–127.
 39. Juutilainen, T.; Hirvensalo, E.; Majola, A.; Partio, E.K.; Pätiälä, H.; Rokkanen, P.; Kinnunen, J. Bone mineral density in fractures treated with absorbable or metallic implants. *Ann. Chir. Gynaecol.* **1997**, *86* (1), 51–55.
 40. Kellomäki, M.; Paasimaa, S.; Törmälä, P. Pliable polylactide plates for guided bone regeneration: Manufacturing and in vitro. *Proc. Inst. Mech. Eng., Part H J. Eng. Med.* **2000**, *214* (6), 615–629.
 41. Koskikare, K.; Hirvensalo, E.; Patiala, H.; Rokkanen, P.; Pohjonen, T.; Törmälä, P.; Lob, G. Intraosseous plating with absorbable self-reinforced poly-L-lactide plates in the fixation of distal femoral osteotomies on rabbits. *J. Biomed. Mater. Res.* **1996**, *30* (4), 417–421.
 42. Majola, A.; Vainionpää, S.; Vihtonen, K.; Mero, M.; Vasenius, J.; Törmälä, P.; Rokkanen, P. Absorption, biocompatibility, and fixation properties of polylactic acid in bone tissue: An experimental study in rats. *Clin. Orthop. Relat. Res.* **1991**, *268*, 260–269.
 43. Mäkelä, P.A.; Ruuskanen, M.; Ashammakhi, N.; Kallioinen, M.; Pohjonen, T.; Serlo, W.; Törmälä, P.; Waris, T. Comparison of self-reinforced poly-L-lactide and steel wire in fixation of sternotomy in rabbits. *Ann. Chir. Gynaecol.* **1999**, *88* (4), 318–321.
 44. Niiranen, H.; Törmälä, P. Bioabsorbable polymer plates coated with bioactive glass spheres. *J. Mater. Sci., Mater. Med.* **1999**, *10* (12), 707–710.
 45. Partio, E.K.; Hirvensalo, E.; Böstman, O.; Rokkanen, P. A prospective controlled trial of the fracture of the humeral medial epicondyle—how to treat? *Ann. Chir. Gynaecol.* **1996**, *85* (1), 67–71.
 46. Peltoniemi, H.H.; Tulamo, R.M.; Toivonen, T.; Hallikainen, D.; Törmälä, P.; Waris, T. Biodegradable semirigid plate and miniscrew fixation compared with rigid titanium fixation in experimental calvarial osteotomy. *J. Neurosurg.* **1999**, *90* (5), 910–917.
 47. Rokkanen, P. Current clinical use of absorbable fracture fixation devices. *Ann. Chir. Gynaecol.* **1991**, *80* (3), 243–244.
 48. Rokkanen, P.U.; Böstman, O.M.; Hirvensalo, E.; Makela, E.A.; Partio, E.K.; Patiala, H.; Vainionpää, S.; Vihtonen, K.; Törmälä, P. Bioabsorbable fixation in orthopaedic surgery and traumatology. *Biomaterials* **2000**, *21* (24), 2607–2613.
 49. Saikku-Backstrom, A.; Tulamo, R.M.; Pohjonen, T.; Törmälä, P.; Raiha, J.E.; Rokkanen, P. Material properties of absorbable self-reinforced fibrillated poly-96L/4 D-lactide (SR-PLA96) rods; A study in vitro and in vivo. *J. Mater. Sci., Mater. Med.* **1999**, *10* (1), 1–8.
 50. Sinisaari, I.; Pätiälä, H.; Böstman, O.; Mäkelä, E.A.; Hirvensalo, E.; Partio, E.K.; Törmälä, P.; Rokkanen, P. Metallic or absorbable implants for ankle fractures: A comparative study of infections in 3,111 cases. *Acta Orthop. Scand.* **1996**, *67* (1), 16–18.
 51. Törmälä, P.; Vasenius, J.; Vainionpää, S.; Laiho, J.; Pohjonen, T.; Rokkanen, P. Ultra-high-strength absorbable self-reinforced polyglycolide (SR-PGA) composite rods for internal fixation of bone fractures: In vitro and in vivo study. *J. Biomed. Mater. Res.* **1991**, *25* (1), 1–22.
 52. Törmälä, P. Biodegradable self-reinforced composite materials; Manufacturing structure and mechanical properties. *Clin. Mater.* **1992**, *10*, 29–34.
 53. Tuompo, P.; Partio, E.; Rokkanen, P. Bioabsorbable fixation in the treatment of proximal tibial osteotomies and fractures. A clinical study. *Ann. Chir. Gynaecol.* **1999**, *88* (1), 66–72.
 54. Vasenius, J.; Vainionpää, S.; Vihtonen, K.; Mero, M.; Makela, A.; Törmälä, P.; Rokkanen, P. Histomorphological study on self-reinforced polyglycolide (SR-PGA) osteosynthesis implants coated with slowly absorbable polymers. *J. Biomed. Mater. Res.* **1990**, *24* (12), 1615–1635.
 55. Ferguson, S.; Wahl, D.; Gogolewski, S. Enhancement of the mechanical properties of polylactides by solid-state extrusion. II. Poly(L-lactide), poly(L/D-lactide), and poly(L/DL-lactide). *J. Biomed. Mater. Res.* **1996**, *30*, 543–551.
 56. Serlo, W.; Kaarela, O.I.; Peltoniemi, H.H.; Merikanto, J.; Ashammakhi, N.A.; Lassila, K. Use of self-reinforced polylactide osteosynthesis devices in craniofacial surgery: A long-term follow-up study. *Scand. J. Plast. Reconstr. Surg. Hand Surg.* **2001**, *35* (3), 285–292.
 57. McVicar, I.; Hatton, P.V.; Brook, I.M. Self-reinforced polyglycolic acid membrane: A bioresorbable material for orbital floor repair. Initial clinical report. *Br. J. Oral Maxillofac. Surg.* **1995**, *33* (4), 220–223.
 58. Pakkanen, M.; Salisbury, A.V.; Ersek, R.A. Biodegradable positive fixation for the endoscopic brow lift. *Plast. Reconstr. Surg.* **1996**, *98* (6), 1087–1091.
 59. Ambrosio, A.M.; Sahota, J.S.; Khan, Y.; Laurencin, C.T. A novel amorphous calcium phosphate polymer

- ceramic for bone repair: I. Synthesis and characterization. *J. Biomed. Mater. Res.* **2001**, *58* (3), 295–301.
60. Devin, J.E.; Attawia, M.A.; Laurencin, C.T. Three-dimensional degradable porous polymer-ceramic matrices for use in bone repair. *J. Biomater. Sci., Polym. Ed.* **1996**, *7* (8), 661–669.
 61. Gomes, M.E.; Reis, R.L.; Cunha, A.M.; Blitterswijk, C.A.; de Bruijn, J.D. Cytocompatibility and response of osteoblastic-like cells to starch-based polymers: Effect of several additives and processing conditions. *Biomaterials* **2001**, *22* (13), 1911–1917.
 62. Ignatius, A.A.; Betz, O.; Augat, P.; Claes, L.E. In vivo investigations on composites made of resorbable ceramics and poly(lactide) used as bone graft substitutes. *J. Biomed. Mater. Res.* **2001**, *58* (6), 701–709.
 63. Marra, K.G.; Szem, J.W.; Kumta, P.N.; DiMilla, P.A.; Weiss, L.E. In vitro analysis of biodegradable polymer blend/hydroxyapatite composites for bone tissue engineering. *J. Biomed. Mater. Res.* **1999**, *47* (3), 324–335.
 64. Murphy, W.L.; Kohn, D.H.; Mooney, D.J. Growth of continuous bonelike mineral within porous poly(lactide-co-glycolide) scaffolds in vitro. *J. Biomed. Mater. Res.* **2000**, *50* (1), 50–58.
 65. Schliephake, H.; Kage, T. Enhancement of bone regeneration using resorbable ceramics and a polymer-ceramic composite material. *J. Biomed. Mater. Res.* **2001**, *56* (1), 128–136.
 66. Ignatius, A.A.; Augat, P.; Claes, L.E. Degradation behavior of composite pins made of tricalcium phosphate and poly(L,D,L-lactide). *J. Biomater. Sci., Polym. Ed.* **2001**, *12* (2), 185–194.
 67. Kasuga, T.; Ota, Y.; Nogami, M.; Abe, Y. Preparation and mechanical properties of polylactic acid composites containing hydroxyapatite fibers. *Biomaterials* **2001**, *22* (1), 19–23.
 68. Vaz, C.M.; Reis, R.L.; Cunha, A.M. Use of coupling agents to enhance the interfacial interactions in starch-EVOH/hydroxylapatite composites. *Biomaterials* **2002**, *23* (2), 629–635.
 69. Lin, F.H.; Chen, T.M.; Lin, C.P.; Lee, C.J. The merit of sintered PDLLA/TCP composites in management of bone fracture internal fixation. *Artif. Organs* **1999**, *23* (2), 186–194.
 70. Andriano, K.P.; Daniels, A.U.; Heller, J. Biocompatibility and mechanical properties of a totally absorbable composite material for orthopaedic fixation devices. *J. Appl. Biomater.* **1992**, *3* (3), 197–206.
 71. Andriano, K.P.; Daniels, A.U.; Smutz, W.P.; Wyatt, R.W. Preliminary biocompatibility screening of several biodegradable phosphate fiber reinforced polymers. *J. Appl. Biomater.* **1993**, *4* (1), 1–12.
 72. Furukawa, T.; Matsusue, Y.; Yasunaga, T.; Nakagawa, Y.; Shikinami, Y.; Okuno, M.; Nakamura, T. Bone bonding ability of a new biodegradable composite for internal fixation of bone fractures. *Clin. Orthop.* **2000**, *379*, 247–258.
 73. Mendes, S.C.; Reis, R.L.; Bovell, Y.P.; Cunha, A.M.; van Blitterswijk, C.A.; de Bruijn, J.D. Biocompatibility testing of novel starch-based materials with potential application in orthopaedic surgery: A preliminary study. *Biomaterials* **2001**, *22* (14), 2057–2064.
 74. Knowles, J.C.; Hastings, G.W.; Ohta, H.; Niwa, S.; Boeree, N. Development of a degradable composite for orthopaedic use: In vivo biomechanical and histological evaluation of two bioactive degradable composites based on the polyhydroxybutyrate polymer. *Biomaterials* **1992**, *13* (8), 491–496.
 75. Ignatius, A.A.; Ohnmacht, M.; Claes, L.E.; Kreidler, J.; Palm, F. A composite polymer/tricalcium phosphate membrane for guided bone regeneration in maxillofacial surgery. *J. Biomed. Mater. Res.* **2001**, *58* (5), 564–569.
 76. Nguyen, T.P.; Dupraz, A. Spectroscopic studies of a multiphasic polymer-ceramic mixture material. *J. Biomater. Sci., Polym. Ed.* **1996**, *8* (2), 141–149.
 77. Qiu, Q.Q.; Ducheyne, P.; Ayyaswamy, P.S. New bioactive, degradable composite microspheres as tissue engineering substrates. *J. Biomed. Mater. Res.* **2000**, *52* (1), 66–76.
 78. Tracy, B.M.; Doremus, R.H. Direct electron microscopy studies of the bone-hydroxylapatite interface. *J. Biomed. Mater. Res.* **1984**, *18*, 719–726.
 79. Neo, M.; Nakamura, T.; Ohtsuki, C.; Kokubo, T.; Yamamuro, T. Apatite formation on three kinds of bioactive material at an early stage in vivo. A comparative study by transmission electron microscopy. *J. Biomed. Mater. Res.* **1993**, *27*, 999–1006.
 80. Neo, M.; Kotani, S.; Nakamura, T.; Yamamuro, T.; Ohtsuki, C.; Kokubo, T.; Bando, Y. A comparative study of ultrastructures of the interfaces between four kinds of surface-active ceramic and bone. *J. Biomed. Mater. Res.* **1992**, *26*, 1419–1432.
 81. Kangasniemi, I.M.; Vähätalo, K.; Happonen, R.P.; Yli-Urpo, A.; de Groot, K. In vivo reactions of Ca,P particle containing surface reactive glasses. *J. Biomed. Mater. Res.* **1994**, *28* (9), 993–1002.
 82. Hench, L.L. Bioactive ceramics. *Ann. N.Y. Acad. Sci.* **1988**, *523*, 54–71.
 83. Ishaug, S.L.; Crane, G.M.; Miller, M.J.; Yasko, A.W.; Yaszemski, M.J.; Mikos, A.G. Bone formation by three-dimensional stromal osteoblast culture in biodegradable polymer scaffolds. *J. Biomed. Mater. Res.* **1997**, *36* (1), 17–28.
 84. Cui, D.; Gao, H.A.B. Advance and prospect of bionanomaterials. *Biotechnol. Prog.* **2003**, *19* (3), 683–692.
 85. Du, C.; Cui, F.Z.; Zhu, X.D.; de Groot, K.A.D.o. Three-dimensional nano-HAp/collagen matrix loading with osteogenic cells in organ culture. *J. Biomed. Mater. Res.* **1999**, *44* (4), 407–415.
 86. Du, C.; Cui, F.Z.; Feng, Q.L.; Zhu, X.D.; de Groot, K.A.D.o. Tissue response to nano-hydroxyapatite/collagen composite implants in marrow cavity. *J. Biomed. Mater. Res.* **1998**, *42* (4), 540–548.

Dendrimers for Drug Delivery

Hu Yang

Department of Biomedical Engineering, Virginia Commonwealth University, Richmond, Virginia, U.S.A.

W. John Kao

School of Pharmacy, University of Wisconsin-Madison, and Department of Biomedical Engineering, College of Engineering, University of Wisconsin-Madison, Madison, Wisconsin, U.S.A.

INTRODUCTION

Dendrimers are a unique class of macromolecules, having a highly branched, three-dimensional architecture with very low polydispersity and high functionality (Fig. 1). A typical dendrimer molecule has an initiator core, several interior layers composed of repeating units, and multiple active terminal groups (Fig. 2). The diameter of a polyamidoamine (PAMAM) dendrimer ranges from 1.5 to 14.5 nm. Dendrimers such as PAMAM dendrimers have attracted growing attention in engineering, materials science, chemistry, and biological science, because they have unparalleled properties over traditional polymers and because a variety of commercialized dendrimers allow researchers to easily construct nanoscale devices. Dendrimers are recognized as the most versatile, compositionally, and structurally controlled nanoscale building blocks. Thus, dendrimers provide critically needed nanoscale starting building blocks for the development of high specialty materials such as light-harvesting agents,^[1] chemical sensors,^[2] catalysts,^[3] and cross-link agents.^[4] Dendrimers have also been investigated in the biomedical field for drug delivery,^[5,6] gene therapy,^[7] and imaging contrast agent delivery.^[8] Several products using dendrimer as platform have been developed and commercialized upon the approval of the FDA. Among them, VivaGelTM (Starpharma, Melbourne, Australia) is a topical microbicide to prevent the transmission of HIV and other sexually transmitted diseases. SuperFect[®] (Qiagen, Valencia, CA) is used for gene transfection of a broad range of cell lines. Alert TicketTM (US Army Research Laboratory, Adelphi, MD) is an anthrax-detection agent. Stratus[®] CS (Dade Behring, Deerfield, IL) is used for cardiac marker diagnostic.

The pharmaceutical and biomedical applications of dendrimers, with an emphasis on polymeric drug delivery, are mainly reviewed. Polymeric drug delivery technologies help improve the efficiency of drug delivery and the biopharmaceutic and pharmacokinetic property of drugs. In a polymeric drug delivery system,

drugs of interest can be either attached to the polymer backbone via a permanent/cleavable bond or physically encased inside the core of a polymeric micelle. According to Langer, polymeric drug delivery systems are expected to continuously maintain drug levels in a therapeutically desirable range, reduce harmful side effects, decrease amount of drug needed, decrease number of dosages and possibly less invasive dosing, improve patient compliance with the prescribed drug regimen, and facilitate drug administration for pharmaceuticals with short in vivo half-lives (e.g., proteins and peptides).^[9] If a polymer backbone has additional sites, these additional sites will be able to accommodate targeting moieties to direct the drug carrier to a desired biological site and allow the modification of the polymer backbone with other functional groups to change the solubility of the system. For example, polar groups can be incorporated to increase the solubility in aqueous phase, or nonpolar groups can be incorporated to enhance the solubility in lipid regions. The unique structure makes dendrimers an ideal building block to create a polymeric drug delivery system with multiple functionalities, which is difficult to be achieved with linear polymers (Fig. 3).

DENDRIMER SYNTHESIS

Dendrimers were introduced by Tomalia^[10] and Newkome^[11] in the early 1980s. The synthesis of early dendrimers was based on the divergent method.^[10,11] In each generation-adding step, the molecular weight and the diameter of the dendrimer were nearly doubled. For the synthesis of an ammonia core dendrimer, a two-step process, exhaustive Michael addition to a suitable amine initiator core with methyl acrylate and extensive amidation of the resulting esters with a large excess of ethylenediamine, is required. The so-called de Gennes dense packing effect is such that the end groups on the surface of the globule become more densely packed as a dendritic molecule grows larger, and eventually the dendrimer reaches its upper

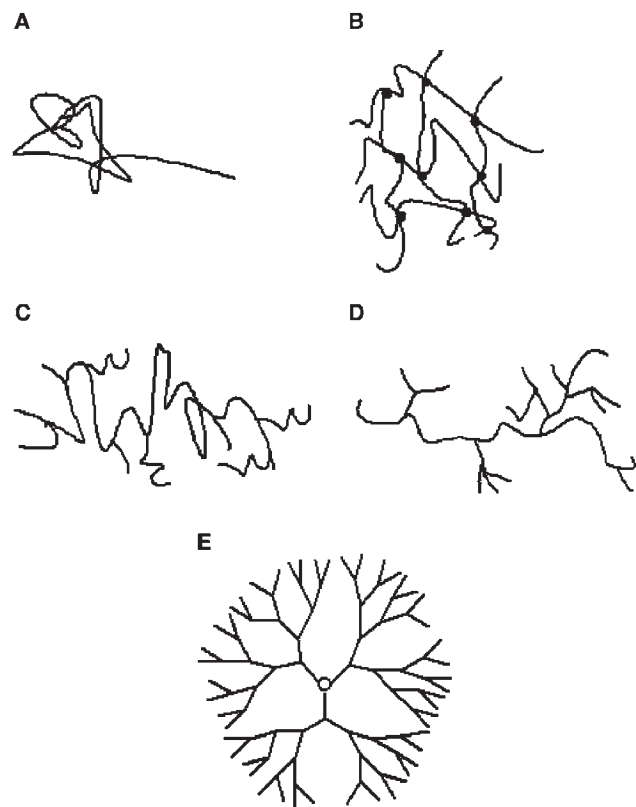


Fig. 1 Simplified architecture of four major classes of polymers: (A), linear; (B), cross-linked; (C), branched; (D), hyperbranched; and (E), dendritic.

generation limit.^[12] Ethylene diamine (EDA)-cored PAMAM dendrimers can only reach up to generation 10 owing to the de Gennes dense packing effect. The divergent method requires tens and even hundreds of simultaneous reactions and produces imperfect samples because of incomplete growth steps and side reactions. In addition, currently available technologies are not advanced enough to isolate pure divergently grown dendrimers from side products of similar structures. Frechet and coworkers^[13] developed the convergent method as a good alternative to the divergent method. Opposite to the divergent method, the reaction begins at the surface of a future dendrimer molecule and proceeds to the core, only requiring two simultaneous reactions to produce next generation of the dendrimer. This method allows intermediate purification, and results in the production of a nearly perfect dendrimer structure. In addition to the divergent and convergent methods, double-stage convergent growth approach,^[14,15] hypercore or branched monomer approach,^[16] double-exponential dendrimer growth approach,^[17] orthogonal coupling,^[18] and self-assembly strategy^[19] have also been developed for reducing synthesis and purification steps and to increase yield. A variety of dendrimers such as poly(propyleneimine) (PPI) dendrimers,^[20] polyesters,^[21]

polyamides,^[22] phosphorus dendrimers,^[23] thiophosphate dendrimers,^[24] melamine dendrimer,^[25] siloxane dendrimers,^[26] and peptide dendrimers^[27] can be obtained by using different initiator cores, repeat units, and surface groups.

DENDRIMER-BASED DRUG DELIVERY SYSTEMS

The delivery of drugs relies on physical or chemical bonding of drugs to the carrier. The dendrimer-based drug delivery systems are discussed in the following sections.

Physical Encapsulation and Bonding

Dendritic box

Meijer and coworkers were the first to entrap guest molecules in a dendritic molecule by using tert-butyloxycarbonyl (t-BOC) groups to produce bulky, sterically crowded amino acid derivatives on the PPI dendrimer surface (Fig. 4). Owing to a densely packed shell on the dendrimer surface, several Rose Bengal dye molecules were stably entrapped inside the dendrimer core, and the guest molecules did not diffuse out of the box^[28] even after prolonged heating or sonication. The Rose Bengal molecules can be liberated by acid hydrolyzing the outer shell to yield the original flexible dendrimer.

Unimolecular dendritic micelles

Polymeric micelles, made of amphiphilic block copolymers, are investigated for drug encapsulation and delivery, because the inner core created by the aggregation of hydrophobic segments can accommodate hydrophobic drugs. The critical micelle concentration (CMC) of polymeric micelles is much lower than that of conventional low-molecular-weight micelles; however, the polymeric micelles are still temperature and concentration dependent. Dendritic unimolecular micelle is another delivery system that can entrap drugs inside. Unimolecular micelles have amphiphilic characteristics, and one single micelle molecule behaves similarly to an assembly of amphiphilic molecules in an aqueous environment. Different from conventional micelles, unimolecular micelles are independent of concentration and temperature.

Dendritic unimolecular micelles can be built through PEGylation on the dendrimer surface (Fig. 5).^[11,29–32] The hydrophobic core of a dendrimer can hold guest molecules and increase the water solubility of hydrophobic drugs.^[33] When polyethylene glycol (PEG) is conjugated to the dendrimer, it helps stabilize

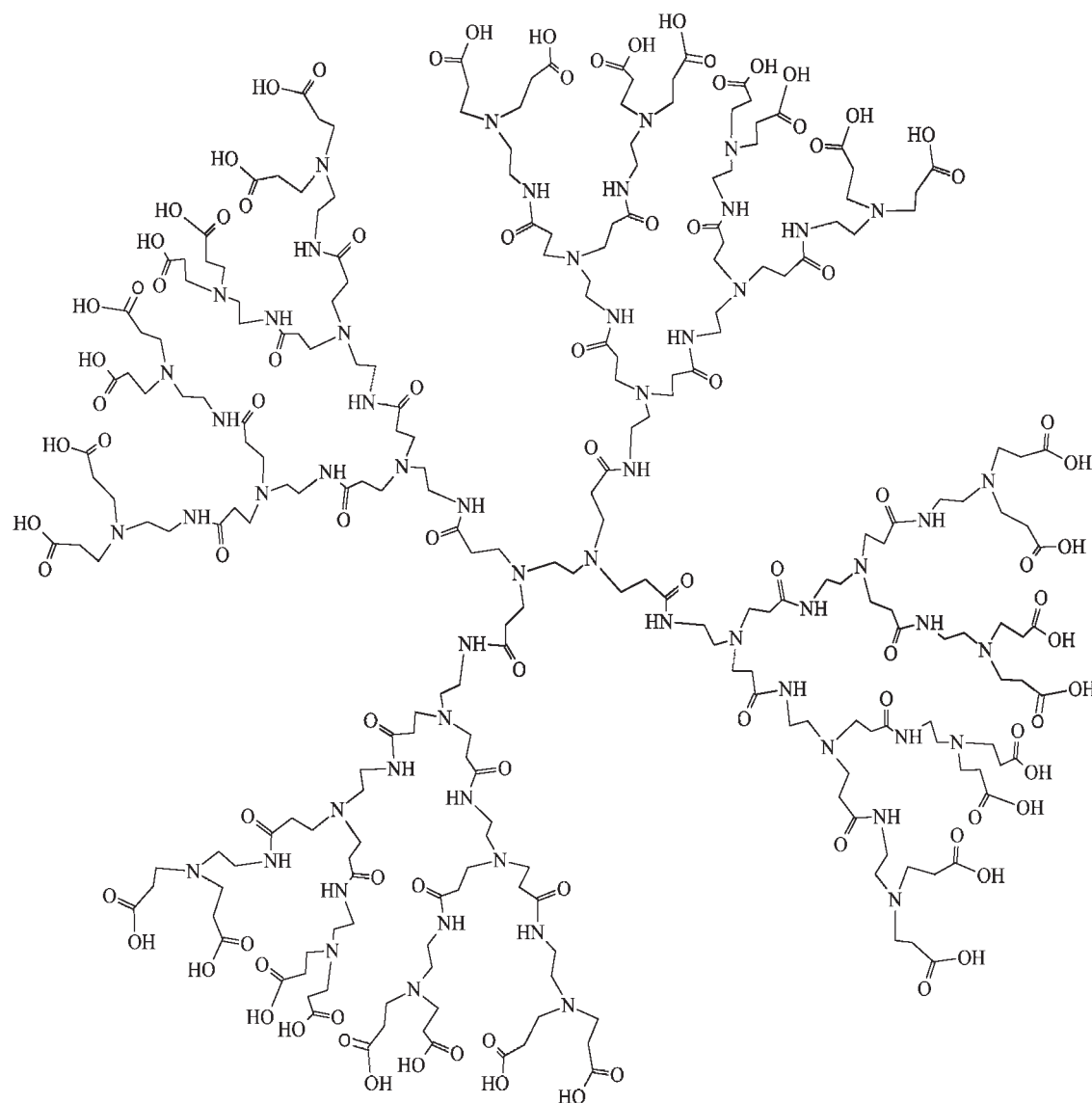


Fig. 2 Schematic structure of a G2.5 PAMAM dendrimer molecule (generally, commercial carboxylate dendrimer is in the carboxylate sodium salt form which is stored in methanol).

entrapped drugs and provide additional sites to retain more drugs. Low generation PAMAM dendrimers have desired biological properties such as nontoxicity and nonimmunogenicity for *in vivo* applications.^[34] However, they are more flexible than higher generations (generation 5 or higher), which may allow some of the surface sites to fold inward, thus reducing their accessibility for reaction and the volume of the inner space.^[35,36] Therefore, low generation dendrimers do not entrap the same amount of guest molecules as high generation dendrimers.^[33] Dendrimer generation and PEG's arm length and arm number at the surface impact the encapsulation of drugs. In the case of the encapsulation of anticancer drugs within PEGylated PAMAM dendrimers done by Kojima et al.,^[37] dendritic micelles with a higher generation dendrimer core

(G4.0 as opposed to G3.0) and longer PEG arms (average molecular weight 2000 Da as opposed to 550 Da) achieved high drug encapsulation. Polyethylene glycol arms help not only the stability of the entrapped molecules but also the solubility of hydrophobic molecules within the PEG layers.^[37,38] Yang, Morris, and Lopina^[39] found that low generation such as G3.0 PAMAM-based dendritic micelles can still enhance the water solubility of hydrophobic compounds, and PEG arm length can be varied to control the extent of the solubility. The scaffold size and flexibility are important factors considering their drug delivery applications. The number of PEG arms and their molecular weight on the surface should be controlled to avoid a decrease in the flexibility of the system. In addition, too many PEG arms would result

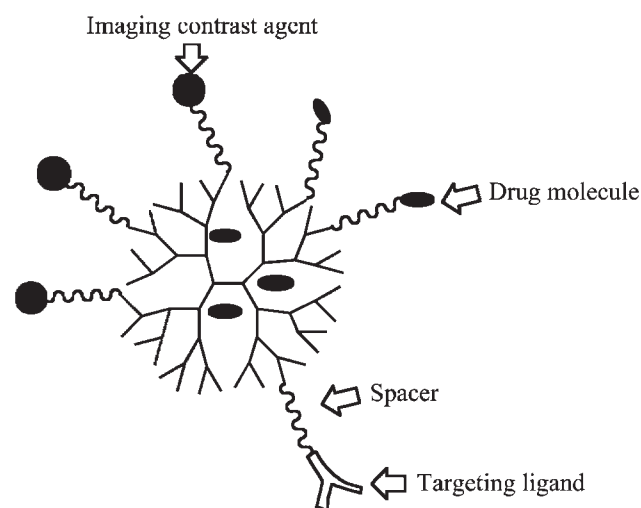


Fig. 3 Schematic drawing of an ideal dendrimer-based multifunctionalized drug delivery system.

in steric hindrance, which reduces further coupling other small molecules of interest to the dendrimer surface. Therefore, for practical *in vivo* applications, the use of high generation or densely PEGylated dendrimers was deliberately avoided.

Dendritic hydrogels

Multiple surface-ending groups make dendritic macromolecules ideal cross-linking agents to obtain dendritic cross-linked hydrogels. The number of the junctions can be controlled by varying the generation of the dendrimer. Gitsov and Zhu^[40] synthesized amphiphilic hydrogels with highly shape persistent cross-linking junctions composed of linear blocks such as PEG and perfectly branched dendritic macromolecules and found that the relative PEG content, the polarity of the medium, and the temperature affected the swelling of the gels. The incorporation of PAMAM dendrimers into poly(vinyl alcohol) (PVA) can lead to the formation of physically cross-linked PVA/

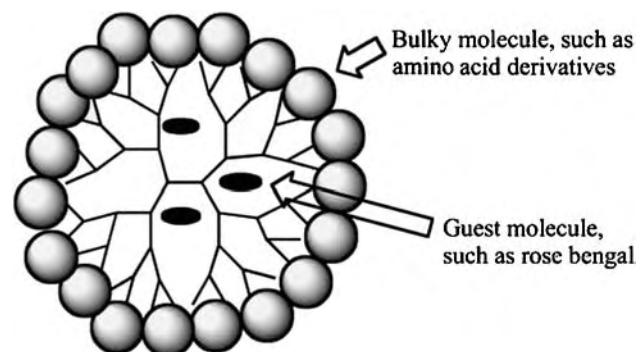


Fig. 4 A dendritic box.

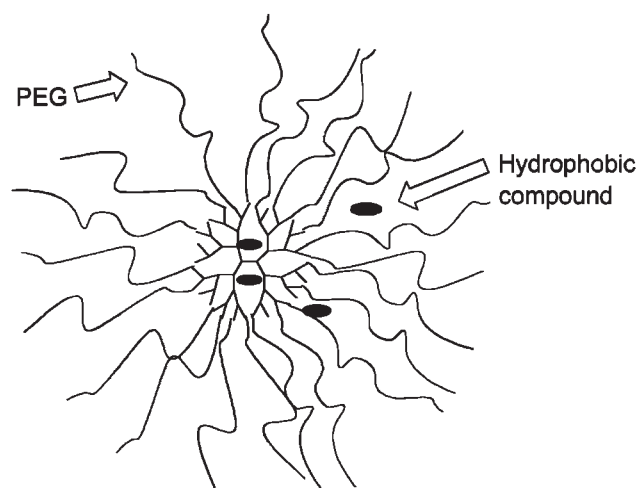


Fig. 5 A unimolecular dendritic micelle.

PAMAM dendrimer hydrogels, and such a hydrogel obtained higher swelling ratios and faster reswelling rate because of the high hydrophilicity of PAMAM dendrimers.^[41] Hydrogels and macroscopic fibers can also form through salt-induced self-assembly of water-soluble, polycationic, phosphorus dendrimers.^[42] Such hydrogels are thermoreversible and their sol-gel transition temperature can be tuned over a wide range of temperatures. Dendrimers may provide hydrogels with inner structures different from conventional cross-linking agents. Dendritic cross-linked hydrogels hold promise for drug delivery and controlled release and need further investigation to obtain thorough understandings of the relationships between hydrogel structure and hydrogel properties.

Dendrimer/DNA complexes

Dendrimers as gene carriers have shown efficiency regarding gene transfection and targeting efficacy *in vitro*. Cationic dendrimers have multiple cationic groups on the surface. Therefore, cationic dendrimers can form complexes with negatively charged DNA. The overall positive charge of the complex allows interaction with negatively charged cell-surface groups (Fig. 6), and can maintain the structural stability of DNA/dendrimer complex.^[43] There are several formulations of the dendrimer/DNA complex that have been investigated, such as PAMAM/DNA,^[43] PEG-PAMAM/DNA,^[44] PAMAM-PEG-PAMAM/DNA,^[45] PPI/DNA.^[46] Gene transfection efficiency depends on the size and amount of dendrimer in the DNA-dendrimer complex. The release of intracellular DNA, stability of DNA in the cytoplasm, unpacking of the DNA-vector complex, and the targeting ability of DNA to the nucleus also affect gene transfection

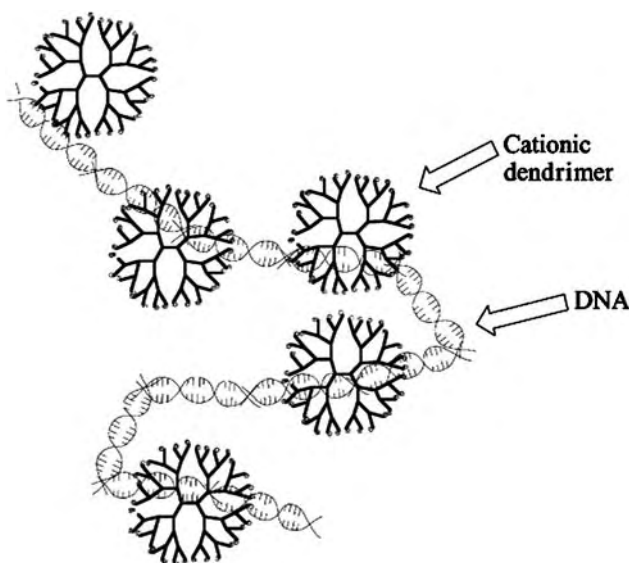


Fig. 6 DNA/cationic dendrimer complex.

efficiency. Endosomal release, nuclear localization, and transcription may be the rate-limiting steps rather than cellular uptake. The compaction of DNA by dendrimers can protect the DNA from nuclease degradation; however, it also inhibits transcriptional activity for DNA/dendrimer complexes.^[47] Therefore, it is more appropriate to consider the structural effects of DNA/dendrimer complexes on the individual events during the gene delivery process.

Chemical Bonding

Dendrimer–drug conjugates

Drug molecules can be either chemically conjugated to the dendrimer surface or physically encased inside a dendrimer core. For chemical conjugations, a good coupling efficiency may be achieved if functional groups are activated prior to coupling. Hydroxyl (OH), carboxyl (COOH), primary amine (NH₂), thiol (SH), and guanidino are commonly found functional groups in drug molecules and polymers. The strategies for the activation of such functional groups are reviewed and discussed by Veronese and Morpurgo.^[48] There are many drug molecules of interest such as anticancer drugs that can be directly conjugated to the dendrimer surface. Baker and coworkers^[49] conjugated folic acid to the dendrimer surface to obtain a targeting agent to deliver effective doses of the anticancer drug methotrexate to malignant cells. Contrast agents and fluorophores can be also conjugated to the dendrimer surface to detect the alterations of tumor vessel permeability^[50] and to amplify signals in immunoassays.^[51] Polyethylene glycol is widely used in drug delivery

owing to excellent solubility in both aqueous and most organic solutions and nontoxic and nonimmunogenic properties. Attachment of PEG to the dendrimer surface can modify the unfavorable properties of dendrimers and minimize dendrimers' toxicity and immunogenicity. In addition, each PEG strand can provide one additional ending group for coupling reaction(s). Yang and Lopina^[52] have developed synthetic approaches to synthesize penicillin-conjugated PEG-PAMAM star polymers, in which penicillin was conjugated to the outmost ending site of PEG based on full generation and half generation PAMAM dendrimers.

Glycodendrimers

Glycodendrimers are built with carbohydrates as building blocks. A typical building block is saccharide residues such as mannose, D-galactose, and lactose that make up the dendrimer's outer surface groups, the central core, or the branching units. Attachment of saccharide residues to the dendrimer surface produces cluster effect to amplify carbohydrate–protein interactions.^[53] Glycodendrimer's structure advantages generate unique properties. Saccharide-functionalized dendrimers are large enough to span multiple lectin-binding sites.^[54] Immunomodulatory and antiangiogenic properties can be obtained by conjugating D(+)-glucosamine and D(+)-glucosamine 6-sulfate to the dendrimer surface. Dendrimer glucosamine and dendrimer glucosamine 6-sulfate can be used together to prevent scar tissue formation.^[55]

Peptide dendrimers

Peptide dendrimers consist of amino acid building blocks.^[56–58] Dendrimer-mediated oligonucleotide delivery successfully inhibited the *in vitro* vascular endothelial growth factor expression and choroidal neovascularization.^[59] PPI dendrimers significantly enhanced the uptake of oligonucleotides by breast, prostate, and ovarian cancer cells.^[60] Dendrimers were reported to prolong the corneal residence time and increase the bioavailability of pilocarpine nitrate and tropicamide in eye drops.^[61] Recently, Baker and coworkers^[62] proposed a novel DNA-assembled PAMAM dendrimer drug delivery model for targeting cancer cells in which dendrimers conjugated with fluorescein and folic acid separately were linked with complementary DNA oligonucleotides.

DNA dendrimers

Using DNA building blocks to construct polymeric materials has attracted much attention, because DNA

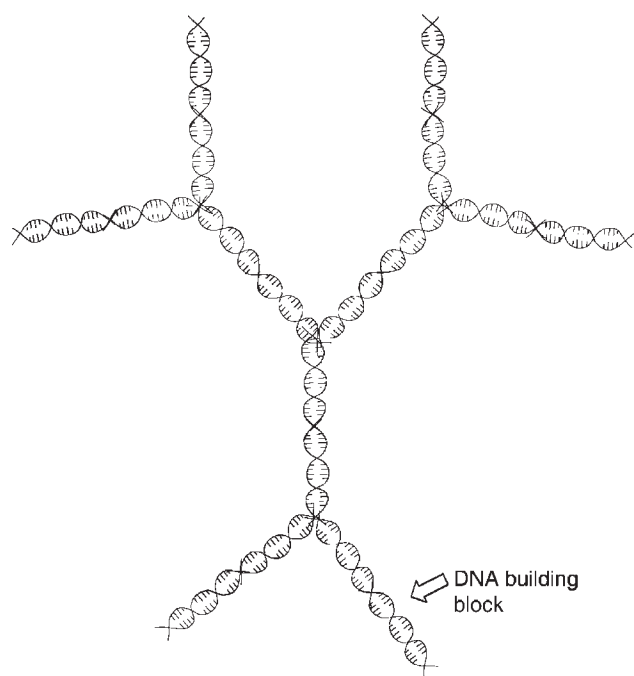


Fig. 7 A DNA dendrimer molecule.

has desirable physiochemical properties. Luo and coworkers^[63] designed and prepared dendrimer-like DNA nanostructures from Y-shaped DNA building blocks for the first time. There are two strategies for the synthesis of the Y-DNA: stepwise synthesis and all-in-one (one-pot) synthesis. In the stepwise synthesis of DNA dendrimers, one arm of a Y-DNA is assembled by two oligonucleotides with partial complementary sequences, and the other two arms of the Y-DNA are assembled by the third oligonucleotide that is complementary to the first two unmatched portions of oligonucleotides. In the onepot approach, a Y-DNA is assembled by mixing equal moles of all three oligonucleotides together and the resulting DNA is stable and almost monodisperse (Fig. 7). Potential applications of such DNA dendrimers may be based on the fact that the multivalent DNA dendrimers can be either isotropic or anisotropic and thus can provide great potential to link other entities.

PROPERTIES OF DENDRIMERS

Starburst™ PAMAM dendrimers are most investigated owing to their early and successful commercialization. The PAMAM dendrimers are water soluble and are called *artificial proteins*, because their size and contour are similar to many important natural globular proteins.^[64] For example, the size of G3.0 PAMAM dendrimer is close to insulin that has a 30 Å diameter,

and G5.0 has an equivalent diameter to hemoglobin of 55 Å. Natural proteins are easily denatured upon the change of pH, temperature, and light. In contrast, dendrimers are much more stable. Their unique properties make them good replacements of natural proteins for drug delivery and gene therapy. The PAMAM dendrimers behave similarly to liposomes. Small lipid vesicles known as liposomes have gained regulatory approval for passive targeted drug delivery,^[65] and are being widely explored for active targeting.^[65–67] Liposomes, however, must be stabilized against leakage, releasing drug before reaching the targeted site. Researchers have exploited the ability to control vesicle parameters including cholesterol content, surface charge, and vesicle size to optimize liposome-based drug delivery. The stability of dendrimers makes them good substitutes of liposome for drug delivery. Dendrimers have shown promise as drug carriers. In order to fully understand the biocompatibility, cytotoxicity, and biodistribution of dendrimers and dendrimer-based delivery systems, *in vitro* and *in vivo* evaluations are necessary. Particularly, good cell viability and morphology, no hemolysis, and other satisfying cellular parameters warrant further studies of dendrimers as carriers for drug delivery.

Roberts, Bhalgat, and Zera^[34] used V79 cells (Chinese hamster lung fibroblasts) and male Swiss-Webster mice to study PAMAM G3.0, G5.0, and G7.0 dendrimers. They found that PAMAM toxicity is dose and generation dependent *in vitro*. There was not apparent evidence of immunogenicity for any tested generation (G3.0, G5.0, or G7.0) in Roberts's study. According to Malik et al.,^[68] cationic dendrimers (with amine end groups) caused hemolysis and cytotoxicity. In contrast, 8-branched and 32-branched anionic carboxylate dendrimers were neither hemolytic nor cytotoxic up to a concentration of 2 mg/ml. Such unsuitable dendrimers as cationic PAMAM dendrimers may be accepted for biological application if the surface is attached to anionic or neutral groups. Twelve branched and thirty-six branched poly(ethylene oxide)-grafted carbosilane (Csi-PEO) dendrimer^[68] at concentrations up to 2 mg/ml did not induce hemolysis or exhibit toxicity to CCRF and HepG2 cells. However, for B16F10 cells, viability was decreased at higher concentrations (over 0.1 mg/ml) of 12-branched or 36-branched Csi-PEO dendrimers. Thirty-six branched dendrimers still allowed higher cell viability than the 12-branched dendrimers. Low generation PAMAM dendrimers (G5.0 or below) appear to have relatively fewer problems for *in vitro* applications than higher generations. The very high polycationic nature of high-generation PAMAM dendrimers (G6.0 or higher) may be responsible for some or all of their toxicity and unusual

biodistribution.^[34] Cationic PAMAM dendrimers may be biologically acceptable if the surface is modified with anionic or neutral group to reduce the toxicity of the “naked” dendrimers.^[68] Other types of dendrimers such as melamine dendrimers have been analyzed in a similar way to PAMAM dendrimers.^[69]

Therapeutics can cross cell membranes or biological barriers with the aid of dendrimers. Anionic PAMAM dendrimers G2.5 and 3.5 showed rapid serosal transfer rates in crossing adult rat intestine *in vitro* and had low tissue deposition.^[70] The transport of G3.0 PAMAM and surface-modified (with lauroyl chains) G3.0 PAMAM across Caco-2 cell monolayers followed endocytosis-mediated cellular internalization.^[71] Based on a study of the dynamics of cellular entry of dendritic polymers into human lung epithelial carcinoma cells, it was found that the cellular entry of dendritic polymers was dependent on the functional end groups and molecular mass.^[72] The electrostatic interaction with negatively charged epithelial cells, probably via the fluid phase pinocytotic route, contributed to a faster cellular entry of cationic G3.0 and G4.0 PAMAM dendrimers. A slow entry of neutral G3.0-PEG and polyol polymers may depend on nonspecific adsorption to the cell membrane and subsequent ingestion by endocytosis.^[72] Thomas et al.^[73] synthesized dendrimer-antibody (60bca, and J591) conjugates and studied their *in vitro* biological properties. These conjugated antibodies bound to the antigen-expressing cells in a time- and dose-dependent fashion, with affinity similar to that of the free antibody. The dendrimer-antibody conjugates were partly internalized. Florence, Sakthivel, and Toth^[74] synthesized lipidic peptide (polylysine) dendrimers based on lysine with 16 surface alkyl (C12) chains and investigated the absorption of a fourth generation dendrimer after oral administration to female Sprague-Dawley rats. They found that the levels of uptake and translocation are lower than those exhibited by polystyrene particles in the range from 50 to 3000 nm.

CONCLUSIONS

The application of dendrimers for drug delivery is promising owing to their unique dendritic architecture. Dendrimers also provide a unique platform to study the mechanism of cellular biology and to develop novel enabling technologies.

ACKNOWLEDGMENT

This work was supported by an NIH Grant HL-077825.

ARTICLE OF FURTHER INTEREST

Transdermal Drug Delivery

D

REFERENCES

- Hahn, U.; Gorka, M.; Vogtle, F.; Vicinelli, V.; Ceroni, P.; Maestri, M.; Balzani, V. Light-harvesting dendrimers: efficient intra- and intermolecular energy-transfer processes in a species containing 65 chromophoric groups of four different types. *Angewandte Chemie, International Edition* **2002**, *41* (19), 3595–3598.
- Wells, M.; Crooks, R.M. Preparation and properties of chemically sensitive dendrimer surfaces. *J. Am. Chem. Soc.* **1996**, *118* (16), 3988–3989.
- Bu, J.; Judeh, Z.M.A.; Ching, C.B.; Kawi, S. Epoxidation of olefins catalyzed by Mn(II) salen complex anchored on PAMAM-SiO₂ dendrimer. *Catalysis Letters* **2003**, *85* (3–4), 183–187.
- Viers, B.D.; Bauer, B.J.; Akpalu, Y.; Grohn, F.; Liu, D.; Kim, G. Hydrogels formed by endlinking PEG to dendrimer crosslink agents. *Polymer Preprints (American Chemical Society, Division of Polymer Chemistry)* **2000**, *41* (1), 728–729.
- Poxon, S.W.; Mitchell, P.M.; Liang, E.; Hughes, J.A. Dendrimer delivery of oligonucleotide. *Drug Deliv.* **1996**, *3* (4), 255–261.
- Yang, H.; Lopina, S.T. Extended release of a novel antidepressant, venlafaxine, based on anionic poly-amidoamine dendrimers and poly(ethylene glycol)-containing semi-interpenetrating networks. *J. Biomed. Mater. Res.* **2005**, *72A* (1), 107–114.
- Eichman, J.D.; Bielinska, A.U.; Kukowska-Latallo, J.F.; Baker, J.R., Jr. The use of PAMAM dendrimers in the efficient transfer of genetic material into cells. *Pharma. Sci. Technol. Today* **2000**, *3* (7), 232–245.
- Kobayashi, H.; Kawamoto, S.; Saga, T.; Sato, N.; Hiraga, A.; Ishimori, T.; Akita, Y.; Mamede, M.H.; Konishi, J.; Togashi, K.; Brechbiel, M.W. Novel liver macromolecular MR contrast agent with a poly-propylenimine diaminobutyl dendrimer core: comparison to the vascular MR contrast agent with the polyamidoamine dendrimer core. *Mag. Reson. Med.* **2001**, *46* (4), 795–802.
- Langer, R. Drug delivery and targeting. *Nature*. **1998**, *392* (Suppl 6679), 5–10.
- Tomalia, D.A.; Baker, H.; Dewald, J.; Hall, M.; Kallos, G.; Martin, S.; Roeck, J.; Ryder, J.; Smith, P. A new class of polymers: starburst-dendritic macromolecules. *Polym. J. (Tokyo)* **1985**, *17* (1), 117–132.
- Newkome, G.R.; Yao, Z.; Baker, G.R.; Gupta, V.K. Micelles. Part 1. Cascade molecules: a new approach to micelles. A [27]-arborol. *J. Org. Chem.* **1985**, *50* (11), 2003–2004.
- De Gennes, P.G.; Hervet, H. Statistics of “starburst” polymers. *J. Phys. Lett.* **1983**, *44* (9), 351–360.

13. Hawker, C.; Frechet, M.J. A new convergent approach to monodisperse dendritic macromolecules. *J. Chemical Soc. Chem. Commun.* **1990**, (15), 1010–1013.
14. Ihre, H.; Hult, A.; Frechet, J.M.J.; Gitsov, I. Double-stage convergent approach for the synthesis of functionalized dendritic aliphatic polyesters based on 2,2-bis(hydroxymethyl)propionic acid. *Macromolecules* **1998**, *31* (13), 4061–4068.
15. L'Abbe, G.; Forier, B.; Dehaen, W. A fast double-stage convergent synthesis of dendritic polyethers. *Chem. Commun. (Cambridge)* **1996**, (18), 2143–2144.
16. Wooley, K.L.; Hawker, C.J.; Frechet, J.M.J. Hyperbranched macromolecules via a novel double-stage convergent growth approach. *J. Am. Chem. Soc.* **1991**, *113* (11), 4252–4261.
17. Kawaguchi, T.; Walker, K.L.; Wilkins, C.L.; Moore, J.S. Double exponential dendrimer growth. *J. Am. Chem. Soc.* **1995**, *117* (8), 2159–2165.
18. Zeng, F.; Zimmerman, S.C. Rapid synthesis of dendrimers by an orthogonal coupling strategy. *J. Am. Chem. Soc.* **1996**, *118* (22), 5326–5327.
19. Zimmerman, S.C.; Zeng, F.; Reichert, D.E.; Kolotuchin, S.V. Self-assembling dendrimers. *Science* **1996**, *271* (5252), 1095–1098.
20. de Brabander-van den Berg, E.M.M.; Meijer, E.W. Poly(propyleneimine) dendrimers: large-scale synthesis via heterogeneously catalyzed hydrogenation. *Angewandte Chemie.* **1993**, *105* (9), 1370–1372; *Angew. Chem. Int. Ed. Engl.* **1993**, *32* (9), 1308–1311.
21. Hawker, C.J.; Lee, R.; Frechet, J.M.J. One-step synthesis of hyperbranched dendritic polyesters. *J. Am. Chem. Soc.* **1991**, *113* (12), 4583–4588.
22. Uhrich, K.E.; Boegeman, S.; Frechet, J.M.J.; Turner, S.R. The solid-phase synthesis of dendritic polyamides. *Polymer Bulletin (Berlin, Germany)* **1991**, *25* (5), 551–558.
23. Launay, N.; Caminade, A.-M.; Majoral, J.-P. Synthesis and reactivity of unusual phosphorus dendrimers a useful divergent growth approach up to the seventh generation. *J. Am. Chem. Soc.* **1995**, *117* (11), 3282–3283.
24. Salamonczyk, G.M.; Kuznikowski, M.; Skowronska, A. A divergent synthesis of thiophosphate-based dendrimers. *Tetrahedron Lett.* **2000**, *41* (10), 1643–1645.
25. Zhang, W.; Simanek, E.E. Dendrimers based on melamine divergent and orthogonal, convergent syntheses of a G3 dendrimer. *Organic Lett.* **2000**, *2* (6), 843–845.
26. Morikawa, A.; Kakimoto, M.; Imai, Y. Convergent synthesis of siloxane starburst dendrons and dendrimers via hydrosilylation. *Macromolecules* **1992**, *25* (12), 3247–3253.
27. Shao, J.; Tam, J.P. Unprotected peptides as building blocks for the synthesis of peptide dendrimers with oxime, hydrazone, and thiazolidine linkages. *J. Am. Chem. Soc.* **1995**, *117* (14), 3893–3899.
28. Jansen, J.F.G.A.; de Brabander van den Berg, E.M.M.; Meijer, E.W. Encapsulation of guest molecules into a dendritic box. *Science (Washington, DC)* **1994**, *266* (5188), 1226–1229.
29. Hawker, C.J.; Wooley, K.L.; Frechet, J.M.J. Unimolecular micelles and globular amphiphiles: dendritic macromolecules as novel recyclable solubilization agents. *J. Chem. Soc. Perkin Trans. 1.* **1993**, *12*, 1287–1297.
30. Liu, M.; Frechet, J.M.J. Preparation, MALDI-TOF analysis, and micelle-like behavior of alkyl-modified poly(propylene imine) dendrimers. *Polym. Bull. (Berlin)* **1999**, *43* (4–5), 379–386.
31. Heise, A.; Hedrick, J.L.; Frank, C.W.; Miller, R.D. Starlike block copolymers with amphiphilic arms as models for unimolecular micelles. *J. Am. Chem. Soc.* **1999**, *121* (37), 8647–8648.
32. Liu, H.; Jiang, A.; Guo, J.; Uhrich, K.E. Unimolecular micelles: synthesis and characterization of amphiphilic polymer systems. *J. Polym. Sci. Part A: Polym. Chem.* **1999**, *37* (6), 703–711.
33. Milhem, O.M.; Myles, C.; McKeown, N.B.; Attwood, D.; D'Emanuele, A. Polyamidoamine starburst dendrimers as solubility enhancers. *Int. J. Pharm.* **2000**, *197* (1–2), 239–241.
34. Roberts, J.C.; Bhalgat, M.K.; Zera, R.T. Preliminary biological evaluation of polyamidoamine (PAMAM) starburst dendrimers. *J. Biomed. Mater. Res.* **1996**, *30* (1), 53–65.
35. Li, J.; Piehler, L.T.; Qin, D.; Baker, J.R., Jr.; Tomalia, D.A.; Meier, D.J. Visualization and characterization of poly(amidoamine) dendrimers by atomic force microscopy. *Langmuir* **2000**, *16* (13), 5613–5616.
36. Bhalgat, M.K.; Roberts, J.C. Molecular modeling of polyamidoamine (PAMAM) starburst dendrimers. *Eur. Polym. J.* **2000**, *36* (3), 647–651.
37. Kojima, C.; Kono, K.; Maruyama, K.; Takagishi, T. Synthesis of polyamidoamine dendrimers having poly(ethylene glycol) grafts and their ability to encapsulate anticancer drugs. *Bioconjugate Chem.* **2000**, *11* (6), 910–917.
38. Sideratou, Z.; Tsiourvas, D.; Paleos, C.M. Solubilization and release properties of pegylated diaminobutane poly(propylene imine) dendrimers. *J. Colloid Interface Sci.* **2001**, *242* (1), 272–276.
39. Yang, H.; Morris, J.J.; Lopina, S.T. Polyethylene glycol-polyamidoamine dendritic micelle as solubility enhancer and the effect of the length of polyethylene glycol arms on the solubility of pyrene in water. *J. Colloid Interface Sci.* **2004**, *273* (1), 148–154.
40. Gitsov, I.; Zhu, C. Amphiphilic hydrogels constructed by poly(ethylene glycol) and shape-persistent dendritic fragments. *Macromolecules* **2002**, *35* (22), 8418–8427.
41. Wu, X.-Y.; Huang, S.-W.; Zhang, J.-T.; Zhuo, R.-X. Preparation and characterization of novel physically cross-linked hydrogels composed of poly(vinyl alcohol) and amine-terminated polyamidoamine dendrimer. *Macromolecular Bioscience* **2004**, *4* (2), 71–75.
42. El Ghzaoui, A.; Gauffre, F.; Caminade, A.-M.; Majoral, J.-P.; Lannibois-Drean, H. Self-assembly of water-soluble dendrimers into thermoreversible hydrogels and macroscopic fibers. *Langmuir* **2004**, *20* (21), 9348–9353.
43. Braun, C.S.; Vetro, J.A.; Tomalia, D.A.; Koe, G.S.; Koe, J.G.; Middaugh, C.R. Structure/function relationships of polyamidoamine/DNA dendrimers as gene delivery vehicles. *J. Pharm. Sci.* **2005**, *94* (2), 423–436.
44. Luo, D.; Haverstick, K.; Belcheva, N.; Han, E.; Saltzman, W.M. Polyethylene glycol-conjugated PAMAM

- dendrimer for biocompatible, high-efficiency DNA delivery. *Macromolecules* **2002**, *35* (9), 3456–3462.
45. Kim, T.I.; Seo, H.J.; Choi, J.S.; Jang, H.S.; Baek, J.U.; Kim, K.; Park, J.S. PAMAM-PEG-PAMAM: novel triblock copolymer as a biocompatible and efficient gene delivery carrier. *Biomacromolecules* **2004**, *5* (6), 2487–2492.
 46. Schatzlein, A.G.; Zinselmeyer, B.H.; Elouzi, A.; Dufes, C.; Chim, Y.T.A.; Roberts, C.J.; Davies, M.C.; Munro, A.; Gray, A.I.; Uchegbu, I.F. Preferential liver gene expression with polypropylenimine dendrimers. *J. Contro. Release* **2005**, *101* (1–3), 247–258.
 47. Bielinska, A.U.; Kukowska-Latallo, J.F.; Baker, J.R., Jr. The interaction of plasmid DNA with polyamidoamine dendrimers: mechanism of complex formation and analysis of alterations induced in nuclease sensitivity and transcriptional activity of the complexed DNA. *Biochim. Biophys. Acta* **1997**, *1353* (2), 180–190.
 48. Veronese, F.M.; Morpurgo, M. Bioconjugation in pharmaceutical chemistry. *IL Farmaco*. **1999**, *54* (8), 497–516.
 49. Quintana, A.; Raczka, E.; Piehler, L.; Lee, I.; Myc, A.; Majoros, I.; Patri, A.K.; Thomas, T.; Mule, J.; Baker, J.R., Jr. Design and function of a dendrimer-based therapeutic nanodevice targeted to tumor cells through the folate receptor. *Pharm. Res.* **2002**, *19* (9), 1310–1316.
 50. Kobayashi, H.; Reijnders, K.; English, S.; Yordanov, A.T.; Milenic, D.E.; Sowers, A.L.; Citrin, D.; Krishna, M.C.; Waldmann, T.A.; Mitchell, J.B.; Brechbiel, M.W. Application of a macromolecular contrast agent for detection of alterations of tumor vessel permeability induced by radiation. *Clin. Cancer. Res.* **2004**, *10* (22), 7712–7720.
 51. Ong, K.K.; Jenkins, A.L.; Cheng, R.; Tomalia, D.A.; Durst, H.D.; Jensen, J.L.; Emanuel, P.A.; Swim, C.R.; Yin, R. Dendrimer enhanced immunosensors for biological detection. *Analytica Chimica Acta* **2001**, *444* (1), 143–148.
 52. Yang, H.; Lopina, S.T. Penicillin V-conjugated PEG-PAMAM star polymers. *J. Biomater. Sci. Polym. Ed.* **2003**, *14* (10), 1043–1056.
 53. Zanini, D.; Park, W.K.C.; Roy, R. Synthesis of novel dendritic glycosides. *Tetrahedron Lett.* **1995**, *36* (41), 7383–7386.
 54. Woller, E.K.; Cloninger, M.J. The lectin-binding properties of six generations of mannose-functionalized dendrimers. *Org. Lett.* **2002**, *4* (1), 7–10.
 55. Shaunak, S.; Thomas, S.; Gianasi, E.; Godwin, A.; Jones, E.; Teo, I.; Mireskandari, K.; Luthert, P.; Duncan, R.; Patterson, S.; Khaw, P.; Brocchini, S. Polyvalent dendrimer glucosamine conjugates prevent scar tissue formation. *Nature Biotechnol.* **2004**, *22* (8), 977–984.
 56. Bracci, L.; Falciani, C.; Lelli, B.; Lozzi, L.; Runci, Y.; Pini, A.; De Montis, M.G.; Tagliamonte, A.; Neri, P. Synthetic peptides in the form of dendrimers become resistant to protease activity. *J. Biol. Chem.* **2003**, *278* (47), 46,590–46,595.
 57. Spetzler, J.C.; Tam, J.P. Unprotected peptides as building blocks for branched peptides and peptide dendrimers. *Int. J. Pept. Protein Res.* **1995**, *45* (1), 78–85.
 58. Sadler, K.; Tam James, P. Peptide dendrimers: applications and synthesis. *J. biotechnol.* **2002**, *90* (3–4), 195–229.
 59. Marano, R.J.; Wimmer, N.; Kearns, P.S.; Thomas, B.G.; Toth, I.; Brankov, M.; Rakoczy, P.E. Inhibition of in vitro VEGF expression and choroidal neovascularization by synthetic dendrimer peptide mediated delivery of a sense oligonucleotide. *Exp. Eye Res.* **2004**, *79* (4), 525–535.
 60. Santhakumaran, L.M.; Thomas, T.; Thomas, T.J. Enhanced cellular uptake of a triplex-forming oligonucleotide by nanoparticle formation in the presence of polypropylenimine dendrimers. *Nucleic Acids Res.* **2004**, *32* (7), 2102–2112.
 61. Vandamme, T.F.; Brobeck, L. Poly(amidoamine) dendrimers as ophthalmic vehicles for ocular delivery of pilocarpine nitrate and tropicamide. *J. Control Release* **2005**, *102* (1), 23–38.
 62. Choi, Y.; Thomas, T.; Kotlyar, A.; Islam, M.T.; Baker, J.R., Jr. Synthesis and functional evaluation of DNA-assembled polyamidoamine dendrimer clusters for cancer cell-specific targeting. *Chem. Biol.* **2005**, *12* (1), 35–43.
 63. Li, Y.; Tseng, Y.D.; Kwon, S.Y.; D’Espaux, L.; Bunch, J.S.; McEuen, P.L.; Luo, D. Controlled assembly of dendrimer-like DNA. *Nat. Mater.* **2004**, *3* (1), 38–42.
 64. Esfand, R.; Tomalia, D.A. Poly(amidoamine) (PAMAM) dendrimers: from biomimicry to drug delivery and biomedical applications. *Drug Discovery Today* **2001**, *6* (8), 427–436.
 65. Forssen, E.; Willis, M. Ligand-targeted liposomes. *Adv. Drug Delivery Rev.* **1998**, *29* (3), 249–271.
 66. Spragg, D.D.; Alford, D.R.; Greferath, R.; Larsen, C.E.; Lee, K.-D.; Gurtner, G.C.; Cybulsky, M.I.; Tosi, P.F.; Nicolau, C.; Gimbrone, M.A., Jr. Immunotargeting of liposomes to activated vascular endothelial cells: a strategy for site-selective delivery in the cardiovascular system. *Proc. Natl. Acad. Sci. USA* **1997**, *94* (16), 8795–8800.
 67. Vandoorpe, J.; Schacht, E.; Dunn, S.; Hawley, A.; Stolnik, S.; Davis, S.S.; Garnett, M.C.; Davies, M.C.; Illum, L. Long circulating biodegradable poly(phosphazene) nanoparticles surface modified with poly(phosphazene)-poly(ethylene oxide) copolymer. *Biomaterials* **1997**, *18* (17), 1147–1152.
 68. Malik, N.; Wiwattanapatapee, R.; Klopsch, R.; Lorenz, K.; Frey, H.; Weener, J.W.; Meijer, E.W.; Paulus, W.; Duncan, R. Dendrimers: relationship between structure and biocompatibility in vitro, and preliminary studies on the biodistribution of 125I-labeled polyamidoamine dendrimers in vivo. *J. Control. Release* **2000**, *65* (1–2), 133–148.
 69. Chen, H.T.; Neerman, M.F.; Parrish, A.R.; Simanek, E.E. Cytotoxicity, hemolysis, and acute in vivo toxicity of dendrimers based on melamine, candidate vehicles for drug delivery. *J. Am. Chem. Soc.* **2004**, *126* (32), 10,044–10,048.
 70. Wiwattanapatapee, R.; Carreno-Gomez, B.; Malik, N.; Duncan, R. Anionic PAMAM dendrimers

- rapidly cross adult rat intestine in vitro: a potential oral delivery system? *Pharmaceutical Res.* **2000**, *17* (8), 991–998.
71. Jevprasesphant, R.; Penny, J.; Attwood, D.; D'Emanuele, A. Transport of dendrimer nanocarriers through epithelial cells via the transcellular route. *J. Control. Release* **2004**, *97* (2), 259–267.
 72. Kannan, S.; Kolhe, P.; Raykova, V.; Glibatec, M.; Kannan, R.M.; Lieh-Lai, M.; Bassett, D. Dynamics of cellular entry and drug delivery by dendritic polymers into human lung epithelial carcinoma cells. *J. Biomater. Sci. Polym. Ed.* **2004**, *15* (3), 311–330.
 73. Thomas, T.P.; Patri, A.K.; Myc, A.; Myaing, M.T.; Ye, J.Y.; Norris, T.B.; Baker, J.R., Jr. In vitro targeting of synthesized antibody-conjugated dendrimer nanoparticles. *Biomacromolecules* **2004**, *5* (6), 2269–2274.
 74. Florence, A.T.; Sakhivel, T.; Toth, I. Oral uptake and translocation of a polylysine dendrimer with a lipid surface. *J. Control. Release* **2000**, *65* (1–2), 253–259.

A. N. Natali
P. G. Pavan
E. L. Carniel

Centre of Mechanics of Biological Materials, University of Padova, Padova, Italy

INTRODUCTION

The biomechanical analysis in dental area represents a relevant task in biomedical research. Universities, research centers, and industrial laboratories show a growing interest in investigating different aspects such as dental biomaterials performances, soft and hard tissue mechanics, dental implants, and orthodontics devices reliability, by means of experimental and numerical approaches.

In this article, according to the necessity of limiting the arguments to be treated, the attention is addressed to the periodontal ligament (PDL), tooth and dental implant mechanics. These topics allow for a preliminary discussion about the response of the soft and hard tissues that characterize the mechanics in dental area and for a direct investigation about the functional behavior of main structural components during masticatory action.

Particular attention is paid to the numerical approach to dental biomechanics, showing the offered potentialities and discussing the refined and reliable results. Data from experimental analysis are also reported as a basis for a reciprocal reliability evaluation. This article is not only aimed at offering an introduction to the investigation of mechanical problems in dental area but it also gives an appropriate description of specific relevant tasks and can represent the starting point for further developments.

BIOMECHANICS OF THE PDL

The PDL

The PDL is a thin layer of soft connective tissue that surrounds the root of a tooth and provides a connection with the alveolar bone. The transmission of loads, as occlusal and masticatory or orthodontics loads, from tooth to alveolar bone is regulated by the PDL according to its structural configuration, composed of ground material and collagen fibers.^[1] The PDL overall stiffness is much lesser than surrounding bone and tooth materials. As a consequence, the PDL

deformation mostly determines tooth instantaneous and short-term mobility.^[2] Furthermore PDL strain state regulates the activity of cells involved in alveolar bone remodeling process and consequently influences tooth long-term mobility.^[3] It follows that biomechanical characterization of PDL is a fundamental task to analyze tooth response to loads. The biomechanical response of PDL is investigated by means of both experimental and numerical analysis.

Structural Arrangement and Mechanical Behavior of PDL

Alveolar bone, cementum, lamina propria of gingiva, and PDL form the boundary structure of tooth, usually named periodontium. Near the alveolar crest, PDL joins with gingiva, while in the lower portion of its apical region PDL is in direct contact with dental pulp. In crestal, central, and upper apical regions, PDL is directly linked to alveolar bone on one side and cementum on the other side (Fig. 1A). The thickness of human PDL ranges from 0.15 to 0.25 mm, depending on the specific tooth and position around the root and varies with age and healthy condition of the subject.^[1,4]

The PDL is a well-vascularized soft connective tissue consisting of a fibrous stroma, with collagen (varying in the range from 53% to 74% in volume) and elastin fibers embedded in a gel of ground substance. Collagen fibers mainly withstand tensile stresses as most soft connective tissues, while ground substance withstands compressive stresses.

The ground substance is a viscous gel that contains proteoglycans, hyaluronan, glycoproteins, and water. The high liquid content of the ground substance and electrostatic phenomena, mainly involving proteoglycans, determines the volumetric behavior and the consequent compressive stiffness of the PDL. Tissue viscoelastic behavior is mainly dependent on the liquid substance fluxes inside the fibrous network.

Collagen fibers develop through PDL from cementum to alveolar bone. In Fig. 1B, the distribution of the collagen fibers and fibrils, as the principal sub-component of collagen fibers, is reported. Their orientation is influenced by the loads acting on the

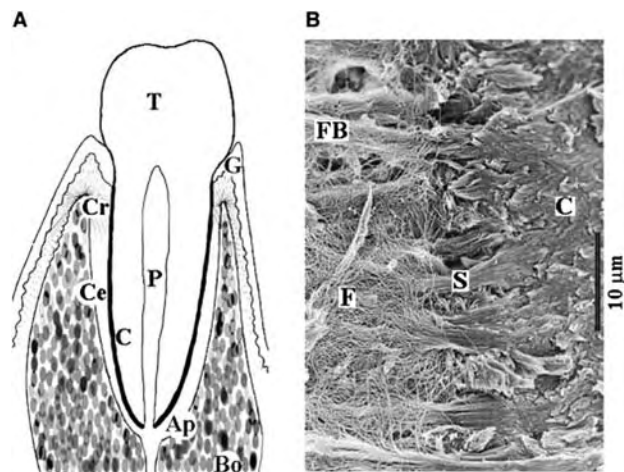


Fig. 1 (A) Schematic representation of tooth (T), alveolar bone (B), gingiva (G), dental pulp (P), cementum (C), crestal region of PDL (Cr), central region of PDL (Ce), apical region (Ap) and (B) configuration of collagen fibrils bundles (F) and fibers (FB) within a human PDL sample; Sharpey fibers (S) connect soft tissue with cementum (C).

tooth. In the unloaded state, they are characterized by a typical crimped conformation, while tensile loading determines uncrimping phenomena. The waviness change under tensile loading contributes to the typical nonlinear stress–strain behavior of the PDL, showing a low stiffness at moderately large strains and much higher stiffness at large strains. By increasing the load, failure phenomena can take place and tissue stiffness decreases, showing a typical softening behavior.

Experimental Analysis of PDL

Experimental investigation of the PDL biomechanical properties is mainly developed following two complementary strategies. To characterize the PDL as a material, it is necessary to develop experimental tests on simple geometry specimens, while the characterization of the PDL as a structure entails experimental testing of the entire tooth unit (tooth, alveolar bone, and PDL).

The former strategy usually requires cutting the tooth unit into slices perpendicular to the tooth longitudinal axis (Fig. 2A).^[5] Subsequently the testing samples are prepared starting from the slices (as indicated in Fig. 2B by white rectangles) to obtain specimens that contain a part of bone, a part of root, and a part of the PDL between them, as reported in Fig. 2C. Samples' cutting is usually developed by rotating metal saws, providing for a continuous perfusion by water or water solutions to reduce tissues alterations. The geometric dimensions of the specimen, with particular regard to the PDL layer, must be carefully evaluated by means of optical microscopes. Experimental tests are developed maintaining the soft tissue inside a physiologic solution, to avoid drying out phenomena. Mechanical tests are usually developed by standard material testing machines, but because of the peculiar properties of biological samples surfaces, it is necessary to develop special clamping fixtures, to avoid slipping or damage of the sample.^[6] Before testing the PDL specimens, preconditioning procedures of the tissue should be developed to stabilize the material response.^[7] Standard

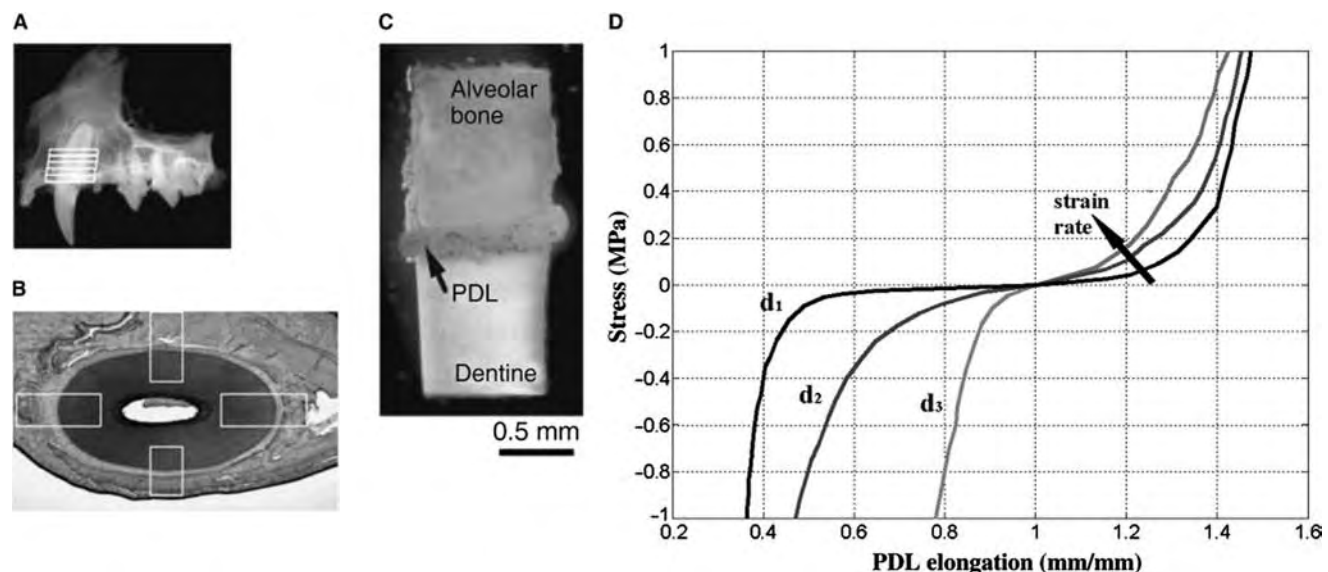


Fig. 2 (A) Experimental tests on tensile and compressive response of cat PDL sample: maxillary canine of cat (white rectangles indicate the transverse sections that are cut out); (B) transverse slice from the maxillary canine (white rectangles indicate the specimens taken for experimental testing); (C) sample including alveolar bone, dentine, and PDL between them; and (D) stress–elongation curves for tensile and compressive tests at different strain rates as d_1 , d_2 , and d_3 .

elongation tests make it possible to obtain information about tissue stiffness and strength.^[8]

Fig. 2D reports data from tensile and compression loading tests, by considering different strain rates, on cat PDL specimens.^[5] Results emphasize the time dependence of PDL mechanical behavior by showing the influence of strain rate on the stress–elongation curve. The different responses of the tissue in tension and compression must be pointed out. A strong dependence on the strain rate is found in compression, while tensile tests are less affected by the strain rate. This behavior is mostly determined by the different fluxes of liquid phases inside the fibers network. Elongation tests at different strain rates are performed also for investigating hysteresis, stress relaxation, and creep response to determine the PDL viscoelastic properties.

Fig. 3 represents data from a tensile loading test on pig PDL specimen.^[6,9] The tensile stress–elongation curve is characterized by strongly nonlinear behavior in the first branch that is correlated with the collagen fibers uncrimping, followed by an almost linear region. During the uncrimping process, mechanical stiffness gradually increases and keeps a nearly constant value in an almost linear range. Afterwards, as previously mentioned, damage phenomena take place and tissue stiffness decreases.

The characterization of PDL as a structure entails the “in vivo” testing of the tooth and the measurement of the corresponding displacements.^[10] To obtain meaningful data about PDL response, the tooth under testing must not have any contact with adjacent teeth.^[11] The procedure requires accurate systems for load application and evaluation of the tooth displacements and rotations. Several methods have been adopted both for load application and for displacements measurement. For example, loads can be applied

by direct contact between tooth and mechanical devices or by pulling a thread that is enveloped around the tooth,^[12] while displacements can be measured by using optical systems^[13] or electric devices.^[12] These techniques make it possible to directly characterize the tooth mobility when physiological loads are applied.

Formulation of Constitutive Models and Numerical Analysis of PDL

Mechanical testing shows that PDL is characterized by nonlinear and time-dependent mechanical response, while histological analysis, with particular regard to the collagen fibers distribution, suggests considering PDL as an anisotropic material. A first approach to constitutive modeling of the PDL has been developed in the framework of linear elasticity.^[14] This approach entails heavy limitations in evaluating the appropriate biomechanical behavior. A more accurate representation of the nonlinear characteristic and almost-incompressible response of the PDL can be performed by using isotropic and anisotropic hyperelastic constitutive models.^[15,16] These models entail the definition of a strain energy function capable to interpret the nonlinear response of the tissue accounting for the almost-incompressible and anisotropic behavior. To define PDL mechanical behavior along the whole stress–elongation curve (Fig. 3E) up to a failure region, elastodamage constitutive models have been formulated.^[17] These models are defined starting from hyperelastic formulation by introducing a specific damage function. In this way it is possible to account for damage phenomena such as collagen fibers rupture. This approach is particularly useful for the analysis of tooth mobility in the case of degeneration owing to ageing or pathologies.

The analysis of the PDL biomechanical behavior must evaluate the time-dependent response, assuming viscoelastic constitutive models in the framework of the hyperelastic or elastodamage formulations.^[18] Another approach to the analysis of PDL mechanics and in particular to the time-dependent behavior is the adoption of the porous media formulation, which accounts for the contribution of the fibrous network and fluid components considering the flux of the liquid phase within the fibrous network.^[19] This approach can describe the interaction between liquid and solid phases of the soft tissue and provides a better micro-mechanical consistency of the model.

Constitutive parameters must be defined within the numerical model formulated. This is usually performed by comparing the experimental data from tests on simple geometry specimens with the results obtained by analytical solutions that interpret the tests themselves,

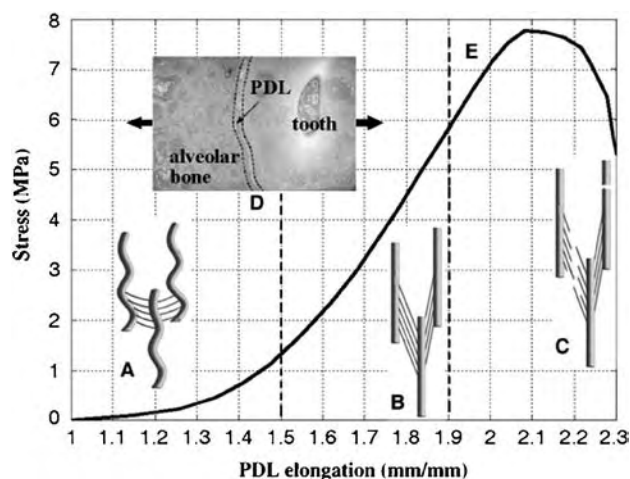


Fig. 3 Configuration of fibrils and interfibrillar bonds (A, B, and C) within a stress–elongation curve (E) for pig PDL sample test in tension (D).

leading to operational procedures that often determine remarkable difficulties.

As constitutive models are defined, the investigation of the biomechanical response of the PDL and the tooth mobility requires the development of solid and numerical models, to perform a finite element analysis (FEA) of the problem. Solid models are usually obtained starting from radiographic or tomographic data of the tooth unit. Fig. 4 reports the result of the FEA of tooth short-term mobility, when load is applied along buccolingual direction (Fig. 4A) and a comparison between experimental and numerical results.^[19,20] The load is kept constant for a short term (2.5 sec) and then relaxed (Fig. 4B). The analysis is performed by a numerical model of the tooth unit (Fig. 4C) and assuming a poroelastic constitutive model for PDL. The evolution of tooth displacement on time is reported in Fig. 4D. It must be confirmed the necessity to compare numerical results with experimental data from “in vivo” tests to provide a reciprocal validation of the procedures.

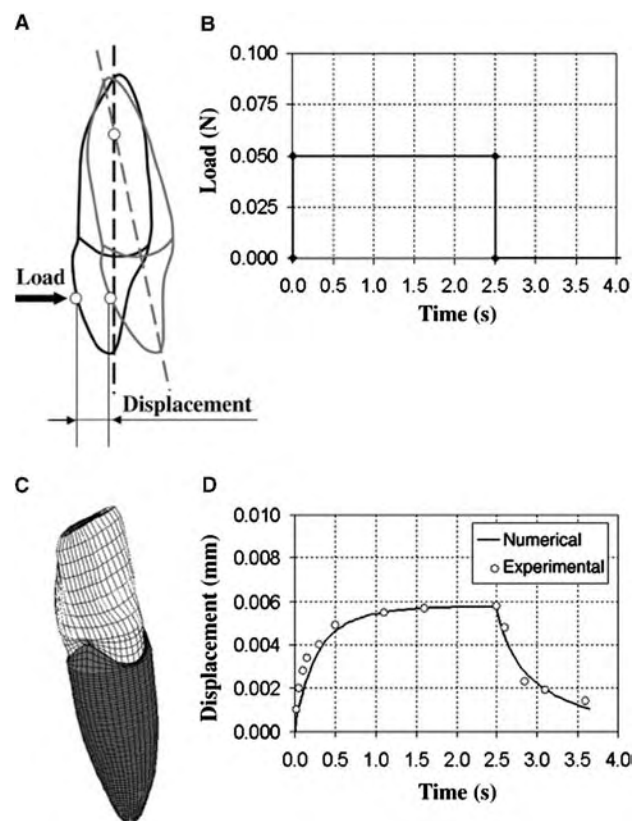


Fig. 4 Porous media analysis of tooth mobility as loading-unloading response of a central incisor: (A) schematic representation of the tooth loading conditions; (B) load history; (C) numerical model of tooth and PDL; and (D) tooth displacement in loading and unloading conditions as comparison between experimental and numerical results.

This formulation of constitutive models for PDL makes it possible to analyze tooth short-term mobility, while the analysis of long-term mobility requires also the introduction of a remodeling law for the adjacent alveolar bone. Several authors have proposed the existence of correlations between PDL strain state and the activity of alveolar bone osteoblasts and osteoclasts,^[3,21] leading to the possibility to investigate orthodontic treatments.

BIOMECHANICS OF DENTAL IMPLANTS

Introduction to Dental Implants Mechanics

Implantology is probably the most efficient procedure for the treatment of partial and total edentulous patients.^[22] This technique is becoming more and more available among dentists and is receiving major attention by researchers investigating the clinical, biochemical, and biomechanical aspects that represent the basis of treatment success.

A single implant can be fixed in bone (mandible or maxilla) to directly support an artificial tooth, while other surgical plans make use of two or more implants inserted in the bone and connected by bridgework that becomes the support for a set of artificial teeth. This prosthetic configuration, mostly adopted in the case of total edentulous patients, is more complex also from a biomechanical point of view, and must be evaluated also with regard to the passive connection (fit) between implants and framework.^[23]

Many aspects that are at the basis of success or failure of a fixed prosthesis are closely related to mechanical issues. The implant system must be capable to guarantee a normal functionality to the patient by the capability to remain fixed and to bear the occlusal loads. These requirements entail a correct design of the implants, of all the secondary components, of the bridgework in multi-implant configurations, and, especially, the evaluation of the way in which the implants get and maintain their fixity in bone tissue. The last aspect is particular important, because success or failure of fixed prostheses also depends on the capability to transmit occlusal loads to the bone in which the implants are placed, inducing appropriate mechanical stress. In fact, an excessive mechanical stress could induce bone resorption. The consequence is usually very heavy for the patient and the risk of bone overloading should be very carefully evaluated. Therefore, the correct clinical treatment should entail also a structural design of the prosthesis in the true sense of the word, and this task can be achieved by using traditional engineering methods, both experimental and numerical.

Configuration of Dental Implants

Dental implants usually have a cylindrical threaded shape offering, whether self-tapping or not, a good mechanical performance with regard to the interaction phenomena occurring between implant and surrounding bone tissue. Other shapes, conical for example, are also present but are not frequently used. Surgical procedures adopted can also be different depending on “one-stage” or “two-stage” technique. With the first technique, the implant is placed in bone with one only surgical intervention and the shape of the implants designed to protrude through the gingival tissues. In the “two-stage” technique, a fixture is placed in bone and is totally covered by soft tissues until it reaches a sufficient fixity in bone. By a second surgical intervention, an abutment is connected to the fixture by using different connection procedures and the whole system finally protrudes through the gingival tissues. The material of dental implants is mostly titanium (for example grade-2, or -4), which offers an optimum biocompatibility and an excellent resistance to corrosion in physiological environment.^[24]

The prosthetic bridgework is designed according to the specific anatomical configuration of every patient and is obtained by using different materials and manufacturing procedures.

Dental Implants Integration

The biomechanical functionality of dental implants is related to their capability of being a fixed support. An essential requirement is the fixity of the implant in bone tissue that is correlated with the osseointegration, which is defined as “a direct structural and functional connection between ordered living bone and the surface of a load carrying implant.”^[25] During the healing phase after the insertion of an implant, the tissue growing is favored by the presence of small ridges and asperities. This configuration can be provided by adopting a proper treatment of the implant surface.^[26]

Plasma-spray techniques or chemical deposition make it possible to obtain a surface roughness on fine scale that promotes such a growth of bone. A proper roughness of the surface can be obtained also with other procedures such as sandblasting, electroerosion, or deposition. Hydroxyapatite coating has been used to offer a surface with a similar chemical composition of bone tissue for its growth. Recent developments in surface treatments have considered the use of specific peptides that favor the adhesion of osteoblasts and the following growth of bone tissue. This technique should improve the tendency of the osseointegration, especially in the very initial phase of healing.

The shape, the choice of appropriate diameter, and length of the endosseous portion of the implants have obviously relevant effects on the mechanical action induced by occlusal loads on the peri-implant bone tissue. This is directly correlated to the maintenance of osseointegration in time and, therefore, on the long-term fixity and success of the implant. The details in the geometrical configuration are also important for their influence on mechanical performances. The use of threaded implants can offer a primary fixity and is nowadays widely used in clinical practice.

Mechanical Testing on Dental Implants

Mechanical testing on dental implants is necessary to reach reliability requirements, mostly with reference to the capability of getting the fixity in mandible and bearing the occlusal loads. It is also necessary to estimate the magnitude and the type of occlusal loads. This estimation is really hard, because the loads depend on several factors such as the patient anatomy and the prosthetic configuration. The type of occlusal forces related to masticatory activity has been studied in different “in vivo” tests^[27,28] and attention has been also paid to correlate the magnitude of these forces to the strength in the components making up the implant, or the loss of marginal bone caused by an excessive stress-strain state of the peri-implant region that can reduce the mechanical function of the implant.^[29-31]

Concerning the mechanical performance of the implant, experimental tests to evaluate the ultimate load and the postelastic behavior of the implant are performed with the aim of estimating the ductility level, i.e., the ability to experience plastic (large) deformations without failure. Generally speaking, implant systems with a good level of ductility should be preferred, while the choice implants showing a brittle failure, i.e., a failure that immediately follows the elastic limit of the structure, should be avoided. These mechanical tests are also capable of giving information about the strength hierarchy of the different components within an implant system.

The loading related to occlusal activity induces cyclical stresses on the different components of the prosthetic systems, leading to failure. It is extremely interesting to evaluate the cycle-dependent strength of the different components of the implant system for an estimation of the long-term reliability of dental implants. It is obviously very difficult to perform “in vivo” fatigue tests, because the total number of loading cycles should be in the order of some millions and also because of the problem to control “in vivo” combination and amplitude of the forces acting at the level of an implant. Fatigue tests “in vitro”^[32] can be performed, by adopting a specific test machine that is

capable of applying cyclical loading on an implant fixed in a synthetic support, usually made up of an acrylic resin or aluminum. A typical failure correlated to cyclical loads can occur in the small connecting screw of those prosthetic systems that are characterized by the join of endosseous fixture and abutment.

The direct application of occlusal loads is not the only reason of the stress induced in bone and in the different components of the prosthetic systems. In fact, in multi-implant prosthetic configurations relevant stress states can be caused by a misfit in the join between the different implants and the connecting bar. The level of stress is directly related to the magnitude and type of misfit and also to the stiffness of the connecting bar. The misfit depends on several factors, mainly concerning the accuracy in manufacturing process adopted. Several experiences have been made to estimate bar-implant misfit, both making use of “in vivo” and “in vitro” techniques. The latter are more suitable for obtaining information about the mechanical state of the prosthetic components. Some “in vivo” tests have been performed to verify the risk of mobility of an implant correlated to excessive stress of peri-implant tissue induced by misfit.^[33] This type of tests including the action of occlusal loads are not performed. This is a limit of the investigation, as it is clear that the effective mechanical state of the prosthetic system and peri-implant regions is the combination of occlusal forces and coaction owing to a possible misfit.

Numerical Analysis of Dental Implants Mechanics

As mentioned in the previous section, experimental testing represents heavy and complex task. In addition, experimental tests performed on animals give rise to ethical problems. In an experimental activity, many

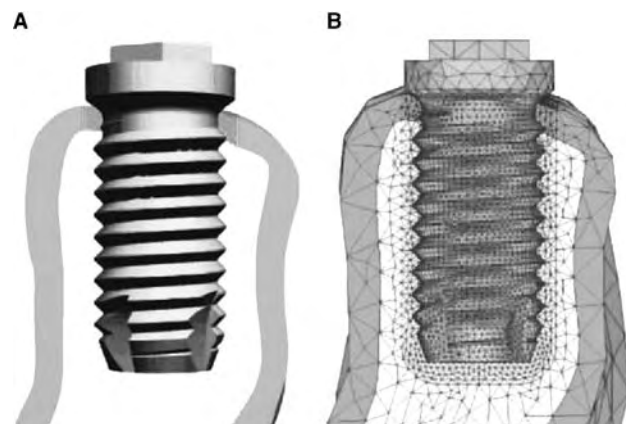


Fig. 5 (A) Three-dimensional virtual solid model and (B) finite element model of a threaded implant inserted in a mandible.

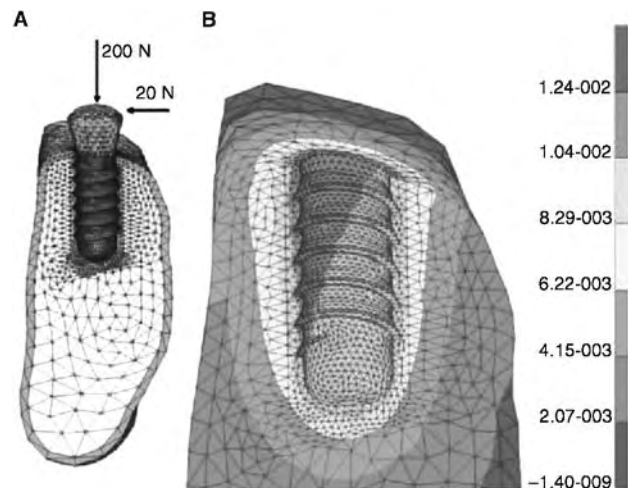


Fig. 6 (A) Finite element model of a single implant inserted in a mandible with indication of typical occlusal forces and (B) magnitude displacement field induced in the mandible by the indicated loading configuration.

parameters such as stress or strain in the peri-implant bone region remain unknown, while only a partial number of data is reasonably achievable. These reasons, together with time and costs requested, favor the use of numerical analysis to investigate the mechanical behavior of implant systems and the functional response of anatomical districts that interact with implant themselves. In the last decade in particular, the use of finite element method has reached wide diffusion because of its capability to provide extended and reliable data on the biomechanical behavior of dental implant systems that can be correlated with data taken from experimental activity for a reciprocal validation. This method makes it possible to simulate the mechanical response of the bone-implant compound, because effective geometry, mechanical properties of the

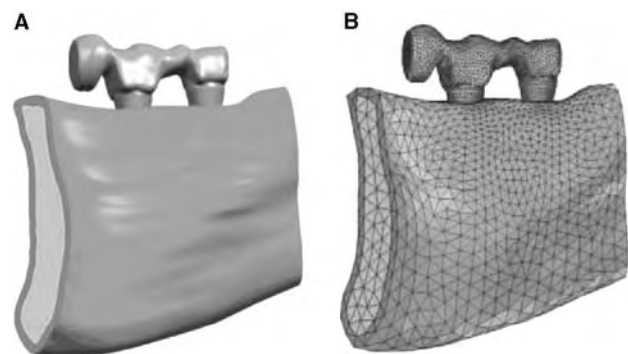


Fig. 7 (A) Three-dimensional virtual solid model and (B) finite element model of a portion of mandible with bridgework connecting two implants inserted in premolar region.

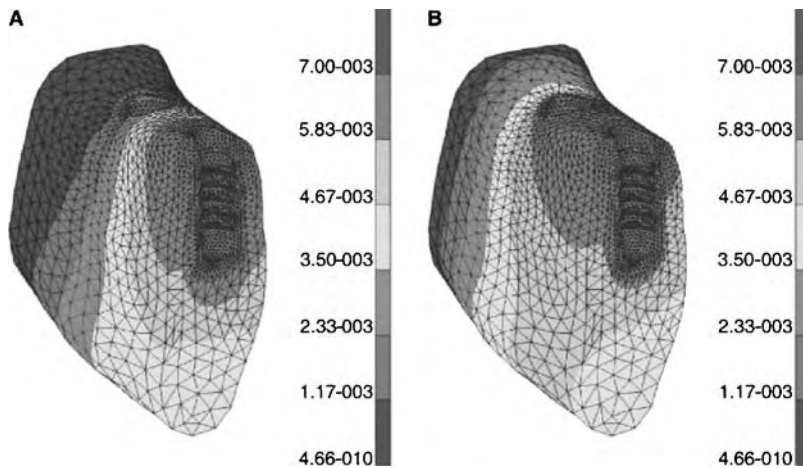


Fig. 8 (A) Magnitude displacement field in the case of an intrusive 200 N force applied on the bridgework at the end of the cantilever and (B) a distributed 300 N load applied along the bridgework between the two implants.

different materials, and loading conditions can be accurately defined in the finite element model.^[34]

Concerning the mechanical response of a single implant fixed in the mandible bone, a FEA can be oriented to investigate several factors such as the variability of implants (length, diameter, etc.) and their placement within the jaw, the effects of the type of loading (intrusive, transversal, etc.), and the influence of the mechanical properties of the bone tissue that interacts with the implant. All these terms affect the biomechanical reliability of the implant. If the analysis is properly performed, the results obtained can lead to reliable evaluation, as it can be analyzed under several configurations that interpret the real working conditions. As a way of example, a three-dimensional geometric model and the corresponding finite element model of a threaded implant with the surrounding region of the mandible are shown in Fig. 5. The displacement induced for a typical combination of occlusal forces acting on another type of implant (Fig 6A) inserted in a mandible is proposed in Fig. 6B.

The large potentiality of the numerical analysis is also evident when the investigation is extended to more complex systems made up of more implants connected by bridgework. Such a case is reported in Fig. 7, which shows virtual solid model and the corresponding finite element model of two implants inserted in a premolar region of a mandible and connected by a bar with a cantilever in the mesial side. The magnitude displacement fields in the mandible bone induced by occlusal loads acting on the bridgework are depicted in Fig. 8. The main interest of this analysis is to evaluate the mechanical function of the bone-implant system, depending on the type of loading. It is interesting to estimate the worst loading condition, usually represented by intrusive or transversal forces acting on the end of the cantilever and inducing large stress on the mesial side of the nearest implant to the cantilever. This can be correlated to the risk of a partial crestal bone loss that is sometimes remarked during clinical follow-up.^[29,35] The

same estimation can be made for the case of an induced misfit between implants and bar. It is noteworthy to remark that the effects of stress induced by misfit and those correlated to occlusal loads can be coupled, thus analyzing the behavior for a very general condition. It is evident that this capability is just an example of the larger opportunities offered by numerical analysis with respect to the experimental testing.

It is essential that the model must be accurately defined in terms of geometric configuration, material properties, and loading conditions. For instance, an important task concerns the definition of the mechanical properties of the peri-implant bone tissue, which changes its morphology during the healing time after the surgical insertion.^[36,37] Experimental data are necessary to define the elastic and strength properties of the bone tissue, including anisotropy characteristics.^[38] In this way, numerical investigation allows for a reliable and detailed definition of the mechanical behavior of the bone implant interface, to characterize the interaction phenomena between bone tissue and the surface of the implant.^[39,40] Moreover, numerical analysis can represent also the basis for a new advanced approach to the design of dental implant to have a better understanding of the biomechanical response.

CONCLUSIONS

The relevance of mechanical aspects in dental area is confirmed by the interest of international research community. It is essential to have a strong cooperation between researchers acting in different areas to be able to reach a better understanding of problems that appears to be very complicated, offering solutions that must be directly transposed to clinical practice. It is necessary to perform experimental and numerical analyses, to consider the results for a reciprocal validation procedure, and to combine medical and

mechanical competences to perform both basic and applied research. This cooperation will determine a consistent improvement of the general knowledge in dental area and it will provide, in particular, better reliability of biomedical devices as far as biomechanical aspects are concerned, also with the aim of addressing the results to industrial activity.

REFERENCES

- Berkovitz, B.K.B.; Moxham, B.J.; Newmann, H.N. *The Periodontal Ligament in Health and Disease*; Mosby-Wolfe: London, 1995.
- Mühlemann, H.R. Ten years of tooth mobility measurements. *J. Periodontol.* **1960**, *31*, 110–122.
- Roberts, W.E.; Huja, S.; Roberts, J.A. Bone modelling: biomechanics, molecular mechanisms, and clinical perspectives. *Semin. Orthod.* **2004**, *10* (2), 123–161.
- Schroeder, H.E. *Orale Strukturbiologie*; Georg Thieme Verlag: Stuttgart, 1987.
- Nishihira, M.; Yamamoto, K.; Sato, Y.; Ishikawa, H.; Natali, A.N. Mechanics of periodontal ligament. In *Dental Biomechanics*; Natali, A.N., Ed.; Taylor & Francis: London, 2003; 20–34.
- Dorow, C.; Krstin, N.; Sander, F.G. Experiments to determine the material properties of the periodontal ligament. *J. Orofac. Orthop.* **2002**, *63*, 94–104.
- Maurel, W.; Wu, Y.; Magnenat Thalmann, N.; Thalmann, D. *Biomechanical Models for Soft Tissue Simulation*; Springer: Berlin, 1998.
- Pini, M.; Zysset, Ph.; Botsis, J.; Contro, R. Tensile and compressive behaviour of the bovine periodontal ligament. *J. Biomech.* **2004**, *37* (1), 111–119.
- Dorow, C.; Krstin, N.; Sander, F.G. Determination of the mechanical properties of the periodontal ligament in a uniaxial tensional experiment. *J. Orofac. Orthop.* **2003**, *64*, 100–107.
- Parfitt, G.J. Measurement of the physiological mobility of individual teeth in an axial direction. *J. Dent. Res.* **1960**, *39* (3), 608–618.
- Mühlemann, H.R. The role of interdental contact points and of activation on tooth mobility. *J. Periodontol.* **1955**, *25*, 125–137.
- Dorow, C.; Schneider, J.; Sander, F.G. Finite element simulation of in vivo tooth mobility in comparison with experimental results. *J. Mech. Med. Biol.* **2003**, *3* (1), 79–94.
- Commer, P.; Bourauel, C.; Maier, K.; Jäger, A. Construction and testing of a computer-based intraoral laser scanner for determining tooth positions. *Med. Eng. Phys.* **2000**, *22*, 625–635.
- Andersen, K.; Mortensen, H.T.; Pedersen, E.H.; Melsen, B. Determination of stress levels and profiles in the periodontal ligament by means of an improved three-dimensional finite element model for various types of orthodontic and natural force. *J. Biomed. Eng.* **1991**, *13*, 293–303.
- Natali, A.N.; Pavan, P.G.; Scarpa, C. Numerical analysis of tooth mobility: formulation of a non-linear constitutive law for the periodontal ligament. *Dent. Mater.* **2004**, *20* (7), 623–629.
- Pietrzak, G.; Curnier, A.; Botsis, J.; Scherrer, S.; Wiskott, A.; Belser, U. A nonlinear elastic model of the periodontal ligament and its numerical calibration for the study of tooth mobility. *Comput. Methods Biomech. Biomed. Eng.* **2002**, *5* (2), 91–100.
- Natali, A.N.; Pavan, P.G.; Carniel, E.L.; Dorow, C. A transversally isotropic elasto-damage constitutive model for the periodontal ligament. *Comput. Methods Biomech. Biomed. Eng.* **2003**, *6*, 329–336.
- Natali, A.N.; Pavan, P.G.; Carniel, E.L.; Dorow, C. Visco-elastic response of the periodontal ligament: an experimental-numerical approach. *J. Connect. Tissue Res.* **2004**, *45*, 222–230.
- Natali, A.N.; Pavan, P.G.; Schrefler, B.A.; Secchi, S. A multi-phase formulation for biomechanical analysis of periodontal ligament. *Meccanica* **2002**, *15*, 345–352.
- Ross, G.G.; Lear, C.S.; DeCou, R. Modelling the lateral movement of teeth. *J. Biomech.* **1976**, *9*, 723–734.
- Kawarizadeh, A.; Bourauel, C.; Zhang, D.; Götz, W.; Jäger, A. Correlation of stress and strain profiles and the distribution of osteoclastic cells induced by orthodontic loading in rat. *Eur. J. Oral Sci.* **2004**, *112*, 140–147.
- Albrektsson, T.; Sennerby, L. State of the art in oral implants. *J. Clin. Periodontol.* **1991**, *18*, 474–481.
- Sahin, S.; Cehre, M.C. The significance of passive framework fit in implant prosthodontics: current status. *Implant Dent.* **2001**, *10*, 895–92.
- Kasemo, B. Biocompatibility of titanium implants: surface science aspects. *J. Prosthet. Dent.* **1983**, *2*, 335–380.
- Brånemark, P.I.; Zarb, G.; Albrektsson, T. *Tissue Integrated Prostheses*; Quintessence: Chicago, 1985.
- Klein, C.P.A.T.; Patka, P.; Wolke, J.G.C.; de Bleeck-Hogervost, J.M.A.; de Groot, K. Plasma-sprayed coatings of tetracalcium-phosphate, hydroxyapatite and TCP on titanium alloy: an interface study. *J. Biomed. Mater. Res.* **1991**, *25*, 53–65.
- Brunski, J.B. Biomechanical forces affecting the bone-dental implant interface. *Dent. Mater.* **1992**, *10*, 153–201.
- Merickse-Stern, R.; Piotti, M.; Sirtes, G. 3-D in vivo force measurements on mandibular implants supporting overdentures. *Clin. Oral Implants Res.* **1996**, *7*, 387–396.
- Rangert, B.; Krough, P.H.; Langer, B.; Van Roekel, N. Bending overload and implant fracture: a retrospective clinical analysis. *Int. J. Oral Maxillofac. Implants* **1995**, *10*, 326–334.
- Hart, R.T. Bone modeling and remodeling: theories and computation. In *Bone Mechanics Handbook*; Cowin, S.C., Ed.; CRC Press: Boca Raton, 2001; 1–42.
- Hoshaw, S.J.; Brunski, J.B.; Cochran, G.V.B. Mechanical loading of Brånemark implants affects interfacial bone modelling and remodelling. *Int. J. Oral Maxillofac. Implants* **1994**, *9*, 345–360.
- Soncini, M.; Pietrabissa, R.; Natali, A.N.; Pavan, P.G.; Williams, K.R. Testing the reliability of dental implant devices. In *Dental Biomechanics*; Natali, A.N., Ed.; Taylor & Francis: London, 2003; 111–131.

33. Carr, A.B.; Gerard, D.A.; Larsen, P.E. The response of bone in primates around unloaded dental implants supporting prostheses with different levels of fit. *J. Prosthet. Dent.* **1996**, *76*, 500–509.
34. Natali, A.N.; Pavan, P.G. Numerical approach to dental biomechanics. In *Dental Biomechanics*; Natali, A.N., Ed.; Taylor & Francis: London, 2003; 211–239.
35. Rodriguez, A.; Aquilino, S.; Lund, P.; Ryther, J.; Southard, T. Evaluation of strain at the terminal abutment site of a fixed mandibular prosthesis during cantilever loading. *Int. J. Prosthodont.* **1993**, *2*, 93–102.
36. Chang, M.C.; Ko, C.C.; Liu, C.C.; Douglas, W.H.; DeLong, R.; Seong, W.J.; Hodges, J.; An, K.N. Elasticity of alveolar bone near dental implant-bone interfaces after one month's healing. *J. Biomech.* **2003**, *36*, 1209–1214.
37. Brånemark, R.; Öhrnell, L.O.; Nilsson, P.; Thomsen, P. Biomechanical characterization of osseointegration during healing: an experimental in vivo study in the rat. *Biomaterials* **1997**, *18*, 969–978.
38. Natali, A.N.; Hart, R.T.; Pavan, P.G.; Knets, I. Mechanics of bone tissue. In *Dental Biomechanics*; Natali, A.N., Ed.; Taylor & Francis: London, 2003; 1–19.
39. Natali, A.N.; Pavan, P.G. A comparative analysis based on different strength criteria for evaluation of risk factor for dental implants. *Comput. Methods Biomech. Biomed. Eng.* **2002**, *5*, 511–523.
40. Hansson, S.; Norton, M. The relation between surface roughness and interfacial shear strength for bone-anchored implants. A mathematical model. *J. Biomech.* **1999**, *32*, 829–83.

Dental Implants

Nirit Tagger-Green

Eli E. Machtei

Unit of Periodontology, Department of Oral and Maxillofacial Surgery, Rambam Medical Center and The Faculty of Medicine, Haifa, Israel

INTRODUCTION

Dental implants are one of the most important revolutions introduced to the field of dentistry in the past 50 years. The concept of implants replacing removable dentures was more readily accepted by patients than by their dentists. Dentists received the new mode of therapy only after long-term clinical data based on research and scientific data were presented.

Today, after 50 years of experience, almost every dental practice—even general practitioners—use dental implants on a daily basis. Treatment that was once preserved for highly selective cases is now being widely used. The costs of dental implants have decreased substantially, and new treatment modalities are now often introduced.

HISTORY OF DENTAL IMPLANTS

Dental implants mimicking natural tooth roots were invented thousands of years ago by the ancient Chinese, Egyptians, and Incas.^[1] The modern era of root form implants began only in the 19th century. Two types of implants from the past are still functional and rarely used, the blade form implant and the subperiosteal implant. The blade-vent implants have become less popular due to low survival rate and the formation of connective tissue around the implant.^[2] The subperiosteal implants may have a long-term survival rate but exhibit soft tissue problems and fractures of the metal framework.^[3]

Use of root form implants began after the pioneering research of Branemark and his coworkers. They inserted submerged titanium implants with a machined surface in the jawbone of dogs. In 1969, they published the first evidence of direct bone anchorage named osseointegration.^[4] A few years later, they published results of 10–15 years of clinical experience with the osseointegrated implants.^[5,6] Branemark's team was the first to suggest the idea of osseointegration between implants and bone, but they could not show histologic evidence because there were no methods available to section intact, undecalcified bone together with the metallic

implant. Histologic sections of the implant area could be prepared only after the implant was removed—probably with some of the tissues surrounding the implant and then only could the bone be inspected.^[7] Schroeder, who was working in Switzerland independently from Branemark, used new techniques to cut through undecalcified bone and implant without separating them first. He demonstrated very clearly direct bone-to-implant contact and called this connection functional ankylosis.^[8] Nevertheless, it is very important to emphasize the great contribution of Branemark and his team, who are considered the pioneers of the field of implantology.

Development of the Surgical Technique

The earliest modern investigations on dental implants were carried out on implants placed according to the method published by Branemark in 1977—the two-stage technique. The two-stage surgical protocol uses a two-piece implant pillar. The main reasons to this approach are to minimize the risk of infection, prevent apical downgrowth of mucosal epithelium, and minimize the risk of premature loading. In addition, a stress-free period of three to six months is recommended before piercing the mucosa with the abutments. Such a stress-free period was considered to be an absolute prerequisite for successful osseointegration.^[9]

Over the years, Branemark's original protocol was reevaluated. Schroeder et al. were the first to show the possibility to achieve osseointegration without submerging the implant. During the last decade, reports have demonstrated a successful clinical treatment outcome using the one-stage surgical approach. The ITI group proposed the use of one-stage surgical technique for oral implantology (Waldenburg, Switzerland). This technique involves nonsubmerged implants and loading, which usually starts earlier than in the submerged technique. Recent developments have brought implants into the next step, immediate loading. This technique not only inserts the implants transgingivally, but it also loads them at the same time. Immediate loading raises the problem of micromovement that, in excess of 100 μm , can induce fibrous tissue formation at the bone-to-implant interface instead of the desired bone

formation.^[10] In order to prevent micromovement during the initial phase of healing, new designs of implants and modification of the surface are being invented. The common aim of the new designs is to try and achieve the best initial stability possible. The results from recent studies indicate predictable results for treating partial or total edentulism with dental implants using the one-stage, immediate-loading technique. According to some studies, the results can be comparable with the success rate of implants placed according to the original protocol. Patient selection is probably the most important factor in achieving high success rates. Adequate dimensions of the alveolar process are mandatory, and patient cooperation should be at the highest level. Multicenter studies are necessary before this treatment modality can be recommended for general use^[11] (Figs. 1 and 2).

OSSEOINTEGRATION

Osseointegration represents a direct connection between bone and the implant without interposed soft tissue layers. A 100% bone-to-implant contact does not occur, and there is neither agreement as to the extent of bone-to-implant contact required for acceptance of the connection as osseointegration nor even on the criteria for definition of the term. Histologically, osseointegration is defined as bone in direct contact with an implant seen at the magnification of the light microscope.^[12] Another definition is based on clinical stability rather than on histologic criteria: "... a process whereby clinically asymptomatic rigid fixation of alloplastic materials is achieved in bone and maintained during functional loading."^[13]



Fig. 1 Clinical picture of dental implants inserted in the traditional method, with cover screws, 4 months after the insertion of the implants. (Courtesy of MIS (Medical Implant System), Shlomi, Israel.)



Fig. 2 A panoramic X-ray demonstrating implants of different methods in both jaws. (Courtesy of MIS (Medical Implant System), Shlomi, Israel.)

In 1981, Albrektsson et al. presented a series of factors that need to be controlled to achieve osseointegration:

1. Biocompatibility.
2. Design.
3. Surface condition of the implant.
4. Status of the host bed.
5. Surgical technique.
6. Loading conditions applied afterwards.^[14]

Osseointegration shares many prerequisites with primary bone fracture healing, such as precise alignment, primary stability, and adequate loading during the healing period. Osseointegration requires a bioinert or bioactive material and a surface configuration that are osteophilic, i.e., attractive for bone deposition. Surface properties of implants influence the elaboration of bone-to-implant contact. The assumption is that a rough surface improves adhesive strength compared with a smooth one. A smooth surface transmits only compressive forces with minor resistance against shear. Mild roughness ($<10\ \mu\text{m}$) augments the resistance against shear. In order to obtain adhesion, there is a need for a chemical bond or microporosity.^[15] Macroporosity ($100\text{--}500\ \mu\text{m}$) favors bony ingrowth and is used in porous coating.

Stages of Osseointegration

Several phases can be distinguished in the process of osseointegration:

- (a) Direct bone healing. This phase is activated by lesion of the pre-existing bone matrix. When the matrix is exposed to extracellular fluid,

noncollagenous proteins and growth factors are set free and activate bone repair.

- (b) Incorporation by woven bone formation. The first type of bone that is formed is woven bone. It is considered a primitive type of bone and is the major type that is formed during the first four to six weeks after surgery.
- (c) Adaptation of bone mass to load. This phase starts at the second month. The microscopic structure of bone changes towards lamellar or parallel-fibered bone. Bone remodeling characterizes the last stage of osseointegration. It starts around the third month and continues for the rest of life. Remodeling is the last stage of osseointegration, and it improves bone quality by replacing necrotic bone and the primitive woven bone. Remodeling also leads to functional adaptation of the bone structure to load by changing the dimensions and orientation of the supporting elements.

IMPLANT DESIGN

Implant design refers to the three-dimensional structure of the implant, including the shape, form, configuration, surface macrostructure, and macroirregularities. The objective of all the designs is to create an implant with a long-term survival and high success. The huge variety of implants available today can be classified in a number of different ways. One possibility is classification based on the interface between the implant and the abutment. Another is based on the shape of the body of the implant and the implant-to-bone surface.^[16]

The Interface Between Implant and Abutment

This interface is described as an internal or external connection. The most common external connection is the hexagonal connection or external hex. The aim of that design is to have an antirotational effect on the prosthetic restoration. There are variations in the height and width of the hex. Other forms of external connection are the octagonal top or octa and the spline interface. The internal connection includes the Morse taper interface and the internal hexagonal and internal octagon.

The connection can be further characterized as a slip-fit joint, in which there is a slight space between the mating parts and the connection is passive, or as a friction-fit joint, in which there is no space between the components. They are actually forced together. Reviewing the prosthetic possibilities and various connections available is beyond the scope of this introduction.^[17]

The Shape of the Implant

The geometry of the body of the endosteal implant can be classified into three basic shapes, a threaded screw (ad modum Branemark), a press-fit cylinder (ad modum IMZ), and a hollow basket cylinder (ad modum ITI). Implants are further categorized as being threaded or nonthreaded and as solid or hollow cylinder.

Changing the geometry of an implant can induce dramatic changes in the functional surface area. The functional surface area is defined as the amount of thread that participates in compressive and tensile load transmission under the action of an axial occlusal load. As the bone density decreases, increasing the surface area can offset the potential increase in stress. Anatomical landmarks, machining limitations, and engineering constraints must be considered when choosing a design. For example, the height and width of the implant must fit the anatomical landmarks of the maxilla and mandible. On the other hand, the minor diameter must be large enough to prevent fatigue fracture.

Implant's threads

Threads are used to maximize initial contact, improve initial stability, enlarge implant surface area, and improve the distribution of forces. The depth, thickness, pitch, face angle, and the helix angle of the threads are all parameters that define the functional thread surface and affect the biomechanical load distribution of the implant. Thread pitch is the distance between two adjacent threads. Decreasing the pitch increases the number of threads on the implant; thus, this increases the surface area of the implant's body. Changes in pitch are not constrained by the anatomy of the jaws, and they create a big increase in the surface of the implant without noticeably modifying the dimensions of the implant.

Thread shape is another important factor in the design of the implant. There are three common shapes of threads, a v-thread, a square thread, and a buttress thread. The v-thread results in a tenfold increase in the in-shear resistance at the implant-to-bone interface compared with the square thread design. Thread depth can also be altered along the implant so that the surface of the implant will be increased or decreased in high- and low-load areas. This design increases the functional surface area in the crestal region. The load transfer is reduced and transferred throughout the length of the implant. Thread shape and functional surface area are becoming more important with increasing numbers of practitioners using the immediate-loading procedure. These are particularly important when anatomical considerations dictate the placement of short implants^[18] (Fig. 3).

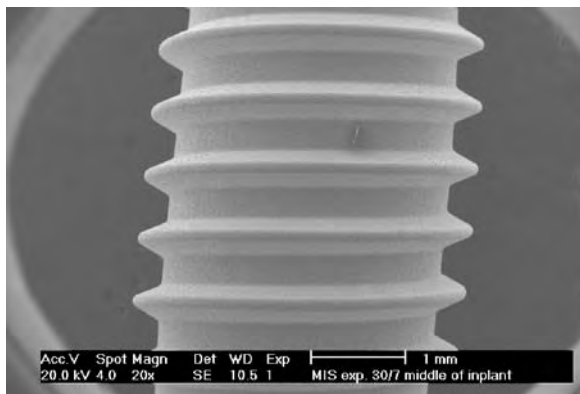


Fig. 3 $\times 20$ magnification of implant threads. (Courtesy of MIS (Medical Implant System), Shlomi, Israel.)

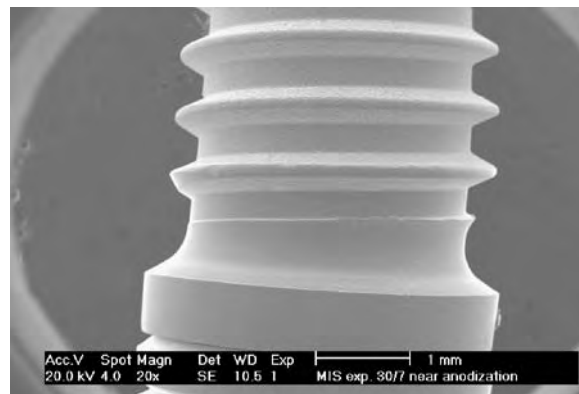


Fig. 4 $\times 20$ magnification of the base of the implant. (Courtesy of MIS (Medical Implant System), Shlomi, Israel.)

The crest module

The crest module of the implant is the transosteal region of the implant and is characterized as a region of highly concentrated mechanical stress. Bone loss due to the design of the area is a common finding if it is not well designed for load bearing. Bone loss is so common that the manufacturers prepare this area to a very smooth finish so that, when bone loss occurs, plaque accumulation will be reduced.

The crest module should be slightly larger than the outer thread diameter. Thus, the crest module seats completely over the osteotomy. The seal created by the larger crest module also helps to achieve greater initial stability of the implant. A larger diameter also increases the surface area, which, in turn, contributes to decreases in stress at the crestal region. A polished collar of minimum height should be designed on the upper (occlusal) part of the crest module—just under the prosthetic part. In implants with a longer polished collar, bone is subjected to excessive shear loading. This overload is manifested by significant bone loss in the crestal region attributed to the lack of effective mechanical interlocking between the smooth coronal region and the bone.

A common clinical finding is bone loss down to the first thread. This bone loss cannot be attributed to any anatomic feature because the distance to the first thread varies from 1.2 to 3.0 mm, according to the brand of implant. An improved design or surface condition of the crest module can reduce the risk for crestal bone loss (Fig. 4).

Implant length

As the length of the implant increases, the overall surface increases as well. A common axiom was to insert as long an implant as possible, but this is being

reevaluated. Quite often, the bone that requires more stabilization and a larger surface area is of bone type D3 or D4. This bone exists mainly in posterior areas of the jaws in which the height of the bone is not sufficient for longer implants. There is a minimum implant length for each bone density, depending on the width and design. As a rule, the softer the bone, the longer should the implant be. Usually, the bone in the posterior areas of the maxillae is softer, and, if necessary, in the mandible repositioning of the inferior alveolar nerve is possible while in the maxilla augmentation of the floor of the maxillary sinus may be needed.

Apical design considerations

Most root form implants are circular in cross section. The problem with the round cross section is that it does not resist torsion or shear forces, and a single round implant does not resist rotational force. As a result, an antirotational feature in the form of a vent or a hole is incorporated, which is usually in the apical region of the implant body. A hole in the apical area increases the surface area, but, when inserting implants simultaneously with sinus elevation, the hole may be filled with fibrous tissue or with mucosa and thus become a source of retrograde contamination. Another way to resist rotation is to create flat sides or grooves along the implant or at its apical part only. Bone that grows against these features prevents rotation of the implant.

A few implants have an active apical part, which means that the implant itself helps while inserting it. This is especially helpful in dense bone (types D1–D2) (Fig. 5).

Implant Design and Distribution of Forces

Finite element analysis (FEA) has been used to show that stresses in cortical bone decrease are in inverse

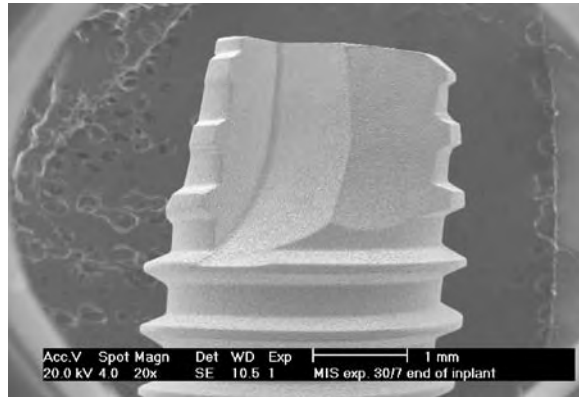


Fig. 5 $\times 20$ magnification of the apical part of the implant. (Courtesy of MIS (Medical Implant System), Shlomi, Israel.)

proportion to an increase in implant diameter with both vertical and lateral load. Using the widest implant is not necessarily the best choice when considering stress distribution to surrounding bone. Occlusal forces are distributed primarily to the crestal bone rather than evenly throughout the entire surface of the implant. In summary, the optimum length and diameter necessary for long-term success depends on the bone support condition. If the bone is in normal condition, the length and width of the implant are not significant factors. But, if the bone is of poor quality, large diameter implants are recommended.

Polished necks of implants do not osseointegrate. This characteristic was investigated by FEA when comparing the bonding created to a smooth neck to that found on implants with retentive elements. Retention elements at the implants neck dramatically decrease the peak interfacial shear stress. The retention elements may counteract marginal bone resorption in accordance with Wolff's law on bone remodeling (Fig. 6A–C).

Surface Characteristics

The study of biomedical implant surfaces and the effect of surface modification have become popular in recent years because of the notion that surface characteristics directly influence the interaction between the bio-material and the tissue. Direct apposition of bone to commercially pure titanium (cpTi) surface results in cellular attachment and implant stability. These characteristics are attributed to the oxide layer that is formed spontaneously on the implant's surface. Many studies attempted to define the nature of the interaction between the oxide layer and the surrounding tissue. Studies were also dedicated to investigating the biologic effects of modifying the surface. The effect of surface roughness on the formation of the oxide layer and cellular adhesion has also been studied.

The oxide layer can exhibit microscopically smooth or rough topographies, depending on the cleaning and passivating process. Surface macroscopic roughness is

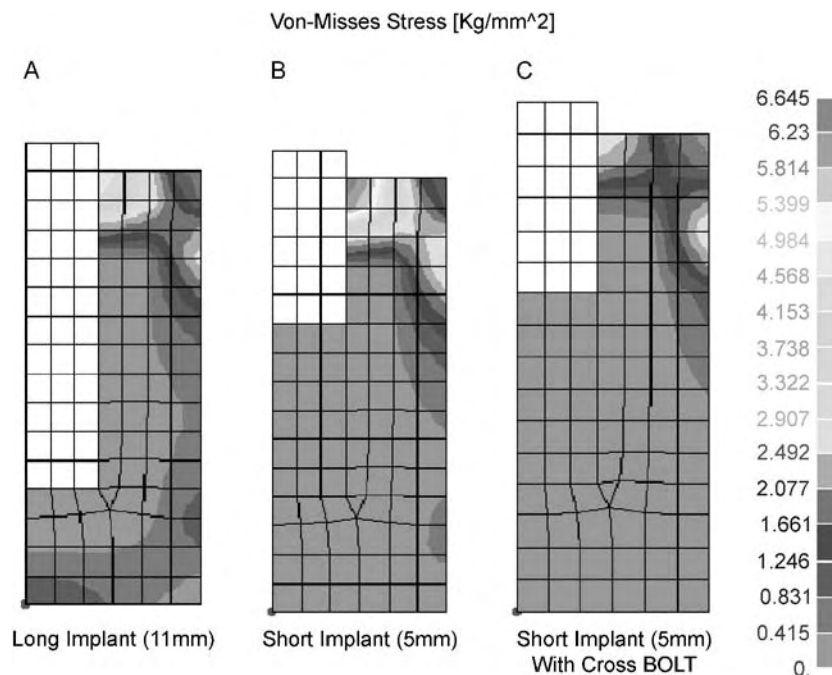


Fig. 6 (A) FEA of a long implant (11 mm). (B) FEA of a short implant (5 mm). (C) FEA of a short tri-cortical implant (5 mm). (Courtesy of MIS (Medical Implant System), Shlomi, Israel.)



normally introduced into the substrate beneath the oxide by mechanical grinding, resorbing by particulate blasting, or a chemical procedure (acid etching). The thickness of the oxide layer remains rather constant since the moment that it is first exposed to air.

Oxide modification may occur during in-vivo exposure. The highest oxide thickness is observed in close relation to bone marrow while the lowest growth appears in contact with cortical bone. Increased quantities of calcium and phosphorous are found in the oxide layer and are the result of ion exchange at the interface.

Microstructure of the implant's surface is important because variations in the surface microstructure influence the stress distribution, retention of the implant in bone, and cell response in the implant. Several authors have discussed the ideal roughness that would provide increased retention and an improved bone response. Roughness could be considered on different levels, macrostructural, microstructural, and ultrastructural. Roughness on these different levels has different effects on living tissues. In order to gain complete growth of bone into a material's irregularities, they have to be at least 100 μm in size. Growth of bone into cavities or pores of this size will give a mechanical interlocking of the material with bone. Bobyn et al. investigated the optimal pore size to get bone ingrowth. They have tested alloys with a pore size of 50–400 μm and found that optimal results are achieved with a pore size of 100 μm .^[19]

The optimal pore size to obtain bone ingrowth depends also on the material. Titanium alloy implants have an optimal pore size of 100 μm . Increased removal torques were found when using titanium-coated screws of stainless steel with a pore size of 10–40 μm . These small pores do not allow maximum ingrowth of bone but provides increased retention based on mechanical interlocking. There are a few mechanisms, besides mechanical interlocking that determine the reaction between bony tissues and biomaterials. Studies in vitro show that surface topography influences the function of cells, matrix deposition, and bone mineralization. Cells are sensitive to microtopography and may use the morphology of the material for orientation and migration. Microtopography also influences the differentiation of mesenchymal cells into fibroblasts, chondroblasts, or osteoblasts.

MATERIALS USED FOR DENTAL IMPLANTS

The choice of metals for dental implants is based on a several factors, including biomechanical properties, previous experience with working and handling the material, and the possibility to sterilize the final product. The implant is a foreign body and does not

behave like a living tissue. Indeed, some metals and metal alloys provoked adverse tissue reaction, and their long-term survival was short. Corrosion is the visible destruction of metals. The corrosion of new metals used in surgery is very small, and very sensitive techniques are needed to detect it. Noble metals like gold and silver have an acceptable resistance to corrosion in air, i.e., about in the middle of a practical scale, but, in physiologic solution, their resistance to corrosion is much lower.

Titanium is a reactive metal. When it is exposed to air or to physiologic solutions, a dense oxide film forms immediately at its surface. The oxide is primarily titanium dioxide with small quantities of Ti_2O_3 and TiO .

This film has a thickness of 4 nm. Its mode of growth is specific in that the oxygen ions migrate towards the metal and react with the titanium at the base of the oxide. This mechanism is unique to titanium, silicon, and zirconium. All are bases with a valence of IV. The advantage of that mode of oxidation is unique in that it does not release metal ions into the electrolyte. This surface feature creates a potent barrier against dissolution of the metal. The oxide on surgical implants is primarily amorphous if it is formed at normal temperature. If cpTi is processed at elevated temperatures or anodized at high voltage, the oxide forms a crystalline atomic structure and can be much thicker.

In serum, the concentration of titanium is about 0.1 μm and, in human skeletal muscle, it gets up to 5 μm . The concentration of titanium in the tissue around implants is about 100 times higher than that of muscle tissue. In spite of this high concentration, no case of systemic or local reaction to titanium has been documented.

Titanium and titanium alloys are the material of choice for dental implants, but the prosthetic parts are usually made of gold alloys, stainless steel, nickel-chromium, and cobalt-chromium alloys. Consequently, there is a chance to have a galvanic action between the dissimilar metals with possible effects on the electrochemical corrosion, oxidation, and even triggering pain.

The stoichiometric composition of commercially pure titanium (cpTi) allows its classification into four grades that vary in the oxide content. Grade 4 has the highest oxide content (0.4%), and grade 1 has the lowest content (0.18%). The properties of the oxide layer are not altered, but there are mechanical differences between the grades. The differences are attributed to the contaminants that are present in very small quantities. Other elements that are found are nitrogen, carbon, hydrogen, and iron. All improve the stability and mechanical or physicochemical properties of the implant. Iron improves the resistance to corrosion, and aluminum increases the strength and decreases the density.

Properties of the Material and Finite Element Analysis (FEA)

Properties of the material greatly influence the stress and strain distribution in a structure. These properties can be modeled in Finite Element Analysis (FEA) as isotropic, transversely isotropic, orthotropic, and anisotropic. In an isotropic material, the properties are the same in all directions. An anisotropic material has different properties in different directions. The assumption is that most of the materials are homogeneous and linear plus they have elastic material behavior characterized by two material constants of Young's modulus and Poisson's ratio. One of the difficulties in calculating FEA is that the surrounding bone is neither homogeneous nor isotropic. These characteristics of the bone present difficulties in assessing the values for ultimate strain and modulus of elasticity during stress or strain.

Surface of the Implant

Hydroxyapatite coating

There is significant controversy concerning the benefit of hydroxyapatite (HA) coating of titanium implants. Some of the reports show the benefit of the HA coating while others claim that the coating compromises the long-term survival of the implant. In-vivo and in-vitro studies have shown that HA is osteoconductive and promotes bone healing at gaps between the implant and the bone.^[20] A retrieval analysis of HA-coated implants revealed that the HA was biocompatible and bonded the surrounding bone without dissolution of the coating. Others clinicians observed that the coating dissolved and separated from the metal core of failing implants. Advocates of HA claim that the coating is stable in a physiologic pH. Some claim that the coating does not provide any benefit over the conventional uncoated implants. It even worsens the survival of implants since microorganisms may adhere to the coating and cause the implant to fail. The results of a metaanalytic review revealed that many of the studies concerning HA-coated implants are anecdotal. Studies were conducted that examined in detail and evaluated the survival by overall percentage or life tables. The survival rates of the coated implants were in the range of survival as the titanium implants. If degradation of the coating results in implant failure, the yearly interval survival rates should decrease with successive follow-ups. Yet, progressive decrease of the yearly interval survival rate was not observed in any study.^[21]

Machined surface

Machined implants exhibit an irregular surface with grooves ridges and pits. Proponents of that system

claim that this is the most conductive surface for cell attachment. Most of the literature reports are based on the average surface roughness (Sa). Sa smaller or equal to 1 μm is considered smooth, and those with surface roughness greater than 1 micrometer is described as rough.

Machined (turned) cpTi is a smooth surface with an Sa value of 0.53–0.96 μm , depending on the manufacturer's protocol, grade of the material, and shape and sharpness of the cutting material. Circumferential parallel lines of 0.1 μm depth and width perpendicular to the long axis of the implant are a common finding in the machined surface.

Plasma-spray coating

Plasma-spray coating (TPS) is one of the most common methods for surface modification. Plasma-spraying may be used on titanium or over HA surface. Porous or rough titanium surface has been produced by plasma-spraying a powder form of molten drops at high temperatures. At temperatures around 15,000 degrees, argon plasma provides high velocity partially molten particles of titanium powder. After fusion, the plasma sprayed layer gets the thickness of 0.04–0.05 mm. Microscopically, the coating shows round or irregular pores that can be connected to each other.

Microtextured titanium (MTX)

A new surface of dental implant is being used by Centerpulse. The MTX surface is grit-blasted with a soluble, biocompatible medium. The implant is then cleansed with a mild acid wash that removes any residual microblast powder. The MTX process produces a uniform 1–2 μm surface roughness that exhibits 44% more surface area than machined titanium. The roughened surface is proven to enhance bone attachment. The MTX process does not weaken the titanium implant body and ensures that the strength of the implant is maintained. Implant threads are not rounded by the MTX process, and cutting grooves remain intact for efficient self-tapping.

Other surface modification

Blasting with particles of varying diameters is another method for surface alteration. In this approach, the implant is bombarded with particles of aluminum oxide (Al_2O_3) or titanium oxide and by abrasion. Among the titanium surfaces, both blasted and TPS surfaces showed a slight increase in bone contact over machined surface. Etching with hydrochloric or sulfuric acid after particle blasting resulted in a more

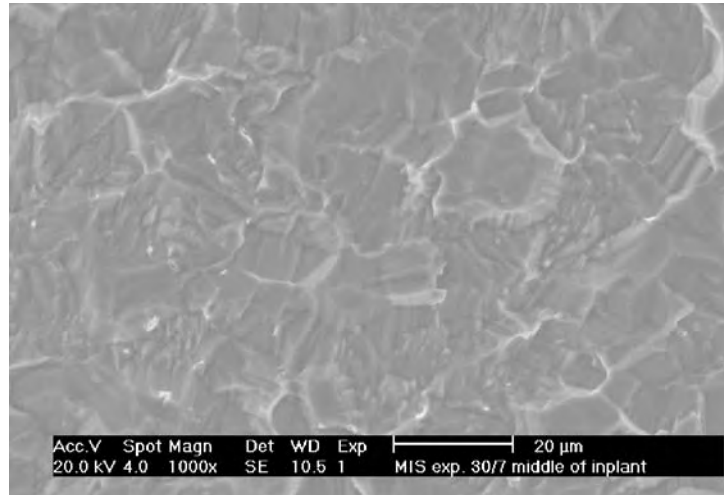


Fig. 7 $\times 1000$ magnification of a dual-etched surface. (Courtesy of MIS (Medical Implant System), Shlomi, Israel.)

dramatic increase in the bone contact level. The dual-etched surface treatment resulted in a 3.5 increase in mechanical pull-out force compared to the original machined surface. Studies *in vivo* showed that bone growth into chambers maintained contact with the titanium surface. In contrast, the machined surface exhibited lifting of the clot from the wall, which resulted in less bone contact (Fig. 7).

Lazzara et al., in a split implant study, evaluated the difference between machined titanium and dual-etched surfaces. They demonstrated that the dual-etched surfaces showed increased bone contact six months after insertion of the implant.^[22] In a study conducted by London et al., histologic comparison was carried out between dual-etched implants to machined TPS and HA surfaces. They compared the bone contact *in vivo* in rabbits. They found that the bone to implant contact varied from 3% after one week for the TPS and up to 60% after four weeks for the dual-etched. DE surfaces showed the highest bone-to-implant contact over the eight weeks. The machined and the TPS implants showed the least bone-to-implant contact. Each of the implant surfaces showed an increase in bone contact percentage over time. After eight weeks, the machined TPS and DE surfaces showed similar bone contact levels.^[23]

An interesting recent innovation in surface technology is the combination of two or more surfaces on the same implant body. The rationale is to achieve improved soft tissue response, stability, and attachment in cortical bone with a machined or etched coronal implant surface and better mechanical locking in medullar bone with a roughened TPS or HA surface in the middle to apical portion of the implant. One design even includes four different surfaces on the same implant body.

Evaluation of the Interface

There are two methods to assess the quality of the osseointegrated interface, biomechanical test and histomorphometric analysis. There are usually three types of biomechanical tests, pull-out, push-out, and torque (Figs. 8 and 9).

Pull-out measurement. The stress values are estimated by dividing the failure load by the total implant surface area. Wong et al. tested the shear strength as a function of various implant surfaces. The shear strength was tested by push-out and pull-out measurements. They showed a nearly linear relationship between the push-out load to failure and surface roughness in microns.^[24]

Push-out test. When performing this test, both the coronal and the apical parts of the implant must be free

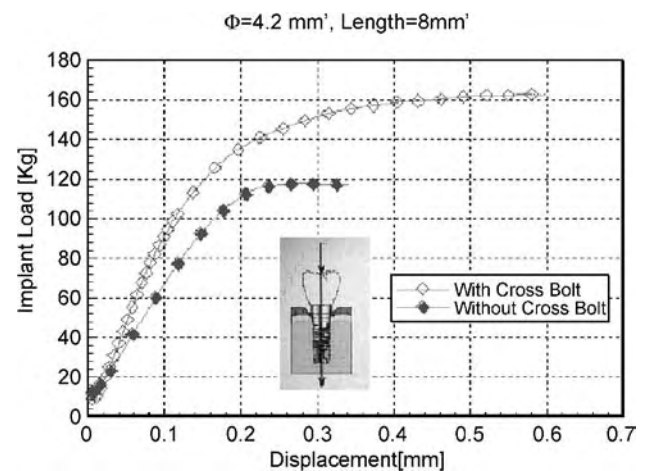


Fig. 8 Graph illustrating axial implant load. (Courtesy of MIS (Medical Implant System), Shlomi, Israel.)

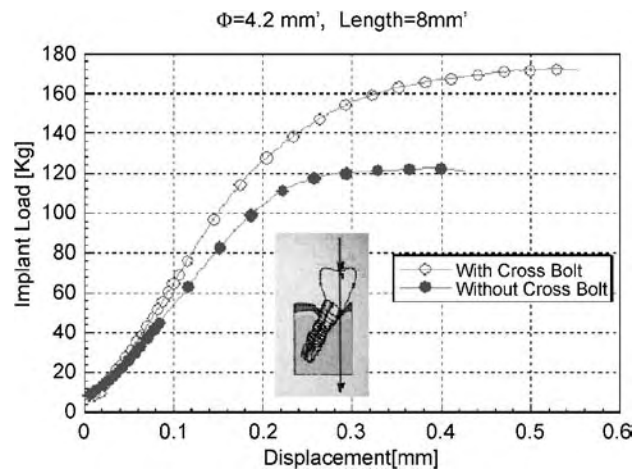


Fig. 9 Graph illustrating 30° implant load. (Courtesy of MIS (Medical Implant System), Shlomi, Israel.)

of bone contacts. If not cleared, high forces are necessary to resist the bone structure between the threads at the apical end of the implant. The coronal portion accepts the applied push-out force, but the apical end must be exposed to allow smooth and free extrusion of the implant. Push-out and pull-out tests are indicated when measuring cylindrical and press-fit implants. Threaded implants are tested more effectively by the torque test.

Torque. The moment of a force about a point tends to produce rotation or bending about that point. This moment refers to as torque or torsional load. This force is the most destructive for the implant system. Torque may be imposed on implants as a result of long cantilevers. A few authors compared various implant systems and surface treatments and found that the removal torque values increase with time.

Histomorphometry

These analyses of the bone-to-implant interface can be carried out in different ways. Investigators often present bone-to-implant contact as a percentage of the total implant length and as a percentage of the three consecutive best threads length. Implant design and material, surface treatment, healing time, and loading conditions are some of the parameters that affect the analytic approach. Thread volume fill and the number of cells in contact with the implant surface are two other important variables.

Difficulties in assessing the bone-to-implant interface by FEA

Most FEA analyses assume a state of optimal osseointegration and perfect bonding between bone and implant. In clinical situations, this doesn't usually

occur. Bone is a porous material with complex microstructure. Part of it is dense cortical bone, and the middle is a softer trabecular bone. The proportion between the two types of bone is dependent on the location of the implant in the jaw. Therefore, the model for FEA of osseointegration is a rather complex algorithm.

Survival of Dental Implants

Since the introduction of endosseous implants, many design modifications have been made. Each element in of implant design may theoretically present certain advantages and disadvantages to the clinician. The status of bone-to-implant contact is also important because it is the basis for the successful long-term osseointegration. All these parameters and many others may influence the success of the implant, but there are rather few objective-independent investigations concerning short- and long-term survival of implants.

No precise definitions of implant success or failure are available. Cases, which qualified as success in one clinical situation, may not be considered successful in a different situation. The American Academy of Periodontology determined in 1996 a list of criteria for assessing implant success. This list is compiled from various studies.^[25]

Common criteria used to determine success of implants:

1. Less than 1 mm mobility in any direction.
2. No radiolucency.
3. Bone loss not greater than 1/3 of implant.
4. Provide functional service for five years in 75% of cases.
5. Absence of persistent and/or irreversible signs and symptoms such as pain, infection, neuropathies, paresthesia, and violation of mandibular canal.
6. Clinically immobile.
7. No peri-implant radiolucency.
8. Bone loss less than 0.2 mm annually after first year of service.
9. 85% success after five years; 80% success after 10 years.
10. Implant design allows satisfactory restoration to patient and dentist.
11. Implant still in mouth.
12. Absence of continuous marginal bone loss.
13. Minimal marginal bone loss.
14. Absence of persistent soft tissue complications.
15. Surgical retrievability.
16. Probing depth less than 4-5 mm; bone loss less than 4 mm.
17. No mechanical failure.

18. Bone loss does not reach apical 1/3 of implant.
19. Nonstandardized radiographs demonstrate less than 50% of implant has bone loss.

In 2000, the American Academy of Periodontology published the report of the three-year results from the Veteran Administration's (VA) multicenter prospective implant study.^[26] This report, known as the Dental Implant Clinical Research Group (DICRG), is considered to be the consensus for dental implants. The DICRG findings can be applied to most endosseous implant systems, dentists of all skills, and most patients.

The first part of the report deals with the survival and stability of six implant designs from placement to 36 months. The study was performed in 30 VA medical centers and two dental schools. More than 2900 implants were placed and restored. Survival was determined for each of the three phases of implant treatment. Phase 1 is from the time of implant placement to uncovering of the submerged device. The second stage is from uncovering to placement and loading of the prosthesis. Phase 3 is from initial loading of the prosthesis to 36 months. The stability of the implant was checked using the periotest[®]. The periotest scale ranges from -8 (clinically firm) to +50 (very loose) so that the more negative a reading, the more stable the implant. There were two approaches to determine survival of the implant. A common approach was to consider as failures only implants that failed after loading. The DICRG considered as failure every implant that had to be taken out, regardless of whether it had been loaded or not. The results show that the largest difference in survival, which includes implants that failed before loading, was for the cpTi implants (21.7%). According to the DICRG criteria, the HA cylinder implants had the highest survival rate (97.5%). When the postloading approach was used, the titanium alloy screw had the highest survival rate (99.4%). The HA cylinder was next (98.6%). The conclusions were that the HA-coated implants had the highest survival rate. Non-HA implants showed increased stability following loading while HA-coated implants showed a slight decrease or no change in stability.

The next investigation was to see if the length and diameter of the implant was related to survival. Various implant lengths were investigated—7 mm, 8 mm, 10 mm, 13 mm, and 16 mm. The diameters investigated were 3–4 mm and 4 + mm. The results were that the stability of the 3–4 mm implants was significantly smaller than that of the group less than 4 mm. The survival was significantly larger for the 16 mm implants (96.4%) than for the 7 mm implants (66.7%) ($p=0.001$).

The next step in the investigation was to assess the survival of implants that were mobile at placement. Orenstein et al. showed that out of 2641 implants that were included in the DICRG survey, 81 (3.1%) were

mobile at the time of placement and that 76 (93.8%) of these were osseointegrated at uncovering, which was compared with 97.5% of the implants that were immobile at placement. The difference was not significant. The conclusion is that implant stability at the time of placement is clearly desirable, but it may not be a prerequisite to osseointegration or to the long-term survival of the implant. HA-coating significantly improved the performance of mobile and immobile implants at placement up to three years after placement.

Truhlar et al. documented changes in stability, as reflected in the PVT values, correlated with various implant surfaces, designs, and bone density. The findings were that bone quality and surface coating of the implant did influence the implant's stability. For bone quality type I and II, the stability improved over time. But, for bone qualities type III and IV, all implants showed a slight decrease in stability over the five years following insertion. Anyway, the PTV at the time of uncovering was the best predictor to the clinical acceptance at the bone-to-implant interface. When comparing the stability of HA coated implants to noncoated implants, HA implants became slightly less stable over time while the noncoated showed an improvement in stability.

Biomechanical Aspects of Dental Implants

All structures and combinations of structures that are exposed to functional loading can be exposed to overload that can lead to mechanical complications and failures. In implant dentistry, such failures can occur in the implant itself or in the prosthetic parts.

Implant versus tooth mobility

The biomechanical situation of an osseointegrated implant is fundamentally different from that of a natural tooth. The initial deflection of an implant is linear and elastic whereas the natural tooth has an initial phase of periodontal compliance followed by a more rigid appearance at the engagement of the alveolar bone. Axial and horizontal mobility of a natural tooth has been estimated to be larger than that of an osseointegrated implant by a factor of 12–100.

The existence of such great differences between the mobility of the tooth and the implant have generated recommendations not to use rigid prosthetic connections between the tooth and the implant, and, if possible, avoid combinations of natural teeth and implants as abutments for a single prosthetic appliances. Recently presented in-vivo data have shown that the biomechanical significance of controlled superstructure misfit has indicated higher compliance for vital alveolar bone.

Force generation

Patients treated with dental implants can be exposed to a range of traumatic situations; yet, the prosthetic and surgical planning of implant placement generally considers only the physiologic loading of implants and implants superstructure. A wide range of loading variations must be taken into account. For example, Bruxism must be considered as a condition that requires special attention. It causes not only increased loading levels but also prolonged loading times and an increased number of loading cycles (Figs. 10 and 11).

In the natural dentition without implants, the mean maximal vertical bite force is of a magnitude of 469 ± 85 N at the region of the canines, 583 ± 99 N at the second premolar region, and 723 ± 138 N at the second molar. Average values of the maximal voluntary bite force are 545 N in men and 383.6 N in women. The angle between the bite force and a plane that is perpendicular to the occlusal plane is typically 3.9 degrees.^[27]

Bite forces after implant treatment

For implants used to replace a single tooth, in-vivo forces ought to duplicate the forces exerted on natural teeth. This is expected because, in both cases, biting would occur on a single, stand-alone crown. Factors such as the width of the crown's occlusal table, the height of the abutment above the bone level, and the angulation of the implant with respect to the occlusal plane affect the value of the lever moment on the implant. When dealing with a prosthesis borne by several implants linked together, leverage effects can be expected because of geometric factors relating to the

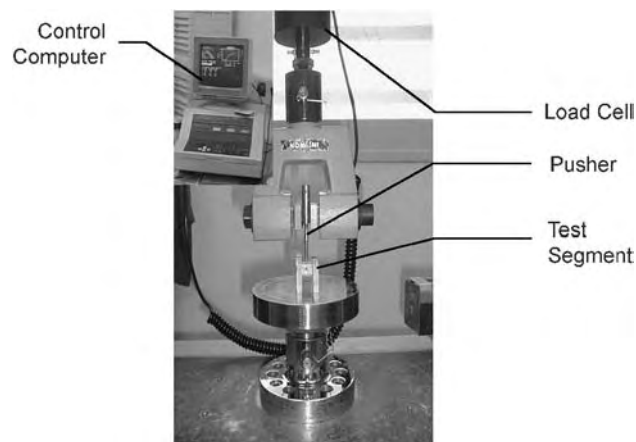


Fig. 10 Machine designed to test mechanical configuration. (Courtesy of MIS (Medical Implant System), Shlomi, Israel.)

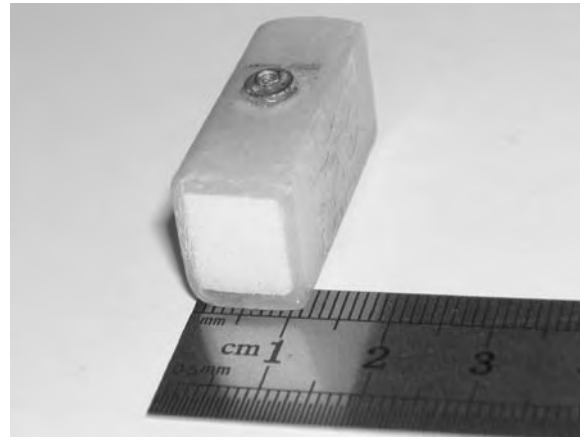


Fig. 11 Artificial bone used for mechanical testing. (Courtesy of MIS (Medical Implant System), Shlomi, Israel.)

restorations. Such factors cause the implant to be subjected to increased bending moments as well as to axial forces that can be tensile and compressive. For fixed prostheses the maximal occlusal force ranged from 35 to 330 N with the largest values at the area of the second premolar.^[28]

Dental implants are subjected to occlusal load when placed in function, and the load may change dramatically according to the patient, location of the implant, and parafunctional habits. Even before loading, the implant is subjected to passive mechanical load from the mucosa and the cover screw. The implants are also subjected to horizontal forces from the tongue and the perioral mucosa. Application of nonpassive prostheses may also generate mechanical forces from the abutments, even without occlusal load.

A force applied to a dental implant usually results in three-dimensional loading that is not necessarily exclusively along the long axis of the implant. Vertical and transverse loads from mastication include axial forces and bending moments and results in stress gradients in the implant and bone. A key factor for the success of the implant is the manner in which stresses are transferred to the surrounding bone. Load transfer depends on the bone-implant interface, the length and diameter of the implant, the shape and characteristics of the implant surface, the prosthesis type, and the quality and quantity of the surrounding bone. Finite Element Analysis (FEA)^a helps researchers to

^aFEA is a technique for obtaining a solution to a complex mechanical problem by dividing the problem domain into a collection of smaller and simpler domains (elements). In other words, FEA is a method where instead of seeking a solution to the entire domain, one formulates the solution functions to each finite element and then combines them to obtain the solution to the entire body.

predict the stress distribution in the contact area of the implant with cortical bone and around the apex of the implant in trabecular bone. It is known that there is some adaptive remodeling of the surrounding bone, but the precise mechanism is not fully understood.

THE FUTURE OF DENTAL IMPLANTS

Dental implants (DI) will become a more common mode for treating edentulism. In order to offer the use of DI to every patient, the surgical procedure will be simplified.

One way to simplify the surgical procedure is by what is called flapless surgery, i.e., inserting the implant without the need to raise a flap. There are many advantages in this procedure for the patient as well as for the surgeon. The procedure is less time consuming, bleeding is minimal, implant placement is expedited, and there is no need to place or remove sutures.^[29] This approach also minimizes proximal alveolar bone loss associated with sectioning of collagen fibers while elevating a surgical flap.^[30] This approach may be aided by using navigation techniques. This technique uses three-dimensional CT-based models for preoperative planning. The models are transferred into the surgical environment by mechanical devices that are being fixed to the bone or teeth. Thus, precise intraoperative verification of operation planning is more feasible by using CT data for the intraoperative navigation of the implantation tools. This technique enables intraoperative control in the deep bone without surgical exploration. A drilling tool can be controlled three-dimensionally by watching its manipulation on a screen that shows three-dimensional reconstruction of the CT scan. With this technique, precise placement of implants without injury to critical anatomic structures should be possible, even in severely atrophic jaws.^[31] These procedures already exist, but their use is rather limited. These techniques and others will enable the general practitioner as well to insert DI more easily and safely. Another field for progress is to expand the possibilities of augmentation of the residual ridge and thus enable the insertion of the DI in extremely resorbed ridges.

Patients are seeking an esthetic, quick, and easy solution. The one-stage immediate loading technique provides a better solution from the patient's point of view. Companies will try to provide implants that are inserted more easily and have a better chance to achieve primary stability.

There are two possibilities in improving primary stability—by modifying the surface of the implant or by a more radical approach such as changing the

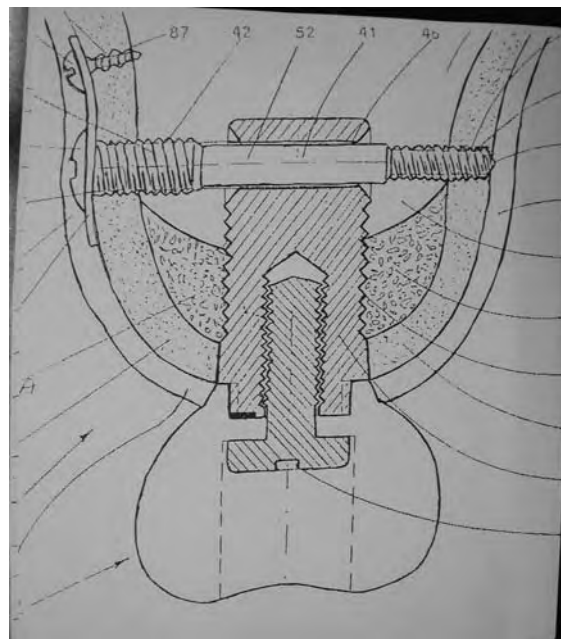


Fig. 12 Illustration of new tri-cortical implant. (Courtesy of MIS (Medical Implant System), Shlomi, Israel.)

form of the implant. For example, the MIS company is developing an implant that gets its primary stability from a screw that is inserted at the apical part of the implant through the cortical plate.

These innovations—or at least part of them—will be seen more and more in the future. In turn, this will enable more patients to restore their dentition with dental implants (Figs. 12 and 13).



Fig. 13 Clinical picture of tri-cortical implant in place immediately postop. (Courtesy of MIS (Medical Implant System), Shlomi, Israel.)

CONCLUSION

Dental implants have changed enormously since they were first introduced. The composition of the alloy has changed although titanium has remained the major component. The surface of the implants and the mode of modification of the surface are being constantly changed. The time required to achieve osseointegration has been shortened and the more adventurous treatments are being implemented.

The revolution was successfully carried out, but the evolution of dental implants has not said its last word. The fruitful cooperation between clinicians, researchers, and bioengineers will continue to bring new technology for the benefit of all patients.

ARTICLES OF FURTHER INTEREST

Bone-Implant Interface; Bone Plates and Screws, Bioabsorbable; Finite Element Analysis; Orthodontic Wires; Osteogenic Progenitor Cells of Bone; Osteoinductive Substances and Materials; Surface Modification; Titanium and Its Alloys

REFERENCES

- Anjard, R. Mayan dental wonders. *Oral Implant* **1981**, *9*, 423.
- Proussaefs, P. Evaluation of two Vitallium blade-form implants retrieved after 13–21 years of function: A clinical report. *J. Prosthet. Dent.* **2002**, *87*, 412–415.
- Bodine, R.L. Forty years of experience with subperiosteal implant denture in 41 edentulous patients. *J. Prosthet. Dent.* **1996**, *75* (1), 33–44. Jan.
- Branemark, P.-I.; Adell, R.; Breine, U.; Hansson, B.O.; Lindstrom, J.; Ohlsson, A. Intra-osseous anchorage of dental prostheses. I. Experimental studies. *Scand. J. Plast. Reconstr. Surg.* **1969**, *3*, 81–100.
- Branemark, P.-I.; Hansson, B.O.; Adell, R.; Breine, U.; Lindstrom, J.; Hallen, O.; Ohman, A. Osseointegrated implants in the treatment of the edentulous jaw. *Scand. J. Plast. Reconstr. Surg.* **1977**, *16* (Suppl.).
- Adell, R.; Lekholm, U.; Rocklar, B.; Branemark, P.-I. A 15-year study of osseointegrated implants in the treatment of edentulous jaw. *Int. J. Oral. Surg.* **1981**, *10*, 387–416.
- Lindhe. Textbook, pp. 852–853.
- Schroeder, W.; van der Zypen, E.; Stich, H.; Sutter, F. The reactions of bone, connective tissue and epithelium to endosteal implants with titanium-sprayed surfaces. *J. Maxillofac. Surg.* **1981**, *9*, 15–25.
- Ericsson, I.; Nilner, K. Early functional loading using Branemark dental implants. *Int. J. Periodontics Restor. Dent.* **2002**, *22*, 9–19.
- Pierrisnard, L.; Hure, G.; Berquins, M.; Chappard, D. Two dental implants designed for immediate loading: A finite element analysis. *Int. J. Oral Maxillofac. Implants* **2002**, *17*, 353–362.
- Andersen, E.; Haanaes, H.R.; Knusten, B.M. Immediate loading of single tooth ITI implants in the anterior maxilla: A prospective 5-year pilot study. *Clin. Oral Implants Res.* **2002**, *13*, 281–287.
- Albrektsson, T.; Zarb, G.A.; Warthington, P.; et al. The long term efficacy of currently used dental implants: A review and proposed criteria of success. *Int. J. Oral Maxillofac. Implants* **1986**, *1*, 1–25.
- Zarb, G.A.; Albrektsson, T. Osseointegration—A requiem for the periodontal ligament? An editorial. *Int. J. Periodontics Restor. Dent.* **1995**, *11*, 88–91.
- Albrektsson, T.; Branemark, P.-I.; Hansson, H.-A.; Lindstrom, J. Osseointegrated titanium implants. Requirements for ensuring a long lasting, direct bone anchorage in man. *Acta Orthop. Scand.* **52**, 155–170.
- Schenk, K.; Buser, D. Implant dentistry. *Periodontology* **2000**, *17*, 1998.
- Binon, P.P. Implants and components: Entering the new millenium. *JOMI* **2000**, *15*, 76–94.
- Sykaras, N.; Iacopino, A.M.; et al. *JOMI* **2000**, *15*, 675–690.
- Strong, J.T.; Misch, C.E.; Nalluri, P. Functional surface area: Thread form parameter optimization for implant body design. *Compendium* **1998**, *19* (3), 4–9. Special issue.
- Bobyn, J.D.; Pilliar, R.M.; Cameron, H.U.; Weatherly, G.C. The optimum pore size for the fixation of porous-surfaces metal implants by the ingrowth of bone. *Clin. Orthop.* **1980**, *150*, 263–270.
- Krauser, J.T. Hydroxyl apatite coated dental implants. Biologic rationale and surgical technique. *Dent. Clin. North Am.* **1989**, *33* (4), 879–903. Oct.
- Lee, J.J.; Rouhfar, L.; Beirne, O.R. Survival of the hydroxyapatite coated dental implants: A meta analytic review. *J. Oral Maxillofac. Surg.* **2000**, *58*, 1372–1379.
- Lazzara, R.J.; Testori, T.; Trisi, P.; Porter, S.S.; Weinstein, R.L. A human histologic analysis of osseointegration and machined surface using implants with two opposing surfaces. *Int. J. Periodontics Restor. Dent.* **1999**, *19* (2), 117–129.
- London, R.M.; Roberts, F.A.; Baker, D.A.; Rohrer, M.D.; O'Neal, R.B. Histologic comparison of a thermal dual-etched implant surface to machined, TPS and HA surfaces: Bone contact in-vivo in rabbits. *Int. J. Oral Maxillofac. Implants* **2002**, *17*, 369–376.
- Wong, M.; Eulenberger, J.; Schenk, R.; Hunziker, E. Effect of surface topology on the osseointegration of implant materials in trabecular bone. *J. Biomed. Mater. Res.* **1995**, *29*, 1567–1576.
- Ann. Periodontol.* **1996**.
- # Year results from the VA multicenter prospective implant study. *Ann. Periodontol.* **December 2000**, *5* (1).
- Raadsheer, M.C.; van Eijden, T.M.G.J.; van Ginkel, F.C.; Prahl-Andersen, B. Contribution of jaw muscle size and craniofacial morphology to human bite force magnitude. *J. Dent. Res.* **1999**, *78*, 31–42.
- Mericske-Stern, R.; Zarb, G.A. In-vivo measurements of some functional aspects with mandibular fixed

- prostheses supported by implants. *Clin. Oral Implants Res.* **1996**, *7*, 153–161.
29. Campelo, L.D.; Camara, J.R.D. Flapless implant surgery: A 10 year clinical retrospective analysis. *Int. J. Oral Maxillofac. Implants* **2002**, *17*, 271–276.
30. Binderman, I.; Bahar, H.; Yaffe, A. Strain relaxation of fibroblasts in the marginal periodontium is the common trigger for alveolar bone resorption: A novel hypothesis. *J. Periodontol.* **2002**, *73*, 1210–1215.
31. Gaggl, A.; Schultes, G. Assessment of accuracy of navigated implant placement in the maxilla. *Int. J. Oral Maxillofac. Implants* **2002**, *17*, 263–270.

D

Diamond and Diamond-Like Carbons

James H. Arps
Geoff Dearnaley

Southwest Research Institute, San Antonio, Texas, U.S.A.

INTRODUCTION

The number of biocompatible materials available for use *in vivo* is severely limited, and carbon in its various forms is one of them. Over the past decade or so, it has become possible to coat materials with adherent films of diamond or diamond-like carbon (DLC) to provide many potential applications in orthopedics, cardiovascular components, guidewires, etc. We shall review the various methods of deposition, their advantages and disadvantages, and the properties of the resulting coatings.

It will be seen that, of these properties, one of particular value is that of smoothness, resulting in much lower wear of a softer counterface. It is here that DLC emerges as a superior coating. On a well-prepared substrate it can be atomically smooth, with roughness (R_a) lower than 1 nm.

The volume of publications on the biomedical applications of DLC, in particular, has grown steadily over the past decade as an outcome of successful demonstrations *in vitro* of its wear resistance, low friction, and biocompatibility. At the same time, increasingly more applications are being found for this versatile form of carbon. It has now reached an exciting stage of development as a biomaterial.

DIAMOND COATINGS

Because of its extreme hardness and generally low coefficient of friction, crystalline diamond may be considered the ideal wear-resistant coating, for example, for orthopedic implants. We shall see, however, that the requirement of smoothness makes this difficult to realize.

There are numerous ways to deposit diamond coatings, but the most widely and successfully used is by microwave plasma-activated chemical vapor deposition (CVD) from a mixture consisting mostly of hydrogen with minor admixtures of a carbonaceous precursor, such as methane.

However, before deposition of diamond onto metal surfaces, it is necessary to carry out a pretreatment in order to improve the nucleation and adherence of the diamond. This is necessary for several reasons, two of which are: 1) the interatomic bonding in diamond is

wholly covalent and distinct from that in a metal; and 2) there is generally a mismatch in the coefficients of thermal expansion (CTE). Thus, the CTE of titanium alloy is about 2.6 times that of diamond, and since deposition temperatures in CVD are typically 700°–800°C, the coating develops a high compressive stress on cooling that may lead to peeling. Mechanical treatments such as scratching and diamond polishing and application of interlayer or bond coats such as titanium nitride or diamond-like carbon have been described by Fu et al.^[1] These interlayers were much less effective than a process of plasma nitriding followed by plasma carbonitriding of the titanium substrates. Scratching of the surface with diamond powder substantially increased the nucleation density and adherence of the coatings.

Rats et al.^[2] describe the coating of Ti-6Al-4V disks with diamond in a tubular reactor crossed by a waveguide connected to a 6-kW microwave generator operating at 2.45 GHz. In some cases, a 1% CH₄/H₂ mixture at 700°C was used, whereas other coatings were made in the presence of 8% CO/H₂ at the somewhat lower temperature of 600°C. The pressure was set at 10 Torr. Raman spectrometry and friction and wear tests showed that the CO/H₂ mixture gave more pure diamond with lower friction. The residual growth stress in the 600°C coatings was lower, –5.2 GPa, compared with –6.2 GPa for those grown at 700°C.

For most tribological applications, however, these coatings all had the drawback of excessive roughness due to the individual diamond crystallites. Although the R_a values were 70 to 80 nm, the peak-to-valley ratio was in the range of 0.5 to 0.5 micron, some of which was due to the prior scratching. This hard abrasive surface produces very rapid wear of all uncoated metallic counterfaces. In view of this, the same research group attempted to produce finer-grained and smoother diamond coatings on Ti-6Al-4V alloy at 600°C.

The preliminary polishing was carried out to an R_a of about 5 nm, and nucleation and growth step procedures were varied in order to achieve coatings about 2 microns thick, with submicron crystalline size. The final roughness values varied from 5 to 140 nm, the lowest figure being obtained by polishing for 6 hours against a diamond-coated flat. Friction coefficients of diamond

against diamond of 0.05 to 0.07 were reported for the finer-grained coatings. The wear of uncoated titanium alloy pins against these diamond-coated disks was very strongly dependent on their surface roughness. Thus, the wear rate increased from $2 \cdot 10^{-9} \text{ mm}^3/\text{Nm}$ to $5 \cdot 10^{-7} \text{ mm}^3/\text{Nm}$ as the roughness increased from 5 nm to 140 nm. When a diamond-coated titanium alloy pin with a 3.5-micron coating was worn against a diamond-coated disk, it was observed that, after only 200 m of sliding an alloy of 13 N, the alloy substrate became exposed. Subsequent wear remained low as the Hertzian contact pressure on the pin surface diminished due to flattening, and the wear of the coated disk was unmeasurably small. The Hertzian pressure is defined as the maximum pressure induced by the applied load beneath the spherical contour of the pin. The authors conclude that these may be the same engineering applications for which adherent diamond coatings on titanium alloys will be useful.

Thus, polycrystalline diamond coatings on titanium are extremely hard, have low friction coefficients, and can be highly adherent, but nevertheless are not ideal for tribological purposes because of their roughness and the considerable difficulty of polishing them smooth. A second problem lies in the high growth stress due to the mismatch in coefficients of thermal expansion, producing a tendency toward decohesion. Attempts have been made to polish CVD diamond coatings using molten rare earth metals, liquid iron at 1250°C , and hyperthermal oxygen ions, but all these methods have their drawbacks and would not be easily applied in a production process. In conclusion, polycrystalline diamond has not emerged as a viable material for coating of orthopedic implants or other wear-resistant biomedical components.

Turning next to the feasibility of applying diamond coatings to cobalt–chromium orthopaedic alloys, we encounter another problem that was first revealed when diamond was deposited onto cobalt-cemented tungsten carbide tools. Cobalt has a catalytic influence that causes the initial layer to consist of graphite, which is weak and results in poor adhesion. It is therefore necessary to remove the cobalt binder from the surface by acid etching, laser ablation, or other procedures before diamond deposition. These methods are obviously not applicable to cobalt alloys, but there is an alternative that so far has not been explored. If silicon is deposited at an elevated temperature around 500°C by ion beam-assisted deposition (IBAD), surface cobalt may be converted completely to cobalt disilicide, CoSi_2 , which is no longer catalytically active. This procedure has been successful in the case of cobalt-cemented tungsten carbide.^[3] However, the problems associated with the roughness of polycrystalline coatings, described earlier, would also pertain to their use as cobalt–chrome or other substrate systems.

In conclusion, the applications of crystalline diamond coatings in the medical field appear to be limited, and it is significant that in his survey of CVD diamond applications and projects for the future, Michael Seal did not include a single biomedical application.^[4]

DLC COATINGS

An alternative way to overcome the problems due to roughness in crystalline diamond coatings, while retaining some of their merits of hardness, low friction, and chemical inertness, is by the use of deposited amorphous DLC. With care, exceptional degrees of smoothness can be achieved, together with good coating adherence under stress because growth stresses in DLC films can be better controlled than in diamond.

As mentioned earlier, DLCs comprise a broad family of compositions and microstructures and all have some degree of hardness arising from the presence of a strong matrix containing three-dimensional carbon–carbon bonds (sp^3 bonding), with a dispersion of nanocrystalline graphite islands (sp^2 bonding). The presence of the sp^3 bonding justifies use of the term “diamond-like.” Figure 1 shows a typical transmission electron microscope cross-section of a DLC film.

There are numerous methods available for the deposition of DLC coatings; these will be considered in turn.

DEPOSITION OF DLC COATINGS

It is now more than thirty years since Aisenberg and Chabot reported in 1971 the results of a direct ion

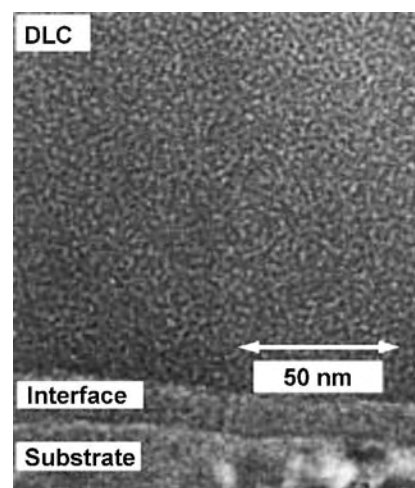


Fig. 1 TEM micrograph of DLC coating.

beam deposition of DLC. Since then, the methods developed include:

- Direct ion beam deposition.
- Dual ion beam sputtering.
- Pulsed laser ablation.
- Filtered cathodic arc deposition.
- Ion beam conversion of condensed precursor.
- Magnetron sputtering.
- Radio-frequency (RF) plasma-activated CVD.
- Plasma source ion implantation and deposition.

Each of these requires the choice of a carbonaceous precursor, such as methane, acetylene, or butane, and some make use of elemental carbon. In other cases, the precursor may be a much more complex molecule chosen to possess the low vapor pressure for a condensate film. Alternatively, the material may contain silicon and oxygen, in the form of a siloxane, or metallic constituents, as in ferrocene or metal-containing phthalocyanines. Some of the methods are versatile enough to produce a wide range of modified DLCs with varying but controllable electrical conductivities, hydrophobic properties, and catalytic behaviors. By comparison with polycrystalline diamond films, the potential exists in DLC to tailor the microstructure and the composition and to grade these where necessary to achieve the required functionality.

We now consider the advantages and limitations of each process in turn before addressing the applications of DLCs in the biomedical field.

Direct Ion Beam Deposition

The principle of this method is to bombard a surface, in vacuum, with energetic ions usually of methane so that, upon impact, the molecule dissociates and most of the hydrogen is liberated. Ion energies of 100 to 750 eV have been used, e.g., by Liu et al.,^[5] the current density of about 2.5 A/cm² corresponds to an arrival rate of approximately $1.5 \cdot 10^{16}$ carbon atoms/cm²/sec, most of which are retained.

With a uniform atom-by-atom arrival, the coatings are very smooth and have a Knoop microhardness of 6000 kg/mm² according to Liu et al.^[5] Although the residual hydrogen content is not stated, this hardness value suggests that it is relatively low. Against a zirconia pin, the wear rates were below 10^{-7} mm³/Nm, and friction coefficients for dry sliding were 0.02 to 0.10, increasing with relative humidity for reasons discussed later.

Dual Ion Beam Sputtering

This is the method developed by Weissmantel in early work, and it makes use of one ion beam, usually of

argon, to sputter carbon onto the substrate while it is simultaneously irradiated with 100–300 eV ions by a second gun. This IBAD process increases considerably the adherence of the sputtered carbon deposit. However, the forward momentum transferred by the ion-assist beam compacts the coating and, if used excessively, will lead to a high compressive stress that can result in coating decohesion. For this reason, some workers have preferred to use higher ion energies and lower fluxes to achieve good coatings with some degree of ion beam mixing at the coating-substrate interface. The process is slow due to the low sputtering coefficient of carbon, but the method is capable of producing relatively pure coatings, although typically with a small percentage of argon.

Pulsed Laser Ablation

Voevodin et al. described a process for the deposition of DLC coatings using KrF excimer laser to vaporize material, in vacuum, from either a graphite target or one consisting of polycarbonate. The latter results in hydrogenated DLC films sometimes designated as α -C:H. Steel substrates were either negatively biased or unbiased with regard to the target. Those coatings deposited under bias, from –100 to –800 V, have somewhat lower friction coefficients. The wear rates of the carbon films against sapphire pins were remarkably low, at about 10^{-9} mm³/Nm, whereas for the softer α -C:H films, the wear rate was around 10^{-6} mm³/Nm at a contact pressure of 0.8 GPa.

A potential drawback to pulsed laser ablation (PLD) is that occasionally droplets or chunks of the target material are ablated and cause a surface roughness that is not easily removed. The internal stress can also be very high, but Wei et al.^[6] have described a functionally-graded DLC coating with reduced compressive stress and improved adhesion.

Ion Beam Conversion Coatings

One of the more versatile methods for depositing DLCs is by the simultaneous ion bombardment and condensation of a low-vapor-pressure compound. The rupture of C–H bonds releases hydrogen into the vacuum leaving a largely carbonaceous coating. Deposition rates are high, up to 10 μ m/hour, and the compressive stress in such coatings is low (1 GPa or less), perhaps due to the progressive removal of hydrogen from within. The precursor materials may be either diffusion pump fluids or solids such as adamantane or coronene. Arps has shown that if ferrocene is used, the coating may contain up to 24 wt.% of iron. Siloxane precursors such as pentaphenyld trimethyl siloxane produce films with a small percentage of

silicon and oxygen, with advantageous properties, such as reduced dependence of friction coefficient or relative humidity. The ion energies used may range from 100 eV to 100 keV. Accounts of the process have been published by Fountzoulas et al.^[7]

It is also possible to incorporate elements such as sulfur, fluorine, or nitrogen by choice of the precursor, and some of these assist in providing a low friction coefficient. The coatings produced by these methods are smooth, but due to the residual hydrogen content, typically 15 wt. %, they are not as hard as amorphous carbon coatings. However, they can still be comparable with hard tool coatings such as titanium nitride or alumina.

Filtered Cathodic Arc Deposition

This is a method for producing very hard, virtually hydrogen-free DLC coatings. An arc is drawn in vacuum between a graphite cathode and an anode, typically at 50–60 A. As illustrated in Fig. 2, adapted from Xu et al.,^[8] a radial magnetic field steers and filters carbon ions and rejects macroparticles. A bias voltage of 100 to 300 V is applied to the substrate.

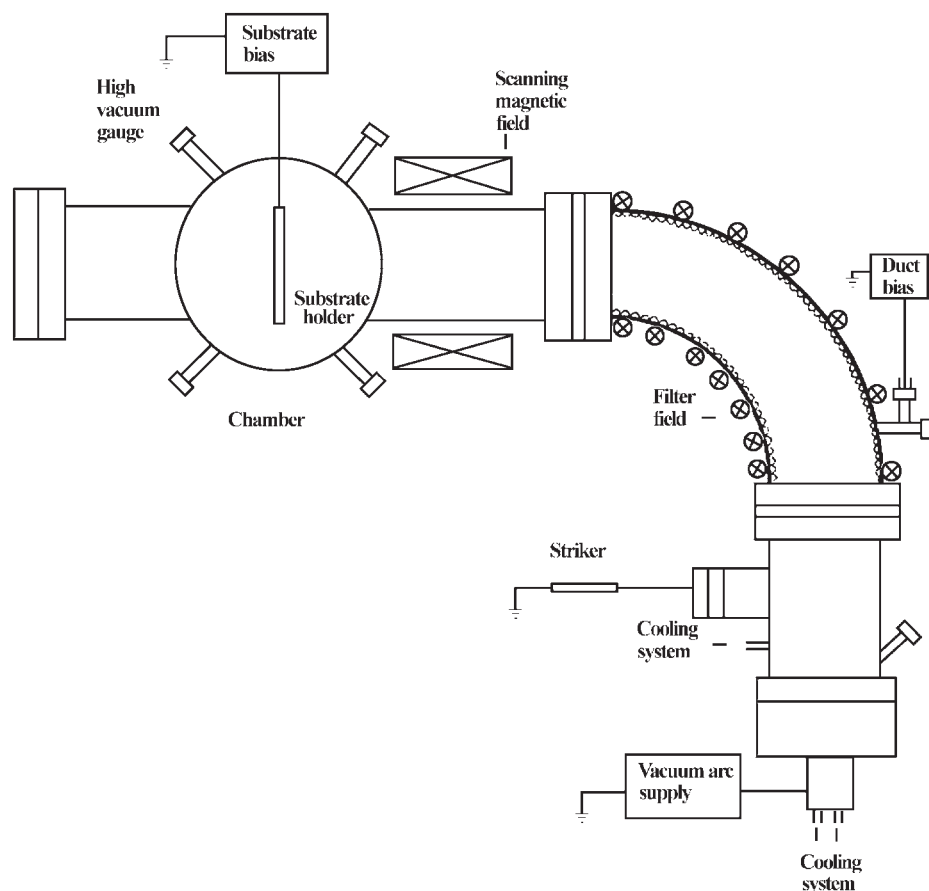


Fig. 2 A filter cathodic arc apparatus for DLC deposition.

That the filtering is effective is shown by the high degree of smoothness of the films. Xu et al.^[8] report an R_a value of about 0.4 nm. Stresses in the coatings are high, -7 to -10 GPa, and this contributes to coating hardness that may reach superhard values of 40 to 50 GPa.

Such hardness is advantageous in resisting abrasive wear, but the high compressive stress means that there will be a tendency toward decohesion unless adhesion is strong. Most of the reports refer to substrates of silicon or quartz, to which carbon adheres strongly.

Plasma Source Ion Implantation and Deposition

In this method, described by Anders et al.,^[9] a plasma is created in the work chamber and the workpieces are immersed in it. Periodic pulsing to around 2 kV accelerates carbon ions from the expanding plasma sheath during the pulse of duration 1 microsecond. The carbon plasma may be formed by a pulsed cathodic arc, with a magnetic filter (as described earlier) to remove macroparticulates. The films are virtually free of hydrogen and are hard. Anders et al. report values up to 45 GPa, with low friction coefficients.^[9]

The compressive stress in such coatings is high, but Anders et al.^[9] described a method for varying the substrate bias during deposition to produce alternating hard and softer layers with a net reduction in stress. Some of these coatings have been the subject of orthopedic testing (to be discussed later), with outstanding results.

Magnetron Sputter Coating

Magnetron sputtering is a widely used process for physical vapor deposition and it can be employed for carbon, but deposition rates are relatively low due to the low sputtering coefficient. Peng et al.^[10] used DC magnetron sputtering as one of the four methods of deposition in a study of the smoothness of DLC coatings. Nonhydrogenated films up to 1 micron thick were prepared by sputtering from a pure graphite target in an argon plasma. A bias voltage was applied to the substrate and could be varied from 200 to 300 V. The argon pressure was around 1 Pa. A literature survey indicates that magnetron sputtering is not widely used for DLC deposition, by comparison with CVD methods or the filtered cathodic arc method.

RF Plasma-Activated CVD

Chemical vapor deposition, involving the thermal dissociation of a selected precursor, is extensively used for coating tools and semiconductor devices. The addition of an electrical plasma lowers process temperatures. It is easy to apply the method for DLC, and Erdemir et al.^[11] described a careful study of the results of using four different source gases; namely, methane, ethane, acetylene, and ethylene, at pressures of 10 to 13 mTorr. The substrate bias was 1600 V, and the substrates were of steel. The lowest friction and wear, against steel balls, were obtained for methane, and the poorest for acetylene. The authors attribute the difference to the hydrogen:carbon ratio, the role of hydrogen possibly being to promote a higher sp^3 content, or else to decorate the coating surface and reduce friction. Methane/hydrogen mixtures in the plasma (1:1) gave even lower friction, the coefficient against steel falling to 0.01. It is a pity that these results were not correlated with determinations of the residual hydrogen present in the coatings, which can be done by an elastic-recoil ion beam technique.

The wear resistance corresponded to a friction behavior of about $9 \dots 10^{-9} \text{ mm}^3/\text{Nm}$ for methane-grown films and two orders of magnitude greater for the acetylene-grown film. In conclusion, methane/hydrogen mixtures are shown to be best for tribological applications of DLC.

THE TRIBOLOGICAL PROPERTIES OF DLC

For a number of biomedical applications, it is desirable to have coatings that provide low friction and low wear under load-bearing conditions. For reasons discussed by Liu and Meletis,^[12] DLC can meet such requirements and provide corrosion resistance as well.

Both experimental and molecular dynamics modeling studies show that, during sliding wear, the outer molecular layers of DLC undergo transformation to graphite, with its basal planes parallel to the surface. This enables shearing to take place at load-bearing asperities, and provides friction coefficients in dry nitrogen as low as 0.01.

It has long been known that the friction and wear behavior of diamond is controlled by a macromolecular film of adsorbate at the surface, and Field et al. have shown the same to be true of DLC. Hydrogen atoms play a key role, as does the formation of a transfer film of DLC on the counterface. Under steady-state conditions, at contact pressures up to 1 GPa, both mating surfaces are decorated by hydrogen atoms that slide over one another with no tendency for adhesion. A very smooth surface then also minimizes the likelihood of plowing and abrasive wear.

In humid air, the friction coefficient of DLC inverses and may rise above 0.3. Gardos et al. attribute this to the possibility of hydrogen being removed due to the relative strength of hydrogen bonding to water molecules. Additions of silicon to DLC substantially reduce the influence of humidity or friction. However, there is evidence that the reduced rate of surface graphitization of DLC in water, or presumably body fluids, will reduce its already low wear rate, according to Jiang et al.^[13]

THE SMOOTHNESS OF DLC COATINGS

The surface roughness of hard coatings has a strong influence on the wear of the counterface, especially when this is a soft material such as ultrahigh-molecular-weight polyethylene (UHMWPE). A careful study has been made by Peng et al.^[10] of the roughness of DLC coatings prepared by three different techniques, and of the effects of postdeposition treatments.

The methods chosen were RF plasma-activated CVD in methane, DC magnetron sputtering in argon, and a carbon ion beam extracted from a cathodic arc discharge. In each case the bias voltage applied to the substrate could be varied up to about 350 V. After deposition, some specimens were subjected to argon plasma sputtering, hydrogen etching in a plasma, or heat treatment at 500°C in vacuum. The substrates were of silicon, were argon-sputter cleaned, and had a roughness of only 0.046 nm.

The coating roughness was similar for all three deposition methods, but depended markedly on the bias voltage, as ion impingement energy, with the lowest values of about 0.04 nm being obtained at a bias of 250–300 V. Heating at 500°C increased roughness, as did argon sputtering and hydrogen etching, which perhaps remove graphitic clusters from certain regions of the surface.

Thus, under preferred conditions, DLC coatings up to 700 nm thick could be as smooth as a carefully prepared silicon crystal substrate.

ORTHOPEDIC APPLICATIONS OF DLC

The load-bearing surfaces of total hip and knee arthroplasties are subject to wear, and it is well recognized that the most serious consequences are those due to the formation of polyethylene (PE) debris at a rate of perhaps 10^{10} particulates per year. These particles are phagocytosed, resulting in granulomatous lesions, osteolysis, and bone resorption, causing pain and aseptic loosening of the prosthesis.

Diamond-like carbon coatings have been explored over the past 10 years as a means for eliminating this problem, with some highly encouraging results. The goal is to demonstrate lower PE wear than occurs against metals and ceramics such as alumina or zirconia. Fractures have occurred in ceramic femoral heads, and so a hard, low-friction coating on metal would potentially be the best solution. Figure 3 shows a DLC-coated example prepared using IBAD.

Some early investigations of DLC by Davidson and Mishra were beset by problems of high compressive stress and decohesion, but the work referred to earlier has done much to reduce the likelihood of coating



Fig. 3 DLC coated and uncoated alumina femoral heads.

delamination, either by bond coats or control of coating stress.

Results over the past five years have indicated the excellent potential for DLC in total joint replacements—but the findings, although all good, have been variable, and it is interesting to arrive at an explanation for this in terms of coating roughness and test procedure.

Saikko et al.^[14] compared femoral heads of CoCr, alumina, and CoCr coated with DLC prepared by an RF plasma discharge in acetylene. Tests were done in bovine serum, with no additives, in an anatomical hip wear simulator operating at 1 Hz. The wear of UHMWPE over 3 million cycles was similar for each type of head.

Sheeja et al.^[15] tested Co–Cr alloy disks coated with DLC by the filtered cathodic arc method and compared them with uncoated Co–Cr against UHMWPE pins in a simulated body fluid. The wear rates were similar, but it was found that the corrosion rate for the coated alloy was reduced by a factor of about 10,000.

In an NSF-funded study, Xu and Pruitt^[16] examined DLC coatings on Ti-6Al-4V alloy against UHMWPE in pin-on-disk tests in water. The DLC was prepared by the plasma source method described earlier. The rate of PE wear was reduced by a factor of 3–4, compared with uncoated alloy.

Oñate et al.^[17] used a knee wear simulation machine to compare Co–Cr, alumina, and Co–Cr coated with DLC from acetylene (as provided by Saikko et al.^[14]). In this case there was a substantial improvement over the uncoated alloy, by a factor of 4, and 40% less wear than against alumina.

A hip joint simulator was used by Lappalainen et al.^[18] to compare DLC coatings on stainless steel, Ti-6Al-4V, and CoCrMo alloy, by the filtered cathodic arc method, with a chromium bond coat. With ultra-smooth coatings with a roughness of 7 nm, the wear rate of UHMWPE was reduced 30- to 600-fold compared with that of the uncoated metal. Corrosion rates in 10% HCl were decreased over 10,000-fold.

Affatato et al.^[19] made use of a hip joint simulator to compare DLC-coated titanium alloy with CoCr and alumina over 5 million cycles in bovine serum with addition of EDTA (ethylene diamine tetraacetic acid) to minimize precipitation of calcium phosphate, which they state can strongly affect friction and wear properties. The DLC, made by PACVD from methane, was comparable to alumina in terms of PE wear.

The variability of these findings suggests that care must be taken in regard to:

- DLC deposition procedure: Methane is superior to acetylene as a source gas.
- Smoothness of the substrate and coating.
- Additive to serum to inhibit calcification.

When conditions are favorable, tests indicate that DLC coatings have the potential to provide the lowest wear rates of UHMWPE—perhaps half those against alumina—without the risks associated with fracture of ceramic components. Corrosion and leaching of metal in body fluids has been shown to be very markedly reduced by DLC coatings.

BIOCOMPATIBILITY OF DLC COATINGS

It would hardly be possible to use DLC as a coating for items *in vivo* unless it had been shown to be biocompatible, and all studies to date agree that this is the case. We will here summarize the evidence.

In an early study, Thomson et al.^[20] used mouse peritoneal macrophages and fibroblasts on DLC prepared by the dual ion beam method, monitoring levels of lactate dehydrogenase (LDH) as a measure of cell viability. There was no indication of cytotoxicity. Similar results were obtained by Allen et al.^[21] using the murine macrophage cell line, IC-21, in a growth medium supplemented with calf serum, on DLC made by plasma-activated CVD. Cell growth studies were made using human synovial fibroblasts and a human osteoblast-like cell line, SaOS-2. Cell growth kinetics were determined by counting after each 24-h incubation period, following disaggregation and staining. Cells grew well on DLC-coated glass and polystyrene with no evidence of abnormal morphology. Cells generally grew faster on DLC, adhered well, and produced extensive filopodia.

Other *in vitro* studies made use of human ML-1 hematopoietic myeloblasts, which are immature and can differentiate into monocytes as macrophages as a result of environmental conditions. There was no difference in cell growth or morphology on DLC coatings (Lu et al.^[22]) made by the direct ion beam method.

Butter and Lettington^[23] reported preliminary *in vivo* studies involving the implantation of DLC-coated pins into soft tissue and femurs of sheep. Much better bonding was observed at DLC rather than metal–tissue interfaces, indicating a lower risk of infection. They also described work done by Mitura on the coating of orthopedic screws with DLC by the RF plasma method. Over a period of 52 weeks, implanted DLC-coated metal showed no evidence of corrosion products or chronic inflammatory reaction.

Blood compatibility tests have been carried out by Dion et al. on DLC-coated Ti6Al4V alloy showing that more albumin adheres to DLC than to medical grade silicone elastomer, by a factor of about 2.2, and this is known to have positive consequences for platelet adhesion. The higher the albumin:fibrinogen ratio, it is believed, the lower will be platelet adhesion and at 1.24 the value for DLC is superior to the 0.76 for silicone.

Ex vivo experiments with whole blood have shown that DLC-coated stainless steel and titanium compare well with LTI carbon and are superior to glass. There is a need, however, for *in vivo* trials, justified by the fact that all biocompatibility tests of DLC-coated surfaces have been successful, with no contrary indications.

In an extension of the biocompatibility of DLC, Steffen et al.^[24] used it as a substrate for heparin following treatment in an ammonia plasma to form amino reaction sites for attachment. The optimum plasma exposure time at a pressure of 3.10^{-3} mbar was around 45 seconds. This treatment with heparin increased blood coagulation times from 25 to over 250 seconds.

Successful *in vivo* tests of the biocompatibility of DLC have been reported by Allen et al.^[25] who implanted DLC-coated CoCr cylinders into intramuscular locations in rats and transcortical sites in sheep. After 3 months, histologic analysis showed that the specimens were well tolerated. Based on this excellent biocompatibility, the authors have initiated a long-term animal study of a DLC-coated knee arthroplasty.

Finally, for some applications, it can be reported that DLC adheres very strongly to silicone elastomers, probably due to the formation of strong Si–C bonds at the interface.

CARDIOVASCULAR APPLICATIONS OF DLC

Heart valves are conventionally made from LTI carbon and, despite great care in manufacture, there remains a very small but finite risk of failure due to fracture initiated at an undetected surface or sub-surface flaw. Therefore, there has been interest in replacing this brittle material with a metal coated with hard carbon that is nonthrombogenic.

Butter and Lettington^[23] suggested the use of NiCr18 stainless steel coated with carbon by a non-line-of-sight method. The plasma conversion method would be a suitable choice. It has been reported that a silicon-containing DLC may be a better coating for cardiovascular purposes, but Butter and Lettington^[23] commented that in tests it showed no difference from the controls. The published studies all show DLC to be a promising biomaterial for heart valve applications, and proprietary work is in progress on this topic.

Arterial stents can induce platelet activation and may initiate thrombosis by shear forces on the flow and by platelet adhesion to the metal. Gutensohn et al.^[26] evaluated *in vitro* the performance of stents coated with DLC. Growth arrays using smooth muscle cells and endothelial cells showed that DLC did not affect proliferation rates, and no cytotoxic effects were

observed. Flow cytometric analyses showed no significant changes in mean channel fluorescence intensity for the structural antigens CD41a and CD42b, whereas by contrast, expression of the activation-dependent antigens CD62p and CD63 increased significantly in uncoated compared with DLC-coated stents.

Release of metal ions into the bloodstream is a matter of concern, and Gutensohn et al.^[26] used atomic adsorption spectrometry to detect a significant release of nickel and chromium ions from noncoated stents into human plasma over a 96-hour period. However, only minimal concentrations of released ions could be detected in the case of DLC-coated stents. Similar results were obtained by inductively-coupled plasma mass spectrometry analysis, and in this case the release of metal ions from DLC-coated stents was virtually undetectable. The authors concluded that the coating of intracoronary stents with DLC may contribute to a reduction in thrombogenicity and consecutively a reduction in the incidence of acute occlusion and in-stent restenosis *in vivo*.

Radioactive coatings have been applied successfully on stents in order to reduce restenosis as a result of the growth of scar tissue. Rhenium isotopes are beta-emitters with a conveniently short half-life, but the soluble perrhenate anion, ReO_4^- , tends to leach out. Evans et al.^[27] described the use of DLC coatings deposited from an RF plasma to form a barrier layer over the rhenium. After one month in 0.9% saline solution, the amount of released radioactivity was determined. The DLC coating reduced the release of activity by an average of a factor of 3, with greater effectiveness obtained at higher bias voltages of 500 V. It was inferred that the greater energy of ion bombardment served to densify the DLC coating and so reduce the leach rate. It was also shown in this study that there was no cytotoxicity for cells in contact with the DLC coating.

Diamond-like carbon coatings on segmented polyurethane were tested for blood compatibility by Alanazi et al.^[28] they showed that the coated polyurethane was superior to an excellent nonthrombogenic polymer, 2-hydroxyethyl methacrylate (HEMA), in tests carried out in a parallel flow chamber. The authors concluded that greater attention should be paid to DLC for use in the medical field.

OTHER BIOMEDICAL APPLICATIONS OF DLC

Four examples will serve to demonstrate the versatility and usefulness of DLC coatings in other biomedical applications. There seems no doubt that others will follow.

Elinson et al.^[29] applied thin coatings of DLC (20–200 nm) to soft contact lenses and contact lens

cases to reduce the problems of biofilm formation. The authors regard microbial contamination of contact lens cases to be as major a problem as that of the lenses themselves, and showed that lenses stored in DLC-coated cases were free from any contamination.

McLaughlin et al.^[30] investigated RF plasma-deposited DLC on stainless steel medical guidewires and found good adherence with a coefficient of friction superior to that of PTFE. Furthermore, the thin DLC coating did not alter the overall stiffness of the guidewires, in contrast to PTFE or silicon coatings that were up to 300 microns thick.

In ophthalmic surgery, the surgeon may need to suture the corneal region while causing the least distortion of the eyeball. Butter and Lettington^[23] referred to experiments that have shown that a DLC coating on the needle reduces the force required to penetrate the corneas of pigs' eyes by about 30%. There may be more such applications in microsurgery, and the authors point out that the dark color of DLC would have the further benefit of reducing reflections from the lights of operation microscopes.

Finally, Schwank et al.^[31] made use of the high electrical resistance of hydrogen-containing DLC, made by the plasma-activated CVD method, to insulate the tip of a new design of Pt/Ir needle for intracellular potential measurement. In order to render the tip of the needle conductive, exposure to localized oxygen plasma at a pressure of 10 bar was used. This may remove the carbon by reactive ion etching, but in addition may transform the film into a more graphitic (sp^2) state.

CONCLUSION

Whereas diamond coatings may have some applications in biomedicine, it is DLC that has emerged over the past decade as a most versatile and useful biomaterial. Harder than most ceramics, bioinert, and with a low friction coefficient, DLC is one of the best materials for orthopedic applications. If deposited with the preferred composition on a very smooth substrate, it offers the lowest wear rate of a UHMWPE counterface, below that observed against polished alumina or zirconia, and without risks associated with fracture of a brittle ceramic.

All studies of the biocompatibility of DLC are in agreement that there is no cytotoxicity, and cell growth is normal on a DLC-coated surface. Its blood compatibility is as good as that of the well-established LTI carbon used in heart valves. Diamond-like carbon coatings on stainless steel have performed very well in *in vitro* studies of hemocompatibility.

Undoubtedly there is more work in progress than has been reported in the open literature, and the stage

has been reached at which in vivo testing of DLC is warranted. With cautious optimism, we look forward to the application of well-bonded DLC coatings in a variety of orthopedic and cardiovascular applications. All of these procedures involving total joint replacements, heart valves, and stenting, for that DLC may provide greater efficacy and an extended service life, are ones that give patients a much higher quality of life.

ARTICLES OF FURTHER INTEREST

Biocompatibility Testing; Implant, Total Hip; Orthopedic Biomaterials; Stents; Surface Coatings; Vascular Grafts; Vascular Grafts: Development Strategies

REFERENCES

1. Fu, Y.; Yan, B.; Loh, N.L. Effects of pretreatments and interlayers on the nucleation and growth of diamond coatings on titanium substrates. *Surf. Coat. Technol.* **2000**, *130*, 173–185.
2. Rats, D.; Vandenbulcke, L.; Boher, C.; Farges, G. Tribological study of diamond coatings on titanium alloys. *Surf. Coat. Technol.* **1997**, *94–95*, 555–560.
3. Dearnaley, G.
4. Seal, M. The Current Status of CVD Diamond Applications and Prospects for the Future. In *Applications of Diamond Films and Related Materials*; Feldman, A., Tzeng, Y., Yarbrough, W.A., Yoshikawa, M., Murakawa, M., Eds.; National Institute of Standards and Technology: Washington, DC, 1995; 3–10.
5. Liu, Y.; Erdemir, A.; Meletis, E.I. A study of the wear mechanisms of diamond-like carbon films. *Surf. Coat. Technol.* **1996**, *82*, 48–56.
6. Wei, Q.; Sankos, J.; Narayan, J. Structure and properties of novel functional diamond-like carbon coatings produced by laser ablation. *Surf. Coat. Technol.* **2001**, *146–147*, 250–257.
7. Fountzoulas, C.G.; Kattamis, T.Z.; Demaree, J.D.; Hirvonen, J.K. Silicon-Containing DLC Coatings Formed by Ion Beam Assisted Deposition and Properties of a New Category of Hard, Wear Resistant Thin Films. In *Applications of Diamond Films and Related Materials*; Feldman, A., Tzeng, Y., Yarbrough, W.A., Yoshikawa, M., Murakawa, M., Eds.; National Institute of Standards and Technology: Washington, DC, 1995; 907–910.
8. Xu, S.; Tay, B.K.; Tan, H.S.; Zhong, L.; Tu, Y.Q.; Silva, S.R.P.; Milne, W.I. Properties of carbon ion deposited tetrahedral amorphous films as a function of ion energy. *J. Appl. Phys.* **1996**, *79*, 7234–7240.
9. Anders, S.; Anders, A.; Brown, I.G.; Wei, B.; Komvopoulos, K.; Ager, J.W.; Yu, K.M. Effect of vacuum arc deposition parameters on the properties of amorphous carbon thin films. *Surf. Coat. Technol.* **1994**, *68–69*, 388–393.
10. Peng, X.L.; Barber, Z.H.; Clyne, T.W. Surface roughness of DLC films prepared using various techniques. *Surf. Coat. Technol.* **2001**, *138*, 23–32.
11. Erdemis, A.; Nilufer, J.B.; Eryilmaz, O.C.; Beschliesser, M.; Fenske, G.R. Friction and wear performance of DLC films grown in various source gas plasmas. *Surf. Coat. Technol.* **1999**, *120–121*, 589–593.
12. Liu, Y.; Meletis, E.I. Evidence of graphitization of DLC films during sliding wear. *J. Mater. Sci.* **1997**, *32*, 3491–3495.
13. Jiang, J.; Zhang, S.; Arnell, R.D. The effect of relative humidity on wear of a diamond-like carbon coating. *Surf. Coat. Technol.* **2003**, *167*, 221–225.
14. Saikko, V.; Ahlroos, T.; Caloni, O.; Keränen, J. Wear simulation of total hip prostheses with polyethylene against CoCr, alumina and DLC. *Biomaterials* **2001**, *22*, 1507–1514.
15. Sheeja, D.; Tay, B.K.; Lau, S.P.; Nung, L.N. Tribological characterisation of DLC coatings on Co–Cr–Mo alloy for orthopaedic applications. *Surf. Coat. Technol.* **2001**, *146–147*, 410–416.
16. Xu, T.; Pruitt, L. Diamond-like carbon coatings for orthopaedic applications. *J. Mater. Sci.: Mater. Med.* **1999**, *10*, 83–90.
17. Oñate, J.I.; Comin, M.; Bracer, I.; Garcia, A.; Viviente, J.L.; Brizuela, M.; Garagorri, N.; Peris, J.L.; Alava, J.I. Wear reduction effect on UHMWPE by application of hard coatings and ion implantation on Co–Cr alloy, as measured in a knee wear simulator machine. *Surf. Coat. Technol.* **2001**, *142–144*, 1056–1062.
18. Lappalainen, R.; Anttila, A.; Heinonen, H. Diamond-coated total hip replacements. *Clin. Orthop. Relat. Res.* **1998**, *352*, 118–127.
19. Affatato, S.; Frigo, M.; Toni, A. An invitro investigation of DLC as a femoral head coating. *J. Biomech. Mater. Res. (Appl. Biomater.)* **2000**, *41*, 221–226.
20. Thomson, L.A.; Law, F.C.; Rushton, N.; Franks, J. Biocompatibility of a DLC coating. *Biomaterials* **1991**, *12*, 37–40.
21. Allen, M.; Law, F.; Rushton, N. The effects of DLC coatings on macrophages, fibroblasts and osteoblast-like cells in vitro. *Clin. Mater.* **1994**, *17*, 1–10.
22. Lu, L.; Jones, M.W.; Wu, R.L.C. DLC as a biocompatible material for cell culture and medical application. *Biomed. Mater. Eng.* **1993**, *3*, 223–228.
23. Butter, R.S.; Lettington, A.H. DLC for biomedical applications. *J. Chem. Vapor Depos.* **1995**, *3*, 182–192.
24. Steffen, H.J.; Schmidt, J.; Gonzalez-Elipe, A. Biocompatible surfaces by immobilization of heparin on DLC films deposited on various substrates. *Surf. Interface Anal.* **2000**, *29*, 386–391.
25. Allen, M.; Myer, B.; Rushton, N. In vitro and in vivo investigations into the biocompatibility of DLC coatings for orthopedic applications. *J. Biomed. Mater. Res.* **2001**, *58*, 319–328.
26. Gutensohn, K.; Beyrhien, C.; Brockmann, M.; Bau, J.; Keuhnl, P. Surface Coating of Stents with DLC: Reduction of Thrombogenicity. In *10th Congress on Thrombosis*; Porto, May 2000. Abstract No. 24.
27. Evans, E.A.; Hafeli, U.; Wusinke, R.; Morrison, P.W. *Mater. Res. Soc. Symp. Proc.* **2000**, *593*, 433–438.

28. Alanazi, A.; Nojiri, C.; Noguchi, T.; Kido, T.; Komatsu, Y.; Hirakuri, K.; Funakubo, A.; Sakai, K.; Fukui, Y. Improved blood compatibility of DLC coated polymeric material. *ASAIO J.* **2000**, 440–443.
29. Elinson, V.M.; Sleptsov, V.V.; Laymin, A.N.; Potray-say, V.V.; Kostaychenko, L.N.; Moussina, A.D. Barrier properties of carbon films deposited on polymer-based devices in aggressive environments. *Diamond and Related Materials* **1999**, 8, 2103–2108.
30. McLaughlin, J.A.; Meenan, B.; Maguire, P.; Jamieson, N. Properties of DLC thin film coatings on stainless steel medical guidewires. *Diamond and Related Materials* **1996**, 5, 486–491.
31. Schwank, M.; Mülder, U.; Hauert, R.; Rossi, R.; Volkert, M.; Wintermantel, E. Production of a micro-electrode for intracellular measurements based on a Pt/Ir needle insulated with amorphous hydrogenated carbon. *Sens. Actuators* **1999**, B56, 6–14.

Dielectric Properties of Tissues

Kenneth R. Foster

Department of Bioengineering, University of Pennsylvania, Philadelphia, Pennsylvania, U.S.A.

INTRODUCTION

This entry presents a nontechnical review of the dielectric properties of tissues, at frequencies ranging from DC to the microwave range. These properties (dielectric permittivity and conductivity) are needed for purposes such as the development of telemetry applications or diagnostic applications based on electrical measurements. They also can provide useful information about the structure and function of tissues. Special attention is paid to the historical development of this field.

HISTORY

The bulk electrical (or dielectric) properties of tissues and cell suspensions have been of interest to many investigators for more than a century, for both technological and scientific reasons.

Spurred both by the usefulness of dielectric studies in examining the physical properties of biological material, and advances in instrumentation, numerous scientists studied the dielectric properties of biological materials intensively during the early years of the 20th century. In the first decade of the century, Rudolf Höber (1886–1961) demonstrated by means of electrical measurements that cells had a thin insulating layer—which was one of the first convincing proofs of the existence of the cell membrane and a scientific discovery of the first order (for an interesting biographical sketch of Höber see Ref. [1]). Following this lead, a number of other distinguished scientists began to study the dielectric properties of tissues and cell suspensions. For example, in the early 1920s Hugo Fricke (1892–1972), the famous Danish biophysicist (and eventual head of Cold Spring Harbor Laboratory in New York, U.S.A.), published an influential series of papers on the electrical conductivity of cell suspensions.

These studies, which involved painstaking measurements of the physical properties of cells and biological materials and the interpretation of the results in terms of quantitative models, might be considered to be early contributions to the modern field of biophysics. Because the information that such measurements can

provide about the electrical characteristics of cell membranes is limited, other investigators began to explore the use of more refined electrical measurements. A major breakthrough was introduced by K.S. (known universally to cell biologists as “Kacy”) Cole (1900–1984), who pioneered the use of impedance techniques to study cellular membranes from the giant squid axon using electrodes placed inside the cells. His simultaneous recording of the action potential and impedance change in a squid axon (published in 1939) was a major development in electrophysiology.

Other important dielectric studies in the first part of the 20th century were conducted by physical chemists who wished to learn about the shape and motion characteristics of proteins. Famous work along those lines was conducted starting in the early 1930s by the American chemist J. L. Oncley (1910–2004) and R. H. Cole (1914–1990) (brother of K. S. Cole).

After the middle of the 20th century, dielectric measurements on bulk samples were largely supplanted by advanced electrophysiological techniques that allowed investigators to study directly the electrophysiology of cell membranes. Other spectroscopic techniques appeared, such as magnetic resonance, which provided more specific information about molecular properties. Dielectric spectroscopy remains an important scientific tool in polymer and colloidal chemistry and other fields.

Apart from their scientific interest, the development of applications of electromagnetic fields in medicine has led to a growing need for accurate data on the dielectric properties of tissues and biological materials. Already in the mid-1930s, groups at Siemens (Erlangen, Germany) and General Electric (Schenectady, New York, U.S.A.) and elsewhere were developing therapeutic applications of radiofrequency energy, including diathermy for treatment of pain and radiofrequency therapy for cancer. This work has continued, with many investigators around the world developing radiofrequency methods for hyperthermia. Developing treatment devices, and modeling the thermal consequences of treatment, requires knowledge of the bulk electrical properties of tissues. Studies on health and safety of radiofrequency energy, which began in earnest in the United States and other countries shortly after the end of World War II, created a need for dielectric

properties of tissues for purposes of dosimetry. Finally, numerous techniques have appeared that depend on measurements of the electrical properties of the body to detect disease and monitor physiological processes.

Herman P. Schwan (1915–2005), a German-born scientist who spent most of his career at the University of Pennsylvania, is undoubtedly the best-known investigator in the dielectric properties of tissues during the mid to the late 20th century.^[2] Schwan developed measurement techniques that remain unsurpassed in accuracy and precision today, and he wrote definitive review articles on the dielectric properties of tissues and their measurement techniques.^[3,4] The present entry is based on a longer review written jointly with Schwan.^[1]

CONDUCTIVITY AND PERMITTIVITY

Electrical conductivity σ and permittivity $\epsilon'\epsilon_0$ are measures, respectively, of the current flow or polarization induced by an electric field in matter. (Analogous properties can be defined for magnetic fields, but will not be considered here.) These properties are usefully defined over distance scales exceeding the size of individual molecules. In complex materials such as tissue, the distance scale of interest depends on the needs of a particular investigation. For example, when studying the absorption of microwave energy for purposes of dosimetry in health studies, the required spatial resolution is usually in millimeters or greater. In that case, the dielectric properties of a tissue can be measured using bulk samples of millimeter or centimeter dimension. For studies of the distribution of an externally imposed field within cells and across cell membranes, information is needed about the dielectric properties over much smaller distances.

The following discussion assumes that the response is linear (i.e., the induced current is proportional to the applied field). This is usually the case when the applied field is small (a few volts per meter or less), in which case the dielectric properties typically reflect physical mechanisms of charge transport in the material as opposed to biological responses. A variety of nonlinear effects are well known, both biological (such as membrane excitation) and nonbiological (e.g., electrical breakdown of cell membranes), which are separate from dielectric effects as considered in this entry. These nonlinear responses generally require quite strong fields, hundreds of volts per meter or more. For technological applications, which typically involve fields that are insufficient to produce nonlinear responses, the linear dielectric properties are usually sufficient.

The conductivity σ and permittivity $\epsilon'\epsilon_0$ are measured by applying an alternating electric field to the sample, and measuring the current that flows in phase

(conductive current) and 90° out of phase (reactive current) with the applied field. The relative permittivity ϵ' (no units) is the permittivity of a medium divided by ϵ_0 , the permittivity of free space (8.85×10^{-12} F/m). The conductivity σ has dimensions siemens per meter (S/m).

These properties are illustrated in terms of an idealized capacitor consisting of two electrodes of area A and separation d (Fig. 1). The region between the plates is filled with material of permittivity ϵ' relative to free space and conductivity σ . A voltage V is impressed between the plates. The electric field between the plates is assumed to be uniform. The capacitance C and the conductance G of this system are:

$$\begin{aligned} C &= \epsilon'\epsilon_0 A/d \\ G &= \sigma A/d \end{aligned} \quad (1)$$

For tissues and other biological materials that contain water and ions, the major contribution to the conductivity is the motion of mobile ions. By contrast, the permittivity reflects the polarization of bound charges in the tissue and other effects. Because the electrical transport mechanisms are time dependent, the current and polarization will both change with time after a field is suddenly imposed. In that case, the capacitance and conductance in Eq. (1) would be values measured after the system has reached a steady state after the imposition of a field.

An alternative approach to characterizing the dielectric properties of a material is to apply sinusoidal fields, in which case ϵ' and σ would be functions of frequency.

It is useful to consider the dielectric properties as parts of a single complex quantity. This can be seen by noting that the capacitance and conductance in Eq. (1) can be combined into a complex admittance Y^*

$$\begin{aligned} Y^* &= G + j\omega C \\ &= \frac{A}{d}(\sigma + j\omega\epsilon'\epsilon_0) \end{aligned}$$

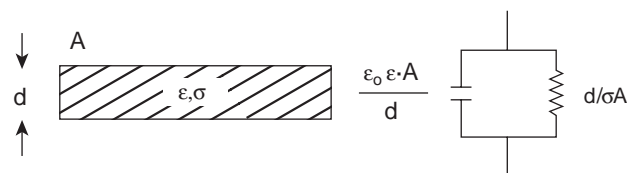


Fig. 1 Equivalent circuit of an idealized parallel plate capacitor filled with material of relative permittivity ϵ' and conductivity σ . The plate area is A and the distance between plates is d . In a real circuit, a small additional shunt capacitance would occur from the stray fields in the air outside the material. (From Ref. [5].)

where

$$\omega = 2\pi f \quad \text{and} \quad j = \sqrt{-1} \quad (2)$$

We can define a complex conductivity σ^* as:

$$\begin{aligned} \sigma^* &= \sigma + j\omega\epsilon'\epsilon_0 \\ &\equiv \sigma' + j\sigma'' \end{aligned} \quad (3)$$

and a complex permittivity ϵ^* as:

$$\begin{aligned} \epsilon^* &= \epsilon' - j\sigma/\omega\epsilon_0 \\ &\equiv \epsilon' - j\epsilon'' \end{aligned} \quad (4)$$

where the asterisks denote complex quantities. Defined in this way, the imaginary part of the complex permittivity (known as the loss factor ϵ'') is proportional to the measured conductivity of the material. An important theoretical result, called the Kramers–Kronig relations, is that the real and imaginary parts of ϵ^* cannot vary independently with frequency.^[5]

The ratio ϵ'/ϵ'' is a measure of the relative contributions of the capacitive to conductive nature of current flow in the material. When this ratio is less than 1, the material is primarily conductive, i.e., the current is nearly in phase with the applied field. This is typically the case with tissues at frequencies below about 1 MHz. In the other limit, when ϵ'/ϵ'' is much less than 1 (which is sometimes the case at high frequencies) the tissue is electrically similar to a dielectric material.

Arrangements similar to that shown in Fig. 1 can be used to measure the dielectric properties of biomaterials, with due consideration of a variety of real-world effects such as electrode artifacts, stray fields, and wave propagation effects at high frequencies. An extensive, and still valuable, review of dielectric measurement techniques was done by Schwan.^[2] Fig. 2 shows a sample cell that was used by Schwan and his colleagues for many years to study the dielectric properties of protein and cell suspensions, in which the capacitor “plates” are small platinum-coated electrodes that are surrounded by the material being measured.^[6] The cell would be calibrated using known dielectric fluids such as alcohol and aqueous electrolyte solutions; the electrodes are coated with electrolytically deposited platinum (called platinum black) to reduce electrode artifacts. Depending on the sample being measured, such a cell would be useful between a few megahertz and several hundred megahertz. At lower frequencies, electrode artifacts would become severe, while at higher frequencies wave propagation effects would become troublesome.

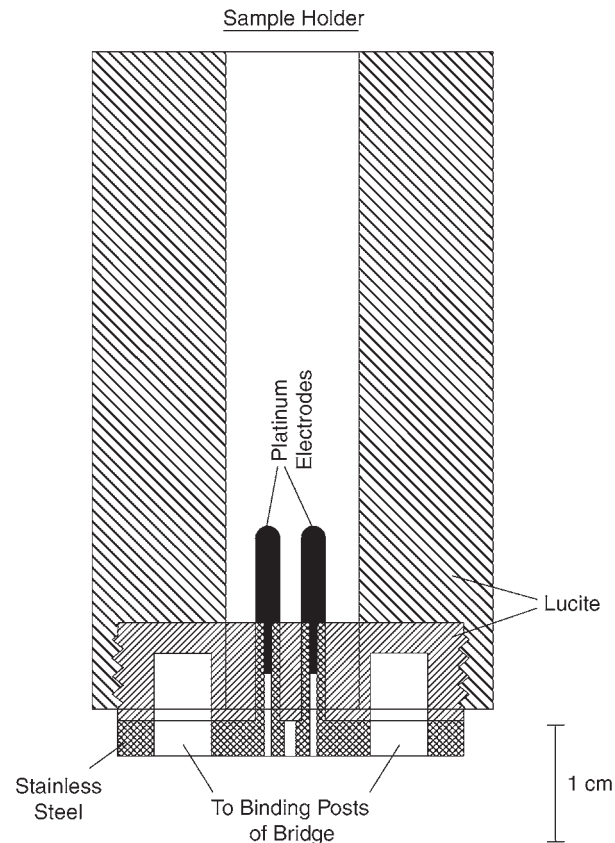


Fig. 2 Schematic of measuring cell used by Schwan for many years in studies of dielectric properties of tissues, cell suspensions, and biopolymer and protein solutions. The electrodes at the bottom of the cell are platinum, covered with a layer of electrolytically deposited platinum black to reduce electrode artifacts. A water jacket for thermal control surrounds the measurement volume. The cell would be used with an impedance bridge, and was usable over the frequency range from about 1 to several hundred MHz. (From Ref. [4].)

DIELECTRIC RELAXATION

The permittivity and conductivity are typically functions of frequency, and studying these properties over a wide frequency range yields a kind of spectroscopic information about the material. At different frequencies, different mechanisms of charge transfer contribute differently to the conduction and polarization.

As a simple example, consider a material whose polarization, in response to a step change in field, relaxes to its final value as a single exponential process with time constant τ . In response to an applied field that oscillates sinusoidally at frequency ω radians per second (or $\omega/2\pi$ Hz), the complex permittivity can be written as follows:^[5]

$$\epsilon^* = \epsilon_\infty + \frac{\epsilon_s - \epsilon_\infty}{1 + j\omega\tau} \quad (5)$$

or, separating into real and imaginary parts,

$$\begin{aligned}\varepsilon' &= \varepsilon_\infty + \frac{\varepsilon_s - \varepsilon_\infty}{1 + (\omega\tau)^2} \\ \varepsilon'' &= \frac{(\varepsilon_s - \varepsilon_\infty)}{1 + (\omega\tau)^2} \omega\tau\end{aligned}\quad (6)$$

In the above expression, ε_s is the value of the permittivity in the limit of low frequencies, while ε_∞ is the limiting value at high frequencies.

Fig. 3 shows the variation in permittivity ε' , loss, and conductivity with frequency for this dispersion. The center frequency of the dispersion is called the relaxation or characteristic frequency f_c , and is equal to $1/(2\pi\tau)$ Hz. In this example, with single time constant relaxation, most of the change in permittivity occurs over a decade of frequency centered about f_c . The conductivity σ increases by an amount

equal to $(\varepsilon_s - \varepsilon_\infty)/(\varepsilon_0\tau)$ over this relaxation. The broad frequency dependence of this relaxation stands in contrast to much narrower resonance peaks that are measured by various spectroscopic techniques including MRI.

Characteristically, the dielectric properties of biological materials are far more complex than the single time constant relaxation given by Eqs. (5) and (6). Various reasons include the presence of multiple relaxation mechanisms and the structural complexity of the material. There are also dielectric relaxation processes with nonexponential relaxation characteristics. As a result, the dielectric properties of these materials, when plotted vs. frequency, exhibit far broader dispersions than would be expected for single time constant relaxation. If the relaxations are widely separated in frequency, multiple dispersion regions might be apparent in the data. More commonly,

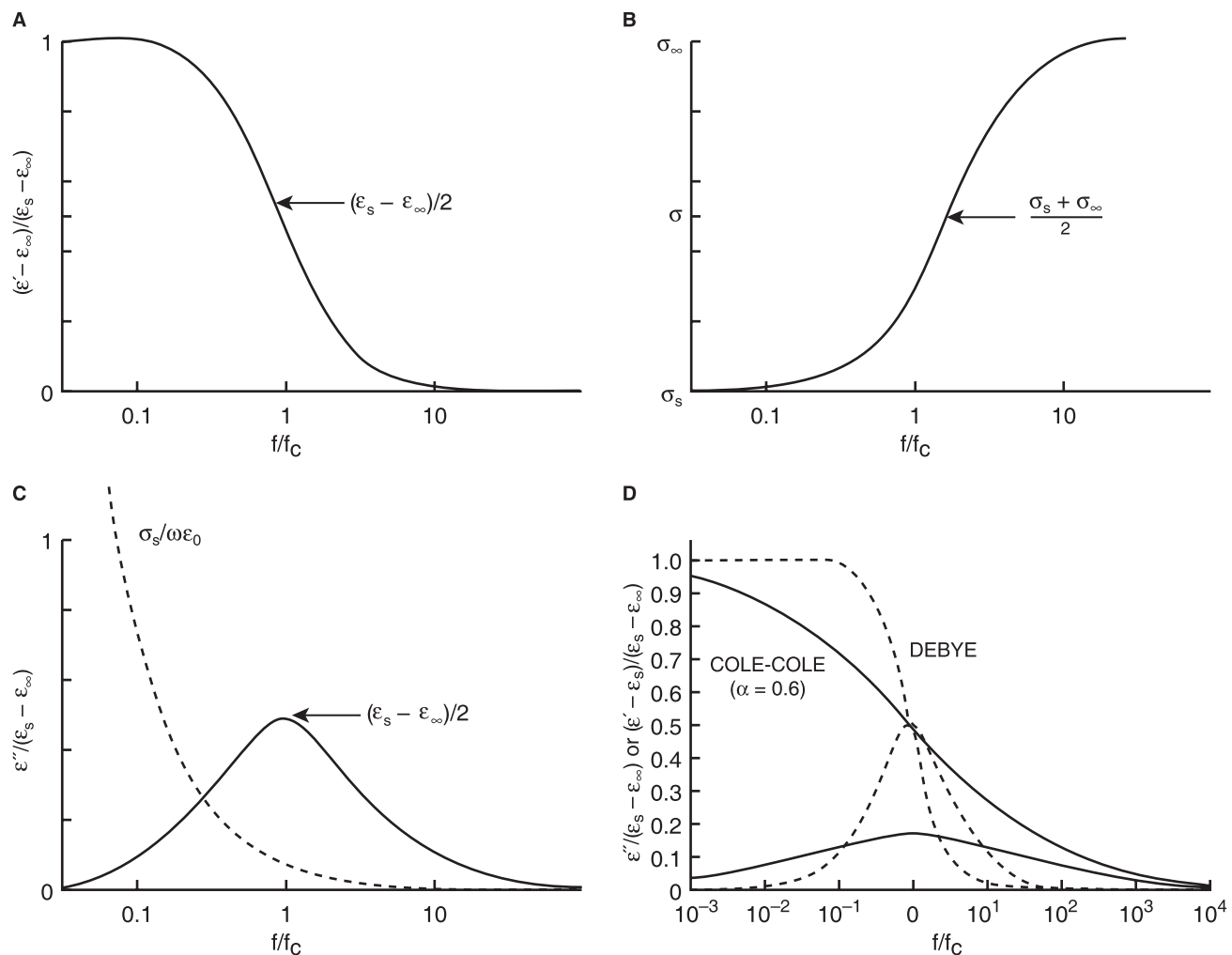


Fig. 3 Single time constant relaxation. (A) Relative permittivity vs. frequency; (B) conductivity vs. frequency; (C) loss vs. frequency showing (broken line) a contribution from a DC conductivity; and (D) comparison of single time constant relaxation with broadened relaxation described by the Cole-Cole equation, representing the effects of a distribution of relaxation times. (From Ref. [5].)

however, the dispersions overlap considerably, which complicates the interpretation of the data.

A particular mathematical form, the Cole–Cole expression, is often successful in empirically describing the dielectric properties of biological materials over a broad frequency range. This equation, named after K. S. Cole and R. H. Cole, who first introduced it in 1941, can be written as:^[7]

$$\varepsilon^* = \varepsilon_\infty + \frac{\varepsilon_s - \varepsilon_\infty}{1 + (j\omega\tau)^{1-\alpha}} \quad (7)$$

In Eq. (7), α is a parameter that characterizes the breadth of the dispersion. In the limit where α approaches zero, this approaches Eq. (5) for single time constant behavior. Fig. 3 compares a single time constant relaxation with that described by the Cole–Cole equation.

The Cole–Cole equation is frequently successful in describing dielectric data from many materials over broad frequency ranges, and experimentalists routinely fit their data to it. Unfortunately, the equation is empirical only, and yields no particular insights into the underlying mechanisms involved.

DIELECTRIC RELAXATION MECHANISMS IN BIOLOGICAL MATERIALS

Several major, and a number of minor, mechanisms contribute to the dielectric properties of tissues and other biological materials. Depending on the material and frequency range of interest, one or another might predominate in its influence.

Dipolar Relaxation

Many molecules and biological structures possess electric dipole moments owing to their distribution of charges. An externally imposed electric field will exert a torque on the dipoles, causing them to partially align, contributing to the polarization in the material. This effect is an important source of polarization in dipolar liquids such as water, and in solutions of dipolar molecules including proteins. A theory, originally attributed to the Dutch chemist Peter Debye (who won the Nobel Prize in chemistry in 1936 for his work on dipole molecules) provides a simple quantitative model for the phenomena.^[8]

The Debye theory considers an ensemble of independent dipoles, each with dipole moment μ . An externally imposed electric field of intensity F will exert a torque on each dipole given by $\mu F \sin(\theta)$, where θ is the angle between the dipole moment and the field. This torque will act to orient the dipole, but it will be

opposed by random thermal agitation. Debye calculated the mean orientation of a molecule in the ensemble to be:

$$\langle \cos(\theta) \rangle \approx \frac{\mu F}{3kT} \quad (8)$$

where k is Boltzmann's constant (1.38×10^{-23} J/K) and T is the absolute temperature (K). Eq. (8) refers to a mean orientation, averaged over many dipoles. At any realistic field strength, this mean orientation is very small; the effects of the torque to align the molecules are far overwhelmed by random thermal agitation.

When a field is suddenly imposed on the ensemble of dipoles, the polarization will develop gradually as the dipoles diffuse in the viscous medium in which they are located. For a spherical molecule of radius a , the mean relaxation time τ can be obtained by considering the diffusion of the sphere in a viscous medium. Application of Stokes' law gives:

$$\tau = \frac{4\pi\eta a^3}{kT} \quad (9)$$

where η is the viscosity (Poise).

The relaxation time τ ranges from picoseconds for small dipolar molecules such as water, to microseconds for large globular proteins in aqueous solution. The corresponding dielectric dispersion, measured with alternating fields, will have a center frequency that ranges from megahertz frequencies (for protein molecules) to 20 GHz (for water). If the molecule is not spherical, as is typically the case with proteins, additional dielectric dispersions will appear that correspond to rotational diffusion along each principal axis of the molecule. Table 1 summarizes the dielectric properties of various proteins, from Oncley.^[9]

Water, the major component by volume of soft tissues, shows a strong dipolar relaxation that is very close to single time constant in nature, with a low-frequency permittivity of 78 and center frequency of about 20 GHz at room temperature. This is well above the frequency range of most other relaxation processes that occur in tissue. Consequently, at frequencies above about 1 GHz, the dielectric properties of tissue are dominated by those of water. This fact underlies the frequent use of dielectric techniques to measure the water content of agricultural and other materials.

Counterion Polarization

Both biological and nonbiological materials characteristically show huge dielectric dispersions at low frequencies. Tissues, for example, show astonishingly large permittivity values at low frequencies, in the range of tens of millions for soft tissues at frequencies

Table 1 Dielectric properties of various proteins at 25°C

Protein	Mol wt ($\times 10^{-3}$) (Da)	$\Delta\epsilon'/\text{g/L}$	μ (D)	$\tau \times 10^8$ (sec)	a/b
Ovalbumin	44	0.10	250	18; 4.7	5
Horse serum albumin	70	0.17	380	36; 7.5	6
Horse carboxyhemoglobin	67	0.33	480	8.4	1.6
Horse serum pseudoglobulin	142	1.08	1100	250; 28	9
β -Lactoglobulin (in 0.25 M glycine)	40	1.51	730	IS; 5.1	4
Myoglobin	17	0.15	170	2.9	

Mol wt, molecular weight (daltons); μ , dipole moment (debye units); τ , rotational relaxation times; a/b , axial ratio; $\Delta\epsilon'/\text{g/L}$, total increase in permittivity per gram of protein per gram of solution, measured at low frequencies above the permittivity of pure water.

of 1 Hz or below. Similar effects are found in colloidal and other nonbiological suspensions, indicating that the phenomenon is essentially nonbiological in origin. The effect is associated with displacements of clouds of mobile charges in the material, and not the orientation of permanent dipoles.

This phenomenon has been subjected to extensive theoretical and experimental work, and the effect is understood in principal even though a detailed analysis of the effect in most biomaterials remains beyond reach. The first important theory for the effect was proposed by the Swiss theoretical chemist G. Schwarz in 1962, who postulated that counterions diffuse along the surfaces of the particles under the influence of the applied field, inducing a dipole moment in the particle when charges of opposite sign accumulate at each end.^[10]

Schwarz's theory is simple and intuitively appealing, and has been cited in hundreds of papers over the years. But scientists now accept a very different explanation of the effect, involving diffusion of ions in the bulk electrolyte surrounding the particle (not, as in Schwarz's model, counterions on the surface of the particle). This diffusion leads to regions in the electrolyte near the surface of the particle where the electrolyte concentration is higher or lower than in the surrounding bulk electrolyte solution.

Because of its importance in understanding colloidal systems, a great many theoretical and experimental studies have been conducted on the effect over the years, typically with reference to nonbiological systems including colloidal suspensions or suspensions of polystyrene microparticles. Important contributions to the problem were published in a 1974 book by Ukrainian investigators Stanislav Dukhin and Vladimir Shilov; however, perhaps because of a flawed translation the work received little attention in the West.^[11] The problem subsequently, attracted the attention of numerous scientists in Argentina, Australia, the Netherlands, Spain, Ukraine, as well as the United States, and other countries. Important contributions to the problem have been recently made by A. Dukhin, the son of S. Dukhin and a distinguished physical

chemist in his own right. An excellent (now dated) review of this modern theory is by Mandel and Odijk.^[12]

While the modern theory is surely more correct than the simple model for "counterion polarization" developed by Schwarz, it is not necessarily more useful to the experimentalist. The theory, which has to take into account both electrical and hydrodynamic forces on the ions, is mathematically very complex. Presently, it can be applied rigorously only to the simplest systems, such as dilute suspensions of polystyrene microspheres. This is a serious impediment to trying to interpret the low-frequency dielectric dispersion that is found in most biomaterials, which are neither structurally simple nor dilute in nature.

Interfacial Polarization (Maxwell-Wagner Effect)

If a material is electrically heterogeneous, a dispersion occurs in the bulk properties from the charging of the interfaces within the material. This phenomenon does not arise from dielectric relaxation in the bulk phases of the material or movement of counterions (which might also be present), but is purely a macroscopic effect arising from the heterogeneity of the material leading to changes in the distribution of electric fields within the material at different frequencies. This effect is named after K. W. Wagner, who published an influential paper on the effect in 1914, extending an earlier result by the famous physicist James Clerk Maxwell. The theory is straightforward but mathematically cumbersome, and is described in more detail in^[5] and many other places.

One example of biological significance is a suspension of membrane-covered spheres that are filled with and surrounded by electrolyte solution, which is a simple model of a cell suspension.^[5] Because the membranes are thin, they have a large capacitance whose impedance decreases inversely with frequency. The membranes are considered to be nonconductive, but (because they are thin) they have a capacitance similar to that of cell membranes.

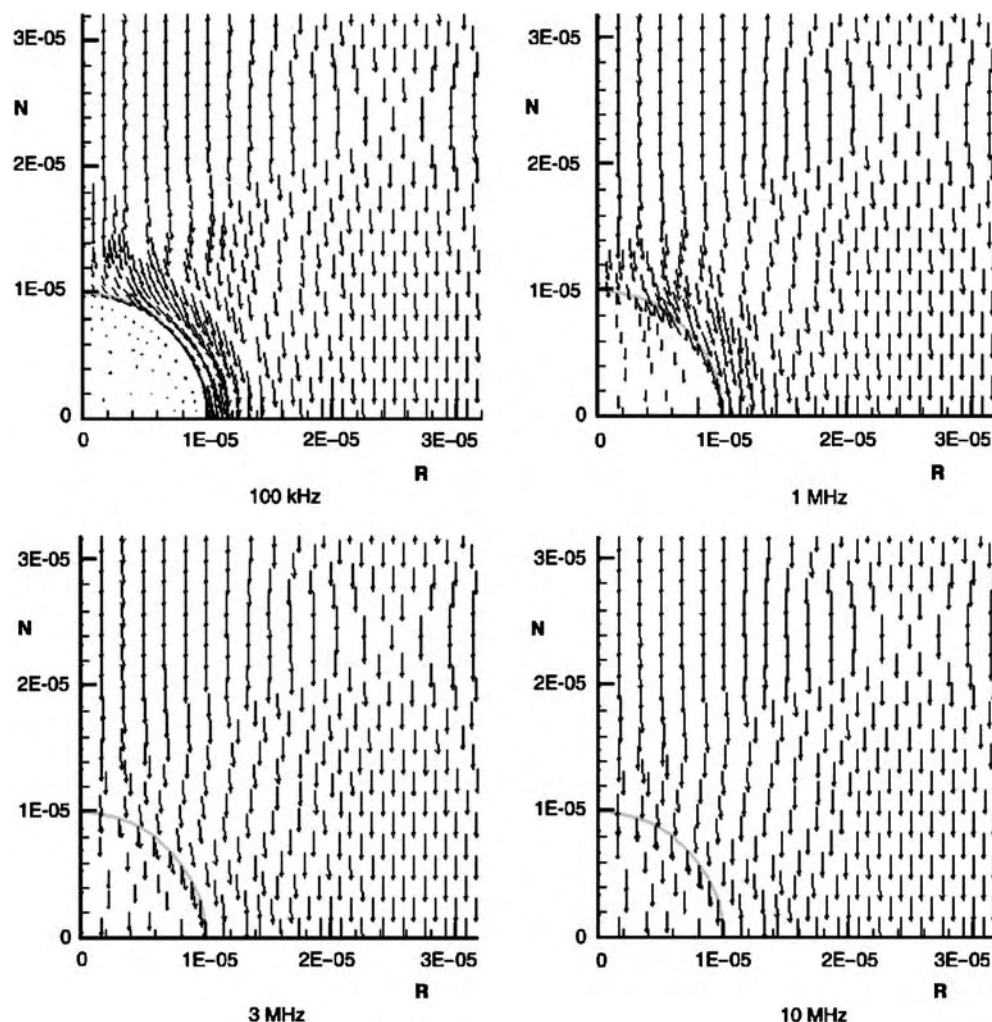


Fig. 4 Finite element calculation, showing paths of currents around and through a membrane covered sphere (approximating a cell) at different frequencies. The sphere was chosen to have properties reminiscent of those of a cell: radius $10\ \mu\text{m}$, conductivity of inner and outer media $1\ \text{S/m}$, membrane capacitance $0.01\ \text{F/m}^2$. The figures show one-quarter of each cell; because of symmetry the patterns of current flow would be repeated in the other three quadrants that are not shown.

Fig. 4 shows a numerical solution for the current flow around a single sphere, obtained using a finite element program. At low frequencies, the current flows around the spheres because of their high impedance. As the frequency is increased, progressively more current flows through the cells. At very high frequencies, the membranes are short-circuited owing to their capacitance, and no longer impede current flow. At low frequencies, the conductivity of the suspension will be low (owing to the presence of the nonconducting sphere) and the permittivity will be high (owing to the collection of charge against the membrane); at very high frequencies the permittivity and conductivity of the suspension will be close to those of the media that are contained by and surround the spheres.

This model can also be solved analytically, but the mathematics involved is cumbersome.^[5] For dilute

suspensions of spheres with properties similar to those of cells, approximate results can be written in the simple form:

$$\epsilon_s \simeq \frac{9pRC_m}{4\epsilon_0} \quad (10)$$

(relative permittivity of the suspension at low frequencies),

$$\sigma_s \simeq \sigma_a(1 - 3p/2) \quad (11)$$

(conductivity at low frequencies), and

$$\tau \simeq RC_m \left(\frac{1}{2\sigma_a} + \frac{1}{\sigma_i} \right) \quad (12)$$

(relaxation time).

In the above expressions, p is the volume fraction of the spheres (assumed to be small), R is the radius of the spheres (m), C_m is the capacitance of the membrane (F/m^2), and σ_a , σ_i are the conductivities of the outer and inner media (S/m).

To give a numerical example, for a suspension with electrical parameters typical of those of cells ($R = 10 \mu m$, $p = 0.1$, $\sigma_a = \sigma_i = 1 S/m$), the static permittivity is about 1000 and the center relaxation frequency is about 1 MHz. Suspensions of cells such as erythrocytes show quite similar properties.

This simple theory is remarkably successful in predicting the dielectric properties of cell suspensions. The theory can be extended to more concentrated suspensions, and to suspensions of nonspherical cells.^[5]

This simple theory has important implications for the dielectric properties of tissues and cell suspensions. The dielectric properties of the suspension depend very little on the conductance of the membrane but depend quite strongly on the membrane capacitance and size of the cells. (The membrane conductance, which usually is of greater interest to cell biologists than the membrane capacitance, is very difficult to measure by straightforward dielectric techniques). Also, the conductivity of the suspension at low frequencies depends strongly on the volume fraction of extracellular liquid, which is the chief pathway for the current through the suspension.

This example has important implications concerning the possible biological effects of electromagnetic fields.^[5] When low-frequency currents are passed through the cell suspension, potentially substantial fields are imposed across the cell membrane, equal (for spherical cells) to the $1.5R$ times the field strength in the surrounding medium. At frequencies substantially above the relaxation frequency, the potential differences that result across the cell membranes are very much smaller. Indeed, at radio-frequencies and above, very little of the applied field is imposed across cell membranes, because they are short-circuited by virtue of their capacitance.

Dielectric Properties of Tissues

Now, after nearly a century of study, a vast literature exists on the dielectric properties of tissues and biological materials have accumulated. Fig. 5 indicates the dielectric permittivity of a typical soft tissue, from an early (1957) review by Schwan.^[3] The terminology invented by Schwan for the major dispersion regions, alpha, beta, and gamma, remains in use. Table 2 summarizes these major dispersion regions and the mechanisms that give rise to them.

Figs. 6–8 show more recent data, from excised sheep tissues at body temperature.^[13]

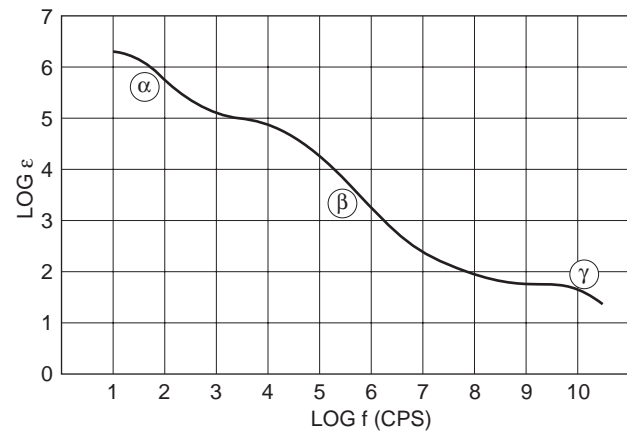


Fig. 5 A schematic illustration of three major dispersion regions found in typical tissues, with permittivity and conductivity values typical of soft tissues.

Several points are worth noting from these figures:

- The tissues all show broad and rather featureless dielectric relaxations that extend over a wide frequency range, which are not clearly separated in frequency. This complicates the interpretation of the data.
- Over most of the frequency range covered by the data, the tissues are primarily resistive ($\epsilon'' > \epsilon'$), i.e., the current and field are in phase. Thus, for many practical applications, the permittivity is not an important factor in determining current flow.
- The tissues show a pronounced dispersion above 1 GHz, which is due to dipolar relaxation of the tissue water content.
- The dielectric properties of all four tissues summarized in Figs. 6–8 show dispersions in the megahertz region which arise, at least in part, from the charging of cellular membranes.
- The dielectric properties of two of the tissues (brain and skeletal muscle) show large, low-frequency

Table 2 Major dispersion regions observed in the dielectric properties of tissues and their origin

Type	Frequency range	Mechanism
α	^a mHz– ^b kHz	Counterion effects; long-range structures in tissue
β	0.1–100 MHz ^c	Maxwell–Wagner effects; charging of cellular membranes
γ	0.1–100 GHz ^d	Dipolar reorientation of water

(From Ref. [3].)

^a1mHz = 10^{-3} Hz

^b1kHz = 10^3 Hz

^c1MHz = 10^6 Hz

^d1GHz = 10^9 Hz

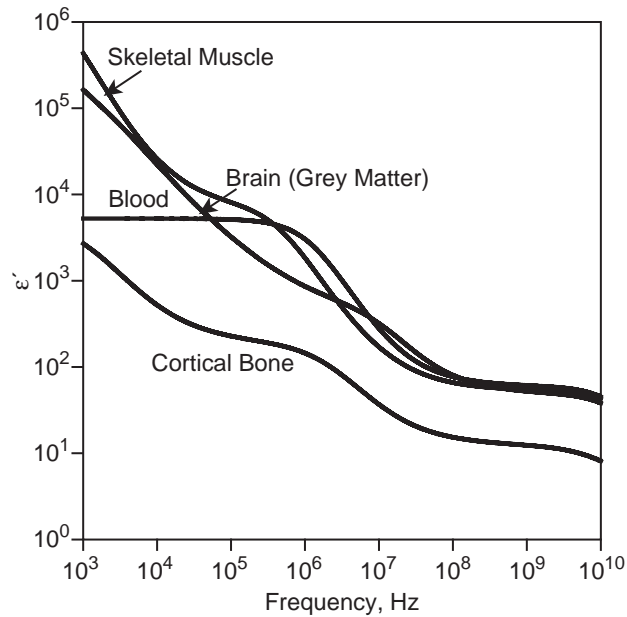


Fig. 6 Permittivity of four representative ovine tissues (skeletal muscle, grey matter of brain, blood, and cortical bone).

dispersions, which arise, at least in part, from counterion effects.

We consider these properties separately in different frequency ranges. The discussion will attempt to relate the theoretical discussion in the preceding section, with the data obtained from tissues.

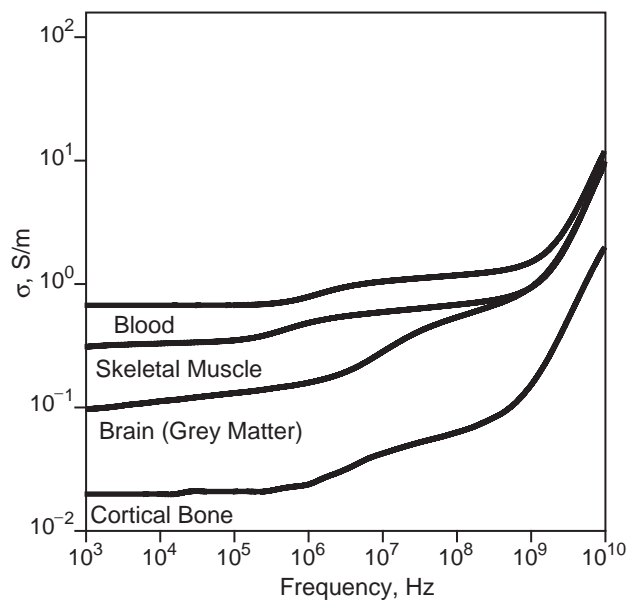


Fig. 7 Conductivity of four representative ovine tissues (skeletal muscle, grey matter of brain, blood, and cortical bone).

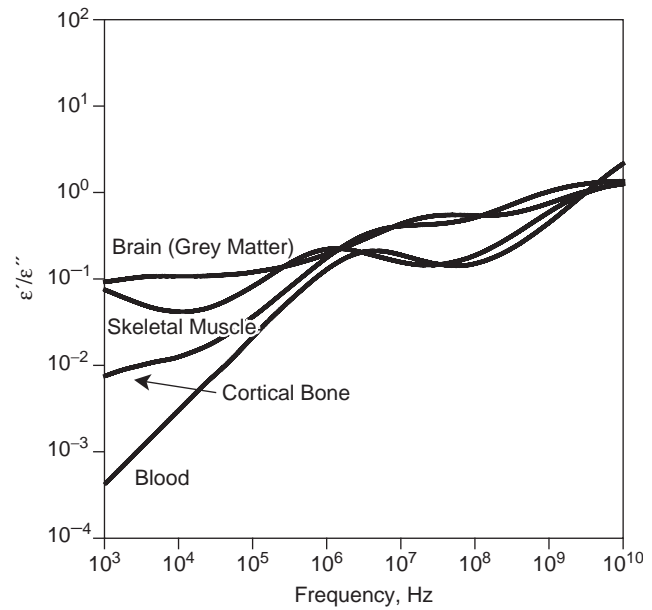


Fig. 8 Ratio of ϵ'/ϵ'' of four representative ovine tissues (skeletal muscle, grey matter of brain, blood, and cortical bone).

Low Frequencies (<0.1 MHz)

At frequencies below the beta dispersion frequency, current is largely excluded from the interior of cells by virtue of the high impedance of their membranes; the current flows primarily through the extracellular space. Indeed, in 1924 Hugo Fricke successfully modeled the electrical conductivity of blood using electrolyte-soaked sand. The conductivity of tissues, which is nearly independent of frequency, at low frequencies (Fig. 7) reflects conduction in the electrolyte in the extracellular space. At low frequencies, this conductivity depends on both the volume of extracellular space and the geometry of the cells [cf. Eq. (11)], but not the electrical characteristics of the cells themselves.

The permittivity of two of the soft tissues (muscle and brain) increases to very high values at low frequencies, far higher than that of blood. These reflect counterion effects, although other effects related to tissue structure (polarization in the sarcotubular system in muscle, for example) might be important also.

All of the tissues show relaxation behavior that is far broader than would be attributed to a single time constant process. Either multiple dielectric relaxation processes are occurring over this frequency range, or the dominant relaxation process cannot be described by single time constant behavior. Several authors have suggested that this broad relaxation arises from structural factors in the tissue, which give rise to a fractal-like behavior of the relaxation.^[14]

Frequencies Between 1 and 100 MHz

In this frequency range, the cell membranes are largely short-circuited owing to their capacitance and do not offer significant barrier to current flow. Thus, the tissues can be considered to be electrically equivalent to suspensions of nonconducting protein (and other solids) in electrolyte. The conductivity of most soft high water content tissues levels off noticeably around 10–100 MHz and comes to approach that of the electrolyte that constitutes 70–80% of soft tissues (Fig. 7). The conductivity of cortical bone in the same figure does not show this behavior because other conduction mechanisms dominate.

Other smaller effects have been identified between 100 MHz and 1 GHz. This includes dipolar relaxation of “bound water” near protein surfaces, and a Maxwell–Wagner effect that contributes a small dispersion at several hundreds of megahertz owing to the dissimilar electrical properties of tissue protein and electrolyte.^[5]

Microwave Frequencies (>1 GHz)

Above 1 GHz, the electrical conductivity of tissues increases dramatically with frequency owing to dipolar relaxation in the tissue water. For typical soft tissues, the conductivity at 3–5 GHz is due in roughly equal proportions to ionic conductivity and to dipolar loss of water. The permittivity is slightly lower than that of pure water owing to the displacement of water by solid content. At frequencies above about 10 GHz, the permittivity starts to decline noticeably, owing to the dipolar relaxation of the tissue water.

DIELECTRIC PROPERTIES—SOURCES OF DATA

Numerous dielectric data from tissues are available in the literature, over widely varying frequency ranges, and of varying reliability (for one extensive table see Ref. [5]). The data in the literature show considerable variability, even in samples of the same tissue type from animals of the same species. The dielectric properties vary with the condition of the samples being measured (e.g., freshly excised tissues vs. autopsy specimens vs. tissues *in vivo*), orientation of the sample being measured (skeletal muscle in particular is highly anisotropic in its properties), age and physical condition of the animal donating the tissues, volume of blood in the tissue, distribution of tissue fluids between intracellular and extracellular space, and other factors. These variations may be significant or not, depending on the needs of an investigator who uses them. While

there have been many attempts to use impedance techniques to monitor physiological events and detect pathology in humans, the lack of specificity of the dielectric properties with respect to biologically important endpoints has been an enduring problem in this field.

For technological applications, a particularly useful set of data have been provided by Gabriel and Gabriel.^[13] These data were taken from similar animals (in most cases, sheep) under similar conditions, over a broad frequency range. The investigators also provided empirical functions, based on the Cole–Cole equation Eq. (7) to enable the user to obtain representative values of the dielectric properties of numerous tissue types over the broad frequency range from 10 Hz to 20 GHz. These data are available on several websites (e.g., <http://www.brooks.af.mil/AFRL/HED/hedr/reports/dielectric/Report/Report.html> and <http://niremf.ifac.cnr.it/tisprop/>), and spreadsheets and other programs can be found on the Internet to facilitate the use of this data set. While not the last word on the subject, this standardized and widely accepted set of data greatly aids in the development and testing of many biomedical applications.

CONCLUSIONS

The dielectric properties of tissues have been of interest to scientists and engineers for many years, from the insights that they provide about the structure and properties of biological materials, for practical applications of electromagnetic fields in medicine, and for studies of possible health and safety problems from exposure to electromagnetic fields. A good theoretical understanding exists of the underlying mechanisms even though the theories may be difficult to apply in detail to specific biomaterials. Extensive collections of data, including one standardized set of data from many tissues that is widely available on the Internet, are of great benefit for the development of applications of electricity in medicine.

REFERENCES

1. <http://www.amphilsoc.org/library/mole/h/hoerber.htm>.
2. Foster, K.R.; Herman, P. Schwan: a scientist and pioneer in biomedical engineering. *Ann. Rev. Biomed. Eng.* **2002**, *4*, 1–27.
3. Schwan, H.P. Electrical properties of tissue and cell suspensions. In *Advances in Biological and Medical Physics*; Academic Press: New York, 1957; Vol. 5, 147–209.
4. Schwan, H.P. Determination of biological impedances. In *Physical Techniques in Biological Research*; Nastuk, W.L., Ed.; Academic Press: New York, 1963; Vol. 6, 323–403.

5. Foster, K.R.; Schwan, H.P. Dielectric properties of tissues—a review. In *Handbook of Biological Effects of Electromagnetic Radiation*, 2nd Ed.; Polk, C., Postow, E., Eds.; CRC Press: Boca Raton, 1995; 25–102.
6. Schwan, H.P. Electrical properties of tissue and cell suspensions. In *Advances in Biological and Medical Physics*; Lawrence, J.H., Tobias, C.A., Eds.; Academic Press: New York, 1957; 147–209.
7. Cole, K.S.; Cole, R.H. Dispersion and absorption in dielectrics: alternating current characteristics. *J. Chem. Phys.* **1941**, *9*, 341–351.
8. Debye, P. *Polar Molecules*; Chemical Catalog Co.: New York, 1929.
9. Oncley, J.L. In *Proteins, Amino Acids, and Peptides*; Cohn, E.J., Edsall, J.T., Eds.; Reinhold: New York, 1943; . chap 22, 543–568.
10. Schwarz, G. A theory of the low frequency dielectric dispersion of colloidal particles in electrolyte solution. *J. Phys. Chem.* **1962**, *66*, 2636–2642.
11. Dukhin, S.S.; Shilov, V.N. *Dielectric Phenomena and the Double Layer in Disperse Systems and Polyelectrolytes*; Greenberg, P., Ed.; John Wiley & Sons: New York, 1974.
12. Mandel, M.; Odijk, T. Dielectric properties of polyelectrolyte solutions. *Ann. Rev. Phys. Chem.* **1984**, *35*, 75–108.
13. Gabriel, S.; Lau, R.W.; Gabriel, C. The dielectric properties of biological tissues: II. Measurements in the frequency range 10 Hz to 20 GHz. *Phys. Med. Biol.* **1996**, *41*, 2251–2269.
14. Raicu, V.; Sato, T.; Raicu, G. Non-Debye dielectric relaxation in biological structures arises from their fractal nature. *Phys. Rev. E.* **2001**, *64*, 21916–21926.

Digital Biomedical Signal Acquisition and Processing: Basic Topics

D

Martin Oswaldo Mendez Garcia

Luca T. Mainardi

Department of Biomedical Engineering, Polytechnic University, Milan, Italy

INTRODUCTION

Biological signals carry useful information for probing, exploring, and understanding the behavior of biological systems under investigation. Often, such information cannot be readily accessed, being masked by noise or buried by other vital signals simultaneously recorded. In these cases, the raw signal has to be processed to yield useful results.

Biomedical signal processing (BSP) deals with representation, transformation, and manipulation of biomedical signals and their informative content. The objectives of BSP may be various, ranging from the enhancement of features (waveforms) of interest, the quantitative analysis of physiological systems (from cells to organs to the whole human organism), to the definition of reliable indexes to be used in clinical practice for diagnostic, therapeutic, and rehabilitation purposes.

In this article, the basic tools in the BSP field are presented. Far from being exhaustive, the article is intended as a review of the fundamental concepts of digital filtering, synchronous averaging, and spectral analysis in the special context of biomedical signals. The processing of biomedical signals is an intriguing, unique challenge in which the need to investigate complex biological systems is often associated with the availability of indirect, noninvasive observations of the phenomena. The article is organized as follows: in section “Acquisition,” a brief introduction to the acquisition of biomedical signals is presented, while classical techniques for filtering and signal averaging are described in section “Signal Enhancement.” Section “Spectral Analysis” is dedicated to fundamental concepts of spectral analysis while the problem of extracting relevant signatures from vital signals is presented in section Event Detection. In each section, a few examples will be used to illustrate the problems related to the analysis of biomedical signals and possible solutions. Only basic approaches will be presented. More advanced methods will be described in a latter article of this encyclopedia.

ACQUISITION

Acquisition deals with the problem of converting physical magnitudes (biological signals) into digital data. Basics steps for signal acquisition are: signal sensing, signal conditioning, and Analog-to-Digital (A/D) conversion. “Sensors” (or transducers) serve as interface between biological and electrical system: They pick up signals representing biological variables and convert them into electric signals, which may be easily handled, processed, and transmitted. Several books on biomedical instrumentation provide detailed descriptions of the various transducers and hardware employed in the acquisition of biomedical signals.^[1,2]

After sensing, the signal must be conditioned to satisfy the requirements of the following hardware. A conditioner stage is used to “amplify” the signal [biomedical signals are quite small in amplitude: 1–10 mV for the electrocardiogram (ECG) in the adults—around 10 μ V in the fetus—from 2 to 100 μ V for the electroencephalogram (EEG)], to “compensate” for some nonlinearity in the sensor characteristics, to adapt impedances, or to reduce the amount of undesired noise. In addition, the signal is limited in band (antialiasing filter) to satisfy the requirement of the sampling theorem (see section “The Sampling Process”). The conditioner stage must preserve the information contained in the original biological signal, whose waveforms are analyzed by the physician to detect signs of pathologies (i.e., displacement of ST-segment in the ECG, peak-and-wave patterns in EEG, and so on). Any distortion introduced by the acquisition procedure may lead to misleading interpretation of the data and may prevent in correct diagnoses. For this reason, requirements on the design of the conditioner stage are usually restrictive: The block must be designed with a “linear-phase” and “constant-modulus” frequency response at least in the band of interest. A linear phase introduces a purely time delay, thus avoiding phase distortion in the signals. Finally, the analog signal is A/D converted into a digital one. The A/D conversion can be conceptually divided into two

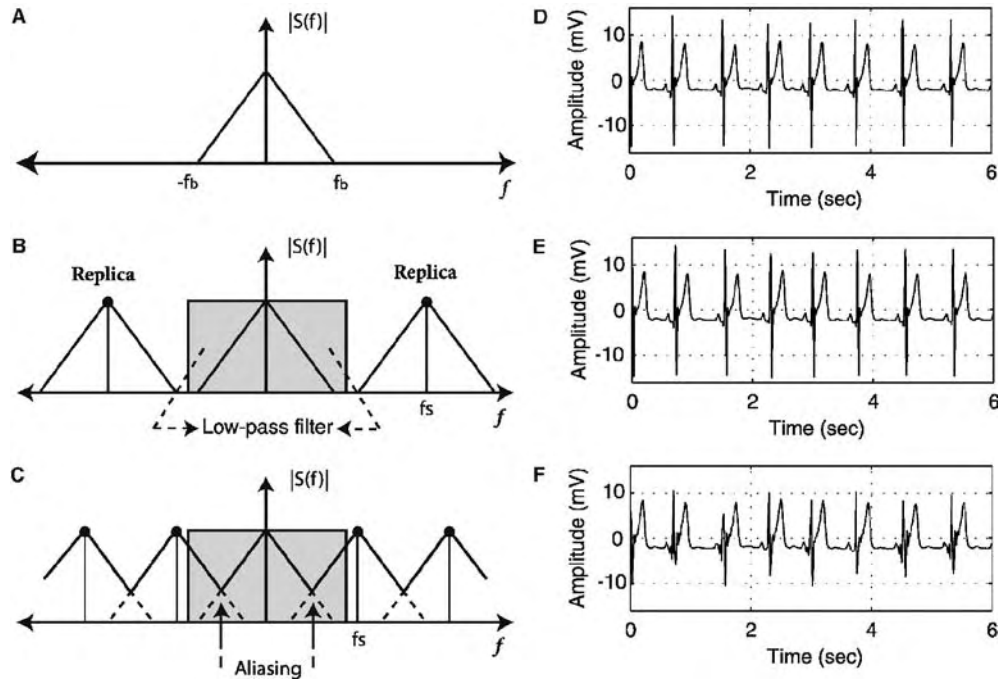


Fig. 1 (A) Fourier transform of a continuous-time band-limited signal; (B) FT of the sampled version of (A) when the sampling theorem is satisfied ($f_s > 2f_b$): the original signal may be recovered with a low-pass filter (rectangle); (C) aliasing effect produced when $f_s < 2f_b$. The replicas overlap with the original spectrum generating a corrupted signal; and (D) original ECG signal (band limited at 250 Hz), and its discrete-time versions sampled at 500 Hz (E) and 50 Hz (F), respectively. Note the alterations in the QRS shape induced by inappropriate sampling rate.

different processes: sampling and quantization. These processes will be briefly discussed from a signal processing point of view.

The Sampling Process

The “Sampling Theorem” (or Shannon Theorem) states that a continuous-time, band-limited signal is entirely represented by its discrete-time (sampled) version, if and only if, the sampling rate is greater than twice the signal bandwidth. When this hypothesis is satisfied, the sampled signal can be processed instead of the original waveform, without losing or destroying any information.

Let us consider a continuous waveform $s(t)$ and its band-limited (between $-f_b$ and f_b) Fourier transform (FT), $S(f)$ (Fig. 1A). The sampled signal $s_s(t)$ is obtained by multiplying the continuous waveform $s(t)$ with an impulse train $i(t)$:

$$\begin{aligned} s_s(t) &= s(t) \cdot i(t) = s(t) \cdot \sum_{k=-\infty}^{\infty} \delta(t - kT_s) \\ &= \sum_{k=-\infty}^{\infty} s(t) \cdot \delta(t - kT_s) \end{aligned} \quad (1)$$

where the $\delta(t)$ is the Dirac delta function, T_s is the sampling interval, and k is an arbitrary integer. In the

frequency domain we obtain

$$\begin{aligned} S_s(f) &= S(f) * I(f) = S(f) * \frac{1}{T_s} \\ &\times \sum_{k=-\infty}^{\infty} \delta(f - kf_s) \\ &= \frac{1}{T_s} \sum_{k=-\infty}^{\infty} S(f - kf_s) \end{aligned} \quad (2)$$

In Eq. (2), $*$ denotes the operation of convolution, and $f_s = 1/T_s$ corresponds to the sampling frequency. It can be appreciated from Eq. (2) that the FT of $s_s(t)$ is an infinite and periodic replication of the FT of $s(t)$. Each replica of $S(f)$ is centered around multiples of f_s . Let us consider the two situations depicted in Fig. 1B and C. In Fig. 1B ($f_s > 2f_b$, i.e., the sampling theorem is satisfied), there are no overlaps among the replicas and the original signal may be easily recovered by low-pass filtering the sampled signal. Conversely, when $f_s < 2f_b$, adjacent replicas overlap and higher frequencies are summed up to lower components: the phenomenon is known as “aliasing.” In the presence of aliasing, the spectral content of the original signal is irreversibly corrupted, and there is no way of recovering it.^[3–5] Fig. 1D depicts a band-limited ECG signal ($f_b = 250$ Hz), while Fig. 1E and F show the same ECG sampled at

500 Hz and 50 Hz, respectively. In Fig. 1E, the signal maintains all the waveform characteristics of the original one, because it was correctly sampled. Conversely, in Fig. 1F, the QRS complexes are altered because the adopted sampling frequency is lower than the maximum frequency contained in the signal. It is worth noting that the appropriate sampling frequency depends either on the frequency content of the signal or on the purpose of the analysis. For example, the recommended ECG sampling rate for diagnostic electrocardiography is 500 Hz (American Heart Association recommendations). The sampling rate can be lowered for long-term acquisition (ambulatory ECG) or increased (>1 kHz) for high-resolution ECG [i.e., for the analysis of ventricular late potential (VLP) or pacemaker functions].^[6]

Quantization

So far, we are supposing that the amplitude of each discrete-time sample is set within a continuous range of values. In practice, this amplitude must be selected among a finite, predefined set of values, depending on the characteristics of the A/D converter (ADC). When a uniform quantizer is used, the ADC voltage range is uniformly divided into equally spaced intervals (the quantization interval Δ is constant: $\Delta = 2^{-b}$, where b is the number of bits used for coding), and the signal amplitude is assigned to the nearest binary level (rounding) or to the closest lower level (truncation). Obviously, this procedure introduces some errors in the description of the signal.

Even if quantization is a highly nonlinear procedure, some of its effects can be modeled statistically. Let us assume that the sampled and quantized signal $s_q(n)$ can be described by the following model

$$s_q(n) = s_s(n) + e(n) \quad (3)$$

Here the effect of quantization is seen as an additive noise term, $e(n)$, superimposed onto the sampled signal. The $e(n)$ is known as “quantization error” and, to deal with a simplified mathematical problem, is usually supposed to satisfy the following hypothesis:

1. $e(n)$ is a white noise with uniform distribution.
2. $e(n)$ is uncorrelated with $s(n)$.

Depending on the adopted coding procedure, the probability density function (PDF) of $e(n)$ is different: in the case of rounding, the PDF is uniform in the range $-\Delta/2 < e(n) < \Delta/2$, with zero mean and the variance (σ_e^2) equal to $\Delta^2/12$; in case of truncation, the PDF is still uniform but within the interval $-\Delta < e(n) < 0$, the mean is $-\Delta/2$ and the variance is unchanged. As

$\Delta = 2^{-b}$, the variance of the error is reduced when the number of bits is increased.^[4]

Because $e(n)$ is seen as an additive noise corrupting the signal, one can estimate the signal-to-noise ratio (SNR) of the quantization process. Defining σ_s^2 as the variance of the signal, we have

$$\begin{aligned} \text{SNR}_{\text{dB}} &= 10 \log_{10} \left(\frac{\sigma_s^2}{\sigma_e^2} \right) \\ &= 6.02b + 10.79 + 10 \log_{10}(\sigma_s^2) \end{aligned} \quad (4)$$

It is interesting to note that SNR increases by 6 dB for each added bit of coding.

In biomedical applications, several quantization procedures can be adopted, including uniform quantization (described above), nonuniform quantization (as logarithmic, exponential, look up table), and semiuniform quantization.^[3-5,7]

SIGNAL ENHANCEMENT

The relevant information contained in the acquired biomedical signal may be not immediately available. It may be masked by “endogenous” and “exogenous” noise sources. Endogenous effects include disturbances related to other biomedical signals simultaneously recorded such as influences of maternal ECG in fetal ECG recordings, muscle artifacts in ECG leads, ECG disturbances in EEG signals, or the case of evoked potentials (EP) hidden by the normal brain activity. Exogenous effects are related to the interaction between the recording instrumentation, the patient, and the environment (disturbances from electrode motions, power line interference, and so on).

The recovery of a signal from background noise is known as “signal enhancement.” For this purpose, several signal-processing techniques have been developed. Two of them will be described in the sections “Digital Filters” and “Signal Averaging.” The former is appropriate when the frequency contents of signal and noise do not overlap; the latter is useful when a deterministic signal is buried in stochastic noise.

Digital Filter

A digital filter is a discrete-time system that transforms an input signal $x[n]$ into an output sequence $y[n]$ in which some desired properties of $x[n]$ are enhanced and emphasized. We may write $y[n] = T[x[n]]$, indicating that a particular transformation $T[\]$ is applied to $x[n]$ to produce $y[n]$. When $T[\]$ is time invariant and satisfies the linear superposition principle, the class of linear, time-invariant (LTI) filters is obtained. The LTI

filters have been widely studied and have important applications in BSP.^[3-5,7-11]

When the properties of a system have to be analyzed, we may observe the system output when known functions are at the input. To define the properties of LTI filters, two input functions are of particular interest: 1) the δ -Dirac function ($\delta[n]$) and 2) complex sinusoids.

The filter output, obtained when an impulse is at input, is known as the “impulse response” (IR). The filter outputs when sinusoids are the input define the “frequency response” of the filter.

The knowledge of the IR makes possible to compute the filter response to any input. To prove that, let us express a generic signal as the sum of weighted and delayed impulses

$$x[n] = \sum_{k=-\infty}^{\infty} x[k] \cdot \delta(n-k) \quad (5)$$

Each $\delta[n-k]$ will generate the same IR $h[n-k]$ (filter is time invariant), even if delayed. The responses can be summed up at the output, so that we obtain

$$y[n] = \sum_{k=-\infty}^{\infty} x[k] \cdot h(n-k) \quad (6)$$

or in the z -domain

$$Y(z) = H(z)X(z) \quad (7)$$

where $H(z)$ is the “transfer function” of the filter. Thus, in a LTI system, the output is computed by convolving the input with the IR. Two important filter properties can be deduced by the IR: stability (any bounded input will produce a bounded output) and causality (the filter will not generate any output before the input is applied). In fact, it can be demonstrated that a filter is stable only if

$$\sum_{k=-\infty}^{\infty} |h(k)| < \infty \quad (8)$$

and causal if and only if

$$h(k) = 0 \quad \text{for } k < 0 \quad (9)$$

Now, let us assume a complex sinusoid, $x[n] = e^{jwn}$, as the input, so that the filter output becomes

$$\begin{aligned} y[n] &= \sum_{k=-\infty}^{\infty} h[k] e^{jw(n-k)} \\ &= e^{jwn} \sum_{k=-\infty}^{\infty} h[k] e^{-jwk} = e^{jwn} H(w) \end{aligned} \quad (10)$$

Thus, the sinusoid in input is still the same at the output, but multiplied by a complex quantity, $H(w)$.

This quantity is known as the “frequency response” of the filter, and it is usually defined in terms of its modulus $|H(w)|$ and phase. Eq. (10) tells us that, according to its pulse frequency, a sinusoid as an input may be amplified [when $|H(w)| > 1$] or attenuated [when $|H(w)| < 1$] by the filter. Therefore, if the signal and noise are separated in the frequency domain, we may design a linear filter to enhance the signal contribution. The ideal filter should have $|H(w)|$ equal to 1 in the signal band and equal to 0 in the noise band.

A LTI filter can be expressed by the following difference equation

$$y[n] = \sum_{m=0}^M b_m x[n-m] - \sum_{k=1}^N a_k y[n-k] \quad (11)$$

where the a_k and b_m are the filter coefficients. If at least one a_k is different from zero, previous output values contribute to the current output. The filter is said to be “recursive.” Conversely, when the a_k values are all zeros, the filter is “non-recursive.”

Applying the z -transform at Eq. (11), after a few rearrangements, the “transfer function” becomes

$$\begin{aligned} H(z) &= \frac{Y(z)}{X(z)} = \frac{\sum_{m=0}^M b_m z^{-m}}{1 + \sum_{k=1}^N a_k z^{-k}} \\ &= b_0 z^{N-M} \frac{\prod_{m=1}^M (z - z_m)}{\prod_{k=1}^N (z - p_k)} \end{aligned} \quad (12)$$

where the z_m and the p_k are, respectively, the zeros and the poles of the transfer function. In terms of modulus and phase, we have

$$|H(e^{jw})| = |b_0| \frac{\prod_{m=1}^M |e^{jw} - z_m|}{\prod_{k=1}^N |e^{jw} - p_k|} \quad (13)$$

$$\angle H(e^{jw}) = \angle b_0 + \sum_{m=1}^M \angle (e^{jw} - z_m) - \sum_{k=1}^N \angle (e^{jw} - p_k) \quad (14)$$

The “pole-zero” form of $H(z)$ is of great help for analyzing the stability of the filter and for evaluating the frequency response characteristics. Stability is verified when all the poles lie inside the unit circle. In addition, the frequency response of a filter can be easily evaluated by noting that $|e^{jw} - z_m|$ is a vector in the z -plane joining the m th zero to the point of the unit circle identified by w . Thus, to obtain an estimate of $H(w)$, we move around the unit circle: When we are close to a zero, $|H(w)|$ will approach zero; conversely, when we are close to a pole, $|H(w)|$ will tend to a peak (the

closer the pole to the unit circle, the sharper the peak).^[9,11]

As previously said, a LTI filter may be completely described by its IR. For finite IR (FIR) filters, the IR is time limited (i.e., it is composed by a limited number of nonzero samples). In a FIR filter, all a_k 's are zeros, and the filter output is a combination of the input sequence only. The FIR filters are always stable (all poles are in the origin) and usually implemented in a nonrecursive form. In addition, the requirement of linear phase can be easily obtained by putting some constraints on the IR coefficients. These have to be symmetrical around their midpoint

$$b_m = \pm b_{M-m} \quad (15)$$

The sign + or - stays in accordance with the symmetry (even or odd) and M value (even or odd). On the contrary, in infinity IR (IIR) filters, the IR oscillates up to infinity with nonzeros values. To support a response of infinite length with a single impulse in input, the IIR filter must contain a feedback path, which feeds the filter when the input has vanished (i.e., at least one a_k must be different from zero). This type of filter is therefore recursive. Finally, in IIR filters, the linear phase response is not guaranteed, and particular attention has to be paid to filter stability.

A filter is designed to satisfy a set of requirements, usually in the frequency response. Those requirements depend on the application, but they usually involve the width of the pass-band (and stop-band), the maximum extent of the transition band, as well as the largest ripple amplitude allowed in the pass-band and stop-band. The aim is to design a filter that mimics the ideal filter response. Different techniques can be used. An IIR filter may be projected using the impulse invariance method, a bilinear transformation or Butterworth, Chebyshev, and elliptic approximations. Window techniques, equiripple approximation, or frequency-sampling methods can be applied for FIR filter design.^[3-5,9]

An example of filter application for the analysis of an EEG signal is depicted in Fig. 2. A filter bank is used to enhance the patterns of rhythmic activity in different frequency bands; from top to bottom: the delta (0–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), and beta (13–30 Hz) rhythms. Rhythm separation was obtained by using band-pass FIR filters. Fig. 2B shows the frequency and phase response of the band-pass filter used to enhance activity in the alpha band. The corresponding zero-pole diagram of the filter is shown in Fig. 2C. Quantitative analysis of the different rhythms of the EEG is important for detecting and quantifying a great number of processes and diseases involving disturbances in brain activity.^[12]

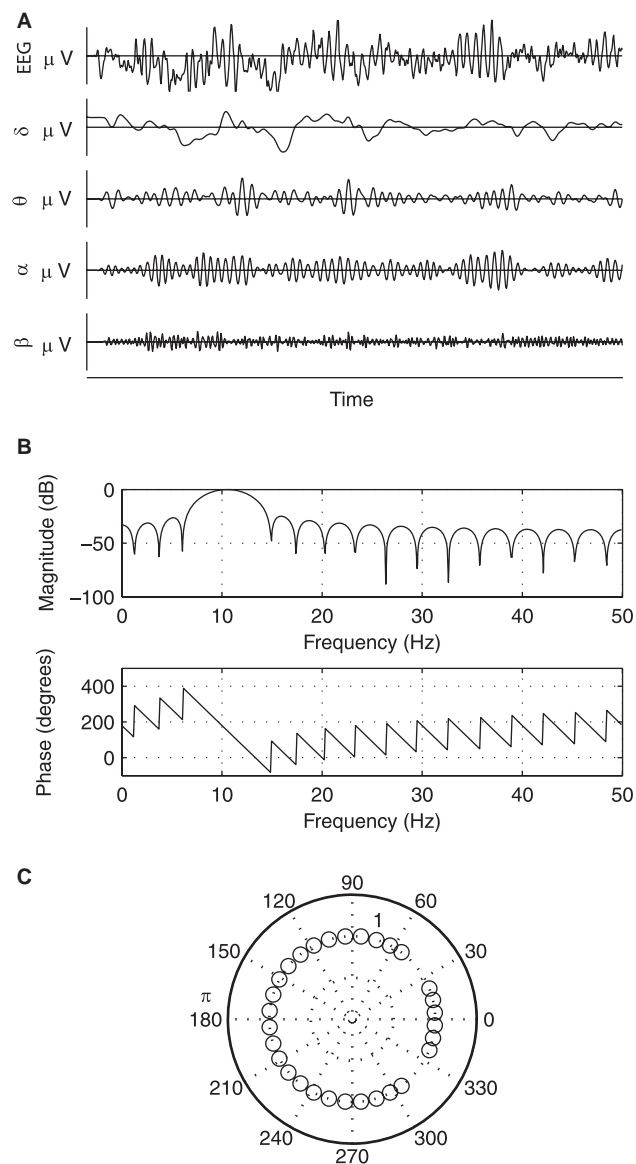


Fig. 2 Separation of EEG rhythms using a bank of band-pass FIR filters: (A) c3 channel of the EEG recorded for the scalp of a subject and its components in δ , θ , α , and β bands; (B) modulus and phase diagrams of the frequency response of the band-pass FIR filter (8–13 Hz) used to separate the α rhythm in the EEG; and (C) zero-poles diagram of the filter. Note the one-to-one correspondence between the position of the zeros on the unit circle and the damping in the modulus of the frequency response as well as the absence of zeros in the slice corresponding to the pulse frequency of the band-pass zone.

Signal Averaging

In many BSP applications, we are faced with the problem of detecting small signal components embedded in wide band noise. A typical example is the enhancement of the EP, i.e., brain potentials evoked by

sensory stimulations of different nature (visual, acoustic, and somatosensory). Usually of lower amplitude than the background EEG, EP (signal) and EEG (noise) have components in the same frequency range, and the problem to separate them cannot be approached by traditional filtering.^[9,12]

When the signal waveform repeats identically at known time instants (i.e., the evoked response occurring after each sensory stimulus) and the noise samples are not correlated (also to the signal), the “signal averaging” technique can be applied to separate signal components from background noise. Let us consider N repetitions (epochs) of a deterministic signal and indicate each repetition as $x_i(n)$ $\{i=1, 2, \dots, N\}$. The signal is embedded in noise $w_i(n)$. We assume that

1. $x_i(n)$ is repetitive and it does not vary epoch-by-epoch [$x_i(n) = x_j(n)$],
2. the noise is a broadband stationary process with zero mean and variance σ^2 , and
3. signal and noise are uncorrelated and additive.

$$y_i[n] = x_i[n] + w_i[n] \quad (16)$$

The averaging technique consists of summing up the signals of the different epochs (after having these properly aligned using efficient trigger points). We have

$$y_e[n] = \frac{1}{N} \sum_{i=1}^N y_i = x[n] + \frac{1}{N} \sum_{i=1}^N w_i[n] \quad (17)$$

where N is the number of epochs, $x[n]$ represents the deterministic signal to be determined, and the second term in Eq. (17) is an estimator of the mean of process $w_i[n]$. This term tends to the true mean of the process (zero, according to our assumptions) when N tends to infinity, while the variance of the estimate decreases with N being equal to σ^2/N . Therefore, after averaging, the original signal components are unaffected, while the variance of noise is decreased to σ^2/N . The net increment in terms of SNR depends on the number of repetitions included in the averaging process. The resulting SNR can be demonstrated as

$$\text{SNR}_e = \sqrt{N} \text{SNR}_i \quad (18)$$

Thus, the original SNR_i is improved by a factor \sqrt{N} . Finally, it may be useful, from a theoretical point of view, to interpret a signal averaging process as a digital filter. For each fixed time sample, the averaging process is analogous to a low-pass moving average

filter whose transfer function is

$$\begin{aligned} H(z) &= \frac{1 + z^{-1} + z^{-2} + \dots + z^{-(N-1)}}{N} \\ &= \frac{1 - z^{-N}}{1 - z} \end{aligned} \quad (19)$$

It is worth noting that this filter does not operate on the time sequence (n is fixed), but on the sequence of aligned samples taken from successive sweeps. The modulus of the frequency response of Eq. (19) presents a main lobe with the first zero located at around $f=1/N$ and secondary lobes spaced by a factor $1/N$ (we are assuming that the repetition frequency of the epoch is 1, without loss of generality). On increasing N , the width of the main (and the secondary) lobe is reduced as well as the amplitude of the secondary lobes. Therefore, as the noise is a broadband process, the larger the value of N , the bigger the noise attenuation. Conversely, the desired signal component (which is constant among sweeps) will not be attenuated by the filter. In fact, the amplitude of the main lobe (located at zero frequency) does not change with N .

An averaging technique is widely used for the analysis of biological signals. In addition to the above-mentioned EP application, other examples can be found in the detection of VLPs,^[6] in the separation between mother and fetal QRS complexes, as well as in the removal of ventricular activity from ECG signal during atrial fibrillation (AF).^[13] An application of this technique to the analysis of VLP is presented in Fig. 3. The VLPs are high-frequency micropotentials, which may occur at the end of QRS complex. The VLPs are signatures of abnormal conduction regions in the ventricles, associated with the substrate responsible for life-threatening arrhythmias. Their amplitude is lower

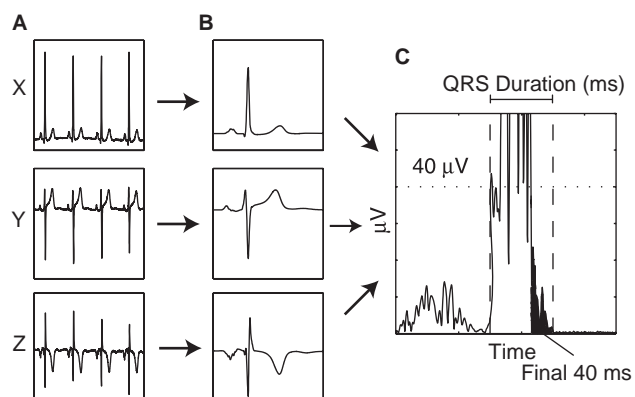


Fig. 3 Detection of the VLP. (A) Orthogonal X, Y, Z ECG leads; (B) averaged leads obtained by applying the signal averaging technique; and (C) plot of the magnitude vector.

than the QRS amplitude and other noise superimposed onto the ECG. To detect the VLP, orthogonal X, Y, Z leads are recorded (Fig. 3A). As the QRS complex and its VLPs should not vary beat-by-beat, a “signal averaging technique” may reduce the superimposed random noise and enhance the repetitive, low-amplitude VLPs. Normal QRS are detected, aligned (cross-correlation), and averaged to build noiseless templates (Fig. 3B). Then a bidirectional filter (high-pass >40 Hz, Butterworth filter) is applied at each averaged lead to eliminate the low frequencies contained in the QRS and ST segment and to enhance VLP. The filtered signals are combined into a vector magnitude, $\sqrt{X^2 + Y^2 + Z^2}$ (Fig. 3C). The black area represents the last 40 msec of the QRS complex, which is the key window for detecting the VLP.^[6]

SPECTRAL ANALYSIS

The power spectral density (PSD) of a given signal describes how the signal power is distributed in frequency. This representation of signal energy shows information that is not directly perceptible in the time domain. However, the PSD cannot be calculated in practice, as it requires infinite integration time. Only an estimate of the true PSD can be obtained by a number of approaches, basically divided into two classes: nonparametric and parametric.

Nonparametric Approaches

The traditional approach to PSD computation is based on FT. The easy handling and fast computation of FT-based estimators make the nonparametric approach widely used. The PSD may be obtained directly for the time series, using the periodogram approach:^[14,15]

$$P(f) = \frac{1}{NT_s} \left| \sum_{n=0}^{N-1} x[n] e^{-j2\pi f n T_s} \right|^2 \equiv \frac{1}{NT_s} X(f)^2 \quad (20)$$

where N represents the number of samples, T_s is the sampling period, and $X(f)$ corresponds to the FT of the time series $x[n]$. As an alternative, on the basis of the Wiener–Khinchin theorem, the PSD estimate can be obtained from the estimate of the autocorrelation function (ACF) of the signal. We have

$$\phi_{xx}[m] = \frac{1}{N} \sum_{n=0}^{N-|m|-1} x[n] x^*[n+m] \quad (21)$$

$$X(f) = T_s \sum_{m=-N+1}^{N-1} \phi_{xx}[m] e^{-j2\pi f m T_s} \quad (22)$$

where $\phi_{xx}[m]$ is the ACF, m represents the time-lag, and $*$ denotes the complex conjugate.

The FT is defined theoretically for an infinite sequence of data. In practice, only finite data sequences are available, and it is necessary to make some hypothesis, sometimes unrealistic, on the samples outside the observed window: they are commonly assumed to be zeros. This assumption is an implicit rectangular windowing of the data (which drastically truncates the sequence), and produces error or discrepancy in the estimated spectrum called “leakage error.” To handle this problem, different windows, which smoothly connect the side sample to zeros, are often used. The cost to pay is a reduction in the frequency resolution. In addition, the estimators, Eq. (20) and Eq. (22), are not statistically consistent (the variance of the estimate does not tend to zero, as the length of the data window goes to infinity). To improve the statistical performance, various methods have been developed, such as Bartlett, Welch, or Daniell methods.^[4,5,15,16] Even in this case, the application of these procedures is paid in terms of frequency resolution.

Parametric Approaches

Parametric approaches assume the signal to be the output of a mathematical model, and the PSD is estimated as a function of model parameters and its inputs. Special classes of models, driven by white noise and characterized by rational transfer functions, include the autoregressive model (AR), the moving average model (MA), and the autoregressive–moving average model (ARMA). These models are appropriate to fit many deterministic and stochastic stationary processes. They provide a general framework for the description of the input–output relationships of the analyzed process/system, independently from the physical, physiological, or anatomical characteristics of the system itself: it is the so-called black-box approach. An ARMA model is described by the following linear difference equation:

$$y[n] = - \sum_{k=1}^p a[k] y[n-k] + \sum_{m=1}^q b[m] x[n-m] + w[n] \quad (23)$$

where $y[n]$ is the output of the causal filter, which models the observed data, and $w[n]$ is a white noise input with zero average and variance λ^2 . The $a[k]$'s and $b[m]$'s are the model coefficients. As evidenced by Eq. (23), an ARMA process model is built

up by an AR model of order p and an MA model of order q . Because an AR model of appropriate order is able to represent any ARMA or MA processes of finite variance (Wold decomposition theorem) and the AR parameters are easier to estimate than the ARMA ones (estimation of AR parameters needs only solving a set of linear equations), the AR model is usually selected to fit the data. In this case, the AR PSD becomes

$$P(f) = \frac{T_s \lambda^2}{\left| 1 + \sum_{k=1}^p a[k] z^{-k} \right|_{z=e^{j2\pi f T_s}}^2} = \frac{T_s \lambda^2}{\left| \prod_{k=1}^p (1 - z_k z^{-1}) \right|_{z=e^{j2\pi f T_s}}^2} \quad (24)$$

which shows that the PSD is completely described by the properties of the noise feeding the model (i.e., its variance) and by the model parameters. In the right hand part of Eq. (24), we put into evidence the poles of the transfer function.

An attractive and special feature of parametric models is the possibility to decompose the spectrum into several components (i.e., bell-shaped curves in Fig. 4C) and to obtain a direct computation of power and frequency associated with each component. We firstly observe that ACF of the signal and its spectrum are related, in the z -domain, by the following expression:

$$R(k) = \frac{1}{2\pi j} \int_{|z|=1} P(z) z^{k-1} dz \quad (25)$$

If the residual integration method is used to compute the above integral, the ACF is decomposed into a sum of both damped sinusoids, originated by a pair of complex conjugated poles, and damped exponential functions related to real poles.^[17] The FT of the k th ACF component provides the pattern of the spectral component. The frequency and power of each component can be estimated from the poles: The argument of the pole gives the central frequency of the component, while the residual γ_k is used to compute the power [the component power is γ_k for real poles or $2\text{Re}(\gamma_k)$ for conjugate pole pairs]. The residuals γ_k are calculated as

$$\gamma_k = z(z - z_k)P(z) \Big|_{z=z_k} \quad (26)$$

If properly applied, either parametric or nonparametric approaches provide reliable estimates. However, it is worth making a comparative discussion of the two approaches. Advantages of the parametric

approach can be summarized as follows:

- It is statistically consistent on short segments of data, and, under general assumptions, the AR PSD is a maximum entropy spectrum (MES).
- Spectral resolution does not depend on the signal length.
- There is no need to window the data.
- A spectral decomposition method is available, which facilitates the postprocessing of the spectrum and the extraction of the relevant spectral parameters.

Conversely, an AR PSD estimate is more complex and more heavy from a computational point of view; it requires some a priori assumptions as to the kind of model (AR, ARMA, or others) and its complexity (i.e., the model order). Regarding the last point, it is worth mentioning that a battery of criteria does exist, which may help in the identification of the optimal model order. They include the Akaike, the final prediction error, and the minimum description length (MDL) criterion. In addition, some a posteriori tests could be applied to verify the assumptions; for example, a test of whiteness may be performed on the model errors to check if the model was able to capture all the information contained in the data.

As a final remark, the selection of the most appropriate PSD estimator depends on the particular application and the peculiar signal characteristics. One method may be suitable for processes with sharp spectral peaks, while others will perform better with broad, smoothed spectra. The use of a priori knowledge should guide the choice of the proper method.

Fig. 4A shows the R–R interval series (tachogram) of a normal subject. The R–R intervals are not constant, but oscillate under the control of the autonomic nervous system (ANS) controlling heart rate. The power spectrum of this signal evidences two main rhythms (Fig. 4B and C) related to the activity of sympathetic and parasympathetic branches of ANS. A high-frequency (HF) component, (in a range between 0.15 and 0.45 Hz), which is synchronous with respiration and modulated by vagal tone. Conversely, a low-frequency (LF) component (centered around 0.1 Hz) is mainly modulated by sympathetic activity. Thus the PSD of the RR series is a noninvasive tool for assessing ANS status.^[18] In Figs. 4B and C, the PSD is computed using a nonparametric approach (Welch method and Hamming window) and by parametric method (AR model of order 12), respectively. The parametric estimate presents a higher frequency resolution and a lower spectral leakage than those obtained by nonparametric method.

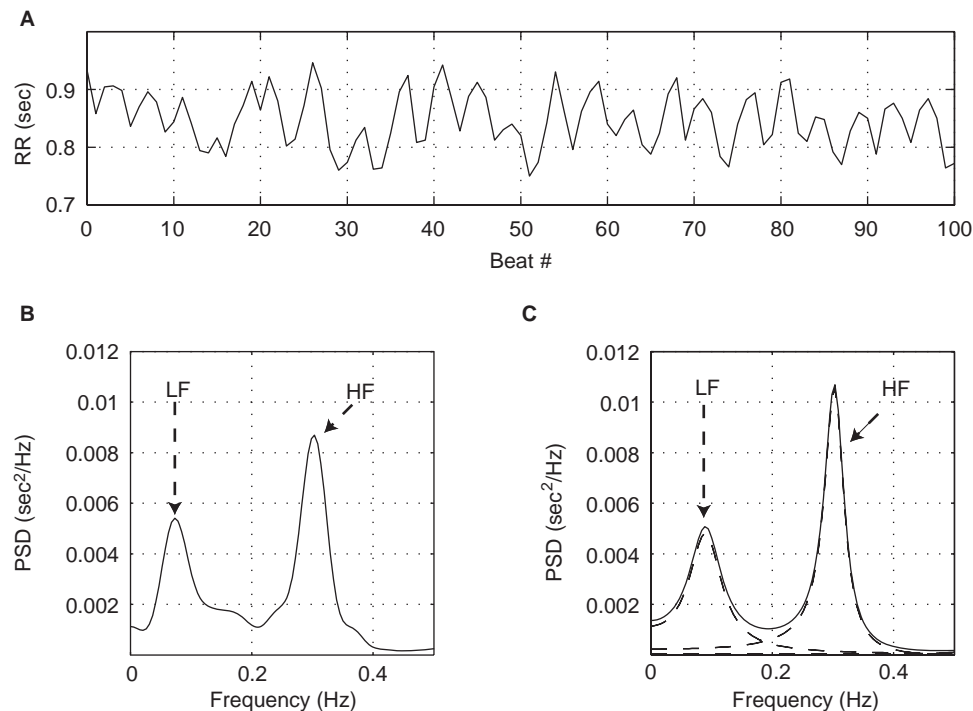


Fig. 4 Example of the spectral analysis of R–R intervals series: (A) the tachogram; PSD obtained; (B) by nonparametric approach; and (C) by parametric approach based on AR modeling. Spectral decomposition is also plotted.

The capability to separate signal power in the frequency domain makes spectral estimation widely used for BSP. The analysis of EEG rhythms (during sleep, anesthesia, or epilepsy), the characterization of heart sounds,^[8] and the quantitative analysis of the free induction decay signal in magnetic resonance spectroscopy (for metabolite quantification)^[19] are only a few representative examples.

EVENT DETECTION

Biomedical signals are characterized by peculiar waveforms, markers of physiopathological events. Detection and classification of such waveforms are of primary importance for a better understanding of physiological processes and to define the degree of pathology in routine clinical practice. In this regard, different techniques and procedures have been developed, which use information on the event (such as morphology, amplitude, or frequency content) to locate it in the signal.^[8]

A well-known example of event detection in BSP is the QRS detector.^[7,8] Location of QRS complexes is usually the fundamental step for subsequent analysis of an ECG signal, such as heart rate computation, beat classification, VLP extraction, and heart rate variability analysis. Detection may be obtained by

observing that QRS complexes have the largest slope with respect to P and T waves. To enhance this, a derivative (or second derivative) filter may be applied. Owing to the inner noisy nature of a derivative filter (which enhances HF components of the signal), the output of derivative steps is usually smoothed (MA filters) prior to other computations. A very popular QRS detection algorithm, reported in several textbooks,^[7,8] is the Pan-Tompkins algorithm. It consists of a few signal-processing steps: 1) a band-pass filter is used to isolate the QRS frequencies (removing the lower frequency related to baseline wandering and higher frequency containing muscle noise); 2) a derivative filter is used to enhance QRS waves; 3) the result is squared to account for different ECG polarity; and 4) the output is low pass filtered. The resulting signal is composed by a set of local maxima, which identify the location of the QRS complex (Fig. 5).

Sometimes the morphology of a certain event is available. In this case, a special filter can be designed to “match” the characteristic of the event. Let us suppose that the reference waveform $x[n]$ is merged in some noise $n[t]$. The resulting signal is passed through a LTI filter whose IR is indicated by $h[t]$. The output signal will be $y[n] = (x[n] + n[n]) * h[t] = x_0[n] + n[n] * h[t]$. Now, among all the possible choice of $h[n]$, we will select the one which maximizes at a certain time instant t_0 , the SNR at the filter output, i.e., the SNR between

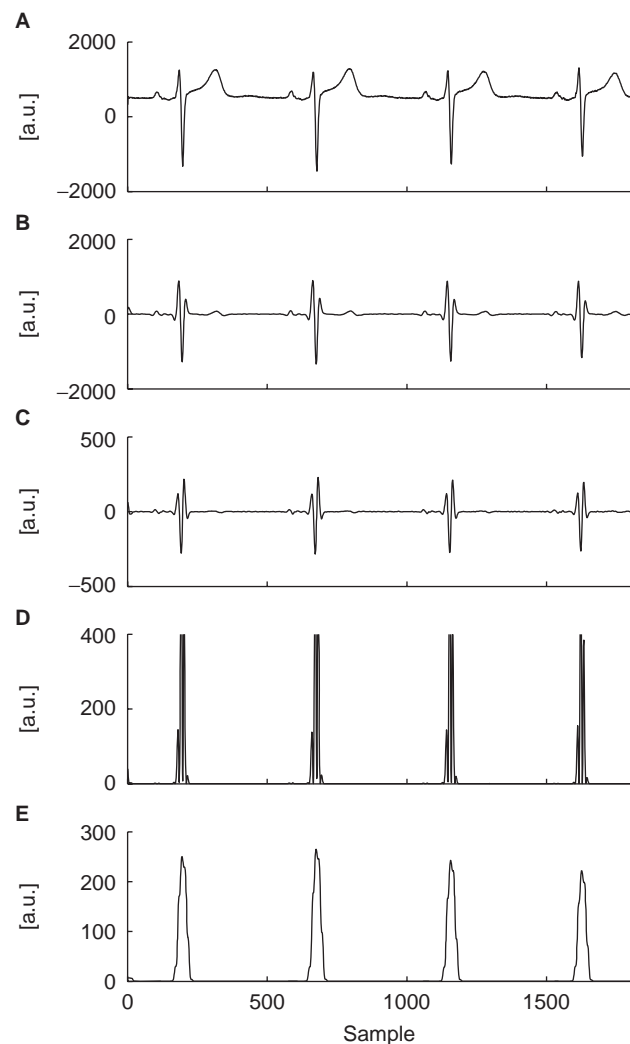


Fig. 5 Processing steps in a QRS detector algorithm: (A) Raw ECG signal; (B) output of the band-pass filter; (C) differentiation; (D) squaring process; and (E) final output after the moving average filter. The signals are realigned, taking into account the filter's delay.

$x_0[n]$ (signal) and $n[n] * h[t]$ (noise). It can be demonstrated^[16] that, when the noise is a white process, the optimal IR is

$$h[n] = \frac{1}{\sigma_n^2} x[t_0 - t] \quad (27)$$

where σ_n^2 is the noise power. Thus, the IR is simply a reversed version of the reference waveforms, scaled and shifted. The resulting filter is known as the “matched filter” (MF). The MF has been applied to detect particular patterns in EEG signals such as spike-and-wave patterns and cyclic alternating pattern in sleep, or to detect components of event-related brain potentials.^[20]

In other situations, the template is unknown, but the occurrence of a similar event is searched in two signals. The “cross-correlation function” (CCF) provides a second order statistical comparison of two, zero-mean sequences as a function of the time shift between them. The CCF is evaluated as

$$\phi_{x,y}(\tau) = \frac{1}{N-1} \sum_{n=1}^N x[n]y[n-\tau] \quad (28)$$

where N is the sequence length, and τ is the lag index ($\tau=0, 1, 2, \dots, N$). See the analogy with Eq. (21). A normalized version of CCF (NCCF) is obtained by dividing it by the product of the standard deviation of the two signals

$$\gamma_{x,y}(\tau) = \frac{\sum_{n=1}^N x[n]y[n-\tau]}{\sqrt{\sum_{n=1}^N x[n]^2 \sum_{n=1}^N y[n]^2}} \quad (29)$$

The NCCF ranges from 0 to 1 and reaches its maximum when similar patterns are present in the two signals. In addition, the lag of the maximum indicates the time delay at which the event occurs in the two signals. The CCF is used to align, temporally, two waveforms (see VLP example), to locate events with known templates (P-wave detection), or to detect common patterns in two signals. For example, NCCF can be used to evaluate the degree of organization of atrial activity during AF.

The frequency-domain counterpart of CCF is the “cross-spectral density” or simply “cross-spectrum” (P_{xy}). It provides both amplitude and phase information about the frequencies held in common between two sequences. Formally, the cross-spectrum and CCF are related as a FT pair. Thus

$$P_{xy}(f) = FT(\phi_{xy}[\tau]) = X(f)Y^*(f) \quad (30)$$

and the squared coherence spectrum (SCS) is defined as

$$\gamma_{xy}(f) = \frac{|P_{xy}(f)|^2}{P_{xx}(f)P_{yy}(f)} \quad (31)$$

where $P_{xx}(f)$ and $P_{yy}(f)$ are the auto-spectrums of the sequences.

A value of SCS close to 1 indicates that same rhythms in two signals have a constant phase relationship, thus they may have a common origin or they can be linearly synchronized. The SCS has been applied to detect the occurrence and the extent of synchronized activity in multisite recordings of neuronal activity (during epilepsy, sleep stages, or during preparation and execution of hand movements) and cardiac electrical activity (during different atrial rhythms). In



addition, it has been used to assess the coupling between cardiovascular variability series.^[18]

Finally, it is worth mentioning that some biomedical signals, such as the electromyogram and phonocardiogram (PCG), do not present clear fingerprints. Nevertheless, these signals are affected by changes in their level of activity or in their energy, which may carry useful information. In these cases, very simple metrics can be used to detect these changes. A useful technique is the “zero-crossing rate” (ZCR). The ZCR measures the complexity of a signal in relation to the number of times it crosses the zero activity line or another reference line. As the ZCR is affected by the offset and LF artifacts, it is advisable first to obtain the derivative of the signal. A typical example of ZCR application is the detection of speech and silence epochs in the voice signal.

The “root mean-square value” (RMS) of a signal gives a global indication of its average power. If RMS is evaluated over windows of M samples, it is possible to detect alterations or variations in a process. This technique was used to identify systolic and diastolic segments in the PCG,^[8] to quantify EMG activity, or to evaluate the energy content of EEG in different bands.

CONCLUSIONS

The fundamental and widely used signal-processing techniques for biomedical signal analyses have been introduced in this article. These classical methods are intended as the basic, powerful tools for enhancing important information from biological signals and as fundamental bricks for building more sophisticated algorithms. A few examples of applications, including feature extraction, noise reduction, event detection, and spectral analysis, have been presented. These examples were selected to demonstrate how traditional techniques of digital signal processing may impact on the problem of extracting informative parameters from biological signals.

ARTICLE OF FURTHER INTEREST

Biopotential Amplifiers

REFERENCES

1. Webster, J. *Medical Instrumentation: Application and Design*, 3rd Ed.; Wiley: New York, NY, U.S.A., 1998.
2. Cobbold, R. *Transducers for Biomedical Measurements*; Wiley & Sons: New York, U.S.A., 1998.
3. Lynn, P.; Fouerst, W. *Introductory Digital Signal Processing with Computer Application*, 2nd Ed.; Wiley & Sons: Chichester, England, 1997.
4. Oppenheim, A.; Schafer, R. *Discrete-Time Signal Processing*, 2nd Ed.; Oppenheim, A., Ed.; Prentice-Hall: New Jersey, U.S.A., 1999.
5. Haddad, R.; Parsons, T. *Digital Signal processing: Theory, Applications and Hardware*; Computer Science Press: New York, U.S.A., 1991.
6. Simson, M. Use of signals in the terminal QRS complex to identify patients with ventricular tachycardia after myocardial infarction. *Circulation* **1981**, *64* (2), 235–241.
7. Tompkins, W. *Biomedical Digital Signal Processing*; Tompkins, W., Ed.; Prentice-Hall: New Jersey, U.S.A., 1993.
8. Rangayyan, R. *Biomedical Signal Analysis: A Case-Study Approach*; Metin, A., Ed.; IEEE press Wiley & Sons: New Jersey, U.S.A., 2002.
9. Challis, R.; Kitney, R. The design of digital filters for biomedical signals processing. *J. Biomed. Eng.* **1982**, *4*, 267–278.
10. Bianchi, A.; Mainardi, L.; Cerutti, S. Signal processing. In *Noninvasive Electrocardiology: Clinical Aspects of Holter Monitoring*; Moss, A., Stern, S., Eds.; W. B. Saunders Company Ltd: London, England, 1996; 11–35.
11. Mainardi, L.; Bianchi, A.; Cerutti, S. Digital biomedical signal acquisition and processing. In *The Biomedical Engineering Handbook*, 2nd Ed.; Bronzino, J., Ed.; CRC Press and IEEE Press: Florida, U.S.A., 2000; Vol. 53, 1–25.
12. Gevins, A.S.; Remond, A. *Handbook of Electro-physiology and Clinical Neurophysiology*; Elsevier: Amsterdam, 1987.
13. Holm, M.; Pehrson, S.; Ingemansson, M.; Sornmo, L.; Johansson, R.; Sandhall, L.; Sunemark, M.; Smideberg, B.; Olsson, C.; Olsson, S.B. Non-invasive assessment of atrial cycle length during atrial fibrillation in man: introducing, validating and illustrating a new ECG method. *Cardiovasc. Res.* **1998**, *38*, 69–81.
14. Kay, S. *Modern Spectral Estimation: Theory and Application*; Signal Processing Series; Oppenheim, A., Ed.; Prentice-Hall: New Jersey, U.S.A., 1988.
15. Marple, S. *Digital Spectral Analysis with Applications*; Oppenheim, A., Ed.; Signal Processing Series; Prentice-Hall: New Jersey, U.S.A., 1987.
16. Cohen, A. *Biomedical Signal Processing*; CRC Press: Boca Raton, FL, 1986.
17. Zetterberg, L. Estimation of parameters for a linear difference equation with application to EEG analysis. *Mathematical Biosci.* **1969**, *5*, 227–275.
18. Malliani, A. *Principles of Cardiovascular Neural Regulation in Health and Disease*; Kluwer Academic Publishers: Boston/Dordrecht/London, 2000.
19. Vanhamme, L.; Sundin, T.; Ven Hecke, P.; Van Huffel, S. MR spectroscopy quantitation: a review of time-domain methods. *NMR Biomed.* **2001**, *14*, 233–246.
20. Serby, H.; Yom-Tov, E.; Inbar, G.F. An improved P300-based brain-computer interface. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2005**, *13*, 89–98.

Digital Biomedical Signal Processing: Advanced Topics

Luca T. Mainardi

Department of Biomedical Engineering, Polytechnic University, Milan, Italy

INTRODUCTION

Biomedical Signal Processing (BSP) deals with extraction and processing of information contained in vital signals recorded from living systems. In the pioneer age, BSP was restricted merely to application of traditional signal processing algorithms for signal-to-noise ratio enhancement and spectral analysis. A review of these basic approaches is provided in another entry of this encyclopedia.^[1] A major breakthrough occurred in the past two decades. Since the 1990s, we have witnessed the explosion of new theoretical tools for signal analysis in an attempt to overcome the limitations of former approaches and to exploit signal features in nonstationarity and nonlinearity contexts. New methodologies became available, ranging from time-scale to time-frequency methods,^[2] higher-order statistics and independent component analysis,^[3] up to nonlinear dynamic approaches.^[4] These tools found their natural application to BSP due to the inner nature of biomedical signals and systems, which are characterized by nonstationarity and may exhibit nonlinear behavior.

The aim of this entry is to provide an introductory review of these modern techniques, putting some emphasis on both methodological aspects and potential application to BSP. The entry is organized as follows: “Wavelet Analysis” introduces wavelet transform as a new tool for signal decomposition and synthesis, and “Spectral Analysis of Nonstationary Signals” describes time-frequency and time-variant approaches as an extension of traditional spectral analysis to nonstationary epochs. The successive section addresses basic concepts of higher-order spectra. Finally, the section “Nonlinear Biomedical Signal Processing” is dedicated to a short overview of popular nonlinear methods. Each section starts by describing basic methodological concepts and ends with explicative examples of possible applications to BSP.

WAVELET ANALYSIS

Signal decomposition and reconstruction have been historically approached using Fourier analysis, which has become ubiquitous for both the analytic and numerical solution of differential equations and for the analysis and treatment of communication signals.

A signal is seen as the sum (possibly infinite) of complex exponentials, and the well-known Fourier Transform (FT) may be written as

$$X(\omega) = \int x(t)e^{-j\omega t} dt = \langle x(t), e^{-j\omega t} \rangle \quad (1)$$

where we put into evidence that FT is obtained as inner product $\langle \cdot \rangle$ between the signal and sinewaves functions of infinite duration. The signal can be reconstructed by the inverse FT

$$x(t) = \int X(\omega)e^{j\omega t} d\omega \quad (2)$$

FT is appropriate when dealing with stationary signals composed by few sinewaves components. However, it is a nonparsimonious transform when signals are nonstationary or characterized by abrupt changes. In the latter case, frequency contents are spread out over the frequency axis, and an infinite number of sinusoidal components is needed. The reason is that complex sinusoids have frequency localization but no time localization. This means that although we may be able to determine all the frequencies present in a signal, we cannot define when they occur.

Theory

To circumvent the previous limitation, we introduce the Continuous Wavelet Transform (CWT)^[5]

$$\begin{aligned} \gamma(\tau, a) &= \int x(t)\psi_{\tau,a}^*(t)dt = \langle x(t), \psi_{\tau,a}^*(t) \rangle \\ &= x(t)^* \psi_{\tau,a}^*(t) \\ \tau, a &\in \Re; \quad a \neq 0 \quad x, \psi \in L^2(\Re) \end{aligned} \quad (3)$$

which measures the similarity between the signal $x(t)$ and a family of functions called wavelets:

$$\psi_{\tau,a}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-\tau}{a}\right) \quad (4)$$

The wavelets are generated from a prototype $\psi(t)$, the mother-wavelet, by scaling (contracting or dilating) and translation. The parameter τ determines the time

shift, while a is the scaling factor that controls the dilation of the wave (coefficient $1/\sqrt{a}$ is used for energy normalization). We would expect that, using an appropriate selection of $\psi(t)$, any general signal can be represented by its decomposition into wavelets, and the signal synthesis may be obtained by summing projections of signal onto wavelets. In particular, if the admissibility condition is satisfied,

$$C_\psi = \int \frac{|\Psi(\omega)|^2}{|\omega|} d\omega < +\infty \quad (5)$$

where $\Psi(\omega)$ if the FT of $\psi(t)$, $x(t)$ can be reconstructed by

$$x(t) = C_\psi^{-1} \int_a \frac{da}{a^2} \int_\tau \gamma(\tau, a) \psi_{\tau, a}(t) d\tau \quad (6)$$

without losing any information.^[6] In Eq. 6, $\gamma(\tau, a)$ represents the weight of every wavelet in the synthesis of $x(t)$. The admissibility condition implies

$$|\psi(\omega)|_{\omega=0}^2 = 0 \quad (7)$$

and for practical application,

$$\int |\psi(t)|^2 dt < +\infty \quad (8) \quad \psi_{j,k}(t) = a_0^{-j/2} \psi(a_0^{-j} t - kT) \quad (9)$$

Therefore, continuous wavelets $\psi(t)$ have zero-mean and finite energy. In practice, they are oscillating around the t axis and vanishing with time (see Fig. 1a). In general, as $|\Psi(\omega)| \rightarrow 0$ when $\omega \rightarrow \infty$ wavelets have a band pass-like spectrum. Other important properties include regularity, null vanishing moments, orthogonality, and finite support.^[5]

Some analogies between FT and CWT exist: they are linear operators, and both transforms can be viewed as a projection of a function into a different domain. For the FT, this new domain contains basis functions that are sines and cosines. For the wavelet transform, this new domain contains more complicated basis functions called wavelets. The main difference lies in the fact that wavelets are usually localized in time, whereas complex exponentials are not (see Fig. 1a).

The CWT is highly redundant, being obtained by continuously translated and dilated functions. Reduction of redundancy may be obtained by discretization of the time-scale parameters ($a-\tau$) according to some sampling grid.^[5] In general $a = a_0^j$, and the translation is discretized with respect to each scale by using $\tau = ka_0^j T$, where j and k are integers, and T is the sampling interval. In this case, the wavelet basis functions are obtained by

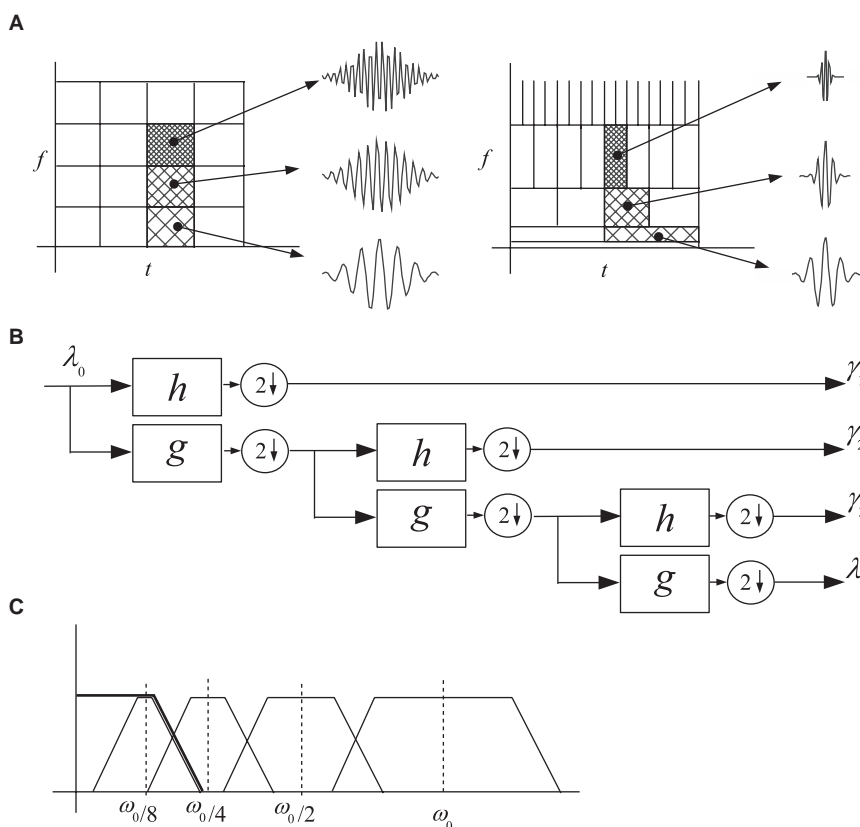


Fig. 1 (A) Partitioning of the time–frequency (t – f) plane using STFT (left) and WT (right). The corresponding basis functions are also reported. Note that time localization of wavelets varies in the t – f plane, preferring time localization for higher-frequency components and frequency localization for lower-frequency ones. (B) Calculation of wavelets coefficients, using an iterative filter bank structure. (C) Partitioning of frequency axis, using wavelets and scaling functions. The scaling function substitutes a infinite numbers of wavelets, and at a certain level (J), it covers the frequency spectrum uncovered by wavelets details at that scale.

Discretization acts like a microscope in which the term a_0^{-j} controls the magnification and $a_0^{-j}T$ controls the movement steps. If one is interested in small details, magnification will be large (j is negative and large), and coherently $a_0^j T$ describes small steps to center the details.^[5] In the most common formulation, $a_0=2$ and $T=1$ are used.

The Discrete Wavelet Transform (DWT) may now be introduced, and the resulting wavelet coefficients are

$$c_{j,k} = \int x(t)\psi_{j,k}^*(t)dt = \langle x(t), \psi_{j,k}^*(t) \rangle \quad (10)$$

Obviously, discretization of time-scale parameters makes sense if we are still able to reconstruct the original signal using the wavelets coefficients according to

$$x(t) = K \sum_j \sum_k c_{j,k} \psi_{j,k}^*(t) \quad (11)$$

where K is a constant playing the same role as C_ψ^{-1} in Eq. 6. In contrast with the nonrestrictive constraints imposed on wavelets in the continuous case (see Eq. 5), only very special choices of $\psi_{j,k}$ are now allowed. In general, the larger the redundancy, the milder the restrictions on the basis function.^[5]

Redundancy may be also reduced by designing the wavelets to be orthogonal. In mathematical terms, the family of wavelets can be selected in such a way that

$$\begin{aligned} \langle \psi_{j,k}, \psi_{j',k'} \rangle &= 1 \quad \text{if } k = k', j = j' \\ \langle \psi_{j,k}, \psi_{j',k'} \rangle &= 0 \quad \text{otherwise} \end{aligned} \quad (12)$$

Eq. 12 means that each wavelet coefficient $\psi_{j,k}$ stores information that is not correlated, and all the wavelets are necessary to reconstruct a general signal. They form an orthonormal basis.

Computation of wavelets coefficients can be obtained by using an iterative filter bank structure.^[6] To explain it, the scaling function $\varphi(t)$ has to be introduced. It is the companion of $\psi(t)$ and the two functions are linked through the following relations^[6]:

$$\begin{aligned} \varphi(t) &= \sqrt{2} \sum_{k=0}^{L-1} h_k \varphi(2t - k) \\ \psi(t) &= \sqrt{2} \sum_{k=0}^{L-1} g_k \varphi(2t - k) \end{aligned} \quad (13)$$

where

$$g_k = (-1)^k h_{L-k-1} \quad (14)$$

Different from the wavelet function, the scaling function has a low-pass properties

$$\int \varphi(t)dt = 1 \quad (15)$$

Because our signal $x(t)$ could be expressed in terms of dilated and translated wavelets (Eq. 11), and taking into account Eq. 13, it turns out that $x(t)$ can also be expressed in terms of scaling functions. In particular, at scale $J=0$, we have

$$x(t) = \sum_k \lambda_k \varphi(t - k) \quad (16)$$

where λ_k 's are the scaling coefficients. Let us suppose to proceed one step head along scale ($j=-1$). At this new scale, the scaling function can be still used to approximate the signal, but we have to add wavelets to preserve the same level of detail:

$$x(t) = \sum_k \lambda_{1,k} \varphi(t - 2^1 k) + \sum_k \gamma_{1,k} \psi(t - 2^1 k) \quad (17)$$

Iterating the procedure through the scales

$$\begin{aligned} x(t) &= \sum_k \lambda_{1,k} \varphi(t - 2^1 k) + \sum_k \gamma_{1,k} \psi(t - 2^1 k) x(t) \\ x(t) &= \sum_k \lambda_{2,k} \varphi(t - 2^2 k) + \sum_k \gamma_{2,k} \psi(t - 2^2 k) \\ &\quad + \sum_k \gamma_{1,k} \psi(t - 2^1 k) \\ &\quad \vdots \end{aligned} \quad (18)$$

we finally obtain

$$x(t) = \sum_k \lambda_{J,k} \varphi(t - 2^J k) + \sum_{j=1}^J \sum_k \gamma_{j,k} \psi(t - 2^j k) \quad (19)$$

which is the multiresolution decomposition of $x(t)$ obtained through wavelets.^[6] In addition, as $\psi(t)$ and $\varphi(t)$ are orthonormal bases, wavelet and scaling coefficients are found through inner product (Eq. 3). Substituting Eq. 13 in Eq. 3, after a few rearrangements we obtain the fundamental relationships

$$\begin{aligned} \lambda_{j-1,k} &= \sum_m h(m - 2k) \lambda_{j,m} \\ \gamma_{j-1,k} &= \sum_m g(m - 2k) \lambda_{j,m} \end{aligned} \quad (20)$$

which shows that wavelets and scaling coefficients at each scale can be computed from those of the previous scales. Eq. 20 describes a filter bank structure whose scheme is reported in Fig. 1b.

Applications

Wavelet analysis found wide fields of application in the analysis of biological signals.^[7] Three main objectives

have been pursued: (1) wave detection, (2) data compression, and (3) time–frequency representation of the signal. In addition, statistical properties of wavelets coefficients can be used to evaluate fractal-like properties and scaling exponents.^[8]

Among the earlier applications of Wavelet Transform (WT) to BSP, we found the analysis of heart sounds and murmurs^[7] generated by the beating heart and the resultant flow of blood through valves. WT was documented to characterize primary heart sounds, resolving some frequency components to which other methods were blind. WT was also used to investigate turbulent sounds and to provide a simple, noninvasive approach for detection of coronary-artery disease or stenosis.

In the analysis of electrocardiograms (ECG),^[9] the use of WT has been proposed both for the detection of main signal waves (QRS, T, and P)^[10,11] and for feature extraction and classification of cardiac patterns.^[12] WT seems particularly suitable for this task because ECG waves are transient events well localized in time (i.e., they resemble wavelets). Wave detection can be performed searching for singularity and discontinuity in the ECG. In this regard, the modulus maxima criterion has largely been employed to locate ECG waves and their fiducial points (onset, offset, and deflections). The idea behind the criterion was first elucidated by Mallat and Zhong.^[13] Let's consider a wavelet obtained as the first derivative of a smoothing (scaling) function:

$$\psi_a(t) = \frac{\partial \theta_a(t)}{\partial t} \quad (21)$$

The wavelets transform of $x(t)$ becomes

$$\gamma(\tau, a) = x(t) * \psi_a(t) = \tau \frac{d}{dt}(x(t) * \theta_a(t)) \quad (22)$$

showing that the maximum of $|\gamma(a, \tau)|$ is located on the sharp variation points (positive or negative) of a smoothed version of the signal. In the presence of important discontinuity, the term $x(t) * \varphi_\tau(t)$ will not vanish among scales (as it does in the presence of high-frequency amplitude fluctuations). Thus, local maxima in $|\gamma(a, \tau)|$ that spread among scales are the hallmark of the discontinuity. An example is presented in Fig. 2, where an ECG signal during normal sinus rhythm is analyzed.

WT was also employed to detect Ventricular Late Potentials (VLP),^[14] which are transient, low-amplitude, high-frequency waves occurring at the end part of QRS complex and signature of abnormal conduction regions in the ventricles. With respect to other time–frequency methods (described in the following paragraph), WT was suggested to have superior accuracy in detection of VLPs.^[15] Finally, ECG compression for storing and transmission has largely been performed using WT.^[16,17] WT is particularly appealing for this task because signals become sparser when projected in the wavelets domain than in the Fourier domain (in other words, significant information is contained in only a few of the wavelet coefficients).

In electroencephalography (EEG),^[7] use of WT has been proposed for seizure detection^[18] and for locating transient waveforms (spikes and sharp waves). Wavelet

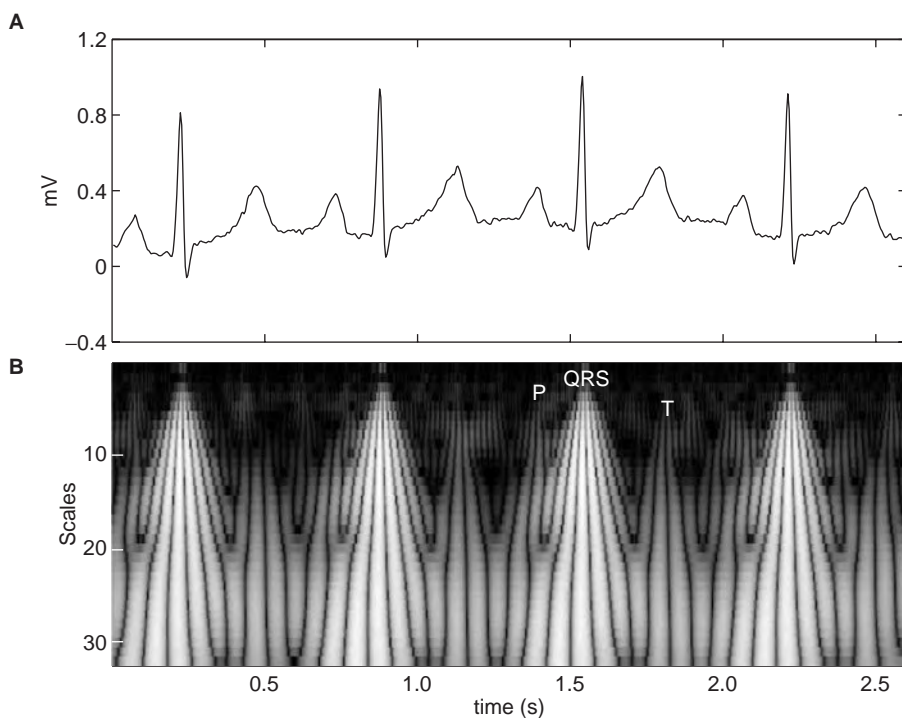


Fig. 2 (A) ECG signal during normal sinus rhythm. (B) Absolute value of CWT using Morlet wavelet. ECG waves (P–QRS–T) are indicated on the CWT. These waves can be detected by searching for CWT maxima spreading over scales.

has also proved to be effective in enhancing Evoked Potentials (EP) from background EEG activity.^[19,20] This can be pursued either through the selection of significant wavelets coefficients^[20] or through the improvement of the convergence rate of synchronous grand-averaging techniques.^[21]

SPECTRAL ANALYSIS OF NONSTATIONARY SIGNALS

Traditional power spectral analysis (employing either parametric or nonparametric approaches) implies the signal to be stationary, at least in the observation windows. Biomedical signals are intrinsically nonstationary, however, and characterized by transient waveforms of various natures and patterns. As a result, the spectral features of these signals evolve with time, and the capability to track these changes is relevant for the characterization of the dynamics of biological systems during both physiological and pathological conditions.

In this regard, signal processing tools are needed to describe how signal power is simultaneously distributed in the time-and-frequency domain. We look for a joint time–frequency (t – f) representation satisfying the marginal properties^[2]

$$\begin{aligned} \int P(t, \omega) dt &= |X(\omega)|^2 \\ \int P(t, \omega) d\omega &= |x(t)|^2 \end{aligned} \quad (23)$$

and where

$$\begin{aligned} E_x &= \int |x(t)|^2 dt = \int |X(\omega)|^2 d\omega \\ &= \iint |P(t, \omega)|^2 dt \quad d\omega \end{aligned} \quad (24)$$

is the signal energy. The most intuitive way to build such distribution is obtained by the Short Time Fourier Transform (STFT)

$$P_{\text{STFT}}(t, \omega) = \left| \int x(\tau) g(t - \tau) e^{-j\omega\tau} d\tau \right|^2 \quad (25)$$

where $g(t)$ is a time-window shifting along the signal. The STFT divides a signal into segments and takes the FT of each resulting section. Selection of the window defines the property of the distribution, and in general, a trade-off exists between time and frequency resolution. In fact, to locate short-time, high-frequency behavior temporarily, narrow time windows are needed. Conversely, to resolve the low-frequency behavior, a wider time window must be employed. This results

from a more general rule that prohibits simultaneously determining time localization, Δt , and frequency resolution, $\Delta\omega$, at arbitrary precision. A lower bound is defined by the Heisenberg inequality^[2]:

$$\Delta\omega \Delta t \geq \frac{1}{2} \quad (26)$$

Another time–frequency representation (TFR) can be obtained by WT: it can be shown that WT preserves energy. We have

$$E_x = C_\psi^{-1} \int_a \frac{da}{a^2} \int_\tau |\gamma(\tau, a)|^2 dt \quad (27)$$

evidencing that the squared modulus of wavelet coefficients can be interpreted in terms of energy t – f distribution. In this case, the parameter a can be used to tune different resolutions by using a larger window for low-frequency components (in favor of frequency resolution) and shorter windows for higher frequency (best location of high-frequency details is pursued) without violating the constraint of Eq. 26.

STFT and WT are also known as linear t – f distributions. In the following, some examples of quadratic t – f distribution will be revised, and finally, the parametric approach to time-variant spectral estimation will be presented.

Quadratic Time–Frequency Representations

An energy distribution can be obtained^[2] from the marginals by inverting Eq. 23

$$P(t, \omega) = \frac{\partial}{\partial t} |X_t(\omega)|^2 \quad (28)$$

where

$$|X_t(\omega)|^2 = \int E[x(t)x^*(t - \tau)] e^{-j\omega\tau} dt \quad (29)$$

is the spectrum of $x(t)$ and where the subscript t means that this spectrum evolves with time. Substituting Eq. 29 into Eq. 28 leads to the Wigner-Ville distribution (WD)

$$\begin{aligned} P_{\text{WD}}(t, \omega) &= \int x\left(t - \frac{\tau}{2}\right) x^*\left(t + \frac{\tau}{2}\right) e^{-j\omega\tau} d\tau \\ &= \int R(t, \tau) e^{-j\omega\tau} d\tau \end{aligned} \quad (30)$$

where the quantity $R(t, \tau) = x(t - \tau/2)x^*(t + \tau/2)$ is known as an instantaneous autocorrelation function (ACF). WD satisfies several desirable properties for a quadratic TFR. It is real-valued; it satisfies time and

frequency marginals; and it preserves time- and frequency-shifts, the instantaneous frequency, and groups delay properties.^[22] Finally, for linear-chirp (i.e., sinusoids whose instantaneous frequency varies linearly with time $f_x(t) = f_0 + \alpha t$), we have

$$P_{\text{WD}}(t, \omega) = \delta(t, t - f_x(t)) \quad (31)$$

Thus, the WD is a line on the t - f plane located on the instantaneous frequency of the signal. The WD, however, is affected by interference terms (IT) when more than one signal component is present. In fact, as any quadratic TRF, WD satisfies the quadratic superposition principle,^[23] which states that for a N -component signal

$$x(t) = \sum_{k=1}^N c_k x_k(t) \quad (32)$$

the corresponding quadratic TRF is

$$P_x(t, \omega) = \sum_{k=1}^N |c_k|^2 P_{x_k}(t, \omega) + \sum_{j=1}^N \sum_{i=j+1}^N c_i c_j^* P_{x_i, x_j}(t, \omega) + c_j c_i^* P_{x_j, x_i}(t, \omega) \quad (33)$$

where $P_{x_k}(t, \omega)$'s are the N signal terms and the $P_{x_i, x_j}(t, \omega)$'s (for $i \neq j$) represent $\binom{N}{2} = N(N-1)/2$ interference terms—one for each pair of signal components. Note that the number of cross-term grows quadratically with the number of signal terms. The presence of cross-terms prohibits the interpretation of the WD as a pointwise time–frequency energy density and widely limits the readability of the TFR. In fact, cross-terms may appear in TFR as intensities in time and frequency locations, where no true components of the signal are present. ITs have precise geometry in the t - f plane, oscillating as a function of the distance between true signal terms; see Hlawatsch and Boudreaux-Bartels^[22] for details. The oscillatory nature of IT suggests performing a 2-D smoothing of the t - f plane to reduce them. If the properties of this smoothing filter do not vary on the t - f plane, the class of quadratic, shift-invariant, time–frequency distribution, known as *Cohen's Class*,^[2] is obtained.

$$P_C(t, \omega) = \int \psi(\tau - t, v - \omega) P_{\text{WD}}(\tau, v) d\tau dv \quad (34)$$

Each member of the class is derived by convolution between P_{WD} and a *kernel* ψ . Thus, every member of Cohen's Class may be interpreted as a 2-D, filtered version of WD^[22] and is unequivocally defined by ψ . This has practical implications. In fact, by placing constraints on the kernel, one obtains a subset of distributions that satisfies a particular property.^[2]

Introducing the ambiguity function,

$$A(\theta, \tau) = \int x\left(t - \frac{\tau}{2}\right) x^*\left(t + \frac{\tau}{2}\right) e^{j\omega t} dt \quad (35)$$

Cohen's Class can be rewritten as

$$C(\theta, \tau) = \phi(\theta, \tau) A(\theta, \tau) \quad (36)$$

where $\phi(\theta, \tau)$ is the 2-D FT of ψ , and the θ - τ plane is known as the correlative plane. In this plane, signal terms will be located around the origin, whereas the oscillating IT will tend to concentrate far from the origin. As evident from Eq. 36, the kernel merely weights the points of the correlative plane. Thus, to reduce the cross-terms effectively, the kernel must vanish when leaving the origin of the plane, and $\phi(\theta, \tau)$ must be the 2-D FT of a low-pass filter. In the analysis of biomedical signals,^[24,25] the Exponential Distribution (ED), the Reduced Interference Distribution (RID), and the Smoothed Pseudo Wigner-Ville (SPWD) are the most frequently used TFR. In particular, the SPWD expression

$$P_{\text{SPWD}}(t, \omega) = \int_{\tau} g(t - \tau) \int_v H(f - v) P_{\text{WD}}(\tau, v) d\tau dv \quad (37)$$

is characterized by an independent smoothing in frequency (obtained by $H(f)$ window) and time (obtained by $g(t)$ domain). The independent smoothing yields practical advantages, resulting in great flexibility in the choice of smoothing criteria, application, and efficient computation.

Autoregressive Time-Variant Models

A different approach to the analysis of time–frequency properties of a nonstationary signal is based on a recursive identification of AutoRegressive (AR) model parameters.^[25] Let us consider an AR process

$$x(t) = \sum_{k=1}^p a_k x(t - k) + w(t) = \varphi(t)^T \theta + w(t) \quad (38)$$

where the a_k are the model coefficients, p is the order, and $w(t)$ is a white-noise process ($w(t) \approx \text{WN}(0, \sigma^2)$). In Eq. 38, $\varphi(t) = [x(t-1), x(t-2), \dots, x(t-p)]^T$ is the observation vector, and $\theta(t) = [a_1, a_2, \dots, a_p]^T$ is the parameter vector. The model coefficients can be estimated by the least-squares (LS) approach, minimizing the following figure of merit:

$$J = \sum_{t=1}^N e(t)^2 \quad (39)$$

where N is the signal length and

$$e(t) = y(t) - \hat{y}(t) = y(t) - \varphi(t)^T \theta \quad (40)$$

is the prediction error—i.e., the difference between the current sample $y(t)$ and the model predicted one, $\hat{y}(t)$. The minimization of J leads to the solution^[26]

$$\begin{aligned} \hat{\theta}(t) &= \left[\sum_{k=1}^N \varphi(k)\varphi(k)^T \right]^{-1} \sum_{k=1}^N \varphi(k)x(k) \\ &= \mathbf{S}(N)^{-1} \mathbf{Q}(N) \end{aligned} \quad (41)$$

When using the LS method, the model is identified on a batch of data. If a new sample is added to sequence, the whole procedure has to be restarted. In addition, in many practical problems, the estimation needs to be performed in real time, sample by sample. Thus, it is necessary to employ recursive identification algorithms, where the estimation at time t is obtained by updating the estimation at time $t-1$ with the new information contained in the sample at time t . The LS method can be easily expressed in a recursive form, which is called recursive least squares (RLS). We have

$$\begin{aligned} \hat{\theta}(t) &= \hat{\theta}(t-1) + \mathbf{K}(t)e(t) \\ \mathbf{K}(t) &= \frac{\mathbf{P}(t-1)\varphi(t)}{1 + \varphi(t)^T \mathbf{P}(t-1)\varphi(t)} \\ e(t) &= y(t) - \varphi(t)^T \hat{\theta}(t) \\ \mathbf{P}(t) &= \left[\mathbf{P}(t-1) - \frac{\mathbf{P}(t-1)\varphi(t)\varphi(t)^T \mathbf{P}(t-1)}{1 + \varphi(t)^T \mathbf{P}(t-1)\varphi(t)} \right] \end{aligned} \quad (42)$$

where $\mathbf{P}(t) = \mathbf{S}^{-1}(t)$. We may observe that the parameter vector $\hat{\theta}(t)$ is updated by a factor that is proportional to the estimation error $e(t)$ through the gain vector $\mathbf{K}(t)$. This gain, which is also updated step by step, defines the property of the algorithm and modulates the extent to which the model errors are reflected in the estimation of $\hat{\theta}(t)$. If the data samples are coming from a nonstationary process, the estimated model suffers from the same problems as the batch estimation: in fact, the updating is based on the whole dataset, and after a while, the algorithm becomes unable to track changes. If we want the model to track the process modification, a forgetting factor has to be introduced to make the updating dependent mainly on the more recent data. The most popular forgetting function is based on the selection of an exponentially decreasing weight to the sequence of prediction error. A new figure of merit can be defined:

$$J = \frac{1}{t} \sum_{i=1}^t \mu^{t-i} e(i)^2 \quad (43)$$

The parameter μ ($0 < \mu < 1$) is the forgetting factor,^[25] and the quantity $1/(1-\mu)$ can roughly be seen as the

memory length of the algorithm, because it indicates approximately how many data samples are taken into account at time t to produce the estimate. The algorithm becomes

$$\begin{aligned} \hat{\theta}(t) &= \hat{\theta}(t-1) + \mathbf{K}(t)e(t) \\ \mathbf{K}(t) &= \frac{\mathbf{P}(t-1)\varphi(t)}{\mu + \varphi(t)^T \mathbf{P}(t-1)\varphi(t)} \\ e(t) &= y(t) - \varphi(t)^T \hat{\theta}(t) \\ \mathbf{P}(t) &= \frac{1}{\mu} \left[\mathbf{P}(t-1) - \frac{\mathbf{P}(t-1)\varphi(t)\varphi(t)^T \mathbf{P}(t-1)}{\mu + \varphi(t)^T \mathbf{P}(t-1)\varphi(t)} \right] \end{aligned} \quad (44)$$

The forgetting factor can be seen as the “tuning knob” of the algorithm. Low values for μ lead to estimates based on a small number of samples, and correspondingly, the algorithm results are quite reactive to changes in the true parameter vector. On the other hand, the averaging property necessary to effectively reduce the effect of the noise cannot be obtained if a short memory length is employed. Therefore, even if the true parameter vector stays constant, the estimates may be characterized by high variability using small values of μ .

The sequence of evolutionary spectra can be estimated from the time-varying parameter vectors as

$$P_{\text{AR}}(t, \omega) = \frac{\sigma^2(t)}{[1 + z^T \theta(t)]_{z=e^{j\omega}}^2} \quad (45)$$

where $z = [z^{-1}, z^{-2}, \dots, z^{-p}]^T$ and where $\sigma^2(t)$ is the variance of the prediction error, which also has to be estimated from the data and updated step by step.

Applications

In EEG processing, quadratic TFRs have been applied to analyze evoked or event-related potentials, as well as to characterize surface and cortical EEG recordings.^[24] The WD of EPs have been demonstrated to be useful for characterizing EP patterns, thanks to the excellent time–frequency resolution of the method.^[27] In addition, if associated with appropriate postprocessing, the RID was able to enhance subtle changes in EP’s response to physiological and psychological stimuli.^[28] At the level of cortical EEG recordings, the RID was used to characterize the recordings during epilepsy of the temporal lobe.^[29] Finally, WD was documented to be helpful in detecting rapid frequency changes related to changes in the level of consciousness of patients during anesthesia—changes that could not be captured by traditional spectral analysis.^[30]

In the analysis of the ECG, the frequency contents of the main waves (P–QRS–T) have been explored. Aims were the detection of P-wave, the analysis of

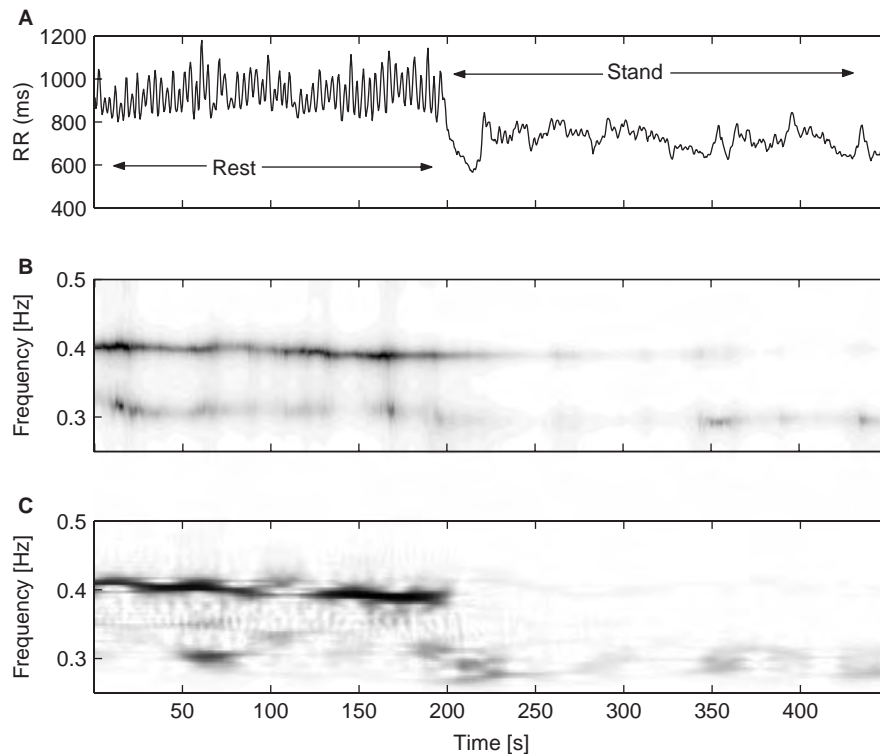


Fig. 3 (A) HRV signal during a rest-to-stand maneuver. Time–frequency representation of the signal obtained by SPWD. (B) Time-smoothing rectangular window $M=25$; frequency-smoothing Gaussian window ($\alpha=1.2$) and AR time-variant model (C), (model order 8, $\omega=0.98$). Both methods provide similar results, which is relevant information to describe the dynamical changes in LF and HF rhythms during the test. A dominant, quite stable, HF component is present during rest, but it decreases during stand. Conversely, LF components become predominant during stand, and their amplitude oscillates in time.

VLP,^[31] or the characterization of normal and pathological complexes.^[32] In particular, higher-frequency (>90 Hz) contents of the QRS, when exploited by SPWD, were augmented in patients after myocardial infarction than in normal subjects, and these components were present for the whole QRS complexes.^[31] In addition, TFR of the QRS and ST–T tracts were used to evaluate the efficiency of thrombolytic treatments.^[32] Finally, the time–frequency resolution properties of SPWD may be used to detect ventricular fibrillation events^[33] or to characterize the dynamic variation of the frequency of atrial activity during atrial fibrillation.^[34]

The beat-to-beat fluctuations of the RR interval (the so-called Heart Rate Variability [HRV]) were also investigated by quadratic TFR.^[25] The power spectrum of this signal evidences two main rhythms: (1) a high-frequency (HF) component, synchronous with respiration (in a range between 0.2 and 0.45 Hz) and modulated by vagal tone, and (2) a low-frequency (LF) component (centered around 0.1 Hz) mainly modulated by sympathetic activity. The dynamical changes of these rhythms were quantified through SPWD during a provocative test (i.e., tilt test) evidencing different responses of LF and HF rhythms: a fast reduction of HF component is followed by a more gradual increase of LF. In addition, LF was less stable in patients undergoing a tilt-induced syncopal event. These patients also experienced higher very-low-frequency (<0.03 Hz) components in the early phase of tilt.^[35] An example of SPWD

of RR series during a rest-to-tilt maneuver is shown in Fig. 3.

Other applications deal with the analysis of electromyograms (EMG) for the evaluation of fatigue in muscle during slow dynamic contractions^[36] or the analysis of electrogastrograms (EGG) for detecting gastric contractions and stomach motility.^[37]

Time-variant estimators have been used mainly in the presence of pseudo-stochastic signals or when noise is predominant. Typical applications deal with the analysis of HRV signals during pharmacologically induced ischemia,^[38] autonomic tests, or syncope events.^[39] In these situations, the evolutionary spectra of Eq. 45 have been documented to track the spectral changes associated with the dynamic variation of ANS control mechanisms.^[25]

In neurophysiology, only a few applications have been suggested to assess the dynamic variation of EEG signals using adaptive AR identification.^[40] Other applications include the analysis of the EGG^[41] and uterine EMG.^[42] In the latter work, several TFR methods were compared. The one based on an adaptive AR model provides the best tracking of main EMG spectral peaks, even if quadratic TFR may provide a more accurate quantification of spectral parameters.

HIGHER-ORDER SPECTRA

The Power Spectrum Density (PSD) of a signal may be obtained as the FT of the autocorrelation function

$(R_x(\tau))$ of the signal^[1]

$$S_x(\omega_1) = \sum_{\tau_1=-\infty}^{\infty} R_x(\tau_1) \exp^{-j\omega_1 \tau_1} \quad (46)$$

and the information carried by PSD is essentially contained in R_x . Being a second-order statistic, R_x provides complete statistical description of Gaussian processes. There are situations, however, in which it could be useful to look beyond PSD introducing higher-order spectra based on higher-order statistic (HOS) of the signal. The key advantage over traditional second-order techniques is that phase information is preserved. PSD is blind to phase, and as a consequence, second-order techniques cannot be used to identify nonminimum phase signals or to check certain types of phase coupling (associated with nonlinearities).^[43] In addition, using the fact that for Gaussian processes, all cumulants of order greater than two are identical to zero, HOS can be used to assess deviation from Gaussianity or to filter out Gaussian noise.

Moments and Cumulants

Let's consider a stochastic variable x with known probability density function $p(x)$. The characteristic function is defined as

$$\begin{aligned} \phi(\gamma) &= \int_{-\infty}^{\infty} p(x) e^{j\gamma x} dx \\ &= 1 + j\gamma \int_{-\infty}^{\infty} xp(x) dx \\ &\quad + \frac{(j\gamma)^2}{2} \int_{-\infty}^{\infty} x^2 p(x) dx + \dots \end{aligned} \quad (47)$$

or

$$\phi(\gamma) = \sum_{k=0}^{+\infty} \frac{(j\gamma)^k}{k!} \int_{-\infty}^{\infty} x^k p(x) dx = \sum_{k=0}^{+\infty} \frac{(j\gamma)^k}{k!} M_x^k \quad (48)$$

where the M_x^k are the k th-order central moment. The characteristic function uniquely defines the statistical distribution and can be used to generate both the n -order moments

$$M_x^k = (-j)^k \frac{\partial^k \phi(\gamma)}{\partial \gamma^k} \Big|_{\gamma=0} \quad (49)$$

and the cumulants

$$C_x^k = (-j)^k \frac{\partial \ln \phi(\gamma)}{\partial \gamma^k} \Big|_{\gamma=0} \quad (50)$$

Intuitively, moments and cumulants are closely related, and a formula exists to compute the latter from the former, and vice versa.^[3]

Let's consider a zero-mean, stationary random process $x(n)$, $\{n=0,1,2,3,\dots\}$. Similarly to Eqs. 49 and 50, we may define the joint k -order moments (cumulant) of the process

$$\begin{aligned} M_x^k &= (-j)^k \frac{\partial^k \phi(\gamma_1, \gamma_2, \dots, \gamma_k)}{\partial \gamma_1 \partial \gamma_2 \dots \partial \gamma_k} \Big|_{\gamma_1=\gamma_2=\dots=\gamma_k=0} \\ C_x^k &= (-j)^k \frac{\partial^k \ln \phi(\gamma_1, \gamma_2, \dots, \gamma_k)}{\partial \gamma_1 \partial \gamma_2 \dots \partial \gamma_k} \Big|_{\gamma_1=\gamma_2=\dots=\gamma_k=0} \end{aligned} \quad (51)$$

In particular

$$\begin{aligned} M_x^k(\tau_1, \tau_2, \dots, \tau_{k-1}) \\ = E[x(k)x(k+\tau_1)x(k+\tau_2)\dots(k+\tau_{k-1})] \end{aligned} \quad (52)$$

represents the k th-order moment, and it depends on time difference τ_i due to the assumption of stationarity. We may observe that M_x^1 and M_x^2 are the mean and the variance of the process. The cumulant functions become^[44]

$$\begin{aligned} n=2 \quad C_x^2(\tau_1) &= M_x^2(\tau_1) \\ n=3,4 \quad C_x^k(\tau_1, \tau_2, \dots, \tau_{k-1}) &= M_x^k(\tau_1, \tau_2, \dots, \tau_{k-1}) \\ &\quad - M_G^k(\tau_1, \tau_2, \dots, \tau_{k-1}) \end{aligned} \quad (53)$$

where $M_G^k(\tau_1, \dots, \tau_{k-1})$ is the k th-order moment of an equivalent Gaussian signal with the same mean and autocorrelation function of $x(n)$. It turns out that $C_x^n(\tau_1, \dots, \tau_{n-1}) = 0$ for Gaussian processes for orders $n=3$ and $n=4$. Thus, cumulants provide a measure of the distance of a random process from Gaussianity.

The C_k^n 's have several interesting properties^[44]: They are scaled quantities, symmetric and additive in their arguments, and blind to additive constants. In addition, the cumulant of the sum of statistical independent variables is the sum of cumulants of the individual quantities. Thus, if a certain process x is corrupted by Gaussian noise w , we have $C_{x+w}^k = C_x^k + C_w^k = C_x^k$, and the resulting cumulants are blind to the superimposed Gaussian noise. It is worth noting that the additive property does not hold for higher-order moments. Therefore, it usually is easier to use cumulants than moments as operators. A more detailed description of cumulants and moments can be found in Nikias and Petropulu.^[3]

Practically speaking, the estimation of higher-order cumulants from sampled data becomes harder when the order is increased, because longer datasets are required to reduce the variance associated with the estimate of joint probabilities. Therefore, computation usually is restricted to third- and fourth-order cumulants.

Higher-Order Spectra

The higher-order spectra (or polyspectra) are defined as^[44]

$$S_x^k(\omega_1, \omega_2, \dots, \omega_{k-1}) = \sum_{\tau_1=-\infty}^{\infty} \dots \sum_{\tau_{k-1}=-\infty}^{\infty} C_x^k(\tau_1, \tau_2, \dots, \tau_{k-1}) \exp(-j \sum_{i=1}^{k-1} \omega_i \tau_i) \quad (54)$$

where

$$|\omega_1| < \pi, |\omega_2| < \pi, \dots, |\omega_{k-1}| < \pi,$$

$$|\omega_1 + \omega_2 + \dots + \omega_{k-1}| < \pi.$$

Polyspectra, therefore, are the $(k-1)$ -dimensional FT of the k th-order cumulant, and the power spectrum (Eq. 46) is in fact a member of the class because $R_x = C_x^2$. Even if polyspectra can be defined in terms of either moments or cumulants, the latter are preferable when dealing with stochastic process. Reasons are contained in the properties of cumulants described above.

In particular, the FT of third-order cumulants is known as bispectrum.^[43]

$$S_x^3(\omega_1, \omega_2) = \sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} C_x^3(\tau_1, \tau_2) \exp(-j(\omega_1 \tau_1 + \omega_2 \tau_2)) = |S_x^3(\omega_1, \omega_2)| \exp(j\psi(\omega_1, \omega_2)) \quad (55)$$

Bispectrum generally is a complex function; it carries information on both amplitude and phase (conversely, the PSD is a real quantity, blind to phase). Taking into account the properties of symmetry of cumulant, the bispectrum is highly symmetric in the ω_1 - ω_2 plane. In particular, 12 regions of symmetry can be evidenced, and the knowledge of S_x^3 in the triangular region (known as the principal region) $\omega_2 \geq 0$, $\omega_1 \geq \omega_2$, $\omega_1 + \omega_2 \leq \pi$ is enough for a complete description of the bispectrum.^[43]

The bicoherence is a normalized bispectrum

$$K_x^3(\omega_1, \omega_2) = \frac{C_x^3(\omega_1, \omega_2)}{\sqrt{S_x^2(\omega_1) S_x^2(\omega_2) S_x^2(\omega_1 + \omega_2)}} \quad (56)$$

that combines the third-order cumulant spectrum and the power spectrum (S_x^2) of a signal. Bicoherence is useful for detection and characterization of nonlinearities in the time series and to discriminate between linear and nonlinear processes. Consider

$$z(n) = \sum_{i=1}^3 \cos(\omega_i n + \varphi_i) \quad (57)$$

where, $\omega_2 > \omega_1 > 0$, $\omega_3 = \omega_1 + \omega_2$, φ_1, φ_2 are independent and uniformly distributed on $0-2\pi$. Let's consider

two situations: (1) φ_3 is uniformly distributed on $0-2\pi$ and independent from, φ_1, φ_2 and (2) the relation $\varphi_3 = \varphi_1 + \varphi_2$ links phases—i.e., a Quadratic Phase Coupling (QPC) does exist between ω_1 and ω_2 harmonics. QPC results in phase relations that are the same as the frequency relations. QPC can arise only among harmonically related components. In the first case, observing that the harmonics are phase independent and that third-order cumulant of complex harmonic is zero, the resulting magnitude bispectrum (evaluated in the principal region) is zero everywhere. Conversely, in case (2), the bispectrum is zero everywhere, but in the point, ω_1, ω_2 indicating that this pair of frequencies are phase coupled. Because only phase-coupled components contribute to the bispectrum, the latter is a useful tool for discriminating quadratic phase-coupling components.^[43]

Applications

The use of polyspectra has attracted much attention for the processing of many biological signals, including EEG analysis, EMG studies, and interaction among rhythms. In most of the cases, interest has been limited to computing bispectrum, assessing the nature of the frequencies contained in the signal, and studying QPC. For the analysis of EEG, the bispectrum has some remarkable properties: it enhances the signal-to-noise ratio for nonGaussian EEG; it identifies nonlinearities and detects phase coupling (which is a common phenomenon in signals generated from the central nervous system); and it quantifies both power and phase information. In particular, HOS analysis of EEG is the basis of the well-known Bispectral Index (BIS), developed by Aspect Medical System (Newton, Massachusetts). BIS is a composite index of parameters^[45] used to assess the depth of anesthesia and the level of unconsciousness. It includes time domain, frequency domain indexes, and a bispectral metric (the SynchFastSlow ratio) computed as the logarithm of the ratio of cumulative values of the bispectrum over two ranges (0.5–47 Hz and 40–47 Hz). Even if the BIS does not work properly in all situations, the method is popular as a result of strong empirical evidence of high probability to predict the absence of consciousness correctly.^[46] Other EEG applications include the possibility of assessing the place and the extent of focal ischemic cerebral injury^[47] or characterizing sleep spindles.^[48]

Higher-order statistics have been also applied for diagnostic assessment and characterization of lung sounds. Harmonic analyses of lung sounds have been combined with the bicoherence index to provide reliable evaluation of pulmonary dysfunctions. Especially in noisy environments, the use of HOS was proved to be advantageous.^[49]

Even if some promising applications have been suggested, HOS analysis of BSP is limited by the difficulty of estimating high-order statistics accurately when short data sequences are available (which is often the case).

NONLINEAR BIOMEDICAL SIGNAL PROCESSING

Chaotic systems are characterized by complex dynamics, determinism, and sensitivity to initial conditions, making long-term forecasting impossible. Several signals and series of biological origin exhibit patterns typical of complex dynamics. The topic is certainly too

wide to be covered fully by this entry. Textbooks^[4] are available to the interested reader. Here, we mention only a few popular signal processing methodologies largely used to investigate biological time series. They are (1) Detrended Fluctuation Analysis (DFA) for exploring scaling behavior and (2) entropy-based metrics for assessing series regularity.

Detrended Fluctuation Analysis

Several signals and series of biological origin display scaling (self-similar) behavior. The concept is evidenced in Fig. 4, where the HRV signal is observed at various time windows (scales): the signal has an irregular

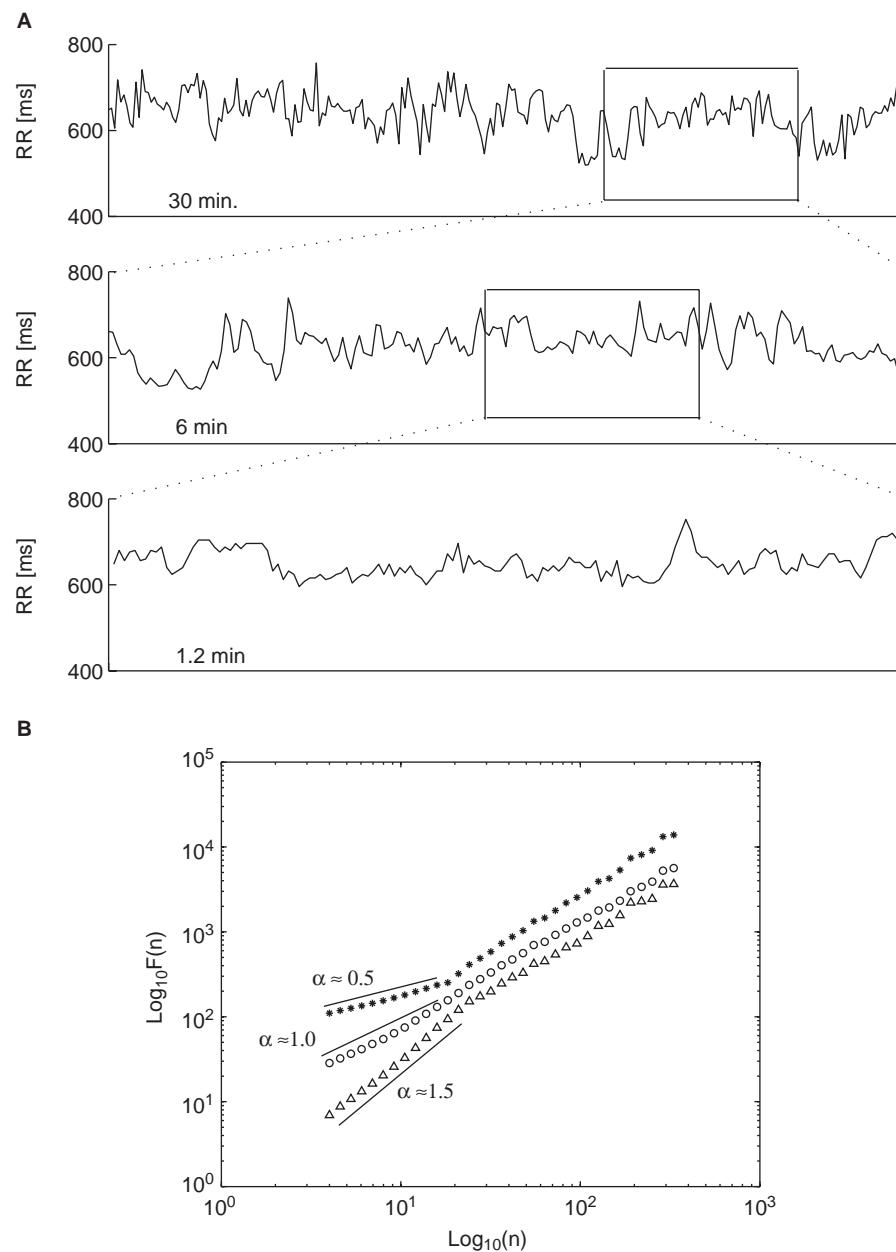


Fig. 4 (A) Evidence of self-similar fluctuations in the dynamic of HRV signal. On different time scales, the static properties of the signal are self-similar (i.e., the signal resembles itself when observed with different time lenses). (B) Examples of scaling properties in the HRV fluctuations and their changes in health (triangle), ageing (circle), and chronic heart failure disease (star).

appearance, which is statistically similar among scales. It is not possible to distinguish a scale from the others, and there is a lack of a characteristic (single) scale for this process. A scale invariance is observed, and the process is self-similar. Far from being just a fancy manifestation, the existence of scaling properties is an useful probe for understanding the inner nature of biological phenomena and their modifications with age and pathologies.^[50] Among the numerous tools developed to assess scaling behavior,^[8] DFA is likely to be the method most used, as its implementation is straightforward, and it is robust and blind to superimposed linear trends.^[51] DFA permits the detection of long-range correlations embedded in a seemingly non-stationary time series, and it also averts the spurious detection of apparent long-range correlations that are artifacts of nonstationarity.^[51]

Let us consider a discrete-time series $x(n)$ $n = \{1, 2, \dots, N\}$ and integrate it:

$$y(k) = \sum_{i=1}^k x(i) - E[x] \quad (58)$$

Next, the integrated time series is divided into boxes of equal length n . In each box of length n , an LS line is fit to the data, representing the trend in that box. Finally, the integrated time series is detrended by subtracting the local trend $y_n(k)$ in each box, and the root mean-square fluctuation of this integrated and detrended time series is calculated by

$$F(n) = \sqrt{E[(y(k) - y_n(k))^2]} \quad (59)$$

The computation is repeated over all time scales (box size) to provide a relationship among $F(n)$, the average fluctuation as a function of box size, and the box size n .

If this relationship is linear on a log–log plot, it indicates the presence of power law scaling—that is,

$$F(n) \sim n^\alpha \quad (60)$$

where α is the scaling exponent. Taking the logarithm, we get a linear relationship

$$\log[F(n)] \sim \alpha \log(n) \quad (61)$$

and the scaling exponent can be evaluated by the slope of a fitting curve.

This exponent provides a measure of the “roughness” of the original time series: the larger the value of α , the smoother the time series.^[50] A value of $\alpha = 0.5$ indicates a white-noise process (complete unpredictability), whereas $\alpha = 1.5$ is characteristic of Brownian noise. Finally, a value of $\alpha = 1$ indicates $1/f^2$ -like noise behavior.

Approximate Entropy

Approximate entropy (ApEn) is a regularity statistic quantifying the unpredictability of fluctuations in a time series. Intuitively, the presence of repetitive patterns of fluctuation in a time series makes it more predictable than a time series in which such patterns are absent. ApEn reflects the likelihood that “similar” patterns of observations will not be followed by additional “similar” observations. A time series containing many repetitive patterns (i.e., a regular and predictable series) has a relatively small ApEn; a less predictable (i.e., more complex) process has a higher ApEn.

Given a sequence S_N consisting of N samples $x(n)$ $n = \{1, 2, \dots, N\}$, we define the pattern of length m $p_m(i) = \{x(i), x(i+1), \dots, x(i+m-1)\}$ as the sequence of m samples that begins at interval i . We now define criterion of similarity r , and two patterns, $p_m(i)$ and $p_m(j)$, are considered similar if the difference between any pair of corresponding measurements in the patterns is less than r —i.e., if for $|x(i+k) - x(j+k)|$.

Now consider the set P_m of all patterns of length m within S_N . We may now define

$$C_{im}(r) = \frac{n_{im}(r)}{N - m + 1} \quad (62)$$

where $n_{im}(r)$ is the number of patterns in P_m that are similar to $p_m(i)$. The quantity $C_{im}(r)$ is the fraction of patterns of length m that resemble the pattern of the same length that begins at interval i . $C_{im}(r)$ is calculated for each pattern in P_m and $C_m(r) = E[C_{im}(r)]$ is defined as the mean of these $C_{im}(r)$ values. The quantity $C_m(r)$ expresses the prevalence of repetitive patterns of length m in S_N . Finally, the approximate entropy of S_N , for patterns of length m and similarity criterion r is defined as^[52]:

$$\text{ApEn}(S_N, M, r) = \ln \left[\frac{C_m(r)}{C_{m+1}(r)} \right] \quad (63)$$

i.e., as the natural logarithm of the relative prevalence of repetitive patterns of length m compared with those of length $m+1$.

Regularity Index

The regularity of a signal may be quantified by using its Conditional Entropy (CE)—the amount of information carried by the signal when its past samples are known.^[53] From a given process $x = \{x(n), n = 1, \dots, N\}$, let's label $x_L(n)$ the $N-L+1$ patterns of length L that may be extracted from x and p_L its probability. In terms of the Shannon entropy $E(x)$, the CE may be

defined as

$$CE(L) = E(x_m) - E(x_{m-1}) \quad (64)$$

which emphasizes the amount of information carried out when the pattern is enlarged from $m-1$ to m . Eq. 64 has several computational advantages, but its major drawback is that estimation of $CE(L)$ is no longer statistically consistent when L increases.^[53] Therefore, when a limited amount of samples is available, the Corrected Conditional Entropy (CCE) must be used,

$$CCE(L) = CE(L) + \text{perc}(L)E(x) \quad (65)$$

where $\text{perc}(L)$ is the percentage of single points in the L -dimensional phase space and $E(1,12)$ is the Shannon Entropy for $L=1$. As $\text{perc}(L)$ increases with L , the corrective term tends to compensate the decrease in $CE(L)$ due to the increasing number of single ($L-1$) patterns giving a null contribution to $CE(L)$. CCE is the sum of two contributions: the first one decreasing and the second one increasing with L . The best estimate of CE is the minimum of $CCE(L)$. An index of regularity may be defined as follows,^[53]

$$R = 1 - \min\left(\frac{CCE(L)}{E(x)}\right) = 1 - \min(\text{NCCE}(L)) \quad (66)$$

where NCCE is the CCE value normalized by $E(x)$, to derive an index that is independent of the different probability distribution of the processes. R tends to be 1 when the process x is periodic (i.e., it is perfectly

predictable), falls to zero in the presence of high complexity (i.e., completely unpredictable) signals, and remains in between the two positions for intermediate processes.

Applications

DFA is a widely used technique for the assessment of long-range correlations in nonstationary time series. The most straightforward and studied examples can be found in the analysis of both beat-to-beat fluctuation of the HRV^[50] and DNA sequences,^[54] but other evidences may be found in the analysis of human walking^[55] or neuron spiking.^[56] An example of DFA applied to the analysis of long-term fluctuations in the heartbeat series is shown in Fig. 4b. Two regions with different slopes are evidenced, suggesting that there are both short-range ($0 < n < 16$) and a long-range scaling exponents. Differences in scaling properties induced by age and pathology have been also reported.^[50] In particular, the low scaling exponents were powerful predictors of mortality among patients surviving myocardial infarction^[57] or congestive heart failure.^[58] In addition, DFA indexes were superior to spectral analysis in the classification of sleep stages and in separation of apnea severity.^[59] In the analysis of DNA sequences, DFA was used to assess long-range correlations in coding and noncoding regions of the gene.^[54]

ApEn has gained much attention as an indicator of in-depth anesthesia monitoring, differentiating between awake and asleep states,^[60] classifying the burst suppression patterns,^[61] or monitoring the transitions

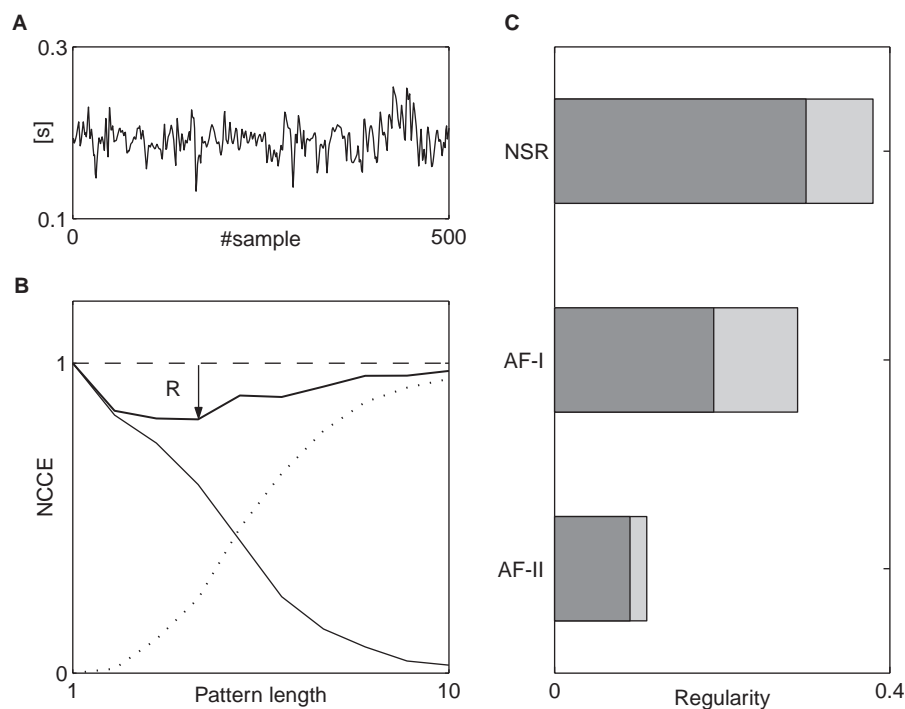


Fig. 5 (A) Local Atrial Period (LAP) series recorded during AF-I episode. (B) The estimated normalized corrected conditional entropy functions (bold line) is obtained as the sum of the conditional entropy term (thin line) and the corrective term (dotted line). The regularity index (R) is obtained as the distance between the minimum of the resulting curve and the unity line (dotted). (C) Mean (\pm std) value of R estimated during Normal Sinus Rhythm (NSR), AF-I, and AF-II. It is interesting to note the decreasing of R passing from NSR to organized (AF-I) and not organized (AF-II) fibrillatory rhythms.

between awareness and unresponsiveness.^[62] In the analysis of HR series, changes of ApEn have been correlated with the aging process^[63] or with the phases of fetal growth.^[64] In an interesting application, entropy-based indexes have been proposed as indicators of the early diagnosis of sepsis in premature newborns in intensive care units.^[65] In addition, altered RR dynamics have been assessed before the spontaneous onset of paroxysmal atrial fibrillation (AF) episodes.^[66] Other applications deal with the evaluation of intracranial pressure^[67] or the analysis of respiratory patterns.^[68]

The Regularity index (R) has mainly been applied to evaluate complexity in cardiovascular beat-to-beat variability signals or in sympathetic nerve discharge^[53] during various stimulations induced by head-up tilt, controlled respiration, or peripheral muscarinic blockade.^[69] In general, systolic arterial pressure (SAP) series is more regular than RR series; regularity of RR, however, is influenced by the above stimulus, whereas regularity of SAP is not.

In addition, R was used to characterize the dynamic of the Local Atrial Period (LAP) series (i.e., the sequences of time intervals computed among successive atrial activations) during Well's types of Atrial Fibrillation (namely, AF-I and AF-II type). An example of LAP series during AF is shown in Fig. 5: LAP series (a) and corresponding NCCE (b) were obtained during a Wells AF-II type episode. When comparing NSR and AF rhythms, it has been documented that $R^{[70]}$ is able to detect the atrial rhythms (see Fig. 5c): regularity is significantly reduced by passing from NSR to AF-I to AF-II. Such a reduction measures the reduced atrial organization in wavefront propagation characteristic of AF-I and AF-II events, and demonstrates a more fragmented atrial activation.

CONCLUSION

We have introduced some advanced signal processing methods that have been used to extract new quantitative information from biological signals. The methods include wavelets transform, time-frequency analysis, and higher-order spectra, as well as some known tools from nonlinear analysis of biological series. The advent of these tools in BSP makes it possible to quantify new signal features and considerably enlarge the number of parameters that can be extracted from recorded vital signals, opening new perspectives for a deeper comprehension of biological phenomena and pathologies. The few examples were selected to emphasize the new available information from the analysis of widely analyzed signals, such as ECG, EEG, and EP. Finally, it is important to point out that a larger number of parameters does not necessarily mean more information on the biological process if we are not able to correctly

handle and integrate information coming from modern and traditional approaches. In fact, the abundance of parameters opens new questions and problems, such as those related to the appropriate fusion of the information for a better understanding of physiological systems or for a better definition of patient status.



ARTICLE OF FURTHER INTEREST

Biopotential Amplifiers

REFERENCES

1. Mendez, M.; Mainardi, L.T. *Digital Biomedical Signal Acquisition and Processing: Basic Topics*, Encyclopedia of Biomaterials and Biomedical Engineering; Taylor & Francis: New York, 2006; 1–11.
2. Cohen, L. Time-frequency distributions—a review. *Proc. IEEE* **1989**, *77*, 941–981.
3. Nikias, C.L.; Petropulu, A.P. *Higher Order Spectra Analysis: A Nonlinear Signal Processing Framework*; Prentice Hall: Englewood Cliffs, NJ, 1993.
4. Akay, M. Dynamic analysis and modeling. In Akay, M., Ed.; *Nonlinear Biological Signal Processing*; IEEE Press: Piscataway, NJ, 2001; Vol. II, 159–213.
5. Rioul, O.; Vetterli, M. Wavelets and signal processing. *IEEE Signal Process.* **1991**, 14–38.
6. Cohen, A.; Kavacevic, J. Wavelets: the mathematical background. *Proc. IEEE* **1996**, *84*, 514–522.
7. Unser, M.; Aldroubi, A. A review of wavelets in biomedical applications. *Proc. IEEE* **1996**, *84*, 626–638.
8. Audit, B.; Bacry, E.; Muzy, J.F.; et al. Wavelet-based estimators of scaling behavior. *IEEE Trans. Inf. Theory* **2002**, *11* (4), 2938–2954.
9. Addison, P.S. Wavelet transform and the ECG: a review. *Physiol. Meas.* **2005**, *26* (5), R155–R199.
10. Li, C.; Zheng, C.; Tai, C. Detection of ECG characteristic points using wavelet transforms. *IEEE Trans. Biomed. Eng.* **1995**, *1*, 21–28.
11. Martinez, J.P.; Almeida, R.; Olmos, S.; Rocha, A.P.; Laguna, P.A. A wavelet-based ECG delineator: evaluation on standard data bases. *IEEE Trans. Biomed. Eng.* **2004**, *4*, 570–581.
12. Senhadji, L.; Carrault, G.; Bellanger, J.J.; Passariello, G. Comparing wavelet transforms for recognizing cardiac patterns. *IEEE Trans. Med. Biol.* **1995**, *13*, 167–173.
13. Mallat, S.; Zhong, S. Characterization of signals from multiscale edges. *IEEE Trans. Pattern Anal. Mach. Intell.* **1992**, *14*, 710–732.
14. Meste, O.; Rix, H.; Caminal, P.; Thakor, N.V. Ventricular late potentials characterization in time-frequency domain by means of a wavelet transform. *IEEE Trans. Biomed. Eng.* **1994**, *7*, 625–634.
15. Khadra, L.; Dickhaus, H.; Lipp, A. Representations of ECG late potentials in the time frequency plane. *J. Med. Eng. Techn.* **1993**, *17*, 228–231.

16. Jalaleddine, S.; Hutchens, C.; Strattan, R.; Coberly, W. ECG data compression techniques—a unified approach. *IEEE Trans. Biomed. Eng.* **1990**, *37*, 329–343.
17. Lu, Z.; Kim, D.Y.; Pearlman, W.A. Wavelet compression of ECG signals by the set partitioning in hierarchical trees algorithm. *IEEE Trans. Biomed. Eng.* **2000**, *47*, 849–855.
18. Khan, Y.U.; Gotman, J. Wavelet based automatic seizure detection in intracerebral electroencephalogram. *Clin. Neurophysiol.* **2003**, *114*, 898–908.
19. Bertrand, O.; Bohorquez, J.; Pernier, J. Time–frequency digital filtering based on an invertible wavelet transform: an application of evoked potentials. *IEEE Trans. Biomed. Eng.* **1994**, *41*, 45–69.
20. Bartnik, E.A.; Blinowska, K.J.; Durka, P.J. Single evoked potential reconstruction by means of wavelet transform. *Biolog. Cybern.* **1992**, *67*, 175–181.
21. Causevic, E.; Morley, R.E.; Wickerhauser, M.V.; Jacquin, A.E. Fast wavelet estimation of weak biosignals. *IEEE Trans. Biomed. Eng.* **2005**, *52*, 1021–1032.
22. Hlawatsch, F.; Boudreaux-Bartels, G.F. Linear and quadratic time–frequency signal representations. *IEEE SP* **1992**, 21–67.
23. Flandrin, P. Some features of time–frequency representations of multicomponent signals, acoustics, speech, and signal processing. *IEEE Int. Conf. ICASSP'84* **1984**, *9*, 266–269.
24. Williams, W.J. Reduced interference distribution: biological application and interpretations. *Proc. IEEE* **1996**, *84*, 1264–1280.
25. Mainardi, L.T.; Bianchi, A.M.; Cerutti, S. Time–frequency and time-varying analysis for assessing the dynamic responses of cardiovascular control. *Crit. Rev. Biomed. Eng.* **2002**, *30*, 175–217.
26. Soderstrom, T.; Stoica, P. *System Identification*; Prentice Hall International: London, U.K., 1989.
27. Morgan, N.H.; Gevins, A.S. Wigner distributions of human event-related brain potentials. *IEEE Trans. Biomed. Eng.* **1986**, *33*, 66–70.
28. Williams, W.J.; Shevrin, H.; Marshall, R.E. Information modeling and analysis of event related potentials. *IEEE Trans. Biomed. Eng.* **1987**, *51*, 737–743.
29. Zaveri, H.P.; Williams, W.J.; Iasemidis, L.D.; Sackellares, J.C. Time–frequency representation of electrocorticograms in temporal lobe epilepsy. *IEEE Trans. Biomed. Eng.* **1992**, *39*, 502–509.
30. Nayak, A.; Roy, R.J.; Sharma, A. Time–frequency spectral representation of the EEG as an aid in the detection of depth of anesthesia. *Ann. Biomed. Eng.* **1994**, *22*, 501–513.
31. Novak, P.; Li, Z.; Novak, V.; Hatala, R. Time–frequency mapping of the QRS complex in normal subjects and in postmyocardial infarction patients. *J. Electrocardiol.* **1994**, *27*, 49–60.
32. Chouvarda, I.; Maglaveras, N.; Boufidou, A.; Mohlas, S.; Louridas, G. Wigner-Ville analysis and classification of electrocardiograms during thrombolysis. *Med. Biol. Eng. Comput.* **2003**, *41*, 609–617.
33. Zhao, Y.; Atlas, L.E.; Marks, R.J., II. The use of cone-shaped kernels for generalized time–frequency representations of nonstationary signals. *IEEE Trans. ASSP* **1990**, *38*, 1084–1091.
34. Stridh, M.; Sornmo, L.; Meurling, C.J.; Olsson, S.B. Characterization of atrial fibrillation using the surface ECG: time-dependent spectral properties. *IEEE Trans. Biomed. Eng.* **2001**, *48*, 19–27.
35. Novak, V.; Novak, P.; Kus, T.; Nadeau, R. Slow cardiovascular rhythms in tilt and syncope. *J. Clin. Neurophysiol.* **1995**, *12*, 64–71.
36. Knaflitz, M.; Bonato, P. Time–frequency method applied to muscle fatigue assessment during dynamic contractions. *J. Electromyogr. Kinesiol.* **1999**, *9*, 337–350.
37. Lin, Z.; Chen, J.D. Time–frequency representation of the electrogastrogram-application of the exponential distribution. *IEEE Trans. Biomed. Eng.* **1994**, *41*, 267–275.
38. Petrucci, E.; Mainardi, L.T.; Balian, V.; Ghiringhelli, S. Assessment of heart rate variability changes during dipyridamole infusion and dipyridamole-induced myocardial ischemia: a time variant spectral approach. *J. Am. Coll. Cardiol.* **1996**, *28*, 924–934.
39. Furlan, R.; Piazza, S.; Dell'Orto, S.; Barbic, F.; Bianchi, A.; Mainardi, L.; Cerutti, S.; Pagani, M.; Malliani, A.; et al. Cardiac autonomic patterns preceding occasional vasovagal reactions in healthy humans. *Circulation* **1998**, *31(98)*, 1756–1761.
40. Muthuswamy, J.; Thakor, N.V. Spectral analysis methods for neurological signals. *J. Neurosci. Methods* **1998**, *31* (83), 1–14.
41. Chen, J. A computerized data analysis system for electrogastrogram. *Comput. Biol. Med.* **1992**, *22*, 45–57.
42. Duchène, J.; Devedeux, D.; Mansour, S.; Marque, C. Analyzing uterine EMG: tracking instantaneous burst frequency. *IEEE Eng. Med. Biol.* **1995**, *41*, 1867–1880.
43. Nikias, C.L.; Mendel, J.M. Signal processing with higher-order spectra. *IEEE Signal Process.* **1993**, *64* (2), 10–37.
44. Mendel, J.M. Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications. *Proc. IEEE* **1991**, *79* (3), 278–305.
45. Rampil, I.J. A primer for EEG signal processing in anesthesia. *Anesthesiology* **1998**, *89*, 980–1020.
46. Glass, P.S.A.; Bloom, M.; Kearsse, L.A., Jr.; et al. Bispectral analysis measures sedation and memory effects of propofol, midazolam, isoflurane and anfen-tanil in healthy volunteers. *Anesthesiology* **1997**, *86*, 836–847.
47. Zhang, J.W.; Zheng, C.X.; Xie, A. Bispectrum analysis of focal ischemic cerebral EEG signal using third-order recursion method. *IEEE Trans. Biomed. Eng.* **2000**, *47*, 352–359.
48. Akgul, T.; Mingui, S.; Schlaassi, R.J.; et al. Characterization of sleep spindles using higher order statistics and spectra. *IEEE Trans. Biomed. Eng.* **2000**, *47*, 997–1009.
49. Hadjileontiadis, L.J.; Panas, S.M. Higher-order statistics: a robust vehicle for diagnostic assessment and characterisation of lung sounds. *Technol. Health Care* **1997**, *5*, 359–374.
50. Goldberger, A.L.; Amaral, L.A.N.; Hausdorff, J.M.; Ivanov, P.Ch.; Peng, C.-K.; Eugene Stanley, H. Fractal dynamics in physiology: alterations with disease and aging. *PNAS* **2002**, *99* (S1), 2466–2472.

51. Peng, C.K.; Havlin, S.; Stanley, H.E.; Goldberger, A.L. Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos* **1995**, *5*, 82–87.
52. Pincus, S.M. Approximate entropy as a measure of system complexity. *Proc. Natl Acad. Sci. U.S.A.* **1991**, *88*, 2297–2301.
53. Porta, A.; Baselli, G.; Liberati, D.; Montano, N.; Cogliati, C.; Gnechi-Ruscione, T.; Malliani, A.; Cerutti, S. Measuring regularity by means of a corrected conditional entropy in sympathetic outflow. *Biol. Cybern.* **1998**, *78*, 71–78.
54. Peng, C.K.; Buldyrev, S.V.; Goldberger, A.L.; Havlin, S.; Mantegna, R.N.; Simons, M.; Stanley, H.E. Statistical properties of DNA sequences. *Physica A* **1995**, *221*, 180–192.
55. Hausdorff, J.M.; Mitchell, S.L.; Firtion, R.; Peng, C.K.; Cudkowicz, M.E.; Wei, J.Y.; Goldberger, A.L. Altered fractal dynamics of gait: reduced stride-interval correlations with aging and huntington's disease. *J. Appl. Physiol.* **1997**, *82*, 262–269.
56. Bhattacharya, J.; Edwards, J.; Mamelak, A.N.; Schuman, E.M. Long-range temporal correlations in the spontaneous spiking of neurons in the hippocampal–amygdala complex of humans. *Neuroscience* **2005**, *131*, 547–555.
57. Tapanainen, J.M.; Thomsen, P.E.; Kober, L.; Torp-Pedersen, C.; Makikallio, T.H.; Still, A.M.; Lindgren, K.S.; Huikuri, H.V. Fractal analysis of heart rate variability and mortality after an acute myocardial infarction. *Am. J. Cardiol.* **2002**, *90*, 347–352.
58. Ho, K.K.; Moody, G.B.; Peng, C.K.; Mietus, J.E.; Larson, M.G.; Levy, D.; Goldberger, A.L. Predicting survival in heart failure case and control subjects by use of fully automated methods for deriving nonlinear and conventional indices of heart rate dynamics. *Circulation* **1997**, *96*, 842–848.
59. Penzel, T.; Kantelhardt, J.W.; Grote, L.; Peter, J.H.; Bunde, A. Comparison of detrended fluctuation analysis and spectral analysis for heart rate variability in sleep and sleep apnea. *IEEE Trans. Biomed. Eng.* **2003**, *50*, 1143–1151.
60. Zhang, X.S.; Roy, R.J.; Jensen, E.W. EEG complexity as a measure of depth of anesthesia for patients. *IEEE Trans. Biomed. Eng.* **2001**, *48*, 1424–1433.
61. Bruhn, J.; Ropcke, H.; Rehberg, B.; Bouillon, T.; Hoefl, A. Electroencephalogram approximate entropy correctly classifies the occurrence of burst suppression pattern as increasing anesthetic drug effect. *Anesthesiology* **2000**, *93*, 981–985.
62. Koskinen, M.; Seppanen, T.; Tong, S.; Mustola, S.; Thakor, N.V. Monotonicity of approximate entropy during transition from awareness to unresponsiveness due to propofol anesthetic induction. *Trans. Biomed. Eng.* **2006**, *53*, 669–675.
63. Pikkujamsa, S.M.; Makikallio, T.H.; Sourander, L.B.; et al. Cardiac interbeat interval dynamics from childhood to senescence: comparison of conventional and new measures based on fractals and chaos theory. *Circulation* **1999**, *100*, 393–399.
64. Lange, S.; Van Leeuwen, P.; Geue, D.; Hatzmann, W.; Grönemeyer, D. Influence of gestational age, heart rate, gender and time of day on fetal heart rate variability. *Med. Biol. Eng. Comput.* **2005**, *43*, 481–486.
65. Moonman, J.R.; Lake, D.E.; Griffin, M.P. Heart rate characteristics monitoring for neonatal sepsis. *IEEE Trans. Biomed. Eng.* **2006**, *53*, 126–132.
66. Vikman, S.; Makikallio, T.H.; Yli-Mayry, S.; Pikkujämsä, S.; Koivisto, A.-M.; Reinikainen, P.; Juhani Airaksinen, K.E.; Huikuri, H.V. Altered complexity and correlation properties of R–R interval dynamics before the spontaneous onset of paroxysmal atrial fibrillation. *Circulation* **1999**, *100*, 2079–2084.
67. Hornero, R.; Aboy, M.; Abasolo, D.; McNames, J.; Goldstein, B. Interpretation of approximate entropy: analysis of intracranial pressure approximate entropy during acute intracranial hypertension. *IEEE Trans. Biomed. Eng.* **2005**, *52*, 1671–1680.
68. Akay, M.; Sekine, N. Investigating the complexity of respiratory patterns during recovery from severe hypoxia. *J. Neural. Eng.* **2004**, *1*, 16–20.
69. Porta, A.; Guzzetti, S.; Montano, N.; Pagani, M.; Somers, V.; Malliani, A.; Baselli, G.; Cerutti, S. Information domain analysis of cardiovascular variability signals: evaluation of regularity, synchronisation and co-ordination. *Med. Biol. Eng. Comput.* **2000**, *38*, 180–188.
70. Mainardi, L.T.; Porta, A.; Calcagnini, G.; Bartolini, P.; Michelucci, A.; Cerutti, S. Linear and non-linear analysis of atrial signals and local activation period series during atrial fibrillation episodes. *Med. Biol. Eng. Comput.* **2001**, *39*, 249–254.

Distal Protection Devices

Thomas Tu
Roger Laham

Cardiovascular Division, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, Massachusetts, U.S.A.

INTRODUCTION

Angioplasty of atherosclerotic lesions can result in the liberation of microscopic and macroscopic debris. Recent evidence suggests this embolization occurs more frequently than was previously thought. The particulate matter is comprised of plaque components, platelets, thrombus, and vasoconstrictors. Large emboli may obstruct the epicardial vessel and can generally be treated by traditional angioplasty techniques. Smaller material that is too large to pass through the capillary network of the distal bed (15–100 micron) becomes lodged in the microvasculature. Embolization leads to microvascular dysfunction by a variety of mechanisms, including occlusion, vasoconstriction, and platelet aggregation. The angiographic appearance of this phenomenon is described as no reflow, which is defined as reduced coronary flow despite the lack of epicardial stenosis.

Embolization and no reflow can result in periprocedural ischemia and microvascular infarction, which is detectable by cardiac enzyme elevation, impaired wall motion, and abnormalities on MRI. Percutaneous interventions associated with impaired flow and periprocedural enzyme elevation are associated with longer hospital stay and increased mortality.

OVERVIEW

Recent advances in percutaneous coronary and peripheral revascularization have focused on the problem of distal embolization during interventions, a process which often results in the phenomenon of no-reflow during coronary intervention with resultant microvascular dysfunction and major adverse cardiac events and cerebral and peripheral ischemia during peripheral and carotid interventions.^[1–7] The use of glycoprotein IIb/IIIa antagonists has been shown to reduce periprocedural infarcts and also improves long-term outcomes.^[10–12] This is thought to be due in part to a reduction in ischemic complications. However, in the subset of degenerated vein grafts, the benefit is not evident, perhaps because the magnitude of

embolization is too great to be mitigated by antiplatelet therapy.^[13]

Angioplasty technique has also been modified in response to concern about microembolization. The use of high pressure inflation to maximally expand stents has been widely adopted and is effective at reducing subsequent restenosis. It is also associated with a higher incidence of embolization, which is likely due to increased vessel trauma. When approaching lesions at high risk of embolization, many operators choose to use less traumatic self-expanding stents and/or foregoing high-pressure postdilation.

Despite these advances, there is still a significant incidence of periprocedural infarction after percutaneous intervention. A number of devices have been developed in order to prevent the distal embolization associated with angioplasty. Together, they are known as distal protection devices.

Currently, the principal applications of distal protection include saphenous vein graft and carotid artery angioplasty.

Saphenous Vein Grafts (SVG)

Saphenous vein grafts are the most common conduits used for aorto-coronary bypass surgery. They exhibit accelerated intimal hyperplasia and atherosclerosis when compared to native coronary arteries and arterial conduits.^[14] Ten years after surgery, 40% of SVGs are occluded and even more have developed severe stenosis.^[15] The plaque is typically diffuse and complex with large amounts of friable atheroma and intraluminal thrombus. Traditional percutaneous intervention on these lesions was often complicated by distal embolization, no reflow, and decreased procedural success.^[8,9] Factors that have been associated with worsened acute success include grafts older than three to four years, multiple lesions or diffuse disease, small graft diameter, and presence of intraluminal thrombus.^[14]

Carotids

Percutaneous carotid artery intervention for carotid stenosis has been performed successfully and is

Supported in part by NIH grants MO1-RR01032 and HL 63609 (RJL).

currently being investigated as a possible alternative for carotid endarterectomy. The major complication of the procedure is stroke, which occurs in 1–3% of the cases, presumably from embolization.^[16] Transcranial Doppler studies have demonstrated a high incidence of microembolization during percutaneous intervention.^[17–19] The clinical significance of this phenomenon is not clear. However, it is likely that substantial reduction of embolization may reduce the incidence of periprocedural stroke as well as more subtle late neurological deficits. There is significant interest in distal protection as the complication rates of carotid artery stenting become similar to surgical endarterectomy.

DEVICES TO REDUCE DISTAL EMBOLIZATION

There are three major categories of devices that have been developed to address the problem of distal embolization: balloon occlusion devices, filter devices, and flow reversal devices (Table 1).

Balloon occlusion devices utilize a compliant balloon mounted on a coronary guidewire. The wire is directed across the target lesion, and the balloon is inflated until coronary blood flow has stopped. The intervention is performed over the balloon wire, and the liberated particles are unable to migrate into the distal vascular bed. These particles are then removed by various aspiration techniques.

Filter devices employ a porous, expandable filter mounted on a coronary guidewire. The collapsed filter is directed across the target lesion. The filter is then deployed, and the coronary intervention can then be performed over the filter wire. Anterograde perfusion is maintained while any debris larger than the filter pore size is trapped. At the completion of the procedure, the filter is retracted into a recovery sheath. The filter with its captured debris is then removed from the body.

Table 1 Types of distal protection devices

Balloon occlusion devices

- Guardwire (Medtronic)
- Triactiv (Kensey-Nash)

Filter devices

- Angioguard (Cordis)
- Filterwire EX (Boston Scientific)
- Mednova Neuroshield (MedNova)
- AccUNET (Guidant)

Flow reversal devices

- Parodi Anti-Emboli System (PAES) (Arteria Medical Science)

Flow reversal devices have been developed particularly for percutaneous intervention on the carotid artery. They occlude anterograde flow proximal to the lesion, which results in a reversal of flow in the target vessel. Liberated particles migrate away from the distal vascular bed and are aspirated by the device.

PERCUSURGE GUARDWIRE (MEDTRONIC)

Device Description

The PercuSurge Guardwire Temporary Occlusion and Aspiration System consists of four principal components: The Guardwire, the Microseal Adapter, the Export Catheter, and the EZ-Flator Inflation Device. The Guardwire design utilizes a 0.014-inch hollow-core wire that is connected to an elastomeric balloon near the distal end. The wire has a radiopaque, shapeable distal tip that facilitates lesion crossing. The Microseal Adapter allows access to the hypotube inside the wire. The EZ-Flator Inflation Device is connected to the Guardwire via the Microseal adapter (Fig. 1A). The balloon can be inflated to a diameter of 3–6 mm using the graded inflation device (Fig. 1B). The Export Catheter consists of a 5 French monorail system with a 1 mm internal diameter central lumen. The catheter is used to aspirate debris from the static blood column formed by the occlusion balloon.

Procedure

A 7 French or larger guiding catheter is used to cannulate the SVG ostium. The Guardwire is prepared by inflating and deflating the balloon with the Inflation Device. The wire is then directed across the target stenosis in a similar fashion as a standard coronary guidewire. The occlusive balloon is positioned 2 to 10 cm distal to the lesion. The proximal end of the Guardwire is placed into the Microseal Adapter. Dilute radiographic contrast is infused until anterograde perfusion has ceased. It is possible to trap a column of contrast in the vessel to allow easy visualization of the lesion during occlusion. The lumen of the Guardwire is then sealed, which allows removal of the Microseal Adapter and Inflation Device. Treatment of the lesion can then proceed in standard fashion by using the Guardwire to direct angioplasty and stenting. Any liberated plaque or thrombus tends to remain in the static column of blood in the occluded vessel. The Guardwire is then used to direct the Export Catheter up to the occlusion balloon. A syringe is used to aspirate 20–40 cc of blood through the export catheter lumen, which removes the trapped debris. The Guardwire is reattached to the Inflation Device using

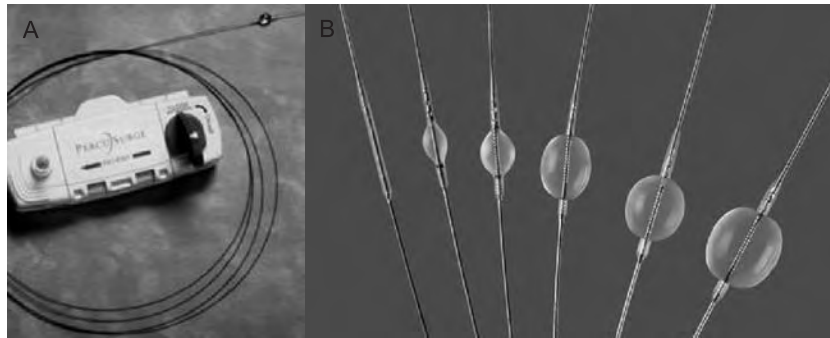


Fig. 1 The Percutaneous Distal Protection system is the only FDA-approved device in the United States. It consists of an Inflation Device with a Microseal Adapter (A) that connects to the Guardwire (B). The balloon can be inflated to a diameter of 3–6 mm using the graded inflation device.

the Microseal Adapter, and the balloon is deflated and thereby restores antegrade flow.

In the hands of skilled operators, the entire procedure can be completed with an occlusion time of three to four minutes. Depending on the size and importance of the vascular territory involved, the patient may have severe ischemia during occlusion, as evidenced by angina, ST-segment changes, hypotension, or arrhythmias. Occlusion time can be reduced by planning the procedure well and performing the intervention quickly. If several balloon/stent treatments are required, the procedure can be performed in stepwise fashion with multiple occlusion and aspiration cycles. Occasionally, the patient cannot tolerate even a short period of balloon occlusion, and a filter device should be substituted.

Clinical Trials

The Guardwire system is the only distal protection device approved by the U.S. Food and Drug Administration (FDA) for treatment of saphenous vein graft stenoses. The initial clinical results were published in the SAFE registry (Saphenous Vein Graft Angioplasty Free of Emboli).^[20] One hundred three patients underwent percutaneous vein graft intervention using the Guardwire balloon occlusion system. Macroscopically visible red and/or yellow debris was extracted in 91% of patients. By core lab analysis, normal coronary flow was present in 98.9% of grafts postintervention versus 83.5% before intervention. No patient developed angiographic evidence of no reflow or distal embolization. Postprocedural infarcts, defined by creatine phosphokinase MB isozyme levels elevated to $>3\times$ normal, were seen in only five patients (5%), and 97 patients (94%) were free of major adverse events at 30 days.

The pivotal SAFER (Saphenous Vein Graft Free of Emboli Randomized) trial^[21] was the first multicenter randomized trial to evaluate use of a distal embolic protection device during stenting of saphenous vein graft lesions. Four hundred six patients were randomly

assigned to stent placement over the shaft of the distal protection device, and 395 were assigned to stent placement over a conventional 0.014-inch angioplasty guidewire (control group). The primary end point—a composite of death, myocardial infarction, emergency bypass, or target lesion revascularization by 30 days—was observed in 65 patients (16.5%) assigned to the control group and 39 patients (9.6%) assigned to the embolic protection device ($P=0.004$). This 42% relative reduction in major adverse cardiac events was driven by myocardial infarction (8.6% versus 14.7%, $P=0.008$) and no-reflow phenomenon (3% versus 9%, $P=0.02$). Clinical benefit was seen even when platelet glycoprotein IIb/IIIa receptor blockers were administered (61% of patients) with composite end points occurring in 10.7% of protection device patients versus 19.4% of control patients ($P=0.008$).

These two studies demonstrated that use of this distal protection device during stenting of venous graft lesions was associated with a highly significant reduction in major adverse events compared with stenting over a conventional angioplasty guidewire. The SAFER study highlighted the importance of distal embolization in causing major adverse cardiac events and the value of embolic protection devices in preventing such complications.

There are several ongoing clinical studies that are evaluating the use of the Guardwire system for carotid angioplasty. The CAFÈ (Carotid Angioplasty Free of Emboli) registry will look at the 30-day success rate of 100 patients who undergo carotid stenting with Guardwire distal protection. MAVERIC (Medtronic AVE Self-Expanding Carotid Stent System with Distal Protection in the Treatment of Carotid Stenosis) is a multicenter registry of 50 patients at high risk of carotid endarterectomy who will undergo carotid stenting with a new stent system and Guardwire distal protection. SHELTER (Carotid Stenting of High-Risk Patients: Extracranial Lesion Trial with Emboli Removal) is a registry of 630 patients with one-year follow-up after carotid stenting with a Monorail Carotid Wall-stent and Guardwire distal protection.

TRIACTIV (KENSEY NASH)

Device Description

The Triactiv device is a balloon occlusion-type distal protection device that acquired its CE Mark in January 2002, which gives approval for use in Europe. It consists of a hollow 0.014-inch stainless steel guidewire with a distal elastomeric balloon. The wire is used to cross the lesion in a similar fashion to a coronary guidewire. A calibrated syringe filled with CO₂ is attached to the proximal end of the wire and is used to inflate the balloon to a diameter of 3.0–5.0 mm. Once vessel occlusion is achieved, the wire is used to guide the stent to the target lesion. After vessel treatment, a 3 French sump catheter is attached to the side of the guidewire using a spiral groove monorail. The sump catheter is advanced to the distal occlusion balloon (Fig. 2). A peristaltic pump is used to infuse saline into the vessel while aspiration of debris is performed through the coronary guide catheter. The active infusion/sumping of clean saline may provide enhanced clearance of embolic debris. The balloon is then deflated, and anterograde perfusion is restored. If postdilation or additional stenting is required, the balloon can be reinflated, and the procedure can be repeated.

Clinical Trials

The Triactiv device is currently being evaluated in a clinical trial called PRIDE (Protection During Saphenous Vein Graft Intervention to Prevent Distal

Embolization). In this U.S. trial, up to 800 patients at up to 50 sites who undergo vein graft intervention will be randomized to standard-of-care versus intervention with Triactiv distal protection. Early results of the trial suggest a low major adverse cardiac event rate and led to approval of its CE Mark.^[22] The trial is expected to be completed sometime in the spring of 2004.

ANGIOGUARD (CORDIS)

Device Description

The Angioguard device utilizes a polyurethane filter mounted on a nitinol skeleton. The filter is attached distally to a floppy wire tip and proximally to a 300 cm, 0.014-inch, medium-support guidewire (Fig. 3). The closed filter, contained in a delivery sheath with a diameter of 0.061 to 0.065 inch and a maximum profile of 4 F to 4.5 F, is advanced through the lesion. A second-generation device (Angioguard XP) has crossing profiles of 3.2–4.1 F. The filter is opened in the target vessel distal to the lesion by removal of the delivery sheath. A filter 0.5 to 1.0 mm larger than the visually estimated distal vessel diameter is chosen to ensure complete apposition of the filter to the vessel wall. Sizes range from 4–8 mm. The filter has numerous, laser-drilled 100 micron holes that allow anterograde passage of red blood cells while the filter is deployed. Liberated debris is trapped by the filter, which is then collapsed by advancement of a retrieval sheath and removed from the artery.

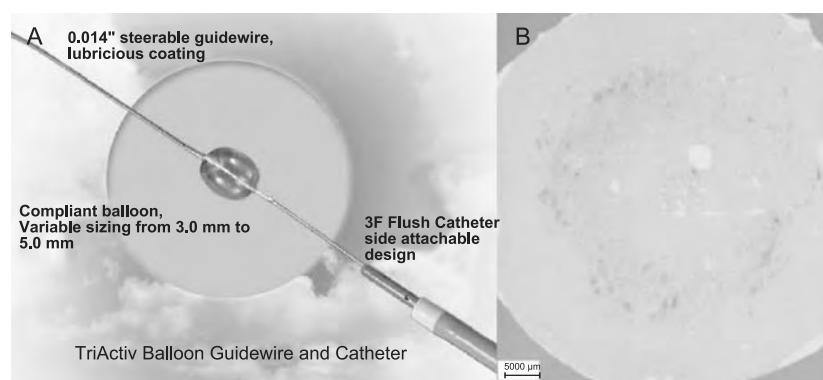


Fig. 2 The Triactive system is currently in clinical trials in the United States and Europe. (A) It consists of a hollow 0.014-inch stainless steel guidewire with a distal elastomeric balloon. A calibrated syringe filled with CO₂ is attached to the proximal end of the wire and is used to inflate the balloon to a diameter of 3.0–5.0 mm. Once vessel occlusion is achieved, the wire is used to guide the stent to the target lesion. After vessel treatment, a 3 French sump catheter is attached to the side of the guidewire using a spiral groove monorail. The sump catheter is advanced to the distal occlusion balloon. A peristaltic pump is used to infuse saline into the vessel while aspiration of debris is performed through the coronary guide catheter. (B) Atherosclerotic debris retrieved with the Triactive system underscores the need for distal protection.



Fig. 3 The Angioguard device utilizes a polyurethane filter mounted on a nitinol skeleton. The filter is attached distally to a floppy wire tip and proximally to a 300 cm, 0.014-inch, medium-support guidewire. The closed filter, contained in a delivery sheath with a diameter of 0.061 to 0.065 inch and a maximum profile of 4 F to 4.5 F, is advanced through the lesion. A second-generation device (Angioguard XP) has crossing profiles of 3.2–4.1 F. The filter is opened in the target vessel distal to the lesion by removal of the delivery sheath. Liberated debris is trapped by the filter, which is then collapsed by advancement of a retrieval sheath and removed from the artery.

Clinical Trials

Initial feasibility and safety data demonstrated a high rate of technical success using the Angioguard device in saphenous vein grafts.^[23] SAPPHERE (Stenting and Angioplasty with Protection in Patients at High Risk for Endarterectomy) was a pivotal study comparing stenting with distal protection to traditional surgical carotid endarterectomy for high-risk patients. The results of this combination registry and randomized control trial were presented at the November 2002 American Heart Association meeting. There were 307 patients in the randomized trial and 416 patients followed in the registry arm enrolled from 29 centers. The Nitinol Precise Stent and Angioguard distal protection device (both from Cordis) were used for the intervention. Entry criteria included symptomatic patients with more than 70% stenosis or asymptomatic patients with over 80% stenosis. In general, high-risk characteristics that allowed enrollment included severe heart failure, unstable angina or recent myocardial infarction, severe lung disease, concomitant coronary and carotid disease, open heart surgery planned within six weeks, or technically difficult surgical anatomy.

Results of the trial showed a significant reduction in major clinical events at 30 days favoring stenting over surgery. Stroke, myocardial infarction, or death occurred in 5.8% of the stent patients versus 12.6% of the surgical patients ($p=0.047$). Thirty-day

complication rates were equivalent between the two groups for transient ischemic attack and bleeding. The surgical group had a 5.3% incidence of cranial nerve injury, which is compared to 0% with stenting. There were 409 patients in the stent registry compared to seven in the surgical registry, which indicates a strong preference for the less invasive intervention.

GUARD (Saphenous Vein Graft Intervention Using Angioguard for Reduction of Distal Embolization) is a randomized safety and efficacy trial for vein graft interventions. Between 300–700 patients will be randomized to percutaneous intervention with or without use of the Angioguard device. The primary outcome will be major adverse clinical events at 30 days.

NEUROSHIELD (MEDNOVA)

Device Description

The MedNova NeuroShield is a filter that is mounted on the distal tip of a 0.014-inch guidewire. The filter contains a preshaped nitinol expansion system that facilitates fluoroscopic visualization, accurate deployment, and wall apposition. The filter guidewire is placed within the delivery catheter and is passed through the target lesion. The delivery catheter is then withdrawn, and the filter is deployed. After the vascular intervention, a retrieval catheter is advanced to envelop the filter. The entire device, including the trapped embolic contents, is then withdrawn.

Clinical Trials

Preliminary clinical experience with Neuroshield device for use in carotid stenting has been published.^[24] Between September 1999 and July 2001, a total of 162 patients (164 hemispheres) underwent carotid stenting with filter protection according to prospective protocols evaluating the filter system at three institutions. Angiographic success was achieved in 162 of the procedures (99%), and filter placement was successful in 154 (94%) procedures. Carotid access was unsuccessful in two cases (1%) and filter placement in eight cases (5%). Of the latter, five procedures were completed with no protection, and three were completed using alternative protection devices. On an intention-to-treat basis, the overall combined 30-day rate of all stroke and death victims was 2% (four events: two minor strokes and two deaths). This includes one minor stroke in a patient with failed filter placement, and stenting was completed without protection. There was one cardiac arrhythmic death and one death from hyperperfusion-related intracerebral hemorrhage. There were no major embolic strokes. From these studies, it appears that carotid artery stenting with Mednova filter

protection is technically feasible and safe. Larger comparative studies are now underway.

Mednova has developed a similar device for use in vein graft interventions called the Cardioshield. It is currently being evaluated in the CAPTIVE study (CardioShield Application Protects during Transluminal Intervention of Vein Grafts by reducing Emboli), which is a randomized trial of 800 patients undergoing vein graft intervention. The primary end point will be major adverse clinical events at 30 days.

FILTERWIRE EX (BOSTON SCIENTIFIC)

Device Description

The Filterwire EX (Boston Scientific Corporation) is another filter-type device currently undergoing clinical investigation. It consists of a nitinol filter loop mounted on a 0.014-inch guidewire. The filter material is made of polyurethane and has 80-micron pores to permit anterograde perfusion during deployment (Fig. 4). The filter rotates freely on the end of the guidewire. The device is housed within a 1.3 mm (0.052 inch) delivery sheath. The loop has an off-center, fish-mouth design that facilitates complete apposition for vessels 3.5–5.5 mm in diameter. After the intervention, the device is retracted into the delivery sheath. The trapdoor action of the filter loop allows retention of the embolic debris during retrieval.

Clinical Trials

Early clinical work has demonstrated feasibility and safety in using the device.^[25] From May 16, 2000, to October 26, 2000, a total of 36 lesions in 35 patients were protected using the Filterwire EX system. Of the 36 lesions, 22 (61%) were located in native coronaries, and 14

(39%) were located in SVG. The Filterwire EX device was delivered, deployed, and removed successfully in 33 procedures. In two procedures, the device could not easily cross extraordinarily angled and tight lesions, and, in one other procedure, the device could not be deployed because of abnormal distal anatomy of the vessel. In every procedure in which the Filterwire EX was deployed, retrieval was successful. Embolic material was obtained in 71% of these cases, and only 2 of 31 patients (6%) meeting eligibility criteria experienced major adverse cardiac events. This appears to be an improved rate in comparison to that expected without protection.

The FIRE (Filterwire During Transluminal Intervention of Saphenous Vein Grafts) trial is enrolling up to 800 patients at 70 sites. It is a safety and efficacy trial of treatment with the Filterwire during percutaneous intervention of saphenous vein grafts.

The BEACH (Boston Scientific/EPI: A Carotid Stenting Trial for High-Risk Surgical Patients) Investigation began in February 2002 and is enrolling up to 776 patients in 40 sites. It is investigating the use of the Carotid Wallstent Monorail Endoprosthesis and Filterwire EX embolic protection in carotid stenting. The primary end point is 30-day combined major adverse clinical events.

The CABERNET (Carotid Artery Revascularization Using the Boston Scientific EPI Filterwire EX and the EndoTex Nexstent) is an international, multicenter study of the safety and efficacy of the self-expanding NexStent used in conjunction with the Filterwire for high-risk carotid angioplasty.

ACCUNET (GUIDANT)

Device Description

The AccUNET device consists of a polyurethane filter supported by a nitinol basket frame. The pore size is



Fig. 4 The Filterwire EX of a nitinol filter loop mounted on a 0.014-inch guidewire. The filter material is made of polyurethane and has 80-micron pores to permit anterograde perfusion during deployment. The filter rotates freely on the end of the guidewire. The device is housed within a 1.3 mm (0.052 inch) delivery sheath. The loop has an off-center, fish-mouth design that facilitates complete apposition for vessels 3.5–5.5 mm in diameter. After the intervention, the device is retracted into the delivery sheath. The trapdoor action of the filter loop allows retention of the embolic debris during retrieval.

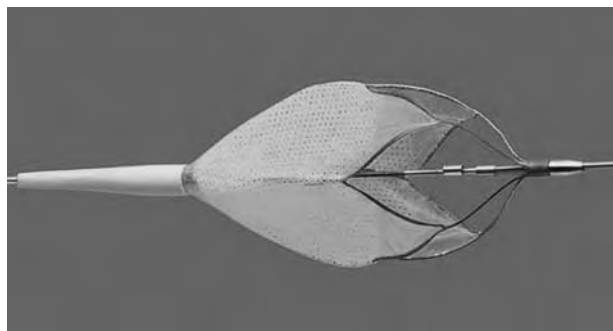


Fig. 5 The Accunet device consists of a polyurethane filter supported by a nitinol basket frame. The pore size is ≤ 120 microns. It is available in sizes ranging from 4–8 mm in diameter. The filter is mounted on an ACS Hi-Torque Balance Heavyweight coronary wire platform, which is 300 cm in length and 0.014 inch in diameter. The sheathed filter is directed beyond a target vascular lesion and is deployed by withdrawal of the delivery sheath. The filter is then collapsed and retracted into the sheath for removal.

≤ 120 microns (Fig. 5). It is available in sizes ranging from 4–8 mm in diameter. The filter is mounted on an ACS Hi-Torque Balance Heavyweight coronary wire platform, which is 300 cm in length and 0.014 inches in diameter. The sheathed filter is directed beyond a target vascular lesion and is deployed by withdrawal of the delivery sheath. The intervention is then performed over the 0.014-inch wire. The filter is then collapsed and retracted into the sheath for removal.

Clinical Trials

The ACCUNET Trial is a 50-patient registry designed to evaluate the safety and efficacy of the Acculink carotid stent used in conjunction with the Accunet distal protection device. Outcomes will be the acute success of the device as well as a 30-day combined outcome of stroke, death, and myocardial infarction.

The device is also being evaluated in CREST (Carotid Revascularization Endarterectomy versus Stenting Trial), which is an NIH-funded prospective, randomized, multicenter clinical trial of carotid endarterectomy (CEA) versus carotid artery stenting (CAS). CREST includes a lead-in phase for credentialing of interventionalists who have previously been trained and evaluated in CAS. Patients with transient ischemic attack, amaurosis fugax, or non-disabling stroke within 180 days of randomization and ipsilateral carotid stenosis $\geq 50\%$ (defined as $\geq 70\%$ by ultrasound or $\geq 50\%$ by angiography) are eligible for randomization. Follow-up includes clinic visits at 1, 6, and 12 months and then every 6 months and phone contact every 3 months for study duration. Primary

end points are a composite of stroke, myocardial infarction, or death within 30 days and stroke ipsilateral to the study artery after 30 days. A total of 2500 patients will be randomized. After review of early device data, the CREST Executive Committee approved the addition of an embolic protection device in the trial. The CREST Embolic Protection Subcommittee analyzed the clinical data from the study of the ACCUNET Embolic Protection System, and the committee's recommendation to the Executive Committee was approved. In the fall of 2001, a protocol revision to include the ACCUNET Embolic Protection Device was approved by the NIH and the FDA. To date, 52 centers have been selected to participate in CREST, and 21 centers are approved to enroll participants. As of the end of November 2001, 79 participants have been enrolled.

PARODI ANTI-EMBOLI SYSTEM

Device Description

The Parodi Anti-Emboli System (PAES) is an 80-cm guiding catheter with a distal occlusion balloon. The balloon is 7 mm long and can be inflated up to 10 mm. It is placed proximal to the lesion and inflated, which thereby reverses flow in the target vessel. In the carotid artery, the balloon is used to occlude the common carotid, which reverses blood flow in the internal carotid artery. A second occlusion balloon is placed in the external carotid artery, which prevents flow of embolic material from the internal to external carotid artery. This system establishes distal protection without having to cross the target lesion and thereby avoids potential embolization caused by disruption of the plaque during device deployment.

Clinical Trials

To investigate the safety and efficacy of the Parodi antiembolism system (PAES), seven centers participated in a nonrandomized, prospective trial of carotid angioplasty and stenting.^[26,27] PAES protection was used in 30 patients with 15 symptomatic ($>70\%$) and 15 asymptomatic ($>80\%$) stenotic carotid arteries. Safety was defined as achieving sufficient brain oxygenation during flow reversal as determined by level of awareness and motor control. The presence of new or enhanced neurological deficits and death were end points. Performance was based on angiographic evidence of successful retrograde flow. The PAES was positioned in all 30 patients, but technical error and access-related difficulties prevented establishment of reversed flow in two. Among the 28 (93%) patients

treated under PAES protection, one patient developed aphasia after flow reversal, which necessitated balloon deflation between subsequent stages of the procedure. Three other adverse events included one case of bradycardia and two cases of hypotension with dysarthria and facial paresis in one and temporary loss of consciousness in the other. All events resolved with appropriate therapy, and there was no change from baseline in the neurological status or brain scans at 24 hours. There were no strokes or neurological deficits at 30 days. From this early experience, it appears that the PAES can be a safe and effective means of providing protection from embolic complications during carotid stenting. This finding needs to be confirmed with larger randomized clinical trials.

ADVANTAGES AND DISADVANTAGES OF EACH DEVICE

Balloon occlusion devices have been used longer than the other distal protection systems. The most experience has been with the Medtronic Guardwire system. There are many advantages of this type of protection device. The balloons typically allow easy and complete occlusion of a vessel. Particles of any size are trapped in the static column of blood. The balloon devices have smaller crossing profiles and greater flexibility than filters and therefore are generally easier to deliver and withdraw.

Disadvantages of the balloon occlusion technique are related to ischemia due to complete occlusion of blood flow. For most interventions, the procedure time is short enough that the ischemia is well-tolerated. In a large-territory vessel or in one without collateral circulation, the occlusion may not be tolerated. Because aspiration of the debris is required, the procedure has multiple steps and is technically challenging. These procedures are best performed with two experienced operators. Occlusion prevents angiographic visualization during the intervention. Typically, operators will either use noncontrast landmarks or attempt to trap a column of contrast with the occlusion balloon. Balloon and stent lengths tend to be longer to ensure lesion coverage given the time and visualization constraints. In internal carotid artery intervention, there is a risk of embolization to the external carotid, which is not protected by the balloon.

Filter devices are advantageous because they maintain antegrade perfusion during the intervention. The procedure requires fewer steps because aspiration is not required. Preservation of antegrade flow also allows continual angiographic visualization. The bulkiness of the current designs limits their ability to be delivered easily, especially in tortuous vessels. The larger crossing profile may be associated with embolization during delivery of the protection device.

The filters trap the largest particles but allow smaller particles to pass through the pores. It is unclear whether the smaller emboli are clinically significant. Large amounts of liberated plaque can overwhelm a filter, which converts it to an occlusion device.

The flow reversal devices are advantageous because the distal protection is established prior to placing equipment across the target lesion. Transcranial Doppler studies have demonstrated evidence of microemboli just with passage of the distal protection filter or balloon. Difficult and tortuous anatomy does not affect the device as it does not have to be placed across the lesion. The flow reversal can remove particles of any size. The intervention can be performed over any choice of guidewire rather than being limited to the characteristics of the distal occlusion platform.

The chief disadvantage of the Parodi device is that it requires an 11 French arterial sheath. Its initial use was associated with a higher incidence of arterial access site complication. Some patients may not tolerate reversal of flow, particularly if they have contralateral carotid disease. The device is somewhat complicated, and use in the carotid requires a second occlusion balloon for the external carotid branch.

MECHANISMS OF FAILED DISTAL PROTECTION

Distal protection devices can fail by a number of different mechanisms. Device-related vessel injury can be caused by the occlusion balloon or filter, which results in vasospasm, dissection, or thrombosis. The bulkier devices, particularly filters, can become entrapped in the freshly placed stent. Balloon occlusion and flow reversal devices can be limited by prolonged ischemia. Filter devices that become occluded with debris are also at risk for this complication. Distal protection can be lost for a number of reasons. The occlusion balloon may rupture or move during wire exchange, which leads to a loss of protective effect. The export catheter may not clear the embolic material completely, which results in embolization when the occlusion balloon is deflated. Filters may incompletely occlude to the vessel wall, which allows emboli to travel through an unfiltered channel. The pore size may allow passage of small but clinically important particles. Incompletely trapped debris may be liberated from the filter during retrieval.

EVOLUTION OF CURRENT DISTAL PROTECTION TECHNOLOGY

Current distal protection technology can be improved to handle an ever-increasing range of clinical situations. All devices can be improved by reducing their crossing

profile and improving steerability. This will improve their deliverability in severely stenotic or tortuous vessels. It may also mitigate embolization prior to device deployment. The devices can also be improved to reduce the local injury related to balloon/filter deployment. Protection of vessels with complex anatomy including bifurcations, large proximal sidebranches, ostial, and distal lesions is particularly challenging. Device-sizing in totally occluded vessels is also difficult. The problem of prolonged ischemic time can be addressed by simplifying the procedure to allow more rapid intervention. Additional research with filter devices can help determine the optimal pore size to maintain antegrade perfusion without sacrificing protective effect. Finally, debris retrieval can be improved with better aspiration techniques and filter retrieval systems.

FUTURE APPLICATIONS OF DISTAL PROTECTION

It is becoming apparent that distal embolization occurs more frequently than was previously realized. As the field of percutaneous vascular interventions grows, so do the potential applications for distal protection devices.

Acute Myocardial Infarction

Acute myocardial infarction is generally caused by rupture of an atherosclerotic plaque with subsequent platelet deposition and thrombus formation. Lesions are at high risk of embolization during intervention. Use of distal protection devices in native coronary arteries is somewhat problematic due to the inability to protect major sidebranches. The use of distal protection is currently being evaluated in the EMERALD trial in which acute myocardial infarction patients are randomized to conventional angioplasty versus use of the Guardwire distal protection system.

Renal Artery Intervention

Percutaneous intervention of the renal arteries is rapidly supplanting surgery for the treatment of renal artery stenosis. The procedure is associated with a substantial risk of postprocedural decline in renal function. Many postulate that this is due to distal embolization. While the scope of this phenomenon is not yet known, some operators are using distal protection devices in an attempt to minimize this complication.

Superficial Femoral Artery Intervention

Revascularization of the superficial femoral artery (SFA) for lower extremity claudication has many

similarities to vein graft intervention. The vessel is long with generally diffuse disease and is without major sidebranches. Distal embolization is uncommon, but it is likely an underrecognized phenomenon. For this reason, some operators propose the use of distal protection devices during SFA angioplasty.

CONCLUSION

Distal protection devices have been successful in reducing the embolic complications associated with angioplasty of high-risk lesions. The three types of devices being developed have differing properties that can be tailored to individual clinical situations. The majority of clinical experience thus far has been with the balloon occlusion type of devices. The filter and flow-reversal devices are not yet approved in the United States, but they are actively being investigated. Distal protection is rapidly becoming standard-of-care for vein graft and carotid interventions. As the devices continue to evolve and improve, they will be useful for an ever-expanding array of percutaneous procedures.

ARTICLES OF FURTHER INTEREST

Cardiac Patch Engineering; Shape Memory Metals; Thrombosis; Vascular Grafts

REFERENCES

1. Topol, E.J.; Yadav, J.S. Recognition of the importance of embolization in atherosclerotic vascular disease. *Circulation* **2000**, *101*, 570–580.
2. Piana, R.N.; Paik, G.Y.; Moscucci, M.; Cohen, D.J.; Gibson, C.M.; Kugelmass, A.D.; Carrozza, J.P., Jr.; Kuntz, R.E.; Baim, D.S. Incidence and treatment of ‘no-reflow’ after percutaneous coronary intervention. *Circulation* **1994**, *89*, 2514–2518.
3. Webb, J.G.; Carere, R.G.; Virmani, R.; Baim, D.S.; Teirstein, P.S.; Whitlow, P.L.; McQueen, C.; Kolodgie, F.D.; Buller, E.; Dodek, A.; Mancini, G.B.; Oesterle, S.N. Retrieval and analysis of particulate debris after saphenous vein graft intervention. *J. Am. Coll. Cardiol.* **1999**, *32*, 468–475.
4. Hori, M.; Gotoh, K.; Kitakaze, M.; Hori, M.; Gotoh, K.; Kitakaze, M.; Iwai, K.; Iwakura, K.; Sato, H.; Koretsune, Y.; Inoue, M.; Kitabatake, A.; Kamada, T. Role of oxygen-derived free radicals in myocardial edema and ischemia in coronary microvascular embolization. *Circulation* **1991**, *84*, 828–840.
5. Abbo, K.M.; Dooris, M.; Glazier, S.; O’Neill, W.W.; Byrd, D.; Grines, C.L.; Safian, R.D. Features and outcome of no-reflow after percutaneous coronary intervention. *Am. J. Cardiol.* **1995**, *75*, 778–782.

6. Koch, K.C.; vom Dahl, J.; Kleinhans, E.; Koch, K.C.; vom Dahl, J.; Kleinhans, E.; Klues, H.G.; Radke, P.W.; Ninnemann, S.; Schulz, G.; Buell, U.; Hanrath, P. Influence of a platelet GPIIb/IIIa receptor antagonist on myocardial hypoperfusion during rotational atherectomy as assessed by myocardial Tc-99m sestamibi scintigraphy. *J. Am. Coll. Cardiol.* **1999**, *33*, 998–1004.
7. Ambrosio, G.; Weissman, H.F.; Mannisi, J.A.; Becker, L.C. Progressive impairment of regional myocardial perfusion after initial restoration of postischemic blood flow. *Circulation* **1989**, *80*, 1846–1861.
8. Hong, M.K.; Mehran, R.; Dangas, G.; Mintz, G.S.; Lansky, A.J.; Pichard, A.D.; Kent, K.M.; Satler, L.F.; Stone, G.W.; Leon, M.B. Creatine kinase-MB enzyme elevation following successful saphenous vein graft intervention is associated with late mortality. *Circulation* **1999**, *100*, 2400–2405.
9. Califf, R.M.; Abdelmeguid, A.E.; Kuntz, R.E.; Popma, J.J.; Davidson, C.J.; Cohen, E.A.; Kleiman, N.S.; Mahaffey, K.W.; Topol, E.J.; Pepine, C.J.; Lipicky, R.J.; Granter, C.B.; Harrington, R.A.; Tardiff, B.E.; Crenshaw, B.S.; Bauman, R.P.; Zuckerman, B.D.; Chaitman, B.R.; Bittl, J.A.; Ohman, E.M. Myonecrosis after revascularization procedures. *J. Am. Coll. Cardiol.* **1998**, *31*, 241–251.
10. Lincoff, A.M.; Califf, R.M.; Topol, E.J. Platelet glycoprotein IIb/IIIa receptor blockade in coronary artery disease. *J. Am. Coll. Cardiol.* **2000**, *35*, 1103–1115.
11. Mathew, V.; Grill, D.E.; Scott, C.G.; Grantham, J.A.; Ting, H.H.; Garratt, K.N.; Holmes, D.R., Jr. The influence of abciximab use on clinical outcome after aortocoronary vein graft interventions. *J. Am. Coll. Cardiol.* **1999**, *34*, 1163–1169.
12. Mak, K.H.; Challapalli, R.; Eisenberg, M.J.; Anderson, K.M.; Califf, R.M.; Topol, E.J. Effect of platelet glycoprotein IIb/IIIa receptor inhibition on distal embolization during percutaneous revascularization of aortocoronary saphenous vein grafts. EPIC Investigators. Evaluation of IIb/IIIa platelet receptor antagonist 7E3 in preventing ischemic complications. *Am. J. Cardiol.* **1997**, *80*, 985–988.
13. Ellis, S.G.; Lincoff, A.M.; Miller, D.; Tchong, J.E.; Kleiman, N.S.; Kereiakes, D.; Califf, R.; Topol, E.J. Reduction in complications of angioplasty with abciximab occurs largely independently of baseline lesion morphology. *J. Am. Coll. Cardiol.* **1998**, *32*, 1619–1623.
14. Liu, M.W.; Douglas, J.S., Jr.; Lembo, N.J.; King, S.B., III. Angiographic predictors of a rise in serum creatinine kinase (distal embolization) after balloon angioplasty of saphenous vein coronary artery bypass grafts. *Am. J. Cardiol.* **1993**, *72*, 514–517.
15. Motwani, J.G.; Topol, E.J. Aortocoronary saphenous vein graft disease: Pathogenesis, predisposition, and prevention. *Circulation* **1998**, *97* (9), 916–931.
16. Roubin, G.S.; New, G.; Iyer, S.S.; Vitek, J.J.; Al-Mubarak, N.; Liu, M.W.; Yadav, J.; Gomez, C.; Kuntz, R.E. Immediate and late clinical outcomes of carotid artery stenting in patients with symptomatic and asymptomatic carotid artery stenosis: A 5-year prospective analysis. *Circulation* **2001**, *103*, 532–537.
17. Jordan, W.D.; Voellinger, D.C.; Doblar, D.D.; Plyushcheva, N.P.; Fisher, W.S.; McDowell, H.A. Microemboli detected by transcranial Doppler monitoring in patients during carotid angioplasty versus carotid endarterectomy. *Cardiovasc. Surg.* **1999**, *7*, 33–38.
18. Rapp, J.H.; Pan, X.M.; Sharp, F.R.; Shah, D.M.; Wille, G.A.; Velez, P.M.; Troyer, A.; Higashida, R.T.; Saloner, D. Atheroemboli to the brain: Size threshold for causing acute neuronal cell death. *J. Vasc. Surg.* **2000**, *32*, 68–76.
19. Ackerstaff, R.G.; Moons, K.G.; van de Vlasakker, C.J.; Moll, F.L.; Vermeulen, F.E.; Algra, A.; Spencer, M.P. Association of intraoperative transcranial doppler monitoring variables with stroke from carotid endarterectomy. *Stroke* **2000**, *31*, 1817–1823.
20. Grube, E.; Schofer, J.; Webb, J.; Schuler, G.; Colombo, A.; Sievert, H.; Gerckens, U.; Stone, G.W.; for the SAFE Trial Study Group. Evaluation of a balloon occlusion and aspiration system for protection from distal embolization during stenting in saphenous vein grafts. *Am. J. Cardiol.* **2002**, *89* (8), 941–945.
21. Baim, D.S.; Wahr, George, B.; Leon, M.B.; Greenberg, J.; Cutlip, D.E.; Kaya, U.; Popma, J.J.; Ho, K.K.L.; Kuntz, R.E.; on behalf of the SAFER Trial Investigators. Randomized trial of a distal embolic protection device during percutaneous intervention of saphenous vein aorto-coronary bypass grafts. *Circulation* **2002**, *105*, 1285–1290.
22. Carrozza, J. In *Early Clinical Trial Results with the Triactiv System*, 13th Annual Symposium on Transcatheter Cardiovascular Therapeutics, Washington, DC, September 11–16, 2001.
23. Grube, E.; Gerckens, U.; Yeung, A.C.; Rowold, S.; Kirchhof, N.; Sedgewick, J.; Yadav, J.S.; Stertzer, S. Prevention of distal embolization during coronary angioplasty in saphenous vein grafts and native vessels using porous filter protection. *Circulation* **2001**, *104* (20), 2436–2441.
24. Al-Mubarak, N.; Colombo, A.; Gaines, P.A.; Iyer, S.S.; Corvaja, N.; Cleveland, T.J.; Macdonald, S.; Brennan, C.; Vitek, J.J. Multicenter evaluation of carotid artery stenting with a filter protection system. *J. Am. Coll. Cardiol.* **2002**, *39* (5), 841–846.
25. Grube, E.; Hauptman, K.; Reifart, N.; Tully, G.; McColl, M. The FilterWire EX embolic protection device for use during percutaneous coronary artery and saphenous vein graft interventions. *Am. J. Cardiol.* **2001**, *88* (Suppl. 5A), 1G–147G.
26. Ohki, T.; Parodi, J.; Veith, F.J.; Bates, M.; Bade, M.; Chang, D.; Mehta, M.; Rabin, J.; Goldstein, K.; Harvey, J.; Lipsitz, E. Efficacy of a proximal occlusion catheter with reversal of flow in the prevention of embolic events during carotid artery stenting: An experimental analysis. *J. Vasc. Surg.* **2001**, *33*, 504–509.
27. Parodi, J.C.; La Mura, R.; Ferreira, L.M.; Mendez, M.V.; Cersosimo, H.; Schonholz, C.; Garelli, G. Initial evaluation of carotid angioplasty and stenting with three different cerebral protection devices. *J. Vasc. Surg.* **2000**, *32*, 1127–1136.

Drug Delivery, Controlled

M.N.V. Ravi Kumar

Department of Pharmaceutics, National Institute of Pharmaceutical Education and Research (NIPER), Mohali, India

A.J. Domb

Department of Medicinal Chemistry and Natural Products, The Hebrew University of Jerusalem, Jerusalem, Israel

INTRODUCTION

Most drugs, whether taken orally or via injection, are delivered to the body systemically. The medication is expected to circulate through the whole body affecting the cells and the organs that are dysfunctional as well as those that are healthy. Systemic drug delivery floods the body with medication and can cause serious side effects. Conventional drug formulations typically provide a prompt release of drug in a bolus form. For drugs that get cleared rapidly from the body, achieving and maintaining the drug concentration within the therapeutic window requires a multiple dosing regimen, often more than once a day. Such an inconvenient dosing regimen can lead to a lack of patient compliance as well as a significant fluctuation in drug levels in the plasma. Recently, several technical advancements have resulted in the development of new technologies capable of controlling the administration of a drug at a targeted site in the body in an optimal concentration-versus-time profile. The term “drug delivery” covers a very broad range of techniques used to get therapeutic agents into the human body. Polymer-based devices are gaining importance due to their capabilities of controlling the rate of drug delivery, sustaining the duration of therapeutic activity, and/or targeting the delivery of drug to a tissue or organ. Controlled drug delivery using polymers can be achieved via various forms depending upon the end use specification, e.g., hydrogels, nanoparticles, microparticles, transdermal, etc. The advancement in this fascinating area of science, which requires interdisciplinary scientific approach, is enormous. Thus, we restrict ourselves to discussing important issues on polymer-based delivery systems and their mechanism of action with representative applications.

WHY DO WE NEED CONTROLLED-RELEASE SYSTEMS?

The ways in which drugs or new biological products are administered have gained increasing attention in the past few decades. Controlled-release systems

provide numerous benefits over the conventional dosage forms. Conventional dosage forms, which are still predominant in pharmaceutical products, are not able to control either the rate of drug delivery or the target area of drug administration or provide an immediate or rapid drug release. This necessitates frequent administration in order to maintain a therapeutic level. As a result, drug concentrations in the blood and tissues fluctuate widely. In conventional dosage forms, the concentration of an administered drug is initially high, which can cause toxic and/or harmful side effects, and then quickly falls down below the minimum therapeutic level with time elapsed. The duration of therapeutic efficacy is dependent upon the frequency of administration, the half-life of the drug, and the release rate of the dosage form. In contrast, controlled-release dosage forms are not only able to maintain therapeutic levels of drug with narrow fluctuations but they also make it possible to reduce the frequency of drug administration, which in turn can minimize side effects.^[1–12] Drug concentration profile in serum depends on the preparation technology, which may generate different release kinetics resulting in different pharmacological and pharmacokinetic responses in the blood or tissues.

Controlled drug-release formulations (CDRFs) offer several advantages over conventional dosage forms. Some of the salient features of controlled-release formulations are described below.

- The drug is released in a controlled fashion that is suitable for the application.
- The frequency of doses could be reduced thereby enhancing patient compliance.
- The drug could be released in a targeted region. This could be achieved either by tailoring the formulation to release the drug in that particular environment or by timed release of the drug. By targeting drugs, the drug efficacy could be maximized.
- The drug could be protected from the physiological environment for a longer duration of time. Thus the effective residence time of the drug could be extended.

However, controlled-release products do not always provide positive effects for every type of formulation design. Negative effects outweigh benefits in the following circumstances.^[13,14]

- Dose dumping
- Less accurate dose adjustment
- Increased potential for first-pass metabolism
- Dependence on residence time in GI tract
- Delayed onset

The limitations of CDRF technology making some drugs unsuitable for formulations are as follows:^[15]

- There is a risk of drug accumulation in the body if the administered drug has a long half-life, causing the drug to be eliminated at a slower rate than it is absorbed.
- Some drugs have a narrow therapeutic index, and thus, need to be administered repeatedly to maintain the serum drug level within a narrow range. Such drugs may not be feasible for CDRF.
- If the GI tract limits the absorption rate of the drug, the effectiveness of the CDRF is limited (for oral controlled release).
- If a drug undergoes extensive first-pass clearance, its controlled release formulation may suffer from lower bioavailability.
- The cost of CDRF may be substantially higher than the conventional form.

POLYMERS IN CONTROLLED DRUG-RELEASE FORMULATIONS

Polymers have been widely used as excipients that adjust the consistency of creams and liquid formulations, as plasma expanders, and as tablet coatings. Generally, pharmaceutical research is focused on the synthesis and testing of new drug molecules, while less effort is made on the final dosage form. Drugs are almost exclusively administered orally or injected, often at a site remote from the target tissue. However, conventional therapeutic systems suffer from various disadvantages while using the recently developed more potent drugs. The disadvantages include adverse side effects, strongly fluctuating drug levels in the body, poor drug efficacy, and poor patient compliance. The concentration, duration, and bioavailability of the pharmaceutical agents cannot be controlled. Controlled-release technology was anticipated to circumvent these problems.^[15] The ambition was to maintain a therapeutic concentration of a drug in the body for a sustained period of time by releasing the agent in a predictable and controllable fashion. Various natural, semisynthetic and synthetic polymers are in use as the

structural backbone for both controlled-release and conventional drug delivery systems^[6,7,15] (various articles in this encyclopedia describe a variety of polymers used for biomedical applications). Polymers selected in the preparation of the dosage form must comply with the following requirements.

- **Safety:** Harmful/toxic impurities must be removed from polymers before their use in CDRFs. The residual monomers, initiators, and other chemicals used in the polymer synthesis/modification must be removed after the polymerization/modification. The chemicals employed in the polymer fabrication processes (i.e., additives, stabilizers, plasticizers, and catalyst residues) are carefully selected to meet regulatory requirements.
- **Physical and mechanical properties:** The polymers must possess the necessary mechanical properties required for the dosage form design, such as elasticity, compatibility, resistance to tensile, swelling, and shear stresses, and resistance to tear and fatigue.
- **Biocompatibility:** The polymer should not cause significant local irritation to the surrounding tissues. If biodegradable, then the polymer degradation by-products must be nontoxic and non-immunogenic and noncarcinogenic.

There are many ways to synthesize new polymers and modify existing polymers. Different monomers (for addition polymerization or condensation polymerization) may be used or existing polymers may be modified. However, only a handful of polymers are used in pharmaceutical drug delivery systems due to their commercial availability, established biocompatibility, and government registration. Most polymers used in pharmaceutical dosage forms were not originally designed for this purpose. However, the production of new, life-saving, genetically engineered drugs (peptides and proteins), that have characteristically short half-lives presents an opportunity for significant research in the area of polymer development in order to prolong their therapeutic effects in the human body.

POLYMER FABRICATION AND DRUG ENCAPSULATION

Encapsulation of the therapeutic agent into a polymer matrix can be brought about by a variety of methods. These methods depend on the desired fabrication of the polymer matrix. Some of the common techniques used for drug encapsulation for different polymer forms are described in this section.

Polymer Films and Rods

Polymer films can be cast by various techniques including dip coating, spin coating, hot-melt casting, and solvent casting. Typical applications for polymer films are in the medical device industry as coatings for controlled-release devices and implants such as stents. These applications require the polymer to be elastomeric to allow flexible films at micron-size thickness. The films can be cast with the drug encapsulated by homogenization/solubilization in the polymer melt/solution. The polymer film could also be extruded using a single or double screw extruder. Screw extrusion is also used to fabricate another form of polymer, i.e., drug encapsulated rods/cylinders. Such extruded rods could be implanted subcutaneously to provide sustained drug delivery. There are several reports describing the preparation techniques and application of polymer films and rods.^[16,17]

Polymer Particulates

For applications in drug delivery, the polymer formulation should be present as an injectable form. From this perspective, the polymer-drug combination could be fabricated as microspheres that can be suspended in an injection vehicle prior to injection. Subcutaneous or intramuscular injections are used for microspheres, whereas smaller particles (in the nano range) can be injected intravenously (Fig. 1). Although, a number of microencapsulation techniques have been developed and reported to date, the choice of the technique depends on the nature of the polymer, the drug, the intended use, and the duration of the therapy.^[18–20]

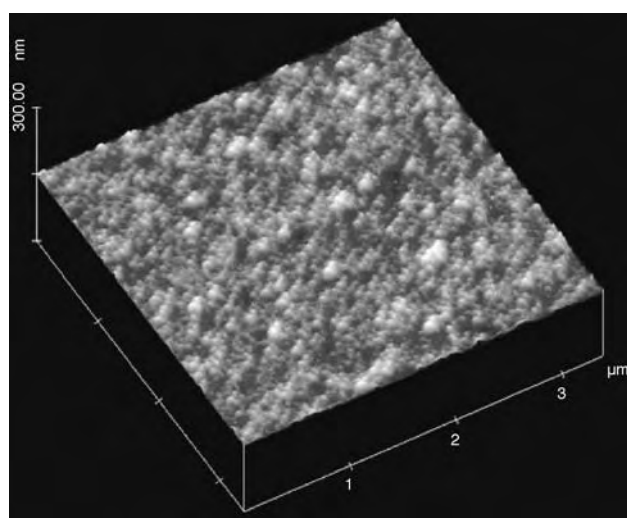


Fig. 1 Atomic force microscopic picture of chitosan nanoparticles. (Photo by M. N. V. Ravi Kumar. Unpublished data.)

Microspheres can be fabricated using a variety of techniques, some of which are described below.

- **Spray drying:** The polymer and drug are dissolved in a common solvent and spray atomized to create microspheres. This technique is useful to create particles in the size range of up to 50 microns, wherein the polymer has a sufficiently high glass transition temperature to allow formation of discrete microparticles during atomization. The size and morphology of microparticles created using the spray drying technique depends on the nature of the polymer as well as the spray dryer operating parameters, including chamber volume, flow rate, and nozzle design.^[21]
- **Emulsion technique:** The drug is emulsified/dispersed in the polymer solution. This emulsion/dispersion is further emulsified in a surfactant bath to allow formation of solid microparticles consisting of the drug encapsulated within the polymer matrix. Variations of this technique include single emulsion and double emulsion microencapsulation. Since microparticles are created during the slow evaporation of the solvent, this technique is also termed as solvent evaporation technique. Solvent evaporation/emulsification techniques are useful when the desired particle size is higher and/or when the drug is not soluble in the organic solvent. A schematic representation of routine emulsion technique to make nano- and microparticles is shown in Fig. 2.
- **Freeze spray atomization:** A suspension of drug in an organic polymer solution is atomized into a liquid nitrogen bath. Absolute ethanol is added to extract the organic solvent from the bath. This process is particularly feasible for proteins since the protein is in a solid, less reactive form, making it less susceptible to damage during processing. Furthermore, the low temperature maintained throughout the process prevents thermal denaturation of the protein.^[22] This process has been utilized in the production of the first commercially available polymer-based, sustained-release protein formulation, Nutropin Depot, marketed by Genentech.

Polymer Gels

Drug encapsulation for injectable gels could be carried out by simply mixing the drug into the polymer gel. To obtain sustained release, the polymer gel should acquire some viscosity upon injection into the body. Such in situ gelation characteristics could be achieved by various mechanisms as:

- **Temperature-induced gelation:** Thermoreversible gelation could be obtained for ABA-type

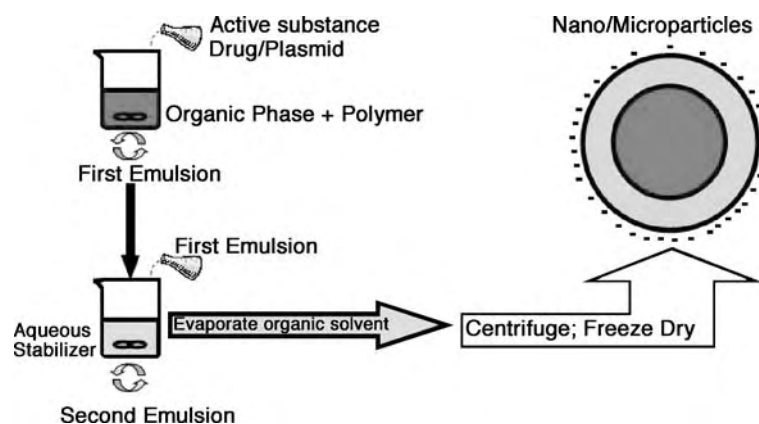


Fig. 2 Routine emulsion technique to make nano- and microparticles.

block copolymers wherein A denotes a hydrophilic segment and B denotes a hydrophobic segment. Because of the alternating block structure, such polymers undergo micellization driven by increasing temperature. The formation of micelles is thermodynamically favorable in ABA block structures, especially when the hydrophobic-hydrophilic balance is appropriately achieved. Examples of such thermoreversible polymers include PEG-PLGA-PEG^[23] and Pluronics, which are ABA-type block copolymers of PEG and PPG. In each of the examples, the polymers could be engineered to be solutions at room temperature and convert into semisolid gels at body temperature. Recently, this concept has also been applied to chitosan gels, which are intact for 7 days when injected subcutaneously in rats (Fig. 3).^[24–26]

- pH-induced gelation: If the polymer is soluble at a certain pH and is insoluble at pH 7.4, then it would undergo pH-induced gelation on injection into the body. Examples of such polymers include synthetic polymers such as polyacrylic acid and natural polymers such as chitosan.
- Solvent-induced gelation: Solvent-induced gelation could be obtained for a water-insoluble polymer dissolved in a biocompatible solvent to create an injectable solution. When the solution is injected, the solvent diffuses out and water from the physiological environment diffuses in. This diffusion process leads to a phase transition for the polymer as it goes from the solvent phase to a nonsolvent phase and forms a semisolid gel.^[27] Several factors such as polymer crystallinity, hydrophilicity, and water uptake govern the sol-gel transition and subsequently the drug release characteristics.

CLASSIFICATION OF CONTROLLED DRUG-RELEASE FORMULATIONS

Drug delivery systems have been classified on the basis of route administration, for example, parenteral, enteral, respiratory, transdermal, etc. (Fig. 4). Controlled-release systems are based on the release mechanisms of erosion, diffusion, or chemical control. For example, under enteral drug delivery systems, release of a drug can be controlled by various mechanisms such as diffusion, osmosis, or a chemically controlled mechanism. In a broad sense, these devices fall into two types: reservoir devices and matrix devices. The former involve the encapsulation of a drug within the polymeric shell, while the latter describes a system in which a drug is well dispersed throughout the polymer matrix.



Fig. 3 Subcutaneous gel-implants had bean-like shapes and remained quite morphologically intact when excised at 3, 6, 12, 24, 48, and 72 hr and 7 days. All implants were intimately integrated within the subdermal fibrous membranes. (Adapted from Ref. [24].)

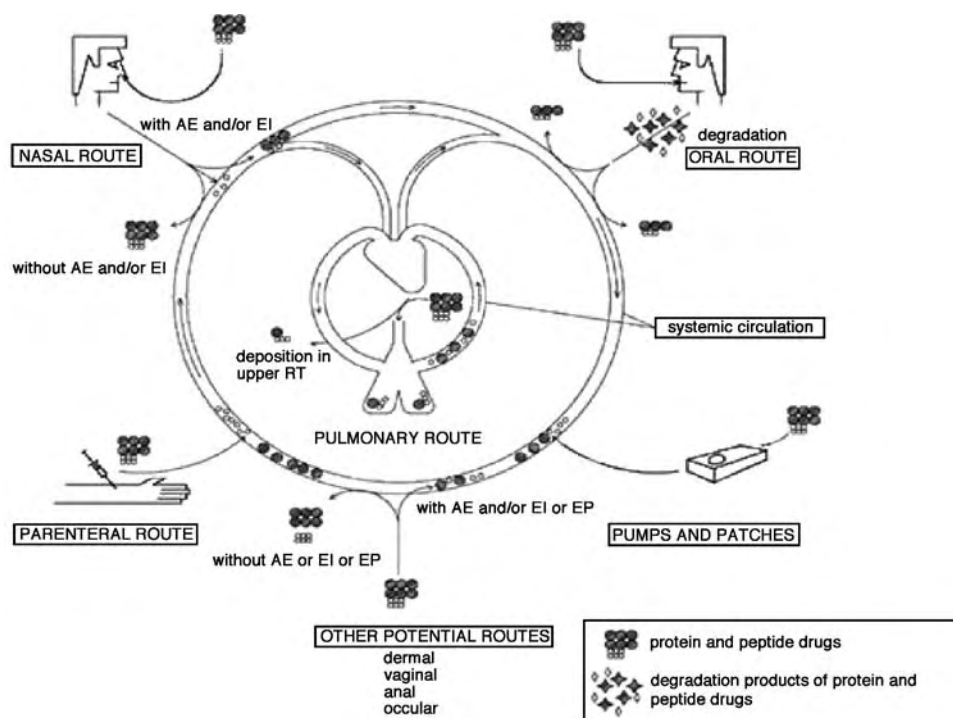


Fig. 4 Clinical and potential routes of administration for therapeutic proteins and peptides. Note that small peptides may be absorbed in limited amounts without absorption enhancers (AE) and/or enzyme inhibitors (EI) via some routes (e.g., nasal). EP, electroporation/iontophoresis (specific for dermal delivery); RT, respiratory tract. (Adapted from Ref. [28].)

Some examples of controlled-release systems are described briefly here.

Diffusion-Controlled

Two types of diffusion-controlled systems have been used including reservoir systems (drug coated by a polymer membrane) and matrix systems (drug dispersed in a polymer matrix). In the reservoir systems, the drug is encapsulated by a polymeric membrane through which the drug is released by diffusion. This polymeric membrane is known as the solution diffusion membrane (implying the mechanism of drug transport) and can be microporous or nonporous. In a nonporous membrane, drug release is governed by the diffusion through the polymer and, thus, release can be controlled by selecting a polymer showing desirable drug solubility and diffusivity in the polymer matrix. In microporous membranes, pores with the size range 1.0 nm to several hundred nm are filled with a drug-permeable liquid or gel medium. Thus, diffusion of the drug through the medium in the pore will dominate the drug-release process. These systems are very useful in the delivery of high-molecular-weight drugs such as protein and peptide drugs. Singh et al.^[29] reported the theoretical considerations of designing and engineering a “barrier coated-reservoir” type of a delivery system for theophylline using poly (vinyl alcohol) (PVA) as

the coating material. After getting the desired theoretical in-vitro release profile, in-vivo studies were carried out on a dog model.

In monolithic systems, the drug is dissolved or dispersed homogeneously throughout the water-insoluble polymer matrix, which may be microporous or nonporous.^[30] Monolithic systems are not suitable for zero-order release; however, it can be achieved by adjusting the physical shape of the device.^[31]

Dissolution-Controlled

Dissolution-controlled systems can also be classified as either reservoir or matrix devices. Polymers used for these devices are generally water-soluble but water-insoluble polymers can also be used as long as they absorb water and disintegrate the drug. In reservoir devices, drug particles are coated with water-soluble polymeric membranes. The solubility kinetics of the membrane depend on the thickness of the membrane and the type of polymer used. Thus, drug release can be achieved and controlled by preparing devices with alternating layers of drug and polymeric coats or by preparing a mixture of particles that have different coating characteristics. Matrix dissolution devices are generally prepared by compressing a powder mix of drug and a water-soluble or water swellable polymer. They can also be made by casting and drying a

polymer solution containing a suitable amount of dissolved or dispersed drug. A variety of other excipients may optionally be included to aid formulation properties. The influence of excipients and formulation factors on the dissolution behavior of methyl hydroxyethyl cellulose (MHEC)-tablets is investigated in Ref. [32]. The use of drugs with higher solubility leads to a slight acceleration of the release due to the contribution of diffusion to the release process (caused by channels formed due to drug solubilization). Furthermore, alterations of the composition of the dissolution medium affect drug release.

Degradation/Erosion-Based Systems

Though early research on polymer-based controlled-release systems involved both degradable and non-degradable polymers, degradable polymers are preferred for parenteral drug delivery applications.^[33] Degradation of the polymer eliminates the need for surgery to recover the spent polymer after the entire drug is released. It also reduces issues related to the long-term safety of the polymer. For biodegradable polymers, release of the drug is often intricately tied up with the polymer degradation profile. The drug can be either physically entrapped in the polymer matrix, wherein it will be released by diffusion and/or erosion of the polymer mass, or chemically attached to the polymer backbone. In such a situation, the drug is released by the enzymatic/hydrolytic cleavage of the chemical bond between the polymer and the drug. A further classification of polymer degradation is established based on the polymer mass erosion patterns. If the polymer is hydrophobic, then it restricts the diffusion of water into its matrix and hence polymer erosion may be restricted to the surface. A more conventional case for the majority of polymers including PLA and PLGA is when erosion occurs throughout the bulk of the polymer matrix (termed bulk erosion). Bulk erosion is associated with a drop in pH within the interior of the matrix. Though such a pH drop is detrimental to pH-labile drugs, it could be utilized to stabilize certain type of basic drugs such as camptothecin.^[34]

Various diffusion models have been proposed for different scenarios and embodiments.^[35] Ritger and Peppas^[36] proposed an empirical equation that has been successful in modeling the drug release for several formulations. This equation correlates the fraction of drug released to the time as:

$$M_t/M_o = kt^n$$

M_t and M_o denote the drug release at time t and total amount of drug in the formulation, respectively. The empirical coefficients k and n are related to the kinetics and diffusion mechanism, respectively. When $n=0.5$,

this equation obeys Fick's law of diffusion. On the other hand, $n=1$ denotes case II diffusion to play a prominent role. The key feature of this model is its simplicity as well as its ability to offer insights into the diffusion mechanisms based on the value of k .

Mathematical models have also been proposed to incorporate polymer degradation kinetics as an integral part of the theoretical framework explaining drug release characteristics.^[37] Similarly, kinetic models have been used to explain degradation profiles of the polymer.

Osmotic Delivery Systems

Osmotic controlled devices compose a core reservoir of drugs, with or without osmotically active salt, coated with a semipermeable membrane. The presence of salt or drug molecules creates an osmotic pressure gradient across the membrane and the diffusion of water into the device gradually forces the drug molecules out through an orifice made in the device. For a durable device, the mechanical strength of semipermeable membrane should be strong enough to resist the stress building inside the device. The drug release rate from the osmotic devices, which is directly dependent on the rate of external water diffusion, can be controlled by the type, thickness, and area of the semipermeable membrane. Alza developed osmotic devices such as an elementary osmotic pump system for oral administration and Alzet an osmotic pump for implant. Singh et al.^[38] discussed osmosis as a phenomenon for controlled drug delivery, along with the device concepts Rose Nelson pump, Higuchi osmotic pump, and Higuchi Theeuwes osmotic pumps.

Ion-Exchange Systems

Polyelectrolytes have been used as cross-linkers to form water-insoluble ion-exchange resins. The drug is bound to the ionic groups by salt formation during absorption and released after being replaced by appropriately charged ions in the surrounding media. For cationic drug delivery, poly(styrene sulfonic acid) and poly(acrylic acid) can be used as anionic ion-exchange resin, where sulfonic and carboxylic group make the complexes with cationic drugs and hydrogen ions and/or other cation such as sodium or potassium ions activate the release of cationic drugs by replacing them from the drug-resin complex. Though cationic ion-exchange resinlike poly(dimethylamino ethyl methacrylate) has been used for the delivery of anionic drug, in which basic groups, namely amino or quaternary amino groups, make a complex with anionic drugs. The interaction of a series of *o-n*-acyl propranolol prodrugs (I, R=C1-9 alkyl or Me3C) with strong cation exchange resins is reported and various variables

that control loading and release profiles were investigated.^[39] The effect of *o-n*-acyl chain length on the loading and release profiles was detected by molecular size; for example, the loading was inhibited for the groups with increased size and release rates were reduced. Again, this enables some control of release profiles but the approach was found to be most suitable for drugs that were active at low doses, which allows full use of these variations without the necessity of large amounts of resin in the delivery system. Sometimes, the ion-exchange resins are additionally coated with a polymer film, such as acrylic acid and methacrylate copolymer or ethylcellulose, to regulate the swelling of the resin and to further control the drug release. The Penkinetic system is an example of the devices based on these mechanisms to deliver dextromethorphan from the ethylcellulose-coated poly(styrene sulfonate).

Polymeric Prodrugs

Many water-soluble polymers possess a functional group to which drug molecules can be covalently attached, and thus these polymers that have no therapeutic effect serve as drug carriers. The drug molecules are gradually released from the polymer by hydrolytic or enzymatic cleavage. If the cleavage occurs by chemical hydrolysis, the drug release depends on the nature of the covalent bonds and pH of the environment, but this release is very slow in the body. If the drug molecule is released by enzymatic hydrolysis, the release is mainly dependent on the concentration of enzymes. Thus, the exact release profile depends on the in-vivo condition, not on the delivery system itself. To be a useful carrier, a polymer should possess certain features:

- The polymer should remain water-soluble even after drug loading.
- The molecular weight of the polymer should be large enough to permit glomerular filtration but small enough to reach all cell types.
- Drug-carrier linkages should be stable in body fluid and yet degradable after capturing in target cells.

This can be achieved by making the linkage degradable by lysosomal enzymes, in which the polymer is nontoxic, non-immunogenic, biocompatible, and degradable by lysosomal enzymes to be eliminated from the body after releasing drugs. Starch derivatives, dextran, poly(aminoacids), PVP, poly(hydroxypropyl methacrylamide), and arabinogalactan have been used as polymeric drug carriers.^[6,7,11,40,41]

Magnetically stimulated systems

The two principal parameters controlling the release rates in these systems are magnetic field characteristics

and mechanical properties of the polymer matrix. It was found that when the frequency of the applied field was increased from 5 to 11 Hz, the release rate of the bovine serum albumin (BSA) from EVAc copolymer matrices increases in linear fashion.^[42] The mechanical properties of the polymeric matrix also affect the extent of magnetic enhancement.^[42] For example, the modulus of elasticity of the EVAc copolymer can be easily altered by changing the vinyl acetate content of the copolymer. The release rate enhancement induced by the magnetic field increases as the modulus of elasticity of EVAc decreases. A similar phenomenon was observed for cross-linked alginate matrices: higher release-rate enhancement for less rigid matrices.^[43] Edelman et al.^[44] also showed that enhanced release rates observed in response to an electromagnetic field (50 G, 60 Hz) applied for 4 min were independent of duration of the interval between repeated pulses.

Photostimulated systems

Photoresponsive gels reversibly change their physical or chemical properties upon photoradiation. A photoresponsive polymer consists of a photoreceptor, usually a photochromic chromophore, and a functional part. The optical signal is captured by the photochromic molecules and then the isomerization of the chromophores in the photoreceptor converts it to a chemical signal. Suzuki and Tanaka^[45] reported a phase transition in polymer gels induced by visible light, where the transition mechanism is due to the direct heating of the network polymer by light.

Ultrasonically stimulated systems

Kost et al.^[46] proposed that cavitation and acoustic streaming are responsible for the augmented degradation and release of biodegradable polymers. Miyazaki et al.^[47] speculate that the ultrasound caused increased temperature in their delivery system, which may facilitate diffusion.

Electrically stimulated systems

In late 1980s Grimshaw et al.^[48] reported four different mechanisms for the transport of proteins and neutral solutes across hydrogel membranes:

- Electrically and chemically induced swelling of a membrane to alter the effective pore size and permeability
- Electrophoretic augmentation of solute flux within a membrane
- Electroosmotic augmentation of solute flux within a membrane

- Electrostatic partitioning of charged solutes into charged membranes

Kwon et al.^[49] studied drug release from electric current-sensitive polymers. Edrophonium chloride, a positively charged solute, was released in an on-off pattern from a matrix device by an electric field. The mechanism was explained as an ion exchange between positive solute and hydroxonium ion, followed by fast release of the charged solute from the hydrogel. The fast release was attributed to the electrostatic force, the squeezing effect, and the electroosmosis of the gel. However, the release of neutral solute was controlled by diffusion affected by swelling and deswelling of the gel.

REPRESENTATIVE APPLICATIONS

Controlled Release of Peptides and Proteins

Sustained-release applications are especially useful for proteins/peptides due to their short half-lives which require frequent injections.^[50] This concept was first utilized commercially in Lupron Depot, which was introduced as a sustained release formulation of a luteinizing hormone-releasing hormone (LHRH) using leuprolide acetate, with poly(lactic acid) as the polymer.^[51] Following the success of Lupron Depot, other sustained release formulations of LHRH analogs have also been commercialized including Zoladex, wherein the LHRH peptide is encapsulated in PLA extruded rods, and Trelstar depot, wherein the peptide is encapsulated in PLGA polymer matrices.

Recently, a once-in-4-weeks formulation of a cyclic peptide Octreotide, encapsulated in biodegradable PLGA-glucose polymer matrix, has also been commercialized under the trade name Sandostatin LAR (Novartis Pharmaceuticals Corporation).^[52] In 2000, a human growth hormone sustained-release formulation (Nutropin Depot) became the first polymer-based sustained release formulation of a therapeutic protein to receive marketing approval from the U.S. Food and Drug Administration. This commercialized formulation encapsulates a zinc-complexed form of recombinant human growth hormone in a PLGA matrix.^[21] Other therapeutic proteins that are being tested for encapsulation in polymer matrices include bone morphogenic protein,^[53] erythropoietin,^[54] and nerve growth factor.^[55]

Controlled Release of Antirestenotic Agents from Stent Coatings

Stents are tiny wire scaffoldlike devices that have become the most successful and widely used innovation

in interventional cardiology of the last decade. These devices are inserted inside blocked sections of coronary arteries and expanded into place using a balloon catheter in a procedure called an angioplasty. In as many as 40 percent of patients receiving angioplasty, a new blockage develops at the site due to scar tissue growth and inflammation, a condition referred to as restenosis. More than 500,000 Americans are treated for restenosis annually. Stents coated with a biocompatible polymer encapsulating an antirestenotic agent have proven to be a successful therapy to reduce or eliminate restenosis. Several drugs including cytostatic agents (such as Rapamune), antiproliferative agents (such as Paclitaxel), and anti-inflammatory agents have been encapsulated in micron-thick films that coat the metallic stent.^[56] The polymers used for this application are required to be elastic, biocompatible, and hemocompatible. Lewis and coworkers^[57] demonstrated the use of phosphorylcholine-based polymers for this application. Various other polymers have also been tested for this purpose.^[58,59]

Controlled-Release Systems for Cancer Treatment

By providing sustained release at the desired site, high doses of toxic drugs can be delivered to the site without introducing the drug into systemic circulation. This results in a major advantage for the administration of chemotherapeutic drugs, wherein the drugs can be encapsulated into polymer matrices and administered directly into the tumor area. Gliadel[®] (Guilford Pharmaceuticals) implant is a product based on the aforementioned concept, encapsulating a chemotherapeutic agent carmustine into a biodegradable polymer, fabricated as a dime-sized wafer. Several wafers are implanted by the surgeon into the brain cavity during a tumor resection surgery. By direct release of the drug in the tumor region in the brain, the problem of overcoming the blood-brain barrier is resolved. Also, high doses of the chemotherapeutic agent are delivered at the tumor site, thus making the therapy significantly more effective than systemic administration. Other chemotherapeutic agents that are also being developed as sustained-release formulations include paclitaxel and cisplatin.^[60]

The biggest success of polymer-controlled release systems in cancer therapy, however, has been in the area of prostate cancer treatment. This therapy is unique because it uses a hormone suppressant rather than a chemotherapeutic agent to minimize cancer cell growth in the prostate. Several products currently in use include Lupron Depot (TAP Pharmaceuticals), Zoladex (Astra-Zeneca), and Trelstar Depot (Debio RP, Pharmacia). Lupron Depot involves PLA

Table 1 Controlled-release products in market

Technology	Product/drug	Application	Organization
Liposome-based formulations	Evacet (Doxorubicin)	Breast cancer and other cancers	Elan with acquisition of Liposome Inc.
One-yearly drug implant	Leuprolide (Viadur)	Prostate cancer	ALZA
Intradose injectable gel	Cisplatin/ epinephrine	Recurrent or refractory head and neck cancer	Matrix Pharmaceuticals Inc.
Biodegradable microsphere for sustained release	Medisorb	Delivery of peptides and small molecules	Alkermes Inc.
Sustained-release injections	ProLease	Delivery of fragile proteins and small bioactive molecules	Alkermes Inc.
Time release oral drug release	Pulsincap	Drug release at predetermined time or location in the GI-tract	RP scherer corporation
Oral controlled-release system to control the release of a specific drug	Geomatrix	Predetermined therapeutic objective for a drug	SkyePharma

microspheres and Zoladex formulation is an extruded rod, whereas Trelstar depot consists of PLGA microspheres. All of these drugs incorporate LHRH analogs.

Sustained Release of Drugs for CNS-Related Disorders

Another area of application for polymer-based controlled-release technologies is for the sustained delivery of drugs for central nervous system (CNS)-related disorders. CNS-related drugs include a wide range of therapeutics including pain management agents, drugs to prevent substance abuse, and drugs for conditions such as schizophrenia, Parkinson's disease, and Alzheimer's disease. Drugs such as lidocaine and bupivacaine have been studied for sustained local anesthesia at the site of surgery.^[61] For this application, the anesthetic agent is encapsulated in a biodegradable polymer and injected in the proximity of the site of pain. The drug is released facilitating high concentrations at the local site without reaching the threshold levels of systemic toxicity. In other cases such as schizophrenia, Parkinson's disease, and Alzheimer's disease, sustained release of the medication may reduce the chance of missed doses and aid in an effective dose regimen.

COMMERCIALY AVAILABLE CONTROLLED-RELEASE SYSTEMS

Controlled release has had a good impact among the various drug delivery technologies due to patient compliances, safety of the drug, and minimum side

effects. It had been estimated that the world's pharmaceuticals sales reached nearly \$400 billion during the year 2000, and that about 12.5%, or \$50 billion, went toward drugs involving special drug-delivery technology. Sales of drug-delivery products are expected to more than double to \$104 billion in 2005. More than 20% of the world's drug-delivery products sold in 2005 are estimated to involve controlled drug-release technologies. Examples of controlled-release technologies available in the market place are given in Table 1.

ARTICLES OF FURTHER INTEREST

Antimicrobial/Antibiotic (Infection Resistance) Materials; Biofunctional Polymers; Chitosan; Controlled Release; Insulin Delivery; Poly(Glycolic Acid); Poly(lactic acid)s; Transdermal Drug Delivery

REFERENCES

1. Ruan, G.; Feng, S.-S. Preparation and characterization of poly(lactic acid)-poly(ethylene glycol)-poly(lactic acid) (PLA-PEG-PLA) microspheres for controlled release of paclitaxel. *Biomaterials* **2003**, *24*, 5037–5044.
2. Sood, A.; Panchagnula, R. Peroral route: An opportunity for protein and peptide drug delivery. *Chem. Rev.* **2001**, *101*, 3275–3303.
3. Courrier, H.M.; Butz, N.; Vandamme, Th.F. Pulmonary drug delivery systems: Recent developments and prospects. *Crit. Rev. Ther. Drug Carr. Syst.* **2002**, *19* (4 & 5), 425–498.

4. Hejazi, R.; Amiji, M. Chitosan-based gastrointestinal delivery systems. *J. Control. Release* **2003**, *89* (2), 151–165.
5. Duvvuri, S.; Majumdar, S.; Mitra, A.K. Drug delivery to the retina: Challenges and opportunities. *Expert Opin. Biol. Ther.* **2003**, *3* (1), 45–56.
6. Uhrich, K.E.; Cannizzaro, S.M.; Langer, R.S.; Shakesheff, K.M. Polymeric systems for controlled drug release. *Chem. Rev.* **1999**, *99*, 3181–3198.
7. Ravi Kumar, M.N.V.; Kumar, N.; Domb, A.J.; Arora, M. Pharmaceutical polymeric controlled drug delivery systems. *Adv. Polym. Sci.* **2002**, *160*, 45–117.
8. Luo, Y.; Prestwich, G.D. Cancer-targeted polymeric drugs. *Curr. Cancer Drug Targets* **2002**, *2* (3), 209–226.
9. Qiu, Y.; Park, K. Environment-sensitive hydrogels for drug delivery. *Adv. Drug Deliv. Rev.* **2001**, *53* (3), 321–339.
10. Panyam, J.; Labhasetwar, V. Biodegradable nanoparticles for drug and gene delivery to cells and tissue. *Adv. Drug Deliv. Rev.* **2003**, *55* (3), 329–347.
11. Edlund, E.; Albertsson, A.-C. Degradable polymer microspheres for controlled drug delivery. *Adv. Polym. Sci.* **2002**, *157*, 67–112.
12. Panchagnula, R. Transdermal delivery of drugs. *Indian J. Pharmacol.* **1997**, *29* (3), 140–156.
13. Ritschel, W.A. Biopharmaceutic and pharmacokinetic aspects in the design of controlled-release peroral drug delivery systems. *Drug Dev. Ind. Pharm.* **1989**, *15* (6–7), 1073–1103.
14. Rubinstein, A.; Robinson, J.R. Controlled drug delivery. *Prog. Clin. Biochem. Med.* **1987**, *4*, 71–107.
15. Vyas, S.P.; Khar, R.K. *Targeted and Controlled Drug Delivery: Novel Carrier Systems*; CBS: New Delhi, 2002.
16. Witt, C.; Kissel, T. Morphological characterization of microspheres, films and implants prepared from poly(lactide-co-glycolide) and ABA triblock copolymers: Is the erosion controlled by degradation, swelling or diffusion? *Eur. J. Pharm. Biopharm.* **2001**, *51* (3), 171–181.
17. Rothen-Weinhold, A.; Gurny, R.; Orsolini, P.; Heimgartner, F. Implants for controlled release of pharmaceutically active principles and method for making same. *PCT Int. Appl.* **1998**. 26 pp. WO 9855101.
18. Ravi Kumar, M.N.V. Nano and microparticles as controlled drug delivery devices. *J. Pharm. Pharm. Sci.* **2000**, *3* (2), 234–258.
19. Jain, R.A. The manufacturing techniques of various drug loaded biodegradable poly(lactide-co-glycolide) (PLGA) devices. *Biomaterials* **2000**, *21*, 2475–2490.
20. Torchilin, V.P. Polymer-coated long-circulating microparticulate pharmaceuticals. *J. Microencapsul.* **1998**, *15* (1), 1–19.
21. Palmieri, G.F.; Bonacucina, G.; Di Martino, P.; Martelli, S. Spray-drying as a method for microparticulate controlled release systems preparation: Advantages and limits. I. Water-soluble drugs. *Drug Dev. Ind. Pharm.* **2001**, *27* (3), 195–204.
22. Johnson, O.L.; Jaworowicz, W.; Cleland, J.L.; Bailey, L.; Charnis, M.; Duenas, E.; Wu, C.; Shepard, D.; Magil, S.; Last, T.; Jones, A.J.; Putney, S.D. The stabilization and encapsulation of human growth hormone into biodegradable microspheres. *Pharm. Res.* **1997**, *14* (6), 730–735.
23. Jeong, B.; Bae, Y.H.; Kim, S.W. In situ gelation of PEG-PLGA-PEG triblock copolymer aqueous solutions and degradation thereof. *J. Biomed. Mater. Res.* **2000**, *50* (2), 171–177.
24. Chenite, A.; Chaput, C.; Wang, D.; Combes, C.; Buschmann, M.D.; Hoemann, C.D.; Leroux, J.C.; Atkinson, B.L.; Binette, F.; Selmani, A. Novel injectable neutral solutions of chitosan form biodegradable gels in situ. *Biomaterials* **2000**, *21* (21), 2155–2161.
25. Ruel-Gariepy, E.; Chenite, A.; Chaput, C.; Guirguis, S.; Leroux, J.-C. Characterization of thermosensitive chitosan gels for the sustained delivery of drugs. *Int. J. Pharm.* **2000**, *203* (1–2), 89–98.
26. Chenite, A.; Buschmann, M.; Wang, D.; Chaput, C.; Kandani, N. Rheological characterization of thermogelling chitosan/glycerol-phosphate solutions. *Carbohydr. Polym.* **2001**, *46* (1), 39–47.
27. Brodbeck, K.J.; Pushpala, S.; McHugh, A.J. Sustained release of human growth hormone from PLGA solution depots. *Pharm. Res.* **1999**, *16* (12), 1825–1829.
28. Agu, R.U.; Ugwoke, M.I.; Armand, M.; Kinget, R.; Verbeke, N. *Respir. Res.* **2001**, *2*, 198–209.
29. Singh, R.; Matharu, P.; Lalla, J.K. A diffusion-controlled drug delivery system for theophylline. *Drug Dev. Ind. Pharm.* **1994**, *20* (7), 1225–1238.
30. Heller, J. Controlled drug release from monolithic systems. *Fidia Res. Ser.* **1987**, *11* (Ophthalmic Drug Delivery), 179–189.
31. Laurencin, C.T.; Koh, H.J.; Neenan, T.X.; Allcock, H.R.; Langer, R. Controlled release using a new bioerodible polyphosphazene matrix system. *J. Biomed. Mater. Res.* **1987**, *21* (10), 1231–1246.
32. Zuleger, S.; Lippold, B.C. Polymer particle erosion controlling drug release. I. Factors influencing drug release and characterization of the release mechanism. *Int. J. Pharm.* **2001**, *217* (1–2), 139–152.
33. Heller, J.; Domb, A.J.; Kumar, N. Polyanhydrides and poly(ortho esters). *Adv. Drug Deliv. Rev.* **2002**, *54* (7), 887–888.
34. Shenderova, A.; Burke, T.G.; Schwendeman, S.P. The acidic microclimate in poly(lactide-co-glycolide) microspheres stabilizes camptothecins. *Pharm. Res.* **1999**, *16* (2), 241–248.
35. Siepmann, J.; Peppas, N.A. Modeling of drug release from delivery systems based on hydroxypropyl methylcellulose (HPMC). *Adv. Drug Deliv. Rev.* **2001**, *48* (2–3), 139–157.
36. Ritger, P.L.; Peppas, N.A. A simple equation for description of solute release. II. Fickian and anomalous release from swellable devices. *J. Control. Release* **1987**, *5* (1), 37–42.
37. Batycky, R.P.; Hanes, J.; Langer, R.; Edwards, D.A. A theoretical model of erosion and macromolecular drug release from biodegrading microspheres. *J. Pharm. Sci.* **1997**, *86* (12), 1464–1477.
38. Singh, P.; Sihorkar, V.; Mishra, V.; Saravanababu, B.; Venkatesan, N.; Vyas, S.P. Osmotic pumps. From present view to newer perspectives in pharmaceutical industry. *East. Pharm.* **1999**, *42* (503), 39–46.

39. Ehrenfreund-Kleinman, T.; Azzam, T.; Falk, R.; Polachek, I.; Golenser, J.; Domb, A.J. Synthesis and characterization of novel water soluble amphotericin B-arabinogalactan conjugates. *Biomaterials* **2002**, *23* (5), 1327–1335.
40. Irwin, W.J.; Belaid, K.A. Drug-delivery by ion-exchange. Part I: Ester prodrugs of propranolol. *Drug Dev. Ind. Pharm.* **1987**, *13* (9–11), 1231–2017.
41. Mehvar, R. Dextrans for targeted and sustained delivery of therapeutic and imaging agents. *J. Control. Release* **2000**, *69* (1), 1–25.
42. Kost, J.; Noecker, R.; Kunica, E.; Langer, R. Magnetically controlled release systems: Effect of polymer composition. *J. Biomed. Mater. Res.* **1985**, *19* (8), 935–940.
43. Saslawski, O.; Couvreur, P.; Peppas, N. Proceedings of 15th International Symposium on Controlled Release of Bioactive Materials. Basel, Heller, J., Harris, F., Lohmann, H., Merkle, H., Robinson, J., Eds.; 1988, 26.
44. Edelman, E.R.; Brown, L.; Taylor, J.; Langer, R. In vitro and in vivo kinetics of regulated drug release from polymer matrices by oscillating magnetic fields. *J. Biomed. Mater. Res.* **1987**, *21* (3), 339–353.
45. Suzuki, A.; Tanaka, T. Photoinduced phase transition of gels. *Nature* **1990**, *346*, 345–347.
46. Kost, J.; Leong, K.; Langer, R. Ultrasound-enhanced polymer degradation and release of incorporated substances. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86* (20), 7663–7666.
47. Miyazaki, S.; Yokouchi, C.; Takada, M. External control of drug release: Controlled release of insulin from a hydrophilic polymer implant by ultrasound irradiation in diabetic rats. *J. Pharm. Pharmacol.* **1988**, *40* (10), 716–717.
48. Grimshaw, P.E.; Grodzinsky, A.J.; Yarmush, M.L.; Yarmush, D.M. Dynamic membranes for protein transport: Chemical and electrical control. *Chem. Eng. Sci.* **1989**, *44* (4), 827–840.
49. Kwon, I.C.; Bae, Y.H.; Kim, S.W. Electrically erodible polymer gel for controlled release of drugs. *Nature* **1991**, *354* (6351), 291–293.
50. Cohen, S.; Yoshioka, T.; Lucarelli, M.; Hwang, L.H.; Langer, R. Controlled delivery systems for proteins based on poly(lactic/glycolic acid) microspheres. *Pharm. Res.* **1991**, *8* (6), 713–720.
51. Okada, H. One- and three-month release injectable microspheres of the LH-RH superagonist leuporelin acetate. *Adv. Drug Deliv. Rev.* **1997**, *28*, 43–70.
52. Davies, P.H.; Stewart, S.E.; Lancranjan, L.; Sheppard, M.C.; Stewart, P.M. Long-term therapy with long-acting octreotide (Sandostatin-LAR) for the management of acromegaly. *Clin. Endocrinol.* **1998**, *48* (3), 311–316.
53. Schrier, J.A.; DeLuca, P.P. Recombinant human bone morphogenetic protein-2 binding and incorporation in PLGA microsphere delivery systems. *Pharm. Dev. Technol.* **1999**, *4* (4), 611–621.
54. Pistel, K.F.; Bittner, B.; Koll, H.; Winter, G.; Kissel, T. Biodegradable recombinant human erythropoietin loaded microspheres prepared from linear and star-branched block copolymers: Influence of encapsulation technique and polymer composition on particle characteristics. *J. Control. Release* **1999**, *59* (3), 309–325.
55. Lam, X.M.; Duenas, E.T.; Cleland, J.L. Encapsulation and stabilization of nerve growth factor into poly(lactico-glycolic) acid microspheres. *J. Pharm. Sci.* **2001**, *90* (9), 1356–1365.
56. Regar, E.; Sianos, G.; Serruys, P.W. Stent development and local drug delivery. *Br. Med. Bull.* **2001**, *59*, 227–248.
57. Lewis, A.L.; Cumming, Z.L.; Goreish, H.H.; Kirkwood, L.C.; Tolhurst, L.A.; Stratford, P.W. Cross-linkable coatings from phosphorylcholine-based polymers. *Biomaterials* **2001**, *22* (2), 99–111.
58. van der Giessen, W.J.; Lincoff, A.M.; Schwartz, R.S.; van Beusekom, H.M.; Serruys, P.W.; Holmes, D.R., Jr.; Ellis, S.G.; Topol, E.J. Marked inflammatory sequelae to implantation of biodegradable and nonbiodegradable polymers in porcine coronary arteries. *Circulation* **1996**, *94* (7), 1690–1697.
59. Hietala, E.-M.; Salminen, U.-S.; Stahls, A.; Valimaa, T.; Maasilta, P.; Tormala, P.; Nieminen, M.S.; Harjula, A.L.J. Biodegradation of the copolymeric polylactide stent: Long-term follow-up in a rabbit aorta model. *J. Vasc. Res.* **2001**, *38* (4), 361–369.
60. Harper, E.; Dang, W.; Lapidus, R.G.; Garver, R.I., Jr. Enhanced efficacy of a novel controlled release paclitaxel formulation (PACLIMER delivery system) for local-regional therapy of lung cancer tumor nodules in mice. *Clin. Cancer Res.* **1999**, *5* (12), 4242–4248.
61. Paavola, A.; Tarkkila, P.; Xu, M.; Wahlstrom, T.; Yliruusi, J.; Rosenberg, P. Controlled release gel of ibuprofen and lidocaine in epidural use—Analgesia and systemic absorption in pigs. *Pharm. Res.* **1998**, *15* (3), 482–487.

Elastin

William T. Brinkman
Karthik Nagapudi

Department of Surgery, Emory University School of Medicine, Atlanta, Georgia, U.S.A.

Elliot L. Chaikof

Department of Surgery, Emory University School of Medicine, Department of Biomedical Engineering, Emory University School of Medicine, and School of Chemical Engineering, Georgia Institute of Technology, Atlanta, Georgia, U.S.A.



INTRODUCTION

Elastomeric proteins are widely distributed among a diverse range of animal species and tissues in which they have evolved precise structures to perform specific biological functions. These proteins, which include abductin,^[1] tropoelastin,^[2,3] bysuss,^[4] silk,^[5,6] and titin all possess rubberlike elasticity. This flexibility enables them to undergo high deformation without rupture, store energy involved in deformation, and then recover to their original state when the stress is removed. Significantly, the resilience of these bioelastomeric proteins dominates low-strain mechanical responses, which minimizes fatigue and failure due to repetitive loads by preventing the dissipation of transmitted energy into heat. In man, elastin is an important component of tissues subjected to repetitive strain, such as the arterial wall, aortic and mitral valves, plus skin and pulmonary alveoli. In certain circumstances, elastin also serves an important physiologic role by facilitating tissue shape recovery upon deformation. A case in point is respiration, which is dependent upon effective air exchange that is linked to rapid alveolar recoil during expiration.

Although elastin fibers are structurally complex and may contain glycoproteins and glycosaminoglycans, the physical properties of the network have been attributed primarily to the elastin protein component produced from the soluble precursor tropoelastin. The ability of this protein to exhibit rubberlike elasticity relates both to its primary and secondary structure as well as to those features such as protein self-assembly and other intermolecular interactions that dictate the formation of true or virtual networks. The objective of this article is to provide an overview of the structure and properties of elastin and elastin-related proteins as well as provide a brief description of their potential applications as novel biomaterials.

STRUCTURE OF THE ELASTIN GENE

While the primary focus of this section is the description of structure-function relationships that

underlie the unique potential of elastin and elastin-related polymers as biomaterials, it should be recognized that recent efforts in tissue engineering have also been directed at manipulating the expression of elastin through metabolic or genetic approaches in order to enhance the functional performance of cell-based tissue engineering constructs. Thus, it is important to note that the elastin gene exists as a single copy localized on chromosome 7q11.1-21.1 where it is characterized by small exons (27–265 bp) that are interspersed between large introns.^[6] Tropoelastin, the soluble noncross-linked precursor of elastin, expression is primarily under pre-translational control with potential control regions in untranslated regions, intronic regions, and promoter regions. But, posttranslational mechanisms that control protein expression have also been described.^[7] For example, decreased tropoelastin synthesis due to mRNA destabilization has been observed in aortic smooth muscle cells isolated from elderly animals.^[8] Furthermore, growth factors such as transforming growth factor- β , insulin-like growth factor I, vitamin D, and interleukin-1 α have all been shown to modulate tropoelastin expression.^[7]

ELASTIN PROTEIN STRUCTURE

Determination of the Primary Structure of Elastin

Defining the primary sequence of elastin was initially difficult due to its unusual physiochemical properties, including its remarkable resistance to enzymatic, chemical, and physical degradation. These properties have been attributed to both its intrinsic molecular structure and its highly cross-linked state. In 1969, Partridge^[9] provided insight into the complex and varied nature of the elastin cross-links by determining that four lysine residues—two from each elastin chain—could combine to form desmosine and isodesmosine crosslinks while two lysine residues were capable of forming an individual cross-link that is termed lysinonorleucine. With insight into cross-link

formation, the soluble precursor, tropoelastin, could be produced through an inhibition of these processes. This was accomplished by raising pigs on a copper-free diet, which limited the activity of lysyl oxidase,^[10,11] the critical copper-dependent enzyme responsible for oxidizing lysine residues. With a soluble precursor protein in hand, elastin's primary structure was determined.

By the early 1970s, Gray and Sandberg^[12] began to elucidate elastin's amino acid sequence with the identification of alanine-lysine rich sites that were presumed to be important for cross-linking and glycine-proline-valine rich sequences. Repeat sequences of Val-Pro-Gly-Val-Gly and Val-Pro-Gly-Gly were frequently identified between potential cross-linking regions, and it was postulated that these sequences were responsible for elastin's unusual extensibility and fatigue resistance. In the process, Gray and Sandberg established that elastin was a fibular protein composed of alternating segments of cross-linked and loosely coiled regions, which they referred to as oiled coils. Thus, while Hoeve and Flory^[13] demonstrated in 1974 that the driving force for the spontaneous recoil of stretched elastin is entropic in origin, which is consistent with classical theory of rubber elasticity, an apparent paradox to this conceptual picture of elastin were these and other experimental results that confirmed that, at a supra-molecular level, the conformation of elastin was not that of a random coil that is typical of polyisoprene or other classic nonprotein based elastomers.

Further molecular insight into elastin's structure was subsequently provided by a comparative analysis of human^[6] and bovine^[14] elastins and their gene structure. Human elastin,^[6] which consists of 786 amino acids and has a molecular weight of approximately 70 kDa, exhibits significant homology with both DNA and amino acid sequences of elastins isolated from chick,^[15] bovine,^[14] and rat^[16] tissue. In all cases, the amino acid composition is remarkable for a high content of Gly, Ala, Val, and Pro; a scarcity of Arg relative to lysine; absence of His, Cys, Met, and Trp; and a very low frequency (<5%) of polar amino acids. The C-terminus of tropoelastin is highly conserved among all species tested. The highly basic terminus region contains the only two Cys residues in elastin and ends with a positively charged RKRK sequence.^[7] The loss of lysine through the formation of cross-links is the main difference between soluble tropoelastin and insoluble elastin.^[12] Indik et al.^[6] was the first to isolate the human elastin cDNA clone and confirmed the presence of coding regions for both hydrophobic domains as well as hydrophilic regions rich in lysine residues, as suggested by protein sequencing studies. Of significance, they demonstrated that these regions were coded by separate exons of the 17.8 kb elastin gene. Although alterations in mature elastin mRNA produced by alternative splicing has been noted, the functional significance of these splicing events currently remains unknown (Fig. 1).

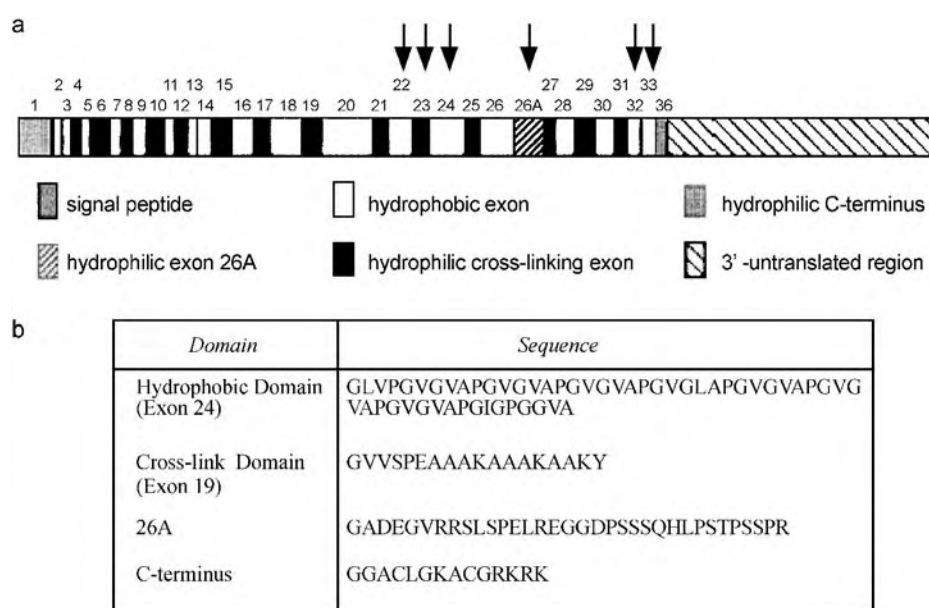


Fig. 1 Structure of human tropoelastin: a) cDNA structure; and b) amino acid sequences of selected domains. Most of the gene consists of exons encoding hydrophobic and cross-linking regions which alternate. Typical sequences are shown in (b). Exon 26A is an unusual hydrophobic exon of unknown function that does not encode any lysine residues. Thus, it is not involved in lysyl oxidase mediated cross-linking. Exon 36 encodes a hydrophilic C-terminus as well as a large three-foot-untranslated region. Exons subject to alternate removal by splicing are marked with an arrow. (Used with permission from Vrhovski and Weiss. From Ref. [7].)

Secondary and Tertiary Structure of Elastin

In applying their knowledge of the primary sequence of elastin, Gray and Sandberg were the first to propose a molecular model for elastin's secondary structure, the β -spiral. Moreover, additional spectroscopic studies confirmed the presence of α -helical and β -turn conformations and a twisted-rope organization imaged by electron microscopy. In an effort to relate a growing understanding of elastin's molecular structure to its mechanical properties, liquid drop and oiled coil models were proposed as an alternative to the random coil model following the identification of both hydrophobic and hydrophilic elastin protein domains. The liquid drop model assumed that, when hydrated, elastin could be viewed as a two-phase system such that hydrophobic globular domains (oil droplets) were connected by hydrophilic cross-links. When subjected to strain, additional hydrophobic regions containing Val and Pro residues would be exposed to surrounding water, which causes a decrease in entropy with a predilection for recoil due to an entropic driving force.^[17,18] In an extension of the liquid drop model, the oiled coil model considered elastin to be composed of globular units composed of broad, coiled hydrophobic regions separated by cross-linked α -helical regions.^[12] Although high-resolution NMR and X-ray crystallography structures have yet to be obtained, subsequent investigations by Urry, Tamburro, et al.^[13,19-24] have elaborated more refined models in order to reconcile the entropic origins for elastin's tensile characteristics and its known molecular structure.

The Tamburro model is largely derived from studies of the GXGGX repeat sequence in which G represents glycine and X is a hydrophobic amino acid.^[25,26] The model suggests that there are two major families of folded or quasifolded structures (type I and type II β -turns, γ -turns, and half-turns) and extended or quasiextended structures (β -sheets, polyproline II conformation) that are dynamically interchanging among themselves and, in the case of β -turns, sliding along the chain.^[25-29] In the stretched chain, a reduction of phase-space volume and the development of hydrophobic interactions allow only high-frequency, low-amplitude motion. In structural terms this would be consistent with less flexible β -pleated sheets. On relaxation, larger amplitude motions are characterized by high dynamic entropy.^[30]

The Urry model is based upon studies with synthetic polymers derived from the pentapeptide unit (VPGVG)_n.^[31] This sequence occurs frequently along the elastin chain, albeit in a different region than that containing the GXGGX sequences. Urry noted by NMR and circular dichroism spectroscopy that proline, due to its rigid imino acid structure, induces a type II β -turn when associated with glycine. As a

consequence, hydrogen bond formation between the valine residues located in the 1st and 4th positions of pentapeptide repeat sequence is facilitated and gives rise to helical β -spiral. One type II β -turn is present per pentameric unit and, as a group, acts as spacers between dipeptide VG segments, which can undergo large-amplitude, low-frequency rocking motions. On chain extension, a decrease in amplitude of the rocking motion causes a large decrease in the entropy of the segment. In turn, this provides the driving force for a return to the relaxed state.

Further insight into elastin's tertiary structure has also been obtained by a consideration of its propensity to coacervate or aggregate in aqueous solution upon heating. The temperature above which coacervation occurs, which is referred to as an inverse transition temperature (T_i), or the lower critical solution temperature (LCST) in the synthetic polymer literature, appears to be associated with folding and assembly of elastin molecules that are individually organized as loose β -spiral chains.^[32] This phenomenon can be explained by the interaction of water molecules with the hydrophobic regions of the β -spiral. In brief, water forms an ordered arrangement around the hydrophobic regions of tropoelastin at low temperature, which maintains the protein in an unfolded state. Upon increasing temperature, the surrounding water becomes more disordered and allows the hydrophobic segments to form a more ordered β -spiral.^[32] This structure is inherently in an equilibrium state since the positive entropy change from the disordering of surrounding water is greater than the negative entropy arising from the formation of the β -spiral.^[33] Since the interactions of water with the elastin chain are dependent upon the polarity of the substituent residues, the inverse transition temperature (T_i) provides a measure of the relative hydrophobicity of the substituent residue.^[31]

TROPOELASTIN ASSEMBLY AND ELASTIN ULTRASTRUCTURE

Insight into the formation of higher order elastin structures in vivo, such as fibers and fiber networks, have been provided by in-vitro studies, which have demonstrated, that, with an increase in temperature to 37°C, tropoelastin exhibits an increasing rate of aggregation or coacervation. It is noteworthy that tropoelastin assembly in aqueous solution at temperatures of up to 50°C is characterized by the formation of fibers with diameters in the 5 nm range and lateral attachments similar to native elastin fibers. As the temperature is increased above 50°C, additional organization of elastin filaments is noted by the appearance of banded fibers with a staggered periodicity of

50 nm.^[34] Since these processes are thermodynamically controlled, they can be reversed by cooling in aqueous solution. All told, coacervation is believed to be a physiologically relevant step in elastin fiber formation since it both concentrates and aligns elastin to enable cross-linking. Moreover, this aggregation behavior has also been demonstrated for synthetic peptides composed of hydrophobic elastin consensus sequences.^[35] Thus, the presence of an inverse temperature transition and associated protein coacervation is also an important consideration in the processing of elastin-based materials and the development of novel applications in the areas of actuator, sensor design, and drug delivery.

Although freeze fracture electron microscopy of purified tropoelastin monomers reveals a globular protein structure with a diameter range of 5–7 nm,^[36] within tissues, elastin is organized as a network of interwoven fibers. The arrangement of which varies within organ systems and amongst various phyla and species. For example, mammalian aorta contains dense elastin fibers while lower vertebrates tend to have loose fiber arrangements.^[37] In general, elastin is most abundant in tissues that are subjected to repetitive strain, such as skin, alveoli, arterial wall, and cartilage, with an organizational pattern that varies in relationship to the applied stress field.^[38] For example, detailed examination of the mammalian arterial wall has demonstrated networks of fenestrated elastin lamellae arranged in concentric layers that are separated by spindle-shaped smooth muscle cells, loose collagen bundles, and proteoglycans.^[39] In the dermis, elastin forms a loose network parallel to the skin surface in which the volume fraction is between 1.0 and 2.5%. Mean fiber diameters range between 0.5 and 3.0 μm .^[40] In ligamentum nuchae, elastin fibers tend to be much thicker (10–20 μm) in which they run parallel to and are separated by collagen fibers.^[41] All of this emphasizes that both the molecular structure of elastin and its organization into higher order assemblies are critical determinants of functional properties.

APPLICATIONS IN BIOENGINEERING

Elastin's capacity for shape and energy recovery as well as its capacity to aggregate or coacervate in aqueous solution with changes in temperature, pH, pressure, salt, organic solutes, and oxidation state of a side chain or functional group provide opportunities for the design of novel materials for drug delivery, biosensors, or tissue engineering.^[42–45] Other potential advantages of elastin-based polymers are their inherent biostability and biocompatibility. For example, the extraordinary metabolic stability of native elastin has been confirmed by ¹⁴C prevalence data and quantitative

autoradiography that suggest very slow elastin turnover with an associated half life of 70 years.^[46] Furthermore, implant studies of radiation cross-linked poly(VPGVG)-related materials have been associated with minimal or no inflammatory response and without evident material degradation.^[47]

Tissue-derived elastin has had relatively limited impact as a biomaterial primarily due to difficulty in generating a pure product that is free of contaminating globular proteins. As a result, model polymers based upon the elastin pentapeptide sequence have been synthesized by solution chemistry and solid-phase approaches and recently by more efficient recombinant genetic engineering methodologies.^[48,49] As materials, these protein polymers have been processed into elastomeric hydrogels of various forms including sheets and tubular constructs by chemical, enzymatic, and γ -irradiation mediated cross-linking of protein solutions.^[32] Recently, the successful generation of fibers from elastin-mimetic peptide polymers has established a unique opportunity for the creation of fiber-reinforced composites and network structures that may lead to the engineering of improved human tissue constructs as well as completely synthetic artificial organs with enhanced clinical performance capabilities.^[50,51] In addition, the incorporation of cell binding sites into elastin-mimetic protein polymers further extends the functionality of these materials. For example, both Panitch et al.^[52] and Nicol et al.^[53] have demonstrated promotion of cell adhesion and growth when such polymers are cast from solution onto otherwise nonadhesive substrates.

The use of these elastin mimetic polymers as drug delivery devices is also under active investigation. By introduction of specific cross-links and amino acid sequences, the rate of degradation and thus drug release can be controlled. In addition, these materials are easily processable and compliant, and can be designed to swell and contract by manipulation of the polymer's inverse temperature transition (T_i).^[32] Motifs may be incorporated that allow for gel-formation, stimuli-responsive characteristics, biodegradation, and biorecognition.^[54] Recently, protein polymers derived from elastin-mimetic protein sequences have been expressed using such recombinant techniques as formation of diblock and triblock copolymers. Blocks that display different phase separation behavior facilitate the generation of a thermally reversible, self-assembled gel.^[55] The capacity of these materials to sense alterations in their chemical microenvironment and release a drug or elicit a mechanical response may enhance the effectiveness of controlled drug delivery.

Finally, elastin and elastin mimetic polymers also can be viewed as molecular machines that are able to convert energy from one form or location to another. Through the manipulation of the inverse temperature

transition (T_i) by either altering relative protein hydrophobicity, chemical-electromagnetic potential, or mechanical force, free energy transduction can be facilitated and controlled.^[33] For example, films of elastin-mimetic proteins have been used to create proton-driven machines by the addition of a glutamic acid residue (VPGE_G).^[33] As mechanochemical engines, elastin protein polymers are only constrained by the specific chemical energy selected to drive the coacervation phenomenon.

ARTICLES OF FURTHER INTEREST

Biologic and Synthetic Apatites; Collagen; Extracellular Matrix Scaffolds; Small-Angle Light Scattering Methods for Soft Connective Tissue Structural Analysis; Tissue Engineering of Blood Vessel; Tissue Engineering Scaffolds

REFERENCES

1. Cao, Q.; Wang, Y.; Bayley, H. Sequence of abductin, the molluscan "rubber" protein. *Curr. Biol.* **1997**, *7* (11), R677–R678.
2. Sandberg, L.; Soskel, N.; Leslie, J. Elastin structure, biosynthesis, and relation to disease states. *N. Engl. J. Med.* **1981**, *304* (10), 566–579.
3. Deming, T. Mussel byssus and biomolecular materials. *Curr. Opin. Chem. Biol.* **1999**, *3* (1), 100–105.
4. Hayashi, C.; Lewis, R. Evidence from flagelliform silk cDNA for the structural basis of elasticity and modular nature of spider silks. *J. Mol. Biol.* **1998**, *275* (5), 773–784.
5. Hayashi, C.Y.S.N.; Lewis, R.V. *Int. J. Biol. Macromol.* **1999**, *24*, 271–275.
6. Indik, Z.; Yeh, H.; Ornstein-Goldstein, N.; et al. Alternative splicing of human elastin mRNA indicated by sequence analysis of cloned genomic and complementary DNA. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84*, 5680–5684.
7. Vrhovski, B.; Weiss, A. *Biochemistry of tropoelastin.* *Eur. J. Biochem.* **1998**, *258*, 1–18.
8. Johnson, D.; Robson, P.; Hew, Y.; Keeley, F. Decreased elastin synthesis in normal development and in long-term aortic organ and cell cultures is related to rapid and selective destabilization of mRNA for elastin. *Circ. Res.* **1995**, *74*, 1107–1113.
9. Partridge, S. Elastin, biosynthesis and structure. *Gerontologia* **1969**, *15*, 85–100.
10. Sandberg, L.; Weissman, N.; Smith, D. The purification and partial characterization of a soluble elastin-like protein from copper deficient porcine aorta. *Biochemistry* **1969**, *8* (7), 2940.
11. Pinnell, S.; Martin, G. The cross-linking of collagen and elastin: Enzymatic conversion of lysine in peptide linkages to alpha-amino adipic-delta-semialdehyde (allysine) by an extract from bone. *Proc. Natl. Acad. Sci. U. S. A.* **1968**, *61*, 708–716.
12. Gray, W.; Sandberg, L.; Foster, J. Molecular model for elastin structure and function. *Nature* **1973**, *246*, 461–466.
13. Hoeve, C.; Flory, P. The elastic properties of elastin. *Biopolymers* **1974**, *13*, 677–686.
14. Yeh, H.; Anderson, N.; Ornstein-Goldstein, N.; et al. Structure of the bovine elastin gene and S1 nuclease analysis of alternate splicing of elastin messenger RNA in the bovine nuchal ligament. *Biochemistry* **1989**, *28*, 2365–2370.
15. Bressan, G.; Argos, P.; Stanley, K. Repeating structures of chick tropoelastin revealed by complimentary DNA cloning. *Biochemistry* **1987**, *26*, 1497–1503.
16. Pierce, R.; Deak, S.; Stolle, C.; Boyd, C. Heterogeneity of rat tropoelastin mRNA revealed by cDNA cloning. *Biochemistry* **1990**, *29*, 9677–9683.
17. Weiss-Fogh, T.; Anderson, S. New molecular model for the long range elasticity of elastin. *Nature* **1970**, *227*, 718–721.
18. Rosenbloom, J.; Abrams, W.; Mecham, R. Extracellular matrix 4: The elastic fiber. *FASEB J.* **1993**, *7*, 1208–1218.
19. Debelle, L.; Tamburro, A. Elastin: Molecular description and function. *Int. J. Biochem. Cell Biol.* **1999**, *31* (2), 261–272.
20. Urry, D.; Shaw, R.; Prasad, K. Polypentapeptide of elastin: Temperature dependence of ellipticity and correlation with elastomeric force. *Biochem. Biophys. Res. Commun.* **1985**, *130* (1), 50–57.
21. Darwin, B.; Alonso, O.; Bennion, B.; Daggett, V. Hydrophobic hydration is an important source of elasticity in elastin-based polymers. *J. Am. Chem. Soc.* **2001**, *123* (11991–11998).
22. Dorrington, K.; Grut, W.; McCrum, N. Mechanical state of elastin. *Nature* **1975**, *255* (476–478).
23. Torchia, D.; Peiz, K. Mobility of elastin chains as determined by ¹³C nuclear magnetic resonance. *J. Mol. Biol.* **1973**, *76*, 419–424.
24. Gosline, J. Hydrophobic interaction and a model for the elasticity of elastin. *Biopolymers* **1978**, *17*, 677–695.
25. Lelj, F.; Tamburro, A.; Villani, V.; et al. Molecular dynamics study of the conformational behavior of a representative elastin building block: Boc-Gly-Val-Gly-Gly-Leu-OMe. *Biopolymers* **1992**, *32*, 161–172.
26. Tamburro, A.; Quantieri, V.; Pandolfo, L.; Scopa, A. Synthetic fragments and analogues of elastin. II. Conformational studies. *Biopolymers* **1990**, *29*, 855–870.
27. Villani, V.; Tamburro, A. Conformational modeling of elastin tetrapeptide Boc-Gly-Leu-Gly-Gly-NMe by molecular dynamics simulations with improvements to the thermalization procedure. *J. Biomol. Struct. Dyn.* **1995**, *12*, 1173–1202.
28. Martino, M.; Bavoso, A.; Saviano, M.; et al. Structure and dynamics of elastin building blocks. Boc-LG-OEt, Boc-VGG-OH. *J. Biomol. Struct. Dyn.* **1998**, *1998* (15), 861–875.
29. Broch, H.; Moulabbi, M.; Vasilescu, D.; Tamburro, A.M. Quantum molecular modeling of the elastic

- tetrapeptide Val-Pro-Gly-Gly. *J. Biomol. Struct. Dyn.* **1998**, *15*, 1073–1091.
30. Villani, V.; Tamburro, A. Conformational chaos of an elastin-related peptide in aqueous solution. *Ann. N.Y. Acad. Sci.* **1997**, *879*, 284–287.
 31. Urry, D.W.; Luan, C.H.; Peng, S.Q. Molecular Biophysics of Elastin Structure, Function and Pathology. In *Symposium on the Molecular Biology and Pathology of Elastic Tissues*; Chadwick, D.J.P.S., Ed.; John Wiley and Sons Ltd.: Windsor Golf Club, Nairobi, Kenya, 1994; Vol. 192, 4–22.
 32. Urry, D.W. Physical chemistry of biological free energy transduction as demonstrated by elastic protein-based polymers. *J. Phys. Chem., B* **1997**, *101*, 11007–11028.
 33. Urry, D.W. Molecular machines: How motion and other functions of living organisms can result from reversible chemical changes. *Angew. Chem., Int. Ed. Engl.* **1993**, *32*, 819–841.
 34. Bressan, G.M.; Castellani, I.; Giro, G.M.; et al. Banded fibers in tropoelastin coacervates at physiologic temperatures. *J. Ultrastruct. Res.* **1983**, *82*, 335–340.
 35. Long, M.M.; Rapka, R.S.; D.V.; et al. Spectroscopic and electron micrographic studies on the repeat tetrapeptide of tropoelastin: Val-Pro-Gly-Gly. *Arch. Biochem. Biophys.* **1980**, *201*, 445–452.
 36. Pasquali-Ronchetti, I.; Fornieri, C.; Baccarani-Contri, M.; Volpin, D. The ultrastructure of elastin revealed by freeze-fracture electron microscopy. *Micron* **1979**, *10*, 89–99.
 37. Sage, H.; Gray, W.R. Studies on the evolution of elastin-II. *Histology. Comp. Biochem. Physiol.* **1980**, *66B*, 13–22.
 38. Pasquali-Ronchetti, I.B.-C.M.; Young, R.D.; Vogel, A.; Steinmann, B.; Royce, P.M. Ultrastructural analysis of skin and aorta from a patient with Menkes disease. *Exp. Mol. Pathol.* **1994**, *61*, 36–57.
 39. Quaglino, D.; Fornieri, C.; Nanney, L.; Davidson, J. Extracellular matrix modifications in rat tissues of different ages. Correlations between elastin and collagen type I mRNA expression and lysyl oxidase activity. *Matrix* **1993**, *13*, 481–490.
 40. Vitellaro-Zuccarello, L.; Cappelletti, S.; Dal Pozzo-Rossi, V.; Sari-Gorla, M. Stereological analysis of collagen and elastic fibers in the normal human dermis: Variability with age, sex and body region. *Anat. Rec.* **1994**, *238*, 153–162.
 41. Pasquali-Ronchetti, I.; Fornieri, C.; Baccarani-Contri, M.; Quaglino, D. Ultrastructure of Elastin. In *Symposium on the Molecular Biology and Pathology of Elastic Tissues*; Chadwick, D.J.P.S., Ed.; John Wiley and Sons Ltd.: Windsor Golf Club, Nairobi, Kenya, 1994; Vol. 192, 31–42.
 42. Nidhi, N.; Chilkoti, A. Interfacial phase transition of an environmentally responsive elastin biopolymer adsorbed on functionalized gold nanoparticles studied by colloidal surface plasmon resonance. *J. Am. Chem. Soc.* **2001**, *123*, 8197–8202.
 43. Rama Rao, G.; Balamurugan, S.; Meyer, D.; Hybrid bioinorganic smart membranes that incorporate protein based molecular switches. *Langmuir* **2002**, *18*, 1819–1824.
 44. Betre, H.; Setton, L.; Meyer, D.; Chilkoti, A. Characterization of a genetically engineered elastin-like polypeptide for cartilaginous tissue repair. *Biomacromolecules* **2002**, *3*, 910–916.
 45. Hyun, J.; Chilkoti, A. Micropatterning biological molecules on a polymer surface using elastomeric microwells. *J. Am. Chem. Soc.* **2001**, *123*, 6934–6944.
 46. Lillie, M.; Gosline, J. The viscoelastic basis for the tensile strength of elastin. *Int. J. Biol. Macromol.* **2002**, *30* (2), 119–127.
 47. Urry, D.; Parker, T.; Reid, M.; Gowda, D. *J. Bioact. Compat. Polym.* **1991**, *6*, 263–282.
 48. McMillan, R.; Lee, T.; Conticello, V. Rapid assembly of synthetic genes encoding protein polymers. *Macromolecules* **1999**, *32* (11), 3643–3648.
 49. McPherson, D.T.X.J.; Urry, D.W. Product purification by reversible phase transition following *Escherichia coli* expression of genes encoding up to 251 repeats of the elastomeric pentapeptide GVGVP. *Protein Expr. Purif.* **1996**, *7*, 51–57.
 50. Nagapudi, K.B.W.; Huang, L.; McMillan, R.A.; Conticello, V.P.; Chaikof, E.L. Photomediated solid-state crosslinking of an elastin-mimetic recombinant protein polymer. *Macromolecules* **2002**, *35*, 1730–1737.
 51. Huang, L.; McMillan, R.; Apkarian, R.; et al. Generation of synthetic elastin-mimetic small diameter fibers and fiber networks. *Macromolecules* **2000**, *33*, 2989–2997.
 52. Panitch, A.; Yamaoka, T.; Fournier, M.; Design and biosynthesis of elastin-like artificial extracellular matrix proteins containing periodically spaced fibronectin CS5 domains. *Macromolecules* **1999**, *32* (5), 1701–1703.
 53. Nicol, A.; Gowda, D.; Parker, T.; Urry, D. Elastomeric polytetrapeptide matrices: Hydrophobicity dependence of cell attachment from adhesive (GGIP)_n to non-adhesive (GGAP)_n even in serum. *J. Biomed. Mater. Res.* **1993**, *27*, 801–810.
 54. Megeed, Z.; Cappello, J.; Ghandehari, H. Genetically engineered silk-elastinlike protein polymers for controlled drug delivery. *Adv. Drug Deliv. Rev.* **2002**, *54* (8), 1075–1092.
 55. Wright, E.; Conticello, V. Self-assembly of block copolymers derived from elastin-mimetic polypeptide sequences. *Adv. Drug Deliv. Rev.* **2002**, *54*, 1057–1074.

Elastomers, Biodegradable

John J. Stankus

*McGowan Institute for Regenerative Medicine, University of Pittsburgh, and
Department of Chemical and Petroleum Engineering, University of Pittsburgh, Pittsburgh,
Pennsylvania, U.S.A.*

Jianjun Guan

*McGowan Institute for Regenerative Medicine, University of Pittsburgh, Pittsburgh,
Pennsylvania, U.S.A.*

William R. Wagner

*McGowan Institute for Regenerative Medicine, University of Pittsburgh,
Department of Chemical and Petroleum Engineering, University of Pittsburgh, and
Department of Bioengineering, University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.*



INTRODUCTION

The attraction of biodegradable elastomers for applications in tissue engineering and drug delivery is the latitude that they provide to the biomaterials engineers in achieving mechanical properties that can approximate soft tissue behavior. To achieve these properties, a variety of synthetic biodegradable elastomers have been developed. Two major groups are discussed, polyesters and polyurethanes, although the majority of examples of the latter contain polyester segments. The synthesis, mechanical properties, and biodegradation are generally addressed for these polymers, and, when appropriate, biocompatibility, processing, and applications are also discussed.

The mechanical properties of elastomers include high breaking strains, low stiffness, and high resilience, which is defined as the ability of a material to quickly return to its original shape after being deformed. Synthetic polymers that possess elastic properties are applicable to a variety of biomaterials applications, including drug delivery vehicles and scaffolds for the engineering of tissues with elastic properties. Many of the shortcomings faced with natural materials are of less concern with synthetic molecules. Immunogenicity of the molecular segments is likely absent, degradation is generally designed to occur through hydrolysis and may be less susceptible to enzymatic environment, and processing of thermoplastic elastomers can achieve varied three-dimensional architectures. Further, the options for molecular design are many and varied to achieve desired mechanical and chemical properties. Among the limitations and concerns with synthetic materials are the potential toxicity of degradation products or remnant synthetic by-products, material stimulation of inflammatory processes, and extra

processing or manipulation that is necessary to impart bioactivity.

POLYESTER BIODEGRADABLE ELASTOMERS

Polyesters have been the most commonly utilized biodegradable synthetic polymers for biomedical applications. The hydrolytically labile ester bond imparts degradability in vivo with the relative accessibility of water to this bond being dependent upon the molecular architecture, in addition to the micro- and macrostructure of the material. Polyglycolide (PGA), polylactide (PLA), polycaprolactone (PCL), and poly[(R)-hydroxy butyrate] (R-PHB) are aliphatic polyesters that have found use in a variety of settings. These polymers are usually crystalline or have high glass transition temperatures (T_g) leading to high tensile strength and low elongation at break (Table 1). Mechanical properties can be softened by copolymerization with other ester monomers to alter the crystalline behaviour of the bulk polymers. Examples are poly(lactide-co-glycolide) (PLGA),^[1] poly(glycolide-co- ϵ -caprolactone) (50/50),^[2] poly(glycolide-co-lactide-co- ϵ -caprolactone),^[3] and PLGA-b-PEG-b-PLGA in which PEG is poly(ethylene glycol).^[4] Most of these copolymers have attractive tensile strengths and high elongation at break, but they exhibit plastic deformation due to their amorphous linear molecular structure. For use as elastomers, cross-linking is necessary to impart resilience. However, the cross-linking of aliphatic polyesters is not readily achieved under ordinary chemical or radiation conditions because of the risk of decomposition.^[5] Chemically cross-linked polyesters can be achieved by employing multifunctional monomers and polycondensation

Table 1 Mechanical properties of biodegradable polyesters

Polymer (molecular weight)	T _g (°C)	T _m (°C)	Tensile strength (MPa)	Tensile modulus (MPa)	Elongation at break (%)
PGA(50,000) ^a	37	223	270	NA	NA
PLLA(50,000) ^b			28	1200	6.0
PDLLA(107,000) ^b			29	1900	6.0
PCL(107,300) ^c	–	63	7.0	343	120.2
R-PHB(189,000) ^c	–	173	22.5	1362.2	2.2
PBS(67,000) ^d	–36.6	115.9	37.2	1901	47

^aPihlajamaki, H.; Makela, E.A.; Ashammakhi, N.; Viljanen, J.; Patiala, H.; Rokkanen, P.; Pohjonen, T.; Tormala, P.; Joukainen, A. Strength retention of drawn self-reinforced polyglycolide rods and fixation properties of the distal femoral osteotomies with these rods. An experimental study on rats. *J. Mater. Sci., Mater. M* **2002**, *13* (4), 389–395.

^bEngelberg, I.; Kohn, J. Physico-mechanical properties of degradable polymers used in medical applications: A comparative study. *Biomaterials* **1991**, *12*, 292–304.

^cTsuji, H.; Suzuyoshi, K. Environmental degradation of biodegradable polyesters 1. Poly(epsilon-caprolactone), poly[(R)-3-hydroxybutyrate], and poly(L-lactide) films in controlled static seawater. *Polym. Degrad. Stabil.* **2002**, *75* (2), 347–355.

^dJin, H.J.; Park, J.K.; Park, K.H.; Kim, M.N.; Yoon, J.S. Properties of aliphatic polyesters with *n*-paraffinic side branches. *J. Appl. Polym. Sci.* **2000**, *77* (3), 547–555.

techniques. Another method is to introduce physical cross-links into the polymer by microphase separation of the backbone into hard and soft segments.

Chemistry

Synthesis of cross-linked polyesters

Cross-linked polyesters can be prepared by polycondensation of trifunctional alcohol and bifunctional acid monomers of low molecular weight as demonstrated recently by Wang et al. with glycerol and sebacic acid.^[6] Here, it was suggested that the presence of hydroxyl groups on at least one monomer could lead to hydrogen bonding interactions that influenced mechanical properties. Cross-linked polyester can also be prepared by polycondensation of multifunctional hydroxyl-telechelic polyester oligomer and multifunctional isocyanate.^[7] Usually, difunctional hydroxyl-telechelic polyester oligomers react with triisocyanate, and trifunctional hydroxyl-telechelic polyester oligomers react

Table 2 Tensile properties of poly(ε-caprolactone-co-δ-valerolactone) networks

Oligomer molecular weight (MPa)	Modulus (MPa)	Tensile strength (%)	Elongation at break
900	2.59	1.96	193
1500	1.95	1.92	221
1900	1.66	1.71	275
2400	1.21	1.50	313
3600	1.04	1.16	407

Data adapted from Storey, et al. 1991. Ref. [7].

with diisocyanate. Multifunctional polyesters are prepared by ring-opening polymerization (ROP) of cyclic monomers initiated by multifunctional alcohol. The monomers used are ε-caprolactone, δ-valerolactone, lactide, and glycolide. Storey et al. described the use of 1,6-hexanediol to initiate the copolymerization of ε-caprolactone with δ-valerolactone to form difunctional hydroxy-terminated poly(ε-caprolactone-co-δ-valerolactone). The polyester oligomers were cross-linked by triphenylmethane triisocyanate (TTI). Trifunctional hydroxyl-telechelic polyester oligomers were synthesized by copolymerization of monomers initiated by trimethylolpropane (Table 2). The networks of polyesters were formed by reacting with toluene-2,6-diisocyanate (TDI). The resulting cross-linked polyesters were amorphous with glass transition temperatures ranging from 21°C to –45°C.^[7,8]

Synthesis of linear polyester elastomers

Linear polyester elastomers have microphase separated structures with crystalline hard segments and amorphous soft segments. In this structure, hard segments act as cross-linkers whereas soft segments provide the flexibility. The properties of the polyester elastomer are controlled by the ratio of hard and soft segments.^[5] In general, linear polyester elastomers are prepared by random or block copolymerization of monomers. The ratio of crystalline and amorphous segments can be controlled by the feed ratio of monomers. Polyester elastomers are usually synthesized by ring-opening polymerization (ROP) instead of condensation polymerization as the former yields high molecular weight polymers under relatively mild conditions. ROP can be carried out with no or very

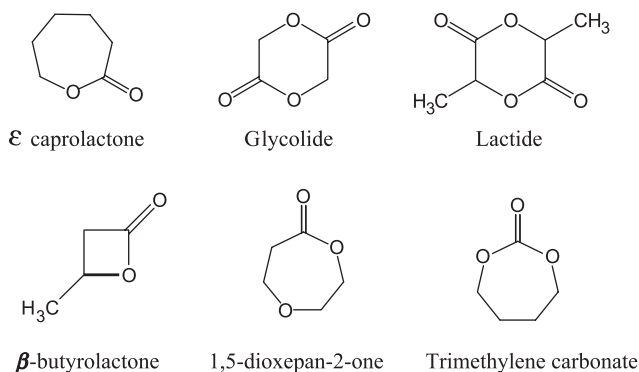


Fig. 1 Structures of various monomers used in biodegradable polyesters.

limited side reactions, which makes it possible to control properties like molecular weight and polydispersity. ROP can be divided into free radical polymerization, anionic polymerization, carbocationic polymerization, and coordinative ring-opening polymerization.^[9] Among them, coordinative ring-opening polymerization is the most popular method because the polymerization conditions are easy to control. Coordinative ring-opening polymerization can be conducted with or without alcohol as an initiator.^[9] Tin(II) or tin(IV) compounds such as stannous octanoate ($\text{Sn}(\text{oct})_2$) and dibutyltin oxide have been used as catalysts.^[9] Lactones (ϵ -caprolactone, [R,S]- β -butyrolactone) and cyclic diesters (lactides, glycolide, trimethylene carbonate) are used as monomers (Fig. 1).

PHA and copolymers

Polyhydroxyalkanoate (PHA) elastomers can be synthesized by bacterial fermentation, transgenic microorganisms, and plants.^[10] By incorporating various functional monomers in the polymer sequence during

the biosynthesis, one can obtain different PHAs with various physical properties ranging from crystalline polymers to elastomers. PHAs can also be obtained by chemical synthesis. Polymerization of the racemic monomer [R,S]- β -butyrolactone with distannoxanes as initiators can obtain an atactic or syndiotactic-rich poly(3-[R,S]-hydroxybutyrate) ([R,S]-PHB) with different stereoregularities.^[11]

For [R,S]-PHB, introduction of a biodegradable hard segment can give a wholly biodegradable creep-resistant elastomer. Hiki et al. prepared ABA block copolymers with [R,S]-PHB as the soft segment (B) and poly(L-lactide) (PLLA) as the hard segment (A).^[5] The polymerization was carried out by a two-step method. In the first step, hydroxy-terminated poly([R,S]- β -butyrolactone) was prepared by the reaction of (R,S)- β -butyrolactone with 1,4-butanediol. In the second step, L-lactide was initiated by hydroxyl groups of [R,S]-PHB to form PLLA blocks (Fig. 2). Two different catalysts were used to prepare [R,S]-PHB and PLLA blocks. Distannoxane was used for [R,S]-hydroxybutyrate polymerization as it is a high transesterification catalyst whereas $\text{Sn}(\text{oct})_2$ was used for polymerization of L-lactide for the chain transfer reaction to occur.^[5]

Polyglycolide, polylactide, poly(ϵ -caprolactone), and copolymers

As mentioned previously, PLA, PGA, and PCL are high-strength materials with low distensibility. Modification of these polymers by incorporating other monomer(s) can alter the mechanical properties from rigid materials to thermoplastic elastomers, which can be used in mechanically challenging environments. Elastomers containing microphase separation can be prepared by random or block copolymerization. Microphase separated poly(glycolide-co- ϵ -caprolactone) has

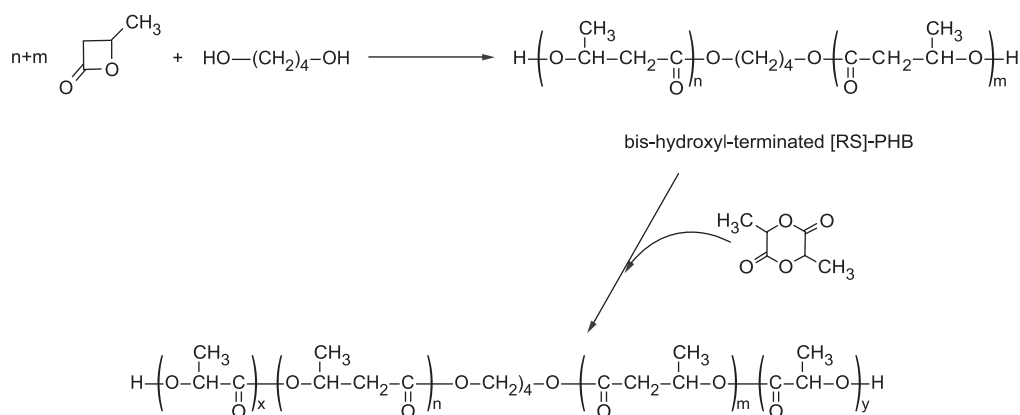


Fig. 2 Ring-opening polymerization of [RS]-beta-butyrolactone and copolymerization with L-lactide. Reproduced from Ref. [5].

been prepared by random polymerization of glycolide and ϵ -caprolactone with zirconium(IV) acetylacetonate or zirconium(IV) chloride as initiators.^[12] Poly(lactide-co- ϵ -caprolactone) elastomers have been prepared by random copolymerization of lactide and ϵ -caprolactone with tin(II) as initiator.^[13] The microphase separated chain structure was dependent on both composition of monomers and reaction temperature.

1,5-dioxepan-2-one (DXO) has been used as a versatile component in biodegradable polymers. Poly (DXO) is amorphous, but introduction of a crystalline segment such as PCL and PLLA can lead to microphase separated elastomers. PCL-b-PDXO-b-PCL block copolymers have been prepared by the sequential addition of different monomers to a living polymerization system initiated with aluminum isopropoxide in tetrahydrofuran (THF) or toluene solution.^[14] PLLA-b-PDXO-b-PLLA block copolymers were prepared by using controlled ROP in a two-step method. In the first step, polymerization of DXO using 1,1,6,6-tetra-*n*-butyl-1,6-distanna-2,5,7,10-tetraoxacyclodecane was processed to obtain PDXO. In a second step, lactide was added and polymerization was conducted at 60°C in chloroform.^[15]

Similar to PDXO, poly(1,3-dioxan-2-one) (PTMC) is amorphous. Random copolymerization of trimethylcarbonate (TMC) and ϵ -caprolactone can obtain a blocklike copolymer since the reactivity of ϵ -caprolactone is higher than TMC.^[16] When the ϵ -caprolactone content is higher than 70%, the copolymer exhibits elastomeric characteristics. Microphase separated block copolymers based on PTMC and PLLA were prepared by Kim et al.^[17] The hydroxy-terminated PLLA-b-PTMC-b-PLLA was prepared by polymerization of *L*-lactide with the PTMC diol chain using Sn(oct)₂ as catalyst. Next, the polymer was chain extended by hexamethylene diisocyanate (HMDI).

Mechanical Properties

Mechanical properties of polyester elastomers depend on several factors, such as the molecular composition, crystallinity, and molecular weight. For cross-linked polyesters, the mechanical properties depend on the cross-linking density and molecular weight between cross-links as demonstrated in the polyester synthesized from glycerol and sebacic acid by Wang et al.^[6] For cross-linked elastomers synthesized from difunctional hydroxy-terminated poly(caprolactone-co-valerolactone) oligomers, it was found that the tensile strength decreased and elongation at break increased with increased oligomer length (Table 2). This trend resulted from increased interactions between molecular chains as the molecular weight between

Table 3 Mechanical properties of poly(trimethylene carbonate-co- ϵ -caprolactone)

ϵ -caprolactone (%)	Young's modulus (MPa)	Tensile strength (MPa)	Elongation at break (%)
75	4.6	0.1	236
86	123	46	1008
90	252	40	906

Data adapted from Pego et al. 2001. Ref. [16].

cross-links increased from decreasing molecular weight of oligomer.^[8]

For random copolymerized linear polyester elastomers of caprolactone, glycolide, or lactide, the feed ratio and reaction temperature can affect the microstructure of resulting polymers and thus their mechanical properties. The microphase separated elastomers can be prepared only at certain monomer ratios^[18] and reaction temperatures.^[5] For example, a copolymer made from 85 wt% of *DL*-lactide and 15 wt% ϵ -caprolactone is a thermoplastic elastomer; whereas a copolymer prepared from 90 wt% of *DL*-lactide and 10 wt% ϵ -caprolactone is a rigid thermoplastic. The poly(glycolide-co- ϵ -caprolactone) copolyesters are elastomers when the reaction temperature is 100°C; whereas more randomized structures are formed when the temperature is 150°C. For copolymers of ϵ -caprolactone and TMC, the microphase-separated structure can be formed at a wide feed ratio range. When the ϵ -caprolactone ratio is higher than 70 wt%, the copolymers are semicrystalline and have microphase separated structures. The increase in ϵ -caprolactone content increased the crystallinity and thus increased the Young's modulus (Table 3).

Block copolyester elastomers possess combined properties of both homopolymers. The elastomeric mechanical properties are enhanced by microphase separation. Because the soft segment provides flexibility and the hard segment provides mechanical strength, tensile strength generally increases with an increased length of hard-segment block and crystallinity, while the strain increases with the length of soft-segment block as shown by Lofgren et al., who synthesized the well-defined PCL-b-PDXO-b-PCL triblock copolyesters.^[14]

Biodegradation

The biodegradation rate of cross-linked polyester elastomers depends on both the nature of monomers and the cross-linking density. The previously discussed cross-linked polyester based on glycerol and sebacic acid possessed low cross-linking density and

crystallinity. This polymer has approximately a 17% weight loss after 60 days and could be totally absorbed in 60 days *in vivo*.^[6] Storey et al. prepared a series of cross-linked amorphous polyesters from triol polyester oligomers and TDI. The triol polyesters were poly(glycolide-co-TMC) (PGTMC), poly(lactide-co-TMC) (PLTMC), and poly(ϵ -caprolactone-co-TMC) (PCTMC). It was found that the degradation rates, in order of fastest to slowest, were PGTMC > PLTMC > PCTMC, which follows the hydrophilicity order of glycolide > lactide > caprolactone.^[7]

For linear polyester elastomers, the degradation rate depends on the chemical composition, the nature of monomers as well as the individual microphases. The degradation rate is mainly influenced by the amorphous phase because of reduced water access to the crystalline phase. Increases in the amorphous content can increase the degradation rate.^[11] As the rate of hydrolytic degradation of the ester linkage is higher than the carbonate linkage,^[9] incorporation of carbonate functionalities can decrease the degradation rate.

BIODEGRADABLE POLYURETHANES

Polyurethanes are the most economically important elastomers for a variety of commercial applications. Thus, it is not surprising that this class of polymers has been the focus of many groups seeking to develop biodegradable elastomers for medical applications. Polyurethane and poly(urethane urea) thermoplastic elastomers derive their attractive mechanical properties from microphase separation between hard and soft segments. The selection of hard and soft segments and the complexity possible within these segments offers polymer chemists a myriad of options in achieving material design objectives. Although urethane and urea linkages are susceptible to degradation *in vivo*, efforts to make biodegradable polyurethanes have incorporated more labile bonds into the polymer backbone, usually by introducing polyester segments.

Polyurethane Synthesis

Polyurethanes are commonly synthesized by one of two methods. The first involves the reaction of bis-chloroformates with diamines. The second—and most important method—is the reaction of diisocyanates with dihydroxy compounds, which has the advantage of no by-products. Diisocyanates can also react with diamines to form polyureas, with carboxylic acids to form polyamides, and with water to form polyurea foam. Side reactions in polyurethane synthesis can occur when unreacted isocyanate groups react with the growing urethane or urea chains to yield allophanate and biuret groups, respectively. These reactions yield polymers with unwanted cross-linking and generally less attractive mechanical properties. Biodegradable polyurethane elastomers can also consist of a network of chemically cross-linked chains. These elastomers can be synthesized by using monomers that are trifunctional or of higher functionalities compared with the difunctional monomers used in synthesizing linear (thermoplastic) polyurethanes.^[20–22] Cross-linked polyurethanes have the disadvantage of not being amenable to processing by thermal or solvent-based techniques.

Polyurethane elastomers are usually constructed from long-chain polyester or polyether diols and diisocyanates with short-chain diols or diamines utilized as chain extenders. The reaction is commonly carried out in two steps with the initial formation of a prepolymer of the polyester and diisocyanate that is subsequently chain-extended to a high molecular weight (Fig. 3).

Alternatively, the reaction can occur in one step (one-shot polymerization) when all of the monomers are mixed together at once. The longer-chain polyester or polyether in the resulting polyurethane is referred to as the soft segment, and the diisocyanate and chain extender are referred to as the hard segment. Common polyester soft segments used in biodegradable polyurethanes include PCL,^[23–25] PLA,^[26] PGA,^[27] and

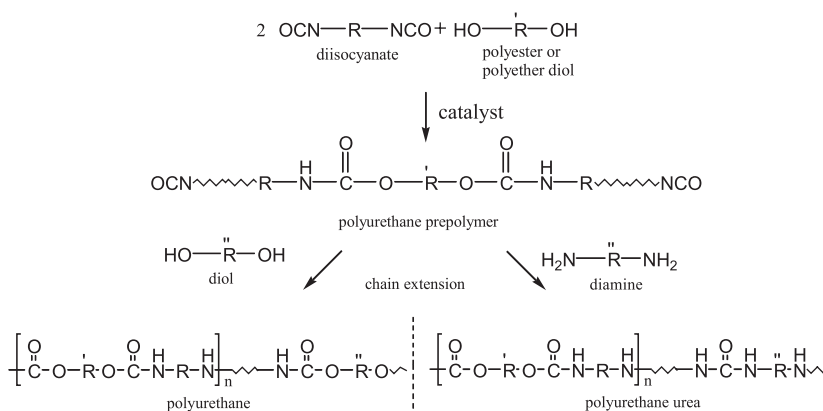


Fig. 3 Two step reaction to form polyurethane or polyurethane ureas.

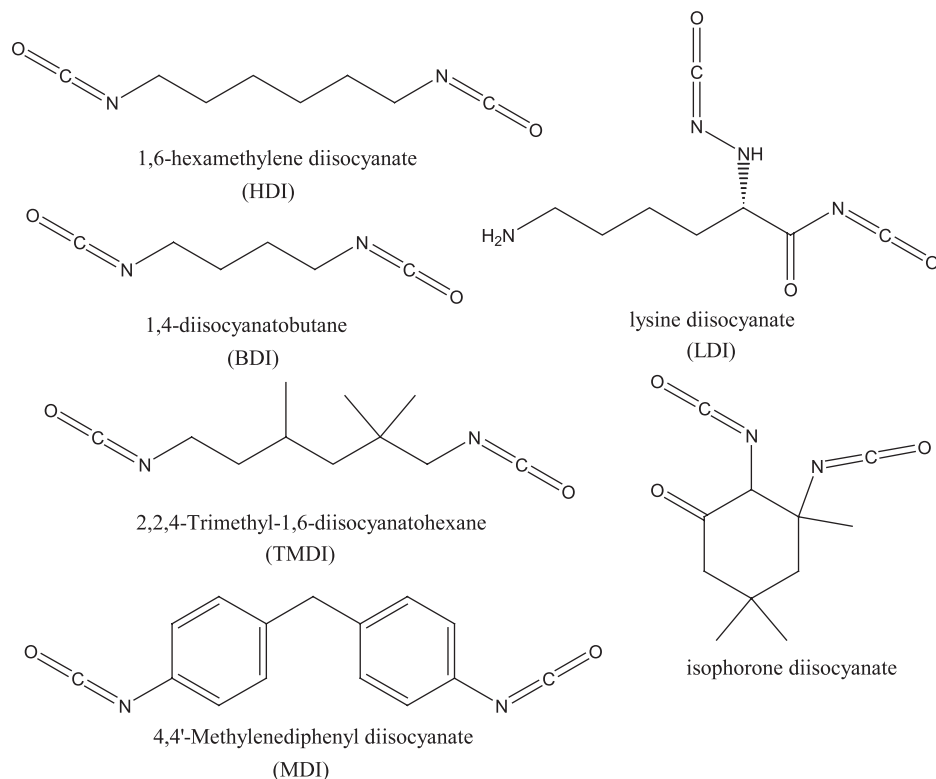


Fig. 4 Chemical structures of diisocyanates used as monomers in biodegradable polyurethanes.

PHB.^[27,28] Common diisocyanates used in constructing biodegradable polyurethanes are 1,6-hexamethylene diisocyanate (HDI),^[29,30] 1,4-diisocyanate (BDI),^[23–25] lysine diisocyanate (LDI),^[31] isophorone diisocyanate, and 2,2,4-trimethyl-1,6-diisocyanate (TMDI).^[27] The chemical structures of these diisocyanates are shown in Fig. 4 along with the chemical structures of 4,4'-methylenediphenyl diisocyanate (MDI) and toluene diisocyanate (TDI), which are commonly employed industrial diisocyanates. However, there is speculation that MDI and TDI degradation products—4,4'-methylene aniline (MDA) and toluene diamine (TDA)—are carcinogenic.^[32–34]

The chemical and physical properties of polyurethanes can be varied by choice of appropriate hard and soft segments. Table 4 summarizes the different biodegradable polyurethanes that have been reported in the literature based on their soft and hard segment composition. The majority of these polymers contain PCL in the soft segment. Gorna et al. bulk-polymerized PCL with aliphatic HDI and 1,4-butanediol (BDO) as a chain extender.^[29] Other studies have introduced PEG along with PCL into the soft segment with or without BDO chain extenders.^[30,35] High-molecular weight polyurethanes based on PCL soft segments with BDI were chain-extended with BDO, BDO-b-BDI-b-BDO blocks, or BDI-b-BDO-b-BDI-b-BDO-b-BDI blocks.^[24,25] The larger-chain extenders were used to prevent transesterification reactions from creating a large size distribution of hard segments. de

Groot et al. synthesized a series of linear polyurethane ureas based on a PCL soft segment and hard segments consisting of either LDI, HDI, or BDI that were chain-extended with putrescine.^[36] Similar poly(ester-urethane ureas) were synthesized by Guan et al. using putrescine and lysine ethyl ester as chain extenders.^[23]

Bioerodible polyurethanes have also been synthesized with L-lactide in the soft segment. Kylma and Seppala investigated comonomers of lactic acid and caprolactone in poly(ester-urethanes) based on HDI.^[26] Condensation copolymerization of lactide and ϵ -caprolactone at different feed ratios took place with stannous octoate as a catalyst. High molecular weight polyurethanes were produced by chain extending with HDI. Polyurethanes based on a 50/50 poly(L-lactide-co- ϵ -caprolactone) prepolymer with BDI were synthesized.^[36] Because the lactide bond is susceptible to aminolysis, chain extension was not possible with putrescine. Furthermore, the use of BDO as a chain extender resulted in transesterification that led to poor mechanical properties. For these reasons, an isocyanate-terminated chain extender, BDO-b-BDI-b-BDO, was used.

Several research groups have constructed biodegradable polyurethanes and poly(urethane ureas) by introducing biodegradable functionalities into the hard segment as well as the soft segment to facilitate hydrolytic degradation. Poly(ester-urethanes) consisting of crystallizable poly(3-R-hydroxybutyrate)-co-(3-R-hydroxyvalerate) (PHB-co-PHV) hard segments

Table 4 Composition and physical properties of linear biodegradable polyurethanes and polyurethane ureas

Soft segment composition	Hard segment composition		Soft segment Tg/C	Tensile strength, MPa	Modulus (MPa)	Ultimate elongation (%)	Reference
	Diisocyanate	Chain extender or other					
PCL ^a	HDI	BDO	− 54.4 to − 40.6	12–63	13–107	460–760	[29]
PCL	LDI	putrescine	− 45	17	40	800	[31]
PCL	BDI	putrescine	− 45	29	52	1042	[31]
PCL	BDI	BDO	− 54	23.1	23.2	843	[25]
PCL	BDI	BDO.BDI.BDO	− 54	44	70	560	[24]
PCL	BDI	BDI.BDO.BDI.BDO.BDI	− 60.4	35	105	650	[24]
(50/50) L-lactide/PCL	BDI	BDO.BDI.BDO	− 5	45	60	560	[36]
PCL ^a	BDI	putrescine	− 53.2 to − 36.5	25–29	54–78	660–686	[23]
PCL ^a	BDI	lysine ethyl ester	− 54.3 to − 39.7	9.2–13	14–38	841–895	[23]
PCL ^a	LDI	p-alanine based diamine ^a	− 51.9 to − 6.2	12.5–30.8	6.6–81.9	618–676	[42]
PCL-PEG ^a	HDI	BDO	− 115.7 to − 101.5	4–48.6	7.3–58.8	120–866	[30]
PCL, PGA, PEG ^a	TMDI	PHB-co-PHV ^a	− 44 to 2	2.6–15.9	50–500	50–1250	[27]
PCL ^a	TMDI	poly(p-dioxanone)diol ^a	− 51 to − 39	17–25	34–90	540–1100	[40]

^aIndicates series of polymers synthesized with different feed concentrations or monomer molecular weights.

and amorphous segments from ϵ -caprolactone, diglycolide, and ethylene glycol homopolymers or copolymers have been reported.^[27,28] Aliphatic TMDI was the diisocyanate employed in the polymerization. Elastomeric polyurethanes have also been synthesized from poly(ϵ -caprolactone)diols and poly(*p*-dioxanone)diols (PDX) reacted with TMDI.^[27] The crystalline PDX was used as a biodegradable hard segment. Skarja and Woodhouse synthesized novel poly-(urethane ureas) based on either PCL or PEG soft segments, HDI or LDI, and with an amino acid-based chain extender based on L-phenylalanine.^[37,38]

Mechanical Properties

The soft segments impart polyurethanes with softness, flexibility, and low temperature properties whereas the hard segments reinforce or strengthen the soft segment to increase the hardness, modulus, tear strength, and high temperature properties of polyurethanes. The hard segments possess a high affinity for hydrogen bonding due to the large number of polar groups present. Urea groups provide even more opportunities

for hydrogen bonding relative to urethane groups. The combination of crystallinity and hydrogen bonding in the phase separation of polyurethanes ties together the polymer chains in physical cross-links. These physical cross-links result in a network of polymer chains that display rubberlike elasticity.

Similar to multiphase polyesters, the factors that affect mechanical properties of biodegradable polyurethanes include hard- and soft-segment composition, degree of crystallization of the hard segment and amount of microphase separation, and molecular weight of the polymer. A summary of mechanical properties for different monomer composition of biodegradable polyurethanes and poly(urethane ureas) is provided in Table 4. Gorna et al. showed that the mechanical properties of polyurethanes with similar molecular weights were dependent on the hard-segment content.^[29] Tensile strength and modulus increased with increasing hard-segment molecular weight or with shorter soft segment while the breaking strain decreased. In an extensive study, the mechanical properties of polyurethanes based on L-lactide and ϵ -caprolactone copolymers were found to be highly dependent on the initial feed ratio of the monomers.

Poly(ester-urethanes) containing a high ratio L-lactide and ϵ -caprolactone were found to be rigid with tensile modulus of 1700–2100 MPa, tensile strength of 36–47 MPa, and breaking strains of 4–7%.^[26] Polymers with low L-lactide to ϵ -caprolactone ratios were found to be highly elastomeric with a maximum tensile strength of 9 MPa and breaking strain greater than 1000%. These properties were found to be direct functions of the glass transition temperatures of the polymers. When the glass transition temperatures were above room temperature, the soft segment would be amorphous and mostly immobile, which would contribute to higher tensile strengths and lower elongation.^[26]

The polyurethanes and poly(urethane ureas) synthesized by de Groot et al. and Spaans et al. showed high flexibility and tensile strength, which results from high molecular weights and phase-separated morphology.^[24,25,31] Poly(ester-urethanes) containing lysine ethyl esters as a chain extender compared to those with putrescine as a chain extender were found to have lower mechanical properties than the latter.^[23] These properties may have been caused by the lower molecular weight of the lysine-extended polymers or because the bulky, lysine-based chain extender may have reduced the chain packing and physical cross-linking of the polymer chains.

Biodegradation

Biodegradation products and rates for polyurethanes can be controlled by the choice of the soft and hard segment. Controlling or predicting degradation rates can be very complicated and is influenced by factors similar to mechanical properties such as crystallinity and polymer molecular weight. The degradation of the common soft-segment PCL usually occurs first by hydrolytic chain scission and then by bulk mass loss of subsequently lower chain lengths.^[39] PCL alone has demonstrated some long degradation times that can extend to the order of years due to its hydrophobicity and tendency to crystallize at higher molecular weights. The in-vitro degradation of PCL-based polyurethanes in aqueous buffers have also demonstrated slow degradation times. Because most studies are conducted over limited time periods, the time scales for complete resorption of most PCL-based degradable polyurethanes both in vitro and in vivo are mostly unknown. Degradation of PCL-based polyurethanes by Gorna et al. exhibited slow degradation times with a 1–2% mass loss at 48 weeks and a 1.1–3.8% mass loss at 76 weeks.^[30] However, poly(ester ether urethanes) composed of PCL-PEG diols exhibited levels of weight loss from 1.6–76% at 48 weeks and 1.6–96% at 76 weeks. The weight loss was found to increase with

higher levels of the more hydrophilic PEG.^[30] A similar in-vitro degradation trend was observed for amino acid chain-extended, PCL-based polymers that exhibited little mass loss after 56 days compared with the PEG-based polymers that exhibited high mass loss in this period.^[37]

The in-vitro degradation of L-lactide/ ϵ -caprolactone copolymer-based polyurethanes was found to be highly dependent on the comonomer ratio. Polymers synthesized from high ratios of L-lactide to ϵ -caprolactone exhibited mass losses of more than 80% after 50 days compared with higher ϵ -caprolactone content polymers exhibiting less than 10% weight loss after 50 days.^[26] These degradation weights were related to the hydrophilicity of the monomers with lactic acid being more hydrophilic than PCL.

A very thorough investigation of the hydrolytic degradation of copoly(ester urethanes) with PHB-co-PHB hard blocks synthesized by Lendlein et al. has shown the degradation rate to be dependent on the amount and distribution of the glycolate ester bonds in the soft segments of the urethane.^[39] Here, the polymer degradation began hydrolytically, which decreased the molecular weight and lowered the breaking strain. Next, weight loss decreased the sample size and left the poly(PHB-PHV) hard segment as the principal remaining material.^[39] Due to the nature of the amorphous segment and crystalline segment of physically cross-linked polyurethanes, it is common for the soft segment to degrade at a much faster rate than the hard segment, which resulted in a nonlinear degradation rate. However, Lendlein and Langer demonstrated a linear weight loss in vitro of their polyurethane based on PCL/PDX due to the high crystallinity of the polyesters.^[40]

Biocompatibility

The use of aliphatic diisocyanates in biodegradable polyurethanes facilitates degradation and avoids the potentially carcinogenic aromatic diisocyanates, MDI and TDI. The use of LDI and BDI as hard segments is of interest because of their putative nontoxic degradation products, ethanol and L-lysine, and putrescine, respectively. A number of reports have characterized the biocompatibility of biodegradable polyurethanes by culturing cells on the polymers and performing in-vivo studies. An extensive biocompatibility study was performed on polymers prepared from PCL soft segments, PHB/PHV diol hard segments with either TMDI or LDI.^[38] Macrophages and fibroblasts were cultured on four different polyurethane compositions as well as on tissue polystyrene (TCPS). With macrophages, a mitochondrial activity assay indicated cell adhesion with polyurethanes from 65–100% of TCPS

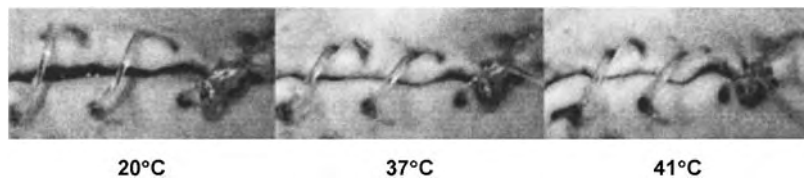


Fig. 5 Degradable shape-memory suture for wound closure. The photo series from the animal experiment shows (left to right) the shrinkage of the fiber when temperature increases. Reproduced from Ref. [40].

and a doubling time of 124–143% of TCPS.^[38] Similar results were found with fibroblasts with adhesion to the polymers 50–85% of TCPS and doubling times 125–140% of TCPS.^[38] These cells preserved their phenotypes as shown by nitric oxide (NO) tumor necrosis factor α (TNF- α) production by macrophages as well as extracellular matrix production by the cells cultured on the polymers. Adherent macrophages also responded to lipopolysaccharide by increasing NO and TNF- α production in a manner similar to cells on control surfaces.^[38]

To encourage cell adhesion, the poly(ester-urethane urea) synthesized by Guan et al. were surface-modified with radio frequency glow discharge followed by coupling of the adhesion peptide RGDS using a BDI spacer. Adhesion after one day of culture on films of the unmodified polymer was 85% of TCPS and, to films of the RGDS-modified polymer, was greater than 160% of TCPS.^[23] This study also demonstrated the biocompatibility of poly(ester urethane urea) degradation products in a cell culture assay.^[23]

Processing and Applications

Physical cross-linking is reversible with heat treatment or solvation, which allows polyurethanes to be processed by such methods as extrusion, injection molding, and solvent casting. Polyurethanes or poly(urethane ureas) consisting of physical cross-links are referred to as linear or thermoplastic elastomers. Examples of some techniques to process polymers into porous scaffolds for cellular ingrowth in tissue engineering applications include freeze drying (lyophilization) and salt leaching. Salt leaching occurs when sieved particles of sodium chloride are added to a polymer solution. The solvent is allowed to evaporate, which leaves dry polymer. The salt present in the polymer can then be leached out with water.^[43] The BDO and BDI-b-BDO-b-BDI-b-BDO-b-BDI chain-extended polyurethanes by Spaans et al. were processed into porous scaffolds by a combination of salt leaching and freeze drying in 1,4-dioxane solvent.^[25] The scaffolds possessed interconnected pores ranging in size from 150–300 μm . This material showed a high compression modulus and was thought suitable for application as a meniscal prosthesis. Scaffolds for tissue engineering of the knee meniscus were created from

polyurethanes based on a 50/50 poly(L-lactide-co- ϵ -caprolactone) prepolymer with BDI.^[37] In a subsequently study, porous scaffolds of the above polymer were constructed by the same method described in the preceding discussion to create two types of scaffolds.^[44] One type of scaffold possessed interconnected micropores of 15–20 μm and a compression modulus of 460 kPa, whereas the other had interconnected micropores of 30 μm and a compression modulus of 74 kPa. Samples were implanted in rats and characterized at 1, 4, 8, and 24 weeks. The scaffold with the smaller pore size displayed incomplete tissue ingrowth as compared to the scaffold with larger pores. However, the scaffold with the larger pore size had a compression modulus that was too low. Both foams were concluded to be unsuitable for meniscal reconstruction.^[44]

Lendlein and Langer designed and characterized shape-memory thermoplastic polyurethanes to be used in minimally invasive surgical procedures.^[40] The elastic polymers were melted and extruded into monofilaments to be used as sutures. The filaments were loosely sutured in dead rats and heated to 41°C. With the increase in temperature came shrinkage of the polymer, which resulted in a tighter wound closure (Fig. 5).

Different compositions of polymers prepared from poly[glycolide-co-(ϵ -caprolactone)-diol and crystalline poly(*R*)-3-hydroxybutyric acid-co-(*R*)-3-hydroxyvaleric acid]-diol with TMDI were melt-extruded into thin hollow tubes for tissue engineering of nerve guidance channels.^[45] Nerve guidance channels are polymer tubes onto which severed nerve endings can be grafted. Borkenhagen et al. studied nerve regeneration in rats at 4, 12, and 24 weeks. Twenty-three out of 26 implants contained regenerated tissue cables composed of myelinated axons and Schwann cells within the lumen.

CONCLUSION

The number of research reports published in the area of biodegradable elastomers is rapidly rising as appreciation for the role of mechanical properties in tissue-biomaterial interactions grows. The negative effects of mechanical property mismatch between an implant, e.g., a stiff vascular graft, and the surrounding tissue have been documented and studied for some time. The importance of mechanical training in the

appropriate development of tissue-engineered constructs is more recently being recognized in several systems.^[46–48] To create new biodegradable elastomeric materials that meet the needs for a variety of clinical applications, researchers are working toward solutions from platforms of naturally occurring materials or elastomers based on synthetic polymers widely utilized for nonmedical applications. The synthetic approach advances as biomaterials engineers work from an increasingly diverse array of biocompatible starting materials and incorporate specific bioactivity. There is a trend towards convergence wherein materials ideally would be created that possess the processing and design flexibility of synthetic approaches and the inherent bioactivity and biocompatibility of natural materials.

ARTICLES OF FURTHER INTEREST

Biocompatibility Testing; Biodegradable Polymers, An Overview; Biofunctional Polymers; Biomimetic Materials; Biorubber/Poly(Glycerol Sebacate); Degradable Polymer Composites; Poly(carbonate)urethanes; Poly(Glycolic Acid); Poly(lactic acid)s; Tissue Engineering of Blood Vessel

REFERENCES

- Cai, Q.; Bei, J.Z.; Wang, S.G. Synthesis and characterization of polycaprolactone (B)-poly(lactide-co-glycolide) (A) ABA block copolymer. *Polym. Adv. Technol.* **2000**, *11* (4), 159–166.
- Kwon, I.K.; Park, K.D.; Choi, S.W.; Lee, S.H.; Lee, E.B.; Na, J.S.; Kim, S.H.; Kim, Y.H. Fibroblast culture on surface-modified poly(glycolide-co-epsilon-caprolactone) scaffold for soft tissue regeneration. *J. Biomater. Sci., Polym. Ed.* **2001**, *12* (10), 1147–1160.
- Cai, Q.; Bei, J.Z.; Wang, S.G. Synthesis and degradation of a tri-component copolymer derived from glycolide, L-lactide, and epsilon-caprolactone. *J. Biomater. Sci., Polym. Ed.* **2000**, *11* (3), 273–288.
- Mohammadi-Rovshandeh, J.; Farnia, S.M.F.; Sarbolouki, M.N. Synthesis and characterization of novel ABA triblock copolymers from L-lactide, glycolide, and PEG. *J. Appl. Polym. Sci.* **1999**, *74* (8), 2004–2009.
- Hiki, S.; Miyamoto, M.; Kimura, Y. Synthesis and characterization of hydroxy-terminated [RS]-poly(3-hydroxybutyrate) and its utilization to block copolymerization with L-lactide to obtain a biodegradable thermoplastic elastomer. *Polymer* **2000**, *41* (20), 7369–7379.
- Wang, Y.D.; Ameer, G.A.; Sheppard, B.J.; Langer, R. A tough biodegradable elastomer. *Nat. Biotechnol.* **2002**, *20* (6), 602–606.
- Storey, R.F.; Herring, K.R.; Hoffman, D.C. Hydroxy-terminated poly(epsilon-caprolactone-co-delta-valerolactone) oligomers: Synthesis, characterization, and polyurethane network formation. *J. Polym. Sci., Polym. Chem. Edn.* **1991**, *29*, 1759–1777.
- Storey, R.F.; Hickey, T.P. Degradable polyurethane networks based on D, L-lactide, glycolide, epsilon-caprolactone, and trimethylene carbonate homopolyester and copolyester triols. *Polymer* **1994**, *35* (4), 830–838.
- Albertsson, A.C.; Varma, I.K. Aliphatic polyesters: Synthesis, properties and applications. *Adv. Polym. Sci.* **2002**, *157*, 1–40.
- Sudesh, K.; Doi, Y. Molecular design and biosynthesis of biodegradable polyesters. *Polym. Adv. Technol.* **2000**, *11*, 865–872.
- Abe, H.; Matsubara, I.; Doi, Y.; Hori, Y.; Yamaguchi, A. Physical-properties and enzymatic degradability of poly(3-hydroxybutyrate) stereoisomers with different stereoregularities. *Macromolecules* **1994**, *27* (21), 6018–6025.
- Dobrzynski, P. Synthesis of biodegradable copolymers with low-toxicity zirconium compounds. II. Copolymerization of glycolide with epsilon-caprolactone initiated by zirconium(IV) acetylacetonate and zirconium(IV) chloride. *J. Polym. Sci., Polym. Chem.* **2002**, *40* (10), 1379–1394.
- Hiljanen-Vainio, M.P.; Orava, P.A.; Seppälä, J.V. Properties of epsilon-caprolactone/DL-lactide (epsilon-CL/DL/LA) copolymers with a minor epsilon-CL content. *J. Biomed. Mater. Res.* **1997**, *34*, 39–46.
- Lofgren, A.; Albertsson, A.C.; Dubois, P.; Jerome, R.; Teyssie, P. Synthesis and characterization of biodegradable homopolymers and blockpolymers and block-copolymers and block-copolymers based on 1,5-Dioxepan-2-one. *Macromolecules* **1994**, *27* (20), 5556–5562.
- Stridsberg, K.; Albertsson, A.C. Controlled ring-opening polymerization of L-lactide and 1,5-dioxepan-2-one forming a triblock copolymer. *J. Polym. Sci., Polym. Chem.* **2000**, *38* (10), 1774–1784.
- Pego, A.P.; Poot, A.A.; Grijpma, D.W.; Feijen, J. Copolymers of trimethylene carbonate and epsilon-caprolactone for porous nerve guides: Synthesis and properties. *J. Biomater. Sci., Polym. Ed.* **2001**, *12* (1), 35–53.
- Kim, J.H.; Lee, J.H. Preparation and chain-extension of P(LLA-b-TMC-b-LLA) triblock copolymers and their elastomeric properties. *Macromol. Res.* **2002**, *10* (2), 54–59.
- Lin, S. Lactide/Caprolactone Polymer, Method of Making the Same, Composites Thereof, and Prostheses Produced Therefrom. US Patent 4,643,734, February 17, 1987.
- Lofgren, A.; Renstad, R.; Albertsson, A.C. Synthesis and characterization of a new degradable thermoplastic elastomer based on 1,5-dioxepan-2-one and epsilon-caprolactone. *J. Appl. Polym. Sci.* **1995**, *55* (11), 1589–1600.
- Bruin, P.; Veenstra, G.J.; Nijenhuis, A.J.; Pennings, A.J. Design and synthesis of biodegradable poly(ester-urethane) elastomer networks composed of non-toxic building blocks. *Makromol. Chem., Rapid Commun.* **1988**, *9*, 589–594.
- Bruin, P.; Smedinga, J.; Pennings, A.J. Biodegradable lysine diisocyanate-based poly(glycolide-co-epsilon-caprolactone)-urethane network in artificial skin. *Biomaterials* **1990**, *11*, 291–295.

22. Storey, R.F.; Wiggins, J.S.; Puckett, A.D. Hydrolyzable poly(ester-urethane) networks from L-lysine diisocyanate and D,L-lactide/ ϵ -caprolactone homo- and copolyester triols. *J. Polym. Sci., Polym. Chem.* **1994**, *32*, 2345–2363.
23. Guan, J.; Sacks, M.S.; Beckman, E.J.; Wagner, W.R. Synthesis, characterization, and cytocompatibility of elastomeric, biodegradable poly(ester-urethane)ureas based on poly(caprolactone) and putrescine. *J. Biomed. Mater. Res.* **2002**, *61*, 493–503.
24. Spaans, C.J.; de Groot, J.H.; Belgraver, V.W.; Pennings, A.J. A new biomedical polyurethane with a high modulus based on 1,4-butanediisocyanate and ϵ -caprolactone. *J. Mater. Sci., Mater. Med.* **1998**, *9*, 675–678.
25. Spaans, C.J.; de Groot, J.H.; Dekens, F.G.; Pennings, A.J. High molecular weight polyurethanes and a polyurethane urea based on 1,4-butanediisocyanate. *Polym. Bull.* **1998**, *41*, 131–138.
26. Kylmä, J.; Seppälä, J.V. Synthesis and characterization of a biodegradable thermoplastic poly(ester-urethane) elastomer. *Macromolecules* **1997**, *30*, 2876–2882.
27. Lendlein, A.; Neuenchwander, P.; Suter, U.W. Tissue-compatible multiblock copolymers for medical applications, controllable in degradation rate and mechanical properties. *Macromol. Chem. Phys.* **1998**, *199*, 2785–2796.
28. Saad, G.R.; Lee, Y.J.; Seliger, H. Synthesis and characterization of biodegradable poly(ester-urethanes) based on bacterial poly(R-3-hydroxybutyrate). *J. Appl. Polym. Sci.* **2002**, *83*, 703–718.
29. Gorna, K.; Polowinski, S.; Gogolewski, S. Synthesis and characterization of biodegradable poly(ϵ -caprolactone urethane)s. I. Effect of the polyol molecular weight, catalyst, and chain extender on the molecular and physical characteristics. *J. Polym. Sci., Part A, Polym. Chem.* **2002**, *40*, 156–170.
30. Gorna, K.; Gogolewski, S. Biodegradable polyurethanes for implants. II. In vitro degradation and calcification of materials from poly(ϵ -caprolactone)-poly(ethylene oxide) diols and various chain extenders. *J. Biomed. Mater. Res.* **2002**, *60*, 592–606.
31. de Groot, J.H.; de Vrijer, R.; Wildeboer, B.S.; Spaans, C.S.; Pennings, A.J. New biomedical polyurethane ureas with high tear strengths. *Polym. Bull.* **1997**, *38*, 211–218.
32. Lamba, N.M.K.; Woodhouse, K.A.; Cooper, S.L. *Polyurethanes in Biomedical Applications*; CRC Press: New York, 1998.
33. Schoental, R. Carcinogenic and chronic effects of 4,4-diamino-diphenyl methane, an epoxy resin hardener. *Nature* **1968**, *219*, 1968.
34. *National Toxicology Program Technical Report on the Carcinogenesis Bioassay of 4,4'-Methylenedianiline Dihydrochloride, NTP-81-143*; NIH Publication No. 82-2504, National Institutes of Health: Bethesda, MD, 1982.
35. Cohn, D.; Stern, T.; González, M.F.; Epstein, J. Biodegradable poly(ethylene oxide)/poly(ϵ -caprolactone) multiblock copolymers. *J. Biomed. Mater. Res.* **2002**, *59*, 273–281.
36. de Groot, J.H.; Spaans, C.J.; Dekens, F.G.; Pennings, A.J. On the role of aminolysis and transesterification in the synthesis of ϵ -caprolactone and L-lactide based polyurethanes. *Polym. Bull.* **1998**, *41*, 299–306.
37. Skarja, G.A.; Woodhouse, K.A. In vitro degradation and erosion of degradable, segmented polyurethanes containing an amino acid-based chain extender. *J. Biomater. Sci., Polym. Ed.* **2001**, *12* (8), 851–873.
38. Saad, B.; Hirt, T.D.; Welti, M.; Uhlschmid, G.K.; Neuenchwander, P.; Suter, U.W. Development of degradable polyesterurethanes for medical applications: In vitro and in vivo evaluations. *J. Biomed. Mater. Res.* **1997**, *36*, 65–74.
39. Lendlein, A.; Colussi, M.; Neuenchwander, P.; Suter, U.W. Hydrolytic degradation of phase-segregated multiblock copoly(ester urethane)s containing weak links. *Macromol. Chem. Phys.* **2001**, *202*, 2702–2711.
40. Lendlein, A.; Langer, R. Biodegradable, elastic shape-memory polymers for potential biomedical applications. *Science* **2002**, *296*, 1673–1676.
41. Skarja, G.A.; Woodhouse, K.A. Synthesis and characterization of degradable polyurethane elastomers containing an amino acid-based chain extender. *J. Biomater. Sci., Polym. Ed.* **1998**, *9* (3), 271–295.
42. Skarja, G.A.; Woodhouse, K.A. Structure-property relationships of degradable polyurethane elastomers containing an amino acid-based chain extender. *J. Appl. Polym. Sci.* **2000**, *75* (12), 1522–1534.
43. Agrawal, C.M.; Ray, R.B. Biodegradable polymeric scaffolds for musculoskeletal tissue engineering. *J. Biomed. Mater. Res.* **2001**, *55*, 141–150.
44. van Tienen, T.G.; Heijkants, R.G.J.C.; Buma, P.; de Groot, J.H.; Pennings, A.J.; Veth, R.P.H. Tissue ingrowth and degradation of two biodegradable porous polymers with different porosities and pore sizes. *Biomaterials* **2002**, *23* (8), 1731–1738.
45. Borkenhagen, M.; Stoll, R.C.; Neuenchwander, P.; Suter, U.W.; Aebischer, P. In vivo performance of a new biodegradable poly(ester urethane) system used as a nerve guidance channel. *Biomaterials* **1998**, *19*, 2155–2165.
46. Butler, D.L.; Goldstein, S.A.; Guilak, F. Functional tissue engineering: The role of biomechanics. *J. Biomech. Eng.* **2000**, *122*, 570–575.
47. Hoerstrup, S.P.; Zund, G.; Sodian, R.; Schnell, A.M.; Grunfelder, J.; Turina, M.I. Tissue engineering of small caliber vascular grafts. *Eur. J. Cardio-thorac. Surg.* **2001**, *20*, 164–169.
48. Davisson, T.; Kunig, S.; Chen, A.; Sah, R.; Ratcliffe, A. Static and dynamic compression modulate matrix metabolism in tissue engineered cartilage. *J. Orthop. Res.* **2002**, *20*, 842–848.

Electric Cell-Substrate Impedance Sensing

Philippe Lam

Genentech, Inc., South San Francisco, California, U.S.A.

INTRODUCTION

Invented by I. Giaever and C. Keese in the early 1980s, electrical cell impedance sensing (ECIS) is a non-invasive technique that permits monitoring live cells in situ and in real time. A small amplitude AC signal is imposed across a pair of electrodes onto which cells are deposited and the impedance of the system measured by a dedicated instrument. ECIS is capable of detecting the behavior of cells at each phase of culture—from attachment, spreading, and growth to mitosis, and death. This versatile technique has been successfully used to study adhesion, cell locomotion, tight junction barrier functions, response to toxicants, membrane electroporation, and various other applications. Ease of implementation and broad applicability make ECIS an ideal tool for any type of adherent cell-based assays.

The following provides an overview of the ECIS method by reviewing some fundamental concepts, discussing the instrumentation used, and briefly presenting some applications.

IMPEDANCE CONCEPT

For electrical circuits involving direct currents (DC), the voltage (E) and current (I) are related by a simple proportionality constant termed the resistance (R), as described by the well-known equation:

$$E = R \times I$$

The voltage, current, and resistance have units of volts (V), amperes (A), and ohms (Ω), respectively.

In systems involving alternating currents (AC), the time-dependent nature of the signal introduces an additional dimension. In this instance, the proportionality factor must describe both magnitude and phase relationships between the voltage and the current as illustrated in Fig. 1. This factor, the impedance (Z), is typically represented in complex notation and can be viewed as the AC analogue of the resistance and has the same unit. The equation above can be rewritten more generally as:

$$E = Z \times I$$

Whereas there are many accepted notations for the impedance, the following are more commonly found in the ECIS literature:

$$Z = a + ib$$

in which i is the imaginary number and a and b are the real and imaginary parts representing pure resistive and pure reactive (capacitive and inductive) contributions, respectively.

With this notation, the total impedance is:

$$|Z| = (a^2 + b^2)^{0.5}$$

If we assume a simple circuit representation for the system comprising an ideal resistor and capacitor in series, the impedance can be written as:

$$Z = R + 1/(2i\pi fC)$$

in which f is the frequency of the excitation signal in hertz (Hz), and C is the capacitance in Farads (F). Representing the impedance in such manner gives some physical meaning to the measured quantities as it is a simple model for the electrode-electrolyte interface in which C represents the double-layer capacitance, and R accounts for the resistance to interfacial charge transfer and double-layer charging current.^[1] In general, the quantities (a , b , R , C) in the above expressions depend on the system under study as well as frequency. It should also be noted that all of the representations of the impedance are equivalent and can be converted to any other notation to facilitate interpretation of the data.^[2]

PRINCIPLE OF ECIS

For conducting electrolytes, the potential measured across an electrochemical cell is typically dominated by the potential drop at the electrode-solution interface. Using AC provides greater sensitivity and versatility and alleviates electrode polarization problems associated with DC methods. The exquisite sensitivity of the ECIS technique hinges upon the dimensions of the sensing and counter electrodes used. By making the area of the sensing electrode small, on the order of 10^{-3} to 10^{-4} cm² (10^5 to 10^4 μ m²), and the counter

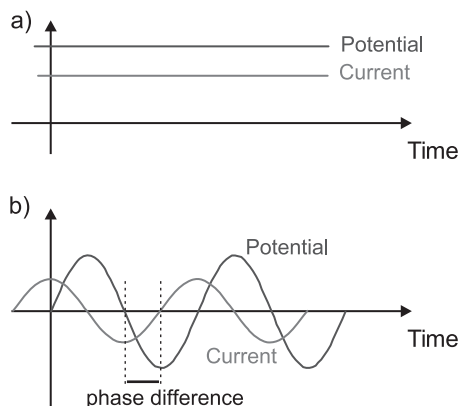


Fig. 1 Current-potential relationship. a) For DC, the potential and current differ only in magnitude: b) for AC, the signals can be different in both magnitude and the relative position of the peaks and valleys (phase difference). As shown, the current is said to be leading the potential by 90° . Because the signal is periodic, we can also say that the current is lagging the potential by 270° .

electrode much larger, approximately 1 cm^2 , the impedance of the system comprising the two electrodes and typical cell culture media will be dominated by the impedance at the sensing electrode-solution interface.^[1,3] This arrangement creates an electrical bottleneck about the sensing electrode. The contributions of the counter electrode and the culture medium represent a very small percentage (typically $<1\%$) of the total system impedance. Because cell membrane is electrically insulating, the presence and activity of cells on the surface of the sensing electrode can easily be detected. Figure 2 illustrates this concept. Only adherent cells

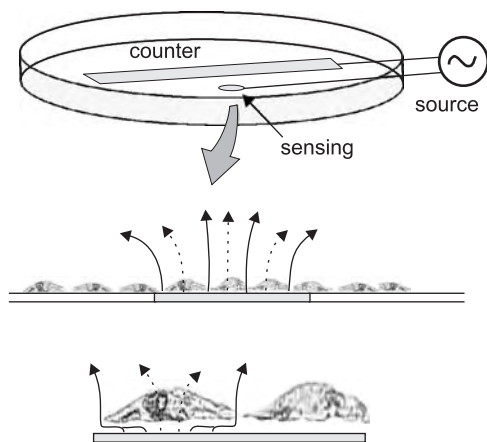


Fig. 2 ECIS electrodes. A schematic of typical electrodes, shown not-to-scale. Cells cover the surface of the whole dish; however, the only bottleneck is at the sensing electrode (shown by the current path in the magnified views). AC current can flow out from the sensing electrode surface from around (solid lines) or through the cells because cell membrane is capacitive (dashed lines).

cause an appreciable increase in impedance. Dead cells or rounded cells that have not attached do not result in significant blockage of the current path and thus are not detected.

INSTRUMENTATION

In principle, common impedance measurement methods such as bridges, LRC (inductance-resistance-capacitance) meters, frequency response analyzers, etc. may be used to collect data from the cell-covered electrodes. However, due to the delicate nature of the test specimen, one must ensure that the instrument is capable of obtaining reliable readings using low excitation potentials and currents so as not to be disruptive to the culture. A suitable instrument must also be able to conduct measurements at various frequencies in order to provide maximum sensitivity depending on the type of cells under study.^[1] Generally, the current across the cell layer should be on the order of a microampere and the potential should be limited to a few millivolts.^[3] Giaever and Keese have shown that, under these conditions, the cells are not affected by the electrical probing and behave identically to cells cultivated in regular culture dishes. By continuously monitoring the impedance response of the culture as a function of excitation frequency over time, the optimum frequency can be selected based upon the time change of the observed impedance signal. Once the desired frequency has been obtained, a typical ECIS experiment would involve probing the culture over time at that particular frequency.

Frequency response analyzers (FRA) such as the Solartron 1260 (Farnborough, United Kingdom) are well suited for ECIS measurements. However, due to high costs, FRAs have not been adopted by the ECIS community. Most of the ECIS work reported in the literature has been conducted on a phase-sensitive lock-in amplifier as described by Giaever and Keese.^[1,3] In this setup, schematized in Fig. 3, the sensing and counter electrodes are connected to an AC signal generator through a current limiting $1\text{ M}\Omega$ resistor in series. The typical impedance of a populated ECIS electrode (sensing, counter) and culture medium is several orders of magnitude less than $1\text{ M}\Omega$ at the frequencies used in the experiments. Therefore, the current through the system (sensing electrode, culture media, counter electrode, and resistor) may be considered to be nearly constant. The impedance of the cell layer can then be computed from the in-phase and out-of-phase voltage readings from the lock-in amplifier.

Using a single instrument, only one electrode pair (sensing, counter) can be measured at a time. However, each measurement can typically be completed in well

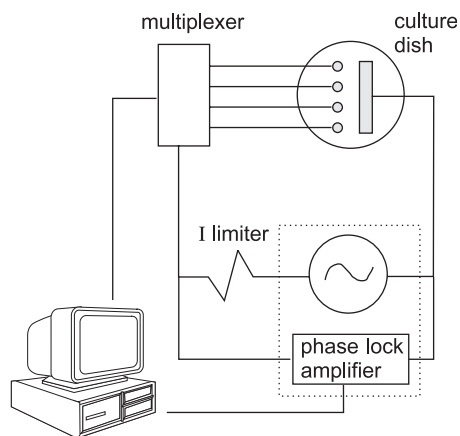


Fig. 3 ECIS system. This depicts a complete system, including multiplexing capability. The culture dish is located inside an incubator. The current-limiting resistor may be of different denomination depending on the application and instrument used. The AC signal source and phase-lock amplifier can be substituted by other impedance-measuring instruments such as a frequency response analyzer.

under one second. Because of the relatively slow nature of biological systems, measurement rates on the order of one per minute are sufficient to capture most of the dynamic behavior of a culture. For faster transients, such as monitoring membrane permeation, higher measurement rates can be used. An additional advantage of this technique is that the probing potential need only be present during the actual measurement. Therefore, cells are subjected to the externally applied potential only for a very brief instant, minimizing the risks of culture disruption.

Because the measurements involve low amplitude signals, care should be taken to reduce the influences of electrical noise. This means using high-quality, short cable runs, proper electrical connectors, and appropriate shielding. It is especially important to eliminate ground loops as well as any DC leakages, as these may cause artifacts in the data and have adverse effects on the culture. Furthermore, it is recommended to use a separate circuit for the power feed to the instrument from that of high-load devices such as water baths and laminar flow hoods.

Construction of the electrode assembly presents some challenges due to the diminutive size of the sensing electrode. Also, the electrode material must be nontoxic, conducive to cellular attachment, and present negligible electrical resistance. According to the guidelines recommended by the inventors, for a sensing electrode area of approximately 10^{-4} cm^2 , the physical dimensions would need to be approximately $120 \mu\text{m}$ in diameter if a disk or approximately $100 \mu\text{m}$ per side if a square. Larger electrodes that are several times the ideal area may be used successfully, but, if made too

large, sensitivity will be lost. These dimensions are easily obtained by common photolithography techniques. In their initial work, Giaever and Keese deposited a thin layer of gold by evaporation through a mask onto a polystyrene substrate. Red wax, which was found to be nontoxic was used to seal electrical connection and limit electrode areas.^[1,3] Gold is highly conductive, provides a good biocompatible surface, and, in a thin layer, is sufficiently transparent to allow observation of the cultured cells by optical microscopy.

A complete ECIS apparatus can now be purchased from Applied Biophysics (Troy, New York). The system comprises the measuring instrument, a multiplexer to allow quasisimultaneous monitoring of up to 16 different electrodes, and a computer with software for control and data acquisition/analysis. Also available are specially designed multiwell gold, disk-shaped microelectrode culture dishes. Photoresist material, in lieu of red wax, is used to limit electrode size to approximately $250 \mu\text{m}$ in diameter.

ECIS APPLICATIONS

The ECIS method was originally conceived to permit in situ and real-time monitoring of adherent cell cultures. The following gives a brief overview of some of the many applications that have been reported in the literature.

In the original account detailing an ECIS experiment, Giaever and Keese demonstrated that the behavior of fibroblast cells undergoing various stages could easily be monitored as a function of time.^[3] They also confirmed the noninvasive nature of the measurements by observing cell morphology and division time. In their experiments, the cells were seeded from a suspension onto a gold-sensing electrode. The attachment, spreading of the cells, growth, replication, and confluence stages of the culture were clearly reflected in the impedance data. Figure 4 depicts typical ECIS data showing the impedance response of a culture at various stages. As cells grow to confluence on the electrode surface, the impedance data typically exhibits a maximum. Thereafter, rapid noiselike fluctuations in the signal can be seen. These fluctuations, which are found to be fractal in nature, can be attributed to the micromotion of live cells.^[4] Such features are absent from either unpopulated electrodes or these occupied by dead cells.

Because of the large impedance change associated with the presence of live cells on the electrode, ECIS provides an excellent method for monitoring cell adhesion. For instance, it has been shown that cells were attaching and spreading quite differently depending on the specific proteins that were precoated onto the culture surface.^[1,5] Fibronectin, laminin,

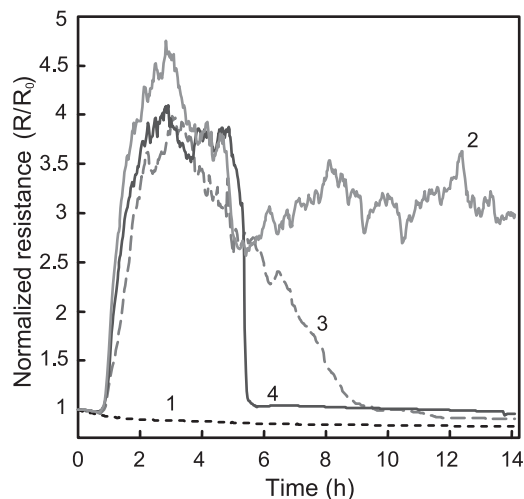


Fig. 4 Typical ECIS data. Normalized resistance data generated from a culture of endothelial cells exposed to a toxicant (0.025% solution of diisobutylphenoxyethoxyethyl dimethyl benzyl ammonium chloride) at various levels. Trace (1) plain media, no cells; (2) control, no toxicant; (3) 1 μ l toxicant; and (4) 5 μ l toxicant added. Media volume is approximately 0.5 ml. A large increase in resistance, starting at about 1 hour, is a result of cell attachment onto the electrodes. The typical ECIS noiselike signal due to cellular micromotion can be seen in healthy cells (Trace 2). The signal from cells exposed to toxicant (Traces 3 and 4) eventually fell to levels of no-cell control, which indicated seriously impaired or dead cells. (Data courtesy of Prof. Gary Bowlin, Dept. of Biomedical Engineering, Virginia Commonwealth University, Richmond, Virginia, U.S.A.)

vitronectin, and bovine serum albumin were coated onto separate ECIS electrodes from 100 μ g/ml protein solutions in saline. A suspension of canine kidney-derived MDCK epithelial cells was then inoculated onto the precoated electrodes, and the impedance of each culture followed over time. A preliminary experiment determined 40 kHz to be the optimum frequency for monitoring attachment and spreading of these cells. For each case, $t_{1/2}$, the time required for the cells to achieve half-coverage of the electrode, was extracted from a plot of the series capacitance versus time. The data revealed that MDCK cells attached and spread most easily onto fibronectin-coated electrodes with $t_{1/2}$ of approximately 1 hour. Laminin and vitronectin coatings yielded $t_{1/2}$ of 3.8 and 4.6 hours, respectively, whereas bovine serum albumin surfaces were the least amenable to cell spreading ($t_{1/2}$ = 10 hours). Traditional cell-substrate interaction assays are more prone to experimental variation as they typically require seeding, attachment, rinsing off weakly or nonattached cells, and then quantification of the surface cells. Using ECIS, the researchers were able to obtain data without disrupting the culture.

The effects of various factors, such as genes, growth factors, signaling molecules, etc., that influence cell attachment/motion can also be observed by ECIS. Charboneau et al. reported that MCF-7 breast cancer cell transfected to express adenocarcinoma of the lung-1 (DAL-1) protein resulted in enhanced cell attachment.^[6] These researchers used ECIS to compare transfected cells to a control by measuring the rate of increase in resistance during attachment as well as total increase in resistance over longer times. Similarly, Schmidt et al. used an ECIS-based assay as evidence that the adaptor protein SETA/CIN85/Ruk and binding partner AIP1 were involved in regulating astrocyte adhesion.^[7] More recently, Sawhney et al. have shown that epidermal growth factor (EGF) receptor mediated cell adhesion and motion.^[8] In these experiments, ECIS was used to detect the micromotion of human colon carcinoma HCT116 cells in response to exogenous EGF. This was accomplished by analyzing the fluctuations in the impedance signal from established cultures over time. The amount of signal fluctuation, observed after the large initial impedance change due to attachment (e.g., Fig. 4, trace #2) is a quantitative measure of cellular movement. As each cell changed position about the surface, the inter-cellular distance as well as the gap between the cell membrane and the electrode also changed. This, in turn, resulted in changing current paths, which leads to measured impedance fluctuations.

Furthermore, it has been shown that alteration of environmental factors such as temperature, pH, and media composition can also be detected from the fluctuation in the signal caused by the micromotion of the cells.^[9] For instance, small pH fluctuations of the culture media can be observed in some on-off type CO₂ controlled incubators. Although these subtle changes are not easily detectable using conventional pH probes, they can affect cellular morphology as seen in ECIS data.^[10] Care must be taken when designing and conducting experiments to minimize effects from factors not under study.

Certain cultures, such as endothelial or epithelial cells, form tight junctions at confluence. Traditionally, the study of tight junctions have been conducted via electrical resistance measurements from cells grown on permeable filter substrates. These populated filters are placed between two media-filled chambers and an electrical signal imposed across the cell layers. An increase in the permeability of the cell layer results in a lower measured resistance. However, the type of filters used (material, pore size) may affect the morphology of the cultured layer; hence, the measurements can be influenced by factors other than tight junctions alone. ECIS provides a more convenient method for such investigations as the culture setup involves only one chamber and eliminates the risk of cells penetrating

into the permeable filter pores that can cause erroneously high tight junctions resistance values. Tiruppathi et al. first reported that endothelial barrier function could be followed in real time by ECIS.^[11] In these experiments, bovine pulmonary microvessel endothelial and bovine artery pulmonary endothelial cells were first grown to confluence on ECIS electrodes. The cells were then challenged with addition of α -thrombin, and the electrical response was measured. Molecules such as α -thrombin are implicated in inflammation and are known to increase the permeability of endothelial cell layers. The results, which were consistent with literature accounts, indicated that α -thrombin caused a reversible decrease in endothelial barrier resistance. These changes were due to cell shape alteration mediated by α -thrombin. A later communication by Lo et al. gave some insights on other factors that can affect measured transepithelial resistance and compared results from cell layers grown on filters and by ECIS.^[12] The researchers reported that certain types of cells, when grown on polycarbonate filters, yielded unusually high measured resistance values. They attributed this to an additional resistance from the restricted current path between the base of the cells and the substrate (see Fig. 2). The smaller the cell-substrate distance, the higher the measured resistance. In the analysis of ECIS data, Lo et al. utilized a mathematical model proposed by Giaever and Keese^[13–15] and obtained parameters such as average cell membrane capacitance, average cell-substrate distance, and inter-cellular junction resistance. The model abstracts cells as being disks of radius r_c located at a distance h from the electrode surface. The current originates from the sensing electrode surface uniformly and follow various paths to the counter electrode (see Fig. 2). Since the excitation is AC in nature, a fraction of the current goes through the cell as capacitive current. The remainder travels horizontally through the small gap between the cell bottom and the electrode surface, vertically through the space between neighboring cells, through the tight junction (if present), and out to the counter electrode via the cell culture medium. By conducting measurements at various excitation frequencies, adjustable model parameters can be fitted to the data. Using this approach, the average cell-substrate distances for MDCK epithelial cells, WI38-VA13 and WI38 fibroblasts cells were estimated to be 1, 13, and 113 nm, respectively. Such parameters would be difficult to obtain via traditional microscopy.

Neutrophil migration through endothelial cell layers occurs at sites of tissue infection or injury. This process is normally studied by microscopy (electron microscopy, immunofluorescence, etc.). However, these conventional techniques are complex, require a great deal of sample preparation, and most do not give real-time data. Burns et al. used ECIS to study human

peripheral blood neutrophil (PMN) migration through human umbilical vein endothelial cell (HUVEC) layers and found that the endothelial tight junctions were not disrupted.^[16] The HUVEC cells were cultured to confluence on ECIS electrodes. IL-8-activated PMNs were then added in a 20 to 1 ratio over the HUVEC layer. In separate electrode wells, addition of unactivated neutrophils was used as a negative control, and histamine addition, which disrupts tight junctions, was the positive control. No noticeable decrease in trans-endothelial electrical resistance was recorded for either of the neutrophil addition cases, whereas histamine caused a significant drop. The ECIS results were in agreement with electron microscopy data. Migration of the neutrophils at the discontinuities present at the junction of three neighboring HUVEC cells may explain why no significant disruption to the tight junctions was detected.

In contrast, Keese et al. showed that deposited metastatic cells were severely disruptive as they invaded an endothelial cell culture.^[17] The process by which cancerous cells migrate from one tumor site to another is important in understanding disease progression. Furthermore, quantifying the metastatic potential of each type of cancer cell can provide valuable information. Multiple lines of cancer cells with varying degrees of metastatic potential (non-metastatic to highly metastatic) were used to challenge HUVEC cell layers grown on ECIS electrodes. Several sublines of Dunning murine prostatic adenocarcinoma cells as well as several human prostatic cell lines were added to separate confluent layers of HUVEC in a 1 to 1 ratio. The impedance at 4 kHz was followed over time. Results indicated that the change in measured resistance was proportional to the degree of metastasis of the invading cells. Highly metastatic cells caused a large decrease in resistance, whereas weakly or non-metastatic cells caused a small or no decrease in resistance. This work shows that ECIS can be a valuable tool for oncology research.

Another ideally suited application for ECIS is that of cell-based, in-vitro toxicology. Giaever and Keese suggested that the impedance response of cultured cells could be used as an indicator of cell health.^[18] Presumably, the behavior of cells at various stages of culture (attachment, spreading, motility, replication, death) would be disrupted by an introduced toxicant, and the change in impedance compared to a control be used to gauge the level of toxicity. In this case, it would be advantageous to determine which culture phase should be monitored in order to obtain the most useful information. A carefully planned experiment may be able to tease out information on the mechanism of action as well as obtain level-of-toxicity data. For instance, one possible scenario would be to mix the toxicant with the cell suspension and monitor the

effects during attachment. Another possibility would be to add the toxicant after establishment of the culture. Keese et al. chose to expose confluent MDCK epithelial and WI-38/VA13 fibroblasts cell layers to various levels of several common detergents such as Tween 20 and sodium lauryl sulfate and measure the resulting cellular micromotion.^[19] The data were analyzed in terms of disruption to the micromotion signal as compared to a control. Using this approach, they were able to rank the detergents in order of effects. Their findings are in agreement with other published reports obtained using conventional animal testing or skin testing. Another study by Xiao et al. employed a different procedure, adding toxicants—either arsenate, cadmium or benzalkonium salts at various concentrations—to a suspension of V79 fibroblast and monitored cellular attachment and growth by ECIS.^[20] They analyzed the response of the challenged cells at 16 hours postinoculation as compared to a control. They showed comparability of ECIS to the traditional neutral red uptake assay. Although test conditions and analysis procedures must be optimized for each case, it is evident that ECIS can be applied to cell-based toxicity assays.

An elegant use of ECIS for monitoring chemotaxis of eukaryotic *Dictyostelium discoideum* cells exposed to a folic acid gradient was reported by Hadjout et al.^[21] Quantitative measurements of cell chemotaxis can be difficult to obtain as there are several issues involved. First, a stable gradient must be established. This can be accomplished by using a gel matrix or a porous filter to eliminate convective and reduce diffusive mixing of the agent under test. Second, one must be able to detect cellular motion. Most conventional methods are based on microscopy, and involve cell staining or tracking with elaborate video systems. Hadjout et al. proposed a novel method whereby they have adapted the commercial ECIS multiwell gold microelectrode dish to permit real-time monitoring of cell migration. In their setup, each well of the electrode array was partially filled with agarose gel, and small plugs—one on each side of the sensing electrode—were removed from the gel. Cells were deposited in one of the newly created openings and chemoattractant was deposited in the other. Impedance was monitored as the cells migrated across the sensing electrode and under the agarose gel. Initially, the electrode was uncovered by cells. The agarose did not contribute significantly to the measured impedance as the gel was ionically permeable. As the cells migrated onto the sensing electrode, an increase in impedance, which was proportional to the number of cells present, was detected. When the cells moved off the electrode, the impedance decreased. The time of arrival of the cells to the electrode, i.e., a measure of cell-moving speed, can be easily obtained as well. In this example, ECIS

was used as an on-off type sensor in a setup that was simple, sensitive, multiplexable, and capable of providing data in real time.

With minor modifications to the instrumentation, in situ electroporation of the cells growing on the sensing electrode can be performed. Furthermore, with computer-controlled setups, the instrument can be reverted back to the regular mode immediately after the electroporation step and begin monitoring cell behavior. In contrast to traditional electroporation chambers in which high voltages (hundreds to thousands of volts) are necessary for efficient permeation, the ECIS asymmetrical electrode configuration permits high current densities and large potential drops at the sensing electrode from low applied potentials. Furthermore, the process can be conducted in regular culture media without risks of heating. Ghosh et al. demonstrated feasibility by using AC pulses to introduce horseradish peroxidase into fibroblast cells.^[22] WI-38/VA-13 human lung fibroblast cells were first cultured to confluence onto ECIS electrodes. After addition of horseradish peroxidase to the culture media, 1–5 V pulses at the same frequency as the measurement signal and 200 ms in duration, were applied to each electrode. The data showed that there was a significant drop in impedance as a result of the high voltage exposure, which was an indication that cell membranes had been temporarily permeabilized. The impedance returned to prepulse value within several seconds. Uptake of the horseradish peroxidase was detected colorimetrically after fixing the cells. Later, Wegener et al. conducted extensive investigations aimed at optimizing electroporation conditions on several types of cells.^[23] They studied the effects of pulse amplitude, pulse duration, and pulse frequency and found optimum conditions to be cell-type dependent. Because only the cells present on the sensing electrode are affected, it is also possible to conduct wound healing experiments by applying a lethal voltage to the targeted cells and observing the surrounding cells repopulate the electrode area.

CONCLUSION

As evidenced by the multitude of applications that have been reported, ECIS is an elegant, versatile, and easily implementable technique that is well suited for a variety of adherent cell-based assays. It offers many advantages over traditional optical microscopy, such as low cost, high sensitivity, reproducibility, noninvasiveness, real-time, in situ monitoring, and multiplexing capabilities. Furthermore, the technique is based on simple, fundamental principles and can be adapted to new applications. For instance, it may be possible to use nonplanar sensing electrodes to probe whether cells

can recognize three-dimensional features or even use ECIS for noncell-based applications to detect molecular interaction between solution species and an immobilized ligand.

Already well-established as a powerful research tool with computer-controlled operation and scalability, this technique can potentially be an attractive solution for automated high-throughput, cell-based assay screening systems.

ACKNOWLEDGMENTS

The author would like to thank Ms. K. Bhawe and Drs. S. Sane, J. Baier-Leach, and G. Wnek for helpful comments and Dr. G. Bowlin for ECIS data.

ARTICLES OF FURTHER INTEREST

Biocompatibility Testing; Biopotential Amplifiers; Biosensors; Cell Culture Assays; Electrical Impedance Imaging

REFERENCES

- Giaever, I.; Keese, C.R. Use of electric fields to monitor the dynamical aspect of cell behavior in tissue culture. *IEEE Trans. Biomed. Eng.* **1986**, *33*, 242–247.
- Macdonald, J.R. Fundamentals of Impedance Spectroscopy. In *Impedance Spectroscopy*; John Wiley & Sons: New York, 1987; 1–26.
- Giaever, I.; Keese, C.R. Monitoring fibroblast behavior with an applied electric field. *Proc. Natl. Acad. Sci. U. S. A.* **1984**, *81*, 3761–3764.
- Giaever, I.; Keese, C.R. Fractal motion of mammalian cells. *Physica, D* **1989**, *39*, 128–133.
- Wegener, J.; Keese, C.R.; Giaever, I. Electric cell-substrate impedance sensing (ECIS) as a noninvasive means to monitor the kinetics of cell spreading to artificial surfaces. *Exp. Cell Res.* **2000**, *259*, 158–166.
- Charboneau, A.L.; Singh, V.; Yu, T.X.; Newsham, I.F. Suppression of growth and increased cellular attachment after expression of DAL-1 in MCF-7 breast cancer cells. *Int. J. Cancer* **2002**, *100* (2), 181–188.
- Schmidt, M.H.H.; Chen, B.; Randazzo, L.M.; Bogler, O. SETA/CIN85/Ruk and its binding partner ALP1 associated with diverse cytoskeletal elements, including FAKs, and modulated cell adhesion. *J. Cell. Sci.* **2003**, *166*, 2845–2855.
- Sawhney, R.S.; Zhou, G.K.; Humphrey, L.E.; Ghosh, P.; Kreisberg, J.I.; Brattain, M.G. Differences in sensitivity of biological functions mediated by epidermal growth factor receptor activation with respect to endogenous and exogenous ligands. *J. Biol. Chem.* **2002**, *277*, 75–85.
- Lo, C.M.; Keese, C.R.; Giaever, I. Monitoring motion of confluent cells in tissue culture. *Exp. Cell Res.* **1993**, *204*, 102–109.
- Lo, C.M.; Keese, C.R.; Giaever, I. pH Changes in pulsed CO₂ incubators cause periodic change in cell morphology. *Exp. Cell Res.* **1994**, *213*, 391–397.
- Tiruppathi, C.; Malik, A.B.; Del Vecchio, P.J.; Keese, C.R.; Giaever, I. Electrical method for detection of endothelial cell shape change in real time. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 7919–7923.
- Lo, C.M.; Keese, C.R.; Giaever, I. Cell-substrate contact: Another factor may influence transepithelial electrical resistance of cell layers cultured on permeable filters. *Exp. Cell Res.* **1999**, *250*, 576–580.
- Giaever, I.; Keese, C.R. Micromotion of mammalian cells measured electrically. *Proc. Natl. Acad. Sci. U. S. A.* **1991**, *88*, 7896–7900.
- Giaever, I.; Keese, C.R. Correction. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 1634.
- Lo, C.M.; Keese, C.R.; Giaever, I. Impedance analysis of MDCK cells measured by electric cell-substrate impedance sensing. *Biophys. J.* **1995**, *69*, 2800–2807.
- Burns, A.R.; Bowden, R.A.; MacDonell, S.D.; Walker, D.C.; Odebunmi, T.O.; Donnachie, E.M.; Simon, S.I.; Entman, M.L.; Smith, C.W. Analysis of tight junctions during neutrophil transendothelial migration. *J. Cell. Sci.* **2000**, *113*, 45–75.
- Keese, C.R.; Bhawe, K.; Wegener, J.; Giaever, I. Real-time impedance assay to follow the invasive activities of metastatic cells in culture. *BioTechniques* **2002**, *33*, 842–850.
- Giaever, I.; Keese, C.R. Toxic? Cells can tell. *Chemtech* **1992**, 116–125.
- Keese, C.; Karra, N.; Dillon, B.; Goldberg, A.; Giaever, I. Cell-substratum interactions as a predictor of cytotoxicity. *In Vitro Mol. Toxicol.* **1998**, *11* (2), 183–192.
- Xiao, C.; Lachance, B.; Sunahara, G.; Luong, J. Assessment of cytotoxicity using electric cell-substrate impedance sensing: Concentration and time response function approach. *Anal. Chem.* **2002**, *74* (22), 5748–5753.
- Hadjout, N.; Laevsky, G.; Knecht, D.A.; Lynes, M.A. Automated real-time measurement of chemotactic cell motility. *BioTechniques* **2001**, *31* (5), 1130–1138.
- Ghosh, P.M.; Keese, C.R.; Giaever, I. Monitoring electroporation in the plasma membrane of adherent mammalian cells. *Biophys. J.* **1993**, *64*, 1602–1609.
- Wegener, J.; Keese, C.R.; Giaever, I. Recovery of adherent cells after in situ electroporation monitored electronically. *BioTechniques* **2002**, *33* (2), 348–357.

Electrical Impedance Imaging

Keith D. Paulsen
Alex Hartov

Thayer School of Engineering, Dartmouth College, Hanover, New Hampshire, U.S.A.



INTRODUCTION

Electrical impedance imaging, which is usually referred to as electrical impedance tomography (EIT) but sometimes called electrical impedance spectroscopic imaging (EIS or EITS) when multiple frequencies are involved, is the process of forming spatially resolved electrical property images of tissue using actively driven source and sensing electrodes located on the body surface. The bulk electrical properties of conductivity (Siemens/meter) and permittivity (Farads/meter) are the intrinsic tissue characteristics that the technique directly recovers, although these parameters are often exploited as surrogates for other physiological or biological processes of medical interest (e.g., temperature,^[1] radiation injury,^[2] etc.). EIT is usually distinguished from other electromagnetic (EM) imaging approaches (e.g., magnetic resonance, microwave tomography, etc.) by the frequency range of the applied EM fields—typically 10 KHz to 1 MHz but sometimes as high as 10 MHz or as low as 100 Hz. It is also recognized by its use of exciting currents/voltages exogenous to the body in contrast to other electrical sensing techniques, which detect the body's intrinsic electrical activity (electrocardiogram, electroencephalogram, etc.).

OVERVIEW

There are three basic requirements for electrical impedance imaging: 1) an externally applied electrical stimulus or set of stimuli that can be controlled by the operator and falls within the relevant frequency range (less than 10 MHz); 2) a recording device to measure the magnitude and phase that is relative to the driving signal of the electrical potential and/or current that develops at multiple locations on the tissue surface as a result of the applied stimulus; and 3) an algorithm for converting the data acquired in concert with the known stimulating signals into an image of the electrical property distribution (conductivity and permittivity) in the tissue region of interest. The first two requirements are embodied in the hardware of the imaging instrument itself; whereas, the third requirement involves subsequent software processing of the basic

measurements provided by the hardware data acquisition system. This article explores the tenets that underpin EIT from the perspective of the hardware and software components necessary to generate impedance, i.e., electrical property, images in tissue. Figure 1 illustrates the conceptual framework for EIT in terms of these requisite system modules.

Bioimpedance measurements have a rich history dating back to the early 1930s and before with seminal contributions from Cole^[3] and Schwan^[4] among others. There is extensive literature on the electrical properties of tissue that has been periodically reviewed and catalogued.^[5] One of the most comprehensive compilations of the knowledge base on bioimpedance techniques, including EIT, can be found in^[6] and is recommended reading for any newcomer to this exciting field. A very recent special issue illustrating the state-of-the-art in EIT can be found in.^[7] For the purposes of this synopsis on EIT, it is sufficient to recognize that the induced current density, \mathbf{J} , (in amps/m²) in tissue is related to an applied time-harmonic electric field, \mathbf{E} , (in volts/m) through the distributed analogue of Ohm's Law:

$$\mathbf{J} = \sigma^* \mathbf{E} \quad (1)$$

in which $\sigma^* = \sigma - i\omega\epsilon$ is the complex-valued conductivity that conveniently—in mathematical terms—incorporates the conduction current component due to ion transfer in tissue with the displacement current contribution resulting from capacitive energy storage due to bound (dipolar) charge at cellular membranes. Here, σ is the electrical conductivity, ϵ is the electrical permittivity, ω is the radian frequency of the applied electrical stimulus. Further, the electric field and electric potential, Φ , are related through the equation:

$$\mathbf{E} = -\nabla\Phi \quad (2)$$

under the assumption that the electromagnetic wavelength is sufficiently large relative to the scale of the imaging geometry that the spatial variation in the induced magnetic field (from the time-harmonic current) is small from the perspective of determining the electric field and/or electric potential via Eq. 2. However, note that, whereas the size of $\omega\epsilon$ is generally small (relative to σ) at EIT frequencies, it is possible to detect a phase

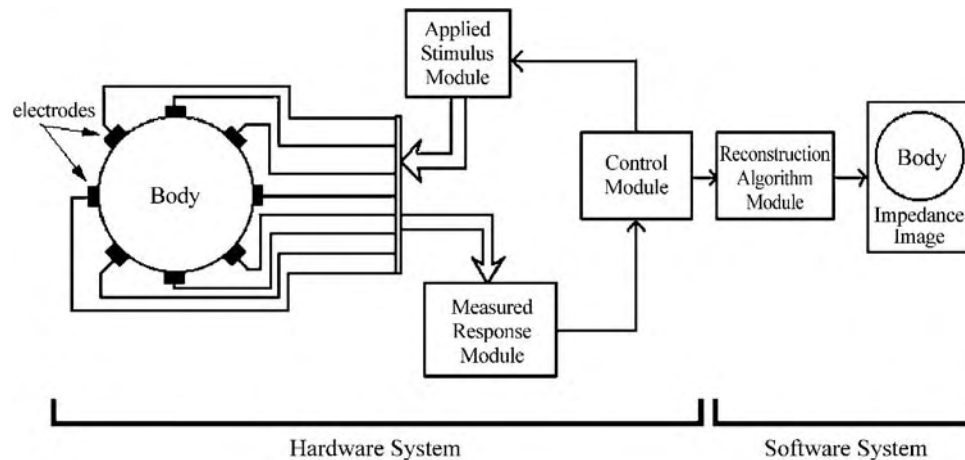


Fig. 1 Conceptual diagram of an EIT system illustrating the primary functional modules needed to generate an image of the body.

difference between an applied current or voltage and a measured voltage or current with well-designed instrumentation, which provides sufficient information for the creation of permittivity images with EIT—even though this does complicate both the hardware and software requirements for electrical impedance imaging.

HARDWARE CONSIDERATIONS

Electrical Impedance Tomography has been developed to image a cross section of the anatomy using signals applied at electrodes on the body surface that define a plane of intersection. As the technology has advanced, it has been extended to three-dimensional configurations in some cases. From the hardware point of view, the implications for developing a three-dimensional system versus a two-dimensional instrument are few—primarily that a three-dimensional system implies a larger number of channels. Typical two-dimensional systems that have been presented in the literature utilize 16 electrodes for breast imaging,^[8] for example, or 32 electrodes for chest imaging.^[9] In this section, some of the basic design decisions for EIT data acquisition, which are the same for either a two-dimensional or a three-dimensional system, will be reviewed.

An additional distinction should be made between tomographic devices and mapping instruments. The latter consist of arrays of electrodes arranged in a planar rectangular or circular configuration that are applied to the area to be examined, which is typically the breast, and measure the load impedance sensed by each electrode. As few as 64 and as many as 256 electrodes have been arranged in handheld probes. The measured impedances are usually represented as shades of gray on a map representing the array configuration. These devices are significantly different in that they do not reconstruct tissue properties in a tomographic

fashion. Two groups have been pursuing these types of devices for breast cancer screening—one in Russia^[10] and one in Israel.^[11] The Israeli group developed the device known as the T-Scan, which is currently marketed in the United States for clinical testing.

It is important to recognize that when an electrode carries a current, a potential will be developed at the electrode-skin interface due to the contact impedance between the two surfaces. This contact impedance can result in erroneous voltage readings. The contact impedance can be reduced by using appropriate electrodes, e.g., Ag–AgCl, and proper skin surface preparation with a conductive gel. However, the artifact is essentially unavoidable. One method to minimize the problem consists of measuring potentials only at electrodes that do not drive currents, which is equivalent to the tetrapolar resistance measurement technique used when lead losses are significant. The measurement approach has been adopted by the pioneering EIT group in Sheffield.^[12] Unfortunately, contact impedance is not the only consideration. From a sensitivity perspective, this approach is less than optimal. For a circular geometry, maximum distinguishability is obtained when trigonometric current patterns are applied on all electrodes simultaneously while voltages are measured at these same sites.^[13] For three-dimensional systems or arbitrary geometries, the optimum current pattern can be computed adaptively.^[14] The effect of the contact impedance is then taken into account during image reconstruction by using a full electrode model.^[15]

Various excitation patterns have been used in EIT. The simplest consists of applying a current between two electrodes and measuring the voltages at all of the remaining channels. Although not optimal, the simplicity of this approach has its appeal. In such a design, a single source can be switched between electrode pairs. With this type of excitation pattern, there is freedom in selecting the electrode pairing. Most designs use

adjacent electrodes as paired partners, which results in the greatest sensitivity occurring at the periphery of the imaged cross section that is nearest the electrodes. The significant lack of sensitivity at the center of the imaged area can be improved by also using patterns in which opposing electrodes are paired. Note that, with a given number of electrodes (N), it is possible to pair them in $N^*(N-1)/2$ ways.

The electronics behind each electrode constitutes a channel and can be designed to apply a desired potential (voltage mode), drive a specific current (current mode), or simply sense the potential at a given location (passive mode). In general, it is much simpler to design a voltage source; however, over extended frequency and load ranges, neither voltage nor current sources behave ideally. As a result, it is preferable to measure both the applied and induced signals. A practical method to achieve measurement duality is to use a sense resistor in the feedback loop of an op-amp buffer. By measuring voltages on both sides of the resistor, the electrode potential is known directly, and the current can be computed. Alternatively, one can use a sense resistor in series with the electrode. This design will result in a potential at the electrode that will vary with the load experienced, which can be corrected by adjusting the driving voltage. The approach is readily implemented in a DSP (digital signal processing)-based design, for example.

Early EIT devices operated at a single frequency, usually in the tens of kilohertz, which made their design relatively simple. Most tissues exhibit significant frequency dependence in their electrical properties, and, in some applications, this dispersion is the basis for discriminating between normal and pathological conditions. The ability to image at multiple frequencies or perform spectroscopy in a spatially resolved manner has therefore become an important goal for EIT. This requirement complicates the level of circuitry involved significantly. It is important to be aware of two critical issues that stem from the desire to realize a multi-frequency system, source design and signal coherence. If current sources are to be used, their output shunt impedance must be very high (on the order of $64\text{ M}\Omega$ to maintain 16 bits of accuracy).^[16] In single frequency circuits, this is achieved by tuning a matching circuit.^[16] The matched circuit approach is not realizable for multiple frequency devices—other than through replication of the current source—and with each tuned to a different frequency. A second difficulty which stems from using high frequencies is signal integrity and coherence. At frequencies above 1MHz, the wire leads delivering the signals to the electrodes couple to ground or to other neighboring leads; hence, signals measured at the sensing point in the electronic circuit cards will be significantly different from signals present at the electrodes.^[17] Below 100 kHz, one can make effective use of

signal guards and shields. A guard consists of a coaxial shield, which is maintained at the same potential as the signal lead it surrounds. This effectively eliminates the capacitive loading between the signal and the shield. Usually, a second grounded shield surrounds the guard to prevent the guard from broadcasting. Unfortunately, at higher frequencies, these shields become ineffective due to small phase and gain errors in the guard-driving circuit. One remedy is to keep the driving and sensing circuits physically located as close as possible to the electrodes. If multiple channels are to be operated simultaneously to produce a spatial pattern with the driving signals, consideration must be given to the relative delay errors between channels in the signal paths. Interchannel propagation differences on the order of ± 5 to ± 10 ns can be expected. At low frequencies, these represent negligible phase or instantaneous amplitude errors; however, at higher frequencies, these can result in significant errors as shown in Table 1.

A compromise must be reached between the speed of measurements, the flexibility of the instrumentation, and its complexity. In the simplest implementation, an EIT system could consist of a single source and a single measurement module that are operated in conjunction with a switching or multiplexing matrix. The source or measurement module would be switched in succession between electrode pairs. Measurements would have to be performed sequentially, which thus increases the acquisition time for recording the data necessary for imaging. Such a system also precludes spatial current distributions that are desirable for maximum sensitivity. Data acquisition speed can be increased by having multiple measurements modules operating in parallel while still using a single source. For a system with maximum flexibility and speed, multiple sources can be used to create arbitrary patterns in combination with parallelized measurements.

For the purpose of image reconstruction, a set of voltages and corresponding currents for each applied pattern is required. In principle, it is of no consequence which quantity (voltage or current) is driven and which is measured. It is generally easier to design a high-fidelity voltage source relative to a current source. It is also simpler to measure a voltage than a current. One additional consideration in deciding whether to use current or voltage sources is the sensitivity of the

Table 1 Effect of differential propagation delays on phase and instantaneous amplitude coherence

Phase and amplitude error based on 10 nS difference				
Frequency (Hz)	50 kHz	100 kHz	1 MHz	10 MHz
Phase (°)	0.180	0.360	3.600	36.000
Amplitude (%)	0.000	0.002	0.197	19.098



measured variable to inaccuracies in electrode placement. For a cylindrical geometry, with electrodes completely covering the end surfaces of a cylinder of length (L) and cross-sectional area (A), the cylinder's resistance (R) can be expressed as: $R = (L \cdot \rho) / A$ in which ρ is the cylinder's resistivity. Using this as an approximation to the relationship between two neighboring electrodes, if a current is applied, the effect on the measured voltage of an error in the distance separating the electrodes will be directly proportional to the size of the error introduced: $V = I \cdot R = I \cdot L \cdot \rho / A$. Conversely, if a voltage is applied, the error in the measured current will go as $1/L$: $I = V / R = V \cdot A / (L \cdot \rho)$.

Differential voltage recordings between neighboring electrodes are potentially more precise than absolute measurements related to a common reference. The reason for this is that differential voltages between neighboring electrodes are usually much smaller, which results in a greater sensitivity for the system given a specific measurement resolution, assuming an identical set of applied currents or voltages. A more important aspect of differential measurements is that they bypass the problem of a common reference for all electrodes that can be noisy, which takes advantage of common mode rejection. Differential measurements are readily achievable at frequencies up to approximately 200 kHz, beyond which they become increasingly inaccurate due to the difficulty in designing differential amplifiers operating at these frequencies. At higher frequencies, absolute measurements are more readily realizable.

When designing a research tool, it is desirable to preserve as much operational flexibility as possible. One approach is to design a modular system that can be operated in any mode.^[18] Each channel consists of a selectable voltage or current source. A relay can also disconnect the driving circuitry, which makes the channel a passive sensing electrode. In order to simplify the hardware design, it is possible to use a voltage source to drive a current by continuously monitoring its current output and adjusting its amplitude.^[19] This is particularly effective with DSP-based designs in which the sensing and adjustment can be performed very rapidly.

Amplitude measurements alone allow one to reconstruct the conductivity map of the tissue cross section of interest. In order to obtain both permittivity and conductivity, it is necessary to obtain the complex voltages and currents at every electrode. Analog methods for recovering in-phase and quadrature (I & Q) components of a signal exist and can be used if a single measurement module is required. Their economical implementation on a per channel basis is difficult to achieve. Alternatives exist that can convert amplitude and phase to DC voltages.^[17] This approach, based on RMS to DC converters and phase detectors and filters, is practical; however, the integration times required for

a precise measurement result in an inherently slower system. As an alternative to these methods, direct sampling of the AC waveforms is possible using fast analog to digital components. The sampled waveforms can be numerically processed to extract magnitude and phase using software-based mixing or quadrature sampling. In order to adequately sample the waveforms, it is not essential to adhere to the Nyquist rate. Under-sampling techniques can be used to obtain effective rates that are very high.^[18] However, the input bandwidth of the sampling subsystem must be higher than the signal frequency. In addition, the sampling needs to be synchronized with the applied signal and across all channels for a coherent data set.

Digital signal processing (DSP) technology makes available some very attractive features for designing an EIT system. With components found commercially, it is possible to create a system that will generate and sample waveforms in lockstep. Furthermore, given the very high clocking frequency of existing devices, it is possible to have timing precision in generating the waveforms of better than 2 ns between channels. Very high-bandwidth, 16-bit analog to digital (A/D) converters are available with actual sampling rates that are on the order of 1 MHz. By combining programmable logic devices and DSP chips, extremely high performance systems can be fabricated. In such a system, the applied signal can be generated using direct digital synthesis. A 256×16 -bit memory buffer is loaded with a numeric sine wave scaled to the desired amplitude. The pattern uses 14 bits that are fed to a fast D/A. An address counter is clocked to scroll through this buffer continuously to generate the output signal. The frequency of the clock cycling through the buffer as well as the number of points used in creating the waveform control the signal frequency ($F_{sig} = F_{cl} / N_{samp-per-cycle}$). The 15th bit is used to clock the sampling A/Ds on each side of the sense resistor, which thus allows synchronous sampling of the voltage and current waveforms.

SAFETY CONSIDERATIONS

Prior to clinical deployment, one must examine issues that pertain to the safety of the patient or imaging subject. Standards have been established regarding the electrical safety of medical devices.^[20,21] The salient feature of these standards is the maximum current that can be allowed to flow through the body. This maximum current is a function of frequency and can be summarized as follows. Below 1 kHz, 1 mA (RMS) or less, from 1 kHz to 100 kHz, the allowable current increases from 1 mA to 10 mA linearly. Beyond that, it remains at 10 mA. When operating at 10 kHz or above, it is relatively simple to couple the signal to the body using a fairly large capacitor in order to suppress DC

and low frequency currents. It should be noted that these safety standards were developed for equipment and were not meant to deliberately apply currents to a subject. A conservative interpretation is that none of the electrodes should be applying a current of magnitude greater than the amount indicated in the guideline. The potential danger of applied currents is of particular concern in chest imaging in which it is possible to induce ventricular fibrillation. The relatively high frequencies (10 kHz and above) of operation of most EIT systems and their comparatively low current densities have proved to be safe. For subjects with pacemakers or other implanted active devices, e.g., neural stimulators, EIT examinations should be avoided due to the risks of direct interference with the electronic operation of these devices.

SOFTWARE CONSIDERATIONS

Perhaps the most critical and difficult aspect of EIT is image reconstruction because the electrical property features of tissue, while attractive candidates for imaging because they offer large intrinsic contrast, are subtly encoded into small changes in surface currents/potentials that are difficult to extract. Further, the sensitivity of surface recordings to electrical property change decreases with distance from the measurement sites, i.e., electrodes, making the depth resolution of EIT relatively poor. It is also important to realize that the measured electrical response is not a linear function of the tissue electrical properties and this nonlinearity, when coupled to the spatial dependence of the sensitivity map to property change, makes for a very challenging and difficult parameter estimation problem.

In general, two classes of image reconstruction approaches have emerged over the years: 1) those that linearize the relationship between electrical property change and change in measured response;^[22] and 2) those that preserve this functional nonlinearity.^[23] In the former case, algorithms are generally fast and analytically motivated but highly approximate and best suited for nonquantitative difference imaging. Typically, the inversion problem has been linearized using variations on the ideas well developed for X-ray CT in which linear inversion is an excellent model. In the case of EIT, the back-propagation of the measured response to the generating source occurs along computed equipotential lines—rather than straight lines as in CT—defined by the governing Laplace equation under assumptions of a homogeneous medium. To develop this further, note that in the absence of any independent sources of time-varying charge within the tissue of interest, \mathbf{J} in Eq. 1 has no divergence, which, when coupled with Eq. 2, produces

$$\nabla \cdot \mathbf{J} = \nabla \cdot \sigma^* \mathbf{E} = \nabla \cdot \sigma^* \nabla \Phi = 0 \quad (3)$$

Equation 3 is readily recognized as the complex-valued Laplace equation in a source-free medium that defines the electric potential everywhere within the tissue. Because the electric potentials spread apart with distance from their source, pixel contributions to the measured response must be weighted appropriately, depending on their location within the equipotential map constituting a source-receive pair. The process leads to images that are typically rather blurred. Various strategies for improving image quality through filtering schemes have been devised.^[24] However, the approach is ultimately limited because the linear approximation of the relationship between measurement change and property change is essentially invalid.

The nonlinear image reconstruction approach is more promising, but it can only be realized at the expense of considerable computational costs that are also accompanied by instabilities due to measurement errors introduced in the data by the hardware acquisition. The functional nonlinearity forces stepwise linearization that is iteratively incremented as part of the solution. Further, because of the complexities associated with solving Eq. 3 when σ^* is arbitrarily defined, image reconstruction algorithms that seek to enforce nonlinearity are numerical in nature.

Perhaps the most popular approach in this regard is the finite element (FE) method,^[25,26] which is the technique that will be summarized here. It is important to recognize that, while finite element reconstruction for EIT has evolved to the point where codes are available for use,^[27,28] new image reconstruction methods continue to emerge from within the research community that hold exciting new prospects for improving EIT.^[29] Nonetheless, the current state-of-the-art in EIT image reconstruction is well represented by the finite element method. The most straightforward approach to nonlinear EIT image reconstruction minimizes the squared difference between measured and computed potentials (or currents) at a discrete set of electrodes:

$$\min \sum_{i=1}^0 (\Phi_i^m - \Phi_i^c)^2 \quad (4)$$

positioned on the body surface around the region of interest. In Eq. 4, the superscripts on Φ refer to measured (M) and computed (C) quantities, and the sum extends over the total number of measurement observations. For simplicity, Φ in Eq. 4 is assumed to be real-valued, and yields recovery of only the real-valued electrical conductivity in Eq. 1. Generalization of Eq. 4 to include permittivity estimation requires the introduction of complex-plane analysis, which is straightforward but serves to unnecessarily complicate the notation needed for the summary presented here. Analogous development for complex-valued image reconstruction can be found in.^[26,30]



Minimization of Eq. 4 requires determination of first-derivative stationarity:

$$\sum_{i=\phi}^0 (\Phi_i^m - \Phi_i^c) \frac{\partial \Phi_i^c}{\partial \sigma} = 0 \quad (5)$$

At this point, a number of mathematical strategies can be invoked. For example, the nonlinear set of equations in Eq. 5 can be solved iteratively for the σ -distribution via Newton's method in standard matrix form in which the second derivative of Φ_i^c with respect to σ is assumed to be small. Alternatively, the nonlinearity in the relationship between the potential and the property distribution can be linearized by Taylor series expansion of Φ^c around the current σ estimate:

$$\Phi^c(\sigma) \cong \Phi^c(\bar{\sigma}) + \partial_\sigma \Phi^c(\bar{\sigma}) \Delta\sigma + HOT \quad (6)$$

in which $\bar{\sigma}$ represents the current estimate, $\Delta\sigma = \sigma - \bar{\sigma}$; ∂_σ denotes partial differentiation with respect to σ ; and *HOT* indicates the higher-order-terms in the Taylor series. Substitution of Eq. 6 into Eq. 5 leads to the classical form for updating an initial conductivity distribution estimate:

$$\Delta\sigma = (J^T J)^{-1} J^T (\Phi^M - \Phi^c) \quad (7)$$

in which J is a matrix of derivatives, $\partial_\sigma \Phi_i^c$, of each observable for $i=1,2,\dots,0$ (e.g., potentials at a set of electrodes) with respect to each parameter for $k=1,2,\dots,M$ defining the conductivity distribution. Equation 7 is numerically unstable due to measurement errors and must be smoothed or regularized. Conceptually, this can be achieved by representing the *HOT* in Eq. 6 in various ways, for example, as:

$$HOT = \lambda (\partial_\sigma \Phi^c)^{-1} \Delta\sigma \quad (8)$$

in which λ is a scaling factor for improving stability, in which case Eq. 7 becomes:

$$\Delta\sigma = (J^T J + \lambda I)^{-1} J^T (\Phi^M - \Phi^c) \quad (9)$$

Reconstruction of finite elements (i.e., iteratively solving Eq. 9) is centered around FE solution of two partial differential equations:

$$\nabla \cdot \sigma \nabla \Phi = 0 \quad (10)$$

that is, Eq. 3, under assumption of the current estimate of the σ distribution and:

$$\nabla \cdot \sigma \nabla \Phi' = -\nabla \cdot \sigma' \nabla \Phi \quad (11)$$

in which the prime indicates differentiation with respect to a localized variation in the conductivity distribution.

Discretization of Eq. 10 by finite elements proceeds through the method of weighted residuals combined with basis function expansion of Φ and σ to yield the algebraic system of equations:

$$\sum_{j=1}^N \Phi_j \left\langle \sum_{k=1}^M \sigma_k \phi_k \nabla \phi_j \cdot \nabla \phi_i \right\rangle = \oint J \phi_i ds \quad (12)$$

for $i=1,2,\dots,N$ in which ϕ_i is a locally defined Lagrange polynomial (typically linear) centered at node i , $J = \sigma \nabla \Phi \cdot \hat{n}$ is the current density integrated as \oint around the boundary, and $\langle \cdot \rangle$ represents domain integration over the imaging body. N is the total number of nodes in the FE mesh on which Φ is computed, i.e., Φ_j is the potential solution at node j , whereas M is the number of parameters that defines the σ distribution (often discretized on a separate mesh such that σ_k is the conductivity at node k). In matrix form, it is convenient to partition Eq. 12 into quantities on the boundary (subscript b) and interior (subscript I) to the imaging field-of-view:

$$\begin{bmatrix} A_{bb} & A_{bI} \\ A_{Ib} & A_{II} \end{bmatrix} \begin{Bmatrix} \Phi_b \\ \Phi_I \end{Bmatrix} = \begin{bmatrix} B_{bb} & 0 \\ 0 & 0 \end{bmatrix} \begin{Bmatrix} J_b \\ 0 \end{Bmatrix} \quad (13)$$

in which the submatrices symbolize integration of the known basis functions and property assumptions depending on their location, i.e., boundary or interior.

Equation 11 is solved analogously for Φ' through FE discretization and yields the algebraic equations:

$$\begin{aligned} & \sum_{j=1}^N \Phi'_j \left\langle \sum_{k=1}^M \sigma_k \phi_k \nabla \phi_j \cdot \nabla \phi_i \right\rangle \\ &= - \sum_{j=1}^N \Phi_j \langle \phi_\ell \nabla \phi_j \cdot \nabla \phi_i \rangle \\ & \quad + \oint \sigma \nabla \Phi' \cdot \hat{n} \phi_i ds + \oint \sigma' \nabla \Phi \cdot \hat{n} \phi_i ds \\ &= - \sum_{j=1}^N \Phi_j \langle \phi_\ell \nabla \phi_j \cdot \nabla \phi_i \rangle + \oint J' \phi_i ds \end{aligned} \quad (14)$$

for $i=1,2,\dots,N$ in which the prime denotes differentiation with respect to the ℓ th parameter, σ_ℓ , in the sum $\sum_{k=1}^M \sigma_k \phi_k$. The matrix system in Eq. 14 is computed for each of the M terms in the sum to fill in the elements composing J in Eq. 9. It is likewise conveniently partitioned into boundary and interior contributions:

$$\begin{bmatrix} A_{bb} & A_{bI} \\ A_{Ib} & A_{II} \end{bmatrix} \begin{Bmatrix} \Phi'_b \\ \Phi'_I \end{Bmatrix} = - \begin{bmatrix} A'_{bb} & A'_{bI} \\ A'_{Ib} & A'_{II} \end{bmatrix} \begin{Bmatrix} \Phi_b \\ \Phi_I \end{Bmatrix} + \begin{bmatrix} B_{bb} & 0 \\ 0 & 0 \end{bmatrix} \begin{Bmatrix} J'_b \\ 0 \end{Bmatrix} \quad (15)$$

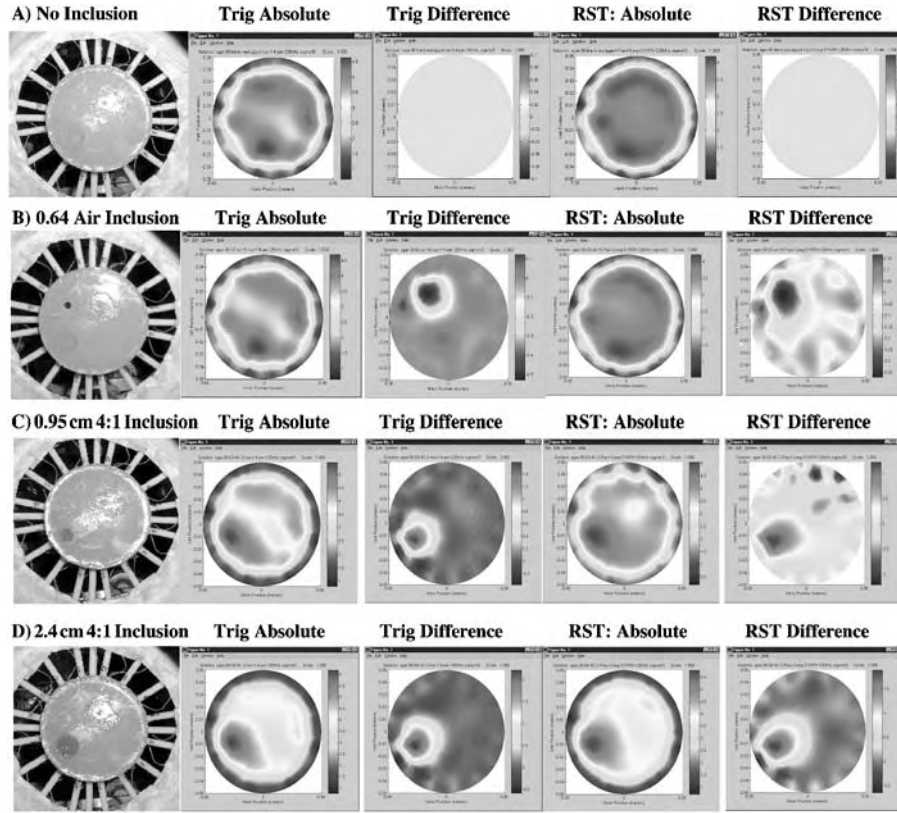


Fig. 2 Representative images obtained from an agar study involving systematic variations in inclusion contrast, size, and depth. From left to right across each row: photograph of the experimental phantom agar setups; absolute conductivity image at 125 kHz using trigonometric patterns; difference conductivity image using trigonometric patterns; absolute conductivity image using RST patterns; and difference conductivity image using RST patterns. The inclusion in the second row is 2 cm from the edge of the agar, whereas the inclusions in the third and fourth rows are 1 cm from the edge.

An important part of EIT image reconstruction is electrode modeling^[25] because the discrete finite size of the electrodes utilized in practical imaging influences the measured potentials and currents. The framework for electrode modeling is readily developed on finite elements starting from the weighted residual weak form given in Eq. 12. As a two-dimensional example, a finite-length electrode might consist of three connected nodes along a small portion of the boundary. In voltage mode, the potential is specified at the electrode; hence, Eq. 12 for i corresponding to the three boundary nodes composing the electrode are removed from the algebraic system in Eq. 13 in favor of the known potential that is uniform across the electrode. These three Galerkin equations contain the relationships for computing the corresponding current, which is compared to the

measured value during the image reconstruction process. To obtain the current, these Galerkin equations are summed over the electrode nodes after the solution for Φ is computed, subject to the potential boundary conditions on the electrode:

$$\begin{aligned} \sum_{i=1}^3 \sum_{j=1}^N \Phi_j \left\langle \sum_{k=1}^M \sigma_k \phi_k \nabla \phi_j \cdot \nabla \phi_i \right\rangle &= \sum_{i=1}^3 \oint J \phi_i ds \\ &= \oint J \sum_{i=1}^3 \phi_i ds = \oint J ds = I \end{aligned} \quad (16)$$

to provide exactly the integrated current density along the electrode length (because $\sum \phi_i = 1$) which produces the value of interest. In Eq. 16, i runs from 1 to 3, which

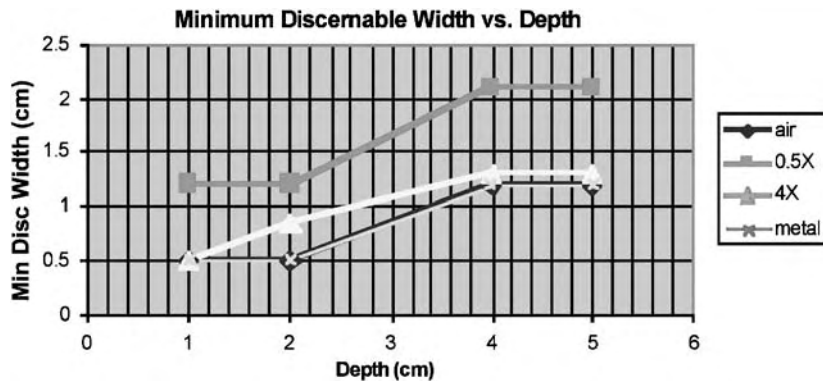


Fig. 3 Graph of minimum discernable width (MDW) as a function of depth for various contrast inclusions embedded in (1X) agar. The graph is based on conductivity difference images at 50 kHz.

represent local node numbers spanning the three nodes that compose the electrode in this illustrative example.

In current mode, the summation in Eq. 16 is performed; however, this single Galerkin equation is retained in the matrix system for the unknown potential within the imaging region and is located on the row associated with a preselected nodal position for the electrode, i.e., the center node, locally numbered 2 in this example. The right-hand side of Eq. 16 is the driving current supplied by the hardware system to the electrode that becomes the known boundary condition for this position on the mesh surface. At this point, the matrix system in Eq. 13 is under-determined because two degrees-of-freedom (potential at local nodes 1 and 3) lack corresponding governing equations. Here, the

boundary constraints:

$$\Phi_1 - \Phi_2 = 0 \tag{17a}$$

$$\Phi_3 - \Phi_2 = 0 \tag{17b}$$

are applied that enforce a uniform potential across the electrode and supply the missing relationships needed to close the algebraic system.

Modeling the contact impedance follows similarly through use of a mixed boundary condition:

$$\sigma \nabla \Phi \cdot \hat{n} = J = \alpha(\Phi - V) \tag{18}$$

in which α is an electrode contact-coupling coefficient relating the current density to the voltage drop across

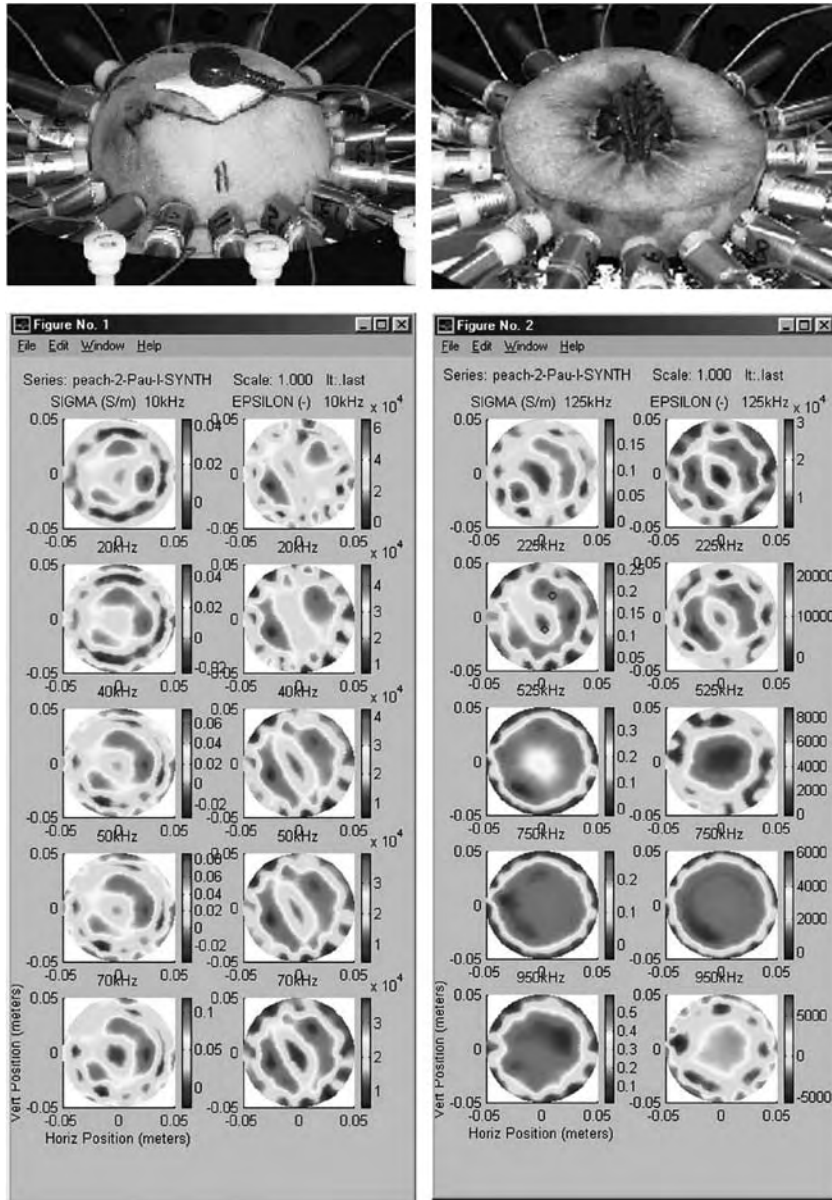


Fig. 4 Photographs of a peach situated in the electrode array during imaging (top, left) and at the end of the experiment after being cut in half (top, right) prior to removal from the array to reveal the orientation of the pit. Corresponding conductivity and permittivity images of the peach from 10–950 kHz using the RST excitation patterns.

the electrode, given V is the voltage recorded at the electrode itself. This condition is substituted into the boundary integral in Eq. 16:

$$\begin{aligned} \sum_{i=1}^3 \oint \alpha(\Phi - v)\phi_i ds &= \oint \alpha\Phi ds - \oint \alpha v ds \\ &= \alpha\Phi L - \alpha VL \end{aligned} \quad (19)$$

in which L is the electrode length. The first term in Eq. 19 is added to the system matrix of unknowns (locally Φ_2 —the potential at the center of the electrode) while the second term becomes the known driving data for the electrode. As in current mode, the missing constraints required to close the algebraic system are Eqs. 17a and 17b, which enforce a uniform potential along the electrode length. Once the potential distribution is calculated, the electrode current is readily computed from Eq. 19.

Note that, if a passive electrode (i.e. no driving current) is utilized in the hardware system configuration, the electrode model still applies. In this case, the right-hand side of Eq. 16 is set to zero, and the constraints in Eqs. 17a and 17b are invoked. The model leads to the situation in which the potential is uniform along the electrode but no current flows out of the system through this portion of the boundary, which are precisely the conditions that occur in practice and are to be emulated in the reconstruction process.

EXAMPLE IMAGES

Typically, prototype EIT instrumentation and concomitant image reconstruction software is evaluated in phantom studies to characterize its performance. A standard benchmark is to image high contrast conductors and insulators of various sizes at a series of depth locations

within a uniform background of either physiological saline or conductive solid agar intended to mimic the electrical properties of typical tissue. Figure 2 shows a series of typical images from a collection of experiments involving various inclusion sizes, locations, and contrasts. Both absolute and difference images are shown with two types of excitation patterns, Trigonometric (trig) and reversed synthesis trigonometric (RST). Standard practice is a complete contrast-detail analysis in which image recovery as a function of the contrast and size of a heterogeneity is investigated. For linear systems, this can be completed without regard to the location of the object. However, in the case of EIT, the response function is nonlinear and spatially dependent. Hence, the quality of the image will depend not only on the size and contrast of a heterogeneity but also on its location.

Figure 3 presents a typical contrast-detail analysis for EIT-based on agar phantoms. It shows the size of an object required for detection that is defined in terms of minimum discernable width (MDW) at a given depth for specific contrast levels. The degradation in depth resolution with EIT is evident. Objects less than 1 cm in diameter were not detectable in absolute EIT images when positioned at the center of a 10-cm diameter agar phantom, even with high contrast. It is also clear from Fig. 3 that the depth resolution of difference images is superior to absolute images.

It is important to recognize that, because agar lacks the cellular membrane structures present in tissue, the permittivity component of impedance is extremely small and not reliably imaged. To generate sufficient capacitance, it is necessary to image organic materials, and a suitable set of candidates for phantoms include fruits and vegetables available from the local market. Figure 4 shows results from a sample experiment using a peach as the object to be interrogated. The images were generated with a 16-electrode array attached to the surface

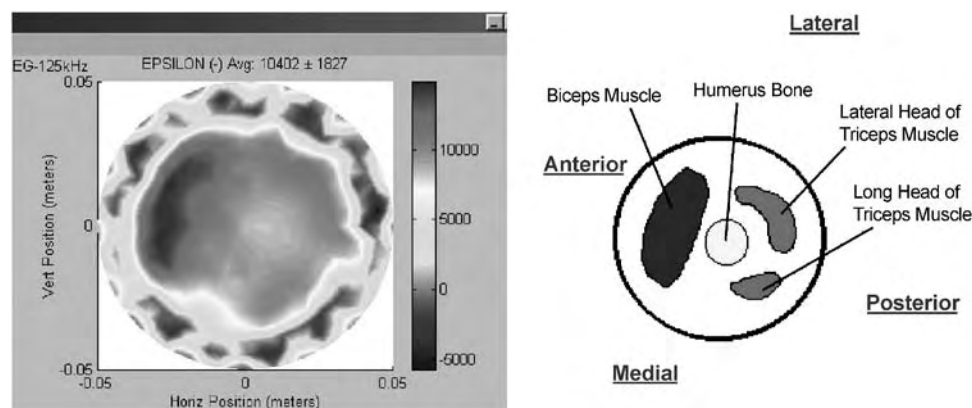


Fig. 5 In-vivo permittivity image of the upper arm accompanied by anatomical sketch indicating that the expected locations of the humerus bone and triceps and biceps muscle masses appear in the correct positions in the EIT image.

of the peach using EIT instrumentation capable of recording both the signal amplitude and phase. The recovered permittivity images are now meaningful and nicely illustrate the details of the peach anatomy. In this particular experiment, multifrequency data acquisition was employed. Both the conductivity and permittivity images indicate the orientation and location of the pit of the peach over the frequency range of 70 kHz to 225 kHz in which this EIT instrument operation is optimal. Independent measurements of the electrical properties show that the reconstructed images are quantitatively accurate to 20% of the true values across the frequency band illustrated here.

As a final in-vivo example, Fig. 5 shows the permittivity image of the upper arm of a human volunteer. The humerus bone and biceps and triceps muscle masses are evident in the image. These results were generated using a multifrequency EIT device with 16 electrodes operating in voltage mode. The anatomical detail and tissue-type discrimination within this image is encouraging.

CONCLUSION

Electrical impedance tomography uses single or multiple low-frequency signals (below 10 MHz) to probe the body from which electrical property distributions are estimated. The system is based on contact electrode arrays that serve as source and sensing elements during the process. Hardware considerations indicate that voltage sources are easier to produce than current sources and simultaneous measurement of both applied and induced signals is recommended, especially in the multifrequency setting. Costs per channel are relatively modest, which allows high-density two-dimensional and three-dimensional arrays to be developed. DSP-based designs are particularly attractive because of their superior speed and flexibility.

Image reconstruction for EIT is particularly challenging because of the nonlinear relationship between sensed signal change and change in the underlying tissue electrical properties. Iterative methods become essential in this regard and, to date, have yielded the highest image quality. Nonlinear image reconstruction centered around finite element technology is particularly popular, and several codes are available on the Internet. Sample images show that the depth resolution with EIT degrades significantly. However, recovery of distributions of both electrical conductivity and permittivity is clearly possible, provided adequate phase shift can be recorded. In-vivo images of the upper arm demonstrate clear permittivity variation between the humerus bone and the triceps and biceps muscle masses. State-of-the-art in EIT can be understood by consulting a recent special issue journal on the topic.^[7]

ARTICLES OF FURTHER INTEREST

Magnetic Resonance Imaging in Temperature Measurement; Microcomputed Tomography and Its Applications; Nuclear Magnetic Resonance Spectroscopy; Optical Coherence Tomography; Positron Emission Tomography (PET); Tissue Engineering of Blood Vessel; Ultrasound

REFERENCES

1. Bolomey, J.C.; Hawley, M.S. Non-Invasive Control of Hyperthermia. In *Clinical Thermology Subseries Thermotherapy*; Gautherie, M., Ed.; Springer: Berlin, 1989; 35–111.
2. Paulsen, K.D.; Osterman, K.S.; Hoopes, P.J. In vivo electrical impedance spectroscopic monitoring of the progression of radiation-induced tissue injury. *Radiat. Res.* **1999**, *152*, 41–50.
3. Cole, K.S. Electric phase angle of cell membranes. *J. Gen. Physiol.* **1932**, *15*, 641–649.
4. Schwan, H.P.; Li, K. Capacitance and conductivity of body tissues at ultra-high frequencies. *Proc. IRE* **1953**, *41*, 1735–1740.
5. Duck, F. *Physical Properties of Tissues: A Comprehensive Reference Book*; Academic Press: London, 1990.
6. Critical Reviews in Biomedical Engineering. Bourne, J.R., Ed.; 1996; Vol. 24. (Issues 4–6).
7. *Special Issue on Electrical Impedance Tomography*; IEEE Trans. Med. Imaging Ed.; Newell, J.C., Isaacson, D., Mueller, J.L., Eds.; 2002; Vol. 21.
8. Osterman, K.S.; Kerner, T.E.; Williams, D.B.; Hartov, A.; Poplack, S.P.; Paulsen, K.D. Multifrequency electrical impedance imaging: Preliminary in vivo experience in breast. *J. Phys. Meas.* **2000**, *21* (1), 99–109.
9. Newell, J.C.; Isaacson, D.; Cheney, M.; Saulnier, G.J.; Gisser, D.G.; Goble, J.C.; Cook, R.D.; Edic, P.M. Impedance images of the chest. *Proc. IEEE-EMBS Conf.* **1992**, *14*, 1752–1753.
10. Cherepenin, V.A.; Karpov, A.Y.; Korjnevsky, A.V.; Kornienko, V.N.; Kultiasov, Y.S.; Ochapkin, M.B.; Trochanova, O.V.; Meister, J.D. Three-dimensional EIT imaging of breast tissues: System design and clinical testing. *IEEE Trans. Med. Imag.* **2002**, *21* (6), 662–667.
11. Assenheimer, M.; Laver-Moskovitz, O.; Malonek, D.; Manor, D.; Nahaliel, U.; Nitzan, R.; Saad, A. The T-Scan technology: Electrical impedance as a diagnostic tool for breast cancer detection. *Physiol. Meas.* **2000**, *22*, 1–8.
12. Barber, D.C.; Brown, B.H. Applied potential tomography. *J. Phys. E, Sci. Instrum.* **1984**, *17*, 723–733.
13. Cheney, M.; Isaacson, D. Distinguishability in impedance imaging. *IEEE Trans. Biomed. Eng.* **1992**, *39* (8), 852–860.
14. Hua, P.; Woo, E.J.; Webster, J.G.; Tompkins, W.J. Improved methods to determine optimal currents in electrical impedance tomography. *IEEE Trans. Med. Imag.* **1992**, *11* (4), 488–495.

15. Somersalo, E.; Cheney, M.; Isaacson, D. Existence and uniqueness for electrode models for electric current computed tomography. *SIAM J. Appl. Math.* **1992**, *52* (4), 1023–1040.
16. Cook, R.D.; Saulnier, G.J.; Gisser, D.; Goble, J.C.; Newell, J.C.; Isaacson, D. ACT3: A high-speed, high-precision electrical impedance tomograph. *IEEE Trans. Biomed. Eng.* **1994**, *41* (8), 713–722.
17. Hartov, A.; Kerner, T.E.; Markova, M.T.; Osterman, K.S.; Paulsen, K.D. Dartmouth's next generation EIS system: Preliminary hardware considerations. *Physiol. Meas.* **2001**, *22*, 25–30.
18. Hartov, A.; Mazzaresse, R.; Reiss, F.; Kerner, T.; Osterman, S.; Williams, D.; Paulsen, K.D. A multi-channel continuously-selectable multi-frequency electrical impedance spectroscopy measurement system. *IEEE Trans. Med. Imag.* **2000**, *47* (1), 49–58.
19. Hartov, A.; Demidenko, E.; Soni, N.; Markova, M.; Paulsen, K. Using voltage sources as current drivers for electrical impedance tomography. *Meas. Sci. Technol.* **2002**, *13*, 1425–1430.
20. IEC International Standard. *Medical Electrical Equipment. Part 1: General Requirements for Safety*; 1993. Document reference number IEC 601-1-2.
21. Association for the Advancement of Medical Instrumentation. *American National Standard. Safe Current Limits for Electromedical Apparatus*; Dec. 2 1993. ANSI/AAMI ES1-1993, Revision of ANSI/AAMI ES1-1985.
22. Barber, D.C.; Brown, B.H. Imaging spatial distribution of resistivity using applied potential tomography. *Electron. Lett.* **1983**, *19* (22), 933–935.
23. Yorkey, J.Y.; Webster, J.G.; Tompkins, W.J. Comparing reconstruction algorithms for electrical impedance tomography. *IEEE Trans. Biomed. Eng.* **1987**, *BME-34* (11), 843–852.
24. Barber, D.C.; Brown, B.H.; Avis, N.J. Image reconstruction in electrical impedance tomography using filtered back-projection. *Proc. 14th Annu. Int. Conf. IEEE-EMBS* **1994**, *14*, 1691–1692.
25. Vauhkonen, P.J.; Vauhkonen, M.; Savolainen, T.; Kaipio, J.P. Three-dimensional electrical impedance tomography based on the complete electrode model. *IEEE Trans. Biomed. Eng.* **1999**, *46* (9), 1150–1160.
26. Paulsen, K.D.; Jiang, H. An enhanced electrical impedance imaging algorithm for hyperthermia applications. *Int. J. Hypertherm.* **1997**, *13* (5), 459–480.
27. Polydorides, N.P.; Lionheart, W.R.B. A MATLAB toolkit for three dimensional electrical impedance tomography: A contribution to the EIDORS project. *Meas. Sci. Technol.* **2002**, in press.
28. Vauhkonen, M.; Lionheart, W.R.B.; Heikkinen, L.M.; Vauhkonen, P.J.; Kaipio, J.P. A Matlab package for the EIDORS project to reconstruct two-dimensional EIT images. *Physiol. Meas.* **2001**, *22*, 107–111.
29. Mueller, J.L.; Siltanen, S.; Isaacson, D. A direct reconstruction algorithm for electrical impedance tomography. *IEEE Trans. Med. Imag.* **2002**, *21*, 555–559.
30. Hartov, A.; Kerner, T.E.; Paulsen, K.D. Simulation of error propagation in finite element image reconstruction for electrical impedance tomography. *Meas. Sci. Technol.* **2001**, *12*, 1040–1049.

Electroactive Polymeric Materials

V. Prasad Shastri

Vanderbilt University, Nashville, Tennessee, U.S.A.

INTRODUCTION

Electroactive materials may be broadly defined as those materials whose bulk or surface properties may be altered reversibly or irreversibly upon exposure to an electrical stimulus. In keeping with the scope of this encyclopedia, only materials that have potential application in drug delivery, biosensing, tissue engineering, and high throughput screening will be discussed. The article is divided into four sections. The first two sections cover conductive and piezoelectric polymers in detail. Electroresponsive gels are discussed briefly in the third section and the last section covers more general topics that are material specific. Electrets are not covered in this article, as examples of these are few and their biomedical applications very limited.

OVERVIEW

Since the early 1980s the application of polymers in biomedical sciences has seen an exponential growth. This growth has been fueled in great part by the demand placed by emerging disciplines, such as drug delivery, tissue engineering, and genetic screening, for novel materials that may be fine-tuned with respect to their bulk and surface characteristics on demand as a function of time. Electroactive materials, by virtue of their ability to predictably respond to electrical stimulus, are therefore ideal candidates.

The role of electrical signals in modulating cellular response is well-recognized.^[1] Among various tissues where bioelectricity plays a role, the nervous system and osseous tissue are the most obvious examples. The neural synapse in essence is a biosensor. An action potential (electrical signal) arriving at the synaptic junction is transmitted across the synapse by depolarization of the adjacent synaptic membrane. This depolarization is triggered by acetylcholine, a biomolecule that is secreted in response to the action potential at the synapse.^[2] Bone is yet another tissue where electrical activity plays an important role in fulfilling its function. In the late 1890s, Wolff demonstrated that the bone responds to mechanical stress by growing.^[3,4] This relationship between applied mechanical stress and bone remodeling (growth) is known as Wolff's Law. It turns out that bone is piezoelectric in

nature. Iwao Yasuda was the first to demonstrate piezoelectricity in bone in the 1950s when he observed that compressive loading along a bone's major axis resulted in the accumulation of electrical charges on the surface along its axis.^[5,6] In a seminal study, Becker and Bassett demonstrated that when bone was mechanically loaded, the concave (bone formation) and convex (bone resorption) faces of bone yielded charges of opposing polarity due to compression and tension at the two respective faces.^[7] Therefore, cellular architecture in bone is able to sense the mechanical stress due to the localized electrical currents generated upon loading of the inorganic hydroxyapatite phase in bone.

Electroactive materials encompass ceramics and plastics that are capable of electro-mechanical and electrochemical transduction. However, two major considerations for biomedical applications are cellular and tissue compatibility and ease of processing. The library of organic chemical bonds and building blocks available for the synthesis of polymers make it a very versatile system capable of satisfying both of the aforementioned requirements. In emerging applications such as tissue engineering (TE), wherein a scaffold is used to guide the generation of neotissue from dissociated mammalian cells, it is plausible that the use of electroactive scaffolds will offer the possibility of manipulating cellular functions and tissue growth in a localized manner via minimally invasive methodologies.

CONDUCTIVE POLYMERS

Ease of processibility and their insulative properties are two important characteristics that distinguish plastics from metals. In the 1970s, Heeger, MacDiarmid, and Shirakawa, demonstrated that polyacetylene, an organic polymer characterized by alternating double bonds (Fig. 1a), can exhibit conductivity approaching that of conductors. Conductivities as high as that of copper (10^6 S/cm) have been reported in polyacetylenes doped with a cation or an anion.^[8,9] Their seminal work, for which they were awarded the 2000 Nobel Prize in Chemistry, led to the birth of an entirely new class of materials and an exciting field called conductive polymers. The origin of conductivity in this interesting class of polymers lies in the high degree of π - π overlap in a conjugated structure, which reduces the band gap

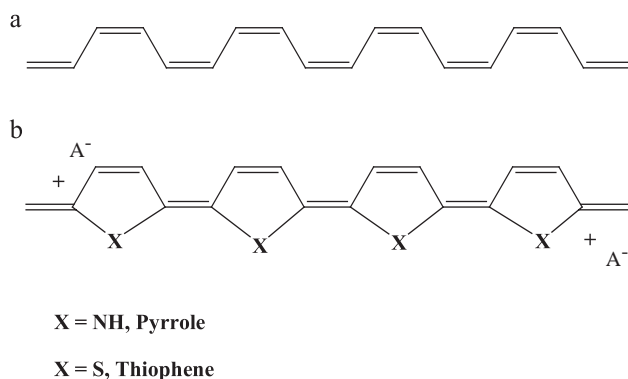


Fig. 1 (a) Chemical structure of oxidized polyacetylene; (b) Chemical structure of oxidized poly(heterocycles).

and facilitates carrier mobility.^[10] Since then several polymers that exhibit a high degree of extended polyconjugation have been shown to possess electrical conductivity in their doped state. They include poly(aniline), poly(heterocycles), poly(*p*-phenylene vinylene) and poly(*p*-phenylene ethynylene).

In 1979, Diaz and coworkers reported the electrochemical synthesis of polypyrrole (Ppy) (Fig. 1b).^[11] With the electrochemical approach, freestanding films of Ppy of varying thickness and crystallinity can be obtained by controlling the passage of current during electrochemical deposition and chemistry of the dopant anion. Three years later, Tourillon and Garnier reported the electrochemical polymerization of poly(thiophene) (Pth) from thiophene.^[12] Electrochemical polymerization offers several advantages over conventional chemical synthesis: 1) Thin films (membranes) of the polymer can be directly deposited on an electrode of interest or surface of a metal implant; and 2) freestanding thin films can be obtained that can then be fabricated into devices. Both Ppy and Pth exhibit reversible electrochemistry, which enables switching of the polymer from the conductive form (oxidized) to the insulator form (neutral) by the application of a small voltage.^[13] Electrical conduction in poly(heterocycles) occur via the interchain hopping and delocalization of radical cations called polarons.^[14] The conductivity in Ppy and Pth is significantly influenced by the crystallinity and long-range order in the polymer, which aids in reducing the conduction band gaps, the latter being more important. Among the various conductive polymers, Ppy and Pth (Fig. 1b) are the most promising for biomedical applications due to ease of their synthesis. However, Ppy is by far the most suited for biosensing and tissue-contacting applications as it can be electrochemically synthesized from aqueous solutions at a relatively low potential of 700 mV versus SEC. Furthermore, its oxidized form is far more stable in both air and water (physiological conditions of pH 7.2 phosphate buffer, 37°C) in comparison to Pth.

Since the mid-1980s Ppy has been extensively studied as an electrochemical transducer for sensing biomolecules in solution.^[15–17] In a Ppy-based biosensor, an oxidative enzyme such as glucose oxidase is entrapped in the film during electrochemical deposition on a platinum or palladium electrode (Fig. 2). When this film is brought into contact with an analyte such as glucose in presence of oxygen, hydrogen peroxide is liberated which is then detected electrochemically at the electrode.^[18–22] The advantage of this system is that it allows for the one-step fabrication of a sensor. However, the slow response time, due to the slow diffusion of hydrogen peroxide through the film to the electrode, and low current densities are limitations. These limitations have been overcome by introducing redox couple mediators, such as ferrocene (Fc) (II) ↔ ferrocinium (Fc⁺) (III), either by coentrapment or by covalent immobilization to the enzyme or polymer, that can act as an electron shunt.^[23–26] In this system, the electron transfer occurs directly from enzyme to the oxidized form of the mediator Fc⁺, yielding Fc which is then followed by cyclic voltammetry. In addition to the Fc ↔ Fc⁺ couple, other charge transfer systems such as tetrathiafulvalene-tetracyanoquinodimethane (TTF-TCNQ)^[27] and 4,4'-bipyridyl mediators have been explored as well.^[28] Polymers bearing such electron mediators are said to be wired and they typically yield sensors with faster response times and higher current densities. Besides enzyme-mediated oxidative reactions, changes in local pH have been leveraged in the development of Ppy-based biosensors. Ppy membrane microarrays containing immobilized penicillinase have been used to detect penicillin by monitoring the increase in Ppy membrane conductivity upon the release of acid at the membrane surface due to enzymatic hydrolysis of the penicillin.^[29] Ppy membranes/films with immobilized oligonucleotides have

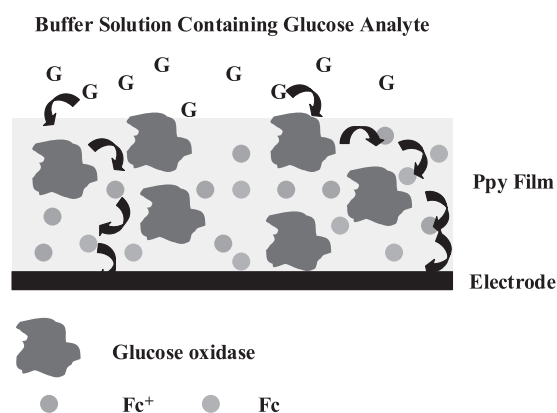


Fig. 2 Enzyme-based conductive polymer biosensor: schematic of a Ppy-based glucose sensor containing glucose oxidase and ferrocene as an electron mediator.

been prepared by electrochemical polymerization of pyrroles covalently bound to oligonucleotides.^[30,31] By individually addressing microelectrodes in an array, surfaces bearing a spatially well-defined array of oligonucleotides have been fabricated. Such arrays might be useful in the development of high throughput devices for genomics and proteomics that are amperometric rather than colorimetric, thus increasing reliability and response time. In addition to Ppy, Pth and polyanilines have been explored in glucose-sensing applications.^[32,33]

In the late 1980s, Miller and coworkers demonstrated that a Ppy film doped with a polyanionic dopant poly(styrenesulfonate) sodium (PSS-Na) can behave as a cation exchanger.^[34,35] The immobilization of PSS-Na results in a film that has a net negative charge in its oxidized state. Because these charges are fixed, reduction of the film results in the binding of cations, which can be released upon oxidation of the film. Based on this principle, Miller and co-workers showed that dopamine, a neurotransmitter whose deficiency leads to lesions in the brain such as those found in Parkinson's disease, may be cathodically bound and anodically released from these Ppy films.^[34] Prezyna and co-workers later showed that these films can irreversibly bind polycations such as poly(lysine-HCl) and histones, which are responsible for compacting DNA in the nucleus of a cell.^[36] However, extensive exploration of Ppy films as drug-delivery vehicles has been limited by unpredictability associated with binding and release kinetics and low drug loading levels. These systems, however, are well suited for highly potent drugs and should not be overlooked especially in the delivery of drugs to the central nervous system. Diseases such as epilepsy and Parkinson's could significantly benefit from self-regulating delivery systems wherein the Ppy electrode serves as a biosensor for disease-predictive markers, thereby reducing the drug delivery response time.

Ppy has been explored in TE applications as well. Wong^[37] and Shastri^[38,39] have shown that by varying the dopant ion chemistry and oxidation state of Ppy, protein adsorption and conformation can be altered resulting in control over cell cycle progression. This has sparked an interest in using Ppy in cell and tissue contacting applications as interactive interfaces for dictating cell behavior and enhancing cell functions. In a seminal study demonstrating the potential of Ppy as an electroactive interface in manipulating cellular functions, the effect of applying of an electrical stimulus through a Ppy substrate on PC-12 cell differentiation was studied.^[40] PC-12 cells respond reversibly to soluble nerve growth factor by differentiation into a neuronal phenotype, a process accompanied by extension of neurites (neuronal processes). In this study, PC-12 cells grown on Ppy-PSS-Na

substrates were subjected to an electrical stimulus applied through the polymer film. It was observed that application of a constant potential or current dramatically improved PC-12 differentiation as assessed by an increase in both neurite numbers and lengths. These studies have also shown that tubular scaffolds fabricated from Ppy-PSS-Na can support the guided regeneration of transected sciatic nerve in rats.^[41] Furthermore, it has been shown that the soft-tissue biocompatibility of Ppy film doped with PSS-Na is far superior to that of most commonly used biodegradable polymers such poly(lactic-co-glycolic acid).^[40] The response to Ppy films is characterized by the total absence of a fibrous capsule even after 14 weeks (Fig. 3). Schmidt and co-workers have shown that vascularization around Ppy films can be enhanced by doping the film with polysaccharides.^[42] Ppy-PSS films have been shown to support the proliferation and osteogenic differentiation of mesenchymal progenitor cells (bone marrow stromal cells).^[43] It was observed that the production of alkaline phosphatase, a marker of osteogenic activity in these cells, was enhanced in cells cultured on Ppy-PSS substrates in comparison to those cultured on tissue culture plastic. These observations suggest that Ppy substrates have good cytocompatibility and should be useful as interactive coatings in tissue regeneration paradigms such as TE. One can also envisage the use of Ppy coatings to enhance cellular interactions towards metal implant surfaces.

However, Ppy synthesized from pyrrole or PNpy synthesized from N-methyl pyrrole-yields a highly cross-linked polymer that has low to no solubility in

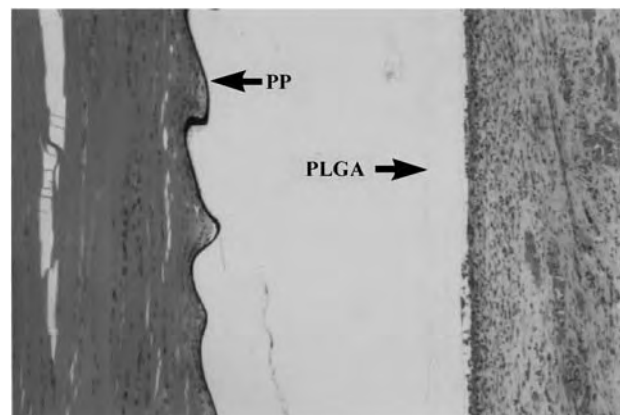


Fig. 3 Hematoxylin-Eosin-stained histological cross-section of Ppy-PSS (PP)-laminated with a biocompatible poly(L-lactic acid-co-glycolic acid) (PLGA) polymer implanted in a intramuscular site in rat for two weeks. Note the absence of fibrous tissue adjacent to the Ppy film. The gap between the Ppy and PLGA phase is an artifact created during processing of the tissue explant. Reproduced from Ref. [40].

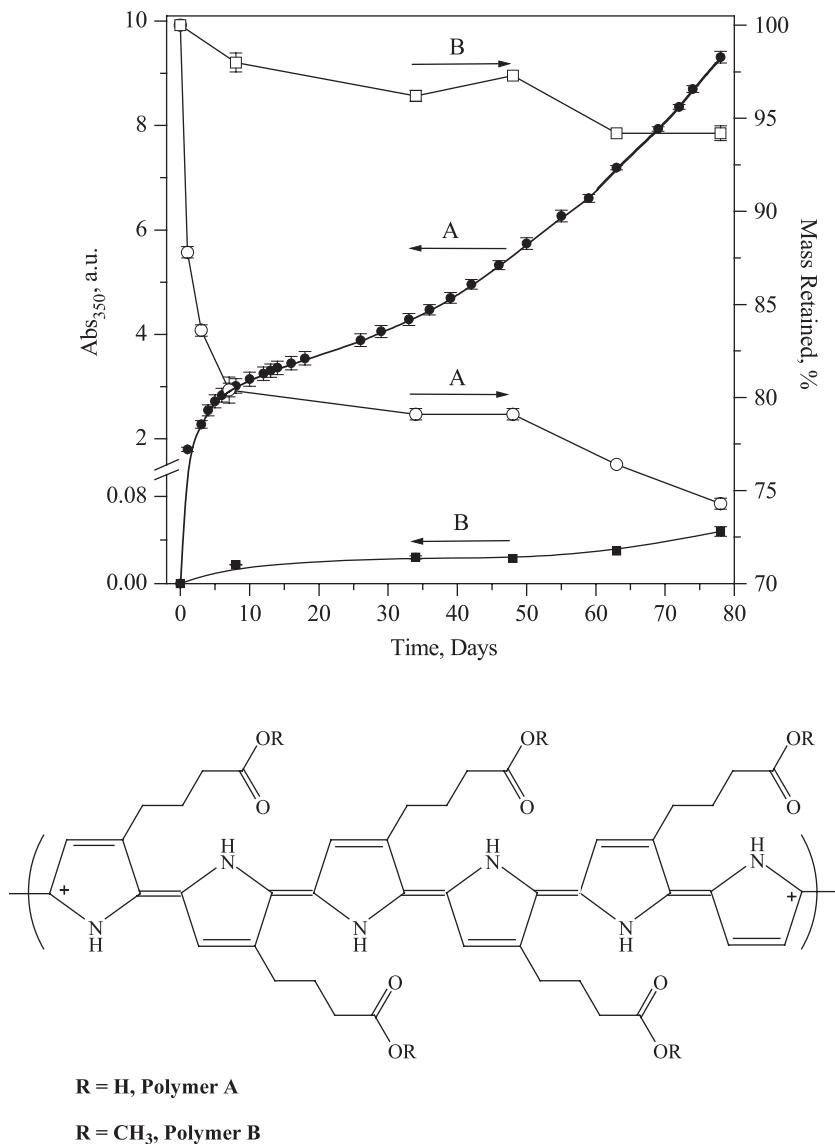


Fig. 4 Degradation characteristics of bioerodible Ppy pellets synthesized from pyrroles substituted in the β -position with ionizable or hydrolyzable side groups (polymer) as followed by dissolution (filled symbols) and mass loss (open symbols). From Ref. [44].

common organic solvents and hence not processible. The cross-linked structure also renders it nondegradable in a biological environment. Imparting degradability in a biological environment should significantly expand the use of Ppy in biomedical uses such as TE and drug delivery. Recently, Shastri and Langer reported the synthesis of conductive bioerodible Ppy from pyrroles substituted in the β -position with ionizable or hydrolyzable side groups.^[44] In this system, by controlling the ratio of the ionizable:hydrolyzable group, polymers with erosion times ranging from a few weeks to potentially over a year may be synthesized (Fig. 4). In addition to being bioerosible, Ppy possessing hydrolyzable groups show excellent solubility in common organic solvents such as tetrahydrofuran and methylene chloride, thus allowing for solution-based processing of these polymers into coatings and scaffolds for cell and tissue contacting applications.

These polymers possess excellent cytocompatibility and have been shown to support the proliferation and differentiation of human marrow-derived mesenchymal progenitor cells into osteoblast lineage.^[44] Another approach that has been explored to synthesize degradable Ppy is the coupling of 2,5-bis-(5-(3-hydroxypropoxycarbonyl)-2-pyrrolyl)thiophene with adipoyl chloride to yield a polymer with degradable ester linkages.^[45] Although this polymer has good tissue compatibility, it has extremely low conductivity which limits its potential use.

PIEZOELECTRIC POLYMERS

The piezoelectric effect was discovered by Jacques and Pierre Curie in the 1880s, when they observed that certain crystalline materials when subjected to tension

or compression became electrically polarized resulting in induced voltages of opposing polarity (compression tension).^[46] They also observed that the magnitude of this polarization was directly proportional to the applied stress. Conversely, minute changes in physical dimensions can be introduced in a piezoelectric material by subjecting it to an electric field. This phenomenon is called converse piezoelectricity. Piezoelectricity therefore may be defined as an electrical polarization induced in a solid in response to a mechanical stress (Figs. 5A and B). In order for a material to exhibit piezoelectric behavior, the crystalline domains (or crystal lattice) should be anisotropic (noncentrosymmetric) and polarizable. A piezoelectric material when heated above its Curie temperature loses its piezoelectric properties permanently. Most common examples of piezoelectric materials include naturally occurring substances such as quartz, polycrystalline ceramics such as barium titanate, lead zirconate titanate, and some inorganic salts (e.g., Rochelle's salt). Examples of synthetic and biological polymers that exhibit some form of piezoelectric behavior include poly(vinylidene difluoride) (PVDF), aromatic polyamides, polysulfones, polyparaxylene, poly(lactic acid), and some biopolymers such as collagen and poly(γ -methyl and γ -benzyl-L-glutamate).

In order to induce piezoelectric behavior in a ceramic or polymer, it first has to be poled in an external electric field. During the poling process, the dipoles in the crystalline regions align to form highly ordered domains resulting in a net dipole moment or polarization. The residual polarization after the electric field

is removed, termed remnant polarization, is responsible for the piezoelectric behavior in the material. The relationship between the induced polarization and applied stress and between the induced strain and applied electrical stimulus is described in terms of coupling constants that relate the two, namely:

d (C/N) = piezoelectric charge constant, the polarization generated by unit mechanical stress or the induced strain per unit electric field;

g (Vm/N) = piezoelectric voltage constant, the electric field generated per unit mechanical stress or the induced strain per applied unit displacement of charge;

C = current in coulombs; N = force in newtons; V = voltage in volts; and m = displacement in meters.

The d and g constants of a piezoelectric material are described using two subscripts, e.g., d_{33} and g_{31} , wherein the first subscript denotes the direction of the induced polarization and the second subscript denotes the direction of the applied stress. The coupling between the induced polarization and the direction of applied stress is described using an X,Y,Z system of notation (Fig. 5C), wherein the Z axis is the direction of the applied external poling field and represents a positive polarization; 1, 2, and 3 represent the stress along X, Y, and Z axes, respectively; and 4, 5, and 6 represent the shear stress about these axes. Therefore, in this system of notation, d_{33} means that the induced polarization is in the direction of the poling (Z axis)

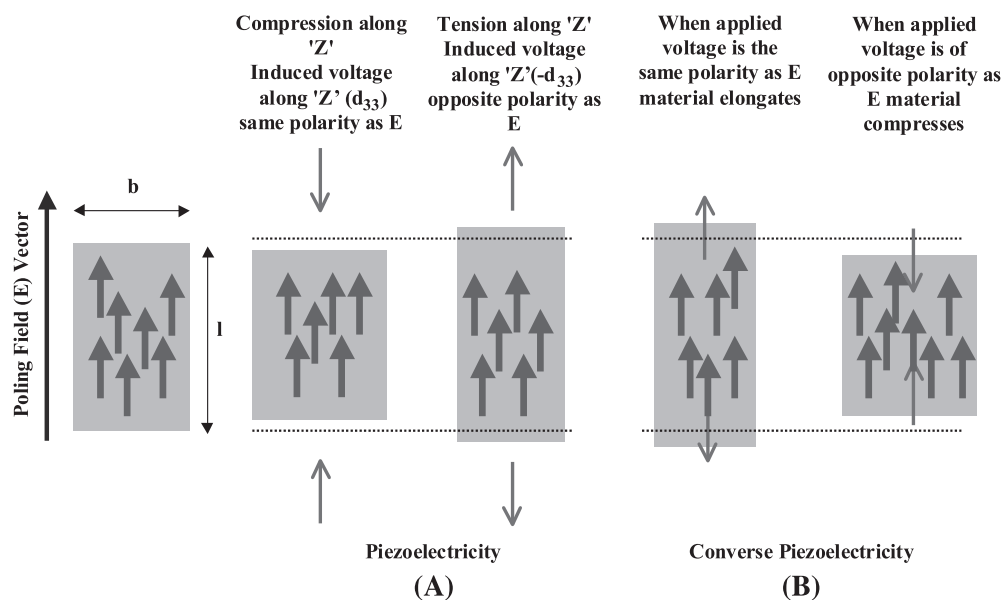


Fig. 5 (A) Schematic representation of the piezoelectric phenomenon in a material when subjected to compressive and tensile stress; (B) Schematic representation of reverse piezoelectric phenomenon in a material when subjected to an external field of the same and opposing polarity as the poling field; (C; see next page) Piezoelectric coupling tensors.

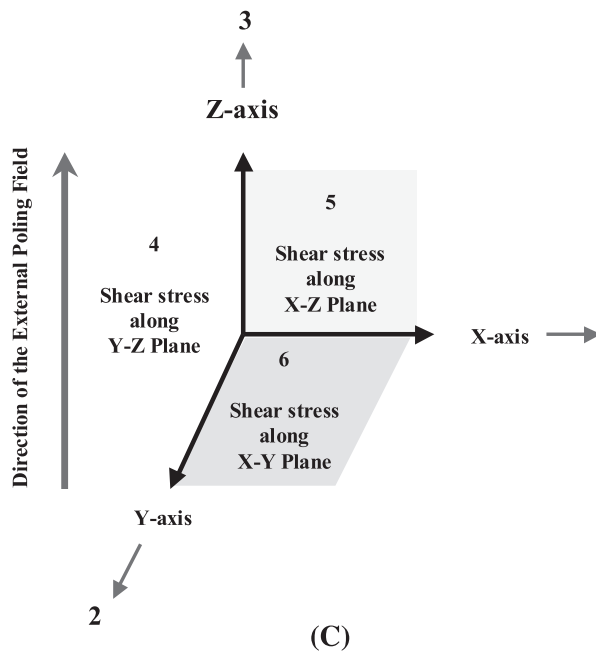


Fig. 5 (Continued)

(positive polarization) when the stress is applied in that same direction. Similarly, d_{25} means the induced polarization is along the Y axis when shear stress is applied along the X-Z plane.

Piezoelectricity in PVDF was discovered by Kawai in 1969.^[47] This discovery was followed by several reports showing that polypeptides and collagen exhibit piezoelectricity when subjected to shear stress.^[48–50] PVDF is a semicrystalline polymer composed of $-\text{CH}_2-\text{CF}_2-$ repeat units that are linked to each other predominantly in the *head-tail* fashion. The presence of the electronegative fluorine atom makes it responsible for the dipole moment of the CF_2 unit.^[51] The crystalline domains, which are responsible for long-range order in the polymer, are structurally anisotropic. The symmetry of the backbone dictates the overall dipole of the unit cell (crystalline domain). The crystalline domains in PVDF exist in four principal forms, called α , β , γ , and δ .^[52] Among these phases, only the β phase exhibits an all-*trans* conformation. This imparts a polar characteristic to this phase due to the alignment of the dipoles, something that is absent in the other phases due to a mixture of *trans*, *gauche*, and anti-*gauche* conformations. However, in the unpoled state, the dipoles cancel each other out due to random orientation of the crystalline domains resulting in a net dipole moment of zero. It has been observed that head-head $-\text{CH}_2-\text{CH}_2-$ and tail-tail $-\text{CF}_2-\text{CF}_2-$ linking of repeat units results in the formation of defects in the crystalline phase of PVDF, a process that favors the formation of the polar β phase.^[52] In fact, copolymers of PVDF with

tri-fluoroethylene (PVDF-TrFE) and tetrafluoroethylene (PVDF-TeFE) yield polymers with predominantly a β -phase backbone. The formation of β phase can also be induced by orienting the polymer chains through mechanical means such as drawing and stretching of fibers and films during extrusion. To impart persistent polarization in PVDF, the polymer is heated below its Curie temperature which is approximately 110°C and then subjected to an external electric field to align the dipoles in the crystalline domains. The polymer is then cooled in the presence of the field to freeze the orientation of the dipoles. The most likely mechanism for the alignment of dipoles based on the kinetic model proposed by Dvey-Aharon involves a 60° rotation of the polymer chain at the grain boundary resulting in the formation of a kink, which then propagates along the polymer chain and into the phase leading to long-range alignment within the phase.^[53] The piezoelectric charge constant, d , for PVDF depends on its configuration (fiber versus film). Films poled across its thickness (Z-axis) show a very strong d_{33} constant of about -30 pC/N and g_{33} constant of around -300 mV-m/N .

Collagen fibrils and other biopolymers exhibit piezoelectric properties when subjected to shear stress. The piezoelectricity in these polymers is thought to originate in the crystalline regions, which are formed when proteins and peptides form ordered structures such as α -helices and β -sheets. Piezoelectricity in these systems is thought to originate in the rotation of the amide bond. The straining of these ordered regions leads to rotation of the polar amide bonds in the backbone, which has a dipole moment of 3.4 debye, resulting in the generation of electrical currents. Biopolymers typically show d_{25} values of a few pC/N . Likewise, in the highly crystalline poly(L-lactic acid) a similar mechanism is presumed to be at play. The d_{25} values in highly oriented films of PLLA can be as large as -10 pC/N .

Among polymers, PVDF, poly(lactic acid), and poly(γ -methyl and γ -benzyl L-glutamate) are the most promising with respect to cell and tissue contacting applications. As discussed earlier, it has been shown that mechanical loading of bone leads to stimulation of bone remodeling and growth.^[5,6] The interplay between mechanical stress and osteoclastic (bone resorption) and osteoblastic (bone formation) activity occurs via the generation of localized electrical charges that in turn trigger cellular activity.^[7] The localized electrical activity is generated by the mechanical deformation of the piezoelectric microcrystalline inorganic hydroxyapatite phase. In addition, collagen fibrils, which are known to exhibit shear piezoelectricity, may play a role as well. The electrical charges generated upon compression and tension-loading in bone are of opposing polarity, which

suggests an intricate coupling between the bone resorption and formation processes through the microcrystalline hydroxyapatite-collagen composite matrix. Orthopaedic metal implants coated with various forms of hydroxyapatite have been shown to enhance new bone formation and integration of the implant.^[54] Whether the observed effect is due to the piezoelectric nature of the coating or due to the calcium-rich surface is still unclear. Similarly, in wound healing, electrical activity at the injury site due to the depolarization of cell membranes is thought to play a role in the migration of cells to the site. Several studies have shown that electromagnetic stimulation can aid in bone and wound healing.^[55–57] Aebischer and coworkers have shown that tubular scaffolds derived from poled PVDF are capable of supporting the regeneration of transected sciatic nerve in rat.^[58,59] They observed that the quality of the regenerated sciatic nerve using poled-PVDF guidance channels was superior to that of controls. They have attributed this outcome to localized electrical stimulation of the regenerating neuronal tissue mass by electrical currents generated as a consequence of minute mechanical deformations of the PVDF conduits. So it is entirely conceivable that, in the near future, scaffolds either composed entirely of or coated with piezoelectric polymers will find uses in wound healing and tissue regenerative applications. One can also envision using these polymers as barrier membranes or microvalves in smart drug delivery systems or artificial pores wherein permeability of the drug across the membrane or through a valve is triggered by an external electric field.

ELECTRO-RESPONSIVE GELS

A hydrogel is a cross-linked network of polymer and imbibed water. When the polymer phase is composed of a polyelectrolyte, the gel is called an ionic gel. If the polyelectrolyte is composed of a weakly ionizable group such as a carboxylic acid, the gel will undergo contraction at pHs below the isoelectric point (pI) of the acid moiety and expansion at pH's above the pI. This phenomenon is due to the ionization of the acid groups at higher pHs that leads to electrostatic repulsion within the gel resulting in an expansion in gel volume.^[60,61]

Pioneering work carried out by Toyochi Tanaka and later by Osada and coworkers have shown that when an ionic gel, irrespective of the nature of the ionizable group, is placed between two electrodes and subjected to a DC field, it undergoes anisotropic contraction with the concomitant expulsion of water. Upon removal of the electric field, the contracted gel expands by taking up water and returning to its resting volume. This results in the conversion of electrical

energy into mechanical strain.^[62,63] The basis for this deformation lies in the electrophoretic redistribution of ions within the cross-linked network. For example, in a cross-linked poly(acrylic acid) gel, because the carboxylic acid groups (negative charges) are fixed to the polymer backbone, upon the application of an electric field the positively charged counterions being mobile migrate toward the cathode. This migration of counterions is accompanied by expulsion of water resulting in shrinkage of the gel at the anode.^[64,65] The converse is true for a positively charged cross-linked network. Interestingly, when a weak polyelectrolyte gel such as poly(acrylic acid) gel is subjected to a mechanical stress, the pH in the local environment changes in response to this stress.^[66] This is thought to occur due to the elongation of polymer chains in response to the stress, resulting in an increase in free energy, which is compensated by ionization of the acid groups to yield protons. It has also been shown that, by a process akin to piezoelectric behavior, mechanical deformation of a weak polyelectrolyte gel can induce an electrical potential. Osada and co-workers have explored this phenomenon to develop tactile sensors that may find uses in robotic devices in remote surgery applications.^[66]

When a swollen cross-linked polyelectrolyte gel is exposed to a surfactant bearing opposite charge, the charges in the network are quenched leading to deswelling of the network. In the presence of an electric field, the de-swelling behavior can yield anisotropic stress within a gel film or fiber, resulting in bending of the gel toward one of the electrodes with concomitant mechanical motion. The gel-bending phenomenon has been leveraged to make a hydrogel-based device called the gel looper that is capable of moving in an electric field.^[67] Such smart polymer gels are bound to find applications in drug delivery and tissue engineering as these systems get more integrated with microelectromechanical system (MEMS)-based devices.

OTHER SYSTEMS

Matrices composed of a complex (salt) of polyelectrolytes of opposing charges have been explored as electroerodible systems for the release of proteins.^[68] The drug delivery matrix which in essence is a polymer complex can be made to dissociate upon exposure to an electrical potential in aqueous media. The dissociation originates at the face, opposing the counter electrode, and proceeds inward towards the working electrode. Although this system has its limitations with respect to ease of use and thickness of the delivery matrix, truly responsive drug delivery can nevertheless be achieved. Magnetite is an oxide of iron with a chemical composition of Fe_3O_4 wherein iron exists in

both +2 and +3 oxidation states ($\text{Fe}_3\text{O}_4 = \text{FeO} + \text{Fe}_2\text{O}_3$). Due to the mixed valence state of iron, magnetite has unpaired electrons, hence a net dipole moment, making it ferromagnetic. Ferrofluid is a dispersion of magnetite colloids (nanoparticles) in a hydrophobic medium such as silicone oil. When a ferrofluid is exposed to an external magnetic field, the magnetite nanoparticles become magnetized and align along the magnetic lines of force. This results in a physical transformation of the suspension from a liquid to a solid. Upon removal of the field the solidification is reversed. Ferrofluids can also be made from cobalt, zinc, and manganese oxides. However, magnetite is the most promising with respect to biological applications. One can envisage many applications for ferrofluids in medicine. Some obvious ones are localization media for drug delivery, plugs for coronary embolism, electroresponsive valves in implantable drug delivery systems, and implantable muscle actuators. However, as currently available, ferrofluids are not biocompatible; hence methodologies to improve biocompatibility of magnetite need to be worked out and alternative media for dispersion and delivery of the nanoparticles need to be developed. Iron oxide nanoparticles are already used in cell sorting and are being explored in cell and tissue imaging. Nevertheless, efforts are already underway to explore this system in localization of chemotherapy.^[69] Other *ex vivo* applications may include carriers for cell sorting, valves in bio-MEMS devices for high throughput applications, cuffs to prevent vascular embolism, and a battlefield tourniquet. As we gain a better understanding of biomolecular processes, the newly acquired knowledge has enabled us manipulate materials using the very tools of nature.^[70] The advent of biomolecular materials engineering as a science and discipline holds immense promise for the future, wherein medical devices are fully integrated with the human body via smart materials and interfaces that can respond to external stimuli such as an electrical signal to deliver drugs on demand^[71] or to change a fate of a cell^[37,40] or an implant surface^[72] to improve clinical outcomes.

ARTICLES OF FURTHER INTEREST

Artificial Muscles; Biologic and Synthetic Apatites; Conductive Polymers; Piezoelectric Materials

REFERENCES

1. Becker, R.O.; Marino, A.A. *Electromagnetism and Life*; State University of New York Press: Albany, NY, 1982.
2. Guyton, E.C. *Text Book of Medical Physiology*, 8th Ed.; W.B. Saunders Company, 1991. Chp. 7.
3. Wolff, J. *Das gesetz der transformation der knochen*; Hirschwald, A., Ed.; Berlin, 1891.
4. Forwood, M.R.; Turner, C.H. Skeletal adaptations to mechanical usage. *Bone* **1995**, *17*, 197s–205s.
5. Yasuda, I. On the piezoelectric activity of bone. *J. Jap. Orthop. Surg. Soc.* **1954**, *28*, 267.
6. Fukada, E.; Yasuda, I. On the piezoelectric effect in bone. *J. Physiol. Soc. Jpn.* **1957**, *12*, 1198.
7. Bassett, C.; Becker, R.O. Generation of electric potentials by bone in response to mechanical stress. *Science* **1962**, *137*, 1063.
8. Chiang, C.K.; Drury, M.A.; Gau, S.C.; Heeger, A.J.; Louis, E.J.; MacDiarmid, A.G.; Park, Y.W.; Shirakawa, H. Synthesis of highly conducting films of derivatives of polyacetylene, (CH)_x. *J. Am. Chem. Soc.* **1978**, *100*, 1013.
9. Kaner, R.B.; MacDiarmid, A.G. Plastics that conduct electricity. *Sci. Am.* **1988**, 106.
10. Skotheim, T.A. *Handbook of Conducting Polymers*; Marcel Dekker: New York, 1986; Vol. 1.
11. Diaz, A.F.; Kanazawa, K.K.; Garidini, G.P. Electrochemical Polymerization of Pyrrole. *J. Chem. Soc. Chem. Commun.* **1979**, 635.
12. Tourillon, G.; Garnier, F. J. Electrochemical synthesis of polythiophenes. *Electroanal. Chem.* **1982**, *135*, 173.
13. Otero, T.F.; Villanueva, S.; Bengoechea, M.; Brillas, E.; Carrasco, J. Reversible redox switching in polypyrrole and polySNS films. *Synth. Met.* **1997**, *85*, 183.
14. Nalwa, H.R. *Handbook of Organic Conductive Molecules and Polymers*; John Wiley & Sons, 1997; volume 2. General conductive polymer ref.
15. Cosnier, S.; Innocent, C.; Jouanneau, Y. Amperometric detection of nitrate via nitrate reductase immobilized and electrically wired at the electrode surface. *Anal. Chem.* **1994**, *66*, 3198.
16. Trojanowicz, M.; Lewenstam, A.; Krawczynski, T.; Lahdesmaki, I.; Szczepek, W. Flow injection amperometric detection of ammonia using a polypyrrole-modified electrode and its applications in urea and creatinine biosensors. *Electroanalysis* **1996**, *8*, 233.
17. Garnier, F.; Youssauffi, H.; Srivastava, P.; Yassar, A. Enzyme recognition by polypyrrole functionalized with bioactive peptides. *J. Am. Chem. Soc.* **1994**, *116*, 8813.
18. Foulds, N.; Lowe, C. Enzyme entrapment in electrically conducting polymers: Immobilization of glucose oxidase in polypyrrole and its application in amperometric glucose sensor. *J. Chem. Soc.* **1986**, *82*, 1259.
19. Bartlett, P.; Whitaker, R. Electrochemical immobilization of enzymes. Part I. Theory. *J. Electroanal. Chem.* **1987**, *224*, 27.
20. Bartlett, P.; Whitaker, R. Electrochemical Immobilization of enzymes. Part II. Glucose Oxidase immobilized in poly(N-methylpyrrole). *J. Electroanal. Chem.* **1987**, *224*, 37.
21. Fortier, G.; Brassard, E.; Belanger, D. Optimization of polypyrrole glucose oxidase biosensor. *Biosens. Bioelectron.* **1990**, *5*, 473.
22. Pandey, P. A new conducting polymer-coated glucose sensor. *J. Chem. Soc. Faraday Trans. 1.* **1988**, *84*, 2259.
23. Caglar, P.; Wnek, G.E. Glucose-sensitive polypyrrole/poly(styrenesulfonate) films containing co-immobilized

- glucose oxidase and (ferrocenylmethyl) trimethylammonium bromide. *J. Macromol. Sci., Pure Appl. Chem.* **1995**, *A32*, 349.
24. Schuhmann, W.; Ohara, T.; Heller, A.; Schmidt, H.-L. Electron transfer between glucose oxidase and electrodes via redox mediators bound with flexible chains to the enzyme surface. *J. Am. Chem. Soc.* **1991**, *113*, 1394.
 25. Foulds, N.; Lowe, C. Immobilization of glucose oxidase in ferrocene modified pyrrole polymers. *Anal. Chem.* **1988**, *60*, 2473.
 26. Willner, I.; Willner, B. Electrical communication of redox proteins by means of electron relay-tethered polymers in photochemical, electrochemical and photoelectrochemical systems. *React. Polym.* **1994**, *22*, 267.
 27. Khan, G.; Ohwa, M.; Wernet, W. Design of stable charge transfer complex electrode for a third generation amperometric glucose sensor. *Anal. Chem.* **1996**, *68*, 2939.
 28. Schuhmann, W.; Huber, J.; Kranz, C.; Wohlsejager, H. Conducting polymer based amperometric enzyme electrode: Towards the development of miniaturized reagentless biosensors. *Synth. Met.* **1994**, *61*, 31.
 29. Nishizawa, M.; Matsue, T.; Uchida, I. Penciliin sensor based on a microarray electrode coated with pH responsive polypyrrole. *Anal. Chem.* **1992**, *64*, 2462.
 30. Livache, T.; Roget, A.; Dejean, E.; Barthet, C.; Bidan, G.; Téoule, R. Preparation of a DNA matrix via an electrochemically directed polymerization of pyrrole and oligonucleotides bearing a pyrrole group. *Nucleic Acids Res.* **1994**, *22*, 2915.
 31. Livache, T.; Roget, A.; Dejean, E.; Barthet, C.; Bidan, G.; Téoule, R. Biosensing effects in functionalized electroconducting conjugated polymer layers: Addressable DNA matrix for the detection of gene mutations. *Synth. Met.* **1995**, *71*, 2143.
 32. Yamato, H.; Ohwa, M.; Wernet, W. Stability of polypyrrole and poly(3,4-ethyldioxythiophene) for biosensor application. *J. Electroanal. Chem.* **1995**, *397*, 163.
 33. Parente, A.; Marques, E.; Azevedo, W.; Diniz, F.; Melo, E.; Lime Filho, J. Glucose biosensor using glucose oxidase immobilized in polyaniline. *Appl. Biochem. Biotechnol.* **1992**, *37*, 267.
 34. Miller, L.L.; Zhou, Q.-X. Poly(N-methylpyrrolylium) poly(styrenesulfonate). A conductive, electrically switchable cation exchanger that cathodically binds and anodically releases dopamine. *Macromolecules* **1987**, *20*, 1594.
 35. Zhou, Q.-X.; Miller, L.L.; Valentine, J.R. Electrochemically controlled binding and release of protonated dimethyldopamine and other cations from poly(N-methyl pyrrole)/polyanion composite redox polymers. *J. Electroanal. Chem.* **1989**, *261*, 147.
 36. Prezyna, L.A.; Qiu, Y.-J.; Reynolds, J.R.; Wnek, G.E. Interaction of cationic polypeptides with electroactive polypyrrole/poly(styrenesulfonate) and poly(N-methylpyrrole)/poly(styrenesulfonate) films. *Macromolecules* **1991**, *24*, 2224.
 37. Wong, J.; Ingber, D.E.; Langer, R. Electrically conductive polymers can noninvasively control cell shape. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 3201.
 38. Shastri, V.R.; Wnek, G.E. Effect of dopant ion on cell growth on polypyrrole thin films. *Polym. Prepr.* **1993**, *34* (2), 70.
 39. Shastri, V.R. *Evaluation of Polypyrrole "Thin Films" as Substratum for Mammalian Cell Culture, Doctoral Dissertation*; Rensselaer Polytechnic Institute: Troy, NY, 1995.
 40. Schmidt, C.E.; Shastri, V.R.; Vacanti, J.P.; Langer, R. Stimulation of neurite outgrowth using electrically conducting polymer. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 8948.
 41. Shastri, V.R.; Schmidt, C.E.; Kim, T.-H.; Vacanti, J.P.; Langer, R. Polypyrrole-A potential candidate for stimulated nerve regeneration. *Mat. Res. Soc. Symp. Proc.* **1996**, *414*, 113.
 42. Collier, J.H.; Camp, J.P.; Hudson, T.W.; Schmidt, C.E. Synthesis and characterization of polypyrrole-hyaluronic acid composite biomaterials for tissue engineering applications. *J. Biomed. Mat. Res.* **2000**, *50*, 574.
 43. Shastri, V.P.; Rahman, N.; Martin, I.; Langer, R. Applications of conductive polymers in bone regeneration. *Mat. Res. Soc. Symp. Proc.* **1999**, *550*, 215.
 44. Zelikin, A.; Lynn, D.M.; Farhadi, J.; Martin, I.; Shastri, V.P.; Langer, R. Erodible conducting polymers for potential biomedical applications. *Angew. Chem. Int. Ed.* **2002**, *41* (1), 141.
 45. Rivers, T.J.; Hudson, T.W.; Schmidt, C.E. Synthesis of a novel, biodegradable electrically conducting polymer for biomedical applications. *Adv. Funct. Mat.* **2002**, *12* (1), 33.
 46. Lines, M.E.; Glass, A.M. *Principles and Applications of Ferroelectric Related Materials*; Clarendon Press: Oxford, 1977.
 47. Kawai, H. The piezoelectricity of poly(vinylidene-fluoride). *Jpn. J. Appl. Phys.* **1969**, *8*, 975.
 48. Shamos, M.H.; Lavine, L.S. Piezoelectricity as a fundamental property of biological tissue. *Nature* **1967**, *213*, 267.
 49. Fakuda, E. Piezoelectricity of biopolymers. *Biorheology* **1995**, *32* (6), 593.
 50. Konikoff, J.J. Origins of osseous bioelectric potentials. A review. *Ann. Clin. Lab. Sci.* **1975**, *5* (5), 330.
 51. Kepler, R.G.; Anderson, R.A. Ferroelectric polymers. *Adv. Phys.* **1992**, *41*, 1.
 52. Lovinger, A.J. Poly(Vinylidene fluoride), in *Developments in Crystalline Polymers*. In *Applied Science*; Bassett, D.C., Ed.; Kluwer Academic Publishers, 1982; 195.
 53. Dvey-Aharon, H.; Sluckin, T.J.; Taylor, P.L. Kink propagation as a model for poling in poly(vinylidene-fluoride). *Phys. Rev., B.* **1980**, *21*, 3700.
 54. Soballe, K.; Overgaard, S.; Hansen, E.S.; Brokstedt-Rasmussen, H.; Lind, M.; Bunger, C.J. A review of ceramic coatings for implant fixation. *J. Long-Term Eff. Med. Implants* **1999**, *9* (1-2), 131.
 55. Wahlstrom, O. Stimulation of fracture healing with electromagnetic fields of extremely low frequency. *Clin. Orthop.* **1984**, *186*, 293.
 56. Weiss, D.S.; Kirsner, R.; Eaglstein, W.H. Electrical stimulation and wound healing. *Arch. Dermatol.* **1990**, *126*, 222.

57. Alvarez, O.M.; Mertz, P.M.; Smerbeck, R.V.; Eaglstein, W.H. The healing of superficial skin wounds is stimulated by external electric fields. *J. Invest. Dermatol.* **1983**, *81*, 144.
58. Aebischer, P.; Valentini, R.; Dario, P.; Domencini, C.; Galletti, P. Piezoelectric guidance channels enhance regeneration in the mouse sciatic nerve after axotomy. *Brain Res.* **1987**, *436*, 165.
59. Valentini, R.; Vargo, T.; Gardella, J.; Aebischer, P. Electrically charged polymeric substrates enhance nerve fiber outgrowth in vitro. *Biomaterials* **1992**, *13*, 183.
60. Tanaka, T. Collapse of gels and the critical endpoint. *Phys. Rev. Lett.* **1978**, *40*, 820.
61. Tanaka, T.; Fillmore, D.J. Kinetics of swelling of gels. *J. Chem. Phys.* **1979**, *70*, 1214.
62. Tanaka, T.; Nishio, I.; Sun, S-T.; Ueno-Nishio, S. Collapse of gels under an electric field. *Science* **1982**, *218*, 467.
63. Osada, Y.; Ross-Murphy, S. Intelligent gels. *Sci. Am.* **1993**, *268*, 82.
64. Osada, Y.; Hasebe, M. Electrically activated mechanochemical devices using polyelectrolyte gels. *Chem. Lett.* **1985**, *1285-1288*, 1985.
65. Gong, J.P.; Nitta, T.; Osada, Y. Electrokinetic modeling of the contractile Phenomena of polyelectrolyte gels-one dimensional capillary model. *J. Phys. Chem.* **1994**, *98*, 9583.
66. Osada, Y.; Gong, J.P.; Sawahata, K.J. Synthesis, mechanism, and application of electro-driven chemomechanical system using polymer gels. *J. Macromol. Sci., Chem.* **1991**, *A28*, 1189.
67. Osada, Y.; Okuzaki, H.; Hori, H. A polymer gel with electrically driven motility. *Nature* **1992**, *355*, 242.
68. Kwon, I.C.; Bae, Y.H.; Kim, S.W. Electrically erodible polymer gel for controlled release of drugs. *Nature* **1991**, *354*, 291.
69. Alexiou, C.; Arnold, W.; Klein, R.J.; Parak, F.G.; Hulin, P.; Bergemann, C.; Erhardt, W.; Wagenpfeil, S.; Lübke, A.S. Locoregional cancer treatment with magnetic drug targeting. *Cancer Res.* **2000**, *60*, 6641.
70. Lee, S.W.; Mao, C.; Flynn, C.E.; Belcher, A.M. Ordering of quantum dots using genetically engineered viruses. *Science* **2002**, *296*, 892.
71. Santini, J.T.; Cima, M.J.; Langer, R. A controlled-release microchip. *Nature* **1999**, *397*, 335.
72. Lahann, J.; Mitragotri, S.; Tran, T.-H.; Kaido, H.; Sundaram, J.; Choi, I.S.; Hoffer, S.; Somorjai, G.; Langer, R. A reversibly switching surface. *Science* **2003**, *299*, 371.

Electrocardiography

Soumyadipta Acharya
Bruce Taylor

Department of Biomedical Engineering, The University of Akron, Akron, Ohio, U.S.A.

INTRODUCTION

The electrocardiogram (ECG or EKG) is a recording of the electrical activity of the heart, as manifested at the surface of the body. This simple noninvasive biomedical signal is a tool of tremendous clinical power, and when used in the proper clinical context, can provide deep insights into the health status of an individual. The material included in this article is intended as a review of some clinical and engineering aspects of the ECG and should not be used as a clinical text.

HISTORICAL PERSPECTIVE OF MODERN ELECTROCARDIOGRAPHY

The discovery that the mechanical activity of the heart was accompanied by an electrical signal was first reported by Kolliker and Mueller in 1856 and Donders in 1872.^[1] These researchers discovered that a frog gastrocnemius muscle and its attached sciatic nerve, when laid over a beating frog heart, would twitch in synchrony with the beating heart. This indicated that some sort of galvanic current was initiated in the frog heart and was sufficient to stimulate the sciatic nerve preparation. The first measurements of the time course of the electrical activity were performed by Burdon-Sanderson in 1879^[2] using a current-sampling “rheotome” and a slowly responding galvanometer. The first recordings of the electrical activity were performed by physiologists Marey and Lippmann in 1876^[1] using a device called a capillary electrometer.

Although he most certainly borrowed some of the ideas from his predecessors, Willem Einthoven is generally accepted as the father of electrocardiography, a distinction for which he received the Nobel Prize in 1924. In his physiology laboratory at the University of Leiden, Einthoven developed the string galvanometer, which permitted the first accurate representations of the ECG. Figure 1 is a reconstruction of Einthoven’s laboratory showing the means by which he was able to obtain sufficiently low electrode resistances to permit an accurate presentation of the electrical activity of the heart.

The string galvanometer consisted of a fine quartz fiber, coated with silver and placed in a strong electromagnetic field. A very small current traversing the silvered quartz filament would cause the filament to deflect slightly. The shadow of the deflecting filament was then projected onto a moving photographic surface, thereby producing the familiar ECG.

One problem of early galvanometers, for displaying the ECG, was the requirement that an electric current drive the measuring device. The source for this current was the parameter being measured, which for the beating heart is in the range of 100 mV. Figure 2 shows a simple electric circuit model of how cardiac electrical activity is measured. The voltage at the indicator, V_i , modified from Ohm’s law, is shown in Fig. 2.

The measured voltage as a function of the internal body resistance and the input resistance of the indicator $R_{indicator}$ is given by the following equation.

$$V_i = \frac{V_s R_{indicator}}{R_{body} + R_{indicator}}$$

From this equation, it can be seen that the indicator voltage V_i decreases as the indicator resistance $R_{indicator}$ decreases. In other words, the lower the input resistance of the indicator, the lower the measured voltage. Conversely, the higher the input resistance, the greater the percentage of V_s that can be detected by the indicator. This is an excellent example of insertion error, where the event we are trying to measure is disturbed by our method of measurement. The best way around this problem is to increase the input resistance $R_{indicator}$ of the indicator to the point where $R_{indicator}$ approaches infinity. In that case, V_i will equal V_s .

The development of the vacuum tube triode by Lee De Forest in 1906 and the efforts by General Electric to incorporate the vacuum tube into the electrocardiograph in 1921 provided the means by which the input resistance of the indicator could be increased to very high levels. This then initiated the modern era of practical electrocardiography that continues to this day. The main difference is that the vacuum tube amplifier has been replaced by integrated circuit amplifiers that are smaller, require less power to operate,



Fig. 1 Early measurements of the electrocardiogram required a large electrode surface area that could be achieved with the use of saline-filled buckets into which the appropriate limbs were immersed.

and have higher input resistance. A basic circuit model of a modern indicator method is shown in Fig. 3.

Once high-input-resistance amplifiers came into use in electrocardiography, the large-surface-area electrode requirement of Einthoven's string galvanometer was eliminated and smaller electrodes could be used. Today's electrodes are approximately 1 cm in diameter and are generally disposable.

AN OVERVIEW OF CARDIAC CELLULAR ELECTROPHYSIOLOGY

The cardiac cells, like all other cells in the body, maintain a negative resting membrane potential with respect to their exterior. This electrical polarization is maintained by a number of ionic pumps on the cellular

membrane, which pump the appropriate ions (Na^+ , K^+ , Ca^{++} , Cl^-) in and out in order to maintain the electronegativity of the interior.^[3] Slight perturbations in the resting state can cause a complete reversal of this negativity, due to the opening of many ionic channels that allow the charged ions to move along their gradient. Depolarization of one cell acts as a perturbation for its immediate neighbors, causing them to depolarize also. This process spreads from cell to cell, resulting in a depolarization wave at the macroscopic level. The open ionic channels close down in a few milliseconds, resulting in a cessation of the transient ionic flow along their gradient and causing the membrane potential to return to its resting state. This process is known as repolarization. Depolarization followed by repolarization is the fundamental electrical activity of the heart recorded in the electrocardiogram. However, the question remains: What initiates this depolarization?

The heart consists of localized areas of a special type of cells known as pacemaker cells. Unlike most other cells in the body, these have the property of autorhythmicity; in other words, they spontaneously depolarize at fixed intervals of time. This unique process occurs due to the presence of special types of "leaky" ionic channels on the membranes of these cells. These channels cause the membrane to gradually lose its electronegativity until the membrane potential reaches the threshold necessary for initiation of an action potential. Each spontaneous depolarization of these pacemaker cells acts as a source for initiating one complete cardiac cycle.

A recording of this wave of depolarization and repolarization using electrodes placed on the exterior of the cell looks quite different from the action potential as recorded by intracellular electrodes. Figure 4 shows a step by step recording of a wave of depolarization followed by repolarization. Note that there is no voltage difference between the two electrodes when the membrane is either fully depolarized or fully repolarized. The ECG is actually a similar recording of the depolarization and repolarization waves of the heart as measured at the surface of the body.

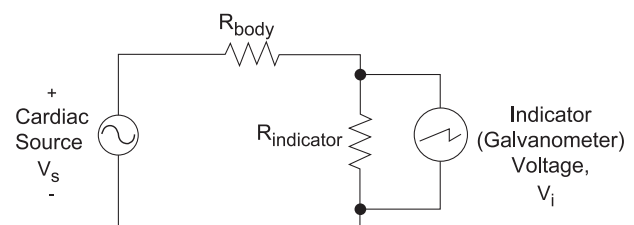


Fig. 2 A simple electrical circuit model of the heart (cardiac source), internal body resistance, and indicator.

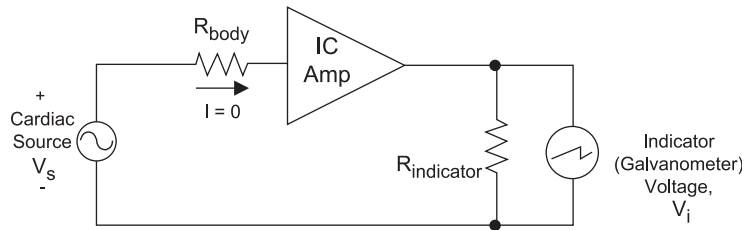


Fig. 3 A basic circuit diagram of a voltage indicator, coupled with an integrated circuit (IC) amplifier that has a very high input resistance and a corresponding input current requirement approaching zero.

THE COMPONENTS OF THE NORMAL ECG

The heart is composed primarily of two types of tissues: the contractile myocardial muscles, which allow a wave of depolarization to pass slowly through them, and the specialized conducting and pacemaker tissues, which allow a rapid conduction of the depolarization wave.^[4,5] Figure 5 shows a cross-section of the heart, illustrating the specialized conducting pathways. The depolarization activity is rapidly propagated along these specialized pathways to the adjoining myocardium.

The ECG is a measurement of the mean, instantaneous, electrical vector of the heart resulting from electrical activity traversing the myocardium. Figure 6 depicts a typical ECG showing the P, Q, R, S, and T waves and some standard durations and segments.

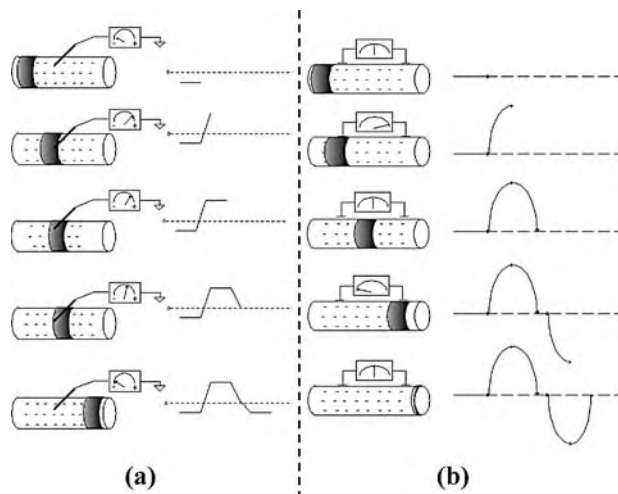


Fig. 4 The appearance of the action potential depends on how it is measured. Part (a) shows the result of placing a microelectrode inside (intracellular) the active cell. In this instance, the depolarization wave—a section of intracellular positive charge (shaded area)—passes the recording electrode, causing a monophasic action potential. A biphasic action potential (b) is produced if the recording electrodes are placed on or near the surface of the active cells.

Atrial Depolarization and the P Wave

The sino-atrial (SA) node, located on the right atrium, acts as the normal pacemaker of the heart.^[4] Though the normal heart rate is 72 beats per minute, the rate at which the SA node fires can vary greatly depending upon modulation by the autonomic nervous system. Each spontaneous firing of the SA node results in a wave of depolarization spreading out across the atria. This depolarization is preferentially conducted via the three specialized atrial tracts. This event is reflected in the ECG as the P wave. Normally, the P wave can have varied shapes, from rounded to sharply peaked. Normal amplitude of the P wave varies from 0 to 0.3 mV and the normal duration is up to 0.1 seconds.^[6]

Atrio-Ventricular Delay

The wave of depolarization can normally move from the atrium to the ventricle only through the atrio-ventricular (AV) node, located at the base of the interventricular septum between the atria and ventricles.^[4] The speed of conduction through the AV node is very slow, resulting in an AV delay. This allows the atria to fully contract and empty the contents into the ventricles before the onset of ventricular contraction. Like the SA node, the AV node is modulated by the autonomic nervous system, resulting in variations in the delay. The P-R interval in the ECG is the total time between the onset of atrial depolarization to the onset of ventricular depolarization. The normal P R interval can vary between 0.12 and 0.2 seconds.^[6]

Ventricular Depolarization and the QRS Complex

After the delay at the AV node, the depolarization wave passes to the AV bundle (also known as the Bundle of His) and from there along the septum to the left and right bundle branches. The left bundle branch has two main divisions, the anterior and the posterior fascicles. Eventually, after successive branching, they terminate on the network of Purkinje fibers, which connect to the ventricular myocardium.^[4,5] Due to the

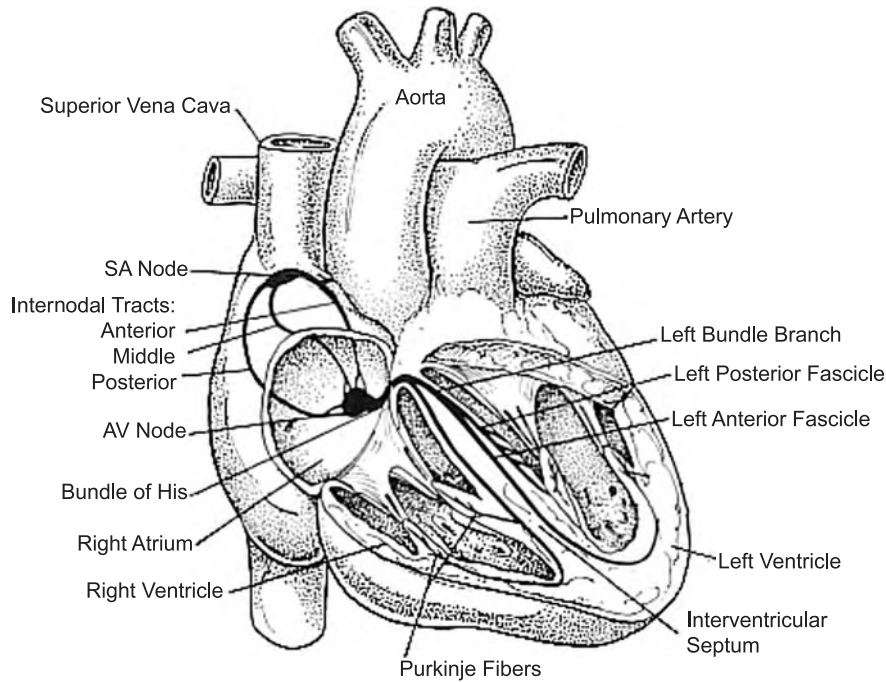


Fig. 5 Depiction of the conduction pathways of the heart. (Reprinted with permission, Biophysical Measurement Series: Electrocardiography by Charles A. Rawlings. SpaceLabs, Inc. 1991.)

presence of this elaborate conducting pathway, the depolarization wave passes very rapidly through the ventricles and is reflected in the ECG as the QRS complex. The Q wave is defined as the first downward deflection and is normally absent in most leads of the ECG. The R wave is the first positive deflection (with or without the Q wave) and the S wave is the negative deflection immediately following the R wave. Both the R and the S wave can vary greatly in amplitude, normally between 0 and 2.8 mV (R wave) and 0 to 0.8 mV (S wave).^[6] The normal QRS duration is between 0.05 and 0.1 seconds.^[6] Occasionally, in some conditions, there is a second upward deflection known as the R' (R prime) wave and a second downward deflection after the R' wave, known as the S' (S prime) wave.

Ventricular Repolarization and the T Wave

The T wave is a manifestation of ventricular repolarization. In most leads, the T wave is an upward deflection in the range of 0 to 0.8 mV with a duration of 0.1 to 0.25 seconds.^[6]

Atrial Repolarization

The atrial repolarization wave is normally not seen in the ECG because it coincides with and is obscured by the electrically dominant QRS complex.

The S-T Segment

This is a flat segment from the end of the QRS complex to the beginning of the T wave. Normally this segment is isoelectric with the P-R segment. Baseline elevation or depression of this segment with respect to the P-R segment can be manifestations of cardiac disease (discussed later).

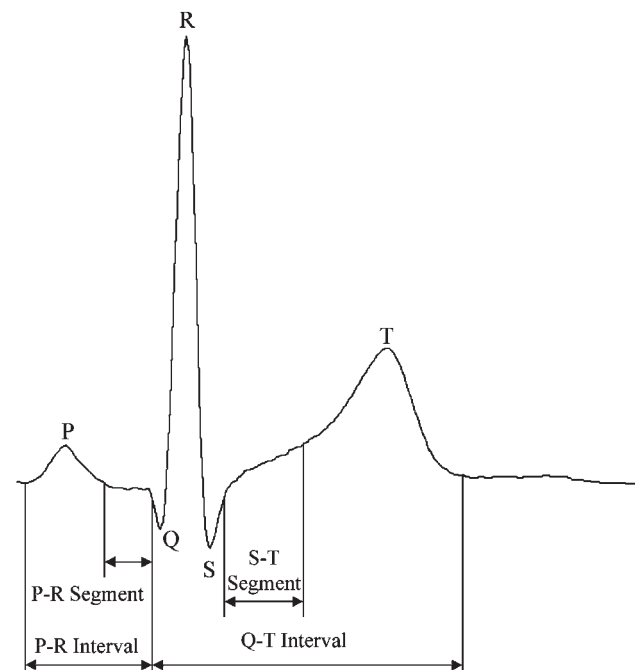


Fig. 6 One cycle of a typical ECG showing P, Q, R, S, and T waves with segments and intervals.

The U Wave

The U wave is a rare occurrence in some ECGs immediately after the T wave, with a similar shape and size but an amplitude that is 5 to 20% of the T wave. It is believed to be caused by the repolarization of the papillary muscles.^[4]

RELATIONSHIP BETWEEN THE ECG AND THE MECHANICAL ACTIVITY OF THE HEART

Depolarization of the myocardial muscle membrane is immediately followed by release of calcium ions into the cells. This is responsible for a complex process known as excitation contraction coupling, resulting in contraction of the myocardial muscle fibers. Figure 7 shows the output from a representation by a simulation based on a circulatory system model modified from a representation by Rideout.^[7] The figure shows the temporal relationships among ECG, left atrial pressure (LAP), left ventricular pressure (LVP), and aortic pressure (AoP) through three beats of the cardiac cycle.

1. The P wave reflects the electrical activity that precedes contraction of the atria. Following this event, the atria contract, increasing LAP and giving the ventricles a final filling “kick” just before contraction of the ventricles. The pressure in the ventricle rises only slightly as a result of the atrial activity because the ventricles, which are still in diastole, are still very compliant.
2. The QRS complex reflects the electrical activity that precedes contraction of the ventricles. Following this event, the ventricles begin to contract (the beginning of ventricular systole), causing a rapid rise in intraventricular pressure. When the intraventricular pressure exceeds the aortic pressure (this is the diastolic pressure of approximately 80 mm Hg in the normal human), the aortic valve opens permitting ejection of the ventricular blood into the aorta. The AoP then follows the LVP through its peak (this is the systolic pressure of approximately 120 mm Hg in the normal human). The systolic pressure of 120 mm Hg and diastolic pressure of 80 mm Hg are the basis for normal human blood pressure to be given as 120/80.
3. The T wave reflects repolarization of the ventricles. Following this event, the ventricles begin to relax (initiation of ventricular diastole) and the AoP also falls, following the ventricular pressure. When the ventricular pressure decreases to a point below the AoP, the aortic valve closes and LVP decreases rapidly toward zero. Aortic Pressure then decreases gradually as a result of arterial runoff as the arterial blood perfuses the rest of the body.

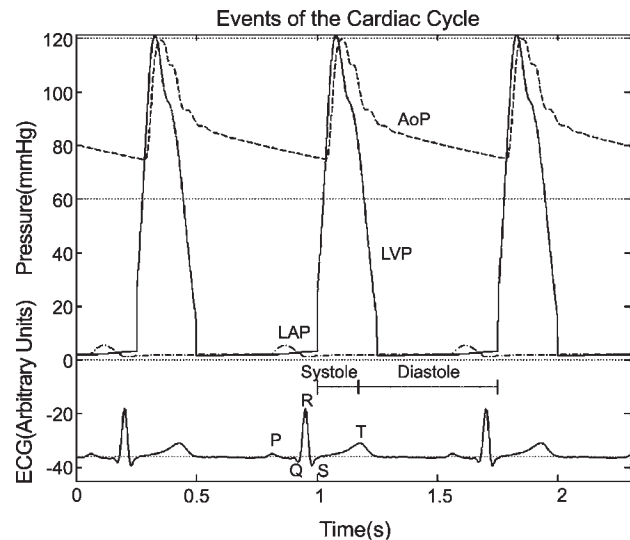


Fig. 7 The output of a circulatory system model showing the relationships among the ECG, left atrial (LAP), left ventricular (LVP), and aortic (AoP) pressures. In this model, the systolic pressure is approximately 120 mm Hg and the diastolic pressure is approximately 75 mm Hg.

4. The time interval between successive R waves, the R-R interval, can be used to estimate the heart rate (HR) where $HR \text{ (beats/min)} = 60 \text{ (seconds per min)} / R\text{-R interval (seconds)}$.

It is important to note here that the electrical activity always precedes the associated mechanical activity. Also note that the repolarization of the atria is associated with an atrial T wave or “T of P”; however, this signal is obscured by the ventricular QRS and is rarely discussed in the literature.

THE 12 LEADS OF THE ECG: DIFFERENT VIEWS OF THE SAME EVENT

Figure 8 shows a three-dimensional object whose shadows look different when light is projected on it from different angles. Similarly, the ECG is a recording of the mean three-dimensional electrical vector of the heart. However, the same event can have different representations, depending on the angle from which this activity is viewed. The standard 12 leads of the ECG “see” the heart from different angles and therefore carry unique information about some particular region of the heart not viewed by the others.

Placement of Electrodes

Electrodes are placed on both the wrists and the ankles and form the basis for the six limb leads that represent

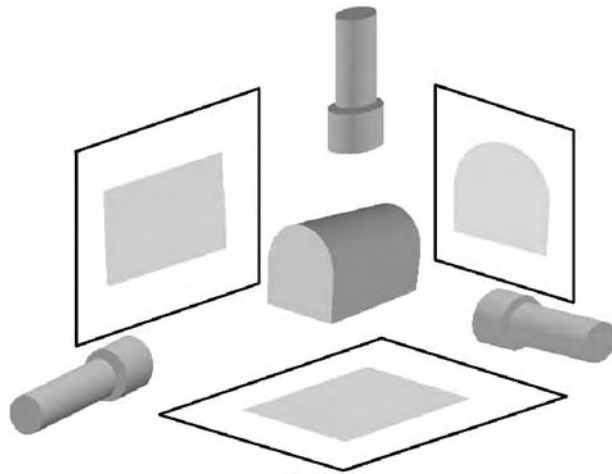


Fig. 8 Flashlights illuminating a three-dimensional-object create different-appearing shadows.

the vertical (frontal) plane. The standard bipolar limb leads (I, II, and III) measure the difference in potential between two limbs.^[4] The augmented unipolar leads (aVL, aVR, and aVF) are measurements between one limb and two of the others.^[4] This configuration increases the size of the potentials by 50% and therefore they are called “augmented” leads. Six more electrodes are placed across the chest wall (Table 1) to form the basis for the precordial leads, V1–V6. These view the heart in the horizontal plane. The different leads are created by a unique combination of the measuring electrodes as shown in Table 2. The particular angles of orientation of the different leads are depicted in Fig. 9.

MEASUREMENT OF THE ECG

With today’s modern instrumentation, the ECG is one of the easier physiological parameters to detect. At the surface of the body, the ECG is an electrical signal in the range of ± 1 mV with a maximum frequency content of approximately 125 Hz in the human. To detect this signal, it is necessary to amplify it by

Table 1 Placement of chest electrodes

Precordial electrodes	Placement on the chest wall
V1	Right 4th intercostal space, right of sternum
V2	Left 4th intercostal space, left of sternum
V3	Midway between V2 and V4
V4	Left 5th intercostal space, midclavicular line
V5	Midway between V4 and V6
V6	Left 5th intercostal space, midaxillary line

Table 2 Connections for standard leads

	Positive electrode	Negative electrode
<i>Standard limb leads</i>		
Lead I	Left arm	Right arm
Lead II	Left leg	Right arm
Lead III	Left leg	Left arm
<i>Augmented unipolar leads</i>		
aVL	Left arm	All other limbs
aVR	Right arm	All other limbs
aVF	Left leg	All other limbs
<i>Precordial (chest) leads</i>		
V1–V6	Corresponding chest electrode	“Common terminal” of all the limb electrodes

approximately $1000 \times$ with a frequency bandwidth requirement of 0 (sometimes referred to as DC for direct current, i.e., nonoscillatory) to 125 Hz. Surface electrodes provide the necessary coupling of the instrumentation with the body. The electrodes are generally made from chlorided silver with NaCl paste to provide electrical conductivity.

This electrode/electrolyte interface carries with it a half-cell potential of approximately 700 mV. Because this value is some 700 times greater than the magnitude of the event we are attempting to measure, some special precautions must be taken. To eliminate this electrode offset potential, a difference amplifier

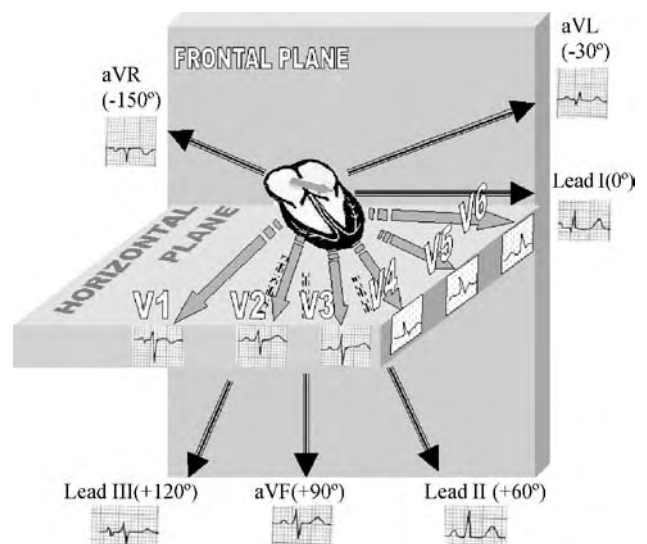


Fig. 9 Twelve leads of the ECG. (Different views of the same electrical vector.)



(also called a differential amplifier) must be used. Because the difference amplifier amplifies the difference between two input quantities, the electrode potentials, if equal in magnitude and phase, will cancel each other out. In reality, however, small differences in the surface characteristics—such as motion artifact—will cause wild swings of the amplified signal.

One means to reduce this effect is to limit the low-frequency cutoff to 0.05 Hz rather than DC. This will eliminate the very-low-frequency changes that accompany electrode artifacts. The useful frequency range of the difference amplifier should therefore be 0.05 to 125 Hz to obtain a high-fidelity recording of the ECG. Without a DC component, the average value of the ECG will be centered at 0/V.

Another important advantage of the difference amplifier is the rejection of common mode inputs such as 60 and 120 Hz noise from the surrounding environment. With battery-powered equipment this periodic noise is not a problem, but if the equipment is powered by the 120-V AC line voltage, as is generally the case, 60 and 120 Hz noise can be a formidable adversary to the equipment designer. Commercial instrumentation amplifiers with common mode rejection ratios (CMRR) of > 100 db are readily available and provide a good rejection of 60 and 120 Hz noise.

ECG Paper

The conventional recording devices plot the ECG onto a special paper strip that has major gridlines (5 mm) and minor gridlines (1 mm). In the horizontal axis (time), each 1 mm represents 0.04 sec and each 5 mm (major grid) represents 0.2 sec. On the vertical axis (volts), each major gridline represents 0.5 mV.

A PRACTICAL ECG AMPLIFIER

A simple practical ECG amplifier is described in the following section. This amplifier can be built by the nominally capable engineer or experimentalist and can be used to investigate the characteristics of the ECG described elsewhere in this article. Most components—such as resistors, capacitors, controls, and operational amplifiers—can be obtained from your local retail outlet. The special components listed can be purchased over the Internet (URLs given).

Partial Parts List

Special components

- INA114 Instrumentation amplifier (Burr-Brown Corp., www.ti.com).
- ISO122 Isolation amplifier (Burr-Brown Corp., www.ti.com).
- AEE00-12Vin DC/DC converter (Astec, Inc., www.astec.com).

Resistors are all quarter-watt $\pm 5\%$ carbon composition resistors. The two 10 k Ω potentiometers (pots) should have a linear taper and, preferably, be wire-wound multiturn pots, for fine-tuning the amplifier. The circuit can be built using a hobby-type breadboard that is also available at your local retail outlet. It should have a dual 15 V supply used to power the nonisolated side of the amplifier.

The circuit shown in Fig. 10 can be used to demonstrate the principles of ECG monitoring. It will permit a reasonably high-fidelity reproduction of the ECG; however, it should not be used for diagnostic purposes.

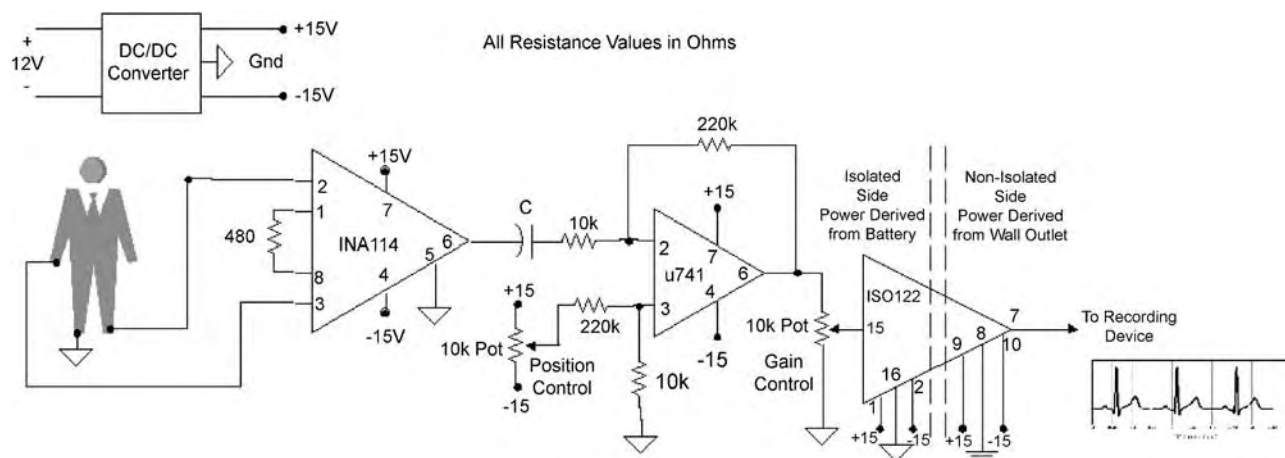


Fig. 10 Circuit diagram for a practical ECG amplifier.

Electrical Safety

The circuit is an electrically isolated circuit that provides a high degree of electrical safety if built properly. The first 2½ stages are powered from a 12 V battery. The DC/DC converter provides the ± 15 V needed to power the components. The INA114 instrumentation amplifier, the $\mu 741$ general purpose operational amplifier (op amp), and the isolated part of the ISO122 isolation amplifier are all powered from this source. The circuit grounds are also all connected to the circuit ground of the DC/DC converter. The output, nonisolated side of the ISO122 isolation amplifier is powered by a separate ± 15 V source whose power is derived from the wall outlet. This isolation method allows the user to connect the ECG amplifier to a subject and connect the output to a display device such as an oscilloscope or computer without fear of compromising the electrical safety of the subject. Under no circumstances should the isolated and nonisolated circuits be wired together. When the device is connected to a subject, there should be no direct connection between the isolated and nonisolated circuitry. *Without this isolation, the subject's electrical safety could be compromised if something were to go wrong or be miswired.*

Circuit Description

The subject is connected to the ECG amplifier using standard ECG disposable electrodes (available from a medical supply outlet, hospital, or urgent care center). The electrodes can be connected to the ECG amplifier using alligator clip leads or snap connectors.

The first stage makes use of the INA114 instrumentation amplifier, which has a differential input, a high CMRR, and a voltage gain of $105 \times$ (using the 480Ω gain resistor). The second stage has a voltage gain of $20 \times$, bringing the total voltage gain to $2310 \times$. The first two stages are coupled through a series capacitance C , whose value should be approximately $318 \mu\text{F}$. This value may be difficult to find, so a smaller capacitor can be used (something in the range of $50 \mu\text{F}$); however, the low-frequency cutoff will be 0.3 Hz, which may eliminate some low-frequency components of the ECG but, for the purposes of this discussion, will not substantially distort the waveform.

To reduce common-mode noise such as 60 Hz interference, the input leads can be twisted together. This will ensure that both inputs of the differential amplifier receive the same noise signal, which will reduce or eliminate this common-mode noise. If needed, a $0.02 \mu\text{F}$ capacitor can be added across the $220 \text{ k}\Omega$ feedback resistor to limit the high-frequency cutoff to 36 Hz, which will reduce 60 Hz noise but may also distort the QRS complex of the ECG.

The last stage uses the ISO122 isolation amplifier to electrically isolate the subject from wall-powered devices such as an oscilloscope or computer. The isolated side is powered from the DC/DC converter, which is connected to the 12 V battery. The nonisolated side is powered from a ± 15 V source whose power is derived from the 120 V AC wall outlet. The rated isolation of the device is 1500 V rms^a and should protect the subject from electrical shock.

To use the device, simply connect three disposable electrodes to the subject. Standard lead II can be obtained using the right arm/left leg leads with the right leg providing circuit ground. Other leads can be selected by moving the location of the electrodes. Standard leads I, II, and III can be used without additional components. The augmented unipolar and precordial leads require a resistor network connected to the input, so it is probably not practical to duplicate those conditions here.

SOME ABNORMALITIES OF THE ECG

The ECG is an invaluable diagnostic, monitoring, and prognostic tool that provides crucial information about numerous pathological states. Presented later is a brief discussion of some. For more detailed information, the reader may refer to any standard textbook on electrocardiography.^[8–10]

Axis Deviation and Hypertrophy

The average direction of the electrical vector of the heart during ventricular depolarization is known as the mean electrical axis. It is defined only in the vertical plane; therefore an estimate of the axis can be made from the shape of the QRS complex in leads I, II, III, aVL, aVF, and aVR. For example, a predominantly positive deflection (R wave) in one of these leads indicates that the mean electrical axis is directed toward that lead. Similarly, a biphasic QRS complex (equal positive and negative deflections) in one of the leads denotes a mean electrical axis that is perpendicular to the direction of that particular lead. Normally, the electrical axis of the heart lies between 0 and 90° . A left-axis deviation commonly occurs in hypertrophy of the left ventricle, which is associated with conditions such as long standing hypertension or aortic valvular diseases. Similarly, a right-axis

^aRMS (or rms) stands for root-mean-square. It is a measure of a sinusoidal voltage (and current) available for power calculations. The line voltage in the home is 120 V_{rms}. The actual peak voltage is 170 V. For a sinusoidal voltage (or current), the rms value is the (peak value)/1.414.

deviation (Fig. 11) may occur in long-standing pulmonary hypertension or severe pulmonary valvular disease. Axis deviation can also occur in conduction defects involving the AV bundle branches. However, in the case of hypertrophy, the amplitude and duration of the QRS complex is increased due to the increased muscle mass, generating more current over a longer period of time.

Heart Blocks

The failure of the AV node to allow normal passage of the depolarization wave from the atria to the ventricles results in varying degrees of "AV block" and can easily be diagnosed from the ECG by looking at the relationship between the P waves and the QRS complexes. In the absence of impulses from the SA node, the AV node itself can act as the pacemaker for the ventricles (but at a rate that is much slower than that at the SA node). In first-degree AV block (Fig. 12a), there is a prolongation of the P-R interval (greater than 0.2 sec). In second-degree AV block (Fig. 12b), some of the atrial depolarization waves fail to cross the AV node resulting in dropped beats. In third-degree or Complete AV block, no depolarization wave is conducted to the ventricles. The atria and the ventricles beat at their own separate rates, resulting in complete dissociation of the P wave and the QRS complexes

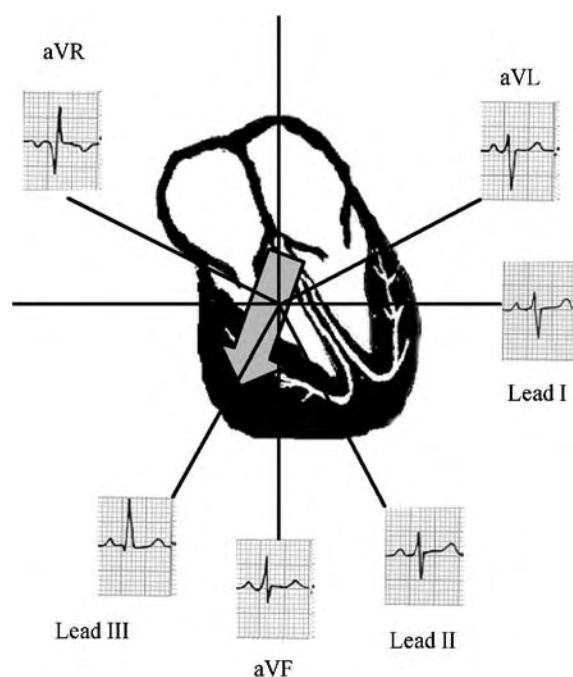


Fig. 11 Right axis deviation caused by right ventricular hypertrophy. Note the predominant R wave in lead III and the biphasic QRS complex in lead II.

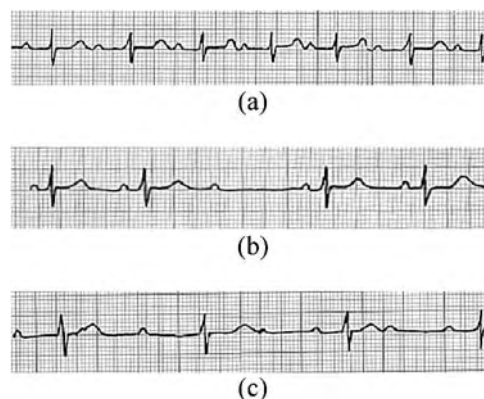


Fig. 12 (a) First-degree heart block. Note the prolonged P-R interval. (b) A type of second-degree heart block with a dropped beat. (c) Third-degree heart block, showing complete dissociation of the atrial and ventricular rhythms.

(Fig. 12c). Conduction blocks can also occur in the left or right bundle branches, which are manifested as axis deviations in the ECG. Some of these heart blocks are potentially life threatening and require implantation of a pacemaker.

Arrhythmias

The heart normally beats with a regular rhythm of 60–100 beats per minute, originating at the SA node. Any deviations from the normal rate, rhythm, or site of origin are known as arrhythmias, and are reflected in the ECG. Depolarization waves arising from locations other than the normal pacemaker sites are known as ectopics. Of these, ventricular ectopic beats (resulting in premature ventricular contraction or PVC) are quite common and produce bizarre QRS patterns in the ECG (Fig. 13b). Another common arrhythmia is paroxysmal supraventricular tachycardia (PSVT), characterized by sudden onset premature ventricular beats originating from the atria or AV node and shown in Fig. 13a. These events can be triggered by alcohol or coffee, quite often spontaneously revert back to a regular rhythm, and do not necessarily indicate a cardiac abnormality.

Fibrillation is another type of arrhythmia caused by the coordinated electrical activity being replaced by random and chaotic impulses generated by the myocardium. This results in a loss of contractile activity. Atrial fibrillation (Fig. 14a) is generally not fatal, because the AV node allows the ventricles to continue with their normal coordinated contractile activity. However, ventricular fibrillation (Fig. 14b) is almost always fatal unless a cardio-pulmonary resuscitation or defibrillation is attempted immediately. Fibrillation is usually the end stage of a variety of

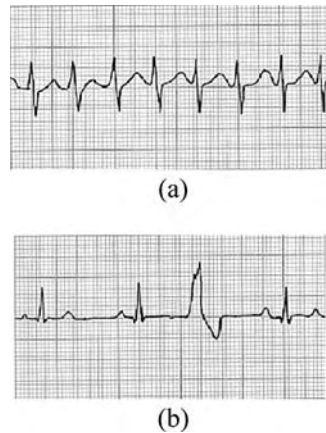


Fig. 13 (a) PSVT. The P waves are buried in the QRS complexes and cannot be seen. (b) PVC. Note the bizarre shape of the ventricular ectopic.

cardiac pathologies. In a healthy heart, a strong electric shock can cause ventricular fibrillation.

Myocardial Infarction

A myocardial infarction (commonly known as the heart attack) occurs due to the injury or death of a localized area of the heart resulting from complete occlusion of one or more coronary arteries. Some of the common ECG findings during an infarction are T wave inversion, S-T segment elevation, and appearance of a new Q wave (Fig. 15).

The S-T segment changes occur because the injured or dead myocardial tissue can no longer repolarize, and therefore events before and after the depolarization process are no longer isoelectric. The Q wave occurs because the dead tissue is electrically silent. This results in the electrical vector being directed away from

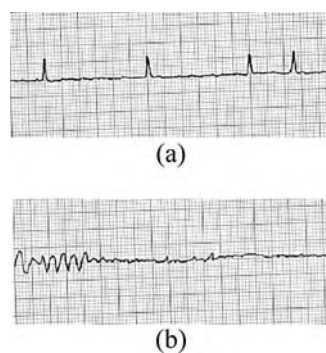


Fig. 14 (a) Atrial fibrillation. Note the absence of the P waves and an irregular and variable ventricular rate due to sporadic passage of impulses through the AV node. (b) Ventricular fibrillation. Absence of any organized electrical activity of the ventricles and a chaotic and random electrical pattern.

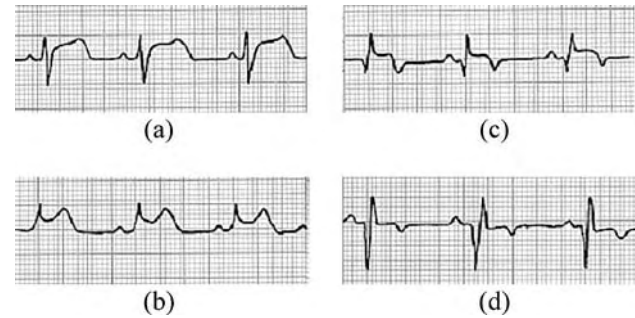


Fig. 15 Some typical ECG changes seen in a myocardial infarction. Note the S-T segment elevation (a, b, c), T wave inversion (c, d), and Q wave (c, d).

the electrodes overlying the dead tissue, causing the strong negative deflection. By examination of the ECG leads in which these changes occur, an estimate of the area of the injury can be made. For example, in an inferior wall infarction these characteristic changes are most pronounced in leads II, III, and aVF (because they “look” at the incident from below), whereas in an anterior wall infarct the changes are more pronounced in the chest leads V1 to V6, depending on the region of the anterior wall involved.

Note that the ECG is always interpreted in the context of the clinical symptoms and other findings, and is seldom used as a stand-alone diagnostic tool.

ARTICLES OF FURTHER INTEREST

Biopotential Amplifiers; Cardiac Bioelectricity; Cardiac Pacemaker; Defibrillators; Tissue Engineering of Cardiac Muscle

REFERENCES

1. Geddes, L.A. Electrocardiography and the Electrocardiograph. In *Cardiovascular Devices and Their Applications*, 1st Ed.; John Wiley & Sons: New York, 1984; 242–243.
2. Hoff, H.E.; Geddes, L.A. The rheotome and its prehistory. *Bull. Hist. Med.* **1957**, *31* (3), 212–347.
3. Ganong, W.F. *Excitable Tissue: Muscle*. In *Review of Medical Physiology*, 19th Ed.; Appleton & Lange: Stamford, Connecticut, 1997; 60–79.
4. Ganong, W.F. Origin of the Heartbeat & the Electrical Activity of the Heart. In *Review of Medical Physiology*, 19th Ed.; Appleton & Lange: Stamford, Connecticut, 1997; 522–538.
5. Guyton, A.C.; Hall, J. Rhythmical Excitation of the Heart. In *Textbook of Medical Physiology*, 9th Ed.; W.B. Saunders Company: Philadelphia, 1996; 121–128.

6. Rawlings, C.A. Traditional Electrocardiographic Voltages. In *Electrocardiography*, 1st Ed.; Spacelabs, Inc.: Redmond, Washington, 1991; 27–47.
7. Rideout, V.C. Simulation of a Complete Cardiovascular Loop. In *Mathematical and Computer Modeling of Physiological Systems*, 1st Ed.; Prentice Hall: New Jersey, 1991; 90–105.
8. Guyton, A.C.; Hall, J. The Normal Electrocardiogram. In *Textbook of Medical Physiology*, 9th Ed.; W.B. Saunders Company: Philadelphia, 1996; 129–134.
9. Thaler, M.S. *The Only EKG Book You'll Ever Need*, 2nd Ed.; J.B. Lippincott Co: Philadelphia, 1995; 10–233.
10. Guyton, A.C.; Hall, J. Electrocardiographic Interpretation of Cardiac Muscle and Coronary Abnormalities: Vectorial Analysis. In *Textbook of Medical Physiology*, 9th Ed.; W.B. Saunders Company: Philadelphia, 1996; 135–148.

Electrogastrography

Robert M. Stern

Department of Psychology, The Pennsylvania State University, University Park, Pennsylvania, U.S.A.

Kenneth L. Koch

Section on Gastroenterology, Wake Forest University School of Medicine, Winston-Salem, North Carolina, U.S.A.



INTRODUCTION

Clinical use of the electrogastrogram (EGG) by gastroenterologists is increasing rapidly because of the ease, quality, and reliability of the recording methods for detecting normal gastric myoelectrical rhythms and gastric dysrhythmias. Increased knowledge about the relationship between gastric dysrhythmias and nausea is evolving. Novel pharmaceutical agents and electrical pacing of the stomach are exciting new therapeutic areas requiring EGG recordings to detect dysrhythmias, develop rational approaches to therapy, and assess results of treatment. The recent increase in the use of the EGG by gastroenterologists has brought with it refinements in both hardware and software.

BRIEF HISTORY

Electrogastrography refers to the cutaneous recording of the gastric pacesetter potentials or slow waves of the stomach. These waves in a healthy human are three cycles per minute (cpm), but do vary between 1 and 15 cpm, which will be described below. On October 14, 1921, after considerable experimentation with rabbits at the University of California in San Francisco, Walter Alvarez^[1] recorded the first human EGG from an elderly woman with an abdominal wall hernia. The woman was so thin that Alvarez could observe three per minute gastric contractions in the upper abdomen with the naked eye that corresponded to the 3cpm electrical waves that he recorded with cutaneous electrodes. Davis,^[2,3] a psychophysicologist, began a series of exploratory studies with the EGG in the mid-1950s. Davis was primarily interested in the interactive effects of psychological and physiological factors on gastric functioning. During the 1970s, Smout, van der Schee, and their colleagues at the University of Rotterdam and Utrecht conducted several validation studies of the EGG and made major contributions in the area of signal analysis. In their landmark 1980 paper, "What is Measured in Electrogastrography?"^[4] they showed that the amplitude of the EGG increases when

contractions occur. They also initiated the use of running spectral analysis (RSA), a technique for displaying EGG data in the frequency and time domain. Today, RSA is used in many laboratories. In 1985, Koch, Stewart, and Stern^[5] recorded simultaneously the EGG and fluoroscopy of the barium-filled stomach. The correspondence between individual EGG waves and antral peristaltic contractions was observed repeatedly during simultaneous EGG–fluoroscopy recordings (Fig. 1). And an ambulatory unit has flown on Space Shuttle flights in an effort to learn more about the relationship of gastric electrical activity and space motion sickness.^[7]

PHYSIOLOGICAL BASIS OF ELECTROGASTROGRAPHY

The anatomical regions of the stomach are the fundus, the body (corpus), the antrum, and the pyloroduodenal area (Fig. 2). Extrinsic innervation of the stomach is provided by the vagus nerve and splanchnic nerves. The pacemaker region is on the greater curvature of the stomach between the fundus and the corpus. From the pacemaker region, spontaneous electrical depolarization and repolarization occur and generate the myoelectrical waves that are termed the gastric pacesetter potentials or slow waves.

Gastric Pacesetter Potentials and Action Potentials

Important anatomical and functional relationships exist among the circular smooth muscle layer, the myenteric neurons, and the interstitial cells of Cajal (ICCs) (Fig. 2, bottom). The ICCs are the pacemaker cells, the cells that spontaneously depolarize and repolarize and set the myoelectrical rhythmicity of the stomach and other areas of the gastrointestinal tract. The interstitial cells are electrically coupled with the circular muscle cells. Low-amplitude rhythmic circular contractions occur at the pacemaker rhythm.

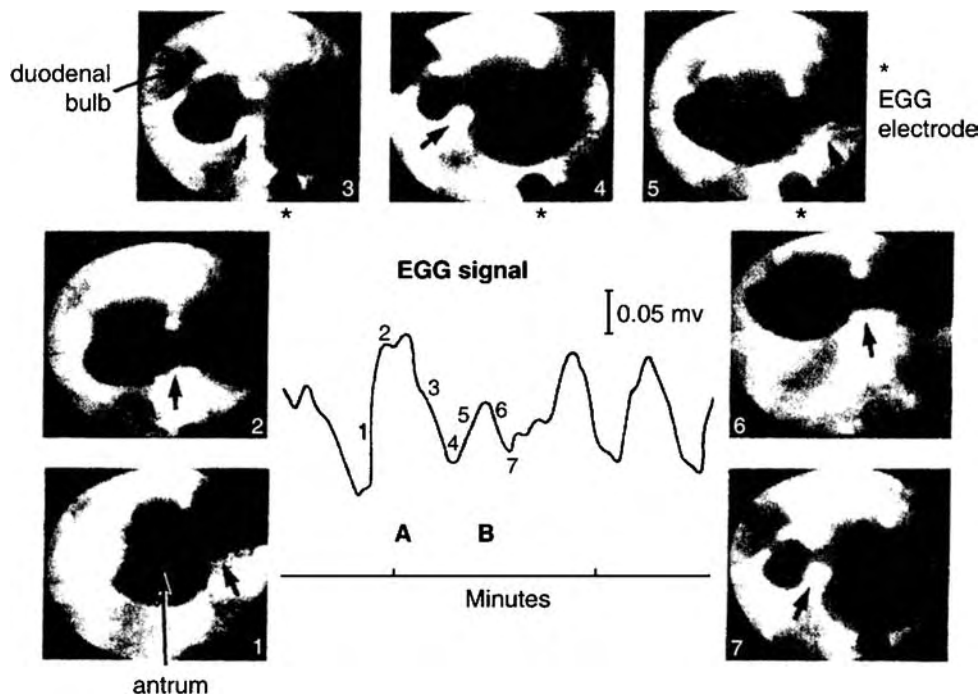


Fig. 1 Simultaneous EGG recording and fluoroscopic imaging of the barium-filled stomach in a healthy subject. Each EGG wave (A and B) correlated with a gastric peristaltic contraction (see radiographic frames 1–4 and 5–7). The small black arrow indicates the migrating peristaltic contraction. Each of the peristaltic waves emptied an aliquot of barium into the duodenum (see frames 3 and 7). These electrocontractile waves are examples of “antral systole.” *One EGG electrode. (From Ref. [6].)

Rhythmicity and contractility of the circular muscle layer are modulated by ongoing excitatory and inhibitory activity of myenteric neurons that synapse with the interstitial cells. The interstitial cells have a variety of other receptors. Electrocontractile activities of the gastric smooth muscle are modified by neuronal and hormonal inputs appropriate for fasting and specific postprandial conditions. Control of rhythmicity may be modulated by a variety of stimuli that affect the interstitial cells and is a focus of intense investigation. For more information about the ICC and the electrical rhythmicity of the stomach, see chapter 2 in Ref. [6].

A human gastric slow wave or pacesetter potential activity generated by the ICCs occurs at a rate of 3 cpm. Electrodes sewn onto the serosa of the stomach record the depolarization and repolarization waves of the pacesetter potentials. The electrical wavefront travels around the circumference of the stomach at a fast rate of speed and migrates slowly toward the antrum at an increasing velocity. As a slow wave disappears in the distal antrum, another slow wave originates in the pacemaker area and begins to migrate toward the antrum approximately every 20 sec. When there is little smooth muscle contractility, these electrical events reflect depolarization and repolarization of the ICCs and some small degree of contractility of the circular muscle cells.

From an *in vivo* electrical viewpoint, the fasting pacesetter potential activity is relatively weak compared with the gastric myoelectrical activity during the postprandial period, when luminal contents and other stimuli augment gastric neuromuscular activity. Human gastric pacesetter potentials migrate around the stomach very quickly and move distally through the antrum in approximately 20 sec resulting in the normal gastric electrical frequency of 3 cpm. It is this moving electrical wavefront that is recorded in the EGG, the gastric myoelectrical activity recorded from electrodes placed onto the surface of the epigastrium. Additional gastric myoelectrical activity occurs when stronger circular muscle contraction occurs (e.g., when vagal efferent activity and release of acetylcholine from the postganglionic cholinergic neurons are elicited in response to ingestion of a meal). Because of the increased gastric myoelectrical activity, the EGG signal during the postprandial peristaltic contractions is generally higher in amplitude in healthy subjects compared with the fasting EGG. Amplitude of the EGG signal may also be affected by the distance of the electrodes from the stomach. These basic gastric myoelectrical activities form the physiological basis for understanding fasting and postprandial EGG patterns recorded in healthy individuals. See chapters 2 and 3 in Ref. [6] for more information about the physiological basis of the EGG.

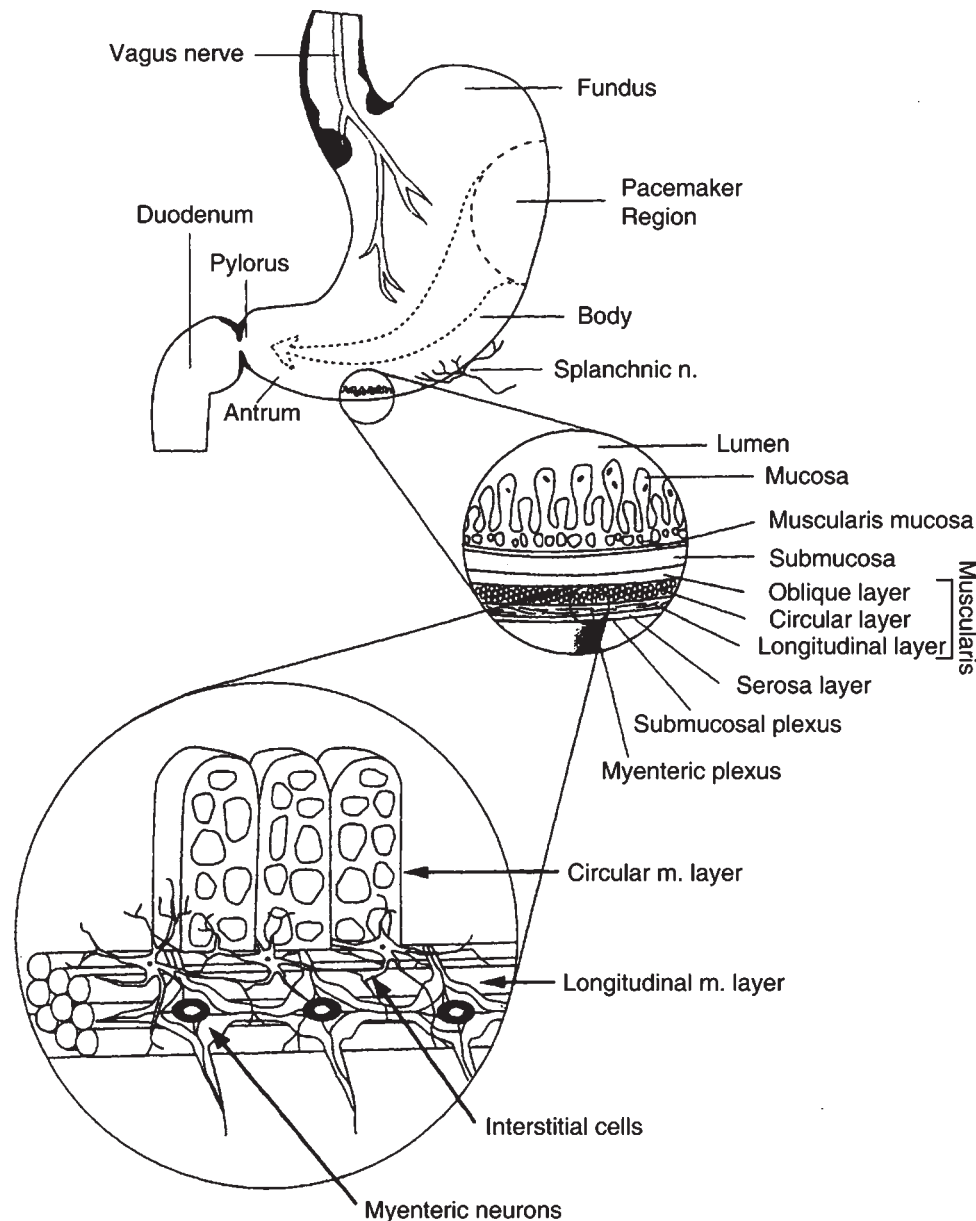


Fig. 2 Anatomical regions of the stomach. The upper figure shows the fundus, the body, and the antrum of the stomach. The pacemaker region is the origin of gastric pacesetter potential activity. The cutout of the gastric wall shown in the middle portion of the figure depicts the lumen of the stomach, the mucosa, and the muscularis of the stomach wall. The main muscle layers of the stomach are the longitudinal and circular muscle layers, with some contribution by the oblique muscle layer. In the lower portion of the figure, the relationship among the circular muscle layer, the ICCs, and the myenteric neurons of Auerbach's plexus are shown. The ICCs synapse with the circular muscle layer as well as the myenteric neurons. The ICCs are the origin of the electrical rhythmicity that is recorded as the gastric pacesetter potentials. The gastric pacesetter potentials coordinate the frequency and the propagation of the circular smooth muscle contractions. (From Ref. [6].)

RECORDING EGG

Electrogastrograms are recorded by placing electrocardiogram (EKG)-type electrodes on the surface of the epigastrium. The EGG is one of several biological signals that can be recorded from the electrodes on the epigastrium. Some of these signals, such as the EKG,

are much stronger than the EGG. The EGG signal is relatively of low amplitude, ranging from approximately 100 to 500 μV . Thus, the EGG signal must be properly amplified and filtered. To reduce baseline drift and to remove unwanted cardiac and respiratory signals, a 0.016 Hz high-pass filter and a 0.25 Hz low pass filter are used. These filters created a bandpass from

approximately 1 to 15 cpm through which the desired gastric myoelectrical signals pass during the EGG recording.

Equipment

Electrodes

High-quality, fresh, disposable electrodes such as those used for EKG recording are recommended. To minimize artifacts in the EGG recording caused by electrode movement on the skin, it is best to use electrodes that adhere very well to the skin, either disposable or reusable. The size of the electrode surface is not important, but the electrical stability of the electrode is important. The electrodes should show little bias or offset potential because the EGG signal is relatively of low amplitude and low frequency.

Recording equipment

A high-quality recording system is needed to amplify and process the 100–500- μ V EGG signal that ranges from 1.0 to 15.0 cpm. Several medical device companies produce complete EGG recording and analysis systems that include appropriate amplifiers and filters with analog-to-digital boards that digitize the EGG signal for analysis with software. Ambulatory EGG equipment is also available, and data are stored on a RAM card for later analysis. A visual display of the raw EGG signal is very important to select artifact-free EGG signal for visual or computer analysis and interpretation.

Recording Procedure

Electrode placement

EGGs are obtained with electrodes arranged for bipolar recordings. Electrodes are placed on the skin surface of the epigastrium over the general area of the corpus and antrum of the stomach. A reference electrode is positioned on the right side of the patient's abdomen. One active electrode should be placed approximately 10 cm cephalad from the umbilicus and 6 cm to the patient's left. It is important to place this electrode below and not on the lowest rib to avoid respiratory signals. The second active electrode should be placed approximately 4 cm above the umbilicus (midway between the umbilicus and xiphoid) on the midline of the abdomen. The reference electrode is placed 10–15 cm to the right of the midline electrode, usually along the midclavicular line and 2–3 cm below the lowest anterior rib on the right side. The recording

sites selected will depend on the nature of the EGG signal desired. From our experience, the electrode locations described above will provide the largest possible amplitude and least artifact in EGGs from most patients. The exact placement of the electrodes is not important if the frequency of the EGG signal is what is of interest.

We have sometimes recorded from two abdominal sites simultaneously and have chosen the more artifact-free channel for analysis and interpretation. Chen et al. have recorded multichannel EGGs, using four active electrodes and a common reference electrode, and reported detecting gastric slow wave propagation^[8] and quantifying the degree of slow wave coupling in normals and patient groups such as individuals with systemic sclerosis.^[9] Single channel EGG recording is usually sufficient for diagnosing gastric dysrhythmias in patients.

The EGG can be recorded from the patient's wrists, but the amplitude will be low compared to recordings from the abdomen because the wrist electrodes are far from the source of the signal, i.e., the gastric pacemaker in the corpus-antrum.^[10] On the other hand, wrist electrodes are recommended when recording from obese patients to minimize the fat layer under the electrodes, which acts as an insulator and decreases the amplitude of the EGG signal. The patient's girth may also affect the amplitude of the EGG signal.

Procedure

Patients should be instructed what and when to eat prior to an EGG recording session because the contents of the stomach will affect gastric neuromuscular activity and hence EGG activity. For our research studies, subjects are instructed to fast for at least four hours prior to the experimental session. For our clinical studies, patients fast after midnight, and then ingest a 200 kcal breakfast of two pieces of toast and 4 oz. of apple juice two hours prior to the EGG test. By controlling the pretest meal, a more standardized baseline EGG is obtained prior to provocative testing.

Prior to the attachment of electrodes, the patient's skin must be prepared by shaving excess hair, gently abrading the skin, and cleaning the area with alcohol to lower electrical resistance and achieve optimal electrode–skin interface. This is a very important part of the EGG recording procedure. If the electrical resistance between the electrode and skin is high, then the EGG signal amplitude will be decreased and artifacts may increase in the signal.

The best EGG recordings will be obtained if the patient reclines about 30° in a comfortable chair. The patient should be instructed to minimize talking and movement during EGG recording to prevent

movement artifacts. If possible, the patient should be in a quiet room separated by some distance from the person doing the testing. Loud noises, crying children, and other stimuli that might disturb the patient and the EGG recording should be avoided. If ambulatory EGG records are desired, then Fetrodes (UFI, Morro Bay, California, U.S.A.) should be used. Fetrodes are miniature preamplifiers that are attached to the electrodes, and they reduce movement artifacts. For more details about the procedure used to record EGG, see chapter 4 in Ref. [6].

ANALYSIS OF THE EGG

The analysis of EGG recordings involves an initial visual inspection of the signal to assess the quality of the signal, identification of artifacts, and selection of the minutes of EGG signal for the computer to analyze. The sections below describe the procedures used to analyze the EGG for clinical and research studies.

Visual Analysis of EGG Signal

All raw EGG recordings should be visually inspected to identify 3-cpm signals, gastric dysrhythmias, and any artifacts in the signal. Certain characteristics of the EGG and qualitative judgments can be made from visual inspection of an EGG record. Artifact-free minutes of the EGG signal must be selected for analysis.

Computer Analysis of the EGG Signal

There are three steps involved in the computer-aided quantitative analysis of the EGG:

1. The EGG recording is filtered digitally to remove unwanted frequencies such as 0.5-cpm ultra-slow drift, respiratory, and cardiac rhythms.
2. A spectral analysis is performed on the selected artifact-free minutes of EGG signal to quantify the different frequencies of interest in the signal.
3. Specific EGG parameters such as percentage distribution of EGG power are calculated.

Filters and the EGG signal

To record a high-quality EGG, myoelectrical activity that occurs at lower or higher frequencies must be filtered out. These filters reduce baseline drift and eliminate most respiratory and cardiac rhythms. Using a 0.016-Hz (~ 1 cpm) high-pass filter and a 0.25-Hz

(15 cpm) low pass filter, a bandpass from 1 to 15 cpm is achieved. The bandpass filter described above allows only frequencies between 1 and 15 cpm to reach the amplifier. The frequencies of interest from the stomach and duodenum are within the 1–15 cpm range: bradygastria (1–2.4 cpm), normal range (2.5–3.7 cpm), tachygastria (3.75–10.0 cpm), and duodenal/respiration (11–15 cpm).

Adaptive filtering

In adaptive filtering, a separate reference signal is used as an error signal or noise, which is subtracted from the primary signal of interest. For further discussion of adaptive filtering, see Ref. [11].

Spectral analysis of the EGG signal

Fourier analysis and RSA are used to determine the frequencies present in the EGG signal.^[12]

Fourier analysis

In the fast-Fourier transform (FFT) analysis of EGG recordings, the power in the bradygastria, normal, tachygastria, and duodenal/respiration ranges is determined. The percentage of the EGG power from 1 to 15 cpm in each of the four relevant frequency ranges is also expressed as the percentage distribution of EGG power. Ratios of the power can be calculated to indicate the changes within a frequency band of interest before and after a test meal or stimulus. The power value itself can also be used as a quantitative measure.

Running spectral analysis

Van der Schee, Smout, and Grashuis^[13] described modifications of the spectral analysis method—the RSA—to quantify frequencies in the EGG signal. The RSA is of interest because it provides EGG frequency and power information over time, whereas, the FFT provides only power as a function of frequency. Figs. 3A and 3B show FFTs of two minutes of EGG signal before and after a meal. Fig. 4 shows an RSA of a 15-min baseline EGG followed by 30 min of EGG signal recorded after the subject ingested 550 ml of water. In the RSA, the power of each of the four relevant frequency ranges in the EGG can be seen in the pseudo-three-dimensional (3-D) plot. The clear peaks at 3 cpm in the RSA (B) correspond with the clear 3-cpm waves seen in the rhythm strip.

The mathematical procedures used to convert the raw EGG signal to an RSA and pseudo-3-D are carried out by the proprietary EGG hardware and

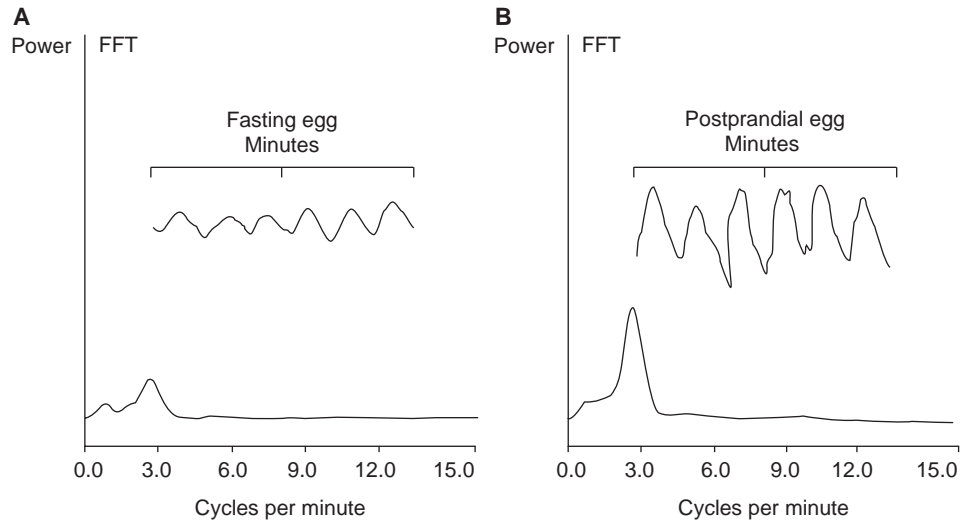


Fig. 3 Two-minute EGG rhythm strips and corresponding FFT recorded from a healthy subject. (A) The EGG recorded during fasting shows a low-amplitude 3-cpm EGG signal. The FFT plot of this signal shows a low-power peak at 3 cpm and (B) the postprandial EGG signal from the same subject after ingesting two pieces of toast and 8 oz. of apple juice shows an increase in amplitude in the 3-cpm EGG waves. The FFT shows a corresponding increase in power at 3 cpm. (From Ref. [6].)

software sold by medical equipment companies, but the first step in any quantification procedure is to insure that quality data are being analyzed. Hence, time must be taken to insure the quality of EGG recordings

before computer analyses are undertaken. The amplifying and recording system should filter out signals below 1.0 cpm and above 15 cpm. With these filter settings the slow rhythms (1.0–2.5 cpm) are recorded

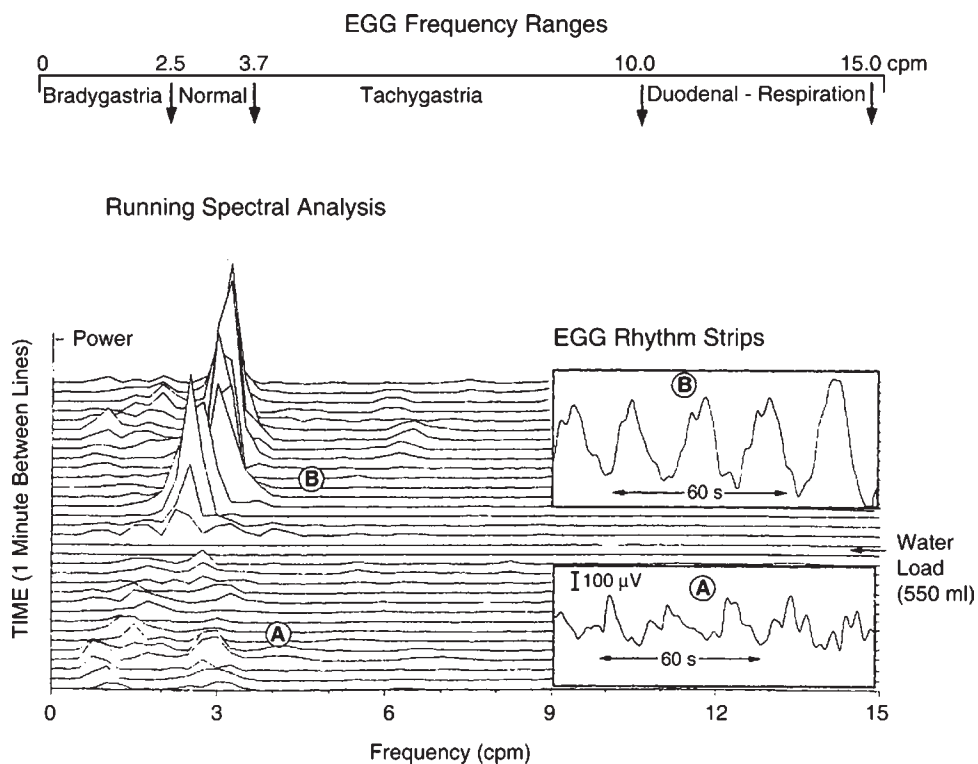


Fig. 4 The RSA is shown with frequency in cpm on the X-axis, time in minutes going from bottom to top on the Y-axis, and power is depicted as the peaks in the Z-axis. During baseline, the subject had low-amplitude 3-cpm EGG waves (see A in EGG rhythm strip). Note the corresponding peaks at 3 cpm in the RSA (A). After 550 ml of water was ingested during the water load test, frequency decreased briefly to approximately 2 cpm, followed by clear 3-cpm peaks of increased power in the RSA (B). The EGG rhythm strip from after water ingestion (B) shows increased amplitude and regularity of 3-cpm activity. Thus, increases in the amplitude of EGG waves are reflected as increases in power in the RSA at that frequency. The frequency ranges of relevance (bradygastria, normal, tachygastria, and duodenal respiration) are shown above the RSA. (From Ref. [6].)

accurately, and the shifts in baseline because of DC potentials that are <0.5 cpm are eliminated. Frequencies higher than 15 cpm are filtered out to avoid contamination of the gastric signal by EKG signals. Respiration signals can obscure the EGG when breathing rates falls near that of tachygastria (3.7–10 cpm) or duodenal slow wave frequencies (12–14 cpm). Removal of respiration signals by setting the low pass filter at 10 cpm or less would also remove duodenal signals. Therefore, a separate respiration tracing is used to select visually and exclude minutes that contain the obvious artifacts related to respiratory rates or movement artifacts.

A/D conversion units typically allow a wide range of sampling rates. For over 20 years we have used 4.267 Hz for our sampling rate because it yields 256 data points per minute, and 256 is a power of 2, a requirement for the spectral analysis we use. Regardless of the analysis method used, we recommend a sampling rate of at least 4 Hz to eliminate the potential for aliasing.^[12] If aliasing is not eliminated prior to computer analysis, some EKG power will be erroneously added to the EGG power. This is referred to as leakage. A high sampling rate for EGG recording such as 4.267 Hz allows complete resolution of all potential frequencies in the EGG signal and thus prevents aliasing.

Once the entire EGG recording has been digitized, the data must be preprocessed to meet the assumptions of FFT analysis. Because the normal 3-cpm frequency in the EGG is a relatively slow frequency, we recommend the analysis of at least four minutes of EGG data. The data are “centered” for spectral analysis around a mean of zero. This is accomplished by subtracting the mean of the data segment from each individual data point. Additionally, the EGG is likely to contain some ultra-slow components (<0.5 cpm) that may reflect a drifting baseline. These ultra-slow components may not be completely filtered out by the bandpass filter. Such extremely low frequency shifts and simple linear trends should be removed. A high-pass digital filter that attenuates frequencies below 0.01 Hz (0.6 cpm) accomplishes the removal of ultra-slow frequency components in the digitized EGG data. One can also remove simple linear trends by fitting a least squares regression line and subtracting it from the data segment. Thus, the analog EGG signal is converted to a digital series of values and placed in a file that is subjected to various quantitative analyses.

After the preprocessing steps for the EGG signal are completed, the data undergo the FFT. Consecutive four-minute EGG data “segments” are overlapped to produce an RSA which displays frequency and power over time. We use a 75% overlap in both clinical and research recordings.^[12] In other words, segment 1

contains minutes 1–4 of EGG data, segment 2 includes minutes 2–5, segment three includes minutes 3–6, etc. Thus, one minute of new EGG signal (e.g., minute 6) is added to the previous three minutes of EGG signal (minutes 3, 4, and 5), thus accomplishing the 75% overlap. An FFT is then performed on each of the overlapped four-minute EGG data segments. These overlapping power spectra are plotted in a pseudo 3-D fashion to determine changes in power at various frequencies over time.

RSAs provide an informative picture of EGG frequency and power changes over time. Electro-gastrogram frequency changes of 60 sec or longer will usually appear as a concordant change at that frequency in the RSA. It should be noted that frequency changes lasting less than 60 sec may not be detected. For such cases adaptive spectral analysis methods are recommended.^[11] Dysrhythmias with a duration more than 60 sec are considered clinically significant and are easily seen in the raw EGG recording and in the RSA.^[13]

Specific EGG Parameters

1. *Percentage distribution of EGG power in the four frequency bands of interest:* The percentage distribution of EGG power in the four relevant frequency bands is the measure that the authors have found most useful in over 20 years of experience in recording and analyzing EGGs in studies involving patients and healthy research participants.^[2–4] The percentage distribution of total EGG power is calculated for each of the frequency bands of interest: bradygastria, normal 3 cpm, tachygastria, and duodenal/respiration. The power estimates for a given frequency range (e.g., the 1–2.5-bradygastria range) are obtained from the spectral analysis. The sum of the bradygastria power for a given time period is divided by the power for the entire frequency range (1–15 cpm) for the same time period, and multiplied by 100%. Thus, the percentage distribution of all the power from 1 to 15 cpm is determined for each of the four relevant frequency ranges: bradygastria, normal, tachygastria, and duodenal/respiration.^[12] The percentage of EGG power in the different frequency bands is altered by provocative stimuli such as the Water Load Test, or a caloric meal, or drug therapy. This measure should not be confused with Percentage of Time with The Dominant Frequency in the Normal (or other) Frequency Bands, which is described below.
2. *Percentage of time with the dominant frequency in the normal (or other) frequency band:* This measure

calculates the percentage of time (within a defined period) that the dominant EGG frequency (the specific frequency with the greatest power as determined by computer analysis) lies within a certain range of frequencies, e.g., 2–4 cpm.^[11] An adaptive filter analysis of the EGG is first computed using one or two minutes of EGG data to construct each power spectrum. A spectrum is considered normal if the dominant (highest) peak lies in the 2–4-cpm range. The percentage of time that the highest peak is in the 2–4-cpm range is determined visually. Thus, if eight one-minute spectra had the highest peak in the 2–4-cpm range in a 10-min period, then the percentage of time the dominant frequency was in the normal range would be 80.

3. *Power ratio:* The EGG activity in the frequency band of interest after a test or therapy is compared with the EGG activity in that same frequency band during the baseline period.^[12] For example, a clinician might calculate the ratio of power in the normal range before and after treatment with a particular drug used to reduce nausea and use the ratio as a measure of improvement. However, an increase in EGG activity in all frequency bands may occur and would make interpretation of the specific ratio difficult.
4. *Dominant frequency and its stability:* The dominant frequency is determined visually by identifying the frequency with the greatest power in the FTT or RSA during a specific time period. Frequencies can change even in healthy subjects. Note that a frequency dip to 1–2 cpm occurs briefly after a meal. Smout, Jebbink, and Samsom^[14] calculated various measures of stability of the dominant frequency and concluded that there was little evidence that either dominant frequency or the stability of the dominant frequency was physiologically related to gastric function.

NORMAL ELECTROGASTROGRAPHIC PATTERNS IN HEALTHY SUBJECTS

Fasting Electrogastrographic Patterns

During prolonged fasting, the stomach and the small intestine proceed through several stereotyped contraction patterns. The clearest 3-cpm EGG signals generally occur during phase I, when there is contractile quiescence, at least as measured by intraluminal pressure catheters. At this time, the 3-cpm EGG activity is the most stable and likely reflects the underlying activity of the ICCs. During phase II, when contractions are irregular, the 3-cpm EGG signal is

unstable. During phase III contractions, 3-cpm EGG waves may show increased amplitude. Thus, increased amplitude of the EGG signal during fasting does not reliably reflect the presence of phase III contractions.

In general, the EGG signal recorded when the stomach was empty is usually a low-amplitude 3-cpm rhythm or unstable signal that may vary from the 3 cpm normal range (2.5–3.7 cpm) to tachygastria (3.7–10.0 cpm) and bradygastria (1.0–2.5 cpm) ranges. For these reasons, investigators have used a variety of meals and drugs to stimulate the fasted or baseline gastric myoelectrical activity. The next section reviews postprandial EGG patterns in healthy individuals.

Postprandial Electrogastrographic Patterns

The stomach must respond appropriately to a vast variety of intraluminal stimuli. In addition, the stomach reacts to visual, olfactory, and gustatory stimuli during the cephalic phase of digestion. The stomach neuromuscular apparatus initially responds to the ingestion of solid foods and liquids with relaxation of the gastric muscular wall of the fundus. The fundus relaxes to receive the volume of food ingested. This important phenomenon is called receptive relaxation.

The relationships between the emptying of barium from the stomach and gastric electrical events recorded from a healthy subject are shown in Fig. 1. The fluoroscopic images of the stomach show that barium suspension is emptied from the corpus antrum into the duodenum by a peristaltic contraction. Sequential radiographs of the stomach during a gastric peristaltic wave are shown in frames 1, 2, 3, and 4. The asterisk indicates EGG recording electrodes positioned on the epigastrium. The small, black arrows indicate a circular muscle contraction that is migrating through the antrum (frames 1 and 2) and begins to empty barium into the duodenal bulb (frames 3 and 4). Electrogastrogram wave A with frames 1–4 was recorded during the progression of the gastric peristaltic wave shown by the radiographs in frames 1 to 4. During the gastric peristaltic wave, the EGG wave peaked (see 2 on EGG wave A) and then the amplitude decreased (see 4 on EGG wave A). Electrogastrogram wave B was recorded during the next gastric peristaltic wave, which is shown in the radiographs in frames 5–7. This peristaltic wave also emptied barium into the duodenal bulb as shown in frame 7. Thus, each gastric peristaltic wave emptied a small aliquot of the barium from the antrum into the duodenal bulb. One EGG wave reflected the myoelectrical activity of a peristaltic wave that migrated from proximal to distal stomach. This sequence repeated itself until the noncaloric barium meal was emptied from the stomach.

The important point is that each EGG wave is a summary of the electrical events that occur during a barium-induced peristaltic wave. The peristaltic wave is the “electrical halo” that is formed as the pacesetter potential linked to plateau or spike potential moves from the corpus to the pylorus beneath the recording electrodes. Thus, the EGG signal as recorded by surface electrodes after a barium meal, a provocative water load test, or a caloric meal reflects the gastric pacesetter potential activity (ICC activity) and the myoelectrical activity associated with circular muscle contraction (action and plateau potentials). These are the electrical events that produce gastric peristalsis, the peristaltic contractions that produce emptying of the barium meal.

CONCLUSIONS

The 3-cpm gastric pacesetter potential is a fundamental electrical phenomenon of the stomach. This low-frequency biorhythm is the basis for normal neuromuscular function of the stomach. In regard to the origins and the various neural and hormonal influences that affect the 3-cpm rhythm, many mysteries remain. Ongoing and future inquiries into the very nature of rhythmicity will provide deeper understanding of gastric myoelectrical activity and the electrical activity detected in the EGG. A delicate balance maintains normal 3-cpm activity. Stomach electrical rhythmicity is rather unstable during fasting, for example, compared with the rhythmic 3-cpm electrical events and contractile events that occur in the postprandial period. What mechanisms produce these fasting and postprandial electrical changes? Are neural or hormonal circuits most critical? Are extrinsic or intrinsic nerves the most important?

Different EGG patterns induced by different meals reflect the different gastric neuromuscular work required to receive, mix, and empty the specific meal. Characteristics of the EGG signal from frequency to amplitude may also correlate with perceptions of stomach activity. Future studies of the EGG may better define these relationships and provide important clues regarding the contribution of the stomach to perceptions of hunger and satiety, perceptions that are important in regard to our understanding of obesity and other eating-related disorders.

ARTICLE OF FURTHER INTEREST

Biopotential Amplifiers

REFERENCES

1. Alvarez, W.C. New methods of studying gastric peristalsis. *J. Am. Med. Assoc.* **1922**, *22*, 1281–1284.
2. Davis, R.C.; Garafolo, L.; Gault, F.P. An exploration of abdominal potentials. *J. Comparative Physiol. Psychol.* **1957**, *50*, 519–523.
3. Davis, R.C.; Garafolo, L.; Kveim, K. Conditions associated with gastrointestinal activity. *J. Comparative Physiol. Psychol.* **1959**, *52*, 466–475.
4. Smout, A.J.P.M.; van der Schee, E.J.; Grashuis, J.L. What is measured in electrogastrography? *Digestive Dis. Sci.* **1980**, *25*, 179–187.
5. Koch, K.L.; Stewart, W.R.; Stern, R.M. Effect of barium meals on gastric electromechanical activity in man. A fluoroscopic-electrogastrographic study. *Digestive Dis. Sci.* **1987**, *32*, 1217–1222.
6. Koch, K.L.; Stern, R.M. *Handbook of Electro-gastrography*; Oxford University Press; New York, 2004.
7. Harm, D.L.; Sandoz, G.R.; Stern, R.M. Changes in gastric myoelectric activity during space flight. *Digestive Dis. Sci.* **2002**, *47*, 1737–1745.
8. Chen, J.D.Z.; Zou, X.P.; Lin, X.M.; Ouyang, S.; Liang, J. Detection of gastric slow wave propagation from the cutaneous electrogastrogram. *Am. J. Physiol.* **1996**, *277*, G424–G430.
9. McNearney, T.; Lin, X.; Shrestha, J.; Lisse, J.; Chen, J.D.Z. Characterization of gastric myoelectrical rhythms in patients with systemic sclerosis using multichannel surface electrogastrography. *Digestive Dis. Sci.* **2002**, *47*, 690–698.
10. Stern, R.M.; Stacher, G. Recording the electrogastrogram from parts of the body surface distant from the stomach. *Psychophysiology* **1982**, *19*, 350.
11. Chen, J. Signal processing and analysis. *Neurogastroenterology* **2000**, *3*, 104–112.
12. Stern, R.M.; Koch, K.L.; Muth, E.R. Gastrointestinal system. In *Handbook of Psychophysiology*, 2nd Ed.; Cacioppo, J.T., Tassinari, L.G., Berntson, G.G., Eds.; Cambridge University Press: New York, 2000; 294–314.
13. Van der Schee, E.J.; Smout, A.J.P.M.; Grashuis, J.L. Application of running spectral analysis to electrogastrographic signals recorded from dog and man. In *Motility in the Digestive Tract*; Wienbeck, Ed.; Raven Press: New York, 1982; 241–250.
14. Smout, A.J.P.M.; Jebbink, H.J.A.; Samsom, M. Acquisition and analysis of electrogastrographic data: The Dutch experience. In *Electrogastrography*; Chen, J.Z., McCallum, R.W., Eds.; Raven Press: New York, 1994; 3–30.

Electromyography

Didier Gamet

Odette Fokapu

Laboratory of Biomechanics and Bioengineering, UMR CNRS 6600, Research Center Royallieu, University of Technology of Compiègne, Compiègne, France

INTRODUCTION

Electromyography is the evaluation of muscles based on detection of the muscle's electrical activity (electromyogram or EMG) and its study (EMG signal processing and interpretation of the results). The electrical activity that precedes the muscle contraction is detectable using simple sensors and is painless for the majority of sensors used. In the case of voluntary muscle contractions, study of the EMG signal assesses the correct operation of the motor ways from the motor cortex to the muscle fiber itself. Thus, electromyography is a powerful method used in several investigative domains, including medical research (as functional neurology, gait and posture analysis), rehabilitation (as neurological rehabilitation, physical therapy), ergonomics (as risk prevention), and sports science (as biomechanics, movement analysis, training control).

Understanding of EMG signal composition is based on the anatomic and functional knowledge of the neuromuscular system. Skeletal muscle is made up of a large number of mostly parallel muscle cells called muscle fibers. Each fiber, as with all excitable cells, has electrical activity. However, to understand how an EMG signal is generated from the electrical activity of the muscle fiber, it is important to consider not only the muscle but also the nerve that transmits the motor command. Thus, the neuromuscular system must be considered an association of several functional units, called motor units.

The review is organized as follows: The motor unit is described in the first section. In the second section, genesis of an EMG signal from motor unit activity is presented. Electromyographic signal processing is described in the third section, including EMG detection, signal conditioning and recording, and data processing. At the output of processing chain, the EMG signal may be represented by several graphic models and characterized by several quantifiers in the time and frequency domain. The study of EMG signals and their changes, presented in the fourth section, is valuable to the understanding of physiological phenomena such as muscle fatigue, as well as to the discovery and

diagnosis of abnormalities in nerve and muscle. However, EMG signal study must take into account the test conditions resulting from measurements of other signals (for example, mechanical signals) in order to obtain reliable and repeatable results, as described in the fifth section. Lastly, in the sixth section, some examples of EMG signal changes are described, notably during muscle contraction and after electrical stimulations as used in clinical EMG.

MOTOR UNIT (MU)

The MU is the functional unit of the muscle (Fig. 1A). The MU is defined as the association of a motoneuron α (nervous cell) and the muscle fibers it innervates.^[1] The MU has the ability to contract itself (motor action or contraction) following a driving order (neuronal excitation) because of excitation–contraction coupling.^[2]

Nervous and muscular cells (muscle fibers) are excitable cells. These cells react to all external events (named excitation) that are mechanical, chemical or electrical in nature. As is the case for all live cells, these cells have electrical polarity on both sides of their cytoplasmic membrane (membrane potential or resting potential). Using microelectrodes (typical sensors in the cell electrophysiological domain) with a very thin tip, it is possible to measure the difference in potential between the outside (taken as an electric reference) and the inside of the cell (this detection type is qualified as intracellular detection). The measured membrane potential is stable with time and ranges between -70 and -90 mV according to cell type. After excitation, cells react with a transitory variation of the electrical polarity of the membrane, called the action potential (AP). The AP (Fig. 2) is identical in amplitude and in duration and is characteristic of the cell type (nervous or muscle). Moreover, AP is propagated along the membrane of these cells (axonal membrane for nervous cells, fiber membrane for muscle fibers). The cell parts considered are long extensions: the axon of the nervous cell whose cell body is located in the ventral horn of the spinal cord or a long cylindrical cell (muscle cell).

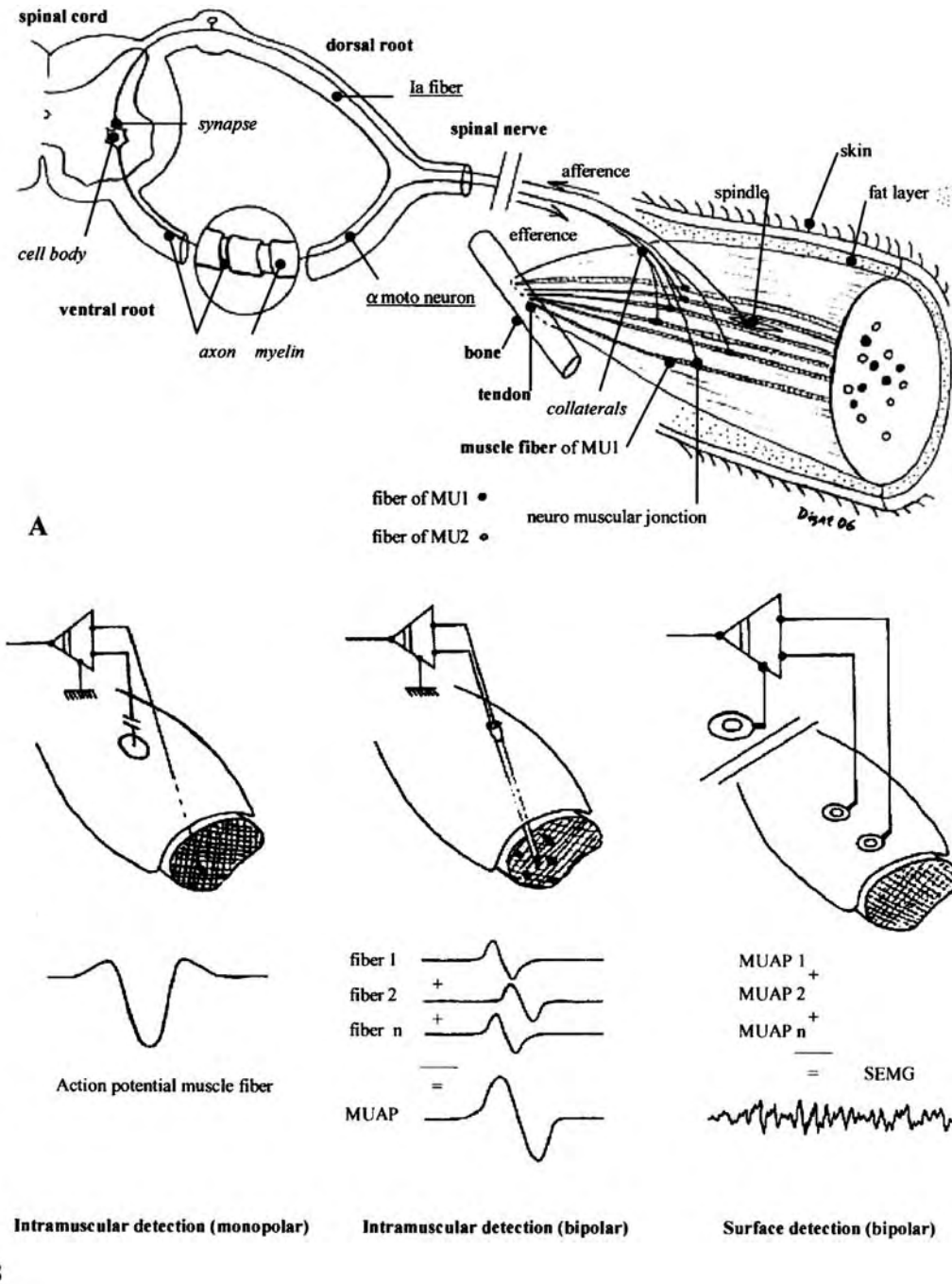


Fig. 1 (A) Schematic anatomical drawing of motor unit (MU) and (B) typical waveforms according to detection modalities. MUAP, motor unit action potential; SEMG, surface electromyogram.

The AP is transmitted from an axon (nervous AP) to a cell muscle (muscle AP fiber) through the neuromuscular junction (NMJ), also called the motor end plate, a chemical synapse whose neurotransmitter is acetylcholine. The single axon is subdivided at its end into several branches or collaterals. Thus, axonal AP is reproduced on each collateral ending with an NMJ and innervates several muscle fibers.

A muscle associated with its motor nerve is described as a set of MUs, from ten MUs for small muscles to several thousands of MUs for large muscles. The muscle fibers of one MU are distributed in the muscle. This distribution determines the concept of MU territory defined by Buchthal et al.^[3] In the first approximation, the MU territory corresponds to the muscle surface (in the muscle cross-section view)

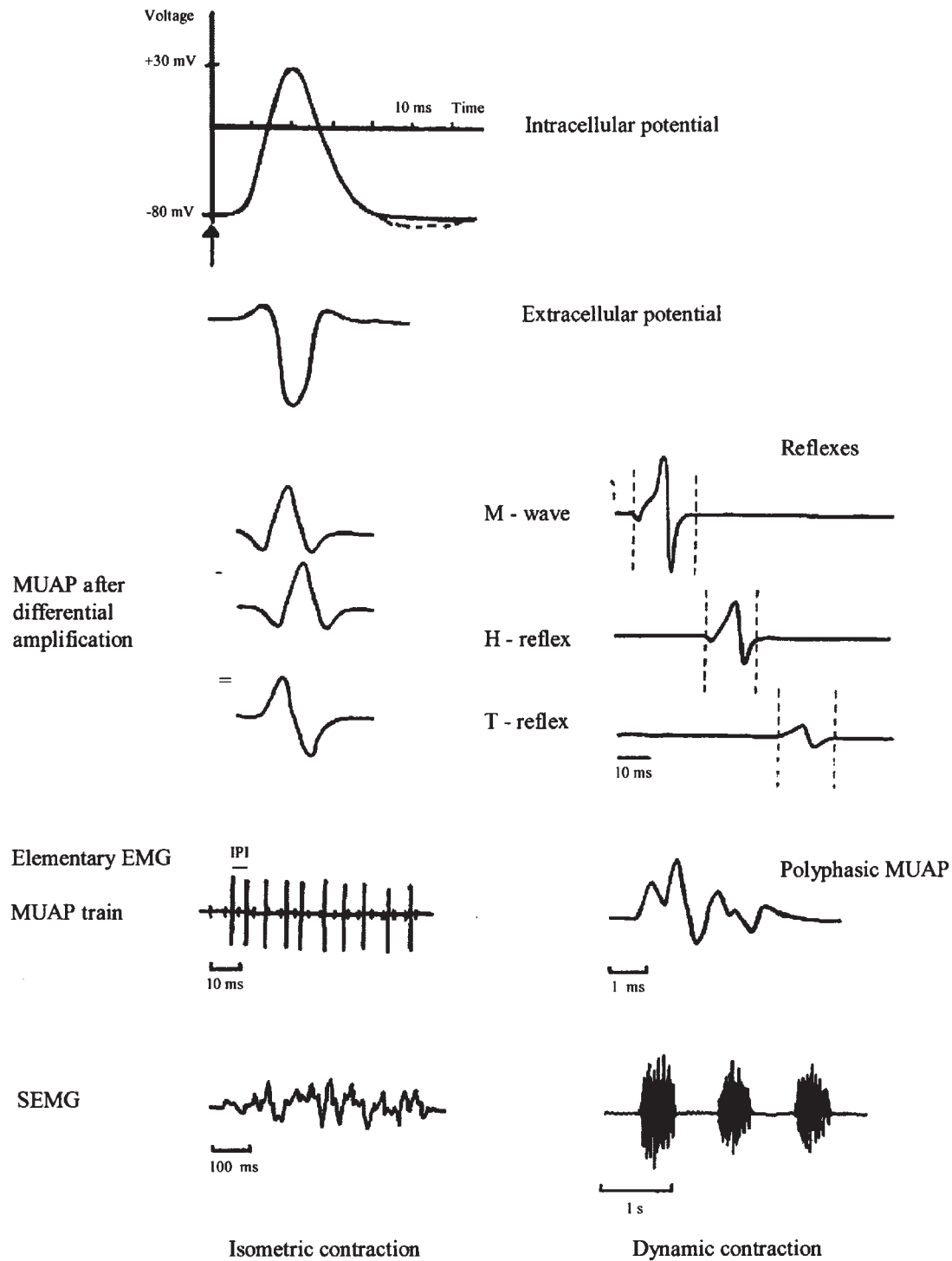


Fig. 2 Typical EMG signals (redrawn from personal results). MUAP, motor unit action potential; SEMG, surface electromyogram; IPI, interpotential interval.

delimited by the most distant fibers. The spatial scattering of fibers is regarded as a uniform or a Gaussian^[4] distribution. In the muscle cross-section, MU territories overlap each other. The rate of innervation of a MU is defined as the number of muscle fibers belonging to it. Muscles that must provide a coarse

effort (for example, in limb muscles) are constituted by MU with a large rate of innervation, whereas muscles that need to ensure a precious and discriminative activity (for example, face muscles) have MUs with a larger number of fibers, but each has a low rate of innervation.

The entire set of MU axons corresponds to the motor component of the mixed nerve that innervates the muscle. The other nervous component is sensitive and drives sensitive information from muscle spindles. The nervous endings on the muscle are a privileged area (identifiable by cutaneous electric stimulation) called the innervation zone, or motor point (MP). The MPs are located at different places on the muscle. Tables of MP localization were constructed previously.^[5]

Type of MU

The type of MU is related to the type of fiber that composes it.^[2] Several types of muscle fibers exist and are classified into two main categories: slow fibers (type I) and rapid fibers (type II). The fast fibers are divided into three groups, IIa, IIb, and IIc. The various fiber classes are determined from criteria or characteristics dependent on their structural, contractile, metabolic, and functional properties. For example, fiber diameter, AP amplitude, and generated force are larger in type II fibers, whereas the excitability threshold, fatigue resistance, and time of twitch contraction are larger for type I fibers. In one MU, the type of muscle fibers that composes it is the same. Categories of MU type are determined based on fiber type, such as slow or fast MU.

In a given muscle, the proportion of different types of MU assigns the properties of the muscle: slow muscle (98% slow MU; for example, soleus muscle), fast muscle (90% fast MU; for example, extraocular muscle), and mixed muscles (slow and fast MU; for example, biceps brachii muscle). During a progressive increase in the intensity of muscle contraction, Henneman^[6] determined that the larger the cell body of the motoneuron, the higher the threshold of excitability will be and consequently the motoneuron will be activated late; this result is called the size principle. Then, consequently, slow MUs are recruited before the fast MUs.

GENESIS OF THE EMG SIGNAL

Muscle Fiber Potential

In electromyography, independent of the method used, electrical activity detection is realized extracellularly. The changes in detected potential are collected outside the cell. The relation between intracellular and extracellular AP is a second derivative with respect to time.^[7] The activity of just one cell can be picked up (Fig. 1B) and represents the muscle fiber action potential (MFAP). The propagation velocity of the

MFAP along the muscle fiber is about 2 to 5 m/sec. The propagation velocity is smaller than that of axons, which is about 45 to 65 m/sec. The propagation velocity in the axons is increased because of saltatory conduction, which is conduction resulting from the presence of a sheath of myelin (Fig. 1A) that surrounds and electrically insulates the axon in small sections. Each myelin section corresponds to a Schwann cell in the case of peripheral nerves.

Motor Unit Action Potential

When a motoneuron is activated, an AP is generated and propagated along the axon and its collaterals. These AP are transmitted by NMJ to muscle fibers and then each muscle fiber generates MFAP. The motor unit action potential (MUAP) is the algebraic sum of MFAP (Fig. 1B). These potentials, detected using intramuscular electrodes, have amplitudes and shapes that depend on the following parameters: electrode–fiber distance, fiber conduction velocity, density of fibers at the site of detection, and length of collateral.

Elementary EMG

Elementary EMG is an intramuscular EMG constituted from algebraic summation of a restricted number of MUAP (from 1 to 4–5 MUs; Fig. 2). If a needle electrode with one wire is used, MFAP may be detected.^[8]

Surface EMG

Surface EMG (SEMG) is a global EMG because a large number of MUAP are detected (Fig. 2). Surface EMG is picked up on the surface of the skin (Fig. 1B) and thus the electromyographic evaluation is reduced to the most external muscles. The potentials then cross various media such as the fat layer (via adipocytes) and the skin layers (via keratocytes). The detected EMG signals are altered in amplitude and frequency by the cell layers. Signals with the greatest amplitudes come from the muscle fibers closest to the electrodes.

EMG Characteristics

Electromyographic signals have different characteristics according to the modalities used to generate them. The two kinds of modalities are stimulation and voluntary contraction performed by the subject. Using electric stimulation (applied to the motor nerve or the muscle itself) or mechanical stimulation (impact on one of the muscle tendons by a hammer), EMG signal is deterministic. At the occurrence of stimulation or

excitation, activation of each MU is synchronous, i.e., all are active at the same time. Using electric stimulation (adjustable in intensity and duration), the size principle is reversed: type II MU are initially excited. During voluntary activity, the EMG signal presents characteristics of a random signal; the activities of MUs are asynchronous. The EMG signal in bipolar detection and collected on the skin is called interferential EMG and is centered on electrical zero (baseline). In some cases of muscle contraction, such as fast or explosive contractions, MUs tend to have synchronous activity.

EMG SIGNAL PROCESSING CHAIN

A typical EMG signal measurement chain^[9-11] consists of three basic units: detection, signal conditioning, and signal acquisition and processing. Because an EMG signal is extremely weak, ranging from microvolts to millivolts, it is easily contaminated by undesirable noise. Thus, the higher the performance of each unit, the better the quality of EMG in terms of signal-to-noise ratio. The Surface Electromyography Not Invasive for the Assessment of Muscles (SENIAM) research group,^[5] supported by the European Union in the biomedical research program (BIOMEDII) proposed in 1999, made recommendations based on practical methods for the use of SEMG in clinical diagnostics. These recommendations help users standardize measurement methods, analysis, and modeling of SEMG.

Detection

The detection of EMG is performed by specific sensor electrodes. These electrodes, which come in direct contact with the muscle or skin, are made of different metals that must be non-toxic for the subject and do not polarize with an electrical current across the electrode (unpolarizable electrode). The interface, or the contact surface, between a metal and the biological environment produces long-lasting polarization effects. In order to minimize polarization, the metal must be covered with a layer of salt of the same metal. Generally, silver chloride (Ag/AgCl) is used for sensor electrodes. The most important characteristics of the interface of the electrode–biological medium are their electric impedance (in ohms) and their filtering effect (in particular the low-pass filter) on the detected signal. The EMG signal depends on these electrical characteristics but also on the electrode position. The detected electrical activity is a signal under high impedance.

This high electric impedance is lowered using an impedance transformer (preamplification circuit).

Three electrodes must be used: two electrodes (active electrodes) are connected to the high impedance inputs (one input for each electrode). The third electrode (the reference electrode) is connected to the low-input impedance (placed in a region without electrophysiological activity that is linked to the ground).

Special care must be taken to realize EMG detection; the amplifier requires electric insulation in order to protect the subject electrically (subject safety) because devices in the acquisition system are linked to the AC power line.

Two modes of detection (monopolar and bipolar) are usually used. The monopolar mode involves only one active electrode and gives information regarding changes in potential under the detection place. The second electrode must be placed on an electrically inactive zone (wrist, ankle, or ear lobe). In the bipolar mode, which is most generally used, two electrodes are placed with a gap between them. Because of the difference in potential, the undesirable remote potentials (parasitic or noise) are eliminated, thus improving the signal-to-noise ratio. A bipolar EMG signal obtained is very different in shape from the monopolar EMG signal; this must be taken into account in the conditioning and treatment of the signal. The bipolar EMG is dependent on the interelectrode distance. A pair of electrodes will detect more or less remote electric activity with a large or small interelectrode distance, respectively. Then active muscle volume (or conductor volume) will be larger or smaller depending on the interelectrode distance. When the interelectrode distance is small, the conducting volume is low and the electrodes have high selectivity. The distance between electrodes must be chosen in such a way that the conductor volume receives no activity from neighboring muscles (concept of “cross-talk”).

KINDS OF ELECTRODES

Two kinds of electrodes are used (Fig. 3): invasive electrodes cross the skin (intramuscular electrodes) and noninvasive electrodes are attached to the skin surface (surface electrodes). Invasive electrodes have electrode needles with one or more Ag wire of detection (coaxial, multiple) and wire electrodes (with or without hooks) allowing detection of the electrical activity during motion with low amplitude and low speed. The first electrodes are mostly for neurophysiological evaluation, whereas the second electrodes are generally dedicated to biomedical research. Non-invasive electrodes, or cutaneous surface electrodes, are widespread in the fields of kinesiology and sport. They consist of a round metal plate a few millimeters in diameter or a rectangular plate of Ag/AgCl (about 2×10 mm). Moreover, special kinds of electrodes can be constructed in relation to the

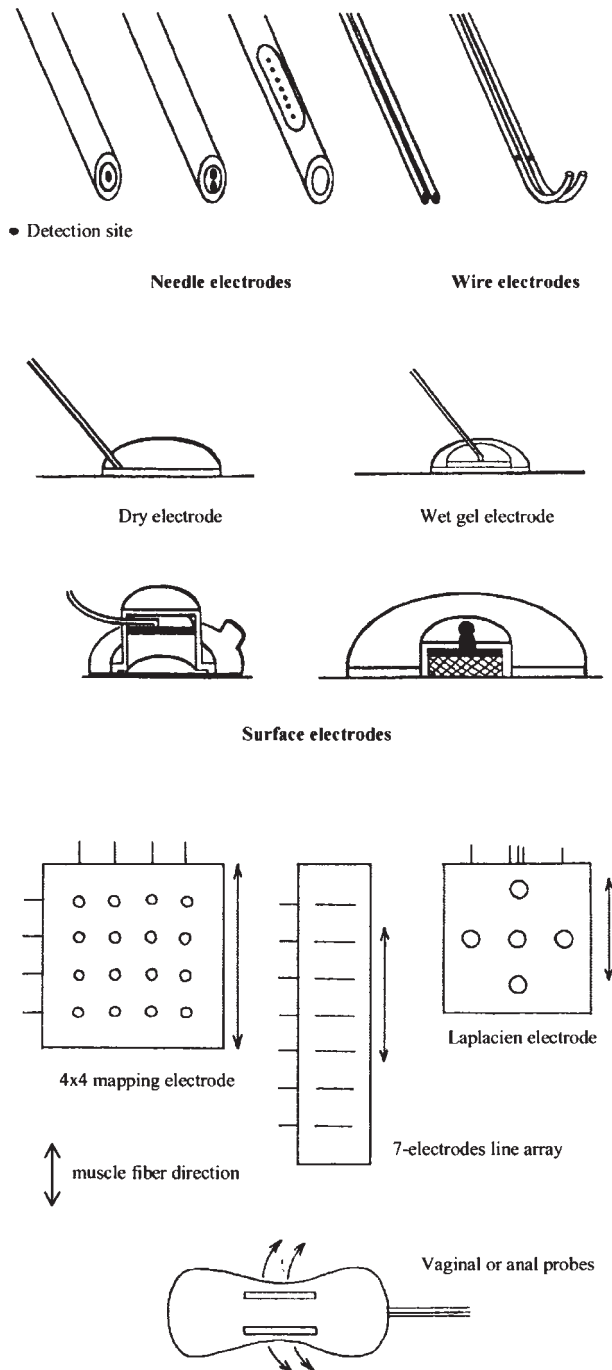


Fig. 3 Different kinds of electrodes (intramuscular, surface).

geometry and position of the muscle (for example, for vaginal and anal probes). In order to ensure the best contact between the skin and the electrode, it is necessary to use a layer of conductive paste. Wet gel electrodes are used when the shape of the electrode makes it possible to maintain the conducting layer. If the detection surface is directly on the skin with only a very thin layer of conductive paste to ensure contact, the

electrodes are called dry electrodes. The advantage of wet gel electrodes compared with dry electrodes is their insensitivity to the movements of wire-linked electrodes to amplifiers during the evaluation.

Electrode Placement in Surface EMG

Skin preparation

When surface electrodes are used, contact impedance with the skin is higher because the skin may be fatty and dirty, have a large cutaneous keratinization, or be senile skin. In such cases, before electrode placement it is necessary to eliminate fatty impregnation of the skin, scaled cells, and abnormally high pilosity as much as possible. The reduction techniques for cutaneous impedance consist of shaving and then abrading the zone on which the electrodes will be fixed using an abrasive paste or an equivalent mixture of alcohol and acetone. Once the electrodes are correctly positioned, interelectrode resistance is measured. Interelectrode resistance must be lower than $5 \text{ k}\Omega$.

Electrode placement

The electrodes are placed in parallel to the orientation of muscle fibers and one of the two sides of the MP. The best placement is in the position the muscle will have during its contraction phase, which does not necessarily coincide with the resting position. The electrodes are fixed using double-face adhesive tape. For most muscles, surface electrodes of 10 mm in diameter and a distance of 20 mm intercenter is recommended by the SENIAM^[5], but for small muscles the size of the electrodes and their distance must be adapted.

Electrode Placement in Elementary Electromyography

The use of intramuscular electrodes is a medical act and requires initial asepsis of the electrode. After the skin is cleaned with alcohol, the skin is locally anesthetized (by spray cooling or local analgesic) before needle puncture.

Electromyographic Multidetector (Fig. 3)

In some elementary EMG uses, the needle contains several detection wires constituting a multielectrode bundle, which makes it possible to detect several channels of signals simultaneously. For SEMG, several electrodes can be arranged, either online with equivalent interelectrode distances or matrix built for mapping EMG investigations.

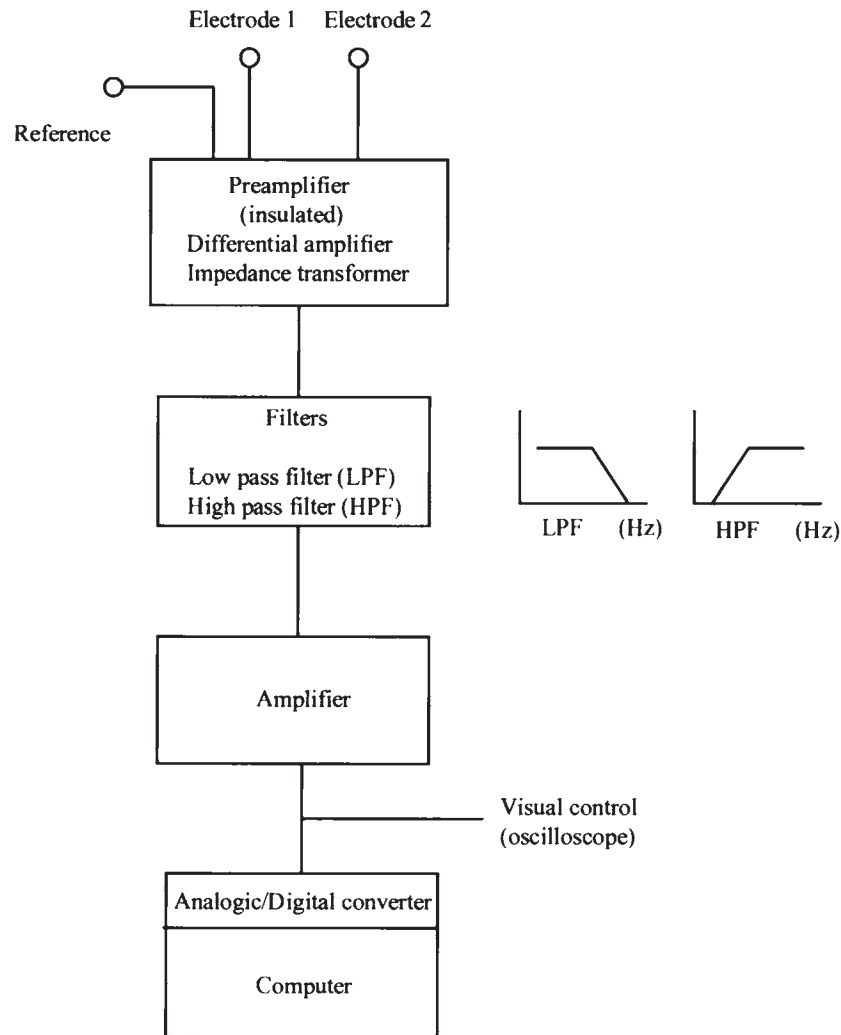


Fig. 4 A typical signal processing chain.

In these cases, the associated signal processing is specifically developed. As an example, an electrode device called a laplacian electrode made with five electrodes is used to perform space filtering. Laplacian detection enhances the depth of detection with high selectivity. A second example is n -electrodes in line with constant interelectrode distances in order to compute the velocity of propagation of muscle potential.

At the present time, the needle and cutaneous electrodes are manufactured industrially and ready for use. The disposable electrodes available for SEMG are wet gel or adhesive gel electrodes.

SIGNAL CONDITIONING AND DATA RECORDING (FIG. 4)

The primary purpose for analog signal conditioning circuitry is to modify sensor output into a form that can be optimally converted to a discrete time digital data stream by data acquisition. The conditioning circuit consists of three consecutive stages:

preamplification, filtering, and amplification. The accuracy of the data recording relies on the precision of these three units.^[9,10]

Preamplification

The preamplification stage boosts the scaled electrical signal received from the electrodes. The most important circuit configuration for amplifying sensor output is the instrumentation amplifier. The EMG preamplifiers are typically built around an instrumentation amplifier with the following requirements: accurate and stable weak gain (usually between 2 and 10), extremely high input impedance (about $101\text{M}\Omega$), extremely low output impedance (about 100 ohms), and an extremely high common mode rejection ration (more than 110 dB). Because signals received from the electrodes have high impedance and low amplitude, the preamplification stage operates as an impedance transformer. Usually, it is located as close as possible to the detection site.

Filtering

The filter section will comb out irrelevant frequencies, electrical noise, and other unwanted interference from the signal, leaving only the part of the converted bio-potential signal one wishes to measure and record. To reduce the undesirable signals (noises) as best as possible, the filtering unit uses both low- and high-pass filters. Noise has an extrinsic origin (electromagnetic disturbances of 50 or 60 Hz, engine in action, wire movements) or an intrinsic origin (other biopotential signals, such as ECG, the amplitude of which is larger than that of an EMG). The high-pass filter removes unbalanced components of half-cell potentials and movement artifacts of the cables, i.e., spurious changes in the signal resulting from movements of the leads and slipping during gesticulation or motion. Low-pass filtering limits the spectral characteristic of the EMG signal to the known bandwidth so the signal can be sampled at the required rate. The filters are analogic devices and are tunable according to frequency bandwidth. A notch filter (which ideally attenuates the frequency for which the filter was built) is currently added to enhance the suppression of electromagnetic disturbance. The bandwidth of each filter is set according to the frequency band of the EMG signal, the value of which depends on the type of EMG (elementary or surface), on electrodes placement, and also on muscle size and type. Typical amplitudes and frequency bands are 1–10 mV and 5–500 Hz for surface EMG and 200–250 μ V and 30–3000 Hz for intramuscular EMG, respectively.

Amplification

Before the signal is sent to the last unit (visualization or processing devices), the EMG signal passes through a final amplification with a very high gain. By placing high-gain amplification at the filter output unit, selective increase in the desired signal is ensured without noise amplification because the majority of the electrical interference has already been filtered. The high gain significantly increases the amplitude of the EMG signal, making it easily detectable and manageable by visualization or numerical acquisition devices. To avoid distortion or saturation, the output signal must not exceed the specific range of the device (example of voltage range: ± 1.4 V, ± 5 V, ± 10 V).

The active sensor is a specific system use for SEMG detection, and includes a preamplification circuit and a filtering circuit placed nearest the two active electrodes. Thus it generates a low self-impedance signal that is ready to be amplified. In some cases, the signal must be brought far from the detection site and be transmitted by radiotelemetry.

Signal Insulation

For human safety, it is necessary to isolate the sensor from the power supply of the conditioning units. The EMG preamplification unit includes electric insulation in order to protect the subject. This occurs in one of two ways: magnetic isolation or optical isolation. Magnetic isolation is primarily used for coupling power from the computer or the wall outlet to the sensor. This is performed through the use of a transformer. Optical isolation is used for coupling the sensor signal to the data acquisition input. This is usually performed through the use of a light-emitting diode and a photodetector.

Recording

Some devices are used for visualization or acquisition of the conditioned EMG. Using an oscilloscope (analog or digital), the amplitude in time variation of the EMG signal can be investigated.

In order to permanently record the signal, tape or graphic recorders are used to produce a trace paper. Tape recorders, which are less used nowadays, make it possible to reread the recorded signal. Graphic recorders using ink jet or thermic paper are commonly used, as for ECG or EEG. In choosing a recorder, particular care must be given to the value of the recorder's bandwidth because it is dependent on technical considerations. During an examination or a test using elementary EMG, it is common to listen to EMG signals via a loudspeaker. The EMG is transformed into a signal compatible with that of loudspeakers. Hearing the activity allows the investigator to have a "third hand": it enables him to position the needle without having to see the signal.

All contemporary analyses and applications of EMG signals are accomplished using algorithms implemented on computers, and necessitates that the signals be expressed as numerical sequences. The process by which the detected signals are converted into these numerical sequences is called analog-to-digital conversion. Analog signals are voltage signals that are analogous to the physical signal they represent. The amplitude of these signals typically varies continuously throughout their range. The analog-to-digital conversion process generates a sequence of numbers. Each number represents the amplitude of the analog signal at a specific point in time. The resulting number sequence is called a digital signal, and the analog signal is said to be sampled. Numerical recording for upcoming EMG signal processing requires devices with an analog-to-digital converter (ADC). Their use is highly varied and widespread. Each application must be

assessed with consideration of the advantages and limitations of the specified ADC.

Knowledge of the signal ranges is strongly recommended—both amplitude and frequency—in order to match the dynamic of the signal as well as the sampling rate as much as possible. It is advisable to sufficiently but not excessively (to avoid an overload) amplify the signal with regard with the voltage range of the AD boards (typical ranges for ADCs are ± 1.25 , ± 2.5 , ± 5 , and ± 10 V) and to choose a sampling rate at least twice the maximum frequency of the signal (according to Shannon's theorem).

DATA PROCESSING—EMG QUANTIFICATION

Data processing and computation of EMG depend on the type of electrode, the type of detection, and the type of study that justifies the use of electromyography.^[9,11–14] Data processing is applied to numerical signals, which are then sampled. Pretreatment may be carried out to enhance the signal-to-noise ratio. For example, a numerical filtering may be applied to the signal in the temporal domain (e.g., filter of Butterworth or Tchebychev) or in the spectral domain to eliminate one frequency, where after the frequency component is reset (zero value) the signal is rebuilt. Another technique is signal averaging, and requires three conditions. The first is the existence of a trigger event, which will correlate with the signal; the second is that the time between the trigger event and the signal is invariable; and the third condition is that it is possible to repeat several times the generation of the signal without changes in the two first conditions. By summation of sampling epochs, only the signal generated at the same time will appear and thus its amplitude will increase, whereas the existing signals not generated by the starting signal will decrease in amplitude. This method, often used in EEG to extract evoked potentials, makes it possible to extract a signal from a composite signal. For example, it is possible to extract the force generated by one MU from the force signal during contraction; the trigger event is the MU potential itself.^[15] Finally, the use of adaptive filtering eliminates undesirable signal but the signal must be recorded independently. For example, if the SEMG of back muscle is contaminated by an ECG signal, to delete the ECG in the SEMG it is necessary to record the subject's ECG on an additional channel in order to apply adaptive filtering.

After pretreatment, the signals can be quantified using numerical parameters. In the case of elementary EMG, MUAP are analyzed individually or in the entire activity during a defined duration. Individually, the shape of the MUAP is characterized by this duration and number of phases, amplitude, and area. In a defined time interval, the firing rate of the potentials, the

mean interpotential interval (IPI), the instantaneous frequencies (inverse of IPI), and the jitter (mean variation in IPI between single-fiber action potential of the same MU) will be studied.

In the case of SEMG, the signal may be studied in the temporal domain (according to time) and in the spectral domain (according to the frequencies which it contains). In the temporal domain, common processing involves wave rectification (absolute value of the signal, knowing the signal is centered on a level mean, generally zero), counting the number of passages to zero, integration (calculation of the signal area with respect to the baseline), and calculation of the signal envelope (the total shape of the signal). Moreover, the envelope may be estimated using a moving average of rectified EMG (average of the amplitude in the moving time window).

In the spectral domain, characterization is carried out on the power density spectrum of the signal. Using a fast calculation algorithm (fast Fourier transform or FFT), the temporal signal is transformed into the spectral domain. Using a calculation associated with FFT, the signal is then represented as a power density spectrum. The signal is graphically represented by a continuous curve (power density) according to frequencies the signal contains. The power density spectrum is characterized by quantifiers such as the root mean square (RMS), corresponding to total power of the signal computed as the spectrum surface, the mean power frequency (frequency corresponding to the barycenter of all the frequencies according to their power), and the median frequency (the frequency that divides the power spectrum into two equal RMS parts). Using the most common parameters, other parameters such as the symmetry of the power density spectrum or flatness with respect to a normal Gaussian curve can be calculated.

If, by calculation of these parameters, it is possible to quantify the EMG signal, interpretation of the results and their changes during the time course must take into account the test set-up in which the signal is generated. For example, according to the level of contraction (maximal or submaximal) or during muscle fatigue the parameters may be different, thus leading to different interpretations.

Specific Computation of EMG Signal

Data computation^[12–14,16] can be used to improve the detected EMG signal or to seek particular information regarding the signal using specific algorithms. For example, new algorithms in filtering methods are elaborated upon to improve the preprocessing step, such as the removal of artifacts in the electrical signal. The methods most used to extract information from the

EMG signal are decomposition of the EMG signal into MUAP trains and classification of MUAP parameters using wavelet transform or neural network methods. Following the application of extraction methods, data are analyzed using statistical analysis (such as mean, variance, or correlation analysis to evaluate the relevance of a given parameter) or multivariate methods (such as principal component analysis or discriminant analysis to classify data and to take a decision from many parameters).

Graphical Representation of EMG Signals

Other characterizations of the EMG signal, such as graphic modes of representation, provide information on the shape, duration, and intensity of the electric activity of the muscle during contraction.

In the monodimensional mode (1-D), observation of the signal (one channel) as a function of time provides information regarding the signal amplitude. However, another representation called graphical superposition of signal epochs provides information regarding shape changes. Graphical superposition is made from a characteristic in each signal (such as trigger event or maximal amplitude). For example, by triggering the visualization of multiphasic MUAP (Fig. 2) in the first phase, it is possible to identify changes in the other phases. The jitter is defined as the delay in propagation on the branches of collateral of the motor nerve. In the same way, the superposition of power density spectra may identify spectral changes in EMG, such as in the case of muscle fatigue processes.

In the two-dimensional mode (2-D), according to the axes (abscissa and ordinate) the parameters used may be homogeneous or heterogeneous. For example, the time (abscissa)–frequency (ordinate) representation called a periodogram demonstrates power spectra for each segment of the raw signal. The possible spectrum change appears upon visual examination of the graph. When several channels of EMG are recorded, one image can be built pixel by pixel using an interpolation between the signals and by transforming the amplitude to color according to a color chart. This corresponds to a 2-D representation or a pseudo 3-D representation. The most current application concerns the EMG mapping (position of electrode in abscissa and in ordinate), or a pixel has a determined color (amplitude), which illustrates the distribution of EMG activity on the surface of the skin. More recently, by using an n-electrode line array (electrode positioned in the abscissa and time in the ordinate with color as amplitude), it is possible to estimate the propagation velocity of electric activity by computing the slope of the colored lines, which appear on the image.

PRINCIPAL FACTORS LEADING TO EMG SIGNAL CHANGE

The principal factors implied in EMG signal changes for a healthy subject^[13,14,16] are as follows:

1. Methodological factors such as those linked to the sensors (chemical properties of the detection surface, size and shape of the electrodes) and sensor placement (orientation of the electrodes compared with the direction of muscle fibers, sensor placement with respect to the motor point, distance and impedance between electrodes);
2. Structural physiological factors such as the space distribution of MUs, the diameter of fibers constituting the MU, and the filter properties of muscle tissue and skin; and
3. Functional physiological factors such as the kind and the level of contraction, the degree of synchronization in MU behavior, muscle metabolism (ionic disorder: pH, balance in release–production of lactate), intramuscular pressure, blood flow, and muscle temperature.

In addition to these factors, it is necessary to consider the age of the subject and the various neuromuscular pathologies^[17,18]

1. Structural factors such as diseases of the central nervous pathways (e.g., brain motor area) or peripheral diseases (neuropathies), structural changes in the neuromuscular junction (such as myasthenia gravis) and in muscle fibers (myopathies); and
2. Functional factors such as those associated with diabetes and obesity or muscle fiber function such as McArdle's syndrome (abnormal glycogen accumulation in muscle tissue).

Finally, electromyography characterizes specific changes in the neuromuscular system that have led to a specialization of this system as this is identified after endurance or pliometric training, after reduced muscle activity (limb immobilization), and during exposition in extreme environments such as space flights of long duration. These neuromuscular changes, which are generally reversible, illustrate the neuromuscular plasticity.

ASSOCIATION OF EMG SIGNALS TO OTHER ELECTROPHYSIOLOGICAL SIGNALS

The study of EMG signal and its changes is dependent on the conditions of signal generation. The EMG signal is evoked from a muscle activity (muscle contraction), which requires voluntary motor control or involuntary solicitation (external stimulation, e.g., the

electromyographical component in the study of reflexes).

Using experimental devices called ergometers, it is possible to normalize or standardize muscle contractions according to a fixed protocol. The detected EMG signal is associated with other signals (such as force, power, position, and speed for mechanical signals) or related to the stimulation parameters (e.g., intensity and duration for electric stimulations).

Two kinds of contractions exist: isometric contractions (the muscle does not modify its length and thus a force torque is measured) and dynamic or anisometric contractions (the muscle works with variable length). The interest in using mechanical devices or ergometers that are more or less complex lies in seeking the most reliable, reproducible, and even standard experimental conditions. Moreover, specific protocols such as the research of maximum (e.g. maximal voluntary force, maximal RMS) are generally proposed and leads to standardization of the instructions given to the subject or to normalization of the calculated parameters values to this reference value. Under test conditions, it is possible to identify changes in the EMG signal in comparison to standard signals or characteristic signals obtained from healthy subjects.

According to the properties of the SEMG signal, the signal is not reproducible but the quantifier values and evolution are (e.g., total energy, mean frequency during muscle fatigue) if test conditions are repeatable. Changes in EMG signal are then perceived by its parameters rather than the signal waveform variations.

EXAMPLES OF PHYSIOLOGICAL CHANGES IN EMG SIGNAL

Level of Contraction

Under isometric conditions of contraction, when a subject increases the intensity of a voluntary contraction in arm flexion, the SEMG of the biceps brachii muscle increases gradually in amplitude, as characterized by an increase in the EMG RMS. The muscle activity is dependent on the number of recruited MU (by spatial recruitment according to the size principle) and on the activity of each MU by their firing rate (temporal recruitment). For various levels of contraction, the MU firing rates differ with respect to MU type; the firing rate of MU ranges from 4–5 to 25–30 Hz for slow MU and up to 60 Hz for fast MU. For small muscles, all the MU are recruited before the intensity of contraction reaches 50% of the maximal voluntary force, whereas for large muscles the entire MU recruitment is obtained when contraction levels range between 70 and 80%.

Muscle Fatigue

Muscle fatigue is a reversible physiological phenomenon that is studied during isometric contractions (with maximal or submaximal level). During the submaximal level, which is maintained until exhaustion, the SEMG is altered and RMS and mean power frequency changes are observed. The increase in RMS is associated with a decrease in the mean power frequency of the signal. The interpretation is that increased MU activity is required to maintain the force and mean power frequency changes are associated with biochemical disorders such as muscle acidification.

In maximal contraction, the number of active MU cannot be increased and the muscle activity is close to the maximal activity. Fatigue is identified by a fall in mean power frequency and force produced.

Body Motion in Cycling, Walking, and Jumping

Using of an ergometer (for cycling), EMG signals (Fig. 2) can be detected with multiple channels and demonstrate patterns of activity that are quantified during body motion. For example, the beginning and end of activities may be quantified in duration and in magnitude with signal RMS. Each motion, such as walking, jumping, or exercising on a bicycle, is associated with mechanical variables (such as position, speed, and acceleration). Interrelations between the EMG quantifiers and the mechanical parameters are studied and may describe anomalies in motor coordination.

CLINICAL ELECTROMYOGRAPHY

Elementary EMG with needles is currently used in clinical electromyography.

Neuromuscular evaluation^[17,18] using electromyography corresponds to the evaluation of the most distal organ (muscle) and, consequently, leads to the identification of signs of suspicious pathologies at various places from the cerebral motor area to the muscle (localization of the excitation–contraction coupling). Electromyography is used to assess the effect of medicinal therapy or to follow up with rehabilitation (e.g., in the case of muscle reinnervation: the MU territories decrease, whereas the rate of innervation increases).

In practice, an EMG signal is generated following a verbal instruction to the subject and signal anomalies with respect to reference signals may be identified. Another manner of generating EMG signal consists of evoking EMG responses following stimulations that are electric (electrodiagnostic) or mechanical (tendinous stimulation or disruption of posture). The

method, which is most often used in the study of reflexes, is called reflexology.

For example, electrostimulation of the sciatic nerve applied to the hollow poplity (area located behind the knee) generates various muscle responses (Fig. 2) according to the intensity of stimulation. With low intensities, the sensitive nervous fibers (Ia fibers from muscle spindles) are the only ones to be excited and by the monosynaptic loop reflex of the homonymous muscle will produce the slow type of MU contraction. The EMG response is called Hoffman's reflex (H reflex). With stronger stimulation intensities, the H reflex gradually decreases and the direct response of the muscle, called the M wave, appears. The evoked EMG response to abrupt percussion on the Achilles' tendon is called tendinous reflex (T reflex). The T reflex occurs later than the H reflex because of the coupling between tendinous mechanical stimulation and the response of the spindles.

CONCLUSION

Electromyography, or the study of EMG signal, provides information on nerve and muscle functionality. First, after drawing up a protocol and deciding whether an ergometer should be used, special care must be given to the choice of muscle investigated, kinds of electrodes, and their placement. Second, when the EMG signal processing chain is constructed, each unit of the chain must be adjusted according to the modality of the EMG signal generation (voluntary contraction or stimulation) and the signal kind (intramuscular or surface). In the third step, the kind of data analysis must be chosen to extract the researched information. In comparison with "normal" waveforms or changes, the EMG signal may reveal abnormalities in nerve or muscle function.

At the present time, new methods for EMG signal processing are in progress. The study of EMG is also associated with other means (for example, Nuclear Magnetic Resonance or NMR) used to investigate nerve and muscle function.

ARTICLE OF FURTHER INTEREST

Biopotential Amplifiers

REFERENCES

- Liddell, E.G.; Sherrington, C.S. Recruitment and some other factors of reflex inhibition. *Proc. R. Soc. London* **1925**, *Ser. B* 97 (686), 488–518.
- Burke, R.E. Motor units: Anatomy, physiology and functional organization. In *Handbook of Physiology: The Nervous System*; Brookhart, J.M., Mountcastle, V.B., Eds.; American Physiological Society: Bethesda, MD, 1981; Vol. 2, 345–422.
- Buchthal, F.; Erminio, F.; Rosenfalck, P. Motor unit territory in different human muscles. *Acta Physiol. Scand.* **1959**, *45*, 72–87.
- Miller-Larsson, A. A model of spatial distribution of muscle fibers of a motor unit in normal human limb muscles. *Electromyogr. Clin. Neurophysiol.* **1980**, *20*, 281–298.
- Hermens, H.J.; Freriks, B.; Merletti, R.; Hägg, G.; Stegeman, D.F.; Blok, J.; Rau, G.; Disselhorst-Klug, C. *SENIAM 8: European Recommendations for Surface Electromyography, Roessingh Research and Development*; Enschede, The Netherlands, 1999.
- Henneman, E. Relations between size of neurons and their susceptibility to discharge. *Science* **1957**, *126*, 1345–1346.
- Radicheva, N.; Gerilovsky, L.; Gydikov, A. Effect of short interstimulus intervals on the intra and extracellular action potentials of isolated frog muscle fibers. *Acta Physiol. Pharmacol. Bulgar.* **1986**, *12* (1), 26–35.
- Ekstedt, J. Human single muscle fiber action potentials. *Acta Physiol. Scand. Suppl.* **1964**, *226*, 1–96.
- Basmajian, J.V.; De Luca, C.J. *Muscles Alive: Their Functions Revealed by Electromyography*, 5th Ed.; Williams & Wilkins: Baltimore, 1985.
- Bretschneider, F.; de Wille, J.R. *Introduction to Electrophysiological Methods and Instrumentation*; Academic Press: London, 2006.
- Desmedt, J.E. Computer-aided electromyograph. In *Progress in Clinical Neurophysiology*; Desmedt, J.E., Ed.; Karger: Basel, 1983; Vol. 10.
- Weitkunat, R. Digital biosignal processing. In *Techniques in the Behavioral and Neural Sciences*; Huston, J.P., Ed.; Elsevier: Düsseldorf, 1991; Vol. 5.
- Duchene, J.; Goubel, F. Surface electromyogram during voluntary contraction: Processing tools and related to physiological events. *CRC Crit. Rev. Biomed. Eng.* **1993**, *21*, 313–397.
- Farina, D.; Merletti, R.; Enoka, R.M. The extraction of neural strategies from the surface EMG. *J. Appl. Physiol.* **2004**, *96* (4), 1486–1495.
- Milner-Brown, H.S.; Stein, R.B.; Yemm, R. The contractile properties of human motor units during voluntary isometric contractions. *J. Physiol.* **1973**, *228* (2), 285–306.
- De Luca, C.J. The use of surface electromyography in biomechanics. *J. Appl. Biomech.* **1997**, *13*, 135–163.
- Gilchrist, J.M.; Sachs, G.M. Electrodiagnostic studies in the management and prognosis of neuromuscular disorders. *Muscle Nerve.* **2004**, *29* (2), 165–190.
- Karpati, G.; Hilton-Jones, D.; Griggs, R.C. *Disorders of Voluntary Muscle*, 7th Ed.; Cambridge University Press: Cambridge, U.K., 2001.

Electron Microscopy

Robert L. Price

Jeff Davis

*Department of Cell and Developmental Biology and Anatomy, School of Medicine,
University of South Carolina, Columbia, South Carolina, U.S.A.*

Michael J. Yost

*Department of Surgery, School of Medicine, University of South Carolina, Columbia, South
Carolina, U.S.A.*

Richard L. Goodwin

*Department of Cell and Developmental Biology and Anatomy, School of Medicine,
University of South Carolina, Columbia, South Carolina, U.S.A.*

Louis Terracio

*Department of Basic Sciences, College of Dentistry, New York University, and Department
of Pediatrics, School of Medicine, New York University, New York, New York, U.S.A.*

INTRODUCTION

As technology has improved and we have moved toward the ability to engineer biomaterials, cells, and tissues, the necessity to image submicron details in a variety of sample types has also increased. Although light microscopy, confocal scanning laser microscopy, and other imaging modes—such as microcomputed tomography and X-ray imaging—can provide significant morphological information regarding samples, it is only through scanning (SEM) and transmission (TEM) electron microscopy that we can image the ultrastructural information required to adequately analyze information at the nanoscale level. As detailed in this article, biomaterials and bioengineering studies involving electron microscopy cover a wide range of applications, including the imaging of cell:cell and cell:extracellular matrix interactions, intracellular events, and nanoscale changes in biomaterials used as tissue substrates and in microspheres used in the delivery of targeted drugs.

INSTRUMENTATION

Although a detailed description of the design and operation of SEM and TEM instruments goes beyond the scope of this article, a basic description of the differences between the two types of microscopes and the types of studies they are used for is appropriate. The primary advantage of electron microscopes over other types of imaging sources is their improved resolution, which allows detail to be resolved routinely at the submicron to nanometer level. In some cases,

resolution at the angstrom level is also possible when the proper instrumentation and samples are used. This high resolution is based on the use of electrons as the imaging source. When compared to other types of imaging sources, such as photons that are used in regular light microscopes and laser-based imaging systems, electrons have very short wavelengths. When the formula that determines resolution is examined, the importance of the wavelength of the imaging source is evident.

$$\text{Resolution} = 0.612\lambda/n(\sin \alpha)$$

In this equation, 0.612 is a constant used in calculating the resolution of various systems; λ represents the wavelength of the illumination source; n represents the refractive index of the medium that the illumination source passes through; and $\sin \alpha$ represents the numerical aperture, or angle, of the lens. Typically, the wavelength of light used in optical systems is longer than 380 nm (blue light), whereas the wavelength of electrons is shorter than 0.005 nm. Although other factors, including aberrations in the lenses, may affect the overall resolution of a system, this fundamental difference in the wavelengths of the illuminating sources is responsible for the improved resolution of electron microscopes. This basic difference in the ability to resolve structures is illustrated in Fig. 1, which shows the structure of cardiac myocytes imaged by confocal scanning laser microscopy and transmission and scanning electron microscopy. For a more detailed discussion of the factors involved in the calculation of resolution, there are several good references available, including Bozzola and Russell.^[1]

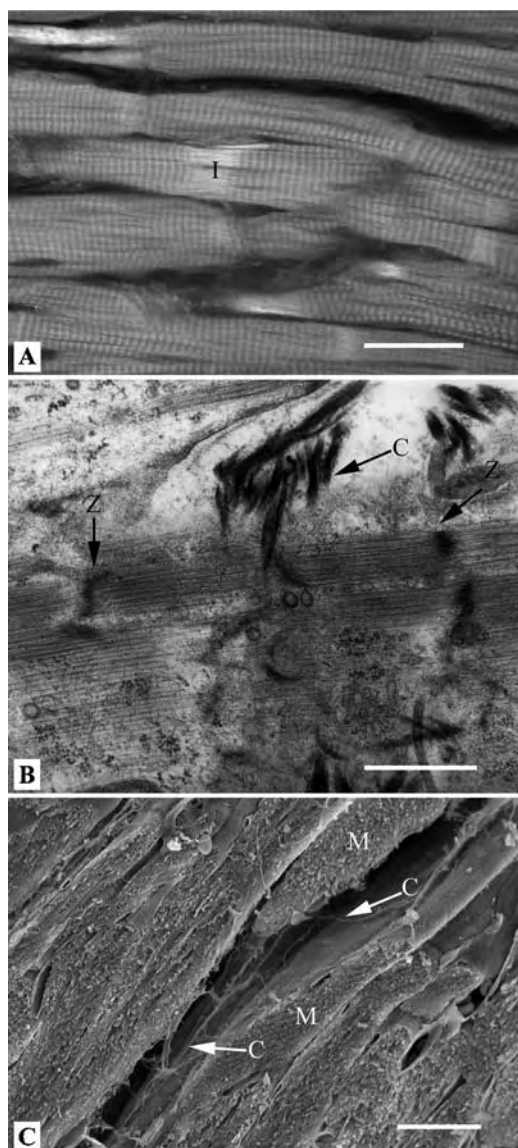


Fig. 1 Comparison of micrographs of cardiac myocytes collected by confocal scanning laser microscopy, TEM, and SEM, illustrating the differences in resolution and types of images collected by the three imaging systems. Part A shows an image of cardiac myocytes stained with phalloidin conjugated to rhodamine. Rhodamine phalloidin is a specific stain for f-actin, which is a primary component of the contractile apparatus of cardiac myocytes. Staining by this method allows imaging of intercalated disks (I) and the characteristic striations of individual sarcomeres in the myocytes. In this image, the bright white striations represent areas with actin and the dark bands represent the Z-disks where no actin is present. Part B illustrates the higher resolution obtainable by TEM, where individual Z-disks and the actin filaments between the Z-disks are easily resolved. It is also possible to determine the interaction of individual sarcomeres with components of the extracellular matrix, such as collagen (C). Part C is an SEM micrograph illustrating the surface imaging of myocytes (M) and the collagen fibrils (C) that interact with them. Scale bars: Part A, 200 μm ; Part B, 1 μm ; Part C, 20 μm .

BRIEF COMPARISON OF SCANNING AND TRANSMISSION ELECTRON MICROSCOPES

A basic diagram comparing the structure of typical SEM and TEM instruments is shown in Fig. 2. In both systems the illuminating source, or gun area, and the column of the microscope are maintained under high vacuum. This provides an unimpeded path for the electrons to pass through without striking molecules of air. The gun region consists of three components: an electron source, a bias shield, and an anode. The electron source is usually a tungsten or lanthanum hexaboride (LaB_6) filament. To generate electrons, a low-current, high-voltage power source is applied to heat the filament until the metal reaches its work function, or the point at which electron emission occurs. High-voltage (accelerating voltage) sources typically operate in the range of 50 to 100 kV, although instruments are available that operate at one million volts. As the accelerating voltage increases, the wavelength of the electrons decreases, providing a higher-resolution system. In addition, electrons have more energy as the accelerating voltage increases, allowing the imaging of thicker specimens.

Surrounding the filament is a bias shield (Wehnelt cylinder) that has a 2- to 3-mm aperture directly in front of the filament tip. The bias shield is maintained at a slightly negative potential with respect to the filament and the electrons emitted from it. This repels the electrons and directs them toward the aperture in the bias shield. Electrons exit through this aperture

Comparison of Transmission and Scanning Electron Microscopes

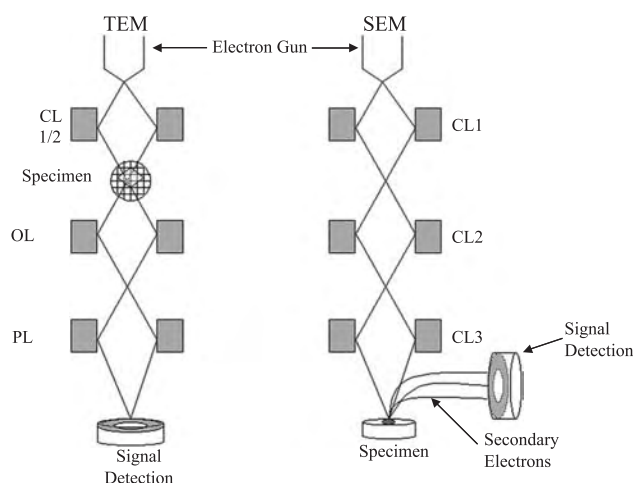


Fig. 2 Comparison of the structure and specimen placement in SEM and TEM instruments. A major difference in the two types of instruments is the placement of the specimen between the condenser (CL) and objective (OL) lenses of the TEM, and at the base of the column in the SEM. Abbreviations: CL, condenser lenses; OL, objective lenses; PL, projector lenses.

E

and are strongly drawn toward the gun anode, which has a positive charge with respect to the electrons. From the anode region the high-energy electrons enter the microscope column, which consists of a series of electromagnetic lenses that are used to control the path of the electrons. Controlling the path of the electrons makes it possible to control the magnification as well as the focus of the system, in order to create an image. Images can be recorded when electrons strike a sheet of film, or as digital media when electrons strike a phosphor-coated surface and are converted to a digital image.

The two types of instruments differ in placement of the specimen, which in the SEM is slightly below the final lens and in the TEM is in a plane between the final condenser and the objective lenses. This leads to the basic difference in the types of images that can be collected in the two types of instruments. In the SEM, the electron beam is condensed to a point of adjustable diameter that is rastered, or scanned, across the specimen. As the electrons interact with the molecular structure of the specimen, a series of interactions occur that result in primary, or backscattered, electrons that have lost little energy; secondary electrons that have interacted with the molecular structure of the specimen and have lower energy than those that were in the original electron beam; and a number of other types of signals, such as X rays. Most SEM studies involving biomaterials and biomedical engineering utilize secondary electrons as the imaging signal. Images collected from SEM instruments are primarily used to determine surface structure and the three-dimensional topography of a specimen that can be several centimeters in size.

For imaging in the TEM, most specimens must be sectioned or otherwise prepared so that electrons can

pass through the specimen to give an image of its internal ultrastructure. Because the goal of TEM is to image the interior of a specimen, instruments are operated at much higher accelerating voltages (50,000 to 1 million volts) than in the SEM (10,000 to 40,000 volts), and the resolution of images is much higher than that typically collected in an SEM. As detailed below, specimen preparation for the two types of microscopes also differs.

Recent advances in the design of electron microscopes include the development of field emission guns that provide a very small, bright source of electrons, which allows much brighter and higher-resolution imaging capabilities for both SEM and TEM instruments. Differential vacuum systems that allow for the examination of wet specimens in the specimen chambers of environmental SEMs have been developed, and the addition of cryostages in both types of instruments allow the elimination of much of the preparation procedure that may damage or change the molecular structure of a specimen. Although these modifications provide specific advantages in the imaging of biomaterials and biological specimens, the added cost of these enhanced instruments may be prohibitive for many applications.

ROUTINE PREPARATION OF BIOLOGICAL SPECIMENS

Specimen preparation varies a great deal depending on the type of specimen to be examined, the imaging mode that will be used, and the laboratory making the preparation. A generalized step-by-step protocol for the preparation of biological samples for SEM and TEM is presented in Fig. 3, but it must be remembered

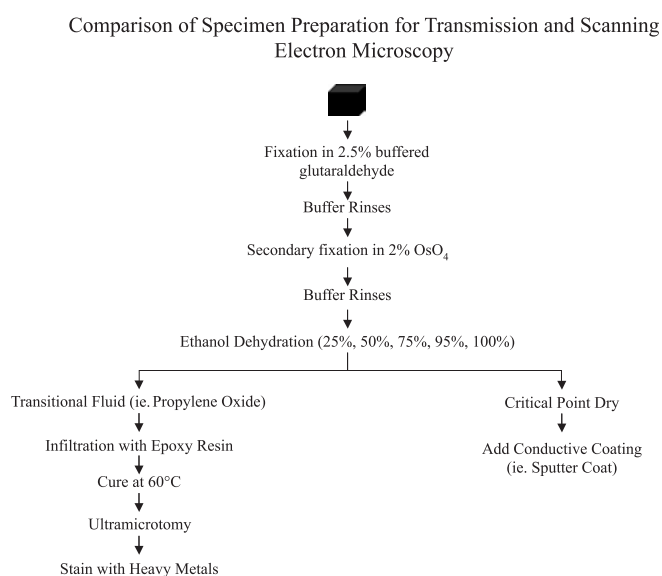


Fig. 3 Diagrammatic representation of the basic steps in specimen preparation for SEM and TEM.

that this protocol will not work equally well for all sample types. However, the routine processing of cells and soft tissues for both SEM and TEM follows a few general principles, or common variations of these principles, that are presented in the following text as a guideline for routine specimen preparation. A good rule of thumb is to search the literature to determine how others have successfully prepared specimens similar to those being examined.

The goal of fixation and subsequent specimen preparation is to rapidly stop and preserve ongoing cellular function and to maintain the three-dimensional structure of the specimen. Rapid fixation in 2 to 3% buffered glutaraldehyde is required for good ultrastructural preservation. Autolytic changes in cells and tissues begin to occur within seconds following death. For deep tissues such as the brain, it may be necessary to perform a cardiovascular perfusion technique to adequately deliver the fixative to the tissue prior to the onset of membrane disruption and the swelling of mitochondria. Following fixation, samples should undergo a series of buffer rinses to remove free aldehyde groups and a secondary fixation in 1 to 2% osmium tetroxide (OsO_4), which is also followed by a buffer rinse. Specimens are then dehydrated through a graded ethanol series. Care should be taken not to extend dehydration in concentrations above 70%, as this may result in extensive shrinkage and disruption of the three-dimensional architecture of the sample. As shown in Fig. 3 and in the following discussion, preparation methods for biological tissues for the two imaging modes diverge following the dehydration steps.

ROUTINE PREPARATION OF BIOLOGICAL SPECIMENS FOR SEM

Imaging of biological SEM samples in a standard microscope column requires that the specimen be completely dried and coated with a conductive substance, typically carbon or gold, prior to examination. Following dehydration in the graded ethanol series, the most common specimen drying method is critical-point drying, during which the ethanol is exchanged with liquid CO_2 at low temperatures. The CO_2 is then heated in a sealed chamber until it reaches its critical temperature and pressure, at which point it rapidly changes from a liquid to a gas, resulting in minimal shrinkage and damage to the specimen. Instruments for automated critical-point drying are available from several companies.

Dried specimens need to be coated with a conductive substance that will dissipate the charge accumulated in the specimen during interaction with the electron beam in the SEM. If the charge is not

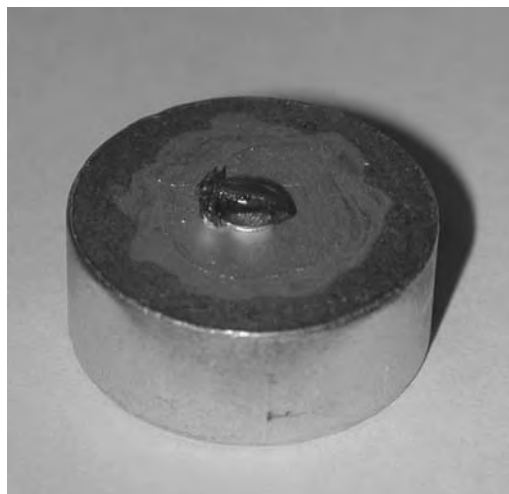


Fig. 4 Prepared SEM specimen (section of heart) that has been critical-point dried, sputter-coated with a thin layer of gold, and mounted on a 24-mm aluminum stub for imaging.

dissipated it will cause very bright areas in the image, which are usually found along the edges of a specimen or associated with thin cellular processes such as cilia and filopodia. Prevention of charging is usually accomplished by the process of sputter-coating a 20- to 30-nm layer of gold or palladium onto the surface of the specimen. Alternatively, for high-resolution studies a layer of these metals or carbon can be evaporated onto the surface of the specimen. Instrumentation for these processes is also available from several companies that manufacture equipment for electron microscopy. Figure 4 shows a prepared specimen ready to be examined by SEM.

VARIATIONS IN THE PREPARATION OF BIOLOGICAL SPECIMENS FOR SEM

As might be expected, the drying and coating of samples often introduces shrinkage artifacts in the three-dimensional structure of samples, which may result in the fracture of membranes and fine cellular processes (Fig. 5). Several modifications of protocols and instrumentation have been introduced to help minimize these effects. A common protocol involves the further stabilization of the cytoskeleton by the addition of repetitive treatments of glutaraldehyde, tannic acid, and OsO_4 , which maintains the structure of samples during drying and renders them conductive without further coating with heavy metals or carbon.^[2] Although this method renders specimens relatively stable and minimizes apparent damage, the repetitive addition of tannic acid and OsO_4 to small cytoskeletal structures may increase the apparent size of cell

components such as microfilaments and extended cell processes. An alternative method, although limited to relatively low magnification and low-resolution work is to examine specimens that have not been coated. Since these samples are not conductive, imaging is often difficult because of charging effects, but the fracture of membranes and cellular processes is minimal and the examination of cellular interaction with the surrounding extracellular environment is often enhanced.

Specialized equipment, including the addition of cryostages and low-vacuum environmental chambers to microscopes, has also been designed to minimize the introduction of drying artifacts in specimens. In experiments that utilize cryo-SEM, specimens may be fixed in glutaraldehyde prior to freezing, but in many cases this is not essential. Freezing is typically

accomplished by high-pressure freezing in liquid nitrogen or by plunge freezing in liquid ethane at -183°C .^[3,4] Specimens are then coated under cryo-conditions and examined in microscopes equipped with a cryostage. This procedure eliminates artifacts induced by the dehydration of samples, but it requires relatively expensive modifications of the microscope.

Alternatively, wet, uncoated samples can be examined in instruments equipped with an environmental stage, by controlling the pressure in the immediate area surrounding the specimen. With biological specimens it is often difficult to reach an equilibrium between a hydrated state and the hydrated condition that results from the low vacuum maintained in the instrument. In addition, the potential charging that results from examining an uncoated specimen often requires an experienced operator to acquire good images. However, these instruments do provide the opportunity to examine specimens that do not have artifacts introduced by the drying and coating processes, and they also allow imaging of dynamic processes such as crystal formation and the microstructure of biomaterials.^[5,6]

PREPARATION OF HARD BIOMATERIALS AND POLYMERS FOR SEM

Stable polymers and hard substances, such as apatite crystals, used as substrates or scaffolds for tissue growth generally require little specimen preparation for SEM, other than a light coating of gold or carbon if

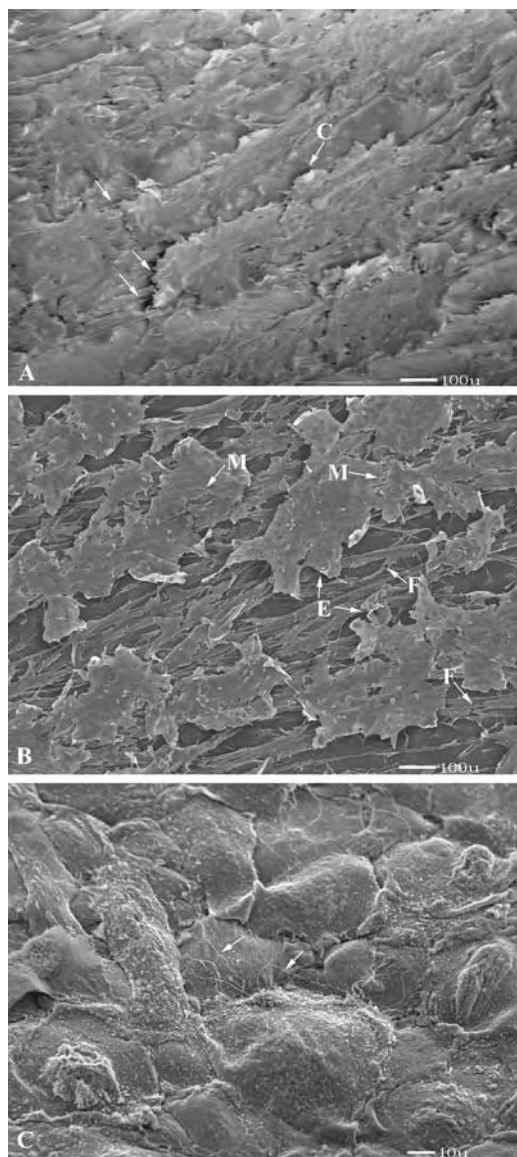


Fig. 5. Comparison of specimen damage and images collected by SEM following three methods of preparation. Part A represents cardiac myocytes that have been prepared as diagrammed in Fig. 3, except they have not been coated with gold. This prevents much of the specimen damage created by sputter-coating, but it requires that the samples be imaged by low-accelerating voltages, which limits resolution. Charging (C) is often a problem when small fibrils and contiguous cells are present. Part B is from an adjacent area of the culture dish seen in Part A, but it has been processed as diagrammed in Fig. 3, including sputter-coating. Although coating with gold allows operation at higher accelerating voltages and the collection of higher resolution images, specimen damage, including the fracture (F) of fine cellular processes and membranes (M), is often present. In addition, the process often results in the artifactual separation of cells and the peeling away of edges (E) from the culture substrate. The sample in Part C was prepared by the OTAOTAO procedure, which involves repeated treatment with OSO_4 and tannic acid to stabilize the cell cytoskeleton. In this procedure, fine cell processes (arrows) remain intact and cell shrinkage is minimal.

they are not inherently conductive samples. Examples of these types of samples include a variety of polymers^[7,8] and crystals.^[9]

ROUTINE PREPARATION OF BIOLOGICAL SPECIMENS FOR TEM

Following ethanol dehydration, biological samples that will be examined by TEM must be infiltrated and embedded in a resinous material that will allow them to be cut into thin sections (typically in the range of 70–200 nm) that are mounted onto a grid substrate and placed in the TEM (Fig. 6). Several resins are available, but the most common are epoxy-based resins for routine ultrastructural work and acrylic resins that may be used when immunocytochemistry is to be performed on the samples. After specimens are embedded, the resins are polymerized with heat or UV irradiation and the specimens are sectioned on an ultramicrotome using a diamond knife. To enhance the contrast of sections and highlight structures of interest, they are frequently stained with lead and uranyl compounds. Although the exact mechanism of staining with these components is not fully understood, it is believed that the uranyl salts react with phosphate and amino groups of nucleic acids and some proteins, and lead stains bind to negatively charged components such as hydroxyl ions and areas that have reacted with OsO₄.^[1]

VARIATIONS IN THE PREPARATION OF BIOLOGICAL SPECIMENS FOR TEM

A variation in the preparation of biological specimens for TEM involves those studies that require immunocytochemical labeling with electron-dense substances

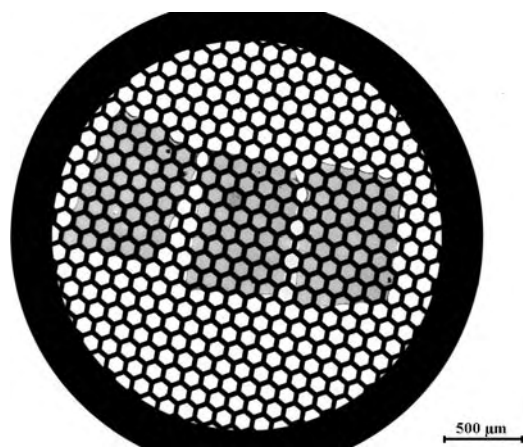


Fig. 6 TEM grid with three thin sections across the center of the grid.

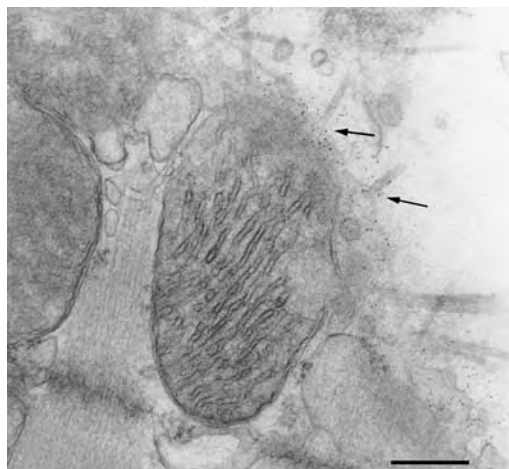


Fig. 7 High-resolution TEM immunocytochemistry showing 5 nm gold particles (arrows) labeling the extracellular matrix component laminin surrounding collagen fibrils in the region that they insert into a cardiac myocyte. Scale bar: 0.25 μm.

such as colloidal gold. These experiments are often used to localize the presence of specific proteins at high resolution (Fig. 7). In these types of studies, the extensive cross-linking of proteins that results from glutaraldehyde fixation and the effects of OsO₄ usually reduces the antigenicity of a specimen to the point that antibodies no longer recognize the antigens of interest. Various modifications of the fixation and dehydration protocol are often required in order to maintain the antigenicity of a sample. Ultrastructural immunocytochemical protocols have been developed for both pre-embedding experiments, in which the antigen:antibody reactions are performed prior to normal processing for TEM, and post-embedding experiments, in which the reaction is performed after resin embedding and sectioning. In both types of experiments, the ultrastructural integrity of the specimen is usually sacrificed to some extent to maintain the antigenicity of the sample by fixing it in a freshly prepared 2–4% solution of phosphate-buffered paraformaldehyde that contains a small concentration (0.1–0.2%) of glutaraldehyde. In general, pre-embedding protocols result in more efficient labeling of surface antigens, and post-embedding protocols should be used for intracellular labeling of antigens or in experiments that involve a large piece of tissue that would prevent penetration of the sample by antibodies and the labeling agent.

PREPARATION OF HARD BIOMATERIALS AND POLYMERS FOR TEM

The aforementioned techniques work well for biological samples and for soft nonbiological samples that

can be embedded in a resin and sectioned with a diamond knife on an ultramicrotome. However, many composite materials and alloys used as biomaterials and in biomedical engineering applications may be too hard to section on a routine basis. Examples of these types of samples are nanocrystals, which may be used for targeted drug delivery as discussed in the following text, or larger materials that may be used as scaffolds for tissue growth. Because of the small size of nanocrystals, it is often possible to spray a concentrated solution of particles directly onto a TEM grid that is coated with a thin plastic formvar or nitrocellulose film. When the solution dries on the grid, the material can be directly imaged by a TEM.^[10]

To image larger pieces of hard biomaterials, it may be necessary to thin the material to a point where the TEM electron beam can pass through the material to form an image. In this case, it may be necessary to use specialized equipment and techniques such as jet polishing, ion milling, or a focused ion beam to remove specific layers of material prior to imaging, or to create particles small enough to prepare as described earlier.^[11,12]

OVERVIEW OF SOME SEM AND TEM APPLICATIONS

Cell and tissue engineering often rely on biomaterials that act as scaffolds to physically support tissue growth and stimulate specific cell functions.^[13] Over the past several years, electron microscopy studies have certainly been important in developing a number of tissue engineering protocols for materials such as the hydroxyapatite composites used for bone and dental implants,^[14-16] the plastics used as a bridge to grow neuronal tissue,^[17] and other naturally occurring biomaterials, such as collagen, that have been used to grow a variety of cells and tissues.^[18-20]

Characterization of the ultrastructural morphology of these biomaterials by electron microscopy often includes imaging and defining features such as porosity,^[16,21-23] the arrangement of fibrous material such as carbon or collagen within a substrate,^[18,20,24,25] and topographical characteristics, such as minute ridges, that may guide cell growth and/or attachment.^[22,26] An extension of these basic studies that characterize biomaterials is the examination of cell and tissue phenotype and growth following exposure to a scaffold material. These studies frequently utilize SEM to characterize cell-spreading patterns, in order to illustrate that cells respond more in an *in vivo* like preparation on one substrate than another. Studies of this nature have shown that structures as small as a few nanometers can significantly alter the morphology and behavior of cells in culture.^[26-28] Other examples

of these types of studies include the application of chondrocytes to a variety of alginate and fibrin matrices to determine if they would promote sufficient chondrocyte proliferation and differentiation to produce a cartilage matrix. In these studies, Perka et al. found that specific mixtures of these biomaterials created stable cartilage transplants for reconstructive surgery.^[29]

In addition to the characterization of nanoparticles,^[11,15] TEM is frequently used to examine the interaction of extracellular matrix components, such as collagen, with implanted biomaterials;^[30] the differentiation and structure of cells in response to culture in the presence of a variety of biomaterials;^[18,25] and inflammatory responses to materials used in the construction of long-term bioengineering scaffolds or implants. For example, when metallic alloys used to repair bones are in contact with physiological fluids they frequently corrode and release small metal particles into the surrounding tissues and blood.^[31] This is often a problem in the use of long-term implants such as those used in hip replacements and in metals used for the internal repair of long bones. Germaine et al. found that nanoparticles of cobalt-chromium and alumina ceramics commonly associated with the wear of hip replacements resulted in the increased death of cells in culture.^[32] Vöggenreiter et al. also demonstrated by TEM that microparticles associated with the wear of stainless-steel and titanium implants in long bones were phagocytosed by cells in the vicinity of the implants, and that ingestion of the particles by the cells resulted in a significant inflammatory response.^[33]

Electron microscopy has also been used recently to examine a number of biomaterials designed to provide the controlled delivery of reagents to a specific tissue site for targeted drug delivery to tumors.^[34-36] As in the examples previously cited, it is important when designing these targeted carriers to be able to characterize the structure of these materials in order to determine factors such as size, shape, and porosity,^[10] as well as biodegradability and the immunological responses of the host.^[37,38] As nano-biomaterials get smaller and are used more frequently, electron microscopy will become even more important in their manufacture and characterization. Indeed, techniques and instrumentation associated with electron microscopy will need to be improved to keep pace with the technology.

SPECIFIC APPLICATIONS IN CARDIOVASCULAR TISSUE ENGINEERING

In our laboratories both SEM and TEM techniques have been used extensively to characterize a cardiovascular tissue-engineering model that involves the

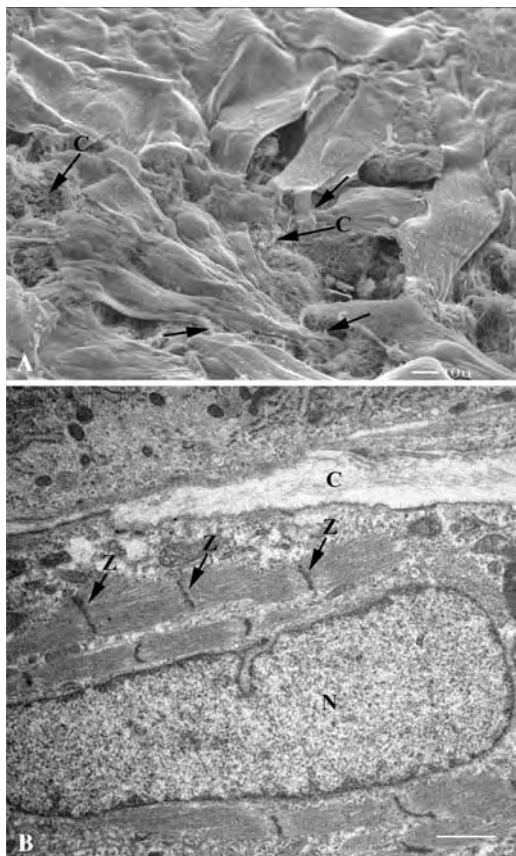


Fig. 8 SEM micrograph of cardiac fibroblasts (Part A) and TEM micrograph of cardiac myocytes (Part B) that have been cultured on collagen tubes used in a tissue engineering model for cardiac tissue. Thin collagen fibrils (C) from the tube are evident in both images. In Part A it is possible to determine fibroblast interaction with the collagen (arrows). In Part B, undifferentiated cardiac myocytes were placed in the tube and cultured for up to six weeks. During the culture period myocytes differentiate and develop mature sarcomeres, as illustrated by the well-developed Z-disks (Z), sarcomeres, and nuclei (N). Scale bars: Part A, 10 mm, ; Part B, 1 μ m.

seeding of cardiac myocytes and/or fibroblasts on an aligned collagen tube (Fig. 8). Descriptions of the collagen tubes, the equipment needed to produce them, and how these tubes are being used to examine the maturation of embryonic cardiac myocytes have recently been published.^[20,25] In this model system it is possible to seed the collagen tubes with undifferentiated embryonic cardiac myocytes, to observe their differentiation over time, and to treat the myocytes with a variety of growth factors and/or mechanical stimulants to induce increased rates of differentiation. This provides a model to study embryonic cardiac development during the period of time that myocytes convert from a hyperplastic to a differentiated hypertrophic phenotype, and also to study the potential to produce a viable cardiac patch with mature functional

cardiac myocytes, which can be implanted in an area of a cardiac infarct. Further development of this cardiac tissue-engineering model may provide a unique mechanism for repair of damaged heart tissue.

CONCLUSION

The Microscopy Society of America has adopted the slogan “We are the eyes of science,” and this statement certainly applies to the use of SEM and TEM in the fields of biomaterials and tissue engineering. As detailed earlier, both types of microscopy can be used to answer a number of questions concerning materials used in engineering, as well as the response of cells and tissues to variations in the materials. As these fields continue to develop, electron microscopy will play a central role in biomedical engineering and applications of the nanosciences in targeted drug and gene delivery for the modeling of cells and tissues.

ARTICLES OF FURTHER INTEREST

Cardiac Patch Engineering; Tissue Engineering of Cardiac Muscle

REFERENCES

1. Bozzola, J.J.; Russell, L.D. *Electron Microscopy*, 2nd Ed.; Jones and Bartlett Publishers: Sudbury, MA, 1999.
2. Murphy, J.A. Non-coating techniques to render biological specimens conductive/1980 update. *Scanning Electron Microsc.* **1980**, *1*, 209–220.
3. Erlandsen, S.L.; Lei, M.; Frethem, C.; Wells, C.L. High resolution cryoSEM: Examination of microbial samples in an in-lens and a below-the-lens FESEM. *Microsc. Microanal.* **2003**, *9* (Suppl. 2), 250–251.
4. Wright, E.R.; Conticello, V.P.; Apkarian, R.P. Cryo, cryo-etch and tandem cryo-HRSEM correlated with cryo-STEM of elastin-mimetic block copolymers. *Microsc. Microanal.* **2003**, *9* (Suppl. 2), 254–255.
5. Neilly, J.P.; Vogt, A.D.; Sciaraffa, L.A.; Ariman, S.Z.; Willcox, D.R.; Fagerland, J.A. Applications of environmental scanning electron microscopy (ESEM) in pharmaceutical development. *Microsc. Microanal.* **2003**, *9* (Suppl. 2), 198–199.
6. Stokes, D.J. Low vacuum & ESEM imaging of biological specimens. *Microsc. Microanal.* **2003**, *9* (Suppl. 2), 190–191.
7. Laitinen, O.; Pihlajamaki, H.; Sukura, A.; Bostman, O. Transmission electron microscopic visualization of the degradation and phagocytosis of a poly-L-lactide screw in cancellous bone: A long-term experimental study. *J. Biomed. Mater. Res.* **2002**, *61* (1), 33–39.
8. Watanabe, J.; Eriguchi, T.; Ishihara, K. Stereocomplex formation by enantiomeric poly(lactic acid) graft-type

- phospholipid polymers for tissue engineering. *Bio-macromolecules* **2002**, *3*, 1109–1114.
9. Zhang, Y.; Zhang, M. Synthesis and characterization of macroporous chitosan/calcium phosphate composite scaffolds for tissue engineering. *J. Biomed. Mater. Res.* **2001**, *55* (3), 304–312.
 10. Hu, Y.; Jiang, X.; Ding, Y.; Ge, H.; Yuan, Y.; Yang, C. Synthesis and characterization of chitosan-poly(acrylic acid) nanoparticles. *Biomaterials* **2002**, *23*, 3193–3201.
 11. Cheng, D.; Tellkamp, V.L.; Lavernia, C.J.; Lavernia, E.J. Corrosion properties of nanocrystalline C–Cr coatings. *Ann. Biomed. Eng.* **2001**, *29*, 803–809.
 12. Roberts, S.; Flatoff, D.; Chiou, W. Low damage sample preparation of semiconductor materials using low energy ion milling. *Microsc. Microanal.* **2003**, *9* (Suppl. 2), 810–811.
 13. Collier, J.H.; Camp, J.P.; Hudson, T.W.; Schmidt, C.E. Synthesis and characterization of polypyrrole-hyaluronic acid composite biomaterials for tissue engineering applications. *J. Biomed. Mater. Res.* **2000**, *50* (4), 574–584.
 14. Dong, Z.L.; Khor, K.A.; Quek, C.H.; White, T.J.; Cheang, P. TEM and STEM analysis on heat-treated and in vitro plasma-sprayed hydroxyapatite/Ti-6Al-4V composite coatings. *Biomaterials* **2003**, *24*, 97–105.
 15. Wang, X.; Li, Y.; Wei, J.; de Groot, K. Development of biomimetic nano-hydroxyapatite/poly(hexamethylene adipamide) composites. *Biomaterials* **2002**, *23*, 4787–4791.
 16. Zhao, F.; Yin, Y.; Lu, W.W.; Leong, J.C.; Zhang, W.; Zhang, J.; Zhang, M.; Yao, K. Preparation and histological evaluation of biomimetic three-dimensional hydroxyapatite/chitosan–gelatin network composite scaffolds. *Biomaterials* **2002**, *23*, 3227–3234.
 17. Brors, D.; Aletsee, C.; Schwager, K.; Mlynski, R.; Hansen, S.; Schäfers, M.; Ryan, A.F.; Dazert, S. Interaction of spiral ganglion neuron processes with alloplastic materials in vitro. *Hear. Res.* **2002**, *167*, 110–121.
 18. Itoh, S.; Takakuda, K.; Kawabata, S.; Aso, Y.; Kasai, K.; Itoh, H.; Shinomiya, K. Evaluation of cross-linking procedures of collagen tubes used in peripheral nerve repair. *Biomaterials* **2002**, *23*, 4475–4481.
 19. Pieper, J.S.; van der Kraan, P.M.; Hafmans, T.; Kamp, J.; Buma, P.; van Susante, J.L.C.; van den Berg, W.B.; Veerkamp, J.H.; van Kuppevelt, T.H. Crosslinked type II collagen matrices: Preparation, characterization, and potential for cartilage engineering. *Biomaterials* **2002**, *23*, 3183–3192.
 20. Yost, M.J.; Baicu, C.F.; Stonerock, C.E.; Goodwin, R.L.; Price, R.L.; Davis, J.M.; Evans, H.; Watson, P.D.; Gore, C.M.; Sweet, J.; Creech, L.; Zile, M.R.; Terracio, L. A novel tubular scaffold for cardiovascular tissue. *Tissue Eng.* **2004**, *10*, 273–284.
 21. Acil, Y.; Springer, I.N.G.; Broek, V.; Terheyden, H.; Jepsen, S. Effects of bone morphogenetic protein-7 stimulation on osteoblasts cultured on different biomaterials. *J. Cell. Biochem.* **2002**, *86*, 90–98.
 22. Dong, J.; Uemura, T.; Shirasaki, Y.; Tateishi, T. Promotion of bone formation using highly pure porous β -TCP combined with bone marrow-derived osteoprogenitor cells. *Biomaterials* **2002**, *23*, 4493–4502.
 23. Miura, H.; Nishibe, T.; Yasuda, K.; Shimada, T.; Hazama, K.; Katoh, H.; Watanabe, S.; Okuda, Y.; Kumada, T. The influence of node-fibril morphology on healing of high-porosity expanded polytetrafluoroethylene grafts. *Eur. Surg. Res.* **2002**, *34* (3), 224–231.
 24. Elias, K.L.; Price, R.L.; Webster, T.L. Enhanced functions of osteoblasts on nanometer diameter carbon fibers. *Biomaterials* **2002**, *23*, 3279–3287.
 25. Evans, H.J.; Sweet, J.K.; Price, R.L.; Yost, M.; Goodwin, R.L. Novel 3D culture system for study of cardiac myocyte development. *Am. J. Physiol., Heart Circ. Physiol.* **2003**, *285*, H570–H578.
 26. Dalby, M.J.; Riehle, M.O.; Johnstone, H.; Affrossman, S.; Curtis, A.S.G. In vitro reaction of endothelial cells to polymer demixed nanotopography. *Biomaterials* **2002**, *23*, 2945–2954.
 27. Dalby, M.J.; Yarwood, S.J.; Riehle, M.O.; Johnstone, H.; Affrossman, S.; Curtis, A.S.G. Increasing fibroblast response to materials using nanotopography: Morphological and genetic measurements of cell response to 13-nm-high polymer demixed islands. *Exp. Cell Res.* **2002**, *276*, 1–9.
 28. Deng, Y.; Zhao, K.; Zhang, X.; Hu, P.; Chen, G. Study on the three-dimensional proliferation of rabbit articular cartilage-derived chondrocytes on polyhydroxyalkanoate scaffolds. *Biomaterials* **2002**, *23*, 4049–4056.
 29. Perka, C.; Spitzer, R.S.; Lindenhayn, K.; Sittinger, M.; Schultz, O. Matrix-mixed culture: New methodology for chondrocyte culture and preparation of cartilage transplants. *J. Biomed. Mater. Res.* **2000**, *49* (3), 305–311.
 30. De Aza, P.N.; Luklinska, Z.B.; Santos, C.; Guitian, F.; De Aza, S. Mechanism of bone-like formation on a bioactive implant in vivo. *Biomaterials* **2003**, *24*, 1437–1445.
 31. Pina-Barba, C.; Bosch, P.; Villarreal, E.; Martin, S.; León, B.; Palma, B.; Torres-Villasenor, G. Bone response to 316L-SS and zinalco implants. *Microsc. Microanal.* **2003**, *9* (Suppl. 2), 1284–1285.
 32. Germain, M.A.; Hatton, A.; Williams, S.; Matthews, J.B.; Stone, M.H.; Fisher, J.; Ingham, E. Comparison of cytotoxicity of clinically relevant cobalt–chromium and alumina ceramic wear particles in vitro. *Biomaterials* **2003**, *24*, 469–479.
 33. Voggenreiter, G.; Leiting, S.; Brauer, H.; Leiting, P.; Majetschak, M.; Bardenheuer, M.; Obertacke, U. Immuno-inflammatory tissue reaction to stainless-steel and titanium plates used for internal fixation of long bones. *Biomaterials* **2003**, *24*, 247–254.
 34. Carrasquillo, K.G.; Ricker, J.A.; Rigas, I.K.; Miller, J.W.; Gragoudas, E.S.; Adamis, A.P. Controlled delivery of the anti-VEGF aptamer EYE001 with poly(lactic-co-glycolic) acid microspheres. *Invest. Ophthalmol. Vis. Sci.* **2003**, *44* (1), 290–299.
 35. Sanborn, T.J.; Messersmith, P.B.; Barron, A.E. In situ crosslinking of a biomimetic peptide–PEG hydrogel

- via thermally triggered activation of factor XIII. *Biomaterials* **2002**, *23*, 2703–2710.
36. Nsereko, S.; Amiji, M. Localized delivery of paclitaxel in solid tumors from biodegradable chitin microparticle formulations. *Biomaterials* **2002**, *23*, 2723–2731.
37. Kim, S.Y.; Lee, Y.M.; Baik, D.J.; Kang, J.S. Toxic characteristics of methoxy poly(ethylene glycol)/poly(ϵ -caprolactone) nanospheres; in vitro and in vivo studies in the normal mice. *Biomaterials* **2003**, *24*, 55–63.
38. Mi, F.; Lin, Y.; Wu, Y.; Shyu, S.; Tsai, Y. Chitin/PLGA blend microspheres as a biodegradable drug-delivery system: Phase-separation, degradation and release behavior. *Biomaterials* **2002**, *23*, 3257–3267.

Electrospinning

Darrell H. Reneker

Haoqing Hou

Department of Polymer Science, The University of Akron, Akron, Ohio, U.S.A.

INTRODUCTION

Electrospinning is a unique process that produces polymer fibers with diameters ranging from a few nanometers to a few microns. The electrospinning process is being used or finding uses in protective clothing,^[1] air filtration,^[2] template nanofibers for preparing metal or polymer nanotubes,^[3,4] biomedical applications including biomedicine,^[5–7] scaffoldings for tissue growth,^[8,9] and drug-delivery systems.^[10] Other possible uses include solar sails, light sails, mirrors for use in space,^[11] and nanoelectronics.^[12]

It is the purpose of this article to describe what polymer nanofibers are, how they are made by an electrohydrodynamical process called electrospinning, and show their potential relevance to the broad range of biomaterial needs outlined above. Hundreds of kinds of nanofibers have been made from polymers, including elastomers, and polymers that can be heated to form carbon, glass, or ceramics. Solvents ranging from paraffin oil to sulfuric acid have been used. Fibers have been electrospun from a shorter list of molten polymers. Nanofibers were used for more than two decades in the air filtration industry,^[2] e.g., to remove particles from the inlet air stream of large turbine and engines. Interest in the applications of nanofibers in filtration is widespread. The interest in the biomaterial applications of polymer nanofibers is not far behind. The technology of making and using nanofibers is not yet fully developed. Many technical and scientific problems need to be solved. It is fortunate that the electrospinning apparatus is inexpensive, simple to operate, fits nicely on a laboratory bench, and can be scaled up to produce large quantities of nanofibers.

ELECTROSPINNING PROCESS AND NANOFIBERS

History of the Electrospinning Process

Although the first patent on electrospinning of polymer fibers was issued to Formhals^[13] in 1934, Lord Rayleigh^[14] in the late 1800s studied the fundamental principles of electrospinning and calculated the maximum amount of charge that a drop of liquid can hold before the electrostatic driving forces overcome the

surface tension of the drop and eject a jet. Zeleny^[15] described and photographed electrospinning phenomena in the early 1910s. In the 1960s, Taylor^[16] studied the deformation and disintegration of water droplets in an electrical field and analyzed the conditions at the point that formed. He showed that a conical interface between two fluids in an electric field is stable if the cone has a semiangle of 49.3 degrees. During the past decade, Reneker and others studied the electrospinning process—both experimentally and theoretically—and also developed its applications.^[17–25] The average fiber diameter can be adjusted throughout a broad range down to a few nanometers^[26,27] by proper selection of the processing parameters.^[11,19] The electrospun fibers can be oriented and ordered as they are collected.^[28,29]

Apparatus and Description of the Electrically Charged Fluid Jet

The apparatus for the electrospinning process is simple, as shown schematically in Fig. 1. The high voltage power supply produces a voltage in the range of 0–30 kV or higher. For safety, the current is often limited to a few hundred microamperes. The schematic shows a jet moving in a horizontal direction from the tip of a syringe needle to the grounded collector. In fact, the direction of the jet can be vertical and directed up or down since the electrical forces on the jet are usually much larger than gravitational forces. For melt electrospinning, the apparatus is similar in principle. An electrical heater maintains the temperature of a reservoir of molten polymer, which is pumped to the tip.

When a strong electrostatic field is applied to the syringe needle or a capillary, a droplet of the polymer solution held at the tip by surface tension is deformed into a conical shape, which is called the Taylor cone. When electrostatic force overcomes the surface tension of the solution, a liquid jet is ejected from the tip of Taylor cone. The high frame rate images in Fig. 2 reveal the change of the shape of the drop to a sharp cone, the ejection of the jet, and the readjustment of the shape of the drop after the flow of the jet is established. The as-formed jet has a relatively large

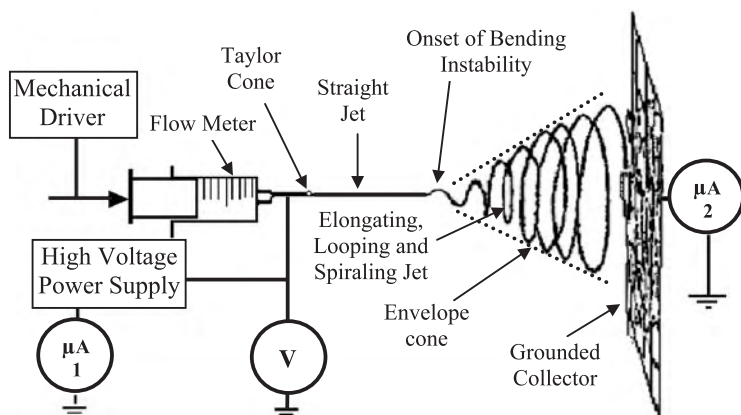


Fig. 1 Schematic of electrospinning setup using polymer solution.

diameter at the beginning and becomes thinner as the electrical forces elongate the jet.

The jet travels in a straight line for some distance in the direction of the applied electric field. Beyond that distance, the path is curved and more complicated. The path of the jet can be observed^[19] stereographically at a high frame rate using an exposure time for each frame that stopped the motion so that the incremental evolution of the path of the jet is observed in successive frames. A mathematical and computer model^[19] developed by Alexander L. Yarin shows that the curved path is a consequence of an electrically driven bending instability of the charged jet. The Yarin model reproduced the path of the jet, its motion, the large elongation that was observed, and other data. In a subsequent paper,^[30] it was shown that including the increase in viscosity, which occurs as the solvent evaporates from the jet, improved the agreement between the observed path and the calculated path.

A typical segment of the jet does not flow along the path shown in a particular stopped-motion image except in the straight part of the path. The trajectory of the

center of mass of a typical segment in the curved part of a jet path moves both outwardly and in the direction of the applied electric field between the tip and collector. There is also an influence from the charge carried on the more distant parts of the jet. All the while, the curved segment is elongated, mostly by the self-repulsion of the electrical charges carried with the segment.

This bending and elongating behavior of segments in the curved part of a jet is not self-evident. Neither splaying nor whipping provide adequate descriptions of the actual path of the jet. These words were used^[18,22,23,25] to attempt to describe the paths^[25,31] deduced from visual observations, ordinary video photography, and single stop-motion images. Splaying and whipping have been used uncritically in the introductions of contemporary papers on electrospinning.

Nanofibers and Nonwoven Nanofiber Sheets

A polymer nanofiber can be formed after solvent evaporates or the melt solidifies while the jet moves

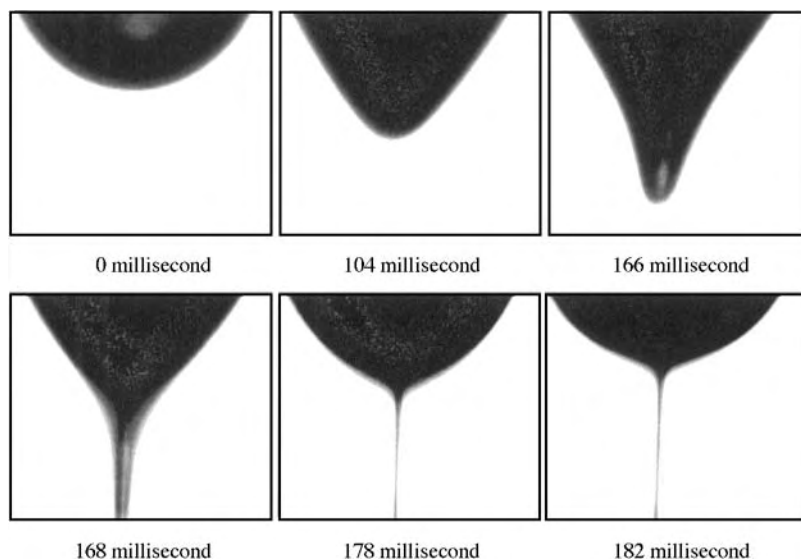


Fig. 2 High-speed images from CCD camera taken as the droplet became a Taylor cone from which a charged jet was ejected. A 7.0% solution of polyethylene oxide (Mw 400,000) in water was used for the experiment.

from the tip to the grounded collector. The diameter of the as-formed polymer nanofiber can be made as small as a few nanometers, as shown in Fig. 3A, by empirical choices of solvents and process parameters. The nanofibers are gathered on the collector in the form of a nonwoven sheet. Residual charges on the collected nonconductive nanofibers repel the oncoming nanofibers and tend to spread the nanofibers uniformly on the collector. Figure 3B shows a scanning electron microscope image of an electrospun nonwoven nanofiber sheet of biodegradable and bioresorbable polylactide. If a collector in the form of a cylinder rotating around its axis with a surface velocity greater than a few meters per second is used, a partially aligned nanofiber sheet is obtained, as shown in Fig. 3C and D.

Parameter Measurements and Nanofiber Characterization

The measurement concepts and methods for characterization of electrospun nanofibers described previously are helpful for control of the electrospinning process. Useful process parameters include: 1) electrical potential difference and electric field; 2) electrical current; 3) solution flow rate; 4) the diameter of the straight segment of the jet; 5) typical three-dimensional paths of the jet; and 6) the distribution of diameters of the nanofibers produced. The following provide practical comments on each of these measurements. Other parameters may also be useful.

Potential difference—in volts—between the tip and the collector

For a flat collector plate with a diameter that is not much smaller than the distance between the tip and the collector, the electric field along a line that passes through the tip and is perpendicular to the collector plate is equal to the potential difference divided by the distance between the tip and the plate. There are important differences from this rule at locations that are typically less than one micron from a conical tip with a half angle of 35 to 50 degrees. This rule only

applies before the jet is created. The jet is charged and has a complicated, changing path, which affects the electric field near other parts of the jet. The electric field at a particular segment can be approximated by calculations based on models of the experimentally observed paths.^[20]

Electrical current carried by a jet

Typical currents carried away from the tip by the electrically charged jet of polymer solution range from less than 100 nanoamperes to around 100 microamperes. The jet can be affected by currents caused by the movement of airborne ions that are sometimes generated inadvertently by corona discharge from parts of the apparatus that are surrounded by high electric fields. The presence of airborne ion currents can be detected by measuring the current between a metal ball placed near the apparatus and the grounded terminal of the power supply. It is usually possible to reduce the corona current by avoiding fine wires and sharp points in the parts of the apparatus to which the high voltage is applied or covering the conductors with insulating tubes.

Flow rate of the polymer solution

Pumps of various sorts can be used to establish controlled rates of flow into the tip although the part of the pump that contains the fluid will need to be at a high electrical potential if the collector plate is connected to ground. Alternatively, the pump and reservoir may be connected to ground, and the electrical potential difference between the tip and the collector established by applying a high voltage to the collector. The flow rate can be determined by the pumping rate, observing the change in the amount of fluid in the reservoir, or measuring the mass of dry fibers collected. Even with a constant flow rate pump, the flow rate of the jet may oscillate with time if the shape of the droplet varies with time. The flow rate of solution to the jet can be adjusted in an experimentally simple way by changing the height of a liquid column of solution

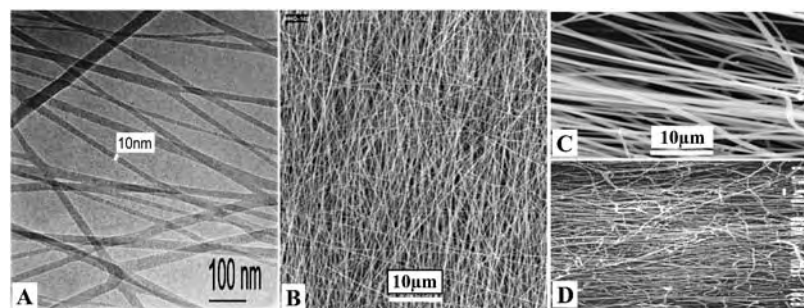


Fig. 3 (A) Transmission electron microscope image of nylon-4,6 electrospun nanofibers; (B) SEM image of PLA electrospun nonwoven nanofiber sheet. (C) and (D) are aligned PLA and collagen fibers. (Reproduced with permission from Ref. [9]. Copyright (2002) American Chemical Society.)

that feeds the droplet, e.g., by using a tilted pipette in place of the syringe pump shown in Fig. 1.

The electrical conductivity

Electrical conductivity of the solution can be measured with any of a wide variety of electrode geometries that are calibrated in a solution of known conductivity using an ohmmeter. Wires in contact with the ends of a column of liquid in a tube of known diameter is convenient since the length of the electrical path is determined by the distance between the ends of the wires. Bubbles of gas produced at the electrodes by electrolysis should be avoided by using low current and short measurement time.

Diameter of the jet

The diameter of a flowing jet in the straight segment before the electrically driven bending instability develops can be measured as a function of position by diffraction of laser light. This method, used by Doshi,^[17] measures the jet diameters in the range from around one micron to a few hundred microns. Video photography of the diffraction patterns follows temporal changes in the diameter of the segment of the jet that is illuminated by the laser beam.

Interference colors generated when white light is scattered in certain directions by a tapered cylindrical polymer fiber that is larger than the wavelength of visible light, which is around 0.5 microns, provide an elegant way to compare the diameters of adjacent segments of a jet when the diameter is in the range of 0.5 to 10 microns.^[32] Photography of the interference colors captures the time dependence of the diameters. The interference colors can be seen even where the jet path becomes complicated after the development of the electrically driven bending instability. The region within the envelope cone (Fig. 1), where the jet diameter is larger than the wavelength of light, is easily distinguished by its color from the faint white region where the jet diameter is smaller.

Three-dimensional path of the jet

Stereo images of the jet can be obtained by photographing the jet through a pair of wedge-shaped prisms with an apex angle of around 10 degrees. The prisms are placed near the jet, which is just far enough away to avoid attracting the charged jet to the prisms. The light that passes through one prism is captured on one side of a frame in a digital video camera, and the light that passes through the other prism is recorded on the other side of the frame. The size of the field that can be

photographed in this way is commensurate with the area of each of the wedge prisms. Stereo pairs captured synchronously with two or more digital video cameras allow the observation of larger or smaller fields and provide an adjustable choice of angles between the stereo views. The stereographic apparatus collects data that can be compared with the computer model described by Reneker et al.^[19,30]

Flash illumination with argon or other gas discharge lamps is very useful since the jet path does not change much during the 50 to 100 microsecond duration of the flash. An easy way to obtain representative images of jet paths is to set up a digital video camera to observe the interference colors described previously, and then illuminate the jet with periodic, unsynchronized sequences of flashes. The flashes from inexpensive strobe lights can be made to occur at a frequency slightly different from the 30 frames per second that is typical of video cameras. Occasional frames, in which the flash occurs while the camera is capturing an image, record images of the jet path at the instant of the flash. This illumination scheme was also used with the prisms described previously to obtain stereographic images of many typical paths in only a few minutes. The evolution of the shape of the jet path cannot be followed in detail by this method. A flash repetition rate of a kilohertz or more may be required to record the incremental changes satisfactorily.

Diameter distribution of nanofibers

Scanning electron microscope images provide data useful for determining the distribution of nanofiber diameters when the diameters are smaller than the wavelength of light. Both hand measurement of images and computational image analysis are useful and practical. Measurements of the intensity of laser light scattered by a nanofiber promises a less cumbersome method to determine the diameter of nanofibers, but this method is not yet fully developed.

Safety

Electrical shock hazards are minimal if the maximum current available from the high voltage power supply is limited to a few hundred microamperes. Large, electrically isolated conducting materials such as metal or solution reservoirs can accidentally form large capacitors. If such a capacitor is charged to a high voltage by the supply, a serious spark or shock could occur. The possibility of forming and charging a capacitor must be recognized and avoided.

The fire hazard of flammable solvents can be limited by keeping the amount of solvent small. Larger quantities of solvents require appropriate precautions

to prevent fire. Scale-up of the electrospinning processes requires careful engineering for safety.

Dangerous accumulations of explosive vapors must be avoided since sparks can occur at any time and sometimes occur in unexpected places. Noxious vapors can be avoided with adequate ventilation or efficient solvent recovery.

BIOMEDICAL APPLICATIONS OF ELECTROSPUN NANOFIBERS

Scaffolds for Tissue Engineering

Nanofibers offer many ways to make highly porous scaffolds to support and guide tissue growth. Nanofibers that provide biodegradability, bioresorbability, and biocompatibility can be made. The highly porous nature of scaffolds of nanofibers allows cell migration and growth as well as transport of nutrients and metabolic waste.^[33] Nanofiber technology will take its place alongside solution casting, gel casting, gas saturation, three-dimensional printing, and the like^[34] as a way to produce scaffolds useful for tissue engineering. By electrospinning a solution of poly(D,L-lactide-co-glycolide) (PLGA) in a solvent mixture (1:1) of tetrahydrofuran (THF) and N,N-dimethylformamide (DMF), Ko et al.^[8] made a scaffold structure in the form of an electrospun fiber sheet of 92% porosity with pore sizes ranging from 2 to 465 μm , as shown in Fig. 4A. The cells lived and divided in this structure, which pushed nanofibers aside to make room for the increasing number of cells. Bowlin et al.^[9] made a scaffold structure of electrospun collagen nanofiber. They seeded aortic smooth muscle cells into an electrospun collagen scaffold. Within seven days, the scaffold was densely populated with the smooth muscle cells (Fig. 4B).

Although the goal of producing a clinically useful tissue scaffold is not yet realized, electrospinning provides a practical approach. A wide variety of polymers can be used. Process parameters can be adjusted to produce nanofibers of different diameter and sheet density. Intricate designs of scaffolds, which require choice and control of the pore size distribution, mechanical properties, and biochemical properties of the scaffold, can be realized with nanofibers.

Nanofibers for Controlled Release of Drugs

Controlled delivery systems are used to improve therapeutic efficacy and safety of drugs by delivering them at a rate dictated by the need of the physiological environment to the site of action.

A new method to prepare drug-delivery matrix by electrospinning was recently reported and patented.^[10,35,36] Poly(lactic acid) (PLA), poly(ethylene-co-vinyl acetate) (PEVA), or their 50/50 blend were electrospun from a chloroform solution that contained a suspension of fine particles of tetracycline hydrochloride. Short-time, drug-release experiments showed that electrospun fibers of PLA, PEVA, or their 50/50 blend have a higher release rate than a comparable film. Electrospun mats of PLA exhibited a release rate about six times higher than that of films with similar compositions. The PEVA mats exhibited a release rate that was about 10 times higher than that of the comparable film after an 80-hour test. The diffusion path of drug from the nanofiber reservoir into the test liquid was much shorter than for the film.

The morphology, porosity, and composition of drug-release matrices can be controlled to optimize the release rate and release period. Drug-release systems made by the electrospinning process could be used not only for human health but also for the protection of crops or forest plants against disease and insect pests. Nanofibers loaded with pesticides can be applied by

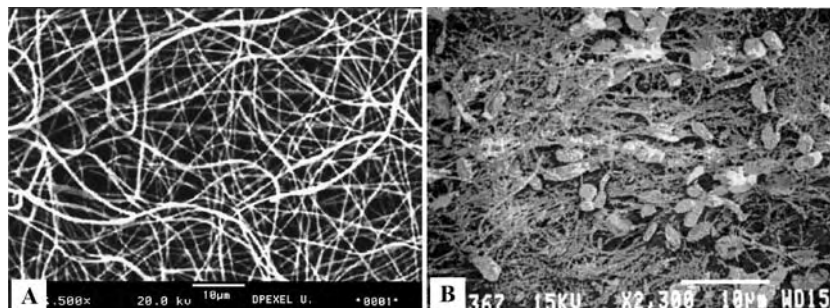


Fig. 4 (A) SEM image of electrospun poly(D,L-lactide-co-glycolide) (PLGA) scaffold structure composed of randomly oriented nanofibers, scale = 10 μm . (Reproduced with permission of John Wiley & Sons, Inc., Ref. [8].) (B) Electrospun collagen scaffold matrix densely populated with smooth muscle cells within 7 days of seeding the cells. Scale = 10 μm . (Reproduced with permission from Ref. [9]. Copyright (2002) American Chemical Society.)

electrospinning nanofibers onto electrically grounded plants. In this method, the plants capture the pesticides efficiently with minimal environmental pollution. Pesticides can be released from the nanofibers gradually over a relatively long time to shield crops or plants from disease. Pesticide retained in nanofibers on the plant leaves can kill insects that eat the leaves.

Wound Dressing

An ideal wound dressing provides: 1) a moist environment; 2) protection from secondary infections; 3) adequate gaseous exchange; 4) absorption of wound fluids; 5) painless and easy removal; and 6) prevention of secondary contamination.^[37] Winter^[38] showed that an occlusive wound dressing that maintains a moist wound environment accelerates healing.

Electrospun nanofibers provide a practical approach to improve wound dressings.^[6,7] Wound dressings can be produced for individual patients according to the location, size, and nature of the wound. The electrospun nanofiber membranes are porous and permeable to gases such as O₂, CO₂, and water vapor. Appropriate process parameters can control the pore size of the membrane and the surface area per unit mass to maintain a moist environment under the wound dressing and control the exchange gases or vapors. Antibacterial drugs can be carried with nanofiber. The surface area of the dressing can be increased even more by making highly porous nanofibers.^[39] The dressings can be made to be biodegradable, bioresorbable, and biocompatible by using appropriate polymers. Nanofiber wound dressings are flexible, conformable, and easy to apply. Complete removal of the dressing, which can be painful, may not be needed.

COATED NANOFIBERS

Nanofibers can be coated with other polymers, metals, carbon, or semiconductors. Ceramic nanofibers and tubes can be made from electrospun polymer nanofibers. Physical vapor deposition,^[4] plasma-enhanced physical sputtering,^[40] chemical vapor deposition,^[3] and the sol-gel process^[41] can be used to modify the surface chemical and physical properties of nanofibers. The applications of such coated nanofibers in biotechnology remain to be investigated. Some examples of coated nanofibers are described here.

Metal and Ceramic Coating

Subnanometer thick coatings of high purity metal can be deposited by physical vapor deposition or plasma-enhanced physical sputtering technique. A 10 nm-thick

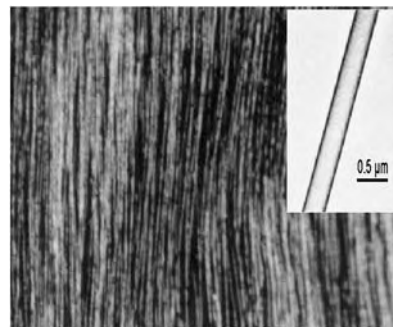


Fig. 5 Optical microscope image of nylon-4,6 coated with platinum by plasma-enhanced physical sputtering technique. The insert is a transmission electron microscope image of a platinum-coated nylon-4,6 nanofilament.

aluminum coating on poly[N,N'-(1,3-phenylene)]isophthalamide nanofibers, as reported by Liu et al.,^[4] was obtained by physical vapor deposition aluminum-coating technique. Subsequently, the aluminum-coated poly[N,N'-(1,3-phenylene)]isophthalamide nanofibers were converted into Al₂O₃ ceramic nanotubes by heating in air. Figure 5 shows platinum-coated nylon-4,6 membrane prepared by plasma-enhanced physical sputtering platinum coating of polymer nanofibers.

Coating with a Second Polymer

The chemical vapor deposition coating technique is very suitable for coating electrospun nanofibers with poly(p-xylylene) (PPX) or its derivatives poly(p-chloroxylylene) (PPX-C) or poly(α -bromo-p-xylylene). Poly(α -bromo-p-xylylene) can be converted into PPV, an electroluminescent polymer, by heating to 200°C. The chemical vapor deposition PPX-coating technique allows a hole-free coating to be formed on the substrate at room temperature and was used by Greiner et al.^[3] to form PPX-coated electrospun fibers. PPX is thermally stable, solvent-resistant, and biocompatible. The PPX-coated polylactide electrospun nanofibers were transformed into PPX nanotubes by thermal degradation or solvent-extraction removal of PLA.^[3] Because of the biocompatibility of PPX, these nanotubes are candidates for use in controlled release of drugs.

CONCLUSION

Nanofibers are promising candidates for biomedical applications. The mechanical, surface, and biochemical properties of nanofibers span a large and useful range. Their high tensile strength, which is combined with

flexibility and high surface area per unit mass, provide many possibilities for the design of new therapeutic materials, devices, and systems. Electrospinning of polymer nanofibers provides a broadly applicable method for producing nanofibers for needed research and development of these possibilities.

ARTICLES OF FURTHER INTEREST

Biomimetic Materials; Burn Dressing; Collagen; Drug Delivery, Controlled; Extracellular Matrix Scaffolds; Fabrics; Tissue Engineering, Microscale; Tissue Engineering Scaffolds

REFERENCES

- Gibson, P.; Schreuder-Gibson, H.; Rivin, D. Transport properties of porous membranes based on electrospun nanofibers. *Colloids Surf., A Physicochem. Eng. Asp.* **2001**, *187–188*, 469–481.
- Graham, K.; Ouyang, M.; Raether, T.; Grafe, T.; McDonald, B.; Knauf, P. In *Polymeric Nanofibers in Air Filtration Applications*, The Fifteenth Annual Technical Conference & Expo of the American Filtration & Separations Society, Galveston, TX, April 9–12, 2002; American Filtration & Separations Society: Falls Church, VA, 2002.
- Bognitzki, M.; Hou, H.; Ishaque, M.; Frese, T.; Hellwig, M.; Schwarte, C.; Schaper, A.; Wendorff, J.H.; Greiner, A. Polymer, metal, and hybrid nano- and mesotubes by coating degradable polymer template fibers (TUFT process). *Adv. Mater.* **2000**, *12*, 637–640.
- Liu, W.; Graham, M.; Evans, E.A.; Reneker, D.H. Poly(meta-phenylene isophthalamide) nanofibers: Coating and post processing. *J. Mater. Res.* **2002**, *17* (12), 1–6.
- Stizel, J.D.; Bowlin, G.L.; Mansfield, K.; Wnek, G.E.; Simpson, G. . *Int. SAMPE Tech. Conf.* **2000**, *32*, 205.
- Smith, D.; Reneker, D.H.; Kataphinan, W.; Dabney, S. PCT Int. Appl. WO 2000-US27775, October 6, 2001.
- Smith, D.; Reneker, D.H. PCT Int. Appl. WO 2000-US27776, October 6, 2001.
- Li, W.J.; Laurencin, C.T.; Caterson, E.J.; Tuan, R.S.; Ko, F.K. Electrospun nanofibrous structure: A novel scaffold for tissue engineering. *J. Biomed. Mater. Res.* **2002**, *60* (4), 613–621.
- Matthews, J.A.; Wnek, G.E.; Simpson, D.G.; Bowlin, G.L. Electrospinning of collagen nanofibers. *Biomacromolecules* **2002**, *3*, 232–238.
- Kenawy, E.R.; Bowlin, G.L.; Mansfield, K.; Layman, J.; Simpson, D.G.; Sanders, E.H.; Wnek, G.E. Release of tetracycline hydrochloride from electrospun poly(ethylene-co-vinylacetate), poly(lactic acid), and a blend. *J. Control. Release* **2002**, *81*, 57–64.
- Fong, H.; Reneker, D.H. Electrospinning and Formation of Nanofibers. In *Structure Formation in Polymeric Fibers*; Salem, D.R., Sussman, M.V., Eds.; Hanser: Munich, 2001; 225–246.
- Norris, I.D.; Shaker, M.M.; Ko, F.K.; MacDiarmid, A.G. Electrostatic fabrication of ultrafine conducting fibers: Polyaniline/polyethylene oxide blends. *Synth. Met.* **2000**, *114*, 109–114.
- Formhals, A. Process and Apparatus for Preparing Artificial Threads. US Patent 1,975,504 October 2, 1934.
- Lord Rayleigh. On the equilibrium of liquid conducting masses charged with electricity. *Philos. Mag. J. (Lond.)* **1882**, *44*, 184–186.
- Zeleny, J. Instability of electrified liquid surfaces. *Phys. Rev.* **1917**, *10*, 1–6.
- Taylor, G. Disintegration of water drops in an electric field. *Proc. R. Soc., A* **1964**, *280*, 383–397.
- Doshi, J.; Reneker, D.H. Electrospinning process and applications of electrospun fibers. *J. Electrostat.* **1995**, *35*, 151–160.
- Reneker, D.H.; Chun, I. Nanometre diameter fibres of polymer, produced by electrospinning. *Nanotechnology* **1996**, *7*, 216–223.
- Reneker, D.H.; Yarin, A.L.; Fong, H.; Koombhongse, S.J. Bending instability of electrically charged liquid jets of polymer solutions in electrospinning. *Appl. Phys.* **2000**, *87*, 4531–4547.
- Yarin, A.L.; Koombhongse, S.; Reneker, D.H. Taylor cone and jetting from liquid droplets in electrospinning of nanofibers. *J. Appl. Phys.* **2001**, *90* (9), 4836–4846.
- Spivak, A.F.; Dzenis, Y.A.; Reneker, D.H. A model of steady state jet in the electrospinning process. *Mech. Res. Commun.* **2000**, *27* (1), 37–42.
- Hohman, M.M.; Shin, M.; Rutledge, G.; Brenner, M.P. Electrospinning and electrically forced jets. I. Stability theory. *Phys. Fluids* **2001**, *13* (8), 2201–2236.
- Hohman, M.M.; Shin, M.; Rutledge, G.; Brenner, M.P. Electrospinning and electrically forced jets. II. Applications. *Phys. Fluids* **2001**, *13* (8), 2201–2236.
- Tsai, P.; Schreuder-Gibson, H.; Gibson, P. Different electrostatic methods for making electric filters. *J. Electrostat.* **2002**, *54*, 333–341.
- Shin, Y.M.; Hohman, M.M.; Brenner, M.P.; Rutledge, G.C. Experimental characterization of electrospinning: The electrically forced jet and instabilities. *Polymer* **2001**, *42*, 9955–9967.
- Hou, H.; Zeng, J.; Reuning, A.; Schaper, A.; Wendorff, J.H.; Greiner, A. Poly(p-xylylene) nanotubes by coating and removal of ultrathin polymer template fibers. *Macromolecules* **2002**, *35*, 2429–2431.
- Fong, H.; Reneker, D.H. Elastomeric nanofibers of styrene-butadiene-styrene triblock copolymer. *J. Polym. Sci., B, Polym. Phys.* **1999**, *37*, 3488–3493.
- Hou, H.; Averdung, J.; Czado, W.; Greiner, A.; Wendorff, J.H. Oriented Mesotubular and Nanotubular Oriented Nonwoven Fabrics and Their Manufacture. German Patent 2000-10053263, October 26, 2000.
- Hou, H.; Averdung, J.; Czado, W.; Greiner, A.; Wendorff, J.H. Oriented Mesotubular and Nanotubular Oriented Nonwoven Fabrics and Their Manufacture. PCT Int. Appl., PIXXD2 WO 0234986, May 2, 2002.
- Yarin, A.L.; Koombhongse, S.K.; Reneker, D.H. Bending instability in electrospinning of nanofibers. *J. Appl. Phys.* **2001**, *89*, 3018–3029.

31. Baumgarten, P.K. Experimental characterization of electrospinning: The electrically forced jet and instabilities. *J. Colloid Interface Sci.* **1971**, *36*, 71.
32. Reneker, D.H.; Xu, H. In *Fluid Polymer Jets During Electrospinning of Nanofibers. Annual APS March Meeting, Indianapolis, IN, March 18–22, 2002*; American Physical Society: College Park, MD, 2002. www.aps.org/meet/MAR02/index.html.
33. Itälä, A.I.; Ylänen, H.O.; Ekholm, C.; Karlsson, K.H.; Aro, H.T. Pore diameter of more than 100 μm is not requisite for bone ingrowth in rabbits. *J. Biomed. Mater. Res., Appl. Biomater.* **2001**, *58*, 679–683.
34. Agrawal, C.M.; Ray, R.B. Biodegradable polymeric scaffolds for musculoskeletal tissue engineering. *J. Biomed. Mater. Res.* **2001**, *55*, 141–150.
35. Ignatious, F.; Baldoni, J.M. Electrospun Pharmaceutical Compositions Comprising a Polymeric Carrier. US 2000-178682, January 28, 2000.
36. Ignatious, F.; Baldoni, J.M. Electrospun pharmaceutical compositions comprising a polymeric carrier. PCT Int. Appl., PIXXD2 WO 0154667, August 2, 2001.
37. Lin, S.Y.; Chen, K.S.; Liang, R.C. Design and evaluation of drug-loaded wound dressing having thermoresponsive, adhesive, absorptive and easy peeling properties. *Biomaterials* **2001**, *22*, 2999–3004.
38. Winter, G.D. Formation of the scab and the rate of epithelialization of superficial wounds in the skin of young domestic pig. *Nature* **1962**, *193*, 257–268.
39. Bognitzki, M.; Czado, W.; Frese, T.; Schaper, A.; Hellwig, M.; Steinhart, M.; Greiner, A.; Wendorff, J.H. Nanostructured fibers via electrospinning **2001**, *13* (1), 70–72.
40. Matossian, J.; Wei, R.; Vajo, J.; Hunt, G.; Gardos, M.; Chambers, G.; Soucy, L.; Oliver, D.; Jay, L.; Taylor, C.M.; Alderson, G.; Komanduri, R.; Perry, A. Plasma-enhanced, magnetron-sputtered deposition (PMD) of materials. *Surf. Coat. Technol.* **1998**, *108–109*, 496–506.
41. Caruso, R.A.; Antonietti, M.; Giersig, M.; Hentze, H.P.; Jia, J. Modification of TiO_2 network structures using a polymer gel coating technique. *Chem. Mater.* **2001**, *13* (3), 1114–1123.

Endoscopy

Jonathan T.C. Liu

Department of Electrical Engineering, Stanford University School of Medicine, Stanford, California, U.S.A.

Tonya Kaltenbach

Stanford University School of Medicine, Stanford, California, U.S.A.

Thomas D. Wang

Bio-X Program, Stanford University School of Medicine, Stanford, California, U.S.A.

Roy M. Soetikno

Stanford University School of Medicine, Stanford, California, U.S.A.

INTRODUCTION

Endoscopy is a medical tool at the forefront in the diagnosis and treatment of human diseases. Endoscopes, inserted through orifices such as the mouth, nose, anus, and urethra, play an instrumental role in the management of diseases of the oropharynx, esophagus, stomach, small intestine, colon, pancreatobiliary tree, larynx, bronchial tree, and urinary system. This article will focus on the use of endoscopes for optical diagnostic purposes in the gastrointestinal (GI) tract. It will present the fundamental concepts of endoscopy, including the optical and mechanical components of a typical endoscope. It will review current endoscopic imaging modalities as well as introduce emerging and novel endoscopic imaging technologies.

The field of endoscopy continues to evolve with the aid of technological innovations and development. Early endoscopes consisted of simple rigid tubes that provided limited views of a few easily accessed organs.^[1] Recent developments have substantially enhanced the capabilities of endoscopes. For example, fiber optic imaging bundles have allowed for the development of flexible instruments that may be guided through tortuous organs to visualize deeply into the body. The development of miniaturized semiconductor detectors have enabled substantial improvements in image resolution and a reduction in instrument size. Conventional endoscopy is based on the detection of diffusely reflected white light from tissue surfaces to reveal neoplasms.^[2] Recent advancements in optical imaging take full advantage of light's properties such as its spectral content and coherence to improve contrast and resolution. Ultrasound imaging has been combined with endoscopy to enable visualization beyond the tissue surface. Novel imaging modalities such as fluorescence imaging and optical coherence tomography (OCT) provide even more informative images. Technological improvements that are enhancing the capability of endoscopes to

visualize, diagnose, and treat human diseases are revolutionizing the practice of medical endoscopy.

ENDOSCOPE COMPONENTS

An upper GI endoscope will be used to illustrate the basic components of an endoscope. An upper GI endoscope enables visualization of the esophagus, stomach, and duodenum. A schematic diagram of the individual components of the endoscope is shown in Fig. 1. The distal tip houses the optics required for illumination and collection of endoscopic images, multipurpose instrument channels, and conduits for providing air, water, and suction.^[2] The bending section consists of a set of user-operated hinges that allow the distal tip to deflect at angles as high as 270°. The insertion tube is encased in a rugged plastic. The proximal end of the endoscope contains an article port into a multipurpose instrument channel, the control knobs for manipulating the hinges in the distal end, a lock for maintaining tip deflection, controls for air, water, and suction, and switches for capturing and storing images. The endoscope is connected to a video processor unit through an umbilical device, which contains the light guide, electrical connectors, and supply lines for air, water, and suction. Endoscope length varies according to the length of the targeted lumen. The total length of the upper endoscope is about 1.5 m.

The Distal Tip of Endoscope

The distal end of an endoscope contains the optics for image acquisition. A cross-sectional view of a forward-viewing endoscope is shown in Fig. 2A, illustrating the location of the objective and illumination lenses, air/water conduit, and detector. A corresponding end-view is shown in Fig. 2B, showing that identical illumination

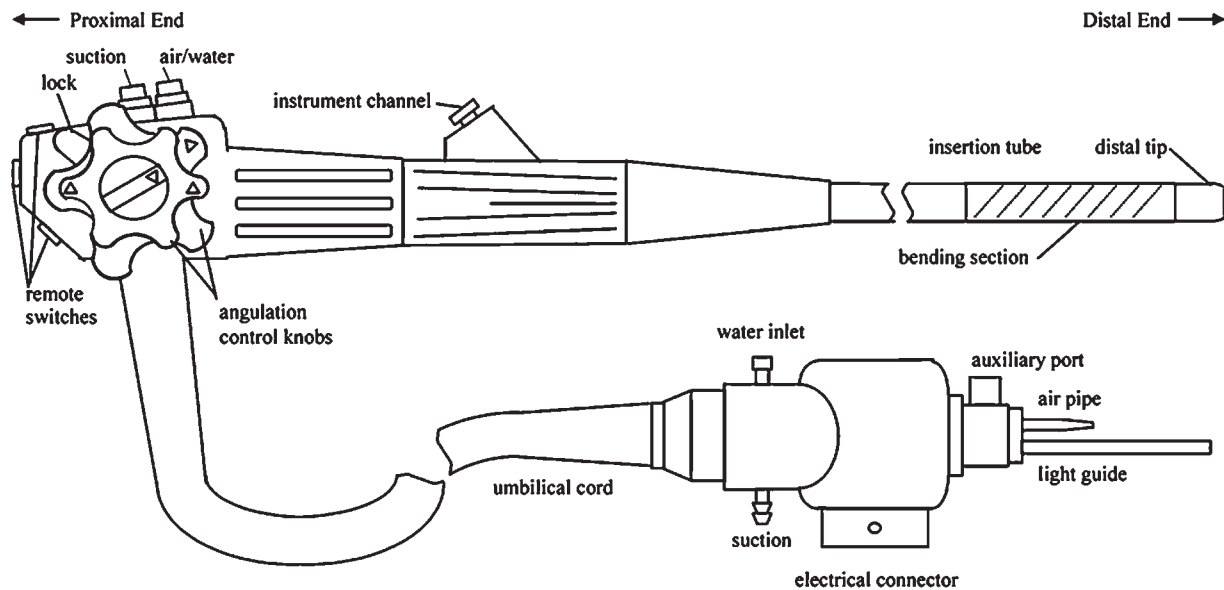


Fig. 1 A basic endoscope.

sources are located on either side of the objective, in this case, to uniformly illuminate the image. The large instrument channel is useful for the delivery of tools such as biopsy forceps, electrical cautery tools, balloon dilators, diagnostic brushes, etc. The instrument channel also acts as the suction tube. The objective generally contains several optical elements such as a diverging lens, intermediate lenses, a pupil, and an achromatic lens for imaging large fields of view and large spectral ranges with minimal aberrations. The air/water channel directs water across the outer surface of the objective to clear debris, or air to insufflate and expand the organ being examined. The first endoscopes used an optical fiber imaging bundle to transmit images to the proximal end of the endoscope for direct viewing through an eyepiece by an endoscopist. Advancements in semiconductor technology have led to the development of video endoscopes in which miniature charge-coupled device (CCD) arrays have replaced the imaging bundles. These CCD arrays are small enough to be positioned in the distal tip directly behind the objective lens.

The Imaging Lens

The objective lens of the endoscope is designed to image a large field of view with high resolution. As it is very difficult to achieve both of these requirements with a simple lens, endoscopes utilize multiple lenses to form an image. The design of these lens systems is complex, and is usually done with an optical ray-tracing program. An example of an optical train of lens elements required to produce an endoscopic image is shown in Fig. 2A. The angle between a ray of light and the normal to the objective is defined as θ . A diverging lens (negative focal length) is utilized to image a large angle θ and therefore to maximize the field of view. A pupil is located behind the objective lens to spatially reject extraneous internal reflections. A series of intermediate lenses, rather than one optical component, helps to focus the image onto the detector with minimal aberrations. An achromatic lens corrects for chromatic aberrations so that all colors in the visible spectrum are precisely focused at the detector. The focal length of the objective is different at the

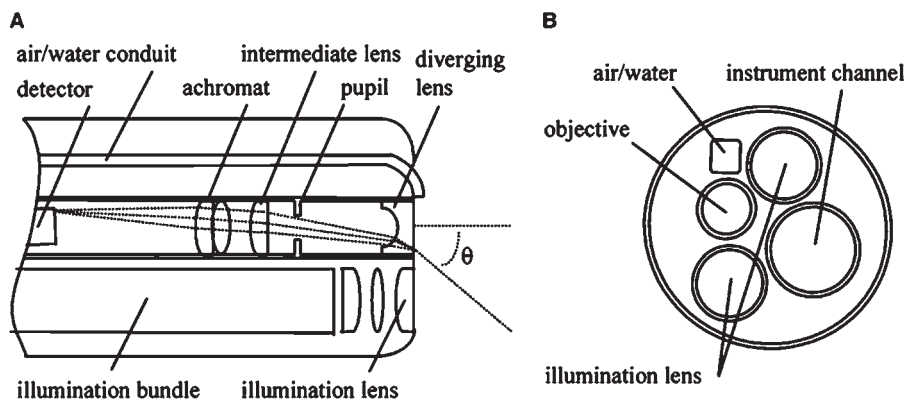


Fig. 2 The distal tip of a conventional video endoscope. (A) Side view showing illumination and detection optics. (B) End view.

center of the image than at the periphery, causing barrel distortion in the image.^[3] For an imaging system with barrel distortion, image magnification is higher at the center of the image than at the edges. The endoscopist can maximize the resolution of a region of interest by placing it in the center of the image where the density of image pixels is the highest (owing to increased magnification).

The Insertion Tube

An exposed cross section of the insertion tube is shown in Fig. 3. The outer layer consists of a durable plastic covering capable of withstanding caustic bodily fluids such as gastric acid and bile as well as the disinfectants used to sanitize the instrument after each use. Under the covering, a wire mesh runs along the length of the tube to provide structural rigidity. Under the mesh, helical steel bands maintain the round shape and provide further mechanical protection. The contents of the insertion tube may include a light guide, imaging bundle, angulation control wires, air and water pipes, a hollow instrument channel, and electrical wires to a CCD camera. The stiffness of the endoscope is designed to vary along its length to facilitate insertion into the GI tract. The distal end can be manipulated to produce the large bending angles that are needed for visualization of tissues located behind the endoscope (retroflexed view), as well as for aiding the endoscopist in traversing tortuous internal organs such as the intestines.

Air, Water, and Suction

A set of buttons and valves are located on the proximal end of the endoscope to deliver air and water or to suction intraluminal contents. Air supplied by a pump within the video processor unit may be emitted from a nozzle on the distal tip. Air is introduced into a lumen such as the colon or stomach to expand the mucosal folds for improved visualization. Air may also be

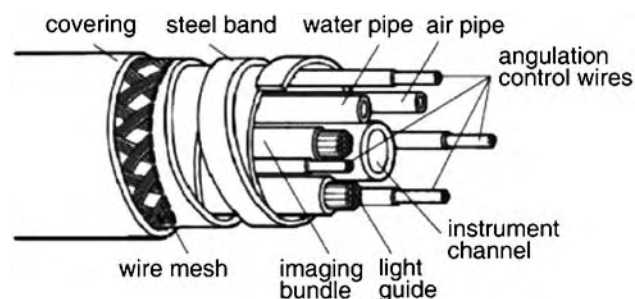


Fig. 3 The contents of the insertion tube include the light guide, imaging bundle, angulation control wires, air and water pipes, and instrument channel. Mechanical support is provided by a wire mesh and steel bands.

suctioned from a distended lumen for decompression. Water may be delivered through a nozzle on the distal tip, and over the surface of the objective to remove obstructing debris. Intraluminal contents such as fluid or stool are aspirated through the endoscope to a collection bottle via an external suction pump. The umbilical cord to the video processor unit contains connections for the light source, water supply, suction pump, and electrical power.

Manipulating the Endoscope

The angulation control of the distal tip is performed by the left hand of the endoscopist, leaving the right hand free to hold and manipulate the insertion tube. These angulations are produced by a durable set of guide wires that deflect the distal end in four directions. If desired, a brake can be applied to lock the position of the deflection. The second, third, and fourth fingers grip the proximal end of the instrument against the palm, leaving the left thumb to control the up/down knob and the left index finger to operate the air/water and suction valves. The right hand is used to torque and advance the insertion tube, to insert accessories down the instrument channel, to control the right/left knobs, and to activate remote switches.

Endoscopes are available with varying points of deflection, diameters, and instrument-channel sizes. These endoscopes are designed for specialized procedures, many of which are experimental. The exciting pace of advancements in endoscopic technology and development suggests that the tools and procedures described in this review will likely undergo further refinements.

Endoscopic Tools

A diverse set of tools are available for performing diagnostic and therapeutic endoscopic procedures such as standard biopsy, polyp removal (polypectomy), tumor resection, and hemostasis. These tools are generally inserted through the instrument channel of an endoscope. Biopsy forceps are often used to obtain samples of GI mucosa or diminutive (<2mm) polyps for histopathological analysis. Larger polyps may be removed by the use of snares that frequently utilize electrocautery to burn through tissues. For more effective and safe removal of sessile, flat, or depressed polyps, endoscopic mucosal resection is employed. Saline solution is injected with a sclerotherapy needle in the submucosa to elevate and separate the lesion from the underlying muscularis layer. Conventional snares and electrocautery devices may then be used to remove the lesion. Clip application is increasingly used in GI endoscopy, particularly in the treatment of GI bleeding and the closure of mucosal defects such as perforations,

anastomotic leaks, and fistulas. They are also used for radiographic marking to direct fluoroscopic stent placement and to target surgical, interventional radiologic, or radiation therapy. Argon plasma coagulation is used therapeutically to cauterize lesions or vasculature by the delivery of a large electrical current through ionized argon gas. There are many valuable tools in the armamentarium of endoscopic accessories. Further refinement of accessory technology and design will continue to strengthen the safety and efficacy of diagnostic and therapeutic endoscopy.

Non-GI Endoscopy

Endoscopy is often used to describe a host of related technologies outside of the GI tract. A brief listing of other common endoscopic modalities will be provided here. In laparoscopy, a small incision is made in the abdominal wall through which a laparoscope is inserted to visualize and perform procedures in the abdominal and pelvic region with minimal invasion. During laparoscopy, carbon dioxide gas is used to insufflate the abdominal region, separating the organs for better visualization. Coronary angioscopes are used to visualize the interior of blood vessels for diagnosing coronary lesions. A transparent balloon is inflated at the tip of the device to provide a clear imaging path to the vessel walls by displacing the blood, which is optically dense. Bronchoscopes are used for close examination of the bronchial-alveolar tree to perform a diagnostic lavage and biopsy. Scopes are also used by otolaryngologists in the naso-oral pharynx (laryngoscopy), by urologists in the bladder and urinary tract (cystoscopy), and by gynecologists in the uterus (hysteroscopy). As with colonoscopy, polyps may be removed and tissue may be cauterized. In short, endoscopic technologies are finding widespread use in an increasing number of medical areas. The overwhelming majority of endoscopic procedures, however, are performed in the GI tract.

ENDOSCOPIC IMAGING

Illumination

A variety of light sources such as metal halide, halogen, and xenon lamps may be used to provide endoscopic illumination. These light sources can provide incident powers of up to 300 W or more at incident wavelengths over the visible and near-infrared spectral regions. Because of the intense heat generated by these light sources, spectral filters are used to block the infrared portion, allowing only visible light to be delivered at the distal tip. The endoscope's light guide consists of an incoherent fiber optic bundle that has a large numerical

aperture (NA) lens at the proximal end to maximize the amount of illumination light coupled from the light source into the light guide. A diverging lens at the distal end is designed to illuminate the field of view of the imaging objective. The individual optical fibers in the light guide have high NAs and are made as large as possible in diameter, without compromising their flexibility and durability, to produce the highest possible packing fraction and illumination efficiency.

Image Detection

In early years, image detection was first performed by the eye of the endoscopist viewing the proximal end of a fiber optic imaging bundle through an eyepiece. Later, a camera was attached to the eyepiece to detect the image and display it on a video monitor. While this method is simple and effective, there are several limitations to fiber-based imaging.^[4] For example, a Moire pattern (interference artifact) is produced by the superposition of the image from the fiber optic bundle and the detector array of the camera. Also, the fiber optic imaging bundle is inefficient in collecting light, requiring high-power light sources. The number of optical fibers packed into the bundle, which affects the image resolution, is limited by instrument size and stiffness. Additionally, individual fibers may break over time, creating gaps in the image. In modern video endoscopes, a miniaturized CCD array located directly behind the objective lens detects the image and overcomes many of these shortcomings.

Image Resolution

The image resolution defines the level of detail that can be clearly seen on an endoscopic image. The transverse resolution is the smallest distance that two objects located perpendicularly to the optic axis (along the surface of the tissue) can be distinguished while the axial resolution defines the smallest distance that two objects located along the optic axis (into the tissue) can be distinguished. For conventional white light endoscopy (WLE), only the transverse resolution parameter is relevant because the light detected is reflected from the tissue surface and does not penetrate below. In other endoscopic methods where the light detected has penetrated with depth into the tissue, such as endoscopic ultrasound (EUS) and OCT, the axial resolution is also relevant. Barrel distortion introduced by the objective lens, as discussed earlier, causes the image resolution to vary across the field of view. Thus, endoscopes are characterized by an average transverse resolution. In Fig. 4, a set of plots shows the average transverse resolution of endoscopic images as a function of distance, d , varying from 5 to 75 mm, the typical range of

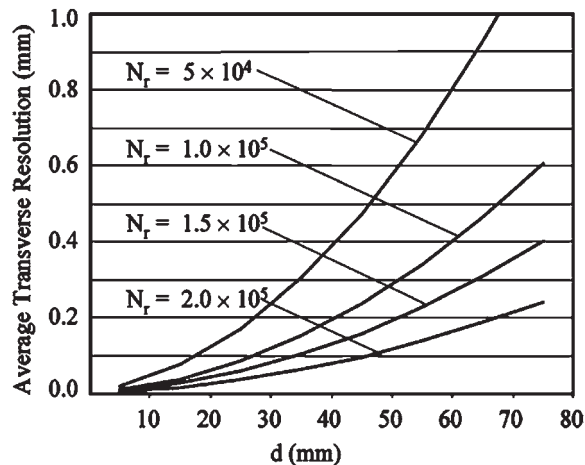


Fig. 4 Plots showing the average transverse resolution of endoscopic images as a function of distance varying from 5 to 75 mm. N_r is the total number of resels in the imaging system.

working distances. The total number of pixels (resolution elements), N_r , varies from 50,000 to 200,000, and reflects typical values found in fiber optic and video endoscopes. The total number of pixels is about 50,000 for fiber optic endoscopes (limited by the number of fibers) and about 100,000 to 200,000 for video endoscopes. Because more pixels can be packed into a CCD array than optical fibers in an imaging bundle, the average resolution for video endoscopes is generally better when imaging similar fields of view. As shown in Fig. 4, the average transverse resolution for both fiber optic and video endoscopes is submillimeter over the depth of focus and improves as the distal end of the endoscope is moved closer to the tissue surface.

ENDOSCOPIC RETROGRADE CHOLANGIOPANCREATOGRAPHY

Conventional endoscopes are too large to directly probe the pancreatic and bile ducts. However, a procedure

called endoscopic retrograde cholangiopancreatography (ERCP) is commonly used to visualize these sites with X-rays. In ERCP, a cannula is inserted into the papilla, the site where the pancreatic and common bile ducts join, and contrast is injected. This contrast is viewed with X-rays onto an external fluoroscopy monitor. Because the papilla, located within the ampulla of Vater, is on the side wall of the duodenum, there is not enough space for a forward-viewing endoscope to maneuver and perform the cannulation. Instead, a side-viewing endoscope, or duodenoscope, is used. The optics for a side-viewing endoscope are designed for imaging tissue located parallel to the axis of the endoscope. In Fig. 5A, a cross-sectional view shows an illumination bundle that is curved to deliver light to the side of the endoscope. The objective collects returning light from the side using a prism to reflect the image to the detector. Furthermore, the side-viewing endoscope has a forceps raiser located in the instrument channel to elevate the catheter during cannulation, as shown in the side-view in Fig. 5B. Only one illumination lens is located on the distal tip. Side-viewing duodenoscopes are about 120 cm in working length, and have a variety of instrument channel sizes, ranging from 2.8 to 4.0 mm. The larger diameter channels are useful for special instruments used in therapeutic biliary endoscopy, such as for endoscopic sphincterotomy, stone extraction, and stent placement. In Fig. 6, a radiograph shows the position of a duodenoscope with contrast agent illuminating the bile and pancreatic ducts. A filling defect in the cystic duct is because of a gall stone.

EUS

EUS combines endoscopy with ultrasound, using a miniaturized ultrasound transducer attached to the distal tip. Ultrasound images are collected with a transducer placed in close proximity to the structure being imaged.^[5] High resolution is obtained with EUS because

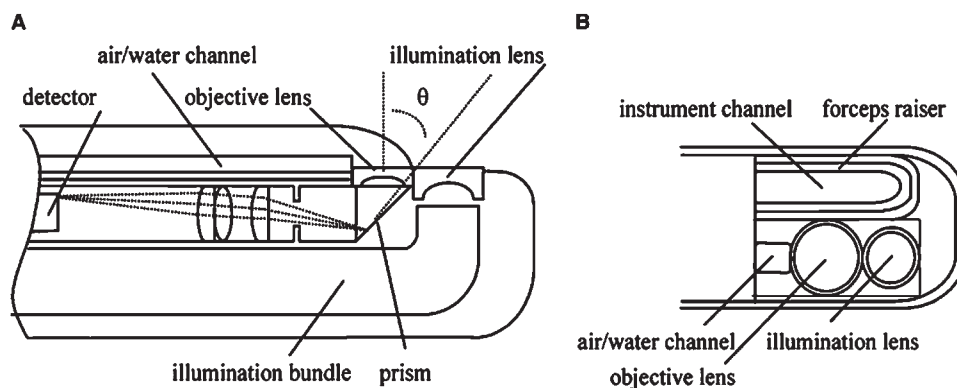


Fig. 5 (A) A cross-sectional view of a side-viewing endoscope shows the optics aligned to view images in the radial direction. (B) A side-view shows one illumination lens and a forceps raiser, utilized to help guide cannulation.

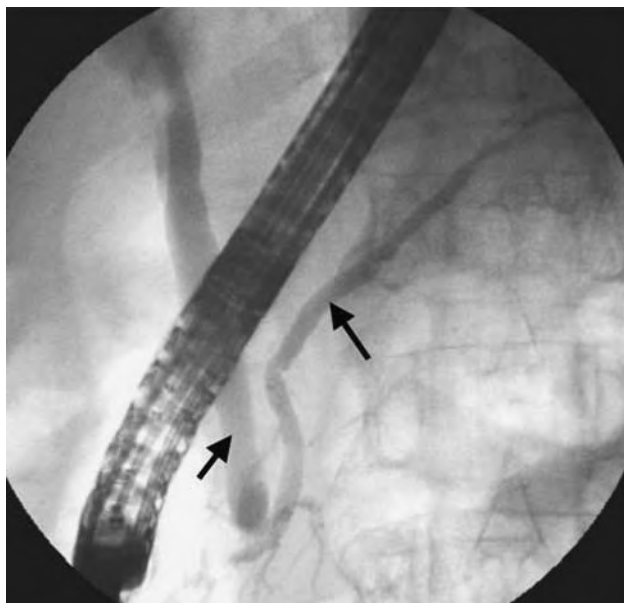


Fig. 6 A radiograph showing the position of a duodenoscope with contrast agent illuminating the bile duct (left arrow) and pancreatic duct (right arrow). The existence of a gall stone is preventing contrast agent from entering the cystic duct, creating a filling defect.

obstructing layers of bone and gas are avoided, allowing for higher acoustic frequencies to be used. In addition, high resolution of the GI wall can be obtained by having the ultrasound transducer placed directly on the wall. The clinical applications of this technology include determining the grade and stage of tumors and guiding fine needle aspiration (FNA) for tumor cytology. Two modes of scanning—radial sector and linear array—are often clinically used in a complementary fashion.^[6,7]

Radial Sector Scanning

Radial sector scanning provides ultrasound images in the plane oriented perpendicular to the axis of the endoscope. A radial sector echoendoscope uses a rotating acoustic mirror that scans images in sectors of 360 or 180°. They are available with 7.5 and 12 MHz transducers that have a focal distance of 30 and 25 mm and

an axial resolution of 200 and 120 μm , respectively, and a penetration depth of about 100 mm. The oblique-viewing echoendoscope, shown in Fig. 7A, has a small radial sector echoprobe attached at the distal tip. Radial scanning is usually performed as a preliminary step to determine the exact size, location, vascularity, and consistency of a lesion. An example of a 360° radial sector scan of the esophagus is shown in Fig. 8A. The layers of the wall, mucosa (m), submucosa (sm), and muscularis propria (mp) can be seen.

Linear Array Scanning

Linear array scanning provides ultrasound images in the plane oriented parallel to the axis of the endoscope, and is useful for guiding FNA procedures. A linear array scanning echoendoscope uses an array of transducers that provides a wedge-shaped image in sectors of 100 to 180°. They are available with 5 and 7.5 MHz transducers that have variable focal distances. The linear array can also provide color-flow mapping and Doppler imaging to locate blood flow in vessels for identifying landmarks and for assessing vascular invasion in tumors. In addition, this scanning mode provides continuous real-time visualization of a needle during EUS-guided FNA for cytology. The oblique-viewing echoendoscope, shown in Fig. 7B, has a curved linear array transducer located on the distal tip with a FNA needle protruding from the instrument channel. Fig. 8B shows an endosonographic view of an EUS-guided FNA of a periaortic lymph node. An arrow points to the needle used for aspiration.

WIRELESS CAPSULE ENDOSCOPY

In addition to video endoscopes, the development of miniaturized detector technology has allowed for images to be collected with a wireless capsule.^[8,9] The optics and electronics are contained within a small capsule, the size of a large vitamin pill, which can be swallowed and passed through the GI tract in several hours. There are no cables, wires, or optical fibers. This method of endoscopy is particularly suited for imaging

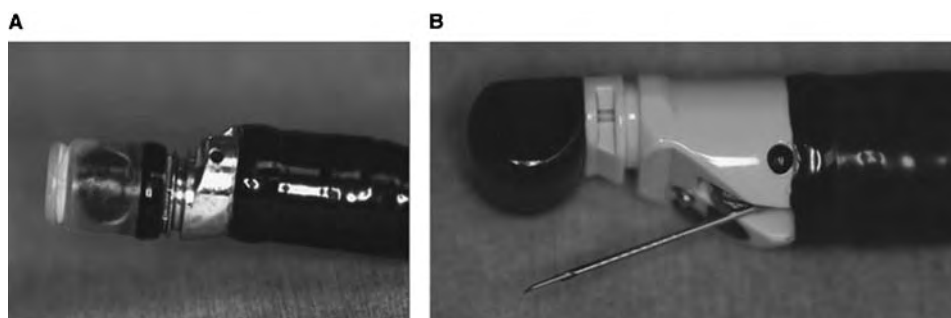


Fig. 7 (A) A mechanical radial EUS endoscope tip. (B) A curved linear array EUS endoscope tip with FNA needle protruding from the instrument channel.

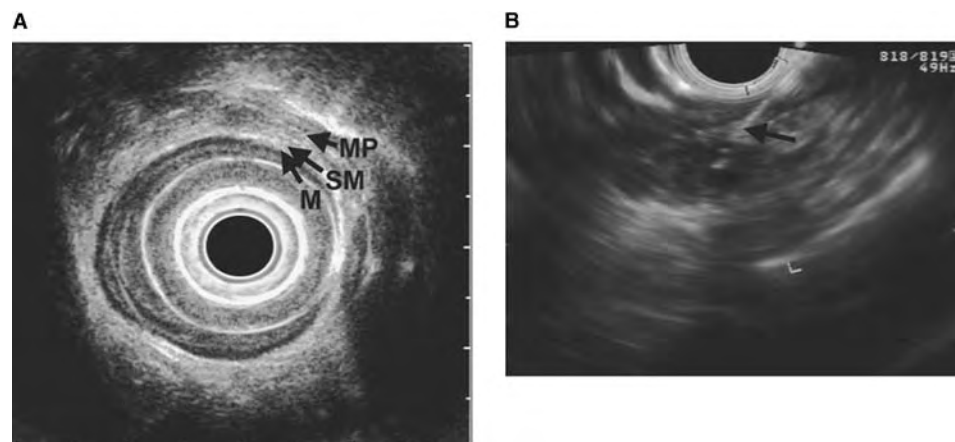


Fig. 8 (A) A radial EUS image of an esophageal mass (on the right side of the wall, mucosa (m), submucosa (sm), and muscularis propria (mp) can be seen. (B) EUS-guided FNA of a lymph node. An arrow points to the needle used for aspiration.

regions of the GI tract that cannot be easily reached by conventional endoscopes, such as the small bowel. The illumination light is provided by light-emitting diodes. Images are collected by a complementary metal oxide silicon detector that produces images of comparable quality to a CCD detector but with greatly reduced power consumption. The images are transmitted from within the body by radio telemetry to aerials worn around the waist, and are stored on solid-state media.

In current devices, images are delivered at a rate of two frames per second. The capsule passes through the human GI tract in an average time of 80 (range of 17–280) minutes. An example of an endoscopic image collected with the wireless capsule is shown in Fig. 9, where an arteriovenous malformation has been found in the jejunum. The anatomic location of the capsule within the small bowel is determined by calculating the time of travel within the small bowel relative to the total small bowel transit time and the strength of the transmitted signal. The wireless capsule has potential for diagnostic use in occult GI bleeding.

NEW AND DEVELOPING IMAGING TECHNOLOGIES

A variety of endoscopic imaging technologies are under development and show great promise, warranting a brief description of their principles and results.

High-Magnification Endoscopy

High-magnification endoscopy visualizes the details of the mucosal surface to differentiate between malignant and benign tissue.^[10,11] While most raised mucosal lesions such as polyps are easily identified with a conventional endoscope, there is increased evidence for the existence of flat lesions with malignant potential.^[12] These lesions have a subtle appearance on WLE. High-magnification endoscopes have special intermediate

lenses that can be inserted into the optical train of the objective to attain image magnifications as high as 170X, although most instruments operate at a magnification around 35X. In the high magnification mode, the depth of field is usually significantly reduced, and the endoscope must be placed very close to the mucosal surface. Also, the field of view decreases significantly, and the images are susceptible to motion artifacts. These endoscopes are often combined with high-resolution CCD detectors to provide transverse resolution as fine as 10 μm . In comparison, the human eye has a transverse resolution of approximately 125 μm . Conventional video endoscopes have CCDs with 100,000 to 200,000 total pixels. The transverse resolution varies with distance as shown in Fig. 4. High-magnification endoscopes can have total pixels as high as 850,000. High-magnification endoscopes are often used in combination with chromoscopy.

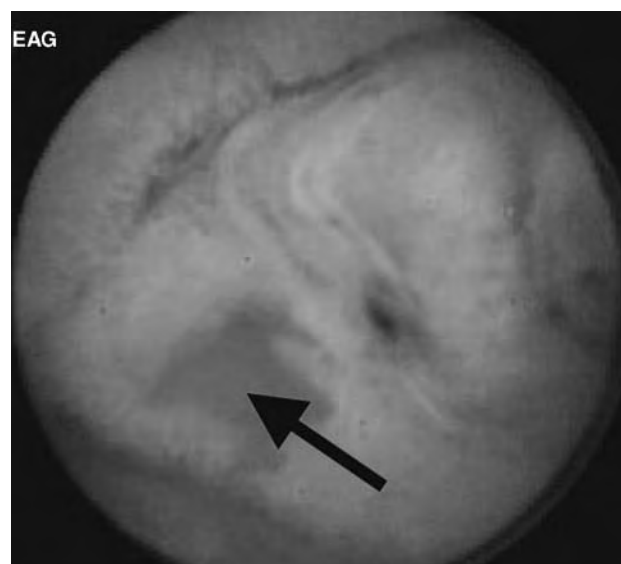


Fig. 9 Wireless capsule image of an arteriovenous malformation (shown by an arrow) in the jejunum.

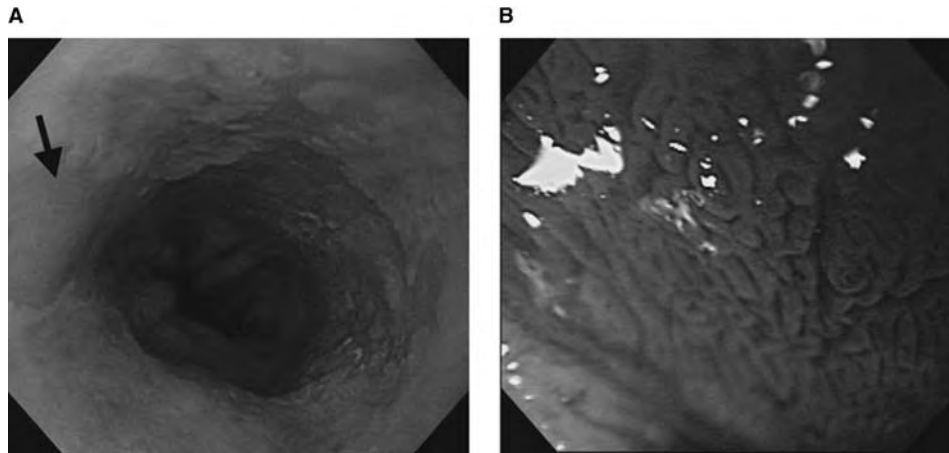


Fig. 10 (A) White light image of Barrett's esophagus in the distal esophagus. (B) A high-magnification view of the Barrett's epithelium.

An example of high-magnification endoscopy is illustrated with images taken of Barrett's esophagus, a premalignant condition for esophageal adenocarcinoma. In Fig. 10A, a standard endoscopic image is shown of the distal esophagus. The arrow points to a tongue of Barrett's esophagus (salmon pink) extending up into normal esophagus (light pink). In Fig. 10B, the Barrett's epithelium is visualized at high magnification. The magnified view shows a fine pit-patterned appearance. These pit patterns have a high correlation with the identification of Barrett's esophagus on histology.

Narrow Band Imaging

Narrow band imaging (NBI), comprising modified optical filters used in the light source of a videoendoscope system, has been applied to enhance image contrast.^[13] In brief, the videoendoscope light source unit consists of a xenon lamp and a rotating disk with three optical filters. The rotating disk and monochrome CCD are synchronized to sequentially generate images in three spectral bands. Using all three bands, a single color endoscopic image is synthesized. The image displayed on the monitor depends greatly on the spectral

characteristics of the optical filters. For example, NBI units have been developed with optical band pass filters at 500, 445, and 415 nm (with a FWHM band pass of 30 nm). In this case, the shorter wavelength falls within the hemoglobin absorption band. Colorectal neoplasms have been shown to alter normal capillary patterns. An NBI system may enhance capillary pattern imaging, thus increasing the ability to detect small or flat and depressed colon polyps. To illustrate the contrast enhancement that NBI affords, Fig. 11A shows a polyp imaged with conventional white light and the same polyp imaged with an NBI system in Fig. 11B.

Chromoscopy

Chromoscopy is an endoscopic technique that involves the application of tissue stains to highlight regions of interest in the mucosa, and is often used in conjunction with high-magnification endoscopy.^[14] The use of these methods is becoming more common for detecting flat lesions with malignant potential, which are not apparent on conventional endoscopy. There are two common types of stains: contrast and vital stains. Contrast stains such as indigo carmine and cresly violet collect

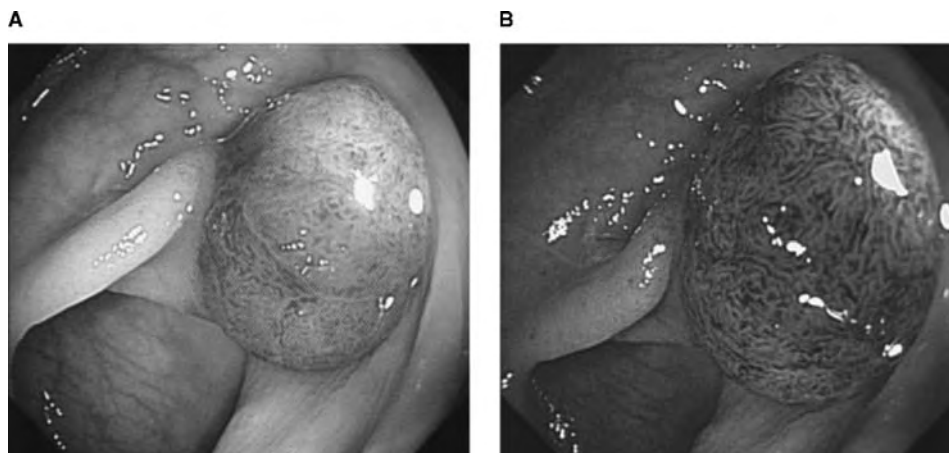


Fig. 11 (A) White light image of an adenomatous polyp. (B) A narrow-band image of an adenomatous polyp, showing improved contrast.

within the surface defects of the lesion and enhance the color contrast. These dyes are useful for identifying flat adenomas and carcinoma in situ in the colon and stomach. They are not absorbed by the mucosa and thus produce no systemic effects to the patient. Vital stains such as Lugol's iodine, methylene blue, and toluidine blue bind to the tissue and produce a chemical change. Lugol's solution contains iodine and potassium iodide and reacts with glycogen to produce a brown-black color. This dye is used to enhance lesions such as squamous cell carcinoma and high-grade dysplasia of the esophagus, which are poor in glycogen. Methylene blue is absorbed by the cytosol of the epithelium of the esophagus, small intestine, and large intestine. This dye is particularly useful for staining the intestinal metaplasia associated with Barrett's esophagus, which turns a bright blue color and exhibits a granular structure. Toluidine blue binds to the nucleus of cells and thereby highlights mucosa with increased DNA synthetic activity. This dye is also used for enhancing regions of Barrett's esophagus to screen for high-grade dysplasia. An example of chromoscopy in which indigo carmine has been applied to detect a small depressed tumor in the colon is shown in Fig. 12. On conventional endoscopy, these tumors exhibit a faint reddish or whitish color, subtle vascular abnormalities, and a slight deformity in the mucosal surface. In Fig. 12A, a sessile polyp is imaged with a slightly depressed lesion located on the top surface. The depression is not well visualized under standard white light colonoscopy. After indigo carmine is sprayed onto the lesion, the borders of the lesion are more clearly defined and the surface appearance is more clearly depressed as shown in Fig. 12B.

Endocytoscopy

Chromoscopy has been combined with high-magnification imaging for *in vivo* pathological diagnosis based on cellular and subcellular morphology of surface

tissues, a technique referred to as endocytoscopy.^[15] In this technique, an imaging device is passed through the instrument channel of a standard endoscope. An observation of nuclear-to-cytoplasm ratios, as well as cell density and staining characteristics, allows for delineation between superficial mucosal tissues that are normal or dysplastic.

Fluorescence Imaging

The endoscopic techniques discussed previously provide images of structural changes in the tissue. New endoscopic methods collecting fluorescence can determine biochemical and molecular changes in the tissue. Autofluorescence can occur from endogenous molecules such as aromatic amino acids, NADH, FAD, and porphyrins. Fluorescence can also be produced by exogenously administered drugs such as hematoporphyrin derivative and 5-aminolevulinic acid (5-ALA)—a technique known as photodynamic diagnosis. Both of these fluorescence methods are used to identify cancer at an early stage.

Autofluorescence

Autofluorescence images can be collected with either a fiber optic or video endoscope. Because no exogenous drugs are administered, these images can be acquired during routine endoscopy, and the patients do not experience photosensitivity after the procedure. However, the fluorescence produced from endogenous molecules is much weaker than that from exogenous agents or from diffuse reflectance, by a factor of 1000 or more, and a laser is required for excitation (illumination). With a fiber optic endoscope, fluorescence can be separated into two or more spectral bands to ratio the images, which corrects for spatial nonuniformities. The optimal spectral regimes for collecting images are determined by fluorescence spectra collected from

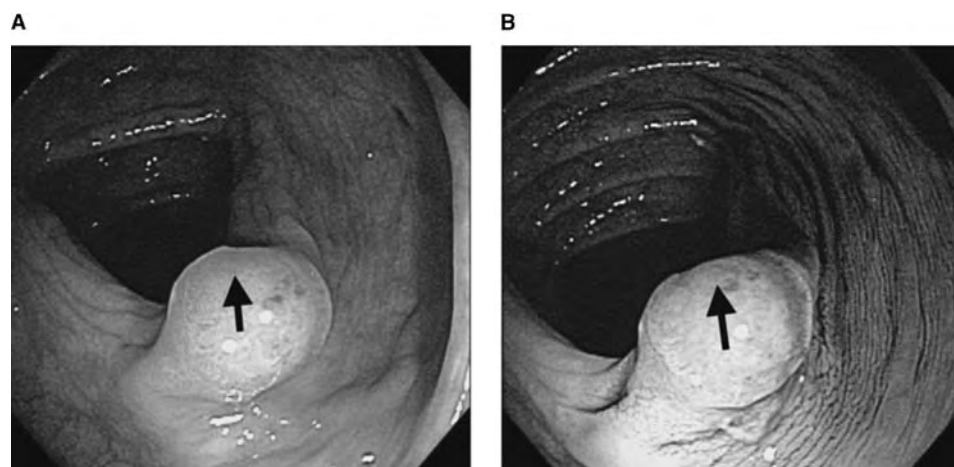


Fig. 12 (A) An slightly depressed lesion on the top surface of a sessile polyp, imaged in the colon with standard colonoscopy. (B) The depression is best visualized using chromoscopy with indigo carmine.

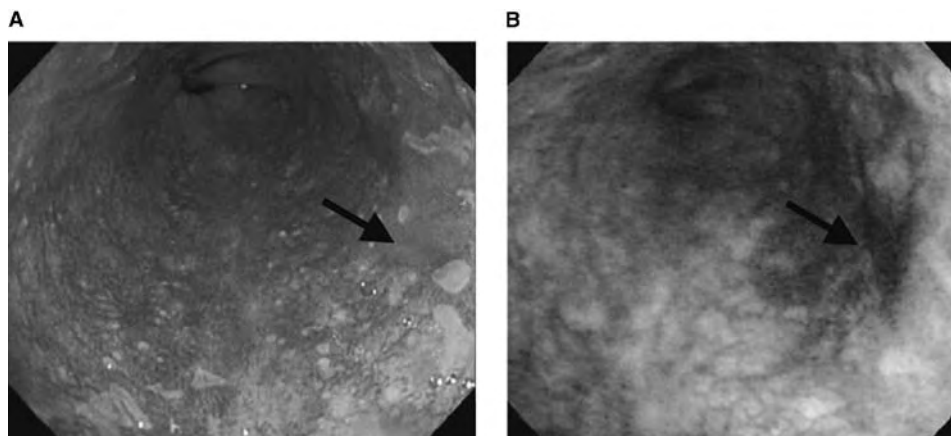


Fig. 13 (A) WLE image of the distal esophagus in a patient with Barrett's esophagus. (B) A corresponding autofluorescence image shows a bluish/purple color on the right side. Biopsy specimens revealed carcinoma in that region. Initial WLE did not reveal the lesion, which is indicated with an arrow in each image.

E

single-point optical fiber probes. An intensified camera is generally required for image detection owing to the low signal levels. A video endoscope can also be used to collect fluorescence. Because the CCD detector is located directly behind the objective, the fluorescence signal is much more intense. Image-processing algorithms are required to correct for spatial nonuniformities, and to identify regions of disease.

An example of an autofluorescence image of a flat colonic adenoma is shown in Fig. 13, along with a corresponding white light endoscopic image of the same region.^[16] This example demonstrates the improved ability to localize neoplastic mucosa with autofluorescence imaging. A xenon light source with a rotary RGB band pass filter sequentially illuminates the mucosa with red, green, and blue light at 20 Hz. Two CCD detectors are utilized, one to collect a white light image (by combining the RGB reflectance signals) and one to collect autofluorescence and reflectance signals from the individual RGB illumination phases. An image-processing algorithm combines the autofluorescence and reflectance signals to provide the high-contrast image shown in Fig. 13B. In these images, normal mucosa and nondysplastic Barrett's esophagus appear green whereas neoplastic lesions appear bluish/purple. In addition to premalignant tissue in other areas of the GI tract, autofluorescence imaging has demonstrated the potential to detect lesions in the bladder, head and neck, larynx, lung, and nasopharynx.

Photodynamic imaging

Photodynamic imaging collects fluorescence generated from externally administered agents that are internally converted and selectively taken up by cells exhibiting elevated metabolic activity. A promising agent is 5-ALA, a porphyrin substrate in the heme biosynthetic pathway, which induces the formation of protoporphyrin IX (PPIX), a compound with high fluorescence efficiency. 5-ALA has demonstrated potential to detect

pre-malignant tissue in the bladder, colon, esophagus, and oropharynx. Excess exogenous application of 5-ALA leads to a three- to sixfold higher intracellular accumulation of PPIX in areas of high-grade dysplasia and malignant lesions compared to the surrounding mucosa. PPIX fluorescence is excited by blue light (peak of 405 nm), and the emission is red at wavelengths above 600 nm. ALA can be administered orally for systemic distribution, and the drug is cleared in about 48 hr, and can produce mild skin photosensitivity and transient elevation of liver enzymes. An example of white light and photodynamic imaging of fluorescence with 5-ALA is shown in Fig. 14.^[17] Six or seven hours prior to upper endoscopy with a fiber optic endoscope, patients ingested an ALA solution at a dose of 15 mg/kg body weight. The light source was a 300 W short-arc violet lamp, with fluorescence at wavelengths above 600 nm being detected with a CCD camera attached to the fiber optic endoscope. The images indicate an improved ability to distinguish an esophageal carcinoma that is not easily detected in this case with WLE.

OCT

OCT performs cross-sectional imaging in a manner similar to that of EUS, except that the coherence (phase) property of light rather than the backscattering behavior of sound is used.^[18] Moreover, OCT has up to 10 to 100 times better axial resolution than EUS at a penetration depth of up to 1 mm. OCT is a technology that is being developed to assess tissue histopathology in vivo—a goal that is often termed “optical biopsy.” OCT has demonstrated the ability to distinguish layers of mucosa in the GI tract with the potential to detect precancerous lesions such as dysplasia. The OCT, in its most basic form, uses an interferometer to analyze light as it is reflected from a layer of tissue in one optical path to light traveling along a reference path of nearly equal length. Because a low-coherence light source is used, an interference signal is only generated when the path

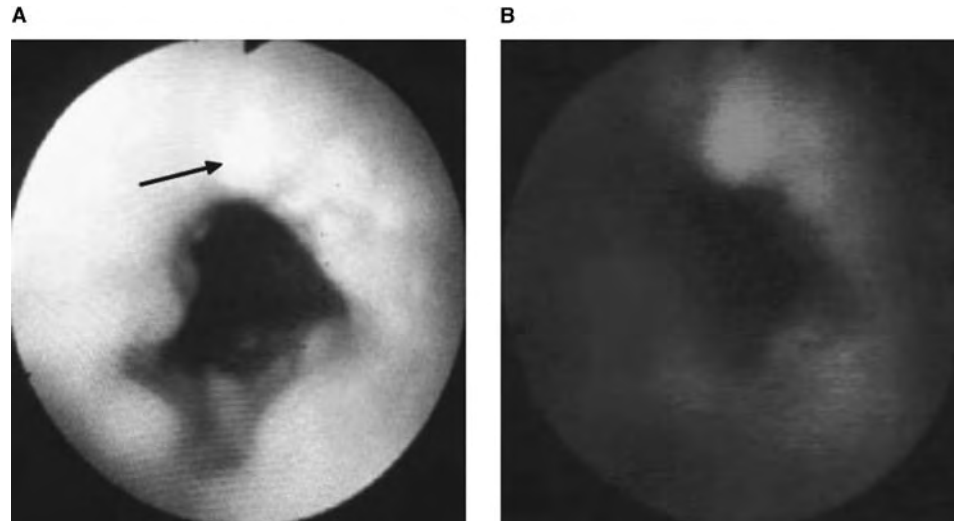


Fig. 14 (A) WLE image of a region containing esophageal carcinoma, which is not noticeable under WLE. (B) A corresponding image showing ALA-induced fluorescence (red) at the site of the tumor.

lengths are matched to within the coherence length of the source (typically on the order of $10\ \mu\text{m}$). As the length of the reference arm is scanned, different tissue depths are probed, enabling vertical image sections to be acquired. As shown in the schematic diagram in Fig. 15A, light from a low-coherence light source, $22\ \text{mW}$ at $\lambda = 1300\ \text{nm}$, is coupled into a single mode optical fiber and divided by a beam splitter to the scanner and reference arms of a Michelson interferometer.^[19] Contrast is provided by variations in the reflectance of cellular and extracellular material present in each tissue layer.

In the OCT system described, the transverse and axial resolutions at $\lambda = 1300\ \text{nm}$ are approximately 25 and $8\ \mu\text{m}$, respectively, and the light penetrates on the order of a millimeter into the tissue. The images are

acquired in a manner similar to radial sector scanning in EUS with image dimensions of 600×600 pixels in depth with a frame rate of up to four revolutions (images) per second. In Fig. 15B, the radial OCT image shows the layers of normal esophagus.^[19] OCT imaging has also demonstrated potential for other diagnostic purposes such as the identification of premalignant tissue in the lung, endometrium, and cervix, measurement of blood flow rates via the Doppler effect, and the characterization of atherosclerotic plaque in coronary arteries.

In Vivo Confocal Microscopy

An histologic analysis by trained pathologists is currently the most reliable method of diagnosing

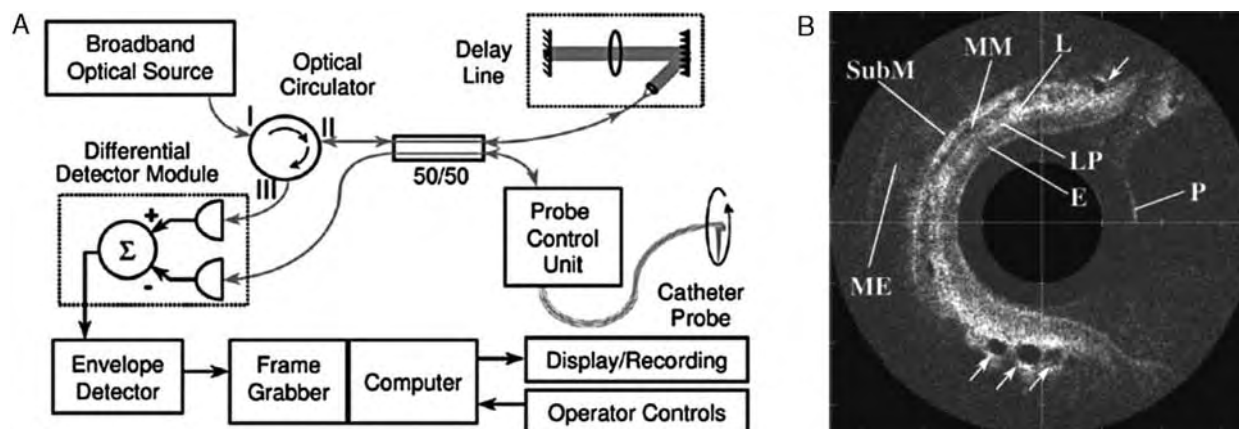


Fig. 15 (A) A block diagram of an endoscopic OCT system. The OCT catheter probe constitutes one arm of the Michelson interferometer (at the 50/50 coupler) whereas the rapid-scanning delay line is the other arm. Grey lines in the diagram are optical and black lines are electronic. (B) A single 4 frame/sec endoscopic OCT image with a radial image depth of $2.5\ \text{mm}$ and 600×600 pixels constructed from 1000 (430-pixel) radial "A-scans." The outer surface of the probe sheath is labeled P. Mucosal structures are labeled as squamous epithelium (E), lamina propria (LP), muscularis mucosa (MM), submucosa (SubM), and muscularis externa (ME). Arrows correspond to blood vessels and glands. Bright inclusions in the lamina propria may correspond to lymph nodes (L).

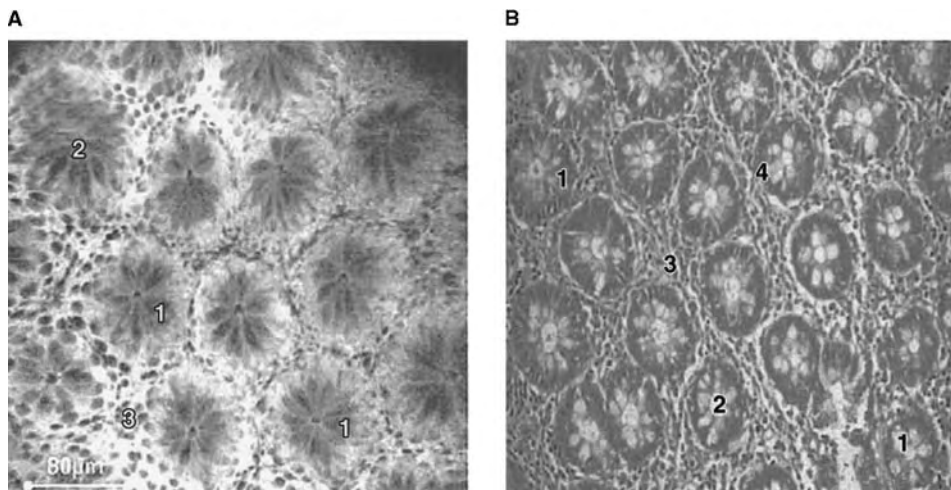


Fig. 16 (A) Confocal endomicroscopic image of rectal mucosa, obtained with fluorescence contrast from intravenous fluorescein injection. (B) Corresponding histology (H&E) showing 1. goblet cells; 2. crypt lumen; 3. lamina propria; and 4. nuclei.

malignancy. Recent technologies such as OCT strive to provide an “optical biopsy” that is in situ and in real time with the same diagnostic capability as histopathology. Confocal microscopy is an optical-sectioning technology, which differs from OCT in that spatial filtering techniques are utilized, rather than coherence gating techniques, for providing optical sectioning and the rejection of out-of-plane scattered light.^[20] This allows for incoherent fluorescence to be used in conjunction with confocal techniques. Much work has been done to develop exogenous optically labeled molecules, which specifically target and bind to biomarkers of neoplasia, in a field known as functional or molecular imaging. The development of in vivo confocal microscopes that are endoscope compatible holds great potential for powerful biomarker-assisted imaging strategies for the early detection of cancer in hollow organs. Various research groups and companies are developing confocal endomicroscopes as clinical tools.^[21] Fig. 16A shows an in vivo confocal image acquired with an endomicroscope developed by Pentax Corporation.^[22] Fluorescence contrast is obtained through the intravenous injection of sodium fluorescein in the patient. The axial (in plane) resolution of the confocal device is $7\ \mu\text{m}$ whereas the lateral resolution is $0.7\ \mu\text{m}$. The excellent resolution and contrast of the image compares well with histology performed on biopsy specimens, shown in Fig. 16B. The potential for technologies such as this, to provide optical histology in vivo, could be of great diagnostic utility, as well as for guiding surgical procedures in real time.

CONCLUSIONS

The field of endoscopy is undergoing vast technological development, with new innovations promising to radically improve the capabilities of endoscopes in the near future. These technological improvements are diverse

and interdisciplinary. Advances in semiconductor technology have allowed detectors to become miniaturized, and have led to the development of video endoscopes and wireless capsules. These endoscopes can view larger areas with higher resolution. At the same time, the diameters of insertion tubes are becoming smaller in size, improving patient comfort. The addition of ultrasound technology to endoscopy has allowed structural information to be collected beyond the tissue surface to depths of several millimeters and even into adjacent tissue structures. New methods are being developed to obtain even more diagnostic information from tissue. OCT provides morphologic information with subcellular resolution for real-time histopathology. Fluorescence spectroscopy and confocal microscopy, using exogenous fluorophores and photosensitizers, have the potential to reveal biochemical and molecular information below the tissue surface. Magnification and chromoscopy provide high-resolution details about the mucosal surface. The wireless capsule is providing endoscopic images from previously uncharted territory such as the small bowel. Active research and development is revolutionizing the practice of diagnostic and interventional endoscopy.

REFERENCES

1. Hopkins, H.H. The development of the modern endoscope. *Nat. News* **1980**, *17* (8), 18–22.
2. Sivak, M.V. *Gastroenterologic Endoscopy*; W.B. Saunders Co: Philadelphia, 2000.
3. Hecht, E. *Optics*, 2nd Ed.; Addison-Wesley: Reading, MA, 1987.
4. Reedy, R.P. Coherent fiber optics test techniques. *Optical Eng.* **1980**, *19* (4), 556–560.
5. Goldstein, A. Physics of ultrasound. In *Diagnostic Ultrasound*; Rumack, C.M., Wilson, S.R., Charboneau, J.W., Eds.; Mosby-Year Book: St. Louis, 1991; 2–18.

6. Menze, J.; Domschke, W.; Brambs, H.J.; Frank, N.; Hatfield, A.; Natterman, C.; Odegaard, S.; Seifert, H.; Tamada, K.; Tio, T.L.; Foerster, E.C. Miniprobe ultrasonography in the upper gastrointestinal tract: state of the art 1995, and prospects. *Endoscopy* **1995**, *28* (6), 508–513.
7. Adain, A.L.; Han-Chuan, T.; Casidy, M.; Schiano, T.D.; Liu, J.; Miller, L.S. High-resolution endoluminal sonography is a sensitive modality for the identification of Barrett's metaplasia. *Gastrointest. Endosc.* **1997**, *46* (2), 147–151.
8. Gong, F.; Swain, P.; Mills, T. Wireless endoscopy. *Gastrointest. Endosc.* **2000**, *51* (6), 725–729.
9. Iddan, G.; Meron, G.; Glukhovskiy, A.; Swain, P. Wireless capsule endoscopy. *Nature* **2000**, *405* (6785), 417.
10. Kudo, S.; Tamura, S.; Nakajima, T.; Yamano, H.; Kusaka, H.; Watanabe, H. Diagnosis of colorectal tumorous lesions by magnifying endoscopy. *Gastrointest. Endosc.* **1996**, *44* (1), 8–14.
11. Siegel, L.M.; Stevens, P.D.; Lightdale, C.J.; Green, P.H.R.; Goodman, S.; Garcia-Carrasquillo, R.J.; Rotterdam, H. Combined magnification endoscopy with chromoendoscopy in the evaluation of patients with suspected malabsorption. *Gastrointest. Endosc.* **1997**, *46* (3), 226–230.
12. Jaramillo, E.; Watanabe, M.; Befrits, R.; Ponce de León, E.; Rubio, C.; Slezak, P. Small, flat colorectal neoplasias in long-standing ulcerative colitis detected by high-resolution video endoscopy. *Gastrointest. Endosc.* **1996**, *44* (1), 15–22.
13. Gono, K.; Obi, T.; Yamaguchi, M.; Ohyama, N.; Machida, H.; Sano, Y.; Yoshida, S.; Hamamoto, Y.; Endo, T. Appearance of enhanced tissue features in narrow-band endoscopic imaging. *J. Biomed. Opt.* **2004**, *9* (3), 568–577.
14. Axelrad, A.M.; Fleischer, D.E.; Geller, A.J.; Nguyen, C.C.; Lewis, J.H.; Al-Kawas, F.H.; Avigan, M.I.; Montgomery, E.A.; Benjamin, S.B. High-resolution chromoendoscopy for the diagnosis of diminutive colon polyps: implications for colon cancer screening. *Gastroenterology* **1996**, *110* (4), 1253–1258.
15. Kumagai, Y.; Monma, K.; Kawada, K. Magnifying chromoendoscopy of the esophagus: in-vivo pathological diagnosis using an endocytoscopy system. *Endoscopy* **2004**, *36* (7), 590–594.
16. Kara, M.A.; Peters, F.P.; ten Kate, F.J.W.; van Deventer, S.J.; Fockens, P.; Bergman, J.J.G.H.M. Endoscopic video autofluorescence imaging may improve the detection of early neoplasia in patients with Barrett's esophagus. *Gastrointest. Endosc.* **2005**, *61* (6), 679–685.
17. Mayinger, B.; Neidhardt, S.; Reh, H.; Martus, P.; Hahn, E.G. Fluorescence induced with 5-aminolevulinic acid for the endoscopic detection and follow-up of esophageal lesions. *Gastrointest. Endosc.* **2001**, *54* (5), 572–578.
18. Huang, D.; Swanson, E.A.; Lin, C.P.; Schuman, J.S.; Stinson, W.G.; Chang, W.; Hee, M.R.; Flotte, T.; Gregory, K.; Puliafito, C.A.; Fujimoto, J.G. Optical coherence tomography. *Science* **1991**, *254* (5035), 1178–1181.
19. Rollins, A.M.; Ung-arunyawee, R.; Chak, A.; Wong, R.C.K.; Kobayashi, K.; Sivak, M.V., Jr.; Izatt, J.A. Real-time in vivo imaging of human gastrointestinal ultrastructure by use of endoscopic optical coherence tomography with a novel efficient interferometer design. *Opt. Lett.* **1999**, *24* (19), 1358–1360.
20. Corle, T.R.; Kino, G.S. *Confocal Scanning Optical Microscopy and Related Imaging Systems*; Academic Press: San Diego, California, 1996.
21. Wang, T.D.; Contag, C.H.; Mandella, M.J.; Chan, N.Y.; Kino, G.S. Confocal fluorescence microscope with dual-axis architecture and biaxial postobjective scanning. *J. Biomed. Opt.* **2004**, *9* (4), 735–742.
22. Kiesslich, R.; Burg, J.; Vieth, M.; Gnaendiger, J.; Enders, M.; Delaney, P.; Polglase, A.; McLaren, W.; Janell, D.; Thomas, S.; Nafe, B.; Galle, P.R.; Neurath, M.F. Confocal laser endoscopy for diagnosing intraepithelial neoplasias and colorectal cancer in vivo. *Gastroenterology* **2004**, *127*, 706–713.

Ergonomics

Brian N. Craig

Department of Industrial Engineering, Lamar University, Beaumont, Texas, U.S.A.



INTRODUCTION

Ergonomics is an interdisciplinary field that incorporates principles from engineering, anatomy, biomechanics, physiology, medicine, anthropometry, human factors, psychology, and safety. Ergonomics, simply stated, is the study of work. Ergo means work, and nomos means to study. More formally, ergonomics can be defined as the science and practice of designing systems to fit people. As a science, it involves the analysis of human performance capabilities, limitations, and other characteristics to develop human-system interface technology in the form of ergonomic design principles, guidelines, specifications, tools, and methods. As a practice, ergonomics applies human-system interface technology to the analysis, design, test, evaluation, standardization, and control of systems to reduce human error and improve human and system performance, safety, health, comfort, productivity, quality, efficiency, and quality of life.^[1]

Working to maximize the effectiveness, efficiency, and safety of human-system interfaces, ergonomics can be viewed as a science of performance. Therefore, ergonomics is also defined as the design/redesign of tools, tasks, equipment, environments, software interfaces, pedagogic approaches, and organizations that enhance performance with many of the principles contributing directly to contemporary business and production management philosophies as well as inventory management systems.

OVERVIEW

Ergonomics is most often recognized as a discipline related to the primary prevention of work-related musculoskeletal disorders (WMSDs); therefore, the purpose of this article is to acquaint the reader with the current status of ergonomics initiatives, discuss WMSD risk factors, discuss elements of an effective ergonomics program, and introduce various ergonomic assessment tools. Work-related musculoskeletal disorders, e.g., muscle strains of the back, shoulder, neck, etc., are generally associated with the manufacturing sector of the economy; however, workers in the service sector and office workers also experience musculoskeletal

discomfort. Therefore, this article will also include a section addressing the principles of office ergonomics.

MAGNITUDE OF WORK-RELATED MUSCULOSKELETAL DISORDERS

According to the Bureau of Labor Statistics (BLS), in excess of 4 out of 10 injuries and illnesses resulting in days away from work in 2001 were sprains or strains, which most often involve the back.^[2] The BLS definition of days away cases is defined as the sum of lost work day cases and restricted work day cases. Even though these trends have been decreasing, these findings are consistent with the previous 10 years of BLS data. The trunk, which includes the shoulder and back, was the most affected part of the body and accounted for 36.5% of the cases. Cases of overexertion and contact with objects and equipment remained the leading cause of a disabling condition.^[2] Also, according to the BLS, the median number of days away from work in 2001 for all cases was six days. At 25 days, carpal tunnel syndrome (CTS) is associated with the most median days away from work. At 18 days, repetitive motions is the most frequent event or exposure associated with a median number of days away from work. The wrist and shoulder injuries resulted in median days away of 13 and 12 days respectively.^[2]

It is estimated that 1 out of every 3 spent for workers' compensation costs is spent on WMSDs, which result in employers paying 15 billion in workers' compensation.^[3] Additional costs associated with these WMSDs, such as training new workers, employee absenteeism and turnover, lost productivity, and lost wages may result in a total cost of 45 billion per year.^[3]

CURRENT ERGONOMIC INITIATIVES

The Occupational Safety and Health Administration (OSHA) proposed a final rule for an Ergonomics Program Standard on November 14, 2000.^[3] The standard became effective on January 16, 2001. This standard, however, was short-lived and was rescinded by the United States Senate on March 6, 2001. Currently, OSHA has assembled a National Advisory Committee on Ergonomics and unveiled a Four-Pronged,

Comprehensive Approach on Ergonomics.^[4] OSHA's four-pronged approach provides for guidelines, enforcement, outreach and assistance, and research.

According to OSHA, ergonomic-related injuries/illnesses are on the decline, and it is the goal of OSHA to accelerate the decline.^[4] A major advantage of the four-pronged approach over the repealed Ergonomic Program Standard is the emphasis on program flexibility and avoiding a one-size-fits-all approach.

Guidelines

OSHA is in the process of developing industry-specific guidelines for a number of industries that are based on current incidence rates and available information about effective and feasible solutions. In March 2003, OSHA released the first industry-specific Ergonomics Guidelines for Nursing Homes^[5] and is currently developing Ergonomics Guidelines for Retail Grocers and for Poultry Processing. In April 2003, OSHA announced plans to develop Ergonomics Guidelines for the Shipyard Industry.^[4] Furthermore, OSHA will encourage other industries to develop ergonomic guidelines to meet their own specific needs.

Enforcement

The primary goal of OSHA is the reduction of injuries and illnesses in the workplace, and, under the OSH Act's General Duty Clause, employers must keep their workplaces free from recognized serious hazards, including ergonomic hazards. According to OSHA, enforcement efforts will not be focused on employers who have implemented effective ergonomic programs or are making good-faith efforts to reduce ergonomic hazards. However, OSHA will conduct inspections for ergonomic hazards and issue citations under the General Duty Clause and issue ergonomic hazard alert letters when appropriate.^[4]

Outreach and Assistance

OSHA will provide assistance to businesses, particularly small businesses, and help them proactively address ergonomic issues in the workplace. OSHA will also provide advice and training on the voluntary guidelines and implementation of a successful ergonomics program. OSHA is developing a set of compliance assistance tools, including Internet-based training and information, to support understanding of guidelines and how to proactively define and address ergonomic problems (Table 1).

Research

OSHA will work closely with the National Institute for Occupational Safety and Health (NIOSH) and through the National Occupational Research Agenda (NORA) process to encourage ergonomic research in needed areas.

For additional information, visit <http://www.cdc.gov/niosh/homepage.html> and <http://www2.cdc.gov/NORA/default.html>. While there is a large body of research available on ergonomics, there are many areas in which additional research is necessary, including gaps identified by the National Research Council and the Institute of Medicine.^[6]

WORK-RELATED MUSCULOSKELETAL DISORDERS RISK FACTORS

Low Back

Due to the prevalence of low back pain and its associated cost, much of the ergonomic research has focused on the low back. Low back pain is an intricate interaction between the affected person and the multitude of potential risk factors that have been linked to occupationally related low back pain. The internal and external risk factors, whether personal, non-occupational, occupational, and/or psychosocial in nature, are dynamic, interdependent, and subject to debate. In addition, personal abilities, skills, attributes, buffers, and thresholds of individuals are also dynamic in nature. Thus, the dose of the injuring agent may vary widely from person to person.

Research has associated musculoskeletal disorders to personal risk factors, including low maximal aerobic power,^[7-12] high body fat/body mass index,^[8,13,14] low levels of flexibility,^[15-17] low strength capabilities,^[18-23] low levels of exercise/physical activity,^[21,24] smoking,^[13,25,26] low general overall muscular fitness,^[16,22,27,28] age, gender,^[13,29-31] and anthropometrics.^[32] However, there are also studies that have failed to demonstrate some of these associations.

In an occupational environment, workers may perform jobs that require high-frequency lifting and lowering, high forces, extended reaches, frequent bending and twisting along with static postures and carrying. There have been several investigations of the relationships manual materials handling may have with occupational musculoskeletal injuries, specifically low back injuries.^[13,33,34] Occupational injuries have been associated with high physical workloads,^[13,30,33,35,36] body motions and postures,^[13,37] and length of employment.^[38] In addition, many nonoccupational risk factors of equivalent magnitude exist at home, work, and play.^[39,40]

Table 1 OSHA's ergonomic outreach and assistance

Outreach and assistance	Examples	Description
OSHA Ergonomic eTools	Baggage handling Beverage delivery Computer workstations Grocery warehousing Healthcare Poultry processing Sewing	Stand-alone, interactive, Web-based information to address ergonomic hazards.
Cooperative Programs	Alliances	OSHA's Alliance Program is a tool for employers, trade or professional organizations, employees, and labor organizations to collaborate with OSHA in effectively sharing information on the usefulness and feasibility of their ergonomics programs within their own sectors.
	Onsite consultation program	OSHA's Onsite Consultation Program is a free service, particularly for small businesses involved in high-hazard industries or involved in hazardous operations.
	Strategic partnership	Organizations can enter into Strategic Partnerships with OSHA to address specific safety and health issues, including ergonomics.
	Voluntary partnership programs	Voluntary Partnership Programs recognizes and promotes effective safety and health management. VPP participants are facilities that have designed and implemented outstanding health programs.
Training and Education	OSHA training institute (OTI) and OTI Education Centers	Provides training in a wide variety of occupational safety and health topics.

There is increasing evidence between the relationship of job and work-related psychosocial risk factors and occupational injury/illness.^[41] Research indicates a 50% increase in health care expenditures for workers that report high levels of stress.^[42] Back pain and discomfort have been associated with job task dissatisfaction,^[36,43-45] psychological job demands,^[36] an employee's lack of social support from supervisors and/or colleagues,^[36,46-48] frequency of job problems,^[36] increased job pressure,^[38] monotonous work,^[38] and lack of job control.^[47]

Upper Extremity Disorders

Currently, the most common method to analyze the ergonomic exposures of a job for the upper extremities involves the use of generic risk factors. These generic contributing conditions include force, repetition, posture, vibration, localized compression, and other considerations.

Not all WMSDs are related to work activities. Other factors such as personal characteristics, nonoccupational, and psychosocial factors have also been associated with ergonomic-related injuries and illnesses.

However, when analyzing jobs or work tasks that may be associated with upper extremity disorders, conditions most commonly considered consider include, but are not limited to:

1. Awkward Postures, e.g., prolonged work with hands above the head or with the elbows above the shoulders; prolonged work with the neck bent; squatting, kneeling, or lifting; handling objects with back bent or twisted; repeated or sustained bending or twisting of wrists, knees, hips or shoulders; forceful and repeated gripping or pinching.
2. Forceful Lifting, Pushing/Pulling, and Carrying, e.g., handling heavy objects; moving bulky or slippery objects; assuming awkward postures while moving objects.
3. Prolonged Repetitive Motion, e.g., keying/keyboarding; using tools or knives; packaging, handling, or manipulating objects.
4. Contact Stress, e.g., repeated contact with hard or sharp objects, such as desk or table edges.
5. Vibration, e.g., overuse of power, pneumatic, or impact hand tools.
6. Other Factors, e.g., cold temperatures and poorly fitting gloves.

ELEMENTS OF AN ERGONOMICS PROGRAM

There are several key elements of a successful ergonomics program. The first and most important is true management commitment up to and including integration of ergonomics into production and management systems. The following provides the minimal elements of a successful ergonomics program.^[49]

1. Management commitment.
2. Employee participation.
3. Training and education.
4. Job hazard identification, including design, redesign, worksite analysis, and prioritization.
5. Job hazard prevention and control (implementation).
6. Follow-up and continual improvement.
7. Medical management and health promotion.

For additional information on ergonomic program effectiveness, the reader is encouraged to review McSweeney et al., "Ergonomic Program Effectiveness: Ergonomics and Medical Intervention."^[50] For more information about Elements of an Ergonomics Program, download the NIOSH document "Elements of an Ergonomics Program" at <http://www.cdc.gov/niosh/ephome2.htm>.

ERGONOMIC ASSESSMENT TOOLS

Although there are numerous job hazard analysis techniques offered in the literature, OSHA chose to include eight job hazard analysis tools in Appendix D of the Ergonomics Program Rule.^[3] These tools included the revised NIOSH Lifting Equation, the psychophysical tables published by Snook and Ciriello, the Strain Index, Appendix B of the Washington State Ergonomics standard, Rapid Upper Limb Assessment (RULA), Rapid Entire Body Assessment (REBA), the General Motors risk factor checklist, and the threshold limit value (TLV) for hand-arm vibration. The first four of these tools are briefly presented in this article along with the University of Michigan's Three-Dimensional Static Strength Prediction Program[™] (3DSSPP) along with checklists and surveys.

The 1991 Revised NIOSH Lifting Equation

The 1991 Revised NIOSH lifting equation is an improvement over the 1981 NIOSH lifting equation. Compared to the 1981 lifting equation, the 1991 equation allows for evaluation of asymmetrical lifting tasks, accounts for lifts with less than optimal hand-object couplings, and provides guidelines for a larger range of

work durations and lifting frequencies.^[9] It is also based on previous physiological, psychophysical, biomechanical, and epidemiological research.

The equation for calculating the recommended weight limit (RWL) is based on a multiplicative model that provides a weighting for each task variable. The weightings are expressed as coefficients that are used in the model to decrease the recommended load constant, i.e., maximum load weight to be lifted. The revised lifting equation established six multipliers that range from zero to one and the load constant as 51 pounds. The six multipliers include the horizontal multiplier (HM), vertical multiplier (VM), distance multiplier (DM), frequency multiplier (FM), asymmetry multiplier (AM), and coupling multiplier (CM). These six multipliers and load constant result in the recommended weight limit (RWL), as shown in the following equation.

$$RWL = (Load\ Constant) \times (HM) \times (VM) \\ \times (DM) \times (FM) \times (AM) \times (CM)$$

The RWL value is used to determine the lifting index (LI). The LI is the ratio of the actual load weight to the RWL. It is the LI that provides a relative estimate of the physical stress associated with manual lifting jobs and is given in the following equation:

$$LI = Load\ Weight / RWL$$

Whether a LI greater than one constitutes a risk is still a subject for debate. Waters et al.^[9] note that it is not possible to quantify the precise degree of risk associated with the increments in the LI. There is uncertainty about whether a LI of one is a reliable threshold for differentiating between an increase in risk and no increase in risk for some fraction of the working population.^[9] Despite the lack of research studies, it is reported that lifting tasks with a LI more than one poses an increased risk for lifting-related low back pain for some fraction of the workforce.^[9] The LI is considered a useful tool for prioritization for job redesign and job design.

The revised lifting equation continues to have limitations, including assumptions of two-handed, smooth, continuous lifting motion; unrestricted lifting posture; adequate foot traction; moderate ambient environment; lifting stable objects; and an object length less than 25 inches. It assumes that a smooth continuous lowering motion can be treated as lifting, other manual handling activities are minimal, and no unexpected heavy loads are present. It also assumes the worker is not lifting while carrying, pushing, or pulling, and the lifting is not done while seated or kneeling.

For more information about the scientific construct and precise use of the revised lifting equation, review the 1994 *Applications Manual for the Revised NIOSH*

Lifting Equation,^[51] or visit the Web site at <http://www.cdc.gov/niosh/pdfs/94-110.pdf>.

Snook & Ciriello's Liberty Mutual Maximum Acceptable Weights and Forces Tables

The Liberty Mutual Research Center has attempted to develop guidelines for the evaluation and design of manual handling activities based on a psychophysical approach.^[52] The results of the experiments conducted by Liberty Mutual have been compiled into a set of tables that depict gender-specific maximum acceptable weights and forces for lifting, lowering, pushing, pulling, and carrying tasks.^[52] The psychophysical approach refers to the method of conducting laboratory studies in which workers control an object's weight or force according to their perceptions of acceptable levels of exertion or fatigue. The worker can adjust the weight or force of the object according to his or her own capabilities and limitations, which thereby establish a maximal acceptable weight of lift, lower, or carry, or maximal acceptable push/pull force.

These maximum acceptable weight tables are intended for job design and hazard assessment.^[52] Whenever the strength requirements of a job exceed the maximum acceptable weights/forces for the worker, the worker may be at increased risk for overexertion injury. Although use of the maximum acceptable weights and forces tables has not undergone significant validation, Snook and associates estimate that up to one-third of compensable low back pain incidents may have been related to performing exertions that exceeded the maximum acceptable weight of 75 percent of females.^[53] Although these tables were derived from data gathered from simulated industrial tasks, such as lift, lower, push, pull, and carry, there is currently no data for twisting and lifting/lowering, type of grasp, duration of job (one hour versus eight hours), one-handed lifting, or sitting and lifting. For more information, visit the following Web site at <http://hsc.usf.edu/~tbernard/ergotools/#oshasuggested>.

The Strain Index

The Strain Index evaluates a job's level of risk for developing a disorder of the distal upper extremity (DUE).^[54] The DUE is defined as the hand, wrist, forearm, and elbow. Through a semiquantitative method, jobs are evaluated by analysis of six task variables, including intensity of exertion, duration of exertion, exertions per minute, hand/wrist posture, speed of work, and duration per day.

The Strain Index involves the direct measurement of duration of exertion, efforts per minute, and duration per day with estimation of intensity of exertion, hand/

wrist posture, and speed of work. The output of the Strain Index method is an SI score. Using the Strain Index involves five steps: 1) collect data on the six task variables; 2) assign ordinal ratings for each task variable (based on their data); 3) determine the multiplier values for each task variable (based on their individual ratings); 4) calculate the SI score (the product of the six multiplier values); and 5) interpret the result.^[54] The right and left sides are analyzed separately.

Interpretation of the SI score is the last step. According to the research, data suggest that an SI score of 5.0 best discriminates between jobs with and jobs without a history of workers developing DUE disorders.^[54-57] An SI of 5.0 can be considered a point value, but the authors suggest the following SI score interpretation. Strain Index scores less than or equal to 3.0 are probably safe. SI scores greater than 3.0, but less than 7.0 may be hazardous. SI scores greater or equal to 7.0 are probably hazardous.^[54] The Strain Index does not consider stresses related to localized mechanical compression and recommends this risk factor to be considered separately. For more information, visit the following Web site at <http://hsc.usf.edu/~tbernard/ergotools/#oshasuggested>.

Washington State Appendix B

The Washington State Department of Labor and Industries published its Ergonomics Rule in May 2000.^[58] Appendix B of that rule outlines specific criteria for analyzing jobs. The primary purpose of this job hazard analysis tool is to reduce employee exposure to specific workplace hazards that can cause or aggravate WMSDs. According to this rule, a WMSD hazard is a physical risk factor that exceeds the criteria in Appendix B.^[58]

Washington State's Appendix B is primarily a checklist to determine if employers possess jobs that are identified as caution zone jobs.^[58] Caution zone jobs are jobs or tasks in which an employee's typical work includes any of the following generic risk factors: awkward postures; high hand force; highly repetitive motions; repeated impact; heavy/frequent/awkward lifting; or moderate to high vibration.^[58] According to the rule, an employer has a choice of analyzing and reducing WMSD hazards using a general performance approach or a specific performance approach.^[58] The general performance approach allows employers to choose their own methods and criteria for analyzing jobs and reducing WMSDs as long as it is as effective as other recommended and widely used methods, such as the Strain Index, RULA, and REBA. If an employer chooses the specific performance approach, each caution zone job needs evaluation for specific physical risk factors that are specified in Appendix B of the ergonomics rule.

Washington State's Appendix B is based on epidemiological evidence of relationships between physical risk factors and WMSDs.^[58] The predictive and external validity of the checklist and its components have not been reported. For more information about Washington State Appendix B, visit the Web site at <http://www.lni.wa.gov/wisha/regs/ergo2000/AppendixBfinal5-19-00.pdf>.

The University of Michigan's 3-D Static Strength Prediction Program

The University of Michigan's 3-D Static Strength Prediction Program™ (3DSSPP) is a job hazard analysis tool that predicts compressive forces acting on the lower lumbar discs.^[37] The 3DSSPP is a computer software program that predicts the proportion of the population that is capable of performing an exertion with reference to various static strength norms and identifies the joint(s) that limits the capacity to perform the task.^[37]

To use the 3DSSPP, a job analyst describes and observes a worker performing a job. The computer software input data includes: 1) the magnitude and direction of the load; 2) whether one or two hands are involved; 3) the general anthropometry of the people to be analyzed; and 4) the postural angles of the body.^[37] The output screen includes graphical and tabular predictions describing the percentage of the male and female populations that are expected to have sufficient static strength at each major joint, the L5/S1 disc compression forces exerted by men and women, and the static coefficient of friction required not to slip along with a statement as to whether the task being performed would result in a loss of forward or backward balance.^[37]

The back compression design limit (BCDL) of 770 pounds does not predict the probability of injury. The BCDL is a job design criterion level that is similar to NIOSH's recommended L5/S1 maximum disc compressive force criteria of 770 pounds.^[37] This design limit is set at 770 pounds because research has suggested that most workers lifting loads below the compressive force criteria were at little risk of lifting-related low back pain.^[9] The criteria for good job design are loads requiring less than 770 pounds of disc compression force and within the static strength capability of 75% of female workers and 99% of male workers.^[9]

Validation studies of the 3DSSPP have been conducted by Chaffin et al.^[59] and Chaffin and Erig.^[60] Research has suggested that the 3DSSPP appears to accurately predict the percent of the population that is capable of performing a large variety of different types of maximal static exertions using the population strength norms in the model,^[61] however, the reader is reminded that the physical capacities of workers vary

depending on other personal, occupational, non-occupational, and psychosocial factors not entered in the model. For more information on the biomechanical basis for low back pain, the reader is encouraged to review Garg, "Occupational Biomechanics and Low-Back Pain."^[62]

Checklists and Surveys

A common tools used for conducting exposure assessments are workstation checklists and body part discomfort surveys. Even though subjective are not known to predict risk of injury, workstation checklists tools provide a quick method of evaluating workstation design, work postures, and work activities. Similarly, even though body part discomfort surveys are subjective and not specific to discomfort arising from the work environment, they may be helpful in some circumstances.

Checklists

In general, checklists are used to itemize certain aspects of the work environment for comparison to recommended criteria. In their most basic form, these checklists simply address the issue of whether or not a particular workstation incorporates specific objects or designs. For example, a line item on a checklist may be as straightforward as "Is Your Seat Pan Height Adjustable? (Yes or No)", or "Does the Operator Perform Intensive Keying for Four or More Hours per Day? (Yes or No)." These basic checklists are primarily initial screening tools used to determine if there needs to be a more detailed workstation exposure assessment. There are numerous checklists available to assess a variety of work environments.^[49]

It is important to remember that checklists are a limited collection of questions and criteria. No single checklist can anticipate the wide variety of complex situations that may be encountered in the work environment. Therefore, checklists should be considered simple screening tools that are used to assess current conditions of the workplace. They do not predict a worker's risk of developing a WMSD.

Surveys

Body part discomfort surveys ask workers to rate the intensity of discomfort for various body parts, such as the feet, lower legs, hands, wrist, lower back, and neck, on a categorical or continuous scale. These results can be applied at an individual level to identify opportunities for workstation evaluation or intervention, or they can also be applied collectively to prioritize work areas

2. Rate the backrest cushion comfort

Unacceptable	Average	Excellent
--------------	---------	-----------

3. Rate the lumbar (low back) support

Unacceptable	Average	Excellent
--------------	---------	-----------

Fig. 1 Excerpt from chair survey.

that need more detailed ergonomic assessments. Other surveys may also be similarly constructed, such as in the case of chair comfort questionnaire. An excerpt from such a survey is given in Fig. 1.^[63]

OFFICE ERGONOMICS OVERVIEW

The purpose of this section is to provide the reader with general knowledge about ergonomic principles for office furniture and workstation design and layout. Some of the voluntary standards and guidelines available for designing proper office workstations will also be mentioned.

Adjustability

The key to worker performance, injury/illness reduction, and improved quality of work life in the office environment is adjustability of the chair and the workstation to accommodate most—if not all—of the users. Since there are variations in most human anatomical characteristics, designing for a range of body sizes and working postures instead of the average person allows for greater workstation adjustability.

Anthropometry is the science of anatomical human measurement. Many ergonomic textbooks provide tables of standardized anthropometric data for a variety of human characteristics.^[64,65] Anthropometric data includes distributional information regarding physical attributes for generalized populations. Examples of such data include standing and seated stature, shoulder and hip breadth, buttock-knee length, and hand span. The design of an ergonomic chair is one example of applied anthropometry. A well-designed ergonomic chair should have the following adjustments to accommodate the 5th percentile female up to the 95th percentile male:^[66,67]

1. Seat pan height adjustment with waterfall front.
2. Seat depth adjustment.
3. Backrest height adjustment.
4. Seat pan angle adjustment, which includes tension adjustment.

5. Backrest angle adjustment that is independent of the seat pan angle.
6. Armrest width adjustment.
7. Armrest height adjustment.
8. Armrest rotation.
9. Adjustable lumbar support.

Furniture and equipment that comprise the modern office workstation include—but are not limited to—the work surfaces, chairs, footrests, monitors, and input devices. The successful design of a workstation depends on the proper interface of the furniture and equipment with the person and his or her work tasks.^[65] Several textbooks cover specific design criteria such as ranges of adjustability; equipment locations, such as heights, distances, etc.; clearances; and zones of reaching.^[65,66]

Generally, when a worker's hands are on the keyboard, the elbow should be in an angle slightly greater than 90° and his or her hands are aligned with the forearm with minimal wrist deviation.^[66] The backrest should conform to the contour of the lower back. The operator is seated all the way back in the chair to properly support the back.^[66] The operator's neck should be comfortably erect approximately 15° forward flexed. The knees are slightly lower than the hips, and the trunk-thigh angle is slightly greater than 90°.^[66] The feet should remain flat on the floor or supported by a footrest.^[66]

The desk should be a thin work surface in order to provide necessary thigh clearance and should preferably have an adjustable surface height with adequate area and depth to support the operator's computer(s) with all needed peripherals, books, telephone, files, etc., at the correct height, orientation to the operator, and distance from the operator. The monitor, depending on its size, should be placed directly in front of the operator on the desk surface at approximately 18–30 inches from the operator. The top of the viewable screen is no higher than the operator's horizontal eye height.

The keyboard should be directly in front of the operator with the pointing device immediately adjacent to the keyboard and located at the same height as the keyboard. To assist in avoiding eyestrain, the operator should make sure that his or her vision is checked by an optometrist or ophthalmologist. Also, light sources should primarily be indirect and nonglare producing. Finally, each operator should periodically rest his or her eyes by looking into the distance. In addition, telephone headsets and document holders should be provided if needed.^[66]

Environmental design includes principles related to lighting, noise, and temperature. There are numerous textbooks and guidelines the reader can consult for more information on office ergonomic principles and guidelines related to these environmental factors.^[64,65,67,68]

Guidelines and Standards

There have been several efforts to generate guidelines for ergonomic office environments, most notably ANSI/HFS 100-1988^[67] and the BSR/HFES 100-2002 (Draft).^[68] These documents provide voluntary guidelines for office environments and offer detailed criteria for workstation design, chair design, general office layout, and monitor features. The purpose of these guidelines is to maximize comfort. It should be noted that, due to the age of the 1988 guideline, many details contained in that document may not be necessarily consistent with current knowledge, theories, or issues regarding the design of office equipment. The 2002 document is still in draft form. However, these guidelines do provide a good starting point for implementing ergonomic designs in an office environment. Copies of ANSI/HFS 100-1988 and the BSR/HFES 100-2002 (Draft) can be obtained on the HFES Web site at <http://hfes.org/Publications/ANSI-HFS-100.html> and <http://hfes.org/publications/HFES100.html>, respectively.

CONCLUSION

This article provides the reader with an understanding of the current status of ergonomics initiatives, provides a general discussion on WMSD risk factors, discusses elements of an effective ergonomics program, and introduces several ergonomic assessment tools. The key principles of office ergonomics are also presented. In addition, numerous references throughout the article and a list of 20 additional references are given at the end of this section. Even though many of the major topics were covered, many topics were not covered. Many of these topics can be found in the references throughout the text as well as in the following list:

1. <http://www.osha.gov/SLTC/ergonomics/index.html>
2. <http://www.cdc.gov/niosh/ergopage.html>
3. <http://www.lni.wa.gov/wisha/ergo/default.htm>
4. <http://www.hfes.org>
5. <http://www.bcpe.org>
6. <http://www.ergoweb.com>
7. <http://www.engin.umich.edu/dept/ioe/C4E/index.html>
8. <http://ergocenter.tamu.edu/nursing/index.html>
9. <http://dept.lamar.edu/industrial/index.html>
10. <http://ergo.human.cornell.edu/>
11. <http://www.louisville.edu/speed/ergonomics/>
12. <http://osuergo.eng.ohio-state.edu/>
13. <http://ergo.human.cornell.edu/ahREBA.html>
14. <http://ergo.human.cornell.edu/ahRULA.html>
15. <http://www.eagle.org/rules/downloads/86-Ergo.pdf>
16. <http://www.eagle.org/rules/downloads/102-HAB.pdf>
17. <http://www.eagle.org/rules/downloads/103-COMF.pdf>
18. <http://www.eagle.org/rules/downloads/105-OFFHAB.pdf>
19. Konz, S. and Johnson, S., *Work Design: Industrial Ergonomics* (5th Edition), Holcomb Hathaway Publishers, Scottsdale, AZ, 2000.
20. Sanders, M. and McCormick, E., *Human Factors in Engineering and Design* (7th ed.), McGraw-Hill, New York, 1993.

ARTICLES OF FURTHER INTEREST

Eye Tracking: Characteristics and Methods; Eye Tracking: Research Areas and Applications; Gait Analysis; Medical Robotics; Telemedicine

REFERENCES

1. Human Factors and Ergonomics Society. *Human Factors and Ergonomics Society Directory and Yearbook; 2002–2003*. Human Factors and Ergonomics Society: Santa Monica, CA.
2. Bureau of Labor Statistics. *Lost-Worktime Injuries and Illnesses: Characteristics and Resulting Days Away from Work, 2001*; United States Department of Labor: Washington, DC, March 2002.
3. Occupational Safety and Health Administration. *Ergonomics Program; Final Rule, 29 CFR 1910, Subpart W, Section 1910.900 Ergonomics Program Standard, Federal Register #: 65:68261-68870*; United States Department of Labor: Washington, DC, 2000.
4. Occupational Safety and Health Administration. *OSHA Effective Ergonomics: Strategy for Success*; U.S. Department of Labor. <http://www.osha.gov/SLTC/ergonomics/index.html> (accessed April 2003).
5. Occupational Safety and Health Administration. *Guidelines for Nursing Homes; Ergonomics for the Prevention of Musculoskeletal Disorders*; U.S. Department of Labor. http://www.osha.gov/ergonomics/guidelines/nursing_home/final_nh_guidelines.pdf (accessed April 2003).
6. National Research Council and the Institute of Medicine. *Musculoskeletal Disorders and the Workplace: Low Back and Upper Extremities*; National Academy Press: Washington, DC, 2001. Panel on Musculoskeletal Disorders and the Workplace. Commission on Behavioral and Social Sciences and Education.
7. National Institute for Occupational Safety and Health. *A Work Practices Guide for Manual Lifting*, NIOSH Technical Report No. 81-122; US Department of Health and Human Services: Cincinnati, OH, 1981.
8. Craig, B.; Congleton, J.; Kerk, C.; Lawler, J.; McSweeney, K. Correlation of injury occurrence data with estimated maximal aerobic capacity and body

- composition in a high-frequency manual materials handling task. *Am. Ind. Hyg. Assoc. J.* **1998**, *59*, 25–33.
9. Waters, T.; Putz-Anderson, V.; Garg, A.; Fine, L. Revised NIOSH equation for the design and evaluation of manual lifting tasks. *Ergonomics* **1993**, *36*, 749–776.
 10. National Institute for Occupational Safety and Health. *Los Angeles Times*, DHHS (NIOSH) Publication No. HETA 90-013-2277; U.S. Department of Health and Human Services: Cincinnati, OH, 1993.
 11. Garg, A.; Saxena, U. Effects of lifting frequency and technique on physical fatigue with special reference to psychophysical methodology and metabolic rate. *Am. Ind. Hyg. Assoc.* **1979**, *40*, 894–903.
 12. Garg, A.; Chaffin, D.; Herrin, G. Prediction of metabolic rates for manual materials handling jobs. *Am. Ind. Hyg. J.* **1978**, *38*, 661–674.
 13. Tsai, S.; Gilstrap, E.; Cowles, S.; Waddell, L.; Ross, C. Personal and job characteristics of musculoskeletal injuries in an industrial population. *J. Occup. Environ. Med.* **1992**, *34*, 606–612.
 14. Jones, B.; Bovee, M.; Knapik, J. Associations Among Body Composition, Physical Fitness, and Injury in Men and Women Army Trainees. In *Body Composition and Physical Performance*; Marriott, B., Ed.; National Academy Press: Washington, DC, 1992; 141–173.
 15. Moore, T. A workplace stretching program: Physiologic and perception measurements before and after participation. *J. Am. Assoc. Occup. Health Nurses* **1998**, *46*, 563–568.
 16. American College of Sports Medicine. *Guidelines for Exercise Testing and Prescription*; Lea & Febiger: Philadelphia, PA, 1995.
 17. Stone, M. Muscle conditioning and muscle injuries. *Med. Sci. Sports Exerc.* **1990**, *22*, 457–462.
 18. Reimer, D.; Halbrook, B.; Dreyfuss, P.; Tibiletti, C. A novel approach to pre-employment worker fitness evaluations in a material-handling industry. *Spine* **1994**, *19*, 2,026–2,032.
 19. Cady, L.; Thomas, P.; Karwasky, R. Program for increasing health and physical fitness in firefighters. *J. Occup. Med.* **1985**, *2*, 269–272.
 20. Keyserling, M.; Herrin, G.; Chaffin, D. Isometric strength testing as a means of controlling medical incidents on strenuous jobs. *J. Occup. Med.* **1980**, *22*, 332–336.
 21. Cady, L.; Bischoff, D.; O'Connell, E. Strength and fitness and subsequent back injuries in firefighters. *J. Occup. Med.* **1979**, *4*, 269–272.
 22. Chaffin, D.; Herrin, G.; Keyserling, M. Pre-employment strength testing: An updated position. *J. Occup. Med.* **1978**, *20*, 403–408.
 23. Chaffin, D. Human strength capabilities and low back pain. *J. Occup. Med.* **1974**, *16*, 505–511.
 24. Hoffman, J.; Chapnik, L.; Shamis, A.; Givon, U.; Davidson, B. The effect of leg strength on the incident of lower extremity overuse injuries during military training. *Mil. Med.* **1999**, *164*, 153–156.
 25. McSweeney, K.P.; Congleton, J.J.; Kerk, C.J.; Jenkins, O.E.; Craig, B.N. Correlation of recorded injury and illness data with smoking, and absolute aerobic capacity. *Int. J. Ind. Ergon.* **1999**, *24*, 193–200.
 26. Sacks, J.; Nelson, D. Smoking and injuries: An overview. *Prev. Med.* **1994**, *23*, 515–520.
 27. Mital, A.; Nicholson, A.; Ayoub, M. *A Guide to Manual Materials Handling*; Taylor & Francis: London, 1993.
 28. Biering-Sorensen, F. Physical measurements as risk indicators for low-back trouble over a one-year period. *Spine* **1984**, *9*, 106–119.
 29. Houtman, I.; Bongers, P.; Smulders, P.; Kompier, M. Psychosocial stressors at work and musculoskeletal problems. *Scand. J. Work, Environ. & Health* **1994**, *20*, 139–145.
 30. Berkowitz, S.; Feuerstein, M.; Lopez, M.; Peck, C. Occupational back disability in U.S. army personnel. *Mil. Med.* **1999**, *164*, 412–418.
 31. Ayoub, M.; Mital, A. *Manual Materials Handling*; Taylor & Francis: London, 1989.
 32. Smedley, J.; Egger, P.; Cooper, C.; Coggon, D. Manual handling activities and risk for low back pain in nurses. *J. Occup. Environ. Med.* **1995**, *52*, 160–163.
 33. Marras, W.; Lavender, S.; Leurgans, S. Biomechanical risk factors for occupationally related low back disorders. *Ergonomics* **1995**, *38*, 377–410.
 34. Spengler, D.; Bigos, S.; Martin, N.; Zeh, J.; Fisher, L.; Nachemson, A. Back injuries in industry. A retrospective study I. Overview and cost analysis. *Spine* **1986**, *11*, 241–245.
 35. Craig, B.; Congleton, J.; Kerk, C.; Amendola, A.; Gaines, W.; Jenkins, O. A prospective field study of the relationship of potential occupational risk factors with occupational injury/illness. *Am. Ind. Hyg. Assoc. J.* **2003**, *64*, 376–387.
 36. Krause, N.; Ragland, D.; Fisher, J.; Syme, S. Psychosocial job factors, physical workload, and incidence of work-related spinal injury/illness: A 5-year prospective study of urban transit operators. *Spine* **1998**, *23*, 507–516.
 37. Chaffin, D.; Andersson, G.; Martin, B. *Occupational Biomechanics*, 3rd Ed.; Wiley and Sons: New York, 1999; pp. 297–300, 339–347.
 38. Bernard, B.; Sauter, S.; Fine, L.; Petersen, M.; Hales, T. Job task and psychosocial risk factors for work-related musculoskeletal disorders among newspaper employees. *Scand. J. Work, Environ. & Health* **1994**, *20*, 417–426.
 39. Leino, P. Does leisure time physical activity prevent low back disorders? A prospective study of metal industry employees. *Spine* **1993**, *18*, 863–871.
 40. Thorbjornsson, C.; Alfredsson, L.; Fredriksson, K.; Koster, M.; Michelsen, H.; Vingard, E.; Torgen, M.; Kilbom, A. Psychosocial and physical risk factors associated with low back pain: A 24-year follow-up among women and men in a broad range of occupations. *J. Occup. Environ. Med.* **1998**, *55*, 84–90.
 41. National Institute for Occupational Safety and Health. *Musculoskeletal Disorders and Workplace Factors, A Critical Review of Epidemiological Evidence for Work-Related Musculoskeletal Disorders of the Neck, Upper Extremity, and Low Back*, U.S. Department of Health and Human Services (NIOSH) Publication No. 97-141, 1997.
 42. Goetzal, R.; Anderson, D.; Whitmer, R.; Oziminkowski, R.; Wasserman, J.; the Health Enhancement Research

- Organization (HERO) Research Committee. The relationship between modifiable health risks and health care expenditures: An analysis of the multi-employer HERO health risk and cost database. *J. Occup. Environ. Med.* **1998**, *40*, 843–854.
43. Bigos, S.; Battié, M.; Spengler, D.; Fisher, L.; Fordyce, W.; Hansson, T. A prospective study of work perceptions and psychosocial factors affecting the report of back injury. *Spine* **1991**, *16*, 1–6.
 44. Holmström, E.; Lindell, J.; Moritz, U. Low back and neck/shoulder pain in construction workers: Occupational workload and psychosocial risk factors. Part 2: Relationship to neck and shoulder pain. *Spine* **1992**, *17*, 672–677.
 45. Svensson, H.; Andersson, G. The relationship of low back pain, work history, work environment, and stress. *Spine* **1989**, *14*, 517–522.
 46. Bigos, S.; Spengler, D.; Martin, N.; Zeh, J.; Fisher, L.; Nachemson, A. Back injuries in industry: A retrospective study III. Employee-related factors. *Spine* **1986**, *11*, 252–256.
 47. Elo, A.; Ilmarinen, J.; Leppanen, A.; Martikainen, R. The differences in psychological work factors and work ability by musculoskeletal disorders among employees in the metal industry. *Adv. Occup. Ergon. Saf. I* **1996**, *2*, 281–285.
 48. Leino, P.; Hänninen, V. Psychosocial factors at work in relation to back and limb disorders. *Scand. J. Work, Environ. & Health* **1995**, *21*, 134–142.
 49. National Institute for Occupational Safety and Health. *Elements of Ergonomics Programs: A Primer Based on Workplace Evaluations of Musculoskeletal Disorders*, DHHS (NIOSH) Publication No. 97-117; U.S. Department of Health and Human Services: Cincinnati, OH, 1997.
 50. McSweeney, K.; Craig, B.; Congleton, J.; Miller, D. Ergonomic program effectiveness: Ergonomics and medical intervention. *Int. J. Occup. Saf. Ergon.* **2002**, *8*, 433–449.
 51. Waters, T.; Putz-Anderson, V.; Garg, A. Revised NIOSH Lifting Equation, NIOSH Publication No. 94-110; US Department of Health and Human Services: Cincinnati, OH, 1994.
 52. Snook, S.H.; Ciriello, V.M. The design of manual handling tasks: Revised tables of maximum acceptable weights and forces. *Ergonomics* **1991**, *34*, 1,197–1,213.
 53. Snook, S.; Campanelli; Hart, J. A study of three preventative approaches to low back injuries. *J. Occup. Med.* **1978**, *20*, 478–481.
 54. Moore, J.; Garg, A. The strain index: A proposed method to analyze jobs for risk of distal upper extremity disorders. *Am. Ind. Hyg. Assoc. J.* **1995**, *56*, 443–458.
 55. Moore, J.S.; Rucker, N.P.; Knox, K. Validity of generic risk factors and the strain index for predicting nontraumatic distal upper extremity morbidity. *Am. Ind. Hyg. Assoc. J.* **2001**, *62*, 229–235.
 56. Knox, K.; Moore, J.S. Predictive validity of the strain index in turkey processing. *J. Occup. Environ. Med.* **2001**, *43*, 451–460.
 57. Rucker, N.P.; Moore, J.S. Predictive validity of the strain index in manufacturing. *Appl. Occup. Environ. Hyg.* **2002**, *17*, 63–73.
 58. *Appendix B: Criteria for Analyzing and Reducing WMSD Hazards for Employers Who Choose the Specific Performance Approach*, WAC 296-62-05174; Washington State Department of Labor and Industries, May 2000.
 59. Chaffin, D.; Freivalds, A.; Evans, S. On the validity of an isometric biomechanical model of worker strength. *IIE Trans.* **1987**, *19*, 280–288.
 60. Chaffin, D.; Erig, M. Three-dimensional biomechanical static strength prediction model sensitivity to postural and anthropometric inaccuracies. *IIE Trans.* **1991**, *23*, 215–227.
 61. Chaffin, D.B. Static Biomechanical Modeling in Manual Lifting. In *The Occupational Ergonomics Handbook*; Karwowski, W., Marras, W.S., Eds.; CRC Press: New York, 1999; 933–944.
 62. Garg, A. Occupational Biomechanics and Low-Back Pain. In *Ergonomics: Low-Back Pain, Carpal Tunnel Syndrome, and Upper Extremity Disorders in the Workplace*; Moore, J.S., Garg, A., Eds.; Occupational Medicine: State of the Art Reviews, Hanley & Belfus, Inc.: Philadelphia, PA, 1992; 609–611.
 63. Craig, B.; Fillyaw, J.; Joubert, B. In *Analysis of Buttock-Thigh Pressure of Five High End Ergonomic Chairs: A Field Study*, Proceedings of the Institute of Industrial Engineers Research Conference, Institute of Industrial Engineers: Norcross, GA, May 2002.
 64. Kroemer, K.; Kroemer, H.; Kroemer-Elbert, K. *Ergonomics: How to Design for Ease and Efficiency*; Prentice Hall: Englewood Cliffs, NJ, 2000; 404–440.
 65. Eastman Kodak Company. *Ergonomic Design for People at Work*; Van Nostrand Reinhold: New York, 1983; Vol. 1, 12–63.
 66. Niebel, B.; Freivalds, A. *Methods, Standards, and Work Design*, 11th Ed.; McGraw-Hill: Boston, 2003; 187–195.
 67. *American National Standard for Human Factors Engineering of Visual Display Terminal Workstations (ANSI/HFS 100-1988)*; Human Factors Society, Ed.; Human Factors Society: Santa Monica, CA, 1988.
 68. *Human Factors Engineering of Computer Workstations (BSR/HFES 100-2002), Draft Standard for Trial Use*; Human Factors Society, Ed.; Human Factors Society: Santa Monica, CA, 2002.

Excitable Tissue, Electrical Stimulation Of

Dominique M. Durand

Neural Engineering Center, Department of Biomedical Engineering, and Department of Neurosciences, Case Western Reserve University, Cleveland, Ohio, U.S.A.

E

INTRODUCTION

Two of the most vital organs in the body, the brain/nervous system and the heart comprise excitable tissue. Patients with disorders in either of these organs can have severe deficits. Although rehabilitation and neural regeneration can provide some improvement,^[1,2] the amount of restoration provided to the patients has been limited. Electrical stimulation of cardiac tissue has been particularly successful to restore function in the heart with implanted pacemakers, and electrical stimulation of the nervous system can provide functional restoration to neurologically impaired individuals. By placing electrodes within excitable tissue and passing current through these electrodes, it is possible to activate pathways in the brain as well as cardiac and skeletal muscles.

Neural prostheses refer to applications for which electrical stimulation is used to replace a previously lost or damaged neural function. Electrical stimulation has been applied to restore neural function in several neural systems.^[3,4] The most successful neural prosthesis is the cochlear prosthesis. Electrical stimulation of the auditory nerves can restore hearing in deaf patients.^[5] Other applications include phrenic nerve stimulation for patients with high-level spinal cord injury to generate diaphragm contractions and can restore ventilation;^[6,7] and electrical stimulation of the visual cortex^[8] or the retina^[9] to restore vision and stimulation of upper/lower extremities.^[10] There are also several applications of electrical stimulation in the brain such as deep brain stimulation of thalamic nuclei to decrease tremor in patients with Parkinson's disease.^[11] Some applications are therapeutic instead of restorative, such as electrical stimulation of the vagus nerve to reduce seizure frequency in patients with epilepsy.^[12,13]

Stimulation of cardiac tissue has generated significant therapeutic benefits. A cardiac pacemaker is a stimulator capable of restoring the normal rhythm of the heart in patients suffering from cardiac disorders such as arrhythmia and heart block.^[14] The sophistication of pacemakers has increased dramatically in the recent years. Starting from the simple asynchronous (free running), synchronous (synchronized to the patient's heart beat) for patients with incomplete heart

block, rate-responsive pacemakers (frequency of pacing varies with metabolic demand) and pacemakers with defibrillators have been released.

BACKGROUND

Electrical excitation results from the interaction between extracellular electric fields and cellular membrane as described by the cable equation.^[15] The interaction can be analyzed by determining the voltages and electrical fields generated by the electrode using the solution of Maxwell equations. The relationship between the applied extracellular field and the transmembrane voltage is described by the source term in the cable equation and is derived below (section "Fundamental Principles of Electrical Stimulation"). The effect of waveform of the stimulation pulse, the electrochemistry taking place at the electrode interface and the tissue damage are also reviewed (section "Membrane-Electric Field Interaction"). In the last section ("Electrode Tissue Interface"), the stimulation of neural tissue is emphasized and stimulation of various other excitable tissues such as skeletal muscles and cardiac tissue are discussed.

FUNDAMENTAL PRINCIPLES OF ELECTRICAL STIMULATION

Anatomy and Physiology

Axonal excitation by applied current is effected when the transmembrane current generated by the electrode depolarizes the membrane sufficiently to activate the sodium channels located in high density within the cell membrane. Once activated, the sodium current will further increase until the membrane reaches an unstable fixed point. This is the point at which a full action potential (~ 100 mV) will develop. Once started, the action will propagate unattenuated either along nerve fibers or cardiac tissue.

Although the dynamics of membrane channels play a major role in the excitation properties of excitable cells, the sodium channels are almost completely closed

at the resting potential. Therefore, the passive properties of the membrane contribute significantly to the determination of the membrane voltage along the axons and to the site of the excitation. The transmembrane voltage is the difference between the intracellular voltage and the extracellular voltage. This extracellular voltage is generated by the current of the electrode and can be estimated with a reduced set of Maxwell equations, the quasistatic formulation.

Electric Fields in Volume Conductors

Although the volume conductor clearly contains material with a high dielectric constant (cell membrane for example), its capacitive properties can be generally neglected when the frequency is below 10 kHz.^[16] Similarly, the inductive properties can be neglected at these low frequencies. Therefore, the volume conductor surrounding the neural excitable tissue can be assumed to be purely resistive with a resistivity varying between 50 and 500 Ωcm in most applications. The resistivity of the volume conductor can vary at different locations or in different directions and is defined at each point of the volume by a resistivity tensor. If the resistivity is the same in all directions, the volume is isotropic. If the resistivity is the same at all points of the volume, then the volume is homogeneous. A simpler solution of the Maxwell equations can be derived for this type of volume conductor and applications.

Quasistatic formulation of Maxwell equations

For frequencies under 10 kHz, both the capacitive and inductive properties can be neglected and a simplified set of equations can be used to calculate the voltage and current in volume conductors.^[16]

$$\text{Conservation of charge: } \nabla \cdot \mathbf{J} = 0 \quad (1)$$

$$\text{Gauss's law: } \nabla \cdot \mathbf{E} = \rho/\varepsilon \quad (2)$$

$$\text{Ohm's law for conductors: } \mathbf{J} = \sigma\mathbf{E} \quad (3)$$

$$\text{Electric field: } \mathbf{E} = -\nabla V \quad (4)$$

where \mathbf{E} is the electric field (V/m) defined as the gradient of the scalar potential V , \mathbf{J} the current density (mA/m^2), the conductivity (inverse of resistivity) in S/m , ρ the charge density in C/m^3 , ε the permittivity of the medium (F/m).

Equivalence between dielectric and conductive media

The current density \mathbf{J} at any point is the sum of a source term \mathbf{J}_s and an ohmic term \mathbf{J}_Ω .

$$\mathbf{J} = \mathbf{J}_\Omega + \mathbf{J}_s = \sigma\mathbf{E} + \mathbf{J}_s \quad (5)$$

Using Eq. (1):

$$\nabla \cdot \mathbf{J} = \nabla \cdot (\sigma\mathbf{E}) + \nabla \cdot \mathbf{J}_s = 0 \quad (6)$$

Assuming an homogeneous volume conductor, $\nabla \cdot (\sigma\mathbf{E}) = \sigma(\nabla \cdot \mathbf{E})$ therefore:

$$\nabla \cdot \mathbf{E} = -\nabla \cdot \mathbf{J}_s/\sigma \quad (7)$$

$$\text{As } \mathbf{E} = -\nabla V,$$

$$\nabla^2 V = \nabla \cdot \mathbf{J}_s/\sigma = -I_v/\sigma \quad (8)$$

where I_v is a volume source current in A/m^3 and ∇^2 is the Laplacian operator. The volume current I_v can be calculated from the knowledge of the distribution of sources in the volume conductor. This equation is the equivalent of the Poisson equation derived for dielectrics:^[17]

$$\nabla^2 V = -\rho/\varepsilon \quad (9)$$

Using the following equivalence:

$$\rho \iff I_v$$

$$\varepsilon \iff \sigma$$

the solution of the Poisson equation for dielectric problems can be applied to the calculation of the current and voltage distribution in volume conductors.

Monopole point source

For the monopole source, the current density \mathbf{J} at any point in the medium located at a distance r from the source can be obtained simply and is equal to the total current crossing a spherical surface with radius r divided by the surface area

$$\mathbf{J} = \frac{I}{4\pi r^2} \mathbf{u}_r \quad (10)$$

where \mathbf{u}_r is the unit radial vector, and r the distance between the electrode and the measurement point. The electric field is then obtained from Eq. (3)

$$\mathbf{E} = \frac{I}{4\pi\sigma r^2} \mathbf{u}_r \quad (11)$$

The electrical field is the gradient of the potential. In spherical coordinates

$$\mathbf{E} = -\frac{dV}{dr}\mathbf{u}_r \quad (12)$$

The potential at point P measured with a reference electrode located at infinity is

$$V = \frac{I}{4\pi\sigma r} \quad (13)$$

Both the current density \mathbf{J} and the electric field \mathbf{E} have a radial distribution with an amplitude inversely proportional to the square of the distance to the source. The potential decay is inversely proportional to the distance. At $r=0$, the potential goes to infinity and this singularity can be eliminated if the electrode is modeled as a sphere with a radius a . Eq. (13) is then valid on the surface of the electrode $r=a$ and for $r>a$.^[18]

Assuming the medium to be homogeneous and linear, using superposition, Eq. (13) can be generalized to n monopolar electrodes with a current I_i located at a distance r_i from the recording point. The voltage is then given by

$$V = \frac{1}{4\pi\sigma} \sum_i \frac{I_i}{r_i} \quad (14)$$

For an axon located in the volume conductor at a distance d from the source I_s (Fig. 1A), the voltage along the axon is given by the following equation and is plotted in Fig. 1B ($d=1$ cm and $I_s=1$ mA)

$$V = \frac{1}{4\pi\sigma\sqrt{d^2 + x^2}} \quad (15)$$

The longitudinal component of the electrical field (E_x) is biphasic and plotted in Fig. 1C.

Bipolar electrodes and dipoles

Using Eq. (14), it is possible to calculate the voltage generated by a current source and a current sink separated by a known distance d . The potential at point P (assuming that the voltage reference is at infinity) is given by

$$V = \frac{1}{4\pi\sigma} \left(\frac{1}{r_1} - \frac{1}{r_2} \right) \quad (16)$$

where r_1 and r_2 are the distance between the measuring points and the electrodes. When the distance d between the two electrodes is small compared to the distance r ,

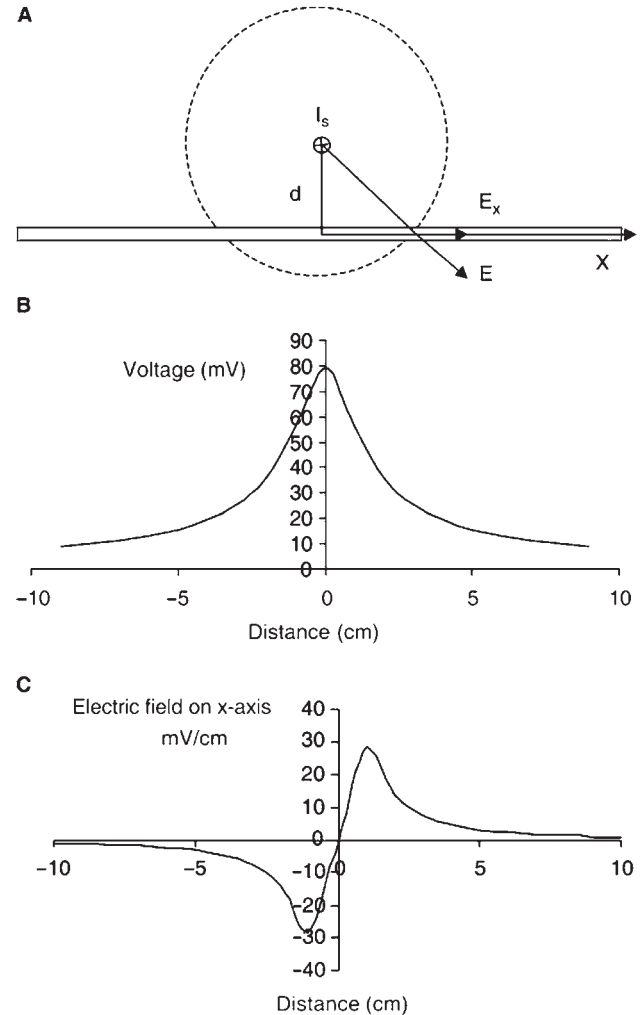


Fig. 1 Electrical stimulation of a fiber. (A) An electrode is located at a distance d from the axon and an anodic stimulus is applied to the electrode. The current flows outside and inside the axon generating depolarization of the membrane. Action potentials are generated underneath the electrode and propagate both orthodromically and antidromically. The fiber is activated by the longitudinal component (E_x) of the electric field \mathbf{E} . Extracellular voltage (B) and electric field (C) generated by 1 mA and measured along an axon located at 1 cm from the electrode (resistivity of the medium is assumed to be $100 \Omega \text{cm}$).

the equation for the potential reduces to

$$V = \frac{Id \cos \theta}{4\pi\sigma r^2} \quad (17)$$

where θ is the angle between the dipole axis and the point of measurement. There is an equipotential line ($\theta=0$) passing between the two electrodes. Therefore, an axon or nerve located along this line (transverse excitation) has a very high threshold for excitation. The voltage generated by a dipole is inversely

proportional to the square of the distance between the source and the recording site. Because the voltage decays rapidly compared to a monopole, the dipole is a much more selective method of stimulation. However, the current threshold for a dipole is significantly higher than for monopoles. The ratio of the monopole voltage to the dipole voltage is proportional to r/d and d is much smaller than r .

Inhomogenous volume conductors

Biological volume conductors are highly inhomogenous and the complexity of the volume conductors requires in most cases numerical solutions using finite element or finite boundary methods. However, the effect of the boundary between two layers of various conductivities can be studied in a simple configuration. Consider two volume conductors region 1 and 2 with conductivities σ_1 and σ_2 separated by an infinite plane. A monopolar electrode is placed in region 1 and potentials are recorded in the same region. Solving for the boundary conditions at the interface, it can be shown that the inhomogenous volume conductor can be replaced by a homogenous volume with conductivity σ_1 by adding a mirror current source with an amplitude equal to^[18]

$$I' = \frac{\sigma_1 - \sigma_2}{\sigma_1 + \sigma_2} I \quad (18)$$

The voltage is then given by Eq. (14) for region 1 only. In the case where layer 1 is a volume conductor such as the body and 2 is air ($\sigma_2 = 0$), the mirror source I' is equal to I . If the stimulation electrode is located at the interface, the zero-current boundary condition is satisfied by simply doubling the amplitude of the current. Similarly, if the recording electrode is located on the surface, the zero conductivity layer will double the size of the recorded potentials. If layer 2 is a perfect conductor ($\sigma_2 = \infty$), the current density must be normal at the boundary. The condition is satisfied if $I' = -I$. This image theory is only applicable in simple cases, but can be useful to obtain approximations when the distance between the recording electrode and the surface of discontinuity is small thereby approximating an infinite surface.^[19]

MEMBRANE-ELECTRIC FIELD INTERACTION

The interaction between the extracellular voltage and axons can be studied with a simple passive membrane model (Fig. 2A).

The membrane resistance is modeled as the rest conductance of the sodium, potassium, and leakage

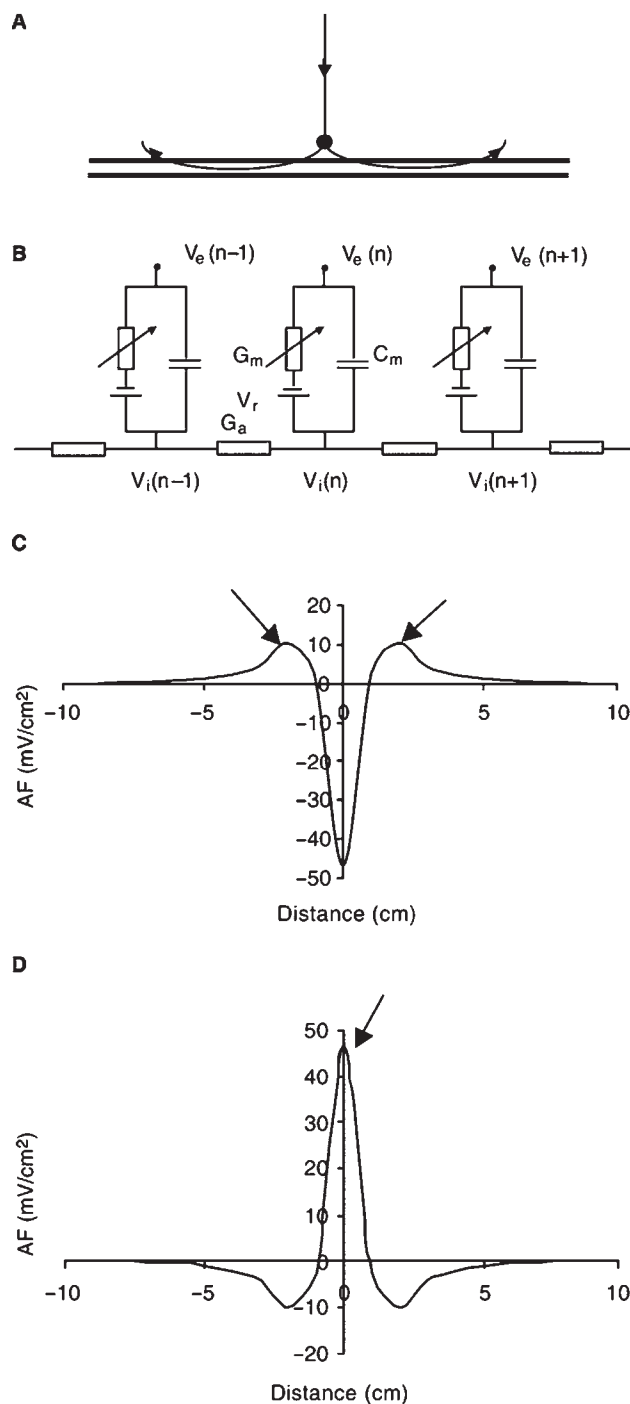


Fig. 2 Activating function for the electrical stimulation of a fiber: (A) The current lines generated by an anode enter the axon underneath the electrode. The current leaves the fibers in the distal regions. Electrical model of the membrane and axonal component of the fibers. (C) Activating function for an anode. The maximum positive value of the function shows the location of depolarization (excitation) of the fiber and is indicated by the two arrows. (D) For a cathode, the maximum value of the activating function is at the origin and is much greater than in (C). Therefore, cathodes have a lower threshold than anodes.

channels using compartmental analysis.^[20] Current flowing from the outside to the inside of the cell increases the voltage across the membrane and causes hyperpolarization. When current flows from the inside to the outside, the membrane is depolarized (Fig. 2).^[21]

A quantitative analysis of the interaction between electric fields and neural tissue can be derived by combining the passive membrane model with the extracellular voltage (V_e) generated by stimulating electrodes. The model makes several assumptions.^[22] Assuming that the presence of the fiber does not affect the extracellular voltage, the extracellular voltage can be calculated using the equations previously derived. A passive electrical model of the axon can be built and the circuit can be solved using numerical methods for compartmental analysis.^[23] These methods have been analyzed using neuronal simulation packages such as Neuron.^[24] Each compartment in Fig. 2B models a length Δx of the cell. Applying Kirchoff's law at each node and taking the limit when Δx goes to zero, one obtains the following inhomogenous cable equation^[25,26]

$$\lambda^2 \frac{\partial^2 V_m}{\partial x^2} - \tau_m \frac{\partial V_m}{\partial t} - V_m = -\lambda^2 \frac{\partial^2 V_e}{\partial x^2} \quad (19)$$

λ is the space constant of the fiber and is determined by the geometric and electric properties of the axon

$$\lambda = \frac{1}{2} \sqrt{\frac{R_m d}{R_a}} \quad (20)$$

where R_m is the specific membrane resistance, R_a the axoplasmic specific resistance, and d the diameter of the axon. The time constant of the axon is given by

$$\tau_m = R_m C_m \quad (21)$$

where C_m is the specific membrane capacitance. The source term of the cable equation is negative and is the product of the square of the space constant with the second spatial derivative of the extracellular voltage. At the onset of a pulse ($t=0$), the voltage on the cable is equal to zero and the change in voltage is proportional to the second-order spatial derivative of the extracellular voltage also known as the activating function $AF(x, t)$ ^[27]

$$\frac{\partial V_m}{\partial t} = \frac{\lambda^2}{\tau_m} \frac{d^2 V_e}{dx^2} \Big|_{t=0} = \frac{\lambda^2}{\tau_m} AF(x, t) \quad (22)$$

The amplitude of the activating function $AF(x, 0)$ is plotted in Fig. 2C and D for a 10 μm axon stimulated by a 1 mA anodic and cathodic current respectively

located 1 cm away from the axon. A positive value of $AF(x, 0)$ indicates membrane depolarization, while a negative value indicates hyperpolarization. The membrane polarization calculated from this equation can be predicted by simply examining the current flow pattern in and out of the axon. This analysis is valid only at the onset of the pulse, because during the pulse, currents will be distributed throughout the cable and will affect the transmembrane potential.^[28] However, for short pulses, these effects are small and the shape of the transmembrane voltage can be predicted using the activation function source. This analysis is valid for long axonal structures with uniform membrane. However, when cell bodies, dendrites, and more realistic axons are simulated, the site of excitation is strongly influenced by the distribution of sodium channels and a more detailed analysis is required to predict the site of excitation.^[29]

Activation of Myelinated Axons

The cable equation derived above needs to be modified for myelinated fibers because the presence of a myelin sheath around the axon forces the current to flow in and out of the membrane only at the nodes of Ranvier. Therefore, the action potential also jumps from node to node (Saltatory conduction). The interaction between myelinated fibers and applied fields can be described using the model shown in Fig. 1A with R_m and C_m replaced by R_n and C_n , respectively. In this simple model, the myelin sheath is assumed to have an infinite resistance. Therefore, R_n and C_n represent the membrane resistance and capacitance at the node of Ranvier only, respectively.

$$R_n = \frac{R_n^s}{\pi dl} \quad C_n = C_n^s \pi dl \quad (23)$$

where R_n^s and C_n^s are the specific membrane capacitance and resistance at the node, d is the inner fiber diameter, and l the width of the node. The axoplasmic resistance R_a is defined as:

$$R_a = \frac{4R_a^s L}{\pi d^2}$$

where R_a^s is the specific axoplasmic resistance and L , the internodal distance. The cable equation for a myelinated nerve can be derived using Kirchoff's law

$$\frac{R_n}{R_a} \Delta^2 V_m - R_n C_n \frac{\partial V_m}{\partial t} - V_m = -\frac{R_n}{R_a} \Delta^2 V_e \quad (24)$$

Δ^2 is the second difference operator ($\Delta^2 V = V_{n-1} - 2V_n + V_{n+1}$), and R_a represents the resistance between

two nodes. The source term for this equation

$$\left(-\frac{R_n}{R_a}\Delta^2 V_e\right) = -\frac{R_n^s}{R_a^s}\frac{d}{4\rho L}\Delta^2 V_e$$

is not explicitly dependant on the axon diameter because $d/L = d/D \times D/L$ ($= 0.007$), where D is the outside diameter of the axon and l is constant. However, because the distance between the node increases with the diameter of the fiber, the second-order difference does indeed depend on the of diameter of the axons.

Effect of Polarity of Applied Stimulus

Activation of axons is determined by the amplitude of the membrane depolarization. The polarity of the membrane polarization can be predicted directly from the sign of the activating function (positive sign indicates depolarization). This activation function is plotted in Fig. 2 and it is clear from the figure that the membrane depolarization (see arrows) generated by the cathode is greater than the depolarization generated from the anode (Figs. 2C and D). Note also that the polarity of the stimulus affects the location of the excitation. Cathodic excitation will produce excitation directly under the cathode, whereas anodic excitation will produce excitation at two sites located away from the anode (see arrows in Fig. 2D). However, the location of excitation also depends on the largest density of sodium channels. For example, anodic stimulation of cortical pyramidal cells has a lower threshold than cathodic stimulation, when applied to the surface of the brain.

Effect of the Space Constant of the Axon

The source term of the cable equation is proportional to the square of the space constant λ . λ^2 is proportional to the diameter of the fiber^[20] for unmyelinated fibers and to the square of the diameter for myelinated fibers.^[19] Therefore, in both cases, the source term is higher for fibers with larger diameters and, therefore, large diameter fibers have a lower threshold. As λ^2 is also dependent on the electrical properties of the axons, it is possible to predict that fibers with a larger membrane resistance or lower axoplasmic resistance will also have lower thresholds.

Recruitment Order of Fibers

The physiological recruitment order by the central nervous system (CNS) is to first recruit the small fibers followed by large ones. As previously indicated,

electrical stimulation recruits large fibers first. Therefore, electrical stimulation produces a recruitment order opposite to the physiological and this effect is known as the reverse recruitment order. Several techniques have been developed to recruit small fibers before large fibers. The first technique uses anodal block, whereby axons activated under the cathode are blocked when passing underneath the anode. This blocking effect requires a long pulse (about 400 μ sec) to maintain hyperpolarization of the membrane. By decreasing the amplitude of the current, the small diameter fibers are first released from the anodic block. Further reduction of the amplitude activates the larger fibers^[29] restoring a normal recruitment order. Another more recent technique places an array of contact along the nerve. When the contact separation of the electrodes is equal to the internodal distance of fiber of a given diameter, the activation function for that fiber is reduced and the fiber does not fire.^[30] By adjusting the polarity and the electrode distance, it is possible to reverse the recruitment order of the nerve activation.

Electrode–Tissue Interface

At the interface between the electrode and the tissue, the charge carrier must change between electrons in the metal and ions in the tissue. This exchange is carried out by chemical reactions that can modify the chemical milieu around the electrode, produce neural damage, and change the impedance. Moreover, the shape of the waveform can influence the threshold for activation as well as the corrosion of the electrode.

Effect of Stimulation Waveform on Threshold

Strength–duration curve

It has been known for a long time that it is the time-change in the applied current and not the continuous application of the external stimulus that can excite. Direct currents cannot excite and even in small amplitudes can cause significant tissue damage. It has also been observed experimentally that the relationship between the pulse width and the amplitude suggests that the total charge injected is the important parameter. The relationship between the amplitude and the width of a pulse required to bring an excitable tissue to threshold is shown in Fig. 3.

The amplitude of the current threshold stimulus (I_{th}) decreases with increasing pulse width (W) and can be modeled by the following relationship^[30]

$$I_{th} = \frac{I_{rh}}{1 - \exp\left(-\frac{W}{T}\right)} \quad (25)$$

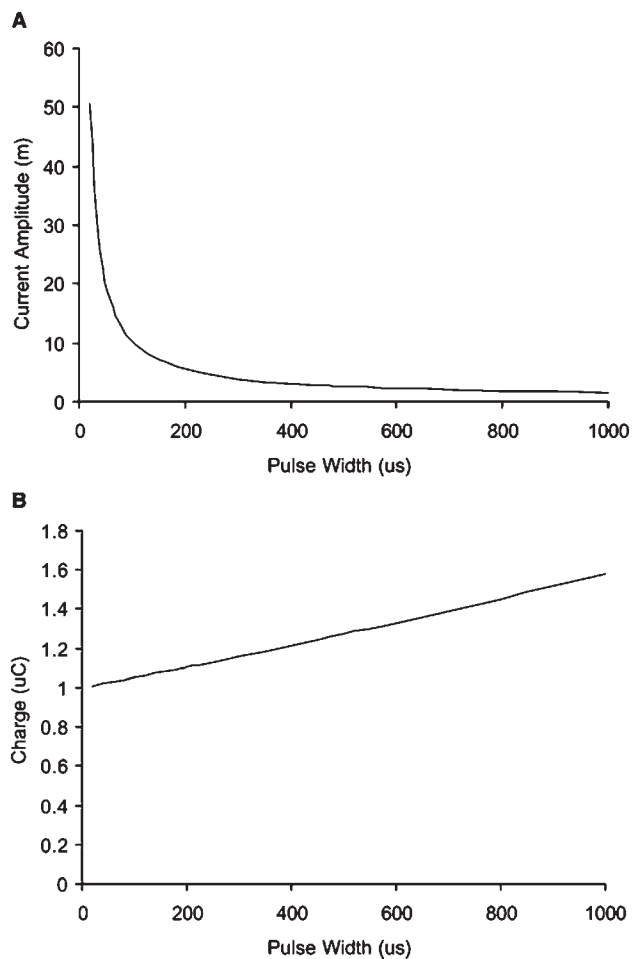


Fig. 3 Strength-duration (A) and charge-duration (B) curves for excitation of neurons. The time constant, chronaxie, and rheobase currents are 1 msec, 0.7 msec, and 1 mA, respectively.

The minimum current amplitude required to cause excitation is known as the rheobase current (I_{rh}). T is the membrane time constant of the axon if the axon is stimulated intracellularly. For extracellular stimulation, T is a time constant that takes into account the extracellular space resistance. The relationship between current amplitude and pulse width can also be derived theoretically using the cable equation by assuming that total charge on the cable for excitation is constant.^[31] The pulse width required to excite the tissue with an amplitude equal to twice the rheobase current is known as the chronaxie (T_{ch}) and is equal to $0.693T$. This strength duration curve can also be modeled by the following simpler equation

$$I_{th} = I_{rh}[1 + (T_{th}/W)]$$

Charge–duration curve

The threshold charge injected $Q_{th} = I_{th} \times W$ is plotted in Fig. 3B and increases with the pulse width.

$$Q_{th} = \frac{I_{rh} W}{1 - \exp(-\frac{W}{T})} \quad (26)$$

The increase in the amount of charge required to fire the axon with increasing pulse width is owing to the fact that for long pulse durations, the charge is distributed along the cable and does not participate directly to raising the membrane voltage at the excitation site. The minimum amount of charge Q_{min} required for stimulation is obtained by taking the limit of Q_{th} ^[32] when W goes to zero and is equal to $I_{rh} \times T$. In practice, this minimum charge can be nearly achieved by using narrow, current pulses. Short pulses have other advantages; they increase the spatial selectivity of stimulation^[33] and fiber diameter selectivity.^[32]

Anodic break

Excitation generated by cathodic current threshold normally takes place at the onset of the pulse. However, long-duration, subthreshold, cathodic or anodic current pulses have been observed experimentally to generate excitation at the end of the pulse. This effect has been attributed to the voltage sensitivity of the sodium channel. The sodium channel is normally partially inactivated at rest. However, when the membrane is hyperpolarized during the long duration pulse, the inactivation of the sodium channel is completely removed. Upon termination of the pulse, an action potential is generated because the inactivation gate has a slow time constant relative to the activation time constant and cannot recover fast enough.^[34] This effect can be observed with an anodic or cathodic pulse because both can generate hyperpolarization. Anodic break can be prevented by avoiding abrupt termination of the current. Pulse shapes with slow decay phases such as exponential or trapezoidal decay shapes have been successfully used.^[35]

ELECTROCHEMISTRY OF STIMULATION

Conduction in a metal is carried by electrons, while in the tissue, current is carried by ions. Although capacitive mechanisms have been explored, capacitive electrodes capable of storing enough charge for stimulation have not yet been developed. Therefore, most electrical stimulation electrodes rely on the Faradaic mechanisms at the interface between the metal and the tissue. Faradaic mechanisms require that

oxidation and reduction take place at the interface.^[36] Faradaic stimulation mechanisms can be divided into reversible and nonreversible mechanisms. Reversible mechanisms occur at or near the electrode potential and include oxide formation and reduction and hydrogen plating. Irreversible mechanisms occur when the membrane is driven far away from its equilibrium potential and include corrosion, hydrogen, or oxygen evolution. Those irreversible processes can cause damage to both the electrode and the tissue because they alter the composition of the electrode surface and can generate toxic products with pH changes in the surrounding tissue. During charge injection, the electrode potential is modified by an amount related to the charge density (total charge divided by the surface area). To maintain the electrode potential within regions producing only minimal irreversible changes, this charge density must be kept below maximum values.^[37] The maximum charge density allowed depends on the metal used for the electrode, the stimulation waveform, the type of electrode used, and the location of the electrode within the body.

Stimulation Waveform

Common biphasic waveforms for stainless steel or platinum use a cathodic pulse followed by an anodic phase. An example of a square balanced biphasic waveform is shown in Fig. 4 and can be easily implemented using a capacitor and switches. This waveform ensures that the charge is balanced because a capacitor is inserted in series with the tissue to be stimulated and the charge injected is then reversed by discharging the capacitor.^[38] Biphasic cathodic-first waveforms have a higher threshold than monophasic waveforms because the maximum depolarization induced by the cathodic pulse is decreased by the following anodic pulse.^[34] A delay can also be added between the cathodic and anodic phase as shown in Fig. 4. However, this time delay can also prevent adequate charge reversal and can be dangerous to the electrode and the tissue. An alternative method is to decrease the maximum amplitude of the anodic phase but increase its duration. This waveform can also be damaging to the electrode because the charge is not reversed fast enough following the cathodic pulse. The various waveforms have been ranked for their effect on tissue damage, corrosion, and threshold of activation.^[19,36]

Electrode Damage

Corrosion of the electrode is a major concern because it can cause pitting, metal dissolution, and tissue damage. Corrosion occurs at the anodic phase of the stimulation during oxidation. Therefore, by using

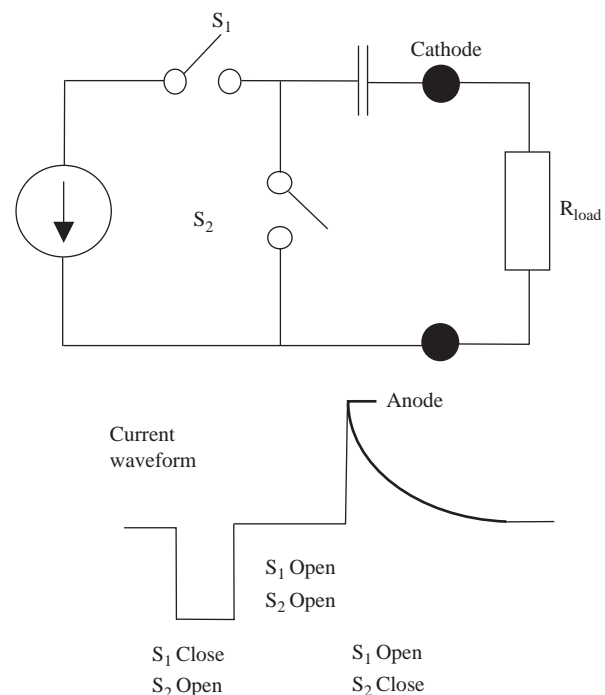


Fig. 4 Method for generating biphasic cathodic-first stimulation waveform. S_1 is closed and a square negative pulse is generated through the load. S_1 is then opened, and when S_2 is closed, the charge stored in the capacitor is discharged and produces the anodic phase.

cathodic monophasic or cathodic-first biphasic waveforms, corrosion can be avoided or minimized. The monophasic anodic waveform must be avoided because it will cause corrosion. Because for most applications, cathodic stimulation has a lower threshold than anodic stimulation (Fig. 2), the monophasic cathodic waveform will be a preferred stimulation waveform as it minimizes both the current to be injected and the corrosion. However, because the current only flows in one direction, the chemical reactions at the interface are not reversed and the electrode is driven into the irreversible region potentially generating tissue damage. A biphasic waveform will avoid both corrosion and tissue damage if the electrode potential is maintained within reversible regions.^[36]

Tissue Damage

Successful restoration of function with electrical stimulation requires that activation of the excitable tissue be carried without causing damage. Tissue damage has been divided in two categories: passive and active. Passive damage refers to surgical trauma and the lack of biocompatibility. Active damage is associated with the effect of current passing through the

electrode, such as physiological and electrochemical processes.^[39]

Surgical access to the implant site can cause significant trauma involving stretching or compression of the tissue. Postsurgical trauma can also trigger compression injury by swelling of the tissue in the presence of electrodes. Moreover, damage to the vascularization of the tissue can induce hypoxic damage. Although mechanical damage can occur, recovery of function and normal histology has been observed.^[40]

The biocompatibility of the material implanted within the tissue also controls the response of the tissue. Both the metal electrode and the insulation material can contribute to the foreign tissue response. This response consists of inflammation and encapsulation of the foreign material by a layer of collagen. This layer forms 3–4 week following the implant and usually remains thin (about 75 μm) unless relative movement of the implant and the tissue is allowed.

A likely cause of tissue damage is the chemical reaction at the interface between the electrode and the tissue. Electrodes operating in the irreversible region can cause significant tissue damage because these processes can modify the pH of the surrounding tissue and generate toxic products. Balanced biphasic waveforms are preferred as the second phase can completely reverse the charge injected into the tissue. Waveforms that have the most unrecoverable charge are the most likely to induce tissue damage. The fact that electrical stimulation can cause damage directly by overdriving the tissue was obtained from experiments reporting a lesser amount of damage when activity in the peripheral nervous system was blocked by lidocaine.^[41] The mechanisms underlying this effect are still unclear, but could include damage to the blood–nerve barrier, ischemia, or a large metabolic demand on the tissue leading to changes in the ionic concentration, such as calcium and potassium concentration both intra- and extracellularly.

STIMULATION OF EXCITABLE TISSUE

Stimulation of Brain Tissue

Electrodes can be placed on the surface of the brain and directly into the brain to activate CNS pathways. Experiments with stimulation of electrode arrays made of platinum and placed on the surface of the brain indicate that damage was produced for a charge density between 50 and 300 $\mu\text{C}/\text{cm}^2$, but that the total charge/phase was also an important factor and should be kept below 3 μC .^[42,43] Intracortical electrodes with a small surface area can tolerate a charge density as high as 1600 $\mu\text{C}/\text{cm}^2$ provided that the charge/phase remains below 0.0032 μC .^[44] More recently, studies

have shown that prolonged stimulation by electrode cortical arrays can significantly decrease the excitability of the neural tissue.^[45]

Stimulation of the Peripheral Nerve

The electrodes used with peripheral nerves are varied and include extraneural designs such as the spiral cuffs or helix electrodes^[46] with electrode contacts directly on the surface of the nerve, or intraneural designs placing electrode contacts directly inside the nerve.^[47,48] Intraneural electrodes can cause significant damage because electrodes are inserted directly into the nerve through the perineurium but have excellent selectivity. Extraneural electrodes are relatively safe because the designs such as the spiral^[49] or helix,^[41] both round designs are self-sizing and allow swelling without compression. However, the selectivity of a round electrode design is not as high as it could be as it is difficult to excite fascicles in the middle of the electrode without first activating axons located in the periphery.^[50] New electrode designs aimed at producing selective stimulation by recruiting only small portions of the nerve have been proposed.^[51,52] An electrode array has been designed to place electrodes directly into the nerve through the epineurium and the perineurium.^[53] Another design takes advantage of the plasticity of the nerve and reshapes the nerve cross section into a flatter configuration. With this flat interface nerve electrode design, electrodes can be placed close to the fascicles without damaging the perineurium.^[40,54] Damage in peripheral nerve stimulation can be caused by the constriction of the nerve as well as neuronal hyperactivity and irreversible reactions at the electrode.^[55]

Stimulation of Muscle Tissue

Muscle tissue can be best excited by electrodes located on the nerve innervating the muscle.^[56] However, for some applications, electrodes can be placed directly on the surface of the skin (surface stimulation),^[57] directly into the muscle (intramuscular electrode),^[58] or on the surface of the muscle (epimysial electrode).^[59] The current thresholds are higher when compared to nerve stimulation unless the electrode is carefully placed near the point of entry of the nerve.^[38] Stainless steel is often used for these electrodes and is safe below 40 $\mu\text{C}/\text{cm}^2$ for coiled wire intramuscular electrodes.^[60]

Stimulation of Cardiac Tissue

Stimulation of cardiac tissue differs in several respects from the stimulation of the nervous system. Nerve

stimulation must be localized to stimulate only a single nerve or a single fascicle. In contrast, stimulation of the heart in any appropriate region will activate the whole heart through the syncytium. Therefore, the electrode design is simpler and the timing or the pulses is of crucial importance to provide appropriate contraction patterns and avoid the set of parameters (timing and amplitude) that can cause fibrillation (vulnerable window). As all cells in the heart are similar in shape and size, the stimulation system does not need to consider issues dealing with fiber diameter recruitment and fatigue. Most electrodes are placed in the endocardium with a reference located on the stimulator. This monopolar stimulation can generate excitation by inducing intracellular voltage (V_m) that depends on the size of the electrode, the conductivity, the current, and the space constant. An approximate formula for the relationship between intracellular voltage V_m (mV) and extracellular current I_a (mA) is given by:^[61]

$$V_m = 34 \times I_a$$

Magnetic Stimulation of Excitable Tissue

Time-varying magnetic fields can induce electric fields in conducting media. Although the conductivity of biological tissue is much lower than that of metals, it is still a good conductor. Moreover, because the conductivity is so much lower than in a metal, the magnetic field can penetrate easily within excitable tissues such as brain or heart tissue. The varying magnetic field B can then induce an electrical field E inside the tissue. The cable equation for magnetic stimulation is given below:^[62]

$$\lambda^2 \frac{\partial^2 V_m}{\partial x^2} - \tau_m \frac{\partial V_m}{\partial t} - V_m = \lambda^2 \frac{\partial E_x}{\partial x}$$

This equation is very similar to that obtained for electrical stimulation and indicates that the nerve is excited by the first spatial derivative of the component of the induced electrical field longitudinal with the axon. The mechanism of action of the magnetic stimulation shares similarities with electrical stimulation.^[63] However, there are several important differences,^[64] which are not reviewed in this article.

CONCLUSIONS

Electrical stimulation of excitable tissue has been highly successful as indicated by the broad application of devices such as the cardiac pacemaker and cochlear

prostheses. Yet it is also clear that the potential for these stimulation devices has not been realized. The sales in the area of neurotechnology^[65] are predicted to be \$7.2 billion in 2004–2008. Cardiac applications are also increasing rapidly as well as combined cardiac and neural applications. The challenge to researchers is to design better electrodes that can 1) be integrated within the tissue and 2) provide a functional interface with the excitable tissue. Recent advances in technology have made it possible to interface directly with the brain.^[66] Yet the mechanical and material properties of the electrodes are still not adequate for long-term implants. In the peripheral nervous system, new designs capable of intimate contact between nerve and electrodes has been shown to be selective for fascicle stimulation^[53,67,68] and for recording.^[69,70] This ability of placing electrodes closer to axons also allows fiber diameter selectivity^[71] for improving the recruitment order of the nerve fibers. Yet cardiac stimulation technology is more mature with significantly larger clinical impact. Neural engineering is by comparison to cardiac engineering still maturing and offers great potential for new discoveries to help patients suffering from a wide range of disorders such as spinal injury, epilepsy, or stroke.

ACKNOWLEDGMENTS

This work was supported by NIH grant # NS32845-10.

ARTICLE OF FURTHER INTEREST

Tissue Engineering of Cardiac Muscle

REFERENCES

1. Grill, W.M.; McDonald, J.W.; Peckham, P.H.; Heetderks, W.; Koesis, J.; Weinrich, M. At the interface: convergence of neural regeneration and neural prostheses for restoration of function. *J. Rehabil. Res. Dev.* **2001**, *38* (6), 633–639.
2. McDonald, J.W.; Sadowsky, C. Spinal-cord injury. *Lancet* **2002**, *359* (9304), 417–425.
3. Hambrecht, F.T. Neural prostheses. *Ann. Rev. Biophys. Bioeng.* **1979**, *8*, 239–267.
4. Grill, W.M.; Kirsch, R.F. Neuroprosthetic applications of electrical stimulation. *Assist. Technol.* **2000**, *12* (1), 6–20.
5. Clark, G.M.; Tong, Y.C.; Patrick, J.F. *Cochlear Prostheses*; Churchill Livingstone, Edinburgh: New York, 1990.
6. Glenn, W.W.; Hogan, J.E.; Loke, J.S.; Ciesielski, T.E.; Phelps, M.L.; Rowedder, R. Ventilatory support by pacing of the conditioned diaphragm in quadriplegia. *N. Eng. J. Med.* **1984**, *310* (18), 1150–1155.

7. Schmit, B.D.; Mortimer, J.T. The effects of epimysial electrode location on phrenic nerve recruitment and the relation between tidal volume and interpulse interval. *IEEE Trans. Rehabil. Eng.* **1999**, *7* (2), 150–158.
8. Brindley, G.S.; Lewin, W.S. The sensations produced by electrical stimulation of the visual cortex. *J. Physiol.* **1968**, *196* (2), 479–493.
9. Margalit, E.; Maia, M.; Weiland, J.D.; Greenberg, R.J.; Fujii, G.Y.; Torres, G.; Pivathaisere, D.V.; O'Hearn, T.M.; Liu, W.; Lazzi, G.; Dagnelie, G.; Scribner, D.A.; de Juan, E., Jr.; Humayun, M.S. Retinal prosthesis for the blind. *Surv. Ophthalmol.* **2002**, *47* (4), 335–356.
10. Mulcahey, M.J.; Smith, B.T.; Betz, R.R.; Triolo, R.J.; Peckham, P.H. Functional neuromuscular stimulation: outcomes in young people with tetraplegia. *J. Am. Paraplegia Soc.* **1994**, *17* (1), 20–35.
11. Dostrovsky, J.O.; Lozano, A.M. Mechanisms of deep brain stimulation. *Mov. Disord.* **2002**, *17* (suppl 3), S63–S68.
12. Rutecki, P. Anatomical, physiological, and theoretical basis for the antiepileptic effect of vagus nerve stimulation. *Epilepsia* **1990**, *31* (suppl 2), S1–S6.
13. George, M.S.; Nahas, Z.; Bohning, D.E.; Kozel, F.A.; Anderson, B.; Chae, J.H.; Lomarev, M.; Denslow, S.; Li, X.; Mu, C. Vagus nerve stimulation therapy: a research update. *Neurology* **2002**, *59* (6 suppl 4), S56–S61.
14. Greatbatch, W.; Seligman, L. Pacemakers. In *Encyclopedia of Medical Devices and Instrumentation*; John Wiley and Sons: New York, 1988.
15. Roth, B.J. Mechanisms for electrical stimulation of excitable tissue. *Crit. Rev. Biomed. Eng.* **1994**, *22* (3–4), 253–305.
16. Plonsey, R. *Bioelectric Phenomena*; McGraw-Hill: New York, 1969.
17. Kraus, J.D. *Electromagnetics*; McGraw-Hill: New York, 1973.
18. P.L., N. *Electric Fields in the brain*; Oxford Press: New York, 1981.
19. Durand, D.M. Electrical stimulation of excitable tissue. In *Handbook of Biomedical Engineering*; Bronzino, J.D., Ed.; CRC Press, 2000; 1, 17.1–17.22.
20. Rall, W. Core conductor theory and cable properties of neurons. In *Handbook of physiology: The Nervous System I*; American Physiological Society, 1979; Chap. 3, 39–96.
21. Ranck, J.B. Which elements are excited in electrical stimulation of mammalian central nervous system: a review. *Brain Res.* **1975**, *98*, 417–440.
22. McNeal, D.R. Analysis of a model for excitation of myelinated nerve. *IEEE Trans. Biomed. Eng.* **1976**, *23* (4), 329–337.
23. Koch. *Methods in Neural Modeling*; MIT Press: Cambridge, MA, 1989.
24. Hines, M. Efficient computation of branched nerve equations. *Int. J. Biomed. Comput.* **1984**, *15* (1), 69–76.
25. Altman, K.W.; Plonsey, R. Development of a model for point source electrical fibre bundle stimulation. *Med. Biol. Eng. Comput.* **1988**, *26* (5), 466–475.
26. Rattay, F. Analysis of models for extracellular fiber stimulation. *IEEE Trans. Biomed. Eng.* **1989**, *36* (7), 676–682.
27. Rattay, F. *Electrical Nerve Stimulation, Theory, Experiments and Applications*; Springer-Verlag: New York, 1990.
28. Warman, E.N.; Grill, W.M.; Durand, D. Modeling the effects of electric fields on nerve fibers: determination of excitation thresholds. *IEEE Trans. Biomed. Eng.* **1992**, *39* (12), 1244–1254.
29. McIntyre, C.C.; Grill, W.M. Excitation of central nervous system neurons by nonuniform electric fields. *Biophys. J.* **1999**, *76* (2), 878–888.
30. Lopicque. Recherches quantitatives sur l'excitation électrique des nerfs traités. *J. Physio. Iaris* **1907**, *9*, 622–635 (from Bartock et al. *J. Physio.* **1983**, *34*, 59–74).
31. Jack, J.J.B.; Noble, D.; Tsien, R.W. *Electrical Current Flow in Excitable Cells*; Clarendon Press: Oxford, 1983.
32. Gorman, P.H.; Mortimer, J.T. The effect of stimulus parameters on the recruitment characteristics of direct nerve stimulation. *IEEE Trans. Biomed. Eng.* **1983**, *30* (7), 407–414.
33. Grill, W.M., Jr.; Mortimer, J.T. The effect of stimulus pulse duration on selectivity of neural stimulation. *IEEE Trans. Biomed. Eng.* **1996**, *43* (2), 161–166.
34. Mortimer, J.T. Electrical excitation of nerve. In *Neural Prostheses Fundamental Studies*; Agnew and McCreery Eds.; 1990; 67–84.
35. Fang, Z.P.; Mortimer, J.T. Selective activation of small motor axons by quasi-trapezoidal current pulses. *IEEE Trans. Biomed. Eng.* **1991**, *38* (2), 168–174.
36. Robblee, L.S.; Rose, T.L. Electrochemical guidelines for selection of protocols and electrode materials for neural stimulation. In *Neural Prostheses Fundamental Studies*; Agnew and McCreery, Eds.; 1990.
37. Shannon, R.V. A model of safe levels for electrical stimulation. *IEEE Trans. Biomed. Eng.* **1992**, *39* (4), 424–426.
38. Mortimer, J.T. Motor prostheses. In *Handbook of Physiology—The Nervous System*; American Physiological Society; 1981; Vol. 3, 155–187.
39. McCreery, D.B.; Agnew, W.F. Mechanisms of stimulation induced neural damage and their relation to guidelines for safe stimulation. In *Neural Prostheses Fundamental Studies*; Agnew and McCreery, Eds.; 1990.
40. Tyler, D.J.; Durand, D.M. Chronic response of the rat sciatic nerve to the flat interface nerve electrode. *Ann. Biomed. Eng.* **2003**, *31* (6), 633–642.
41. Agnew, W.F.; McCreery, D.B. Considerations for safety with chronically implanted nerve electrodes. *Epilepsia* **1990**, *31* (suppl 2), S27–S32.
42. Pudenz, R.H.; Bullara, L.A.; Talalla, A. Electrical stimulation of the brain. I. Electrodes and electrode arrays. *Surg. Neurol.* **1975**, *4* (1), 37–42.
43. Pudenz, R.H.; Bullara, L.A.; Dru, D.; Talalla, A. Electrical stimulation of the brain. II. Effects on the blood-brain barrier. *Surg. Neurol.* **1975**, *4* (2), 265–270.
44. Agnew, W.F.; Yuen, T.G.; McCreery, D.B.; Bullara, L.A. Histopathologic evaluation of prolonged intracortical electrical stimulation. *Exp. Neurol.* **1986**, *92* (1), 162–185.

45. McCreery, D.B.; Agnew, W.F.; Bullara, L.A. The effects of prolonged intracortical microstimulation on the excitability of pyramidal tract neurons in the cat. *Ann. Biomed. Eng.* **2002**, *30* (1), 107–109.
46. Naples, G.G.; Mortimer, J.T.; Yuen, T.G.H. Overview of peripheral nerve electrode design and implantation. In *Neural Prostheses: Fundamental Studies*; Agnew, W.F., McCreery, D.B., Eds.; Prentice Hall: Englewood Cliffs, NJ, 1990; 107–145.
47. Nannini, N.; Horch, K. Muscle recruitment with intrafascicular electrodes. *IEEE Trans. Biomed. Eng.* **1991**, *38* (8), 769–776.
48. Rutten, W.L.; van Wier, H.J.; Put, J.H. Sensitivity and selectivity of intraneural stimulation using a silicon electrode array. *IEEE Trans. Biomed. Eng.* **1991**, *38* (2), 192–198.
49. Naples, G.G.; Mortimer, J.T.; Scheiner, A.; Sweeney, J.D. A spiral nerve cuff electrode for peripheral nerve stimulation. *IEEE Trans. Biomed. Eng.* **1988**, *35* (11), 905–916.
50. Choi, A.Q.; Cavanaugh, J.K.; Durand, D.M. Selectivity of multiple-contact nerve cuff electrodes: a simulation analysis. *IEEE Trans. Biomed. Eng.* **2001**, *48* (2), 165–172.
51. Veraart, C.; Grill, W.M.; Mortimer, J.T. Selective control of muscle activation with a multipolar nerve cuff electrode. *IEEE Trans. Biomed. Eng.* **1993**, *40* (7), 640–653.
52. Tyler, D.J.; Durand, D.M. A slowly penetrating interfascicular nerve electrode for selective activation of peripheral nerves. *IEEE Trans. Rehabil. Eng.* **1997**, *5*, 51–61.
53. Branner, A.; Stein, R.B.; Normann, R.A. Selective stimulation of cat sciatic nerve using an array of varying-length microelectrodes. *J. Neurophysiol.* **2001**, *85* (4), 1585–1594.
54. Tyler, D.J.; Durand, D.M. Functionally selective peripheral nerve stimulation with a flat interface nerve electrode. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2002**, *10* (4), 294–303.
55. McCreery, D.B.; Agnew, W.F.; Yuen, T.G.; Bullara, L.A. Damage in peripheral nerve from continuous electrical stimulation: comparison of two stimulus waveforms. *Med. Biol. Eng. Comput.* **1992**, *30* (1), 109–114.
56. Popovic, D.; Gordon, T.; Rafuse, V.F.; Prochazka, A. Properties of implanted electrodes for functional electrical stimulation. *Ann. Biomed. Eng.* **1991**, *19* (3), 303–316.
57. Myklebust, J.B.; Cusick, J.F.; Sances, A.J.; Larson, S.J. *Neural Stimulation*. RC Press, Boca Raton, 1985.
58. Caldwell, C.W.; Reswick, J.B. A percutaneous wire electrode for chronic research use. *IEEE Trans. Biomed. Eng.* **1975**, *22* (5), 429–432.
59. Grandjean, P.A.; Mortimer, J.T. Recruitment properties of monopolar and bipolar epimysial electrodes. *Ann. Biomed. Eng.* **1986**, *14* (1), 53–66.
60. Mortimer, J.T.; Kaufman, D.; Roessman, U. Intramuscular electrical stimulation: tissue damage. *Ann. Biomed. Eng.* **1980**, *8* (3), 235–244.
61. Malmivuo, J.; Plonsey, R. *Bioelectromagnetism*; 1995.
62. Basser, P.J.; Wijesinghe, R.S.; Roth, B.J. The activating function for magnetic stimulation derived from a three-dimensional volume conductor model. *IEEE Trans. Biomed. Eng.* **1992**, *39* (11), 1207–1210.
63. Roth, B.J.; Basser, P.J. A model of the stimulation of a nerve fiber by electromagnetic induction. *IEEE Trans. Biomed. Eng.* **1990**, *37* (6), 588–597.
64. Nagarajan, S.S.; Durand, D.M.; Hsuing-Hsu, K. Mapping location of excitation during magnetic stimulation: effects of coil position. *Ann. Biomed. Eng.* **1997**, *25* (1), 112–125.
65. Neurotech Business Report. The Market of Neurotechnology. 2004–2008, 2004.
66. Donoghue, J.P. Connecting cortex to machines: recent advances in brain interfaces. *Nat. Neurosci.* **2002**, *5* (suppl), 1085–1088.
67. Leventhal, D.K.; Durand, D.M. Chronic measurement of the stimulation selectivity of the flat interface nerve electrode. *IEEE Trans. Biomed. Eng.* **2004**, *51* (9), 1649–1658.
68. McDonnall, D.; Clark, G.A.; Normann, R.A. Selective motor unit recruitment via intrafascicular multi-electrode stimulation. *Can. J. Physiol. Pharmacol.* **2004**, *82* (8–9), 599–609.
69. Lawrence, S.M.; et al. Acute peripheral nerve recording characteristics of polymer-based longitudinal intrafascicular electrodes. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2004**, *12* (3), 345–348.
70. Yoo, P.B.; Durand, D.M. Selective Fascicular Recording of the Hypoglossal Nerve Using a Multi-Contact Nerve Cuff Electrode. *Annals of Bioemical Engineering* **2003**, *32*, 511–519.
71. Lertmanorat, Z.; Durand, D.M. A novel electrode array for diameter-dependent control of axonal excitability: a simulation study. *IEEE Trans. Biomed. Eng.* **2004**, *51* (7), 1242–1250.

Extracellular Matrix Scaffolds

Stephen F. Badylak

McGowan Institute for Regenerative Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.



INTRODUCTION

The extracellular matrix (ECM) is a complex mixture of structural and functional proteins that are arranged in a unique, three-dimensional ultrastructure that is specific to each tissue and organ. These proteins serve many functions, including the provision of structural support and tensile strength, provide attachment sites for cell surface receptors, and act as a reservoir for cell signaling factors that modulate diverse host processes such as angiogenesis and vasculogenesis, cell migration, cell proliferation and orientation, inflammation, immune responsiveness, and wound healing. Stated differently, the ECM is a vital, dynamic, and indispensable component of all tissues and organs and is nature's natural scaffold for tissue and organ morphogenesis, maintenance, and reconstruction following injury. For these very reasons, the ECM is an excellent scaffold for tissue engineering applications. Unfortunately, it is not yet possible to recreate the ECM by synthetic methods. Therefore, ECM scaffolds must be prepared by processes that utilize a mammalian—usually xenogeneic—tissue source and methods that produce an acellular, sterile device for medical and surgical applications.

OVERVIEW

With the discovery that the ECM was necessary for the conversion of myoblasts to myotubes^[1] and structural proteins such as collagen and glycosaminoglycans played a functional role in salivary gland morphogenesis,^[2] it became apparent that the ECM was more than a passive bystander in the events of tissue and organ development and the pathophysiologic host response to tissue injury. Cytokines, growth factors, and potent functional proteins were found to reside within the ECM. The ECM was quickly recognized as a critically important conduit for exchange of information between cells. The concept of dynamic reciprocity between the ECM and intracellular cytoskeletal and nuclear elements has become widely accepted.^[3–5] Selected forms of the ECM have been successfully used as a scaffold for tissue engineering applications in recent years in both preclinical animal

studies and human clinical applications. The success is likely due to the structure-function relationships of this naturally occurring bioscaffold and its lack of recognition as a foreign material against which an inflammatory or fibrotic reaction must be raised. Both the three-dimensional ultrastructure and the composition of the ECM scaffolds are likely important in the biologic remodeling response.

The ECM is not a static structure. The composition and structure of the ECM are a function of age of the host, the location within specific tissues and organs, and the physiologic demands of the particular tissue.^[6–8] For example, the ECM that constitutes the basement membrane of epithelial tissues and organs, such as the liver, is an excellent substrate for cell growth, especially epithelial cell growth. These basement membrane ECMs tend to be comprised of proteins, including laminin, collagen type IV, and entactin, which are not nearly as abundant in ECMs that derive from submucosal tissue sources. In contrast, the ECM derived from submucosal tissues is often well vascularized, contains primarily type I collagen and site-specific glycosaminoglycans, and a wide variety of growth factors, including basic fibroblast growth factor (bFGF), vascular endothelial cell growth factor (VEGF), and epidermal growth factor (EGF). In other words, even though all ECMs share the common features of providing structural support and serving as a reservoir of growth factors and cytokines, their composition and structure can be markedly different.

The ECM presents its growth factors efficiently to resident cell surface receptors, protects the constituent growth factors from degradation, and modulates the synthesis^[9–12] of these growth factors by the local cell population. In this manner, the ECM affects local concentrations and biologic activity of growth factors and cytokines and makes the ECM an ideal scaffold for tissue repair and reconstruction. In one sense, the ECM can be considered as a localized, controlled release biomaterial that is regulated by host homeostatic control mechanisms. The consequence of the differences between the basement membrane versus submucosal tissue source of the ECM upon in vivo tissue remodeling when these ECMs are used as scaffolds for tissue repair has yet to be determined.

COMPONENTS OF THE EXTRACELLULAR MATRIX THAT SUPPORT TISSUE RECONSTRUCTION

Scaffolds for tissue reconstruction and replacement must have both appropriate structural and functional properties for the intended application. Many synthetic scaffolds are coated with one or more biologically derived components of the ECM, such as fibronectin or collagen type IV, in an attempt to make them more cell friendly. The distinction between structural and functional proteins is becoming increasingly difficult. Domain peptides of proteins originally thought to have purely structural properties have been found to have physiologically significant and potent modulating effects upon cell behavior. For example, the RGD peptide that promotes adhesion of numerous cell types was first identified in the fibronectin molecule,^[13,14] a molecule originally described for its structural properties. Biologically active peptides have since been identified in other dual function proteins including laminin, entactin, fibrinogen, types I and VI collagen, and vitronectin.^[15] Thus, the inclusion of biologically derived components within or upon synthetic scaffold materials affords both structural and functional properties that generally favor a constructive host response (Table 1).

The ECM is a degradable bioscaffold for tissue reconstruction when presented in its natural state. The structural and the functional components of these scaffolds must be thought of as transient due to the rapid rate of degradation of ECM scaffolds in vivo.^[16,17] Once again, the concept of ECM scaffolds as temporary controlled-release vehicles for naturally occurring growth factors is attractive.

The most abundant protein within the ECM is collagen. More than 20 distinct types of collagen have been identified, and the type and relative amounts of each that reside within the ECM are tissue-specific. The primary structural collagen in mammalian tissues is type I collagen. This protein has been well characterized and is ubiquitous across the animal and plant kingdom.^[18] Collagen has maintained a highly conserved amino acid sequence through the course of evolution and, for this reason, allogeneic and xenogeneic sources of type I collagen have been considered to have low antigenic potential. Bovine type I collagen is perhaps the most widely used biologic scaffold for therapeutic applications. Purified forms of collagen are typically chemically cross-linked prior to clinical use; thus, this renders them biologically inert.

Collagen types other than type I exist in naturally occurring ECM, albeit in much lower quantities. These alternative collagen types each provide distinct

Table 1 Major structural components of scaffolds derived from extracellular matrix collagens

Collagen type	Role
I,III	Most abundant collagenous protein that is responsible for the formation of collagen fibrils
IV	Protein forming a basement membrane network structure that permits other basement membrane components to interact; allows for cell attachment
VI	Basement membrane protein responsible for the attachment of the cells as well as anchoring the collagen fibers, nerves, and blood vessels to the surrounding connective tissue
VII	Basement membrane protein responsible for anchoring of fibrils to basement membrane
<i>Glycoproteins</i>	
Fibronectin	Adhesive glycoprotein that interacts with many ECM components; most notable of which is the ability to interact with a cells' fibronectin receptors which results in cell adhesion and migration
Laminin	Adhesive glycoprotein responsible for cell adhesion and migration in the basement membrane; pivotal role in cell surface receptors, both integrins and nonintegrin receptors
Nidogen/Entactin	Pivotal glycoprotein in basement membrane that binds to laminin; paramount for the connection of the networks of laminins and collagen type IV in the basement membrane
Thrombospondins	Adhesive glycoproteins that binds calcium and heparin; modulate cell attachment, migration, and proliferation
Vitronectin	Adhesive glycoprotein that binds to collagens and heparinlike glycosaminoglycans and acts as a cell adhesion molecule; regulates proteolytic degradation of matrix.
<i>Proteoglycans</i>	
Perlecan	Most abundant proteoglycan found in the basement membrane; binds to cell surface and ECM proteins
Decorin	Basement membrane proteoglycan that binds to (decorates) collagen fibrils present in the ECM

mechanical and physical properties to the ECM and contribute to the utility of the intact ECM—as opposed to isolated components of the ECM—as a scaffold for tissue repair. For example, type IV collagen is present within the basement membrane of all vascular structures and is an important ligand for endothelial cells. Thus, collagen type IV is frequently used as a coating for scaffolds intended to have a blood surface interface. Type VII collagen is an important component of the anchoring fibrils of keratinocytes to the underlying basement membrane of the epidermis. Type VI collagen functions as a connector of functional proteins and glycosaminoglycans to larger structural proteins such as type I collagen, which helps to provide a gellike consistency to the ECM. Type III collagen exists within selected submucosal ECMs, such as the submucosal ECM of the urinary bladder, in which less rigid structure is demanded for appropriate function. The diversity of collagens and their distinctive spatial organization is likely responsible for the distinctive biologic activity of ECM scaffolds when placed *in vivo*. This complex composition and structure is also an exemplary representation of the difficulty in recreating such a composite *in vitro*. The ECM is a source of numerous types of collagen, and the relative concentrations and orientation of these collagens to each other provide an ideal environment for cell growth both *in vitro* and *in vivo*.

Laminin is a complex adhesion protein found in the ECM, especially within ECM derived from basement membrane structures.^[19] This trimeric cross-linked polypeptide exists in numerous forms that are dependent upon the particular mixture of peptide chains, (such as $\alpha 1$, $\beta 1$, and $\gamma 1$).^[20,21] Laminin plays a particularly important role in the formation and maintenance of vascular structures and appears to be important in the development of embryonic stem cells. When considering the ECM as a scaffold for tissue repair, laminin is one of the principle reasons that the ECM is an attractive option for vascular reconstruction.^[22,23] Vascularization of scaffolds for tissue repair is one of the rate limiting steps in the field of tissue engineering. Proteins such as laminin and type IV collagen are receiving close attention as an important component of endothelial cell-friendly scaffold materials.

Fibronectin is a major component of ECM and is second only to collagen in quantity within the ECM. As stated earlier, fibronectin has both structural and functional roles *in vivo*. Fibronectin exists both in soluble and tissue isoforms and possesses many desirable properties of a tissue repair scaffold, including ligands for adhesion of many cell types.^[24,25] The fibronectin component of the ECM scaffold derived from the porcine small intestinal submucosa (SIS) and urinary bladder submucosa (UBS) has been shown to be partially responsible for the adhesion of endothelial

cells during *in-vivo* remodeling of these xenogeneic bioscaffolds.^[26,27] The cell-friendly characteristics of this protein have made it an attractive ligand for use as a coating protein upon various synthetic scaffold materials to promote host biocompatibility.

Vitronectin, a multifunctional glycoprotein, binds to components in the ECM such as collagens, glycosaminoglycans, and plasminogen activation inhibitor-1. In addition to binding to ECM components, vitronectin can also control cell attachment, spreading, and migration due to the presence of the Arg-Gly-Asp (RGD) sequence in vitronectin. Not only is vitronectin a key player in ECM scaffold's response to cells, but it can also regulate proteolytic degradation of the matrix through binding to plasminogen activation inhibitor-1.^[28] Vitronectin may play an important role in wound healing^[29] since vitronectin is intimately involved with integrin $\alpha_v\beta_3$, which plays a paramount role in angiogenesis.^[30]

Glycosaminoglycans (GAGs) are important components of ECM and play important roles in binding of growth factors and cytokines, water retention, and the gel properties of the ECM. The heparin-binding properties of numerous cell surface receptors and growth factors such as fibroblast growth factor family and vascular endothelial cell growth factor make the heparin-rich GAGs attractive components of scaffolds for tissue repair. The GAG components of the ECM scaffolds include the naturally occurring mixture of chondroitin sulfates A and B, heparin, heparan sulfate, and hyaluronic acid.^[31] Hyaluronic acid has been extensively investigated as a scaffold for dermal reconstruction and a carrier for numerous cell types and growth factors.

The characteristic of the intact ECM that distinguishes it from other scaffold materials is its diversity of structural proteins and associated bioactive molecules and their unique spatial distribution. Although cytokines and growth factors represent a very small component of the ECM composition, they are potent modulators of cell behavior. Many different growth factors reside within the ECM, including VEGF, EGF, bFGF, keratinocyte growth factor (KGF), transforming growth factor beta (TGF-beta), hepatocyte growth factor (HGF), and platelet-derived growth factor (PDGF). These factors often exist in multiple isoforms, and each has its specific biologic activity. Purified forms of growth factors and biologic peptides have been investigated in recent years as therapeutic means of encouraging blood vessel formation (VEGF), inhibiting blood vessel formation (angiostatin), stimulating deposition of granulation tissue (PDGF), and encouraging epithelialization of wounds (KGF). However, true controlled release of these proteins has been difficult to achieve, and the inability to turn the factor on and off as needed during

the course of tissue repair has limited the success of this approach. An advantage of utilizing the ECM in its native state as a scaffold for tissue repair is the presence of all of the attendant growth factors and their inhibitors in the relative amounts that exist in nature and, perhaps most importantly, in their native three-dimensional ultrastructure. To date, it is not possible to recreate the structure of the native ECM *in vitro*. It is important to note that almost all attempts to chemically cross-link naturally occurring ECMs has resulted in destruction of the constructive scaffold remodeling properties. Cross-linking invariably changes the host response to one of chronic inflammation, frustrated degradation of the scaffold, and a fibrotic reaction.

SOURCES OF EXTRACELLULAR MATRIX AND HOST RESPONSE

ECM exists in all tissues organs. However, sufficient ECM for use as a therapeutic bioscaffold can be harvested from relatively few sources. The dermis of the skin, submucosa of the small intestine and urinary bladder, pericardium, basement membrane and stroma of the decellularized liver, and the decellularized Achilles tendon are all potential sources of ECM. The methods used to process the material will determine its size, shape, practical utility, and the host response (Fig. 1).

Chemical and nonchemical means of cross-linking ECM proteins have been utilized extensively in an

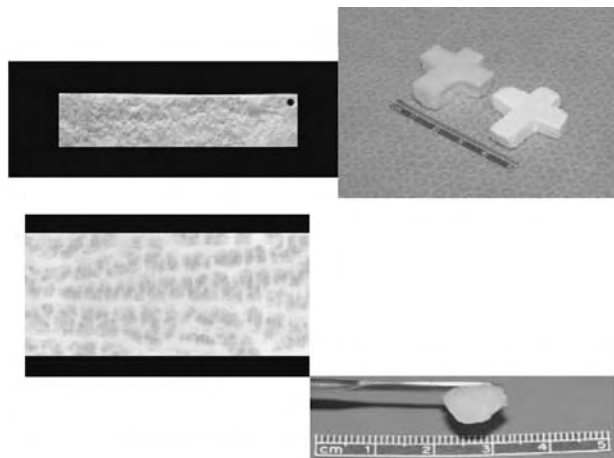


Fig. 1 Extracellular matrix can be engineered to many different forms. The upper left picture shows a multi-laminated (4-ply) sheet. The lower left picture shows a single-layer lyophilized sheet. The picture in the lower right shows a gel form of ECM. The picture in the upper right shows a powdered particulate form that can be shaped into any three-dimensional configuration.

effort to modify the physical, mechanical, or immunogenic properties of naturally derived scaffolds.^[32] Chemical cross-linking methods generally involve aldehydes or carbodiimide. Photochemical means of protein cross-linking have also been investigated. The end result of most chemical cross-linking methods is the transformation of a biologically interactive scaffold material into a relatively inert bioscaffold material. The host response to purified collagen of ECM that has been heavily cross-linked is typically encapsulation of fibrous connective tissue, inhibition of scaffold degradation, and inhibition of cellular infiltration into the scaffold. The physical properties of ECM are often enhanced or strengthened by cross-linking, but mineralization of such cross-linked materials is a common long-term complication. These properties are counter-intuitive to many current approaches in the field of tissue engineering, especially those approaches in which cells are seeded upon scaffolds prior to or at the time of implantation.

The intact, nonmodified ECM or purified components of the ECM, such as collagen, fibronectin, or hyaluronic acid, are currently being investigated as cell-friendly scaffolds or scaffold coatings. These proteins and glycosaminoglycans promote cell infiltration and rapid scaffold degradation, deposition of host-derived neomatrix, and, eventually, constructive tissue remodeling.

EXTRACELLULAR MATRIX SCAFFOLDS FOR TISSUE REPAIR

The ECM plays an important role in embryonic stem cell (ESC) development, tissue differentiation, and organogenesis. The ability of the ECM ligands to modulate cell phenotypes and cell behavior also likely has important implications for other biologic processes such as inflammation, immunity, and wound healing. It is well accepted that developing fetuses and newborns heal wounds differently—generally better—than mature individuals and certainly more quickly and efficiently than aged individuals. It is plausible that the composition and structure of the ECM is at least somewhat responsible for the differences noted in the wound healing process as a function of age. Recently, scaffolds derived from naturally occurring—usually porcine-derived—ECM have been used for therapeutic applications.

There is abundant literature on the use of modified ECM scaffolds, especially chemically cross-linked biologic scaffolds, for tissue repair and replacement. Porcine heart valves, decellularized and cross-linked human dermis (Alloderm™), and chemically cross-linked purified bovine type I collagen (Contigen™) are examples of such products currently available

for use in humans. Similarly modified ECM scaffolds have been used for the reconstitution of the cornea, skin, cartilage and bones plus nerve regeneration.

Porcine-derived ECM scaffolds that have not been modified, except for the decellurization process and terminal sterilization, have been successfully used for the repair of numerous body tissues, including musculotendinous structures,^[33,34] lower urinary tract reconstruction,^[35–40] dura mater replacement,^[41,42] vascular reconstruction,^[43–45] and the repair of full- and partial-thickness skin wounds.^[46] The remodeling process in all of these applications has been remarkably similar. Immediately following implantation in vivo, there is an intense cellular infiltrate consisting of equal numbers of polymorphonuclear leukocytes and mononuclear cells. By 72 hours after implantation, the infiltrate is almost entirely mononuclear cell in appearance with early evidence for neovascularization. These mononuclear cells have not been definitively characterized to date. Between the third and fourteenth day, the number of mononuclear cells increases, and vascularization intensifies. There is also a progressive degradation of the xenogeneic scaffold with an associated deposition of host-derived neomatrix. Following the fourteenth day, the mononuclear cell infiltrate diminishes, and there is the appearance of site-specific parenchymal cells that orient along lines of stress, i.e., a microenvironmental factor.^[47] These parenchymal cells consist of fibroblasts, smooth muscle cells, skeletal muscle cells, and epithelial cells that depend upon the site in which the scaffold has been placed. It has been shown that circulating, marrow-derived progenitor cells participate in this remodeling process when ECM scaffolds are used.^[48] Environmental factors, such as mechanical loading, have also been shown to be important in the remodeling of ECM scaffolds.^[44] Of note, there is an absence of tissue necrosis and scar tissue formation during the remodeling of these xenogeneic ECM scaffolds.

ECM scaffolds derived from the urinary bladder submucosa (UBS) have been used for reconstruction of the lower urinary tract and has shown remodeling and healing that is atypical for adult mammals.^[49–54] The UBS scaffolds have been either allogeneic or xenogeneic in origin and have been used either alone or with cultured autologous cells. Sections of urethra, ureter, and urinary bladder have shown near normal reconstitution with formation of organized and innervated smooth muscle. There is a substantial body of literature developing that supports the use of intact ECM as a scaffold for tissue repair. More than 150,000 human patients have now been implanted with xenogeneic ECM scaffold derived from the porcine small intestinal submucosa for a variety of applications.

SUMMARY

The extracellular matrix is a complex structure that exists in dynamic reciprocity with the resident cell population. The ECM plays a critical role in cell, tissue, and organ development, including the development and differentiation of embryonic stem cells. The cause-effect relationship between the ECM and cell behavior and the ECM and biologic events such as organogenesis, inflammation, and wound healing is only partially understood. However, this naturally occurring scaffold has many attractive features that have at least, to date, been difficult to match with synthetic materials.

ARTICLES OF FURTHER INTEREST

Allografts; Biofunctional Polymers; Biologic and Synthetic Apatites; Biomimetic Materials; Burn Dressing; Collagen; Elastin; Electrospinning; Matrix Metalloproteinases; Microporous Materials; Tissue Engineering; Tissue Engineering, Microscale; Tissue Engineering Scaffolds

REFERENCES

1. Hauschka, S.D.; Konigsberg, I.R. The influence of collagen on the development of muscle colonies. *Proc. Natl. Acad. Sci. U.S.A.* **1966**, *55*, 119–126.
2. Wessells, N.K.; Cohen, J.H. Effects of collagenase on developing epithelia in vitro: Lung, ureteric bud, and pancreas. *Dev. Biol.* **1968**, *18*, 294–309.
3. Bissell, M.J.; Hall, H.G.; Parry, G. How does the extracellular matrix direct gene expression? *J. Theor. Biol.* **1982**, *99*, 31–68.
4. Boudreau, N.; Myers, C.; Bissell, M.J. From lamini to lamin: Regulation of tissue-specific gene expression by the ECM. *Trends Cell Biol.* **1995**, *5*, 1–4.
5. Ingber, D. Extracellular matrix and cell shape: Potential control points for inhibition of angiogenesis. *J. Cell. Biochem.* **1991**, *47*, 236–241.
6. Laurie, G.W.; Horikoshi, S.; Killen, P.D.; Degui-Real, B.; Yamada, Y. In situ hybridization reveals temporal and spatial changes in cellular expression of mRNA for a laminin receptor, laminin, and basement membrane (type IV) collagen in the developing kidney. *J. Cell Biol.* **1989**, *109*, 1351–1362.
7. Martins-Green, M.; Bissell, M.F. Cell-extracellular matrix interactions in development. *Semin. Dev. Biol.* **1995**, *6*, 149–159.
8. Baldwin, H.S. Early embryonic development. *Cardiovasc. Res.* **1996**, *31*, E34–E45.
9. Bonewald, L.F. Regulation and regulatory activities of transforming growth factor beta. *Crit. Rev. Eukaryot. Gene Expr.* **1999**, *9*, 33–44.
10. Kagami, S.; Kondo, S.; Loster, K.; Reutter, W.; Urushihara, M.; Kitamura, A.; Kobayashi, S.; Kuroda,

- Y. Collagen type I modulates the platelet-derived growth factor (PDGF) regulation of the growth and expression of beta 1 integrins by rat mesangial cells. *Biochem. Biophys. Res. Commun.* **1998**, *252*, 728–732.
11. Roberts, R.; Gallagher, J.; Spooncer, E.; Allen, T.D.; Bloomfield, F.; Dexter, T.M. Heparan sulphate bound growth factors: A mechanism for stromal cell mediated haemopoiesis. *Nature* **1988**, *332*, 376–378.
 12. Sjaastad, M.D.; Nelson, W.J. Integrin-mediated calcium signaling and regulation of cell adhesion by intracellular calcium. *BioEssays* **1997**, *19*, 47–55.
 13. Pierschbacher, M.D.; Ruoslahti, E. Cell attachment activity of fibronectin can be duplicated by small synthetic fragments of the molecule. *Nature* **1984**, *309*, 30–33.
 14. Yamada, K.M.; Kennedy, D.W. Dualistic nature of adhesive protein function: Fibronectin and its biologically active peptide fragments can autoinhibit fibronectin function. *J. Biol. Chem.* **1984**, *99*, 29–36.
 15. Humphries, M.J.; Mould, A.P.; Yamada, K.M. Matrix Receptors in Cell Migration. In *Receptors for Extracellular Matrix*; McDonald, J.A., Mecham, R.P., Eds.; Academic Press, 1991; 195–253.
 16. Badylak, S.F.; Kropp, B.T.; McPherson, H.L.; Snyder, P.W. SIS: A rapidly resorbable bioscaffold for augmentation cystoplasty in a dog model. *Tissue Eng.* **1998**, *4*, 379–387.
 17. Rickey, F.A.; Elmore, D.; Hillegonds, D.; Badylak, S.; Record, R.; Simmons-Byrd, A. Regeneration of tissue about an animal-based scaffold: AMS studies of the fate of the scaffold. *Nucl. Instrum. Methods Phys. Res.* **2000**, *172*, 904–909.
 18. Vanderrest, M.; Garrone, R. Collagen family of proteins. *FASEB J.* **1991**, *5*, 2814–2823.
 19. Schwarzbauer, J.E. Basement membranes: Putting up the barriers. *Curr. Biol.* **1999**, *9*, R242–R244.
 20. Timpl, R. Macromolecular organization of basement membranes. *Curr. Opin. Cell Biol.* **1996**, *8*, 618–624.
 21. Timpl, R.; Brown, J. Supramolecular assembly of basement membranes. *BioEssays* **1996**, *18*, 123–132.
 22. Ponce, M.; Nomizu, M.; Delgado, M.C.; Kuratomi, Y.; Hoffman, M.P.; Powell, S.; Yamada, Y.; Kleinman, H.K.; Malinda, K.M. Identification of endothelial cell binding sites on the laminin gamma-1 chain. *Circ. Res.* **1999**, *84*, 688–694.
 23. Werb, Z.; Cu, T.H.; Rinkenberger, J.L.; Coussens, L.M. Matrix-degrading proteases and angiogenesis during development and tumor formation. *Acta Pathol. Microbiol. Immunol. Scand.* **1999**, *107*, 11–18.
 24. Schwarzbauer, J.E. Fibronectin: From gene to protein. *Curr. Opin. Cell Biol.* **1991**, *3*, 786–791.
 25. Miyamoto, S.; Katz, B.Z.; Lafrenie, R.M.; Yamada, K.M. Fibronectin and integrins in cell adhesion, signaling, and morphogenesis. *Ann. N.Y. Acad. Sci.* **1998**, *857*, 119–129.
 26. McPherson, T.B.; Badylak, S.F. Characterization of fibronectin derived from porcine small intestinal submucosa. *Tissue Eng.* **1998**, *4*, 75–83.
 27. Hodde, J.; Record, R.; Tullius; Badylak, S. Fibronectin peptides mediate HMEC adhesion to porcine-derived extracellular matrix. *Biomaterials* **2002**, *23*, 1841–1848.
 28. Gechtman, Z.; Belleli, A.; Lechpammer, S.; Shaltiel, S. The cluster of basic amino acids in vitronectin contributes to its binding of plasminogen activator inhibitor-1: Evidence from thrombin-, elastase-, and plasmin-cleaved vitronectins and anti-peptides and antibodies. *Biochem. J.* **1997**, *325*, 339–349.
 29. Varner, J.A.; Cheresch, D.A. Integrins and cancer. *Curr. Opin. Cell Biol.* **1996**, *8*, 724–730.
 30. Hammes, H.P.; Brownlee, M.; Jonczyk, A.; Sutter, A.; Preissner, K.T. Subcutaneous injection of a cyclic peptide antagonist of vitronectin receptor-type integrins inhibits retinal neovascularization. *Nat. Med.* **1996**, *2*, 529–533.
 31. Hodde, J.P.; Badylak, S.F.; Brightman, A.O.; Voytik-Harbin, S.L. Glycosaminoglycan content of small intestinal submucosa: A bioscaffold for tissue replacement. *Tissue Eng.* **1996**, *2*, 209–217.
 32. Badylak, S.F. Modification of Natural Polymers: Collagen. In *Methods of Tissue Engineering*; Atala, A., Lanza, R.P., Eds.; Academic Press, 2002; 505–514. Chapter 42.
 33. Aiken, S.W.; Badylak, S.F.; Toombs, J.P.; Shelbourne, K.D.; Hiles, M.C.; Lantz, G.C.; Van Sickle, D. Small intestinal submucosa as an intra-articular ligamentous repair material: A pilot study in dogs. *Vet. Comp. Orthop. Traumatol.* **1994**, *7*, 124–128.
 34. Badylak, S.F.; Arnoczky, S.; Plouhar, P.; Haut, R.; Mendenhall, V.; Horvath, C. Naturally-occurring ECMs as scaffolds for musculoskeletal repair. *Clin. Orthop. Relat. Res.* **1999**, *367S*, S333–S343.
 35. Kropp, B.P.; Sawyer, B.D.; Shannon, H.E.; Ripy, M.K.; Badylak, S.F.; Adams, M.C.; Keating, M.A.; Rink, R.C.; Thor, K.B. Characterization of small intestinal submucosa-regenerated canine detrusor: Assessment of reinnervation, in vitro compliance and contractility. *J. Urol.* **1996**, *156*, 599–607.
 36. Kropp, B.P.; Ripy, M.K.; Badylak, S.F.; Adams, M.C.; Keating, M.A.; Rink, R.C.; Thor, K.B. Regenerative urinary bladder augmentation using small intestinal submucosa: Urodynamic and histopathologic assessment in long term canine bladder augmentations. *J. Urol.* **1996**, *155*, 2098–2104.
 37. Atala, A.; Guzman, L.; Retik, A.B. A novel inert collagen matrix for hypospadias repair. *J. Urol.* **1999**, *162*, 1148–1151.
 38. Piechota, H.J.; Dahms, S.E.; Nunes, L.S.; Dahiya, R.; Lue, T.F.; Tanagho, E.A. In vitro functional properties of the rat bladder regenerated by the bladder acellular matrix graft. *J. Urol.* **1998**, *159*, 1717–1724.
 39. Piechota, H.J.; Dahms, S.E.; Probst, M.; Gleason, C.A.; Nunes, L.S.; Dahiya, R.; Lue, T.F.; Tanagho, E.A. Functional rat bladder regeneration through xenotransplantation of the bladder acellular matrix graft. *Br. J. Urol.* **1998**, *81*, 548–559.
 40. Piechota, H.J.; Gleason, C.A.; Dahms, S.E.; Dahiya, R.; Nunes, L.S.; Lue, T.F.; Tanagho, E.A. Bladder acellular matrix graft: In vivo functional properties of the regenerated rat bladder. *Urol. Res.* **1999**, *27*, 206–213.
 41. Cobb, M.A.; Badylak, S.F.; Janas, W.; Boop, F.A. Histology after dural grafting with small intestinal submucosa. *Surg. Neurol.* **1996**, *46*, 389–394.

42. Cobb, M.A.; Badylak, S.F.; Janas, W.; Simmons-Byrd, A.; Boop, F.A. Porcine small intestinal submucosa as a dural substitute. *Surg. Neurol.* **1999**, *51* (1), 99–104.
43. Sandusky, G.E.; Lantz, G.C.; Badylak, S.F. Healing comparison of small intestine submucosa and ePTFE grafts in the canine carotid artery. *J. Surg. Res.* **1995**, *58*, 415–420.
44. Prevel, C.D.; Eppley, B.L.; McCarty, M.; Harruff, R.; Brock, C.; Badylak, S.F. Experimental evaluation of small intestine submucosa as a microvascular graft material. *J. Microsurg.* **1994**, *15*, 588–591.
45. Badylak, S.F.; Lantz, G.; Coffey, A.; Geddes, L.A. Small intestinal submucosa as a large diameter vascular graft in the dog. *J. Surg. Res.* **1989**, *47*, 74–80.
46. Prevel, C.D.; Eppley, B.L.; Summerlin, D.J.; Jackson, J.R.; McCarty, M.; Badylak, S.F. Small intestinal submucosa (SIS): Utilization as a wound dressing in full-thickness rodent wounds. *Ann. Plast. Surg.* **1995**, *35*, 381–388.
47. Hodde, J.P.; Badylak, S.F.; Shelbourne, K.D. The effect of range of motion on remodeling of small intestinal submucosa (SIS) when used as an Achilles tendon repair material in the rabbit. *Tissue Eng.* **1997**, *3*, 27–37.
48. Badylak, S.F.; Park, K.; McCabe, G.; Yoder, M. Marrow-derived cells populate scaffolds composed of xenogeneic extracellular matrix. *Exp. Hematol.* **2001**, *29*, 1310–1318.
49. Chen, F.; Yoo, J.J.; Atala, A. A cellular collagen matrix as a possible “off the shelf” biomaterial for urethral repair. *Urology* **1999**, *54*, 407–410.
50. Dahms, S.E.; Piechota, H.J.; Dahiya, R.; Gleason, C.A.; Hohenfellner, M.; Tanagho, E.A. Bladder acellular matrix graft in rats: Its neurophysiologic properties and mRNA expression of growth factors TGF-alpha and TGF-beta. *Neurourol. Urodyn.* **1998**, *17*, 37–54.
51. Dahms, S.E.; Piechota, H.J.; Dahiya, R.; Lue; Tanagho, E.A. Composition and biomechanical properties of the bladder acellular matrix graft: Comparative analysis in rat, pig and human. *Br. J. Urol.* **1998**, *82*, 411–419.
52. Dahms, S.E.; Piechota, H.J.; Nunes, L.; Dahiya, R.; Lue, T.F.; Tanagho, E.A. Free ureteral replacement in rats: Regeneration of ureteral wall components in the acellular matrix graft. *Urology* **1997**, *50*, 818–825.
53. Probst, M.; Piechota, H.J.; Dahiya, R.; Tanagho, E.A. Homologous bladder augmentation in dog with the bladder acellular matrix graft. *Br. J. Urol.* **2000**, *85*, 362–371.
54. Yoo, J.J.; Meng, J.; Oberpenning, F.; Atala, A. Bladder augmentation using allogenic bladder submucosa seeded with cells. *Urology* **1998**, *51*, 221–225.

Eye Tracking: Characteristics and Methods

Daniel C. Richardson

Department of Psychology, Stanford University, Stanford, California, U.S.A.

Michael J. Spivey

Department of Psychology, Cornell University, Ithaca, New York, U.S.A.

INTRODUCTION

Eye movements are fundamental to the operation of the visual system. The term eye movement research relates to a patchwork of fields more diverse than the study of perceptual systems, however. Due to their close relation to attentional mechanisms, eye movements can provide insight into cognitive processes such as language comprehension, memory, mental imagery, and decision making. Eye movement research is of great interest in the study of neuroscience and psychiatry, as well as ergonomics, advertising, and design. Since eye movements can be controlled volitionally, to some degree, and tracked by modern technology with great speed and precision, they can now be used as a powerful input device, and they have many practical applications in human-computer interactions.

CHARACTERISTICS OF EYE MOVEMENTS

Eye movements are arguably the most frequent of all human movements.^[1] Large ballistic scanning movements called saccades typically occur 3–4 times every second. As one early researcher put it, “there seems to be an almost ceaseless twitching, as if rest for more than an instant were the one thing not to be endured.”^[2] Indeed, virtually all animals with developed visual systems actively control their gaze using eye or head movements.^[3] This frenetic movement is a consequence of the enormous amount of visual information that is available to an organism. Rather than devote resources to processing it all in detail, evolution appears to have selected a solution whereby small portions of the visual world are inspected in a rapid sequence.^[4] Consequently, the human eye monitors a visual field of about 200 degrees, but receives detailed information from only 2 degrees.^[5] This region, about the size of a thumbnail at arm’s length, is called the fovea. It is jerked around at speeds of up to 500° a second, during which its sensitivity drops to near blindness levels.^[6,7] During the 200–300 milliseconds it is at rest, however, over 30,000 densely packed photoreceptors in the fovea provide high-acuity color vision.

HISTORY AND METHODS

The first data regarding eye movements were obtained either through introspection or by the experimenter observing a subject’s eye, using a mirror, telescope, or peephole. These methods were dubious, of course, because any feature of the eye being studied could be obscured by the eye doing the studying. The first significant advance, therefore, was the invention of mechanical devices that would translate the eye’s movements into permanent, objective records of its motion.^[8]

At the end of the nineteenth century, eye movement research served a pressing theoretical need. As Delabarre wrote, “[m]any problems suggest themselves to the psychologist whose solution would be greatly furthered by an accurate method of recording the movements of the eye.”^[9] Phenomena such as visual illusions and aesthetic preferences were commonly explained away in terms of eye movements, and yet there were few data beyond introspection to support these hypotheses. In some of the first empirical studies, Javal^[10] used mirrors to observe the eye movements of subjects while reading, and was the first to note that the eyes moved in a series of “jerks.” These fixations were counted by placing a microphone on a closed eyelid while the subject read monocularly. Each time the bulge of the moving cornea bumped the microphone, a saccade could be counted, according to Lamare and Javal.^[11] An approximate measure of the location of these fixations could be obtained by inducing an afterimage in the subject’s eye with a bright light, and then asking subjects to report the location of the afterimage as they read.^[12]

Such observational methods, however, were limited by the accuracy and memory of the person making the observations.^[13] An objective record of the motion of the eyes was required. Since “plaster of Paris will attach itself firmly and immovably to any moist surface,” Delabarre^[9] was able to fix a small, moulded cap to a sufficiently “cocainized eye.” A wire ran from the cap to a lever, which drew the horizontal movements of the eye on the smoked surface of a kymograph cylinder. With his lids propped open, the subject (usually Delabarre himself) could read text through a hole drilled in the cap. It was reported that the cap “will not detach itself

until it becomes thoroughly soaked with tears.” Delabarre writes, “As to whether there is any danger to the eye to be feared from using it in this manner, I cannot say with assurance.” He reports that he was able to record his own eye movements for up to an hour, and suffered no ill effects after a week’s recuperation.

Huey^[14] simultaneously developed a similar mechanical method to study the behavior of subjects reading an article in the magazine *Cosmopolitan*. By varying the distance between the subject and the text, he found that “the number of jerks was shown to be a function of the matter read rather than of the arc described by the eye’s rotation.” Although the temporal resolution of his apparatus was too poor to measure the speed of the eye movements, he conjectured that they may be so fast that “we really do not see foveally what we read except at the few points on the ordinary line where the eye pauses.”

Huey and Delabarre were able to gain valuable first insights into the function and characteristics of eye movements, yet their mechanical devices were criticized for impeding motion and straining the eye. To overcome these flaws, Dodge and Cline^[15] invented a device to produce “a group of what we may justly claim to be the first accurate measurements of the angle of velocity of the eye movements under normal conditions.” Dodge’s method, as it became known, used photography to record the movements of the eye accurately and noninvasively, and the same basic technique continued to be used into the 1970s.^[8]

If the eye were a perfect sphere and rotated about its center, a light ray would be reflected at a constant angle despite rotations. Given that the eye has neither of these characteristics, however, the reflection of a ray of light bouncing off the bulge of the cornea will move as the eyes move. In Dodge’s first device, a vertical line of light was bounced off the cornea and fell on a horizontal slit. Behind this slit was a photographic plate that moved vertically, regulated by the escape of air from a cylinder. When developed, this plate showed time on the y -axis and horizontal motion of the eye on the x -axis.

Further developments in the 1920s, in labs at Chicago and Stanford, allowed two photographic recordings to be made simultaneously. In this way, head position could be recorded by reflecting a light off a bead on a pair of spectacles,^[16] or the horizontal eye movements of one eye could be recorded with the vertical movements of another, producing the first two-dimensional eye movement records.^[17] Later technology allowed the reflection beam from a single eye to be split, its vertical and horizontal components measured and recombined in the form of a fixation dot recorded on a film reel. This methodology allowed researchers, including Buswell,^[18] to produce some of the first two-dimensional scan paths of subjects inspecting images. Although a thread of research continued for a couple of

decades investigating the relationship between mental imagery and eye movements,^[19–23] the vast bulk of eye movement research in the first half of the century investigated the processes, habits, and individual and cultural differences involved in reading.^[11,24–27]

The 1960s saw a renaissance of the turn-of-the-century, invasive techniques of Delabarre^[9] and Huey^[14] for recording eye movements. Researchers found that rather than using sticky plaster of Paris, a device could be tightly clamped to the eye using suction. Yarbus^[28] used a tiny valve to withdraw fluid from under a contact lens, and Fender^[29] found that sodium bicarbonate would osmose through the tissue of the eye and create negative pressure. A tiny plane mirror could be attached to the surface of the contact lens, and its reflection could be recorded in much the same way as a corneal reflection. To avoid a tear film clouding the lens, researchers also mounted the mirrors on stalks that protruded out from the eye.^[29] As well as reflecting with mirrors, these stalks could produce their own light source if fitted with small lamps^[30] or glowing radioactive tritium.^[31] Finally, a nonoptical method employed a scleral search coil: a contact lens embedded with two orthogonal wire coils that would perturb a magnetic field surrounding the subject’s head.^[32] Due to the discomfort produced by these contact lens methods, all have dropped out of use except the last, and that is primarily used in animal research.

In the early 1970s, a host of techniques were developed in which the eye was scanned with a television camera, and certain distinct features were electronically detected and localized. These methods are most sensitive to high contrast, and so one technique scanned the image of the eye for the limbus, the boundary between the white sclera and the coloured iris. If small electronic photoreceptors are aligned near the limbus, their output will vary according to the amount of white sclera exposed. This method will give a very rapid measure of horizontal eye movement.^[33] Unfortunately, the iris is large and often obscured by the eyelid, and vertical eye movements in particular are difficult to track with this method. An alternative is to scan for the lack of reflectance from the pupil (dark-pupil tracking), although there can be low contrast between this black circle and dark-brown irises. If the pupil is lit directly from the front, the light will bounce off the back of the retina and appear very bright (bright-pupil tracking), as it does in poorly taken flash photographs. This bright circle can then be more easily detected by a scanning technique.^[34]

All the methods described so far for recording eye movements are more precisely stated as methods for recording movements of the eye in relation to the head. In order to infer where the subject was looking in the world, researchers needed to ensure that the head was absolutely stationary, by employing severe methods of

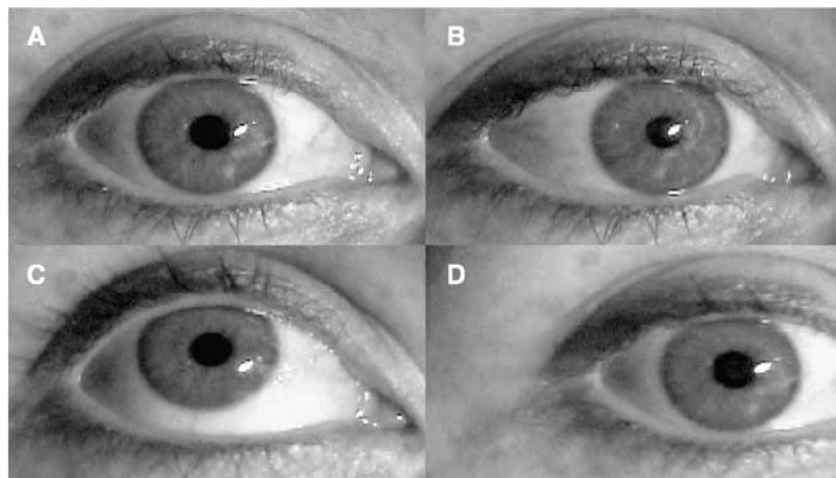


Fig. 1 The corneal reflections produced by different eye-head positions. The corneal reflection appears as a bright white dot, just to the side of the pupil (A). The relative positions of the pupil and the corneal reflection change when the eye rotates around its vertical (B) and horizontal (C) axes. This relationship does not change, however, when the head moves and the eye is stable (compare A and D).

restraint. These draconian means were obviated by the innovation, in the 1970s, of simultaneously measuring two optical characteristics of the moving eye. Since these features behaved differently under head movement and eye rotation, their differential could be used to calculate the point of regard, the place in the world where the subjects was actually looking (Fig. 1). Although such devices still needed to restrain the head with a bitebar or chinrest, they allowed slight movements of the head to be deconfounded from eye movements, and so produced more accurate gaze tracking.

As described above, Merchant and Morrisette^[34] employed a scanning method to detect the center of a brightly lit pupil. The same technique was also used to find the smaller, brighter corneal reflection. Since the position of the corneal reflection in relation to the center of the pupil remains constant during head translation, but moves with eye rotation, point of regard can be extrapolated. Like contemporary systems,^[35] the Honeywell oculometer was able to calculate what location on a screen was being fixated, while being so noninvasive that the subject was often not aware of its presence in the room.

A similar logic lay behind the development of the dual Purkinje image eye tracker.^[36] Light bouncing off the eye produces a series of reflections. The first, and the brightest, is the corneal reflection. A second image is reflected off the rear surface of the cornea, and another two by the front and rear of the lens. These four Purkinje images all have different motions in relation to eye rotation. The dual Purkinje image eye tracker measures the disparity between the first and the fourth images by adjusting a series of mirrors with servomotors until the two images are superimposed upon electronic photoreceptors. The degree to which the mirrors have to be moved is directly related to eye rotation and is independent of head translation, although the head still must be held in place with a chinrest or bitebar so that the eye can be detected by the equipment. The

advantage of the Purkinje eye tracker is that since it is limited only by the speed of the servomotors, it is remarkably fast and accurate. The continuous analog signal was sampled at a rate of 300 Hz in initial incarnations, and modern computers can sample at up to 1000 Hz.

The balance between obtaining a high-precision record of an observer's point-of-regard and allowing natural head and body movements is where much of the technological advancement in eye-tracking takes place in the current state of the art. Whereas systems like the dual Purkinje eye trackers require that a subject's head be immobilized by a chinrest or bitebar, new headband-mounted eyetrackers point an additional scene camera at the subject's field of view from the subject's perspective, thus allowing point-of-regard to be superimposed on the scene camera's image regardless of where or how the subject moves^[37-39] (Fig. 2). Recent table-mounted remote optical eye trackers have also been developed that allow some natural head movement while the



Fig. 2 A subject wearing a light, head-mounted eye tracker. The eye monitor, with crosshairs on the pupil and corneal reflection visible, is shown at the top right. The view from the scene camera, with the subject's point of gaze superimposed, can be seen at the bottom right.

subject sits in front of a computer screen for two-dimensional stimulus presentation.

One particularly important, and frequently used, method in eye tracking concerns not just the eye tracker itself but also the yoking of the eye-position signal with real-time stimulus presentation: gaze-contingent display paradigms. The visual system receives highly detailed information from the small part of the visual field that the eye is fixating at that moment, and yet the target for each detailed fixation is planned on the basis of information gathered from low-resolution peripheral vision. This relationship between peripheral information and saccades has intrigued researchers. Its influence could only be assessed, however, if subjects were instructed to hold their eyes still, or if stimuli were presented tachistoscopically before an eye movement could occur. The next advance in eye movement research came when a cluster of researchers in the 1970s were able to overcome these limitations by coupling the system that was measuring eye movements with the system that was presenting stimuli to the subjects.^[40–42] During the 30–50 milliseconds that the eyes take to saccade to a new location, the visual system's sensitivity is greatly reduced. To trigger a change in the visual stimulus during that brief period, the saccade must be detected within a few milliseconds of onset, the appropriate stimulus change calculated by the computer, and the screen display refreshed. Using a limbus reflection system that sampled eye position every millisecond, McConkie and colleagues^[41,43] were able to produce saccade-contingent display control. In such experiments, the characters of a line of text could be changed while a subject was reading, and hence the amount of information that a subject detected peripherally could be assessed.

The advent of gaze-contingent paradigms was a highly significant advance in eye movement research. It has had a far greater impact, perhaps, in its practical applications. The technology provides not merely a device a scientist can use to study a subject's eye movement behavior, but also a way for the subject's eye movement behavior to manipulate a device.

CONCLUSION

Dodge and Cline's^[15] revolutionary invention allowed researchers, for the very first time, to make a permanent, objective record of a subject's eye movements using a noninvasive method. Although eye tracking techniques were refined in the following decades, there were no comparable breakthroughs until the advent of digital technology and image processing in the 1970s. Such advances have placed fewer constraints on experimental subjects, with the result that modern researchers are able to record the eye movements of freely moving

subjects carrying out everyday tasks. We speculate that further developments will produce smaller, cheaper eye trackers that may well populate a range of human-computer interfaces. The potential ubiquity of such devices underscores the need for a solid understanding of the cognitive and perceptual processes that govern human eye movements.

ARTICLES OF FURTHER INTEREST

Ergonomics; Eye Tracking: Research Areas and Applications; Medical Robotics

REFERENCES

1. Bridgeman, B. Conscious vs unconscious processes: The case of vision. *Theory Psychol.* **1992**, 2 (1), 73.
2. Stratton, G.M. Symmetry, linear illusion, and the movements of the eye. *Psych. Rev.* **1906**, 13 (2), 82.
3. Land, M.F. *Eye Movement Research: Mechanisms, Processes and Applications. Studies in Visual Information Processing*, 6; Findlay, J.M., Walker, R., Eds.; Elsevier Science: New York, NY, 1995; 63–76.
4. Treue, S. Neural correlates of attention in primate visual cortex. *Trends Neurosci.* **2001**, 24, 295.
5. Levi, D.M.; Klein, S.A.; Aitsebaomo, A.P. Vernier acuity, crowding and cortical magnification. *Vis. Res.* **1985**, 25 (7), 963.
6. Matin, E. Saccadic suppression: A review and an analysis. *Psychol. Bull.* **1974**, 81 (12), 899.
7. Thiele, A.; Henning, P.; Kubischik, M.; Hoffmann, K.P. Neural mechanisms of saccadic suppression. *Science* **2002**, 295 (5564), 2460.
8. Taylor, S.E. *EDL Research and Information Bulletin*; McGraw-Hill: New York, 1971; 9.
9. Delabarre, E.B. A method of recording eye-movements. *Am. J. Psychol.* **1898**, 9 (4), 572.
10. Javal, E. Essai sur la Physiologie de la Lecture. *Ann. Ocul.* **1879**, 81, 61.
11. Tinker, M.A. A photographic study of eye movements in reading formulae. *Genet. Psychol. Monogr.* **1928**.
12. Erdmann, B.; Dodge, R. *Psychologische Untersuchungen über das Lesen, auf Experimenteller Grundlage*; Halle, 1898.
13. Dodge, R. Recent studies in the correlation of eye movement and visual perception. *Psychol. Bull.* **1906**, 3 (3), 85.
14. Huey, E.B. Preliminary Experiments in the Physiology and Psychology of Reading. *Am. J. Psychol.* **1898**, 9 (4), 575.
15. Dodge, R.; Cline, T.S. The angle velocity of eye movements. *Psychol. Res.* **1901**, 8 (2), 145.
16. Miles, W.R.; Shen, E. Photographic recording of eye movements in the reading of Chinese in vertical and horizontal axes; method and preliminary results. *J. Exp. Psychol.* **1925**, 8, 344.

17. Gilliland, A.R. Photographic methods for studying reading. *Visual Educ.* **1921**, 2, 21.
18. Buswell, G.T. *How People Look at Pictures: A Study of the Psychology and Perception in Art*; Univ. Chicago Press: Oxford, England, 1935; 198.
19. Totten, E. Eye movement during visual imagery. *Comp. Psychol. Monogr.* **1935**.
20. Perky, C.W. An experimental study of imagination. *Am. J. Psychol.* **1910**, 21, 422.
21. Stoy, E.G. A preliminary study of ocular attitudes in thinking of spatial relations. *J. of Gen. Psych.* **1930**.
22. Clark, H. Visual imagery and attention: an analytical study. *Am. J. Psychol.* **1916**, 27 (4), 461.
23. Goldthwait, C. Relation of eye movements to visual imagery. *Am. J. Psychol.* **1933**, 45.
24. Tinker, M.A. The study of eye movements in reading. *Psychol. Bull.* **1946**, 43.
25. Walker, R.Y. Eye-movements of good readers. *Proc. Iowa Acad. of Sci.* **1933**, 40.
26. Stone, L.G. Reading reactions for varied types of subject matter: an analytical study of the eye-movements of college freshmen. *J. Exp. Educ.* **1941**, 10.
27. Jacobson, J.Z.; Dodwell, P.C. Saccadic eye movements during reading. *Brain Lang.* **1979**, 8 (3), 303.
28. Yarbus, A.L. *Role of Eye Movements in the Visual Process*; Nauka: Oxford, England, 1965; 167.
29. Fender, D.H. Contact lens stability. *Biomed. Sci. Instrum.* **1964**, 2, 43.
30. Byford, G.H. A sensitive contact lens photoelectric eye movement recorder. *IRE Trans. Bio-Med. Electron.* **1962**, 9, 236.
31. Nayrac, P.; Milbled, G.; Parquet, P.H.J.; Leclercq, M.; Dhedin, G. Un nouveau procédé d'enregistrement des mouvements oculaires. Application aux tests de tracking. *Lille Med.* **1969**, 14 (685-687).
32. Robinson, D. A. A method of measuring eye movement using a scleral search coil in a magnetic field. *IEEE Trans. Bio-Med. Electron.* **1963**, 10, 137.
33. Young, L.R. *Biomedical Engineering Systems*; Clynes, M., Milsom, J.H., Eds.; McGraw-Hill: New York, 1970.
34. Merchant, J.; Morrissette, R.; Porterfield, J.L. Remote measurement of eye direction allowing subject motion over one cubic foot of space. *IEEE Trans. Bio-Med. Electron.* **1974**, 21 (4), 309.
35. Lambert, R.H.; Monty, R.A.; Hall, R.J. High-speed data processing and unobtrusive monitoring of eye movements. *Behav. Res. Meth. Instrum.* **1974**, 6 (6), 525.
36. Cornsweet, T.N.; Crane, H.D. Accurate two-dimensional eye tracker using first and fourth Purkinje images. *J. Opt. Soc. Am.* **1973**, 63.
37. Ballard, D.H.; Hayhoe, M.M.; Pelz, J.B. Memory representations in natural tasks. *J. Cogn. Neurosci.* **1995**, 7 (1), 66.
38. Land, M.F.; Lee, D.N. Where do we look when we steer. *Nature* **1994**, 369 (6483), 742.
39. Tanenhaus, M.K.; Spivey Knowlton, M.J.; Eberhard, K.M.; Sedivy, J.C. Integration of visual and linguistic information in spoken language comprehension. *Science* **1995**, 268 (5217), 1632.
40. Reder, S.M. On-line monitoring of eye-position signals in contingent and noncontingent paradigms. *Behav. Res. Meth. Instrum.* **1973**, 5.
41. McConkie, G.W.; Rayner, K. The span of the effective stimulus during a fixation in reading. *Percept. Psychophys.* **1975**, 17 (6), 578.
42. Rayner, K. Parafoveal identification during a fixation in reading. *Acta Psychologica.* **1975**, 39 (4), 271.
43. McConkie, G.W.; Zola, D.; Wolverton, G.S.; Burns, D.D. Eye movement contingent display control in studying reading. *Behav. Res. Meth. Instrum.* **1978**, 10 (2), 154.

Eye Tracking: Research Areas and Applications

Daniel C. Richardson

Department of Psychology, Stanford University, Stanford, California, U.S.A.

Michael J. Spivey

Department of Psychology, Cornell University, Ithaca, New York, U.S.A.



INTRODUCTION

In this brief review of the research fields and practical applications of eye movement technology, we will begin by considering studies of the relationship between eye movements and attentional mechanisms. Typically, in such experimental paradigms, a subject is required to fixate one location and direct their attention to another while psychophysical and neuropsychological measures are taken. We will then examine how the relationship plays out as observers view more complex, natural scenes. The relationship between visual perception and language has historically interested researchers, and we will review studies of eye movements during spoken language comprehension and production, as well as during the more specialized case of reading. In addition to perceptual demands, language processing is a cognitive task, and we shall see how eye movements are also implicated in cognitive processes such as memory, imagination, reasoning, and decision-making. Perhaps as a result of this tight integration with a spectrum of perceptual and cognitive processes, eye movements are richly intertwined with behavior in a variety of task domains, from social interaction to jet plane navigation. Moreover, differences between the eye movements of individuals can reveal differences in aptitude, expertise, and even pathology.

SPATIAL ATTENTION AND SCENE PERCEPTION

Attending to a location in space aids processing of a stimuli at that location,^[1] and has neural correlates, such as modulating the activity and sensitivity of neurons across the visual cortex.^[2] The link between attending to a location and launching a ballistic eye movement (a saccade) to it is intuitive, yet it is equally clear that while fixing the eyes on one location, spatial attention can be directed to other locations. This spotlight of covert attention^[1] was thought to be a mechanism behaviorally and neurally dissociable from

eye movement control. However, it is highly likely that spatial attention and saccade planning are closely coupled during natural, unconstrained eye movement.^[3] There is behavioral evidence that covert attention directed in one direction can lead to deviations in orthogonal saccades,^[4] and there is neuropsychological evidence from single-cell recordings suggesting that attention and eye movement systems utilize overlapping neural networks.^[5] Moreover, planning a saccade toward a location improves processing at that location, regardless of whether or not the saccade is launched.^[6–8] Indeed, evidence shows that microstimulation of neurons in the frontal eye fields can cause both a saccade to a certain location^[9] and, with a lower level of stimulation, an absence of eye movement—but improved stimulus detection at that location.^[10,11]

How are these coupled mechanisms of attention and eye movement deployed when a scene is viewed? Early eye movement recordings showed that a subject's fixations would center on interesting or informative areas of the image, leaving blank or uniform regions uninspected.^[12] These findings have been recently replicated on a large scale. Wooding and colleagues installed an autonomous eyetracker in a public museum in London, and collected data from over 5000 subjects looking at works from the National Gallery.^[13,14] They too found that only a small set of regions in a work of art were reliably fixated by viewers. Contemporary research aims to quantify what determines such regions of interest, and has identified two main components.^[15] First, statistical properties of the image such as high spatial frequency^[16,17] and local contrast^[18] have been found to be closely correlated with fixation likelihood. These “bottom-up” properties can be analyzed to produce a salience map of any given image, which may correspond to a neural representation in the visual cortex.^[19,20]

A second influence on scene perception comes “top-down,” from knowledge, memories, beliefs, or goals that the viewer may bring to the image. In an early demonstration of this effect, Yarbus^[21] presented various subjects with Ilya Repin's painting *An Unexpected Visitor*, and asked them different questions,

such as “What are the ages of the people in the painting?” “What had the family been doing before the visitor arrived?” and “How long had the visitor been away?” The scanpaths differed dramatically according to the question. The semantic content of the scene itself can also attract attention. If there are out-of-place objects in a scene, such as a farmyard animal in a hospital, these anomalies are frequently fixated.^[22]

In order to scan a scene efficiently, some form of short-term memory for saccade location must be retained.^[23,24] It is not clear exactly what kind of memory representation is constructed from these successive fixations.^[25–28] Although observers are surprisingly poor at detecting large changes across scenes under some circumstances,^[29–33] there is evidence that eye fixation patterns are related to this form of scene recognition.^[34–36]

Eye movements have contributed to our understanding of the functioning of the visual system and attentional mechanisms. They have also provided insight into abnormal brain functioning. In the early years of eye movements research, Diefendorf and Dodge^[37] observed a deficit in smooth-pursuit eye movements in patients with praecox dementia. The relationship between this disorder, now known as schizophrenia, and eye movements was rediscovered in the 1970s.^[38,39] During a saccade the eye moves in a rapid ballistic motion, but the human eye can also rotate slowly and smoothly when tracking a moving object. Schizophrenic patients, however, find it difficult to maintain these smooth-pursuit movements, and they launch more predictive saccade eye movements than control subjects. On tasks that demand such predictive eye movements, however, schizophrenic subjects perform faster than controls.^[40,41] In an antisaccade task, subjects must inhibit an eye movement toward a suddenly appearing stimulus and make a saccade in the opposite direction, something that schizophrenics find relatively difficult.^[42–44] Schizophrenics also show different exploratory eye movements and scanpaths when viewing faces and stimuli with emotional valence.^[45,46] These eyemovement differences are of great interest to researchers attempting to relate schizophrenia to dysfunction in specific brain mechanisms,^[47] and to clinicians diagnosing and treating the disorder. For example, abnormalities in eye movements can be seen by the age of six in children susceptible to schizophrenia, more than ten years before the highest risk for onset of psychosis.^[48]

READING

The general characteristics of eye movements during perception have been studied in great depth during the process of reading.^[49,50] In general, eye fixations during

reading last approximately 200–250 milliseconds.^[51] Reading saccades span on average about two degrees of visual angle, although this can be better expressed here in terms of a span of seven to nine letter spaces, since the number of letters covered remains largely invariant despite differences in text size or distance.^[52] The chances of an individual word being fixated vary according to whether it is a content word (85%) or a function word (35%),^[53] and in relationship to the length of the word, with two- or three-letter words being skipped 75% of the time, but eight-letter words fixated almost always.^[54] Eye movements also vary as a function of the legibility of the text,^[55–57] syntactic difficulty of the text,^[58–60] conceptual difficulty of the text,^[61–63] and whether it is being read silently or out loud.^[64]

Saccades typically land between the beginning and the middle of a word.^[65] A reader will gather information from text at and around this fixation point. The size of this perceptual span has been investigated using a gaze-contingent paradigm.^[66] The text appears normal at the reader’s fixation point and within a window of a certain character length, but beyond that the text is garbled. With a small window, reading is impaired. Reading is unaffected, however, when the window extends 3–4 letters to the left of fixation and 14–15 letters to the right, suggesting that there is a perceptual span of around 18 characters centered asymmetrically around the fixation point.^[67,68] This asymmetry in the perceptual span is reversed for languages that are printed left to right,^[69] and rotated for those that run vertically.^[70] Although readers typically move their eyes forward when reading, approximately 10–15% of saccades move backward, fixating previous letters or words. These regressive saccades are thought to be related to difficulties in processing an individual word, or difficulties in processing the meaning or structure of a sentence. In these cases, readers can accurately refixate the part of the text that generated confusion.^[71,72]

These features of eye movements during reading—gaze durations, saccade lengths, and occurrence of regressions—can be used to infer moment-by-moment cognitive processing of a text by the reader.^[73] Details of the cognitive processes of pronoun resolution and co-reference, word frequency, lexical ambiguity, syntactic ambiguity, and discourse factors can all be gleaned from analyses of eye movement patterns. Rayner^[49] has reviewed the research in this area. For this reason, eye movements can reveal significant differences between individuals, as well as being an important data source about language processing. Differences in eye movement patterns can be distinguished between fourth- and fifth-graders, successful and unsuccessful third-graders or undergraduates, dyslexics and nondyslexics, and even graduate students

and professors.^[74–78] As a consequence, the study of eye movements while reading has considerable practical applications in education psychology.^[79]

LANGUAGE IN A VISUAL AND SOCIAL CONTEXT

Language use, more often than not, occurs within rich visual contexts, and the interplay between linguistic processes and visual perception is of increasing interest to psycholinguists and vision researchers.^[80] Eye movement research provided some of the first compelling demonstrations of the real-time interdependence between spoken language processing and visual perception.^[81] In these paradigms, a subject wears a headband-mounted eyetracker and typically follows spoken instructions to move objects around on a table or on a computer screen. One of the important findings with this paradigm demonstrated how spoken word recognition is an incremental process, at the timescale of milliseconds, that is profoundly influenced by the visual context. For example, subjects facing a display containing a bag of candy, a candle, an envelope, and a spoon, and being instructed to pick up the candy, occasionally look first at the candle (because of the overlap in the first few phonemes between “candy” and “candle”) and then at the bag of candy to pick it up.^[82] When the object with the phonologically competing name—candle—is not present in the display, subjects look more accurately and more quickly to the bag of candy. (The same results are found with a wide variety of objects and object names, e.g., penny and pencil, towel and tower.) The timing of the eye movements suggests that, in a sufficiently constrained visual context, spoken word recognition can be achieved in the listener before the word has even finished being spoken. This kind of real-time interaction between visual and linguistic processing is also seen in eye movement patterns during syntactic processing,^[83] spoken language production,^[84] and even natural, unscripted conversation.^[85]

Eye movements also reveal social characteristics and processes of face-to-face communication. Humans are remarkably sensitive to changes in where others are looking. Gibson and Pick^[86] demonstrated that a subject can detect a one-millimeter displacement of a viewer’s gaze away from the bridge of the subject’s nose, which is at the limit of the visual acuity of the eye. The gaze direction of others is a powerful attentional cue, yet is only beginning to be studied empirically.^[87] In analysis of face-to-face communication, however, research has mainly focused on one variable—eye contact between a speaker and a listener. Mutual gaze has two consequences. First, it correlates with many interpersonal factors, such as attractiveness,

authority, and perceived sincerity.^[178] Secondly, the timing of mutual gaze is used to coordinate dialogue. Bavelas and colleagues^[88,89] found that, a listener typically will look more at the speaker than vice versa, but the speaker will periodically cast looks at the listener, creating a period of mutual gaze. At this point, the listener is likely to respond with a nod or “mhm,” prompting the speaker to continue. In this way, eye movements are employed as part of the composite signal used in the collaborative act of conversation.^[90]

MEMORY, IMAGERY, AND DREAMING

Because of their selective and rapid sampling of the visual world, eye movements must recruit memory processes in order to build up some representation of the visual world.^[91] Conversely, it has also been found that memory processes often recruit eye movements.^[92] Richardson and Spivey^[93] demonstrated that subjects will systematically fixate particular empty spaces when questioned about the semantic content of a linguistic event that had previously taken place at that location. Subjects were presented with video clips of four people relating short pieces of factual information. The clips appeared in each of four regions of a grid. While viewing an empty grid, subjects answered a question about one of those facts. The location which had previously been associated with that information received significantly more fixations. This result holds even when the locations move around before the question period, and parallel behavior has been shown in infants as young as six months.^[94] These results suggest that the location of cross-modal events is spatially indexed by the brain, so that relevant locations can be refixated when a memory of the event is activated.

Contemporary neuroscience suggests that memory representations recapitulate perceptual processes to some degree.^[95–98] Researchers have long been intrigued by the relationship between eye movements during a perceptual experience and eye movements that occur when one remembers, imagines, or dreams about that experience. Early empirical investigations found that the frequency of eye movements increases during mental imagery, particularly that of a spatial nature,^[99–103] and an increase in rapid fluttering of the eyes while sleeping correlates with vividness of dreams.^[104–106] A specific account of this relation during picture recognition, known as scanpath theory,^[107] held that a mental representation of a given perceptual experience is composed of a sequence of sensory and motor activities. A picture is recognized partially by replaying the memory of this sequence of eye movements and visual stimulation

and comparing it to the present stimulus. The predictions of this model were borne out by the finding that scanpaths of subjects inspecting an abstract pattern bore a strong similarity to scanpaths of subjects looking at a blank display and recalling the pattern.^[108]

Recent work suggests that eye movements are engaged not only by a memory of a specific perceptual experience, but also by cognitive acts of imagination. Spivey and colleagues^[92,109] asked subjects to listen to short narratives describing events extending either horizontally or vertically (for example, a story about a train pulling out of the station and a story about the apartments in a 40-story building). Subjects' eye movements were recorded while they looked at a blank white screen, or while their eyes were closed. The directionality of saccades corresponded to the directionality of the narratives, suggesting that lower-level motor actions such as eye movements can accompany higher-level cognitive processes. This is because motor actions may be fundamental components of the mental state, rather than being separate functions that are triggered by a mental state.

The notion of a systematic relationship between eye movements and mental activity has percolated into a wider realm, where it has devolved into a series of popular myths and dubious therapeutic techniques. Bandler and Grinder^[110] proposed the theory of neurolinguistic programming (NLP). They observed that when patients were visualizing they tended to look up, when they remembered sounds they looked horizontally, and when they looked downward and to the left they were "accessing their feelings." NLP theory stated that these eye movements to particular locations corresponded to areas of the brain that were being used. Although Bandler and Grinder's observations, theories, and grasp of neuroanatomy have been discredited,^[111-116] NLP is still the basis of a large therapeutic industry.

There have also been claims that eye movements can be used to help alleviate the symptoms of post-traumatic stress syndrome. Shapiro^[117,118] developed the technique known as eye movement desensitization and reprocessing, whereby patients recalled a traumatic episode while following a moving stimulus (such as the therapist's hands) with their eyes. The claim is that this looking behaviour would disrupt the eye movement pattern associated with the traumatic memory, diminishing the visual aspect of the memory and desensitizing the patient to the memory. It is hotly contested whether such methods produce objective improvements.^[119,120] Although they have been empirically shown to help some patients,^[121-123] it could be the case that improvement is due to the general therapeutic practices employed, rather than to the specific effect of eye movement manipulation.^[124]

EYE MOVEMENTS IN THE SERVICE OF COMPLEX TASKS

The research we have discussed so far investigates relatively passive acts of perception and information processing. But eye movement research can inform the study of more complex behaviors, such as selecting a commercial product, playing cricket, or flying a jet plane. In these cases, eye-tracking technology can reveal what different sources of information the subject is using, how frequently they are sampled, and how they affect decisions.

Consumer behavior rests on an interplay between immediate perceptual factors and decision-making processes that can be seen in eye movement behavior. Early in its development, it was realized that eye movement technology could be used to identify factors that determine the amount of attention allocated to different advertisements.^[125-127] Indeed, when advertisements are placed next to vital traffic signs, their over-effectiveness and distracting qualities can also be seen in eye movements.^[128] Such measures have predictive value, since it is the case that gaze durations between competing advertisements correlate with later product choice.^[129,130]

Good visual design is important in a wide range of activities. The eye movements of radar operators,^[131,132] gunners,^[133] and pilots^[134,135] have been recorded and used to evaluate operator skill and equipment design. Similarly, a range of studies employed eye tracking to study the effectiveness of different instrumentation panels,^[55,136-140] and later, different computer-to-user interfaces.^[141-144] Such research has even been used to investigate the ways in which people experience Japanese formal gardens.^[145]

Clear instrumentation, efficient perception, and rapid decision-making are crucial while driving or piloting a vehicle. Early research indicated that eye movements away from the direction of heading were related to steering errors of pilots,^[146] and that saccade frequency was employed as a measure of fatigue in long-distance truck drivers.^[147] Finer analyses investigated what information sources people use, and how frequently they were sampled during different driving conditions.^[148-151] Particularly difficult perceptual/motor tasks, such as cornering, have come under close scrutiny. One or two seconds before average car drivers enter a curve, they direct their gaze to the tangent point on the inside of each bend and leave it there for approximately three seconds.^[152] The direction of the tangent point relative to the car heading is a good predictor of the curvature of the road. Although this anticipatory behavior is not present in novice drivers,^[153,154] in professional race-car drivers it is advanced to the point that the head is rotated proportionally to the estimated car rotation speed,

such that the head is brought in line with the expected tangent points.^[155]

Differences in oculomotor responses between professional and amateur athletes^[156,157] can be seen, and saccadic performance has been claimed to correlate with batting averages in little-leaguers.^[158] More specific analyses have examined differences in scanpaths between novices and experts. For example, skilled golfers are less likely to look at the club during putting, and more likely to look at the ball and the hole.^[159] Land and McLeod^[160] studied the scanpaths of cricket batsmen, who have to judge the trajectory of a fast bowl with an unpredictable bounce and spin based on half a second of viewing time. The batsmen fixated the ball at its time of release and then made a saccade to where they anticipated the bounce. The better batsmen were faster at making this predictive eye movement.

Eye movements not only reveal the strategies of highly trained sportsmen; they can also reveal the more mundane strategies we employ while carrying out everyday tasks, such as making a sandwich. Land and Hayhoe^[161] found that eye movements are tightly linked with moment-to-moment goals and subtasks. This is in contrast to the passive perception of a static scene or picture, when eye movements may be drawn by areas of high contrast or spatial frequency.^[15] Such task-related fixations have been examined in detail using the block-copying task developed by Ballard, Hayhoe, and colleagues.^[162–164]

In these eye tracking experiments, subjects were given the job of constructing a pattern of colored blocks. The pattern was displayed in the model area of a board. Participants took blocks from the resource area, and copied the pattern in the workspace area. The participants' hand actions were recorded, and a headband-mounted eyetracker recorded their eye movements to obtain a window on the strategy used in the task. One method participants could use was to look at the model area and memorize the pattern; each block in turn could be located in the resource area, and then placed in the workspace. A second method, less memory-intensive, would be to remember the color and location of one block from the model, collect it from the resource area, place it in the workspace, and then consult the model again for the next block. The strategy most often used by participants, however, entailed the minimal possible memory demands. Participants would commonly fixate the model, then fixate and pick up a correctly colored block from the resource area, fixate the model again, and then place the block in the workspace. Thus, two fixations per block were made on the model—one to extract color information and one to extract relative spatial location information.^[162] This is a strategy of indexing, whereby just the location of an object is maintained in working

memory, and other properties can be “looked up” as they are needed.^[163] In a computerized, gaze-contingent version of the block-copying task, the color of a block was changed during a subject's saccade.^[164] The subjects rarely noticed this property change, further demonstrating that subjects rely on fixating external information rather than mentally storing it.^[165]

INTERACTIVE APPLICATIONS

The relationship between eye movements and cognitive processing has been employed as a tool to improve ergonomic design and computer interfaces.^[166] The emergence of gaze-contingent eye tracking has allowed eye movements to be further employed as a means to interact with such systems.^[167] Since fixations indicate what parts of the visual world are salient to the user, eye movement information can be used to tailor visual displays. Regions that are not fixated can be rendered at a far lower level of detail,^[168–170] and hence eye tracking can be used to make valuable computation shortcuts. Given the speed and frequency of naturally occurring eye movements, and their close relation to perceptual and cognitive processing, they also have the potential to be an efficient and expressive input device.

In the first applications of this technology, eye gaze replaced the hand-operated mouse, allowing users to select icons and menus in a graphical user interface. It takes very little training for naïve users to employ their gaze as a means to control a computer.^[171] Although gaze is also faster than a mouse, there is no clear equivalent of a button press. Blinks were found to be unnatural for users, so dwell times were used as a selection criterion. Initially, this led to the so-called Midas touch problem: Users could not distinguish between items they wanted to visually inspect and items they wanted to select, so everything they looked at would be chosen. If a long enough dwell time is used, typically 600–1000 ms, this can be avoided. If the dwell time is very carefully calibrated to be just longer than a normal fixation, then users report an eerie sense that the system knows what they want to do before they do.^[172]

This technology can be employed by users who are unable to operate standard input devices because of physical disability.^[173] One such application is eye typing, in which users can enter characters by fixating on the keys of a virtual keyboard.^[174] There are currently limitations to such implementations. Accuracy issues determine the size of buttons on the screen, such that a whole alphabet cannot always be employed at once.^[173] It can take severely disabled users many months to learn this form of interaction,^[175] because they have had little experience actively controlling devices. The speed of eye typing is typically a character a

second.^[176] In order to be a successful, real-time communication tool, eye typing systems have been developed in which the user selects a word or phrase from a hierarchical menu.^[177]

CONCLUSION

Eye movements are driven both by properties of the visual world and by processes in a person's mind. Uniquely poised between perception and cognition, eye movements are an invaluable tool for psychologists. Compared to the single data point provided by a button-press reaction time, the eye movements of a subject can provide researchers with a rich, dynamic data source concerning the temporal dynamics and psychological processes that led up to the response. These properties are also of great value to designers and engineers, because they allow for detailed measurements of how a user is interacting with a device. Since the technology has become highly efficient, such information can now be fed back into devices in real time, and the movements of a user's eyes can be used to issue commands or tailor computational processes. Although such applications are currently in their infancy, this most frequent of all human behaviors could turn out to provide the most fluid and expressive interface between humans and computers.

ARTICLES OF FURTHER INTEREST

Ergonomics; Eye Tracking: Characteristics and Methods; Medical Robotics

REFERENCES

1. Posner, M.I. Orienting of attention. *Q. J. Exp. Psychol.* **1980**, *32* (1), 3.
2. Treue, S. Neural correlates of attention in primate visual cortex. *Trends Neurosci.* **2001**, *24*, 295.
3. Findlay, J.M.; Gilchrist, I.D. *Eye Guidance in Reading and Scene Perception*; Underwood, G., Ed.; Elsevier Science, Ltd.: Oxford, England, 1998; 295–312.
4. Sheliga, B.M.; Riggio, L.; Craighero, L.; Rizzolatti, G. Spatial attention—Determined modifications in saccade trajectories. *Neuroreport: Int. J. Rapid Commun. Res. Neurosci.* **1995**, *6* (3), 585.
5. Corbetta, M. Frontoparietal cortical networks for directing attention and the eye to visual locations: Identical, independent, or overlapping neural systems? *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 831.
6. Shepherd, M.; Findlay, J.M.; Hockey, R.J. The relationship between eye movements and spatial attention. *Q. J. Exp. Psychol., Human Exp. Psychol.* **1986**, *38* (3-A), 475.
7. Sheliga, B.M.; Riggio, L.; Rizzolatti, G. Orienting of attention and eye movements. *Exp. Brain Res.* **1994**, *98* (507–522).
8. Hoffman, J.E.; Subramaniam, B. The role of visual attention in saccadic eye movements. *Percept. Psychophys.* **1995**, *57* (6), 787.
9. Robinson, D.A.; Fuchs, A.F. Eye movements evoked by stimulation of frontal eye fields. *J. Neurophysiol.* **1969**, *32*.
10. Moore, T.; Armstrong, K.M. Selective gating of visual signals by microstimulation of frontal cortex. *Nature* **2003**, *421* (6921), 370.
11. Moore, T.; Fallah, M. Control of eye movements and spatial attention. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (3), 1273.
12. Buswell, G.T. *How People Look at Pictures: A Study of the Psychology and Perception in Art*; Univ. Chicago Press: Oxford, England, 1935; 198 pp.
13. Wooding, D.S. Eye movements of large populations: II. Deriving regions of interest, coverage, and similarity using fixation maps. *Behav. Res. Meth. Instrum. Comput.* **2002**, *34* (4), 518.
14. Wooding, D.S.; Muggelstone, M.D.; Purdy, K.J.; Gale, A.G. Eye movements of large populations: I. Implementation and performance of an autonomous public eye tracker. *Behav. Res. Meth. Instrum. Comput.* **2002**, *34* (4), 509.
15. Henderson, J.M. Human gaze control in real-world scene perception. *Trends Cogn. Sci.* **2003**, *7*, 498.
16. Mannan, S.K.; Ruddock, K.H.; Wooding, D.S. The relationship between the locations of spatial features and those of fixations made during visual examination of briefly presented images. *Spat. Vis.* **1996**, *10* (3), 165.
17. Mannan, S.K.; Ruddock, K.H.; Wooding, D.S. Fixation patterns made during brief examination of two-dimensional images. *Perception* **1997**, *26* (8), 1059.
18. Reinagel, P.; Zador, A.M. Natural scene statistics at the center of gaze. *Network* **1999**, *10*, 314.
19. Li, Z. A saliency map in primary visual cortex. *Trends Cogn. Sci.* **2002**, *6*, 9.
20. Itti, L.; Koch, C. A saliency—Based search mechanism for overt and covert shifts of visual attention. *Vision Res.* **2000**, *40* (10–12), 1489.
21. Yarbus, A.L. *Role of Eye Movements in the Visual Process*; Nauka: Oxford, England, 1965; 167 pp.
22. Loftus, G.R.; Mackworth, N.H. Cognitive determinants of fixation location during picture viewing. *J. Exp. Psychol. Hum. Percept. Perform.* **1978**, *4* (4), 565.
23. Posner, M.I.; Rafal, R.D.; Choate, L.S.; Vaughan, J. Inhibition of return: Neural basis and function. *Cogn. Neuropsychol.* **1985**, *2* (3), 211.
24. Leek, E.C.; Reppa, I.; Tipper, S.P. Inhibition of return for objects and locations in static displays. *Percept. Psychophys.* **2003**, *65* (3), 388.
25. Irwin, D.E. Integrating information across saccadic eye movements. *Curr. Dir. Psychol. Sci.* **1996**, *5* (3), 94.
26. Irwin, D.E. Information integration across saccadic eye movements. *Cogn. Psychol.* **1991**, *23* (3), 420.

27. Peterson, M.S.; Kramer, A.F.; Wang, R.F.; Irwin, D.E.; McCarley, J.S. Visual search has memory. *Psychol. Sci.* **2001**, *12* (4), 287.
28. Henderson, J.M.; Hollingworth, A. Global transsaccadic change blindness during scene perception. *Psychol. Sci.* **2003**, *14* (5), 493.
29. Simons, D.J.; Levin, D. Change blindness. *Trends Cogn. Sci.* **1997**, *1*, 261.
30. Rensink, R.A.; O'Regan, J.K.; Clark, J. Image flicker is as good as saccades in making large scene changes invisible. *Perception* **1995**, *24*, 26.
31. O'Regan, J.K.; Deubel, H.; Clark, J.; Rensink, R.A. Picture changes during blinks: Looking without seeing and seeing without looking. *Vis. Cognit.* **2000**, *7*.
32. O'Regan, J.K.; Rensink, R.A.; Clark, J.J. Change-blindness as a result of "mudsplashes". *Nature* **1999**, *398*.
33. Bridgeman, B.; Hendry, D.; Stark, L. Failure to detect displacement of the visual world during saccadic eye movements. *Vis. Res.* **1975**, *15* (6), 719.
34. Henderson, J.M.; Hollingworth, A. The role of fixation position in detecting scene changes across saccades. *Psychol. Sci.* **1999**, *10* (5), 438.
35. Hollingworth, A.; Schrock, G.; Henderson, J.M. Change detection in the flicker paradigm: The role of fixation position within the scene. *Mem. Cogn.* **2001**, *29* (2), 296.
36. Nelson, W.W.; Loftus, G.R. The functional visual-field during picture viewing. *J. Exp. Psychol. Hum. Learn. Mem.* **1980**, *6* (4), 391.
37. Diefendorf, A.R.; Dodge, R. An experimental study of the ocular reactions of the insane from photographic records. *Brain* **1908**, *31*, 451–492.
38. Holzman, P.S.; Proctor, L.R.; Hughes, D.W. Eye-tracking patterns in schizophrenia. *Science* **1973**, *181*.
39. Holzman, P.S.; et al. Eye-tracking dysfunctions in schizophrenic patients and their relatives. *Arch. Gen. Psychiat.* **1974**, *31* (2), 143.
40. McDowell, J.E.; Clementz, B.A.; Wixted, J.T. Timing and amplitude of saccades during predictive saccadic tracking in schizophrenia. *Psychophysiology* **1996**, *33* (1), 93.
41. Karoumi, B.; Ventre Dominey, J.; Dalery, J. Predictive saccade behavior is enhanced in schizophrenia. *Cognition* **1998**, *68* (3), B81.
42. Sereno, A.B.; Holzman, P.S. Antisaccades and smooth pursuit eye movements in schizophrenia. *Biol. Psychiat.* **1995**, *37* (6), 394.
43. Gooding, D.C. Antisaccade task performance in questionnaire—Identified schizotypes. *Schizophr. Res.* **1999**, *35* (2), 157.
44. Fukushima, J.; et al. Disturbances of voluntary control of saccadic eye movements in schizophrenic patients. *Biol. Psychiat.* **1988**, *23* (7), 670.
45. Williams, L.M.; Loughland, C.M.; Green, M.J.; Harris, A.W.F.; Gordon, E. Emotion perception in schizophrenia: An eye movement study comparing the effectiveness of risperidone vs. haloperidol. *Psychiat. Res.* **2003**, *120* (1), 13.
46. Shimizu, T.; et al. Comparison of eye-movement patterns in schizophrenic and normal adults during examination of facial affect displays. *Percept. Motor Skills* **2000**, *91* (3, Pt 2), 1045.
47. Radant, A.D.; Claypoole, K.; Wingerson, D.K.; Cowley, D.S.; Roy Byrne, P.P. Relationships between neuropsychological and oculomotor measures in schizophrenia patients and normal controls. *Biol. Psychiatry* **1997**, *42* (9), 797.
48. Ross, R.G. Early expression of a pathophysiological feature of schizophrenia: Saccadic intrusions into smooth-pursuit eye movements in school-age children vulnerable to schizophrenia. *J. Am. Acad. Child Adolesc. Psychiat.* **2003**, *42* (4), 468.
49. Rayner, K. Eye movements in reading and information processing: 20 Years of research. *Psychol. Bull.* **1998**, *124* (3), 372.
50. Rayner, K. Eye movements in reading and information processing. *Psychol. Bull.* **1978**, *85* (3), 618.
51. Pollatsek, A.; Rayner, K.; Collins, W.E. Integrating pictorial information across eye movements. *J. Exp. Psychol. Gen.* **1984**, *113* (3), 426.
52. Morrison, R.E.; Rayner, K. Saccade size in reading depends upon character spaces and not visual angle. *Percept. Psychophys.* **1981**, *30* (4), 395.
53. Carpenter, P.A.; Just, M.A. *Eye Movements in Reading: Perceptual and Language Processes*; Rayner, K., Ed.; Academic Press: New York, 1983; 275–307.
54. Rayner, K.; McConkie, G.W. What guides a reader's eye movements? *Vis. Res.* **1976**, *16* (8), 829.
55. Kolers, P.A.; Duchnicky, R.L.; Ferguson, D.C. Eye movement measurement of readability of CRT displays. *Hum. Factors* **1981**, *23* (5), 517.
56. Morrison, R.E.; Inhoff, A.W. Visual factors and eye movements in reading. *Visible Lang.* **1981**, *15* (2), 129.
57. Tinker, M.A. Legibility and eye movement in reading. *Psychol. Bull.* **1927**, *24*.
58. Ferreira, F.; Clifton, C. The independence of syntactic processing. *J. Mem. Lang.* **1986**, *25*, 348.
59. Rayner, K.; Carlson, M.; Frazier, L. The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *J. Verbal Learn. Verbal Behav.* **1983**, *22* (3), 358.
60. Trueswell, J.C.; Tanenhaus, M.K.; Kello, C. Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *J. Exp. Psychol. Learn. Mem. Cogn.* **1993**, *19* (3), 528.
61. Tinker, M.A. A photographic study of eye movements in reading formulae. *Gen. Psychol. Monogr.* **1928**, *3*.
62. Rayner, K.; Sereno, S.C.; Morris, R.K.; Schmauder, A.R. Eye movements and on-line language comprehension processes. *Lang. Cogn. Processes* **1989**, *4* (3–4), 21.
63. Jacobson, J.Z.; Dodwell, P.C. Saccadic eye movements during reading. *Brain Lang.* **1979**, *8* (3), 303.
64. Levy Schoen, A. *Eye Movements: Cognition and Visual Perception*; Fisher, D.F., Monty, R.A., Senders, J.W., Eds.; Erlbaum: Hillsdale, NJ, 1981; 299–316.
65. Morris, R.K.; Rayner, K.; Pollatsek, A. Eye movement guidance in reading: The role of parafoveal letter and space information. *J. Exp. Psychol. Hum. Percept. Perf.* **1990**, *16* (2), 268.

66. McConkie, G.W.; Rayner, K. The span of the effective stimulus during a fixation in reading. *Percept. Psychophys.* **1975**, *17* (6), 578.
67. McConkie, G.W.; Rayner, K. Asymmetry of the perceptual span in reading. *Bull. Psychon. Soc.* **1976**, *8* (5), 365.
68. Rayner, K.; McConkie, G.W.; Zola, D. Integrating information across eye movements. *Cogn. Psychol.* **1980**, *12* (2), 206.
69. Pollatsek, A.; Bolozky, S.; Well, A.D.; Rayner, K. Asymmetries in the perceptual span for Israeli readers. *Brain Lang.* **1981**, *14* (1), 174.
70. Osaka, N.; Oda, K. Effective visual field size necessary for vertical reading during Japanese text processing. *Bull. Psychon. Soc.* **1991**, *29* (4), 345.
71. Kennedy, A.; Murray, W.S. Spatial coordinates and reading: Comments on Monk. *Q. J. Exp. Psychol.* **1987**, *39A*, 649.
72. Murray, W.S.; Kennedy, A. Spatial coding in the processing of anaphor by good and poor readers: Evidence from eye movement analyses. *Q. J. Exp. Psychol. Hum. Exp. Psychol.* **1988**, *40* (4-A), 693.
73. Just, M.A.; Carpenter, P.A. A theory of reading: From eye fixations to comprehension. *Psychol. Rev.* **1980**, *87* (4), 329.
74. Dixon, W.R., Jr. Studies of the eye-movements in reading of university professors and graduate students. *Microfilm Abstr.* **1948**, *8*.
75. Robinson, F.P. The role of eye movement habits in determining reading efficiency. *Proc. Iowa Acad. Sci.* **1932**, *39*.
76. Eurich, A.C. Fourth-and fifth-grade standards for photographic eye-movement records. *J. Genet. Psychol.* **1933**, *43*.
77. Eurich, A.C. The photographic eye-movement records of successful and unsuccessful college students. *J. Appl. Psychol.* **1933**, *17*.
78. Miles, W.R.; Bell, H.M. Eye-movement records in the investigation of study habits. *J. Exp. Psychol.* **1929**, *12*.
79. McKane, P.F.; Maples, W.C.; Sellars, P.; McNeil, M. A comparison of auditory/language therapy with school visual support procedures in a public school setting. *J. Optim. Vis. Dev.* **2001**, *32* (2), 83.
80. *The Integration of Language, Vision, and Action: Eye movements and the Visual World*; Henderson, J.M., Ferreira, F., Eds.; Psychology Press: New York, 2004.
81. Tanenhaus, M.K.; Spivey Knowlton, M.J.; Eberhard, K.M.; Sedivy, J.C. Integration of visual and linguistic information in spoken language comprehension. *Science* **1995**, *268* (5217), 1632.
82. Allopenna, P.D.; Magnuson, J.S.; Tanenhaus, M.K. Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *J. Mem. Lang.* **1998**, *38* (4), 419–439.
83. Spivey, M.J.; Tanenhaus, M.K.; Eberhard, K.M.; Sedivy, J.C. Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. *Cognit. Psychol.* **2002**, *45* (4), 447.
84. Griffin, Z.M.; Bock, K. What the eyes say about speaking. *Psychol. Sci.* **2000**, *11* (4), 274.
85. Brown-Schmidt, S.; Campana, E.; Tanenhaus, M.K. *World-Situated Language Processing: Bridging the Language as Product and Language as Action Traditions*; Trueswell, J.C., Tanenhaus, M.K., Eds.; MIT Press: Cambridge, 2004.
86. Gibson, J.J.; Pick, A.D. Perception of another persons looking behavior. *Am. J. Psychol.* **1963**, *76* (3), 386.
87. Kingstone, A.; Smilek, D.; Ristic, J.; Friesen, C.K.; Eastwood, J.D. Attention, researchers! It is time to take a look at the real world. *Curr. Dir. Psychol. Sci.* **2003**, *12* (5), 176.
88. Bavelas, J.B.; Coates, L.; Johnson, T. Listener responses as a collaborative process: The role of gaze. *J. Commun.* **2002**, *52* (3), 566.
89. Bavelas, J.B.; Chovil, N. Visible acts of meaning—An integrated message model of language in face-to-face dialogue. *J. Lang. Soc. Psychol.* **2000**, *19* (2), 163.
90. Clark, H.H. *Using Language*; Cambridge University Press: Cambridge, 1996; pp. xi, 432.
91. Irwin, D.E.; Zacks, J.L.; Brown, J.S. Visual memory and the perception of a stable visual environment. *Percept. Psychophys.* **1990**, *47* (1), 35.
92. Spivey, M.J.; Geng, J.J. Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychol. Res./Psychol. Forsch.* **2001**, *65* (4), 235.
93. Richardson, D.C.; Spivey, M.J. Representation, space and hollywood squares: Looking at things that aren't there anymore. *Cognition* **2000**, *76*, 269.
94. Richardson, D.C.; Kirkham, N.Z. Multi-modal events and moving locations: Evidence for dynamic spatial indexing in adults and six month olds. *J. Exp. Psychol. Gen.* **2004**, *133* (1), 46–62.
95. Kosslyn, S.M.; Behrmann, M.; Jeannerod, M. The cognitive neuroscience of mental imagery. *Neuropsychologia* **1995**, *33* (11), 1335.
96. Kan, I.P.; Barsalou, L.W.; Solomon, K.O.; Minor, J.K.; Thompson-Schill, S.L. Role of mental imagery in a property verification task: fMRI evidence for perceptual representations of conceptual knowledge. *Cogn. Neuropsychol.* **2003**, *20* (3–6), 525–540.
97. Barsalou, L.W. Perceptual symbol systems. *Behav. Brain Sci.* **1999**, *22* (4), 577.
98. Martin, A. *Handbook of Functional NeuroImaging of Cognition*; Cabeza, R., Kingstone, A., Eds.; MIT Press: Cambridge, 2001; 153–186.
99. Totten, E. Eye movement during visual imagery. *Comp. Psychol. Monogr.* **1935**, *11*.
100. Perky, C.W. An experimental study of imagination. *Am. J. Psychol.* **1910**, *21*, 422.
101. Stoy, E.G. A preliminary study of ocular attitudes in thinking of spatial relations. *J. Gen. Psychol.* **1930**, *4*.
102. Clark, H. Visual imagery and attention: An analytical study. *Am. J. Psychol.* **1916**, *27* (4), 461.
103. Goldthwait, C. Relation of eye movements to visual imagery. *Am. J. Psychol.* **1933**, *45*.
104. Goodenough, D.R.; Shapiro, A.; Holden, M.; Steinschriber, L. A comparison of “dreamers” and “non-dreamers”: Eye movements, electroencephalograms, and the recall of dreams. *J. Abnorm. Soc. Psychol.* **1959**, *59*.

105. Roffwarg, H.P.; Dement, W.C.; Muzio, J.N.; Fisher, C. Dream imagery: Relationship to rapid eye movements of sleep. *Arch. Gen. Psych.* **1962**, *7*.
106. Antrobus, J.S.; Antrobus, J.S. Rapid eye movements and rapid eye movement periods. *Psychophysiology* **1969**, *6*.
107. Noton, D.; Stark, L. Scanpaths in eye movements during pattern perception. *Science* **1971**, *171*.
108. Brandt, S.A.; Stark, L.W. Spontaneous eye movements during visual imagery reflect the content of the visual scene. *J. Cogn. Neurosci.* **1997**, *9* (1), 27.
109. Spivey, M.J.; Tyler, M.; Richardson, D.C.; Young, E. Paper presented at the 22nd Annual Conference of the Cognitive Science Society, Philadelphia, 2000.
110. Bandler, R.; Grinder, J. *The Structure of Magic*; Science and Behavior Books: Palo Alto, CA, 1975; pp. xv, 225.
111. Gumm, W.B.; Walker, M.K.; Day, H.D. Neuro-linguistics programming: Method or myth? *J. Couns. Psychol.* **1982**, *29* (3), 327.
112. Poffel, S.A.; Cross, H.J. Neurolinguistic programming: A test of the eye-movement hypothesis. *Percept. Mot. Skills* **1985**, *61* (3, Pt 2), 1262.
113. Farmer, A.; Rooney, R.; Cunningham, J.R. Hypothesized eye movements of neurolinguistic programming: A statistical artifact. *Percept. Mot. Skills* **1985**, *61* (3, Pt 1), 717.
114. Elich, M.; Thompson, R.W.; Miller, L. Mental imagery as revealed by eye movements and spoken predicates: A test of neurolinguistic programming. *J. Couns. Psychol.* **1985**, *32* (4), 622.
115. Salas, J.A.; de Groot, H.; Spanos, N.P. Neuro-linguistic programming and hypnotic responding: An empirical evaluation. *J. Ment. Imag.* **1989**, *13* (1), 79.
116. Jupp, J.J. A further empirical evaluation of neuro-linguistic primary representational systems (PRS). *Couns. Psychol. Q.* **1989**, *2* (4), 441.
117. Shapiro, F. *Eye Movement Desensitization and Reprocessing: Basic Principles, Protocols, and Procedures*; Guilford Press: New York, NY, 1995; . pp. xviii, 398. .
118. Shapiro, F. EMDR 12 years after its introduction: Past and future research. *J. Clin. Psychol.* **2002**, *58* (1), 1.
119. Oswalt, R.; Anderson, M.; Hagstrom, K.; Berkowitz, B. Evaluation of the one-session eye-movement desensitization reprocessing procedure for eliminating traumatic memories. *Psychol. Rep.* **1993**, *73* (1), 99.
120. Acierno, R.; Hersen, M.; Van Hasselt, V.B.; Tremont, G.; Meuser, K.T. Review of the validation and dissemination of eye-movement desensitization and reprocessing: A scientific and ethical dilemma. *Clin. Psychol. Rev.* **1994**, *14* (4), 287.
121. Forbes, D.; Creamer, M.; Rycroft, P. Eye movement desensitization and reprocessing in posttraumatic stress disorder: A pilot study using assessment measures. *J. Behav. Ther. Exp. Psychiatry* **1994**, *25* (2), 113.
122. Silver, S.M.; Brooks, A.; Obenchain, J. Treatment of Vietnam War veterans with PTSD: A comparison of eye movement desensitization and reprocessing, bio-feedback, and relaxation training. *J. Trauma. Stress* **1995**, *8* (2), 337.
123. Wilson, S.A.; Becker, L.A.; Tinker, R.H. Eye movement desensitization and reprocessing (EMDR) treatment for psychologically traumatized individuals. *J. Consult. Clin. Psychol.* **1995**, *63* (6), 928.
124. Hyer, L.; Brandsma, J.M. EMDR minus eye movements equals good psychotherapy. *J. Trauma. Stress* **1997**, *10* (3), 515.
125. Brandt, H.F. Ocular photography as a scientific approach to the study of the psychological aspects of seeing. *Illum. Eng. Inc. Trans. Illum. Eng. Soc. N.Y.* **1944**, *39*.
126. Macnamara, J.J. A new method for testing advertising effectiveness through eye movement photography. *Psychol. Rec.* **1941**, *4*.
127. Karlake, J.S. The Purdue eye-camera: A practical apparatus for studying the attention value of advertisements. *J. Appl. Psychol.* **1940**, *24*.
128. Boersema, T.; Zwaga, H.J.G. *Visual Search*; Brogan, D., Ed.; Taylor & Francis: Philadelphia, PA, 1990; 151–157.
129. Lohse, G.L. Consumer eye movement patterns on yellow pages advertising. *J. Advert.* **1997**, *26* (1), 61.
130. Treisman, J.; Gregg, J.P. Visual, verbal, and sales responses to print ads. *J. Advert. Res.* **1979**, *19* (4), 41.
131. Gerathewohl, S.J. Eye movements during radar operations. *J. Aviat. Med.* **1952**, *23*.
132. Wallis, D.; Samuel, J.A. Some experimental studies of radar operating. *Ergonomics* **1961**, *4*.
133. Brues, A.M.; Damon, A. *Eye movement in sighting as related to design of turret sighting panels*; 1946. (AAF, Tech. Rep. 1943; Publ. Bd., No. 5183.).
134. Milton, J.L. Analysis of pilots' eye movements in flight. *J. Aviat. Med.* **1952**, *23*.
135. Kamyshev, I.A.; Lazarev, V.G. Some features of changes in a pilot's fixations in reading instruments during actual flight. *Vopr. Psychologii.* **1969**, *15*.
136. Steinke, T.R. An evaluation of map design and map reading using eye movement recordings. *Dissert. Abstr. Int.* **1980**, *40* (8-B), 3639.
137. Haider, E.; Luczak, H.; Rohmert, W. Ergonomics investigations of workplaces in a police command-control centre equipped with TV displays. *Appl. Ergon.* **1982**, *13* (3), 163.
138. Swanston, M.T.; Walley, C.E. Factors affecting the speed of acquisition of tabulated information from visual displays. *Ergonomics* **1984**, *27* (3), 321.
139. Leermakers, M.A.; Boschman, M.C. Eye movements, performance and visual comfort using VDTs. *IPO Ann. Prog. Rep.* **1984**, *19*.
140. Moray, N.; Rotenberg, I. Fault management in process control: Eye movements and action. *Ergonomics* **1989**, *32* (11), 1319.
141. MacGregor, J.N.; Lee, E. Menu search: Random or systematic? *Int. J. Man-Mach. Stud.* **1987**, *26* (5), 627.
142. Yamamoto, S.; Kuto, Y. A method of evaluating VDT screen layout by eye movement analysis. *Ergonomics* **1992**, *35* (5–6), 591.
143. Deffner, G. *Eye Movement Research: Mechanisms, Processes and Applications. Studies in Visual Information Processing*, 6; Findlay, J.M., Walker, R., Eds.; Elsevier Science: New York, NY, 1995; 479–490.

144. Menozzi, M. Visual ergonomics of head-mounted displays. *Jpn. Psychol. Res.* **2000**, *42* (4), 213.
145. Ohno, R.; Hata, T.; Kondo, M. *Handbook of Japan United States Environment Behavior Research: Toward a Transactional Approach*; Wapner, S., Demick, J., Eds.; Plenum Press: New York, NY, 1997; 163–182.
146. Martin, G. Eye movements as the cause of faulty steering of airplanes, automobiles, and bicycles. *Graefes Arch. Ophthalmol.* **1940**, *142*.
147. Specht, H. Fatigue and hours of service of interstate truck drivers. V: Eye movements and related phenomena. *Public Health Bulletin (Wash. DC)* **1941**, 265.
148. Mourant, R.R.; Rockwell, T.H. Mapping eye-movement patterns to the visual scene in driving: An exploratory study. *Hum. Factors* **1970**, *12*.
149. Mourant, R.R.; Rockwell, T.H. Strategies of visual search by novice and experienced drivers. *Hum. Factors* **1972**, *14*.
150. Robinson, G.H.; Erickson, D.J.; Thurston, G.L.; Clark, R.L. Visual search by automobile drivers. *Hum. Factors* **1972**, *14*.
151. Kito, T.; Haraguchi, M.; Funatsu, T.; Sato, M.; Kondo, M. Measurements of gaze movements while driving. *Percept. Mot. Skills* **1989**, *68* (1), 19.
152. Land, M.F.; Lee, D.N. Where do we look when we steer. *Nature* **1994**, *369* (6483), 742.
153. Dishart, D.C.; Land, M.F. *Eye Guidance in Reading and Scene Perception*; Underwood, G., Ed.; Elsevier Science, Ltd.: Oxford, England, 1998; 419–429.
154. Underwood, G.; Chapman, P.; Brocklehurst, N.; Underwood, J.; Crundall, D. Visual attention while driving: Sequences of eye fixations made by experienced and novice drivers. *Ergonomics* **2003**, *46* (6), 629.
155. Land, M.F.; Tatler, B.W. Steering with the head: The visual strategy of a racing driver. *Curr. Biol.* **2001**, *11* (15), 1215.
156. Harbin, G.; Durst, L.; Harbin, D. Evaluation of oculomotor response in relationship to sports performance. *Med. Sci. Sports Exerc.* **1989**, *21* (3), 258.
157. Lenoir, M.; Crevits, L.; Goethals, M.; Wildenbeest, J.; Musch, E. Are better eye movements an advantage in ball games? A study of prosaccadic and antisaccadic eye movements. *Percept. Mot. Skills* **2000**, *91* (2), 546.
158. Trachtman, J.N. The relationship between ocular motilities and batting average in little leaguers. *Am. J. Opt. Arch. Am. Acad. Opt.* **1973**, *50*.
159. Vickers, J.N. Gaze control in putting. *Perception* **1992**, *21* (1), 117.
160. Land, M.F.; McLeod, P. From eye movements to actions: How batsmen hit the ball. *Nat. Neurosci.* **2000**, *3* (12), 1340.
161. Land, M.F.; Hayhoe, M. In what ways do eye movements contribute to everyday activities? *Vis. Res.* **2001**, *41* (25–26), 3559.
162. Ballard, D.H.; Hayhoe, M.M.; Pelz, J.B. Memory representations in natural tasks. *J. Cogn. Neurosci.* **1995**, *7* (1), 66.
163. Ballard, D.H.; Hayhoe, M.M.; Pook, P.K.; Rao, R.P.N. Deictic codes for the embodiment of cognition. *Behav. Brain Sci.* **1997**, *20*, 723.
164. Hayhoe, M.M.; Bensinger, D.G.; Ballard, D.H. Task constraints in visual working memory. *Vis. Res.* **1998**, *38* (1), 125.
165. Spivey, M.J.; Richardson, D.C.; Fitneva, S. *Interfacing Language, Vision, and Action*; Henderson, J.M., Ferreira, F., Eds.; Academic Press: San Diego, 2004.
166. Kramer, A.F.; McCarley, J.S. Oculomotor behaviour as a reflection of attention and memory processes: Neural mechanisms and applications to human factors. *Theor. Issues Ergon. Sci.* **2003**, *4* (1–2), 21.
167. Duchowski, A.T. A breadth-first survey of eye-tracking applications. *Behav. Res. Meth., Instrum. Comput.* **2002**, *34* (4), 455.
168. Watson, B.; Walker, N.; Hodges, L.F.; Worden, A. Managing level of detail through peripheral degradation: Effects on search performance with a head-mounted display. *ACM Trans. Comput.-Hum. Interact.* **1997**, *4*, 323.
169. Santella, A.; DeCarlo, D. Abstracted Painterly Renderings Using Eye-Tracking Data. In *Second International Symposium on Non-Photorealistic Animation and Rendering (NPAR)*; ACM Press: New York, NY, U.S.A., 2002.
170. Ohshima, T.; Yamamoto, H.; Tamura, H. *Paper Presented at the VRAIS, Los Alamitos, CA*; 1996.
171. Stampe, D.M.; Reingold, E.M. *Eye Movement Research: Mechanisms, Processes and Applications*; Findlay, J.M., Walker, R., Kentridge, R.W., Eds.; Elsevier: Amsterdam, 1995; 467–478.
172. Jacob, R.J.K. The use of eye movements in human-computer interaction techniques: What you look at is what you get. *ACM Trans. Inf. Syst.* **1991**, *9* (3), 152.
173. Frey, L.A.; White, K.P.; Hutchinson, T.E. Eye-gaze word processing. *IEEE Trans. Syst. Man Cybern.* **1990**, *20* (4), 944.
174. Majaranta, P.; Raiha, K.-J. *Eye Tracking Research & Applications: Proceedings of the Symposium on ETRA 2002*; ACM Press: New York, 2002; 15–22.
175. Gips, J.; Dimattia, P.; Curran, F.X.; Olivieri, P. *Interdisciplinary Aspects on Computers Helping People with Special Needs*; Klaus, J., Auff, E., Kremser, W., Zagler, W., Eds.; Oldenburg: Vienna, 1996.
176. Kahn, D.A.; Heynen, J.; Snuggs, G.L. Eye-Controlled Computing: The VisionKey Experience. In *Proceedings of the Fourteenth International Conference on Technology and Persons with Disabilities (CSUN'99)*; ACM Press: New York, NY, U.S.A., 1999.
177. Chapman, J.E. Use of an Eye-Operated Computer System in Locked-in Syndrome. In *Proceedings of the Sixth Annual International Conference on Technology and Persons with Disabilities*; ACM Press: New York, NY, U.S.A., 1991.
178. Argyle, M.; Cook, M. *Gaze and Mutual Gaze*; Cambridge University Press: Cambridge.

Fabrics

Frank K. Ko

Fibrous Materials Laboratory, Department of Materials Science and Engineering, Drexel University, Philadelphia, Pennsylvania, U.S.A.



INTRODUCTION

Fabrics are fibrous structures manufactured by the interlacing, interlooping, intertwining, or direct bonding of oriented or random, discrete or continuous, fibers into 2-D or 3-D textile assemblies. Textile surgical implants are fibrous structures placed into the human body for a period of time to perform a specific clinical function. Surgical implants can be classified according to their end uses or by their functions. Based on the location of use in the body, surgical implants can be categorized as cardiovascular and thoracic implants, orthopaedic implants, or implants for plastic and constructive surgery. The functions of these implants can include replacement of parts of the anatomy, such as hip and joints; simulation of a congenitally absent portion of the anatomy, such as facial prostheses; correction of congenital, traumatic, or pathological deformity (spinal plates); and rectifying an organ's mode of operation (pacemaker), as well as therapeutic functions and drug delivery.

In this article, the wealth of fiber architecture in fabric forms and their characteristics, which are critical for the selection and design of implants for surgery, are introduced. A review of the basic components of fabrics and the various classes of fabric structures is provided. The characteristics of the various fabrics are illustrated in performance maps, and potential applications of fabrics for tissue engineering scaffolds are illustrated.

OVERVIEW

Even though the use of surgical implants dates back to ancient Egyptian days, the modern use of surgical implants did not begin until the 19th century. The use of textile- and fabric-based surgical implants has greatly expanded since the development of synthetic fibers in the 1940s and 1950s. Some well-established textile surgical implants include artificial arteries, hernia mesh, sutures, bone fixation plates, hip joints,^[1,2] and more recently, material for the regeneration of tissues and organs.^[3,4]

It is well established that the fundamental building blocks of living systems are fibrous in nature. From the

1.5-nm double-helix strand of DNA molecules (including cytoskeleton filaments with diameters around 30 nm) to sensory cells such as hair cells and the rod cells of the eyes, fibers form the extracellular matrices—the multifunctional structural backbone—for tissues and organs.^[5] Whereas collagen fibrils constitute the structural backbone of connective tissues such as tendons and ligaments, specific junctions between various cells conduct electrical and chemical signals that result from various kinds of stimulation. The signals direct normal functions of the cells such as energy storage, information storage and retrieval, tissue regeneration, and sensing. Figure 1 illustrates the hierarchical fibrous nature of biological tissues.^[5] The architecture of these natural fibrous structures is a blueprint for the design and fabrication of replacement body parts and the creation of scaffolds for tissue regeneration.

Man-made or regenerated fibrous structures in fabric form are attractive candidates for surgical implants. Because of their large length-to-diameter ratio and flexibility, metallic, ceramic and polymeric materials in fibrous form can be organized into textile structures in linear, two-dimensional, and three-dimensional structures. The enormous possibilities of combinations of material properties and structural geometry can result in structures exhibiting one or more of the following characteristics:

- High strength-to-weight ratio.
- High voluminosity.
- Controlled modulus, extensibility, flexibility, pliability.
- Conforming behavior.
- Reactive or nonreactive to body environment.
- Controlled porosity and permeability.
- Ease of suturing, clamping, or joining.
- Compatible surface texture for tissue adhesion.
- Structural similarity to body tissues and organs.

Available fibers can range from the very strong aramid or gel-spun polyethylene to the extremely extensible elastomers. The properties of some of the candidate fibers are shown in Table 1.^[6] The stress-strain behavior of a wide range of fibers tested in ambient and physiological conditions is shown in Fig. 2.^[7,10] As a result of their higher molecular orientation, the tensile

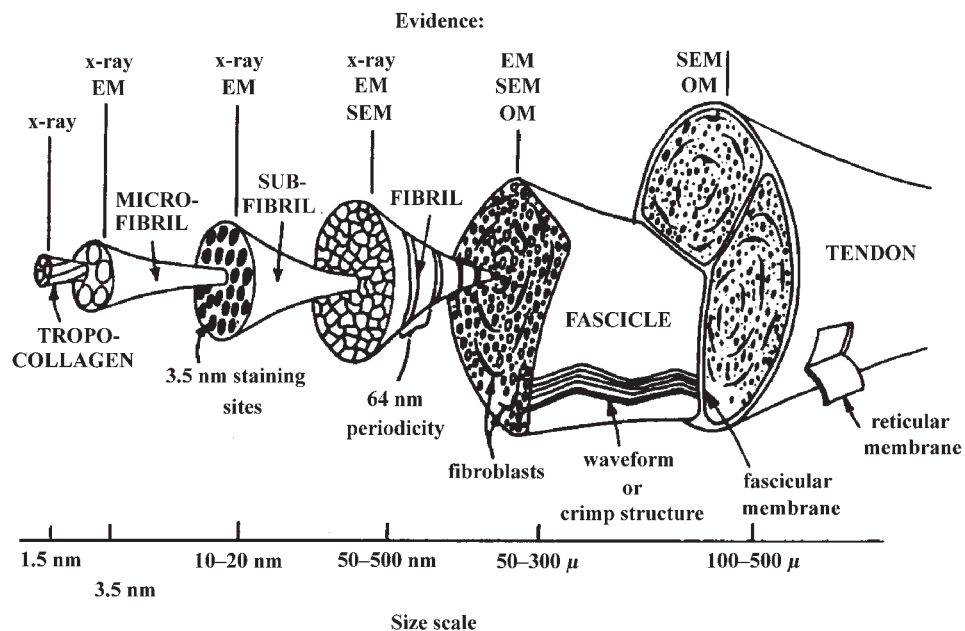


Fig. 1 Hierarchical organization of fibrous structures of biological tissues. (From Ref. [5].)

properties of the high-tenacity fibers are quite insensitive to the physiological environment.

Although biomaterials that are nonreactive to body environment are desirable for some surgical applications, the emergence of tissue engineering has resulted in a trend in surgical implants toward materials that respond to the body environment in a programmable manner. A large number of absorbable polymers have been developed in recent years. These absorbable polymers serve a temporary purpose: Once the wound or fracture heals, the implant, suture, or other structure disappears. Table 2^[8] gives a summary of some of the synthetic absorbable polymers and their applications.

The various fibers can exist in monofilament, multifilament (with or without twist), textured, and staple fiber forms to create a wide variety of surface geometries and extensibilities. Figure 3a illustrates the different types of linear fibrous assemblies available for use. In Fig. 3b, the stress-strain behavior of the various linear fiber assemblies is illustrated.^[9]

The fibers or yarns can be converted into a tremendous array of 2-D and 3-D fabric structures, as shown in Fig. 4.^[10] An implant designer can utilize, for example: 1) highly covered and bulky nonwoven structures; 2) dimensionally stable plain-weave fabrics; 3) weft-and-warp-insertion warp knits (warp-knitted structures with inserted linear yarns integrated into the fabric during fabric formation, along the machine direction (warp) and across the machine direction (weft), respectively); and 4) highly conformable weft knits and braids, as well as net-shaped three-dimensional fabrics, which provide a means to organize fibrous networks similar to the structural features of various body parts.

STRUCTURAL COMPONENTS OF FABRICS

Textile Units

Fabrics are characterized by fiber material (polyester or polylactidglycolic [PLAGA]), fiber geometry (crimp,

Table 1 Properties of candidate materials for prostheses

Fiber	Tenacity (g/denier)	Breaking elongation	Young's modulus (g/denier)
Polytetrafluoride (Teflon)	1	45%	5
Polyester (Dacron 55)	6	30%	30
Polyamide (Nylon 65)	6	45%	20
Aramid (Kevlar 29,49)	22	3-6%	450-1000
Acrylic (Orlon)	2.5	25%	10
Stainless steel (315 L)	3	1%	300
Polypropylene (Olefin)	5	50%	30

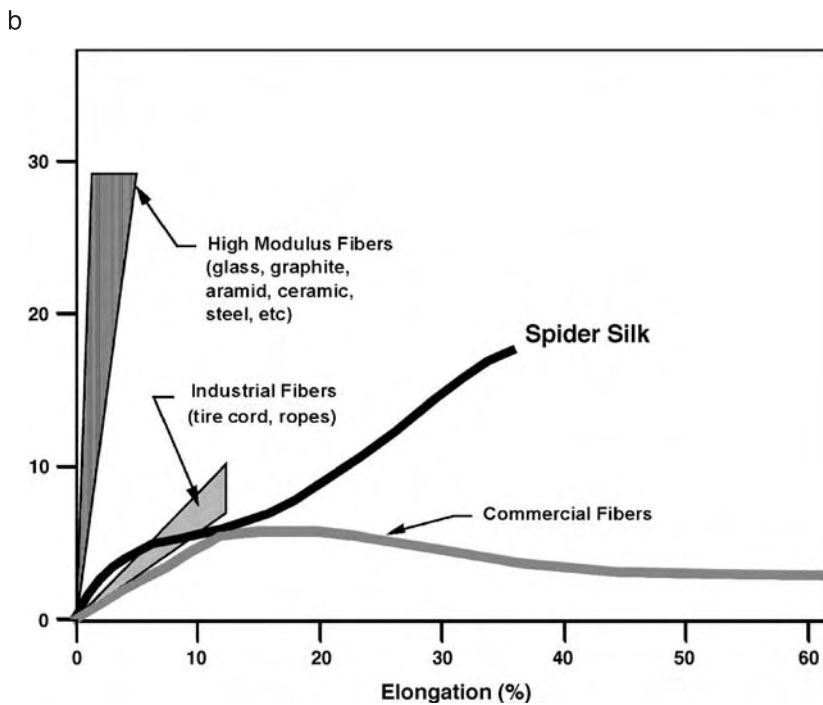
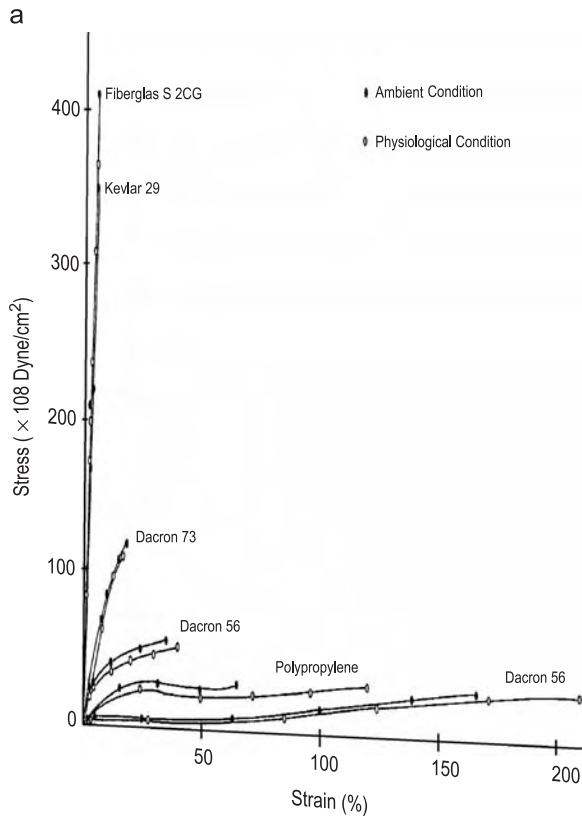


Fig. 2 a. Stress-strain behavior of a wide range of fibers tested in ambient and physiological conditions.^[7] b. Stress-strain properties of various groups of fibers. (From Ref. [10].)

cross-sectional shape, fineness), yarn geometry (monofilament or textured multifilament), fabric construction (woven, knit, braid), and fabric weight (mass per unit area). Similar to soft biological tissues, the compressible nature of fibrous assemblies makes it difficult, if not impossible, to precisely measure the dimension of

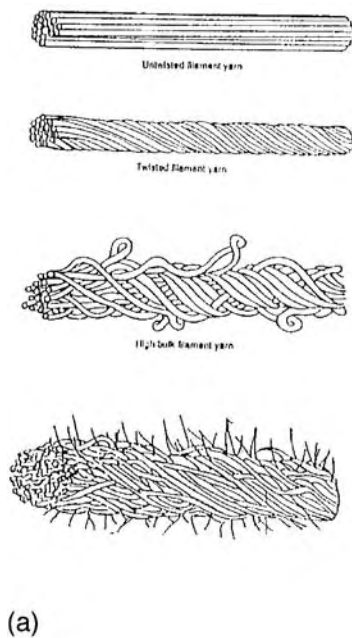
fibrous materials. The cross-sectional area of fibers and linear fiber assemblies are more conveniently expressed in terms of the traditional textile unit of linear density: the denier. One denier is the thickness of a fiber with mass of one gram per 9000 meters. Alternatively, the linear density can be expressed as the SI unit kg/m, or

Table 2 Absorbable Fibers

Polymer	Product
Polylactic acid (PLA)	Suture Maxillofacial surgery Controlled-release device
Polyglycolic acid (PGA) (DEXON)	Suture Bone plate, pins, screws Gauzes, sponge
PLA/PGA copolymer (Vicryl)	Suture
Polycaprolactone and copolymer	Hernia
Polydioaxone (DDS)	Suture
Composite-PLA/PGA, ceramic and carbon fibers	Bone fixation Bone repair
Polyester-amides	Suture
Isomorphic copolyoxalates and polyalkylene oxalates	Suture Controlled-release device Surgical aid
Polyethylene oxide- polyethylene telephthalate copolymers	Artificial artery Tendon

tex. The equivalent cross-sectional area or specific cross-sectional area has the following expression:^[11]

$$A_{sp} = \frac{\text{tex}}{10^5 \times \text{density}} \text{cm}^2 = \frac{\text{denier}}{9 \times 10^5 \times \text{density}} \text{cm}^2 \quad (1)$$



With the specific cross-sectional area expressed in terms of linear density, one can obtain the specific stress and other quantities in mechanics by normalizing of force by A_{sp} (to obtain force/area [MPa or psi]) or by linear density, such as force/linear density (g/denier; g/tex; N/tex). For fabric structures (planar and 3-D assemblies), the normalization of properties can be carried out using the following relationships for specific stress:^[12]

$$\frac{N}{\text{tex}} = \frac{\frac{\text{force (N)}}{\text{width (mm)}}}{\text{fabric areal density (g/m}^2\text{)}} \quad (2)$$

$$1 \text{ N/tex} = 1 \text{ GPa/density (g/cc)}.$$

Fibers

A textile fiber is an entity of matter that is processable in a textile manufacturing process and characterized by slenderness, flexibility, and a high length-to-diameter ratio greater than 1000:1. Beside flexibility, the unique characteristic of textile fibers is a combination of strength, light weight, and toughness. As shown in Fig. 2, the strength of fibers ranges from 1 g/den* to 40 g/den for super-high-strength fibers, with elongation ranging from 1% to several hundred percent for elastomers.

Depending on chemical composition and internal structure, fibers have different stress-strain properties. The stress-strain curves of natural and man-made fibers

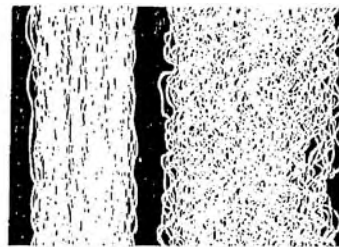


Figure 5.1 Identical stems of nylon filaments showing increase in apparent bulk or volume produced by texturing. (Courtesy of Fibers Division, Monsanto Chemical Co)

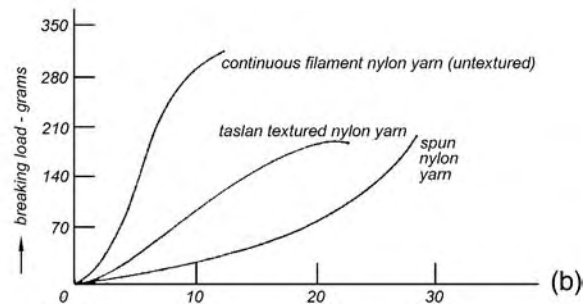


Fig. 3 Linear fiber assemblies and textured yarns. (From Ref. [9].)

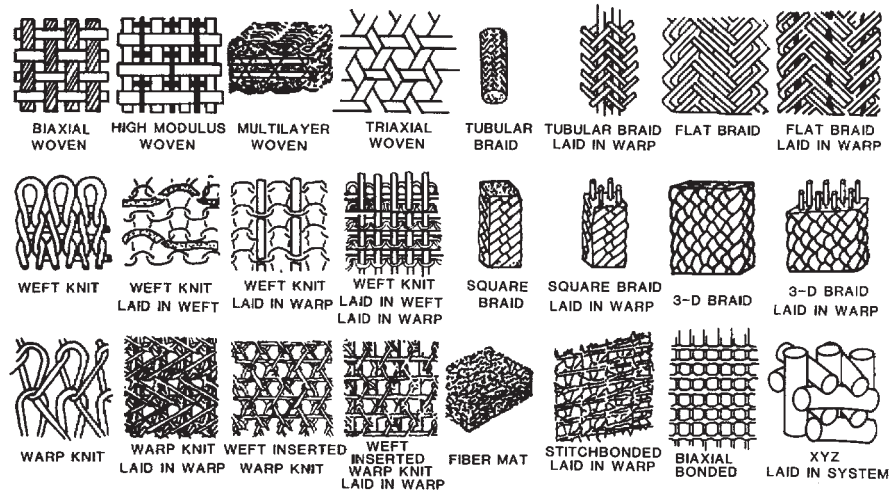


Fig. 4 Fiber architecture for surgical implants. (From Ref. [16].)



shown in Fig. 2 exhibit the wide range of material properties available at our disposal.^[10] It is of interest to observe that the fiber with the best combination of strength and toughness is spider silk.^[13]

Yarns

A yarn is a linear assembly of discrete textile fibers or continuous filaments in a form suitable for knitting, weaving, or intertwining in order to form textile fabrics. Depending upon the nature of the fiber (natural or man-made, long or short, fine or coarse) and the end-use requirements, various yarn formation systems are used. The fiber material properties and the resulting yarn geometry (twist, denier, etc.) determine the performance properties of a yarn and, consequently, dictate the processing methods and the intended applications.

Yarns can be classified according to their form, by the formation system, and by physical properties and

performance characteristics (Fig. 3). The three basic forms of yarns are:

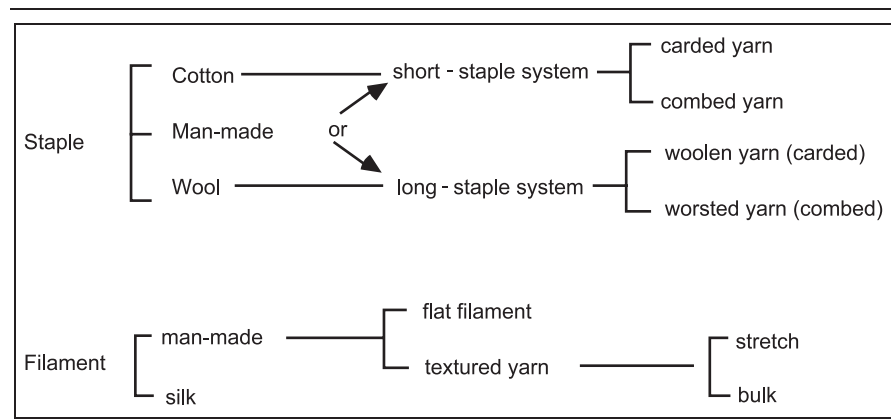
- Monofilament, a solid, single strand of unlimited length.
- Multifilament, a multitude of continuous-filament fibers twisted or textured together.
- Staple (spun), a yarn made from staple (discrete-length) fibers; spun yarn.

Table 3 provides a summary of the various types of yarns produced by different yarn formation systems. For more information on the manufacturing technology of yarns, readers are referred to the texts by Goswami et al.^[14] and Kaswell.^[15]

FABRIC STRUCTURES AND PROPERTIES

There is a large family of fiber architectures available for surgical implants (Fig. 4).^[16] The design and selection of these fiber architectures for surgical

Table 3 Classification of yarns based on yarn formation systems



implants or tissue engineering scaffolds can be carried out on the fiber level as well as on the fabric structure level. These fabrics can be made from yarns (yarn-to-fabric structures) or directly from fibers (fiber-to-fabric structures), resulting in a wide range of dimensional scale, fiber tortuosity, and fabric porosity as characterized by the fiber volume fraction-orientation map described in the section on fabric performance characteristics.

Yarn-to-Fabric Structures

In the family of yarn-to-fabric structures, it can be seen that interlooped structures such as weft knits tend to be more porous, more voluminous, and bulkier and thicker. Interlaced structures such as woven fabrics tend to be less porous, less bulky, and thinner. The performance of these fabrics is determined to a great extent by the linearity and interlacing density of the yarns in the fabrics. For example, the weft knits have high extensibility and are extremely comfortable and compressible. As a result of the openness of the loop geometry to a knitted structure, weft knits are highly permeable. The high linearity of the multiaxial (non-crimped) warp knit structure allows only limited extensibility and low toughness. The permeability of woven and multiaxial warp knit fabrics is lower than that of weft knit fabrics because of the high coverage. In general, based on structural geometry, the family of yarn-to-fabric structures can be classified into two major groups: 2-D planar fabrics and 3-D fabrics (wherein more than two yarns are placed in the through-thickness direction).

Structural Geometry of 2-D Fabrics

Woven fabrics

Fabricated by the interlacing of yarns, woven fabrics can be produced in hundreds of possible combinations. From the in-plane fiber orientation point of view, woven fabrics can be divided into biaxial and triaxial woven structures. Biaxial weaves consist of 0° and 90° yarns interlaced in various repeating patterns or topological unit cells. The three basic weave geometries from which many other patterns evolve are the plane, satin, and twill weaves. A schematic diagram for the various views of these three basic weaves is shown in Fig. 5. These three fabrics are distinguished by the frequency of yarn interlacing and the linearity of the yarn segments. The plane weave has the greatest number of interlacing yarns and the satin weave has the least number of yarns interlacing, with the twill weave somewhere in between. Accordingly, the plane weave has a higher level of structural integrity and

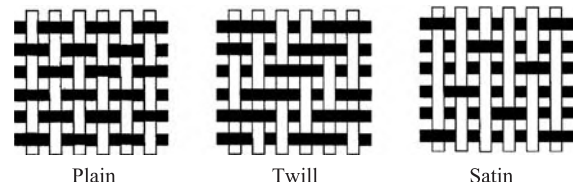


Fig. 5 Three basic weaves.

greater ductility as a result of the crimp geometry produced by yarn interlacing. On the other hand, the satin weave has the highest level of fiber-to-fabric strength and modulus translation efficiency due to the low level of yarn interlacing and high level of yarn linearity. The low level of yarn integration in the satin weave also allows freedom of yarn mobility, which contributes to higher fiber packing density and consequently to higher fiber volume fraction.

Although cane weaving for cane chairs has been known for a long time, a machine-made triaxial weave wasn't available until Norris Dow's development of triaxial machine weaving in the early 1970s.^[17] The unique feature of triaxial weave is the $90 \pm 60^\circ$ hexagonal yarn orientation in one plane, resulting in a high level of in-plane shear resistance. High levels of isotropy and dimensional stability can be achieved with triaxial weave at very low fiber volume fraction. Figure 6 shows a schematic diagram of two triaxial-weave geometries.

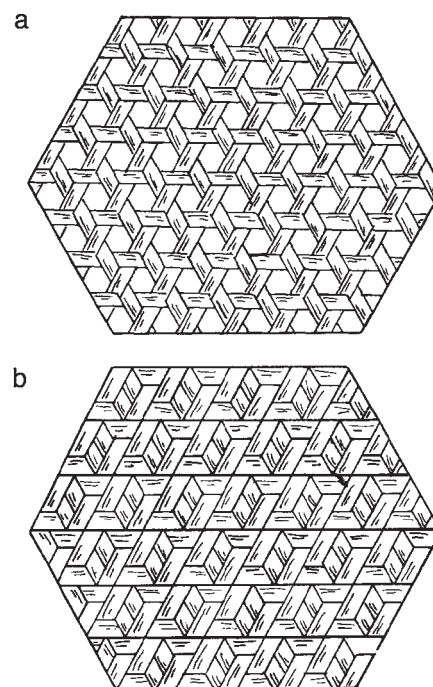


Fig. 6 a. Structural geometry of triaxially woven fabrics (basic weave). b. Structural geometry of triaxially woven fabrics (biplane weave).

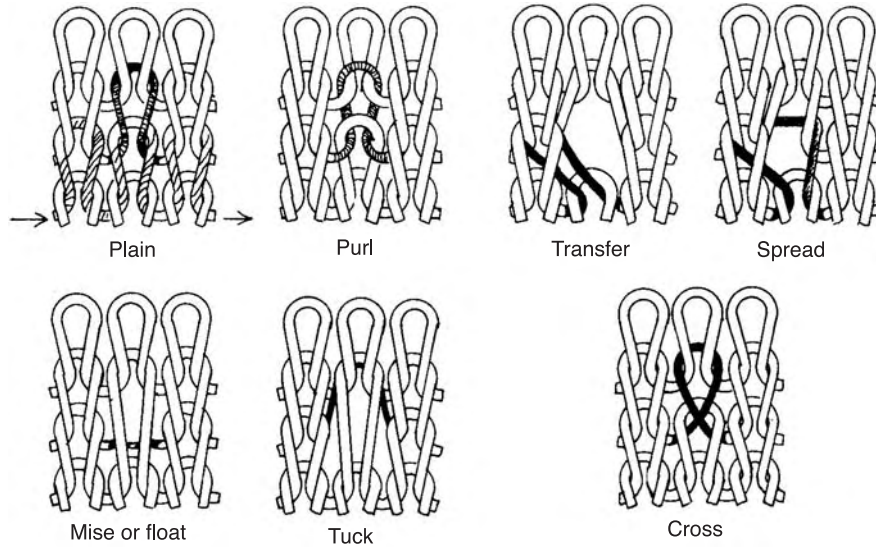


Fig. 7 Weft-knitted structures.

Knitted fabrics

Knitted fabrics are interlooped structures wherein the knitting loops are produced by the introduction of the knitting yarn either in the cross-machine direction (weft knit), as in Fig. 7, or along the machine direction (warp knit), as in Fig. 8, producing a large number of stitch geometries. By controlling the stitch (loop) density, a wide range of pore geometries can be generated. Because of their interlooped nature, the maximum fiber packing density of knitted structures is lower than that of the woven fabrics. The high level of conformability and porosity of weft-knitted fabrics makes them ideal candidates for vascular implants.

Braided fabrics

By intertwining three or more yarn systems together, braided fabrics can be produced in flat or tubular form (Fig. 9). The bias-interlacing nature of the braided fabrics make them highly conformable and shear resistant. To enhance reinforcement in the 0° direction, triaxial braiding can be employed to introduce 0±θ yarns, as shown in Fig. 10. Because of the torque-free nature of braided fabrics, they are used extensively for sutures and are increasingly being considered for orthopaedic applications, such as ligaments and tendons.

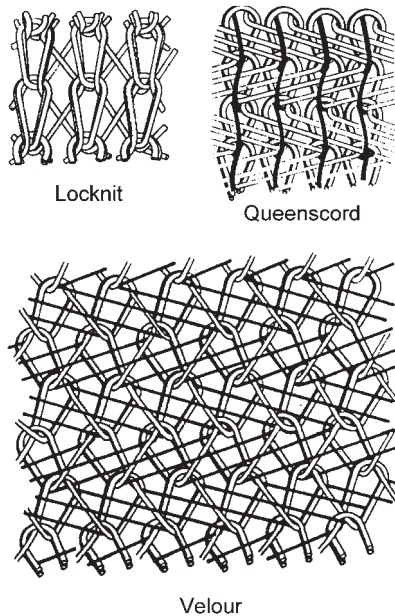


Fig. 8 Warp-knitted structures.

Structural Geometry of 3-D Fabrics

The structural geometry of the 3-D fabrics will be discussed according to the four basic methods of textile

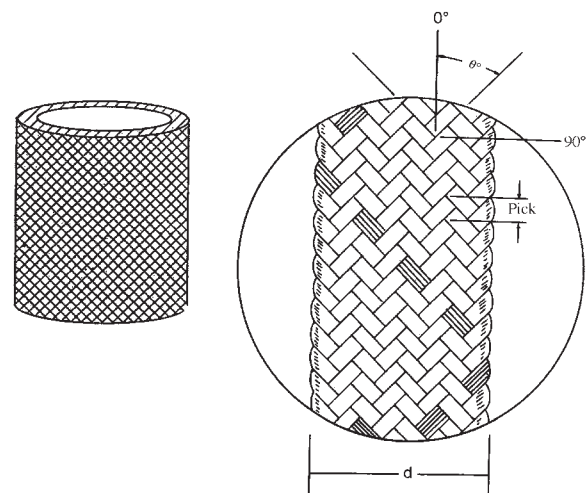


Fig. 9 2-D braided structures.

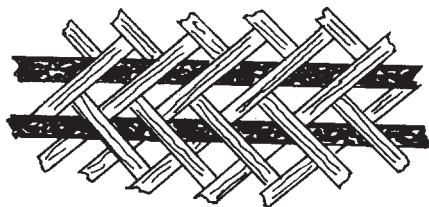


Fig. 10 Triaxially braided structure.

manufacturing: weaving, orthogonal, knitting, and braiding.

Woven 3-D fabrics

Three-dimensional woven fabrics are produced using the multiple-warp weaving method, which has long been used to manufacture double and triple cloth for bags, webbings, and carpets. Fabrics have been successfully woven with as many as 17 layers. Various fiber architectures can be produced by the weaving method, including solid orthogonal panels (Fig. 11a), variable-thickness solid panels (Fig. 11b), and core structures simulating a box-beam (Fig. 11c) or trusslike structure (Fig. 11d). Furthermore, by proper manipulation of the warp yarns, as exemplified by the angle-interlock structure (Fig. 11e), the through-thickness yarns can be organized into a diagonal pattern. To provide in-plane reinforcement in the bias direction, a new process was developed by Dow^[17] as an extension to the triaxial weaving technology, producing multi-layer triaxial fabrics as shown in Fig. 12.

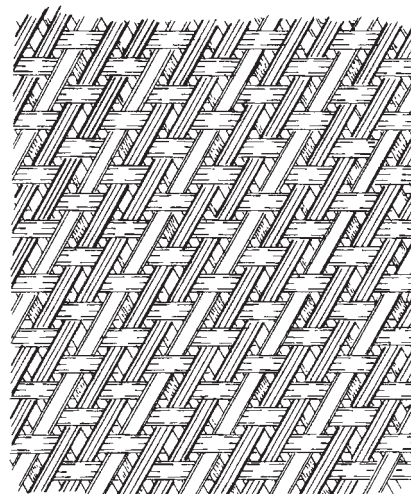


Fig. 12 3-D triaxial weaves (dow).

Orthogonal 3-D fabrics

Pioneered by aerospace companies such as General Electric,^[18] the orthogonal 3-D fabric technology was developed further by Fiber Materials Incorporated.^[19] Progress in automation of the orthogonal 3-D fabric manufacturing process has been made in France by Aerospatiale^[20,21] and SEP,^[22] and in Japan by Fukuta at the Research Institute for Polymers and Textiles.^[23,24] The structural geometries resulting from the various processing techniques are shown in Fig. 13.^[16] Figures. 13a and 13b show the single-bundle

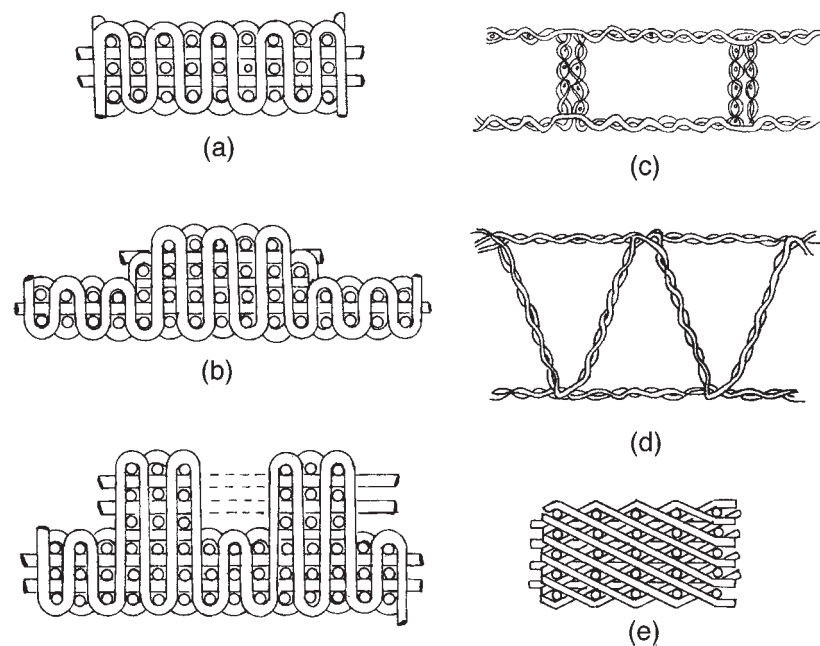


Fig. 11 3-D woven fabrics. (a) Solid shape; (b) variable thickness panel; (c) Box beam; (d) Truss structure; (e) Angle interlock.

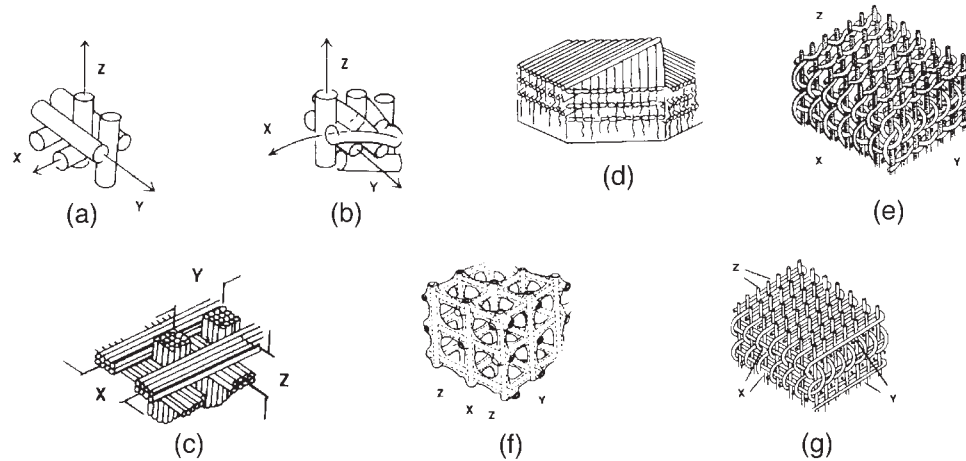


Fig. 13 3-D orthogonal fabrics. (a) Rectangular single bundle shape; (b) Cylindrical single bundle shape; (c) Rectangular multiple bundle shape; (d) Multidirectional 3-D fiber placement; (e) Flexible structure; (f) Open lattice structure; (g) Rigid structure.

XYZ fabrics in a rectangular and cylindrical shape. Figure 13c demonstrates the multiple-yarn bundle possibilities in the various directions. In Fig. 13d the multidirectional fiber placement in the plane of the 3-D structure is shown. Although most of the orthogonal 3-D structures consist of linear yarn in all of the directions, introduction of the planar yarns in a nonlinear manner, as shown in Fig. 13e, f, and g can result in an open lattice, or a flexible and conformable structure. Some of these 3-D lattice structures have been used for producing scaffolds for composite bone implants.^[24]

Knitted 3-D fabrics

Three-dimensional knitted fabrics are produced by either warp knitting or weft knitting. An example of a weft knit is the nearly net-shaped structure knitted under computer control by the Pressure Foot[®] process^[25] (Fig. 14). In a collapsed form, this preform has been used for carbon-to-carbon structures. The unique feature of the weft-knit structures is their conformability and their highly porous nature. When additional reinforcement is needed in the 0° and 90° directions, linear laid-in yarns can be placed inside the knitting loops, as shown in Fig. 14.

Multiaxial warp (MWK) fabric

MWK systems consist of warp (0°), weft (90°), and bias ($\pm\theta$) yarns held together by a chain or tricot stitch through the thickness of the fabric, as illustrated in Fig. 15.^[26] The major distinctions of these fabrics are the linearity of the bias yarns, the number of axes, and the precision of the stitching process. Depending on the number of guidebars available and the yarn

insertion mechanism, the warp-knit fabric can consist of predominantly uniaxial, biaxial, triaxial, or quadraxial yarns. The Mayer Textile Corporation, utilizing a multiaxial magazine weft insertion mechanism, produces the latest commercial nonimpaled MWK fabric. The attractive feature of this system is precision yarn placement, with four layers of linear or nonlinear bias yarns plus a short fiber mat arranged in a wide range of orientations. Furthermore, stitches are formed without piercing through the reinforcement yarns (hence the term, nonimpaled) at a production rate of 100 meters per hour. Another MWK is the impaled LIBA system shown in Fig. 16. This fabric has six layers of linear yarns assembled in various stacking sequences and

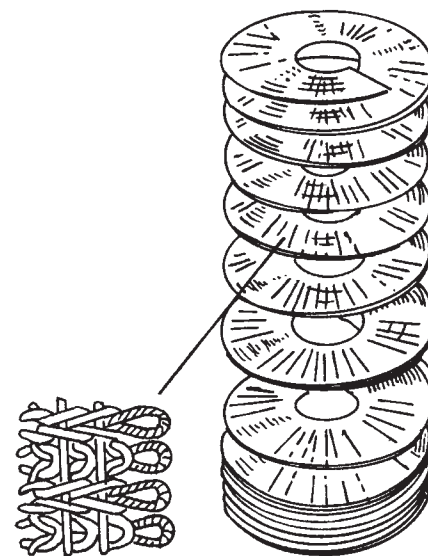


Fig. 14 3-D knitted fabrics (pressure foot[®]).

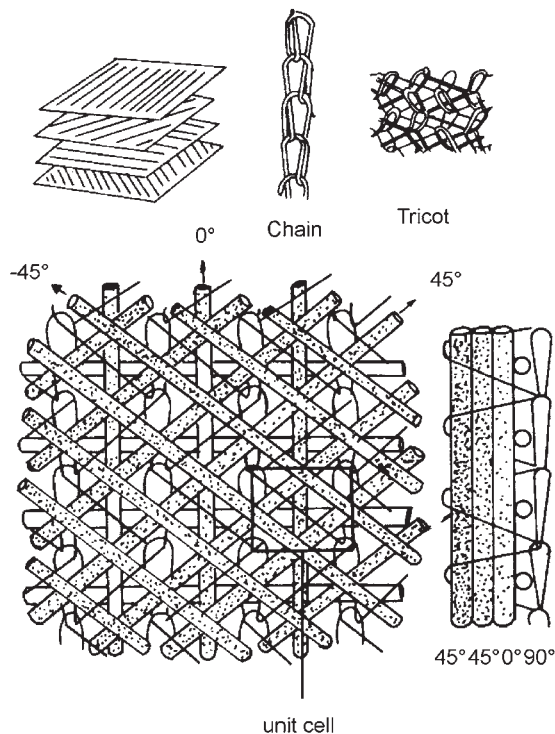


Fig. 15 Multiaxial warp knit (MWK).

stitched together by knitting needles piercing through the yarn layers.

3-D braided fabrics

Three-dimensional braiding technology is an extension of the well-established 2D braiding technology, wherein the fabric is constructed by the intertwining or orthogonal interlacing of two or more yarn systems to form an integral structure. A wide variety of solid, complex structural shapes can be produced in this manner, resulting in 3-D interconnected fibrous structures.^[27] Figure 17 shows two basic loom setups in circular and rectangular configurations. The 3D braids are produced by a number of processes, including the

track and column method (Fig. 18), the 2-step method (Fig. 19), and a variety of displacement braiding techniques. The basic braiding motion includes the alternate X and Y displacement of yarn carriers, followed by a compacting motion. Formation of shapes is accomplished by properly positioning the carriers and joining various rectangular groups through selected carrier movements. The 3-D braided structure has been used for various orthopaedic implants, and more recently for tissue-engineered ligament scaffolds.^[28]

Fiber-to-Fabric Structures

Fiber-to-fabric structures are generally known as nonwoven fabrics. The simplicity of the manufacturing process—by direct conversion of fibers to fabrics, thus bypassing the yarn formation stage—has great appeal to the apparel and medical textile industries because of the high productivity and cost savings. A majority of surgical gowns are made from nonwoven fabrics. One of the obstacles preventing more extensive use of nonwovens for primary garment applications is the lack of drapability/conformability, or the paperlike nature, of nonwoven fabrics. There is an increasing variety of fabrication processes for nonwoven fabrics. Of the two major classes of nonwovens, chemically and mechanically bonded nonwovens, the mechanically bonded systems (such as needled felts) and spunlaced systems tend to be more voluminous, bulkier, thicker, and nonporous. The chemically bonded nonwovens tend to be more paperlike, thin, low in bulk, and less porous. As a result, the mechanically bonded nonwovens are more permeable, more compressible, and have higher extensibility. The toughness of nonwoven felts is a complex interaction of bonding density, fiber tenacity, and fiber orientation. Because of the high strength of the continuous filaments and the strength of the bonds, a spunbonded system features high levels of shear resistance, tear resistance, and toughness. Needled felt made from strong fibers (e.g., Kevlar), on the other hand, can be quite tough and capable of

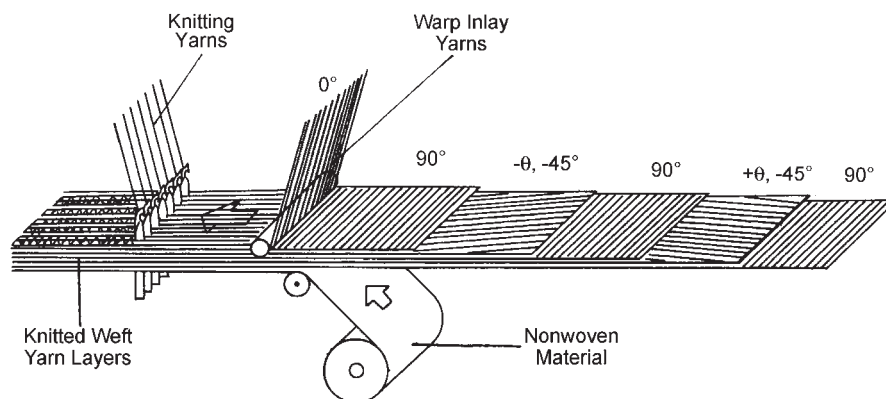


Fig. 16 Multiaxial warp knit (MWK): LIBA system.

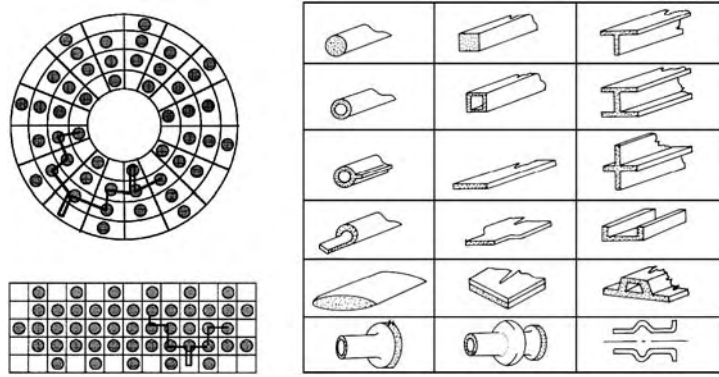


Fig. 17 3-D braiding equipment and net-shaped structures.

arresting fragments from a ballistic projectile when a sufficient level of entanglement is created in the manufacturing process.

Since the 1950s, many of the textile structures illustrated in Fig. 3 have been used for surgical implants, with considerable success. Some of the structures and their respective applications are summarized in Table 4. These fibrous structures are mostly from nonabsorbable polymers consisting of fibers larger than 10 μm in diameter, although some microfibers are increasingly being used. Unfortunately, these fibers are developed primarily for clothing rather than for medical applications. Recently, with the rediscovery of the electrostatic spinning process, a new family of nonwovens and yarns consisting of ultrafine fibrils have been fabricated, with fiber diameters ranging from a few nanometers to a few micrometers. The extension of the basic building block of a fabric to the nanoscale has created new design options for numerous medical fabrics. One method that is capable of direct formation of nanofiber-based yarns and fabrics is the electrospinning process. Electrospinning technology has been detailed in numerous publications,^[29] and a brief description is provided here.

Electrospun yarns and fabrics

The electrostatic spinning (electrospinning) process is a simple, nonmechanical process developed over sixty years ago.^[30] Because of its simplicity and its capability of generating nanoscale fibers from polymer solutions and melts under ambient conditions, there has been a revival of interest in the electrospinning process to create structures that have ultrahigh surface

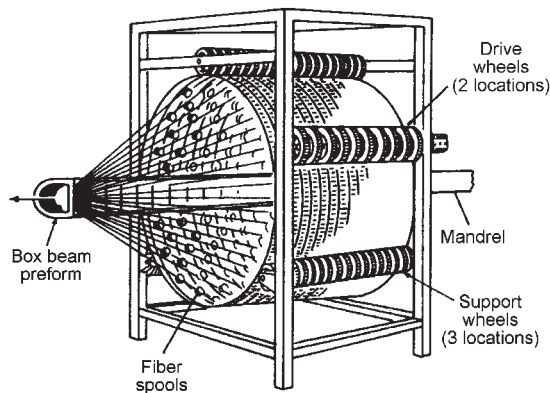


Fig. 18 Track and column 3-D braiding.

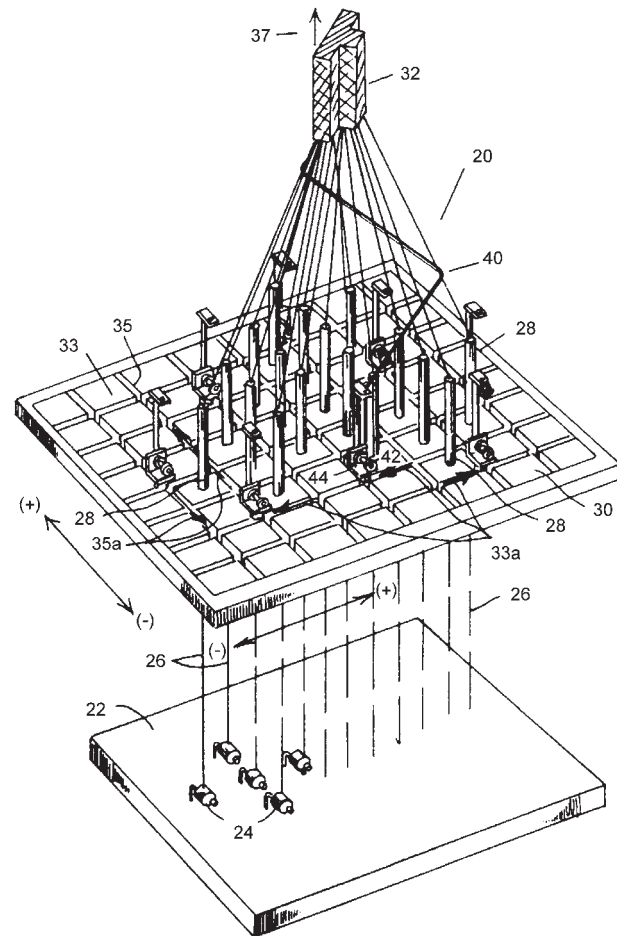


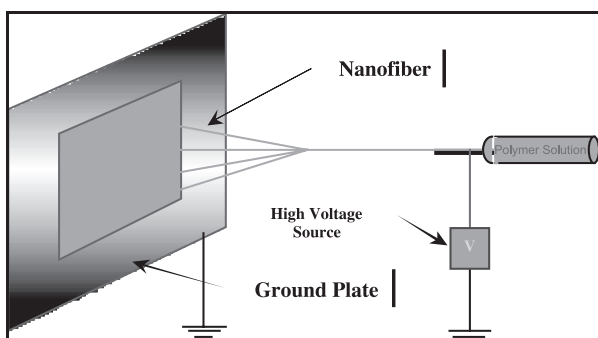
Fig. 19 Two-step 3-D braiding method.



Table 4 Textile structures for surgical implants

Application	Material	Structure	
		Yarn	Fabric
Arteries	Polyester	Textured	Weft/warp knit
	Dacron 56		Straight tube bifurcation
	Teflon		Plain-woven
			Straight tube, nonwoven
Tendon	Polyester	Low-twist filament	Plain-woven narrow tape coated with silicon rubber
	Dacron 56		
	Kevlar		
Hernia repair	Polypropylene	Monofilament	Tricot jersey knit
Esophagus	Regenerated collagen	Multifilament	Plain weave
Heart valve	Polyester	Multifilament	Plain weave
	Dacron 56		Knit
Patches	Polyester	Textured	Knitted velour
	Dacron 56		
Sutures	Polyester	Monofilament	Braid
	Nylon	Multifilament	Woven tapes
	Regenerated collagen		
	Silk		
Ligament	Polyester	Multifilament	Braid
	Teflon		
	Polyethylene		
Bone and joints	Carbon in thermoset or thermoplastic matrix	Multifilament	Woven
			Braid

areas.^[31–33] In the electrospinning process, as illustrated in Fig. 20, an electric field is generated between an oppositely charged polymer fluid and a collection screen. A polymer solution is added to a glass syringe that has a capillary tip, and an electrode is placed in the solution, with another connection made to a

**Fig. 20** The electrospinning process.

ground-fiber-collecting metal screen. As the electrical potential is increased, the charged polymer solution is attracted to the screen. When the voltage reaches a critical value, the electric field overcomes the surface tension of the polymer cone formed on the capillary tip of the syringe and a jet of ultrafine fibers with diameters ranging from 2–500 nm is produced. As the charged polymer jets are splayed, the solvent in the fiber quickly evaporates and the fibers are accumulated randomly, or laid parallel on the surface of the collection screen. This results in a porous, nonwoven film, or a yarn composed of nano- to micron-scale fibers. Varying electric field strength, polymer solution concentration, and duration of electrospinning can control fiber diameter and film thickness. In addition, coelectrospinning can be carried out by mixing one or more kinds of functional nanoparticles, nanoplatelets, or nanotubes during solution preparation, resulting in ultrafine nanocomposite fibers that have a multitude of electronic and structural functions.^[34]

FABRIC PERFORMANCE CHARACTERISTICS

Fabric performance characteristics are a result of the interaction between fiber (material properties), yarn and fabric geometry, and finishing treatment. As outlined earlier, textile structures in fabric form can be characterized in terms of geometric and performance properties. The performance maps provide an overview of the range of behavior of various fabrics as a function of four geometric parameters and four performance parameters.

The geometric parameters include:

1. Porosity: the amount of open space in a unit volume of the fabric. As the fiber diameter and yarn diameter increase, the structure tends to be porous. The porosity of a fabric is inversely proportional to the areal coverage, or cover factor, of a fabric. A porous fabric tends to be lighter and more permeable.
2. Surface texture: the surface geometry of a fabric, characterized by the smoothness of the surface, which in turn is governed by fiber and yarn diameter. Modular fiber or yarn lengths are the geometric repeating units of the fabric.
3. Voluminosity: a reflection of the bulkiness of a fabric for a given areal density (mass per unit area). A fabric tends to be more voluminous if the fiber and yarn diameter is larger and the freedom of fiber mobility in the geometric repeating unit is high. Voluminosity is directly related to fiber thickness, in that a voluminous fabric tends to be thick.
4. Thickness of the fabric: similar to voluminosity and related to fiber and yarn diameter. The larger the fiber and yarn diameter, the thicker and bulkier is the fabric.

The performance parameters include:

1. Permeability: the ease of air or liquid flow through a fabric. The permeability of a fabric is higher when the porosity of the fabric is high. Porosity and fiber volume fraction (1-porosity) are related to packing efficiency, which is influenced by fiber diameter and fiber cross-sectional geometry. Permeability is a strong function of fiber or yarn diameter for a given fiber architecture (fiber orientation).
2. Compressibility: the ability of a fabric to resist transverse (through the thickness) compression. A voluminous fabric tends to be more compressible. On the other hand, compressibility decreases with the stiffness of fiber and yarn, which is significantly influenced by fiber diameter. As fiber diameter increases, the bending stiffness and longitudinal compressive stiffness of the fiber increases geometrically.
3. Extensibility: a measure of the ability of a fabric to stretch and conform. Fabric extensibility is affected by fabric geometry and inherent fiber-bending elongation. A yarn that consists of finer fibers tends to have a higher potential for fabric extensibility.
4. Toughness: a measure of the durability of a fabric. As reflected in the areas under the stress-strain curve of a fabric, a high-strength fabric with high elongation at break usually produces high toughness. Fabrics having high compliance and extensibility are usually tougher.

Summaries of the geometric and performance properties of the yarn-to-fabric structures, and the fiber-to-fabric structures in the form of performance maps, are shown in Figs. 21–24. These performance

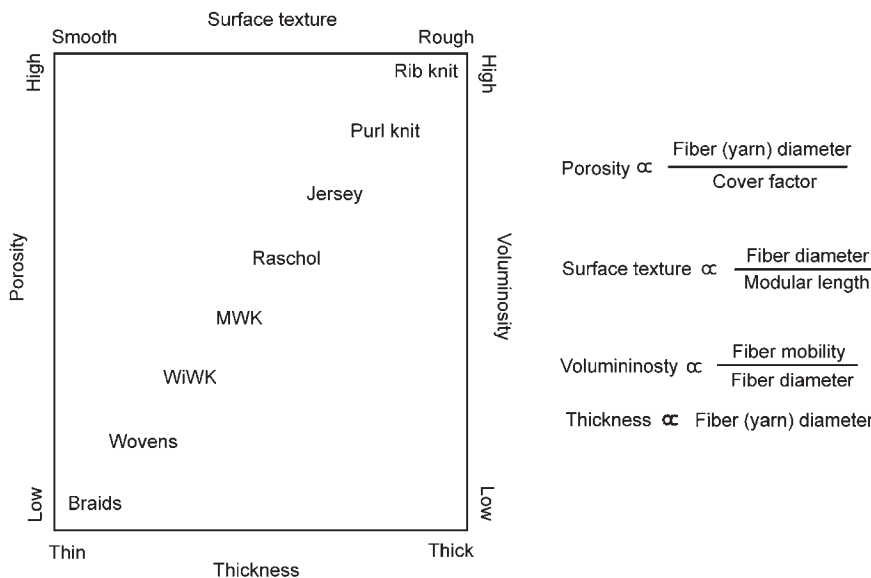


Fig. 21 Geometric properties of yarn-to-fabric structures.

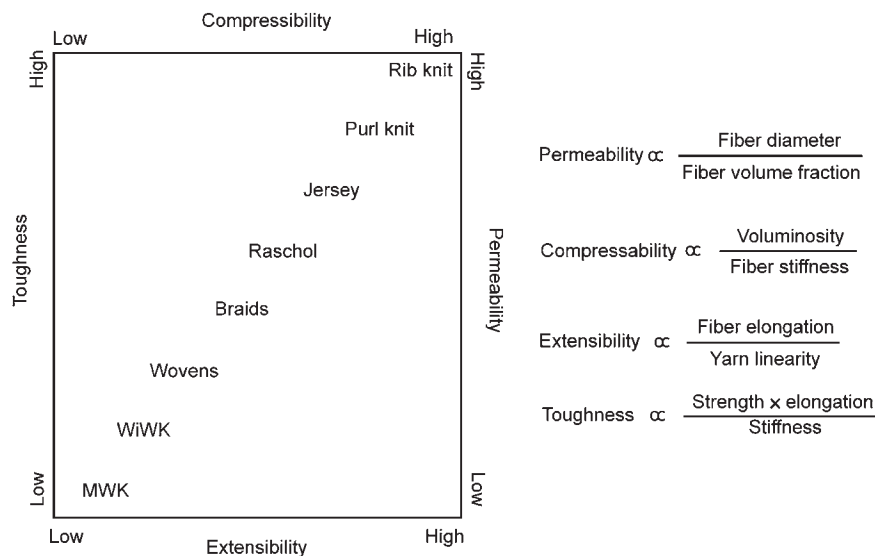


Fig. 22 Performance properties of yarn-to-fabric structures.

maps show that the geometric parameters play an important role in the structural and physical properties of fabrics. One key geometric factor is the fineness of the component fiber.

FABRIC SCAFFOLDS FOR TISSUE ENGINEERING

In addition to traditional applications for surgical implant development, fabric structures are excellent candidates for the emerging field of tissue engineering, or regenerative engineering. As articulated succinctly by Vacanti and Mikos,^[36] the key challenges in tissue engineering are: 1) the synthesis of new cell adhesion-specific materials; and 2) development of fabrication methods for the manufacturing of reproducible

three-dimensional synthetic or natural biodegradable polymer scaffolds, with tailored properties. These properties include dimensional porosity, pore size distribution, and connectivity; mechanical properties for load-bearing applications; and rate of degradation. In this article, we introduce the large family of three-dimensional fiber architectures, as well as the range of fiber-size scales available for tissue engineering.

The geometry and size scale effect and their potential for tissue-engineering scaffolds are demonstrated through various fibrous structures.^[37] Specifically, experiments were carried out using osteoblasts isolated from neonatal rat calvaria and grown to confluence in Ham's F-12 medium (GIBCO), supplemented with 12% Sigma fetal bovine on polylacticglycoic acid (PLAGA) sintered spheres, 3-D braided 20 mm filament bundles, and nanofibrils.

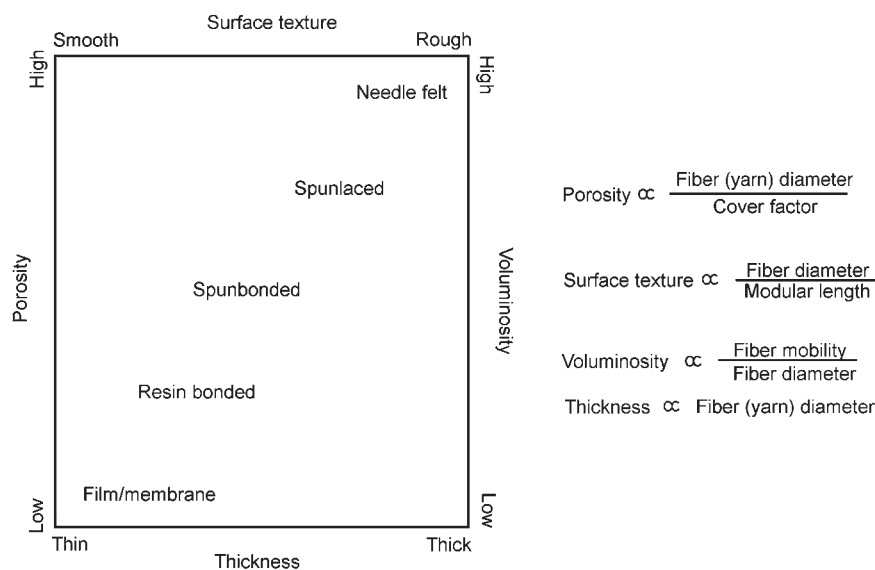


Fig. 23 Geometric properties of fiber-to-fabric structures (nonwoven fabrics).

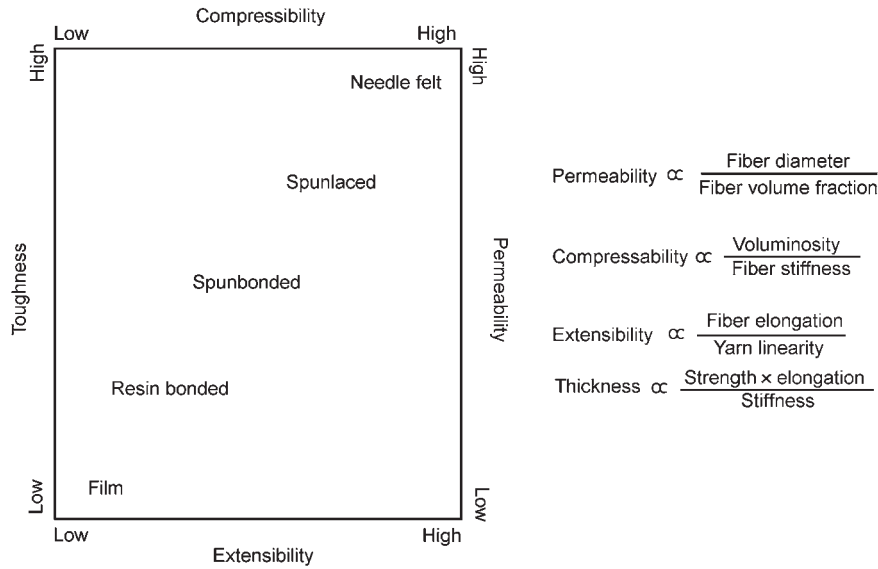


Fig. 24 Performance properties of fiber-to-fabric structures (nonwoven fabrics).

Four matrices were fabricated for the cell culture experiments:

1. 150–300 mm PLAGA sintered spheres.
2. Unidirectional bundles of 20 mm filaments.
3. 3-D braided structure consisting of 20 bundles of 20 mm filaments.
4. Nonwoven structure consisting of nanofibrils.

The porosity of the microsphere structure, the nanofibrous specimens, and the 3-D braided structures were characterized by mercury porosimetry measurement to determine their pore size distributions. Cells were seeded on the UV sterilized PLAGA matrices at a density of 100,000 cells per cm^2 . The osteoblasts were cultured on the matrices for durations ranging from one day to 21 days. These specimens were prepared according to established procedures by fixing in gluteraldehyde and dehydrating through a series of ethanol dilutions. The seeded cells were labeled with [^3H]-

thymidine, and the thymidine uptake was measured at 1, 3, 7, 10, and 11-day intervals.

Pore Size Distribution

The pore size and distribution of the various specimens were characterized using a Micrometrics Autopore III porosimeter. As shown in Fig. 25, a bimodal distribution of pores was observed for the 3-D braided structures, showing the existence of large interstitial pores on the order of 250 μm and interfiber pores having a diameter on the order of 30 μm . At the other extreme, as shown in Fig. 26, the nanofiber structures show a predominant concentration of pores with an average pore size of 14 μm . The pore surface of the nanofibrous structures was 0.823 m^2/g , an order of magnitude greater than that in the 3-D braided structure at 0.0045 m^2/g . The sintered spheres show a single-mode pore diameter distribution over the range

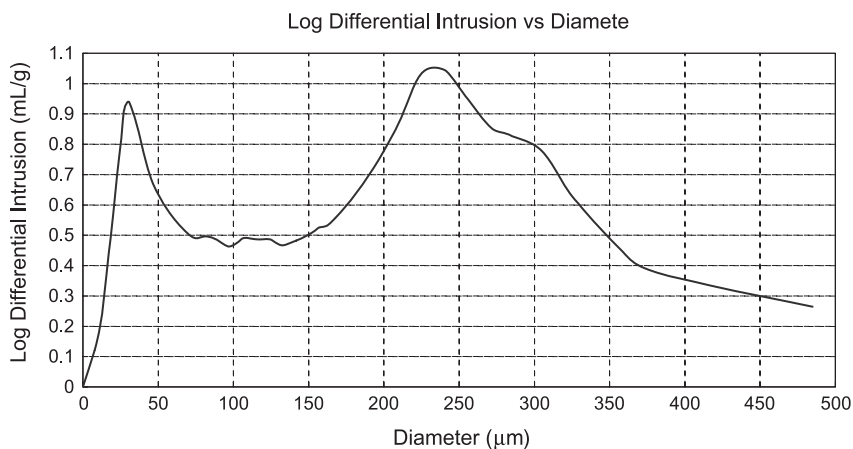


Fig. 25 Pore size distribution of braided structure.

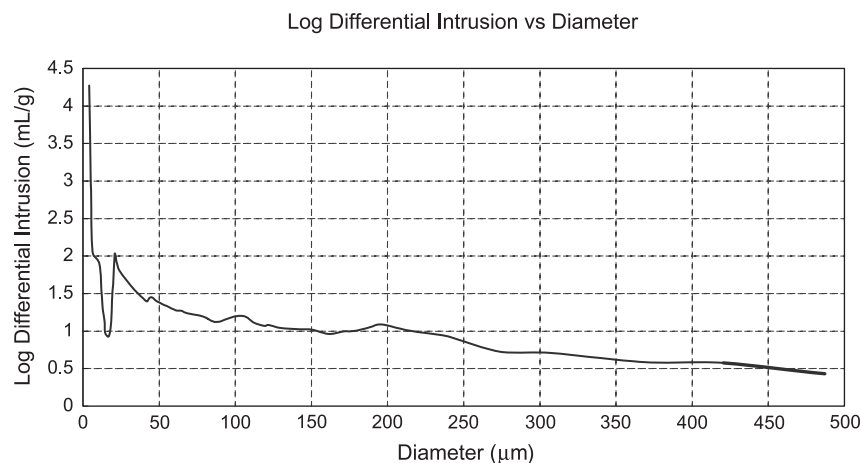


Fig. 26 Pore size distribution of nanofibrous nonwoven structure.

of 200–600 μm, depending on the sphere diameter, as shown in Fig. 27.

Cell Proliferation

Cell proliferation was quantitatively measured by uptake of radio-labeled methyl [^3H] thymidine (Sigma, St. Louis MO). During the last 6 hours of cell incubation, 10 μCi of [^3H] thymidine was added to each sample well. Incubation was terminated by removal of the radioactive media. Cells were fixed to the matrices by soaking in methanol for 10 minutes. Samples were then rinsed twice with Hank's Balanced Salt Solution (HBSS). Uptake of [^3H] thymidine into cellular DNA was assessed by dissolution of the material in 2N perchloric acid. Uptake was measured using an EC&G-Wallac 1027 Beta scintillation counter.

The cell proliferation, as shown in Fig. 28, is expressed in terms of the amount of [^3H] thymidine uptake as a function of time. It can be seen that there is a consistent increase of cell population with time. The nanofibrous structure demonstrated the most proliferate cell growth whereas the tightly woven 3-D

braided structure showed the least proliferation. On the other hand, the cell growth, in fused microsphere structures is between that of the 3-D braid and the nanofibrous structures, showing a surprising drop after the 10th day.

Cell-Scaffold Architecture Interaction

Experimental results show that cell growth pattern is related to the relative dimensions between the components in the scaffold and the cells. Scanning electron microscopy pictures were taken from the cell-matrix systems, which were prepared in a series of Freon 113 dilutions (25%, 50%, 75%, and 100%). The SEM samples were sputter-coated with gold (Denton Desk-1 Sputter Coater). An Amray 3000 SEM using an accelerating voltage of 20 kV was employed to take the SEM photographs shown in Fig. 29. It can be seen that in responding to the large spheres, wherein the cells are more than $10\times$ smaller than the spheres (Fig. 29a), the cells tend to spread over the surface of the sphere before connecting to the adjacent spheres and eventually forming an interconnected cellular network. In the case

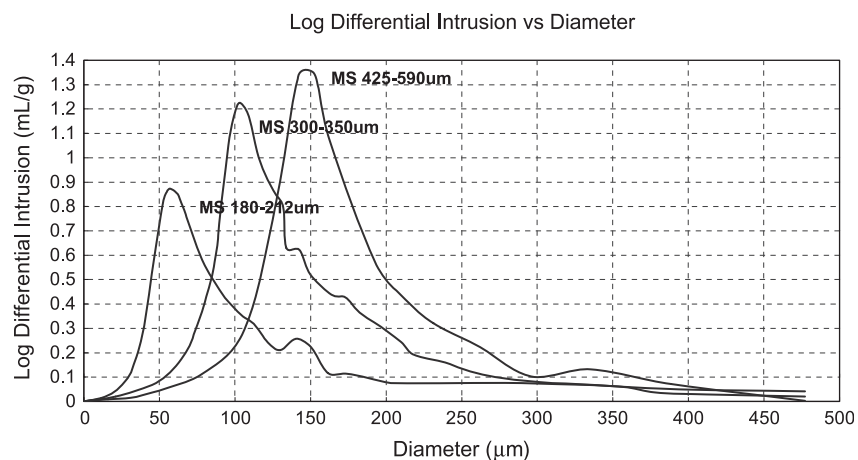


Fig. 27 Pore size distribution of sintered spheres.

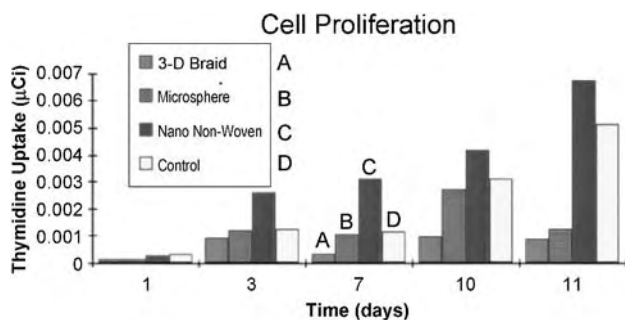


Fig. 28 Cell proliferation of fibrous scaffolding.

of 20 μm filaments in unidirectional bundles and 3-D braid (Fig. 29b), wherein the cells are about the same order of magnitude in dimension, the cell–matrix interactions appear to be similar. The cells tend to slide off the matrix at the moment of seeding. For those cells remain on the surface of the substrates they tend to grow around the filaments and bridge onto the adjacent filaments along the length of the fibers. The most intensive cell deposition was seen in the nanofibril nonwoven structure (Fig. 29c), wherein the cells are more than an order of magnitude larger than the individual fibrils. Extensive cell spreading was observed along the length of the fibrils and through the thickness. This suggests the feasibility of organizational cell growth.

This study showed that fiber-based scaffolds are capable of supporting cell attachment and proliferation. The use of fiber technology in tissue engineering holds several advantages over a number of nonfibrous 3-D structures. Mainly, the ability to impart high levels of structural organization to the matrix may allow for precise control of matrix structure. The 3-D braided and nonwoven matrices demonstrate a range of 3-D fiber architectures that can be designed and produced. The braided matrix consisted of highly organized PLAGA yarns woven into a 3-D structure. Although the nonwoven matrix was the result of randomly oriented microfibers, the structure was highly uniform.

Both the four-step, 3-D braiding technique and the electrospinning process are promising fabrication methods showing high levels of versatility. The use of hierarchical fiber architecture-based design enables the tailoring of a broad spectrum of substrates with a wide range of mechanical properties, and the creation of structures capable of supporting cell proliferation, as well as the microdiffusion of nutrients throughout the structure.

The ability to manufacture a variety of different matrices/scaffolds and the ability to maintain precise control over matrix fabrication are extremely important factors in the design of a tissue-engineered scaffold. Based on the results of the study reported in this article and in an increasing body of work,^[28,38] it seems that fiber-based scaffolds show promise as replacements for orthopaedic tissues, as well as for a broad range of biological tissues such as skin, arteries, muscles, and others.

SUMMARY AND CONCLUSIONS

This article provides an overview of the structural components of fabrics and a description of the various 2-D and 3-D fabrics, along with a gallery of the various fabric structures. The potential application of fabrics for scaffolds in the emerging field of tissue engineering is also reviewed in light of the development of nanofibers and nanofiber-based fabric structures.

Fibrous materials in fabric forms are an important structural backbone for surgical implants and tissue-engineering scaffolds. Analogous to natural tissues and organs, the fabric structures are organized into hierarchical fiber architecture, thus forming a scaffold for cell adhesion, migration, and proliferation.

Successful development of surgical implants and medical textiles requires a thorough knowledge of the structure and properties of fabrics. In addition to their use as surgical implants for the replacement or regeneration of body parts, fabrics are also important

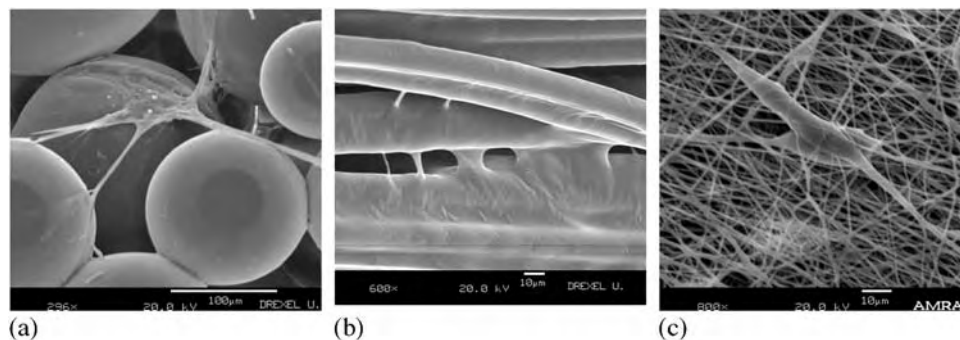


Fig. 29 Cell-fiber architecture interaction: a, sphere; b, braid; c, nonwoven.

substrates for drug delivery and a wide range of medical textiles, including numerous hospital supplies ranging from nurses' uniforms to surgical gowns, drapes, and masks.

ARTICLES OF FURTHER INTEREST

Biocompatibility Testing; Collagen; Electrospinning; Poly(Glycolic Acid); Poly(lactic acid)s; Polymers; Tissue Engineering Scaffolds; Vascular Grafts

REFERENCES

- Williams, D.F.; Roaf, R. *Implants in Surgery*; W.B. Saunders Company: Philadelphia, 1973.
- Lee, H.; Neville, K. *Handbook of Biomedical Plastics*; Pasadena Technology Press, 1971.
- Nerem, R.; Sambanis, A. Tissue engineering: From biology to biological substitutes. *Tissue Eng.* **1995**, *1* (1).
- Principles of Tissue Engineering*; Lanza, R.P., Langer, R., Chick, W.L., Eds.; R.G. Landes Company and Academic Press, Inc., 1997.
- Baer, E.; Kastelic, J.; Galeski, A.J. *Connect. Tissue Res.* **1978**, *6*, 11.
- Ko, F.K. Fiber Reinforced Composites in Medical Applications. Proceedings of the Fiber Producer Conference, Clemson University, Oct. 24, 1984; 30–33.
- Winston, S.M. Viscoelastic Behavior of Fibers under Physiological Conditions. M.S. Thesis; Cornell University, 1978.
- Ozer, Z. Synthetic Absorbable Polymers for Implantation. Textile Engineering Seminar, Philadelphia College of Textiles and Science, October, 1982.
- Piller, B. *Bulk Yarns—Production, Processing and Applications*; SNTL: Prague, 1973.
- Ko, F.K. Medical applications for textile structures. *Text. Asia* **April, 1997**, *4*.
- Mechanics of Flexible Fiber Assemblies*; Hearle, J.W.S., Thwaites, J.J., Amirbayyat, J., Eds.; NATO ASI, Ser. E Appl. Sci., Sijthoff & Noordhoff 1980; Vol. 38.
- Hearle, J.W.S.; Grosberg, P.; Backer, S. *Structural Mechanics of Fibers, Yarns, and Fabrics*; Wiley-Interscience, 1969; Vol. 1.
- Ko, F.K. Engineering Properties of Spider Silk Fibers. In *Natural Fibers, Plastics, and Composites*; Wallenberger, F.T., Weston, N., Eds.; Kluwer Academic, 2000.
- Goswami, B.G.; Martindale, J.G.; Scardino, F.L. *Textile Yarns, Technology, Structure and Applications*; John Wiley and Sons: New York, 1977; 273–337.
- Kaswell, E.R. *Wellington-Sears Handbook of Industrial Textiles*; Wellington-Sears: New York, 1963.
- Ko, F.K. Preform architecture for ceramic matrix composites. *Ceram. Bull.* **1989**, *68* (2).
- Dow, N.F.; Tranfield, G. Preliminary investigations of feasibility of weaving triaxial fabrics (Dow weave). *Tex. Res. J.* **1970**, 986–998.
- Stover, E.R.; Mark, W.C.; Marfowitz, I.; Mueller, W. Preparation of an Omniweave-Reinforced Carbon-Carbon Cylinder as a Candidate for Evaluation in the Advanced Heat Shield Screening Program. AFML TR-70283, Mar. 1979.
- Herrick, J.W. Multidimensional Advanced Composites for Improved Impact Resistance. 10th National SAMPE Technical Conference, Oct. 17–19, 1977.
- Pastenbaugh, J. Aerospace Technology. 3rd Textile Structural Composites Symposium, Philadelphia, PA, June 1–2, 1988.
- Bruno, P.S.; Keith, D.O.; Vicario, A.A., Jr. Automatically woven three dimensional composite structures. *SAMPE Q.* **July 1986**, *17* (4), 10–16.
- O'Shea, J. Autoweave : A Unique Automated 3-D Weaving Technology. 3rd Textile Structural Composites Symposium, Philadelphia, PA, June 1–2, 1988.
- Fukuta, K.; Aoki, E. 3-D Fabrics for Structural Composites. 15th Textile Research Symposium, 1986.
- Fukuta, K.; Aoki, E.; Onooka, R.; Magatsuka, Y. Application of latticed structural composite materials with three dimensional fabrics to artificial bones. *Bull. Res. Inst. Polym. Tex.* **1982-2**, *131*, 151–159.
- Williams, D.J. New knitting methods offer continuous structures. *Adv. Compos. Eng.* Summer **1978**, 12–13.
- Ko, F.K.; Pastore, C.M.; Yang, J.M.; Chou, T.W. Structure and Properties of Multidirectional Warp Knit Fabric Reinforced Composites. In *Composites '86: Recent Advances in Japan and the United States*, Proceedings, Japan-U.S. CCM-III, Tokyo, Kawata, K., Umekawa, S., Kobayashi, A., Eds.; 1986; 21–28.
- Ko, F.K.; Pastore, C.; Head, A. *Atkins and Pearce Handbook of Industrial Braids*; Drexel University, 1989.
- Cooper, J.A. Design, Optimization and In Vivo Evaluation of a Tissue Engineered Anterior Cruciate Ligament Replacement. In *Ph.D. Thesis*; Drexel University, 2002.
- Lyons, J.; Ko, F. Nanofiber Technology. In *Encyclopedia of Nanoscience and Nanotechnology*; Nalwa, H.S., Ed.; American Scientific, 2004; Vol. 1–10.
- Formhals, A. US Patent 1,975,504, Oct. 2, 1934.
- Reneker, D.H.; Chun, I. Nanometre diameter fibres of polymer, produced by electrospinning. *Nanotechnology* **1996**, *7*, 216–223.
- Norris, I.D.; Shaker, M.M.; Ko, F.K.; MacDiarmid, A.G. Electrostatic Fabrication of Ultrafine Conducting Fibers: Polyaniline/Polyethylene Oxide Blends. In *Synthetic Metals*; Elsevier, 2001.
- MacDiarmid, A.G.; Jones, We., Jr.; Norris, I.D.; Gao, J.; Johnson, A.T., Jr.; Pinto, N.J.; Hone, J.; Han, B.; Ko, F.K.; Okusaki, H.; Llanguno, M. Electrostatically-Generated Nanofibers of Electronic Polymers. In *Synthetic Metals 10*; Elsevier, 2002.
- Ko, F.; Gogotsi, Y.; Ali, A.; Naguib, N.; Ye, H.; yang, G.; Li, C.; Willis, P. Electrospinning of continuous carbon nanotube-filled nanofiber yarns. *Adv. Mater.* **July 17, 2003**, *15* (14).
- Li, W.J.; Laurencin, C.T.; Caterson, E.J.; Tuan, R.S.; Ko, F.K. Electrospun nanofibrous structure: A novel

- scaffold for bioengineering. *J. Biomed. Mater. Res.* **March 25, 2002**, *58*, 613–621. Wiley Interscience
36. Sukigara, S.; Gandhi, M.; Ayutsede, J.; Micklus, M.; Ko, F. Regeneration of *Bombyx mori* silk by electrospinning- Part 1: Processing parameters and geometric properties. *Polymer* **2003**, *44*, 5721–5727.
37. Vacanti, C.A.; Mikos, A.G. Letter from the Editors. *Tissue Eng.* **1995**, *1* (1).
38. In Ko, F.K.; Laurencin, C.T.; Borden, M.D.; Reneker, D. In *The Dynamics of Cell-Fiber Architecture Interaction*, Proceedings, Annual Meeting, Biomaterials Research Society: San Diego, April 1998.
39. Fertala, A.; Han, W.B.; Ko, F.K. Mapping critical sites in collagen II for rational design of gene-engineered proteins for cell-supporting materials. *J. Biomed. Mater. Res.* **2001**, *57*, 48–51.
40. Boland, E.D.; Simpson, D.G.; Wnek, G.E.; Bowlin, G.L. Electrospinning of biopolymers for tissue engineering scaffolds. *Polym. Prepr.* **2003**, *44* (2), 92.

Fiber Optic Biosensors

Ilko K. Ilev

U.S. Food and Drug Administration, Center for Devices and Radiological Health, Office of Science and Engineering Laboratories, Division of Physics, Rockville, Maryland, U.S.A.

Israel Gannot

National Institute of Health, NICHD, Biomedical Stochastic Physics, Bethesda, Maryland, U.S.A., and Department of Biomedical Engineering, Faculty of Engineering, Tel-Aviv University, Tel-Aviv, Israel

INTRODUCTION

The field of fiber optic biosensors (FOBSs) is an intensively active area of modern biophotonics technologies because these sensors are potential alternatives to conventional medical methods for diagnostics and monitoring of diseases. Fiber optic biosensors offer a minimally invasive, highly sensitive, and accurate means for measurement and control of various physical and biomedical quantities. Biosensor techniques and devices utilizing minimally invasive FOBSs technology are rapidly finding their way into the mainstream for early disease diagnostics, laser therapy, and in situ optical biosensing and bioimaging.

As a key component in FOBSs systems, biomedical optical fibers (OFs) are an attractive and extensively used medium for both effective laser radiation delivery and advanced biosensor design used in medical diagnostics and therapeutic devices. Optical fibers possess a number of unique properties, such as low attenuation losses, micrometer-scale core diameters, long interaction length, flexibility, immunity to external interference, electrically passive operation, delivery of nonionizing radiation, high sensitivity, and can be easily modified for multiparameter sensor measurements. These advanced features make the biomedical OFs an ideal medium for design and development of novel highly effective biosensing techniques and devices.

In this entry, we present the fundamental principles and features of the most significant and widely used fiber optic-based biosensors. This includes description of a basic FOBS system as well as key optical elements of this system, such as typically used biomedical lasers and incoherent light sources, various biosensing OFs including their advanced properties and critical parameters, and effective light modulator and detector FOBS systems. A major part of this entry is focused on current status as well as future trends in FOBSs applications. The vast variety of both conventional and recently developed FOBSs are summarized in three basic categories: passive (or intrinsic) FOBSs, active (or extrinsic, usually cellular/intracellular) FOBSs, and

direct FOBSs. Among these categories, we consider the ultimate advantages and specific limitations of some of the most effective and well-developed FOBS systems, such as fiber optic-based chemical, evanescent, combined, clinical, and nanobiosensors.

FUNDAMENTALS OF FOBSs

Principal FOBS Design

Biosensors are self-contained analytical devices that are capable of detecting and measuring biologically derived signals. A biosensor device is composed of a specific analyte (a biological sensing substance that needs to be measured), which is in direct spatial contact with a physical transducer that produces either discrete or continuous electrical signals proportional to the analyte concentration. Biosensors can be based on an extensive variety of transducer and detection approaches, such as electrochemical, thermometric, piezoelectric, magnetic, acoustic, optical, etc. Among these technologies, optical biosensors form a major group of biosensing techniques and devices that display specific advanced features in terms of sensitivity, minimal invasiveness, and immunity to all forms of external electrical interference. One of the most effective and extensively used optical biosensors is fiber optic-based biosensing system.^[1-6]

Fiber optic biosensors are analytical optical biosensor devices in which a fiber optic system is utilized as a key sensor component that provides both effective input/output laser radiation delivery and biosensing transduction. A principal design of a basic FOBS system is illustrated in Fig. 1. Using flexible fiber optic delivery links, the input laser emission is effectively and controllably transmitted to a remote biomedical analyte to be tested, and next, the output signal laser emission is returned to a detection system for extracting and analyzing essential biochemical information about the analyte's testing properties.

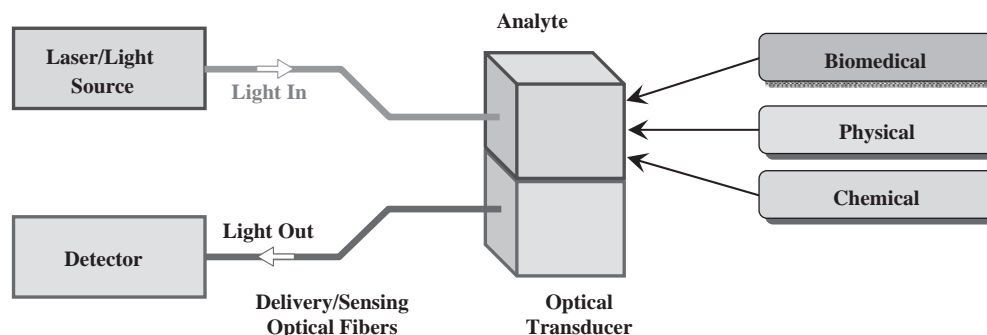


Fig. 1 A principal FOBS design.

Basic Components of a FOBS System

The principal FOBS design described (see Fig. 1) consists of the following basic components: a laser/light source, delivery/sensing OFs, an analyte, an optical transducer, and a detector. Fundamental features and critical parameters of these FOBS components are summarized below.

Laser/light source

A wide selection of both coherent (lasers) and incoherent (e.g., broadband lamps, light-emitting diodes) light sources is available for various FOBS applications. Depending on the specific requirements of a certain FOBS system, the laser sources should possess different features, for example, continuous-wave or pulsed (including ultrashort pulses) generation; a single narrow bandwidth wavelength or a tunable emission in a broadband spectral range; and emission in various spectral regions including the ultraviolet (UV), visible (VIS), and infrared (IR). For some FOBS applications, where a more portable and robust sensor design is required, incoherent light sources (e.g., light-emitting diodes) are preferable because of their significant advantages over other FOBS light sources, such as compactness, low power consumption, selective wavelength emission, simple operation, and relative inexpensiveness. To be able to use these laser/light sources in the FOBS setups they should satisfy specific requirements, the most important of which are intensity-stabilized emission (in some cases power fluctuation less than 0.5% is required), low noise, and compatibility with fiber optic components.

Delivery/sensing OFs

Optical fibers are a key component in FOBS systems because they demonstrate a number of unique features, such as low attenuation losses, effective delivery of nonionizing radiation, small core diameters, long

interaction length, flexibility, immunity to external interference, and enhanced sensitivity in sensor designs. Based on these advanced properties, the OFs play a crucial role in FOBS systems performing simultaneously several significant functions. 1) Flexible OFs provide effective delivery of both the input laser emission to a remote analyte and the output light signal to a detection system. 2) Optical fibers serve as biosensing transducers, which convert detection event into a measurable signal. This OF function can include either a passive mode (when the changes in the testing analyte's properties cause change of the sensing OF properties) or an active mode (when the OF tip serves as a biochemical probe including sensing reagents or molecules placed on the tip). 3) Optical fibers are used for the formation of laser emission with specific features, for example, special laser beam mode distribution (i.e., fundamental Gaussian beam profile using single-mode OFs), energy/power parameters, or spectral characteristics.

Depending on specific FOBS applications and design requirements, an enormous variety of both conventional and recently developed OFs are available. These OFs can be classified in various groups using different criteria. 1) Optical fiber material—silica glass, quartz, plastic, fluoride, chalcogenide, germanium-doped, silver-halide, etc. Silica glass and quartz OFs have low attenuation in the UV (spatially quartz OFs), VIS, and near-IR; however, these OFs have low numerical aperture and are fragile, which presents a safety concern. Plastic OFs have bigger aperture (for better coupling efficiency), and they are mechanically robust, flexible, and inexpensive, but have low UV transmission. The other OFs (fluoride, chalcogenide, germanium-doped, silver-halide) are specialized for mid-IR laser delivery and biosensing (e.g., for biochemical and temperature sensing). 2) Solid-core or hollow OFs—hollow OFs are basically utilized for either mid-IR or powerful UV, VIS, and near-IR laser delivery, but these OFs have poor bending attenuation losses. 3) Optical fiber mode distribution and core dimensions—small core diameter (3–10 μm) single-mode

or large core (50–1000 μm) multimode OFs. Single-mode OFs provide a smooth Gaussian beam profile for precise collimation and focusing, but the coupling efficiencies and absolute laser energy/power delivered is lower than those achieved with large core multimode OFs. 4) Optical fiber configuration—single-input/output OFs, OF bundles, or OF couplers. The single-fiber FOBS configuration ensures a relatively better signal-to-noise ratio than those of the OF couplers. However, the OF couplers combine various sensor functions (input laser coupling, spatial signal separation, and output coupling) in one single element.

Analyte

Target analyte or bioreceptor is the biologically sensitive component of FOBS systems (see Fig. 1) that needs to be measured. It is capable of producing a biochemical signal, which is converted by the transducer into an electrical response. The analyte may be a biological material (e.g., enzymes, protein, DNA, organelles, cellular structures, microbial cells, or bulk mammalian or plant tissue), a biologically derived material (created by biological engineering), or a biomimetic mimicking of bioreceptors.^[1–3] The bioreceptor should have direct contact with the transducing system so that the transmittance should be proportional to the reaction rate with a high range of linearity. Bioreceptors should also have a high sensitivity for the specific compound.

Optical transducer

Optical transducer is a FOBS component that converts detection event into either discrete or continuous measurable electrical signals proportional to the analyte concentration. Typically, OFs or other OF components used in FOBS systems serve as biosensing transducers. Depending on their principle of operation, these transducers can be summarized in three categories: 1) passive (or intrinsic) OF transducers—the changes in the testing analyte's properties affect the optical properties (e.g. intensity, spectral attenuation, and reflectance) of the sensing OFs, i.e., the light modulation is inside the sensing OF; 2) active (or extrinsic) OF transducers—usually, the OF tip serves as a biochemical probe that can involve sensing reagents or molecules placed on the tip and in this case, the light modulation is outside the sensing OF; and 3) direct OF transducers—the OFs are used as a plain transducer to deliver the input laser emission to a remote bulk tissue sample, to return the optical signal to the detection system, and to obtain direct information (e.g., fluorescence or absorption spectra, or reflectance) about the

tissue properties based on the changes in the tissue optical properties.

Detector

Fiber optic biosensor designs include various types of photodetectors. Typically, these FOBS components should satisfy a number of basic requirements, the most important of which are spectral response, sensitivity, noise, and response time. Photodiodes are usually preferable FOBS detectors because they need compact and simple circuits. In some cases, when more sensitive detectors are needed, photomultipliers with specific spectral response and sensitivity should be used.

Advantages and Limitations of FOBSs

Almost three decades have passed since the first fiber optic sensor, used to detect ammonia, appeared in 1976.^[7] Since then, these sensors have advanced both in their diversity and in their accuracy. Among the broad range of fiber optic sensor applications including industry, technology, and environment, the field of fiber optic-based biosensors has one of the best prospects in the future because the FOBSs demonstrate significant advantages that can be summarized as follows. 1) Fiber optic biosensors are non- and minimally invasive techniques and devices. There is no need to take away samples or alter the living system to make samples. 2) Fiber optic biosensors are completely immune to all forms of electrical and electromagnetic interferences, which make them safer than electrochemical biosensors. 3) They are self-contained and do not require external reference signals. 4) Because of low attenuation losses of the sensing OFs, the FOBS allow remote monitoring in hazardous environment and one could perform the sensor measurements from great distances only connected with small core diameter OFs. 5) Fiber optic biosensors system can be easily modified for multiparameter sensor measurement because they can respond simultaneously to more than one analyte. 6) Fiber optic biosensors can be miniaturized and are completely portable, which allows these sensors to be used for in vivo monitoring.

Besides these advantages, FOBSs exhibit specific limitations such as: 1) they usually have a limited sensor's life because of the biological character of their design, for example, biosensors with immobilized dyes have limited long-term stability; 2) they are composed of biological material, which causes problems in sterilization or may be inactivated by strong chemical compounds; 3) they have a limited dynamic range compared to electronic sensors; 4) they have limited sensitivity because the signal levels are proportional to

the small amount of immobilized biological material; 5) they are affected by daylight and light from other sources; and 6) they will work only if appropriate reagents can be developed.

Despite these difficulties, considerable efforts have been made recently to improve some FOBS features in terms of multianalyte capability, integration/miniaturation, implantable devices, and living cell/tissue monitoring.^[2,3,6]

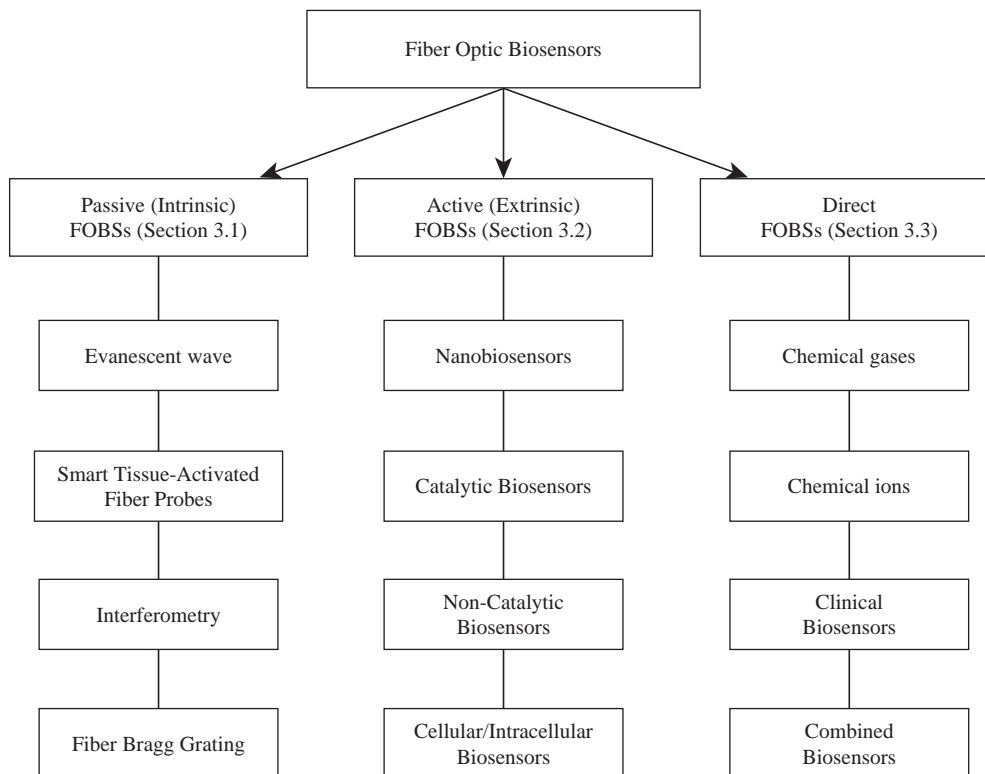
Classification of FOBSs

The enormous variety of both conventional and recently developed FOBSs can be classified using different criteria. For example, with regards to possible measuring principles and testing analyte's parameters, the FOBS classification may include the following categories: absorbance, luminescence, fluorescence, reflectance, scattering, ellipsometry, evanescence, etc. Using a different criterion related to the testing light characteristics, the classification may include the following FOBS types: intensity (or amplitude), frequency (or wavelength), polarization, phase-modulation, etc. However, a more complete classification (see Table 1) consists of three basic categories that are similar to the various types of optical transducers considered in the section "Optical Transducer."

1) Passive (or intrinsic) FOBSs, where the optical properties (e.g., intensity, spectral attenuation, wavelength, and reflectance) of the sensing OFs are modulated by the changes in the testing analyte's properties, that is, the light modulation occurs inside the sensing OF when an external perturbation causes changes in the boundary conditions for the OF. These passive FOBSs may include, for instance, evanescent wave, interferometric, smart FOBSs, etc. 2) Active (or extrinsic) FOBSs, where the OF tip usually serves as a biochemical probe that can include sensing reagents or molecules placed on the OF tip. In this case, the light modulation occurs outside the sensing OF when the testing analyte interacts with the sensing reagents. Basically, these types of FOBSs are related to cellular/intracellular biosensing. 3) Direct FOBSs, where the OFs are used for delivering both the input laser emission (with well-defined optical parameters) to a remote bulk tissue sample and the backward optical signal to the detection system. Next, by analyzing the acquired signals we can get direct information (e.g., fluorescence or absorption spectra, reflectance, or chemiluminescence) about the optical properties of the testing biomedical analytes. These three basic FOBS categories are illustrated in Table 1 and among these categories, some of the most effective and well-developed FOBS systems will be considered in the next section.



Table 1 A. basic FOBSs classification



FIBER OPTIC BIOSENSOR APPLICATIONS

Based on the FOBS fundamental principles and properties described in the section “Fundamentals of Fiber Optic Biosensors,” these fiber optic-based biosensors have been studied and developed for a broad range of biomedical applications. The most effective and well established of these FOBS applications can be summarized into three major categories, identical with the classification presented in the section “Classification of FOBS,” which are passive, active, and direct FOBSs.

Passive FOBSs

Fiber optic biosensor systems operate in a passive sensing mode when the optical properties of the laser/light emission inside the sensing OFs (e.g., light intensity, refractive index, reflectance, spectral attenuation, etc.) are modulated by changes of the analyte–fiber boundary conditions. This FOBS principle is commonly applied for large groups of FOBSs including evanescent wave, smart tissue-activated, and interferometric FOBSs.

Evanescent wave FOBSs

The fundamental waveguiding principle of OFs is based on the effect of total internal reflectance. However, under conditions of total internal reflectance, the transmission coefficients for the transverse electromagnetic field are nonzero, that is, although the light energy is totally reflected, an electromagnetic field penetrates from the interface to the lower refractive index medium. This field, called evanescent wave, decays exponentially with distance from the surface, typically over a distance, which is about a quarter of a wavelength. Decay of the evanescent wave field has been characterized by the penetration depth, d_p , the distance from the interface at which the intensity falls to $1/e$ of its value at the interface, which is defined by the dependence.^[7–9]

$$d_p = \frac{\lambda}{2\pi(n_1^2 \sin^2 \theta_i - n_2^2)^{1/2}}$$

where λ is the light wavelength, θ_i is the incidence angle of light entering the core/cladding boundary, and n_1 and n_2 are the core and cladding refractive indexes, respectively. The evanescent wave FOBSs are based on this evanescent wave phenomenon. Because of the limited penetration depth of the evanescent wave, detection is a surface-sensitive measurement, which has been exploited by a number of biosensor types, such as surface plasmon resonance sensors, interferometers, resonant mirrors, etc.

An effective technique for enhancing the evanescent wave effect is to modify a conventional cylindrical OF by introducing a double-tapered section along the fiber length (see Fig. 2). It will result in decreasing the incident angle of the internally propagated laser beam as well as in increasing the number of reflections inside the fiber. This will create a much larger area of evanescent waves into a medium in which the OF is immersed or the tissue it touches. It has been found that the sensitivity is almost linearly dependent on the reciprocal value of the fiber diameter (D) at the pressed zone. This phenomenon has been used by a few research groups to measure the concentrations of various components in biological media. For example, Katzir’s group, which is developing FOBS applications based on silver-halide OF in the IR range, has measured various blood components (e.g., total protein and cholesterol).^[8] Lucas’s group has used its chalcogenide-based OFs to measure cell metabolism information by attaching their evanescent wave FOBS to tissue.^[9] DNA or protein molecules were immobilized inside the air holes of a two-dimensional photonic band gap OF and specific capture of DNA or protein molecules was obtained. Evanescent wave sensing of the specific binding events by fluorescence detection was analyzed.

Smart tissue-activated FOBSs

A large group of passive FOBSs is based on modulation of reflectance properties of the sensing OFs used. Recently, a simple approach for an on–off switching modulation of reflectance/transmittance properties of so-called smart tissue-activated OF probes has been demonstrated (see Fig. 3A).^[11,12] It is a single-fiber

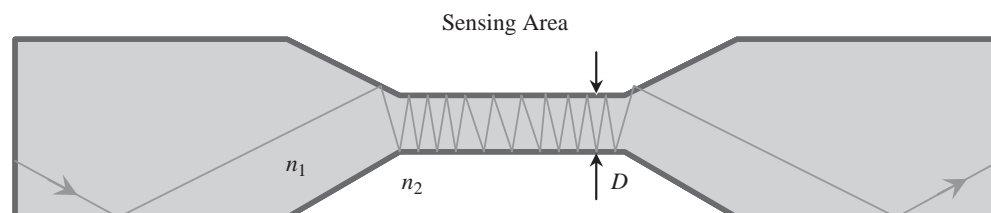


Fig. 2 A double-tapered sensing fiber for evanescent wave spectroscopy.

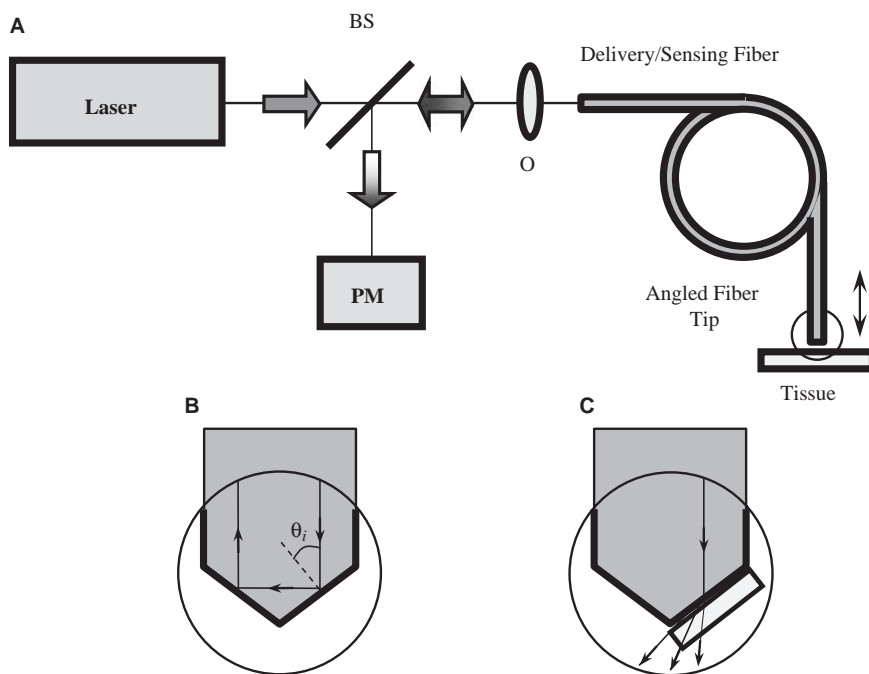


Fig. 3 An FOBS arrangement of (A) using a spatially shaped smart tissue-activated OF tip that provides either a total internal reflectance (B) or frustrated total internal reflectance (C). *Abbreviations:* BS, cube beam-splitter; PM, power meter; O, 10× microscope objective.

delivery/sensing technique based on the principle of frustrated total internal reflectance at the fiber–tissue boundary caused by the refractive index change of the surrounding medium. The key element of the system is the fiber tip shown in Figs. 3B and 3C. To get the on–off switching mode, the OF tip has a specially shaped angled (usually 45°) prism-type profile. Because of the effect of frustrated total internal reflectance, the angled tip acts as a smart, tissue-activated fiber probe. It provides a safe way for laser delivery, which includes only two states of tissue illumination. 1) The first is the off-state or no-tissue illumination mode. In this case (see Fig. 3B), the fiber tip is out of the tissue area, i.e., the surrounding medium is the air that leads to total internal reflection of the forward laser beam at the fiber–tip–air boundary. As a result, the laser emission is backreflected at the angled tip and the tissue area is not illuminated. 2) The second is the on-state or maximum tissue illumination mode. Here (see Fig. 3C), the angled OF tip is placed on the absorbing tissue area, which has a refractive index that is significantly larger than that of the air and close to the refractive index of the OF core. Then, the critical angle for total internal reflection increases and becomes greater than the shaped angle of the OF tip. Thus, the OF tip becomes transparent (or tissue-activated) for laser emission because of the frustrated total internal reflectance. The forward laser emission passes through the OF tip and the energy is coupled into the absorber. In this way, the smart fiber optic probe can recognize the absorbing tissue area and the on–off switching mode of the laser power delivery into the precise tissue area is achieved. The smart OF delivery/sensing system provides a high

spatial resolution in the micrometer and sub-micrometer range ($<2\mu\text{m}$) when the OF tip is moved toward the absorbing tissue area. The results obtained confirm the technique’s potential for on-the-spot laser delivery and sensing (when the backreflectance signal is analyzed), which can be applied for precise tissue diagnostics and laser treatment.

Recently, an attractive type of passive FOBS system has been demonstrated based on fiber optic Fabry–Perot interferometer.^[13] This interferometer is assembled on a single-mode fiber tip by multilayered coatings using an electrostatic self-assembly technique. This biosensor has applications for a precise breath gas analysis. Another FOBS based on modulation optical characteristics of the delivery/sensing OF is fiber Bragg grating FOBS.^[4] The operating principle of these sensors involves implantation of fiber gratings with specific properties into the sensing OF. These biosensors are typically used for precise measurement of temperature and strain.^[14,15]

Active FOBSs

When the biochemical sensing reactions employed for analyte testing possess no intrinsic optical properties, biosensors use an indicator/reactant for indirectly generated optical changes. In this case of active biosensing, the specific response of various indicators is monitored. These indicators might be either chemically (dyes) or biologically (enzymes, antibodies, and proteins) sensitive interfaces. In general, the indicators are

immobilized at the sensing OF tip and they interact with the light guided by OF nanobiosensor probes.

Indicator FOBSs

Indicator-based FOBSs depend on the recognition properties of the most biological components. These components are responsible for the generation of the sensing signal monitored by the FOBS detecting elements. The FOBS biological components can be divided into two categories: catalytic and noncatalytic.^[1,2,5,6]

The catalytic biosensors include enzymes, microorganisms, subcellular organelles, and tissue. These are also known as metabolism biosensors and can be used for continuous-monitoring metabolites by achieving a steady-state concentration of a transducer-detectable species. Enzymes are still indispensable and the most commonly employed biological components of FOBSs because of some their essential advantages, such as high sensitivity and selectivity, fast response time, and potential for multiparameter sensing. However, in some cases the use of enzymes is not feasible because they might be unstable. Then, the use of microorganisms, cells, and subcellular organelles are referable as they usually contain natural and more stable enzymes. The most commonly applied enzymes in FOBS systems are the oxidoreductases, which catalyze the oxidation of compounds using oxygen; the esterases, decarboxylases, and deaminases, which produce acids, CO₂, and NH₃, respectively; and the hydrolases, which catalyze the hydrolysis of compounds. The enzyme-based FOBS systems operate using various optical measurement methods, such as absorbance and reflectance, fluorescence, and chemiluminescence.

The noncatalytic or affinity FOBSs are based on specific binding between the target molecules and the bioreceptors (for instance, antibodies, antigens, nucleic acids, or hormone receptors), which causes physicochemical changes that will be measured by the transducer. These types of biosensors are applicable to single-use disposable devices for measurement of hormones, drugs, toxins, cancer markers, and viruses.

Fiber optic nanobiosensors

Nanotechnology is an emerging growth field in modern science and technology. Over the past several years, rapid progress in nanotechnology has led to the development of novel optical nanobiosensors, which are sensors with dimensions on nanometer scale. This has opened up new horizons for single-cell and intracellular sensing and measurements.^[16–19] After the first optical nanosensor application for intracellular monitoring in 1992, there has been a great impetus to obtain

quantitative chemical information at the cellular and intracellular level in the nanometer scale range.^[16] In this range, optical sensing and imaging techniques have a major drawback related to the fundamental “Rayleigh resolution limit” that is theoretically half the wavelength of the operating radiation. An effective way to avoid the resolution limit is a novel fiber optic-based nanobiosensor approach. It includes the use of a fiber nanometer-sized probe with a specially sharpened tapered tip. The tip is covered with a metal coating except at its very end where a small light-transmitting aperture, having diameter from 20 to 1000 nm, is formed. Such nanobiosensors have potential for direct probing and chemical analysis within individual cells and within the subcellular organelles. Laser emission having specific characteristics can be injected into the cell membrane, mitochondria, and other absorber side inside the cell. Fiber optic nanobiosensors have recently been developed for intracellular measurement of pH, calcium, oxygen, and absorption of glucose molecules by living cells. Because of some unique properties of the nanobiosensor fiber probes, including minimally invasive monitoring, micro- and nanoscale dimensions, effective laser delivery over relatively long distances, flexibility, immunity to external interference, and potential for multiparameter sensing, the fiber optic-based nanobiosensors demonstrate high sensitivity, fast response time, reproducibility, and long-term stability.

Direct FOBSs

Fiber optic biosensor systems operate in a direct sensing mode when the analyte–transducer biochemical reactions lead to products and intermediates that possess specific optical properties (e.g., absorbance, reflectance, fluorescence, luminescence, etc.), which can be directly (without any indicators unlike the active FOBS, see the section “Active Fiber Optic Biosensors”) monitored by the biosensor, that is, these FOBS measure directly a native optical quality. Among the wide variety of available direct FOBSs techniques, two major groups of well-developed and commercialized FOBSs are considered below: direct chemical and combined FOBSs.

Chemical FOBSs

The principles of chemical FOBS systems are based on a direct monitoring of various optical parameters and characteristics of the analyte to be tested, such as absorbance, reflectance, fluorescence, refractive index change, or light scattering, and aimed at measurement of different chemical components, such as oxygen, pH, carbon dioxide, ammonia, blood gasses, glucose,

etc.^[1-6]. Both the commercially available and the recently developed chemical FOBS techniques and devices can be classified into several major categories as follows:

1. Gas FOBSs, which include basic monitoring and measuring of oxygen and other blood gases. Oxygen biosensors or oximetry is a large group of FOBS techniques utilized for measurement of the degree of oxygen saturation, which refers to the amount of oxygen carried by the hemoglobin (the oxygen-carrying pigment) in relation to its maximum capacity. This parameter is measured optically by exploiting the different absorption spectra of deoxyhemoglobin (Hb) and oxyhemoglobin (HbO₂) in the VIS/near-IR region. The measurement is performed at two specific wavelengths: one, in the VIS (e.g., around 660 nm) where there is a large difference in light absorbance between Hb and HbO₂, and the other, in the near-IR (e.g., >805 nm), where there is no significant difference between the Hb and HbO₂ absorbance. A disadvantage of this method is its limited potential for the measurement of gas mixtures with oxygen concentrations larger than 20%, as routinely used in anesthesia. An improved non-invasive biosensor technique for oxygen monitoring is pulse oximetry. It is based on continuous monitoring of arterial oxygen saturation. This technique is a rapidly growing practice in clinical medicine because of its capability to provide continuous, safe, and effective monitoring of blood oxygenation at the patient's bedside without the need of a preliminary calibration. With regard to the blood gases, FOBSs, the research and development efforts are focused on effective disposable and, in particular, intravascular FOBSs that can provide continuous information on different gas concentrations, including oxygen and carbon dioxide partial pressures, and pH.
2. Ion FOBSs, which involve all types of inorganic ions, such as proton (pH), cations, and anions. Biosensors for pH monitoring are among the earliest and the best-established FOBS systems. The interest in these biosensors is due to the importance of precise pH measurement, which is directly related to accuracy in the diagnosis of properly functioning organs and systems within the body. Currently, fluorescence and absorption spectroscopy are the most widely used techniques for pH measurements in biomedical applications. In comparison to the fluorescence approach, the absorption biosensors are simple and easy to use, but not as sensitive. Absorption FOBS systems are also difficult to miniaturize. In contrast, the fluorescence method is more sensitive and can be used

for small-size sensors and/or low-concentration measurements. Further developments of the pH biosensors for practical use are focused on the design of novel OF probes with improved parameters (e.g., providing an accuracy of >0.001 pH units) and potential for pH imaging inside living cells using fluorescence nanobiosensors.

3. Clinical FOBSs, which, in addition to the gas and ion FOBS systems described, include other important types of biosensors, such as glucose, pharmaceutical, DNA, and cholesterol FOBSs. Development of non- or minimally invasive methods for blood glucose sensing in human subjects is a problem of major importance for clinical laboratories. Among a large number of various approaches applied for noninvasive glucose detection, optical sensing techniques, including FOBS systems based on absorption, fluorescent, and Raman spectroscopy, and polarimetry, have predominantly been investigated because of their advantages, such as no dependence on chemical reactions or enzymatic lifetimes and no fluorescent molecule relating specifically to glucose. Recently, near-IR and mid-IR spectroscopy has been also studied as an alternative noninvasive technique with a number of potential advantages: no required reagents, multicomponent chemical sensing from a single spectrum, and existence of spectral windows where hemoglobin, melanin, and water absorption intensities are low. However, more precise determination of specificity and calibration is needed, and yet no commercialized noninvasive glucose product is available.

F

Combined FOBSs

The use of a single-sensor method provides basically only relative measurements of the analyte's properties. In many cases, for either absolute measurements or comparative analyte testing, combined sensor methods should be utilized. Recent research efforts have been focused, for instance, in combining various FOBS techniques with high-resolution bioimaging systems, such as near-field and confocal microscopy, and optical coherence tomography.

A combined FOBS technique for direct sensing and monitoring of single-cell and intracellular components is shown in Fig. 4A. This design combines simultaneously a dual-nanobiosensor-probe technique and a high-resolution near-field microscopy that provides direct spectroscopic measurements and monitoring of cellular/intracellular components. It is a dual-probe arrangement that includes two identical nanobiosensor channels. One channel is used for precise delivery of the input light emission from either a coherent laser

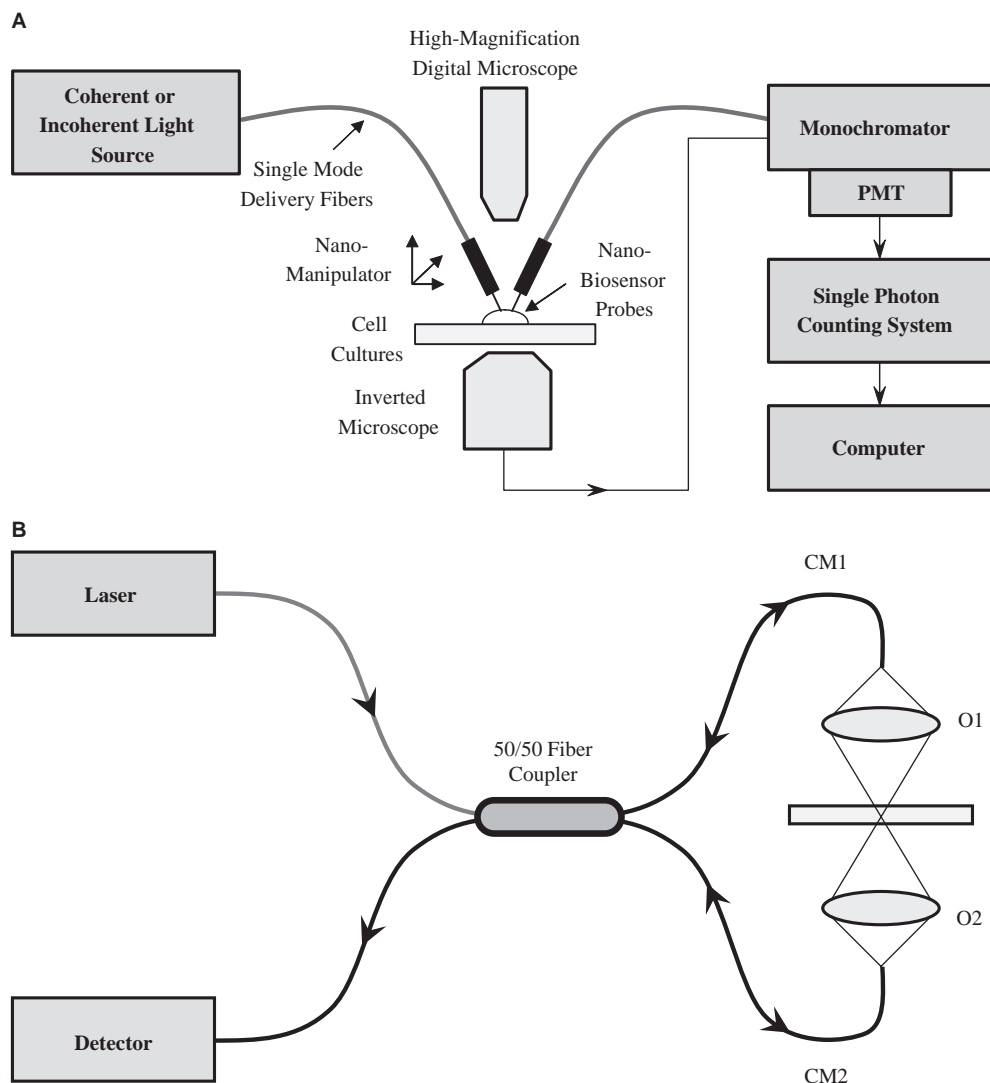


Fig. 4 Principal designs of a combined dual-nanobiosensor-probe setup (A) and a combined dual-confocal FOBSs (B).

source or a broadband incoherent source covering the UV, VIS, and near-IR spectral range. The second channel is used for detecting the output signals from the sensing of various cellular chemical components. In both the channels, small core diameter (3–10 μm) single-mode (for specific laser wavelengths) OFs equipped with tapered nanoscale tips are used as nanobiosensor probes. In addition, the nanobiosensor design is compatible with an ultrasensitive detection technique based on near-field microscopy. In this case, the input OF tapered nanobiosensor probe is used to inject light into specific parts of a living cell and the signal spectra from a near-field spatial zone are detected using a high-numerical-aperture ($\text{NA} > 1$) microscope. The ultimate advantage of the near-field technique equipped with nanosensor probes is the potential for breaking the subwavelength diffraction limit for high-resolution microscopic and broadband

spectroscopic analysis of single-cell components. Moreover, the combined FOBS technique (Fig. 4A) is compatible with a number of basic spectroscopic methods, such as absorption, fluorescence, and Raman spectroscopy.

A noninvasive combined dual-confocal FOBS technique for multiparameter sample sensing is illustrated in Fig. 4B.^[20,21] This is a dual-confocal sensor system that includes two identical confocal microscopes (CM1 and CM2) used as independent confocal sensor channels. The microscopes are based on a fiber optic-type apertureless confocal arrangement that excludes either the input or the output aperture, unlike the conventional pinhole-type confocal microscope. The combined dual-confocal FOBS design provides direct measurement in absolute units of refractive index and thickness of optically transparent and non-transparent media. Because of the typically high spatial

sensitivity of the fiber optic-based confocal approach, the dual-confocal sensor ensures high accuracy in spatially locating the image confocal points and objects ($<1\ \mu\text{m}$), and therefore simultaneously measuring the refractive index and thickness in absolute units. Such a system can be potentially used for precise measurement of optical properties of the various tissue layers and determining small changes in these properties between cancer and noncancer tissue areas.

CONCLUSIONS

Fiber optic biosensors have been under development for 30 yr. Nonetheless, the technology continues to evolve with new breakthroughs in optics, biochemistry, bioengineering, and nanotechnologies. Recent study and development efforts have been focused on significant improvement of some FOBS critical features including multianalyte detection potential, integration/miniaturation, and living cell/tissue monitoring.

ARTICLE OF FURTHER INTEREST

Biosensors

REFERENCES

- Monk, D.; Walt, D. Optical fiber-based biosensors. *Anal. Bioanal. Chem.* **2004**, *379* (7–8), 931–945.
- Wolfbeis, O. Fiber-optic chemical sensors. *Anal. Chem.* **2004**, *76* (12), 3269–3284.
- Marazuela, M.; Moreno-Bondi, M. Fiber-optic biosensors—an overview. *Anal. Bioanal. Chem.* **2002**, *372* (5–6), 664–682.
- Lee, B. Review of the present status of optical fiber sensors. *Opt. Fiber Technol.* **2003**, *9* (1), 57–79.
- Paddle, B. Biosensors for chemical and biological agents of defence interest. *Biosens. Bioelectron.* **1996**, *11* (11), 1079–1113.
- Mehrvan, M.; Bis, C.; Scherer, J.; Young, M.; Luong, J. Fiber-optic biosensors—trend and advances. *Anal. Sci.* **2000**, *16* (5), 677–692.
- David, D.; Wilson, M.; Ruffin, D.S. Direct measurement of ammonia in ambient air. *Anal. Lett.* **1976**, *9* (3), 389–393.
- Raichlin, Y.; Fel, L.; Katzir, A. Evanescent-wave infrared spectroscopy with flattened fibers as sensing elements. *Opt. Lett.* **2003**, *28* (23), 2297–2299.
- Hocde, S.; Loreal, O.; Sire, O.; Boussard-Pledel, C.; Bureau, B.; Turlin, B.; Keirsse, J.; Leroyer, P.; Lucas, J. Metabolic imaging of tissues by infrared fiber-optic spectroscopy: an efficient tool for medical diagnosis. *J. Biomed. Opt.* **2004**, *9* (2), 404–407.
- Jensen, J.; Hoiby, P.; Pedersen, L.H.; Carlsen, N.L.; Bjarklev, A.; Hansen, T. Evanescent wave sensing using a hollow-core photonic crystal fiber. *Proceedings SPIE* 5317, San Jose, CA, 2004; 139–146.
- Ilev, I.; Waynant, R.; Byrnes, K.; Anders, J. On-off laser delivery into a precise tissue area using smart tissue-activated fiber probes. *IEEE J. Sel. Top. Quantum Electron.* **2003**, *9* (2), 331–336.
- Ilev, I.; Waynant, R. All-fiber-optic sensor for liquid level measurement. *Rev. Sci. Instrum.* **1999**, *70* (10), 2551–2555.
- Zhang, Y.; Shih, H.; Cooper, K.; Wang, A. Miniature fiber-optic multicavity Fabry–Perot interferometric biosensor. *Opt. Lett.* **2005**, *30* (9), 1021–1023.
- Bhatia, V.; Campbell, D.; Claus, R.; Vengsarkar, A. Simultaneous strain and temperature measurement with long-period gratings. *Opt. Lett.* **1997**, *22* (9), 648–650.
- Han, Y.; Lee, S.; Kim, C.; Kang, J.; Paek, U.; Chung, Y. Simultaneous measurement of temperature and strain using dual long-period fiber grating with controlled temperature and strain sensitivities. *Opt. Express* **2003**, *11* (5), 476–481.
- Tan, W.; Shi, Z-Y.; Kopelman, R. Development of submicron chemical fiber optic sensors. *Anal. Chem.* **1992**, *64* (10), 2985–2990.
- Tan, W. Optical measurements on the nanometer scale. *Trends Anal. Chem.* **1998**, *17*, 501–513.
- Barker, S.; Thorsrud, B.; Kopelman, R. Nitrite- and chloride-selective fluorescent nano-optodes and in vitro applications to rat conceptuses. *Anal. Chem.* **1998**, *70*, 100–104.
- Niedre, M.; Patterson, M.; Wilson, B. Direct near-infrared luminescence detection of singlet oxygen generated by photodynamic therapy in cells in vitro and tissues in vivo. *Photochem. Photobiol.* **2002**, *75* (4), 382–391.
- Ilev, I.; Waynant, R.; Byrnes, K.; Anders, J. Dual-confocal fiber-optic method for absolute measurement of refractive index and thickness of optically transparent media. *Opt. Lett.* **2002**, *27*, 1693–1695.
- Ilev, I.; Waynant, R. A simple submicron confocal microscope with a fiber-optic output. *Rev. Sci. Instrum.* **2000**, *71*, 4161–4164.

Fibrin

Sam L. Helgerson

BioScience R&D, Baxter BioScience, Fremont, California, U.S.A.

Thomas Seelich

BioSurgery R&D, Baxter BioScience, Vienna, Austria

James P. DiOrio

Baxter Technology Resources, Round Lake, Illinois, U.S.A.

Bill Tawil

BioSurgery Marketing, Baxter BioScience, Westlake Village, California, U.S.A.

Katharina Bittner

BioSurgery European Marketing, Baxter BioScience, Munich, Germany

Reiner Spaethe

BioSurgery R&D, Baxter BioScience, Vienna, Austria

INTRODUCTION

Fibrin clots are self-assembling biopolymer networks with important properties that can be exploited in the design of novel biomaterials. This article will summarize key aspects of the biochemical, biomechanical, and biological properties of fibrin clots formed by human plasma-derived fibrinogen, thrombin, and associated proteins. Also, recent work exploring new applications for fibrin-based biomaterials will be highlighted focusing on the delivery of therapeutic agents and cells for wound-healing and tissue-regeneration applications.

OVERVIEW

The process of fibrin clot formation is diagrammed in Fig. 1. Fibrinogen is a plasma protein composed of two sets of three polypeptide chains ($A\alpha$, $B\beta$, γ) that are linked by disulfide bridges.^[1,2] The $A\alpha$, $B\beta$, and γ chains have 610, 461, and 411 amino acid residues, respectively. The rod-shaped fibrinogen protein is approximately 45 nm long with two compact outer D domains each joined by flexible-coiled coil segments to the inner E domain.^[3] The conversion of fibrinogen to fibrin is catalyzed by the serine protease thrombin, generated by activation of prothrombin during the blood coagulation cascade. Thrombin binds to the central E domain of fibrinogen and catalyzes the removal of two sets of small N-terminal fibrinopeptides, FpA and FpB, from the $A\alpha$, and $B\beta$, chains, respectively. The resulting fibrin monomers then self-associate in a half-staggered manner (22.5 nm repeat spacing) with the end D domain of one molecule interacting with the central E domain of another to form linear

protofibrils. These protofibrils further self-assemble into branched fiber bundles, resulting in the formation of a characteristic biopolymer gel material with an open three-dimensional porous network structure.

The structure of a fibrin clot as seen by scanning electron microscopy (SEM) is shown in Fig. 2A. The fibrin network structure and porosity are clearly evident. A higher magnification view of an individual fiber bundle as seen by transmission electron microscopy (TEM) is shown in Fig. 2B. The 22.5-nm repeat pattern indicating the well-ordered assembly of the fibrin monomers within the fiber can be seen (arrow). The fibrin clot shown in Fig. 2 was prepared from a commercially available fibrin sealant product. However, the structural characteristics of the fiber network are similar to clots formed in vitro from purified fibrinogen samples or under more physiological conditions that mimic the formation of plasma clots in vivo. This represents a key aspect in the versatility of fibrin-based biomaterials. The overall structure of the biopolymer gel is determined primarily by the thrombin-mediated conversion of fibrinogen to fibrin and the subsequent self-assembly of the fiber network. The presence of other naturally occurring accessory proteins, e.g., factor XIII, fibronectin, and plasminogen, can then be used to add important biomechanical and biological properties to the final fibrin-based biomaterial with relatively small effects on the overall physical structure.

HISTORY OF FIBRIN-BASED BIOMATERIALS

The development of fibrin-based biomaterials has been closely linked with the use of fibrin sealant preparations. Fibrin sealants are products currently approved

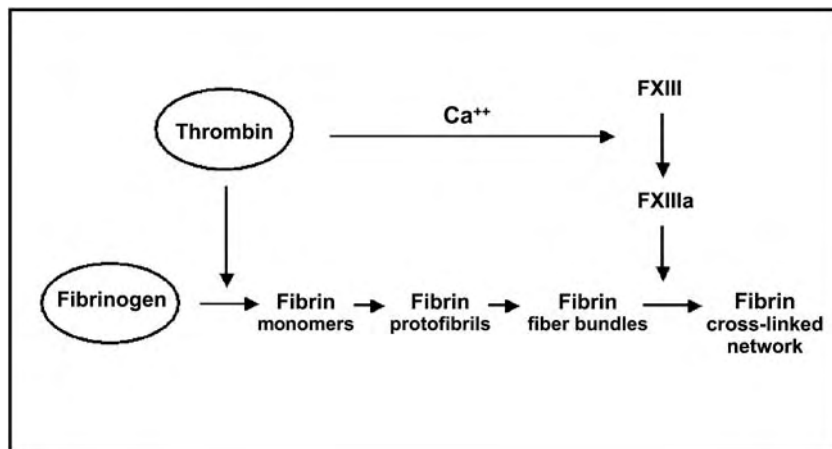


Fig. 1 Diagram of fibrin clot assembly.

for use in surgical procedures to achieve rapid hemostasis and to seal tissue surfaces together.^[2] They consist of two human plasma-derived components: 1) a highly concentrated fibrinogen complex composed

primarily of fibrinogen and fibronectin along with catalytic amounts of factor XIII and plasminogen and 2) thrombin and calcium ions. The rapid mixing of the two components produces a fibrin-based biopolymer that acts as a strong, bioresorbable adhesive coating especially useful in surgery on soft tissues and organs. Surgical procedures where fibrin sealant products are used include cardiovascular, thoracic, neurological, and reconstructive techniques. Fibrin-based preparations for the experimental development and testing of new biomaterials can be readily prepared from the fibrinogen complex and thrombin components of fibrin sealants.

The natural blood coagulation process provides the starting paradigm for how bleeding can be stopped, tissue parts sealed together, and normal wound healing achieved through the use of fibrin-based biomaterials. It took until the beginning of the 19th century to develop the first scientific concepts of the blood coagulation process. In 1905, P. Morawitz presented his model of blood coagulation, which at that time was partly based on hypothesis. He correctly perceived that by the action of calcium ions and tissue fluid, the prothrombin present in blood is converted into active thrombin, which in turn converts fibrinogen into fibrin.^[4] In 1909, S. Bergel postulated, based on animal experiments, that the physiological function of fibrin is not only to stop bleeding and seal wound edges together, but also to stimulate subsequent wound healing. He recommended the use in surgery of a fibrin-powder preparation prepared from equine or bovine blood.^[5] As early as 1915 and 1916, E. G. Grey^[6] and S. C. Harvey^[6] used fibrin tampons and fibrin patches to control bleeding from parenchymatous organs. In 1943, R. T. Tidrick and E. P. Warner,^[6] as well as E. P. Cronkite,^[6] reported on the use of either citrated human plasma or human fibrinogen clotted with bovine thrombin for the fixation of skin grafts in humans. Although these early attempts already suggested the basic advantages of using fibrin for wound

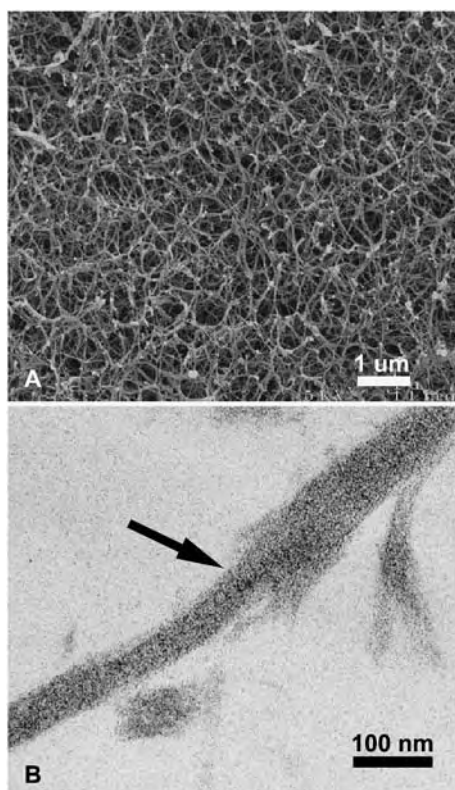


Fig. 2 Electron micrographs demonstrating the key structural features of fibrin clots. A) The overall clot structure including the branching of individual fibers to form an open three-dimensional porous network structure as seen by SEM. B) The 22.5-nm repeat spacing (arrow) of individual fibrin protein monomers making up the assembled fibers as seen by TEM. The clot was prepared with the fibrinogen complex and thrombin components of the Fibrin Sealant product Tisseel^R at final concentrations of 50 mg total protein/ml and 2 U/ml, respectively.

closure, such as excellent tissue tolerance, improved wound healing, and complete in-vivo adsorption, there was a relatively high rate of failure due to poor adhesive strength and durability of the fibrin preparations. The reason was probably the lack of sufficiently concentrated fibrin solutions together with the lack of insight into the underlying biochemistry of the new method. Thus, none of the authors reported the use of calcium chloride or of a fibrinolysis inhibitor in the fibrin-sealant mixture, which are now known to be essential for quality and efficacy in sealing applications.

It was due to these unsatisfactory results that the new technique was not pursued further for some thirty years. Instead, other approaches to achieve hemostasis by use of fibrin in combination with thrombin were investigated. In 1944, E. A. Bering, Jr.^[7] reported on the preparation of fibrin foams to be used as hemostats in combination with thrombin. In 1968, M. Gerendás^[8] described in a review article the preparation and clinical application of bovine-derived fibrin foam, fibrin powder, and fibrin films, as well as a novel slowly resorbable fibrin implant material. The latter was obtained from a mixture of fibrin powder and plasticizer (polyvalent alcohol) under high pressure and temperature in specific molds. A variety of implants were produced and excellent clinical results were reported, including in hip arthroplasty to cover the femoral head and in the regeneration of cartilage.

In parallel with these developments, ongoing basic research led to a better understanding of the reactions of coagulation and fibrinolysis. In 1948, the fibrin stabilizing factor (fXIII) was discovered by K. Laki and L. Lorand, followed by numerous studies on the biochemical and biological properties of this enzyme, including the importance of the fXIII-mediated cross-linking of fibrin to the mechanical strength and resistance to fibrinolytic degradation of the fibrin clot.^[9] A new era in the use of fibrin sealants began in 1972, when H. Matras and coworkers^[10] successfully used a highly concentrated human plasma-derived fibrinogen solution in combination with fXIII, thrombin, and calcium chloride to seal severed nerves in animal experiments. In 1973, H. P. Spängler^[11] used such a preparation for the first time clinically in an emergency situation in heart surgery to successfully stop severe bleeding that could not be controlled by other conventional means.

The efficacy of this biomaterial in heart surgery as well as other surgical applications opened the way for the development of the first commercial fibrin sealant, Fibrinkleber Human Immuno^R, which entered the market in 1978, followed by an improved version in 1980 available under the trade names Tisseel^R and Tissucol^R in different countries. Further developments of this product have included the continued

enhancement of viral safety profiles and the improvement of handling properties without compromising the required biochemical and physiological properties. Due to its pioneering position, this product may well be regarded as the conceptual prototype of fibrin sealant products now approved or under development.^[12] Fibrin sealant preparations have a long history and continue as an important platform in the development of innovative biomaterials.

FIBRIN CLOT STRUCTURE AND BIOMECHANICAL PROPERTIES

Many factors such as the ratio of thrombin to fibrinogen, the ionic strength and divalent cation concentration of the medium, or the presence of certain plasma proteins, e.g., serum albumin, regulate the rate of fibrin polymerization, and thus, the resulting clot structure.^[1,13-15] The rate of release of the fibrinopeptides FpA and FpB, as well as other factors affecting the lateral association of the fibers, regulate fiber thickness, degree of branching, and clot porosity.^[16] These possibilities to regulate fibrin formation and assembly can be employed to control the physical structure of the final fibrin-based biomaterial and have been shown to be important for wound healing applications.^[17-19]

Examples of different fibrin clot structures that can be prepared by varying some of these parameters are shown in Figs. 3 and 4. In Fig. 3, SEM micrographs are shown for fibrin clots formed using two different thrombin concentrations, A) 25 U/ml or B) 0.25 U/ml, to clot a fibrinogen complex preparation (5 mg total protein/ml). The fiber structure formed was much finer at the higher thrombin concentration with less porosity and many more branch points compared to the thicker, more laterally associated fibers at the lower thrombin concentration. In Fig. 4, fibrin clots were formed at three different salt concentrations, A) 0.07 M, B) 0.13 M, or C) 0.20 M, using the same fibrinogen complex protein (22.5 mg/ml) and thrombin (75 U/ml) concentrations. As the salt concentration was increased, thinner fibers with more branch points were formed leading to a fibrin network of lower porosity.

Fibrin clots are biopolymers with viscoelastic physical characteristics. The structure and composition of the final clot directly affect the biomechanical properties as measured by elastic modulus, breaking strength, and adhesiveness. Measurements of the mechanical properties are useful in predicting the performance of a fibrin-based preparation under shear force conditions in vivo.^[2] The elastic modulus increases as a power log of the fibrinogen concentration, as shown with fibrin clots formed from purified fibrinogen.^[15] An increase of elastic modulus, tensile

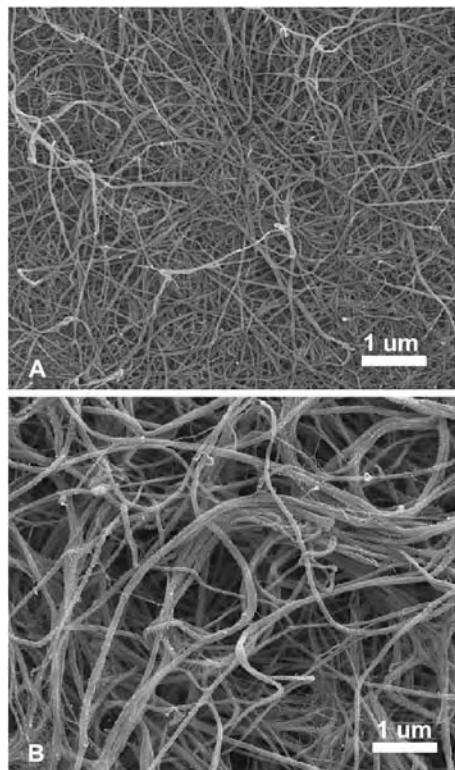


Fig. 3 The effect of varying the thrombin potency on fibrin clot structure. Thrombin at A) 25 U/ml or B) 0.25 U/ml was used to clot a fibrinogen complex preparation at 5 mg total protein/ml (final concentrations). Note the relatively thinner individual fibers and decreased overall porosity for the fibrin clot formed using higher versus lower thrombin potency.

strength, and adhesiveness, as a function of fibrinogen concentration, has also been seen for clots formed from fibrin sealant preparations.^[2,20]

Factor XIII-mediated cross-linking of the fibrin monomers that make up the fiber bundles also contributes considerably to the strength characteristics of the fibrin network. In the presence of thrombin and calcium ions, fXIII is activated to fXIIIa and catalyses the formation of covalent bonds between fibrin chains, resulting in the formation of both γ - γ and α - α chain polymers.^[1] This stabilizes the fibrin network and creates an insoluble fibrin clot, i.e., one that is insoluble in solutions of chaotropic agents such as monochloroacetic acid or urea. The elastic modulus increases as a function of the fXIII content, as shown with fibrin formed from purified fibrinogen. fXIII can increase the elastic modulus fivefold under physiological conditions.^[21] This has also been seen for clots formed from fibrin sealant preparations^[2,20] where α - α polymer formation has been shown to contribute greatly to the tensile strength.^[22] The ionic strength of the fibrin sealant formulation can also affect the degree of cross-linking.^[6] The change in elastic modulus seen

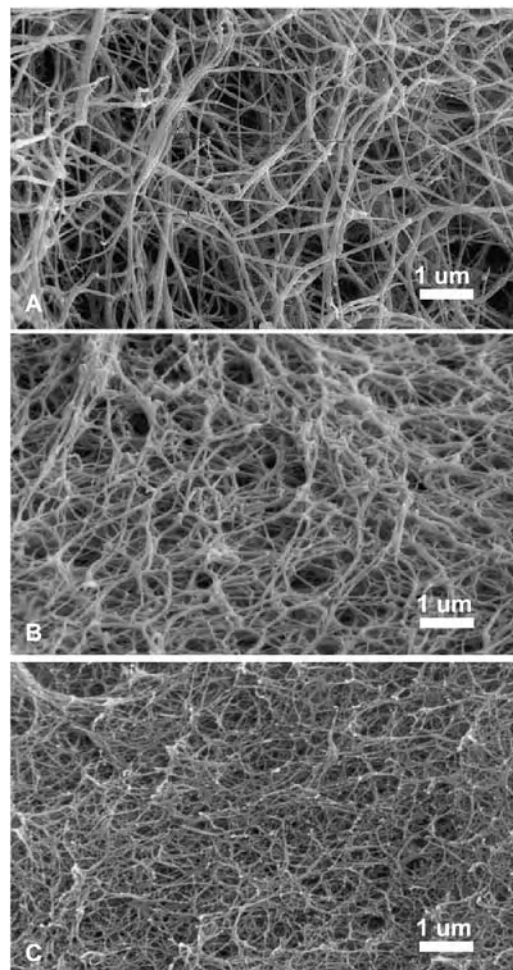


Fig. 4 The effect of varying the salt concentration of the medium on fibrin clot structure. The NaCl concentration during clot formation was A) 0.07 M, B) 0.13 M, or C) 0.20 M. The clots were prepared with the fibrinogen complex and thrombin components at final concentrations of 22.5 mg total protein/ml and 75 U/ml, respectively. Note the relatively thinner individual fibers and decreased overall porosity in the fibrin clot as the salt concentration of the medium increases.

with increasing levels of fXIII does not seem to correlate with significant changes in the clot structure. Similar fiber thickness was seen for clots formed with and without various levels of fXIII.^[20] Fiber mass-to-length ratios were shown to be similar with and without fXIII, but an increased density of fibers in the network was seen with the addition of fXIII.^[23]

FIBRIN BIOMATRIX FOR TISSUE REGENERATION

An appropriate biomatrix is a key structural element for tissue regeneration since a provisional extracellular matrix is required to support the highly organized

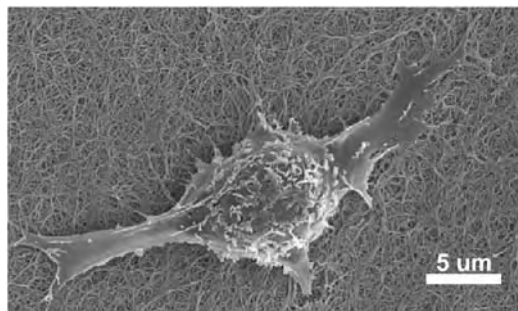


Fig. 5 Human dermal fibroblast cell adhered to and growing on a fibrin clot surface as seen by SEM. The clot was prepared with fibrinogen complex and thrombin components at final concentrations of 5 mg total protein/ml and 5 U/ml, respectively. The human dermal fibroblast cells were seeded onto the clot surface and then grown under standard cell culture conditions for 24 hours at 37°C.

remodeling and repair processes involved in forming new tissue. Fibrin-based biomaterials provide a biocompatible physical structure and many of the important biological activities needed for successful tissue regeneration. Fibrin is the natural provisional matrix for wound healing and tissue repair. The fibrin clot provides a structural scaffold for the adhesion, proliferation, and migration of cells important in wound healing (see, for example, Fig. 5). Also, fibrin and the associated proteins making up a physiological clot have intrinsic biological activities that support, and in some cases control, the cellular differentiation important in wound healing. Finally, fibrin clots are readily remodeled and resorbed through normal fibrinolytic processes as cells deposit the tissue-specific extracellular matrix components during the regeneration of functional tissues.

Biological Activities of Fibrin(ogen) and Associated Proteins

The first function of fibrin(ogen) in the overall process of blood coagulation and wound healing is to rapidly control bleeding at the site of injury. Once hemostasis is achieved, the fibrin clot serves as the foundation of a complex provisional matrix both to initiate and support subsequent tissue repair.^[24,25] The fibrin(ogen) protein has several specific domains that actively control cell behavior during wound healing. The fibrinopeptides FpA and FpB released by thrombin during clot formation, and fibrin degradation products resulting later during fibrinolysis, are chemoattractants and activators of leukocytic cells. The clot provides a physical scaffold for the migration of neutrophils and monocytes into the wound as part of the initial inflammatory response to protect the injured tissue.

Fibrin has peptide domains that facilitate cell adhesion to the clot, serving as binding sites for integrin receptors on cells involved in tissue repair. Similar binding sites on proteins associated with the clot (fibronectin), or deposited during extracellular matrix remodeling (collagens), play primary roles in cell migration. The release of Fibrinopeptide B from the β -chain of fibrinogen exposes a new N-terminal peptide domain on fibrin that serves as a unique binding site for the vascular endothelial cell surface receptor VE-cadherin.^[26] Fibrin-specific signaling through this receptor causes the vascular endothelial cells in injured blood vessels to migrate, proliferate, differentiate, and form new blood capillaries. The fibrin clot then serves as the physical scaffold for this angiogenesis at the wound site. The resulting neovascularization is critical for successful tissue regeneration.

The biological activities inherent in the final clot are not limited to fibrin, but include a complex mixture of other proteins that bind to the fibrin matrix.^[1] This is true both for physiological fibrin clots and for those made from fibrin sealant preparations. Thrombin remains bound reversibly to fibrin after the clot forms and is protected from inhibition by plasma factors such as antithrombin III. Thus, the fibrin clot is a storage matrix for thrombin that could affect wound healing through interactions with protease-activated receptors (PAR) present on epithelial, fibroblast, endothelial, leukocytes, and mast cells.^[27] Factor XIIIa, which stabilizes the clot by covalently cross-linking fibrin monomers, also cross-links other plasma proteins to fibrin. The most important of these are the plasma-derived proteins fibronectin, which is also normally present in fibrin sealant preparations, and α_2 -antiplasmin and plasmin activator inhibitor 2 (PAI-2), which help to control subsequent fibrinolysis of the clot.^[1] The cross-linked plasma fibronectin then provides specific receptor binding sites for cells involved in wound healing (keratinocytes, fibroblasts, vascular endothelial cells) to adhere to and migrate through the fibrin matrix, thus playing a role similar to fibronectins present in the extracellular matrix of tissues.^[28]

The biological activities inherent in fibrin-based biomaterials can be enhanced by incorporating novel therapeutic agents to control cell proliferation, migration, or differentiation during tissue regeneration. Certain growth factors, e.g., basic fibroblast growth factor (bFGF), the insulin-like growth factor-1/insulin-like growth factor binding protein-3 complex (IGF-1/IGFBP-3), and the vascular endothelial growth factor isoform VEGF₁₆₅, specifically bind to fibrin(ogen).^[1,29] These bioactive proteins can be delivered using fibrin-based biomaterials without further modifications.^[30,31] Other therapeutic proteins or drug molecules that do not specifically bind can be added to the biomaterial

preparations and delivered by freely diffusing from the biopolymer matrix.^[32] However, this latter approach does not allow for careful control over the timing or extent of delivery. Methods are being developed to modify either the fibrin-based matrix or the bioactive substance such that therapeutic proteins, drugs, or DNA can be delivered in a controlled manner.^[33,34]

Applications of Fibrin-Based Materials in Tissue Regeneration

Applications for fibrin-based biomaterials in tissue regeneration include the delivery of tissue-specific cells in order to stimulate the wound healing process. The regeneration of damaged skin based on the combined use of keratinocytes and a fibrin matrix was first described in 1988.^[35] Keratinocytes isolated from skin biopsies were cultivated under physiological conditions and then applied as keratinocyte/fibrin matrix suspension, resulting in the neof ormation of epidermal tissue.^[36] Pellegrini et al.^[37] have shown that the relative percentage of epidermal stem cells found in vivo is maintained when isolated keratinocytes were cultivated on a fibrin matrix and that the subsequent transplantation efficiency for the keratinocytes when applied as an autograft was high, reproducible, and permanent. The use of fibrin microbeads is another approach to deliver fibroblasts for skin regeneration^[38] and might be useful for other cell types as well.

In addition to skin wound-healing applications, further investigations have been done in vitro as well as in vivo to combine fibrin-based biomaterials with other tissue-specific cells. Wechselberger et al.^[39] showed that cultured urothelial cells could be successfully transplanted via fibrin glue onto a prefabricated structure. Based on this approach, Bach et al.^[40] were able to show the in-vivo formation of a new multilayered urothelium at the site of transplantation. Taken together, both studies point out the promising potential for an urothelial cell/fibrin matrix suspension in urethral reconstruction procedures. In a recent study, Wechselberger et al.^[41] extended the use of cell/fibrin matrix suspensions to the delivery of tracheal epithelial cells and preadipocytes. Other cell types including chondrocytes,^[42] human myofibroblasts,^[43] and osteoprogenitor cells,^[44] have also been grown in fibrin gels and tested for tissue regeneration applications.

The combination of fibrin-based biomaterials with other biomaterials provides possibilities for tissue regeneration approaches, either to add additional biological activities or to improve the biomechanical properties of the composite matrix. The combination of fibrin-based preparations with inorganic bone ceramics, e.g., hydroxyapatite and β -tricalcium phosphate,

has been tested in preclinical and clinical orthopedic procedures. It has been shown that a fibrin matrix enhances the osteogenic properties of the ceramic materials including osteointegration, bone remodeling, and new bone formation.^[45,46] These effects may be explained by a stimulatory effect of the fibrin matrix on osteoblastic precursor cells and mesenchymal cells to enhance osteogenesis. The ability of the fibrin matrix to promote angiogenesis and vascular development is another important property of fibrin-based biomaterials in the context of bone regeneration.

F

CONCLUSION

The ultimate goal of tissue regeneration therapies is the controlled formation of entire tissue and organ systems in vivo, thereby avoiding the problems of donor transplant shortages and tissue rejection reactions. One barrier to tissue regeneration strategies is the complexity and variety of the cells required for the task, whereas another is the limitation of the biomaterials currently available. Fibrin-based biomaterials can address problems involved in the uniform distribution of cells and nutrients throughout the scaffold, the delivery of growth factors and cytokines to induce specific cellular growth, migration, and differentiation responses, and revascularization of the newly formed tissue. In addition, the proven hemostatic and tissue-adhesive properties of fibrin sealants can be utilized in applying the fibrin-based biomaterials in vivo, while the inherent wound-healing properties of fibrin clots can enhance the successful integration of the constructs. Continuing advances in fibrin-based biomaterials including methods for delivery of cells and therapeutic agents provide a basis for creating the next generation of biomaterials^[47] designed to achieve site-directed stimulation of tissue regeneration processes.

ARTICLES OF FURTHER INTEREST

Bioadhesion; Biologic and Synthetic Apatites; Biological Adhesives from Nature; Burn Dressing; Fibrin Sealants; Thrombosis; Tissue Engineering Scaffolds

REFERENCES

1. Mosesson, M.W.; Siebenlist, K.R.; Meh, D.A. The structure and biological features of fibrinogen and fibrin. *Ann. N.Y. Acad. Sci.* **2001**, *936*, 11–30.
2. Sierra, D.H. Fibrin sealant adhesive systems: A review of their chemistry, material properties, and clinical applications. *J. Biomater. Appl.* **1993**, *7*, 309–352.

3. Doolittle, R.F.; Yang, Z.; Mochalkin, I. Crystal structure studies on fibrinogen and fibrin. *Ann. N.Y. Acad. Sci.* **2001**, *936*, 31–43.
4. Morawitz, P. Die chemie der blutgerinnung (The chemistry of blood coagulation). *Erg. Physiol.* **1905**, *4*, 307–422.
5. Bergel, S. Ueber die wirkungen des fibrins (About the effects of fibrin). *Dtsch. Med. Wochenschr.* **1909**, *35*, 663–665.
6. Seelich, T. Tissucol^R (Immuno, Vienna): Biochemistry and methods of application. *J. Head Neck Pathol.* **1982**, *3*, 65–69.
7. Bering, E.A., Jr. Chemical, clinical, and immunological studies on the products of human plasma fractionation. XX. The development of fibrin foam as hemostatic agent and for use in conjunction with human thrombin. *J. Clin. Invest.* **1944**, *23*, 586–590.
8. Gerendás, M. Fibrin Products as Aids in Hemostasis and Wound Healing. In *Fibrinogen*; Laki, K., Ed.; Marcel Dekker, Inc.: New York, 1968; 277–316.
9. Board, P.G.; Losowsky, M.S.; Miloszewski, K.J.A. Factor XIII: Inherited and acquired deficiency. *Blood Rev.* **1993**, *7*, 229–242.
10. Matras, H.; Dinges, H.P.; Lassmann, J.; Mamoli, B. Zur nahtlosen interfazikulären nerventransplantation im tierexperiment (Regarding the sutureless interfascicular nerve transplantation in animal experiments). *Wien. Med. Wochenschr.* **1972**, *122*, 517–523.
11. Odar, J. Die Physiologische Gewebeklebung mit Fibrin—Prinzip, Einsatzgebiete, Klebetechniken, Kosten und Nutzen (Physiological Gluing of Tissues with Fibrin—Principle, Indications, Gluing Techniques, Costs and Benefits). In *Techniken und Methoden der Modernen Medizin (Techniques and Methods of Modern Medicine)*; Odar, J., Ed.; Steinkopff Verlag: Darmstadt, 1990; 75–84.
12. Jackson, M.R.; MacPhee, M.J.; Drohan, W.N.; Alving, B.M. Fibrin sealant: Current and potential clinical applications. *Blood Coagul. Fibrinolysis* **1996**, *7*, 737–746.
13. Blomback, B. Fibrinogen and fibrin—Proteins with complex roles in hemostasis and thrombosis. *Thromb. Res.* **1996**, *83* (1), 1–75.
14. Mosesson, M.W.; DiOrio, J.P.; Muller, M.F.; Shainoff, J.R.; Siebenlist, K.R.; Amrani, D.L.; Homandberg, G.A.; Soria, J.; Soria, C.; Samama, M. Studies on the ultrastructure of fibrin lacking fibrinopeptide B (β -FIBRIN). *Blood* **1987**, *69* (4), 1073–1081.
15. Ferry, J.D.; Morrison, P.R. Preparation and properties of serum and plasma proteins VIII. The conversion of human fibrinogen to fibrin under various conditions. *J. Am. Chem. Soc.* **1947**, *69*, 388–400.
16. Weisel, J.W. Fibrin assembly. Lateral aggregation and the role of the two pairs of fibrinopeptides. *Biophys. J.* **1986**, *50*, 1079–1093.
17. Amrani, D.L.; DiOrio, J.P.; Delmotte, Y. Wound healing. Role of commercial fibrin sealants. *Ann. N.Y. Acad. Sci.* **2001**, *936*, 566–579.
18. Redl, H.; Schlag, G.; Dinges, H.P. Vergleich zweier fibrinkleber. Einfluß ionischer zusätze auf fibrinstruktur sowie morphologie und wachstum menschlicher fibroblasten (Comparison of two fibrin sealants: Effect of ionic additives on fibrin structure as well as on morphology and growth of human fibroblasts). *Medwelt* **1985**, *36*, 769–776.
19. Schlag, G.; Redl, H.; Turnher, M.; Dinges, H.P. The Importance of Fibrin in Wound Repair. In *Fibrin Sealant in Operative Medicine Volume 1 Otorhinolaryngology*; Schlag, G., Redl, H., Eds.; Springer-Verlag: Berlin, 1986; 3–12.
20. Khare, A.; Woo, L.; Mclean, A.; Stewart, J.E.; DiOrio, J.P.; Amrani, D.A.; Helgerson, S. Mechanical characterization of fibrin gels. *Blood Coagul. Fibrinolysis* **1998**, *9* (7), 105.
21. Roberts, W.W.; Lorand, L.; Mockros, L.F. Viscoelastic properties of fibrin clots. *Biorheology* **1973**, *10*, 29–42.
22. Redl, H.; Schlag, G.; Dinges, H.; Kuderna, H.; Seelich, T. Background and Methods of Fibrin Sealing. In *Biomaterials*; Winter, D., Gibbons, D.F., Plenk, H., Jr., Eds.; John Wiley and Sons: New York, 1982; 669–676.
23. Carr, M.E.; Gabriel, D.A.; McDonagh, J. Influence of factor XIII and fibronectin on fiber size and density in thrombin-induced fibrin gels. *J. Lab. Clin. Med.* **1987**, *110* (6), 747–752.
24. Martin, P. Wound healing—Aiming for perfect skin regeneration. *Science* **1997**, *276*, 75–81.
25. Clark, R.F. Fibrin and wound healing. *Ann. N.Y. Acad. Sci.* **2001**, *936*, 355–367.
26. Martinez, J.; Ferber, A.; Bach, T.L.; Yaen, C.H. Interaction of fibrin with VE-Cadherin. *Ann. N.Y. Acad. Sci.* **2001**, *936*, 386–405.
27. Stukova, S.M. Thrombin as a regulator of inflammation and reparative processes in tissues. *Biochemistry (Mosc.)* **2001**, *66*, 14–27.
28. Danen, E.H.J.; Yamada, K.M. Fibronectin, integrins, and growth control. *J. Cell. Physiol.* **2001**, *189*, 1–13.
29. Sahni, A.; Francis, C.W. Vascular endothelial growth factor binds to fibrinogen and fibrin and stimulates endothelial cell proliferation. *Blood* **2000**, *96*, 3772–3778.
30. Albes, J.M.; Klenzner, T.; Kotzerke, J.; Thiedemann, K.U.; Schafers, H.-J.; Borst, H.-G. Improvement of tracheal autograft revascularization by means of fibroblast growth factor. *Ann. Thorac. Surg.* **1994**, *57*, 444–449.
31. Hasimoto, J.; Kurosaka, M.; Yoshiya, S.; Hirohata, K. Meniscal repair using fibrin sealant and endothelial cell growth factor. *Am. J. Sports Med.* **1992**, *20*, 537–541.
32. MacPhee, M.J.; Singh, M.P.; Brady, R., Jr.; Akhyani, N.; Liau, G.; Lasa, C., Jr.; Hue, C.; Best, A.; Drohan, W. Fibrin Sealant: A Versatile Delivery Vehicle for Drugs and Biologics. In *Surgical Adhesives and Sealants Current Technology and Applications*; Sierra, D.H., Saltz, R., Eds.; Techomic Publishing Co.: Lancaster, PA, 1996; 109–120.
33. Schense, J.C.; Bloch, J.; Aebischer, P.; Hubbell, J.A. Enzymatic incorporation of bioactive peptides into fibrin matrices enhances neurite extension. *Nat. Biotechnol.* **2000**, *18*, 415–419.
34. Andree, C.; Voigt, M.; Wenger, A.; Erichsen, T.; Bittner, K.; Schaefer, D.; Walgenbach, K.J.; Borges, J.; Horch, R.E.; Eriksson, E.; Stark, G.B. Plasmid gene

- delivery to human keratinocytes through a fibrin-mediated transfection system. *Tissue Eng.* **2001**, *7*, 757–766.
35. Hunyadi, J.; Farkas, B.; Bertenyi, C.; Olah, J.; Dobozy, A. Keratinocyte grafting: A new means of transplantation for full-thickness wounds. *J. Dermatol. Surg. Oncol.* **1988**, *14*, 75–78.
 36. Horch, R.; Bannasch, H.; Kopp, J.; Andree, C.; Stark, G.B. Single-cell suspension of cultured human keratinocytes in fibrin-glue reconstitute the epidermis. *Cell Transplant* **1998**, *7*, 309–317.
 37. Pellegrini, G.; Ranno, R.; Stracuzzi, G.; Bondanza, S.; Guerra, L.; Zambruno, G.; Micali, G.; de Luca, M. The control of epidermal stem cells (holoclones) in the treatment of massive full-thickness burns with autologous keratinocytes cultured on fibrin. *Transplantation* **1999**, *68*, 868–879.
 38. Gorodetsky, R.; Clark, R.A.F.; An, J.; Gailit, J.; Levdansky, L.; Vexler, A.; Berman, E.; Marx, G. Fibrin microbeads (FMB) as biodegradable carriers for culturing cells and for accelerating wound healing. *J. Invest. Dermatol.* **1999**, *112*, 866–872.
 39. Wechselberger, G.; Schoeller, T.; Stenzl, A.; Ninkovic, M.; Lille, S.; Russell, R.C. Fibrin glue as delivery vehicle for autologous urothelial cell transplantation onto a prefabricated pouch. *J. Urol.* **1998**, *160*, 583–586.
 40. Bach, A.D.; Bannasch, H.; Galla, T.J.; Bittner, K.M.; Stark, G.B. Fibrin glue as matrix for cultured autologous urothelial cells in urethral reconstruction. *Tissue Eng.* **2001**, *7*, 45–53.
 41. Wechselberger, G.; Russell, R.C.; Neumeister, M.W.; Schoeller, T.; Piza-Katzer, H.; Rainer, C. Successful transplantation of three tissue-engineered cell types using capsule induction technique and fibrin glue as a delivery vehicle. *Plast. Reconstr. Surg.* **2002**, *110*, 123–129.
 42. van Susante, J.; Buma, P.; Schuman, L.; Homminga, G.; van den Berg, W.; Veth, R. Resurfacing potential of heterologous chondrocytes suspended in fibrin glue in large full-thickness defects of femoral articular cartilage: An experimental study in the goat. *Biomaterials* **1999**, *20*, 1167–1175.
 43. Ye, Q.; Zünd, G.; Benedikt, P.; Jockenhoevel, S.; Hoerstrup, S.P.; Sakyama, S.; Hubbell, J.A.; Turina, M. Fibrin gel as a three-dimensional matrix in cardiovascular tissue engineering. *Eur. J. Cardio-Thorac. Surg.* **2000**, *17*, 587–591.
 44. Tholpady, S.S.; Sclosser, R.; Spotnitz, W.; Ogle, R.C.; Lindsey, W.H. Repair of osseous facial critical-size defect using augmented fibrin sealant. *Laryngoscope* **1999**, *109*, 1585–1588.
 45. Daculsi, G.; Bagot d’Arc, M.; Cottreal, M. Calcium Phosphate Ceramics (Triosite®) in Otology: Evaluation in Animals and Applications in Humans. In *Fibrin Sealing in Surgical and Nonsurgical Fields, Wound Healing (1)*; Schlag, G., Redl, H., Eds.; Springer Verlag: Berlin, 1994; 121–126.
 46. Bagot d’Arc, M.; Corlieu, P.; Daculsi, G. Mastoid Cavity Filling for Bone Reconstruction with a Mixture of Fibrin Glue and Ceramic Granules. In *Fibrin Sealing in Surgical and Nonsurgical Fields, Neurosurgery, Ophthalmic Surgery, ENT (5)*; Schlag, G., Redl, H., Eds.; Springer Verlag: Berlin, 1994; 174–179.
 47. Hench, L.H.; Polak, J.M. Third-generation biomedical materials. *Science* **2002**, *295*, 1014–1017.

Fibrin Sealants

Marcus E. Carr, Jr.

*Central Virginia Center for Coagulation Disorders, Virginia Commonwealth University,
Richmond, Virginia, U.S.A.*

Erika J. Martin

*Department of Internal Medicine, Virginia Commonwealth University,
Richmond, Virginia, U.S.A.*

Heather Ambrose

*Coagulation Special Studies Laboratory, Central Virginia Center for Coagulation Disorders
Virginia Commonwealth University, Richmond, Virginia, U.S.A.*

INTRODUCTION

When injury occurs, the clotting process is triggered, and, within a few minutes, hemorrhage ceases. Fluid blood is converted to a solid plug by the production of the three-dimensional fibrin polymer matrix. The importance of this event has been recognized for greater than a century, and efforts to utilize exogenous fibrin to aid hemostasis in bleeding patients have been attempted for virtually the same length of time. Over the past 30 years, these efforts have resulted in products whose appropriate utilization appears to be of benefit in a variety of clinical and surgical settings. In addition, the importance of fibrin for appropriate wound healing and the potential to utilize this biological matrix to deliver drugs, growth factors, and even cellular components have been increasingly realized. This brief review will summarize the evolution of fibrin clinical products, the evidence for their efficacy, and the recent trends toward new applications.

BACKGROUND

Coagulation Pathway and Fibrin Formation

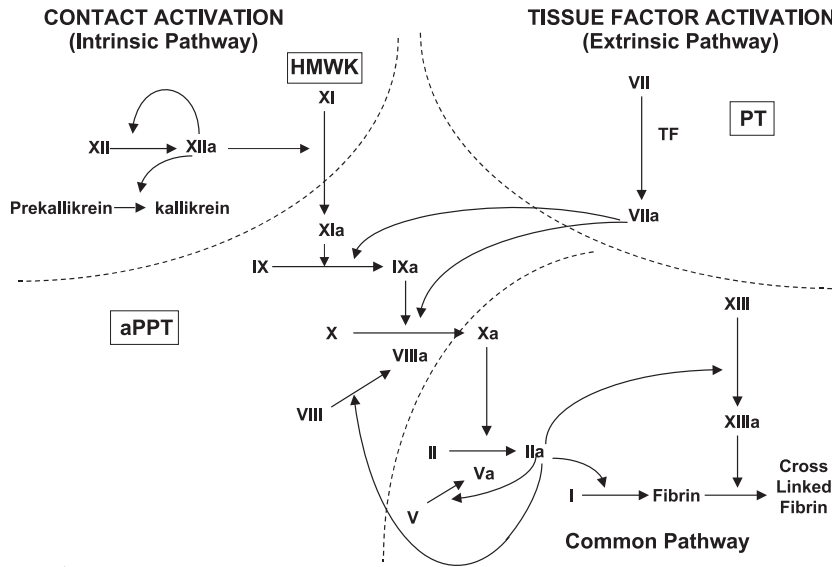
Early in the 20th century, Morowitz brought to an end the first epoch of coagulation research when he correctly postulated that the addition of material released from damaged tissue to blood would result in the formation of the enzyme thrombin.^[1] Once formed, thrombin cleaves fibrinogen which results in the spontaneous production of the fibrin clot. Prior to this hypothesis, parameters, such as changes in temperature and exposure to air, were offered as potential explanations for the mysterious solidification of blood upon its removal from the body. Dr. Morowitz's rudimentary pathway served as the framework for the vast amount of research that resulted in the discovery of multiple clotting factors and their placement in a

logical sequence that was subsequently described as the clotting cascade (Fig. 1A). This work continues, and recent modifications of the mechanism serve to emphasize the critical role played by platelets and the importance of the thrombin burst (Fig. 1B) for normal hemostasis.^[2]

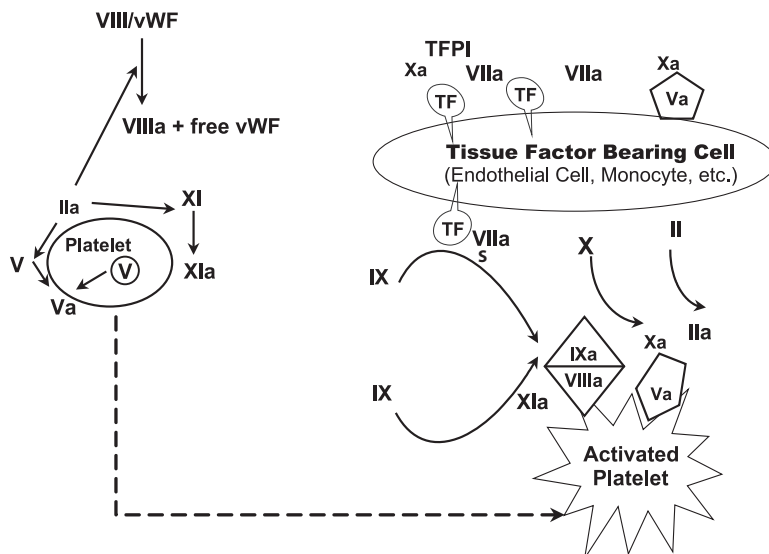
Once thrombin is formed, irreversible activation of platelets occurs. Fibrinogen is converted to fibrin monomer, and factor XIII is activated. Fibrinogen is composed of six extended polypeptide chains representing three distinct types— α , β , γ . Each of which is present as a pair ($\alpha_2\beta_2\gamma_2$). Fibrinogen's molecular shape is that of an extended, slightly bent rod containing two identical end nodules called D domains and one central nodule called E domain. Thrombin activates fibrinogen by cleaving two sets of peptides, fibrinopeptides A and B, from the amino terminal ends of the α and β chains. Due to the folding of the molecule, these peptides are removed from the central E domain. Upon removal of the fibrinopeptides, the new molecular species, which is termed fibrin monomer, spontaneously assembles into protofibrils and, subsequently, into fibers of the fibrin network (Fig. 2). The initial assembly is solely accomplished by hydrogen bonding. Such structures are inherently weak and subject to rapid removal by the fibrinolytic system. This situation is rectified by the introduction of covalent bonds between the assembled fibrin molecules through a process of transglutamination mediated by activated factor XIII. Factor XIII normally circulates as a pro-enzyme, but it is activated by thrombin during clotting (Fig. 2).

Development of Fibrin as a Clinical Agent

While our understanding of the coagulation apparatus continues to evolve, the potential for using fibrin to promote hemostasis in bleeding patients was recognized early on. In 1909, Bergel first reported use of



Panel A



Panel B

Fig. 1 Panel A: the clotting cascade emphasizing the intrinsic, extrinsic, and common pathways. Panel B: the new cascade emphasizing the critical roles of tissue-presenting cells and platelets in the initial activation and critical burst of thrombin production.

dried plasma as a source of fibrinogen and fibrin fleece as an aid to control hemorrhage in surgery.^[3] In 1915, Grey used fibrin patches to stop bleeding during cerebral surgery,^[4] and, in 1940, the adhesive properties of fibrin were exploited to reapproximate peripheral nerves.^[5] In 1944, the process of mixing fibrinogen with thrombin to prepare a fibrin adhesive to anchor skin grafts was described.^[6] For the next 30 years, the use of dry fibrin continued to be intermittently reported, but enthusiasm for the material was hindered by the variable quality of the product and the potential complications of viral exposure. Interest was renewed in the mid-1970s with the development of commercial processes for isolation of large quantities of virally inactivated fibrinogen. Subsequently, multiple products were developed and placed into clinical practice in Europe, Japan, and Canada (Table 1).

Fibrin's Desirable Properties

Fibrin adhesives, such as sealants, glues, etc., have multiple properties that make them desirable as hemostatic agents and tissue repair matrices. Since these fibrin products are of human origin, they are biocompatible and biodegradable. Since they are the normal injury repair matrix, they do not induce an inflammatory response, foreign body reaction, extensive fibrosis, or tissue necrosis. There is increasing evidence that fibrin sealants promote tissue growth and repair^[7] and may actually promote new vessel formation.^[8]

Commercially Available Products

Several companies have developed fibrin products for clinical use (Table 1). While these products have been



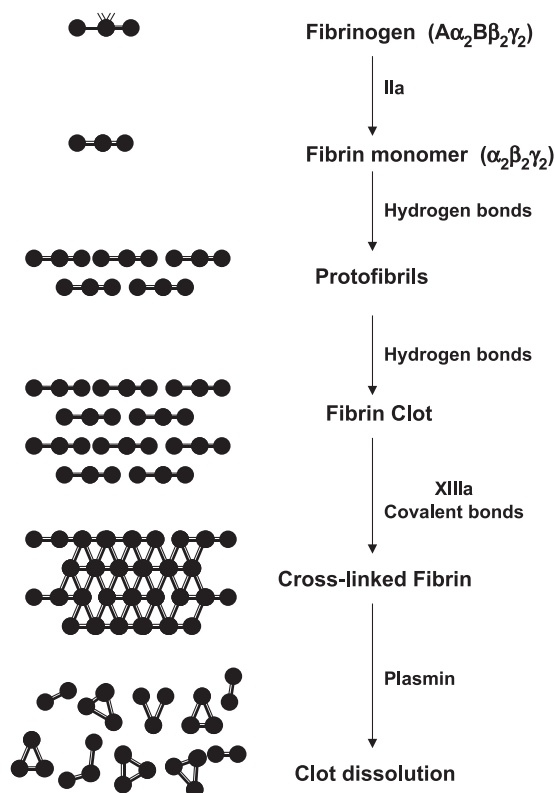


Fig. 2 Activation of fibrinogen and steps in the formation of the fibrin network.

available in Europe, Canada, and Japan for a number of years and in worldwide use for greater than a decade, their introduction in the United States occurred only recently. The delayed U.S. Federal Drug Administration (FDA) approval stemmed from concerns over viral safety and the FDA requirement that each component contained in the materials must add substantial benefit. In 1978, due to reported transmission of the hepatitis B virus, the FDA revoked the license to use fibrinogen concentrates in the United States. Since the majority of commercial fibrin adhesives started from human cryoprecipitate or Cohn Fraction I, concerns persisted despite the development of several effective viral inactivation treatments. The ever-increasing volume of literature regarding the use of fibrin products in multiple types of surgeries prompted the use of homemade fibrin sealants by many U.S. medical facilities.^[9-11] These locally produced materials were of variable quality and generally produced by combining cryoprecipitate with bovine thrombin and calcium—with or without the addition of an antifibrinolytic agent. U.S. fibrin glues, while effective in some clinical settings, carried several undesirable characteristics. First, the amount of fibrinogen and other proteins, such as factor XIII, fibronectin, etc., used to make the clots varied due to variations in cryoprecipitate used. Fibrinogen

concentration is a critical determinant in the structural properties of the formed fibrin. Second, the bovine thrombin typically used in U.S. fibrin was contaminated with bovine factor five. In some patients, exposure to bovine factor five produced antibodies that cross-reacted with human factor five, which led to acquired factor five deficiencies.^[12] Such deficiencies resulted in severe bleeding in some patients. Third, no simple quality control measures were typically employed to test the properties of the fibrin formed. Finally, although generally safe, cryoprecipitate from other humans continued to carry some risk of viral exposure. Virally inactivated human thrombin has become available through commercially available fibrin sealant kits; however, it is still not available as a stand-alone product. Whereas procedures for producing autologous fibrin glue that uses virally inactivated human thrombin have been developed,^[13] the pressure to proceed along these lines has recently decreased with FDA approval for the use of Tisseel[®] (Baxter-Immuno AG, United States of America) in the United States.

Production and Composition of Commercial Fibrin Adhesives

Components generally found in all fibrin adhesive products include fibrinogen, which is from concentrate; thrombin, which is initially bovine and later human, calcium; and a fibrinolytic inhibitor, which is mostly bovine aprotinin (Table 1). The fibrinogen is of human origin and initially isolated from frozen plasma. The first isolation procedure involves either thawing of the frozen plasma at 2°C (cryoprecipitate) or precipitation of plasma proteins with 10% ethanol (Cohn fraction I). The later process tends to yield materials enriched with fibronectin and factor XIII. This starting material is further purified by additional precipitation steps. Once purified and concentrated, the protein mixture undergoes viral inactivation with one of several processes including two-step vapor heat (at 60°C and 80°C), pasteurization, or solvent-detergent treatment. Viral inactivation is followed by additional purification and concentration steps, sterile filtration, and lyophilization. While the final concentrations of fibrinogen and factor XIII vary between products (Table 1), all products contain at least 70 milligrams of fibrinogen and 10 units of factor XIII per milliliter. Most products use bovine aprotinin as the antifibrinolytic, and all use virally inactivated human thrombin. Some products are packaged and supplied as vials of lyophilized protein concentrates. When reconstituted, one vial yields a solution of fibrinogen, factor XIII, and aprotinin. The other yields a thrombin plus calcium solution. The separate solutions are rapidly mixed through a variety of devices. One device yields a spray to cover

Table 1 Commercially available fibrin sealant products

Manufacturer	Trade name	Chemical composition			Viral inactivation		
		Fibrinogen (mg/ml)	Thrombin (IU/ml)	Factor XIII (U/ml)	Aprotinin (KIU/ml)	Fibrinogen (Human)	Thrombin (Human)
Baxter-Immuno AG (Austria)	Tisseel ^a	70–110	500	10–50	3000	TSVH ^b	TSVH
Hemacure (Canada)	Hemaseel APR ^c	75–115	500	None	3000	TSVH	TSVH
Aventis Behring (Germany)	Beriplast P	65–115	400–600	40–80	1000	P ^d	P
LFB-Lille (France)	Biocol	127	558	11	3000	SD ^e	SD
Kaketsuken Pharmaceutical (Japan)	Bolheal	80	250	75	1000	DH ^f	DH
Omrrix Bioharmaceuticals (Israel)	Quixil	60–100	1000	None	None ^g	SD	SD

^aUSFDA approved as hemostatic aid.^bTSVH two step vapor heat.^cAs Tisseel VH kit Baxter-Immuno AG USA.^dP-pasteurization.^eSD-solvent detergent.^fDH-dry heat.^gContains tranexamic acid (92 mg/ml).

large surface areas while a dual-syringe device allows production of a significant amount of clot directly on an affected area.^[14] Recently dried, preformed fibrin bandages have been developed for direct application in the setting of massive hemorrhage.^[15]

CLINICAL APPLICATIONS FOR FIBRIN SEALANTS

Problems with the Evidence for the Clinical Utility of Fibrin Sealants

While the literature on the use of fibrin sealants is extensive, it is of relatively poor quality. Most trials have been uncontrolled and have employed differing products that use a variety of protocols to assess effectiveness. These difficulties have resulted in a number of instances in which trials designed to address the same clinical question have resulted in absolutely opposite conclusions. Trials testing the effectiveness of fibrin sealant for reducing seroma formation in breast surgery found that fibrin sealant was effective^[16] and ineffective.^[17] Trials testing whether fibrin sealant could reduce fistula formation in pancreatic surgery found that it did^[18] and did not.^[19] Trials testing the ability of fibrin sealant to reduce bleeding in patients with hemophilia undergoing tooth extraction were both positive^[20] and negative.^[21] Finally, despite several series that indicated that fibrin sealants improved the take and adhesion of skin grafts, a controlled trial failed to demonstrate benefit.^[22] With these conflicting results in mind, the following sections briefly summarize results from the literature.

Use in Neurosurgery

The properties and ease of application of fibrin sealants have led to their use in several neurosurgical applications. They have been used as an adjunct to suturing and reduce cerebrospinal fluid (CSF) leaks. The local inflammatory effects seen with suturing may be avoided theoretically with the use of fibrin sealants. The technically difficult access to microvascular anastomoses during transnasal surgeries makes the role of fibrin sealants attractive for enhancement of hemostasis intraoperatively and prevention of post-operative hemorrhage. Unfortunately, the literature on fibrin usage in neurosurgery consists mainly of small series and case reports plus a few randomized clinical trials.

Al-Yamany described a series of 25 patients who underwent 29 transcortical approaches for the resection of intraventricular and/or paraventricular lesions. None of the 25 patients developed a clinically

symptomatic subdural fluid collection after the application of fibrin sealant at the time of the procedure. Additionally, none of the patients developed new seizure activity or experienced progression of a pre-existing seizure disorder.^[23] Lee reported 26 patients undergoing neurosurgery for various intracranial conditions, including brain tumors, lipomyelomeningoceles, arteriovenous malformation, and dural tear.^[24] A commercially available fibrin sealant formulation was applied in conjunction with Teflon, cotton, or muslin pieces to reinforce aneurysmal clippings in nine patients. No recurrent bleeding or adverse reaction was reported during a 25-month follow-up in the group receiving fibrin sealant.

The potential role of fibrin sealant as biomaterial and drug delivery systems in neurosurgery has been investigated in animal models. Iwakawa evaluated fibrin sealant as an intraspinal delivery system for application of glial cell line-derived neurotrophic factor (GDNF) for regeneration of cut dorsal root in adult rats. When compared with nerves treated with GDNF-free fibrin sealant, rats treated with GDNF embedded in fibrin that was directly applied to the cut dorsal root experienced enhanced nerve regeneration into the spinal cord. These results suggest a possible role for fibrin sealant as a therapeutic drug delivery system in neurosurgical patients.^[25]

Use in Oral Surgery

Patients who are on long-term anticoagulation therapy or have hemorrhagic disorders such as hemophilia have increased risk of hemorrhagic complications with oral surgery. For patients on long-term anticoagulation therapy, it must be temporarily discontinued to reduce risk of bleeding perioperatively. Such intervention may increase risk of thrombosis and depending on the initial indication for anticoagulant therapy, may be particularly hazardous. The use of fibrin sealant is appealing for such situations if the therapy cannot only reduce perioperative bleeding but also eliminates the need for the interruption of anticoagulant therapy.

Fibrin sealants may reduce the incidence of post-operative bleeding for patients with hemorrhagic disorders undergoing oral surgical procedures.^[20] Rokocz examined 80 patients with a history of hemorrhagic disorders, who were between the ages 3 of 75, that were undergoing dental extractions.^[20] One hundred thirty-five tooth extractions were performed without preventative factor replacement therapy. Complete initial local hemostasis was achieved with fibrin sealant application (Berioplast[®]) after extraction. Nine of 12 patients with severe hemophilia experienced secondary bleeding. In those patients with severe hemophilia,

improved hemostasis was achieved by increasing the concentration of aprotinin in the fibrin sealant compound along with the addition of swish and swallow tranexamic acid before and after extractions.^[20]

Fibrin sealant application may reduce the risk of postoperative hemorrhage for patients on chronic anticoagulation therapy undergoing oral surgery. Martinowitz reported no secondary bleeding in 40 patients on chronic anticoagulation therapy with preoperative INR values ranging from 2.5 to 4.5 who were treated with Beriplast[®] to attain postextraction hemostasis.^[26] Bodner reported similar results in 69 patients undergoing oral surgery who were on chronic anticoagulation with INR values ranging from 1.0–5.0.^[27] All patients were treated with fibrin sealant to achieve local hemostasis, and none of the patients interrupted anticoagulation therapy before the procedures. Minor postoperative bleeding occurred in only three (4.3%) patients.^[28]

Use in Endoscopy

Acute upper gastrointestinal (GI) bleeding occurs in more than 100 per 100,000 persons each year and can be associated with mortality rates approaching 15%. Subsequent bleeding episodes will occur in 5 to 30% of patients following the initial presentation.^[28] Endoscopy is an effective tool in visualizing the etiology of bleeding, and sclerotherapy can be used to stop active bleeding and prevent recurrence. Fibrin sealants seem an ideal therapy since they are composed of physiologic materials involved in coagulation and healing. In addition, fibrin sealants may be applied serially. This is a significant advantage over traditional agents that may lead to tissue necrosis and destruction with multiple applications.

Available literature indicated that fibrin sealants may be an effective therapy in achieving hemostasis and reducing risk of additional bleeding in patients with acute upper GI bleeding. Rutgeerts reported 805 patients with acute upper GI bleeding who were treated with endoscopy and local therapy with either polidocanol or fibrin sealant (Beriplast[®]).^[29] The trial was designed to detect the difference between a single injection with sclerosant polidocanol in the ulcer bed versus local therapy with a single application of fibrin sealant versus repeated application of fibrin sealant as prophylaxis once daily for five days after the initial bleeding episode. Repeated injections of polidocanol were not utilized due to the risk of tissue necrosis with subsequent injections. The group that received repeated applications of fibrin sealant experienced a 33.3% reduction in the risk of recurrent bleeding, which was a 40.8% reduction in treatment failure rates, and a lower incidence of clinically relevant bleeding

($p=0.011$).^[29] An additional subgroup analysis revealed a significantly lower risk of bleeding in patients with nonbleeding visible vessels who were treated with repeated applications of fibrin sealant. Although not reaching statistical significance, a trend toward lower risk of recurrent bleeding was seen in the group with spurting vessels who were treated with repeated applications of fibrin sealant.^[29] Zimmer et al. examined upper GI bleed secondary to esophageal varices and randomized patients to receive either fibrin sealant or polidocanol. Fibrin sealant therapy was superior with regards to rebleeding risk and overall mortality at 28 days follow-up.^[30]

While these studies indicated a positive benefit to using fibrin sealants in the setting of acute upper GI bleeding treated with endoscopy, other studies have shown conflicting data. In 2002, Lin et al. reported a trial involving 51 patients who were presenting with acute upper GI bleeding with either actively bleeding or nonbleeding visible vessels that was visualized with endoscopy. Patients were randomized to either sclerosis with endoscopic epinephrine injections or fibrin sealant application.^[31] The rate of rebleeding was significantly higher in the epinephrine group—56% versus 15% in the fibrin sealant group ($p=0.003$).^[31] The volume of blood transfusions, number of surgeries, length of hospital stay and mortality was not significantly different between the two groups. The trial concluded that fibrin sealant may achieve better initial hemostasis than epinephrine injections, but it does not change overall outcomes.^[31] Similarly, Heldwein examined the role of fibrin sealant for treatment of acute upper GI blood loss secondary to arterial bleeding.^[32] All patients were treated with either photocoagulation or fibrin sealant application, and all patients achieved initial hemostasis. However, there was no significant difference between the two groups with regards to incidence of rebleeding, ultimate hemostasis, or hospital mortality.^[32]

Use in Skin Grafting

Fibrin sealant therapy may have multiple benefits in the setting of skin grafting procedures. Blood loss during engrafting procedures in burn patients can be profound. Topical fibrin sealants may have an important role in minimizing hemorrhage during such procedures. These agents may improve engraftment in slightly infected wounds while the combination of fibrin sealant and sutures seal parenchymal surfaces better than suturing alone.^[33,34] When used alone, fibrin sealant resulted in a 95–100% successful skin engraftment in such difficult graft sites as eyelids and neck with no hematoma formation and a 50%

reduction in operative time.^[34] Future clinical applications in this setting are under investigation.

Use in Urologic Procedures

Few trials have examined the use of fibrin sealants in the urologic setting. In a 30-patient trial, postoperative application of fibrin sealant to the prostatic bed significantly reduced postoperative bleeding in patients undergoing transurethral resection of the prostate when compared to placebo.^[35] Although they remain largely experimental at this time, use of fibrin sealants for other urologic procedures, such as urethral anastomosis and closure of bladder fistulae, have been reported.^[36,37]

Use in Ophthalmology

Clinical applications of fibrin sealants in ophthalmology include repair of conjunctival wounds, sealing of lens perforation, eyelid surgery, lachrymal repair, and closing of corneal ulcers reducing the need for corneal implants. Fibrin sealants are appealing in this setting due to their performance in moist environments and the potential to avoid sutures.

Use in Abdominal Surgery

Despite good surgical technique, complications such as infection, abscess formation, and sepsis can arise due to bleeding and fluid leakage from suture sites and parenchymal tissue where sutures cannot be used. Fibrin sealant, when used in conjunction with sutures, staples, or collagen fleece, could decrease and prevent leaks as well as enhance healing after abdominal surgery. In a multicenter study of patients undergoing esophageal, gastric, colonic, or rectal resection, fibrin sealant was applied to closure sites in 37 patients. After two weeks, 30 patients (81.1%) showed no evidence of leakage upon radiological examination. Minor or major suture failure was recorded on seven anastomoses.^[38]

Hepatic and Pancreatic Surgery

Fibrin sealant provides a potentially effective therapy to achieve hemostasis of parenchymatous organs in which sutures alone may not control and may even aggravate bleeding and fluid leakage.^[39] Their properties make fibrin sealants suitable adjunctive hemostatic agents in liver resection, biopsy, and transplantation. In a study of 77 patients who underwent elective liver resection for benign and malignant lesions, Noun et al. reported significantly lower mean fluid drainage and bilirubin concentrations during

the three postoperative days in the group treated with a single dose of fibrin sealant. Additionally, the abdominal closure site was judged as dry in 97% of patients treated with fibrin sealant versus 81% not treated with fibrin sealant ($p=0.016$). There was no difference in postoperative morbidity and mortality between the two groups.^[40] In a study in which fibrin sealant was sprayed onto the cut liver surfaces of eight patients with hepatic cancer and cirrhosis who underwent hepatectomy, no postoperative bile leakage and abscess formation was noted.^[41] In laparoscopic liver biopsies, postpuncture bleeding was prevented by injecting fibrin sealant into or near the biopsy site.^[42,43]

Pancreatic fistulas are most commonly derived as complications of elective surgical procedures on the pancreas and as sequelae of pancreatitis or pancreatic trauma. They are also a leading cause of morbidity after pancreaticoduodenectomy.^[44] The application of fibrin sealant during pancreatic surgery has been used with varying degrees of success in preventing complications associated with fluid leakage, tissue necrosis, and fistula formation. In one study, Suzuki et al. reported a 25% decrease in incidence of pancreatic fistula formation after applying fibrin sealant to seal the pancreatic stump.^[45] However, in another study by D'Andrea et al. of 97 patients undergoing pancreatic surgery, no statistically significant differences in the incidence of fistula formation was seen between the control group and the group treated with fibrin sealant.^[46] Although differences between groups were not statistically significant, the incidence of pancreatic fistula was significantly lower than in the first study. Differences in success rates in these studies may be due to differences in fibrin sealant formulation, techniques of fibrin application, and/or differences in surgical technique.

Cardiovascular Surgery

Fibrin sealants have been used in conjunction with sutures in a wide range of cardiovascular procedures to limit perioperative bleeding and postoperative blood loss. Fibrin sealants have been used to promote hemostasis after femoral artery catheterization and in procedures such as vascular grafts, valve replacements, repair of septal defects, repair of vessel/heart chamber ruptures, and prosthetic implantations. Fibrin sealant as an adjunct to sutures has been used with a success rate of at least 94% in cardiopulmonary bypass procedures.^[47,48] In 1982, Borst et al. reported a trial of 340 patients undergoing extracorporeal circulation in which fibrin sealant was applied as an adjunct to sutures. Fibrin-presealed fabric was inserted in 60 heparinized patients. Perfect sealing was reported in 95% of these cases.^[47] Jackson et al. reported a case in which fibrin sealant was used as an adjunct to sutures

in a patient with heparin-induced thrombocytopenia that was anticoagulated with heparinoid. Fibrin sealant was applied at the mediastinal wound surfaces and the anastomosis graft site, and complete hemostasis was achieved at these sites.^[49] In a multicenter study of 333 patients undergoing re sternotomy or reoperation after cardiac surgery, Rousou et al. reported a 92.6% success rate for fibrin sealant (Tisseel[®]) in controlling bleeding within five minutes versus a success rate of 12.4% for other conventional topical hemostatic agents. Additionally, fibrin sealant also controlled 82% of bleeding episodes that were initially treated with other topical agents.^[50] In 1991, Seguin et al. evaluated the use of fibrin sealant in a study of three groups of patients ($n=42$) undergoing surgery to correct type A acute aortic dissection. Fibrin sealant was used in 18 patients to seal the aortic stumps and reinforce the suture sites. In the other two groups, intraluminal sutureless graft or a Dacron prosthesis was used to replace the ascending aorta. Fibrin sealant significantly ($p<0.05$) reduced perioperative blood loss as well as perioperative mortality.^[51] In 1994, Seguin et al. suggested the use of fibrin sealant in lieu of sutures for aortic valve repair in order to prevent bleeding from the aortic wall caused by suturing.^[52] In a study of 17 patients undergoing carotid endarterectomy, mean blood loss from the arteriotomy suture lines was reduced by 130 milliliters. Plus, the time to achieve hemostasis was reduced by 13.5 minutes in the group treated with fibrin sealant.^[53]

The results of experimental studies suggest that fibrin sealants may improve hemostasis after femoral catheterization. Ismail et al. and Kipshidze et al. have reported on the efficacy and safety of applying fibrin sealant percutaneously after cardiac catheterization.^[54,55] Both studies reported less bleeding from the access site and fewer complications related to the procedure. Fibrin sealants may also be an effective means to deliver antibiotics to reduce the risk of infection at vascular graft sites. In-vitro experiments and studies in animal models have demonstrated the potential of adding a variety of antimicrobial agents to fibrin sealant products. In a model system in which 5 mg/cm³ tobramycin was combined with a fibrin sealant product, a significant reduction in vascular graft infection and pseudoaneurysm formation was reported.^[56] When combined with fibrin sealant, dibekacin sulfate has also been reported to be effective as a means of preventing vascular infection without affecting the strength of the sealant.^[57]

Thoracic Surgery

Air leakage and hemorrhage leading to infections, pleural adhesions, fistula formation, and infections

continue to be common complications of thoracic surgery. These complications can be overcome by the use of fibrin sealants. Fibrin sealants have been successfully used in numerous thoracic procedures, including assisting sternal closure plus sealing bronchial stumps, tracheal anastomoses, and pulmonary lesions, after surgical dissection in pleurodesis/decortication, tumor resection, and lobectomy/pneumonectomy.

In 1993, Mouritzen et al. reported on the efficacy of fibrin sealant (Beriplast[®] P) in a study of 114 patients undergoing pulmonary resections and pneumonectomies. Fifty-nine patients received sutures alone, and 55 patients received fibrin sealant as adjunct to sutures. Decreased air leakage, fewer postoperative complications, improved airway tolerance pressure testing, and a reduction in hospital stay were reported in the patients treated with fibrin sealant after suturing.^[58] Two newer methods of applying fibrin sealant (Beriplast[®] P) in thoracic surgery have shown better adhesive properties than conventional methods. One method consists of rubbing thrombin solution in the lung tissue before applying fibrinogen. A second method involves soaking fibrinogen into an absorbable mesh and subsequently applying it to the lung. Both methods have shown satisfactory results. The second method, however, had higher adhesive properties and a shorter duration of air leaks in a 35-patient clinical trial.^[59] In another proposed application, applying large amounts of diluted fibrin sealant was no more effective than chemical agents in the treatment for intractable pneumothorax.^[60]

FUTURE DIRECTIONS

While the hemostatic properties of fibrin will undoubtedly continue to drive most applications, the potential to use fibrin as a delivery system for drugs and growth factors has been postulated and is being actively employed. In this setting, the fibrin network is implanted to deliver high drug concentrations directly to the area of interest. The potential to enhance blood vessel formation in areas of ischemic risk by implanting fibrin networks enriched with angiogenesis factor is but one example. Equally intriguing is the possibility of accomplishing gene transfer by the implantation of genetically altered cells embedded and growing in fibrin structures. These possibilities will be facilitated by new techniques for producing bioengineered fibrin matrices with predefined structural characteristics through processes such as electrospinning. The primary limiting step at the moment is the supply of absolutely safe starting materials. Fibrinogen has been cloned, and, once it can be economically produced in large enough quantities, the potential applications for fibrin matrices appear compelling.



CONCLUSION

Fibrin sealants are used worldwide as adjuvants to control hemostasis. Although the products vary somewhat in their composition, they all contain high concentrations of fibrinogen and thrombin. Fibrin sealants are biodegradable and tissue-compatible. They also do not appear to cause significant inflammatory response. Since all current products are made from human plasma, a major concern regarding their utilization continues to be the potential for transmission of disease-causing materials. These concerns may be met in the short term by development of autologous products and/or better methods of pathogen inactivation. In the long term, utilization of recombinant products could solve the problem. Since fibrin is the normal three-dimensional tissue repair matrix, fibrin sealants are prime targets for development as implantable drug and tissue delivery systems.

ARTICLES OF FURTHER INTEREST

Bioadhesion; Biological Adhesives from Nature; Burn Dressing; Fibrin; Nanoscale Neurosurgery; Thrombosis

REFERENCES

- Morawitz, P. Die Chemie der Blutgerinnung (The chemistry of blood coagulation). *Erg. Physiol.* **1905**, *4*, 307–422.
- Monroe, D.M.; Hoffman, M.; Roberts, H.R. Platelets and thrombin generation. *Arterioscler. Thromb. Vasc. Biol.* **2002**, *22*, 1381–1389.
- Bergel, S. Über Wirkungen des Fibrins. *Dtsch. Wochenschr.* **1909**, *35*, 633–665.
- Grey, E.G. Fibrin as a hemostatic in cerebral surgery. *Surg. Gynecol. Obstet.* **1915**, *21*, 452–454.
- Seddon, H.J.; Medawar, P.B. Fibrin sutures of human nerves. *Lancet* **1942**, *ii*, 87–92.
- Cronkite, E.P.; Lozner, E.L.; Deaver, J. Use of thrombin and fibrinogen in skin grafting. *JAMA* **1944**, *124*, 976–978.
- Kram, H.B.; Nathan, R.C.; Mackabee, J.R.; Klein, S.R.; Shoemaker, W.C. Clinical use of nonautologous fibrin glue. *Am. Surg.* **1988**, *54*, 570–573.
- Dvorak, J.F.; Harvey, V.S.; Estrella, P.; Brown, L.F.; McDonagh, J.; Dvorak, A.M. Fibrin containing gels induce angiogenesis. Implications for tumor stroma generation and wound healing. *Vox Sang.* **1987**, *57*, 673–686.
- Spotnitz, W.D.; Mintz, P.D.; Avery, N.; Bithell, T.C.; Kaul, S.; Nolan, S.P. Fibrin glue from stored human plasma. An inexpensive and efficient method for local blood bank preparation. *Am. Surg.* **1987**, *53*, 460–462.
- Martinowitz, U.; Saltz, R. Fibrin sealant. *Curr. Opin. Hematol.* **1996**, *3*, 395–402.
- Kjaergard, H.K.; Trumbull, H.R. Vivostat system autologous fibrin sealant: Preliminary study in coronary bypass grafting. *Ann. Thorac. Surg.* **1998**, *66*, 482–486.
- Nichols, W. Adverse Antibody-Mediated Reactions to Topical Bovine Thrombin and Fibrin Glue. In *Symposium on Fibrin Sealant: Characteristics and Clinical Use*; Uniformed Services University of the Health Sciences: Bethesda, 1994; 5–10.
- Kevy, S.V.; Jacobson, M.S.; Blasetti, L. An Automated Device for the Rapid Preparation of Autologous Fibrin/Platelet Gel, an Effective, Convenient, Bioactive Wound Sealant. *Cambridge Symposium on Tissue Sealants: Current Practice, Future Uses*, La Jolla, 1996.
- Redl, H.; Schlag, G.; Dinges, H.P. Methods of fibrin seal application. *Thorac. Cardiovasc. Surg.* **1973**, *30*, 223–227.
- Holcomb, J.B.; Pusateri, A.E.; Hess, J.R.; Hetz, S.P.; Harris, R.A.; Tock, B.B.; Drohan, W.N.; MacPhee, M.J. Implications of new dry fibrin sealant technology for trauma surgery. *Surg. Clin. North Am.* **1997**, *77*, 943–952.
- Moore, M.M.; Nguyen, D.H.D.; Spotnitz, W.D. Fibrin sealant reduces serous drainage and allows for earlier drain removal after axillary dissection: A randomized prospective trial. *Am. Surg.* **1997**, *63*, 97–102.
- Vaxman, F.; Kolbe, A.; Stricher, F.; Zund, D.; Volkmar, P.; Gros, D.; Grenier, J.F. Does fibrin glue improve drainage after axillary lymph node dissection? *Eur. Surg. Res.* **1995**, *27*, 346–352.
- Suzuki, Y.; Kuroda, Y.; Morita, A.; Fujimo, Y.; Tanoika, Y.; Kawamura, T.; Saiton, V. Fibrin glue sealing for the prevention of pancreatic fistulas following distal pancreatectomy. *Arch. Surg.* **1995**, *130*, 952–955.
- D'andrea, A.A.; Costantino, V.; Sperti, C.; Pedrazzoli, S. Human fibrin sealant in pancreatic surgery: A prospective randomized study. *Ital. J. Gastroenterol.* **1994**, *26*, 283–286.
- Rakocz, M.; Mazar, A.; Varon, D.; Spierer, S.; Blinder, D.; Martinowitz, U. Dental extractions in patients with bleeding disorders: The use of fibrin glue. *Oral Surg. Oral Med. Oral Pathol.* **1993**, *75*, 280–282.
- Hoots, K.; McLeod, J.; E2qqers, E.; Berg, J.; Lee, M.; Gomperts, E. Pilot study to evaluate efficacy of fibrin sealant(human) on hemostasis in hemophiliacs undergoing tooth extraction (abst.). *Blood* **1993**, *82*, 598.
- Achauer, B.M.; Miller, S.R.; Lee, T.E. The hemostatic effect of fibrin glue on graft donor sites. *J. Burn Care Rehabil.* **1994**, *15*, 24–28.
- Al-Yamany, M.; Del Maestro, R.F. Prevention of subdural fluid collections following transcortical intraventricular and/or paraventricular procedures by using fibrin adhesive. *J. Neurosurg.* **2000**, *92*, 406–412.
- Lee, K.-C.; Park, S.-K.; Lee, K.-S. Neurosurgical applications of fibrin adhesive. *Yonsei Med. J.* **1991**, *32*, 53–57.
- Iwakawa, M.; Mizoi, K.; Tessler, A.; Itoh, Y. Intraspinal implants of fibrin glue containing glial cell line-derived neurotrophic factor promote dorsal root

- regeneration into spinal cord. *Neurorehabil. Neural Repair* **2001**, *15*, 173–182.
26. Martinowitz, U.; Mazar, A.L.; Taicher, S.; Varon, D.; Gitel, S.N.; Ramot, B.; Rakocz, M. Dental extraction for patients on oral anticoagulant therapy. *Oral Surg. Oral Med. Oral Pathol.* **1990**, *70*, 274–277.
 27. Bodner, L.; Weinstein, J.M.; Baumgarten, A.K. Efficacy of fibrin sealant in patients on various levels of oral anticoagulant undergoing oral surgery. *Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endo.* **1998**, *86*, 421–424.
 28. Longstreth, G.F. Epidemiology of hospitalization for acute upper gastrointestinal hemorrhage: A population-based study. *Am. J. Gastroenterol.* **1995**, *90*, 206–210.
 29. Rutgeerts, P.; Rauws, E.; Wara, P.; Swain, P.; Hoos, A.; Solleder, E.; Halttunen, J.; Dobrilla, G.; Richter, G.; Prassler, R. Randomised trial of single and repeated fibrin glue compared with injection of polidocanol in treatment of bleeding peptic ulcer. *Lancet* **1997**, *350*, 692–696.
 30. Zimmer, T.; Rucktaschel, F.; Stolzel, U.; Liehr, R.M.; Schuppan, D.; Stallmach, A.; Zeitz, M.; Weber, E.; Riecken, E.O. Endoscopic sclerotherapy with fibrin glue as compared with polidocanol to prevent early esophageal variceal rebleeding. *J. Hepatol.* **1998**, *28*, 292–297.
 31. Lin, H.J.; Hsieh, Y.H.; Tseng, G.Y.; Perng, C.L.; Chang, F.Y.; Lee, S.D. Endoscopic injection with fibrin sealant versus epinephrine for arrest of peptic ulcer bleeding: A randomized, comparative trial. *J. Clin. Gastroenterol.* **2002**, *35*, 218–221.
 32. Heldwein, W.; Avenhaus, W.; Schonekas, H.; Kaess, H.; Muller-Lissner, S.; Hasford, B.; Hasford, J. Injection of fibrin tissue adhesive versus laser photocoagulation in the treatment of high-risk bleeding peptic ulcers: A controlled randomized study. *Endoscopy* **1996**, *28*, 756–760.
 33. Dunn, C.J.; Goa, K.L. Fibrin sealant: A review of its use in surgery and endoscopy. *Drugs* **1999**, *58*, 863–886.
 34. Lilius, P. Fibrin adhesive: Its use in selected skin grafting. Practical note. *Scand. J. Plast. Reconstr. Surg. Hand Surg.* **1987**, *21*, 245–248.
 35. Luke, M.; Kvist, E.; Andersen, F.; Hjortrup, A. Reduction of post-operative bleeding after transurethral resection of the prostate by local instillation of fibrin adhesive (Beriplast). *Br. J. Urol.* **1986**, *58*, 672–675.
 36. McKay, T.C.; Albala, D.M.; Gehrin, B.E.; Castelli, M. Laparoscopic ureteral reanastomosis using fibrin glue. *J. Urol.* **1994**, *152*, 1637–1640.
 37. Bach, A.D.; Bannasch, H.; Galla, T.J.; Bittner, K.M.; Stark, G.B. Fibrin glue as matrix for cultured autologous urothelial cells in urethral reconstruction. *Tissue Eng.* **2001**, *7*, 45–53.
 38. Muto, T.; Ishibiki, K.; Hatakeyama, K.; et al. Efficacy and safety of B1 91.021.3 in patients receiving anastomotic surgery of the digestive tract: A multicenter study. *Shinyaku To Rinsho* **1994**, *43*, 2274–2283.
 39. Spotnitz, W.D.; Falstrom, J.K.; Rodeheaver, G.T. The role of sutures and fibrin sealant in wound healing. *Surg. Clin. North Am.* **1997**, *77*, 651–669.
 40. Noun, R.; Elias, D.; Balladur, P.; Bismuth, H.; Parc, R.; Lasser, P.; Belghiti, J. Fibrin glue effectiveness and tolerance after elective liver resection: A randomized trial. *Hepatogastroenterology* **1996**, *43*, 221–224.
 41. Wakasugi, J.; Shimada, H. Application of fibrin sealant in liver surgery. *Biomed. Prog.* **1994**, *7*, 33–35.
 42. Ishitani, M.B.; McGahren, E.D.; Sibley, D.A.; Spotnitz, W.D.; Rodgers, B.M. Laparoscopically applied fibrin glue in experimental liver trauma. *J. Pediatr. Surg.* **1989**, *24*, 867–871.
 43. Thiele, H.; Berg, P.L.; Frick, B.; Kalk, J.F. Fibrin-glue injection—a new method of haemostasis after laparoscopic liver biopsy. *Dtsch. Med. Wochenschr.* **1989**, *114* (31–32), 1196–1198.
 44. Pessaux, P.; Tuech, J.J.; Arnaud, J.P. Prevention of pancreatic fistulas after surgical resection. A decade of clinical trials. *Presse Med.* **2001**, *30*, 1359–1363.
 45. Suzuki, Y.; Kuroda, Y.; Morita, A.; Fujino, Y.; Tanioka, Y.; Kawamuro, T.; Saitoh, Y. Fibrin glue sealing for the prevention of pancreatic fistulas following distal pancreatectomy. *Arch. Surg.* **1995**, *130*, 952–955.
 46. D'Andrea, A.A.; Constantino, V.; Sperti, C.; Pedrazzoli, S. Human fibrin sealant in pancreatic surgery: Is it useful in preventing fistulas? A prospective randomized study. *Ital. J. Gastroenterol.* **1994**, *26*, 283–286.
 47. Borst, H.G.; Haverich, A.; Walterbush, G.; Maatz, W. Fibrin adhesive: An important hemostatic adjunct in cardiovascular operations. *J. Thorac. Cardiovasc. Surg.* **1982**, *84*, 548–553.
 48. Matthew, T.L.; Spotnitz, W.D.; Kron, I.L.; Daniel, T.M.; Tribble, C.G.; Nolan, S.P. Four years' experience with fibrin sealant in thoracic and cardiovascular surgery. *Ann. Thorac. Surg.* **1990**, *50*, 40–44.
 49. Jackson, M.R.; Danby, C.A.; Alving, B.M. Heparinoid anticoagulation and topical fibrin sealant in heparin induced thrombocytopenia. *Ann. Thorac. Surg.* **1997**, *64*, 1815–1817.
 50. Rousou, J.; Levitsky, S.; Gonzalez-Lavin, L.; Cosgrove, D.; Magilligan, D.; Weldon, C.; Hiebert, C.; Hess, P.; Joyce, L.; Bergsland, J.; et al. Randomized clinical trial of fibrin sealant in patients undergoing re sternotomy or reoperation after cardiac operations. A multicenter study. *J. Thorac. Cardiovasc. Surg.* **1989**, *97*, 194–203.
 51. Seguin, J.R.; Frapier, J.-M.; Colson, P.; Chaptal, P.A. Fibrin sealant improves surgical results of type A acute aortic dissections. *Ann. Thorac. Surg.* **1991**, *52*, 745–749.
 52. Seguin, J.R.; Picard, E.; Frapier, J.M.; Chaptal, P.A. Aortic valve repair with fibrin glue for type A acute aortic dissection. *Ann. Thorac. Surg.* **1994**, *58*, 304–307.
 53. Milne, A.A.; Murphy, W.G.; Reading, S.J.; Ruckley, C.V. Fibrin reduces suture line bleeding during carotid endarterectomy: A randomised trial. *Eur. J. Vasc. Endovasc. Surg.* **1995**, *10*, 91–94.
 54. Ismail, S.; Combs, M.J.; Goodman, N.C.; Teotia, S.S.; Teates, C.D.; Abbott, R.D.; Fechner, R.E.; Nolan, S.P.; Powers, E.R.; Spotnitz, W.D. Reduction of femoral arterial bleeding post catheterization using percutaneous application of fibrin sealant. *Catheter. Cardiovasc. Diagn.* **1995**, *34*, 88–95.

55. Kipshidze, N.; Ferguson, J.J., III; Macris, M.P.; Clubb, F.; Cloy, M.; Horn, J.; Sperl, F.; Sahota, H.; Keelan, M.H., Jr.; Nikolaychik, V.V. Percutaneous application of fibrin sealant to achieve hemostasis following arterial catheterization. *J. Invasive Cardiol.* **1998**, *10*, 133–141.
56. Ney, A.L.; Kelly, P.H.; Tsukayama, D.T.; Bubrick, M.P. Fibrin glue-antibiotic suspension in the prevention of prosthetic graft infection. *J. Trauma* **1990**, *30*, 1000–1006.
57. Tanemoto, K.; Fujinami, H. Experimental study on bacterial colonization of fibrin glue and its prevention. *Clin. Ther.* **1994**, *16*, 1016–1027.
58. Mouritzen, C.; Dromer, M.; Keinecke, H.-O. The effect of fibrin glueing to seal bronchial and alveolar leakages after pulmonary resections: A randomised trial. *Eur. J. Cardiothorac. Surg.* **1993**, *7*, 75–80.
59. Morikawa, T.; Katoh, H. Improved techniques of applying fibrin glue in lung surgery. *Eur. Surg. Res.* **1999**, *31*, 180–186.
60. Kinoshita, T.; Miyoshi, S.; Katoh, M.; Yoshimasu, T.; Juri, M.; Maebeya, S.; Naito, Y. Intrapleural administration of a large amount of diluted fibrin glue for intractable pneumothorax. *Chest* **2001**, *119*, 671–672.

Finite Element Analysis

Jennifer S. Wayne

Biomedical Engineering, Virginia Commonwealth University, and Orthopaedic Research Laboratory, Virginia Commonwealth University, Richmond, Virginia, U.S.A.



INTRODUCTION

Finite element analysis (FEA), also known as the finite element method, is a numerical solution technique that allows users to quantify how a field variable can vary in space within a particular geometry. These field variables in engineering problems can take the form of displacement in stress analyses, temperature in heat transfer/thermal analyses, velocity/pressure in fluid mechanics, and current in electromagnetics, for example. Problems can be time-independent or vary in time as well as in space. Coupling of more than one field variable into a finite element analysis can also be performed to provide answers to more complex problems.

BACKGROUND

Engineers, mathematicians, and physicists have all contributed to what we now know as FEA.^[1] The 1950s and 1960s saw some of the first applications of FEA in engineering. It was originally applied by the aeronautical industry to study aircraft wing behavior. Since then, it has provided information on such topics as stresses in structures (roofs, bridges, concrete, spacecraft, dams, foundations, etc), heat flow in rocket nozzles and internal combustion engines, temperature distribution in nuclear reactor components, and movement of fluid in containers and aquifers.^[2] In the bioengineering field, one might therefore envision applying FEA to understand the stresses that bones are subjected to in the body and the effect on the bone after total joint arthroplasty (such as implanting a total hip replacement); heat transfer and temperature patterns during heating and cooling (including cryopreservation) of different cells and tissues (skin, brain, liver, breast); or blood flow in normal arterial vessels and changes that occur to the flow in the presence of an atherosclerotic plaque. Thus, FEA has a broad application to a number of engineering problems.

One of the major reasons for the wide use of FEA is its versatility. There are virtually no restrictions on the geometry, boundary conditions, applied loading, or material behavior of the object under study.^[3] Different objects of different sizes and with different material

properties can be combined to analyze the behavior of an entire structure. Thus, FEA provides solutions to problems that could not be solved otherwise. The finite element structure also physically resembles the original structure rather than being an oversimplified shape. As FEA usefulness has caught on, the size of analyses has increased and continues to do so, thus demanding computer implementation of FEA procedures. It is fortunate that computer efficiency, power, memory, and storage capacity have dramatically improved over the last few decades and now allow complex finite element problems to be solved not only on supercomputers and mainframes but even on desktop computers. Hundreds of FEA software packages are commercially available (although only a few dozen are extensively used), making application of the technique even more widespread.

This article will focus on the steps involved in completing a finite element study and conclude with examples of FEA applications to bioengineering problems.

OUTLINE OF A FINITE ELEMENT PROBLEM

A finite element analysis progresses through three main stages: preprocessing, analysis, and postprocessing (Fig. 1). Each stage involves many substages, each of which are described in more detail later in this article. Any FEA begins with representing the physical problem by a mathematical model or idealization. This means essentially that the geometry of the physical problem, the properties of the material in the geometry, and the environment around the geometry (boundary conditions, applied loads) must be expressed mathematically.

PREPROCESSING

Geometric Information

This stage begins with the decision of whether a 1-D, 2-D, or 3-D representation will accurately reflect the

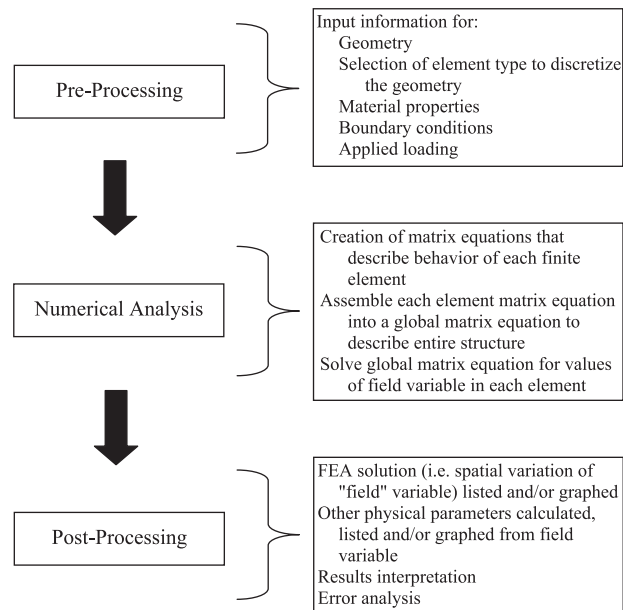


Fig. 1 General steps followed in a finite element analysis.

physics of the problem. Coordinate values are assigned to the geometry of the problem that serve as the initial input into the FEA software.

Discretization

A patchwork of finite elements that interconnect to form the geometry is then created (Fig. 2). The process of generating this finite element mesh is called discretization. If the space is one-dimensional, the elements may be bar elements to form trusses or beam elements to form frames, for example. If the space is two-dimensional, triangles or quadrilateral elements are most commonly used to discretize the space. If the space is three-dimensional, tetrahedral or brick elements may be implemented. Numerous types of elements (shell elements, plate elements, planar elements, 3-D solid elements, etc) often exist within various FEA software packages to lend versatility. It is also conceivable to discretize a geometry by combining two or more element types in one, two, or three dimensions.

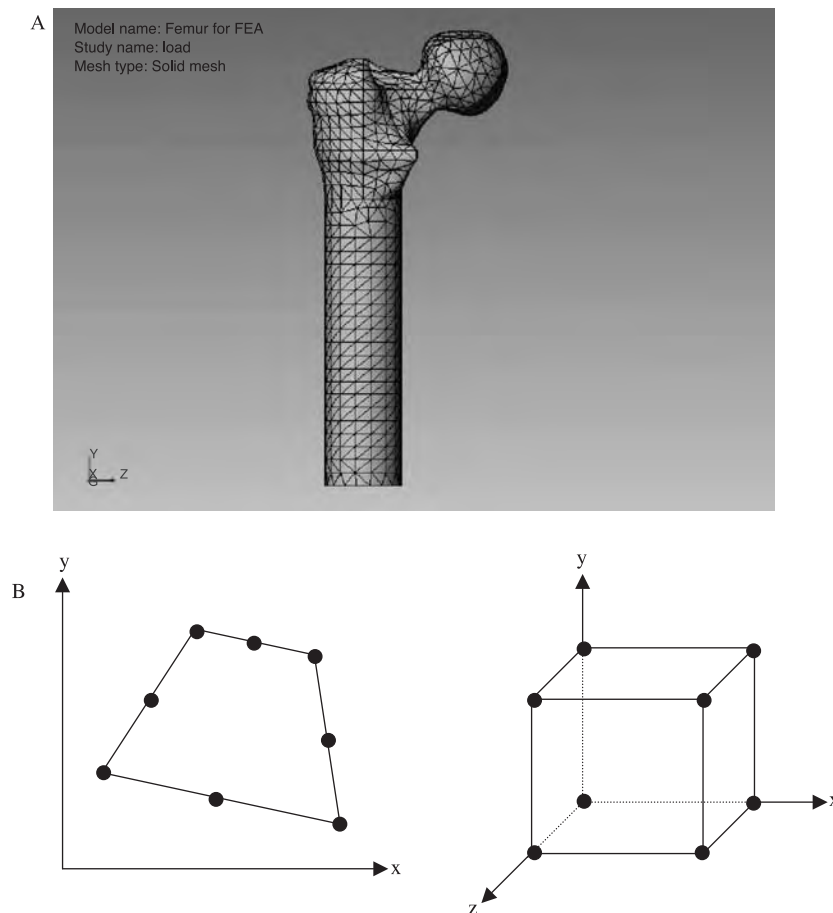


Fig. 2 A) The concept of discretization, or creating a patchwork of interconnected finite elements, is visually displayed on a solid model of the proximal femur or upper thigh bone. Model was created with SolidWorks 2001 Plus. B) Representation of two types of elements, 2-D (quadrilateral) and 3-D (brick) finite elements, defined with respect to a given coordinate system, which could discretize a given geometry. Black circles denote nodes, typically positioned at corners and/or edges, where the field variable of interest is specified.

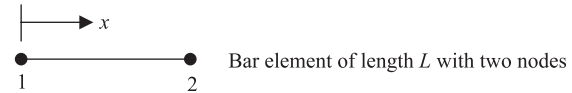
The number of elements in the mesh and their size are dependent on the degree of spatial variation of the field variable that is expected for the given problem. If large variations are expected in some regions of the geometry, smaller and more elements would be used in those regions to better approximate the field variable. A mesh can also be refined (more and smaller elements) after an analysis if the results are not deemed accurate enough. As the number of elements (and thus nodes) in a given mesh increase, the FEA results typically converge to the true solution of the problem. An infinite number of elements is not computationally efficient or feasible, however, and thus mesh refinement is completed until the level of convergence and accuracy is acceptable.

All elements in a finite element mesh are connected to one another at specific locations within the element. These locations are designated as nodes; nodes are positioned at the corners of the elements, but additional ones may be included along the edges of the element or within the element's interior. The field variable associated with the problem is calculated at the nodes during the analysis stage. The number of nodes assigned to a particular element determines how the field variable can vary within that element. For example, elements with only corner nodes permit only linear variation of the field variable in any given direction (x, y, z) because two points in a specific direction define a line. If an element has corner nodes as well as mid-side nodes, then a quadratic variation is possible for the field variable (three points can define a quadratic curve). Connecting the nodes from element to element creates a piecewise interpolation of the field variable over the entire geometry of the problem.

Mathematical equations are used to express the interpolation (i.e., variation) of the field variable in an element. These equations are called interpolation, or shape, functions. Shape functions depend on coordinate direction(s) (x, y, z) and can be linear equations or quadratic equations, depending on the number of nodes within the element. To do so, each node is assigned its own shape function with the special feature that the shape function must have value 1 when the coordinates of the node are inserted into the shape function's equation and 0 when the coordinates of any of the remaining nodes are inserted into the equation. The shape functions then permit the field variable to be calculated at any location (x, y, z coordinates) within the element, simply by knowing the values of the field variable at the nodes for the element. Multiplying the shape functions by the nodal values of the field variable provides the value of the field variable anywhere in the element.

A one-dimensional example of a bar element is most useful to describe shape functions. Let this bar element have nodes only at the corners (i.e., ends) of the bar and be numbered 1 and 2. This element thus has two shape

functions, designated as N_1 and N_2 . The subscript refers to the node number that the shape function represents. For an element with four nodes, there would be four shape functions: N_1, N_2, N_3 and N_4 . For the two-noded bar element in this one-dimensional example, variation in only one coordinate direction (x) is allowed.



The shape functions that would represent this element are: $N_1 = (L-x)/L$ and $N_2 = x/L$. Note that when the coordinates of node 1 ($x=0$) are inserted into the equations, $N_1 = 1$ and $N_2 = 0$. When the coordinates of node 2 ($x=L$) are inserted into the equation, $N_1 = 0$ and $N_2 = 1$. Let the symbol u represent the field variable of interest in this problem; then u_1 and u_2 are the values of the field variables at the nodes. The expression $u = N_1u_1 + N_2u_2$ permits the calculation of the field variable anywhere along the length of the element u by inserting the coordinates of the position of interest into the shape functions. This multiplication can be represented as the matrix calculation $u = [N] \{u\}^e$, where $[N]$ is a row vector of the two shape functions for the element, $[N] = [N_1 N_2]$ and $\{u\}^e$ is a column vector of nodal values for the element $\{u\}^e = \{u_1 u_2\}^T$. The value of the field variable at the nodes can also be recovered from this equation, e.g. $u = u_1$ when coordinates of node 1 ($x=0$) are used and $u = u_2$ when coordinates of node 2 ($x=L$) are used.

This same approach is applied to shape functions for any element type in two- and three-dimensional geometries. Now, each shape function would depend on x and y (in 2-D problems) or on $x, y,$ and z (in 3-D problems). Let the symbols u and v represent the value of the field variable in the x and y directions, respectively, in a 2-D element. Then the matrix expression allows the calculation of the field variable anywhere in the element. The symbols $u_i, v_i,$ and w_i represent the value of the field variable at node i in the $x, y,$ and z directions. The repeating dots drawn into the matrix equation indicate that the matrix of shape functions continues similarly for the number of nodes within the element. This matrix equation can be more easily expressed as $\{u\} = [N] \{u\}^e$.

$$\begin{Bmatrix} u \\ v \end{Bmatrix} = \begin{bmatrix} N_1 & 0 & N_2 & 0 & N_3 & 0 & \dots \\ 0 & N_1 & 0 & N_2 & 0 & N_3 & \dots \end{bmatrix} \begin{Bmatrix} u_1 \\ v_1 \\ u_2 \\ v_2 \\ u_3 \\ v_3 \\ \vdots \end{Bmatrix}$$

Assignment of Material Properties

Because each element is defined individually, each element is assigned its own material properties. These material properties are specific to the problem under study, e.g., in thermal problems, the material property of thermal conductivity is important whereas in structural problems, Young's Modulus is important. These properties govern how the field variable is related to the forces driving the problem through relationships called governing, or constitutive relationships. These relationships are typically represented as differential equations. For example in one-dimensional thermal problems, the heat conduction equation $q = -Ak(dT/dx)$, is dominant in thermal problems. Here, q = heat flow; A = cross-sectional area; k = thermal conductivity; T = temperature (the field variable); x = axial coordinate; and dT/dx = differential of temperature in the x -direction. The equation states that the amount of heat flow through a bar depends on the size of the bar (area), its thermal conductivity, and the temperature gradient. In one-dimensional structural analyses, the linear form of the stress-strain relationship, $\sigma = E(du/dx)$, is dominant. In these equations, σ = axial stress; E = Young's Modulus; u = displacement (the field variable); and du/dx = differential of displacement in the x -direction or strain in the x -direction. The equation states that the amount of stress applied to a bar is linearly proportional to the strain (or displacement gradient) via Young's Modulus. If the geometry of the problem is greater than one-dimensional, then the stress-strain relationship is represented by $\{\sigma\} = [E] \{e\}$, where the stress vector, $\{\sigma\}$, contains normal and shear stresses in the coordinate directions; the strain vector, $\{e\}$, contains normal and shear strains in the coordinate directions; and $[E]$ is a matrix containing the material properties (including Young's Modulus).

As presented earlier, the value of the field variable anywhere in an element is determined by the shape functions and the matrix of shape functions through the matrix equation, represented in 1-D as $u = [N] \{u\}^e$. In order to implement the governing relationship and material properties into an analysis, it is necessary to calculate the gradient (differential) of the field variable, du/dx . This is accomplished by taking the derivative of the shape functions with respect to x . In structural analyses, this process creates a new matrix called the strain-displacement matrix, $[B]$. In 1-D, $[B] = [dN_1/dx \ dN_2/dx]$. Thus, $du/dx = [B] \{u\}^e$. In 2-D, there are two normal strains ($e_{xx} = du/dx$ in the x direction and $e_{yy} = dv/dy$ in the y direction), and one shear strain ($\gamma_{xy} = du/dy + dv/dx$). Thus, the strain-displacement matrix $[B]$ provides calculation of strains from nodal values of the displacement through

$$\begin{Bmatrix} du/dx \\ dv/dy \\ du/dx + dv/dy \end{Bmatrix} = \begin{bmatrix} dN_1/dx & 0 & dN_2/dx & 0 & dN_3/dx & 0 & \dots \\ 0 & dN_1/dy & 0 & dN_2/dy & 0 & dN_3/dy & \dots \\ dN_1/dy & dN_1/dx & dN_2/dy & dN_2/dx & dN_3/dy & dN_3/dx & \dots \end{bmatrix} \times \begin{Bmatrix} u_1 \\ v_1 \\ u_2 \\ v_2 \\ u_3 \\ v_3 \\ \vdots \end{Bmatrix}$$

or $\{e\} = [B] \{u\}^e$.

Boundary Conditions and Applied Loading

With the geometry, mesh, and material properties prescribed to the finite element model, boundary conditions and loads are assigned. Boundary conditions refer to specifying the environment the geometry is subjected to in space. These require designating the value of the field variable(s) at specific nodes within the geometry. In thermal problems, for example, a boundary condition may imply that the temperature is specified on one region within the geometry. In structural analyses, a region may be restrained from moving in a particular direction or forced to move a certain amount in a particular direction.

The applied loads are those physical quantities that cause the field variable of the problem to change from its original or equilibrium state. In thermal problems, the applied load could be a heat source delivering a flux of heat to the geometry on one of its regions. In structural analyses, the applied load could be a stress that results in forces on one of the regions of a geometry.

NUMERICAL ANALYSIS

This stage begins with the processing of all input information from the preprocessing stage and creating finite element matrix equations to represent the physics of the problem. The behavior of each element with its prescribed material properties is cast into matrix form using the laws of physics and a finite element formulation technique.^[1,3] In thermal analyses, the matrix would be called an element conductivity matrix, whereas in structural analyses, the matrix is referred to as an element

stiffness matrix. To illustrate this concept, let $[k]^e$ be the stiffness of an element in a structural analysis problem. This matrix is calculated by integrating the following matrix multiplications over the volume of the element, dV^e :

$$[k]^e = \int [B]^T [E] [B] dV^e$$

Recall that $[B]$ is the strain-displacement matrix, and $[E]$ is the matrix of material properties. Then $[k]^e \{u\}^e = \{f\}^e$ for each element. Here $\{f\}^e$ represents the load vector containing the forces acting on that particular element, as prescribed by the applied load.

It is rare in FEA that finite element integrals can be calculated exactly. More commonly, the integral must be numerically approximated by a mathematical quadrature. Gauss quadrature is frequently used in FEA to numerically integrate finite element equations but can be applied to numerically integrate any mathematical expression. The general manner in which this is accomplished is to calculate the value of the integrand (i.e., $[B]^T [E] [B]$ in the element stiffness matrix) at a specified number of locations within the element (Gauss points), multiply the value of each integrand at the Gauss points by a weighting factor, and sum up the contributions from all Gauss points. The accuracy of this numerical integration depends on the number of Gauss points specified and the complexity of the mathematical expression in the integrand (e.g., a ratio of polynomials is most common in FEA).

To illustrate the use of Gauss quadrature for numerical integration, consider the integral

$$\int \cos(x) dx$$

specified in one direction (x), with the definite limits of $x = -1$ to $x = 1$. Integrating this expression directly gives the exact value of $[\sin(1) - \sin(-1)] = 1.6829$. For this one-dimensional integration, the Table 1 lists possible Gauss points at different locations within the interval of the defined limits.^[3] As the number of Gauss

points increases, the percent error characteristically decreases, converging to the exact solution.

With the stiffness matrix and load vector known for each element, the contributions of all elements to the behavior of the geometry as a whole must be assembled. This is accomplished by knowing which elements are connected at which nodes and summing their individual contributions. This results in a global (i.e., representing the entire geometry) equation,

$$[K]\{U\} = \{F\}$$

Here, $[K]$ is the global stiffness matrix, assembled from the summed contributions of each element's stiffness matrix; $\{U\}$ is the vector of field variables at the nodes for the entire mesh; and $\{F\}$ is the global force vector, assembled from the summed forces acting on each element. With this global stiffness matrix $[K]$, the boundary conditions (or prescribed values of the field variable at the nodes) are imposed on the matrix. This alters some of the entries in the global stiffness matrix because it is not necessary to solve for those prescribed nodal values.

The set of simultaneous equations, $[K]\{U\} = \{F\}$, must now be solved to obtain unknown nodal values of the field variables within $\{U\}$. Direct solvers, such as Gaussian elimination, are most commonly implemented in an FEA software package to solve for the unknowns. Sometimes, iterative solvers become necessary due to the physics of the problem.

POSTPROCESSING

Display of Results

Upon completion of the numerical analysis stage, the values of the field variables at all nodes within the finite element mesh are known and can be provided to the user. Additional physical parameters such as gradients (of temperature or displacement for strain and then stress, for example) can also be calculated from the

Table 1 Possible Gauss points at different locations within the interval of defined limits

Order of Gauss quadrature	Location of Gauss points over the interval $\{-1 \text{ to } 1\}$	Weight factors	Value of numerical integration	Resulting error (%)
1	0	2	$2\cos(0) = 2$	18.84
2	$\pm 1/\sqrt{3}$	1	$1[\cos(1/\sqrt{3}) + \cos(-1/\sqrt{3})] = 1.6758$	-0.42
3	$\pm\sqrt{0.6}$	5/9	$5/9[\cos(\sqrt{0.6}) + \cos(-\sqrt{0.6})] + 8/9[\cos(0)] = 1.683$	0.006
	0	8/9		

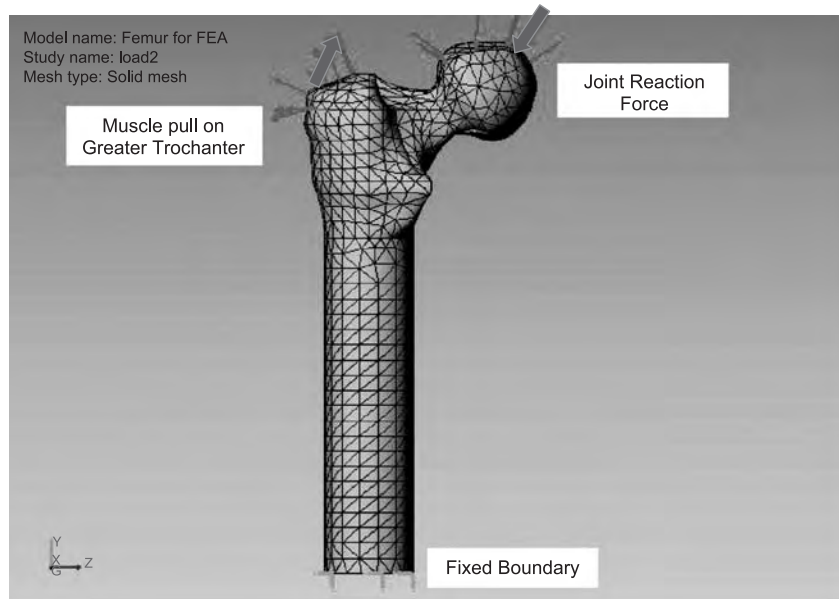


Fig. 3 Schematic drawing of the geometry in Fig. 2A and plausible boundary conditions and applied loads acting on the geometry. Fixed (immovable) boundary conditions are represented by green arrows; pink arrows represent joint reaction forces acting on the femoral head and tension created by the abductor muscles on the greater trochanter. Finite element analysis was performed with CosmosWorks.

constitutive relationships specified by the mathematical model. With the improvement in computer graphics, graphical display is the most common approach to view results. The finite element information for the femur as described in Figs. 2 and 3 were numerically analyzed with an FEA software package to provide a representative picture of what a graphical display might resemble (Fig. 4).

Error Analysis

A vital step in any finite element analysis is to analyze the computed results critically and make decisions of

their accuracy. Care must be taken so that potential, controllable errors are eliminated. Error in this sense means disagreement between the finite element results and the exact solution of the problem. One difficulty in error assessment is that the exact solution is typically not known; this is the reason that an FEA was undertaken. Sometimes, an exact solution may be obtained on a simplified, but similar problem to which some comparison of the FEA results can be made. Or, experimental data might exist for the FEA problem and comparisons made. Thus, the finite element modeler must thoroughly review FEA results to ensure the highest accuracy has been obtained by the model.

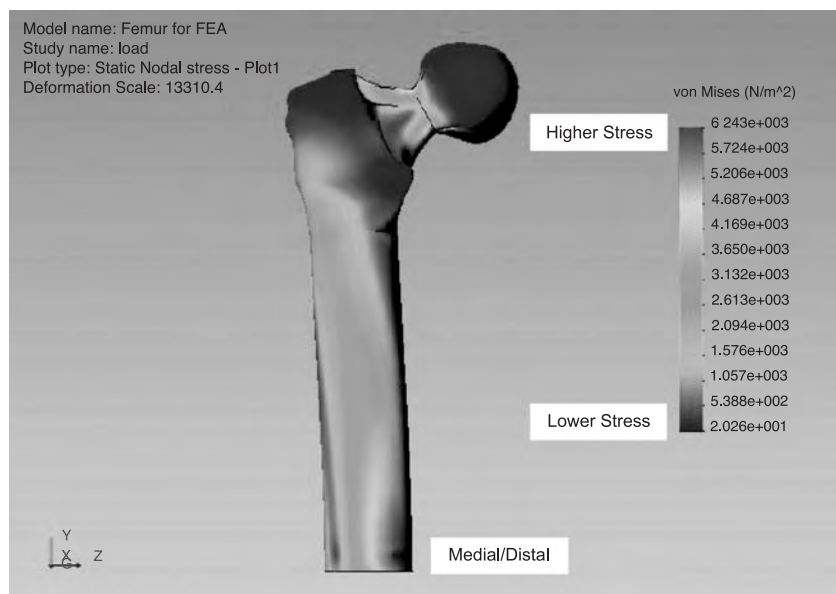


Fig. 4 Graphical display of the deformed geometry (enlarged by a scale factor for descriptive purposes) from the finite element mesh of Fig. 2 under the action of the boundary conditions and applied loads of Fig. 3. The color contours within the geometry represent the magnitude of an averaged stress distribution (von Mises stresses in Pascals) that exists within the geometry. Note the color bar scale on the right; red represents areas under high stress (superior/inferior femoral neck, medial distal femur), while blue shading indicates areas of low stress. Finite element analysis was performed with CosmosWorks.

Sources of error can be classified into a few categories:^[3] modeling error, user error, software bugs, discretization error, and numerical error. Modeling error refers to how well the the mathematical model (geometry, material properties, boundary conditions, applied loading, etc.) represents the physical problem. Assumptions about the physical problem are made in the model out of necessity: Perhaps enough is not known about certain parts (boundary conditions, applied loads, for example) of the physical problem or material behavior must be approximated. Modeling error thus refers to those approximations of the problem that are done intentionally with the best representation applied to answer the question about the physical problem.

Errors made by the software user, once the mathematical model of the physical problem has been created, can be classified as user error. User error would include such issues as choosing the wrong element type to discretize the geometry (bar versus beam elements, for example), allowing element shapes/sizes to have poor aspect ratios or angles, not using an appropriate degree of numerical integration (quadrature), and mistakes in data entry. Another potential source of error is the existence of bugs in the software. Commercial vendors spend an extraordinary amount of time in debugging routines to eliminate this type of error, but software users should not be blind to the potential existence of software bugs.

Discretization error refers to how accurately the mesh and refinement in key regions of the geometry can capture the behavior of the object in the physical problem. The number of nodes per element (and thus the shape functions that define the piecewise interpolation within the geometry), the number of nodes in the entire mesh, and the number of elements in the geometry influence discretization error. Finally, numerical error classifies how the computer used for the FEA treats numbers. Information about numbers in the far decimal places is lost when the computer truncates or rounds off numbers. Additionally, as numbers are multiplied, divided, added, or subtracted, information may again be lost in the far decimal places. Thus, the number of decimal places stored should be as large as allowed by the computer.

A certain amount of error is expected in a finite element solution, as any FEA is an approximation of the physical problem. However, minimizing sources of error that can be controlled provides for a great degree of confidence in FEA results.

BIOENGINEERING APPLICATIONS

Some of the earliest applications of FEA to bioengineering problems were performed in the 1970s. The first symposium to present many of these applications occur-

red in 1980^[4] and covered a wide variety of topics. These included simulations of pulsatile blood flow in normal and abnormal arterial vessels, and numerous stress analyses of the lung, heart, normal and bioprosthetic heart valves, bone, and fixation devices for fractured bone. It is quite astounding to see the breadth of problems treated at that time with FEA still in its beginning stages of use. In these initial applications, it was more common for studies to use software that was internally developed rather than commercially available software.

Since these early applications, FEA in bioengineering has seen as dramatic a rise in use as in other engineering fields. Today, studies of similar problems originally described in 1980 remain, while new applications continue to appear. An important reason for the increasing use of FEA in all fields is the commercial availability of powerful FEA software with large element selections, extensive material models, and interactive preprocessing and post-processing graphical packages. Thus, the level of complexity of the problems and accuracy of the mathematical models have increased, particularly with respect to geometric representation and material behavior of the physical problem. Imaging techniques such as laser scanning, computed tomography (CT), and magnetic resonance imaging (MRI) are often employed to provide accurate, three-dimensional information about the shape and size of objects of interest. Numerous material models are incorporated into software packages that allow more realistic representation of the material behavior/properties of the physical problem. A few specific fields that have experienced extensive finite element applications will be highlighted. The clinical utility of FEA is growing.

Dentistry

Teeth are subjected to amazingly high forces of mastication (chewing) and yet are capable of withstanding these forces for many decades. Finite element simulations of a tooth during its normal function demonstrate how stresses and strains are distributed in the different layers of the tooth (enamel, dentin, etc.) and into the bone supporting the tooth.^[5] Finite element analyses of a tooth with a composite restoration (i.e., dental filling) illustrate how stress distributions are altered by the presence of the restoration and can provide comparative information on the behavior of the tooth with different materials for the composite. Thermal analyses of normal teeth and composite restorations while exposed to hot and cold environments (e.g., liquids)^[6] provide insight into the behavior of these physical systems.

Cardiovascular System

The cardiovascular system receives quite a lot of attention from finite element modelers. So many

life-threatening diseases can affect this system that stress analyses and fluid mechanics applications provide insightful information into normal function, function affected by disease, and the efficacy of treatment procedures. Blood flow through various vessels of the cardiovascular system is frequently analyzed. Coronary arteries provide the oxygenated blood to the heart itself and are critical elements in a normal heart. Stenosed coronary arteries can thus present major problems to the patient, the mechanics of which can be visualized with finite element simulations.^[7] Similarly, the aorta is the major conduit for blood leaving the heart to be distributed to the rest of the vascular system. Velocity patterns and shear stress distributions in the vascular wall of the ascending aorta can serve as predictive tools for the clinician.^[8] Some exciting work has been accomplished in modeling the heart itself. The left ventricle, the chamber of the heart responsible for forcing blood into the rest of the vascular system, has received much attention because of its importance. Models that simulate the mechanics of the normal left ventricle and an infarcted (due to heart attack) ventricle are enlightening.^[9-11]

Musculoskeletal System

Articular cartilage

Articular cartilage is the soft, connective tissue that covers the ends of long bones in joints. Its presence allows for smooth, painfree movement of our joints. Articular cartilage is very durable and can remain healthy for the length of one's life. This is in the face of enormous joint forces it is subjected to during activities of daily living. However, arthritis is a major health problem and destroys joint cartilage until severe disability results. FEA has been applied to study this tissue in order to understand various issues that contribute to its longevity and how the mechanics of articular surfaces become altered by the presence of damaged cartilage.^[12,13] To represent the material behavior of articular cartilage better, new mathematical models have been implemented in the study of this tissue, implemented coupled field variable analyses and new material behaviors.^[14,15] Simulations of surgical treatment of deteriorating cartilage have also been performed.^[16]

Bone and bone remodeling

Very exciting applications of FEA examine bone and bone remodeling. All biological tissues have the ability to remodel depending on the input signals they receive. For example, muscles adapt to increased activity (i.e., increased stress), such as weightlifting, by increasing their capacity and size. Muscles also respond to decreased activity (i.e., decreased stress), such as the zero

gravity experienced in space, by decreasing in capacity and size. The same is true with bone. However, bone is one of the few biological tissues whose potential for remodeling has been successfully described by mathematical models, some of which are called strain-adaptive remodeling theories. These remodeling theories have been put into practice with the finite element analysis to show how bone would change its geometry and/or material properties due to changes in applied loads/stresses that the bone experiences.^[17-19] The major effort in applying this combined FEA/bone remodeling theory is toward describing bone behavior around total joint replacements. A total joint implant can survive joint forces for many years, but loosening of the implant is a frequent occurrence, in part due to how the bone remodels around the replacement after implantation because of the changed stress environment experienced by the bone. The combined FEA/bone remodeling theory has the potential to serve as a clinical tool that allows the design engineer and clinician to predict where bone will remodel and how to design the implant, or implantation technique better, to extend the life of the total joint replacement.

CONCLUSION

This article has provided a basis for understanding a finite element analysis. Numerous books have been written with much detail on the technique and can be used for further reference. A small sampling of such books are given in the reference section.^[1-3,20-22] It must be remembered that FEA is an approximate solution technique used to solve a mathematical model. This mathematical model is an idealized representation of the actual physical problem under consideration. One can reap many benefits by knowing the behavior of the physical problem. Finite element results can show how a particular device will perform in its intended application. Inference can be made on whether it will perform adequately or fail. In a straightforward manner, geometry, material properties, boundary conditions, and applied loads can be varied in a finite element analysis to see how such design changes can alter the behavior of the problem. With attention given to the mathematical model and its incorporation into finite element equations, an accurate solution of the physical problem via finite element analysis can be obtained.

ACKNOWLEDGMENTS

The assistance of Joseph Iaquinto for the solid model of the femur and of Scott Sell for finite element images is greatly appreciated.

ARTICLES OF FURTHER INTEREST

Cartilage Regeneration; Composites; Dental Implants; Modeling, Biomedical; Osteoinductive Substances and Materials

REFERENCES

1. Huebner, K.H.; Dewhurst, D.L.; Smith, D.E.; Byron, T.G. *The Finite Element Method for Engineers*, 4th Ed.; John Wiley & Sons, Inc.: New York, 2001.
2. Rao, S.S. *The Finite Element Method in Engineering*, 3rd Ed.; Butterworth Heinemann: Boston, 1999.
3. Cook, R.D.; Malkus, D.S.; Plesha, M.E.; Witt, R.J. *Concepts and Applications of Finite Element Analysis*, 4th Ed.; John Wiley & Sons, Inc.: New York, 2002.
4. *Finite Elements in Biomechanics*; Gallagher, R.H., Simon, B.R., Johnson, P., Gross, J., Eds.; John Wiley & Sons, Inc.: New York, 1982.
5. Hubsch, P.F.; Middleton, J.; Knox, J. A finite element analysis of the stress at the restoration-tooth interface, comparing inlays and bulk fillings. *Biomaterials* **2000**, *21* (10), 1015–1019.
6. Toparli, M.; Aykul, H.; Sasaki, S. Temperature and thermal stress analysis of a crowned maxillary second premolar tooth using three-dimensional finite element method. *J. Oral Rehabil.* **2003**, *30* (1), 99–105.
7. Shim, E.B.; Kamm, R.D.; Heldt, T.; Mark, R.G. Numerical analysis of blood flow through a stenosed artery using a coupled, multiscale simulation method. *Comput. Cardiol.* **2000**, *27*, 219–222.
8. Leuprecht, A.; Perktold, K.; Kozerke, S.; Boesiger, P. Combined CFD and MRI study of blood flow in a human ascending aorta model. *Biorheology* **2002**, *39* (3–4), 425–429.
9. Needleman, A.; Rabinowitz, S.A.; Bogen, D.K.; McMahon, T.A. A finite element model of the infarcted left ventricle. *J. Biomech.* **1983**, *16* (1), 45–58.
10. Costa, K.D.; Hunter, P.J.; Rogers, J.M.; Guccione, J.M.; Waldman, L.K.; McCulloch, A.D. A three-dimensional finite element method for large elastic deformations of ventricular myocardium: I—Cylindrical and spherical polar coordinates. *J. Biomech. Eng.* **1996**, *118* (4), 452–463.
11. Costa, K.D.; Hunter, P.J.; Wayne, J.S.; Waldman, L.K.; Guccione, J.M.; McCulloch, A.D. A three-dimensional finite element method for large elastic deformations of ventricular myocardium: II—Prolate spheroidal coordinates. *J. Biomech. Eng.* **1996**, *118* (4), 464–472.
12. Wayne, J.S. Load partitioning influences the mechanical response of articular cartilage. *Ann. Biomed. Eng.* **1995**, *23*, 40–47.
13. Dunbar, W.L., Jr.; Un, K.; Donzelli, P.S.; Spilker, R.L. An evaluation of three-dimensional diarthrodial joint contact using penetration data and the finite element method. *J. Biomech. Eng.* **2001**, *123* (4), 333–340.
14. Wayne, J.S.; Woo, S.L.-Y.; Kwan, M.K. Application of the u-p finite element method to the study of articular cartilage. *J. Biomech. Eng.* **1991**, *113*, 397–403.
15. Li, L.P.; Buschmann, M.D.; Shirazi-Adl, A. A fibril reinforced nonhomogeneous poroelastic model for articular cartilage: Inhomogeneous response in unconfined compression. *J. Biomech.* **2000**, *33* (12), 1533–1541.
16. Wu, J.Z.; Herzog, W.; Hasler, E.M. Inadequate placement of osteochondral plugs may induce abnormal stress-strain distributions in articular cartilage—finite element simulations. *Med. Eng. Phys.* **2002**, *24* (2), 85–97.
17. Hart, R.T.; Davy, D.T.; Heiple, K.G. A computational method for stress analysis of adaptive elastic materials with a view toward applications in strain-induced bone remodelling. *J. Biomech. Eng.* **1984**, *106*, 342–350.
18. Huiskes, R.; Weinans, H.; Grootenboer, H.J.; Dalstra, M.; Fudala, B.; Sloof, T.J. Adaptive bone remodelling theory applied to prosthetic-design analysis. *J. Biomech.* **1987**, *20*, 1135–1150.
19. Kerner, J.; Huiskes, R.; van Lenthe, G.H.; Weinans, H.; van Rietbergen, B.; Engh, C.A.; Amis, A.A. Correlation between pre-operative periprosthetic bone density and post-operative bone loss in THA can be explained by strain-adaptive remodelling. *J. Biomech.* **1999**, *32*, 695–703.
20. Zienkiewicz, O.C.; Taylor, R.L. *The Finite Element Method*, 5th Ed.; Butterworth Heinemann: Oxford, 2000.
21. Bathe, K.J. *Finite Element Procedures*; Prentice-Hall: New Jersey, 1996.
22. Hughes, T.J.R. *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*; Prentice-Hall: New Jersey, 1987.

Flow Cytometry

J. Paul Robinson

Immunopharmacology and Biomedical Engineering, Purdue University, West Lafayette, Indiana, U.S.A.

INTRODUCTION

Flow cytometry is a technology that has impacted both basic cell biology and clinical medicine in a very significant manner. The essential principle of flow cytometry is that single particles suspended within a stream of liquid are interrogated individually in a very short time as they pass through a light source focused at a very small region. The optical signals generated are mostly spectral bands of light in the visible spectrum, which represent the detection of various chemical or biological components, mostly fluorescence. A key aspect of flow cytometers is that because they can analyze single particles/cells, it is possible to separate particles/cells into populations based upon a statistical difference of any of 10 to 20 variables that can be measured on each particle/cell. Using these statistical analyses, it is possible to separate these populations electronically and identify them using multivariate analysis techniques.

The most common detection system in flow cytometry uses fluorescent molecules that are attached by one means or another to the particle of interest. If the particle is a cell, such as a white blood cell, for example, the fluorescent probe might be membrane bound, cytoplasmic, or attached to nuclear material. It is a common practice to use monoclonal or polyclonal antibodies that recognize specific receptors on cells. By conjugating fluorescent molecules to these antibodies, it is possible to monitor both the location and number of these conjugated antibodies as they bind to cell receptors. Particles of almost any nature can be evaluated by flow cytometry. They can be very small, even below the resolution limits of visible light, because they can be detected by their fluorescent signatures. Similarly, depending on the structure of the flow cell and fluidics, particles as large as several thousand microns can be evaluated.

The key advantage of flow cytometry is that a very large number of particles can be evaluated in a very short time; some systems can run particles at rates approaching 100,000 particles per second while collecting 10 to 20 parameters from each particle. Finally, the principle of cell sorting in flow cytometry allows this technology to separate single particles/cells physically from mixed populations. Thus single particles can be physically placed into a defined location for further analysis and, if necessary, this process can be performed

under sterile conditions. This capability makes flow cytometry a valuable tool for rare event (1:100,000 or even 1:1,000,000) analysis.

In 1983 Shapiro noted that multiparameter flow cytometry was now a reality in the field^[1] because of the availability of commercial instruments. Since that time, the field has expanded well beyond anything that was then considered possible. Today's instruments have the capacity to measure 10–15 spectral bands simultaneously together with a variety of scatter signals. With modern computers it is possible to perform complex multiparametric analyses virtually instantaneously, allowing time to make sorting decisions after measurements are made. The result of this technology is that it is now possible to generate clinical diagnostic information rapidly from complex heterogeneous mixtures of samples such as human blood and to perform this in real time.^[2]

OVERVIEW

Basic Principles

The basic principles of flow cytometry arise from some very old ideas generated early in the 20th century and of course follow the principles of laminar flow defined by Reynolds in the late 19th century. Some 50 years later, Maldavan designed an instrument (although it is not clear that he actually constructed it) that could have identified single cells using a microscope and a photodetector.^[3] In the 1940s Papanicolaou demonstrated that he could identify as cancerous cells from cervical cancer by observing the staining patterns obtained by staining tissues with specifically designed stains.^[4] This suggested several directions of research, primarily using image analysis techniques for the identification of abnormal cells. The limited capability of computers and imaging technology at that time made this quite difficult and resulted in a movement toward single-cell analysis, as opposed to image processing and recognition. It was in the 1960s that Louis Kametsky began the drive to design and build single-cell analyzers. While working at IBM's Watson Labs, Kametsky was interested in using optical character recognition techniques to identify cancer cells. Because of the lack of computation, this became a difficult goal and in place

of image-based technology,^[5] Kamensky focused on single-cell analysis and the design of a cytometer that measured absorption and scatter and shortly thereafter added the ability to sort cells using fluidic switching.^[6] At the same time, Fulwyler was trying to solve a problem generated by the study of red blood cells using a single-cell analysis system. It had become apparent that a bimodal distribution of red blood cells observed using a Coulter volume detector suggested two different types of red blood cells, contrary to accepted medical understanding. Fulwyler had heard of Richard Sweet's development of high-speed chart recorders using electrostatic drop generation.^[7] Fulwyler visited Sweet's laboratory and essentially utilized this technology to design and build a cell sorter to separate red blood cells.^[8] Ironically, upon completion of the instrument, it took only a few hours to recognize that the supposed bimodal distribution was related to spatial orientation rather than to inherent red blood cell variability (Fulwyler, personal communication). Amazingly, this finding of great significance was never published since it was immediately obvious that sorting of white blood cells was an opportunity not to be missed. The history of the development of cell sorting is well covered by Shapiro.^[9]

Fluidic Systems

Reynolds formulated the relationship for fluid flow as $Re = vd\rho/\eta$ where Re is the Reynolds number (a dimensionless number), v the average velocity, d the tube diameter, ρ the fluid density, and η a velocity coefficient. Below a Reynolds number of 2300, flow will be laminar, a necessary factor for quality optical measurements in flow cytometry. Maintenance of nonturbulent flow requires careful design of fluidic systems in flow cytometers, particularly the flow cell components. Cells are hydrodynamically focused in a core stream encased within a sheath (Fig. 1). This sheath-flow principle was derived from the work of Moldavan and subsequently Crosland-Taylor,^[10] who designed a system similar to most used today in which an insertion rod (needle) deposits cells within a flowing stream of sheath fluid (usually water or saline), creating a coaxial flow that moves from a larger to a smaller orifice, creating a parabolic velocity profile with a maximum at the center of the profile. The general design of such a system is shown in Fig. 1. Because of the hydrodynamic focusing effect, cells that are injected through the injection tube remain in the center of the core fluid, thus allowing very accurate excitation with subsequent excellent sensitivity and precision of measurement within the flowing stream. There is a small differential pressure between the sheath and the sample (which is the core); the sample is 1 to 2 PSI above the sheath, forcing alignment

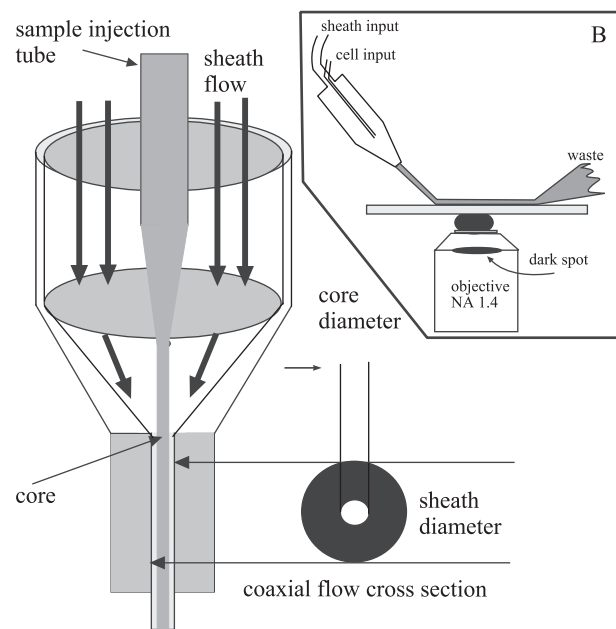


Fig. 1 Shown here is the basic structure of a typical flow cell. Sheath fluid flows through a large area and under pressure is forced into a much smaller orifice. In the center of the cell is an injection tube that injects cells or particles into the center of the flowing stream, forcing the cells to undergo hydrodynamic focusing, which will result in laminar flow if Reynolds number does not exceed 2300. Shown is the coaxial cross section of the sheath and core. B shows an alternative flow cell, an axial flow system typically used in microscope-based flow cytometers. In this system the laminar stream flows across a coverglass to a waste collector on the opposite side.

of cells in single file throughout the core. If the pressure is increased too much, the core diameter will increase, destabilizing the flowing cells and reducing the accuracy and precision of the measurement. If a highly accurate system is required, multiple sheaths can be used to create very stable flow streams, but this is generally not used in commercial systems.

A crucial component of the flow cytometer is the design of the flow cell, in which the fluid flows from a very large area to a very constrained channel. The velocity, which is proportional to the square of the ratio of the larger to smaller diameter, increases significantly within the smaller channel. Within this channel, the velocity profile is parabolic with a maximum at the center of the stream and almost zero at the walls of the vessel. This becomes a crucial issue in flow cytometry when biological specimens are used because these samples contain proteins, and surface binding will eventually increase turbidity and destroy the hydrodynamic nature of the flow. While Reynolds number remains less than 2300, laminar flow occurs. The acceleration at the core of the vessel is an important aspect of flow cytometers. Since particles are injected into the

center of the flowing stream, the highly accelerated central core creates spatial separation of particles within this rather long core stream. This separation creates the ability to analyze the signals from single cells more accurately. Once particles are accurately identified and are spatially separated within the core, it is possible to separate them physically in a process known as particle sorting (discussed later). An alternative system to the one described earlier uses axial flow where cells are shot onto the surface of a microscope objective with a regular nozzle to obtain laminar flow, flow across the objective in a laminar flow, and are extracted from the system on the other side of the objective. This is shown in Fig. 1B and is similar to systems designed by Harald Steen and others. There are several advantages of this system, such as high numerical aperture microscope objectives providing excellent resolution and signal to noise, and the ability to use a regular arc lamp for the light source. This system has extraordinary sensitivity for forward scatter and is the most sensitive system available. It was initially designed to be optimized for very small particles such as microorganisms.

Optical Systems

Most flow cytometers use lasers as excitation sources. In the earliest systems, mercury lamps were used; however, in the late 1960s relatively large water-cooled ion lasers were identified as the most desirable source of coherent light at 488 nm, the best excitation wavelength for fluorescein. These high-cost, large, and inefficient light sources shaped the design of the instruments themselves, making them enormous constructs often

taking 60 to 80 sq. feet of floor space and requiring high volume cooling water and high current levels. More recently, however, with the advent of solid-state lasers, the footprint of flow cytometers has been significantly reduced. Further, in the mid-1980s, there was an emerging market for flow cytometers that did not sort. These instruments were known as analyzers and are now commonly referred to as benchtop instruments. This is somewhat of a misnomer, as the third generation of sorters is almost indistinguishable from the benchtop analyzers of the past.

As already indicated, the key to the efficiency and sensitivity of current flow cytometers is the laser-based coherent light source. The chief criterion for selection of a laser is the excitation wavelength. The beam should be segmented in a transverse emission mode (TEM) of TEM 00, although in some circumstances a mixed TEM 00 and TEM 01 mode does not preclude the usefulness of such a beam mode (Fig. 2A). The excitation source must match the absorption spectra of the fluorochromes of interest. One reason that early systems used large water-cooled argon-ion lasers was that multiple lines could be obtained from these lasers. The argon laser was selected, as it was the only coherent source of excitation satisfactory for the most used fluorochrome in the field—fluorescein. The argon-ion laser could produce lines in the UV (350 nm), deep blue (457 nm), blue (488 nm), and blue-green (514 nm) regions, making this a very useful light source. The light source needs to be focused to a spot and a desired shape. This is accomplished by using a beam-shaping optic to obtain the desired crossed-cylindrical beam shape. For reasons explained in Fig. 2, the most desirable beam shape is an elliptical beam of

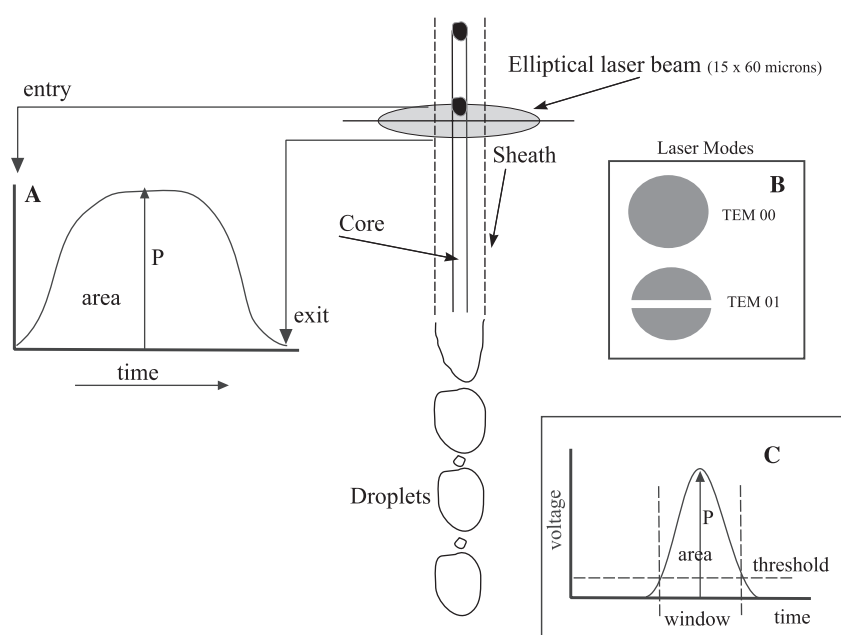


Fig. 2 As cells pass through the interrogation point, they create a pulse that can be characterized as shown above. At the point of entry into the laser beam, the pulse rises to a peak and holds for as long as the cell is in the stream. Once the cell begins to leave the laser beam profile, the signal returns to zero. The maximum signal is the peak, and the time taken for entry and exit of the beam is the time of flight (TOF). It is common to measure the total area under the curve (integral signal) for total fluorescence. Shown in B are the beam profiles most commonly used in flow cytometry. Most desirable is TEM 00; however, it is possible to mix the TEM 00 and the TEM 01 modes. In C are shown the definitions of each component of the signal from a cell passing through an elliptical beam.

approximately 15 by 60 microns. This produces a beam with a large, relatively flat cross-section that reduces the variation in intensity of the excitation spot should the particle move around within the excitation area. Reducing the beam even further would have the effect of slit-scanning the traveling particle.

Electronic Systems

Flow cytometers collect a lot of data very quickly. In fact, they are in a class of instruments that push the limits of data collection. For example, it is currently possible to collect at least 11 fluorescent spectral bands simultaneously together with at least two scatter signals on thousands of cells per second, creating a multivariate analysis problem.^[11] The key principle of flow cytometry is that every particle is identified individually and classified into a category or population member according to multivariate analysis solutions. Every particle that passes the interrogation point would be collected on every detector, which would cause a serious overload collection problem. To solve this, a circuit is included called a discriminator, which can be set to exclude signals lower than a preset voltage (Fig. 2C). On many current instruments it is possible to use discriminators on any or all detectors. That is to say, multiple detectors must register a preset signal level or nothing is collected by the data collection system. Once a discriminator setting is satisfied, this detector triggers the entire data collection system and all identified detectors will measure the signal. Frequently, for particles of bacteria to most animal cell size (1–20 microns), a forward-angle light-scatter signal is used to discriminate the presence of a measurable particle. However, it is also useful to use a fluorescence detector if one wishes to detect only particles of a certain level of fluorescence.

The most frequently recognizable detection system in flow cytometers is that of fluorescence. The initial detection system used in the earliest instruments was Coulter volume, based on the original patent of Wallace Coulter,^[12] whereby the principle of impedance changes was transferred from cell-counting instruments to flow cytometers. In addition to impedance, light scatter was also measured. Current systems have taken a rather complex pathway for the measurement of fluorescence.

Linear amplifiers produce signals that are proportional to their inputs, and while it is possible to amplify this signal, most immunofluorescence applications have huge dynamic ranges that are beyond amplification in the linear domain. For this reason, logarithmic amplifiers with scales covering three to five decades are required. This is particularly useful for samples in which some cells exhibit very small amounts of signal, while others have signals four orders of magnitude larger.

Detectors

It has become standard design to utilize a PMT for each spectral wavelength desired. In most pre-1990 instruments, a maximum of four or five spectral bands was collected. However, beginning in the last decade of the 20th century, it became evident that 5–10 spectral signatures were desirable. Each spectral band is collected by a PMT strategically placed within an optical system of which there are many current designs. Figure 4 shows several different optical layouts currently used in commercial systems. It is now evident that the biological requirements are in the range of 10–15 spectral bands. Next-generation systems will include either a vast number of PMTs, avalanche photodiodes, or multichannel PMTs, in addition to high-speed cameras. The disadvantages of the multichannel PMT is that detection sensitivity is reduced and it is not currently possible to adjust the sensitivity of each channel as can be achieved with individual PMTs. The advantage is that the complexity and number of optical components are reduced.

Most cytometers use photomultiplier tubes (PMTs) as detectors for both fluorescence and scatter. The pulse of a particle crossing the excitation beam will depend upon the beam shape, beam intensity, and particle size, as well as the velocity of the particle. Systems running at 10 m/s will cross a 10-micron beam in 1 microsecond, or a 5-micron beam in only 500 ns. The majority of instruments prior to publication of this article were designed around analog detection, rather than digital electronics. Essentially, once the threshold voltage is met (based on the discriminator circuit described earlier) the signal (usually 0–10 volts) is fed into an analog-to-digital converter (ADC) circuit, called a comparator circuit, whose purpose is to identify and signal the presence of a measurable signal that is used to trigger the rest of the detection systems. This is a binary decision only. Once a decision to collect is made, several measurements for each variable are made, such as peak, integral, and time-of-flight. There are several complications that can cause problems in the detection electronics. For example, if two particles pass the interrogation point at very close intervals, both signals must be aborted if this time is shorter than the reset time for the electronics. Another circuit is required to make this decision.

To further complicate the electronics, many systems use two or more laser beams delayed by a few microseconds. Each particle must be correctly analyzed by each laser, so data from the first beam must be stored while waiting for the same particle to pass the second beam, and so on. If the beam separation is large enough, several cells might be analyzed by the first beam before the first cell passes the second beam. This rather complex system is not necessary on simpler analysis systems, but it is absolutely necessary on more advanced

multilaser cell sorters. In addition, the time taken for all the analysis components is finite, which essentially sets the maximum analysis rate of the flow cytometer. The faster the system, the shorter the dead time must be. For example, to analyze 100,000 cells per second, a dead time of less than 10 microseconds would be necessary. In fact, depending on how many events must actually be analyzed to have 100,000 cells per second, the dead time would need to be considerably shorter.

Spectral Compensation

When a particle or cell contains fluorophores of multiple spectral bands, the identification and analysis become considerably more complex. For example, a detector with a band pass filter designed to collect fluorescence from FITC (525 nm) and another detector designed to collect signals at 550 nm (PE) will register photons in both detectors. It is impossible to determine which detector is detecting the real photons from FITC. This is not a problem if a single fluorophore is being collected, but when two or more fluorophores with close emission bands are present, it is necessary to identify which fluorophore was the real emitter of the photons. To achieve this, it is necessary to perform spectral compensation whereby a percentage of signal from one detector is subtracted from the other. As the number of fluorophores increases, so too does the complexity of the spectral compensation. A complex set of circuits must be designed that allows for a percentage of each signal to be subtracted from every other detector. Naturally, there are some instances where there is no overlap, but with six or seven detectors competing for signals from the narrow spectral emission range available from a single excitation source, it is absolutely necessary to compensate for spectral overlap. While this can be performed perfectly well in software off-line,^[13] if the goal of the analysis is to sort a certain population of cells physically, the compensation must be performed in real time between the time the cell passes the excitation beam and when the cell reaches the last time available for a sort or abort decision to be made. Compensation in flow cytometry is very complex and requires a large number of controls to establish appropriate compensations setting and photomultiplier voltages. As fluorescent dyes increase in number and spectral proximity, the need for complex spectral compensation circuitry also increases. This is far more complex than anything currently available in image analysis systems.

Cell Sorting

The principle of cell sorting was included in instruments designed by Fulwyler,^[8] Kamensky,^[6] and also

Dittrich^[14] in order to analyze a cell of interest definitively. It was Fulwyler, however, who identified the technique developed by Richard Sweet^[7] for electrostatic droplet separation for use in high-speed inkjet printers as the ideal technology for cell sorting. This evolved into the technique of choice for virtually all commercial cell sorters. This is shown in Fig. 2 and also Fig. 3. This idea was implemented into a commercial system by Herzenberg's group^[15] in the early 1970s. As already noted, the initial reason for Fulwyler's implementation was the desire to separate what were apparently two distinct populations of red blood cells that appeared on analysis based on Coulter volume measurements. The principle of electrostatic sorting is based on the ability to first identify a cell of interest based on measured signals, identify its physical position with a high degree of accuracy, place a charge on the stream at exactly the right time, and then physically collect the sorted cell into a vessel.

The technology of high-speed sorting has been recently well defined by van den Engh,^[16] who discusses in detail the complex issues. In brief, the speed and accuracy of a cell sorter are based on a number of factors. Firstly, despite the initial discussion pointing out that a fully stable laminar flow is required for accurate analysis, for cell sorting the stream must be vibrated by a piezoelectric device to generate droplets. As described by van den Engh, it is necessary to have high-speed electronics and to match the nozzle diameter, sheath pressure, and droplet generation frequency to obtain stable droplet generation and thus high-speed cell sorting. The principle that governs the generation of droplets has been characterized by Kachel;^[17] the wavelength of the undulations is $\lambda = v/f$, where λ = the undulation wavelength, v = the stream velocity, and f = the modulation frequency.

When $\lambda = 4.5d$ (d = exit orifice = stream diameter), the system is optimized for maximum droplet generation. Thus, the optimal generation frequency is given by $f = v/4.5d$. If a system is designed to accommodate this optimal droplet formation, as demonstrated by Pinkel,^[18] the jet velocity is proportional to the square root of the jet pressure. Thus, an optimal system to sort events at 20,000 Hz such that drops are separated by 4.5 stream diameters and flowing at 10 ms would make each drop 200 microns apart. As the number of sorted drops increases, the diameter must decrease, with the obvious conclusion that the speed of high-speed sorters will eventually be partially regulated by the size of the particle to be sorted and the velocity that the stream can achieve without destroying the sample. This is particularly important for biological particles such as cells.

High-speed sorters are essentially sorters that are designed to operate at sort speeds in excess of 20,000 particles per second. To accomplish this, higher pressures must be placed on the sample stream. When

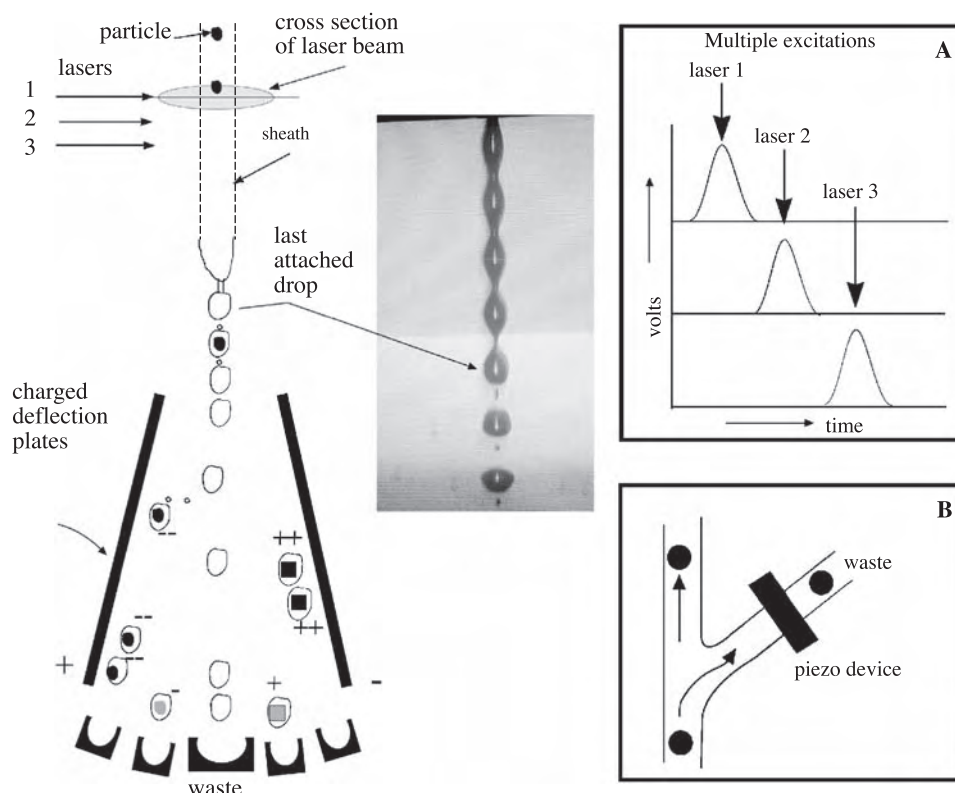


Fig. 3 This figure shows the principle of electrostatic cell sorting based on Sweet's inkjet printer technology. In this figure, a stream of liquid intersects a laser beam (or multiple laser beams 1, 2, 3). The stream is vibrated by a piezo-electric crystal oscillator at frequencies from 10,000 to 300,000 Hz depending upon the orifice size, stream velocity, nature of the stream, and particle size. Typically 30–50,000 Hz is used to create droplets at the same frequency. Once a cell/particle is identified as desirable, a charge is placed on the stream that remains with the last drop (last attached drop) that leaves the stream. Using a computation method, this drop is sorted by being attracted toward a plate almost parallel with the stream and containing opposite charges in the vicinity of 5000 volts. Each droplet containing a desirable particle can be placed into one of several containers (shown is a four-way sorting system). In the center of the figure is a video image of the droplets strobed at the same frequency as the droplet formation. A shows the pulses of 3 different lasers as a particle passes by each beam separated in space. Thus, a particle will pulse from each laser a few microseconds apart. This way, signals from each laser can be individually analyzed. B is an alternative sorting system using fluid switching techniques. In this system the waste stream is blocked momentarily to allow a desired cell to pass into the sorting pathway.

systems exceed 40,000 cells per second, the key issue becomes analysis time—obviously the limiting factor since complex analysis must precede a sorting decision. The maximum speed of droplet formation is therefore not the limiting factor in design of a high-speed flow cytometer. As discussed in van den Engh,^[16] the primary issue is the high pressures that must be used to create very high-speed droplet formation. At droplet frequencies of 250,000/second, the jet pressure must approach 500 PSI, a significantly higher value than can be designed safely in most systems. Thus, if pressures are limited to around 100 PSI, a droplet rate of around 100,000 is closer to the realistic range. This then is the real limitation to current high-speed sorting systems.

Poisson statistics enter the equation at this time as well since it is impossible to predict exactly when any particle is going to pass the interrogation point. This adds uncertainty in the analysis and, as discussed

previously, it is crucial to ensure that no measurements take place as two cells try to pass through the interrogation point at or near the same time. Thus there is a relationship between particle concentration and coincidence detection.

Cell sorting has become a very important component of flow cytometry. In particular, the isolation of CD34 human hematopoietic stem cells by flow sorting specifically for transplantation purposes has revolutionized capabilities in transplantation.^[19] Naturally, to perform such a sort, all components of the instrument that come in contact with the cells must be sterile.

Another issue that relates to sorting is the potential dangers involved in sorting certain samples, particularly human samples that may be infected with AIDS, or more commonly, hepatitis virus. This is an area that can cause considerable tension between operators and researchers wanting to sort materials from infected

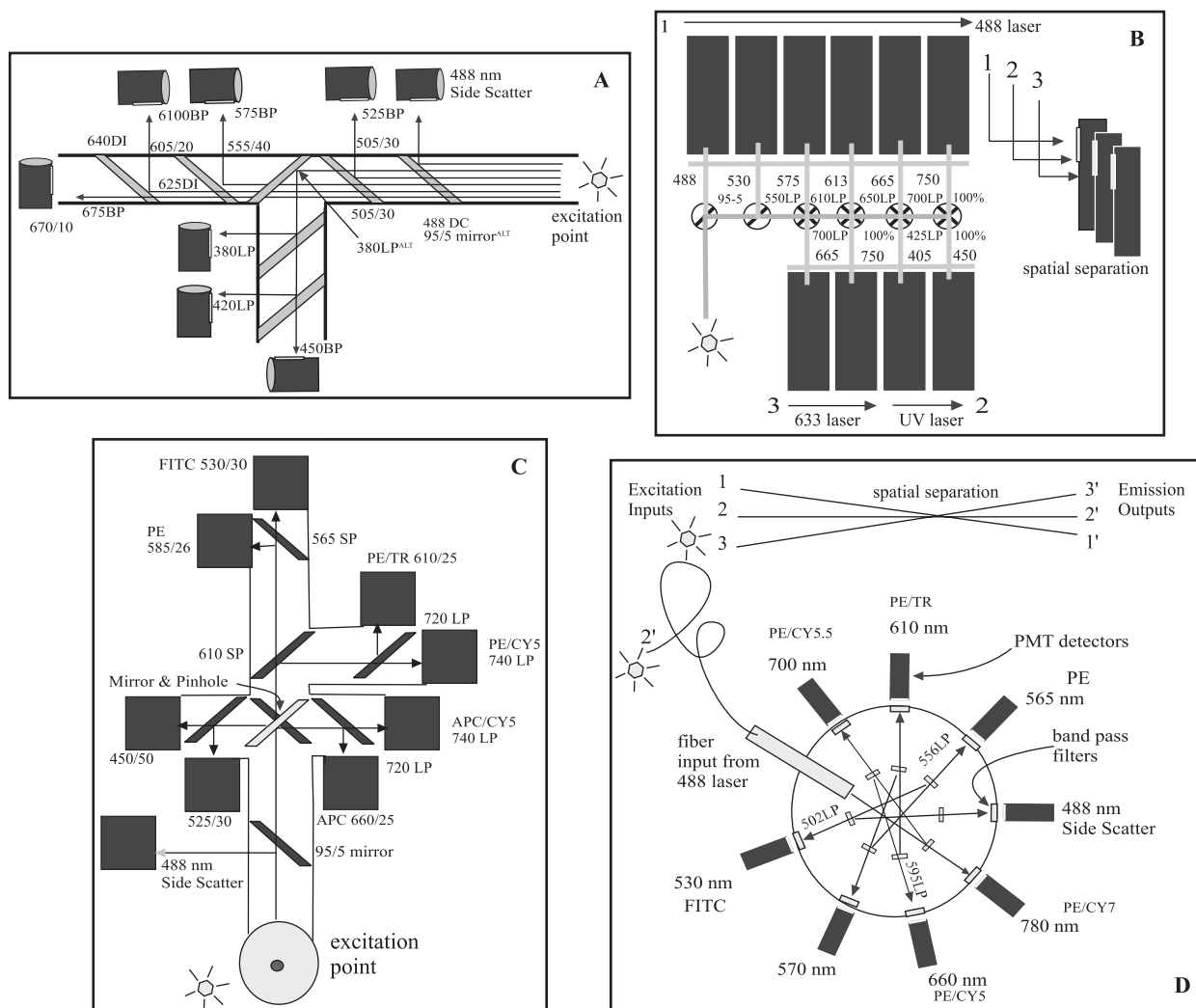


Fig. 4 A–D represent the optical tables of several commercial flow cytometers. A: Beckman-Coulter ALTRA showing the position of 8 PMTs; B shows the Dako-Cytomation CYAN instrument, which has 10 detectors placed in such a way that there are three beams with slightly different trajectories; C shows the Becton-Dickinson Vantage system in a typical configuration showing nine detectors; D is the more recent Becton-Dickinson ARIA system using an innovative PMT array with eight PMTs in a ring, which allows the emission signal to bounce around the ring. There are an additional six detectors on this system (not shown) that come from the first and third lasers (see diagram). In all cases A–D above, each PMT has a narrow bandpass filter immediately in front of the PMT in addition to the dichroic mirrors that are used to direct the various emission spectra.

patients. Because aerosols are generated in the normal operation of a flow cytometer, complex biosafety systems must be employed to reduce the potential of infection. There is a significant literature on the dangers posed by both microbes and carcinogenic molecules such as fluorescent dyes that are used to label many cells.^[20]

APPLICATIONS

Clinical Sciences

One of the largest applications of flow cytometry is in the clinical sciences, where the primary measurements

are of fluorochrome-conjugated antibodies bound to cellular receptors. This is generally referred to as immunophenotyping since many of the cell types being studied are immune cells such as lymphocytes. In fact, almost every possible human cell has been evaluated by flow cytometry. By far the most significant cell populations are peripheral blood cells such as red blood cells (RBC), white blood cells (lymphocytes, monocytes, and polymorphonuclear leukocytes), and platelets. Each of these populations presents some specific challenge in assay performance, but overall these cells are very amenable to flow analysis. A complex system of receptor identification has been developed within immunology to identify cellular receptors, which are referred to as Cluster of Differentiation (CD) antigens of which at

the time of writing there were 166 such classifications. These are based on similarity of antibody binding to specific receptors. Therefore, by conjugating fluorescent molecules to antibodies that recognize specific receptors, a population of cells binding that antibody and therefore that fluorescent molecule can be identified. With certain clinical syndromes, it is evident that a specific pattern will emerge when identifying which cells bind to certain antibodies. One of the most significant findings in the early 1980s was that the identification of certain subsets of human T cells was important for the monitoring of the clinical status of AIDS patients^[21]

(Fig. 5A and B). This significantly increased the utility of flow cytometry and drove the need for simple-to-operate, reliable clinical benchtop analyzers for basic two- and three-color immunofluorescence. These instruments now represent the great majority of flow cytometers in the field.

Cell biology

Some of the earliest studies of cell function investigated neutrophil function by measuring phagocytosis of microorganisms.^[22] This is an excellent example of the

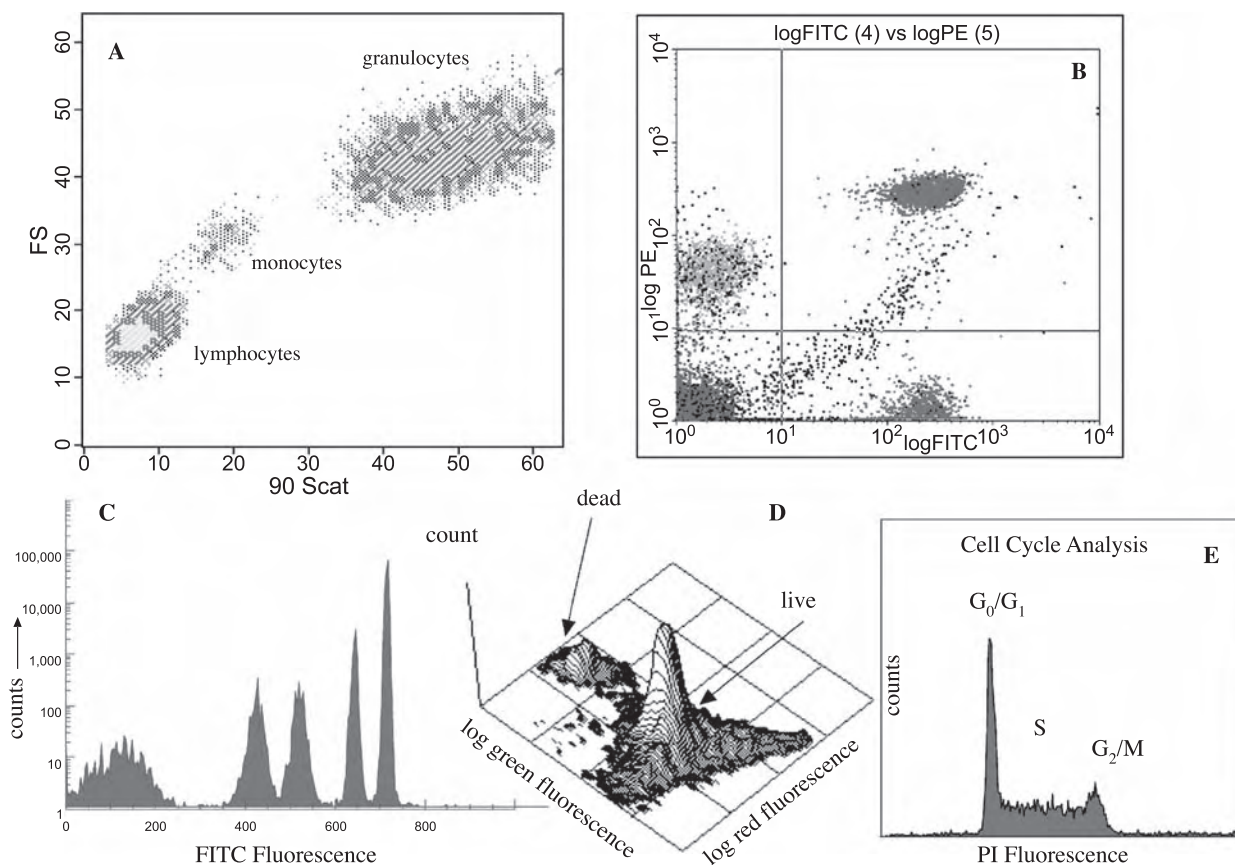


Fig. 5 A. When a cell passes through a laser beam, it scatters light. That light is measured on a detector and the resulting signal can provide information about the cells. Forward angle scatter (FS) is a measure of cell size. Side scatter (90° scat) is a measure of cellular components or granularity. In this dotplot of forward-versus-side scatter, human white blood cells can be differentiated without any other probes. Here is shown the separation of lymphocytes, monocytes, and granulocytes.

B. Gating strategies allow identification of populations of cells such as lymphocytes, shown in (A), the fluorescence emission of conjugated antibodies can be further separated to divide the lymphocytes into four distinct populations. In two-parameter space, the populations can easily be divided into four populations: those cells that are double negative, double positive, and single positive for each color. C. Calibration beads with fluorescent molecules attached to their surface are used to create quantitative measures for flow cytometry. This histogram has five peaks, the lowest peak being negative cells and the other four peaks represent four levels of fluorescence. From this histogram a standard curve can be obtained for quantitation of particles being labeled with this probe. D. This isometric display shows a plot of bacteria as observed by flow cytometry.

Pseudomonas aeruginosa is broth treated with 10× MIC of the antibiotic Imipenem for two hours and stained with BacLight Live/Dead kit. The log green fluorescence is Syto 9 and the log red fluorescence is PI. Positive PI fluorescence represents damage to the cell membrane, an indication of cell death. E. Propidium Iodide (PI) can also be used to study the cell cycle. In this case, the membrane is slightly damaged to allow penetration of the dye. PI binds to DNA in a stoichiometric manner such that there is a direct relationship between DNA content and PI fluorescence.

value of flow cytometry, which can identify individual cells by their size, structure, or specific identifiers such as cell receptors, and simultaneously measure the nature and number of microorganisms that were internalized via the process of phagocytosis. There are many applications of flow cytometry used for studying unique properties of cells that cannot easily be studied with any other technology. For example, real-time, single-cell production of oxygen radicals is frequently evaluated by using flow cytometry. There are a number of well accepted techniques from the earliest studies^[23,24] to more recent ones whereby both cells and organelles have been studied by flow cytometry.^[25,26] A huge number of applications exist in the field of DNA ploidy research (Fig. 5E). The ability to identify the rate of cell division and to monitor the effect of various therapeutic drugs is of great interest. Studies of the cell cycle by flow cytometry have provided a great deal of information on the nature of cell division and more recently apoptosis.^[27–29]

Microbiology

The study of microbes and their behavior is ideally suited to flow cytometry,^[30] however, there is an apparent disconnect between the capability of flow cytometry to answer microbial-related questions and its use in the field. Early studies quickly focused on the possibilities of developing flow-based assays for such time-consuming assays in the clinical environment as antibiotic resistance. With the growth of resistant organisms, determination of antibiotic resistance would be a desirable measure, but one that is rarely if ever performed outside the hospital environment. Even in the medical microbiology laboratory, it is still considered uneconomic, despite the clear demonstration that both organism identification and antibiotic sensitivity can be determined within a couple of hours. Unfortunately, the current cost far exceeds the pennies-per-test conditions set by current medical practices. This is definitely one of the potentials for microbiology-specific instruments that should markedly reduce testing costs for such applications (Fig. 5D).

Many studies of microbial kinetics have been performed using flow cytometry, including growth curves, reproduction studies, and metabolic requirements. In addition, exciting new studies are beginning to demonstrate new opportunities of flow cytometry together with advanced imaging tools for studying growth of microorganisms in complex 3-D environments such as biofilms.^[31]

Plant and Animal Science

Although a great majority of flow cytometry is related to human and laboratory animal systems, there are

some excellent examples of studies of plant systems. For example, it was recognized very early in the use of flow cytometry that cell cycle could be easily analyzed by flow cytometry,^[32] and this stimulated a number of cytometry-related plant-based studies.^[33] Pollen, for example, is perfectly suited to flow cytometry, as are plant chromosomes, even though they are somewhat more difficult to extract. A number of flow sorting experiments performed on plant systems to identify gene expression from transgenic tobacco plants^[34] demonstrated the efficacy of using this technology.

Pharmaceuticals

One of the more recent applications of flow cytometry is high-throughput screening. While there are many technologies that have far greater sample throughput, flow cytometry is one of the few technologies that can identify and analyze individual cells in multiple parameters. Recently, the concept of high-throughput cytometry was introduced, and initial reports suggest the possibility of achieving as many as 100,000 samples per day,^[35] something that approaches the needs of pharmaceutical manufacturers. The clear advantage of flow cytometry over other technologies, such as imaging cell-culture plates, is that with flow cytometry, a large number of parameters can be analyzed on each and every cell. The disadvantage is that flow cytometry, even with high-speed systems, is very much slower than automated image-processing systems.

Reproductive Medicine

Sperm analysis has proved the value of flow cytometry and especially the cell sorting capacity. There are several approaches to analysis of sperm. One utilizes the ability of DNA dyes such as Hoechst 33342 to bind to sperm DNA without inflicting damage,^[36] another uses antibodies to the H-Y antigen.^[37] The ability of flow cytometry to sort human sperm for sex selection raises a number of ethical questions. It is clearly well within the means of this technology to sex-select human sperm; although to date there are no published reports of this having been done, the topic is heavily discussed.^[38]

Calibration Issues

Because flow cytometry is defined as a quantitative technology, it is important to have calibration standards. These were primarily developed by Schwartz^[39] and others to allow reproducibility of clinical assays. Schwartz developed the concept of Molecule Equivalents of Soluble Fluorescein (MESF units). Using a mixture of beads with known numbers of fluorescent

molecules, it is possible to create a standard curve based on a least-squares regression based on the median fluorescence intensity of each bead population. This value is then converted into MESFs (Fig. 5C) from which comparisons can be made from different instruments or the same instrument on different days. Future instruments will most likely provide data in units such as MESFs rather than “arbitrary fluorescence values,” as are frequently observed in present-day publications. It would seem highly desirable to provide more quantitative data for comparison purposes.

CONCLUSIONS

The technology of flow cytometry has made a significant impact on many fields. There are few technologies that can evaluate so many parameters on such small samples in such short time periods. The principle of evaluating each and every cell or particle that passes through the laser beam and then producing a highly correlated data set is unique to flow cytometry. The combination with multivariate analysis and subsequent ability to separate cells physically by the process of cell sorting gives this technology some unique characteristics. It has been almost 40 years since flow cytometry first demonstrated its importance in medical research. Since that time, well over 60,000 publications have highlighted its usefulness. It was identified as one of the most important technologies in the early 1980s upon the recognition of AIDS. The ability of flow cytometry to identify and quantify the T cell population subsets CD4 and CD8 lymphocytes identified it as a most important technology in the diagnosis and monitoring of AIDS patients. Similarly, the ability of flow cytometry to make complex multivariate analyses of bone marrow to identify the CD34+ cells and subsequently sort and purify them has been a vital resource in transplantation immunology.

ARTICLES OF FURTHER INTEREST

Hematopoietic Stem Cells and Assays; Optics, Biomedical

REFERENCES

- Shapiro, H.M. Multistation multiparameter flow cytometry: A critical review and rationale. *Cytometry* **1983**, *3*, 227–243.
- Robinson, J.P.; Durack, G.; Kelley, S. An innovation in flow cytometry data collection and analysis producing a correlated multiple sample analysis in a single file. *Cytometry* **1991**, *12*, 82–90.
- Moldavan, A. Photo-electric technique for the counting of microscopical cells. *Science* **1934**, *80*, 188–189.
- Papanicolaou, G.N.; Traut, R. The diagnostic value of vaginal smears in carcinoma of the uterus. *Am. J. Obstet. Gynecol.* **1941**, *42*, 193.
- Kamentsky, L.A.; Melamed, M.R.; Derman, H. Spectrophotometer: New instrument for ultrarapid cell analysis. *Science* **1965**, *150*, 630–631.
- Kamentsky, L.A.; Melamed, M.R. Spectrophotometric cell sorter. *Science* **1967**, *156*, 1364–1365.
- Sweet, R.G. High frequency recording with electrostatically deflected ink jets. *Rev. Sci. Instrum.* **1965**, *36*, 131–136.
- Fulwyler, M.J. Electronic separation of biological cells by volume. *Science* **1965**, *150*, 910–911.
- Shapiro, H.M. *Practical Flow Cytometry*, 3rd Ed.; Wiley-Liss: New York, NY, 1994.
- Crosland-Taylor, P.J. A device for counting small particles suspended in fluid through a tube. *Nature* **1953**, *171*, 37–38.
- De Rosa, S.C.; Roederer, M. Eleven-color flow cytometry. A powerful tool for elucidation of the complex immune system. *Clin. Lab. Med.* **2001**, *21*, pp. vii, 697–712.
- Coulter, W.H. High speed automatic blood cell counter and cell size analyzer. *Proc. Natl. Electron. Conf.* **1956**, *12*, 1034–1042.
- Bagwell, C.B.; Adams, E.G. Fluorescence spectral overlap compensation for any number of flow cytometry parameters. *Ann. N.Y. Acad. Sci.* **1993**, *677*, 167–184.
- Dittrich, W.; Gohde, W. Impulsfluorometrie bei Einzelzellen in Suspensionen. *Z. Naturforsch.* **1969**, *24b*, 360.
- Bonner, W.A.; Hulett, H.R.; Sweet, R.G.; Herzenberg, L.A. Fluorescence activated cell sorting. *Rev. Sci. Instrum.* **1972**, *43*, 404–409.
- van den Engh, G. High Speed Cell Sorting. In *Emerging Tools for Single-Cell Analysis*; Durack, G., Robinson, J.P., Eds.; Wiley-Liss: New York, 2000; 21–48.
- Kachel, V.; Kordwig, E.; Glossner, E. Uniform lateral orientation of flat particles in flow through systems caused by flow forces. *J. Histochem. Cytochem.* **1977**, *25*, 774–780.
- Pinkel, D.; Stovel, R. Flow Chambers and Sample Handling. In *Flow Cytometry: Instrumentation and Data Analysis*; Dean, P.N., Ed.; Academic Press: New York, 1985.
- Andrews, R.G.; Singer, R.G.; Bernstein, I.D. Precursors of colony forming cells in humans can be distinguished from colony-forming cells by expression of the CD33 and CD34 antigens and light scatter properties. *J. Exp. Med.* **1989**, *169*, 1721–1731.
- Nicholson, J.K.A. Immunophenotyping specimens from HIV-infected persons: Laboratory guidelines from the Centers for Disease Control and Prevention. *Cytometry* **1994**, *18*, 55–59.
- Fahey, J.L.; Giorgi, J.; Martínez-Maza, O.; Detels, R.; Mitsuyasu, R.; Taylor, J. Immune pathogenesis of AIDS and related syndromes. *Ann. Inst. Pasteur., Immunol.* **1987**, *138*, 245–252.
- Bassoe, C.-F.; Solsvik, J.; Laerum, O.D. Quantitation of Single Cell Phagocytic Capacity by Flow Cytometry.

- In *Flow Cytometry IV*; Laerum, O.D., Lindmo, T., Thorud, E., Eds.; Universitetsforlaget: Oslo, 1980; 170.
23. Bass, D.A.; Parce, J.W.; DeChatelet, L.R.; Szejda, P.; Seeds, M.C.; Thomas, M. Flow cytometric studies of oxidative product formation by neutrophils: A graded response to membrane stimulation. *J. Immunol.* **1983**, *130*, 1910–1917.
 24. Rothe, G.; Valet, G. Flow cytometric analysis of respiratory burst activity in phagocytes with hydroethidine and 2',7'-dichlorofluorescein. *J. Leukoc. Biol.* **1990**, *47*, 440–448.
 25. Robinson, J.P.; Carter, W.O.; Narayanan, P.K. Oxidative Product Formation Analysis by Flow Cytometry. In *Methods in Cell Biology: Flow Cytometry*; Darzynkiewicz, Z., Robinson, J.P., Crissman, H.A., Eds.; Academic Press: New York, 1994; 437–447.
 26. Li, N.; Ragheb, K.; Lawler, G.; Sturgis, J.; Rajwa, B.; Melendez, J.A.; Robinson, J.P. DPI induces mitochondrial superoxide-mediated apoptosis. *Free Radic. Biol. Med.* **2003**, *34*, 465–477.
 27. Crissman, H.A.; Tobey, R.A. Cell-cycle analysis in 20 minutes. *Science* **1974**, *184*, 1297–1298.
 28. Darzynkiewicz, Z.; Traganos, F.; Sharpless, T.K.; Melamed, M.R. Cell cycle-related changes in nuclear chromatin of stimulated lymphocytes as measured by flow cytometry. *Cancer Res.* **1977**, *37*, 4635–4640.
 29. Ormerod, M.G. Investigating the relationship between the cell cycle and apoptosis using flow cytometry. *J. Immunol. Methods* **2002**, *265*, 73–80.
 30. Steen, H.B. Flow cytometry of bacteria: Glimpses from the past with a view to the future. *J. Microbiol. Methods* **2000**, *42*, 65–74.
 31. Sincock, S.A.; Robinson, J.P. Flow cytometric analysis of microorganisms. *Methods Cell Biol.* **2001**, *64*, 511–537.
 32. Galbraith, D.W.; Harkins, K.R.; Maddox, J.M.; Ayres, N.M.; Sharma, D.P.; Firoozabady, E. Rapid flow cytometric analysis of the cell cycle in intact plant tissues. *Science* **1983**, *220*, 1049–1051.
 33. Brown, S.C.; Bergounioux, C. Plant Flow Cytometry. In *Flow Cytometry: Advanced Research and Clinical Application*; Yen, A., Ed.; CRC Press: Boca Raton, 1990; 195–212.
 34. Harkins, K.R.; Jefferson, R.A.; Kavanagh, T.A.; Bevan, M.W.; Galbraith, D.W. Expression of photosynthesis-related gene fusions is restricted by cell type in transgenic plants and in transfected protoplasts. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 816–820.
 35. Kuckuck, F.W.; Edwards, B.S.; Sklar, L.A. High throughput flow cytometry. *Cytometry* **2001**, *44*, 83–90.
 36. Libbus, B.L.; Perreault, S.D.; Johnson, L.A.; Pinkel, D. Incidence of chromosome aberrations in mammalian sperm stained with Hoechst 33342 and UV-laser irradiated during flow sorting. *Mutat. Res.* **1987**, *182*, 265–274.
 37. Peter, A.T.; Jones, P.P.; Robinson, J.P. Fractionation of bovine spermatozoa for sex selection: A rapid immunomagnetic technique to remove spermatozoa that contain the H-Y antigen. *Theriogenology* **1993**, *40*, 1177–1185.
 38. Meseguer, M.; Garrido, N.; Remohi, J.; Simon, C.; Pellicer, A. Gender selection: Ethical, scientific, legal, and practical issues. *J. Assist. Reprod. Genet.* **2002**, *19*, 443–446.
 39. Schwartz, A.; Fernández-Repollet, E. Development of clinical standards for flow cytometry. *Ann. N.Y. Acad. Sci.* **1993**, *677*, 28–39.

Foreign Body Response

John Hunt

UKBioTEC, U.K. Centre for Tissue Engineering, Clinical Engineering, University of Liverpool, Liverpool, United Kingdom



INTRODUCTION

The term foreign body response presents itself as a simple nontechnical phrase that belies its complexity, interactions, and possibly dependence with other host responses. In terms of materials and implanted materials, the search for definitions can be limited to the organism being mammalian and the foreign material being that which has been introduced by deliberate intervention into or onto the body to address an improvement in the quality of life for the host. In this context, *The Williams Dictionary of Biomaterials* defines the foreign body response as the “Overall response of a host to the presence of a foreign body.” This is necessarily substantiated with notes that develop this: “The foreign body reaction is neither a single event nor a simple process but a broad concept and multifactorial phenomenon.”^[1] The supplementary note provides a significant insight towards the breadth and depth of the foreign body response.

The foreign body response without a body could be described as inflammation and wound healing. The components and factors involved have the potential to interact, stimulate, and act in synergy, in parallel, or as positive and negative feedbacks. It is often unclear when implantation is significantly invasive (involving a full-thickness incision wound of the skin) at what point the host response becomes the foreign body response and should be attributed to the implant, or the response is in actual fact a result of the surgical procedure itself. This dilemma highlights the overlap and involvement of the same systems and components from the host. This is compounded by the dual effect the combination of surgery and the application of a material can have on the overall host response.

The foreign body response is the human body's first line of defense against invasion and can be short or long term (minutes to years). The invading object is termed a foreign body originating from outside of the host. This can include many temporary invasions, for example, dust particulates in the lung that stimulate the involuntary reflexes of coughing or the peristaltic motion and wavelike movements from the cilia of the epithelium that line the bronchioles of the respiratory tract when moving inhaled particulates out of the lung and air passages into the trachea for removal by swallowing,

coughing, or in saliva. These reflexes do have some significance for the introduction and application of some materials and these examples would particularly relate to asthma treatments, inhaled materials, and further-to-fast delivery vectors for pharmaceutical agents.

The most commonly encountered foreign body response in material terms relates to the invasion of tissues by implanted foreign materials and the host's methods for dealing with these. In basic evolutionary terms, this involves both innate (nonspecific) immunity and adaptive (specific or acquired) immunity mechanisms of increasing complexities to combat micro-organism invasion, infection, and wounding. A good example to relate the principles of the foreign body response first hand would be to consider a body's response to a splinter (macroscopic fragment of material).

From the evolution of our species, many of the mechanisms involved in our defense can be related to our animal needs and the fight-or-flight, life-or-death scenario of our ancestors. Fundamentally as animals, if we become ill or injured, the body's first priority is to enable us to maintain mobility to escape danger and then to recover quickly to acquire food. As a complex high society far removed from our animal origins, we are unsatisfied with some of the potential longer term outcomes of these impressive rapid repairs. They leave scar tissue (skin and soft tissue) and possibly misshaped bone (skeletal); a serious consequence of these two end stages is that they can result in the body being susceptible to further injury or with a shortened life expectancy due to long-term stresses induced on the body because of vigorous unguided healing. Modern medicine in most circumstances avoids these extremes and they generally rarely become a life expectancy issue. In material terms, prior treatment including good operating room preparation can in most circumstances stabilize the host, and the material can be deployed in safe, sterile conditions. Nevertheless, these points remain functional issues for both device and host.^[2]

WHAT IS INVOLVED IN THE FOREIGN BODY RESPONSE?

The foreign body response can involve and encompass many details from host responses that can be included

in other terms: blood-contacting devices, thrombus formation, and complement activation,^[3] stimulation of inflammation,^[4] recruitment of cells from blood,^[5] and the release of protein signaling factors that act both locally and systemically.^[6,7] A foreign body response to a material can equally include first instance (seconds) effects such as protein adsorption^[8] or much later effects, up to years after implantation, such as chronic granuloma.^[9] Each of the individual host response components affect the outcome and performance in material terms, and the potential of many factors as single events has been and continues to be researched and documented in increasing detail.^[10]

The foreign body response as a route to respond to a breach of the skin-tissue barrier, which is the host's front line of defense, is a good specific area in which to develop and discuss many of the significant aspects of the host's foreign body response. As the skin is breached, there is an immediate response, a fantastic and complex sequence of events are triggered to combat the inevitable loss of blood from the system. This is non-specific and acts effectively as a quick sealant system despite the complexity of the sequence of events. Platelets are activated and aggregate at the site of injury, forming a temporary platelet plug. Fibrinogen stimulates the platelets to clump and the platelets change shape to a flattened disclike morphology. A fibrin mesh forms by the activation of fibrinogen by thrombin to trap the platelets. There are two potential pathways which lead to the formation of a fibrin clot: the intrinsic (in the absence of tissue injury) and extrinsic (in response to tissue injury) pathways.^[11,12] The formation of a thrombus also effectively reduces the rate of invasion by microorganisms by limiting the free access to what many organisms find ideal material to live in and on.^[13]

The foreign body response in its first phases utilizes the host's innate defense mechanisms that have evolved to deal with the invasion of microorganisms and are triggered by changes in the hemostasis. The complement system comprises a set of proteins circulating in the serum of blood called the complement proteins. These exist in a stable (albeit fragile) inactivated state in normal circulating blood; they are very sensitive to change. Once activated, the complement system fires in a cascade in which one component activates the next, taking only minutes to coat materials or microorganisms.^[14] The binding of C3b (an activated component of complement) or IgG (immunoglobulin G) to the surface of a foreign bacteria or material labels the particle as nonhost and targets it for attack by phagocytic cells.^[15,16] The phagocytes form the next waves of the foreign body response and these are also termed inflammatory responses in other contexts.^[17] Escaping from the blood into tissue when the capillaries are damaged, inflammatory cells can have

autocrine and paracrine effects, activating themselves and also stimulating the recruitment from undamaged blood vessels with increased vascular permeability of further phagocytic cells and stimulating the release of factors from resident tissue mast cells.^[18] A significant role of the cells is to phagocytose (cell eat) material labeled as foreign or to phagocytose damaged cellular material. This response appears to be potentially quite specific and can be related to opsonization, but it has equally been shown that this response is nonspecific: Once these cells are activated, they will try to phagocytose most things in the area, as there is an abundance of IgG^[19] and activated complement.

Opsonization is the process of coating microorganisms with plasma proteins to increase their visibility as target molecules and to enhance the adhesion of phagocytic cells in preparation for phagocytosis. There are two main opsonins: IgG antibodies and C3b, which bind to the surfaces of microorganisms. Phagocytic cells have membrane receptors for IgG (Fc piece) and C3b. IgG antibody binds to microorganisms because the Fab portion recognizes bacterial epitopes. C3b binds because during activation of C3 a thiol bond is exposed that permits the formation of ester or amide bonds (covalent binding) with the bacterial surface structures.^[20]

Neutrophils and monocytes/macrophages are the two most important blood-derived phagocytic cells. Neutrophils and monocytes circulate in the blood and migrate into the tissues at the postcapillary venule through the process of diapedesis (monocytes are the immature circulating form of the macrophage cell). Initially they adhere to endothelium and then migrate through intercellular junctions. In the tissues, cells migrate toward foreign materials by chemotaxis, which is defined as directed movement along a concentration gradient of a chemical agent (for example, interleukin 8, IL-8). Phagocytes recognize their targets by specific sugar residues (e.g., mannose or Lipopolysaccharides), but binding is greatly enhanced if the organism is opsonized with IgG and or C3b. Phagocytes possess Fc and C3b receptors and there is cooperation between these receptors.^[21] A material or organism opsonized with IgG and C3b is more effectively phagocytosed.^[22-24]

PHAGOCYTOSIS OF MATERIAL

In the early phase, material is phagocytosed by large numbers of neutrophils, which are found in high numbers in normal circulating blood—in fact, they are the most abundant cells and get trapped in the coagulation response, and are also the fastest migrating cells—so they can move into an area from the cardiovascular system quickly. These nonproliferative cells are on a suicide mission to phagocytose bacteria; once engulfed, the phagocytic vacuole is fused with

lysosomal vacuoles in the cytoplasm. Lysosomal vacuoles contain many destructive and degradative agents. Lysozyme is an enzyme that specifically breaks down the cell wall of Gram-negative bacteria, effectively lysing the cell. These vacuoles also contain collagenase, neutral proteases, acid hydrolyses, and proteolytic enzymes, in fact: a good range of aggressive agents to attack bacteria cell membranes. These agents are nonspecific and the cell controls their action only by localization while the agents are contained within vacuoles inside the cell cytoplasm. If released into the extracellular fluid, these mechanisms can have significant implications for the host tissue. These cells also have the capacity to produce oxygen free radicals (electron-rich molecules) to compromise biological membranes by introducing free highly unstable and therefore reactive species into an area.^[25] These oxygen radicals are nonspecific and can react with the host tissue causing cell death and localized tissue damage, the evolutionary idea being to clear an area of invading microorganisms and infected tissue completely before they reproduce into a significant population. This free radical production and release also has much further reaching implications and can lead to accelerated degradation of implanted materials, or actually presents a means by which previously considered biostable materials can be demonstrated to degrade. So significant is this potential that ISO 10993 testing standards require materials to be tested for degradation in free-radical-rich solutions to evaluate this parameter.

After neutrophils have performed their phagocytic function, they themselves die and become debris that requires clearing from tissue. This leads on to the recruitment of a second slower moving phagocytic cell called the macrophage and this is the first cell that has been historically associated with a foreign body response.^[26] This cell can be long-lived in tissue, which is important only because a small percentage (<5%) have the potential to proliferate. The macrophage can exert significant effects in tissues by triggering many further responses.^[27] Equipped with the same innate mechanisms as neutrophils to phagocytose and release free radicals as well as degradative enzymes, these cells also come with the ability to present antigen to other cells and can be called antigen-presenting cells (APC) and be involved in the specific immune response to infection or invasion.

In an inflammatory context a sustained macrophage response is typical of a chronic inflammatory response and many implanted materials can and do endure a chronic inflammatory response for the duration of their implanted working life. In terms of a foreign body response, this sustained macrophage response to an implanted material forms the end-point foreign body response. This is the stage in the host response where for many scientists the foreign body response

begins,^[28] and this can be specifically attributed to the accumulation and persistence of macrophages in an area and their development into multinucleated giant cells. The multinucleated giant cell, or foreign body giant cell (FBGC) as it is also called, has been the subject of much research and debate and still remains poorly understood, perhaps due to its susceptibility to spontaneous lysis when compromised and further activated; these cells have been observed and been found to be absent in almost identical situations and experiments. The research debate about how these cells form and why they form also continues. There is good evidence and theory that foreign body giant cells exist in the context of the following explanation of their origins, and this term FBGC can perhaps be accepted: The macrophage like the neutrophil, when in contact with an unrecognized material or a material labeled as foreign by opsonization, will attempt to phagocytose the material. In the case of individual bacteria, this is easily possible but becomes more of a challenge for clustered populations and even more of a challenge in the context of the scale of many implanted prosthesis. Phagocytic cells actually perform the process of phagocytosis of a foreign material by enveloping it and forming a vacuole around it, which is derived from the cell's own plasma membrane and could be called invagination. The cell extends cytoplasmic extensions around the particle called pseudopodia and/or lamellapodia; these extend around the particle, remaining in contact with it as they develop. At some point, the cell hopes to meet itself around the other side and once again fuse its plasma membrane and complete the sphere that becomes an internalized vacuole that contains the material. The whole process is very similar to the way the single-celled eukaryotic animal the amoeba acquires its nutrients. Clearly, a single cell cannot stretch itself infinitely and there is a point at which it cannot increase in area any further. The macrophage and the neutrophil respond by attempting to achieve this process but if unsuccessful (frustrated phagocytosis), they will release their vacuole contents extracellularly and hopefully onto the surface of the target substance or material; attack then proceeds on the substance outside of the cell. This has critical implications for the cell and the surrounding host tissue. Many of the host defense mechanisms have the ability to affect living tissue and will attack the surrounding cells and matrix, breaking this up also and creating more debris.^[29] There is evidence that when macrophages contact materials, they can attach and form a sealed interface with the material and actually create an enclosed environment underneath themselves in which to release their vacuoles. This can be compared to the mechanism by which osteoclasts attach to the surface of bone and resorb it. In this case, macrophages will attempt impressive aggressive extremes and

can reduce the local pH to a strong acid pH of around 3. This is a very hostile environment indeed for most invading pathogenic microorganisms and certainly for some implanted materials. It has variable levels of success for implanted materials and these are, of course, material specific.^[30]

FOREIGN BODY GIANT CELL RESPONSE

We should return to the concept of the FBGC and develop this further. If macrophages that are going through this process of frustrated phagocytosis come into contact with each other, there is the potential for them to combine forces and their plasma membranes to fuse to form a double nucleated cell that has only one complete cytoplasm. This can continue to involve further adjacent cells and 6-, 10-, and even 20-nuclei giant cells have been observed; researchers have reported on FBGCs with over 200 nuclei. This is hard to accept; to actually be able to see and measure a single cell body of this size as one and determine that it is in fact a single whole unit would be very difficult. Given FBGCs' sensitivity to change and lytic response, FBGCs on this scale are unlikely to be observable.^[31] The macrophage and giant cell response can be the long-term end stage of the body's response to a material, and this situation can persist for years in a fragile balance that can be triggered to return to an acute inflammatory response by seemingly small changes in the material, maybe by factors such as the material's position, shape, or chemistry.^[6,32]

FIBROBLAST

Historically, in the context of the pursuit of a truly inert material, the end-stage acceptance point has often been the encapsulation of the material by fibroblasts and their extracellular matrix, which is predominantly collagen. This is a poor end point, in function and design terms but it has been the fate of many promising and actually widely used implanted devices. The macrophage and fibroblast in chronic inflammation will release chemotatic factors that will recruit fibroblasts to an area seemingly to accept the presence of a material that cannot be removed.^[33] It perhaps no longer carries a threat or stimulus in terms of carrying microorganisms, and many of the more aggressive nonspecific mechanisms are switched off in macrophages, and fibroblasts displace them as the principal cell and start laying down collagen-based extracellular matrix.^[34] This is predominantly type I and type III collagen, and the ratio varies depending on movement of the device and the age of the response, more type III being present at early time periods than later. There are

very few materials in which this can be accepted as the final biocompatible response,^[33] for example, pacemaker leads or the pacemaker driver unit although there are limitations even to these examples. In both cases, a positioning of the device in soft tissue is required, but the device itself has no mechanical or delivery implications that rely on holding tissue in position or articulating. This is somewhat oversimplified and can sometimes lead to a secondary set of inflammatory complications from physical irritation or possibly in other cases even carcinomas.^[35] An example of this fibrous encapsulation resolution point being a serious problem is in the region between the femur and the femoral stem component of a hip prosthesis. The thinnest layer of fibrous soft tissue in this region presents a zone in which loads cannot be transferred efficiently; slipping and movement under load can occur and progressive loosening follows. To complicate this example with other issues, it is not accepted by all research groups as to which comes first, the soft tissue or the loosening, and a good case could be presented for either scenario.^[36]

This complex scenario needs to be updated a little, however, and studies must begin to address the non-contacting form of cellular communication. In doing this, some of the detail becomes controversial because there is conflicting or parallel evidence.^[37] Cells communicate (on the cellular scale) over large distances in the human body using protein messages, signals, and cell-membrane-bound protein receptors.^[38] Cells also communicate with each other at a direct-contact level using proteins, but the two routes are not mutually exclusive. The potential to involve cells from any other part of the body in another part is clearly evident when cells can produce and release proteins into tissue that find their way into the cardiovascular and lymphatic systems and get transported around the body. There are many protein groups identified and many being redefined from previously classified ones.^[39] Growth factors, chemokines, and cytokines form the main groups with direct significance in the foreign body response, in addition to of course, the already mentioned complement proteins and the proteins that compose the coagulation cascade.^[40] The list of significant factors that could be mentioned is extensive and continues to be extended, but examples should certainly include platelet-derived growth factor, Interleukin-1, 4, 6, 8, 10, 12, and 15, Tumour necrosis factor α ,^[41] and transforming growth factors β -1 and 3.^[42,43]

WHAT IS THE FOREIGN BODY RESPONSE?

The first prerequisite: for a foreign body response is a material; the second is an initiation of the wound-healing inflammatory response. The third is persistence

in terms of time. Much of what has been discussed as being involved in the foreign body response also applies to the various phases of an inflammatory response and also to the phases of wound healing. The foreign body response exists only for as long as there is a foreign body present to be responded to. If one is not present, then the cellular response would be described as an inflammatory response. Much of the detail is identical and it is only at the point where the origin of the stimulus can be defined that the foreign body response finds a distinction and therefore perhaps a definition. In historical terms, the resolution of this in cell terms to fibroblast encapsulation of an implanted material signifies the final fate and best case acceptance of the material with the host. Even at this point comparisons can be found in inflammation at the chronic inflammation phase and subsequent fibrous encapsulation.^[44]

WHAT STIMULATES A FOREIGN BODY RESPONSE?

Even the smallest particulates from an implanted material can cause inflammation. If the particulates persist and with them inflammation, this will be termed the foreign body response at some time point and discussed in that context. This becomes much more significant from a materials perspective as the dimensions or complexity of a device increase,^[45] which leads to the bigger questions of what causes a foreign body response? What is required to stop it, or control it? The implications are considered much more serious if developing materials and implanting devices are said to provoke a foreign body response rather than inflammation. In reality, as discussed, there is very little difference.^[46] It is true that a material or device can fail due to the foreign body response and as such is a serious issue.^[47] A material can become completely encapsulated in many layers of fibrous tissue such that it becomes a rigid structure or hardened structure. This new structure may become inflammatory because of these new physical properties and start to manifest itself as a problem for the patient in terms of pain and or swelling. Another alternative is that the encapsulation effectively restricts movement of the material and the same scenario follows.^[48] The recurrence or stimulus to cause inflammation is a common issue for a presented failing device due to a foreign body response. The hardening of the soft tissue surrounding soft tissue breast prosthesis is perhaps one of the clearest examples of a device failing due to a foreign body response and the extremes that this can lead to. In these cases, in actual fact, there is often very little wrong with the material the device is made from, but it is a situation in which inflammation has

progressed to a point where there has been a significant amount of fibroblast activity and fibrous extracellular matrix tissue formation.^[49] This matrix has become less flexible than the surrounding tissue and therefore caused some mechanical mismatches to occur in the surrounding tissue, which have led to inflammation by irritation.

The body is very susceptible to mechanical stimulus and not particularly well adapted to constant abrasion/friction. Necessarily there are some exceptions such as articular cartilage, but in general, unless conditioned and hardened, even the skin when abraded will blister, delaminate, and break down. It is true that there are many factors in material terms that could affect inflammation in a stimulatory manner. Research has demonstrated the significance of implant size, shape, and form with respect to the host response.^[50] This has provided increasing insight into how materials could affect the host when contacting soft tissue.^[51] The issues are all too obvious: Soft tissue and hard implants present a serious problem in an environment where even the slightest mismatch in the transfer of forces is sufficient to cause inflammation, often by the simple mechanism of irritation by friction.^[52] Good stability and fixation are well understood for orthopedic applications, and even here materials demonstrate limitations in terms of stress shielding and bone resorption. In soft tissue, however, the importance and significance is often lost.^[53]

A body is in motion as a loose water-based structure, hung on a semirigid frame (skeleton); any obstruction to the smooth fluid motion within tissues and of tissues themselves will cause inflammation. The skeleton itself is vulnerable to attack by inflammation, but in normal circumstances the hard tissue of bone does not interfere with movements in the surrounding soft tissue. The term movement is not exclusive to contraction and relaxation in muscles; all tissues move or have the potential to be moved, stretched, or compressed.^[54] Our splinter example in the skin is the simple case with which to clarify the significance of even the smallest additional object in soft tissue. A fragment of material can be introduced into the skin and the wound completely healed over with no consequences of infection. The particle can go unnoticed until the skin in that area is compressed or moved, which results in the onset of inflammation, possibly pain if firing a nerve, but more likely pain will result as an indirect result of inflammation. In this case inflammation is caused by the material being moved differently compared to the surrounding cells and causing localized cellular damage, which actually breaks up the cells mechanically. The dead/dying cells stimulate the body's defense mechanisms and cells infiltrate to investigate and clean up, causing swelling and pain to follow.

Poly(lactic acid) (PLA) devices have been designed to be mechanically strong for a finite period of time to enable bone healing. The device then starts to degrade and becomes progressively weaker introducing increasing load and function to the healing bone and actually aiding in the complete bone healing. A positive design benefit is that because it is degradable, the device does not have to be removed. The devices have demonstrated good initial success as fixation devices, with good stabilization of the bone.^[55] Quite unexpectedly, though, at a later time, sometimes years, the degrading plate causes inflammation,^[56] irritation, and pain and the patient re-presents with swelling at the site of the plate. This has been attributed to a significant pH change in the surrounding tissue due to the degradation of the PLA to lactic acid^[57] or due to the release of particulates during the materials degradation.^[58,59] By either mechanism this would demonstrate the failing of a device by the onset of inflammation and irritation and certainly should be termed a significant foreign body response.^[60]

CONTROLLING THE FOREIGN BODY RESPONSE

There are pharmaceutical approaches to controlling inflammation and the foreign body response, but these have yet to be realized in terms of adding crucial secondary function to implanted materials. The introduction and successful delivery of a pharmaceutical agent as a secondary role from a device is more easily envisaged in tissue-engineering applications compared to applying delivery systems to materials that historically have been mainly inorganic solids and in general impermeable.

Materials are required that enable the surrounding tissue to move and to move the material as required in terms of temporary displacements. The concept of a significant inflammatory response and the resolution by collagen encapsulation is not an acceptable scenario for tissue-engineered devices, where good immediate and long-term communication between the introduced cells and the cellular support services of the host in terms of providing nutrients and removing waste is essential. The potential to trigger any component of the inflammatory response during implantation or application as well as part of the function is an ever present possibility that is brought to the forefront when the device or material is genuinely a foreign body. If the material does not cause localized cell damage by physically damaging the surrounding cells, a significant step forward will have been made towards avoiding the establishment of an environment that does not favor the foreign body response, though the

question of how to maintain a directed function will still be an issue.

CONCLUSION

From the instant a material is placed in contact with the body, there is the stimulation and onset of the foreign body response. What determines whether this will develop into something that will cause failure of the material itself or significant inflammation and discomfort depends on what is triggered in the inflammatory and wound-healing pathways and how prolonged the stimulus for this is maintained. There is a point at which the tissues of a body seem to give up aggressively combating a material, which is the final fibrous encapsulation. This is a very poor end point and should in a healthy body lead to its own set of stimuli that will perpetuate the inflammatory response.

Acute inflammation, chronic inflammation, and fibrous encapsulation all form part of the foreign body response. Blood coagulation, complement activation, and the production and release of growth factors and cytokines also form part of the foreign body response. These factors all have significance in other body functions and are not unique to a foreign material. The pathways and triggers can easily be initiated by other factors. These factors get attributed to a foreign body response only when there is a foreign body present as the defining stimulus.

The host's foreign body response is inevitable, and can as result of this be predicted and anticipated. Many of the factors considered detrimental in terms of inflammation and damage to the host and material themselves can be utilized and brought to bear positive and beneficial effects on the material and device or cell that they contain.^[61] Attempts to defeat the foreign body response have been met with little success. The efficient and correct progression is to utilize this response and develop and design materials accordingly.

ARTICLES OF FURTHER INTEREST

Biocompatibility Testing; Biomaterials: Protein-Surface Interactions; Blood-Material Interactions; Host Reactions; Infection of Medical Devices; Thrombosis

REFERENCES

1. Williams, D.F. *The Williams Dictionary of Biomaterials*; Liverpool University press: Liverpool, U.K., 1999.
2. Duskova, M.; Kozak, J.; Mazanek, J.; Smahel, Z.; Vohradnik, M. Augmentation of facial skeleton with ceramics in congenital disorders and in post-traumatic

- or postoperative deformities: A preliminary report. *Eur. J. Plast. Surg.* **2000**, *23* (2), 57–63.
3. Bahulekar, R.; Tamura, N.; Ito, S.; Kodama, M. Platelet adhesion and complement activation studies on poly(N-alkyl mono and disubstituted) acrylamide derivatives. *Biomaterials* **1999**, *20* (4), 357–362.
 4. Pu, F.R.; Williams, R.L.; Markkula, T.; Hunt, J.A. Effects of plasma treated PET and PTFE on expression of adhesion molecules by human endothelial cells in vitro. *Biomaterials* **2002**, *23* (11), 2411–2428.
 5. Pu, F.R.; Williams, R.L.; Markkula, T.; Hunt, J.A. Expression of leukocyte-endothelial cell adhesion molecules on monocyte adhesion to human endothelial cells on plasma treated PET and PTFE in vitro. *Biomaterials* **2002**, *23* (24), 4705.
 6. Griffiths, M.M.; Langone, J.J.; Lightfoote, M. Biomaterials and granulomas. *Methods* **1996**, *9* (2), 295–304.
 7. Zeller, J.M. Surgical implants. Physiological response. *AORN J.* **1983**, *37* (7), 1284–1291.
 8. Keogh, J.R.; Velander, F.E.; Eaton, J.W. Albumin-binding surfaces for implantable devices. *J. Biomed. Mater. Res.* **1992**, *26* (4), 441–456.
 9. Collier, T.O.; Anderson, J.M. Protein and surface effects on monocyte and macrophage adhesion, maturation, and survival. *J. Biomed. Mater. Res.* **2002**, *60* (3), 487–496.
 10. Pawlaczyk, K.; Kuzlan-Pawlaczyk, M.; Anderstam, B.; Heimburger, O.; Bergstrom, J.; Waniewski, J.; Breborowicz, A.; Lindholm, B. Effects of intraperitoneal heparin on peritoneal transport in a chronic animal model of peritoneal dialysis. *Nephrol. Dial. Transplant.: Off. Publ. Eur. Dial. Transpl. Assoc.—Eur. Renal Assoc.* **2001**, *16* (3), 669–671.
 11. Courtney, J.M.; Lamba, N.M.; Sundaram, S.; Forbes, C.D. Biomaterials for blood-contacting applications. *Biomaterials* **1994**, *15* (10), 737–744.
 12. Broberg, M.; Eriksson, C.; Nygren, H. GpIIb/IIIa is the main receptor for initial platelet adhesion to glass and titanium surfaces in contact with whole blood. *J. Lab. Clin. Med.* **2002**, *139* (3), 163–172.
 13. Nimeri, G.; Ohman, L.; Elwing, H.; Wettero, J.; Bengtsson, T. The influence of plasma proteins and platelets on oxygen radical production and F-actin distribution in neutrophils adhering to polymer surfaces. *Biomaterials* **2002**, *23* (8), 1785–1795.
 14. Blondin, C.; Chaubet, F.; Nardella, A.; Siquin, C.; Jozefonvicz, J. Relationships between chemical characteristics and anticomplementary activity of fucans. *Biomaterials* **1996**, *17* (6), 597–603.
 15. Kopp, R.; Mottaghy, K.; Kirschfink, M. Mechanism of complement activation during extracorporeal blood-biomaterial interaction: Effects of heparin coated and uncoated surfaces. *ASAIO J.* **2002**, *48* (6), 598–605.
 16. Tang, L.; Liu, L.; Elwing, H.B. Complement activation and inflammation triggered by model biomaterial surfaces. *J. Biomed. Mater. Res.* **1998**, *41* (2), 333–340.
 17. Davey, P.C.; Zuzel, M.; Kamiguti, A.S.; Hunt, J.A.; Aziz, K.A.; Kamiguti, A.S. Activation-dependent proteolytic degradation of polymorphonuclear CD11b. *Br. J. Haematol.* **2000**, *111* (3), 934–942.
 18. Tang, L.; Jennings, T.A.; Eaton, J.W. Mast cells mediate acute inflammatory responses to implanted biomaterials. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95* (15), 8841–8846.
 19. Schlosser, M.; Wilhelm, L.; Urban, G.; Ziegler, B.; Ziegler, M.; Zippel, R. Immunogenicity of polymeric implants: Long-term antibody response against polyester (Dacron) following the implantation of vascular prostheses into LEW.1A rats. *J. Biomed. Mater. Res.* **2002**, *61* (3), 450–457.
 20. Wettero, J.; Askendal, A.; Bengtsson, T.; Tengvall, P. On the binding of complement to solid artificial surfaces in vitro. *Biomaterials* **2002**, *23* (4), 981–991.
 21. Toufik, J.; Carreno, M.-P.; Jozefowicz, M.; Labarre, D. Activation of the complement system by polysaccharidic surfaces bearing carboxymethyl, carboxymethylbenzylamide and carboxymethylbenzylamide sulphonate groups. *Biomaterials* **1995**, *16* (13), 993–1002.
 22. Andersson, J.; Ekdahl, K.N.; Larsson, R.; Nilsson, U.R.; Nilsson, B. C3 adsorbed to a polymer surface can from an initiating alternative pathway convertase. *J. Immunol.* **2002**, *168* (11), 5786–5791.
 23. Jenney, C.R.; Anderson, J.M. Adsorbed IgG: A potent adhesive substrate for human macrophages. *J. Biomed. Mater. Res.* **2000**, *50* (3), 281–290.
 24. Kao, W.J. Evaluation of leukocyte adhesion on polyurethanes: The effects of shear stress and blood proteins. *Biomaterials* **2000**, *21* (22), 2295–2303.
 25. Kaplan, D.S.; Picciolo, G.L.; Mueller, E.P. Real-time monitoring of oxidative products from in vitro cell-biomaterial interaction using chemiluminescence. *Cell Transplant* **1995**, *4* (4), VIII.
 26. Hunt, J.A.; Abrams, K.R.; Williams, D.F. Modelling the pattern of cell distribution around implanted materials. *Anal. Cell. Pathol.: J. Eur. Soc. Anal. Cell. Pathol.* **1994**, *7* (1), 43–52.
 27. Xing, S.; Waddell, J.E.; Boynton, E.L. Changes in macrophage morphology and prolonged cell viability following exposure to polyethylene particulate in vitro. *Microsc. Res. Tech.* **2002**, *57* (6), 523–529.
 28. Anderson, J.M. Inflammatory response to implants. *ASAIO Trans./Am. Soc. Artif. Int. Org.* **1988**, *34* (2), 101–107.
 29. Hunt, J.A.; Remes, A.; Williams, D.F. Stimulation of neutrophil movement by metal ions. *J. Biomed. Mater. Res.* **1992**, *26* (6), 819–828.
 30. Hunt, J.A.; Flanagan, B.F.; McLaughlin, P.J.; Strickland, I.; Williams, D.F. Effect of biomaterial surface charge on the inflammatory response: Evaluation of cellular infiltration and TNF[alpha] production. *J. Biomed. Mater. Res.* **1996**, *31* (1), 139–144.
 31. Brodbeck, W.G.; Shive, M.S.; Colton, E.; Ziats, N.P.; Anderson, J.M. Interleukin-4 inhibits tumor necrosis factor-alpha-induced and spontaneous apoptosis of biomaterial-adherent macrophages. *J. Lab. Clin. Med.* **2002**, *139* (2), 90–100.
 32. Hagerty, R.D.; Salzman, D.L.; Kleinert, L.B.; Williams, S.K. Cellular proliferation and macrophage populations associated with implanted expanded polytetrafluoroethylene and polyethyleneterephthalate. *J. Biomed. Mater. Res.* **2000**, *49* (4), 489–497.
 33. Yaszay, B.; Trindade, M.C.D.; Lind, M.; Goodman, S.B.; Smith, R.L. Fibroblast expression of C-C

- chemokines in response to orthopaedic biomaterial particle challenge in vitro. *J. Orthop. Res.: Off. Pub. Orthop. Res. Soc.* **2001**, *19* (5), 970–976.
34. Williams, R.L.; Hunt, J.A.; Tengvall, P. Fibroblast adhesion onto methyl–silica gradients with and without preadsorbed protein. *J. Biomed. Mater. Res.* **1995**, *29* (12), 1545–1555.
 35. James, S.J.; Pogribna, M.; Miller, B.J.; Bolon, B.; Muskhelishvili, L. Characterization of cellular response to silicone implants in rats: Implications for foreign-body carcinogenesis. *Biomaterials* **1997**, *18* (9), 667–675.
 36. Baldwin, L.; Flanagan, B.F.; McLaughlin, P.J.; Parkinson, R.W.; Hunt, J.A.; Williams, D.F. A study of tissue interface membranes from revision accord knee arthroplasty: The role of T lymphocytes. *Biomaterials* **2002**, *23* (14), 3007–3014.
 37. Brodbeck, W.G.; Nakayama, Y.; Matsuda, T.; Colton, E.; Ziats, N.P.; Anderson, J.M. Biomaterial surface chemistry dictates adherent monocyte/macrophage cytokine expression in vitro. *Cytokine* **2002**, *18* (6), 311–319.
 38. Harrington, P.M.; Newton, D.J.; Williams, C.M.M.; Hunt, J.A.; Dearman, R.J.; Kimber, I.; Coleman, J.W.; Flanagan, B.F. Eotaxin and eotaxin receptor (CCR3) expression in Sephadex particle-induced rat lung inflammation. *Int. J. Exper. Pathol.* **1999**, *80* (3), 177–185.
 39. Jiang, X.; Marois, Y.; Traore, A.; Tessier, D.; Dao, L.H.; Guidoin, R.; Zhang, Z. Tissue reaction to polypyrrole-coated polyester fabrics: An in vivo study in rats. *Tissue Eng.* **2002**, *8* (4), 635–647.
 40. Kao, W.J.; Liu, Y.; Gundloori, R.; Li, J.; Lee, D.; Einerson, N.; Burmania, J.; Stevens, K. Engineering endogenous inflammatory cells as delivery vehicles. *J. Control. Release: Off. J. Control. Rel. Soc.* **2002**, *78* (1–3), 219–233.
 41. Thomas, A.; Harding, K.G.; Moore, K. Alginates from wound dressings activate human macrophages to secrete tumour necrosis factor-[alpha]. *Biomaterials* **2000**, *21* (17), 1797–1802.
 42. Corry, D.C.; Delucia, A., III; Zhu, H.; Radcliffe, R.R.; Brevetti, G.R.; El-Khatib, H.; Vance, S.J.; Moyer, B.R.; Cotts, W.G.; Richenbacher, W.E. Time course of cytokine release and complement activation after implantation of the HeartMate left ventricular assist device. *ASAIO J.* **1998**, *44* (5), M347–M351.
 43. Suska, F.; Esposito, M.; Gretzer, C.; Källtorp, M.; Tengvall, P.; Thomsen, P. IL-1[Alpha], IL-1[Beta] and TNF-[Alpha] secretion during in vivo/ex vivo cellular interactions with titanium and copper. *Biomaterials* **2003**, *24* (3), 461–468.
 44. Babensee, J.E.; Stein, M.M.; Moore, L.K. Interconnections between inflammatory and immune responses in tissue engineering. *Ann. N.Y. Acad. Sci.* **2002**, *961*, 360–363.
 45. Kirkpatrick, C.J.; Krump-Konvalinkova, V.; Unger, R.E.; Bittinger, F.; Otto, M.; Peters, K. Tissue response and biomaterial integration: The efficacy of in vitro methods. *Biomol. Eng.* **2002**, *19* (2–6), 211–217.
 46. Schindler, R.; Senf, R.; Frei, U. Influencing the inflammatory response of haemodialysis patients by cytokine elimination using large-pore membranes. *Nephrol. Dial. Transplant.* **2002**, *17* (1), 17–19.
 47. Ratner, B.D. Reducing capsular thickness and enhancing angiogenesis around implant drug release systems. *J. Control. Release* **2002**, *78* (1–3), 211–218.
 48. Yan, J.Y.J.; Cooke, F.W.; Vaskelis, P.S.; von Recum, A.F. Titanium-coated Dacron velour. A study of interfacial connective tissue formation. *J. Biomed. Mater. Res.* **1989**, *23* (2), 171–189.
 49. Winkler, P.A.; Herzog, C.; Weiler, C.; Krishnan, K.G. Foreign-body reaction to silastic burr-hole covers with seroma formation: Case report and review of the literature. *Pathol. Res. Pract.* **2000**, *196* (1), 61–66.
 50. Ma, N.; Petit, A.; Yahia, L.; Huk, O.L.; Tabrizian, M. Cytotoxic reaction and TNF-[Alpha] response of macrophages to polyurethane particles. *J. Biomater. Sci., Polym. Ed.* **2002**, *13* (3), 257–272.
 51. Rhodes, N.P.; Hunt, J.A.; Williams, D.F. Macrophage subpopulation differentiation by stimulation with biomaterials. *J. Biomed. Mater. Res.* **1997**, *37* (4), 481–488.
 52. Ward, W.K.; Slobodzian, E.P.; Tiekotter, K.L.; Wood, M.D. The effect of microgeometry, implant thickness and polyurethane chemistry on the foreign body response to subcutaneous implants. *Biomaterials* **2002**, *23* (21), 4185–4192.
 53. Lobler, M.; Sa[ss], M.; Kunze, C.; Schmitz, K.-P.; Hopt, U.T. Biomaterial implants induce the inflammation marker CRP at the site of implantation. *J. Biomed. Mater. Res.* **2002**, *61* (1), 165–167.
 54. Held, M.; Dost, P.; Erdtmann, M.; Baumann, H.; Jahnke, K. Biocompatibility testing of new polymers in a moving implant bed. *ORL* **2000**, *62* (5), 247–250.
 55. Bos, R.R.; Rozema, F.R.; Boering, G.; Nijenhuis, A.J.; Pennings, A.J.; Verwey, A.B.; Nieuwenhuis, P.; Jansen, H.W. Degradation of and tissue reaction to bio-degradable poly(L-lactide) for use as internal fixation of fractures: A study in rats. *Biomaterials* **1991**, *12* (1), 32–36.
 56. Bergsma, J.E.; Rozema, F.R.; Bos, R.R.M.; Boering, G.; De Bruijn, W.C.; Pennings, A.J. Biocompatibility study of as-polymerized poly(L-lactide) in rats using a cage implant system. *J. Biomed. Mater. Res.* **1995**, *29* (2), 173–179.
 57. Khabbaz, F.; Karlsson, S.; Albertsson, A. Py-GC/MS an effective technique to characterizing of degradation mechanism of poly (L-lactide) in the different environment. *J. Appl. Polym. Sci.* **2000**, *78* (13), 2369–2378.
 58. von Recum, H.A.; Cleek, R.L.; Mikos, A.G.; Eskin, S.G. Degradation of polydispersed poly(L-lactic acid) to modulate lactic acid release. *Biomaterials* **1995**, *16* (6), 441–447.
 59. Van Dijk, M.; Tunc, D.C.; Smit, T.H.; Higham, P.; Burger, E.H.; Wuisman, P.I.J.M. In vitro and in vivo degradation of bioabsorbable PLLA spinal fusion cages. *J. Biomed. Mater. Res.* **2002**, *63* (6), 752–759.
 60. Iwasaki, Y.; Sawada, S.-I.; Ishihara, K.; Khang, G.; Hai, B.L. Reduction of surface-induced inflammatory reaction on PLGA/MPC polymer blend. *Biomaterials* **2002**, *23* (18), 3897–3903.
 61. Werkmeister, J.A.; Edwards, G.A.; White, J.F.; Casagrande, F.; Hunt, J.A.; Williams, D.F.; Ramshaw, J.A.M. In vivo evaluation of modified mandrel-grown vascular prostheses. *J. Biomed. Mater. Res.* **1999**, *47* (3), 316–323.

Functional MRI: Applications

Hui Mao

Department of Radiology, Emory University School of Medicine, Atlanta, Georgia, U.S.A.



INTRODUCTION

The development of functional magnetic resonance imaging (fMRI) has made tremendous contributions to the field of neuroscience in advancing our understanding of brain functions. Functional magnetic resonance imaging is now widely accepted as a noninvasive tool for identifying brain regions that are associated with certain perceptual, cognitive, emotional, and behavioral functions, such as visual, sensorimotor, language, and memory. The most commonly used fMRI technique detects brain activities from MRI signal changes that are blood oxygenation level dependent (BOLD).^[1,2] Blood oxygenation level dependent MRI takes advantage of the paramagnetic properties of deoxyhemoglobin molecule, an endogenous MRI contrast agent in the blood, and its concentration change that is induced by the brain activation. With currently available MRI equipment, brain functions can be studied noninvasively at high spatial resolution of less than a millimeter and good temporal resolution of less than a second. Since its introduction in the early 1990s, applications of fMRI have grown continuously from studying various brain functions in the basic neuroscience research to assisting presurgical planning in neurosurgery and to providing additional diagnostic information in the clinical management of patients who have functional disorders due to neurological diseases and mental illness.^[3,4] These applications have brought us new knowledge of human brain functions, better understanding of brain diseases and disorders, and improvement in medical practices. Developing fMRI technology and applications has become an important area in the field of biomedical engineering, as it receives and facilitates the multidisciplinary efforts from neuroscience, physics, engineering, computer science, as well as clinical practices of neurology, psychiatry, neurosurgery, and more. This entry provides a brief review of current applications of fMRI and its potential development in the future.

MAPPING THE FUNCTIONAL BRAIN

Identifying the brain region that is responsible for a particular function is an important approach to the better understanding of how a human brain works and

the improvement of the diagnosis and treatment of brain diseases and disorders. Before neuroimaging methods become available, the knowledge of the functional localization of the human brain was largely obtained based on the studies of patients who suffered various functional disorders due to brain injuries and brain lesions or patients who underwent brain surgery for treatment. The relationship between the specific brain structure and its function was determined through the examination of the damaged area and the characteristics of the functional impairment. For example, an injury to the occipital lobe may cause various degrees of impaired vision, suggesting the occipital cortex is critical to the visual functions. In the surgical treatment of patients with severe brain diseases, such as epilepsy or brain tumor, the region that is affected by the disease has to be removed. In addition to giving patients neurological tests before the surgery, during the open-skull surgery neurosurgeons often carefully evaluate brain functions and their relations to the area that needs to be removed. By directly stimulating or inhibiting the targeted brain regions point by point, one can record the behavioral responses from the patient to obtain spatial localizations of certain functions. Functions as well as functional impairment of the patient can be further examined following the surgery. Developed by physicians, such as Paul Broca, whose name is used as the identification of one of the brain regions responsible for the language function, this approach allowed mapping the general functional organization of the human brain and is still practiced in the neurobehavioral and neuropsychological exams, and in the intraoperative functional tests. However, the use of this approach is not without some significant limitations. For example, it is not widely accessible for research because of the availability of the patients. As an intraoperative procedure, it can be time-consuming and can only cover limited areas of the brain. Furthermore, characterization of functional areas is done, in most instances, on patients with a certain degree of functional impairment, therefore some functional tests cannot be performed naturally. Because the spatial arrangement of functional areas can be altered or reorganized as results of the brain diseases and the plasticity of the brain, there is considerable concern that the information on the functional organization obtained from patients may not be represented correctly in the “normal” brain. Other

noninvasive and nonimaging brain mapping methods, such as electrophysiological recording methods of electroencephalography (EEG) or electromyogram (EMG), suffer from limited spatial resolution and brain coverage. Thus, the number of applications of functional brain mapping in research and clinical practices is very limited.

Although it has been known that increased neuronal activity leads to increases of regional cerebral blood flow (rCBF), regional cerebral blood volume (rCBV), and blood oxygenation, the fundamental biochemical and physiological mechanisms of these processes are still under intense investigation. Nevertheless, there is an empirical correlation between local neuronal activity in the brain and local physiological changes in hemodynamics. It has been generally accepted that brain activation, hemodynamics, and metabolic processes are coupled in both space and time. The coupling of the neuronal activity to the alterations in the vascular system is the basis for a number of functional or neuroimaging methods currently used in studying brain function, e.g., positron emission tomography (PET), single photon emission computed tomography (SPECT), and optical imaging and fMRI (Table 1).^[5,6] Development of neuroimaging methods provides the way to visualize and study the brain noninvasively in patients and healthy persons. Using neuroimaging methods, brain mapping can be done in the conscious state of a person, while the entire brain is examined to determine the functional organization and connections of the brain.

Methods of fMRI include various imaging techniques, such as arterial spin label (ASL) perfusion imaging and contrast agent enhanced dynamic imaging for measuring changes of rCBF, and magnetic resonance spectroscopic imaging for measuring changes of regional metabolic activity; however, the most

commonly used technique is based on the detection of BOLD signal change.^[7-9] Blood oxygenation level dependent fMRI allows robust detection of brain activation and can be readily performed on the MRI scanners conveniently available in many research and healthcare institutions. Increase in neuronal activity leads to a transient increase of local blood flow and blood oxygenation, which causes the change in the blood concentration of deoxyhemoglobin, a paramagnetic species that causes subtle change of magnetic susceptibility or T_2^* effect of proton transverse relaxation time of the brain tissues in the territory of the microvasculature. A statistical comparison of images obtained during an activated condition with images obtained during a control condition can be used to reveal activated brain regions that are specific to a particular mental task. It should be noted that BOLD signal is dependent on both the underlying physiological events and the imaging physics.^[10] The activation-induced BOLD signal measured with fMRI is considered to be an indirect measurement of the change in the blood deoxyhemoglobin concentration coupled with the local synaptic activity in the vicinity. Although it is sensitive, BOLD contrast is not a directly quantifiable measure of neuronal activity; for example, it does not have units of “ml/min” or “activity/sec.” However, measuring the time course of hemodynamic changes and the relationship between these physiological changes to the underlying neuronal activity will determine both the spatial resolution and the accuracy of the fMRI technique.^[11] After a decade of development, fMRI has now been widely recognized as one of the most desirable functional imaging methods with high spatial resolution and good temporal resolution for functional brain mapping when compared to other neuroimaging modalities.

Table 1 Comparison of some imaging modalities and their measurements

Imaging modality	Hemodynamic parameters measured	Contrast used for imaging	Spatial resolution (mm) ^a	Temporal resolution (sec)
PET	rCBF, rCBV	Radioactive tracer (¹⁵ O or ¹⁸ F)	5–20	0.06–0.12
SPECT	rCBF	Radioactive tracer (⁹⁹ Tc or ¹²³ I)	10–20	>0.3
Optical ^b (near infrared)	Blood oxygenation	Hemoglobin and deoxyhemoglobin	2–10	<1
fMRI	Blood oxygenation, rCBF	Deoxyhemoglobin or magnetically “labeled” blood	2–4	0.1–3

^aSpatial resolution is compared at in-plane resolution. A typical slice thickness used for fMRI exam is 4–5 mm.

^bOptical imaging, e.g., near-infrared spectroscopy, in particular, has technical limitation in brain coverage; however, it can measure deoxyhemoglobin concentration directly and quantitatively. It has attracted increased attention to develop this method for studying brain function.

LOCALIZING BRAIN ACTIVATION WITH BOLD fMRI

The image acquisition technique used for BOLD fMRI is the combination of the gradient echo method and single-shot echo planar imaging (EPI), designed for T_2^* weighted contrast and a rapid image readout. Using this technique, fMRI exams can be routinely performed at the field strength of 1.5–3 T with the temporal resolution of 1–3 sec, depending on the acquisition parameter of time of repetition (TR), and the spatial resolution of 2–3 mm (in plane voxel size). With higher field strength, more powerful gradient system, and parallel imaging technique, the spatial resolution and temporal resolution can be further improved.

As other neuroimaging modalities, fMRI experiments rely on the ability to detect time-dependent task-evoked signal changes and to extract regions of activation using statistical methods.^[12] Any area with signal changes that correlate with the neurological stimuli can be identified and compared with the anatomy of the brain. In general, there are two types of task paradigms used in fMRI studies: block (“boxcar”) design and event-related design. Block design uses a boxcar waveform that typically has a control condition, e.g., resting, and a task condition. However, it is also possible to test more than two conditions during one scan. Given instructions and stimuli, a person who receives fMRI scan will execute different tasks in different conditions that are alternating, while the MRI scanner records images over a period of time. Block design offers a relatively simple experimental method and robust detection of brain activation as the BOLD signal rising from the brain responses to stimuli in each block is accumulated in that block. Therefore, it is commonly used in clinical applications of fMRI, which are done mostly on the individual patient with restraints of the scan time and patient tolerance. In addition to its simplicity, block design is well suited to clinical fMRI, because the primary goal in most clinical applications is to specifically localize an activated brain region rather than to look for the pattern of connectivity throughout the brain. Event-related design is similar to the experimental designs used in the conventional electrophysiological and neurophysiological tests, such as event-evoked EEG or EMG. It uses very brief stimuli, typically <1 sec each in duration against a baseline control condition. Because most of the event-related design experiments use very brief stimuli, they can be used to track the flow of neural process through the brain. They can be also used to investigate differences in the onset of neural activity evoked by different types of stimuli. Furthermore, because the stimuli can appear randomly, they are statistically more efficient than blocked designs and can minimize the effect of mental preparation such as prediction and planning. Consequently, event-related

fMRI is considered as a more suitable approach to address a number of important cognitive questions that focus on the temporal behavior of the hemodynamic response.^[13] Today, most fMRI applications in studying cognitive processes use event-related experimental design. However, its application in clinical fMRI faces several technical challenges. First, complicated statistical models are often needed for deriving the functional maps, but this is becoming feasible with the inevitable increase in computer processing speed. Second, substantial validation and development remains to be done before event-related designs become widespread in clinical practice. Nevertheless, because event-related designs offer greater flexibility, they will likely gain wider acceptance in the future.

Postprocessing of fMRI data generally follows the steps of: 1) realigning the functional images recorded in the time series to correct for motion artifacts; 2) coregistering functional images to the anatomical images; 3) statistical analysis to correlate functional image time series to the design paradigm; and 4) superimposing a functional map on high-resolution images for anatomical assignments. Because data analysis of fMRI strongly relies on the statistical power, in many fMRI applications, results are obtained by averaging data from many individuals in the selected population and presented in the normalized standard brain atlas. This group analysis approach is very effective when studying group differences in brain functions and is widely used in research applications. However, because many clinical applications of fMRI are usually for investigating an individual patient case by case, it is not generally necessary, and appropriate in many circumstances, to normalize the images and functional maps to a standard brain atlas in those applications.

APPLICATIONS OF fMRI

Propelled by rapid advancement of MRI technology, applications of fMRI have been developed and are widely spread over many areas of neurosciences and clinical management of brain diseases and disorders. New applications and interventions based on fMRI are emerging continuously. Here, we attempt to highlight several major applications.

Studying Basic Functions and Functional Organization of the Brain

Perhaps, the most important application of fMRI should be attributed to its role in studying basic brain functions. Development of fMRI allowed opening up a window to the insights of the brain where our knowledge and understanding are still limited. Since it was

introduced, fMRI has been used mainly by neuroscientists to map brain functions and led to the tremendous growth of our knowledge about the brain in the past decade. In recent years, fMRI is applied largely in studying cognitive functions and functional organization of the brain.^[14,15] Cognitive functions or “higher-level functions,” such as memory, emotion, and perception, are processed by different but interrelated brain regions in a synergic fashion. Neuronal and physiological responses of cognitive activities are usually subtle and difficult to detect. The ability to simultaneously scan the entire brain with high spatial resolution and deep tissue penetration makes fMRI the desirable technique for neuroscientists to interrogate the complex neuronal network associated to cognitive functions.^[16] Functional magnetic resonance imaging not only can be used to test the classic views and hypotheses of cognitive functions that were obtained from studies of patients or animal models, but also makes significant additions to advance our understanding of

those functions. For example, Haist and colleagues used fMRI to investigate the activation pattern in memory consolidation, one of the important aspects of the memory function, using a test that subjects viewing the faces of people who were famous primarily within a particular decade, as well as faces of nonfamous people. They found the selective activation for famous faces in the hippocampus only for responding to the faces of famous people of the most recent decade, while activation in the entorhinal cortex was most prominent for faces of the most recent decade, but less prominent for faces famous in the less recent decade, and absent for the remote decades.^[17] This pattern of activation (shown in Fig. 1) is consistent with the once-debatable classic view of the prolonged memory consolidation initially derived from the study of patients with defects in hippocampal regions. Furthermore, new information obtained from their study suggested that the prolonged consolidation occurs in two stages, an initial, relatively brief stage mediated by the hippocampus and a more

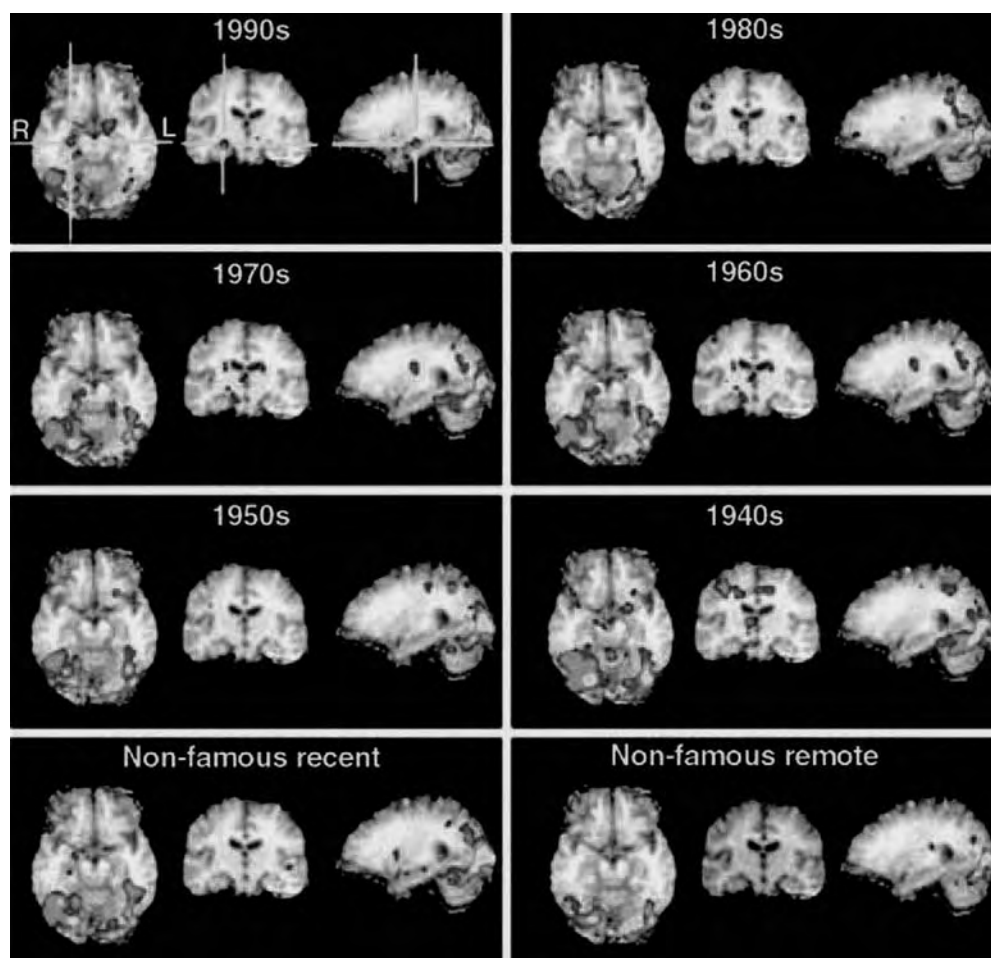


Fig. 1 Brain activation patterns from the famous faces remote memory test and fMRI indicated the right hemisphere hippocampal activation (showed in the crosshairs). Images are shown in the neurological convention in which the right hemisphere is the right side of the image. (From Ref. [17].)

lasting stage mediated by the adjacent entorhinal cortex.

So far the vast amount of fMRI studies in cognitive functions have derived new information that advanced our understanding of brain functions and also brought some light to the mechanism of some brain diseases and disorders.

Improving the Diagnosis of Mental Disorders

Because of its close ties to neuroscience, fMRI has been used to investigate the neural basis of mental illness since fMRI was introduced. Most of the major psychiatric disorders, e.g., schizophrenia, major depression, and attention deficit hyperactivity disorder, appear to represent disordered neural systems widely distributed in the brain. Because of this, it is thought that there is more than one single brain region associated with a particular mental illness. Using fMRI, it is now possible to observe how cognitive tasks are performed. Such “in-action” information is not available from traditional neuropsychological tests. With proper experimental design, brain activation patterns identified by fMRI can reveal the mechanisms of the psychological processes, such as strategy or compensatory. Patients with psychological disorders may demonstrate different neural networks from normal controls as the cognitive processes in patients may be altered by specific neuropsychological deficits.

There are a few consistent findings that have emerged, which hold promise for the development of clinically useful fMRI in psychiatry. For example, the anterior cingulate cortex has become a focal point for the pathophysiology of major depression. There are some evidences suggesting that depression is associated with hyperactivity in this region and that effective pharmacological treatment is associated with a normalization of activity. Similarly, the anterior cingulate has been shown to be less active in schizophrenic patients.^[18] On the other hand, fMRI is very useful in testing existing hypotheses derived from the previous neuropsychological studies.^[19] The hypothesis of deficits in recruitment of frontal lobe structures during higher-level executive functions in attention deficit hyperactivity disorder (ADHD) patients has been confirmed with great specificity using fMRI. Rubia et al. demonstrated that the frontal lobe in ADHD adolescent patients exhibits task-specific hypoactivity during complex higher-level executive tasks, but appears normal during a lower-level finger tapping task.^[20] However, in all of these examples the difference between patients and controls was statistical. Using fMRI for individual diagnosis of mental illness is not state of the art. Because psychiatry does not currently have any reliable diagnostic test, the development of a reliable test that

could be used on an individual basis would represent a major breakthrough.

Assisting Neurosurgery in Presurgical Planning

The ability to noninvasively localize functional foci of the brain makes fMRI particularly useful for preoperative planning in neurosurgery. The primary use of fMRI is to identify cortical functional areas so that their functions can be preserved, if possible, during brain biopsy, resective surgery, neuroradiological embolization, or radiation therapy.^[21,22] When removal of such areas becomes necessary for the best of the surgical treatment, functional mapping can be used to determine the surgical margin and to assess the potential functional deficits so that other treatment options can be explored or proper postsurgery rehabilitation can be planned.

In the clinical management of brain tumor patients, fMRI can play an important role in mapping of eloquent functional areas to guide stereotactic biopsy and tumor resection, and to assess functional plasticity in relation to the lesion.^[23] One of the examples shown here involved a patient with a brain tumor located in the left parietal lobe. Surgical removal of the tumor posed the potential risk of impairing the movement of the right limbs. Functional magnetic resonance imaging was then used to identify the left hemisphere motor area as the tumor sits close to the left motor stripe that controls the right-hand side of the body. Patient was tested with a series of motor tasks including the right-handed sequential finger tapping, hand movement with the fist open and close, wrist flexion, etc. This exam allowed localizing the areas that control right-hand movements, as shown in Fig. 2, and identifying the possible surgical margin between the functional area and the tumor boundary. In the other case, a growing tumor is located in the left frontal area of a patient, an area important for many cognitive functions, including the language processes. Presurgical preparation included fMRI exam to determine the language area and its spatial relation to the tumor. This was accomplished by using a confrontation naming task, in which the patient was instructed to generate a sentence to name each animal shown in the picture in one condition while counting numbers in the other. The results from fMRI showed that the tumor appeared to infiltrate into the language areas (Fig. 3), revealing the risk of disrupting language function if the surgical treatment applies. There are, however, potential limitations of fMRI in the study of brain tumors. The presence of edema, tumor mass, and radiation-induced tissue damage affects the microvasculature surrounding the lesion. One must be careful, especially in the interpretation of data from an fMRI exam for the guidance

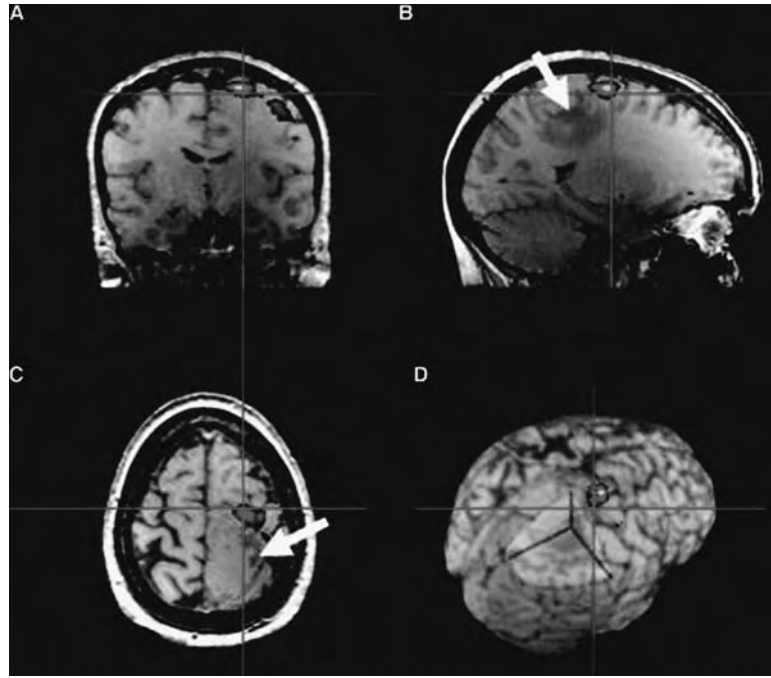


Fig. 2 Preoperative fMRI identified the left motor area of a patient who had a brain tumor located in the left parietal area. Motor activation was the response of a right-handed finger tapping. The activation map was overlaid on T1 weighted images of coronal (A), sagittal (B), and transverse (C) views as well as a three-dimensional rendering of the brain (D). The left motor area was found close to the tumor (indicated by arrows). Images are shown in the radiological convention in which the left hemisphere is in the right-hand side of the image.

of surgical procedures, especially for areas directly adjacent to tumors.

In the area of surgical treatment of epilepsy, there have been significant efforts made in developing fMRI method to replace the traditional intracarotid amytal procedure, also known as the Wada test. The Wada test involves inserting a catheter into the carotid of the patient to infuse anesthetic amytal to partially paralyze the selected part of the brain in order to localize the language and the memory areas. In comparison, fMRI provides not only a safe alternative but also a high spatial and anatomic resolution in localizing functional areas. In addition, fMRI procedure can be conveniently

done in conjunction with the routine preoperative MRI. Some studies have developed task paradigms, such as picture naming, word generation, or simple rhyme detection to determine hemispheric dominance for language in epileptic patients.^[24,25] Results derived from fMRI examinations were compared with those obtained on the same patients using the standard Wada test, stereotactic intracerebral EEG stimulations and recordings, as well as outcomes of the temporal lobectomy. These studies demonstrated that fMRI examination generates robust responses in the language areas and is feasible for determining hemispheric language dominance. As a result, fMRI has been used

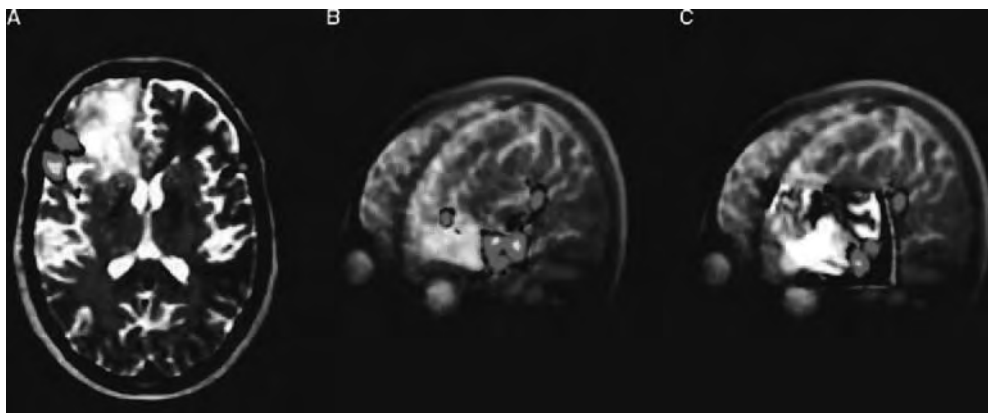


Fig. 3 Functional magnetic resonance imaging was used to determine the left frontal language area and its relation to the tumor for a brain tumor patient in the planning of a biopsy. (A) A transverse sectional view of the brain showing the activated area in color and bright tumor mass in the T2 weighted image; (B) three-dimensional rendering of the brain with activated areas shown in color; and (C) a cutoff in the three-dimensional rendering view showing the bright tumor mass and its relation to the activated area. Images are shown in the neurological convention.

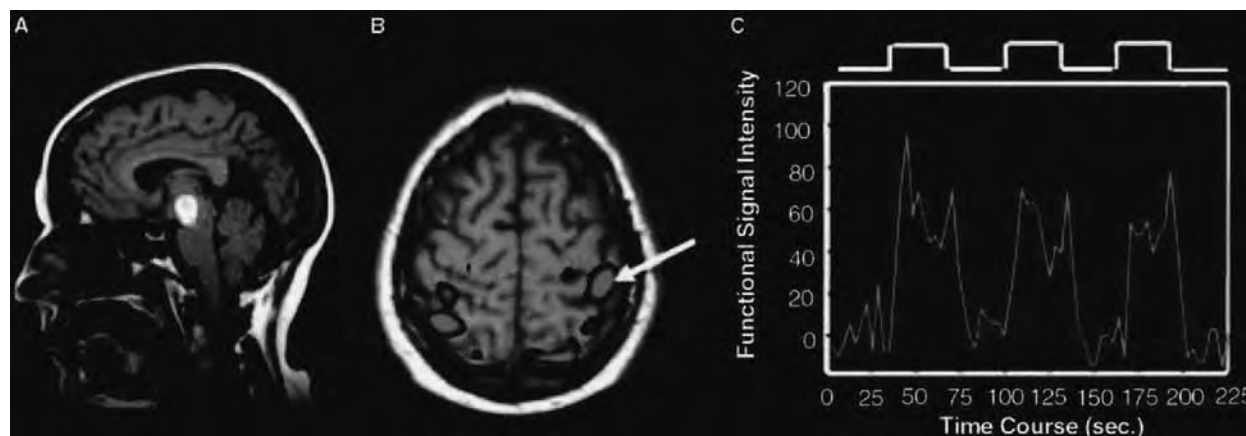


Fig. 4 A locked-in patient suffered from a brainstem stroke (highlighted in a T1 weighted sagittal image) (A) and demonstrated activated motor area in the posterior parietal lobule when performing the imagery finger tapping (B); blood oxygenation level dependent signal time course (C) was recorded in the area (indicated by an arrow), which was selected for implanting a neurotrophic electrode that extracts brain signals for interface with a computerized assisting device.

in many institutions for lateralization of the language area in presurgical planning for patients with temporal lobe epilepsy, although the complete replacement of the Wada test has not yet been accomplished.

Other neurosurgical interventions may use fMRI for functional guidance. For example, in the effort of developing brain–computer interface for patients who are quadriplegic or “locked-in” with loss of the means of communication, because of brain injuries or other neurological disorders, such as amyotrophic lateral sclerosis (also known as Lou Gehrig’s disease); fMRI can be used to identify the optimal site for the implantation of a neurotrophic electrode that picks up and converts neuronal activation into digital signals.^[26] Functional magnetic resonance imaging combined with an imaginary finger-tapping task was able to find an activated area in the motor area of a patient who was locked-in because of a brainstem stroke. The implantation of neurotrophic electrodes at the location identified by the fMRI exam (Fig. 4) established the direct interface between the patient’s brain and a device that was used to assist the patient in communication.

Monitoring Functional Changes Related to Diseases and Treatments

Another application of fMRI is to monitor functional recovery, disease progression, treatment efficacy, and the response to medication.^[27,28] For example, fMRI has been used to map the short- and long-term cortical reorganization after a spinal cord injury (SCI). Significant reorganization of the motor cortex was found in completed SCI patients over the time after the injury, revealing the characteristics of brain plasticity of SCI (Fig. 5). This type of information is not previously available, but can be important to the design of rehabilitation therapy. In investigating the recovery of motor ability after brain injury, such as stroke, fMRI is also found to be useful in studying the important underlying mechanism and in developing rehabilitation intervention to improve the recovery of patients. Functional magnetic resonance imaging was applied to monitor the changes of activation pattern that may represent the response of the constraint-induced therapy for motor recovery after stroke.^[29] Applications of

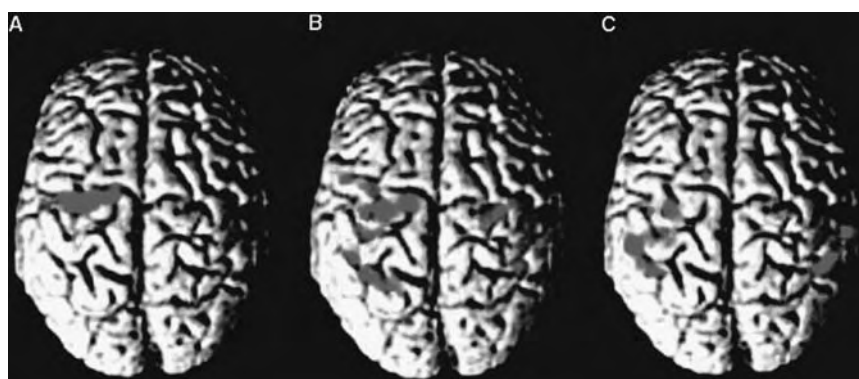


Fig. 5 Functional reorganization of the left motor area was observed in the patients with short-term and long-term SCI. (A) Control; (B) 2 yr after SCI; and (C) more than 10 yr after SCI.

fMRI to study neurological impairments that are not associated with structural abnormalities, such as learning problems, dyslexia, and motor neuron diseases, are currently in progress.

There is a growing interest to look for the brain activation pattern as a possible “marker” for the diagnosis of neurodegenerative diseases, such as Alzheimer’s disease, dementia, as well as mild traumatic brain injuries. As other functional imaging modalities, fMRI provides a promising approach to identify the specific activation pattern or disruption of the normal activation pattern that is associated to the disease and the severity of the disease.

Measuring the Neurological Response to Pharmacological Effects

There is an emerging application of fMRI in studying pharmacological problems in recent years. The ability of fMRI to detect brain activation in vivo can be also applied to investigate the neurological and physiological changes associated to the pharmacological effect of drugs, for instance, the changes evoked by the treatment of the psychotropic drug.^[30,31] Several studies have demonstrated the feasibility and potential of using fMRI to assess the drug effect on the brain function and the specific brain region. Using fMRI, significant BOLD changes in the medial geniculate nucleus, mediodorsal thalamic nucleus, mesencephalic trigeminal nucleus, and parietal cortex can be observed in rats administered with dizocilpine MK-801, an *N*-methyl-d-aspartate receptor antagonist that is a neuroprotective compound with well-documented effects on CBF and metabolism.^[32] Although pharmacological fMRI is largely used in the study of animal models at this time, such application in human is proven to be feasible. However, one of the major challenges of pharmacological applications of fMRI is to quantify the pharmacological effect of drugs. There is a significant limitation in using BOLD fMRI because of the complicated source of the BOLD signal. It is expected that the development of quantitative fMRI methods, such as perfusion fMRI, will facilitate the applications of pharmacological fMRI in the future.

FUTURE DEVELOPMENT OF fMRI APPLICATIONS

Very High-Field MRI Scanner

Both MR signal intensity and BOLD contrast increase as a function of the field strength. With increased signal-to-noise and T_2^* weighting, functional imaging can be performed at higher temporal resolution and sensitivity.

Although the majority of fMRI studies have been done using 1.5 T MRI scanner, it has been demonstrated that both functional and anatomic information can be obtained at high spatial and temporal resolution with high-field magnets (3.0 T). This promoted the rapid development of 3.0 T clinical scanners in recent years. Using 3.0 T scanners, which are currently available from all major manufacturers, anatomical images can be obtained with submillimeter resolution. With the endless effort of developing MRI engineering and technologies, 7.0 T whole-body MRI scanners are becoming available from major MRI scanner manufactures at the end of 2005. This will further enhance the ability of fMRI applications to perform reliable functional studies on individual patients and improve the efficiency of fMRI exams. Once available, such very high-magnetic field systems are capable of performing high-resolution fMRI, which can be used for mapping functional brain organization ranging from large cortical networks to small nuclei, and even to cellular layers.

Quantitative fMRI

Perhaps, the biggest challenge to BOLD fMRI is its limited ability to quantify the BOLD signal itself. However, quantitative fMRI is particularly important to clinical fMRI if it is to assess disease, prognosis, and treatment. A method with great potential for quantitative fMRI is the ASL technique. This is a type of dynamic MRI that can measure the rCBF change associated with brain activation noninvasively. Several technical developments, including the correct measurement of transit and trailing times, improvements in temporal resolution, and data processing, remain to be done. In any case, ASL perfusion MRI is expected to provide a quantitative measurement of activation-induced physiological changes, and can be extremely useful in the applications that study the severity of diseases or disorders, treatment efficacy, and pharmacological effect. Another promising method that provides quantitative measurement of rCBF is dynamic perfusion imaging with bolus injection of exogenous contrast agents, such as gadolinium chelate or superparamagnetic iron oxide nanoparticles. These agents can greatly expand the capability of fMRI and provide accurate measurement of task-induced rCBF change. Applications of susceptibility contrast physics and standard “bolus-tracking” kinetic principles allow the calculation of relative CBV maps and mean transit time parameters, which are related directly to the hemodynamic response.

Integrated fMRI

There are increasing needs in high spatial and temporal resolution to study cognitive functions, functional

connectivity, and neuronal network. This has led efforts to integrate fMRI with other imaging technologies or modalities.

As fMRI can only indirectly measure the neuronal activity via physiological response, its temporal resolution is slower than the time scale of actual neuronal activity. To improve the temporal resolution, approaches of combining MRI and fMRI with EEG or MEG have been developed.^[32] Because electrophysiology recording allows true real-time measurements of the neuronal activity evoked by a stimulus, the combination of event-related fMRI and electromagnetic recording of event-related potential (ERP) may offer the millimeter spatial resolution and the millisecond temporal resolution when used for studying cognitive processes. Furthermore, functional connectivity can be also addressed with the help of diffusion tensor imaging and diffusion tensor-derived white matter tractography, which allow delineating the structural connection between different brain regions that are activated in the network of functional processes.^[33] It is expected that these developments will enhance capability of fMRI and expand its applications, and will particularly benefit the diagnosis and treatment of brain diseases.

CONCLUSIONS

Functional magnetic resonance imaging provides a noninvasive and widely applicable method for functional brain mapping. With excellent spatial and temporal resolution, it allows neuroscientists to study the various functions and their organization of human brain. Development of fMRI applications not only advanced our knowledge and understanding of brain functions in great deal, but also provided clinical neuroimaging tools for improving our understanding, diagnosis, and treatment of brain diseases and disorders. As neuroscience research and fMRI technology continue to progress, it is expected that the applications of fMRI will continue to grow in the future.

REFERENCES

- Ogawa, S.; Tank, D.W.; Menon, R.S.; Ellerman, J.M.; Kim, S.G.; Merkle, H.; Ugurbil, K. Intrinsic signal changes accompanying sensory stimulation—functional brain mapping with magnetic resonance imaging. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 5951–5955.
- Kwong, K.K.; Belliveau, J.W.; Chesler, D.A. Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 5675–5679.
- Detre, J.A.; Floyd, T.F. Functional MRI and its applications to the clinical neurosciences. *Neuroscientist* **2001**, *7*, 64–79.
- Sunaert, S.; Yousry, T.A. Clinical applications of functional magnetic resonance imaging. *Neuroimaging Clin. N. Am.* **2001**, *11*, 221–236.
- Mandzia, J.; Black, S.E. Neuroimaging and behavior: probing brain behavior relationships in the 21st century. *Curr. Neurol. Neurosci. Rep.* **2001**, *1* (6), 553–561.
- Ronald, L.V.H.; Ronald, S.T. *Functional Cerebral SPECT and PET Imaging*, 1st Ed.; Lippincott Williams & Wilkins: Philadelphia, 2000.
- Gonzalez-At, J.B.; Alsop, D.C.; Detre, J.A. Cerebral perfusion and arterial transit time changes during task activation determined with continuous arterial spin labeling. *Magn. Reson. Med.* **2000**, *43*, 739–746.
- Li, T.Q.; Haefelin, T.N.; Chan, B.; Kastrup, A.; Jonsson, T.; Glover, G.H.; Moseley, M.E. Assessment of hemodynamic response during focal neural activity in human using bolus tracking, arterial spin labeling, and BOLD techniques. *Neuroimage* **2000**, *12*, 442–451.
- Serafini, S.; Steury, K.; Richards, T.; Corina, D.; Abbott, R.; Dager, S.R.; Berninger, V. Comparison of fMRI and PEPSI during language processing in children. *Magn. Reson. Med.* **2001**, *45*, 217–225.
- Menon, R.S.; Kim, S.G. Spatial and temporal limits in cognitive neuroimaging with fMRI. *Trends Cogn. Sci.* **1999**, *3*, 207–216.
- Buxton, R.B.; Frank, L.R. A model for the coupling between cerebral blood flow and oxygen metabolism during neural stimulation. *J. Cereb. Blood Flow Metab.* **1997**, *17*, 64–72.
- Friston, K.J.; Jezzard, P.; Turner, R. Analysis of functional MRI time series. *Hum. Brain Mapp.* **1994**, *1*, 153–171.
- Rosen, B.R.; Buckner, R.L.; Dale, A.M. Event-related functional MRI: past, present, and future. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 773–780.
- Gaillard, W.D. Functional MR imaging of language, memory, and sensorimotor cortex. *Neuroimaging Clin. N. Am.* **2004**, *14* (3), 471–485.
- Culham, J.C.; Kanwisher, N.G. Neuroimaging of cognitive functions in human parietal cortex. *Curr. Opin. Neurobiol.* **2001**, *11*, 157–163.
- Phelps, E.A. Human emotion and memory: interactions of the amygdala and hippocampal complex. *Curr. Opin. Neurobiol.* **2004**, *14* (2), 198–202.
- Haist, F.; Bowden Gore, J.; Mao, H. Consolidation of human memory over decades revealed by functional magnetic resonance imaging. *Nat. Neurosci.* **2001**, *4* (11), 1139–1145.
- Carter, C.S.; MacDonald A.W., III; Ross, L.L.; Stenger, V.A. Anterior cingulate cortex activity and impaired self-monitoring of performance in patients with schizophrenia: an event-related fMRI study. *Am. J. Psychiatry* **2001**, *158*, 1423–1428.
- Roth, R.M.; Saykin, A.J. Executive dysfunction in attention-deficit/hyperactivity disorder: cognitive and neuroimaging findings. *Psychiatr. Clin. North. Am.* **2004**, *27* (1), 83–96.
- Rubia, K.; Russell, T.; Bullmore, E.T.; Soni, W.; Brammer, M.J.; Simmons, A.; Taylor, E.; Andrew, C.; Giampietro, V.; Sharma, T. An fMRI study of reduced left prefrontal activation in schizophrenia during normal



- inhibitory function. *Schizophr. Res.* **2001**, *52* (1–2), 47–55.
21. Lee, C.C.; Ward, H.A.; Sharbrough, F.W.; Meyer, F.B.; Marsh, W.R.; Raffel, C.; So, E.L.; Cascino, G.D.; Shin, C.; Xu, Y.; Riederer, S.J.; Jack, C.R., Jr. Assessment of functional MR imaging in neurosurgical planning. *AJNR Am. J. Neuroradiol.* **1999**, *20* (8), 1511–1519.
 22. Haberg, A.; Kvistad, K.A.; Unsgard, G.; Haraldseth, O. Preoperative blood oxygen level-dependent functional magnetic resonance imaging in patients with primary brain tumors: clinical application and outcome. *Neurosurgery* **2004**, *54* (4), 902–914.
 23. Roux, F.E.; Ibarrola, D.; Tremoulet, M.; Lazorthes, Y.; Henry, P.; Sol, J.C.; Berry, I. Methodological and technical issues for integrating functional magnetic resonance imaging data in a neuronavigational system. *Neurosurgery* **2001**, *49* (5), 1145–1156.
 24. Lehericy, S.; Cohen, L.; Bazin, B.; Samson, S.; Giacomini, E.; Rougetet, R.; Hertz-Pannier, L.; Le Bihan, D.; Marsault, C.; Baulac, M. Functional MR evaluation of temporal and frontal language dominance compared with the Wada test. *Neurology* **2000**, *58* (8), 1625–1633.
 25. Woermann, F.G.; Jokeit, H.; Luerding, R.; Freitag, H.; Schulz, R.; Guertler, S.; Okujava, M.; Wolf, P.; Tuxhorn, I.; Ebner, A. Language lateralization by Wada test and fMRI in 100 patients with epilepsy. *Neurology* **2003**, *61* (5), 699–701.
 26. Kennedy, P.; Andreasen, D.; Ehirim, P.; King, B.; Kirby, T.; Mao, H.; Moore, M. Using human extra-cortical local field potentials to control a switch. *J. Neural Eng.* **2004**, *1* (2), 72–77.
 27. Elbert, T.; Rockstroh, B. Reorganization of human cerebral cortex: the range of changes following use and injury. *Neuroscientist* **2004**, *10* (2), 129–141.
 28. Loubinoux, I.; Carel, C.; Pariente, J.; Dechaumont, S.; Albucher, J.F.; Marque, P.; Manelfe, C.; Chollet, F. Correlation between cerebral reorganization and motor recovery after subcortical infarcts. *Neuroimage* **2003**, *20*, 2166–2180.
 29. Seritan, A.L.; Mendez, M.F.; Silverman, D.H.; Hurley, R.A.; Taber, K.H. Functional imaging as a window to dementia: corticobasal degeneration. *J. Neuropsychiatry Clin. Neurosci.* **2004**, *16* (4), 393–399.
 30. Leslie, R.A.; James, M.F. Pharmacological magnetic resonance imaging: a new application for functional MRI. *Trends Pharmacol. Sci.* **2000**, *21* (8), 314–318.
 31. Houston, G.C.; Papadakis, N.G.; Carpenter, T.A.; Hall, L.D.; Mukherjee, B.; James, M.F.; Huang, C.L. Mapping of brain activation in response to pharmacological agents using fMRI in the rat. *Magn. Reson. Imaging* **2001**, *19* (7), 905–919.
 32. Rossini, P.M.; Dal Forno, G. Integrated technology for evaluation of brain function and neural plasticity. *Phys. Med. Rehabil. Clin. N. Am.* **2004**, *15* (1), 263–306.
 33. Ramnani, N.; Behrens, T.E.; Penny, W.; Matthews, P.M. New approaches for exploring anatomical and functional connectivity in the human brain. *Biol. Psychiatry* **2004**, *56*, 613–619.

FURTHER READING

- Detre, J.A.; Floyd, T.F. Functional MRI and its applications to the clinical neurosciences. *Neuroscientist* **2001**, *7*, 64–79.
- Huettel, S.A.; Song, A.W.; McCarthy, G. *Functional Magnetic Resonance Imaging*, 1st Ed.; Sinauer Associates: Sunderland, MA, 2004.
- Marshall, J.C.; Fink, G.R. Cerebral localization, then and now. *Neuroimage* **2003**, *20* (suppl. 1), S2–S7.
- Richard, B.B. *Introduction to Functional Magnetic Resonance Imaging: Principles and Techniques*, 1st Ed.; Cambridge University Press: New York, 2001.
- Ugurbil, K.; Toth, L.; Kim, D.S. . How accurate is magnetic resonance imaging of brain function? *Trends Neurosci.* **2003**, *26* (2), 108–114.

Gait Analysis

Richard Baker

Hugh Williamson Gait Laboratory, Royal Children's Hospital, Melbourne, Victoria, Australia

INTRODUCTION

Gait analysis is the study of walking. In a biomedical context it is now assumed to refer to the instrumented measurement of walking and the interpretation of such data. There is now a broad consensus on how clinical gait analysis is conducted and is this that is outlined in this article. After a brief historical review there is a description of the different sorts of data used and of how these are measured and processed. This is followed by an overview of the process of clinical gait analysis. The final section reviews some recent developments.

HISTORY

Scientific interest in walking was first documented by Aristotle (384–322 B.C.^[1]), although it was not until 1680 that the first experimental data were reported by Giovanni Borelli (1608–1679^[2]). In the nineteenth century, Eduard (1795–1874) and Wilhelm (1804–1901) Weber^[3], Etienne Jules Marey (1830–1904^[4]), Edward Muybridge (1830–1904^[5]), and Willhem Braune (1831–1892) and Otto Fischer (1861–1917^[6]) all had a role in developing techniques to document walking and made progressive steps in describing the process of walking. A team headed by Verne Inman (1905–1980) and Howard Eberhart (1906–1993) at the Biomechanics Laboratory of the University of California at Berkeley shortly after the Second World War advanced experimental techniques and theoretical understanding about as far as possible before the advent of modern computers made the processing and study of data from individual patients a realistic undertaking.^[7]

Modern clinical gait analysis involves instrumented data capture and biomechanical modelling and interpretation of these data to assist the clinical management of the subject. Although various centers had a role in developing these elements of analysis during the mid-twentieth century, they were first brought fully together at Newington Children's Hospital (Hartford, CT)^[8] in the early 1980s under the direction of orthopedic surgeon Jim Gage.

The last two decades have seen some stability in the operation of clinical gait analysis based broadly on the

protocols and systems developed at Newington. These techniques were developed to analyze the walking of children with cerebral palsy, and clinical use of these techniques is still most widespread among this patient group. Use with other pathologies is increasing but the same general approach to capture, processing, presentation, and interpretation of data is used. This approach can now justifiably be referred to as conventional clinical gait analysis. In describing methodologies below the implementation of the Newington model will be focused on.

DATA USED IN GAIT ANALYSIS

Conventional clinical gait analysis is based on the collection of four core types of data: clinical examination, basic temporal-spatial data, joint kinematics, and joint kinetics. The clinical examination of the patient includes assessments of joint range of motion, muscle strength and tone, and skeletal abnormalities. This is the realm of the physiotherapist or orthopedic surgeon and will not be discussed further here. Many laboratories also conduct electromyography (EMG) tests and/or oxygen consumption measurements, but these separate subjects are not discussed in this article.

Basic Terminology

Walking is a cyclic process and is broken down into gait cycles for analysis. Specific phases and events within each cycle can be identified. Various systems for subdividing the gait cycle have been suggested; Perry's^[9] is described here (see Fig. 1). In able-bodied walking, the heel is the first part of the foot to strike the ground. This is not necessarily the case for people with pathology, and the term initial contact is preferred. Conventionally, the gait cycle starts with initial contact on one side and continues until the next initial contact on the same side. This is also referred to as a stride. A step starts with initial contact on one side and continues until initial contact on the other side. Two steps thus constitute a stride.

For each limb, the gait cycle can be divided into two phases: the stance phase, when the foot is in contact



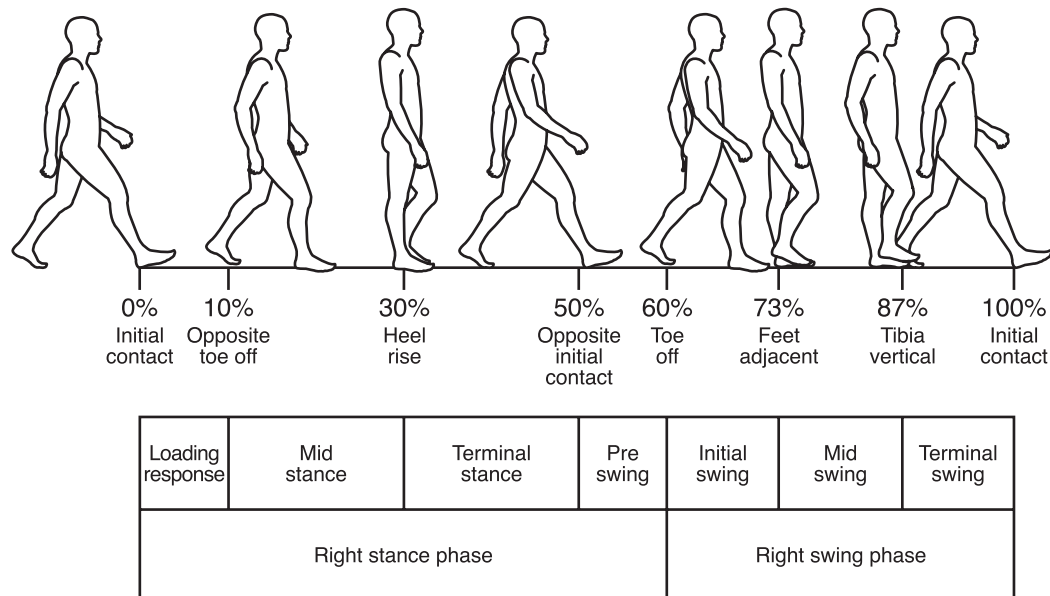


Fig. 1 Terminology to describe the events of the gait cycle.

with the ground, and the swing phase, when it is not. In able-bodied subjects, the stance phase lasts for around 60% of the gait cycle and the swing phase lasts for about 40% of the cycle. It is almost always the toe that is the last part of the foot to leave the ground. This is generally referred to as toe off, although foot off is preferable.

The cycle can be further subdivided by considering what is happening to the other limb. Opposite toe off occurs at around 10% of the gait cycle for able-bodied walking and opposite initial contact occurs at 50% of the cycle. Further divisions of the cycle are depicted in Fig. 1.

Temporal-Spatial Parameters

Definitions

These are the basic time and distance measurements that describe walking. Cadence is the number of steps taken in a given time, generally per minute (occasionally it is reported as strides per minute, which is obviously half the number of steps). Stride length is the distance between the position of any point on the body at the start of the gait cycle and that at the end; by definition this lies in the direction of progression. The base of support (or stride width) is a measure perpendicular to the direction of progression of the separation between the stance phase position of one foot and that of the other. There is no consensus on a precise definition of this. Stride time is that required to complete a gait cycle. The ratio of stride length to stride time is the walking speed; the ratio of

speed to stride length is the cadence. These five parameters are defined for the whole gait cycle and are independent of the side considered.

Step length is the distance between the point of initial contact for one leg and that of subsequent initial contact for the other leg. The component of this along the direction of progression is generally quoted, but occasionally the total distance is used. Step time is that from initial contact to opposite initial contact. Stance time is that between initial contact and toe off; swing time is the time between toe off and initial contact. Double support is the time during which both feet are in contact with the ground, and occurs twice during the gait cycle. If the gait pattern is asymmetric, then the foregoing terms will be different for left and right steps.

Measurement

Temporal-spatial parameters can be directly measured by various methods, but in most conventional gait analysis they are derived from measurements of trajectories of markers used for kinematic analysis. Typically, the position of a marker on the foot will be taken as the reference for stride and step length, and stride width. Detecting the precise instant of initial contact and toe off does not greatly affect measurement of the distance parameters, but can affect measures of the temporal parameters. Force plate data are the most accurate means of detecting these instants if they are available. Autocorrelation techniques on marker trajectories can be used to estimate positions of these events for subsequent gait cycles once they have been identified for any single cycle if the trajectory is

repeatable. Manual identification of the events is required if force plate data are not available or if the walking pattern is not repeatable.

Processing

Little processing is required for temporal-spatial data, although reliability is increased by averaging data, from several strides. Normalization of data is an issue not yet resolved. Both temporal and spatial parameters vary with growth. Dividing lengths by body height or leg length is common but may not fully normalize data.^[10] There is no consensus on the normalization of temporal data. A common approach is to report data with reference to normative data by age cohort, but this makes it difficult to interpret data for a particularly tall or short person within the cohort.

Kinematics

Definitions

Kinematics is the study of movement. Conventional gait analysis focuses on movement at each joint throughout the gait cycle. The most commonly used models in gait analysis,^[8,11,12] all of which are based on the model first fully developed at Newington, assume three degrees of rotational movement at each joint but allow for no translation. Joint movements in three planes are considered, and are the components of a Cardan decomposition of the three-dimensional orientation of the distal with respect to the proximal segment (mathematics more fully described by Wu).^[13]

Cardan angles are sequence-dependent, and for the joints of the extremities, the flexion–adduction–internal rotation sequence is universally applied. It is necessary to describe the orientation of one segment with respect to another, as well as to describe the orientation of at least one segment with respect to the laboratory reference system. The pelvis is used for this. Although the Cardan sequence equivalent for the extremities (tilt–obliquity–rotation) has conventionally been used for the pelvis, it has recently been suggested that the opposite sequence (rotation–obliquity–tilt) gives measurements that are closer to the clinical definition of these terms.^[14]

Measurement

Most contemporary gait analysis systems detect the position of markers placed on the subject. The most widespread systems are based on video camera technology.^[15] Measurement volume is illuminated with a strobed light source in either the visible or infrared

spectrum. The markers are covered in retroreflective tape that reflects light back to the camera. Using pulses of very brief duration, very high-intensity light can be used without risk of damaging eyesight. High-intensity light can eliminate any blurring caused by movement of the markers during capture of the image. Because of the brightness of the reflection from the markers, threshold detection can separate the image of the markers from any background detail. Modern systems generally use some form of pattern recognition in this process to pick up only circular reflections and determine their centroids. Capture rates of 50–120 Hz are fairly common in modern systems and much higher rates are possible for specific applications.

Detection generally occurs in a firmware coordinate generator, the output of which is the two-dimensional coordinates of each of a number of markers in the image plane of the camera. Most systems now incorporate some form of correction factor for non-linear optical or electronic systems within the camera. If more than one camera sees a given marker and the cameras' position, orientations, and magnification factors (focal lengths) are known, then the three-dimensional coordinates can be reconstructed. A calibration procedure to determine these parameters is generally conducted before each test session.

To determine joint angles it is first necessary to define a number of segments on each of which a minimum of three points are defined. The Newington model was developed at a time when system resolutions were low and processing time critical. It is thus based on the minimum number of markers required to define seven segments of the lower limbs (two feet, two legs, two thighs, and one pelvis). The model is hierarchical in that the definition of distal segments depends on the prior definition of the proximal segments. Although modern systems are capable of dealing with many more markers, more complex models have not yet been accepted into widespread routine use.

In order to define joint angles it is first necessary to define a meaningful coordinate system (CS) for each segment. This is referred to as the anatomical CS. To place markers accurately, their position must be defined with reference to bony landmarks, and they must be placed in positions that are visible to the cameras and not likely to be knocked off by the impingement of other body segments. It is thus rarely possible to place markers in positions that allow direct determination of the anatomical CS. It is more common for a technical CS to be defined directly from marker positions. A transformation is then applied to the technical CS to obtain the anatomical CS. This transformation can be determined on the basis of assumptions about the anatomy, anthropometric measurements, or data collected during a static calibration trial.



Processing

The data provided by the coordinate generator are a series of two-dimensional marker coordinates in the image plane of each camera. The first stage in reconstructing the three-dimensional marker positions is to decide which of the images in which of the cameras have been created by which of the markers. Once this has been determined, the data from camera calibration can be used to determine the three-dimensional coordinates. It is then necessary to link each of the detected markers with one of the markers on the subject. This process is now either semi- or fully-automated.

Markers on a given segment are then used to determine a technical CS with a defined origin (generally one of the markers) and axis system (generally using a Gram-Schmidt orthogonalization procedure). This technical CS is then transformed to determine the anatomical reference system for the segment.

In the Newington model,^[8] (Fig. 2), the anatomical CS for the pelvis has its origin at the midpoint of the anterior superior iliac spines (ASIS) of the pelvis, its medial-lateral axis from the right ASIS to the left ASIS, the anterior-posterior axis perpendicular to this and in the plane containing it and the midpoint of the posterior superior iliac spines (PSIS). The proximal-distal axis is orthogonal to the other two axes. Markers

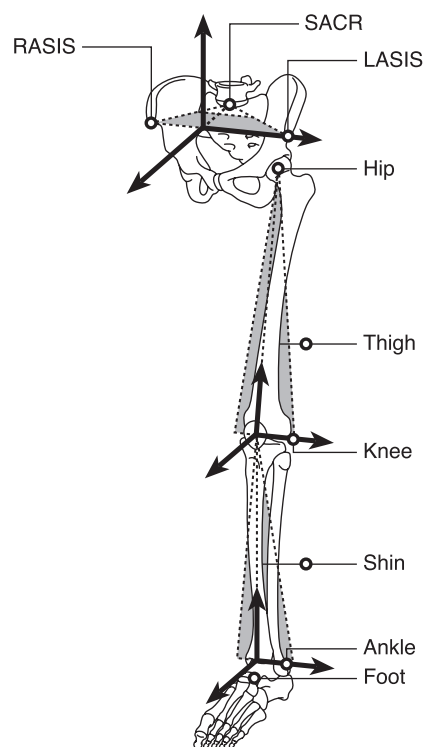


Fig. 2 Definition of Co-ordinate Systems in the Newington model (see text for full description).

(LASIS, RASIS, and SACR) are placed over each ASIS and the mid-point of the PSIS. The only transformation required to obtain the anatomical CS from the technical CS is thus to correct for the marker centers' distance from the bony landmarks.

Locating markers on bony landmarks on the femur is problematic. There is significant skin movement over the greater trochanter (particularly during transverse plane rotations of the hip), and the medial epicondyle is subject to being knocked off by the opposite limb. The only useful landmark is the lateral epicondyle (Knee). Another point of the femur is obtained by recognizing that the center of the femoral head is coincident with the center of the acetabulum in most people. The center of the acetabulum can be estimated on the basis of regression equations relating it to the origin of the pelvis in the anatomical pelvic CS. The third marker is generally on a short wand strapped to the lateral aspect of the thigh.

The anatomical axis system for the femur has its origin at the knee joint center (approximated by the midpoint of the medial and lateral femoral epicondyles). Its proximal-distal axis is from the knee joint to the hip joint center, and its medial-lateral axis is perpendicular to this and in a plane containing the knee axis (approximated by the line from medial to lateral epicondyle). The anterior-posterior axis is orthogonal to both. The transformation from the technical frame to the anatomical frame requires a rotation and translation. Two approaches to quantifying the transverse plane component of the rotation are used. The first is to place the thigh marker (Thigh) accurately so that the coronal plane of the femur's technical CS is coplanar with the anatomical CS in which no rotation is required. The second is to place additional markers to mark the knee axis during a static calibration and to use the data from this to calculate the appropriate degree of rotation.

Although the tibia does have more reliable bony landmarks than the femur, the method for defining it is analogous to the method for the femur. One marker is placed on the lateral malleolus (Ankle); another on a wand on the lateral aspect of the lower leg (Shin); and the knee joint centre, as defined in the femoral CS, is taken as the third point. The ankle joint axis is assumed to pass from the medial to the lateral malleolus, and the joint center is the midpoint of these. This joint center is the origin of the tibial CS, and the axes are then defined as for the femur. Again, either exact placement of the lower leg marker can be used to ensure that the technical and anatomical tibial coronal planes are aligned, or a rotation offset between the two can be calculated during the static test. This is calculated to ensure that knee rotation is zero after a clinical measurement of tibial torsion has been taken into account.

At the time the Newington model was devised, system resolution and the minimum marker size required for detection were such that only one marker on the forefoot could be detected. The Newington model thus allows only the long axis of the foot to be defined, and does not define rotations about this. The line between a forefoot marker (Foot) and the ankle joint center as defined in the tibial CS provides the technical reference for this. Both sagittal and transverse plane offsets are calculated during a static calibration file to quantify the transformation to the anatomical equivalent.

When the anatomical CS for each of the segments has been determined, then the joint angles must be determined. The Cardan angle decomposition is the sequence of rotations about orthogonal axes fixed in an imagined moving segment, starting aligned with the proximal segment and ending aligned with the distal segment. This can also be visualized as the moving segment being rotated about an axis fixed in the proximal segment, an axis mutually perpendicular to this and an axis fixed in the distal segment, and then about that axis in the distal segment. This is the Grood and Suntay^[16] representation, which is often viewed as different from the Cardan representation. It should be noted that the two are simply different ways of visualizing the same underlying mathematical quantities.^[17]

The Cardan angles are derived from the direction cosine matrix, expressing the relative orientation of the anatomical CS for the segment distal to the joint with respect to the segment proximal to it. There are generally two solutions to the resulting equations, and the one with the middle rotation lying in the range -90° to $+90^\circ$ is chosen. At the point at which the middle angle is equal to 90° there is a singularity with the equations, giving an infinite number of solutions. Close to this point, the three angles that result from the Cardan decomposition are increasingly sensitive to small perturbations in orientation, and a clinical interpretation of the results is not meaningful. This is called gimbal lock and is analogous to the ambiguity in defining the directions north, south, east, and west as one gets very close to one of the poles. The effect can be ignored if the middle angle lies in the range -75° to $+75^\circ$ (which is almost always the case in gait analysis). The gimbal lock effect, is however, a major concern in representing movement at other joints, particularly the shoulder.

Kinetics

Definitions

Kinetics is the study of the forces and moments that bring about movement. In clinical gait analysis,

the moments that muscles generate about joint centers are assumed to be the cause of movements of the joints, and hence of body segments. Determining the forces and moments generated by specific internal structures (e.g., tendons, ligaments, joint surfaces) is an indeterminate problem. Although various optimization approaches have been used to estimate these forces and moments, conventional gait analysis tends to be based on the external moments developed about the joints that can be determined uniquely. These moments are generally described as the three components about the axes of the orthogonal CS aligned with the anatomical CS of the segment distal to the joint but with its origin at the joint center. Note that the components of joint kinematics and kinetics are defined about different axes systems; care is needed in correlating features between them.

There is no consensus on whether moments should be described as internal (the moment applied by the muscles to the joints) or external (the moment arising as a consequence of the ground reaction). The magnitude of the moments is completely unaffected by the choice of internal vs. external, and valid arguments exist for each. One convention should be adopted and used consistently, and this should be clearly documented whenever data are presented.

It is also common to see joint powers presented. These are the scalar product of joint angular velocity and joint moment. Although joint angles are defined about nonorthogonal axes, angular velocity is defined as the change in angular position with time in the limit as the time interval approaches zero. As this limit is approached, the components of the change in angular position approach orthogonality, thus angular velocity is a vector quantity and its components are about the same axes as the joint moment. Some systems present three components of joint power, which have no physical significance (power being a scalar variable), but which can be interpreted as having some clinical significance.

Measurement

The principal requirement for determining joint kinetics is measures of the force system applied by the floor on the foot (which is equal and opposite to that applied by the foot on the floor). The full six components of the force system must be recorded. These are probably most easily visualized as the three components of the ground reaction force, the moment about the vertical axis, and the coordinates of the center of pressure (the point on the force plate's surface at which the ground reaction can be considered to be acting). Most force plates designed to make such measurements are based on either piezoelectric

or strain-gauged technology. These plates must be designed so the minimum resonant frequency is considerably higher than the maximum frequency components of the ground reaction, which are generally assumed to be of the order of 10–20 Hz.

Processing

Two steps are required to calculate joint moments from force plate data. The first is to describe the ground reaction in relevant CS; the second is to take into account the gravitational and inertial effects on the segments distal to the joint being considered. This is known as inverse dynamics and requires calculation of the acceleration of these segments and an estimation of their inertial parameters (mass, position of center of mass, and radii of gyration about the principal axes). The acceleration of the segments is calculated from the kinematic analysis. The inertial parameters are taken from tables scaled to body weight and segment length. Early U.S. military data^[18–20] are most commonly used, despite concerns over their relevance. (The data are drawn from elderly cadavers, not from the typical subjects of most clinical gait analyses).^[21–24]

THE PROCESS OF GAIT ANALYSIS

Conduct of Tests

The measurement volume of most clinical gait analysis systems is a few meters in length and around two meters in width and height. It is important to record the cyclic part of walking after initiation and before the process of stopping commences. There is little hard evidence as to how many steps are involved, but most labs allow subjects a minimum of four steps before entering and after leaving the measurement volume.

The conditions of the analysis (i.e., use of footwear, splints, walking aids) will depend on the reason for the analysis. Assessments of functional performance require the subject to be walking under conditions most similar to those of everyday life. An analysis to plan surgery, however, may focus on the intrinsic capabilities of the patient, in which case walking in bare feet with minimal aids is likely to be most useful.

The inverse dynamic calculations of joint moments will not be valid if the foot has only partially hit the force plate or if anything else (e.g., the other foot or a walking aid) makes contact with the force plate at the same time. The calculation will be valid if a walking aid is used, provided it does not make contact with the force plate, and care is taken in interpretation of the data. Subjects are generally not told about the force plates, because their walking pattern may be affected

by targeting the plates. Data are only used from those walks on which clean force plate contact is made (although this itself may lead to some bias in the data analyzed). It is most common to ask subjects to walk at a comfortable walking speed, although there is some evidence that the speed selected may not be representative of the speed they use to walk in other situations.

Presentation of Data

All gait variables (i.e., those that vary with time through the gait cycle) are conventionally presented in the same graphical format (see Fig. 3). Time is plotted along the horizontal axis as a percentage of the gait cycle, starting at foot contact on one side and ending at foot contact on the same side. The gait variable is plotted on the vertical axis. Plot areas often have an aspect ratio of close to 3:4. Keeping a fixed aspect ratio makes interpretation of gait data more intuitive. Gait events are also often plotted. A vertical line across the full height of the plot is often used to mark toe off and therefore distinguish the stance phase (to the left) and swing phase (to the right). Opposite toe off and opposite initial contact are often represented as tick marks along the top or bottom of the plot area.

Data representing measurements of a particular gait variable recorded for able-bodied subjects are generally plotted on graphs to enable comparison with data

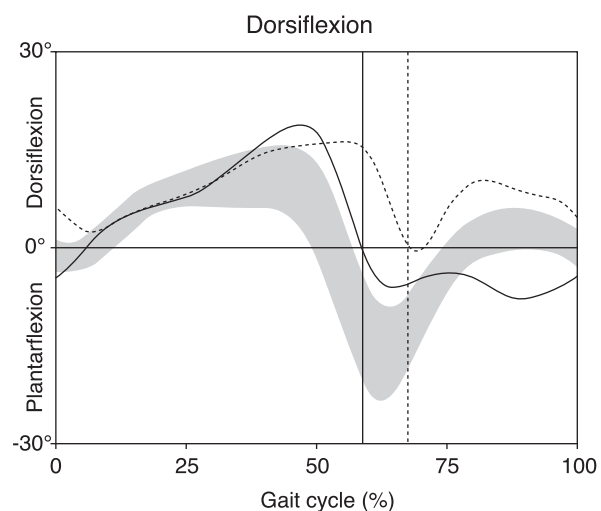


Fig. 3 Graph format for all gait variables. Time is plotted along the x-axis from initial contact to the next initial contact for the same foot. Units are % of gait cycle. The variable is plotted on the y-axis using the relevant units. The left trace is plotted as a solid line and the right unit dashed. Toe off for both sides is marked by the vertical line across the whole graph. The ± 1 standard deviation range about the ensemble average for a cohort of able-bodied comparisons is plotted in light grey. Graphs are all plotted using the same aspect ratio (0.75 in this case).

from a patient. They are often represented as a band representing the ensemble average of data for a number of subjects and a range of a specified number of standard deviations.

Data for the left and right sides are generally plotted on the same axes using different colors. Because the gait cycles for left and right sides start at different times, a point on the graph representing a specific instant during the left-gait cycle represents a different instant in time than does the same point on the graph for the right side. It is particularly important to understand this in relation to pelvic angles because there is only one pelvis, and any differences between the traces for the right and left sides are a consequence of the fact that the same point on the horizontal axis of

the graphs represents the orientation at different times for the left trace and the right trace. It is important to understand the repeatability of a subject's gait pattern, often accomplished by plotting data from several walking trials on the same axes.

Figure 4 shows graphs of individual gait variables, presented as an array on the page. This allows comparison of variables in the same plane for different joints and also of different planes for the same joint. A common arrangement is for the different joints to be represented by different rows (say, pelvis, hip, knee, and for ankle going down the page) and for the different components of the Cardan angle decomposition to be represented by different columns (say, sagittal, coronal, and transverse).

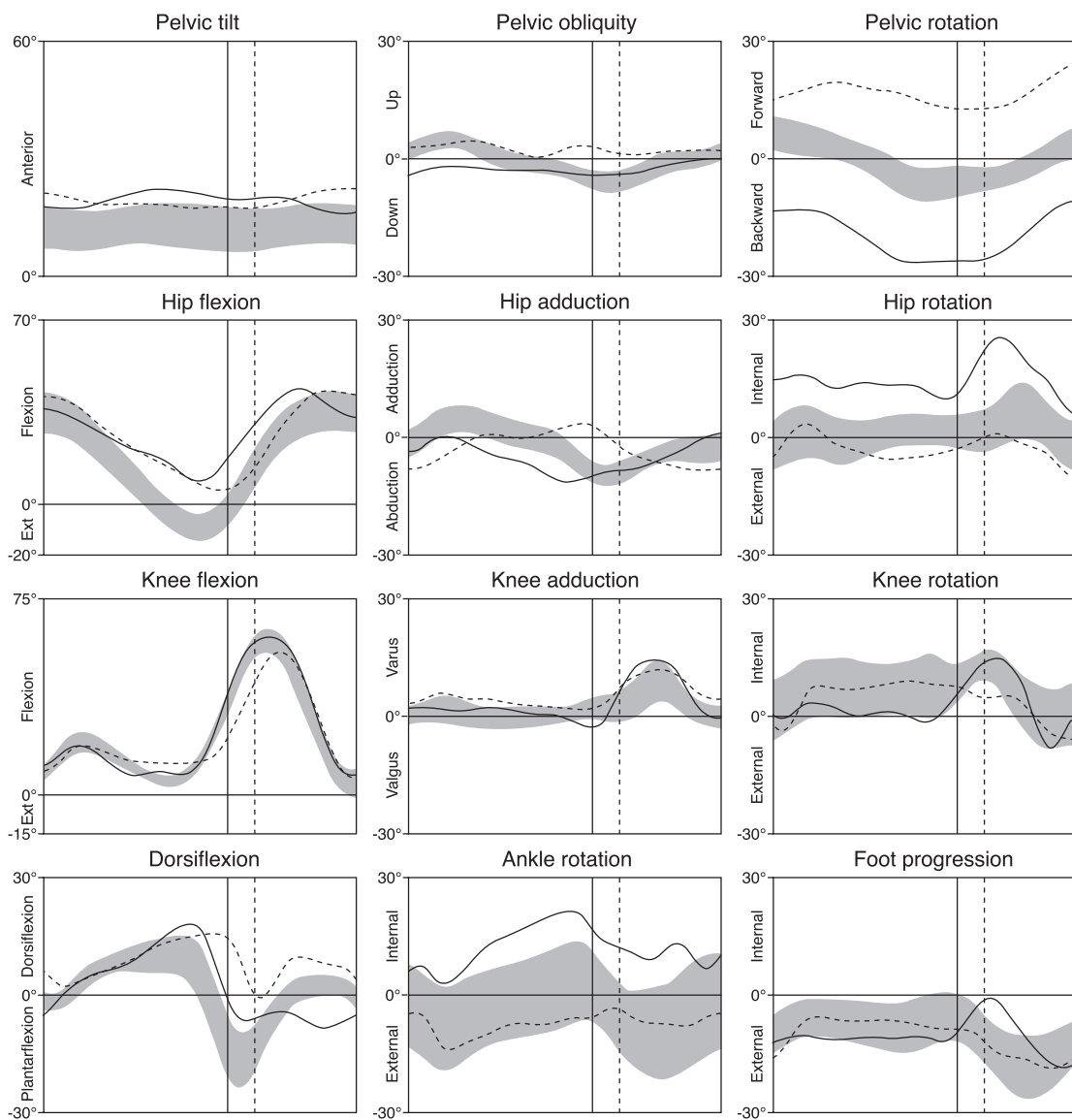


Fig. 4 Layout of graphs for kinematic data. Rows represent different joints (pelvis, hip, knee, ankle in descending order). Columns represent the different planes (sagittal, coronal, and transverse from left to right).

Interpretation of Data

Temporal-spatial data are most often interpreted in comparison to normative data to give an overall assessment of performance. Comparison with subjects' performance at earlier analyses can also provide useful information, but the effects of the subjects' growth on these parameters must always be taken into account. Sometimes comparison of cadence and stride length can give some indication of what is limiting walking speed. Comparison of data for left and right sides can give some indication of gait asymmetry. Subjects will generally spend less time on the weaker or more painful side and will have a shorter step length. Various symmetry indices have been proposed to quantify this.

Detailed interpretation of kinematic and kinetic data is dependent on the diagnosis of the patient, their own functional abilities, and the specific question that the analysis is conducted to address. Although this is beyond the scope of this article, some general guidelines can be given. Interpretation should be carried out following a structured plan. The exact order in which the data are examined is not greatly important, but interpretation should cover at least the issues discussed in the next section.

The first stage of any data interpretation must be an assessment of how reliable the data are. Comparison of data from different trials is required to establish how consistent a subject's gait pattern is. Gait analysis depends on examining a small number of walking trials, assuming that these data are representative of the patient's overall pattern of walking. Evidence of marked variability from walk to walk invalidates the entire approach.

It is also important to be critical of how valid the data are. Certain features intrinsic to the kinematic data can be considered. For example, are there discontinuities in the traces (very rare in human walking)? Is there evidence of cross talk between the knee sagittal and coronal plane kinematics? Is there an appreciable ankle moment in the swing phase? It is also important to ensure that the data is consistent with observation of the patient walking either directly or using video recordings. Gait analysis gives more data than video recordings. The data may explain what is seen but it should never contradict it. For example a coronal plane video recording may clearly show that the knee is being brought across the midline. This may be a result of hip adduction or a combination of pelvic rotation and hip flexion. Gait analysis can be used to determine which of these is happening but must be invalid if it records that the knee is being taken away from the midline. Certain gait data are plotted with respect to the laboratory axis system (pelvic angles and foot progression); these can be straightforwardly compared with video recordings.

If some traces are deemed to be invalid then other traces will necessarily be invalid also. For example, cross talk between knee sagittal and coronal kinematics indicates unreliable transverse plane hip rotations. A swing phase ankle moment almost always indicates invalid moment and power data for other levels. This may be restricted to the swing phase or apply to the entire gait cycle.

The next stage of interpretation is to identify which features of the gait pattern are abnormal. A gait trace can be abnormal in having a more or less fixed offset from the mean trace for able-bodied norms. A gait trace may have the same pattern but exaggerated or diminished features. Finally, the trace may show features that are not present at all in the able-bodied gait. Although the identification of abnormal features may appear to be an objective process, some judgment is required in deciding how large a feature must be to be deemed to be real. Some purists claim that a divergence from the able-bodied mean of at least 2 standard deviations (SD) is required. It must be remembered, however, that much of the variability in the comparative able-bodied data is a consequence of measurement error; rigorously applying the 2-SD rule will mean that some abnormalities are passed over. It must also be appreciated that a trace may be abnormal if the pattern is abnormal, even if it lies entirely within the able-bodied range. Gait traces should be analyzed systematically to identify these abnormalities.

The next stage is to link the identified abnormalities into patterns, where possible. Links can be between different levels; across planes; and between kinetic, kinematic, and EMG data. Flexed hips, knees, and dorsiflexed ankles, for example, represent crouch gait. This almost always results in increased internal knee extensor moments. Any explanation of the abnormalities should concentrate on explaining the overall pattern rather than the individual features.

Once abnormalities have been identified and linked, an attempt must be made to explain them in terms of the patient's pathology. A key factor here is the incorporation of data from the physical examination and EMG studies. Restricted joint movement is often a result of fixed contractures or may be caused by under- or overactivity in specific muscles. Such explanations are rarely either completely objective or completely certain; some indication of the degree of both objectivity and certainty should be recorded, most commonly by the words chosen to describe the features in the report.

It can be argued that this biomechanical explanation of observations ends formal gait analysis, and that decisions as to how the patient should be managed are the realm of the clinician to whom any report is made. It is, common however, for referring clinicians to seek an expert clinical opinion at the time of the

referral. In this case the biomechanical explanation and the clinical recommendations should be clearly distinguished. It is also important that the personnel performing the analysis give clinical recommendations only if they are fully qualified to do so.

RECENT DEVELOPMENTS

Conventional gait analysis has now been in use for around 20 years. In addition to employing these procedures, many centers have done considerable work incorporating other developments into their testing. The most important of these are described in the following section, although none has yet been accepted into widespread clinical practice.

Optimization Techniques in Modeling Gait Data

Perhaps the greatest challenge to clinical gait analysis is obtaining valid basic data. For example, the Newington model^[8] can be very sensitive to small errors in marker placement; the skin on which markers are mounted may move with respect to underlying bone; there is some intersubject variability of the bony anatomy; and the validity of some of the underlying assumptions of clinical gait analysis can be challenged.

The most promising responses to these challenges are optimization approaches that relate marker positions to the underlying anatomy. There have been several attempts to apply such techniques to individual joints,^[25–29] but recent global optimization models applied to the whole body and throughout the measurement period are now proving successful, and are being incorporated into commercially available systems.^[30]

Foot and upper body modeling

For some pathologies the movement of the upper body is an important constituent of gait. For others the assumption that the foot acts as a single segment is invalid. Many detailed models of the trunk^[31–45] and feet^[46–50] have been developed, but none has achieved widespread acceptance. Foot models generally require the placement of additional markers close together on the foot, and are only becoming clinically applicable as measurement technology improves.

Muscle length modeling

In many pathologies for which gait analysis is used, the underlying problem is how the muscles are functioning. Kinematic traces for individual joints offer considerable

insight into how uni-articular muscles contract and elongate, but making inferences about multijoint muscles is much more difficult. By modeling the anatomy of the musculoskeletal system and driving this model by joint kinematics it is possible to estimate the lengths of multijoint muscles, thus giving insight into how they function. Muscle length modeling is probably the most widely used of the techniques discussed here.^[51–57]

Moment arm modeling

The moment that a particular muscle can develop is proportional to its moment arm at a particular joint. In some patients this is critical to the functioning of muscles, and is described by the term, lever arm disease. In two dimensions the moment arm is the perpendicular distance from the joint to the line of action of the muscle. This does not extend to three dimensions, in which case the most common definition is the partial derivative of muscle length with respect to joint angle (component of Cardan angle decomposition). This can be obtained by modeling the path of muscles as a function of the musculoskeletal anatomy and joint angle.^[52,5358–68]

Modelling internal loads

Conventional gait analysis is restricted to consideration of external joint loads. In specific patients, knowledge of internal loads is essential; for example, in determining disease processes in osteoarthritic patients. Based on assumptions about joint anatomy, muscle physiology, and an optimization approach, it is now possible to estimate internal joint loads in certain structures. Duda has established such an approach for the hip,^[69] which has since been validated by measurements from an instrumented hip prosthesis.^[70,71] Similar approaches have been developed for the knee.^[72–74]

Customizing models to individual patients

The Newington model makes various assumptions about the subject. More complex models require more assumptions to be made, and are generally more sensitive to variability between subjects. Various imaging techniques—principally, magnetic resonance imaging (MRI) and computer assisted tomography (CAT scans)—offer the potential to customize models to individual patients. Although considerable research in these areas is ongoing, the cost of obtaining MRI, and CAT scans will probably prevent their widespread clinical use in the immediate future.^[51,60]



Center-of-mass modeling

There has been a long history of interest in how the center of mass moves during walking. Although the insight that center-of-mass modeling gives into the interpretation of gait analysis data has probably been overestimated, there are certain cases where a knowledge of its trajectory does help explain the gait pattern. Models tend to be based on kinematic models involving estimations of the mass and center-of-mass position of each segment of the body. Some studies have validated these models by comparing them with the second integral of the total ground reaction.^[75–81]

Forward simulation

There has been considerable debate (which stimulated the Weber brothers as long ago as 1835) over whether the swing phase of gait is passive or requires muscle action. It is now clear that muscle activity is required, but the balance between initial conditions and ongoing muscle activation in controlling the swing phase is often a key issue in relation to pathological gait. Forward simulation of swing phase motion allows these two to be separated and allows better informed interpretation of gait data.^[82–93]

Induced acceleration analysis

Because the external joint moments act within either an open- or closed-chain structure, there is not a simple relationship between these and joint angular velocities. Methods have recently been established to determine how much of the angular acceleration observed at a particular joint can be attributed to either the external moment generated at that on another joint or, to the tension generated in a specific muscle. Although there are specific technical problems, particularly in modelling contact of the foot with the ground, these methods are perhaps the most exciting recent advances in this field.^[87,94–97]

CONCLUSION

Clinical gait analysis is an increasingly important part of the management of patients with musculoskeletal disorders affecting the lower limbs. The procedures established in Newington in the early 1980s have become widely implemented and can now appropriately be referred to as conventional clinical gait analysis. Many other techniques are now being developed, but none has so far had widespread clinical acceptance. Many of these techniques will make a considerable impact on the interpretation of data for specific

patients, and it is hoped that the next decade will see them implemented in routine clinical practice.

ARTICLE OF FURTHER INTEREST

Ergonomics

REFERENCES

1. Aristotle. *Aristotle: Parts of Animals, Movement of Animals, Progression of Animals*; Harvard University Press, 1968.
2. Borelli, G.A. *On the Movement of Animals*; Springer Verlag: Berlin, 1989.
3. Weber, W.E.; Weber, E. *Mechanics of the Human Walking Apparatus*; Springer-Verlag: Berlin, 1991.
4. Braun, M. *Picturing Time: The Work of Etienne-Jules Marey (1830–1904)*; University of Chicago Press: Chicago, 1992.
5. Muybridge, E. *Muybridge's Complete Human and Animal Locomotion: All 781 Plates From the 1887 Animal locomotion*; Dover Publications: New York, 1979.
6. Braune, W.; Fischer, O. *The Human Gait*; Springer-Verlag: Berlin, 1987.
7. Klopsteg, P.E.; Wilson, P.D. *Human Limbs and Their Substitutes*; McGraw-Hill: New York, 1954.
8. Davis, R.; Ounpuu, S.; D, T.; JR, G. A gait analysis data collection and reduction technique. *Hum. Mov. Sci.* **1991**, *10*, 575–587.
9. Perry, J. *Gait Analysis*; SLACK: Thorofare, 1992.
10. O'Malley, M.J. Normalization of temporal-distance parameters in pediatric gait. *J. Biomech.* **1996**, *29* (5), 619–625.
11. Kadaba, M.; Ramakrishnan, H.; Wootten, M. Measurement of lower extremity kinematics during level walking. *J. Orthop. Res.* **1990**, *8* (3), 383–391.
12. Kadaba, M.; Ramakrishnan, H.; Wootten, M.; Gainey, J.; Gorton, G.; Cochran, G. Repeatability of kinematic, kinetic, and electromyographic data in normal adult gait. *J. Orthop. Res.* **1989**, *7* (6), 849–860.
13. Wu, G. Kinematics Theory. In *Gait Analysis: Theory and Application*; Craik, R., Oatis, C., Eds.; Mosby-Year Book: St Louis, 1995.
14. Baker, R. Pelvic angles: A mathematically rigorous definition which is consistent with a conventional clinical understanding of the terms. *Gait Posture* **2001**, *13* (1), 1–6.
15. Koff, D. Joint Kinematics: Camera-Based Systems. In *Gait Analysis: Theory and Practice*; Craik, R., Oatis, C., Eds.; Mosby-Year Book: St. Louis, 1995.
16. Grood, E.; Suntay, W. A joint coordinate system for the clinical description of three-dimensional motions: Application to the knee. *Trans. ASME. J. Biomech. Eng.* **1983**, *105*, 136–143.
17. Chao, E.Y. Justification of triaxial goniometer for te measurement of joint rotation. *J. Biomech.* **1980**, *13*, 989–1006.

18. Chandler, R.; Clauser, C.; McConville, J.; Reynolds, H.; Young, J. *Investigation of Inertial Properties of the Human Body. (AMRL Technical Report: 74-137)*; Wright-Patterson Air Force Base: Ohio, 1975.
19. Clauser, C.; McConville, J.; Young, J. *Weight Volume and Centre of Mass of Segments of the Human Body (AMRL Technical Report)*; Wright-Patterson Air Force Base: Ohio, 1969.
20. Dempster, W. *Space Requirements of the Seated Operator (WADC Technical Report: 550-159)*; Wright-Patterson Airforce Base: Ohio, 1955.
21. De Leva, P. Adjustments to Zatsiorsky-Seluyanov's segment inertia parameters. *J. Biomech.* **1996**, *29* (9), 1223-1230.
22. Jensen, R.K. Body segment mass, radius and radius of gyration proportions of children. *J. Biomech.* **1986**, *19* (5), 359-368.
23. Zatsiorsky, V.; Seluyanov, V. Estimation of the Mass and Inertia Characteristics of the Human Body by Means of the Best Predictive Regression Equations. In *Biomechanics IX-B*; Winter, D., Norman, R., Wells, R., Hayes, K., Patla, A., Eds.; Human Kinetics: Champaign, IL, 1985.
24. Zatsiorsky, V.; Seluyanov, V.; Aruin, A.; Prilutzky, B. Methods of Determining Mass-Inertial Characteristics of Human Body Segments. In *Contemporary Problems of Biomechanics*; Chernyi, G., Regirer, S., Eds.; CRC Press: Massachusetts, 1990.
25. Frigo, C.; Rabuffetti, M. Multifactorial estimation of hip and knee joint centres for clinical application of gait analysis. *Gait Posture* **1998**, *8*, 91-102.
26. Gamage, S.; Lasenby, J. New least squares solutions for estimating the average centre of rotation and the axis of rotation. *J. Biomech.* **2002**, *35*, 87-93.
27. Halvorsen, K.; Lesser, M.; Lundberg, A. A new method for estimating the axis of rotation and the center of rotation. *J. Biomech.* **1999**, *32*, 1221-1227.
28. Leardini, A.; Cappozzo, A.; Catani, F.; Toksvig-Larsen, S.; Petitto, A.; Sforza, V.; Cassanelli, G.; Giannini, S. Validation of a functional method for the estimation of hip joint centre location. *J. Biomech.* **1999**, *32* (1), 99-103.
29. Piazza, S.; Okita, N.; Cavanagh, P. Accuracy of the functional method of hip joint center location: Effects of limited motion and varied implementation. *J. Biomech.* **2001**, *34*, 967-973.
30. Lu, T.-W.; O'Connor, J. Bone position estimation from skin marker co-ordinates using global optimisation with joint constraints. *J. Biomech.* **1999**, *32*, 129-134.
31. Bartonek, A.; Saraste, H.; Eriksson, M.; Knutson, L.; Cresswell, A.G. Upper body movement during walking in children with lumbo-sacral myelomeningocele. *Gait Posture* **2002**, *15* (2), 120-129.
32. Callaghan, J.P.; Patla, A.E.; McGill, S.M. Low back three-dimensional joint forces, kinematics, and kinetics during walking. *Clin. Biomech.* **1999**, *14* (3), 203-216.
33. Cappozzo, A. Analysis of the linear displacement of the head and trunk during walking at different speeds. *J. Biomech.* **1981**, *14* (6), 411-425.
34. Cappozzo, A. The forces and couples in the human trunk during level walking. *J. Biomech.* **1983**, *16* (4), 265-277.
35. Chapman, M.; Kurokawa, K. Some observations on the transverse rotations of the human trunk during locomotion. *Bull. Prosthet. Res.* **1969**, *10*, 38-59.
36. Crawford, N.; Yamaguchi, G.; Dickman, C. Methods for determining spinal flexion/extension, lateral bending, and axial rotation from marker co-ordinate data: Analysis and refinement. *Hum. Mov. Sci.* **1996**, *15*, 55-78.
37. Crosbie, J.; Vachalathiti, R.; Smith, R. Patterns of spinal motion during walking. *Gait Posture* **1997**, *5* (1), 6-12.
38. Gregersen, G.G.; Lucas, D.B. An in vivo study of the axial rotation of the human thoracolumbar spine. *J. Bone Joint Surg.* **1967**, *49* (2), 247-262.
39. Khoo, B.C.; Goh, J.C.; Bose, K. A biomechanical model to determine lumbosacral loads during single stance phase in normal gait. *Med. Eng. Phys.* **1995**, *17* (1), 27-35.
40. Krebs, D.E.; Wong, D.; Jevsevar, D.; Riley, P.O.; Hodge, W.A. Trunk kinematics during locomotor activities. *Phys. Ther.* **1992**, *72* (7), 505-514.
41. Sartor, C.; Alderink, G.; Greenwald, H.; Elders, L. Critical kinematic events occurring in the trunk during walking. *Hum. Mov. Sci.* **1999**, *18*, 669-679.
42. Stokes, V.P.; Andersson, C.; Forssberg, H. Rotational and translational movement features of the pelvis and thorax during adult human locomotion. *J. Biomech.* **1989**, *22* (1), 43-50.
43. Thorstensson, A.; Nilsson, J.; Carlson, H.; Zomlefer, M.R. Trunk movements in human locomotion. *Acta. Physiol. Scand.* **1984**, *121* (1), 9-22.
44. Thurston, A.J.; Harris, J.D. Normal kinematics of the lumbar spine and pelvis. *Spine* **1983**, *8* (2), 199-205.
45. Waters, R.L.; Morris, J.; Perry, J. Translational motion of the head and trunk during normal walking. *J. Biomech.* **1973**, *6* (2), 167-172.
46. Carson, M.; Harrington, M.E.; Thompson, N.; O'Connor, J.; Theologis, T. Kinematic analysis of a multi-segment foot model for research and clinical applications: A repeatability analysis. *J. Biomech.* **2001**, *34*, 1299-1307.
47. Carson, M.; Thompson, N.; Theologis, T. A four segment in vivo foot model for clinical gait analysis. *Gait Posture* **1998**, *8*, 61-81.
48. Kidder, S.; Abuzzahab, F.; Harris, G.; J, J. A system for the analysis of foot and ankle kinematics during gait. *IEEE T. Rehabil. Eng.* **1996**, *4* (1), 33-38.
49. Leardini, A.; Benedetti, M.; Catani, F.; Simoncini, L.; Giannini, S. An anatomically based protocol for the description of foot segment kinematics during gait. *Clin. Biomech.* **1999**, *14* (8), 528-536.
50. Smith, P.; Humm, J.; Hassani, S.; Harris, G. Three Dimensional Motion Analysis of the Pediatric Foot and Ankle. In *Pediatric Gait*; Harris, G., Smith, P., Eds.; IEEE Press, 2000.
51. Arnold, A.; Blemker, S.; Delp, S. Evaluation of a deformable musculoskeletal model for estimating muscle-tendon lengths during crouch gait. *Ann. Biomed. Eng.* **2001**, *29* (3), 1-25.
52. Delp, S.L.; Loan, J.P. A graphics-based software system to develop and analyze models of musculoskeletal structures. *Comput. Biol. Med.* **1995**, *25* (1), 21-34.



53. Delp, S.L.; Loan, J.P.; Hoy, M.G.; Zajac, F.E.; Topp, E.L.; Rosen, J.M. An interactive graphics-based model of the lower extremity to study orthopaedic surgical procedures. *IEEE T. Bio-Med. Eng.* **1990**, *37* (8), 757–767.
54. Eames, N.W.; Baker, R.; Cosgrove, A. Defining gastrocnemius length in ambulant children. *Gait Posture* **1997**, *6*, 9–17.
55. Eames, N.W.; Baker, R.; Hill, N.; Graham, K.; Taylor, T.; Cosgrove, A. The effect of botulinum toxin A on gastrocnemius length: Magnitude and duration of response. *Dev. Med. Child. Neur.* **1999**, *41* (4), 226–232.
56. Frigo, C.; Nielsen, J.; Crenna, P. Modelling the triceps surae muscle-tendon complex for the estimation of length changes during walking. *J. Electromyogr. Kinesiol.* **1996**, *6*, 191–203.
57. Vasavada, A.N.; Delp, S.L.; Maloney, W.J.; Schurman, D.J.; Zajac, F.E. Compensating for changes in muscle length in total hip arthroplasty. Effects on the moment generating capacity of the muscles. *Clin. Orthop. Rel. Res.* **1994**, *302*, 121–133.
58. Arnold, A.S.; Asakawa, D.J.; Delp, S.L. Do the hamstrings and adductors contribute to excessive internal rotation of the hip in persons with cerebral palsy? *Gait Posture* **2000**, *11* (3), 181–190.
59. Arnold, A.S.; Delp, S.L. Rotational moment arms of the medial hamstrings and adductors vary with femoral geometry and limb position: Implications for the treatment of internally rotated gait. *J. Biomech.* **2001**, *34* (4), 437–447.
60. Arnold, A.S.; Salinas, S.; Asakawa, D.J.; Delp, S.L. Accuracy of muscle moment arms estimated from MRI-based musculoskeletal models of the lower extremity. *Comput. Aided. Surg.* **2000**, *5* (2), 108–119.
61. Delp, S.; Maloney, W. Effects of hip center locations on the moment-generating capacity of the muscles. *J. Biomech.* **1993**, *26* (4/5), 485–499.
62. Delp, S.; Ringwelski, D.A.; Carroll, N.C. Transfer of the rectus femoris: Effects of transfer site on moment arms about the knee and hip. *J. Biomech.* **1994**, *27* (10), 1201–1211.
63. Delp, S.; Statler, K.; Carroll, N.C. Preserving plantar flexion strength after surgical treatment for contracture of the triceps surae: A computer simulation study. *J. Orthopaed. Res.* **1995**, *1* (1), 96–104.
64. Delp, S.L.; Bleck, E.E.; Zajac, F.E.; Bollini, G. Biomechanical analysis of the Chiari pelvic osteotomy. Preserving hip abductor strength. *Clin. Orthop. Rel. Res.* **1990**, *254*, 189–198.
65. Delp, S.L.; Hess, W.E.; Hungerford, D.S.; Jones, L.C. Variation of rotation moment arms with hip flexion. *J. Biomech.* **1999**, *32*, 493–501.
66. Delp, S.L.; Wixson, R.L.; Komattu, A.V.; Kocmond, J.H. How superior placement of the joint center in hip arthroplasty affects the abductor muscles. *Clin. Orthop. Relat. Res.* **1996**, *328*, 137–146.
67. Delp, S.L.; Zajac, F.E. Force- and moment-generating capacity of lower-extremity muscles before and after tendon lengthening. *Clin. Orthop. Rel. Res.* **1992**, *284*, 247–259.
68. Free, S.A.; Delp, S.L. Trochanteric transfer in total hip replacement: Effects on the moment arms and force-generating capacities of the hip abductors. *J. Orthopaed. Res.* **1996**, *14* (2), 245–250.
69. Duda, G.N.; Schneider, E.; Chao, E.Y. Internal forces and moments in the femur during walking. *J. Biomech.* **1997**, *30* (9), 933–941.
70. Bergmann, G.; Deuretzbacher, G.; Heller, M.; Graichen, F.; Rohlmann, A.; Strauss, J.; Duda, G.N. Hip contact forces and gait patterns from routine activities. *J. Biomech.* **2001**, *34* (7), 859–871.
71. Heller, M.O.; Bergmann, G.; Deuretzbacher, G.; Durselen, L.; Pohl, M.; Claes, L.; Haas, N.P.; Duda, G.N. Musculo-skeletal loading conditions at the hip during walking and stair climbing. *J. Biomech.* **2001**, *34* (7), 883–893.
72. O'Connor, J.; Feikes, J.; Gill, R.; Zavatsky, A. Mobility of the Knee. In *Daniel's Knee Injuries*; D, P., J., O.C., W, A., Eds.; Lippincott, Williams & Wilkins: Philadelphia, 2003.
73. O'Connor, J.; Gill, R.; Zavatsky, A. Stability of the Knee. In *Daniel's Knee Injuries*; Pedowitz, D., O'Connor, J., Akesson, W., Eds.; Lippincott, Williams & Wilkins: Philadelphia, 2003.
74. O'Connor, J.; Lu, T.-W.; Wilson, D.; Feikes, J.; Lardini, A. Computer methods in biomechanics and biomedical engineering—Kinematic pairs, mechanisms or flexible structures? *Comp. Methods Biomech. Biomed. Eng.* **1997**, *1*, 123–150.
75. Cavagna, G.A. Force platforms as ergometers. *J. Appl. Physiol.* **1983**, *39* (1), 174–179.
76. Eames, M.; Cosgrove, A.; Baker, R. Comparing methods of estimating the total body centre of mass in three-dimensions in normal and pathological gaits. *Hum. Mov. Sci.* **1999**, *18*, 637–646.
77. Saini, M.; Kerrigan, D.; Thirunarayan, M.; Duff-Rafaele, M. The vertical displacement of the centre of mass during walking: A comparison of four measurement methods. *J. Biomech. Eng.* **1998**, *120*, 133–139.
78. Shimba, T. An estimation of centre of gravity from force plate data. *J. Biomech.* **1984**, *17*, 53–60.
79. Tesio, L.; Lanzi, D.; Detrembleur, C. The 3-D motion of the centre of gravity of the human body during level walking. I. Normal subjects at low and intermediate walking speeds. *Clin. Biomech.* **1998**, *13* (2), 77–82.
80. Tesio, L.; Lanzi, D.; Detrembleur, C. The 3-D motion of the centre of gravity of the human body during level walking. II lower limb amputees. *Clin. Biomech.* **1998**, *13* (2), 83–90.
81. Whittle, M. Three-dimensional motion of the center of gravity of the body during walking. *Hum. Mov. Sci.* **1997**, *16*, 347–355.
82. Anderson, F.C.; Pandy, M.G. Dynamic optimization of human walking. *J. Biomech. Eng.* **2001**, *123* (5), 381–390.
83. Anderson, F.C.; Pandy, M.G. Static and dynamic optimization solutions for gait are practically equivalent. *J. Biomech.* **2001**, *34* (2), 153–161.
84. Anderson, F.C.; Ziegler, J.M.; Pandy, M.G.; Whalen, R.T. Application of high-performance computing to numerical simulation of human movement. *J. Biomech. Eng.* **1995**, *117* (1), 155–157.
85. Jonkers, I.; Spaepen, A.; Papaioannou, G.; Stewart, C. An EMG-based, muscle driven forward simulation of

- single support phase of gait. *J. Biomech.* **2002**, *35* (5), 609–619.
86. Kerrigan, D.; R, R.; PO, R. The modelling of adult spastic paretic stiff-legged gait swing period based on actual kinematic data. *Gait Posture* **1998**, *7* (2), 117–124.
87. Neptune, R.R.; Kautz, S.A.; Zajac, F.E. Contributions of the individual ankle plantar flexors to support, forward progression and swing initiation during walking. *J. Biomech.* **2001**, *34* (11), 1387–1398.
88. Pandy, M.G.; Anderson, F.C. Dynamic simulation of human movement using large-scale models of the body. *Phonetica* **2000**, *57* (2–4), 219–228.
89. Piazza, S.J.; Delp, S.L. The influence of muscles on knee flexion during the swing phase of gait. *J. Biomech.* **1996**, *29* (6), 723–733.
90. Riley, P.O.; Kerrigan, D.C. Torque action of two-joint muscles in the swing period of stiff-legged gait: A forward dynamic model analysis. *J. Biomech.* **1998**, *31* (9), 835–840.
91. Tashman, S.; Zajac, F.E.; Perakash, I. Modeling and simulation of paraplegic ambulation in a reciprocating gait orthosis. *J. Biomech. Eng.* **1995**, *117* (3), 300–308.
92. Veltink, P.H.; Franken, H.M.; Van Alste, J.A.; Boom, H.B. Modelling the optimal control of cyclical leg movements induced by functional electrical stimulation. *Int. J. Artif. Organs* **1992**, *15* (12), 746–755.
93. Yamaguchi, G.T.; Moran, D.W.; Si, J. A computationally efficient method for solving the redundant problem in biomechanics. *J. Biomech.* **1995**, *28* (8), 999–1005.
94. Anderson, F.; Pandy, M. Individual muscle contributions to support in normal walking. *Gait Posture*. **2003**, *17* (2), 159–169.
95. Kepple, T.; Siegel, K.; Stanhope, S. Relative contributions of the lower extremity joint moments to forward progression and support during gait. *Gait Posture* **1997**, *6*, 1–8.
96. Schwartz, M.; Lakin, G. The effect of tibial torsion on the dynamic function of the soleus during gait. *Gait Posture* **2003**, *17* (2), 113–118.
97. Siegel, K.; Kepple, T.; Stanhope, S. Joint moment control of mechanical energy flow during normal gait. *Gait Posture* **2004**, *19* (1), 69–75.

Gene-Activated Matrix

Neil Davies

Cardiovascular Research Unit, University of Cape Town, Cape Town, South Africa

INTRODUCTION

The gene-activated matrix (GAM), a system in which a gene delivery vector is contained within a biodegradable scaffold (Fig. 1),^[1,2] is one of those fertile regions in science in which the paths of two major disciplines, namely gene therapy and tissue engineering, cross. Gene therapy has until recently been considered a treatment whereby genetic diseases could be cured through the introduction of normal genes into the patient to supersede the effects of their malfunctioning counterparts. This understanding of gene therapy resulted in the first human clinical trials in 1990 for the treatment of patients with severe combined immunodeficiency resulting from an adenosine deaminase deficiency.^[3] However, there is another sphere in which gene therapy may prove useful, one that does not specifically require the replacement of a defective gene but rather the production of a desired protein within a specific tissue or disease site. Thus, the gene is delivered by a suitable vector to the defective site resulting in the localized production of a therapeutic protein.

This latter mode of gene therapy is particularly pertinent to the field of tissue engineering. Tissue engineering is principally concerned with the replacement of tissue that has lost function due to injury or disease. In one approach toward this goal, cells are cultured *in vitro* on biodegradable polymeric scaffolds to form neoorgans that are then implanted into the body at the necessary anatomical site. This strategy has great potential and has already resulted in the generation of a clinically approved tissue-engineered human skin equivalent, Apligraf (Organogenesis Inc.; Canton, MA). This bilayering of allogeneic keratinocytes and fibroblasts, respectively on and within a bovine collagen I lattice, has been approved in the United States and been used effectively in the treatment of surgical wounds and venous ulcers.^[4] However, keratinocytes and fibroblasts do not produce an allogeneic T-cell response, enabling this product to sidestep a major hurdle in the use of living cells, the patient's immune response.^[5] Additionally, the need for several weeks of *ex-vivo* culture of the neo organ, the difficulties in preservation of the neo organ, and the sheer number of replacements that are needed (e.g., 570,000 aortocoronary bypass procedures per annum,^[6]

suggest that in the short- to medium-term, other approaches must be explored. The implantation of acellular scaffolds to enhance an organ's ability to spontaneously regenerate is one such tactic. A central component of *in-vivo* tissue regeneration is the supply of proteins, which are involved in driving the process of regeneration (e.g., growth factors) to the required site. It is for this purpose that the GAM was developed.

In this article, the need for gene-activated matrices will be presented, as well as brief overviews of the necessary central tenets: the requirement of a wound bed, gene delivery vectors, and biodegradable polymers. Finally, the advances that have been made in this new field will be summarized.

THE NEED FOR GENE-ACTIVATED MATRICES

Cytokines and growth factors are proteins secreted by cells to induce target cells to carry out specific actions such as migration or proliferation. Many growth factors have been shown to play pivotal roles in the regeneration of tissue, for example, bone morphogenetic proteins (BMP) in bone regrowth, vascular endothelial growth factor (VEGF) in the growth of new blood vessels, and platelet-derived growth factor BB (PDGF-BB) in the healing of skin wounds. Therefore, they are clear targets for exploitation in the pursuit of engineered tissue regeneration.

However, several issues have impeded progress in the use of recombinant growth factors in the repair of damaged tissue. The delicate nature of the complex 3-D structures of the proteins results in very short half-lives in the body as demonstrated by VEGF having a clearance half-life of less than 1 hour after injection.^[7] This phenomenon essentially precludes the systemic delivery of growth factors due to the large doses that would be required. Such doses are naturally associated with the potential for toxicity and have indeed limited the amount of basic fibroblast growth factor (bFGF) and VEGF that could be administered in phase 1 clinical trials due to their potential to induce severe hypotension.^[7] A logical alternative would be the use of a system that allows controlled and sustained delivery of the growth factor. Though this strategy is being actively pursued, it too is hampered by the

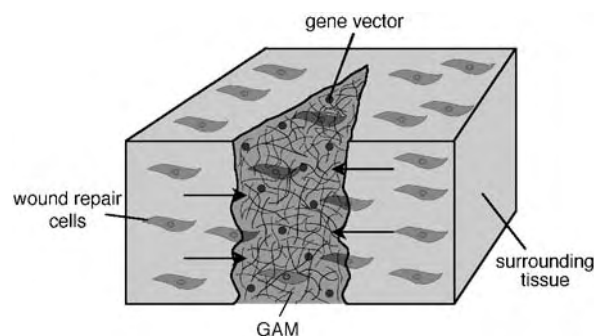


Fig. 1 The central components of a GAM—a biodegradable matrix containing gene-delivery vectors—is positioned within a wound. The matrix retains the gene vectors within the wound while the wound is invaded by wound repair cells. The localized concentration of vector promotes gene delivery to the invading cells and the cells that do take up the gene then generate a sustained and focused production of the encoded bioactive product.

fragility of the polypeptide, as the generally available processes to manufacture controlled-release systems expose the growth factor to solvents and/or sonication.^[2] Therefore, this approach still requires fairly high dosages and again raises the issue of toxicity. Pertinent to all of the above is the not-inconsiderable cost involved in the manufacture of recombinant growth factors.

A clear avenue through which the above problems can possibly be circumvented would be, as indicated above, to utilize the tools of gene therapy and deliver the gene encoding for a growth factor directly to the site requiring tissue regeneration. The cells that take up the gene would then serve as in-vivo bioreactors producing the required growth factor. In an indirect manner termed ex-vivo gene therapy, cells are harvested from the patient, the gene delivered in vitro, and the cells transplanted back into the defective site. This is technically demanding due to the need to harvest cells, and simpler approaches involving the direct delivery of the gene would be desirable. Systemic delivery of a gene vector through injection into the blood system, though potentially suitable for the treatment of metastatic diseases such as cancer, clearly would not be useful for tissue regeneration due to the elevated risks associated with delivery of the gene to cells throughout the body. A more directed approach where the VEGF gene in an aqueous vehicle was directly injected into muscle tissue has been shown to be effective in generating the growth of new blood vessels in an ischemic hindlimb of a rabbit.^[8] Indeed, in a phase 1 clinical trial, the same gene was safely delivered via injection into the ischemic myocardium of five patients with some indication of increased reperfusion.^[9] However, where tissue regeneration into an anatomical space is

required, it would be desirable to retain the gene within the actual defective site, thus preferentially targeting ingrowing cells. This would allow for the production of the growth factor within the defective site, generating chemotactic gradients to attract desired cell types into the site. A gene delivered in an aqueous vehicle into the site would probably be poorly retained and evidence indicating this will be detailed below. The provision of a biodegradable scaffold that retains the gene delivery vector but allows the invasion of cells, thereby bringing these cells in direct contact with the gene vector, is the core of the GAM concept.^[10] Though several biodegradable scaffolds and gene delivery vectors have been employed in GAMs, a fundamental component is the harnessing of the cells involved in the wound healing response.

WOUND HEALING

It is difficult to envisage a situation where tissue regeneration is required that will not involve the formation of a wound. The wound will have either been caused by injury or will almost certainly have resulted from surgical intervention. Even intramuscular injections cause sufficient tissue damage to generate a wound repair response.^[11] The wound healing program is highly conserved and follows a very similar pattern in traumatic injury, pathology-induced injury, and injury caused by surgery.^[10] The sequence of events have been most closely studied for cutaneous wound healing and found to consist of three temporally overlapping phases: inflammation, granulation tissue formation, and tissue remodeling.^[12] The inevitable leakage of blood into the wound site results in the formation of a clot that is composed mainly of fibrin and contains many activated platelets. The degranulation of the activated platelets releases a large variety of growth factors that form chemotactic gradients for the inflammatory response neutrophils and monocytes. Neutrophils cleanse the wound area of foreign particles and are rapidly removed by activated macrophages within the first three–four days of the wound healing response. Fibroblasts start to migrate into the wound after three–four days along with the invasion of blood vessels, resulting in the formation of granulation tissue. After laying down a collagen-rich matrix, fibroblasts begin to convert to a contractile myofibroblast phenotype that is believed to be necessary for wound contraction. During tissue remodeling, the cells within the wound undergo apoptosis, leaving behind a relatively acellular scar.

It is the aforementioned cell types that will come into contact with the gene delivery vector contained in the implanted scaffold.



GENE DELIVERY VECTORS

In brief, gene delivery vectors can be viewed as falling into three classes: naked DNA, complexed DNA, and viral vectors. Naked DNA essentially refers to plasmid DNA expression vectors, which are circular DNA constructs consisting of the DNA sequence for the gene of interest and other necessary sequences that include transcriptional promoters and regions that impart stability to the transcribed mRNA. Plasmid DNA is the simplest vector to produce and probably has the optimal safety profile of presently available vectors, but is also the least effective. This is not surprising as eukaryotic cells are well protected against the uptake of foreign nucleic acids by a number of cellular barriers. These compose the plasma membrane, which is negatively charged (repellent to the large negatively charged plasmid molecules), the sequestration of plasmids away from the nucleus through internalization within endosomes, and the nuclear membrane, which does not allow passive diffusion of particles of a greater molecular weight than 60 kDa.^[13]

Therefore, intensive research has been directed toward overcoming these obstacles in order to increase the efficiency of transfection, the process whereby eukaryotic cells take up foreign DNA.^[13] Plasmid DNA has been complexed with a variety of reagents (lipids, polymers, and peptides). The most widely employed complexing agents are cationic liposomes and condensing reagents such as poly(ethyleneimine) and poly(L-lysine), which respectively form lipoplexes and polyplexes with the DNA. Both facilitate interaction of the negatively charged DNA with the negatively charged plasma membrane. They are also presumed to improve transmembrane transport through diffusion of the lipoplexes and via reduction in the size of the DNA molecule in the polyplexes. The complexing reagents also permit the coupling of cell receptor ligands (e.g., epidermal growth factor) that not only enhance the binding of the complex and its subsequent uptake by endosomes but also allow for the potential of cell-specific targeting. In addition, peptides that disrupt the endosome (e.g., bacterial protein toxins) and nuclear localization signals can be bound to lipoplexes and polyplexes.

GAMs that carry plasmid DNA are therefore a subset of the class of complexed DNA vectors. In part, plasmid DNA was utilized in the initial design of GAMs due to its simplicity of production and good safety profile^[1] but also due to the desire to increase the efficiency of plasmid DNA, as a gene therapy vector was central to the development of GAMs. There is a subtle difference in the rationale behind the GAM approach to improved gene delivery from that of other complexed DNA strategies, thus placing GAMs in a distinct subset. The emphasis is on the retention of the gene within a specific locale to await the arrival of

target cells rather than the active delivery of genes to these target cells. Thus, it has been hypothesized that the GAM not only allows for delivery of the therapeutic gene to the desired site in tissue engineering but also that the elevated local concentration and sustained persistence of the gene would enhance uptake.^[14] In a canine tibia defect model, plasmid DNA complexed within a collagen sponge was retained at the site of implantation for at least six weeks as determined by the polymerase chain reaction (PCR).^[1] Similar results were found for a collagen-based GAM implanted into transections in rat optic nerves.^[15] In both cases, expression of the delivered plasmid DNA was also found to persist for at least six weeks as evidenced by the detection of delivered gene-specific mRNA by reverse transcriptase PCR. The high level of invading cells transfected by a reporter gene, β -galactosidase, in the above canine tibia defect model strongly indicates that there is elevated efficiency of delivery by GAM. The exact mechanisms that resulted in 20–50% of the invasive cells taking up the reporter gene are unknown, but it has been speculated that the high mitotic index of cells involved in the wound healing response could have contributed,^[2] with the breakdown of the nuclear membrane during cell division possibly facilitating the plasmid DNA reaching the nucleus.

Viral vectors have been driven by evolution to become highly efficient at bypassing the previously mentioned cellular barriers toward foreign DNA ingress. It is therefore not surprising that the majority of clinical gene-therapy trials have employed them as the vector of choice. The retrovirus, the herpes simplex virus, the adenovirus, and the adeno-associated virus (AAV) are the most commonly used viral vectors.^[16] All of these have been engineered to render them replication-deficient, permitting the initial infection of the target cell but preventing escape of the virus from the target to other cells.

The retroviruses are single-stranded RNA viruses, the majority of which are based on the Moloney murine leukemia virus. These viruses are able to infect only actively dividing cells, limiting their usefulness in gene therapy. However, due to their low immunogenicity resulting from the complete removal of viral proteins, they are the most widely used vectors in clinical trials. Retroviruses can semirandomly insert themselves into the host genome, which results in greater persistence of expression but also raises the risk of insertional mutations. This can potentially occur when the insertion activates a protooncogene or disables a tumor suppressor gene. More recently, vector systems based on the lentivirus genus (which includes the human immunodeficiency virus) of the retroviruses have been developed. These vectors are able to infect dividing and nondividing cells, but their use is still complicated by their pathogenicity.

The herpes simplex virus is a double-stranded DNA virus that is able to take up to 30 kilo bases (kb) of foreign DNA. This complex virus is also able to infect dividing and nondividing cells, but research has been hampered by the vector's cytotoxicity.

The original adenovirus vector was based on the virus that causes the common cold in humans. The recombinant replication-deficient double-stranded DNA virus is able to infect a wide variety of cells, both dividing and nondividing, with excellent efficiency, and the virus is easily produced to high titres (1×10^{13} viral particles/ml). This form of vector does not result in persistent expression, as it does not integrate into the host genome. The first generation versions have been plagued by cytotoxicity and immunogenicity problems related to the expression of viral proteins that remain in the vector backbone. Thus, much research has been directed towards overcoming this problem, resulting in the gutless vectors from which all adenovirus genes have been removed, leaving only the sequences necessary for replication and packaging of the genome. This has, however, complicated the manufacture of recombinant virus due to the need for a helper virus.

AAVs are single-stranded DNA viruses that are nonpathogenic for humans. In their wild-type form they specifically integrate into chromosome 19, but this specificity has been lost in recombinant viruses. Therefore, though recombinant AAVs still have the potential for persistent expression, they are also capable of insertional mutagenesis. These viruses are able to infect both dividing and nondividing cells and are substantially less immunogenic than the adenovirus (though the generation of neutralizing antibodies may restrict readministration). Until recently, their use has been somewhat curtailed by manufacturing difficulties and their limited ability to carry foreign DNA (<5 kb). The former drawback has been surmounted by the development of helper-free systems and a single-step purification methodology that allows for the generation of titres of 1×10^{12} viral particles/ml that are free of replicative virus. Recent exploitation of the need of AAV to form concatemers during their infection of a cell has effectively doubled their payload of foreign DNA.^[16]

So far, adenoviruses are the only viral vectors that have been employed in GAMs. This is probably due in part to their ease of production and high efficiency. In addition, the transient expression of the therapeutic protein that results from adenoviral infection might be sufficient and indeed desirable for the regeneration of tissue.^[17] Conceptually, the GAM would serve a similar role for viral vectors as for naked plasmid DNA, that is, the delivery of the vector to a desired anatomical site with the consequent generation of bioeffectors in the invading wound repair tissue. However, another consideration more pertinent to viral vectors is

the retention of the virus at the site of delivery, thereby increasing the safety profile by reducing infections at distal sites.^[18] In-vitro release studies found 20% of loaded adenoviruses were released from 0.15% collagen GAMs over 48 hours and that this release could be reduced to 10% by increasing the collagen content to 2.6%.^[17,19] When adenoviruses were reacted with biotinylated antibodies specific to the hexon protein of the virus and subsequently polymerized into 0.21% collagen gels containing biotin-binding avidin, the release was reduced to 5% at 48 hours.^[18] However, in a direct comparison between adenoviruses injected either directly or as an antibody-complexed GAM into porcine right ventricles, the biodistribution was found to be identical. For each delivery method, adenovirus was found in the lungs of seven out of eight animals. No evidence for distribution was determined in the liver, kidneys, or left ventricles. The GAM preparation did show significantly more efficient delivery of a reporter gene, green fluorescent protein, to the target tissue of the right ventricle than the aqueous solution. Though the biodistribution results were not significantly different for the antibody-complexed GAM compared to aqueous delivery, this approach holds significant promise, as relatively simple modifications such as a higher-specificity antibody or covalent binding of the avidin to the GAM polymer should greatly improve in-vivo retention.^[18] This aspect in the development of GAMs will become increasingly critical with the employment of viral vectors with the potential for persistent gene expression and hence an increased need for accurate delivery to their target cells. With reference to these persistent expression-type viral vectors, the very small AAV (25-nm diameter, compared to the 90-nm diameter adenovirus) may prove particularly difficult to retain within a GAM.

BIODEGRADABLE SCAFFOLDS

Several biodegradable scaffolds such as collagen, gelatin, carboxymethylcellulose, and poly(lactide-co-glycolide) have been utilized for GAMs. Collagen, a triple helix polypeptide that is the major component of mammalian connective tissue, can be conveniently gelled by elevating the pH to neutral and the temperature to 37°C. In addition, it has been clinically approved for use in a tissue-engineering application (bovine collagen: see the discussion of Appligraf in the Introduction). These considerations have probably resulted in this biopolymer being most commonly employed in GAMs, and it has been used in lyophile sponge, paste, and gel formulations.^[2] Fibrin, another biologically derived polymer, is a potentially interesting candidate that has not yet been reported on in the GAM literature. This polymer, which is the major



constituent of the clot present in the early stages of wound healing, has been found to allow substantially more rapid ingrowth of wound repair cells than collagen,^[20] and human-derived fibrin has been clinically approved.

Biologically derived biopolymers are compromised by their potential to transmit copurified pathogens, batch variation, and their limited ability to be tailored to fit specific requirements. For these reasons, biodegradable synthetic polymers are being intensively investigated.^[21] Focuses of research into synthetic polymers include temporal control of degradation, mechanical properties, and derivatization of polymers with specific signals such as adhesive peptides and growth factors. A promising approach for future use in GAMs is the development of gels that are polymerized with oligopeptides that are substrates for degradative enzymes (e.g., plasmin, collagenase), thus rendering the synthetic matrix degradable by cell-based enzymatic activity.^[22] The synthetic linear aliphatic polyester poly(lactide-co-glycolide) has been used in a GAM-based strategy to deliver a plasmid encoding for platelet-derived growth factor B (PDGF-B) in a rat subcutaneous model.^[14] This approach should perhaps be more strictly classified as a sustained release preparation, as no cells were observed actually to invade the poly(lactide-co-glycolide) sponge and take up a plasmid encoding for the reporter gene β -galactosidase contained within the gel. However, high levels of β -galactosidase expression were observed in the wound repair tissue peripheral to the implant and poly(lactide-co-glycolide)-complexed PDGF-B plasmid significantly enhanced the formation of granulation tissue in the subcutaneous pocket relative to that detected for directly injected PDGF-B plasmid.

TISSUE REGENERATION

GAMs have been used in facilitating the regeneration of several different tissues types, as detailed in the following sections.

Hard Tissue

Bone has the ability to regenerate itself but only if the defect is under a critical size and complex fractures with large voids are therefore at risk of nonunion. Regeneration of bone in critical defects introduced into rat femurs was the first study to assess the viability of GAMs in guiding tissue growth.^[23] Fang and colleagues implanted lyophilic collagen sponges into the defects that delivered 0.5–1.0 mg plasmid DNA encoding for either a secreted form of parathyroid hormone (hPTH 1-34), bone morphogenetic protein-4

(BMP-4), or a combination of these plasmids. The single-plasmid GAMs resulted in bone growth and a functional union across a 5-mm defect within nine weeks where fibrous scarring would have normally occurred. Bridging of the gap was observed at four weeks after delivery of the combination of hPTH 1-34 and BMP-4, these two proteins being known to interact in a synergistic manner. In a follow-up study in a larger animal model (canine tibia critical defect), the exact mechanics of therapeutic protein delivery were examined in greater detail.^[1] As described above, plasmid DNA and RNA transcription persisted for six weeks and 20–50% of the invading cells were transfected. These cells were identified as fibroblasts on the basis of their morphology. At two weeks there was a local content of approximately 2.9 pg hPTH 1-34 in gaps that had received 1.0 mg of hPTH 1-34-encoding plasmid, pMAT-1. Bone growth in a 2.0-cm critical defect in the tibia began when 40.0 mg of pMAT-1 DNA was delivered. This amount allowed for only 25% healing within four weeks and it required 100.0 mg pMAT-1 DNA to generate a 60–75% filling by six weeks. Complete gap filling was not observed in this size defect. It is intriguing that these very high dosages of plasmid were needed to generate a therapeutic effect, as the lower dosage (1 mg) resulted in transfection of a significant percentage of the invading cells. With the maximal loading of plasmid, 1.6-cm gaps were completely bridged in 18 weeks and 1.0-cm gaps in eight weeks. Most strikingly, no bone growth was observed for control sponge implants for either of these two different size defects.

In a preliminary study into lumbar interbody fusions, a lyophilic collagen sponge delivering hPTH 1-34 plasmid DNA was also found to stimulate bone growth above that of control sponges.^[24] These results further substantiate the conclusion from the two aforementioned studies that the production of bone growth factors by wound repair cells present within a defect can induce substantial growth of new bone.

Soft Tissue

Skin

Cutaneous wounds in adults essentially heal as reepithelialized scar tissue due to the respectively potent and feeble regenerative capabilities of the epidermis and underlying dermis.^[25] Venous stasis, pressure, and diabetes mellitus can result in chronic ulcers where even the migration of the epidermis is derailed. The large numbers of these chronic wounds (2 million annually (U.S.A.)^[4]) and the resultant high financial costs have driven research in this area. Growth factors applied as proteins have met with limited success and this

has been ascribed in part to the protease-rich environment of the wound bed.^[26] Therefore, the potential of the GAM for sustained, localized production of a therapeutic protein has made it an interesting candidate in this instance.

Approval of PDGF-BB protein by the U.S. Food and Drug Administration as a treatment for improving the healing of diabetic foot ulcers naturally resulted in this growth factor being the initial candidate for GAM delivery. The effect of the delivery of the gene as a plasmid from lyophilic collagen sponges on wound healing was studied in an ischemic rabbit dermal ulcer model.^[19,27] This model mimics an important component of the human diabetic ulcer, the reduced supply of oxygen and other nutrients resulting from peripheral vascular disease. When 1.0-mg PDGF-B plasmid DNA in lyophilized collagen was placed within 6-mm wounds in ischemic rabbit ears, a five-fold increase in granulation tissue formation and a two-fold increase in epithelialization relative to collagen alone were observed after 10 days. Similar results were found for the delivery of PDGF-A plasmid DNA, a closely related growth factor that forms both homodimers and PDGF-B heterodimers.

GAM-based delivery of PDGF-B gene has also been considered in the context of a highly efficient adenoviral vector, replication-deficient adenovirus serotype 5.^[17,19] Initial investigations into the efficacy of this system in aiding wound healing were carried out in a rodent subcutaneous polyvinyl alcohol (PVA) sponge implant model. Adenoviral vectors in a 0.15% collagen gel were directly injected into PVA sponges four days after their subcutaneous implantation, and the sponges were explanted up to 12 days after viral/collagen injection. PDGF-BB expression, as determined by immunohistochemistry, continued for up to eight days. Six days after treatment, granulation tissue content was increased three- to four-fold by the administration of 2×10^{10} PDGF-B encoding adenoviruses relative to collagen gel alone. Interestingly, a single dose of 100- μ g PDGF-BB protein resulted in significantly less granulation tissue formation even though explanted sponges that had received the above loading of adenovirus contained only 54-ng PDGF-BB protein upon explant. Three sequential doses (alternate days) of 10- μ g protein were needed to generate a similar amount of tissue and this suggests that it is indeed the persistent expression of PDGF-B that is responsible for the effectiveness of the GAM. The efficiency of the adenovirus was demonstrated, as the majority of cells (leucocytes and fibroblasts) at one day after delivery contained virus and expressed PDGF-B mRNA. The PVA sponge model was further utilized to assess the effect of conjugating bFGF to the adenoviral vector.^[26] The targeting of adenoviruses to fibroblast-growth-factor receptor-expressing cells had been

previously shown to increase the potency of the viral vector. Similarly, in a combination of GAM technology and viral targeting, the injection of a 0.15% collagen solution containing 2×10^9 bFGF-conjugated adenoviruses encoding for PDGF-B generated granulation tissue formation equivalent to that obtained with 2×10^{10} nonconjugated adenoviruses. The enhanced tissue formation was not due to any direct actions of bFGF.

In application of PDGF-B gene-encoding adenoviruses in collagen gels or lyophilic sponges to dermal wounds in normoxic pigs and ischemic rabbits, similar improvements in wound healing were observed to those reported in the aforementioned study.^[17,19] Significantly, when an equivalent dosage (2×10^{10} viral particles) was delivered in an aqueous buffer to a wound in the rabbit model, no increase in tissue formation above that of controls was observed. The seepage of the vector from the delivery site was evidenced by fibroblast proliferation and PDGF-BB production in the underlying dermis.

Arterioles

Ischemia caused by vascular disturbances is a devastatingly prevalent condition and can result in, for example, lower limb loss, coronary heart disease, and peripheral neuropathy. It is little wonder, therefore, that growth factor-driven therapeutic neovascularization has received such great attention.

The potency of GAMs in stimulating arteriolization was investigated in a rat model where 3×5 -mm plugs were removed from the rectus femoris.^[17] The defects were filled with collagen-gelatin (1% each) mixtures containing either adenovirus or plasmid DNA encoding for bFGF, FGF6, and PDGF-B. After excision at 21 days, 5×10^9 bFGF-encoding adenoviruses resulted in a six-fold increase in arterioles growing into the wound space relative to matrix alone. Similar results were observed for 1.0-mg loadings of bFGF and FGF6 plasmid DNA, but again no response was observed for 60- μ g bFGF protein. PDGF-B plasmid DNA stimulated the growth of capillaries but did not result in any significant arteriolization, a somewhat puzzling result, as PDGF-BB is a potent chemoattractant for smooth muscle cells.^[28] As evidence of the pleiotropic nature of the FGF proteins, increased muscle regeneration was apparent with an approximately twenty-fold increase in density of a marker for myotube proliferation observed for GAMs carrying either bFGF or FGF6 plasmid DNA.

In a subsequent study, the utility of the aforementioned findings for bFGF in the treatment of ischemic myocardium was determined.^[29] This study was particularly pertinent, as the stimulation of



arteriogenesis rather than capillary growth is considered the preferred form of neovascularization for restoring myocardial perfusion.^[7] Ischemia in the myocardium of pigs was produced by surgically restricting the flow in the left circumflex artery. A GAM of collagen-gelatin (1% each) with an adenovirus encoding for bFGF was injected into the ischemic area (20 injections of 5×10^{10} particles each, one injection/cm²), and the contractility of the region was determined at six weeks by echocardiography analysis and magnetic resonance imaging. Contractility was significantly improved compared to that found in mock-injected ischemic regions and correlated with a five-fold increase in arteriolization.

Nerve

Central nervous system tissues almost always show limited healing and a major cause is the inability of adult neurons to proliferate. Berry and colleagues investigated the efficacy of GAMs in promoting nerve healing in surgically severed rat optic nerves.^[15] The placement of a collagen paste containing a reporter gene encoded on plasmid DNA between the proximal and distal ends resulted in retrograde transport of transfected DNA to the nerve cell body in the retina. A GAM containing a cocktail of plasmids encoding for several neurotrophic growth factors significantly increased the number of surviving active retinal ganglion cells to 2.25% relative to the 0.37% for a GAM containing a reporter gene-encoding plasmid.

CONCLUSION

GAMs have been shown to be capable of promoting growth in a wide variety of tissues. In several instances, the presentation of growth factor genes embedded in a matrix was demonstrably more efficient than the delivery of the gene in an aqueous vehicle or the growth factor protein itself. Therefore, the central premise that invading wound healing cells would be able to serve as bioreactors for sustained production of a therapeutic protein is a sound one.^[30]

The future of the GAM will of course incorporate an increasing range of growth factors and combinations thereof. The delivery of other entities such as antisense DNA and intracellular signaling proteins will become increasingly more common. Recently, an antisense plasmid against thrombospondin 2 was employed in a collagen coating for biomaterials resulting in alterations in the foreign body reaction.^[31] The utilization of improvements in transfection such as endosomal disruption as they become more established would be appropriate, particularly in the nonviral

GAMs. Greater persistence of gene expression may be desirable in certain instances such as particularly large defects in bone. This would be best achieved by the use of vectors with a long-term effect such as AAV or retroviruses. The direct targeting of longer-lived wound repair cells such as fibroblasts through the use of receptor-mediated targeting or cell-specific promoters may also prove useful. Long-term expression would most likely necessitate the combination of a regulatable gene expression system with the GAM.^[2] The use of these systems that employ oral uptake of drugs such as tetracycline or rapamycin to control chimeric transcriptional activators^[32] would also allow for temporal control so that pulsatile delivery of a therapeutic protein could be achieved if required. When the physiological sequences in which different growth factors are supplied during tissue regeneration become more clearly understood, it could be envisaged that the required order of delivery of two or more growth factors might be matched by regulating their genes with distinct drugs.

The challenging arena of human clinical trials still awaits GAMs as many other gene therapies that have shown great promise in animal models have failed at this critical stage. The biodistribution and biotransformation of the genes and their products will of necessity require greater attention. Although these are formidable obstacles, the significant progress made thus far holds great promise for the GAM.

ARTICLES OF FURTHER INTEREST

Biologic and Synthetic Apatites; Burn Dressing; Gene Therapy; Poly(Glycolic Acid); Poly(lactic acid)s; Tissue Engineering of Bone; Tissue Engineering of Skin

REFERENCES

1. Bonadio, J.; Smiley, E.; Patil, P.; Goldstein, S. Localized, direct plasmid gene delivery in vivo: Prolonged therapy results in reproducible tissue regeneration. *Nat. Med.* **1999**, *5* (7), 753–759.
2. Bonadio, J. Tissue engineering via local gene delivery: Update and future prospects for enhancing the technology. *Adv. Drug Deliv. Rev.* **2000**, *44* (2–3), 185–194.
3. Anderson, W. Human gene therapy. *Science* **1992**, *256* (5058), 808–813.
4. Clark, R.; Singer, A. Wound Repair: Basic Biology to Tissue Engineering. In *Principles of Tissue Engineering*, 2nd Ed.; Lanza, R., Langer, R., Vacanti, J., Eds.; Academic Press: San Diego, 2000; 857–878.
5. Nickoloff, B.; Turka, L. Immunological functions of non-professional antigen-presenting cells: New insights from studies of T-cell interactions with keratinocytes. *Immunol. Today* **1994**, *15* (10), 464–469.

6. National Hospital Discharge Survey. In *National Center for Health Statistics, Vital Health Statistics*; 13; 113–151.
7. Simons, M.; Bonow, R.; Chronos, N.; Cohen, D.; Giordano, F.; Hammond, H.; Laham, R.; Li, W.; Pike, M.; Sellke, F.; Stegmann, T.; Udelson, J.; Rosengart, T. Clinical trials in coronary angiogenesis: Issues, problems, consensus: An expert panel summary. *Circulation* **2000**, *102* (11), E73–E86.
8. Tsurumi, Y.; Takeshita, S.; Chen, D.; Kearney, M.; Rossow, S.; Passeri, J.; Horowitz, J.; Symes, J.; Isner, J. Direct intramuscular gene transfer of naked DNA encoding vascular endothelial growth factor augments collateral development and tissue perfusion. *Circulation* **1996**, *94* (12), 3281–3290.
9. Losordo, D.; Vale, P.; Symes, J.; Dunnington, C.; Esakof, D.; Maysky, M.; Ashare, A.; Lathi, K.; Isner, J. Gene therapy for myocardial angiogenesis: Initial clinical results with direct myocardial injection of phVEGF165 as sole therapy for myocardial ischemia. *Circulation* **1998**, *98* (25), 2800–2804.
10. Bonadio, J. Tissue engineering via local gene delivery. *J. Mol. Med.* **2000**, *78* (6), 303–311.
11. Doukas, J.; Blease, K.; Craig, D.; Ma, C.; Chandler, L.; Sosnowski, B.; Pierce, G. Delivery of FGF genes to wound repair cells enhances arteriogenesis and myogenesis in skeletal muscle. *Molec. Ther.* **2002**, *5* (5 Pt. 1), 517–527.
12. Singer, A.; Clark, R. Cutaneous wound healing. *N. Engl. J. Med.* **1999**, *341* (10), 738–746.
13. Uherek, C.; Wels, W. DNA-carrier proteins for targeted gene delivery. *Adv. Drug Deliv. Rev.* **2000**, *44* (2–3), 153–166.
14. Shea, L.; Smiley, E.; Bonadio, J.; Mooney, D. DNA delivery from polymer matrices for tissue engineering. *Nat. Biotechnol.* **1999**, *17* (6), 551–554.
15. Berry, M.; Gonzalez, A.; Clarke, W.; Greenlees, L.; Barrett, L.; Tsang, W.; Seymour, L.; Bonadio, J.; Logan, A.; Baird, A. Sustained effects of gene-activated matrices after CNS injury. *Mol. Cell. Neurosci.* **2001**, *17* (4), 706–716.
16. Kay, M.; Glorioso, J.; Naldini, L. Viral vectors for gene therapy: The art of turning infectious agents into vehicles of therapeutics. *Nat. Med.* **2001**, *7* (1), 33–40.
17. Doukas, J.; Chandler, L.; Gonzalez, A.; Gu, D.; Hoganson, D.; Ma, C.; Nguyen, T.; Printz, M.; Nesbit, M.; Herlyn, M.; Crombleholme, T.; Aukerman, S.; Sosnowski, B.; Pierce, G. Matrix immobilization enhances the tissue repair activity of growth factor gene therapy vectors. *Hum. Gene Ther.* **2001**, *12* (7), 783–798.
18. Levy, R.; Song, C.; Tallapragada, S.; DeFelice, S.; Travis Hinson, J.; Vyavahare, N.; Connolly, J.; Ryan, K.; Li, Q. Localized adenovirus gene delivery using antiviral IgG complexation. *Gene Ther.* **2001**, *8*, 659–667.
19. Chandler, L.; Doukas, J.; Gonzalez, A.; Hoganson, D.; Gu, D.; Ma, C.; Nesbit, M.; Crombleholme, T.; Herlyn, M.; Sosnowski, B.; Pierce, G. FGF2-Targeted adenovirus encoding platelet-derived growth factor-B enhances de novo tissue formation. *Molec. Ther.* **2000**, *2* (2), 153–160.
20. Dvorak, H.; Harvey, V.; Estrella, P.; Brown, L.; McDonagh, J.; Dvorak, A. Fibrin containing gels induce angiogenesis. Implications for tumor stroma generation and wound healing. *Lab. Invest.* **1987**, *57* (6), 673–686.
21. Pachence, J.; Kohn, J. Biodegradable Polymers. In *Principles of Tissue Engineering*, 2nd Ed.; Lanza, R., Langer, R., Vacanti, J., Eds.; Academic Press: San Diego, 2000; 263–277.
22. Hubbell, J. Bioactive biomaterials. *Curr. Opin. Biotechnol.* **1999**, *10* (2), 123–129.
23. Fang, J.; Zhu, Y.; Smiley, E.; Bonadio, J.; Rouleau, J.; Goldstein, S.; McCauley, L.; Davidson, B.; Roessler, B. Stimulation of new bone formation by direct transfer of osteogenic plasmid genes. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93* (12), 5753–5758.
24. Patil, P.V.; Graziano, G.P.; Bonadio, J.; Goldstein, S.A. Interbody fusion augmentation using localized gene delivery. *Trans. Orthop. Res. Soc.* **2000**, *25*, 360.
25. Parentau, N.; Hardin-Young, J.; Ross, R. Skin. In *Principles of Tissue Engineering*, 2nd Ed.; Lanza, R., Langer, R., Vacanti, J., Eds.; Academic Press: San Diego, 2000; 879–890.
26. Chandler, L.; Gu, D.; Ma, C.; Gonzalez, A.; Doukas, J.; Nguyen, T.; Pierce, G.; Phillips, M. Matrix-enabled gene transfer for cutaneous wound repair. *Wound Repair Regen.* **2000**, *8* (6), 473–479.
27. Tyrone, J.; Mogford, J.; Chandler, L.; Ma, C.; Xia, Y.; Pierce, G.; Mustoe, T. Collagen-embedded platelet-derived growth factor DNA plasmid promotes wound healing in a dermal ulcer model. *J. Surg. Res.* **2000**, *93* (2), 230–236.
28. Hellstrom, M.; Kal, N.M.; Lindahl, P.; Abramsson, A.; Betsholtz, C. Role of PDGF-B and PDGFR-beta in recruitment of vascular smooth muscle cells and pericytes during embryonic blood vessel formation in the mouse. *Development* **1999**, *126* (14), 3047–3055.
29. Horvath, K.; Doukas, J.; Lu, C.-Y.; Belkind, N.; Greene, R.; Pierce, G.; Fullerton, D. Myocardial functional recovery after fibroblast growth factor 2 gene therapy as assessed by echocardiography and magnetic resonance imaging. *Ann. Thorac. Surg.* **2002**, *74*, 481–487.
30. Bonadio, J. Local gene delivery for tissue regeneration. *E-biomed J. Regen. Med.* **2000**, *1*, 25–29.
31. Kyriakides, T.; Hartzel, T.; Huynh, G.; Bornstein, P. Regulation of angiogenesis and matrix remodeling by localized, matrix-mediated antisense gene delivery. *Molec. Ther.* **2001**, *3* (6), 842–849.
32. Agha-Mohammadi, S.; Lotze, M. Regulatable systems: Applications in gene therapy and replicating viruses. *J. Clin. Invest.* **2000**, *105* (9), 1177–1183.



Gene Therapy

Viraj P. Mane

William McCormack

Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas, U.S.A.

Brendan Lee

Department of Molecular and Human Genetics, Baylor College of Medicine, Howard Hughes Medical Institute, Baylor College of Medicine, and Houston, Texas, U.S.A.

INTRODUCTION

The term “therapy” has been described as treatment of illness or disability. However, gene therapy promises far more than traditional medical therapy. As opposed to treating a condition, gene therapy offers to cure it by correcting, at the genetic level, a congenital deficiency or alteration of some critical function. The first gene replacement therapy clinical trial occurred in 1990 involving Ashanthi DeSilva, who suffered from ADA-SCID, or adenosine deaminase severe combined immunodeficiency. Since the 1990 ADA-SCID trial, patients have enrolled in 918 gene therapy protocols, with the vast majority taking place in the United States.

Gene therapy has a straightforward goal: to introduce genetic information into a host to replace or correct a biological function. Important components of gene therapy include the nature of the vector, the nature of the therapeutic gene, the target cell or tissue (host), and the pathophysiology of the disease process. These factors interact to achieve a biologic effect that can be characterized by a time course (transient v permanent), by a measure of efficacy often directly correlated with efficiency of transduction, and by toxicity. An important feature of the pathophysiology is whether the disease process is functionally a cell-autonomous or non-cell-autonomous one. If correction in one cell can functionally correct a phenotype by complementing the function of neighboring cells, then it may be viewed as non-cell-autonomous process. However, if a phenotypic deficiency cannot be cross-corrected in this way, then it may be viewed as cell-autonomous. This is not always a clear distinction. A simple example of a non-cell-autonomous pathophysiology is a deficiency of a secreted protein such as in hemophilia. In this case, correction of a small number of target cells to overexpress the missing plasma protein may be sufficient to correct the clinical phenotype.

THE DISEASE

Gene therapy can be broadly divided into therapy for three different disease domains. They include gene

replacement therapy of genetic deficiencies, gene therapy for cancer, and gene therapy for infectious diseases. Gene therapy for infectious disease has focused on genetic vaccination. Like a traditional vaccine, genetic vaccination seeks to introduce or stimulate the host immune response by expressing a pathogenic epitope. Genetic vaccination generally initiates a specific host immune response against a single antigenic peptide of a pathogen, obviating exposure to the pathogen in its entirety. The de novo expression of antigenic peptide in a target cell may influence the type of host immune response (Th1 v Th2) and determine whether a cellular or an antibody response is mounted. Th1 (T-helper cell type 1) is characterized by a proinflammatory predominantly cellular response. Th2 responses are partially antagonistic to Th1 and are predominantly humoral (antibody-mediated). This approach may also be applied to the domain of cancer. In this case, cancers often express unique tumor-associated antigens (TAA) and the host immune response could be primed to respond to such antigens by genetic vaccination. Cancers often arise from loss of function of tumor suppressor genes, and gene replacement therapy can target replacement of such a gene in an effort to restore growth control.

Another approach draws upon pharmacological manipulations to convert a systemically delivered nontoxic prodrug into its toxic counterpart by prodrug-converting enzymes that are delivered to tumor cells by gene therapy. The transduced cancer cells expressing the enzyme will subsequently convert the prodrug into a lethal compound in a cell-specific manner and thus achieve a cancer-specific suicide gene therapy. Another approach of gene therapy relevant for cancer as well as infectious disease is immunotherapy, which resembles genetic vaccination because it also primes the immune system. In this case, the goal is to introduce genes that encode cytokines or signaling molecules that facilitate autocrine and/or paracrine signaling. Numerous cytokines have differential properties associated with cell maturation, cell proliferation and apoptosis, transcriptional repression or

enhancement, recruitment of inflammatory cells, etc.; thus, the immunomodulatory cytokine gene(s) is chosen in a context-dependent fashion.

In contrast to cancer and infectious disease gene therapy, where the patient immune system is an ally, gene replacement therapy for genetic disease must often overcome the host immune system to achieve long-term correction without toxicity. In this situation, the goal is to replace a missing protein in a sufficient number of target cells, at the same time avoiding immune clearance of the vector or the transgene. This contrasts with cancer and infectious disease therapy, in which the goal is to stimulate clearance of a pathogenic target by directly transducing either the pathogenic cell or a host cell to facilitate the clearance.

THE VECTOR

Vectors, or delivery systems, are an integral part of gene therapy in that the vector is the vehicle that achieves gene transfer into target cells or tissues. Depending on the disease, the vector may be delivered systemically or locally to tissues. Depending on the disease pathophysiology, different vectors are chosen to satisfy specific requirements relating to: 1) vector tropism for target tissue; 2) cloning capacity; 3) persistence in target cells (episomal v integrating vectors); 4) production yields; 5) immunogenicity/toxicity; and 6) cost (Table 1). The disease pathophysiology will dictate the qualities required for a vector system. This implies that there is no general vector system that will satisfy all disease processes. Instead, the features of each vector system will be suited for a particular group of conditions. Vectors can be broadly classified as nonviral v viral vectors. In the latter, the natural life cycle of the virus is co-opted and engineered to generate replication deficient versions that will express a therapeutic transgene.

Nonviral Vectors

Naked DNA

One of the simplest methods of gene transfer is the administration of a plasmid containing the gene of interest. Purified plasmid can be introduced *in vivo* by injection into tissue, or by delivery into a vein (which often favors hepatocyte transduction).^[1] Kou-Gi et al.^[2] found that injection of plasmids encoding human VEGF (vascular endothelial growth factor) or HIF-1 (hypoxia-inducible factor 1) into rat heart resulted in detectable tissue-specific expression of the transgene for up to 14 days. This study, like many others, relied upon commercially available plasmid

constructs and plasmid purification kits. Thus, production is both simple and affordable. The transduction efficiency can be enhanced by precipitation of the DNA with calcium phosphate before delivery; this process has been widely found to increase entry of DNA and RNA into cells.^[3] With regard to genetic deficiencies, gene transfer holds a distinct advantage over recombinant protein administration, because the efficacious yield of recombinant protein may be artificially reduced *in vivo* due to intra- and extracellular enzymes and proteases.^[4] The Wolff and Kay labs^[5] have also validated, in mice, the concept of “hydrodynamic” tail vein injections toward enhanced naked DNA transfection (especially of the liver). These large-volume, high-speed, high-pressure injections of plasmid have been found to dramatically enhance hepatocyte transduction (maximum of 40%), with consistent but transient acute toxicities. The proposed mechanism involves the accumulation of fluid from the overloaded heart back to the liver.

Although one study found high-level expression of rat erythropoietin for up to 12 weeks in the liver following rat tail-vein injection of an erythropoietin-expressing plasmid,^[1] the majority of naked DNA gene therapy applications are complicated by low-level transduction as well as transient expression.^[6] A possible reason may be that naked DNA must contend with extra- and intracellular nucleases in the recipient tissue.^[7] Successful transfection is further complicated by the necessity of several concurrent events: nucleic acid entry into the cell, followed by unimpeded transport into the nucleus, as well as recruitment of transcription machinery to the transgene. Even if these conditions are met, the percentage of transduced cells is halved following each round of cell division, unless the DNA has integrated into the host genome. Furthermore, transfection with plasmids encoding noncytotoxic transgenes was found to induce Fas-mediated apoptotic cell death in a murine cell line.^[8] In this Fas study, apoptosis of transfected cells may logically account for short-term transgene expression. Although transfection efficiency is low, naked DNA transfer may still be adequate for genetic vaccination, in which transient expression may generate sufficient priming of the immune system.^[6,7] A major advantage of naked DNA delivery is absence of a robust immune response, although unmethylated CpG sequences (cytosine-guanine dinucleotides that are a hallmark of foreign bacterial DNA) can sometimes stimulate a host response.

Gene gun delivery

A variation on naked DNA transfection is ballistic or “gene gun” delivery. In this system, the DNA is

Table 1 Characteristics of viral vectors

Recombinant viral vector	Genome	Duration of expression	Advantages	Disadvantages	Applications
Adeno-associated virus (AAV)	Single-stranded DNA (+ and - polarity)	Persistent	Low immunogenicity; integrates into genome; different serotypes offer different cell tropisms	Low production yields; lower transduction than efficiency than adenovirus; smaller packaging capacity; potential integration risk	Infects actively dividing and quiescent cells, epithelial cells, and hepatocytes
Retrovirus	Two (+) sense RNA molecules	Persistent	Integrates into genome; low immunogenicity	Nonspecific integration; transduces only actively dividing cells; production yields low; lower transgene capacity	Hematopoietic cell transduction; cancer gene therapy
Lentivirus	Two (+) sense RNA molecules	Persistent	Transduces both dividing and nondividing cells	Potentially higher risks due to HIV genome; nonspecific integration	Immunotherapy of T cells; genetic vaccination against HIV; CNS transduction
Herpesvirus	Double-stranded DNA	Persistent	Tropism for neurons; large transgene capacity; infection may be asymptomatic; persists episomally in neuronal nuclei	Substantial safety risks	Neurodegenerative disease (Parkinson's)
Adenovirus	Double-stranded DNA	Transient; can be longer term	Large transgene capacity; high production yields; multiple cell-type transduction, especially high-level hepatocyte transduction; non-integrating	Immunogenic; acute toxicity; short-term expression	Transduction of hepatocyte, muscle, endothelium, etc.

precipitated onto chemical particles (usually gold), which are then loaded onto a pressurized gun. The gun propels the particles into tissue at high speed via release of pressurized gas. Although gene guns have shown success at introducing DNA and chemicals into cells, it has the drawback of limited penetration, i.e., the particles do not embed into deeper layers of tissue. Hence, regarding *in vivo* applications, the gene gun currently is somewhat limited to transduction of superficial tissue layers. Transgene expression may be additionally restricted to cells with gold particles lodged specifically in the nucleus as opposed to cytoplasm.^[9]

Electroporation

Electroporation is historic practice of using pulses of electricity to temporarily disrupt cell membranes. The application of an electric field generates a voltage potential across the cell membrane. Normally, the membrane acts as an electrical insulator between the cytoplasm and extracellular space. However, when the insulating capacity of the membrane is exceeded, it is reversibly permeabilized. During this period, negatively charged nucleotides can travel across the membrane. One mouse muscle study found up to a 70-fold increase in reporter luciferase expression by plasmid injection followed by electroporation, as compared to plasmid injection alone.^[10] However, this same study found substantial development of muscle injury resulting from the combination of plasmid injection plus electroporation, as well as a characteristic necrosis of fiber cells and infiltration of muscle by inflammatory cells. Whereas many naked DNA-electroporation applications target muscle, a recent study found that injection of naked DNA into mouse melanoma followed by electroporation led to high reporter gene expression coupled to drastic tumor regression. The antitumor activity of naked DNA (which in this study encoded a reporter, but no therapeutic transgene) was rationalized by citing a possible innate immune response against cells harboring bacterially-derived unmethylated plasmid DNA (especially CpG sequences, mentioned previously).^[8,11] Such results suggest that the naked DNA/electroporation treatment combination might be more relevant to cancer tissues where injury and stimulation of host immune response are less problematic, because enhanced transgene expression coupled to destruction of transduced cells (“suicide” gene therapy) may be desirable in a tumor application. It should be kept in mind that electroporation requires an unbroken membrane to hold an electric potential.^[12] Therefore, cell membranes and organelle membranes are amenable to electroporation, but nuclear pores prevent a potential from forming

across the membrane of the nucleus. Thus, electroporation can enhance entry of exogenous DNA via the cell’s outer membrane, but offers no enhancement of actual DNA import into the nucleus.

Liposomes

Liposomes have been shown to enhance nucleic acid incorporation into cells and help mask the nucleic acids from extracellular nucleases. Liposomes are rounded structures composed of lipid bilayers, whose hydrophilic interior contains the nucleic acid of interest. The DNA is integrated into enclosed liposomes via sonication of the lipids.^[13] The bilayer structure allows fusion with the target cell’s outer membrane, followed by subsequent entry into endosomes. Endosomal uptake, however, leads to lysosome-mediated degradation of the endosomal contents; lysosome-disrupting peptides have been shown to reduce this degradation.^[14] The efficacy of liposome-mediated gene transfer has been noted in several studies. In a growth-enhancement application, keratinocyte growth factor (KGF)-encoding liposomes were found to stimulate epidermal growth and neo-angiogenesis *in vivo*, in rat models of thermal skin injury.^[4] In a growth-suppressive application, alternatively, liposomes containing a plasmid encoding human interferon-beta were reported to have initiated tumor regression in human esophageal cancer cell lines.^[15] This growth suppression correlated to that seen upon addition of exogenous interferon-beta cytokine. Enhanced cell specificity of liposomes, and even for naked DNA, can be achieved by attaching targeting peptide ligands. This approach was utilized to generate a liposome whose membrane incorporated integrin-targeting peptides.^[16]

Synthetics

One approach that has been aggressively pursued is the application of synthetic molecular networks for the containment and controlled release of naked DNA. These occur in various forms including capsules, beads, magnetized particles, and gels. One murine intramuscular transfection experiment found enhanced transgene expression when injection of plasmid DNA, entrapped in a proprietary polymeric network, was coupled to electroporation.^[17] Whereas some synthetic vectors have demonstrated cell-type specificity,^[18] others are designed for controlled release of DNA from a matrix.^[19]

In general, nonviral methods are characterized by relatively low transduction efficiencies, but minimal host immune response and toxicity from the delivery vehicle. They have been shown to be especially



efficacious in genetic vaccination strategies for cancer and infectious disease gene therapy, and for gene replacement therapy of non-cell-autonomous processes to generate highly bioactive molecules that are needed in small concentrations.

Viral Vectors

Viral vectors take advantage of millions of years of evolution that have optimized wild-type viruses' ability to infect host cells, co-opt the cellular machinery, and produce new virions. Because of this, viral gene therapy vectors are often superior to nonviral vectors for efficiency of transduction and gene expression. This is not unexpected, given the evolutionary history of viruses. They have been under constant pressure to successfully maintain and propagate their life cycle in the face of ever-evolving host immune mechanisms. As a result, viral gene therapy researchers have an ample variety of vectors, each with particular characteristics that make them attractive reagents: often high levels of transgene expression, varied range of host cell tropism, and ease of vector production in cell systems. However, major modifications must be made to prevent them from becoming pathogenic to the host. This involves making them conditionally replication-defective. The virus must be able to replicate in a producer system or cell line, but not when delivered into the host. Conditional replication is generally achieved by deleting essential viral genes necessary for replication or packaging of the virus. Hence, these viral vectors should be able to facilitate transgene expression but be unable to autonomously replicate within a host organism. To be able to generate such vectors *in vitro*, producer cell lines are designed to efficiently complement the missing essential factors of the virus life cycle *in trans*. This can be done transiently or stably in the producer cells. This is most simply done by transfecting DNA plasmids expressing the essential genes into the producer cell lines. As of now, the most widely used viral vectors are based on oncoretrovirus and adenovirus. However, other viruses such as herpesvirus are increasingly being studied.

Adeno-associated virus (AAV)

Adeno-associated viruses do not cause known disease in humans. Categorized within the parvoviridae family, adeno-associated virus offers at least three advantageous features for gene therapy: integration into the mammalian genome, absence of associated pathogenicity in humans, and wide range of cellular tropism. However, these viruses are among the smallest and are subsequently constrained by one of the lowest transgene packaging limits. Furthermore,

AAVs' tropism for the relatively widespread heparan sulfate proteoglycan (HSPG) receptor complicates attempts at achieving cell-specific transduction.^[20]

Single-stranded DNA genome molecules of either polarity are found in an AAV capsid. Both ends of the molecule comprise palindromic sequences that form base-paired hairpin structures, essential for priming replication and packaging. The conditional replication of wild type AAV is possible only via coinfection by either adenovirus or herpesvirus (helper viruses) or by UV irradiation,^[21] all of which likely stimulate cellular components necessary for viral replication initiation in the host cell.^[22] In the absence of these costimulators, the AAV genome integrates into the host genome, also via the hairpin-shaped DNA ends. This establishes a period of latency that is reversible by exposure to helper virus.^[23] There are two AAV genes that code for Rep (replication, transcription, and integration) and Cap (capsid protein) proteins.^[23] By deleting these genes, transgenes up to 5 kb in size can be inserted. To generate the viral vectors, the Rep and Cap genes are often provided *in trans* by cotransfection of plasmids, along with helper adenovirus infection. As an alternative to adenovirus infection, the early genes of adenovirus may be provided on a plasmid via transfection. One limitation of the recombinant AAV production system is the often mediocre yield when compared to adenovirus and the possibility of helper adenovirus contamination.^[22] Although wild-type AAV has been shown to exhibit site specificity for integration into a locus on human chromosome 19, a study on engineered AAV vectors found that genomic integration was nonhomologous and random.^[24] This introduces a rare but tangible possibility of AAV integration into a locus whose disruption may lead to oncogenesis. Adeno-Associated Virus preferentially transduces dividing cells, and has been used in several gene transfer approaches, such as epithelial^[25] and hepatocyte transductions. The hepatocyte transduction experiments were particularly significant because they employed a single administration of recombinant AAV vector to achieve sustained and therapeutic levels of human Factor IX in mice.^[26]

Non-cell-autonomous diseases such as hemophilia B (or factor IX deficiency) are well suited for use of AAV because transduction of a small percentage of target cells (for example, 5% of hepatocytes) may be sufficient for a clinically relevant effect. Importantly, AAV transduction may be optimized by using a serotype that offers stronger infection (and potentially less toxicity) to a desired cell type. For example, data from Chao et al.^[27] suggest that AAV serotype 1 expresses Factor IX in mouse muscle in quantities that are several logs higher in magnitude than the more often used serotype 2. Wilson's group^[28] found that of serotypes 1–8, AAV serotype 8 demonstrates optimal

hepatocyte gene transfer capacity in murine models of hemophilia B. Whereas AAV is less efficient at transducing tissues such as liver, it does hold the advantage of sustained expression as well as transduction of both dividing and non-dividing cells. Significantly, its cytotoxic effect upon transduced cells is substantially less than with many other viral vectors.^[25]

Retrovirus

Retroviruses have been correlated with several human inflammatory diseases, such as rheumatoid arthritis,^[29] as well as leukemia and AIDS. Retroviruses, whose genome consists of two positive-sense RNA single strand molecules,^[6] are integrated into the host genome following error-prone reverse transcription of the RNA molecule into double-stranded DNA. Integration is found to be nonspecific *in vivo*.^[30] Although retroviruses are relatively large, the transgene insertional capacity of retroviruses is only about 7.5 kb. The retroviral genome encodes three genes: gag (core proteins), pol (reverse transcriptase and integrase enzymes), and env (envelope proteins). These genes can all be deleted and replaced by a therapeutic transgene. As with AAV, these deleted genes must be provided in trans, often by cotransfection of different plasmids. Retroviral expression is contingent upon integration into actively dividing cells, placing a limitation on *in vivo* applications. However, this preference for active cell division may be an advantage in cancer gene therapy,^[22] in which targeting cells of high mitotic index would be desirable. Although genomic integration should in principle facilitate persistent transgene expression, retroviral transgene expression is often transient. This is due in part to vector inactivation by circulating complement proteins^[22] and by the suppression of retroviral gene expression, at the post-transcriptional level, by host interferon cytokines.^[31] Retroviral vectors are the most often used viral vector type in gene therapy trials focusing on hematopoietic cells because they have a particular tropism for this tissue. In fact retroviral vectors were used in the first gene transfer marking trials in the late 1980s.

The majority of retroviral constructs are based on the Moloney murine leukemia virus, or MMLV, virulent in both mouse and human tissue.^[22] The retroviridae family also includes lentiviruses. These are the most recently discovered members of the family, which include the human and simian immunodeficiency viruses (HIV, SIV). Unlike other retroviruses, lentiviruses transduce both dividing and nondividing cells. The natural tropism of lentiviruses for CD4+T cells makes them attractive for immunotherapy research. Importantly, lentiviral vectors may potentially serve as a preventive therapeutic agent against virulent HIV.^[6]

Like that of AAV, the random integration of retroviruses into a host genome raises the possibility of oncogenesis via tumor suppressor gene disruption, or by activation of a protooncogene. This risk may be higher than in AAV because of the presence of transcriptionally active long terminal repeat (LTR) sequences that may act as local enhancers of transcription. This risk of insertional oncogenesis has been realized. Two French patients enrolled in a gene therapy trial for X-SCID recently developed leukemialike symptoms. This adverse event likely arose because of insertion of the retrovirus within the LMO-2 gene, which is responsible for blood cell replication.

Herpes virus

Oral and genital herpes are untreatable conditions in the human population, caused by two herpes simplex viruses, HSV-1 and HSV-2. Although the herpesvirus may lay dormant in the person's neurons, outbreaks of sores do occur with various periodicity. These sores are the manifestation of "viral shedding" from infected skin cells. Importantly, a person does not have to be in the midst of a herpes outbreak to be potentially transmissible.^[72]

Herpesvirus (HSV) double-stranded DNA genomes can reach 235 kb, although the herpes simplex viral genome is 152 kb. The genomic sequence is characterized by two large unique regions flanked by short repeats. About half of the approximately 81 total genes can be removed from engineered HSV vectors, allowing massive transgene inserts up to 50 kb.^[22] In addition to their large size relative to other viral vectors, HSV particles follow a more complicated physical packaging arrangement: The DNA is wound upon a scaffold in the core, which is surrounded by an icosahedral capsid. Another protein-rich layer, named tegument, surrounds the capsid and is in turn encapsulated by the outer envelope. In the HSV life cycle, transcription occurs in a temporal pattern, and the following transcriptional units have been assigned: alpha (immediate early), beta (delayed early), and gamma (late).^[22] Herpes simplex virus vectors have been designed along two related paths: plasmid amplicons, and attenuated recombinant virus.^[22] The amplicons are circular DNA molecules, not actual virus, and resemble generic plasmids: an origin of replication, promoter, transgene, and a selectable marker. The key differences are that the replication origin derives from HSV, not bacteria, and an HSV packaging signal must also be included. Vector production consists of cell transfection with the amplicon in the presence of a helper virus that provides, in trans, all missing HSV components. Recent studies have confirmed high levels of transgene expression in cancer



cell lines from HSV amplicons.^[32,33] Recombinant HSV viral vectors, like the other engineered viral vectors, are attenuated via deletion of genes essential for viral replication, i.e., immediate early (alpha) genes.^[22] For vector production, missing factors are provided in trans via packaging cell lines. Viral gene removal alleviates vector-induced toxicity.

HSV is infective in multiple types but leads to degradation of both host and viral mRNA in nonneuronal cell types. This degradation is facilitated by HSV's virion host shutoff (Vhs) protein, which was recently found to have active nuclease activity.^[34] Neurons, then, are the only true HSV latency-permissive cells. Although HSV latency is restricted to nondividing neurons, HSV's lytic function has a fortuitous preference for dividing cells expressing activated Ras.^[35] This makes HSV vectors well suited for cancer suicide gene therapy. The HSV thymidine kinase (TK) enzyme has been widely employed in suicide therapy (in non-HSV and in nonviral vector applications) via enzymatic conversion of the prodrug gancyclovir (or its related compounds). The HSV-TK enzyme converts the prodrug into a toxic metabolite in HSV-TK-transduced cancer cells, but also kills neighboring cells to some degree via the "bystander effect."^[36] Herpes simplex virus-1 vectors offer a unique combination of large transgenic carrying capacity, neuronal specificity, and persistent latency in the recipient (potentially obviating the need for repeat vector administration). Consequently, HSV vectors have been developed for gene therapy in Parkinson's disease to express tyrosine hydroxylase and neurotrophic factors that have therapeutic potential in this setting.^[22,23] Similarly, in a rat study, nerve growth factor (NGF) was delivered via HSV vector to a superior cervical ganglion that then underwent axotomy. Herpes simplex virus-mediated expression of NGF was found to abrogate the tyrosine hydroxylase deactivation normally associated with axotomy. Indeed, the HSV-transduced ganglion had higher levels of tyrosine hydroxylase than control ganglia 10 days after the axotomy.^[37]

Adenovirus

Adenoviral vectors have been among the most flexible of gene transfer vectors. Wild-type adenovirus is pathogenic and can cause upper respiratory tract symptoms, such as those associated with the common cold. However, it can also cause more serious disease including conjunctivitis, pharyngitis, and hepatitis.^[22] Of more than 40 human serotypes, Ad2 and Ad5 have found the most applications in gene transfer. These serotypes have not been associated with human oncogenesis. Applications for adenoviruses are second only to retroviruses in human gene therapy settings.^[38] The

36-kb linear genome of adenovirus is composed of double-stranded DNA with each 5' end bound by a 55 kD terminal protein (TP). Terminal protein primes viral replication and also functions in packaging signal (ψ)-mediated insertion of viral genomes into capsids.^[39] Packaging constraints limit the size of an engineered genome to 105% of a wild-type genome's size.^[22] The genome is flanked by inverted terminal repeats, with the packaging signal at the left end, and is enveloped by an icosahedral capsid composed of hexon and penton peptide bases. From each of the icosahedron's 12 vertices extends a fiber peptide, whose globular end facilitates recognition by host cell receptors, such as CAR (Coxsackie and adenovirus receptor). The fiber's shaft is also implicated in binding to cellular heparan sulfate glycosaminoglycan receptors. The penton bases are involved in host cell entry by adenovirus: They are recognized by integrin receptors on the cell membrane and initiate endocytosis. Furthermore, a toxic function of pentons may facilitate rupturing of adenovirus-containing endocytic vesicles, enhancing migration of intact adenoviral cores from the cytoplasm into the nucleus via the nuclear pore.^[22]

Like herpesvirus, adenoviral transcription occurs in a temporal pattern after the initial infection. E1A, the sole gene referred to as "immediate early" and the first-transcribed protein, is a transcriptional regulatory factor responsible for initiating other early genes' transcription, including E1B and E2 through E4. The E1B gene product binds p53, the cell cycle regulator, and, by preventing p53 activity, renders the infected cell replication-competent. E2 encodes the DNA polymerase, E3 downregulates MHC antigen presentation (and is referred to as the stealth gene), and the E4 protein disrupts the cytoskeletal network, making the infected cell more amenable to lysis and release of virus.^[22]

The final late phase of transcription includes genes for structural proteins. The majority of adenoviral vectors have been rendered replication-deficient by deletion of the E1 gene. For vector production purposes, a packaging cell line (HEK293) has been developed that expresses the adenoviral E1 gene products in trans. The viral E1 deletion prevents replication of the viral genome as well as expression from the early transcriptional units. Nevertheless, leaky transcription still occurs from early and late transcription units, contingent upon multiplicity of infection.^[22] The term "first generation" denotes replication-deficient adenoviral vectors with deletions in E1 or E1/E3. The deletion of E3 increased overall transgene carrying capacity but removes the stealth functions associated with that gene, rendering these vectors more immunogenic than E1-only deletions. Specifically, E3 gene products were shown to downregulate inflammatory chemokines^[25] and reduce Fas-mediated apoptosis among lymphocytes; hence,

recent studies have suggested returning the E3 gene in E1/E3-deleted adenoviral vectors.^[40]

Adenoviruses have a broad host cell range, and can infect hepatocytes especially efficiently. In fact, they are the only vector capable of achieving near complete transduction of the liver with a single intravascular injection. Unfortunately, of all the vectors, they are also the most immunogenic. This can be related to three phases of toxicity. First is an acute toxicity associated with interaction between vector and host cell (endothelial, antigen presenting cells, platelets) and a cytokine burst. This occurs within hours of a systemic injection. It is likely followed by an intermediate toxicity from hours to days of injection that involves augmentation of the systemic inflammatory immune response with a further amplification of cytokines and also dose-related thrombocytopenia. Finally, chronic toxicity can be seen within days to weeks after treatment and is due to a host cell-mediated immune response. Ways to avoid the acute toxicity include limiting the dose and focusing on local tissue delivery.

Substantial immune-mediated cytotoxicity against cells presenting viral antigens has provided the impetus to develop “gutless” or helper-dependent adenoviral vectors devoid of all viral genes. Although the inverted terminal repeats and packaging signal (ψ) remain, the rest of the viral genome is replaced by human genomic “stuffer” DNA.^[6] The issue of viral gene complementation in trans becomes more salient, because the E1-expressing 293 cell line is incapable of propagating gutted adenoviral vectors that are missing all early and late genes. Thus, a helper adenovirus has been engineered to provide all necessary viral genes in trans, hence the application of the “helper-dependent” nomenclature. The helper virus genome has also been engineered with a removable packaging signal, such that recombinant vector genomes are packaged into viral capsids whereas helper virus genomes are not.

Numerous gutless adenoviral vectors have displayed prolonged transgene expression and reduced immunogenicity as compared to their first-generation counterparts.^[41,42] An important caveat is that our understanding of the *in vivo* performance of most vectors has been limited to cell culture and mouse models. How they function *in vivo* in humans may be significantly different. After having addressed the variables of pharmacogenetic variation to any biological therapy in an outbred genetically diverse population, careful risk-benefit considerations should be made in every case of human gene therapy.

HOST RESPONSES TO GENE THERAPY

An aspect common to all gene therapy systems, both viral and nonviral, is their potential to trigger a host

response. For this reason, it has not been uncommon for researchers to develop methods of high level transduction and expression *in vitro* that do not translate to *in vivo* success.

Immune Responses

All organisms currently utilized for gene therapy employ robust, highly evolved systems to respond to foreign material. These response systems consist of at least two pathways: innate immune and adaptive immune responses. The innate immune response employs nonspecific effector molecules and cells to recognize and eliminate the majority of foreign particles that enter the host. The interaction between effector and target can occur literally within minutes, and is sufficient to clear many infections. Pathogens that evade or overwhelm the innate system, however, are subject to elimination by the adaptive immune response. The adaptive system employs specific antigen-targeting cells that undergo clonal expansion. The specification and expansion of these cells requires a latency period of up to several days before an attack by mature targeting cells commences. Although the timing and pathogen-specificity are drastically different between the innate and adaptive responses, several effector components are common to both.

Innate immunity

Some of the earliest responses to foreign material are unassociated with the immune system *per se*. As mentioned earlier, for example, host intra- and extracellular nucleases and proteases may lead to considerable degradation of therapeutic nucleic acids or recombinant proteins, respectively. Innate immunity comprises effectors that act on classes of foreign particles, as opposed to adaptive immunity's recognition of a specific antigenic epitope. Complement proteins, for example, bind to particles and cells circulating within the plasma. “Self” cells express complement control proteins that cleave complement proteins into inactive isoforms. Pathogenic surfaces that fail to express these control proteins are coated by active forms of complement. This active complement is a recruitment factor for circulating macrophages. Macrophages, in addition to responding to active complement on pathogenic surfaces, also display receptors with affinity for general features of bacteria (particularly lipopolysaccharide, a common bacterial cell wall component) and yeast. Additionally, macrophages that ingest foreign particles and present their antigenic epitopes are referred to as APCs, antigen presenting cells. These APCs are a critical link between innate immunity and the transition to antigen-specific



adaptive immune responses. Furthermore, these macrophages initiate cytokine cascades that recruit additional macrophages and initiate inflammatory host responses. The inflammatory response is characterized by a local widening of blood vessels leading to a lower velocity of blood flow, which in turn allows leukocytes and other blood components to aggregate at the infection site. The interferon cytokines are active against viral replication, fulfilling the dual roles of activation of a ribonuclease (RNA degradation) along with increasing major histocompatibility complex (MHC) class I expression. Increased MHC expression increases presentation of viral antigens by infected cells. Natural killer (NK) cells conversely target those cells with abnormally low MHC expression, and employ cytotoxic effector molecules against targets. While innate effector functions are general, not pathogen-specific, they may act instantly against many pathogens and can clear certain infections within a few days or even a few hours.^[43] Importantly, activation of innate immunity and the associated systemic inflammatory immune response can cause significant morbidity and mortality for the host.

Adaptive immunity

Adaptive immunity is composed of T and B lymphocyte cellular responses. T and B lymphocytes derive from the thymus and the bone marrow, respectively. Both are generated in an inactive, or naïve, state. Activation of either is dependent upon the combined interaction with antigen and costimulatory signals. Activation and proliferation of these lymphocytes takes up to several days, highlighting the significance of the early innate immune responses in attenuating pathogenic activity until adaptive responses are mounted.

In the case of T cells, exposure to antigen followed by interaction with an APC (the costimulator), displaying the same antigen, leads to a tighter association between the cells. Mature APCs are found in the lymphoid organs. The association between the two cells leads to clonal expansion of that T cell into a T cell subtype: Th1 and Th2 are helper T cell subtypes involved with macrophage and B-cell stimulation. Another subtype is the cytotoxic or killer T cell. Killer T cells employ lytic enzymes against target cell membranes.

B lymphocytes contribute the humoral or antibody-mediated response. The first step toward activation is also exposure to antigen. The costimulator for a B cell is an already-armed helper T cell, which provides stimulatory signals that induce hyperproliferation of B cells and somatic hypermutation (which maximizes antibody diversity) at their immunoglobulin loci; these

processes occur in germinal centers of secondary lymphoid organs. These primed B cells return to the systemic circulation where they can secrete antibodies that carry out several functions: They inhibit pathogen entry into host cells (neutralization), they may inactivate pathogen-generated toxins, and they coat particles to mark them for phagocytosis (opsonization). Hence, in any gene transfer protocol, acute toxicity of the vector must be considered. Later, toxicity directed at the vector-transduced cells or at the therapeutic transgene (which can itself act as a neoantigen) needs to also be carefully considered.

Clinical Trials in Gene Therapy

There have been numerous early phase clinical trials of gene therapy. The diseases targeted are diverse and include cystic fibrosis, Duchenne muscular dystrophy, severe combined immunodeficiencies, heart failure, cancer, and hemophilia.^[44–50,69]

Cancer

Cancer accounts for the majority of gene therapy clinical trials. Two-thirds of the gene therapy protocols that have been reviewed by the National Institutes of Health Recombinant DNA Advisory Committee (NIH RAC) are for the treatment of cancer.^[51] Two examples of cancer gene therapy approaches are prodrug activation and tumor-suppressor gene replacement. Both modalities have been employed for the treatment of prostate cancer.

In the prodrug approach, a gene is delivered to the prostate via a viral vector that has the ability to convert an inactive nontoxic prodrug into an active toxic metabolite that can destroy prostate cancer cells. Because the active drug is confined to the prostate, the patient is spared the potentially severe side effects that could result if it were delivered systemically. In one such trial, the prodrug used was gancyclovir, which is converted to its active form by the herpes simplex viral gene, thymidine kinase (HSVtk).^[52] This gene was delivered via an adenoviral vector injected into the prostates of 52 patients who were subsequently given two-week courses of systemic gancyclovir. Prostate-specific antigen (PSA) was measured as an endpoint. PSA levels rise with progression of prostate cancer and fall when therapy has been effective. Side effects were mild and nearly half of the patients had a decline in PSA.^[53,54]

p53 is a tumor suppressor gene; mutations in this gene have been associated with more severe disease in patients with prostate cancer.^[55] In vitro studies demonstrated that prostate cancer cells that are transfected with an adenoviral vector encoding normal

p53 undergo programmed cell death (apoptosis).^[56] Thirty patients with prostate cancer were given intraprostatic injections of the adenoviral vector encoding p53 and followed over time. Analysis of preinjection prostate biopsies demonstrated minimal p53 expression at baseline in the cancer cells. After injection, biopsies showed p53 expression and evidence of apoptosis.^[52] Side effects were minimal.

Genetic Disease

Several coagulation factors circulate within the bloodstream, and comprise catalytic enzymes or inactive protein precursors necessary for proper blood coagulation. Hemophilia A and B are disorders of coagulation that result from a deficiency or abnormality in the plasma coagulation factors VIII and IX, respectively.^[57] There are three ongoing gene therapy trials for hemophilia. Two are directed at hemophilia A and one at hemophilia B. All use different modes of gene delivery.

A muscle-directed approach was employed in a clinical trial of adeno-associated virus delivering factor IX for the treatment of hemophilia B.^[58,59] Three adults with severe hemophilia B were treated. Safety analysis revealed no germline transmission of vector and no development of inhibitory antibodies to factor IX. Analysis of muscle biopsies by PCR and Southern blot found the presence of vector sequences, and immunohistochemistry revealed evidence of FIX (factor IX) expression. However, only modest improvements in the clinical endpoints of circulating FIX and frequency of exogenous factor IX infusions required were noted.

Another study used a retroviral vector to deliver factor VIII to patients with hemophilia A. Retroviral vectors require cells to be dividing for effective transduction. Preclinical studies using this vector to deliver transgene to FVIII-(factor VIII)-deficient mice demonstrated detectable FVIII production in 6 of 13 treated animals with inhibitor antibody development in the other seven.^[60] Expressing animals maintained FVIII production for over 14 months. It should be noted that neonatal mice were used in these experiments. At this age there is rapid proliferation of liver cells leading to effective retroviral transduction. The clinical trial underway will determine if these results can be translated to adult patients.^[61]

An *ex vivo* approach has been used to deliver transfected autologous fibroblasts to six patients with severe hemophilia A.^[62] In this design a skin biopsy was taken from each patient and transfected with plasmids encoding FVIII. Cells producing FVIII were identified and propagated *in vitro* by clonal expansion. These cells were then surgically implanted into the

omentum of the patients. No serious adverse events were reported and inhibitors to FVIII were not detected. In four of six patients, FVIII levels in the serum rose slightly. This rise was correlated with a decrease in episodes of bleeding and exogenous factor requirements.

Whereas none of these studies resulted in production of sufficient factor to cure affected patients, they all demonstrate that effective gene transfer for this disease is possible. Moreover, the therapies were well tolerated without significant toxicity.

To date, the only trial that has unambiguously achieved a clinical cure involves an inherited immunodeficiency. Unfortunately, the therapy was also associated with significant toxicity directly attributable to the gene transfer. X-linked severe combined immunodeficiency (X-SCID) is an inherited disorder characterized by severe infections which usually presents in males early in life.^[63] It is caused by dysfunction of T and NK lymphocytes resulting from abnormal cytokine receptors required for the differentiation of these cells.^[64] Traditionally, the only treatment option available to patients with this deadly disease was allogeneic bone marrow transplantation (BMT). However, BMT may only partially restore immune function and can lead to serious complications.^[47] For this reason, a gene therapy clinical trial was initiated in France.

Bone marrow cells from five boys with X-SCID were harvested and transfected *ex vivo* with a retroviral vector carrying the normal form of the receptor gene. The cells were then expanded in culture and transplanted back into the boys. The patients were then followed for 2.5 years. The patients had few immediate side effects and a restoration of immunity sufficient to allow for a normal life.^[65] Six more patients eventually entered the trial and 9 of 11 were effectively cured. This was the first real clinical success for gene therapy and proved that the technique could cure a life-threatening disease.

However, long-term complications became evident when one of the boys developed a leukemia-like illness in which one type of his immune cells began to proliferate uncontrollably.^[66] The reason appears to be secondary to insertional mutagenesis in one clone from the treated cells. When an integrating virus such as a retrovirus is used in a stem cell population, the point of integration cannot be controlled. If the virus integrates into a gene that is involved in cell growth, uncontrolled growth can result. This is what happened in this case. Several months later a second boy developed a similar illness.^[67] In both cases, the retrovirus integrated into or near the LMO-2 gene, a gene known to be involved in lymphoproliferation.^[68] When gene therapy is applied to genetic disease, the long-term follow-up may differ from cancer trials where the baseline



comorbidity and mortality may obscure cumulative toxicity attributable to gene therapy. At the same time, the underlying disease pathophysiology may confound the evaluation of toxicity. The SCID phenotype has a negative survival impact on noncorrected cells; hence, there is a selective pressure for corrected cell clones that have a proliferative advantage. It may not be surprising that, over time, those cells with activated growth signals such as in the LMO-2 integrants will appear.

While the clinical application of gene therapy has had setbacks, significant milestones have been achieved. Each trial serves to incrementally expand our knowledge base. Continued efforts to increase our understanding of the complex mechanisms involved are ongoing. The long-term goal is the development of a system or systems that can safely deliver therapeutic genes to patients who need them.

ARTICLES OF FURTHER INTEREST

Chitosan; Controlled Release; Gene-Activated Matrix; Host Reactions

REFERENCES

- Maruyama, H.; Higuchi, N.; Nishikawa, Y.; Kameda, S.; Iino, N.; Kazama, J.J.; Takahashi, N.; Sugawa, M.; Hanawa, H.; Tada, N.; Miyazaki, J.; Gejyo, F. High-level expression of naked DNA delivered to rat liver via tail vein injection. *J. Gene Med.* **2002**, *4*, 333–341.
- Shyu, K.G.; Wang, M.T.; Wang, B.W.; Chang, C.C.; Leu, J.G.; Kuan, P.; Chang, H. Intramyocardial injection of naked DNA encoding HIF-1 α /VP16 hybrid to enhance angiogenesis in an acute myocardial infarction model in the rat. *Cardiovasc. Res.* **2002**, *54*, 576–583.
- De Lucca, F.L.; Sawan, F.M.; Watanabe, M.A.; de Souza, L.R. Effect of the calcium phosphate-mediated RNA uptake on the transfer of cellular immunity of a synthetic peptide of HIV-1 to human lymphocytes by exogenous RNA. *Mol. Cell. Biochem.* **2001**, *228*, 9–14.
- Jeschke, M.G.; Richter, G.; Hofstadter, F.; Herndon, D.N.; Perez-Polo, J.R.; Jauch, K.W. Non-viral liposomal keratinocyte growth factor (KGF) cDNA gene transfer improves dermal and epidermal regeneration through stimulation of epithelial and mesenchymal factors. *Gene Ther.* **2002**, *9*, 1065–1074.
- Herweijer, H.; Wolff, J.A. Progress and prospects: Naked DNA gene transfer and therapy. *Gene Ther.* **2003**, *10*, 453–458.
- <http://www-ermm.cbuc.cam.ac.uk/99000691h.htm> (accessed January 2003).
- Glasspool-Malone, J.; Steenland, P.R.; McDonald, R.J.; Sanchez, R.A.; Watts, T.L.; Zabner, J.; Malone, R.W. DNA transfection of macaque and murine respiratory tissue is greatly enhanced by use of a nuclease inhibitor. *J. Gene Med.* **2002**, *4*, 322–323.
- de Carvalho Bittencourt, M.; Saas, P.; Fresnay, S.; Yerly-Motta, V.; Ferrand, C.; Perruche, S.; Duperrier, A.; Herve, P.; Tiberghien, P.; Chalmers, D.E. Exposure to exogenous DNA can modify the sensitivity of the Fas apoptotic pathway. *J. Gene Med.* **2002**, *4*, 14–24.
- Nishikawa, M.; Huang, L. Nonviral vectors in the new millennium: Delivery barriers in gene transfer. *Hum. Gene Ther.* **2001**, *12*, 861–870.
- Hartikka, J.; Sukhu, L.; Buchner, C.; Hazard, D.; Bozoukova, V.; Margalith, M.; Nishioka, W.K.; Wheeler, C.J.; Manthorp, M.; Sawdey, M. Electroporation-facilitated delivery of plasmid DNA in skeletal muscle: Plasmid dependence of muscle damage and effect of poloxamer 188. *Mol. Ther.* **2001**, *4*, 407–415.
- Heller, L.C.; Coppola, D. Electrically mediated delivery of vector plasmid DNA elicits an antitumor effect. *Gene Ther.* **2002**, *9*, 1321–1325.
- Somiari, S.; Glasspool-Malone, J.; Drabick, J.J.; Gilbert, R.A.; Heller, R.; Jaroszeski, M.J.; Malone, R.W. Theory and in vivo application of electroporative gene delivery. *Mol. Ther.* **2000**, *2*, 178–187.
- Lee, M.J.; Cho, S.S.; You, J.R.; Lee, Y.; Kang, B.D.; Choi, J.S.; Park, J.W.; Suh, Y.L.; Kim, J.A.; Kim, D.K.; Park, J.S. Intraperitoneal gene delivery mediated by a novel cationic liposome in a peritoneal disseminated ovarian cancer model. *Gene Ther.* **2002**, *9*, 859–866.
- Baru, M.; Nahum, O.; Jaaro, H.; Sha'anani, J.; Nur, I. Lysosome-disrupting peptide increases the efficiency of in-vivo gene transfer by liposome-encapsulated DNA. *J. Drug Target.* **1998**, *6*, 191–199.
- Tsunoo, H.; Komura, S.; Ohishi, N.; Yajima, H.; Akiyama, S.; Kasai, Y.; Ito, K.; Nakao, A.; Yagi, K. Effect of transfection with human interferon-beta gene entrapped in cationic multilamellar liposomes in combination with 5-fluorouracil on the growth of human esophageal cancer cells in vitro. *Anticancer Res.* **2002**, *22*, 1537–1543.
- Hailes, H.; Hurley, C.; Tabor, A.; Barker, S.; Hart, S. Synthesis and Biological Properties of Cationic Lipids in a Lipid/Peptide/DNA Transfection Vector. In *Molecular Therapy*, American Society of Gene Therapy, Boston, Massachusetts, June 5–9; Kellog, K., Rishizky, P., Wood, D., Eds.; Elsevier Science: New York, 2002; 226–275.
- Wang, D.; Roy, K.; Hedley, M.; Barman, S. Enhanced Intramuscular Gene Expression by DNA Delivery via Combined Polymeric Networks/Electroporation. In *Molecular Therapy*, American Society of Gene Therapy, Boston, Massachusetts, June 5–9; Kellog, K., Rishizky, P., Wood, D., Eds.; Elsevier Science: New York, 2002; 217–272.
- Pun, S.; Cheng, J.; Bellocq, N.; Jensen, G.; Davis, M. Cyclodextrin Polymer-Based Gene Delivery Systems for Hepatocyte Targeting. In *Molecular Therapy*, American Society of Gene Therapy, Boston, Massachusetts, June 5–9; Kellog, K., Rishizky, P., Wood, D., Eds.; Elsevier Science: New York, 2002; 219–669.

19. Koskinen, M.; Sailyloja, E.; Ahola, M.; Ahonen, M.; Kahari, V.; Salonen, J. Silica Xerogel Carrier as Encapsulating Material for Controlled Release of Recombinant Adenoviruses. In *Molecular Therapy*, American Society of Gene Therapy, Boston, Massachusetts, June 5–9; Kellog, K., Rishizky, P., Wood, D., Eds.; Elsevier Science: New York, 2002; 224–684.
20. Ponnazhagan, S.; Mahendra, G.; Kumar, S.; Thompson, J.A.; Castillas, M., Jr. Conjugate-based targeting of recombinant adeno-associated virus type 2 vectors by using avidin-linked ligands. *J. Virol.* **2002**, *76*, 12900–12907.
21. Berns, K.I. Parvovirus replication. *Microbiol. Rev.* **1990**, *54*, 316–329.
22. Latchman, D.S.; Coffin, R.S. Viral vectors for gene therapy in Parkinson's disease. *Rev. Neurosci.* **2001**, *12*, 69–78.
23. Fink, D.J.; Poliani, P.L.; Oligino, T.; Kriskey, D.M.; Goins, W.F.; Glorioso, J.C. Development of an HSV-based vector for the treatment of Parkinson's disease. *Exp. Neurol.* **1997**, *144*, 103–121.
24. Rutledge, E.A.; Russell, D.W. Adeno-associated virus vector integration junctions. *J. Virol.* **1997**, *71*, 8429–8436.
25. Lesokhin, A.M.; Delgado-Lopez, F.; Horwitz, M.S. Inhibition of chemokine expression by adenovirus early region three (E3) genes. *J. Virol.* **2002**, *76*, 8236–8243.
26. Snyder, R.O.; Miao, C.H.; Patijn, G.A.; Spratt, S.K.; Danos, O.; Nagy, D.; Gown, A.M.; Winther, B.; Meuse, L.; Cohen, L.K.; Thompson, A.R.; Kay, M.A. Persistent and therapeutic concentrations of human factor IX in mice after hepatic gene transfer of recombinant AAV vectors. *Nat. Genet.* **1997**, *16*, 270–276.
27. Chao, H.; Monahan, P.E.; Liu, Y.; Samulski, R.J.; Walsh, C.E. Sustained and complete phenotype correction of hemophilia B mice following intramuscular injection of AAV1 serotype vectors. *Mol. Ther.* **2001**, *4*, 217–222.
28. Gao, G.P.; Alvira, M.R.; Wang, L.; Calcedo, R.; Johnston, J.; Wilson, J.M. Novel adeno-associated viruses from rhesus monkeys as vectors for human gene therapy. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 11854–11859.
29. Brand, A.; Griffiths, D.J.; Herve, C.; Mallon, E.; Venables, P.J. Human retrovirus-5 in rheumatic disease. *J. Autoimmun.* **1999**, *13*, 149–154.
30. Bushman, F.D. Integration site selection by lentiviruses: Biology and possible control. *Curr. Top. Microbiol. Immunol.* **2002**, *261*, 165–177.
31. Ghazizadeh, S.; Carroll, J.M.; Taichman, L.B. Repression of retrovirus-mediated transgene expression by interferons: Implications for gene therapy. *J. Virol.* **1997**, *71*, 9163–9169.
32. Tolba, K.A.; Bowers, W.J.; Eling, D.J.; Casey, A.E.; Kipps, T.J.; Federoff, H.J.; Rosenblatt, J.D. HSV amplicon-mediated delivery of LIGHT enhances the antigen-presenting capacity of chronic lymphocytic leukemia. *Mol. Ther.* **2002**, *6*, 455–463.
33. Delman, K.A.; Zager, J.S.; Bennett, J.J.; Malhotra, S.; Ebright, M.I.; McAuliffe, P.F.; Halterman, M.W.; Federoff, H.J.; Fong, Y. Efficacy of multiagent herpes simplex virus amplicon-mediated immunotherapy as adjuvant treatment for experimental hepatic cancer. *Ann. Surg.* **2002**, *236*, 337–342. Discussion 342–333.
34. Everly, D.N., Jr.; Feng, P.; Mian, I.S.; Read, G.S. mRNA degradation by the virion host shutoff (Vhs) protein of herpes simplex virus: Genetic and biochemical evidence that Vhs is a nuclease. *J. Virol.* **2002**, *76*, 8560–8571.
35. Farassati, F.; Yang, A.D.; Lee, P.W. Oncogenes in Ras signaling pathway dictate host-cell permissiveness to herpes simplex virus 1. *Nat. Cell. Biol.* **2001**, *3*, 745–750.
36. van Dillen, I.J.; Mulder, N.H.; Vaalburg, W.; de Vries, E.F.; Hospers, G.A. Influence of the bystander effect on HSV-tk/GCV gene therapy. A review. *Curr. Gene Ther.* **2002**, *2*, 307–322.
37. Federoff, H.J.; Geschwind, M.D.; Geller, A.I.; Kessler, J.A. Expression of nerve growth factor in vivo from a defective herpes simplex virus 1 vector prevents effects of axotomy on sympathetic ganglia. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 1636–1640.
38. <http://www.wiley.co.uk/wileychi/genmed/clinical> (accessed February 2004).
39. Webster, A.; Leith, I.R.; Nicholson, J.; Hounsell, J.; Hay, R.T. Role of preterminal protein processing in adenovirus replication. *J. Virol.* **1997**, *71*, 6381–6389.
40. McNees, A.L.; Garnett, C.T.; Gooding, L.R. The adenovirus E3 RID complex protects some cultured human T and B lymphocytes from Fas-induced apoptosis. *J. Virol.* **2002**, *76*, 9716–9723.
41. Reddy, P.S.; Sakhuja, K.; Ganesh, S.; Yang, L.; Kayda, D.; Brann, T.; Pattison, S.; Golightly, D.; Idamakanti, N.; Pinkstaff, A.; Kaloss, M.; Barjot, C.; Chamberlain, J.S.; Kaleko, M.; Connelly, S. Sustained human factor VIII expression in hemophilia a mice following systemic delivery of a gutless adenoviral vector. *Mol. Ther.* **2002**, *5*, 63–73.
42. Toietta, G.; Pastore, L.; Cerullo, V.; Finegold, M.; Beaudet, A.L.; Lee, B. Generation of helper-dependent adenoviral vectors by homologous recombination. *Mol. Ther.* **2002**, *5*, 204–210.
43. Janeway, C.; Travers, P.; Walport, M.; Capra, D. *Immunobiology: The Immune System in Health and Disease*, 4th Ed.; Elsevier Science Limited: New York, 1999.
44. Mayer-Kuckuk, P.; Banerjee, D.; Kemeny, N.; Fong, Y.; Bertino, J.R. Molecular therapies for colorectal cancer metastatic to the liver. *Mol. Ther.* **2002**, *5*, 492–500.
45. Harjai, K.J.; Chowdhury, P.; Grines, C.L. Therapeutic angiogenesis: A fantastic new adventure. *J. Interv. Cardiol.* **2002**, *15*, 223–229.
46. Rutanen, J.; Markkanen, J.; Yla-Herttuala, S. Gene therapy for restenosis: Current status. *Drugs* **2002**, *62*, 1575–1585.
47. Otsu, M.; Candotti, F. Gene therapy in infants with severe combined immunodeficiency. *BioDrugs* **2002**, *16*, 229–239.
48. Bragonzi, A.; Conese, M. Non-viral approach toward gene therapy of cystic fibrosis lung disease. *Curr. Gene Ther.* **2002**, *2*, 295–305.



49. Thioudellet, C.; Blot, S.; Squiban, P.; Fardeau, M.; Braun, S. Current protocol of a research phase I clinical trial of full-length dystrophin plasmid DNA in Duchenne/Becker muscular dystrophies. Part I: Rationale. *Neuromuscul. Disord.* **2002**, *12* (Suppl), S49.
50. Klink, D.T.; Glick, M.C.; Scanlin, T.F. Gene therapy of cystic fibrosis (CF) airways: A review emphasizing targeting with lactose. *Glycoconj. J.* **2001**, *18*, 731–740.
51. Zwiebel, J.A. Cancer gene and oncolytic virus therapy. *Semin. Oncol.* **2001**, *28*, 336–343.
52. Morris, M.J.; Scher, H.I. Novel therapies for the treatment of prostate cancer: Current clinical trials and development strategies. *Surg. Oncol.* **2002**, *11*, 13–23.
53. Herman, J.R.; Adler, H.L.; Aguilar-Cordova, E.; Rojas-Martinez, A.; Woo, S.; Timme, T.L.; Wheeler, T.M.; Thompson, T.C.; Scardino, P.T. In situ gene therapy for adenocarcinoma of the prostate: A phase I clinical trial. *Hum. Gene Ther.* **1999**, *10*, 1239–1249.
54. Shalev, M.; Kadmon, D.; Teh, B.S.; Butler, E.B.; Aguilar-Cordova, E.; Thompson, T.C.; Herman, J.R.; Adler, H.L.; Scardino, P.T.; Miles, B.J. Suicide gene therapy toxicity after multiple and repeat injections in patients with localized prostate cancer. *J. Urol.* **2000**, *163*, 1747–1750.
55. Navone, N.M.; Troncoso, P.; Pisters, L.L.; Goodrow, T.L.; Palmer, J.L.; Nichols, W.W.; von Eschenbach, A.C.; Conti, C.J. p53 protein accumulation and gene mutation in the progression of human prostate carcinoma. *J. Natl. Cancer Inst.* **1993**, *85*, 1657–1669.
56. Yang, C.; Cirielli, C.; Capogrossi, M.C.; Passaniti, A. Adenovirus-mediated wild-type p53 expression induces apoptosis and suppresses tumorigenesis of prostatic tumor cells. *Cancer Res.* **1995**, *55*, 4210–4213.
57. Kaufman, R.J. Advances toward gene therapy for hemophilia at the millennium. *Hum. Gene Ther.* **1999**, *10*, 2091–2107.
58. Kay, M.A.; et al. Evidence for gene transfer and expression of factor IX in hemophilia B patients treated with an AAV vector. *Nat. Genet.* **2000**, *24*, 257–261.
59. High, K.A. AAV-mediated gene transfer for hemophilia. *Ann. N. Y. Acad. Sci.* **2001**, *953*, 64–74.
60. VandenDriessche, T.; Vanslebrouck, V.; Goovaerts, I.; Zwinnen, H.; Vanderhaeghen, M.L.; Collen, D.; Chuah, M.K. Long-term expression of human coagulation factor VIII and correction of hemophilia a after in vivo retroviral gene transfer in factor VIII-deficient mice. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 10379–10384.
61. High, K.A. Gene therapy: A 2001 perspective. *Haemophilia* **2001**, *7* (Suppl 1), 23–27.
62. Roth, D.A.; Tawa, N.E., Jr.; O'Brien, J.M.; Treco, D.A.; Selden, R.F. Factor VIII Transkaryotic therapy study group. Nonviral transfer of the gene encoding coagulation factor VIII in patients with severe hemophilia A. *N. Engl. J. Med.* **2001**, *344*, 1735–1742.
63. Scriver, C.R. *The Metabolic and Molecular Bases of Inherited Disease*; McGraw-Hill: New York, 2001.
64. Fischer, A.; de Saint-Basile, G.; Disanto, J.P.; Hacein-Bey, S.; Sharara, L.; Cavazzana-Calvo, M. [Gene therapy for hereditary immunodeficiencies]. *C. R. Seances Soc. Biol. Fil.* **1996**, *190*, 77–93.
65. Hacein-Bey-Abina, S.; Le Deist, F.; Carlier, F.; Bouneaud, C.; Hue, C.; De Villartay, J.P.; Thrasher, A.J.; Wulffraat, N.; Sorensen, R.; Dupuis-Girod, S.; Fischer, A.; Davies, E.G.; Kuis, W.; Leiva, L.; Cavazzana-Calvo, M. Sustained correction of X-linked severe combined immunodeficiency by ex vivo gene therapy. *N. Engl. J. Med.* **2002**, *346*, 1185–1193.
66. Check, E. Cancer fears cast doubts on future of gene therapy. *Nature* **2003**, *421*, 678.
67. Check, E. Second cancer case halts gene-therapy trials. *Nature* **2003**, *421*, 305.
68. Neale, G.A.; Rehg, J.E.; Goorha, R.M. Disruption of T-cell differentiation precedes T-cell tumor formation in LMO-2 (rhombotin-2) transgenic mice. *Leukemia* **1997**, *11* (Suppl 3), 289–290.
69. <http://www.prochoiceforum.org.uk/ri6.asp> (accessed January 2003).
70. <http://www.ncbiotech.org/biotech101/timeline.cfm> (accessed December 2002).
71. <http://www.bentham.org/cgt1-1/bronte/brontems.htm> (accessed November 2002).
72. <http://www.niaid.nih.gov/factsheets/stdherp.htm> (accessed March 2003).
73. <http://www-micro.msb.le.ac.uk/335/peel/peel1.html> (accessed November 2002).

Glucose-Responsive Hydrogels

Nicholas A. Peppas

Christie D. Bures

*Departments of Chemical and Biomedical Engineering and Division of Pharmaceutics,
University of Texas at Austin, Austin, Texas, U.S.A.*

INTRODUCTION

Hydrogels are hydrophilic polymer networks that can absorb large amounts of water, but remain insoluble because of the presence of cross-links, entanglements, or crystalline regions.^[1] Hydrogels often exhibit good biocompatibility, because of their high water content. They have been used in various biomedical applications including biosensors, contact lenses, and drug delivery systems because of these characteristics.^[2] Hydrogels can also be sensitive to the conditions of the external environment, because of the presence of certain functional groups along the polymer chains. The swelling behavior of these materials may be dependent on pH, temperature, ionic strength, or even glucose concentration. In particular, there has been increased research interest in the development of glucose-sensitive polymeric systems that deliver appropriate amounts of insulin in response^[2,3,24] to changing glucose levels, so as to mimic the natural response of the body. This could lead to better control of blood glucose levels in diabetic patients. This approach involves an enzyme–substrate reaction that results in a pH change and a pH-sensitive polymer that responds to the change. Glucose reacts with glucose oxidase (GOD) forming gluconic acid (GlucA) and, thus decreasing the pH of the environment. With the change in pH, the gel swells or collapses depending on the characteristics of the particular polymer of the system. Insulin is released from this system with the change in the size of the pores of the polymer.^[3]

In the last few years, extensive work has been carried out with poly(methacrylic acid) (PMAA), poly(ethylene glycol) (PEG), poly(diethylamino ethyl methacrylate) as well as more complex structures including complexation hydrogels of poly(methacrylic acid-g-ethylene glycol) [P(MAA-g-EG)]. Interpolymer complexes are formed in these hydrogels because of hydrogen bonding between the hydrogens of the carboxylic group of the PMAA and the oxygens on the ether groups of the PEG chains. At low pH values, there is sufficient protonation of the carboxylic acid groups causing complexes to form. This results in a collapse of the gel because of the increased hydrophobicity in the polymer network. At high pH values, complexes break as the carboxylic groups

become ionized. This results in an expansion of the gel as electrostatic repulsion is produced within the network. In our specific hydrogel systems, the two species involved in the complexation are bound together in the same polymer. This allows for the reversible formation of complexes at appropriate conditions. Therefore, these materials allow for pH-sensitive solute permeation, a property that is important for drug release.^[4] When incorporating P(MAA-g-EG) with GOD, a “squeezing gel” is formed. At high concentrations of glucose, the GOD-catalyzed reaction of glucose produces GlucA, resulting in a decrease in the pH of the environment around the gel. Glucose-sensitive hydrogels collapse in response to the decrease in pH. Thus, insulin could be squeezed out of the network.

There are several important parameters controlling the behavior of glucose-responsive hydrogel system. The equilibrium and dynamic degrees of swelling are important parameters for calculating network mesh size and molecular weight between cross-links under different experimental conditions. The permeability characteristics of the gels are also important when considering the system for insulin release.

HYDROGELS AS DRUG CARRIERS

Hydrogels

Hydrogels can be described as materials that have the ability to swell in water and retain a significant amount of water within its structure. These materials are capable of absorbing water because of the presence of hydrophilic functional groups^[5] such as $-\text{OH}$, $-\text{COOH}$, $-\text{CONH}_2$, $-\text{CONH}$, and $-\text{SO}_3\text{H}$.

Hydrogels can be classified as neutral or ionic in nature. The swelling of neutral hydrogels is governed by the water–polymer thermodynamic mixing contribution to the overall free energy as well as an elastic polymer contribution. Equilibrium swelling is reached when these two forces are equal.^[1]

In ionic hydrogels, the swelling is governed by thermodynamic mixing, elastic–retractive forces, and ionic interactions between charged polymer and free ions. The swelling behavior is affected by an osmotic

G

force that develops as the charged groups on the polymer are neutralized by mobile counterions. Electrostatic repulsion is also produced between fixed charges and mobile ions inside the gel, affecting the swelling of ionic gels.^[6] The equilibrium swelling ratios of ionic hydrogels are often an order of magnitude higher than those of neutral gels, because of the presence of intermolecular interactions including coulombic, hydrogen bonding, and polar forces.^[7]

Polyelectrolytes have ionized macromolecular chains with more than one ionizable group in their backbone. Upon cross-linking, polyelectrolytes form three-dimensional networks that exhibit high degrees of swelling because of ionization and ion hydration. In these gels, ionization of the fixed charges of the backbone contributes to swelling. Polymer chains uncoil^[8] as electrostatic repulsion is produced between adjacent fixed charges.

Properties Affecting the Swelling of Ionic Hydrogels

There are many properties of ionic hydrogels, which contribute to the swelling of the network. Peppas and Khare^[9] have presented some of these properties, including the ionic content, ionization equilibrium considerations, nature of counterions, and the nature of the polymer.

An increase in the ionic content of the gel increases the hydrophilicity leading to faster swelling and a higher equilibrium degree of swelling. Anionic hydrogels, which contain carboxylic groups, swell at a pH higher than the gel pK_a , because of ionization within the network. At a pH lower than the pK_a , the gel is in a unionized state and does not swell to a high extent. The opposite behavior occurs for cationic gels that contain amine groups. In an ampholyte, the gel contains both acidic and basic groups. In this case, the isoelectric pH determines the transitional pH of swelling of the gel. A polyampholyte is highly swollen at high alkaline pH and low acidic pH because of ionization. However, near the isoelectric pH, the gel swells only to a moderate degree.

Ionization equilibrium considerations also affect the swelling behavior of ionic hydrogels. Fixed charges in the network lead to the formation of an electric double layer of fixed charges and counterions in the gel. According to Donnan equilibrium, the chemical potential of the ions inside the gel is equal to that of the ions outside the gel in the swelling medium. Donnan exclusion prevents the sorption of coions, because of electroneutrality resulting in a higher concentration of counterions in the gel phase than in the external swelling agent. The efficiency of coion exclusion, or an increase in the Donnan potential, increases with

decreasing solution concentration. Increasing ionic content of the gel also increases the efficiency of coion exclusion.

The nature of counterions also impacts the swelling of ionic hydrogels. Polyelectrolyte gels generally prefer counterions of higher valence. The counterion of higher charge is more strongly attracted to the gel, as the concentration inside the gel decreases. Therefore, the preference for higher valence counterions increases with dilution and for ionic gels with high internal molality. Polyelectrolyte gels also prefer counterions with smaller solvated equivalent volume.

The degree of swelling may be affected by the nature of the polymer. An increase in the hydrophobicity of the network can sharpen the transition pH of swelling. This leads to a decrease in the pH range over which the transition swelling behavior occurs. At constant pH, an increase in the ionic strength causes a decrease in the Donnan potential and water uptake resulting in a reduction of swelling.

ENVIRONMENTALLY RESPONSIVE HYDROGELS

There has recently been increased research in the preparation and characterization of materials responsive to changing environmental conditions. Some of the environmental conditions that can affect hydrogel swelling include pH, ionic strength, temperature, and glucose concentration.

In cross-linked networks that contain weakly acidic (anionic) and/or basic (cationic) pendant groups, water sorption can result in the ionization of these pendant groups, depending on the solution pH and ionic composition. The gel then acts as a semipermeable membrane to the counterions.^[10] This influences the osmotic balance between the hydrogel and the external solution through ion exchange, depending on the degree of ionization and ion-ion interactions. For ionic gels containing weakly acidic pendant groups, the equilibrium degree of swelling increases as the pH of the external solution increases,^[11] while the degree of swelling increases as the pH decreases for gels containing weakly basic pendant groups.^[12]

Some hydrogels undergo volume phase transition with a change in the temperature of the environmental conditions. The reversible volume change at the transition depends on the degree of ionization and the components of the polymer chains. There is usually a negligible or small positive enthalpy of mixing, which opposes the process. However, there is also a large gain in the entropy, which drives the process. The opposite behavior is observed in aqueous polymer solutions. This type of behavior is related to polymer phase separation, as the temperature is raised to a critical value

known as the lower critical miscibility or solution temperature (LCST).^[2] Tanaka^[13] investigated the effect of temperature on polyacrylamide gels. Networks showing lower critical miscibility temperature tend to shrink or collapse, as the temperature is increased above the LCST. The gels were found to swell upon lowering the temperature below the LCST.

Environmentally sensitive hydrogels have been studied as possible controlled insulin release systems for the treatment of diabetes. Controlling insulin release with respect to the amount of glucose in the bloodstream is of great importance for an effective insulin delivery device.^[14] An approach to producing materials that are sensitive to glucose involves the incorporation of a pH-sensitive hydrogel with immobilized GOD. GOD acts as a glucose sensor, as it produces GlucA by an enzymatic reaction with glucose. The GlucA produced lowers the pH of the medium, resulting in a change in the degree of swelling of the gel because of its pH sensitivity.^[15] This type of system will be discussed in greater detail later in the entry.

Controlled release systems are capable of delivering a drug at a specific rate over a desirable time period. Some of the advantages of controlled release over conventional drug therapies include maintenance of the drug in a desirable therapeutic range, localized drug delivery to a particular site of the body, protection of drugs that are rapidly destroyed by the body, better patient compliance, and increased comfort. Controlled release systems based on polymers or polymeric materials release drugs by diffusion, chemical reaction, or solvent activation.^[16]

Diffusion-controlled release systems can be classified as reservoir (membrane) and matrix (monolithic) systems. In a reservoir device, a polymeric membrane surrounds a core of drug. Drug release is controlled by the drug diffusion through the surrounding polymer membrane. A constant concentration gradient of drug can be maintained in the device if there is a saturated concentration of the drug inside the reservoir.

In a matrix system, the drug is uniformly dissolved or dispersed throughout the polymer matrix. Matrix systems exhibit a continuously decreasing release rate over time, because of a decrease in the concentration gradient at the polymer surface while the drug is released. It is therefore difficult to achieve zero-order release with matrix systems with conventional geometric shapes such as films, spheres, and cylinders. However, zero-order release can be obtained, however, by careful design of a matrix system. Increase of the surface area can compensate for lowering of the concentration gradient. This results in a suppression of the decrease in release rate with time. Near constant release has been achieved by releasing drug from the center pole of a hemispherical matrix that was coated with a non-permeable membrane.^[17]

Some examples of chemically controlled systems include bioerodible and biodegradable systems and pendent chain systems. In certain biodegradable systems, the polymer undergoes a decrease in molecular weight and eventually disintegrates or dissolves in the dissolution medium. As the polymer dissolves, drug can be released into the medium. Degradation is often hydrolytic and no enzymes are involved for polymers and copolymers of lactic, glycolic, and hydroxybutyric acids. Natural polymers including starch and cellulose are also biodegradable, but require enzymes for degradation. In erodible systems, the release of the active substance is controlled by surface or bulk erosion of the matrix.^[18] In pendent chain systems or prodrugs, the drug molecule is chemically bonded to a polymer backbone and is released through hydrolytic or enzymatic cleavage. The rate of hydrolysis controls the rate of drug release. An advantage of this type of drug release is the ability to target a drug to a particular cell type or tissue. But a disadvantage is the inability to hydrolyze exactly. Natural polymers, including polysaccharides, and synthetic polymers such as polylysine and copolymers of 2-hydroxypropylamide have been investigated for use as drug carriers in such systems. By modifying the structure of these polymers with the incorporation of sugar residues or sulfonyl units, a specific affinity can be obtained.^[9]

Solvent-activated systems include osmotic-controlled and swelling-controlled release systems. Osmotic systems contain a dry core and a semipermeable membrane that controls the permeation of water. The overall rate of drug release is controlled by the rate of water influx and remains constant provided the drug concentration across the membrane is constant. Advantages of osmotic-controlled systems include the capability of providing extended time-independent release and higher delivery rates that can be achieved by solution-diffusion systems. In swelling-controlled systems, the drug that is dispersed in the polymer diffuses out as water uptake occurs and the polymer swells. The drug release rate is dependent on both water diffusion into the polymer and chain relaxation. Continued swelling of this system results in the drug diffusing out at a faster rate. The rate of swelling of the network controls the overall drug release rate. The time dependence of the rate of drug release can be determined depending on the rate of water diffusion and chain relaxation.^[9]

SELF-REGULATED INSULIN DELIVERY DEVICES

Self-Regulation

Current methods of insulin delivery provide inadequate control of blood glucose levels in diabetic patients. This



has motivated research for the development of novel self-regulated devices that deliver insulin as a result of a triggering process because of high glucose concentration.

Diabetes mellitus is a disease that affects approximately 14 million people in U.S.A. It is characterized by the body's inability to completely and continuously metabolize glucose, resulting in poor regulation of blood sugar levels in the body. This poor control is attributed to a decrease in insulin secretion from the β -cells of the pancreas. In Type I (insulin-dependent) diabetes, this decrease in insulin secretion arises from a loss of β -cells. In Type II (non-insulin-dependent) diabetes, the decrease in insulin secretion is attributed to a defective response of the β -cells to glucose.^[19] Insulin is a hormone that accelerates glucose transport into most cells of the body, causing an increase in the rate of glucose metabolism, a decrease in blood glucose concentrations, and an increase in glycogen stored in tissues.^[20]

Control of blood glucose levels in diabetics is important in preventing health problems that could occur from hyperglycemia (high blood glucose levels) and hypoglycemia (low blood glucose levels). Normal blood glucose levels range from 70 to 120 mg/dl. During prolonged hyperglycemia, there is an increase in the oxidation of fats causing the accumulation of ketone substances, the by-product of fat metabolism, in the blood and urine. One breakdown product of the ketone substances, acetoacetic acid, is toxic and can cause diabetic coma in extreme untreated cases.^[21] Hypoglycemic conditions can also be dangerous. It is important to maintain blood glucose levels sufficiently high to provide the nutrition that is required by certain tissues and organs that depend on glucose as their sole source of energy, including the brain, retina, and germinal epithelium.^[20]

In a normal patient, blood glucose levels remain relatively constant at their basal value. High glucose levels signal the β -cells of the pancreas to increase insulin secretion and the α -cells to decrease glucagon secretion. The opposite occurs with low glucose levels. Insulin increases glucose uptake in fat and muscle tissues and suppresses the production of glucose from the liver, while glucagon stimulates the production of glucose. Glucose uptake in the brain is not insulin dependent.^[22]

In a diabetic patient, the β -cells of the pancreas do not produce sufficient insulin to provide for good control. Therefore, it is necessary for a diabetic to take insulin in some form to maintain glucose levels within the normal range. It is also important that diabetics control their diets and their activities.^[20]

The need for better control of blood glucose levels in diabetics has led to increased research in the development of controlled release devices that deliver appropriate amounts of insulin in response^[23,24] to changing

glucose levels, so as to mimic the body's natural response. Insulin release systems have been proposed, which consist of a pH-sensitive hydrogel with GOD and an enclosed insulin solution. Glucose diffuses into the hydrogel, where GOD catalyzes its conversion to GlucA, thereby decreasing the pH of the environment. Then the membrane permeability changes in response to the decrease in pH. It is owing to this change in permeability that insulin release varies with changing glucose concentrations.^[25]

Horbett et al.^[26] prepared a self-regulated, closed loop, feedback-controlled insulin delivery system based on an amine-containing hydrogel. In alkali solutions, the pendent NR_2 groups remained neutral, yielding a membrane that was unswollen and impermeable. In an acidic solution, however, the $-\text{NR}_2$ groups were protonated to $-\text{NR}_2\text{H}^+$, causing the hydrogel to swell and become insulin permeable. The pH change that created the acidic environment was produced by incorporating GOD in the membrane. The resulting pH in the membrane depended on the glucose concentration in the fluid that permeated the hydrogel. However, this system produced only an increase in permeability of a factor of 1.5 with glucose concentrations ranging from 50 to 500 mg/dl. This change was too small to be useful for an insulin delivery device.

Ito et al.^[27] produced a system with increased insulin permeability. In this work, poly(acrylic acid) chains were grafted on a porous cellulose membrane yielding a pH-sensitive membrane. GOD was immobilized onto the membrane. In the absence of glucose and at a neutral pH, electrostatic repulsion of the carboxyl groups resulted in poly(acrylic acid) grafts that were fully extended. The extended chains blocked the pores of the membrane. In the presence of glucose, GlucA was produced, protonating the PAA grafts. This resulted in coiling of the graft chains, because of decreased electrostatic repulsion. The permeability of insulin increased 1.7 times in the presence of glucose with a pH difference of only -0.4 pH units.

Goldraich and Kost^[25] investigated hydrogel matrices that were prepared by chemical polymerization of solutions containing 2-hydroxyethyl methacrylate, *N,N*-dimethyl-aminoethyl methacrylate, tetraethylene glycol dimethacrylate (TEGDMA), ethylene glycol, and water solutions of GOD or insulin. They found that hydrogels with high amine content and low cross-linking agent concentrations were the most sensitive to pH. As the pH decreased, the number of protonated amines on the gel increased. A higher charge density was possible with higher amine content on the polymer chain. The charges localized counterions in the vicinity of the gel. Thus, an increase in the osmotic pressure and swelling of the gel occurred with this high concentration of counterions. They also observed that hydrogels with immobilized GOD displayed faster and higher swelling and release

rates at higher glucose concentrations. Swelling and release kinetics were also responsive to step changes in glucose concentrations. Swelling and release rates increased after a step in glucose concentration from 0 to 500 mg% glucose. Some problems encountered with this system included slow release response to changes in glucose concentration, possible encapsulation after implantation, and low physiological concentrations of oxygen.

Heller et al.^[28] studied the development of a bioerodible insulin delivery device that releases insulin in response to the concentration of external glucose. The release of insulin from this device was modulated by pH changes resulting from the glucose–GOD reaction. They prepared a bioerodible polymer that could reversibly change erosion rates in response to small changes in the pH of the environment by incorporating tertiary amine groups into a linear poly(ortho ester). The incorporation of the tertiary amine groups in the polymer backbone led to increased polymer hydrophilicity and swelling because of repulsion of charged groups. Therefore, water penetration into the matrix was facilitated resulting in accelerated insulin release and polymer erosion. They found excellent control over release when insulin was homogeneously dispersed in the polymer. The polymer was observed to respond almost instantaneously to a decrease in pH, and insulin release was stopped as the pH increased. Their results suggested that a device could be designed to allow for extended, repetitive release before exhaustion as only a small fraction of available insulin was released in response to moderate pH pulses.

Ishihara, Kobayashi, and Shionohara^[29] investigated a glucose-responsive polymeric membrane that was composed of two polymeric membranes. One was a polyacrylamide membrane containing GOD, which acted as a sensor of glucose, forming hydrogen peroxide by an enzymatic reaction. The second membrane was a redox polymer with a nicotinamide moiety that served as a regulator of the insulin permeation rate through an oxidation reaction with the hydrogen peroxide. In the presence of glucose, the permeation rate of insulin was found to increase immediately. Problems with controlling the permeation rate of insulin reversibly were encountered with this system.

A self-regulating insulin delivery system based on the competitive and complementary binding behavior of concavalin A with glucose and glycosylated insulin was examined by Jeong and Kim.^[30] They developed an implant consisting of a regenerated cellulose membrane with the necessary diffusional properties for the exchange of glucose and glycosylated insulin. They found this type of system could approximate the functional response of the intact pancreas. Problems they encountered included inadequate insulin delivery to peripheral and portal blood supplies.

Of particular interest have been pH-sensitive graft copolymers of P(MAA-g-EG). These systems exhibit interpolymer complexation because of hydrogen bonding. Incorporation of these pH-sensitive materials with immobilized GOD yields a glucose-sensitive hydrogel that can be investigated for insulin-release applications.

The hydrogels of interest are graft copolymers of PMAA and PEG. In this system, interpolymer complexes form between PMAA chains and PEG grafts. The complexation occurs because of hydrogen bonding between hydrogens on the carboxylic acid groups of PMAA and the oxygens of the ether groups of the PEG chains. This complexation only forms at a pH low enough to allow for sufficient protonation of the carboxylic acid groups. When complexes form, the polymer is in a collapsed state because of the increased hydrophobicity in the polymer network. At high enough pH values, hydrogen bonds break as the carboxylic acid groups become ionized. As a result, the polymer swells to a high degree as electrostatic repulsion is produced within the network.^[4]

This hydrogel system is unique in that the two species that are involved in the complexation are bound together in the same polymer. This allows for the reversible formation of complexes at appropriate conditions. It is with this complexation behavior that these materials allow for pH-sensitive permeation of solutes, which is of great importance for the release of drugs.

Previous work based on the interpolymer complexation between PMAA and PEG was carried out by Osada and Takeuchi^[31] in which they examined the effects of treating PMAA with PEG. They observed dilations and contractions of the system owing to reversible complexation of PMAA with PEG. In the presence of a small amount of PEG, a PMAA membrane could contract and dilate many times to over 90% of its length. They found PMAA membranes to exhibit an increase in water permeation and flow rate when treated with PEG. When the membrane was rinsed with an alkali solution, the complex dissociated, PEG was removed, and the membrane recovered its initial low water permeability. Flow tests conducted with hemoglobin and albumin solutions showed that these proteins were capable of permeating the membrane without decreasing the water permeability or clogging the membrane. These problems were encountered with untreated PMAA membranes.

The effect of the PEG chain length used to treat PMAA membranes was investigated by Osada.^[32] It was found that PEGs with low molecular weight of 600 and 1000 produced rapid but small contractions and easily attained equilibrium. PEG of molecular weight 2000 yielded a rapid and pronounced contraction. PEGs of molecular weight greater than 2000 showed considerable contractions over a long period of time. This work indicated that the membrane contraction

was closely related to the chemical structure of the polymer and their chain length.

Klier, Scranton, and Peppas^[33] investigated the preparation and characterization of P(MAA-g-EG) networks. Self-associating networks were prepared by the copolymerization of methacrylic acid with poly(ethylene glycol) monomethacrylate (PEGMA) in the presence of the cross-linking agent and TEGDMA. They found the swelling of the networks to depend on swelling solution pH, swelling temperature, copolymer composition, and network structure. Nuclear Overhauser enhancement measurements indicated that graft copolymers of methacrylic acid and PEG formed complexes for a wider range of concentration and PEG molecular weights than the two ungrafted homopolymers. Copolymer networks were also found to swell to a lower extent than homopolymer networks owing to complex formation between PEG and PMAA segments.

Swelling Behavior of P(MAA-g-EG) Hydrogels

Previous work with P(MAA-g-EG) hydrogels has been carried out to characterize their swelling behavior as a function of pH. Bell and Peppas^[4] have done extensive work to characterize this swelling behavior by performing equilibrium swelling studies as a function of pH and oscillatory swelling studies as a function of time and pH.

Equilibrium studies indicated that at low pH values, the polymer network was collapsed and the swelling ratio was low. This behavior was observed up to about pH 4. In the region above pH 4 to about pH 8, the swelling ratio of the gels increased with increasing pH. In this range, the complexes broke, and the carboxylic acid groups of the PMAA were ionized. Network chains were driven apart by electrostatic repulsion and swelling resulted. However, above pH 8, the gels collapsed to an extent. This was because of the acid groups becoming neutralized by cations resulting in diminished electrostatic repulsion. From these equilibrium swelling studies, mesh sizes of the network were calculated as a function of pH. At low pH, the mesh sizes of the various samples were small, between 3 and 9 Å, whereas at high pH in the uncomplexed or expanded state, they were in the range of 240–350 Å. The mesh sizes between the complexed and uncomplexed states for all samples increased by 96–99%. These results showed that this material could possibly be used in controlled drug delivery and membrane separations.

In the oscillatory swelling studies, the swelling ratio of the gels was monitored as a function of time, as the pH of the surrounding solution was changed from acidic to basic over several cycles. In a basic solution, the gel swelled in a linear manner and did not approach equilibrium over the timescale of the experiment. When

transferred to an acidic solution, the gel collapsed abruptly. This swelling/collapsing behavior was repeated over four cycles. Mesh sizes of the networks were also calculated under oscillatory conditions. The mesh sizes responded rapidly to pH changes. Maximum mesh sizes in the expanded states ranged from 11 to 27 Å, whereas in the collapsed states they ranged from 4 to 9 Å.

The relaxation times of the polymer chains were also determined from the oscillatory studies. Average relaxation times were significantly greater during decomplexation in pH 10 than during complexation in pH 2. Relaxation times during the collapse of the gels owing to complexation were between five and eight times faster than those of expansion of the network because of decomplexation.

The permeability of these materials was also determined. It was found that the permeability of membranes of P(MAA-g-EG) was higher in the uncomplexed state than in the complexed state as the expanded, uncomplexed membranes contained larger mesh sizes available for solute permeation. The permeability of the membranes in the uncomplexed states was lower in membranes that contained longer PEG grafts. This was attributed to PEG grafts that were free to dangle in the mesh space of the polymer network in the uncomplexed state. The longer grafts interfered with solute diffusion through the network more than shorter grafts.^[4]

Glucose-Sensitive P(MAA-g-EG) Hydrogels

P(MAA-g-EG) hydrogels are of particular interest in this research because of their pH sensitivity. As described previously, glucose-sensitive hydrogels can be produced by the incorporation of a pH-sensitive hydrogel with immobilized GOD. A “squeezing gel” is expected when incorporating P(MAA-g-EG) with GOD, as shown in Fig. 1. At high concentrations of glucose, the GOD–glucose reaction produces GlucA, resulting in a decrease in the pH of the environment. The hydrogels are expected to collapse abruptly with the decrease in pH. It is with this collapse that insulin could be “squeezed out” as shown in Fig. 2. As the glucose concentration decreases by the action of the released insulin, less GlucA is produced, resulting in an increase in pH of the environment. The hydrogel is expected to swell with this pH change. Owing to the swelling/collapsing behavior of this system, insulin could be released in a pulsatile manner. A basal dose could be released from the gel in the swollen state at normal glucose concentrations, whereas an increased dose could be released as the gel collapses at high glucose concentrations. This mechanism of insulin release is a hypothesis to be tested in this research.

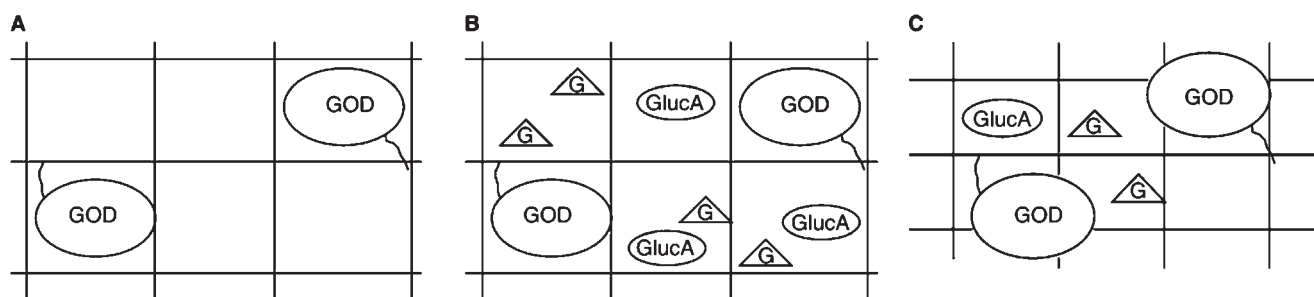


Fig. 1 Mechanism of action of glucose-sensitive P(MAA-g-EG) with GOD, glucose (G), and GlucA. (A) In the absence of glucose, at physiological pH, the hydrogel is in the swollen state. (B) In the presence of glucose, the GOD–glucose reaction produces GlucA causing a decrease in the pH. (C) The hydrogel collapses in response to the decrease in pH.

In previous work,^[34] graft copolymers of pH-sensitive P(MAA-g-EG) were investigated for use in such a hydrogel system. P(MAA-g-EG) hydrogels were synthesized and then characterized at a small physiological pH range. Complexation in this polymer network occurred by hydrogen bonding between the hydrogens of the carboxylic group of PMAA and the oxygens of the PEG chains. The collapse (complexation) of the hydrogels at pH 4.0 occurred much faster than the expansion (decomplexation) at pH 6.6. Glucose-sensitive hydrogels of P(MAA-g-EG) were synthesized by immobilizing GOD onto the polymer network. This was accomplished by first activating the enzyme, and then polymerizing P(MAA-g-EG) in the presence of the activated enzyme. It was determined that the GOD in the glucose-sensitive P(MAA-g-EG) hydrogels remained active throughout the activation and polymerization steps. The swelling behavior of the glucose-sensitive hydrogels was characterized as a function of time, pH, and glucose concentration. The glucose-sensitive hydrogels were shown to exhibit swelling and collapsing behavior for conditions varying by 1 pH unit, indicating that the hydrogel system has the capability of responding to small changes in glucose concentration. Swelling and collapsing behavior of glucose-sensitive P(MAA-g-EG) did not occur with varying glucose concentrations of 50 and 100 mg/dl, because the transition pH of the hydrogel was too high. Equilibrium swelling studies were carried out, and the transition pH was determined to occur between pH 4.3 and 4.9. However, by modifying some of the parameters of the P(MAA-g-EG) hydrogels, this transition pH could be shifted to a value that would yield desirable glucose-responsive behavior.

The specific goals of many of these newer devices are the development of new regimens for the treatment of diabetes in two specific cases:

1. A closed-loop feedback system that responds to changes in blood glucose concentration, the latter being based upon an implantable glucose sensor and permits adjustment of the rate of insulin release provided by the pump.

2. A Feedback system that responds to changes in blood glucose concentration and pH, leading to abrupt *swelling* and *deswelling* of a controlled drug delivery device based on a glucose- and pH-sensitive hydrogel that releases insulin at triggered intervals.

We have developed^[35–44] a series of novel self-regulated, glucose- and pH-sensitive gels for insulin delivery. In preparation for these systems, we have experimented with novel hydrogels in which the swelling ratio and the resulting mesh size change reversibly as a function of environmental parameters such as pH or temperature. These reversible changes allow for the release of drugs or the permeation of solutes depending on surrounding environmental conditions. PMAA exhibits interpolymer complexation with PEG, as the protons of the carboxylic acid groups on PMAA form hydrogen bonds with the ether groups on the PEG chain. This complexation forms only at a pH low enough to insure substantial protonation of the carboxylic acid groups. Complexation of free chains of PMAA with PEG in solution has been studied. We have also shown that poly(dimethyl aminoethyl methacrylate) (PDMAEM) and PEG [henceforth designated as P(DMAEM-g-EG) and shown in Fig. 2] exhibit the same type of hydrogen bonding, except that the pH dependence is such that the

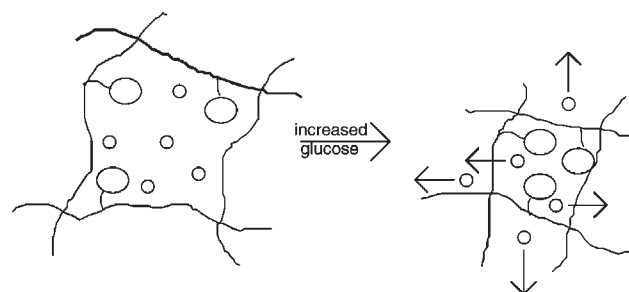


Fig. 2 Mechanism of insulin release from glucose-sensitive P(MAA-g-EG). With increased glucose concentration, complexation occurs resulting in a decrease in the mesh size and a “squeezing out” of insulin from the network.

systems decomplex at low pH and complex at high pH values.

In our early research, Doyle et al.^[45] have studied GOD-immobilized P(MAA-g-EG) hydrogels for pH sensitivity. In the hydrogels mentioned so far, insulin is released by diffusion. In this study, however, the hydrogels are expected to release insulin by a rapid squeezing effect. Oscillatory pH-swelling studies show a promising rapid collapse that can be expected to release insulin by squeezing. Also, a model-based design procedure has been suggested to obtain a hydrogel having the most desirable release characteristics. In another paper,^[34] the authors have characterized the equilibrium and the dynamic swelling behavior of these anionic-grafted hydrogels. The characteristic swelling behavior has been attributed to the formation of reversible complexes within the matrix.

Graft P(MAA-g-EG) and poly(dimethyl aminoethyl methacrylate-g-ethylene glycol) (P(DMEAEM-g-EG)) copolymers were synthesized by free-radical solution polymerization of MAA and dimethyl aminoethyl methacrylate (DMEAEM) purified by vacuum distillation, and PEGMA. TEGDMA was added as a cross-linking agent in the amount of 2 wt% monomer. The reaction mixture was diluted with a 50:50 ethanol:water mixture and a 50:50 mixture of ammonium persulfate, and sodium metabisulfite was used as a redox initiator in the amount of 0.025 wt% monomer. The reaction was carried out in polypropylene vials under nitrogen for 24 hr at 37°C. The cylindrical-shaped polymer samples were dried in air and cut into 0.5 mm thick disks using a diamond blade rotary saw. All disks were washed in distilled/deionized water for a week and dried in air. In addition, we have been able to carry out the same reaction under ultraviolet (UV) irradiation conditions at room temperature, using 0.5% dimethylphenyl acetophenone as a photoinitiator. The latter provides free radicals for the reaction, is partially incorporated in the growing chains but does not affect the swelling behavior, and is known to be a non-toxic initiator.^[34,40,41,46]

Hydrogels of the previous comonomers containing GOD have been prepared. GOD was first functionalized. The ensuing monomer was reacted with diethyl aminoethyl methacrylate (DMEAEM) and PEGMA under the same conditions as above. In the preliminary studies, we have concentrated only on copolymers consisting of 1:1 ratio of DEAEM and PEG monomethacrylate. Upon polymerization, such systems exhibit cationic behavior with hydrogen-bonding characteristics similar to those of P(MAA-g-EG), albeit with a reversal in behavior. Swelling studies were performed in 50 ml of constant pH solution at 37°C; the pH was adjusted by adding sodium hydroxide, hydrochloric acid, or sodium acetate.^[35,47,48] Dynamic and equilibrium swelling was determined gravimetrically

until an equilibrium value was reached. The swelling response of the gels under varying pH conditions was obtained by equilibrating samples in an acidic solution at 37°C and then placing them into a basic solution (still at 37°C) for 45 min following weight changes gravimetrically. The cycle was repeated several times.

To examine the pH sensitivity, the equilibrium swelling behavior of P(MAA-g-EG), P(DEAEM-g-EG), and GOD-containing hydrogels were investigated as a function of pH. The presence of hydrogen bonding in the complexes caused the network to be less hydrophilic because the carboxylic acid groups on the PMAA main chains were used in the complexes. P(MAA-g-EG) gels expand in acidic environment, but collapse at pH values higher than about 6.8. This swelling ratio changes from q of 1.2 (17% water) at pH = 7.2 to $q = 17.5$ at pH = 4.2 (95% water) indicates a change of the insulin diffusion coefficient by about two orders of magnitude. This means that the associated flux will be about 100 times higher at the pH values where GlucA is produced. On the other side P(DEAEM-g-EG) gels exhibit the opposite behavior, as shown in Figs. 3 and 4. Fig. 3 shows the shrinking behavior of glucose-oxidase-bound networks based on P(DEAEM-g-EG) as a function of pH. Clearly, as the pH decreases, because of the enzymatic reaction, there is a drastic swelling of the networks, allowing the release of the incorporated insulin. While this pH dependence is clearly a function of the pH of the ionizable unit (here DEAEM), the level of cross-linking, as defined by the nominal cross-linking ratio X (in moles of cross-linking agent per moles of all reactants) has also a significant effect on the swelling, as shown in Fig. 4. It

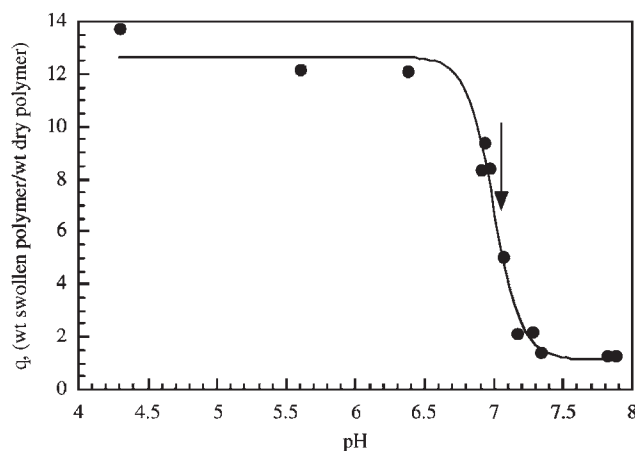


Fig. 3 Typical equilibrium-swelling characteristic obtained from poly(diethylaminoethyl methacrylate-g-ethylene glycol) hydrogels immobilized with GOD and catalase. The arrow indicated the location of the transition pH. This characteristic curve is specifically for a hydrogel sample containing $X = 0.015$, 50 mol% DEAEM to PEGMA 200 and 5.413×10^{-4} g of GOD/g of polymer.

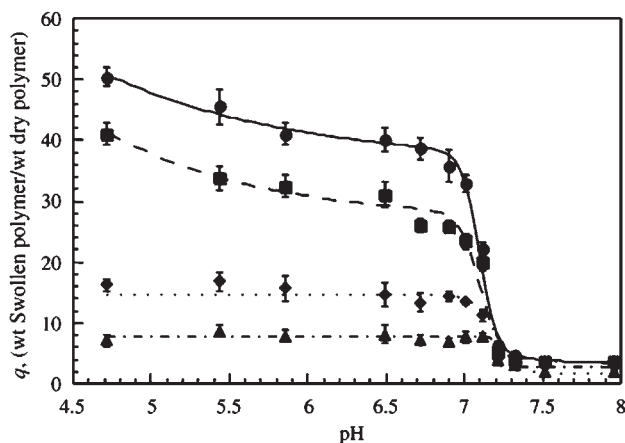


Fig. 4 Variation of equilibrium swelling characteristics based on the cross-linking ratio used in the formation of the network. The graphs shown are for $X=0.005$ (●), 0.01 (■), 0.02 (◆), and 0.04 (▲).

is also interesting to note that the mesh size of such hydrogels changes significantly as they expand or collapse at different pH values as shown in Fig. 5. When analyzing the behavior of glucose-sensitive hydrogels for their response to varying glucose concentrations,^[49–53] the GOD of the hydrogel reacts with glucose to produce GlucA. At high concentrations of glucose, the pH of the glucose solution may be decreased at about 3.0. But even with low concentrations of glucose, enough GlucA is produced to decrease the pH of the glucose solution to 3.0.

The glucose-sensitive hydrogels synthesized were tested to determine if the GOD remained active with the synthesis techniques used.^[54] In a disk of 3-mm thickness, a drop of 3 pH units was caused when placed in a 100-mg/dl solution of glucose for 15 min. Over this 15-min interval, the GOD of the polymer reacted with

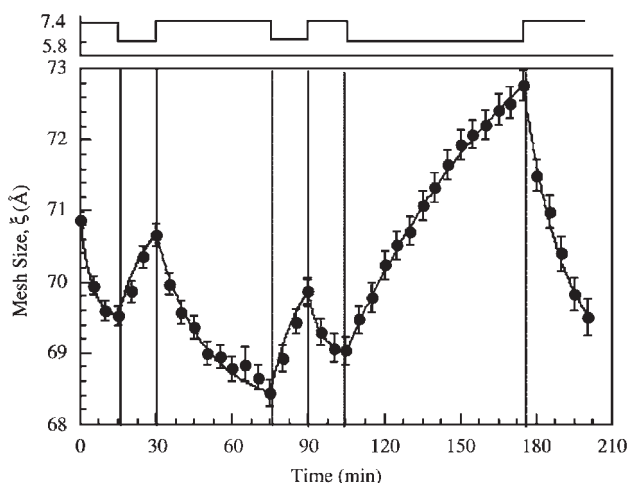


Fig. 5 Mesh sizes of P(DEAEM-g-EG) gel discs under varying pH conditions at 37°C are calculated.

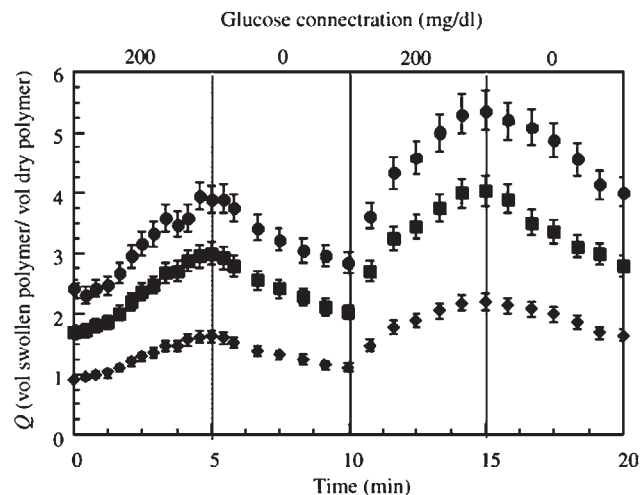


Fig. 6 Pulsatile swelling studies on P(DEAEM-g-EG) microparticles for changes in glucose concentrations in the swelling medium. The results have been shown for different nominal cross-linking ratio of the polymer, namely, $X=0.02$ (●), $X=0.03$ (■), and $X=0.04$ (◆).

the glucose in the surrounding solution to produce GlucA. Thus, GOD remained active throughout the activation of the enzyme and polymerization procedures. Typical results of this effect are shown in Fig. 6, which shows the effect of glucose concentration on the swelling and release.

We have shown that P(MAA-g-EG) and P(DEAEM-g-EG) hydrogels exhibited pH-sensitive swelling behavior owing to the formation and dissociation of complexes. This interpolymer complexation was because of hydrogen bonding between the hydrogels on the carboxylic acid groups of the MAA or the amino groups of DEAEM and the etheric groups on the PEG chains and resulted in pH-sensitive swelling behavior.

CONCLUSIONS

Glucose-sensitive systems based on glucose-responsive hydrogels have a series of unique characteristics that make them exceptional candidates for insulin delivery:

1. They respond fast and abruptly to glucose concentration.
2. They expand and contract abruptly because of pH changes.
3. They are mechanically strong.
4. They can be prepared in simple ways using UV polymerization and can contain large amounts of insulin dispersed uniformly in the gels.

Recent research has shown that these systems are excellent candidates for new glucose sensors, although they will require further studies for proof of



reproducibility of glucose response over hundreds and thousands of cycles, and ability of the carriers to release the same amount of insulin even under decreasing insulin concentration gradients.

Such promising systems provide also an excellent starting point for the study of a large generation of enzyme-sensitive sensors based on hydrogels.

ARTICLES OF FURTHER INTEREST

Biocatalysis; Biosensors

REFERENCES

- Kudela, V. Hydrogels. In *Encyclopedia of Polymer Science and Technology*; Mark, H.F., Kroschwitz, J.I., Eds.; Wiley-Interscience: New York, NY, 1985; Vol. 7, 783–807.
- Peppas, N.A. Physiologically responsive gels. *J. Bioact. Compat. Polym.* **1991**, *6*, 241–246.
- Heller, J. Modulated release from drug delivery devices. *Crit. Rev. Therap. Drug Carrier Syst.* **1993**, *10*, 253–305.
- Bell, C.L.; Peppas, N.A. Water, solute and protein diffusion in physiologically responsive hydrogels of poly(methacrylic acid-g-ethylene glycol). *Biomaterials* **1996**, *17*, 1203–1218.
- Flory, P.J. *Principles of Polymer Chemistry*; Cornell University Press: Ithaca, NY, 1953; 576–590.
- Katchalsky, A.; Lifson, S.; Eisenberg, H. Equation of swelling for polyelectrolyte gels. *J. Polym. Sci.* **1951**, *7*, 571–574.
- Oppermann, W. Swelling behavior and elastic properties of ionic hydrogels. In *Polyelectrolyte Gels: Properties, Preparation, and Applications*; Harland, R.S., Prud'homme, R.K., Eds.; American Chemical Society: Washington, DC, 1992; 159–170.
- Mandel, M. Polyelectrolytes. In *Encyclopedia of Polymer Science and Technology*; Mark, H.F., Kroschwitz, J.I., Eds.; Wiley-Interscience: New York, NY, 1985; Vol. 11, 739–829.
- Peppas, N.A.; Khare, A.R. Preparation structure, and diffusional behavior of hydrogels in controlled release. *Adv. Drug Del. Rev.* **1993**, *11*, 1–35.
- Gehrke, S.H.; Lee, P.I. Hydrogels for drug delivery systems. In *Specialized Drug Delivery Systems*; Tyle, P., Ed.; Marcel Dekker: New York, NY, 1990; 333–392.
- Michaeli, I.; Katchalsky, A. Potentiometric titration of polyelectrolyte gels. *J. Polym. Sci.* **1957**, *23*, 683–696.
- Siegal, R.A.; Firestone, B.A. pH-dependent equilibrium swelling properties of hydrophobic polyelectrolyte copolymer gels. *Macromolecules* **1988**, *21*, 3254–3529.
- Tanaka, T. Phase transitions in gels and a single polymer. *Polymer* **1979**, *20*, 1404–1412.
- Ishihara, K.; Matsui, K. Glucose responsive insulin release from polymer capsule. *J. Polym. Sci.: Polym. Lett. Ed.* **1986**, *24*, 413–417.
- Ishihara, K.; Kobayashi, M.; Ishimaru, N.; Shinohara, I. Glucose induced permeation control of insulin through a complex membrane consisting of immobilized glucose oxidase and a poly(amine). *Polymer J.* **1984**, *16*, 625–631.
- Peppas, N.A.; Langer, R. New challenges in biomaterials. *Science* **1994**, *263*, 1715–1720.
- Okano, T.; Yoshida, R. Intelligent polymeric materials for drug delivery. In *Biomedical Applications of Polymeric Materials*; CRC Press: Boca Raton, FL, 1993; 407–428.
- Jacobs, I.C.; Mason, N.S. Polymer delivery systems concepts. In *Polymeric Delivery Systems: Properties and Applications*; El-Nokaly, M.A., Piatt, D.M., Charpentier, B.A., Eds.; American Chemical Society, Washington, DC, 1993; Vol. 520, 1–5.
- Ashcroft, F.M.; Ashcroft, S.J.H., Eds.; *Insulin: Molecular Biology to Pathology; IRL*; Oxford: NY, New York, 1992.
- Guyton, A.C. *Textbook of Medical Physiology*; Saunders: Philadelphia, PA, 1971.
- Decoursey, R.M. *The Human Organism*; McGraw Hill: New York, NY, 1974.
- Wallum, B.J.; Kahn, S.E.; McCulloch, D.K.; Porte, D. Insulin secretion in the normal and diabetic human. In *International Textbook of Diabetes Mellitus*; Alberti, K.G.M.M., DeFronzo, R.A., Keen, H., Zimmet, P., Eds.; Wiley: New York, NY, 1992; 285–301.
- Krall, L.P.; Beaser, R.S. *Joslin Diabetes Manual*, 12th Ed.; Lea and Febiger: Philadelphia, PA, 1989.
- Kennedy, F.P. Recent developments in insulin delivery techniques: current status and future potential. *Drugs* **1991**, *42*, 213–227.
- Goldraich, M.; Kost, J. Glucose-sensitive polymeric matrices for controlled drug delivery. *Clinical Materials* **1993**, *13*, 135–142.
- Horbett, T.A.; Rattner, B.D.; Kost, J.; Singh, M. A bioresponsive membrane for insulin delivery. In *Recent Advances in Drug Delivery Systems*; Anderson, J.M., Kim, S.W., Eds.; Plenum: New York, NY, 1983; 209–220.
- Ito, Y.; Casolaro, M.; Kono, K.; Imanishi, Y. An insulin-releasing system that is responsive to glucose. *J. Control. Release* **1989**, *13*, 195–203.
- Heller, J.; Chang, A.C.; Rodd, G.; Grodsky, G.M. Release of insulin from pH-sensitive poly(ortho esters). *J. Control. Release* **1990**, *13*, 295–302.
- Ishihara, K.; Kobayashi, M.; Shionohara, I. Control of insulin permeation through a polymer membrane with responsive function for glucose. *Makromol. Chem. Rapid Commun.* **1983**, *4*, 327–331.
- Jeong, S.Y.; Kim, S.W. Self-regulating insulin delivery systems. III. In vivo studies. *J. Control. Release* **1985**, *2*, 143–152.
- Osada, Y.; Takeuchi, Y. Water and protein permeation through polymeric membrane having mechanochemically expanding and contracting pores. Function of chemical valve. *J. Polym. Sci.: Polym. Lett. Ed.* **1981**, *19*, 303–308.
- Osada, Y. Effects of polymers and their chain lengths on contraction of poly(methacrylic acid) network. *J. Polym. Sci.: Polym. Lett. Ed.* **1980**, *18*, 281–286.

33. Klier, J.; Scranton, A.B.; Peppas, N.A. Self-associating networks of poly(methacrylic acid-g-ethylene glycol). *Macromolecules* **1990**, *23*, 4944–4949.
34. Dorski, C.M.; Doyle, F.J.; Peppas, N.A. Glucose-responsive, complexation hydrogels. *Polym. Prepr.* **1996**, *37* (1), 475–476.
35. Podual, K.; Doyle, F.J., III; Peppas, N.A. Dynamic behavior of glucose-oxidase-containing microparticles of poly(ethylene glycol)-grafted cationic hydrogels in an environment of changing pH. *Biomaterials* **2000**, *21*, 1439–1450.
36. Langer, R.; Peppas, N.A. Advances in biomaterials, drug delivery, and bionanotechnology. *AIChE J.* **2003**, *49*, 2990–3006.
37. Foss, A.C.; Goto, T.; Morishita, M.; Peppas, N.A. Development of acrylic-based copolymers for oral insulin delivery. *Europ. J. Pharm. Biopharmac.* **2004**, *57*, 163–169.
38. Peppas, N.A. Kinetics of smart hydrogels. In *Reflexive Polymers and Hydrogels: Understanding and Designing Fast-Responsive Polymeric Systems*; Yui, N., Mrsny, R., Park, K., Eds.; CRC Press: Boca Rton, FL, 2004; 99–113.
39. Morishita, M.; Goto, T.; Peppas, N.A.; Joseph, J.I.; Torjman, M.C.; Munsick, C.; Nakamura, K.; Takayama, K.; Lowman, A.M. Mucosal insulin delivery systems based on complexation polymer hydrogels: effect of particle size on insulin enteral absorption. *J. Control. Release* **2004**, *97*, 115–124.
40. Peppas, N.A. Devices based on intelligent biopolymers for oral protein delivery. *Intern. J. Pharmac.* **2004**, *277*, 11–17.
41. Podual, K.; Doyle, F.J., III; Peppas, N.A. Glucose-sensitive cationic hydrogels: preparation characterization and modeling of swelling properties. In *Biomaterials Carriers for Drug Delivery, and Scaffolds for Tissue Engineering*; Peppas, N.A., Mooney, D.J., Mikos, A.G., Brannon-Peppas, L., Eds.; AIChE: New York, NY, 1997; 190–192.
42. Podual, K.; Doyle, F.J., III; Peppas, N.A. Preparation and characterization of the dynamic response of cationic copolymer hydrogels containing glucose oxidase. *Polymer* **2000**, *41*, 3975–3983.
43. Podual, K.; Doyle, F.J., III; Peppas, N.A. Glucose-sensitivity of glucose oxidase-containing cationic copolymer hydrogels having poly(ethylene glycol) grafts. *J. Control. Release* **2000**, *67*, 9–17.
44. Parker, R.S.; Rabinovitch, K.L.; Doyle, F.J., III; Peppas, N.A. Control analysis of pancreas models for optimal insulin delivery. In *Biomaterials Carriers for Drug Delivery, and Scaffolds For Tissue Engineering*; Peppas, N.A., Mooney, D.J., Mikos, A.G., Brannon-Peppas, L., Eds.; Aiche: New York, NY, 1997; 319–321.
45. Doyle, F.J., III; Dorski, C.M.; Harting, J.E.; Peppas, N.A. Control and modeling of drug delivery devices for treatment of diabetes. *Proc. Am. Cont. Conf.* **1995**, *1*, 776–780.
46. Podual, K.; Peppas, N.A.; Doyle, F.J., III. Insulin release from pH-Sensitive cationic hydrogels. *Proceed. Intern. Symp. Control. Rel. Bioact. Mater.* **1998**, *25*, 56–57.
47. Schwarte, L.M.; Peppas, N.A. Preparation and characterization of peg-containing pH-sensitive, cationic hydrogels for drug delivery applications. *Polym. Prepr.* **1997**, *38* (2), 596–597.
48. Parker, R.S.; Doyle, F.J., III; Peppas, N.A. A model-based algorithm for blood glucose control in type 1 diabetic patients. *Ieee Trans. Biomed. Eng.* **1999**, *46*, 148–157.
49. Peppas, N.A. Is there a future in glucose-sensitive, responsive insulin delivery. *J. Drug Deliv. Sci. Techn.* **2004**, *14*, 247–256.
50. Schwarte, L.M.; Podual, K.; Peppas, N.A. Cationic hydrogels for controlled release of proteins and other macromolecules. In *Tailored Polymeric Materials for Controlled Delivery Systems ACS Symposium Series*; ACS: Washington, DC, 1999; 709, 56–66.
51. Parker, R.S.; Ward, J.H.; Doyle, F.J., III.; Peppas, N.A. Robust control in diabetes using a physiological model. *Aiche J.* **2000**, *46*, 2537–2549.
52. Dorski, C.M.; Doyle, F.J., III; Peppas, N.A. Preparation and characterization of glucose-sensitive P(MAA-g-EG) hydrogels. *Polym. Mater. Sci. Eng. Proceed.* **1997**, *76*, 281–282.
53. Hassan, C.M.; Doyle, F.J., III; Peppas, N.A. Dynamic behavior of glucose-responsive poly(methacrylic acid-g-ethylene glycol) hydrogels. *Macromolecules* **1997**, *30*, 6166–6173.
54. Schwarte, L.M.; Peppas, N.A. Novel poly(ethylene glycol)-grafted cationic hydrogels: preparation, characterization and diffusive properties. *Polymer* **1998**, *39*, 6057–6066.



Glucose Sensors

Jeffrey I. Joseph

Marc J. Torjman

The Artificial Pancreas Center, Department of Anesthesiology, Jefferson Medical College of Thomas Jefferson University, Philadelphia, Pennsylvania, U.S.A.

INTRODUCTION

Proper function of an implantable glucose sensor requires fundamental understanding and control of the biological processes that occur at the sensor-tissue interface.^[1] Scientists have only recently begun to unravel the complex biological interactions that occur between the biomaterial surface and the tissues.^[2,3] Systematic studies are currently underway to optimize sensor size, geometry, surface structure, surface biochemistry, method of transduction, and method of implantation for a specific sensor in a specific location of implantation. Each type of invasive and minimally invasive glucose sensor requires a unique set of physical and biochemical properties to withstand the rigors of the *in vivo* environment. This article will summarize the design and biomaterial issues most important for optimal function of the following glucose sensors.

1. Needle-type, enzyme-based electrochemical sensor (biocatalytic): for short-term implantation in the dermis or subcutaneous tissue.
2. Dialysis catheter ISF glucose monitoring system with an external glucose sensor (electrochemical with enzyme, pressure/viscosity): for short-term implantation in the subcutaneous tissue.
3. Enzyme-based electrochemical sensor (biocatalytic) with telemetry: for long-term implantation in the subcutaneous tissue.
4. Affinity chemical/optical sensor (using concavalin A): for long-term implantation in the subcutaneous tissue.
5. Enzyme-based electrochemical/oxygen sensor: for long-term implantation in the bloodstream (intravascular catheter).
6. Near-infrared absorption spectroscopy sensor: for long-term implantation around a vascular tissue (perivascular sensor head).

CLINICAL SIGNIFICANCE

Diabetic patients treated with intensive insulin therapy and near-normal blood glucose (BG) control live

longer and healthier lives. Long-term tight BG control is achieved in only a select group of patients that modify their insulin doses according to meal size, meal composition, and six or more finger-stick BG measurements per day. Unfortunately, the great majority of diabetics using insulin do not self-monitor frequently enough to achieve tight BG control, without significantly increasing their risk for hypoglycemia.^[4] Confusion, seizure, coma, and death can occur when there is a mismatch between the delivered dose of insulin and the immediate metabolic needs of the patient.

Safe diabetes management demands that the sensor output signal be an accurate estimate of the actual glucose level in the blood or interstitial fluid (ISF). Implantable sensors have the potential to measure glucose on a near-continuous basis (1440 measurements per day). Real-time analysis of glucose data can provide the patient with key information, such as the rate of glucose rise or fall in relation to meals, exercise, sleep, and illness (Fig. 1). Trend analysis can provide an early warning of hypoglycemia, well before the onset of clinical symptoms. Patients are able to more precisely titrate the appropriate dose of insulin using frequent glucose data. Complications due to chronic hyperglycemia will be prevented. Glucose sensors must be accurate, robust, easy to calibrate, and patient-friendly to be useful in the clinical setting. An array of sensors can be used to improve accuracy and reliability.

The current goal of sensor research is the clinical application of an automated device that can measure and display glucose values in real time, with alarms for hypoglycemia. The ultimate goal of sensor research is an automated device often called the artificial pancreas (AP) that integrates a glucose sensor, an insulin pump, and a computer controller (Fig. 2). Insulin delivery can be adjusted on a minute-by-minute basis according to the immediate metabolic needs of the diabetic patient. Blood and ISF glucose AP monitoring systems are able to precisely control the delivery of intravenous or subcutaneous tissue insulin (CSII: continuous subcutaneous insulin infusion therapy), with improved short-term BG control.^[5] The feasibility of continuous glucose monitoring with closed-loop insulin delivery has been demonstrated in ambulatory and hospitalized humans.

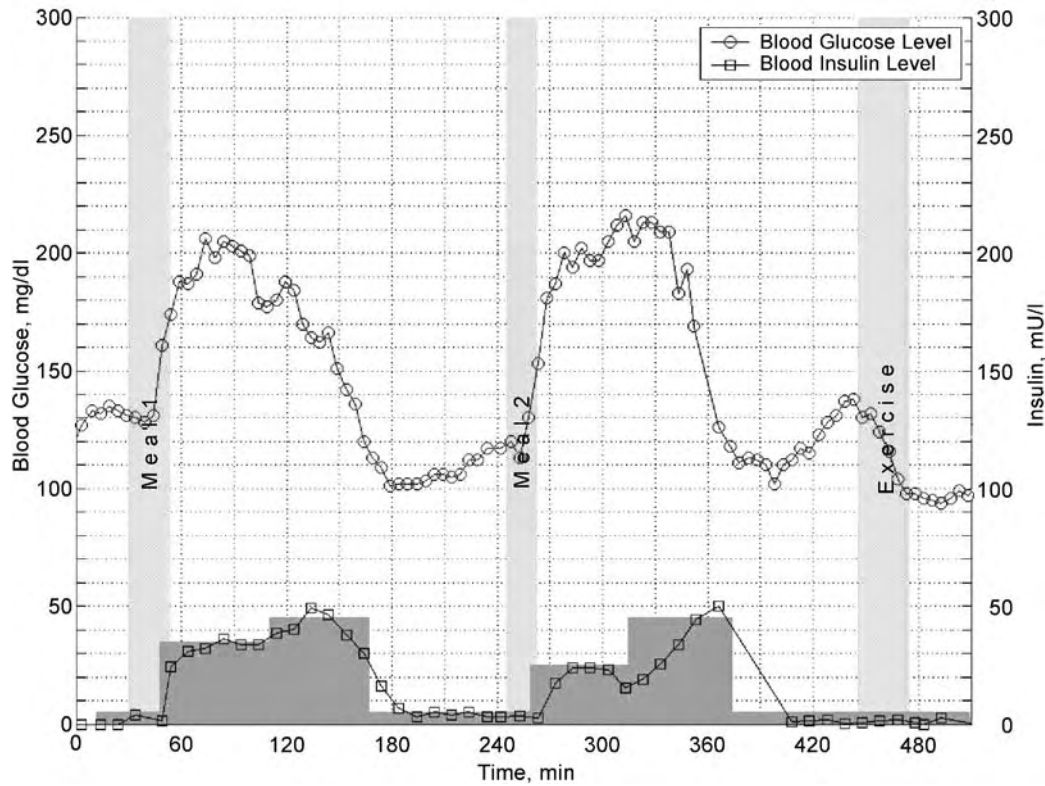


Fig. 1 Real-time display of frequent BG data (mg/dl), demonstrating the rate of glucose rise and fall in relation to meals and exercise, measured using the Metracor VIA-GLU enzyme-based electrochemical sensor.



**INTERSTITIAL FLUID (ISF)
GLUCOSE SENSORS**

Needle-Type, Enzyme-Based Electrochemical Sensor (Biocatalytic) for Short-Term Implantation in the Dermis or Subcutaneous Tissue

The glucose concentration within the ISF of the dermis and subcutaneous tissue generally parallels the glucose concentration within the blood.^[6] The most advanced needle-type sensors use an enzyme, selective membranes, and electrodes to measure glucose in the ISF.^[7] The patient can easily self-insert the sensing device through the skin with minimal discomfort. Depending upon the length of the sensor, the distal tip resides within either the dermis, the loose connective tissue just below the dermis, or the adipose tissue. The sensor's proximal portion is connected to a power source and electronics module with telemetry. The absolute glucose value and the glucose value trends can be visualized on a wristwatch or belt-worn display (Fig. 3).

Needle-type sensors have functioned well within the subcutaneous tissue for weeks to months. The Food and Drug Administration (FDA), however, requires

that percutaneous sensors be removed every three days, to minimize the risk of infection. More than 100 sensors are therefore required in order to provide continuous ISF glucose monitoring for a period of one year.

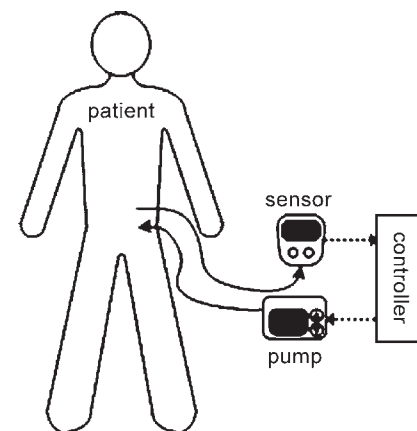


Fig. 2 The artificial pancreas (AP) combines a glucose sensor, an insulin infusion pump, and a computer controller to precisely control the delivery of insulin, based upon the absolute glucose concentration and the rate of rise and fall over time.



Fig. 3 Needle-type glucose sensor with arterioles, capillaries, and venules of subcutaneous tissue. The pictured sensor is manufactured by Medtronic-MiniMed as the Continuous Glucose Monitoring System-CGMS. (Used with permission from Medtronic-MiniMed Corporation, www.medtronic-minimed.com.)

Typical needle-type sensors use semipermeable membranes and an enzyme (glucose oxidase or glucose dehydrogenase) to provide specificity and selectivity in the detection and measurement of β -D-glucose. Sensors tested *in vitro* with glucose solutions demonstrate high sensitivity (change in current per change in glucose concentration), linearity over the physiological range (40 to 600 mg/dl), specificity for glucose, and stability over time. When implanted in the body, however, there is an immediate loss of stability characterized by a significant loss of sensitivity, followed by a period of relative stability, followed by further loss of sensitivity.^[8]

Tissues react to the trauma of sensor implantation and the presence of foreign biomaterial surfaces by activating the coagulation and immune systems.^[9] Platelets, red blood cells, fibrous proteins, and tissue fluids coat the surface and fill the pore structure of the implanted sensor's outer membrane. Neutrophils and macrophages engulf bacteria, injured cells, damaged ECM, and sensor debris. Fibroblasts and macrophages manufacture new ECM in response to the healing wound and the chemical/physical characteristics of the adjacent biomaterial surface.^[10] A foreign body capsule (FBC) develops, initially as a thin layer of fibrous protein (10 to 50 microns) that can severely slow the inward diffusion of glucose. The thin layer of exudative fluid, commonly located between the FBC and the sensor surface, contains dead cells, white blood cells, and numerous products of inflammation. Sensor function requires that glucose molecules be able to freely diffuse from capillary blood through the ISF, though the FBC, through the inflammatory exudate, and through one or more porous membranes to interact with the sensor's transducer.^[11] Biological changes within the membrane pore structure are collectively known as biofouling.^[1]

Percutaneous insertion of a needle-type sensor damages cells, blood vessels, and ECM. Red blood

cells, platelets, plasma proteins, and cellular debris mix with tissue fluid and ECM proteins to form the environment of the fresh wound.^[12] Cells and proteins (fibrinogen, albumin, vitronectin, and fibronectin) adhere to the surface of the sensor's outer membrane, eventually filling the voids of the interconnecting pore structure. Activated platelets release cytokines and growth factors that are chemotactic for neutrophils and macrophages. Neutrophils arrive by the thousands to remove bacteria and cellular debris. Activated monocytes enter the wound and change phenotype to migratory tissue macrophages—cells that are programmed to remove cellular/ECM debris (by phagocytosis and proteolysis) and produce/secrete new ECM. Macrophages release cytokines and growth factors that promote the inward migration and growth of new vascular tissue, containing functional arterials, capillaries, and venules.^[13]

The process of angiogenesis starts immediately with activation of the coagulation and immune systems. The new wound develops a functional vascular system over three to six days.^[14] Prior to new vessel ingrowth, the environment immediately adjacent to the sensor may lack essential nutrients (oxygen, glucose, amino acids, fatty acids, co-enzymes), and may have an excess amount of cellular waste products (hydrogen ions, lactic acid, urea nitrogen, water). Neutrophils leave the region, and by day three the wound is filled with macrophages, fibroblasts, organizing ECM, and the start of a vascular network. The new tissue has a high metabolic rate due to the large number of dividing, growing, moving, synthesizing, secreting, and digesting cells. Control of this complex process depends upon the regulation of cell membrane receptors (RGD and similar ligands) and the chemical composition of the newly formed ECM, recently secreted by the macrophages and fibroblasts. The new ECM develops into a three-dimensional scaffold or gel—consisting of collagen, fibronectin, vitronectin, and other compounds and optimized by cytokines, growth factors, and cell-to-cell communication—for the regeneration of new healthy tissue, rather than the formation of fibrous scar.^[10,12] Even the most biocompatible materials develop a thin layer of fibrous tissue (5 to 10 microns) within 24 hours of implantation.^[15] Sensor movement promotes fibrous protein layer thickening.^[16] The fibrous connective tissue of the developing FBC hinders glucose diffusion to the sensing region, altering sensitivity and prolonging time constants.

Figure 4 illustrates a typical needle-type glucose sensor with an outer membrane, an enzyme layer, an inner membrane, an electrolyte solution, and insulated metal electrodes. The sensor's outer membrane is designed for three functions: 1) to interface with the fluid, cells, and ECM of the interstitial tissue; 2) to slow the diffusion of glucose into the underlying enzyme layer

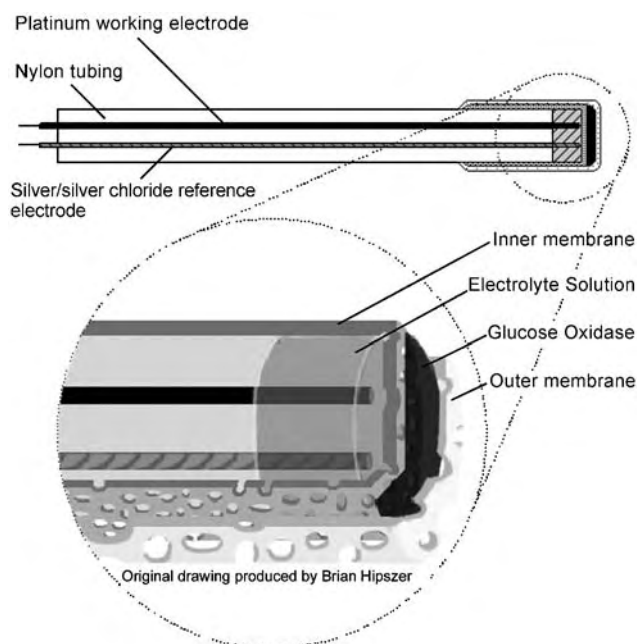
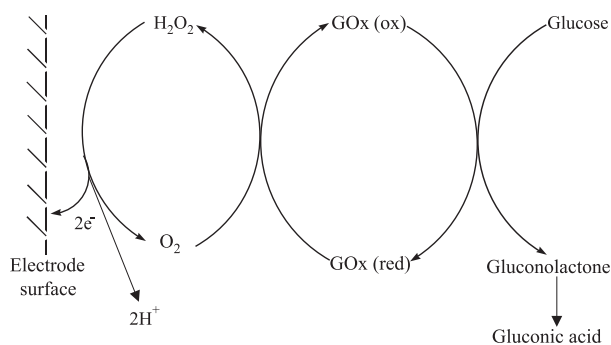


Fig. 4 Typical enzyme-based electrochemical glucose sensor with an outer membrane, an enzyme layer, an inner membrane, an electrolyte solution, and insulated metal electrodes. The basic design applies to all types of sensors that use an enzyme. (Original drawing by Brian Hipszer.)

to a rate much slower than the rate of the enzyme-catalyzed reaction, ensuring that oxygen will always reach the enzyme in excess of glucose (minimizing the effect of tissue hypoxia); and 3) to protect the underlying glucose oxidase enzyme from leaching, degradation, and fouling.

Glucose molecules diffuse through the outer membrane into the gel-like enzyme layer, to bind with the active site of the glucose oxidase enzyme. The following series of redox reactions occur within the enzyme layer: 1) oxidation of β -D-glucose by glucose oxidase to form gluconolactone; 2) reduction of gluconolactone to gluconic acid; 3) reduction of glucose oxidase enzyme; and 4) reoxidation of glucose oxidase by the reduction of oxygen to hydrogen peroxide:



Hydrogen peroxide and oxygen exit the enzyme layer through an inner membrane, to reach the platinum and silver/silver-chloride electrodes. The inner membrane is designed for two functions: 1) to protect the underlying electrodes from fouling and degradation; and 2) to inhibit the inward diffusion of chemical analytes other than hydrogen peroxide (ascorbic acid, uric acid, acetaminophen, and glutathione) that may be oxidized at the platinum electrode. Small-pore membranes made from cellulose acetate, PTFE, silicone, or polyethylene perform well for these functions. Hydrogen peroxide is oxidized at the platinum electrode, which is biased at +700 mV relative to the silver/silver-chloride reference electrode: producing electrons, protons, and oxygen. Two electrons are formed for each molecule of hydrogen peroxide oxidized. An electronic circuit is used to measure a current or voltage. Electrons reaching the reference electrode combine with oxygen and protons to produce water.^[11] The sensor's output signal is converted to a glucose value by correlating the current or voltage to one or more reference BG measurements.^[9]

Alternate embodiments of the needle-type sensors use a mediator such as ferrocene, or a ferrocene derivative, to transfer electrons directly to the working electrode, eliminating the need for oxygen as an electron acceptor.^[17] Catalase has been added to the enzyme layer to prevent the toxic accumulation of hydrogen peroxide.^[18]

Prior to *in vivo* testing, sensors should be rugged enough to function long-term in physiologic solutions at body temperature. Hardware issues leading to premature sensor failure include outer membrane delamination, loss of hermetic seal with corrosion (short-circuit), lead detachment, and enzyme degradation following sterilization.^[11] Biologic issues leading to sensor malfunction and failure include membrane fouling/degradation, enzyme fouling/degradation, and electrode fouling (passivation).^[1]

Outer membranes manufactured from porous polyurethane, cellulose acetate, Nafion, PHEMA (hydrogel), albumin, collagen, fibronectin, silicone, polyethylene, PTFE, ePTFE, Dacron, epoxy, alginate/polylysine, and Parylene demonstrate adequate biocompatibility, but do poorly at providing a hermetic or biologic seal.^[11] Enzyme fouling can be caused by toxic molecules such as oxygen radicals, activated complement, protein degradation products, heavy metals, and hydrogen peroxide. Membrane and enzyme biofouling are reversible. Explanted sensors regain near-normal function when soaked in saline. Macrophages have the capacity to secrete more than a dozen different compounds capable of damaging the glucose oxidase enzyme. For example, macrophages are known to release potent oxidative species from the enzyme myeloperoxidase, which produces hypochlorite (bleach) that can

destroy the glucose oxidase enzyme activity and cause rapid sensor failure. Electrode fouling (passivation) occurs when small molecules and ions pass through the inner membrane to coat the electrode's metal surface.

The biological response to sensor implantation is often complex and variable. Drift in sensitivity requires periodic recalibration of the sensor's output signal to a reference BG measurement.^[8] Very frequent recalibration provides a close correlation of the sensor signal with BG measurements. The clinical success of any sensor, however, requires a low frequency of recalibration, estimated to be four finger-stick blood samples or fewer in 24 hours. This type of recalibration is feasible if sensor drift occurs slowly and consistently. Unpredictable *in vivo* sensor performance cannot be compensated for with intermittent blood samples and recalibration.^[11]

Unique biomaterials are being developed to improve sensor stability and longevity. The wound surrounding traditional membranes develops a classic inflammatory reaction that leads to FBC formation. Sensors typically fail 14 to 30 days after implantation when vascularized granulation tissue is replaced with fibrous connective tissue. Methods to prevent inflammation and FBC formation have been explored. Membranes with an interconnecting pore structure provide a large surface area (scaffold) for cell and ECM adhesion. Pore size and porosity may greatly influence blood vessel density, the location of vessels within the three-dimensional scaffold, and long-term stability of the vascular network. Large-pore scaffolds (greater than 8 microns) permit deep penetration of capillaries and supporting ECM. Membranes with 1 to 3 micron pore size tend to develop a FBC that is thinner, less dense, and more vascular compared to capsules surrounding smooth membranes.^[19] Implanted membranes with greater than 0.8 micron pore size have remained vascular for more than one year in the subcutaneous tissues of rodents.^[20]

Although studies reveal that thinner fibrous capsules allow faster diffusion of small analytes, capsules as thin as 10 microns have been shown to significantly slow the diffusion of glucose.^[15] Membranes made from soft hydrophilic hydrogel polymers (such as PHEMA) produce the least inflammation and thinnest FBC. Scaffolds can be functionalized with receptor ligands (RGD and others) to enhance cell and ECM adhesion.^[21] Steroid and nonsteroidal compounds (dexamethasone, indocine, and cyclosporin) that suppress white blood cell (immune) function have been delivered directly to the sensor-tissue interface in an attempt to improve sensor stability.^[22] Drugs that inhibit the ability of macrophages and fibroblasts to synthesize and secrete collagen have great potential to prevent FBC formation. Growth factors that promote angiogenesis have been incorporated into membranes to promote and maintain a high capillary

density. Degradable microspheres have been developed that release anti-inflammatory drugs and growth factors into the porous membrane at the appropriate time interval (controlled release) following implantation.^[22] This area of research is exciting and clinically relevant.

Dialysis Catheter ISF Glucose Monitoring System with External Glucose Sensor

Microdialysis catheters (probes) can be inserted into the subcutaneous tissue to obtain continuous samples of dialysate containing ISF glucose. A glucose-free physiologic solution (dialysate) is pumped from an external reservoir into the dialysis probe. Glucose molecules pass from the adjacent tissue fluids, through the dialysis membrane pores, into the glucose-free dialysate. If pressure gradients are well balanced at the catheter-tissue interface, glucose will move into (and out of) the porous catheter by simple diffusion. Microdialysis techniques do not measure glucose directly in the ISF environment, in contrast to needle-type sensors. The concentration of glucose in the dialysate and ISF will be identical when complete equilibrium across the dialysis membrane has been reached. Dialysate is continuously pumped into (and/or suctioned out of) the dialysis probe through tubing, to an external (flow-through) glucose sensor. Dialysate waste is stored in an external reservoir (Fig. 5). Three highly sensitive and novel glucose-sensing mechanisms have been developed.^[23]

1. Small volumes of dialysate are continuously delivered to an external enzyme-based electrochemical sensor with a flow-through sample reservoir. Issues related to membrane/enzyme biofouling and degradation are minimized, because the dialysate fluid contains only small molecules and does not contain white blood cells, degraded ECM, and many of the toxic compounds of inflammation. As a result, continuous glucose sensors based upon microdialysis remain relatively stable over a 24-hour period. Sensor drift does occur due to ongoing inflammation and wound healing at the catheter-tissue interface, and to biofouling of the glucose sensor's outer membrane, enzyme, and electrodes. Water-soluble compounds with molecules as small as those of glucose, lactate, glycerol, and amino acids can easily pass through the 20 kDa pore size dialysis membrane to reach the glucose sensor. Toxic products of inflammation (oxygen radicals, ions, complement) can still reach the sensor to adversely affect membrane and enzyme performance. This type of sensor is being developed by A. Menarini Diagnostics (Florence, Italy, <http://www.menarini.com>).



Fig. 5 Dialysis catheter ISF glucose monitoring system with an external enzyme-based electrochemical glucose sensor, designed for short-term implantation in the subcutaneous tissue. Note percutaneous insertion and the large size of the dialysis catheter and monitoring system. (Used with permission of Disetronic Medical, www.disetronic.com.)

2. The preceding system can be modified to continuously infuse a solution of glucose oxidase (GO) enzyme into the outflow channel of the dialysis probe. The external glucose sensor does not require an enzyme layer because the incoming dialysate has already been exposed to the enzyme's active site. This technique improves sensor stability by always supplying an excess amount of fresh GO enzyme. The biocompatibility issues and limitations discussed for the preceding system apply. This type of sensor is being developed by Roche Diagnostics GmbH (Mannheim, Germany, <http://www.roche.com>).
3. This microdialysis-based system uses a novel dialysate solution containing a mixture of high-molecular-weight dextran and concavalin A, an affinity receptor with specificity for β -D-glucose. Molecules of ISF glucose diffuse through the porous membrane to interact with the dialysate solution. The viscosity of this solution can rapidly increase or decrease in direct response to a change in the glucose concentration. Glucose-sensitive dialysate is continuously pumped into (and out of) the dialysis probe at a very slow rate (5 μ l/min) to allow complete equilibrium with ISF glucose. Pressure sensors are located at the inflow and outflow channels to continuously measure the pressure differential across the dialysis probe. The change in viscosity can be strongly correlated to a change in the glucose concentration of ISF fluid. Although this type of system provides superior stability and rapid response, there is a risk that dextran and concavalin A will leach out of the

dialysis probe to contaminate the subcutaneous tissue. Both compounds can initiate an allergic reaction, and concavalin A is considered a carcinogen. This type of sensor is being developed by Disetronic Medical (Burgdorf, Switzerland, <http://www.disetronic.com> or <http://www.adicol.org>).

Incomplete equilibrium occurs when the dialysate does not have sufficient time to equilibrate fully with ISF glucose. This may occur when membrane biofouling and FBC formation slow the rate of glucose diffusion into the dialysate. Microliter volumes of dialysate are pumped into and out of the catheter at specific flow rates (0.3–10 μ l/min). Slow rates of perfusion increase the relative recovery of ISF glucose, defined as the concentration of glucose in the outgoing dialysate compared to the concentration of glucose in the fluid surrounding the probe. Relative recovery from ISF is constant, independent of changes in ISF glucose concentration, as long as the perfusion conditions remain unchanged.^[23] Fortunately, many of the factors that influence *in vivo* glucose recovery (membrane surface area/volume, pore size, pore density, surface chemistry, ECM chemistry, dialysate chemistry, flow rate of dialysate, pressure gradients, temperature) remain relatively stable over short periods of time.

Subcutaneous tissue reacts to insertion of the dialysis probe with acute inflammation, wound repair, and FBC formation: similar to that found at the membrane-tissue interface of needle-type sensors.^[2] A break-in or waiting period, ranging from 2 to 12 hours, is required before initial sensor calibration and the onset of clinical monitoring. This break-in period is characterized by variable glucose recovery, most probably due to biofouling of the probe's pore structure with platelets, red blood cells, and fibrin. Remarkably, the processes of acute inflammation, wound healing, and FBC formation at the catheter-tissue interface affect short-term sensor stability only slightly. Studies have reported that calibration to a reference BG measurement is typically required no more often than once every 24 hours.^[23]

Dialysis catheter ISF glucose monitoring systems are limited by the following factors: 1) medical personnel are required for dialysis probe insertion; 2) the physical lag time is long (approximately 31 minutes); 3) the dialysis probe is large and painful to insert; 4) dialysis probes implanted percutaneously must be removed every three to four days to prevent infection (FDA requirement); 5) 90 to 100 dialysis probes are needed to continuously monitor ISF glucose levels for one year; 6) probes and sensors are expensive; 7) the belt-worn sensor/pump/electronics module is large and heavy; 8) the risk of infection, allergic reaction, and cancer may be increased; 9) progressive biofouling of

the dialysis membrane and GO enzyme leads to sensor instability; and 10) sensors eventually fail when the vascular granulation tissue that surrounds the probe is replaced with avascular fibrous tissue (scar).

Enzyme-Based Electrochemical Sensor (Biocatalytic) with Telemetry for Long-Term Implantation in the Subcutaneous Tissue

The ISF glucose sensor under development by Dexcon Corporation (San Diego, California, <http://www.dexcominc.com>) has been given long-term testing in ambulatory patients with type I diabetes. The sensor has been miniaturized into the shape of a rounded cylinder with an enzyme-based electrochemical sensor at the sensing region. The small, implantable package contains the enzyme sensor, electronic circuitry, power source, antennae, and telemetry. A belt-worn electronics module receives and displays sensor data in real time. Key to the clinical success of this long-term implantable sensor is a precisely engineered sensor-tissue interface that remains highly vascular for the life of the sensor. The vascular FBC has been able to provide ISF to the enzyme surface for more than 11 months with rapid kinetics (5–10 minute delay) and stable sensitivity, after an initial time period for capsule formation (15–30 days). After formation of a stable vascular FBC, external calibration can be extended to monthly intervals. The miniature sensor has been designed with sufficient enzyme capacity and low power requirements to function longer than 24 months *in vivo*. Below the angiogenic scaffold is a bioprotective membrane layer, designed to keep macrophages and their corrosive compounds away from the underlying urethane membrane and enzyme. The membrane's thickness and pore size have been designed to supply oxygen in excess of glucose (30:1 ratio of oxygen to glucose molecules) to the underlying glucose oxidase enzyme.^[24]

The challenge for scientists is to maintain functional capillaries for the long term in the region adjacent to the sensor's outer membrane. This sensor is subject to the same immune and wound-healing response following subcutaneous tissue implantation as other sensor technologies. Under the leadership of Jim Brauker, Ph.D., Dexcon Corporation has optimized membrane structure and chemistry (pore size, pore density, pore connectivity; membrane surface area, thickness, texture, and chemistry) to produce a FBC that maintains long-term vascularity. Studies in animals have shown that following subcutaneous tissue implantation, vascular granulation tissue completely fills the pore structure of large-pore membranes, including the boundary region between laminated large- and small-pore membranes. Porous membranes tested

in animals have maintained a functional capillary network for more than one year.^[20] Sensors constructed with an outer covering of Dacron velour (90-micron pore size) develop a well-vascularized FBC that can supply adequate glucose and oxygen to the sensing region for many months. The sensor can provide linear glucose tracking over the 35- to 800-mg/dl range.^[24] Some scientists hypothesize that an ongoing inflammatory process is required to maintain a long-term vascular FBC. Dacron is known to induce chronic inflammation, characterized by a persistent population of macrophages and the formation of multinucleated giant cells. Human trials to evaluate the safety and efficacy of the implantable glucose sensor are under way. The fact that the sensor-tissue interface of diabetic patients with poor glycemic control may heal differently than that of nondiabetics, or of diabetics with good BG control, complicates the process of sensor development.^[5] Sensors eventually fail when the vascular granulation tissue that surrounds the sensing region is replaced with avascular fibrous tissue.

Affinity Chemical/Optical Sensor for Long-Term Implantation in Subcutaneous Tissue

A novel ISF glucose sensor has been developed by BioHybrid Technologies (Shrewsbury, Massachusetts, <http://www.bio.org/memberProfile/biohybrid.html>) that is based on changing fluorescence within a hydrogel bead (1.5 × 4 mm) implanted chronically in the subcutaneous tissue. Glucose is able to diffuse from ISF into the bead, to compete for glucose binding sites on donor-labeled glycosylated substrate and acceptor-labeled glucose binding ligand. Glucose competes for the glucose binding sites and alters the nonradiative fluorescence resonance energy transfer (FRET) efficiency. An external hand-held laser is used to illuminate the subcutaneous sensor (hydrogel bead) with visible light. Energy reflected from the implanted sensor is detected with a CCD array and analyzed to determine the degree of energy transfer. The sensor has been able to track changes in ISF glucose for more than two months. The long-term biological issues discussed for the enzyme-based sensors apply to this technology as well. Although the PHEMA hydrogel bead causes minimal inflammation at the sensor-tissue interface, biofouling and FBC formation can lead to premature sensor failure. Slowed glucose diffusion into the hydrogel further worsens the time delay due to slow glucose binding/unbinding within the hydrogel.^[25] There is an additional risk of contaminating the subcutaneous tissue with toxic compounds, such as concavalin A, if there is loss of capsule integrity (degradation or breakage due to mechanical factors or macrophage actions).

BLOOD GLUCOSE (BG) SENSORS

Enzyme-Based Electrochemical/Oxygen Sensor for Long-Term Implantation in the Bloodstream (Intravascular Sensor)

Intravascular (IV) sensors for chronic implantation resemble a modern pacemaker, except that the distal end of the flexible lead (sensing region) remains free-floating in the superior vena cava (Fig. 6). The sensor lead is connected to an implanted electronics module that contains a microprocessor, power source, antennae, and telemetry. Data are transmitted to an external data manager with a flat-panel display.

Key to the IV sensor's long-term stability and function is the ability of blood flow to keep the sensor surface clean of platelets, cells, plasma proteins, and blood clot. Allowing the sensor to move freely in the vena cava blood flow optimizes the shear forces generated along the sensor's outer surface. This mechanism prevents significant accumulation of platelets, cells, and protein on or within the sensor's outer porous membrane. Clots that do form may be dissolved by fibrinolysis, or they may break loose to become lodged within the arterioles and capillaries of the lungs. Fortunately, small emboli are rapidly dissolved by fibrinolysis.^[26] The long-term clinical success of implanted pacemaker leads and drug delivery catheters in the superior vena cava provides a strong basis for the safety of this location for the IV glucose sensor. Catheter tips that stay mobile within the central portion of the vena cava remain relatively free of adhered proteins and cells for the life of the sensor.

Tips that contact the endothelium can become mechanically attached to the vena cava wall by an overgrowth of fibrous tissue. On the proximal end, the sensor lead travels from the subcutaneous tissue



Fig. 6 Enzyme-based electrochemical/oxygen sensor for long-term implantation with the tip of the flexible lead (sensing region) free-floating in the superior vena cava. Intravenous (IV) sensor developed by Medtronic-MiniMed Corporation. (Used with permission from Medtronic-MiniMed Corporation, www.medtronic-minimed.com.)

through the wall of the subclavian vein to enter the bloodstream. A layer of fibrous tissue (with or without endothelial cells) extends 3 to 5 cm from the vessel wall to encase the sensor lead.^[18] Removal of a sensor lead from the vena cava is therefore associated with significant surgical risk, due to vessel wall adhesion at two locations. Methods must be developed to remove chronically implanted leads safely, because more than 12 sensors will be required to provide continuous BG monitoring over the lifetime of a young type I diabetic (assuming 5-year sensor function).

University of California (San Diego) scientists Gough, Lucisano, Armour, McKean, and others developed an intravascular BG sensor that can function long-term in the ambulatory setting. The tip of the sensor contains immobilized enzymes glucose oxidase and catalase, layered over a potentiostatic oxygen sensor. Oxygen not consumed by the enzyme reactions is electrochemically reduced at the oxygen sensor to produce a current that is proportional to the glucose concentration. The implanted system requires a second sensor, similar to the glucose sensor but without enzymes, to simultaneously measure the concentration of oxygen in venous blood. The difference between the two sensor output signals provides an accurate estimate of the BG concentration.

The novel mechanical design guarantees that oxygen diffusion to the enzyme layer far exceeds the diffusion of glucose. An outer membrane of porous cross-linked albumin is used to cover the distal end of the enzyme layer, enhancing hemocompatibility and limiting the inward diffusion of glucose (relative to oxygen). The sides of the enzyme layer are surrounded by silicone rubber that is permeable to oxygen, but impermeable to glucose. Oxygen is therefore supplied to the enzyme region in sufficient quantity to avoid stoichiometric oxygen limitation. Glucose oxidase is loaded in excess of the amount calculated for long-term function, and catalase is coimmobilized in excess to prevent hydrogen peroxide-mediated enzyme inactivation. The pore-free hydrophobic membrane placed between the enzyme layer and the electrodes permits the free inward diffusion of oxygen, but it inhibits the passage of interfering substances (ascorbate, acetaminophen) and biochemical toxins that may foul the electrode. Each innovation has been designed to enhance glucose sensor performance and extend operational lifetime.^[18]

A similar enzyme-based electrochemical/oxygen IV glucose sensor is being developed by Medtronic-MiniMed (Northridge, California, <http://www.medtronic-minimed.com>). Sensors implanted in the superior vena cava of 10 ambulatory patients with type 1 diabetes functioned well for an average of 287 days (range: 89–431) with satisfactory correlation to reference BG samples. Complications such as thrombosis,



pulmonary embolism, or vena cava injury (during lead removal) did not occur.^[27]

Near-Infrared Absorption Spectroscopy Sensor for Long-Term Implantation Around a Vascular Tissue (Perivascular Sensor Head)

Long-term implantable optical sensors based upon absorption spectroscopy of near-infrared (NIR) light may overcome the diffusion limitations of traditional enzyme-based glucose sensors.^[28–30] NIR light can easily penetrate any layer of cells or proteins that coat the optical windows. It is not necessary for glucose or other analytes to diffuse through one or more membranes to reach an enzyme or electrode surface. The region between the light source and the sample detector defines a gap, or path length, that can be filled with blood, ISF, or a combination of cells, ECM, and ISF (tissue). The molecular bonds of glucose have a unique absorption signature that can be detected in the near-infrared region of the electromagnetic spectrum.^[30] The optical sensor is able to produce high-quality NIR spectra (intensity versus wavelength) using blood, ISF, or tissue as the sample. Numerous spectra are collected over a wide range of blood chemistries to develop a robust library of spectra, which can be mathematically analyzed to produce a calibration model. Subsequent spectra are mathematically compared (at multiple wavelengths) to the calibration data set, to provide an accurate prediction of the glucose concentration. Multivariate statistical methods such as partial least squares regression and filtering techniques can be used to construct the calibration model. The monitoring system must be very sensitive and stable to measure the very subtle optical changes caused by glucose that occur in the NIR spectrum. The optics must be very efficient, because much of the incident light is scattered.^[29] Biological challenges depend upon the tissue fluid being sampled,^[28] as in the following examples.

1. Placing the light source and detector on opposite sides of a vascular membrane provides complex NIR absorption spectra of cells, ECM, and ISF. The membrane between the sensor halves (gap) must remain healthy and vascular for the life of the sensor.
2. Placing the light source and detector opposite one another, on the outside walls of a blood vessel, provides NIR absorption spectra of blood and two blood vessel walls. There is a large body of literature describing successful long-term implantation of flow probes (electromagnetic and ultrasound) around the outside wall of both arteries and veins.^[31–33]

3. Placing the light source and detector within the bloodstream provides NIR absorption spectra of blood. This sample site eliminates many issues related to time lags and correlation with ISF glucose. Blood flow and shear forces are used to prevent thrombosis and maintain a thin and stable protein/cell coating over the optical windows.

The long-term implantable NIR glucose sensor being developed by Animas Corporation (Fraser, Pennsylvania, <http://www.animacorp.com>) contains a miniature optical sensor head, mechanically coupled to a vein, that measures NIR spectra of blood. The implantable system consists of a perivascular sensor head, a flexible lead, and an electronics/optics/telemetry module. An external electronics module (wristwatch) will receive, analyze, and display glucose data in real time.^[28–30]

CONCLUSION

Each of the implantable sensors requires a unique set of physical and biochemical properties to withstand the rigors of the *in vivo* environment. Systematic studies are under way to determine the optimal sensor size, geometry, surface structure, surface biochemistry, method of transduction, and method of implantation for a specific sensor in a specific location of implantation. Future efforts will focus on the development of novel membranes with chemistry and textures that: 1) promote the ingrowth of healthy vascular tissue; 2) prevent the formation of an avascular FBC; and 3) better protect the underlying enzyme and electrodes from fouling. Controlled delivery of anti-inflammatory and angiogenic factors into the outer membrane shows great promise for optimizing *in vivo* sensor performance and preventing sensor malfunction due to biofouling/degradation of membranes, enzymes, and electrodes.

ACKNOWLEDGMENTS

A special thank you to my wife Celia and children Mark and Amanda for providing support and guidance during preparation of this article. As always, Brian Hipszer provided expert and timely review, in addition to using his artistic talents to produce three-dimensional drawings.

ARTICLES OF FURTHER INTEREST

Biocatalysis; Biosensors; Insulin Delivery

REFERENCES

1. Wisniewski, N.; Moussy, F.; Reichert, W.M. Characterization of implantable biosensor membrane biofouling. *Fresenius' J. Anal. Chem.* **2000**, *366* (6–7), 611–621.
2. Sharkawy, A.A.; Klitzman, B.; Truskey, G.A.; Reichert, W.M. Engineering the tissue which encapsulates subcutaneous implants. I. Diffusion properties. *J. Biomed. Mater. Res.* **1997**, *37* (3), 401–412.
3. Gifford, R.; Kehoe, J.J.; Alterman, M.A.; Wilson, G.S. *Characterization by MALDI-MS of Proteins/Peptides Adhering to Subcutaneously-Implanted Glucose Biosensors. Abstract*; Society for Biomaterials 29th Annual Meeting Transactions; 2003.
4. Hirsch, I.B.; Farkas-Hirsch, R.; Skyler, J.S. Intensive insulin therapy for treatment of type 1 diabetes. *Diabetes Care* **1990**, *13* (12), 1265–1283.
5. Joseph, I.J.; Torjman, M.C. The artificial pancreas symposium: Glucose monitoring, insulin delivery, and feedback control. *Diabetes Technol. Ther.* **1999**, *1* (3), 323–328.
6. Rebrin, K.; Steil, G.M. Can interstitial glucose assessment replace blood glucose measurements? *Diabetes Technol. Ther.* **2000**, *2* (3), 461–472.
7. Mastrototaro, J.J. The MiniMed continuous glucose monitoring system. *Diabetes Technol. Ther.* **2000**, *2* (S1), S13–S18.
8. Thome-Duret, V.; Gangnerau, M.N.; Zhang, Y.; Wilson, G.S.; Reach, G. Modifications of the sensitivity of glucose sensor implanted into subcutaneous tissue. *Diabetes Metab.* **1996**, *22* (3), 174–178.
9. Anderson, J.M. Mechanisms of inflammation and infection with implanted devices. *Cardiovasc. Pathol.* **1993**, *2* (3S), 33S–41S.
10. Long, M.W. Tissue Microenvironments. In *The Biomedical Engineering Handbook*, 2nd Ed.; Bronzino, J., Ed.; CRC Press LLC: Boca Raton, 2000; Vol. 2, 118–1–118–17.
11. Reichert, W.M.; Sharkawy, A.A. Biosensors. In *Handbook of Biomaterials Evaluation*, 2nd Ed.; von Recum, A.F., Ed.; Edwards Brothers: Ann Arbor, 1999; 439–460.
12. Clark, R.A.F. Wound Repair: Lessons for Tissue Engineering. In *Principles of Tissue Engineering*; Lanza, R., Langer, R., Chick, W., Eds.; R.G. Landes Company, 1997; 737–768.
13. Johnson, R.B., Jr. Current concepts: Immunology. Monocytes and macrophages. *N. Engl. J. Med.* **1988**, *318* (12), 747–752.
14. Arenberg, D.A.; Strieter, R.M. Angiogenesis. In *Molecular and Cellular Basis of Inflammation*; Serhan, C.N., Ward, P.A., Eds.; Humana Press: New Jersey, 1998; 29–49.
15. Wood, R.C.; LeCluyse, E.L.; Fix, J.A. Assessment of a model for measuring drug diffusion through implant-generated fibrous capsule membranes. *Biomaterials* **1995**, *16* (12), 957–959.
16. Woodward, S.C. How fibroblasts and giant cells encapsulate implants: Considerations in design of glucose sensors. *Diabetes Care* **1982**, *5* (3), 278–280.
17. Fisher, U. Fundamentals of glucose sensors. *Diabet. Med.* **1991**, *8* (4), 309–321.
18. Armour, J.C.; Lucisano, J.Y.; McKean, B.D.; Gough, G.A. Application of chronic intravascular blood glucose sensor in dogs. *Diabetes* **1990**, *39* (12), 1519–1526.
19. Campbell, C.E.; von Recum, A.F. Microtopography and soft tissue response. *J. Invest. Surg.* **1989**, *29* (1), 51–74.
20. Brauker, J.H.; Carr-Brendel, V.E.; Martinson, L.A.; Crudele, J.; Johnston, W.D.; Johnson, R.C. Neovascularization of synthetic membranes directed by membrane microarchitecture. *J. Biomed. Mater. Res.* **1995**, *29* (12), 1517–1524.
21. Pierce, M.; Lichlyter, D.; Grant, S. *Engineering Surface Modifications to Enhance Sensor Compatibility*; Society for Biomaterials 29th Annual Meeting Transactions; 2003.
22. Hickey, T.; Kreutzer, D.; Burgess, D.J.; Moussy, F. Dexamethasone/PLGA microspheres for continuous delivery of an anti-inflammatory drug for implantable medical devices. *Biomaterials* **2002**, *23* (7), 1649–1656.
23. Heinemann, L. Continuous glucose monitoring by means of the microdialysis technique: Underlying fundamental aspects. *Diabetes Technol. Ther.* **2003**, *5* (4), 545–561.
24. Updike, S.J.; Shults, M.C.; Gilligan, B.J.; Rhodes, R.K. A subcutaneous glucose sensor with improved longevity, dynamic range, and stability of calibration. *Diabetes Care* **2000**, *23* (2), 208–214.
25. McShane, M.J.; Russell, R.J.; Pishko, M.V.; Cote, G.L. Glucose monitoring using implanted fluorescent microspheres. *IEEE Eng. Med. Biol. Mag.* **2000**, *19* (6), 36–45.
26. Hanson, S.R.; Harker, L.A.; Ratner, B.D.; Hoffman, A.S. In vivo evaluation of artificial surfaces with a nonhuman primate model of arterial thrombosis. *J. Lab. Clin. Med.* **1980**, *95* (2), 289–304.
27. Renard, E.; Shah, R.; Miller, M.; Kolopp, M.; Costalat, G.; Bringer, J. Sustained Safety and Accuracy of Central IV Glucose Sensors Connected to Implanted Insulin Pumps and Short-term Closed-loop Trials in Diabetic Patients, American Diabetes Association 63rd Scientific Sessions, New Orleans, Louisiana, June 13–17, 2003. <http://www.attenddinteractive.com/shows/ada0302/index.cfm> (accessed June 2003).
28. Joseph, J.I.; Torjman, M.C.; Moritz, M.; Liu, J.B.; Murphy, M.E.; DeStefano, M. *Long-Term Vessel Patency Following Perivascular Sensor Implantation*; Sixth World Congress Biomaterials; May 15–20, 2000; Vol. 1. Abstract # 231.
29. Crothall, K.D.; Joseph, J.I.; Tojman, M.C.; Jahn, L.G.; Moritz, M.J.; Sullivan, M.E.; Goldstein, B.J.; Butoi, B. Development of an implantable NIR spectroscopic blood glucose sensor: In vivo results in dogs. *Diabetes* **1999**, *48* (S1). Abstract #470.
30. Crothall, K.D.; Jahn, L.G.; Torjman, M.C.; Butoi, B.; Goldstein, B.J.; Joseph, J.I. Development of an implantable NIR spectroscopic blood glucose sensor: In



- vitro results with whole blood. *Diabetes* **1999**, *48* (S1). Abstract #454.
31. Vatner, S.F.; Higgins, C.B.; White, S.; Patric, T.; Franklin, D. The peripheral vascular response to severe exercise in untethered dogs before and after complete heart block. *J. Clin. Invest.* **1971**, *50*, 1950–1960.
 32. Van Citters, R.L.; Franklin, D.L. Cardiovascular performance of Alaska sled dogs in exercise. *Circ. Res.* **1969**, *24*, 33–42.
 33. Bednarik, J.A.; May, C.N. Evaluation of a transit-time system for the chronic measurement of blood flow in conscious sheep. *J. Appl. Physiol.* **1995**, *78*, 524–530.

Glues

Shojiro Matsuda

Research and Development Department, Gunze Limited, Ayabe, Kyoto, Japan

Yoshito Ikada

Department of Clinical Engineering, Suzuka University of Medical Science, Suzuka, Mie, Japan



INTRODUCTION

Surgical glues are widely used to seal small holes present in diseased soft tissues such as lung (sealants), to stop bleeding from damaged tissues (hemostatic agents), and to bond two separated tissues (surgical adhesives). The biomaterials currently used for these surgical purposes include fibrin glues, α -cyanoacrylates, and GRF (gelatin-resorcinol-formaldehyde) glues. All of them are applied to tissues in the liquid state (sol), followed by gel formation. This sol-gel transition takes place rapidly on the tissues. Fibrin glue, which comprises fibrinogen and thrombin, is clinically used for the purpose of sealing, hemostasis, and bonding, while cyanoacrylate monomers are utilized mostly as tissue adhesives. GRF with higher gel strength and bonding strength than fibrin glue has a unique application, that is, ceasing of bleeding from acute aortic dissection. Although these currently available glues are greatly contributing to surgical operations, they have disadvantages. For instance, fibrin glues have a potential risk of viral infection in spite of careful screening of plasma, because fibrinogen and thrombin are harvested from human plasma. Low bonding strength of the gelled fibrin to tissues is also a weak point. Cyanoacrylates are instantly allowed to polymerize, forming a gel in the presence of moisture under strong adhesion to tissues, but yield toxic formaldehyde as a result of hydrolytic degradation of the polymer main-chain. As formaldehyde is a major component of the GRF glue, attempts have been made to develop new liquid-type biomaterials that are free of virus, prion, and carcinogen, and can transform into a gel having a high potential of strong tissue adhesion in the presence of water. Glues should be absorbed after having fulfilled their task so as not to induce undesirable reactions as a foreign body.

REQUIREMENTS FOR SURGICAL GLUE

Sutures have been the most widely used method for wound closure, but suturing has shortcomings, such as the requirement of highly skillful procedures and a

long time to close wound and the postoperative removal of nonabsorbable suture. It is, therefore, desired to develop surgical glues that can overcome these problems. Various conventional glues are known for home and industrial use, but only three glues, namely, fibrin glues, α -cyanoacrylates, and GRF, are widely available for medical use.

Requirements for these surgical glues are listed in Table 1. As they are applied to the human body, they must be safe, nontoxic, and free from risk of infectious transmission. They should be preferably bioabsorbable and not remain long as a foreign body. Not only must the applying solution and gelled glue be safe, but also the degradation products as well. In addition, they must not hinder the natural healing process.

Most surgical glues are gelled in situ, applicable on rough surfaces, and act as tissue anchors. Appropriate gelling time and mechanical properties of formed gel are important. They should bond rapidly to the surrounding living tissues in the presence of water or aqueous fluids, such as blood, and keep the adhesion strength. Similar to the glue itself, degradation products of the glue must be safe. Besides the chemical nontoxicity, their mechanical properties are also important. If the cured glue is harder than tissue, it may induce undesirable tissue reactions. Good handling (ease of use and gelation speed) is also important.

FIBRIN GLUE

Fibrin glue, which consists of human fibrinogen, thrombin, factor XIII, and bovine aprotinin, has been widely used in Europe, Canada, and Japan for about two decades and available in the United States since 1998. Figure 1 shows the mechanism of gelation of fibrin glue. When fibrinogen solution and thrombin solution are mixed at the site of injury, fibrin gel formation occurs through the physiological coagulation and the tissue adhesion cascade. Thrombin catalyzes partial hydrolysis of fibrinogen to form fibrin monomers if calcium ion is present. Thrombin also activates factor XIII into factor XIIIa in the presence of calcium ions. Factor XIIIa reacts with fibrin monomers to form

Table 1 Requirements for surgical glues

1. In situ curable from liquid-state through polymerization, chemical cross-linking, or solvent evaporation
2. Rapidly curable under wet physiological conditions
3. Strongly bondable to tissues
4. Nontoxic for both agents and their degradation products
5. Tough and pliable as natural tissues
6. Biodegradable

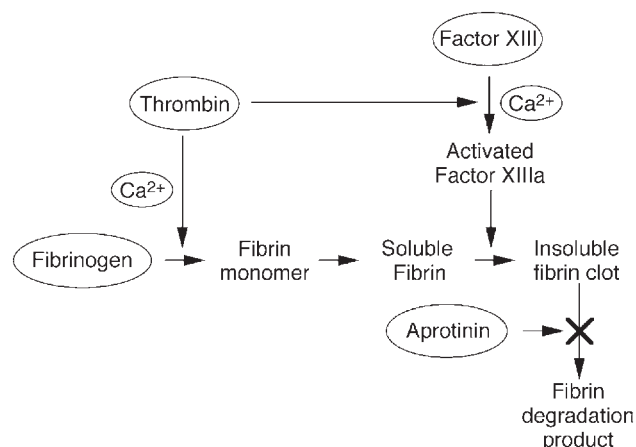
insoluble fibrin clots. Aprotinin inhibits fibrinolysis (degradation of fibrin clot). The fibrin clot will be totally resorbed by fibrinolysis and phagocytosis.

Fibrinogen and thrombin are dissolved separately in medium before surgical operation. Each of two solutions is filled in a separate dual syringe, which is set to Y-connector and applying needle or a spray system.

Application of Fibrin Glue

Surgical fields for fibrin glue and its application purposes are shown in Table 2. Mostly they work as sealants, hemostatic agents, and surgical adhesives. Hemostasis is also effectively carried out by suturing, electrocauterization, infrared coagulation, compression, and tamponade. Very small blood vessels, fragile tissues, and small spaces are difficult to apply with these devices, but liquid-type surgical glues are useful in these cases and also for sealing if they have a proper adhesion strength to tissue, and a proper physical-barrier strength in the presence of water.

The application activity of fibrin glue depends on surgical procedures.^[1] Cardiovascular surgery is the most common among them and its major purpose is hemostasis and suture line sealing. Successful local hemostasis of bleeding after open-heart surgery can reduce blood loss, operative time, and the need for

**Fig. 1** Mechanism of gelation of fibrin glue.**Table 2** Clinical applications of fibrin glue

Cardiovascular surgery	Hemostasis of bleeding from mediastinum and small blood vessels, reinforcement of suture line, sealing of vascular grafts and anastomoses.
Thoracic surgery	Sealing of air leaks from raw lung surface, and pulmonary and bronchial staple lines.
Gastroenterological surgery	Hemostasis and reinforcement of intestinal anastomosis, prevention of bile leakage, and hemostasis of hepatectomy, pancreatic fistula, and gastroduodenal ulcers.
Neurosurgery	Sealing of cerebrospinal fluid leaks, nonsutured peripheral nerve repair.
Gynecologic surgery	Anastomosis of fallopian tube, adhesion of peritoneum.
Plastic surgery	Closure of skin wound, attachment of skin grafts, fixation of tissue flaps.
Orthopaedic surgery	Fixation of bone fragments, hemostasis during total knee arthroplasty.
Urologic surgery	Anastomosis of urinary tube and adhesion of kidney, pelvis renalis, and ureter.

resternotomy in these high-risk patients. Fibrin glue is highly effective as a tissue sealant in diffuse, low-pressure bleeding such as coronary vein, right-ventricle/pulmonary-artery conduits, and right-ventricular patches. Fibrin glue is used routinely whenever suturing is impossible, dangerous, or difficult, and for spontaneous hemostasis to save waiting time. Suture holes, staple line, anastomoses, fistulas, and raw surfaces are effectively treated with fibrin glue, similarly to aortic and coronary suture line. It is also used for sealing vascular grafts. Perfect hemostasis is quite important in pediatric cardiac surgery because of the high incidence of coagulopathy after open-heart procedures and increased risks associated with the use of blood and blood products.

Fibrin glue is also very effective for reoperative procedures, because mediastinal bleeding and excessive non-suturable bleeding from adhesions are highly problematic. The adhesion of tissue induces complications during reoperation, prolonging the operation time and increasing the need for blood transfusion, and hence preventing the formation of adhesion is very important.

In thoracic surgery, fibrin glue is applied to cease hemorrhage and air leaks following pulmonary resection

and decortications. It is also used for the treatment of bronchopleural fistulae and pneumothorax. To reduce air leaks after stapled pulmonary resection, several kinds of predget, such as poly(glycolic acid) (PGA) fabric and bovine pericardium, have been clinically applied. Fibrin glue is also used for this purpose with PGA fabric.

In the treatment of bleeding from highly vascularized parenchymal tissues, such as elective hepatic resection and pancreatic fistula, and from anastomotic sites, such as recurrent bleeding from gastroduodenal ulcers, fibrin glue is effectively applied. Many other applications such as dural closure are known to prevent cerebrospinal-fluid (CSF) leaks, burn bleeding after debridement, and bleeding during the total knee arthroplasty (TKA) procedure, and control artery or aneurysm closure in endovascular surgery.

Future and Improvement of Fibrin Glue

Since fibrin glue is made from human and bovine proteins, potential risk of viral transmission cannot be fully denied, but there have been reported no proven cases of viral transmission associated with the use of commercial fibrin sealant. To circumvent this risk, the development of recombinant thrombin and fibrinogen is currently ongoing. Other sources, well-purified materials, and new synthetics are studied as alternatives to fibrin glue.

For the treatment of acute aortic dissection, fibrin glue does not have sufficient gel strength and bonding strength, and GRF and other glues are superior to fibrin glue. It is also claimed that fibrin glue is readily washed away even by weak bleeding, as no compression is possible on the bleeding area. In this case, collagen sheet is effective for the scaffold of fibrin glue.

Other clinical applications under study include the prevention of postoperative adhesion, endoscopic procedure for minimal invasive surgery, drug and cell delivery vehicle, and scaffolding of cell culture.

Fibrin glue sheet

TachoComb (Nycomed Arzneimittel, Germany) is a ready-to-use hemostatic agent.^[2] As this spongelike sheet is prepared by coating human fibrinogen, bovine thrombin, and bovine aprotinin on the surface of equine collagen fleece, it does not need to solidify, warm, or set to syringe. This dry sheet can be applied directly to traumatized tissues, in contrast to the liquid application of fibrin glue. TachoComb has mechanical stability, can be effectively used for the treatment of diffuse bleeding of the thoracic wall, and is absorbed within a few weeks. Bleeding site in the thoracic apex, especially after lysis of adhesion between lung and parietal pleura, is very difficult to treat with fibrin glue and other conventional methods. By using TachoComb, this bleeding is

effectively hemostased.^[3] Parenchymal leak of the lung after decortication can also be treated with this glue, when the visceral pleura is removed or lacerated. Hemostasis is effective and well tolerated as a secondary procedure in urological, ENT (ear, nose, and throat), hepatic, and vascular surgery when primary hemostatic measures are proved insufficient. TachoComb may be used to control oozing hemorrhage from surgical sites in gynaecological patients effectively and prevent adhesion formation at the application site, thus offering an effective method for preventing adhesion-induced infertility.

Adhesion prevention

Fibrin glue is also expected as an adhesion-prevention material. It is applicable to damaged tissue, fully covering it to avoid contact with another tissue. If the injured tissue comes in contact with another tissue, adhesion formation occurs during wound healing. Since fibrin glue is a kind of hydrogel with a high water content, water-soluble nutrients can pass through this hydrogel without affecting wound healing. The use of fibrin glue results in a significant decrease in postoperative adhesion of a rabbit model compared with the control side in the uterine horns and the ovaries.^[4]

Drug release vehicle

Because fibrin glue can form gel under a physiological condition, it can be easily mixed with drugs under a liquid condition and be applied to any tissue. Local delivery of drug has been carried out to avoid complications of systemic administration. The delivery vehicles for local delivery must stay at the applied place without diffusion into any undesired location. Thus, tissue adhesives are effective for sustained release of drugs. Bioabsorbable materials are applicable for sustained release of water-soluble drugs for a relatively short duration because of their high rate of diffusion. Poorly soluble drugs can be released from hydrogels such as fibrin glue as a result of vehicle degradation. Antibiotics for the treatment of large-area burns, antiproliferative drugs for chemotherapeutics, and growth factors such as bone morphogenetic protein (BMP) and basic fibroblast growth factor (bFGF) for tissue regeneration can be incorporated into fibrin glue for this drug delivery system.^[5]

Applications in tissue engineering

Fibrin monomer can be polymerized into moldable gels to encapsule cells and to provide three-dimensional scaffold of cell culture. Because fibrin glue is bioabsorbable, it is applicable to tissue engineering. When



chondrocytes trapped in fibrin glue gels are transplanted subcutaneously in nude mice, the actively proliferating cells can well form cartilaginous matrix.^[6] Urothelial cells, tracheal epithelial cells, and preadipocytes cultured in vitro were successfully transplanted onto a prefabricated capsule and tubelike structure using fibrin glue as a delivery vehicle and native cell expansion vehicle.^[7-9]

CYANOACRYLATE

When cyanoacrylate monomers come into contact with moisture, they change into a polymer. The detailed mechanism of polymerization of 2-cyanoacrylates is not fully understood, but it seems very probable that they polymerize via an ionic mechanism in which a water molecule acts as coinitiator. The liquid-state monomers polymerize to a solid within a short time period in the presence of water and are stored in one tube as a single component. Very high stiffness of cured glues, low viscosity of monomers, and toxic formaldehyde released during their degradation are drawbacks. However, cyanoacrylates have been used for wound closure, cartilage and bone grafting, coating of corneal ulcers, coating of aphthous ulcers, embolization of gastrointestinal varices, and embolization in neurovascular surgery.

Wound Closure

Surgical glues are preferable to skin sutures owing to good cosmetic outcome and pain reduction, especially in cosmetic surgery and surgery for children. Cyanoacrylate is the only surgical glue that is used for dermal wound closure instead of suturing. Tissue adhesives can save time during wound repair, provide flexible water-resistant protective coating, and eliminate the need for suture removal. Since the 1970s, n-butyl 2-cyanoacrylate (NBCA) with negligible histotoxicity and good bonding strength has been clinically applied and shown to have acceptable cosmetic outcome, but with too high bonding strength and too rapid reaction. Dermabond (Ethicon, Inc., Somerville, New Jersey) is 2-octyl cyanoacrylate that forms a strong bond across wound edges, allowing normal healing. The long-term cosmetic outcome of Dermabond was comparable to that of traditional methods of repair.^[10]

Blood Vessel Embolic Material

Endovascular neurosurgery is one of the strategies for treatments of various neurovascular diseases such as cerebral and spinal arteriovenous malformations (AVMs), dural arteriovenous fistulas, and cerebral aneurysms. A microcatheter is advanced into or close to a lesion and then an embolic material is administered

through it to obliterate the lesion. This less invasive method is accepted as an attractive alternative therapy making up a conventional surgical technique. Requirements for blood vessel embolic materials are shown in Table 3.

Various biomaterials have been used for embolization, but each material has some problems. Polymer solutions have ever been used as liquid-type materials for embolization, including poly(ethylene-co-vinyl alcohol), cellulose acetate, cationic methacrylate copolymer (Eudragit-E), and poly(vinyl acetate) emulsion. When polymer solution is given into a blood vessel, the organic solvent diffuses into the blood and surrounding tissues, thus forming polymeric precipitates and emboli. However, as most of them contain organic solvents, such as dimethyl sulfoxide and ethanol, they exert toxic effects on the surrounding tissues. Cyanoacrylate monomers that instantly polymerize upon contact with blood are used as liquid embolic material. Among the monomers, NBCA has been preferentially used in Europe and North America. NBCA is injected into the diseased part of the blood vessel through a long thin catheter and then quickly becomes solid, blocking blood flow into the diseased part of the blood vessel.

Although NBCA has been successfully used as embolic material, several troubles have been reported. Its instantaneous polymerization and quite low degradation rate are suitable for embolization of diseased lesions, but its exceptionally high adhesive force sometimes causes adhesion between the tip of the microcatheter and the artery. As a result, it becomes impossible to withdraw the microcatheter from the blood vessel and, in addition, the inner space of the microcatheter is occluded with the rapidly polymerized material when intermittently injected. Therefore, intermittent injection should be avoided and the microcatheter should be withdrawn immediately after completion of embolization. The surgeon is required to accumulate many experiences so that she or he can acquire skillful techniques to carry out embolization safely and effectively when NBCA is used.

To overcome these problems, isostearyl 2-cyanoacrylate (ISCA), which carries a long hydrophobic side isostearyl group with lower reactivity than n-butyl

Table 3 Requirements for blood vessel embolic materials

-
1. In situ curable from liquid-state (one liquid) through polymerization or solvent evaporation
 2. Curable in proper time in the presence of blood
 3. Tightly bondable to blood vessel wall tissue and not adhesive to microcatheter
 4. Nontoxic
 5. Tough and pliable as natural tissues
 6. Preferably radiopaque
-

group, was synthesized.^[11] Its adhesive force was low enough not to cause these problems. However, as its polymerization rate was too low to obliterate a vascular lesion with a rapid blood flow, ISCA was mixed with NBCA. The adhesive force of this mixture became extremely low, compared with that of NBCA, and adhesion between the tip of microcatheter and artery was not problematic. The viscosity of the mixture was low enough to use as embolic material. Tissue reactions toward the mixture were milder than those of NBCA and radio-angiography became possible by mixing further with Lipiodol. The evaluation of this new embolic material with a rabbit renal artery showed that the obliteration effect of the mixture from ISCA and NBCA was excellent to allow the use of this embolic material for clinical applications.

GRF GLUE

Gelatin is a widely used bioabsorbable material, especially as surgical glue. This natural polymer is extracted from biological tissues, and has many functional groups. Gelatin is soluble in water at temperatures higher than 35°C, and can form gel by lowering temperature below 25°C or by cross-linking. Gelatin gel is as elastic and soft as natural tissue. Braunwald reported that a gelatin-resorcinol mixture cross-linked with a combination of formaldehyde (GRF) and sometimes glutaraldehyde gave satisfactory bonding strength in hepatic and renal tissues of dogs.^[12] Formaldehyde undergoes condensation with resorcinol to give a three-dimensionally cross-linked water-insoluble resin. GRF is a two-liquid system glue. One component is aqueous solution of gelatin, resorcinol, and calcium chloride, and the other is an aqueous solution of formaldehyde and glutaraldehyde. After warming the polymer solution to 45°C and mixing 1 ml of polymer solution with two or three drops of aldehyde solution, gel is formed within about 1 minute.

GRF has largely been used in Europe to treat acute aortic dissections because of its ability to reinforce the delicate structure of the aortic wall.^[13] Every acute dissection involving the ascending aorta (Stanford type A) should undergo emergency surgical repair. 75% of Type A AAD patients die within two weeks. The surgical techniques vary depending on the clinical status of patients or the anatomical patterns observed. Furthermore, surgery is generally difficult because of very brittle aortic tissues, especially in advance-aged patients. Bachet et al. applied GRF glue for 212 patients with an emergency operation for type A aortic dissection. Their experience extending over more than 23 years proved that GRF glue is extremely useful, making the procedure much easier and safer. GRF is

said to adhere to the false lumen and shows excellent solidity and hemostasis of suture sites.^[14]

Since highly concentrated solutions of formaldehyde and glutaraldehyde are directly applied to an injured tissue, its toxicity is of major concern. Formaldehyde will be released from the cured glue when gelatin undergoes enzymatic hydrolysis and, therefore, surgical application of GRF glue is limited to special cases such as where high bonding strength is imperative.

OTHER SURGICAL GLUES

As mentioned earlier, fibrin glue, cyanoacrylates, and GRF glue show risk of viral infection, toxicity, low-adhesion strength, and long curing time. Thus many studies have been carried out to replace these glues. The curing modality of most surgical glues studied is either in situ cross-linking or in situ polymerization of monomer. Less toxic and bioabsorbable polymers such as gelatin, albumin, hyaluronic acid, chitin, poly(ethylene glycol), and poly(L-glutamic acid) have been studied and glutaraldehyde, formaldehyde, water-soluble carbodiimide, *N*-hydroxysuccinimide, and photoreactive agents have been used as cross-linkers. These agents are mostly multifunctional cross-linkers and are directly used at the application site or immobilized to structural polymers. Several studies will be described below.

Aldehyde System

Aldehyde can react with primary and secondary amines through Schiff's base. Since glutaraldehyde has two aldehyde groups in one molecule, it can react with two amines. If they are from different polymer molecules, glutaraldehyde will produce intermolecular cross-linking. Thus, by mixing the solution of polymer with amines in its side-chain and glutaraldehyde solution, gelation will occur. Natural polypeptides, such as collagen, gelatin, albumin, and enzymes, have amines that are reactive with aldehyde. So, glutaraldehyde is also used as a tissue fixation agent for micrograph observation. As aldehydes with high reactivity show cytotoxicity, too much volume and too high concentration of aldehyde solution may induce untoward tissue reactions.

Aldehyde-polysaccharides and amine-gelatin

Gelatin and polysaccharides were employed to prepare the hemostatic glue, which does not release any low molecular weight substances. To this end, gelatin was modified with ethylenediamine using water-soluble carbodiimide to introduce additional amino groups into the original gelatin, while dextran and hydroxyethyl-starch were oxidized by sodium periodate to convert

1,2-hydroxyl groups into dialdehyde groups.^[15] Upon mixing the two polymer components in aqueous solution, Schiff's base was formed between the amino groups in the modified gelatin and the aldehyde groups in the modified polysaccharides, thus resulting in intermolecular cross-linking and gel formation. The fastest gel formation took place within 2 seconds, and its bonding strength to porcine skin was 225 gf/cm² when 20 wt% of amino-gelatin (55% amine group) and 10 wt% of aldehyde-HES (>84% dialdehyde) aqueous solutions were mixed. In contrast, the gelation time and bonding strength of fibrin glue was 5 seconds and 120 gf/cm², respectively, when the same system was applied for the comparative evaluation.

Albumin glue

Albumin is a natural protein present in blood serum and used as structural polymer for a surgical glue similar to GRF. A glue consisting of 45% bovine serum albumin and 10% glutaraldehyde formed gel within 2 to 3 minutes. Herget et al. reported that this glue was effective as a sealant for bronchial anastomoses and parenchyma lesions in sheep.^[16] Bronchial anastomosis and parenchymal tissue repair could be sealed successfully against air leakage with this adhesive. Healing was not complicated much by foreign body reaction and tissue granulation.

Adhesion of aldehyde-gelatin film

It was reported that a glutaraldehyde (GA)-modified dry gelatin film could adhere to tissues.^[17] As described above, GA has two reactive aldehyde groups in one molecule and hence can act as a cross-linker of collagen, other proteins, and biological soft tissues. In a similar manner, GA can introduce cross-links into gelatin, making it insoluble even in hot water, where the physical cross-link of gelatin is destroyed. It is highly possible that at the gelation reaction with GA, one aldehyde group of a GA molecule will react with an amino group of gelatin, while the other aldehyde of the same GA molecule will not be able to find an appropriate amino group with which to react, thus remaining free. When a gelatin film was treated with 0.5 M GA solution at 60°C, free aldehyde groups were introduced in the film by up to 150 mmol/g. The dried gelatin film could strongly adhere to the porcine skin as strongly as 250 gf/cm², whereas the native gelatin film before GA treatment showed bonding strength of 40 gf/cm². The dried film could probably adsorb and remove the water existing between the GA gelatin film and the biological tissue, thus enabling aldehyde groups on the film and the amino groups on the natural tissue to approach each other and result in Schiff's

base formation. This reaction scheme was supported by the fact that GA-treated films did not demonstrate high bioadhesion when the remaining aldehydes were quenched with glycine or reduced with NaBH₄.

WSC and NHS System

Carbodiimide is a bifunctional cross-linker, similar to GA. 1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC) and 1-cyclohexyl-3-(2-morpholinoethyl) carbodiimide metho-p-toluenesulfonate (CMC) are available as water-soluble carbodiimide (WSC). WSC is a coupling agent to form an amide bond between amino and carboxyl groups in a short time period in aqueous solution. Therefore, WSC functions as a catalyst to cross-link molecules and no WSC-related fragments are incorporated in the structure of resulting cross-linked molecules. An urea derivative is produced from WSC after the catalytic reaction and is much less toxic than the original WSC.

N-Hydroxysuccinimide (NHS) is a well-known coupling agent that can be immobilized to the carboxyl group of a host polymer through ester bonding. WSC also functions as a catalyst of ester formation. Immobilized NHS-ester can easily react with amines of other molecules through ester-exchange reaction. No NHS-related fragments are incorporated in the reacted molecules as NHS is released after the reaction. Such ester formation is used for the cross-linking of aqueous polymer solution to induce gel formation.

Gelatin, PLGA, and WSC

Gelatin and poly(L-glutamic acid) (PLGA) are natural and synthetic polypeptides, respectively, and are soluble in water and biodegradable. Gelatin is a heat-denatured product of collagen and is obtained commonly by extraction from bovine or porcine bone and skin. This natural polymer contains reactive groups such as –COOH and –NH₂ in the side-chain while PLGA has only –COOH groups. An addition of WSC to aqueous solution containing both of these polymers induces gelation, because the amide bond formation between –COOH of PLGA and –NH₂ of gelatin will form intermolecular cross-linking.^[18] The gelation time decreased with an increase in the WSC concentration to the comparable level of fibrin glue. When applied to the dog spleen injured by needle pricking, the WSC-modified gelatin-PLGA hydrogel exhibited superior hemostatic capability to fibrin glue. The WSC-modified gelatin-PLGA hydrogel strongly adhered to the surface of the dog spleen, whereas the fibrin hydrogel was easily detached from the spleen surface.

NHS-PLGA and gelatin

WSC is as cytotoxic as glutaraldehyde and the urea, a byproduct of this glue, also has slight toxicity. To circumvent the toxicity of WSC and urea, a new gelatin-PLGA glue that does not use WSC was designed.^[19] NHS-ester yields stable products upon reaction with primary or secondary amines and the coupling reaction proceeds efficiently at the physiological pH. An attempt was made to synthesize a new NHS derivative of PLGA, that will induce prompt gelation when mixed with aqueous solution of gelatin. NHS-esters can also react with amines of natural protein, resulting in tissue adhesion. The gelation scheme between gelatin and this PLGA derivative is shown in Fig. 2. Esterification of PLGA with NHS was performed by adding WSC to the reaction mixture. After 20 h at an ambient temperature, the whole reaction mixture was poured into dehydrated acetone to precipitate the resulting NHS ester of PLGA (NHS-PLGA). The NHS-PLGA could be synthesized at high yields and was found to be stable for an extended time without losing the ability to cross-link with gelatin when stored under a dry-cold condition. This NHS-PLGA could spontaneously form a gel with gelatin in aqueous solution within a short time, comparable to fibrin glue, when gelation was allowed to proceed at pH 8.3. The bonding strength of NHS-PLGA-gelatin glues to natural tissues was higher than that of fibrin glue. With the increase of NHS-ester introduced to PLGA, the gelation time became shorter and the bonding strength increased.

NHS-PEG and SH-PEG

Poly(ethylene glycol) (PEG) is a highly hydrophilic synthetic polymer and has found wide applications in medicine. A new rapidly synthetic tissue sealant can be synthesized from tetra-succinimidyl and tetra-thiol-derivatized PEG. Wallace et al. studied the *in vitro* and

in vivo efficacy of this glue by dissolving the two reagents in aqueous buffer and subsequently spraying it on a tissue surface.^[20] This glue exhibited good adhesion to collagen membrane, PTFE graft, and carotid artery *in vitro*. The gel maintained fluid pressure of 125 mmHg, which is fivefold greater than the capillary blood pressure and one-half that observed in hypertension. Bleeding in rabbit arteries was stopped immediately in five out of six trials. A significant reduction was observed for the time until to hemostasis and blood loss, compared to the control. Carotid artery and subcutaneous implant data in rabbits showed that the formulation was compatible with biological tissues.

NHS-PEG and albumin

Kobayashi et al. reported the ability of an albumin-based hydrogel sealant (ABHS) to prevent air leakage through a suture line after pulmonary surgery.^[21] This glue is a hybrid of synthetic and natural products and has two basic components: PEG disuccinimidylsuccinate (cross-linker) and human serum albumin. The terminal reactive ester of the PEG-derived cross-linker and the free amine groups in albumin react to form a gel. When the rat lung was used as an air-leak model, the average burst pressure of the ABHS-treated group was higher than that of fibrin glue throughout the observation period. Histological examination of the incision at day 14 revealed that no sealant was visible at the incision site, without any evidence of adverse tissue reaction.

Photocurable System

With the use of a photosensitizer, UV and visible light can initiate radical polymerization. If dye, azide, and acrylic macromer are used at a suitable combination, polymerization is initiated under a mild condition,

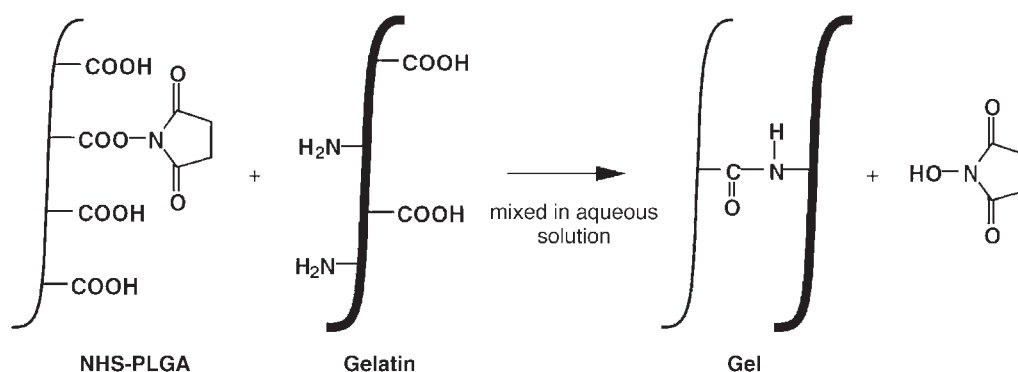


Fig. 2 Gelation scheme between gelatin and NHS-PLGA.

enabling its application as surgical glue. Water-soluble macromers, such as PEG diacrylate, become insoluble upon light irradiation through polymerization. Optical fiber or other devices are used for light irradiation. UV irradiation should be avoided because of its toxic reaction.

PEG diacrylate

Hubbell et al. reported the photoinitiated polymerization of a biodegradable tissue adhesive in 1993.^[22] This is made from the macromer that consists of a copolymer of water-soluble PEG with the two terminals capped with biodegradable oligomer such as glycolide and lactide oligomers. The two terminals of this macromer are further capped by acrylates to provide an unsaturated hydrocarbon required for photoinitiation by UV or visible light in the presence of eosin Y, triethanolamine, and 1-vinyl 2-pyrrolidinone. Depending on the type and molecular weight of the capped oligomer and the molecular weight of PEG, the resulting sealant has a wide range of biodegradation kinetics ranging from 1 day to 6 months.

Tanaka et al. evaluated the efficacy of this glue using a canine acute descending aortic dissection model.^[23] When the glue was applied to the false cavity for reinforcing and fusing the dissected layers and also to the suture line under a wet condition, hemostasis was easily achieved during surgery. All false cavities were perfectly thrombosed, causing no deleterious effects related to the glue, indicating that this glue can be used easily and effectively for the treatment of acute aortic dissection.

Photoreactive-gelatin and PEG-diacrylate

Nakayama et al. reported on a photocurable glue composed of gelatins partially derivatized with photoreactive xanthene dyes (fluorescein, eosin, and rose bengal) and a hydrophilic difunctional PEG macromer.^[24] When a dye-derivatized gelatin (20 wt%), a PEG diacrylate (10 wt%), and ascorbic acid (0.3 wt%) were dissolved in saline solution and irradiated with visible light, a swollen gel was formed within a few tenths of a second due to dye-sensitized photo-cross-linking and photograft polymerization. When a rat liver injured on laparotomy was coated with this hemostatic glue and irradiated with visible light through an optical fiber, the coated viscous solution was immediately converted to a swollen gel and hemostasis was completed. According to a histologic examination at the 7th day after surgery, little gelatin remained in the injured region, scarring with little necrosis occurred, and inflammatory cell infiltration from the surrounding tissue and tissue regeneration proceeded well.

Hyaluronic acid and albumin indocyanine

A photocurable tissue sealant composed of hyaluronic acid, albumin, and indocyanine was studied. For instance, 0.4 ml of hyaluronic acid, 0.2 ml of albumin, and three drops of indocyanine green dye gave a sealant with a peak absorbance of 805 nm, which matches the wavelength (808 nm) of a small, hand-held diode laser. Because tissues do not absorb any light of this wavelength, laser energy is focused in the sealant, minimizing collateral thermal damage. Auteri et al. reported the efficacy of this glue using a canine model of esophageal closure.^[25] The tissue sealant was applied to the edges of a hand-sewn closure, and then exposed to diode laser. The end-point was visible shrinking and desiccation of the sealant, which required about 2 minutes. At all time points, closures with this glue had significantly higher bursting pressure than the control closures. Histologic study revealed trace thermal injury with regeneration of intact mucosal lining by 7 days.

Azide chitosan lactose

Ono et al. studied the efficacy of a photo-cross-linkable chitosan glue, to which both azide and lactose moieties were introduced (Az-CH-LA).^[26] Introduction of the lactose moieties yielded a much more water-soluble chitosan at the neutral pH. UV irradiation to Az-CH-LA produced an insoluble hydrogel within 60 s. This hydrogel firmly bonded two pieces of sliced ham with each other at appropriate Az-CH-LA concentrations. The bonding strength of the chitosan hydrogel prepared from 30–50 mg/mL of Az-CH-LA was similar to that of fibrin glue. Compared to the fibrin glue, the chitosan hydrogel more effectively sealed air leakage from pinholes on an isolated small intestine and aorta and from incisions on an isolated trachea. Neither Az-CH-LA nor its hydrogel showed any cytotoxicity. Furthermore, all mice studied survived for at least 1 month after implantation of 200 μ L of photo-cross-linked chitosan gel and intraperitoneal administration of 1 mL of 30 mg/mL Az-CH-LA solution.

CONCLUSION

Although numerous studies have been carried out to develop surgical glues, as described in the preceding discussion, no perfect glue has been developed yet. Several surgical glues such as fibrin glue, cyanoacrylate, and GRF glue are clinically used, but they have limitations. Requirements for surgical glues such as viscosity, gelation time, adhesion strength, and duration time (biodegradation time), are different depending on the application site and purpose. Handling

of glue is also important for the clinical application. Warming and light irradiation may be hazardous and the toxicity of agents should be taken into consideration. Reactive chemicals have some risks of toxicity and natural biomaterials have a risk of virus infection.

Surgical glues were studied first to replace sutures, but they are mostly used also for hemostasis and sealing. In the future, they may be applicable to less invasive surgery and tissue engineering. However, intimate collaboration between material scientists, clinicians, and industries is needed for the development of new good glues.

ARTICLES OF FURTHER INTEREST

Bioadhesion; Drug Delivery, Controlled; Fibrin; Fibrin Sealants; Photopolymerization; Sutures; Thrombosis

REFERENCES

- Sierra, D.H.; Saltz, R. *Surgical Adhesives and Sealants: Current Technology and Applications*; Technomic: Pennsylvania, 1996.
- Agus, G.B.; Bono, A.V.; Mira, E.; Olivero, S.; Peilowich, A.; Homdrum, E.; Benelli, C. Hemostatic efficacy and safety of TachoComb in surgery. Ready to use and rapid hemostatic agent. *Int. Surg.* **1996**, *81* (3), 316–319.
- Hollaus, P.; Pridun, N. The use of Tachocomb in thoracic surgery. *J. Cardiovasc. Surg. (Torino)* **1994**, *35* (6 Suppl 1), 169–170.
- Takeuchi, H.; Toyonari, Y.; Mitsuhashi, N.; Kuwbara, Y. Effects of fibrin glue on postsurgical adhesions after uterine or ovarian surgery in rabbits. *J. Obstet. Gynaecol. Res.* **1997**, *23* (5), 479–484.
- Woolverton, C.J.; Fulton, J.A.; Salstrom, S.J.; Hayslip, J.; Haller, N.A.; Wildroutd, M.L.; MacPhee, M. Tetracycline delivery from fibrin controls peritoneal infection without measurable systemic antibiotic. *J. Antimicrob. Chemother.* **2001**, *48* (6), 861–867.
- Sims, C.D.; Butler, P.E.; Cao, Y.L.; Casanova, R.; Randolph, M.A.; Black, A.; Vacanti, C.A.; Yaremchuk, M.J. Tissue engineered neocartilage using plasma derived polymer substrates and chondrocytes. *Plast. Reconstr. Surg.* **1998**, *101* (6), 1580–1585.
- Wechselberger, G.; Russell, R.C.; Neumeister, M.W.; Schoeller, T.; Piza-Katzer, H.; Rainer, C. Successful transplantation of three tissue-engineered cell types using capsule induction technique and fibrin glue as a delivery vehicle. *Plast. Reconstr. Surg.* **2002**, *110* (1), 123–129.
- Bach, A.D.; Bannasch, H.; Galla, T.J.; Bittner, K.M.; Stark, G.B. Fibrin glue as matrix for cultured autologous urothelial cells in urethral reconstruction. *Tissue Eng.* **2001**, *7* (1), 45–53.
- Rainer, C.; Wechselberger, G.; Bauer, T.; Neumeister, M.W.; Lille, S.; Mowlavi, A.; Piza, H.; Schoeller, T. Transplantation of tracheal epithelial cells onto a pre-fabricated capsule pouch with fibrin glue as a delivery vehicle. *J. Thorac. Cardiovasc. Surg.* **2001**, *121* (6), 1187–1193.
- Bruns, T.B.; Worthington, J.M. Using tissue adhesive for wound repair: A practical guide to dermabond. *Am. Fam. Phys.* **2000**, *61* (5), 1383–1388.
- Oowaki, H.; Matsuda, S.; Sakai, N.; Ohta, T.; Iwata, H.; Sadato, A.; Taki, W.; Hashimoto, N.; Ikada, Y. Non-adhesive cyanoacrylate as an embolic material for endovascular neurosurgery. *Biomaterials* **2000**, *21*, 1039–1046.
- Braunwald, N.S.; Gay, W.J.; Tatoes, C. Evaluation of crosslinked gelatin as a tissue adhesive and hemostatic agent: An experimental study. *Surgery* **1966**, *59*, 1024–1030.
- Hata, M.; Shiono, M.; Orime, Y.; Yagi, S.; Yamamoto, T.; Okumura, H.; Kimura, S.; Kashiwazaki, S.; Choh, S.; Negishi, N.; Sezai, Y. The efficacy and mid-term results with use of gelatin resorcin formalin (GRF) glue for aortic surgery. *Ann. Thorac. Cardiovasc. Surg.* **1999**, *5* (5), 321–325.
- Bachet, J.; Goudot, B.; Dreyfus, G.; Brodaty, D.; Dubois, C.; Delentdecker, P.; Teimouri, F.; Guilmet, D. Surgery of acute type A dissection: What have we learned during the past 25 years? *Z. Kardiol.* **2000**, *89* (Suppl 7), 47–54.
- Mo, X.; Iwata, H.; Matsuda, S.; Ikada, Y. Soft tissue adhesive composed of modified gelatin and polysaccharides. *J. Biomater. Sci., Polym. Ed.* **2000**, *11*, 341–351.
- Herget, G.W.; Kassa, M.; Riede, U.N.; Lu, Y.; Brethner, L.; Hasse, J. Experimental use of an albumin-glutaraldehyde tissue adhesive for sealing pulmonary parenchyma and bronchial anastomoses. *Eur. J. Cardio-Thorac. Surg.* **2001**, *19* (1), 4–9.
- Matsuda, S.; Iwata, H.; Se, N.; Ikada, Y. Bioadhesion of gelatin films crosslinked with glutaraldehyde. *J. Biomed. Mater. Res.* **1999**, *45*, 20–27.
- Otani, Y.; Tabata, Y.; Ikada, Y. A new biological glue from gelatin and poly (L-glutamic acid). *J. Biomed. Mater. Res.* **1996**, *31*, 157–166.
- Iwata, H.; Matsuda, S.; Mitsuhashi, K.; Itoh, E.; Ikada, Y. A novel surgical glue composed of gelatin and N-hydroxysuccinimide activated poly(L-glutamic acid): Part 1. Synthesis of activated poly(L-glutamic acid) and its gelation with gelatin. *Biomaterials* **1998**, *19*, 1869–1876.
- Wallace, D.G.; Cruise, G.M.; Rhee, W.M.; Schroeder, J.A.; Prior, J.J.; Ju, J.; Maroney, M.; Duronio, J.; Ngo, M.H.; Estridge, T.; Coker, G.C. A tissue sealant based on reactive multifunctional polyethylene glycol. *J. Biomed. Mater. Res.* **2001**, *58* (5), 545–555.
- Kobayashi, H.; Sekine, T.; Nakamura, T.; Shimizu, Y. In vivo evaluation of a new sealant material on a rat lung air leak model. *J. Biomed. Mater. Res.* **2001**, *58* (6), 658–665.
- Sawhney, A.S.; Pathak, C.P.; Hubbell, J.A. Bioerodible hydrogels based on photopolymerized poly (ethylene



- glycol)-co-Poly (α-hydroxy acid) diacrylate macromers. *Macromolecules* **1993**, *26*, 581.
23. Tanaka, K.; Takamoto, S.; Ohtsuka, T.; Kotsuka, Y.; Kawauchi, M. Application of AdvaSeal for acute aortic dissection: Experimental study. *Ann. Thorac. Surg.* **1999**, *68* (4), 1308–1313.
 24. Nakayama, Y.; Matsuda, T. Newly designed hemostatic technology based on photocurable gelatin. *ASAIO J.* **1995**, *41* (3), M374–M378.
 25. Auteri, J.S.; Oz, M.C.; Jeevanandam, V.; Sanchez, J.A.; Treat, M.R.; Smith, C.R. Laser activation of tissue sealant in hand-sewn canine esophageal closure. *J. Thorac. Cardiovasc. Surg.* **1992**, *103* (4), 781–783.
 26. Ono, K.; Saito, Y.; Yura, H.; Ishikawa, K.; Kurita, A.; Akaike, T.; Ishihara, M. Photocrosslinkable chitosan as a biological adhesive. *J. Biomed. Mater. Res.* **2000**, *49* (2), 289–295.

Glycosaminoglycans and Proteoglycans

Jan C. Schagemann

Hsi-Wei Chung

Eike H. Mrosek

Shawn W. O'Driscoll

Gregory G. Reinholz

Department of Orthopedic Surgery, Cartilage and Connective Tissue Research Laboratory, Mayo Clinic College of Medicine, Rochester, Minnesota, U.S.A.

G

INTRODUCTION

Glycosaminoglycans (GAGs) are polysaccharides that are present on cell surfaces or in the extracellular matrix as free molecules or proteoglycans (PGs). GAGs are especially abundant in articular cartilage. Hyaluronan (hyaluronic acid; HA), chondroitin 4- and 6- sulfates, keratan sulfates (KSs) I and II, heparin, heparan sulfate, and dermatan sulfate are the most physiologically significant GAGs known. GAGs are involved in many biological functions such as embryonic development (e.g., synovial joints), wound healing, cell signaling, growth factor regulation, protein trafficking, etc. GAGs such as heparin (anticoagulant) and HA (orthopedic, ophthalmology, and cosmetic use) are also widely used clinically, and chondroitin sulfate (CS) is commonly used as a dietary supplement. Mutations in genes involved in the biosynthesis and degradation pathways of GAGs result in a number of hereditary disorders. In addition, GAGs have been implicated in diseases such as cancer, metastasis, viral infections, and amyloid diseases such as Alzheimer's disease.^[1] The main focus of this review will be on the role of GAGs and PGs in synovial joints and articular cartilage specifically.

A complex milieu that includes collagens, GAGs, PGs, proteases, cytokines, growth factors, and mechanical stimuli directs the development, function, and maintenance of synovial joints and articular cartilage. Articular cartilage is the smooth glistening white and remarkably durable hyaline tissue that covers the bone in our joints and supports the movement of joints by greatly reducing friction between bones. In articular cartilage, a sparsely distributed population of specialized cells (chondrocytes) synthesize different matrix components and assemble and organize them into a highly ordered macromolecular framework consisting largely of Type II collagen and the PG aggrecan. The unique properties of free and protein-bound GAGs provide the shock-absorbing and lubricating properties necessary for the high demands placed on joints.

GLYCOSAMINOGLYCAN STRUCTURE

GAGs are long unbranched polysaccharides with repeating disaccharide units containing uronic acid (glucuronic or iduronic) or hexose (galactose) and hexosamine (galactosamine or glucosamine). GAGs form long chains of negative charges owing to disaccharide units that contain negatively charged carboxylate or sulfate groups. The majority of GAGs are covalently bound to protein cores to form PGs. GAGs extend perpendicularly from the core protein attachments in a brush-like manner. HA is unique among the GAGs in that it does not contain a sulfate group and does not appear to covalently bind to protein. HA is composed of repeating units of *N*-acetylglucosamine and glucuronic acid. Although it is not covalently linked to protein cores, it is an essential backbone for aggregating PGs such as aggrecan.

Chondroitin-, heparan-, and dermatan sulfates are predominantly linked to serines of the protein core via tetrasaccharide. CS contains a series of 30–50 repeating disaccharides of *N*-acetylgalactosamine and glucuronic acid in β 1-3 linkage, either unsulfated or sulfated in 4 or 6 position of the galactosamine residue. Dermatan sulfate is a variation through iduronic acid instead of glucuronic acid by epimerization of C5 carboxylic group and sulfation at C2 position of glucuronic acid. Heparan sulfate can consist of glucosamine and either glucuronic or iduronic acid including *N*-sulfated *N*-acetylglucosamine acetyl groups. Compared to heparin, heparan sulfate contains more *N*-acetyl groups and fewer *N*-sulfate or *O*-sulfate groups. KS is composed of *N*-acetylglucosamine and galactose instead of uronic acid in β 1-4 linkage. This GAG binds covalently to protein cores, either *N*-linked to asparagines or *O*-linked to serine or threonine. In addition to mannose, fucose and, sialic acid, KS's sulfate content may vary and occurs in both galactose and hexosamine. Interestingly, evidence suggests that GAGs contain information in the form of a "sulfation code," which may specify protein function and location.^[2]

PGS

Articular cartilage PGs are a group of complex and fascinating molecules with functional and structural diversity (Fig. 1). Originally named owing to their GAG content, they can be categorized as a protein core-encoding molecule family. Generally, they are composed of a central protein core, oligosaccharides, and GAG chains. Although articular cartilage PGs provide a broad combinatorial potency and functional specificity, they can be categorized in two major groups: large aggregating PGs such as aggrecans (Figs. 1 and 2) and versican, and small, leucine-rich repeat PGs (SLRPs) such as decorin, biglycan, fibromodulin, and lumican. Additional PGs relevant to cartilage include among others perlecan and superficial zone protein (SZP)/lubricin.

PGs play either structural roles or may be involved in cell regulation, and contribute 25–35% of the dry weight of healthy articular cartilage. They fulfill a variety of biological functions such as matrix organization, cell adhesion, lubrication, ionic filtration, and growth and growth factor modulation. Many of them differ between their deposition and structure in

anatomical site, age, or biomechanical stress. Differences in composition, organization, and function can be found among the cartilage layers and cell-related matrix compartments. PGs resist compression (Fig. 2): negatively charged GAG chains of aggrecan in particular contribute more than 50% of the equilibrium compressive stiffness of articular cartilage.^[3,4] Moreover, they are crucial for molecule passing/rate and tissue fluid. Negative charges repel other negatively charged chains and attract ions, which maintain the fluid within the matrix and regulate electrolyte concentrations related to the Donnan osmotic pressure.

Importantly, the extracellular matrix protects chondrocytes from mechanical stress and preserves phenotype. In this context, habitually loaded cartilage regions may show high aggrecan deposition, whereas dynamic loading and mobilization further increase aggrecan concentration combined with restitution of biomechanical properties.^[3] Physical stimulation alters both matrix production and quality and function of newly synthesized PGs.

It is important to note that zonal organization of articular cartilage has functional significance (Fig. 3 a₂ and b₁).^[3] Besides the PG-rich pericellular matrix

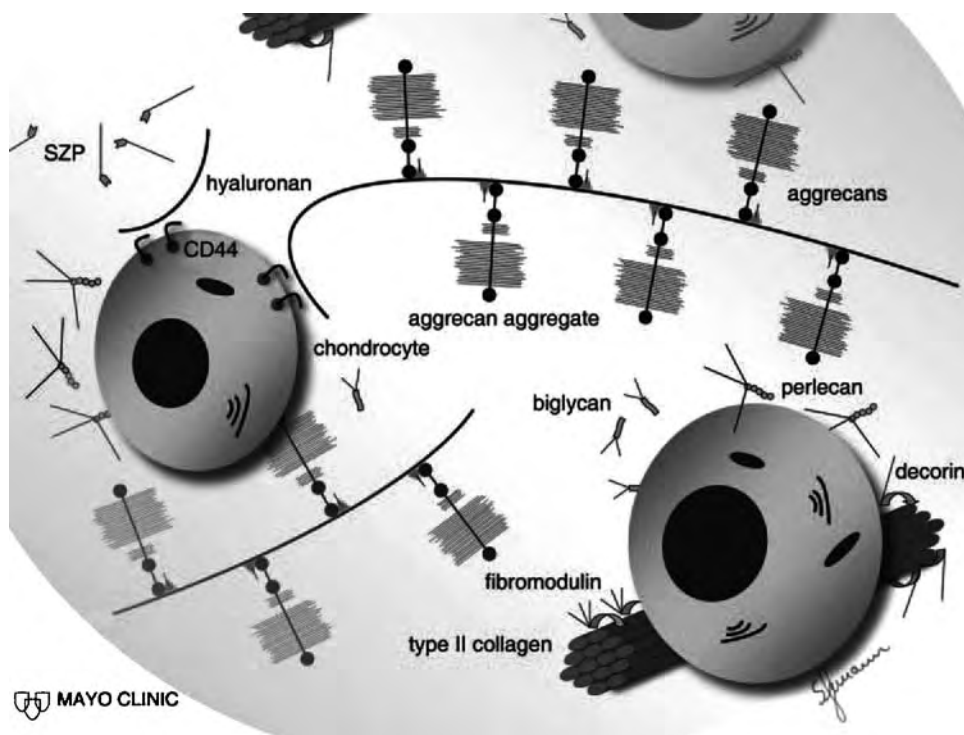


Fig. 1 Schematic overview of articular cartilage PGs. Aggrecan macromolecules bind non-covalently with HA molecules to assemble into large aggregates. HA forms a backbone with multiple palisading aggrecan molecules. Binding of HA to cell surface receptor CD44 provides pericellular matrix homeostasis. Moreover, abundance of SLRPs are present in the articular cartilage matrix. Decorin and fibromodulin are considered to play a role in organizing and stabilizing the collagen meshwork, whereas biglycan can be found in the pericellular matrix. Perlecan and its protein core are characterized by five distinct domains. Perlecan is enriched in the pericellular matrix. SZP is specifically synthesized by superficial chondrocytes and synoviocytes.

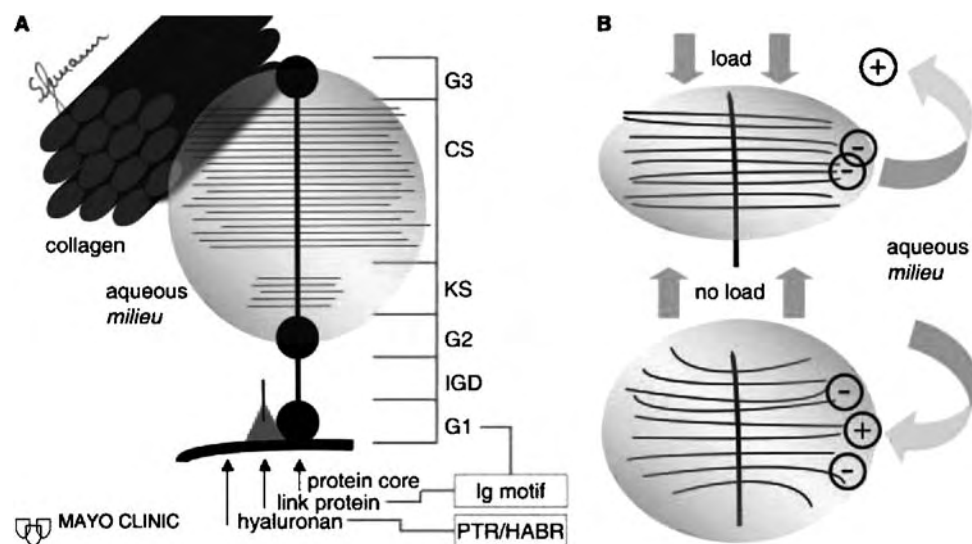


Fig. 2 (A) Schematic structure of the multimodular macromolecule aggrecan characterized by its protein core consisting of three distinct globular domains (G1–3) and related interglobular domains. The complex G1 domain is responsible for non-covalent interactions with HA [PG tandem repeat (PTR)/HA binding region (HABR)] and link protein [immunoglobulin (Ig) motif], which stabilizes the connection between aggrecan monomers and HA. KS and CS attachment domains are located between G2 and G3. GAGs are posttranslationally, covalently linked to the protein core filament and extend perpendicularly in a brush-like manner. Aggrecans can contain up to 60 KS chains and 100 CS chains. CS side chains bind to collagen fibrils allowing for cross-linked matrix meshwork. (B) GAG chains are negatively charged. Hence electrostatic repulsion between chains provides fully extended conformation and compressive stiffness as well as molecule passing/rate and tissue fluid equilibrium.

(Fig. 3a₃ and b₃), low PG concentration is apparent in the superficial zone (Fig. 3a₂ and b₂). Presumably, superficial cells degrade PGs more rapidly and synthesize less matrix compared to the deeper zones where higher PG concentrations can be found.^[3] However, the clarity of the zonal differences in PG content depends on the histological methods used (compare Fig. 3a₂ and b₁). It is also important to note that interaction of all components acting in concert is pivotal for the viscoelastic properties of articular cartilage, which allow it to withstand forces.

Aggregating PGs

Among the large aggregating PGs, the prototypical multimodular macromolecule aggrecan has been extensively studied owing to its abundance in articular cartilage, major contribution to compressive resistance, and degradation in cartilage disorders. Aggrecan (2×10^3 kDa) is the major PG in cartilage. It fills most of the interfibrillar space contributing to about 90% of the total matrix PGs mass. Aggrecans bind non-covalently with HA molecules on the protein core via link proteins and form large, multimolecular aggregates (Fig. 1). The aggrecan protein core (~ 230 kDa) consists of three globular domains (G1–3), with distinct motifs and folded structures and related interglobular domains in between (Fig. 2). The N-terminal globular domain G1

is responsible for interactions with HA. It contains an Ig motif and two PTR units. The Ig motif mediates the interaction between aggrecan and link protein, and the two PTRs formed of disulfide bonding structures provide functional sites of the aggrecan-HA binding within the HABR.^[5,6] The short, relatively stiff and inflexible interglobular domain IGD is located between G1 and G2. It contains proteolytic cleavage sites to a variety of proteinases.^[4] The G2 domain is unique to aggrecan. Although tandem repeats similar to G1 could be found, no binding capacity to other macromolecules has been observed. The C-terminal globular, multifunctional domain G3 is produced by alternative splicing of exons during posttranscriptional processing and contains three modules with homology to epidermal growth factor, complement reactive protein, and C-type lectin.^[6] Presumably, G3 facilitates endoplasmic reticulum (ER)-Golgi complex trafficking and GAG chain binding and enhances product secretion.^[4,6] Its lectin-like properties may explain extracellular matrix interactions and intracellular synthesis regulation.^[6,7] G2 and G3 domains are separated by an extended carbohydrate-rich domain that includes KS and CS attachment domains.^[8,9] The KS domain is not the only site for KS chain attachment in the protein core. The CS attachment domain is the largest domain of aggrecan and notably encoded by a single exon. Aggrecan can contain up to 60 KS chains (5–15 kDa each) and up to 100 CS chains (~ 20 kDa each), which are posttranslationally,

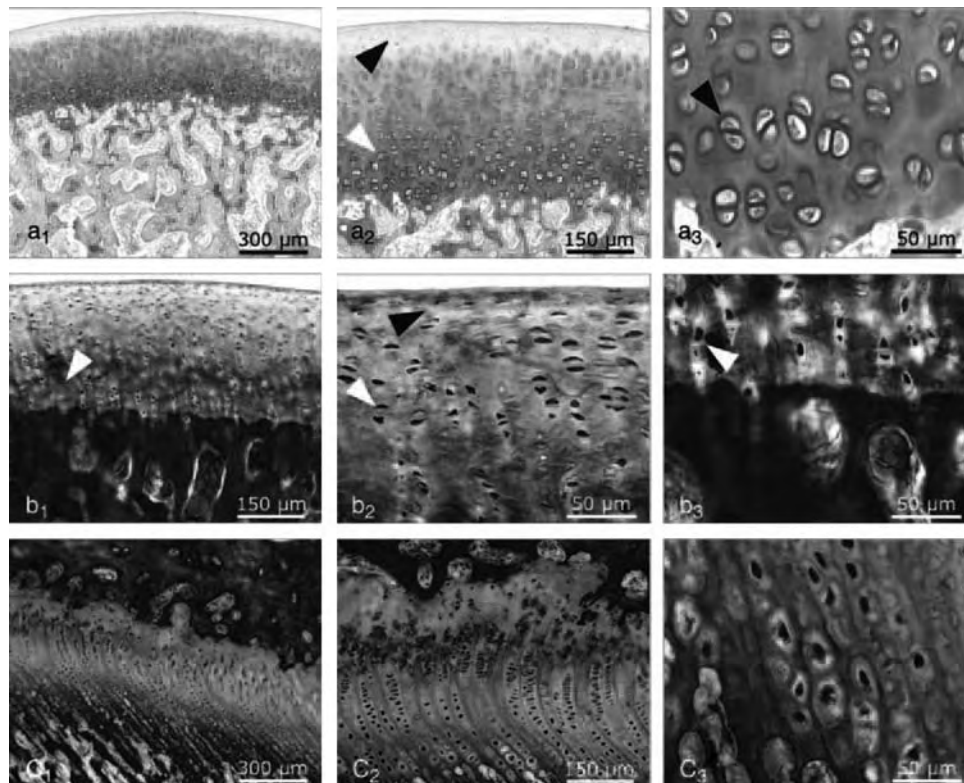


Fig. 3 Safranin O/fast green stained histology obtained from femoral condyles of skeletally immature New Zealand white rabbits using (a) standard paraffin sections, (b) and (c) the Exakt™ system. Series (a) and (b) show articular cartilage while series (c) shows undeveloped growth plate before fusion that is highly enriched with PGs. Articular cartilage shows a characteristic zonal morphology through subchondral, trabecular bone (a₁ and b₁). Low PG content can be found in the superficial zone (a₂ and b₂; black arrows) compared to higher amounts in deeper zones (white arrows). In particular, the pericellular matrix is highly stained for PGs (a₃ and b₃). Notice different Safranin O intensities while comparing histology processed either with the Exakt™ system or with standard protocols: decalcified paraffin sections (a₂) depict highest Safranin O intensities within the extracellular matrix of the deep and calcified zone (white arrow). In contrast, this could not be seen in non-decalcified sections (b₁) using the Exakt™ system (white arrow). Moreover, histological processing via the Exakt™ system leads to less shrinking artifacts and higher structural preservation.

covalently linked to the protein core filament. The CS side chains in aggrecan bind to collagen fibrils (namely Type II in cartilage) to form a cross-linked matrix. Both CS and KS are negatively charged; hence, electrostatic repulsion between chains and along chains allows for PGs to achieve a fully extended conformation.

Link proteins are small non-collagenous glycoproteins, which are highly homologous to the N-terminus of aggrecan. They stabilize the connection between aggrecan monomers and HA,^[10] which forms a backbone with palisading aggrecan molecules. Link proteins seem to play a role in directing aggregates assembling within the extracellular matrix, whereas HA can bind to the cell surface receptor CD44 providing pericellular matrix homeostasis.^[5] Length of the central HA can vary from several hundred to more than 10,000 nm.^[3] Spatial aggregation enables PG anchorage within the matrix, which prevents displacement and provides an organized and stabilized connection and entrapment to the collagenous meshwork.

Versican, a large CS PG, is also expressed in articular cartilage, but at a much lower level than aggrecan. It has been shown that versican represents a splice variant of the larger PG-M.^[11] Versican with its G1 and G3 domains shares this general structure with aggrecan and other members of the HA binding PG or hyalactan family including also neurocan and brevican. Versican contains cysteine-rich regions at both N- and C-termini and a long potential GAG attachment region in between. It may function as a cell adhesion-mediating molecule during initiation of matrix assembly and is no longer expressed after the deposition of aggrecan in the developing cartilage anlagen.^[5]

Small Proteoglycans

SLRPs are characterized by a central leucine-rich repeat domain and can eventually be subdivided into dermatan/CS and KS-substituted SLRPs. Their variety

provides a broad combinatorial potency and functional specificity.

Biglycan contains two dermatan or CS chains, and decorin contains one dermatan or CS chain. Both have short N-terminal propeptides and contain smaller protein cores than aggrecans.^[3,5] Two potential articular cartilage KS SLRPs are lumican and fibromodulin, which contain up to four KS chains.^[5] Decorin is the most abundant SLRP and is increased during degeneration or senescence, along with fibromodulin and lumican. Like dermatan sulfate SLRPs, fibromodulin is present in articular cartilage throughout life. Fibromodulin contains KS just during chondrogenesis, whereas it is present as a glycoprotein in adults.^[12] On the contrary, lumican is only present in adults, predominantly as a glycoprotein.^[13]

Although these molecules contribute little-to-overall PG mass, tissue volume, or mechanical properties, there may be an equal or greater number of SLRP molecules present compared to large PGs. SLRPs do interact with other macromolecules such as Type II collagen and influence cell function. Decorin, present in the interterritorial matrix and enriched in the superficial zone, and fibromodulin are considered to play a role in organizing and stabilizing the collagen meshwork,^[5] whereas biglycan can be found in the pericellular matrix and may interact with Type VI collagen. One of the most important functions of SLRPs is the ability to bind growth factors and cytokines such as decorin and biglycan to transforming growth factor- β (TGF- β).^[9,14]

Restricted expression of the dermatan sulfate PG epiphygan can be noticed during cartilage maturation.^[5] Owing to its single CS chain on the 2 α chain (NC3 domain), Type IX collagen is also considered to be a PG. Other small PGs remain to be discussed.

Other Proteoglycans

Perlecan is a large heparan/CS hybrid PG present in all basement membranes.^[15] Although cartilage is devoid of basement membranes, perlecan is also present in cartilage and enriched mainly in the pericellular matrix suggesting a role in chondrocyte-matrix attachment.^[16] Its protein core seems to be involved in the control of important biological functions such as cell adhesion, maintenance of the extracellular matrix, and binding to growth factors. The perlecan gene encodes a protein core (~467 kDa) consisting of five distinct domains with multiple functions and a single N-terminal GAG attachment domain. The N-terminal domain is unique to perlecan and may play a role in chondrocyte differentiation and cartilage development,^[17] whereas the other domains share sequence similarities with other

cell surface and extracellular matrix proteins.^[5] Domain 1 contains three potential GAG attachment sites.

The SZP, a ~345 kDa PG, is specifically synthesized by superficial chondrocytes and synoviocytes. It has one CS chain attachment domain and is identical to megakaryocyte stimulating factor.^[8] Similarities or homology to lubricin, which is responsible for joint lubrication and frictionless motion have been described.^[18,19] Moreover, potential growth promoting, cytoprotective, and matrix binding properties associated with certain domains as well as influence on synovial cell proliferation have been suggested.^[8,19]

GLYCOSAMINOGLYCAN AND PROTEOGLYCAN BIOSYNTHESIS

PG biosynthesis is a complex process, with multiple steps including numerous posttranslational modifications that are context dependent. The synthesis of aggrecan, for example, presumably involves more than 30 enzymes,^[5] and is still being explored. Intracellular regulatory systems monitor the concerted synthesis of protein core domains, domain folding, and GAG composition and length. CS can serve as a prototype for the GAG synthesis owing to similarities in sugar sequences, which link to the protein core.

Xylosyltransferase recognizes GAG chain consensus sites on the protein core and catalyzes GAG chain initiation that begins in the late ER and continues into the Golgi complex by addition of xylose to serine residues embedded in a specific peptide sequence. Elongation of the linkage region is catalyzed by three different and specific glycosyltransferases forming a glycosidic bond by transferring sugar residues from a donor nucleotide sugar to the non-reducing end of the accepting sugar. The subsequent repeating polymer is catalyzed by both *N*-acetylgalactosaminyltransferase and glucuronosyltransferase.^[9] Synthesis of heparan sulfate requires additional modifying enzymes such as *N*-deacetylase/*N*-sulfotransferase, glucuronic acid C5 epimerase, iduronic acid 2-O-sulfotransferase, and glucosamine 6-O- and 3-O-sulfotransferases.^[9]

During GAG chain elongation, nascent protein core substrates are traversing intracellular secretory pathways. Cotranslational routing and translocon targeting and passing are complex processes. N-linked mannose-rich oligosaccharides are added cotranslationally to the initial unfolded nascent protein core, which is released into the ER lumen.^[6] Molecular chaperones survey folding and advancing to distal compartments. Until a fully glycosylated and sulfated aggrecan molecule is synthesized, xylosylated precursor protein cores are translocated and undergo further modifications and

trimming in various Golgi compartments where both mono- and oligosaccharides are added.

Sulfation of GAG chains in either 4 or 6 hexosamine position concomitant with or shortly following GAG polymerization completes the complex synthesis of a mature aggrecan molecule.^[6] Finally, motifs are properly folded, certain domains are decorated with GAG chains, and hence aggrecans are exported to the extracellular matrix into unique macromolecular arrays.

HA occurs as a free molecule or as the backbone of PG aggregates. Three different HA synthases (HAS) have been characterized with respect to chromosomal location and amino acid sequence.^[20–23] Mice deficient for HAS2 die early in embryogenesis owing to multiple, most prominently cardiac defects that are compatible with key functions of HAS for epitheliomesenchymal transformations.^[24] HAS proteins are thought to be located largely interior to the cell membrane and have been shown to produce HA chains of different molecular masses.^[25] To date, all three HAS isoforms were found, HAS2 being the most, HAS3 the least abundant mRNA in articular chondrocytes, and HAS1 being highly expressed in synovial cells.^[26] Increased numbers of HAS expressing cells have also been detected in neocartilage following cartilage repair techniques.^[27]

CARTILAGE PROTEOGLYCAN DEGRADATION AND AGING

Biomechanical properties of articular cartilage are inseparably dependent on matrix integrity. Throughout life, cartilage undergoes internal remodeling. Chondrocytes sense matrix alterations, degradation, and mechanical demands, as they are able to respond limitedly by synthesizing and reorganizing appropriate types and amounts of matrix components. Senescence, degenerative diseases, altered joint activity, and post-traumatic disorders decisively alter matrix morphology and chondrocyte function, and remarkably the capacity to maintain and replenish regular PGs and amounts of GAGs.

Decreased staining of PGs in superficial cartilage layers are, among others, the first microscopically detectable changes of degeneration.^[28] In the beginning, the inert macromolecular framework is altered at molecular level.^[3] Simultaneous to collagenous alterations, aggrecan aggregation decreases owing to reduced binding capacity of PTR and HABR to HA molecules.^[4] Aberrant integrity or uncoupling of CD44, HA, and aggrecan deteriorates pericellular matrix homeostasis, matrix turnover, and PG retention, and can even initiate chondrocytic chondrolysis.^[5] GAG chains alter in length and charge density or even lose their connection to the protein

core. This break down can cause swelling, increased permeability, and further vulnerability. Non-aggregated PGs, unbound to HA, and tissue softening are the direct correlates of (early) osteoarthritis. Cs-Szabo et al.^[29] summarized the changes found in osteoarthritic and rheumatoid cartilage. Regular PGs remain in diseased cartilage; however, KS and chondroitin-6-sulfate epitopes of aggrecan were missing and in their place was an increased concentration of chondroitin 4-sulfate and fetal epitopes. Subsequently, molecular pattern is disordered owing to elevated proteolytic activity and impaired PG turnover. Biochemical changes are accompanied by chondrocyte hypertrophy and upregulation of Type X collagen and alkaline phosphatase. Altered biomechanical properties make cartilage even more susceptible to further damage. Anabolic repair, namely matrix neosynthesis associated with cell proliferation and cell clusters can temporarily provide satisfactory stages. Yet, catabolic processes of a variety of proteinases result finally in decline of chondrocytic repair response and destructive progression.^[3]

The detailed mechanisms of degeneration and pathogenesis and the impact of different factors in aging are not yet completely understood and cannot generally be equated with osteoarthritis. Moreover, a broad range of causes for degradation such as genetic predisposition, inflammation and hormonal and metabolic disorders should be taken into account.

Proinflammatory Cytokines

Release of nitric oxide and other proinflammatory cytokines such as interleukin -1 and -6 (IL-1, IL-6), and tumor necrosis factor- α (TNF- α) upregulates expression of several matrix metalloproteinases (MMPs), and inhibits synthesis of their inhibitors namely tissue inhibitors of metalloproteinases (TIMPs) and of PGs such as aggrecans and the SZP.^[30] IL-1 plays one of the crucial roles in this cascade by suppressing matrix synthesis in low concentrations and by stimulating production of proteolytic enzymes in high concentrations.^[31] Thereby, the matrix turnover balance, which is also regulated by bone morphogenetic proteins (BMPs), is shifted from anabolism to catabolism. This may also be owing to decorin, which accumulates during degeneration and can bind anabolic cytokines.^[3]

Recently, work by Poole^[32] includes IL-1 and TNF- α antagonists as a useful therapeutic approach to the management of osteoarthritis. Decreased response to catabolic factors, increased rate of PG turnover and synthesis, and a greater inherent amount of PG explain why some joints are less predisposed to osteoarthritis.^[31,33]

Owing to polymorphisms in the region encoding aggrecan's CS1 domain, individuals with short alleles have low CS content aggrecans and may be more susceptible to degradation.^[34] Moreover, aging is characterized by an increased amount of aggrecan monomers that lack the G3 domain probably owing to proteolytic processing, by declined cell response to anabolic cytokines and the synthesis of smaller aggrecans, fewer functional link proteins and thus irregular PG aggregates.^[11,35]

Proteoglycan Cleavage

Besides MMPs' physiological role in matrix turnover controlled, e.g., by cathepsin B or TIMPs, their over-expression enhances matrix degradation and destabilization of the macromolecular framework. MMPs such as MMP-3 (stromelysin), and aggrecanase, a member of the ADAMTS protein family (a disintegrin and metalloproteinase with thrombospondin motifs), are known to cleave aggrecan at the IGD, KS and CS attachment domains, and even at the carboxyl tail.^[36] How far other PGs are degraded remain unresolved.^[11] Two major aggrecan cleavage sites are located between Asn341-Phe342 and Glu373-Ala374 of the IGD.^[4] Cleavage results in a separation of the G1 domain from the whole GAG attachment region generating specific fragments, which could hence be found in both cartilage and synovial fluid to be useful as markers for tissue destruction.^[4] MMPs additionally affect other matrix molecules resulting in, e.g., fibronectin fragments that are shown to modulate PG synthesis and degradation as well.^[31]

In this context, severe impact or overload may trigger deregulation of MMPs and can cause degradation.^[30] It was shown that static compression can reversibly inhibit synthesis and downregulate gene expression of among other aggrecan core protein and link protein.^[3,31] Immobilization or reduced loading can result in decreased matrix synthesis too.

Additionally, it is believed that other mechanisms, proteolytic and non-proteolytic, cleaving HA renders portions likely contribute to PG fragment release with concomitant fragmentation and release of link protein, smaller HA or irregular molecule complexes. This might be owing to HA depolymerization, reactive oxygen, hyaluronidase, or IL-1 β interactions, which are still questionable.^[7]

SLRPs and Degradation

It was shown that discoordinated mRNA expression and imbalanced matrix accumulation of SLRPs and eventually other aggregating PGs may contribute to osteoarthritic degeneration,^[11,37] even though SLRPs'

role in degeneration is unclear. They show disease- and age-related proteolytic processing resulting in SLRPs increase and accumulation of, e.g., non-glycanated fragments. N-terminal propeptides of decorin and biglycan can potentially be removed by procollagen-C proteinase, however, they appear relatively resistant to cytokine-induced proteolytic modifications.^[8,38]

MEASUREMENT OF GAG CONTENT

GAG assays are useful in the clinical setting to quantify the amount of GAGs synthesized within biological samples. As GAGs are found ubiquitously in the body, GAG assays are also used for cancer (tumor samples), cardiovascular disease (artery samples), Rhesus compatibility, renal (urine), and metabolic disease (urine samples) studies.^[39,40] With respect to orthopedic and cartilage research, GAG assays allow for the determination of efficacy of tissue regeneration techniques by providing a quantifiable measure of cartilage matrix. These assays are proving to be of value in the detection, study, and treatment of osteoarthritis, especially with respect to improved understanding of the pathology of the disease and how it can best be detected and managed.

Histological staining methods are commonly used to demonstrate the presence of GAGs in histological sections. These include viz. Thionin, Safranin O (Fig. 3), Toluidine Blue O, Dimethylmethylene Blue (DMMB), Cuproline Blue, Cupromeronic Blue, and *N,N'*-Diethylpseudoisocyanine. Cartilage is largely made up of water. Just as water molecules are attracted to the negative charges of the sulfate and carboxyl groups of GAGs, these same charges also attract dye molecules, which are usually cations. The different extent of staining or dye uptake is correlated with the amount of GAG present in a tissue sample. As PGs are usually found concentrated around chondrocytes within the extracellular matrix, staining is hence most intense in this region (Fig. 3a₃ and b₃).

Some dyes like Toluidine Blue are metachromic. This is a process of color change induced by the negative charges in GAGs, causing aggregation of dye molecules and subsequent change in dye absorption. Toluidine blue stains nucleic acids blue (the orthochromatic color), but sulfated polysaccharides purple (the metachromatic color).

Using these histological stains, indirect quantification of GAGs in cartilage samples can be estimated. Kiraly et al.^[41] found Thionin and Safranin O to be excellent cationic dyes for the histochemical quantification of cartilage matrix PGs using semiquantitative estimations by microspectrophotometry in sections of articular cartilage. Other groups like O'Driscoll et al.^[42] and Martin et al.^[43] developed their own

in-house computer-based techniques for quantitative analysis of Safranin O stained histological sections, using low magnification light microscopy images. Burkhardt, Hwa, and Ghosh^[44] developed a method for the quantitation of GAG content in 6 μm histological sections that involved direct biochemical analysis of GAGs of unstained sections, which were then compared against a standard scanned image of the same section. This method could be applied when specific information about GAG content of different zones in a joint or body is required. It also allows for subtle changes in PG distribution in different cartilage zones to be detected more readily.

These methods have been developed based on the changes in dye absorption spectrum. Dyes commonly used are Alcian Blue, Toluidine Blue, Azure A, B,^[45] and 1, 9-DMMB. Early methods using these dyes have usually been laborious and complicated. However, modifications have been made over the years to make these processes simpler and easier to use. The main dyes used in GAG assays today usually involve DMMB.

Farndale, Sayers, and Barrett^[46] reported a method using DMMB for rapid estimation of sulfated GAGs in cartilage samples. Cartilage samples are digested with papain and mixed with the dye to allow dye complex formation with sulfated GAGs present in the tissue sample. Absorbance at 535 nm is then measured by spectrophotometer, and the values compared to a standard curve prepared using whale chondroitin 4-sulfate. This general method was later modified to obtain better specificity for sulfated GAG detection.^[47,48] Commercial kits are now available (Blyscan by Biocolor Ltd), which allow for quick and easy estimation of GAGs in tissue samples based on Farndale's technique that had unreliable results at levels of GAG less than 5 $\mu\text{g}/\text{ml}$. Recent studies by Barbosa et al.^[49] have improved this and other assay methods to detect submicrogram quantities of GAGs in samples. This modified method is inexpensive, quick, and specific.

Current biochemical methods also incorporate electrophoretic techniques for improved detection of GAGs in cartilage cultures as described below.^[50,51]

Horner^[52] was the first to describe a use for agarose gel electrophoresis for GAG quantification in 1967. This was adapted to achieve the separation of the major sulfated GAGs produced by cells in culture by Funderburgh and Chandler^[53] in 1978. Volpi and Maccari^[50] subsequently improved the detection of GAGs on agarose gels to the submicrogram level. Their work has included immunodetection using the antichondroitin-6-sulfate antibody that specifically recognizes intact chondroitin-6-sulfate, sequential staining with Toluidine Blue and Stains-All to obtain more sensitive visualization of GAGs with a detection

limit at submicrogram level greater than other conventional procedures.^[51,54] Electrophoresis procedures are generally used for pharmaceutical quality control to detect contaminations of polysaccharides species present in GAG preparations such as heparin.

IMAGING OF GAGs WITH MAGNETIC RESONANCE IMAGING

The use of magnetic resonance imaging (MRI) is a well-established technique for the monitoring of cartilage diseases in clinical routine. However, imaging and evaluation of particular cartilage matrix components was not possible for a long time. The introduction of delayed gadolinium enhanced MRI of cartilage (dGEMRIC) in the 1990s offered a revolutionary, non-invasive technique to monitor the maturation of neocartilage matrix following reconstructive surgery such as autologous chondrocyte implantation/transplantation (ACI/ACT).^[55-57]

GAGs form a dense negatively charged entity within the cartilage matrix. An injected negatively charged MRI contrast agent called gadopentetate dimeglumine [$\text{Gd}(\text{DTPA})^{2-}$] will therefore distribute in areas of GAG depletion (Fig. 4). It penetrates into cartilage from both the joint cavity through the synovial fluid and the subchondral bone. It then shows a distinct concentration-dependent effect on the T1 parameter of MRI. Measurement of T1 serves as an index for the concentration of [$\text{Gd}(\text{DTPA})^{2-}$] and therefore inversely allows for GAG concentration estimation within the cartilage matrix. "Delayed" refers to the fact that it takes time for the contrast agent to penetrate into articular cartilage.^[58-61]

GLYCOSAMINOGLYCAN DISORDERS

GAG disorders cause serious abnormalities in a wide variety of organs as they are ubiquitously present in the body. The majority of these disorders are caused by faulty degradation of PGs such as mucopolysaccharidoses (MPS).^[62] However, some defects in PG biosynthesis also exist. Studies involving experimental animals have shown impaired processing of cartilage PG core protein in nanomelic chickens.^[63] Defective PG biosynthesis in humans is responsible for diseases like macular corneal dystrophy and a variant of spondyloepiphyseal dysplasia.^[64] A Progeroid variant of Ehler Danlos Syndrome has also been described as being owing to defective biosynthesis of dermatan sulfate or decorin.^[65,66] These patients suffer from an aged appearance, short stature, osteopenia, and possess Ehler Danlos syndrome characteristics such as joint hypermobility and muscle hypotonia. Quentin

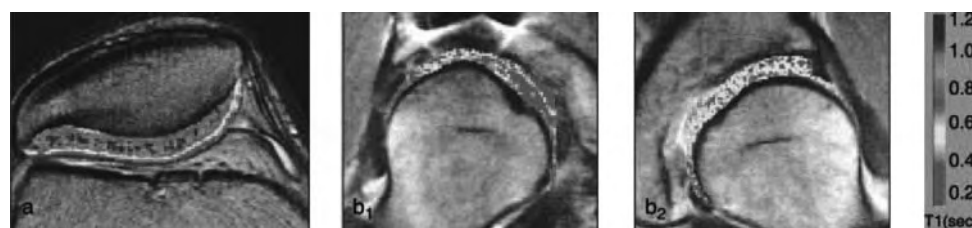


Fig. 4 dGEMRIC. Images postprocessing (usually >1 hr) with contrast distribution map are based on T1 values and allow for quantitative assessments of GAG deposition. Image (a) reflects normal structural GAG distribution in different cartilage zones of the patella. Notice high GAGs/PG concentrations in the middle and deep zone (blue/green) compared to low densities in the superficial zone (yellow/red). In contrast, degenerative diseases such as hip dysplasia (**b₁** and **b₂**) lead gradually to loss of GAG content (yellow/red) and zonal organization.

et al.^[67] found the genetic defect, which results in a deficiency in galactosyltransferase I, which is involved in the synthesis of PG linkage regions.

Lysosomal storage diseases describe a heritable group of heterogeneous human disorders characterized by the accumulation of undigested macromolecules intralysosomally, resulting in an increase in the size and number of these organelles and ultimately in cellular dysfunction and clinical abnormalities. Lysosomes are subcellular organelles responsible for the physiologic turnover of cell constituents containing catabolic enzymes requiring a low optimum pH to function. Lysosomal storage diseases are generally classified by the accumulated substrate and they include sphingolipidoses, glycoproteinases, mucopolisidoses, MPSs, and others.

MPSs result from abnormal degradation of GAGs resulting in organ accumulation and eventual dysfunction. Because GAGs are normal components of the cornea, cartilage, bone, connective tissues, and the reticuloendothelial system, these organs are also targets for excessive GAG storage. Ten known enzyme deficiencies give rise to six distinct MPSs (Table 1). All of these disorders, excepting Hunter's syndrome, are inherited in an autosomal recessive manner.

A major feature of the MPS disorders is abnormal skeletal development leading to short stature, dysostosis multiplex, and degenerative joint disease. Ultrastructurally, the cytoplasm of MPS-affected chondrocytes is filled with membrane-bound vacuoles containing undegraded GAGs. Clusters of hypertrophic chondrocytes form in these disorders, leading to the disruption of the normal cartilage organization. A variety of mutations are described, and correlation of genotype with disease severity is beginning to emerge from mutation analysis.

Simonaro et al.^[68] examined primary cultured chondrocytes and articular cartilage from cats and rats with MPS Type VI (Maroteaux–Lamy disease) and found markedly increased numbers of apoptotic chondrocytes in the MPS animals as compared to age-matched controls. The prevalence of apoptotic cells

was correlated with increased levels of nitrite and TNF- α . Marked PG depletion also was observed in the MPS epiphyses and there was also presence of excess of Type II collagen.

In general, MPS are progressive disorders, characterized by involvement of multiple organs including brain, liver, spleen, heart, and blood vessels, and many are associated with coarse facial features, clouding of the cornea, and mental retardation. Diagnosis can often be made by examination of urine, which reveals increased concentration of GAG fragments.

Enzyme replacement therapies exist for some forms of MPS with the exceptions of MPS III (though trials are ongoing to evaluate efficacy of central nervous system administration of enzyme), MPS IV, and MPS VII. Aldurazyme (Laronidase) is an approved once weekly treatment for the non-neurological symptoms of MPS I after studies demonstrated its efficacy and safety in 2004.^[69] Phase III trials using human enzyme replacement I2S have been completed with good outcomes for MPS II (Hunter's). The drug is currently awaiting Food and Drug Approval (FDA) approval in

Table 1 Summary of known enzyme deficiencies giving rise to six distinct MPS groups

MPS I	Hurler syndrome Schele syndrome Hurler–Schele syndrome
MPS II	Hunter syndrome (mild or severe forms)
MPS III	San Filippo syndrome (A, B, C or D)
MPS IV	Morquio syndrome (A or B)
MPS VI	Maroteaux–Lamy syndrome (mild, intermediate or classic severe)
MPS VII	Sly syndrome

Diseases lead to abnormal degradation of GAGs and become apparent by accumulation in organs, and can result in dysfunction and eventually failure. GAGs are ubiquitous components within the body; therefore, various tissues and organs are targets for excessive GAG storage owing to MPS. All of these disorders, except Hunter's syndrome, are inherited in an autosomal recessive manner.

U.S.A. Naglazyme has recent FDA approval for use in the treatment of MPS VI (Maroteaux–Lamy).

CLINICAL EFFICACY OF CS, GLUCOSAMINE, AND HYALURONAN IN THE TREATMENT OF OSTEOARTHRITIS

As explained in detail previously, GAGs are key components of extracellular articular cartilage matrix. They undergo constant degradation, removal, and replenishing.

In the disease process of osteoarthritis, the equilibrium of GAG degradation and resynthesis are heavily disturbed. Insufficient GAG synthesis (either of lower quality or insufficient quantity) ultimately causes overall cartilage degradation owing to altered matrix composition.^[70]

Glucosamine plays an important role in the synthesis of two of the most abundant GAGs, namely CS and HA. Therefore, glucosamine as well as CS and HA has been used in the past as dietary supplements for the treatment of “cartilage diseases.” This “new way to treat osteoarthritis” was initiated based on the theory that loss or insufficient synthesis of these key molecules in cartilage could simply be corrected by oral supplementation or even intra-articular injections.^[71,72] These nutraceuticals are commonly extracted from animal tissues, namely beef trachea.^[70]

Over time, many concerns from researchers in the field of orthopedics and rheumatology arose, questioning the efficacy of these new “drugs.” Many industry-funded studies, which were published in questionable journals stated highly positive effects and even postulated the potency to restore damaged articular cartilage and stop the disease process early on. A book entitled “The Arthritis Cure” finally provoked the initiation of serious studies in the field for the evaluation of the supplements efficacy.^[73–75]

Whereas in some European countries, these nutraceuticals have already achieved approval by the drug administration authorities, FDA approval has not been issued yet for any of these dietary supplements which are readily available “over-the-counter.”^[70,72]

The bioavailability and pharmacokinetics of glucosamine and CS were among the initial concerns raised by researchers over the use of these supplements. Various studies have investigated serum levels of glucosamine and CS after oral ingestion. Results to date are still controversial and published pharmacokinetic information appears to be incomplete. Oral daily intake of 1500 mg glucosamine seems to be absorbed to up to 87%. A single daily dose is recommended.^[76] Almost 40% of the unaltered molecule seems to be excreted through the urinary system. Additionally, a high first pass metabolism in the liver is apparent.

Nevertheless, radiolabeled glucosamine has been detected in all joint tissues following oral administration.^[70,77–82]

CS’s gastrointestinal absorption rate appears to be very much lower compared to glucosamine’s. However, with only 10–15% absorption after oral administration it can still be detected in human joint tissues. The recommended oral daily intake shall be 1500–2000 mg, based on recent bioavailability studies.^[70,78,81,83]

Because these nutraceuticals are not regulated in U.S.A., the quality of the products appears to be highly variable. The European market offers few products approved by drug administration agencies. The actual content of chondroitin and glucosamine sulfate in nutraceuticals may be significantly different from what product labels claim. According to a study of Adebowale et al.,^[84] deviations from label specifications ranged from 0–115%. These results again underline the need for a quality control and regulation of these products.

Glucosamine has been reported to inhibit disease progressing mediators such as IL-1, which are known to increase aggrecanase activity. It could also be shown that it increases synthesis of PGs in chondrocyte monolayer cultures.^[70]

Numerous studies on clinical improvement of symptoms after oral administration of glucosamine/CS in patients suffering from osteoarthritis have been published within the last decade. Almost all of the early studies until 2000 have shown positive results.^[70,73] A meta-analysis published in 2000 analyzed these studies and reported that the majority of these studies were of low or even unacceptable quality. Higher and high-quality studies showed lower significance levels and more recent, independent studies even null or negative effects.^[70,73,74,85] McAlindon et al. reported that the manufacturer supported most of the studies reporting positive effects. Reviewing these studies, different authors came to the conclusion that oral supplementation of CS and glucosamine should be part of the standard non-surgical armamentarium of an orthopedic surgeon for the symptomatic treatment of osteoarthritis.^[86–88] Skepticism arose especially from experts in the field and they called for more rigorous trials.^[70,73]

In fact, research in the field did greatly intensify within the last few years: Richy^[89] evaluated 500 publications from 1980 to the current which investigated the morphological and symptomatic efficacy of glucosamine and CS with 15 of them matching his quality criteria for randomized, double-blind, placebo-controlled studies. He found both significant morphological and symptomatic efficacy for glucosamine compared to placebo. CS instead showed only significant symptomatic efficacy. Both oral supplements were

tolerated with almost no major adverse effects, demonstrating the safety of these drugs.

McAlindon et al.^[74] performed an internet study among 205 individuals with osteoarthritis. They concluded that glucosamine has no more clinical efficacy than placebo for the treatment of osteoarthritis.

A discontinuation trial performed by Cibere et al. studied the effects of discontinuation of glucosamine in osteoarthritis patients who were already taking these supplements. The design was placebo-controlled, randomized, and double-blind over 6 mo and included four centers with 137 patients. No significant differences were found in disease flare between the placebo and glucosamine group.^[90] They also studied the effect of glucosamine discontinuation on serum and urine Type II collagen degradation fragments in the same patient group. Again, statistically no significant effect was found.^[91]

Another highly interesting recent study investigated the effect of CS on the inhibition of joint space narrowing on osteoarthritis patients. Three hundred patients were recruited and were given a daily oral dose of 800 mg of CS over 2 yr. Loss of joint space as well as pain and function measures was recorded. Again, no symptomatic or inhibitory efficacy could be revealed.^[92]

Conclusively, it still remains to be shown whether or not these nutraceuticals should have a place in the symptomatic, non-surgical treatment of osteoarthritis. In fact, results from a large multicenter trial are awaited shortly.

The term “viscosupplementation” describes the intra-articular injection of HA, a technique that has been performed since the 1970s.^[93] This treatment option for osteoarthritis reached FDA approval and has also been recommended by the American College of Rheumatology as an alternative for pain relief in osteoarthritis. Currently, two preparations are available and approved for the treatment of pain associated with knee osteoarthritis in U.S.A.: Hylan G-F20 (Synvisc[®]) and sodium hyaluronate (Hyalgan[®] and Supartz[™]). They differ greatly in molecular weights and concentrations.

The mechanisms by which intra-articular HA unfolds its effects still remain cryptic. HA half-life within the joint cavity is known to be short (15–36 hr); however, its effects seem to last much longer.^[72]

Commercial preparations are not identical with the body's own HA but chemically altered to improve viscoelastic properties and extend the molecules' half-life within the joint. In fact, different preparations with various molecular weights were compared to placebo or corticosteroid injections and efficacy could be shown for all of the different molecular weights.^[72,94,95]

Still, some studies do not support the opinion that HA injections truly resemble a worthy treatment

option.^[96] It could be demonstrated that this treatment has no measurable effects.^[97] However, Lo et al. found that intra-articular HA injections have a small effect with respect to pain relief when compared to placebo. They also concluded that high and highest molecular weight HA preparations may be more efficacious than low molecular weight HA.^[98] Additional studies suggest that intra-articular application of HA may inhibit or decelerate structural degradation of articular cartilage. Jubb et al. found a slower disease progression in patients with a radiologically milder disease at baseline. However, overall treatment effect after one year could not be observed.^[99]

In addition, the long-term effectiveness of HA is still unknown, and there is a great need for studies revealing prolonged effects and possible side effects.^[72] So far, no reports about significant adverse effects have been reported. Nevertheless, intra-articular injections always bear the risk of causing joint cavity empyemas.

Overall, viscosupplementation seems to be an alternative treatment with beneficial effects for patients suffering from therapy-resistant knee pain owing to osteoarthritic changes. However, contradictory results have been reported, and structural, disease progress inhibiting effects have not been shown to date.

CONCLUSIONS

GAGs are a very diverse family of molecules with multiple physiological functions. Variations in structure and modifications to the backbone such as sulfation add to the complexity of GAGs and may serve to regulate cell fate and function. It is not surprising, therefore, that GAGs are implicated in many diseases and targets of therapeutic approaches. GAGs are especially important for normal joint development and homeostasis. Because GAGs are abundant in articular cartilage, they are commonly used as indicators of cartilage integrity and chondrocyte activity. CS and glucosamine have become very popular nutritional supplements for joint health, although the scientific evidence regarding the efficacy of these supplements remains controversial. In addition, GAGs have been employed in a number of cartilage repair and tissue-engineering techniques. Although there is an abundance of information accumulating about GAGs, there is clearly much more to learn about the role of GAGs in normal tissue and disease and to develop their clinical potential.

ACKNOWLEDGMENTS

We would like to thank Dr. M. Yaszemski and collaborators from the Bone Histomorphometry



Laboratory (D. Jewison, J. Burgess, and L. Peterson) for histological contributions and for using the Exakt™ system, and to J. Fitzsimmons for technical and histological support. Fig. 4 was kindly donated by Drs. D. Burstein and K. Amrami from the Department of Radiology (Mayo Clinic College of Medicine, Rochester, Minnesota, U.S.A.).

ARTICLES OF FURTHER INTEREST

Tissue Engineering of Cartilage; Tissue Engineering of Skin

REFERENCES

- Li, J.P.; Galvis, M.L.; Gong, F.; Zhang, X.; Zcharia, E.; Metzger, S.; Vlodaysky, I.; Kisilevsky, R.; Lindahl, U. In vivo fragmentation of heparan sulfate by heparanase overexpression renders mice resistant to amyloid protein A amyloidosis. *Proc. Natl. Acad. Sci. USA* **2005**, *102* (18), 6473–6477.
- Gama, C.I.; Hsieh-Wilson, L.C. Chemical approaches to deciphering the glycosaminoglycan code. *Curr. Opin. Chem. Biol.* **2005**, *9* (6), 609–619.
- Buckwalter, J.A.; Mankin, H.J.; Grodzinsky, A.J. Articular cartilage and osteoarthritis. *Instr. Course Lect.* **2005**, *54*, 465–480.
- Kiani, C.; Chen, L.; Wu, Y.J.; Yee, A.J.; Yang, B.B. Structure and function of aggrecan. *Cell Res.* **2002**, *12* (1), 19–32.
- Knudson, C.B.; Knudson, W. Cartilage proteoglycans. *Semin. Cell Dev. Biol.* **2001**, *12* (2), 69–78.
- Luo, W.; Guo, C.; Zheng, J.; Chen, T.L.; Wang, P.Y.; Vertel, B.M.; Tanzer, M.L. Aggrecan from start to finish. *J. Bone Miner. Metab.* **2000**, *18* (2), 51–56.
- Sztrolovics, R.; Recklies, A.D.; Roughley, P.J.; Mort, J.S. Hyaluronate degradation as an alternative mechanism for proteoglycan release from cartilage during interleukin-1 β -stimulated catabolism. *Biochem. J.* **2002**, *362* (Pt 2), 473–479.
- Roughley, P.J. Articular cartilage and changes in arthritis: noncollagenous proteins and proteoglycans in the extracellular matrix of cartilage. *Arthritis Res.* **2001**, *3* (6), 342–347.
- Schwartz, N. Biosynthesis and regulation of expression of proteoglycans. *Front. Biosci.* **2000**, *5*, D649–655.
- Shi, S.; Grothe, S.; Zhang, Y.; O'Connor-McCourt, M.D.; Poole, A.R.; Roughley, P.J.; Mort, J.S. Link protein has greater affinity for versican than aggrecan. *J. Biol. Chem.* **2004**, *279* (13), 12,060–12,066.
- Sztrolovics, R.; Grover, J.; Cs-Szabo, G.; Shi, S.L.; Zhang, Y.; Mort, J.S.; Roughley, P.J. The characterization of versican and its message in human articular cartilage and intervertebral disc. *J. Orthop. Res.* **2002**, *20* (2), 257–266.
- Roughley, P.J.; White, R.J.; Cs-Szabo, G.; Mort, J.S. Changes with age in the structure of fibromodulin in human articular cartilage. *Osteoarth. Cartil.* **1996**, *4* (3), 153–161.
- Grover, J.; Chen, X.N.; Korenberg, J.R.; Roughley, P.J. The human lumican gene. Organization, chromosomal location, and expression in articular cartilage. *J. Biol. Chem.* **1995**, *270* (37), 21,942–21,949.
- Hildebrand, A.; Romaris, M.; Rasmussen, L.M.; Heinegard, D.; Twardzik, D.R.; Border, W.A.; Ruoslahti, E. Interaction of the small interstitial proteoglycans biglycan, decorin and fibromodulin with transforming growth factor beta. *Biochem. J.* **1994**, *302* (Pt 2), 527–534.
- Tesche, F.; Miosge, N. Perlecan in late stages of osteoarthritis of the human knee joint. *Osteoarth. Cartil.* **2004**, *12* (11), 852–862.
- SundarRaj, N.; Fite, D.; Ledbetter, S.; Chakravarti, S.; Hassell, J.R. Perlecan is a component of cartilage matrix and promotes chondrocyte attachment. *J. Cell Sci.* **1995**, *108* (Pt 7), 2663–2672.
- French, M.M.; Gomes, R.R., Jr.; Timpl, R.; Hook, M.; Czymmek, K.; Farach-Carson, M.C.; Carson, D.D. Chondrogenic activity of the heparan sulfate proteoglycan perlecan maps to the N-terminal domain I. *J. Bone Miner. Res.* **2002**, *17* (1), 48–55.
- Jay, G.D.; Tantravahi, U.; Britt, D.E.; Barrach, H.J.; Cha, C.J. Homology of lubricin and superficial zone protein (SZP): products of megakaryocyte stimulating factor (MSF) gene expression by human synovial fibroblasts and articular chondrocytes localized to chromosome 1q25. *J. Orthop. Res.* **2001**, *19* (4), 677–687.
- Flannery, C.R.; Hughes, C.E.; Schumacher, B.L.; Tudor, D.; Aydelotte, M.B.; Kuettner, K.E.; Caterson, B. Articular cartilage superficial zone protein (SZP) is homologous to megakaryocyte stimulating factor precursor and is a multifunctional proteoglycan with potential growth-promoting, cytoprotective, and lubricating properties in cartilage metabolism. *Biochem. Biophys. Res. Commun.* **1999**, *254* (3), 535–541.
- Spicer, A.P.; Augustine, M.L.; McDonald, J.A. Molecular cloning and characterization of a putative mouse hyaluronan synthase. *J. Biol. Chem.* **1996**, *271* (38), 23,400–23,406.
- Spicer, A.P.; Seldin, M.F.; Olsen, A.S.; Brown, N.; Wells, D.E.; Doggett, N.A.; Itano, N.; Kimata, K.; Inazawa, J.; McDonald, J.A. Chromosomal localization of the human and mouse hyaluronan synthase genes. *Genomics* **1997**, *41* (3), 493–497.
- Spicer, A.P.; Olson, J.S.; McDonald, J.A. Molecular cloning and characterization of a cDNA encoding the third putative mammalian hyaluronan synthase. *J. Biol. Chem.* **1997**, *272* (14), 8957–8961.
- Spicer, A.P.; McDonald, J.A. Characterization and molecular evolution of a vertebrate hyaluronan synthase gene family. *J. Biol. Chem.* **1998**, *273* (4), 1923–1932.
- Camenisch, T.D.; Spicer, A.P.; Brehm-Gibson, T.; Biesterfeldt, J.; Augustine, M.L.; Calabro, A., Jr.; Kubalak, S.; Klewer, S.E.; McDonald, J.A. Disruption

- of hyaluronan synthase-2 abrogates normal cardiac morphogenesis and hyaluronan-mediated transformation of epithelium to mesenchyme. *J. Clin. Invest.* **2000**, *106* (3), 349–360.
25. Itano, N.; Sawai, T.; Yoshida, M.; Lenas, P.; Yamada, Y.; Imagawa, M.; Shinomura, T.; Hamaguchi, M.; Yoshida, Y.; Ohnuki, Y.; Miyauchi, S.; Spicer, A.P.; McDonald, J.A.; Kimata, K. Three isoforms of mammalian hyaluronan synthases have distinct enzymatic properties. *J. Biol. Chem.* **1999**, *274* (35), 25,085–25,092.
26. Recklies, A.D.; White, C.; Melching, L.; Roughley, P.J. Differential regulation and expression of hyaluronan synthases in human articular chondrocytes, synovial cells and osteosarcoma cells. *Biochem. J.* **2001**, *354* (Pt 1), 17–24.
27. Mrosek, E.; Erggelet, C.; McDonald, J.A.; Kurz, H. Hyaluronan synthases in normal and regenerating joint cartilage. *Cells Tissues Organs* **2003**, *173* (2), 93–104.
28. Mrosek, E.H.; Lahm, A.; Erggelet, C.; Uhl, M.; Kurz, H.; Eissner, B.; Schagemann, J.C. Subchondral bone trauma causes cartilage matrix degeneration: an immunohistochemical analysis in a canine model. *Osteoarth. Cartil.* **2006**, *14* (2), 171–178.
29. Cs-Szabo, G.; Roughley, P.J.; Plaas, A.H.; Glant, T.T. Large and small proteoglycans of osteoarthritic and rheumatoid articular cartilage. *Arthritis Rheum.* **1995**, *38* (5), 660–668.
30. Pearle, A.D.; Warren, R.F.; Rodeo, S.A. Basic science of articular cartilage and osteoarthritis. *Clin. Sports Med.* **2005**, *24* (1), 1–12.
31. Kuettner, K.E.; Cole, A.A. Cartilage degeneration in different human joints. *Osteoarth. Cartil.* **2005**, *13* (2), 93–103.
32. Poole, A.R. Biochemical/immunochemical biomarkers of osteoarthritis: utility for prediction of incident or progressive osteoarthritis. *Rheum. Dis. Clin. North Am.* **2003**, *29* (4), 803–818.
33. Kuettner, K.E. Biochemistry of articular cartilage in health and disease. *Clin. Biochem.* **1992**, *25* (3), 155–163.
34. Doege, K.J.; Coulter, S.N.; Meek, L.M.; Maslen, K.; Wood, J.G. A human-specific polymorphism in the coding region of the aggrecan gene. Variable number of tandem repeats produce a range of core protein sizes in the general population. *J. Biol. Chem.* **1997**, *272* (21), 13,974–13,979.
35. Martin, J.A.; Brown, T.D.; Heiner, A.D.; Buckwalter, J.A. Chondrocyte senescence, joint loading and osteoarthritis. *Clin. Orthop. Relat. Res.* **2004**, *427* (Suppl), S96–103.
36. Lee, V.; Chen, L.; Paiwand, F.; Cao, L.; Wu, Y.; Inman, R.; Adams, M.E.; Yang, B.B. Cleavage of the carboxyl tail from the G3 domain of aggrecan but not versican and identification of the amino acids involved in the degradation. *J. Biol. Chem.* **2002**, *277* (25), 22,279–22,288.
37. Cs-Szabo, G.; Melching, L.I.; Roughley, P.J.; Glant, T.T. Changes in messenger RNA and protein levels of proteoglycans and link protein in human osteoarthritic cartilage samples. *Arthritis Rheum.* **1997**, *40* (6), 1037–1045.
38. Sztrolovics, R.; White, R.J.; Poole, A.R.; Mort, J.S.; Roughley, P.J. Resistance of small leucine-rich repeat proteoglycans to proteolytic degradation during interleukin-1-stimulated cartilage catabolism. *Biochem. J.* **1999**, *339* (Pt 3), 571–577.
39. Whiteman, P.; Henderson, H. A method for the determination of amniotic-fluid glycosaminoglycans and its application to the prenatal diagnosis of Hurler and Sanfilippo diseases. *Clin. Chim. Acta.* **1977**, *79* (1), 99–105.
40. Danes, B.S.; Queenan, J.T.; Gadow, E.C.; Cederqvist, L.L. Antenatal diagnosis of mucopolysaccharidoses. *Lancet* **1970**, *1* (7653), 946–947.
41. Kiraly, K.; Lapvetelainen, T.; Arokoski, J.; Torronen, K.; Modis, L.; Kiviranta, I.; Helminen, H.J. Application of selected cationic dyes for the semiquantitative estimation of glycosaminoglycans in histological sections of articular cartilage by microspectrophotometry. *Histochem. J.* **1996**, *28* (8), 577–590.
42. O'Driscoll, S.W.; Marx, R.G.; Fitzsimmons, J.S.; Beaton, D.E. Method for automated cartilage histomorphometry. *Tissue Eng.* **1999**, *5* (1), 13–23.
43. Martin, I.; Obradovic, B.; Freed, L.E.; Vunjak-Novakovic, G. Method for quantitative analysis of glycosaminoglycan distribution in cultured natural and engineered cartilage. *Ann. Biomed. Eng.* **1999**, *27* (5), 656–662.
44. Burkhardt, D.; Hwa, S.Y.; Ghosh, P. A novel microassay for the quantitation of the sulfated glycosaminoglycan content of histological sections: its application to determine the effects of Diacerhein on cartilage in an ovine model of osteoarthritis. *Osteoarth. Cartil.* **2001**, *9* (3), 238–247.
45. Thuy, L.P.; Nyhan, W.L. A new quantitative assay for glycosaminoglycans. *Clin. Chim. Acta.* **1992**, *212* (1–2), 17–26.
46. Farndale, R.W.; Sayers, C.A.; Barrett, A.J. A direct spectrophotometric microassay for sulfated glycosaminoglycans in cartilage cultures. *Connect. Tissue Res.* **1982**, *9* (4), 247–248.
47. Farndale, R.W.; Buttle, D.J.; Barrett, A.J. Improved quantitation and discrimination of sulphated glycosaminoglycans by use of dimethylmethylene blue. *Biochim. Biophys. Acta.* **1986**, *883* (2), 173–177.
48. Chandrasekhar, S.; Esterman, M.A.; Hoffman, H.A. Microdetermination of proteoglycans and glycosaminoglycans in the presence of guanidine hydrochloride. *Anal. Biochem.* **1987**, *161* (1), 103–108.
49. Barbosa, I.; Garcia, S.; Barbier-Chassefiere, V.; Caruelle, J.P.; Martelly, I.; Papy-Garcia, D. Improved and simple micro assay for sulfated glycosaminoglycans quantification in biological extracts and its use in skin and muscle tissue studies. *Glycobiology* **2003**, *13* (9), 647–653.
50. Volpi, N.; Maccari, F. Detection of submicrogram quantities of glycosaminoglycans on agarose gels by sequential staining with toluidine blue and Stains-All. *Electrophoresis* **2002**, *23* (24), 4060–4066.
51. Volpi, N.; Maccari, F.; Titze, J. Simultaneous detection of submicrogram quantities of hyaluronic acid and dermatan sulfate on agarose-gel by sequential staining with

- toluidine blue and Stains-All. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **2005**, *820* (1), 131–135.
52. Horner, A.A. Electrophoresis of acidic mucopolysaccharides in agarose gel. *Can. J. Biochem.* **1967**, *45* (7), 1009–1013.
53. Funderburgh, J.L.; Chandler, J.W. An agarose gel electrophoretic method for analysis of sulfated glycosaminoglycans of cultured cells. *Anal. Biochem.* **1978**, *91* (2), 464–472.
54. Maccari, F.; Volpi, N. Direct and specific recognition of glycosaminoglycans by antibodies after their separation by agarose gel electrophoresis and blotting on cetylpyridinium chloride-treated nitrocellulose membranes. *Electrophoresis* **2003**, *24* (9), 1347–1352.
55. Bashir, A.; Gray, M.L.; Boutin, R.D.; Burstein, D. Glycosaminoglycan in articular cartilage: in vivo assessment with delayed Gd(DTPA)(2-)-enhanced MR imaging. *Radiology* **1997**, *205* (2), 551–558.
56. Bashir, A.; Gray, M.L.; Burstein, D. Gd-DTPA2- as a measure of cartilage degradation. *Magn. Reson. Med.* **1996**, *36* (5), 665–673.
57. Bashir, A.; Gray, M.L.; Hartke, J.; Burstein, D. Non-destructive imaging of human cartilage glycosaminoglycan concentration by MRI. *Magn. Reson. Med.* **1999**, *41* (5), 857–865.
58. Tiderius, C.J.; Olsson, L.E.; Leander, P.; Ekberg, O.; Dahlberg, L. Delayed gadolinium-enhanced MRI of cartilage (dGEMRIC) in early knee osteoarthritis. *Magn. Reson. Med.* **2003**, *49* (3), 488–492.
59. Williams, A.; Gillis, A.; McKenzie, C.; Po, B.; Sharma, L.; Micheli, L.; McKeon, B.; Burstein, D. Glycosaminoglycan distribution in cartilage as determined by delayed gadolinium-enhanced MRI of cartilage (dGEMRIC): potential clinical applications. *AJR. Am. J. Roentgenol.* **2004**, *182* (1), 167–172.
60. Gillis, A.; Bashir, A.; McKeon, B.; Scheller, A.; Gray, M.L.; Burstein, D. Magnetic resonance imaging of relative glycosaminoglycan distribution in patients with autologous chondrocyte transplants. *Invest. Radiol.* **2001**, *36* (12), 743–748.
61. Gillis, A.; Gray, M.; Burstein, D. Relaxivity and diffusion of gadolinium agents in cartilage. *Magn. Reson. Med.* **2002**, *48* (6), 1068–1071.
62. Neufeld, E.F. Lysosomal storage diseases. *Annu. Rev. Biochem.* **1991**, *60*, 257–280.
63. O'Donnell, C.M.; Kaczman-Daniel, K.; Goetinck, P.F.; Vertel, B.M. Nanomelic chondrocytes synthesize a glycoprotein related to chondroitin sulfate proteoglycan core protein. *J. Biol. Chem.* **1988**, *263* (33), 17,749–17,754.
64. Mourao, P.A.; Kato, S.; Donnelly, P.V. Spondyloepiphyseal dysplasia, chondroitin sulfate type: a possible defect of PAPS—chondroitin sulfate sulfotransferase in humans. *Biochem. Biophys. Res. Commun.* **1981**, *98* (2), 388–396.
65. Yamaguchi, Y.; Ruoslahti, E. Expression of human proteoglycan in Chinese hamster ovary cells inhibits cell proliferation. *Nature* **1988**, *336* (6196), 244–246.
66. Rosenberg, L.C.; Choi, H.U.; Tang, L.H.; Johnson, T.L.; Pal, S.; Webber, C.; Reiner, A.; Poole, A.R. Isolation of dermatan sulfate proteoglycans from mature bovine articular cartilages. *J. Biol. Chem.* **1985**, *260* (10), 6304–6313.
67. Quentin, E.; Gladen, A.; Roden, L.; Kresse, H. A genetic defect in the biosynthesis of dermatan sulfate proteoglycan: galactosyltransferase I deficiency in fibroblasts from a patient with a progeroid syndrome. *Proc. Natl. Acad. Sci. USA* **1990**, *87* (4), 1342–1346.
68. Simonaro, C.M.; D'Angelo, M.; Haskins, M.E.; Schuchman, E.H. Joint and bone disease in mucopolysaccharidoses VI and VII: identification of new therapeutic targets and biomarkers using animal models. *Pediatr. Res.* **2005**, *57* (5 Pt 1), 701–707.
69. Wraith, J.E.; Clarke, L.A.; Beck, M.; Kolodny, E.H.; Pastores, G.M.; Muenzer, J.; Rapoport, D.M.; Berger, K.I.; Swiedler, S.J.; Kakkis, E.D.; Braakman, T.; Chadbourne, E.; Walton-Bowen, K.; Cox, G.F. Enzyme replacement therapy for mucopolysaccharidosis I: a randomized, double-blinded, placebo-controlled, multinational study of recombinant human alpha-L-iduronidase (laronidase). *J. Pediatr.* **2004**, *144* (5), 581–588.
70. Felson, D.T.; McAlindon, T.E. Glucosamine and chondroitin for osteoarthritis: to recommend or not to recommend? *Arthritis Care Res.* **2000**, *13* (4), 179–182.
71. Lyman, S.; Sherman, S.; Dunn, W.R.; Marx, R.G. Advancements in the surgical and alternative treatment of arthritis. *Curr. Opin. Rheumatol.* **2005**, *17* (2), 129–133.
72. Fajardo, M.; Di Cesare, P.E. Disease-modifying therapies for osteoarthritis: current status. *Drugs Aging* **2005**, *22* (2), 141–161.
73. McAlindon, T.E.; LaValley, M.P.; Gulin, J.P.; Felson, D.T. Glucosamine and chondroitin for treatment of osteoarthritis: a systematic quality assessment and meta-analysis. *Jama* **2000**, *283* (11), 1469–1475.
74. McAlindon, T.; Formica, M.; LaValley, M.; Lehmer, M.; Kabbara, K. Effectiveness of glucosamine for symptoms of knee osteoarthritis: results from an internet-based randomized double-blind controlled trial. *Am. J. Med.* **2004**, *117* (9), 643–649.
75. McAlindon, T.E.; Biggee, B.A. Nutritional factors and osteoarthritis: recent developments. *Curr. Opin. Rheumatol.* **2005**, *17* (5), 647–652.
76. Persiani, S.; Roda, E.; Rovati, L.C.; Locatelli, M.; Giacobelli, G.; Roda, A. Glucosamine oral bioavailability and plasma pharmacokinetics after increasing doses of crystalline glucosamine sulfate in man. *Osteoarth. Cartil.* **2005**, *13* (12), 1041–1049.
77. Setnikar, I.; Giacchetti, C.; Zanollo, G. Pharmacokinetics of glucosamine in the dog and in man. *Arzneimittelforschung* **1986**, *36* (4), 729–735.
78. Deal, C.L.; Moskowitz, R.W. Nutraceuticals as therapeutic agents in osteoarthritis. The role of glucosamine, chondroitin sulfate, and collagen hydrolysate. *Rheum. Dis. Clin. North Am.* **1999**, *25* (2), 379–395.
79. Biggee, B.A.; McAlindon, T. Glucosamine for osteoarthritis: part II, biologic and metabolic controversies. *Med. Health R I.* **2004**, *87* (6), 180–181.
80. Biggee, B.A.; Blinn, C.M.; McAlindon, T.E.; Nuite, M.; Silbert, J.E. Low levels of human serum glucosamine after ingestion of glucosamine sulphate relative to capability for peripheral effectiveness. *Ann. Rheum. Dis.* **2006**, *65* (2), 222–226.

81. Adebowale, A.; Du, J.; Liang, Z.; Leslie, J.L.; Eddington, N.D. The bioavailability and pharmacokinetics of glucosamine hydrochloride and low molecular weight chondroitin sulfate after single and multiple doses to beagle dogs. *Biopharm. Drug Dispos.* **2002**, *23* (6), 217–225.
82. Biggee, B.A.; Blinn, C.M.; McAlindon, T.E. Human serum glucosamine and sulfate levels after ingestion of glucosamine sulfate. *Arthritis Rheum.* **2004**, *50*, 657.
83. Conte, A.; Volpi, N.; Palmieri, L.; Bahous, I.; Ronca, G. Biochemical and pharmacokinetic aspects of oral treatment with chondroitin sulfate. *Arzneimittelforschung* **1995**, *45* (8), 918–925.
84. Adebowale, A.O.; Cox, D.S.; Liang, Z.; Eddington, N.D. Analysis of glucosamine and chondroitin sulfate content in marketed products and the Caco-2 permeability of chondroitin sulfate raw materials. *J. Am. Nutraceutical Assoc.* **2000**, *3*, 37–44.
85. Rindone, J.P.; Hiller, D.; Collacott, E.; Nordhaugen, N.; Arriola, G. Randomized, controlled trial of glucosamine for treating osteoarthritis of the knee. *West J. Med.* **2000**, *172* (2), 91–94.
86. Hungerford, D.S.; Jones, L.C. Glucosamine and chondroitin sulfate are effective in the management of osteoarthritis. *J. Arthroplasty* **2003**, *18* (3 Suppl 1), 5–9.
87. Morelli, V.; Naquin, C.; Weaver, V. Alternative therapies for traditional disease states: osteoarthritis. *Am. Fam. Physician.* **2003**, *67* (2), 339–344.
88. Hungerford, D.; Navarro, R.; Hammad, T. Use of nutraceuticals in the management of osteoarthritis. *J. Am. Nutraceutical Assoc.* **2000**, *3*, 23–27.
89. Richy, F.; Bruyere, O.; Ethgen, O.; Cucherat, M.; Henrotin, Y.; Reginster, J.Y. Structural and symptomatic efficacy of glucosamine and chondroitin in knee osteoarthritis: a comprehensive meta-analysis. *Arch. Intern. Med.* **2003**, *163* (13), 1514–1522.
90. Cibere, J.; Kopec, J.A.; Thorne, A.; Singer, J.; Canvin, J.; Robinson, D.B.; Pope, J.; Hong, P.; Grant, E.; Esdaile, J.M. Randomized, double-blind, placebo-controlled glucosamine discontinuation trial in knee osteoarthritis. *Arthritis Rheum.* **2004**, *51* (5), 738–745.
91. Cibere, J.; Thorne, A.; Kopec, J.A.; Singer, J.; Canvin, J.; Robinson, D.B.; Pope, J.; Hong, P.; Grant, E.; Lobanok, T.; Ionescu, M.; Poole, A.R.; Esdaile, J.M. Glucosamine sulfate and cartilage type II collagen degradation in patients with knee osteoarthritis: randomized discontinuation trial results employing biomarkers. *J. Rheumatol.* **2005**, *32* (5), 896–902.
92. Michel, B.A.; Stucki, G.; Frey, D.; De Vathaire, F.; Vignon, E.; Bruehlmann, P.; Uebelhart, D. Chondroitins 4 and 6 sulfate in osteoarthritis of the knee: a randomized, controlled trial. *Arthritis Rheum.* **2005**, *52* (3), 779–786.
93. Altman, R.D. Status of hyaluronan supplementation therapy in osteoarthritis. *Curr. Rheumatol. Rep.* **2003**, *5* (1), 7–14.
94. Kirchner, M.; Marshall, D. A double-blind randomized controlled trial comparing alternate forms of high molecular weight hyaluronan for the treatment of osteoarthritis of the knee. *Osteoarth. Cartil.* **2006**, *14* (2), 154–162.
95. Fuchs, S.; Monikes, R.; Wohlmeiner, A.; Heyse, T. Intra-articular hyaluronic acid compared with corticoid injections for the treatment of rhizarthrosis. *Osteoarth. Cartil.* **2005**, **2006**, *14* (1), 82–88.
96. Felson, D.T.; Anderson, J.J. Hyaluronate sodium injections for osteoarthritis: hope, hype, and hard truths. *Arch. Intern. Med.* **2002**, *162* (3), 245–247.
97. Henderson, E.B.; Smith, E.C.; Pegley, F.; Blake, D.R. Intra-articular injections of 750 kD hyaluronan in the treatment of osteoarthritis: a randomised single centre double-blind placebo-controlled trial of 91 patients demonstrating lack of efficacy. *Ann. Rheum. Dis.* **1994**, *53* (8), 529–534.
98. Lo, G.H.; LaValley, M.; McAlindon, T.; Felson, D.T. Intra-articular hyaluronic acid in treatment of knee osteoarthritis: a meta-analysis. *Jama* **2003**, *290* (23), 3115–3121.
99. Jubb, R.W.; Piva, S.; Beinat, L.; Dacre, J.; Gishen, P. A one-year, randomised, placebo (saline) controlled clinical trial of 500-730 kDa sodium hyaluronate (Hyalgan) on the radiological change in osteoarthritis of the knee. *Int. J. Clin. Pract.* **2003**, *57* (6), 467–474.



Gradient Surfaces: Preparation, Characterization, and Interactions with Biological Species

Hai Bang Lee

Nanobiomaterials Laboratory, Korea Research Institute of Chemical Technology, Yuseong, Daejeon, Korea

Moon Suk Kim

Korea Research Institute of Chemical Technology, Yuseong, Daejeon, Korea

Gilson Khang

Chonbuk National University, Jeonju, Korea

Jin Ho Lee

Hannam University, Daejeon, Korea

INTRODUCTION

Over the last three decades, the importance of biomaterials has been recognized in the biomedical market and thus several specialized biomaterials were provided as commercial grades to the medical market in several countries. The biomaterials have ranged from synthesized materials of metals, ceramics, and polymers to those including biological materials. The newest definition of biomaterials is “either naturally occurring materials in living organisms or materials designed to repair humans.”^[1]

Although the biomaterials may have excellent bulk properties such as strength and elasticity, they may show relatively poor surface properties, e.g., poor wear resistance and limited biocompatibility. It may be necessary to make a compromise between the bulk and surface properties to fulfill biocompatibility. The properties of the biomaterial surface, such as wettability (hydrophilicity/hydrophobicity), chemistry, charge, and roughness, have been shown to be critical for biocompatibility.

Some groups have studied the interactions of different types of biological species with various solid substrates (mainly polymer) with different wettabilities to correlate the relationship between surface wettability and compatibility of biological species. The interactions with biological species mainly concern phenomena like protein adsorption,^[2] cellular adhesion and spreading,^[3] and inflammatory reactions^[4] that may take place under biological circumstances.

The modification of biomaterials by various surface treatment methods has been focused for decades as an interesting topic in the field of surface engineering. Large numbers of research groups have extensively studied the effect of the surface wettability on the interactions of biological species with solid substrates. The popularly used solid substrates with different

wettabilities are listed in Table 1. The water contact angle is an indicator of wettability of material surfaces, which is usually measured using a goniometer (static method) or the Wilhelmy plate apparatus (dynamic method).^[13]

Among various solid substrates, the importance of polymer surface wettability for cell adhesion was reported by Weiss in 1960,^[14] and was later confirmed in many other studies. Kaelble and Moacanin studied bioadhesion in relation to surface energy.^[15] Baier et al.^[16] and Van der Valk et al.^[17] showed some relationships between surface free energy and cell adhesion: Materials with low surface energies showed low cell attachment. Grinnell demonstrated that surface wettability influences the adhesion and proliferation of different types of mammalian cells.^[18] Cell adhesion occurred preferentially on water-wettable substrates.

Schakenraad et al.^[5] and Absolom, Hawthorn, and Chang^[19] reported that human fibroblasts or endothelial cells are spread poorly on hydrophobic substrates and more extensively on more hydrophilic substrates. Tamada and Ikada found that the maximal adhesion of fibroblasts occurs on surfaces having moderate water wettability and the adhesion decreases on surfaces with both higher and lower water wettabilities.^[6,20] van Wachem et al. also found that human endothelial cells adhere optimally on moderately wettable polymers.^[7] Jansen, van der Waerden, and de Groot observed that human fibroblasts are sensitive to surface wettability, whereas epithelial cells may not be.^[21] van Kooten et al. determined the strength of adhesion and detachment of fibroblasts from substrates with different wettability.^[22] Ease of detachment seemed to decrease with increasing wettability.

Much attention has also been paid to the interactions of blood proteins with various polymers having different wettabilities. Uyen et al. studied the

Table 1 Water contact angles of various solid surfaces

Solid surface	Water contact angle ^a	References
Glass	10–32	[5–7]
Cellulose	13–18	[6,7]
Germanium	17–24	[8]
Tissue culture polystyrene (TCPS)	32–60	[5,7,9]
Polyhydroxyethyl methacrylate (PHEMA)	33–38	[8]
Polyurethane (PU)	35–39	[7,8]
Polyvinyl alcohol (PVA)	35–47	[6]
Tissue culture polyethylene terephthalate (TCPET)	42–46	[7]
Ceramics	44	[10]
Cellulose acetate	55	[11]
Polymethyl methacrylate (PMMA)	57–71	[5,7,12]
Nylon 6	59–63	[6]
Polyethylene terephthalate (PET)	59–68	[6,7,12]
Poly-L-lactic acid (PLLA)	68–74	[7]
Polystyrene (PS)	73–94	[6–8,12]
Polyvinyl chloride (PVC)	75	[9]
Polyvinylidene fluoride (PVDF)	75	[5]
Polycarbonate (PC)	83–95	[5]
Polysulfone (PSf)	86	[5,7]
Polyethylene (PE)	87–100	[5,6,12]
Polypropylene (PP)	89–96	[6,12]
Polydimethyl siloxane (PDMS)	100	[5]
Paraffin	101	[5]
Polytetrafluoro ethylene (PTFE)	105–119	[5,6]

^aMeasured by contact angle goniometer.

adsorption of albumin to substrates with a broad range of wettabilities.^[23] They showed that the amount of adsorbed protein is the highest on the most hydrophobic material and decreases with increasing wettability of the substrates. Lu and Park studied the extent of conformational changes of fibrinogen adsorbed on substrates with different wettabilities.^[8] They showed that the amount of adsorbed protein increases as the surface becomes more hydrophobic and the adsorbed protein undergoes a larger degree of conformational changes as the surface hydrophobicity increases.

One main problem derived from the studies using different kinds of substrates is that the surfaces are heterogeneous both chemically and physically (different surface chemistry, charge, roughness, rigidity, crystallinity, etc.), which may result in considerable variation. Another methodological problem is that such studies are often tedious because a large number of samples must be prepared to cover the range of the desired wettability. They may also involve the strong possibility of methodological error because the experiment for each sample is carried out separately.

Many studies have recently been focused on the preparation of surfaces with properties that change gradually along the material length. Such a “gradient surface” is of particular interest for basic studies of the interactions between biological species and surfaces because the effect of a selected property like wettability can be examined in a single experiment on one surface.^[24] Many papers have been published about the preparation of wettability gradient surfaces and their uses in studying the interaction phenomena of biological species like proteins, cells, or enzymes. Moreover, several techniques for preparing chemical composition gradients on various substrates have been developed, and such gradients have been used for further experiments and applications such as biomolecular interactions,^[25] cell-motility studies,^[26] or diagnostics.^[27]

In the forthcoming sections, brief summaries on the preparation methods of wettability gradient surfaces, introduction of biofunctional groups on gradient surfaces, characterizations, and applications of gradient surface in the biomedical fields will be given.

PREPARATION OF WETTABILITY GRADIENT SURFACES

The preparation of wettability gradient surfaces^[28–42] has been described by several groups, as reviewed in detail below.

Gas or Solvent Diffusion

The gas diffusion technique is based on exposing silicon wafers to the vapor of a volatile silane with the vapor concentration being diffusion controlled. A water contact angle hysteresis between 6° and 8°^[43] is typical for gas diffusion gradients, which is low compared to the hysteresis encountered on solvent diffusion gradients. Unfortunately, the wettability steepness cannot be easily controlled.^[43,44]

Meanwhile, the solvent diffusion technique usually uses silane dissolved in a solvent such as xylene, and the silane solution diffuses to the xylene region.^[30]



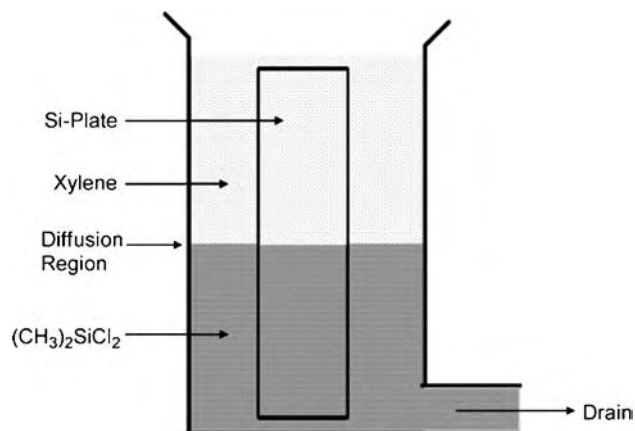


Fig. 1 Schematic diagram showing the diffusion method used to prepare surface gradients of methyl groups. (From Ref. [30].)

Silane is simultaneously bound to a silicon wafer or glass surface and subsequently is coupled to the surface (Fig. 1). The experiment is stopped by removing the solution through the drain. Thus a gradient of coupled silane is formed along the length of the silicon wafer. The gradient of these functional groups can be used as a template for a second reaction step to introduce other biomolecules or even biomaterial grafts.^[45] Liedberg et al. also prepared molecular gradients with various alkanethiols on gold particles using the solvent diffusion technique to design model surfaces for fundamental studies of protein adsorption and immobilization phenomena.^[46,47]

The main advantage of the diffusion technique is that this technique is easily utilized and gives exactly parallel and continuously varying chemical composition profiles on a single substrate surface along the sample length, while the disadvantage is that this technique can only be applied to hydrophilic inorganic substrates like silicon, silica, quartz, or glass.^[36]

Radio-Frequency Plasma Discharge

Various gases such as oxygen and nitrogen are used for the plasma treatments of various substrates. An increased surface energy of substrates results from the plasma treatment. A method for preparing wettability gradients on various polymer surfaces via radio-frequency plasma discharge (RFPD) treatment was developed by Pitt,^[48] Golander and Pitt,^[49] and our group.^[12,50–52]

The wettability gradients were produced by moving a cover over the polymer surface during exposure to RFPD (Fig. 2). As the exposure time to the plasma on the polymer surface gradually increases along its length, a gradient surface is produced. The technique can be applied to any polymer such as PE, PP, PET,

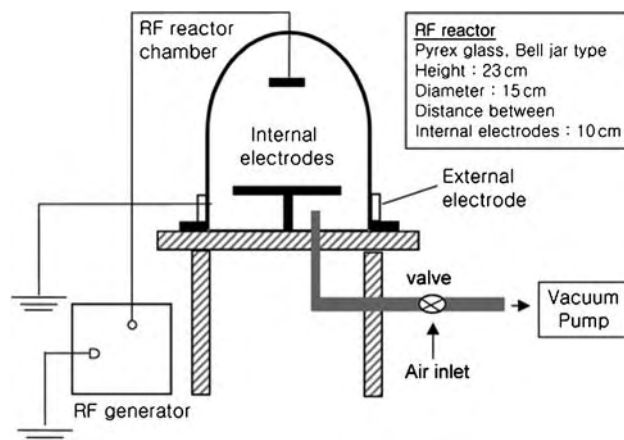


Fig. 2 Schematic diagram showing RFPD apparatus. (From Ref. [50].)

PMMA, and PS. It also allows a wide choice of different gases including water vapor^[50] or oxygen.^[51,52] By this treatment, the polymer surfaces are oxidized gradually along the sample length with increasing plasma exposure time, and thus the wettability gradients are created. The polymer surfaces prepared by the plasma treatment showed a wettability gradient with the water contact angle ranging from 80–90° to 35–50°, depending on the treatment conditions (Fig. 3).^[50–52]

Fig. 4 shows the possible mechanism by which the oxidation that occurs on a polymer surface like PE or PS is affected by the plasma treatment. The plasma discharge treatment on the polymer surface may produce carbon radicals from the hydrocarbon backbone, followed by the formation of unstable hydroperoxides through rapid binding with oxygen in the reactor chamber. The unstable hydroperoxides are easily

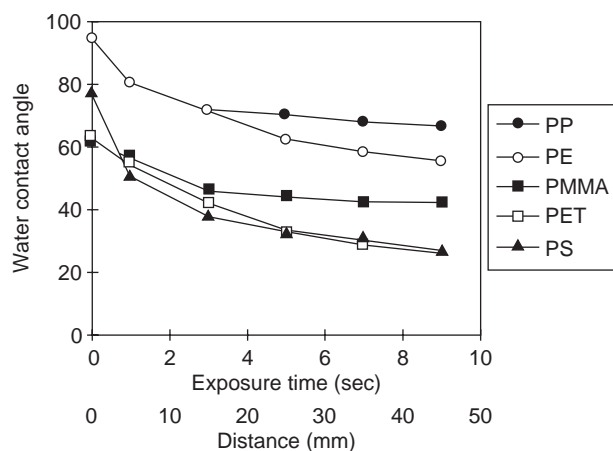


Fig. 3 Water contact angles of gradient surfaces as a function of water vapor plasma treatment (plasma exposure time, up to 10 sec). (From Ref. [50].)

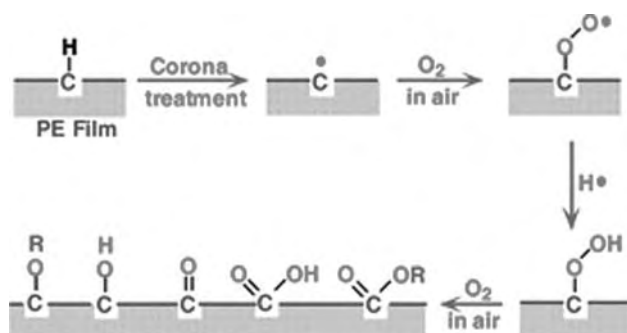


Fig. 4 Plausible mechanism of the formation of oxygen-based functionalities on a polymer surface by oxygen plasma discharge treatment. (From Ref. [50].)

decomposed to produce various oxygen-based polar functionalities (hydroxyl group, ester, ketone, aldehyde, carboxylic acid, carboxylic ester, etc.) by reaction with additional oxygen in the reactor chamber or air.^[53–58] It seems that the oxygen-based functional groups produced on the polymer surface increase with increase in the plasma exposure time, resulting in the formation of wettability gradient on the surface. The main disadvantage of the plasma treatment method is that it is usually carried out in vacuum and may also yield anomalous surface chemistries that are, in addition, not always stable over time.^[59,60]

Corona Discharge

A new method for preparing wettability gradients on polymer surfaces using the corona discharge treatment was developed by our group (Fig. 5).^[60,61] A radio frequency (RF) corona discharge apparatus was used for the preparation of wettability gradient surfaces.^[60] The corona discharge apparatus consists of two key parts that produce the wettability gradient on polymer surfaces: a knife-type electrode onto the sample with

increasing power (from 10 to 35 W at 100 kHz) and a movable sample bed (usually 1.0 cm/sec). By this treatment, the sample surface is continuously exposed to the corona with increasing power. Thus, the surface is oxidized gradually along the sample length, resulting in the wettability gradient with various functional groups (Fig. 4). Corona discharge has also been used to prepare gradients by biomaterial grafts.^[62–64] This method is easier to apply as it can be done under atmospheric conditions.

INTRODUCTION OF BIOFUNCTIONAL GROUPS ON GRADIENT SURFACES

Introduction of various biofunctional groups on biomaterial surfaces may be important to improve biocompatibility or induce an appropriate response in biological systems. The introduction can be performed by employing physical or chemical immobilization of bioactive molecules on various gradient substrates, as reported by several groups.

Hydroxy Groups on Gradient Surface

Various polymer surfaces, including PE, PP, PS, PET, and PMMA, were modified by water vapor plasma discharge treatment. It was observed that by the water vapor plasma treatment, the wettability of the polymer surfaces increases largely and almost all functional groups produced on the surfaces are hydroxyl groups. The changes in chemical structure on the polymer surfaces by water vapor plasma treatment were investigated using the measurement of water contact angles and electron spectroscopy for chemical analysis (ESCA). Among the polymers, PE showed the highest hydroxyl group density on the surface (Table 2).

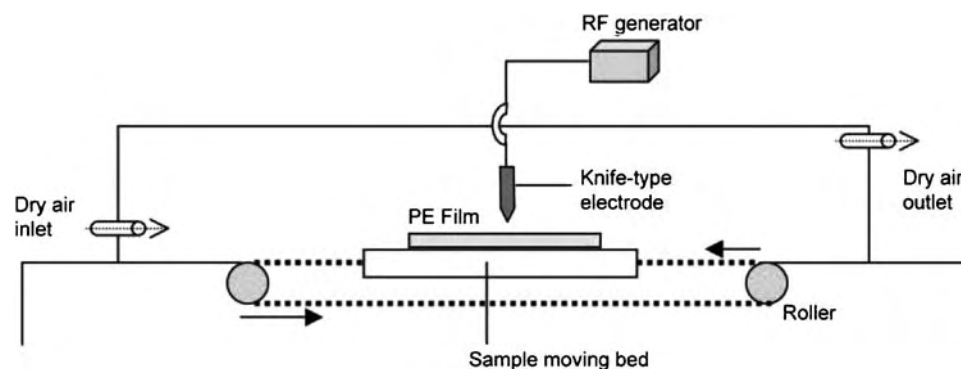


Fig. 5 Schematic diagram showing the corona discharge apparatus for the preparation of a wettability gradient on polymer surfaces by the corona discharge treatment. (From Ref. [60].)

Table 2 ESCA analysis of H₂O plasma-treated polymer surfaces

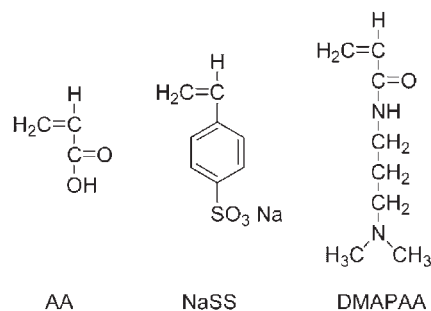
Substrate	Oxygen content (%)	C-OH/C-C (%)
PE	19.2	13.0
PP	10.3	5.4
PS	17.1	7.8
PET	37.5	7.1
PMMA	34.8	6.6

(From Ref. [50].)

Water vapor plasma treatment seems a simple and effective method to produce hydroxyl groups on the polymer surfaces.^[12,50]

Ionic Functional Groups on Gradient Surface

The gradient surfaces with ionic functional groups were prepared by the graft copolymerization of ionic vinyl monomers such as acrylic acid (AA), sodium *p*-styrene sulfonate (NaSS), and *N,N*-dimethyl amino-propyl acrylamide (DMAPAA) after the corona discharge treatment (Fig. 6).^[65-67] The peroxides produced on the corona-treated PE surface act as

**Fig. 6** Chemical structures of chargeable monomers. (From Ref. [65].)

an initiator for graft copolymerization. Ionic vinyl monomers were graft-copolymerized through the peroxides on the corona-treated PE surface. Acrylic acid was graft-copolymerized on the PE surface (PE-COOH) in aqueous solution at 70°C without any additives. Hydroxyl group-grafted (PE-CH₂OH) and amide group-grafted (PE-CONH₂) surfaces were prepared by substituting the -COOH groups grafted on the PE surface to -CH₂OH and -CHNH₂ groups, respectively, via -COCl groups (Fig. 7). The amine groups grafted on the surface (PE-CH₂NH₂) were induced from amide group-grafted surface. PE-COOH or PE-NaSS is negatively chargeable and PE-CONH₂ or PE-DMAPAA is positively chargeable in

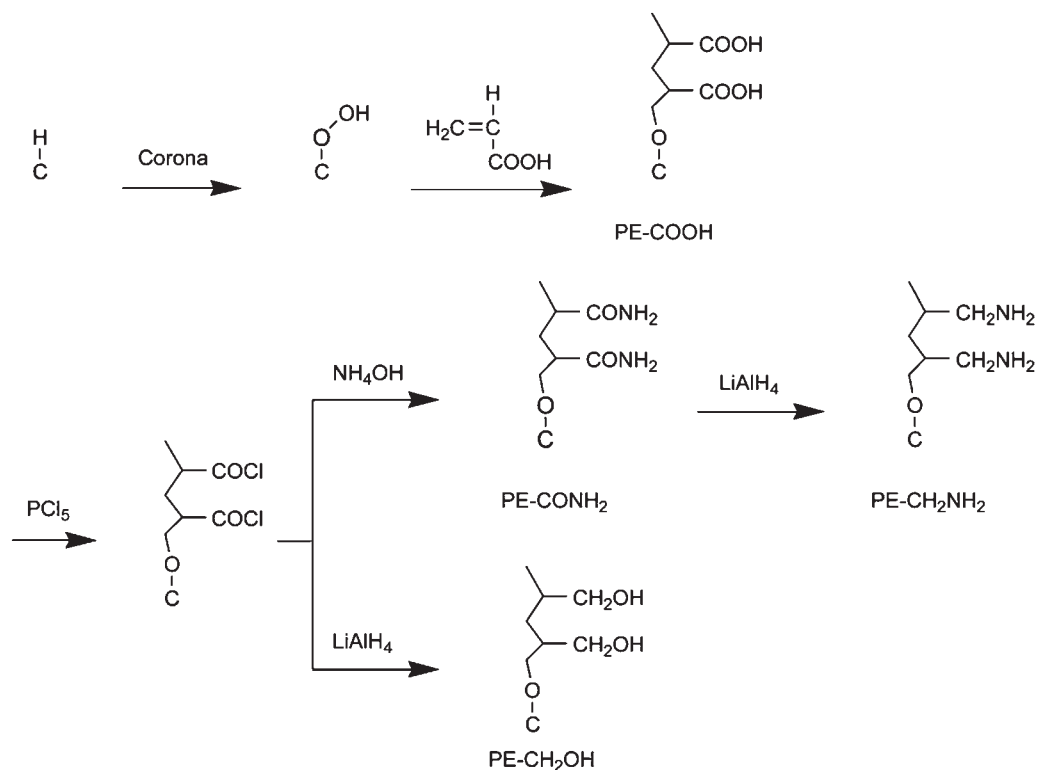
**Fig. 7** Schematic diagram showing the formation of functional group-grafted PE surfaces by corona discharge treatment followed by graft copolymerization and subsequent reactions. (From Ref. [66].)

Table 3 Comparison of water contact angles of functional group-grafted PE surfaces

Surface	Contact angle (°)
PE-COOH	22.8
PE-CH ₂ OH	34.8
PE-CONH ₂	23.2
PE-CH ₂ NH ₂	28.4
PE-NaSS	35.6
PE-DMAPAA	36.8

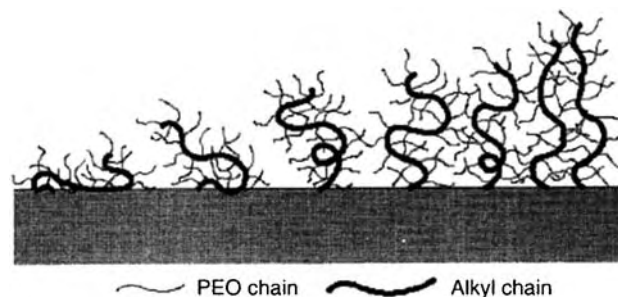
(From Ref. [67].)

phosphate-buffered saline or plasma solution at pH 7.3–7.4. The gradient surfaces with ionic groups were characterized by the measurement of water contact angles, ESCA, and Fourier-transform infrared spectroscopy in the attenuated total reflectance mode (FTIR-ATR). All these measurements indicated that the ionic groups were grafted on the PE surfaces with gradually increasing density. However, the wettability of the different ionic group-grafted surfaces was not much changed by the kind of functional groups grafted (Table 3).

Polyethylenimines (PEIs) that have long spacers can be used for the immobilization of the surface, and consequently allow flexible and multifunctional sites on the surface in the biological environment. In addition, PEI in water is cationic owing to protonation. Thus chemical modification of the surface using PEI was performed.^[68,69] The carboxylic groups on the corona-treated gradient PE surface were activated by the *N*-ethyl-*N'*-(3-dimethylaminopropyl) carbodiimide, followed by the reaction with PEI. The peak assignable to C=O functional groups on the corona-treated PE surface disappeared at 1735/cm while a new peak was observed at around 1586/cm after reaction with PEI. In ESCA, the gradual increase in height of the carbon 1s peak owing to the attribution of C-N species within the PEI was observed along PE length. In addition, nitrogen 1s peak was observed in the survey scan spectrum and the intensity also gradually increased along PE length, indicating nitrogen content increase on the PE surface. Thus PEI was gradually modified on the gradient PE surface.

Polyethylene Oxide on Gradient Surface

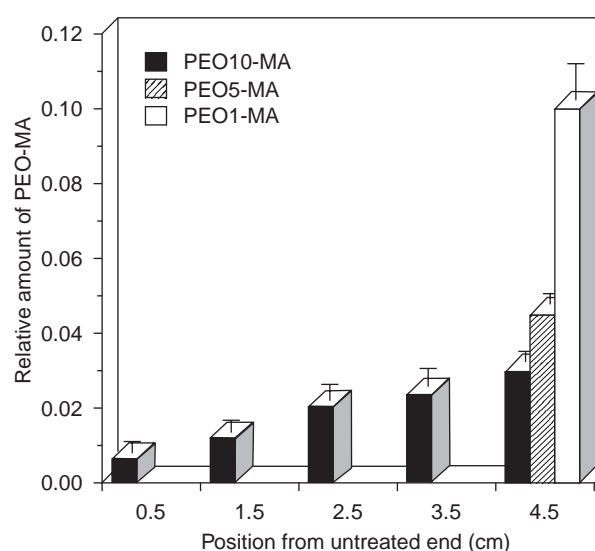
Particularly effective polymers for “protein-resistant” (or “blood-compatible”) surfaces appear to be polyethylene oxide (PEO) owing to its minimum interfacial free energy with water, hydrophilicity, high surface mobility and steric stabilization effects, and unique solution properties and molecular conformation in water.^[70,71] Comb-like PEO gradient surfaces were

**Fig. 8** Schematic diagram showing a comb-like PEO gradient produced on polymer surfaces. (From Ref. [64].)

prepared by graft copolymerization of poly(ethylene glycol) monomethacrylate macromers (PEO-MA) on PE surface (Fig. 8).^[64,72] The macromers with different PEO repeat units, 1, 5, and 10, were used. The comb-like PEO chain density was changed gradually along the sample lengths by the corona discharge treatment with gradually increasing power (Fig. 9). The grafted amount of PEO-MA increased as the PEO chain length in the PEO-MA decreased owing to the steric interference of longer PEO side chains in PEO-MA macromers. The prepared comb-like PEO gradient surfaces were characterized by the measurement of water contact angles and FTIR-ATR (Fig. 10), and ESCA (Fig. 11). All these measurements indicated that the PEO chains are grafted on the PE surface with gradually increasing density of PEO.

Phospholipid Polymer on Gradient Surface

Ishihara et al. reported the synthesis of 2-methacryloyloxyethyl phosphoryl choline (MPC) polymers with

**Fig. 9** Relative amount of PEO-MA grafted on PE surfaces. (From Ref. [72].)

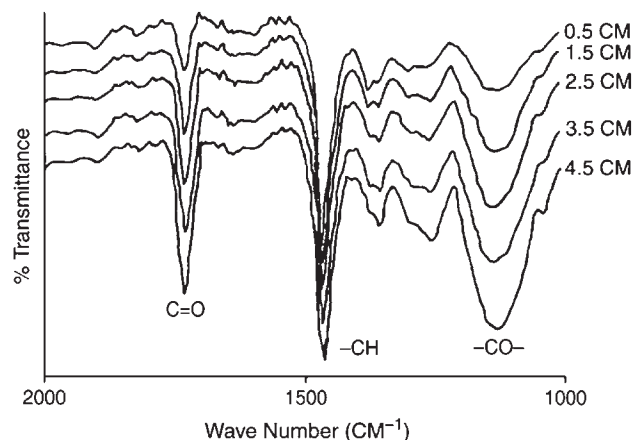


Fig. 10 FTIR-ATR spectra of PEG10-MA grafted PE surface along the sample length. Numbers labeled on the spectra (0.5–4.5) represent the positions from the untreated end of the gradient surface. (From Ref. [64].)

excellent blood compatibility and applied them to biomedical fields.^[73,74] They also prepared the gradient PE surface, grafted gradually with the poly [ω -methacryloyloxyalkyl phosphorylcholine (MAPC)] by corona discharge treatment to examine the surface characteristics, i.e., density, mobility, and orientation of the poly(MAPC) (Fig. 12). The P/C ratio of the PE surface grafted with poly(MAPC) gradually increased along the sample length (Fig. 13). The orientation and mobility of the poly(MAPC) grafted on the PE surface was evaluated using 1,6-diphenyl-1,3,5-hexatriene. The orientation of poly[6-methacryloyloxyhexyl phosphorylcholine (MHPC)], which has six methylene chains between the phospholipid polar group and the backbone, was higher than that of other poly (MAPC)s. The mobility of the poly(MAPC) decreased with an increase in the methylene chain length in the MAPC unit.

CHARACTERIZATION OF GRADIENT SURFACES

Water Contact Angle

Chemical gradient surfaces are nearly always characterized by their wettability using contact angle methods, especially because the measurement of contact angles is one of the most sensitive techniques to study a surface with a depth of information within a few atomic layers.^[75,76] Contact angles can be measured by a variety of methods,^[77,78] however, not all are suitable to measure wettabilities as a function of position. Sessile drop technique can be used in case of gradients with a long extension (Fig. 14)^[12,59,80,81] and involves placing microliter droplets over the length of the

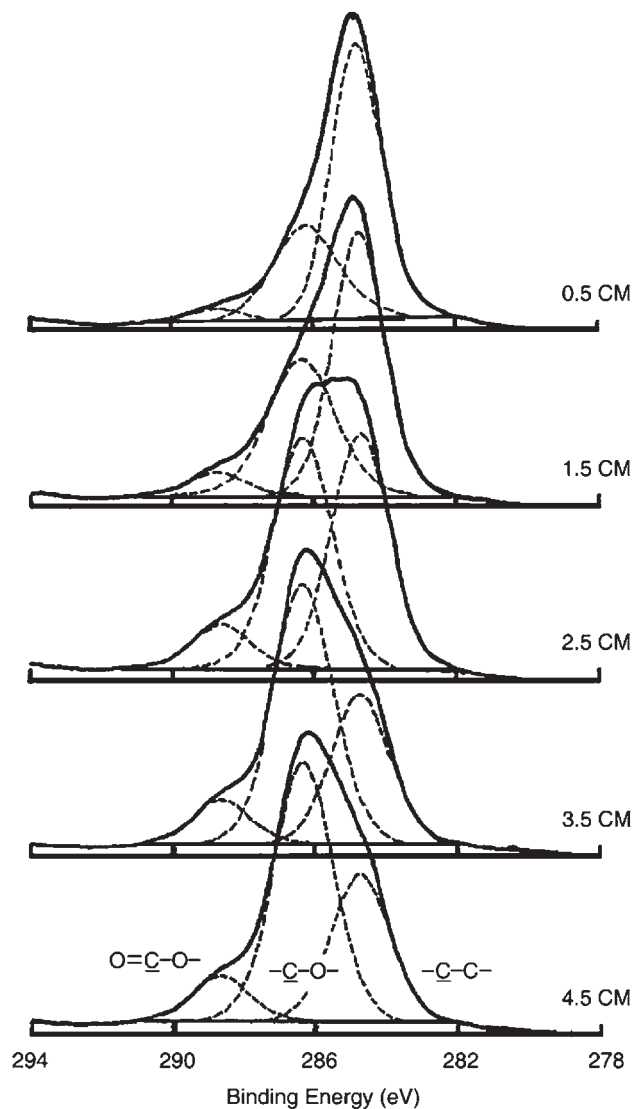


Fig. 11 ESCA carbon 1s core-level spectra of PEG10-MA grafted PE surface along the sample length. (From Ref. [64].)

gradient surface and the measurement of the contact angle of the droplet with the surface by simple goniometry. However, variations in position-bound wettability may occur over the dimensions of the droplet, which makes the sessile drop method less suitable for application on steep gradients.

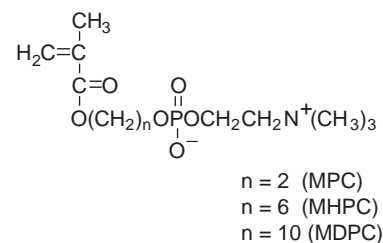


Fig. 12 Chemical structures of MAPC. (From Ref. [74].)

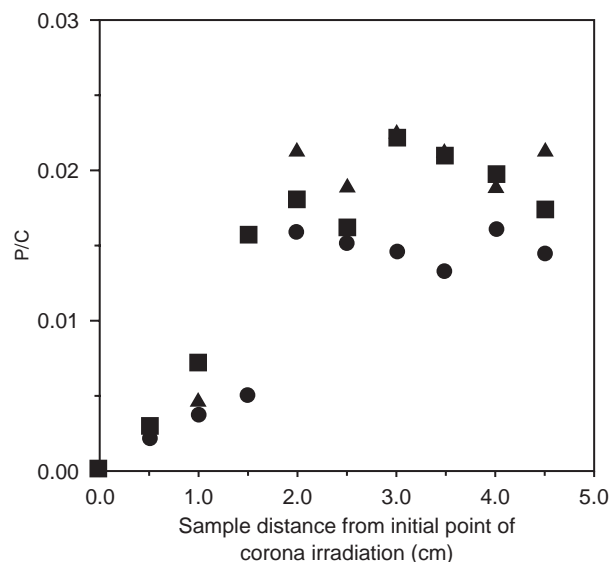


Fig. 13 Changes in P/C ratio of the gradient PE surface grafted with poly(MAPC) along the sample distance: (●) poly(MPC); (■) poly(MHPC); (▲) poly(MDPC). (From Ref. [74].)

Pitt^[48] and Golander and Pitt^[49] also prepared the wettability gradients on various polymer surfaces (PE, PS, PDMS, and PTFE) via RFPD treatment by exposing the polymer sheets continuously to oxygen, ammonia, or sulfur dioxide plasma. They demonstrated that the water contact angles on the plasma-treated polymer surface decrease monotonically as the plasma exposure time increases. They showed that a wide range in wettability is produced by manipulating

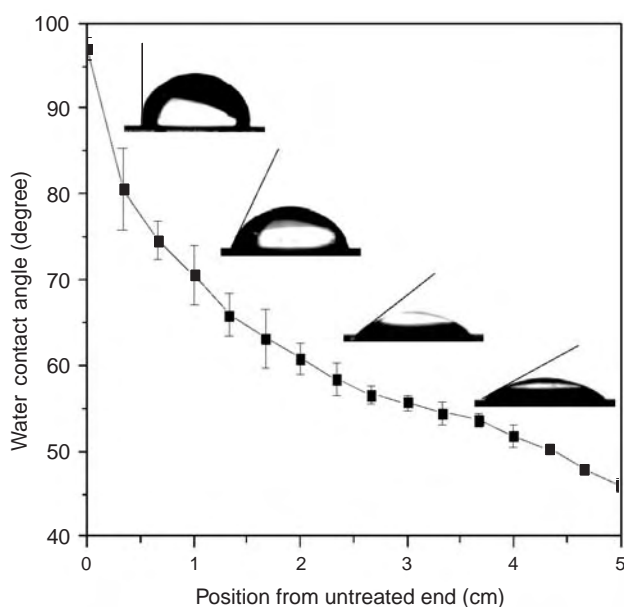


Fig. 14 Contact angle of gradient PE surface measured by sessile drop method. (From Ref. [79].)

the RF power, the gas species, and the duration of plasma exposure.

Ellipsometry

Ellipsometry measures the change in polarization of light upon reflection against a surface, which can subsequently be employed to calculate, for example, the thickness of adsorbed films on a surface.^[82] Thus, ellipsometry was employed as the main technique for monitoring the gradient thickness as a function of distance of the samples.

Recently, the spatial resolution of ellipsometry could be observed even at nanometer scales required to determine the variation in density of functional compounds over the length of a gradient surface. Liedberg et al., for instance, by using ellipsometry and infrared (IR), measured steadily increasing thicknesses of adsorbed alkanethiol with different chain lengths of ethylene oxide on gradient surfaces (Fig. 15).^[47] The thickness of the nanometer scale obtained from the gradient is in good agreement with the thickness of the corresponding single component. In addition, ellipsometry is frequently applied to determine the mean position-bound thickness of proteins adsorbed to the gradient surfaces.^[30,32–36]

ESCA

The ESCA, also called x-ray photoelectron spectroscopy (XPS), is useful for the characterization of surface properties although it has a relatively low spatial resolution of about several square millimeters up to 1–2 μm^2 for small spot, scanning XPS instruments and is not always adequate to detect adsorbed monomolecular layers. The depth of their information is

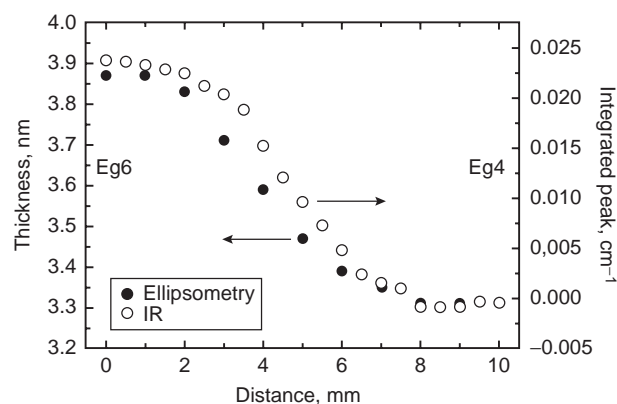


Fig. 15 Ellipsometry thickness of oligo(ethylene glycol)-terminated alkanethiols (Eg4 and Eg6 represent the ethylene oxide unit) on gradient surface. (From Ref. [47].)

within the nanometer range, which is better than that of IR spectroscopy but worse than that of ellipsometry.

As many gradients are prepared using polymer materials, carbon contamination of gradient surfaces, which is almost unavoidable in ESCA, poses a severe complicating factor in the ESCA characterization of gradient surfaces. Although ESCA does not provide absolute quantification, it is very sensitive to the small changes in the surface composition.^[83]

The changes in chemical structure of polymer surfaces by RFPD treatment were investigated by ESCA.^[50,79] Fig. 16 shows ESCA carbon 1s core-level scan spectra of the corona-treated PE surface. The hydrophobic side of the gradient (position, 0.5 cm) showed an alkyl carbon ($-C-C-$) peak at a binding energy, ~ 285 eV. The positions with increasing corona power showed new peaks at higher binding energies, indicating the formation of carbon-oxygen functionalities. The peaks of the high binding energy region correspond to $-C-O-$ at ~ 286.6 eV (e.g., hydroxyl or ether groups), $-C=O$ at ~ 287.9 eV (e.g., ketone or aldehyde groups), and $O=C-O-$ at ~ 289.1 eV (e.g., carboxylic acid or ester groups). The oxygen-based functional groups increased with increasing corona power. In addition, the survey scan spectrum provides information about the elemental composition of surfaces for all elements present (except hydrogen and

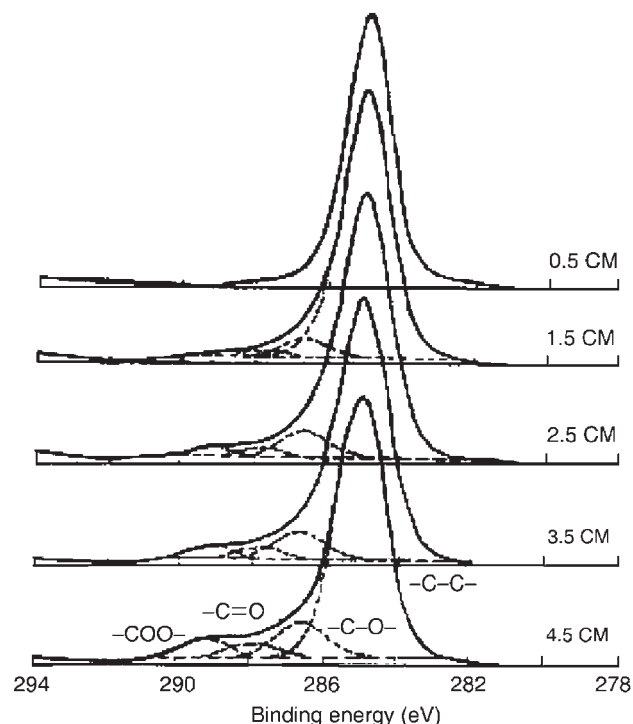


Fig. 16 ESCA carbon 1s core-level spectra of a corona-treated PE surface along the sample length. Numbers labeled on the spectra (0.5–4.5) represent the positions from the untreated end of the gradient surface. (From Ref. [79].)

Table 4 ESCA analysis of a corona-treated PE surface along the sample length

Position from untreated end (cm)	Atomic (%)				O/C ratio
	$-C-C-$ (~ 285.0) ^a	$-C-O-$ (~ 286.6) ^a	$-C=O-$ (~ 288.0) ^a	$-O=C-O-$ (~ 289.1) ^a	
0.5	100	0	0	0	0.00
1.5	92	5	2	1	0.09
2.5	88	8	2	2	0.13
3.5	86	8	3	3	0.16
4.5	80	11	4	5	0.25

^aBinding energy, eV.
(From Ref. [79].)

helium). The elements are identified from their binding energies, which are unique for each element. In the case of corona treatment, hydroxyl or ether groups ($-C-O-$) were produced more largely on the surface than other functional groups, as seen in Table 4.

Infrared Spectroscopy

IR spectroscopy in the attenuated total reflection (ATR) mode is sometimes used to characterize gradient surfaces although the surface sensitivity of IR spectroscopy is within the micrometer range, which is often too deep to detect the adsorbed monomolecular layers on gradient surfaces. For all that, microscopic IR spectroscopy in the ATR mode is technically feasible, like for ellipsometry.

We^[50,60–67,70–74,79] prepared oxygen-rich gradients on PE surface by a corona-discharge treatment and subsequently determined the position-bound infrared spectra by FTIR-ATR. Regions of the gradient that had been exposed longer to the corona discharge had stronger C=O absorption bands, indicating the gradually increasing peak intensity (Fig. 17). Thus, IR spectroscopy in general is a suitable technique to characterize gradient surfaces.

Atomic Force Microscopy

Atomic force microscopy (AFM) has been established as a technique for high-resolution surface imaging. The ability to map the forces between the scanning tip and a surface within nanometer resolution makes it a unique tool to analyze properties of molecules on a surface, although the dangers of creating artifacts by the scanning tip are huge.^[84] The AFM studies to characterize gradient surfaces have been done by our group.^[74] As shown in Fig. 18, AFM measurements showed the corona-treated PE surface increased in

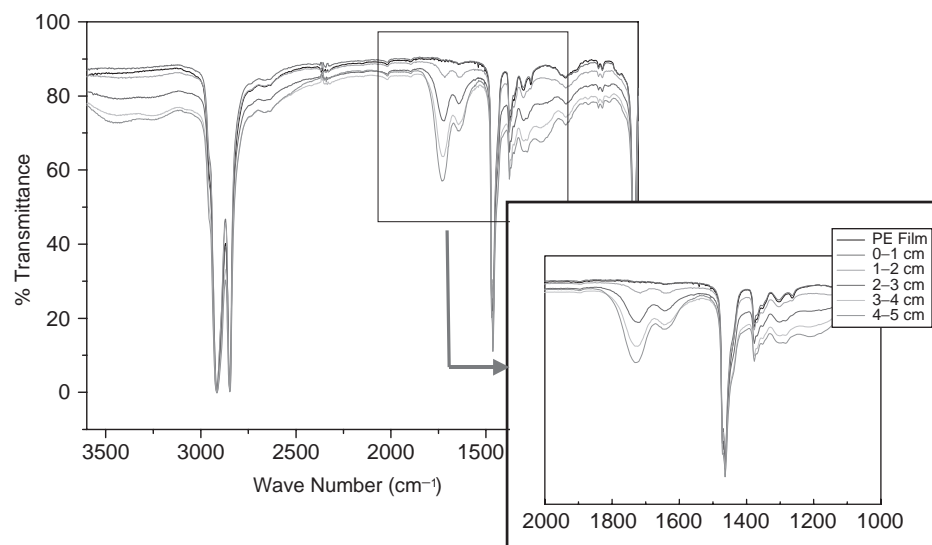


Fig. 17 FTIR-ATR spectra along the PE length after corona treatment. (From Ref. [79].)

roughness, from 10 nm at position 0.5 cm to 100 nm at position 4.5 cm, after corona treatment, indicating that the PE surface changes in roughness after corona treatment, as the corona power increased along the PE length. Thus, AFM can be employed as the unique technique for gradient surface imaging.

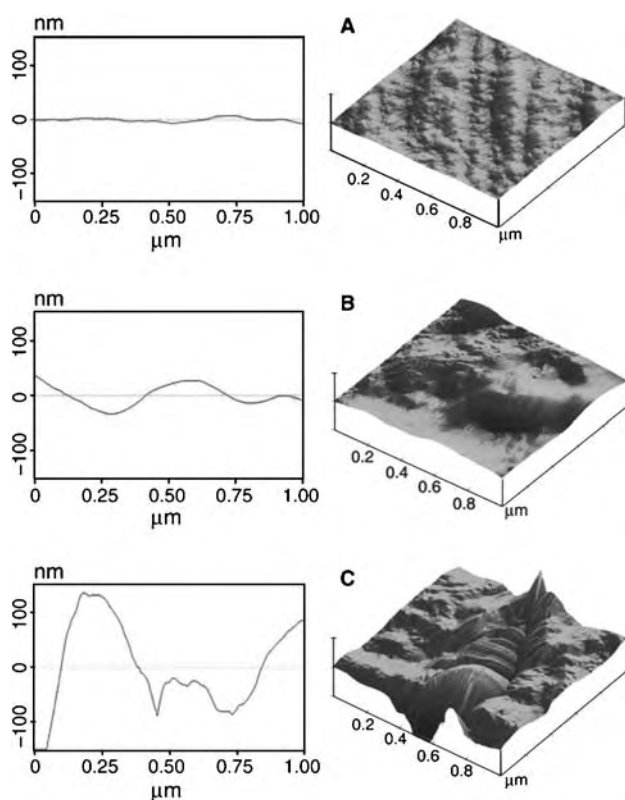


Fig. 18 AFM images at the position (A) 0.5 cm, (B) 2.5 cm, and (C) 4.5 cm of gradient PE surface. (From Ref. [69].)

Other Methods

Characterization of the gradient surfaces was carried out with total internal reflection fluorescence (TIRF) spectroscopy in the presence of 1-anilinonaphthalene-8-sulfonate (ANS) as a surface probe.^[38,85] Strong enhancement of ANS fluorescence is found at the hydrophobic side of the gradient and no fluorescence enhancement at the hydrophilic side. Thus TIRF can be used to characterize the wettability of the gradient silica or glass surface.

Time-of-flight secondary ion mass spectrometry (TOF-SIMS) is a surface analytical technique that is capable of detecting elemental and molecular species with a spatial resolution of $<1 \mu\text{m}$. Because of these characteristics and the imaging capabilities of TOF-SIMS, the technique is ideal for compositional analysis on surface. Roberson et al. reported the analysis of functional groups present on the gradient surface by using TOF-SIMS.^[86] Fig. 19 shows the positive secondary ion trends observed for the silane monolayer gradient surface. The signals for the oxidized species, SiOH^+ and Si_2O^+ , increase with the increasing gradient nature of the surface.

INTERACTIONS OF GRADIENT SURFACES WITH BIOLOGICAL SPECIES

Recent developments in the design of biocompatible materials have focused on biomimetic materials that are able to recognize biomolecules originating from the extracellular matrix or specific cell responses. Gradient modification of biocompatible materials is one of the simplest methods to make biomimetic materials that possess biomolecule-recognizing properties. Thus

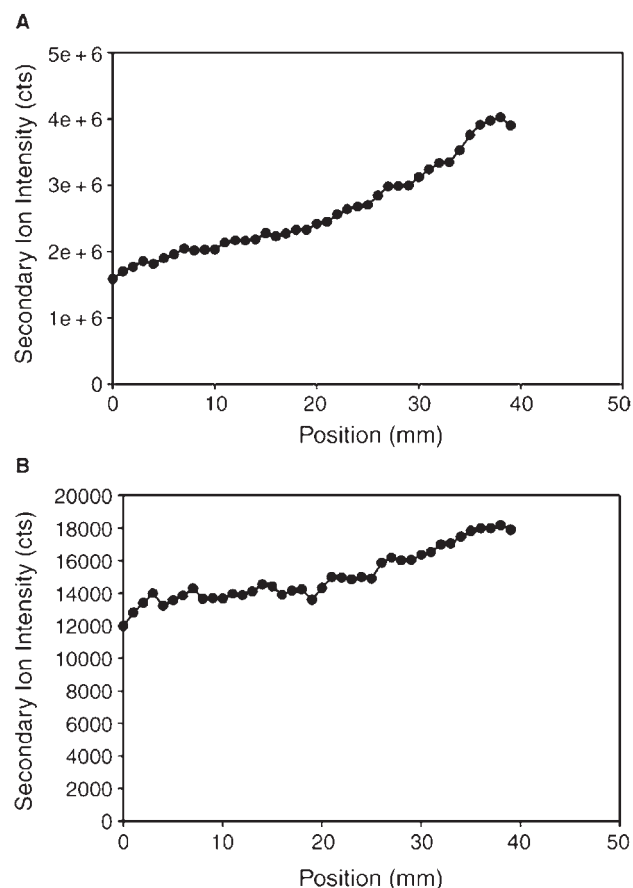


Fig. 19 Positive secondary ion trend of (A) SiOH^+ and (B) Si_2O^+ observed along the length of the silane-treated sample. (From Ref. [86].)

gradient surfaces can offer great potential to examine behaviors and biological interactions of biological species such as proteins and cells. It was observed that the adhesion of biological species to the gradient surfaces decreases gradually with the increasing surface density of functional groups. It seems that surface functional groups and their charge character, as well as wettability, play important roles for the adhesion of biological species. This may be related to the relationship between hydrophilicity and hydrophobicity of the surfaces. Thus, the study for the rapid and accurate recognition of binding behavior of biological species at a specific location on biocompatible materials *in vivo* or *in vitro* presents a challenging target. An overview will be given of protein adsorption and cell adhesion studies on gradient surfaces.

Interaction with Proteins

Elwing et al. studied the surface energy dependence of the adsorption or exchange of single proteins like fibrinogen (molecular weight, 340,000 Da) and

γ -globulin (molecular weight, 160,000 Da) on the solid substrates using the wettability gradient surfaces.^[30,31] The protein adsorption or exchange reactions were quantified along the gradient using ellipsometry or corresponding antibodies. From the study of exchange reactions between fibrinogen and γ -globulin on the fibrinogen-preadsorbed gradient surface, they found that fibrinogen is partially exchanged by γ -globulin only on the hydrophilic side of the gradient, whereas on the hydrophobic side, there appears to be no exchange at all. They also observed that fibrinogen adsorbs in a larger amount on the hydrophobic side of the gradient. The binding energy of fibrinogen was much larger on the hydrophobic side of the gradient; thus exchange reactions were more difficult there. They have also used the wettability gradients to study the surface hydrophilicity-induced conformational changes of adsorbed proteins.^[33] The adsorption of protein and its subsequent interaction with antibodies were measured using enzyme-linked immunosorbent assay and ellipsometry. The interaction force between the hydrophobic surface and the protein molecule gives it an extended structure, covering a relatively large area of the surface. On a hydrophilic surface, however, the forces acting between the surface and the protein molecule may be smaller and, because the conformational change will be smaller, a large repulsive force will act between the protein molecules. Therefore, the packing of adsorbed protein molecules will be less than on a hydrophobic surface and a smaller number of adsorbed proteins will be found on the hydrophilic surface.

We examined binding properties of the human albumin as a model protein adsorbed onto corona-treated PE gradient surfaces.^[61] Albumin also has big hydrophobic patches on its surface. The albumin-adsorbed gradient surface was analyzed by ESCA. The nitrogen signal from the surface was used as an indicator of the protein adsorption because it was observed that little nitrogen is incorporated onto the surface by corona treatment in air. It is mainly derived from peptide bonds of the adsorbed protein. Fig. 20 shows that the nitrogen peaks (binding energy, ~ 399 eV) from ESCA survey scan spectra decreased gradually along the wettability gradient, indicating that the albumin adsorption increased gradually with the increasing hydrophobicity of the surface. This is owing to the increased hydrophobic interactions of the protein molecules with the hydrophobic sections of the gradient surface, as discussed above. The nitrogen content determined by ESCA on the hydrophobic section (water contact angle, 98°) was about 9.3%, which is nearly the value of the monolayer coverage of albumin.^[87] The albumin-adsorbed surfaces were also analyzed by FTIR-ATR. It was observed that the amide II (1500–1600/cm) absorption bands from the

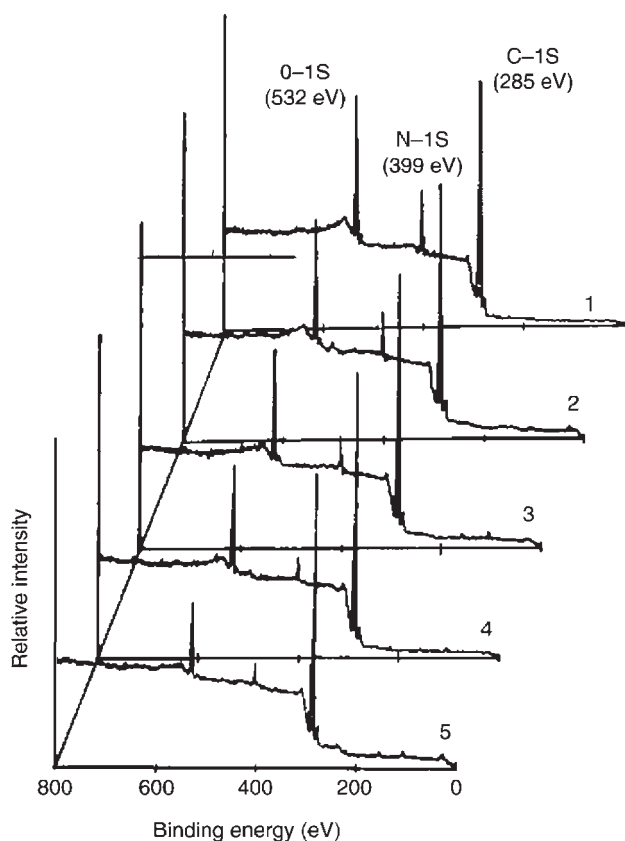


Fig. 20 ESCA survey scan spectra of corona-treated PE surface along the gradient after albumin adsorption (one hour adsorption in 1 mg/ml albumin solution). Numbers labeled on the spectra (1–5) represent the positions from the untreated hydrophobic end of the gradient surface at 0.5, 1.5, 2.5, 3.5, and 4.5 cm, respectively. (From Ref. [61].)

protein adsorbed on the surface also decrease gradually along the wettability gradient. For the desorption study, the albumin-adsorbed PE gradient surfaces were exposed to a nonionic polymeric surfactant, Tetric 1504, which contains starlike 4 PEO (ethylene glycol unit, 26)/polypropylene glycol (propylene glycol unit, 29) blocks in its structure. The exposure of the albumin-adsorbed gradient surface to the surfactant solution resulted in partial displacement of the protein (Fig. 21). This displacement was much greater on the hydrophobic sections of the gradient surface than on the hydrophilic ones. The results of increasing protein adsorption and desorption toward the hydrophobic sections of the corona-treated PE gradient surface agree with those in the works of others, even though the functional groups introduced on the substrates are different.^[30,31,34,39]

The relative amount of proteins adsorbed onto comb-like PEO-grafted PE surfaces was also evaluated.^[64,72] The proteins adsorbed on the control PE surface was much higher than those on PEO-MA-grafted surfaces, indicating the larger amount of

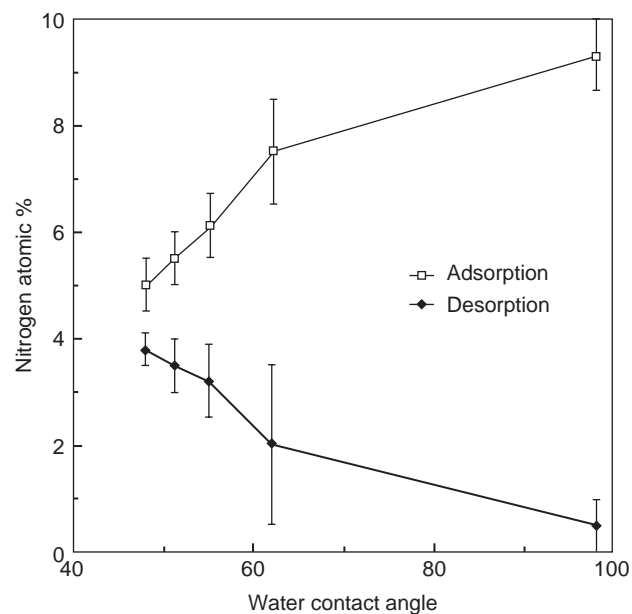


Fig. 21 Adsorption and desorption behavior of albumin on corona-treated PE gradient surface. Nitrogen atomic percent represents the amount of the protein adsorbed (for adsorption) and remaining (for desorption) on the surface (one hour adsorption in 1 mg/ml albumin solution; 30 min desorption in 1 mg/ml Tetric 1504 solution after one hour adsorption in 1 mg/ml albumin solution; $n = 3$). (From Ref. [61].)

protein adsorption on the control surface. This is also owing to the hydrophobic interaction of the protein molecules with the hydrophobic PE surface. Fig. 22 compares the relative adsorbed amount of proteins on the PEO-MA-grafted PE surfaces. It was

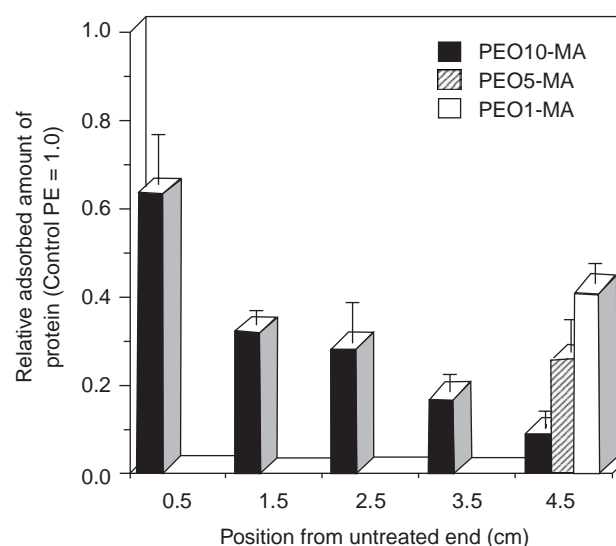


Fig. 22 Relative adsorbed amount of proteins on PEO-MA-grafted PE surfaces ($N = 3$). (From Ref. [72].)

determined as follows: Relative adsorbed amount of proteins = $(N\% \text{ of PEO-MA-grafted surface}) / (N\% \text{ of control surface})$. The surfaces grafted with PEO-MA having longer PEO chains showed less protein adsorption. PEO10-MA-grafted surface at position 4.5 cm was very effective in the prevention of protein adsorption. In the case of PEO10-MA-grafted surface, the amount of adsorbed proteins was reduced more than 90% as compared with the control surface while the PEO1-MA-grafted surface showed about a 50% reduction.

We also reported the gradient PE surface displaying fluorescein intensity of proteins on a gradient surface.^[68,69] Among several binding systems, the binding affinity of the avidin–biotin interaction ($K_D \sim 10^{-15} \text{ M}$) is one of the highest to occur in nature as a non-covalent bond.^[88] Thus, the biotinylated gradient PE surface was prepared by the reaction of PEI and biotin after corona treatment with increasing power along the PE length. The gradient biotinylated PE surface was incubated in a solution of streptavidin–fluorescein isothiocyanate (FITC). The binding of the streptavidin–FITC to the gradient biotinylated PE surfaces was characterized by fluorescence microscopy. The fluorescence microscope image, which corresponds to a binding of FITC with streptavidin on gradient biotinylated PE surfaces, is shown in Fig. 23. The luminosity of fluorescence increased gradually on the surface when corona power for the surface increased along the sample length. Plotting of the length treated by corona vs. the fluorescence intensity of the streptavidin–FITC reveals a gradient as shown in Fig. 24. The longer the length treated by corona, the smaller the water contact angle value and the stronger the fluorescence intensity, which could well explain the gradient PE surface.

Cellular Interactions

The studies on cellular interactions with gradient surfaces were carried out, while distinguishing between numbers of adhering cells, their spread areas, and growth rates. Carter, in 1965,^[89] was probably the first to study the directional migration of mouse fibroblasts on steep palladium gradients, referred to at that time as “adhesion gradients.”

Recently, we reported the property of Chinese hamster ovary (CHO) cells, fibroblasts, and endothelial cells (ECs) cultured on the wettability gradient PE surfaces.^[79] Fig. 25 shows the cell adhesion (after one day) behavior on the wettability gradient surfaces. The CHO cells are recognized as highly surface adhesive and showed large amount of cell attachment compared to fibroblasts and EC. As the surface wettability increased along the sample length, the cells

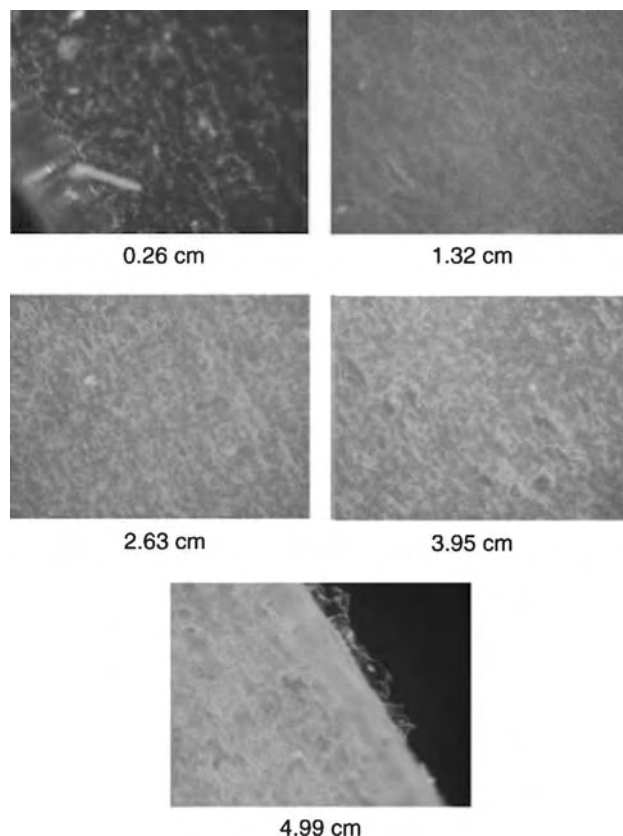


Fig. 23 Change of fluorescence images on gradient PE surface from 0 to 5 cm after binding of streptavidin–FITC. (From Ref. [69].)

adhered on the surface increased and then decreased, regardless of the cell types used. The cells were adhered more on the positions with moderate hydrophilicity of the wettability gradient surface than on more

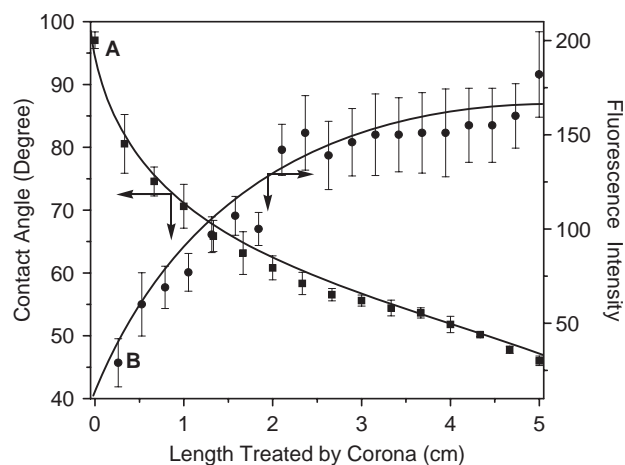


Fig. 24 Plot of (A) water contact angle vs. length treated by corona and (B) fluorescence intensity vs. length treated by corona. (From Ref. [69].)

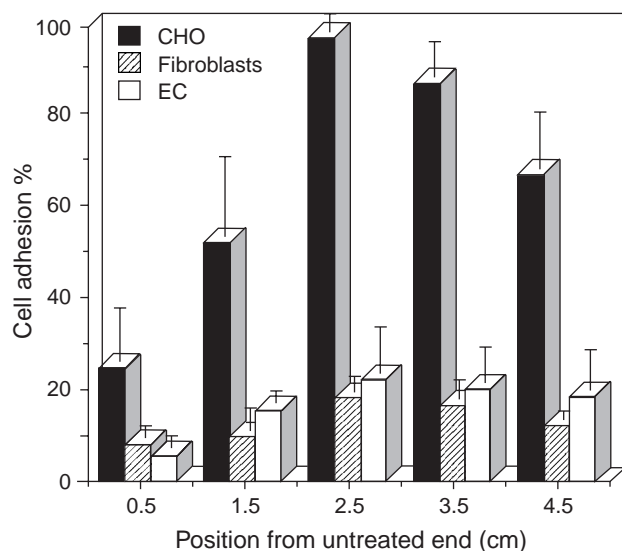


Fig. 25 Different types of cell adhesion on wettability gradient PE surface after a one-day culture (number of seeded cells, $4 \times 10^4/\text{cm}^2$; $n=3$). (From Ref. [79].)

hydrophobic or hydrophilic positions. The maximum adhesion of the cells appeared at around position 2.5 cm (water contact angle, 57° ; Fig. 24). The cell morphology was also changed along the wettability gradient, as observed by SEM. The cells, regardless of the cell types used, were spread and flattened more on the positions with moderate hydrophilicity of the wettability gradient surface than on more hydrophobic or even hydrophilic ones after a one-day culture (Fig. 26). The cells after a two-day culture were almost flattened on all positions of the gradient except the hydrophobic one (position, 0.5 cm). When the number of cells attached on the wettability gradient surfaces were compared after a one- and two-day culture, the cell growth behavior on the surfaces after a two-day culture was almost similar to that of cell adhesion after one day—more attachment of the cells on the positions with moderate hydrophilicity of the wettability gradient surface than on the more hydrophobic or hydrophilic positions. For CHO cells, the large number of attached cells on the surface after a two-day culture was owing to the large amount of cell adhesion after one day compared to other cells. The growth rate of the cells up to two days was in the order of fibroblasts, CHO cells, and ECs. The fact that cells are more adhered, spread, and grown on the moderately hydrophilic surfaces was also observed by other research groups^[6,7,20,22] as they cultured ECs, HeLa S3, or fibroblasts onto various polymer substrates with different surface wettabilities.

The adhesive strength of ECs attached on polymer surfaces with different hydrophilicity was also

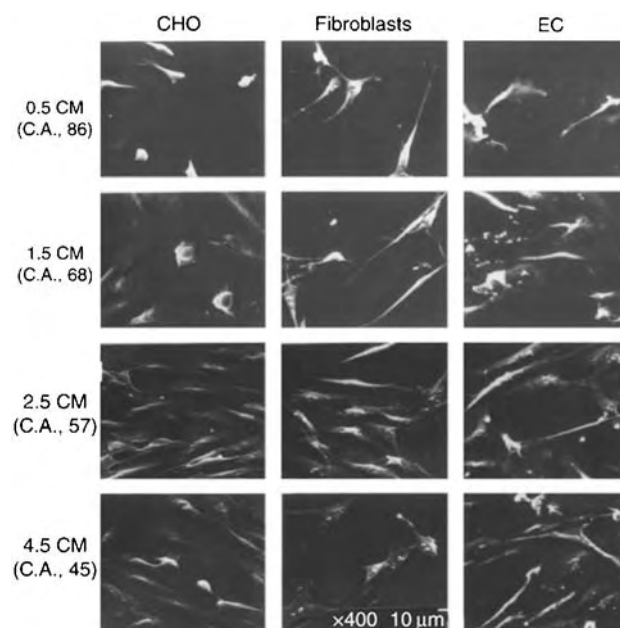


Fig. 26 SEM pictures of different types of cells adhered on wettability gradient PE surfaces after a one-day culture (original magnification, $\times 400$). C.A. means water contact angle ($^\circ$). (From Ref. [79].)

investigated using wettability gradient PE surfaces prepared by corona discharge treatment.^[90] The EC-attached wettability gradient surfaces were mounted on parallel plate flow chambers in a flow system prepared for cell adhesiveness test. Three different shear stresses (150, 200, and 250 dyne/cm^2) were applied to the flow chambers and each shear stress was maintained for 120 minutes to investigate the effect of shear stress and surface hydrophilicity on the EC adhesion strength (Fig. 27). It was observed that the ECs were adhered more onto the positions with moderate hydrophilicity of the wettability gradient surface than onto the more hydrophobic or hydrophilic positions. The maximum adhesion of the cells appeared at around water contact angles of 55° . The EC adhesion strength was higher on the hydrophilic positions than on the hydrophobic ones. However, the maximum adhesion strength of the cells also appeared at around water contact angles of 55° . More than 90% of the adhered cells remained on that position after applying the shear stress— 250 dyne/cm^2 for two hours—whereas the cells were completely detached on the hydrophobic position (water contact angle, about 86°) within 10 minutes after applying the same shear stress. Fig. 28 shows a series of sequential images of ECs adhered on hydrophobic (0.5 cm) and hydrophilic (2.5 cm) positions of the wettability gradient surface at shear stress, 250 dyne/cm^2 . We can see that almost all cells were detached from the hydrophobic position during

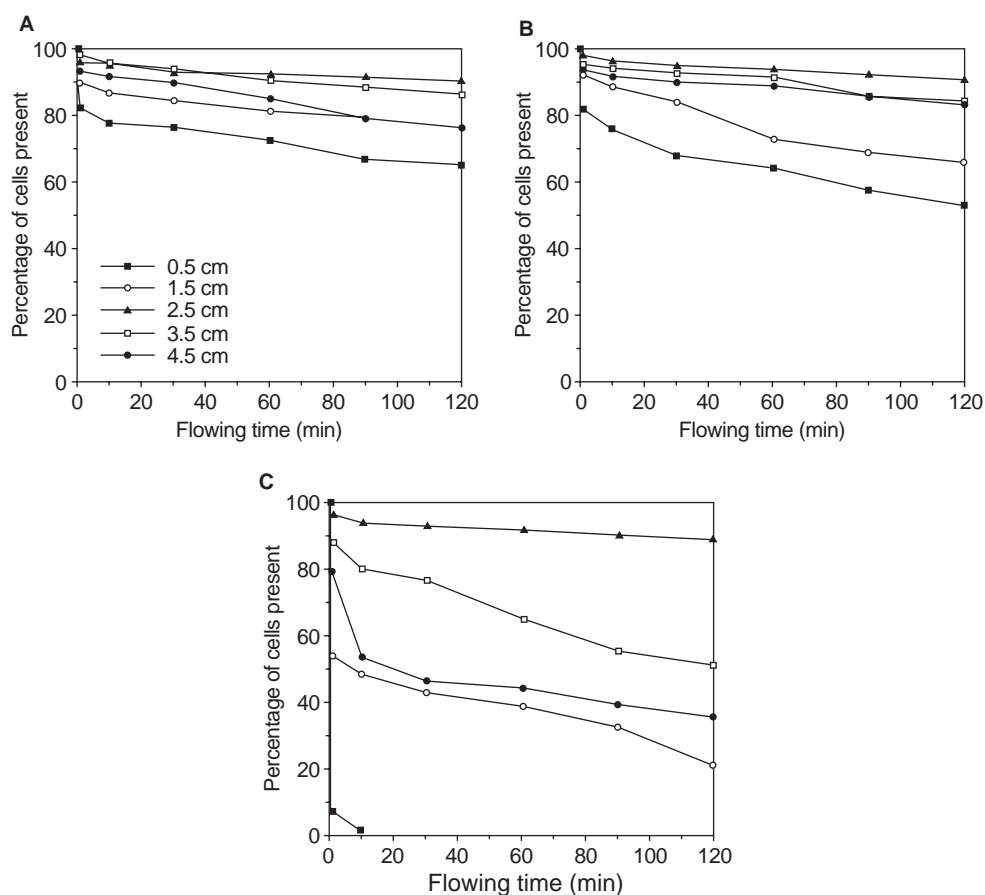


Fig. 27 The percentage of cells present on the positions of wettability gradient PE surfaces as a function of flowing time for different shear stresses: (A) 150 dyne/cm²; (B) 200 dyne/cm²; and (C) 250 dyne/cm². (From Ref. [90].)

the first one minute after applying the shear stress. It seems that surface hydrophilicity plays a very important role for cell adhesion strength.

The human normal diploid fibroblasts (WI-38 cells) were cultured on the gradient PE sheet grafted with poly(MAPC) (Fig. 12). The adhesion behavior of the WI-38 cells was different depending on the concentration of the seeded WI-38 cells.^[73,74] Fig. 29 shows SEM pictures of the gradient PE surface grafted with poly(MAPC) after 24 hr contact with WI-38. A number of adherent cells on the gradient PE surface grafted with poly(MPC) and poly(MDPC) were observed similar to the original PE surface. On the other hand, the gradient PE surface grafted with poly(MHPC) effectively reduced WI-38 cell adhesion. The number of adherent WI-38 cells on the gradient PE surface grafted with poly(MAPC) is summarized in Fig. 30.

By increasing the concentration of the seeded WI-38 cells, the number of adherent WI-38 cells on the gradient PE surface grafted with the poly(MAPC) was increased. In a comparison among the poly(MAPC)s, the poly(MHPC) most effectively reduced

cell adhesion. On every PE surface grafted with the poly(MAPC), the number of adherent cells increased with an increase in the concentration of the seeded cells. On the gradient PE surface grafted with

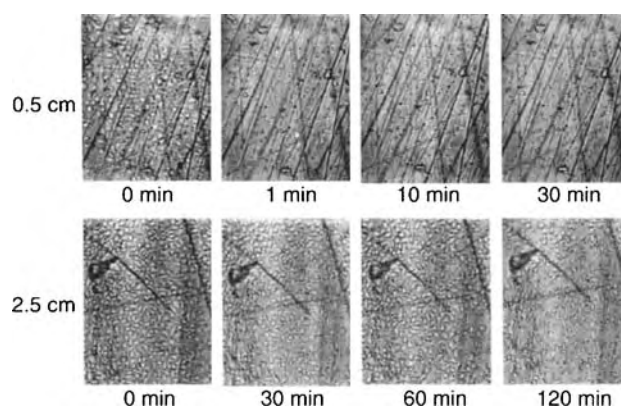


Fig. 28 CCD images of cells adhered on hydrophobic (0.5 cm) and hydrophilic (2.5 cm) positions of wettability gradient PE surfaces as a function of flowing time (shear stress, 250 dyne/cm²). (From Ref. [90].)

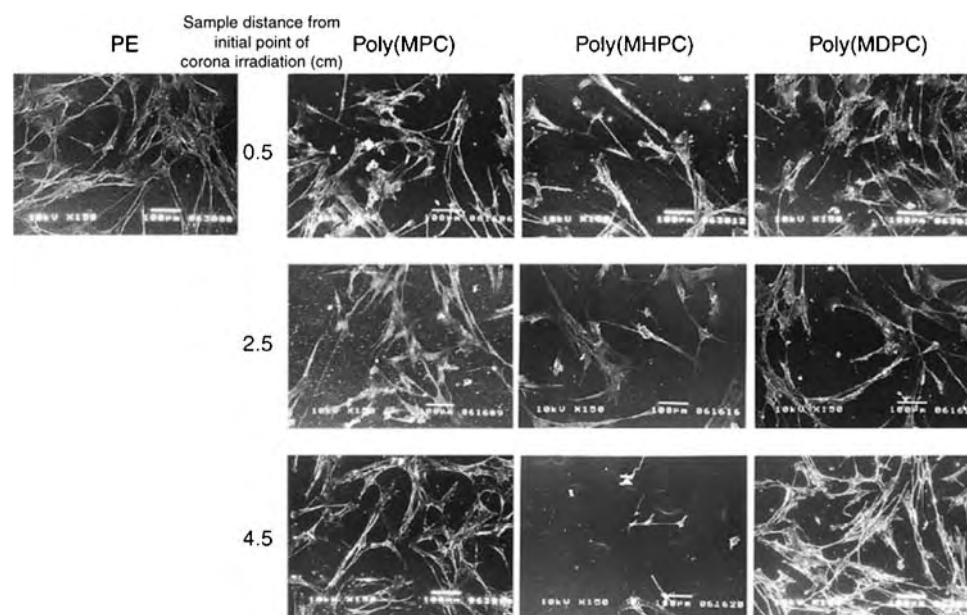


Fig. 29 SEM pictures of the gradient PE sheet grafted with the poly(MAPC) at positions 0.5, 2.5, and 4.5 cm from the starting point of corona treatment, after contact with WI-38 for 24 hr. $[WI-38] = 5.0 \times 10^4$ cells/ml. (From Ref. [74].)

poly(MHPC), adhesion of the WI-38 cells was reduced even when the concentration of the seeded WI-38 cells was high. The cell adhesion can be nonspecifically reduced by modification with a highly oriented phosphorylcholine group.

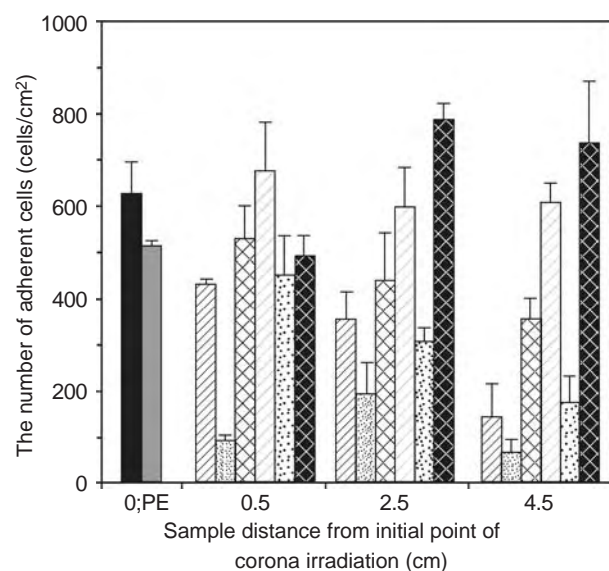


Fig. 30 The number of adherent WI-38 cells on the gradient PE sheet determined with SEM observation; ■ PE; ▨ poly(MPC); ▩ Poly(MHPC); ▤ poly(MDPC). $[WI-38] = 3.0 \times 10^4$ cells/ml, ■ PE; ▨ poly(MPC); ▩ poly(MHPC); ▤ poly(MDPC). $[WI-38] = 5.0 \times 10^4$ cells/ml. (From Ref. [74].)

CONCLUSIONS

The interactions of different types of proteins or cells with various solid substrates depend mainly on surface characteristics such as wettability (hydrophilicity/hydrophobicity or surface free energy), chemistry, charge, roughness, and rigidity. We focused on the surface wettability (particularly on wettability gradient surfaces) in this review. The wettability gradient surfaces have been prepared to study the interactions of proteins or cells since 1987, by many research groups. They were prepared onto silicon or glass substrates by the diffusion method and onto various polymer substrates by plasma discharge treatment and by corona discharge treatment. In most studies, a region of the gradient could be identified where the number of adhering proteins and cells, and their corresponding spread areas were maximal, usually at water contact angles with a specific hydrophilic point.

The major advantage of the use of wettability gradient surfaces as a tool to study protein adsorption/desorption and cell adhesion/proliferation on solid substrates is that several wettabilities are effectively compared at the same time and on the same surface. This is convenient and time saving. The wettability gradient can be used to systematically investigate the interactions of the different types of biological species in terms of the surface wettability of solid substrates. Other specific functional groups or charge gradient surfaces can be produced using plasma or corona discharge treatment and the following graft reactions, which give us useful information on the interaction

behaviors of biological or other species continuously related to the surface chemistry, positive or negative charge, and density of solid substrates. We believe that this method will be useful in study of the analysis of functional biological species on the gradient surface through one experiment because the designed gradient PE surface is experimentally simpler than using a series of single surfaces.

REFERENCES

1. <http://en.wikipedia.org/wiki/Biomaterial>, www.biology-online.org/dictionary/biomaterials; http://www.lexicon-biology.com/biology/definition_29.html.
2. Gristina, A.G. Biomaterial-centered infection: microbial adhesion versus tissue integration. *Science* **1987**, *237*, 1588–1595.
3. Vroman, L. *Biocompatible Polymers, Metals and Composites*; Szycher, M., Ed.; Technomic: Lancaster, 1983; 81 pp.
4. Gumbiner, B.M. Cell adhesion: the molecular basis of tissue architecture and morphogenesis. *Cell* **1996**, *84*, 345–357.
5. Schakenraad, J.M.; Busscher, H.J.; Wildevuur, C.R.H.; Arends, J. The influence of substratum surface free energy on growth and spreading of human fibroblasts in the presence and absence of serum proteins. *J. Biomed. Mater. Res.* **1986**, *20*, 773–784.
6. Tamada, Y.; Ikada, Y. Effect of preadsorbed proteins on cell adhesion to polymer surfaces. *J. Colloid Interf. Sci.* **1993**, *155*, 334–339.
7. van Wachem, P.B.; Beugeling, T.; Feijen, J.; Bantjes, A.; Detmaers, J.P.; van Aken, W.G. Interaction of cultured human endothelial cells with polymeric surfaces of different wettabilities. *Biomaterials* **1985**, *6*, 403–408.
8. Lu, D.R.; Park, K. Effect of surface hydrophobicity on the conformational changes of adsorbed fibrinogen. *J. Colloid Interf. Sci.* **1991**, *144*, 271–281.
9. Johnson, S.D.; Anderson, J.M.; Marchant, R.E. Biocompatibility studies on plasma polymerized interface materials encompassing both hydrophobic and hydrophilic surfaces. *J. Biomed. Mater. Res.* **1992**, *26*, 915–935.
10. Joschek, S.; Nies, B.; Krotz, R.G.; Opferich, A. Chemical and physicochemical characterization of porous hydroxyapatite ceramics made of natural bone. *Biomaterials* **2000**, *21*, 1645–1658.
11. Oh, E.; Luner, P.E. Surface free energy of ethylcellulose films and the influence of plasticizers. *Int. J. Pharm.* **1999**, *188*, 203–219.
12. Lee, J.H.; Park, J.W.; Lee, H.B. Cell adhesion and growth on polymer surfaces with hydroxyl groups prepared by water vapor plasma treatment. *Biomaterials* **1991**, *12*, 443–448.
13. Andrade, J.D.; Smith, L.M.; Gregonis, D.E. The contact angle and interface energetics. In *Surface and Interfacial Aspects of Biomedical Polymers. Vol. 1. Surface Chemistry and Physics*; Andrade, J.D., Ed.; Plenum Press: New York, 1985; 249–292.
14. Weiss, L. The adhesion of cells. *Int. Rev. Cytol.* **1960**, *9*, 187–225.
15. Kaelble, D.H.; Moacanin, J. A surface energy analysis of bioadhesion. *Polymer* **1977**, *18*, 475–482.
16. Baier, R.E.; Meyer, A.E.; Natiella, J.R.; Natiella, R.R.; Carter, J.M. Surface properties determine bioadhesive outcomes: methods and results. *J. Biomed. Mater. Res.* **1984**, *18*, 337–355.
17. van der Valk, P.; van Pelt, A.W.J.; Busscher, H.J.; De Jong, H.P.; Wildevuur, C.R.H.; Arends, J. Interaction of fibroblasts and polymer surfaces: relationship between surface free energy and fibroblast spreading. *J. Biomed. Mater. Res.* **1983**, *17*, 807–817.
18. Grinnell, F. Cellular adhesiveness and extracellular substrates. *Int. Rev. Cytol.* **1978**, *53*, 65–144.
19. Absolom, D.R.; Hawthorn, L.A.; Chang, G. Endothelialization of polymer surfaces. *J. Biomed. Mater. Res.* **1988**, *22*, 271–285.
20. Tamada, Y.; Ikada, Y. Cell attachment to various polymer surface. In *Polymer in Medicine II*; Chellini, E., Giusti, P., Migliaresi, C., Nicolais, L., Eds.; Plenum Press: New York, 1986; 101–115.
21. Jansen, J.A.; van der Waerden, J.P.C.M.; de Groot, K. Fibroblast and epithelial cell interactions with surface-treated implant materials. *Biomaterials* **1991**, *12*, 25–31.
22. Van Kooten, T.G.; Schakenraad, J.M.; van der Mei, H.C.; Busscher, J.H. Influence of substratum wettability on the strength of adhesion of human fibroblasts. *Biomaterials* **1992**, *13*, 897–904.
23. Uyen, H.M.W.; Schakenraad, J.M.; Sjollem, J.; Noordmans, J.; Jongbloed, W.L.; Stokroos, I.; Busscher, H.J. Amount and surface structure of albumin adsorbed to solid substrata with different wettabilities in a parallel plate flow cell. *J. Biomed. Mater. Res.* **1990**, *24*, 1599–1614.
24. Ruardy, T.G.; Schakenraad, J.M.; van der Mei, H.C.; Busscher, H.J. Preparation and characterization of chemical gradient surfaces and their application for the study of cellular interaction phenomena. *Surf. Sci. Rep.* **1997**, *29*, 1–30.
25. D'Souza, S.F. Immobilization and stabilization of biomaterials for biosensor applications. *Appl. Biochem. Biotechnol.* **2001**, *96*, 225–238.
26. Upadhyaya, A.; van Oudenaarden, A. Biomimetic systems for studying actin-based motility. *Curr. Biol.* **2003**, *13*, R734–R744.
27. Fields, G.B.; Lauer, J.L.; Dori, Y.; Forns, P.; Yu, Y.C.; Tirrell, M. Protein-like molecular architecture: biomaterial applications for inducing cellular receptor binding and signal transduction. *Biopolymers* **1998**, *47*, 143–151.
28. Walker, D.S.; Garrison, M.D.; Richert, W.M. Protein adsorption to HEMA/EMA copolymers studied by integrated optical techniques. *J. Colloid Interface Sci.* **1993**, *157*, 41–49.
29. Lydon, M.J.; Minett, T.W.; Tighe, B.J. Cellular interactions with synthetic polymer surfaces in culture. *Biomaterials* **1985**, *6*, 396–402.
30. Elwing, J.; Welin, S.; Askendahl, A.; Nilsson, U.; Lundstrom, I. A wettability gradient method for studies

- of macromolecular interactions at the liquid/solid interface. *J. Colloid Interface Sci.* **1987**, *119*, 203–207.
31. Elwing, J.; Askendal, A.; Lundstrom, I. Protein exchange reactions on solid surfaces studied with a wettability gradient method. *Progr. Colloid Polym. Sci.* **1987**, *74*, 103–107.
 32. Elwing, H.; Askendal, A.; Lundstrom, I. Competition between adsorbed fibrinogen and high-molecular-weight kininogen on solid surfaces incubated in human plasma (the Vroman effect): Influence of solid surface wettability. *J. Biomed. Mater. Res.* **1987**, *21*, 1023–1028.
 33. Elwing, H.; Nilsson, B.; Svensson, K.E.; Askendahl, A.; Nilsson, U.R.; Lundstrom, I. Conformational changes of a model protein (complement factor 3) adsorbed on hydrophilic and hydrophobic solid surfaces. *J. Colloid Interface Sci.* **1988**, *125*, 139–145.
 34. Elwing, H.; Askendal, A.; Lundstrom, I. Desorption of fibrinogen and γ -globulin from solid surfaces induced by a nonionic detergent. *J. Colloid Interface Sci.* **1989**, *128*, 296–300.
 35. Welin-Klintstrom, S.; Wikstrom, M.; Askendal, A.; Elwing, H.; Lundstrom, I.; Karlsson, J.O.; Renvert, S. Proteolytic degradation of fibrinogen layers adsorbed on hydrophilic and hydrophobic surfaces. *Colloids Surf.* **1990**, *44*, 51–60.
 36. Elwing, H.; Golander, C.G. Protein and detergent interaction phenomena on solid surfaces with gradients in chemical composition. *Adv. Colloid Interface Sci.* **1990**, *32*, 317–339.
 37. Welin-Klintstrom, S.; Askendal, A.; Elwing, H. Surfactant and protein interactions on wettability gradient surfaces. *J. Colloid Interface Sci.* **1993**, *158*, 188–194.
 38. Hlady, V.; Golander, C.; Andrade, J.D. Hydrophobicity gradient on silica surfaces: a study using total internal reflection fluorescence spectroscopy. *Colloids Surf.* **1988**, *33*, 185–190.
 39. Golander, C.G.; Lin, Y.S.; Hlady, V.; Andrade, J.D. Wetting and plasma-protein adsorption studies using surfaces with a hydrophobicity gradient. *Colloids Surf.* **1990**, *49*, 289–302.
 40. Hlady, V.; Lin, J.N.; Andrade, J.D. Spatially resolved detection of antibody-antigen reaction on solid/liquid interface using total internal reflection excited antigen fluorescence and charge-coupled device detection. *Biosens. Bioelect.* **1990**, *5*, 291–301.
 41. Hlady, V. Spatially resolved adsorption kinetics of immunoglobulin G onto the wettability gradient surface. *Appl. Spectrosc.* **1991**, *45*, 246–252.
 42. Carley, J.F.; Kitze, P.T. Corona-discharge treatment of polymeric films. II. Chemical studies. *Polym. Eng. Sci.* **1980**, *20*, 330–338.
 43. Lopez, G.P.; Biebuyck, H.A.; Frisbie, C.D.; Whitesides, G.M. Imaging of features on surfaces by condensation figures. *Science* **1993**, *260*, 647–649.
 44. Zhao, H.; Beysens, D. From droplet growth to film growth on a heterogeneous surface: condensation associated with a wettability gradient. *Langmuir* **1995**, *11*, 627–634.
 45. Lin, Y.S.; Hlady, V.; Golander, C.G. The surface density gradient of grafted poly(ethylene glycol): preparation, characterization and protein adsorption. *Colloids Surf. B Biointerfaces* **1994**, *3*, 49–62.
 46. Liedberg, B.; Tengvall, P. Molecular gradients of omega-substituted alkanethiols on gold: preparation and characterization. *Langmuir* **1995**, *11*, 3821–3827.
 47. Riepl, M.; Ostblom, M.; Lundstrom, I.S.; Svensson, C.T.; Denier van der Gon, A.W.; Schaferling, M.; Liedberg, B. Molecular gradients: an efficient approach for optimizing the surface properties of biomaterials and biochips. *Langmuir* **2005**, *21*, 1042–1050.
 48. Pitt, W.G. Fabrication of a continuous wettability gradient by radio frequency plasma discharge. *J. Colloid Interface Sci.* **1989**, *133*, 223–227.
 49. Golander, C.G.; Pitt, W.G. Characterization of hydrophobicity gradients prepared by means of radio frequency plasma discharge. *Biomaterials* **1990**, *11*, 32–35.
 50. Lee, J.H.; Park, J.W.; Lee, H.B. Characterization of hydroxyl group gradient surfaces prepared by water vapor plasma treatment. *Polymer (Korea)* **1990**, *14*, 646–652.
 51. Lee, H.B.; Andrade, J.D. Cell adhesion on gradient surfaces. *Trans. 3rd World Biomaterials Congr.* **1988**, *7*, 43.
 52. Lee, H.B. Application of synthetic polymers in implants. In *Frontiers of Macromolecular Science*; Saegusa, T., Higashimura, T., Abe, A., Eds.; Blackwell Scientific Publications: Oxford, 1989; 579–584.
 53. Bamford, C.H.; Ward, J.C. The effect of the high-frequency discharge on the surfaces of solids. I. The production of surface radicals on polymers. *Polymer* **1961**, *2*, 277–293.
 54. Morosoff, N.; Crist, B.; Bumgarner, M.; Hsu, T.; Yasuda, H. Free radicals resulting from plasma polymerization and plasma treatment. *J. Macromol. Sci. Chem.* **1976**, *A10*, 451–471.
 55. Mitchell, J., Jr.; Perkins, L.R. Determination of hydroperoxide groups in oxidized polyethylene. *Appl. Polym. Symp.* **1967**, *4*, 167–173.
 56. Clark, D.T.; Dilks, A. ESCA applied to polymers. XXII. RF glow discharge modification of polymers in pure oxygen and helium-oxygen mixtures. *J. Polym. Sci., Polym. Chem. Ed.* **1979**, *17*, 957–976.
 57. Steinhauser, H.; Ellinghorst, G. Corona treatment of isotactic polypropylene in nitrogen and carbon dioxide. *Angew. Makromol. Chem.* **1984**, *120*, 177–191.
 58. Iwata, H.; Kishida, A.; Suzuki, M.; Hata, Y.; Ikada, Y. Oxidation of polyethylene surface by corona discharge and the subsequent graft polymerization. *J. Polym. Sci. Polym. Chem. Ed.* **1988**, *26*, 3309–3322.
 59. Chan, C.M.; Ki, T.M.; Hiraoka, H. Polymer surface modification by plasmas and photons. *Surf. Sci. Rep.* **1996**, *24*, 1–54.
 60. Lee, J.H.; Kim, H.G.; Khang, G.; Lee, H.B.; Jhon, M.S. Characterization of wettability gradient surfaces prepared by corona discharge treatment. *J. Colloid Interface Sci.* **1992**, *151*, 563–570.
 61. Lee, J.H.; Lee, H.B. A wettability gradient as a tool to study protein adsorption and cell adhesion on polymer surfaces. *J. Biomater. Sci., Polymer Ed.* **1993**, *4*, 467–481.



62. Kim, H.G.; Lee, J.H.; Lee, H.B.; Jhon, M.S. Dissociation behavior of surface-grafted poly(acrylic acid): effects of surface density and counterion size. *J. Colloid Interface Sci.* **1993**, *157*, 82–87.
63. Lee, J.H.; Kim, H.W.; Pak, P.K.; Lee, H.B. Preparation and characterization of functional group gradient surfaces. *J. Polym. Sci. A Polym. Chem.* **1994**, *32*, 1569–1579.
64. Jeong, B.J.; Lee, J.H.; Lee, H.B. Preparation and characterization of comb-like PEO gradient surfaces. *J. Colloid Interface Sci.* **1996**, *178*, 757–763.
65. Lee, J.H.; Lee, J.W.; Khang, G.; Lee, H.B. Interaction of cells on chargeable functional group gradient surfaces. *Biomaterials* **1997**, *18*, 351–358.
66. Lee, J.H.; Jung, H.W.; Kang, I.J.; Lee, H.B. Cell behavior on polymers surfaces with different functional groups. *Biomaterials* **1994**, *15*, 705–711.
67. Lee, J.H.; Khang, G.; Lee, J.W.; Lee, H.B.; Platelet adhesion onto chargeable functional group gradient surfaces. *J. Biomed. Mater. Res.* **1998**, *40*, 180–186.
68. Kim, M.S.; Seo, K.S.; Khang, G.; Lee, H.B. First preparation of biotinylated gradient polyethylene surface to bind photoactive caged streptavidin. *Langmuir* **2005**, *21*, 4066–4070.
69. Kim, M.S.; Seo, K.S.; Khang, G.; Lee, H.B. Preparation of a gradient biotinylated polyethylene surface to bind streptavidin-FITC. *Bioconjug. Chem.* **2005**, *16*, 245–249.
70. Burnham, N.L. Polymers for delivering peptides and proteins. *Am. J. Hosp. Pharm.* **1994**, *51*, 210–218.
71. Greenwald, R.B.; Choe, Y.H.; McGuire, J.; Conover, C.D. Effective drug delivery by PEGylated drug conjugates. *Adv. Drug Deliv. Rev.* **2003**, *55*, 217–250.
72. Lee, J.H.; Jeong, B.J.; Lee, H.B. Plasma protein adsorption and platelet adhesion onto comb-like PEO gradient surfaces. *J. Biomed. Mater. Res.* **1997**, *34*, 105–114.
73. Iwasaki, Y.; Ishihara, K.; Nakabayashi, N.; Khang, G.; Jeon, J.H.; Lee, J.W.; Lee, H.B. Platelet adhesion on the gradient surfaces grafted with phospholipid polymer. *J. Biomater. Sci. Polym. Ed.* **1998**, *9*, 801–816.
74. Iwasaki, Y.; Sawada, S.; Nakabayashi, N.; Khang, G.; Lee, H.B.; Ishihara, K. The effect of the chemical structure of the phospholipid polymer on fibronectin adsorption and fibroblast adhesion on the gradient phospholipid surface. *Biomaterials* **1999**, *20*, 2185–2191.
75. Bain, C.D.; Whitesides, G.M. Depth sensitivity of wetting: monolayers of omega-mercapto ethers on gold. *J. Am. Chem. Soc.* **1988**, *110*, 5897–5898.
76. Troughton, E.B.; Bain, C.D.; Whitesides, G.M.; Nuzzo, R.G.; Allara, D.L.; Porter, M.D. Monolayer films prepared by the spontaneous self-assembly of symmetrical and unsymmetrical dialkyl sulfides from solution onto gold substrates: structure, properties, and reactivity of constituent functional groups. *Langmuir* **1988**, *4*, 365–385.
77. Johnson, R.E.; Dettre, R.H. *Surface and Colloid Science*; Matijevic, E., Ed.; Wiley Interscience: New York, 1969; Vol. 2, 85 pp.
78. Adamson, A.W. *Physical Chemistry of Surfaces*, 3rd Ed.; Wiley: New York, 1976.
79. Lee, J.H.; Khang, G.; Lee, J.W.; Lee, H.B. Interaction of different types of cells on polymer surfaces with wettability gradient. *J. Colloid Interface Sci.* **1998**, *205*, 323–330.
80. Ruardy, T.G.; Moorlag, H.E.; Schakenraad, J.M.; van der Mei, H.C.; Busscher, H.H. Growth of fibroblasts and endothelial cells on wettability gradient surfaces. *J. Colloid Interface Sci.* **1997**, *188*, 209–217.
81. Ueda-Yukoshi, T.; Matsuda, T. Cellular responses on a wettability gradient surface with continuous variations in surface compositions of carbonate and hydroxyl groups. *Langmuir* **1995**, *11*, 4135–4140.
82. Azzam, R.M.A.; Bashara, N.M. *Ellipsometry and Polarized Light*; Elsevier: Amsterdam, 1987; 269 pp.
83. Andrade, J.D. X-ray photoelectron spectroscopy (XPS). In *Surface and Interfacial Aspects of Biomedical Polymers Vol. 1. Surface Chemistry and Physics*; Andrade, J.D., Ed.; Plenum Press: New York, 1985; 105–195.
84. Lin, J.N.; Drake, B.; Lea, A.S.; Hansma, P.K.; Andrade, J.D. Direct observation of immunoglobulin adsorption dynamics using the atomic force microscope. *Langmuir* **1990**, *6*, 509–511.
85. Hlady, V.; van Waggenen, R.A.; Andrade, J.D. Total internal reflection intrinsic fluorescence (TIRF) spectroscopy applied to protein adsorption. In *Surface and Interfacial Aspects of Biomedical Polymers. Vol. 2. Protein Adsorption*; Andrade, J.D., Ed.; Plenum Press: New York, 1985; 81–119.
86. Roberson, S.V.; Fahey, A.J.; Sehgal, A.; Karim, A. Multifunctional TOF-SIMS: combinatorial mapping of gradient energy substrates. *Appl. Surf. Sci.* **2002**, *200*, 150–164.
87. Lee, J.H. Interactions of PEO-containing polymeric surfactants with hydrophobic surfaces. Ph. D. thesis, University of Utah, Salt Lake City, 1988.
88. Green, N. Avidin. 1. The use of (14-C) biotin for kinetic studies and for assay. *Biochem. J.* **1963**, *89*, 585–591.
89. Carter, S.B. Principles of cell motility: the direction of cell movement and cancer invasion. *Nature* **1965**, *208*, 1183–1187.
90. Lee, J.H.; Lee, S.J.; Khang, G.; Lee, H.B. The effect of fluid shear stress on endothelial cell adhesiveness to polymer surfaces with wettability gradient. *J. Colloid Interface Sci.* **2000**, *230*, 84–90.

Hard Tissue: Biomaterial Interactions

Feza Korkusuz

Medical Center and Department of Physical Education and Sports, Middle East Technical University, Ankara, Turkey

Petek Korkusuz

Department of Histology and Embryology, Hacettepe University, Sıhhiye, Ankara, Turkey

INTRODUCTION

Bone and cartilage are vulnerable to injury and replacement of these tissues by biomaterials, synthetic and treated natural materials used to replace and/or regenerate functioning tissues, has been successfully used for years. Recent advancements in science elaborated the Tissue Engineering concept where regeneration of injured tissues is intended by the combined application of: a) biomaterial-based scaffolds; b) cells grown in culture; c) systemic and/or local hormones/mediators; and recently d) genetic modulators. Tissue engineering products and basically biomaterials of various shapes and forms have been widely used in musculoskeletal diseases and disorders for decades. Metals of steel, cobalt–chrome, and titanium in pure and/or alloy form, ceramics of hydroxyapatite (HA) and calcium phosphate, and polymers such as polymethyl-methacrylate are currently used in bone, cartilage, and joint replacement. Biocompatibility of biomaterials is simply defined as the *in vivo* functioning of the implant without causing local and/or systemic intolerable responses. Implants used for hard tissue replacement should also possess appropriate mechanical properties as load bearing of bone and lubrication in joints are among the basic physiological functions of these tissues. Research to improve the biocompatibility of biomaterials and tissue engineering products is still going on as bone and cartilage will, nevertheless, cause a response even when “inert” materials are implanted.^[1] Local and systemic responses elaborated by biomaterials depend on various factors including surface topography, heterogeneity, and chemical and physical properties. The tissue–implant interface is, therefore, the hot spot of research for the better understanding of the response to biomaterials and tissue engineering products. Industry has also focused on improving biomaterial surfaces, the front line of tissue–implant interactions.

Eleven million patients in the U.S.A. receive at least one implant each year. These implants can be classified as: a) artificial joints; b) fracture fixation devices; c) artificial heart valves; d) intraocular lenses; e) internal heart stimulators/defibrillators; f) ventilation tubes; g) infusion pumps; h) dental implants; i) silicon

implants; and j) artificial blood vessels. Fracture fixation devices and artificial joints cover almost 50% of these implants and its impact on the implant applied individual and society, excluding its impact on quality of life, productivity, sociology, and psychology, is very high.^[2] Furthermore, it is well appreciated that the demand for fracture fixation and joint replacement devices is expected to increase within the next few decades because of aging population.

As a basic concept, a time, type, and dose-dependent systemic and/or cellular response will be elaborated by any biomaterial when implanted into living tissues.

METALS

Metals and their alloys are widely used as fracture fixation and joint replacement devices. Biocompatibility of metals depends on their: a) type; b) producer and production technique; c) composition; d) mechanical properties; e) structure; and f) geometry. Low carbon concentration in stainless steel is essential for its biocompatibility. Chromium in the microstructure of stainless steel provides corrosion resistance by forming a passive layer on the surface of the implant. Passive layer of titanium oxide on titanium implants provides a significantly better corrosion resistance than stainless steel and cobalt–chromium implants. Titanium alloys have a lower elastic modulus that improves the structural stiffness of the implant compared to steel and cobalt alloys.^[3] One structural disadvantage of titanium alloys is its vanadium content that can be toxic when released, and recent studies^[4] focus on replacing vanadium with niobium to improve their biocompatibility.

Interaction of the implant and the host mainly occurs through its surface, and a) structure and b) coating of the surface are major determinants of biocompatibility. Theoretically, smoother the implant’s surface, lesser the response to it. In recent years, ion implantation techniques were developed and used to obtain smoother surfaces on metallic surfaces. In contrast, various surface modification techniques are developed to obtain a larger surface area with the aim of tissue integration into the implant. Chemical methods are



used to macroscopically and microscopically modify the surface of metallic implants. For example, surface modified titanium implants using phosphoric acid were presented to increase covalent bonding with biological molecules.^[5] One recent *in vitro* study^[6] presented significantly better interaction between coarse-blasted, acid-etched titanium surfaces and osteoblasts. Other methods include: a) sphere;^[7] b) wire mesh; and c) porous surface coatings for surface expansion of metals with the aim of better tissue integration. An optimal pore size of implants for osteoblast but not fibroblast in-growth is assumed to be between 150 and 450 μm . Interconnected pores are preferred for better tissue integration. Other method that enhances tissue integration^[8] and biocompatibility^[9,10] in cementless implant application is HA coating. The thickness and the method of HA coating are important in preventing delamination.^[11] In artificial hip replacement immersion techniques for HA coating of the acetabular component are preferred to overcome the delamination problem. Furthermore, HA coating of the femoral side prevents wear particle migration into the bone-implant interface and avoids early implant failure.^[12,13] Other study, however, revealed that ceramic particles detached from the coating may immigrate into the artificial joint space and can cause wear debris from the polymer.^[14] It is assumed that novel methods such as calcium ion implantation^[15,16] may replace conventional methods in the near future to overcome these problems.

Coating of titanium implants with bioactive molecules other than ceramics for better tissue integration is getting popular. Studies using fibronectin^[17] or type 1 collagen^[18] confirmed the improved osteointegration and cellular attachment. One recent study^[19] regarding dental implants presented the importance of epithelial cells in implant integration. In that study, polarization of cells was superior on the titanium surface while the number of cells was higher when the implant was coated with HA. When evaluated in details, serum extracellular matrix deposition on the substrate is followed by cellular attachment. This event in implant-bone integration can be called "Osteoblast-Substrate Adhesion."^[19] Cellular attachment to implants is modulated by integrin-mediated (extracellular matrix receptor) signal transduction.^[20] Once when cells are attached to the implant they start to spread their focal contact area. Fibronectin and/or laminin play important roles at this stage of cell attachment. Basically fibronectin promotes cell adhesion and reorganization of actin microfilaments within osteoblasts. Surface chemistry of the material^[19] will keep influencing cytoskeletal molecules including vinculin, actin, and focal adhesion kinase following cellular attachment. Cellular attachment is followed by migration and secretion of cytokines and extracellular matrix elements. At

this stage, surface roughness promotes osteogenic differentiation of less mature cells. The optimal rate and shape of surface roughness however is not determined. It is well noted that increasing surface area will improve osteogenic differentiation. One recent study^[21] using scanning electron microscopy and double immune labeling of vinculin and actin presented marked synergistic effect of micro- and nanoscale topography on cell proliferation. On the other hand, this may increase the incidence of particle detachment from the surface. The so-called "Particle Disease" may lead to implant loosening and other severe disorders defined in the literature.^[22-25] Optimization studies of surface area and roughness in the benefit of osteoblast attachment and differentiation and particle detachment should be therefore taken into consideration. Adhesion molecules lead to polarization and colony formation of cells in *in vitro* conditions. *In vitro* conditions may not necessarily simulate *in vivo* conditions as agonist and antagonist factors of cellular adhesion and colony formation may differ when a material is implanted into the tissue.

Maintaining vascularity of the implant applied bone site during surgery is as important as material and surface characteristics of the implant for integration. Tissue damage will initiate a cascade of apoptosis, osteoclastic activity, bone resorption (osteolysis), and implant failure. Furthermore, an ill-vascularized bone is prone to infection. A well-vascularized bone, on the other hand, may establish strong bonds with the implant through adhesion molecules such as fibronectin, laminin, fibrinogen, and vitronectin owing to the establishment of a metabolically active tissue.

Microscopical and molecular interaction between the implant and the hosting bone is broadly evaluated in material science and dental surgery.^[26] Microparticles produced or migrating into the bone-implant interface initiate a cascade of inflammatory reaction according to their type, size, and amount.^[27] As a general rule, when the components of the implant (i.e., metal, polymer, and ceramic) increase, the incidence of wear debris will also increase. Particles from metals are detached by corrosion, abrasion, adhesion, and/or delamination. Interfaces such as the implant-bone, implant-cement, and cement-bone are the sites of wear production. Initial response to the debris is elicited and mediated by osteoclasts.^[28] Release of large amounts of osteolytic factors such as interleukin-1 (IL-1),^[29] interleukin-6 (IL-6), and tumor necrosis factors alpha and beta (TNF- α and - β) from macrophages in response to wear debris will eventually lead to chronic inflammation and implant failure.^[30] Cell types in periprosthetic membranes surrounding failed hip arthroplasties of 10 patients were macrophages and fibroblasts.^[31] These fibroblasts differentiated peripheral blood mononuclear cells into mature osteoclasts. Inhibition of inflammatory

cytokines^[31] and/or type 1 IL-1 receptors^[29] are proposed to control periprosthetic osteolysis.

Metals and basically cobalt–chrome alloys may also discharge ions which could mediate cytokine release and hypersensitivity reaction.^[32] A hypersensitivity test before metal implantation is therefore recommended when available. One other cellular response to metals implanted into bone is apoptosis that is significantly higher in failed ones than in stable ones.^[33]

BIOCERAMICS

Bioceramics are basically divided into a) fast resorbing; b) slow resorbing; and c) injectable ceramics.^[34] High temperature sintered (>1000°C) HA is a slow resorbing bioceramic with a 1–3% degradation rate per year. Calcium phosphates' degradation rate, on the other hand, is 35–40% per year. These ceramics alone or as a composition are used as scaffolds or void fillers in non-weight bearing sites for osteoconduction. Apart from osteoconduction, these ceramics were effectively used as antibiotic^[35] or as anticancer drug^[36] carriers in bone sites. One recent study^[37] focused on the improvement of mechanical determinants of HA–calcium phosphate

ceramics that may enable their usage in load bearing sites.

Ceramics have also been used in hip joint replacement as the femoral head and acetabular cup. Their gliding properties are superior to metals and polymers. Load bearing is balanced as long as they do not present microfractures. They obviously last longer than any other material and do not cause wear debris.

Ceramics that are implanted as porous scaffolds to bone without cells and growth factors are involved in a cascade of events similar to the molecular and cellular events seen in fracture healing. This cascade includes: a) hematoma; b) inflammation; c) neovascularization; d) osteoclastic resorption; and e) new bone formation. Osteoclastic resorption and new bone formation are coupled with the integration process. First line of cells to get in touch with the ceramics is assumed to be monocytes that stimulate IL-1, IL-6, and TNF- α secretion. In dental ceramic implants, endothelial cells are also involved in this process. Endothelin is effective in neovascularization and enhances ceramic integration and bone healing. It was observed that when 1 μ g of endothelin was applied to bone together with HA and calcium phosphate composite (60/40) lamellar bone formation was obtained (Figs. 1A and 1B). Bioceramics

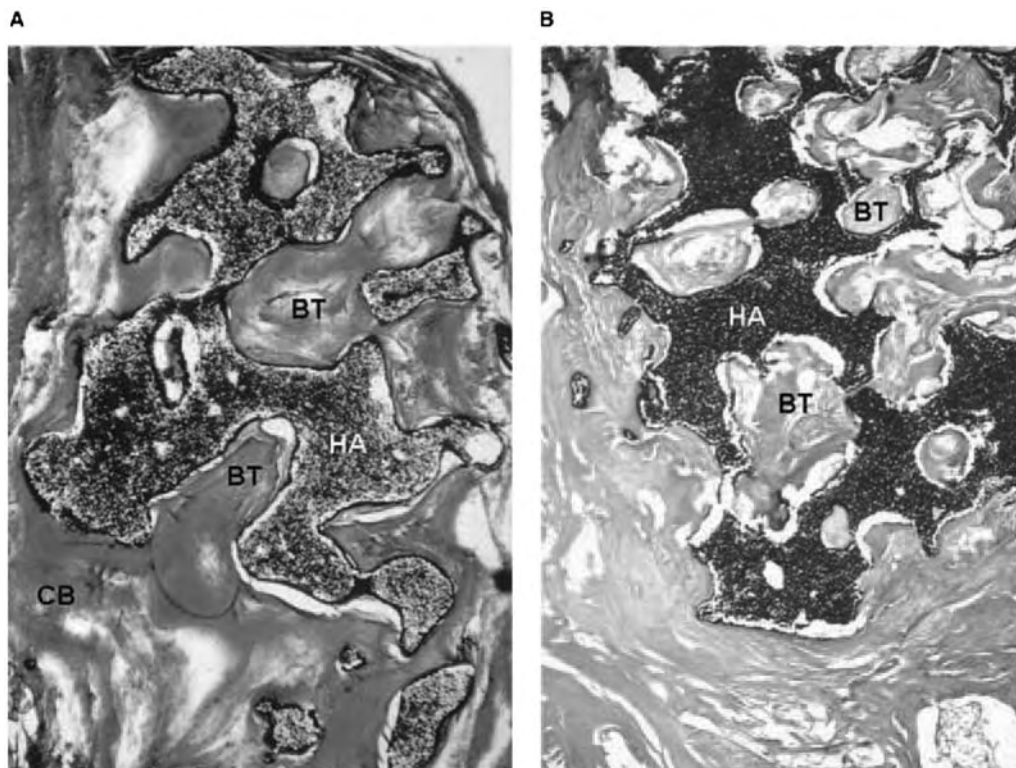


Fig. 1 (A) ET-1 antagonist bositentan loaded HA/tricalcium phosphate (60/40) composite implanted into rabbit critical size segmental tibial defect; six weeks. 250 \times , Masson trichrome staining and (B) HA/tricalcium phosphate (60/40) composite implanted into rabbit critical size segmental tibia defect without active molecule; six weeks. BT: new bone trabecules; HA: hydroxyapatite; 250 \times , hematoxylin and eosin staining.

are defined as bioactive materials. They cause no or minimal tissue reaction when implanted into bone, and chronic inflammation is almost never observed. Giant cells and macrophages are seldom seen around ceramic implants. New bone formation begins from the outer surface and proceeds to the center of the ceramic. The presence of free calcium ions on the surface of HA ceramic, but not carbonate, sulfate, and sodium ions, significantly enhances protein binding in a dose-dependent manner.^[38] Proteoglycans and glycosaminoglycans of bone were proposed to have important roles in the mineralization process.^[39] Bone sialoprotein and osteopontin were also presented^[40] to bind to HA ceramics for the regulation of bone mineralization.

POLYMERS

Polymers are promising materials to be used in bone^[41] and cartilage^[42,43] engineering. They can effectively be used in controlled delivery of cells and mediators to the injury site and can aid in regeneration. Polylactides, polyglycolides, and polyhydroxybutyrates (PHBV) are effectively used for these reasons. Macrophages increase in number in about 12 weeks of polymer application into the bone.^[44] Advancements in bone morphogenic protein (BMP)-2 (Fig. 2) and -7^[41,45] delivery and chondrocyte^[43] transfer using polymers were experimentally achieved in the last few years.

Scaffolds for tissue engineering, whether ceramic or polymer in nature, should be: a) biocompatible;

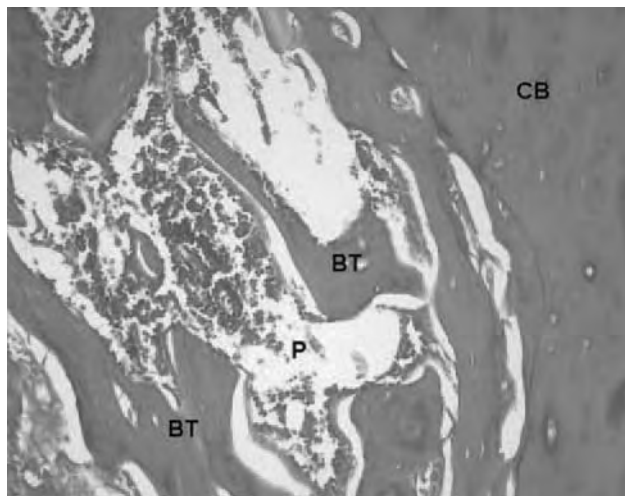


Fig. 2 One microgram BMP-2 loaded PHBV implanted rat tibial cortical defect. Note that a partial cortical defect was established to prevent the migration of cells and mediators from the bone marrow; three weeks. BT: new bone trabeculae; P: polymer; CB: cortical bone; 200 \times , hematoxylin and eosin staining.

b) porous; c) cell and mediator friendly; d) resorbable; e) mechanically suitable; f) inexpensive; and g) surgically easy applicable. Characterization of these implants before experimental or clinical application is essential, i.e., a) scanning electron microscopy; b) electron diffraction spectroscopy; c) porosimetry; d) pore size distribution; e) internal void volume assessment; and f) degradation profile are usually assessed.

The sources of osteoblasts and chondrocytes are many. Bone marrow is the source of stem cells and therefore widely used to obtain osteoblast precursors. The cambium layer of the periosteum has the potential to contain osteoblast and chondrocyte precursors. Cartilage tissue itself is also a major source for chondrocytes. Adipocytes have also attracted attention recently as they can be induced into osteoblast-like cells *in vitro* under appropriate stimulus.

Local mediators, except for common features, may vary from tissue to tissue. These features include: a) time and dose-dependent action; b) possible side and adverse effects even when applied locally; and c) triggering antagonistic mechanisms and losing activity after local application. Various mediators may have different effects at different stages of bone and cartilage regeneration. For example, one mediator that is effective at the proliferation stage may lose its effect at the differentiation stage. Besides local mediators and cells, controlled mechanical stimulation for musculoskeletal tissue healing and regeneration is additionally essential.

Factors that influence bone regeneration are complex (Fig. 3). Among the factors are: a) mechanical and physical forces; b) systemic hormones; and c) neuro-mediators. All these three factors control bone cell proliferation and differentiation through local mediators. Local mediators stimulate osteoblasts and/or osteoclast formation/activity/neogenesis; the former causing bone formation and the latter mediating bone resorption. In normal bone, osteoblastic and osteoclastic activity is coupled and balanced. The balance between osteoblastic formation and osteoclastic resorption is controlled, among other factors, by osteoprotegerin and its ligand.^[46] When applied to bone externally the type, time, and dose of local mediators should be well defined.

Cartilage is a highly differentiated tissue, and tissue engineering is recently used for the repair of its defects.^[47] Polymers alone^[43] or in combination with ceramics^[47] are preferred for cartilage replacement; the former is used as chondrocyte carrier for the replacement of the cartilage layer and the latter for the replacement of subchondral bone. Some biomaterials were further used to transfer transforming growth factor beta and mesenchymal stem cells for cartilage repair.^[48] Such implants were also used in clinical practice in limited cases to replace the injured cartilage.^[49,50]

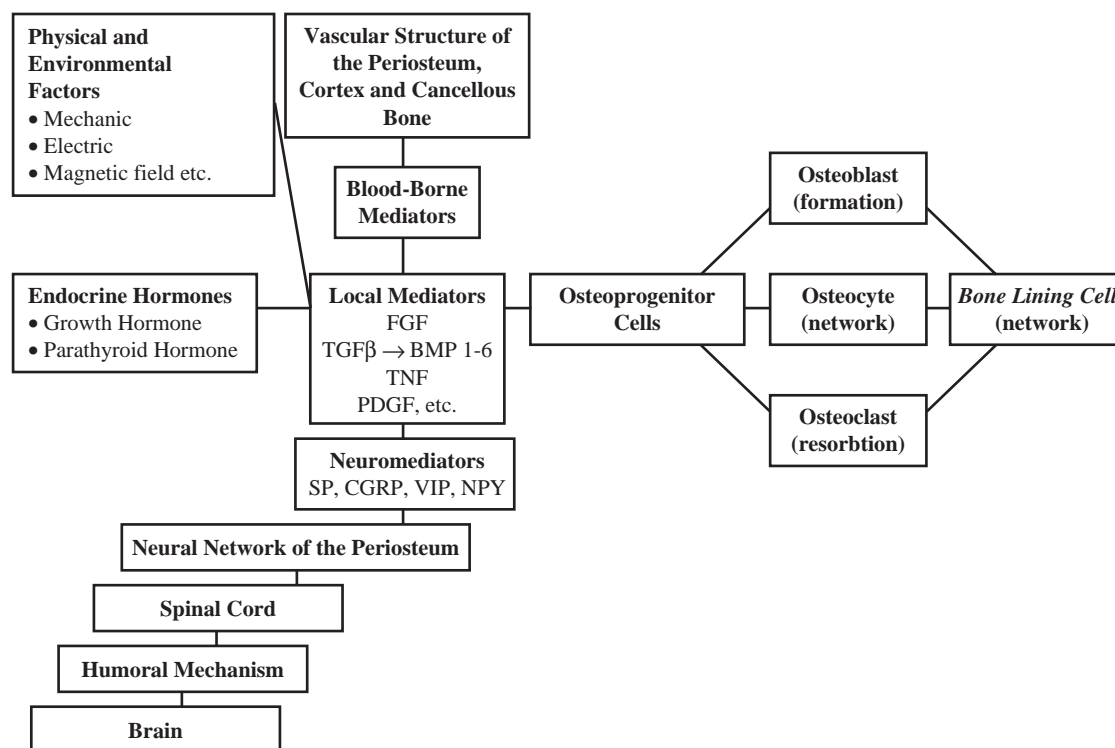


Fig. 3 Systems that control bone repair and regeneration.

CONCLUSIONS

As a conclusion, biomaterials are frequently and widely used in bone and cartilage tissue. Metals are the most common used biomaterials in bone, and extensive research is focused on surface coating or engineering of metals. Advanced techniques in surface coating are assumed to decrease particle release from metal implants and increase biocompatibility. Another approach with metal implants is to decrease their profile and contact area with bone. Ceramic and polymer based implants are likely to be more frequently used as tissue engineering products leading to complete regeneration of bone and cartilage tissue, and such products are already used in clinical applications.

ARTICLES OF FURTHER INTEREST

Biocompatibility Testing; Biologic and Synthetic Apatites; Tissue Engineering of Bone; Titanium and Its Alloys

REFERENCES

- Santavirta, S.; Gristina, A.; Kontinen, Y.T. Cemented versus cementless hip arthroplasty: a review of prosthetic biocompatibility. *Acta Orthop. Scand.* **1992**, *63*, 225–232.
- Praemer, A.; Furner, S.; Rice, D.P. *Musculoskeletal Conditions in the United States*; American Academy of Orthopaedic Surgeons: Illinois, U.S.A., 1992; 127–141.
- Wright, T.M.; Li, S. Biomaterials. In *Orthopaedic Basic Science: Biology and Biomechanics of the Musculoskeletal System*, 2nd Ed. Ed.; Buckwalter, J.A., Einhorn, T.A., Simon, S.R., Eds.; American Academy of Orthopaedic Surgeons: Illinois, U.S.A., 2000; 181–215.
- Spirano, S.; Bosetti, M.; Bronzoni, M.; Verne, E.; Maina, G.; Bergo, V.; Cannas, M. Surface properties and cell response of low metal ion release Ti-6Al-7Nb alloy after multi-step chemical and thermal treatments. *Biomaterials* **2005**, *26*, 1219–1229.
- Nanci, A.; Wuest, J.D.; Peru, L.; Brunet, P.; Sharma, V.; Zalzal, S.; McKee, M.D. Chemical modification of titanium surfaces for covalent attachment of biological molecules. *J. Biomed. Mater. Res.* **1998**, *40*, 324–335.
- Wieland, M.; Textor, M.; Chehroudi, B.; Brunette, D.M. Synergistic interaction of topographic features in the production of bone-like nodules on Ti surfaces by rat osteoblasts. *Biomaterials* **2005**, *26*, 1119–113026.
- Hofmann, A.A.; Bloebaum, R.D.; Bachus, K.N. Progression of human bone ingrowth into porous-coated implants. *Acta Orthop. Scand.* **1997**, *68*, 161–166.
- Rashmir-Raven, A.M.; Richardson, D.C.; Aberman, H.M.; DeYoung, D.J. The response of cancellous and cortical canine bone to hydroxyapatite-coated and uncoated titanium rods. *J. App. Biomaterials* **1995**, *6*, 237–242.
- Coathup, M.J.; Blunn, G.W.; Flynn, N.; Williams, C.; Thomas, N.P. A Comparison of bone remodeling

- around hydroxyapatite-coated, porous-coated and grit-blasted hip replacements retrieved at post-mortem. *J. Bone Joint Surg.* **2001**, *83-B*, 118–123.
10. Montanaro, L.; Arciola, C.R.; Campoccia, D.; Cervellati, M. In vitro effects of MG63 osteoblast-like cells following contact with two roughness-differing fluorohydroxyapatite-coated titanium alloys. *Biomaterials* **2002**, *23*, 3651–3659.
 11. Nimb, L.; Gotfredsen, K.; Steen-Jensen, J. Mechanical failure of hydroxyapatite-coated titanium and cobalt-chromium-molybdenum alloy implants. *Acta. Orthop. Belg.* **1993**, *59*, 333–338.
 12. Rahbek, O.; Overgaard, S.; Lind, M.; Bendix, K.; Bunker, C.; Soballe, K. Sealing effect of hydroxyapatite coating on peri-implant migration of particles. *J. Bone Joint Surg.* **2001**, *83*, 441–447.
 13. Coathup, M.J.; Blackburn, J.; Goodship, A.E.; Cunningham, J.L.; Smith, T.; Blunn, G.W. Role of hydroxyapatite coating in resisting wear particle migration and osteolysis around acetabular components. *Biomaterials* **2005**, *26*, 4161–4169.
 14. Bauer, T.W.; Taylor, S.K.; Jiang, M.; Medendorp, S.V. An indirect comparison of third-body wear in retrieved hydroxyapatite-coated, porous and cemented femoral components. *Clin. Orthop. Rel. Res.* **1994**, *298*, 11–18.
 15. Sul, Y.T.; Johansson, C.B.; Albrektsson, T. Oxidized titanium screws coated with calcium ions and their performance in rabbit bone. *Int. J. Oral Maxillofac. Implants* **2002**, *17*, 625–634.
 16. Nayab, S.N.; Jones, F.H.; Olsen, I. Effects of calcium ion implantation on human bone cell interaction with titanium. *Biomaterials* **2005**, *26*, 4717–4727.
 17. Kornu, R.; Maloney, W.J.; Kelly, M.A.; Smith, R.L. Osteoblast adhesion to orthopaedic implant alloys: effects of cell adhesion molecules and diamond-like carbon coating. *J. Orthop. Res.* **1996**, *14*, 871–877.
 18. Roehlecke, C.; Witt, M.; Kasper, M.; Schulze, E.; Wolf, C.; Hofer, A.; Funk, R.W. Synergistic effect of titanium alloy and collagen type I on cell adhesion, proliferation and differentiation of osteoblast-like cells. *Cells Tissues Organs* **2001**, *168*, 178–187.
 19. Goto, T.; Yoshinari, M.; Kobayashi, S.; Tanaka, T. The initial attachment and subsequent behaviour of osteoblasts and oral epithelial cells on titanium. *Biomed. Mater. Eng.* **2004**, *14*, 537–544.
 20. Gornowicz, G.; McCarthy, M.B. Response of human osteoblasts to implant materials: integrin-mediated adhesion. *J. Orthop. Res.* **1996**, *14*, 878–887.
 21. Zinger, O.; Anselme, K.; Denzer, A.; Habersetzer, P.; Wieland, M.; Jeanfils, J.; Hardouin, P.; Landolt, D. Time-dependent morphology and adhesion of osteoblastic cells on titanium model surfaces featuring scale-resolved topography. *Biomaterials* **2004**, *25*, 2695–2711.
 22. Urban, R.M.; Jacobs, J.J.; Tomlinson, M.J.; Gavriloic, J.; Black, J.; Peoch, M. Dissemination of wear particles to the liver, spleen and the abdominal lymph nodes of patients with hip or knee replacement. *J. Bone Joint Surg.* **2000**, *82-A*, 457–477.
 23. Mak, K.H.; Wong, T.K.; Poddar, N.C. Wear debris from total hip arthroplasty presenting as an intrapelvic mass. *J. Arthroplasty* **2001**, *16*, 674–676.
 24. Doherty, A.T.; Howell, R.T.; Ellis, L.A.; Bisbinas, I.; Learmonth, I.D.; Newson, R.; Case, C.P. Increased chromosome translocations and aneuploidy in peripheral blood lymphocytes of patients having revision arthroplasty of the hip. *J. Bone Joint Surg.* **2001**, *83-B*, 1075–1081.
 25. Keel, S.B.; Jaffe, K.A.; Petur Nielsen, G.; Rosenberg, A.E. Orthopaedic implant-related sarcoma: a study of twelve cases. *Modern Pathol.* **2001**, *14*, 969–977.
 26. Ellingsen, J.E.; Lyngstadaas, S.P. *Bio-Implant Interface*; CRC Press: Boca Raton, U.S.A., 2003.
 27. Korkusuz, P.; Korkusuz, F. Hard tissue-biomaterial interactions. In *Biomaterials in Orthopedics*; Yaszemski, M.J., Trantolo, D.J., Lewandrowski, K.U., Hasýrcý, V., Altobelli, D.E., Wise, D.L., Eds.; Marcel Dekker Inc: New York, U.S.A., 2004; 1–210.
 28. Pipino, F. The bone-prosthesis interaction. *J. Orthopaed. Traumatol.* **2000**, *1*, 3–9.
 29. Epstein, N.J.; Warme, B.A.; Spanogle, J.; Ma, T.; Bragg, B.; Smith, R.L.; Goodman, S.B. Interleukin-1 modulates periprosthetic tissue formation in an intramedullary model of particle-induced inflammation. *J. Orthop. Res.* **2005**, *23*, 501–510.
 30. Nakashima, Y.; Sun, D.H.; Trindade, M.C.D.; Chun, L.E.; Song, Y.; Goodman, S.B.; Schurman, D.J.; Maloney, W.J.; Smith, R.L. Induction of macrophage C-C chemokine expression by titanium alloy and bone cement particles. *J. Bone Joint Surg.* **1999**, *81-B*, 155–162.
 31. Sabokbar, A.; Itonaga, I.; Sun, S.G.; Kudo, O.; Athanasou, N.A. Arthroplasty membrane-derived fibroblasts directly induce osteoclast formation and osteolysis in aseptic loosening. *J. Orthop. Res.* **2005**, *23*, 511–519.
 32. Granchi, D.; Ciapetti, G.; Stea, S.; Savarino, L.; Filippini, F.; Sudanse, A.; Zinghi, G.; Montanaro, L. Cytokine release in mononuclear cells of patients with Co-Cr hip prosthesis. *Biomaterials* **1999**, *20*, 1079–1086.
 33. Stea, S.; Visentin, M.; Granchi, D.; Cenni, E.; Ciapetti, G.; Sudanese, A.; Toni, A. Apoptosis in peri-implant tissue. *Biomaterials* **2000**, *21*, 1393–1398.
 34. Fleming, J.E.; Cornell, C.N.; Muschler, G.F. Bone cells and matrices in orthopedic tissue engineering. *Orthop. Clin. N. Am.* **2000**, *31*, 357–374.
 35. Korkusuz, F.; Uchida, A.; Shinto, Y.; Araki, N.; Inoue, K.; Ono, K. Experimental implant-related osteomyelitis treated by antibiotic-calcium hydroxyapatite ceramic composites. *J. Bone Joint Surg.* **1993**, *75-B*, 111–114.
 36. Uchida, A.; Araki, N.; Shinto, Y.; Yoshikawa, H.; Ono, K.; Kurosaki, E. The use of hydroxyapatite ceramic in bone tumour surgery. *J. Bone Joint Surg.* **1990**, *72-b*, 298–302.
 37. Koç, N.; Timuçin, M.; Korkusuz, F. Fabrication and characterization of porous tricalcium phosphate ceramics. *Ceramics Int.* **2004**, *30*, 205–211.
 38. Roy, M.E.; Nishimoto, S.K. Matrix Gla protein binding to hydroxyapatite is dependent on the ionic environment: calcium enhances binding affinity but phosphate and magnesium decrease affinity. *Bone* **2002**, *31*, 296–302.

39. Rees, S.G.; Hughes-Wassell, D.T.; Waddington, R.J.; Embery, G. Interaction of bone proteoglycans and proteoglycan components with hydroxyapatite. *Biochim. Biophys. Acta* **2001**, *1568*, 118–128.
40. Goldberg, H.A.; Warner, K.J.; Li, M.C.; Hunter, G.K. Binding of bone sialoprotein, osteopontin and synthetic polypeptides to hydroxyapatite. *Connect. Tissue Res.* **2001**, *42*, 25–37.
41. Sendil-Keskin, D.; Tezcaner, A.; Korkusuz, P.; Korkusuz, F.; Hasirci, V. Collagen-chondroitin sulfate-based PLLA-SAIB-coated rhBMP-2 delivery system for bone repair. *Biomaterials* **2005**, *26*, 4023–4034.
42. Köse, G.T.; Korkusuz, F.; Korkusuz, P.; Hasırcı, V. In vivo tissue engineering of bone using Poly(3-hydroxybutyric acid-co-3-hydroxyvaleric acid) and collagen scaffolds. *Tissue Eng.* **2004**, *10*, 1234–1250.
43. Torun-Köse, G.; Korkusuz, F.; Özkul, A.; Soysal, Y.; Özdemir, T.; Yıldız, C.; Hasirci, V. Tissue engineered cartilage on collagen and PHBV matrices. *Biomaterials* **2005**, *26*, 5187–5197.
44. Korkusuz, F.; Korkusuz, P.; Ekşioğlu, F.; Gürsel, I.; Hasirci, V. In vivo response to biodegradable controlled antibiotic release systems. *J. Biomed. Mater. Res.* **2001**, *55*, 217–228.
45. Coskun, S.; Korkusuz, F.; Hasirci, V. Hydroxyapatite reinforced poly(3-hydroxybutyrate) and poly(3-hydroxybutyrate-co-3-hydroxyvalerate) based biodegradable composite bone plate. *J. Biomater. Sci. Polym. Ed.* **2005**, *16*, 1485–1502.
46. Idris, A.I.; van't Hof, R.J.; Greig, I.R.; Ridge, S.A.; Baker, D.; Ross, R.A.; Ralston, S.H. Regulation of bone mass, bone loss and osteoclast activity by cannabinoid receptors. *Nat. Med.* **2005**, *11*, 774–779.
47. Lynn, A.K.; Brooks, R.A.; Bonfield, W.; Rushton, N. Repair of defects in articular joints. Prospects for material-based solutions in tissue engineering. *J. Bone Joint Surg.* **2004**, *86-B*, 1093–1099.
48. Gao, J.; Dennis, J.E.; Solchaga, L.A.; Awadallah, A.S.; Goldberg, V.M.; Caplan, A.I. Tissue-engineered fabrication of an osteochondral composite graft using rat bone marrow-derived mesenchymal stem cells. *Tissue Eng.* **2001**, *7*, 363–371.
49. Agung, M.; Ochi, M.; Adachi, N.; Uchio, Y.; Takao, M.; Kawasaki, K. Osteochondritis dissecans of the talus treated by the transplantation of tissue-engineered cartilage. *Arthroscopy* **2004**, *10*, 1075–1080.
50. Alhadlaq, A.; Mao, J.J. Tissue-engineered osteochondral constructs in the shape of an articular condyle. *J. Bone Joint Surg.* **2005**, *87-A*, 936–944.



Hard Tissue Elastic Properties

J. Lawrence Katz

*Co-Director, Center for Research on Interfacial Structure and Properties,
University of Missouri–Kansas City, Kansas City, Missouri, U.S.A.*

INTRODUCTION

The principal protein of bone is type 1 collagen, which along with other proteins, glycoproteins, and protein-polysaccharides, comprises about 50% of bone's volume.^[1,2] The other major constituent of bone is a calcium phosphate in the form of a crystalline carbonateapatite.^[3] Bone is an anisotropic, heterogeneous, inhomogeneous, non-linear, thermorheologically complex viscoelastic material, making it one of the most interesting materials known in terms of structure–property relationships. In addition, bone exhibits electromechanical effects, presumed to be because of streaming potentials, both in vivo and in vitro when wet.^[4]

Currey^[5] states unequivocally that he thinks, “the most important feature of bone material is its stiffness.” This statement constitutes the rationale that it is only the elastic and viscoelastic properties of compact cortical bone as well as the elastic properties of trabecular bone that will be covered in this article. Although bone is such a complex material, we will take advantage of some simplification by modeling its behavior as a linear elastic material with additional indication, at the end of this article, of how to take into account its viscoelastic properties.

STRUCTURE/COMPOSITION OF BONE

To understand the structure–property relationships in bone, it is necessary to consider it as a complex hierarchical structural material.^[6] An artist's sketch of this hierarchical arrangement for compact cortical femoral bone is shown in Fig. 1. It is convenient to categorize the structure on four levels: molecular, ultrastructural (nanostructural), microstructural, and macrostructural (whole bone). Fig. 2 further illustrates the hierarchical subdivisions within mammalian bone. At the molecular structural level, we have the tropocollagen molecule and apatitic mineralites (Ap). The tropocollagen molecule is $\sim 1.5 \times 280$ nm and is made up of three individual left-handed helical polypeptide (alpha) chains coiled into a right-handed helix.^[2] The mineralites have been found to be carbonate-substituted hydroxyapatite, generally thought to be

non-stoichiometric.^[3] Atomic force microscopy (AFM) measurements have shown them in mature bovine bone to be about $0.61 \times 10 \times 12$ nm in size,^[9] while those isolated from young bovine bone are about $2.0 \times 6.0 \times 9.0$ nm in size.^[10] At the ultrastructural level, the collagen and Ap are closely associated and formed into a microfibrillar composite that is then assembled into fibers that range from approximately 3 to 5 μ m thick. At the microstructural level, these fibers are either randomly arranged (woven bone) or organized into concentric lamellar groups (secondary osteons or haversian systems) or linear lamellae (plexiform bone).^[5,11] This generally is the tissue level of structure that is used when making measurements of elastic moduli either by mechanical testing or ultrasonic wave propagation techniques. Bone, at this level, is found in two different architectural structures—dense and porous.^[5] Dense bone is found, for example, in the shafts of long bone and is usually termed “compact or cortical bone.” The porous, or spongy type of bone is found, for example, in the articulating ends of long bones, and is called cancellous bone. It is important to note that both cortical and cancellous bone comprise the same basic materials in essentially the same hierarchical arrangements. The macroscopic or whole bone comprises the tissue level components organized into both the cortical shell around the medullary canal and the cancellous bone in the articulating regions.^[5] The mechanical properties of the whole bone results from the contribution of each of these hierarchical levels.^[6]

The composition of bone can vary depending upon a large number of factors: species; type of bone, e.g., woven and cancellous; location; age; sex; etc. Some data for overall composition of both human and bovine cortical bone by weight is given in Table 1.

ELASTIC PROPERTIES

As indicated above in the Introduction, for the purposes of this article, it is reasonable, as a first approximation, to model bone as an anisotropic, linear elastic solid with Hooke's law as the appropriate constitutive equation. Although bone is a viscoelastic material, at the quasistatic strain rates used in

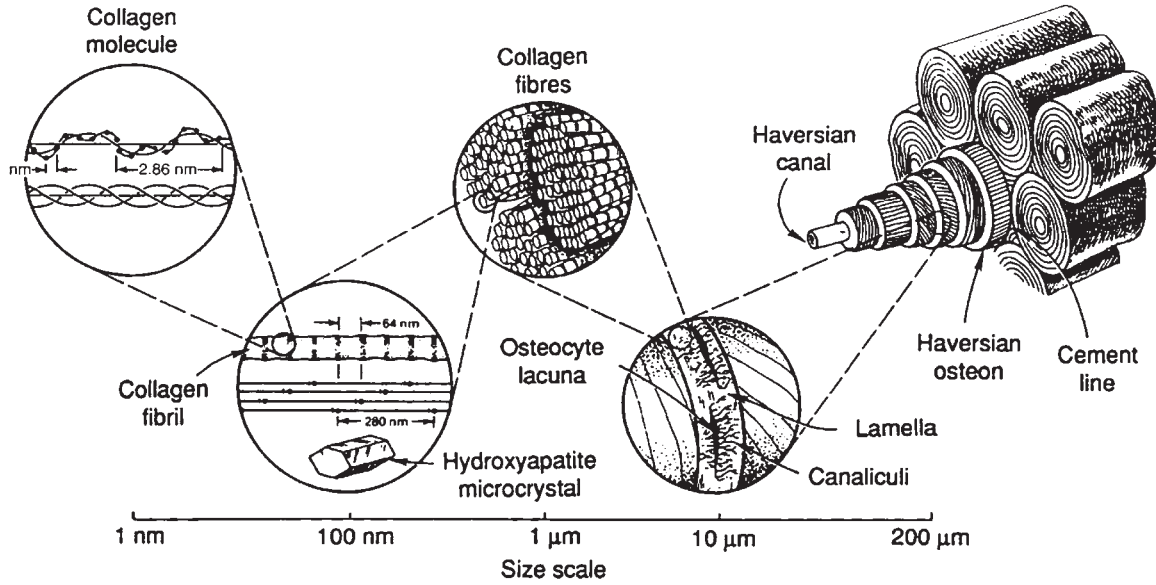


Fig. 1 Hierarchical levels of structure of human bone. (From Ref. [7], permission sought from CRC Press, Boca Raton, FL.)

mechanical testing and the near zero strain rate ultrasonic frequencies used experimentally, linear elasticity is a reasonable approximation. It must be kept in mind that this holds if the range of frequency (time) studied is relatively small. If, for example, one were to compare the results of the quasistatic mechanical tests with the ultrasonic tests, dispersion (variation of modulus with frequency) associated with viscoelasticity becomes evident, as seen in Table 2. Standard international units are used throughout this article. In tensor notation, Hooke's law is written as

$$\sigma_{ij} = C_{ijkl} \varepsilon_{kl} \tag{1}$$

where σ_{ij} and ε_{kl} are the second rank stress and infinitesimal second rank strain tensors, respectively, and C_{ijkl} is the fourth rank elasticity tensor. Using a reduced notation, where: 11 → 1, 22 → 2, 33 → 3, 23 → 4, 31 → 5, and 12 → 6 in Eq. (1), we can rewrite Eq. (1) as

$$\sigma_i = C_{ij} \varepsilon_j \quad i, j = 1 - 6 \tag{2}$$

where the C_{ij} are the stiffness coefficients (elastic constants), and the S_{ij} are known as the compliance coefficients.

Lang,^[18] Katz and Ukraincik,^[19] and Yoon and Katz,^[20,21] based on their ultrasonic wave propagation measurements, assumed cortical bone to be transversely isotropic with the bone axis of symmetry (the 3 direction) as the unique axis of symmetry; this requires five independent C_{ij} (or S_{ij}) to fully describe its anisotropic elastic properties. The matrix form of Eq. (2) for transverse isotropy is given in Eq. (3), while the corresponding matrix form for the compliances is given in Eq. (4).

$$\begin{Bmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \\ \sigma_{yz} \\ \sigma_{zx} \\ \sigma_{xy} \end{Bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & 0 \\ C_{12} & C_{11} & C_{13} & 0 & 0 & 0 \\ C_{13} & C_{13} & C_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & (C_{11} - C_{12})/2 \end{bmatrix} \times \begin{Bmatrix} \varepsilon_{xx} \\ \varepsilon_{yy} \\ \varepsilon_{zz} \\ \varepsilon_{yz} \\ \varepsilon_{zx} \\ \varepsilon_{xy} \end{Bmatrix} \tag{3}$$

$$\begin{Bmatrix} \varepsilon_{xx} \\ \varepsilon_{yy} \\ \varepsilon_{zz} \\ \varepsilon_{yz} \\ \varepsilon_{zx} \\ \varepsilon_{xy} \end{Bmatrix} = \begin{bmatrix} S_{11} & S_{12} & S_{13} & 0 & 0 & 0 \\ S_{12} & S_{11} & S_{13} & 0 & 0 & 0 \\ S_{13} & S_{13} & S_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & S_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & S_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & 2(S_{11} - S_{12}) \end{bmatrix} \times \begin{Bmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \\ \sigma_{yz} \\ \sigma_{zx} \\ \sigma_{xy} \end{Bmatrix} \tag{4}$$



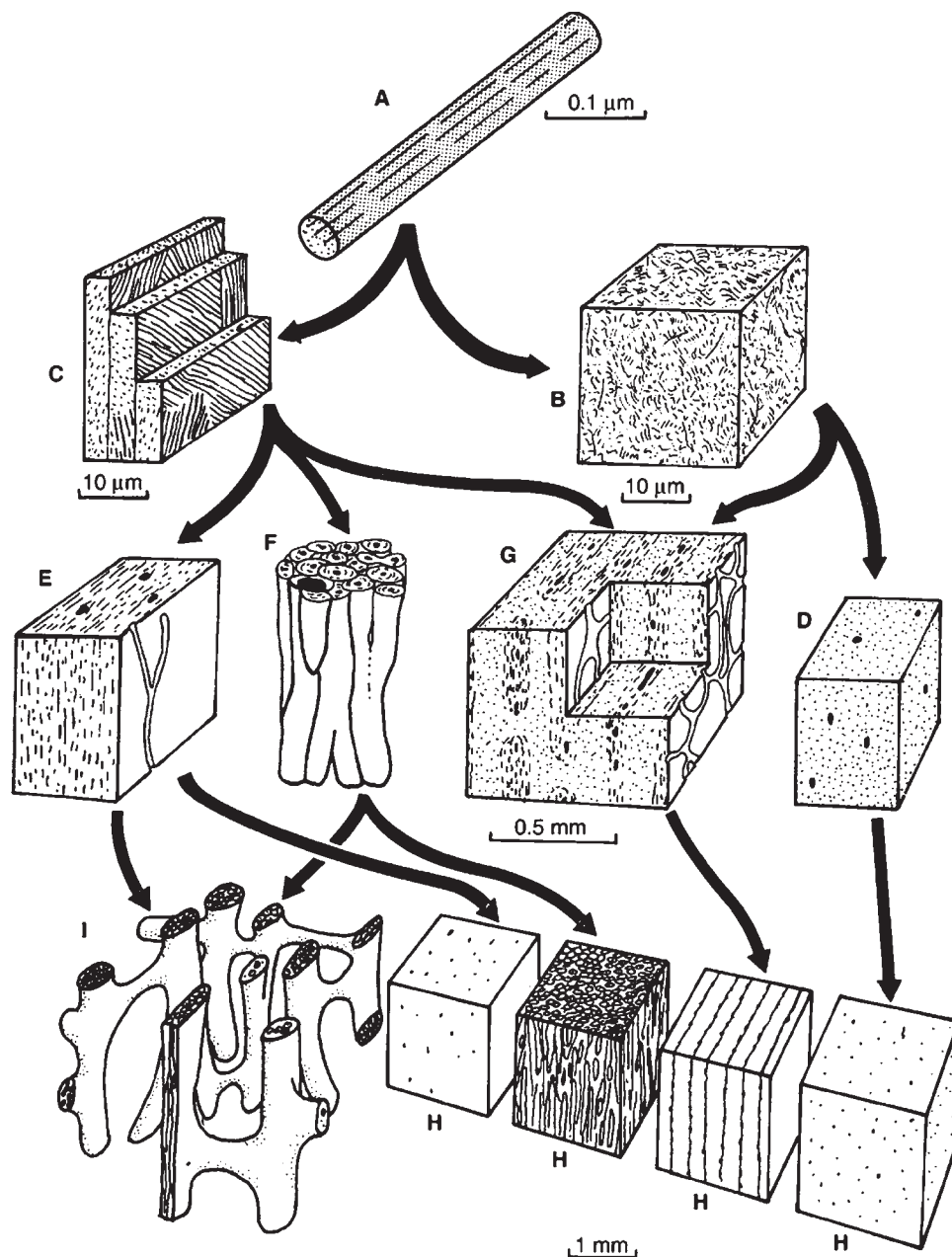


Fig. 2 Diagram showing mammalian bone at different levels. Diagram showing the structure of mammalian bone at different levels. Bone at the same level is drawn at the same magnification. The arrows show what types may contribute to structures at higher levels: (A) collagen fibril with associated mineral crystals; (B) woven bone, the collagen fibrils are arranged more or less randomly. Osteocytes are not shown; (C) lamellar bone, there are separate lamellae, and the collagen fibrils are arranged in "domains" of preferred fibrillar orientation in each lamella. Osteocytes are not shown; (D) woven bone, blood channels are shown as large black spots. At this level woven bone is indicated by light dotting; (E) primary lamellar bone, at this level lamellar bone is indicated by fine dashes; (F) haversian bone, a collection of Haversian systems, each with concentric lamellae round a central blood channel. The large black area represents the cavity formed as a cylinder of bone is eroded away. It will be filled in with concentric lamellae and will form a new Haversian system; (G) laminar bone, two blood channel networks are exposed. Note how layers of woven and lamellar bone alternate; (H) compact bone of the types shown at the lower levels; and (I) cancellous bone. (From Ref. [8], permission sought from Princeton University Press.)

On the other hand, Van Buskirk and Ashman^[16] assumed cortical bone to have orthotropic symmetry based on their ultrasonic measurements that showed small differences in the elastic properties between the

radial (1 direction) and transverse (2 direction) axes; this requires nine independent C_{ij} (or S_{ij}) to fully describe its anisotropic elastic properties. Yoon and Katz^[20,21] deemed that this was essentially because of

Table 1 Composition of adult human and bovine cortical bone

Species	%H ₂ O	Ap%	Dry weight collagen	GAG ^a	References
Bovine	9.1	76.4	21.5	ND ^b	[12]
			Human	7.3	67.2
21.2	0.34	[13,14]			

^aGlycosaminoglycan.^bNot determined.

the apparent gradient in porosity from the periosteal to the endosteal sides of bone, and therefore, that this defect did not alter the basic transverse isotropic symmetry, so that there was a saving in modeling to need only five instead of nine constants in any calculations. While plexiform (young) bovine bone is widely accepted to be orthotropic, in both human and mature bovine haversian bone, any deviations from transverse isotropy is small and is generally neglected in modeling calculations. The matrix for the transverse isotropic C_{ij} is as presented in Eq. (3).

The corresponding matrix can be written for the compliance coefficients, S_{ij} , based on Eq. (4) following

$$\varepsilon_i = S_{ij}\sigma_j \quad i, j = 1 - 6 \quad (5)$$

where the S_{ij} th compliance is obtained by dividing the $[C_{ij}]$ stiffness matrix, minus the i th row and j th column, by the full $[C_{ij}]$ matrix and vice versa to obtain the C_{ij} in terms of the S_{ij} . It is important to recognize that, although Young's modulus in the bone axis, 3, direction, $E_{33} = 1/S_{33}$, where $E_{33} \neq C_{33}$, because C_{33} and S_{33} are not reciprocals. In general, C_{ij} and S_{ij} are not simple reciprocals for either transverse isotropy or orthotropic symmetry, or even for the isotropic case.

Ultrasonic wave propagation techniques provide a non-destructive means for in vitro measurements of bone tissues' elastic anisotropy;^[16,18,20,21] they are measurements done near the limit of zero strain at a high strain rate. However, rate is not the best measure in

the linear domain because at a given frequency doubling the amplitude entails doubling the rate. In the linear domain, this has no effect in that damping and modulus are independent of amplitude. Also, the mechanical damping is at a minimum near 1 Hz; so in that region the bone is most closely approximated as linear. Thus, non-linearities present in the material are not likely to be easily assessed with this technique, which is one reason why Hooke's law works so well when analyzing ultrasonic measurements. As for Hooke's law in ultrasonic measurements, there is still attenuation; so Hooke's law is extended to complex moduli with ultrasonics as well as with lower frequency cyclic tests. The combination of Hooke's law with Newton's second law results in a wave equation that yields the following relationship involving the stiffness matrix

$$\rho V^2 U_m = C_{mrsn} N_r N_s U_n \quad (6)$$

where ρ is the density of the medium, V is the wave speed, and U and N are unit vectors along the particle displacement and wave propagation directions, respectively, so that U_m , N_r , etc. are direction cosines.

Thus to find the five transverse isotropic elastic constants, at least five independent measurements are required e.g., a dilatational longitudinal wave in the 3 and 1 or 2 directions, a transverse wave in the 13, 23, and 12 planes, etc. Technical moduli can then be calculated from the full set of C_{ij} . Redundant measurements should be made both for transverse isotropic and orthotropic symmetry to ensure good statistics.

There are several major advantages of the ultrasonic measurements technique over mechanical testing: 1) The former can be done with specimens too small for the latter technique; 2) ultrasonic reproducibility of measurements is greater than for mechanical testing; and 3) the full set of either five or nine coefficients can be measured on a smaller number of specimens using ultrasound, than is possible with mechanical testing. Thus, at present, ultrasonic techniques are preferred in studies of elastic anisotropy in both human and other mammalian bones.

Table 2 Elastic stiffness coefficients for various human and bovine bones measured with ultrasound except for Knets mechanical tests

References	Bone type	C_{11}	C_{22}	C_{33}	C_{44}	C_{55}	C_{66}	C_{12}	C_{13}	C_{23}
[15]	Human tibia	11.6	14.4	22.5	4.91	3.56	2.41	7.95	6.10	6.92
[16]	Human femur	20.0	21.7	30.0	6.56	5.85	4.74	10.9	11.5	11.5
[16]	Bovine femur	14.1	18.4	25.0	7.00	6.30	5.28	6.34	4.84	6.94
[17]	Bovine femur plexiform	22.4	25.0	35.0	8.20	7.10	6.10	14.0	15.8	13.6
[17]	Bovine femur haversian	21.2	21.0	29.0	6.30	6.30	5.40	11.7	12.7	11.1

All values in GPa.

(From Ref. [15].)



In addition to the bulk wave measurements of individual elastic constants as described above, it is possible to obtain Young's modulus directly. If the wavelength of the sound is larger than the specimen cross section, e.g., low-frequency ultrasound with samples of small cross sections. For this case, an extensional longitudinal (bar) wave of velocity V is propagated yielding Young's modulus, E , directly, Eq. (7)

$$V^2 = E/\rho \quad (7)$$

This is analogous to a uniaxial mechanical test experiment in which the slope of the stress/strain curve yields E .

The precision of the ultrasonic measurement technique provides values of the C_{ij} , confirming the anisotropic property difference between bovine haversian and plexiform (lammelar) bones. These differences in C_{ij} between the two types of bone tissue are hypothesized to be owing essentially to the differences in their microstructural organization (Fig. 3). At this level of structure, the textural symmetry has dimensions comparable to those of the ultrasound wavelengths used in the experiment. In addition, the molecular and ultrastructural levels of organization in both types of tissues are essentially identical. These results reinforce the importance that it is necessary to understand bone as a hierarchical ensemble to model its elastic properties as a material/structural composite system. The human femur data^[16] support this description of bone tissue. The nine individual C_{ij} measured for the human femur treated as an orthotropic material are such that their results are consistent with a near-transverse isotropic symmetry. However, the nine C_{ij} measured for bovine femoral bone clearly show the influence of the orthotropic microtextural symmetry of that tissue's plexiform structure.

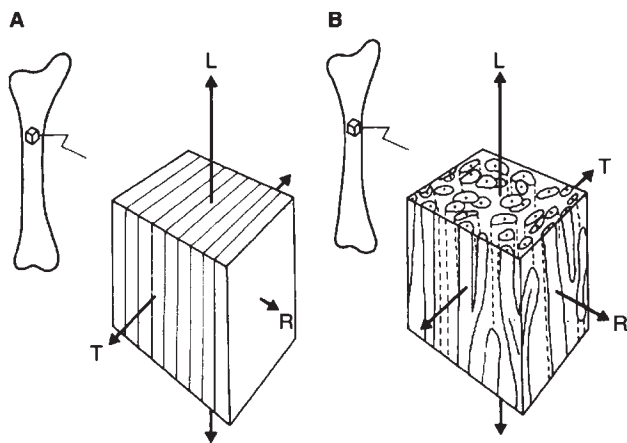


Fig. 3 Diagram showing how lamellar (plexiform) bone (A) differs more between radial and tangential directions (R and T) than does haversian bone and (B) the arrows are vectors representing the various directions. (From Ref. [8], permission sought from Princeton University Press.)

The C_{ij} are simply related to "technical" elastic moduli such as Young's modulus, E , shear modulus, G , bulk modulus, K , and others. Thus, it is possible to describe the moduli along any given direction. In place of the full equations for the most general anisotropy, which are too long to present here, given in the following are the simplified equations for the case of transverse isotropy, e.g., Ref. [22]. For this symmetry, Young's modulus is

$$1/E(\gamma_3) = S'_{33} = (1 - \gamma_3^2)2S_{11}/[\gamma_3^4 S_{33}/\gamma_3^2(1 - \gamma_3^2) \times (2S_{13} + S_{44})] \quad (8)$$

where $\gamma_3 = \cos \phi$ and ϕ is the angle made with respect to the 3 (bone) axis.

The shear modulus (rigidity modulus or torsional modulus for a circular cylinder) is

$$1/G(\gamma_3) = 1/2(S'_{44} + S'_{55}) \\ = S_{44} + (S_{11} - S_{22}) - 1/2S_{44}(1 - \gamma_3^2) \\ + 2(S_{11} + S_{33} - 2S_{13} - S_{44})\gamma_3^2(1 - \gamma_3) \quad (9)$$

where, again $\gamma_3 = \cos \phi$.

The bulk modulus (reciprocal of the volume compressibility), independent of ϕ , is

$$1/K = S_{33} + 2(S_{11} + S_{22} + 2S_{13}) \\ = [C_{11} + C_{12} + 2C_{33} - 4C_{13}]/ \\ [C_{33}(C_{11} + C_{12}) - 2C_{13}^2] \quad (10)$$

The S_{ij} in Eqs. (9) and (10) can be converted to the C_{ij} through the following set of transformation equations:

$$S_{11} = (C_{22}C_{33} - C_{23}^2)/\Delta; \\ S_{22} = (C_{33}C_{11} - C_{13}^2)/\Delta; \\ S_{33} = (C_{11}C_{22} - C_{12}^2)/\Delta; \\ S_{12} = (C_{13}C_{23} - C_{12}C_{33})/\Delta; \\ S_{13} = (C_{12}C_{23} - C_{13}C_{22})/\Delta; \\ S_{23} = (C_{12}C_{13} - C_{23}C_{11})/\Delta \\ S_{44} = 1/C_{44}; \quad S_{55} = 1/C_{55}; \quad S_{66} = 1/C_{66} \quad (11)$$

where the determinant, Δ , of the C_{ij} matrix is being taken

$$\Delta = \begin{vmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{vmatrix} \\ = C_{11}C_{22}C_{33} + 2C_{12}C_{23}C_{13} \\ - (C_{11}C_{23}^2 + C_{22}C_{13}^2 + C_{33}C_{12}^2) \quad (12)$$

The full equations for general symmetry can be found in Yoon and Katz.^[20] There also is available a considerable amount of data on the mechanical properties of cancellous (trabecular) bone, including measurements of the elastic properties of single trabeculae. Keaveny and Hayes^[23] presented an analysis of 20 years of studies on the mechanical properties of trabecular bone. These earlier studies generally used mechanical testing of bulk specimens of a size reflecting a cellular solid, i.e., of the order of cubic centimeter or larger. In these studies, both the modulus and strength of trabecular bone were strongly correlated to the apparent density, where apparent density, ρ_a , is defined as the product of individual trabeculae density, ρ_t , and the volume fraction of bone in the bulk specimen, V_t , and is given by $\rho_a = \rho_t V_t$.

A wide range of Young's moduli, E , resulted from these measurements, ranging from approximately 10 MPa to the order of 1 GPa depending on the apparent density. Keaveny and Hayes^[23] showed that these results could be correlated to the apparent density in g/cc by a power law relationship, $E = 6.13\rho_a^{1.44}$, with an $r^2 = 0.62$ calculated for 165 specimens.

As early as 1975, Townsend, Rose, and Radin^[24] machined individual trabeculae and measured their moduli in buckling, yielding a modulus of 11.4 GPa (wet) and 14.1 GPa (dry). New techniques of micromechanical measurements of bone allowed the determination of the elastic properties of the individual trabeculae in addition to the measurements of bulk properties. Nanoindentation and ultrasonic techniques such as scanning acoustic microscopy have been used for these studies, as well as by other mechanical testing methods. Ultrasound measurements^[25,26] have yielded values commensurate with the measurements of Townsend, Rose, and Radin (Table 3).^[24] Rho et al.^[29] using nanoindentation obtained average values of Young's modulus ranging from 15.0 to 19.4 GPa depending on orientation. These values were somewhat lower compared to 22.4 GPa for osteons and 25.7 GPa for the interstitial lamellae in cortical bone (Table 3). Turner et al.^[30] compared nanoindentation and acoustic microscopy at 50 MHz on the same specimens of trabecular and cortical bone from a common human donor (Table 3). The anisotropy ratio of longitudinal modulus to transverse modulus for cortical bone was similar for both modes of measurement even though the nanoindentation measurements resulted in Young's moduli greater than those measured by acoustic microscopy by 4–14%; the trabecular values are given in Table 3. Acoustic microscopy at 400 MHz has been used to measure the moduli of both human trabecular and cortical bone,^[25] yielding results comparable to those of Turner et al.^[30] for both types of bone (Table 3).

Table 3 Elastic moduli of trabeculae measured by different experimental methods

References	Method	Average modulus
[24]	Buckling	11.4 (Wet)
	Buckling	14.1 (Dry)
[27]	Uniaxial tension	0.760
[28]	4-point bending	5.72
[25]	Ultrasound	13.0 (Human)
	Ultrasound	10.9 (Bovine)
[26]	Ultrasound	14.8
	Tensile test	10.4
[29]	Nanoindentation	19.4 (Longitudinal)
	Nanoindentation	15.0 (Transverse)
[30]	Acoustic microscopy	17.5
	Nanoindentation	18.1
[31]	Acoustic microscopy	17.4

All moduli in GPa.

Such studies provide the data for micromechanical analyses using material properties measured on the microstructural level. They also indicate that nanoscale measurements such as those provided by AFM is the next step necessary to analyze the mechanics of bone on the smallest unit of structure shown in Fig. 1.

ELASTIC ANISOTROPY

A full set of either five or nine C_{ij} (S_{ij}) does describe the anisotropy of that particular specimen of cortical bone. However, there is no simple way of comparing the relative anisotropy between different measurements on one specimen or on different specimens of the same bone or between different species or even between experimenters' measurements by trying to relate the individual C_{ij} (S_{ij}) between sets of measurements. Katz and Meunier^[32,33] adapted a procedure from crystal physics that provided a method for obtaining two scalar mechanical quantities—the compressive and shear anisotropy for crystals with cubic or hexagonal symmetry—to calculate those quantities for transverse isotropic bone tissue. They then extended the calculation to bone tissues exhibiting orthotropic symmetry.^[34] For both cases, the percentage compressive (Ac^*) and shear (As^*) elastic anisotropy are given, respectively, by

$$Ac^*(\%) = 100(K^V - K_R)/(K^V + K_R)$$

$$As^*(\%) = 100(G^V - G_R)/(G^V + G_R) \quad (13)$$



where K^V and K_R are the Voigt^[35] (uniform strain across an interface) and Reuss^[36] (uniform stress across an interface) bulk moduli, respectively, and G^V and G_R are the Voigt and Reuss shear moduli, respectively. Voigt moduli represent the upper bound values while Reuss moduli represent the lower bound values. The equations for K^V , K_R , G^V , and G_R are provided for both transverse isotropy and orthotropic symmetry in the Appendix.

Haversian bones, whether human or bovine, have both their compressive and shear scalar anisotropy factors considerably lower than the respective values for plexiform bone. Plexiform bone is stiffer than haversian bone; it is also more anisotropic.

MODELING ELASTIC BEHAVIOR

The first attempt to model bone as a composite material composed of a simple linear superposition of collagen and apatite was made by Currey.^[37] He followed this later by taking into account the orientation of the apatite crystallites^[38] using a model proposed by Cox^[39] for fiber-reinforced composites. Katz^[40] and Piekarski^[41] independently using Voigt and Reuss or even Hashin-Shtrikman^[42] composite modeling showed that using such linear combinations of either elastic moduli or elastic compliances failed to provide a realistic description of the behavior of cortical bone as a composite material. The reason for these failures is that they were based on considerations of material properties alone. This is comparable to trying to determine the properties of the Eiffel Tower by simply modeling the composite material properties without considering void spaces and the interconnectivity of the structure.^[43] It is clear for both bone and the Eiffel Tower that, in addition to the material properties, consideration of their hierarchical structural organization must be introduced into the modeling; i.e., these systems are complex structural/material composites.

Based on the realization that the simple linear rule of mixtures could not explain the behavior of bone as a composite material,^[40] Katz^[44] hypothesized that haversian bone should be modeled as a hierarchical/structural material composite. Thus, he adapted a hollow fiber composite model of Hashin and Rosen^[45] for compact cortical bone that he published in a number of papers.^[6,46] Bonfield and Grynpas^[47] reported, in *Nature*, their study of the elastic properties of cortical bone specimens cut at angles of 0°, 5°, 10°, 20°, 30°, 40°, 50°, 70°, 80°, 85°, and 90° with respect to the long bone axis, using extensional (longitudinal) ultrasonic wave propagation in both wet and dry bovine femoral cortical bone specimens. They compared their experimental results for Young's moduli with the

theoretical curve predicted by Currey's^[38] 1969 model; the comparison is shown in Fig. 4. The bump in the ultrasonic measurements between 60° and 80° may be because of the mixing in of stiffer plexiform specimens with the other specimens. The clear lack of agreement between theory and experiment led them to "conclude, therefore that an alternative model is required to account for the dependence of Young's modulus on orientation."^[47] Katz and coworkers^[48–51] applying their hierarchical material-structure composite model were able to explain the data in Fig. 4 by considering different amounts of apatitic crystallites aligned parallel to the long bone axis (Fig. 5). More sophisticated modeling has progressed beyond this initial attempt at hierarchical structural/material micromechanical modeling, e.g., finite element micromechanical computations^[52] or homogenization theory.^[53] Further improvements will require the inclusion of more definitive information on the structural organization of collagen and apatite mineralites at the molecular-ultrastructural level. In addition to obtaining more definitive structural information, it will also be necessary to obtain constitutive mechanical properties at each level of the hierarchy before a complete mechanical model of bone is arrived at. That these concerns have generated a considerable range of research activities during the past several decades can be seen in

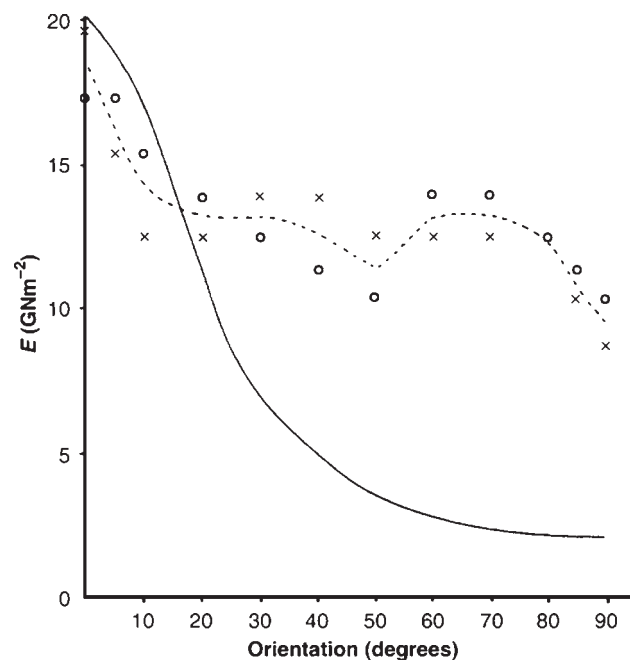


Fig. 4 Variation in Young's modulus of bovine femur specimens, E , with the orientation of specimen axis to the long axis of the bone for wet (o) and dry (x) conditions compared with the theoretical curve (—) predicted from a fiber-reinforced composite model. (From Ref. [47], permission sought from Nature.)

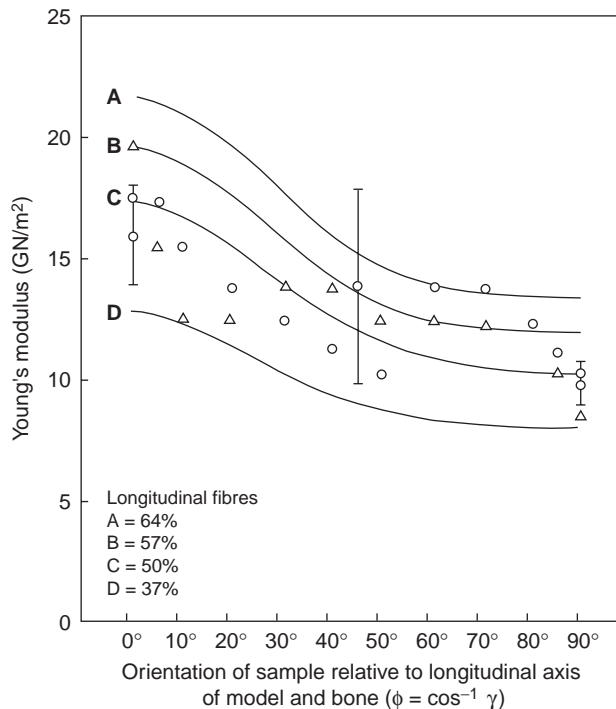


Fig. 5 Predictions of Katz's hierarchical structural/material composite model for cortical bone compared with the experimental data of Bonfield and Grympas. Each curve represents a different lamellar configuration within a single osteon, with longitudinal fibers: A, 64%; B, 57%; C, 50%; D, 37%; and the rest of the fibers assumed to be horizontal. (From Refs. [47–51], permission sought from Nature.)

just a few of the examples cited here.^[9,10,54–66] Considerable research in these areas continues at a rapid pace.

VISCOELASTIC PROPERTIES

Although, the emphasis in this article has been on bone as a linear elastic material, it is clear that bone, along with all other biologic tissues, is a viscoelastic material. Thus, Hooke's law for linear elastic materials must be replaced by a constitutive equation that includes the time dependency of the material properties. The behavior of an anisotropic linear viscoelastic material may be described by using the Boltzmann superposition integral as the constitutive equation^[67]

$$\sigma_{ij}(t) = \int_{-\infty}^t C_{ijkl}(t-\tau) \{d\varepsilon_{kl}(\tau)/d\tau\} d\tau \quad (14)$$

where $\sigma_{ij}(t)$ and $\varepsilon_{kl}(\tau)$ are the time-dependent second rank stress and strain tensors, respectively, and $C_{ijkl}(t-\tau)$ is the fourth rank relaxation modulus tensor. This tensor has 36 independent elements for the lowest

symmetry case and 12 non-zero independent elements for an orthotropic solid. As for Hooke's Law in linear elasticity, a reduced notation is used here as well, i.e., 11 → 1, 22 → 2, 33 → 3, 23 → 4, 31 → 5, and 12 → 6. Applying Eq. (14) to the case of an orthotropic material, e.g., plexiform bone, in uniaxial tension (compression) in the 1 direction,^[67] the following set of equations is obtained

$$\begin{aligned} \sigma_1(t) = \int_{-\infty}^t & [C_{11}(t-\tau) \{d\varepsilon_1(\tau)/d\tau\} \\ & + C_{12}(t-\tau) \{d\varepsilon_2(\tau)/d\tau\} \\ & + C_{13}(t-\tau) \{d\varepsilon_3(\tau)/d\tau\}] d\tau \end{aligned} \quad (15)$$

$$\begin{aligned} \sigma_2(t) = \int_{-\infty}^t & [C_{21}(t-\tau) \{d\varepsilon_1(\tau)/d\tau\} \\ & + C_{22}(t-\tau) \{d\varepsilon_2(\tau)/d\tau\} \\ & + C_{23}(t-\tau) \{d\varepsilon_3(\tau)/d\tau\}] d\tau = 0 \end{aligned} \quad (16)$$

$$\begin{aligned} \sigma_3(t) = \int_{-\infty}^t & [C_{31}(t-\tau) \{d\varepsilon_1(\tau)/d\tau\} \\ & + C_{32}(t-\tau) \{d\varepsilon_2(\tau)/d\tau\} \\ & + C_{33}(t-\tau) \{d\varepsilon_3(\tau)/d\tau\}] d\tau = 0 \end{aligned} \quad (17)$$

for all t .

Setting the integrands in both Eqs. (16) and (17) to 0 and solving them simultaneously for $\{d\varepsilon_2(\tau)/d\tau\}$ and $\{d\varepsilon_3(\tau)/d\tau\}$ and substituting these values in Eq. (15) yields

$$\sigma_1(t) = \int_{-\infty}^t E_1(t-\tau) \{d\varepsilon_1(\tau)/d\tau\} d\tau \quad (18)$$

where, for convenience, adopting the notation $C_{ij} \equiv C_{ij}(t-\tau)$, Young's modulus is given by

$$\begin{aligned} E_1(t-\tau) = & C_{11} + C_{12} \{ [C_{31} - (C_{21}C_{33}/C_{23})] / \\ & \{ (C_{21}C_{33}/C_{23}) - C_{32} \} \} \\ & + C_{13} \{ [C_{21} - (C_{31}C_{22}/C_{32})] / \\ & \{ (C_{31}C_{22}/C_{23}) - C_{23} \} \} \end{aligned} \quad (19)$$

For the case of uniaxial tension (compression), only nine independent orthotropic tensor components are involved, the three shear components being equal to zero. If dealing with compliances rather than with moduli, a compliance tensor element can be measured in a tensile creep experiment. Still, this time-dependent Young's modulus is still a rather complex function to calculate. Using the inverse form of the Boltzmann integral would constitute the compliance formulation. It is easier to measure the C_{ij} using ultrasonic wave propagation; it is easier to measure Young's modulus, E , in quasistatic mechanical testing. The modulus and

compliance formulas are equally valid; one does conversions between them if needed.

For bone being driven by a strain at a frequency ω , with a corresponding sinusoidal stress lagging by an angle δ , the complex Young's modulus $E^*(\omega)$ may be expressed as

$$E^*(\omega) = E'(\omega) + iE''(\omega) \quad (20)$$

where $E'(\omega)$, the stress-strain ratio in phase with the strain is known as the storage modulus, and $E''(\omega)$, the stress-strain ratio 90° out of phase with the strain is known as the loss modulus. $\tan \delta$ is equal to the ratio of the loss modulus to the storage modulus. Usually, data are presented by a graph of the storage modulus along with a graph of $\tan \delta$, both against frequency. A clear indication of the frequency (time)-dependent nature of bone is shown in Fig. 6 where values for $\tan \delta$ are plotted over 13 decades of frequency. For a more complete development of viscoelasticity, including some of the biological applications, see Lakes.^[7]

To better understand the response of bone to applied loads, it is important to know its rheologic properties. The viscoelastic properties of various long bones have been studied by a number of investigators, e.g., Sedlin,^[68] Smith and Keiper,^[69] Laird and Kingsbury,^[70] Lugassy and Korostoff,^[71] and Black and Korostoff.^[72] Unfortunately, none of these early measurements were performed over a sufficient range of frequency (or time) to completely define the viscoelastic properties measured. Thus, it is not possible to mathematically transform one property such as creep into another such as stress relaxation, or vice versa, to compare the results of different viscoelastic experiments on different bones.^[67] In the interest of completeness, it would also be appropriate to make

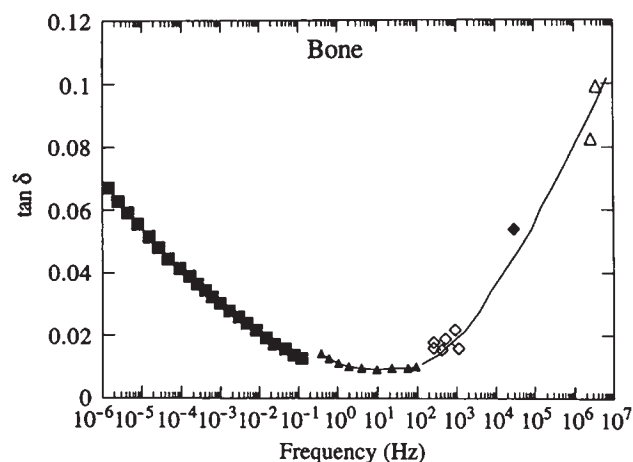


Fig. 6 $\tan \delta$ vs. frequency over 13 decades. (From Ref. [7], permission sought from CRC press, Boca Raton, FL.)

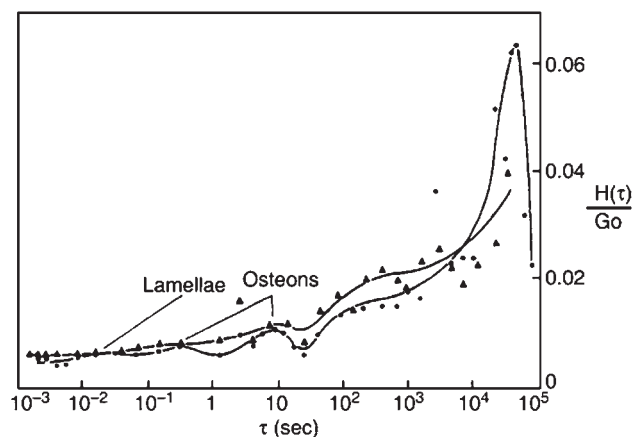


Fig. 7 Comparison of relaxation spectra for wet human bone, specimens 5 and 6, in simple torsion at $T = 37^\circ\text{C}$. First approximation from relaxation and dynamic data. (\bullet) Human tibial bone, specimen 6 and (\blacktriangle) human tibial bone, specimen 5. (From Ref. [77], permission sought from Elsevier.)

some mention of the literature on bulk fluid flow in bone because this is a rapidly growing area of the literature on bone mechanics, see for example Ref. [73].

Lakes and coworkers^[74,75] performed the first viscoelastic measurements of both human and bovine wet bone over an extended frequency range. He measured both the uniaxial and biaxial properties, using both dynamic and stress relaxation techniques, over eight decades of frequency (time).^[74,75] Although, in the normal range of function, bone could be treated as a linear viscoelastic material, the results of these experiments showed that over the eight decades bone, as a material, was both non-linear and thermorheologically complex, i.e., time-temperature superposition could not be used to extend the range of the viscoelastic measurements. A non-linear constitutive equation was developed based on these measurements.^[76] It should be noted that bone has a weak non-linearity at physiological strains, in contrast to tendon and ligament, which are highly non-linear at physiological strains.

Lakes also obtained relaxation spectrums for both human and bovine cortical bone (Fig. 7). Fig. 8 shows the contributions of several mechanisms to the loss tangent of cortical bone; it is interesting that almost all the major loss mechanisms occur at frequencies (times) at or near to those in which there are "bumps" in the relaxation spectrums, indicating possible strain energy dissipation.

Gottesman and Hashin^[78] adapted Katz and coworkers' adaptation^[48-51] of the Hashin-Rosen hollow fiber composite model to perform a viscoelastic calculation using the same major assumptions.

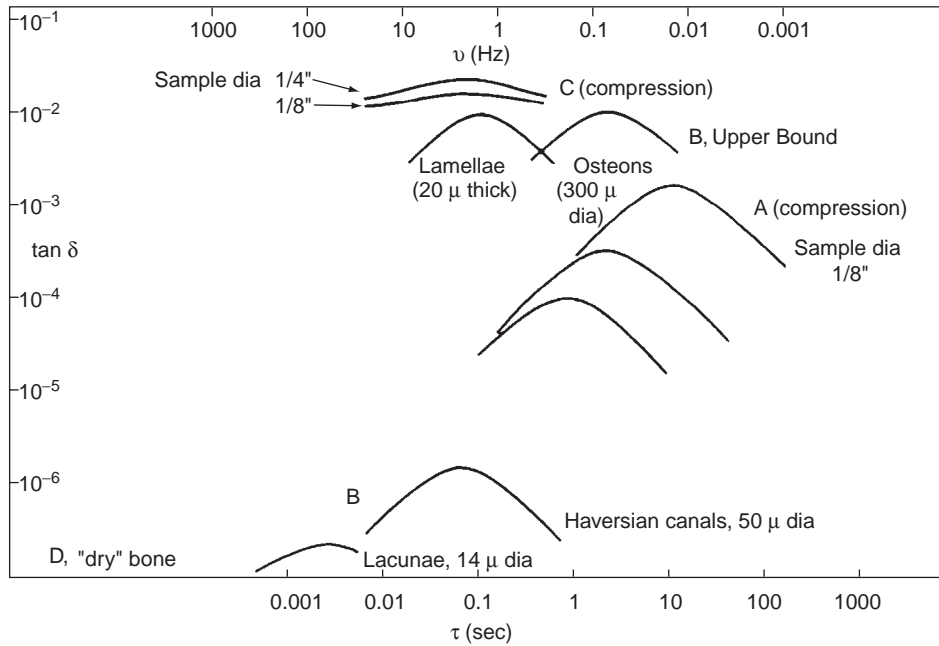


Fig. 8 Contributions of several relaxation mechanisms to the loss tangent of cortical bone. (A) Homogeneous thermoelastic effect; (B) inhomogeneous thermoelastic effect; (C) fluid flow effect; and (D): Piezoelectric effect. (From Ref. [77], permission sought from Elsevier.)



CONCLUSIONS

Although this article is titled, “Hard Tissue Elastic Properties,” it essentially has dealt only with cortical and cancellous (trabecular) bone. There is significant interesting research on other hard tissues, namely dentin and enamel, in the dental area where there are both similarities and considerable differences in structure and properties depending on which level of scale is being examined within these hierarchical structures as well. Indeed, as an example of similarity, the same calculations of scalar shear and compressive anisotropy factors performed for bone as a transverse isotropic material^[33] has been applied to wet and dry dentin^[79] based on resonant ultrasound spectroscopy measurements.^[80] While wet dentin exhibits transverse isotropic symmetry, the calculation shows that dry dentin is isotropic. This may be of clinical significance as there are patients with decreased saliva output, or “dry mouth” syndrome, in which there is an associated drying of tooth structure leading to serious problems. Interestingly, there is a prior small angle x-ray diffraction of human dentin experiment,^[81] which provides independent experimental evidence to support this calculation of isotropy based on the resonant ultrasonic data. Results of this experiment led to the conclusion that the collagen near the pulp appeared to have isotropic symmetry with mild anisotropy in the mid-dentin region.^[81] As the Aps’ orientations are in register with the collagen fiber orientations, this implication supports the above calculation of isotropy for the dry dentin.

Also, this article has not considered other key mechanical properties of bone, such as fracture and fatigue. They have considerable importance both in

basic understanding of bone behavior as well as in clinical practice. A complete understanding of bone formation, resorption, and remodeling clearly requires a comprehensive knowledge of both stiffness and strength and their relationship to fluid flow^[82] and cellular processes.^[83] These are complex problems and are at the cutting edge of present-day hard tissue research.

APPENDIX

The Voigt and Reuss moduli for both transverse isotropic and orthotropic symmetry are given below:

Voigt Transverse Isotropic

$$K^V = [2(C_{11} + C_{12}) + 4(C_{13} + C_{33})]/9 \tag{A1}$$

$$G^V = [C_{11} + C_{12} - 4C_{13} + 2C_{33} + 12(C_{44} + C_{66})]/30 \tag{A2}$$

Reuss Transverse Isotropic

$$K_R = [C_{33}(C_{11} + C_{22}) - 2C_{13}^2]/[C_{11} + C_{22} - 4C_{13} + 2C_{33}] \tag{A3}$$

$$G_R = 5\{C_{33}(C_{11} + C_{12}) - 2C_{13}^2\}C_{44}C_{66}/2\{[C_{33}(C_{11} + C_{12}) - 2C_{13}^2](C_{44} + C_{66}) + [C_{44}C_{66}(2C_{11} + C_{12}) + 4C_{13} + C_{33}]/3\} \tag{A4}$$

Voigt Orthotropic

$$K^V = \{C_{11} + C_{22} + C_{33} + 2(C_{12} + C_{13} + C_{23})\}/9 \quad (\text{A5})$$

$$G^V = \{C_{11} + C_{22} + C_{33} + 3(C_{44} + C_{55} + C_{66}) - (C_{12} + C_{13} + C_{23})\}/15 \quad (\text{A6})$$

Reuss Orthotropic

$$K_R = \{\Delta/(C_{11}C_{22} + C_{22}C_{33} + C_{33}C_{11}) - 2(C_{11}C_{23} + C_{22}C_{13} + C_{33}C_{12}) + 2(C_{12}C_{23} + C_{23}C_{13} + C_{13}C_{12}) - (C_{12}^2 + C_{13}^2 + C_{23}^2)\} \quad (\text{A7})$$

$$G_R = 15/\{4[(C_{11}C_{22} + C_{22}C_{33} + C_{33}C_{11} + C_{11}C_{23} + C_{22}C_{13} + C_{33}C_{12}) - (C_{12}^2 + C_{12}C_{23} + C_{23}^2 + C_{23}C_{13} + C_{13}^2 + C_{13}C_{12})]/\Delta + 3(1/C_{44} + 1/C_{55} + 1/C_{66})\} \quad (\text{A8})$$

where Δ is given in Eq. (13).

ARTICLES OF FURTHER INTEREST

Tissue Engineering of Bone; Zirconia Ceramics

REFERENCES

- Glimcher, M.J. Composition, structure, and organization of bone and other mineralized tissues and the mechanism of calcification. *Handbook of Physiology—Endocrinology VII*; Williams and Wilkins Co.: Baltimore, MD, 1976; 25–116.
- Voet, D.; Voet, J.G. *Biochemistry*, 2nd Ed.; John Wiley & Sons, Inc.: New York, 1995.
- Le Geros, R.Z. *Calcium Phosphates in Oral Biology and Medicine*; Karger: Basel, Switzerland, 1991.
- Johnson, M.; Katz, J.L. Electromechanical effects in bone. In *Handbook of Bioengineering*; Skalak, R., Chen, S., Eds.; McGraw-Hill Book Co.: New York, 1987; 3.1–3.11.
- Currey, J.D. *Bones: Structure and Mechanics*; Princeton University Press: Princeton, NJ, 2002.
- Katz, J.L. Hierarchical modeling of compact haversian bone as a fiber reinforced material. In *1976 Advances in Bioengineering*; Mates, R.E., Smith, C.R., Cowin, S.C., Eds.; A.S.M.E.: New York, 1976; 17–18.
- Lakes, R.S. *Viscoelastic Solids*; CRC Press: Boca Raton, FL, 1999.
- Wainwright, S.A.; Briggs, W.D.; Currey, J.D.; et al. *Mechanical Design in Organisms*; Princeton University Press: Princeton, NJ, 1982.
- Eppell, S.J.; Tong, W.; Katz, J.L.; et al. Shape and size of isolated bone mineralites using atomic force microscopy. *J. Orthop. Res.* **2001**, *19*, 1027–1034.
- Tong, W.; Glimcher, M.J.; Katz, J.L.; et al. Size and shape of mineralites in young bovine bone measured by atomic force microscopy. *Calc. Tissue Int.* **2003**, *72* (5), 592–598.
- Frasca, P.; Harper, R.A.; Katz, J.L. A new method of bone preparation for collagen fiber orientation studies by means of scanning electron microscopy. *J. Dent. Res.* **1976**, *55*, 372–375.
- Herring, G.M. Methods for the study of the glycoproteins and proteoglycans of bone using bacterial collagenase. Determination of bone sialoprotein and chondroitin sulphate. *Calcif. Tissue Res.* **1977**, *24*, 29–36.
- Pellegrino, E.D.; Blitz, R.M. The composition of human bone in uremia. *Medicine* **1965**, *44*, 397–418.
- Vejlens, L. Glycosaminoglycans of human bone tissue: I. Pattern of compact bone in relation to age. *Calcif. Tissue Res.* **1971**, *7*, 175–190.
- Knets, I.V. Mechanics of biological tissues, a review. *Polymer Mech. (Translation of Mekhanika Polimerov)* **1978**, *13*, 434–440.
- Van Buskirk, W.C.; Ashman, R.B. The elastic moduli of bone. In *Mechanical Properties of Bone*; Cowin, S.C., Ed.; A.S.M.E.: New York City, 1981; AMD Vol. 41, 131–143.
- Katz, J.L.; Lipson, S.; Yoon, H.S.; et al. The effects of remodeling on the elastic properties of bone. *Calc. Tissue Int.* **1984**, *36*, S31–S36.
- Lang, S.B. Elastic coefficients of animal bone. *Science* **1969**, *165*, 287–288.
- Katz, J.L.; Ukraincik, K. On the anisotropic elastic properties of hydroxyapatite. *J. Biomech.* **1971**, *4*, 221–227.
- Yoon, H.S.; Katz, J.L. Ultrasonic wave propagation in human cortical bone. I. Theoretical considerations for hexagonal symmetry. *J. Biomech.* **1976**, *9*, 407–412.
- Yoon, H.S.; Katz, J.L. Ultrasonic wave propagation in human cortical bone. II. Measurements of elastic properties and microhardness. *J. Biomech.* **1976**, *9*, 459–464.
- Lekhnitskii, S.G. *Theory of Elasticity of an Anisotropic Body*; Mir Publishers: Moscow, 1981.
- Keaveny, T.M.; Hayes, W.C. A 20-year perspective on the mechanical properties of trabecular bone. *J. Biomech. Eng.* **1993**, *115*, 535–542.
- Townsend, P.R.; Rose, R.M.; Radin, E.L. Buckling studies of single human trabeculae. *J. Biomech.* **1975**, *8*, 199–201.
- Ashman, R.B.; Rho, J.Y. Elastic modulus of trabecular bone material. *J. Biomech.* **1988**, *21*, 177–181.
- Rho, J.Y.; Ashman, R.B.; Turner, C.H. Young's modulus of trabecular and cortical bone material;

- ultrasonic and microtensile measurements. *J. Biomech.* **1993**, *26*, 111–119.
27. Ryan, S.D.; Williams, J.L. Tensile testing of rodlike trabeculae excised from bovine femoral bone. *J. Biomech.* **1989**, *22*, 351–355.
 28. Choi, K.; Goldstein, S.A. A comparison of the fatigue behavior of human trabecular and cortical bone tissue. *J. Biomech.* **1992**, *25*, 1371–1381.
 29. Rho, J.Y.; Roy, M.E.; Tsui, T.Y.; et al. Elastic properties of microstructural components of human bone tissue as measured by indentation. *J. Biomed. Mat. Res.* **1999**, *45*, 48–54.
 30. Turner, C.H.; Rho, J.Y.; Takano, Y.; et al. The elastic properties of trabecular and cortical bone tissues are similar results from two microscopic measurement techniques. *J. Biomech.* **1999**, *32*, 437–441.
 31. Burrell, S.; Katz, J.L. Scanning acoustic microscopy study of human cortical and trabecular bone. *Ann. Biomed. Eng.* **2001**, *29*, 1–9.
 32. Chung, D.H.; Buessem, W.R. The elastic anisotropy of crystals. In *Anisotropy in Single Crystal Refractory Compounds*; Vahldiek, F.W., Mersol, S.A., Eds.; Plenum Press: New York, 1968; Vol. 2, 217–245.
 33. Katz, J.L.; Meunier, A. The elastic anisotropy of bone. *J. Biomech.* **1987**, *20*, 1063–1070.
 34. Katz, J.L.; Meunier, A. A generalized method for characterizing elastic anisotropy in solid living tissues. *J. Mat. Sci. Mat. Med.* **1990**, *1*, 1–8.
 35. Voigt, W. *Lehrbuch der Kristallphysik*; Teubner: Berlin, Germany, 1910.
 36. Reuss, A. Berechnung der Fließgrenze von mischkristallen auf grund der plastizitätsbedingung für einkristalle, z. *Zeitschrift für Angewandte Mathematik und Mechanik* **1929**, *9*, 49–58.
 37. Currey, J.D. Three analogies to explain the mechanical properties of bone. *Biorheol.* **1964**, *2*, 1–10.
 38. Currey, J.D. The relationship between the stiffness and the mineral content of bone. *J. Biomech.* **1969**, *2*, 477–480.
 39. Cox, H.L. The elasticity and strength of paper and other fibrous materials. *Br. Appl. Phys.* **1952**, *3*, 72–79.
 40. Katz, J.L. Hard tissue as a composite material: part I. Bounds on the elastic behavior. *J. Biomech.* **1971**, *4*, 455–473.
 41. Piekarski, K. Analysis of bone as a composite material. *Int. J. Eng. Sci.* **1973**, *1*, 557–565.
 42. Hashin, Z.; Shtrikman, S.J. A variational approach to the theory of elastic behavior of multiphase materials. *J. Mech. Phys. Sol.* **1963**, *11*, 127–140.
 43. Lakes, R.S. Materials with structural hierarchy. *Nature* **1993**, *361*, 511–515.
 44. Katz, J.L. Elastic properties of calcified tissues. *Isr. J. Med. Sci.* **1971**, *7*, 439–441.
 45. Hashin, Z.; Rosen, B.W. The elastic moduli of fiber reinforced materials. *J. Appl. Mech.* **1964**, *31*, 223–232.
 46. Katz, J.L. The structure and biomechanics of bone. In *Mechanical Properties of Biological Materials*; Currey, J.D., Vincent, J.F.V., Eds.; Cambridge University Press: Cambridge, U.K., 1980; 137–168.
 47. Bonfield, W.; Grynias, M.D. Anisotropy of Young's modulus of bone. *Nature* **1977**, *270*, 453–454.
 48. Katz, J.L. On the anisotropy of Young's modulus of bone. *Nature* **1980**, *283*, 106–107.
 49. Katz, J.L. Composite material models for bone. In *Mechanical Properties of Bone*; A.S.M.E.: New York, 1981; AMD-Vol. 45, 171–184.
 50. Katz, J.L.; Maharidge, R.L.; Yoon, H.S. The estimation of inter-osteonal mechanical properties from a composite model for haversian bone. In *Biomechanics Current Interdisciplinary Research*; Perren, S.M., Scheider, E., Eds.; Martinus Injhoff: Dordrecht, 1985; 179–184.
 51. Katz, J.L.; Yoon, H.S.; Maharidge, R.L. Calculation of inter-osteonal mechanical properties for haversian bone based on a hierarchical composite model. *Biomechanics Symposium*; A.S.M.E.: New York, 1985; AMD-Vol. 68, FED-Vol. 21, 33–35.
 52. Hogan, H.A. Micromechanics modeling of haversian cortical bone properties. *J. Biomech.* **1992**, *25* (5), 549–556.
 53. Aoubiza, B.; Crolet, J.M.; Meunier, A. On the mechanical characterization of compact bone structure using the homogenization theory. *J. Biomech.* **1996**, *29*, 1539–1547.
 54. Swedlow, D.B.; Frasca, P.; Harper, R.A.; et al. Scanning and transmission electron microscopy of calcified tissues. *Biomater. Med. Dev. Art. Org.* **1975**, *3*, 121–153.
 55. Frasca, P.; Harper, R.A.; Katz, J.L. Isolation of single osteons and osteon lamellae. *Acta Anat.* **1976**, *95*, 122–129.
 56. Frasca, P.; Harper, R.A.; Katz, J.L. Collagen fiber orientation in human secondary osteons. *Acta Anat.* **1977**, *98*, 1–13.
 57. Lees, S.; Davidson, C.L. The role of collagen in the elastic properties of calcified tissues. *J. Biomech.* **1977**, *10*, 473–486.
 58. Frasca, P.; Harper, R.A.; Katz, J.L. Mineral and collagen fiber orientation studies in human haversians. *J. Dent. Res.* **1978**, *57*, 526–533.
 59. Frasca, P.; Harper, R.A.; Katz, J.L. A new technique for studying collagen fibers and ground substances in bone with scanning electron microscopy. *Microsc. Acta* **1978**, *80*, 211–214.
 60. Weiner, S.; Traub, W. Organization of hydroxyapatite crystals within collagen fibrils. *FEBS letters* **1986**, *206*, 262–266.
 61. Weiner, S.; Traub, W. Bone structure: from angstroms to microns. *FASEB* **1992**, *6*, 879–885.
 62. Wagner, H.D.; Weiner, S. On the relationship between the microstructure of bone and its mechanical stiffness. *J. Biomech.* **1992**, *25*, 1311–1320.
 63. Hasegawa, K.; Turner, C.H.; Burr, D.B. Contribution of collagen and mineral to the elastic anisotropy of bone. *Calc. Tissue Int.* **1994**, *55*, 381–386.
 64. Landis, W.J.; Hodgins, K.J.; Arena, J.; et al. Structural relations between collagen and mineral in bone as determined by high voltage electron microscopic tomography. *Microsc. Res. Tech.* **1996**, *33*, 192–202.
 65. Rho, J.Y.; Kuhn-Spearing, L.; Zioupos, P. Mechanical properties and the hierarchical structure of bone. *Med. Eng. Phys.* **1998**, *20*, 92–102.
 66. Rho, J.Y.; Currey, J.D.; Zioupos, P.; et al. Microstructural elasticity and regional heterogeneity in



- human femoral bone of various ages examined by nano-indentation. *J. Biomech.* **2002**, *35*, 189–198.
67. Lakes, R.S.; Katz, J.L. Interrelationships among the viscoelastic functions for anisotropic solids: application to calcified tissues and related systems. *J. Biomech.* **1974**, *7*, 259–270.
 68. Sedlin, E.D. A rheological model for cortical bone. Study of the physical properties of human femoral samples. *Acta. Orthop. Scan.* **1965**, *83*, 1–77.
 69. Smith, R.; Keiper, D. Dynamic measurement of viscoelastic properties of bone. *Am. J. Med. Elec.* **1965**, *4*, 156–160.
 70. Laird, G.W.; Kingsbury, H.B. Complex viscoelastic moduli of bovine bone. *J. Biomech.* **1973**, *6*, 59–67.
 71. Lugassy, A.A.; Korostoff, E. Viscoelastic behavior of bovine femoral cortical bone and sperm whale dentin. In *Research in Dental and Medical Materials*; Korostoff, E., Ed.; Plenum Press: New York, 1969; 1–17.
 72. Black, J.; Korostoff, E. Dynamic mechanical properties of viable human cortical bone. *J. Biomech.* **1973**, *6*, 435–438.
 73. Knothe-Tate, M.L.; Knothe, U.; Niederer, P. Experimental elucidation of mechanical load-induced fluid flow and its potential role in bone metabolism and functional adaptation. *Am. J. Med. Sci.* **1998**, *316*, 189–195.
 74. Lakes, R.S.; Katz, J.L.; Sternstein, S.S. Viscoelastic properties of wet cortical bone. I. Torsional and biaxial studies. *J. Biomech.* **1979**, *12*, 657–678.
 75. Lakes, R.S.; Katz, J.L. Viscoelastic properties of wet cortical bone. II. Relaxation mechanisms. *J. Biomech.* **1979**, *12*, 679–687.
 76. Lakes, R.S.; Katz, J.L. Viscoelastic properties of wet cortical bone. III. A nonlinear constitutive equation. *J. Biomech.* **1979**, *12*, 689–698.
 77. Lakes, R.S.; Katz, J.L. Viscoelastic properties of bone. In *Natural and Living Materials III: Structure-Property Relationships in Biomaterials*; Hastings, G.W., Ducheyne, P., Eds.; CRC Press: Boca Raton, FL, 1984; 61–87.
 78. Gottesman, T.; Hashin, Z. Analysis of viscoelastic behavior of bones on the basis of microstructure. *J. Biomech.* **1979**, *13*, 89–96.
 79. Katz, J.L.; Kinney, J.H.; Spencer, P.; et al. Elastic anisotropy of bone and dentitional tissues. *J. Mat. Sci. Mat. Med.* **2005**, *16* (9).
 80. Kinney, J.H.; Gladden, J.R.; Marshall, G.W.; et al. Resonant ultrasound spectroscopy measurements of the elastic properties of human dentin. *J. Biomech.* **2004**, *37* (4), 437–441.
 81. Kinney, J.H.; Pople, J.A.; Marshall, G.W.; et al. Collagen orientation and crystallite size in human dentin: a small angle x-ray scattering study. *Calcif. Tissue. Int.* **2001**, *69*, 31–37.
 82. Pollack, S.R. Streaming potentials in bone. In *Bone Mechanics Handbook*, 2nd Ed.; Cowin, S.C., Ed.; CRC Press: Boca Raton, FL, 2001; 24-1–24-22.
 83. Singh, S.; Katz, J.L. Electromechanical properties of bone: a review. *J. Bioelect.* **1988**, *7* (2), 219–238.

Healing of Bone and Connective Tissues

William L. Murphy

Departments of Biomedical Engineering and Pharmacology, University of Wisconsin, Madison, Wisconsin, U.S.A.

Kelley Grorud

Department of Natural Science, Edgewood College, Madison, Wisconsin, U.S.A.

Ray Vanderby, Jr.

Orthopedic Surgery, University of Wisconsin, Madison, Wisconsin, U.S.A.



INTRODUCTION

Connective tissues (CTs) bind, support, and protect the various structures of the body. Derived from embryonic mesoderm, each tissue has an extensive extracellular matrix (ECM) containing relatively few cells. The ECM gives each tissue its particular character. Examples include ligament, tendon, cartilage, meniscus, fascia, intervertebral disc, bone, dermal layer of skin, adipose, iris, etc. Blood is also classified as a CT but is not included herein. In a broad sense, CTs are composed of tissue specific cells (e.g., osteoblasts and osteoclasts in bone, fibroblasts in ligament and tendon, chondrocytes in cartilage, etc.) and their ECMs, which are composed of collagen fibrils, elastin fibrils, proteoglycans, and other glycoproteins, in varied combinations and proportions (Fig. 1 and Table 1). Collagen provides tensile strength and stiffness, and elastin allows large amounts of stretch and recovery (i.e., elasticity) in certain CTs. Many of the other proteins help hydrate the “ground substance” within the ECM. Hydration is important for transport of nutrition and wastes and for mechanical behavior and function. In addition to these organic components, bone is highly mineralized to provide compressive strength and stiffness. Bone also provides a storage site for minerals and a medium (marrow) for the development and storage of cells and cytokines.

Mechanical stability and proximity of the damaged tissue ends are prerequisites for CT healing. An unstable bony fracture will not mineralize. Ruptured intracapsular ligaments floating in synovial fluid have no opportunity to heal. With stability and proximity satisfied, the biological processes have an opportunity to proceed. This mechanical prerequisite is assumed to be satisfied from this point forward and only the biological processes are considered hereafter.

The potential of CTs to heal varies greatly. Ligaments outside of synovial capsules (e.g., medial and lateral collateral ligaments, MCL, and LCL) heal reasonably well, although typically with scar tissue and

greater joint laxity before bearing load. Tendons typically heal if the ruptured ends remain in stable proximity, but matrix proliferation during healing can produce adhesions that impair function. Some CTs have much lower healing potential. Partial thickness lesions in articular cartilage, for example, typically do not heal. Full thickness lesions can spontaneously heal, but usually with a functionally inferior fibrocartilage. Bone, on the other hand, is quite unique in that it can regenerate itself without scar or functional degradation.

Because of their specific cells, structures and functions, CTs do not heal precisely like other tissues. The CT healing response is directed from chemical signals released from local cells, surrounding ECM, and vasculature. The number and type of cells that infiltrate a wound, and the type of local ECM, greatly influence the quality of the repair. Partial thickness lesions in articular cartilage, for example, have no local blood supply to initiate and support a healing response. When soft CTs heal, they have common wound healing processes and conditions. Bone has additional processes that allow it to regenerate its calcified matrix.

This chapter covers two broad topics: healing requirements as they pertain to soft CTs and healing requirements of bone. Some sub-topics in these sections overlap, showing commonalities that underly all CT healing. Others are tissue specific. Within these topics are concepts of deficient healing. By recognizing these deficiencies, targets become apparent for tissue engineering solutions. The scope of this chapter is very limited. Fortunately, recent literature probes healing of each CT and each CT subgroup considered herein in much greater depth.^[3,15,16]

SOFT CONNECTIVE TISSUE HEALING

A wound or injury disrupts the normal structure and function of these tissues. The wound healing process is

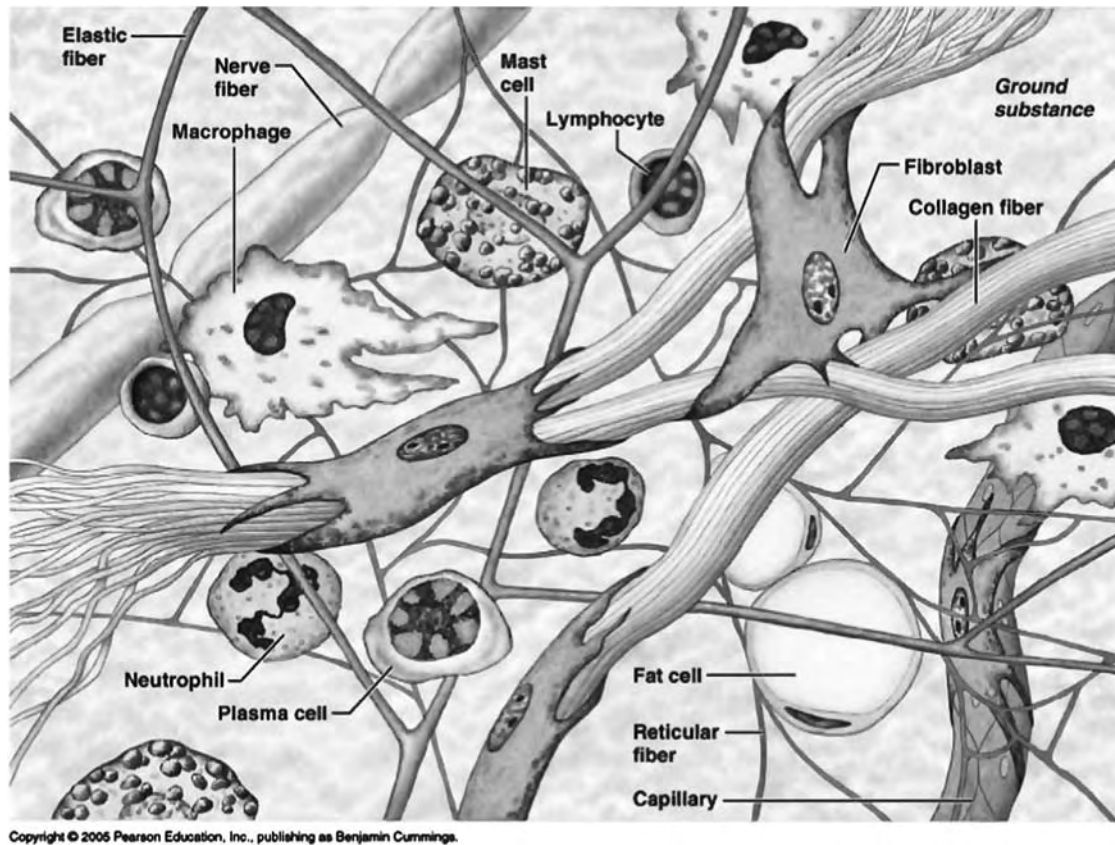


Fig. 1 Components of soft connective tissues.^[3] Soft connective tissues are comprised of tissue-specific cells (such as fibroblasts, chondrocytes, or adipocytes) and extracellular matrix (ECM). The ECM contains numerous proteins, many of which are fibrous in nature. Shown here are: (1) collagen, which provides tensile strength to the tissue, (2) elastin, which allows the tissue to stretch, and (3) reticular fibers, which are small collagen fibers that act as anchors for various cells. Notice that inflammatory cells and blood components (platelets and white blood cells) can also be found within the matrix of some soft CTs.

complex and dynamic as tissues repair with anatomical connections to restore function. Four overlapping stages in wound healing begin immediately following an injury: hemostasis, inflammation, proliferation, and remodeling (Fig. 2). Progression through these stages (which herein is termed “normal” healing) closes the wound and repairs the damage; however, restoring and rebuilding damaged tissue may continue for years post-injury. Herein, “normal” healing indicates the process by which a wounded CT largely recovers its normal function, but not necessarily its structure. Healing “well” means that function is recovered. Alternatively, “poor” healing indicates that the CT does not recover its normal function.

HEMOSTASIS (IMMEDIATELY POST INJURY)

Hemostasis refers to the process that stops bleeding or blood flow. In most CTs, this is the first process to occur during wound healing. From broken vessels,

platelets enter the wound site, contacting the exposed collagen ends.^[15] Platelets aggregate around the damaged matrix. When attached to the matrix, platelets release clotting factors (i.e., chemicals that signal the clotting of any broken vessels in the area).^[6] A fibrin clot, or hematoma, forms as a plug in broken vessels, which also provides a provisional matrix to anchor other cell types in the wound site.^[6]

Platelets then release growth factors into the wound site.^[15] Growth factors are a class of chemicals that regulate cell growth, cell migration, ECM production, enzyme production, and cell differentiation. In particular, platelets release platelet-derived growth factor (PDGF) and transforming growth factor-beta-1 (TGF β -1).^[6,12] Both of these factors stimulate migration of cells into the wound; PDGF is a chemo-attractant for inflammatory cells including neutrophils and macrophages,^[6] and TGF β -1 is a chemo-attractant for smooth muscle cells and fibroblasts.^[6,15] A list of growth factors and other chemicals important to CT healing is found in Table 2.

Table 1 Classification of CTs

Classification	Celltype	Specific CT	Example	Organization	Body location
CT proper	Fibroblast	Loose	Areolar	Disorganized matrix; varied collagen fiber diameter; loosely packed	Surrounding many of the body's organs
			Adipose	Tightly packed cells with some matrix between	Fat insulation throughout the body
			Reticular	Very small diameter collagen fibers; loosely packed; disorganized matrix	Bone marrow as an anchor for developing blood cells
CT proper	Fibroblast	Dense	Irregular	Densely packed collagen fibers in a disorganized fashion	Dermal layer of skin; joint capsules
			Regular	Densely packed collagen in aligned matrix	Ligament and tendon
Cartilage	Chondrocyte		Hyaline	Complex organization of type-II collagen; high water content	Articular cartilage, which covers long bone ends
			Fibrocartilage	Cartilage organized along lines of loading	Intervertebral discs; meniscus; symphyses
			Elastic	Cartilage containing elastic fibers as well	Ears
Bone	Osteocyte			Woven collagen fibers surrounded by mineralized matrix	Any bone within skeleton
	Osteoblast				
	Osteoclast				

INFLAMMATION (IMMEDIATE TO 2–5 DAYS)

Neutrophils, a type of inflammatory cell, follow signals from platelets and dominate the wound site within the first 24 h after injury.^[19] The role of the neutrophils is to clean the wound site so repair can follow. Neutrophils are phagocytic, that is, they ingest foreign materials, bacteria, damaged host cells, and damaged matrix.^[15,19] Mast cells, a second type of inflammatory cell, are also present in the wound within the first 24 h. Mast cells contain granules that are filled with enzymes, including histamine and active amines. When released, these chemicals increase permeability in the local blood vessels. Increased permeability allows fluid and large inflammatory cells to slip out of the local vessels and into the injured region.^[6,15,21]

On the second day following an injury, monocytes dominate the wound site. These cells quickly become macrophages, which are inflammatory cells that are essential to healing. Macrophages, highly phagocytic

inflammatory cells, have two essential roles. First, they engulf large debris-filled neutrophils, damaged matrix, and any leftover bacteria or foreign debris.^[6,15,19] Second, they release PDGF and TGF β -1 and -2, which further attract fibroblasts and smooth muscle cells to the wound site.^[15] In addition, macrophages also release other growth factors that stimulate cell proliferation and angiogenesis (new blood vessels formation), including vascular endothelial growth factor (VEGF), fibroblast growth factor (FGF-1 and -2) and interleukin-1 (IL-1).^[15,19] Signaling processes during inflammation are summarized in Fig. 2.

PROLIFERATION (2 DAYS TO 3 WEEKS)

Proliferation is the major reconstruction phase of wound healing. During this phase fibroblast proliferation and matrix rebuilding dominate. Fibroblasts



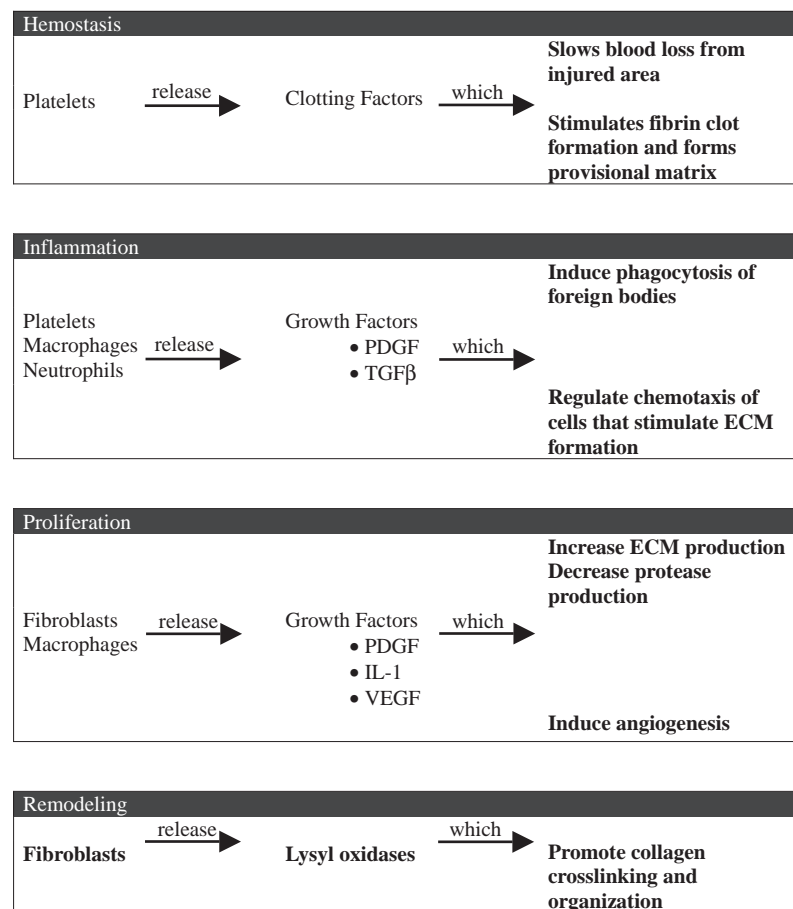


Fig. 2 Stages of healing. There are four overlapping stages of healing that must be coordinated spatially and temporally. Each stage utilizes a specific type of cell (column 1) which releases a chemical mediator (column 2) to perform a specific biological function (column 3). Growth factors often act as mediators in the healing process.

in the wound serve two roles. First, they contract the wound, bringing wound ends closer together. Second, they produce a disorganized ECM to replace the damaged tissue.

Fibroblasts in the wound site attach to the provisional fibrin matrix and produce type I and type III collagens, which are self-assembling, fiber-forming proteins. These new fibrils and fibers connect to the damaged ends of the wound, providing tensile strength. Fibroblasts also produce proteoglycans, ECM proteins that aid in fibrillogenesis and tissue hydration. In fact, many CT injuries remain overly hydrated for several weeks following this proliferative phase. The proliferative phase ECM is often called granulation tissue due to its lack of microstructural organization.

The major signal that regulates this exuberant ECM proliferation is TGFβ-1 (Table 2, Fig. 2).^[15] This growth factor is released from neutrophils, platelets, and macrophages and has three major effects on the ECM deposition by fibroblasts. TGFβ-1 increases the transcription of collagen, proteoglycans, and fibronectin (all ECM proteins),^[6,15] it decreases secretion of proteases (proteins that degrade matrix proteins),^[6] and it increases the secretion of protease inhibitors,

including tissue inhibitors of metalloproteinases (TIMPs) to protect the newly forming matrix.^[6]

With increased cell number and activity in the wound, there is an increased demand for oxygen and nutrients. Local signals, including low pH, low oxygen tension, and increased lactate stimulate cells in the wound to produce VEGF, FGF, and TGFβ-1 and -2.^[6] These growth factors stimulate angiogenesis (new vessel formation) to help keep up with oxygen demands.

REMODELING (3 WEEKS TO 2 YEARS)

The proteins laid down during the proliferative phase are disorganized. The matrix is usually composed of small diameter collagen fibers, and it is weaker than the surrounding healthy tissue.^[15] Remodeling of this weak, disorganized matrix begins within days of the injury, but it can continue for several years after the injury.^[7,10] During remodeling the weak collagen fibers mature. A chemical within the matrix, lysyl oxidase, aids in the cross-linking of these collagen fibers, which is associated with increases in strength and stability of

Table 2 Growth factors important to CT healing

Growth factor	Cell source	Target cell	Biological activity
TGF β -1, β -2	Platelet, macrophage, fibroblast	Inflammatory cells, fibroblasts, osteocyte	Chemotaxis, proliferation, matrix production
PDGF	Platelet, macrophage, fibroblast, vascular smooth muscle	Neutrophil, macrophage, fibroblast, vascular smooth muscle	Chemotaxis, proliferation, matrix production
FGF -1, -2	Macrophage, fibroblast	Fibroblast, chondrocyte, endothelial cells, osteocyte	Proliferation, chemotaxis, angiogenesis
IGF	Fibroblast, macrophage	Fibroblast	Proliferation, collagen synthesis
IL-1	Macrophage, neutrophil	Macrophage, fibroblast	Proliferation; collagenase synthesis, chemotaxis
CTGF	Fibroblast	Fibroblast	Matrix production
VEGF	Macrophage	Endothelial cells	Angiogenesis
BMP -2, -4, -7, -9	Bone precursor cells	Osteocytes	Cell proliferation, matrix production



the fibers (Fig. 2).^[6] Matrix in most uninjured CTs is strong and highly organized. By comparison, scar is weaker and more disorganized. The maximum strength of a naturally healing wound rarely reaches more than 80% of that of the surrounding healthy tissue.^[15] Stiffness and strength are compromised in disorganized scar tissue because recruitment of collagen fibers is less optimal from a microstructural perspective to bear loads and resist failure.

HEALING POTENTIAL OF SPECIFIC CONNECTIVE TISSUES

Specific types of connective tissues have similar healing principles, but each type of tissue has its own potential for healing (Fig. 3). In general, CTs that heal well have ample blood supply, innervation, and a randomly organized matrix. Tissues that heal poorly lack their own blood supply and innervation, and they can have (in their normal tissue) a highly organized matrix that is difficult to reproduce during healing.

Connective Tissues that Heal Well

Areolar and dense irregular connective tissues generally heal well. These tissues, which make up the covering layers for many organs and joints, and the deep (dermal) layers of the skin, have a relatively disorganized ECM (meaning that collagen fibers do not have a strong preferential direction), and they have variability in collagen fiber diameters. The major cell type for these two tissues is the fibroblast. Fibroblasts in these tissues produce type I and type III collagens, the same types that are produced during healing. In general, these tissues can bear functional loads in multiple planar directions.^[16] Areolar and dense irregular connective tissues are highly vascular and innervated, which gives these tissues ready access to plasma proteins, nutrients, oxygen, and inflammatory cells.^[16]

Connective Tissues that Heal Moderately Well

The healing potential of dense regular CT, such as ligament and tendon, varies by environment, loading

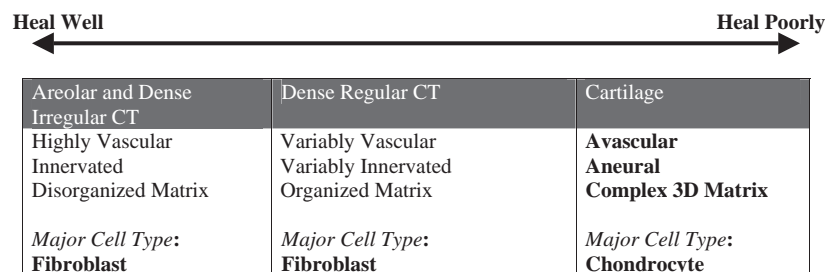


Fig. 3 Healing potential of soft CTs. Each CT heals to a different degree. Some heal very well, while others (such as articular cartilage) heal poorly or not at all. In general, healing potential correlates with the degree of innervation and vascularization of the tissue, and with the dominant cell type that is found within the tissue.

and complexity of the tissue. For example, ligaments that have a complex environment or reside within a joint capsule, such as the anterior cruciate ligament, do not heal well and often require surgical reconstruction.^[5] However, the extra-capsular ligaments (e.g., medial collateral ligament) heal without surgical intervention.

Most of these dense regular CTs heal moderately well with time. They have a highly organized matrix with tightly packed collagen fibers that align along an axis of load bearing. Examples are ligaments outside of joint capsules or tendons. These dense regular CTs have fewer fibroblasts than dense irregular or areolar CTs. Dense regular CTs tend to be poorly vascularized and moderately innervated, with the majority of nerves and vessels found in the outermost layers of the tissue (e.g., peritenon). During healing, a poorly organized matrix fills the injury, disrupting the normal fiber organization and reducing mechanical strength and stiffness. Long term, disorganization can leave the tissue at risk for further damage and compromise the overall health of a joint. Dense regular CTs do not heal to their original strength and stiffness. The resulting deficit can exist for many years following an injury.

Connective Tissues that Heal Incompletely

An example of a CT with severely limited healing potential is articular cartilage.^[9] Even small cartilage defects rarely heal on their own. Damage to articular cartilage can lead to premature osteoarthritis, which then causes pain and disrupts normal joint function. Mature cartilage differs from most other CTs in many ways. First, it does not have a blood, lymph, or nervous supply of its own.^[9] Second, in healthy cartilage the predominant cell type is the chondrocyte, which is responsible for maintaining the large portions of type II collagen and proteoglycans that are present in the normal tissue. These chondrocytes are surrounded by a dense ECM which keeps them sheltered even from immune recognition.^[9] Even though the chondrocytes continue to produce ECM throughout their lifespan, they do not respond to injury.^[18] Unless a cartilage injury involves the underlying bone, it only stimulates a brief and functionally insufficient response in cell proliferation and ECM production.^[18,22] This limited activity results in incomplete filling of the defect. When subchondral bone is involved in the injury it provides a blood supply to initiate a healing cascade. However, the injured region becomes infiltrated with fibroblasts (rather than chondrocytes), which fill the defect with fibrocartilage made of type I and type III collagens (rather than the original type II matrix) and with reduced level of aggrecan (a large proteoglycan to maintain hydration, even under large compressive

loads).^[22] The repair is less able to distribute compressive loads and resist wear than the highly specialized, articular cartilage tissue. Consequently, it is less able to resist joint degeneration and osteoarthritis. Larger cartilage defects frequently require surgical reconstruction or even joint replacement.^[22]

HEALING OUTCOMES

The “normal” healing process can only proceed when all of the cells, growth factors, and other chemicals are well coordinated both temporally and spatially. Alterations to this ‘normal’ cascade of events can result in delayed, incomplete, or pathological healing. Four categorical outcomes are associated with CT healing. They are normal repair (detailed above), excessive healing (extensive scarring), deficient healing (functionally deficient tissue), and complete regeneration.^[6] Excessive and deficient healing are pathological healing responses that occur when the orderly processes and conditions of healing are disrupted or absent.

Excessive Healing

Excessive healing, sometimes called fibrosis, occurs when too much ECM is deposited into the wound site. The result is an altered structure with pathological function. Clinical examples of excessive healing include keloids, hypertrophic scar, adhesions, contractures, cirrhosis, and nerve fiber blockage.^[6] Wounds with excessive healing generally express 2–3 times more TGF β -1 and 2–3 times more type I collagen than normal tissue.^[6,15] These wounds may also have an overabundance of myofibroblasts (causing excessive contracture) and increased mast cell density.

Deficient Healing

Deficient healing is the opposite of fibrosis. Insufficient ECM deposition into a wound leaves the tissue mechanically weakened. A common clinical example is a chronic ulcer, which is a nonhealing dermal wound. Deficient wounds may develop due to over-activity of inflammatory cells in the wound site.^[6] Ulcers are characterized by increased neutrophils, collagenase (an enzyme that degrades collagen), and reactive oxygen species. Deficient wounds do not heal until the inflammatory process is under control.^[6] Left untreated, an open wound site is prone to infection.

Regeneration

In contrast to wound repair, regeneration is the process by which injured tissue is replaced exactly as it was

before the injury. Most adult human tissues do not have complete regenerative capacities; however some adult organs will partially or completely regenerate when injured, including the liver, epidermis, some nerves, and bone. Fetal tissue can repair through regeneration.^[20] Interestingly, fetal tissue does not undergo the inflammatory process that takes place in adult tissues.^[20] By studying differences between adult and fetal tissues, science may learn how to regenerate injured adult CTs.

Tissue Engineering of Soft Connective Tissues

Tissue engineering methodologies are actively being sought to address healing deficiencies in soft CTs.^[9] For example, due to the limited of healing in articular cartilage (AC) and its critical role in joint motion, the need for an engineered repair tissue is particularly high. However, engineering AC repair is problematic because mature chondrocytes are slow to build or remodel tissue and because the ECM is micro-structurally and functionally complex. Cartilage ECM contains large quantities of type II collagen and proteoglycans. The resulting matrix attracts and holds water within the ECM, which stiffens the tissue for impact loads and also distributes contact pressure.^[9] Healthy cartilage is nearly wear-resistant and nearly frictionless. Fluid transport slowly allows more deformation (creep) during sustained loads.^[9] To engineer a repair tissue that mimics this behavior, the correct composition, hydration and organization of the ECM must be synthesized. Tissue engineering must then focus on (1) ways to attract, place, or differentiate chondrocytic cells into a cartilage defect with or without a matrix scaffold, and (2) ways to stimulate these cells to produce an appropriate ECM with collagen II and aggrecan. Undifferentiated mesenchymal cells can differentiate into chondrocytes and form cartilage with appropriate signaling. For example, mesenchymal cells exposed to TGF β -1 can be induced to differentiate into chondrocytes and form a cartilaginous matrix. In addition, the new chondrocytes must remodel and sustain the newly engineered tissue. Bone morphogenic protein-2 (BMP-2) and BMP-7 are potent stimulators of matrix proteins in cultured chondrocytes. However, recent studies show that the biological and mechanical behaviors of healthy AC are significantly altered in engineered cartilage.^[9] The compressive modulus of most engineered cartilage is substantially lower than healthy cartilage, although some improvement can be gained by controlled mechanical loading. Furthermore, gene expression of chondrocytes in healthy AC (e.g., type II collagen and aggrecan) is altered in engineered cartilage, including the undesirable upregulation of type I collagen.^[9]

While AC is only one example of tissue engineering, similar problems persist when engineering other CTs. Most engineered soft CTs have reduced potential for remodeling and integration into the surrounding healthy tissues. The result is a material that remains somewhat isolated from the native tissue. The remodeling of bone (discussed later) suggests that soft CT healing would benefit from a more optimal remodeling process. The “normal” healing process (Fig. 2) shows that temporally controlled signals to specific cell types are requisite to promote the “normal” healing cascade. A provisional matrix must be provided or generated to stabilize the wound until remodeling can restore a functional result in the neo-tissue. Angiogenesis is essential for healing of most tissues and its temporal stimulation appears necessary. Mechanical loading must be controlled, since it is stimulatory to the healing process and functional recovery of the CTs. All of these issues represent targets for tissue engineering to promote, accelerate, or improve soft CT healing.

BONE HEALING

Bone fracture healing involves an exquisite coordination of multiple cell and signal types. The early stages of this process are superficially similar to soft connective tissue healing (described above). However, unique processes in bone healing are required to form a dynamic mineralized tissue. To limit the scope of this discussion, only long bone fracture healing is considered. Long bones are perhaps the best characterized type of bone fracture, and their healing process includes aspects of both major classes of bone formation (i.e., endochondral ossification and intramembranous ossification).

CHARACTERISTICS OF BONE FRACTURE HEALING

Anatomy and Physiology of Long Bone Healing: General Definitions

Long bones include a long cylindrical shaft called the diaphysis, which is terminated on each end by an epiphysis. The diaphysis includes an inner bone marrow cavity, or medullary canal, surrounded by an outer cortex of compact or cortical bone. The epiphysis is surrounded by a shell of compact bone and contains an inner cavity filled with trabecular or spongy bone. A fibrous and cellular region, termed the periosteum, covers the outer surface of the cortex. Periosteum contains an outer fibrous layer, and an inner cellular layer, which houses bone precursor cells that become operative during fracture healing. There are two



general classes of bone formation. The first is intramembranous ossification, which primarily occurs during development of calvarial bones, and involves bone formation directly upon or within fibrous membranes that initially contain undifferentiated mesenchymal cells (Fig. 4). The second process is endochondral ossification, which occurs in growth and repair of long bones, and involves bone formation within a cartilage template matrix (Fig. 5).

Similarities to Soft Connective Tissue Healing: Soft Callus Formation

The early stages of fracture healing in long bones involve an array of spatially and temporally regulated events that culminate in stabilization of the defect with a fracture callus. The process can be separated into three processes (Fig. 6(1–3)), which overlap significantly, and in many cases occur simultaneously: (1) acute inflammatory response (hematoma), (2) bridging of the defect with soft callus (chondrogenesis), and (3) intramembranous ossification to form the hard callus. Each stage is vital for fracture stabilization. The first two stages have clear connections between these processes and the soft CT healing processes described previously. During the acute “inflammatory phase,” blood vessels are damaged and the fracture site is partially deprived of oxygen and nutrients that typically flow through the marrow and the bone matrix. In addition, the bone defect contains debris as a result of the trauma, and there is typically associated damage to the surrounding soft tissues. These events invoke an influx of inflammatory cells, including neutrophils, platelets, and macrophages, which initiate formation of a fibrous matrix, causing the periosteum to bulge outward. This mass of blood, debris, and cellular components, called a hematoma, forms within hours of fracture. In concert with this process, committed osteoblast precursors in the periosteum begin to proliferate and differentiate into mature osteoblasts. Concomitant with this inflammatory response, osteoblasts begin to form thin plates of woven bone in the periosteal region. This new “flat” bone tissue is formed via intramembranous ossification and is called bony callus.

In the days following this acute inflammatory response, factors secreted by infiltrating cells (e.g., macrophages, neutrophils) invoke ingrowth and proliferation of differentiated chondrocytes and undifferentiated mesenchymal cells, which are induced to differentiate into chondrocytes. This initiates a “proliferation phase” for cells in the area delineated by the initial hematoma, leading to synthesis of a cartilage soft callus. The soft callus and the bony callus stabilize

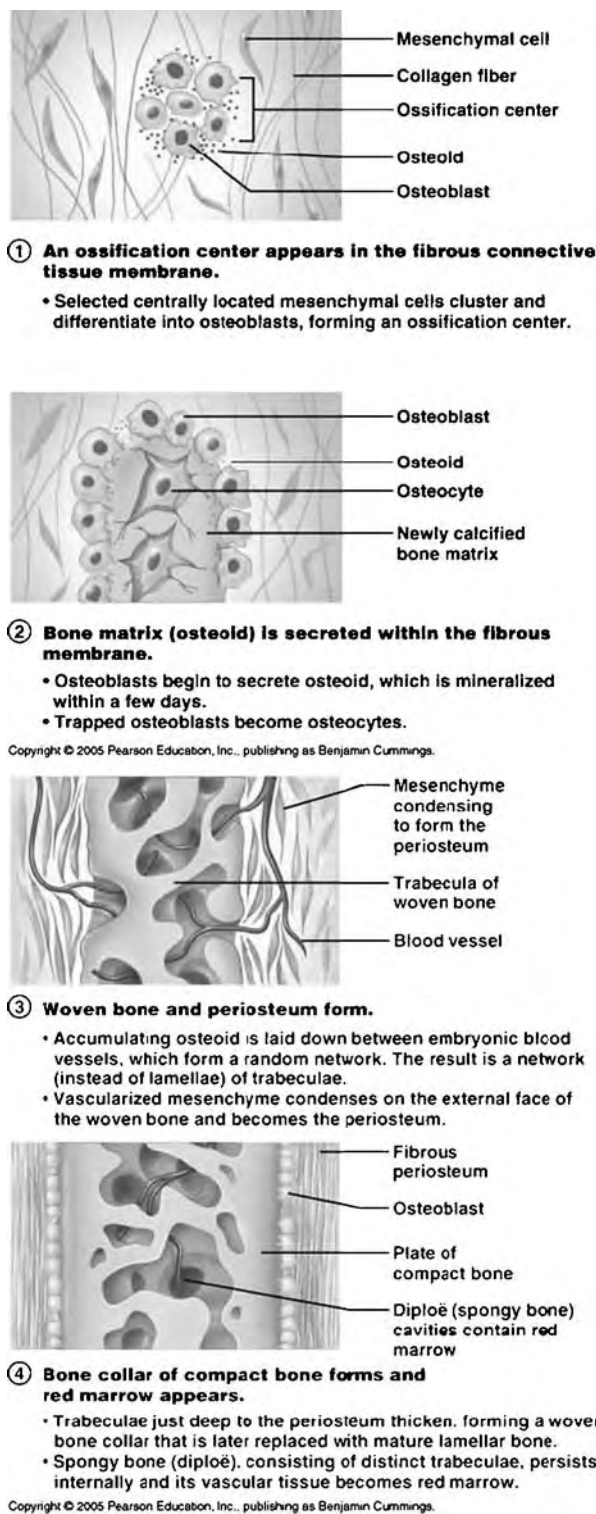
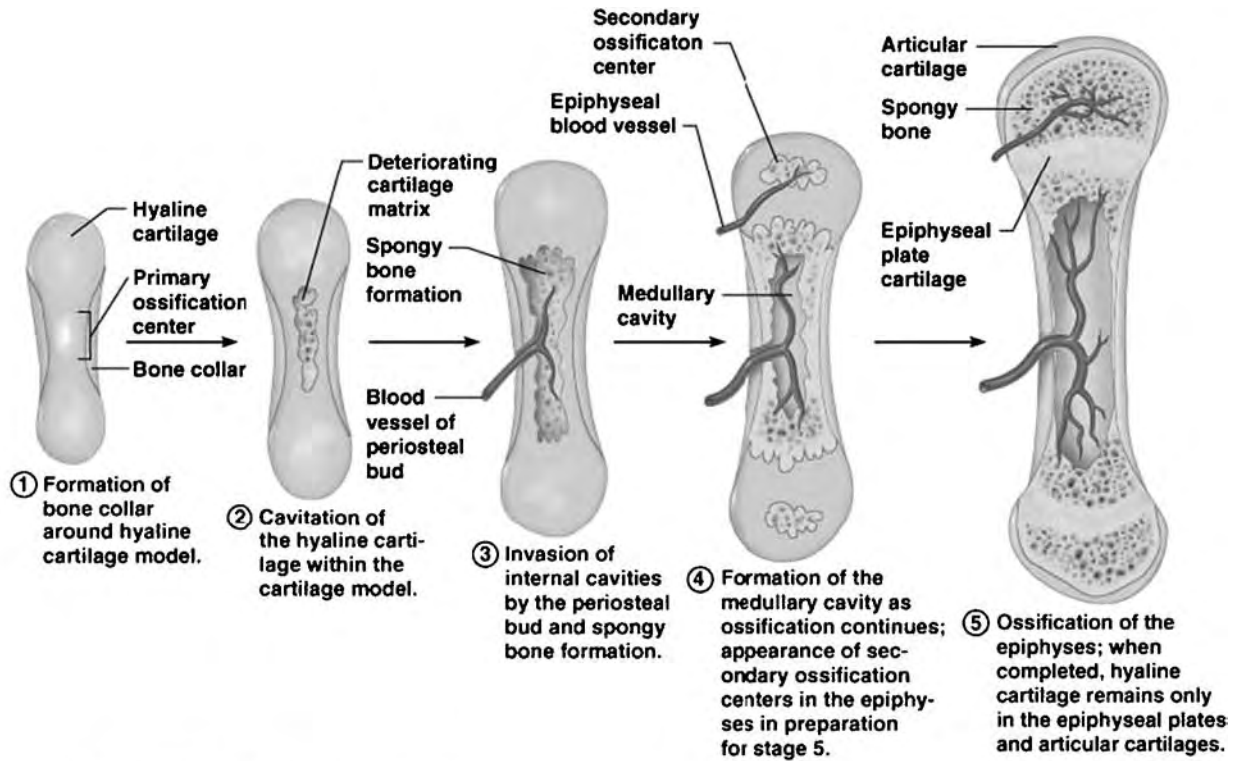


Fig. 4 Intramembranous ossification.^[16] Intramembranous ossification is a process that leads to the formation of the flat bones of the skull. During this process, bone forms within a fibrous membrane. A similar process occurs during the early stages of fracture healing to form a hard callus around the injured area.



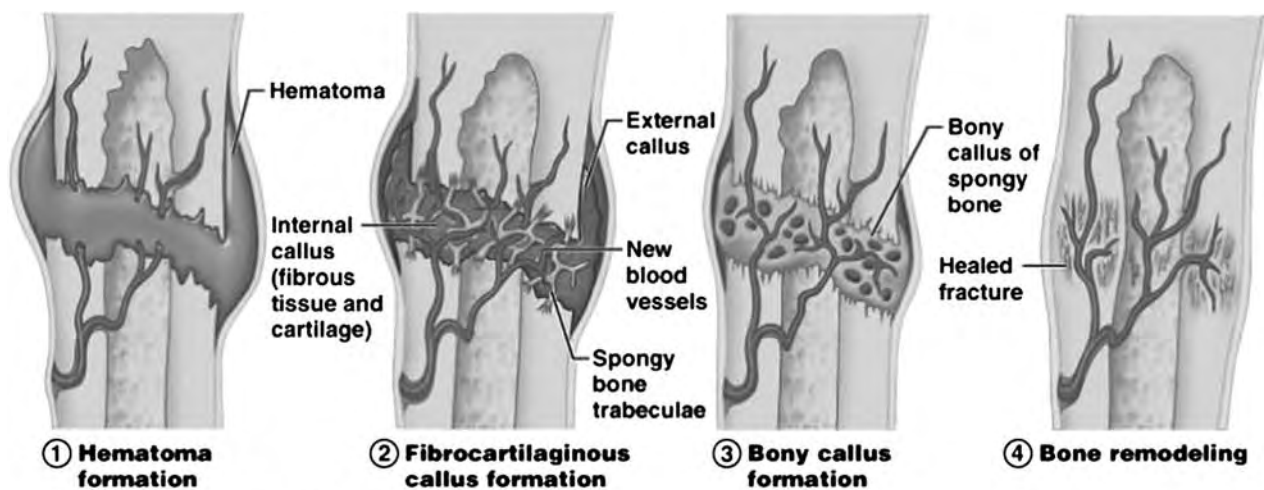
Copyright © 2005 Pearson Education, Inc., publishing as Benjamin Cummings.

Fig. 5 Endochondral ossification.^[16] Endochondral formation is a process that leads to the formation of long bones during development. During this process, cartilage forms to lay the structure of a bone. This cartilage calcifies and is then replaced by woven bone. A similar process takes place during fracture healing. In later stages of fracture healing, endochondral ossification replaces the soft callus that forms within the fracture site.

the fracture and provide a template for subsequent endochondral ossification. Remarkably, the entire, well-coordinated process of acute inflammation and bony and soft callus formation typically occurs within the first

2–3 weeks after bone fracture, and a variety of cells and inductive soluble signals orchestrate these events.

There are similarities between soft CT healing and the early stages of bone fracture. First, an initial



Copyright © 2005 Pearson Education, Inc., publishing as Benjamin Cummings.

Fig. 6 Fracture healing.^[16] Fracture healing occurs in overlapping stages, similar to the stages that occur during soft CT healing. Bone is unique, however, in that fracture healing includes ossification, the replacement of soft tissue with mineralized bone. In addition, the remodeling stage of fracture healing often leads to complete regeneration of bone into the site of repair.

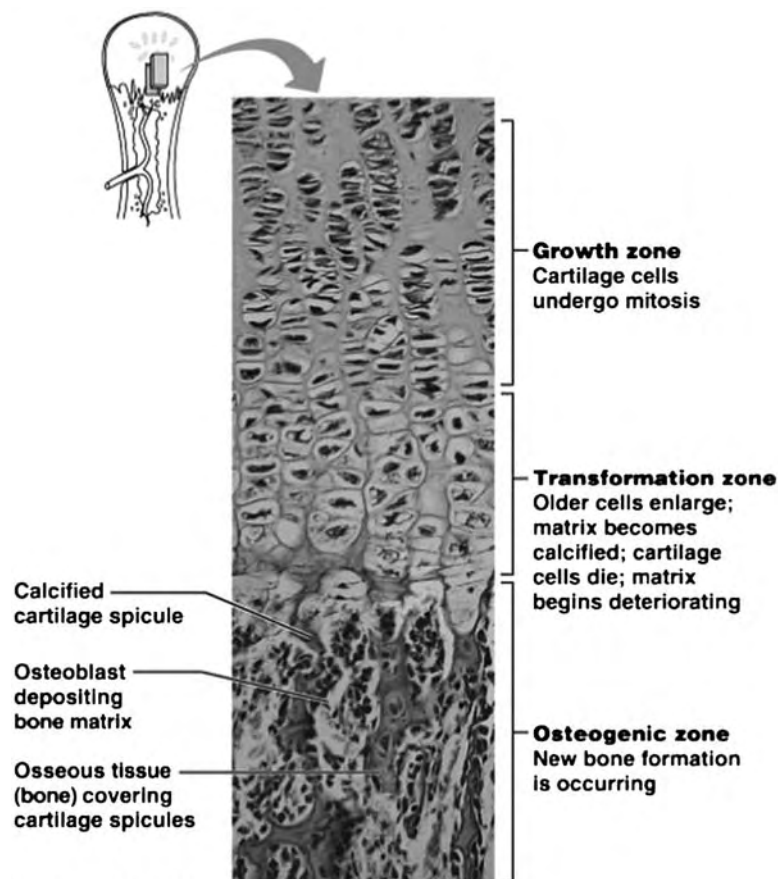


inflammatory response of the tissues to damage occurs. The initial 7–10 days after fracture involve formation of a hematoma, followed by inflammation and widespread proliferation of highly cellular fibrous tissues. The inflammation and proliferation processes are broadly similar to soft CTs. Second, many of the specific cell types involved in soft connective tissue healing are also operative during early bone healing. For example, undifferentiated fibroblasts (often termed “mesenchymal cells”) that reside in the bone marrow infiltrate bone defects within 24 h after fracture, and similar mesenchymal precursor cells enter bone defects from the periosteum within 72 h post fracture. These cells can differentiate into multiple cell types depending on environmental circumstances. Third, a provisional matrix is quickly established, providing a framework for subsequent remodeling of tissue and functional restoration.

Unique Characteristics of Bone Healing: Primary and Secondary Bone Formation

Despite the similarities between soft and hard connective tissue healing, there are also contrasts unique to bone. The primary distinctions stem from the ability

of bone cells to: (1) form mineralized tissue, either within a cartilage-based template matrix or within fibrous membranes (primary bone formation); and (2) remodel primary bone and ultimately regenerate the structure that existed prior to fracture (secondary bone formation). Primary bone formation in long bones occurs via endochondral ossification, and results in formation of a matrix of woven bone. Approximately 10 days post fracture, chondrocytes in the soft fracture callus enlarge or “hypertrophy” to form vesicular structures, and they begin to degenerate and stop synthesizing cartilage matrix. These hypertrophic chondrocytes then begin to degrade their existing matrix via secretion of proteases. This process leads to calcification of the matrix in a spatially-organized manner. The calcified cartilage then provides a substrate for synthesis of primary bone tissue by infiltrating osteoblasts, which form woven bone upon calcified cartilage substrates. The enlarged chondrocytes that remain then begin to undergo apoptosis, leading to a replacement of calcified cartilage with woven bone. The complete process of primary bone formation after fracture typically occurs in four weeks. The endochondral ossification that occurs during fracture healing is broadly similar to bone formation at the growth plate during long bone development (Fig. 7), in which



Copyright © 2005 Pearson Education, Inc., publishing as Benjamin Cummings.

Fig. 7 Growth plate formation.^[16] During fracture healing, bone replaces cartilaginous structures that are initially laid down to structurally reinforce the area and restore normal anatomy. As bone replaces cartilage, the chondrocytes undergo changes, including hypertrophy. These changes are similar to the ones that take place during growth plate development (shown in this figure).

the region between the diaphysis and the epiphysis contains stacked layers of proliferating chondrocytes, hypertrophic chondrocytes, apoptotic chondrocytes, and primary woven bone.

An intriguing aspect of bone tissue repair is the ability of bone cells to remodel primary bone into a structure that is indistinguishable from the initial, intact bone. This process requires the coordinated activity of osteoclasts, which resorb the existing primary woven bone matrix, and osteoblasts, which synthesize new secondary bone tissue. The factors that influence bone remodeling are complex, extensive, and beyond the scope of this chapter. It is important to note, however, that this optimized remodeling process is a key aspect of bone healing that does not occur to the same extent in other CT healing.

CELLS AND SIGNALS IN THE FRACTURE CALLUS

Identity of Precursor Cells

Mesenchymal cells that migrate into the bone defect and differentiate into chondrocytes to form the cartilage-based soft callus, are thought to originate in the marrow, periosteum, and endosteum. These cells are capable of differentiating down several skeletal lineages; therefore, the signaling environment plays a prominent role in bone healing. Cellular components that direct inflammatory processes during the first hours and days post fracture are similar to those described earlier for soft CTs. Several of the key growth factor-producing cell types and corresponding growth factors are described below.

Key Soluble Signals

The acute inflammatory phase of fracture healing is characterized by infiltration of degranulating platelets, monocytes, and macrophages, each of which secrete important soluble signals. Degranulating platelets secrete platelet-derived growth factors (PDGFs), which promote growth of macrophages into the defect. In turn, monocytes and macrophages secrete fibroblast growth factors (FGFs) -1 and -2 and transforming growth factors (TGFs) - β 1 and - β 2, and thereby stimulate proliferation of fibroblasts, chondrocytes, and osteoblasts, as well as mesenchymal cells. Each soluble factor, along with vascular endothelial growth factor (VEGF), has a significant effect on the generation of a vascular supply within the defect (see “role of angiogenesis” section below).

CELLS AND SIGNALS DURING DEVELOPMENT OF PRIMARY BONE TISSUE

Bone fracture healing is a complex process that is tightly regulated in time and space by molecular and cellular effectors. The following provides an overview of signals vital to bone fracture healing, emphasizing signals that can be manipulated in tissue engineering.

Identity of Precursor Cells

Primary cell types that direct fracture healing are: (1) chondrocyte precursors from periosteum, endosteum and bone marrow, (2) osteoblast precursors from periosteum, (3) differentiated osteoblasts from bone matrix, and (4) undifferentiated mesenchymal cells derived from marrow and the periosteum. Chondrocytes are present in the defect area in the first days post fracture creating a cartilage template, which is ultimately remodeled and replaced with woven primary bone. Early in fracture healing osteoblast precursors from periosteum play a key role to form the bony callus. Osteoblasts also infiltrate the cartilage template to create woven bone via endochondral ossification. Mesenchymal cells are not committed to differentiate into a specific cell type, and therefore their ability to form cartilage callus or primary bone is strongly dependent on their microenvironment. For example, mesenchymal cells exposed to bone morphogenetic protein-2 (BMP-2) may be induced to differentiate into osteoblasts. A single injection of BMP-2 into a defect promotes generation of primary bone. In contrast, mesenchymal cells exposed to TGF β 1 may be induced to differentiate into chondrocytes and form cartilage. These examples demonstrate the importance of environmental factors for proper differentiation of mesenchymal cells during bone fracture healing.

Mechanical Signals

Mechanical forces influence remodeling of adult bone and, in turn, regulation of bone strength over time. The mechanobiology of bone is often summarized by Wolff's Law, which indicates that bone adapts to its mechanical environment by dynamic remodeling. The influence of mechanics on fracture healing is perhaps even more dramatic. Too much relative motion at the fracture site can result in a non-union, moderate motion in a stabilized fracture allows “normal” healing (Fig. 6), and very restricted motion can induce healing without robust callus formation. Mechanical stimulation is also vital to proper cartilage and bone development, and the mechanical properties of bone can be correlated to the load applied during development. Specifically, the mechanical environment influences



local gene expression by cells within a bone defect, and it is an important determinant of blood vessel ingrowth after bone fracture.^[4] Furthermore, mechanical forces have a pronounced effect on bone formation by osteoblasts and osteoblast precursor cells in cell and tissue culture.^[4] This evidence helps explain the role of mechanical forces during important fracture healing.

Role of Angiogenesis

Recent studies indicate that the growth of new blood vessels from a pre-existing vascular supply is particularly important for proper growth, regeneration, and healing of bone tissue.^[8,17] Blockage of angiogenesis inhibits endochondral ossification in the growth plate and also inhibits fracture healing. In addition, delivery of VEGF, a stimulator of angiogenesis, from materials in a critical bone defect has a positive effect on bone regeneration. There are several possible explanations for the positive effects of vascular ingrowth on bone generation. The vasculature may be vital to survival and proper function of pre-existing osteoblasts or may include cells that provide stimulatory signals to osteoblasts (e.g. endothelial cells). Alternatively, the vascular supply may be vital for differentiation of bone precursor cells or for recruitment of blood-borne stem cells capable of forming bone.^[13] For these reasons, and perhaps others, factors that influence ingrowth of a functional vascular supply are vital for proper growth, repair, and regeneration of bone tissue. These factors include the mechanical stability of a defect, as well as a variety of growth factors.

Roles in immune stimulation and fibroblast proliferation, several mediators of inflammation in the fracture callus also support fracture healing by inducing blood vessel growth into the callus. Growth factors that stimulate angiogenesis during fracture healing include FGF-2, PDGF-BB, BMP-2, BMP-4, BMP-7, and VEGF.^[1] VEGF is a particularly potent stimulator of angiogenesis, and it is synthesized by mesenchymal cells, osteoblasts, and chondrocytes during the first 7 days following fracture. Angiogenic factors also have direct effects on bone-forming cells. It is therefore difficult to specifically delineate the effects of growth factor-induced angiogenesis on bone healing. However, in the absence of VEGF bone does not heal or grow properly, and this effect has been largely attributed to inhibition of angiogenesis.

Key Soluble Signals

Compelling evidence in previous sections indicates that the mechanical environment and angiogenesis strongly influence fracture healing. However, the most potent direct effectors of bone regeneration may be

osteogenic growth factors. Soluble growth factors have widely variable effects on the cells that direct bone repair, as they can influence migration, proliferation, differentiation, and angiogenesis within a bone defect. The primary function of the factors secreted by cells during the inflammatory phase of fracture healing is to stimulate infiltration of appropriate cell and tissue types. Macrophages and platelets produce factors (e.g., PDGFs, TGF β s, FGFs), which invoke infiltration and proliferation of mesenchymal cells and chondrocytes in the fracture callus. These factors also stimulate angiogenesis, along with VEGF and tumor necrosis factor α (TNF α). The net effect of all of these factors is to load the fracture callus with chondrocytes, blood vessels, and undifferentiated mesenchymal cells, which creates an environment primed for the subsequent formation of primary bone.

Once a mature fracture callus has formed, cells produce several factors that more directly influence the activity of bone-forming cells (e.g., osteoblasts, and osteoblast precursor cells). Members of the TGF β superfamily of proteins, including TGF β -1, TGF β -3 and BMPs -2, -3, -4, -6, and -9, are particularly important in bone repair and regeneration.^[1,11] Although these factors affect chondrocyte and osteoblast proliferation as well as angiogenesis, their primary function is to induce differentiation of osteoprogenitor cells into functional osteoblasts. These factors exert their effects through serine/threonine kinase receptors on mesenchymal cells and pre-osteoblasts, setting off a signal transduction cascade that leads to expression of bone-specific transcription factors (e.g., Runx2/Cbfa1). These osteoblast precursor cells are then induced to become mature, bone-forming osteoblasts.

IMPLICATIONS FOR ENGINEERING TISSUE REGENERATION

Clinical Problems and Early Successes

The remarkable process of bone fracture healing portrayed in this chapter can be considered an optimal scenario, which is accurate in the large majority of fractures. However, failure of bone defects to heal (nonunion) occurs often enough to be a significant clinical problem. Bone loss can also be due to factors distinct from fracture, such as cancer resections, growth defects, and birth defects. The need for novel bone repair and replacement therapies is substantial, and continues to grow rapidly.

By understanding the healing processes, factors that affect bone healing may be exploited to promote de novo growth of bone tissue. For example, it is clear that protein growth factors are a vital component of the bone healing process, due to their effects on blood

vessel ingrowth and/or function on bone-forming cell types. These observations suggest that delivery systems should target protein growth factors to bone defect sites to improve or accelerate healing.^[14] Similarly, the influence of mechanics on bone fracture healing provides a mimetic basis for biomaterial scaffolds to support bone regeneration.

Opportunities in Mimicry of Natural Bone Healing

Despite significant progress towards directed bone regeneration, key aspects of the natural healing process have proven difficult to mimic (e.g., the timing of signaling molecules). Early secretion of growth factors related to chemotaxis and proliferation (e.g., FGFs, PDGFs) is followed by later secretion of factors that induce osteogenesis by bone precursor cells (e.g., VEGF, BMPs, TGF β). Similarly, early presence of immune cells and undifferentiated mesenchymal cells is followed by later presence of mature chondrocytes and osteoblasts. These characteristics of natural bone healing suggest that properly timed delivery of key growth factors or cells to a defect could significantly enhance healing or regeneration. The clear role of timing in proper bone healing represents an opportunity to develop approaches in which growth factor delivery, the presence of specific cell types, or the presence of blood vessels are temporally controlled.

Bone does not grow or heal properly when angiogenesis is artificially blocked.^[8,17] The importance of blood vessel ingrowth presents an opportunity to develop new approaches that promote angiogenesis in bone defects. Finally, participation of mesenchymal “stem” cells in natural bone healing suggests that these cells can be used in tissue engineering to regenerate bone (as well as other CTs) if they are properly delivered and exposed to appropriate signals.^[2] It is clear that multiple cell types derived from mesenchymal cells are involved in the early stages of bone healing and that mature bone contains multiple cell and tissue types.

Broader Implications for Regenerative Medicine

Despite the unique aspects of bone healing, there are central themes in the process that apply to healing of other CTs. The importance of precursor cells, the controlled presence of specific soluble signals, and the rapid promotion of angiogenesis are not unique to bone, or even unique to CTs. Precursor cells are a vital component of all tissue development processes and have been implicated in most regenerative processes in adult animals. Timed production and controlled concentration of inductive molecules (e.g., growth factors) is commonplace during development and regeneration of

a variety of tissue types, including neural, vascular, and cardiac tissues. In addition, angiogenesis is necessary to provide nutrients during development and regeneration of nearly all tissue types. Therefore, the lessons learned in a comprehensive analysis of bone healing apply to most CTs and to many other tissues. Approaches that mimic the complex process of bone healing using tissue engineering techniques are easily adaptable to healing of other CTs, as well as many other tissues.

CONCLUSIONS

This chapter provides a brief overview of wound healing in CTs and covers two broad topics: (1) healing stages and processes for soft CTs, and (2) the healing stages and processes for bone. The specific requirements all CT healing are discussed, including: (1) mechanical stability of the damaged tissues to initiate the healing process, (2) tissue specific and temporally staged cell and ECM signaling, (3) available cells (undifferentiated mesenchymal stem cells or tissue specific differentiated cells), (4) vascular and neural supply to support the entire process. The specific stages of soft CT healing are discussed, including: (1) hemostasis, (2) inflammation, (3) proliferation, and (4) remodeling. Also discussed are the unique processes of bone healing, including an optimized remodeling process that effectively regenerates the original tissue. Tissues that heal incompletely, or do not heal, have alterations to the “normal” processes or conditions, which in turn, can become targets for tissue engineering. But, these targets require an in-depth knowledge of conditions and signaling requirements, many of which are not completely defined. As these processes, as well as the biomimetic issues, are better elucidated, tissue engineering solutions will evolve in ways that benefit healing in all CTs.

ACKNOWLEDGMENTS

This review was funded in part by the NIH AR49266 (to RV) and NIH AR052893 (to WLM).

ARTICLES OF FURTHER INTEREST

Bioactive Glass; Biologic and Synthetic Apatites; Tissue Engineering of Bone

REFERENCES

1. Barnes, G.L.; Kostenuik, P.J.; Gerstenfeld, L.C.; Einhorn, T.A. Growth factor regulation of fracture repair. *J. Bone Miner. Res.* **1999**, *14*, 1805–1815.



2. Caplan, A.I.; Bruder, S.P. Mesenchymal stem cells: building blocks for molecular medicine in the 21st century. *Trends Mol. Med.* **2001**, *7*, 259–264.
3. Clark, R.A.F. *The Molecular and Cellular Biology of Wound Repair*, 2nd Ed., New York: Plenum, 1996; p. 611.
4. Cullinane, D.M.; Salisbury, K.T.; Alkhiary, Y.; Eisenberg, S.; Gerstenfeld, L.; Einhorn, T.A. Effects of the local mechanical environment on vertebrate tissue differentiation during repair: does repair recapitulate development? *J Exp. Biol.* **2003**, *206*, 2459–2471.
5. Daniel, D.M.; Stone, M.L.; Riehl, B. Ligament surgery, the evaluation of results. In *Knee Ligaments: Structure, Function, Injury, and Repair*; Daniel, D.M., Akeson, W.H., O'Connor, J.J., Eds.; Raven Press: New York, NY, 1990.
6. Diegelmann, R.F.; Evans, M.C. Wound healing: an overview of acute, fibrotic, and delayed healing. *Front. Biosci.* **2004**, *9*, 283–289.
7. Frank, C. Medial collateral ligament healing. A multidisciplinary assessment in rabbits. *Am. J. Sports Med.* **1983**, *11*, 379–389.
8. Gerber, H.P.; Vu, T.H.; Ryan, A.M.; Kowalski, J.; Werb, Z.; Ferrara, N. VEGF couples hypertrophic cartilage remodeling, ossification and angiogenesis during endochondral bone formation. *Nat. Med.* **1999**, *5*, 623–628.
9. Hardingham, T.; Tew, S.; Murdoch, A. Tissue engineering: chondrocytes and cartilage. *Arthritis Res.* **2002**, *4*, S63–S68.
10. Indelicato, P.A. Non-operative treatment of complete tears of the medial collateral ligament of the knee. *J. Bone Joint Surg. Am.* **1983**, *65*, 323–329.
11. Kang, Q.; Sun, M.H.; Cheng, H. Characterization of the distinct orthotopic bone-forming activity of 14 BMPs using recombinant adenovirus-mediated gene delivery. *Gene Ther.* **2004**, *11*, 1312–1320.
12. Kim, W.J.; Gittes, G.K.; Longaker, M.T. Signal transduction in wound pharmacology. *Arch. Pharm. Res.* **1998**, *21*, 487.
13. Kuznetsov, S.A.; Mankani, M.H.; Gronthos, S.; Satomura, K.; Bianco, P.; Robey, P.G. Circulating skeletal stem cells. *J. Cell Biol.* **2001**, *153*, 1133–1140.
14. Leach, J.K.; Mooney, D.J. Bone engineering by controlled delivery of osteoinductive molecules and cells. *Expert Opin. Biol. Ther.* **2004**, *4*, 1015–1027.
15. Lorenz, H.P.; Longaker, M.T. Wounds: biology, pathology, and management. In *Surgery: Basic Science and Clinical Evidence*; Norton, J.A., Bollinger, R.R., Chang, A.E., Lowry, S.F., Mulvihill, S.J., Pass, H.I., Thompson, R.W., Eds.; Springer: New York, 2000; 77–88.
16. Marieb, E.N. *Anatomy and Physiology*; Pearson/Benjamin Cummings: San Francisco, 2005.
17. Murphy, W.L.; Simmons, C.A.; Kaigler, D.; Mooney, D.J. Bone regeneration via a mineral substrate and induced angiogenesis. *J. Dent. Res.* **2004**, *83*, 204–210.
18. O'Driscoll, S.W. The healing and regeneration of articular cartilage. *J. Bone Joint Surg. Am.* **1998**, *80*, 1795–1812.
19. Park, J.E.; Barbul, A. Understanding the role of immune regulation in wound healing. *Am. J. Surg.* **2004**, *187*, 11S–16S.
20. Redd, M.J.; Cooper, L.; Wood, W.; Stramer, B.; Martin, P. Wound healing and inflammation: embryos reveal the way to perfect repair. *Trans. R. Soc. Lond.* **2004**, *359B*, 777–784.
21. Roberts, H.R.; Tabares, A.H. Overview of the coagulation reactions. In *Molecular Basis of Thrombosis and Hemostasis*; High, K.A., Roberts, H.R., Eds.; Dekker: New York, 1995; 35–50.
22. Sellers, R.; Peluso, D.; Morris, E.A. The effect of recombinant human bone morphogenic protein-2 (rhBMP-2) on the healing of full thickness defects of articular cartilage. *J. Bone Joint Surg. Am.* **1997**, *79*, 1452–1463.

Hearing Aids

Ian C. Bruce

Department of Electrical and Computer Engineering, McMaster University, Hamilton, Ontario, Canada

INTRODUCTION

A hearing aid is an electronic device for processing and amplifying sounds to compensate for hearing loss. The primary objective in hearing aid amplification is to make all speech sounds audible, without introducing any distortion or making sounds uncomfortably loud. Modern analog electronics have allowed hearing aids to achieve these goals and provide benefit to many sufferers of hearing loss. However, even if a hearing aid can reestablish audibility of speech sounds, normal speech intelligibility is often not fully restored, particularly when listening in a noisy environment. The development of digital hearing aids has opened the possibility of more sophisticated processing and amplification of sound. Speech-processing algorithms in digital hearing aids are being developed to better compensate for the degradation of speech information in the impaired ear and to detect and remove competing background noise.

NORMAL HEARING

The mammalian ear can be divided into three sections: the outer, middle, and inner ears (Fig. 1). The primary function of the outer ear is to funnel sound waves to the tympanic membrane (eardrum) in the middle ear. Vibrations of the eardrum caused by sound pressure waves are transferred to the oval window of the cochlea (inner ear) by the ossicles (articulated bones) of the middle ear. The mechanical system formed by the middle ear helps to create an acoustic impedance match, such that sound waves in the low-impedance air filling the external auditory canal are efficiently transmitted to the high-impedance fluid filling the cochlea, rather than being reflected back out of the ear.

Taking a cross section of the cochlea (Fig. 2) shows that it is divided into three compartments—the scala tympani, the scala media, and the scala vestibuli—each of which extends from the base to the apex of the cochlea. Sitting on the basilar membrane (dividing the scala tympani and scala media) is the organ of Corti, which houses the transducer cells of the cochlea, the outer and inner hair cells. The outer hair cells detect vibrations of the basilar membrane and produce amplification of those vibrations through a mechanism

known as electromotility.^[3] The inner hair cells synapse onto the auditory nerve fibers, which project to the auditory brainstem.

The transduction process from sound waves into neural impulses in the auditory nerve is illustrated in Fig. 3A. Motion of the oval window caused by middle ear movement creates a pressure difference across the basilar membrane. This pressure difference generates a wave that travels along the cochlea from the base toward the apex, causing displacement of the cilia of hair cells. Displacement of the cilia allows ionic currents to flow into a hair cell, modifying the receptor potential (i.e., the hair cell's transmembrane potential). In inner hair cells, this releases neurotransmitters to initiate neural impulses (discharges) in auditory nerve fibers, thus providing information to the brain about the sound vibrations received at the ear.

The mechanics of the basilar membrane are such that high-frequency sounds produce traveling waves that peak near the base of the cochlea and low-frequency sounds peak near the apex, as illustrated in Fig. 3B. Because each auditory nerve fiber synapses onto only one inner hair cell, each auditory nerve fiber encodes information about the basilar membrane vibration at one position in the cochlea. Consequently, auditory nerve fibers inherit the frequency-tuning properties of the cochlea, and each auditory nerve fiber has a best (or characteristic) frequency to which it responds.

Following from the transduction process described above, the auditory periphery is often regarded as a bank of band-pass filters. However, the electromotile behavior of healthy outer hair cells introduces important nonlinearities into the basilar membrane vibrations and the subsequent representation of sounds by auditory nerve fiber discharges. The predominant nonlinearity is compression, in which the basilar membrane vibrations do not grow linearly with sound intensity but rather are compressed for sound pressure levels (SPLs) between approximately 30 and 90 dB SPL.^[5] A related cochlear nonlinearity is suppression, in which the response to a sound frequency component is reduced because of the presence of a more intense component at a nearby frequency.^[5] Further nonlinearities are introduced by the inner hair cell and its synapse with an auditory nerve fiber. Auditory nerve responses exhibit adaptation that accentuates the



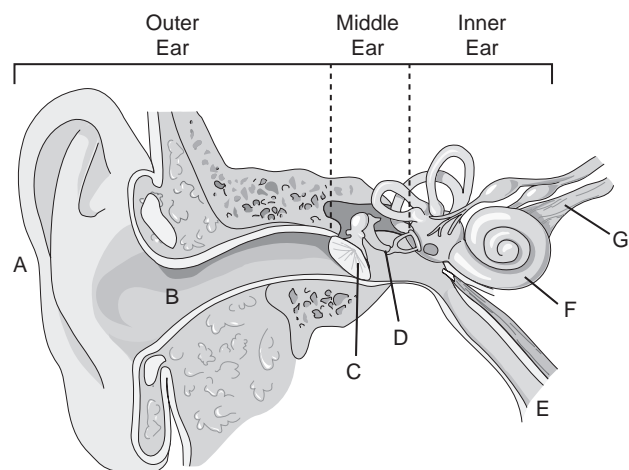


Fig. 1 Anatomy of the human ear, showing the division of the outer, middle, and inner ears. (A) pinna; (B) external auditory canal; (C) tympanic membrane (eardrum); (D) ossicles; (E) eustachian tube; (F) cochlea; and (G) auditory (cochlear) nerve. (From Ref. [1].)

onsets and offsets of sound components, and rectification and saturation are observed in auditory nerve discharge rates. This collection of nonlinearities is important in forming the neural representation of speech sounds.^[4,6]

HEARING LOSS

The two most common forms of hearing loss are conductive loss, which refers to dysfunction of the outer or middle ear, and sensorineural loss, in which

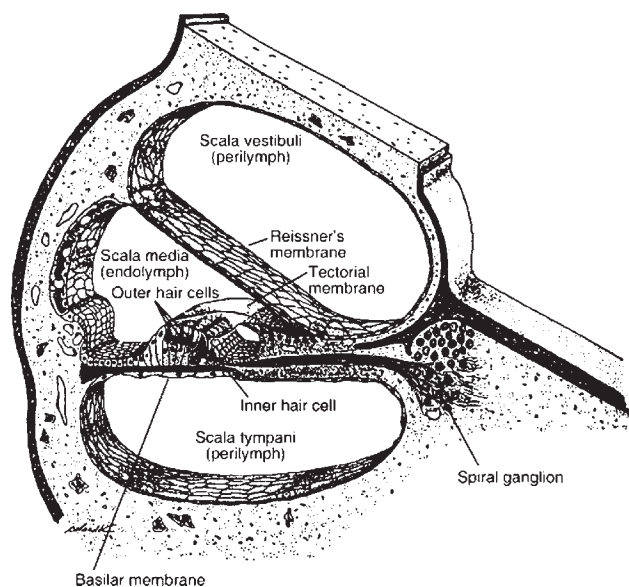


Fig. 2 A cross section of the human cochlea. (From Ref. [2].)

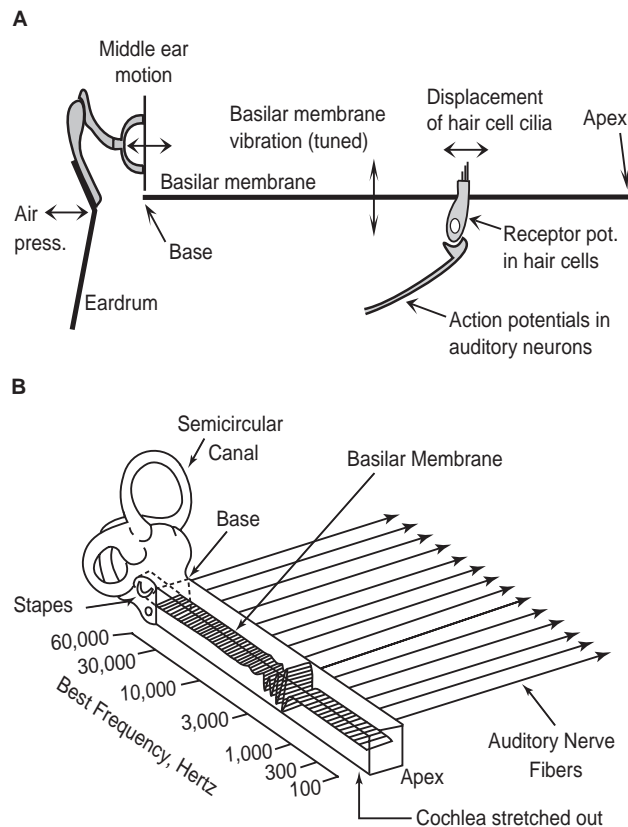


Fig. 3 (A) Schematic summarizing the transduction process in the mammalian cochlea, and (B) Illustration of the tonotopic organization of the cochlea. (From Ref. [4].)

the inner ear or the auditory pathways of the brain are impaired. The outer and middle ears are largely linear systems, and consequently, conductive dysfunction typically produces a simple linear reduction in the amplitude of acoustic signals as they are transmitted to the inner ear. Thus, loud sounds become quieter and quiet sounds may become inaudible. In contrast, the effects of sensorineural loss are more multifaceted. Because the cochlea is a highly nonlinear system, any impairment of cochlear structures (particularly the inner and outer hair cells) can lead to substantial distortion in the neural representation of acoustic signals, in addition to loss of audibility. Common causes of sensorineural impairment include exposure to loud sounds, aging, disease, head injury, and ototoxic drugs. At present, there are no medical or surgical cures for most forms of sensorineural hearing loss.

The loss of audibility of quiet sounds, either from conductive or sensorineural loss, is quantified by hearing thresholds. The hearing threshold is a measure of the SPL required for a tone of a particular frequency to be just audible to a listener. Hearing thresholds are computed using a decibel scale relative to the hearing thresholds of normal hearing listeners, referred to as decibel hearing level (dB HL). A pure-tone audiogram

is a plot of hearing thresholds as a function of tone frequency with -10 dB HL at the top and 110 dB HL at the bottom. An example audiogram is given in Fig. 4. By convention, the hearing thresholds for the left ear (LE) are plotted as Xs and those for the right ear (RE) as Os.

The degree of hearing loss is often categorized according to the pure tone average (PTA), which is the average of the hearing thresholds at 500 , 1000 , and 2000 Hz, the most important frequencies for understanding speech. A typical classification scheme for adults is:

0–25 dB HL	Normal limits
25–40 dB HL	Mild loss
40–55 dB HL	Moderate loss
55–70 dB HL	Moderate-to-severe loss
70–90 dB HL	Severe loss
90+ dB HL	Profound loss

Alternative classification schemes are also used, particularly for categorizing hearing impairment in children. The example audiogram given in Fig. 4 shows a bilateral, asymmetrical hearing loss, with a mild loss in the LE (PTA = 33 dB HL) and a moderate-to-severe high-frequency loss in the RE (PTA = 61 dB HL). The elevation of hearing thresholds with sensorineural loss can be well explained by the loss of auditory nerve

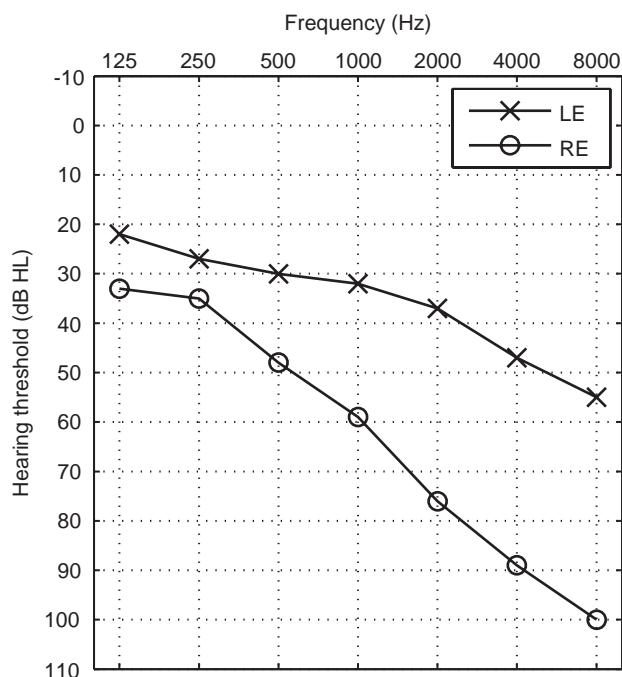


Fig. 4 An example audiogram showing hearing thresholds for the left ear (LE) and right ear (RE) of an individual with hearing loss.

sensitivity with dysfunction of the inner and outer hair cells.^[4,6]

There are several other aspects of sensorineural hearing impairment that are not captured by the audiogram, including reduction in the dynamic range of hearing, decreased frequency resolution, and decreased temporal resolution. Many of these perceptual impairments can be explained from the direct effects of damaged inner and outer hair cells on the transduction process in the auditory periphery, but the effects of sensorineural hearing loss on the central auditory pathways of the brain should also be taken into consideration.

The reduction in dynamic range for a hearing impaired listener is often referred to as “loudness recruitment.” In this phenomenon, the audibility threshold is elevated at a frequency suffering from hearing loss, but the intensity at which a tone at that frequency becomes uncomfortably loud is relatively normal. That is, there is an abnormally steep growth of loudness with stimulus intensity. This causes difficulties for prescribing gain in a linear amplification scheme for hearing aids: If quiet sounds are amplified enough to restore audibility, then loud sounds may be amplified to uncomfortably loud levels. It has long been presumed that loudness recruitment is owing to the more rapid growth of auditory nerve discharge rates with sound level because of loss of cochlear compression. However, recent physiological data contradict this theory and suggest that central mechanisms are required to explain loudness recruitment.^[7]

Decreased frequency and temporal resolution and reduced nonlinearities (such as compression and suppression) in an impaired cochlea lead to distortions in the neural representation of a single sound^[4,6,8] and to difficulties in separating out different sound sources in the complex acoustic “soundscape” in which we operate daily. For example, it has been shown that after loss of audibility is accounted for, individuals with hearing loss still require a higher signal-to-noise ratio (SNR) to understand speech with the same accuracy as those without hearing loss.^[9] Linear amplification schemes cannot compensate for such distortions; more sophisticated speech-processing algorithms will be required.

HISTORY OF HEARING AIDS

The first devices for amplifying speech were acoustic ear trumpets, a range of which was available by the 1600s. These used a tapered horn to collect sounds and funnel them to the ear. In addition to providing some amplification, ear trumpets could increase the SNR if the opening of the trumpet was pointed toward the talker. Around 1900, carbon hearing aids were



developed that consisted of a carbon microphone, a battery, and a magnetic receiver (a miniature loud-speaker), which could produce a gain of around 20 to 30 dB. A pair of carbon hearing aids joined in series, with the coupled magnetic receiver and carbon microphone referred to as a carbon amplifier, could generate even greater gain. In 1920, the carbon amplifier was replaced by a vacuum tube electronic amplifier, which had superior amplification capabilities, but the vacuum tubes required two large batteries to operate. It was not until 1944 that vacuum tubes and batteries became small enough to manufacture a one-piece body-worn hearing aid. During this time, carbon microphones were superseded by magnetic and piezoelectric microphones.

Immediately following the invention of the transistor, head-worn hearing aids using transistor amplifiers instead of vacuum tubes (and consequently smaller batteries) were introduced in the early 1950s. The behind-the-ear (BTE) hearing aid (as shown in the top panel of Fig. 5) was the dominant style for several decades. The invention of the integrated circuit and the electret microphone in the 1960s and 1970s lead to further miniaturization of hearing aid components and the increasing use of in-the-ear (ITE) style hearing aids

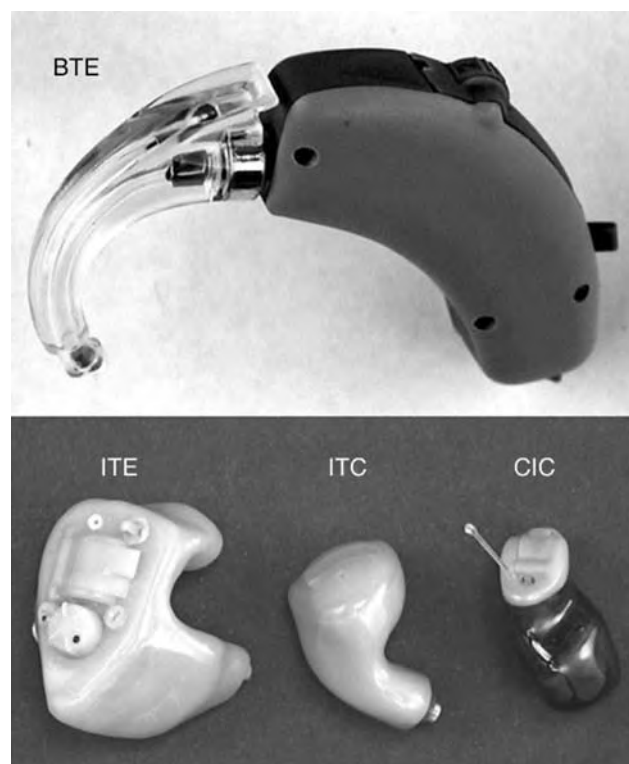


Fig. 5 Modern hearing aid styles: behind-the-ear (BTE); in-the-ear (ITE); in-the-canal (ITC); and completely-in-the-canal (CIC) Not shown is the earmold that connects to the BTE ear hook via a tube to deliver amplified sounds to the ear canal. (From Ref. [10].)

(shown in the bottom panel of Fig. 5). Continued reductions in component sizes lead to the introduction of smaller in-the-canal (ITC) and completely-in-the-canal (CIC) hearing aids (shown in Fig. 5) in the 1980s and 1990s. In the same time period, steady advances in the electronics of microphones, amplifiers, and receivers produced improvements in the fidelity of the audio processing by analog hearing aids. Recently, analog amplifiers have been replaced by digital signal processor (DSP) chips, allowing much more sophisticated amplification and processing of sounds by hearing aids.

MODERN HEARING AID STYLES

Current hearing aids are available in a range of styles from BTE to CIC. The typical locations of components in an ITC and a BTE hearing aid are illustrated in Fig. 6. The BTE aid delivers the amplified sound to the ear through an earmold (or earshell), which connects to the BTE ear hook via a sound tube.

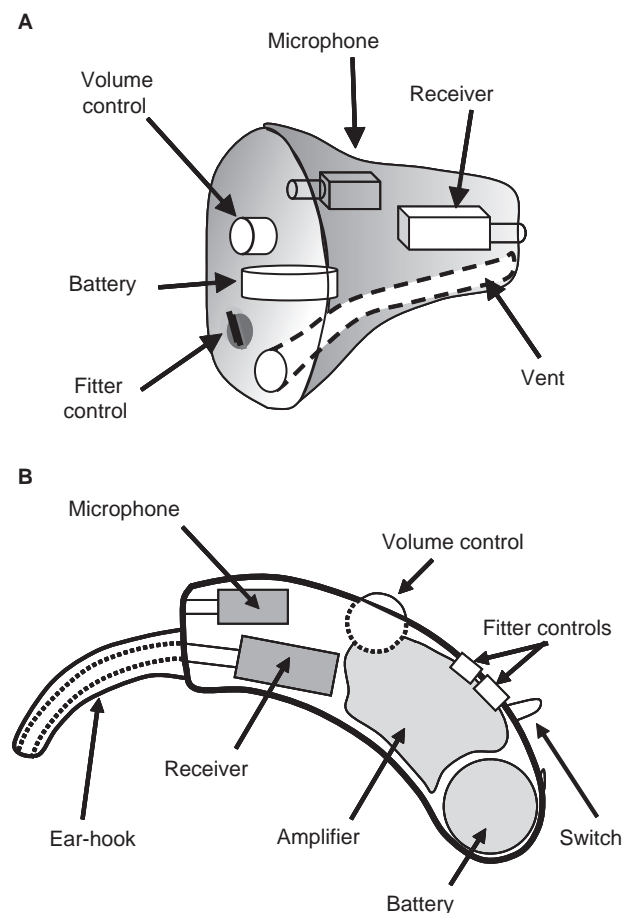


Fig. 6 Typical component placement in an ITE A) and BTE B) hearing aid. Note that the amplifier is not shown in the ITE diagram. (From Ref. [10].)

Molds or shells are typically custom-made from an impression of the hearing aid user's ear, which leads to a comfortable fit in the ear canal and minimization of sound leakage around the mold or shell. Air vents are desirable in an ITE shell or BTE earmold to prevent moisture build-up in the residual canal (i.e., the air between the hearing aid and the ear drum) and to reduce the occlusion effect, in which low-frequency bone-conducted sounds (particularly the hearing aid user's own speech) resonate in the residual canal. However, venting provides an additional pathway for sounds to enter the ear and for amplified sounds coming from the receiver to exit the ear. This alters the effective frequency response of the hearing aid and provides an additional pathway for acoustic feedback from the receiver to the microphone, which can lead to feedback oscillations (whistling).

The smaller devices are less conspicuous and consequently are more cosmetically appealing to many hearing aid users. However, the larger devices can typically provide superior amplification and sound processing capabilities, because more powerful electronics can be packaged in the hearing aid. Another advantage of the larger styles is the greater physical separation between the receiver sound outlet and the microphone, which may help reduce acoustic feedback oscillations.

LINEAR AMPLIFICATION, COMPRESSION, AND FEEDBACK SUPPRESSION

The simplest approach to amplification in hearing aids is to apply a linear gain at each frequency that is dependent on the hearing loss at that frequency. As a rule of thumb, a linear gain of 75% of any conductive hearing loss and around 50% of any sensorineural loss should be applied. These general rules have been refined greatly through experimental testing to give a number of different amplification prescription formulas, which audiologists use as a starting point when fitting hearing aids. Recent results with a computational model of the auditory periphery suggest that these fitting rules are effectively restoring the average auditory nerve response to speech.^[11]

However, linear amplification may not be optimal, because of distortions that can be produced by waveform peak clipping in the amplifier or receiver at high input sound levels and because of loudness recruitment in an ear with sensorineural impairment. Reduction of the amplifier gain as a function of the input or output signal level via a fast-acting automatic gain control (AGC) is referred to as compression; slow-acting AGC can be used as an automatic volume control. Peak clipping at high levels can be avoided via compression limiting, in which the gain is reduced so that the output

level does not exceed some specified limit. Compression that acts over the normal range of speech sound levels is referred to as wide-dynamic-range compression (WDRC). An example of an input/output curve for a compression system is shown in Fig. 7A, with the gain vs. input level shown in panel B. Here, a gain of 30 dB is applied for all input signals that are lower than 40 dB SPL, the knee-point (threshold) of the WDRC. Above 40 dB SPL, the gain of the amplifier is reduced by 10 dB for each 20 dB increase in the input level, corresponding to a compression ratio of 2:1. At an input level of 100 dB SPL, the gain has consequently been reduced to 0 dB, i.e., no

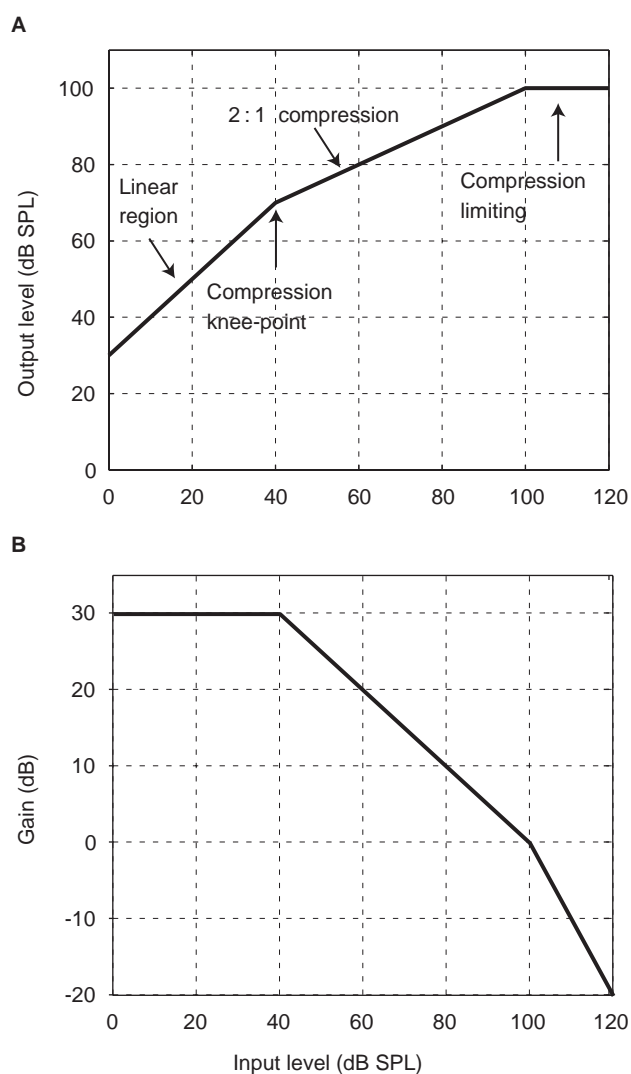


Fig. 7 Example static characteristics of a compression system. (A) Input/output curve for a system with wide-dynamic-range compression (WDRC) with a 2:1 compression ratio for input levels above 40 dB SPL and compression limiting at an output level of 100 dB SPL, and (B) The amplifier gain vs. input-level function required to achieve the compression characteristics described in panel A.

amplification is applied. Compression limiting is applied for input levels above 100 dB SPL, so that the output level cannot exceed 100 dB SPL.

Compression can be applied identically to all frequencies, referred to as single-channel compression, or individually to different frequency bands, referred to as multiband compression. The motivation for multiband compression is that the amount of hearing loss may vary substantially between different frequency bands, and thus the desired compression characteristics may differ between frequency bands. However, multiband compression will tend to flatten out the frequency spectra of sounds, possibly making it harder to distinguish different sounds. Some multiband compression schemes have been developed to try to reduce spectral flattening.^[12,13]

In addition to the static compression characteristics described by the input/output curve, the rate at which the gain is adjusted with changes in the input (or output) signal level determines the dynamic characteristics of a compression system. The time it takes for a compressor to react to an increase in the signal level is referred to as the attack time, and the release time corresponds to the time taken to react to a decrease in signal level. By adjusting the attack and release times, as well as the static compression characteristics, it is possible for compression systems to avoid distorted and uncomfortably loud signals, to reduce the intensity differences between phonemes or syllables, to provide automatic volume control, to increase sound comfort, to normalize loudness, to maximize intelligibility, or to reduce background noise.^[10] Unfortunately, the required compression parameters are very different for each of these goals, and consequently any one compression scheme tends to provide benefit in some but not all aspects of compensating for hearing impairment.

A further use of AGC is to suppress acoustic feedback oscillations, if they can be reliably detected. Reduction of feedback oscillations can also be achieved via automatic phase control, via an internal (i.e., electronic) feedback path to cancel the external (i.e., unintentional acoustic) feedback path, or via frequency shifting by the amplifier to prevent the build-up of oscillations at a particular frequency. With the advent of digital hearing aids, more effective feedback suppression has been obtainable. This has permitted larger vents in earmolds and shells, reducing occlusion effects and providing more comfort to hearing aid users.

ADVANCED PROCESSING ALGORITHMS

To this date, no amplification and compression scheme has been found that can fully compensate for the

effects of sensorineural impairment in individuals with moderate-to-profound hearing loss. Consequently, a number of signal-processing algorithms have been developed to make the task of listening easier for hearing aid users. Although some of these processing strategies could be implemented with analog electronics, the processing power and flexibility of DSPs have greatly increased the number and effectiveness of advanced processing algorithms that can be realized in hearing aids.

One class of algorithm now widely used in hearing aids is referred to as single-microphone noise reduction. These algorithms attempt to detect and separate the components of the signal received by the hearing aid microphone that can be attributed to a talker (to whom the listener wishes to attend) and to “background noise” (which the listener wishes to ignore). Single-microphone noise reduction typically works well when the background noise sources have acoustic properties that are very different from speech and do not vary much over time, but these algorithms have difficulty when the noise sources are speech-like or vary rapidly, as is the case in most real-world listening conditions. A further problem with noise reduction is removing the noise components from a signal without degrading or distorting the target speech components, which could lead to loss of speech intelligibility. Consequently, single-microphone noise reduction algorithms in hearing aids are typically aimed at improving listening ease and comfort, rather than increasing speech intelligibility.

However, noise reduction that leads to intelligibility improvements can be obtained through the use of directional microphones and multiple-microphone algorithms known as beamformers. In this class of algorithms, sounds arriving from a desired direction (typically the direction the hearing aid user is facing) pass through unaltered, but sounds arriving from other directions are attenuated. Some hearing aids utilize microphones with fixed directionality patterns, while others electronically combine signals from multiple microphones such that the directionality can be switched. Adaptive beamformers attempt to detect noise sources and dynamically vary the directionality pattern to cancel them out.

For hearing aid users with very sharply sloping high-frequency losses, transposition of high-frequency components to lower frequencies can provide some improved speech clarity and intelligibility. However, only moderate shifts in frequency are possible, or the frequency spectra of sounds will become too distorted. An added benefit of frequency transposition is the reduction in acoustic feedback oscillations, as mentioned previously.

In various listening environments, different amplification, compression, noise reduction, and directionality

schemes may be required. Even before the introduction of fully digital hearing aids, digital memories for analog hearing aid parameters enabled users to manually switch between different parameter sets (referred to as “programs”) in different listening situations. Some hearing aids now use acoustic classification algorithms to detect the listening environment and automatically switch to the preferred program for that situation. The accuracy of algorithms in implementing automatic volume control and feedback suppression has meant that many hearing aids are now available without a manual volume control.

HEARING AID USAGE

Hearing aids can be a viable form of treatment for individuals with any degree of hearing loss, although those with profound losses are good candidates for cochlear implants. However, the audiogram itself is typically not a good guide to how much benefit will be gained from hearing aid use. In some cases, individuals with hearing thresholds small enough to be classified as “normal” may benefit from hearing aids if they experience hearing difficulties in everyday life. The prevalence of hearing loss in most developed countries is reported to be around 10% of the population. However, regular hearing aid usage is often as low as 20% in the hearing-impaired population. In addition, many cases of mild-to-moderate loss may be undiagnosed or not acknowledged by the sufferer. Overall user satisfaction with hearing aids remained around 59% from 1991 to 2001.^[14] While problems such as hearing aid fit, comfort, and price contribute to customer dissatisfaction, the major concern is lack of benefit in understanding speech, particularly in noisy listening environments. This indicates that there is substantial room for improvement in amplification and processing of speech by hearing aids.

FUTURE DIRECTIONS IN HEARING AIDS

Because linear amplification and conventional compression schemes do not appear to fully compensate for the effects of sensorineural hearing loss in many hearing aid users, there is growing interest in modeling the effects of cochlear impairment to develop sophisticated nonlinear amplification algorithms that maximally restore the neural representation of sounds to normal.^[4,6,8,11,13,15,16] The difficulties in applying this approach include how accurate the model of the ear needs to be, what amplification schemes should be considered, and how impairment of the neural representation should be measured. Simplified

cochlear models can lead to more straightforward amplification schemes, but important features of the normal and impaired neural representation of sounds may be neglected. On the other hand, a more detailed neural representation can lead to problems with choosing a suitable amplification scheme and a suitable metric of impairment. The latter difficulty has led to several investigations into how well models of the ear and subsequent neural processing can predict human speech intelligibility.^[17,18] As we gain a greater understanding of how the ear and brain process speech in normal and impaired cases, we will better know what approaches to amplification can optimally compensate for a particular hearing impairment.

Many hearing aid users have a bilateral loss and consequently wear a hearing aid in each ear. Hearing aid manufacturers are now developing wireless data links between digital hearing aids to facilitate coordinated processing by the devices in each ear. At low data-transfer rates, this will enable coordinated device programming and synchronized program switching. As higher data rates become feasible, it will be possible to transmit audio signals between hearing aids to perform beamforming and other binaural signal processing based on the signals received by the microphones in each ear. A high-rate wireless data system will also permit transmission of high-quality audio signals to a hearing aid from other audio devices such as telephones, televisions, car and home stereos, public address systems, and MP3 players.

CONCLUSIONS

Hearing loss can be a debilitating impairment in daily life, greatly affecting social interaction and making many work situations difficult. Until a cure for sensorineural hearing loss is discovered, hearing aids remain the primary treatment option for the hearing-impaired population. Continued innovations over more than a century have led to modern hearing aids that can provide benefit for many hearing-impaired individuals. However, hearing aids still cannot fully compensate for the effects of hearing loss to the same degree that, for example, eye glasses can compensate for the most common forms of visual impairment. Further improvements in amplification and speech processing by hearing aids are under investigation, while at the same time medical treatments for hearing loss are being pursued.

ARTICLE OF FURTHER INTEREST

Tinnitus Devices



REFERENCES

1. Clark, G.M. *Cochlear Implants: Fundamentals and Applications*; Springer-Verlag: New York, NY, 2003.
2. Nolte, J. *The Human Brain: An Introduction to its Functional Anatomy*, 3rd Ed.; Mosby: St. Louis, MO, 1993.
3. Brownell, W.E.; Bader, C.R.; Bertrand, D.; de Ribaupierre, Y. Evoked mechanical responses of isolated cochlear outer hair cells. *Science* **1985**, *227* (4683), 194–196.
4. Sachs, M.B.; Bruce, I.C.; Miller, R.L.; Young, E.D. Biological basis of hearing-aid design. *Ann. Biomed. Eng.* **2002**, *30*, 157–168.
5. Robles, L.; Ruggero, M.A. Mechanics of the mammalian cochlea. *Physiol. Rev.* **2001**, *81*, 1305–1352.
6. Bruce, I.C.; Sachs, M.B.; Young, E.D. An auditory-periphery model of the effects of acoustic trauma on auditory nerve responses. *J. Acoust. Soc. Am.* **2003**, *113* (1), 369–388.
7. Heinz, M.G.; Issa, J.B.; Young, E.D. Auditory-nerve rate responses are inconsistent with common hypotheses for the neural correlates of loudness recruitment. *JARO* **2005**, *6* (2), 91–105.
8. Bondy, J.; Bruce, I.C.; Dong, R.; Becker, S.; Haykin, S. Modeling intelligibility of hearing-aid compression circuits. *Conference Records of the Thirty-Seventh Asilomar Conference on Signals, Systems and Computers*. IEEE Press: Piscataway, NJ, 2003; Vol. 1, 720–724.
9. Plomp, R.; Duquesnoy, A.J. A model for the speech-reception threshold in noise without and with a hearing aid. *Scand. Audiol. Suppl.* **1982**, *15*, 95–111.
10. Dillon, H. *Hearing Aids*; Thieme Medical Publishers: New York, NY, 2001.
11. Bondy, J.; Becker, S.; Bruce, I.C.; Trainor, L.J.; Haykin, S. A novel signal-processing strategy for hearing-aid design: neurocompensation. *Signal Proc.* **2004**, *84* (7), 1239–1253.
12. White, M.W. Compression systems for hearing aids and cochlear prostheses. *J. Rehabil. Res. Dev.* **1986**, *23* (1), 25–39.
13. Bruce, I.C. Physiological assessment of contrast-enhancing frequency shaping and multiband compression in hearing aids. *Physiol. Meas.* **2004**, *25*, 945–956.
14. Kochkin, S. MarkeTrak VI: 10-year customer satisfaction trends in the US hearing instrument market. *Hearing Rev.* **2002**, *9* (10), 14–25.
15. Kates, J. Toward a theory of optimal hearing aid processing. *J. Rehab. Res.* **1993**, *30* (1), 39–48.
16. Chabries, D.M.; Anderson, D.V.; Stockham, T.G., Jr.; Christiansen, R.W. Application of a human auditory model to loudness perception and hearing compensation, *Proceedings of the 1995 International Conference on Acoustics, Speech, and Signal Processing (ICASSP-95)*; IEEE Press: Piscataway, NJ, 1995; Vol. 5, 3527–3530.
17. Bondy, J.; Bruce, I.C.; Becker, S.; Haykin, S. Predicting speech intelligibility from a population of neurons. In *Advances in Neural Information Processing Systems 16*, NIPS 2003 Conference Proceedings; Thrun, S., Saul, L., Schölkopf, B., Eds.; MIT Press: Cambridge, MA, 2004; 1409–1416.
18. Elhilali, M.; Chi, T.; Shamma, S.A. A spectro-temporal modulation index (STMI) for assessment of speech intelligibility. *Speech Commun.* **2003**, *41* (2–3), 331–348.

BIBLIOGRAPHY

- Dillon, H. *Hearing Aids*; Thieme: New York, NY, 2001.
- Valente, M.; Hosford-Dunn, H.; Roeser, R.J. *Audiology: Treatment*; Thieme: New York, NY, 2000.

Hearing Mechanisms

Martin L. Lenhardt

Biomedical Engineering, Otolaryngology and Emergency Medicine, Virginia Commonwealth University, Richmond, Virginia, U.S.A.

INTRODUCTION

Helen Keller felt her deafness was far more of a handicap than her blindness in that deafness resulted in her social isolation. As the population ages, hearing loss, if not adequately remediated, will continue to socially isolate by impairing communicative interaction. Noise exposure is clearly a contributing factor to hearing loss, but so are ototoxic drugs, diseases, unhealthy lifestyles, and aging. Hearing aids offer the hope of enhancing social communication by providing appropriate level of amplification to the damaged ear. Simple amplification has been only partially successful, since only about 20 percent of the people who could benefit from aids actually use them. It was commonly accepted dogma that the rejection of hearing aids by the general population was attributed to cosmetics—too big, too bulky, too visible—but the true culprit was less-than-expected performance, more specifically, poor performance in ambient or reverberant noise when communicative demands were the highest. Microelectronics and miniaturization are ushering in instruments that are small with excellent performance. Design inspirations for innovation in size and performance have come from unexpected sources, one of which is from how unique animals have used sound to survive and prosper for hundreds of thousands of years. State-of-the-art digital acoustical processing intermixed with biologically inspired designs are beginning to propel hearing aids from something in your grandmother's drawer to vogue, wearable computing devices.

OVERVIEW

Biologically inspired approaches are providing new insights into amplification. The hearing aid is now becoming more than an amplifier; it is a digital device to preprocess speech before it is actually fed into the ear. Inspiration is all around us. For example, a small fly is a source for innovation in hearing aid microphone design. The fly, *Ormia ochracea*, can localize sound with about the same precision as humans using two small eardrums that are so close they almost touch. These eardrums, unlike those in humans, are

torsionally coupled or uncoupled such that the fly obtains information on the average sound displacement impinging on the eardrums as well as the difference. Differential coupling between two different modes of resonance of each drum provides the localization cues. The fly, sensitive to acoustic particle displacement, has only about 100 neurons for coding auditory information, contrasted to 30,000 in normal hearing humans. Neural thresholds vary from 55 to 95 dB sound pressure level (SPL); these thresholds are characteristic of moderate to severe deafness in human terms. Using silicon MEMS technology, a 1- by 2-mm silicon nitride microphone diaphragm has been developed to emulate the fly's hearing system,^[1] which, when connected in an array, can potentially advance hearing aid performance in noise, a very real human problem in understanding speech.

It is important to recognize that human auditory experience is very much different from that of all other animals in that much of our brain is specialized for processing speech and language. Acoustic detection and processing strategies, honed by selective environmental pressure to survive and reproduce, are variations and deviations from hearing mechanics evolved over millions of years. Vertebrates, especially those with restricted hearing by human terms, may offer novel biologically-based hearing aid designs. It is helpful at this point to define hearing in a universal vertebrate sense.

Hearing is the detection of far-field pressure waves, but also near-field particle movement as well as vibratory displacement. The receptor must, by definition of acoustic or auditory, be specialized and most sensitive to sound/vibration. So mechanical disturbances that are either distant (pressure), close (near field), or very close (substrate vibration) and that act on a specialized receptor can result in hearing. It is not helpful to define the ear as solely a hearing organ since the inner ear contains sensors that are stimulated by gravity, orientation, acceleration, and vibration. It is this multiplicity and redundancy in inner-ear function that offers hope for listeners with hearing impairment.

Fish have inner ears with four receptors that can detect sound. The macula neglecta is sensitive to near-field displacement, the saccule or utricle (in the hearing family) is a specialized sound detector, and some



fish have a lagena with adaptation for sound pressure detection. Excluded from the list is the lateral line sensitive to bulk fluid displacement and the variety of tactile receptors that could be activated by vibration. Why are there multiple sound detectors? Part of the answer lies in the very conservative feature of the sound sensors in the ear, the hair cell. Hair cells from fish, amphibians, reptiles, birds, or humans are very similar. The mechanical coupling is quite different; however, if enough energy is applied to the ear, in the proper way, some receptor will be driven. This leads to the obvious question: Do humans have more receptors in their ears that can be recruited into the hearing process? Answering comes by way of two examples from studies of deafened animals. Destroying the cochlea in guinea pigs does not create total deafness. The saccule, a vestibular detector of linear acceleration in the z plane, responds to sound and its presence is uncovered with cochlear destruction. The saccule responds to a wide range of sounds at levels exceeding the loss due to poor bodily coupling to airborne sound.^[2] The human saccule responds best to frequencies between 200 and 400 Hz at about 100 dB.^[3] The saccule has the frequency sensitivity to code some aspects of speech as the envelope, prosody, and possibly the fundamental and first harmonic, so why are there not saccular hearing aids?

Pigeons with ablated cochleae too can maintain hearing if a new window is created in one of the semicircular canal walls. The semicircular canals also contain hair cell receptors specialized for rotation detection. The window allows bulk fluid movement induced by sound. Since the fluid is incompressible, creating a new window or fenestration is needed to allow cilia displacement. The fenestration of the canal wall in deaf humans produces near normal hearing. Postoperatively, when stimulated by bone conduction, all patients with this procedure had only mild hearing losses (35–40 dB HL). These patients quickly habituated to initial dizziness, but had very poor speech discrimination.^[4] Clearly, the canal nerves are not wired to the auditory areas of the brain as with the saccule, but the restoration of hearing is most impressive.

In canal fenestration, sound, when presented to the skull as bone conduction, was the most effective mode of hearing. Bone conduction is a phylogenetic old means of hearing that has fallen into disuse in the hearing industry over the last 30 years. In the mid-1960s, almost half the hearing aids manufactured were bone conduction devices incorporated into glasses frames. Performance was not the issue, but as frames became smaller and lighter, they were less cosmetically adaptable as amplification platforms. Bone conduction hearing aids are still produced for middle-ear abnormalities and a few glasses configurations are also

available. One obvious advantage of bone conduction glasses is the application of microphone arrays for better processing of speech in noise. Certainly a saccular hearing aid is a possibility and the most effective sound route could be by bone conduction. The bone transducer is usually mounted to the skin by a band, but it can also be placed on a titanium screw permanently affixed into the skull. Bone-anchored hearing devices are commercially available and provide an alternative to skin placement.

Hearing on land started with bone conduction. When vertebrates moved from an aquatic habitat to a terrestrial one, the only option for hearing available was detecting substrate vibration. Vibrations were conducted through the bones of the body to the inner ear, i.e., bone conduction. The mechanism of hearing by bone conduction is experienced by all of us when we couple our heads to a vibrating surface. The process of hearing is different hearing by conventional air conduction. The head is first set into vibration and the bones (ossicles) in the middle ear lag behind due to inertia. The resulting phase difference between the head and ossicles results in mechanical stimulation of the cochlea. The eardrum, in bone conduction hearing, acts as a damper. All primitive reptiles, including the extinct forms, likely heard by bone conduction.

Reptiles in the mammalian line of evolution were small, thin-boned, and nocturnal. There was selective environmental pressure to detect the other world of sound in the much less dense medium of air. Mammals evolved air conduction hearing through a series of processes that turned the output of bone conduction hearing, the eardrum, into a sensitive input for aerial sound pressure. Driving aerial hearing was the need to detect species-specific vocalizations and localize them in space at a distance. Hearing evolved into a distance sensor in mammals and birds. The central auditory nervous system evolved as well into parallel neural pathways to separate and selectively attend to one signal embedded in ambient noise. Hearing takes place in our brains not in our ears. So the phylogenetic story of hearing is a change from bone conduction to air conduction and a progressive expansion of sound sensitivity at ever higher frequencies. Interestingly, the natural history of human hearing loss is a gradual collapsing of both sensitivity and frequency range. So losing hearing is much like sliding back down the evolutionary ladder.

THE EAR

The middle ear is a remnant of when vertebrates made the transition from water to land. Stem reptiles, the presumed common ancestor to present-day reptiles, birds, and mammals, had massive middle-ear bones.

Adding mass to an inertial bone conduction system improved substrate detections but rendered these animals effectively deaf to airborne sound. There were survival and reproductive advantages for reptiles to detect sounds in the air. This was accomplished by a gradual thinning and stiffening of the eardrum and lightening of the ossicle. It was the relative size of the drum in reference to the cochlear window (oval) that allowed for adequate impedance matching of vibration in air with that of the cochlear fluid/basilar membrane. Humans recover approximately 35 dB of the sound that would be lost due to the impedance mismatch. The middle-ear is not, however, a perfect transformer and is most effective at or near its resonance and much less effective above that point. Middle-ear resonance can be modeled as a Helmholtz resonator. The middle ear and the ear canal have been shown by modeling and direct measurements to possess a coupled resonance of about 2.9 kHz. The frequency response of the canal alone and the middle ear are depicted in Fig. 1. Note that the middle ear acts as a low-pass filter, attenuating frequencies above resonance at about 27 dB per octave. A similar high-frequency slope is seen in lizards, birds, and some mammals. Hearing aids, when coupled to the canal, eliminate the canal resonance gain and generally have a similar output-frequency slope. The middle ear attenuates high-frequency consonant sounds in reference to low-frequency sound and limits high-frequency emphasis hearing aid gain. Looking ahead, it is the high-frequency sounds that are the salient speech cues in regard to speech intelligibility, and it is these speech sounds that are attenuated by a low-pass middle-ear filter and sensory loss.

Resting on the basilar membrane in the inner ear are two types of sensory hair cells (see Fig. 3 in the article

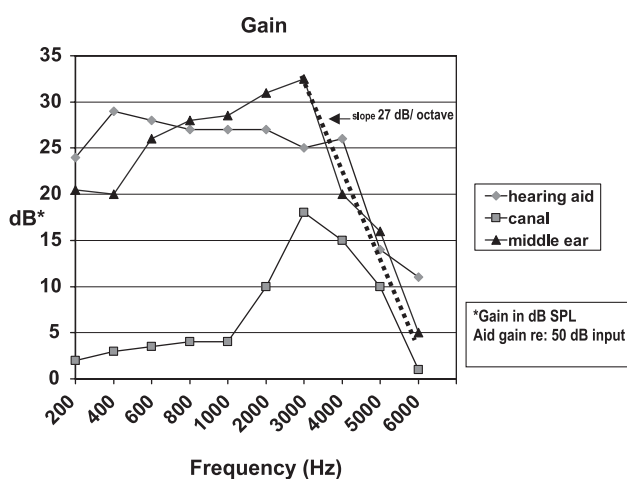


Fig. 1 The frequency responses of the human ear canal and middle ear are contrasted with that of a typical hearing aid with a 50-dB SPL input. Note all act as effective low-pass filters.

Inner Ear Implants, p. 849), outer (OHCs), and inner (IHCs). The IHCs are sensory, but the OHCs are sensory and motor. OHCs change their volumes when stimulated by the brain, becoming active cochlear tuners/amplifiers. The basilar membrane is nonlinear in its vibratory response as a result of the contractive action of OHCs.^[5] That is, OHCs change the BM net stiffness, resulting in mechanical amplification to the IHCs. Without the motor action of OHCs, IHCs would respond only over a narrow intensity range. The OHCs allow for an expansion of sound dynamic range through compression and, conversely, a loss of OHCs produces hearing loss and loss of the range compression. IHCs respond over the range of about 70 dB (50 to 120 dB). IHCs in conjunction with OHC, respond over a range of 120 dB, that is, have a gain of 120 dB. Thus, OHCs compress the sounds coming into the ear to the IHCs, allowing an expanded dynamic range of about 110 dB, i.e., from just detectable to almost painful. Irreversible OHC loss is the first sign of cochlear injury resulting in hearing loss. In the normal cochlea, the OHCs control the growth of loudness and are an adaptive component for analyzing dynamic acoustic changes over time, the nature of our sound environment.

CORRECT FOR AUDIBILITY OR LOUDNESS

If there is a loss of OHCs resulting in less compression, there will be a reduction in hearing range, a hearing loss. Simple amplification to correct for this threshold or range deficit was assumed to be sufficient. Fixed or linear amplification was found to be insufficient and it represented a faulty strategy in all but the most severely deafened in whom auditory expectations were low. For a hearing aid to meet the impaired listeners' demands, it must keep soft sound audible but still soft, and loud sounds loud but not uncomfortable. So how do sound intensity and loudness relate in regard to hearing aid function?

Sound level is measured in decibels (dB). The decibel is a logarithmic unit defining the ratio between two sound levels. The reference sound unit used in hearing and electroacoustics is 20 micro Pascals (μPa). SPL in $\text{dB} = 20 \log (\text{pressure}/\text{reference})$. The sound pressures that the human ear can detect are also measured in dB SPL. Clinically, the mean threshold of a sample of normal listeners is defined as 0 dB hearing level (HL). Loudness is not measured in SPL, but rather tones. For the sake of brevity, loudness will be simplified by substituting dBHL in the loudness match data. In a loudness matching task, a listener will match the loudness of a tone with a reference tone at levels at/above threshold. For those with hearing loss, a small intensity change will produce a larger loudness



perception than normal ears. Loudness growth patterns in impaired ears are key factors in successful use of amplification.

In normal listeners, loudness of a 50-dB HL tone is matched to a tone of the same frequency at about 50 dB HL. A 70-dB HL tone is matched with a 70-dB HL tone, and so on. With varying degrees of hearing loss, loudness judgments differ from the norm. Note in Fig. 2 that those with a mild loss will match a tone at their elevated threshold (40 dB HL) to a tone just detectable (0 dB HL). They will match a tone of 30 dB above their threshold with a tone of 60 dB HL and tones at and above 40 dB above threshold will be matched in loudness just like a normal-hearing listener. Thus, this mildly hearing-impaired listener has an elevated threshold plus a rapid gain in loudness perception for soft sounds and normal loudness perception at high intensities. The presence of such loudness recruitment goes beyond fixed gain strategy of compression for loudness normalization.

A listener with a very high threshold of audibility will experience a very narrow range of rapid loudness growth, but will rate sound as uncomfortable at about the same level as the normal and mildly impaired listeners. Thus, each listener has different loudness needs. Those with mild losses require no gain at high levels, while those with severe and profound losses do. It is also evident that, even in these cases, there is loudness recruitment near the threshold of pain (120 dB HL). All hearing-impaired listeners have impairment for soft sound, but only some have impairment for loud

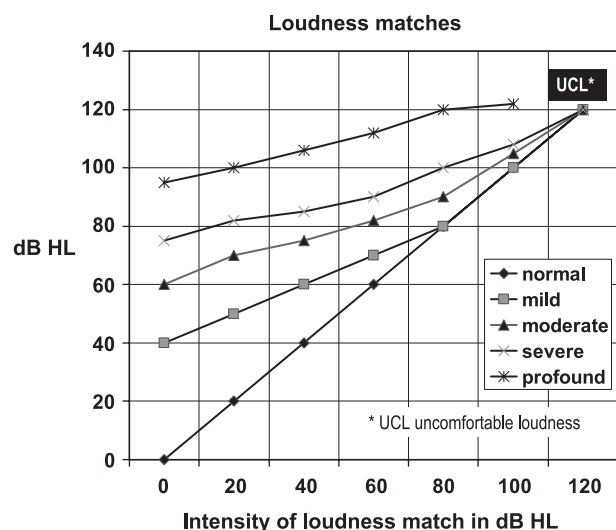


Fig. 2 Normal-hearing listener loudness balance is compared with listeners with varying degrees of hearing loss. The level on the abscissa is matched with the listener's actual threshold (ordinate). Note how loudness growth is hearing level-specific, due to the degree of hair cell loss.

sounds. Note in Fig. 2 that thresholds are only one defining characteristic of hearing loss. The need for loud sound amplification is only present in those with severe to profound hearing loss. These listeners have lost all of their OHCs and many of their IHCs. Listeners with less hair cell damage require little or no amplification for loud sounds since their loudness perception is normal. The relationship of hearing loss type and soft versus loud sound amplification is summarized in Table 1. In the past, the concept of fixed amplification was selected since the assumed impairment was lack of audibility. Hearing aids often delivered too much loudness and that was limited by either peak clipping or automatic gain control, neither of which was satisfactory to hearing-impaired listeners. What is evident is that the impairment in loudness and linear amplification is simply the wrong treatment strategy. When viewed from the perspective of loudness, why amplify normal loudness perception for loud sounds only because there is impairment in detecting soft sounds?

If soft sounds are amplified but remain soft, most sounds remain comfortable, and if loud sounds are not uncomfortable except at intense levels, then hearing aids would be mimicking the actions of the two types of hair cells in the inner ear. Hearing aid compression circuits functionally substitute for missing or damaged OHCs and adjust gain across the wide dynamic range of hearing. Linear growth in loudness is not physiologically normal. Linear amplification only fits the needs of some severely hearing-impaired listeners and only over a reduced range. Fixed amplification is certainly one factor in aid rejection. It has been evident over the last decade that hearing aids should mimic the loudness compression of OHCs in attempting to normalize amplified listening. One representative compression scheme is presented in Fig. 3. Note that there is a rapid growth in loudness for soft sounds and a reduced loudness range for loud sounds. The knee point can be shifted in intensity for different responses for different frequencies, allowing programmability

Table 1 The need for amplification of loud sounds is a function of the severity of the hearing loss

Hearing loss	Loudness and hearing loss	
	Amplification	
	Soft sounds	Loud sounds
No loss	No	No
Mild	Yes	No
Moderate	Yes	Some
Severe	Yes	Yes
Profound	Yes	Yes

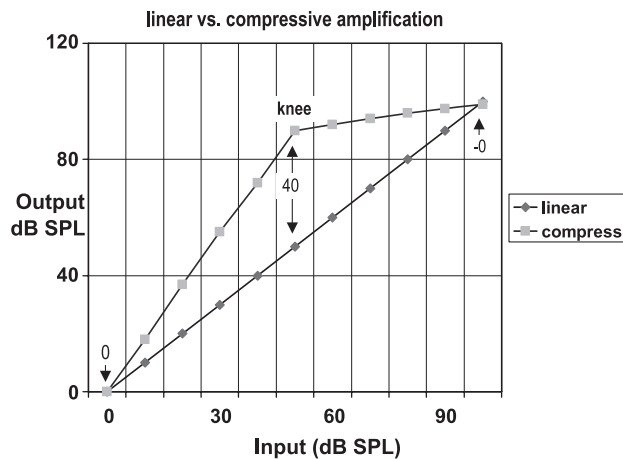


Fig. 3 Linear amplification is contrasted with compressive amplification for one frequency. Compressive amplification can be adjusted to a listener's growth in loudness.

over a wide band of frequency channels, resulting in dynamic range compression. This process is termed wide dynamic range compression (WDRC). WDRC compensates for the loss of range compression with missing OHCs. Multiple filter band compression in WDRC was used more than on frequency channel. In most schemes, there is the adaptation stage, followed by band pass filter, and an adaptive output stage.

FIT THE NERVE, NOT THE AUDIOGRAM

Inner-ear biology reveals that 120 dB of sound intensity is compressed into the dynamic range (~ 50 dB) of IHCs by the motoric response of OHCs (Fig. 4). This range compression allows loudness coding of up to 120 dB in auditory nerve fibers (95% of which innervate IHCs). Auditory nerve fibers are all characterized by a best frequency, that is, a frequency that stimulates the fiber at the lowest intensity level. Fibers are also characterized by sharp high-frequency cutoff and a broad low-frequency response. As a result, fibers respond to a limited number of frequencies at or near threshold, but a much broader number of lower frequencies at high sound intensities. For this reason, loud sound perception may not be impaired for some listeners with hearing loss. The low-frequency areas of fibers are called tails and there is considerable tonotopic overlap (spatial alignment by best frequency) in adjacent fiber-frequency response. This nerve fiber overlap contributes to the 25 critical bands in the ear. Critical bands are band pass filters about 1.2 mm wide encompassing about 1300 hair cells and the nerve population that innervates them. For example, if two tones delivered to the ear fall within the same critical band, then the perceived loudness increases slightly in

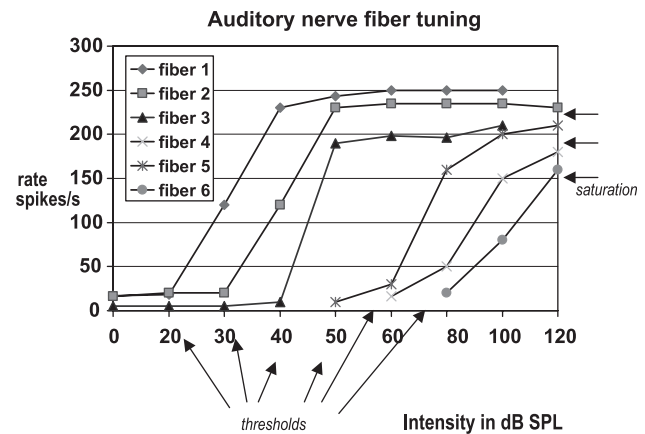


Fig. 4 Response patterns of six theoretical auditory nerve fibers are depicted. Thresholds, firing rates, and saturation rates differ, allowing complex coding in the nerve.

reference to that of one tone. If the second tone falls in a different critical band, the resulting sensation will be much louder. In the former, there is competition for the same nerve fibers; in the latter, two distinct populations of fibers are responding. Loudness is therefore related to the place of stimulation in the cochlea as well as the intensity of stimulation. The concept of multiple-channel hearing aids is an attempt to imitate the critical bands in the speech frequencies.

Speech can be described acoustically as a fundamental frequency (about 120 Hz in men and 220 Hz in women) excited by the vocal folds with harmonics (formants) induced by subsequent vocal tract filtering. The acoustic energy in speech is focused in the fundamental and lower formant frequencies, but most of the information in speech is carried by high-frequency consonant sounds. Noise and silent gaps also play a key role in speech intelligibility, notably in stops, fricatives, and nasals. Since the basilar membrane applies a spatial mapping to frequency, such that energy peaks in speech will be represented as displacement maxima at different place locations on the basilar membrane, the transform of speech acoustical peaks to basilar membrane displace maxima is the speech place code for the fundamental (f_0) and the first three to four formants (f_1 , f_2 , f_3 , f_4). For lower formants and f_0 , there is phase locking of the auditory neurons to the speech frequencies, providing the basis of a second neural code, which becomes more central in hearing-impaired listening (see Fig. 4 in the article *Inner Ear Implants*, p. 849).

Hearing loss is generally high frequency, and the loss of OHCs reduces the tuning for the high-frequency speech components. With progressive hearing loss there is a proportional loss of speech information. When natural compression is lost, auditory nerve fibers are less sensitive and become more broadly tuned. Almost all frequencies can stimulate a fiber in a



damaged cochlea without functioning OHCs if the sound intensity is intense. Broad tuning is only exacerbated by linear amplification.

In the normal cochlea, f_1 and f_2 have distinct places of maximal excitation and this place code is well represented in the nerve. After hearing loss, i.e., loss of OHCs, f_1 spreads across much of the cochlea and f_2 shifts away from its expected place in a normal cochlea. f_2 is poorly represented in the place code, but some phase locking is present. The robust place and phase cues for f_2 are lost after hearing loss. In other words, f_3 is replaced by a broadly represented f_1 and overlaps the place code and critical bands for f_2 . This is one reason speech becomes intelligible. Amplification does to a limited degree restore the place and phase coding for f_1 , but does not restore the place code for the other formants. The neural substrate for speech perception is reduced for the hearing-impaired listeners. If the formant peaks of f_2 and f_3 are amplified, but not the harmonics between f_1 and f_2 , f_2 coding in the nerve will be partly restored.^[6] Speech can be maximally coded in the compromised nerve to provide linguistic information the brain is lacking.

Auditory nerve fibers are not homogenous; they are characterized by different response thresholds, dynamic ranges, bandwidths, and spontaneous firing rates. All contribute to speech coding. Manufacturers provide guidance to fitting a product based on a patient's audiogram falling within the recommended range of hearing loss. Such an approach is assumed to offer dispensers quick access to the appropriate instrument. Unfortunately, selecting an aid by the audiogram configuration only addresses the degree of threshold loss, not the loudness loss. It is the patient's loudness requirements that must be normalized by an aid. Ideally, one would fit the nerve and not the audiogram, so is WDRC sufficient in providing loudness signal processing parameters?

In WDRC, soft sounds are amplified more than loud sounds. In an example shown in Fig. 3, linear amplification is contrasted with rapid level-dependent tuning using nonlinear compressive amplification. In this example, a one-third order time varying narrow filter is used at lower levels and a linear first-order broad filter is used at high levels. Compression acts on the amplitude of the signal; as the input changes the output changes, but by a lesser amount. The ratio of the input to output is the compression ratio, which generally varies between 2 and 3. The objective is to mimic auditory nerve-fiber function at low and high levels within a critical band. The compressor should not introduce nonlinear distortion in each band, a problem with low-level amplification. Compression can be used in the speech preamp as well as the output.

WDRC automatically adjusts the gain across the audible range. For listeners with a mild hearing loss,

the dynamic range, defined as the difference between threshold and uncomfortable levels, may be 80 dB, whereas in a listener with profound loss, the dynamic range may be 30 dB or less. Hearing-aid software samples input intensity, then scales the output by selected compression to compensate for listener-specific loudness recruitment (growth in loudness). Loudness recruitment is usually frequency-dependent, so more than one band of compressive amplification is typically used with different compression ratios set in each band. Rapid multiband compression sets dynamic output levels to keep loudness from reaching uncomfortable levels using software programs, which are enabled to detect periodicities in speech envelope. WDRC aids in suppressing speech echoes in reverberant space, although this remains a problem in meeting high intelligibility expectations (see beamforming later in this article).

More complicated are the fitting strategies using high-frequency emphasis to address the issue of degradations in the neural speech code after OHC loss. Aids must provide sufficient gain to make inaudible sounds audible. Using one strategy as an example, low-level gain (Gain_L) is set initially by subtracting 20 dB from the hearing level. In the case of a 30-dB HL threshold at 2000, the gain would be 10 ($\text{Gain}_L = \text{HL} - 20$ dB). For thresholds above 60 dB HL, only half the gain is used as a rule of thumb, since the range between severely elevated threshold and uncomfortably loud is very small. Sufficient gain is needed to be comfortable, without reaching UCL. For a listener with a threshold of 70 dB HL, comfortable may require only 20 dB or less of gain above that need for detection of soft sounds. There are a number of gain-fitting programs that predict an average target gain formulated from specific assumptions about amplification and ear function. This is the starting point for fine tuning. It is almost a requirement to sample the gain in the ear canal between the hearing aid and eardrum to verify the fit. Additional fine-tuning can be very important once the aid is seated into the geometry of the ear, requiring true individual settings. This is accomplished by inserting a probe microphone between the hearing aid and the eardrum (Fig. 5). The frequency band gain compression characteristics can be adjusted as needed. Future user adjustment will be desired as the brain accommodates to amplified speech.

Directional microphones can suppress noise from behind a listener by using the microphone null with adaptive spectrum subtraction techniques. Just 6 dB suppression in the ambient noise can potentially increase discrimination by almost 40%. If an array of four microphones is used, combining fixed (time invariant) and adaptive (adjusts continuously), beamforming can result in an additional (<12 dB) noise suppression. The adaptive beamforming algorithm calculates the sum and difference of the incoming

Verifying the gain



Fig. 5 A method of verifying the real ear gain of a hearing aid is depicted. An ear canal probe is placed between the eardrum and aid to measure the exact SPL. An external microphone provides the reference.

signals. The difference will be primarily noise in spectrum subtraction from the sum (mainly speech). This is based on the assumption that speech and noise are uncorrelated; hence, the noise would be random in reference to any part of the speech spectrum. Noise will likely not be time-locked to speech and time delays due to reverberation can further reduce the noise. There are certainly various parameters to alter in an attempt to maximize the value of fixed, adaptive, and combination beamforming. An example of one early version using an eyeglass platform is depicted in Fig. 6.

BRAIN AIDS

Unlike the peripheral ears of mammals, the central auditory systems are plastic and rapidly reprogram in response to the conditions of the cochlea. Insults (noise, ototoxic drugs, infections, trauma, etc.) to the cochlea that cause loss of hearing (OHCs and IHCs) trigger reprogramming in the brain, especially in the neocortex.

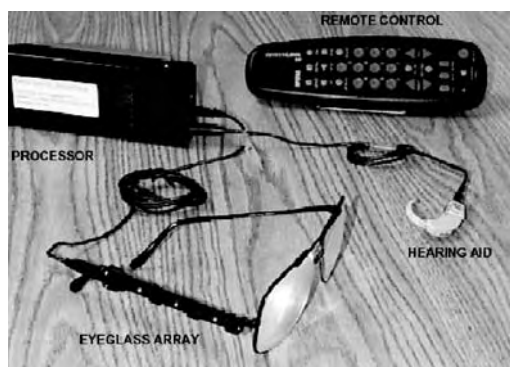


Fig. 6 Key elements to an adaptive beamforming hearing aid. (Courtesy Planning Systems, Inc., Reston, Virginia.)

Cortical neurons are tonotopically organized, as is the cochlea due to the filtering characteristic of the basilar membrane on which the hair cells lie. Cortical neurons with best frequencies tuned to the corresponding damaged area of the cochlea change their frequency response by reprogramming lower in frequency. With increasing high-frequency hearing loss, more and more neurons are tuned to lower frequencies. The auditory brain seeks stimulation. Reprogramming raises two issues in regard to hearing aids.

- Fitting with a device that is not adaptive and programmable seems futile in the light of neural reprogramming in the brain, which is the basis of auditory processing, attention, and learning.
- Aids should not be fitted once; periodic reprogramming appears necessary. A digital device should be viewed as a speech processor coupled to a dynamic neural processor. The lack of refitting, as it were, may be a main contributing factor to lower-than-expected acceptance of hearing aids after purchase.

An interesting and related auditory reprogramming phenomenon is termed late-onset auditory deprivation. Patients using a single hearing aid in the presence of binaural hearing loss exhibit a curious loss of speech discrimination in the unaided ear. That is, speech discrimination declines in the nonamplified ear and discrimination is stable or improved in the aided ear. If the unaided ear is subsequently fitted, speech discrimination improves in the newly fitted ear to equal that of the first aided ear. The most salient explanation is that the brain reprograms to process the new speech information based on enhanced peripheral coding. The brain is favoring the amplified channel over the unamplified ear. Even if this is initially a matter of selective attention, brain reorganization seems to be involved. These data argue for binaural fitting, but also attest to the plasticity of the brain and its effects on speech processing and the need to keep assessing the hearing-aid fit.

LOOKING TO THE FUTURE

A plethora of unique auditory adaptations are present in reptiles.^[7] Turtles, for example, have been on this planet for more than 200 million years. They have outlived many species including the dinosaurs. All turtles have a unique hearing structure in their inner ears: fibroelastic strands that couple the middle ear bone to the wall of the saccule. Humans, having reduced the substrate coupling and increased the distance from their ears to the earth with bipedal posture, have reduced saccular stimulation. If, however, silicone fiber elastic strands are surgically placed through the footplate of the stapes and connected to the saccular



wall, the threshold for saccular hearing would be approximately 35 dB SPL. Functional air conduction hearing can be obtained and it may serve as an adjunctive medical procedure to a cochlear implant or a stand-alone method of hearing remediation.^[8] A saccular coupled hearing aid could be designed using some of the digital concepts highlighted in this review.

Another innovation is frequency transposition, which captures the high-frequency parts of the speech spectrum and lowers it into the audible spectrum. The concept is attractive in that speech is being frequency-compressed into the hearing-impaired ear since cues associated with the high-frequency sound are transposed. This approach has not generated substantial support. Failure is often attributed to inability to discriminate the new frequency transposed cues because discrimination is not very good in the low frequencies. While this is true for normal-hearing listeners, those with high-frequency loss have a larger low-frequency neural area due to reprogramming. Learning to use these aids takes time and training may be an important variable. Recall that f_0 and f_3 in vowels is shifted toward the high-frequency base of the cochlea and masked by a broadly represented f_1 . What happens when high frequencies are overlapped in f_0 and f_1 ? Place cues may be compromised, but some temporal information might persist. Software prespeech processing should be able to improve the outcomes with transposition aids. Ultrasonic modulating can also squeeze speech into the deaf ear. Tones from 20 to 100 kHz can be detected by bone conduction, but not by air conduction due to the filtering action of the middle ear. The deaf can detect these ultrasonic tones too, as long as there is some vestibular function present. Ultrasonic speech stimulates the auditory cortex and is an emerging hearing-aid technology.^[9,10]

Automatic speech recognition (ASR) systems are capable of reasonable performance today, although inherent variability remains. Speech features such as voicing, place of articulation, nasalization, etc., form contrasts that, in isolation, can identify phonemes. Feature detection in discourse is more problematic since acoustic features overlap. The use of short spectra features of speech on the order of 10–20 ms, as in short-term FFTs, can be applied to resynthesis of speech. A classification system would not have to construct the original speech, only a representation. ASR may not have the accuracy and speed for resynthesis today, but the potential to adapt ASR to the neural coding of speech may be the next breakthrough in hearing aids. With ASR, high-frequency consonants could be enhanced, vocal fundamental frequency and its harmonics reduced in intensity to limit expanded formant place code confusion and upward spread of masking. ASR could preprocess speech tailored to the listener. ASR algorithms should

increase the speech-to-noise ratio and reduce speech errors introduced by interspeaker variability.

CONCLUSION

The concept that hearing aids should be fitted to loudness loss and not hearing threshold is far from new, but is revived and obtainable through the use of digital signal processing techniques. The loudness of speech is proportional to the total number of auditory nerve fibers that are activated. Fibers fire synchronously by spatial (place on the basilar membrane) and temporal (time/phase locking) coding. The challenge for digital hearing aids is to preprocess or condition the speech signal to maximally stimulate the remaining sensorineural elements in the ear and brain. Biologically inspired innovation will certainly lead to novel approaches to aural rehabilitation. The reexamination of bone conduction hearing as a mechanism to deliver preprocessed sound to the ear with actuators' arrays, possibly subcutaneously implanted, on the skin of the head and neck is just taking place. Fly ears have stimulated research into miniature microphones needed for arrays in small, completely in the canal aids, which allow extracting speech from the ambient noise and preprocessing it for better neural stimulation. After all, we hear with our brains, not with our ears.

ARTICLES OF FURTHER INTEREST

Biopotential Amplifiers; Inner Ear Implants; Tinnitus Devices

REFERENCES

1. Miles, R.N.; Sundermurthy, S. *In A Biologically Inspired Directional Microphone Concept for Hearing Aids International Hearing Aid Conference*: Lake Tahoe, CA, 2000.
2. Didier, A.; Cazal, Y. Acoustic responses recorded from the saccular bundle on the eighth nerve of the guinea pig. *Hear. Res.* **1989**, *37* (2), 123–237.
3. Sheykholeslami, K.; Kaga, K. The otolithic organ as a receptor of vestibular hearing revealed by vestibular-evoked myogenic potentials in patients with inner ear anomalies. *Hear. Res.* **2002**, *165* (1–2), 62–67.
4. Ribaric, K.; Kekic, B.; Dergenc, R. On the capability of the vestibular apparatus to perceive sound stimuli. *Acta Oto-Laryngol.* **1992**, *112* (2), 221–224.
5. He, D.Z.; Dallos, P. Somatic stiffness of cochlear outer hair cells is voltage-dependent. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *6*, *96* (14), 8223–8228.
6. Miller, R.L.; Calhoun, B.M.; Young, E.D. Contrast enhancement improves the representation of /e/-like

- vowels in the hearing-impaired auditory nerve. *J. Acoust. Soc. Am.* **1999**, *106*, 2693–2708.
7. Wever, E.G. *The Reptile Ear*; Princeton University Press, 1978.
 8. Lenhardt, M.L. Eardrum Saccular Coupling: Novel Form of Hearing. In *Biomedical Engineering: Recent Developments*; Vossouggghi, J., Ed.; Medical and Engineering Publisher: Washington, DC, 2002; 51–52.
 9. Lenhardt, M.L.; Skellett, R.; Wang, P.; Clarke, A.M. Human ultrasonic speech perception. *Science* **1991**, *5*, 253 (5015), 82–85.
 10. Imaizumi, S.; Hosoi, H.; Sakaguchi, T.; Watanabe, Y.; Sadato, N.; Nakamura, S.; Waki, A.; Yonekura, Y. Ultrasound activates the auditory cortex of profoundly deaf subjects. *NeuroReport* **2001**, *5*, 12 (3), 583–586.

Heart Biomechanics

Jeffrey W. Holmes

Department of Biomedical Engineering, Columbia University, New York, New York, U.S.A.

INTRODUCTION

The heart is a remarkably efficient and durable mechanical pump composed of an extraordinarily complex biologic material. The field of cardiac biomechanics seeks to understand how the composition, structure, electrical activation, geometry, and hemodynamic environment of the heart interact to produce known physiologic behavior, and to apply this knowledge to understand and modify pathophysiologic responses that arise, as various disease processes affect the heart. Reviewing all that has been learned about heart biomechanics in a few pages is clearly impossible. Rather, this article is intended to serve as an introduction to the field, and therefore has the following goals: 1) to introduce the reader to the essential features of normal cardiac physiology that cardiac mechanics seeks to explain; 2) to introduce a general framework for analysis and modeling of problems in cardiac mechanics; 3) to point the reader to a number of more-detailed reviews of specific aspects of cardiac mechanics that have been well-studied; and 4) to highlight a few of the more exciting cardiac mechanics problems currently under investigation.

THE HEART AS A PUMP

The basic physiologic function of the heart is to pump blood, and well before most of the biochemical, structural, and physiologic details of muscle function were known, physiologists were already studying the pump function of the heart. One of the intriguing paradoxes of cardiac mechanics is that despite the immense underlying complexity and the fact that most known kinetic, mechanical, and geometric relationships in the heart are highly nonlinear, the overall pump function of the heart can be described with surprising accuracy by a few simple curves.

Ventricular Function Curves

With each beat, the chambers of the heart contract and relax, while a series of valves ensure one-way flow of blood. During the portion of the cardiac cycle termed diastole, when the heart is filling, the right ventricle

receives blood returning from the body through the veins into the right atrium (venous return), while the left ventricle receives oxygenated blood from the lungs through the left atrium. During the portion of the cardiac cycle termed systole, when the ventricles are contracting, valves prevent backflow into the atria and instead direct right ventricular outflow into the pulmonary artery and on to the lungs and left ventricular outflow into the aorta and out to the body. One way of understanding pump function of the heart is to look at the relationship between the input to that pump and its output. Possible inputs of interest might be the volume of blood entering the heart or the net pressure driving blood into the heart; outputs of interest might be the volume of blood pumped, the amount of pressure generated, or the mechanical work done by the heart. In fact, most of these different inputs and outputs have been plotted against one another by cardiac physiologists and used to analyze heart function. Two of the most popular plots, named after early physiologists who described or popularized them, are Starling's cardiac output curve, which considers the heart as a whole and plots right atrial pressure (input pressure) against cardiac output (output volume per minute), and Sarnoff's ventricular function curve, which focuses on the left ventricle and plots end-diastolic volume or pressure against mechanical work per beat (stroke work).^[1] Regardless of the choice of variables, the message is the same: The heart is able to increase its output dramatically in response to small changes in input over the normal physiologic range, until eventually a plateau is reached reflecting the maximal pumping capacity of the heart under a particular set of conditions (Fig. 1). The very important physiologic and clinical implication of this finding is that normal hearts under normal circumstances are able to pump blood as fast as it returns from the body, and the primary regulator of cardiac output is actually venous return.

Pressure–Volume Curves

What is the basis for the dramatic responsiveness of the heart to variations in input flows or pressures? A great deal of insight can be gained by plotting left ventricular pressure against left ventricular volume

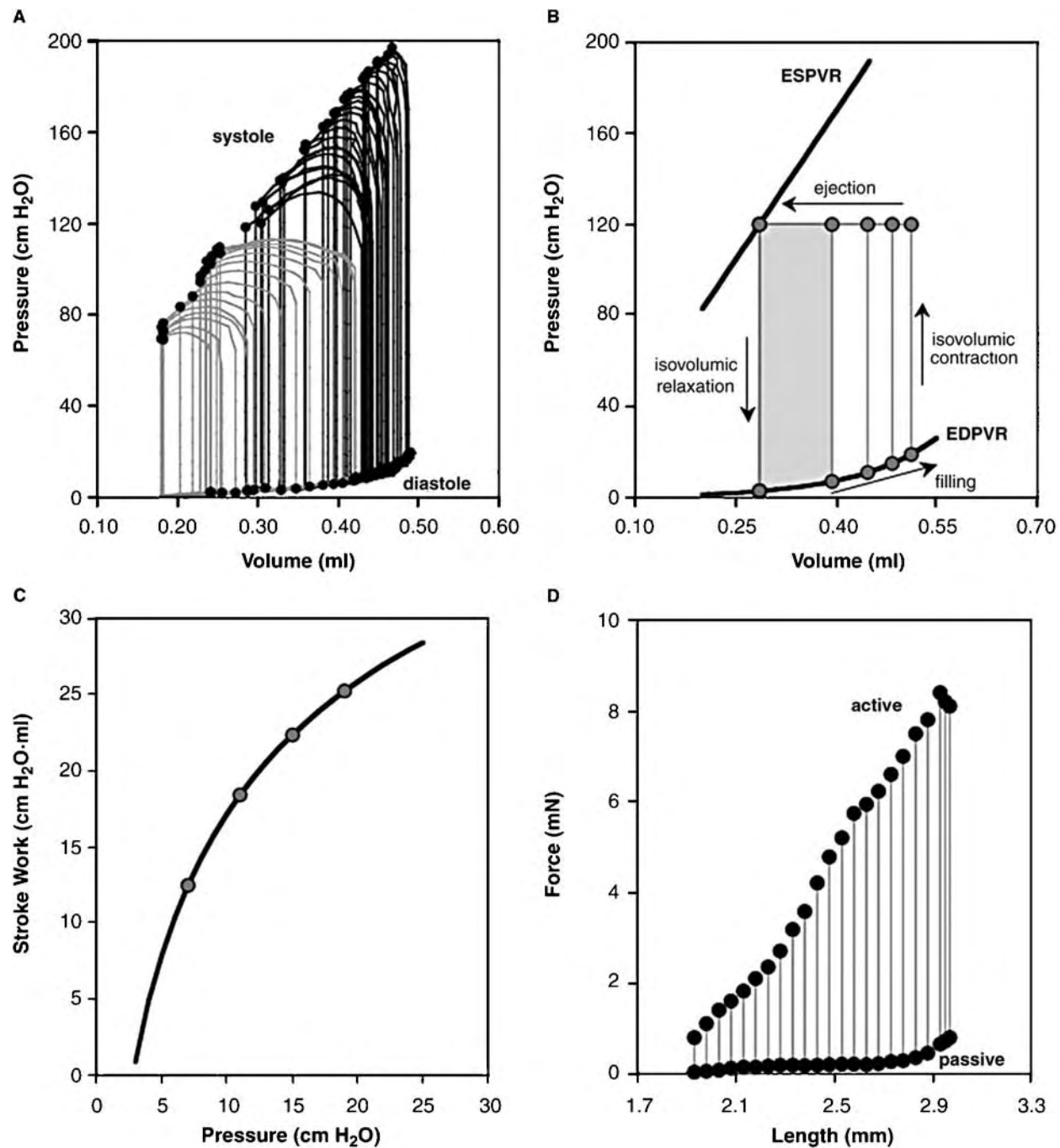


Fig. 1 Analysis of left ventricular pump function. Panel A: Pressure–volume loops from an open-chest anesthetized rat during occlusion of the inferior vena cava (gray) or the ascending aorta (black) to vary hemodynamic loading. Circles indicate end-diastolic and end-systolic corners of PV loops. Panel B: Simulated response to changing filling pressure while holding systolic pressure constant, based on linear ESPVR and exponential EDPVR fitted to data from panel A. As filling pressure increases, stroke volume and stroke work (indicated for first loop by shaded area) increase rapidly at first, then more slowly because of the exponential shape of EDPVR. Panel C: Plotting stroke work against filling pressure yields Sarnoff's ventricular function curve. Panel D: Isometric force–length data from an isolated rat right ventricular papillary muscle illustrate the similarity between muscle force–length relationships and chamber pressure–volume relationships.

over the course of the entire cardiac cycle. In the pressure–volume plane, the cardiac cycle is reflected in a counterclockwise loop: filling at low pressures, followed by isovolumic contraction, then ejection at

higher pressure, and finally isovolumic relaxation to return to the starting point (Fig. 1). While these curves were first produced for the frog heart by Otto Frank in 1898, their popularity for studying and teaching

cardiac physiology is largely because of the work of Drs. Kiichi Sagawa, Hiroyuki Suga, and colleagues beginning in the late 1960s. For the interested reader, their 1988 monograph on the pressure–volume relationship remains one of the best-written and most complete sources of history, data, and modeling regarding the use of pressure–volume loops in analyzing cardiac function.^[1]

The essential feature of cardiac function revealed by analysis in the pressure–volume plane is that the large number of trajectories that may be traced under different hemodynamic loading conditions is bounded by two simple curves, the end-diastolic pressure–volume relationship (EDPVR) on the lower right and the end-systolic pressure–volume relationship (ESPVR) on the upper left (Fig. 1). Restricting the discussion for this article to the left ventricle, the EDPVR reflects the behavior of the left ventricle during filling and is therefore determined by passive material properties of the myocardium plus a number of other factors such as ventricular geometry discussed later in this article. The ESPVR reflects the behavior of the left ventricle during systole and is therefore determined by the active material properties of maximally contracted myocardium plus other factors such as geometry. If all other factors are held constant for a given heart, experimental changes in hemodynamic loading conditions such as the inflow to the ventricle or the resistance to ejection will produce a range of different pressure–volume loops having their lower-right and upper-left corners on or very near the EDPVR and ESPVR (Fig. 1). Changes in these relationships can therefore be used to index and understand changes in cardiac function. In particular, the slope of the ESPVR has proved useful as an index of cardiac contractility. For example, administration of a drug that increases the contractile capacity of the myocardium (contractility) will shift the ESPVR of a normal ventricle upward and to the left without impacting the EDPVR.

While the utility of pressure–volume loops is hopefully evident, the question of how the ESPVR and EDPVR explain the ventricular function curves we examined earlier has still not been answered. The two keys to this question are that the shapes of these two relationships are different, and that under normal conditions arterial blood pressure is maintained within a narrow range. While the ESPVR is nearly linear in most situations, the EDPVR is highly nonlinear, with little pressure required to fill the ventricle at low volumes but much more pressure required to continue to fill the ventricle once it reaches higher volumes. Imagine starting from a very low end-diastolic pressure and gradually increasing it while maintaining end-systolic pressure at a constant value. The first PV loop would show substantial filling, then ejection to the ESPVR and relaxation (Fig. 1B). The next loop would

show slightly more filling (larger end-diastolic volume), then ejection to the same point on the ESPVR, and therefore a larger stroke volume and stroke work (area inside the PV loop); the fact that increased filling leads to increased ejection over the normal working range of the left ventricle is known in physiology as the Frank–Starling effect. Continuing the experiment, the trend would continue until the exponential shape of the EDPVR began to limit the increase in filling obtained with the next step in end-diastolic pressure, and therefore the resulting increase in stroke volume and stroke work (Fig. 1C). At some point, the EDPVR becomes so steep that additional increases in end-diastolic pressure no longer produce any change in left ventricular output, resulting in the plateau seen on ventricular function curves.

CARDIAC MUSCLE MECHANICS

As we work our way into this discussion of cardiac mechanics, it may seem we are simply replacing one question with another. The shape of observed ventricular function curves can be explained using the EDPVR and ESPVR, but how do the EDPVR and ESPVR arise? One more step in this discussion will bring us to the fundamental questions that have long occupied cardiac physiologists and biomedical engineers and that continue to define the field of cardiac mechanics today.

Force-Length Curves in Isolated Cardiac Muscle

Much of what is known about activation, contraction, and energetics of striated muscle was originally learned in isolated muscle preparations. Physiologists placed isolated muscles into baths allowing control of factors such as temperature, pH, and oxygenation, electrically stimulated these muscles to contract, and measured various features of muscle performance such as velocity of contraction, heat output, etc. From a mechanics standpoint, the most important feature of these experiments was the way in which the muscles were loaded and the terms that arose to describe forces acting on a muscle during contraction. To stretch a relaxed muscle to a longer length, small weights were hung from the muscle; as these weights represented the force the muscle experienced prior to contraction, they were referred to as the “preload.” In many experiments, an additional weight was positioned so that once contraction began, the muscle was required to lift both the initial (preload) and the additional load to shorten. The load experienced by the muscle only after it began to contract came to be known as the “afterload.” It is important to note that these loads

were always quantified in terms of force normalized by muscle cross-sectional area; although the term “load” has persisted in the physiology and clinical literature, in fact preload and afterload as originally defined refer to stresses experienced by the muscle.

Several basic relationships that emerged from this early work remain central to our understanding of the mechanics and energetics of isolated striated muscle. These include the relationship between systolic and diastolic force and muscle length (force–length curves), between afterload and muscle shortening velocity (force–velocity curve), and between energy liberation and afterload. Because of space limitations, only the force–length curve of isolated cardiac muscle will be discussed here. However, excellent reviews are available for those interested in more detail, regarding other aspects of the physiology of isolated cardiac muscle: Hunter, McCulloch, and ter Keurs^[2] published a mathematical model of cardiac muscle mechanics, which includes extensive comparison to available experimental data, while a slightly older review by Cooper^[3] is an excellent starting point for obtaining more information on the energetics of isolated cardiac muscle.

As unstimulated isolated cardiac muscle is gradually stretched beyond slack length, it is initially very compliant. However, as length increases the force required to achieve the next step in stretch rises exponentially, yielding a passive force–length curve very similar to the EDPVR of the left ventricle (Fig. 1D). If the muscle is stimulated to contract at each of these lengths and the isometric twitch force recorded, the resulting active force–length curve is very similar to the left ventricular ESPVR (Fig. 1D). Thus, it is apparent that much of the behavior of the left ventricle discussed above and reflected in the EDPVR and ESPVR likely results from basic properties of cardiac muscle. Much of the field of cardiac mechanics is currently devoted either to understanding these muscle properties and how they change in different situations or to understanding how to integrate these local muscle properties to quantitatively predict regional and global ventricular function *in vivo*.

A FRAMEWORK FOR CARDIAC MECHANICS ANALYSIS

At the essence of most solid mechanics and mechanics of materials problems is the desire to relate loading, deformation, and material properties to completely understand the response of a material or structure under expected loading conditions. Consider, for example, the simple beam problems familiar to any student who has taken an introductory strength of materials course. Shear and moment diagrams allow

the calculation of stresses on any cross-section of the beam given the type and location of applied loads; these stresses are the most useful specification of the load experienced by the material. Because the relationship between applied stresses and the resulting deformation (i.e., the material properties) is very simple in many engineering materials, predicting the resulting deformations from the computed stresses is straightforward. Simple material properties also make it easy to invert the problem, observing deformation, and computing what stresses must have been acting to produce that deformation.

In cardiac mechanics, the basic problem is the same: relating loading, deformation, and material properties to understand the response of the structure in various situations (Fig. 2). However, both the problem of determining what stresses act on the material under a given set of loading conditions and the problem of specifying the material properties are much more complex. Most of the field of cardiac biomechanics deals in one way or another with answering these two fundamental questions.

Determining the Stresses Acting on Myocardium

Unlike the simple beam-bending problem, in the case of the heart, the relationship between applied loads and the resulting stresses acting on a given region of the material (the myocardium) is quite difficult to establish and has several important determinants (Fig. 2). Although a similar discussion would be appropriate for the right ventricle, we consider here only the better-studied left ventricle. It is obvious that hemodynamic loading conditions affect the stress in the left ventricular wall; increasing the pressure inside the left ventricle must increase the associated wall stresses. However, the specific relationship between these two depends heavily on ventricular geometry. The left ventricle is a thick-walled, elliptical pressure vessel. In general, stresses are tensile in planes parallel to the left ventricular surface (around the circumference or along the longitude at any point in the wall), and compressive in the radial direction; the magnitude of these stresses depends in part on global geometric factors such as the ratio of the different radii of the ellipse to one another and the ratio of these radii to the wall thickness.^[4] The situation is further complicated by two additional factors. First, there is residual stress in the ventricular wall. In other words, when the left ventricle is completely unloaded, with no pressure difference between the inside and outside, the stresses in the wall are still nonzero.^[5] Second, because the material properties of the myocardium are not the same everywhere in the ventricular wall, different layers of the wall may bear



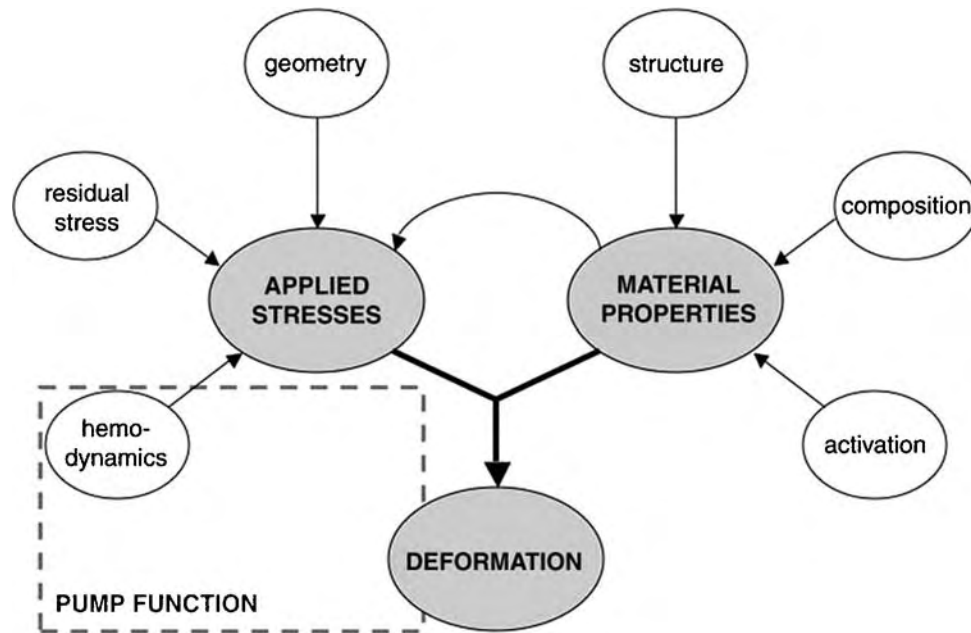


Fig. 2 General approach to problems in heart biomechanics. The overall goal is to relate loading, deformation, and material properties to understand the heart's response to a given situation or disease process. Remarkably, pump function of the heart can be described by a few simple relationships between hemodynamic loading and the volume of blood pumped by the heart. This type of analysis covers only the region of the diagram outlined by the dotted gray lines and describes the behavior of the heart without explaining the underlying mechanisms. More detailed analyses are complicated by a number of factors shown on the diagram and explained in more detail in the text. In particular, both determining the stresses acting on myocardium and appropriately describing the material properties of myocardium are difficult problems that occupy much of the current work in the field; much of the difficulty arises from the fact that these two problems interact (indicated by curved arrow in the center of the diagram).

different amounts of the overall load placed on the left ventricle, just as springs with different spring constants experience the same stretch but different loads when acting in parallel. The consequence of this last fact is that even given residual stress, loading, and geometry, stresses cannot be completely predicted without first knowing material properties.

One of the achievements of the field of cardiac mechanics over the past two decades has been the development of comprehensive, quantitative models of left ventricular geometry and fiber architecture. The groundwork for these models was laid by a pair of classic papers by Streeter and Hanna^[6,7] describing the left ventricle as a nested set of elliptical shells. However, the modern geometrically sophisticated finite-element models that currently allow iterative solution of the equations of equilibrium for a ventricle with prescribed active and passive material properties owe the majority of their development to the cardiac mechanics groups at the University of Auckland (New Zealand) and the University of California, San Diego. These extremely detailed, painstakingly developed models have been published for the dog,^[8] pig,^[9] and rabbit,^[10] pictures and examples of their use to investigate cardiac mechanics problems are available on the websites of these two groups.

Determining Material Properties of Myocardium

Like most biologic materials, passive myocardium is nonlinear, anisotropic, and viscoelastic. In addition, passive myocardium exhibits several more complex material behaviors under appropriate test conditions including preconditioning, hysteresis, and strain softening.^[11] However, the complexity of passive myocardium is only the beginning, as myocardium actively contracts with each heartbeat, cycling through a wide range of different material properties from those of passive myocardium to those of maximally contracting "active" myocardium and back again. The active material properties are both time- and history dependent and are again anisotropic, with more active stress generated along the local muscle fiber direction, but some active stress also generated perpendicular to the local fiber direction.^[12] Given such complexity, it is not surprising that much of the work to date in the field of cardiac mechanics has focused on describing, modeling, and understanding the passive and active material properties of myocardium.

There are two general approaches to determining the material properties of myocardium (Fig. 3).^[13] The first, more familiar from traditional engineering analysis, is to excise a specimen and test it in a

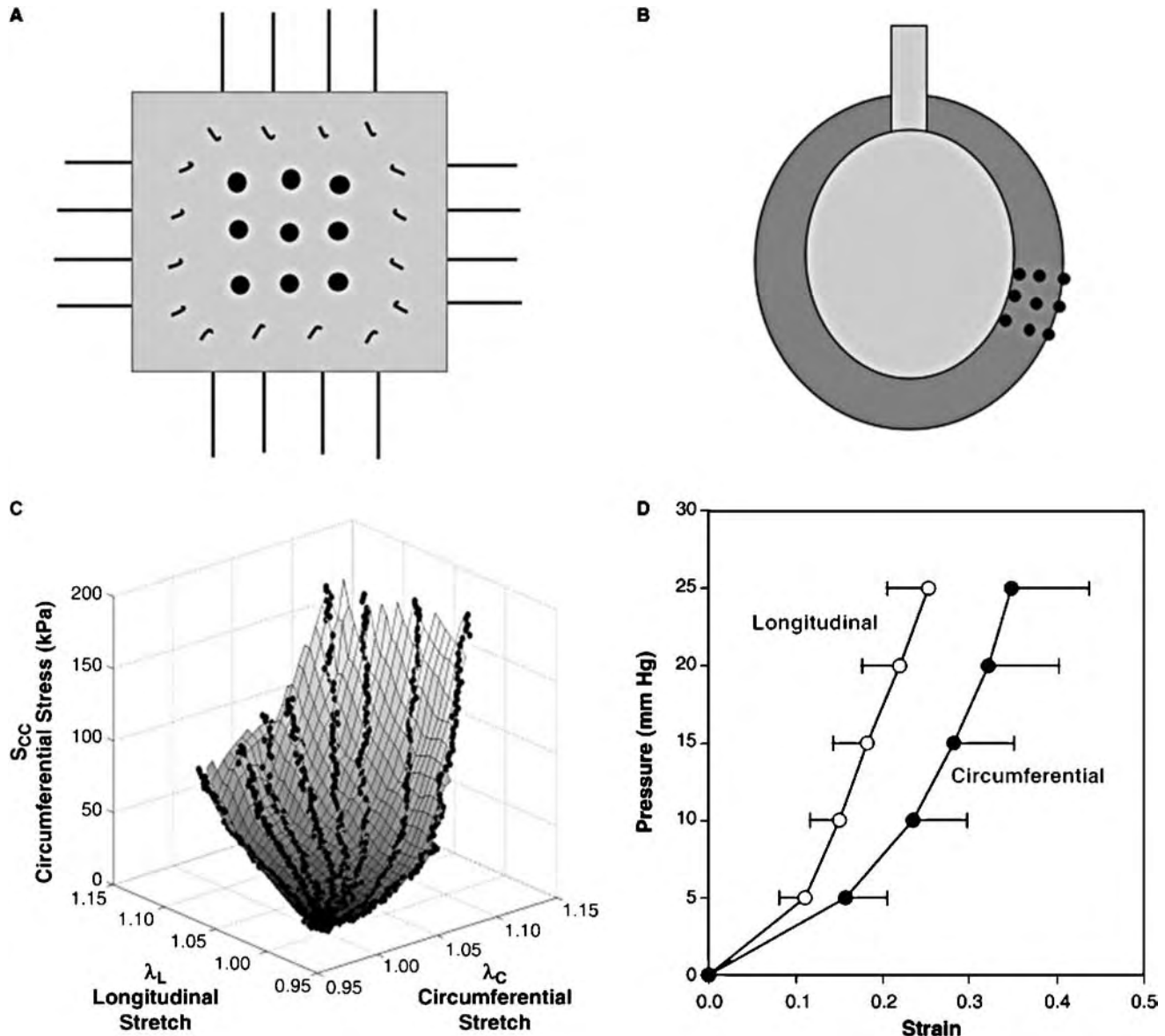


Fig. 3 Approaches to determining material properties of heart tissue. Panel A: For biaxial testing, an excised sheet of myocardium is connected to a testing device with sutures and hooks and subjected to multiple computer-controlled test protocols; deformation is tracked with markers and forces are measured. Panel B: Alternately, the intact heart may be fitted with a balloon that is inflated to known pressures and volumes while deformations in the wall are measured with implanted markers. Panel C: Biaxial test data (circles) and fitted stress surface from healing rat infarct. Biaxial testing provides more direct information on stress–strain relationships over multiple test protocols, at the expense of excising the tissue. Panel D: Data from inflation tests on isolated pig hearts. Pressure–strain data are obtained in the intact heart but only for a single test protocol; finite element modeling is required to estimate stress–strain relationships from the measured pressure–strain data.

computer-controlled testing device.^[14–17] Because myocardium is an orthotropic material, triaxial testing would be required to completely describe its three-dimensional material properties; however, these tests are very difficult to perform. Instead, most investigators have assumed the myocardium to be incompressible and performed biaxial testing. This approach has the obvious benefits that both the loading placed on the myocardium and its deformation can be accurately

measured. For passive myocardium, the main technical difficulty is preventing injury to the specimen that might cause edema, permanent binding of some muscle crossbridges (contracture), or other alterations in the material properties to be measured. For active myocardium, the technical challenges are even greater, as the myocardium cannot easily be maintained at a particular level of partial activation (representing a certain time during systolic contraction) while it is

tested, then shifted to another level of activation, and tested again. The best information about active systolic properties to date was obtained from biaxial testing of myocardium that was tetanized or placed in barium contracture to simulate maximal activation.^[12]

The second approach to determining material properties of myocardium is to leave the heart intact, measure deformation under a range of hemodynamic loading conditions, and attempt to deduce the material properties using a sophisticated model that accounts for all the other factors that impact deformation such as geometry, residual stress, etc.^[18,19] The obvious advantage of this approach is that it leaves the heart intact, so that material properties are much less likely to be disturbed by the testing.^[13] The disadvantage is that the stresses must be computed from the model and their accuracy is therefore only as good as the model; in addition, as the stresses depend on the material properties, the material properties must be optimized using an iterative procedure until the model-predicted deformations at a given loading state agree with the observed deformations. Because the match is never exact, deformation is three-dimensional, and the function describing the material properties typically has several constants, finding the true optimum (a global minimum for the optimization instead of just a local minimum) for a problem like this is very difficult. As with *ex vivo* biaxial testing, it is difficult but possible to inflate and deflate an isolated arrested heart and use the imposed pressures and resulting deformations to obtain a description of the material properties of passive myocardium; however, obtaining information across the entire range of possible active states that occur during a cardiac cycle in a beating heart is extremely difficult and little progress has been made in this area.

To this point, we have discussed only the methods used to obtain information on material properties, not how to specify those properties mathematically or to relate those properties to the structure and composition of the tissue. Constitutive modeling of myocardium is one of the more active and exciting frontiers of current cardiac mechanics research and is discussed in a separate section below.

Measuring Global and Regional Deformation in the Heart

In the attempt to relate loading, deformation, and material properties to understand the response of the heart in any given situation (Fig. 2), the measurement of deformation is the most straightforward. Global measures of deformation such as cavity volume can be obtained using a variety of imaging techniques including angiography (X-ray), ultrasound, and MRI as

well as using newer methods such as the conductance catheter. Measurement of regional deformation has been accomplished using sonomicrometers, implanted radiopaque markers tracked by high-speed X-ray, and more recently by MRI tagging. While these methods are a necessary part of any analysis, from the discussion above it should be clear that in many situations measuring deformation alone provides very little information on the underlying mechanics. For example, although it has been known for 30 years that a segment of myocardium tracked with sonomicrometers or markers shifts its pressure–strain curve to the right during acute ischemia, the fundamental question of whether this shift reflects a change in the material properties of the ischemic myocardium, a change in loading of the ischemic region, or a combination of these effects remains unanswered.^[20]

CURRENT FRONTIERS IN HEART BIOMECHANICS

While cardiac muscle and chamber physiology have been extensively studied by physiologists over many decades, and the molecular biology of heart disease is currently the primary focus of most cardiac research, the relatively small and relatively new field of cardiac biomechanics is just now entering an exciting phase where the basic tools and approaches have been developed and application to a range of important problems is possible. Here, we briefly outline two of the many promising current frontiers in cardiac mechanics.

Structural Constitutive Modeling

An equation that specifies the relationship between stresses and strains in a given material is called a constitutive equation. The development of these mathematical models of the material properties of biologic tissues is often referred to as constitutive modeling. Constitutive modeling of myocardium is one of the exciting frontiers of current cardiac mechanics research and was recently reviewed by Costa, Holmes, and McCulloch.^[13]

Historically, there have been three different approaches to constitutive modeling of passive myocardium. The first approach is phenomenologic. In this approach, the constitutive equation is treated simply as an equation to be fitted to experimental data, a concise representation of the material properties of passive myocardium. As passive stress–strain curves for myocardium are typically exponential in shape, most investigators have used exponential or power-law equations to describe these properties mathematically.

A typical approach is to use a form known as a Fung exponential, popularized for biologic tissues by Y.C. Fung. In this approach, the myocardium is usually modeled as an incompressible hyperelastic material described by a potential function called a strain energy function that relates stresses and strains (t_{ij} =Cauchy stresses, F_{ij} =components of the deformation gradient tensor, C_{ij} =components of the right Cauchy–Green deformation tensor, and p =hydrostatic pressure term used to enforce incompressibility) [Eq. (1)]

$$t_{ij} = -p\delta_{ij} + 2F_{iA} \frac{W}{C_{AB}} F_{jB} \quad (1)$$

Next, the strain energy function W is assumed to be some exponential function of the components E_{ij} of the Lagrangian finite strain tensor. For example, Guc-cione, McCulloch, and Waldman used the following form, which assumes the material is transversely isotropic with its primary material axis oriented along the local muscle fiber direction, in their model-based determination of passive canine myocardial material properties (f =fiber, c =cross-fiber, r =radial, and C and b_1 – b_3 are constant coefficients) [Eq. (2)]

$$W = \frac{1}{2} C(e^Q - 1);$$

$$Q = b_1 E_{ff}^2 + b_2 (E_{cc}^2 + E_{rr}^2 + 2E_{cr}E_{rc}) + 2b_3 (E_{fc}E_{cf} + E_{fr}E_{rf}) \quad (2)$$

One of the problems that arise when fitting these equations to data, especially in the case of iteration through a complex finite element model as discussed above, is the fact that most of the equations that have been chosen ad hoc by various investigators contain constants that are not independent of one another. One of the more novel contributions in recent years was made by Criscione, who developed constitutive equations with parameters that are linearly independent of one another to facilitate robust parameter estimation from experimental data.^[21,22]

The second general approach to constitutive modeling of myocardium is to try to select mathematical forms for the constitutive equations that are preferred because the forms of the equations themselves reflect some underlying properties of the tissue. Some of the best work of this kind was done by Humphrey, Strumpf, and Yin^[16,17] who formulated their strain energy function W as a polynomial function of the invariants of the right Cauchy–Green deformation tensor C and then designed computer-controlled biaxial tests to vary one invariant at a time and determine the appropriate order for each term in the polynomial. A very different approach that still falls in the same general category has been taken by the group at University of Auckland. They argued that the limit

to distensibility of the myocardium in any direction is set by collagen fibers that become essentially rigid once they are fully extended, and that the mechanics would therefore be best described using an equation that has poles in each direction (a “pole-zero” constitutive law) to reflect this structural feature.^[23]

This brings us to the final, most challenging, but potentially most exciting approach to constitutive modeling of passive myocardium. This approach, termed “structural” constitutive modeling, attempts to build a constitutive equation based on measured composition and microstructure. Early work in this area was done by Lanir and coworkers, with a specific formulation for myocardium published by Horowitz.^[24] Some formulations of this type postulate multiple fiber families and end up with so many parameters that it becomes impractical to obtain them by fitting experimental data. However, the most aggressive forms of this approach take the philosophy that as little as possible should be fitted, with most of the parameters based on direct measurements of tissue structure. We have taken this approach to constitutive modeling of myocardial scar tissue^[13] and fibroblast-populated collagen gels, and Sacks^[25] has employed measured fiber orientations in structural models of heart valves and pericardium. Matching experimental data using this approach is generally more difficult than with a phenomenologic approach; however, it directly challenges us to understand how tissue composition and organization give rise to mechanical properties, rather than just describing those properties in curves and equations.

Growth and Remodeling of the Heart

The geometry, structure, and composition of the heart are not static; the heart can adapt and remodel over time. Remodeling of the heart occurs during normal physiologic growth, during exercise training, and during many pathophysiologic processes. In particular, hemodynamic overload is known to induce growth of cardiac myocytes in distinctive patterns that depend on the type of overload. Experimentally increasing the pressure the left ventricle must generate with each beat, causes myocytes in the wall to grow preferentially wider; integrated over the entire left ventricle, the net result of this pressure-overload hypertrophy is an increased wall thickness and decreased ratio of radius to wall thickness. By contrast, increasing the volume of blood the heart must pump per minute causes myocytes in the wall to grow preferentially longer, the left ventricular cavity radius to enlarge, and the radius-to-wall thickness ratio to increase. It has long seemed evident that local mechanical signals sensed by the cells are likely important regulators of myocyte growth and



remodeling,^[26] but the specific mechanical stimuli that are sensed have not been clearly identified.^[27]

Cardiac biomechanics is making two important contributions to the understanding of cardiac growth and remodeling. First, the fact that the stresses in the left ventricular wall depend not only on hemodynamics but also on ventricular geometry has turned out to be of immense clinical importance for patients with heart failure. As discussed above, it is not possible to completely separate the effect of ventricular geometry on wall stresses from other factors such as material properties. However, it is clear that for a given cavity pressure and set of material properties, increasing the cavity radius or decreasing the left ventricular wall thickness will increase circumferential and longitudinal stresses in the wall. Therefore, the ratio of cavity radius to wall thickness is often used as a rough index of the relationship between cavity pressure and wall stress. In patients with dilated heart failure, the radius-to-wall thickness ratio is much higher than normal. The consequence is that myocytes feel a much larger systolic stress, or afterload, than myocytes in a normal heart at a given systolic cavity pressure; these myocytes are at a mechanical disadvantage when they try to contract. This means that patients with heart failure actually have two problems. The first is cellular: The cells have impaired calcium handling and force generation. But the second is geometric: Even if a dilated heart were composed entirely of normal myocytes, its function would be depressed purely because of the mechanical disadvantage imposed by abnormal geometry.

This insight has been very important to understanding and planning new therapies for dilated heart failure that have emerged in recent years. For example, a surgery known as the Batista procedure attempted to remove the geometric disadvantage for patients with severe heart failure by excising a longitudinal wedge of myocardium and joining the edges of the resulting defect, instantaneously reducing left ventricular radius.^[28] Although that surgery has recently fallen out of favor, a number of other procedures and devices currently under testing are based on the same concept of reducing or constraining cavity radius as a way to recover or preserve left ventricular function. Another strategy that appeared promising was temporary mechanical unloading of the left ventricle with left ventricular assist devices (LVADs). It was discovered that when patients in end-stage heart failure are placed on temporary LVAD support while awaiting heart transplantation, the ventricular remodeling process appeared to reverse, with dramatic reductions in heart and myocyte size over about a month of LVAD support.^[29] While some beneficial remodeling of cellular function also occurred during LVAD support, it turned out that the ratio of radius to wall thickness did not change in these hearts, suggesting that the

fundamental mechanical disadvantage associated with abnormal geometry in failing hearts is not improved by LVAD-induced reverse remodeling.^[30]

The availability of therapies like LVADs and myocardial constraint devices has opened up another opportunity for the field of cardiac mechanics to contribute to improved treatment strategies for heart failure. While most of the attention surrounding hypertrophy in the cardiology community is currently focused on the molecular mechanisms and signaling pathways underlying hypertrophy and remodeling, engineers have focused more on identifying the mechanical stimuli that initiate these cascades. Until recently, it could be argued that the molecular approaches were much more likely to lead to therapeutic advances as the pathways and mechanisms, once identified, could serve as targets for pharmacologic blockade or modification. However, with the advent of new therapies allowing the mechanical environment of the myocytes to be prescribed, the reverse argument may be made: If the mechanical inputs to the heart can be controlled and the inputs needed to achieve a particular pattern of remodeling specified, remodeling can be directed to achieve desired changes even before the underlying mechanisms are understood.

In the past, the primary approach that cardiac mechanics investigators have taken in identifying the mechanical stimuli responsible for cardiac growth and remodeling is to hypothesize which stimuli drive growth and remodeling, then set up adaptive models to determine whether the hypothesized “growth law” produces similar responses to those observed experimentally during various types of hemodynamic overload.^[31,32] These growth models are very useful screens for identifying potentially important mechanical stimuli. The success of one of these models in reproducing observed growth responses illustrates that a particular combination of stimuli could drive hypertrophy and remodeling, but does not prove that these stimuli in fact do so. A combination of growth models and direct *in vitro* tests of mechanical perturbations will be required to definitively identify the relevant stimuli and quantify the relationships between those stimuli and resulting growth.

CONCLUSIONS

The field of cardiac mechanics seeks to apply the methods and ideas of solid mechanics to understand the function of normal and diseased hearts. Dramatic advances have been made in measuring deformation in the heart, estimating stresses acting on the myocardium using sophisticated finite element models, and characterizing the material properties of myocardium.

The tools of cardiac mechanics analysis are now increasingly being applied to more fundamental questions such as how tissue structure gives rise to material properties and how mechanical signals drive growth and remodeling of the normal and diseased heart.

FURTHER READINGS

Owing to space limitations, only a small fraction of the relevant literature is referenced here, often giving preference to recent reviews rather than original sources on the argument that these reviews are more convenient starting places for a reader to begin learning about the field of heart biomechanics. In particular, I would like to recommend the following books and special journal issues as the best places to quickly obtain a more detailed overview of the problems, methods, and ideas that comprise the field. First, the history and basic principles of cardiac physiology, as well as some of the best early engineering analysis of cardiac function, are beautifully reviewed in the book “Cardiac Contraction and the Pressure-Volume Relationship” published by Sagawa et al. in 1988.^[1] Second, the book “Theory of Heart: Biomechanics, Biophysics, and Nonlinear Dynamics of Cardiac Function” contains a collection of articles reflecting the state of the art in cardiac mechanics research at the time of its publication in 1991.^[33] Next, in 1998 the journal *Progress in Biophysics and Molecular Biology* published a special issue on a range of topics in cardiac mechanics (Volume 69, Issues 2–3, pp. 151–572). In 2000, the *Journal of Elasticity* published a special issue on soft tissue mechanics in the cardiovascular system that contained several articles on the heart and on constitutive modeling (Volume 61, Numbers 1–3). Most recently, in 2001 the journal *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* published an excellent special issue on computational models of the heart (Volume 359, Number 1783, pp. 1047–1337). Finally, to place readings on heart biomechanics into clinical context, it may be useful to consult a general textbook on cardiovascular medicine such as Braunwald’s *Heart Disease*.

ACKNOWLEDGMENTS

The author would like to acknowledge colleagues in the Cardiac Biomechanics Group who provided data for this article and/or critical readings during its development: Evren Azeloglu, Dr. Kevin Costa, Greg Fomovsky, Charlie Haggart, and T. Alex Quinn.

ARTICLES OF FURTHER INTEREST

Bioreactors; Blood Vessel Mechanics; Tissue Engineering of Cardiac Muscle; Tissue Engineering of Heart Valves

REFERENCES

1. Sagawa, K.; Maughan, L.; Suga, H.; Sunagawa, K. *Cardiac Contraction and the Pressure-Volume Relationship*; Oxford University Press: Oxford, 1988.
2. Hunter, P.J.; McCulloch, A.D.; ter Keurs, H.E. Modelling the mechanical properties of cardiac muscle. *Prog. Biophys. Mol. Biol.* **1998**, *69* (2–3), 289–331.
3. Cooper, G. Load and length regulation of cardiac energetics. *Ann. Rev. Physiol.* **1990**, *52*, 505–522.
4. Yin, F.C.P. Ventricular wall stress. *Circ. Res.* **1981**, *49* (4), 829–842.
5. Omens, J.H.; Fung, Y.-C. Residual strain in rat left ventricle. *Circulation Research* **1990**, *66* (1), 37–45.
6. Streeter, D.D., Jr.; Hanna, W.T. Engineering mechanics for successive states in canine left ventricular myocardium. I. Cavity and wall geometry. *Circ. Res.* **1973**, *33* (6), 639–655.
7. Streeter, D.D., Jr.; Hanna, W.T. Engineering mechanics for successive states in canine left ventricular myocardium. II. Fiber angle and sarcomere length. *Circ. Res.* **1973**, *33* (6), 656–664.
8. LeGrice, I.; Hunter, P.; Young, A.; Smaill, B. The architecture of the heart: a data-based model. *Phil. Trans. Royal Soc. Lond. A* **2001**, *359*, 1217–1232.
9. Stevens, C.; Remme, E.; LeGrice, I.; Hunter, P. Ventricular mechanics in diastole: material parameter sensitivity. *J. Biomech.* **2003**, *36* (5), 737–748.
10. Vetter, F.J.; McCulloch, A.D. Three-dimensional analysis of regional cardiac function: a model of rabbit ventricular anatomy. *Prog. Biophys. Mol. Biol.* **1998**, *69*, 157–183.
11. McCulloch, A.D. Cardiac biomechanics. In *The Biomedical Engineering Handbook*, 2nd Ed.; Bronzino, J.D., Ed.; CRC Press: Boca Raton, FL, 2000; 28.21–28.26.
12. Lin, D.H.; Yin, F.C. A multiaxial constitutive law for mammalian left ventricular myocardium in steady-state barium contracture or tetanus. *J. Biomech. Eng.* **1998**, *120* (4), 504–517.
13. Costa, K.D.; Holmes, J.W.; McCulloch, A.D. Modelling cardiac mechanical properties in three dimensions. *Phil. Trans. Royal Soc. Lond. A* **2001**, *359*, 1233–1250.
14. Demer, L.L.; Yin, F.C.P. Passive biaxial mechanical properties of isolated canine myocardium. *J. Physiol. Lond.* **1983**, *339*, 615–630.
15. Dokos, S.; Smaill, B.H.; Young, A.A.; LeGrice, I.J. Shear properties of passive ventricular myocardium. *Am. J. Physiol. Heart Circ. Physiol.* **2002**, *283* (6), H2650–H2659.



16. Humphrey, J.D.; Strumpf, R.K.; Yin, F.C.P. Determination of a constitutive relation for passive myocardium: I. A new functional form. *J. Biomech. Eng.* **1990**, *112*, 333–339.
17. Humphrey, J.D.; Strumpf, R.K.; Yin, F.C.P. Determination of a constitutive relation for passive myocardium: II. Parameter estimation. *J. Biomech. Eng.* **1990**, *112*, 340–346.
18. Guccione, J.M.; McCulloch, A.M.; Waldman, L.K. Passive material properties of intact ventricular myocardium determined from a cylindrical model. *J. Biomech. Eng.* **1991**, *113*, 42–55.
19. Omens, J.H.; MacKenna, D.A.; McCulloch, A.D. Measurement of strain and analysis of stress in resting rat left ventricular myocardium. *J. Biomech.* **1993**, *26* (6), 665–676.
20. Holmes, J.W.; Borg, T.K.; Covell, J.W. Structure and mechanics of healing myocardial infarcts. *Ann. Rev. Biomed. Eng.* **2005**, *7*, 223–253.
21. Criscione, J.C.; Sacks, M.S.; Hunter, W.C. Experimentally tractable, pseudo-elastic constitutive law for biomembranes: I. Theory. *J. Biomech. Eng.* **2003**, *125* (1), 94–99.
22. Criscione, J.C.; Sacks, M.S.; Hunter, W.C. Experimentally tractable, pseudo-elastic constitutive law for biomembranes: II. Application. *J. Biomech. Eng.* **2003**, *125* (1), 100–105.
23. Hunter, P.J. Myocardial constitutive laws for continuum mechanics models of the heart. *Adv. Exp. Med. Biol.* **1995**, *382*, 303–318.
24. Horowitz, A.; Lanir, Y.; Yin, F.C.P.; Perl, M.; Sheinman, I.; Strumpf, R.K. Structural three-dimensional constitutive law for the passive myocardium. *J. Biomech. Eng.* **1988**, *110*, 200–207.
25. Sacks, M.S. Incorporation of experimentally-derived fiber orientation into a structural constitutive model for planar collagenous tissues. *J. Biomech. Eng.* **2003**, *125* (2), 280–287.
26. Grossman, W.; Jones, D.; McLaurin, L.P. Wall stress and patterns of hypertrophy in the human left ventricle. *J. Clin. Invest.* **1975**, *56*, 56–64.
27. Holmes, J.W. Candidate mechanical stimuli for hypertrophy during volume overload. *J. Appl. Physiol.* **2004**, *97* (4), 1453–1460.
28. Batista, R.J.; Verde, J.; Nery, P.; Bocchino, L.; Takeshita, N.; Bhayana, J.N.; Bergsland, J.; Graham, S.; Houck, J.P.; Salerno, T.A. Partial left ventriculectomy to treat end-stage heart disease. *Ann. Thorac. Surg.* **1997**, *64* (3), 634–638.
29. Burkhoff, D.; Holmes, J.W.; Madigan, J.; Barbone, A.; Oz, M.C. Left ventricular assist device-induced reverse ventricular remodeling. *Prog. Cardiovasc. Dis.* **2000**, *43*, 19–26.
30. Barbone, A.; Oz, M.C.; Burkhoff, D.; Holmes, J.W. Normalized diastolic properties after left ventricular assist result from reverse remodeling of chamber geometry. *Circulation Supplement* **2001**, *104* (suppl I), I-229–I-232.
31. Arts, T.; Prinzen, F.W.; Snoeckx, L.H.E.H.; Rijcken, J.M.; Reneman, R.S. Adaptation of cardiac structure by mechanical feedback in the environment of the cell: a model study. *Biophys. J.* **1994**, *66*, 953–961.
32. Lin, I.E.; Taber, L.A. A model for stress-induced growth in the developing heart. *J. Biomech. Eng.* **1995**, *117* (3), 343–349.
33. Glass, L.; Hunter, P.; McCulloch, A. *Theory of Heart: Biomechanics, Biophysics, and Nonlinear Dynamics of Cardiac Function*; Springer-Verlag: New York, 1991.

Heart Valve, Bioprosthetic

Peter Zilla

Paul Human

Deon Bezuidenhout

Christiaan Barnard Department of Cardiothoracic Surgery, University of Cape Town Medical School, and Groote Schuur Hospital, Cape Town, South Africa



INTRODUCTION

In today's world, the extraordinary composer Gustav Mahler would not have died of aortic valve disease. Instead, he would have become one of the more than 250,000 estimated annual recipients of a prosthetic heart valve.

This ability to replace a diseased heart valve with a prosthetic one has dramatically reduced the mortality associated with heart valve disorders. Unfortunately, this shining success story of modern medicine is only one side of the coin. The other one is a sobering complication rate associated with all contemporary heart valve prostheses. While the life-extending effect of replacement valves is undisputed, there is a high price to be paid in lost quality of life. If the prosthetic valve is made of titanium and/or pyrolytic carbon it requires life-long anticoagulation. In spite of decades of experience, anticoagulation is still a dangerous treatment. Suboptimal anticoagulation carries the threat of major clot formation, either leading to catastrophic immobilization of the valve leaflets or the downstream occlusion of a major blood vessel. If this blood vessel supplies the brain, the consequence is a stroke. In contrast, therapeutic and suprathreshold anticoagulation has a significantly increased risk of hemorrhage. In the younger population group, this risk prevents otherwise perfectly healthy individuals from living a physically active sports life. In the elderly, this bleeding risk often comes on top of hypertension and fragility of blood vessels. The consequence is often a life-threatening intracerebral hemorrhage. Against this background, there was great enthusiasm when heart valve prostheses were introduced that were made of bovine or porcine tissue and did not require anticoagulation. Unfortunately, it soon became evident that these bioprosthetic heart valves eventually degenerate and fail. While the relatively short life span of such tissue valves is mostly compatible with the natural life expectancy of older patients, it is seen to be

prohibitive in younger recipients. Moreover, the process of bioprosthetic degeneration is also much faster in young patients than in senescent ones.

Disappointingly, most of the limitations of replacement valves were already apparent 20 years ago. Since then, no fundamental breakthrough has occurred that addresses the underlying mechanisms of failure. On the contrary, the basic principles of tissue valve production remain largely unchanged. Improvements were rather of a symptomatic than of a principal nature. Post-fixation treatments aiming at the suppression of detrimental leaflet calcification, for instance, were partly successful but failed to provide a satisfying explanation concerning the mechanism of their effect. Therefore, the challenge with regard to improved tissue valves lies in our ability to unravel the key mechanisms determining the interaction of porcine or bovine tissue with the human recipient.

When elucidating these key mechanisms, one needs to keep in mind that there is already a conceptual contradiction that overshadows the idea of bioprosthetic heart valves: On the one hand, nature has provided us with the most compelling design of heart valve leaflets. These tissue membranes are only a few hundred micrometers thick but capable of withstanding the harsh mechanical forces of 3 billion heart cycles in the course of a lifetime. On the other hand, by being forced to choose animal tissue for the production of replacement valves, many of the advantages of native tissue are being lost again. The reason for this unwanted consequence lies primarily in the necessity to mask the immunogenicity of the animal tissue. If untreated, such heart valves would instantly be subjected to acute rejection and fail within days. The chemical cross-linking processes, however, which were adopted to eliminate immunogenicity, defy many of the reasons for choosing native tissue. The delicate mechanical function of natural leaflets, for instance, with their optimal stress-reducing design is partly lost by the cross-linking process. Moreover, these

cross-links were almost exclusively glutaraldehyde-based throughout the past three decades. The highly complex chemistry of this bifunctional aldehyde introduces chemical components that are thought to contribute to the propensity of the prosthetic leaflets to calcify. Most importantly, the delicate leaflet structure that physiologically withstands disproportional mechanical forces relies on the repair capabilities of the living tissue. Most regrettably, the cross-linking process, which is required to reduce immunogenicity, turns the living tissue into a nonvital structure devoid of such repair mechanisms.

In spite of all these critical observations, however, tissue heart valves are a blessing for many patients. Seen against the background of an increasingly critical attitude toward the complications of mechanical heart valve prostheses, bioprosthetic valves have captured a continually growing market share. In contrast to the exuberant enthusiasm of the 1970s and early 1980s, however, the positive driving forces behind tissue valves are much more sober today. Both clinicians and researchers have come to the realization that the only heart valve prosthesis that will ever be truly durable and free of serious complications will be one that consists of living and physiologically functioning recipient tissue. Yet, for those who have observed the slow pace of tissue engineering over the past 20 years, it is evident that no such heart valve will be clinically available for the majority of patients for a long time. This leaves us with the task of optimizing our dual armament of mechanical and biological heart valve prostheses for the foreseeable future. This is a formidable task, given the fact that conservatism has prevailed in this field for the past three decades. Characteristically, the only major deviation from the concept of the 1970s was the step toward unstented bioprostheses. Therefore, optimization needs to question more than the choice of a sewing ring, the transvalvular gradient or the incorporation of another anticalcification molecule. Optimization needs to start with the reassessment of the clinical niches for tissue valves. As much as the indications for tissue valves were too generously applied in the late 1970s and early 1980s, today's narrow limitation to the elderly represents resignation rather than optimization. Equally important to decide is whether we have achieved the main objective of tissue valve treatment, namely the suppression of xenograft immunogenicity, in the first place. If the answer is no, we ought to embrace this aspect with bold openness, bringing in the scientific tools that were not available when the idea of bioprosthetic heart valves was initially conceived. One such promising tool is today's ability to extract cellular antigenicity largely from the tissue. Another is the deeper insight into the complex world of cross-linking chemistry. An array of new approaches to tissue

fixation not only widens the spectrum of antigen masking, but by chemically engineering cross-links, tissue fixation becomes more controllable, thereby allowing the addressing of adverse aspects such as tissue stiffness or cytotoxicity.

CLINICAL NEED FOR TISSUE VALVES

The number of heart valve replacements performed annually has continued to rise over the past few years. While an estimated 175,000 heart valves were replaced in 1995,^[1] it was thought to be more than 250,000 five years later. Apart from this absolute increase in annual procedures, the share of bioprosthetic valves rose from 20% in 1995 to 40% in 2000 and is currently believed to be at about 50% of the overall sales. There is reason to believe that this share of the annual heart valve market represents a ceiling—at least as long as there is no fundamentally improved concept behind tissue valves. In return, such an improved concept requires research and development resources that would become available only if there was a commercial incentive. Therefore, an assessment of the short- to medium-term growth potential of the bioprosthetic heart valve market could provide us with an indication for the likelihood of such a development.

If one looks at the current indications for tissue valves, the majority are implanted into first-world patients of more than 65 years of age. Within this age group, practically all patients are already receiving bioprosthetic valves. A certain proportion is also used in young patients in developing countries where anticoagulation is difficult. However, sales in these regions are often subsidized. Thus, revenue with tissue valves is almost exclusively generated in the geriatric patients of Europe, North America, and Japan. It is therefore unlikely that the demand in the developing world will represent a market force in the near future. This is particularly regrettable in view of the needs of third world countries. Of a world population of almost 6.5 billion people, 5.5 billion live in developing countries. Typically, their emerging middle class is increasingly successful in persuading their governments to establish open-heart surgery because it's easier than tackling rheumatic fever. The combination of a young age at which these patient typically require heart valve replacement and the often difficult access to anticoagulation control represents the strongest theoretical motivation for the development of a truly durable bioprosthesis. Still, as long as the buying power of these countries remains weak, revenue generation will not reach a level sufficient to create a strong incentive for research and development.

Having conceded the limitations of revenue growth in the developing world, what are the chances that the

first world market grows beyond its core indication of geriatric aortic valve disease? Apart from geoeconomic limitations there are also anatomical limits to a market growth. Since tricuspid- and pulmonary-valve replacements make up a fraction of procedures, a potential expansion of indications will always be restricted to the aortic and mitral valve. Considering the fact that in developed countries diseased mitral valves are increasingly repaired and thus preserved, one can narrow the present picture even further: The majority of revenue-generating bioprosthetic tissue valves are being implanted into elderly, first-world patients suffering from aortic stenosis. There is certainly a good chance that tissue valves may additionally be implanted into aortic stenosis patients below the currently observed age limit of 60–65 years, should improvements modestly increase the life span of contemporary bioprostheses. Extrapolating the developments of the past three decades, such modest improvements are likely. In more than two-thirds of aortic stenosis patients currently falling under the age limit for tissue valves, the etiology of the disease is either degenerative or on the basis of an underlying bicuspid malformation. Even if younger than 65 years at the time of operation, these patients still tend to be beyond their prime when they need a replacement valve. The situation is different in the remaining third of patients with aortic stenosis, where the etiology is rheumatic. In contrast to patients with degenerative or bicuspid aortic stenosis, these patients often become symptomatic at a distinctly younger age. In a first-world environment, it would require fundamentally improved bioprosthetic heart valves to justify the implantation of tissue valves into this age group. Similarly, a majority of aortic incompetence patients falls into this category. Two-thirds are either of rheumatic etiology or caused by infective endocarditis. These patients usually become symptomatic between the 4th and the 5th decade. The remaining third of patients with aortic incompetence are suffering from a disease that affects the entire aortic root. If it is a connective tissue disease such as Marfan's disease, patients often become symptomatic as early as in their thirties. The smaller group of ankylosing spondylitis patients are normally in their mid-forties and that with syphilitic aortitis in their early fifties when symptoms may require surgery. The prospects for contemporary tissue valves look even less promising in mitral valve disease. As mentioned before, an increasing number of mitral valve disorders needing surgical intervention do not end up with a replacement valve. If the underlying condition is mitral stenosis, the etiology is rheumatic and the mostly female patients need a therapeutic procedure in their 4th and 5th decade. In pure mitral stenoses, balloon valvuloplasties have largely replaced open valvulotomies. Those needing a

valve replacement because of mixed mitral stenosis and incompetence are far too young for bioprostheses, even if a moderately improved durability of current tissue valves can be achieved. This leaves mitral incompetence as a last possible indication. Here, the age at which the disease becomes symptomatic again excludes current bioprosthetic heart valve concepts categorically: If the etiology is rheumatic, those needing a valve replacement often need it in their fourth decade. Apart from the young age group of these patients, surgical repair is increasingly considered for rheumatic mitral incompetence, even if it still remains a matter of controversy. Similarly, infective endocarditis mainly occurs in young to middle-aged patients. Thus, only mitral valve prolapse and ischaemic mitral incompetence are conditions that need surgical intervention from the 5th to 6th decade onwards, but most of these valves are also repairable.

What becomes obvious from such an assessment is the fact that a quantum leap in bioprosthetic heart valve development would be required to significantly increase the current market share of tissue valves. The only possible way of challenging this conclusion is to revisit our current approach to risk factor assessment. In order to do so, we first need to remind ourselves that there is only one reason for choosing a mechanical heart valve prostheses: the fear of reoperation as a consequence of tissue valve failure. In the decision-making process regarding the choice of prosthetic heart valves, this fear of a reoperation may bias our overall assessment of risk factors. There is reason to believe that such a bias does exist. Although dependent on the experience of a center, the mortality associated with reoperations seems lower than generally assumed. In centers routinely performing heart valve reoperations, the surgical mortality is only marginally higher than in first procedures.^[2] Held against the often life-threatening complications in patients with mechanical heart valves the reevaluation of risk factors may dramatically expand the indications for bioprosthetic heart valves. If such a reevaluation of risk factors resulted in the acceptance of one reoperation in a patient's lifetime, the majority of heart valve recipients in the developed world would be candidates for tissue valves. Furthermore, bioprosthetic heart valves have one other potentially life saving advantage: They usually degenerate gradually. In contrast to the catastrophic emergency a clotted or stuck mechanical valve presents, reoperations can be largely planned in tissue valves, although acute leaflet ruptures were also reported in early bioprostheses, albeit as extremely rare events.^[3,4]

Nevertheless, as long as the reoperation risk is seen to be higher than the risks of anticoagulation, the tissue valve market remains near its ceiling. With a life expectancy of close to 80 years in the developed world



and a mean life span of approximately 15 years for tissue valves, the margins for expansion are narrow. With a trend towards the lowering of the age threshold for tissue valves from 65 to 60, most of the nonrheumatic aortic stenosis patients needing surgery may eventually receive a bioprosthetic heart valve. Moderate improvements on the scale seen in the past 20 years may add the few extra years in durability, which would benefit those aortic stenosis patients blessed with particular longevity. For the vast majority of other indications, current tissue valves would need to at least double their life span. To achieve this goal, a fundamentally improved approach to tissue preservation would be required. Given the fact that most of the commercial producers of bioprostheses are also the market leaders in mechanical heart valves, the incentive for such an extraordinary research effort seems low. Seen against the background of a 30% risk of developing a thromboembolic complication within 20 years and an equally high 30% risk of experiencing a hemorrhage with a mechanical prosthesis,^[5] this prospect is rather dim.

HISTORICAL REVIEW

One of the driving forces behind open-heart surgery was the desire to replace diseased heart valves. When the pioneering efforts of Gibbon, Lillehei, and Kirklin in the first half of the 1950s eventually made it possible to access these valves openly, no prosthetic replacements were available. It took a few more years until simple mechanical prostheses enabled the first aortic valve replacement in 1960 and the first mitral valve replacement in 1961.^[6,7] The drive toward a biological replacement valve that does not require anticoagulation occurred almost simultaneously. As early as in 1961, Robert Frater began experimentally using autogenous pericardium for free-hand valves or parts of valves. Shortly thereafter, he developed a pantaloons-shaped substitute for the posterior mitral cusp out of autogenous pericardium that soon found its way into clinical practice.^[8,9] In 1962, Ross reported the first successful replacement of an aortic valve with an allogenic cadaver valve.^[10] The possibility of an antigenic response to these homografts was anticipated and soon confirmed by Mohri^[11] and Davies.^[12,13] However, since only the aortic wall portion of homograft valves seemed to degenerate with clinical consequence, not the leaflets themselves, the incentive toward an immune-quiescent tissue valve was not very strong. As a consequence, the push toward autogenic bioprostheses by Donald Ross^[14] was not carried by a strong enough clinical demand. Although the use of the patient's own pulmonary valve to replace his diseased aortic valve was first described in 1967, it took

more than two decades until the idea of autografts gained a significant following. On the contrary, the routine introduction of open-heart surgery created a worldwide bottleneck for allografts. It was therefore soon realized that this increasing demand for bioprosthetic heart valves could be satisfied only if animal tissue (xenografts) were acceptable. By denaturing the xenogenic proteins with mercurial solutions, freeze-drying, or formalin treatment, the masking of antigenicity was thought to be sufficient to allow the use of porcine tissue.^[15] This led to the first successful xenograft replacement of an aortic valve in a human in 1965.^[16] Although mechanical issues such as the mounting of the fragile muscular part of the right coronary cusp were subsequently resolved, it soon became evident that remnant immunogenicity was unacceptably high.^[17] It was concluded that fixation with monofunctional formaldehyde resulted in gradual detanning^[18] of the tissue, thereby unmasking immune and degradation sites. This led to the introduction of glutaraldehyde as a bifunctional aldehyde that resulted in distinctly longer durability of the prostheses than the one observed in early formalin-fixed valves. Two laboratories began independently experimenting with the use of glutaraldehyde for xenograft preservation: Robert and Carpentier in Paris and Hancock and Nimni in Los Angeles. The first such glutaraldehyde-preserved porcine aortic valve was clinically implanted in 1969. When it became obvious that glutaraldehyde-preserved xenograft valves showed signs of dystrophic calcification a multitude of anticalcification treatments were added without thoroughly understanding the mechanisms behind tissue mineralization. Assuming a central role of phospholipids in tissue calcification, one of the early anticalcification treatments aimed at the removal of lipids. Pioneered at the Hancock laboratories under the name "T-2 procedure" and essentially using a household detergent containing anionic surfactants to remove lipids after completion of fixation, it achieved only a moderate decrease in calcification.^[19] Alternatively, anticalcification treatments aimed at neutralizing the free aldehyde functionalities of glutaraldehyde and its reaction products, primarily via the formation of Schiff's bases with terminal aldehyde groups. Into this category fell amino-diphosphonates (ADP),^[20] amino-acid treatment such as L-glutamic acid developed by Grimm together with Sorin, and Urazole developed by Zilla with Medtronic.^[21] Although experimentally successful, only amino-oleic-acid found its way into broad use in a commercial second-generation tissue valve.^[22]

Apart from fixation and postfixation treatment, bioprosthetic concepts also differed in the source of tissue used. Although a wide range of material was initially evaluated, from human dura mater to fascia lata promoted by Åke Senning in Zurich, it soon boiled

down to the use of two main tissue sources: pericardium and aortic valve leaflets. In the earliest days of bioprosthetic heart valve replacement, pericardial substitutes were autogenous^[9] and aortic leaflet tissue was either allogenic^[10] or autogenic.^[14] Under the pressure of demand, Alain Carpentier initiated the almost uninterrupted advance of porcine valve bioprostheses in 1968. It did not take long, however, until the advantages of fully engineering the design of a tissue valve led to glutaraldehyde-fixed bovine pericardial prostheses. A boom in pericardial xenografts ensued, resulting in 10,000 clinical implants of Ionescu-Shiley valves within two years at Denton Cooley's Texas Heart Institute alone.^[23] Due to one single design error—namely the so-called “alignment stitch” at the commissure—these valves failed prematurely and discredited bovine pericardium for years. Although Carpentier addressed this issue together with Edwards from the beginning of the 1980s onward by mounting the tissue underneath the stent and not above, the surgical community remained skeptical until favorable clinical seven-year results were available.^[24] However, the woes of pericardial valves were not over yet. The flexing pericardial cusps rubbed against the cover fabric and accelerated ruptures.^[25] Eventually, by the mid 1990s it seemed as if pericardial valves were back apparently with results equal to those of porcine valves.^[26] This rehabilitation of pericardial valves allowed another development in the 1990s, which touched on a taboo of more than two decades: the possibility that insufficiently suppressed immunogenicity played a major role in xenograft failure. Autogenics developed a kit that allowed the surgeon to produce a bioprosthetic heart valve out of autologous tissue in the operating room.^[27,28] Unfortunately, the original concept of living autologous heart valve tissue soon had to be abandoned when the environmental stimulation of pericardial cells led to contracted and scarred leaflets. The subsequent inclusion of a fixation step—albeit brief and superficial—also eliminated the advantages of living tissue while reintroducing mechanical disadvantages. Although early failures led to the demise of the Autogenics concept, this approach for the first time addressed remnant antigenicity as a key issue. These ups and downs of pericardial heart valves are not unparalleled in the camp of porcine aortic valve prostheses. Inspired by the ability to produce lower profile valves with pericardium, St. Jude Medical developed a low profile version of a porcine xenograft. The increased angle of the free edge led to an increase in leaflet stress and the subsequent early failures of these valves resulted in the abandoning of the concept. While this failure eroded the bastion of porcine valve concepts through a design shortcoming, a seemingly novel alternative design approach recaptured clinical following in return. Initially driven by Biocor and

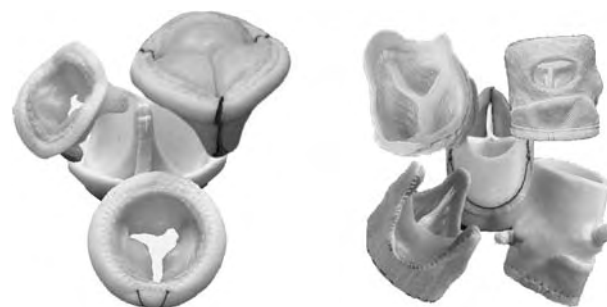


Fig. 1 Representation of modern bioprosthetic heart valves, with stented (left) and stentless models (right).

Tyrone David together with St. Jude (Toronto SPV), the aim was to reduce transvalvular gradients by eliminating the stent. Soon this concept was emulated by Medtronic (Freestyle) and Baxter (Prima) with the difference of representing more of a root prosthesis in the latter two (Fig. 1). While the advantage of a less obstructed outflow orifice remains undisputed, the other advantage of stentless valves, namely mobile commissures, soon gets lost by the distinct calcification of the aortic wall tissue of these prostheses. Nevertheless, stentless valve concepts have gained a significant following particularly in Europe and explants so far confirm that cusp mineralization and tissue degeneration at periods up to eight years are minimal.^[29–32]

All in all, the first generation of bioprosthetic heart valves appeared on stage with a bang but withdrew into early retirement soon after. Today's tissue valves differ relatively little from those of the late 1960s. The hailed new concept of stentless aortic valve prostheses is an adaptation of the homograft approach of the early 1960s to xenografts. Apart from fixing leaflet tissue in a way that preserves the natural crimp and from adding anticalcification treatments after fixation, bioprosthetic heart valves of the third millennium continue to be made of glutaraldehyde-fixed porcine aortic leaflets or bovine pericardium suspended on stents that do not dramatically differ from earlier designs. Only time will tell whether these minor improvements have a significant impact on the longevity of the prostheses. The long observation periods necessary for arriving at a conclusion certainly do not aid the speedy development of new concepts that would be required for a major breakthrough.

FAILURE OF BIOPROSTHETIC HEART VALVES

Although improvements to bioprosthetic heart valves were only moderate and incremental during the past three decades, the trend regarding primary tissue



failure points downward. In one of the biggest early series of first-generation porcine tissue valves implanted into 1648 patients between 1970 and 1980, a 10-year freedom from structural failure of 79% was found in the aortic position and 72% in the mitral.^[33] A comparable study was performed 12 years later, using second-generation porcine valves that had already undergone an anticalcification treatment.^[34] Ten-year freedom from structural failure was found in 92% of aortic valve replacements and 81% of mitral valve replacements. Although this 9% and 13% reduction in structural failure may overstate the actual situation because patients were almost five years younger in the first series, one can extrapolate from experimental studies^[35] that the introduction of more effective anticalcification treatments and the improved processing conditions in third-generation tissue valves have further reduced the incidence of structural failure. Although long-term clinical studies are not yet available, explants for nonvalve-related failures indicate that leaflet calcification is distinctly less than in previous series. In the mechanically less challenging aortic position, 10-year freedom from structural failure is thus not far from 100%, provided the patient selection remains above the cut-off age of 65 years. Recent reports on contemporary pericardial valves estimated the freedom from structural valve failure to be 97% at 17 years in aortic position and 94% at 14 years in mitral position.^[36]

The overwhelming failure mode of bioprosthetic heart valves is primary tissue degeneration. In the majority of cases, primary tissue failure leads to valve incompetence due to tears. In the mechanically more stressed mitral position, 75% of failed porcine prostheses showed rupture of a free cusp edge and 43% of the cusp belly.^[37] In only a small minority of bioprostheses does stenosis predominate as a result of excessive calcific degeneration. Overall, less than half of all failed first-generation tissue valves showed signs of calcification while cuspal perforations were the prevalent mode of failure.^[38,39] The relatively successful introduction of anticalcification treatments—at least as far as cusps are concerned—in second-generation prostheses promises that the balance between calcification and tissue degradation will shift even more toward degradation.

Generally, bioprosthetic degeneration starts long before implantation. During the initial 48–72 hours of transport between the abattoir and the plant, serious autolytic changes occur in the tissue, leading to the loss of tissue integrity primarily on a cellular level.^[40] The ability of these preimplantation damages of the bioprosthetic materials to augment calcific degeneration was demonstrated by various authors.^[41,42] Hypoxic injuries to cells concentrate a large amount of calcium in protrusions of the plasma membrane prior to

blebbing and shedding of membrane vesicles.^[43] Even under in-vitro conditions, hypoxically injured valvular interstitial cells calcified while uninjured controls did not.^[44] This in-vitro calcification emulated the one found in explanted bioprostheses by being localized in mitochondria and extracellular blebs. Mitochondria exposed to hypoxia were shown to undergo cyclic changes of swelling and shrinkage, followed by flocculation of their matrix. These densities are formed by lipolysis and signal the irreversibility of cell injury.^[45] Mitochondrial calcification is known to occur in association with these flocculent densities.^[46]

Upon implantation, plasma insudation into the tissue leads to the unphysiological deposition of blood components both inside the biomaterial and on its surface. Leukocyte adhesion and infiltration dominated by macrophages,^[47] destruction or separation of collagen fibers,^[48] and calcification follow. Apart from degenerative damages, these alterations also lead to a stiffening and thickening of the cusp tissue^[47] with the concomitant further loss of mechanical function. Other than the overgrowth of thick anastomotic intimal hyperplasia, tissue generally referred to as pannus, no endothelialization occurs with host endothelium, even after years of implantation and regardless of whether the prosthesis is stented^[49] or stentless.^[50] This lack of a nonthrombogenic surface endothelium facilitates microthrombi formation. After two months of implantation, extensive fibrin layers are found on the blood surface that are gradually replaced by a thinner fibrin coverage^[48] and mononuclear cells. Secondary mineralization of these surface appositions is referred to as extrinsic calcification. However, as much as all these aspects contribute to the overall failure of bioprosthetic tissue, there are principally four phenomena that are observed in connection with failing tissue valves: calcification, inflammation, mechanical damage, and pannus overgrowth, or a combination thereof (Fig. 2). Although all four are distinct events often occurring in isolation, they are capable of potentiating each other if they coexist. Calcification, for instance, is augmented in areas of mechanically induced collagen fractures. Similarly, inflammation is locally enhanced

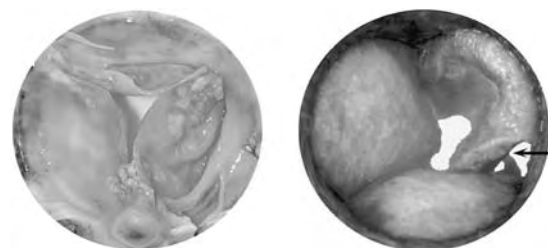


Fig. 2 Typical appearance of failed tissue valves with calcification and pannus outgrowth (left) and degenerative tears (right) as the predominant lesion.

at sites of mechanical injuries and underneath the shear-force protection of pannus tissue.

Dystrophic Calcification

Prior to the advent of stentless xenografts, calcification was primarily a phenomenon of the cusps, and to a much lesser degree, of the wall remnants at the commissures. Although cusp calcification may occur at any site, it is usually seen in the commissural region and the basal region of the cusp belly. With the emergence of stentless xenograft valves, however, aortic wall calcification—a failure mode previously restricted to homografts—also became an issue. In spite of anticalcification treatments, wall calcification remains uninhibited and serious, defying one of the main arguments for stentless valves, namely commissure mobility. In contrast to the preferential location of cusp calcification, wall calcification affects the entire tubular segment of root prostheses.

Histologically, the differences in calcification patterns between leaflets and aortic wall tissue are even more distinct. While wall calcification of glutaraldehyde-fixed bioprostheses affects the inner and the outer layer of the tissue in a sandwichlike manner, cusp calcification occurs predominantly centrally in the spongiosa (Fig. 3). Differences are also seen on an ultrastructural level although the primary nucleation sites are mainly membranes of cell remnants or vesicular debris, regardless of the type of bioprosthetic tissue. In cusps and in pericardium, mineral deposits grow relatively uninhibited from the cellular nucleation sites into the loose collagenous structures of the vicinity. Overall, the morphology of calcium deposits is remarkably similar in porcine cusps and bovine

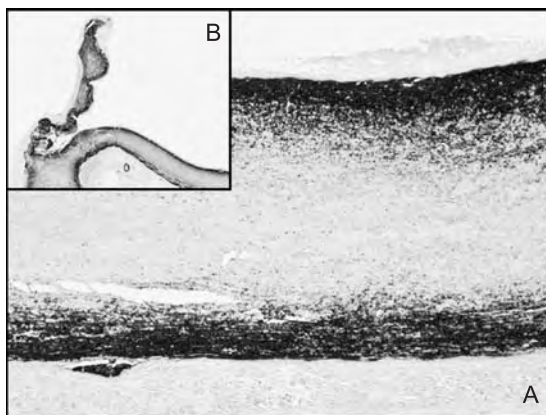


Fig. 3 Characteristic calcification pattern of a glutaraldehyde-fixed root prosthesis; calcium deposits are near the surfaces in the aortic wall (A), but centrally located in the leaflet spongiosa (B), often filling the entire cross-sectional area of the cusp.

pericardium.^[51] In aortic wall tissue, extracellular matrix (ECM) components such as elastin^[52–54] and collagen also participate in the calcification process, yet to a much lesser degree. However, the presence of large numbers of membranous vesicles in close proximity to elastin suggests that even part of the ECM calcification is facilitated by cell membranes.^[55]

Inflammatory Degradation

Every pathologist experienced with bioprosthetic explants is aware of the tattered and eroded appearance of explanted cusps (Fig. 4), with inflammatory cells often residing at the bottom of these excavations. Yet, inflammatory processes wait to be widely recognized as dominant pathogenetic factors of tissue failure partly because “burned out” explants sometimes do not even show traces of inflammatory cells. However, these are usually late observations at the end of the failure process and do not justify the trivialization of the problem. In the vast majority of explanted tissue valves, inflammatory cells were found either to cover the surfaces at almost monolayer density^[47] or focally infiltrate into the tissue.^[56] Since experimental work has shown that the initial macrophage- and plasma-cell-dominated process soon gives way to a more macrophage- and foreign body giant cell-dominated phenomenon, a flair of harmlessness ensued. Given the destructive potential of giant cells to erode even synthetic materials, it becomes obvious that these war-formations of macrophages are serious culprits of destruction. While multinucleation may lead to material failure only in synthetic implants, it appears to endow macrophages with an added powerful resorption capability to degrade and thus present extracellular components.^[57,58] Nevertheless, in a significant number of explants, mononuclear macrophages are still the predominant cell type among those inflammatory cells infiltrating below the prosthetic surface.^[59] In bovine pericardial heart valve prostheses, for instance, macrophages were regularly found



Fig. 4 Eroded surface of the cusp edge of a bioprosthetic heart valve, typically caused by chronic mononuclear inflammation.

invading and focally degrading the prosthetic collagen.^[32,60] Similarly, clinically implanted porcine valves also showed inflammatory cells inside the bioprosthetic tissue.^[47] Conspicuously, dense infiltrates of inflammatory cells were found in all valves that had failed due to tissue tearing.^[59,61] Given the particularly high levels of MMP-9 and other proteases secreted by macrophages,^[62] on top of insufficiently cross-linked degradation sites in the tissue, inflammatory degeneration emerges as a main player. Macrophage-rich areas are known to secrete MMPs constantly and produce lysis of the underlying substrate. The lytic activity also diffuses away into the surrounding tissue.^[63] The recently demonstrated synergism between proteolytic activity and mechanical strain in the degradation of pericardial tissue further highlights the detrimental role of inflammation in bioprosthetic tissue degeneration.^[64] Macrophage-mediated tissue degradation is further evidenced by the direct proof of collagen phagocytosis.^[65] In clinical series, as many as 82% of failed valves showed signs of collagen phagocytosis by macrophages by transmission electron microscopy.^[66] Others also found the presence of cross-banded collagen fibrils in cytoplasmic vacuoles of macrophages.^[67] Since some of these fragments had a periodicity of 100 nm, it appears as if not only collagen I but also collagen IV is degraded by macrophages. These fibrils were located in both electron-translucent and electron-dense vacuoles and colocalization with phosphatase activity identified some of the vacuoles as lysosomal. Although one could argue that macrophage infiltration is a consequence of collagen degeneration rather than its cause, numerous studies have demonstrated that phagocytosis of intact collagen by macrophages is a regular occurrence.^[68,69]

Proponents of the theory that tissue inflammation is unrelated to an immune response often relied on the claim that glutaraldehyde itself was proinflammatory. Grabenwöger has decisively refuted this hypothesis by demonstrating that macrophage infiltration as well as collagen phagocytosis and degradation also occur in tissue where extraction processes had thoroughly eliminated free as well as dissociable glutaraldehyde.^[68] At the same time, only those intact collagen fibers that were coated by a flocculent material of remarkable electron density were phagocytosed. This observation may well relate to an opsonization process. Since bioprosthetic tissue remains largely nonendothelialized and thus deprived of the VECAM/MCP-1 mechanism of leukocyte recruitment, opsonization is one key to local inflammation. Opsonization may occur either through preformed antibodies or as a result of more unspecific plasma proteins binding to the surfaces. Tang^[70] has shown that the strongest adhesion effect by plasma proteins is exerted by surface-bound fibrinogen, followed by IgG and albumin. As much as

unspecific opsonization with plasma proteins is likely, opsonization through preformed antibodies can be assumed to also play a role. A humoral immune response against tissue heart valves and the presence of circulating specific antibodies against glycoproteins of the bioprosthetic material^[71] has also been described, together with histological evidence for a certain degree of a low-level xenograft response. Although milder than in uncross-linked xenograft tissue,^[15] signs of xenograft rejection in the form of polymorphnuclear infiltrates were found both experimentally^[65] and in clinical explants.^[61] In 1984, Levy's and Schoen's groups had already provided a link between insufficient cross-linking and inflammation. By increasing the glutaraldehyde concentration from 0.1% to 2.0% they were able to mitigate an otherwise intense mononuclear and giant cell reaction.^[72] In analogy to Nimni's work of the 1980s in which enhanced cross-linking through long-range diamine extensions of glutaraldehyde cross-links achieved a mitigation of remnant immunogenicity, diamine extension also led to a distinct suppression of inflammation in our own experiments.^[73] In conventionally fixed, low-glutaraldehyde-treated porcine valves, in which we had demonstrated a low cross-link density, the inflammatory reaction was distinct, particularly in the fibrosa of leaflets. Infiltrating cells were identified as macrophages in their majority, but in the primate model the percentage of polymorphnuclear granulocytes was significant^[65] (Fig. 5). The preferential infiltration of the fibrosa of leaflets was in accordance with other studies that also found collagen-destructive macrophages primarily in the fibrosa.^[74] In contrast, inflammation could practically be eradicated when GA concentrations were increased and additional long-range bonds introduced. While spots of foreign body giant cells were occasionally still found in the adventitial area of this highly cross-linked tissue, leaflets were completely free of inflammation.^[75]

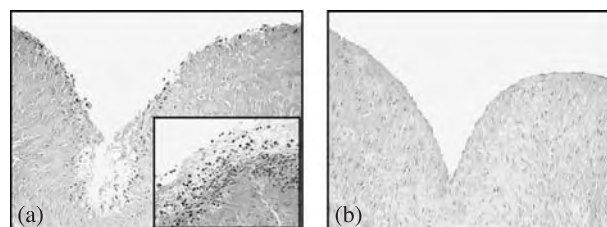


Fig. 5 Histological sections of glutaraldehyde-fixed porcine aortic root tissue after 6 weeks of experimental implantation. (a) Leaflet, 0.2% glutaraldehyde; (b) leaflet, 3% glutaraldehyde + L-Lysine. Note the distinct inflammatory infiltration and erosion of the fibrosa in the 0.2% group as opposed to a regularly unaffected, noneroded leaflet surface in the 3% glutaraldehyde-plus-diamine group.

One of the strongest alternative mechanisms for macrophage adhesion apart from opsonization is low shear stress. While the high shear stress in main arteries ($15\text{--}20\text{ dynes/cm}^2$)^[76] down-regulates adhesion molecules on the endothelial surface, the low shear stress in capillaries ($0.5\text{--}2.0\text{ dynes/cm}^2$) rather serves to up-regulate them, thereby facilitating diapedesis. This pathway is the likely force behind the strong inflammatory cell recruitment seen on the adventitial side of root prostheses.^[65] In the absence of an endothelium, as is the case on the blood surface of bioprosthetic heart valves, shear stress itself provides a similar, albeit mitigated, effect on its own. One would therefore expect direct macrophage adherence to flow-protected areas of the blood surface of bioprosthetic heart valves as a consequence of low shear stress. Flow-protected, low shear-stress areas on the blood surface include the surface of the fibrosa of the leaflets. Against the widespread belief that turbulences cause particularly high shear forces on this inflow surface, the opposite seems to be true: Bellhouse and Talbot used a quasisteady laminar model to describe the pulsatile vortex motion within the sinus. From these calculations, the shear stresses within the sinus are lower than those experienced on the ventricularis side of the cusp. Even if one needs to caution that these results are for laminar steady flow, and the Hill Vortex Model is a simple in-viscid flow model, which describes the complex nature of the pulsatile flow as quasisteady, it represents a valuable tool for such a flow situation. Calculations for arterial wall shear stresses are known to be extremely difficult and according to Ku,^[77] tend to represent estimates that may have errors as high as 20–50%. In the case of bioprosthetic heart valves, these estimates at least replace perceptions to the contrary by the strong evidence that the shear forces within the vortex-flow area of the sinus—including the fibrosa of leaflets—are lower than those on the ventricularis side. These theoretical considerations correspond with histological observations: Infiltrating inflammatory cells are found on the fibrosa surface rather than the ventricularis surface of leaflets^[65,74] as well as the inside of the lower sinus portion and the adventitial surface of root prostheses.^[65,73] In analogy, it is not surprising that the highest macrophage density is regularly found in completely flow-protected areas such as the interface between the prosthesis and surface thrombi^[37,78] or the prosthesis and pannus tissue.^[39] In contrast, inflammatory cells are practically always absent on the blood surface of the supra-commissural aortic wall of root prostheses,^[78] where high shear forces are further augmented by the systolic jet through the valve orifice.^[79]

Concluding these observations, it is unlikely that future valve designs will address these low shear-stress regions of bioprosthetic heart valves. Therefore, the

key to the suppression of inflammation seems to lie in mitigated opsonization. The successful reduction^[72] or elimination^[73] of inflammatory infiltrates through higher cross-link density demonstrated that cross-linking chemistry again holds a key to a pivotal problem of tissue valve failure.

Mechanical Injury

Native aortic valves function on the basis of a delicate combination of macro- and micromechanics. One of the basic macromechanical principles is the avoidance of cusp-bending at acute angles. A distinct dilatation of the annulus during systole results in the triangular stretching of cusp tissue, thereby avoiding bending at the commissures of more than 60° ^[80] (Fig. 6). Micro-mechanically, a sophisticated trilayered ultrastructural architecture ascertains optimal stress reduction. In principle, this trilayered architecture is based on a relatively smooth lamina ventricularis on the outflow side and a folded lamina fibrosa on the inflow side, separated by the sliding gap of the lamina spongiosa. When leaflets bend during valve opening, compressive stresses are generated on the inner surface of the bend and tensile stresses on the outside surface.^[81] When native leaflets flex radially, very little tension is generated in the outer ventricularis, since the already deeply folded fibrosa can be readily accommodated by the loose connection between the two that passes through the spongiosa.^[81] Moreover, even collagen fibrils—the steel cables of cusp support—have a series of complex built-in structural aspects that act as shock absorbers. Although collagen fibrils themselves are stiff and have a limited extensibility of 1–2% before breaking,^[55]

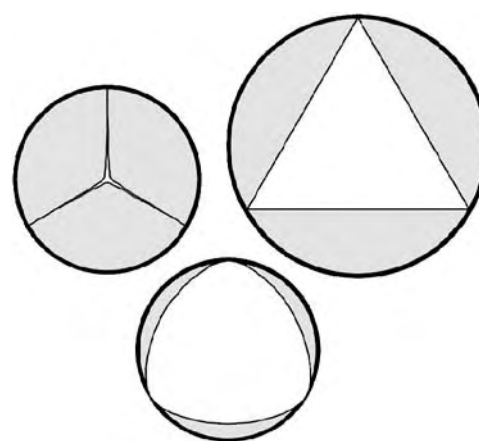


Fig. 6 Root mechanics of the aortic valve. While a native root shows a marked annular distension during systole and thereby a more triangular opening of the leaflets (right, top), the fixed annular diameter of prosthetic valves results in a much more acute bending angle of the leaflets during the outflow phase (bottom).

waviness allows aortic cusp collagen to strain up to 17%. Macrocrimping adds an additional 23%, pushing the overall strain acceptance to 40%.^[82] Together with the recoil effect of the elastic elements of the cusp tissue, this micromechanical shock-absorber principle leads to a 50% increase in cross-sectional cusp thickness in the relaxed state.^[83]

Bioprosthetic heart valves—stented or stentless—defy both main principles of valve mechanics: By either stent-mounting, or through glutaraldehyde fixation, annular dilatation during systole is either eliminated or dramatically diminished. As a consequence, the bending angle becomes acute and reverse curvature and buckling ensues. Vesely has hypothesized that tissue buckling may be an important failure mechanism of bioprosthetic heart valves.^[84] By eliminating the sliding gap between the ventricularis and the fibrosa as a result of cross-linking, a major means of stress distribution is lost. Furthermore, the loss of macrocrimping of the lamina fibrosa critically reduces stress acceptance. While macrocrimping was already lost as a result of pressure fixation of cusp tissue during manufacture of earlier generations of tissue valves, low-pressure fixation (St. Jude Medical Toronto Stentless Porcine valve: 2–4 mmHg) or zero-pressure fixation (Medtronic Freestyle Valve: 0 mmHg) was thought to preserve this crucial element of valve architecture. Observations by Sacks^[85] and those of Christie,^[86] however, suggest that the collagen fiber architecture of 0 mmHg-fixed heart valves, although locked in place by chemical fixation, may not be maintained over a sufficient number of cycles to be clinically beneficial. After only 6 months of implantation, radial and circumferential extensibility decreased by 30% and 80%, respectively.^[86]

In view of the dramatic loss of physiological valve mechanics, it is not surprising that major ultrastructural injuries such as severe collagen breakdown are regularly observed in the major stress zones of bioprosthetic heart valves.^[47] Collagen fibrils also lose their compact arrangement as early as after two months and are distinctly separated after two years.^[48] Not only does injured and fragmented collagen lead to an augmentation of local macrophage activity, but also to facilitated protease digestion. Cyclic fatigue of glutaraldehyde-fixed porcine aortic valve bioprostheses was shown to lead to a progressive loss of helicity of cusp collagen as well as loss of glycosaminoglycans from the cuspal ECM.^[87] While the first exposes collagen to both tissue-bound but still active^[88] gelatinases and proteases released by host inflammatory cells, the latter further contributes to tissue stiffness.

In clinically implanted bioprosthetic heart valves, nothing highlights the contribution of mechanical damage to the overall failure of tissue valves more than the distinctly higher failure rate in mitral than in aortic

position where the loading pressure is diastolic rather than systolic.

Pannus Overgrowth

With the emergence of stentless valve prostheses, myointimal host-tissue overgrowth became a significant reason for concern due to the fact that the suture line is closer to the cusps than in stented valves. Yet, overgrowth of this so-called pannus tissue is also a feared complication of stented and even mechanical heart valve prostheses. In the mitral position, for instance, a significant proportion of failed tissue valves showed pannus overgrowth on the inflow aspect.^[59]

Although the outgrowth of anastomotic intimal tissue onto the prosthetic surface is a multi-factorial event, inflammation certainly contributes to the provision of a sustained growth signal for myointimal tissue. One of the elements of inflammation in bioprosthetic heart valves is the chronic foreign body reaction against the sewing ring material. The dominant cells in this reaction are macrophages that are known to secrete an extended range of powerful cytokines, including IL-1, PDGF, bFGF, etc.^[89] Furthermore, Greisler demonstrated that the interaction of Dacron material—the main component of most sewing rings—with macrophages leads to a particularly strong up-regulation of growth factors such as bFGF.^[90] However, although it was shown that a foreign body reaction against inert materials such as Dacron provides an ongoing stimulus for macrophages to secrete growth factors, an additional specific immune response directed against the bioprosthetic tissue may significantly augment this phenomenon. Apart from enabling the presentation of foreign peptides to the immune system of the host, the augmented digestion of collagen by giant cells—wedged between the pannus and the bioprosthetic tissue—releases even more high-molecular weight peptides. These peptides possess a great chemotactic attraction for fibroblasts.^[91] Since giant cells are also known to secrete IL-1 β and bFGF in abundance—sometimes even after years^[92,93]—fibroblasts and myofibroblasts of the pannus tissue receive a distinct local stimulus for proliferation. In addition to this growth factor stimulus, the interaction of macrophages with T-helper cells leads to further IL-1 secretion, which augments the production of PDGF by SMCs.^[94,95] On top of up-regulating PDGF, the strongest known mitogen for smooth muscle cells, macrophage-derived IL-1 itself also regulates the growth, proliferation, and protein synthesis of fibroblasts.^[96–103]

Immune-activated macrophages are known to induce the vascular myointimal proliferation in heart transplant recipients, which ultimately leads to

ischaemic graft failure.^[104] Similarly, macrophages have been implicated in a hyperplastic intimal response in homograft heart valves, a phenomenon that was distinctly less pronounced in isografts.^[105] If the homograft recipients were immunologically pre-sensitized, this myointimal thickening was seen to be particularly prominent,^[105] further demonstrating the pivotal role immune processes appear to play in myointimal proliferation.

In order to tie insufficient masking of bioprosthetic antigens to myointimal pannus out-growth, however, one needs to correlate cross-link efficiency with macrophage activation and pannus outgrowth. For the interaction of biomaterials with macrophages, the triple sequence of events, namely adherence, spreading, and secretion of inflammatory mediators such as TNF- α , IL-1 β ,^[106] and IL-6,^[107] has been well established. Since spreading is additionally linked to the macrophage's ability to up-regulate its hydrolytic degradation potential, spreading is an essential element of antigen presentation by macrophages. Simply by increasing the glutaraldehyde concentration from 0.2% to 3.0%, a significant reduction of the spreading area of macrophages was observed.^[108] Since both groups had undergone thorough detoxification,^[21] this phenomenon cannot be explained by the cytotoxic effect of glutaraldehyde. Similarly, diamine-enhanced cross-linking at 3% glutaraldehyde not only resulted in the near abolition of an inflammatory response, but also in a significant suppression of pannus outgrowth. After six weeks of sheep implantations of aortic root prostheses, the length of the proximal pannus overgrowing the leaflet base was five times shorter in the enhanced, cross-linked group than in the conventional 0.2% GA-fixed roots (Fig. 7). Tissue was

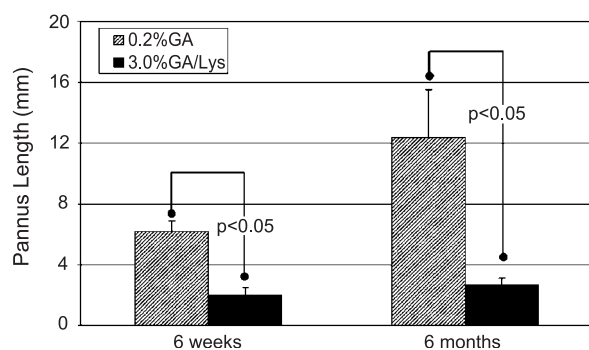


Fig. 7 The suppression of inflammation is associated with a significant mitigation of pannus outgrowth: aortic root prostheses fixed at either 0.2% glutaraldehyde or L-Lysine-enhanced 3% glutaraldehyde showed a 3–4.5-times shorter pannus outgrowth onto the prosthetic tissue after 6 weeks and 6 months of implantation, respectively. Preoperatively, all root-prostheses had been thoroughly detoxified through glutaraldehyde pacification.

again thoroughly detoxified before implantation. When giant cells at the interface of pannus and prosthetic tissue were correlated with cross-sectional pannus area or pannus length, a direct correlation emerged between cross-link and giant cell density.^[73]

Therefore, in view of the dramatic mitigation of pannus outgrowth through better cross-linking of tissue, it seems important to extend the alert-list of immune and inflammatory phenomena that may contribute to bioprosthetic heart valve failure to the potentially leaflet-immobilizing phenomenon of anastomotic pannus outgrowth.

THE CRUX OF THE MATTER

In the past 40 years, bioprosthetic heart valves meant a blessing for hundreds of thousands of patients. Many of the initial patients, however, may have been better off with a mechanical prosthesis because they were young and their first-generation tissue valves calcified rapidly. The disappointment of the surgeons of these early implants was huge and many opinion leaders of today are veterans of this pioneering era. The resulting paradigm reserving bioprosthetic heart valves for the geriatric patient group has led to a biased assessment of risks associated with anticoagulation in mechanical valves. A paradigm shift on the basis of a consensus regarding the acceptance of a reoperation as a trade off for a better quality of life seems overdue.

Similarly, decades of tissue valve research were obsessed with calcific degeneration because first-generation prostheses failed on the basis of dystrophic mineralization. An urgent refocusing on the real culprit of bioprosthetic degeneration, namely inflammatory degradation and inflammation-triggered pannus overgrowth, seems to hold a key to a quantum leap in the development of longer lasting tissue valves.

ARTICLES OF FURTHER INTEREST

Allografts; Collagen Fixation; Heart Valve Failure, Bioprosthetic; Heart Valves, Mechanical; Tissue Engineering of Heart Valves

REFERENCES

1. *The World Medical Market Fact File 1997*; MDIS Publications Ltd., 1997.
2. Jones, J.M.; O'Kane, H.; Gladstone, D.J.; Sarsam, M.A.; Campalani, G.; MacGowan, S.W.; Cleland, J.; Cran, G.W. Repeat heart valve surgery: Risk factors for operative mortality. *J. Thorac. Cardiovasc. Surg.* **2001**, *122*, 913–918.



3. Curcio, C.; Commerford, P.; Rose, A.; Stevens, J.; Barnard, M. Calcification of glutaraldehyde-preserved porcine xenografts in young patients. *J. Thorac. Cardiovasc. Surg.* **1981**, *81*, 621–625.
4. Rose, A.; Forman, R.; Bowen, R. Calcification of glutaraldehyde-fixed porcine xenograft. *Thorax* **1978**, *33*, 111–114.
5. Arom, K.V.; Emery, R.W.; Nicoloff, D.M.; Petersen, R.J. Anticoagulant related complications in elderly patients with St. Jude mechanical valve prostheses. *J. Heart Valve Dis.* **1996**, *5*, 505–510.
6. Harken, D.; Soroff, H.; Taylor, W.; Lefemine, A.; Gupta, S.; Lunzer, S. Partial and complete prostheses in aortic insufficiency. *J. Thorac. Cardiovasc. Surg.* **1960**, *40*, 744–762.
7. Starr, A.; Edwards, M. Mitral replacement: Clinical experiences with a ball-valve prosthesis. *Ann. Surg.* **1961**, *154*, 726.
8. Frater, R.; Berghuis, J.; Brown, A.; Ellis, F. Autogenous pericardium for posterior mitral leaflet replacement. *Surgery* **1963**, *84*, 260–268.
9. Frater, R.; Berghuis, J.; Brown, A.; Ellis, F. The experimental and clinical use of autogenous pericardium for the replacement and extension of mitral and tricuspid valve cusps and chordae. *J. Cardiovasc. Surg.* **1965**, *6*, 214–228.
10. Ross, D. Homograft replacement of the aortic valve. *Lancet* **1962**, *2*, 487.
11. Mohri, H.; Reichenbach, D.; Barnes, R.; Nelson, R.; Merendino, K. Studies of antigenicity of the homologous aortic valve. *J. Thorac. Cardiovasc. Surg.* **1967**, *54*, 564–572.
12. Davies, H.; Missen, G.; Blandford, G.; Roberts, C.; Lessof, M.; Ross, D. Homograft replacement of the aortic valve. A clinical and pathologic study. *Am. J. Cardiol.* **1968**, *22*, 195–217.
13. Davies, H.; Lessaf, M.; Roberts, C.; Ross, D. Homograft replacement of the aortic valve: Follow-up studies in twelve patients. *Lancet* **1965**, *1*, 926.
14. Ross, D. Replacement of aortic and mitral valves with a pulmonary autograft. *Lancet* **1967**, *2*, 956–958.
15. Duran, C.; Gunning, A.; Whitehead, R. Experimental aortic valve heterotransplantation. *Thorax* **1967**, *22*, 510–518.
16. Binet, J.; Carpentier, A.; Langlois, J.; Duran, C.; Colvez, P. Implantation of heterogenic valves in the treatment of aortic cardiopathies. *C. R. Hebd. Seances Acad. Sci., Ser. D, Sci. Nat.* **1965**, *261*, 5733–5734.
17. Carpentier, A.; Lemaigre, G.; Robert, L.; Carpentier, S.; Dubost, C. Biological factors affecting long-term results of valvular heterografts. *J. Thorac. Cardiovasc. Surg.* **1969**, *58*, 467–483.
18. Rosenberg, N.; Henderson, J.; Douglas, J.; Lord, G.; Gaughran, E. Use of enzyme-digested heterografts as segmental arterial substitutes. III. Follow-up observations on two-year-old aortic implants. *Arch. Surg.* **1958**, *76*, 275.
19. Lentz, D.; Pollock, E.; Olsen, D.; Andrews, E. Prevention of intrinsic calcification in porcine and bovine xenograft materials. *Trans. - Am. Soc. Artif. Intern. Organs* **1982**, *28*, 494–497.
20. Dewanjee, M.; Solis, E.; Lanker, J.; Mackey, S.; Lombardo, G.; Tidwell, C.; Ellefsen, R.; Kaye, M. Effect of diphosphonate binding to collagen upon inhibition of calcification and promotion of spontaneous endothelial cell coverage on tissue valve prostheses. *ASAIO Trans.* **1986**, *32*, 24–29.
21. Zilla, P.; Fullard, L.; Trescony, P.; Meinhart, J.; Bezuidenhout, D.; Gorlitzer, M.; Human, P.; von Oppell, U. Glutaraldehyde detoxification of aortic wall tissue: A promising perspective for emerging bioprosthetic valve concepts. *J. Heart Valve Dis.* **1997**, *6*, 510–520.
22. Girardot, M.; Girardot, J.; Schoen, F. Alpha amino oleic acid, a new compound, prevents calcification of bioprosthetic heart valves. *Trans. Soc. Biomater.* **1991**, *14*, 114(abstract)
23. Carpentier, A. Bovine Pericardium or Porcine Valves? Pericardial Bioprostheses Mini-Symposium, Luzern, Switzerland, 1989.
24. Perier, P. Seven Years Clinical Follow-Up of Aortic Valve Replacement with the Carpentier-Edwards Pericardial Valve, Pericardial Bioprostheses Mini-Symposium, Luzern, Switzerland, 1989.
25. Gabbay, S.; Kadam, P.; Factor, S.; Cheung, T. Do heart valve bioprostheses degenerate for metabolic or mechanical reasons? *J. Thorac. Cardiovasc. Surg.* **1988**, *95*, 208–215.
26. Grunkemeier, G.; Bodnar, E. Comparative assessment of bioprosthesis durability in the aortic position. *J. Heart Valve Dis.* **1995**, *4*, 49–55.
27. Love, J.; Schoen, F.; Breznock, E.; Shermer, S.; CS, L. Experimental evaluation of an autologous tissue heart valve. *J. Heart Valve Dis.* **1992**, *1*, 232–241.
28. Goetz, W.; Lim, H.; Lansac, E.; Weber, P.; Duran, C. A temporarily stented, autologous pericardial aortic valve prosthesis. *J. Heart Valve Dis.* **2002**, *11*, 696–702.
29. Butany, J.; deSa, M.; Feindel, C.; David, T. The Toronto SPV bioprosthesis: Review of morphologic findings in eight valves. *Semin. Thorac. Cardiovasc. Surg.* **1999**, *11* (4 Suppl.), 157–162.
30. Goldman, B.; Christakis, G.; David, T.; Rakowski, H.; Bach, D.; Wood, J.; Goldman, S.; Pepper, J.; Yacoub, M.; Verrier, E.; Petracek, M. Will stentless valves be durable? The Toronto valve (TSPV) at 5 to 6 years. *Semin. Thorac. Cardiovasc. Surg.* **1999**, *11*, 42–49.
31. Westaby, S.; Jin, X.; Katsumata, T.; Arifi, A.; Braidley, P. Valve replacement with a stentless bioprosthesis: Versatility of the porcine aortic root. *J. Thorac. Cardiovasc. Surg.* **1998**, *116*, 477–484.
32. Dahm, M.; Lyman, W.; Schwell, A.; Factor, S.; Frater, R. Immunogenicity of glutaraldehyde-tanned bovine pericardium. *J. Thorac. Cardiovasc. Surg.* **1990**, *99*, 1082–1090.
33. Cobanoglu, A.; Jamieson, W.; Miller, D.; McKinley, C.; Grunkemeier, G.; Floten, H.; Miyagishima, R.; Tyers, G.; Shumway, N.; Starr, A. A tri-institutional comparison of tissue and mechanical valves using a patient-oriented definition of “treatment failure”. *Ann. Thorac. Surg.* **1987**, *43*, 245–253.

34. David, T.; Armstrong, S.; Sun, Z. The Hancock II bioprosthesis at ten years. *Ann. Thorac. Surg.* **1995**, *60*, S229–S234.
35. Gott, J.; Pan, C.; Dorsey, L.; Jay, J.; Jett, G.; Schoen, F.; Girardot, J.; Guyton, R. Calcification of porcine valves: A successful new method of antimineralization. *Ann. Thorac. Surg.* **1992**, *53*, 207–215discussion 216
36. Cosgrove, D.; Frater, R.; Scott-Page, U. Clinical communique Edwards Lifesci. LLC.
37. Maxwell, L.; Gavin, J.; Barratt-Boyes, B. Uneven host tissue ongrowth and tissue detachment in stent mounted heart valve allografts and xenografts. *Cardiovasc. Res.* **1989**, *23*, 709–714.
38. Schoen, F.; Hobson, C. Anatomic analysis of removed prosthetic heart valves: Causes of failure of 33 mechanical valves and 58 bioprostheses, 1980 to 1983 *Human Pathol.* **1985**, *16*, 549–559.
39. Camilleri, J.; Pornin, B.; Carpentier, A. Structural changes of glutaraldehyde-treated porcine bioprosthetic valves. *Arch. Pathol. Lab. Med.* **1982**, *106*, 490–496.
40. Zilla, P.; Zhang, Y.; Human, P.; Koen, W.; von Oppell, U. Improved ultrastructural preservation of bioprosthetic tissue. *J. Heart Valve Dis.* **1997**, *6*, 492–501.
41. Human, P.; Weissenstein, C.; Trantina, A.; Zilla, P. Fixation-related autolysis and bioprosthetic aortic wall calcification. *J. Heart Valve Dis.* **2001**, *10*, 656–665.
42. Simionescu, D.; Lovekamp, J.; Vyavahare, N. Degeneration of bioprosthetic heart valve cusp and wall tissues is initiated during tissue preparation: An ultrastructural study. *J. Heart Valve Dis.* **2003**, *12*, 226–234.
43. Lemasters, J.; DiGiuseppi, J.; Nieminen, A.; Herman, B. Blebbing, free Ca²⁺ and mitochondrial membrane potential preceding cell death in hepatocytes. *Nature* **1987**, *325*, 78–81.
44. Kim, K. Cell death and calcification of canine fibroblasts in vitro. *Cells Mater.* **1994**, *4*, 247–261.
45. Trump, B.; Valigorsky, J.; Dees, J.; Mergner, W.; Kim, K.; Jones, R.; Pendergrass, R.; Garbus, J.; Cowley, R. Cellular changes in human disease. A new method of pathological analysis. *Human Pathol.* **1973**, *4*, 89–109.
46. Trump, B.; Berezsky, I.; Laiho, K.; Osorino, A.; Mergner, W.; Smith, M. The role of calcium in cell injury. A review. *Scanning Electron Microsc.* **1980**, *II*, 437–462.
47. Ferrans, V.; Spray, L.; Billingham, M.; Roberts, W. Structural changes in glutaraldehyde-treated porcine heterografts used as substitute cardiac valves. *Am. J. Cardiol.* **1978**, *41*, 1159–1184.
48. Ashraf, M.; Bloor, C. Structural alterations of the porcine heterograft after various durations of implantation. *Am. J. Cardiol.* **1978**, *41*, 1185–1190.
49. Anderson, H. Calcification processes. *Pathol. Annu.* **1980**, *15*, 45–75.
50. Morea, M.; De, P.R.; Galloni, M.; Gastaldi, L.; di, S.M. Mitral valve replacement with the Biocor stentless mitral valve: Early results. *J. Heart Valve Dis.* **1994**, *3*, 476–482.
51. Ishihara, T.; Ferrans, V.; Jones, M.; Boyce, S.; Roberts, W. Structure of bovine parietal pericardium and of unimplanted Ionescu-Shiley pericardial valvular bioprostheses. *J. Thorac. Cardiovasc. Surg.* **1981**, *81*, 747–757.
52. Schoen, F.; Levy, R.; Hilbert, S.; Bianco, R. Antimineralization treatments for bioprosthetic heart valves. Assessment of efficacy and safety. *J. Thorac. Cardiovasc. Surg.* **1992**, *104*, 1285–1288.
53. Chen, W.; Kim, J.; Schoen, F.; Levy, R. Effect of 2-amino oleic acid exposure conditions on the inhibition of calcification of glutaraldehyde cross-linked porcine aortic valves. *J. Biomed. Mater. Res.* **1994**, *28*, 1485–1495.
54. Zilla, P.; Weissenstein, C.; Bracher, M.; Zhang, Y.; Koen, W.; Human, P.; von Oppell, U. High glutaraldehyde concentrations reduce rather than increase the calcification of aortic wall tissue. *J. Heart Valve Dis.* **1997**, *6*, 502–509.
55. Scott, M.; Vesely, I. Aortic valve cusp microstructure: The role of elastin. *Ann. Thorac. Surg.* **1995**, *60*, S391–S394.
56. Schoen, J.; Shemin, R.; Cohn, L. Pathologic evaluation of a unileaflet pericardial bioprosthesis after implantation as tricuspid and mitral valve replacements in sheep. *Trans. Soc. Biomater.* **1986**, *9*, 76.
57. Sterling, H.; Saginario, C.; Vignery, A. CD44 occupancy prevents macrophage multinucleation. *J. Cell Biol.* **1998**, *143*, 837–847.
58. Vignery, A.; Niven-Fairchild, T.; Ingbar, D.; Caplan, M. Polarized distribution of Na⁺K⁺-ATPase in giant cells elicited in vivo and in vitro. *J. Histochem. Cytochem.* **1989**, *35*, 1265–1271.
59. Valente, M.; Bortolotti, U.; Thiene, G. Ultrastructural substrates of dystrophic calcification in porcine bioprosthetic valve failure. *Am. J. Pathol.* **1985**, *119*, 12–21.
60. Dahm, M.; Husmann, M.; Eckhard, M.; Prufer, D.; Groh, E.; Oelert, H. Relevance of immunologic reactions for tissue failure of bioprosthetic heart valves. *Ann. Thorac. Surg.* **1995**, *60*, S348–S352.
61. Nistal, F.; Garcia, S.E.; Artinano, E.; Duran, C.; Gallo, I. Comparative study of primary tissue valve failure between Ionescu-Shiley pericardial and Hancock porcine valves in the aortic position. *Am. J. Cardiol.* **1986**, *57*, 161–164.
62. Fini, M.; Girard, M.; Matsubara, M.; Bertlett, J. Unique regulation of the matrix metalloproteinase, gelatinase B. *Investig. Ophthalmol. Vis. Sci.* **1995**, *36*, 622–632.
63. Galis, Z.; Sukhova, G.; Libby, P. Microscopic localization of active proteases by in situ zymography: Detection of matrix metalloproteinase activity in vascular tissue. *FASEB J.* **1995**, *9*, 974–980.
64. Ellsmere, J.; Khanna, R.; Lee, J. Mechanical loading of bovine pericardium accelerates enzymatic degradation. *Biomaterials* **1999**, *20*, 1143–1150.
65. Trantina, Y.A.; Weissenstein, C.; Human, P.; Zilla, P. Stentless bioprosthetic heart valve research: Sheep versus primate model. *Ann. Thorac. Surg.* **2001**, *71*, S422–S427.
66. Stein, P.; Wang, C.; Riddle, J.; Magilligan, D.J. Leukocytes, platelets, and surface microstructure of spontaneously degenerated porcine bioprosthetic valves. *J. Card. Surg.* **1988**, *3*, 253–261.



67. Everts, V.; van, D.Z.E.; Creemers, L.; Beertsen, W. Phagocytosis and intracellular digestion of collagen, its role in turnover and remodelling. *Histochem. J.* **1996**, *28*, 229–245.
68. Grabenwoger, M.; Grimm, M.; Eybl, E.; Leukauf, C.; Muller, M.; Plenck, H.J.; Bock, P. Decreased tissue reaction to bioprosthetic heart valve material after L-glutamic acid treatment. A morphological study. *J. Biomed. Mater. Res.* **1992**, *26*, 1231–1240.
69. Grabenwoger, M.; Grimm, M.; Eybl, E.; Kadletz, M.; Havel, M.; Kostler, P.; Plenck, H.; Bock, P.; Wolner, E. New aspects of the degeneration of bioprosthetic heart valves after long-term implantation. *J. Thorac. Cardiovasc. Surg.* **1992**, *104*, 14–21.
70. Tang, L.; Ugarova, T.; Plow, E.; Eaton, J. Molecular determinants of acute inflammatory responses to biomaterials. *J. Clin. Invest.* **1996**, *97*, 1329–1334.
71. Human, P.; Zilla, P. Characterization of the immune response to valve bioprostheses and its role in primary tissue failure. *Ann. Thorac. Surg.* **2001**, *71*, S385–S388.
72. Sherman, F.; Schoen, F.; Hawley, M.; Nichols, J.; Levy, R. Collagen crosslinks: A critical determinant of bioprosthetic heart valve calcification. *Trans. - Am. Soc. Artif. Intern. Organs* **1984**, *30*, 577–581.
73. Human, P.; Zilla, P. Inflammatory and immune processes: The neglected villain of bioprosthetic degeneration? *J. Long-Term Eff. Med. Implants* **2001**, *11*.
74. Ferrans, V.; Boyce, S.; Billingham, M.; Jones, M.; Ishihara, T.; Roberts, W. Calcific deposits in porcine bioprostheses: Structure and pathogenesis. *Am. J. Cardiol.* **1980**, *46*, 721–734.
75. Zilla, P.; Bezuidenhout, D.; Weissenstein, C.; Van der Walt, A.; Human, P. Diamine extension of glutaraldehyde crosslinks mitigates bioprosthetic aortic wall calcification in the sheep model. *J. Biomed. Mater. Res.* **2001**, *56*, 56–64.
76. Sprague, E.; Mohan, S.; Nerem, R. Shear stress regulation of monocyte/endothelial interactions. *J. Vasc. Surg.* **1999**, *29*, 1138–1140.
77. Ku, D. Blood flow in arteries. *Annu. Rev. Fluid Mech.* **1997**, *29*, 399–434.
78. Zilla, P.; Weissenstein, C.; Human, P.; Dower, T.; von Oppell, U. High glutaraldehyde concentrations mitigate bioprosthetic root calcification in the sheep model. *Ann. Thorac. Surg.* **2000**, *70*, 2091–2095.
79. Yoganathan, A.; Ellis, J.; Healy, T.; Chatzimavroudis, G. Fluid dynamic studies for the year 2000. *J. Heart Valve Dis.* **1998**, *7*, 130–139.
80. Thubrikar, M.; Boshier, L.; Nolan, S. The mechanism of opening of the aortic valve. *J. Thorac. Cardiovasc. Surg.* **1979**, *77*, 863–870.
81. Vesely, I.; Noseworthy, R. Micromechanics of the fibrosa and the ventricularis in aortic valve leaflets. *J. Biomech.* **1992**, *25*, 101–113.
82. Kastelic, J.; Baer, E. Deformation in tendon collagen. *Symp. Soc. Exp. Biol.* **1980**, *34*, 397–420.
83. Allen, D.; DiDio, L.; Zacharias, A.; Fentie, I.; McGrath, A.; Puig, L.; Pomerantzeff, P.; Zerbini, E. Microscopic study of normal parietal pericardium and unimplanted Puig-Zerbini pericardial valvular heterografts. *J. Thorac. Cardiovasc. Surg.* **1984**, *87*, 845–855.
84. Vesely, I.; Mako, W. Comparison of the compressive buckling of porcine aortic valve cusps and bovine pericardium. *J. Heart Valve Dis.* **1998**, *7*, 34–39.
85. Sacks, M. The biomechanical effects of fatigue on the porcine bioprosthetic heart valve. *J. Long-Term Eff. Med. Implants* **2001**, *11*, 231–247.
86. Christie, G.; Barratt-Boyes, B. Time-Dependent Changes to the Leaflet Elasticity of the Medtronic Intact Valve in-vivo, World Symposium on Heart Valve Disease, London, UK, 1999.
87. Vyavahare, N.; Ogle, M.; Schoen, F.; Zand, R.; Gloeckner, D.; Sacks, M.; Levy, R. Mechanisms of bioprosthetic heart valve failure: Fatigue causes collagen denaturation and glycosaminoglycan loss. *J. Biomed. Mater. Res.* **1999**, *46*, 44–50.
88. Simionescu, D.; Simionescu, A.; Deac, R. Detection of remnant proteolytic activities in unimplanted glutaraldehyde-treated bovine pericardium and explanted cardiac bioprostheses. *J. Biomed. Mater. Res.* **1993**, *27*, 821–829.
89. Bonfield, T.; Anderson, J. Functional versus quantitative comparison of IL-1 beta from monocytes/macrophages on biomedical polymers. *J. Biomed. Mater. Res.* **1993**, *27*, 1195–1199. [published erratum appears in *J. Biomed. Mater. Res.* 1994 Jan;28(1);137].
90. Greisler, H.; Dennis, J.; Endean, E.; Ellinger, J.; Friesel, R.; Burgess, W. Macrophage/biomaterial interactions: The stimulation of endothelialization. *J. Vasc. Surg.* **1989**, *9*, 588–593.
91. Ciapetti, G.; Verri, E.; Granchi, D.; Cenni, E.; Gamberini, S.; Benetti, D.; Mian, M.; Pizzoferrato, A. In vitro assessment of phagocytosis of bovine collagen by human monocytes/macrophages using a spectrophotometric method. *Biomaterials* **1996**, *17*, 1703–1707.
92. Anderson, J.; Miller, K. Biomaterial biocompatibility and the macrophage. *Biomaterials* **1984**, *5*, 5–10.
93. Forsyth, E.; Aly, H.; Neville, R.; Sidawy, A. Proliferation and extracellular matrix production by human infragenicular smooth muscle cells in response to interleukin-1 beta. *J. Vasc. Surg.* **1997**, *26*, 1002–1007. discussion 1007–1008.
94. Clinton, S.; Libby, P. Cytokines and growth factors in atherogenesis. *Arch. Pathol. Lab. Med.* **1992**, *116*, 1292–1300.
95. Raines, E.; Dower, S.; Ross, R. Interleukin-1 mitogenic activity for fibroblasts and smooth muscle cells is due to PDGF-AA. *Science* **1989**, *243*, 393–396.
96. Nathan, C. Secretory products of macrophages. *J. Clin. Invest.* **1987**, *79*, 319–326.
97. Oppenheim, J.; Kovacs, E.; Matasushima, K.; Durum, S. There is more than one interleukin 1. *Immunol. Today* **1986**, *7*, 45.
98. Dinarello, C. Interleukin 1. *Rev. Infect. Dis.* **1984**, *6*, 51.
99. Freundlich, B.; Bomalski, J.; Neilson, E.; Jimenez, S. Regulation of fibroblast proliferation and collagen synthesis by cytokines. *Immunol. Today* **1986**, *7*, 303.
100. Grinnell, F. Fibronectin and wound healing. *Am. J. Dermatopathol.* **1982**, *4*, 185–188.
101. Oppenheim, J.; Gery, I. Interleukin 1 is more than an interleukin. *Immunol. Today* **1982**, *3*, 113.

102. Schmidt, J.; Mizel, S.; Cohen, D.; Green, I. Interleukin 1, a potential regulator of fibroblast proliferation. *J. Immunol.* **1982**, *128*, 2177–2182.
103. Tsukamoto, Y.; Hesel, W.; Wahl, S. Macrophage production of fibronectin, a chemoattractant for fibroblasts. *J. Immunol.* **1981**, *127*, 673–678.
104. Wewers, M.; Marsh, C. Role of the antibody in the pathogenesis of transplant vascular sclerosis: A hypothesis. *Transpl. Immunol.* **1997**, *5*, 283–288.
105. Green, M.; Walsh, M.; Dare, A.; Hogan, P.; Zhao, X.; Frazer, I.; Bansal, A.; O'Brien, M. Histologic and immunohistochemical responses after aortic valve allografts in the rat. *Ann. Thorac. Surg.* **1998**, *66*, S216–S220.
106. Miller, K.; Huskey, R.; Bigby, L.; Anderson, J. Characterization of biomedical polymer-adherent macrophages: Interleukin 1 generation and scanning electron microscopy studies. *Biomaterials* **1989**, *10*, 187–196.
107. Bonfield, T.; Colton, E.; Marchant, R.; Anderson, J. Cytokine and growth factor production by monocytes/macrophages on protein preadsorbed polymers. *J. Biomed. Mater. Res.* **1992**, *26*, 837–850.
108. Dower, T.; Adler, U.; Davids, L.; Zilla, P. Increasing cross-linking efficiency mitigates macrophage activation on bioprosthetic tissue. *Cardiovasc. Pathol.* **1998**, *7*, 295.

Heart Valve Failure, Bioprosthetic

Peter Zilla

Paul Human

Deon Bezuidenhout

Christian Barnard Department of Cardiothoracic Surgery, University of Cape Town Medical School, and Grootte Schuur Hospital, Cape Town, South Africa

INTRODUCTION

The most sobering aspect in the search for causative factors in the degeneration of bioprosthetic heart valves is the necessity to capitulate before the single most important etiological factor: the absence of living tissue in all contemporary bioprosthetic heart valves. However, as long as tissue engineering concepts remain a future promise, this lack of vitality will remain an integral part of bioprostheses.

The second most powerful determinant affecting the fate of tissue heart valves is not even related to the bioprosthesis itself: It is the environment provided by the host. While only half of first generation tissue valves have failed after 15 years of implantation into recipients of more than 60 years of age, practically all of them have failed after 5 years in children.^[1] This dramatic early degeneration in children obliterates all differences one encounters in the various valve types in adults: Regardless of whether the prosthesis was a homograft, a pericardial xenograft, or a porcine xenograft, calcific degeneration became significant before five years of implantation.^[1] In a less dramatic but nevertheless still very distinct fashion bioprostheses also degenerate in an accelerated way in patients on dialysis.

Against this framework, the quest for understanding the basic triggers of prosthesis-related tissue valve failure has been largely unsuccessful. Within the chain of events leading to severe tissue degeneration the focus was often on isolated downstream aspects rather than the bigger picture. Although each of these aspects undisputedly plays a role in the gradual loss of the function of the bioprosthetic tissue, a more comprehensive mechanism behind this development has yet to emerge. Singling out phospholipids in the cell membranes or poorly suppressed membrane-associated alkaline phosphatase molecules reveals mosaic stones but keeps the mural hidden. To get to the bottom of etiological factors behind bioprosthetic tissue failure, one may need to revisit some fundamental principles behind the concept.

ADVERSE EFFECTS OF GLUTARALDEHYDE CROSS-LINK CHEMISTRY

The core principle behind xenograft heart valves is the necessity to mask antigenicity through cross-linking. Complacency regarding the relatively effective cross-linking agent glutaraldehyde (GA) represents one of the major areas of neglect of the past three decades. Not only did this conservative approach to tissue fixation lead to a lack of research into understanding and engineering of optimized cross-links, but it also led to an almost fatalistic acceptance of the adverse effects of commercial glutaraldehyde cross-linking. Whether these adverse aspects were the alleged procalcific effects of glutaraldehyde reaction products, the inability of cross-linking procedures to mask antigenicity and calcium-nucleation sites sufficiently, or the mechanical damage resulting from tissue stiffness; cross-link chemistry had been sacrificed on the altar of anti-calcification treatments. This complacent continuation with one and the same cross-linking process for 35 years was partly based on an illusion, namely the perception that glutaraldehyde ideally represents the keep-it-simple principle. Being a short di-aldehyde, even surgeons could comprehend the perceived straight-forward bridge-formation of this molecule between two amine groups in the tissue. When it slowly became clear that we only saw the tip of the iceberg, too much already hinged on this central element of bioprosthetic heart valve production to change the course easily. Today, we know that the GA-fixation process is much more complex than initially thought. By now it is generally accepted that the Schiff's base intermediates that are initially formed by the reaction of the aldehyde groups with the ϵ -amino groups of L-Lysine or hydroxylysine residues are only the starting point for a variety of subsequent reactions.^[2] However, in spite of decades of use of GA for the production of bioprosthetic heart valves, many of these reactions remain in the realm of speculation or hypothetical interpretation. Cheung and Nimni have suggested that the formed Schiff bases are stable and that cross-links

might be formed after aldol condensation followed by the formation of another Schiff base.^[2] These pioneers of glutaraldehyde cross-link chemistry also suggested that α - β -unsaturated Schiff's bases form, after which cross-linking can be achieved by reaction of L-Lysine ϵ -amine groups with the α - β -unsaturated Schiff's base (Michael addition), or by reaction of the aldehyde endgroup of the unsaturated Schiff's base with amine groups. A Mannich-type reaction between the protonated Schiff's base and the GA-related enol resulting in the formation of a secondary amine was also suggested.^[3,4] Aldol condensation and subsequent reaction with the ϵ -amino groups of proteins can result in the formation of aliphatic cross-links. The reaction of the secondary amines with other GA molecules followed by ring closure would lead to a six-membered dihydropyridine.^[3] Oxygen present in the cross-linking solution may oxidize dihydropyridinium-type cross-links resulting in the formation of substituted quaternary pyridinium-type cross-links.^[5] Last but not least, glutaraldehyde's propensity to autopolymerize after ring-formation to hemiacetals adds a further complicating aspect, namely the ability of tissue to store large amounts of the dialdehyde in the form of poly-hemiacetals and to release free glutaraldehyde after depolymerization later on. This vaguely understood and highly complex chemistry of glutaraldehyde makes it very difficult to establish any correlations between the chemical modifications that have taken place and the resulting biological behavior, particularly glutaraldehyde's suspected role in bioprosthetic tissue calcification. One directly assessable effect of glutaraldehyde, however, is tissue stiffness, which can increase up to twentyfold compared with native tissue.^[6] Considering the fact that commercially used concentrations of GA are still incapable of fully masking antigenicity, this is a high price to pay. Another easily assessable disadvantage of glutaraldehyde

is its toxicity—even at low concentrations—which is potentiated by the reservoir effect of polyhemiacetals in the tissue.

In the past various attempts have been made to achieve optimization of cross-linking chemistry *without* abandoning glutaraldehyde fixation altogether. In order to mask immune and degradation sites better, enhanced fixation was successfully attempted. By either increasing concentrations of GA alone or in combination with a bridging diamine such as L-Lysine in order to create additional longer-range cross-links between terminal aldehyde groups, calcification could be distinctly suppressed and degradability mitigated, but tissue stiffness further increased^[7] (Fig. 1).

Some clarity with regard to the alleged procalcific effect of GA could be reached through the blocking of residual aldehydes under low pH conditions. Grimm successfully used L-glutamic acid for leaflet detoxification, whereas our group tried to identify the most optimal detoxification agent and the most optimal reaction conditions in the more challenging aortic wall tissue. When we compared 12 different amino reagents from four chemical groups (low pKa aromatic amines, amino acids, low pKa N-heterocyclic compounds, and amino sugars), all 12 amino reagents were able to detoxify aortic wall tissue satisfactorily, with Urazole satisfying most of the demands on the process.^[8] Alternatively, diphosphonates, which are believed to act twofold on calcification, were used: Aldehyde stabilization occurs through binding via Schiff's bases and subsequent stabilization with NaBH₄, while restriction of crystal growth is thought to occur through direct diphosphonate binding to developing hydroxyapatite nucleation sites.^[9] The precise mechanism of the anticalcification effect of α -amino-oleic-acid (AOA[®])—the anticalcification substance incorporated into Medtronic's Mosaic and Freestyle valves—is not fully understood either, but it is thought that it also

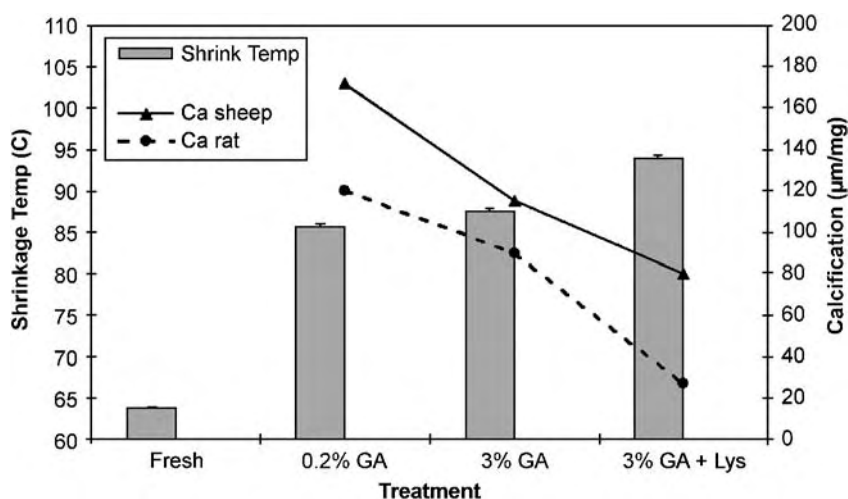


Fig. 1 Correlation between pronase resistance of aortic wall tissue and calcification in the sheep and the rat model. By increasing the cross-link density through increasing glutaraldehyde concentrations and eventually through the additional introduction of diamine bridges (L-lysine), the degradation resistance goes up while tissue calcification decreases dramatically.



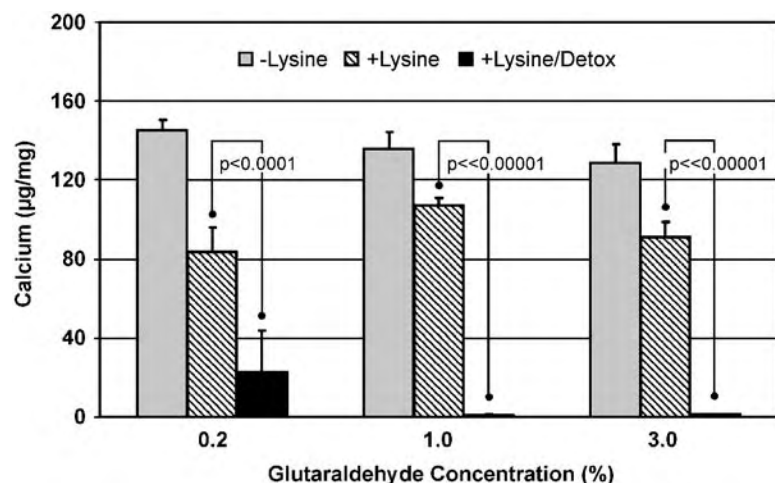


Fig. 2 Dramatic effect of glutaraldehyde detoxification on leaflet calcification. In contrast to aortic wall tissue where higher cross-link density is crucial, detoxification even at low glutaraldehyde concentrations practically eliminates calcification in leaflets.

binds with its 2-amino groups to free aldehyde groups. Residual aldehyde is pivotal for the binding of AOA, which is only effective on leaflet tissue, not on aortic wall.^[10] The fact that these detoxification treatments did not only remove tissue toxicity to an extent that endothelial cells could grow on the tissue^[8] but also significantly mitigated tissue calcification^[7,9,11] (Fig. 2) clearly proves that glutaraldehyde and its reaction products themselves contribute to calcification. The observation that leaflet calcification could be practically abolished without reversing GA cross-links refuted previous suspicions that GA cross-links themselves were the culprits of calcification.^[12]

Thus, one of the main thrusts of an engineering approach to cross-linking will be an optimization of cross-link efficiency *and* of tissue softness on the basis of a nontoxic cross-linking procedure. It is predictable that such an engineered cross-linking procedure will not be glutaraldehyde based.

FAILURE TO INACTIVATE CALCIUM BINDING SITES

Another principle behind bioprosthetic heart valve failure lies in the recognition of the fact that once nucleation is initiated, crystals tend to grow. Therefore, a deeper understanding of nucleation sites is required than the mere recognition that membrane phospholipids represent natural interaction sites for calcium ions. Recent insight into the structure and function of active calcium binding sites as well as into their ability to remain unaffected by contemporary cross-linking highlights the importance of a molecular focus on nucleation sites.^[13]

Although cellular elements were recognized as primary calcification sites for a long time, the process of calcification was mainly explained with passive

processes. The breakdown of an active gradient of a 10,000-times higher calcium concentration in the extracellular space and its uninhibited influx as a consequence of cell death and disintegration was seen as a main mechanism of early nucleation.^[14,15] It was hypothesized that phosphorus-containing sites such as phospholipids and proteolipids as well as enzymatic systems metabolizing high energy phosphates passively attract calcium ions.^[15] When it was discovered that alkaline phosphatase activity was still detectable after glutaraldehyde fixation,^[16] it did not lead to a revisiting of cross-link chemistry. It did, however, lead to an attempt to explain the anticalcification effect of Al^{+++} and Fe^{+++} with a competitive effect on the Zn^{++} and Mg^{++} binding sites of alkaline phosphatase.^[17]

A second look at the molecular events surrounding the calcium transport through the cell membrane and the molecular composition of active calcium binding sites made it clear that cross-linking approaches based on the utilization of ϵ -amino groups would not be able to prevent active calcium binding. By adapting the cross-link chemistry of the fixation process to the molecular needs of these calcium binding sites, membrane nucleation sites could practically be abolished.^[13]

Calcium plays an important role as an activator in the signaling of contraction in vascular smooth muscle cells. Calcium enters the cell via a passive voltage-gated channel and binds to intracellular calmodulin. The intracellular calcium-calmodulin complex activates phosphorylation of myosin and hence triggers smooth muscle cell contraction.^[18] In living cells, the intracellular free calcium concentration is approximately 10^{-7} M while the extracellular calcium concentration is approximately 10^{-3} M. The 10,000-times lower intracellular calcium concentration across the cell membrane is maintained by the action of the energy-dependent calcium pump Ca^{++} -ATPase, a transmembrane protein that can transport two Ca^{++} ions across the membrane

against a calcium gradient at the expense of the hydrolysis of one ATP molecule.^[19,20] Thus, calcium enters the cell by diffusion through gradient-dependent passive channels but is actively ejected from the cell via ATP pumps.^[21] In bioprosthetic tissue, cells are rendered nonviable by the fixation process. Apart from the loss of compartmental integrity of cells due to hypoxic disruptions prior to fixation, energy is not available in the cross-linked membranes and the calcium extrusion process is impaired. A longstanding theory held that membrane phospholipids bathing in the high concentrations of calcium form the nucleation sites for mineralization.^[22] It appears as if the Ca^{++} ATPase rather than phospholipids holds a key to membrane calcification.

The catalytic site of Ca^{++} ATPase comprising the ATP binding site and phosphorylation site were thought to be located on the cytoplasmic side with the high-affinity calcium binding site being buried in the membrane. Using site directed mutagenesis, two Ca^{++} binding sites were elucidated. Glu and Asp are associated with one site, Glu and Thr with the other, and Asp with both. This suggests that the two sites are distinguishable and that the spatial arrangement of Glu, Glu, Asp, Thr, and Asp can be superimposed in a membrane ion channel. Site-directed mutagenesis also revealed that an aspartate in the phosphate binding site is crucial for its catalytic effect.

A key to understanding calcium accumulation in the membrane was provided by the discovery that glutaraldehyde cross-links the ATP binding site of Ca^{++} ATPase but does not affect phosphorylation. Calcium can thus proceed with binding to the Ca^{++} binding sites but the subsequent transport across the membrane is blocked.^[23,24] This implies that lysyl residues are part of the ATP binding site. Alternatively, when carboxyl groups were cross-linked using carbodiimide-based fixatives, the Ca^{++} ATPase ability to phosphorylation was abolished.^[25] It was hypothesized that carbodiimide cross-links were introduced to the carboxyl groups on aspartate and glutamate side chains, thereby altering the tertiary structure of the Ca^{++} ATPase at Asp and Glu-rich Ca^{++} binding sites.

Similarly, calmodulin—a calcium modulator protein—binds up to four calcium ions, thereby transforming its structure and becoming activated. The activated calmodulin is able to interact with the myosin light chain kinase, causing contraction through phosphorylation. Calmodulin is a single chain of 148 amino acids, 25% of these, or 37, being either aspartate or glutamate.^[26] It is thought that the negatively charged carboxyl groups on aspartate and glutamate are crucially involved in calcium binding. Carbodiimide-based cross-linking of those aspartate and glutamate residues would then alter the spatial configuration of the protein, thereby inhibiting its calcium binding activity and

hence terminating the down-stream activation of phosphorylation. A further example for a molecule with a calcium binding site that is not accessible to glutaraldehyde but to carbodiimide-based carboxyl group fixation is alkaline phosphatase (AP). AP contains a calcium binding site whereby the calcium binding is coordinated by the carboxyl groups of Glu 218, Glu 274, Asp 289, and Phe 273, plus a water molecule. Mutations of any of these three—Glu 218, Glu 274, or Asp 289—result in a severe form of hypophosphatasia.^[27] Since the aspartate and glutamate amino acids are essential for the calcium binding potential of AP, carbodiimide-based cross-linking of carboxyl groups is hypothesized to prevent the escalation of phosphate concentrations through impairment of the phosphate hydrolytic mechanism.

If these membrane-located calcium binding molecules were really the mediators of membrane calcification, and the carboxyl group cross-linking of the aspartates and glutamates in their calcium binding domain were capable of blocking this reaction, it should be possible to suppress membrane bound nucleation sites in bioprosthetic tissue. In order to prove this, an extensive transmission electron microscopy study was performed whereby 16,000 individual calcium particles in bioprosthetic tissue either cross-linked with glutaraldehyde or carbodiimide were analyzed. The outcome of this analysis was impressive: Carboxyl group cross-linking dramatically suppressed membrane-bound nucleation sites, even if a carbodiimide treatment was added to a glutaraldehyde-based tissue fixation^[13] (Fig. 3).

The consequences of this observation are serious: First of all, by using a carboxyl group-based cross-linking procedure, the overall calcification of bioprosthetic tissue could be dramatically reduced. This is in accordance with the known phenomenon of a dominance of cell-associated calcification. Secondly, by eliminating the problem of cells acting as primary nucleation sites, decellularization attempts become less important. Thirdly, by eliminating the membranous nucleation sites, immediate fixation becomes less of an issue. Due to hypoxia-induced cell injuries, delayed fixation had significantly contributed to the calcification of bioprosthetic tissue.

REMNANT TISSUE IMMUNOGENICITY

The most fundamental principle behind the concept of xenograft prostheses is the assumption that immunogenicity is sufficiently suppressed. Although scientific proof has never been provided, there was a widely prevailing clinical perception that this assumption reflected the reality. Today, there is increasing



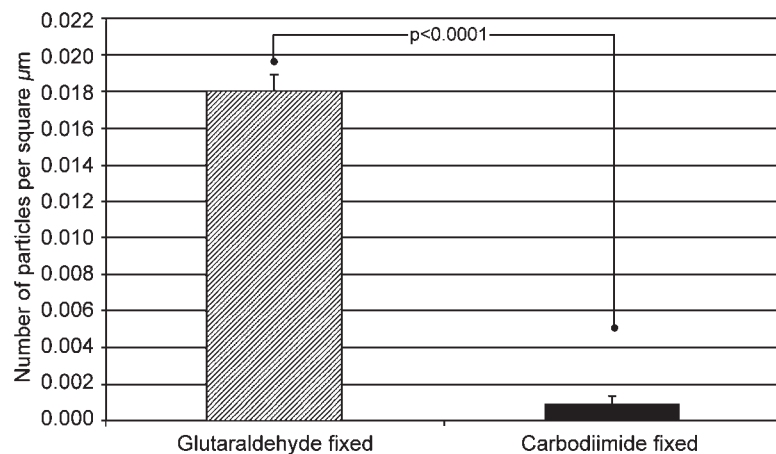


Fig. 3 Assessment of early calcification sites by transmission electron microscopy-based image analysis. There is a dramatic suppression of nucleation sites if carboxyl-group cross-linking (carbodiimide fixation) is used as opposed to amino group cross-linking (glutaraldehyde fixation).

recognition that the defense system of the recipient does in fact damage the implant through immune mechanisms.^[28-31]

It has long been known that the low-dose glutaraldehyde treatment applied to the fixation of bioprosthetic heart valves only reduces immunogenicity but does not abolish it. The low-grade fixation used in commercial valve preparation fails to alter membrane bound receptors or structural glycoproteins significantly,^[32] thus eliciting both cellular and humoral immune responses.^[30] Various groups, for instance, have demonstrated the generation of specific antibodies against pericardium fixed in 0.2% glutaraldehyde.^[30,33] The main criticism against these studies was the fact that their subcutaneous administration of immunogens hardly reflects the low-grade exposure to the immune system of intracardially placed bioprosthetic tissue. Yet, circulating antibodies were regularly detected in human recipients of bioprosthetic heart valves. In patients with porcine valves, specific antibodies were detected in as many as 58%.^[34] In homograft patients, 82% of recipients presented with a specific antibody response, 92% of which produced IgG against HLA class I antigens.^[35] In view of the role preformed antibodies play in the opsonization of antigen, the detection of preformed IgG antibodies in 67% of patients may certainly also have significance.^[29] Together, these IgG antibodies bind to the poorly masked tissue thereby inducing monocytes to both take up residence and differentiate into macrophages.^[36,37] In an attempt to mitigate tissue antigenicity, Nimni successfully increased the cross-link density in tissue.^[38] Subsequently, our own group investigated the ability of improved cross-linking of tissue to suppress specific circulating antibodies. While a significant antibody response was found against 0.2% GA-fixed tissue, there were practically no detectable graft-specific circulating antibodies in the enhanced fixation group (Fig. 4). When some of these circulating antibodies were identified, the strongest response was found against porcine fibronectin.

This discovery of a bioprosthetic immunogen other than the much-debated MHC molecules may be of particular importance in view of recent discussions surrounding MHC involvement in heart valves. Furthermore, the identification of fibronectin as immunogen also affirms previous studies in which extracellular matrix proteins such as collagen and elastin were found to elicit an immune response.^[39-41] It is against this background that we must see recent reports describing heart valve tissue as immune-privileged. The term immune-privileged was coined by the authors of a study that failed to detect class I antigens on valvular endothelial cells.^[42] This finding was not only refuted by others who did find class I antigens on valvular endothelial cells,^[43,44] but its overall relevance may be

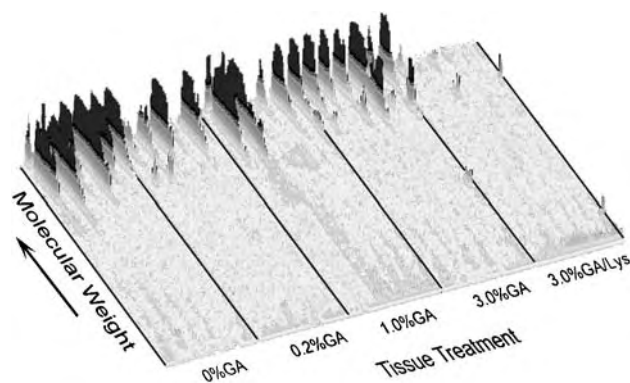


Fig. 4 Densitometric scan of Western blot analysis showing IgG response to porcine fibronectin in New Zealand White rabbits following immunization and boosting with homogenates of glutaraldehyde-fixed porcine aortic wall tissue in Freund's incomplete adjuvant. It is obvious that conventional (0.2% GA) and higher glutaraldehyde fixation (1.0% GA; 3.0% GA) were incapable of suppressing a specific humoral antibody (Ab) response. Only high-density cross-linking with diamine-enhanced 3% glutaraldehyde abolished xeno-antigen recognition leading to a specific Ab response.

diminished in view of the observation of class II antigens rather than class I antigens contributing to bioprosthetic valve failure^[45] and extracellular matrix proteins being capable of inducing a strong specific antibody response.

The most crucial step in proving the involvement of an immune response in tissue valve failure going beyond inflammatory destruction, however, was the demonstration of a link between an immune response and calcific degeneration. In the past few years, both Vincentelli's and our own group have provided this proof independently. Vincentelli et al. demonstrated in 1998^[31] that xenogenic tissue fixed in 0.65% GA calcified 35 times more in the same animal model compared to fixed autologous tissue. Our group found a direct correlation between cross-link density, inflammation, and calcification in both the rat^[46] and the sheep model.^[47] By subsequently demonstrating threefold higher calcification in aortic wall tissue that was preincubated in serum containing high levels of graft-specific antibodies^[48,49] we were eventually able to provide a direct immune-calcification link (Fig. 5).

For many years, this missing link between an immune response and tissue calcification was the main argument for contesting a role of immune mechanisms in bioprosthetic heart valve degeneration. Still, after strong evidence for such a link has now been provided, it will take a while and further experiments to lessen the opinion-forming impact of a study published in the early 1980s by Bob Levy's group. These authors described bioprosthetic calcification in both T-cell deficient mice^[50] and in cage models that prevented the direct interaction of the bioprosthetic tissue with the cellular immune system.^[51] In the meantime, deeper

insight into immune mechanisms has discovered T-cell independent (TI) humoral immune responses, whereby TI type 2 antigens may activate CD5+ B-lymphocytes in the congenitally athymic nude mouse.

The categorical exclusion of immunogenicity as a contributing factor to bioprosthetic tissue degeneration in the 1980s and early 1990s certainly contributed to the fact that observations in this direction started late and the mechanisms involved are still unidentified. However, it is surprising that the obvious conclusion of a link between insufficiently masked immunogenicity and bioprosthetic degeneration had not been drawn earlier. Differences in mid-to long-term responses between allografts and autografts highlighted the role of immune mechanisms in bioprosthetic graft failure particularly strongly and as early as 30 years ago. Although the collagenous elements of aortic roots do not differ between allografts and autografts, only the aortic wall collagen of allografts but not of autografts calcifies after implantation. This observation alone should have raised early suspicions regarding the possibility of immune calcification.

Nevertheless, once the focus had shifted, supportive data kept coming in. In allografts, for instance, with their milder immune differences, trends in clinical studies began to suggest a role for immune degeneration. Yacoub found an increased prevalence of valve degeneration in patients who had circulating antibodies against MHC molecules.^[35,44] O'Brien's group also found an association between allograft valve dysfunction and HLA-DR mismatch, equally suggesting that elements of the antidonor immune response penetrate and damage homografts.^[45] In xenografts, an interesting study by Eishi et al. in

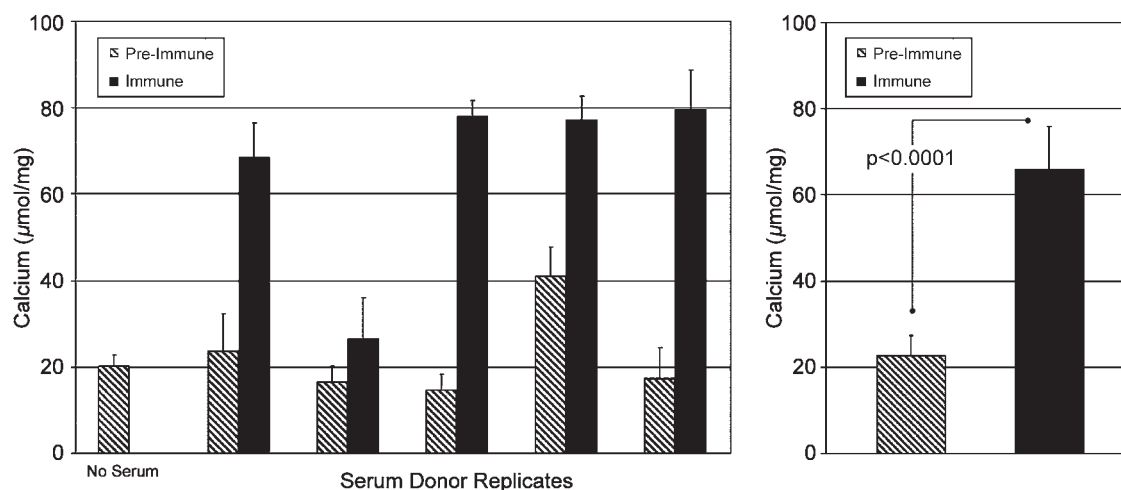


Fig. 5 Influence of a specific immune reaction between circulating antibodies and bioprosthetic aortic wall tissue on dystrophic calcification: After immunization of rabbits against aortic wall tissue, bioprosthetic tissue was incubated with either autologous preimmunization serum or autologous postimmunization serum. Calcification of the bioprosthetic tissue was distinctly higher if exposed to specific antibodies.

1996^[52] demonstrated in Takayasu patients requiring aortic valve replacement that the administration of steroids significantly suppressed calcification.^[52] Dahm et al. also contributed strong clinical evidence for the role of immune mechanisms in bioprosthetic valve failure. While 100% of patients with failed tissue valves had circulating antibodies against the bioprosthetic material, specific antibodies were found in only 25% of patients with intact valves.^[29] Further information contributing to the puzzle comes from Chauvaud.^[53] This report describes the observation that autologous pericardium failed to degenerate and calcify in a relatively large clinical series. Experimentally, it was also a rather recent observation that bioprosthetic tissue calcifies significantly less in the immune-tolerant Fischer rat than the other strains.^[54]

In spite of all the evidence, however, the lack of a meaningful hypothesis as to how an immune response augments tissue calcification emphasizes that we are still in a phase of observation rather than basic understanding. The main challenge ahead will thus be to both identify more tissue antigens eliciting a specific reaction and elucidate the exact mechanisms through which antibody binding facilitates degeneration. Once again, the medium- to long-term key to prevention seems to lie in either antigen removal or improved antigen masking. The discovery of the extracellular matrix being an antigen site that elicits a strong response makes it likely that cellular antigen removal without concomitant cross-linking or removal of ECM antigens will not succeed.

WEAKNESSES IN VALVE DESIGN

In porcine bioprostheses, valve design deals with the optimal suspension of cusps and commissural tissue on a frame in order to minimize potentially damaging stress and strain and to enable the surgeon to attach the valve to the patient's annulus with relative ease using a single row of sutures. In pericardial tissue valves, however, the term design implies significantly more than optimal suspension because the leaflet shape does not necessarily have to emulate native cusps and can thus be an integral part of the engineering efforts.

The overall goal of valve design is the minimization of stress concentrations within the cuspal tissue that may lead to early failure. At the beginning of such an engineering effort stands the assessment of the mechanics of native heart valves. Particularly, flexural stresses play an important role in limiting long-term durability. This highlights the necessity of an engineering approach that integrates material properties and cusp mechanics in a sophisticated way. Therefore, characterization and quantification of bioprosthetic leaflet motion during the opening and closing of the

valve is as important as the in-depth understanding of the fatigue life, compliance, and pliability of the prosthetic biomaterial.

Valve design is expected to profit significantly from technological progress both in computer modeling and in the development of test systems of recent years. Novel optical-based valve-imaging systems on the basis of a structured laser-light projection technique can obtain high resolution images of the complete valve leaflet-surface dynamically.^[55] One of the main shortcomings of previous modeling approaches has also been addressed: Mechanical properties of both native and bioprosthetic tissue were mostly evaluated on the basis of uniaxial testing that provided important qualitative information on the basis of non-physiological loading. After Mayne^[56] and Christie^[57] proposed equibiaxial testing that applies an equal level of tension to each test axis, Sacks developed full multiprotocol biaxial mechanical tests that include the entire range of physiological loading and thus provide sufficient data for constitutive computer modeling.^[58] Although these distinctly improved tools promise to reduce the influence of designer flaws on heart valve failure significantly, Vesely cautions against unanticipated failure modes^[59] that will always evade prediction. He argues that design flaws become evident only after a mode of failure has been identified and, as an example of unanticipated failure mode, mentions the chaotic vibrations that caused the crumbling of the Tacoma Narrows Bridge in 1940. Although bridges had been built for centuries, none of them had failed through this mode.

Vesely also used the example of a suspension bridge to demonstrate the main difference in challenge regarding the design of porcine and pericardial valves: Aortic valve cusps have a similar structure to a suspension bridge, where the main structural cables have the familiar upward curve as they approach the supporting towers. As the cable suspension becomes more vertical, the tension in the main cable decreases. In contrast, the almost cylindrical leaflet configuration in pericardial valves produces a nearly horizontal free edge that would imply extremely high tensile forces were it not for the fact that pericardial valve designers have been able to take advantage of the fact that the three leaflets lean against each other during coaptation.

Historically, probably the largest designer flaw was the so-called alignment stitch in the Ionescu-Shiley first-generation bovine pericardial valves. In recognition of the stress concentration around sutures through the pericardium, many designers have moved towards clamping the pericardial tissue between two solid surfaces. Possibly for cosmetic reasons, the Ionescu-Shiley valve employed a stitch near the free edge of the leaflet, exactly in the region that experienced particularly high membrane stresses. A large

proportion of these valves failed early, and pathological examinations indicated that the tear most likely originated at the site of the alignment stitch.^[60]

The second most consequential designer mistake was the attempt to change the height : diameter ratio (aspect ratio) of a porcine bioprosthesis in order to make the prosthesis as appealingly low profile as a pericardial valve. This BioImplant valve of the St. Jude range reduced the aspect ratio from 0.7 to 0.4.^[61] This dramatic reduction of the angle of the free edge of the porcine valve cusp led to a distinctly higher geometry-related stress. Early failures also forced this design into retreat.^[62]

A more recent development in valve design was the move towards stentless bioprostheses. The two main considerations behind this concept were on the one hand the reduction of outflow resistance due to a direct and indirect obstructive effect of the sewing ring and the stent, and on the other hand the theoretical advantage of mobile commissures. Since the aortic wall calcification occurring in all stentless models soon defied the advantage of mobile commissures, the only remaining advantage was the bigger orifice. The balance between a longer crossclamp time and a technically more challenging surgery on the one side and a larger orifice on the other has kept a lid on the gains in market share of stentless bioprosthetic heart valves.

Another shortcoming of test and prediction systems became apparent when Carbomedics-Sulzer launched their Photofix valve. Although photofixed valves performed well under in-vitro durability tests, they developed premature tears during clinical tests—most likely due to mechanical abrasion of the leaflets on the Dacron cloth of the sewing ring. It became obvious that the high-frequency in-vitro fatigue tests do not lead to the same leaflet loading as in-vivo implantations.

Overall, it appears as if both better insight into major design shortcomings of the past and significantly improved test and computation means have shifted the balance between designer flaws and biology-related reasons of failure further toward the latter.

WHERE TO?

If resistance toward paradigm shifts has stood in the way of a more progressive development in the field of bioprosthetic heart valves during the past 35 years, we are challenged to identify the creaks in the existing paradigms. We are also challenged to abandon the convenience and protection accepted paradigms provide. The main convenience of accepting an existing paradigm is to consent with the rest of the research community to take on digestible small bites rather than rock the boat and drown. By accepting this overall

approach, we can often achieve relatively much with regard to symptomatic improvements with relatively little pains. By not challenging, for instance, whether contemporary xenografts have managed to eliminate all detrimental aspects of xenogenicity and by accepting our decades-old approach to cross-linking as a tolerable compromise, we were able to focus happily on anticalcification treatments, minimally invasive surgery, or another few square millimeters of increased orifice areas.

What becomes increasingly clear, however, is the need to return openly to the drawing board. When venturing out initially to develop tissue heart valves, the main challenge clearly on everyone's mind was: How can we use animal tissue by making it unrecognizable as xenogenic and still preserve most of its mechanical function? Today we know that we partially failed in both: Albeit in a mitigated form, contemporary bioprosthetic heart valves continue to be recognized as foreign, while the cross-linking approach introduced a long list of disadvantages. So what do we need to do to make bioprosthetic heart valves unrecognizable as xenogenic and still preserve their mechanical function? The most likely answer is to eliminate xenogenicity and active calcification sites by rationally engineering cross-links in combination with support strategies that are done for the right reasons. One such support strategy could be decellularization of bioprosthetic tissue. The right reason for this is the elimination of cells as main carriers of antigens and calcification sites. A wrong reason may be to use it as a short cut to tissue engineering in the hope of cellular repopulation. Ingrowing connective tissue cells follow developmental signals in the matrix. A postdevelopmental mature matrix sends out the opposite signal: "Stay out, we are fine." Nothing demonstrates this principle better than homografts. These perfect emulations of an aortic root matrix remain mostly acellular for decades. Moreover, the demonstration of a degree of antigenicity residing in the extracellular matrix that is capable of triggering detrimental host responses makes it unlikely that decellularization will suffice without a concomitant cross-linking procedure. The need for such a concomitant cross-linking procedure, however, would almost certainly extinguish the last hope of repopulation with connective tissue cells because degradability is a cornerstone of remodeling. At the same time, deeper insight into the molecular mechanisms behind membrane-bound calcification sites and the ability for the first time to cross-link these sites effectively makes it less of an urgency to remove cellular remnants from the tissue. This condenses the focus area for a break-through development even further: cross-link engineering with particular emphasis on its molecular effect on antigen, calcification, and degradation sites on the one hand and its mechanical



effect on pliability, compliance, and break resistance on the other.

ARTICLES OF FURTHER INTEREST

Allografts; Collagen Fixation; Heart Valve, Bioprosthetic; Heart Valves, Mechanical; Host Reactions; Tissue Engineering of Heart Valves

REFERENCES

1. Odell, J.; Gillmer, D.; Whitten, I.; Vythilingun, S.; Vanker, E. Calcification of Tissue Valves in Children: Occurrence in Porcine and Bovine Pericardial Bioprosthetic Valves. In *Biologic and Bioprosthetic Valves*; Bodnar, E., Yacoub, M., Eds.; Yorke Medical Books: New York, 1985; 259.
2. Cheung, D.; Perelman, N.; Ko, E.; Nimni, M. Mechanism of crosslinking of proteins by glutaraldehyde III. Reaction with collagen in tissues. *Connect. Tissue Res.* **1985**, *13*, 109–115.
3. Hardy, P.; Nicholls, A.; Rydon, H. The nature of crosslinking of proteins by glutaraldehyde. 1. Interaction of glutaraldehyde with the amino-groups of 6-aminohexanoic acid and of alpha-N-acetyl-lysine. *J. Chem. Soc., Perkin* **1976**, *1*, 958–962.
4. Woodroof, E. Use of glutaraldehyde and formaldehyde to process tissue heart valves. *J. Bioeng.* **1978**, *2*, 1–9.
5. Johnson, T. Glutaraldehyde fixation chemistry: Oxygen-consuming reactions. *Eur. J. Cell Biol.* **1987**, *45*, 160–169.
6. Sacks, M. The biomechanical effects of fatigue on the porcine bioprosthetic heart valve. *J. Long-Term Eff. Med. Implants* **2001**, *11*, 231–247.
7. Weissenstein, C.; Human, P.; Bezuidenhout, D.; Zilla, P. Glutaraldehyde detoxification in addition to enhanced amine cross-linking dramatically reduces bioprosthetic tissue calcification in the rat model. *J. Heart Valve Dis.* **2000**, *9*, 230–240.
8. Zilla, P.; Fullard, L.; Trescony, P.; Meinhart, J.; Bezuidenhout, D.; Grolitzer, M.; Human, P.; von Oppell, U. Glutaraldehyde detoxification of aortic wall tissue: A promising perspective for emerging bioprosthetic valve concepts. *J. Heart Valve Dis.* **1997**, *6*, 510–520.
9. Dewanjee, M.; Solis, E.; Lanker, J.; Mackey, S.; Lombardo, G.; Tidwell, C.; Ellefsen, R.; Kaye, M. Effect of diphosphonate binding to collagen upon inhibition of calcification and promotion of spontaneous endothelial cell coverage on tissue valve prostheses. *ASAIO Trans.* **1986**, *32*, 24–29.
10. Chen, W.; Schoen, F.; Levy, R. Mechanism of efficacy of 2-amino oleic acid for inhibition of calcification of glutaraldehyde-pretreated porcine bioprosthetic heart valves. *Circulation* **1994**, *90*, 323–329.
11. Grabenwoger, M.; Grimm, M.; Eybl, E.; Leukauf, C.; Muller, M.; Plenck, H.J.; Bock, P. Decreased tissue reaction to bioprosthetic heart valve material after L-glutamic acid treatment. A morphological study. *J. Biomed. Mater. Res.* **1992**, *26*, 1231–1240.
12. Golomb, G.; Schoen, F.; Smith, M.; Linden, J.; Dixon, M.; Levy, R. The role of glutaraldehyde-induced crosslinks in calcification of bovine pericardium used in cardiac valve bioprostheses. *Am. J. Pathol.* **1987**, *127*, 122–130.
13. Han, R.; Zilla, P.; Hendricks, M.; Torrianni, M.; Human, P. Carboxyl-group crosslinking is essential to suppress cellular calcification in bioprosthetic tissue. *J. Thorac. Cardiovasc. Surg.* **2003**. (Submitted).
14. Kim, K. Pathologic Calcification. In *Pathobiology of Cell Membranes*; Trump, B., Arstilla, A., Eds.; Academic Press: New York, 1983; 117–155.
15. Schoen, F.; Tsao, J.; Levy, R. Calcification of bovine pericardium used in cardiac valve bioprostheses. Implications for the mechanisms of bioprosthetic tissue mineralization. *Am. J. Pathol.* **1986**, *123*, 134–145.
16. Levy, R.; Schoen, F.; Flowers, W.; Staelin, S. Initiation of mineralization in bioprosthetic heart valves: Studies of alkaline phosphatase activity and its inhibition by AlCl₃ or FeCl₃ preincubations. *J. Biomed. Mater. Res.* **1991**, *25*, 905–935.
17. Hirsch, D.; Schoen, F.; Levy, R. Effects of metallic ions and diphosphonates on inhibition of pericardial bioprosthetic tissue calcification and associated alkaline phosphatase activity. *Biomaterials* **1993**, *14*, 371–377.
18. Somlyo, A.P.; Somlyo, A.V. Signal transduction and regulation in smooth muscle. *Nature* **1994**, *372*, 231–236.
19. Mintz, E.; Guillain, F. Ca²⁺ transport by the sarcoplasmic reticulum ATPase. *Biochim. Biophys. Acta* **1997**, *1318*, 52–70.
20. Canet, D.; Forge, V.; Guillain, F.; Mintz, E. Ca²⁺ translocation across sarcoplasmic reticulum ATPase randomizes the two transported ions. *J. Biol. Chem.* **1996**, *271*, 20566–20572.
21. Anderson, H. Calcification processes. *Pathol. Annu.* **1980**, *15*, 45–75.
22. Schoen, F.J. Future directions in tissue heart valves: Impact of recent insights from biology and pathology [see comments]. *J. Heart Valve Dis.* **1999**, *8*, 350–358.
23. McIntosh, D.; Ross, D.; Champeil, P.; Guillain, F. Crosslinking the active site of sarcoplasmic reticulum Ca²⁺-ATPase completely blocks Ca²⁺ release to the vesicle lumen. *Proc. Natl. Acad. Sci. U. S. A.* **1991**, *88*, 6437–6441.
24. Ross, D.; Davidson, G.; McIntosh, D. Mechanism of inhibition of sarcoplasmic reticulum Ca²⁺-ATPase by active site cross-linking. Impairment of nucleotide binding slows nucleotide-dependent phosphoryl transfer, and loss of active site flexibility stabilizes occluded forms and blocks E₂-P formation. *J. Biol. Chem.* **1991**, *266*, 4613–4621.
25. Coan, C.; Jakobs, P.; Ji, J.; Murphy, A. Sarcoplasmic reticulum calcium ATPase. Labeling of a putative Mg²⁺ site by reaction with a carbodiimide and a spin-label. *FEBS Lett.* **1993**, *335*, 33–36.
26. Cheung, W. Calmodulin. *Sci. Am.* **1982**, *246*, 62–70.

27. Mornet, E.; Stura, E.; Lia, B.A.; Stigbrand, T.; Menez, A.; Le, D.M. Structural evidence for a functional role of human tissue nonspecific alkaline phosphatase in bone mineralization. *J. Biol. Chem.* **2001**, *276*, 31171–31178.
28. Human, P.; Zilla, P. Inflammatory and immune processes: The neglected villain of bioprosthetic degeneration? *J. Long-Term Eff. Med. Implants* **2001**, *11*.
29. Dahm, M.; Husmann, M.; Eckhard, M.; Prufer, D.; Groh, E.; Oelert, H. Relevance of immunologic reactions for tissue failure of bioprosthetic heart valves. *Ann. Thorac. Surg.* **1995**, *60*, S348–S352.
30. Dahm, M.; Lyman, W.; Schwell, A.; Factor, S.; Frater, R. Immunogenicity of glutaraldehyde-tanned bovine pericardium. *J. Thorac. Cardiovasc. Surg.* **1990**, *99*, 1082–1090.
31. Vincentelli, A.; Latremouille, C.; Zegdi, R.; Shen, M.; Lajos, P.; Chachques, J.; Fabiani, J. Does glutaraldehyde induce calcification of bioprosthetic tissues? *Ann. Thorac. Surg.* **1998**, *66*, S255–S258.
32. Carpentier, A.; Lemaigre, G.; Robert, L.; Carpentier, S.; Dubost, C. Biological factors affecting long-term results of valvular heterografts. *J. Thorac. Cardiovasc. Surg.* **1969**, *58*, 467–483.
33. Nimni, M. The cross-linking and structure modification of the collagen matrix in the design of cardiovascular prosthesis. *J. Card. Surg.* **1988**, *3*, 523–533.
34. Sheikh, K.; Tascon, M.; Nimni, M. Autoimmunity in patients with Hancock valve implant and bypass heart surgery. *Fed. Proc. Am. Soc. Exp. Biol.* **1980**, *39*, 472.
35. Smith, J.; Hornick, P.; Rasmi, N.; Rose, M.; Yacoub, M. Effect of HLA mismatching and antibody status on “homovital” aortic valve homograft performance. *Ann. Thorac. Surg.* **1998**, *66*, S212–S215.
36. Wewers, M.; Marsh, C. Role of the antibody in the pathogenesis of transplant vascular sclerosis: A hypothesis. *Transpl. Immunol.* **1997**, *5*, 283–288.
37. Jenney, C.; Anderson, J. Adsorbed serum proteins responsible for surface dependent human macrophage behavior. *J. Biomed. Mater. Res.* **2000**, *49*, 435–447.
38. Nimni, M.; Cheung, D.; Strates, B.; Kodama, M.; Sheikh, K. Chemically modified collagen: A natural biomaterial for tissue replacement. *J. Biomed. Mater. Res.* **1987**, *21*, 741–771.
39. Werkmeister, J.; Ramshaw, J.; Ellender, G. Characterisation of a monoclonal antibody against native human type I collagen. *Eur. J. Biochem.* **1990**, *187*, 439–443.
40. Werkmeister, J.; Ramshaw, J. Multiple antigenic determinants on type III collagen. *Biochem. J.* **1991**, *274* (Pt. 3), 895–898.
41. Gibson, M.; Hughes, J.; Fanning, J.; Cleary, E. The major antigen of elastin-associated microfibrils is a 31-kDa glycoprotein. *J. Biol. Chem.* **1986**, *261*, 11429–11436.
42. Kadner, A.; Chen, R.; Mitchell, R.; Adams, D. Lack of ABH-antigen expression on human cardiac valves. *J. Heart Valve Dis.* **2000**, *9*, 512–516.
43. Green, M.; Walsh, M.; Dare, A.; Hogan, P.; Zhao, X.; Frazer, I.; Bansal, A.; O’Brien, M. Histologic and immunohistochemical responses after aortic valve allografts in the rat. *Ann. Thorac. Surg.* **1998**, *66*, S216–S220.
44. Melina, G.; Rubens, M.; Birks, E.; Bizzarri, F.; Khashgani, A.; Yacoub, M. A quantitative study of calcium deposition in the aortic wall following Medtronic Freestyle compared with homograft aortic root replacement. A prospective randomized trial. *J. Heart Valve Dis.* **2000**, *9*, 97–103.
45. Dignan, R.; O’Brien, M.; Hogan, P.; Passage, J.; Stephens, F.; Thornton, A.; Harrocks, S. Influence of HLA matching and associated factors on aortic valve homograft function. *J. Heart Valve Dis.* **2000**, *9*, 504–511.
46. Zilla, P.; Weissenstein, C.; Bracher, M.; Zhang, Y.; Koen, W.; Human, P.; von Oppell, U. High glutaraldehyde concentrations reduce rather than increase the calcification of aortic wall tissue. *J. Heart Valve Dis.* **1997**, *6*, 502–509.
47. Zilla, P.; Weissenstein, C.; Human, P.; Dower, T.; von Oppell, U. High glutaraldehyde concentrations mitigate bioprosthetic root calcification in the sheep model. *Ann. Thorac. Surg.* **2000**, *70*, 2091–2095.
48. Human, P.; Zilla, P. Characterization of the immune response to valve bioprostheses and its role in primary tissue failure. *Ann. Thorac. Surg.* **2001**, *71*, S385–S388.
49. Human, P.; Zilla, P. Graft-specific antibody implicated in calcification of valve bioprosthetic tissue. *J. Cardiovasc. Pathol.* **2000**, *9*, 205.
50. Levy, R.; Schoen, F.; Howard, S. Mechanism of calcification of porcine bioprosthetic aortic valve cusps: Role of T-lymphocytes. *Am. J. Cardiol.* **1983**, *52*, 629–631.
51. Levy, R.; Schoen, F.; Levy, J.; Nelson, A.; Howard, S.; Oshry, L. Biologic determinants of dystrophic calcification and osteocalcin deposition in glutaraldehyde-preserved porcine aortic valve leaflets implanted subcutaneously in rats. *Am. J. Pathol.* **1983**, *113*, 143–155.
52. Eishi, K.; Ishibashi, U.H.; Nakano, K.; Kosakai, Y.; Sasako, Y.; Kobayashi, J.; Yutani, C. Calcific degeneration of bioprosthetic aortic valves in patients receiving steroid therapy. *J. Heart Valve Dis.* **1996**, *5*, 668–672.
53. Chauvaud, S.; Jebara, V.; Chachques, J.; el Asmar, B.; Mihaileanu, S.; Perier, P.; Dreyfus, G.; Relland, J.; Couetil, J.; Carpentier, A. Valve extension with glutaraldehyde-preserved autologous pericardium. Results in mitral valve repair. *J. Thorac. Cardiovasc. Surg.* **1991**, *102*, 171–177, discussion 177–178.
54. Nimni, M.; Myers, D.; Ertl, D.; Han, B. Factors which affect the calcification of tissue-derived bioprostheses. *J. Biomed. Mater. Res.* **1997**, *35*, 531–537.
55. Iyengar, A.; Sacks, M. Dynamic Imaging of BHV Cuspal Motion Using Laser Projection. The First Joint BMES/EMBS Conference: IEEE, Atlanta, USA, 1999.
56. Mayne, A.; Christie, G.; Smaill, B.; Hunter, P.; Barratt, B.B. An assessment of the mechanical properties of leaflets from four second-generation porcine bioprostheses with biaxial testing techniques [see comments]. *J. Thorac. Cardiovasc. Surg.* **1989**, *98*, 170–180.



57. Christie, G.; Barratt-Boyes, B. Age-dependent changes in the radial stretch of human aortic valve leaflets determined by biaxial testing. *Ann. Thorac. Surg.* **1995**, *60*, S156–S159.
58. Billiar, K.; Sacks, M. Biaxial mechanical properties of the natural and glutaraldehyde treated aortic valve cusp—Part I: Experimental results. *J. Biomech. Eng.* **2000**, *122*, 23–30.
59. Vesely, I. The influence of design on bioprosthetic valve durability. *J. Long-Term Eff. Med. Implants* **2001**, *11* (3&4), 137–149.
60. Walley, V.; Keon, W. Patterns of failure in Ionescu-Shiley bovine pericardial bioprosthetic valves. *J. Thorac. Cardiovasc. Surg.* **1987**, *9393*, 925–933.
61. Liotta, D.; Bracco, D.; Ferrari, H.; Bertolozzi, E.; Pisanu, A. Low profile bioprosthesis for cardiac valve replacement: Early clinical results. *Card. Dis. Bull. Tex. Heart Inst.* **1977**, *4*, 371–382.
62. Ius, P.; Thiene, G.; Minarini, M.; Valente, M.; Bortolotti, U.; Talenti, E.; Carsarotto, D. Low-profile porcine bioprosthesis (Liotta): Pathologic findings and mode of failure in the long-term. *J. Heart Valve Dis.* **1996**, *5*, 323–327.

Heart Valve Mechanics

Ajit Yoganathan

Biomedical Engineering, Georgia Institute of Technology, Atlanta, Georgia, U.S.A.

Anna M. Fallon

Chemical Engineering, Georgia Institute of Technology, Atlanta, Georgia, U.S.A.

Jorge H. Jimenez

Biomedical Engineering, Georgia Institute of Technology, Atlanta, Georgia, U.S.A.

INTRODUCTION

This chapter presents a detailed description of the function, hemodynamics, and mechanics of the four valves in the human heart. Heart valve function is essential to human life as a fundamental component of the cardiovascular system. The entry is divided into two main sections for the two main types of valves: semilunar and atrioventricular. Each of these sections is subdivided into valve structure and function, hemodynamics, and valve mechanics. As a whole, the entry provides a comprehensive description of each of the valves. This information is based mainly on experimental data due to the lack of reliable computational or analytical data. Despite the many years of research that have been devoted to these valves, there is still controversy over the detailed function and mechanics.

HEART VALVES

The human heart uses four distinctive valves to control pulsatile blood flow through its chambers and into the circulatory system. These valves are categorized into two major groups according to their location and structure: atrioventricular valves, which control blood flow between the atria and the ventricles; and semilunar valves, which control blood flow to the lungs and the systemic circulation. On the right side of the heart, blood flow between the atrium and the ventricle is controlled by the tricuspid valve, and flow from the ventricles into the lungs via the pulmonary arteries is controlled by the pulmonary valve. These valves are subjected to pressures up to 40 mmHg. On the left side of the heart, the flow from the atrium to the ventricle is controlled by the mitral valve, and flow to the systemic circulation via the aorta is controlled by the aortic valve. These valves are subjected to higher loads than those on the right side, with pressures up to 150 mmHg. The four heart valves open and close systematically to enable forward flow. Only during

isovolumetric relaxation and contraction are all four valves closed simultaneously for 0.02–0.06 s.

SEMILUNAR VALVES

Semilunar Valve Structure

The aortic valve is composed of three semilunar cusps, or leaflets, contained within a connective tissue sleeve. A thin fibrous layer called the intravalvular curtain separates the aortic valve cusps from the anterior leaflet of the mitral valve. When the aortic valve is closed, the cusps are perfectly aligned and separate the left ventricle from the aorta (Fig. 1a). Along with the three leaflets, there are three recesses, which form a bulging cavity behind the leaflets called the sinuses of Valsalva. At the lower margin the sinuses become continuous with the left ventricle, and at the upper margin they join with the sinotubular junction, which is the inlet of the aorta just downstream of the sinus. This area surrounding the valve is called the aortic root (Fig. 1b). Each valve cusp and corresponding sinus are named according to their anatomical locations within the aorta. Two of the sinuses give rise to coronary arteries that branch off the aorta, providing blood flow to the heart itself. The right coronary artery is located at the right anterior sinus, and the left coronary is located at the left anterior sinus. The third sinus is called the noncoronary or right posterior sinus. Because the cumulative circumferential length of the aortic valve cusps is greater than the annular radius, a small overlap of tissue from each leaflet protrudes and forms a coaptation surface when the valve is closed. This overlapped tissue, called the lunula, helps ensure that the valve is sealed properly. When the valve is open, the leaflets extend to the upper edge of the sinuses of Valsalva. The sinuses of Valsalva aid efficient valve closure by sustaining helical flow patterns behind the leaflets and dampen mechanical stresses imposed by the pressure difference when the valve is fully closed. The overall elasticity of the aortic root provides



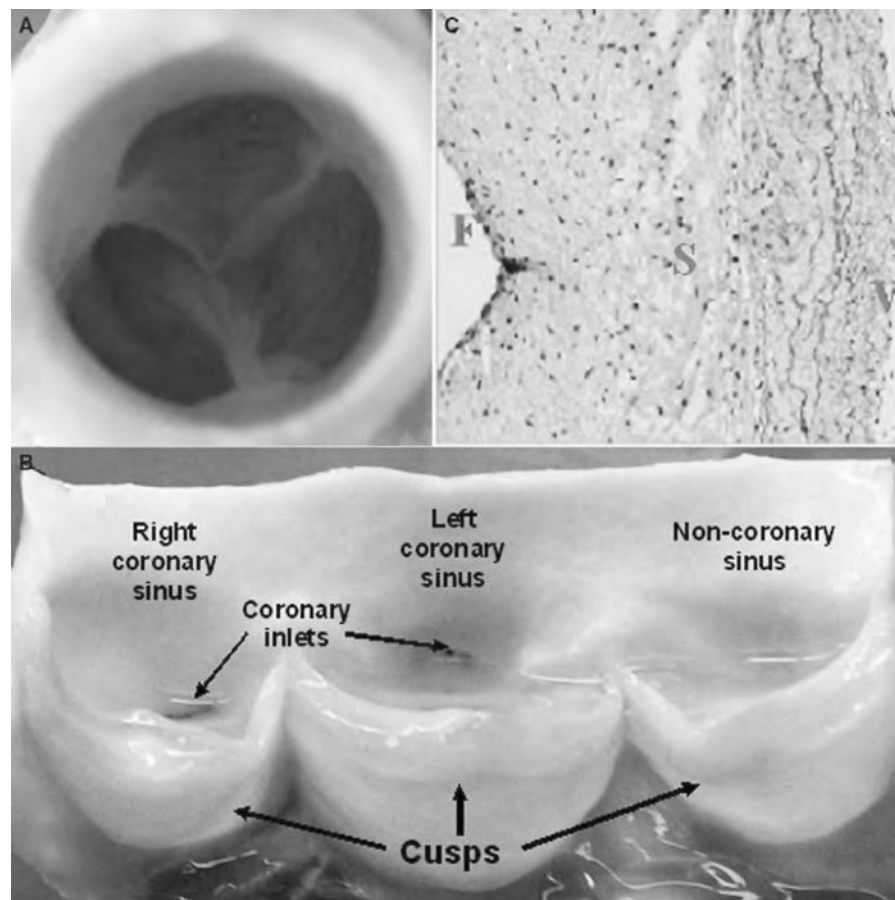


Fig. 1 (A) Closed aortic valve in an unloaded root. (B) Photograph of the aortic root, with its components. (C) Movat staining of porcine aortic valve leaflets: collagen = yellow, elastic fibers and cell nuclei = black, sGAG = light blue. F indicates the fibrosa side and V the ventricularis side. The three layers that compose the leaflet are identified in this image.

mechanical compliance, which reduces the mechanical loading on the valve.

The average aortic leaflet thickness is about 0.6 mm and varies considerably over the surface.^[1] The aortic valve leaflet appearance and size also vary depending on the leaflet. The posterior leaflet, for example, is generally thicker, has a larger surface area, and weighs more than the right or left leaflet,^[2] and the average width of the right aortic leaflet is greater than that of the other two.^[3] As shown in Fig. 1c, aortic valve leaflets have a complex, composite, trilayered structure consisting of the fibrosa, spongiosa, and ventricularis.^[4] Each of the leaflets is lined with endothelial cells and has a dense collagenous core adjacent to the high-pressure aortic side. The side adjacent to the aorta is termed the fibrosa and is the major fibrous layer within the belly of the leaflet. This layer bears most of the mechanical load of the leaflet. Beneath the fibrosa lies the spongiosa, which is a loose, watery connective tissue of varying thickness, fiber composition, and cellularity. This layer normally contains low quantities of collagen and elastin but large quantities of glycoaminoglycans (GAGs). The GAGs are responsible for the semifluid properties of the spongiosa, which enable this layer to reduce overall leaflet stresses due to

compliance mismatch between the layers.^[4,5] The ventricular surface of the leaflet is smooth in contrast to the ridged aortic surface to maintain laminar blood flow during systole.

The valve cusps are attached to a fibrous ring, or annulus, embedded in the ventricular septum. The aortic annulus separates the aorta from the left ventricle. According to echocardiographic measurements in healthy subjects the aortic root diameter is 35 ± 4.2 mm at end systole and 33.7 ± 4.4 mm at end diastole.^[6] High-resolution 3-D echocardiographic measurements of human aortic annular dynamics have shown that the maximum aortic orifice area occurs during early systole,^[7] and in vivo studies in sheep using radiopaque markers have shown this change in annular area to be $32 \pm 8\%$.^[8] The valve annulus also undergoes translation, primarily parallel to the aortic axis, which can be described as a line through the center of the aortic valve that runs parallel to the ascending aorta at the valve outlet. The aortic annulus moves toward the ventricle during systole and then recoils back toward the aorta as the ventricle fills during diastole. The annulus also experiences a slight side-to-side translation with a magnitude of approximately one-half the displacement along the aortic axis.

The overall anatomy of the pulmonic valve is similar to that of the aortic valve, but the dimensions of various structural components are different. On average, pulmonic leaflets are thinner than aortic leaflets, with an average thickness of 0.49 mm vs 0.67 mm.^[9] The main differences are that the sinuses are smaller in the pulmonary artery, and the pulmonic valve annulus is slightly larger than that of the aortic valve. M-mode echocardiography has shown a 3.15 ± 0.82 mm pre-systolic posterior motion of the pulmonary valve known as the A-wave. Although the origin of this motion is still under debate, it is likely associated with the motion of adjacent structures.^[10] In general, the motion of the pulmonic valve has not yet been fully described using more advanced imaging techniques.

Semilunar Valve Hemodynamics

The aortic valve opens during systole, when the ventricle is contracting, and then closes during diastole, as the ventricle relaxes. In healthy individuals, blood flows through the aortic valve accelerating to a peak value of 1.35 ± 0.35 m/s.^[11] The valve closes near the end of the deceleration phase of systole, with very little reverse flow through the valve. The adverse axial pressure difference causes the low-inertia flow in the developing boundary layer along the aortic wall to decelerate and then to reverse direction, resulting in vortices in the sinuses behind the aortic valve leaflets.^[12] This action forces the belly of the leaflets away from the aortic wall and toward the closed position. When this force is coupled with the vortices that push the leaflet tips toward the closed position, a very efficient and fast closure is obtained. In vitro studies have shown that the axial pressure difference alone is sufficient to close the valve.^[12] Thus, without the vortices in the sinuses, the valve still closes, but its closure is not as efficient as when the vortices are present.

The velocity profile at the level of the aortic valve annulus is relatively flat. There is a slight skew towards the septal wall (less than 10% of the center-line velocity), however, caused by the orientation of the aortic valve relative to the long axis of the left ventricle.^[13] The flow patterns just downstream of the aortic valve are of particular interest because of their complexity and relationship to arterial disease. Highly skewed velocity profiles and corresponding helical flow patterns have been observed in the human aortic arch using magnetic resonance phase velocity mapping.^[13]

The pulmonary valve flow profile is similar to that of the aortic valve, but the velocity magnitude is lower. Typical peak velocities at the valve outlet for healthy adults measured with 2-D and Doppler echocardiography are 0.75 ± 0.15 m/s.^[14] During acceleration the peak velocity is observed inferiorly, with this peak flow

rotating counterclockwise throughout the remainder of the ejection phase.^[15] The mean spatial profile is relatively flat, although a region of reverse flow occurs in late systole, which may be representative of flow separation. Typically, there is only a slight skew to the profile, and the peak velocity is generally within 20% of the spatial mean throughout the cardiac cycle. Secondary flow patterns can also be observed in the pulmonary artery and its bifurcation. In vitro laser Doppler anemometry experiments have shown that these flow patterns are dependent on the valve geometry and, thus, can be used to evaluate the function and fitness of the heart valve.^[16]

The ventricular surface of the leaflets experiences elevated shear stress during systole, when blood flows past the leaflets. On the aortic surface, oscillatory shear stress is experienced due to the vortices in the sinuses. To estimate the shear stresses on the leaflet surface, Laser Doppler Velocimetry (LDV) measurements were performed using a polymeric (polyurethane) trileaflet valve under pulsatile conditions. Weston et al.^[17] conducted two component LDV measurements inside a polyurethane valve as well as downstream of the valve. The maximum wall shear stress inside the valve was 79 dynes/cm² at a peak systolic flow rate of 22.5 L/min.

Semilunar Valve Mechanics

The aortic valve is exposed to a complex and harsh mechanical environment during the cardiac cycle. The leaflets are exposed to varying magnitudes of pressure, shear stress, and bending stress as they cyclically load and unload. Like most biological tissues, the aortic valve leaflets are anisotropic, nonhomogeneous, and viscoelastic. The collagen fibers within the fibrosa and ventricularis are crimped and entangled with no apparent directional preference in the unstressed state, but when a stress is applied, the elastin and collagen fibers become oriented primarily in the circumferential direction, with a lower concentration of elastin and collagen oriented in the radial direction.^[4] Vesely and Noseworthy^[18] found that both the ventricularis and fibrosa were stiffer in the circumferential direction than in the radial direction. The ventricularis was found to be more extensible radially than circumferentially, however, whereas the fibrosa had uniform extensibility in both directions. In lower concentrations, elastin fibers are oriented orthogonal to the collagen, and it is this mixed fiber structure that accounts for the anisotropic properties of the leaflets. The variation in thickness and composition across the leaflets is responsible for their nonhomogeneous material properties. Although the aortic valve leaflet as a whole is asymmetric in its distensibility, the basal region tends to be relatively isotropic, and the central region shows



the greatest degree of anisotropy.^[19] The elastin in the ventricularis consists of continuous amorphous sheets or compact meshes, whereas elastin in the fibrosa consists of complex arrays of large tubes that extend circumferentially across the leaflet.^[20] Mechanical testing of elastin structures from both the fibrosa and ventricularis has shown that the purpose of elastin in the aortic valve leaflet is to restore a specific collagen fiber orientation during cyclic loading.^[21] The viscoelastic properties of the valve are actually dominated by the elastic component over the current range of *in vitro* testing conditions, so that the viscous effects, which are largely responsible for energy dissipation, are small.^[4] The aortic valve leaflets are slightly stiffer than the pulmonary valve leaflets, although the extensibility and relaxation rates of the two tissues are similar.^[22] In biaxial studies of pulmonary and aortic valve leaflets,^[23] radial stretch was greater for the pulmonary leaflets than for the aortic leaflets, and circumferential stretch was similar. Thus, the ratio of radial to circumferential stretch was 6.0 ± 1.1 for the aortic leaflets and 9.0 ± 1.8 for the pulmonary leaflets.

Aortic valve leaflets have been shown *in vivo* to change length during the cardiac cycle.^[4] The cusps are longer during diastole than during systole, in both the radial and circumferential directions. Because the collagen fibers are oriented circumferentially, the valve is relatively stiff in this direction; therefore, the length variation is greatest in the radial direction. Because elastin is the primary element in the radial direction, the valve can undergo a great deal of strain, ranging from 20 to 60%.^[4,19] The strain in the circumferential direction is only about 10% due to the deformation restrictions imposed by the higher concentration of collagen fibers.^[19] The stretch in both directions results in an increased valve surface area during diastole. During systole, the shortening of the valve leaflets helps increase the effective orifice area during the ejection phase.

The aortic valve leaflets also undergo bending in the circumferential direction during the cardiac cycle. During diastole, when the valve is closed, each leaflet is convex toward the ventricular side. During systole, when the valve is open, the curvature changes, and each leaflet is concave toward the ventricle. This bending is localized on the valve cusp near the wall of the aorta, and this location is often thicker than the rest of the leaflet. Thubrikar used radiopaque markers to calculate the *in vivo* bending strains in canine aortic valves. The bending strains in the circumferential direction—calculated from the modulus of elasticity, thickness, and radius of the leaflet—were found to be 2% during systole and 2.2% during diastole.^[4] Because smooth bending requires internal tissue shearing, Talman and Boughner^[24] tested the shear stress vs shear strain characteristics both in the circumferential and

the radial directions of porcine aortic valve cusps. These results showed that the tissue behaves nonlinearly over the strain range. The average moduli at near-zero strains were less than 0.3 kPa and increased to more than 20 kPa at strains of approximately 0.9.

Although the pressure differential across the closed valve induces a large load on the cusps, the fibrous network within the cusps is effectively supported by the aortic wall and annulus. *In vivo* studies using radiopaque markers placed on canine aortic valve leaflets were conducted to estimate the leaflet stresses, using equations for membrane stress assuming a cylindrical geometry. The membrane stresses in the circumferential direction of the leaflet were 16.7 kPa during systole and 240.3 kPa during diastole.^[4] A finite element formulation was used to analyze the stresses based on a pressure gradient across the valve of 114.7 mmHg and a human aortic valve leaflet thickness of 0.6 mm.^[1] The maximum principal stress was found to be 220 kPa, which is comparable to the *in vivo* values. In a different study, the total diastolic stress in a valve leaflet has been estimated to be 250 kPa for a strain of 15%.^[4] The most recent finite element models include fluid coupling, aortic root compliance, and material fiber architecture. DeHart et al. used this type of model to demonstrate the effect of aortic root compliance on valve opening and closing configurations.^[25] This model predicts maximum tensile and compressive stresses of 91 and 8 kPa, respectively, during the cardiac cycle. The results from this study demonstrated that the presence of a compliant aortic root and directional fiber reinforcement will reduce the stress on the leaflets. In a comparative study of human aortic and pulmonary valves, the ultimate stress in the circumferential direction was 2.78 and 1.74 MPa for the pulmonary and aortic valves, respectively.^[26] The ultimate stress in the radial direction for the two valves was practically identical at 0.29 and 0.32 MPa for the pulmonary and aortic valves, respectively.^[26]

ATRIOVENTRICAL VALVES

Atrioventricular Valve Structure

Both the mitral valve and tricuspid valves are complex units comprised of an annulus, leaflets, chordae tendineae, and papillary muscles (PM). The annuli of these valves are located between the atrium and ventricle on the left and right sides of the heart for the mitral and tricuspid valves, respectively. The PM extend from the underlying myocardium. From the “belly” of each PM, several chordae tendineae extend, which insert into the leaflets or near the annulus, as well as into the left or right ventricular wall. This

arrangement provides continuity between the valve and ventricular wall to enhance ventricular function.

Although the annulus has been described anatomically as an incomplete structure, recent findings have redefined it as a vital component that promotes efficient valve closure and ventricular filling.^[27] The mitral and tricuspid annuli are elliptical rings composed of dense collagenous tissue surrounded by muscle. The anterior section of the mitral annulus connects to the aortic valve through the intravalvular curtain and is delimited by two cartilaginous masses called the annular trigones. In past years, researchers thought that this section of the mitral annulus was rigid, but a recent study has shown that it changes in length during the cardiac cycle.^[28] The sphincteric action of the annulus aids ventricular filling by expanding during diastole and facilitates leaflet coaptation by contracting during systole. Both human and animal studies of mitral annular dynamics have shown mitral annular contraction at the onset of left atrial contraction that continues until peak ventricular systole. In human studies, the annular area ranges from approximately 7.5 cm² during systole to approximately 10 cm² during diastole.^[27] Studies of the tricuspid annulus have also shown a change in annular area during the cardiac cycle, with a maximum annular area of 11.3 ± 1.8 cm² and a minimum area of 7.6 ± 1.4 cm².^[29]

The mitral annulus has a 3-D curvature that persists during the entire cardiac cycle.^[30] Based on recent studies, this geometry has been described as a 3-D saddle or a nonplanar ellipse. Using 3-D echocardiography, saddle heights of 1.2 ± 0.2 cm have been measured in humans.^[30] Although the geometry of the tricuspid annulus has not been fully described in humans, a recent study in sheep showed a saddle-shape annulus similar to that of the mitral valve.^[31]

The mitral and tricuspid annuli not only contract and bend, but also displace across the apical-basal axis of the ventricles during the cardiac cycle. During systole, the mitral and tricuspid annuli move apically 10 and 15 mm,^[26,32] respectively, from their most basal positions during diastole. The movement of the annulus toward the atrium during diastole may play a role in ventricular filling, possibly increasing the efficiency of blood flow into the ventricle.^[33]

Although there is valve-to-valve anatomical variability in mitral valve leaflets, several features are observed consistently in all normal mitral valve specimens (Fig. 2). The leaflet section of the mitral valve consists of a continuous veil of tissue,^[34] which inserts into the muscular annulus ring of the mitral orifice. The two major leaflet sections—the anterior and the posterior leaflets—are separated by two commissural sections located in the anterolateral and posteromedial sections of the annulus ring that are characterized by fanlike chordae tendineae insertions. This commissural section

is part of the posterior system; thus, the posterior leaflet in 92% of cases^[34] contains three scallops. The major central scallop is usually called the posterior leaflet; the other two are identified as the posteromedial commissural scallop and the anterolateral commissural scallop.

The anterior leaflet has a significantly larger area than the posterior leaflet and covers most of the mitral orifice during coaptation. The mitral valve has on average a leaflet surface area two times larger than the mitral orifice.^[35] The residual tissue enables the two leaflets to overlap during valve closure, producing a tight seal. Because the anterior leaflet covers most of the mitral orifice, it is subjected to a larger load than the posterior leaflet. Three major zones can be identified on both leaflets: a rough zone in the distal section, a clear zone in the center of the leaflet, and a basal zone at the annulus. Both the rough and basal zones have chordae tendineae insertions, but the clear zone is membranous, relatively smooth, and free of chordae tendineae insertions.^[34]

Unlike the mitral valve, the tricuspid valve has three leaflets: an anterior leaflet, a posterior leaflet with a variable number of scallops, and a septal leaflet (Fig. 3). The tricuspid valve is larger and structurally more complicated than the mitral valve, and the separation of the valve tissue into distinct leaflets is less pronounced than in the mitral valve. The leaflet surface is similar to that of the mitral valve, and the basal zone is present in all three leaflets.^[36]

The mitral and tricuspid leaflets are both composed of three layers: an atrial endothelial layer called the atrialis, an intermediate layer of fibrous material called the spongiosa, and a ventricular endothelial layer called the ventricularis. The atrioventricular valve leaflet tissue is primarily collagen-reinforced endothelium but also contains various types of cells, non-myelinated nerve fibers, and blood vessels, making them complex biological systems. Recent research has shown that interstitial cells in atrioventricular valve leaflets have a contractile phenotype.^[37] Therefore, the function of these leaflets may not be passive, although most researchers believe that the contractile nature of these cells is most likely related to leaflet tone.

The main function of the chordae tendineae is to prevent leaflet prolapse during ventricular systole. These “tension wires,” in addition to holding the leaflets in place during systole, maintain the geometrical conformation of the mitral valve and play an important role in left ventricular function. The chordae tendineae extend from both PM and insert into different sections of the leaflets. The chordae are characterized using a nomenclature based on insertion location with marginal, intermediate, and basal chordal insertion groups. This nomenclature also accounts for differences in composition, size, mechanical



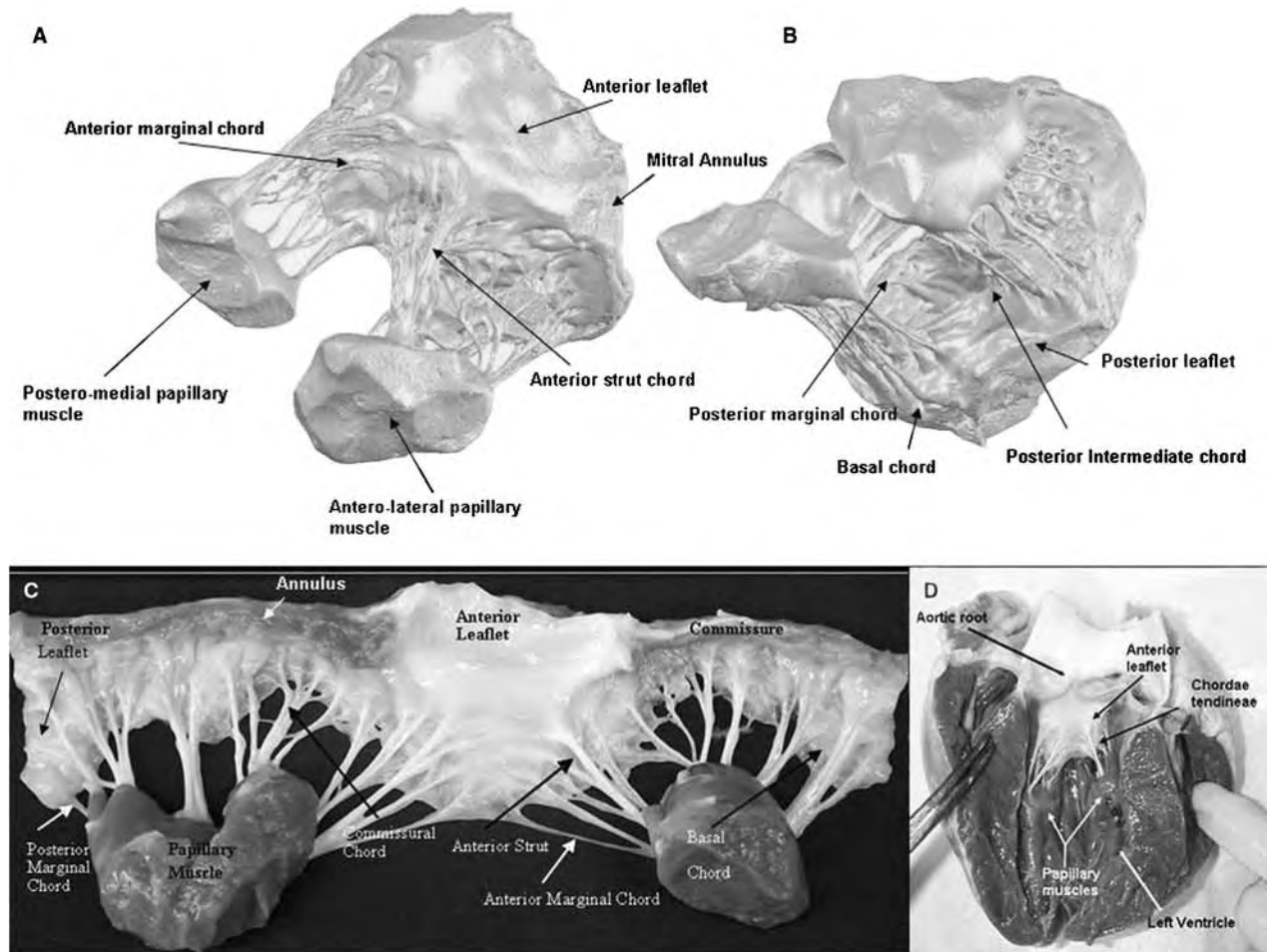


Fig. 2 Anatomy of the mitral valve. (A) Anterior view of a porcine mitral valve from a microCT reconstruction. (B) Posterior view of the same microCT reconstruction. (C) Photograph of an extended porcine mitral valve. (D) Photograph of a mitral valve within the left ventricle.

characteristics, and function. The marginal chordae extend from the PM or branch off from other chords and insert into the free margin of the anterior or posterior leaflet. The intermediate chordae originate at the PM and insert into the body of the ventricular surface of the leaflets. The basal chordae originate at the PM and insert adjacent to the posterior annulus. Normally, only two intermediate, or strut chordae, insert into the body of the anterior leaflet. The anterior strut chordae are subjected to tension during both diastole and systole. During systole, they bear a significant portion of the load on the anterior leaflet, thus preventing its prolapse, whereas during diastole, they restrict anterior motion of the leaflet, thus preventing the anterior leaflet from extending into the aortic outflow tract.^[38]

The chordae tendineae consist of an inner core of collagen surrounded by loosely meshed elastin and collagen fibers with an outer layer of endothelial cells. Recent studies of porcine chordae have also shown vascularization in some of the chordae. Both veins and

arteries run longitudinally through the intermediate layer of the chordae in a helical manner.^[39]

Porcine and human mitral valves are similar in structure, and they have similar chordae tendineae lengths.^[40] Thus, because there are limited data on the cross-sectional area of human chordae, most studies rely on porcine data. Using a porcine model, Sedransk et al. observed that on average the marginal chordae were 68% thinner than the basal chordae and that the chordae on the posterior leaflet were 35% thinner than their counterparts on the anterior leaflet of porcine mitral valves.^[41] By observing excised porcine chordae through a calibrated optical microscope, Liao et al. also found that assuming circular cross-sectional areas and uniform thickness, the average cross-sectional areas were $0.38 \pm 0.18 \text{ mm}^2$, $0.71 \pm 0.25 \text{ mm}^2$, and $2.05 \pm 0.40 \text{ mm}^2$ for the marginal, basal, and intermediate chordae, respectively.^[42]

In the mitral valve, the two PM (Fig. 2) protrude from the left ventricle with their tips pointing to their

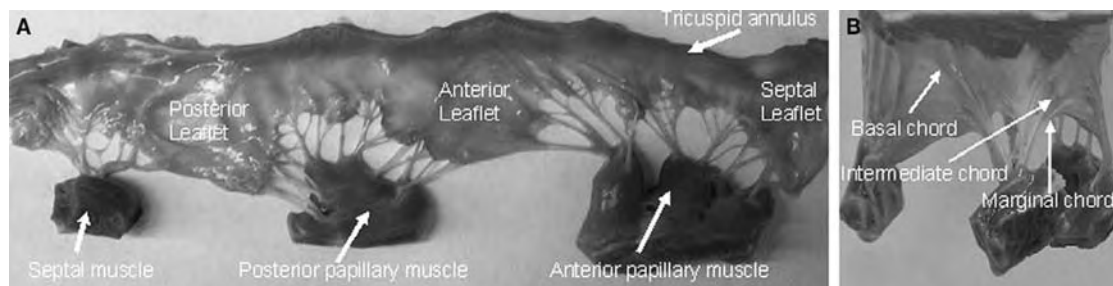


Fig. 3 Anatomy of the tricuspid valve. (A) Photograph of an extended tricuspid valve sectioned at the middle of the septal leaflet. The leaflets and PM are identified in the structure. (B) Photograph of an excised tricuspid valve identifying major chordal types.

respective commissures,^[43] thus directing their chordae to relatively symmetrical insertions into the valve. The tricuspid valve has three PM (Fig. 3A and B). The anterior papillary muscle, which is the largest, attaches to the valve at the commissure between the anterior and posterior leaflets. The posterior PM is located between the posterior and septal leaflets, and the smallest PM—the septal muscle—sometimes is not even present. Using transesophageal echocardiography, both the length and cross-sectional area of normal human mitral PM have been measured in vivo.^[44] These results show that in humans, the mitral valve PM contract approximately 4 mm during systole.

Atrioventricular Valve Hemodynamics

During isovolumic relaxation, the pressure in the left atrium exceeds that in the left ventricle, causing the mitral valve cusps to open. Blood then flows through the open valve from the left atrium to the left ventricle during diastole. The initial filling is enhanced by the active relaxation of the ventricle, which helps maintain a positive transmitral pressure gradient. A peak in the mitral flow curve, called the E-wave, occurs during the early filling phase, with normal peak velocities ranging from 50 to 80 cm/s.^[45] Following active ventricular relaxation, the fluid begins to decelerate, and the mitral valve partially closes. In late diastole, the atrium contracts, and the blood accelerates through the valve again to a secondary, lower velocity peak, termed the A-wave, with normal E/A velocity ratios ranging from 1.5 to 1.7.^[45]

The tricuspid flow profile is similar to that of the mitral valve, although the velocities in the tricuspid valve are significantly lower because it has a larger valve orifice. The peak early (E-wave) and late (A-wave) flow velocities across the tricuspid valve have been measured with Doppler echocardiography to be 0.51 and 0.35 m/s, respectively.^[46] Therefore, the E/A velocity ratios for the tricuspid valve are similar to those of the mitral valve.^[46] By contrast, the timing of

tricuspid valve opening is slightly different when compared with the mitral valve. Because the peak pressure in the right ventricle is less than that in the left ventricle, the right ventricular pressure falls below the right atrial pressure faster than the corresponding time for the left side of the heart. Thus, there is a shorter right ventricular isovolumic relaxation time, causing the tricuspid valve to open earlier. Additionally, tricuspid valve closure occurs after mitral valve closure, because the electrical stimulation of the left ventricle precedes that of the right ventricle.^[14]

Mitral valve fluid dynamics studies using magnetic resonance imaging (MRI) have shown that a large anterior vortex is normally present at the onset of partial valve closure as well as following atrial contraction.^[47] Bellhouse^[48] first suggested in an in vitro model that vortices generated by ventricular filling aid the partial closure of the mitral valve following early diastole and that without the strong outflow tract vortices, the valve would remain open at the onset of ventricular contraction. Later in vitro experiments suggested, however, that both flow deceleration and partial valve closure were due to an adverse pressure differential in mid-diastole, even in the absence of a ventricular vortex.^[49] Thus, although the vortices may provide additional closing effects in the initial stage, the adverse pressure gradient appears to be dominant in mitral valve closure. A more unified theory of valve closure includes the importance of chordal tension, flow deceleration, and ventricular vortices, with chordal tension being a necessary condition for the other two.^[50]

Atrioventricular Valve Mechanics

The mechanics of the mitral and tricuspid valves are complex, and there are limited studies on this subject. More studies have been performed on the mitral valve, because it sustains higher loads and, consequently, is more prone to disease. Although limited information is available on tricuspid valve mechanics, they are expected to be similar to those of the mitral valve.

The leaflet loading depends on the transmitral pressure, annular geometry and dynamics, coaptation geometry, chordae tendineae tension distribution, and contact forces between the leaflets. The curvature generated by the billowing of the leaflets during systole has been shown to be of mechanical importance, because it reduces the stress on the anterior leaflet.^[51] Additionally, a simplified computational model of the anterior leaflet has shown that the secondary curvature induced by the mitral annulus saddle shape is also a stress-relieving characteristic.^[52] This model proposes that a saddle height-to-commissural diameter ratio of approximately 20% generates the minimum stress configuration for the central region of the anterior leaflet under systolic loading. In a nonlinear, fluid-coupled, finite element model of the mitral valve, Einstein et al.^[53] predicted that the peak principal stress (254 kPa) during the cardiac cycle would occur in the belly of the anterior leaflet during peak systole (Fig. 4).

Sacks et al.^[54] and He et al.^[55] described *in vitro* not only the material characteristics of the anterior leaflet under physiological loads, but also how the underlying collagen matrix influences valve mechanics during coaptation by controlling directional strain through collagen fiber locking (Fig. 5). During valve closure, the anterior leaflet stretches in both the circumferential and the radial directions. Because the collagen fibers in the central region of the leaflet are arranged mainly in the circumferential direction, the leaflets become significantly stiffer in that direction as the collagen fibers uncrimp. A plateau at principal circumferential strains of 1.4 and radial strains of 1.2 indicates that the collagen fibers are uncrimped, thus limiting further deformation induced by the transmitral pressure. No material creep was observed during this plateau at constant transmitral pressure. Peak strains persist until the transmitral pressure decreases, after which the material returns to its relaxed conformation. Similar strain behavior was observed in the central region of the posterior leaflet,^[56] with some notable differences. The strain rates were lower in the posterior leaflet than those measured in the anterior leaflet, and a compressive strain at the beginning of valve closure was observed. Grashow et al.^[57] recently performed biaxial testing on the central region of the mitral valve anterior leaflet under physiological loading conditions. This study not only demonstrated the nonlinear response and anisotropy of the central anterior leaflet, but also found that the material response was independent of strain rate. As a result, the material was described as quasielastic and anisotropic. All these studies have been limited to the central regions of the mitral leaflets, which are relatively homogeneous. Small angle light scattering,^[54] however, has demonstrated that other regions of the anterior leaflet have more complex

collagen fiber distributions. Therefore, different material responses are expected for different regions of the leaflets.

Leaflet coaptation geometry is highly dependent on chordal insertion and chordal tension distribution. Nazari et al.^[58] described theoretically how valve leaflet stress distribution is directly related to chordae tendineae distribution. During coaptation, the load on the leaflets is transferred sequentially to increasingly larger chordae, leading to an optimal mechanical configuration during peak systole. He et al.^[59] defined a characteristic triangular structure between chordae that is fundamental to valve function. This triangle is formed when a smaller chordae branches off a larger chordae and the chordae insert to form a triangle between the two chordae and the leaflet. The disruption of this “He” triangle can lead to mitral regurgitation. Lomholt et al.,^[60] using an *in vivo* porcine model, showed that the dynamic tension on the intermediate chordae was three times as large as that on the marginal chordae, with a peak systolic tension of 0.7 N for the intermediate chordae and an average tension of 0.2 N on the marginal chordae. More recently, Jimenez et al.^[61] confirmed *in vitro* that the intermediate chordae hold significantly larger loads than the marginal chordae, with remarkably similar forces to those found *in vivo* (intermediate chordae=0.66 N, marginal chordae=0.22 N). Additionally, they found no significant differences between porcine and human chordae tendineae loads. Recent studies have also demonstrated that the saddle shape of the annulus results in a more even distribution of loads among the various chordae tendineae.^[62]

Recent studies have revealed differences in the mechanical characteristics of the chordae, which may be associated with their function. Uniaxial tensile tests have shown that porcine marginal chordae fail in response to significantly lower tension than the basal chordae. Because the marginal chordae have a smaller cross-sectional area, however, the basal chordae have significantly lower failure stress.^[41] Investigation of the tensile properties of chordae tendineae showed a nonlinear stress–strain relationship.^[63] The smaller chordae with cross-sectional areas of 0.001–0.006 cm² had a modulus of 2×10^9 dynes/cm², whereas larger chordae with cross-sectional areas of 0.006–0.03 cm² had a modulus of 1×10^9 dynes/cm². A more recent study confirmed that the extensibility of the chordae increases with diameter, whereas the modulus of elasticity decreases with increasing diameter.^[42] In this study, transmission electron microscopy observations showed that thicker chordae had a smaller collagen crimp period, which allowed increased strain on thicker chordae before the fibers locked. Although the cross-sectional area occupied by the collagen fibers was constant (49.2% on average), the thicker chordae had

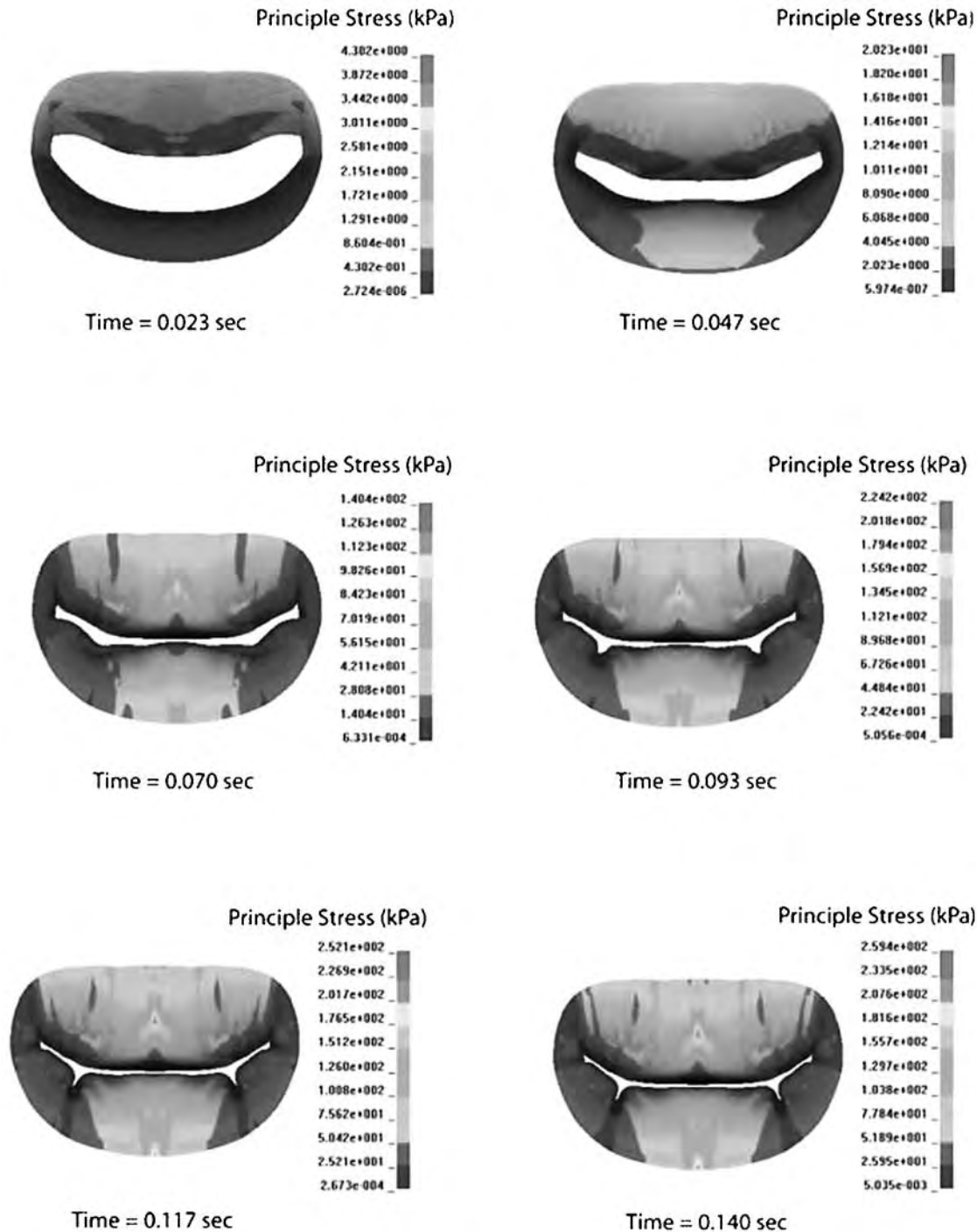


Fig. 4 Time course of principal stress on the mitral leaflets and annulus during the cardiac cycle. Initially, when the valve is fully opened, the largest stresses are concentrated around the trigones of the mitral annulus. As the valve is loaded by the transmitral pressure, the principal stress is transferred to the belly of the anterior leaflet during closure (Einstein 2005).

a larger average fibril diameter with a lower fibril density.

PM contraction, in conjunction with leaflet geometry and chordal loading, determines the force distribution on atrioventricular valves. An *in vitro* study of porcine mitral valves showed PM loads on the order of 4.4 N for

the posteromedial PM and 4.1 N for the anterolateral PM.^[64] This model did not simulate PM contraction; therefore, these results represent only the force present in the PMs due to valve closure without accounting for annular motion, ventricular motion, or PM contraction. PM material characteristics have also been studied using

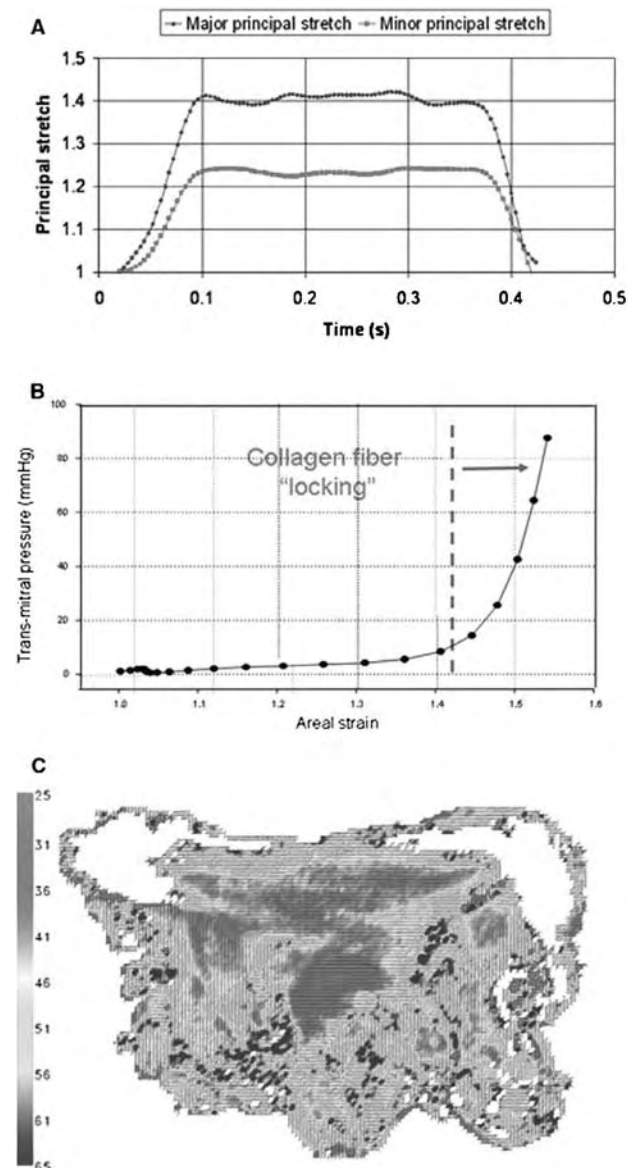


Fig. 5 (A) Graph of the major and minor principal stretches on the belly of the anterior leaflet during valve closure. As observed, the tissue is highly anisotropic. (B) Plot of trans-mitral pressure (load) vs the areal strain, demonstrating the nonlinear response of mitral leaflet tissue to physiological loads. The rapid rise in the curve is explained by collagen fiber locking. (C) Collagen fiber orientation map of the mitral valve anterior leaflet. Red represents circumferentially oriented fibers, and blue represents radially oriented fibers. Source: From Springer (see Ref. [54]).

animal tissue. Pinto and Fung^[65] demonstrated in unstimulated rabbit PM tissue that the relaxation function is independent of stretch ratio for strains under 30%. Creep tests with unstimulated PM tissue have also shown large elongation (creep strain) under a constant load. Under cyclic uniaxial testing, unstimulated PM tissue reached a steady-state hysteresis loop after

preconditioning that was unaffected by strain rate—typical of a pseudoelastic response. For both uniaxial loading and unloading, the stress on unstimulated PM tissue varies as an exponential function of strain.^[66] All these characteristics describe the viscoelastic nature of the unstimulated PM tissue.

The force from active myocardial contraction around the mitral annulus has not been measured directly. However, Einstein et al.^[53] using a nonlinear, fluid-coupled, finite element model of the mitral valve, estimated that the largest stresses around the annulus ring occurred during late diastole and early systole in the trigonal areas, which are the most rigid areas of the annulus. The elevated stresses in the trigone areas are transferred to the bellies of the leaflets during valve closure (Fig. 4).

SUMMARY

The four valves in the heart are complex biological systems, which have developed to serve specific functions. Their geometry, material characteristics, and biological components interact to provide extremely efficient functional characteristics and long-term durability under complex mechanical conditions. Heart valves are an example of how nature through evolution is one of the most elegant designers.

FUTURE DIRECTIONS

The understanding of heart valve mechanics is instrumental to the development of new technology to address the growing problem of heart valve disease in our aging population. Currently, most valve pathologies are addressed through palliative repair techniques or replacement with biological or mechanical prostheses. Both of these options are invasive, however. Current repair techniques have shown substandard long-term results, and current valve prostheses expose patients to life-threatening complications. A better understanding of the tissue mechanics, hemodynamics, and biological structure of both healthy and pathological valves is critical to the development of both effective and enduring repair techniques, as well as improved bioprosthetic valve choices such as tissue-engineered valves. Additionally, new imaging techniques and computerized models may aid in future surgical planning and the development of improved repair procedures.

ACKNOWLEDGMENTS

National Heart, Lung, and Blood Institute Grant # 52009. National Heart, Lung, and Blood Institute Grant # 067622.

ARTICLES OF FURTHER INTEREST

Artificial Heart Fluid Dynamics: Positive Displacement Pumps; Bioreactors; Blood Vessel Mechanics; Tissue Engineering of Heart Valves

REFERENCES

- Cataloglu, A.; Gould, P.L.; Asce, M.; Clark, R.E. Refined stress analysis of human aortic heart valves. *J. Eng. Mech. Div.* **1976**, *102*, 135–151.
- Silver, M.A.; Roberts, W.C. Detailed anatomy of the normally functioning aortic valve in hearts of normal and increased weight. *Am. J. Cardiol.* **1985**, *55*, 454–461.
- Vollebergh, F.E.M.G.; Becker, A.E. Minor congenital variations of cusp size in tricuspid aortic valves: Possible link with isolated aortic stenosis. *Br. Heart J.* **1977**, *39*, 1006–1011.
- Thubrikar, M. *The Aortic Valve*; CRC Press: Boca Raton, FL, 1990.
- Christie, G. Anatomy of aortic heart valve leaflets: The influence of glutaraldehyde fixation on function. *Eur. J. Cardiothorac. Surg.* **1992**, *6* (suppl 1), S25–S33.
- Gramiak, R.; Shah, P.M. Echocardiography of the normal and diseased aortic valve. *Radiology* **1970**, *96*, 1.
- Handke, M.; Jahnke, C.; Hienrichs, G.; et al. New three-dimensional echocardiographic system using digital radiofrequency data—visualization and quantitative analysis of aortic valve dynamics with high resolution. *Circulation* **2003**, *107*, 2876–2879.
- Timek, T.A.; Green, G.R.; Tibayan, F.A.; et al. Aorto-mitral annular dynamics. *Ann. Thorac. Surg.* **2003**, *76*, 1944–1950.
- David, H.; Boughner, D.R.; Vesely, I.; Gerosa, G. The pulmonary valve: Is it mechanically suitable for use as an aortic valve replacement? *ASAIO J.* **1994**, *40*, 206–212.
- Green, S.E.; Popp, R.L. The relationship of pulmonary valve motion to the motion of surrounding cardiac structures: A two-dimensional and dual M-mode echocardiographic study. *Circulation* **1981**, *64* (1), 107–112.
- Otto, C.M. Evaluation and management of chronic mitral regurgitation. *N. Engl. J. Med.* **2001**, *345* (10), 740–746.
- Reul, H.; Talukdar, N. Heart valve mechanics. In *Quantitative Cardiovascular Studies Clinical and Research Applications of Engineering Principles*; Hwang, N.H.C., Gross, D.R., Patel, D.J., Eds.; University Park Press: Baltimore, MD, 1979; 527–564.
- Kilner, P.J.; Yang, G.Z.; Mohiaddin, R.H.; Firmin, D.N.; Longmore, D.B. Helical and retrograde secondary flow patterns in the aortic arch studied by three-directional magnetic resonance velocity mapping. *Circulation* **1993**, *88* (part 1), 2235–2247.
- Oh, J.K.; Appleton, C.P.; Hatle, L.K.; Nishimura, R.A.; Seward, J.B.; Tajik, A.J. The noninvasive assessment of left ventricular diastolic function with two-dimensional and doppler echocardiography. *J. Am. Soc. Echocardiogr.* **1997**, *10*, 246–270.
- Sloth, E.; Houlind, K.C.; Oyre, S.; et al. Three-dimensional visualization of velocity profiles in the human main pulmonary artery using magnetic resonance phase velocity mapping. *Am. Heart J.* **1994**, *128*, 1130–1138.
- Sung, H.W.; Yoganathan, A.P. Axial flow velocity patterns in a normal human pulmonary artery model: Pulsatile in vitro studies. *J. Biomech.* **1990**, *23* (3), 210–214.
- Weston, M.W.; LaBorde, D.V.; Yoganathan, A.P. Estimation of the shear stress on the surface of an aortic valve leaflet. *Ann. Biomed. Eng.* **1999**, *27*, 572–579.
- Vesely, I.; Noseworthy, R. Micromechanics of the fibrosa and the ventricularis in aortic valve leaflets. *J. Biomech.* **1992**, *25*, 101–113.
- Lo, D.; Vesely, I. Biaxial strain analysis of the porcine aortic valve. *Ann. Thorac. Surg.* **1995**, *60* (suppl II), S374–S378.
- Scott, M.J.; Vesely, I. Morphology of porcine aortic valve cusp elastin. *J. Heart Valve Dis.* **1996**, *5*, 464–471.
- Vesely, I. The role of elastin in aortic valve mechanics. *J. Biomech.* **1998**, *31*, 115–123.
- Leeson-Dietrich, J.; Boughner, D.; Vesely, I. Porcine pulmonary and aortic valves: A comparison of their tensile viscoelastic properties at physiological strain rates. *J. Heart Valve Dis.* **1995**, *4*, 88–94.
- Christie, G.W.; Barratt-Boyes, B.G. Age-dependent changes in the radial stretch of human aortic valve leaflets determined by biaxial testing. *Ann. Thorac. Surg.* **1995**, *60*, S156–S159.
- Talman, E.A.; Boughner, D.R. Internal shear properties of fresh porcine aortic valve cusps: Implications for normal valve function. *J. Heart Valve Dis.* **1996**, *5*, 152–159.
- DeHart, J.; Baaijens, F.P.; Peters, G.W.; Schreurs, P.J. A computational fluid–structure interaction analysis of a fiber-reinforced stentless aortic valve. *J. Biomech.* **2003**, *36* (5), 699–712.
- Stradin, P.; Lacis, R.; Ozolanta, I.; et al. Comparison of biomechanical and structural properties between human aortic and pulmonary valve. *Eur. J. Cardiothorac. Surg.* **2004**, *26* (3), 634–639.
- Timek, T.A.; Miller, D.C. Experimental and clinical assessment of mitral annular area and dynamics: What are we actually measuring? *Ann. Thorac. Surg.* **2001**, *72*, 966–974.
- Parish, L.M.; Jackson, B.M.; Enomoto, Y.; Gorman, J.H., III. The dynamic anterior mitral annulus. *Ann. Thorac. Surg.* **2004**, *78* (4), 1248–1255.
- Tei, C.; Pilgrim, J.P.; Shah, P.M.; Ormiston, J.A.; Wong, M. The tricuspid valve annulus: Study of size and motion in normal subjects and in patients with tricuspid regurgitation. *Circulation* **1982**, *66* (3), 665–671.
- Kaplan, S.R.; Bashein, G.; Sheehan, F.H.; et al. Three-dimensional echocardiographic assessment of annular shape changes in the normal and regurgitant mitral valve. *Am. Heart J.* **2000**, *139*, 243–250.
- Hiro, M.E.; Jouan, J.; Pagel, M.R.; et al. Sonometric study of the normal tricuspid valve annulus in sheep. *J. Heart Valve Dis.* **2004**, *13* (3), 452–460.
- Miller, D.; Farah, M.G.; Liner, A.; Fox, K.; Schluchter, M.; Hoit, B.D. The relation between quantitative right



- ventricular ejection fraction and indices of tricuspid annular motion and myocardial performance. *J. Am. Soc. Echocardiogr.* **2004**, *17* (5), 443–447.
33. Alam, M.; Rosenhamer, G. Atrioventricular plane displacement and left ventricular function. *J. Am. Soc. Echocardiogr.* **1992**, *5*, 427–433.
 34. Kalmanson, D. *The Mitral Valve: A Pluridisciplinary Approach*; Publishing Science Group: Acton, MA, 1976; [Chapters 1–5], 3–45.
 35. He, S.; Lemmon, J.D.; Weston, M.W.; Levine, R.A.; Yoganathan, A.P. Mitral valve compensation for annular dilation: In vitro study into the mechanisms of functional mitral regurgitation with an adjustable annulus model. *J. Heart Valve Dis.* **1999**, *8*, 294–302.
 36. Silver, M.D.; Lam, J.H.C.; Raganathan, N.; Wigle, E.D. Morphology of the human tricuspid valve. *Circulation* **1971**, *43*, 333–348.
 37. Chester, A.H.; Misfeld, M.; Yacoub, M.H. Receptor-mediated contraction of aortic valve leaflets. *J. Heart Valve Dis.* **2000**, *9*, 250–255.
 38. Ritchie, J.; Jimenez, J.H.; He, Z.; Sacks, M.S.; Yoganathan, A.P. The material properties of the native porcine mitral valve chordae tendineae: An in vitro investigation. *J. Biomech.* **2006**, *39* (6), 1129–1135.
 39. Ritchie, J.; Warnock, J.N.; Yoganathan, A.P. Structural characterization of the chordae tendineae in native porcine mitral valves. *Ann. Thorac. Surg.* **2005**, *80* (1), 189–197.
 40. Kunzelman, K.S.; Cochran, R.P.; Verner, E.D.; Eberhart, R.D. Anatomic basis for mitral valve modeling. *J. Heart Valve Dis.* **1994**, *3*, 491–496.
 41. Sedransk, K.L.; Allen, J.G.; Vesely, I. Failure mechanics of mitral valve chordae tendineae. *J. Heart Valve Dis.* **2002**, *11*, 644–650.
 42. Liao, J.; Vesely, I. A structural basis for the size-related mechanical properties of mitral valve chordae tendineae. *J. Biomech.* **2003**, *36* (8), 1125–1133.
 43. Rusted, I.E.; Schiefley, C.H.; Edwards, J.E. Studies of the mitral valve: I. Anatomical features of the normal mitral valve and associated structures. *Circulation* **1952**, *6* (6), 825–831.
 44. Madu, E.C.; Baugh, D.S.; D’Cruz, I.A.; Johns, C. Left ventricular papillary muscle morphology and function in left ventricular hypertrophy and left ventricular dysfunction. *Med. Sci. Monit.* **2001**, *7* (6), 1212–1218.
 45. Weyman, A.E. In *Principles and Practices of Echocardiography*; Lea & Febiger: Philadelphia, PA, 1994.
 46. Pye, M.P.; Pringle, S.D.; Cobbe, S.M. Reference values and reproducibility of Doppler echocardiography in the assessment of the tricuspid valve and right ventricular diastolic function in normal subjects. *Am. J. Cardiol.* **1991**, *67* (4), 269–273.
 47. Kim, W.Y.; Walker, P.G.; Pederson, E.M.; Poulsen, J.K.; Houliand, K.C.; Oyre, S. Left ventricular blood flow patterns in normal subjects: A quantitative analysis of three-dimensional magnetic resonance velocity mapping. *J. Am. Coll. Cardiol.* **1995**, *26* (1), 224–238.
 48. Bellhouse, B.J. The fluid mechanics of a model mitral valve and left ventricle. *Cardiovasc. Res.* **1972**, *6*, 199–210.
 49. Reul, H.; Talukder, N.; Muller, E.W. Fluid mechanics of the natural mitral valve. *J. Biomech.* **1981**, *14* (5), 361–372.
 50. Yellin, E.L.; Peskin, C.; Yoran, C.; et al. Mechanisms of mitral valve motion during diastole. *Am. J. Physiol.* **1981**, *241*, H389–H400.
 51. Arts, T.; Meerbaum, S.; Reneman, R.; Corday, E. Stresses in the closed mitral valve: A model study. *J. Biomech.* **1983**, *16*, 539–547.
 52. Salgo, I.S.; Gorman, J.H., III; Gorman, R.C.; et al. Effect of annular shape on leaflet curvature in reducing mitral leaflet stress. *Circulation* **2002**, *106*, 711–717.
 53. Einstein, D.R.; Kunzelman, K.S.; Reinhall, P.G.; Nicosia, M.A.; Cochran, R.P. Non-linear fluid-coupled computational model of the mitral valve. *J. Heart Valve Dis.* **2005**, *14* (3), 376–385.
 54. Sacks, M.S.; He, Z.; Baijens, L.; et al. Surface strains in the anterior leaflet of the functioning mitral valve. *Ann. Biomed. Eng.* **2002**, *30*, 1281–1290.
 55. He, Z.; Sacks, M.S.; Baijens, L.; Wanant, S.; Shah, P.; Yoganathan, A.P. Effects of papillary muscle position on the in vitro dynamic strain on the porcine mitral valve. *J. Heart Valve Dis.* **2003**, *12*, 488–494.
 56. He, Z.; Ritchie, J.; Grashow, J.S.; Sacks, M.S.; Yoganathan, A.P. In vitro dynamic strain behavior of the mitral valve posterior leaflet. *J. Biomech. Eng.* **2005**, *127* (3), 504–511.
 57. Grashow, J.S.; Yoganathan, A.P.; Sacks, M.S. Biaxial stress-stretch behavior of the mitral valve anterior leaflet at physiologic strain rates. *Ann. Biomed. Eng.* **2006**, *34* (2), 315–325.
 58. Nazari, S.; Carli, F.; Bnfi, C.; et al. Patterns of systolic stress distribution on mitral valve anterior leaflet chordal apparatus. *J. Cardiovasc. Surg.* **2000**, *41*, 193–202.
 59. He, S.; Weston, M.W.; Lemmon, J.; Jensen, M.; Levine, R.A.; Yoganathan, A.P. Geometric distribution of chordae tendineae: An important anatomic feature in mitral valve function. *J. Heart Valve Dis.* **2000**, *9*, 495–501.
 60. Lomholt, M.; Nielsen, S.L.; Hansen, S.B.; Andersen, N.T.; Hasenkam, J.M. Differential tension between secondary and primary mitral chordae in acute in-vivo porcine model. *J. Heart Valve Dis.* **2002**, *11*, 337–345.
 61. Jimenez, J.H.; Soerensen, D.D.; He, Z.; Ritchie, J.; Yoganathan, A.P. Effects of a papillary muscle position on chordal force distribution: An in vitro study. *J. Heart Valve Dis.* **2005**, *14*, 295–302.
 62. Jimenez, J.H.; Soerensen, D.D.; He, Z.; He, S.; Yoganathan, A.P. Effects of a saddle shaped annulus on mitral valve function and papillary muscle position. *Ann. Biomed. Eng.* **2003**, *31*, 1171–1181.
 63. Lim, K.O.; Bouchner, D.P. Mechanical properties of human mitral valve chordae tendineae: Variation with size and strain rate. *Can. J. Physiol. Pharmacol.* **1975**, *53*, 330–339.
 64. Jensen, M.O. In *Stentles Mitral Valve Fixation: Impact on Hemodynamic Performance [Master’s Thesis, Department of Biomedical Engineering]*; Georgia Institute of Technology: Atlanta, GA, 2000.
 65. Pinto, J.G.; Fung, Y.C. Mechanical properties of the heart muscle in the passive state. *J. Biomech.* **1973**, *6*, 596–616.
 66. Fung, Y.C. In *Biomechanics: Mechanical Properties of Living Tissues*; Springer: New York, 1981.

Heart Valves, Mechanical

Ajit P. Yoganathan

Zhaoming He

Hwa-Liang Leo

Wallace H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology, Atlanta, Georgia, U.S.A.

Anna Fallon

Chemical Engineering, Georgia Institute of Technology, Atlanta, Georgia, U.S.A.



INTRODUCTION

The first reported mechanical heart valve surgery, which involved a caged ball valve design, was performed by Judson Chesterman in 1955 in England. The patient passed away 14 hours later as a result of mechanical failure of the heart valve. This was followed by the first long-term success in 1961, when Dr. Charles Hufnagel implanted caged ball heart valves into eight patients for aortic insufficiency.^[1] Six of the eight patients survived the operation. This marked the first breakthrough in the design of the artificial heart valve, which is capable of reducing turbulence and blood trauma through the use of biocompatible materials and ingenious designs. During over 40 years of heart valve research and development, the design of heart valves has diversified into a myriad of more than a dozen brands, each varying in mode of operation. The most commonly used designs are the Starr-Edward caged-ball valve, the Bjork-Shiley and Medtronic-Hall tilting disc valves, and the St. Jude Medical and CarboMedics bileaflet valves. Unfortunately, even after many years of experience and success, artificial heart valves are still plagued by the problems of thrombosis and thromboembolism, tissue overgrowth, infection, anticoagulant-related hemorrhage, and valve failure due to material fatigue or chemical change.^[2] This article describes the mechanical durability of the aforementioned valve types, their fluid dynamic characteristics, and the implication of thrombus formation and deposition as a result of the flow structures that arise.

TYPES OF MECHANICAL HEART VALVES

Generally, mechanical heart valves (MHVs) are divided into caged-ball valves, tilting disc valves and bileaflet valves (Fig. 1). The stationary supporting metal struts and the mobile occluders used in these valve designs are typically made of pyrolytic carbon and/or titanium that may be coated with pyrolytic carbon. The material used for the sewing ring cuff is usually Teflon. The only ball-and-cage valve presently available in the United States is

the Starr-Edwards 1260 aortic caged ball model, used since 1966. Its design consists of a radiolucent silicone ball encased by a closed Stellite-alloy cage. The radiopaque wire struts emanate from the valve housing and are spaced 120 degrees apart. They converge at the distal end of the valve forming the apex of the cage. The tilting disc design consists mainly of a circular occluder mounted and secured on a metal strut that controls the range of occluder motion. A slight obstruction downstream of the minor orifice is inherent in the tilting disc design as a result of the protruding metal strut and occluder disc. The Medtronic-Hall and the Bjork-Shiley monostrut tilting disc valves represent the tilting disc designs presently used. The bileaflet prosthetic designs currently available in the United States are the St. Jude Medical (SJM) and CarboMedics valves. The bileaflet valve consists of two semicircular leaflets that are attached to the valve housing with hinge mechanisms. The bileaflet valve design provides a relatively large orifice area for a small valve size; consequently, there is a smaller pressure drop across the valve.

DURABILITY OF MECHANICAL HEART VALVES

The valve design configuration affects the load distribution and dynamics of the valve components, which, in conjunction with the material properties, determines the durability—notably wear and fatigue life. The flow through valves results in little wear in terms of shear on the surface of valve components because blood contains no rigid particles. The transvalvular pressure generated at and after valve closure predominantly determines the load on valve components such as the housing ring, strut, seat stop, and occluder. The impact of the occluder on the housing ring or seat stop and hinges plays a primary role in structural failure of MHVs. Polymers such as silicone rubber, polypropylene, Delrin, and Teflon used in early valves were rejected due to poor wear and fatigue performance. Since the mid-1970s the prevalent materials used in MHVs are titanium, chrome-cobalt alloy, and pyrolytic carbon. Pure pyrolytic carbon was recently

developed to be stronger and more biocompatible than current silicone-alloyed pyrolytic carbon.^[3]

The failure modes of valves are grouped into two categories: failure by intrinsic mechanisms or by extrinsic mechanisms. Failures by intrinsic mechanisms are inherently related to design and material properties such as fatigue fracture, and to abrasive wear or biochemical attack. Extrinsic mechanisms are externally generated interference with the motion of a valve's occluding mechanism. These include massive thrombotic obstruction causing impingement of occluders on the adjacent anatomy, or lockup.^[4] The intrinsic failure mechanisms will be discussed first.

Before pyrolytic carbon was used for MHVs, occluders fabricated from polymers had to be designed to rotate and distribute the wear. The early caged ball valve has a long clinical history. There have been failure reports of uncured silicone rubber balls degrading due to absorption of lipids, but properly cured balls have worked well, with no documented structural failures. The early Bjork-Shiley valves used a Delrin occluder, but had problems with heat sterilization, deformation, and high wear rate. Presently, all tilting disc valves use pyrolytic carbon for the disc and titanium or graphite carbon coated with pyrolytic carbon for the housing ring. There have been strut fractures and failures of the Bjork-Shiley convexo-concave valve design, which led to the withdrawal of this design from clinical use in the mid-1980s.

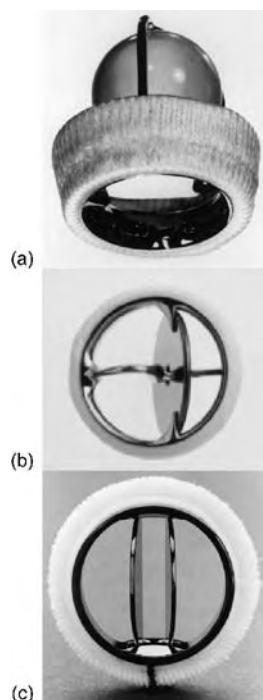


Fig. 1 (a) Starr-Edwards caged-ball valve, (b) Medtronic-Hall tilting disc valve, (c) St. Jude Medical bileaflet valve.

All bileaflet valves have a pivot-hinge mechanism so the mechanical loading is mainly applied to the pivot-hinge and housing ring. There have been some cases of leaflet escape and a cracked orifice ring, but the wear in the hinge area is not a problem. Because there is so little wear between pyrolytic carbon parts, the design of current bileaflet MHVs can ensure a lifetime of operation in patients. However, in reality there are still some incidences of structural failure. The hinge regions of bileaflet and tilting disc valves are vulnerable to localized thrombosis, which in turn may restrict occluder motion. In addition, the rigid circular orifice ring of a MHV is an unnatural configuration for a heart valve, which may lead to increased loading on the orifice ring or sewing ring.

MECHANICAL HEART VALVE WEAR AND FATIGUE FAILURE

Wear is loss of solid material mass as a result of external force on the material. It usually occurs in the contact area between solid materials. There are two types of wear mechanisms that occur on the housing and leaflets in the mechanical heart valve: impact wear and friction wear.^[5] Impact wear occurs between the pivot and hinge stop in the hinge regions of bileaflet valves, between the free edge of the leaflet or ball and the valve ring or cage as a result of the opening and closing action of the valve, and between the seat stop and disc in tilting disc valves. Friction wear occurs predominantly between the leaflet pivots and the bottom of the hinge cavities in the housing of bileaflet valves and between the disc and monostrut, or similar occluder retention mechanism, in tilting disc valves. Recently, cavitation damage that occurs inside or near the gap between the leaflet free edge and the housing at valve closing, has been identified as another wear mechanism that will be discussed in the next section. Impact, friction, and cavitation should be taken into account in the design of MHVs. In early valve designs, occluders made of polymeric materials demonstrated significant wear in clinical use, ultimately causing valve dysfunction. The poor wear performance of polymeric occluders greatly reduced the durability of early MHVs. The polymeric materials were soon replaced by pyrolytic carbon, which has excellent wear performance and biodegradability. Pyrolytic carbon was used to coat the occluder and housing ring surfaces to improve the durability of the MHV. It is estimated from wear tests that MHVs coated with pyrolytic carbon can operate for the lifetime of the patient—more than 50 years.

Metals are prone to fatigue failure. Their polycrystalline nature contains structural characteristics that may produce dislocations under mechanical loading. These dislocations can migrate when subjected

to repeated loading cycles and can accumulate at intercrystalline boundaries, ultimately leading to the formation of microfissures. These fissures grow slowly until they reach a critical size that can propagate and cause fracture. This fatigue failure typifies metal fracture under cyclic loading. There have been reports of fatigue failure of the valves that use titanium and chrome–cobalt alloy. The detailed reasons behind these failures remain undetermined, but it is evident that existing internal flaws, surface defects, and welds can increase the propensity for fatigue failure.

Pyrolytic carbon has a turbostatic structure that has no crystalline order between the atomic layers in the lattice, so mobile dislocation cannot exist. Accordingly, fatigue has not been a problem for pyrolytic carbon, although recent data suggests that cyclic fatigue-crack growth occurs in graphite/pyrolytic carbon composite material.^[6] This work suggests a fatigue threshold as low as 50% of the fracture toughness, so cyclic fatigue is viewed as an essential consideration in the design and life prediction of heart valves constructed from pyrolytic carbon. The FDA now requires detailed characterization of the pyrolytic carbon materials used in the various valve designs.

MECHANICAL HEART VALVE CAVITATION

Surface pitting and erosion observed on certain clinically explanted MHVs manufactured from pyrolytic carbon may be attributed to cavitation that occurs under normal physiological conditions in the human body. Substantial pitting gives rise to rough occluder surfaces, which may act as potential regions for thrombus formation and probable sites of crack growth in the valve components. Therefore, cavitation is another potential source of structural failure. Cavitation is the rapid formation and collapse of vapor-filled bubbles, usually as a result of a transient reduction in local pressure below liquid vapor pressure.^[7] The violent collapse of cavitation bubbles (also known as implosion) generates an intense shock wave and a high-speed re-entrant jet that can cause damage to the nearby solid boundary surface. The cavitation damage typified by pitting on the component surfaces may ultimately cause MHV failure.

There are three essential factors that must exist concurrently in order for MHV cavitation to be provoked: regional low pressure, nuclei, and sufficient time. Of these three factors, regional low pressure is key to the initiation of cavitation. It is usually associated with the closing motion of the MHV and always occurs on the inflow side of the mitral valve. There is substantial evidence demonstrating that MHV cavitation is related to jet flow through the gap between the occluder (disc or leaflet) and seat stop or valve housing ring, when the occluder impacts the housing ring.

Cavitation is closely linked to valve closing fluid dynamics, which in turn is related to valve design. There are two types of cavitation: vortex and transient bubble cavitation.^[8] Vortex cavitation is initiated by low pressure in the core of the vortex. Jets occur at valve closing, and the mixing layer of jet flow teems with vortices that initiate vortex cavitation. Jets can be subdivided into squeeze and transvalvular jets, both of which may initiate vortex cavitation. Squeeze jets occur between the occluder and the housing when the leaflets of tilting disc valves or bileaflet valves impact the seat stop or the wall of the housing ring. Transvalvular jets are generated by transvalvular pressure at valve closing. Sometimes the tension wave generated during valve closure is strong enough by itself to initiate cavitation typified by bubbles. This is called transient bubble cavitation. Both vortices and tension waves can initiate cavitation. They play a primary role in the inception of cavitation, but their overall contributions are not clearly understood. Investigation into the mechanisms of cavitation in MHVs is an active area of research.



APPLICATIONS OF FLUID MECHANICS

Blood flowing through prosthetic valves encounters a reduction in the cross-sectional area within the valve housing. This sudden reduction in flow area inevitably leads to a pressure drop across the valve. The viscosity of the fluid coupled with the turbulent effects causes an irreversible loss of pressure, which leads to a total loss in fluid mechanical energy. In general, caged ball valves have a greater pressure drop compared to other valve configurations. Porcine prosthetic valves also have smaller orifice areas compared to their mechanical valve counterparts, and hence generally produce a greater pressure drop. The valve configurations that produce the least amount of resistance to flow are the bileaflet and tilting disc designs. These two designs provide a relatively streamlined configuration that minimizes the impediment to forward blood flow.

The effective orifice area (EOA) is a measure of describing how well a valve design utilizes its primary or internal orifice area. A higher EOA corresponds to a smaller pressure drop and a smaller energy loss. The EOA of a particular valve can be calculated by measuring its pressure drop and flow rate:

$$\text{EOA}(\text{cm}^2) = \frac{Q_{\text{rms}}}{51.6(\Delta p)}$$

Q_{rms} is the root mean square systolic/diastolic flow rate (cm^3/s) and Δp is the mean systolic/diastolic pressure drop (mm Hg). For example, the EOA of a mitral valve can be calculated by measuring the diastolic flow

rate (Q_{rms}) and the mean diastolic pressure drop across the valve. Tables 1 and 2 list in-vitro EOA values for various valve types in the aortic and mitral positions, respectively.

Regurgitant flow is the sum of the reverse flow occurring during valve closure and the leakage flow after valve closure. The amount of regurgitant flow depends on the valve configuration and the closing dynamics of the leaflets. Almost all valve configurations have a certain amount of regurgitant flow. In the bileaflet mechanical valve, for example, leakage jets were observed originating in the hinge regions. With their soft and flexible leaflets, bioprosthetic heart valves have a smaller degree of regurgitant flow than their mechanical valve counterparts. An improperly designed or malfunctioning valve may display an increase in the regurgitant flow, so it is important that emphasis be

placed on the quantification of this retrograde flow. Tables 1 and 2 list regurgitation values for various valve types in the aortic and mitral positions, respectively.

IN-VITRO FLOW FIELD MEASUREMENTS WITH VARIOUS MECHANICAL VALVES

The flow field distribution of a Starr-Edwards 1260 aortic caged-ball valve design is characterized by an annular jet that flows over the silicone ball, impacts the aorta wall, and then follows the aorta farther downstream.^[9] Peak aortic values of approximately 2 m/s were measured and are comparable to those observed in other valve designs. The maximum peak forward velocity obtained 12 mm downstream of the valve was

Table 1 In-vitro hemodynamic data for common aortic valve prostheses^a

Valve type	Valve	Size	Regurgitant volume (ml/beat)	EOA ^b (cm ²)	PI	
Ball-and-Cage	Starr-Edwards 1260	27	5.5	1.75	0.30	
		25	4.3	1.62	0.33	
		21	2.5	1.23	0.36	
Tilting Disc	Björk-Shiley Convexo-concave	27	8.5	2.59	0.45	
		25	7.3	2.37	0.48	
		21	5.5	1.54	0.45	
		Björk-Shiley Monostrut	27	9.2	3.34	0.58
			23	6.9	2.00	0.48
			21	5.9	1.45	0.42
	Medtronic-Hall	19	5.5	1.07	0.38	
		27	9.6	3.64	0.64	
		25	8.4	3.07	0.62	
Bileaflet	St. Jude Medical Standard	23	7.3	2.26	0.54	
		20	6.2	1.74	0.51	
		27	10.8	4.09	0.71	
		23	8.3	2.24	0.54	
		21	6.8	1.81	0.52	
		19	6.8	1.21	0.43	
	St. Jude Medical Regent	29	13.5	4.98	0.75	
		23	10.3	3.47	0.83	
		21	9.0	2.81	0.81	
		17	6.3	1.56	0.69	
		CarboMedics	27	7.5	3.75	0.65
			23	6.51	2.28	0.55
21	3.4		1.66	0.48		
		19	3.0	1.12	0.40	

EOA = Effective Orifice Area; PI = Performance Index.

^aHR = 70 bpm; CO = 5 L/min typical; 120/80 mm Hg aortic pressure.

^bEOA computed from Eq. 8.

Table 2 In-vitro hemodynamic data for common mitral valve prostheses^a

Valve type	Valve	Size	Regurgitant volume (ml/beat)	EOA ^b (cm ²)	PI	
Ball-and-Cage	Starr-Edwards 1260	30		1.99	0.28	
		28		1.66	0.27	
		26		1.56	0.29	
Tilting Disc	Björk-Shiley Standard	31		3.40	0.45	
		29	6.7	2.79	0.42	
		25		2.15	0.44	
	Björk-Shiley Convexo-concave	29	10.1	2.96	0.45	
		27	8.3	2.28	0.40	
		25		2.03	0.41	
	Medtronic-Hall		31	10.0	3.53	0.47
			29	10.0	3.53	0.53
			27	8.9	2.73	0.47
25			7.2	2.23	0.45	
Bileaflet	St. Jude Medical Standard	31	13.1	3.67	0.48	
		29	10.9	3.40	0.51	
		27	9.7	2.81	0.49	

^aHR = 70 bpm; CO = 5 L/min typical.

^bEOA computed from Eq. 8.

2.2 m/s, decreasing to 1.8 m/s 30 mm farther downstream. A large wake was observed immediately downstream of the ball. This wake develops early during the acceleration phase and then grows into a region of reverse flow of 0.2 to 0.25 m/s that persists to the end of systole. High turbulent shear stresses found in regions of high-velocity gradients range from 1800 to 3500 dynes/cm² and are located at the edges of the jets. The ball seats nicely into the cage, thereby eliminating a large amount of regurgitant flow that comes mainly from the reverse flow during closure.

Two jets are typically observed in the Bjork-Shiley Monostrut and Medtronic-Hall tilting disc configurations, one from the major orifice and another from the minor orifice.^[9,10] The velocity magnitudes and turbulent shear stresses of the major orifice jet are typically larger than those of the minor orifice jet. The orientation and shape of these jets with respect to the streamwise direction changes with the position of the occluder disc as the valve opens and closes. At peak systole 7 mm downstream of the Bjork-Shiley Monostrut tilting disc valve, the velocities of the major orifice and minor orifice jets were measured at 2.1 m/s and 2 m/s, respectively.^[10] Regions of flow separation were observed between the jets and adjacent to the wall in the major orifice region. A stagnation region, which persists during the acceleration and deceleration phases, was observed adjacent to the wall in the minor orifice region. High turbulent shear stresses are located

in the jet shear layers and in the wakes of the disc and struts. The turbulent shear stress values are between 1200 and 3500 dynes/cm² and tend to be higher immediately distal to the valve.^[10] Leakage jets were observed to emanate from the perimeter of the valve through the gap between the occluder and valve housing.^[11] A study by Meyer et al.^[12] reported a large spike in the regurgitant flow during the early phase, which was attributed to the squeeze flow phenomena. The velocities and turbulent shear stresses measured were as high as 3.7 m/s and 10,000 dynes/cm², respectively.

The St. Jude Medical and CarboMedics bileaflet valve configurations divide the forward flow into three regions: one central orifice flow and two lateral flows. The primary differences between the bileaflet valves presently used are in the hinge designs, internal orifice diameter of the valve housing, the occluder opening angle, the occluder traveling arc, and the curvature of the leaflets. The hinge design determines the occluder opening angle and traveling arc and influences the washout, local flow dynamics, and flow patterns generated within the hinge.

Studies show that three jets characterize antegrade flow: one central jet and two lateral jets.^[9,10,13] Typically, the lateral jets have higher velocity magnitudes than the central jet. Taken 8 mm downstream of a SJM bileaflet valve and along the centerline of the valve, the velocities of the lateral jet and the central jet measured

during peak systole were 2.25 m/s and 1.95 m/s, respectively.^[13] Flow separation was observed adjacent to either side of the channel wall. Three jets were still present 13 mm downstream, but the central jet dissipated 20 to 30 mm downstream, whereas the lateral jets persisted for several centimeters more before merging with the surrounding, slower-moving fluid.^[13]

High turbulent shear stresses occur at regions of high-velocity gradients, such as at the edge of the lateral jets and in the region distal to the valve leaflets. Bileaflet designs exhibit lower turbulent shear stresses than the tilting disc and ball-and-cage designs. Peak aortic values were typically higher than those in the mitral valve because of the latter's lower transvalvular velocity. Leakage jets were observed exiting from the hinge region, between closed leaflets, and through the gap between the valve housing and the leaflets of the SJM and CarboMedics bileaflet valve designs.^[14,15] Two jets were observed emitting from each hinge and one from the gap between the leaflets after valve closure. Side jets from the valve periphery were occasionally observed, but they were not fixed and tended to change location with each cardiac cycle. The velocity and turbulent shear stress magnitudes of these leakage jets, as a result of the constricted flow regions, were typically at or above 2 m/s and 5000 dynes/cm², respectively.

A certain degree of regurgitant flow is usually incorporated into the design of a bileaflet MHV to ensure washout of critical areas of the valve, such as the hinges and regions between the leaflet edges and housing. This washout is intended to prevent flow stasis that may be created in the vicinity of the hinge regions and upstream of the valve housing during valve closure, in order to minimize or eliminate the formation of thrombosis, which still poses a significant problem.^[16] Recent studies have shown the importance of the bileaflet MHV hinge geometry in influencing the characteristics of leakage flow structures, such as the leakage velocities and turbulent shear stresses, within and near the hinge region.^[14,15] The constricted flow area at the hinge region of these MHVs, coupled with the pressure difference across the bileaflet valves during closure, is a location in which the fluid can be subjected to high turbulent shear stresses. These studies have suggested the hinge mechanism as the most critical part of the bileaflet MHV because it directly influences valve durability, functionality, and thrombus formation. SJM and CarboMedics prosthetics models have similar hinge geometries, except that in the former the hinge mechanism is streamlined and has a curved profile, and in the latter the hinge geometry has sharper corners and less streamlined edges. The peak velocities observed during the leakage phase in the SJM Regent valve and the CarboMedics CPHV models were 1.5 m/s and 3.17 m/s, respectively.^[14,15]

BLOOD–MHV INTERACTIONS

Shear stress can lead to thromboemboli formation by inducing blood cell damage, thus exposing tissue factor to the plasma, and/or by activating the platelets directly. The shear stress magnitudes shown to cause damage to both red blood cells (RBCs) and platelets have been quantified. RBCs are vulnerable to lethal damage at shear stresses of 1500 to 4000 dynes/cm², and by shear stresses as small as 10 to 100 dynes/cm² in the presence of foreign surfaces. Sublethal damage occurs at 500 dynes/cm².^[16] Platelets have been shown to undergo damage after exposure to 100 to 500 dynes/cm².^[16] Platelet damage is also dependent upon the duration of the shear stress. Platelet damage increases linearly with time during exposure to a constant shear stress implying that shear-induced platelet damage is cumulative.^[16] Shear-induced platelet activation has been studied in viscometers and is also a function of magnitude and duration.^[16] At 150 dynes/cm², platelet aggregates are formed after 300 seconds. Aggregates form at 600 dynes/cm² after 30 seconds, and at 6500 dynes/cm² in less than 5 milliseconds.^[16] These studies also show that as the magnitude increases, formed aggregates tend not to separate when the force is discontinued.^[16] This behavior points to a different mechanism for platelet activation at higher levels of shear stress.

Clot initiation caused by the presence of MHVs can occur via three basic pathways: 1) phospholipid and/or tissue factor exposure to blood during cell damage caused by excessive shear stress; 2) platelet activation caused by exposure to excessive shear stress; and 3) contact activation by foreign materials. Clotting occurs in three steps: initiation, amplification, and propagation. In the tissue factor pathway, a cell or cells are ruptured during the initiation phase to expose tissue factor (TF) to the blood. In the plasma Factor (f) VII binds to TF, thus activating it to fVIIa. This complex activates fX and fIX, and fXa can activate fV. On the cell surface, fXa can bind with fVa to produce small amounts of thrombin, which can activate platelets and fVIII during amplification.^[17] During amplification, platelets adhere to the damaged tissue and are activated. The thrombin generated earlier enhances platelet adhesion, thus fully activating platelets and also activating fV, fVIII, and fXI.^[17] Large-scale thrombin generation takes place during the propagation phase, leading to the formation of fibrin that enmeshes the platelets and stabilizes the clot.^[17]

At levels above 60–80 dynes/cm², stable platelet aggregation occurs in a cone and plate viscometer.^[18] The exact mechanism behind this pathway is not as clearly defined as the TF pathway, but the basic steps are as follows. During initiation, the binding of vWF in the plasma to GPIb on the platelet membrane leads to

a transmembrane flux of Ca^{2+} .^[18] This influx of Ca^{2+} activates the platelet and leads to the amplification phase, during which GPIIb–IIIa interacts with adhesive proteins to mediate platelet–platelet interaction and subsequent aggregation.^[18] GPIIb–IIIa usually binds to fibrinogen, but under high-shear conditions it binds to vWF to mediate platelet aggregation.^[18] Thus, this pathway may not be dependent upon fibrinogen, because vWF both activates the platelets and supports platelet cohesion. The mechanism by which propagation occurs is not clear.

Contact-activated clotting is initiated by the binding of fXII to a negatively charged surface, thus forming fXIIa.^[19] The proteins prekallikrein (PK), fXI, and high-molecular-weight kininogen (HK) are then activated by fXIIa, and PK and fXIa bind to the foreign surface through HK.^[19] On the surface, fXIIa binds PK to kallikrein and fXI to fXIa, and kallikrein generates additional fXIIa. Kallikrein also hydrolyzes HK into its activated forms and releases bradykinin.^[19] Thus, HKa–PK and HKa–fXI complexes are formed on the surface near fXIIa. During amplification, fXIIa acts on these complexes to form surface-bound Hka–kallikrein and Hka–fXIa complexes.^[19] The Hka–fXIa complex activates fIX to fIXa, which acts in a cascade to form thrombin, which in turn activates platelets. Kallikrein is released to activate more fXIIa and convert existing fXIIa to fXIIa, which is still activated but is no longer bound to the surface, so that it can activate more PK.^[19] In addition, fibrinogen is adsorbed to the material surface. A buildup of a layer of proteins on the material helps mediate platelet adhesion and subsequent tissue growth.^[19]

When a valve is implanted, a layer of thrombus usually forms on the sewing cuff. Over time the thrombus either may organize into avascular fibrous tissue or may evolve into granulation tissue and growth of capillaries to form vascular tissue.^[20] This tissue eventually becomes a layer of endothelium of variable completeness and uncertain functionality, which may or may not exhibit anticoagulant activity and may even show procoagulant activity.^[20] The extent of tissue growth, as well as the type, depends upon the location of the valve. Lower pressures lead to greater thickness and more extensive vascularization.^[20] In animal studies, early anticoagulant therapy has been correlated with a reduced tissue thickness. If tissue growth is excessive, it may overgrow the margin of the sewing ring and interfere with valve function.^[20] Some factors leading to this overgrowth (pannus) formation are fabric damage, low-flow conditions, inadequate or inconsistent anticoagulation therapy, the degree of foreign body interaction, and fibroblast activity.^[20] Over time, extension of this overgrowth onto the MHV mechanism can cause narrowing of the orifice, interfere with occluder motion, or act as a site for emboli formation. Reducing the flow

through the orifice results in increased shear stress through the valve and increased areas of stagnation downstream of the valve, predisposing the valve to gross thrombosis. The worst problems are seen in the tricuspid position, most likely because the pressures are lowest here.^[20] Contact activation of platelets by pyrolytic carbon has also been documented, although the importance of this activation in thromboemboli formation is not known.^[21]

In-vitro blood studies with both porcine and human blood have correlated blood damage with both forward flow and leakage flow through MHVs. In a study with porcine blood using a Bjork-Shiley monostrut disc valve, hemolysis was shown to increase with time during both forward flow and leakage flow for the mitral and aortic positions.^[22] Furthermore, in comparisons of hemolysis between the aortic and mitral positions, the mitral position was shown to be more damaging than the aortic position during forward flow, whereas the aortic position was more damaging during leakage flow.^[22] This difference can be attributed to higher turbulent stresses during backflow in the aortic position than in the mitral position, due to a higher transvalvular pressure.

Cavitation was also studied, and it was found that when it occurred its contribution increased with the duration of the cavitation. If the duration was long enough, cavitation became the dominant hemolysis mechanism.^[22] Hemolysis contributions during forward flow and leakage flow were about the same, but considering that leakage affects only 2–5% of the flow, this contribution is significant.^[22] Leakage flow studies on the Medtronic Parallel valve and the SJM standard, high-leaker, and low-leaker prototype valves, using markers for platelet activation and blood element damage in human blood, showed that platelets were activated and damaged in a time-dependent fashion, with activation and damage increasing with time.^[23] Additionally, the high-leaker prototype valve showed more platelet activation and blood damage than the SJM standard valve.^[23] This result shows the effect of leakage magnitude on blood damage and platelet activation. Although a larger leakage area results in lower shear stresses during the leakage phase, a higher leakage flow through the valve exposes more of the blood to high shear stresses encountered during leakage flow. Thus, there is an optimal leakage area that minimizes blood damage.

It has been shown conclusively that elevated shear stress can activate platelets. However, transiently activated platelets would not lead to thrombus or emboli formation if they were not afforded an opportunity to adhere to a surface and/or to one another. In order for a thrombus or emboli to be of sufficient pathological size, millions of platelets must be activated and form an aggregate. Studies using a 10-micron filter to retain platelet aggregates while applying a predetermined pressure leading to elevated shear stress have shown

that activation and aggregation can occur with 8 milliseconds.^[24] Thus, activated platelets that are brought together could aggregate in a very short time period.

In a model of a stenosis with a subsequent enlargement, radiolabeled platelets were used to determine the locations of highest adhered platelet density at different physiological flow rates.^[25] In this model, the stenosis generates a region of high shear stress in which platelets are activated, and the enlargement induces flow separation and consequent recirculation regions.^[25] As expected, this study showed that net platelet density was elevated in areas of flow recirculation and reattachment.^[25] Furthermore, the extent of the platelet density correlated with the length of the recirculation region.^[25] Thus, under the effect of shear stress alone, emboli are most likely to develop in recirculation areas, and they may or may not attach in areas of stagnant flow to become thrombi. Similar recirculation regions are observed in the immediate vicinity of MHVs in the aortic and mitral positions. These regions of recirculation allow platelets to aggregate, with the primary adhesion occurring downstream. Under the influence of tissue overgrowth or an exposed sewing ring, platelets may also attach at these locations as a result of contact activation or shear stress activation and subsequent adhesion.

Studies focusing on the fluid mechanics properties of the Starr-Edwards ball valve, the Bjork-Shiley tilting disc valve, the Medtronic-Hall tilting disc valve, the SJM bileaflet valve, and the CarboMedics bileaflet valve have correlated areas of high shear stress, stagnation, and flow separation to thrombus formation.^[26] The Starr-Edwards ball valves are vulnerable to high-wall shear stresses that can lead to endothelial cell damage. Additionally, high turbulent shear stresses during forward flow, coupled with high-velocity reverse flow relative to the surrounding stagnant flow, create regions of separation that can lead to platelet aggregation.^[26] Explanted valves have demonstrated thrombus formation in the cage area, lending credence to the thrombogenicity of this type of flow field combination.^[26]

The designs of the Bjork-Shiley and Medtronic-Hall tilting disc valves lead to a wake formation behind the valve struts and disc, and these regions exhibit flow separation due to a combination of high velocity and nearly stagnant flows.^[26] Valve explants have demonstrated thrombus formation in this stagnation region on the underside of the disc. Bileaflet valves are also vulnerable to high turbulent shear stresses during both forward and leakage flows. They also exhibit asymmetric separated regions with flow reversal around the perimeter of the housing, with larger spatial dimensions near the hinges adjacent to the lateral jets as well as high regurgitant flow volumes that lead to high shear stresses.^[26] Coupled with adjacent stagnant flow, these high shear stresses could lead to thrombus

formation in the hinge area.^[26] The Medtronic Parallel valve demonstrated such thrombus formation in clinical trials. It passed all preclinical trials, including animal blood studies, in forward flow. However, once this valve was implanted in clinical trials, many patients experienced thromboembolic complications, and valve explants showed thrombus deposition in the inflow region of the hinge area.^[26] The St. Jude and CarboMedics bileaflet valves have also demonstrated thrombus formation in the hinge area.^[20,27]

Because MHVs have been implanted successfully since the 1960s, much patient data is available detailing the success or failure rates of many valve types. According to a collection of such data, it can be seen that valvular thrombosis is most common in the mitral position.^[28] Additionally, bileaflet valves and single disc valves have about the same risk for thromboembolism and valve thrombosis. The ball valve appears to have a slightly lower risk for valve thrombosis, and thrombosis occurs gradually in this valve.^[28] One advantage of the bileaflet valve over the single disc valve may be that one leaflet may function even if the other does not. However, both valves can unexpectedly cease functioning if the hinge is obstructed.^[28]

FUTURE DIRECTIONS

Future emphasis in prosthetic valve research is to improve valve design in terms of better hemodynamic performance, such as lower peak forward and regurgitant flows with reduced turbulent shear stresses and velocity gradients. Improving hemodynamic performance would also reduce blood damage and platelet activation caused by elevated turbulent shear stresses. Progress must also be made toward developing improved materials that are not only durable and biocompatible, but that also enable valve designs to better approximate native valve function. These improvements must be intimately coupled with progress in computational, in-vitro, and animal testing methodologies.

ARTICLES OF FURTHER INTEREST

Heart Valve Failure, Bioprosthetic; Heart Valve, Bioprosthetic; Pyrolytic Carbon; Tissue Engineering of Heart Valves

REFERENCES

1. Hufnagel, C.A.; Gillespie, J.F.; Conrad, P.W.; Mercier, C.; Evangelist, F.A. Cardiac valve replacement—Its current status. *Med. Ann. D. C.* **1966**, *35* (9), 457–462.

2. Starr, A.; Fessler, C.L.; Grunkemeier, G.; He, G.W. Heart valve replacement surgery: Past, present and future. *Clin. Exp. Pharmacol. Physiol.* **2002**, *29* (8), 735–738.
3. Ely, J.L.; Emken, M.R.; Accuntius, J.A.; Wilde, D.S.; Haubold, A.D.; More, R.B.; Bokros, J.C. Pure pyrolytic carbon: Preparation and properties of a new material, On-X carbon for mechanical heart valve prostheses. *J. Heart Valve Dis.* **1998**, *7* (6), 626–632.
4. Bokros, J.C.; Haubold, A.D.; Akins, R.J.; Campbell, L.A.; Griffin, C.D.; Lane, E. The Durability of Mechanical Heart Valve Replacements: Past Experience and Current Trends. In *Replacement Cardiac Valves*; Bodnar, E., Frater, R.W.M., Eds.; Pergamon Press Inc.: New York, 1991; 21–48.
5. Hasenkam, J.M.; Pasquino, E.; Stacchino, C.; Arru, P.; Vallana, F.; Paulsen, P.K. Wear patterns in the Sorin Bicarbon mechanical heart valve: A clinical explant study. *J. Heart Valve Dis.* **1997**, *6* (2), 105–114.
6. Ritchie, R.O.; Dauskart, R.H.; Yu, W. Cyclic fatigue-crack propagation, stress-corrosion, and fracture-toughness in pyrolytic carbon-coated graphite for prosthetic heart valve applications. *J. Biomed. Mater. Res.* **1990**, *24*, 189.
7. Knapp, R.J.; Daily, J.W.; Hammitt, F.G. *Cavitation*; McGraw Hill Book Co.: New York, 1970; 1–2.
8. He, Z.M.; Xi, B.S.; Zhu, K.Q.; Hwang, N.H.C. Mechanism of mechanical heart valve cavitation: Investigation using tilting disk valve model. *J. Heart Valve Dis.* **2001**, *10*, 666–674.
9. Yoganathan, A.P.; Woo, Y.R.; Sung, H.W.; Jones, M. Advances in prosthetic heart valves: Fluid mechanics of aortic valve designs. *J. Biomater. Appl.* **1988**, *2* (4), 579–614.
10. Yoganathan, A.P.; Sung, H.W.; Woo, Y.R.; Jones, M. In vitro velocity and turbulence measurements in the vicinity of three new mechanical aortic heart valve prostheses: Bjork-Shiley monostrut, omni-carbon, and duromedics. *J. Thorac. Cardiovasc. Surg.* **1988**, *95* (5), 929–939.
11. Kini, V.; Bachmann, C.; Fontaine, A.; Deutsch, S.; Tarbell, J.M. Integrating particle image velocimetry and laser Doppler velocimetry measurements of the regurgitant flow field past mechanical heart valves. *Artif. Organs* **2001**, *25* (2), 136–145.
12. Meyer, R.S.; Deutsch, S.; Bachmann, C.B.; Tarbell, J.M. Laser Doppler velocimetry and flow visualization studies in the regurgitant leakage flow region of three mechanical mitral valves. *Artif. Organs* **2001**, *25* (4), 292–299.
13. Woo, Y.R.; Yoganathan, A.P. Pulsatile flow velocity and shear stress measurements on the St. Jude bileaflet valve prosthesis. *Scand. J. Thorac. Cardiovasc. Surg.* **1986**, *20* (1), 15–28.
14. Leo, H.L.; He, Z.; Ellis, J.T.; Yoganathan, A.P. Microflow fields in the hinge region of the CarboMedics bileaflet mechanical heart valve design. *J. Thorac. Cardiovasc. Surg.* **2002**, *124* (3), 561–574.
15. Ellis, J.T.; Travis, B.R.; Yoganathan, A.P. An in vitro study of the hinge and near-field forward flow dynamics of the St. Jude Medical Regent bileaflet mechanical heart valve. *Ann. Biomed. Eng.* **2000**, *28* (5), 524–532.
16. Yoganathan, A.P.; Wick, T.M.; Reul, H. Influence of Flow Characteristics of Prosthetic Valves on Thrombus Formation. In *Thrombosis, Embolism and Bleeding*; Butchart, E.G., Bodnar, E., Eds.; ICR Publishers: United Kingdom, 1992; 123–148.
17. Hoffman, M.; Monroe, D.M., III. A cell-based model of hemostasis. *Thromb. Haemost.* **2001**, *85*, 958–965.
18. Ruggeri, Z.M. Mechanisms of shear-induced platelet adhesion and aggregation. *Thromb. Haemost.* **1993**, *70* (1), 119–123.
19. DeLa Cadena, R.A.; Wachtfogel, Y.T.; Colman, R.W. Contact Activation Pathway: Inflammation and Coagulation. In *Hemostasis and Thrombosis: Basic Principles and Clinical Practice*, 3rd Ed.; Colman, R.W., Hirsh, J., Marder, V.J., Salzman, E.W., Eds.; J. B. Lippincott Company: Philadelphia, 1994; 219–240.
20. Butchart, E.G. Thrombogenicity, Thrombosis and Embolism. In *Thrombosis, Embolism and Bleeding*; Butchart, E.G., Bodnar, E., Eds.; ICR Publishers: United Kingdom, 1992; 172–205.
21. Goodman, S.L. Sheep, pig, and human platelet–material interactions with model cardiovascular biomaterials. *J. Biomed. Mater. Res.* **1999**, *45*, 240–250.
22. Lamson, T.C.; Rosenberg, G.; Geselowitz, D.B.; Deutsch, S.; Stinebring, D.R.; Frangos, J.A.; Tarbell, J.M. Relative blood damage in the three phases of a prosthetic heart valve flow cycle. *ASAIO J.* **1993**, *39*, M626–M633.
23. Travis, B.R. The Effects of Bileaflet Prosthesis Pivot Geometry on Turbulence and Blood Damage Potential. In *Ph.D. Thesis*; Georgia Institute of Technology, 2001.
24. O'Brien, J.R. Shear-induced platelet aggregation. *Lancet* **1990**, *335*, 711–713.
25. Schoepfoerster, R.T.; Oynes, F.; Nunez, G.; Kapadvanjalwa, M.; Dewanjee, M.K. Effects of local geometry and fluid dynamics on regional platelet deposition on artificial surfaces. *Arterioscler. Thromb.* **1993**, *13*, 1806–1813.
26. Yoganathan, A.P.; Travis, B.R. Fluid Dynamics of Prosthetic Valves. In *The Practice of Clinical Echocardiography*, 2nd Ed.; Otto, C.M., Ed.; , 2000.
27. Chang, B.-C.; Lim, S.H.; Kim, D.K.; Seo, J.Y.; Cho, S.Y.; Shim, W.H.; Chung, N.; Kim, S.S.; Cho, B.K. Long-term results with St. Jude medical and carbomedics prosthetic heart valves. *J. Heart Valve Dis.* **2001**, *10*, 185–195.
28. Grunkemeier, G.L.; Li, H.-H.; Naftel, D.C.; Starr, A.; Rahimtoola, S.H. Long term performance of heart valve prostheses. *Curr. Probl. Cardiol.* **2000**, *25* (2), 73–156.



Hematopoietic Stem Cells and Assays

Athanassios Mantalaris

Department of Chemical Engineering & Chemical Technology, Imperial College London, London, United Kingdom

Nicki Panoskaltis

Department of Haematology, Imperial College London, London, United Kingdom

J. H. David Wu

Departments of Chemical Engineering, Biomedical Engineering, and Microbiology & Immunology, University of Rochester, Rochester, New York, U.S.A.

INTRODUCTION

Lineage fidelity of a stem cell may not be as strictly governed as was once thought. This lineage plasticity is most easily studied with hematopoietic stem cells because, even in humans, they are readily obtainable from blood and bone marrow, and the assays to define them are well-characterized, having been the subject of studies for over 30 years. Because of the ease with which they can be obtained from humans, hematopoietic stem cells (HSCs) are an ideal source of cells, which, if truly plastic, could be exploited for the therapeutic repair of damaged and diseased tissues of any type. However, to harness this potential, these cells must be easily identified and isolated from different sources, effectively expanded *ex vivo*, and evaluated functionally by assays that test the desired potential. In order to study and eventually utilize HSCs for the treatment of human diseases, these cells must be expanded *ex vivo* without change in the identifiable and functional phenotype. The largest impediment to the successful development of these expansion methods has been the difficulty in accurate identification and enumeration of HSCs.

HEMATOPOIETIC STEM/PROGENITOR CELL ASSAYS

The “gold standard” for assaying the HSC is the ability of a single cell to reconstitute long-term hematopoiesis in a myeloablated host.^[1] A variety of *in vitro* and *in vivo* assays have also been developed to identify stem cells (Table 1). Phenotypic characterization using flow cytometry is useful not only for the initial isolation of cells that are enriched for HSCs, but also because it has the potential to allow the rapid, relatively inexpensive determination of culture status.^[2] It is widely accepted that the cell-surface glycoprotein CD34 is expressed on almost all human cells with

discernable hematopoietic potential.^[3] Recently, additional markers have been used to fractionate the CD34⁺ cells further and to identify more primitive stem cells.^[4] In particular, cells expressing CD34⁺/CD38⁻ or CD34⁺/Thy-1⁺ phenotype are highly enriched for very primitive hematopoietic cells.^[5,6] However, the identification of CD34⁻ cells with hematopoietic and reconstituting potential has also been reported,^[7] and these cells may be able to lose or gain the expression of CD34 depending on the cell cycle.^[8,9] The combination of these cell-surface markers with fluorescent dyes that indicate cell proliferative status has become a method of choice for isolating HSCs. Hoechst-33342 staining has been used to define a small subset of murine and human bone marrow (BM) cells (Side Population or SP), which are highly enriched for long-term reconstituting HSCs,^[10] and the mitochondrial dye rhodamine-123 has been used to further subdivide primitive stem cells.^[11] Unfortunately, the relationship between the HSC phenotype and its function or genotype is based on uncultured cells; this relationship rapidly dissociates in *in vitro* culture environments^[12] making the dependability of these phenotypic markers as a predictor of biological potential doubtful.^[4] More recently, the expression of endoglin, an ancillary transforming growth factor β receptor, has been identified on murine stem cells and may be used as a functional marker.^[13] In human mobilized peripheral blood, (MPB) side scatter low cells that expressed high levels of aldehyde dehydrogenase (SSC^{lo}ALDH^{hi}) generated colony-forming units (CFUs) and multiple cell lineages both before and after long-term culture and correlated positively with time to engraftment in humans.^[14] However, reconstitution of a myeloablated host with these SSC^{lo}ALDH^{hi} cells alone would be required prior to the use of ALDH as a human HSC marker.

In vitro functional assays have been developed to allow specific lineage-restricted and multi-, bi-, and monopotent subpopulations of hematopoietic

Table 1 HSC^a and HPC^b assays

Assay	Species	Target cell population/comments
<i>Phenotypic markers</i> (Flow cytometry)		
• CD34 ⁺ , Thy-1 ⁺ , Lin ⁻ , c-kit ^{lo} , CD38 ⁻	Human	Subset of HSCs; questionable dependability of phenotypic markers as predictor of biological activity
• Lin ⁻ , Sca-1 ^{hi} , Thy-1.1 ^{lo} , c-kit ⁺	Murine	Same as above
• Side population (SP) defined by Hoechst-33342 staining	Human/murine	Enriched HSCs
<i>In vitro</i> (Long-term culture)		
• Colony-forming assay (CFU, BFU, HPP-CFC)	Human/murine	HPC
• LTC-IC	Human/murine	Primitive HPCs; do not enumerate HSCs
<i>In vivo</i> (Animal models)		
• CFU-S	Murine	Primitive HPC and subset of HSCs (depending on the assay endpoint) forming colonies on the host spleen upon transplantation; exhibition of highly variable self-renewal and differentiation of even the most phenotypically homogeneous CFU-S populations
• Long-term repopulation assay (LTRA) (Syngeneic/xenogeneic transplantation of HSC into SCID, NOD/SCID mice, or sheep, etc.)	Human/murine	HSC; gold standard of HSC assay; xenogeneic microenvironment may produce a less conducive environment for homing, growth, and differentiation of human HSCs

^aHematopoietic stem cell.

^bHematopoietic progenitor cell.

progenitor cells to be detected and quantified. These methods require the provision of appropriate soluble, cell, or matrix-bound factors essential to the support and stimulation of the progenitors being assessed. The CFU, burst-forming unit (BFU), and high-proliferative potential colony-forming cell (HPP-CFC) assays enumerate committed hematopoietic progenitor cells, which form colonies of differentiated cells on semisolid culture media.^[15,16] These progenitors are therefore operationally defined as CFUs with the suffix indicating their lineage commitment. For example, CFU-GM is a progenitor committed to the granulocytic and macrophage lineages. Although some progenitors with multi-lineage differentiation potential can be detected, these progenitors display only limited self-renewal capacity *in vitro*.

An *in vitro* assay that has shown greater specificity for quantifying hematopoietic cells that appear to be more primitive than most colony forming cells (CFCs) is the long-term culture-initiating cell (LTC-IC) assay in which cells are co-cultured with hematopoietic supportive stromal feeders, with or without exogenous cytokines.^[5] The LTC-IC assay fulfils the HSC criterion of being able to initiate and sustain myelopoiesis, but not the two other criteria of self-renewal and long-term lymphopoiesis.^[17]

The colony-forming unit-spleen (CFU-S) assay is an *in vivo* clonal assay in which bone marrow or spleen cells are transplanted into lethally irradiated recipients whose spleens are harvested and analyzed after 8 (CFU-S₈) and 12 (CFU-S₁₂) days.^[18] The colonies of hematopoietic cells appearing on the spleen surface represent different populations of cells where CFU-S₈ cells are predominantly unipotential and CFU-S₁₂ cells are multipotential. Although transplanted HSCs do not generate CFU-S₈, at least a subset of HSCs produce CFU-S₁₂.

The most stringent assay for measuring HSC activity is the long-term repopulating assay (LTRA) because it requires donor cells to fulfill the criteria that define HSCs—the ability of self-renewal, differentiation, and reconstitution of hematopoiesis in the transplant host. In the murine system, syngeneic mice are used as donors and hosts to avoid rejection. In xenogeneic transplants, human HSCs are transplanted in immune-incompetent animals^[19–21] to demonstrate not only multilineal differentiation but also self-renewal and repopulating ability. However, since human HSCs need to repopulate a xenogeneic microenvironment that may be less conducive to HSC homing, growth, and differentiation than the human microenvironment, these assays may not enumerate all human HSCs.^[17]



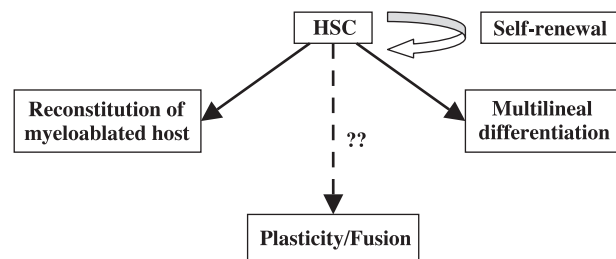


Fig. 1 Definition of hematopoietic stem cells (HSCs).

Recently, the use of hematopoietic stem cells to reconstitute liver, heart, brain, and muscle in animals,^[22–25] as well as in humans receiving these cells for bone marrow reconstitution,^[26] has raised the possibility of plasticity of cells with a pluripotent rather than a defined multipotent capacity (reviewed by A. Liebmann-Vinson and S. C. Presnell in the article titled *Stem Cells* elsewhere in this encyclopedia). This apparent plasticity is controversial and may simply be the result of cell fusion in the setting of tissue damage^[27–29] or of opportunistic changes in the cell cycle in the context of empty hematopoietic niches.^[30,31] In either case, the definition of the hematopoietic stem cell may be changing to not only a cell that is capable of self-renewal, differentiation, and reconstitution but also one that can either reconstitute or fuse with cells of other tissues as well (Fig. 1). This cell fusion can occur both in the presence or in the absence of tissue damage^[26,32] and may be affected by progeny of HSC such as macrophages since, at least in a murine knockout model for the liver enzyme fumarylacetoacetate hydrolase, replacement of the enzyme by the host HSCs took place only after engraftment had occurred.^[33–35]

CELL SOURCE

From a clinical and bioprocess aspect, the two major issues for ex vivo expansion are the source of the hematopoietic cells and the extent of purification. Several sources of hematopoietic stem and progenitor cells are used for clinical therapies including BM, MPB, and umbilical cord which are collected by various procedures from the patient or from human leukocyte antigen-(HLA-)compatible donors.^[33,36,37] There are significant differences in the quantity and quality (in terms of their proliferative and engraftment abilities and responses to hematopoietic growth factors) of the hematopoietic samples collected from these sources, which also serve as the starting material for ex vivo expansion, although there is currently no consensus as to the best starting population.^[38] BM and MPB are considered to be the best sources of stem and progenitor cells for clinical HSC transplantation. Over the last 10 years, however, umbilical cord blood (CB) has been investigated as an alternative source of

HSCs for the allogeneic transplantation to patients lacking an HLA-matched marrow donor.^[39] CB may be an ideal universal stem cell source because it can be collected easily and provides a potential source of HSCs to patient groups underrepresented in transplant registries. Compared with recipients of adult stem cell sources, those who receive CB may be more tolerant of HLA-mismatch, because they exhibit less graft-versus-host disease (GvHD), a benefit that may be attributed to the increased number of immature dendritic cells and naïve T cells contained within the graft.^[40,41] CB also has lower levels of contamination with common viral pathogens.^[38] However, even though CB is rich in stem and progenitor cells, the number of stem cells obtained is low due to the small volume collected (~100 ml). Hence, the vast majority of CB recipients have been children with an average weight of 20 kg. A nucleated cell dose of $1.5 \times 10^7/\text{kg}$ (or $1.7 \times 10^5 \text{ CD34}^+$ cells/kg) in the CB HSC graft defines a threshold below which an inferior clinical outcome may be achieved.^[42] Hence, the widespread use of cord blood HSCs clinically will depend upon the development of methods for the successful expansion of HSCs ex vivo without compromising engraftment, differentiation capability, and maintenance of the stem cell pool.

CB may be more responsive to growth factors and may be capable of greater expansion than the BM or MPB HSCs, but it is unclear whether the regenerative stem cell pool can be equally maintained. In vitro expansion protocols using the cytokines stem cell factor (SCF), granulocyte colony stimulating factor (G-CSF), and megakaryocyte growth and development factor (MGDF) in one- and two-step culture systems have been successful at expanding committed and primitive progenitor cells, but not at expanding and renewing the HSC population.^[43,44] A recent Phase 1 clinical trial augmented CB HSC transplantation in 28 patients with CB cells expanded in vitro and found no significant improvement in engraftment but did find a reduction in the number of $\text{CD34}^+/\text{lin}^-$ cells available after in vitro culture, suggesting failure of stem cell expansion in this system.^[45] Optimization of ex vivo expanded CB products with the goal of enhancing HSC self-renewal and expansion would provide an ideal stem cell for not only HSC transplantation protocols, but also other in vivo protocols for repair of non-hematopoietic tissues. Addition of proteins, such as Wnt,^[46] or improvement of the scaffold in the culture environment may provide the signals required for HSC expansion.

STEM CELL ENRICHMENT

The ideal phenotype or purity of the starting cell population(s) for the inoculation of ex vivo expansion cultures has not yet been characterized.

Unfractionated hematopoietic cells derived from both adult and umbilical cord sources contain mature blood cells and are developmentally heterogeneous. For clinical use, different methods have been used to enrich for hematopoietic stem and precursor cells.^[37] Some clinical trials have used only CD34-enriched allogeneic grafts with the advantage of harboring fewer T cells and therefore less chance of the patient dying from GvHD, but with the disadvantage of delaying engraftment resulting in prolonged neutropenia and an increased risk of dying from infection.

CD34⁺ cell fractions from different sources may possess different proliferation and expansion potentials.^[47,48] Although the number of CD34⁺ cells may be less in CB compared with adult sources, Kim et al.^[48] indicated that CD34⁺ cell fractions derived from CB may possess higher numbers of CFCs with a greater proliferative capacity. In comparison with grafts selected for CD34⁺ cells, transplantation of whole BM mononuclear cells (MNCs) can support rapid neutrophil and platelet engraftment and long-term stem cell engraftment, presumably due to the presence of accessory cells, such as stromal cells, which also enhance ex vivo expansion cultures.^[49,50] Despite the 10-fold expansion of CFU-GM in these studies using BM MNCs and stromal cell feeders, the earlier progenitor/stem cell pool (represented by LTC-IC and CD34⁺/Lin⁻ cells) was reduced to less than half of the input amount.^[49,50] In contrast, a bioreactor culture will support the same number of total cells and CFU-GMs and will have the same volumetric culture requirements for clinical-scale ex vivo expansion regardless of the extent of purification of the input cells.^[51,52]

MESENCHYMAL STEM CELLS

BM stroma provides structural and instructional support for hematopoiesis. Considerable progress has

been made in the identification and characterization of precursors for the stromal tissue of BM, although this very small (0.01–0.0001% of the total nucleated cells) and heterogeneous population still lacks conclusive morphologic and molecular characterization.^[53,54] Nevertheless, evidence for a developmental hierarchy of stromal cells within the BM is strong, pointing to the existence of multipotent mesenchymal stem cells (MSCs), capable of forming a spectrum of specialized mesenchymal and a variety of connective tissues (Fig. 2).^[54–56] In addition to BM, MSCs have been identified in CB^[57] as well as in peripheral blood,^[58] thus widening the potential availability of MSCs for clinical use if these cells can be positively identified and enriched from these sources.

In vivo analysis of MSCs has been augmented by the development of in vitro functional assays, (colony-forming unit-fibroblast; CFU-F), which demonstrate the ability of clonogenic progenitor stromal cells to form a variety of differentiated cell types.^[59,60] Despite the fact that antibodies to several cell surface antigens that recognize MSCs exist, there is a lack of monospecific and unique molecular probes to identify these cells unequivocally in situ. The monoclonal antibody STRO-1, which identifies an as yet uncharacterized cell-surface antigen expressed by a small heterogeneous population of adult BM MNCs, appears to be a suitable candidate.^[61] Further research is clearly required before we can fully exploit the biological potential of MSCs for therapeutic purposes. However, these cells are already in use therapeutically in clinical trials, including for the replacement of marrow stromal tissue damaged by chemotherapy as part of cancer treatment,^[62] as vesicles for gene therapy in the treatment of mesenchymal tissues,^[63] and for the ex vivo expansion for autologous repair of bone and cartilage defects as well as osteoporosis.^[54,64] The capability of these cells to inhibit ongoing immune responses, such as GvHD and graft rejection is also being exploited in stem cell and organ transplantation protocols.

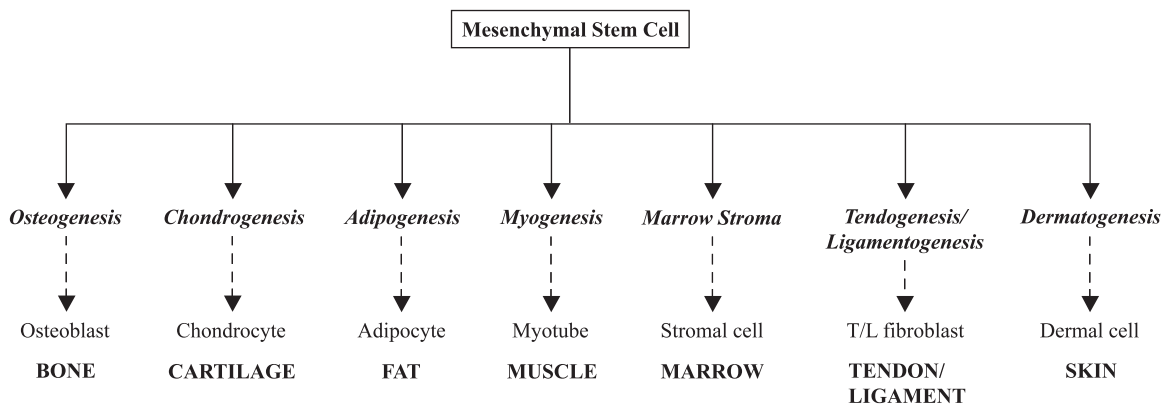


Fig. 2 Mesenchymal differentiation pathway. The differentiation process involves stepwise cellular transitions from the mesenchymal stem cell (MSC) to highly differentiated end-stage phenotypes.



CONCLUSIONS

The immense therapeutic potential of stem cells requires the existence of assays for the accurate identification, isolation, and evaluation of stem cells. With the limited understanding, the investigators in the field still imply properties of stem cells to mixed populations; ideally, to claim a function for a cell population, a homogeneous population should be purified and functionally tested. With the advent of techniques for expansion, approaches that use a combination of cell types, including HSCs and MSCs, in different proportions may be possible for future cellular therapy, each cellular mixture formulated to treat a unique tissue and a specific disease.

ACKNOWLEDGMENT

This work was supported by grants to JHDW from NSF BES-9631670 and NASA NAG9-1360. AM would like to thank the "Alexander S. Onassis" Foundation for providing a scholarship.

ARTICLES OF FURTHER INTEREST

Biomaterials Immune Response; Culture Systems; Histogenesis; Stem Cells; Tissue Engineering of Bone Marrow

REFERENCES

1. Quesenberry, P.; Colvin, G.A. Hematopoietic Stem Cells, Progenitor Cells, and Cytokines. In *Hematology*; Beutler, E., et al., Eds.; McGraw-Hill: New York, 2001; 153–174.
2. Madlambayan, G.J.; Rogers, I.; Casper, R.F.; Zandstra, P.W. Controlling culture dynamics for the expansion of hematopoietic stem cells. *J. Hematother. Stem Cell Res.* **2001**, *10*, 481–492.
3. Krause, D.S.; Fackler, M.J.; Civin, C.I.; May, W.S. CD34: Structure, biology, and clinical utility. *Blood* **1996**, *87*, 1–13.
4. Shih, C.-C.; DiGiusto, D.; Forman, S.J. Ex vivo expansion of transplantable human hematopoietic stem cells: Where do we stand in the year 2000? *J. Hematother. Stem Cell Res.* **2002**, *9*, 621–628.
5. Sutherland, H.J.; Lansdorp, P.M.; Henkelman, D.H.; Eaves, A.C.; Eaves, C.J. Functional characterization of individual human hematopoietic stem cells cultured at limiting dilution on supportive marrow stromal layers. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 3584–3588.
6. Craig, W.; Kay, R.; Cutler, R.L.; Lansdorp, P.M. Expression of Thy-1 on human hematopoietic progenitor cells. *J. Exp. Med.* **1993**, *177*, 1331–1342.
7. Osawa, M.; Hanada, K.; Hamada, H.; Nakauchi, H. Long-term lymphohematopoietic reconstitution by CD34-low/negative hematopoietic stem cells. *Science* **1996**, *273*, 242–245.
8. Sato, T.; Laver, J.H.; Ogawa, M. Reversible expression of CD34 by murine hematopoietic stem cells. *Blood* **1999**, *94* (8), 2548–2554.
9. Matsuoka, S.; Ebihara, Y.; Xu, M.-J.; Ishii, T.; Sugiyama, D.; Yoshino, H.; Ueda, T.; Manabe, A.; Tanaka, R.; Ikeda, Y.; Nakahata, T.; Tsuji, K. CD34 expression on long-term repopulating hematopoietic stem cells changes during developmental stages. *Blood* **2001**, *97* (2), 419–425.
10. Goodell, M.; Brose, K.; Paradis, G.; Conner, A.; Mulligan, A. Isolation and functional properties of murine hematopoietic stem cells that are replicating in vivo. *J. Exp. Med.* **1996**, *183*, 1797–1806.
11. Spangrude, G.J.; Johnson, G.R. Resting and activated subsets of mouse multipotent hematopoietic stem cells. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 7433–7477.
12. Zandstra, P.W.; Conneally, E.; Petzer, A.L.; Piret, J.M.; Eaves, C.J. Cytokine manipulation of primitive human hematopoietic cell self-renewal. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 4698–4703.
13. Chen, C.-Z.; Li, M.; de Graaf, D.; Monti, S.; Gottgens, B.; Sanchez, M.-J.; Lander, E.S.; Golub, T.R.; Green, A.R.; Lodish, H.F. Identification of endoglin as a functional marker that defines long-term repopulating hematopoietic stem cells. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99* (24), 15468–15473.
14. Fallon, P.; Gentry, T.; Balber, A.E.; Boulware, D.; Janssen, W.E.; Smilee, R.; Storms, R.W.; Smith, C. Mobilized peripheral blood SSC^{lo} ALDH^{br} cells have phenotypic and functional properties of primitive haematopoietic cells and their number correlates with engraftment following autologous transplantation. *Br. J. Haematol.* **2003**, *122*, 99–108.
15. Williams, D.E.; Hangoc, G.; Cooper, S.; Boswell, H.S.; Shaddock, R.K.; Gillis, S.; Waheed, A.; Urdal, D.; Broxmeyer, H.E. The effects of purified recombinant murine interleukin-3 and/or purified natural murine CSF-1 in vivo on the proliferation of murine high- and low-proliferative potential colony-forming cells: Demonstration of in vivo synergism. *Blood* **1987**, *70*, 401–403.
16. Eaves, C.J.; Eaves, A.C. Stem cell kinetics. *Bailliere's Clin. Haematol.* **1997**, *10*, 233–257.
17. Verfaillie, C.M. Ex Vivo Expansion of Hematopoietic Stem Cells. In *Hematopoiesis: A Developmental Approach*; Zon, L.I., Ed.; Oxford University Press: New York, 2001; 119–129.
18. Till, J.E.; McCulloch, E.A. A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. *Radiat. Res.* **1961**, *14*, 213–222.
19. McCune, J.M.; Namikawa, R.; Kaneshima, H.; Shultz, L.D.; Lieberman, M.; Weissman, I.L. The SCID-hu mouse: Murine model for the analysis of human hematolymphoid differentiation and function. *Science* **1998**, *24*, 1632–1639.
20. Larochelle, A.; Vormoor, J.; Hanenberg, H.; Wang, J.C.; Bhatia, M.; Lapidot, T.; Moritz, T.; Murdoch, B.;

- Xiao, X.L.; Kato, I.; Williams, D.A.; Dick, J.E. Identification of primitive human hematopoietic cells capable of repopulating NOD/SCID mouse bone marrow: Implications for gene therapy. *Nat. Med.* **1996**, *2*, 1329–1337.
21. Civin, C.I.; Almeida-Porada, G.; Lee, M.J.; Olweus, J.; Terstappen, L.W.; Zanjani, E.D. Sustained, retransplantable, multilineage engraftment of highly purified adult human bone marrow stem cells in vivo. *Blood* **1996**, *88*, 4102–4109.
 22. Orlic, D.; Kajstura, J.; Chimenti, S.; Jakoniuk, I.; Anderson, S.M.; Li, B.; Pickel, J.; McKay, R.; Nadal-Ginard, B.; Bodine, D.M.; Leri, A.; Anversa, P. Bone marrow cells regenerate infarcted myocardium. *Nature* **2001**, *410*, 701–705.
 23. Wang, X.; Ge, S.; McNamara, G.; Hao, Q.-L.; Crooks, G.M.; Nolte, J.A. Albumin-expressing hepatocyte-like cells develop in the livers of immune-deficient mice that received transplants of highly purified human hematopoietic stem cells. *Blood* **2003**, *101*, 4201–4208.
 24. Eglitis, M.A.; Mezey, E. Hematopoietic cells differentiate into both microglia in the brains of adult mice. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 4080–4085.
 25. Majka, S.M.; Jackson, K.A.; Kienstra, K.A.; Majesky, M.W.; Goodell, M.A.; Hirschi, K.K. Distinct progenitor populations in skeletal muscle are bone marrow derived and exhibit different cell fates during vascular regeneration. *J. Clin. Invest.* **2003**, *111*, 71–79.
 26. Korbling, M.; Katz, R.L.; Khanna, A.; Ruifrok, A.C.; Rondon, G.; Albitar, M.; Champlin, R.E.; Estrov, Z. Hepatocytes and epithelial cells of donor origin in recipients of peripheral-blood stem cells. *N. Engl. J. Med.* **2002**, *346* (10), 738–746.
 27. Wagers, A.J.; Sherwood, R.I.; Christensen, J.L.; Weissman, I. Little evidence for developmental plasticity of adult hematopoietic stem cells. *Science* **2002**, *297*, 2256–2259.
 28. Terada, N.; Hamazaki, T.; Oka, M.; Hoki, M.; Mastalerz, D.M.; Nakno, Y.; Meyer, E.M.; Morel, L.; Petersen, B.E.; Scott, E.W. Bone marrow cells adopt the phenotype of other cells by spontaneous cell fusion. *Nature* **2002**, *416*, 542–545.
 29. Ying, Q.-L.; Nichols, J.; Evans, E.P.; Smith, A.G. Changing potency by spontaneous fusion. *Nature* **2002**, *416*, 545–548.
 30. Quesenberry, P.; Colvin, G.A.; Lambert, J.-F. The chiaroscuro stem cell: A unified stem cell theory. *Blood* **2002**, *100* (13), 4266–4271.
 31. Cairns, J. Somatic stem cells and the kinetics of mutagenesis and carcinogenesis. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99* (16), 10567–10570.
 32. Corbel, S.Y.; Lee, A.; Yi, L.; Duenas, J.; Brazelton, T.R.; Blau, H.M.; Rossi, F.M.V. Contribution of hematopoietic stem cells to skeletal muscle. *Nature Med.* **2003**, *9* (12), 1528–1532.
 33. Collins, P.; Miller, W.M.; Papoutsakis, E.T. Ex vivo culture systems for hematopoietic cells. *Curr. Opin. Biotechnol.* **1996**, *7* (2), 223–230.
 34. Wang, X.; Willenbring, H.; Akkari, Y.; Torimaru, Y.; Foster, M.; Al-Dhalimy, M.; Lagasse, E.; Finegold, M.; Olson, S.; Grompe, M. Cell fusion is the peripheral source of bone-marrow-derived hepatocytes. *Nature* **2003**, *422*, 897–901.
 35. Vassilopoulos, G.; Wang, P.-R.; Russel, D.W. Transplanted bone marrow regenerates liver by cell fusion. *Nature* **2003**, *422*, 901–904.
 36. Nielsen, L.K. Bioreactors for hematopoietic cell culture. *Ann. Rev. Biomed. Eng.* **1999**, *1*, 129–152.
 37. McAdams, T.; Sandstrom, C.E.; Miller, W.M.; Bender, J.G.; Papoutsakis, E.T. Ex vivo expansion of primitive hematopoietic cells for cellular therapies: An overview. *Cytotechnology* **1995**, *18*, 133–146.
 38. McAdams, T.A.; Winter, J.N.; Miller, W.M.; Papoutsakis, E.T. Hematopoietic cell culture therapies (Part II): Clinical aspects and applications. *Trends Biotech.* **1996**, *14* (10), 388–396.
 39. Cairo, M.S.; Wagner, J.E. Placental and/or umbilical cord blood: An alternative source of hematopoietic stem cells for transplantation. *Blood* **1997**, *90*, 4664–4671.
 40. De La Selle, V.; Gluckman, E.; Bruley-Rosset, M. Newborn blood can engraft adult mice without inducing graft versus host disease across non H-2 antigens. *Blood* **1996**, *87*, 3977–3984.
 41. Sorg, R.V.; Kogler, G.; Wernet, P. Identification of cord blood dendritic cells as an immature CD11c-population. *Blood* **1999**, *93* (7), 2302–2307.
 42. Grewal, S.S.; Barker, J.N.; Davies, S.M.; Wagner, J.E. Unrelated donor hematopoietic cell transplantation: Marrow or umbilical cord? *Blood* **2003**, *101* (11), 4233–4244.
 43. McNiece, I.; Kubegov, D.; Kerzic, P.; Shpall, E.; Gross, S. Increased expansion and differentiation of cord blood products using a two-step expansion culture. *Exp. Hematol.* **2000**, *28*, 1181–1186.
 44. McNiece, I.; Briddell, R. Ex vivo expansion of hematopoietic progenitor cells and mature cells. *Exp. Hematol.* **2001**, *29*, 3–11.
 45. Jaroscak, J.; Goltry, K.; Smith, A.; Waters-Pick, B.; Martin, P.L.; Driscoll, T.A.; Howrey, R.; Chao, N.; Douville, J.; Burhop, S.; Fu, P.; Kurtzberg, J. Augmentation of umbilical cord (UCB) transplantation with ex vivo-expanded UCB cells: Results of a phase I trial using the AastromReplicell System. *Blood* **2003**, *101* (12), 5061–5067.
 46. Willert, K.; Brown, J.D.; Danenberg, E.; Duncan, A.W.; Weissman, I.; Reya, T.; Yates, J.R., III; Nusse, R. Wnt proteins are lipid-modified and can act as stem cell growth factors. *Nature* **2003**, *423*, 448–452.
 47. Mayani, H.; Lansdrop, P.M. Biology of human umbilical cord blood-derived hematopoietic stem/progenitor cells. *Srem. Cells* **1998**, *16*, 153–165.
 48. Kim, D.; Fujiki, Y.; Fukushima, T.; Ema, H.; Shibuya, A.; Nakauchi, H. Comparison of hematopoietic activities of human bone marrow and umbilical cord blood CD34 positive and negative cells. *Stem Cells* **1999**, *17*, 286–294.
 49. Koller, M.R.; Oxender, M.; Jensen, T.C.; Goltry, K.L.; Smith, A.K. Direct contact between CD34⁺/Lin⁻ cells and stroma induces a soluble activity that specifically increases primitive hematopoietic cell production. *Exp. Hematol.* **1999**, *27*, 734–741.



50. Bachier, C.R.; Gokmen, E.; Teale, J.; Lanzkron, S.; Childs, C.; Franklin, W.; Shpall, E.; Douville, J.; Weber, S.; Muller, T.; Armstrong, D.; LeMaistre, C.F. Ex-vivo expansion of bone marrow progenitor cells for hematopoietic reconstitution following high-dose chemotherapy for breast cancer. *Exp. Hematol.* **1999**, *27*, 615–623.
51. Sandstrom, C.E.; Bender, J.G.; Papoutsakis, E.T.; Miller, W.M. Effects of CD34⁺ cell selection and perfusion on ex vivo expansion of peripheral blood mononuclear cells. *Blood* **1995**, *86*, 958–970.
52. Koller, M.R.; Manchel, I.; Newsom, B.S.; Palsson, M.A.; Palsson, B.O. Bioreactor expansion of human bone marrow: Comparison of unprocessed, density-separated and CD34-enriched cells. *J. Hematother.* **1995**, *4*, 159–169.
53. Simmons, P.J.; Gronthos, S.; Zanneting, A.C.W. The Development of Stromal Cells In *Hematopoiesis: A Developmental Approach*; Zon, L.I., Ed.; Oxford University Press: New York, 2001; 718–726.
54. Haynesworth, S.E.; Reuben, D.; Caplan, A.I. Cell-based tissue engineering therapies: The influence of whole body physiology. *Adv. Drug Deliv. Rev.* **1997**, *33*, 3–14.
55. Owen, M.E. The Marrow Stromal Cell System. In *Marrow Stromal Cell in Culture*; Beresford, J.N., Owen, M.E., Eds.; Cambridge University Press: Cambridge, UK, 1988; 88–110.
56. Caplan, A.I.; Bruder, S.P. Mesenchymal stem cells: Building blocks for molecular medicine in the 21st century. *Trends Mol. Med.* **2001**, *7* (6), 259–264.
57. Gutierrez-Ramos, J.; Reyes-Maldonado, E.; Mayani, H. Characterisation of the adherent cells developed in Dexter-type long-term cultures from human umbilical cord blood. *Stem Cells* **2000**, *18*, 46–52.
58. Huss, R.; Lange, C.; Weissinger, E.M.; Kolb, H.J.; Thalmeier, K. Evidence of peripheral blood-derived, plastic-adherent CD34^{-/low} hematopoietic stem cell clones with mesenchymal stem cell characteristics. *Stem Cells* **2000**, *18*, 252–260.
59. Owen, M.E. Lineage of osteogenic cells and their relationship to the stromal system. *Bone Mineralisation Res.* **1985**, *3*, 1–25.
60. Friedenstein, A.J.; Gorskaja, J.F.; Kulagina, N.N. Fibroblast precursors in normal and irradiated mouse hematopoietic organs. *Exp. Hematol.* **1976**, *4*, 267–274.
61. Simmons, P.J.; Torok-Storb, B. Identification of stromal cell precursors in human bone marrow by a novel monoclonal antibody. STRO-1. *Blood* **1991**, *78*, 55–62.
62. Gronthos, S.; Simmons, P.J. The biology and application of human bone marrow stromal cell precursors. *J. Hematother.* **1996**, *5*, 15–23.
63. Onyia, J.E.; Clapp, D.W.; Long, H.; Hock, J.M. Osteoprogenitor cells as targets for ex vivo gene transfer. *J. Bone Mineral. Res.* **1998**, *13*, 2–30.
64. Young, R.G.; Butler, D.L.; Weber, W.; Caplan, A.I.; Gordon, S.L.; Fink, D.J. Use of mesenchymal stem cells in collagen matrix for Achilles tendon repair. *J. Orthop. Res.* **1998**, *16*, 406–413.

Hemocompatible Materials

Agnese Magnani

Federica M. Piras

Department of Chemical and Biosystem Sciences and Technologies, University of Siena, Siena, and University Center "Colle di Val d'Elsa," Colle di Val d'Elsa (SI), Siena, Italy

INTRODUCTION

Contact of blood with a foreign surface triggers the activation of a number of proteases of the *intrinsic pathway* of the coagulation cascade, which terminates in the formation of clots of fibrin, originated by thrombin-catalyzed cleavage of fibrinogen. The clot formation is the most important concern preventing even more widespread applications of biomaterials that contribute significantly to the quality and effectiveness of the world's health care system. Various strategies have been used to improve the blood compatibility of artificial materials although a truly hemocompatible surface does not exist yet.

All synthetic foreign materials are blood compatible in the sense that they do not have the intrinsic capacity of inducing the specific, evolution-selected mechanisms that lead to the thrombus formation. These mechanisms occur directly only with natural materials such as collagen and tissue factors. In contrast, artificial materials acquire thrombogenicity only after first interacting with blood, and the event responsible for the transformation of the inert polymer to an active surface is the interaction of plasma proteins with the surface that then initiates foreign-surface thrombosis. Although these statements pose an oversimplification, they suggest a critical role for protein adsorption in thrombosis on foreign materials. Processes such as protein adsorption, platelet and complement activation, and fibrin polymerization are all surface-induced phenomena and depend to a large extent on the properties and composition of the outer few atomic layers of a material. The successful design of materials for use as blood-contacting devices, thus demands both advanced techniques for surface analysis and adequate criteria for surface property determination, with the intent of correlating the molecular events occurring at the interface with the chemistry and the hemocompatibility of the biomaterial. Fundamental progress in design of new materials with improved hemocompatibility will, moreover, be determined by a deeper understanding of the nature of interfacial phenomena, mainly protein and platelet interactions with surfaces.

In the following discussion, the various surface and interface aspects of blood-contacting materials will be considered together with some of the different strategies commonly used for improving material hemocompatibility.

SURFACE MODIFICATION STRATEGIES FOR IMPROVING HEMOCOMPATIBILITY

Various strategies have been followed for enhancing material hemocompatibility with most of them related to surface modification approaches;^[1] in particular: 1) strategies for creating inert materials and 2) strategies for developing active materials.

Inert Materials

Much of the effort in biomaterials research over the past 35 yr has been directed toward the development of materials that do not react with platelets and coagulation factors. A number of approaches, often conflicting, have been developed and researchers have primarily focused their efforts on modifying surfaces of existing polymeric materials. This approach is reasonable as it is only the surface chemistry of the material that should dictate its biological responses. Despite the success in reducing protein and cellular deposits on some materials, a truly hemocompatible surface has not been achieved.

A popular method to improve the blood compatibility of materials is to increase their surface hydrophilicity by surface-incorporation of hydrogels, which permit the retention of large amounts of water providing the surface with minimal interfacial energy similar to biological tissues. Although the surface-incorporation of hydrogels produces protein-resistant surfaces, a large *ex vivo* study of a variety of materials, containing hydrogels incorporated at their surface, concluded that hydrogels do not possess low thrombogenicity.^[1] Surfaces have also been coated with hydrophilic polymers such as poly(vinyl pyrrolidone)



(PVP). The PVP is generally used to increase lubricity and ease catheter insertion, but its utilization also provided beneficial effects on thrombogenicity and bacterial adhesion.^[1]

Immobilization of the water-soluble synthetic polymer poly(ethylene glycol) (PEG) is a very popular means of making a biomaterial surface more protein and cell resistant. The protein-surface interactions have theoretically been modeled in the presence of PEG, concluding that steric repulsion by surface-bound PEG was largely responsible for the prevention of protein adsorption. The PEG and block copolymers poly(ethyleneoxide)/poly(propyleneoxide) (PEO/PPO) have been inserted into surfaces^[1] or combined to other strategies^[2] in an attempt to increase surface hemocompatibility. While the *in vitro* results have looked very promising, the lack of correlation between *in vitro* and *ex vivo* studies is of concern. More recent efforts with plasma-deposited tetraglyme^[3] have led to surfaces which adsorb very low amount of fibrinogen, suggesting that previous attempts at using PEG modification have not been successful because of the inability to achieve the desired ultralow levels of adsorbed protein.

The observation that surfaces coated with albumin did not support protein adsorption and platelet adhesion led many investigators to use either albumin coating or surfaces alkylation, which enhances albumin affinity, to improve materials hemocompatibility. Albumin adsorption is stated to lower material thrombogenicity, as it does not possess the peptide sequences to enable interaction with platelets and leukocytes or the enzyme receptors in the coagulation cascade. Albumin has been extensively immobilized on polymeric surfaces, and cellulose membranes (Cuprophane), used as dialysis membrane materials, were alkylated with long carbon chains^[1] to enhance albumin adsorption.

Based on the hypothesis that a surface similar to the external phospholipid membrane of cells should be blood compatible, various approaches have been pursued to incorporate phospholipids into surfaces, including phosphorylcholine coated or blended materials. *In vivo* observations proved that platelet adhesion was significantly reduced on phosphorylcholine-containing surfaces.^[4] The mechanism of action of such materials is however unclear, although the low protein adsorption observed on these surfaces has been attributed to their high free-water fraction.

The blending of a copolymer, consisting of polar and nonpolar blocks, to a base polymer has been considered a useful means of improving hemocompatibility of materials. Copolymers with an amphiphathic structure, added in low concentration, demonstrated the ability of migrating to the base polymer surface during and after fabrication. The

polycaprolactone-polysiloxane block copolymers have been used to lower the thrombogenicity of cardiopulmonary bypass and hemodialysis components by using surface-modified additives (SMA)-blended polymers or SMA-coated surfaces. A clinical evaluation of the effect of SMA on cardiopulmonary bypass circuits demonstrated a reduction in platelet interactions with no effect on complement activation.^[5]

The incorporation of fluorine into materials is another investigated strategy to improve blood compatibility. Fluorine groups are believed to inhibit protein adsorption and platelet adhesion because of their low surface energy. Fluoroalkyl groups as chain extender have been used to realize nonthrombogenic polyurethane, and a number of fluorinated SMA have been developed.^[1]

Plasma treatment is another surface modification technique, which results in a covalently attached coating on the surface of a substrate. Various materials were treated with a variety of monomers in plasma reactors, but the comparison of many plasma-modified surfaces did not allow to identify a modification chemistry that was superior to the base material in terms of platelet or leukocyte activation.^[6]

A number of synthetic polymers have been synthesized or modified^[7,8] to prepare polymers with heparin-like activities and then incorporated into polymer matrices or surfaces to enhance their hemocompatibility.^[7-10]

Self-assembled surface layers have been envisioned as useful templates to nucleate or organize ordered, designed biomaterials.^[11] Self-assembled monolayers (SAMs) of alkylsilanes on poly(dimethylsiloxane) (PDMS) have been used as model system.^[12] They have been demonstrated to lower platelet and fibrinogen deposition than surfaces composed of hydrophilic head groups in a canine *ex vivo* arteriovenous shunt model.

Active Materials

The limited successes of the previously discussed surface treatments for enhancing hemocompatibility have encouraged researchers to pursue also other strategies. The most popular and the one that started the field of bioactive materials is the heparinization of surfaces, although the incorporation of other antithrombotic and antiplatelet agents into material surfaces is gaining popularity.

Depending on its molecular weight and structure, heparin is able to inhibit, in association with its cofactor, antithrombin III, the serine proteases: thrombin and factor X. Heparinized surfaces should also help to minimize platelet activation. Heparin has been covalently, ionically, and physically incorporated

to various substrates utilizing a number of chemistries. The main concern has been that once immobilized, the heparin should be able to assume its native conformation and interact with antithrombin III. If heparin is bound ionically to the surface, then it can be slowly released over time. Selected techniques to produce materials that release heparin at biologically significant rates have been investigated.^[1] The thromboresistance of materials based on controlled release appears to be because of the microenvironment of heparin in solution at the blood–material interface. A commercial procedure to ionically bind heparin (Durafllo II) has been used to coat cardiopulmonary bypass circuits and other devices. While the heparin coating is able to reduce thrombus formation, during clinical trials, its effect on complement, contact activation, and inflammation has been contradictory.^[1]

To make the activity of immobilized heparin longer than that reached with ionic linkages, the molecule has been covalently immobilized to material surfaces. The conformation of the attached heparin and the point of attachment are critical factors determining the catalytic efficiency of the immobilized biomolecule. Heparin has been covalently bound to surfaces by different methods.^[1] The most popular chemistry for covalently binding heparin to surfaces of plastics, glass, and steel is the Carmeda method^[13a] The technique used provided highly stable coatings with low thrombogenicity that have been demonstrated to retain good efficacy *in vivo*. In this case, a beneficial effect on *in vivo* bacterial colonization of treated polyurethane central venous catheters was also observed.^[1]

Recently, a comparison of heparinized materials obtained by the Durafllo and Carmeda methods showed no dramatic differences between the two surfaces in clinical trials.^[13b]

Bioline coating is a newer surface-heparinization method that is used to cover components of extracorporeal perfusion circuits.^[14] The effectiveness of this surface coating is associated with platelet preservation, amelioration of inflammatory response, and reduction of fibrinolytic activity during cardiopulmonary bypass. Surface heparinization limited to the oxygenator and the arterial filter had similar results as totally surface-heparinized circuits.^[15]

Agents other than heparin have been surface immobilized to inhibit thrombin, and thereby enhance material hemocompatibility. Among these, hirudin has some advantages compared to heparin, because it has no influence on platelet function, has no immune-mediated platelet-activating activity, and, most importantly, does not require the presence of endogenous cofactors such as antithrombin III. Thus, hirudin has been covalently bound to different surfaces with the aim of improving their blood compatibility

and its controlled release has been exploited.^[1] Another approach involves the immobilization of thrombomodulin, an endothelial cell-associated protein that inhibits thrombin by the activation of protein C. Thrombomodulin has been immobilized on different surfaces and both its anticoagulant activity and the ability to reduce platelet adhesion have been demonstrated.^[1,16]

The inevitable presence of platelets on biomaterials dictated the interest to incorporate antiplatelet agents into materials. Prostaglandins, prostacyclins, and their conjugates with heparin have been incorporated into various material surfaces.^[1] Antiplatelet agents based on inhibiting fibrinogen binding to activated platelet GPIIb/IIIa receptors have been used to reduce surface thrombogenicity via drug release. A novel approach was to exploit endogenous S-nitrosoproteins in plasma to produce NO from immobilized cysteine to minimize platelet adhesion on polyurethane and poly(tetrafluoroethylene) (PTFE).^[17]

Surface thrombogenicity can also be lowered by the lysis of fibrin clots on artificial surfaces. Fibrinolysis can be promoted by inducing surface generation of plasmin. Plasminogen has been directly immobilized, or selectively adsorbed by the immobilized lysine, onto surfaces and then converted to plasmin, so that the surfaces acquire fibrinolytic activity.^[1]

It is intuitive to believe that the ideal surface for vascular grafts/artificial hearts will consist of an intact luminal endothelial cell layer. Dacron [polyethyleneterephthalate (PET)] and PTFE have been seeded with endothelial cells in a preliminary clotting step,^[18a] and encouraging results have been obtained *in vivo* experiments with dogs.^[18b] Several attempts have been made to modify materials to give them a higher affinity for endothelial cells. For example endothelial cells have been seeded on CO₂ plasma-treated polystyrene precoated with substances able to promote endothelial cell adhesion.^[18c] Seeding this surface with endothelial cells (ECs) significantly reduced the number of adhered platelets. Moreover, polystyrene and PTFE surfaces modified by NH₃ plasma facilitated the attachment of ECs.^[18d] Surfaces precoated with fibrinogen or fibronectin^[18e] or containing Arg-Gly-Asp (RGD) peptides^[18f] resulted in higher number of adhered cells.

The recent development of stents for maintaining patency in coronary arteries after angioplasty introduces metallic surfaces to the circulatory system. In lowering metal thrombogenicity, the interest was firstly directed toward heparin-coated stents. These have been then supplanted by rapamycin-eluting stents that have the effect of eliminating *in-stent* restenosis for at least six months after deployment.^[19] Other strategies have been used to enhance the stent hemocompatibility by coating the metal surface with polyurethane and



parylene, or hyaluronic acid, or antiplatelet agents to reduce platelet adhesion.^[1]

SURFACE ANALYSIS^[20–22]

Determination of Surface Properties

To achieve a more complete understanding of the molecular-level interactions between foreign materials and biological systems, surface properties, (surface structure, surface hydrophobicity/hydrophilicity ratio, and surface charge) and subtle surface parameters (surface compliance, functional group distribution and orientation, surface domain distribution, surface dynamics, surface contamination, etc.) should be addressed.

Surfaces have some peculiar characteristics. They are: 1) uniquely reactive; 2) inevitably different from the bulk; thus, traditional techniques used for bulk structure analysis are not suitable for surface determination; 3) readily contaminated; and 4) dynamic—surfaces can restructure in response of the local environment. Enrichment or depletion of surface components may occur as a result of both diffusion processes from/to the inside of the matrix of the material and reactions with the environment adjacent to the material surface. For these reasons, the techniques that can directly probe this interface are of extreme importance.

Surface analysis of materials should provide data on elemental composition of the surface region, surface

functional groups, extent of the preferential orientation of the functional groups, elemental depth profiles, surface crystalline order, surface domain structure, surface energy, and surface topography.

Surface Property Measurements

Many methods are available for surface characterization, with each providing a unique piece of information about the nature of the surface. Table 1 summarizes the characteristics of many common surface analysis methods.

Sample analysis is guided by two general principles:

1. All methods used to analyze surfaces may also alter the surface; the surface analyst must be aware of potential damage of the method.
2. More than one surface analysis method should be used whenever possible to reduce the possibility of artifacts and to obtain the several pieces of information necessary to build a complete picture of the surface. Information from multiple surface analysis methods is often corroborative and synergistic in developing an understanding of surface structure.

Specific Methods for Surface Analysis

Only the most common surface analytical methods are considered here. As many review articles have been published on each method,^[20–22] they are briefly

Table 1 Surface analytical techniques for biomaterials surface characterization

Technique	Principle	Depth analyzed	Lateral resolution
<i>Microscopic methods</i>			
SEM	Secondary electrons emission from a surface struck by an electron beam	5 Å	40 Å
AFM	Measurement of the forces between a tip and a surface	5 Å	1 Å
<i>Spectroscopic/spectrometric methods</i>			
ATR-IR spectroscopy	Molecular vibrations induced by the absorption of IR radiation	1–10 μm	
XPS	Emission of electrons of characteristic kinetic energy induced by x-rays	10–250 Å	8–150 μm
SIMS	Secondary ions emission induced by primary ions bombardment	10 Å–1 μm ^a	500 Å
<i>Thermodynamic methods</i>			
Contact angle	Liquid wetting of surfaces for estimating surface energy	3–20 Å	1 mm ^b

^aStatic SIMS ~ 10 Å; dynamic SIMS to 1 μm.

^bThe size of a small drop is 1 mm.

discussed only in terms of the information they provide. These methods can be grouped as

1. Microscopic methods
2. Spectroscopic/spectrometric methods
3. Thermodynamic methods

Microscopic methods

The simplest picture of a sample surface is provided by the optical microscopic observation, which tells us how homogeneous the surface is. The utility of optical microscopy can be extended by the use of specialized methods including polarized microscopy, interference and differential contrast microscopy, fluorescence microscopy, and confocal microscopy.

A more detailed picture of surfaces is provided by scanning electron microscopy (SEM). The SEM image of surfaces provides 3-D-like topography. The major disadvantage of this microscopy is that nonconductive materials need to be coated with a very thin layer of metal, typically Au. The energetic electrons can damage beam-sensitive samples, including polymers and biological structures; thus, modern instruments utilize a field emission source that can operate at very low accelerating voltages (100 eV).

A much more advanced image of surfaces at the molecular level is obtained by scanning tunneling microscopy (STM), atomic force microscopy (AFM), and the so-called scanning probe microscopies (SPM, a generalized term to indicate the AFM modes). The STM measures electrical current and therefore is well suited for conductive and semiconductive surfaces. However, images of biomolecules on conductive surfaces can be obtained. In AFM, instead of recording tunneling current, the deflection of a tip mounted on a flexible cantilever owing to surface forces between the tip and the sample surface is measured. As the AFM measures forces, it can be used with both conductive and nonconductive surfaces. The AFM can operate in contact, semicontact, and noncontact modes. Moreover, operating the AFM in liquid opens the possibility of observing biological samples in a physiological environment, surface restructuring of polymeric materials in aqueous media, and interfacial phenomena such as protein adsorption on solid surfaces. The SEM and AFM are very useful when used in conjunction with other surface analysis methods; e.g., data on surface roughness are very useful in the interpretation of x-ray photoelectron spectroscopy (XPS), secondary ion mass spectrometry (SIMS), and contact angle measurements, which are known to be strongly affected by surface roughness and texture.

Spectroscopic/spectrometric methods

Infrared (IR) spectroscopy provides information on the molecular structure of materials observing the vibrations of molecular units excited by the absorption of specific frequencies of energy in the IR range. The attenuated total reflection infrared method (ATR-IR) has been widely used for biomaterial surface characterization. In ATR, the IR spectroscopy is coupled with the physical phenomenon of total internal reflection to enable the study of molecular vibrations within the surface regions of the material. The ATR provides molecular structure, conformation, crystalline, and secondary bonding information about the sample surface region. It has been successfully used to study the solid–aqueous interface *in situ*^[23,24] by monitoring interfacial phenomena such as protein adsorption to surfaces with the intent of correlating the molecular events occurring with the surface chemistry and hemocompatibility.

Other “surface sensitive” IR techniques are: external reflection IR that provides extreme surface sensitivity and gives information on the orientation of molecules at the surface; photoacoustic methods based on IR heating and sample expansion for probing the surface; diffuse reflectance IR, whose advantage is that it requires little or no sample preparation prior to analysis; and finally, IR microscopy and IR imaging, which combines the surface sensitivity with lateral resolution.

The XPS is an important technique for the surface characterization and analysis of biomaterials. This technique, based upon the photoelectric effect, provides an elemental analysis of the top 10–100 Å of any solid surface that is vacuum compatible. It has several advantages, i.e., high information content, the speed of analysis, the low damage potential, and the ability to analyze most samples with no specimen preparation. The disadvantages are the need for vacuum compatibility and the possibility of sample damage by x-rays. The vacuum limitation can be overcome by cooling the sample at liquid nitrogen temperatures. The angle-dependent XPS mode gives a nondestructive depth profiling analysis.

The SIMS produces a mass spectrum of the outermost 10 Å of a surface. Like XPS, it requires complex instrumentation and an ultrahigh vacuum system. The method provides unique information complementary to XPS for an understanding of surface composition. Secondary ion mass spectrometry methods are distinguished in static and dynamic SIMS depending on the ion dose used. Dynamic SIMS is surface erosive and enables the analyst to perform a depth profile of a specimen extending from the outermost atoms to a micron or more into the sample. However, because of the damaging nature of the high-flux beam, only atomic fragments can be detected. Static SIMS induces



minimal surface destruction so that intact molecular fragments can be detected. These molecular fragments contain specific chemical information closely related to the surface composition. Spectral interpretation is assisted by specialized methods.

The advances that have been made in SIMS instrumentation [e.g., development of time-of-flight (ToF) mass analyzers] and data analysis methods (e.g., multivariate analysis) in the past 20 yr have resulted in static ToF-SIMS now being an important and widely used surface analysis technique for characterization of biomaterial surfaces.^[25]

Thermodynamic methods

Contact Angle Measurements. This method provides the surface energy of surfaces. The surface energy is directly related to its wettability and it is a useful parameter that frequently correlates with biological interactions. Experimentally, there are a number of ways to measure the contact angle. The contact angle methods are unique in that the equipment required is relatively simple and inexpensive. Although interpretation of the results obtained is dependent on a number of assumptions, each of which is somewhat controversial, a first-order interpretation is possible and has proven to be useful in many areas of surface science.

Newer methods

Nowadays, there are other surface characterization techniques that are becoming important for

biomaterial surfaces characterization. They are summarized in Table 2.

PROTEIN ADSORPTION

Protein adsorption is the first event occurring at the blood–material interface. The process occurs under complex conditions owing to the fact that the plasma phase of blood contains many different proteins and each protein is present at a characteristic concentration that can differ greatly from that of the other proteins (Table 3).

Several models have been developed to describe the adsorption of proteins onto solid surfaces.^[26,27] In these models both the thermodynamic and/or the kinetic parameters have been evaluated. The very rapid initial protein adsorption typically observed appears to be diffusion limited. The thermodynamics of protein adsorption is not easily described in terms of the familiar concepts of free energy, enthalpy, and entropy because these properties are measurable in reversible processes, whereas protein adsorption appears to be essentially irreversible. The irreversibility of the process can be explained in terms of protein size. Multiple bonding between the relatively large contact face of the protein molecule and the surface can in fact occur, and it is unlikely that all these bonds will break simultaneously. The irreversible adsorption isotherms cannot, however, be correctly analyzed to calculate a true equilibrium constant and free energy.

A more reliable insight into the thermodynamics can be obtained by the heats of adsorption. The studies

Table 2 New surface analytical techniques for biomaterials surface characterization

Technique	Principle
Sum frequency generation (SFG)	The SFG is an optical probe of surfaces and interfaces, where light is generated at a frequency, which is the sum of the frequencies of two incident optical fields.
Surface-enhanced Raman spectroscopy (SERS)	The SERS provides greatly enhanced Raman signal from Raman-active analyte molecules that have been adsorbed onto certain specially prepared metal surfaces.
High resolution electron energy loss spectroscopy (HREELS)	Vibrational spectroscopy of a highly surface-localized region, under ultrahigh vacuum.
Ion scattering spectroscopy (ISS)	Elastically reflected ions probe the outermost atomic layer of the sample surface.
Rutherford backscattering spectroscopy (RBS)	The RBS is based upon bombarding a sample with energetic ions and measuring the energy of the backscattered ions.
Extended X-ray absorption fine structure (EXAFS)	The EXAFS refers to oscillatory structures that appear in the high energy side of characteristics X-ray absorption edges of elements.
Scanning Auger microscopy (SAM)	An electron beam is scanned over the surface and the electrons excited from the surface are energy analyzed to detect Auger peaks.
Surface plasmon resonance (SPR)	An SPR is excited at a metal/dielectric interface by a monochromatic light beam.
Synchrotron methods	By accelerating electrons to near the speed of light in a large, circular ring, energies covering a broad range of the electromagnetic spectrum (IR to X-rays) are emitted.

Table 3 Some plasma proteins

Protein	mg/100 ml serum	MW
Albumin	3500–5500	69,000
IgG	800–1800	150,000
IgA	90–450	150,000–200,000
IgM	60–250	900,000
IgE	~300	190,000
Fibrinogen ^a	200–450	340,000
vWF ^a	1.5–5.0	80,000

^aAdhesive plasma proteins responsible for platelet interaction.

in this direction^[26] support the importance of entropic factors in the adsorption process, which arise from changes in water binding to the surface and the protein. The irreversibility of protein adsorption does not prevent, however, the exchange of a previously adsorbed protein with other bulk proteins.

In the understanding of the adsorption events occurring at the blood–solid surface interface, four principles should be mainly taken into account.

1. Adsorption is essentially limited to a monolayer, i.e., the protein molecules that directly interact with the surface are within a monolayer and the excess proteins on the surface should be considered as a deposit.
2. The driving forces determining the composition of the adsorbed protein layer are the intrinsic surface activity of each plasma protein and its bulk concentration in the plasma.

3. Surfaces vary in their ability to adsorb selectively and cause changes in the reactivity of proteins.
4. The ability of an adsorbed protein to influence platelet adhesion and activation depends not only on the amount of that protein, but also on its reactivity, as determined by changes in its orientation or conformation.^[28–30]

Monolayer Model

The evidence for the existence of a monolayer of adsorbed proteins comes most clearly from studies with single-protein solutions in which saturation can often be observed in the adsorption isotherm (Fig. 1). Usually the plateau value falls within the range expected for a close-packed monolayer of protein ($0.1\text{--}0.5\ \mu\text{g}/\text{cm}^2$ depending on the diameter and orientation assumed for the protein). Moreover, studies on the adsorption from complex protein mixtures or plasma also give values that fall in this range, supporting the monolayer model. The competition between the different plasma proteins adsorbing to a surface reflects the availability of a limited number of sites, and the multilayer deposition model would not be expected to lead to the observed protein discrimination by the surface. There are some limitations in the use of the monolayer concept even though it appears to be a good “zero-order” approximation to what really happens in blood. These limitations concern protein–protein interactions that may occur in whole blood, and the fact that both the orientation of the protein and the true surface area available for adsorption per

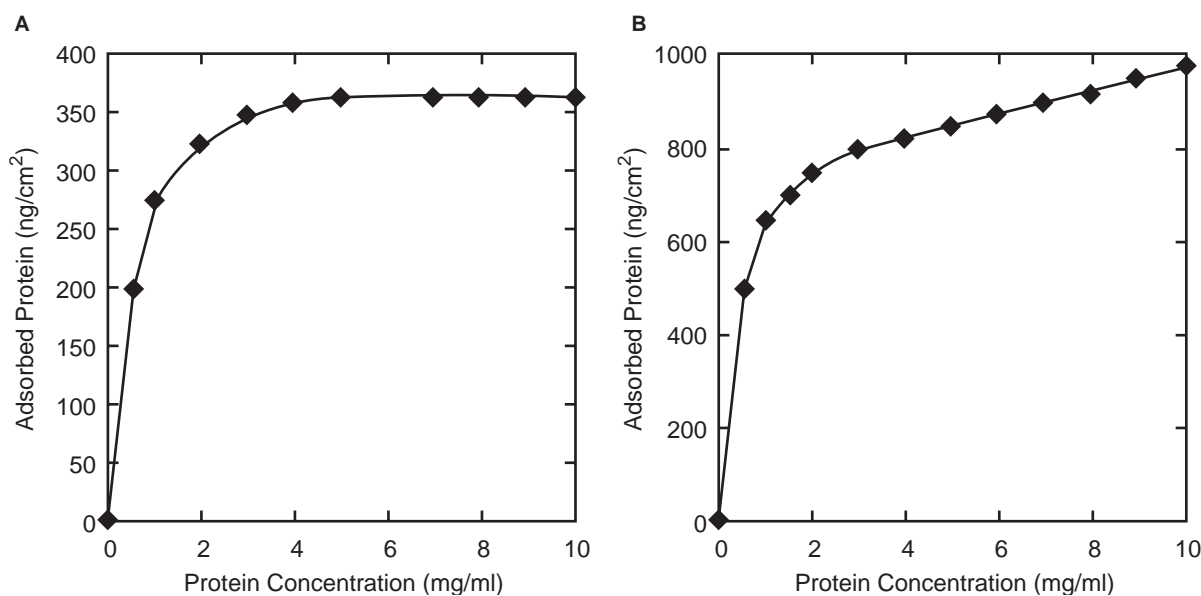


Fig. 1 Monolayer concept: Langmuir (A) and Freundlich (B) adsorption isotherms. The amount of the adsorbed protein is plotted vs. the bulk protein concentration in the solution from which adsorption took place.



protein molecule are not known. Both the classical Langmuir isotherm (true plateau) and the Freundlich isotherm (slow rise) (Fig. 1) have been observed for protein adsorption to solid surfaces.^[29] The monolayer model does not give any information about protein orientation and fibrin formation.

Competitive Adsorption

Adsorption from blood plasma is selective and the surface phase results enriched in certain proteins. Because the number of adsorption sites per unit area is limited, the solid surface can only accommodate a small fraction of the total protein typically present in the bulk phase. As the plasma proteins vary greatly both in their concentration in the bulk phase and in their intrinsic abilities to adsorb to surfaces, some proteins will be adsorbed preferentially. Despite the fact that fibrinogen is only the third most concentrated protein in plasma, after IgG and albumin, materials exposed to blood are typically enriched in fibrinogen in the adsorbed phase. An opposite behavior is observed for albumin owing to its low intrinsic surface activity, although it is the most abundant protein in plasma. The adsorption of fibrinogen from plasma, on the other hand, exhibits some unusual behaviors. On certain surfaces, fibrinogen adsorption is maximum at intermediate dilutions of plasma (Fig. 2). In addition, fibrinogen adsorption from whole or moderately diluted plasma is higher at the very early adsorption times. This is termed the “Vroman effect”^[30] and it is a very clear example of the unique effects of competitive adsorption on both the steady state and the transient

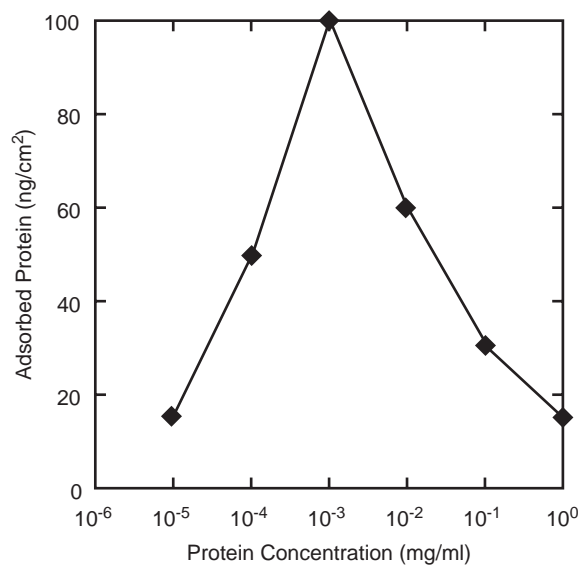


Fig. 2 Representation of the “Vroman effect” for fibrinogen adsorption from plasma.

composition of the adsorbed layer that forms from plasma.

However, studies of surfaces exposed to plasma showed that several different proteins are present in the adsorbate even if enrichment in certain proteins can occur. The concept of a complete “domination” of the adsorbed layer by only one protein (such as fibrinogen), thus, is not very realistic.^[29] Moreover, both the overall surface features and the overall structural stability of a protein have to be considered as key variables for the adsorption process.^[29]

The Influence of Surface Chemistry

The outcome of the competitive process of adsorption is an adsorbed film that is richer in some proteins than others. Because each protein has a different tendency to adsorb onto each type of surface, the outcome of the protein competition is surface chemistry dependent, and the composition of the adsorbed film may be different for each type of biomaterial. Moreover, the composition of the adsorbed layer may change strongly with time. These are the fundamental facts that are believed to underlie at least some of the differences in blood compatibility among various biomaterials, and many of the strategies for improving biomaterial hemocompatibility have been based upon these observations. Most of them resulted in attempts to produce biomaterials that would preferentially adsorb the “right” protein (e.g., an inert one such as albumin), excluding more reactive proteins, e.g., fibrinogen, that tend to favor thrombus formation, or attempts to prevent the formation of the natural-forming mixed layer by the preadsorption of an inert protein to the surface. Efforts have also been made to evade the entire question of the adsorbed protein layer composition by developing nonfouling surfaces. Surfaces bearing polyethylene oxide (PEO) have been shown in many studies to be particularly effective in making a biomaterial surface more protein and cell resistant.^[31] Recent studies with plasma-deposited tetraglyme, as already mentioned, have led to surfaces with ultralow-adsorbed fibrinogen.^[3] Also heparinized polymeric surfaces through an ionic bond have been demonstrated to adsorb low amounts of proteins, without inducing any conformational changes in the adsorbed species.^[24]

Influence of the Conformation and Orientation of the Adsorbed Proteins on Their Bioactivity

Proteins in the adsorbed phase may undergo both noncovalent structural transitions and covalent changes. Although the covalent changes (e.g., fibrin formation and cross-linking, plasmin-catalyzed degradation of adsorbed fibrinogen, and cleavage of adsorbed factor

XII) are undoubtedly a major determinant in the response of blood to synthetic materials, the non-covalent transitions of the adsorbed proteins also play an important role in the interaction with blood.

When proteins adsorb to solid surfaces, they may change their native conformation because of their relatively low structural stability and their tendency to unfold to allow further bond formation with the surface. In spite of these conformational changes, many proteins retain at least some of their activity in the adsorbed state.^[23] In addition, proteins may undergo time-dependent transitions upon their adsorption to foreign surfaces. Many indirect evidences suggest the existence of multiple “states” of adsorbed proteins, which implies a transition step in the formation of such mixed populations.^[29] The concept of transition from a reversibly adsorbed protein to a more tightly held one is included in several models of protein adsorption from single-protein solutions at the solid–liquid interface, and it is also used to explain the time-dependent fibrinogen displacement occurring at the plasma–surface interface.^[30] During the residence time, i.e., the time the protein spends adsorbed on the surface before its “state” is probed, many orientational transitions may occur if the protein continues to interact with the surface provoking the unfolding of the protein and largely reducing the adsorbate elutability.^[32]

Finally, the orientation of proteins in the adsorbed phase is an important concept because proteins are not uniform across their surface. As proteins are not very free to rotate once adsorbed, owing to multiple bonding, a fixed array of the protein surface is exposed to the bulk phase. The degree of uniformity in such orientation is not known, calculations of protein–surface interaction energies in a wide array of orientations of the adsorbed protein against a theoretical plane^[33] suggest that the orientations favoring the greatest possible degree of contact between the face of the protein and the plane are preferred.

These findings are important because they appear to suggest a way that differences in hemocompatibility among different biomaterials may be generated (e.g., by altering the availability of epitopes on adhesive proteins for the cell- and platelet-surface receptors). These processes are important even though the *in vivo* significance of all these changes is not known, and whether the transition changes play a significant role in determining blood compatibility is not clear.

PLATELET INTERACTION WITH BIOMATERIAL SURFACES: INFLUENCE OF THE ADSORBED PROTEINS

The adhesion, activation, and aggregation of platelets, followed in some cases by embolization of platelet

thrombi, is an accepted description of the sequence of events that occur upon exposure of foreign surfaces to flowing blood. The platelet-surface interactions with adsorbed proteins are based on the presence of adhesion receptors in the platelet membrane that bind to certain plasma proteins. Some proteins present in plasma (Table 3) first adsorb on the surface and then mediate platelet adhesion through receptor binding. The proteins strongly adhesive for platelets are fibrinogen, fibronectin, vitronectin, and von Willebrand factor (vWF). The vWF is believed to play a dominant role in platelet binding to subendothelial tissues under high shear conditions, while fibrinogen mediates adhesion to foreign surfaces.^[29] The “platelet Vroman effect” (a maximum in platelet adhesion on surfaces preadsorbed with diluted plasma) also supports the importance of platelet–fibrinogen interactions in mediating platelet adhesion to foreign surfaces.^[29]

Exactly what properties of the material and/or the adsorbed protein layer are critical in causing such activation and what role is played by locally generated concentrations of soluble agonist such as thrombin remain unclear. It is thought, however, that the behavior of fibrinogen, the major protein in plasma known to promote platelet adhesion to foreign surfaces, is especially important.^[34] Studies of the time- and surface-dependent transitions of fibrinogen and the subsequent observation that these transitions can influence the ability of platelets to adhere^[30] are all consistent with a specific mechanism of the adhesion and activation of platelets. The rate and the degree of transformation of adsorbed platelet-adhesive proteins such as fibrinogen appear to be controlled by the surface properties of the polymer, so that the reactivity of a surface to the platelets is in turn dictated by the degree to which the adhesive protein has been converted from a reactive form to an unreactive form. That is, the properties of the surface to which the adhesive protein is adsorbed influence both the “amount” and the “bioreactivity” of the protein itself, which, in turn, affect the platelet response.

CONCLUSIONS

The complexity of the interactions between materials and blood makes the development of hemocompatible surfaces very difficult. Biomaterials research is focused on developing materials that have fewer risks and greater benefits. Stents that “actively” prevent restenosis are a great example of how surface modification of metals can result in a dramatic clinical effect. Inert materials are protein and platelet resistant, but they do not prevent thromboembolic phenomena *in vivo*. On the other hand, incorporating anticoagulants into



surfaces reduces thrombin production, but this may not be sufficient to prevent platelet adhesion and activation. Thus, which approach will ultimately be successful is impossible to predict. Perhaps, the failure to produce an ideal blood-compatible surface merely reflects our limited understanding of the complex blood-materials interaction.

Progresses in design hemocompatible materials will be then determined by a deeper understanding of both the surface properties of materials and the interfacial phenomena originating at the blood-material interface.

ARTICLES OF FURTHER INTEREST

Biocompatibility Testing; Biomaterials: Protein-Surface Interactions; Thrombosis

REFERENCES

- Sefton, M.V.; Gemmel, C.H. Nonthrombogenic treatments and strategies. In *Biomaterials Science. An Introduction to Materials in Medicine*, 2nd Ed.; Ratner, B.D., Hoffman, A.S., Schoen, F.J., Lemons, J.E., Eds.; Elsevier Academic Press: San Diego, 2004; 456–470.
- Kim, K.; Kim, C.; Byun, Y. Preparation of a PEG-grafted phospholipid Langmuir-Blodgett monolayer for blood-compatible material. *J. Biomed. Mater. Res.* **2000**, *52*, 836–840.
- Shen, M.; Pan, Y.V.; Wagner, M.S.; Hauch, K.D.; Castner, D.G.; Ratner, B.D.; Horbett, T.A. Inhibition of monocyte adhesion and fibrinogen adsorption on glow discharge plasma deposited tetraethylene glycol dimethyl ether. *J. Biomat. Sci. Polym. Ed.* **2001**, *12*, 961–978.
- Lewis, A.L.; Hughes, P.D.; Kirkwood, L.C.; Leppard, S.W.; Redman, R.P.; Tolhurst, L.A.; Stratford, P.W. Synthesis and characterization of phosphorylcholine-based polymers useful for coating blood filtration devices. *Biomaterials* **2000**, *21*, 1847–1859.
- Gu, Y.J.; Boonstra, P.W.; Rijnsburger, A.A.; Haan, J.; van Oeveren, W. Cardiopulmonary bypass circuit treated with surface-modifying additives: a clinical evaluation of blood compatibility. *Ann. Thorac. Surg.* **1998**, *65*, 1342–1347.
- Sefton, M.V.; Sawyer, A.; Gorbet, M.; Black, J.P.; Cheng, E.; Gemmel, C.; Pottinger-Cooper, E. Does surface chemistry affect thrombogenicity of surface modified polymers? *J. Biomed. Mater. Res.* **2001**, *55*, 447–459.
- Barbucci, R.; Cialdi, G.; Magnani, A. Novel Heparin-Like Sulfated Polysaccharides. European Patent-PCT WO 95/25751; Sulfated Hyaluronic Acid Esters. U.S. Patent-6051701, April 18, 2000.
- Park, H.D.; Lee, W.K.; Ooya, T.; Park, K.D.; Kim, Y.H.; Yui, N. Anticoagulant activity of sulfonated polyrotaxanes as blood-compatible materials. *J. Biomed. Mater. Res.* **2002**, *60*, 186–190.
- Magnani, A.; Lamponi, S.; Consumi, M.; Barbucci, R. Biological performance of two materials based on sulphated hyaluronic acid and polyurethane. *J. Mater. Chem.* **1999**, *9*, 2393–2398.
- Barbucci, R.; Magnani, A.; Rappuoli, R.; Lamponi, S.; Consumi, M. Immobilization of sulphated hyaluronan for improved biocompatibility. *J. Inorg. Biochem.* **2000**, *79* (1–4), 119–125.
- Ratner, B.D. Surface modification of polymers: chemical, biological and surface analytical challenges. *Biosens. Bioelectron.* **1995**, *10*, 797–804.
- Silver, J.H.; Lin, J.C.; Lim, F.; Tegoulia, V.A.; Chaudhury, M.K.; Cooper, S.L. Surface properties and hemocompatibility of alkyl-siloxane monolayers supported on silicon rubber: effect of alkyl chain length and ionic functionality. *Biomaterials* **1999**, *20*, 1533–1543.
- a. Larm, O.; Larsson, R.; Olsson, P. A new non-thrombogenic surface prepared by selective covalent binding of heparin via a modified reducing terminal residue. *Biomater. Med. Dev. Art. Org.* **1983**, *11*, 161–173; b. Ovrum, E.; Tangen, G.; Oystese, R.; Ringdal, M.A.L.; Istad, R. Comparison of two heparin-coated extracorporeal circuits with reduced systemic anticoagulation in routine coronary artery by pass operations. *J. Thorac. Cardiovasc. Surg.* **2001**, *121*, 324–330.
- Wendel, H.P.; Ziemer, G. Coating-techniques to improve the hemocompatibility of artificial devices used for extracorporeal circulation. *Eur. J. Cardiothorac. Surg.* **1999**, *16*, 342–350.
- Palatianos, G.M.; Foroulis, C.N.; Vassili, M.I.; Astras, G.; Triantafillou, K.; Papadakis, E.; Lidoriki, A.A.; Iliopoulou, E.; Melissari, E.N. A prospective, double-blind study on the efficacy of the bioline surface-heparinized extracorporeal perfusion circuit. *Ann. Thorac. Surg.* **2003**, *76*, 129–135.
- Han, H.S.; Yang, S.L.; Yeh, H.Y.; Lin, J.C.; Wu, H.L.; Shi, G.Y. Study of a novel human thrombomodulin immobilized substrate: surface characterization and anticoagulation activity evaluation. *J. Biomater. Sci. Polym. Ed.* **2001**, *12*, 1075–1089.
- Duan, X.; Lewis, R.S. Improved haemocompatibility of cysteine-modified polymers via endogenous nitric oxide. *Biomaterials* **2002**, *23*, 1197–1203.
- a. Herring, M.B.; Baughman, S.; Glover, J.; Kesler, K.; Jesseph, J.; Dillay, R.; Evan, A.; Gardner, A. Endothelial seeding of Dacron and polytetrafluoroethylene grafts: the cellular events of healing. *Surgery* **1984**, *96*, 745–754; b. Zilla, P.; Detsch, M.; Meinhart, J.; Endothelial cell transportation. *Semin. Vasc. Surg.* **1999**, *12*, 52–63; c. Bos, G.W.; Scharenborg, N.M.; Poot, A.A.; Engbers, G.H.; Beugeling, T.; van Aken, W.G.; Feijen, J. Blood compatibility of surfaces with immobilized albumin-heparin conjugate and effect of endothelial cell seeding on platelet adhesion. *J. Biomed. Mater. Res.* **1999**, *47*, 279–291; d. Sipehia, R. The enhanced attachment and growth of endothelial cells on anhydrous ammonia gaseous plasma modified surface of polystyrene and poly (tetrafluoroethylene). *Biomater. Art. Cells Art. Org.* **1990**, *18*, 437–446; e. Van Wachen, P.B.;

- Vreciks, C.N.; Bengeling, T.; Feijen, J.; Bautjes, A.; Detmers, J.P.; van Aken, W.G. The influence of protein adsorption on interactions of cultured human endothelial cells with polymers. *J. Biomed. Mater. Res.* **1987**, *21*, 701–718; f. Massia, S.P.; Hubbel, J.A. Covalently attached GRGD on polymer surfaces promotes biospecific adhesion of mammalian cells. *Ann. N. Y. Acad. Sci.* **1990**, *589*, 261–270.
19. Morice, M.C.; Serruys, P.W.; Sousa, J.E.; Fajader, J.; Ban Hayashi, E.; Perin, M.; Colombo, A.; Schuler, G.; Barragan, P.; Guagliumi, G.; Molnar, F.; Falotico, R. RAVEL Study Group. Randomized study with the sirolimus-coated Bx velocity balloon-expandable stent in the treatment of patients with de novo native coronary artery lesion. A randomized comparison of a sirolimus-eluting stent with a standard stent for coronary revascularization. *N. Engl. J. Med.* **2002**, *346*, 1773–1780.
 20. Vickerman, J.C. *Surface Analysis. The Principal Techniques*; John Wiley & Sons Ltd: Baffins Lane, Chichester, England, 1997.
 21. Ratner, B.D. Surface properties and surface characterization of materials. In *Biomaterials Science. An Introduction to Materials in Medicine*, 2nd Ed.; Ratner, B.D., Hoffman, A.S., Schoen, F.J., Lemons, J.E., Eds.; Elsevier Academic Press: San Diego, 2004; 40–59.
 22. Hauch, K.D. Microscopy for Biomaterials Science. In *Biomaterials Science. An Introduction to Materials in Medicine*, 2nd Ed.; Ratner, B.D., Hoffman, A.S., Schoen, F.J., Lemons, J.E., Eds.; Elsevier Academic Press: San Diego, 2004; 396–411.
 23. Magnani, A.; Barbucci, R. Fourier transform attenuated total reflection infrared spectroscopy (ATR/FT-IR): applications to proteins adsorption studies. In *Test Procedures for the Blood Compatibility of Biomaterials*; Dawids, S., Ed.; Kluwer Academic Publisher: The Netherlands, 1993; 171–184.
 24. Barbucci, R.; Magnani, A. Conformation of human plasma proteins at polymer surfaces: the effectiveness of surface heparinization. *Biomaterials* **1994**, *15* (12), 955–962.
 25. Belu, A.M.; Graham, D.J.; Castner, D.G. Time-of-flight secondary ion mass spectrometry: techniques and applications for the characterization of biomaterials surfaces. *Biomaterials* **2003**, *24*, 3635–3653.
 26. Norde, W.; Lyklema, J. Why proteins prefer surfaces. *J. Biomat. Sci. Polym. Ed.* **1991**, *2*, 183–202.
 27. Gray, J.J. The interaction of protein with solid surfaces. *Curr. Opin. Struct. Biol.* **2004**, *14*, 110–115.
 28. Magnani, A.; Peluso, G.; Margarucci, S.; Chittur, K.K. Protein adsorption and cellular/tissue interactions. In *Integrated Biomaterials Sciences*; Barbucci, R., Ed.; Kluwer Academic/Plenum Publishers: New York, 2002; 669–689.
 29. Horbett, T.A.; Brash, J.L. Proteins at interfaces: an overview. In *Proteins at Interfaces II: Fundamentals and Applications*, ACS Symposium Series 602; Horbett, T.A., Brash, J.L., Eds.; American Chemical Society: Washington, DC, 1995, 1–25.
 30. Slack, S.M.; Horbett, T.A. The Vroman effect: a critical review. In *Proteins at Interfaces II: Fundamentals and Applications*; Horbett T.A., Brash, J.L., Eds.; ACS Symposium Series 602; American Chemical Society: Washington, DC, 1995, 112–128.
 31. Harris, J.M. Introduction to biotechnical and biomedical applications of poly(ethylene glycol). In *Poly(Ethylene Glycol) Chemistry: Biotechnical and Biomedical Applications*; Harris, J.M., Ed.; Plenum Press: New York, 1992; 1–14.
 32. Chinn, J.A.; Posso, S.E.; Horbett, T.A.; Ratner, B.D. Post-adsorptive transitions in fibrinogen adsorbed to polyurethanes: changes in antibody binding and sodium dodecyl sulphate elutability. *J. Biomed. Mater. Res.* **1992**, *26*, 757–778.
 33. Lu, D.R.; Park, K. Protein adsorption on polymer surfaces: calculation of adsorption energies. *J. Biomat. Sci. Polym. Ed.* **1990**, *1*, 243–260.
 34. Barbucci, R.; Lamponi, S.; Magnani, A.; Fibrinogen conformation and platelet reactivity in relation to material-blood interaction: effect of stress hormones. *Biomacromolecules* **2003**, *4* (6), 1506–1513.

Hemodynamics, Macrocirculatory

Baruch B. Lieber

Departments of Biomedical Engineering and Radiology, University of Miami, Coral Gables, Florida, U.S.A.

Chander Sadasivan

Hemodynamics Laboratory, Department of Biomedical Engineering, University of Miami, Coral Gables, Florida, U.S.A.

INTRODUCTION

The vascular network efficiently transports blood to and from the capillaries in organ systems where the transfer of oxygen and metabolic waste products occurs. Although the arterial tree consists of successive branches of decreasing diameter, the total cross-sectional area at each branching generation increases because of the larger number of vessels present in each generation. Newton's concept of fluid viscosity is a crucial parameter, and its exclusion from hemodynamic models gives inaccurate results. The rheological properties of both blood and the vessel wall must be incorporated into any hemodynamic theory that aims to adequately predict physiological flow. The simpler mathematical models assume steady flow and, although blood flow is pulsatile, useful information can be gathered from these steady flow models. Under certain simplifying assumptions, models of pulsatile flow employing either the linearized or nonlinear Navier–Stokes equations have been solved along with the governing equations of wall motion. The flow velocity patterns derived from such models are in good agreement with measured values. Owing to pulsatility of flow, pressure pulses are generated in the vessel walls, which propagate down the arterial tree as waves. These waves are attenuated as they progress down the length of a vessel and are reflected at branching points or locations where impedance mismatch occurs.

The structure of the venous system is similar to the arterial system, except for the fact that it is under much lower pressure. Negative transmural pressures may develop across the venous vessel walls, which may lead to their collapse. Even under normal operating conditions, the veins are partially collapsed, and the flow through such tubes of noncircular cross-sectional area is different from that in the arteries. Flow patterns have been implicated in the development of many vascular diseases such as atherosclerosis and aneurysms. Disease states may alter the local hemodynamics such as to escalate the deleterious effects.

PROPERTIES OF THE CIRCULATION

Although the human circulatory system has been of interest to man for many centuries, William Harvey's treatise on the subject (*Anatomical Exercise on the Mode of the Heart and Blood in Animals*, published in 1628) is generally accepted as one of the earlier milestones toward a scientific study of cardiovascular physiology.^[1,2] Contrary to the prevailing notion at that time, Harvey correctly inferred from surgical observations that blood constantly recirculates within the body, moving from the heart to the arteries and back to the heart from the veins. Stephen Hales, who reported his various investigations on circulatory function in his *Haemastatics* (published in 1733), is also regarded as a pioneer of the field of hemodynamics.^[1,2]

Application of fluid mechanics principles and analysis requires information about the basic geometrical environment in which such fluid motion takes place. It is therefore essential to have quantitative knowledge of dimensions such as the lengths or cross-sectional areas of blood vessels at various parts of the circulation. Many studies have been conducted to accumulate such information, and Table 1 lists the geometrical parameters of the canine circulation gathered from such studies. The aorta, which is the prominent trunk of the vascular network, displays a significant taper, terminating with a cross section that is about one-sixth of that at its origin.^[1] Because the number of vessels at any given generation of branching in the general arterial tree is much larger than the number at a previous generation, the total cross-sectional area of vessels increases as one moves away from the heart. Between the secondary branches such as the femoral and common carotid arteries and the vessels of 100 μm diameter, the total cross-sectional area increases three- to fivefold.^[1] The network of blood vessels that carry blood from the heart to the lung and back is called the pulmonary circulation, which contains 12% of the total volume of blood. The system of vessels transporting blood to all the other tissue (excluding the heart) is called the systemic circulation, and it carries 83% of the total blood volume.

Table 1 Model of vascular dimensions in a 20-kg dog

Class	Vessels	Mean diameter (mm)	Number of vessels	Mean length (mm)	Total cross-section (cm ²)	Total blood volume (ml)	Percentage of total volume
<i>Systemic</i>							
1	Aorta	(19–4.5)	1	(2.8–0.2)	60		
2	Arteries	4.000	40	150.0	5.0	75	
3	Arteries	1.300	500	45.0	6.6	30	
4	Arteries	0.450	6,000	13.5	9.5	13	11%
5	Arteries	0.150	110,000	4.0	19.4	8	
6	Arterioles	0.050	2.8×10^6	1.2	55.0	7	
7	Capillaries	0.008	2.7×10^9	0.65	1357.0	88	5%
8	Venules	0.100	1.0×10^7	1.6	785.4	126	
9	Veins	0.280	660,000	4.8	406.4	196	
10	Veins	0.700	40,000	13.5	154.0	208	
11	Veins	1.800	2,100	45.0	53.4	240	
12	Veins	4.500	110	150.0	17.5	263	67%
13	Venae cavae	(5–14)	2	(0.2–1.5)	92		
Total	1406						
<i>Pulmonary</i>							
1	Main artery	16.00	1	28.0	2.0	6	
2	Arteries	4.000	20	10.0	2.5	25	3%
3	Arteries	1.000	1,550	14.0	12.2	17	
4	Arterioles	0.100	1.5×10^6	0.7	120.0	8	
5	Capillaries	0.008	2.7×10^9	0.5	1,357.0	68	4%
6	Venules	0.110	2.0×10^6	0.7	190.0	13	
7	Veins	1.100	1,650	14.0	15.7	22	
8	Veins	4.200	25	100.0		35	5%
9	Main veins	8.000	4	30.0	6		
Total	200						
<i>Heart</i>							
Atria	2	30					
Ventricles	2	54	5%				
Total	84						
Total circulation	1690	100%					

(From Ref. [1].)

VISCOSITY OF BLOOD

A fluid is defined as a substance that deforms continuously when subjected to a shearing force.^[3] Consider two parallel plates separated by a distance h that contain fluid in between them. A tangential force F is exerted on the upper plate while keeping the bottom plate stationary such that the upper plate moves with a constant velocity of u cm/sec relative to the lower plate. Then, the fluid near the top plate will move at velocity

u while the fluid near the bottom plate will be stationary, i.e., a velocity gradient u/h is set up in the fluid between the plates. This velocity gradient, also represented by du/dy , is called the rate of shear. If the area of the plates is A , then the shearing stress is the shearing force divided by the area (F/A). Fluids in which the shear stress is directly proportional to the shear rate are called Newtonian fluids.

$$\tau = \mu \dot{\gamma} \quad (1)$$

where, τ is the shear stress (N/m^2), $\dot{\gamma}$ is the shear rate (sec^{-1}), and μ is the constant of proportionality called the coefficient of viscosity. The units of μ are N sec/m^2 , which is also referred to as Poise. Viscosity, referred to by Sir Isaac Newton as a “lack of slipperiness,” is a fundamental property of fluids that represents a kind of internal friction between adjacent fluid layers.^[1] Non-Newtonian fluids are those in which the shear stress varies nonlinearly with the shear rate, and in such cases, the ratio of shear stress to shear rate at any given shear rate is called the apparent viscosity. Certain fluids initially resist deformation to increasing shear stress, but eventually deform at a particular stress value, i.e., they have a yield stress. Some fluids that have a yield stress and a subsequent nonlinear relationship between τ and $\dot{\gamma}$ can be described by empirical relationships such as Casson’s equation

$$\sqrt{\tau} = \sqrt{\tau_y} + k\sqrt{\dot{\gamma}} \quad (2)$$

where, τ_y is the yield stress and k represents plastic viscosity. Blood is a complex fluid consisting of formed elements suspended in plasma. The formed elements consist of 95% red blood cells, 4.9% platelets, and about 0.13% white blood cells. When a sample of blood is centrifuged, the red blood cells collect at the bottom of the tube where they occupy approximately 45% by volume of the total blood sample. This volume percent of red blood cells in blood is called hematocrit. Human plasma has a density of about 1.035 gm/cm^3 , and its viscosity coefficient varies between 0.011 and 0.016 Poise.^[3] Tests of whole blood have shown that it more or less follows Casson’s equation, and at high shear rates (above 50/sec), the apparent viscosity approaches an asymptotic value of approximately 0.035 Poise.^[3] It is thus reasonable to assume that blood acts as a Newtonian fluid at high shear rates, which in general is true in the macrocirculation. Blood viscosity is dependent on temperature and plasma protein content, and increases with increasing hematocrit. The yield stress of blood increases with fibrinogen concentration and hematocrit.^[3]

VESSEL WALL PROPERTIES

To better understand the flow of blood, it is essential to understand the function of the biological conduits that carry it. Pulsatile flow in the circulation continually distends and contracts the vessel wall, and knowledge of the behavior of the vascular wall under such forces provides crucial information in analyzing hemodynamic states. The response of the vessel wall is related to its structure as defined by the cellular and extracellular components that comprise the wall material. All vessels in the macrocirculation have a similar structure and

possess more or less the same components, with only the proportion of these components changing at various parts of the circulation. The wall can be divided into three concentric layers, the intima, the media, and the outermost layer, the adventitia. The inner surface of the intima is lined with a single-cell thick layer of endothelial cells that in some sense serves as the biochemical mediator between blood and the vessel wall. One interesting fact about the endothelial cells is that they alter their morphology in accordance with the shear stress they experience, and they orient themselves in the direction of these stresses. In parts of the circulation, such as arterial bifurcations, where no uniform flow pattern exists, they are more round in structure, disoriented, and proliferated.^[4] The endothelial cell layer is supported by a thin layer of subendothelial connective tissue, and the outermost layer of the intima is the internal elastic lamina, which is a fenestrated sheet of elastic tissue. The media is composed of alternating layers of smooth muscle cells and elastic tissue with interspersed collagen fibrils. The elastic fibers, which are the cause of the elasticity of the vessel walls, contain elastin and microfibrillar components. The external elastic lamina is a layer of the media that demarcates the media from the adventitia. The adventitia itself is composed of dense fibrous tissues in which the nerves, blood vessels, and lymphatics that supply the vessel wall are embedded. Approximately 70% of the vascular wall is composed of water.^[1]

The arteries are classified as elastic or muscular, depending on the composition of the media. The aorta, the arteries that branch from the aortic arch, and the iliac arteries are called elastic as they contain well-defined alternating layers of elastic lamellae and smooth muscle cells.^[4] The elastic lamellae are connected to each other by thin strands of interlamellar elastic tissue, thus forming a three-dimensional (3-D) network through the vessel wall. Further down the arterial tree from these large vessels, however, the composition of medial layer contains a greater proportion of smooth muscle cells, and thus most of the arteries in the circulation are muscular. Smooth muscle fibers run obliquely between the elastic lamellae in elastic arteries, but are more circumferentially oriented in the muscular arteries. The vessel walls in the pulmonary circulation are thinner and less muscular than those in the systemic circulation, because the pulmonary circulation is a lower pressure system. Veins have a very thin intima and media with a small proportion of smooth muscle cells, while the adventitia is thick, containing abundant collagen.^[4] Fig. 1 shows the proportion of elastic, muscular, and fibrous tissues in various vessels. It can thus be appreciated that the vessel wall is far from being a homogenous and isotropic (same properties in all directions) material, as it contains various components in different orientations.

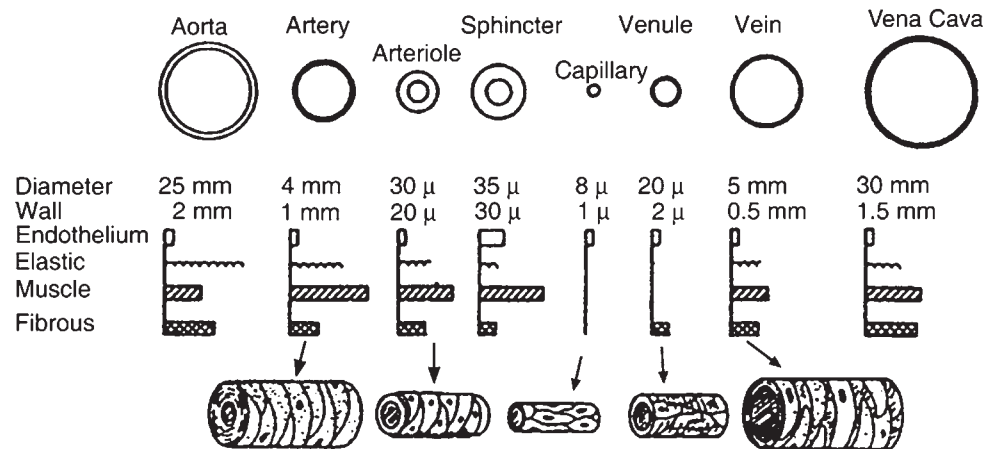


Fig. 1 Size and composition of blood vessel walls. (From Rushmer, R.F. *Organ Physiology: Structure and Function of the Cardiovascular System*; Saunders: Philadelphia, 1972; 135.)

The mechanical response of blood vessels to hemodynamic changes depends primarily on the orientation and interconnectivity of three components: elastin, collagen, and muscle.

The mechanical properties of materials are usually characterized by recording their stress–strain relationship. Those materials that deform when a force is applied on them, but instantaneously return to their original configuration when the force is removed are called elastic materials. Elastic materials that have a linear stress–strain relationship are called Hookean materials, and the ratio of stress to strain is called the elastic modulus. If the stress–strain relation is nonlinear, an incremental elastic modulus is quoted, which represents the slope of the stress–strain curve at a given length. The stress–strain curve of vascular tissue is nonlinear. Certain materials behave as a solid, possessing elasticity, strength, and stability of form, as well as a liquid, displaying characteristics such as flow depending on time, temperature, rate, and amount of loading. Such materials, which include blood vessel walls, that demonstrate both viscous and elastic behavior are called viscoelastic. If a blood vessel is subjected to a sudden increase in pressure (step increase), then it takes a finite amount of time for the radius of the vessel to reach a steady value. If subjected to a constant load, the vessel wall will continue to extend (called creep behavior), and if it is held at a constant distension, the stress decays over time (called stress relaxation). These are properties of all viscoelastic materials. Three variables, force, length, and thickness are measured to characterize the viscoelasticity of the vessel wall.^[1] To characterize the time-dependent, viscous properties of vessel walls, they are subjected to sinusoidal pressures, and the corresponding oscillations in radial distension are recorded. In contrast to an elastic material, viscoelastic materials will display a phase lag between these oscillations. This phase lag is a

measure of the viscosity of the material. Because of this wall viscosity, the stress–strain ratio for blood vessels is characterized by a complex viscoelastic modulus E_c .

$$E_c = |E_c|e^{i\phi} = |E_c| \cos \phi + i|E_c| \sin \phi \quad (3)$$

where, $|E_c|$ is the magnitude of the modulus, which represents the ratio of amplitudes of the pressure–radius curves, and ϕ is the phase lag between these two curves (positive ϕ represents pressure leading radial distension).

FLUID DYNAMICS CONSIDERATIONS

To develop an understanding of blood circulation, many fluid dynamic models have been constructed. It is useful, as is common in science, to initially consider the simplest of models, and then to progress to more complex ones only after the preceding stage has been suitably assimilated.

Steady Flow

One basic model developed in the mid-1850s, which nevertheless provides useful conclusions, is based on experiments done by Poiseuille on flow through small-bore tubes. The empirical expression (called Poiseuille's law) he derived relating the rate of flow through the tube to the pressure gradient across the tube length and the fourth power of the tube radius was independently arrived at by Hagen using theoretical formulations, and is now known as the Hagen–Poiseuille equation:

$$Q = \frac{\pi R^4 (P_1 - P_2)}{8\mu L} \quad (4)$$

where, Q represents the average volumetric flow rate, R is the radius of the tube, P_1 is the pressure at the beginning, and P_2 the pressure at the end of a tube of length L , and μ is the coefficient of viscosity. The equation is only valid for steady, fully developed (i.e., far away from the entrance), laminar flow of a Newtonian fluid through a straight, rigid cylindrical tube, and under those conditions, it also depicts that the velocity profile across the diameter of the tube will be parabolic in nature.

True Poiseuille flow is not established anywhere in the circulation,^[5] and it is helpful to evaluate the effects and validity of the assumptions made in the derivation to understand why this is so. Flow in the circulation is pulsatile and, thus the assumption of steady flow, i.e., the assumption that the fluid is not subjected to any acceleration or deceleration is not valid. Fluid dynamic models of pulsatile flow will be considered below. The simplification that the tube wall is rigid is invalidated as vessel walls are viscoelastic and distensible. The assumption of Newtonian behavior is reasonable in relatively large arteries as the shear rates in these arteries are high enough, and so is the condition that the fluid in contact with the tube wall does not move if the wall does not move (no-slip).

The derivation of the Poiseuille equation also assumes that each particle of liquid moves at a constant velocity in a path parallel to the cylinder wall, which implies that the fluid moves in a series of concentric laminae, with each successive lamina slipping against the viscous friction of the lamina outside it.^[1,2] When flow proceeds in such an organized and deterministic fashion, it is considered laminar. Under certain conditions, however, the orderly pattern of laminar flow is disturbed, and an irregular, random motion sets in. Such nondeterministic and unorganized flow is called turbulent flow, and portions of fluid move both axially and radially, forming eddies and vortices.^[1] At this point, it may be useful to introduce a nondimensional parameter that is widely used in fluid mechanics, called the Reynolds number (Re).

$$Re = \frac{\rho VD}{\mu} \quad (5)$$

where ρ represents the density of the fluid, V is the mean velocity of flow, D is a characteristic length (e.g., tube diameter), and μ is the coefficient of viscosity. The Re represents the ratio of inertial forces to viscous forces, i.e., if Re is large, the inertial forces predominate, whereas if Re is small, viscous forces will dominate fluid flow. The Re also provides an index to classify the flow as laminar or turbulent. In general, flow is said to become turbulent at Re above 2300, approximately. This is not a decisive threshold though, and nonlaminar flow may set in at Re below this

value.^[1] Some of the mean values of Re in the circulation of a 70 kg man are 1500 in the ascending aorta, 640 in the abdominal aorta, 200 in the femoral artery, 1400 in the inferior vena cava, and 1600 in the main pulmonary artery.^[1] Although none of these values are above 2300, these represent the averages, and the peak Re in the large vessels can be considerably higher (9400 in ascending aorta and 7800 in main pulmonary artery). Thus, transient turbulent flow may develop in the ascending aorta and the main pulmonary artery. Flow in all other vessels remains laminar.

The fact that vessels in the circulation are short, curved, and/or branched goes against the assumption of fully developed flow, and the effects of these factors have been investigated.^[1,4,5] When blood enters a vessel, the velocity profile immediately after the entrance is blunt (Fig. 2A), and there exists a steep velocity gradient near the wall owing to viscous friction while the inner core flow is considered inviscid. This outer portion of fluid, which contains the velocity gradient, is called the boundary layer. As the flow progresses down the vessel, the viscous effects propagate toward the axis of the tube, thereby retarding successive layers of fluid. To maintain a constant flow rate across the cross section, the core fluid accelerates. Flow is considered fully developed when the boundary layer reaches the axis of the vessel and a parabolic velocity profile is established. The length of vessel required to establish a fully developed flow is called the entrance length or inlet length, and the flow in this region is said to display entrance effects. The Poiseuille equation is not valid in the entrance region, and as most vessels are not long and straight, entrance effects tend to predominate in the circulation. Entrance effects also exist during start-up flow, which may be the case in the ascending aorta, as the flow almost stops during end of diastole and “starts” again during systole.^[4]

Blood vessels are rarely straight, and their curvature has a marked effect on flow owing to the generation of secondary flow motions that distort the velocity profiles deduced from Poiseuille flow.^[4,5] If it is assumed that the flow is fully developed before it enters the curve (Fig. 2B), then the fluid at the center has a higher velocity than that at the wall. Consequently, as it enters the bend, the central fluid has a higher centrifugal inertia and will resist following the curvature, more than the fluid closer to the wall. This results in the movement of the faster fluid toward the outer wall, where it displaces the slower moving fluid toward the inner wall. The resulting axial velocity profile is skewed toward the outer wall. This movement also sets up “secondary” flows (inset of Fig. 2B), over and above the “primary” axial flow.^[1,4] The flow in the curved part of the tube thus follows helical streamlines rotating in two different directions. It can be seen that the outer wall experiences higher wall shear stress. If

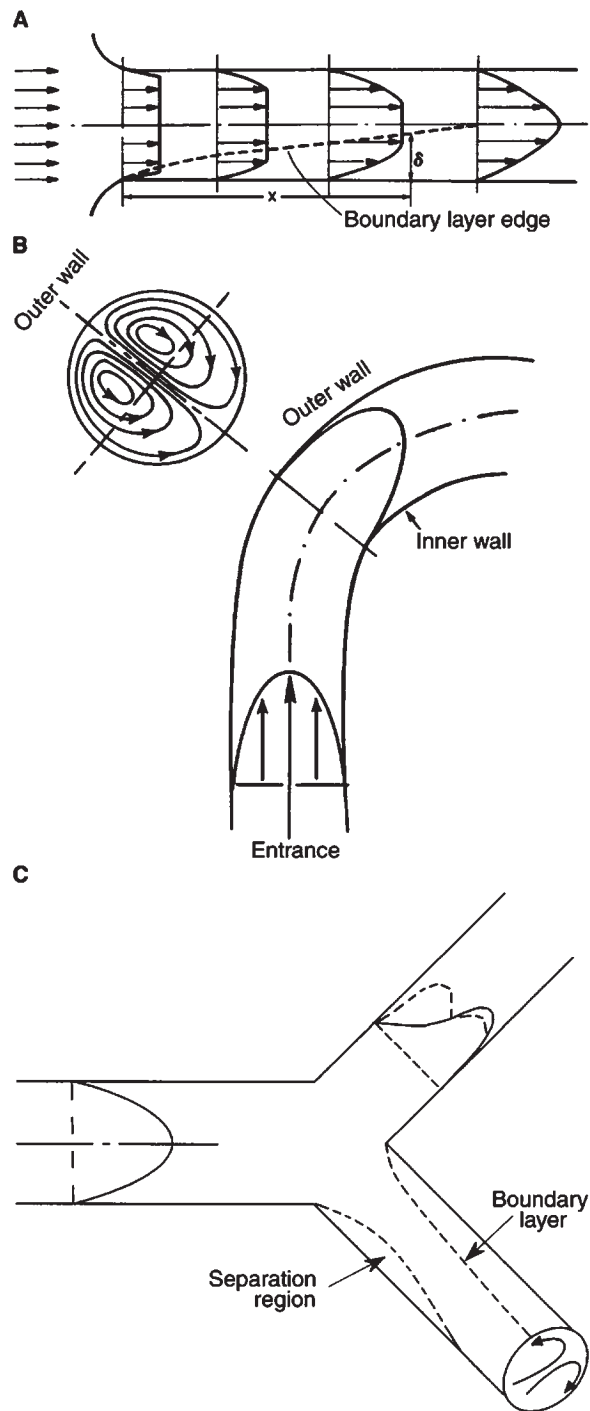


Fig. 2 The effect of complex geometries on flow. (A) Entrance length, δ is the boundary layer thickness at some distance x ; (B) steady flow in a curved tube, inset shows secondary flows; and (C) flow through a symmetric bifurcation, the solid line in the top branch is the velocity profile in plane, and the dotted line is the velocity profile in the perpendicular plane. [Panels (A) and (B) from Caro, C.J.; Pedley, T.J.; Schroter, R.C.; Seed, W.A. *The Mechanics of the Circulation*; Oxford University Press: Oxford, U.K., 1978. Panel (C) from Pedley, T.J. *Fluid Mechanics of Large Blood Vessels*; Cambridge University Press: Cambridge, U.K., 1980.]

the flow at the entrance to the curved tube is blunt, as in the case of the ascending aorta leading to the aortic arch, then all the fluid across the cross section has essentially the same inertia. In this case, the flow lines are not displaced outward by the centrifugal force, but the pressure at the outer wall increases. This increase in pressure at the outer wall must be balanced by a decrease in kinetic energy (Bernoulli's theorem, described below), and the higher velocities are now established at the inner wall.^[1,4,6] It may be noted that blood flow through entrance regions and curved segments is further complicated by the unsteadiness of flow in the circulation, possible noncircular vessel cross sections, and wall distensibility.^[3] In pulsatile flow, the entrance region also becomes transient.^[6] The steep velocity gradients in entrance regions and the secondary flows in curved tubes increase the rate of energy dissipation, which is manifested by an increase in pressure drop for the given flow rate.^[5]

Vessel branching is very common in the vasculature and for the most part, the cross-sectional areas of the daughter branches are smaller than that of the parent vessel, but the sum of their areas is larger than the parent vessel cross section. The same applies for the vessels that converge, i.e., the veins and the vertebral arteries combining into the basilar.^[1] The simplest form of branching is a symmetric bifurcation, as shown in Fig. 2C. The fluid dynamic effect of vessel branching is in some sense a combination of the effects of entrance length and curvature.^[5] Thus, the walls along the apex of the bifurcation (also called flow divider region) experience a higher wall shear stress, because the velocity profile is skewed toward the outer curvature. In contrast, the opposite (inner) walls experience low wall shear stresses, with possible flow separation and reversal. Flow separations are more likely if the distribution of flow rate in the daughter branches is unequal or if the bifurcation is not symmetrical, as in the case of a side branch coming off a parent vessel. As would be expected, secondary flows are also observed during experimental studies in bifurcation models.

Pulsatile Flow

The fact that the heart contracts during systole to eject blood into the vasculature and relaxes during diastole to fill implies that the flow through the circulation is time dependent. This pulsatility generates complex flow patterns that cannot be deduced from steady flow models. Hemodynamic models incorporating pulsatile flow have been developed, the most popular amongst them being the Womersley models. Womersley^[7] developed a series of mathematical models of blood flow that sequentially incorporated higher degrees of complexity, moving from the analysis of pulsatile flow in a

long rigid tube to that of flow in elastic tubes, and eventually considering the tethering of blood vessel walls by perivascular tissue and wall viscoelasticity. It may be helpful to examine his model of pulsatile flow in a rigid tube, to appreciate the process of incorporating pulsatility and its consequences.

The general mathematical equations describing fluid dynamics are called the Navier–Stokes equations, which are based on Newton’s second law (conservation of momentum).

$$\rho \vec{a} = \rho \vec{g} - \nabla p + \mu \nabla^2 \vec{v} \quad (6)$$

where, ρ is the density of the incompressible fluid, a is the acceleration, g is gravitational acceleration, p is the pressure, v is velocity, and μ is the viscosity. The term on the left represents the inertial forces, the first term on the right represents body forces (gravitational forces), the second term represents pressure forces, and the last term, viscous forces. For flow through a tube, it is convenient to express the equations in terms of cylindrical coordinates. If, for example, the direction along the axis of the tube is z , the radial direction is r , the circumferential direction is θ , t represents time, and the velocities in these three directions are v_z , v_r , and v_θ , respectively; the Navier–Stokes equation along the z direction can be expanded and written as

$$\begin{aligned} -\frac{\partial p}{\partial z} = & \rho \left(\frac{\partial v_z}{\partial t} + v_r \frac{\partial v_z}{\partial r} + \frac{v_\theta}{r} \frac{\partial v_z}{\partial \theta} + v_z \frac{\partial v_z}{\partial z} \right) \\ & - \mu \left(\frac{\partial^2 v_z}{\partial r^2} + \frac{1}{r} \frac{\partial v_z}{\partial r} + \frac{1}{r} \frac{\partial^2 v_z}{\partial \theta^2} + \frac{\partial^2 v_z}{\partial z^2} \right) \end{aligned} \quad (7)$$

where the body forces have been neglected. Similar equations can be written for the r and θ directions. These equations are the same as Eq. (6), rearranged such that they equate the decrease in pressure per unit length (left hand side) to the inertial forces (first term on right) minus the viscous forces (second term on right). The acceleration in the inertial term includes not only the rate of change of velocity with time (called local acceleration), but also the effect of convection of fluid particles (convective acceleration). The convective acceleration terms contain the products of velocities and their derivatives, and are therefore nonlinear. In addition to the Navier–Stokes equations, most fluid dynamic models also employ the principle of conservation of mass (the continuity equation). This equation is written in cylindrical coordinates as

$$\frac{\partial v_r}{\partial r} + \frac{v_r}{r} + \frac{1}{r} \frac{\partial v_\theta}{\partial \theta} + \frac{\partial v_z}{\partial z} = 0 \quad (8)$$

Eqs. (7) and (8) are the equations that govern fluid motion when large changes in temperature are negligible, such that conservation of energy need not be

employed. In general, the three components of velocity (v_z , v_r , and v_θ) and the pressure p are unknown for a given flow. The three Navier–Stokes equations (in three directions) and the equation of continuity give four equations with which to determine these four unknowns. After conducting an order of magnitude study, Womersley neglected the nonlinear terms, thereby facilitating a solution by considering the linearized Navier–Stokes equations. Many such linearized models apart from Womersley’s have been developed.^[1,4] For the first Womersley model, which analyzes the axisymmetric, laminar flow of an incompressible Newtonian fluid in a rigid cylindrical tube, the radial velocity component drops out of the equations (from axisymmetry) and so does the radial motion (from tube rigidity). The flow is irrotational and, hence, the circumferential component also drops out of the equations. Only the axial component of velocity remains, and the reduced continuity equation states that this velocity will be constant along the axis. The axial velocity is then a function of radius and time alone. For these assumptions, the Navier–Stokes equation in the z direction reduces to

$$-\frac{\partial p}{\partial z} = \rho \left(\frac{\partial v_z}{\partial t} \right) - \mu \left(\frac{\partial^2 v_z}{\partial r^2} + \frac{1}{r} \frac{\partial v_z}{\partial r} \right) \quad (9)$$

Further, it was assumed that the pressure gradient is a sinusoidal function of time only, expressed as

$$\frac{\partial p}{\partial z} = P e^{i\omega t} \quad (10)$$

where, P is a constant representing the amplitude of the pressure gradient pulse, and ω is the angular frequency. It follows then that the velocity will also be of similar form

$$v_z = w e^{i\omega t} \quad (11)$$

where, w , the magnitude of the velocity, is a function of r only and not of time. Substituting Eqs. (10) and (11) in Eq. (9),

$$-P e^{i\omega t} = i\omega \rho w e^{i\omega t} - \mu \left(\frac{\partial^2 w}{\partial r^2} + \frac{1}{r} \frac{\partial w}{\partial r} \right) e^{i\omega t} \quad (12)$$

and introducing a nondimensional radius $\hat{r} = r/R$, where R is the radius of the tube, the above equation becomes

$$\frac{\partial^2 w}{\partial \hat{r}^2} + \frac{1}{\hat{r}} \frac{\partial w}{\partial \hat{r}} - i\alpha^2 w = \frac{PR^2}{\mu} \quad (13)$$

where, $\alpha^2 = R^2 \omega \rho / \mu$. α is an important nondimensional parameter in pulsatile flow and is now known as the

Womersley parameter. Similar to the Re , it represents the ratio of transient inertial forces to viscous forces and is also referred to as the unsteady Re or unsteadiness parameter. For small values of α , the inertial forces are negligible and the viscous forces predominate. As α increases, the inertial forces are dominant, which implies that most of the fluid can be regarded as inviscid, and the velocity profile across the tube will be blunt except for the regions near the wall, which will have a steep velocity gradient. Some values of α in the circulation are: 21 in the ascending aorta, 12 in the abdominal aorta, 4 in the femoral artery, 17 in the inferior vena cava, and 20 in the main pulmonary artery^[1] under resting conditions. It may be mentioned in passing that in fluid dynamics experiments, scaled models are frequently employed, because they are either easier to work with or allow measurements with higher resolution. The correspondence of the flow in these scaled models is validated by matching up the nondimensional parameters (called dynamic similarity). For pulsatile flow studies of blood flow, both the Re and Womersley number, as well as the shape of the pulsatile waveform, must be matched.

Given boundary conditions from axisymmetry and no-slip at the wall, Eq. (13) can be solved for the axial velocity component.

$$v_z = \frac{PR^2}{i\mu\alpha^2} \left\{ 1 - \frac{J_0(i^{3/2}\alpha\hat{r})}{J_0(i^{3/2}\alpha)} \right\} \quad (14)$$

The instantaneous flow rate Q at any given cross section can then be obtained by integrating this expression

$$Q = 2\pi R^2 \int_0^1 w\hat{r}d\hat{r} \\ = \frac{\pi R^2 P}{i\omega\rho} \left\{ 1 - \frac{2J_1(i^{3/2}\alpha)}{i^{3/2}\alpha J_0(i^{3/2}\alpha)} \right\} e^{i\omega t} \quad (15)$$

where, J_0 and J_1 are Bessel functions of order zero and one, respectively. Bessel functions are standard solutions to certain forms of differential equations [such as Eq. (13)] that define many problems involving cylindrical or spherical coordinates such as heat conduction in a cylindrical object or modes of vibration of a thin circular membrane.

The advantage of assuming a sinusoidal pressure gradient during the derivation is that any pressure gradient curve can be expressed as a summation of sinusoidal components (Fourier series expansion). Then, velocity profiles and instantaneous flow rates can be obtained for each term in the Fourier series expansion (harmonics) from Eq. (14) [Eq. (15) for flow], and these can be added up to give the complete axial velocity profile across the tube, or the

instantaneous flow rate at any cross section. Fig. 3A compares the instantaneous flow rate calculated using this model to that measured in a canine femoral artery, by following the motion of a gas bubble in the artery

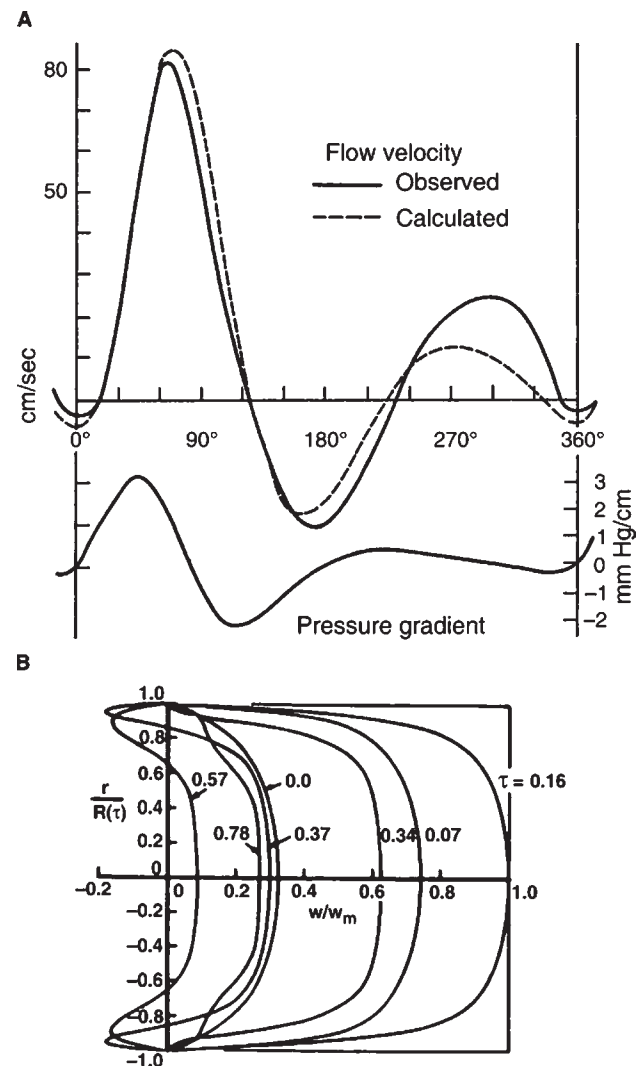


Fig. 3 Flow and velocity computations from mathematical models of pulsatile flow. (A) Measured flow rates in a canine femoral artery compared to the values calculated from Womersley's initial model. The bottom graph is the pressure gradient curve derived from measurements of pressure. (B) Axial velocity profiles in a canine descending aorta computed by nonlinear theory from measurements of pressure difference signals. The radial distance (r) was normalized by the inside radius [$R(\tau)$], where τ was the normalized time (t/T). T represents the time period of flow and w_m was the maximum centerline velocity measured with hot-film anemometry. The axial velocity w was normalized by w_m . Backflow can be noticed at $\tau = 0.37$ and 0.57 . [Pane (A) from Ref. [2]. Pane (B) from Ling, S.C.; Atabek, H.B.; Letzing, W.G.; Patel, D.J. Nonlinear analysis of aortic flow in living dogs. *Circ. Res.* 1973, 33 (2), 198–212.]

by means of high-speed cinematography. The pressure gradient that was used in the solution was calculated from measurements of pressure.^[2] As can be noted, the two graphs are in good agreement.

If the walls were assumed to be distensible, i.e., if elastic or viscoelastic wall material is assumed, then the governing equations of motion for the wall must be developed, and they must be solved in concert with the equations of fluid motion. Such models allow the analysis of pulse waves, which will be discussed in the next section. Theories where the nonlinear components of the Navier–Stokes equations are included in the problem definition have also been developed,^[4] although oscillatory flow components calculated from such models have been found to differ only slightly from linear models.^[1] The linear models, however, overestimate the steady flow term.^[1] Axial velocity profiles calculated by a nonlinear model based on measurements of pressure gradient (pressure difference over a finite length) in a canine descending aorta are shown in Fig. 3B.

Arterial Pulse Waves

The ventricle ejects approximately 85 ml of blood during each systole^[1] into the proximal aorta, which distends that portion of the vessel. During diastole, the elastic recoil of the aorta pushes the blood into the next distal section, which in turns distends and recoils, and thus a pressure wave travels down the arterial tree. The observation by Hales that this process is similar to a compression chamber (“Windkessel” in German), and the subsequent mathematical expression of this observation led to the development of the Windkessel theory.^[1] This theory considers the arterial tree as an elastic storage chamber into which ventricular outflow enters, and the outflow of this chamber faces a resistance and pressure gradient derived from Poiseuille flow. Equating the net rate of flow into the chamber with the rate of storage, an expression of pressure within the chamber and the ventricular stroke volume can be calculated. Although the Windkessel theory is useful for calculations of stroke volume, its validity in determining pressure pulse curves is generally not accepted today, primarily because the theory assumes instantaneous transmission of the pulse. In reality, the pressure pulse travels with a finite speed through the arterial wall, with the speed of transmission being dependent on the vessel wall properties and the interaction between the vessel wall and the flow of blood within.^[3]

Pulse wave propagation in a conduit carrying inviscid fluid was investigated by Thomas Young in about 1808, and was given relevant mathematical expression by Moens and Korteweg in 1878. The

expression for the velocity of propagation (also called wave velocity) in a long, thin-walled, elastic tube in which an incompressible, inviscid fluid is flowing is now called the Moens–Korteweg relationship.

$$c_0 = \sqrt{\frac{hE}{2R\rho}} \quad (16)$$

where, c_0 is the wave velocity, ρ is fluid density, R is the tube radius, h is the thickness of the wall ($h \ll R$), and E is the incremental circumferential elastic modulus.

Womersley^[7] derived a complex expression for the wave velocity, considering the effect of fluid viscosity in a freely moving elastic tube.

$$\frac{c}{c_0} = \sqrt{2(1 - \Phi)} \left[1 + k/2 - \Phi(k/2 + \sigma - 1/4) \right. \\ \left. \pm \sqrt{\left\{ 1 + k/2 - \Phi(k/2 + \sigma - 1/4) \right\}^2 + (1 - \sigma^2)(\Phi^2 - (1 - 2k)\Phi - 2k)} \right]^{-1/2} \quad (17)$$

where, $\Phi = \frac{2J_1(\alpha i^{3/2})}{\alpha i^{3/2} J_0(\alpha i^{3/2})}$, and $k = \frac{h\rho}{R\rho_0}$.

Here, c_0 is the wave velocity for inviscid flow [Eq. (16)], h is the wall thickness, R is the vessel radius, ρ is the density of the wall material, ρ_0 is the fluid density, and α is the Womersley parameter. The complex wave velocity c describes both speed of propagation (pulse velocity) and attenuation.^[1] The estimates for pressure wave velocity in the human vasculature are 5 m/sec in the ascending aorta, 8 m/sec in the femoral artery, 12 m/sec in the popliteal artery, and 7.5 m/sec in the carotid.^[1] Pulse waves are attenuated as they travel down a vessel wall and Womersley described this attenuation as

$$e^{2\pi(z/\lambda)(\text{imag}(c_0/c)/\text{real}(c_0/c))} \quad (18)$$

where, $\lambda = 2\pi c/\omega$ is the wave length. The attenuation in terms of the transmission per wavelength is therefore

$$e^{2\pi(\text{imag}(c_0/c)/\text{real}(c_0/c))} \quad (19)$$

A wave is thus attenuated by this amount, after traveling a distance of one wavelength.

Expressions for resistance and impedance of the vascular system can be written by electrical analogy. For steady flow,

$$\Omega = \frac{\Delta P}{Q} \quad (20)$$

where, Ω is the vascular resistance, ΔP is the mean pressure gradient, and Q is the mean flow. The vascular

resistance provides a measure of the extent by which the system resists flowing under the applied pressure gradient.^[1] By similar reasoning, for pulsatile flow,^[8]

$$Z_n = \frac{\bar{P}_n}{\bar{Q}_n} = \frac{|P_n|e^{i(\omega_n t + \phi_n)}}{|Q_n|e^{i(\omega_n t + \varphi_n)}} \quad (21)$$

where, Z_n represents vascular impedance that is dependent on frequency, and the subscript n refers to the n th harmonic of the corresponding waveform. The n th harmonic of the vascular impedance is then

$$Z_n = |Z_n|e^{i\theta_n} \quad (22)$$

where, the magnitude of the impedance $|Z_n| = \frac{|P_n|}{|Q_n|}$, and the phase lag $\theta_n = \phi_n - \varphi_n$

Three kinds of vascular impedance are relevant in physiologic flow: a) characteristic impedance is the ratio of oscillatory pressure to oscillatory flow, when there are no reflected waves present in the vessel segment being considered; b) longitudinal impedance is the ratio of oscillatory pressure gradient to oscillatory flow; and c) input impedance of a region of the circulation is represented by the ratio of one harmonic of the pressure at the input to the region divided by the corresponding harmonic of flow [Eq. (21)]. The input impedance of the femoral artery, for example, is the impedance of the entire vasculature supplied by it. The characteristic impedance of the human aorta is of the order of 50 dyn sec/cm⁵, and that of the femoral artery is 7000 dyn sec/cm⁵.^[1]

If pressure waves at various distances along the arterial tree are measured, it is found that the amplitude of the pressure wave actually increases down the arterial length (Fig. 4). This is owing to wave reflections that occur at branching points, or essentially, at any point on a vessel where there is a change in the impedance. It can also be noted from the figure that the amplitude of the flow velocity pulse decays, eventually terminating in steady flow in the finer vessels.

If the vascular network has the same impedance at each point (impedance matching), then there will be no wave reflections, and the wave energy will be transmitted efficiently.^[9] If we consider a wave traveling down an artery that ends in a bifurcation, a part of this wave will be reflected back upon reaching the bifurcation point. The other part of this wave energy will be transmitted into the two daughter branches. If the parent tube has a characteristic impedance of Z_0 , and the daughter branches have impedances Z_1 and Z_2 , then perfect impedance matching requires

$$\frac{1}{Z_0} = \frac{1}{Z_1} + \frac{1}{Z_2} \quad (23)$$

As was mentioned earlier, perfect impedance matching is not achieved in the circulation and wave

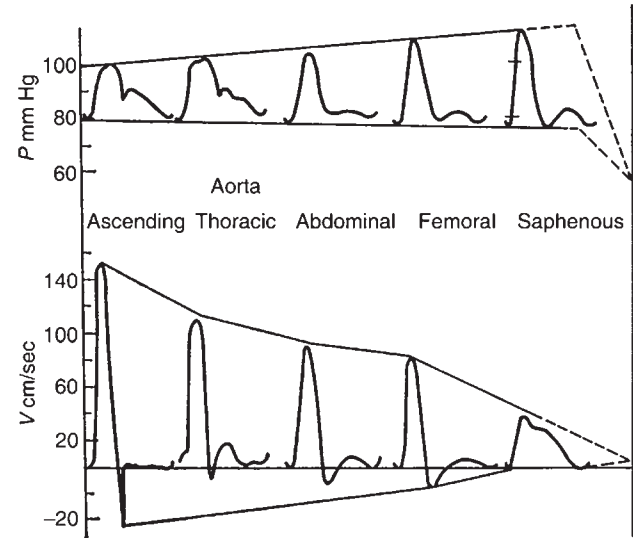


Fig. 4 A series of pressure and flow velocity pulse waves recorded along the arterial tree. The change in shape that occurs can be noted. The oscillatory pressure starts to decrease after the saphenous artery, which is represented by the dotted lines. The corresponding dotted lines in the flow pulses represent the continued decrease to steady flow in the microcirculation. (From Ref. [2].)

reflections occur. The reflection coefficient is given as

$$\mathfrak{R} = \frac{Z_0^{-1} - (Z_1^{-1} + Z_2^{-1})}{Z_0^{-1} + (Z_1^{-1} + Z_2^{-1})} \quad (24)$$

The pressure (or flow) wave recorded at any point in the network can be regarded as the sum of a forward traveling (antegrade) wave and a backward traveling (retrograde) wave. It is possible to mathematically extract the antegrade and retrograde components from such measurements.^[8] It should be noted that although the mode of wave propagation that is usually considered (axially traveling waves of radial distension) is the most physiologically relevant, other modes do exist. These include: a) longitudinal mode, where the vessel wall compresses and elongates in the longitudinal direction as the wave propagates; b) torsional mode, involving twisting motions of the vessel wall; and c) flexural mode, where the wave propagates by lateral displacements of the tube.^[1]

Venous Hemodynamics

The reason that the veins are considered storage reservoirs and capacitance vessels can be appreciated from Table 1 as they contain approximately 70% of the systemic blood volume. During blood loss, it is

H

venous blood volume and not arterial blood volume that decreases to maintain adequate arterial perfusion pressure.^[8] The pressure in a vein is considerably lower than the pressure in an artery at the same location, and venous pressures rarely exceed approximately 15 mmHg. In general, the flow through a vessel is considered a unique function of the pressure gradient across the tube. The pressure outside the vessels in our body is approximately atmospheric, except possibly for the vessels in the chest. The difference between the inner pressure and outer pressure is approximately 3 mmHg at the exit of the vena cavae.^[5] If we then consider veins above the heart, or the veins in an arm raised up, this pressure difference may become negative, thus collapsing the vein. This situation does not arise in arteries, because the arterial system is highly pressurized and the pressure difference stays positive, keeping arteries patent. It thus becomes important to consider transmural pressures (difference between the internal and external pressures) when considering pressure–flow relationships in veins. The collapse of venous walls is studied by observing pressure–radius relationships, as shown in Fig. 5A. This relationship in the excised segment of canine vena cava shows that collapse can occur to extremely small cross-sectional areas under negative transmural pressure, thereby dramatically affecting the flow. Two other factors effectively reduce the transmural pressure in veins: the “skeletal muscle pump,” where voluntary contraction of skeletal muscles outside the veins can push blood toward the heart, and the presence of valves in the mid-sized veins that serve to prevent backflow and bear the weight of columns of venous blood.^[1,6]

The Starling resistor (Fig. 5B) is possibly the most popular experimental device to study flow through collapsible tubes. It consists of a segment of collapsible tube whose ends are connected to rigid tubes. The segment itself is contained within a chamber whose pressure (p_e) can be varied. The chamber pressure represents the external pressure, and with this device, the variations in flow through the collapsible tube because of alterations in pressure-differences (e.g., $p_1 - p_2$ or $p_1 - p_e$) can be studied. One interesting outcome of such experiments is that for certain pressure differences, when $p_1 > p_e > p_2$, the flow through the tube might be completely independent of the outlet pressure p_2 , varying only with $p_1 - p_e$. This is called the Starling resistor phenomenon, sluicing, flow limitation, or waterfall effect. This effect has been reported in venous blood flow from extrathoracic to intrathoracic veins in dogs.^[4]

Pressure and flow pulses have also been studied in veins. The pressure in the right atrium of man is about 2 mmHg and the amplitude of pressure fluctuations is about 5 mmHg, whereas the pressure in the venules is about 18 mmHg with the amplitude of fluctuations being around 0.3 mmHg.^[6] The velocity waves in

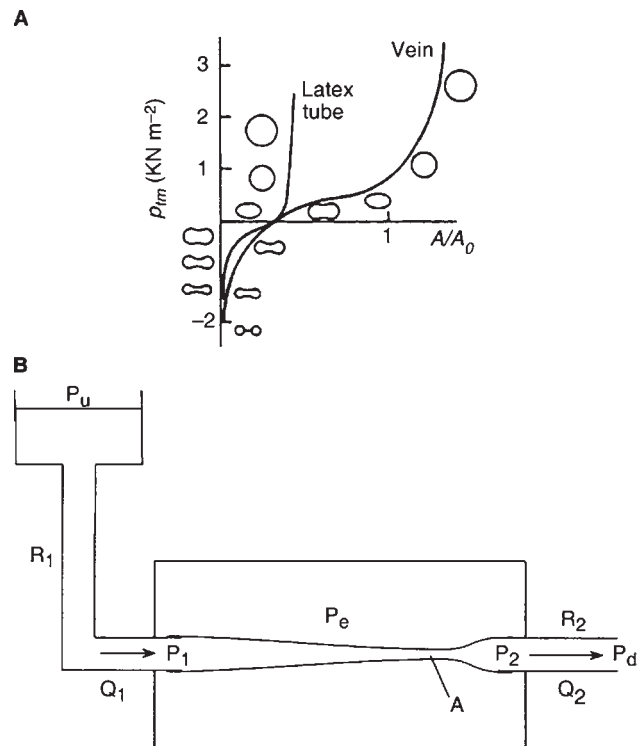


Fig. 5 (A) Transmural pressure (p_{tm}) as a function of area (A) normalized by the area at zero transmural pressure (A_0) of an excised segment of canine abdominal vena cava. The curve for a latex tube is also shown for comparison. (B) Sketch of a Starling resistor. R , p , and Q represent resistance, pressure, and flow, respectively. [Pane (A) from Moreno, A.H.; Katz, A.I.; Gold, L.D.; Reddy, R.V. *Mechanics of distension of dog veins and other very thin-walled tubular structures*. *Circ. Res.* **1970**, 27 (6), 1069–1080. Pane (B) from Ref. [5].]

human vena cava show two main oscillations that are out of phase with the pressure oscillations in the right atrium.^[6] In an experiment conducted by applying pressure waves to venous segments to study their propagation, it was found that the wave speed at any given transmural pressure was independent of the frequency of oscillation, implying that venous wave propagation is essentially nondispersive. The attenuation per wavelength was estimated to be in the range of about 65–90%.^[6]

PATHOLOGY

Abdominal aortic aneurysms are abnormal dilations of the abdominal aorta that carry a serious risk of rupture. Discrete parts of the arterial tree are predisposed to gathering atherosclerotic fatty deposits, thereby constricting these vessels. Blood flow through such sudden dilations or constrictions can be understood by a simple model of steady flow of an inviscid, incompressible fluid, called Bernoulli's theorem. This is

based on the principle of conservation of energy. The total energy per unit volume at any point may be given as the sum of the pressure energy, kinetic energy, and the gravitational energy. If we consider the cross section of the normal vessel wall as station 1, having mean velocity v_1 and area A_1 , and the cross section immediately after a sudden enlargement as station 2, having velocity v_2 and area A_2 , then the flow of blood (Q) must be the same at both these stations, according to the principle of conservation of mass. Therefore,

$$A_1 v_1 = Q = A_2 v_2 \quad (25)$$

As A_2 is larger than A_1 , it follows that v_2 is smaller than v_1 . Conservation of energy states that the total energies at these two stations must also be equal.

$$P_1 + 1/2\rho v_1^2 + \rho g h_1 = P_2 + 1/2\rho v_2^2 + \rho g h_2 \quad (26)$$

where, P represents the pressure and ρ , the density of the fluid; the second term represents the kinetic energy, and the third term is the hydrostatic energy. If both stations are at the same height, and as v_2 is less than v_1 , it follows that P_2 is greater than P_1 . In other words, the kinetic energy at station 1 is converted to pressure energy at station 2, and the flow velocity reduces. Owing to this pressure rise, the fluid decelerates, and being unable to follow the wall at the dilatation, the flow separates resulting in recirculating eddies. This recirculation region has biochemical implications as blood is essentially a suspension of formed elements in plasma, and the blood particles keep circulating in one place for a longer period of time (increased residence time), possibly causing adverse biochemical reactions there. The same conditions may apply at the exit of a stenosis, while the converse happens as flow enters a stenosis. The pressure energy is traded for kinetic energy and the flow velocity increases as it goes through the stenosed section.

The carotid sinus is a frequent site of atherosclerosis, and it is now thought that low and oscillating wall shear stresses that are produced at this location owing to complex flow patterns are an atherogenic factor. Hemodynamics has also been implicated in the growth and rupture of intracranial aneurysms, which are sacular abnormalities found mostly at the bifurcations of vessels in the brain. These are just a few of the examples that signify the importance of understanding the normal and abnormal states of hemodynamic values and flow patterns in various parts of the circulation.

CONCLUSIONS

Hemodynamic factors have been implicated in a variety of vascular diseases, and many models, both theoretical and experimental, have been developed by researchers to glean the complex flow variations that occur in

physiologic flow. Some of these steady and pulsatile flow models have been discussed here along with their results. Representative dimensions of the vasculature and the rheology of blood and the vessel wall have been mentioned. Arterial pulse waves and flow in collapsible veins have been briefly reviewed.

Although this article covers the fundamental theoretical concepts of hemodynamics, computational fluid dynamics and medical imaging have rapidly advanced our ability to calculate detailed 3-D flows in anatomically accurate, patient-specific geometries of the vasculature. Numerical computing has thus become sufficiently complex to be regarded as a separate field in itself. Any degree of accuracy in determining flow fields does not, however, mitigate the fact that the vascular system is biochemical in nature. Biochemical interactions have not yet been adequately incorporated into hemodynamic models, and this is the subject of current research. Needless to say, the description of hemodynamic concepts presented here is a very brief account of the present state of knowledge in the science, and the interested reader is directed to the references for further depth.

ARTICLES OF FURTHER INTEREST

Artificial Heart Fluid Dynamics: Positive Displacement Pumps; Bioreactors; Blood Vessel Mechanics

REFERENCES

1. Milnor, W.R. *Hemodynamics*; 2nd Ed.; Williams and Wilkins: Baltimore, MA, 1989.
2. McDonald, D.A. *Blood Flow in Arteries*, 2nd Ed.; The Williams and Wilkins Co.: Baltimore, MA, 1974.
3. Chandran, K.B. *Cardiovascular Biomechanics*; New York University Press: New York, NY, 1992.
4. Patel, D.J.; Vaishnav, R.N. *Basic Hemodynamics and Its Role in Disease Processes*; University Park Press: Baltimore, MA, 1980.
5. Pedley, T.J. Arterial and venous fluid dynamics. In *Cardiovascular Fluid Mechanics*; 1st Ed.; Pedrizzetti, G., Perktold, K., Eds.; Springer Wien: New York, 2004; 1–72.
6. Fung, Y.C. *Biodynamics: Circulation*; Springer-Verlag: New York, NY, 1984.
7. Womersley, J.R. *An Elastic Tube Theory of Pulse Transmission and Oscillatory Flow in Mammalian Arteries*; Wright Air Development Center Technical Report TR 56-614; U.S. Air Force: Ohio, 1957.
8. Li, J.K.-J. *Dynamics of the Vascular System*; Li J.K.-J., Series Ed.; Series on Bioengineering & Biomedical Engineering; World Scientific Publishing Co.: Singapore, 2004; Vol. 1.
9. Lieber, B.B. Arterial macrocirculatory hemodynamics. In *The Biomedical Engineering Handbook*; 2nd Ed.; Bronzino, J.D., Ed.; CRC Press: Boca Raton, FL, 2000; 30-1–30-10.



Heparin-Binding to Improve Biocompatibility

Rolf Larsson

Department of Oncology, Radiology and Clinical Immunology, University Hospital, Uppsala, Sweden

INTRODUCTION

Heparin is one of the most negatively charged biomolecules known to be present in the body. It is always bound to tissues, most abundantly to mast cells, as a proteoglycan in which heparin molecules are covalently linked to a protein core. Despite the fact that soluble heparin, detached from the protein core, has been used as an efficient anticoagulant for more than five decades, the true physiological role of heparin is still not fully clarified.^[1-4] In the mid-1970s, it was discovered that the anticoagulant mechanism of heparin involves a catalytic interaction with antithrombin.^[5]

The concept of binding heparin to an artificial surface emerged by serendipity in the early 1960s.^[6] A carbon surface that was sterilized by immersion in benzalkonium chloride was subsequently soaked in a heparin/saline solution as a measure of surgical precaution, and a benzalkonium-heparin complex was inadvertently deposited. This discovery led to a number of different heparin surfaces with major emphasis on easiness of preparation, production costs, and related technical aspects but without considering the biological mechanisms of heparin (examples given in the following text). Concomitantly with the unravelling of the biological mechanisms of the interaction between heparin and antithrombin, a new method for covalent immobilization of heparin was developed involving partial depolymerization of heparin to introduce terminal aldehyde groups to avoid compromising the antithrombin binding capacity.^[7] This approach formed the basis for the heparin surface now known as Carmeda Bioactive Surface (CBAS), which is widely published.^[8] Another permanent heparin surface built on a macromolecular conjugate of unfractionated heparin allowing the antithrombin binding capacity to be adjusted, incidentally also developed in Sweden, is known as Corline Heparin Surface, CHS.^[9-12,68] This article covers a brief overview of some important and relevant biological interactions of heparin with emphasis on immobilized heparin, which is followed by a catalogue of the most prominent heparin surfaces describing their chemical constitution and the rationale for their use. Finally, some clinical applications are exemplified.

THE CATALYTIC ROLE OF HEPARIN

Heparin exerts its anticoagulant effect via a unique pentasaccharide sequence normally present in one-third of a commercial preparation of heparin.^[13-15] This high affinity fraction holds a high specific anticoagulant activity (typically, 250 IU/mg), whereas the residual low affinity fraction displays very low activity (typically, 10–20 IU/mg). The pentasaccharide sequence in heparin, which may also be present in heparan sulphate, binds to antithrombin whereby a structural reorientation is induced to the effect that the binding site of antithrombin for a number of serine proteases (e.g., thrombin) is more readily accessible.^[16] Binding of heparin to antithrombin thus accomplishes a thousandfold increase of the inhibition rate without altering the stoichiometry.^[16] Once an antithrombin-protease complex has been formed, heparin acquires a lower binding affinity and heparin can depart from the complex to find a new antithrombin molecule to bind to.^[17,18] This explains the catalytic role of heparin, which may occur both with soluble heparin and with heparin adequately immobilized to a solid surface. Due to their strong anionic character, heparin and related sulphated glucosaminoglycans exert a number of additional biological interactions.^[1-4]

ANTICOAGULANT ACTIVITY OF THE VASCULAR ENDOTHELIUM

The endothelial lining on the natural vascular wall exposes specific antithrombin binding sites linked to the presence of heparan sulfate.^[20] The density differs among different types of vessels, such that capillaries expose 10–12 pmol/cm², whereas the corresponding figure for aortic tissue is much lower, typically 2–3 pmol/cm².^[21] Experiments on fresh vessel segments with intact endothelium, as well as with vessels in which the endothelial cell layer had been experimentally detached, have revealed that factor XIIa, the initiator of the contact activation system, is bound to both types of vascular linings, but is immediately inhibited only if the endothelium is intact.^[21] The sub-endothelial structures also readily attract and activate platelets^[22] and white blood cells (Fig. 1).^[23]

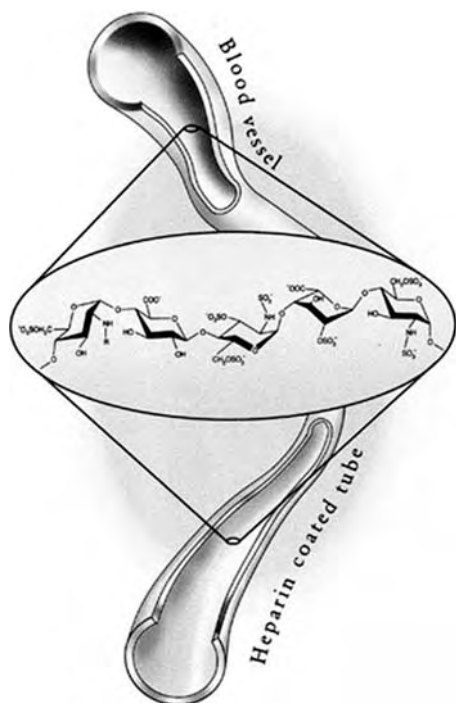


Fig. 1 An artist's view of a common denominator of the vascular endothelium and a heparin-modified surface: the antithrombin binding sequence.

BIOLOGICAL EFFECTS OF IMMOBILIZED HEPARIN

From a principal point of view, it is important to distinguish between effects induced by slow release of heparin from a surface and effects related to interactions with immobilized heparin. The former situation may merely be viewed as a rather complicated way of administering heparin locally for short-term use, whereas immobilized heparin may act as a catalyst over an extended period of time. The following discussion of biological effects related to heparin surfaces will emphasize interactions with immobilized heparin, confined to the interface between blood and a modified surface, rather than discussing possible effects that are likely to occur in the bulk phase of blood due to release of heparin. It is, however, difficult to compare the vast literature on the effects of heparin surfaces, due to the fact that the chemical nature and characteristics of the various heparin surfaces are often poorly described. The characteristics related to performance may differ quite extensively depending on the chemical constitution of the modified heparin surfaces.

Platelet Adhesion and Activation

Early investigations have established that platelet adhesion is dramatically reduced and virtually

eliminated at a heparin-coated surface,^[24,25] which has been verified in a number of later studies.^[9,12,26–30] It was demonstrated that fibrinogen did not adsorb to a heparin surface exposed to plasma, but if fibrinogen was adsorbed to the heparin surface from a pure fibrinogen solution prior to contact with blood, platelet adhesion promptly occurred.^[25] In a recent study, it was shown that activation of platelets, as indicated by release of β -thromboglobulin, decreased with increasing capacity of the heparin surface to bind antithrombin.^[12] Numerous reports have verified by, e.g., scanning electron spectroscopy that heparin surfaces are conspicuously free from deposits including platelets (Fig. 2). If platelet adhesion is not affected by surface modification with heparin, as reported in one study,^[31] the most likely explanation would be that the particular method used did not result in a coherent coating.

White Blood Cell Adhesion and Activation

Microscopic investigations, as illustrated in Fig. 2, imply that not only platelet adhesion but also adhesion of WBC is greatly reduced, which has been confirmed in several studies.^[28,32–34] In a clinical study of coronary bypass, oxygenators that were either unmodified or modified with the Corline Heparin Surface were investigated with respect to attachment of cells. The number of platelets, granulocytes, and monocytes retrieved in the eluates from the two different types of oxygenators were significantly reduced in the group with heparin-coated surfaces.^[11] Experimental and clinical studies have demonstrated less activation of WBC as indicated by reduced levels of markers for activation of WBC as, e.g., myeloperoxidase,^[10,35] lactoferrin,^[35] and the integrin CD11b.^[10]

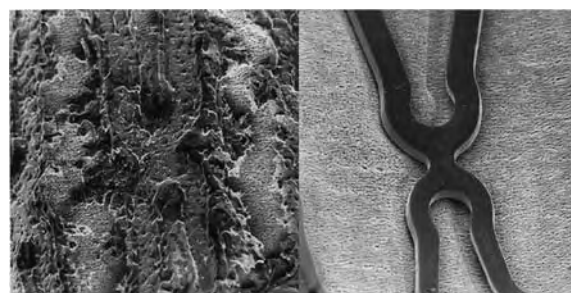


Fig. 2 Unmodified (left), and CHS-modified (right) Stent Graft after exposure to fresh human blood for one hour at 37°C. The specimens were gently rinsed with saline, fixed with glutaraldehyde and subsequently dried at CO₂ critical point, covered with gold, and finally examined by scanning electron microscopy (magnification $\times 50$).

H

Interaction with the Enzymatic Coagulation System

Early studies with an ionic heparin surface stabilized with glutaraldehyde demonstrated that immobilized heparin is able to bind and inhibit thrombin,^[36] but a sulphated surface devoid of the anticoagulant effect of heparin proved to be thrombogenic due to the fact that thrombin would still be bound but protected from inhibition.^[25,36] In the first report on the method of covalent end-point binding of heparin, it was reported that ordinary heparin could be irreversibly attached by ionic binding under certain conditions, resulting in firmly bound heparin, which, however, proved to be thrombogenic due to binding of thrombin in a form that could not be inhibited.^[7] These findings demonstrate that certain requirements have to be met in order to obtain the desired capacity for inhibition of thrombin. A number of studies have been published to show that the heparin surface prepared by end-point attachment of partially degraded heparin inhibits thrombin, coagulation factor Xa, and factor XIIIa (for review, see Ref. [37]). Also the heparin surface based on a macromolecular heparin conjugate has proved to reduce coagulation activation *in vitro* effectively based on tests with fresh human blood,^[9,12] and to prevent contact activation.^[38] In a clinical study on coronary bypass patients, it was demonstrated that activation of coagulation, which normally occurs during bypass despite systemic administration of heparin, could be reduced by applying the Corline Heparin Surface to the complete bypass system.^[10,11]

Contact Activation

The contact activating system is generally considered to be activated by negatively charged surfaces primarily by adsorption of factor XII, which will then be converted to the active enzyme FXIIa. The two other proenzymes of the contact activation system, factor XI and prokallekrein, together with the cofactor high molecular mass kininogen may then become activated leading to activation of coagulation and inflammatory reactions. Several studies have shown that FXII binds to a heparin surface, most likely due to its negative charge, but that any FXIIa that is formed is immediately inhibited, provided that the heparin surface displays adequate antithrombin binding capacity.^[39,40] If, however, the surface is devoid of antithrombin binding capacity, as was demonstrated by preparing a heparin surface with low affinity heparin,^[40] the adsorbed FXIIa could not be effectively inhibited. Again, we see an example of a surface that does contain heparin but does not qualify as a nonthrombogenic surface. As

pointed out previously, FXII is adsorbed to the vascular endothelium but FXIIa is rapidly inhibited as long as the endothelium remains intact. Any damage to the endothelium will render the surface incapable of inhibiting FXIIa, which implies that the whole contact activation system may be activated.^[21] There seems to be a close link between contact activation and activation of platelets as described in a recent study. Selective inhibition of contact activation induced reduced activation of platelets and also caused less thrombin generation. The latter effect could also be achieved by blocking the platelet receptor GPIIb/IIIa.^[41]

Complement and Inflammatory Response

The effects exerted by surface-attached heparin on complement and inflammatory response have been broadly investigated. The vast majority of studies are related to coronary bypass surgery. In general, heparin surfaces reduce complement activation^[35,42-46] and also minimize activation of granulocytes as evidenced by less release of lactoferrin and myeloperoxidase.^[35] In this context, it is interesting to note that selective blockade of complement results in considerable reduction of adhesion of granulocytes to foreign surfaces.^[47] In a recent publication, a link between chemical constitution of a heparin surface and complement activation was established. Activation of complement was shown to decrease as the antithrombin binding capacity increased.^[12]

Protein Interactions

Due to the strong anionic character and the presence of specific binding domains, heparin binds and interacts with a large number of peptides, proteins, lipids, etc., both in a specific and nonspecific manner.^[48] Antithrombin, a natural inhibitor of several serine proteases in the coagulation system, binds to a specific pentasaccharide domain in heparin, whereby the rate of inhibition is greatly accelerated.^[16] Surfaces modified with heparin using selective coupling methods have been shown to bind antithrombin, which is clearly related to functional integrity.^[12,40] Using plasmon resonance, it could be demonstrated that the binding of several proteins including antithrombin was strongly affected by the method used to immobilize heparin.^[49] A few, nonlimiting examples of additional peptides and proteins that bind to heparin are a variety of growth factors,^[50] vitronectin,^[51,52] laminin,^[53] lipoprotein lipase,^[54-56] complement factors,^[57-59] protamine,^[60] kininogens,^[61] coagulation factor XII,^[62] and tissue factor pathway inhibitor.^[63]

TYPES OF HEPARIN SURFACES

Carmeda Bioactive Surface (CBAS™)

In the first publication on this heparin surface based on single end-point covalent attachment of heparin, the priming layer onto which heparin was attached was composed of polyethyleneimine (PEI) irreversibly adsorbed to a sulphated polyethylene surface.^[7] Interestingly, two quite different priming layers could be accomplished simply by changing the pH of the PEI solution used during surface adsorption of PEI. At an acidic pH, the PEI layer proved to bind unmodified heparin only by electrostatic interaction firmly enough so that heparin could not be eluted from the surface layer. However, this particular heparin surface was found to be thrombogenic due to the fact that thrombin that was adsorbed to the surface appeared to be protected against inhibition. When the pH of the PEI solution was adjusted to 9, the ionic interaction with heparin was very weak and unmodified heparin was readily eluted from the surface. By using heparin that had been partially depolymerized by mild treatment with nitrous acid, thereby introducing terminal 2,5-anhydromannose residues exposing a reactive aldehyde, heparin could be covalently immobilized while retaining adequate mobility to avoid compromising the antithrombin binding sequence of heparin.^[7] Thrombin adsorbed to this surface with upright heparin proved to be effectively inhibited in the presence of antithrombin in sharp contrast to the surface with laying heparin (PEI at pH 4 followed by unmodified heparin). This discovery laid the foundation for what later came to be known as CBAS™. A priming layer based on sulphated polyethylene is, however, not very useful for practical and commercial purposes. Hence, the CBAS™ involves a different priming layer composed of a sandwich ending with PEI. The underlying layers are composed of alternating layers of PEI treated with glutaraldehyde and dextran sulphate, respectively.

Pharmacia HSM Intraocular Lens

Intraocular lenses (IOLs) made of poly(methyl methacrylate) are available with a covalent heparin coating (Heparin Surfaced Modified, HSM). This coating is a modification of the CBAS™ basically to improve the stability towards mechanical abrasion. Heparin is covalently linked to PEI, being irreversibly adsorbed to PMMA after prior treatment with ammonium persulphate. To achieve adequate thickness, the sequence of PEI and heparin is repeated. The HSM surface has proved to be stable, as demonstrated by the fact that there was no detectable loss of heparin

following two years implantation in the anterior eye chamber of rabbits,^[64] and to reduce inflammatory response.^[65,66]

Corline® Heparin Surface (CHS)

The most conspicuous component of this rather new heparin surface is a macromolecular conjugate composed of a carrier chain to which a large number of heparin molecules (typically seventy) have been covalently linked.^[9–12,67,68] The heparin conjugate is then irreversibly bound, basically by multiple ionic interaction, to surfaces that have been primed by a polymeric amine compound (PAV, proprietary substance supplied by Corline Systems AB, Uppsala, Sweden).

Trillium™ Biopassive Surface

This coating is commercialized by Medtronic on their line of Affinity oxygenators under a license from Biointeractions Ltd., Reading, U.K. The Trillium coating is composed of a priming layer (unknown composition) to which polyethylene oxide (PEO) chains have been attached. Heparin molecules are covalently bound to the PEO chains and in order to increase the negative charge sulphate/sulfonate groups have been introduced in addition to the negative charges provided by heparin.^[69–72] According to information put forward by Biointeractions Ltd., the sulphate/sulfonate groups should be expected to improve the nonthrombogenic character by increased binding of antithrombin.^[73] There are, however, no data available on the antithrombin binding capacity of the Trillium™ surface. With reference to reports showing that negatively charged surfaces that are not based on properly bound heparin induce coagulation,^[7,25,36] and to the detailed information of the molecular mechanisms responsible for heparin-antithrombin binding (c.f. previous text), caution should be exercised in anticipating increased antithrombin binding due to nonspecific sulfate/sulfonate groups.

Duraflo II

This coating, developed by Baxter,^[74] has been widely used in connection with products for extracorporeal circulation. The concept is based on the same approach used for preparing a surface with benzalkonium-heparin complex,^[75] but Duraflo II takes advantage of a branched surfactant carrying quarternary ammonium groups^[76] instead of benzalkonium chlorid as a way to decrease the rate of decomposition of the heparin-surfactant complex in contact with blood. Basically, a heparin-surfactant complex is precipitated



from an aqueous solution, and the precipitate is dried and then dissolved in an organic solvent. The products to be modified are then subjected to a dip-dry process. These types of heparin surface have not been designed to bind antithrombin, and there are no reports to show that antithrombin effectively interacts with a Duraflo II surface. Due to the fact that these types of ionically bound heparin surfaces are not stable in contact with blood, they are confined to short-term use.

Bioline Coating™

According to the manufacturer (Jostra AG, Hechingen, Germany) this heparin coating takes advantage of both ionic and covalent bonding of heparin to natural polypeptides,^[77] but there are no scientific reports that describe the chemical constitution or molecular structure of this approach. The main application in which Bioline is used is bypass systems intended for short-term use.^[78] To the best of our knowledge, the Bioline Coating™ does not bind any appreciable amounts of antithrombin.

T-NCVC Coating

This recent development carried out jointly by Toyobo Corporation and National Cardiovascular Center in Japan employs an ionic complex of heparin and aliphatic coupling reagent composed of several long-chain dialkyl groups.^[79] The coating is applied by a dip-dry process. No information is available on the antithrombin binding capacity of this surface. Encouraging results indicating long term stability at extracorporeal circulation in goats have been presented.^[79]

Hepaface (Terumo Co., Tokyo, Japan)

Heparin is N-desulphated using sulfuric acid to generate primary amine groups. Heparin is then linked by glutaraldehyde to PEI adsorbed to the substrate surface after treatment with ozone gas.^[80,81] Already from this description, it can be deduced that this surface should not be expected to interact favorably with antithrombin due to the fact that the aminosulfate groups of heparin are essential for its anticoagulant activity via binding to antithrombin.

CLINICAL APPLICATIONS OF HEPARIN-MODIFIED SURFACES

A heparin surface does not exist as a specific product but needs to be carried along by some sort of

medical device. This implies that very few, if any, products have been developed based on a specification that a heparin surface should be included to improve blood compatibility. The use of heparin surface modification has rather emerged by a recognized need to improve blood compatibility of existing devices that have been developed relying on available construction materials, such as polymers and metals.

The most wide-spread use of heparin surfaces is to be found in the area of extracorporeal circulation as applied in open-heart surgery. There are numerous reports from clinical studies that confirm favorable performance with less inflammatory and procoagulant response of bypass systems being modified with a heparin surface,^[11,12,35,45,82] but there are also other reports in which the investigators failed to demonstrate any improvements with respect to less formation of thrombin by adding the heparin surface.^[83,84] As pointed out above, it is important to realize that a number of different heparin surfaces have been involved in these studies and that it is impossible to draw any general conclusion that would be valid for all types of heparin surfaces.

ECMO, extracorporeal membrane oxygenation, is an example of a bypass procedure that is supposed to operate over a more extended period of time, typically several weeks. The patients undergoing this treatment are often susceptible to bleeding and would be better helped if systemic use of heparin could be reduced to a minimum. There are reports that have demonstrated that ECMO can indeed be maintained over a prolonged period of time with very little or no heparin given to the patient, provided that the blood-contacting surfaces have been modified with permanently bound heparin.^[85]

Following encouraging preclinical reports of implanting intraocular lenses made of polymethyl (methacrylate) surface modified with heparin,^[28,64,65] such IOLs are now commercially available. Clinical studies have confirmed less incidence of inflammatory adverse reactions following implantation of heparin-surface-modified IOLs.^[66] Heparin coating of stainless steel stents has been demonstrated to reduce thrombogenicity *in vitro*^[9,27,86] and *in vivo*.^[29,87] Clinical studies have confirmed less incidence of subacute thrombosis and there seems to be a trend toward reduced rate of restenosis but the evidence regarding the latter aspect is less convincing.^[88-91] Further aspects on this topic can be found in the following reviews: Refs. [67] and [92].

Additional examples of medical devices that have been modified with immobilized heparin are vascular grafts,^[93,94] indwelling catheters including decreased incidence of bacterial contamination,^[95,96] and intravascular sensors.^[97,98]

CONCLUSIONS

Heparin-binding to improve biocompatibility is a long story with many facets, which has been condensed into some major aspects in this article. Due to the catalytic mode of action, heparin offers unique opportunities for use as an immobilized substance, and it is indeed gratifying to know that this mode of action is involved in regulating the hemostatic balance of the circulating blood continuously in contact with the immense endothelial surface of the blood vessels. It follows that it is important to identify the nature and chemical constitution of surfaces designed and promoted as heparin surfaces. Adequate interaction between immobilized heparin and antithrombin is an essential prerequisite for any surface that qualifies to be considered as a heparin surface.

ACKNOWLEDGMENTS

A long standing collaboration with professor Per Olsson is gratefully acknowledged. Financial support has been provided by the Swedish Board for Technical Development and the Swedish Foundation for Strategic Research.

ARTICLES OF FURTHER INTEREST

Biocompatibility Testing; Biofunctional Polymers; Blood-Material Interactions; Cardiac Catheters; Host Reactions; Surface Coatings; Surface Modification; Thrombosis; Vascular Grafts; Vascular Grafts: Development Strategies; Vascular Grafts: Host Interactions

REFERENCES

- Lindhahl, U.; Lidholt, K.; Spillmann, D.; Kjellén, L. More to "heparin" than anticoagulation. *Thromb. Res.* **1994**, *75*, 1–32.
- Salmivirta, M.; Lidholt, K.; Lindahl, U. Heparan sulfate: A piece of information. *FASEB J.* **1996**, *10*, 1270–1279.
- Lindhahl, U. What else can "heparin" do? *Haemostasis* **1999**, *Suppl. S1*, 38–47.
- Rabenstein, D.L. Heparin and heparan sulfate: Structure and function. *Nat. Prod. Rep.* **2002 Jun**, *19* (3), 312–331. Review.
- Rosenburg, R.D.; Damus, P.S. The purification and mechanism of action of human antithrombin–heparin cofactor. *J. Biol. Chem.* **1973**, *248*, 6490.
- Gott, V.L.; Daggett, R.L. Serendipity and the development of heparin and carbon surfaces. *Ann. Thorac. Surg.* **1999**, *68* (3 Suppl.), S19–S22.
- Larm, O.; Larsson, R.; Olsson, P. A new non-thrombogenic surface prepared by selective covalent binding of heparin via a modified reducing terminal residue. *Biomater. Med. Dev. Artif. Organs* **1983**, *11* (2 & 3), 161–173.
- www.carmeda.com/bibliografi.
- Christiansen, K.; Larsson, R.; Emanuelsson, H.; Elgue, G.; Larsson, A. Heparin coating of the stent graft—Effect on platelets, coagulation and complement activation. *Biomaterials* **2001**, *22*, 349–355.
- Johnell, M.; Elgue, G.; Larsson, R.; Larsson, A.; Thelin, S.; Siegbahn, A. Coagulation, fibrinolysis, and cell activation in patients and shed mediastinal blood during coronary artery bypass grafting with a new heparin-coated surface. *J. Thorac. Cardiovasc. Surg.* **2002**, *124*, 321–332.
- Johnell, M.; Elgue, G.; Thelin, S.; Larsson, R.; Siegbahn, A. Cell adhesion and tissue factor up-regulation in oxygenators used during coronary artery bypass grafting are modified by the Corline heparin surface. *Scand. Cardiovasc. J.* **2002**, *36*, 351–357.
- Andersson, J.; Sanchez, J.; Nilsson-Ekdahl, K.; Elgue, G.; Nilsson, B.; Larsson, R. Optimal heparin surface concentration and antithrombin binding capacity as evaluated with human non-anticoagulated blood in vitro. *J. Biomed. Mater. Res.* **2003**, *67 A* (2), 458–466.
- Lam, L.H.; Silbert, J.E.; Rosenberg, R.D. The separation of active and inactive forms of heparin. *Biochem. Biophys. Res. Commun.* **1976**, *69*, 570.
- Höök, M.; Björk, I.; Hopwood, J.; Lindahl, U. Anticoagulant activity of heparin: Separation of high-activity and low-activity heparin species by affinity chromatography on immobilized antithrombin. *FEBS Lett.* **1976**, *66*, 90–93.
- Andersson, L.-O.; Barrowcliffe, T.W.; Holmer, E.; Johnson, E.A.; Sims, G.E.C. Anticoagulant properties of heparin fractionated by affinity chromatography on matrix-bound antithrombin III and by gel chromatography. *Thromb. Res.* **1976**, *9*, 575–583.
- Björk, I.; Olson, S.T.; Shore, J.D. Molecular Mechanisms of the Accelerating effect of Heparin on the Reactions Between Antithrombin and Clotting Proteinases. In *Heparin, Chemical and Biological Properties, Clinical Applications*; Lane, D.A., Lindahl, U., Eds.; Edward Arnold: London, 1989; 229–255.
- Carlström, A.-S.; Liedén, K.; Björk, I. Decreased binding of heparin to antithrombin following the interaction between antithrombin and thrombin. *Thromb. Res.* **1977**, *11*, 785–797.
- Jordan, R.; Beeler, D.; Rosenberg, R. Fractionation of low molecular weight heparin species and their interaction with antithrombin. *J. Biol. Chem.* **1979**, *254*, 2902–2913.
- Hricovini, M.; Guerrini, M.; Bisio, A.; Torri, G.; Petitou, M.; Casu, B. Conformation of heparin pentasaccharide bound to antithrombin III. *Biochem. J.* **2001**, *359*, 265–272.
- de Agostini, A.I.; Watkins, S.C.; Slayter, H.S.; Yusufian, H.; Rosenberg, R.D. Localization of anticoagulant active heparan sulfate proteoglycans in vascular endothelium: Antithrombin binding on



- cultured endothelial cells and perfused rat aorta. *J. Cell Biol.* **1990**, *111*, 1293–1304.
21. Sánchez, J.; Olsson, P. On the control of the plasma contact activation system on human endothelium comparison with heparin surface. *Thromb. Res.* **1999**, *93*, 27–34.
 22. Ware, J.A.; Colter, B.S. Platelet Morphology, Biochemistry and Function. In *Williams Hematology*, 5th Ed.; Beutler, E., Lichtman, M.A., Colter, B.S., Kipps, T.J., Eds.; McGraw-Hill: New York, 1995; 1161–1201.
 23. Mulivor, A.W.; Lipowsky, H.H. Role of glycocalyx in leucocyte-endothelial cell adhesion. *Am. J. Physiol. Heart Circ. Physiol.* **2002**, *283*, H1282–H1291.
 24. Larsson, R.; Rosengren, Å.; Olsson, P. Determination of platelet adhesion to polyethylene and heparinized surfaces with the aid of bioluminescence. *Thromb. Res.* **1977**, *11*, 517–530.
 25. Larsson, R.; Eriksson, J.C.; Lagergren, H.; Olsson, P. Platelet and plasma coagulation compatibility of heparinized and sulphated surfaces. *Thromb. Res.* **1979**, *15*, 157–167.
 26. Weber, N.; Wendel, H.P.; Ziemer, G. Hemocompatibility of heparin-coated surfaces and the role of selective plasma protein adsorption. *Biomaterials* **2002**, *23*, 429–439.
 27. Bickel, C.; Rupprecht, H.J.; Darius, H.; Binz, C.; Hauröder, B.; Kummener, F.; Meyer, J. Substantial reduction of platelet adhesion by heparin-coated stents. *J. Interv. Cardiol.* **2001**, *14*, 407–413.
 28. Larsson, R.; Selén, G.; Björklund, H.; Fagerholm, P. Intraocular PMMA lenses modified with surface-immobilized heparin: Evaluation of biocompatibility in vitro and in vivo. *Biomaterials* **1989**, *10*, 511–516.
 29. Chronos, N.; Markou, C.; Kocsis, J.; Llanos, G.; Hanson, S. Surface heparinization profoundly decreases acute thrombosis on Crown and Mini Crown stents in the baboon arteriovenous shunt model. *J. Am. Coll. Cardiol.* **1998**, *31* (Suppl. A), 413.
 30. Mollnes, T.E.; Videm, V.; Christiansen, D.; Bergseth, G.; Riesenfeld, J.; Hovig, T. Platelet compatibility of an artificial surface modified with functionally active heparin. *Thromb. Haemost.* **1999**, *82*, 1132–1136.
 31. Keuren, J.F.W.; Wielders, S.J.H.; Willems, G.M.; Morra, M.; Lindhout, T. Fibrinogen adsorption, platelet adhesion and thrombin generation at heparinized surfaces exposed to blood. *Thromb. Haemost.* **2002**, *87*, 742–747.
 32. Solberg, R.; Scholtz, T.; Videm, V.; Okkenhaug, C.; Aasen, A.O. Heparin coating reduces cell activation and mediator release in an in vitro venovenous bypass model for liver transplantaton. *Transpl. Int.* **1998**, *11*, 252–258.
 33. Borowiec, J.W.; Jaramillo, A.; Venge, P.; Nilsson, L.; Thelin, S. Effects of heparin-coating of cardiopulmonary bypass circuits on leucocytes during simulated extracorporeal circulation. *Cardiovasc. Surg.* **1997**, *5* (6), 568–573.
 34. Barstas, R.M.; Hamers, M.J.; Möller, A.S.; Sakariaassen, K.S. Monocyte procoagulant activity induced by adherence to an artificial surface is reduced by end-point immobilized heparin-coating of the surface. *Thromb. Haemost.* **1998**, *79*, 302–305.
 35. Fosse, E.; Moen, O.; Johnson, E.; Semb, G.; Brockmeier, V.; Mollnes, T.E.; Fagerhol, M.K.; Venge, P. Reduced complement and granulocyte activation with heparin-coated cardiopulmonary bypass. *Ann. Thorac. Surg.* **1994**, *58*, 472–477.
 36. Larsson, R.; Olsson, P.; Lindahl, U. Inhibition of thrombin on surfaces coated with immobilized heparin and heparin-like substances: A crucial non-thrombogenic principle. *Thromb. Res.* **1980**, *19*, 43–54.
 37. Olsson, P.; Sánchez, J.; Mollnes, T.E.; Riesenfeld, J. On the compatibility of end-point immobilized heparin. *J. Biomater. Sci., Polym. Ed.* **2000**, *11*, 1261–1273.
 38. Sánchez, J.; Lundquist, P.B.; Elgue, G.; Larsson, R.; Olsson, P. Measuring the degree of plasma contact activation induced by artificial materials. *Thromb. Res.* **2002**, *105*, 407–412.
 39. Sánchez, J.; Elgue, G.; Riesenfeld, J.; Olsson, P. Studies of adsorption, activation and inhibition of factor XII on immobilized heparin. *Thromb. Res.* **1998**, *89*, 41–50.
 40. Sánchez, J.; Elgue, G.; Riesenfeld, J.; Olsson, P. Inhibition of the plasma contact activation system of immobilized heparin: Relation to surface density and antithrombin binding sites. *J. Biomed. Mater. Res.* **1997**, *36*, 37–42.
 41. Hong, J.; Larsson, A.; Elgue, G.; Nilsson Ekdahl, K.; Larsson, R.; Nilsson, B. Contact between biomaterial and whole blood: The sequence of events that leads to thrombin generation. *J. Clin. Lab. Med.* **2000**, *138*, 139–145.
 42. Pekna, M.; Larsson, R.; Formgren, B.; Nilsson, U.R.; Nilsson, B. Complement activation by polymethyl methacrylate is minimized by end-point heparin attachment. *Biomaterials* **1993**, *14*, 189–192.
 43. Gong, J.; Larsson, R.; Nilsson Ekdahl, K.; Mollnes, T.E.; Nilsson, U.; Nilsson, B. Tubing loops as a model for cardiopulmonary bypass circuits: Both the biomaterial and the blood-gas phase interphases induce complement activation in an in vitro model. *J. Clin. Immunol.* **1996**, *16*, 223–230.
 44. Videm, V.; Mollnes, T.E.; Garred, P.; Aasen, A.; Svennevig, J.L. Biocompatibility of extracorporeal circulation: In vitro comparison of heparin-coated and uncoated oxygenator circuits. *J. Thorac. Cardiovasc. Surg.* **1991**, *101*, 654–660.
 45. Videm, V.; Svennevig, J.L.; Fosse, E.; Semb, G.; Österud, A.; Mollnes, T.E. Reduced complement activation with herapin-coated oxygenator and tubings in coronary bypass operations. *J. Thorac. Cardiovasc. Surg.* **1992**, *103*, 806–813.
 46. Kopp, R.; Mottaghy, K.; Kirschfink, M. Mechanism of complement activation during extracorporeal blood-biomaterial interaction: Effects of heparin coated and uncoated surfaces. *ASAIO J.* **2002**, *48*, 598–605.
 47. Nilsson, B.; Larsson, R.; Hong, J.; Elgue, G.; Nilsson Ekdahl, K.; Sahu, A.; Lambris, J.D. Compstatin inhibits complement and cellular activation in whole blood in two models of extracorporeal circulation. *Blood* **1998**, *92*, 1661–1667.
 48. Conrad, H.E. *Heparin-Binding Protients*; Academic Press: New York, 1998.

49. Osmond, R.I.W.; Kett, W.C.; Skett, S.E.; Coombe, D.R. Protein-heparin interactions measured by BIAcore 2000 are affected by the method of heparin immobilization. *Anal. Biochem.* **2002**, *310*, 199–207.
50. Casu, B.; Lindahl, U. Structure and biological interactions of heparin and heparan sulfate. *Adv. Carbohydr. Chem. Biochim.* **2001**, *57*, 159–206.
51. Underwood, P.A.; Kirkpatrick, A.; Mitchell, S.M. New insights into heparin binding to vitronectin: Studies with monoclonal antibodies. *Biochem. J.* **2002**, *365*, 57–67.
52. Schwarz, I.; Seger, D.; Shaltiel, S. Vitronectin. *Int. J. Biochem. Cell Biol.* **1999**, *31*, 539–544.
53. Timpl, R.; Tisi, J.F.; Andac, Z.; Sasaki, T.; Hohenester, E. Structure and function of laminin LG modules. *Matrix Biol.* **2000**, *19*, 309–317.
54. Lookene, A.; Chevreuil, P.; Ostergaard, P.; Olivecrona, G. Interaction of lipoprotein lipase with heparin fragments and with heparan sulfate: Stoichiometry, stabilization, and kinetics. *Biochemistry* **1996**, *35*, 12155–12163.
55. Rosenberg, R.D.; Schworak, N.W.; Liu, J.; Schwartz, J.J.; Zhang, L. Heparan sulfate proteoglycans of the cardiovascular system. Specific structures emerge but how is synthesis regulated. *J. Clin. Invest.* **1997**, *99*, 2062–2070.
56. Connelly, P.W. The role of hepatic lipase in lipoprotein metabolism. *Clin. Chim. Acta* **1999**, *286*, 243–255.
57. Calabrese, C.G.; Recondo, E.F.; Fernandez de Recondo, M.E. Antithrombin and first complement protein recognize the same active heparin fraction. *Thromb. Res.* **2002**, *15*, 537–541.
58. Blom, A.M. Structural and functional studies of complement inhibitor C4b binding. *Biochem. Soc. Trans.* **2001**, *30*, 978–982.
59. Zipfel, P.F.; Skerka, C.; Hellwage, J.; Jokiranta, S.T.; Meri, S.; Brade, V.; Kraiczy, P.; Noris, M.; Remuzzi, G. Factor H family proteins: On complement, microbes and human diseases. *Biochem. Soc. Trans.* **2001**, *30*, 971–978.
60. Carr, J.A.; Silverman, N. The heparin-protamine interaction. A review. *J. Cardiovasc. Surg. (Torino)* **1999**, *40*, 659–666.
61. Colman, R.W. Role of the light chain of high molecular weight kininogen in adhesion, cell-associated proteolysis and angiogenesis. *Biol. Chem.* **2001**, *382*, 65–70.
62. Kaplan, A.P.; Joseph, K.; Silverberg, M. Pathways for bradykinin formation and inflammatory disease. *J. Allergy Clin. Immunol.* **2002**, *109*, 195–209.
63. Xu, X.; Takano, R.; Nagai, Y.; Yanagida, T.; Kamei, K.; Kato, H.; Kamikubo, Y.; Nakahara, Y.; Kumeda, K.; Hara, S. Effect of heparin chain length on the interaction with tissue factor pathway inhibitor (TFPI). *Int. J. Biol. Macromol.* **2002**, *30*, 151–160.
64. Larsson, R.; Selén, G.; Formgren, B.; Holst, A.S. Long-term stability of heparin surface-modified intraocular lenses in vivo. *J. Cataract Refract. Surg.* **1990**, *18*, 247–251.
65. Spångberg, M.; Kihlström, I.; Björklund, H.; Bjurström, S.; Lydahl, E.; Larsson, R. Improved biocompatibility of intraocular lenses by heparin surface modification: A 12-month implantation study in monkeys. *J. Cataract Refract. Surg.* **1990**, *16*, 170–177.
66. Trocme, S.D.; Li, H. Effect of heparin-surface-modified intraocular lenses on postoperative inflammation after phacoemulsification: A randomized trial in a United States patient population heparin-surface-modified lens study group. *Ophthalmology* **2000**, *107*, 1031–1037.
67. van der Giessen, W.; van Beusekom, H.M.M.; Larsson, R.; Serruys, P.W. Heparin-coated coronary stents. *Curr. Interv. Cardiol. Rep.* **1999**, *1*, 234–240.
68. Kristensen, E.M.E.; Rensmo, H.; Larsson, R.; Siegbahn, H. Characterisation of a heparin surface using photoelectron spectroscopy and quartz crystal microbalance. *Biomaterials* **2003**, *24*, 4153–4159.
69. Mueller, X.M.; Tevæarai, H.T.; Jegger, D.; Augstburger, M.; Goddar, G.; von Segesser Antithrombotic properties of trillium coated connectors. *ASAIO J.* **2002**, *48*, 483–486.
70. Murphy, J.A.; Savage, C.M.; Alpard, S.K.; Deyo, D.J.; Jayroe, J.B.; Zwischenberger, J.B. Low-dose versus high-dose heparinization during arteriovenous carbon dioxide removal. *Perfusion* **2001**, *16*, 460–468.
71. Ereth, M.H.; Nuttall, G.A.; Clarke, S.H.; Dearani, J.A.; Fiechtner, B.K.; Rishavy, C.R.; Buda, D.A.; Shaw, T.A.; Orszulak, T.A.; Oliver, W.C., Jr. Biocompatibility of trillium biopassive surface-coated oxygenator versus uncoated oxygenator during cardiopulmonary bypass. *J. Cardiothorac. Vasc. Anesth.* **2001**, *15*, 545–550.
72. Palanzo, D.A.; Zarro, D.L.; Manley, N.J.; Montesano, R.M.; Quinn, M.; Elmore, B.A.; Gustafson, P.A.; Castagna, J.M. Effect of Carmeda bioactive surface coating versus Trillium biopassive surface coating of the oxygenator on circulating platelet count drop during cardiopulmonary bypass. *Perfusion* **2001**, *16*, 279–283.
73. Sandhu, S.; Luthra, A. New biointeracting materials. *Med. Device Technol.* **Oct.**, **2002**.
74. Hsu, L.C. Heparin-coated cardiopulmonary bypass circuits: Current status. *Perfusion* **2001**, *16*, 417–428.
75. Cramer, R.; Moore, R.; Amplatz, K. Reduction of the surgical complication rate by the use of a hypothermogenic catheter coating. *Radiology* **1973**, *109*, 585–588.
76. Hsu, L.C.; Tong, S.D. Ionic Heparin Complex. US patent 4.871.357, 1989.
77. Bader, K.E. Biocompatibility in extracorporeal circulation: The heparin coating method Bioline coating. *Can. Perfus. Can.* **1997**, *9*, 9–16.
78. Tayama, E.; Hayashida, N.; Akasu, K.; Kosuga, T.; Fukunaga, S.; Akashi, H.; Kawara, T.; Aoyagi, S. Biocompatibility of heparin-coated extracorporeal bypass circuits: New heparin bonded Bioline system. *Artif. Organs* **2000**, *24*, 618–623.
79. Nishinaka, T.; Tatsumi, E.; Taenaka, Y.; Katagiri, N.; Ohnishi, H.; Shioya, K.; Fukuda, T.; Oshikawa, M.; Sato, K.; Tsukiya, T.; Homma, A.; Takewa, Y.; Takano, H.; Sato, M.; Kashiwabara, S.; Tanaka, H.; Sakai, K.; Matsuda, T. At least thirty-four days of animal continuous perfusion by a newly developed extracorporeal membrane oxygenation system without systemic anticoagulants. *Artif. Organs* **2002**, *26*, 548–551.
80. Nojiri, C.; Hagiwara, K.; Yokoyama, K.; Kuribayashi, E.; Hidaka, K.; Ishida, N.; Horiuchi, K.; Oshiyama, H.; Nogawa, A.; Kido, T.; Kijama, T.; Akutsu, T. Evaluation of a new heparin bonding process in process in prolonged extracorporeal membrane oxygenation. *ASAIO J.* **1995**, *41*, M561–M567.



81. Usui, A.; Hiroura, M.; Kawamura, M. Heparin coating extends the durability of oxygenators used for cardiopulmonary support. *Artif. Organs* **1999**, *23*, 840–844.
82. Svenmarker, S.; Häggmark, S.; Jansson, E.; Lindholm, R.; Appelblad, M.; Sandström, E.; Åberg, T. Use of heparin-bonded circuits in cardiopulmonary bypass improves clinical outcome. *Scand. Cardiovasc. J.* **2002**, *36*, 241–246.
83. Ernofsson, M.; Thelin, S.; Siegbahn, A. Thrombin generation during cardiopulmonary bypass using heparin-coated or standard circuits. *Scand. J. Thorac. Cardiovasc. Surg.* **1995**, *29*, 157–165.
84. Ovrum, E.; Holen, E.A.; Tangen, G.; Brosstad, F.; Abdelnoor, M.; Ringdal, M.; Oystese, R.; Istad, R. Completely heparinized cardiopulmonary bypass and reduced systemic heparin: Clinical and hemostatic effects. *Ann. Thorac. Surg.* **1995**, *60*, 365–371.
85. Bindslev, L.; Bohm, C.; Jolin, A.; HambræusJonzon, K.; Olsson, P.; Ryniak, S. Extracorporeal carbon dioxide removal performed with surface-heparinized equipment in patients with ARDS. *Acta Anaesthesiol. Scand., Suppl.* **1991**, *95*, 125–130. Discussion 130-1.
86. Kocsis, J.; Lunn, A.; Mohammad, S. Incomplete expansion of coronary stents: Risk of thrombosis and protection provided by a heparin coating. *J. Am. Coll. Cardiol.* **1996**, *27* (Suppl. A), 84A.
87. Hardhammar, P.A.; vanBeusekom, H.M.; Emanuelsson, H.U.; Hofma, S.H.; Albertsson, P.A.; Verdouw, P.D.; Boersma, E.; Serruys, P.W.; van der Giessen, W.J. Reduction in thrombotic events with heparin-coated Palmaz-Schatz stents in normal porcine coronary arteries. *Circulation* **1996**, *93*, 423–430.
88. Serruys, P.W.; Emanuelsson, H.; van der Giessen, W.; Lunn, A.C.; Kiemeney, F.; Macaya, C.; Rutsch, W.; Heyndrickx, W.; Suryapranata, G.; Legrand, H.; Goy, V.; Materne, J.J.; Bonnier, P.; Morice, H.; Fajadet, M.C.; Belardi, J.; Colombo, J.; Garcia, A.; Ruygrok, E.; de Jaegere, P.; Morel, M.A. Heparin-coated Palmaz-Schatz stents in human coronary arteries. Early outcome of the benestent-II pilot study. *Circulation* **1996**, *93*, 412–422.
89. Serruys, P.W.; van Hout, B.; Bonnier, H.; Legrand, V.; Garcia, E.; Macaya, C.; Sousa, E.; van der Giessen, W.; Colombo, A.; Seabra-Gomes, R.; Kiemeney, F.; Ruygrok, P.; Ormiston, J.; Emanuelsson, H.; Fajadet, J.; Haude, M.; Klugmann, S.; Morel, M.A. Randomised comparison of implantation of heparin-coated stents with balloon angioplasty in selected patients with coronary artery disease. (Benestent II). *Lancet* **1998**, *352*, 673–681.
90. Wohrle, J.; Al-Khayer, E.; Grotzinger, U.; Schindler, C.; Kochs, M.; Hombach, V.; Hoher, M. Comparison of the heparin coated vs. the uncoated Jostent—No influence on restenosis or clinical outcome. *Eur. Heart J.* **2001**, *22*, 1808–1816.
91. Shin, E.K.; Son, W.J.; Sohn, M.S.; Jin, D.K.; Park, G.S.; Koh, K.K.; Ahn, T.H.; Choi, I.C. Efficacy of heparin-coated stents in early setting of acute myocardial infarction. *Catheter. Cardiovasc. Interv.* **2001**, *52*, 306–312.
92. Kocsis, J.F.; Llanos, G.; Holmer, E. Heparin-coated stents. *J. Long-Term Eff. Med. Implants* **2000**, *10*, 19–45.
93. Walpoth, B.H.; Rogulenko, R.; Tikhvinskaia, E.; Gogolewski, S.; Schaffner, T.; Hess, O.M.; Althaus, U. Improvement of patency rate in heparin-coated small synthetic vascular grafts. *Circulation* **1998**, *98* (19 Suppl.), II319–II323.
94. Wissink, M.J.; Beernink, R.; Pieper, J.S.; Poot, A.A.; Engbers, G.H.; Beugeling, T.; van Aken, W.G.; Feijen, J. Binding and release of basic fibroblast growth factor from heparinized collagen matrices. *Biomaterials* **2001**, *22*, 2291–2299.
95. Foley, P.; Barthel, C.; Brausa, H. Effect of covalently bound heparin coating on patency and biocompatibility of long-term indwelling catheters in the rat jugular vein. *Comp. Med.* **2002**, *52*, 243–248.
96. Appelgren, P.; Ransjo, U.; Bindslev, L.; Espersen, F.; Larm, O. Surface heparinization of central venous catheters reduces microbial colonization in vitro and in vivo: Results from a prospective, randomized trial. *Crit. Care Med.* **1996**, *24*, 1482–1489.
97. Nilsson, E.; Arnander, C. Long-term monitoring of arterial pO₂ in burned patients. *Clin. Physiol.* **1984**, *4*, 13–21.
98. Nagaoka, S.; Mikami, M.; Shimizu, Y. Antithrombogenic pO₂ sensor for continuous intravascular oxygen monitoring. *Biomaterials* **1990**, *11*, 414–418.

Hip Biomechanics

Farid Amirouche

Department of Mechanical Engineering, University of Illinois at Chicago, Chicago, Illinois, U.S.A.

Mark Gonzalez

Department of Orthopedic Surgery, University of Illinois at Chicago, Chicago, Illinois, U.S.A.

Luke Aram

DePuy Orthopedics/Johnson & Johnson Company



THE NORMAL HIP, ARTHRITIS, AND THE ARTIFICIAL HIP

The hip is an essential component to human locomotion whose mechanical unit function can be explained by the law of mechanics. The hip joint is composed of the femur (ball) and acetabulum (socket). Although neither are perfect spheres or congruent, they usually function as an effective ball and socket joint. The three bones of the pelvis form the acetabulum: the ilium (wing), the ischium, and the pubis (Fig. 1).

Numerous ligaments and tendons traverse the healthy hip and thus it is the most difficult joint in the body to dislocate.^[1] When a person stands symmetrically on both legs, the hips support 62% of the body weight. In a static equilibrium position, little muscular force is needed to maintain equilibrium, and thus each hip supports 31% of the body weight.^[2] When a person stands on one leg, the hip supports the upper body as well as the other leg, or 81% of the total body weight (K). However, Maquet^[2] formulates that the joint reaction force in the hip is a little more than three times the body weight owing to abductor muscles (M) counterbalancing the eccentric forces of gravity (Fig. 2). So the resultant force inside the hip joint (R) is the vector sum of the (K) and (M):

$$R = \sqrt{K^2 + M^2 + 2KM \cos(\widehat{KM})} \quad (1)$$

During gait, kinematics is computed by tracking markers on the body. Once the body is in motion the inertia forces change the dynamics of the pelvis/leg joint forces. Using Newton's and Euler's equations the inverse dynamic problem results in the evaluation of those reaction forces at the hip. The forces at the hip joint during gait were estimated by Maquet to vary between 138% body weight during the middle of the leg swing, and can reach 433% at the time of impact during heel strike.

In the late 1980s, a group of engineers and surgeons at Case Western led by Davy implanted a total hip with a load cell, battery, and transmitter built into the neck of an artificial hip.^[4] One month after surgery, they recorded joint forces different than those previously calculated. During single leg stance, the hip force was 1.0 times body weight, more than three times that calculated by Maquet. During single leg stance, 2.1 times the body weight was recorded, 30% less than Maquet (Table 1).

For this study, we assumed that the maximum force exerted to the hip during gait is approximately three times the body weight or 1500 N.

SURGICAL TECHNIQUE FOR TOTAL HIP REPLACEMENT

Cementless total hip arthroplasty (THA) is coated with a porous surface (200–400 μm pore size) that allows bone ingrowth providing long-term integration and fixation of the prosthesis. To ensure ingrowth into the prosthesis there must be: 1) minimal motion between the bone and prosthesis and 2) close apposition between the prosthesis and the bone. A 3- to 10-in. incision is made parallel to the femur, through the skin and the musculature of the hip. The hip can be dislocated anteriorly or posteriorly.^[5] The femoral neck is cut, the head removed, and the canal reamed with cylindrical reamers. The proximal femur is then shaped with broaches to accept the femoral component. The prosthesis must be placed in tight apposition to permit ingrowth, but care must be taken not to oversize the prosthesis and fracture the bone.^[6]

Reaming of the acetabulum must be precise to guarantee a tight interference fit of the cup into the bone. Bone will not ingrow into a cup exhibiting gross motion. The acetabular reamer looks like a spherical cheese grater with a handle that attaches to a reamer (Fig. 3).

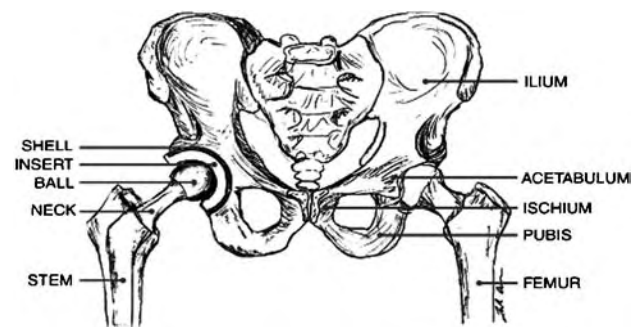


Fig. 1 An illustration of the human pelvis with a healthy hip joint on the right and an artificial hip joint on the left. The pubis, ischium, ilium, and acetabulum form the healthy pelvis. The shell and insert make up the acetabular component of the total hip. The ball, neck, and stem make up the femoral component.

Reaming is performed with reamers of increasingly larger diameters to remove all cartilage and attain the proper depth. Then the surgeon aligns the reamer to a position correlating to 45° of abduction from the horizontal and approximately 25° of anteversion.^[7] Research suggests that the spherical cut has diametric errors averaging 2.1%.^[8] This leads to a limited contact of the bone with the cup that some estimate to be as low as 20% to 60% of the total surface area of the acetabulum.^[9] After reaming is complete, the surgeon implants the cup with an impactor. An alignment guide is used to achieve 45° of inclination and 0° to 30° of anteversion. If the cup is placed in too much anteversion, the risk of posterior impingement increases, and too little

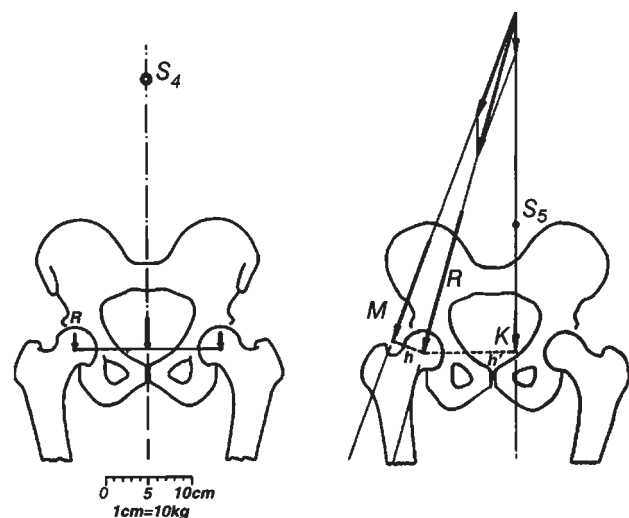


Fig. 2 Person standing with weight supported symmetrically on both legs. The center of gravity of the head, trunk, and upper limbs is denoted by S_4 . R denotes the force exerted on the hip.

Table 1 Calculated and measured peak hip forces during different activities

Activity	Hip forces calculated by Maquet	Hip forces measured by Davy
Dual leg stance	31% BW	100% BW
Single leg stance	300% BW	210% BW
Gait cycle	433% BW	
Walking with crutch		280% BW
Walking with parallel bars		240% BW
Climbing stairs		260% BW

anteversion runs a risk of anterior impingement.^[9] Impingement of the femoral neck on the acetabular component produces a shear on the bone prosthesis interface that is detrimental to its stability (Fig. 4).

The reaming can be performed 2 mm less in diameter than the diameter of the cup (under-reaming) to improve the interference fit. Under-reaming in very osteoporotic bone can cause an acetabular fracture and is to be avoided. After implantation of the acetabular cup, the polyethylene acetabular liner is inserted.

A femoral trial is now placed into the femur and a trial femoral head is used to judge the stability of the implant. The neck length of and offset of the femoral head can now be used to adjust stability of the prosthesis and leg length. The appropriate femoral component is now inserted into the femur. The modular head of appropriate length is locked onto the Morse taper of the femoral trunion. The prosthesis is reduced and the joint closed in layers.

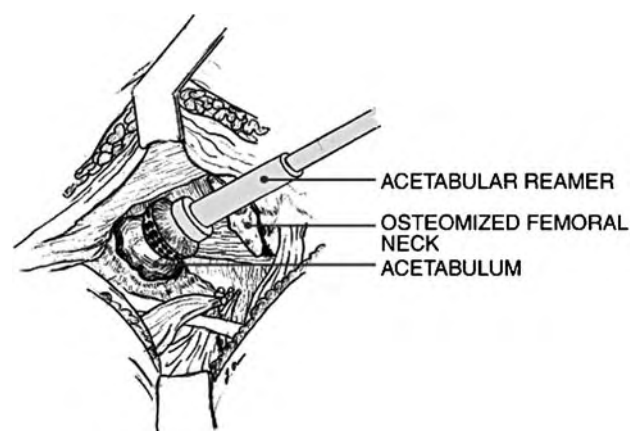


Fig. 3 Reaming of the acetabulum. The acetabular reamer is placed inside a drill, and the bone is reamed using successively larger reamers until the desired diameter is reached.



Fig. 4 An illustration of a femoral stem impinging on an acetabular cup.

IMPLANT FIXATION

Modern implant fixation is generally fostered by cementation with methylmethacrylate and by cementless fixation. Various implant designs rely on different design elements to achieve fixation. These designs can be generally categorized as follows.

1. **Cemented**—Implants rely on polymethylmethacrylate cement that mechanically interlocks with both the metallic implant and the bone. Cement pressurization during the surgery helps the cement to interdigitate with the trabecular bone.
2. **Cementless**—Implants rely on bony ingrowth into a porous coated metallic implant (pore size 200–400 μm). This is most commonly done by sintering titanium or cobalt chromium bead or fiber mesh onto the implant.
3. **Hybrid**—A cementless cup is implanted with a cemented stem.

An abundance of clinical studies show excellent performance of both cemented and uncemented femoral components. Cemented femoral stems are technique intensive and long-term survival of the prosthesis is dependent on an adequate cement mantle (2–3 mm) surrounding the prosthesis. Cemented stems are immediately stable and are preferred by some surgeons in older patients where osteoporosis may limit stability and ingrowth into a cementless prosthesis.

There have been several studies showing progressive loosening of cemented acetabular components, compelling a majority of surgeons in the U.S.A. to prefer cementless fixation for the acetabular cup. In older patients, some surgeons prefer a cemented stem with a cementless cup termed “hybrid fixation.” Most

successful implants rely on a porous material that is conducive for bony ingrowth and a geometry that enables the surgeon to attain initial press fit. Popular acetabular designs include a simple hemisphere cup, a tri-spike design, and a cup that includes holes for adjunctive screw fixation.

Adjunctive screws are commonly used to enhance the initial stability of a hip cup. However, adjunctive screws have three potential failure modes. One, the screw holes are direct pathways for polyethylene wear particles to enter the bone–prosthesis interface. Two, the screw itself is another wear surface, scratching the polyethylene liner and the metal shell, contributing to the number of wear particles in the joint space. Three, the use of screw fixation leads to an uneven load distribution, and often an uneven bone remodeling. In scenarios where bone prosthesis contact is deficient or bone is severely osteoporotic, adjunctive screws may be necessary to achieve implant stability.

Biomaterial

The chief mode of failure of the modern THA is aseptic loosening.^[10] The hip articulation is relatively concentric producing fine submicron polyethylene wear particles. Wear debris is opsonized by macrophages, initiating an inflammatory cascade that creates osteolysis. Osteolysis degrades both the bone–prosthesis interface and the bone–cement interface producing loosening. Improvements in prosthesis design and surgical technique have led to reduction in wear. Improvements in polyethylene manufacture, such as cross-linking, and the introduction of alternate bearing surfaces such as metal-on-metal and ceramic-on-ceramic may further improve total hip longevity in the future.

Material selection is an important factor in the development of wear particles. Titanium is commonly used to fabricate femoral stems and cups because of high strength, low stiffness, and osteophilic properties. Titanium alloys however, are not used for femoral head fabrication because of poor wear characteristics.^[11] A polished cobalt chrome head articulating on an ultra-high molecular weight polyethylene (UHMWPE) liner has improved wear characteristics. The UHMWPE is an inert and nontoxic material. In its bulk form it is harmless to the body. In particle form it is an initiator of osteolysis. It is estimated that one million UHMWPE wear particles are produced from a total hip in a single step.^[12] It is also estimated that the average individual takes one million steps a year. The volume of wear produced is relatively small, but the large number of submicron particles appear to be the primary cause of osteolysis and the eventual loosening of a previously well-fixed prosthesis.

In 1977, Willert was among the first to hypothesize that local macrophage response to wear particles

contributes to aseptic loosening.^[13] Goldring, in 1983, showed that cells within the periprosthetic membrane produces large amounts of “bone resorbing” factors around loose implants.^[14] In 1992, Quinn and others demonstrated that wear particle-induced giant cells are capable of absorbing bone in vitro. More recent studies have documented the exact biochemical pathways of bone resorption and its relationship to UHMWPE particles.^[15]

One way to reduce the effects of osteolysis on the success of a total hip is to limit the exposure of the bone to wear particles. Femoral stems designed with circumferential coating have a better survival than partially coated stems because complete ingrowth around the fully coated stems precludes the passage of wear particles into the bone–prosthesis interface. Acetabular cups with no screw holes or with screw-in plugs for the holes to limit the egress of particulate matter into the bone–cup interface have been developed.

Implant Design

The femoral component has two essential components: the ball and the stem. The ball is typically a highly polished cobalt chromium sphere, which provides an excellent articulating surface. The stem provides the fixation interface between the femur and the prosthesis.

The acetabular component is typically composed of a metal shell and a UHMWPE insert. The UHMWPE insert is an essential component to the success of the joint. Excessive wear of the polyethylene component and subsequent aseptic loosening is the leading cause of implant failure. Therefore researchers have concentrated on reducing wear by decreasing contact stresses, evenly distributing contact stresses, and improving the wear characteristics of the polyethylene liner. Work has also been done to diminish the effects of wear particles by manipulating the size of the particle produced and preventing the particles from entering the bone–prosthesis interface by removing screw holes and sealing the insertion hole.

It was once believed that a low-friction hip joint applied very little torque to the hip cup. Research more recently suggests that impingement (Fig. 4) of the femoral stem on the rim of the acetabular cup produces a torque on the cup.^[16] Also, radiographic evidence exists that the femoral component commonly subluxes when a patient sits and then recenters during stance initiation.^[17] During recentering, the cup is eccentrically loaded and thus a torque is produced (Fig. 5). Newer cup designs incorporate rigid locking mechanism designs of the liner in the outer metal cup to counteract rotational and torsional forces.

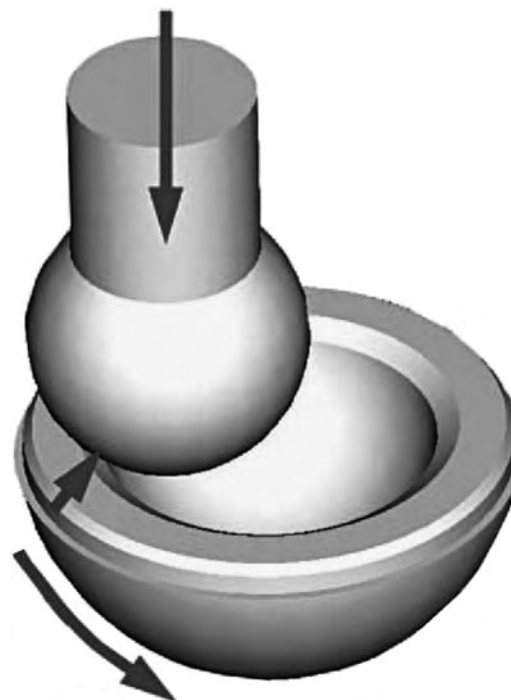


Fig. 5 An illustration of a femoral that has been subluxed from the cup and is recentering, applying a torque to the cup.

THE RELATIONSHIP BETWEEN HIP CUP STABILITY, MICROMOTION, FIXATION, AND OSTEOLYSIS

If initial stability is not achieved, the normal physiological movements of the patient will produce excessive micromotion between the bone and the prosthesis, preventing bony ingrowth. Without bony ingrowth, loosening of the hip cup will lead to early failure.^[18] The obvious and most common solution is to add adjunctive screw fixation to the acetabular component. While adjunctive screws do decrease micromotion by 20% to 30%,^[19] they may indirectly lead to the increase of osteolysis.^[20]

THE RELATIONSHIP BETWEEN PELVIC OSTEOLYSIS AND ADJUNCTIVE SCREWS

A recent study involving 132 primary cementless hip arthroplasties at Tulane University reported an acetabular revision rate of 5% at five to eight years owing to acetabular osteolysis (Fig. 6).^[20] “Abrasion of the screw head against the back-side of the polyethylene liner was seen in all retrieved cases and may have contributed to the development of the lytic lesions seen.” The materials used in the uncemented THA implant selected for this study can be divided in two groups, metallic



Fig. 6 Well-fixed THA with osteolytic lesion of the superior acetabulum as a result of polyethylene wear.

biomaterials (femoral head, femur stem, and acetabular cup) and polymeric biomaterial (liner). Table 2 shows the material properties of biomaterials involved in the present study.

Implant design has been shown to have an adverse effect on polyethylene wear. With certain designs, a gap may be found between the polyethylene liner and the metal cup component. Implants are manufactured with exact dimensions drawn from computer aided design (CAD) drawings under some tolerance conditions. In the manufacturing process, reproduction of the exact dimensions of the proposed design is not always achieved. The actual implant may be slightly smaller or larger than the design specifications, and a gap clearance would occur between the liner and the metal cup. Therefore, the liner would be unsupported over the weight-bearing dome location of the gap of the prosthesis and would then transfer high stresses to the rim. In some designs of liner, the polyethylene thickness at the rim may be smaller than in the dome surface; therefore, greater stresses are generated in the thinner

parts of the liner. When the polyethylene yield stress is exceeded as might be the case in other designs, a failure in the liner rim may occur with eventual fracture and penetration of the femoral head against the acetabular cup.

Dislocation after THA

After aseptic loosening, dislocation of the femoral head within the acetabular liner is the most significant complication that occurs after THA.^[21–23] The incidence of dislocation after THA varies from study to study, but in general it is in the range of 1% to 10% of the THA patients (Fig 7). It is important to highlight that the incidence of dislocation after THA revision is significantly greater. The most significant preoperative risk factor for dislocation after THA is previous hip surgery. Other factors include operative technique (leg length and component position), gender (greater dislocation cases in females over males), mental disorders, neuromuscular disease, alcoholism, and patient noncompliance.

Femoral Head Subluxation

During THA many supporting structures for the retention of the femoral head within the acetabulum are removed. Previous studies have determined that this lack of supporting structures may lead to separation of the femoral head within the polyethylene liner.^[24,25] Subluxation of the femoral head can be defined as the separation of the femoral head within the liner during abduction/adduction activities. While total dislocation after THA does not necessarily imply a relocation of the femoral head just following the dislocation, subluxation is always followed by the relocation of the femoral head within the liner.

In previous studies using fluoroscopy techniques, separation between the femoral head and the liner during subluxation were computed.^[25] In this study 20 subjects were analyzed. Five subjects had a normal hip, 10 subjects were implanted with an unconstrained THA, and five subjects were implanted with a THA using a constrained acetabular component. All THA

Table 2 Selected properties of biomaterials

Material	Young's modulus, E (GPa)	Tensile strength (MPa)	Poisson's ratio, ν (unitless)	Density, ρ (g/cm ³)
Stainless steel	190	586–1351	0.33	8.8
Cobalt-chromium	210–253	655–1896	0.28	7.8
Titanium alloy	110	760	0.34	4.5
UHMWPE	0.975	≥ 35	0.46	0.94
Cortical bone	15–30	30–70	< 0.5	1.9



Fig. 7 Patient with bilateral THA. The prosthesis on the right of the picture is reduced concentrically and the prosthesis on the left of the picture is dislocated posteriorly.

subjects were implanted by the same surgeon; none of them reported any signs of hip instability and none had suffered a total dislocation postoperatively.

To measure the separation distances between the femoral head and the liner socket, THA subjects were analyzed using fluoroscopy imaging and posterior three-dimensional (3-D) fitting. Based on the 3-D model fitting, axial separation of the femoral head was measured. Additionally, during subluxation, the femoral head appeared to remain in contact with and pivot on the polyethylene liner, potentially creating excessive loads in the region of contact. The average separation detected in the subjects analyzed by Dennis et al. was 3.3 mm (range 1.9–5.2 mm).

The first objective of this investigation was to evaluate the stability of the acetabular liner within the acetabular shell when subjected to physiological cycling loads. The second objective of the experiment was to validate a finite element (FE) model of the cementless modular acetabular component that is going to be described in following sections.

To perform the experiment presented in this study, a complete setup was designed. To satisfy the design requirements of the experiment, both commercially available and custom-made testing instruments were used.

Linear variable differential transducers (LVDT) were used to measure the micromotion and deformation that occurred in the liner when subjected to load. Six LVDT sensors were positioned in contact with the liner at specific locations. Three of the sensors touch the liner perpendicularly to the outer flat surface, while the other three are in contact with the liner perpendicularly from the side, as shown in Fig. 8. To avoid interferences

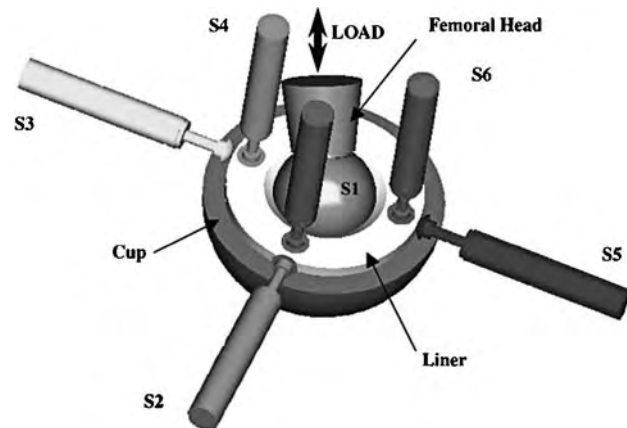


Fig. 8 Relative orientation of the LVDT sensors with respect to the liner.

of the loader with the LVDT sensors, the positioning plate was rotated 30° with respect to the loader.

A custom-made loader was designed to transfer load from the load unit to the acetabular liner in a manner that resembled the way loads are applied in the actual components. A commercially available femoral head (DePuy, Inc., Warsaw, Indiana) was press fitted in a metal rod, which was connected to the testing machine load cell.

Experiment Setup

The following commercially available (DePuy, Inc., Warsaw, Indiana) cementless modular acetabular implant components were tested:

- Seven UHMWPE liners, 28 mm inner diameter and 54 mm outer diameter
- Two acetabular cups, 66 mm outer diameter
- One femoral head, 28 mm outer diameter

The experimental procedure was as follows. Upon positioning the acetabular liner into the acetabular shell using the specific tool designed for such purpose, the latter is screwed to the holding/positioning jig from its bottom. The positioning jig was designed to hold the acetabular shell at a constant angle of 45° with respect to the horizontal, simulating a 45° inclination angle. For this experiment the anteversion angle was 0°. Two acetabular shells (66 mm) were used during the experiments. The first had only a polar fenestration and the second had polar fenestration and three screw holes. At this point, the sensor positioning plate containing six calibrated LVDT sensors was attached to the positioning jig. The positioning plate must be screwed tightly to minimize motion. The metal tip of each

LVDT sensor was calibrated at zero position when it came in contact with the liner.

The femoral head was inserted into the liner socket and adjusted until a perfect conformity was achieved; this was done without applying significant load on the liner to eliminate initial micromotion and deformations of the liner.

FE Model and Validation

A 3-D 3300-element FE model of the femoral head, liner, and acetabular cup was developed based on the CAD previously created in Pro-E (DePuy, Johnson and Johnson; PINNACLE) using ANSYS/LS-DYNA (Fig. 9) shows, respectively, the FE models for the acetabular cup, liner, and femoral head used in the FE explicit analysis. The three components were put in contact and an assembly was created. The diameter of the articulating surface femoral head/liner measured 28 mm while the liner back-side diameter measured 54 mm; a 12 mm diameter polar fenestration was considered in the metal acetabular cup. The acetabular cup was positioned with 45° inclination and 25° anteversion with respect to the reference coordinate system.^[26]

The FE mesh of the acetabular cup, liner, and femoral head consisted of eight-node hexahedron solid elements. Owing to the hemispherical shape of the contacting areas between all the components, two main types of mapped mesh can be performed, polar and nonpolar mesh. A polar mesh design was selected to be the one that better solves the nodal “penetration” problem. The pattern of mesh is the same in all the contacting areas (outer surface of the femoral head/liner front side and liner back-side and acetabular cup) so that the nodes placed in the contacting surfaces do not interfere in the target area. When meshing 3-D solids, the elements of the mesh that cannot be seen

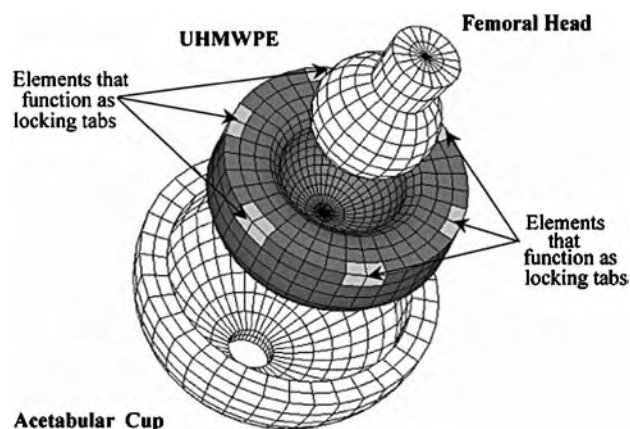


Fig. 9 FE element models of the femoral head, liner, and acetabular cup. The model shows the nodes that were constrained to simulate the liner-locking tabs.

from the outside may have irregular shapes. Owing to the axial symmetry of the liner, acetabular cup, and femoral head, a cross section of those was first meshed. Planar elements were then used (shell 163), having perfect control of the shape of the elements. The major problem of polar mesh design occurs in the change of shape of the elements placed in the poles, where the hexahedron elements are converted to tetrahedron elements.^[27] Singularities occur in tetrahedron elements when the angle between adjacent facets is smaller than a certain angle. In this case, a singularity nonconvergence problem may occur. Material properties of the liner were based on the information obtained in the available literature, Young's modulus = 975 MPa, Poisson's ratio = 0.46.^[28]

It is important to note that the definition of contacts in an explicit analysis differs from the way they are usually defined in an implicit analysis. In the latter analyses, contact is represented by specific contact elements. For explicit analyses there are no contact elements. The definition of the contact is done defining the contact surfaces instead. Also, the type of contact between surfaces and some other parameters related with the contact type must be defined.

To adequately describe the interaction between the implant components, an automatic surface-to-surface contact was selected. This type of contact is established when a surface of one body penetrates the surface of another body (such as the case of the femoral head impacting against the liner). Also, this type of contact is the most effective when dealing with bodies of arbitrary geometry.

The frictional coefficient used for contact (μ_c) is determined from the static friction coefficient, the dynamic friction coefficient, and the exponential decay coefficient. The frictional coefficient is assumed to be dependent on the relative velocity of the surfaces in contact (V_{rel}). Different values for the friction coefficient between the metal and UHMWPE components were studied. The values ranged from 0.083 to 0.2. The friction coefficient data was based on the information obtained through two ways: published literature in the field of biomechanics of human prosthesis,^[28] and general literature regarding the material properties of UHMWPE^[29] when subjected to friction experiments against different materials and contacting shape surfaces. The minimum value found in the published literature^[30] was $\mu = 0.083$ and a maximum value of $\mu = 0.2$.

$$\mu_c = FD + (FS - FD)e^{-DC \times V_{rel}} \quad (2)$$

A similar loading profile to the one used in the experiment setup was created in ANSYS/LS-DYNA to validate the experimental results in terms of stability/micromotions of the liner. The maximum force applied

over the liner during the experiment was 10,000 N. Because there is no evidence that this force can actually be achieved in the human body (even during worst case scenarios such as falling situations), it did not seem necessary to reach that force to validate the FE. At the center of gravity of the femoral head, a dynamic load was applied, ranging from 0 to 1775 N, applied linearly. Based on previous studies of the gait model describing the hip forces and orientations,^[3,30] the value of 1775 seemed perfectly reasonable to be reached in most cases during normal activities.

To reproduce the experimental conditions as closely as possible, both the orientation of the applied load (vertical) and the position of the acetabular component (45° anteversion) were the same in the FE model as well as in the experiment setup. The values ranged from 0.083 to 0.2. Taking the minimum value as a reference, comparisons of the results obtained in the FE model and the experiment, in terms of micromotions, stresses, and deformations of the liner were performed. The friction coefficient value that better fitted both results was selected.

The main steps are as follows:

- *Apply loads and obtain the solution.* Unlike implicit analyses, all loads in an explicit analysis must be time dependent in nature. For this reason, all loads are applied using a pair of array parameters, one corresponding to the time and the other one corresponding to the loading condition.
 - *Review the results.* To review the results obtained in LS-DYNA the postprocessor available in ANSYS was used. Reviewing the results implies getting the deformation, stresses, strains, etc. of nodes and elements. Two different explicit loading profiles were applied to the prosthesis and the consequences over the liner were analyzed. Micromotions, maximum stresses, and deformations were measured.
1. *Gait Loading Profile.* To compute the behavior of the liner when subjected to full gait cycle, a gait loading profile was created and applied to the FE model described and validated. Perfect conformity between the liner and the femoral head was assumed. Gait model data was based on the published literature of hip muscle forces during gait.^[2,31,43] A time–force array was created based on the force data and input in the FE model. The liner-locking mechanism as well as the conformity between contacting surfaces follow the same pattern as described in the FE model (Fig. 10). The resultant forces applied at the center mass of the femoral head are shown in Fig. 11. The analysis was performed using LS-DYNA.
 2. *Impact Loading Profile (Subluxation).* Subluxation occurring in patients with THA has been

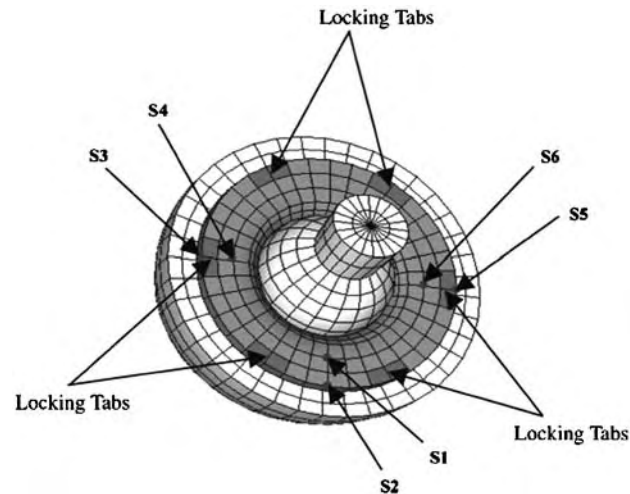


Fig. 10 Relative positions of the liner-locking tabs and the nodes that represent contacting points of the LVDT sensors.

studied *in vivo* using fluoroscopy techniques.^[25] Subluxation is studied in terms of the separation of the femoral head and the liner during abduction/adduction activities. The separation values ranged from $\Delta y = 0$ to 5.3 mm.

Most of the data provided highlight the axial separation (Δy , Fig. 12). Therefore, the values for the radial separation (Δx , Fig. 12) were calculated based on the condition that a portion of the femoral head always remains in contact with the liner. Δx values were calculated using simple trigonometric equations, ranging from $\Delta x = 0$ to 1.1 mm. The different subluxation cases with the corresponding values of Δx and Δy separation are shown in Table 3.

As depicted by Dennis et al.^[25] the average axial separation detected over all the studied patients was $\Delta y = 3.3$ mm. For this value, a radial separation of $\Delta x = 0.3$ mm was calculated. To create a pattern of study, 12 subluxation cases were considered with a displacement increment of 0.5 mm. An extrapolation was made to get results for the average separation case.

A loading profile was developed to simulate the effect of the relocation of the femoral head within the liner after subluxation occurred. For all subluxation cases, according to Mow et al. a maximum peak force of 2132 N was applied at an angle of 23.4° lateral and 5.7° posterior.^[32–35] The load was applied at the center of the dislocated femoral head. In all subluxation cases the time interval considered for the duration of the load was $\Delta t = 0.1$ sec. The magnitude and orientation of the input force were assumed constant during the analysis, owing to the small duration of the impact.

Certain conditions must be given as an input while performing a dynamic analysis, such as the initial

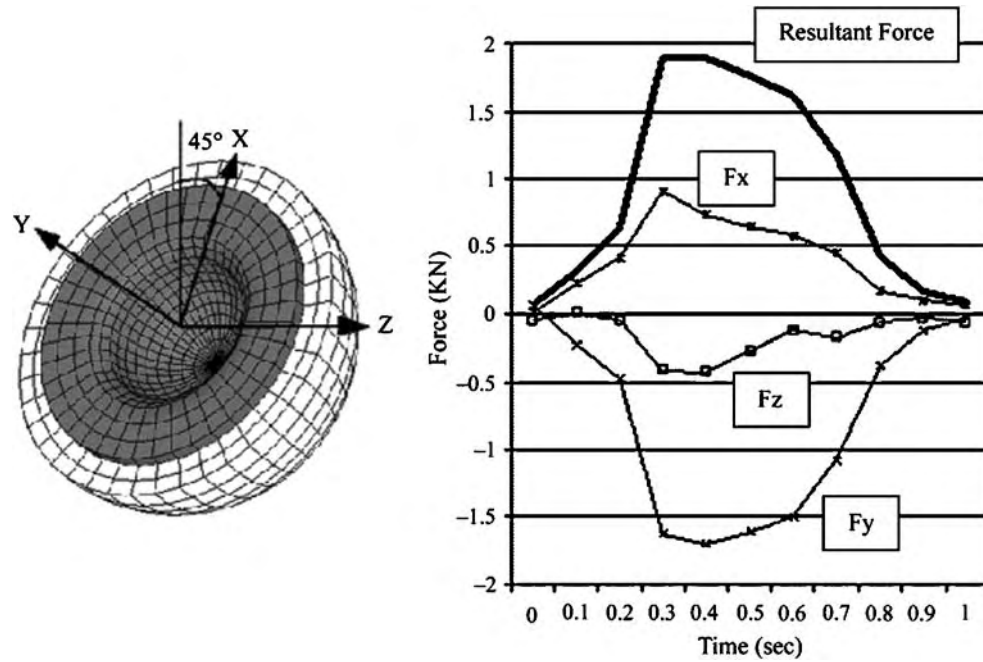


Fig. 11 Gait loading profile applied to the femoral head for the gait analysis.

velocity and acceleration of all the bodies involved in the problem. The analysis time Δt starts when the femoral head is moving backwards to relocate its position within the liner. Therefore, the initial velocity and acceleration was considered to be negligible. The same constrain conditions as described in the FE model were applied for this loading profile.

Experiment Results: Liner Micromotion and FE Model Validation

The maximum physiological loads experienced in the hip are within a range of 1700 to 2200 N. As described before in this study, the maximum load applied to each liner specimen during the validation experiment was 10,000 N, which is beyond the attainable force in the body. The objective of applying such load profiles was to determine the local deformation as well as the range

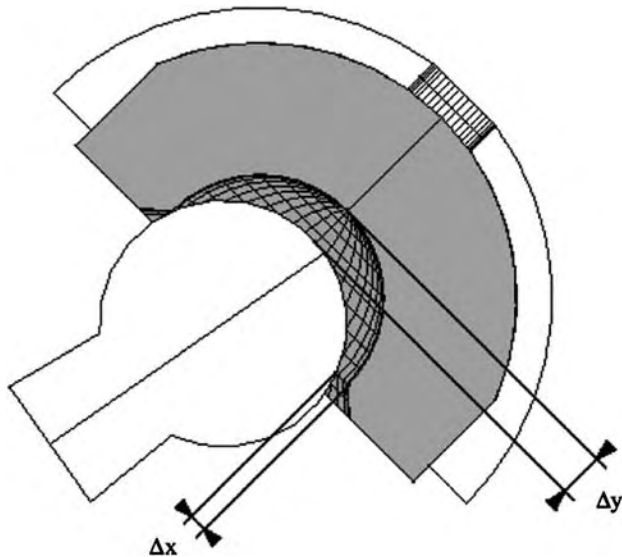


Fig. 12 Separation of the femoral head occurred during subluxation. Δx represents the radial separation and Δy represents the axial separation.

Table 3 Subluxation cases with the corresponding values of radial (Δx) and axial separation (Δy)

Subluxation case	Δx , radial separation (mm)	Δy , axial separation (mm)
dis-1	0	0
dis-2	0	0.5
dis-3	0.036	1
dis-4	0.081	1.5
dis-5	0.144	2
dis-6	0.225	2.5
dis-7	0.325	3
dis-8	0.445	3.5
dis-9	0.584	4
dis-10	0.743	4.5
dis-11	0.923	5
dis-12	1.126	5.5



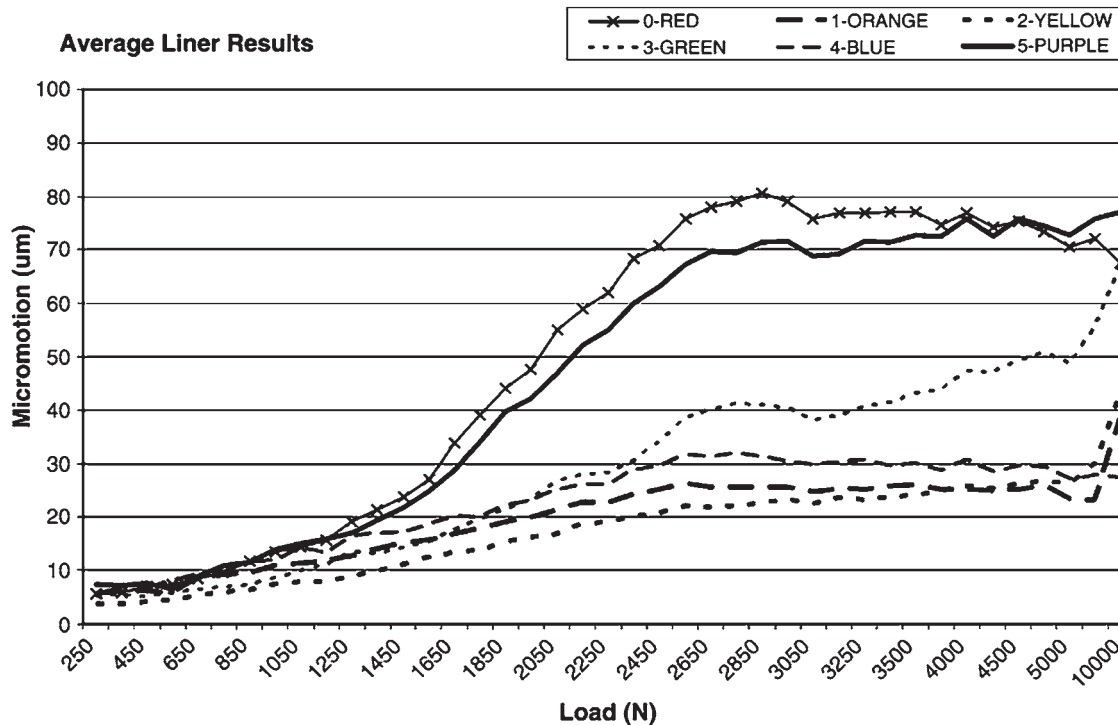


Fig. 13 Average micromotion values of the seven liner specimens analyzed for each LVDT sensor. The maximum micromotion values for the validation load range are S1 = 75.85 μm , S2 = 38.41 μm , S3 = 42.65 μm , S4 = 67.32 μm , S5 = 31.94 μm , and S6 = 77.02 μm .

at which there will be a jump in the deformation curve corresponding to permanent deformation.

The average results in terms of micromotions for the seven liners analyzed are shown in Fig. 13. To perform the FE model, validation loads were applied in the same fashion as it was done in the experiment. The force values applied in the FE model was restricted only to the force range from 0 to 1775 N. According to the experimental setup, the micromotion results corresponded to six specific points, where the sensors' tips came in contact with the liner. Those micromotion values did not depict the rigid body motion of the liner, but they serve as a reference when compared with the results obtained from the FE model. Those six points were identified in the FE model.

Because the LVDT sensors only measure displacements along their longitudinal axes, those directions were identified with simple trigonometric calculation in the FE model through superposition of the experimental axes. The micromotions (displacements) were computed in the FE model of the nodes positioned at the reference points, along the calculated direction. The micromotion results obtained in the FE analysis for the two different friction coefficients analyzed are shown in Fig. 14.

Gait Loading Results

During gait the liner is subjected to certain loading conditions that can lead to relative motion of the liner

within the acetabular cup (micromotion of the liner's back-side surface) and relative motion of the femoral head within the liner (micromotion of the liner's front side surface). The maximum micromotions occurring during a full gait cycle were computed for the front- and back-side surfaces of the liner at each of the time steps of the cycle. Table 4 shows those results.

To determine the areas where such micromotion occurred, a vector representation of the micromotion tracks at each step time is shown in Fig. 15, for back-side surfaces of the liner. In this figure, the micromotion tracks are represented as a series of drawings of the liner at each step time, accompanied with a general drawing of the FE model of the acetabular component to understand the spatial position of the liner with respect to the applied load. The arrows that appear in each liner represent the micromotion tracks at each element centroid.

Subluxation Loading Results

Twelve different subluxation cases were studied. Each subluxation case was defined by a certain separation of the femoral head with respect to the liner. The maximum stresses and micromotions occurred in the liner when the femoral head relocations within the liner (impact situation) were computed.

The period of time (Δt) for the application of the relocating load used in all subluxation cases was

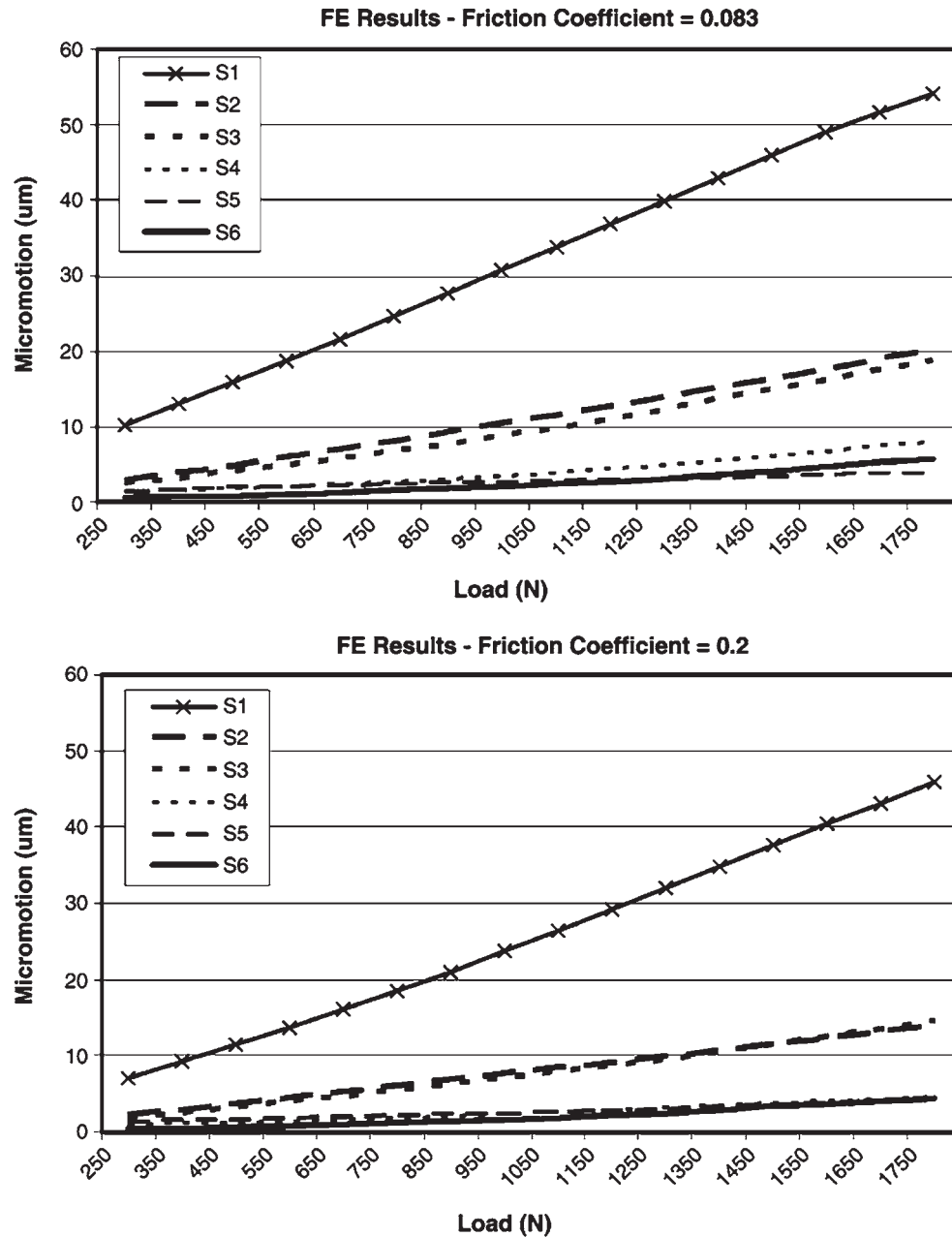


Fig. 14 Micromotion results obtained in the FE model for two different friction coefficients.

viewed to be reasonable to capture the complete effects of impact. The average subluxation case ($\Delta y = 3.3$ mm, $\Delta x = 0.3$ mm) achieved a maximum stress of 44 MPa. In addition, the average stress results achieved for the average dislocation case was 17.5 MPa at $t = 0.015$ sec.

Fig. 16 represents the maximum stresses that occurred in the liner during the relocation process for the average subluxation case (total time $t = 0.1$ sec). The stresses occurred in the liner at three different time intervals of the average subluxation case. The first time step ($t = 0.015$ sec) represents the stresses in the instant the femoral head first impacts with the liner. The second time step ($t = 0.02$ sec) represents the bouncing of

the femoral head after the first impact. The third time ($t = 0.03$ sec) interval represents the second impact of the femoral head with the liner.

Computing the displacement of the nodes in the back-side surface of the liner, with respect to the nodes of the inner surface of the acetabular cup, allowed us to obtain the micromotion values and the wear vectors during impact relocation. The wear vector orientation are based on the micromotion tracks. The maximum micromotion achieved in the average subluxation case was 0.324 mm.

Similar to the gait loading case, it is important to consider the areas where the maximum micromotion

Table 4 Maximum micromotions occurred in the front- and back-side surface of the liner during gait loading profile

Load step	Time (sec)	Front-side micromotion (μm)	Back-side micromotion (μm)
1	0	31.2	12.1
2	0.1	33.0	23.7
3	0.2	39.0	34.2
4	0.3	82.4	81.7
5	0.4	108.72	92.2
6	0.5	138.78	97.1
7	0.6	154.8	102.8
8	0.7	169.38	117.58
9	0.8	180.45	121.22
10	0.9	178.11	107.92
11	1	175.95	101.22

occurred during the relocation process of the femoral head after subluxation. The way to observe graphically those areas is through the use of tracks (vectors) that represent the micromotion magnitude and orientation for each time step. Liner micromotion corresponds to the fourth subluxation case ($\Delta x = 0.081$ mm and $\Delta y = 1.5$ mm) at time $t = 0.02$ sec, moment at which the femoral head strikes the polyethylene liner for the first time. To have a better idea of the different solicitations to which the liner is subjected after subluxation, a comparison between the micromotion tracks in the subluxation analysis and the micromotion tracks in the gait analysis were also made. Fig. 14 shows an example of the micromotion tracks in the back-side of the liner's surfaces during gait ($t = 0.4$ sec).

During relocation, the micromotion tracks follow a steady path, while during gait it can be observed that at the points where the micromotion are higher, the micromotion tracks do not follow any specific path as evident as during relocation. The main reason for this behavior may be that during the relocation of the femoral head with the liner the impact force is much higher than the contacting forces that occur afterwards. Also, the orientation of the impact force is almost constant during the entire relocation process; therefore micromotion during relocation is mainly owing to one resultant force, which is constant in magnitude and orientation. During gait, the resultant force involved in the contact of the femoral head and the liner varies in magnitude and orientation. Therefore, there is not a main contributing load (in terms of magnitude and orientation) responsible for such micromotion tracks, creating an unsteady micromotion tracks distribution over the liner's surfaces.

FURTHER DISCUSSION OF THE RESULTS

Based on the micromotion values obtained experimentally for each liner specimen (Fig. 12), it can be observed that the sensors positioned perpendicularly to the outer flat surface of the liner (as shown in Fig. 16) recorded the most significant micromotion. The latter makes sense, because this is the direction along which the liner is less constrained.

The three sensors positioned perpendicularly to the side of the liner (Fig. 16) recorded very little micromotion. These micromotions were produced because of local deformations from the bulging effect of the polyethylene when subjected to compressive loads. In most of the cases the result obtained in the FE model, for both friction coefficients, are between the maximum and minimum values obtained in the experiment. Considering the average values for the experiment results (based on the maximum and minimum micromotion values), we can conclude that the FE model friction coefficient that better fit with those values is $\mu = 0.083$. The latter was the friction coefficient used for the remaining FE analysis in which contact was involved.

$$\begin{aligned}
 & \begin{bmatrix} S2_{T_{i+1}} & 0 & S6_{T_{i+1}} \\ 0 & S4_{T_{i+1}} & 0 \\ S1_{T_{i+1}} & S3_{T_{i+1}} & S5_{T_{i+1}} \\ 1 & 1 & 1 \end{bmatrix} \\
 & = \begin{bmatrix} \cos \theta_y & 0 & \sin \theta_y & T_x \\ \sin \theta_x \cos \theta_y & \cos \theta_x & -\sin \theta_x \cos \theta_y & T_y \\ -\cos \theta_x \sin \theta_y & \sin \theta_x & \cos \theta_x \sin \theta_y & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
 & \times \begin{bmatrix} S2_{T_i} & 0 & S6_{T_i} \\ 0 & S4_{T_i} & 0 \\ S1_{T_i} & S3_{T_i} & S5_{T_i} \\ 1 & 1 & 1 \end{bmatrix} \quad (3)
 \end{aligned}$$

$$e = \cos \left[\frac{1}{2} (\theta_x + \theta_z) \right] \cos \left(\frac{1}{2} \theta_y \right) \quad (4)$$

Using the transformation matrix shown in Eq. (3) (T_i represents the position of the LVDT sensor tips at time i , and T_{i+1} represent the position of the LVDT sensor tips at time $i+1$), the rigid solid movements of the liner within the acetabular cup were computed (three rotations θ_x , θ_y , and θ_z and three translations T_x ,

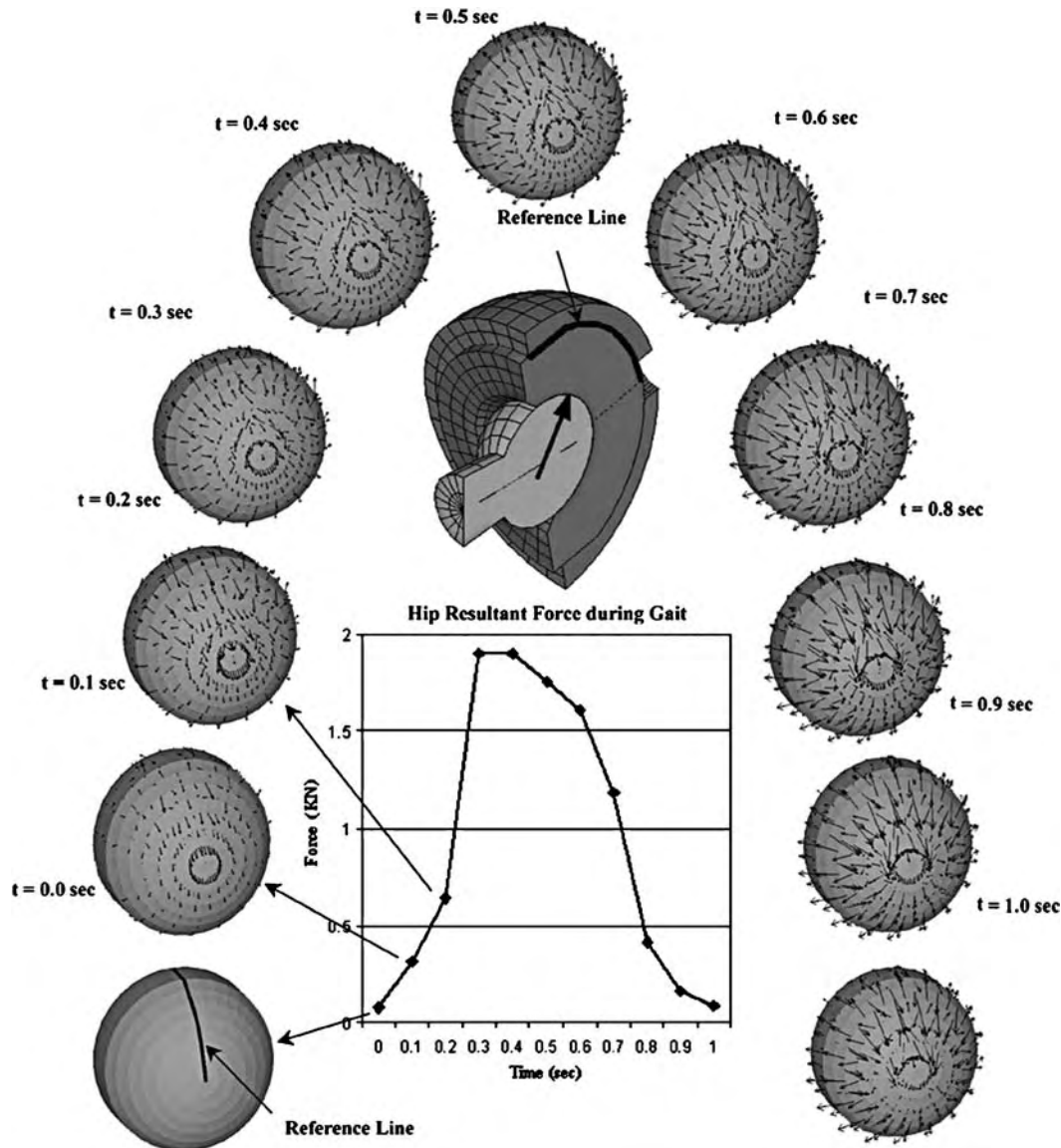


Fig. 15 Back-side liner micromotions vectors when subjected to gait loading profile. The graph represents the resultant force applied to the femoral head during gait. Reference line represents a reference on the liner position with respect to the resultant force.

T_y , and T_z). Using this transformation, Euler angles [Eq. (4)] and the vertical displacement of the liner were computed. Table 5 shows the Euler angle value, averaged for the seven-liner specimen, for the load range used in the validation process (0–1750 N). Table 5 also shows the displacement of the outer flat surface of the liner, which was computed based on the values given by Eq. (3).

From Table 5 it can be observed that the maximum rotation angle of the liner within the acetabular cup induced by the applied load is in every case lesser than 0.1° . Furthermore, the maximum vertical displacement of the outer surface of the liner is smaller than $30\ \mu\text{m}$. Both results demonstrate that the liner remains stable for the load range studied.

The maximum micromotion and stresses obtained during the relocation of the femoral head after subluxation appeared to be much higher than the ones obtained for the gait analysis. The maximum load applied in the subluxation case is a 12% larger than the maximum load applied during gait. However, the maximum micromotion during subluxation is almost three times larger than the one obtained during gait. Similarly, the maximum stress value during subluxation is almost seven times the stress value obtained during a full cycle of gait. Therefore, it can be speculated based on the results obtained, that during gait, the femoral head and the liner remain in contact at all times. However, during the relocation of the femoral head after subluxation, peak forces are achieved owing to the impacting conditions.

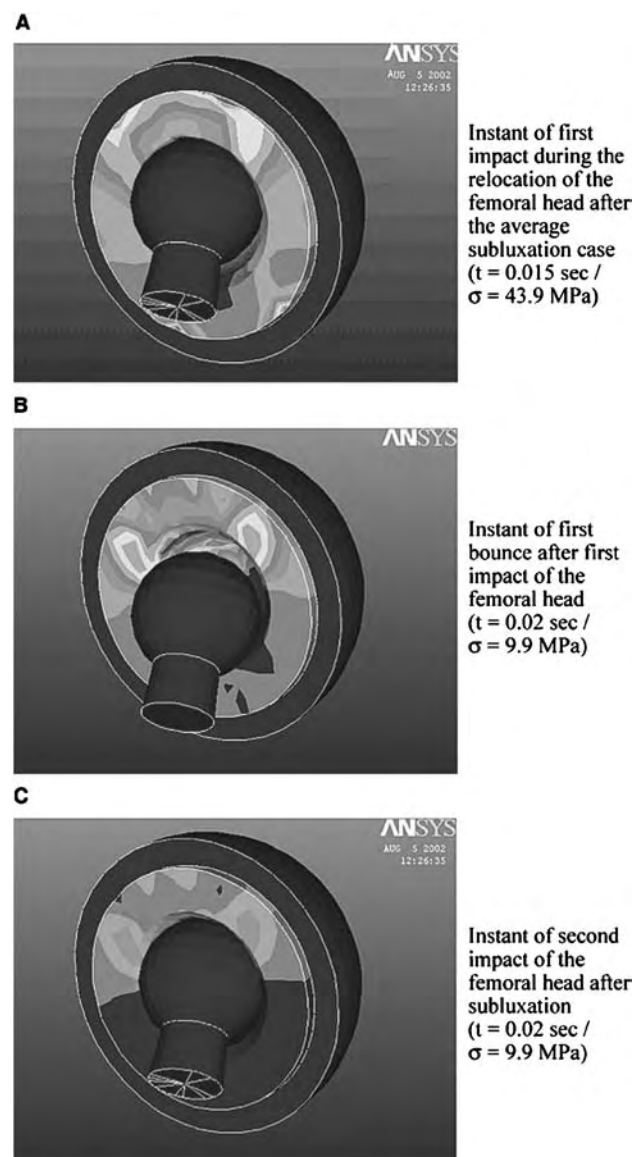


Fig. 16 Stresses occurred in the liner at three different time intervals of the average subluxation case. The first time step ($t = 0.015 \text{ sec}$) represents the stresses in the instant the femoral head first impacts with the liner. The second time step ($t = 0.02 \text{ sec}$) represents the bouncing of the femoral head after the first impact. The third time ($t = 0.03 \text{ sec}$) interval represents the second impact of the femoral head with the liner.

Past studies have described the loads applied to the hip joint in different activities such as gait, stair climbing, falling, and running. There is very little information on the loads generated in the relocation of the femoral head after subluxation. As previously described, the loading condition used for the subluxation cases correspond to a peak force of 2132 N. This is the force achieved in the hip joint at the moment when the leg strikes the ground during walking.

Table 5 Average rotation angle and vertical displacement of the liner within the acetabular cup

Load (N)	Euler angle ($^{\circ}$)	Rotational displacement (μm)	Vertical displacement (μm)
250	0.0141	8.11	1.02
350	0.0151	8.69	2.35
450	0.0163	9.40	3.29
550	0.0168	9.68	4.02
650	0.0216	12.47	5.74
750	0.0233	13.41	7.31
850	0.0269	15.51	8.84
950	0.0309	17.78	10.63
1050	0.0328	18.88	12.42
1150	0.0349	20.11	14.38
1250	0.041	23.60	16.18

CONCLUSIONS

This article provided an overview into the biomechanics of the hip, focusing on the THA and the engineering and clinical problems associated with such a procedure. The clinical issues such as implant loosening, wear, impingement, and dislocation are described and illustrations are provided for each case.

This article also provides the modeling of the liner/cup interface and bone interface using the FE method and shows how such a tool can be used and validated with experimental data.

ACKNOWLEDGMENTS

The partial support of DePuy Orthopaedic is greatly appreciated. We also extend our gratitude and thanks to Francisco Romero at Biomechanics Research Laboratory (BRL) at UIC. We also like to thank Ivan Zivkovic for his technical assistance and the reviewers for their comments and valuable support.

ARTICLE OF FURTHER INTEREST

Tissue Engineering of Bone

REFERENCES

1. Bergmann, G.; Graichen, F.; Rohlmann, A. Hip joint loading during walking and running, measured in two patients. *J. Biomech.* **1993**, *26*, 969–990.
2. Bergmann, G.; Graichen, F.; Rohlmann, A. Hip joint forces during loading carrying. *J. Clin. Ortho.* **1997**, *335*, 190–210.

3. Barrack, R.L.; Folgueras, A.; Munn, B.; et al. Pelvis lysis and polyethylene wear at 5–8 years in an uncemented total hip. *Clin. Orthop.* **1997**, *355*, 211–217.
4. Engh, C.A.; O'Conner, D.; Jasty, M.; et al. Quantification of implant micromotion, strain shielding, and bone resorption with porous-coated anatomic medullary locking femoral prostheses. *Clin. Orthop.* **1992**, *285*, 12–29.
5. Maquet, P.G.J. *Biomechanics of the Hip*; Springer-Verlag: Berlin, Germany, 1985.
6. Alexander, E.J.; Bregler, C.; Andriachi, T.P. Limb segment pose from range data streams through homogeneous factorization. *Proceedings of 2001 ASME Bioengineering Conference* **2001**, *50*.
7. Davy, D.T.; Kotzar, G.M.; Brown, R.H. Telemetric force measurements across the hip after total arthroplasty. *J. Bone Joint Surg. Am.* **1988**, *70* (1), 45–50.
8. Mulroy, R.D.; Harris, W.H. The effect of improved cementing techniques on component loosening in hip replacement: an 11-year radiographic review. *J. Bone Joint Surg.* **1990**, *72B*, 757–760.
9. Agins, H.J.; Alcock, N.W.; Bansal, M.; et al. Metallic wear in failed titanium-alloy total hip replacements. A histological and quantitative analysis. *J. Bone Joint Surg.* **1988**, *70* (3), 347–356.
10. Archibeck, M.J.; Jacobs, J.J.; Black, J. Alternate bearing surfaces in total joint arthroplasty: biologic considerations. *Clin. Orthop.* **2000**, *1* (379), 12–21.
11. Arthritis Foundation, 2004 <http://www.arthritis.org/>.
12. Bellemans, J. Osseointegration in porous coated knee arthroplasty. The influence of component coating type in sheep. *Acta Orthop. Scand. Suppl.* **1999**, *288*, 1–35.
13. Buma, P.; van Loon, P.J.; Versleyen, H.; et al. Histological and biomechanical analysis of bone and interface reactions around hydroxyapatite-coated intramedullary implants of different stiffness: a pilot study on goat. *Biomaterials* **1997**, *18* (18), 1251–1260.
14. Charnley, J. *Low Friction Arthroplasty of the Hip*; Springer Verlag: New York, NY, 1978.
15. Charnley, J. *Low Friction Arthroplasty, Theory and Practice*; Springer Verlag: New York, NY, 1979.
16. Charnley, J. *Acrylic Cement in Orthopaedic Surgery*; E. and S. Livingstone: Edinburgh, Scotland, 1970.
17. Cobb, T.K.; Morrey, B.F.; Ilstrup, D.M. Effect of the elevated-rim acetabular liner on loosening after total hip arthroplasty. *J. Bone Joint Surg. Am.* **1997**, *79* (9), 1361–1364.
18. Davy, D.T.; Kotzar, G.M.; Brown, R.H.; et al. Telemetric force measurements across the hip after total arthroplasty. *J. Bone Joint Surg. Am.* **1998**, *70* (1), 45–50.
19. Dorr, L.D.; Wan, Z.; Cohen, J. Hemispheric titanium porous coated acetabular component without screw fixation. *Clin. Orthop.* **1998**, *351*, 158–168.
20. El Maraghy, A.W.; Schemitsch, E.H.; Waddell, J.P. Acetabular blood flow during total hip arthroplasty. *Can. J. Surg.* **2000**, *43* (3), 197–201.
21. Engh, C.A. *Surgical Technique: Anatomical Medullary Locking Total Hip System with Porocoat*; DePuy Orthopaedics: Warsaw, IN, 1992.
22. Friedman, R.J.; Black, J.; Galante, J.O.; et al. Current concepts in orthopaedic biomaterials and implant fixation. *J. Bone Joint Surg. Am.* **1993**, *75* (7), 1086–1109.
23. Goldring, S.R.; Schiller, A.L.; Roelke, M.; et al. The synovial-like membrane at the bone-cement interface in loose total hip replacements and its proposed role in bone lysis. *J. Bone Joint Surg. Am.* **1983**, *65* (5), 575–584.
24. Goodman, S.B.; Song, Y.; Chun, L.; et al. Effects of TGF beta on bone ingrowth in the presence of polyethylene particles. *J. Bone Joint Surg. Br.* **1999**, *81* (6), 1069–1075.
25. Jaramaz, B.; Nikou, C.; DiGioia, A.M. Sensitivity of impingement limits to error in cup placement. *Proc. Ann. Meeting ORS* **1998**.
26. Jasty, M.; Burke, D.; Harris, W.H. Biomechanics of cemented and cementless prostheses. *Chir. Organi. Mov.* **1992**, *77* (4), 349–358.
27. Kapandji, I.A. *The Physiology of the Joints*; Churchill Livingstone: Edinburgh, Scotland, 1987.
28. Kienapfel, H.; Sprey, C.; Wilke, A.; Griss, P. Implant fixation by bone ingrowth. *J. Arthroplasty* **1999**, *14* (3), 355–368.
29. Learmonth, I.D.; Lee, M.B.; Bitounis, V.C. Acetabular replacement with the cementless Harris-Galante porous cup. *J. Bone Joint Surg.* **2001**, *83B* (suppl II), 143.
30. Leunig, M.; Hertel, R. Thermal necrosis after tibial reaming for intramedullary nail fixation. A report of three cases. *J. Bone Joint Surg. Br.* **1996**, *78* (4), 584–587.
31. Macdonald, W.; Carlsson, L.V.; Charnley, G.J.; et al. Inaccuracy of acetabular reaming under surgical conditions. *J. Arthroplasty* **1999**, *14* (6), 730–737.
32. Malchau, H.; Herberts, P.; Ahnfelt, L. Prognosis of total hip replacement in Sweden. Follow-up of 92,675 operations performed 1978–1990. *Acta Orthop. Scand.* **1993**, *64* (5), 497–506.
33. Manley, M.T.; Capello, W.N.; D'Antonio, J.A.; et al. Fixation of acetabular cups without cement in total hip arthroplasty. A comparison of three different implant surfaces at a minimum duration of follow-up of five years. *J. Bone Joint Surg. Am.* **1998**, *80* (8), 1175–1185.
34. Maquet, P.G.J. *Biomechanics of the Hip*; Springer-Verlag: Berlin, Germany, 1985.
35. Morrey, B.F. Instability after total hip arthroplasty. *Orthop. Clin. N. Am.* **1992**, *23*, 237–248.
36. McCollum, D.E.; Gray, W.J. Dislocation after total hip arthroplasty: causes and prevention. *Clin. Orthop.* **1990**, *261*, 159–170.
37. Woo, R.Y.G.; Morrey, B.F. Dislocation after total hip arthroplasty. *J. Bone Joint Surg. Am.* **1982**, *64*, 1295–1306.
38. Lombardi, A.V.; Mallory, T.H.; Dennis, D.A.; Komistek, R.D.; Fada, R.A.; Northcut, E.J. An in vivo determination of total hip arthroplasty pistoning during activity. *J. Arthroplasty* **2000**, *15*, 702–709.
39. Dennis, D.A.; Komistek, R.D.; Northcut, E.J.; Ochoa, J.A.; Ritchie, A. In vivo determination of hip joint separation and the forces generated due to impact loading conditions. *J. Biomech.* **2001**, *34*, 623–629.
40. McKellop, H.A.; Campbell, P.; Park, S.H.; et al. The origin of submicron polyethylene wear debris in total hip arthroplasty. *Clin. Orthop.* **1995**, *311*, 3–20.
41. Moskowitz, R.W. *Osteoarthritis, Diagnosis and Medical/Surgical Management*, 2nd Ed. Ed.; W.B. Saunders: Philadelphia, 1992.



42. Müller, C.H.R.; Rahn, B.A.; Pfister, U.; et al. Intramedullary pressure, strain on the diaphysis and increase in cortical temperature when reaming the femoral cavity—a comparison of blunt and sharp reamers. *Injury* **1993**, *24* (3), S22–S30.
43. Nakata, K.; Ohzono, K.; Masuhara, K.; et al. Acetabular osteolysis and migration in bipolar arthroplasty of the hip: five- to 13-year follow-up study. *J. Bone Joint Surg. Br.* **1997**, *79* (2), 258–264.
44. Teoh, S.H.; Chan, W.H.; Thampuran, R. An elastoplastic finite element model for polyethylene wear in total hip arthroplasty. *J. Biomech.* **2002**, *35*, 323–330.
45. Kurtz, S.M.; Edidin, A.A.; Bartel, D.L. The role of backside polishing, cup angle, and polyethylene thickness on the contact stresses in metal-backed acetabular components. *J. Biomech.* **1997**, *30*, 639–642.
46. Northcut, E.J.; Komistek, R.D.; Dennis, D.A.; et al. “In vivo” determination of hip joint separation that may lead to impulsive loading conditions. *J. Biomech.* **2001**, *34* (5), 623–629.
47. Symplastics Inc. Ramex-UHMEPE. Ultra High Molecular Weight Polyethylene Material Properties.
48. Brand, R.A.; Pedersen, D.R.; Davy, D.T. Comparison of hip force calculations and measurements in the same patient. *J. Arthrop.* **1994**, *9*, 45–51.
49. Morlock, M.; Sheneider, E.; Blumh, A.; Vollmer, M.; Bergmann, G.; Muller, V.; Honl, M. Duration and frequency of every day activities in total hip patients. *J. Biomech.* **2001**, *34*, 873–881.
50. Bergmann, G.; Deuretzbacher, G.; Heller, M.; Graichen, F.; Rohlmann, A.; Strauss, J.; Duda, G.N. Hip contact forces and gait patterns from routine activities. *J. Biomech.* **2001**, *34*, 859–871.
51. Ipavec, M.; Brand, R.A.; Pedersen, D.R.; Mavcic, B.; Kralj-Iglic, V.; Iglic, A. Mathematical modeling of stress in the hip during gait. *J. Biomech.* **1999**, *32*, 1229–1235.
52. Herberts; Malchau; Garellick. The Annual Report of the Swedish National Hip Arthroplasty Register; <http://www.jru.orthop.gu.se/2003> (accessed).
53. Ong, K.; Gunsallus, K.; Ryan, M.; et al. Comparison of the initial stability of hemispherical and elliptical acetabular cups. *Proceeding of 2001 ASME Bioengineering Conference* **2001**, *50*.
54. Pellicci, P.M.; Tria, A.J.; Garvin, K.L. Orthopaedic knowledge update. *J. Am. Acad. Ortho. Surg.* **1999**.
55. Perona, P.G.; Lawrence, J.; Paprosky, W.G.; et al. Acetabular micromotion as a measure of initial implant stability in primary hip arthroplasty. *J. Arthroplasty* **1992**, *74*, 537–547.
56. Pilliar, R.M.; Lee, J.M.; Maniopoulos, C. Observations on the effect of movement on bone ingrowth into porous surfaced implants. *Clin. Orthop.* **1986**, *208*, 108–113.
57. Ranawat, C.S.; Peters, L.E.; Umlas, M.E. Fixation of the acetabular component: the case for cement. *Clin. Orthop.* **1997**, *344*, 207–215.
58. Saleh; Thongtrangan; Schwarz Osteolysis: medical and surgical approaches. *Clin. Orthop.* **2004**, *427*, 138–147.
59. Schmalzried, T.P.; Kwong, L.M.; Jasty, M.; et al. The mechanism of loosening of cemented acetabular components in total hip arthroplasty: Analysis of specimens retrieved at autopsy. *Clin. Orthop.* **1992**, *274*, 60–78.
60. Schmalzried, T.P.; Guttman, D.; Grecula, M.; Amstutz, H.C. The relationship between the design, position, and articular wear of acetabular components inserted without cement and the development of pelvic osteolysis. *J. Bone Joint Surg.* **1994**, *76* (A), 677–688.
61. Spears, I.R.; Pfeleiderer, M.; Schneider, E.; et al. Interfacial conditions between a press-fit acetabular cup and bone during daily activities: implications for achieving bone in-growth. *J. Biomech.* **2000**, *33* (11), 1471–1477.
62. Van Lenthe, G.H.; Van Den Bergh, J.P.W.; Hermus, A.R.M.M.; Huijkes, R. The prospects of estimating trabecular bone tissue properties from the combination of ultrasound, dual-energy x-ray absorptiometry, microcomputer tomography, and microfinite element analysis. *J. Bone Miner. Res.* **2001**, *16* (3), 550–555.
63. Walde; Claus; Mohan, et al. Advancement and separation of apex hole eliminators with cementless duraloc 100 cups. A report of nineteen cases. *J. Bone Joint Surg. Am.* **2004**, *86-A* (10), 2251–2256.
64. Wessinghage, D.; Kisslinger, E. Long-term results after cemented total hip arthroplasty in chronic polyarthritis. *Orthopade.* **1998**, *27* (6), 381–391.
65. Willert, H.G. Reaction of the articular capsule to wear products of artificial joint prostheses. *J. Biomed. Mater. Res.* **1977**, *11* (2), 157–164.
66. Wolfe, F.; Zwillich, S.H. The long-term outcomes of rheumatoid arthritis: a 23-year perspective, longitudinal study of the total joint replacement and its predictors in 1,600 patients with rheumatoid arthritis. *Arthritis Rheum.* **1998**, *41* (6), 1072–1082.
67. Won, C.H.; Hearn, T.C.; Tile, M. Micromotion of cementless hemispherical acetabular components: does press-fit need adjunctive screw fixation? *J. Bone. Joint. Surg.* **1995**, *77B*, 484–489.
68. Finkelstein, J.A.; Anderson, G.I.; Waddell, J.P.; et al. A study of micromotion and appositional bone growth to a canine maderporic-surfaced femoral component. *J. Arthroplasty* **1994**, *9* (3), 317–324.
69. Maloney, W.J.; Paprosky, W.; Engh, C.; et al. Surgical treatment of pelvic osteolysis. *Clin. Orthop.* **2001**, *393*, 78–84.
70. Maloney, W.J.; Galante, J.O.; Anderson, M.; et al. Fixation, polyethylene wear, and pelvic osteolysis in primary total hip replacement. *Clin. Orthop.* **1999**, *369*, 157–164.

Hip Joint: Overuse Injuries

Marko Pećina
Ivan Bojanić
Alan Ivković

Department of Orthopaedic Surgery, School of Medicine, University of Zagreb, Zagreb, Croatia

INTRODUCTION

Painful syndrome in the hip joint area is generally considered the most frequent overuse syndrome in some athletic activities, e.g., soccer. When considering the area and anatomic structures affected by overuse in the hip joint region, together with the locations and characteristics of pain, the location might be precise, but it could also be a case of diffuse vague pain in the groin region, small pelvis, and upper leg regions. The term syndrome is fully justified; indeed, the symptoms are numerous and so are the causes of pain in the locations mentioned. To understand the painful overuse syndromes in the hip joint area, one should bear in mind all the muscles that insert or originate in the hip region. These muscles are the abdominal group of muscles, adductors muscles, hamstring muscles, hip external rotator muscles, iliopsoas muscle, rectus femoris muscle, and others (sartorius, pyramidalis, cremaster muscles). In this entry we will describe the most important and the most frequent overuse syndromes in the hip joint area.

GROIN PAIN SYNDROME

The modern medical literature abounds with terms defining pain in the groin region: necrotic osteitis pubis, anterior pelvic joint syndrome, traumatic pubic osteitis, Pierson syndrome, gracilis muscle syndrome, pubic stress symphysis, pubic symphysis osteoarthropathy, symphysis, pubic chondritis, and post-traumatic necrosis of the pubic bone.^[1,2] These terms and expressions have commonly been used to describe adductor tendinitis, rectus abdominis tendinitis, avulsion injuries of the adductor tendons, postoperative changes without infection in the symphysis region, and a number of arthrotic changes. All of the above-mentioned terms clearly show that the painful groin syndrome cannot be attributed solely to one particular anatomic structure in the region but, rather, that the approach to the syndrome, in prevention, diagnosis, or treatment, should be based on the idea of multiple causative factors contributing to development of the

syndrome. It is precisely the approach we have assumed, although we distinguish between painful groin syndrome in a narrow sense and painful groin syndrome in a wider sense, for which “pain in the groin” seems to be a more appropriate term because the pain may arise from different pathological condition in and around the groin region. Painful groin syndrome in a narrow sense primarily implies tendinitis of the adductor muscles—the long adductor muscle and the gracilis muscle in the first instance, as well as the abdominal muscles, especially the rectus abdominis and pyramidalis muscles (see Fig. 1). Long-lasting tendinitis of these muscles may evolve into a general and vague picture of diffuse pain in the pubic bones and the symphysis region.

Groin Strain

This is the painful groin syndrome in a narrow sense. The cause of the painful groin strain lies in the disproportion between the strength of the abdominal wall muscles and the strength of the lower extremity muscles, as well as in an uneven load on all attachments in the symphysis and groin region. Groin strain is an injury or, more precisely, a lesion characteristically occurring in soccer players because the symphysis region is the point of attachment and the point of origin for groups of muscles with different functions, which are especially active during soccer playing. In addition to soccer players, these injuries have also been reported in ice hockey, water polo, handball players, fencers, high jumpers, bowlers, skaters, hurdlers, swimmers, and triathletes. Painful groin syndrome is characterized by pain in the inguinal and lower abdominal region. Often the pain develops gradually, and the athlete is unaware of its connection with any injury. As time goes by, the pain becomes more severe and irradiates into the upper leg adductor zone, the pubic region, and the perineum, spreading toward the hips and the anterior abdominal wall. The pain restricts certain motions during training and the game, especially reducing the speed of the player. Furthermore, the motion performed when entering or exiting a car may cause pain.





Fig. 1 Test for straight abdominal muscles: lift legs from the surface to form a 45° angle.

Similar to other overuse syndromes, painful groin syndrome is also characterized by specific developmental stages. In the most developed stage coughing, sneezing, defecation, urination, and sexual activity also cause pain in the groin. Clinical examination reveals pain during palpation of the pubic bone occurring above the gracilis and adductor longus muscle attachments. When testing adductors, the patient lies on his or her back with slightly abducted lower extremities so that the fist may be inserted between the knees (see Fig. 2).

The patient is then asked to hold the fist tightly by contracting the adductor muscles of the upper leg. The motion causes pain in the typical site of the upper leg adductor muscles attachment on the pubic bone. In tendinitis of the rectus abdominis muscles, pain following palpation appears in the region of the attachment to the ramus superior pubis. In a great number of patients, inspection of the abdominal wall reveals weakness of the oblique abdominal muscle. The so-called sausage and spindle-like prominences above the iliac crest and the inguinal ligament (in the literature commonly referred to as Malgaigne's sign) appearing during contraction of the abdominal muscles are also signs of abdominal wall weakness. Radiographic images of the pelvis, showing pubic bones, symphysis, and both hips, are necessary. These enable the physician to rule out arthrotic changes in the hip, stress fractures of the pubis ramus and pubic symphysis, pelvic avulsion injuries, and femoral stress fractures as causes of pain; it also enables rough analysis of the sacroiliac joints and the caudal segment of the lumbosacral spine (discogenic pain, spondylolysis). Radiologically identified changes in the symphysis and pubic bone may be classified into four stages that correspond to the developmental stages of the painful groin syndrome's clinical picture. Scintigraphy has proved highly valuable in the diagnosis. It helps to identify

stress fractures and avulsions in the pubic branch region. Ultrasound diagnostics are very useful in differential diagnosis of acute trauma from overuse injuries. If inguinal hernia is suspected as a cause of pain, herniography should be performed. Magnetic resonance imaging (MRI) is a valuable method for evaluating discrete and ambiguous pelvic pain in athletes.^[3] Treatment of the groin strain is as complex as the cause of its development. In our opinion, treatment should always commence with nonoperative procedures; only when these fail should strictly selective surgery be contemplated, i.e., depending on the cause of the syndrome. The following are some common principles of nonoperative treatment of painful groin syndrome regardless of its cause: alleviate pain and control inflammation in the myotendinous apparatus; hasten healing of the myotendinous apparatus, and control further activities. We emphasize that it is of utmost importance to commence treatment as early as possible, i.e., when the first symptoms occur. Not paying attention to initial symptoms and continuation with athletic activity of the same intensity are the most common mistakes. Nonoperative treatment consists of local cryotherapy, anti-inflammatory nonsteroidal medications, stretching and strengthening exercises,



Fig. 2 Test for determining tenderness of the adductor muscles.

and some authors recommend local corticosteroid injections. Athletes with weak abdominal muscles should work on strengthening the muscles. Generally speaking, the purpose of rehabilitation is to preserve the strength, elasticity, and contractual abilities of the tissues along with the improvement of mechanical and structural characteristics of the tissues. At the stage when isometric training may be performed without pain, dynamic strength training may be introduced. Isokinetic training is actually the best, although it requires special equipment (Cybex). At the end of rehabilitation athletes may return to their specific training. The common error is that many tend to resume full athletic engagement too soon, which may necessitate repeating the treatment. Operative treatment of painful groin syndrome, to a certain extent, depends on the attitude in approaching this complex problem. Authors who consider tendinitis the most important feature of the syndrome recommend operative treatment in terms of desinsertion, i.e., cutting (tenotomy) of the adductors, gracilis, rectus abdominis, pyramidalis, and other muscles. Authors who believe that painful groin syndrome is primarily caused by disproportion in strength between the abdominal wall muscles and the upper leg muscles, especially emphasizing weak abdominal muscles (oblique and transverse abdominal muscles in particular) recommend surgical strengthening of the anterior abdominal wall by the old or new methods applied for inguinal hernia. We agree with the opinion: "It is fairer to enrich the poor than to impoverish the rich," meaning that it is better to strengthen or support the abdominal muscles than to weaken the adductor muscles by tenotomy. Nevertheless, on the basis of our experiences, we have realized that sometimes the muscular balance may be achieved only by a combination of "enriching the poor and impoverishing the rich." It should be pointed out that when discussing treatment of painful groin syndrome, whether nonoperative or operative, the approach should always be individualized. In this regard, it should be emphasized that procedures may and should be combined and there is no unique scheme for painful groin treatment. Prevention is of utmost importance, especially in young athletes.

Osteitis Pubis

The term osteitis pubis indicates that inflammatory changes are taking place, although in the majority of cases, no causative agent has been isolated. A number of different terms have been used in the past with the purpose of avoiding the connotation of inflammation (e.g., osteitis necroticans pubis, syndroma rectus-adductor, traumatic inguino-leg syndrome, anterior

pelvic joint syndrome, groin pain in athletes, osteoarthropathy of the pubic symphysis, traumatic pubic osteitis, etc.). Osteitis pubis was first reported in an athlete (fencer) in 1932 by Spinelli, and after that it has been reported in soccer players, football players, long distance runners, and long distance walkers.^[4,5] Osteitis pubis is seen particularly in those sports requiring sprinting and sudden changes of direction. The patients complain of pain in the lower abdomen and in the groin, but cannot link it to any known injury. At a later stage, the pain spreads in a fan-like manner into the adductor region. The principal characteristics of the pain are that its onset is linked to the athletic activity, then it diminishes and gradually disappears upon resting. The pain may also occur or become more intense when coughing, sneezing, or laughing. Palpations during clinical examination cause pain along the pubic bone and the symphysis itself. Radiography of the pelvis shows no changes in the initial phase of the disorder, but later on, the alterations are visible in the symphyseal region (fraying and sclerosis).^[6] Scintigraphic scan shows increased bilateral accumulation of radionuclides around the pubic bone even in the initial stage of the disease. Scintigraphy helps in differential diagnosis, because in the case of avulsion or stress fracture in the pubic branch region, the accumulation of radionuclides is unilateral. In the case of a negative scintigraphic scan and positive clinical signs, adductor longus and rectus muscle tendinitis is usually in question. The increased signal intensity seen on MRI is due to pubic bone marrow edema, such that an association exists between the clinical features of osteitis pubis and the MRI finding of pubic bone marrow edema. The treatment recommended is reduced athletic activity or complete rest if the pain is too severe. Some authors recommend symphyseal corticosteroid injections as a useful technique for treatment of osteitis pubis in athletes.^[7,8] Prospects for recovery in athletes with osteitis pubis are very good. When athletic activities are avoided, the disease subsides spontaneously. However, the duration of the recovery process may be variable. Mulhall et al. reported an outcome with symphyseal curettage in cases refractory to conservative management of osteitis pubis in professional soccer players.^[9] Williams, Thomas, and Downes treated operatively rugby players with osteitis pubis and vertical instability at the pubic symphysis.^[10] Operative treatment consisted of arthrodesis of the pubic symphysis by bone grafting supplemented by a compression plate.

ILIOPSOAS TENDINITIS AND BURSITIS (PSOAS SYNDROME)

Iliopsoas muscle tendinitis is caused by frequent repetitive flexions in the hip joint. It is the cause of

anterior hip pain and anterior snapping hip in weight lifters, skiers, oarsmen, football players, long and high jumpers, and hurdlers. Bursitis of the iliopsoas is more common in gymnasts, wrestlers, and uphill runners.^[11,12] The principal clinical sign is pain in the groin region on the anterior aspect of the hip joint, in the small pelvis, and sometimes in the back as well; this is accounted for by the fact that tendinitis and bursitis of the iliopsoas develop as a result of hypertonic muscle and prolonged strain as a consequence of disturbed biomechanical relations in the lumbar spine and in the hip joint. Initially, the pain is mild in intensity, but the repetitive strain makes the pain increasingly intense to the degree that athletic activity becomes impossible. Often the athlete cannot find the link between the pain and injury. Palpation of the tendinous attachment to the small trochanter causes pain. During flexion of the upper leg in the hip joint against resistance, pain occurs at the point of the muscle attachment (see Fig. 3). The motion should be attempted in a sitting position because then the iliopsoas muscle is the chief and only flexor of the upper leg in the hip joint. Between the iliopsoas tendon and the anterior aspect of the hip joint articular capsule, there is a nut-sized synovial sac (bursa); when irritated, it produces a clinical picture of a typical bursitis with pain during and after the athletic



Fig. 3 Test for iliopsoas muscle function.

activity. In differential diagnosis of the iliopsoas tendinitis and bursitis should be considered: tendinitis of the adductor longus muscle and rectus femoris muscle; rupture and avulsion of the iliopsoas muscle, rectus femoris muscle, and adductor longus muscle; and labrum lesion of the hip joint. In addition to clinical diagnosis, radiographs and ultrasound should be taken and MRI especially in patients with clinical suspicion of labral lesion of the hip joint.^[13] Tendinitis and bursitis of the iliopsoas muscle are treated nonoperatively. The conservative treatment depends of the stage of the disorder. During the initial stage training should not be completely discontinued but only reduced in intensity, but at the later stage, a rest period of 3–4 weeks is necessary. During this time, the athlete may engage in substitute activities (swimming, bicycling). The anti-inflammatory nonsteroidal drugs are administered orally and Renstrom and Peterson recommend an additional local infiltration of corticosteroids followed by a 2-week rest.^[14] The stretching and strengthening exercises may be introduced, although care should be taken that the patient feels no pain.

TENDINITIS OF THE RECTUS FEMORIS MUSCLE

Tendinitis of the rectus femoris muscle usually occurs as a result of prolonged, repetitive, and a sudden increase in strengthening exercise training or intensive goal-shooting training, e.g., in soccer. The athlete feels pain in the region above the hip joint at the point where the rectus femoris muscle attaches anteriorly and inferiorly to the iliac spine. Pain appears on palpation of rectus femoris muscle tendon attachment and can be also provoked by resisting upper leg flexion and lower leg extension. In differential diagnosis, the following should be taken into consideration: rupture or avulsion of the rectus femoris muscle or iliopsoas muscle, tendinitis of the iliopsoas muscle or adductor longus muscle, stress fracture of the femoral neck, changes in the hip joint such as arthrosis, osteochondritis dissecans, loose body in a joint, or bursitis.^[14,15] Calcific tendinitis at the site of origin of the rectus femoris muscle is a rare lesion.^[16] Pain in the groin region may develop as a consequence of the rupture at the attachment point or in the proximal third of the rectus femoris muscle. At the moment of injury, the pain is so severe that the athlete is compelled to discontinue activity, while in tendinitis of the rectus femoris muscle the athlete is unable to associate the onset of pain with an injury. Ultrasound will confirm the diagnosis and enable evaluation of the severity of injury (see Fig. 4). Radiographs should also be taken to rule out bone avulsion. Complete rupture, especially if in combination with avulsion fracture, is treated



Fig. 4 Sonogram showing rupture with hematoma (black arrow) of rectus femoris muscle during contraction.

operatively.^[17] Partial rupture is treated non-operatively: RICE (rest, immobilization, compression, elevation) and nonsteroidal anti-inflammatory medication are prescribed. Rehabilitation in terms of strengthening the muscles may begin when pain and swelling disappear. Treatment of rectus femoris muscle tendinitis is very similar to treatment of tendinitis of the iliopsoas muscle.

HIP EXTERNAL ROTATOR SYNDROME

The external hip rotator syndrome is rare, but a complicated problem, with regard to both diagnosis and treatment.^[18] It is most frequently found in hurdlers and ballet dancers where the symptoms consist of a vague pain in the buttocks and the greater trochanter area. The inflexibility of the hip external rotators, primarily the piriformis, gemellus superior, and gemellus inferior muscles, together with the internal rotation contracture of the hip, is thought to lead to chronic overuse of these muscles. Tendinitis of the quadratus femoris muscle is described as a separate entity.^[19] To locate the exact area of the pain, the upper leg is maximally and externally rotated during physical examination; pressure is applied to the greater trochanter area (fossa trochanterica), producing intense pain. Differential diagnosis must take into account the possibility of piriformis muscle syndrome and trochanteric bursitis, which can be confirmed by ultrasound. Piriformis muscle syndrome is based on the anatomical relationships of the muscle and the sciatic nerve; direct and indirect trauma, muscular hypertrophy or anatomical nerve variations, inflammation, and local ischemia probably combine to induce the piriformis muscle syndrome.^[20-22] Frequently,

diagnosis of piriformis syndrome requires eliminating the other causes of sciatic pain. Nonoperative treatment of short external hip rotator syndrome is based on stretching exercises for the afflicted muscle groups and neighboring muscles, with the aim of increasing their flexibility and concurrently the flexibility of the whole hip. Other nonoperative measures include rest, ice, and nonsteroidal anti-inflammatory drugs. Steroid injections can also be administered.

GLUTEUS MEDIUS SYNDROME

Gluteus medius syndrome is caused by overuse and inflammation of the gluteus medius muscle or of the sciatic nerve, which is immediately adjacent. The pain involves the buttock, lateral hip, groin, or the sacroiliac joint. The patient may also report dysesthesia and numbness in the buttock or posterior aspect of the thigh. Prolonged sitting may exacerbate pain. Activities with high hip abductor demand have been associated with gluteus medius syndrome, especially martial arts, cycling, aerobics, and dancing. During physical examination, hip abductor strength should be compared with the contralateral extremity. A full neurological exam is essential to rule out neurological causes. Palpation in the buttock region and over the gluteus medius muscle can discover trigger points or tenderness to palpation on or around the muscle.^[23] Other physical findings in gluteus medius syndrome may include limited stretch range of muscle, presence of a taut, palpable band in muscle, and referred pain produced by direct palpation of trigger points. Magnetic resonance is the imaging technique of choice in differentiating the cause of hip pain in athletes. Differential diagnosis of the gluteus medius syndrome includes sciatica, radiculopathy, greater trochanteric bursitis, iliotibial band syndrome, quadratus femoris tendinitis, sacroiliac joint dysfunction, and especially piriformis syndrome, but also pelvic compartment syndrome.^[19] Stretching exercises of the gluteus medius, gluteus maximus, and gluteus minimus and the piriformis muscle are recommended in conservative treatment of the gluteus medius syndrome. Physical therapy program includes cold packs, ultrasound, and other modalities. Injection with anesthetics and corticosteroids has been diagnostic as well as therapeutic. In rare cases, surgical debridement of a chronic tendinosis is an available option.

SNAPPING HIP SYNDROME

Snapping hip syndrome is a symptom complex characterized by an audible snapping sensation, which is usually, but not necessarily, associated with pain



Table 1 Different snapping hip syndromes

Localization	Etiology
Medial (internal)	Iliopsoas tendon over iliopectineal eminence
	Iliofemoral ligament over femoral head
	Iliopsoas tendon over the anterior-inferior iliac spine
	Iliopsoas tendon over the bony bridge on the lesser trochanter
Lateral (external)	Iliotibial band over the greater trochanter
Posterior	Tendinous origin of the long head of the biceps femoris muscle

during certain movements of the hip joint. A number of different etiologies, both intra-articular and extra-articular, have been described.^[24] Intra-articular causes for the snapping hip include loose bodies, osteocartilaginous exostosis, osteochondromatosis, and subluxation of the hip. The most frequent extra-articular cause of the snapping hip is the snapping of the iliotibial band over the greater trochanter. There are other extra-articular etiologies of the snapping hip. These include the snapping of the iliopsoas tendon over the iliopectineal eminence of the pelvis, the iliofemoral ligaments, the femoral head, and the tendinous of the origin of the long head of the biceps femoris muscle over the ischial tuberosity. Taking into account the localization of the cause for snapping hip, we recognize three different snapping hip syndromes: medial (internal), lateral (external), and posterior (see Table 1).^[1]

Medial snapping hip syndrome is frequently found in ballet-dancers.^[25] The snapping of the iliopsoas tendon occurs when the hip is extended from a flexed, abducted, and externally rotated position. In addition to an audible snapping sound, this movement is also, in some cases, accompanied by a painful sensation in the frontal hip area. Differential diagnosis includes consideration of psoas muscle syndrome, which consists of psoas bursitis, enthesitis, and tenosynovitis of the iliopsoas muscle. Nonoperative treatment of medial snapping hip syndrome consists of reduction or temporary cessation of painful activities and movements, nonsteroidal anti-inflammatory medication, and stretching exercises.^[26] Surgical treatment is very rare and consists of resection of the lesser trochanter bony prominence and/or partial or total release of the iliopsoas tendon. Lateral snapping hip syndrome is the most frequent and the best known. It is caused by the slipping of the iliotibial band over the posterior part of the greater trochanter and is frequent in runners, dancers, and basketball players.^[26] It can be asymptomatic or accompanied by pain, which is especially intense in those cases that develop trochanteric bursitis. In addition to the audible snapping phenomenon, the symptomatic form is characterized by pain in the region of the greater trochanter, which sometimes radiates to the buttock and/or lateral thigh. The

snapping can be best observed by placing the palm of one's hand on the greater trochanter area of the patient during walking. Another maneuver that will consistently reproduce the patient's symptoms is to have the patient stand on the healthy leg and imitate running motions on the affected side. The most common cause of the slipping of the iliotibial band over the greater trochanter is its excessive tightness (shortening), which can result from leg length discrepancies, pelvic tilt, or certain activities such as habitual running on the side of the road. Tightness of the iliotibial band is diagnosed using Ober's test (see Fig. 5).

Excessive tightness of the iliotibial band causes greater friction between the band, the greater trochanter, and the bursa that is normally located there; this in turn leads to chronic trochanteric bursitis. Characteristic symptoms of this bursitis include intense pain in the greater trochanter area and limping. Conventional sonographic studies can identify signs of tendinitis, bursitis, or synovitis. Dynamic sonographic studies reveal the cause of snapping hip in most patients.^[21] Nonoperative treatment of lateral (external) snapping hip syndrome consists of rest, avoidance of painful movements, iliotibial band stretching exercises, strengthening of the hip flexors, buttocks, hamstrings,

**Fig. 5** Ober's test for assessing the flexibility of the iliotibial band.

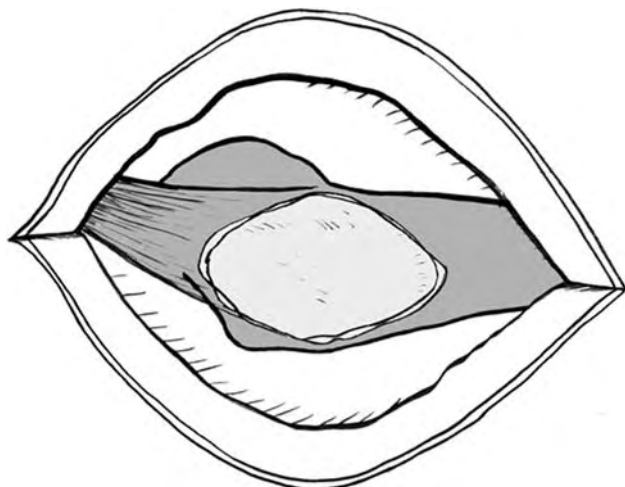


Fig. 6 Elipsoid excision of the iliotibial band overlying the greater trochanter.

hip rotation muscles, and quadriceps, nonsteroidal anti-inflammatory medication, and correction of the disturbed biomechanical relationships of the lower extremities (arch supports, adequate shoes, etc.). Injection of local anesthetics and soluble corticosteroids is also sometimes effective.^[27] In refractory cases surgical treatment is indicated. Numerous operative procedures have been described. We have successfully used the operative procedure described by Zoltan, Clancy, and Keene in our patients.^[28] We have termed this procedure fenestration of the iliotibial band (see Fig. 6).^[24]

Posterior snapping hip syndrome is caused by snapping of the tendinous origin of the long head of the biceps femoris muscle over the proximal part of the ischial tuberosity. Rask has described cases in which the subluxation of this tendon was accompanied by an audible snapping sound, a condition that he termed snapping bottom.^[29] Differential diagnosis must include consideration of developed bursitis in this area and enthesitis of the tendon, which we classify as part of the hamstring syndrome. Treatment is usually nonoperative.

HAMSTRING SYNDROME

Hamstring syndrome is a condition that affects muscles of the posterior thigh, primarily biceps femoris, semimembranosus and semitendinosus muscles. It is frequent in sprinters, hurdlers, and both long and high jumpers, but can also be found in other athletes, especially those who engage in rapid acceleration and short, intense sprinting, such as baseball players, football players, tennis players, water-skiers, dancers, and others.^[30-32] Hamstring syndrome is caused either by the overuse of the muscles at their insertion on the

ischial tuberosity or by the partial or total rupture of an individual muscle. The biceps femoris is the most commonly injured hamstring muscle and the semitendinosus is the second most commonly injured. Although hamstring injuries often involve one muscle injured proximally, multiple muscles were involved in 33% of athletes and the injuries were distal in 40% of athletes. All intramuscular injuries occurred at the musculotendinous junction, either in the ends of the muscle or within the muscle belly.^[33] Hamstring syndrome is a very common sports injury and numerous studies have been devoted to determining the causative factors and possible preventive measures of this typical self-induced injury. One of the reasons for this attention is the well-documented propensity of this syndrome to recur (as epitomized by the old saying “once a strain, always a strain”). Most authors today agree that the basic cause for the development of hamstring syndrome is muscle strength imbalance between the hamstring muscle group and the quadriceps femoris muscle. Hamstring syndrome will develop, in various forms, when there is relative weakness of the hamstring muscle group in relationship to the quadriceps femoris muscle. Studies have indicated that the hamstring muscles should have at least 60% of the strength of the quadriceps in athletes. Hamstring syndrome can also develop as a result of strength deficit between the right and left hamstring muscle groups. A 10% or greater muscle strength deficit between the two sides is also thought to be a predisposition for hamstring strain. In addition to the relative or absolute weakness of the hamstring muscles, another and fairly frequent cause of hamstring syndrome is decreased flexibility of these muscles. The flexibility of the hamstring muscles is determined by a number of individual characteristics, including sex, age, type and duration of athletic activity, and the specific type of training programs (during the season and off-season).

Flexibility of the hamstring muscles is commonly measured by Well’s “sit-and-reach” test (see Fig. 7). A more appropriate test for the flexibility of the hamstring muscle group has been described by Wallace (see Fig. 8). Aided by various types of testing equipment, of which the Cybex II is the most commonly used, the strength of the hamstring muscle and the hamstring-to-quadriceps muscle strength ratio are isokinetically determined. This test also allows evaluation of the muscle strength torque curve, which typically decreased in size.

Clinical features of hamstring syndrome that have resulted from overuse of the hamstring muscles at the point of their insertion at the ischial tuberosity include constant pain in the lower gluteal area, which radiates down the posterior thigh to the popliteal area. The pain typically increases during the performance of forcible or sudden movements, which stretch the





Fig. 7 Sit and reach test.

hamstring muscles. Another characteristic complaint is pain felt while sitting, e.g., while driving a car or sitting during lectures. The pain is often relentless, causing the patient to change position or stand up for relief. During physical examination, the pain can be reproduced by applying pressure to the ischial tuberosity area. Various clinical tests that cause the stretching of the hamstring muscles also induce pain at the site of the ischial tuberosity (Well's and Wallace's tests). Pećina and Bojanić use the more exact "shoe wiping" test, which consists of having the patient imitate movements typically used when wiping shoes on a doormat.^[1] The hip and knee are in a position of maximal extension while the hamstring muscles are in retroflexion in the hip. While performing these movements, the hamstring muscles are stretched and, at the same time, in motion, which gives an additional value to this test; this enables the patient to score positive even when scoring negative on other tests. A positive score is indicated by pain in the ischial tuberosity area or along one of the hamstring muscles. Any diagnosis of hamstring syndrome should include a neurological examination because differential diagnosis includes ruling out the possibility of spinal sciatica. Radiological, electrodiagnostic, CT, and MRI examinations can be used to achieve this. Piriformis syndrome presents a special problem in differential diagnosis.^[34] Magnetic resonance imaging is very helpful in detecting the site, extent, and characteristics of the injury, and can also predict the time period an athlete will be disabled and help decide the best treatment planning.^[33] Treatment of the hamstring syndrome can best be summarized with the old saying "the best treatment

is prevention." Treatment of hamstring syndrome that has developed as a result of overuse injury is usually nonoperative. It consists of a rehabilitation program of stretching exercises and strengthening of the hamstring muscles. Physical procedures commonly applied to



Fig. 8 Wallace's sign.

all chronic overuse syndromes are also applicable. Surgical treatment as recommended by Puranen and Orava consists of cutting and dividing the tendinous tissue near the site of origin of the hamstring muscles without loosening the muscles from the ischial tuberosity.^[32] After division, the tendon ends should be completely separated from each other. The sciatic nerve must also be freed from any tendinous parts of the hamstring muscle. According to Kujala, Orava, and Jarvinen after first aid with rest, compression, cold, and elevation, the treatment of hamstring muscle injury must be tailored to the grade of injury (slight, moderate, severe).^[35] Urgent surgical treatment is recommended in cases with total or nearly total soft tissue hamstring muscle insertion rupture.^[36]

CONCLUSIONS

In sports traumatology the hip joint pathology is not as abundant as it is in the rest of the orthopedic surgery. Nevertheless, the presence of numerous structures such as muscles, tendons, or synovial bursas in the pelvic region and hips, along with great dynamic and static forces exerted during physical activity, makes it prone to numerous athletic injuries. In contrast to the knee joint, which is known for acute injuries, the hip joint is particularly prone to overuse injuries and syndromes. Painful syndromes belonging to the overuse group are generally considered the most frequent in some athletic activities, e.g., soccer; however, neither the site nor the pain are precisely defined. Because of this these injuries are easily overlooked or not treated correctly. Therefore, it is of utmost importance to be familiar with the principles of epidemiology, etiology, pathophysiology, and diagnostics principles to treat the patients successfully and without recidives, which are not uncommon. The main treatment for hip overuse injuries is rest from the offending athletic activity, a concept known as “relative rest.” In the rare occasions when conservative treatment fails, surgical treatment is advised.

ARTICLES OF FURTHER INTEREST

Bone Fracture Fixation; Knee Joint: Overuse Injuries; Stress Fractures

REFERENCES

1. Pećina, M.M.; Bojanić, I. *Overuse Injuries of the Musculoskeletal System*, 2nd Ed.; CRC Press: Boca Raton, 2003.
2. Renstrom, P.A. Tendon and muscle injuries in the groin area. *Clin. Sports Med.* **1992**, *11*, 815–831.
3. Barile, A. Groin pain in athletes: role of magnetic resonance. *Radiol. Med.* **2000**, *100*, 216–222.
4. Spinelli, A. Una nuova malattia sportiva: la pubalgia degli schermitori. *Ortop. Traumatol. Appar. Mot.* **1932**, *4*, 111–127.
5. Bass, C.J.; Connell, D.A. Sonographic findings of tensor fascia lata tendinopathy: another cause of anterior groin pain. *Skeletal Radiol.* **2002**, *31*, 143–148.
6. Pearson, R.L. Osteitis pubis in a basketball player. *Phys. Sportsmed.* **1988**, *16*, 69–72.
7. Holt, M.A.; Keene, J.S.; Graf, B.K.; Helwig, D.C. Treatment of osteitis pubis in athletes. Results of corticosteroid injections. *Am. J. Sports Med.* **1995**, *23*, 601–606.
8. O'Donoghue, D.H. *Treatment of Injuries to Athletes*, 4th Ed.; W. B. Saunders: Philadelphia, 1984.
9. Mulhall, K.J.; McKenna, J.; Walsh, A.; McCormack, D. Osteitis pubis in professional soccer players: a report of outcome with symphyseal curettage in cases refractory to conservative management. *Clin. J. Sport Med.* **2002**, *12*, 179–181.
10. Williams, P.R.; Thomas, D.P.; Downes, E.M. Osteitis pubis and instability of the pubic symphysis. When nonoperative measures fail. *Am. J. Sports Med.* **2000**, *28*, 350–355.
11. Biundo, J.J., Jr.; Irwin, R.W.; Umpierre, E. Sports and other soft tissue injuries, tendinitis, bursitis, and occupation-related syndromes. *Curr. Opin. Rheumatol.* **2001**, *13*, 146–149.
12. Johnston, C.A.; Wiley, J.P.; Lindsay, D.M.; Wiseman, D.A. Iliopsoas bursitis and tendinitis. A review. *Sports Med.* **1998**, *25*, 271–283.
13. Ekberg, O.; Sjoberg, S.; Westlin, N. Sports-related groin pain: evaluation with MR imaging. *Eur. Radiol.* **1996**, *6*, 52–55.
14. Renstrom, P.; Peterson, L. Groin injuries in athletes. *Br. J. Sports Med.* **1980**, *14*, 30–36.
15. Peterson, L.; Renstrom, P. *Sports Injuries: Their Prevention and Treatment*; Martin Dunitz: London, 1988.
16. Pope, T.L., Jr.; Keats, T.E. Case report 733. Calcific tendinitis of the origin of the medial and lateral heads of the rectus femoris muscle and the anterior iliac spin (AIIS). *Skeletal Radiol.* **1992**, *21*, 271–272.
17. O'Connell, M.J.; Powell, T.; McCaffrey, N.M.; O'Connell, D.; Eustace, S.J. Symphyseal cleft injection in the diagnosis and treatment of osteitis pubis in athletes. *Am. J. Roentgenol.* **2002**, *179*, 955–959.
18. Hunter, S.C.; Poole, R.M. The chronically inflamed tendon. *Clin. Sports Med.* **1987**, *6*, 371–387.
19. Klinkert, P., Jr.; Porte, R.J.; de Rooij, T.P.; de Vries, A.C. Quadratus femoris tendinitis as a cause of groin pain. *Br. J. Sports Med.* **1997**, *31*, 348–349.
20. Srinivasan, A.; Schuricht, A. Long-term follow-up of laparoscopic preperitoneal hernia repair in professional athletes. *J. Laparoendosc. Adv. Surg. Tech. A.* **2002**, *12*, 101–106.
21. Choi, Y.S.; Lee, S.M.; Song, B.Y.; Paik, S.H.; Yoon, Y.K. Dynamic sonography of external



- snapping hip syndrome. *J. Ultrasound Med.* **2002**, *21*, 753–758.
22. Dederich, R. Die Schnappende Huftte erweiterung des Tractus Iliotibialis durch Z-plastik. *Z. Orthop.* **1983**, *121*, 168–170.
 23. Njoo, K.H.; Van der Does, E. The occurrence and inre-rater reliability of myofascial trigger points in the quadratus lumborum and gluteus medius: a prospective study in non-specific low back pain patients and controls in general practice. *Pain* **1994**, *58*, 317–323.
 24. Pećina, M.; Bojanić, I.; Hašpl, M. The snapping hip. *Hip Int.* **1994**, *4*, 133–136.
 25. Reid, D.C. Prevention of hip and knee injuries in ballet dancers. *Sports Med.* **1988**, *6*, 295–307.
 26. Schaberg, J.E.; Harper, M.C.; Allen, W.C. The snapping hip syndrome. *Am. J. Sports Med.* **1984**, *12*, 361–365.
 27. Melamed, H.; Hutchinson, M.R. Soft tissue problems of the hip in athletes. *Sports Med. Arthrosc. Rev.* **2002**, *10*, 168–175.
 28. Zoltan, D.J.; Clancy, W.G., Jr.; Keene, J.S. A new operative approach to snapping hip and refractory trochanteric bursitis in athletes. *Am. J. Sports Med.* **1986**, *14*, 201–204.
 29. Rask, M.R. Snapping bottom: subluxation of the tendon of the long head of the biceps femoris muscle. *Muscle Nerve* **1980**, *3*, 250–251.
 30. Askling, C.; Tengvar, M.; Saartok, T.; Thorstensson, A. Sports related hamstring strains—two cases with different etiologies and injuries sites. *Scand. J. Med. Sci. Sports* **2000**, *10*, 304–307.
 31. Orchard, J.W. Intrinsic and extrinsic risk factors for muscle strain in Australian football. *Am. J. Sports Med.* **2001**, *29*, 300–303.
 32. Puranen, J.; Orava, S. The hamstring syndrome. *Am. J. Sports Med.* **1988**, *16*, 517–521.
 33. De Smet, A.A.; Best, T.M. MR imaging of the distribution and location of acute hamstring injuries in athletes. *Am. J. Roentgenol.* **2000**, *174*, 393–399.
 34. Pećina, M.; Krmpotić-Nemanić, J.; Markiewitz, A.D. *Tunnel Syndromes. Peripheral Nerve Compression Syndromes*, 3rd Ed.; CRC Press: Boca Raton, 2001.
 35. Kujala, U.M.; Orava, S.; Jarvinen, M. Hamstring injuries. Current trends in treatment and prevention. *Sports Med.* **1997**, *23*, 397–404.
 36. Orava, S.; Kujala, U.M. Rupture of the ischial origin of the hamstring muscles. *Am. J. Sports Med.* **1995**, *23*, 702–705.

Histogenesis

Julie Glowacki

Department of Orthopedic Surgery, Brigham and Women's Hospital, Harvard Medical School, and Harvard School of Dental Medicine, Boston, Massachusetts, U.S.A.

INTRODUCTION

Histogenesis, or the formation of tissue, is accomplished by cells in various normal and pathological settings. Attempts to use new ideas of regenerative medicine and bioengineering to replace missing or lost parts rely on information gathered from different sources. Research in developmental biology shows that many species use the same genes to regulate the complex steps involved in tissue formation. Cell–cell interactions and cell–matrix interactions play critical roles in tissue formation both in embryogenesis and in wound healing. One of the best approaches to designing tissue replacements is to understand the processes of histogenesis during embryogenesis. Progress has been made by identifying factors that account for derangements of embryogenesis in mutant and recombinant mice and in clinical malformations. In-vitro models of histogenesis have used adult somatic cells and putative stem cells to identify the key elements in growing and maintaining tissues.

Tissue engineering offers the potential for consistent and practical management of patients in need of replacement tissues. Tissue engineering includes both cell-free and cell-based approaches. With cell-free therapies, agents or materials are introduced to stimulate the body to recapitulate the process of controlled tissue growth and maintenance. With cell-based therapies, manipulation of autogenous cells with signals provided by a scaffold or accompanying factors may require an intermediate in-vitro phase or proceed directly to reconstruct the missing volume in vivo. One challenge is to apply the growing body of information about normal and pathological processes for controlled therapeutic applications. The goal is to apply first principles from multidisciplinary approaches into rational design solutions. The translation of basic science knowledge to regenerative medicine is in its infancy and is likely to encounter many shifts as innovative ideas for modifiers of histogenesis are discovered and integrated into therapeutic possibilities.

TERMINOLOGY

The term histogenesis is derived from the Greek roots *histon*, meaning web or tissue, and *genesis*, meaning

generation or production. Its literal definition is given as “the formation and differentiation of tissues.” It may be more useful to separate the two processes into histodifferentiation and histogenesis. Thus, histodifferentiation refers to the process in which apparently indifferent, or unspecialized, groups of cells attain the distinctive features of adult form and function. Early stages of embryonic development are often characterized by the morphological appearance of recognizable tissue types, i.e., histomorphogenesis. Once the first signs of a specialized tissue type appear, further increase in the mass and form of the tissue is possible by its hypertrophy, or growth. Thus, histogenesis would indicate the accrual of more tissue mass of the specific differentiated type. The distinction between these two terms becomes especially useful in discussing the mechanisms and regulation of tissue repair processes. Histonomy, incorporating the Greek root *nomos* for law, consolidates the two processes into the laws that govern development and structure of tissues.

Histology deals with the microscopic structure of plant and animal parts and morphological evidence of their functions. It is usually restricted to light microscopic methods. Various techniques are used to gain information about tissue functions, including histochemical, histoenzymatic, and immunohistologic localization of specific molecular components. Histophysiology is the science that integrates the microscopic features of tissues in relation to tissue function. Histopathology is the science of disease states as they evidence themselves in tissues. New methods, such as in situ hybridization, offer powerful means to identify cellular markers. Gene transcripts that are not plentiful can be amplified and detected by RT-PCT/in situ hybridization. Alternatively, laser capture microdissection, or manual collection of specific cells with micromanipulators, can be used for identification of DNA or messenger RNA molecules within tissue sections or to compare cells or regions in tissues.

HIERARCHY

A histioblast is the generic term for a cell capable of forming a tissue. Classical usage reserves the term histiocyte for fixed or wandering macrophages within tissues and the term histoblast for certain epithelial



cells in insect larvae. Specialized histioblasts produce the characteristic tissue, for example, fibroblasts, osteoblasts, angioblasts, hepatoblasts, and neuroblasts. The suffix *-cyte* is often restricted to the mature cell that appears to be terminally differentiated and involved in the tissue function, but without conspicuous activity in synthesizing more tissue. One notable exception to this distinction between *-blast* and *-cyte* is the term *fibroblast*, which is used for both the cells producing fibrous tissue as well as the stromal cell residing in mature connective tissue; *fibrocyte* is rarely used. Another notable conventional usage is the term *chondrocyte*, which is used for both the cell producing cartilage tissue and the cell residing in cartilage; *chondroblast* is rarely encountered. There is no logic for these exceptions.

A tissue is an aggregate of cells that are specialized to perform a certain function. Tissue types can be organized on the basis of their primary functions (Table 1). There is some overlap in functions, but an extreme case is found with bone tissue.^[1] As a tissue, bone furnishes important mechanical properties to the skeleton. The mechanical roles of the skeleton are to provide support of the organism, locomotion, and protection of vital organs, such as the brain and heart. These mechanical demands require stiffness to resist large forces. The mechanical character is exquisitely achieved by the biphasic nature of the extracellular matrix, containing an organic component, largely collagen, and inorganic calcium phosphate mineral in the form of poorly crystalline carbonatohydroxyapatite. The metabolic role of bone is its contribution to calcium homeostasis. The extracellular mineral can be mobilized when calcium needed for vital bodily functions such as clotting of body fluids, neuromuscular irritability, growth and regeneration, maintenance of mucous coverings, intercellular contacts, and ameboid and ciliary motion. Bone's role in calcium metabolism is accomplished by the integrated activities of its three cell types: the osteoblast, osteocyte, and osteoclast, each directly or indirectly regulated by the three major calcitropic hormones: parathyroid hormone, 1,25-dihydroxyvitamin D, and calcitonin. Of note, the mechanical function can be sacrificed, to the devastating extent of risk for fracture, in situations of

Table 1 Types of tissue functions

Function	Examples
Structural	Connective tissue, muscle, bone, cartilage, epidermis, dermis
Metabolic	Pulmonary, hepatic, renal, pancreatic, gastrointestinal, blood, marrow, bone
Circulatory	Myocardium, vascular
Sensory	Neural, ocular

calcium deprivation, such as rickets, osteomalacia, osteoporosis, and aging. Processes are designed to ensure that calcium homeostasis is maintained as stored calcium is removed, without a failsafe limit, by resorption of bone matrix and reduction of new bone formation.

A simple tissue may be made of a single cell type, for example, cartilage. Cartilage comprises cells called chondrocytes and an extracellular matrix of fibrillar collagen type II, highly charged molecules of aggrecan, and water. It is avascular and has no nerves. Cartilage tissues function as a lubricating surface in joints and as a temporary (and readily resorbed) scaffold for bone formation.

Epithelial tissue may be simple or complex, but is usually bounded by basement membrane. Most tissues, however, consist of several cell types. Connective tissues, for example, include blood vessels, nerves, and specialized tissue cells, such as osteocytes, myocytes, or tenocytes. Connective tissues often have a large amount of extracellular matrix whereas other tissue types often have a high cell-to-matrix ratio. Thus, tissues constitute the level of organization between that of cells and organs.

An organ is a structure comprised of two or more tissues that are organized to carry out a specific function, e.g., kidneys, lungs. Development of organs involves the integrated activities of different cell and tissue types. Patterning is far more complex in organogenesis than in histogenesis. Skin can be considered either an organ or a tissue depending on the context. Likewise, bone can be discussed as a tissue or as a specific organ, e.g., femur, mandible. Authors generally attempt to avoid confusion in readers by specifying which level is intended.

TISSUE COMPLEXITY AND VARIABILITY

In unicellular organisms, specialized functions occur in different compartments. Mitochondria, for example, efficiently organize energy metabolism. Development of multicellular organisms directs the expression of specific phenotypes for different cells and patterning for histogenesis and organogenesis. The generation of cellular diversity is called differentiation.

It is insufficient to characterize a tissue type on the basis of generalized constituents and organization. A tissue type can display an array of arrangements in order to realize distinct properties. For example, skeletal muscle tissue anatomy is variable and highly dependent upon the energy requirements of the specific musculotendinous organ. The basic motor unit includes a peripheral nerve, the myoneuronal junction, and the muscle fibers innervated by that nerve cell. Each muscle fiber is a multinucleated syncytium

formed by the fusion of myoblasts during development. Histologically, groups of muscle fibers are enclosed by a fibrous epimysium and packed with endomysium through which neurovascular bundles course. Different types of sensory fibers serve as receptors of the stretch reflex, proprioception, vibration, and pain. Red and white skeletal muscle fiber types demonstrate an extreme range of function, the former with relatively slow contraction, but slower to fatigue. Examples of the red type are the intercostal muscles, masticatory musculature, and muscle of the eyeball. The red are thinner than the white fibers and more richly vascularized. White fibers compose the muscles of the extremities. Thus, it is evident that discussions of muscle tissue and its design requirements need to specify the precise function needed.

Another example of the wide variety of tissue arrangement is presented by bone tissue. Bone displays a variety of gross and histological patterns. At the macroscopic level, it can be compact, as in the cortical tubes in the midshafts of long bones, or more porous, as in the cancellous, often marrow-containing, inner portion of vertebrae, flat bones, and ends of long bones. Microscopically, human cortical bone shows a complicated pattern of osteons, or units, with a central vascular canal and radiating capillaries within concentric rings of matrix and network of intercommunicating osteocytes. This contrasts with the thin trabecular arrangement of cancellous bone. These arrangements result in different mechanical properties^[2] and metabolic^[3] consequences. Strength and stiffness differ in cortical bone depending on the direction of measurement. This is termed mechanical anisotropy. Whereas the bone is required to provide strength, it needs to be remodeled and renewed for rejuvenation of both its mineral and organic components.^[4] This is achieved by focal tunneling and refilling, or remodeling, within the tissue. Remodeling is visible histologically as cutting cones with osteoclasts at the advancing tip and a layer of osteoblasts and a capillary in the process of laying down concentric lamellae in new osteons. In contrast, the thin trabeculae (approximately 200 micrometers thick) in cancellous bone show morphological evidence of osteoclastic resorption and refilling of lacunae on their surfaces. The trabecular tissue is lighter due to its porosity and provides a greater surface for resorption. These histological features are a static image of the dynamic nature of bone, as it adapts to an acute or chronic mechanical or physiological stimulus.^[5,6]

EMBRYONIC HISTOGENESIS

Embryonic tissues are organized either as epithelia, in which cells are held together in layers, or as

mesenchyme, in which cells are enmeshed in an extracellular matrix. That pattern is also followed in compound structures. The ectodermal lineage gives rise to the outer epithelium of the body, to brain and neural structures, and parts of the tooth, adrenal gland, and skull and branchial cartilages. The endodermal lineage gives rise to gastrointestinal, urinary, and pharyngeal tissues and derivatives of the pharyngeal pouches, such as the thymus and parathyroids. The mesoderm give rise to musculoskeletal and connective tissue structures and components of viscera, hemangioblastic tissues, cardiac tissues, and adrenal cortex. The germ line cells that give rise to gametes are considered a unique lineage from primordial germ cells. In the avian and mammalian embryo, basal laminae first appear beneath the ectoderm and endoderm during the formation of the primitive streak. Mesenchymal cells migrate from the streak between ectoderm and endoderm, filling the space between and forming the third germ layer, the mesoderm.

Cells and tissues proceed along pathways of differentiation once they are determined, or committed to the lineage. That status can be demonstrated by transfer to another region of the embryo. A cell or tissue is said to be specified when, even if it does not appear to express its characteristic features, its lineage is imprinted and can be revealed *in vitro*. Current interest in the ability of adult cells to develop into different lineages widens the prospect of reparative medicine with so-called stem cell therapy.

Elegant studies in chick embryos by Saunders illustrate the process of formation of complex tissues and the phenomenon of inductive specificity.^[7] A striking lesson is learned from transfer and recombination experiments. Different parts of the chick's skin have different cutaneous structures: wide feathers, narrow feathers, scales, and claws. Saunders' experiments showed that the mesoderm instructs the ectoderm which structures to produce. When immature mesoderm from the foot is combined with ectoderm from the wing, the skin develops scales and claws. Thus, the mesenchyme regulates different genes in the responding cells of the ectoderm. Inductive interactions can occur in three ways: by diffusion from one cell toward another, by contact between cells, or by mediators in extracellular matrix. One approach to tissue engineering is to mimic the information provided by the embryonic mesenchyme by delivery of the signal either in soluble form or in a scaffold.

The dominant paradigm of morphogenesis involves differential cell affinities, either for the surfaces of other cells or for the molecules of the extracellular matrix. Cell adhesion molecules (CAMs) comprise two major classes: calcium-dependent glycoproteins called cadherins and calcium-independent immunoglobulin CAMs. Cell receptors for extracellular matrix



molecules constitute a family of integrins that integrate the extracellular and intracellular scaffolds.

Extracellular matrix molecules play important roles in organizing histogenesis. The basal laminae are made of collagen type IV and provide a structure to which the epithelial and mesenchymal cells can adhere. Collagen type II, a fibrillar collagen, first appears around the notochord and not only provides support, but also contributes to inductive effects of the notochord on growth and differentiation. Thin fibrils of collagen type I provide both structural and informational roles. Other extracellular proteins present in the early embryo are fibronectin, laminin, and tenascin. More information is available on the importance of extracellular hyaluronan and other proteoglycans on morphogenesis and differentiation. Both in-vivo and in-vitro studies provide evidence for cell–matrix interactions through specific binding receptors, such as CD44 for hyaluronan. Hyaluronan has excellent biochemical properties for tissue expansion. Accumulation of hyaluronan increases swelling pressure because of its hydrophilic nature and, conversely, its degradation leads to reduced swelling pressure, allowing recondensation. The embryonic cornea is an excellent example of this process. Heparin and other related compounds have been shown to inhibit cell proliferation and heparin-binding growth factors can modulate proliferation as matrix is remodeled. Branching morphogenesis is common in the development of many structures (mammary gland, lung, salivary gland, and renal tissue) and classical studies showed that digestion with hyaluronidase blocks branching. Toole's studies on limb development indicate the importance of changes in hyaluronan and of its organizing the pericellular matrix during mesoderm condensation and chondrogenesis.^[8]

Although much attention is paid to molecular regulators of pattern formation, signals involved in embryonic histogenesis are often mechanical in nature. The effect of motion on joint cavitation and of shear on blood vessels are examples of mechanical regulation in vivo. In addition, alteration of cell–matrix interactions have been shown to involve tension-dependent changes in cell shape and cytoskeletal mediators of cell function.^[9]

Three-dimensional, if not four-dimensional, development has been studied in models of dorsal/ventral and right/left axes in various species. There has been remarkable evolutionary conservation of gene organization and transcriptional expression. Examination of distribution of gene expression in very early embryos shows that neural crest cells, for example, are already committed to form structures specific to their location well before they begin to migrate. Gene knockout experiments offer the means to identify the patterning genes that regulate these events. Deletion of one of the

Hox genes, for example, by gene targeting gives rise to a mouse with lumbar vertebra that look like thoracic vertebrae with a rib.^[10] Experiments of nature, i.e., malformations, and studies with teratological agents have given us additional insights into patterns of development as well as coordination of development of different organs that are linked in congenital abnormalities. Rapid progress has been made in identifying different genes that control tissue and organ development, and mutations in many have been identified in congenital malformations. A remarkable illustration is a missense mutation in *MSX1* in all affected members of a family with absent second premolars and third molars.^[11]

As important as are embryonic cell differentiation and activity for histogenesis, programmed cell death, or apoptosis (from the Greek *apo*, off, and *ptosis*, falling or dropping), occurs as tissues remodel into functional units. Apoptosis is a way to change temporary structures or parts of structures, but has yet to be exploited in tissue engineering.

POSTNATAL HISTOGENESIS

Histodifferentiation is generally regarded as the process in embryonic development by which a tissue develops into its mature form. There are many examples, however, in which dramatic changes occur in tissues and organs after embryogenesis. In amphibian metamorphosis, explosive changes occur not only in visible signs (loss of gills and tail, gain of limbs and mucus glands), but in liver enzymes and hemoglobin. Also under hormonal direction, mammalian puberty is heralded by development of secondary sexual structures and by spurts of growth. There are dramatic changes in tissues of the womb during pregnancy and following parturition. Some cancers and diseases such as psoriasis are also characterized by continuous growth and histogenesis.

Postnatal histogenesis is usually thought of as that occurring in normal adult wound repair and regeneration. Repair is the process of response to an injury, usually as a scar, while regeneration is the restoration of the injured parts with characteristics indistinguishable from the original material. Postnatal skin is not capable of regeneration, but repairs quickly by scar formation. Scar is a poorly organized tissue that can be regarded as the quick way to seal the skin and avoid risk of invasion and infection. Experimental studies indicate that wound healing in fetal animals is fundamentally distinct from postnatal responses.^[12] Cutaneous wounds made in the midgestational period heal with true regeneration of epithelium and dermis, exactly like the surrounding tissues. Current thinking attributes the differences between fetal regeneration

and postnatal repair to environmental and cellular factors. Evidence implicates Transforming Growth Factor- β (TGF- β) as mediating fibrosis in the adult pattern.^[13] Full understanding of fetal regenerative mechanisms is likely to result in novel approaches to clinical management of wounds without scarring and adhesions.

Soft tissue wound healing begins with an acute inflammatory response to the injury. In general, the degree of the acute response is proportional to the extent of the injury, with complex reactions noted in cases of polytrauma. The cascade of responses includes immediate vasoconstriction, followed by vasodilation, leukocytic diapedesis, gathering of polymorphonuclear neutrophils, and predominance of mononuclear leukocytes. The clinical signs associated with this process are redness, swelling, warmth, and tenderness. Cell-derived factors and proteolytic enzymes mediate and amplify the cellular reactions, culminating with the beginning of repair, usually visible within three days after injury. Epithelial cells at the wound margin begin to proliferate and migrate. As necrotic tissue and clot (and bacteria, if present) are removed by granulocytes and macrophages, fibroblastic cells invade along fibrin fibers, accompanied by new capillary sprouts. Granulation tissue is characterized as rich in delicate small vessels. The name derives from the granular appearance of this tissue when it is incised. In the fibroblastic phase of matrix formation, capillaries regress and collagen synthesis diminishes with loss of fibroblasts. Changing patterns of cytokines control the initiation and mitigation of each phase. The matrix is modified by myofibroblasts and wound contraction.^[14] Intense lytic activity results in tissue compaction and continuous accumulation of fibrillar collagen gives rise to a gradual increase in tensile strength to nearly 70% as in intact skin. If scar formation is viewed as a rapid response of skin intended to seal openings against invading organisms, its efficiency for doing so clearly outweighs its lack of fidelity to the original structure.

Angiogenesis or neovascularization is a critical element in wound healing. Poorly vascularized tissues, like cartilage, cornea, and the white zone of the meniscus, are usually not able to repair large defects. In the case of penetrating wounds of the cornea, vascularization can occur, but is associated with formation of scar with disordered collagen fibers of random size that are opaque and cause varying degrees of disability. Clinicians have noted especially robust bone regeneration in sites with a modest degree of infection. This has been attributed to local inflammatory reaction to infection. Whether novel osteogenic regulators or combinations mediate this occurrence and whether the phenomenon can be harnessed is unknown.

Research on the factors that mediate the phases of wound healing of soft, hard, or visceral tissues has led

to trials to promote responses to injury, especially in compromised settings. Some of the tested agents are platelet-derived growth factor (PDGF), TGF- β , epidermal growth factors (EGFs), fibroblast growth factors (FGFs), insulin-like growth factors (IGFs), and morphogenic factors, such as the bone morphogenic proteins (BMPs). Although each may show promise in experimental model systems and small, controlled trials, therapeutic potential and optimal delivery remain to be realized. There may be a theoretical shortcoming to this strategy. Evidence suggests that addition of a single factor may enhance a phase of healing, but may disrupt the required downregulation of its action required for subsequent coordinated events. Negative feedback normally serves to progress from one stage of healing to the next and may be disorganized by application of a single agent. Several factors appear to act together and in tandem during progress from one phase to another. Another simple strategy that appears to be gaining popularity is the use of platelet-rich plasma (PRP), concentrated from autologous blood with cell separators at the time of surgery.^[15] Simply put, PRP is intended to amplify and accelerate the effects of growth factors (known and unidentified) found in platelets and normally involved as initiators of all wound healing. There is little risk of toxicity or immunoreactions with these preparations, and improvements in the technology advance quickly toward achieving greater concentrations with greater speed.

Tissue specificity is achieved in postnatal histogenesis by the available chemical mediators and cells. Sources of new cells capable of neohistogenesis are required, as are sources of neovascularization. These points are highlighted by the failure of repair and regeneration in postirradiation tissue. Irradiated tissue may be among the great challenges for tissue engineering. Periodontal tissue regeneration is an example of corrections needed in several tissue types, i.e., the periodontal ligament, cementum, and alveolar bone. In-vitro studies show that a combination of two growth factors, PDGF and IGF-I, is chemotactic for both periodontal ligament cells and osteoblasts and is a stimulus for both cell types.^[16]

ENGINEERED HISTOGENESIS

Defining the constituents of a tissue type and its three-dimensional organization is a starting point for understanding the elements needed for specific biological functions. It is incomplete and flawed, however, when reflecting on how to engineer dynamic tissue substitutes. One intellectual paradigm is to examine the ways that body parts operate in different species, to identify their character and limitations, and to design



substitutes. Synthetic materials have achieved some clinical success. For example, dental implants replace lost dental roots and can support prosthetic teeth. Long-term success of dental implants is due in part to their immediate mechanical stability and in part to osteointegration as the jaw bone remodels around them and incorporates them, indifferent to the fact that they are, in fact, foreign bodies. This is possible because of the development of biocompatible materials and attentive design used to manufacture them. Use of large metallic prostheses, for example, the long components inserted well into the proximal femur, seem like an extraordinary measure for replacing the function of the thin shell of articular cartilage that has eroded due to osteoarthritis. Current approaches attempt to activate the body's inherent regenerative capacity.

Different models exist for promoting histogenesis clinically (Fig. 1). For some applications, implantation of biocompatible scaffolds can act as a support for directing and containing tissue growth. There are many natural and synthetic resorbable materials with good biocompatibility and tissue compatibility that can be modified to have the porosity and mechanical properties needed for specific applications. Guided tissue regeneration (GTR) is an example in which a periodontal membrane barrier, either resorbable or permanent and requiring removal, maintains a space for

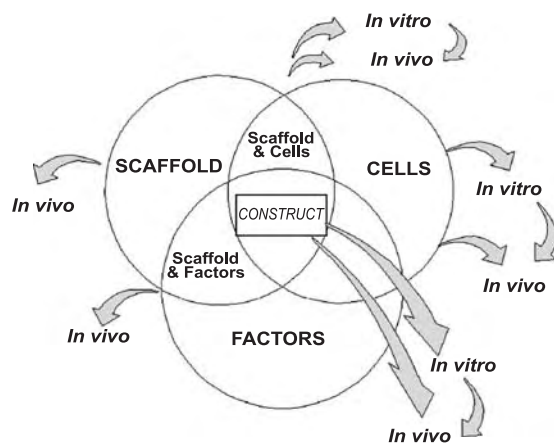


Fig. 1 Schematic for the design of engineered tissues. SCAFFOLD, CELLS, and FACTORS may be used separately or together in a construct to enhance neohistogenesis. The SCAFFOLD may serve three functions: 1) as a cell-free implant, it may support the ingrowth of needed tissue; 2) it may be a delivery system for factors, genes, or drugs; and 3) it may be a carrier for cells directly or following in-vitro growth and maturation. Autogenous CELLS may be harvested for direct transplantation or for intermediate in-vitro culture. FACTORS may be used for in-vitro or in-vivo delivery to direct the expansion of the needed tissue or to promote in-vivo integration.

osteogenesis while preventing epithelial invasion. Addition of bioactive signals, either as genes, soluble factors, or tethered agents, is an attractive way to guide a tissue response. An example is provided by a proof-of-principle study in which an RGD-containing peptide (i.e., with the tripeptide sequence arginine–glycine–aspartic acid that serves as the extracellular ligand for cell attachment) was used to coat a porous polylactic–glycolic acid scaffold and promoted the early stages of osseous ingrowth in a rat tibial wound.^[17] Scaffolds for cell-free use are designed to stimulate neohistogenesis, often by mimicking signals for anabolic processes (also called biomimetic implants).

For bone reconstruction, an alternate to synthetic replacements is a biological approach. Bone tissue is one of the few tissues that is capable of true regeneration. In situations where there is mobility, poor vascularity, or loss of tissue viability secondary to steroid treatment, for example, nonunion can result. There is extensive experience with various methods to promote osteogenesis. Research on the principles underlying the mechanisms by which bone regeneration is enhanced has led to advanced therapeutic approaches.^[18] Surgeons are often successful with the practice of curettage to bleeding bone to stimulate an indolent wound site. Autogenous bone grafting, or transfer of bone or particulate marrow, often from the iliac crest, is considered the gold standard for treating osseous deficiencies. Use of allogeneic, i.e., banked, segments of bone and bone-derived materials has increased because of their wide availability through accredited tissue banks.

Demineralized bone, properly prepared, is safe and consistent for osseous reconstruction.^[19–21] It invokes the mechanism of induced osteogenesis,^[22] in which nonosseous, connective, or muscle tissue is stimulated to form cartilage that is replaced by bone with hematopoietic marrow.^[23] The Bone Morphogenetic Protein (BMP) members of the TGF- β superfamily have been associated with the osteoinductive activity of demineralized bone.^[24] Use of demineralized bone implants or the BMPs may be considered examples of biological approaches to tissue engineering. The material stimulates the body to generate bone tissue by a process that resembles aspects of endochondral osteogenesis seen in embryogenesis and in fracture healing. In light of the schemes of tissue engineering (Fig. 1), demineralized bone may be viewed as a natural composite of an insoluble scaffold with factors or signals that induce osteogenesis in vivo. In-vitro studies to examine the molecular changes in cells targeted by osteoinductive material may reveal differences between embryonic skeletogenesis and postnatal induced osteogenesis that can be exploited for optimizing histogenesis.^[25]

Transplantation of cells from one part of the body, either directly or after expansion *in vitro*, is another approach for tissue construction (Fig. 1). It is clear from experimental^[26] and clinical^[27] studies that adult chondrocytes can be expanded *in vitro* and used to produce cartilage *in vivo*. Cartilage histogenesis has been investigated extensively, in large measure, because of its relative simplicity: cartilage is homogeneous in cell type and devoid of blood vessels. Adult articular cartilage has poor capacity for healing and requires innovative solutions for congenital and acquired degenerative cartilage lesions.^[28] Various scaffolds have been used to deliver chondrocytes and/or provide a template for cell distribution and matrix accumulation, including alginate gel,^[29] porous collagen sponges,^[30] and peptide hydrogels.^[31] There is great potential for stem cells as alternatives to harvesting normal chondrocyte cells for resurfacing joints and replacing damaged intervertebral disks and menisci.^[32] Marrow cells have been shown to be a source of chondrocyte progenitors/precursors and skin fibroblasts have been chondroinduced by culture with demineralized bone.^[33]

Standard concerns (Fig. 1) exist for engineering bone: source of cells, nature of the carrier, and effects of regulatory factors or genes.^[34] Bone tissue engineering has additional demands that highlight the differences between cartilage and bone: bone is a more rigid and vascularized tissue that is required to bear and adapt to great load. For *in-vitro* generation of functional tissues like bone, perfusion of medium improves the viability of cells within the mass and results in greater production of matrix.^[35] Marrow cells have the potential for differentiation into osteoblasts and supporting osteogenesis *in vivo*.^[36] Histogenesis of vascularized tissues can be approached by stimulating ingrowth of adjacent vasculature with angiogenic signals or by including capillary networks in the construct, for example, for stereolithography.

CONCLUSION

Histogenesis concerns the formation of tissue either in embryonic development, in spontaneous attempts to repair wounds, and in deliberate (or engineered) fabrication of missing parts. Understanding the mechanisms that the body uses to generate simple or complex tissues often provides direction for enhancing tissue regeneration. There are three considerations in designing a construct for engineered tissue: the source of cells, if any; the nature of the carrier or scaffold; and use, if any, of genes, factors, or adjuvants. Autogenous cells, often expanded *in vitro*, have been useful for cartilage tissue engineering. Stem/precursor/progenitor cells have shown potential for more complex tissues

like bone. Scaffolds may be a passive means to deliver and contain cells or to present factors or a micro-environment for histogenesis. The scaffolds can be modified to provide biological signals to augment repair and integration. The simplest solution for tissue deficiency would be to use an off-the-shelf scaffold made of a resorbable material that will have specific signaling factors on its surfaces. Thus, the scaffold could convey important regulatory information to reactive tissue, ensuring that it developed as desired. Other design elements could control the vascularization of the construct. It is unlikely that a combination of agents applied together would recapitulate the orderly sequence required for the generation of functional tissues. It would be desirable if the implant contained a single component that triggers the entire cascade of tissue development.

Many materials, design approaches, and manufacturing methods are amenable to tissue engineering for repair and regeneration. It is unlikely that a single approach will solve all problems of tissue and organ regeneration. Within the domain of musculoskeletal tissue and organ engineering, there are likely to be multiple solutions depending on the particular application. Fundamental studies of differentiation, growth, and postnatal tissue maintenance are likely to continue to inform design strategies. Less reliance on two-dimensional models of monolayer cell culture is needed for deconstructing and reconstructing the three-dimensional geometry of cellular interactions as they occur in tissues. Progress in materials design to ensure safe incorporation of new tissue must parallel advances in cell and tissue biology.

ARTICLES OF FURTHER INTEREST

Extracellular Matrix Scaffolds; Hematopoietic Stem Cells and Assays; Host Reactions; Stem Cells; Tissue Engineering; Tissue Engineering Scaffolds

REFERENCES

1. Murray, M.; Mankin, H.J.; Glowacki, J. The Musculoskeletal System. In *The Physiological Basis of Surgery*, 3rd Ed.; O'Leary, J.P., Ed.; Lippincott, Williams & Wilkins: Philadelphia, 2002; 577–595.
2. Martin, R.B.; Burr, D.B. *Structure, Function, and Adaptation of Compact Bone*; Raven Press: New York, 1989.
3. Mundy, G.R. Bone Remodeling. In *Primer on the Metabolic Bone Diseases and Disorder of Mineral Metabolism*; Favus, M.J., Ed.; Lippincott, Williams & Wilkins: Philadelphia, 1999; 30–38.
4. Glowacki, J. Mechanisms of Biom mineralization. In *Disorders of Bone and Mineral Metabolism*; Coe, F., Favus, M., Eds.; Lippincott, Williams & Wilkin: Philadelphia, 2002; 227–234.

5. Wolff, J. *The Law of Bone Remodeling*; Springer-Verlag: Berlin, 1986. Translated by P. Maquet.
6. Martin, R.B.; Burr, D.B.; Sharkey, N.A. *Skeletal Tissue Mechanics*; Springer: New York, NY, 1998.
7. Saunders, J.W.; Cairns, J.M.; Gasseling, M.T. The role of apical ectodermal ridge in the differentiation of the morphological structure of and inductive specificity of limb parts of the chick. *J. Morphol.* **1957**, *101*, 57–88.
8. Toole, B.P. Hyaluronan in Morphogenesis. *Semin. Cell Dev. Biol.* **2001**, *12*, 79–87.
9. Ingber, D.E. Mechanochemical Switching Between Growth and Differentiation by Extracellular Matrix. In *Principles of Tissue Engineering*; Lanza, R., Langer, R., Chick, W., Eds.; RG Landes Company: Austin, TX, 1997; 89–100.
10. Le Mouellic, H.; Lallemand, Y.; Brulet, P. Homeosis in the mouse induced by a null mutation in the *Hox-3.1* gene. *Cell* **1992**, *69*, 251–264.
11. Vastardis, H.; Karimbux, N.; Guthua, S.W.; Seidman, J.G.; Seidman, C.E. A human *MSX1* homeodomain missense mutation causes selective tooth agenesis. *Nat. Genet.* **1996**, *13*, 417–421.
12. Hess, A. Reaction of mammalian fetal tissue to injury. II. Skin. *Anat. Rec.* **1954**, *19*, 435–453.
13. Shah, M.; Foreman, D.M.; Ferguson, M.W. Neutralizing antibody to TGF-beta 1,2 reduces cutaneous scarring in adult rodents. *J. Cell. Sci.* **1994**, *107*, 1137–1157.
14. Clark, R.A.F. Wound Repair: Lessons for Tissue Engineering. In *Principles of Tissue Engineering*; Lanza, R., Langer, R., Chick, W., Eds.; R.G. Landes Company: Austin, TX, 1997; 737–768.
15. Marx, R.E. Platelet-Rich Plasma: A Source of Multiple Autologous Growth Factors for Bone Grafts. In *Tissue Engineering. Application in Maxillofacial Surgery and Periodontics*; Lynch, S.E., Genco, R.J., Eds.; Quintessence Books: Chicago, 1999; 71–82.
16. Giannobile, W. Periodontal Tissue Regeneration by Polypeptide Growth Factors and Gene Transfer. In *Tissue Engineering. Application in Maxillofacial Surgery and Periodontics*; Lynch, S.E., Genco, R.J., Eds.; Quintessence Books: Chicago, 1999; 231–243.
17. Eid, K.; Chen, E.; Griffith, L.; Glowacki, J. Effect of RGD-coating on osteocompatibility of PLGA-polymer disks in a rat tibial wound. *J. Biomed. Mater. Res.* **2001**, *57*, 224–231.
18. Glowacki, J. Tissue Response to Bone-Derived Materials. In *Bone Grafts: From Basic Science to Clinical Application*; Habal, M.B., Reddi, A.H., Eds.; W.B. Saunders: Philadelphia, 1992; 84–92.
19. Glowacki, J.; Kaban, L.B.; Murray, J.E.; Folkman, J.; Mulliken, J.B. Application of the biological principle of induced osteogenesis for craniofacial defects. *Lancet* **1981**, (i), 959–963.
20. Upton, J.; Glowacki, J. Hand reconstruction with allograft demineralized bone: 26 Implants in 12 patients. *J. Hand Surg.* **1992**, *17A*, 704–713.
21. Rosenthal, R.K.; Folkman, J.; Glowacki, J. Demineralized bone implants for non-union fractures, bone cysts, and fibrous bone lesions. *Clin. Orthop. Relat. Res.* **1999**, *364*, 61–69.
22. Urist, M.R. Bone: Formation by autoinduction. *Science* **1956**, *150*, 893–899.
23. Reddi, A.H.; Huggins, C.B. Biochemical sequences in the transformation of normal fibroblasts in adolescent rats. *Proc. Natl. Acad. Sci. U. S. A.* **1972**, *69*, 1601–1605.
24. Reddi, A.H. Bone morphogenetic proteins: From basic science to clinical applications. *J. Bone Jt. Surg., Am.* **2001**, *83-A*, S1–S6.
25. Yates, K.E.; Glowacki, J. Altered expression of connective tissue genes in postnatal chondroinduced human dermal fibroblasts. *Connect. Tissue Res.* **2003**, *44*, 1–7.
26. Grande, D.A.; Pitman, M.I.; Peterson, L.; Menche, D.; Klein, M. The repair of experimentally produced defects in rabbit articular cartilage by autologous chondrocyte transplantation. *J. Orthop. Res.* **1989**, *7*, 208–218.
27. Brittberg, M.; Lindhal, A.; Nilsson, A.; Ohlsson, C.; Isaksson, O.; Peterson, L. Treatment of deep cartilage defects in the knee with autologous chondrocyte transplantation. *New Engl. J. Med.* **1994**, *331*, 889–895.
28. Glowacki, J. In vitro engineering of cartilage. *J. Rehabil. Res. Dev.* **2000**, *37*, 171–177.
29. Hauselmann, H.J.; Fernandes, R.J.; Mok, S.S.; Schmid, T.M.; Block, J.A.; Aydelotte, M.B.; Kuettner, K.E.; Thonar, E.J. Phenotypic stability of bovine articular chondrocytes after long-term culture in alginate beads. *J. Cell Sci.* **1994**, *107*, 17–27.
30. Mizuno, S.; Tateishi, T.; Ushida, T.; Glowacki, J. Hydrostatic fluid pressure enhances matrix synthesis and accumulation by bovine chondrocytes in three-dimensional culture. *J. Cell. Physiol.* **2002**, *193*, 319–327.
31. Kisiday, J.; Jin, M.; Kurz, B.; Hung, H.; Semino, C.; Zhang, S.; Grodzinsky, A.J. Self-assembling peptide hydrogel fosters chondrocyte extracellular matrix production and cell division: Implications for cartilage tissue repair. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 9996–10001.
32. Glowacki, J. Engineered cartilage, bone, joints, and menisci. *Cells Tissues Organs* **2001**, *169*, 302–308.
33. Mizuno, S.; Glowacki, J. Chondroinduction of human dermal fibroblasts by demineralized bone in three-dimensional culture. *Exp. Cell Res.* **1996**, *227*, 89–97.
34. Mueller, S.M.; Glowacki, J. Construction and regulation of three-dimensional bone tissue in vitro. In *Bone Engineering*; Davies, J.E., Ed.; Em squared, Inc.: Toronto, 2000; 473–487.
35. Mueller, S.M.; Mizuno, S.; Gerstenfeld, L.C.; Glowacki, J. Medium perfusion enhances osteogenesis by murine osteosarcoma cells in three-dimensional collagen sponges. *J. Bone Miner. Res.* **1999**, *14*, 2118–2126.
36. Krebsbach, P.H.; Kuznetsov, S.A.; Bianco, P.; Robey, P.G. Bone marrow stromal cells: Characterization and clinical application. *Crit. Rev. Oral Biol. Med.* **1999**, *10*, 165–181.

Host Reactions

Kristin J. Pawlowski

*National Aeronautics and Space Administration (NASA), Langley Research Center,
Hampton, Virginia, U.S.A.*

INTRODUCTION

The introduction of a medical device into a human host can cause injury to the body and induce a variety of host reactions that are aimed at healing the injury and restoring homeostasis to the body. Essentially, the body views the device as a foreign material, and the overall goal is to protect itself from any potential damage and toxicity caused by this foreign material. The extent of injury and response is dependent upon what is implanted (size, shape, and chemical and physical properties), the site of implantation, and the surgical procedure used (injection, insertion, or surgical implantation). Patient factors are also important; each patient is different, making each case unique. In turn, the extent of response often defines biocompatibility of the material.

The overall response of the body to introduction of a foreign element can be divided into four phases. Injury is the initial event. Acute and then chronic inflammation are the next phases. These phases are coupled to other response mechanisms such as blood coagulation, platelet aggregation, and other blood-material interactions. In addition, immunological responses are invoked such as those involving T and B cells, the complement system, and hypersensitivity reactions. Outcome of these responses ultimately determines the extent of tissue colonization of the material and acceptance of the implant into the body environment. If tissue colonization prevails and healing begins, granulation tissue begins to form at the foreign element's surface. Following this, a series of foreign body reactions commences and may continue for the life of the implant. This phase depends primarily on the form and topography of the implanted material. In time, these physical properties could also lead to formation of tumors around the implant. Finally, if healing has continued, remodeling, which typically involves fibrosis or fibrous encapsulation, ensues, signifying final healing or walling-off of the element.

INFLAMMATION

Inflammation is the reaction of vascularized living tissue to local injury.^[1] This reaction is presented clinically as redness, heat, and swelling at the site of the wound.^[4] Immediately following injury, regardless of

implant site, alterations to vascular flow, caliber, and permeability occur.^[1,4] These alterations open up the vascular system and, in effect, set up a delivery system for inflammation.^[2,4] Fluid, proteins, and blood cells needed for host response to injury can then travel toward and access the injured tissue.

Acute Inflammation

The acute phase of inflammation typically lasts only several minutes to days.^[1] Polymorphonuclear leukocytes (PMNs) and mononuclear leukocytes predominate during this phase of inflammation and their primary responsibility is to phagocytose microorganisms and other foreign materials present at the injury site. This function is facilitated by factors called opsonins located naturally in blood plasma that coat a foreign element and contain receptors that then allow leukocytes to attach to and phagocytose the element.^[5] The key opsonins are Immunoglobulin G and complement fragment C3b.^[1] Typically, the biomaterial is too large for a cell to phagocytose; this situation is known as frustrated phagocytosis.^[3] In these cases, leukocytes release proteases, free radicals, and lysosomal enzymes that attempt to slowly degrade the material.^[1,3] There is evidence that the larger the foreign element, the larger the amount of these substances released by leukocytes, suggesting that the degree of inflammation is modulated by implant size.^[3]

Chronic Inflammation

Persistent presence of inflammatory stimuli leads to chronic inflammation. These stimuli include not only the chemical and physical properties of the biomaterial but also its motion.^[3] This phase is generally short in duration as well and is usually characterized by infiltration of macrophages, monocytes, lymphocytes, foreign body giant cells (FBGCs), fibroblasts, and endothelial cells.^[1] Presence of these cell types indicates the initiation of healing.

Lymphocytes mediate antibody and cell-mediated reactions.^[3] Monocytes are recruited to the wound site by damage debris such as elastin and collagen fragments and endothelial cell pieces, as well as injury-response



factors such as interleukins and platelet-derived factors.^[3] Once at the site of injury, monocytes differentiate into macrophages, which start to remove any debris left at the site.^[3,4] Macrophages also secrete several substances such as chemotactic factors, coagulation factors, and complement components that are important for healing.^[1] FBGCs are multinucleated cells thought to form by fusion of macrophages and, in that capacity, contribute to the healing process and attempts by the body to isolate the foreign element.^[3] Fibroblasts and endothelial cells begin to release materials and organize themselves for the rebuilding of connective and vascular tissue.

GRANULATION TISSUE

If healing has initiated, granulation tissue can be seen as early as 3–5 days following implantation of a material. Granulation tissue contains four main components. Macrophages clear debris and begin to break down blood clots. Neutrophils attack bacteria that are present or are attempting to infiltrate the area.^[4] New small blood vessels, which have developed from proliferation, maturation, and organization of endothelial cells, begin to form. Finally, fibroblasts begin to synthesize components of connective tissue such as collagens and proteoglycans.^[1,2] In addition, some fibroblasts present have features of smooth muscle cells; these cells are thought to cause the tissue contraction seen during wound healing.^[1] The resulting tissue, called granulation tissue, is a transitional compilation of cells and tissue that are either in the process of being degraded from initial inflammation and clotting responses or are being built to contribute to the healing process.

FOREIGN BODY REACTION

The foreign body reaction consists of granulation tissue components, foreign body giant cells (FBGCs), and macrophages to varying degrees. This reaction and the extent of involvement of each component is highly variable depending on the form and topography of the biomaterial surface.^[1] The components of the foreign body reaction often form a layer that covers the surface of the biomaterial. This layer may persist throughout the life of the implant and is typically enclosed within the final fibrous capsule.

REMODELING

Remodeling is the last event in healing or repairing of the implant site and can involve regeneration of native

tissues present at the site prior to placement of the material or replacement with connective tissue known as the fibrous capsule.^[2] The predominant process is decided by two factors: the proliferative capacity of the cells in the tissue or organ receiving the foreign material and the extent of injury to the tissue framework, the supporting structure of parenchymal cells, of the implant site.^[1] Regeneration with the same or similar tissue and restoration of function is rare. Fibrosis or fibrous encapsulation is the more typical route and is the final walling-off of the implanted biomaterial. Granulation tissue that began to form earlier in the healing process builds to form a capsule that surrounds the implant and its associated foreign body reaction layer, isolating it from the local tissue environment.^[1,4] At this point, it is possible for normal local processes to resume, if the degree of injury or defect allows.

COMPLEMENT SYSTEM

The primary role of the complement system is the direct killing, opsonization, and clearance of microorganisms.^[5] The system is initiated by the presence of bacteria or immune response complexes and, once initiated, cascades to eventually involve at least 20 proteins.^[5,6] The various components are located in plasma and were named as they were discovered and found to be necessary for antibody-mediated bactericidal activity.^[1] As the system was elucidated, the components were placed in the order in which they occur in the cascade.^[4]

Two pathways can be initiated, depending on the initiator. However, the most pivotal step in both pathways is the activation of complement component C3. The classical pathway, discovered first, is typically activated by specific antibody-coated targets or antigen-antibody complexes.^[4] Specifically, immunoglobulins IgM, and most IgG subclasses, are capable of activating this pathway. The alternative pathway is triggered by more general initiators and is, therefore, the more commonly employed system of defense against infectious agents, as it is capable of clearing a wider variety of potential pathogens without the presence of specific antibodies.^[5] Liposaccharides of Gram-negative bacteria, a factor in cobra venom, and plasmin are some examples of initiators of the second pathway.^[4]

Classical Pathway

The first step in the classical pathway, as illustrated in Fig. 1, is the binding of activated antibodies to C1. This binding allows C1 to cleave C4 and C2, which results in three products. The most important is

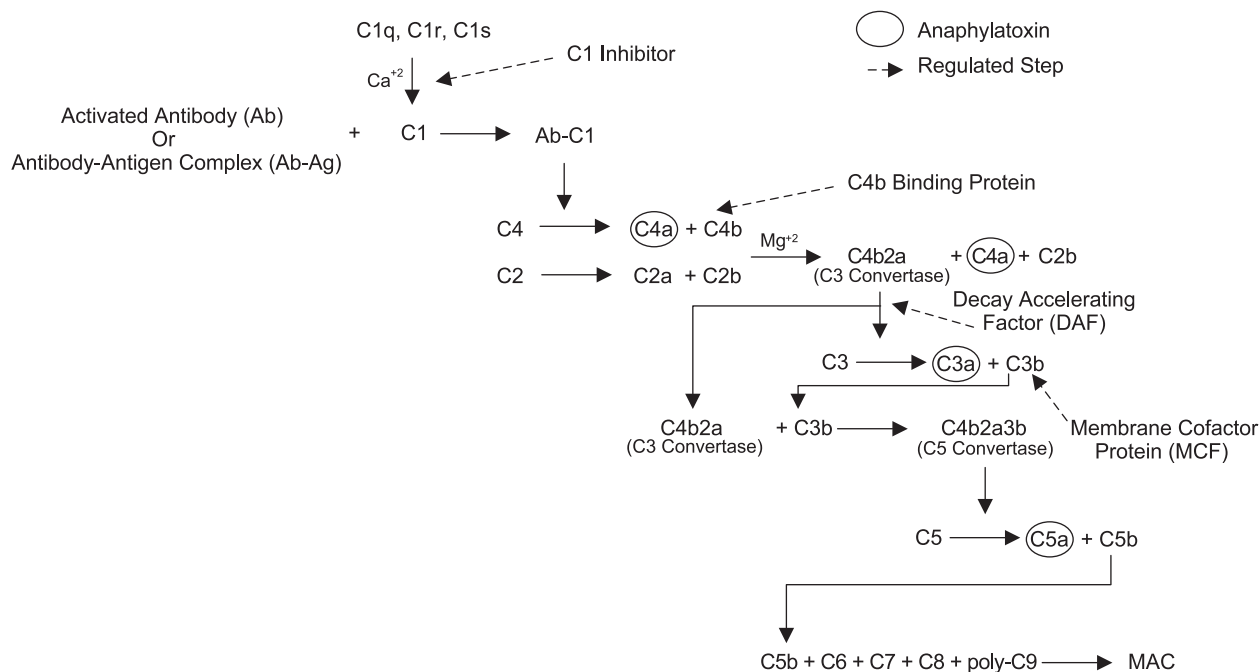


Fig. 1 The classical pathway of the complement system.

the C4b2a complex known as C3 convertase.^[4] C4a, a weak anaphylatoxin, and C2b, a kinin-like molecule that also causes vasodilation and increased vascular permeability, are also released.^[6] C3 convertase then cleaves C3 into C3a and C3b. This step is important for amplification of the classical pathway: One molecule of C3 convertase has the ability to cleave hundreds of molecules of C3.^[4] C3a, another anaphylatoxin, then causes the release of histamine from various cells, resulting in mast cell degranulation and increases in vascular permeability.^[4] Clinically, individuals with C3 deficiency are prone to severe pyrogenic infections such as pneumococcal pneumonia and meningococcal meningitis.^[6] Lack of C3 to bind to microorganisms leads to a large decrease in phagocytic activity.

This leaves C3b to activate C5 through C9. C5 is cleaved to C5a, another anaphylatoxin that has the same effects as C3a and acts as a chemotactic factor for phagocytic cells, and C5b. C5b, C6, C7, C8, and C9 then form a variety of complexes with one another to produce a series of reactions. C8 and C9 attach to and cause lysis of antibody-coated cells. C5b combines with C6 to begin formation of the membrane attack complex (MAC), eventually composed of C5b–C9. MAC inserts itself into the cell membrane of microorganisms, disrupting the cell membrane proteins and causing formation of hydrophilic membrane channels, eventually leading to cell lysis.^[6] One molecule of this complex is potent enough to penetrate and destroy a cell.^[4]

The classical pathway can also be initiated via antibody-independent activation. Mannose-binding lectin (MBL) can bind to mannose sugars on the surface of an antigen or microorganism. Once bound, MBL interacts with two mannan-binding lectin-associated serine proteases called MASP and MADSP2. This complex then cleaves C4 and C2, and the cascade continues exactly as the classical pathway. This pathway is often called the Lectin Pathway of complement activation.^[7,8]

Alternative Pathway

The alternative pathway, in contrast, is activated by direct interaction with pathogens and is, therefore, more important to the initial stages of the host response than the classical pathway. As shown in Fig. 2, the first step is activation of factor P (properdin) by alternative pathway triggering agents. In addition, hydrolysis of C3 is required to advance the cascade; interaction with water causes a change in conformation of C3, allowing it to contribute to the alternative pathway.^[4] Activated factor P then triggers factor D, which subsequently cleaves factor B and allows it to interact with altered C3. The proteolytic complex formed, C3bBb, is a C3 convertase that can, therefore, cleave more C3, starting a feedback amplification loop that intensifies the alternative and the classical pathway responses.^[6] C3b is created in the process and goes on to initiate formation of the MAC, as in the classical pathway.

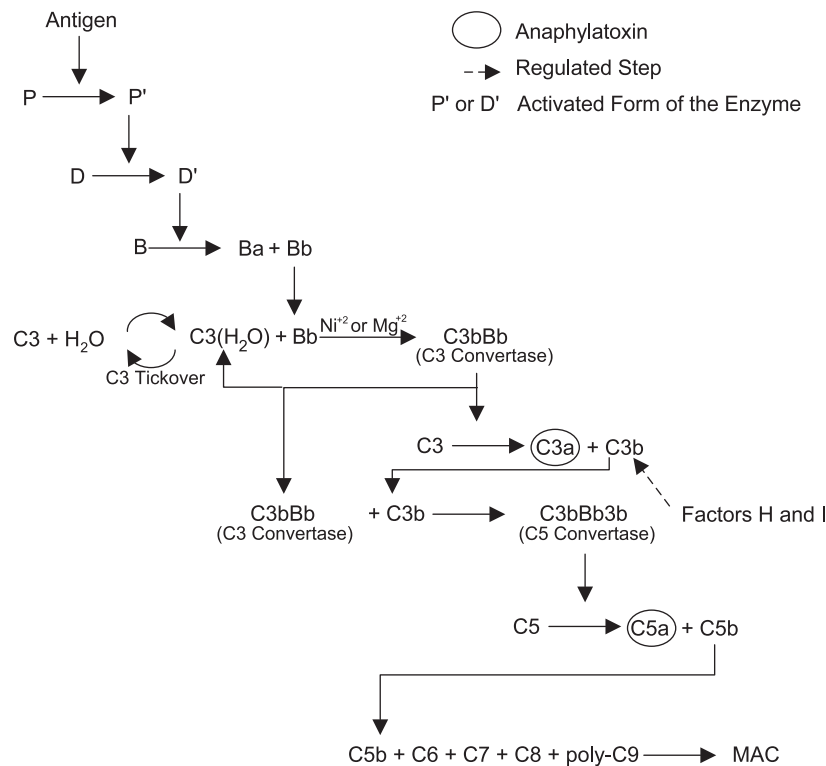


Fig. 2 The alternative pathway of the complement system.

Biologic Activities of Complement Factors and the Complement System

To summarize, there are five major active complement components. Factors C3a, C4a, and C5a are all called anaphylatoxins. Though C4a is probably the least important, these factors cause mast cells and basophils to degranulate and release histamine, leading to increases in vascular permeability. These responses are presented clinically as vasodilation, redness, swelling, and possibly smooth muscle contraction and mucus secretion in the airways.^[6] C5a is also an important chemotactic factor for PMNs, monocytes, and eosinophils.^[4] C3b is an opsonin, which binds to foreign elements that are present and contains receptors that then allow leukocytes to attach to and phagocytose the element. C3b also aids antibody-dependent cell-mediated lysis.^[6] Finally, C5b–C9 compose the MAC, which causes target cell lysis, as described above.

Creation and action of these factors is a key element in the success of the implanted biomaterial. In attempts to rid the body of the biomaterial, the complement factors aid destruction of any bacterial cells associated with the implant. These factors also help guide inflammatory and immune cells that are en route to the area in response to implantation and facilitate their actions, ensuring that the proper responses can continue to completion.

RACE FOR THE SURFACE: BACTERIAL VERSUS TISSUE COLONIZATION

Bacterial infection is a large risk factor for medical device implantation. The rate of infection can range from less than 0.5% for artificial hips to 100% for total artificial heart replacements.^[9] These rates are typical even with the use of systemic antibiotics. Infection can be triggered from an improperly handled implant or due to problems at the wound or incision site. The discussion below will focus primarily on infection due to the presence of the implant in the body.

Bacteria that most frequently populate medical device surfaces are *Staphylococcus aureus*, *Staphylococcus epidermidis*, and *Pseudomonas aeruginosa*.^[10] *S. aureus* is most frequently seen at metal implant surfaces, and *S. epidermidis* is typically seen in infections of polymeric implants. *P. aeruginosa* is most frequently associated with keratitis resulting from extended-wear contact lenses.

Prevention of bacterial colonization through tissue cell integration is the best approach to avoidance of infection. If a bacterial layer is allowed to adhere to the device surface before tissue cells can colonize, it is very difficult to rid the surface of the bacterial layer. Once adhered, bacteria are strong, aggressive, and capable of modifying the environment to support their livelihood. Infection ensues, and initiation of normal healing is

severely hampered, if not impossible. Often, relief comes only with outside interventions such as reoperation.^[10]

To avoid this catastrophe, tissue cells are involved in a race for the surface with bacteria immediately upon implantation of a biomaterial.^[10] If tissue cells are able to populate the device surface first, a protective layer composed of tissue cells and a supporting matrix of proteins and polysaccharides is formed. This layer is resistant to typical infection routes and provides an optimal environment for the function of other host mechanisms.^[9]

Developers of biomaterials take several approaches to promoting population of tissue cells, avoiding bacterial adhesion, and ensuring proper integration of the material into the body. Paramount is choice of a biocompatible material. Meticulous attention to and use of the sterile technique during implantation is another step in the prevention of bacterial colonization.^[11] Traditionally, treatment with systemic antibiotics prior to and following implantation has been the method of management. Surface modifications to the material and overall design of the implant are newer methods utilized to avoid initiation of infection. Texturing or addition of an outer layer of microbeads is one method that has been utilized successfully with orthopedic implants to improve the process of integration with bone. Preimplantation cell seeding and coating with local antibiotics or proteins that promote tissue cell attachment are other approaches.

As mentioned, compatible material choice is one of the first steps in implant development. The choices of materials are currently evolving with the advent of tissue-engineered materials, which consist partially or completely of materials that originate in the body.^[12] Host response to these materials is minimal, since the material is recognized as native. Already xenografts and homografts are used in some devices such as tissue heart valves. Tissue-engineering designs are becoming more and more complex with the development of composite materials that incorporate tissue and cellular components. The goal is seamless integration of these materials into the body so that they become a member of the functional and remodeling environment of the implant site.

BLOOD-MATERIAL INTERACTIONS

Another consideration for most implants and especially if the implantation site involves direct interaction with the bloodstream, such as a vascular graft or stent, is the response of blood components to the material. This response is often referred to as blood-material interaction (BMI). Implantation of a material or device into a blood-contacting site causes injury to the site.

More serious, the material itself is likely to be thrombogenic and induce response from the blood. The initial response is activation of blood cells known as platelets. When in a stable state, platelets are smooth and discoid in shape. However, they are easily activated, at which point these cells change shape, become sticky, release pseudopods, and secrete granule contents that activate additional platelets.^[1,13] These changes lead to aggregation of platelets, blood coagulation, and, eventually, formation of thrombus on the material surface. More serious, this thrombus could develop and occlude a vessel or be shed as an embolus and travel to the heart or brain, blocking blood flow and oxygenation.^[14]

Platelets become activated by interaction with sub-endothelial collagen located in blood vessel walls. This collagen becomes accessible when the endothelial layer is damaged by biomaterial implantation. Activated platelets have exposed surface receptors that interact with particular plasma proteins such as fibronectin, fibrinogen, vitronectin, and von Willebrand factor that are already adsorbed to device surfaces. The platelet receptor GPIIb/IIIa can bind to these plasma proteins.^[1,13,15] In addition, fibrinogen can serve as a platelet-platelet bridge, allowing platelet aggregation to proceed.^[16]

Aggregation is followed by coagulation. As in the complement system, two pathways for coagulation exist. These are illustrated in Fig. 3. The intrinsic pathway is activated by contact with and adsorption of factors contained in the vascular system onto a negatively charged surface. The extrinsic pathway is initiated by tissue factor (TF), which is expressed by white blood cells in response to tissue fragments in the blood and exposed vascular layers created by damage.^[1,16] Both pathways lead to a common point, however, and the result is fibrin formation and hemostasis.

The intrinsic pathway begins with cleavage of factor XII to factor XIIa. Contact activation then allows factor XIIa to convert factor XI to factor XIa, which, with the help of calcium ions, cleaves factor IX to factor IXa. With the additional help of calcium ions and factor VIII provided by platelets, factor IXa activates factor X.^[1] This step requires a platelet phospholipid surface. In the extrinsic pathway, TF activates factor VII, a vitamin K-dependent reaction. Factor VIIa results and activates factor X with help from circulating calcium ions.^[16]

With activation of factor X, the common pathway of coagulation begins. Factor Xa causes conversion of prothrombin to thrombin, a reaction that requires calcium and factor V as the platelet receptor for factor Xa, as well as platelet phospholipids.^[16] Thrombin then detaches from the platelet surface and activates a number of additional coagulation components, advancing as well as amplifying the coagulation cascade.



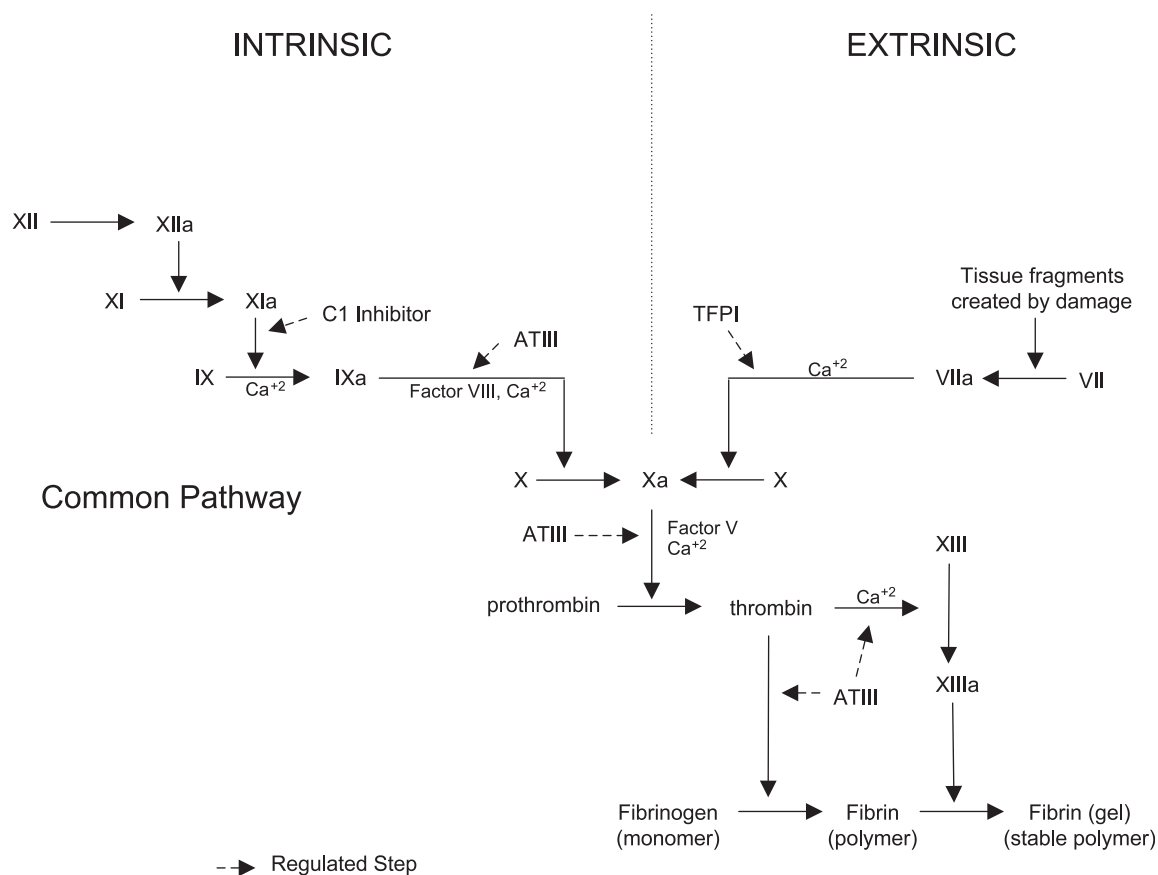


Fig. 3 The coagulation cascade.

First, it further activates factors VIII and V. Second, with the help of calcium, thrombin converts factor XIII to factor XIIIa. Third, it directly causes polymerization of the fibrinogen monomer to the fibrin polymer. Subsequently, activated factor XIIIa helps to stabilize the fibrin polymer by inducing crosslinking. Hemostasis is now well under way.

When the clot is no longer needed, fibrinolysis ensues. Plasminogen is activated to plasmin by tissue-Plasminogen Activator (tPA), which attaches to and digests the fibrin clot.^[16] Fibrin degradation products are then washed away from the clot site.

Throughout the coagulation cascade, there are a number of regulation factors that block its progression if necessary. Antithrombin III (ATIII) is the most active. It can block action of factor IXa on factor X, factor Xa on prothrombin, and thrombin on factor XIII and on fibrinogen. Clinically, action of ATIII can be enhanced by heparin.^[16] C1 inhibitor can block conversion of factor IX by neutralizing factor XIa. Tissue Factor Pathway Inhibitor (TFPI) is capable of blocking conversion of factor X.

Design of successful biomaterials requires attention to and minimization of these blood-material interactions. Researchers have developed a number of

strategies toward this goal. The obvious first step is to choose a blood-compatible material. Following this selection, form, topography, and possible surface modifications to the material are the next steps. One example is the use of expanded poly(tetrafluoroethylene) for the replacement of blood vessels. Replacement of large vessels with this material has been very successful; however, if the inner diameter of the prosthesis is too small, blood clot formation causes complete blockage of flow through the implant. One method designed to avoid this is systemic application of heparin or other anticoagulants prior to implantation and during healing. Another approach is the addition of an endothelial layer to the inside of the prosthesis.^[17] In native blood vessels, the endothelial layer provides a smooth contact surface for flowing blood. Presence of this layer in implanted prostheses may provide the same effect.

THE IMMUNE SYSTEM

The function of the immune system is to protect the body from any foreign presence, called an antigen, that could eventually lead to disease. Biomaterials and their

breakdown or wear products are one possible type of antigen. The body utilizes several mechanisms to carry out the steps of the immune response. The components of the response are organized in the lymphatic system and are in constant communication via the blood and lymphatic networks.^[6] Collectively, these components work together to protect the body from almost any antigen. First, the antigen must be identified as foreign; this is accomplished by antigen-presenting cells (APCs) such as macrophages, endothelial cells, and Langerhans' cells. These cells have the ability to break down the antigen into a form that the rest of the immune response system can recognize and attack. Macrophages also possess phagocytic capabilities and can lyse tumor cells and produce cytokines, as discussed in the following section.

Once the antigen is recognizable, lymphocytes respond and attempt to destroy the antigen. There are three types of lymphocytes. Initially, natural killer (NK) cells provide a generalized response, usually to tumor cells and virus-infected cells, during the time that the other two types of lymphocytes are being stimulated and amplified.^[4] These other lymphocytes, T cells and B cells, are actually programmed to respond to specific antigens. T cells are responsible for cell-mediated immune response. B cells are responsible for humoral immune response.

Cell-Mediated Immune Response

T cells act by producing cytokines and other factors that have the ability to fight a specific antigen. A T cell binds to an antigen via a receptor on its surface that recognizes an antigen only if the antigen is attached to an APC and has bound to a complex called the major histocompatibility complex (MHC), discussed later. Binding of the T-cell receptor to the antigen-APC-MHC complex requires presence of magnesium ions; furthermore, T cells cannot send a destruction signal to the antigen without the presence of calcium ions.^[6]

Cytokines, as discussed above, regulate the cellular response of the immune system. Their work usually involves communication between leukocytes; thus, many cytokines are called interleukins (ILs). Other categories of cytokines are interferons (IFNs), which focus on antiviral, immunomodulatory, and anti-proliferative activities, and tumor necrosis factors (TNFs), which regulate cytotoxic activity against tumor cells exclusively.^[1] Other cytokines function as growth factors, chemokines, and regulatory factors. For example, erythropoietin is a hematopoietic growth factor responsible for the development of red blood cells.^[6]

In general, cytokines have much in common. They are potent in the picomolar range, act through very specific cell surface receptors, are approximately

100–200 amino acids long, are typically glycosylated, and have overlapping synergistic or antagonistic activities. The most studied cytokine is IL-1, which has been found to promote coagulation and cellular accumulation at inflammatory sites and to further promote wound repair and fibrosis by inducing fibroblast cells to make collagen types I, III, and IV.^[1]

Humoral Immune Response

B cells are stimulated by attachment to a specific antigen via B-cell surface ligands called antibodies. Antibodies, also called immunoglobulins (Ig), bind specific antigens and, in effect, facilitate the work of the B cells. Once bound to the antigen, B cells can differentiate into plasma cells, which produce more antibodies against the particular antigen that is present. T cells can also stimulate B cells, leading to production of antibodies and further processing of these antibodies for even more specific actions or purposes.^[1]

As discussed above, antibodies bind to antigens and facilitate B-cell stimulation. This antigen-antibody complex is also an initiator of the classical pathway of the complement system. Antibodies have additional functions, as well. They can act as linkers between two antigens causing them to cluster; this corrals the antigens, making it easier for neutrophils and macrophages to phagocytose them. They have antibacterial activity and are, therefore, capable of neutralizing bacterial toxins or coating bacteria to aid in the recognition of the antigen for phagocytes. Finally, antibodies can combine with viruses and prevent them from entering cells.^[4]

There are five types of antibodies or immunoglobulins, but the basic structure of each is the same. As illustrated in Fig. 4, they are composed of two light and two heavy chains, connected by hydrogen and disulfide bonds. Only the end portion of the structure is variable, giving rise to the different types of antibodies. IgM antibodies are found predominantly in the intravascular space, as they are too big to gain access to other locations.^[4] They are typically involved in the

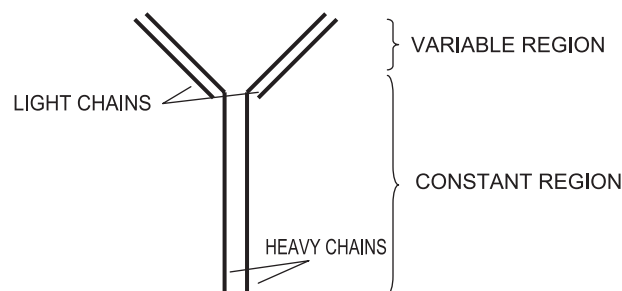


Fig. 4 Basic structure of an antibody.



primary immune response and are the most efficient antibody in activation of the complement system. IgD antibodies are observed in trace amounts in the serum and may play a role in B-cell maturation. Both IgM and IgD antibodies are located on the surface of B cells, the typical receptor for specific antigen, as discussed previously. IgE antibodies are usually found bound to immune cells and likely have a role in responding to the presence of parasites and anaphylaxis initiators. IgA antibodies are located in saliva, tears, and the mucosal coatings of the respiratory, gastrointestinal, and genitourinary tracts, preventing attack of these delicate surfaces.^[4,6] IgA antibodies are also effective in inducing the alternative complement pathway. The fifth type of antibody, IgG, is typically found in serum and is the only antibody with the capacity to cross into the placenta. Therefore, most antibodies in newborn babies are of type IgG and are, therefore, called maternal antibodies.^[4]

SYSTEMIC TOXICITY AND HYPERSENSITIVITY REACTIONS

At times, an immune response to an antigen can cause systemic toxicity and lead to damage to the tissues of the host. These types of reactions are called hypersensitivity reactions. It is difficult to predict or test the types and severity of immune responses invoked by the presence of a biomaterial because they are highly dependent upon the individual and the dose, composition, nature, and location of the implant.^[1] However, hypersensitivity reactions are and should be a concern for biomaterial and implant designers.

There are four types of systemic immune reactions.^[4] Types I, II, and III are fast-acting and mediated by immunoglobulins and the humoral immune response. Type IV takes longer to develop and is controlled by T cells and other cell-mediated responses. Type I, also called immediate-type or anaphylaxis, is usually induced by pollens, drugs, food, dust, and insect stings or bites. The anaphylactic response begins with interaction between an IgE-type antibody and the antigen. Subsequently, the antibody portion of this complex attaches to host cells, resulting in stimulation of the complement system and production of anaphylatoxins. Clinically, symptoms of Type I hypersensitivity can be hives, fever, asthma, and/or vomiting due to release from cells of substances such as histamine, serotonin, and other vasoactive substances.^[1] Immune cytotoxic or Type II reactions are similar to Type I reactions; however, the antigen attaches to the host cells and leads to interaction with IgG- and IgM-type antibodies, leading to the destruction of the cell or tissue via complement activation, phagocytosis, or NK

cell activity. Instances of stimulation of either Type I or Type II hypersensitivity reactions by biomaterials are rare.^[1]

Type III reactions are classed as immune complex hypersensitivity. The mechanisms are similar to those of Type II responses. These reactions occur following constant or recurring exposure to the same antigen, resulting in high levels of antibody to that antigen in the serum. Type III reactions can be divided into two categories, separated mainly based on the size of the immune complex involved.^[4] Arthus-type reaction is the precipitation of large antigen-antibody complexes that deposit in blood vessels, the kidneys, and the heart, especially at bifurcations and locations of high turbulence. The result is formation of anaphylatoxins, destruction of blood vessel walls, and activation of platelets, possibly leading to occlusive thrombi and even emboli. Serum sickness-type reaction is a result of excess circulating small and soluble antigen or antibody that can also deposit in blood vessel walls, leading to complement activation and necrotizing vasculitis.^[4] With the exception of biodegradable or drug delivery systems, Type III hypersensitivity due to biomaterials is unlikely.^[1]

Type IV or delayed-type hypersensitivity develops slowly, often over a 2–3-day period. Antibody is not employed in these reactions; they are cell-mediated. Delayed-type reactions are evident clinically as four main conditions. Most familiar is contact dermatitis such as poison ivy. Second, persistent infection from bacteria, viruses, or fungi can result in chronic infections such as the lesions observed with tuberculosis patients. Third, known as autoimmune disease, an immune response against self-antigens of the host can be initiated if the immune system cannot discriminate between foreign antigens and native components. Finally, graft rejection and deep tissue reaction is the fourth type of condition resulting from Type IV hypersensitivity. These reactions involve cytotoxic T cells and NK cells and are the result of a complex immunologic response to cell surface antigens of the graft. The major histocompatibility complex (MHC), the most significant factor in modern immunology, mediates this complex response.^[6] The MHC is a cluster of genes on human chromosome 6,^[4] and the protein coded by it is located on the cell membrane of APCs and almost all nucleated mammalian cells, depending on the MHC molecule class.^[6] It is this complex that determines if tissues and organs will be compatible between donors and recipients; in order for compatibility, the MHC genes of both must be the same. MHC molecules have a role in T-cell activation by forming the ligand between the antigen-APC and the T cell.

TUMORIGENESIS

Though it is often difficult to link it as a direct cause, there is evidence that implanted biomaterials can cause tumor growth and induce cancer. This situation is a major concern for biomaterial researchers. Formation of a tumor results from neoplasia, which is defined as excessive and uncontrollable cell proliferation. Neoplasia is likely caused by excessive foreign body reaction, and the cells that develop have immunity to the body's natural defense systems. If the growth cannot infiltrate the tissues or spread to other areas of the body, it is called benign; however, if it can access the lymphatic or blood circulation and spread, it is called malignant or cancerous.^[1] Solid materials with high surface area are most tumorigenic, and initiation of neoplasia depends primarily on physical not chemical characteristics of the implant.^[1] Though general studies have been performed for years, more recent research is focused on a deeper understanding of the mechanisms of biomaterial-induced tumorigenesis and how to avoid or minimize it.^[18–21]

CONCLUSION

The response of a host to implantation and presence of a medical device or biomaterial is complex. The severity of the response depends upon interrelations between many systems and the clues that these systems take from the implanted material. As a biomedical researcher, a keen understanding of these systems is essential in striving to minimize host response. Fortunately, information about host responses is constantly being enriched, guiding researchers toward optimal solutions to biomaterial and device challenges.

ARTICLES OF FURTHER INTEREST

Antimicrobial/Antibiotic (Infection Resistance) Materials; Biocompatibility Testing; Biologic and Synthetic Apatites; Biomaterials Immune Response; Biomaterials: Protein–Surface Interactions; Blood–Material Interactions; Corrosion; Foreign Body Response; In Vivo Cardiovascular Models; Infection of Medical Devices; Vascular Grafts: Host Interactions

REFERENCES

1. Anderson, J.; Gristina, A.; Hanson, S.; Harker, L.; Johnson, R.; Merritt, K.; Naylor, P.; Schoen, F. Host Reactions to Biomaterials and Their Evaluation. In *Biomaterials Science: An Introduction to Materials in Medicine*; Ratner, B., Hoffman, A., Schoen, F., Lemons, J., Eds.; Academic Press, Inc.: San Diego, 1996; 165–214.
2. Silver, F. *Biomaterials, Medical Devices, and Tissue Engineering: An Integrated Approach*; Chapman & Hall: London, 1994.
3. Shankar, R.; Greisler, H. Inflammation and Biomaterials. In *Implantation Biology: The Host Response and Biomedical Devices*; Greco, R., Ed.; CRC Press: Boca Raton, FL, 1994; 67–80.
4. Trowbridge, H.; Emling, R. *Inflammation: A Review of the Process*; Quintessence Publishing Co., Inc.: Chicago, 1997.
5. *Immunobiology of the Complement System: An Introduction for Research and Clinical Medicine*; Ross, G., Ed.; Academic Press, Inc.: Orlando, FL, 1986.
6. *Immunobiology and Inflammation: Basic Mechanisms and Clinical Consequences*; Sigal, L., Ron, Y., Eds.; McGraw-Hill, Inc.: New York, 1994.
7. www.med.sc.edu:85/ghaffar/complement.htm. (accessed July 2003).
8. www.medschool.vcu.edu (accessed July 2003).
9. Gristina, A.; Giridhar, G.; Myrvik, Q. Bacteria and Biomaterials. In *Implantation Biology: The Host Response and Biomedical Devices*; Greco, R., Ed.; CRC Press: Boca Raton, FL, 1994; 131–148.
10. Gristina, A. Bacterial-centered infection: Microbial adhesion versus tissue integration. *Science* **1987**, *237*, 1588–1595.
11. Schierholz, J.; Beuth, J. Implant infections: A haven for opportunistic bacteria. *J. Hosp. Infect.* **2001**, *49*, 87–93.
12. *Tissue Engineering and Biodegradable Equivalents: Scientific and Clinical Applications*; Lewandrowski, K., Wise, D., Trantolo, D., Gresser, J., Yaszemski, M., Altobelli, D., Eds.; Marcel Dekker, Inc.: New York, 2002.
13. Sheppeck, R.; LoGerfo, F. Blood and Biomaterials. In *Implantation Biology: The Host Response and Biomedical Devices*; Greco, R., Ed.; CRC Press: Boca Raton, FL, 1994; 55–65.
14. Hanson, S.; Ratner, B. Testing Biomaterials: Testing of Blood-Materials Interactions. In *Biomaterials Science: An Introduction to Materials in Medicine*; Ratner, B., Hoffman, A., Schoen, F., Lemons, J., Eds.; Academic Press, Inc.: San Diego, 1996; 228–238.
15. Choi, E.; Callow, A. The Effect of Biomaterials on the Host. In *Implantation Biology: The Host Response and Biomedical Devices*; Greco, R., Ed.; CRC Press: Boca Raton, FL, 1994; 39–53.
16. Colman, R.; Clowes, A.; George, J.; Hirsh, J.; Marder, V. Overview of Hemostasis. In *Hemostasis and Thrombosis: Basic Principles and Clinical Practice*; Colman, R., Hirsh, J., Marder, V., Clowes, A., George, J., Eds.; Lippincott Williams & Wilkins: Philadelphia, 2001; 3–16.
17. Bowlin, G.; Schmidt, S.; Rittgers, S.; Pawlowski, K. Endothelialization of Vascular Prostheses. In *Biomaterials Engineering and Devices: Human Applications: Fundamentals and Vascular and Carrier Applications*; Wise, D., Trantolo, D., Lewandrowski, K., Gresser, J., Cattaneo, M., Yaszemski, M., Eds.; The Humana Press, Inc.: Totowa, NJ, 2000; 85–118.



18. Kirkpatrick, C.; Alves, A.; Kohler, H.; Kriegsmann, J.; Bittinger, F.; Otto, M.; Williams, D.; Eloy, R. Bio-material-induced sarcoma: A novel model to study preneoplastic change. *Am. J. Pathol.* **2000**, *156* (4), 1455–1467.
19. Nakaoka, R.; Tsuchiya, T.; Nakamura, A. Studies on the mechanisms of tumorigenesis induced by polyetherurethane in rats: Production of superoxide, tumor necrosis factor, and interleukin 1 from macrophages cultured on different polyetherurethanes. *J. Biomed. Mater. Res.* **2000**, *49*, 99–105.
20. Nakaoka, R.; Tsuchiya, T.; Nakamura, A. Studies on the tumor-promoting activity of polyethylene: Inhibitory activity of metabolic cooperation of polyethylene films containing an antioxidant. *J. Long-Term Eff. Med. Implants* **1995**, *5* (4), 253–262.
21. Kinoshita, Y.; Kuzuhara, T.; Kobayashi, M.; Ikada, Y. Reduction in tumor formation on polyethylene by collagen immobilization. *J. Long-Term Eff. Med. Implants* **1995**, *5* (4), 275–284.

Hyaluronan

Jennie Baier Leach

Department of Chemical Engineering, University of Texas, Austin, Texas, U.S.A.

Christine E. Schmidt

Department of Biomedical Engineering and Texas Materials Institute, University of Texas, Austin, Texas, U.S.A.

INTRODUCTION

Hyaluronan (HA, also called hyaluronic acid or hyaluronate), a high molecular weight glycosaminoglycan found in all mammals, is a unique and highly versatile biopolymer. HA plays a vital role in embryonic development, extracellular matrix homeostasis, wound healing, and tissue regeneration. Although it is clear that HA is critically involved in these processes, the exact mechanisms by which it influences cellular behavior are not yet elucidated. Complicating such studies is the fact that the behavior and cellular influence of HA is highly dependent on its concentration and molecular weight. Therefore, the biology and chemistry of HA is a current area of extensive scientific research. Biomaterials made from derivatized and cross-linked HA have received a great deal of attention in the bioengineering community and have been used in a wide variety of applications (e.g., orthopedic, cardiovascular, ophthalmology, and dermatology, as well as general applications in tissue engineering, surgery, and drug delivery). HA is naturally derived, nonimmunogenic and also has multiple sites for modification and inherent biological activities; therefore, this unique biopolymer has great potential for creative application in biomedical engineering. This article reviews HA's natural structure and function, and biomaterials made from derivatized and cross-linked HA, as well as a number of clinical, surgical, and bioengineering applications of HA.

THE NATURAL STRUCTURE AND FUNCTION OF HYALURONAN

Physicochemical and Structural Properties

Hyaluronan, an extracellular matrix (ECM) component, is a high molecular weight glycosaminoglycan composed of disaccharide repeats of N-acetylglucosamine and glucuronic acid (Fig. 1). This relatively simple structure is conserved throughout all mammals, suggesting that HA is a biomolecule of considerable importance.^[1] In the body, HA occurs in the salt form,

hyaluronate, and is found in high concentrations in several soft connective tissues, including skin, umbilical cord, synovial fluid, and vitreous humor. Notable amounts of HA are also found in lung, kidney, brain, and muscle tissues. In commercial production, HA is commonly extracted from rooster comb and human umbilical cord or manufactured in large quantities by bacterial fermentation.^[2]

The cellular synthesis of HA is a unique and highly controlled process.^[3] Most glycosaminoglycans are made in the cell's Golgi networks. HA, however, is synthesized at the plasma membrane and immediately extruded out of the cell and into the ECM. This process is carried out by a group of proteins called HA synthases, which are located in the cell membrane. For a review of HA synthases, see Ref. [3].

HA's structure imparts unique physicochemical and biological properties that depend on molecular weight.^[2] When extracted from tissues, HA is polydisperse in size, with an average molecular weight of several million. In physiological solution, HA conforms to a stiffened random coil with a contour length of about 2.5 μm for a 1×10^6 molecular-weight chain of about 2650 disaccharide repeats. Secondary hydrogen bonds form along the axis of HA, imparting stiffness and creating hydrophobic patches that allow the formation of ordered structures. HA solutions are highly viscoelastic; in other words, when shear rate is increased, the HA chains align in the direction of flow, resulting in a decreased solution viscosity. This shear thinning effect can be seen when delivering HA through a syringe.

HA is a highly hydrophilic polymer. Each glucuronic acid unit contains one carboxyl group, giving rise to HA's polyanionic character at physiological pH. In the presence of water, HA molecules can expand in volume up to 1000 times and form loose hydrated matrices.^[4] Because of this property, HA is thought to play several roles in the ECM, including space filler, lubricant, and osmotic buffer.^[4] By forming a hydrated polymer network, HA can act as a sieve, restricting the movement of pathogens, plasma proteins, and proteases.^[1,4] In addition, HA's polyionic structure is able to scavenge free radicals, which mediates inflammation by



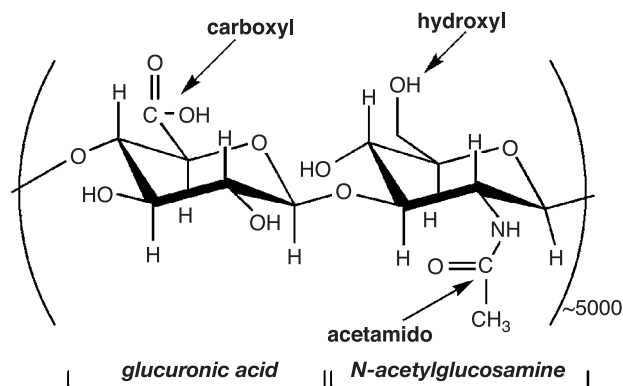


Fig. 1 The structure of native HA. HA is a naturally derived polymer composed of disaccharide repeats of glucuronic acid and N-acetylglucosamine. The molecular weight of native HA is typically several million. Each disaccharide repeat of HA contains three possible modification sites: the hydroxyl, carboxyl, and acetamido groups.

imparting an antioxidant effect.^[1] For a thorough review of HA's structural and physicochemical properties, see Ref. [2].

Roles in the Extracellular Matrix and Interactions with Hyaladherins

HA plays several important organizational roles in the ECM by binding with cells and other components through specific and nonspecific interactions. For example, HA specifically interacts with several proteoglycans (e.g., aggrecan, versican) through HA-binding motifs,^[2] and modulates the organization of fibrin,

fibronectin, and collagen networks.^[5] In addition to contributing to the homeostasis of ECM structure and hydration, HA is also involved in a number of more complex signaling events relating to cell migration, attachment, and metastasis. Such processes are mediated through a group of proteins called hyaladherins. These proteins can be categorized into three types: those that are soluble, those that bind HA to other ECM molecules, and those that act as cellular receptors for HA.

Two hyaladherins, CD44 and RHAMM (Receptor for HA-Mediated Mobility) have been extensively studied. CD44 has been implicated as the major cell surface receptor for HA, and is expressed on a variety of cell types (e.g., leukocytes, fibroblasts, epithelial cells, keratinocytes, and some endothelial cells).^[6] The CD44 receptor is associated with various cellular processes, including adhesion, migration, proliferation, and activation, as well as HA degradation and uptake.^[1,4] (See Fig. 2 for a schematic of the signaling mechanisms associated with CD44.) In a number of cell types, the RHAMM hyaladherin has been found on the cell surface, as well as in the cytosol and nucleus.^[7] RHAMM has been implicated in regulating cellular responses to growth factors and plays a role in cell migration, particularly for fibroblasts and smooth muscle cells.^[7,8] For a comprehensive review of hyaladherins, see Ref. [8].

HA's Complex Influence on Cell Migration

As mentioned previously, HA may influence cellular behavior not only through direct mechanisms (e.g., specific binding with cellular hyaladherins), but also by indirect means (through altering the physical properties

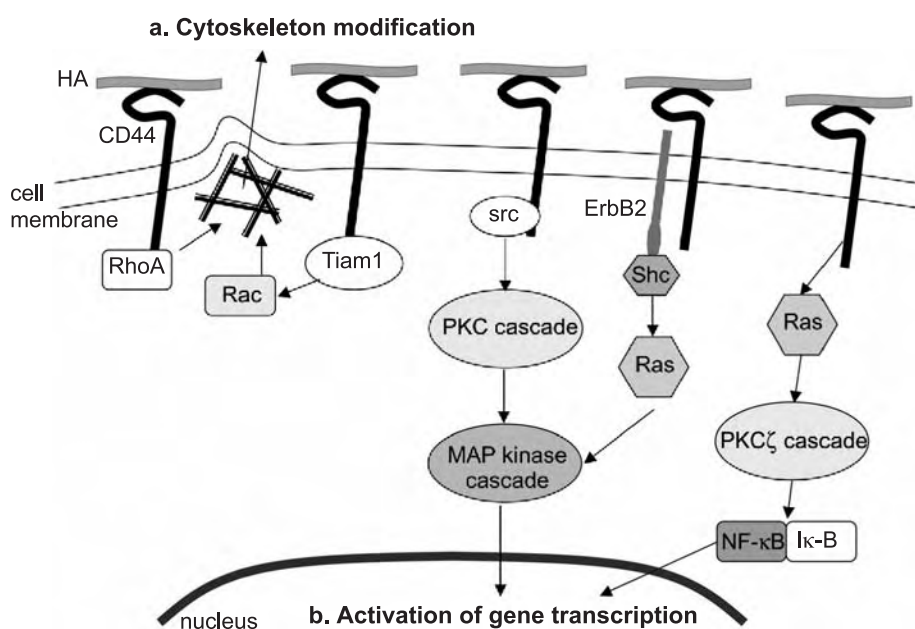


Fig. 2 Signaling pathways of the HA receptor CD44. The binding of HA and CD44 can result in signaling cascades that affect a) the modification of cytoskeletal proteins, leading to changes in cell migration, and b) the activation of gene transcription, resulting in changes in cell proliferation. Cytoskeletal assembly may be modified through signaling involving Rac and RhoA. The protein kinase C (PKC) and MAP kinase cascades may be involved in altering gene expression inside the nucleus. (Adapted from Ref. [3].)

of the ECM). HA's effect on cell migration is a remarkable illustration of this complexity. For instance, the binding of cellular hyaladherins to HA is involved in migration for a variety of cell types.^[5] Yet, it is evident that HA may also indirectly aid cell migration by contributing to more open and hydrated spaces in the ECM or by otherwise remodeling the environment through interactions with collagen and fibrin.^[3] Furthermore, as the main component of pericellular coats (matrices that are formed around migrating and proliferating cells), HA nonspecifically facilitates the detachment of cells from the ECM.^[8] Given such complexity, distinguishing between the effects of HA's biological activity and physicochemical properties is not straightforward.

HA Degradation

In mammals, the enzymatic degradation of HA results from the action of three types of enzymes: hyaluronidase (hyase), β -D-glucuronidase, and β -N-acetyl-hexosaminidase. Throughout the body, these enzymes are found in various forms, intracellularly and in serum. In general, hyase cleaves high molecular weight HA into smaller oligosaccharides while β -D-glucuronidase and β -N-acetyl-hexosaminidase further degrade the oligosaccharide fragments by removing nonreducing terminal sugars. Besides the degradation reaction, testicular hyase can catalyze the reverse reaction, transglycosylation. Therefore, HA is not simply degraded by this enzyme into disaccharide fragments. On the contrary, incubation of HA in a testicular hyase solution yields a mixture of fragment sizes, primarily made up of tetrasaccharides with smaller amounts of hexa-, octa-, and disaccharides.^[9] The cell types primarily responsible for hyase synthesis are macrophages, fibroblasts, and endothelial cells; consequently, each of these cell types have been associated with HA degradation in the body.^[10] In addition to enzymatic degradation, HA can also be degraded by reactive oxygen intermediates, a mechanism that has been implicated as a source of HA fragments at sites of inflammation.^[11] For a review of other nonenzymatic means of HA degradation (e.g., degradation induced by free radicals, ultrasound, pH, and temperature treatments), see Ref. [2].

HA in Wound Repair and Angiogenesis

High concentrations of HA can be found in many tissues undergoing remodeling, regeneration, and morphogenesis.^[8] Furthermore, in the early stages of wound healing, a matrix composed largely of fibrin and HA is formed, which aids in fibroblast and endothelial cell migration into the injury site. This matrix also supports granulation tissue formation, particularly in fetal tissue,

where HA has been implicated in aiding scar-free wound healing.^[1] For this reason, exogenous HA has been investigated in wound healing applications and has been found to be effective even in challenging conditions such as chronic wounds.^[1]

The exact role of HA in wound healing is very complicated. In a comprehensive review of HA and wound repair, Chen and Abatangelo suggest that HA plays a multifaceted role in each wound healing stage (i.e., inflammation, granulation tissue formation, re-epithelization, and remodeling).^[1] For example, HA aids in setting off the early stages of inflammation and initiating the wound healing response, and also moderates the later stages of this potentially damaging process. Furthermore, the granulation tissue that subsequently forms remains rich in HA, aiding cell migration, proliferation, and organization of the ECM.

Perhaps one of the most interesting roles of HA in wound healing is the ability of HA oligosaccharides to promote angiogenesis.^[10] Degradation products of 4–20 disaccharide units have been shown to stimulate in-vivo capillary growth and induce in-vitro endothelial proliferation, migration, and tube formation. The complex nature of HA, however, is again demonstrated: Although low molecular weight HA promotes angiogenesis, high molecular weight HA has been shown to inhibit new blood vessel growth. Researchers have attempted to better discern the direct and indirect effects of HA molecular weight on angiogenesis, but an exact understanding of this process remains to be elucidated. For a recent review of the role of HA in angiogenesis, see Ref. [10].

The Role of HA in Pathological Processes

Factors such as HA concentration and molecular weight distribution have been correlated with the development of certain pathological conditions and can therefore be used as diagnostic markers for disease. Increased serum concentrations of HA can be associated with the progression of certain inflammatory processes, such as rheumatoid arthritis, and have been used to evaluate the success of kidney and liver transplants.^[4] Changes in the properties of HA have also been associated with certain vascular diseases, including atherosclerosis and restenosis.^[12] Furthermore, high concentrations of HA have been linked to tumor development and in some cases can be used as a negative predictor of cancer patient survival.^[12] Presumably, one of the roles of HA in cardiovascular disease and tumor growth is to increase the free volume and fluidity of the ECM, allowing increased cell proliferation and migration. However, a detailed understanding of the role of HA in these conditions is far from complete.^[12]

DERIVATIZED AND CROSS-LINKED HA BIOMATERIALS

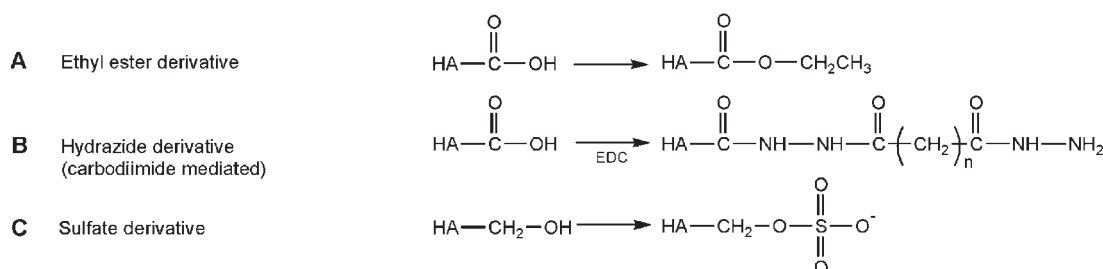
HA offers many unique advantages as a building block for biomedical devices. Specifically, HA is enzymatically degradable, naturally derived, and nonimmunogenic. Unlike proteins, which are easily denatured, HA is a carbohydrate polymer and can withstand demanding chemical modification procedures without losing its biological activity. Purified HA has found clinical uses in ophthalmology, osteoarthritis, and wound-healing applications.^[2] HA has many other potential applications, particularly as a biomaterial to deliver cells or proteins, but its use is limited because of its high water solubility and rapid degradation in the body. Thus, researchers have investigated methods to prolong the residence time of HA, using chemical modifications to yield materials with greater in-vivo stability. Two main approaches for

creating HA biomaterials have been utilized: derivatization and cross-linking (summarized in Fig. 3). Both approaches are used to modify chemically one or more of HA's three available reactive groups (hydroxyl, carboxyl, and acetamido; Fig. 1), while attempting to maintain biocompatibility and biological activity. This section provides an overview of several HA derivatization and cross-linking methods; however, this discussion is not comprehensive. For further information on modification strategies, including HA composites and blends, see Refs. [2] and [13].

HA Derivatives

Many studies focus on three main types of HA derivatization strategies: esterification, carbodiimide-mediated modification, and sulfation. These simple

HA Derivatives



Crosslinked HA

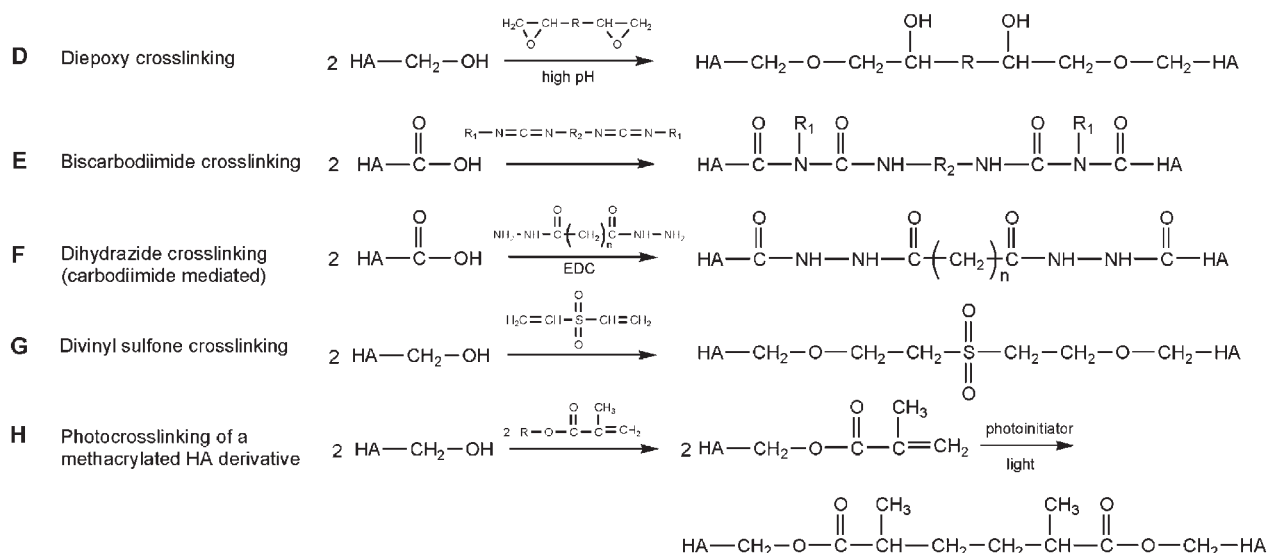


Fig. 3 Derivatized and cross-linked HA biomaterials. To tune HA's physicochemical properties for specific applications, HA has been modified in several ways. Uncross-linked HA derivatives, including modifications based on A) esterification,^[14] B) carbodiimide-mediated reactions,^[15] and C) sulfation^[18,19] have been created. Cross-linked HA has also been produced using several cross-linking agents, including D) diepoxies,^[20,21] E) biscarbodiimides,^[17] F) bifunctional amine cross-linkers,^[16] and G) divinyl sulfone,^[25,28] as well as H) photocross-linking methods. (From Refs. [31,32].)

modification strategies can be used to create HA derivatives that are suitable for direct application or for subsequent cross-linking procedures. This section will describe uncross-linked HA derivatives and the next section (“Cross-linked HA Biomaterials”) will include a discussion of cross-linking methods for these materials.

Ester derivatives

To create more hydrophobic forms of HA that have added rigidity and are less susceptible to enzymatic degradation, researchers have esterified HA's carboxyl groups.^[14] The modification is carried out through an alkylation step of HA with an alkyl halide, yielding derivatives with 0–100% modifications of the available carboxyl groups. As the percent of HA esterification increases, these materials become more rigid and more hydrophobic. An increased hydrophobic nature means that the HA derivatives are less soluble in water and less susceptible to enzymatic degradation. Fidia Advanced Biopolymers have created a range of HA esters, including ethyl esters (Hyaff-7; Fig. 3A) and benzyl esters of HA (Hyaff-11). Through a variety of processing steps, these HA esters can be shaped into fibers, membranes, sponges, and microspheres.

Carbodiimide-mediated derivatives

Reactions mediated by carbodiimides (e.g., 1-ethyl-3-(3'-dimethylaminopropyl)carbodiimide or EDC) are a common means of covalently binding the carboxyl group of one bioactive molecule with an amine of another (e.g., reactions between amino acids during peptide synthesis). However, carbodiimide reactions are very sensitive to pH and often result in the formation of an unreactive intermediate, acylurea, from carboxyl groups. Therefore, depending on the reaction conditions, carbodiimide-mediated reactions on HA can result in very low coupling yields. (For excellent discussions of the challenges to carbodiimide-mediated modifications of HA, see Refs. [15–17]). Pouyani and Prestwich gained a detailed understanding of these reactions, and as a result, chose hydrazide derivatization (Fig. 3B) as a means of increasing the utility of carbodiimide in HA modifications.^[15] Prestwich and coworkers have continued to develop a wide variety of HA-hydrazide derivatives for use as drug delivery devices, cellular probes, and HA-protein conjugates.^[13]

Sulfated derivatives

To create a blood-compatible molecule that mimics heparin, researchers have modified HA with sulfate

groups (Fig. 3C; Ref. [18]). The sulfation is completed through a reaction of sulfur trioxide pyridine with HA's hydroxyl groups, yielding a range of sulfation from one to four sulfur groups incorporated per HA disaccharide.^[19] This sulfation modification creates HA derivatives that are structurally and chemically similar to heparin;^[18,19] cardiovascular applications of these materials are described in more detail in a later section (see “Applications of HA Biomaterials”).

Cross-Linked HA Biomaterials

As already mentioned, HA presents many inherent advantages as a foundation for biomaterials. In addition to derivatization, cross-linking is another means of engineering HA's physicochemical properties. Depending on the cross-linking molecule and reaction chemistry, a wide variety of HA materials can be created, ranging from films with relatively low water content to highly swelling hydrogels. Most HA cross-linking methods fall into either of two general schemes: a one-step procedure consisting of the exposure of HA to a cross-linker, or a two-step procedure in which a highly reactive HA derivative is first synthesized and then cross-linked in a subsequent reaction. Following is a brief listing of common HA cross-linking techniques; see Refs. [2] and [13] for more detail.

Diepoxy cross-linking

Almost 40 years ago, Laurent and coworkers created diepoxy-cross-linked HA hydrogels using polyethylene glycol diglycidyl ether.^[20] Others have extended the utility of the diepoxy chemistry to other cross-linkers, including ethylene glycol diglycidylether (a bifunctional cross-linker) and polyglycerol polyglycidylether (a trifunctional cross-linker).^[21] Interestingly, detailed studies of this chemistry have shown that at low pH, diepoxy compounds form ester linkages between carboxyl groups, while at high pH, they form ether linkages between hydroxyl groups (Fig. 3D; Ref. [22]). Utilizing this fact, diepoxyoctane has been used to cross-link HA in a two-step double-cross-linking treatment: HA is reacted with diepoxyoctane at low pH, forming ester cross-links between HA chains, and then a second diepoxyoctane reaction is subsequently carried out at high pH, forming ether cross-links.^[22]

Carbodiimide-mediated cross-linking

As discussed above, carbodiimide-mediated reactions are commonly used for reactions between carboxylic acids and amines. These reactions, however, are highly sensitive to pH and readily form an unreactive acylurea



moiety when applied to HA reactions; therefore, the applications for which carbodiimides can be used with HA are somewhat limited. Fortunately, several researchers have found specific conditions under which carbodiimide-mediated reactions can allow HA cross-linking.

The simplest of these methods uses solely carbodiimide to induce inter- and intramolecular cross-links on HA;^[23] in other words, these materials can be synthesized without the use of exogenous cross-linkers. To form cross-links, the carbodiimide-mediated reaction is performed on films of at least 70 weight-percent HA in acetone- or ethanol-water solutions. Similar reactions carried out with HA in solution did not yield cross-linked materials, perhaps due to the formation of unreactive acylureas. Challenged by this tendency of carbodiimide-mediated reactions to form unreactive acylureas on HA, Kuo, et al. created a biscarbodiimide cross-linking technique (Fig. 3E; Ref. [17]). Several biscarbodiimides were synthesized and covalently bound to HA through a mechanism similar to the carbodiimide-mediated creation of an acylurea, producing aromatic or aliphatic cross-links between HA molecules.

Finally, several efforts have utilized carbodiimides to mediate reactions between HA's carboxyl groups and the amines of bifunctional cross-linkers (Fig. 3F). Bullpitt and Aeschlimann made use of reactions mediated with carbodiimide and 1-hydroxybenzotriazole (HOBt) to couple activated amines ($pK_a < 8.0$; e.g., adipic dihydrazide) to HA's carboxyl groups; carbodiimide and N-hydroxysulfosuccinimide (sulfo-NHS) were used for coupling reactions between HA and simple primary amines ($pK_a > 9.0$; e.g., 1,4-diaminobutane dihydrochloride).^[16] Similarly, Vercryse and coworkers have synthesized a wide variety of polyvalent hydrazide cross-linkers (2-6 hydrazides per cross-linker) for use in carbodiimide-mediated cross-linking reactions.^[24]

Aldehyde cross-linking

Formaldehyde and glutaraldehyde have long been used to cross-link proteins for tissue preservation. Formaldehyde is used to create Biomatrix's Hylan-A, a cross-linked form of HA that is water-soluble, but more viscous and elastic than native HA.^[25] Glutaraldehyde has been used to cross-link HA to yield materials with high resistance toward degradation. Tomihata and Ikade used glutaraldehyde in acidic acetone-water solutions to create cross-linked HA films of low water content.^[26] Hu et al. compared glutaraldehyde and carbodiimide-mediated reactions and found that glutaraldehyde yielded more highly

cross-linked materials suitable for use as three-dimensional tissue scaffolds.^[27]

Divinyl sulfone cross-linking

Insoluble HA hydrogels can be formed by cross-linking HA with divinyl sulfone (Fig. 3G), yielding materials such as Biomatrix's Hylan-B gel.^[25,28] At high pH, divinyl sulfone creates sulphonyl-bis-ethyl crosslinks between HA's hydroxyl groups. Depending on the reaction conditions, materials ranging from soft gels to firm solids (e.g., membranes and tubes) can be formed. Compared to native HA, Hylan-B has an extended residence time in vivo, particularly in sites where low mechanical force is imposed on the implant.

Photocross-linking

While the cross-linking methods discussed above yield a wide variety of stabilized materials, most of these techniques proceed under physiologically incompatible conditions. Furthermore, many of these methods cross-link immediately upon mixing of the HA with the cross-linker. Because HA forms highly viscous solutions, it is often difficult to obtain homogenous materials in this manner. Photocross-linking, however, is a method that reacts only upon exposure to the appropriate wavelength of light; therefore, the reactants can be thoroughly mixed before cross-linking is initiated.^[29] Furthermore, photocross-linking methods are capable of proceeding under physiological conditions without detrimental effects to copolymerized bioactive molecules or encapsulated cells. To this end, HA has been modified with several photocross-linkable groups, including cinnamoyl, coumarin, and thymine;^[30] methacrylic anhydride;^[31] glycidyl methacrylate;^[32] and styrene.^[33] (See Fig. 3H for a general schematic of photocross-linking methacrylated HA derivatives.) Additionally, sulfated HA has been modified with 4-azidoaniline, yielding an HA derivative suitable for patterning features as small as 100 μm ,^[34] which may be useful for directing cell growth or adhesion in new biomaterials.

APPLICATIONS OF HA BIOMATERIALS

HA's intrinsic physicochemical and biological properties suggest that this distinctive molecule is suitable for application in clinical therapies, diagnostics, tissue engineering scaffolds, and drug delivery devices. Unmodified, derivatized, and cross-linked HA have proven beneficial in a variety of applications, which are briefly described in the following sections. For comprehensive reviews of medical applications of HA, see

Table 1 FDA-approved HA products

Application	Tradename	Approval type	Company
Osteoarthritis	Hyalgan	Premarket	Fidia
	Synvisc	Premarket	Biomatrix/Genzyme
	Supartz	Premarket	Seikagaku
	Hylashield	510 k	Biomatrix
Ophthalmology	Healon	Premarket	Pharmacia
	Amvisc	Premarket	Bausch and Lomb
	Coase, Shellgel, Staarvisc	Premarket	Anika Therapeutics
	Amo Vitrax, Vitrax	Premarket	Allergan/Medtronic
	Provisc, Viscoat	Premarket	Alcon Laboratories
Wound healing	Bionect	510 k	Fidia
	Ialuset	510 k	IBSA
	Adcon	Premarket	Gliatech
Postsurgical adhesions	Intergel	Premarket	LifeCore Biomedical
	Seprafilm	Premarket	Genzyme
	Hyalomatrix	510 k	Fidia
Surgical scaffolding	Hylasine	510 k	Biomatrix
	Deflux	Premarket	Q-Med

(Adapted from Refs. [2,13] and [35].)

Refs. [2] and [13]. Table 1 provides a list of HA products currently approved by the FDA for clinical use.

Orthopedic Applications

HA plays a vital role in the development of cartilage, the maintenance of the synovial fluid, and the regeneration of tendons.^[8] High concentrations of HA can be found in the ECM of all adult joint tissues, including the synovial fluid and the outer layer of cartilage.^[36] In part because of its viscoelastic nature and ability to form highly hydrated matrices, HA acts in the joint as a lubricant and a shock absorber.^[37] In diseases such as osteoarthritis, the concentration and molecular weight of the HA naturally present in the joint are decreased, contributing to stiffness and pain. Viscosupplementation treatments aim to treat these conditions, and to this end, a variety of HA materials have been successfully applied as clinical therapies.^[36,37]

Researchers have also investigated cross-linked HA as cell delivery scaffolds for cartilage and bone tissue engineering. Esterified HA (Hyaff, Fidia Advanced Biopolymers) supports the growth of chondrocytes^[38] and the differentiation of bone marrow-derived mesenchymal progenitor cells into chondrocytes and osteoblasts.^[39] In fact, when compared to a well characterized delivery vehicle (porous calcium phosphate ceramic), Hyaff materials allowed greater amounts of bone and cartilage to be formed *in vivo* (as

determined from the image analysis of stained tissue sections).^[39]

Antiadhesion Applications

As HA is highly hydrophilic, it is a polymer that is well suited for applications requiring minimal cellular adhesion. Postoperative adhesions, which form between adjacent tissue layers following surgery, impede wound healing and often require additional surgical procedures to repair successfully. Barriers made from cross-linked HA have been effectively used to prevent such adhesions from forming.^[30] Furthermore, the adhesion of bacteria onto biomaterials can induce infections and great risk to the patient; with this in mind, esterified HA has also been used to prevent bacterial adhesion to dental implants, intraocular lenses, and catheters.^[40]

Cardiovascular Applications

In a manner related to its antiadhesive properties, HA has also proven to be effective for increasing the blood compatibilities of cardiovascular implants such as vascular grafts and stents. For example, biomaterial surfaces treated with cross-linked HA have been associated with reduced platelet adhesion and thrombus formation.^[25,34,41] Furthermore, sulfated HA derivatives



can act as heparin mimics,^[18,19] in fact, HA derivatives with higher degrees of sulfation are associated with increased abilities to prevent blood coagulation (as measured by longer times required for whole blood clotting).^[19]

Cross-linked HA is also a promising biomaterial for cardiac tissue engineering. In the embryo, HA is required for the normal development of the valves and chambers within the heart.^[42] Building upon this information, studies by Masters and Anseth have indicated that photopolymerized HA hydrogels are suitable materials for constructing tissue-engineered heart valves (K.S. Masters and K.S. Anseth, Department of Chemical Engineering, The University of Colorado at Boulder, 2003, personal communication). Valvular interstitial cells, which are the primary cell type found in heart valves, were found to adhere to and proliferate upon the HA hydrogels. Furthermore, the hyaladherins CD44 and RHAMM are present on the valvular interstitial cells, and these receptors were thought to be involved in promoting cellular proliferation in response to degraded HA hydrogel fragments.

Ophthalmology

HA, a natural component of the vitreous humor of the eye, has found many successful applications in ophthalmologic surgery.^[2] In part because it forms viscoelastic and highly hydrated matrices, HA is particularly useful as a space-filling matrix in the eye; thus, intraocular injection of HA during surgery is used to maintain the shape of the anterior chamber. Furthermore, HA solutions also serve as a viscosity-enhancing component of eye drops and an adjuvant to eye tissue repair.

Dermatology and Wound-Healing Applications

HA is naturally present in high concentrations in the skin and soft connective tissues. Therefore, HA is an appropriate choice for a matrix to support dermal regeneration and augmentation. For example, Prestwich and coworkers found that cross-linked HA hydrogel films accelerate the healing of full-thickness wounds (Fig. 4), presumably by providing a highly hydrated and nonimmunogenic environment that is conducive to tissue repair.^[43] Hyaff scaffolds cultured *in vitro* with keratinocytes and fibroblasts have been used to create materials similar to skin, including two distinct epidermal and dermal-like tissue layers.^[38] Moreover, as a result of its ability to form hydrated, expanded matrices, HA has also been successfully used in cosmetic applications such as soft tissue augmentation.^[44]

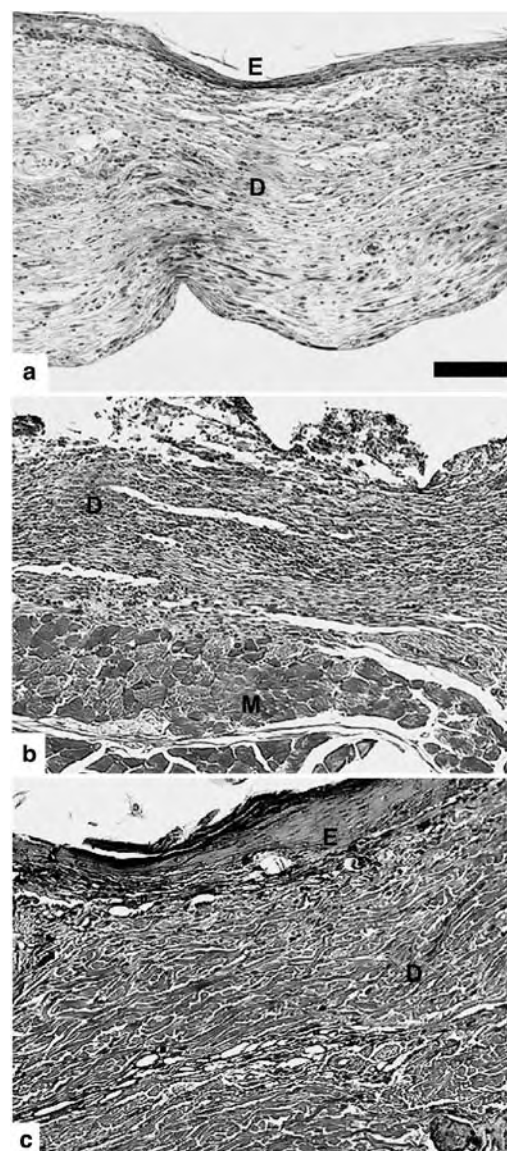


Fig. 4 Cross-linked HA hydrogel films to assist wound healing. Full-thickness wounds were created in the skin of Balb/c mice and a) treated with a cross-linked HA hydrogel film (adipic dihydrazide-modified HA that was cross-linked with polyethylene glycol dialdehyde) and covered with a wound dressing (Tegaderm, a thin polyurethane membrane dressing that is impermeable to microorganisms, 3M Health Care, St. Paul, MN), or b) covered with the wound dressing alone. After 10 days, the HA films were completely degraded and four out of six of the HA-treated wounds were associated with the formation of an epithelial layer and a collagen-rich dermal layer. The tissues treated with the wound dressing alone were less consistent and had reduced reepithelialization. Tissue extracted from an area peripheral to the wound is shown in c) as a normal tissue control. The sections are stained with Masson's Trichrome. D = dermis, E = epidermis, M = muscle. Scale bar = 100 μ m. Images reprinted from Ref. [43] with permission from Elsevier Science.

Neural and Glial Applications

HA treatments have also been associated with improved peripheral nerve regeneration. Injections of unmodified HA into a nerve guide have been associated with increased levels of physiological and functional regeneration.^[45] Furthermore, when strands of cross-linked HA are coated with polylysine, Schwann cells are able to attach and proliferate on these materials.^[27] It is possible that the presence of HA enables increased glial cell migration, more rapid neurite outgrowth or a more highly organized ECM between the nerve stumps. However, the exact mechanism by which HA may aid this process is not well understood.

General Tissue-Engineering and Surgical Applications

HA's unique biological properties have inspired the development of a variety of novel therapies for general surgical applications. For example, work in our lab has created photocross-linked scaffolds to support potentially a broad range of tissue engineering applications.^[32] Following cross-linking, these hydrogels maintained HA's biological activity and promoted similar levels of vascularization as fibrin positive controls. We have also used HA as a dopant during the synthesis of polypyrrole, an electrically conductive polymer that alone can promote repair in a variety of systems.^[46] When compared to polypyrrole films without HA, composite films of HA and polypyrrole were associated with higher levels of vascularization in subcutaneous implants. Therefore, HA-polypyrrole materials could be suitable for tissue-engineering applications that may benefit from electrical stimulation and enhanced angiogenesis. Other researchers have also utilized HA's antioxidant properties to assist in the prevention of inflammation. In-vitro studies have indicated that several materials, including unmodified HA, Hyaff, and a steroid-ester of HA, possess antioxidant activities.^[47,48] Such materials could aid in healing conditions such as chronic wounds and rheumatoid arthritis.

Drug Delivery

Because of its biocompatibility, biodegradability, and readily modified chemical structure, HA has been extensively investigated in drug-delivery applications. A variety of commercially available preparations of HA derivatives and cross-linked HA materials have been developed for drug delivery; these materials are created in forms such as films, microspheres, liposomes, fibers, and hydrogels. For excellent reviews on applications of HA in drug delivery, see Refs. [2] and [13].

CONCLUSIONS AND FUTURE DIRECTIONS

Through multidisciplinary discoveries about the structure, properties, biological activity, and chemical modification of this unique polymer, HA has found success in an extraordinarily broad range of biomedical applications. Future clinical therapies of HA-derived materials critically rely on a more detailed understanding of the effects of HA molecular weight and concentration and how this biomolecule specifically interacts with cells and ECM components in the body. Future directions of these materials will require finely tuned and controllable interactions between HA and its environment. Work in these areas is underway; for example, adhesive peptide sequences have been covalently bound to HA materials.^[49] Also, environmentally responsive materials have been synthesized from HA. These materials can be created to swell or degrade in response to inflammation,^[21] electrical stimulation,^[50] and heat.^[51] For an excellent online resource regarding the biology, modification chemistries, and various applications of HA, see the Glycoforum website.^[52]

H

ARTICLES OF FURTHER INTEREST

Burn Dressing; Cartilage Regeneration; Extracellular Matrix Scaffolds; Orthopedic Biomaterials; Tissue Engineering of Heart Valves; Tissue Engineering Scaffolds; Tissue Engineering of Bone Marrow

REFERENCES

1. Chen, W.Y.; Abatangelo, G. Functions of hyaluronan in wound repair. *Wound Repair Regen.* **1999**, *7* (2), 79–89.
2. Lapcik, L., Jr.; Lapcik, L.; De Smedt, S.; Demeester, J.; Chabreck, P. Hyaluronan: Preparation, structure, properties, and applications. *Chem. Rev.* **1998**, *98* (8), 2663–2684.
3. Lee, J.Y.; Spicer, A.P. Hyaluronan: A multifunctional, megaDalton, stealth molecule. *Curr. Opin. Cell Biol.* **2000**, *12* (5), 581–586.
4. Laurent, T.C.; Fraser, J.R. Hyaluronan. *FASEB J.* **1992**, *6* (7), 2397–2404.
5. Nehls, V.; Hayen, W. Are hyaluronan receptors involved in three-dimensional cell migration? *Histol. Histopathol.* **2000**, *15* (2), 629–636.
6. Isacke, C.M.; Yarwood, H. The hyaluronan receptor, CD44. *Int. J. Biochem. Cell Biol.* **2002**, *34* (7), 718–721.
7. Cheung, W.F.; Cruz, T.F.; Turley, E.A. Receptor for hyaluronan-mediated motility (RHAMM), a hyaladherin that regulates cell responses to growth factors. *Biochem. Soc. Trans.* **1999**, *27* (2), 135–142.

8. Toole, B.P. Hyaluronan in morphogenesis. *Semin. Cell Dev. Biol.* **2001**, *12* (2), 79–87.
9. Roden, L.; Campbell, P.; Fraser, J.R.; Laurent, T.C.; Pertoft, H.; Thompson, J.N. Enzymic pathways of hyaluronan catabolism. *Ciba Found. Symp.* **1989**, *143*, 60–76.
10. West, D.C.; Fan, T.-P.D. Hyaluronan Oligosaccharides Promote Wound Repair. In *The New Angiotherapy*; Fan, T.-P.D., Kohn, E.C., Eds.; Humana Press: Totowa, NJ, 2002; 177–188.
11. Noble, P.W. Hyaluronan and its catabolic products in tissue injury and repair. *Matrix Biol.* **2002**, *21* (1), 25–29.
12. Toole, B.P.; Wight, T.N.; Tammi, M.I. Hyaluronan–cell interactions in cancer and vascular disease. *J. Biol. Chem.* **2002**, *277* (7), 4593–4596.
13. Vercruyse, K.P.; Prestwich, G.D. Hyaluronate derivatives in drug delivery. *Crit. Rev. Ther. Drug Carr. Syst.* **1998**, *15* (5), 513–555.
14. Campoccia, D.; Doherty, P.; Radice, M.; Brun, P.; Abatangelo, G.; Williams, D.F. Semisynthetic resorbable materials from hyaluronan esterification. *Biomaterials* **1998**, *19* (23), 2101–2127.
15. Pouyani, T.; Prestwich, G.D. Functionalized derivatives of hyaluronic acid oligosaccharides: Drug carriers and novel biomaterials. *Bioconjug. Chem.* **1994**, *5* (4), 339–347.
16. Bulpitt, P.; Aeschlimann, D. New strategy for chemical modification of hyaluronic acid: Preparation of functionalized derivatives and their use in the formation of novel biocompatible hydrogels. *J. Biomed. Mater. Res.* **1999**, *47* (2), 152–169.
17. Kuo, J.-W.; Swann, D.A.; Prestwich, G.D. Chemical modification of hyaluronic acid by carbodiimides. *Bioconjug. Chem.* **1991**, *2*, 232–241.
18. Barbucci, R.; Magnani, A.; Casolaro, M.; Marchettini, N.; Rosi, C.; Bosco, M. Modification of hyaluronic acid by insertion of sulfate groups to obtain a heparin-like molecule. Part I. Characterization and behavior in aqueous solution towards H⁺ and Cu²⁺ ions. *Gazz. Chim. Ital.* **1995**, *125* (4), 169–180.
19. Magnani, A.; Albanese, A.; Lamponi, S.; Barbucci, R. Blood-interaction performance of differently sulphated hyaluronic acids. *Thromb. Res.* **1996**, *81* (3), 383–395.
20. Laurent, T.C.; Helling, K.; Gelotte, B. Crosslinked gels of hyaluronic acid. *Acta Chem. Scand.* **1964**, *18* (1), 274–275.
21. Yui, N.; Okano, T.; Sakurai, Y. Inflammation responsive degradation of crosslinked hyaluronic acid gels. *J. Control. Release* **1992**, *22* (2), 105–116.
22. Zhao, X.B.; Fraser, J.E.; Alexander, C.; Lockett, C.; White, B.J. Synthesis and characterization of a novel double crosslinked hyaluronan hydrogel. *J. Mater. Sci., Mater. Med.* **2002**, *13*, 11–16.
23. Tomihata, K.; Ikada, Y. Crosslinking of hyaluronic acid with water-soluble carbodiimide. *J. Biomed. Mater. Res.* **1997**, *37* (2), 243–251.
24. Vercruyse, K.P.; Marecak, D.M.; Marecek, J.F.; Prestwich, G.D. Synthesis and in vitro degradation of new polyvalent hydrazide cross-linked hydrogels of hyaluronic acid. *Bioconjug. Chem.* **1997**, *8* (5), 686–694.
25. Balazs, E.A.; Bland, P.A.; Denlinger, J.L.; Goldman, A.I.; Larsen, N.E.; Leshchiner, E.A.; Leshchiner, A.; Morales, B. Matrix engineering. *Blood Coagul. Fibrinolysis* **1991**, *2* (1), 173–178.
26. Tomihata, K.; Ikada, Y. Crosslinking of hyaluronic acid with glutaraldehyde. *J. Polym. Sci., A, Polym. Chem.* **1997**, *35*, 3553–3559.
27. Hu, M.; Sabelman, E.E.; Tsai, C.; Tan, J.; Hentz, V.R. Improvement of Schwann cell attachment and proliferation on modified hyaluronic acid strands by polylysine. *Tissue Eng.* **2000**, *6* (6), 585–593.
28. Band, P.A. Hyaluronan Derivatives: Chemistry and Clinical Applications. In *Chemistry, Biology and Medical Applications of Hyaluronan and Its Derivatives*; Laurent, T.C., Ed.; Portland Press: London, 1998; 33–42.
29. Anseth, K.S.; Burdick, J.A. New directions in photopolymerizable biomaterials. *Mater. Res. Soc. Bull.* **2002**, *27* (2), 130–136.
30. Matsuda, T.; Moghaddam, M.J.; Miwa, H.; Sakurai, K.; Iida, F. Photoinduced prevention of tissue adhesion. *ASAIO J.* **1992**, *38* (3), M154–M157.
31. Smeds, K.A.; Pfister-Serres, A.; Hatchell, D.L.; Grinstaff, M.W. Synthesis of a novel polysaccharide hydrogel. *J. Macromol. Sci., Part A, Pure Appl. Chem.* **1999**, *A36* (7 and 8), 981–989.
32. Leach, J.B.; Bivens, K.A.; Patrick, C.W., Jr.; Schmidt, C.E. Photocrosslinked hyaluronic acid hydrogels: Natural, biodegradable tissue engineering scaffolds. *Biotechnol. Bioeng.* **2003**, *82*, 578–589.
33. Matsuda, T.; Magoshi, T. Preparation of vinylated polysaccharides and photofabrication of tubular scaffolds as potential use in tissue engineering. *Biomacromolecules* **2002**, *3* (5), 942–950.
34. Chen, G.; Ito, Y.; Imanishi, Y.; Magnani, A.; Lamponi, S.; Barbucci, R. Photoimmobilization of sulfated hyaluronic acid for antithrombogenicity. *Bioconjug. Chem.* **1997**, *8* (5), 730–734.
35. <http://www.fda.gov/cdrh>. (accessed December 2003).
36. Pozo, M.A.; Balazs, E.A.; Belmonte, C. Reduction of sensory responses to passive movements of inflamed knee joints by hylan, a hyaluronan derivative. *Exp. Brain Res.* **1997**, *116* (1), 3–9.
37. Peyron, J.G. A new approach to the treatment of osteoarthritis: Viscosupplementation. *Osteoarthr. Cartil.* **1993**, *1* (2), 85–87.
38. Brun, P.; Abatangelo, G.; Radice, M.; Zacchi, V.; Guidolin, D.; Daga Gordini, D.; Cortivo, R. Chondrocyte aggregation and reorganization into three-dimensional scaffolds. *J. Biomed. Mater. Res.* **1999**, *46* (3), 337–346.
39. Solchaga, L.A.; Dennis, J.E.; Goldberg, V.M.; Caplan, A.I. Hyaluronic acid-based polymers as cell carriers for tissue-engineering repair of bone and cartilage. *J. Orthop. Res.* **1999**, *17* (2), 205–213.
40. Pavesio, A.; Renier, D.; Cassinelli, C.; Morra, M. Anti-adhesive surfaces through hyaluronan coatings. *Med. Device Technol.* **1997**, *8* (7), 20–21, 24–27.

41. Kito, H.; Matsuda, T. Biocompatible coatings for luminal and outer surfaces of small-caliber artificial grafts. *J. Biomed. Mater. Res.* **1996**, *30* (3), 321–330.
42. Camenisch, T.D.; Schroeder, J.A.; Bradley, J.; Klewer, S.E.; McDonald, J.A. Heart-valve mesenchyme formation is dependent on hyaluronan-augmented activation of ErbB2-ErbB3 receptors. *Nat. Med.* **2002**, *8* (8), 850–855.
43. Kirker, K.R.; Luo, Y.; Nielson, J.H.; Shelby, J.; Prestwich, G.D. Glycosaminoglycan hydrogel films as bio-interactive dressings for wound healing. *Biomaterials* **2002**, *23* (17), 3661–3671.
44. Duranti, F.; Salti, G.; Bovani, B.; Calandra, M.; Rosati, M.L. Injectable hyaluronic acid gel for soft tissue augmentation. A clinical and histological study. *Dermatol. Surg.* **1998**, *24* (12), 1317–1325.
45. Seckel, B.R.; Jones, D.; Hekimian, K.J.; Wang, K.K.; Chakalis, D.P.; Costas, P.D. Hyaluronic acid through a new injectable nerve guide delivery system enhances peripheral nerve regeneration in the rat. *J. Neurosci. Res.* **1995**, *40* (3), 318–324.
46. Collier, J.H.; Camp, J.P.; Hudson, T.W.; Schmidt, C.E. Synthesis and characterization of polypyrrole-hyaluronic acid composite biomaterials for tissue engineering applications. *J. Biomed. Mater. Res.* **2000**, *50* (4), 574–584.
47. Cortivo, R.; Brun, P.; Cardarelli, L.; O'Regan, M.; Radice, M.; Abatangelo, G. Antioxidant effects of hyaluronan and its alpha-methyl-prednisolone derivative in chondrocyte and cartilage cultures. *Semin. Arthritis Rheum.* **1996**, *26* (1), 492–501.
48. Moseley, R.; Leaver, M.; Walker, M.; Waddington, R.J.; Parsons, D.; Chen, W.Y.; Embery, G. Comparison of the antioxidant properties of HYAFF-11p75, AQUACEL and hyaluronan towards reactive oxygen species in vitro. *Biomaterials* **2002**, *23* (10), 2255–2264.
49. Glass, J.R.; Dickerson, K.T.; Stecker, K.; Polarek, J.W. Characterization of a hyaluronic acid-Arg-Gly-Asp peptide cell attachment matrix. *Biomaterials* **1996**, *17* (11), 1101–1108.
50. Tomer, R.; Dimitrijevic, D.; Florence, A.T. Electrically controlled release of macromolecules from cross-linked hyaluronic acid hydrogels. *J. Control. Res.* **1995**, *33*, 405–413.
51. Ohya, S.; Nakayama, Y.; Matsuda, T. Thermo-responsive artificial extracellular matrix for tissue engineering: hyaluronic acid bioconjugated with poly (N-isopropylacrylamide) grafts. *Biomacromolecules* **2001**, *2* (3), 856–863.
52. <http://www.glycoforum.gr.jp/science/hyaluronan/hyaluronanE.html>. (accessed December 2002).

Hybrid Vascular Prostheses

Jerome A. Werkmeister

CSIRO, Molecular and Health Sciences, Clayton South, Victoria, Australia

Glenn A. Edwards

Department of Veterinary Science, University of Melbourne, Werribee, Victoria, Australia

John A. M. Ramshaw

CSIRO, Molecular and Health Sciences, Clayton South, Victoria, Australia

INTRODUCTION

Synthetic and biologically based hybrid prostheses are reasonably successful when used as large-diameter peripheral vascular replacements (>5–6 mm), although outcomes do depend on the location of the replacement vessel and the status of the patient. However, treatment of both peripheral vascular disease (PVD) and coronary artery disease, the largest cause of mortality in the United States and Europe, is severely hampered by the inadequacy of the current synthetic and hybrid prostheses when used in these smaller-diameter applications. The autogenous saphenous vein and internal mammary artery continue to be the most widely used vascular grafts for small-caliber vessel reconstruction, but these substitutes are often not available or are unsuitable.

The predominant synthetic materials used in PVD, including below-the-knee procedures, are woven polyethylene terephthalate (Dacron™) and expanded polytetrafluoroethylene (ePTFE), while the latter is a poor but often only choice for cardiac surgery. The relative success of the autogenous vessel is largely due to the fact that it is a living, natural vessel with all the necessary attributes required of a blood vessel. A number of different biological strategies have been undertaken to mimic these natural qualities, which are required for reconstructing a functional vascular replacement. This review will outline some of the different approaches that have been explored to bridge the gap between the purely synthetic approach and the use of natural blood vessels. In particular, these hybrid approaches include 1) biological modification of conventional synthetic prostheses (sealants); 2) endothelial cell (EC) seeding of synthetic and biosynthetic prostheses; 3) mandrel-guided fibro-collagenous tubes (FCTs); and 4) tissue engineered vascular substitutes. The review does not include developments in the use of natural biomimetic scaffolds—for example, allografts, bovine ureter, bovine carotid heterografts, acellular matrix vessels, small intestinal submucosal substitutes—which may have the potential to remodel into hybrid blood vessels after implantation.^[1–3]

BIOLOGICAL SEALANTS FOR SYNTHETIC VASCULAR PROSTHESES

Synthetic polymer materials have been used successfully for a number of vascular procedures. The majority of these materials commercially available are composed of polyester textile materials (either knitted or woven) or ePTFE. The main disadvantage with the porous textile materials is their permeability to blood. While ePTFE is microporous and impermeable to whole blood, there is bleeding from the needle holes immediately postimplantation, or when the prosthesis is used as a vascular access device and regularly punctured in patients receiving dialysis for renal failure. The concept of coating or sealing these synthetic prostheses was introduced primarily to reduce this porosity and reduce bleeding at the time of implantation.^[4] In addition, these coatings provide a means of surface passivation that can prevent early thrombogenic responses.

Originally, preclotting of the prosthesis involved soaking the prosthesis in the patient's own blood. More recently, a number of materials including albumin, collagen, gelatine, fibrin, and alginate have been used as alternate biological compatible components to precoat woven or knitted polyester and ePTFE, to produce what has been colloquially referred to as a "biohybrid" device.^[5,3] The coating is intended to be a temporary sealant that resorbs with time to leave a patent porous prosthesis that allows sufficient host tissue in-growth to sustain its long-term function.

There are potential problems that might arise with biological coatings. For example, the concerns on the use of collagen in vascular coatings include immunogenicity, thrombogenicity, and complement activation. The potential immunogenic nature of collagen has been extensively reviewed.^[6] These issues can be minimized by the use of cross-linking agents to treat the collagen. Thus, some prostheses are lightly cross-linked with formaldehyde [e.g., Hemashield graft (Boston Scientific, Massachusetts, U.S.A.); Dialine warp-knitted polyester prosthesis (Cardial S.A., St. Etienne, France)], while others are more strongly cross-linked with glutaraldehyde [e.g., Hemaguard (InterVascular,

Florida, U.S.A.]). Although all of these devices are successful in preventing prominent transinterstitial bleeding after restoration of circulation, the thrombogenicity and particularly the wound healing responses vary between the different cross-linked collagen coated prostheses.

The patency of the vessels is largely attributable to the polyester component and appears to be independent of the coating. By and large, the polyester devices are prone to thrombosis and dilatations, and these consequences are not altered, favorably or adversely, by the additional treatment.^[7] The degree of cross-linking will have obvious effects on the rate of resorption of the protein material and in some animal studies, this has resulted in different healing responses. The lightly cross-linked sealants allowed good neointimal formation and rapid tissue integration in the inner capsule, while the more cross-linked prostheses showed a delayed resorption and reduced tissue integration.^[8,9]

Clinical studies using collagen-coated polyester prostheses have largely confirmed the safety of these biohybrid prostheses, showing acceptable patency and minimal occlusions or other complications such as thrombogenicity and adverse immunological responses.^[2,7,10] “Doubly” sealed collagen devices, which include an inner luminal coating, can be more prone to activating platelets and thrombus formation, while the more highly cross-linked prostheses may have less platelet deposition and subsequent thrombus formation.^[11,12]

A number of gelatine-sealed polyester (Gelseal™ and Gelsoft™ Vascutek Terumo) and ePTFE (SEALPTFE, Vascutek Terumo Ltd) prostheses are also currently available on the market. An *in vivo* study in a dog model showed that the use of a gelatine coating on an ePTFE prosthesis significantly reduced the bleeding from suture holes in the immediate post-operative period. The coating did not appear to effect the development of neointima, and the healing was similar to control ePTFE prostheses.^[13]

ENDOTHELIAL SEEDING OF PROSTHESES

Although there are a number of factors that contribute to the poor patency of conventional synthetic, biohybrid, and biosynthetic (see below) devices, in small-diameter vascular replacements in low-flow areas, thrombosis and subsequent occlusion is the key factor in failure. At the biological level, the predominant cells that are critical to preventing this process are the ECs that form a continuous surface monolayer in the natural blood vessel and are absent in “off-the-shelf” prostheses.

A minimal prerequisite to manufacturing a functional vascular prosthesis in culture or in animals is the process of actively isolating and expanding ECs and using these ECs to form a lining into the lumen of the device. Synthetic prostheses do not spontaneously form EC linings in humans, except for an occasional region near the anastomotic junction. The process of seeding should therefore, in theory, reduce the thrombogenicity of these artificial vessels. Unfortunately, to date, this tedious process has not been reliable for commonly used synthetic prostheses. Various strategies and tissue sources for retrieving autologous ECs have been reported, but to date there is still only limited optimism in the potential to use these cells reliably in the clinic. There are currently two broad strategies that are used in seeding: one involves a single stage (extraction of cells and immediate use), the other involves a two-stage process that includes prolonged cell culture prior to use. A variety of differing techniques that use a selection of different tissue sources (e.g., vein, artery, fat) have been used to retrieve ECs; these methods have been reviewed elsewhere.^[14] The limited availability of ECs is a key issue. More recently, alternate sources have been examined, including bone marrow derived cells and umbilical cord blood. Also, gene therapy approaches to introduce and optimize EC function have been developed.

Overall, cell seeding on Dacron and ePTFE prostheses has had disappointing outcomes, although there have been occasional reports of improved clinical patency with both ePTFE and Dacron grafts.^[15–17] These have used high-cell-density seeding (sodding) to counteract the hostile pulsatile environment and subsequent loss of seeded cells in culture. In the clinical study using two-stage EC seeding onto ePTFE, the high-density seeding was carried out using RGD-enriched fibrin glue.^[16] Even then, there is variability caused by the degradation or disattachment of the biological substrate. As a result of the variability of EC seeding, attempts have been made to improve the cell attachment and retention of seeded cells onto Dacron and ePTFE surfaces under arterial flow. Most of these approaches have involved the attachment of proteins, peptides, or surface chemical modification by methods such as plasma glow discharge. These included passivating proteins like albumin or albumin–heparin, which minimize platelet adhesion and activation, as well as proteins and peptides that could enhance EC binding in culture—gelatin, collagen, fibronectin, laminin, and RGD.^[15] Other approaches have included preclotting these surfaces with blood, plasma, serum fibrin glue, gelatin, and collagen without EC seeding, as discussed above.

In addition there have been developments in alternate polymer materials. One of the disadvantages with the ePTFE grafts is the issue of compliance mismatch.



The MyoLink graft is a compliant poly(carbonate-urea) urethane tube that has similar elasticity to the natural artery and has relatively better cell retention than ePTFE. Covalent attachment of RGD, heparin, or a combination of RGD and heparin to the surface can improve the initial binding of EC, but after 6 hr of pulsatile flow, the RGD/heparin was less thrombogenic.^[18] For single-stage cell seeding the initial time and extent of EC binding is critical. Fibronectin, or RGD-based peptides derived from fibronectin, is a natural integrin receptor on the EC, but fibronectin and peptides are normally susceptible to cleavage and proteolysis. Fibronectin-like engineered protein polymers containing multiple copies of RGD, in combination with heparin, have shown improved retention of ECs as well as the smooth muscle cells (SMCs) that are critical to the integrity of the blood vessel.^[19] Recently, a novel hyaluronan-based biomaterial (Hyaff-11) has been used and EC seeded onto these biodegradable scaffolds can produce a fully functional matrix.^[20] Also, fibrin gel, which is the only substrate that has been used clinically to retain ECs, could be used as a carrier for ECs and SMCs to allow rapid remodeling within cylindrical tissue constructs to generate a developing vessel that produces collagen and elastin fibers—a concept that is discussed later, in the section on tissue engineered prostheses.^[16,21]

The concept of EC seeding has also been extended to “off-the-shelf” porcine acellular matrix tubes and biosynthetic grafts [Omniflow Vascular Prosthesis™ (OVP)—discussed later]. The OVP is a hybrid device of collagen and Dacron that is produced *in vivo* and has a natural matrix coverage on the lumen surface, and so does not require further surface modifications to allow cell attachment and retention. Endothelial cell seeding on the OVP has demonstrated that the natural unmodified luminal surface can form a confluent monolayer of EC within 24–48 hr and that these cells are retained under physiological shear stress.^[22]

MANDREL-GUIDED FCTS

The basis for the development of FCTs as vascular replacements stems from the knowledge that implantation of synthetic materials into a host will lead to a foreign body fibrous tissue response, with the extent and nature of this response often dictated by the material itself. In the development of FCTs a variety of temporary mandrel supporting scaffolds have been used. Later, in the evolution of these FCTs, a variety of reinforcement synthetic meshes have been used to further strengthen the developing tube. These latter FCTs are true hybrid vascular devices integrating natural collagen with synthetic polymer and are

distinct from synthetic grafts coated with collagen or other biological molecules as sealants.

The Autogenous FCT

In 1953, Pierce and Baltimore reported on the first production and autogenous implantation of an FCT in dogs, using subcutaneously implanted polypropylene tubes. The 10 aortic implants remained patent up to 30 mo. The development of these autogenous “self-generating” tubes in soft connective tissue continued with a variety of different polymeric and metallic mandrel supports, although none of these could prevent the eventual development of aneurysms and vessel rupture. The early development of FCT devices has been reviewed previously.^[2,3]

Hufnagel suggested that implanted ePTFE mandrels could be surrounded with a polyester (Dacron) mesh to reinforce the FCT, and at the same time Assefi and Parsonnet used a reinforcing polypropylene mesh to generate an FCT that remained patent without signs of aneurysmal changes. In both instances, the reinforced mesh was infiltrated by inflammatory cells and became an integral part of the new collagenous network of the FCT, resulting in a true hybrid graft. These early studies were encouraging and formed the basis of the studies of Sparks and his “Sparks mandril” (sic).^[23] Sparks initially made use of slender cylindrical stainless steel mandrels. These so-called “tissue dies” consisted of an outer perforated tubular shell of stainless steel that would allow cell infiltration and an inner polished tubular mandrel of stainless steel that was wrapped around with Dacron, all over a central tie rod through the length of the mandrel.^[23] The tissue that formed over 1–6 mo was vascularized and was rich in collagen, and importantly the knitted Dacron had become an integral part of the hybrid FCT. Substitutions were made to replace the metallic die with a silicone mandrel. Studies from Parsonnet and colleagues demonstrated good long-term patency over 4 yr with no signs of aneurysm formation in these FCTs, which had been manufactured using a polyester (Dacron) or polypropylene mesh support over a silicone mandrel template. Similarly, Sparks, using this modified approach to generate FCTs, reported high patency rates in animal and limited human trials. The Sparks FCT, which was patented in 1973, was commercialized by Edwards Laboratories and used by many surgical groups. Following these favorable animal and early clinical results, the “Sparks mandrel” remained popular until subsequent clinical reports described poor results. For example, Hallin and Sweetman reported 100% patency in the ilio-iliac and ilio-femoral sites, but only 26% (10/38) in the femoropopliteal and 0% in the femorotibial and axillofemoral positions. Similar clinical findings were

also reported from other groups, showing very low patency rates of around 18%. The reasons for failure of these vessels were predominantly early thrombosis, aneurysm formation, and late thrombosis due to luminal narrowing. Studies from one of these groups had included a number of additional modifications to the design and manufacture of these FCTs. These variations included using xenogenic species like dogs, goats, and pigs, using fresh and frozen FCTs, and glutaraldehyde and enzyme treatment of the FCTs. After the initial promising animal studies, the use of autogenous FCTs in humans, based largely on the Sparks mandrel, did not continue and the technology was withdrawn from the market in 1976.^[2,3]

The Peritoneal-Grown Autologous Graft

A further development of the Sparks mandrel concept is the autologous cell–tissue hybrid blood vessel that is generated by peritoneal implantation.^[24] This approach is based on the prior knowledge that any foreign body that is implanted into the peritoneal cavity will induce a foreign body reaction that results in a tissue comprising collagen, myofibroblasts and a single outer layer of mesothelial cells. It is the presence of this layer of mesothelial cells that could have an impact on the generation of a nonthrombogenic graft. Typically, the peritoneal-grown vessel is reversed inside out to generate what seems to mimic a natural artery with, after reversal, an inner mesothelial cell lining, media comprising myofibroblasts, and an outer collagen-rich adventitia. Mesothelial cells, while not entirely the same as ECs, do have some common biological characteristics including secretion of prostacyclin and fibrinolytic activity. In initial studies in rats and rabbits using this approach, patency rates of 67% and 70% were obtained over 4 mo. Subsequent animal trials in dogs have shown improvements, but to date there is no indication of long-term performance and it is still too early for clinical trials.

An additional benefit of this artificial autologous hybrid graft is the ability to remodel under arterial shear stress.^[25] It is postulated that further differentiation of cells and tissue maturation within the hybrid graft occurs, a process that has not been demonstrated in human arteries. Alternatively, some of the animal experiments in rabbits were conducted without everting the graft so that the mesothelial cell layer remained on the outside. In these cases, ECs were found on the inner luminal surface, which could indicate migration of cells from the adjacent host artery.^[26]

Current FCT Vascular Prostheses

At present there is no autologous FCT in the clinic. However, a use of this technology to provide a

xenogenic off-the-shelf hybrid vessel has been successful and a commercially available device, the OVP (Bio Nova International, North Melbourne, Australia), is in current clinical use. The OVP is a glutaraldehyde-stabilized xenogenic FCT that is manufactured in sheep using a polyester (Dacron) covered silicone mandrel.^[27] This strategy of using a xenogenic host, a polyester reinforced mandrel support, and an effective stabilization process has generated a device that has proved to be effective for peripheral arterial replacement. The stabilization of the collagen in the hybrid device minimizes the extent of cross-species immunological reactivity to allow off-the-shelf transplantation into animals and into humans. The stabilization also allows the collagenous matrix to persist for long implantation times. Also, significant amounts of new collagenous tissue begin to deposit within the vessel wall soon after implant, and continue to accumulate with time, providing a very durable and patent prosthesis. The development of the OVP has been described in previous reviews.^[2,28,29] During production, the Dacron mesh within the OVP becomes fully encased with collagen during its implantation in sheep. The prosthesis comprises an organized extracellular matrix of collagen type I and a high content of collagen type III. While elastin is notably absent, the OVP does contain other microfibrillar elastic proteins, including fibrillin. Collagen is found in a circular array in regions surrounding the polyester fibers, but further away from the silicone mandrel, the collagen forms an elongated, highly aligned layered structure. The inner region of the OVP is somewhat less dense comprising loose connective tissue, and the thickness of this layer varies inversely with the presence of nodes of polyester fiber bundles. The OVP has a burst strength greater than 1688 mm Hg and maintains excellent physical integrity following both *in vitro* mechanical and fatigue testing.^[2,28,29]

Following the initial reports of the use of this prosthesis in animal models, it has undergone an extensive series of *in vitro* and *in vivo* tests, and is now used successfully in clinical applications.^[2,28,29] The patency of this collagen-based polymer composite prosthesis is comparable to, or better than, other available materials. Clinical studies have shown excellent long-term patency, and a study using the more recent OVP II as a femoropopliteal bypass showed a primary patency of 83% and a secondary patency of 91% over 3 yr.^[30] For use as a small-diameter (<3–4 mm) vascular replacement, the OVP has yet to be fully tested. The major issues here would be to generate a luminal surface that was a deterrent for platelet activation and thrombosis, but that was favorable for endothelialization. Further improvements to the manufacture of the device can produce an inner surface that is less undulated and more uniformly covered with



collagen, which can effectively protect the prosthesis against potential deleterious polyester breaking through the inner surface.^[31]

TISSUE ENGINEERED VASCULAR SUBSTITUTES

Tissue engineering in cell culture is an emerging multidisciplinary technology that combines basic engineering concepts with biological sciences to generate a functional tissue equivalent that can replace and allow regeneration of damaged tissues or organs. Tissue engineering should enable the fabrication of a regenerated blood vessel that is almost indistinguishable from the damaged one. The choice of cells and natural or synthetic polymer scaffolds that can direct the growth and differentiation of a cell type(s) into the appropriate tissue architecture is critical to the regenerated tissue. The basis of this approach arises from early studies on hybrid biomaterial scaffolds, as well as the concept of seeding these scaffolds with ECs *in vitro*. The distinction can be made in the choice of the biomaterial scaffold that is used with the early approaches mainly using biostable synthetic polymers, natural tissue scaffolds, or combinations of these to induce remodeling or cell seeding. The current approaches are predominantly using biodegradable polymer scaffolds and a selection of different cell types and seeding strategies.

The Basis for Tissue Engineering of Vessels

Following from the initial concept of engineering tissue-like structures by cell-based collagen contraction, Weinberg and Bell in 1986 paved the way for further refinements and developments of a tissue engineered vascular prosthesis (TEVP).^[32] In this instance, the tissue comprised autologous cell types remodeled in a collagen matrix to meet the demands of a functional living blood vessel. Weinberg and Bell produced the first biological TEVP using porcine skin collagen or rat tail tendon collagen with bovine aortic ECs, SMCs, and adventitial fibroblasts.^[32] The results were encouraging but the final prototype lacked the necessary mechanical strength to withstand moderate pressures, even with the inclusion of a supportive Dacron sheath(s). A number of factors may have contributed to the lack of appropriate mechanical strength. These included the random or longitudinally orientated SMCs compared with the circumferential arrangement in natural arteries, the switching on in culture of metalloproteinases that could be degrading the tissue, and the somewhat low densities of SMC and collagen, only around one-fourth to one-eighth of those of

natural vessels. The investigations that followed this initial approach seem to have taken some of these factors into consideration, but still there remains a huge gap in bringing this technology into the clinic.

Current TEVPs

There are three distinct layers that need to be designed and cultivated to mimic the natural artery: an intimal layer composed of a continuous monolayer of ECs designed to enhance hemocompatibility, a medium composed of SMCs designed to provide strength and elasticity, and an adventitial layer composed of fibroblasts designed to augment the durability and integration of the vessel.

Overall, three strategies have been documented including EC seeding alone or within a collagen or fibrin gel, layer-by-layer reconstruction of the EC-seeded intima and SMC-seeded media, and finally a three layered structured tissue with the additional fibroblast-seeded adventitia.^[33,34] Other differences in strategies have arisen from the use of a supportive synthetic mesh in some developments but not in others to generate a mechanically sound finished vessel.^[33,34] The layer-by-layer approaches are similar in conceptual design, and rely on stepwise sequential generation of layers of each type of tissue, using temporary tubular supports to generate the necessary vessel architecture.

A Japanese group from Kyushu University with others developed a self-organized tubular hybrid prosthesis using a selection of mold designs, collagen and SMC seeding densities that formed a media matrix after contraction, and that produced burst strengths around 100 mm Hg.^[35] Implantation of the vessels in the posterior vena cava of dogs from where the cells were derived resulted in massive bleeding due to vessel rupture within a few days. With a supportive outer Dacron mesh, 9 of the 14 implants remained patent at various times of explant from 1 to 24 weeks after implantation. In the patent explants, elastic lamellae were present, SMCs were contractile in nature, and there were signs of good tissue augmentation.^[33] Indeed, while EC retention is critical to the thrombogenicity of the engineered vessel, the remodeling of cells and tissues is equally critical to withstand the physiological shear stress. The supramolecular matrix assemblies of collagen and elastin, and the circumferential orientation of the SMCs are influenced by the different strategies used to build these vessels—as expected, more rapid and efficient remodeling occurs when all cells are used in a three-layer approach.^[36] In addition, various approaches have been used to influence the phenotypic switching and orientation of SMCs, including peristaltic flow, the use of perfusion

bioreactors, collagen contraction along the axis of the supporting scaffold, and magnetic fields.^[36]

Work from Laval University, using human vascular cells and human collagens similar to those described by others, failed to generate a mechanically sound vessel.^[37] While the use of a supportive Dacron mesh had been shown to augment the structural strength within the cultured vessel, a later study has outlined a modified method using similar sequential buildup of the various vessel layers, but without the need for any synthetic or exogenous biomaterial support.^[33,34] The SMCs were cultured with ascorbic acid and generated a cohesive cellular sheet, which was then further cultured around an appropriate temporary support to generate the media, prior to further seeding with ECs. The final cultured prototype was well organized with circumferentially orientated SMCs, numerous matrix proteins including elastin and desmin, and a functional EC layer. The three-layered vessel had a burst strength of over 2000 mm Hg, which was far in advance of the original tissue engineered vessel of Weinberg and Bell.^[32] Short-term implantation of these “human vessels” was performed in dogs and, as expected, there was only 50% patency owing to the xenogenic nature of the vessel. A later study, using similar culturing and testing of autologous ovine tissue engineered pulmonary arteries, has shown no early signs of aneurysmal changes.^[35,38] The tissue engineering strategy in this case was different in that the implanted vessel comprised a completely bioresorbable scaffold of poly (caprolactone-colactic acid) copolymer. The use of biodegradable synthetic scaffolds may offer advantages to further developments in this area—polyglycolic acid and polyglycolic acid-polyhydroxyalkanoate are polymers that have been used for two- or three-layer tissue engineering approaches.^[39] Alternatively, natural tissue scaffolds like elastin could be used with the current collagen-based multicell seeding approaches to increase the tensile strength of the cultured vessel.^[40]

CONCLUSIONS

A wide range of hybrid vascular prostheses have now been established, and from these a number of clinically effective devices have emerged. Typically, these devices work well at larger diameters, but there is still a significant unmet medical need for small-diameter prostheses, for example, for coronary artery replacement. This is where the emerging technologies of tissue engineering may produce significant benefits. However, there are still a number of issues that need to be addressed before the tissue engineering approach becomes the benchmark for further developments. The most obvious one is the long time that is needed to

produce one of these vessels. The system is highly advantageous in that it is designed for autologous arterial substitution. What is required is further research to elucidate matters like the amount of biopsy required, the source of materials that best suits the application, and optimization and quality assurance systems to guarantee generation of a vessel with the same biological features from batch to batch.

ARTICLES OF FURTHER INTEREST

Biocomposites; Tissue Engineering of Blood Vessel



REFERENCES

1. Griesler, H.P. *Medical Intelligence Unit. New Biologic and Synthetic Vascular Prostheses*; R.G. Landes Co.: Austin, TX, 1991; 1–91.
2. Ramshaw, J.A.M.; Edwards, G.A.; Werkmeister, J.A. Tissue-polymer composite vascular prostheses. In *Encyclopedic Handbook of Biomaterials and Bioengineering*; Wise, D.L., Trantolo, D.J., Altobelli, D.E., Yaszemski, M.J., Gresser, J.D., Schwartz, E.R., Eds.; Marcel Dekker: New York, 1995; Part B, Vol. 2, 953–978.
3. Werkmeister, J.A.; Edwards, G.A.; Ramshaw, J.A.M. Collagen-based vascular prostheses. In *Biomaterials Engineering and Devices: Human Applications*; Wise, D.L., Trantolo, D.J., Lewandowski, K-U., Gresser, J.D., Cattaneo, M.V., Yaszemski, M.J., Eds.; Human Press: Totowa, NJ, 2000; Vol. 1, 121–136.
4. Yates, S.G.; Barros D'sa, A.A.B.; Berger, K.; Fernandez, L.G.; Wood, S.J.; Rittenhouse, E.A.; Davis, C.C.; Mansfield, C.C.; Mansfield, P.B.; Sauvage, L.R. The preclotting of porous arterial prostheses. *Ann. Surg.* **1978**, *188*, 611–622.
5. Gundry, S.R.; Behrendt, D.M. A quantitative and qualitative comparison of fibrin glue, albumin, and blood as agents to pretreat porous vascular grafts. *J. Surg. Res.* **1987**, *43*, 75–77.
6. Werkmeister, J.A.; Ramshaw, J.A.M. Immunology of collagen-based biomaterials. In *Handbook of Biomaterials Engineering*; Wise, D.L., Ed.; Marcel Dekker Inc.: New York, 2000; 739–759.
7. Freischlag, J.A.; Moore, W.S. Clinical experience with a collagen impregnated knitted Dacron vascular graft. *Ann. Vasc. Surg.* **1990**, *4*, 449–454.
8. Scott, S.M.; Gaddy, L.E.; Sahmel, R.; Hofmann, H. A collagen coated vascular prosthesis. *J. Cardiovasc. Surg.* **1987**, *28*, 498–504.
9. Jonas, R.A.; Ziemer, G.; Schoen, F.J.; Britton, L.; Castaneda, A.R. Biological sealants and knitted Dacron conduits: comparison of collagen and fibrin glue pretreatments in circulatory models. *Ann. Thorac. Surg.* **1987**, *44*, 283–290.
10. Schmiedt, W.; Neufang, A.; Scholl, E.; Schmid, F.X.; Oelert, H. Immune response to gelatin- and

- collagen-impregnated aortic Dacron grafts. A randomized study. *Vasc. Surg.* **1996**, *30*, 513–518.
11. Konrad, P.; Dougan, P.; Bergqvist, D. Acute thrombogenicity of collagen coating of Dacron grafts: an experimental study in sheep. *Eur. J. Surg.* **1992**, *6*, 67–72.
 12. Guidoin, R.; Marceau, D.; Couture, J.; Rao, T.J.; Merhi, Y.; Roy, P.-E.; de la Faye, D. Collagen coatings as biological sealants for textile arterial prostheses. *Biomaterials* **1989**, *10*, 156–165.
 13. Edwards, G.A.; Beck, C.; Charles, J.A. 2001. Evaluation of a gelatin-sealed ePTFE vascular prosthesis in a dog model. Proceedings 25th Anniversary Meeting of European Society for Biomaterials, London, U.K., Sep 12–14, 2001; 106.
 14. Tiwari, A.; Salacinski, H.J.; Hamilton, G.; Seifalian, A.M. Tissue engineering of vascular bypass grafts: role of endothelial cell extraction. *Eur. J. Vasc. Endovasc. Surg.* **2001**, *21*, 193–201.
 15. Seifalian, A.M.; Tiwari, A.; Hamilton, G.; Salacinski, H.J. Improving the clinical patency of prosthetic vascular and coronary bypass grafts: the role of seeding and tissue engineering. *Artif. Organs* **2002**, *26*, 307–320.
 16. Deutsch, M.; Meinhart, J.; Fischlein, T.; Preiss, P.; Zilla, P. Clinical autologous in vitro endothelialization of infrainguinal ePTFE grafts in 100 patients: a 9-year experience. *Surgery* **1999**, *126*, 847–855.
 17. Leseche, G.; Ohan, J. Enhanced patency of venous Dacron grafts by endothelial cell seeding. *Ann. Vasc. Surg.* **1995**, *9*, 325–326.
 18. Tiwari, A.; Salacinski, H.J.; Punshon, G.; Hamilton, G.; Seifalian, A.M. Development of a hybrid cardiovascular graft using a tissue engineering approach. *FASEB J.* **2002**, *16*, 791–796.
 19. Rashid, S.T.; Salacinski, H.J.; Button, M.J.; Fuller, B.; Hamilton, G.; Seifalian, A.M. Cellular engineering of conduits for coronary and lower limb bypass surgery: role of cell attachment peptides and pre-conditioning in optimising smooth muscle cells (SMC) adherence to compliant poly(carbonate-urea)urethane (MyoLink) scaffolds. *Eur. J. Vasc. Endovasc. Surg.* **2004**, *27*, 608–616.
 20. Turner, N.J.; Kielty, C.M.; Walker, M.G.; Canfield, A.E. A novel hyaluronan-based biomaterial (Hyaff-11) as a scaffold for endothelial cells in tissue engineered vascular grafts. *Biomaterials* **2004**, *25*, 5955–5964.
 21. Swartz, D.D.; Russell, J.A.; Andreadis, S.T. Engineering of fibrin-based functional and implantable small-diameter blood vessels. *Am. J. Physiol. Heart Circ. Physiol.* **2005**, *288*, H1451–H1460.
 22. Birchall, I.E.; Lee, V.W.; Ketharanathan, V. Retention of endothelium on ovine collagen biomatrix vascular conduits under physiological shear stress. *Biomaterials* **2001**, *22*, 3139–3144.
 23. Sparks, C.H. Autogenous grafts made to order. *Ann. Thorac. Surg.* **1969**, *8*, 104–113.
 24. Hoenig, M.R.; Campbell, G.R.; Rolfe, B.E.; Campbell, J.H. Tissue-engineered blood vessels: alternative to autologous grafts? *Arterioscler. Thromb. Vasc. Biol.* **2005**, *25*, 1128–1134.
 25. Campbell, J.H.; Efendy, J.L.; Campbell, G.R. Novel vascular graft grown within recipient's own peritoneal cavity. *Circ. Res.* **1999**, *85*, 1173–1178.
 26. Daly, C.D.; Campbell, G.R.; Walker, P.J.; Campbell, J.H. In vivo engineering of blood vessels. *Front. Biosci.* **2004**, *9*, 1915–1924.
 27. Ketharanathan, V.; Christie, B.A. Glutaraldehyde-tanned ovine collagen conduits as vascular xenografts in dogs. *Arch. Surg.* **1980**, *115*, 967–969.
 28. Edwards, G.A.; Roberts, G. Development of an ovine collagen-based composite vascular prosthesis. *Clin. Mater.* **1992**, *9*, 211–223.
 29. Werkmeister, J.A.; White, J.F.; Ramshaw, J.A.M. Evaluation of the Omniflow collagen-polymer vascular prosthesis. *Med. Progr. Technol.* **1994**, *20*, 231–242.
 30. Koch, G.; Gutsch, S.; Pasher, O.; Fruhwirth, J.; Glanzer, H. Analysis of 274 Omniflow vascular prostheses implanted over an eight year period. *Aust. N.Z. J. Surg.* **1997**, *67*, 637–639.
 31. Ramshaw, J.A.M.; Casagrande, F.; White, J.F.; Edwards, G.A.; Hunt, J.A.; Williams, D.F.; Werkmeister, J.A. Effects of mesh modification on the structure of a mandrel grown biosynthetic vascular prosthesis. *J. Biomed. Mater. Res.* **1999**, *47*, 309–315.
 32. Weinberg, C.B.; Bell, E. A blood vessel model constructed from collagen and cultured vascular cells. *Science* **1986**, *231*, 397–400.
 33. Hirai, J.; Matsuda, T. Venous reconstruction using hybrid vascular tissue composed of vascular cells and collagen: tissue regeneration process. *Cell Transplant* **1997**, *5*, 93–105.
 34. L'Heureux, N.; Paquet, S.; Labbe, R.; Germain, L.; Auger, F.A. A completely biological tissue-engineered human blood vessel. *FASEB J.* **1998**, *12*, 47–56.
 35. Shinoka, T.; Shumtim, D.; Ma, P.X.; Tanel, R.E.; Isogai, N.; Langer, R.; Vacanti, J.P.; Mayer, J.E. Creation of viable pulmonary artery autografts through tissue engineering. *J. Thorac. Cardiovasc. Surg.* **1998**, *115*, 536–545.
 36. Matsuda, T. Recent progress of vascular graft engineering in Japan. *Artif. Organs* **2004**, *28*, 64–71.
 37. L'Heureux, N.; Germain, L.; Labbe, R.; Auger, F.A. In vitro construction of a human blood vessel from cultured vascular cells: a morphologic study. *J. Vasc. Surg.* **1993**, *17*, 499–509.
 38. Shinoka, T.; Imai, Y.; Ikeda, Y. Transplantation of a tissue-engineered pulmonary artery. *New Engl. J. Med.* **2001**, *344*, 532–533.
 39. Shum-Tim, D.; Stock, U.; Hrkach, J.; Shinoka, T.; Lien, J.; Moses, M.A.; Stamp, A.; Taylor, G.; Moran, A.M.; Landis, W.; Langer, R.; Vacanti, J.P.; Mayer, J.E., Jr. Tissue engineering of autologous aorta using a new biodegradable polymer. *Ann. Thorac. Surg.* **1999**, *68*, 2298–2304.
 40. Berglund, J.D.; Nerem, R.M.; Sambanis, A. Incorporation of intact elastin scaffolds in tissue-engineered collagen-based vascular grafts. *Tissue Eng.* **2004**, *10*, 1526–1535.

Hydrogels

Junji Watanabe
Yoshihiro Kiritoshi
Kwang Woo Nam
Kazuhiko Ishihara

*Department of Materials Engineering, School of Engineering, The University of Tokyo,
Tokyo, Bunkyo-ku, Japan*

INTRODUCTION

Hydrogels refer to certain materials that are able to swell under conditions of excess water and hold a large amount of water in the wet state. Hydrogels generally consist of three-dimensional polymer networks that are cross-linked chemically and/or physically. The chemical properties of hydrogels are determined by the polymer backbone, the functional side chain in the monomer unit, and the cross-linking agent. The physical properties—for example, mechanical strength and swelling ratio—are controlled by the cross-link density. In general, hydrogels do not show any biodegradable behavior. Biodegradable hydrogels may be easily designed using natural polymers that are susceptible to enzymatic degradation or synthetic polymers that possess hydrolyzable moieties.

Biomaterials and biomedical engineering practices involve drug delivery carriers, tissue engineering scaffolds, and biomedical devices. Hydrogels provide adequate materials for these purposes. The research area of biomaterials is a multidisciplinary field composed of materials science, chemical engineering, medical engineering, and pharmacology. In particular, a biomaterial can play an important role as a bio-matrix and a bio-interface. Hydrogels provide excellent properties for improving conventional biomaterials, using such elements as copolymerization and polymer blends. As drug carriers, hydrogels can achieve stimuli-responsive drug release (for example, using pH or temperature) synchronized with enzymatic or non-enzymatic degradation. Hydrogels may be loaded with a variety of drugs and solubilizers or excipients. One major advantage of using hydrogels as a three-dimensional tissue reconstruction scaffold is the ability of these highly hydrophilic cross-linked polymer networks to swell and absorb sufficient amounts of cell culture medium. The permeability of the medium is important to the exchange of water-soluble molecules such as nutrients and gases. Widely used hydrogel medical devices include not only disposable materials, such as syringes, sample bags, and sample tubes, but

also artificial organs such as blood vessels. The most promising applications may involve lenses, muscle models, and artificial cartilage. Contact lenses are a major application of hydrogel products due to their high transparency.

This chapter focuses not only on fundamental principles but also on the use of recent anomalous hydrogels as biomaterials. Hydrogels can be classified as chemically cross-linked hydrogels for use as soft materials, and physically cross-linked hydrogels made of synthetic polymers, natural polymers, and their hybrids. In particular, a series of bio-inspired phospholipid polymers show unique properties for application as drug delivery carriers, tissue engineering scaffolds, and biomedical devices.

CHEMICAL GELS

The chemical gels, which are cross-linked by covalent bonds, are synthesized by polymerization of monomers with some difunctional monomers, called cross-linking agents, such as ethylene glycol dimethacrylate (Fig. 1a). The gel is affected by changes in pH, temperature, solvents, and other factors in the surroundings. A chemical gel shrinks or swells, showing a volume phase transition. That is, controlling the polymer network by a change in the surroundings will lead to a change in the gel volume. This phenomenon makes the gel especially useful as a functional material for medical devices.

Soft Contact Lenses

The foreign-body reaction often becomes a problem when artificial materials come in contact with the human body. A significant difference in the surface free energy between the material and the body is the cause of this undesirable reaction. A hydrogel can avoid this problem because the human body consists of network structures including biomacromolecules and water in a high percentage, which corresponds to the hydrogel



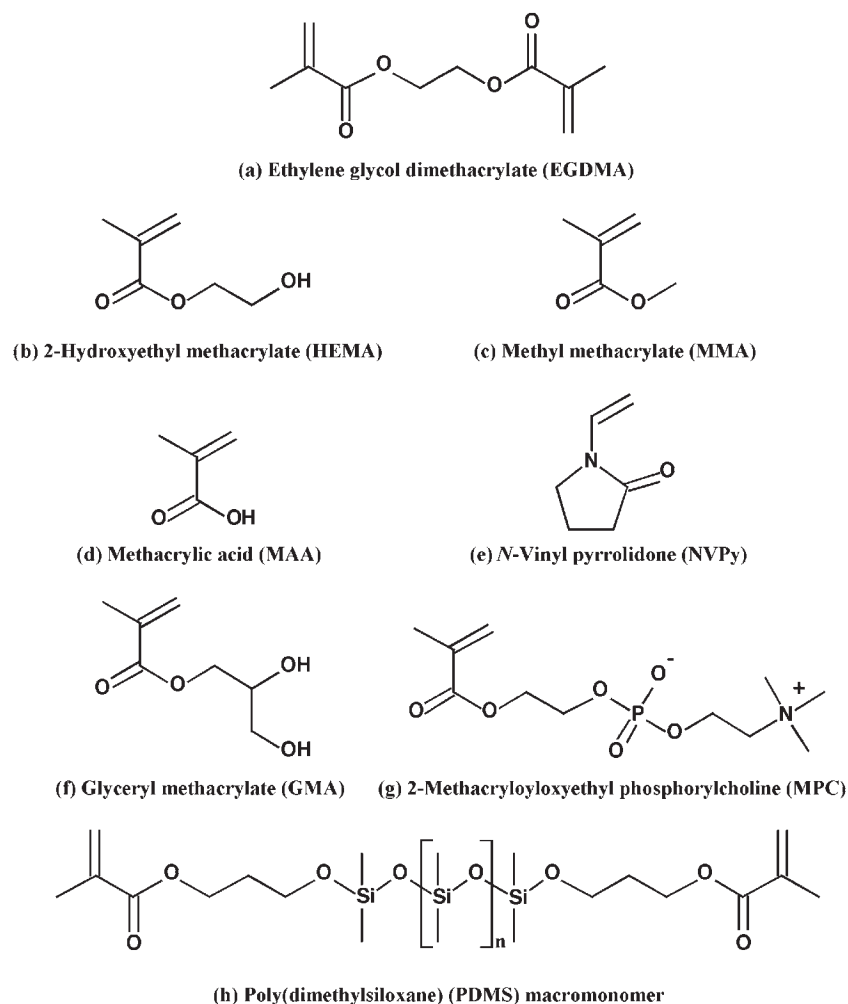


Fig. 1 Chemical structures of various monomers typically used for contact lens material.

structure itself. The surface free energies of the hydrogel and the human body are almost the same, so that the nonspecific interaction between the hydrogel and cell or protein is low, inhibiting protein adsorption following the foreign-body reaction. Therefore, the hydrogel is the most suitable biomaterial.

One of the remarkable properties of a hydrogel is softness, the property of an elastomer such as a rubber. Moreover, hydrogels are characterized by transparency and transmissibility of molecules through water flow. For the hydrogels in practical use, the soft contact lens (SCL) is the ideal manufactured article to take advantage of these properties. The eye naturally needs clarity for vision and the cornea needs oxygen by dermal respiration from the atmosphere. The properties of a hydrogel comply with these requirements.

Background on soft contact lens material

Poly(2-hydroxyethyl methacrylate) (HEMA, Fig. 1b), the first appearance of a soft biomaterial, was a great achievement by Otto Wichterle^[1] and can not

be omitted from a history of the SCL. Amazingly, this invention is still the main material used for SCLs some 40 years later. Moreover, poly(HEMA) and its copolymers have widened the range of application to include such products as intraocular lenses. Before poly(HEMA), poly(methyl methacrylate) (MMA, Fig. 1c) was known not to cause critical harm to the human body, a fact discovered when a pilot injured in the eye by a broken aircraft windshield made of poly(MMA) experienced no inflammation. As a consequence of this episode, poly(MMA) was used as the first plastic material for the hard contact lens (HCL). Poly(MMA) lenses not only have good optical properties for clear vision, including greater clarity than glass, but also properties of high strength, light weight, and ease of molding. However, poly(MMA) lenses have limitations of discomfort and prohibited long-term use because of low oxygen permeability. Wichterle thought that the introduction of hydrophilicity into the polymer could lead to amelioration of the defects of poly(MMA), so poly(HEMA) was the first material with hydrophilicity introduced in the polymer chain.

Increasing oxygen permeability

It is said that in addition to the anterior epithelium of the cornea, the endothelium of the cornea undergoes damage after long-term use of contact lenses having low oxygen permeability. As the physiology of the cornea has become understood, the oxygen permeability of contact lenses has been regarded as the important factor for homeostasis of the cornea. Without an adequate supply of oxygen, adverse events such as corneal edema may occur, followed by a number of possible physiological responses. Therefore, competitive efforts have resulted to develop materials with high performance in terms of regarding oxygen permeability.

The expression DK , where D represents the diffusion coefficient and K the partition coefficient of oxygen—inherent properties for oxygen transport through a material—has become familiar among contact lens manufacturers and users these days. There is a further relationship between oxygen flow and lens thickness; The transmissibility of a particular lens is thickness related and is given by the expression DK/t , where t refers to the thickness of the hydrogel.^[2] In the case of SCLs, there are two approaches to increasing oxygen permeability based on DK/t , that is, to reduce the thickness of the SCL or to increase the water content in the hydrogel, because oxygen permeability increases with water content. In the former case, a number of hyperthin poly(HEMA) SCLs (i.e., a center thickness of less than 0.05 mm) have been produced; however, such lenses turned out to be fragile, marked by poor mechanical properties, difficult handling, and difficult manufacture. In the latter case, the hydrophilicity of the polymer was a concern. Actually, the hydrophilicity of poly(HEMA) is too low to provide a water content of about 40%, indicating that not enough water is present to for carry sufficient oxygen to the cornea.

Water content may be adjusted by using various hydrophilic monomers to overcome poor oxygen permeability. For example, methacrylic acid, *N*-vinyl pyrrolidone, and glyceryl methacrylate are used for this purpose (Fig. 1d–f). However, the DK of pure water is less than necessary to supply the minimum oxygen needed for the cornea. There is naturally a limit to increasing oxygen permeability via increased water content in the hydrogel. The water content in the gel also plays an important role in controlling the mechanical properties of the hydrogel. The higher the water content in the gel, the lower the tensile strength. High water content and high tensile strength are ideal conditions for SCL material; however, a hydrogel cannot have it both ways.

Thus, the same problem that is associated with hyperthin SCLs, fragility, applies. Further, such lenses

tend to bind organic materials such as tear proteins, microbes from handling, or inorganic salts, especially those of calcium, all of which can affect oxygen permeability.

Silicone hydrogel

In the HCL field, a silicone-containing lens, poly(dimethylsiloxane) (PDMS, Fig. 1h), is famous for possessing a high DK and many other properties such as optical excellence, softness and flexibility similar to a hydrogel, and high tensile strength. The bulkiness of the siloxane group and the chain mobility are responsible for the high diffusibility of oxygen through siloxane-containing materials.^[3] Designing a hydrogel based on PDMS will lead to high performance in terms of oxygen permeability and good elasticity for comfortable wearing; however, the hydrophobicity of the siloxane component makes this difficult to realize. When a PDMS macromonomer is polymerized with some hydrophilic monomers, phase separation occurs and the resultant polymer is opaque.

Several strategies have been considered for improving the hydrophilicity of PDMS. Some macromonomers including PDMS and hydrophilic segments such as polyethylene oxide (PEO) in the molecule were synthesized and polymerized with some hydrophilic monomers such as dimethylacrylamide. The incorporation of a nonpolar fluorosubstituent (CF_2H) had a dramatic effect on the increased solubility with hydrophilic monomers.^[4] Those hydrogels based on PDMS macromonomers showed some improvement; however, the siloxane component, which tends to concentrate at the surface of the lens, and the low surface energy of PDMS result in very poor tear wetting, causing binding of tear lipids and lens adhesion to the cornea.

Adverse effects of fouling on the lens surface

With daily wearing of contact lenses, protein or lipid depositions create a biofilm on any kind of lens surface. Proteins have charge distributions, and they attract one another. This could result in the lens losing its ocular properties, and oxygen permeability will be disrupted. Moreover, it is possible for hydrophilic lenses to absorb chemicals into their naturally porous structure. This fouling of the SCL causes adverse events such as adherence of microbes to the lens. Bacteria can colonize in the lens and on its surface, so that the lens itself can become a potential pollutant source, a bacteria bed. This is the first stage in contact lens-associated microbial keratitis, which may ultimately lead to corneal ulceration and to blindness, in the worst case. Thus, lens cleaning and sterilizing are



crucial components of daily care, especially for SCLs. However, an inadequately cleaned and sterilized lens could also trigger these adverse events. Also, any sterilizing process, such as heat, peroxide, or chemical disinfection, can harm the lens material, leading to deterioration. Disposable lenses are a way of avoiding these risks. However, there is the problem of the cost and several other risks due to improper usage and mishandling, such as using the lens for more than a certain period.

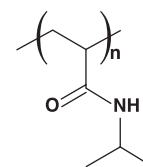
Requirements for ideal lens material

Surface fouling is inevitably associated with any kind of lens. This fouling has unfavorable effects on both the lens and its user. Unfortunately, there are no SCLs using conventional materials that do not have problems. A matter of primary importance in providing easily handled and problem-free SCLs is biocompatibility, that is, the absence of negative effects due to contact with the constituents of a living organism and the absence of toxic or injurious effects on biological systems.

Poly(2-methacryloyloxyethyl phosphorylcholine) [poly(MPC), Fig. 1g], a bio-inspired material, was modeled on the surface of a natural biological membrane. The MPC polymers are highly effective in inhibiting the near-instantaneous protein adsorption and subsequent denaturation that can lead to thrombus formation, indicating true biocompatibility.^[5,6] Recently, a poly(MPC) hydrogel was made that showed a high swelling ratio, meaning high hydrophilicity.^[7] Using this MPC and HEMA, a biomimetic SCL was made.^[8] This SCL reduced not only protein adsorption but also the risk of infection by bacteria, exhibiting significantly lower bacterial adhesion *in vitro* compared to the standard poly(HEMA) SCL. Moreover, there has been an attempt to make a hybrid hydrogel having the high oxygen permeability of a silicone hydrogel and the surface of a biomembrane characteristic of the poly(MPC). The potential of a poly(MPC) hydrogel has not been fully realized; however, the poly(MPC) hydrogel assists in widening the range of usage in association with all biomaterials that need true biocompatibility.

Volume Phase Transition

Since Tanaka reported the volume phase transition of a hydrogel in solvent,^[9] many research studies on this phenomenon have been initiated, especially for medical devices such as drug delivery systems and tissue engineering. It is desirable to release the drug from the gel by shrinkage or swelling on stimulation. For example, if a hydrogel containing a drug changes volume according



Poly(*N*-isopropyl acrylamide (NIPAAm))

Fig. 2 Chemical structure of typical polymer used for gel system employing volume phase transition.

to a change in body temperature, patients can automatically receive the drug when they have a fever. The poly(*N*-isopropyl acrylamide) [poly(NIPAAm)] hydrogel is well qualified for this use (Fig. 2). Poly(NIPAAm) becomes insoluble in water due to dehydration of the polymer chain when the water temperature is over 32°C, which is well known as a lower critical solution temperature (LCST).^[10] With this property, the poly(NIPAAm) gel could swell under 32°C but shrink above 32°C, showing a volume phase transition.

Gel with functions

Phenylboronic acid and its derivatives are well known to form covalent complexes with polyol compounds including glucose. Taking advantage of this property, Kataoka et al. have made a hydrogel (NB gel) by polymerization of NIPAAm with 3-acrylamidophenylboronic acid and a small amount of cross-linker, which shows swelling-deswelling behavior in response to a change in glucose concentration (Fig. 3a).^[11]

The NB gel showed a volume phase transition at 22°C and pH 9.0; however, the temperature rose with increased glucose concentration from 0 to 5 g/L. In the case of a 1 g/L glucose concentration (the same level as for a healthy human), the NB gel would show a remarkable change in swelling in response to glucose in the range of 25 to 30°C. In fact, the repeated on/off regulation of insulin release from the NB gel in synchrony with a change in the concentration of external glucose was confirmed. Although there is still a problem with the difference in pH and temperature in the human body, this type of glucose-responsive gel would be good news for patients under diabetes treatment.

Miyata et al. made a hydrogel that can change its swelling ratio to increase the transmissibility of molecules in response to a specific antigen (Fig. 3b).^[12] They first prepared the monomers, the modified antigen and modified antibody, using rabbit immunoglobulin G (rabbit IgG) and goat anti-rabbit IgG (GAR IgG) as the antigen and antibody, respectively. They then synthesized the polymer GAR IgG using vinyl(GAR IgG) and acrylamide (AAm). In the presence of this polymerized

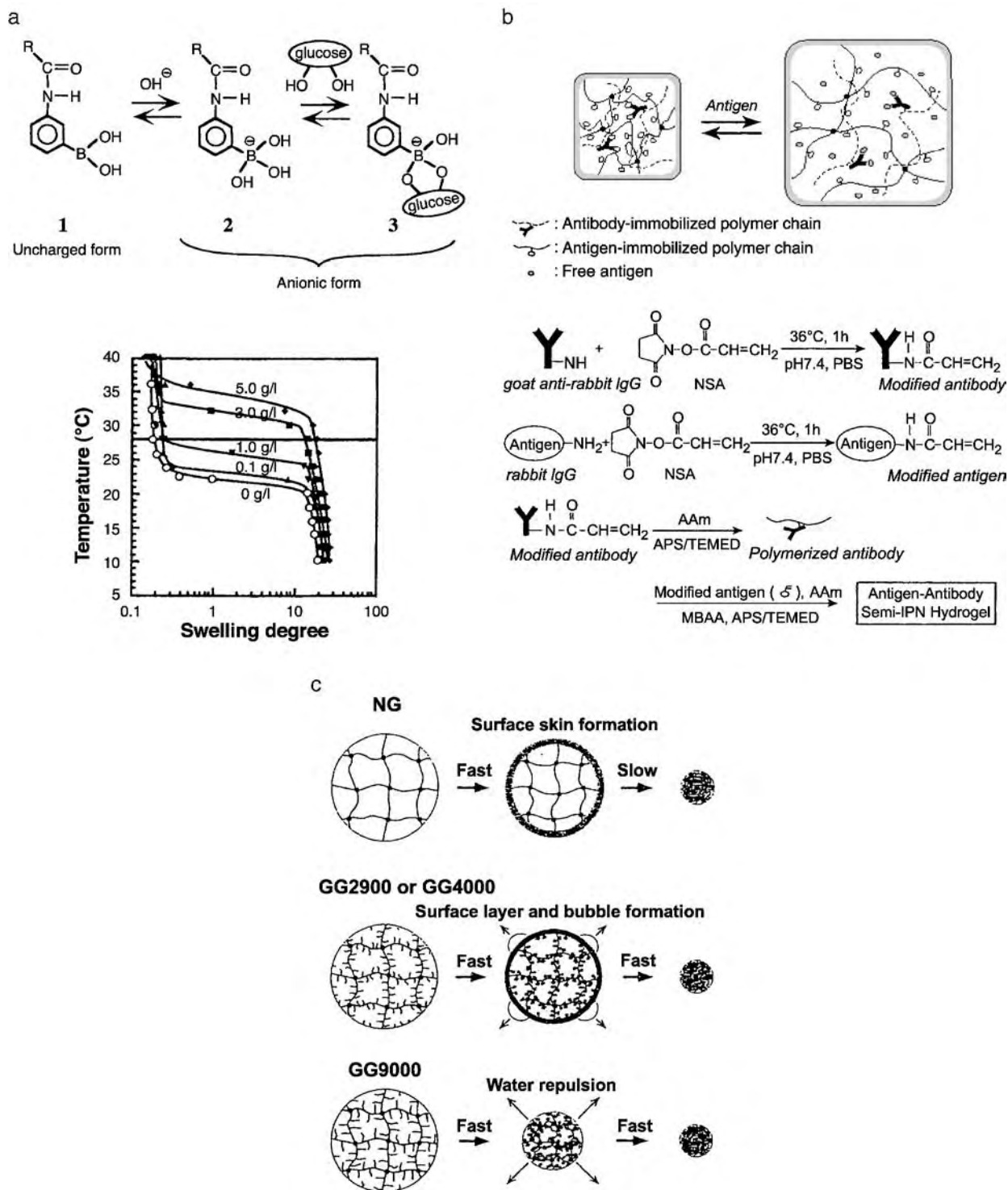


Fig. 3 (a) Schematic illustration of reversible change in the glucose bonding with (alkylamido)phenyl boronic acid according to the surroundings, and change in the swelling ratio of NB gel in response to the change in glucose concentration.

(Reprinted with permission from *J. Am. Chem. Soc.* **1998**, *120*, 12694–12695. Copyright 1998 American Chemical Society.)

(b) Diagram of a suggested mechanism for the swelling of an antigen–antibody semi-IPN hydrogel in response to a free antigen. Synthesis of the modified antigen and antibody monomers, containing rabbit IgG and GAR IgG as the antigen and antibody, respectively, followed the synthesis of semi-IPN hydrogel using those monomers. (Reprinted with permission from *Nature* **1999**, *399*, 766–769. Copyright 1999 Macmillan Publishers Ltd.) (c) Schematic illustration of deswelling mechanism of non-grafted gel (NG) and comb-type grafted poly(NIPAAm) gels (GG) having different lengths of grafted chains, with average molecular weights of 2900, 4000, and 9000, above their phase transition temperature. (Reprinted with permission from *Macromolecules* **1995**, *28*, 7717–7723. Copyright 1995 American Chemical Society.)

GAR IgG, an antigen-antibody semi-interpenetrated polymer network (semi-IPN) hydrogel was continuously prepared by the copolymerization of the vinyl(rabbit IgG), AAm, and some cross-linker. This semi-IPN type of hydrogel showed an abrupt increase in the swelling ratio on the addition of free rabbit IgG; however, it showed no reaction in the case of goat IgG, meaning that the hydrogel recognized a specific antigen. They suggest that the binding of the GAR IgG with rabbit IgG polymerized in the antigen-antibody hydrogel was much weaker than that to native rabbit IgG, due to denaturation of the modified rabbit IgG, causing the intra-chain grafted antigen-antibody binding to be dissociated by exchange of the grafted antigen for free antigen. Although the ratio of the volume change is small, by this unique antigen-specific type of hydrogel, it is possible to use the hydrogel to release a drug in response to the antigen as a result of the expression of cancer cells.

Quick response hydrogel

For this to be an effectively quick and sharp response to the stimulation, the volume phase transition of the hydrogel is necessary. However, the hydrogel volume change is dominated by the character of hydrogel; that is, the characteristic time of swelling is proportional to the square of the linear dimension of the hydrogel and is also proportional to the diffusion coefficient of the hydrogel network.^[13] Several research studies have been undertaken to overcome this limit. Forming pores in the hydrogel is one method to allow water to flow well. There are many ways to form a porous hydrogel—for example, preparing a hydrogel with a water-soluble porosigen, such as sodium chloride, sucrose, PEO, and so on, which can be removed after gelation. However, the pore size of the hydrogel depends on the size of the porosigens.

Park et al. successfully created a superporous hydrogel that quickly absorbed water and reached a high swelling equilibrium.^[14] They introduced many interconnected pores into the hydrogel that behaved as open channels to allow water flow through a capillary action. They utilized gas bubbles, CO₂, generated by mingling NaHCO₃ with acid, to form the pores in the hydrogel. The key point is how to maintain the bubbles during gelation, because the bubbles tend to fade away. They found the answer by mingling the pre-gel solution and a foam stabilizer, [poly(ethylene oxide)-poly(propylene oxide)-poly(ethylene oxide)] tri-block copolymers] (PEO-PPO-PEO) together to impart a high viscosity to the pre-gel solution, which helped to stabilize the foam during gelation. They obtained a superporous hydrogel showing up to 200 times faster absorption compared with an ordinary hydrogel.

Okano et al. have created a hydrogel system that quickly shrinks into tiny pieces with a change in temperature (Fig. 3c).^[15] They introduced hydrophobic chains into the matrix of poly(NIPAAm) to enhance the agglutinability, causing a hydrophobic interaction among the polymer chains. First, they made some poly(NIPAAm) macromonomers with different molecular weights by using telomerization of a NIPAAm monomer with 2-aminoethanethiol as a chain transfer agent and a continuous condensation reaction of amino semitelechelic poly(NIPAAm) with *N*-acryloxysuccinimide. They then synthesized several hydrogels by radical polymerization of the NIPAAm monomer and a NIPAAm macromonomer with a cross-linker, yielding a comb-type grafted polymer hydrogel in which the molecular architecture is different from that of the normal poly(NIPAAm) hydrogel, though the composition is the same. This comb-type grafted polymer gel could shrink much faster than the normal poly(NIPAAm) gel when the surrounding temperature climbs, and the rate of response depended on the length of the grafted polymer chains; that is, a hydrogel having a longer grafted polymer chain shrank faster.

Reentrant volume phase transition

Recently, a strange volume phase transition has been reported for the poly(MPC) hydrogel.^[6] Generally, the hydrogel swells in a good solvent and shrinks in a poor solvent. Water and ethanol are good solvents for poly(MPC); however, the poly(MPC) hydrogel shrank in certain water/ethanol mixtures. This phenomenon, in which the hydrogel once shrank and then reswelled according to the composition of the solvent, is called reentrant volume phase transition. The water or ethanol condition around polymer chains was altered by the introduction of ethanol or water, which triggered the phase change of the whole system.

The poly(NIPAAm) gel also showed this phenomenon in some water-miscible alcohol-aqueous solutions,^[16] but in the case of a poly(MPC) gel, the swelling and shrinking in response to the alcohol composition is dependent on the hydrophobicity of the alcohol. This means that the poly(MPC) gel has a molecular recognition property. By using this unique characteristic of the poly(MPC) gel and in combination with the functions of the hydrogel mentioned above, novel smart biomaterials with high biocompatibility can be realized.

PHYSICAL GELS

Physical gels, which are often called pseudo gels, are continuous, disordered, three-dimensional networks

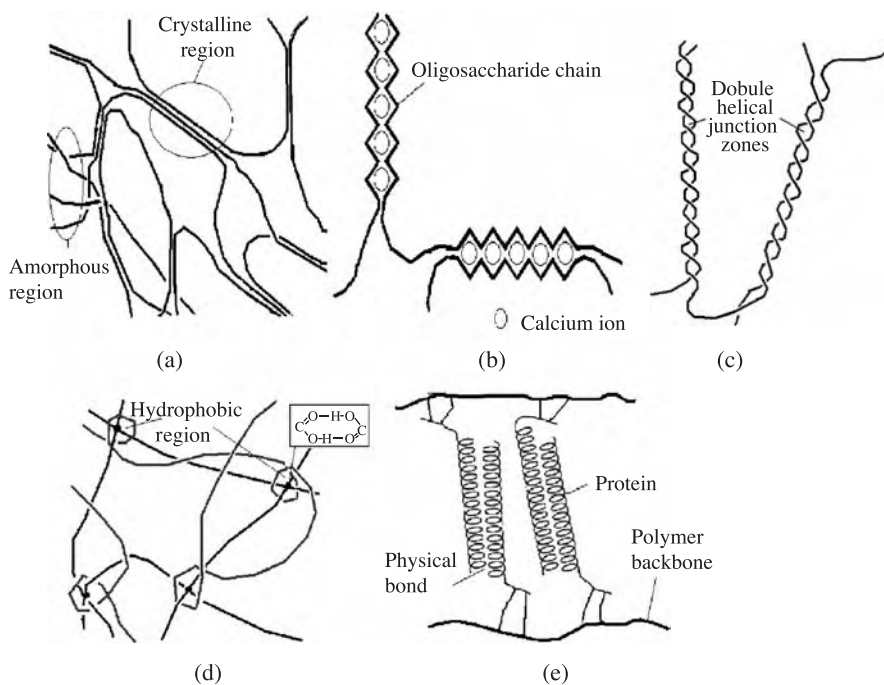


Fig. 4 Physical gels with multiple junctions. (a) Aggregation of the chains by crystallization. (b) Egg-box model created by specific ionic interaction. (c) Helical junction zone. (d) Hydrogen bond occurring in hydrophobic domain. (e) Hybrid hydrogel formation by coiled-coil proteins.

formed by associative forces capable of forming non-covalent cross-links.^[17] The noncovalent cross-links are formed by weaker interactions, and the cross-links are reversible. These interactions include hydrogen bonds, ionic interaction, hydrophobic association, stereo-complex formation, coiled-coil interaction, antigen-antibody interaction, crystalline segment cross-linking, and solvent complexation. These interactions would bring the polymer chains together and induce formation of a stable structure by forming a junction zone that would occur as shown in Fig. 4. These junction zones maintain the ordered structure inside the gels. A change in temperature, pH, or the addition of salts, for example, would form or modify the junction zones. The junction zones that had been formed between chain segments tended to stabilize as the segment number increased. Because the cross-linked junctions are not covalently bonded, breakdown of the network may occur due to a change in the environment (reversible phenomena).

Synthetic Polymers

Synthetic polymers that can be utilized for a hydrogel base material have several requirements. They need to be stimuli sensitive, but the most important requirement is biocompatibility, for they are destined to be used for biomaterials.

Poly(vinyl alcohol)

Linear poly(vinyl alcohol) (PVA) is a one of the most actively used synthetic polymers for physical

hydrogels. PVA is a polymer that has a thermo-reversible property. PVA hydrogel can be prepared by simply lowering the temperature. Due to its thermo-reversible property, the hydrogel will form a crystallized zone when the temperature is lowered. Cyclic freeze-thaw or blending a hydrophilic polymer such as poly(sodium L-glutamate) and poly(allylamine) HCl salt together with PVA increases the crystallinity and mechanical properties due to a complex between the PVA and the hydrophilic polymer.^[18]

Block copolymers

PEO, PPO, and poly(lactic acid) (PLA) are the materials most actively applied to biomedical and biotechnology because of their excellent physicochemical and biological properties. They are prepared in the form of block or triblock copolymers. Micelles are formed at low concentrations in aqueous solution, and thermoreversible gels are formed at high concentrations. This block copolymer hydrogel also shows LCST behavior, which allows it to be used for drug delivery. Multiblock copolymers of PEO and poly(butylene terephthalate) (PBT) can be prepared by melt condensation of PEO, butanediol, and dimethyl terephthalate. A phase-separated structure is formed with hard PBT domains, resulting in thermally reversible cross-linking, dispersed in the soft amorphous phase of hydrophilic PEO and amorphous PBT (Fig. 4a).^[19] The PEO-PPO-PEO triblock, also known as Pluronic, dissolves in cold water, forming a viscous solution. When the temperature is increased, the solution gels



due to phase transition as the polymer concentration is raised above a critical value. The increment in temperature leads to an increase in the micelle volume fraction (ϕ_m). When $\phi_m > 0.53$, the system becomes a gel by micelle packing (hard-sphere crystallization). A PLA–PEO–PLA block copolymer hydrogel can be prepared simply by immersing the polymer film or tablet in water. Phase separation of PEO and PLA induces the gelation. It has been reported that when the PEO to PLA ratio is strictly controlled, immersion into an organic solvent followed by the addition of water may produce a softer and more hydrophilic hydrogel.^[20]

Acrylamide polymers

One of the most commonly investigated acrylamide polymers for physical gelation is the linear homopolymer poly(NIPAAm). Poly(NIPAAm) is often formed as a gel by chemical cross-linking. However, its precipitation ability above 32°C due to LCST makes subject to the physical gelation process. Generally, an acrylamide polymer including poly(NIPAAm) becomes physically gelled when it is copolymerized with another polymer having a different property. Transparent aqueous polymer solutions of NIPAAm copolymers containing small amounts of acrylic acid are readily polymerized using free radical initiation in benzene and turn opaque as the temperature increases. When the temperature reaches 30–34°C, the solution turns into a gel. As more heat is added to the hydrogel, the gel collapses.^[21] Block and star copolymers of PEO and poly(NIPAAm) form aqueous solutions at low temperatures and transform to relatively strong elastic gels upon heating, multiple-arm copolymers appear to form gels via a physical cross-linking mechanism, and diblock copolymers gel due to a micellar aggregation mechanism.

Acrylic polymers

Acrylic polymers possess intra/interchain hydrogen bonds. Poly(acrylic acid) (PAA) and poly(methacrylic acid) (PMAA) form a complex with PEO. Swelling under oscillatory pH conditions reveals the dynamic sensitivity of poly(methacrylic acid-*graft*-ethylene glycol) gels and shows that network collapse (complexation) occurs more rapidly than network expansion (decomplexation). Hydrophobic interactions are very important for stabilizing the complexes. Poly(organophosphazene) is an organometallic polymer containing double-bonded phosphorus and nitrogen on the main chain with two organic side groups attached to the phosphorus atom. Polyphosphazene hydrogel, like other synthetic hydrogels, can be divided into two

major classes: neutral or non-ionic hydrogels and ionic hydrogels. Ionic hydrogels are formed by phosphazene polyelectrolytes, while non-ionic polyphosphazene hydrogels are based on water-soluble polymers ranging from ultimately hydrophilic, such as polyphosphazene containing glycosyl and glyceryl side groups, to less hydrophilic, including poly[bis(methylamino)phosphazene]. The difference in solubility is believed to be partially due to the exposure of skeletal hydrogen bonding to water.^[22]

Biomolecules

Oligosaccharides

Oligosaccharides are naturally occurring materials that can be modified for use as suitable biomaterials. Simple modification or a change in the environment may cause gelation phenomena.

Alginate. Alginate is one of most common oligosaccharides used for hydrogel synthesis. The cross-linking is carried out by injecting calcium ions with alignment of each sequence to form an array of cavities called an egg box. In this formation, carboxylate and oxygen atoms are organized into a cavity at room temperature and physiological pH (Fig. 4b). Monovalent cations as well as divalent cations such as magnesium ions do not induce gelation, whereas Ba^{2+} and Sr^{2+} ions induce very strong gelation. High concentrations of α -L-gluconic acid chain sequences lead to brittle gels, while low concentrations lead to more elastic gels.

Dextran. The reaction for the gelation of dextran is due to crystallization caused by the association of chains through hydrogen bonding, but gelation also occurs when potassium ions are added. Although dextran does not have ionic groups on the chain, the cage that is formed by six oxygen atoms from these repeating glucose units yields a cavity that is a perfect fit for potassium ions.^[23] However, instability in water makes it unsuitable for biomedical use.

Chitosan. Physical cross-linking has been used to prepare chitosan-based gels by the *O*- and *N*-acetylation of chitosan and the attachment of C_{10} -alkyl glycosides to chitosan. Acetylation yields either “rigidly solidified” or “fragile” gels with high- and low-molecular-weight chitosan, respectively, while the attachment of C_{10} -alkyl glycosides forms gels in acidic media at elevated temperature ($> 50^\circ\text{C}$).^[24]

Gellan and Pectin. Gellan can form either elastic or brittle gels, depending on cation concentration. Gellan

gels are soft and easily deformable below the critical calcium, concentration and, brittle above. Using a salt of ions other than calcium, such as Mg^{2+} or K^+ , may induce a relatively stronger gellan gel.

Pectins are divided into high-methoxyl (HM) pectins and low-methoxyl (LM) pectins according to the degree of esterification to pectin accomplished in the presence of methanol.^[17] In LM pectins, gel formation is influenced by the number and sequence of consecutive carboxyl groups and by the concentration of calcium. LM pectin + Ca^{2+} gels conform to the “egg box” model, while HM pectin gelation is governed by hydrogen bonds and hydrophobic interactions.

Agarose and Carrageenan. Agarose is soluble in water at temperatures above 65°C and, depending on the degree of hydroxyethyl substitution on its side chains, it forms a gel in the range of 17–40°C. Agarose gels are stable and do not swell at constant temperature or reliquefy until heated to 65°C. A commercially available agarose hydrogel is SeaPrep, a polysaccharide derived from red algae that melts at 50°C and gels at 7–18°C.

Carrageenan is gelled by salts such as KCl, but it can form a gel under salt-free conditions also. There are three types of carrageenan: kappa, iota, and lambda. Among these three, only the kappa type shows gel-forming characteristics in water. Potassium and calcium ions induce the gelation, while sodium ions do not. In the iota type, gelation occurs in the presence of calcium ions. At low temperature and high salt concentration, a double helical arrangement of the parallel polysaccharide chains can be formed (Fig. 4c). The temperature for gelation is about 25°C, and remelting occurs at 35°C.

Peptides, proteins, and glycoconjugates

Peptide, protein, and glycoconjugate hydrogels are formed by specific interactions called lock-and-key interactions, but in a different way from polysaccharides. The specific interaction may occur only when the molecules maintain three-dimensional structures such as coiled–coil interaction, or hydrogen-bonded β -sheets or strands.

Fibrin Gels. Fibrin plays an important role in natural wound healing and is used as an adhesive sealant in surgery. It is formed by the enzymatic polymerization of fibrinogen. Cross-linking of the fibrin oligomer can be executed by factor XIIIa. The main application of fibrin is as a scaffold matrix for tissue engineering. The fixation of the gel with poly(L-lysine) during the culturing period was introduced to solve the problems of gel shrinking and low stiffness. Inner tension in the gel

as a result of the tendency to shrink and the fixation may have led to the collagen synthesis.

Actin Gels. Globular actin monomer (G-actin) polymerizes to form an actin filament (F-actin), which forms a gel. Cross-linking of the actin network by α -actin can transform isotropic entangled networks into heterogeneous networks of microgels, whose macroscopic viscoelasticity is a complex of contributions from both dense and loose mesh networks.^[25] In the case of G-actin in vitro, the presence of a salt (K^+ , Na^+ , Ca^{2+} , Ba^{2+} , Li^+ , etc.) may activate the gelation process.

Lectin-Mediated Gels. Lehr elucidated the mechanisms of lectin-cell interactions and explored the pharmaceutical potential of lectin-modified drug carrier systems at epithelial biological barriers.^[26] The weak interaction between lectins and glycoconjugates can induce easy dissociation behavior at neutral pH, making lectin suitable for delivery systems.

Biomimetic Polymers

Biomimetic polymers are synthetic polymers that mimic biomolecules. The merit of this approach is that the advantages of synthetic polymers and biomolecules can be achieved at the same time.

Phospholipid polymers

A phospholipid is a substance that comprises the outline of the cell wall and is commonly mimicked by investigators. Its high permeability and biocompatibility have interested investigators for a long time. Ishihara et al. prepared MPC-based polymers, poly(MPC-*co*-methacrylic acid) (PMA) and poly(MPC-*co*-*n*-butyl methacrylate) (PMB) aqueous solution, to synthesize the hydrogel.^[27] The gelation occurs due to hydrogen bonds between the carboxyl groups of the PMA in the hydrophobic domain of the PMB (Fig. 4d). The formation can take place under either high or low aqueous conditions at room temperature. Thermoreversible and pH-reversible phenomena have also been observed, so that the phospholipid hydrogels can be utilized as peptide drug carriers.

Hybrid gels

Hybrid hydrogels refer to hydrogel systems that contain two or more components of distinct classes of molecules. The main purpose of a hybrid polymer is to create stimuli responsiveness to a change in the environment. Numerous research studies are underway

on covalently bonding a stimuli-sensitive polymer with double-stranded deoxyribonucleic acid (DNA). DNA forms double or triple strands with its complementary base pair by way of hydrogen bonding. This formation is reversible and is affected by a change in physical or chemical stimuli. Kopecek et al. designed hybrid gels composed of coiled-coil protein domains and a polymer backbone (Fig. 4e). The coiled-coil proteins were engineered by control of amino acid sequences. The coiled-coil formation is regulated by temperature, pH, and ionic strength.^[28]

Biomimetic/Synthetic Polymer Hybrid Gels

Biomimetic/synthetic polymer hybrid gels are good candidates for application in biomaterials and biomedical engineering. Most biomimetic polymers are capable of highly sophisticated functions. The hybrid technique is one of the promising approaches for further advancement in the area. In this section, recent biomimetic/synthetic polymer hybrid gels are discussed.

Bio-inspired/biodegradable polymer hybrid

A novel bio-inspired material based on the structure of the cell membrane has been proposed by Ishihara et al. They first succeeded with MPC having a phospholipid polar group.^[5] To fabricate the cell membrane surface, the MPC was copolymerized with *n*-butyl methacrylate (BMA) as a comonomer, and the resulting copolymer was coated onto a substrate. For a cell culture scaffold, further materials design would be

necessary for the MPC polymer. A new concept has been proposed based on a biodegradable polymer hybrid. An enantiomeric poly[L-(D-)lactic acid] [PL(D)LA] macromonomer has been newly synthesized and copolymerized with MPC and BMA.^[29] The resulting copolymers would spontaneously form the phospholipid polymer scaffold by stereocomplexation (Fig. 5a). The scaffold is a porous structure formed by a NaCl leaching technique. The enantiomeric PL(D)LA segment was utilized not only as a physical cross-link point but also as a protein adsorption domain in cell culture. One of the advantages was water intrusion into the phospholipid polymer scaffold, which was evaluated by the static contact angle of the water. The contact angle of the phospholipid polymer scaffold quickly decreased. This result suggests that water uptake was enhanced by the MPC unit. This is quite an important property for a porous scaffold to be used in three-dimensional tissue reconstruction.

Host/guest supramolecular gels

The field of supramolecular chemistry has been explored for constructing new molecular architectures such as molecular association by intermolecular forces and inclusion complexation. Supramolecular science is a new research area, and the study of supramolecular assemblies has progressed due to the study of biomimetic chemistry and molecular recognition in relation to such areas as enzymatic catalysis reactions. One of the major supramolecular structures is a polyrotaxane. Polyrotaxanes, in which many α -cyclodextrins (α -CDs) are threaded onto a PEO chain capped with bulky end

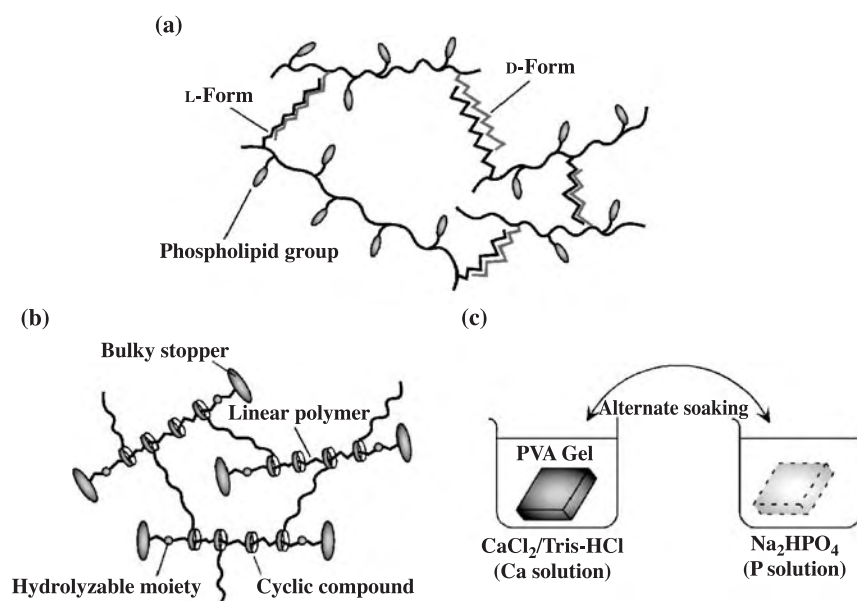


Fig. 5 Schematic illustrations of biomimetic/synthetic polymer hybrid gels. (a) Bioinspired/biodegradable polymer hybrid. (b) Host/guest supramolecular gels. (c) Organic/inorganic composite gels.

groups, have recently been demonstrated to be a molecular necklace. A biodegradable polyrotaxane consisting of α -CDs, PEG, and benzyloxycarbonyl (Z)-L-phenylalanine (L-Phe) introduced via ester linkages has been prepared as an implantable material for tissue engineering.^[30] The advantage of the polyrotaxane structure is its rapid and perfect degradation under physiological conditions. The specific structure of the polyrotaxanes may enable us to maintain the shape of the scaffold as a cell support and to achieve complete hydrolysis.

Using these perspectives, Yui et al. have designed PEG hydrogels cross-linked by hydrolyzable polyrotaxanes (Fig. 5b).^[31] In the hydrogel design, cyclic molecules in a polyrotaxane can be used as a cross-linking point. The cyclic molecules could be easily dethreaded onto the linear polymeric chain in response to terminal ester hydrolysis; then the cross-link density of the PEG hydrogel would decrease. The expected characteristics of the hydrogels include rapid and perfect disappearance based on supramolecular dissociation via terminal hydrolysis. PEG hydrogels cross-linked by the polyrotaxane were prepared and characterized to estimate gel erosion and identify the mechanism. It is assumed that terminal hydrolysis of the polyrotaxane cross-links can trigger dethreading of the α -CDs to induce a unique erosion behavior of the hydrogels.

Organic/inorganic composite gels

Mineralized tissues such as bone and dentin are indispensable elements in our body. For example, human bone consists of inorganic salts and organic substances such as collagen. The mineralized tissue has a well-organized structure, with inorganic salts, calcium, and phosphate. For this reason, organic/inorganic composites are good candidates for mineralized tissue reconstruction. In one technique, a hydroxyapatite composite has been produced by conventional flame spraying. The technique is limited to metal implants. In the case of organic substances such as polymeric gels, it is impossible to coat hydroxyapatite because of the heating process. A novel technique for preparation of organic/inorganic composite, especially hydroxyapatite, has been investigated. The hydroxyapatite was prepared by an alternative soaking process between $\text{CaCl}_2/\text{tris-HCl}$ (Ca solution) and Na_2HPO_4 (P solution) (Fig. 5c).^[32] Polymeric gels, such as PVA and poly(HEMA), were used as a template. The apatite crystal was formed not only on the surface but also in the gel through two subsequent steps.^[32] The apatite content and apatite formation rate were correlated with the gel swelling ratio. Three-dimensional gels/hydroxyapatite composites are of great importance in the fabrication of bonelike tissue.

CONCLUSION

This chapter summarizes the basic principles and recent topics concerning hydrogels as biomaterials. Hydrogels represent a diverse cross section of biomaterials, e.g., drug delivery carriers, tissue engineering scaffolds, and medical devices. Various chemical and physical properties are controlled by changing chemical reagents such as monomers and polymers and by the preparation conditions of the hydrogels. Hydrogels represent one of the most promising approaches for the fabrication of novel biomaterials.



ARTICLES OF FURTHER INTEREST

Blood-Material Interactions; Chitosan; Corneal Implants; Fibrin; Fibrin Sealants

REFERENCES

1. Wichterle, O.; Lim, D. Hydrophilic gels for biological use. *Nature* **1960**, *185*, 117–118.
2. *Contact Lenses*, 4th Ed.; Phillips, A.J., Speedwell, L., Eds.; Butterworth-Heinemann: Oxford, 1997.
3. Nicolson, P.C.; Vogt, J. Soft contact lens polymers; an evolution. *Biomaterials* **2001**, *22*, 3273–3283.
4. Künzler, J.; Ozark, R. Methacrylate-capped fluoro side chain siloxanes: Synthesis, characterization, and their use in the design of oxygen-permeable hydrogels. *J. Appl. Polym. Sci.* **1997**, *65*, 1081–1089.
5. Ishihara, K.; Ueda, T.; Nakabayashi, N. Preparation of phospholipid polymers and their properties as polymer hydrogel membrane. *Polym. J.* **1990**, *22* (5), 355–360.
6. Ishihara, K.; Nomura, H.; Mihara, T.; Kurita, K.; Iwasaki, Y.; Nakabayashi, N. Why do phospholipid polymers reduce protein adsorption? *J. Biomed. Mater. Res.* **1998**, *39* (3), 323–330.
7. Kiritoshi, Y.; Ishihara, K. Preparation of cross-linked biocompatible poly(2-methacryloyloxyethyl phosphorylcholine) gel and its strange swelling behavior in water/ethanol mixture. *J. Biomater. Sci., Polym. Ed.* **2002**, *13*, 213–224.
8. Andrews, C.S.; Denyer, S.P.; Hall, B.; Hanln, G.W.; Lloyd, A.W. A comparison of the use of an ATP-based bioluminescent assay and image analysis for the assessment of bacterial adhesion to standard HEMA and biomimetic soft contact lenses. *Biomaterials* **2001**, *22*, 3225–3233.
9. Tanaka, T. Collapse of gels and the critical endpoint. *Phys. Rev. Lett.* **1978**, *40*, 820.
10. Schild, H.G. Poly(N-Isopropylacrylamide): Experiment, theory and application. *Prog. Polym. Sci.* **1992**, *17*, 163–249.
11. Kataoka, K.; Miyazaki, H.; Bunya, M.; Okano, T.; Sakurai, Y. Totally synthetic polymer gels responding to external glucose concentration: Their preparation

- and application to on-off regulation of insulin release. *J. Am. Chem. Soc.* **1998**, *120*, 12694–12695.
12. Miyata, T.; Asami, N.; Uragami, T. A reversibly antigen-responsive hydrogel. *Nature* **1999**, *399*, 766–769.
 13. Tanaka, T.; Fillmore, D.J. Kinetics of swelling of gels. *J. Chem. Phys.* **1979**, *70*, 1214–1218.
 14. Chen, J.; Park, H.; Park, K. Synthesis of superporous hydrogel: Hydrogels with fast swelling and superabsorbent properties. *J. Biomed. Mater. Res.* **1999**, *44*, 53–62.
 15. Kaneko, Y.; Sakai, K.; Kikuchi, A.; Yoshida, R.; Sakurai, Y.; Okano, T. Influence of freely mobile grafted chain length on dynamic properties of comb-type grafted poly(N-isopropylacrylamide) hydrogel. *Macromolecules* **1995**, *28*, 7717–7723.
 16. Mukae, K.; Sakurai, M.; Sawamura, S.; Makino, K.; Kim, S.W.; Ueda, I.; Shirahama, K. Swelling of poly(N-isopropylacrylamide) gels in water-alcohol(C1–C4) mixed solvents. *J. Phys. Chem.* **1993**, *97*, 737–741.
 17. Park, K.; Shalaby, W.S.W.; Park, H. Physical Gels. In *Biodegradable Hydrogels for Drug Delivery*; Technomic Publishing Co.: Lancaster, PA, 1993; 99–140.
 18. Shaheen, S.M.; Yamaura, K. Preparation of theophylline hydrogels of atatic poly(vinyl alcohol)/NaCl/H₂O system for drug delivery system. *J. Control. Release* **2002**, *81*, 367–377.
 19. Bezemer, J.M.; Radersma, R.; Grijpma, D.W.; Dijkstra, P.J.; Feijen, J.; van Blitterswijk, C.A. Zero-order release of lysozyme from poly(ethylene glycol)/poly(butylene terephthalate) matrices. *J. Control. Release* **2000**, *64*, 179–192.
 20. Molina, I.; Li, S.; Martinez, M.B.; Vert, M. Protein release from physically crosslinked hydrogels of PMA/PEO/PLA triblock copolymer-type. *Biomaterials* **2000**, *22* (3), 363–369.
 21. Bae, Y.H.; Vernon, B.; Han, C.K.; Kim, S.W. Extracellular matrix for a rechargeable cell delivery system. *J. Control. Release* **1998**, *53*, 249–258.
 22. Andrianov, A.K.; Payne, L.G. Protein release from polyphosphagene matrices. *Adv. Drug Deliv. Rev.* **1998**, *31*, 185–196.
 23. Watanabe, T.; Ohtsuka, A.; Murase, N.; Barth, P.; Gersonde, K. NMR studies on water and polymer diffusion in dextran gels. Influence of potassium ions on microstructure formation and gelation mechanism. *Magn. Reson. Med.* **1996**, *35*, 697–705.
 24. Holme, K.R.; Hall, L.D. Chitosan derivatives bearing c10-alkyl glycoside branches: A temperature-induced gelling polysaccharide. *Macromolecules* **1991**, *24*, 3828–3833.
 25. Tempel, M.; Isenberg, G.; Sackmann, E. Temperature-induced sol gel transition and microgel formation in alpha-actinin cross-linked actin networks—A rheological study. *Phys. Rev. E.* **1996**, *54* (2), 1802–1810.
 26. Lehr, C.-M. Lectin-mediated drug delivery: The second generation of bioadhesives. *J. Control. Release* **2000**, *65*, 19–29.
 27. Nam, K.; Watanabe, J.; Ishihara, K. Characterization of the spontaneously forming hydrogels composed of water-soluble phospholipid polymers. *Biomacromolecules* **2002**, *3* (1), 100–105.
 28. Wang, C.; Steward, R.J.; Kopecek, J. Hybrid hydrogels assembled from synthetic polymers and coiled-coil protein domains. *Nature* **1999**, *397*, 417–420.
 29. Watanabe, J.; Eriguchi, T.; Ishihara, K. Stereocomplex formation by enantiomeric poly(lactic acid) graft-type phospholipid polymers for tissue engineering. *Biomacromolecules* **2002**, *3* (5), 1109–1114.
 30. Watanabe, J.; Ooya, T.; Yui, N. Effect of acetylation of biodegradable polyrotaxanes on its supramolecular dissociation via terminal ester hydrolysis. *J. Biomater. Sci., Polym. Ed.* **1999**, *10* (12), 1275–1288.
 31. Ichi, T.; Watanabe, J.; Ooya, T.; Yui, N. Controllable erosion time and profile in poly(ethylene glycol) hydrogels by supramolecular structure of hydrolyzable polyrotaxane. *Biomacromolecules* **2001**, *2* (1), 204–210.
 32. Taguchi, T.; Kishida, A.; Akashi, M. A study on hydroxyapatite formation on/in the hydroxyl groups-bearing nonionic hydrogels. *J. Biomater. Sci., Polym. Ed.* **1999**, *10* (1), 19–32.

Hydroxyapatite

Alexandra Elizabeth Porter

Department of Materials Science and Metallurgy, The Nanoscience Centre Cambridge, University of Cambridge, Cambridge, U.K.

Nelesh Patel

Serena Best

Department of Materials Science, University of Cambridge, Cambridge, U.K.

H

INTRODUCTION

Hydroxyapatite (HA) is utilized as a bone-grafting material owing to its similarity in composition with the inorganic component of hard tissues (bone and tooth mineral). Synthetic, stoichiometric HA has the formula $\text{Ca}_{10}(\text{PO}_4)_6(\text{OH})_2$ (Fig. 1) and belongs to a broad group of calcium-containing minerals based around the phosphate (PO_4^{3-}) group and is characterized by a calcium/phosphorus (Ca/P) molar ratio of 1.67. However, synthetic HA differs in composition from bone mineral, as, in addition to calcium, phosphorus, oxygen, and hydrogen ions, bone mineral contains a number of ionic substitutions in its lattice and can also be non-stoichiometric with $\text{Ca} \neq 1.67$.^[1]

HA is a bioactive material and, as such, encourages direct bone apposition when implanted into a skeletal site (Fig. 2). It is used clinically in granular, porous, and dense forms, as coatings on metallic implants^[2] and as a filler in composite materials.^[3–5]

This entry provides an overview of the structure and properties of HA and ionically substituted apatites and details the methods used to characterize the material chemically, physically, and biologically. The entry discusses the preparation and processing of HA and substituted apatites, for clinical application, and addresses the research that has been performed to characterize bone response to these materials.

HA PROPERTIES

Physical and Chemical Properties

During the 1980s, there was a great deal of interest in the physical and chemical properties of HA. Clearly the methods by which the materials are prepared will influence the values obtained, but the properties may broadly be summarized as shown in Tables 1 and 2.

Crystallography

HA possesses a hexagonal Bravais lattice and a $\text{P6}_3/m$ space group with cell dimensions of $a = 9.4211$ and $c = 6.8810 \text{ \AA}$.^[6] Fig. 1 shows a schematic diagram of the crystal structure of HA. Table 3 illustrates the atomic fractional coordinates and occupancy values of the atoms in the HA unit cell.^[7]

CHEMICALLY AND PHYSICALLY MODIFIED HA

One disadvantage of HA implants, in comparison with bioactive glasses and glass ceramics, is that the reactivity of HA with existing bone is relatively low^[8,9] and this has implications for the time required for patient recovery. One approach toward improving the osseointegration rates of HA has been to incorporate biological entities such as growth factors, proteins, and cells into the HA implant.^[10–12] As an alternative to modifying the biology of the implant, HA may also be chemically doped with small amounts (up to 20 mol%) of elements (e.g., carbonate and zinc) that are commonly found in physiological bone.^[1] These substitutions play a role in influencing the dissolution rate of apatites, which enhance the proliferation of human osteoblast-like cells in vitro^[13] and may favor osseointegration.^[14]

Substituted Apatites

The crystal structure of HA can accommodate substitutions by various other ions. Substitutions can occur for the Ca^{2+} , PO_4^{3-} , and OH^- groups present in HA. The ionic substitutions can affect the lattice parameters, crystal morphology, crystallinity, solubility, and thermal stability of HA. Table 4 summarizes the qualitative effects of some common ionic substitutions on the lattice parameters and crystallinity of HA.^[15]

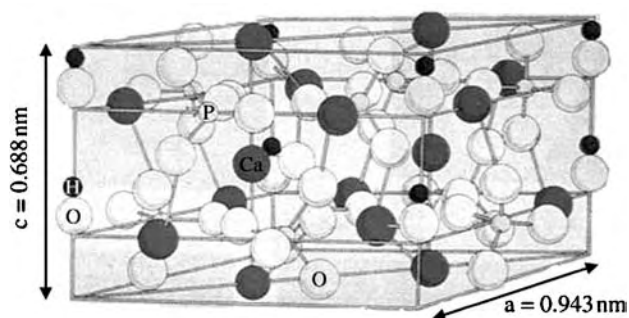


Fig. 1 The structure of HA. (Adapted from Ref. [7].)

Cationic substitutions

Cationic substitutions occur in the sites normally occupied by the calcium atoms. Elements such as sodium, magnesium, potassium, strontium, and manganese can be accommodated in the HA lattice, although, as highlighted in Table 4, these substitutions have not been observed to increase crystallinity. Imbalances in the charges of the substituting ion can cause disorder within the crystal structure of HA. The difference in valency caused by such a substitution requires a reduction in anionic charge to maintain charge balance.^[16]

Anionic substitutions

Fluorapatite and chlorapatite are common examples of anionically substituted HA. They display a similar structure to HA, but the F^- and Cl^- ions substitute for OH^- .

Carbonate Substitution. The most extensively studied synthetic substituted HA is carbonate HA (CHA), as

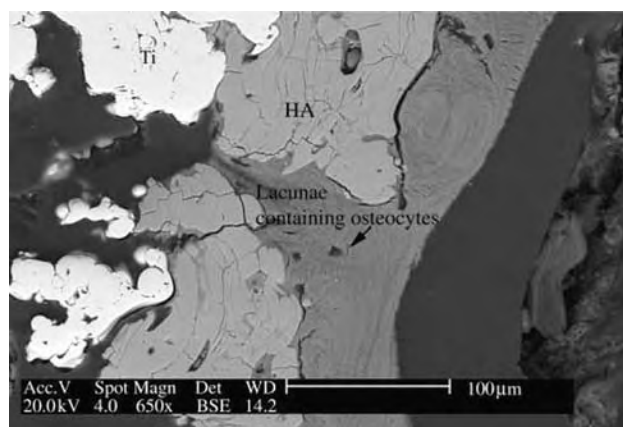


Fig. 2 Environmental scanning electron microscope image, illustrating bone growth directly along the surface of an HA-coated titanium hip replacement at 261 days in vivo. Image taken in backscattered electron mode. (From Ref. [35].)

Table 1 Chemical and physical data for HA

Property	
Lattice parameter (\AA)	$a = 9.4211$, $c = 6.8810$
Bravais lattice	Hexagonal
Space group	$P6_3/m$
Coefficient of thermal expansion ($10^{-6} \times \text{W/m/K}$)	13.3
Solubility product K_{sp}	2.34×10^{-59}
Density (g/cm^3)	3.156
Thermal conductivity (W/m/K)	0.72
Index of bioactivity, $I_B (100/t_{50bb})^a$	3.1

^a t_{50bb} is the time required for more than 50% of the interface to be bonded to bone.

(From Refs.[6,7] and Hench, L.L.; Wilson, J. *An Introduction to Bioceramics*. Singapore: World Scientific Press, 1993, 1–24; Cofino, B.; Fogarassy, P.; Millet, P.; Lodini, A. Thermal residual stresses near the interface between plasma-sprayed hydroxyapatite coating and titanium substrate: finite element analysis and synchrotron radiation measurements. *J. Biomed. Mater. Res.* **2004**, *70A* (1), 2–27; Markovic, M.; Fowler, B.O.; Tung, M.S.; Lagergren, E.S. Composition and solubility product of a synthetic calcium hydroxyapatite. Chemical and thermal determination of ca/p ratio and statistical analysis of chemical and solubility data. In *Mineral and Scale Formation*, Proc. of ACS Symposium; Z. Amjad, Ed.; Plenum: New York, 1995, 271–282.)

carbonate is the most abundant substitution in bone mineral (3–8 wt%).^[1,17,18] Carbonate ion substitution has been shown to increase rates of bone apposition around dense HA implants as compared to pure HA.^[19] The increased bioactivity of CHA has been attributed to be owing to the greater solubility of the CHA.^[20,21]

There are two types of carbonate substitution proposed in the literature, these being the substitution of CO_3^{2-} for OH^- (Type A)^[20] and CO_3^{2-} for PO_4^{3-} (Type B).^[22] In the Type A substitution, the exchange

Table 2 Mechanical properties of HA

Property	
Compressive strength	100–900 MPa
Tensile bend strength	40–120 MPa
Young's modulus	70–120 Gpa
Fracture toughness, K_{Ic}	1.0–1.5 $\text{Mpa m}^{1/2}$
Poisson ratio	0.23
Vickers hardness	4500–8000 MN/m^2

(From Aoki, H. *Science and Medical Applications of Hydroxyapatite*; Japanese Association of Apatite Science, 1991; Best, S.M. Preparation and Characterisation of Synthetic Hydroxyapatite. Ph.D. Thesis, University of London, 1990.)

Table 3 Occupancy, atomic fractional coordinates, and isotropic displacement parameters, B, for HA

Atom	Wyck. ^a	s.o.f ^b	x	y	z	B/10 ⁴ pm ²
O1	6h	1.000000	0.328280	0.484600	0.250000	1.399970
O2	6h	1.000000	0.587580	0.465200	0.250000	1.399970
O3	12i	1.000000	0.343300	0.257880	0.070500	1.399970
O4	4e	0.500000	0.000000	0.000000	0.200800	1.399970
P1	6h	1.000000	0.398200	0.368200	0.250000	2.500000
Ca1	4f	1.000000	0.333290	0.666700	0.001380	0.899990
Ca2	6h	1.000000	0.246600	0.993080	0.250000	0.899990
H1	4e	0.500000	0.000000	0.000000	0.061700	0.000000

^aWyck., Wyckoff crystallographic factors.

^bs.o.f., site occupancy factor.

(From Ref. [7].)

of the larger planar CO₃²⁻ group for the smaller linear OH⁻ group results in an expansion in the *a*-axis and a contraction in the *c*-axis and occurs when HA is treated with dry CO₂ at 1000°C.^[17] In the Type B substitution, CO₃²⁻ substitutes for PO₄³⁻, in HA,^[18] this causes a decrease in the length of the *a*-axis, while the *c*-axis increases as the carbonate content is raised.^[17] The change in lattice parameter has been attributed to the planar carbonate ion being smaller than the tetrahedral phosphate ion.

Studies have also highlighted that the two types of carbonate substitution, A- and B-type, can occur simultaneously resulting in a mixed AB-type CHA.^[18,24–26] This mode of substitution occurs almost exclusively by aqueous precipitation reactions and results in a more complex substitution mechanism. As for the B-type substitution, the mixed AB-type substitution involves the cosubstitution of sodium or ammonium ions from the reaction mixture to maintain charge balance. The calcium phosphate apatite that

Table 4 Qualitative effects of ionic substitutions in HA

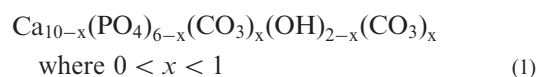
Substituting ions	Ionic radii (nm)	<i>a</i> -axis	<i>c</i> -axis	Crystallinity
HA [Ca ²⁺ ; PO ₄ ³⁻ ; OH ⁻]	[0.099; 0.154; 0.097]	0.943	0.688	
Strontium, Sr ²⁺ for Ca ²⁺	0.112	(+)	(+)	(nc)
Barium, Ba ²⁺ for Ca ²⁺	0.134	(+)	(+)	(-)
Lead, Pb ²⁺ for Ca ²⁺	0.120	(+)	(+)	(-)
Potassium, K ⁺ for Ca ²⁺	0.133	(nc)	(nc)	(nc)
Sodium, Na ⁺ for Ca ²⁺	0.097	(nc)	(nc)	(nc)
Lithium, Li ⁺ for Ca ²⁺	0.068	(nc)	(nc)	(nc)
Magnesium, Mg ²⁺ for Ca ²⁺	0.066	(-)	(-)	(-)
Cadmium, Cd ²⁺ for Ca ²⁺	0.097	(-)	(-)	(-)
Manganese, Mn ²⁺ for Ca ²⁺	0.080	(-)	(-)	(-)
Zinc, Zn ²⁺ for Ca ²⁺	0.074	(+)	(+)	(-)
Aluminium, Al ³⁺ for Ca ²⁺	0.051	(+)	(+)	(-)
Fluoride, F ⁻ for OH ⁻	0.136	(-)	(nc)	(+)
Chloride, Cl ⁻ for OH ⁻	0.181	(+)	(-)	(nc)
Carbonate, CO ₃ ²⁻ for PO ₄ ³⁻	0.131	(-)	(+)	(-)
Carbonate, CO ₃ ²⁻ for OH ⁻	0.131	(+)	(-)	(-)

(+), increase; (-), decrease, and (nc), no change.

(From Adapted from Ref. [15].)

constitutes bone mineral is considered to be of a mixed AB-type (predominantly B-type) substitution.^[27] It is, therefore, this more complex substitution type that is of most interest when developing a synthetic carbonate-substituted HA as a bone-substitute material.

Driessens, Verbeeck, and Kiekens^[28] produced evidence for a sodium-containing mixed AB-type CHA by an aqueous precipitation involving CaHPO₄ and CaCO₃ solutions. Na₂CO₃ was introduced into the reaction mixture as the source of carbonate ions. Carbonate contents of up to 15.4 wt% were determined by gravimetric analysis of CO₂ evolved from placing the samples in acidic aqueous solutions. Barrelet, Best, and Bonfield^[29] synthesized sodium-containing mixed AB-type CHAs over a wide range of reaction temperatures (3–90°) and carbonate concentrations (up to 640 mM sodium bicarbonate). Barrelet, Best, and Bonfield proposed an AB-type carbonate substitution mechanism [Eq. (1)], whereby the carbonate ions occupy both A and B sites in equal proportions.

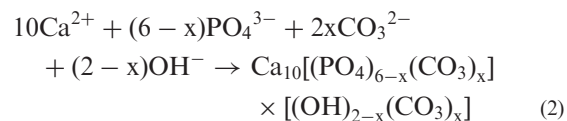


Vignoles et al.^[30] also reported that, where carbonate content is low, A-type carbonate substitution occurs, and that with increasing carbonate content, B-type sites are favored to the exclusion of all A-type substitution.

Although the chemical and structural properties of CHAs prepared by aqueous precipitation methods have been studied extensively, it is clear that as long as there is the cosubstitution of sodium or ammonium ions with the carbonate ions, the true role of only carbonate substitution on the chemical, physical, and biological properties of synthetic B-type (or AB-type) CHAs cannot be determined.

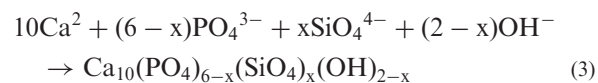
Recent studies by Gibson and Bonfield^[31] described a novel aqueous precipitation method to produce a pure, sodium and ammonium ion-free, AB-type CHA. The experimental process devised by Gibson and Bonfield involved two fundamental steps. Firstly, the Ca/P molar ratio of the reactants was deliberately calculated as being greater than the stoichiometric value of 1.67. By making the reaction mixture calcium rich, but in the presence of carbonate ions, an overall Ca/(P + CO₃) molar ratio could be maintained by the substitution of carbonate ions. The second aspect involved the introduction of carbonate ions into the reaction mixture. Instead of using an inorganic carbonate compound such as sodium carbonate or ammonium carbonate, carbonated water (carbonic acid) was used as the source of carbonate ions. As described in Eq. (2), Gibson and Bonfield^[31] were able to produce a mixed AB-type CHA

free of any additional cation substitutions, with carbonate contents of up to 4.1 wt%.



Silicate Substitution. Recent studies have focused on the development of silicate-substituted HA (SiHA).^[6,19,32] The rationale for these studies was the results of series of experiments by Carlisle,^[33] suggesting that silicon plays a critical role in the bone-calcification process. In vivo studies, comparing the rates of bone apposition to HA and SiHA ceramic implants, demonstrated a significant increase in amount of bone apposition and organization to around silicon-substituted HA (SiHA) implants, illustrating their potential as bone graft materials.^[19]

Several attempts have been made to prepare SiHA by a variety of synthesis routes. It is important that the ionic substitution of silicon does not result in the thermal instability of the SiHA, where sintering would result in the decomposition of the SiHA to undesirable second phases. Gibson, Best, and Bonfield^[6] produced phase-pure SiHA by an aqueous precipitation of a calcium-containing solution and a phosphate-containing solution at high pH, using silicon acetate as the source of silicate ions. They proposed the following substitution mechanism given in Eq. (3) below



Structural analysis has demonstrated that silicate (SiO₄⁴⁻) ions can substitute for the PO₄³⁻ sites in HA.^[6,19] The substitution of silicate ions in HA is feasible as the ionic bond lengths of a P–O bond (0.157 nm) with tetrahedral coordination are comparable to a similarly coordinated Si–O bond (0.166 nm).^[34] The effect of silicate substitution on the crystal structure is to cause a decrease in the *a*-axis and an increase in the *c*-axis of the unit cell of HA.^[6,35]

CHEMICAL SYNTHESIS

There are numerous methods for the preparation of synthetic apatites, which can be generally considered in three groups: aqueous reactions, solid-state reactions, and hydrothermal reactions. HA powders produced by aqueous reactions tend to be calcium deficient as a result of cationic substitution, which results in the formation of α -tricalcium phosphate (TCP) upon

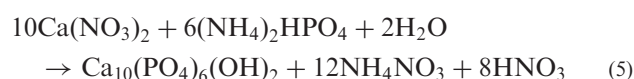
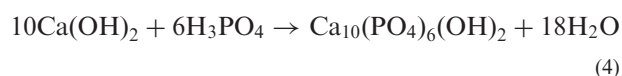
sintering and, if conditions are highly basic, carbonate substitution may occur.

Aqueous Techniques

These may be divided into two groups, precipitation and hydrolysis methods.

Precipitation

By far the most common method used to prepare HA is aqueous precipitation, not only because it is closer to the *in vivo* situation, but because it is a relatively simple procedure that can give a wide variety of apatite particle sizes and morphologies. The precipitation method involves mixing reactants together in the presence of water (usually at a controlled temperature, atmosphere, and pH 10 and leaving the resulting precipitate to age under conditions of continuous stirring for periods of up to 12 hr. If the pH is not well controlled, HA powders produced by precipitation methods often result in powders with deviations from stoichiometry (i.e., Ca/P \neq 1.67) resulting in additional secondary phases such as TCP or calcium oxide (CaO). Once aged, the precipitate is thoroughly washed, filtered, and dried. The two most common methods quoted in the literature are that of Akao, Aoki, and Kato,^[36] [Eq. (4)], and Hayek and Newesely,^[37] [Eq. (5)].



In the first method, [Eq. (4)], HA is prepared from the reaction of acid, H_3PO_4 , and alkali, $\text{Ca}(\text{OH})_2$ solutions. In the second method [Eq. (5)], aqueous solutions of calcium nitrate and ammonium phosphate are prepared (ammonium hydroxide may be added to control the solution pH). The phosphate solution is then added drop-wise to the stirred calcium solution. This method does have drawbacks in that other ions such as ammonium are necessary to attain the required pH and these ions are incorporated into the structure. There have been few investigations concerning the possible effect of these ions on mechanical properties or biocompatibility.

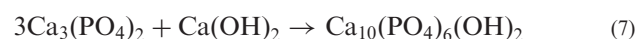
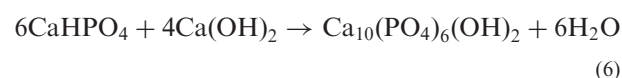
Hydrolysis

Hydrolysis methods involve the hydrolysis of acid calcium phosphates (e.g., dicalcium phosphate dihydrate) in carbonate, fluoride, chloride, ammonium,

sodium, or potassium hydroxide solutions depending on the desired composition of apatite.^[38–41] This method has drawbacks in that it introduces additional ions into the reaction procedure.

Solid-State Reactions

In solid-state reactions HA is produced via the stoichiometric mixing of reactants at specific temperatures. Examples of this method are given in Eqs. (6) and (7). The calcium compounds are mixed, formed, and then sintered above 950°C.^[42]



This procedure is prone to the production of materials that are inhomogeneous at the final processing stage. Also it is difficult to estimate the time required to produce a homogeneous end product.

Hydrothermal Transformations

The third method of synthesizing HA is that of hydrothermal transformation. This process employs the use of elevated temperatures, pressures, and controlled atmospheres to convert one substance into another via an exchange reaction. The reactions in Eqs. (5) and (6) could be carried out hydrothermally by heating at approximately 250–300°C under steam pressure of 80–100 MPa. Although this route is relatively time consuming, it has the added benefit of the option to preserve the original architecture, as illustrated by the production of porous HA from calcium carbonate corals.^[38]

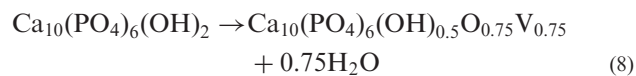
SINTERING HA

The most common method used to sinter HA is heating in air by means of an electric furnace at atmospheric pressure. Recent studies into the sintering behavior of HA have concentrated on the effects of temperature, time, and atmosphere on the levels of density and decomposition.

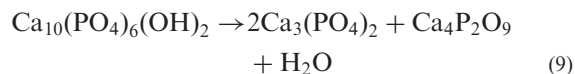
Stoichiometric HA can thermally decompose by two mechanisms. At relatively low sintering temperatures, dehydroxylation can occur, resulting in the formation of oxyHA (OHA) and water according to the reaction mechanism described by Eq. (8).^[43] At relatively high sintering temperatures, HA decomposes to TCP and



tetracalcium phosphate (TeCP) according to the mechanism described in Eq. (9).^[44]



where V = vacancy



The densification and phase purity of HA powders is dependent on heating rate and sintering dwell time. In addition to powder characteristics and sintering conditions, variations from stoichiometry in the starting apatite play a major role on the densification and decomposition process of HA. An excess of calcium (Ca/P > 1.67) will favor the formation of a CaO phase. In contrast, if the Ca/P molar ratio is < 1.67, then TCP phases are likely to form.

PREPARATION OF HA, CHA, AND SiHA CERAMICS FOR CLINICAL APPLICATION

HA has achieved application in a number of forms, including dense monolithic ceramics, porous scaffolds into which the host bone can regenerate and heal, granules (dense and porous), as coatings on metallic stems, fillers (either in fibrous^[45,46] or particulate form^[3]) in composites, and finally as calcium phosphate bone cements.^[47] As a ceramic material, HA exhibits relatively poor tensile strength and toughness and, therefore, clinical usage has mainly be concentrated on where the material would confer bioactivity to otherwise nearly inert implants or in situations where mechanical performance is not an issue. The following section describes several frequently employed processing techniques.

Dense Monolithic Ceramics

The conventional method of preparing dense ceramics involves three fundamental aspects. The first step involves the preparation and processing of powder with a suitable particle size for compaction. The second step involves the compaction of the powder to form a green body of desired size and shape and, finally, the green body is consolidated into a dense ceramic by sintering.

The density of ceramics is typically described in terms of a percentage of the maximum theoretical density, this being 3.156 g/cm³ in the case of HA. The process of sintering is unlikely to lead to a fully dense material—the final product will contain some microporosity.

Porous Scaffolds

HA scaffolds may be designed with inherent porosity, such that they mimic the architecture of natural trabecular bone. A number of techniques may be employed to fabricate porous HA scaffolds:

- “Burn-out.”
- Hydrothermal conversion methods.
- Sponge impregnation.
- Foaming methods.

All of these methods have their relative advantages and disadvantages, including control of pore size and interconnectivity and the requirement for the optimization of structural integrity of the struts between the pores. HA scaffolds with macroporosity greater than 100 μm have been utilized clinically for application such as spinal fusions, bone tumors, fractures, and in the replacement of failed or loose joint. More recently, there has been a trend toward engineering bone grafts with interconnected porosity between the macropores of the scaffolds. Pore interconnections with size 30–100 μm act as pathways between the macropores to favor cellular and vascular penetration, assuring improved bone ingrowth into the pores of the scaffold.^[47]

Granular Implants

An advantage in the use of granular implants compared to blocks or cylinders is that granules can fill voids of complex geometry, avoiding the need to cut and shape the implant during surgery. Furthermore, the morphology and the manner by which the granules can be “packed” into a defect site can be optimized to produce voids/spaces for bone ingrowth.

Biomedical Coatings

The philosophy behind adding a coating to a biomedical device is to improve either the wear properties or the biological properties of the implant. A number of different applications for ceramic coatings have been investigated and these have been used in dental implants and hip and knee replacement prostheses. Commercially, the most important production method for coatings is via plasma spraying.

One of the major innovations in the last 20 years has been to plasma spray the femoral stems of hip prostheses with HA (Fig. 3). The clinical results of HA-coated implants reveal that they have much longer life times after implantation than uncoated devices and they have been found to be particularly beneficial for younger patients.



Fig. 3 Coated femoral component of a hip prosthesis (Furlong).

Although the plasma-sprayed coatings are highly successful, they may have inherent flaws and are sometimes subject to problems of reproducibility in terms of crystallinity and mechanical properties. Other issues include the potentially poor adherence of the coating to the substrate, non-uniformity of the thickness of the deposited layer, and the integrity of the structure and composition of the coating. In addition, long-term *in vivo* studies have shown that the separation or loss of plasma-sprayed HA coating was progressive with time *in vivo*, which may cause cell-mediated osteolysis and subsequent implant loosening.^[48,49] Therefore, the need to make a more adherent, thinner, and uniform bioactive coating on metals with high mechanical strength is imperative. For these reasons, a number of other sputtering techniques such as magnetron sputtering^[50–52] and “low temperature” techniques are also being investigated, including electrophoresis and bone-like apatite coatings through simulated body-fluid treatments. Apatite coatings may be produced by soaking the metal stem in simulated body fluid (SBF) containing physiologically relevant ions (NaCl, KCl, $K_2HPO_4 \cdot 3H_2O$, $MgCl_2 \cdot H_2O$, CaCl₂, and Na₂SO₄) in

distilled water at pH 7.4.^[53] Apatite layer-coated titanium bone-bonding implants have been shown to bond firmly to the substrate and bone,^[54] with significant improvement over non-coated metallic stems. SBF-coated implants therefore offer an attractive alternative to plasma-sprayed implants for orthopaedic applications.

Fillers in Composites

One approach toward improved mechanical performance of HA ceramics is to design composite materials comprising both the ceramic HA phase and a polymeric phase.^[3–5] Composites of polymer and ceramic can confer favorable mechanical properties, including strength via the ceramic phase, toughness and plasticity via the polymer phase, and graded mechanical stiffness. An area of interest in recent years has been the development of calcium phosphate/collagen biocomposites. Because of their compositional similarity to natural bone, the main focus of their application has been the repair of bone defects. A number of initiatives have sought to develop HA/collagen biocomposites with a high degree of structural similarity to mature bone,^[4,5,54] with some of these producing materials whose resemblance to bone extends even to the nanometer scale. Such biomimetic composites have been produced by precipitating HA on to the collagen fibril template under alkaline conditions^[55] Future research effort will now aim to use a bottom-up approach to develop macroscale biomimetic scaffolds.^[56]

Calcium Phosphate Bone Cements

In the 1980s, the idea of a new bone substitute material was introduced, and the materials were referred to as calcium phosphate bone cements. These materials offer the potential for *in situ* moulding and injectability. There are a variety of different combinations of calcium compounds (e.g., α -TCP, and dicalcium phosphate), which are used in the formulation of these bone cements, but the end product is normally based on a calcium-deficient HA.^[57,58]

CHARACTERIZATION OF HA CERAMIC IMPLANTS

The characterization of HA implants has typically involved chemical, physical, and biological techniques. The results obtained from such studies are inter-related, and to understand the mechanisms underlying the bioactivity of HA ceramics, it is important to perform comprehensive studies that incorporate chemical, physical, and biological assessment. The aim



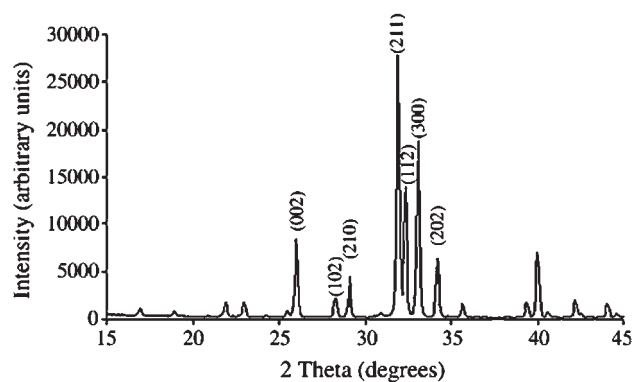


Fig. 4 XRD trace of phase pure HA sintered at 1200°C. (From Ref. [45].)

of this section is to provide a general overview into some of the techniques used to characterize HA, CHA, and SiHA ceramics.

Chemical Characterization

Chemical characterization involves the assessment of crystal structure, phase purity, chemical content, and the type and amount of ionic substitution present within the sample. The methods commonly used to characterize HA ceramics include X-ray diffraction (XRD), X-ray fluorescence spectroscopy (XRF), inductively coupled plasma spectroscopy (ICP), C–H–N elemental analysis and infrared spectroscopy.

XRD

XRD is a versatile technique that can be used to determine the lattice parameters within a crystal structure and thus identify the phase or phases present within a material. XRD provides information about the phases rather than the combination of elements and can be used to assess the phase purity of a material. Fig. 4 shows typical XRD traces for sintered HA.

The scan is comparable to the International Centre for Diffraction Data (ICDD) standard diffraction pattern for HA and was performed over a 2θ range of 25–40°, with a step size of 0.02° and a dwell time of 5 sec. This 2θ range is typically selected because it is in this region that the major HA peaks and the main impurity peaks are present. In addition to phase analysis, XRD can also be used to qualitatively assess the degree of crystallinity of a sample as indicated by the sharpness of the peaks in the XRD pattern. Similarly, stresses within a crystal lattice can be evaluated by peak broadening and peak shifting.

ICP, XRF, and carbon hydrogen nitrogen elemental analysis

ICP and XRF are commonly used to determine the levels of impurities and the levels of ionic substituents that may be present within a sample. ICP can measure most elements, but is best suited to measuring low concentrations, such as trace elements, or significantly diluted bulk compositions such as fluorine and chlorine in HA. XRF is also used to measure most elements ($Z > 10$), but is best suited for element concentrations > 0.1 wt% or 1000 ppm, such as the levels of elements that can form oxides at high temperatures.

Carbon hydrogen nitrogen elemental analysis is a useful technique to determine the levels of carbonate within a sample. This technique is based on an analysis of the gases that result from the combustion of a sample. The levels of carbon, hydrogen, and nitrogen can be determined by the amount of CO_2 , H_2O , and NO_x that pass through three pairs of thermal conductivity detectors.

Vibrational spectroscopy

Infrared spectroscopy is a non-destructive technique that enables the presence of certain bonds in a material to be established. Infrared spectroscopy is particularly useful in determining the presence of carbonate groups and their probable locations in the lattice of carbonate-substituted HA. There are two types of vibrational spectroscopy techniques: Fourier transform infrared (FTIR) spectroscopy (Fig. 5) (Table 5) and Raman spectroscopy (Fig. 6) (Table 6). FTIR is concerned with the vibrations of polar bonds where a change in dipole moment occurs as a result in vibration. Raman spectroscopy is associated with vibrations of non-polar covalent bonds where polarization can be induced.

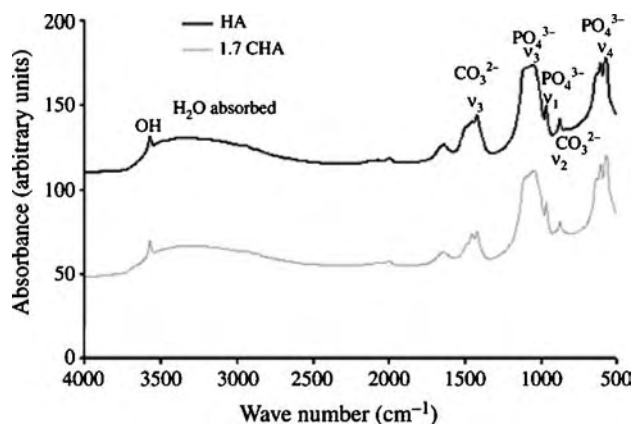


Fig. 5 FTIR spectra of sintered HA and 1.7 wt% CHA.

Table 5 FTIR analysis of sintered HA and CHA samples

	OH (stretch)	CO ₃ ν_3	PO ₄ ν_3	PO ₄ ν_1	CO ₃ ν_2	CO ₃ ν_4	OH (vibration)	PO ₄ ν_4	PO ₄ ν_2
HA	3572	—	1020, 1068, 1086	964	—	—	648	590, 617	478
1.70CHA	3571	1419, 1473, 1508, 1550, 1570	1018, 1083, 1145	968	883	759	648	567, 590	451, 474

All values correspond to the wave numbers (/cm) of the different hydroxyl, carbonate, and phosphate bands (unpublished data).

FTIR spectroscopy

FTIR spectroscopy allows the identification of the different functional groups present within the apatite structure of the samples, i.e., the phosphate, carbonate, and hydroxyl groups. The analysis technique is non-destructive and produces peaks/bands according to the way in which the bonds within functional groups vibrate in response to infrared radiation at specific characteristic frequencies. A common method of determining the type of carbonate substitution in HA is by FTIR. In addition to phosphate and hydroxyl bands, the FTIR spectra of CHA consist of additional peaks because of carbonate ions. The vibrations owing to carbonate ions may be either stretching or bending types. The carbonate ion is a trigonal planar ion that has four normal modes of vibration of which ν_2 (out of plane bending), ν_3 (asymmetric stretching), and ν_4 (planar bending) are infrared active and the symmetric ν_1 is not infrared active. A typical FTIR trace for HA is shown in Fig. 5. Both carbonate ion and silicate ion substitution affect the infrared spectral properties of HA. A-type carbonate substitution is characterized by the presence of a ν_2 peak at 878/cm and ν_3 peaks at 1450/cm and 1550/cm. B-type carbonate substitution is

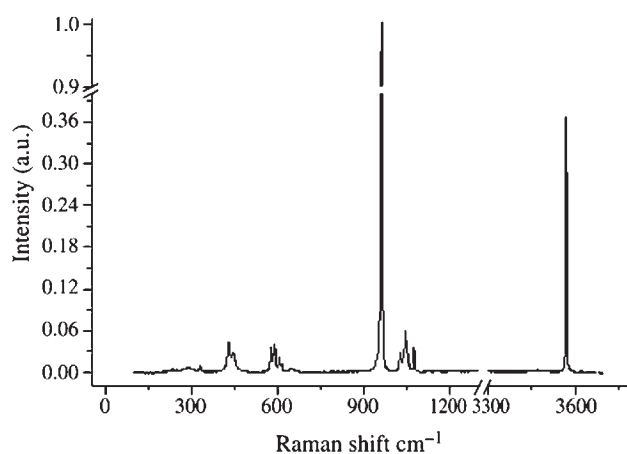


Fig. 6 Raman trace for pure HA sintered at 1200°C for 2 hr.

characterized by the presence of a ν_2 peak at 875/cm and ν_3 peaks at 1410/cm and 1460/cm.^[59]

Physical Characterization

Physical properties such as micro- and macrostructure have been reported to influence the bioactivity of HA implants. For example, the bioactivity of porous HA is dependent on the geometry and size of the pores.^[60] The microstructure (in particular the grain size) may affect the solubility and the mechanical properties of HA in vivo.

BONE RESPONSE TO SYNTHETIC HA

Osseointegration may be defined as the ability of a material to support bone growth directly along its surface when placed in the vicinity of bone or differentiated bone-forming cells (as highlighted previously, in Fig. 2). A wide range of animal models have been used to demonstrate the efficacy of HA implants in vivo. Several studies have investigated the HA/host-tissue interface in vivo, and the majority of these reported that bone “bonds” directly on the surface of HA implants. Although the true definition of bone “bonding” to HA implants has not been resolved, the term has generally been applied where bone directly apposes an implant surface without the formation of an interfering fibrous layer. Considerable efforts have been made to unravel the mechanisms of material and biological responses that underlie bone-bonding capacity. It is commonly believed that the mechanism of bone bonding to HA implants may be explained by the proposed “dissolution–reprecipitation” theory for apatite nucleation. Dissolution of calcium and phosphate ions from the outermost surface of the HA implants leads to a local increase in concentration of these ions that favors biological (carbonated) apatite precipitation heterogeneously on proteins in the vicinity of the coating or directly on the coating surface itself (Figs. 7 and 8).^[8,61–64,65–72] This process may



Table 6 Raman analysis of sintered HA^a

		PO ₄ ν ₂	PO ₄ ν ₄	OH ⁻ vibrational	PO ₄ ν ₁	PO ₄ ν ₃	OH ⁻ stretch
	OH ⁻ translational	Wave number (/cm)					
HA	329	431	579	630	962	1028.5	3573
		447	590			1040	
			607			1047	
			614			1052	
						1076	

^aUnpublished data.

occur through the surrounding environment (extra-cellular fluid) and may also be actively mediated by osteoclast cells^[73] Subsequently, this modified surface rapidly absorbs more proteins and promotes cell adhesion, particularly osteoblasts.

Studies have reported that the degradation processes of calcium phosphates are primarily influenced by chemical composition,^[8,35,63,65,66] crystallinity,^[67,68] porosity^[47] and microstructure of the materials.^[35,69,70] As Ca:P ratio decreases, both the solubility and the speed of hydrolysis of calcium phosphates increase. The rate of biodegradation of TCPs and HA between pH 6.2 and 7.2 follow the order^[70]

TCP > TeCP ≫ HA

Daculsi, LeGeros, and Mitre^[14] reported that the dissolution of calcium and phosphate ions from HA ceramics was initiated at microstructural defects

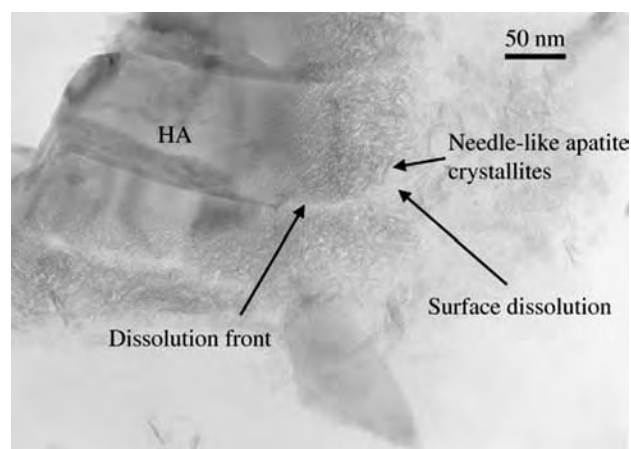


Fig. 7 TEM micrograph of grains of 0.8 wt% SiHA within the implant at six weeks in vivo. *Black arrow* shows surface dissolution of the grains. (From Ref. [73].)

such as dislocations and grain boundaries. Electron microscopy studies have reported that ionic substitutions modify the series of events occurring at the interface between the HA and the surrounding bone by altering the grain boundary structure of the ceramic.^[71–73] For, example, it has been suggested that substitution of silicate ions into the HA lattice modifies the sintering properties of HA and decreases the grain size (Table 2).^[71] The smaller grain size and formation of a less-ordered grain boundary phase in the SiHA increase the solubility and may affect both the morphology and the timing of apatite deposition at the bone–implant interface^[72,73] Despite the preponderance of evidence supporting the improved performance of physically and chemically modified HA implants as compared to HA, the mechanism by these modifications increases the bioactivity of HA remains unresolved.

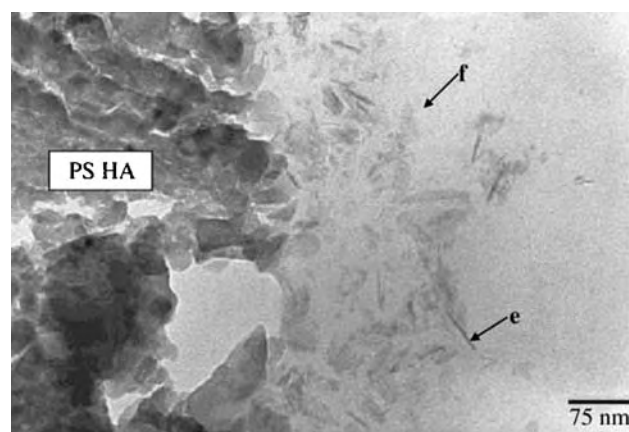


Fig. 8 Transmission electron micrograph of the interface between as-deposited PSHA coating and embedded biological material opposing the coating at 3 hrs in vivo. Showing plate-like apatite crystallites viewed en face (f) and on edge (e). (From Ref. [48].)

CONCLUSIONS

Synthetic HA offers the potential for bone repair and regeneration, and has been in clinical use for a number of years. The future for the material will be in the development of new compositions either with enhanced bioactivity, or the ability to be resorbed in vivo. Alongside, the development of more chemically and biochemically active materials will be the continued search for improvements in the design of structures suitable for use as tissue-engineering scaffolds.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the Engineering and Physical Sciences Research Council for funding the research on HA and substituted apatites.

ARTICLES OF FURTHER INTEREST

Bioactive Glass; Biologic and Synthetic Apatites; Biomimetic Materials; Tissue Engineering of Bone

REFERENCES

- Driessens, F.C.M. The mineral in bone, dentin and tooth enamel. *Bull. Soc. Chim. Belg.* **1980**, *89*, 663.
- de Groot, K.; Geesink, R.; Klein, C.P. Plasma sprayed coatings of hydroxyapatite. *J. Biomed. Mater. Res.* **1984**, *21*, 1375–1387.
- Bonfield, W.; Grynblas, M.D.; Tully, A.E.; Bowman, J.; Abram, J. Hydroxyapatite reinforced polyethylene—a mechanically compatible implant. *Biomaterials* **1981**, *2*, 185–186.
- Song, J.; Saiz, E.; Bertozzie, C.R. A new approach to mineralization of biocompatible hydrogel scaffolds: an efficient process toward 3-dimensional bonelike composites. *J. Am. Chem. Soc.* **2003**, *125* (5), 1236–1243.
- Lynn, A.K.; Nakamura, T.; Patel, N.; Porter, A.E.; Renouf, A.C.; Laity, P.R.; Best, S.M.; Cameron, R.E.; Shimizu, Y.; Bonfield, W. Composition-controlled nanocomposites of apatite and collagen incorporating silicon as an osseopromotive agent. *J. Biomed. Mater. Res. A.* **2005**, *74* (3), 447–453.
- Gibson, I.R.; Best, S.M.; Bonfield, W. Chemical characterization of silicon-substituted hydroxyapatite. *J. Biomed. Mater. Res.* **1999**, *44*, 422–428.
- Kay, H.A.; Young, R.A.; Posner, A.S. The crystal structure of hydroxyapatite. *Nature* **1964**, *204*, 1050–1052.
- Ducheyne, P.; Radin, S.; King, L. The effect of calcium-phosphate ceramic composition and structure on in vitro behaviour. I. Dissolution. *J. Biomed. Mater. Res.* **1993**, *27*, 25–34.
- Oonishi, H.; Hench, L.L.; Wilson, J.; Suginara, F.; Tsuji, E.; Kushitani, S.; Iwaki, H. Comparative bone growth behavior in granules of bioceramic materials of various sizes. *J. Biomed. Mater. Res.* **1999**, *44*, 31–43.
- Mankani, M.H.; Kuznetsov, S.A.; Shannon, B.; Nalla, R.K.; Ritchie, R.O.; Qin, Y.; Robey, P.G. Canine cranial reconstruction using autologous bone marrow stromal cells. *Am. J. Pathol.* **2005**. *In press*.
- Mastrogiacomo, M.; Muraglia, A.; Komlev, V.; Peyrin, F.; Rustichelli, F.; Crovace, A.; Cancedda, R. Tissue engineering of bone: search for a better scaffold. *Orthod Craniofac. Res.* **2005**, *8* (4), 277–284.
- Hing, K.A. Bioceramic bone graft substitutes: influence of porosity and chemistry. *Int. J. Appl. Ceram. Technol.* **2005**, *2* (3), 184–199.
- Gibson, I.R.; Hing, K.A.; Best, S.M.; Bonfield, W. Enhanced in vitro cell activity and surface apatite layer formation on novel silicon-substituted hydroxyapatites. 12th International Symposium on Ceramics in Medicine; Ohgushi, H., Hastings, G.W., Yoshikawa, T., Eds.; Nara, Japan, 1999; 191–194.
- Daculsi, G.; LeGeros, R.Z.; Mitre, D. Crystal dissolution of biological and ceramic apatites. *Calcif. Tissue Int.* **1989**, *45*, 95–103.
- LeGeros, R.Z. Apatites in biological systems. *Prog. Cryst. Growth Char.* **1981**, *4*, 1–45.
- de Maeyer, E.A.P.; Verbeeck, R.M.H. Possible substitution mechanisms for sodium and carbonate in calciumhydroxyapatite. *Bull. Soc. Chim. Belg.* **1993**, *102*, 601–609.
- LeGeros, R.Z. Effect of carbonate on the lattice parameters of apatite. *Nature* **1965**, *205*, 403–404.
- Nelson, D.G.A.; Featherstone, J.D.B. Preparation analysis and characterization of carbonated apatites. *Calcif. Tissue Int.* **1982**, *34*, 569–581.
- Patel, N.; Gibson, I.R.; Hing, K.A.; Best, S.M.; Revell, P.A.; Bonfield, W. A comparative study on the in vivo behaviour of hydroxyapatite and silicon substituted hydroxyapatite granules. *J. Mater. Sci. Mater. Med.* **2002**, *13*, 1199–1206.
- Porter, A.E.; Patel, N.; Brooks, R.; Best, S.; Rushton, N.; Bonfield, W. Effect of carbonate-substitution on the ultrastructural characteristics of hydroxyapatite implants. *J. Mater. Sci. Mater. Med.* **2005**, *16*, 1–9.
- Doi, Y.; Shibutani, T.; Moriwake, Y.; Kajimoto, T.; Iwayama, Y. Sintered carbonate apatites as bioresorbable bone substitutes. *J. Biomed. Mater. Res.* **1997**, *39*, 603–610.
- Bonel, G. Contribution a l'étude de la Carbonation des Apatites carbonatées du type A. *Ann. Chim.* **1972**, *47*, 65–88.
- Borneman-Staronkevich, I.V. On some isomorphous substitutions in apatite. *Compt. Recd. Acad. Sci. URSS* **1939**, *22*, 113.
- Doi, K.; Moriwaki, Y.; Okazaki, M.; Takashi, J.; Joshin, J. Carbonate apatites from aqueous and non-aqueous media studied by E.S.R., I.R. and X-Ray diffraction: effect of NH_4^+ ions on crystallographic parameters. *J. Dental Res.* **1982**, *61*, 429–434.



25. Barralet, J. Processing and sintering of carbonate hydroxyapatite. Ph.D. Thesis. 1995, University of London.
26. Driessens, F.C.M.; Verbeeck, R.M.H.; Heijligers, H.J.M. Some physical properties of Na- and CO₃-containing apatites synthesized at high temperatures. *Inorganica. Chemica. Acta* **1983**, *80*, 19–23.
27. Emerson, W.H.; Fischer, E.D. The infrared absorption spectra of carbonate in calcified tissue. *Arch. Oral Biol.* **1962**, *7*, 671–683.
28. Driessens, F.C.M.; Verbeeck, R.M.H.; Kiekens, P. Mechanism of substitution in carbonated apatites. *Zeitschrift fur Anorganische und allgemeine Chemie*, **1983**, *504*, 195–200.
29. Barralet, J.; Best, S.M.; Bonfield, W. Carbonate substitution in precipitated hydroxyapatite: an investigation into the effects of reaction temperature and bicarbonate ion concentration. *J. Biomed. Mater. Res.* **1998**, *41*, 79–86.
30. Vignoles, M.; Bonel, G.; Holcomb, D.W.; Young, R.A. Influence of preparation conditions on the composition of type b carbonated hydroxyapatite and on the localization of the carbonate ions. *Calcif. Tissue. Int.* **1998**, *43*, 3–40.
31. Gibson, I.R.; Bonfield, W. Novel synthesis and characterization of an AB-Type carbonate-substituted hydroxyapatite. *J. Biomed. Mater. Res.* **2002**, *59*, 697–708.
32. Kim, S.R.; Lee, J.H.; Riu, D.H.; Jung, S.J.; Lee, Y.J.; Chung, S.C.; Kim, Y.H. Synthesis of Si, Mg substituted hydroxyapatites and their sintering behaviors. *Biomaterials* **2003**, *24*, 1389–1398.
33. Carlisle, E.M. Silicon: a possible factor in bone calcification. *Science* **1970**, *167*, 279–280.
34. Shannon, R.D.; Prewitt, C.T. Effective ionic radii in oxides and fluorides. *Acta Cryst* **1969**, *B25*.
35. Porter, A.E.; Best, S.M.; Bonfield, W. Comparison of the ultrastructure of hydroxyapatite and silicon-substituted hydroxyapatite for biomedical applications. *J. Biomed. Mater. Res.* **2003**, *68A* (1), 133–141.
36. Akao, M.; Aoki, H.; Kato, K. Mechanical properties of sintered hydroxyapatite for prosthetic applications. *J. Mater. Sci.* **1981**, *28*, 809.
37. Hayek, E.; Newesely, H. Pentacalcium. Monohydroxyorthophosphate. *Inorganic Syst.* **1963**, *7*, 121–128.
38. LeGeros, R.Z.; LeGeros. Dense Hydroxyapatite. In *Introduction to Bioceramics*; Hench, L.L., Wilson, J., Eds.; World Scientific: Singapore, 1993; 139–180.
39. Monma, H.; Ueno, S.; Kanazawa, T. Properties of hydroxyapatite prepared by the hydrolysis of tricalcium phosphate. *J. Chem. Tech. Biotechnol.* **1981**, *31*, 15–24.
40. Rootare, H.M.; Craig, R.G. Characterization of hydroxyapatite powders and compacts at room temperature and after sintering at 1200°C. *J. Oral Rehab.* **1978**, *5*, 293–307.
41. Fang, Y.; Agrawal, D.K.; Roy, D.M.; Roy, R. Fabrication of porous hydroxyapatite ceramics by microwave. *J. Biomed. Mater. Res.* **1981**, *7*, 490–494.
42. Trombe, J.C.; Montel, G. Some features of the incorporation of oxygen in different oxidation states in the apatite lattice—I. On the existence of calcium and strontium oxyapatites. *Inorg. Nucl. Chem.* **1978**, *40*, 15–21.
43. Wang, P.E.; Chaki, T.K. Sintering behaviour and mechanical properties of hydroxyapatite and dicalcium phosphate. *J. Mater. Sci. Mater. Med.* **1993**, *4*, 50–158.
44. Aizawa, M.; Shinoda, H.; Uchida, H.; Itatani, K.; Okada, I.; Matsumoto, M.; Morisue, H.; Matsumoto, H.; Toyama, Y. Development and biological evaluation of apatite fiber scaffolds with large pore size and high porosity for bone regeneration. *Key Eng. Mater.* **2003**, *240–242*, 647–650.
45. Aizawa, M.; Ueno, H.; Itatani, K.; Okada, I. Development of scaffolds for tissue engineering using single-crystal apatite fibres and their biological evaluation by osteoblastic cell. *Trans. Mater. Res. Soc. Jpn.* **2003**, *28*, 849–852.
46. Fernandez, A.; Planell, J.A.; Best, S.M. Precipitation of carbonated apatite in the cement system α -Ca₃(PO₄)₂-Ca(H₂PO₄)₂-CaCO₃. *J. Biomed. Mater. Res.* **1999**, *47*, 466–471.
47. Hing, K.A.; Annaz, B.; Saeed, S.; Revell, P.A.; Buckland, T. Microporosity enhances bioactivity of synthetic bone graft substitutes. *J. Mater. Sci. Mater. Med.* **2005**, *16*, 467–475.
48. Porter, A.E.; Taak, P.; Hobbs, L.W.; Blunn, G.W.; Coathup, M.J.; Spector, M. Bone bonding to hydroxyapatite and titanium surfaces on human femoral stems retrieved at autopsy. *Biomaterials* **2004**, *25* (21), 5199–5208.
49. Bloebaum, R.D.; Dupont, J.A. Osteolysis from a press-fit hydroxyapatite-coated implant. A case study. *J. Arthroplasty* **1993**, *8* (2), 195–202.
50. van Dijk, K.; Schaeken, H.G.; Wolke, J.C.; Maree, C.H.; Habraken, F.H.; Verhoeven, J.; Jansen, J.A. Influence of discharge power level on the properties of hydroxyapatite films deposited on Ti6Al4V with RF magnetron sputtering. *J. Biomed. Mater. Res.* **1995**, *29* (2), 269–276.
51. Thian, E.S.; Huang, J.; Best, S.M.; Barber, Z.H.; Bonfield, W. A new way of incorporating silicon in hydroxyapatite (Si-HA) as thin films. *J. Mater. Sci. Mater. Med.* **2005**, *16* (5), 411–415. . .
52. Porter, A.E.; Rea, S.M.; Galtrey, M.; Best, S.M.; Barber, Z. Production of thin film silicon-doped hydroxyapatite via sputter deposition. *J. Mater. Sci.*, **2004**, *39*, 1895–1898.
53. Kokubo, T.; Kushitani, H.; Sakka, S. Solutions able to reproduce *in vivo* surface-structure changes in bioactive glass-ceramic A-W. *J. Biomed. Mater. Res.* **1990**, *24*, 721–734. . .
54. Yan, W.Q.; Nakamura, T.; Kawanabe, K.; Nishigochi, S.; Oka, M.; Kokubo, T. Apatite layer-coated titanium for use as bone bonding implants. *Biomaterials* **1997**, *18* (17), 1185–1190.
55. Lawson, A.C.; Czernuszka, J.T. Collagen-calcium phosphate composites. *Proc. Inst. Mech. Engl. (H)* **1998**, *212* (6), 413–425.
56. Song, J.; Malathong, W.; Bertozzi, C.R. Mineralization of synthetic polymer scaffolds: a bottom-up approach

- for the development of artificial bone. *J. Am. Chem. Soc.* **2005**, *127* (10), 3366–3372.
57. Fernandez, E.; Gil, F.X.; Ginebra, M.P.; Driessens, F.C.M.; Planell, J.A.; Best, S.M. Calcium phosphate bone cements for clinical applications, Part I, Solution Chemistry. *J. Mater. Sci. Mater. Med.* **1999**, *10*, 169–176.
 58. Fernandez, E.; Gil, F.X.; Ginebra, M.P.; Driessens, F.C.M.; Planell, J.A.; Best, S.M. Calcium phosphate bone cements for clinical applications, Part II, precipitate formation during setting reactions. *J. Mater. Sci. Mater. Med.* **1999**, *10*, 177–184.
 59. Hing, K.A. Assessment of Porous Hydroxyapatite for Bone Replacement. Ph.D. Thesis. University of London, 1996.
 60. Hing, K.A.; Best, S.M.; Tanner, K.E.; Bonfield, W.; Revell, P.A. Mediation of bone ingrowth in porous hydroxyapatite bone substitutes. *J. Biomed. Mater. Res.* **2004**, *68*, 187–200.
 61. Weng, J.; Liu, Q.; Wolke, J.G.C.; Zhang, X.; de Groot, K. Formation and characteristics of the apatite layer on plasma-sprayed hydroxyapatite coatings in simulated body fluid. *Biomaterials* **1995**, *18*, 1027–1035.
 62. Radin, S.R.; Ducheyne, P. The effect of calcium phosphate ceramic composition and structure on in vitro behavior. II. Precipitation. *J. Biomed. Mater. Res.* **1993**, *27*, 35–45.
 63. Ducheyne, P.; Qiu, Q. Bioactive ceramics: The effect of surface reactivity on bone formation and bone cell function. *Biomaterials* **1999**, *20*, 2287–2303.
 64. Barralet, J.; Akao, M.; Aoki, H.; Aoki, H. Dissolution of dense carbonate apatite subcutaneously implanted in Wistar Rats. *J. Biomed. Mater. Res.* **1999**, *49*, 176–182.
 65. Hankermeyer, C.R.; Ohashi, K.L.; Delaney, D.C.; Ross, J.; Constantz, B.R. Dissolution rates of carbonated hydroxyapatite in hydrochloric acid. *Biomaterials* **2002**, *23* (3), 743–750.
 66. Porter, A.E.; Hobbs, L.W.; Benezra Rosen, V.; Spector, M. The ultrastructure of the plasma-sprayed hydroxyapatite-bone interface predisposing to bone bonding. *Biomaterials* **2002**, *23*, 725–733.
 67. de Bruijn, J.D.; Flach, J.S.; de Groot, K.; van Blitterswijk, C.A.; Davies, J.E. Analysis of the bony interface with various types of hydroxyapatite in vitro. *Cells Mater.* **1993**, *3*, 115–127.
 68. Layani, J.D.; Cuisnier, F.J.G.; Steuer, P.; Voegel, J.C.; Mayer, I. High-resolution electron microscopy study of synthetic carbonate and aluminum containing apatites. *J. Biomed. Mater. Res.* **2000**, *50*, 199–207.
 69. Webster, T.J.; Ergun, C.; Doremus, R.H.; Siegel, R.W.; Bizios, R. Enhanced osteoclast-like cell functions on nanophase ceramics. *Biomaterials* **2001**, *22*, 1327–1333.
 70. Klein, C.P.; de Bleeck-Hogervorst, J.M.; Wolke, J.G.; de Groot, K. Studies of the solubility of different calcium phosphate ceramic particles in vitro. *Biomaterials* **1990**, *11* (7), 509–512.
 71. Gibson, I.R.; Best, S.M.; Bonfield, W. Effect of silicon substitution on the sintering and microstructure of hydroxyapatite. *J. Am. Ceram. Soc.* **2002**, *85* (11), 2771–2777.
 72. Porter, A.E.; Patel, N.; Skepper, J.N.; Best, S.M.; Bonfield, W. Effect of sintered silicate-substituted hydroxyapatite on remodelling processes at the bone-implant interface. *Biomaterials* **2004**, *25* (16), 3303–3314.
 73. Porter, A.E.; Patel, N.; Skepper, J.N.; Best, S.M.; Bonfield, W. Comparison of in vivo dissolution processes in hydroxyapatite and silicon-substituted hydroxyapatite bioceramics. *Biomaterials* **2003**, *24*, 4609–4620.



Hydroxyapatite Coatings

Yunzhi Yang

*Departments of Biomedical Engineering and Orthopedic Surgery/Campbell Clinic,
University of Tennessee Health Science Center, Memphis, Tennessee, U.S.A.*

Kazuhisa Bessho

*Department of Oral and Maxillofacial Surgery, Graduate School of Medicine,
Kyoto University, Sakyo-ku, Kyoto, Japan*

Kyo-Han Kim

*Department of Dental Biomaterials, College of Dentistry, Kyungpook National University,
Jung-gu, Daegu, South Korea*

Sang-won Park

*Department of Prosthodontics, College of Dentistry, Chonnam National University,
Dong-Ku, Gwang-ju, South Korea*

Joo L. Ong

*Departments of Biomedical Engineering and Orthopedic Surgery/Campbell Clinic,
University of Tennessee Health Science Center, Memphis, Tennessee, U.S.A.*

INTRODUCTION

The use of plasma-sprayed hydroxyapatite (HA)-coated implants in the clinics over the past 10 years has initiated substantial controversy in the field of dental implantology. One of the greatest concern about the use of plasma-sprayed HA coated implants is the general lack of long-term documentation on implant survival. Over the years, vast amount of in vivo research has indicated biocompatibility with HA-coated implants and a better performance when compared to noncoated titanium (Ti) implants. The better performance of plasma-sprayed HA implants is because of rapid initial osseointegration, and the high bone contact and interfacial strength between implants and bone.^[1-4] However, the advantages of plasma spraying HA on Ti implants are mired in considerable controversy owing to a lack of correlation between implant success and the specific surface properties of HA. It is known that the properties of HA coatings are altered during the coating process. Although, there is a lack of appreciation for the effects of altered HA coating properties on bone cell activity, the rationale for coating implants with HA grew out of the desire to combine the strength of Ti implant with a bioactivity HA surface.^[5-10] However, there are problems with plasma-sprayed HA coatings, including variation in the coating-metal adhesion strength, nonuniformity in coating thickness between vendors, alteration in structural and chemical properties during the coating process, and nonuniformity in coating density.^[1-10] As a result of these problems, there is controversy over whether or not HA coatings are beneficial to implant success. However, these problems are not shortcomings inherent in the rationale for

HA coating, but rather in the plasma spray technology currently used for depositing HA coatings on implant surfaces. This article will provide the current research on plasma-sprayed HA implants, including some plasma spraying techniques, in vivo study, clinical applications, as well as future developments.

PLASMA SPRAYING

Plasma spraying is the current commercial method for producing HA coatings on implant surfaces. As shown schematically in Fig. 1, a d.c. electric arc is struck between two electrodes, and an ionized high temperature gas (up to 30,000°C) is resulted, as a stream of gases passes through this electric arc. The increase in gaseous temperature resulted in an expansion of the gas and causes the carrier gas stream to pass through the arc at approximately the speed of sound. The coating powder, suspended in the carrier gas stream, is then fed into the plasma flame. The heat content of the plasma flame, and thus the ability to increase the temperature of a particle, is strongly dependent on the type of gas used. In addition, the longer the particle remains in the plasma flame, the higher will be the plasma temperature. Ideally, only a thin outer layer of each powder particle gets into the molten plastic state, which undergoes some phase transitions. This molten plastic state is critical to ensure an adhesive and dense coating. As a result, properties of plasma-sprayed coatings is highly dependent on particle size, type of gas used, speed of the plasma, the distance between the implants and plasma nozzle, and the cooling rate for the desired structural and chemical properties.

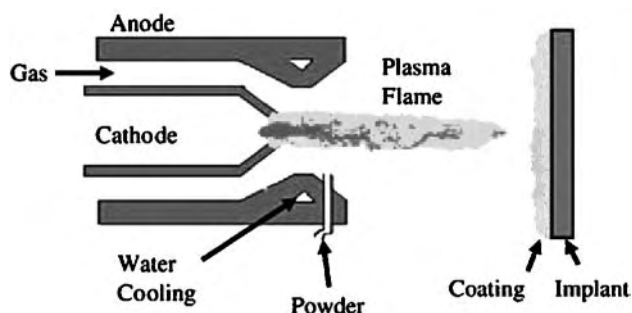


Fig. 1 Schematic representation of the plasma spraying process.

Apatite is known to be thermodynamically unstable during plasma spraying at high temperatures. When HA is plasma sprayed in air, calcium oxide is usually formed and its concentration will affect the integrity of plasma-sprayed HA coatings. A poor coating integrity is resulted from the high concentration of calcium oxide reacting with water. In addition, an increase in calcium oxide can also be resulted from the presence of a gas carrier having high hydrogen concentration. The use of hydrogen gas as a carrier resulted in a lower flame velocity and a more flame enthalpy as compared to the use of nitrogen gas, thereby resulting in more melting and decomposition of HA particles.

Plasma-sprayed HA coatings on orthopedic and dental implants usually have a thickness of 50–100 μm . Coatings become brittle when coating thickness exceeds 80 μm . However, a fast resorption may occur with very thin HA coatings. In addition, significant porosity within the HA coatings weakens the coating's mechanical strength and increases the resorption rate. It is desirable to have plasma-sprayed coatings that are completely dense and consist of pure HA. As such, the calcium-to-phosphorus ratio in the plasma-sprayed coating should be as similar to natural HA ($10/6 = 1.67$). Although mixtures of HA with tricalcium phosphate (TCP) have been investigated, clinical applications of these mixed coatings are rare. This is mainly attributed to the faster degradation rate of TCP as compared to HA and the fear that these mixed coatings may result in an unpredictable or uncontrolled degradation of the entire coating. In addition to TCP, coatings of calcium phosphates phases such as brushite, fluorapatite, and brushite have also been investigated, although there are no reported advantages to justify their clinical use.

IN VIVO STUDY

Although there are numerous reports indicating direct bone contact with plasma-sprayed HA-coated implants,^[11] only a few of these reports quantified and compared the percentages of bone contact with

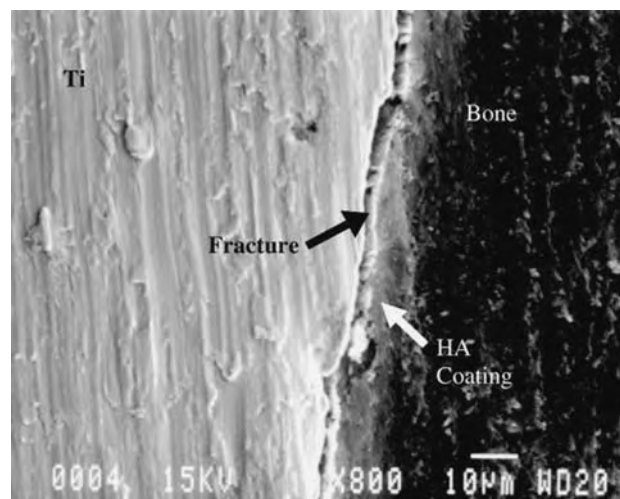


Fig. 2 Representative scanning electron micrograph of plasma-sprayed hydroxyapatite (HA) after one year loading in a dog's mandible followed by pull-out testing. Fracture (indicated by the black arrow) took place between the titanium (Ti) substrate and the HA coating during pull-out testing. The white arrow shows the plasma-sprayed HA coating.

implants. Animal studies have also demonstrated a higher bone implant contact for plasma-sprayed HA as compared to noncoated Ti implant surfaces, although no significant difference in bone contact length between the plasma-sprayed HA and plasma-sprayed titanium (TPS) surfaces has been reported.^[11]

In addition to the above observations with regard to bone contact length, our studies have also indicated no significant difference in mean ultimate interfacial strength between plasma-sprayed HA, TPS, and noncoated Ti implants at 12 weeks after implant placement and at one year after loading.^[11] As shown in Fig. 2, scanning electron micrographs of the mechanically tested implants at one year after loading indicated a weakness at the coating-substrate interface. From the above observations, the lack of correlation between bone contact percentage and interfacial strength suggested a different bonding mechanisms for plasma-sprayed HA and TPS implants, with a chemical bonding occurring between bone and plasma-sprayed HA surfaces, whereas a mechanical interlocking occurring between bone and TPS surfaces (Fig. 3).

APPLICATION OF HA COATING IN CLINICS

The HA is commonly plasma sprayed on implants used in dentistry, total hip arthroplasty, total knee arthroplasty, total elbow arthroplasty, and total shoulder arthroplasty. The long-term clinical success of plasma-sprayed HA-coated endosseous dental implants has been excellent, with a survival rates in excess of 90% for

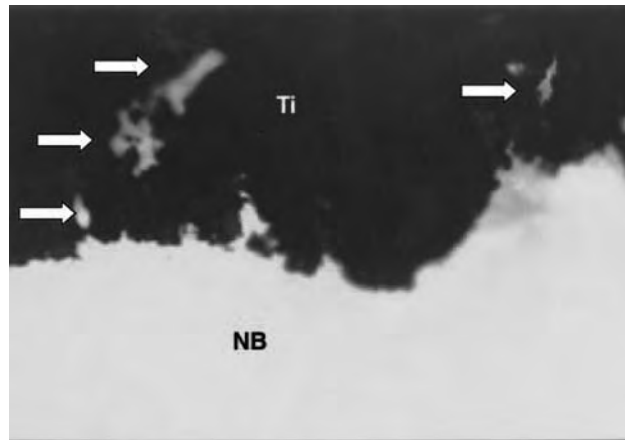


Fig. 3 Representative fluorescence micrograph of the plasma-sprayed Ti implant-bone interface at eight weeks after implantation (original magnification, $\times 100$; arrowhead = NB).

implants at greater than five to eight years' follow-up. In a clinical study at three study sites evaluating the five-year postrestoration safety and efficacy of 2062 threaded, plasma-sprayed HA-coated implants placed and restored among 720 patients, the surgical success rate at exposure was reported to be 98.3%.

From the clinical point of view, HA coatings have excellent success. However, peri-implant bone loss, rapid bone resorption, and a higher failure rate for HA-coated implants have also been reported. During a five years follow-up (mean 7.9 yr, range, 5.3–9.1 yr), one (1%) out of the 131 HA-coated threaded cups, two (2%) out of the 109 porous-coated press-fit cups, and 21 (11%) out of the 188 plasma-sprayed HA-coated press-fit cups were revised owing to aseptic loosening.^[11]

The prerequisites for achieving bone-dental implant osseointegration include the use of biocompatible materials, primary stability of the implant, atraumatic surgical procedure, and surgical protocol with unloaded submerged implants during three to six months healing.^[12] In 1985, Furlong emphatically insisted on fulfilling conditions of contact, stillness, and stressing in order for plasma-sprayed HA-coated implant to integrate with bone.^[13] In addition to the presence of plasma-sprayed HA coatings, implant design and suitable surgical procedures are equally critical in providing the initial total implant immobility (primary fixation), so that bone ingrowth (secondary fixation) can occur.

RETRIEVAL STUDIES

Implant retrieval studies are important in not only highlighting the success of the implant, but also in providing failure information that would allow the improvement of future implants. In examining six clinically successful

plasma-sprayed HA-coated cementless acetabular components that were retrieved at autopsy between 3.3 and 6.6 yr after implantation, bone growth onto the implant was indicated for all cups and with a bone-implant contact of $36.5 \pm 13.5\%$.^[14] The extent and thickness of plasma-sprayed HA coatings were reduced in specimens obtained from old patients and in specimens associated with a longer duration of implantation. Osteoclastic degradation of plasma-sprayed HA coatings was observed. Although loose HA granules were observed near the coatings, no adverse tissue reaction to these granules was reported.

In examining five-year plasma-sprayed HA-coated implants retrieved at autopsy from three patients who died from nonhip-related causes, Bauer et al.^[15] described excellent clinical performance of all involved implants with implantation time varying between 4 and 25 mo. An average coating thickness of $40 \mu\text{m}$ was observed on all plasma-sprayed HA-coated explants, with no observation of coating delamination and no correlation between length of implantation and coating thickness. The lack of correlation between coating thickness and length of implantation suggested minimal early loss of coating to extracellular fluid. A uniform layer of bone, ranging from 32% to 78%, was reported over the HA coating circumference, with more bone fixation observed in the distal parts of the coating area. The average area of osseointegration was observed to be greater than 60% of available HA surface. In the 25 mo retrieved sample, focal areas of osteoclastic resorption on the plasma-sprayed HA coatings were observed. Once bony coverage of the coating was obtained, the process of physicochemical dissolution was suggested to be arrested. However, as observed in the 25 mo retrieved sample, a cellular process of coating degradation caused by osteoclastic activity was observed, and that the osteoclastic resorption of plasma-sprayed HA coatings was replaced by bone. This suggested that with longer implantation period, the plasma-sprayed HA coating will be replaced by autologous bone, while preserving the quality of osseointegration. There was no fibrous tissue formation in these HA-resorption areas. In addition, degradation of the plasma-sprayed HA coating was more pronounced in the proximal femur as compared with the more cortical middle and distal stem area, suggesting a relationship between coating degradation and rate of bone remodeling. However, the quality and amount of fixation in all retrieved samples were not affected by degradation of the plasma-sprayed HA coatings.

CURRENT PROBLEMS

Aside from the weak coating-metal interface, vast amount of investigations have indicated that there are

no accepted standards for producing HA coatings on implant surfaces and that all HA coatings are not identical.^[1,9,10] As such, variability in coating quality has resulted in many conflicting animal and clinical observations. Although plasma-sprayed HA coatings on implant surfaces have been indicated to enhance implant stability, such findings have not been universal. These equivocal findings likely stem from the fact that HA coatings are poorly characterized. Typical plasma-sprayed HA coatings have been reported to consist primarily of partially dehydrated HA, with amorphous calcium phosphate and other more soluble phases such as TCP produced during the high-temperature deposition process. Although plasma spraying has the ability to produce coating crystallinity in the range 30% to 70%, the crystallinity for plasma-sprayed HA coatings produced under normal deposition process is approximately 65%. In addition, the use of plasma spraying technology often results in coating properties that differs from the starting HA powders.

The ratio of HA to TCP has been reported to be crucial for bone regeneration. It has also been reported that the dissolution rate of an HA coating is dependent on the biochemical calcium phosphate phase of the coating. As coating crystallinity increases, dissolution rate of the coatings decreases. In addition, increasing the amorphous and TCP concentration within the coatings are thought to predispose HA coatings to dissolution. In addition to the coating integrity, the local concentrations of calcium and phosphate in the pericellular environment are known to influence cellular response. Enrichment of the calcium and phosphate ions in the localized biological environment, as a result of dissolving CaP coatings, may promote the formation of bone apatite-like minerals. However, a high calcium concentration is also known to induce apoptosis of bone-forming cells and to promote the formation of osteoclasts.^[16] As an example of the influence of indirect dissolution on bone biological response, the effect of coating crystallinity on cellular response is shown in Fig. 4. From the figure, it was observed that varying the crystalline content of the HA coatings significantly affects the bone-implant interfacial strength. Lacefield and Soballe have also identified several other coating properties that affect bone response, including surface texture, porosity, and the presence of trace elements.^[17,18] As such, there is a need to address the optimal crystallinity, porosity, texture, and presence of trace elements within the coatings to maximize implant success.

FUTURE DEVELOPMENTS

There is no standard guideline for depositing HA coatings on medical devices. At present, most of the coatings were performed using the plasma spraying

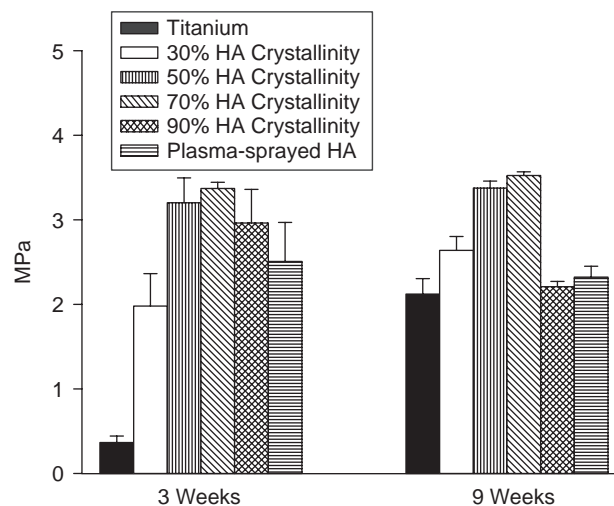


Fig. 4 Push-out strength at the implant–bone interface using a rat femur model.

technology because of its rapid deposition rate and sufficiently low cost. However, despite these advantages and the reported clinical success, there are problems associated with plasma-sprayed HA coatings. These problems included poor adhesion and variation in bond strength between the coatings and metallic substrates, as well as alterations in HA properties as a result of the coating process. The nature of the substrate plays an important role in the adhesion between plasma-sprayed coatings and metallic substrates. The bonding of plasma-sprayed coatings to the metallic substrate appears to be entirely mechanical in nature, and a highly roughened substrate surface is needed to ensure coating–metal adhesion.

To overcome the problems associated with plasma spraying, experimental deposition processes such as electrophoretic codeposition, sputter deposition, and high-velocity oxy-fuel combustion spray deposition have been investigated.^[10] Although these experimental processes have been successful to overcome some of the problems associated with plasma spraying, there are limitations with all of these processes. For example, sputtering has been reported to have ability for controlling the physical and chemical properties of HA coatings produced on dental and orthopedic implants.^[12] The advantages of radio-frequency magnetron sputter-coating over other coating processes include a highly dense coating with high metal-implant bonding strength (40 MPa for sputtered coatings vs. 9 MPa for plasma-sprayed HA coatings), and a highly uniform and continuous coating.^[9,10] Another important advantage of the sputtering process is that the highly adhesive coatings can be performed on mirror-polished surface.^[10] However, despite these advantages, the sputtering process has a very low sputtering rate, and thus is not economical.

As the biological apatite is nanosize, the synthesized nano-HA is expected to be recognized as biocompatible. The nano-HA should be involved in the natural bone remodeling process rather than being phagocytized like polymer debris.^[19] At present, there is a huge interest in nanotechnology including the production of nano-HA that can be used to optimize osseointegration. It has already been reported that nano-HA supported the attachment and the spread of human osteoblast cells.^[20] Although this study and many other studies on nano-HA have shown promise, more studies are needed before nano-HA can be used in clinical applications.

CONCLUSIONS

New promising coating processes are continuously being investigated. Investigators are presently exploring nanotechnology and biomimetics in the hope to optimize coating properties for rapid osseointegration. These investigations have included better materials characterization, understanding bone responses to implant surfaces of different properties, and the use of various biological factors. Although it is generally accepted that plasma-sprayed HA and TPS coatings improve bone strength and initial osseointegration, optimum coating properties required to achieve maximum bone response are yet to be achieved. As such, coating methodology, the use of HA and other calcium phosphate phases, together with well characterized implant surfaces in cell culture, animal, and clinical studies, should continue to be investigated to provide basic understanding of bone-implant interactions, and that an improved HA-coated implants are needed to optimize osseointegration.

ACKNOWLEDGMENTS

The studies presented were funded by NIH/NIDCR (DE12276), NIH/NIAMS (AR46581), and collaboration was also supported in part by a Grant-in-Aid for Scientific Research B (No. 15390612) of the Japanese Ministry of Education, Science, Sports and Culture. The authors would also like to thank Friatec AG, Germany for some of the implants used in the book chapter.

ARTICLES OF FURTHER INTEREST

Bioactive Glass; Biologic and Synthetic Apatites

REFERENCES

1. Lemons, J.E. Hydroxyapatite coatings. *Clin. Orthop. Rel. Res.* **1988**, 235, 220–223.
2. Lacefield, W.R. Hydroxylapatite coatings. In *Bioceramics: Material Characteristics Versus In vivo Behavior*; Ducheyne, P., Lemons, J.E., Eds.; The New York Academy of Science: New York, 1988; 72–80.
3. Ducheyne, P.; Cuckler, J.M. Bioactive ceramic prosthetic coatings. *Clin. Orthop. Rel. Res.* **1992**, 276, 102–114.
4. Furlong, R.J.; Osborn, J.F. Fixation of hip prostheses by hydroxyapatite ceramic coatings. *J. Bone Joint Surg.* **1991**, 73B, 741–745.
5. Geesink, R.G.T. Osteoconductive coating for total joint arthroplasty. *Clin. Orthop. Rel. Res.* **2002**, 395, 53–65.
6. Havelin, L.I.; Engesaeter, L.B.; Espehaug, B.; Furnes, O.; Lie, S.A.; Vollset, S.E. The Norwegian arthroplasty register, 11 years and 73,000 arthroplasties. *Acta Orthop. Scand.* **2000**, 71, 337–353.
7. D'Antonio, J.A.; Capello, W.N.; Manley, M.T.; Geesink, R. Hydroxyapatite femoral stems for total hip arthroplasty. *Clin. Orthop. Rel. Res.* **2001**, 393, 101–111.
8. de Groot, K.; Wolke, J.G.C.; Jansen, J.A. State of the art: hydroxyapatite coatings for dental implants. *J. Oral Implantol.* **1994**, 20, 232–234.
9. Yang, Y.; Ong, J.L.; Bessho, K. Plasma-sprayed hydroxyapatite-coated and plasma-sprayed titanium-coated implants. In *Biomaterials in Orthopedics*; Yaszemski, M.J., Trantolo, D.J., Lewandrowski, K., Hasirci, V., Altobelli, D.E., Wise, D.L., Eds.; Marcel Dekker, Inc.: New York, 2004; 401–423.
10. Yang, Y.; Kim, K.H.; Ong, J.L. A review on calcium phosphate coatings produced using a sputtering process—an alternative to plasma spraying. *Biomaterials* **2005**, 26, 327–337.
11. Manley, M.T.; Capello, W.N.; D'Antonio, J.A.; Geesink, R.G.T. Fixation of acetabular cups without cement in total hip arthroplasty. *J. Bone Joint Surg.* **1998**, 80A, 1175–1185.
12. Adell, R.; Lekholm, V.; Rockler, B. A 15-year study of osseointegrated implants in the treatment of the edentulous jaw. *Int. J. Oral Surg.* **1981**, 10, 387–416.
13. Furlong, R. Proximal femoral bone loss and increased rate of fracture with a proximally hydroxyapatite-coated femoral component. *J. Bone Joint Surg.* **2001**, 83B, 46.
14. Tonio, A.J.; Therin, M.; Doyle, C. Hydroxyapatite-coated femoral stems. Histology and histomorphometry around five components retrieved at post mortem. *J. Bone Joint Surg.* **1999**, 81B, 148–154.
15. Bauer, T.W.; Stulberg, B.N.; Jiang, M.; Geesink, R.G.T. Uncemented acetabular components: histologic analysis of retrieved components. *J. Arthroplasty* **1993**, 8, 167–177.
16. Adams, C.S.; Mansfield, K.; Perlot, R.L.; Shapiro, I.M. Matrix regulation of skeletal cell apoptosis. Role of calcium and phosphate ions. *J. Biol. Chem.* **2001**, 276, 20,316–20,322.
17. Lacefield, W.R. Characterization of hydroxyapatite coatings. *J. Oral Implantol.* **1994**, 20, 214–220.
18. Soballe, K.; Hansen, E.S.; Brockstedt-Rasmussen, H.; Burgess, A. Torsional stability of HA-coated and grit

- blasted Ti dental implants. *J. Bone Joint Surg.* **1993**, *75B*, 270–278.
19. Driessens, F.C.M.; Boltong, M.G.; Khairoun, I.; De Maeyer, E.A.P.; Ginebra, M.P.; Wenz, R.; Planell, J.A.; Verbeeck, R.M.H. Applied aspects of calcium phosphate bone cement application. In *Biomaterials Engineering and Devices. Human Applications. Volume 2*; Wise, D.L., Trantolo, D.J., Lewandrowski, K.U., Gresser, M.V., Yaszemski, M.J., Eds.; Humana Press: New Jersey, Totowa, 2000; 253–260.
20. Huang, J.; Best, S.M.; Bonfield, W.; Brooks, R.A.; Rushton, N.; Jayasinghe, S.N.; Edirisinghe, M.J. In vitro assessment of the biological response to nano-sized hydroxyapatite. *J. Mater. Sci. Mater. Med.* **2004**, *15*, 441–445.

Implant, Total Hip

Nadim James Hallab
Joshua Jacobs

Department of Orthopedic Surgery, Rush Medical College, Rush-Presbyterian-St. Lukes Medical Center, Chicago, Illinois, U.S.A.

INTRODUCTION

Total hip arthroplasty (THA) is enormously successful in restoring mobility and quality of life to millions of individuals. This article covers a history of total hip arthroplasty, current technology, clinical concerns, and future developments. The history of total hip arthroplasty is particularly pertinent to biomaterials science because it is one of the best illustrations of how an implant first used over a century ago has achieved such high success, primarily because of advances in biomaterials. In general, total hip arthroplasty evolved from interpositional membranes inserted between the femoral head and acetabulum, to cups that fit over the femoral head, to short-stemmed prostheses, and finally to long-stemmed total hip prostheses. Clinical issues associated with biomaterial degradation basically break down to four: 1) the amount of material released from the implant; 2) where the material is transported and in what quantity; 3) the chemical form of released degradation products (e.g., inorganic precipitate versus soluble organometallic complex); and 4) the pathophysiological interactions and consequences of biomaterial degradation.

The clinical concerns portion of this article focuses on what is known (and is of orthopedic clinical concern) regarding biomaterial degradation (through wear and electrochemical corrosion), dissemination of debris, and consequent local/systemic effects. Currently total hip arthroplasty is typified by a metal alloy femoral stem with metal femoral head that articulates on a polymer acetabular cup. This article focuses on the metals that compose THA and why particular alloys are used for specific components. However, there exist many variations on this archetype, including variations in materials and geometry. Despite the enormous success of modern THA, current efforts to improve implant performance center on the development of new materials and on small, incremental changes in implant design.

HISTORY OF TOTAL HIP ARTHROPLASTY

The earliest attempts to restore mobility to painful and deformed hip joints took place in the 1820s (White,

1822, and Barton, 1827) and centered on simply removing the affected femoral and acetabular bone involved. This evolved in the 1830–1880s into ghastly attempts to restore mobility using interpositional membranes between the femoral head and acetabulum, where materials such as wooden blocks and animal (e.g., pig) soft tissue were tried. The first prosthetic hip replacement dates to 1890, when Gluck published a description of a carved ivory femoral head replacement using bone cement like materials such as pumice and plaster of paris to secure the implants in place.^[1,2]

Most popular was the interpositional membrane strategy that continued into the twentieth century with the use of new implant materials in the early 1900s (1900–1920), including organic materials (e.g., pig bladders and peri-implant soft tissues) and inorganic materials such as gold foil. The use of an individual's own soft tissues was the most popular method of interpositional membrane hip surgery. The limited success of this procedure prevented widespread use, although the treatment of painful, disfigured, and frozen (ankylosed) hip joints with other implant materials remained commonplace into the 1920s.

Mold arthroplasty

In 1923 Marius Smith-Peterson was credited with ushering in the modern era of total joint replacement with the development of mold arthroplasty (Fig. 1), made of glass and inspired by a shard of glass found in a patient's back, surrounded in a benign, synovial membrane. This mold, or cup, arthroplasty was designed as a cup that fit between the femoral head and the acetabular cup, and articulated on both surfaces using a cartilagelike layer. Despite the fanfare, his desire was simply to develop a better interpositional membrane than had been in use for the previous 100 years. The efforts of Smith-Peterson and colleagues over the years 1923–1938 were spent improving the fracture resistance of the glass mold arthroplasty cup design, using materials such as early polymers (e.g., celluloid or phenol-formaldehyde Bakelite or Formica) and improved glass (e.g., Pyrex). It was not until 1939, when the first metal—a cobalt alloy termed Vitallium—became available and was used by Venable, Stuck, and Beach,

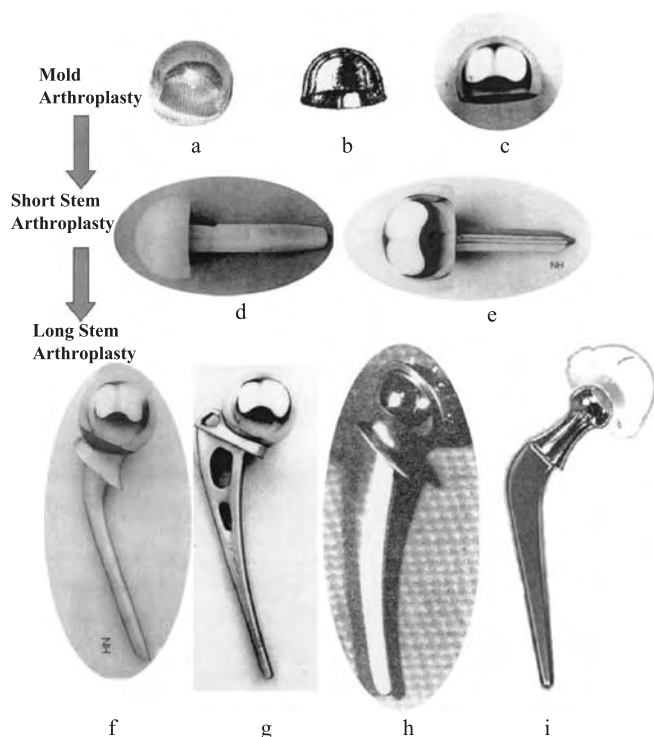


Fig. 1 Array of implants representing the history of THA evolution. (a) Marius Smith-Peterson in 1923 is credited with ushering in the modern era of total joint replacement with his mold arthroplasty made of glass; (b) Venable, Stuck, and Beach used a cobalt alloy in 1939 named Vitallium, possessing superior corrosion resistance; (c) Vitallium alloy facilitated further design modifications of the mold arthroplasty by Otto E. Aufranc; (d) the Judet brothers in Paris, in 1946 used polymethylmethacrylate to manufacture short-stemmed prostheses; (e) a Judet brothers Vitallium (cobalt-chrome alloy) prosthesis, which replaced acrylic in short-stem designs in 1951; (f) 1950 designs of Frederick R. Thompson; and (g) Austin T. Moore, which ushered in the era of long-stemmed prostheses; (h) the 1951 McKee-Farrar prosthesis evolved quickly to incorporate a true spherical femoral head undercut at the neck to reduce impingement of the head on the rim of the acetabular prosthesis and provide for greater range of mobility; (i) a 1962 Charnley low-friction arthroplasty remains the prototype of total hip arthroplasty performed today.

that the corrosion resistance of the hip arthroplasty provided sufficient biocompatibility and performance (Fig. 1).

In 1937 Venable, Stuck, and Beach published a landmark article that was the first to analyze in systematic fashion the electrolytic effects of various metals on bone and tissues (e.g., aluminum, copper, iron, nickel, lead, gold, magnesium, silver, stainless steel, and other alloys), and arrived at the conclusion that Vitallium (a cobalt chromium alloy) was superior to other metals in corrosion resistance and in the mechanical

properties required for an implant.^[3,4] By observing the effects of corrosion and proposing guidelines for performance, Venable, Stuck, and Beach set the standard by which future metallic alloys would be selected for use in hip and other types of implants.

The superior material properties of the Vitallium alloy facilitated further design modifications of the mold arthroplasty by Otto E. Aufranc (Fig. 1). Aufranc removed the rim of the Smith-Peterson mold (often the cause of adhesions, cup freezing, and subsequent pain and immobility), and machined matching curves on the inner and outer surface to meet at the rounded outer edge. Despite the high short-term success rates (<4 years) reported by Aufranc (>82%), the overall failure rate remained high (>50%). Another design modification of the mold arthroplasty in the 1940–1950s was fixation of the mold to the acetabulum rim with screws by physicians such as Albee-Pearson and Gaenslen. Although used in only four cases, Gaenslen reported using a cobalt alloy mold fixed to the acetabulum and another fixed to the femoral head, creating a metal-on-metal total hip replacement. The popularity of mold arthroplasties endured into the 1970s when they remained touted as the treatment of choice for traumatic arthritis of the hip by leading orthopedic surgeons.^[5] Back in the 1930s, however, the natural progress in THA development was the progression from mold arthroplasty to short-stem prosthesis.

Femoral head prostheses/short-stem prostheses

Femoral head prostheses were first made of such materials as ivory (Gluck, 1890) and rubber (Delbet, 1919), and were cemented for stability.^[1,2] At about the same time, these replacement heads were first fitted with a short stem by Earnest Hey Groves, who used an ivory nail to replace the articular surface of the femur. These types of implants were rare and remained unpopular compared to mold arthroplasties until 1937 when Harold Bohlman, using the work of Venable and Stuck, designed a corrosion-resistant cobalt–chrome alloy femoral head replacement with a short stem. This design was popularized by the Judet brothers in Paris in 1946, who used polymethylmethacrylate (presumed to be biologically inert *in vivo*) to manufacture short-stemmed prostheses (Fig. 1). Initial good results were soon replaced with problems of implant fracture and excessive wear debris, and by the early 1950s, these implants were losing favor and being removed by surgeons. Vitallium (cobalt–chrome alloy) eventually replaced acrylic in several other short-stem designs. However, there were sound short-stem designs as early as 1938, when Wiles introduced the cobalt alloy femoral shell attached to the femur with a central nail. This design was popularized by Peterson in 1950, when he

used a similar Vitallium shell design with a central nail and a plate attached to the nail for added stability. Others (including J. Thompson, 1951, and Rossignal, 1950) adopted and adapted the Judet brothers' design using Vitallium. Rossignal designed large threads onto the stem to aid in fixation. These short-stem designs were subject to what was deemed high shear stress, and resulted in early loosening and failure in some patients. Short-stem designs were gradually replaced by longer stem designs that provided less stress concentration.

Long-stem prostheses

Long-stem prostheses continued the trend established by short-stemmed prostheses (i.e., increasingly more weight-bearing forces were transferred to the femur though an intramedullary stem). The pattern for long-stem prosthesis was established in 1940 by Bohlman in collaboration with Austin T. Moore, in which they implanted a 12-inch Vitallium prosthesis that replaced the femoral head and had long supports screwed into the outside of the femoral shaft.^[6] Although there were innovations in long-stem design in the 1940s (such as the doorknob design of Earl McBride, where a threaded stem was screwed into the intramedullary canal of the femur for fixation and load transfer), these designs were not popular. It was not until the 1950 designs of Frederick R. Thompson and Austin T. Moore that long-stemmed prostheses became popular (Fig. 1). These designs were cast in Vitallium (cobalt-chrome alloy) and required the removal of the femoral head but only part of the neck. The design of Moore differed from Thompson's in that it had fenestrations through the implant to allow bone growth, and it had a rear vane to enhance rotational stability. Initially these implants were used without bone cement. The success of the designs of the Thompson and Moore prostheses is proved by their continued use with only slight variations. Despite the excellent design of these early long-stemmed prostheses, they were primarily successful when used to replace diseased femoral heads, and did not work well when acetabular reaming was required. This inadequacy prompted the development of the total hip replacement arthroplasty.

Total hip replacement arthroplasty

Philip Wiles is credited with the first total hip arthroplasty in 1938, when he used a stainless steel ball secured to the femur with a bolt and a stainless acetabular liner secured with screws.^[7] The results of this design were disappointing because of the poor corrosion resistance of early stainless steel in vivo, the high stress concentrations of short-stemmed prostheses, etc. An adaptation of this design that proved successful was

developed by G.K. McKee and J. Watson-Farrar in 1951. They used a stainless steel cup and long-stemmed prosthesis (Thompson stem) that rapidly failed due to the poor corrosion resistance of the stainless steel. The cup was changed to cobalt-chrome alloy with greater success. The McKee-Farrar prosthesis evolved quickly to incorporate a true spherical femoral head, undercut at the neck to reduce the impingement of the head on the rim of the acetabular prosthesis and to provide for greater range of mobility (Fig. 1).^[8]

The next milestone in the evolution of modern total hip arthroplasty was the advent/popularization of acrylic dental bone cement, first used by Sven Kiar in 1950 to attach a plastic prosthesis to bone.^[9] Later that year, the Hospital for Joint Diseases in New York used polymethylmethacrylate (acrylic) bone cement as a means of fixation in total hip arthroplasties.^[10,11] The development of acrylic bone cement dramatically reduced the rates of loosening associated with metal-metal total hip arthroplasty. The Stanmore metal-metal design, which used a horseshoe-shaped cup, was popular but led to excessive wear and was replaced by a complete cup. McKee and Watson-Farrar adapted their design to facilitate bone cement with a land-minelike studded acetabular cup designed to maximize mechanical fixation.

The 1950s marked the introduction and popularization of the total hip arthroplasty, in which replacement arthroplasty became simple and reliable enough to be practiced on a wide scale by the average orthopedic surgeon. However, the squeaking reported to occur in Judet's and in some later metal prostheses was identified by Charnley to be a result of the relatively high frictional forces in the joint. These high-torque and frictional forces resulted in the generation of significant metallic debris that (purportedly) resulted in early loosening. In 1960 Charnley developed a low-friction arthroplasty device using shells of Teflon on the femoral and acetabular sides, which resulted in early/immediate failures because of massive debonding and wear debris. This was quickly followed by a thick-walled Teflon acetabular component articulating on a small head designed to reduce the shearing forces and torque. However, this design also generated excessive wear debris, which produced immediate and severe inflammation and failure of the prosthesis. Charnley then replaced the Teflon with high-density polyethylene that was not as friction-free as Teflon but was 1000 times more wear-resistant. This prototype of total hip arthroplasty, developed in 1962, was the basis for future designs that remain the most popular form of total hip arthroplasty performed today (Fig. 1).

The basic design of Charnley was modified by Muller with variable neck sizes and larger heads. During the same period, metal-on-metal designs by Smith, Ring,^[12] and others were unsuccessful challengers to the basic

Charnley metal-on-polymer design. Other currently adopted design modifications were developed by Ling, Aufranc, Turner, Amstutz, Harris, and Galante, which include such innovations as femoral prosthesis geometrical modification for increases in stability and mobility, modular components for increased customization, porous coatings, surface texturing/coating to increase fixation and bone ingrowth, etc. Charnley is often deified in orthopedic literature as the metaphorical spark that lit the flames of innovation in prosthetic design. This is typical surgeon-centered glorification. Other implant designs that predate Charnley (such as the all-metal McKee-Farrar THA implant) have enjoyed success rates similar to those reported by Charnley. More important, total hip arthroplasty has evolved over the last century through the innovation of many investigators, and has benefited most significantly from recent advances in areas of materials technology, biomechanics, biochemistry, immunology, infectious diseases, thrombosis, and pharmacology.

TOTAL HIP ARTHROPLASTY, CURRENT TECHNOLOGY

Today the total hip implant remains much as it was in the 1970s, albeit with more variety of implant materials and geometries (Fig. 2). The current THA is typically constructed of a titanium or cobalt–chromium alloy femoral stem (cemented with PMMA or press-fit into place), connected to a modular cobalt–chrome alloy or ceramic head that articulates on an ultrahigh molecular-weight polyethylene (UHMWPE) or ceramic acetabular cup fitted into a titanium or cobalt–chromium cup liner that is cemented, screwed, or press-fit into place (Fig. 2). Despite this simple archetype of the total hip replacement, there exist hundreds of variations on this theme (Fig. 3) offered to today's orthopedic surgeons, with little in terms of absolute guidelines as to which type of implant is the best (or which of more than 10 major manufacturers produces the best) for well-defined orthopedic disease states. However, there are some general guidelines. Typically implants in older individuals (>80 years of age) are cemented into place with bone cement, because the chance for revision is minimal when compared to younger individuals (<60 years), and removing bone cement is technically challenging and may compromise the availability of bone stock. Generally, there are choices of surface roughness, coating, geometry, material composition, etc., and each manufacturer claims that its product is superior to the rest. This, in combination with the lack of much publicly available information tracking the performance of each type of implant, precludes accurate scientific analysis of which implant materials and which design perform best. Additionally, stiff competition between



Fig. 2 Common type of total hip arthroplasty—titanium alloy stem, with cobalt-based alloy (ASTM F-75) modular head bearing on an ultrahigh molecular-weight polyethylene (UHMWPE) liner within a titanium alloy cup. Also shown is a ceramic head and three acetabular sockets with various surfaces for both cemented and cementless fixation. (Photograph courtesy of Zimmer, Inc.)

manufacturers and the extent of commercial promotion required to compete in the marketplace have resulted in a dizzying array of new implants released each year, each claiming to be improved over the previous year's model (see Fig. 3). These claims are clearly baseless, because the typical total hip replacement enjoys a success rate of over 90% at 7 years. In most cases, therefore, a minimum of 7 to 10 years must transpire before such claims can be substantiated; even then, proof of



Fig. 3 Hundreds of THA components designed by dozens of manufacturers present an array of choices to modern orthopedic surgeons, charged with the responsibility to select the best-performing implants.

superior performance is compromised by myriad external factors such as surgeon, region of the country, average activity of patient populations, etc. This conflict between science, marketing, and market share may represent the single greatest obstacle to the scientific determination of superior implant design and progress. The responsibility rests with the U.S. Food and Drug Administration (FDA) to prevent the zeal and economic pressure of the market from undermining implant design in a regressive fashion.

Polymers

Total hip arthroplasty typically consists of an acetabular cup of ultrahigh molecular-weight polyethylene (UHMWPE) articulating against a femoral ball of cobalt–chromium alloy. The resultant wear of the polyethylene bearing produces billions of submicron-sized wear particles annually, predominantly in the 0.5-micron range. Extra cross-linking of polyethylene (using chemical and radiation techniques) has only recently improved its abrasion resistance in hip arthroplasty. Wear testers have shown that the wear resistance of UHMWPE is improved by cross-linking with gamma irradiation at 2.5 to 10 Mrad. However, this can negatively affect such physical properties as tensile strength. Therefore, care is taken and processes patented to minimize negative oxidation effects while preserving high-wear characteristics. The ultimate performance of new, highly cross-linked polyethylene and the subsequent oxidation remain an area of active research.

In order to maximize the performance characteristics of polyethylene, it is cross-linked prior to fabrication into an acetabular cup. Typically an extruded bar of polyethylene is cross-linked using conventional gamma irradiation and then heated to reduce free radicals. The remaining oxidized surface layer of the irradiated bar is discarded prior to or during acetabular cup machining. There have been claims of negligibly small polyethylene debris produced by hip simulators when using highly cross-linked UHMWPE, due to the increase in surface wear resistance.

Ceramics

Alumina (Al_2O_3) and Zirconia (ZrO_2) ceramics have been used in THA for the past 30 years. The first ceramic couple (Alumina/Alumina) was implanted in 1970 by Pierre Boutin. The theoretical advantage of hard-on-hard articulating surfaces was low wear. Ceramics, because of their ionic bonds and chemical stability, are also relatively biocompatible. Initial concerns about fracture toughness and wear have been addressed by lowering grain size, increasing purity, lowering porosity, and improving manufacturing techniques (e.g., hot

isostatic pressing, HIP). Early failures of these couples were plagued with both material-related and surgical errors. The very low wear rates combined with steadily decreasing rates of fracture (now estimated to occur at the rate of 1 in 2000 over 10 years) have resulted in the growing popularity of all ceramic bearings. In general, ceramic debris is purportedly chemically stable; thus, most problems are associated with the amount of debris and not with particle toxicity.

Zirconia was introduced in 1985 as a material alternative to Al_2O_3 for ceramic femoral heads, and has been gaining preference because of its mechanical properties superior to Alumina. Femoral heads of Zirconia can typically withstand 250 kN (or 25 tons), a value exceeding that possible with Alumina or metal femoral heads.

Metals

Metals remain the central material component of state-of-the-art total hip arthroplasties. Metals provide appropriate material properties such as high strength, ductility, fracture toughness, hardness, corrosion resistance, formability, and biocompatibility necessary for use in load-bearing roles required in fracture fixation and total joint arthroplasty (TJA). Implant alloys were originally developed for maritime and aviation uses where mechanical properties such as high strength and corrosion resistance are paramount. Three principal metal alloys are used in orthopedics, particularly in total joint replacement: 1) titanium-based alloys; 2) cobalt-based alloys; and 3) stainless steel alloys. Alloy-specific differences in strength, ductility, and hardness generally determine which of these three alloys is used for a particular application or implant component. However, it is the high corrosion resistance of all three alloys, more than anything, that has led to their widespread use as load-bearing implant materials. These material properties of metals (Table 1) are due to the miraculous nature of the metallic bond, molecular microstructure, and elemental composition of metals.

Stainless steel alloys

Stainless steels were the first metals to be used in orthopedics in 1926. However, it was not until 1943, when ASTM 304 was recommended as a standard implant alloy material, that steels became reliable as an implant alloy. All steels are composed of iron and carbon, and may typically contain chromium, nickel, and molybdenum. Trace elements such as manganese, phosphorus, sulfur, and silicon are also present. Carbon and the other alloy elements affect the mechanical properties of steel through alteration of its microstructure.

Table 1 Approximate weight percent of different metals and other element within major orthopedic alloys

Alloy ^b	Ni	N	Co	Cr	Ti	Mo	Al	Fe	Mn	Cu	W	C	Si	V
Stainless steel														
(ASTM F138)	10-15.5	<0.5	a	17-19	a	2-4	a	61-68	a	<0.5	<2.0	<0.06	<1.0	a
CoCrMo alloys														
(ASTM F75)	<2.0	a	61-66	27-30	a	4.5-7.0	a	<1.5	<1.0	a	a	<0.35	<1.0	a
(ASTM F90)	9-11	a	46-51	19-20	a	a	a	<3.0	<2.5	a	14-16	<0.15	<1.0	a
(ASTM F562)	33-37	a	35	19-21	<1	9.0-11	a	<1	<0.15	a	a	a	<0.15	a
Ti alloys														
CPTi (ASTM F67)	a	a	a	a	99	a	a	0.2-0.5	a	a	a	<0.1	a	a
Ti-6Al-4V (ASTM F136)	a	a	a	a	89-91	a	5.5-6.5	a	a	a	a	<0.08	a	3.5-4.5
T45iNi	55	a	a	a	45	a	a	a	a	a	a	a	a	a

^aIndicates less than 0.05%.

^bAlloy compositions are standardized by the American Society for Testing and Materials (ASTM vol. 13.01).



The form of stainless steel most commonly used in orthopedic practice is designated 316LV (American Society for Testing and Materials (ASTM) F138). 316 classifies the material as austenitic, L denotes the low carbon content, and V denotes the vacuum under which it is formed. The carbon content must be kept at a low level to prevent carbide (chromium-carbon) accumulation at the grain boundaries.

Although the mechanical properties of stainless steels are generally less desirable than those of the other implant alloys (lower strength and corrosion resistance), stainless steels do possess greater ductility, indicated quantitatively by a threefold greater percentage of elongation at fracture when compared to other implant metals. This aspect of stainless steel allows it to remain popular as a material for cable fixation components in total knee arthroplasty, and as a low-cost alternative to Ti and Co alloys.

Cobalt-chromium alloys

Of the many Co-Cr alloys available, currently only two are predominantly used as implant alloys (Table 1). These are cobalt-chromium-molybdenum (CoCrMo), which is designated ASTM F-75 and F-76, and cobalt-nickel-chromium-molybdenum (CoNiCrMo), designated ASTM F-562. Other Co-alloys approved for implant use include one that incorporates tungsten (W) (CoCrNiW, ASTM F-90) and another that incorporates iron (CoNiCrMoWFe, ASTM F-563). Co-Ni-Cr-Mo alloys, which contain large percentages of Ni (25–37%), promise increased corrosion resistance, yet raise concerns of possible toxicity and/or immunogenic reactivity (discussed later) from released Ni. The biologic reactivity of Ni released from Co-Ni-Cr alloys is cause for concern under static conditions, and due to their poor frictional (wear) properties, Co-Ni-Cr alloys are also inappropriate for use in articulating components. Therefore, the dominant implant alloy used for total joint components is CoCrMo (ASTM F-75).

Cobalt alloys are generally cast into their final shape because they are susceptible to work-hardening at room temperatures. That is, the improvements in strength and hardness gained by cold-working are not worth the loss in fracture toughness. Thus, Co-Cr-Mo alloy hip implant components are predominantly manufactured using lost wax (investment) casting methods.

Although Co-Cr-Mo alloys are the strongest, hardest, and most fatigue-resistant of the metals used for joint replacement components, care must be taken to maintain these properties because finishing treatments may function to reduce them (Table 2). For example, sintering of porous coatings onto femoral or tibial TJA Co-Cr-Mo stems can decrease the fatigue strength of the alloy from 200–250 MPa to 150 MPa after heating (annealing) the implant at 1225°C.

Titanium alloys

Titanium alloys were developed in the mid 1940s for the aviation industry and were first used in orthopedics around the same time. Two post-World War II alloys—commercially pure titanium (CPTi) and Ti-6Al-4V—remain the dominant titanium alloys used in implants. Commercially pure titanium (CPTi, ASTM F67) is 98–99.6% pure titanium. Although CPTi is most commonly used in dental applications, the stability of the oxide layer formed on CPTi (and consequently its high corrosion resistance) and its relatively higher ductility (i.e., the ability to be cold-worked) compared to Ti-6Al-4V have led to the use of CPTi in porous coatings (e.g., fiber metal) of TJA components. Generally, joint replacement components (i.e., TJA stems) are made of Ti-6Al-4V (ASTM F-136) rather than CPTi because of its superior mechanical properties (Table 2).

Titanium alloys are particularly good for THA components because of their high corrosion resistance compared with stainless steel and Co-Cr-Mo alloys. A passive oxide film (primarily of TiO₂) protects both Ti-6Al-4V and CPTi alloys. Generally, Ti-6Al-4V has mechanical properties that exceed those of stainless steel, with a flexural rigidity less than stainless steel and Co-Cr-Mo alloys. The torsional and axial stiffness (moduli) of Ti alloys are closer to those of bone and theoretically provide less stress-shielding than do Co alloys and stainless steel. However, titanium alloys are particularly sensitive to geometrical factors, notch sensitivity in particular. This reduces the effective strength of a component by increasing the material's susceptibility to crack propagation through the component. Care is therefore taken in both the design geometry and fabrication of Ti alloy components. Perhaps the greatest drawback to Ti alloys is their relative softness compared to Co-Cr-Mo alloys (Table 2), and their relatively poor wear and frictional properties. Ti-6Al-4V is >15% softer than Co-Cr-Mo alloys and also results in significantly more wear than Co-Cr-Mo when used in applications requiring articulation, e.g., TKA or THA femoral heads. Ti alloys are thus seldom used as materials when hardness or resistance to wear are primary concerns.

New Alloys and Surface Coatings

The quest for new THA metal alloys with improved biocompatibility and mechanical properties is ongoing. The use of Ti alloys, Co-Cr-Mo alloys, or stainless steels in a specific application generally involves trade-offs of one desirable property for another. Examples of this are the sacrifice of chemical inertness for hardness (wear resistance), as is the case with Ti alloy for Co-Cr-Mo in TJA bearing surfaces, and the compromise

Table 2 Mechanical properties of implant alloys

Implant alloy	ASTM designation	Trade name and company (examples)	Elastic modulus (GPa)	Yield strength (MPa)	Ultimate strength (MPa)	Fatigue strength (endurance limit) (MPa)	Hardness (HVN)	Elongation at fracture (%)
Stainless steel								
	ASTM F138	Protosul S30-Sulzer	190	792	930	241-820	130-180	43-45
CoCrMo alloys								
	ASTM F75	Alivium-Biomet CoCrMo-Biomet	210-253	448-841	655-1277	207-950	300-400	4-14
		Endocast SIL-Krupp						
		Francobal-Benoist Girard						
		Orthochrome-Depuy						
		Protosul 2-Sulzer						
		Vinertia-Deloro						
		Vitalium Howmedica						
		Zimaloy-Zimmer						
		Cobalt-Chromium-Smith and Nephew						
	ASTM F90	Vitalium W-Howmedica	210	448-1606	1896	586-1220	300-400	10-22
	ASTM F562	HS251-Haynes Stellite	200-230	300-2000	800-2068	340-520	8-50 (RC)	10-40
		MP35N-Steel Corp.						
	ASTM 1537	TJA 1537-Allevac Metasul-Sulzer	200-300	960	1,300	200-300	41 (RC)	20
Ti alloys								
cpTi	ASTM F67	CSTi-Sulzer	110	485	760	300	120-200	14-18
		Pure Titanium-Smith and Nephew						
Ti-6Al-4V	ASTM 136	Isotan-Aesculap Werke	116	897-1034	965-1103	620-689	310	8
		Protosul 64WF-Sulzer						
		Tilastan-Waldemar Link						
		Tivaloy 12-Biomet						
		Tivanium-Zimmer						
		Titanium alloy-Smith and Nephew						
Ti5Al2.5Fe	^a	^a	100-110	780	860	300-725	310	7-13
Ti45Ni	^a	<i>Nitinol-Nitinol Medical Technology</i>	28-110	621-793	827-1172	<200	40-62 (RC)	1-60

RC = Rockwell hardness scale.

HVN = Vickers hardness number, kg/mm.

^aNo current ASTM standard.



of strength for ductility when using stainless steel instead of Ti and Co–Cr–Mo alloys for bone fixation cables. Although new alloys claim to be just that—new—they are often merely variations of the three categories of implant metals previously described (and already approved for use as implant materials). These improved alloys usually contain only the minor addition of new elements to protect assertions of substantial equivalence to existing ASTM- and FDA-approved alloys, thereby easing the burden of regulatory approval. These new alloys generally fall under one of four categories—titanium alloys, cobalt alloys, stainless steels, and (less approved) refractory metals.

New titanium alloys

One new group of Ti alloys put forward for orthopedic component use employs molybdenum at concentrations greater than 10%. The addition of Mo acts to stabilize the BCC (beta) phase at room temperature; these alloys are thus referred to as beta Ti alloys. These beta Ti alloys promise 20% lower moduli, which are closer to bone and thus provide better formability with maintenance of other mechanical properties typical of Ti-6-4.

Other attempts at improving traditional Ti-6Al-4V alloys seek to improve biocompatibility and mechanical properties by the substitution of vanadium (V, a relatively toxic metal) with other less toxic metals. Two such Ti alloys are Ti-5Al-2.5Fe and Ti-6Al-17Nb, which substitute Fe and Nb for V, respectively. These alloys have properties similar to traditional Ti-6-4, yet they claim higher fatigue strength and a lower modulus, thus enhancing bone-to-implant load transfer (Table 2).

New cobalt alloys

Some new cobalt alloys are identical in composition to traditional alloys, but use novel processing techniques to manipulate the microstructure of the implant materials to improve their mechanical properties. One such example recently patented, TJA-1537, although compositionally identical to ASTM F-75, claims enhanced wear resistance and fatigue strength through elimination of carbide, nitride, and second-phase particles (Allegheny Technologies). These particles normally form at the grain boundaries within a standard F75 CoCrMo alloy, and act to decrease wear and fatigue resistance. Other new Co alloys under development for use in orthopedics seek to improve biocompatibility by eliminating Ni, and improve mechanical properties by reducing carbon content, thus avoiding carbide precipitation at grain boundaries.

New stainless steels

The relatively poor corrosion resistance and biocompatibility of stainless steels compared to Ti and Co–Cr–Mo alloys provide incentive for development of improved stainless steels. New alloys such as BioDur 108 (Carpenter Technology Corp.) attempt to solve the problem of corrosion with an essentially nickel-free austenitic stainless alloy. This steel contains a high nitrogen content to maintain its austenitic structure, and boasts improved levels of tensile yield strength and fatigue strength, and improved resistance to pitting and crevice corrosion compared to nickel-containing alloys such as Type 316L (ASTM F138).

Refractory metals

Metals characterized as refractory include zirconium (Zr), tantalum (Ta), molybdenum (Mo), and tungsten (W) because of their relative chemical stability (passive oxide layer) and high melting points. Because of their stability, they are also highly corrosion-resistant. Corrosion resistance generally (although not always) correlates with biocompatibility because more stable metal alloys tend to be less chemically active and less participatory in biologic reactions. Additionally, some of these refractory metals (Mo and W) also possess high hardness properties, which makes them ideally suited for wear-resistance applications. However, these materials are difficult to form and machine. TJA components can be manufactured through the process of hot isostatic pressure using alloys such as copper to catalyze the binding of the refractory powder into a solid component. But removing binding metals that remain within components as grain boundary impurities is problematic and remains a technical barrier to the use of these alloys as implant materials.

CLINICAL CONCERNS OF MODERN TOTAL HIP IMPLANTS

Implant Degradation

Wear

The generation of wear debris and subsequent tissue reaction are central to the longevity of total joint replacements. In fact, particulate debris are currently extolled as the primary factor affecting the long-term performance of joint replacement prostheses, and as the primary source of orthopedic biomaterial degradation (based on overall implant mass or volume lost). Particulate debris generated by wear, fretting, or fragmentation induce an inflammatory reaction that, at a certain

point, promotes a foreign-body granulation tissue response that has the ability to invade the bone-implant interface. This commonly results in progressive local bone loss that threatens the fixation of both cemented and cementless devices alike.

Wear involves the loss of material in particulate form as a consequence of relative motion between two surfaces. Two materials placed together under load will only come into contact over a small area of the higher peaks or asperities. Electro-repulsive and atomic binding interactions occur at the individual contacts and, when the two surfaces slide relative to one another, these interactions are disrupted. This results in the release of material in the form of particles (wear debris). The particles may be lost from the system, be transferred to the counter face, or remain between the sliding surfaces. Three primary processes can cause wear: 1) abrasion, by which a harder surface plows grooves in the softer material; 2) adhesion, by which a softer material is smeared onto a harder counter surface, forming a transfer film; 3) fatigue, by which alternating episodes of loading and unloading result in the formation of subsurface cracks that propagate to form particles shed from the surface. During an initial wearing-in period, the relative motion of surfaces causes a large number of asperities to break, resulting in a high wear rate. After this initial period, the actual contact area increases and the two surfaces can be said to have adapted to one another. Over time, the wear rates decrease and eventually become linearly dependent on the contact force and sliding distance represented by the steady-state wear equation:

$$V = KFx$$

where V is volumetric wear (mm^3/year), K is a material constant of the material couple, F is the contact force (N), and x is the distance of relative travel (mm).^[10]

In general, the softer of two bearing materials will wear more rapidly. In the most popular joint replacement couple (a metal-on-polymer pair), the polymer wears almost exclusively, whereas in the case of a metal-on-ceramic pair, the metal wears to a greater extent. The in-vitro wear rates for the socket (in hip joint simulation studies) range from 0 to $3000 \text{ mm}^3/\text{year}$ depending on the type of couple employed and the environment (e.g., lubricant). A great deal of variability is associated with in-vivo wear rates of orthopedic biomaterials, which are generally measured by radiographic follow-up studies. Radiographic wear measurements are expressed as linear wear rates, whereas in-vitro studies generally report volumetric wear. Volumetric wear can be directly related to the number of wear particles released into periprosthetic fluids (typically on the order of 1×10^9 particles per year). The most common wear couple for hip and knee arthroplasty currently in use in the United

States is a cobalt-based alloy head (most commonly a Co–Cr–Mo alloy ASTM F-75) bearing on an ultrahigh molecular-weight polyethylene (UHMWPE) cup or liner. The wear rates of this couple are generally on the order of $0.1 \text{ mm}/\text{year}$, with particulate generation as high as 1×10^6 particles per step or per cycle. Clinically, implant wear rates have been found to increase with physical activity and weight of the patient, size of the femoral head (32 vs. 28 millimeters), roughness of the metallic counterface, and oxidation of the polyethylene.

Corrosion

Electrochemical corrosion occurs to some extent on all metallic surfaces, including THA implants. This is undesirable for two primary reasons: 1) the degradative process may reduce the structural integrity of the implant; and 2) the release of the products of degradation is potentially toxic to the host. Metallic biomaterial degradation may result from electrochemical dissolution phenomena or wear, but most commonly occurs through a synergistic combination of the two. Electrochemical processes include generalized corrosion uniformly affecting an entire surface, and localized corrosion affecting either areas of a device relatively shielded from the environment (crevice corrosion), or seemingly random sites on the surface (pitting corrosion). Additionally, these electrochemical and other mechanical processes interact, potentially causing premature structural failure or accelerated metal release (e.g., stress corrosion cracking, corrosion fatigue, and fretting corrosion), (Fig. 4). Corrosion of orthopedic biomaterials is a multifactorial phenomenon, dependent on five primary factors: 1) geometric variables (e.g., taper crevices in modular component hip prostheses), 2) metallurgical variables (e.g., surface microstructure, oxide structure



Fig. 4 Modular THA junction taper connection showing corrosion of taper connections. Macrograph of deposits of CrPO_4 corrosion particle products on rim of a modular cobalt-chrome femoral head.

and composition), 3) mechanical variables (e.g., stress and relative motion), 4) solution variables (e.g., pH, solution proteins, enzymes); and 5) the mechanical loading environment (e.g., degree of movement, contact forces). Current investigational efforts to minimize the corrosion of orthopedic biomaterials deal directly with the complex interactions of these factors.^[13]

Metal ion release

There has always been concern with the release of chemically active metal ions from implants into surrounding tissues. Although these metal ions may stay bound to local tissues, they may also bind to protein moieties and then be transported via the bloodstream or lymphatics to remote organs. Normal human serum levels of prominent implant metals are approximately: 1–10 ng/ml Al, 0.15 ng/ml Cr, <0.01 ng/ml V, 0.1–0.2 ng/ml Co, and <4.1 ng/ml Ti. Following total joint arthroplasty, levels of circulating metal (Al, Cr, Co, Ni, Ti, and V) have been shown to increase (Table 3).^[14–17] The values in this table show that following successful primary total joint replacement there are measurable elevations in serum and urine Co, Cr, and Ti. Transient elevations of urine and serum Ni have been noted immediately following surgery. However, urine and serum Al and V concentrations have not been found to be greatly elevated in patients with TJA (Table 3). In-vitro investigations indicate that specific metals in ionic form can affect the functionality of a variety of peri-implant cells such as fibroblasts, osteoclasts, macrophages, and lymphocytes within the ranges of metal concentrations reported to exist in periprosthetic tissue. Generally, the most toxic metal ions are Ni, Fe, Cu, Mn, and V, whereas others such as Na, Cr, Mg, Mo, Al, Ta, and Co demonstrate relatively less cellular reactivity in vitro.

Particle distribution

Generally, metal particles found disseminated beyond the periprosthetic tissue are <10 μm in size. Although variables influencing accumulation of wear debris in remote organs are not clearly identified, numerous case reports document the presence of metallic, ceramic, or polymeric wear debris from hip and knee prostheses in regional and pelvic lymph nodes. In fact, metallic wear particles have been detected in the para-aortic lymph nodes in up to 70% of patients with total joint replacement components (Fig. 5).^[18,19]

Hypersensitivity Responses

Dermal metal hypersensitivity is a phenomenon that purportedly affects about 10–15% of the population.

The incidence of metal sensitivity among patients with TJA implants is approximately 25% (roughly twice as high as that of the general population), and the average incidence of metal sensitivity among patients with a failed implant (i.e., one in need of revision surgery) is approximately 50–60%.^[20] Some patients with metallic orthopedic implants experience excessive eczema or other inflammatory reactions directly associated with implanted metallic materials.^[20,21] This reactivity has in some cases been shown to satisfy Koch's postulate (that exposure or implantation induces reaction that disappears upon removal of the metal stimulus and reappears upon reexposure).^[20–26] In addition to direct immunogenic responses, metal degradation products may mediate indirect immunologic effects that manifest as hypersensitivity-like responses in more predisposed (i.e., susceptible) individuals. It is likely that cases involving implant-related metal sensitivity have been underreported due to the difficulty of diagnosis. The biologic mechanisms by which in-vivo metal sensitivity occurs, remain incompletely characterized. Current efforts using Lymphocyte Transformation Testing are being made to establish whether metal hypersensitivity may elicit an overaggressive immune response that contributes to implant failure.^[20] These and other improvements in immunologic testing methods will likely enhance future assessment of patients susceptible to hypersensitivity responses.

Carcinogenesis

The carcinogenic potential of the metals used in TJA remains an area of concern and has been documented in animal studies. Whether metal release from orthopedic implants is carcinogenic remains conjectural because causality has not been established in human subjects and the actual numbers of tumors associated with orthopedic implants are likely underreported. However, the incidence of cancer at the site of implantation is rare.

FUTURE DEVELOPMENTS AND CONCLUSIONS

Current strategies to address the problem of THA failure are primarily aimed at decreasing the periprosthetic particulate burden and its effects. Recently there has been a great deal of innovation regarding stronger, more wear-resistant polyethylene. These more highly cross-linked UHMWPE polymers show demonstrable decrease in polyethylene wear with potential for less particulate-induced bioreactivity/osteolysis and therefore greater implant performance. In the same vein, femoral heads with diameters of 32 mm have been associated with increased volumetric polyethylene wear; to combat this, smaller 28 mm heads are currently

Table 3 Approximate average concentrations of metal in human body fluids with and without total joint replacement (ng/ml or ppb)

		Ti	Al	V	Co	Cr	Mo	Ni
Serum	Normal	2.7	2.2	<0.8	0.18	0.05–0.15	a	0.4–3.6
	THA	4.4	2.4	1.7	0.2–0.6	0.3	a	<9.1
	THA-F	8.1	2.2	1.3	a	0.2	a	a
	TKA	3.2	1.9	<0.8	a	a	a	a
	TKA-F	135.6	3.7	0.9	a	a	a	a
Urine	Normal	<1.9	6.4	0.5	a	0.06	a	a
	TJA	3.55	6.53	<0.4	a	0.45	a	a
Synovial fluid	Normal	13	109	5	5	3	21	5
	TJA	556	654	62	588	385	58	32
Joint capsule	Normal	723	951	122	25	133	17	3996
	TJA	1540	2053	288	1203	651	109	2317
	TJA-F	19173	1277	1514	821	3329	447	5789
Whole blood	Normal	17	13	6	0.1–0.1.2	2.0–4.0	0.5–1.8	2.9–7.0
	TJA	67	218	23	20	110	10	29

Concentrations of metal in body tissue of humans with and without total joint replacements ($\mu\text{g/g}$)

		Cr	Co	Ti	Al	V
Skeletal Muscle	Normal	<12	<12	a	a	a
	TJA	570	160	a	a	a
Liver	Normal	<14	120	100	890	14
	TJA	1130	15200	560	680	22
Lung	Normal	a	a	710	9830	26
	TJA	a	a	980	8740	23
Spleen	Normal	10	30	70	800	<9
	TJA	180	1600	1280	1070	12
Pseudocapsule	Normal	150	50	<65	120	<9
	TJA	3820	5490	39400	460	121
Kidney	Normal	<40	30	a	a	a
	TJA	<40	60	a	a	a
Lymphatic tissue	Normal	690	10	a	a	a
	TJA	690	390	a	a	a
Heart	Normal	30	30	a	a	a
	TJA	90	280	a	a	a

Normal: Subjects without any metallic prosthesis (not including dental).

THA: Subjects with well-functioning total hip arthroplasty.

THA-F: Subjects with a poorly functioning total hip arthroplasty (needing surgical revision).

TKA: Subjects with well-functioning total knee arthroplasty.

TKA-F: Subjects with a poorly functioning total knee arthroplasty (needing surgical revision).

TJA: Subjects with a well-functioning total joint arthroplasty.

TJA-F: Subjects with a poorly functioning total joint arthroplasty (needing surgical revision).

^aNot tested.

(From Refs. [14–17].)

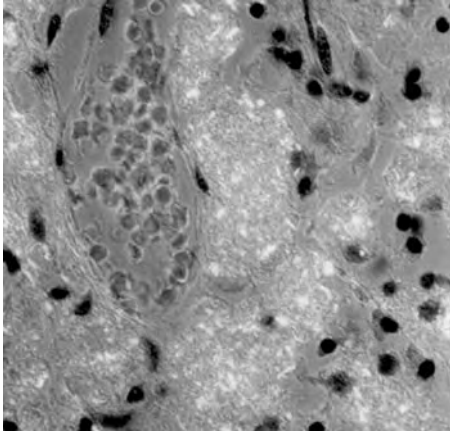


Fig. 5 Polarized light micrograph (190 ×) of paraaortic lymph node demonstrates abundance and morphology of birefringent polyethylene particles within macrophages.

extolled as more biocompatible. Manufacturing flaws such as fusion defects and foreign body inclusions have also been suggested as contributing to adverse polyethylene wear properties. The elimination of polyethylene is another approach being investigated clinically in various centers. With the realization that early problems may have been related to the design and not the articulation, there has been renewed interest in the application of metal-metal and ceramic-ceramic bearings. Future designs that attempt to reduce wear include improved tolerances between polyethylene inserts and their metal backing, improved surface finish on metallic concave surfaces, secure locking mechanisms, and the avoidance of holes on the convex portion of the acetabular prosthesis.

Electrochemical corrosion of orthopedic implants remains a significant clinical concern. Although the freely corroding implants used in the past have been replaced with modern corrosion-resistant super alloys, deleterious corrosion processes have been observed in certain clinical settings. Attention to metallurgical processing variables, tolerances of modular connections, surface-processing modalities, and appropriate material selection can diminish corrosion rates and minimize the potential for adverse clinical outcomes. A need to further investigate the mechanical-electrochemical interactions of metal oxide surfaces in implants persists. Characterization of the stresses and motion needed to fracture passivating oxide films and the effects of repeated oxide abrasion on the electrochemical behavior of the interface (and ultimately the implant) remain avenues of active investigation.

Although beyond the scope of this article and generally a realm inaccessible to nonclinically-oriented biomedical engineers, it is worthwhile to note THA implant performance is intimately related to the skill of the

surgical technician. Improvements to surgical methods are significant in that greater ease, standardization, and timeliness of properly executed rigid fixation will more adequately facilitate bone ingrowth and performance by minimizing motion between the bone and implant, resulting in reduced debris and longer implant life. For example, methodological improvements in PMMA bone cement installation to facilitate a uniform-thickness cement mantle will diminish the likelihood of cement mantle defects (debris generation), a cause of osteolysis.

THA implants are one of the great success stories of modern medicine and biomedical engineering, with well over 300,000 performed every year, allowing mobility, independence, and a quality of life that would otherwise be drastically compromised. However, in certain situations mechanical and electrochemical degradative processes can compromise the clinical outcome and potentially lead to adverse local or systemic effects. Continued awareness and investigation of these issues will enhance understanding of implant degradation phenomena and lead to improvements in biomaterial and implant performance.

ARTICLES OF FURTHER INTEREST

Alumina; Bone Cement; Corrosion; Diamond and Diamond-Like Carbons; Knee Joint Replacement; Orthopedic Biomaterials; Titanium and Its Alloys; Ultra-High Molecular Weight Polyethylene (UHMWPE); Zirconia Ceramics

REFERENCES

1. Walker, P.S. *Human Joints and Their Artificial Replacements*; Charles C. Thomas: Springfield, 1978.
2. Stillwell, W.T. *The Art or Total Hip Arthroplasty*; Grune & Stratton, Inc.: Orlando, 1987.
3. Charnley, J. *Low Friction Arthroplasty of the Hip, Theory and Practice*; Springer-Verlag: Berlin, 1979.
4. Venable, C.S.; Stuck, W.G.; Beach, A. The effects on bone of the presence of metals; based upon electrolysis. An experimental study. *Ann. Surg.* **1937**, *105*, 917.
5. Harris, W.H. Traumatic arthritis of the hip after dislocation and acetabular fractures: Treatment by mold arthroplasty. An end-result study using a new method of result evaluation. *J. Bone Jt. Surg., Am.* **1969**, *51*, 737-755.
6. Moore, A.T. Metal hip joint: A case report. *J. Bone Jt. Surg., Am.* **1943**, *25*, 688.
7. Wiles, P. The surgery of the osteoarthritic hip. *Br. J. Surg.* **1953**, *45*, 488.
8. McKee, G.K.; Watson-Farrar, J. Replacement of the arthritic hips to the McKee-Farrar replacement. *J. Bone Jt. Surg., Br.* **1943**, *48*, 245.

9. Charnley, J. The bonding of prosthesis to bone by cement. *J. Bone Jt. Surg., Br.* **1964**, *46*, 518.
10. Wilson, J.N.; Scales, J.T. Loosening of total hip replacements with cement fixation. Clinical findings and laboratory studies. *Clin. Ortop.* **1970**, *72*, 145–160.
11. Charnley, J. Anchorage of the femoral head prosthesis to the shaft of the femur. *J. Bone Jt. Surg., Br.* **1960**, *42*, 28.
12. Ring, P.A. Complete replacement arthroplasty of the hip by the ring prosthesis. *J. Bone Jt. Surg., Br.* **1968**, *50*, 720.
13. Black, J. *Orthopaedic Biomaterials in Research and Practice*; Churchill Livingstone: New York, 1988.
14. Jacobs, J.J.; Gilbert, J.L.; Urban, R.M. Corrosion of Metallic Implants. In *Advances in Orthopaedic Surgery Vol. 2*; Stauffer, R.N., Ed.; Mosby: St. Louis, 1994.
15. Jacobs, J.J.; Skipor, A.K.; Urban, R.M.; Black, J.; Manion, L.M.; Starr, A.; Talbert, L.F.; Galante, J.O. Systemic distribution of metal degradation products from titanium alloy total hip replacements: An autopsy study. *Trans. Orthop. Res. Soc. (New Orleans)* **1994**, 838.
16. Jacobs, J.J.; Silvertown, C.; Hallab, N.J.; Skipor, A.K.; Patterson, L.; Black, J.; Galante, J.O. Metal release and excretion from cementless titanium alloy total knee replacements. *Clin. Ortop.* **1999**, *358*, 173–180.
17. Jacobs, J.J.; Skipor, A.K.; Patterson, L.M.; Hallab, N.J.; Paprosky, W.G.; Black, J.; Galante, J.O. Metal release in patients who have had a primary total hip arthroplasty. A prospective, controlled, longitudinal study. *J. Bone Jt. Surg., Am.* **1998**, *80*, 1447–1458.
18. Urban, R.B.; Jacobs, J.; Gilbert, J.L.; Rice, S.B.; Jasty, M.; Bragdon, C.R.; Galante, G.O. Characterization of Solid Products of Corrosion Generated by Modular-Head Femoral Stems of Different Designs and Materials. In *STP 1301 Modularity of Orthopedic Implants*; Marlowe, D.E., Parr, J.E., Mayor, M.B., Eds.; ASTM: Philadelphia, 1997.
19. Urban, R.M.; Jacobs, J.J.; Tomlinson, M.J.; Gavrilovic, J.; Black, J.; Peoc'h, M. Dissemination of wear particles to the liver, spleen, and abdominal lymph nodes of patients with hip or knee replacement. *J. Bone Jt. Surg., Am.* **2000**, *82*, 457–476.
20. Urban, R.M.; Jacobs, J.J.; Gilbert, J.L.; Galante, J.O. Migration of corrosion products from modular hip prostheses. Particle microanalysis and histopathological findings. *J. Bone Jt. Surg., Am.* **1994**, *76*, 1345–1359.
21. Merle, C.; Vigan, M.; Devred, D.; Girardin, P.; Adessi, B.; Laurent, R. Generalized eczema from vitallium osteo-synthesis material. *Contact Dermatitis* **1992**, *27*, 257–258.

Implanted Medical Devices, Power Sources For

Erik R. Scott
Craig L. Schmidt
Paul M. Skarstad

Energy and Component Center, Medtronic, Inc., Brooklyn Center, Minnesota, U.S.A.

INTRODUCTION

Electrically powered implanted devices treat and diagnose a variety of conditions, providing greatly enhanced quality of life to those who use them. Most of these implants derive their power from batteries that are neither rechargeable nor replaceable. Therefore, the battery's finite energy supply often limits the service life of an implant, and battery depletion necessitates a surgical procedure and replacement of the device. It is desirable, from the standpoint of patient risk and expense to society, to reduce the frequency of such replacement events. Thus, the power source is a critical component of an implanted device.

For any electrical device (implanted or not), power and energy requirements are the main criteria for selection of a power source. However, the requirements of implanted power sources do not end with power and energy. For example, implanted devices must be of small enough size and appropriate shape that they can be continuously worn with minimal intrusion on activity and comfort. Because the battery is often the largest component of a device, having a shape that packages efficiently within the device is crucial. However, the most important requirement of implanted power sources is to provide safe and reliable operation. This is especially true for devices that are critical to human health or life. This article will briefly describe the function, design, and applications of power sources for a majority of devices commercially available today: batteries based on lithium chemistry.

HISTORY OF POWER SOURCES FOR IMPLANTABLE MEDICAL DEVICES

The first successful use of an electrically powered implantable device came in the early 1960s with the introduction of the implantable cardiac pacemaker.^[1,2] Throughout the 1960s and early 1970s, the zinc/mercuric oxide battery was the power source used in most such devices.^[3,4] Although this battery system enabled the birth of implantable cardiac pacing, it had a number of drawbacks. These included: generation of hydrogen gas, which prevented hermetic sealing of the

pacemaker; little warning of depletion at end-of-service; and significant self-discharge at body temperature. Although significant efforts were made to overcome these limitations, two different approaches were made in the late 1960s to develop other power sources with higher longevity than zinc/mercuric oxide batteries. These include nuclear power sources that used plutonium to heat thermoelectric generators^[5] and cadmium/nickel oxide (NiCAD) secondary batteries that could be recharged transcutaneously by electromagnetic induction.^[6] Both of these approaches were successful in creating reliable systems with significantly improved longevity. However, both also had serious drawbacks. These drawbacks included regulatory requirements for tracking nuclear power sources and, in particular, for ensuring proper disposal of explanted devices. For the NiCAD system, the need to recharge placed a significant burden on the patient in managing the device.

The advent in the early 1970s of highly reliable, stable, and small primary batteries based on lithium as the anode material revolutionized cardiac pacing and obviated the previously mentioned approaches. The first lithium-powered pacemaker was implanted in 1972, using a lithium/iodine battery invented by Schneider and Moser^[7,8] and improved by Wilson Greatbatch and coworkers.^[9,10] Although the inherent power capability of this system was low, it was well matched to the power requirements of cardiac pacing. Moreover, the system provided a deliverable energy density as high as 1 Wh/cm³, more than a twofold increase in deliverable energy density over alternatives available at the time. The absence of gas evolution meant both battery and pacemaker could be hermetically sealed. In addition, the system provided high reliability, and had a versatile shape factor.

Several other lithium-based battery systems were developed for use in cardiac pacemakers within a few years of the lithium/iodine system. These included lithium/thionyl chloride, lithium/manganese dioxide, lithium/silver chromate, lithium/cupric sulfide, and others.^[11-13] All had energy densities close to that of lithium/iodine. However, these battery systems all gave way to the lithium/iodine battery system, which has remained overwhelmingly the most frequently used

power source for this application since the early 1980s. It is likely that in the 30-year history of the lithium/iodine battery, more than five million cardiac pacemakers have been implanted with this power source.

In the 1980s new applications of implantable pulse generators were developed that required higher power than lithium/iodine batteries could produce efficiently. These included implantable neurological stimulators and implantable drug delivery systems. Most devices of these types have used batteries based on lithium/thionyl chloride. Recently, these devices have turned increasingly to lithium batteries based on organic electrolytes with higher deliverable energy densities at application rates. These battery systems include lithium/silver vanadium oxide, lithium/carbon monofluoride, and lithium with blends of carbon monofluoride and silver vanadium oxide.

The 1980s and 1990s also saw development of implantable cardioverter-defibrillators (ICDs) that treat ventricular fibrillation, a major cause of death.^[14] These devices monitor and analyze the electrical activity of the heart and can function as cardiac pacemakers, drawing tens of microwatts of power. If an ICD detects a potentially lethal arrhythmia, the battery must deliver power in the range 5–10 W to charge high-voltage output capacitors quickly so that a life-saving shock can be delivered to the heart within a few seconds to terminate the arrhythmia. The requirements of long-term stability at low power coupled with the ability instantaneously to deliver power as much as five orders of magnitude higher place severe demands on a battery system. The earliest implantable defibrillators used lithium/vanadium oxide batteries as power sources.^[15] This system was supplanted in the late 1980s by lithium/silver vanadium oxide (Li/SVO), developed for the application by Wilson Greatbatch, Ltd.^[16,17]

The revolution in readily accessible computational power in 1980s and 1990s fostered development of accurate, physically based predictive models of battery performance to aid in the development of optimal custom batteries for implantable devices. Such models have been described in some detail for the lithium/iodine pacemaker batteries^[18] and for lithium/silver vanadium oxide ICD batteries.^[19] Development of these models required quantitative understanding of underlying chemical processes in these battery systems.

As pacemakers were developed in the 1990s that included microprocessors, memory, sensors of various types, and long-range telemetry, it became increasingly difficult to meet the demand for pulse power using lithium/iodine batteries. Moreover, new therapies were developed with higher power requirements, including treatment of atrial fibrillation and congestive heart failure. New medium-rate power sources were needed that had energy density comparable to that of lithium/iodine, with up to two orders of magnitude greater

pulse power capability. Lithium batteries with silver vanadium oxide cathodes have been used, but they fall short of matching lithium/iodine for energy density. Carbon monofluoride cathodes have also been used. These have comparable energy density and twice the rate capability of lithium/iodine. However, the medium-rate system that has come to be used most frequently uses a blend of carbon monofluoride and silver vanadium oxide as the active cathode material.^[20] Batteries based on this cathode system have improved energy density, rate capability, and end-of-service characteristics, compared with carbon monofluoride alone. They have found application in devices that treat heart failure and atrial fibrillation and are likely to find application in other devices.

BASICS OF BATTERIES

Electrochemical Power Generation

A battery consists of one or more electrochemical cells that convert chemical energy into electrical energy. For simplicity in this article, the term battery implies a single cell unless otherwise noted. The three most fundamental electrical characteristics of a battery are the amount of power that it delivers, its operating voltage, and the energy it delivers over its lifetime. In the battery industry, power and energy are frequently expressed in watts (W) and watt-hours (Wh), respectively. Power, voltage, and energy are determined by the relationship between battery chemistry, design, and the battery's interaction with the device circuit. To understand this relationship, it is necessary to briefly review the basic electrochemistry of batteries. For more detailed treatments of general battery science, please see introductory chapters in references.^[21–24]

Electrochemical cells contain at least two electrodes in contact with an ion-conducting electrolyte. Electron transfer reactions (also known as redox reactions) at the two electrodes of a cell result in an electromotive force that is the basis for a battery's generation of electrical energy. During discharge, oxidation occurs at the negative electrode (anode) while reduction occurs at the positive electrode (cathode). Lithium is an attractive anode material due to its highly electro-positive nature, large specific capacity, and stability in a large number of organic electrolytes. Lithium batteries are made using a variety of cathode materials, depending on battery application. An illustrative example is the lithium/iodine (Li/I₂) battery, which is presently used to power bradycardia pacemakers.^[25] The anode and cathode reactions are:



and



The total cell reaction is the sum of these two half-cell reactions:



It should be clear from Eqs. 1a-c that the rates of the reduction and oxidation reactions are perfectly balanced; the number of electrons liberated at the anode is identical to the number consumed at the cathode. This flow of electrons between the electrodes is the current that powers a device.

In addition to electronic flow, discharge also involves transport of ionic species from one electrode to the other, to preserve electroneutrality. The movement of ions takes place through the electrolyte. For a lithium battery, the electrolyte may consist of an organic solvent system into which a suitable lithium salt is dissolved, or it may be a solid-state Li^+ conductor. Aqueous solvents are incompatible with Li batteries due to the high reactivity of Li with water. Two desirable properties of an electrolyte are stability over the range of potentials found in the battery, and sufficiently high ionic conductivity to support device application currents without causing significant energy losses. Another function of the electrolyte layer is to prevent electrical contact between anode and cathode, which would otherwise cause an electrical short that could quickly deplete the battery. To prevent such shorts, liquid electrolyte battery systems typically incorporate the liquid into a porous polymer film, known as a separator.

Voltage and Current: Chemical Thermodynamics and Kinetics

A difference in potential energy (i.e., voltage) between electrons released at the anode and those consumed at the cathode determines the maximum amount of work that the electrons can do in a circuit. The electrical potential is determined by the change in molar free energy (ΔG_{rxn}) of the whole cell reaction according to:

$$V_{oc} = \frac{-\Delta G_{rxn}}{nF} \quad (2)$$

where V_{oc} is the open circuit (i.e., zero current) potential, n is the number of equivalents of charge passed per mole of species reacted, and F is Faraday's constant (96485 C/mol).

When a battery experiences a current load, the operating voltage is always lower than the open-circuit value. A simple expression of this effect is given in Eq. 3. Here, the battery voltage (V_b) under load is given as the open-circuit value (V_{oc}), minus the ohmic drops resulting from current flow (I) through the sum of the series internal resistive elements (R_i):

$$V_b = V_{oc} - I \sum_i R_i \quad (3)$$

It is often practical to use ideal resistors to describe the internal resistive elements of a battery. However, it should be understood that while batteries do include some components (e.g., current collectors and electrolyte) that closely approximate ideal resistors, other elements of batteries do not always behave in a linear and stationary fashion. These so-called nonohmic resistances include charge transfer resistance and concentration polarization resistance. The charge transfer resistance is related to the reaction kinetics at an electrode that is described by the Butler-Volmer equation:^[26]

$$i = i_o \left[\exp \frac{-\alpha F \eta}{RT} - \exp \frac{(1-\alpha) F \eta}{RT} \right] \quad (4)$$

where i is the current density, i_o is the exchange current density, and α is the charge transfer coefficient. The overpotential, η , is a measure of the deviation of the electrode potential from equilibrium. In the low overpotential range (less than a couple tenths of a volt, corresponding to low to moderate power output), the current-voltage relationship is effectively linear, resulting in a constant value of the charge transfer resistance. However, at higher power levels, the I-V relationship becomes exponential, resulting in a charge transfer resistance that decreases with applied current. Another nonohmic resistance results from concentration polarization, which is the buildup and depletion of ions within the electrolyte due to mass transfer limitations within the electrolyte or the electrodes themselves. Unlike the charge transfer resistance, the concentration polarization resistance increases with current, as well as with time of operation. Despite being nonohmic in a general sense, both charge transfer resistance and concentration polarization resistance may often be practically treated as ohmic over the normal operating range of a given device.

Another attribute of the electrical response of batteries is capacitance, which results from charge accumulation (both electronic and ionic) at interfaces. The unique combination of resistances and capacitances of a battery comprises its impedance, which, along with the open-circuit voltage, determines a battery's response to time-varying power demands.

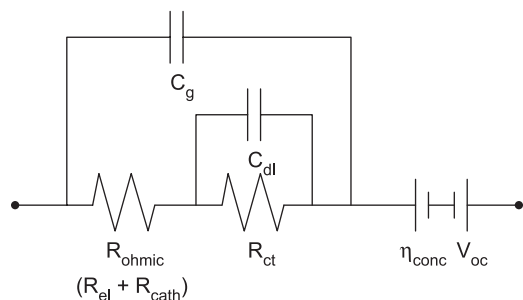


Fig. 1 Equivalent circuit for a Li/I₂ battery. R_{ohmic} is the ohmic resistance due to the electrolyte and cathode, C_g is the geometric capacitance, R_{ct} is charge transfer resistance, and C_{dl} is the double-layer capacitance. The concentration polarization is represented as an overpotential η_{conc} rather than a resistance. (Adapted from Ref. [27].)

Characterization of impedance in the form of a battery equivalent model allows accurate prediction of battery performance in a device. An example of an equivalent circuit used to model a Li/I₂ battery is shown in Fig. 1.^[27] An important aspect of battery impedance is that it is usually not constant over a battery's lifetime. The degree by which impedance changes depends primarily on cell chemistry. For a lithium/silver vanadium oxide (Li/SVO) cell, the resistance increases by about a factor of two,^[19] while for a Li/I₂ cell, the increase is more than two orders of magnitude.^[28]

Power, Capacity, and Longevity

Electrical power is the product of voltage and current. Current demands on medical batteries can range by about six orders of magnitude, from about 10 μA to 10 A, depending on device. By contrast, the operating voltage covers a narrow range, between about 2 and 4 V. Because voltage is nearly constant compared with current, power is roughly proportional to current, and the two terms are loosely interchangeable for describing rate of battery discharge.

The power capability of a battery is determined by the battery voltage and internal resistance. If the internal resistance is approximated as an ohmic quantity, the peak power that can be delivered by the battery is

$$P_{\text{max}} = \frac{V_{\text{oc}}^2}{4R_{\text{int}}} \quad (5)$$

where V_{oc} is the open-circuit voltage of the battery and R_{int} is the internal resistance. To achieve this peak power in practice requires a close matching of circuit impedance with that of the battery. However, most devices operate below this peak level.

A battery's delivered capacity is the integral of current over its total time of operation, and is

commonly expressed in ampere-hours (Ah). The maximum theoretical capacity is equal to the lesser of the stoichiometric capacities of the cathode or anode. Stoichiometric capacity for an electrode is computed from the formula:

$$Q_s = \frac{nFm}{3.6M} \quad (6)$$

where n is the number of equivalents of charge passed for mole of active species reacted, F is Faraday's constant (96485 C/eq), m is the mass of active material present (g), and M is the active formula weight (g/mol). However, the practical deliverable capacity is always less than Q_s . Factors that reduce the deliverable capacity include losses due to parasitic reactions (self-discharge), imperfect utilization of active material, and load-dependent depression of the battery voltage.

Battery capacity determines the longevity of a device, which is the length of time over which the operating voltage of its battery remains above that needed to operate the device. Both longevity and capacity are directly proportional to the energy delivered by a battery, which can be computed from the area under a discharge curve (voltage versus capacity). Discharge curves are given for several battery chemistries in Fig. 2. A general feature of most of the chemistries shown is that the voltage reduces as the current drain (i.e., power) increases, due to resistive and polarization losses. For most chemistries, this results in a concomitant decrease in deliverable capacity. Two exceptions to this are lithium/thionyl chloride (Fig. 2B) and, to a lesser extent, the lithium/iodine (Fig. 2A) systems. Both of these experience significant capacity loss due to self-discharge at lower currents.

Primary Versus Secondary Batteries

Primary batteries can be discharged only once. Secondary batteries may be recharged. The ability of a battery to be recharged depends on whether the cell reactions are practically reversible. Many redox systems, while chemically reversible, are not reversible in a practical sense due to phase changes that occur during charge and discharge. Therefore, the most highly energetic batteries (on a per-cycle basis) are not rechargeable. However, because secondary batteries can be recharged many times, they deliver much more total energy over their lifetime than do primary cells. As secondary cells age, they experience a gradual decrease in the per-cycle energy. Because of this wear-out process, known as capacity fade, even a device with a rechargeable battery will have a limited service life.

The majority of implanted batteries today are primary cells. This is because in most applications,

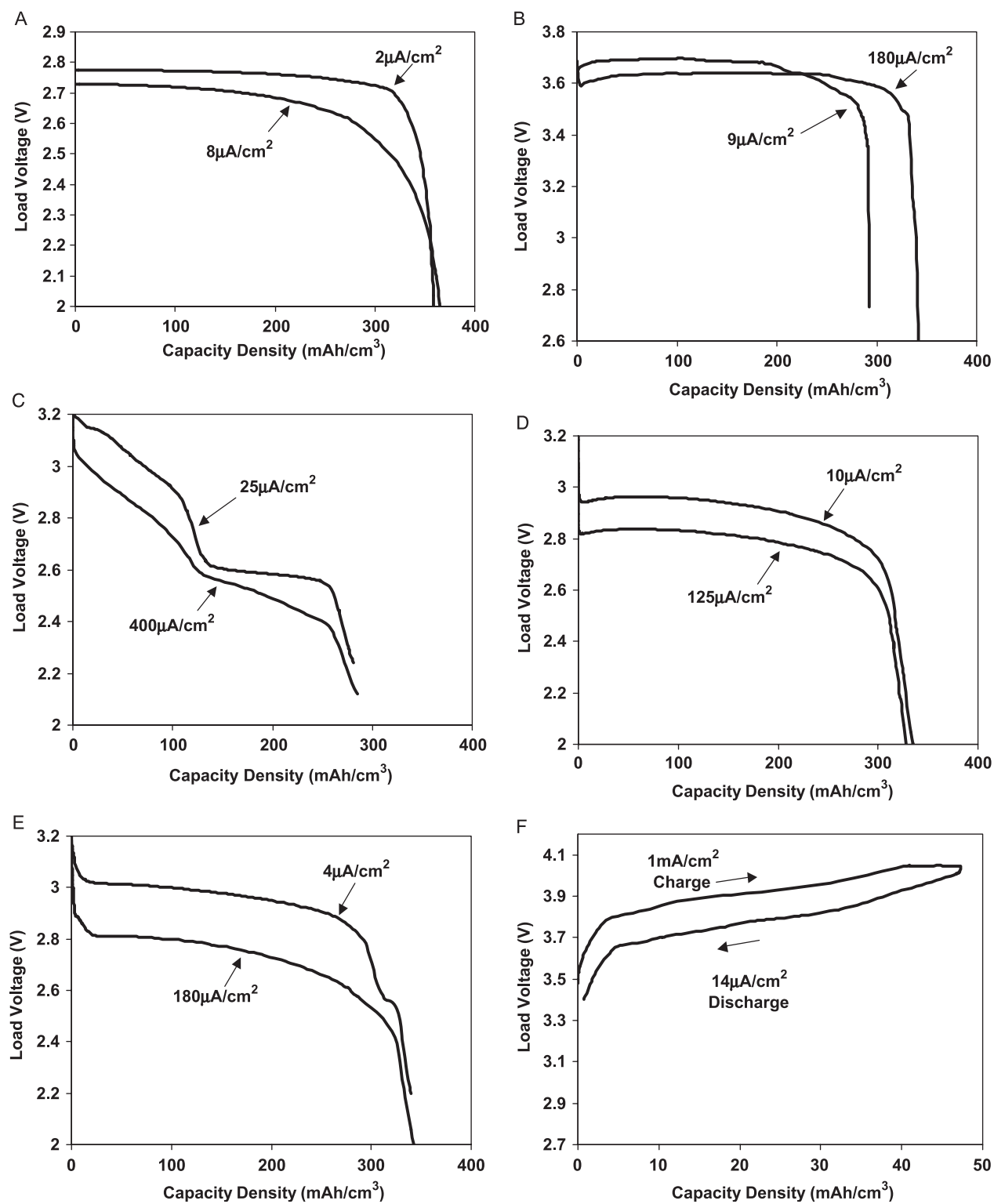


Fig. 2 Discharge curves for five primary battery chemistries and one rechargeable chemistry. Capacity is normalized to battery volume (mAh/cm³) and to electrode area (μA/cm²). A) Li/I₂ at 2 and 8 μA/cm²; B) Li/SOCl₂ at 9 and 180 μA/cm²; C) Li/SVO at 26 and 400 μA/cm²; D) Li/CF_x at 10 and 125 μA/cm²; E) Li/CF_x/SVO hybrid cathode at 4 and 180 μA/cm²; F) Li ion rechargeable, charged at 1 mA/cm² and discharged at 14 μA/cm².

primary batteries have sufficiently high energy density that they provide acceptable longevity in reasonably sized packages, without the overhead that is associated with a rechargeable device. This overhead includes the volume associated with the recharge circuit as well as the inconvenience of needing to recharge on a regular basis. One challenge with implanting a rechargeable power source is ensuring that the recharge schedule provides continuity of therapy for the patient. This concept is most important if the device sustains one's life. Two examples of secondary batteries in critical applications are the left ventricular assist device (LVAD) and the totally implantable artificial heart. Both of these devices receive operating power (and battery recharge power) nearly continuously from an external appliance, and the batteries serve a backup function.^[29,30]

Important Aspects of Implanted Batteries

Implantable batteries must meet a set of requirements that, taken in their entirety, distinguish them from all other batteries. Most important among these is safe, reliable (i.e., predictable), and stable performance at 37°C, over a period that can last up to 10 years. For devices whose function is critical to life or health, the requirement of predictability also mandates some form of a detectable warning that the battery's end-of-service is approaching. This warning enables continuation of therapy by giving the clinician a window of opportunity to replace the device before battery depletion occurs. Other requirements include containment of reactive and potentially corrosive materials, volumetrically efficient packaging within a device, and absence of excessive swelling during use. The above requirements translate into a set of physical characteristics that generally include the following:

- The materials, design, and device requirements have been characterized to the point of creating scalable, predictive models of performance longevity under application conditions.
- As battery depletion approaches, there is a change in some measurable intrinsic property (e.g., a slope in voltage or impedance) that can be interpreted as an indication of a certain minimum amount of capacity remaining.
- The materials are stable and are sealed within a hermetic package.
- The combination of chemical system and mechanical design results in adequate power density, energy density, and shape factor, when configured into the device's allotted size and shape.
- There is an acceptably low rate of self discharge.

MECHANICAL DESIGN

Power capability and energy density are two critically important characteristics of an implantable battery. Both are determined by the choice of battery chemistry and mechanical design. Unfortunately, optimization of a mechanical design for increased power typically results in decreased energy density, and vice versa.

As was mentioned previously, a battery's power capability is proportional to the square of its open-circuit voltage and inversely proportional to its internal resistance (Eq. 5). Voltage is determined by the battery chemistry, but the internal resistance is largely influenced by the mechanical design, varying inversely proportional to the electrode surface area. High-power batteries, incorporating electrodes with large surface area, contain a greater proportion of inert material (e.g., current collector and separator) than do high-energy batteries. This leads to an inverse relationship between power capability and energy density for a given battery chemistry. Thus, the mechanical design of a battery must be optimized to meet the specific power requirements for a given application without unduly reducing energy density.

Implantable batteries can be generally categorized into one of two general design types. Batteries that have low- to moderate-power requirements are typically designed similar to Fig. 3, while batteries requiring high power are similar to Fig. 4. There are certain characteristics, however, that are common to both designs. The batteries are hermetically sealed to

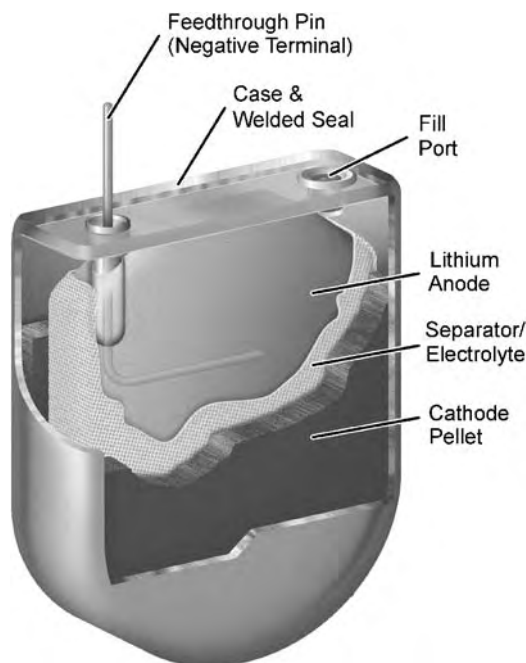


Fig. 3 Schematic of low-to-moderate power (single plate) battery design.

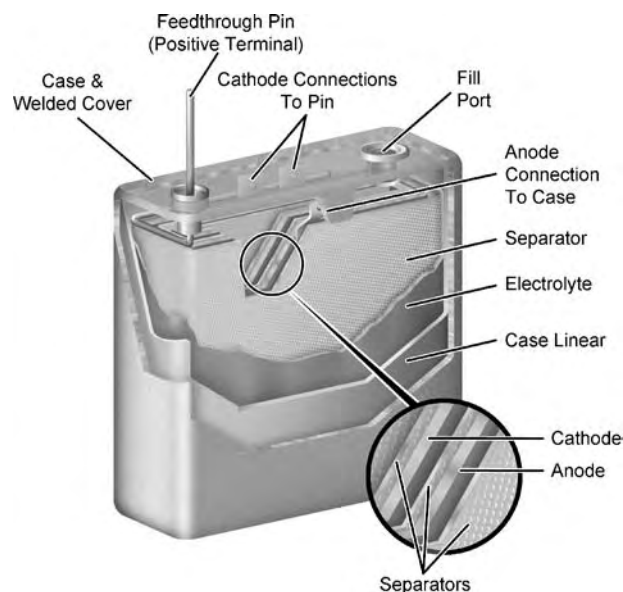


Fig. 4 Schematic of coiled prismatic design used for high-power primary and lithium ion designs.

prevent leakage of electrolyte and possible damage to other components of the implanted device. The case and cover are made of corrosion resistant materials, typically stainless steel or titanium. The electrical feedthrough consists of a metal ferrule welded to the battery cover, a corrosion resistant glass insulator, and a metal pin. The cover (with feedthrough) is welded to the battery case, and the electrolyte (or in some cases, a liquid cathode material) is filled through a small port located in the cover. A metal button is welded over the fill-port to complete the hermetic seal.

The low- to moderate-power batteries are characterized by relatively thick electrodes with surface areas ranging from about 3 to 15 cm² (see Fig. 3). The example shown has a central anode with a cathode on both sides (a case-positive design). However, the battery can also be constructed with a central cathode (case-negative) or with a single anode and cathode (either case polarity). For simplicity, these designs are all referred to as single plate. The cathode is often a pressed powder pellet. The active cathode material is often mixed with a conductivity enhancer (usually carbon) and a binder material to facilitate the formation of an electrically-conductive pellet. Alternatively, the cathode can be a material that is added as a liquid and allowed to solidify in situ (e.g., the cathode of a Li/I₂ battery, Fig. 5). In all cases, the anode is lithium metal. A number of different methods are used to establish electrical contact with the cathode (i.e., current collection). These include direct contact of the cathode with the case, pressing the cathode into a metal cup or ring that is then welded to the case or

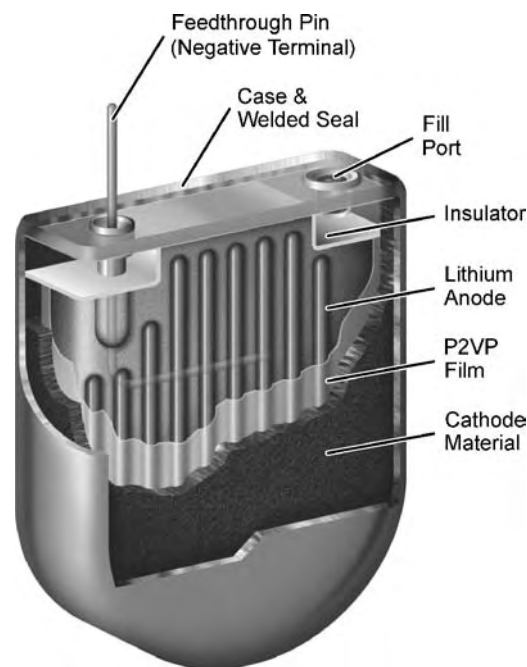


Fig. 5 Schematic of Li/I₂ battery.

feedthrough pin, or by embedding in the cathode a metal grid that is then welded to the case or feedthrough pin.^[31,32] Since lithium is highly conductive, current collection is easily accomplished by embedding the feedthrough pin in the lithium (central anode design), or by pressing onto the lithium a metal grid with a tab that can be welded to either the case or the feedthrough pin. With the exception of the lithium/iodine battery, all of these designs use a thin, porous separator material that prevents direct electrical contact of the anode and cathode. A major advantage of the single plate design is that the profile of the plate can be any shape. An especially popular cell shape is the so-called D-shape, which uses a deep-drawn can with a rounded bottom (see Fig. 3). This shape packages efficiently in the rounded profile of an implanted device.

The coiled prismatic design of Fig. 4 is used for high-power primary batteries, which are characterized by very thin electrodes whose surface areas range from about 40 to 90 cm².^[33] The cathode is formed from a powdered mixture of active material plus conductivity enhancer and binder. This mixture is pressed onto a metal grid current collector. The function of the cathode current collector is twofold: it serves as a substrate for mechanical attachment of the active material and it uniformly distributes the electric field across the large cathode area. The current collector has one or more tabs that are welded to either the case or the feedthrough pin. The anode is lithium and incorporates a metal grid current collector similar to

the cathode. With the high conductivity of lithium, however, the primary function of the current collector is to preserve electrical continuity as the lithium is consumed during discharge. A microporous separator material is placed between the electrodes, and this assembly is then wound into a coil and placed into a case.

Rechargeable lithium ion cells have a coiled construction that is similar to the high-power design described earlier, although the package may be cylindrical rather than prismatic. With lithium ion cells, both electrodes are prepared from a slurry containing the active materials, a conductivity enhancer, and a binder. The slurry is coated onto a thin metal foil current collector and then dried and compressed. Tabs are welded to the foil current collector to permit a subsequent weld to either the case or the feedthrough pin. Again, a microporous separator material is placed between the electrodes before they are wound into a coil. Surface areas of these cells can be in excess of 100 cm².

A limitation of the coiled battery design in Fig. 4 is that the cell designs are typically limited to rectangular prisms that may not package efficiently in a device. There are variations of the coiled design that can get around this limitation, however. These include the use of multiple cathode and anode plates, or individual cathode plates with a continuous anode.^[34]

BATTERY CHEMISTRIES

Five primary battery chemistries and one rechargeable chemistry comprise the majority of power sources for implanted devices. A common feature of the primary batteries is a lithium anode, for reasons mentioned previously. Lithium batteries are made using a variety of cathode materials, depending on battery application. The following discussion describes the chemistry of lithium primary batteries as well as the lithium ion rechargeable battery.

Lithium/Iodine

The Li/I₂ battery is among the highest energy density batteries (about 1.0 Wh/cm³). The net cell reaction is given in Eq. 1c and a schematic view is shown in Fig. 5. The cathode material is prepared by thermally reacting iodine with a small amount of poly-2-vinylpyridine (P2VP). This reaction produces a two-phase material composed of crystalline iodine and a conductive, tarlike liquid. Molten cathode is filled through the fill-port and allowed to solidify in situ. The battery case serves as the cathode current collector. A unique aspect of the lithium/iodine battery is that no separator

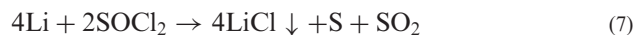
material is placed between the anode and the cathode. Upon filling the cell with cathode, the lithium and iodine react to form a layer of lithium iodide at the interface. LiI is an ionic conductor that serves as both an electrolyte and a barrier to continued reaction between anode and cathode. Therefore, no other separator is needed. Should the LiI barrier be damaged for any reason, it would simply reform. The simple battery design and the self-healing nature of the LiI electrolyte impart extreme robustness and reliability to this battery.

As the Li/I₂ battery is discharged, the LiI layer increases in thickness and becomes a major contributor to the internal resistance of the battery. Coating the lithium anode with a thin film of P2VP dramatically reduces the resistance of the lithium iodide layer. Nonetheless, the resistance of the electrolyte is large and increases with discharge. Depletion of the iodine in the cathode also causes large resistance increases. The battery resistance can reach several thousand ohms, limiting the battery to extremely low-power applications. Pacemakers, which require only a few tens of microwatts for typical operation, are the main application for lithium/iodine batteries.

The Li/I₂ battery has an open-circuit voltage of 2.8 V throughout discharge. Virtually all loss of voltage during discharge results from the continuously increasing internal resistance. Discharge curves for the battery are shown in Fig. 2A. The relatively large loss of voltage and delivered capacity, for even very small increases in current, is characteristic of a battery with a high internal resistance. The increased slope in discharge voltage near the end of its service life is a useful feature, allowing for battery voltage to be a reliable indicator of impending depletion.

Lithium/Thionyl Chloride

The lithium/thionyl chloride (Li/SOCl₂) battery has a porous carbon cathode and a SOCl₂:LiAlCl₄ active material/electrolyte. It is constructed similar to the battery in Fig. 3. An unusual feature is that the thionyl chloride is both the electrolyte solvent and the active cathode material. The cell reaction is



The sulfur and sulfur dioxide that are formed are initially soluble in the thionyl chloride. As thionyl chloride is consumed, there is a slight buildup of pressure from the sulfur dioxide and some precipitation of sulfur in the porous carbon cathode. The lithium chloride is not soluble and precipitates in the cathode throughout discharge.

The lithium anode is protected from excessive direct reaction with the thionyl chloride by a thin lithium chloride layer that forms at the interface, analogous to the lithium/iodine battery. This results in excellent shelf-life. However, when the battery is discharged after a quiescent period, the resistive lithium chloride layer produces a brief voltage drop (known as voltage delay) until the layer is removed by discharge of the lithium. During continuous discharge, the prolonged reaction of thionyl chloride with the freshly exposed lithium can lead to significant self-discharge. Fortunately, the total capacity lost to self-discharge plateaus as the discharge rate decreases.

Implantable lithium/thionyl chloride batteries have power capabilities in the tens of mW range. The open-circuit voltage is 3.65 V throughout discharge and the typical operating voltage is in the 3.5 to 3.6 V range. The energy density can reach 1.0 Wh/cm³, but self-discharge reduces the delivered energy density by about 15% at typical application discharge rates. Figure 2B shows typical discharge curves for this cell, illustrating that the delivered capacity is decreased at lower current. Another trait of the thionyl chloride system, a rapid loss of voltage at the end of discharge, gives little warning of battery depletion. This is one reason that this chemistry has generally been limited to other than life support applications.

Lithium/Silver Vanadium Oxide

The cathode of this battery is a porous material composed of silver vanadium oxide (SVO), carbon, and a binder. The electrolyte is typically 1M LiAsF₆ in a mixture of equal parts (by volume) of propylene carbonate and dimethoxyethane (i.e., PC/DME). Details of the cathode reduction products are not definitively established, but the reaction stoichiometry of the cell is



The battery is constructed for either medium-power applications (e.g., neurological stimulators) using the design in Fig. 3, or for high-power defibrillator applications using a design similar to Fig. 4.^[35] The power capability ranges from tens of mW for the medium-power design to about 10 watts for the high-power design. Energy densities are about 0.8 Wh/cm³ for the medium-power design and about 0.6 Wh/cm³ for the high-power design. The open-circuit voltage of the battery has two plateaus. The first occurs at 3.25 V and the second at 2.5 V, as shown in Fig. 2C. The battery performance is extremely stable with the exception of a small, time-dependent resistance increase that occurs in the second half of discharge.^[19]

Lithium/Carbon Monofluoride

The cathode for this battery is a porous material comprised of carbon monofluoride (CF_x where x ≈ 1), carbon, and a binder. The most common electrolyte is LiBF₄ in δ-butyrolactone (GBL) or PC/DME. Carbon monofluoride is among the highest energy density cathode materials (on a gravimetric basis). The cathode material has only moderate rate capability, so the implantable designs are limited to those similar to Fig. 3. The simplified net cell reaction is



The LiF precipitates in the cathode as the cell is discharged.

The power capability of implantable Li/CF_x batteries is on the order of tens of mW. Battery performance is very stable over the life of implanted device and energy densities in excess of 0.9 Wh/cm³ can be achieved. The theoretical open-circuit voltage of the cell is about 4.5 V, although this value is never attained in practice. Typical load voltages are in the range of 2.9 to 3.0 V throughout most of discharge. Discharge curves are shown in Fig. 2D.

Lithium/CF_x/SVO Hybrid Cathode

These cells are similar to the Li/CF_x batteries described earlier, except that the active cathode material is a mixture of CF_x and SVO. The advantage of this hybrid over CF_x is improved power capability, improved energy density (over 1.0 Wh/cm³), and a stepped voltage near end-of-discharge that is useful in signaling a necessary replacement. The high energy density and good power capability make this chemistry suitable for many applications, including possibly replacing the Li/I₂ cell for pacemakers in the future. Discharge curves are shown in Fig. 2E.

Lithium-Ion Rechargeable

Lithium-ion cells use lithium intercalation compounds as the active materials. As the battery is charged and discharged, lithium ions are exchanged between the positive and negative electrodes. A coiled prismatic cell design similar to Fig. 4 is typical, but coiled cylindrical cells are also used. The active negative material is some form of carbon, frequently graphite. A variety of positive materials are used, but the most common is lithium cobalt oxide (LiCoO₂). The cell reaction for a LiCoO₂ battery is



The open-circuit voltage at full charge is about 4.1 to 4.2 V and decreases throughout discharge to about 3 V. The average cell voltage over the discharge cycle is about 3.7 V. Recharge times can be less than two hours, but actual times may be limited by other factors, such as the maximum rate that the recharge energy can be delivered. A low-capacity fade rate makes the cycle life for this chemistry very high. Typically, thousands of charge/discharge cycles can be completed with most of the original capacity remaining.^[29,30] In addition, self-discharge of a lithium ion cell is low compared to other rechargeable batteries. A typical charge/discharge curve is shown in Fig. 2F.

COMMERCIAL APPLICATIONS

Given a variety of medical grade battery systems available, the choice of power source for a given device is driven by the match between the device's set of power, longevity (energy) and size requirements, and the corresponding capabilities of the battery. These capabilities result from the combination of a battery chemistry with its mechanical design. Because a battery's energy and size are inversely related to each other, the design process includes balancing the competing needs for a small, comfortable device with adequate power and longevity. Table 1 lists approximate ranges of power and longevity for several representative types of implanted devices currently available.

The devices can be classified into three categories, according to power needs:

- Low-power (10 to 100 μ W) devices that provide sensing or low-level electrical stimulation, such as bradycardia pacing.
- Medium-power (100 μ W to 10 mW) devices that pump drugs or stimulate the spinal cord or brain.
- High-power (≥ 10 W) devices that pump blood defibrillate the heart.

Within the classification scheme above, the ICD is unique because it is both a low- and high-power device. Most of its life is spent in a low-power mode (sensing and low-level pacing). But when needed, it may deliver many watts of power to a high-voltage capacitor for several seconds in preparation for applying a defibrillation shock to a heart. Three devices in Table 1 receive their power from external sources: the LVAD, artificial heart, and cochlear implant. For all three, the power is transmitted from outside the body, sometimes using wired connections (certain LVADs), but usually using transcutaneous inductive coupling. With inductive coupling, a transmitting coil located outside the body generates an alternating magnetic field that induces a

current in a receiving coil on the implant. The receiving coil converts the magnetic field to an electric current that is conditioned for use to power a device (or recharge a battery). The first two of these devices (i.e., LVAD and artificial heart) are critical to sustaining life and therefore, employ rechargeable batteries as backup when the external power is interrupted. For cochlear implants, the reason for an external power source is different. This device employs a continuous RF telemetry link between the implant and its external signal processor, which is also used to power the device with no added inconvenience. Therefore, because of the small size and high power density requirements of cochlear implants, the battery has been excluded from the commercial designs thus far. For externally powered devices, longevity is not driven by battery capacity, and is therefore omitted from Table 1.

Data from Table 1 have been transformed into a map of power and energy (Fig. 6) by making the simplifying assumption that the devices are used at the midpoint of the stated power and longevity ranges. The three externally-powered devices are represented by broad arrows at the right-hand side of the graph. There is an absence of battery-powered devices that operate at power levels above 10 mW for sustained times on Fig. 6. This is not a result of a shortage of therapies that would potentially require such power levels. Rather, the 10 mW threshold indicates the limit of present battery technology to provide sustained high power in small packages and over practically long lifetimes.

FUTURE POWER SOURCES

To a large extent, power sources are significant factors that limit the ability to decrease size and increase longevity of implanted devices, and there are many driving forces to push power source technology forward. These include: the continual development of new therapies; development of nontherapeutic device features (such as continuous long-range telemetry of patient data); and the therapeutic and competitive marketing advantages of smaller, longer-lived devices. It is therefore likely that there will be continued demand for power sources with ever-greater power and energy density. Advancements in power source technology are occurring at a rapid pace in many nonmedical fields (such as consumer electronics devices and electric vehicles). Some of these technologies are expected to find their way into medical implants, provided that they can be shown to meet the unique requirements of implanted sources outlined earlier. The latest power source technology is more likely to gain its foothold in power-hungry devices

Table 1 Common commercially implanted electric devices and their power sources

Disease treated or diagnosed	Device	Range of typical values			Battery chemistry	Battery design
		Power (w)	Longevity			
Bradycardia, heart failure (ventricular dysynchrony), sleep apnea	Pacemaker	10–100 μ W	7–10 yr		I ₂ , SVO, SVO/CF _x hybrid	Primary single-plate
Ventricular tachycardia, atrial fibrillation, heart failure	ICD	10–100 μ W (background), 10 W (peak)	5–8 yr		SVO	Primary coiled or multiplate
Syncope, epilepsy	Cardiac and hemodynamic monitors	10–100 μ W	1–3 yr		SOCl ₂	Primary single-plate
Parkinson's, essential tremor	Deep brain stimulator (DBS)	100 μ W–1 mW	3–8 yr		SVO	Primary single-plate
Pain, spasticity, diabetes	Implantable drug pump	100 μ W–1 mW	4–10 yr		SOCl ₂	Primary coiled
Sensorineural hearing loss	Cochlear implant	100 μ W–1 mW	n/a*		n/a	n/a
Chronic pain of trunk and limbs	Spinal cord stimulator	1–10 mW	< 1 yr–8 yr		SVO, SOCl ₂ , transcutaneous induction*	Primary single-plate
Heart failure	Left Ventricular Assist Device (LVAD), artificial heart	10–100 W	n/a*		Li ion (backup only)	Secondary coiled

*At present, commercial cochlear implants have no internal battery and receive power continuously via transcutaneous inductive transmission from an external source. The LVAD and artificial heart use either induction or a transcutaneous feedthrough for primary power, but they also include a secondary battery for backup purposes.

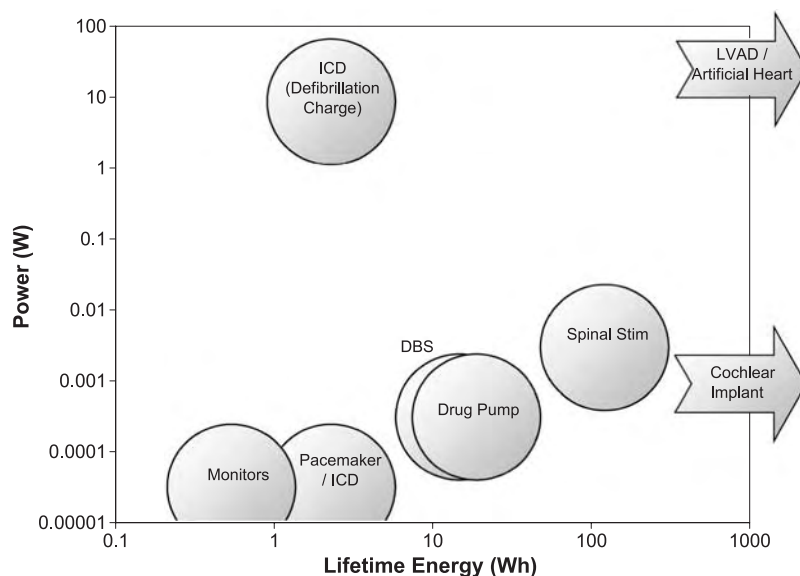


Fig. 6 Map of power versus energy for several classes of implanted devices. Externally powered devices are represented as arrows on the right-hand axis.

that are not life-sustaining. An example could be neurostimulators for pain management. For such devices, rechargeables may offer a substantial increase in device longevity, with the possibility of reduced device size. As the batteries move down in size and up in reliability, the degree of inconvenience of a rechargeable device is diminished. It would therefore not be surprising to see an increased prevalence of such cells in medical implants that will increase over time. A particularly exciting area is that of rechargeable lithium ion cells. These batteries are showing steady improvements in the energy density and reliability of^[30] since their commercialization in the early 1990s.

Even more advanced power technologies may exist over the horizon. Thin-film, solid-state batteries, in development for many years, appear to be approaching commercial use. Such cells can theoretically be made very thin (with thickness on the order of tens of μm) and have been shown to withstand extremely great numbers of cycles with little wear-out.^[36] Miniaturized fuel cells are being reported in the literature, which promise high energy density, but the problems of refilling them in an implant have yet to be addressed. Further down the road, we may see systems that harvest energy from the body itself. These may include piezoelectric systems that scavenge kinetic energy from motion of the body, or fuel cells that consume chemical fuel in circulating body fluids.^[37]

ACKNOWLEDGMENT

The authors wish to thank Dr. Esther Takeuchi for helpful discussions regarding LVAD devices.

ARTICLES OF FURTHER INTEREST

Cardiac Assist Devices; Cardiac Pacemaker; Defibrillators

REFERENCES

1. Ekstrom, S.; Johansson, L.; Lagergren, H. Handling av Adam-Stokes syndrom med intracardiell pacemaker elektrod (Treatment of Adam-Stoke's Syndrome with Intracardial Pace-Maker Electrode). *H. Opusc. Med.* **1962**, *7*, 1–3.
2. Greatbatch, W.; Chardack, W. A transistorized implantable pacemaker for the long-term correction of complete atrioventricular block. *Proc. New Engl. Res. Eng. Meet. (NEREM)* **1959**, *1* (8), 113.
3. Reuben, S.R. Alkaline Dry Cell. US Patent 2,422,045, June 10, 1947.
4. Salkind, A.; Reuben, S. Mercury Batteries for Pacemakers and Other Implanted Devices. In *Batteries for Implantable Biomedical Devices*; Owens, B.B., Ed.; Plenum Press: New York, 1986; 261–274.
5. Purdy, D.L. Nuclear Batteries for Implantable Applications. In *Batteries for Implantable Biomedical Devices*; Owens, B.B., Ed.; Plenum Press: New York, 1986; 285–350.
6. Fischell, R.E.; Schulman, J.H. A rechargeable power system for cardiac pacemakers. *Proc. 11th Intersoc. Energy Convers. Eng. Conf.* **1976**, 163.
7. Moser, J.R. Solid State Lithium-Iodine Primary Battery. US Patent 3,660,163, May 2, 1972.
8. Schneider, A.A.; Moser, J.R. Primary Cells and Iodine Containing Cathodes. US Patent 3,674,562, July 4, 1972.
9. Mead, R.T. Solid State Battery. US Patent 3,773,557, November 20, 1973.

10. Mead, R.T.; Greatbatch, W.; Rudolph, F.W. Lithium-Iodine Battery Having Coated Anode. US Patent 3,957,533, May 18, 1976.
11. Gabano, J.-P.; Broussley, M.; Grimm, M. Lithium Solid Cathode Batteries for Biomedical Implantable Applications. In *Batteries for Implantable Biomedical Devices*; Owens, B.B., Ed.; Plenum Press: New York, 1986; 181–212.
12. Skarstad, P.M. Lithium-Liquid Oxidant Batteries. In *Batteries for Implantable Biomedical Devices*; Owens, B.B., Ed.; Plenum Press: New York, 1986; 215–225.
13. Holmes, C.F. The bournier lecture: Electrochemical power sources—An important contributor to modern health care. *J. Power Sources* **1997**, *65*, xv–xx.
14. Mirowski, M.; Mower, M.M.; Reid, P.R. The automatic implantable defibrillator. *Am. Heart J.* **1980**, *100*, 1089–1092.
15. Horning, R.J.; Viswanathan, S. High Rate Lithium Cell for Medical Application. In *Proc. 29th Power Sources Conf., Atlantic City, N.J.*; The Electrochemical Society: Pennington, NJ, 1980; 64–66.
16. Liang, C.C.; Bolster, M.E.; Murphy, R.M. Metal Oxide Composite Cathode Material for High Energy Density Batteries. US Patent 4,310,609, January 12, 1982.
17. Liang, C.C.; Bolster, M.E.; Murphy, R.M. Metal Oxide Composite Cathode Material for High Energy Density Batteries. US Patent 4,391,729, July 5, 1983.
18. Schmidt, C.L.; Skarstad, P.M. Development of a Physically-Based Model for the Lithium-Iodine Battery. In *Power Sources 13*; Keily, T., Ed.; International Power Sources Symposium Committee: Leatherhead, UK, 1991; 347–362.
19. Crespi, A.M.; Schmidt, C.L.; Norton, J.; Chen, K.; Skarstad, P.M. Modeling and characterization of the resistance of lithium/SVO batteries for implantable cardioverter defibrillators. *J. Electrochem. Soc.* **2001**, *148*, A30–A37.
20. Schmidt, C.L.; Skarstad, P.M. The future of lithium and lithium-ion batteries in implantable medical devices. *J. Power Sources* **2001**, *97–98*, 742–746.
21. *Modern Batteries*, 2nd Ed.; Vincent, C.A., Scrosati, B., Eds.; Arnold: London, 1997.
22. *Handbook of Batteries*, 3rd Ed.; Linden, D., Reddy, T.B., Eds.; McGraw-Hill: New York, 2002.
23. Untereker, D.F.; Shepard, R.B.; Schmidt, C.L.; Crespi, A.M.; Skarstad, P.M. Power Systems for Implantable Pacemakers, Cardioverters and Defibrillators. In *Clinical Cardiac Pacing and Defibrillation*, 2nd Ed.; Ellenbogen, K.A., Kay, G.N., Wilkoff, B.L., Eds.; W.B. Saunders Co.: Philadelphia, 2000; 167–193.
24. *Batteries for Implantable Biomedical Devices*; Owens, B.B., Ed.; Plenum Press: New York, 1986.
25. Greatbatch, W.; Holmes, C.F. The lithium-iodine battery: A historical perspective. *PACE* **1992**, *15*, 2034–2036.
26. Bard, A.J.; Faulkner, L.R. *Electrochemical Methods*; John Wiley and Sons: New York, 1980; 103.
27. Schmidt, C.L.; Skarstad, P.M. Development of an equivalent-circuit model for the lithium/iodine battery. *J. Power Sources* **1997**, *65*, 121–128.
28. Schmidt, C.L.; Skarstad, P.M. Development of a Physically Based Model for the Lithium-Iodine Battery. In *Power Sources 13*; Keily, T., Baxter, B.W., Eds.; International Power Sources Symposium Committee: Leatherhead, England, 1991; 347–361.
29. MacLean, G.K.; Aiken, P.A.; Adams, W.A.; Mussivand, T. Preliminary evaluation of rechargeable lithium-ion cells for an implantable battery pack. *J. Power Sources* **1995**, *56*, 69–74.
30. Okamoto, E.; Watanabe, K.; Hashiba, K.; Inoue, T.; Iwazawa, E.; Momoi, M.; Hashimoto, T.; Mitamura, Y. Optimum selection of an implantable secondary battery for an artificial heart by examination of the cycle life test. *ASAIO J.* **2002**, *48*, 495–502.
31. Sunderland, W.C.; Rorvick, A.W.; Merritt, D.R.; Schmidt, C.L.; Haas, D.P. Electrochemical Cell. US Patent 5,716,729, February 10, 1998.
32. Elliot, W.B.; Paulot, W.B.; Smesko, S.A. Electrode Edge Design. May 18, 1999.
33. Howard, W.G.; Kelm, R.W.; Weiss, D.J.; Crespi, A.M.; Berkowitz, F.J.; Skarstad, P.M. High Reliability Electrochemical Cell and Electrode Assembly Therefore. US Patent 5,439,760, August 8, 1995.
34. Keister, P.P.; Mead, R.T.; Muffoletto, B.C.; Takeuchi, E.S.; Ebel, S.J.; Zelinsky, M.A.; Greenwood, J.M. Non-aqueous Lithium Battery. US Patent 4,830,940, May 16, 1989.
35. Takeuchi, E.S.; Quattrini, P.J.; Greatbatch, W. Lithium/silver vanadium oxide batteries for implantable defibrillators. *Pacing Clin Electrophysiol.* **1988**, *11* (11), 2035–2039.
36. Bates, J.B.; Dudney, N.J.; Neudecker, B.; Ueda, A.; Evans, C.D. Thin-film lithium and lithium-ion batteries. *Solid State Ionics* **2000**, *135*, 33–45.
37. Chen, T.; Barton, S.C.; Binyamin, G.; Gao, Z.; Zhang, Y.; Kim, H.H.; Heller, A. A miniature biofuel cell. *J. Am. Chem. Soc.* **2001**, *123*, 8630–8631.

Infection of Medical Devices

Lori L. Burrows

Hospital for Sick Children Research Institute, Toronto, Ontario, Canada

Antoine E. Khoury

Hospital for Sick Children, Toronto, Ontario, Canada

INTRODUCTION

The past century has seen tremendous advances in medicine that have translated into longer, healthier lives for patients. However, changes in delivery of healthcare, mainly from in-hospital care to increased outpatient or transitional care, mean that hospitalized patients are generally sicker than in the past and require more interventions. Aided by excellent innovations in the biomaterial and bioengineering fields, modern medicine makes frequent use of medical devices, which range from temporary venous catheters to long-term, completely implanted artificial joints. It is likely that most individuals in developed countries will use biomaterials and medical devices at some point in their lives. However, despite the advantages offered by the use of biomedical devices, infection remains a serious limitation to their use and a source of their malfunction.

OVERVIEW

Infection of Medical Devices... Magnitude of the Problem

While medical devices, particularly in the hospital (nosocomial) setting, are common and necessary forms of medical intervention, they can also be a major contributor to patient morbidity. For example, urinary tract infections (UTIs) compose approximately 40% of nosocomial infections in the U.S., and, of these, 80% are catheter-associated urinary tract infections (CAUTIs). Approximately 25% of hospitalized patients have urinary catheters placed with an incidence of CAUTI of 5% per day and with related bacteremia in 4% of patients.^[1] Medical devices are heavily used in the seriously ill patients that populate intensive care units (ICUs). A recent study of nosocomial infections in medical-surgical ICUs in the United States showed that the majority of reported infections were pneumonias (31%) (83% of which were associated with mechanical ventilation), UTIs (23%) (97% of which were CAUTIs), and primary blood-stream infections

(BSIs) in 14% of patients, related in 87% of cases to the presence of central lines.^[2] Table 1 lists the most commonly used devices in the United States with their estimated rates of infection and attributable mortality.^[3]

Costs Associated with Medical Device Infections

There are considerable socioeconomic consequences of device-related nosocomial infections, which are commonly manifested in increased hospital stays with associated costs. Nosocomial BSIs and pneumonia are associated with up to 35% and 30% attributable mortality, respectively.^[4] BSIs are directly responsible for 62,500 deaths per year in the United States, a number that appears to be increasing over time.^[5] A CAUTI is estimated to increase hospital stays by 1–4 days and costs nearly \$600 per day. A bloodstream infection is estimated to increase hospital stays by 7–21 days with increased costs of up to \$40,000. Pneumonia is expected to increase hospital stays by 7–30 days with approximately \$5,000 in added costs. These numbers do not consider the further, potentially dramatic, increases in cost that could occur if the organism(s) causing these device-related infections are resistant to antibiotics. The latter scenario is becoming increasingly likely as the incidence of resistance increases—especially since the presence of invasive devices is a well-documented risk factor for infection with antimicrobial-resistant microbes.^[6]

Both mortality and morbidity caused by device infections are important to consider. Infections of some devices, such as mammary or penile implants, are rarely life-threatening but cause significant morbidity through disfigurement and associated psychological trauma.^[3] Infections of both urinary catheters and joint prostheses are associated with low attributable mortality, but the latter have an increased substantial medical and economic impact. Replacement of an infected prosthesis is itself a risky proposition since the

Table 1 The magnitude of the problem of device-associated infections

Device	Estimated no. inserted in the United States per year	Rate of infection, %	Attributable mortality ^a
Bladder catheters ^b	>30,000,000	10–30	Low
Central venous catheters ^{b,c}	5,000,000	3–8	Moderate
Fracture fixation devices ^b	2,000,000	5–10	Low
Dental implants ^d	1,000,000	5–10	Low
Joint prostheses ^b	600,000	1–3	Low
Vascular grafts ^b	450,000	1–5	Moderate
Cardiac pacemakers ^{b,d}	300,000	1–7	Moderate
Mammary implants, in pairs ^c	130,000	1–2	Low
Mechanical heart valves ^d	85,000	1–3	High
Penile implants ^{b,d}	15,000	1–3	Low
Heart assist devices ^d	700	25–50	High

^aSemiquantitative scale for attributable mortality: low, <5%; moderate, 5%–25%; high, >25%.

^bNumbers estimated by analysis of market reports.

^cNumbers estimated by review of the medical literature.

^dNumbers estimated by personal communication with personnel from device manufacturing companies.

^eNumbers estimated by review of data provided by medical associations.

(From Ref. [3].)

failure rate of reimplanted devices is higher than that seen for first-time implants.

FACTORS INFLUENCING INFECTION OF MEDICAL DEVICES

Bacterial Factors

Biofilms

What leads to infection related to medical devices? The innocuous, biocompatible materials from which modern devices are constructed provide an inert surface for microorganisms to colonize. Once microorganisms gain access to the surface or lumen of a device, they attach and, through cell division, expand into multicellular, polysaccharide-enclosed communities called biofilms. Biofilms are the preferred mode of growth for bacteria in natural systems, offer the microorganisms protection from predation and noxious substances, and likely increase access to nutrients through cooperative behavior and metabolism. Though biofilm formation has long been recognized as a problem in industrial systems, only more recently has the medical significance of this process been appreciated.^[7]

Etiological agents of medical device infection

Both bacteria and fungi are capable of biofilm formation and can colonize a wide variety of surfaces.

Depending on the source of the colonizing microflora and the length of exposure of the device to the relevant environment, biofilms can be composed of a single species, which is common in medical device infections, or may be a mixture of organisms. The sources of microorganisms include those naturally present on the skin and mucosal surfaces of the patient as well as those introduced from the surroundings. Environmental flora includes organisms present on healthcare workers, in tap water, or on surfaces in the patient's vicinity. The majority of medical device infections are due to pathogens of low virulence to healthy hosts, including staphylococci and enterococci, viridans streptococci, and Gram negatives such as *Escherichia coli*, *Klebsiella spp.*, *Pseudomonas aeruginosa*, and *Proteus mirabilis*.^[1] Strain-specific differences in biofilm formation are typically found within any species, and those strains that elaborate factors that enhance biofilm formation are more commonly isolated from device infections than strains lacking these virulence features. For example, staphylococci expressing the *ica* locus, which is associated with exopolysaccharide production, are more frequently recovered than *ica*-negative strains.^[8] Of note, fungi such as *Candida*, which are typically isolated from immunocompromised patients treated with antibacterials, are also capable of forming biofilms on medical devices.^[9] These fungi are often isolated from line infections in patients who have undergone antibacterial therapy—especially those being treated with peritoneal dialysis or with long-term venous access catheters.

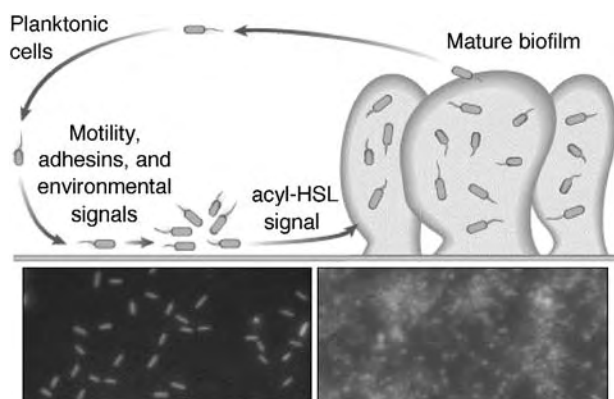


Fig. 1 Formation of a biofilm. (From Ref. [10].)

The process of biofilm formation

Biofilm formation occurs via a series of predictable events, as shown in Fig. 1.^[10] Microorganisms encounter a surface and interact with it via self-produced adhesins, which can be either proteinaceous, polysaccharide, or both. The magnitude and stability of initial attachment can be influenced by the physicochemical properties of both the bacterial and biomaterial surfaces. However, modelling of the initial interaction by representing the organisms as simple spheres subject to Van der Waals, ionic, and hydrophobic interactions is generally an oversimplification due to the variety of adhesins that can be elaborated and to their ability to actively interact with surfaces at a distance from the cell body.^[11] Attempting to predict the ability of specific microorganisms to interact with various surface types is difficult since the conditions (e.g., temperature, pH, flow) under which adherence occurs can influence the results, and anti-biofilm interventions that are successful *in vitro* may not work under clinical conditions.^[3]

In vitro studies of biofilm development have shown that, after attachment, the microorganisms begin to divide and form a monolayer of attached cells that develops into mounded microcolonies of cells and eventually into mature, polysaccharide-encased mushrooms or towers of cells separated by water channels. The open structure of the biofilm allows nutrients to penetrate and waste to exit. The structure of a mature biofilm is influenced by its composition and environment, but it is strikingly reproducible under a given set of conditions. The process of biofilm formation has been likened to a form of microbial development in which single cells progress through reproducible states into a differentiated, three-dimensional community. The ability of the cells to communicate with one another via small, diffusible signaling molecules helps to orchestrate biofilm development. Disruption of

signaling results in defects in biofilm structure and physiology.^[12]

Interestingly, a mature bacterial biofilm behaves as a viscoelastic or rubbery material and can be deformed under flow. Biofilms formed under high-shear stress are actually physically stronger and more firmly attached to the substratum than those formed under low shear.^[13] Bacteria can be disseminated from a biofilm either as individual free-swimming cells, broken fragments, or clumps of biofilm. Physicians are familiar with the consequences of the disintegration of biofilms (vegetations) on artificial heart valves. The released fragments lodge in small capillary beds and lead to embolisms (an embolic shower).^[14]

Treatment challenges associated with biofilms

The most striking and medically relevant phenotype of microorganisms growing in biofilms is their markedly increased tolerance of antimicrobial agents and immune system components. Numerous studies have demonstrated that biofilms can resist up to thousands of times the minimal inhibitory concentration (MIC) of antimicrobials that can kill planktonic (free-swimming or unattached) cells of the same strain.^[7] Similarly, biofilms are resistant to complement, antibodies, and professional phagocytes due to their physical bulk and the production of antiphagocytic exopolysaccharides. This tolerant phenotype is not a result of mutation, since cells that are mechanically dispersed from a biofilm can be killed by standard antibiotic concentrations and are susceptible to killing by immune cells. For this reason, phenotypic or biofilm-growth related resistance is more correctly referred to as tolerance to distinguish it from genetically-encoded antimicrobial resistance.^[15]

A number of hypotheses have been advanced to explain the ability of biofilms to withstand high levels of antibiotics.^[15] They include:

1. Slow growth with concomitant reduction in susceptibility to antibiotics that affect physiological pathways.
2. Reduced exposure due to the diffusion barrier created by the exopolysaccharide matrix surrounding the cells.
3. Escape from the normal process of programmed cell death by a subset of biofilm cells.
4. Biofilm-specific alterations in gene expression and physiology resulting in a more resistant phenotype.

In reality, it is likely that a combination of these factors results in the observed tolerance of biofilms to typical concentrations of antibiotics. Older biofilms that have been allowed to mature and fully develop their

three-dimensional architecture demonstrate markedly increased tolerance compared with younger biofilms.

Because of their resistant phenotype, biofilms are difficult to treat *in situ*. Although antibiotic therapy can often reduce the amount of biofilm present and kill bacteria that have dispersed from the nidus of infection, withdrawal of therapy inevitably leads to recrudescence. For this reason, prevention of initial biofilm formation on a device is more practical and preferable to attempting treatment of the fully developed and drug-tolerant form.

Device-Related Factors

Site of device use and associated risk of infection

There are several factors that influence the susceptibility of a medical device to colonization with microorganisms. Exposure of the device to the external milieu is a risk factor. Devices with an exit site, such as venous or urinary (Foley) catheters, are at higher risk of colonization compared with completely implanted devices. Dental devices are particularly subject to continuous exposure to a variety of microorganisms in a warm, moist environment that promotes colonization. Semi-permanent devices with exit sites, such as peritoneal dialysis catheters or Hickman lines are also at risk, especially prior to successful tissue integration at the cuff. However, even completely implanted devices are vulnerable to contamination at the time of implantation. Some implanted devices, such as prosthetic heart valves, can become latently infected via hematogenous seeding.

Physical factors; design and composition

The physical design or construction of a medical device can influence its risk of colonization. Complex devices offer increased surface area for colonization as do devices with lumens. The material of which a device is constructed can influence colonization in that its physicochemical properties, such as hydrophobicity, may reduce or increase adherence of specific types of microorganisms. Duration of use is an important variable; time indwelling is positively correlated with incidence of infection. The texture of the surface material (roughness) can influence colonization; rough surfaces are more readily colonized than smooth surfaces.^[16]

Host Factors

Conditioning films

Regardless of their surface composition and characteristics, most types of materials are rapidly covered with

host molecules after introduction in a process known as conditioning film deposition.^[17] Formation of a conditioning film on a device surface provides alternative receptors for pathogenic microorganisms and attenuates the effect of antibiofilm coatings designed to reduce bacterial adhesion. The properties of a conditioning film varies and depends on the site of introduction of a device, *i.e.*, whether it is in contact with blood, urine, or saliva, all of which contain unique components. The site of device placement also influences the rate of conditioning film deposition as does its exposure to specific volumes and flow rates of host fluids. Many pathogenic bacteria produce receptors for host molecules such as fibronectin, which tend to accumulate on foreign bodies exposed to human fluids. Receptors for human molecules promote adherence of the bacteria to a coated surface and, in turn, encourage further deposition of human proteins onto the surface of the adherent bacteria. A noteworthy example of this phenomenon is the development of large bacterial biofilms or vegetations on prosthetic heart valves.

Other host factors

What about other host factors? The immune status of the individual receiving the medical device can influence the rate of colonization.^[18] For example, low numbers of skin microorganisms introduced with a new venous catheter might ordinarily be scavenged by a healthy immune system, but they may survive to become irreversibly attached to a device in an immunocompromised individual. Similarly, the presence of perioperative antibiotics in the circulation of the patient may reduce biofilm formation by killing transient microorganisms before they are able to attach to a device and develop into a biofilm. Depending on the site, the presence of the device will initially elicit an immune response against the foreign body, which may coincidentally mop up small numbers of organisms before they can form a resistant biofilm. However, long-term presence of the medical device can impair local immune responses and increase the risk that bacteria may be able to colonize.^[19] The site of device implantation can influence infection rates. Sites that are moist or difficult to clean will harbor increased numbers of microorganisms with increased risk of device contamination.

COLONIZATION VERSUS INFECTION

While colonization of a medical device by bacteria or fungi is common, if not inevitable, it does not lead inexorably to disease. In fact, the mechanisms by which biofilm formation leads to disease are not completely understood. Examination of implants such

as intrauterine devices retrieved from asymptomatic women showed they can be heavily contaminated with bacteria that presumably migrated from the plastic tail used for vaginal location of the device for removal.^[1] It is difficult to predict whether colonization of a device will lead to overt disease, but the combination of decreased host immune status, increased duration of exposure to a colonized implant, and level of pathogenicity of the colonizing strain(s) will affect the progression of infection.

Some disease manifestations are more easily understood, including the development of urinary tract and bloodstream infections from cells or cell clusters that become dislodged from indwelling catheters. Other diseases related to biofilms may present a more subclinical and chronic picture, such as the development of extensive scarring or capsule formation surrounding a device that is colonized with a bacterial biofilm of low pathogenicity. Such chronic irritation of the immune system has been implicated in the formation of thickened capsules of scar tissue around silicone breast implants.^[20,21]

An area of growing concern is the increasing prevalence of multidrug-resistant bacteria associated with medical device infections. Since the biofilm mode of growth already affords considerable protection against antimicrobials, the development of biofilms by drug-resistant organisms has the potential to lead to completely intractable infections. In fact, the formation of biofilms can encourage the development of drug resistance since the bacteria within biofilms that are unsuccessfully treated with antibiotics may be exposed to suboptimal drug levels that present opportunities for resistant mutants to be selected. The exchange of mobile genetic resistance factors such as plasmids, transposons, and bacteriophages has been shown to be enhanced in the biofilm environment, which is probably due to the physical proximity of the bacteria that facilitates transfer.^[22]

DETECTION OF MEDICAL DEVICE-RELATED INFECTIONS

Diagnosis of a medical device-related infection is not always straightforward. As mentioned above, bacterial colonization of a device can occur without evidence of infection. This scenario is common in the case of urinary catheters, particularly in the short term. Simple colonization of a device can precede symptoms of infection by days and is probably due to the requirement for sufficient numbers of microorganisms to accumulate before the immune response produces symptoms, which may include reddening of tissue at exit sites, fever, leukocytosis, and pain.

It is relatively easy to detect microorganisms using semiquantitative staining or plate-count techniques on a

suspicious device that has been removed from the patient. However, accurate diagnosis of potential infection of in situ devices is much more difficult and key to the difficult clinical decision of whether a device should be removed. For totally implanted devices, such a decision may mean increased morbidity to the patient from surgical procedures required to effect device removal. Analysis of CVCs removed for suspected infection show that as few as 15% may actually be infected,^[23] which emphasizes the need to establish a specific diagnosis of device-associated infection. Traditional approaches to confirm a suspected infection of totally implanted devices include assays of erythrocyte sedimentation rate, C-reactive protein level, site aspiration, advanced radiologic imaging, and nuclear medicine scans.

One of the techniques recently developed to determine whether CVCs are an infectious nidus without removal of the device is the use of differential blood samples taken both from the catheter and a peripheral site. If the sample from the catheter contains 5 to 10 times more organisms on quantitative plate counts or becomes culture-positive more than two hours before the peripheral sample, the catheter is deemed to be the source of the infection.^[24] This technique is limited because it requires that the organism causing the infection be culturable under standard laboratory conditions and that it grows within a reasonable timeframe for differentiation to be made. However, if these criteria are met, differential time-to-positivity has a high sensitivity and specificity for diagnosis of the common agents causing catheter-related sepsis.

One novel approach to detection of implant infections in situ is the recent development of technetium 99m-labelled antimicrobial agents, including ciprofloxacin (commercialized under the brand name Infec-ton), fluconazole, or antimicrobial peptides.^[25-29] These substances are injected prior to standard nuclear imaging of a patient to differentiate suspected bacterial or fungal infections from sterile inflammatory processes, and function by accumulating in the target organisms, if present. This approach has proven to have reasonable sensitivity and specificity in identifying device-related infections and has also been used to monitor the success of long-term antimicrobial therapy in patients with confirmed implant infections. This type of non-invasive imaging modality is likely to play a much larger role in future diagnoses of implant-related infections.

APPROACHES TO PREVENTING INFECTION OF MEDICAL DEVICES

Infection Control and Design Considerations

The most important antibiofilm approach in medicine is prevention. Stringent infection control is important

in preventing colonization of the device at placement and during its contact with the patient. Also, thoughtful design of medical devices to minimize the potential contact with bacteria when possible helps to reduce the incidence of infection. The usefulness of such measures was apparent early on with the redesign of urinary catheters from open drainage systems that were highly susceptible to luminal colonization to completely closed drainage systems with aseptic access ports and antimicrobials in the drainage bag to prevent retrograde access to the lumen. While closed drainage systems have significantly reduced infections associated with catheter lumens, they are not as useful at preventing ascension of bacteria to the bladder along the external surface of the catheter.

Central venous catheters are used for a variety of interventions including the administration of blood products, medications, nutrient solutions, or supportive fluids as well as permitting the sampling of blood from the patient. They are also the most common source of blood-stream infections (BSI). Skin flora, particularly coagulase-negative staphylococci, are the primary pathogens. They colonize the external surface of the catheter from the skin or enter the lumen via the catheter hub. Use of stringent asepsis during catheter placement by trained personnel is key to reduction of infection rates. Redesign of CVCs in recent years to include ports that can be flushed or locked with antiseptic or antibiotic solutions has had a significant impact in reducing BSI.

Interventions Targeting Colonization Reduction

Non-biomaterial based, and largely unsuccessful, strategies to reduce colonization of urinary catheters have included: disinfection of the meatus prior to catheter insertion or use of antimicrobial lubricants; irrigation of the catheter and bladder with antimicrobial agents; or using systemic antibiotic prophylaxis of catheterized patients.^[30] Systemic ciprofloxacin therapy was successful in preventing catheter blockage in one study^[31] and was possibly due to binding of the drug to the device surface. Ciprofloxacin is secreted in urine in an active form and has been shown to prevent adhesion of bacteria to catheters that have been pretreated by soaking in the drug.^[32] A recent study demonstrated that oral therapy with fluoroquinolones led to deposition of the drug on ureteral stents at levels above the minimal inhibitory concentrations for common uropathogens such as *Escherichia coli*, *Pseudomonas aeruginosa*, *Enterococcus faecalis*, and *Staphylococcus aureus*.^[33] The accumulation of systemically or orally administered antibiotics onto urinary catheters or stents may therefore contribute to the

perceived usefulness of this therapy at preventing CAUTI.

In the case of CVCs, application of topical antimicrobial ointments at the exit site to prevent line infections has met with limited success. Povidone-iodine,^[34] mupirocin,^[35] and polysporin^[36] all help to reduce infections in hemodialysis patients, particularly in long-term situations. Compared to a urinary catheter, the exit site of a CVC is in a more exposed and accessible location. Also, CVCs have cuffs that allow tissue integration in the longer term. These considerations make peripheral antimicrobial therapy more practical and successful.

Surface Modification

There have been many creative attempts to stop biofilms from forming or remove biofilms from surfaces. Unfortunately, no single method has been successful. Modification of the physicochemical properties of surfaces to reduce bacterial adhesion has been largely ineffective since microorganisms display a range of cell surface properties and the presence of conditioning films quickly negates any short-term gains in colonization prevention. Because of their resistance to biocides, mechanical removal and treatment with harsh chemicals are the industrial standard for remediation of biofilms. Since these tactics are generally not practical in medical systems, other measures are necessary.

Alteration of the base materials of which urinary catheters are made does not appear to significantly affect bacterial colonization, but it has been shown to influence the rate of catheter encrustation. Specifically, silicone leads to less encrustation than latex, siliconized latex, or Teflon.^[37] Since the risk of catheter infection increases with indwelling time, patients that are able are encouraged to use intermittent self-catheterization (IC) with a clean catheter. The practice of IC has led to modifications in catheter surface coatings to increase their hydrophilicity and lubricity for ease of insertion.^[38,39] While available data suggests that use of hydrophilic catheters for IC may reduce bacteriuria and urethral stricture formation, there are no prospective randomized multicenter trials that address these issues. One of the most popular and prevalent methods for biofilm reduction has been coating or impregnation of surfaces with antimicrobial agents. This approach is based on the acknowledgement that bacterial adhesion cannot currently be completely prevented. Instead, as a pre-emptive strike, the surface of a device is modified to elute high local concentrations of antimicrobial that kill any adherent organisms before they have a chance to develop into a resistant biofilm. This tactic has been widely applied to

the most frequently used devices, such as urinary (Foley) catheters and central venous catheters (CVCs).

Urinary Catheters

With a focus on prevention of biofilm formation on urinary catheters, a variety of surface modifications have been tried over the past two decades. The most successful experimental approaches have been coating of the inner and outer surfaces or impregnation of the catheter material with single or combinations of antimicrobials or with various forms of silver. Antimicrobial catheters that were efficacious in prevention of bacterial colonization *in vitro*, and have been put into widespread commercial distribution, include nitrofurazone-impregnated catheters,^[40] minocycline/rifampin-impregnated catheters,^[41] and silver-coated catheters.^[42] Other surface modifications that have been tested include antiseptics such as chlorhexidine,^[43] tobramycin,^[44] and ciprofloxacin.^[45]

Unfortunately, there are few large, randomized controlled trials that examine the usefulness of antimicrobial-treated urinary catheters. A recent review of completed trials on silver catheters found that majority of studies done to date were not designed with sufficient quality standards to permit the formulation of recommendations on the use of silver catheters.^[46] Another potentially contentious issue is the markedly increased cost of modified catheters compared with standard catheters. This is clearly an area that requires further research efforts to demonstrate the cost-effectiveness of treated catheters and justify their widespread use.

CVCs

CVCs coated or impregnated with silver, silver/chlorhexidine, silver-carbon-platinum, or minocycline/rifampin have been tested clinically. Similar to urinary catheters, as described above, there are as yet few large-scale, well-designed clinical trials on which usage recommendations can be based. To date, the minocycline/rifampin combination appears to demonstrate good efficacy and potential cost savings when patients have the catheter in place for at least two weeks.^[47] Issues with these catheters include duration of antimicrobial release, particularly important in long-term devices, and the mechanism of action, whether bacteriocidal or bacteriostatic.

ADJUNCT APPROACHES TO TREATMENT OF MEDICAL DEVICE INFECTIONS

In recent years, there have been attempts to use non-chemical modes of treatment or adjunct treatments for

remediation of biofilms on medical devices. While these approaches can be shown to be efficacious in a laboratory setting, it is not yet practical for them to enter widespread clinical use in their current form. These techniques include exposure of biofilms to ultrasonic or electrical energy.

Several studies have shown that ultrasound (US) can potentiate the antibiofilm activity of antimicrobials, both *in vitro* and when tested in an animal model.^[48] However, the energy levels required for biofilm disruption may be high enough to result in collateral tissue or device damage, and targeted delivery of US to the affected device can be difficult to achieve. Interestingly, the ability of sonic energy to disrupt biofilms on surfaces has prompted the development of novel medical devices, specifically dental instruments and toothbrushes that employ sonication or ultrasonication in conjunction with mechanical shear forces to remove dental plaque from teeth.

Work done in the early 1990s showed that biofilms placed into low-voltage electrical fields demonstrated increased susceptibility to antimicrobials compared with control biofilms, which is termed the bioelectric effect.^[49] The iontophoretic effect of low-voltage fields can help antibiotics penetrate biofilm matrices. Also, it has been hypothesized that the electrical field causes increased generation of oxygen within the biofilm with concomitant increases in bacterial growth and associated enhancement of their susceptibility to antimicrobial agents.^[50] The electrical field can generate protons, hydroxyl ions, and reactive oxygen species within the biofilm structure, all of which may contribute to bacterial killing.

FUTURE APPROACHES TO PREVENTION AND THERAPY OF MEDICAL DEVICE INFECTION

Novel Polymers

Among the approaches currently under development for the treatment of biofilm infections is manipulation of polymer design to incorporate antibiotics that are released only in response to infection. For example, polyurethanes containing integral ciprofloxacin residues embedded via chemical linkages susceptible to inflammatory cell enzymes such as cholesterol esterase have been developed and demonstrated to release therapeutic levels of drug *in vitro*.^[51] These smart polymers release drug in the presence of an inflammatory process in proportion to the level of inflammation present. Once the drug kills the infecting organism, the inflammatory response and related release of antibiotic subsides. This type of construct may reduce the overall amount of antibiotic released into the environment by a loaded device.

Non-Antibiotic Approaches

Inhibitors of bacterial communication

Although device coatings containing antimicrobial agents are currently one of the more useful antibiofilm approaches, concerns over increased resistance to these agents has prompted the search for nonantibiotic options. Alternatives to antibiotics or metals as catheter coatings that may become useful in the future include compounds that interfere with bacterial cell–cell communication (quorum sensing), thus prevents or disrupts the ability of the organisms to generate a biofilm. Such compounds have been identified in the secretions of marine plants, which naturally inhibit the formation of biofilms on their surfaces through the production of halogenated furanones.^[52] Use of peptides that naturally suppress quorum sensing in staphylococci has been demonstrated to reduce virulence and biofilm formation by these bacteria.^[53] There is some speculation that modulation of quorum sensing circuitry may be useful in promoting the dispersal of biofilm cells by mimicking the natural signals produced by the organisms during that phase of biofilm development. While this is an attractive idea, encouragement of biofilm dispersal by use of quorum sensing signals or mechanistic means, such as exopolysaccharide-degrading enzymes, could have deleterious consequences in patients if large amounts of organisms are released into the circulation, potentially causing sepsis or inflammation in response to bacterial components such as lipopolysaccharide or peptidoglycan.

Essential nutrient limitation

Recent studies of the effects of lactoferrin^[54] and catecholamine inotropes^[55] on biofilm formation have demonstrated that manipulation of iron availability in the host plays an important role in reducing biofilm formation. Chelation of iron by lactoferrin reduced biofilm formation by *P. aeruginosa* by decreasing the number of cells binding irreversibly to a biomaterial surface. In contrast, the ability of *S. epidermidis* to utilize iron complexed with inotropes like norepinephrine enhanced biofilm formation on CVCs. These studies imply that development of medical device materials or adjunct treatments that scavenge iron or other essential growth factors may result in decreased biofilm formation.

Matrix disruption

Another potential target to reduce or prevent biofilm formation is the polysaccharide matrix surrounding and protecting the biofilm cells. This component

plays many roles in biofilm physiology. Reducing or preventing its synthesis may result in sufficient impairment of biofilm defenses that they may be susceptible to more conventional treatment strategies. An example of this approach is the demonstrated effect of macrolide antibiotics, which can inhibit biofilm formation at subminimal inhibitory concentration levels by reducing polysaccharide synthesis.^[56] Combination therapy using low levels of macrolides with conventional levels of bacteriocidal antimicrobials may aid in reducing biofilm infections.

Probiotics

One novel concept of interest related to the reduction or prevention of medical device infection is bacterial interference. This idea is based on the observation that commensal bacteria (normal flora) can interfere with or prevent the establishment of pathogens or transients in the body. It may be possible to use commensal or probiotic bacteria or their products to block or remediate the adhesion of pathogenic bacteria to medical devices. As an example, application of some strains of probiotics to artificial voice prostheses in a model throat system was shown to reduce the presence of the fungal pathogen *Candida* by up to 96%^[57] and mixed bacterial/fungal biofilm formation by up to 100%.^[58]

CONCLUSIONS

Medical device use will continue to grow as our population ages and as new devices are introduced into widespread use. As infection remains a primary complication of these devices, innovative approaches to prevention and treatment of biofilms are required. Also, diagnostic techniques that can accurately detect device-associated infection without removal or interference with the device will be extremely useful in preventing unnecessary interventions and guiding therapy. There are indications that antimicrobial catheters can have an impact on patient care—both in reducing morbidity and in saving healthcare dollars. However, large, well-designed clinical trials are necessary to justify their widespread usage outside of high-risk environments.

ARTICLES OF FURTHER INTEREST

Antimicrobial/Antibiotic (Infection Resistance) Materials; Biomaterials Immune Response; Foreign Body Response; Host Reactions; Surface Coatings; Surface Modification

REFERENCES

1. Donlan, R.M.; Costerton, J.W. Biofilms: Survival mechanisms of clinically relevant microorganisms. *Clin. Microbiol. Rev.* **2002**, *15* (2), 167–193.
2. Richards, M.J.; Edwards, J.R.; Culver, D.H.; Gaynes, R.P. Nosocomial infections in combined medical-surgical intensive care units in the United States. *Infect. Control Hosp. Epidemiol.* **2000**, *21* (8), 510–515.
3. Darouiche, R.O. Device-associated infections: A macroproblem that starts with microadherence. *Clin. Infect. Dis.* **2001**, *33* (9), 1567–1572.
4. Jarvis, W.R. Selected aspects of the socioeconomic impact of nosocomial infections: Morbidity, mortality, cost, and prevention. *Infect. Control Hosp. Epidemiol.* **1996**, *17* (8), 552–557.
5. Pittet, D.; Wenzel, R.P. Nosocomial bloodstream infections. Secular trends in rates, mortality, and contribution to total hospital deaths. *Arch. Intern. Med.* **1995**, *155* (11), 1177–1184.
6. Safdar, N.; Maki, D.G. The commonality of risk factors for nosocomial colonization and infection with antimicrobial-resistant *Staphylococcus aureus*, enterococcus, gram-negative bacilli, *Clostridium difficile*, and *Candida*. *Ann. Intern. Med.* **2002**, *136* (11), 834–844.
7. Costerton, J.W.; Stewart, P.S.; Greenberg, E.P. Bacterial biofilms: A common cause of persistent infections. *Science* **1999**, *284* (5418), 1318–1322.
8. Heilmann, C.; Schweitzer, O.; Gerke, C.; Vanittanakom, N.; Mack, D.; Gotz, F. Molecular basis of intercellular adhesion in the biofilm-forming *Staphylococcus epidermidis*. *Mol. Microbiol.* **1996**, *20* (5), 1083–1091.
9. Hawser, S.P.; Douglas, L.J. Biofilm formation by *Candida* species on the surface of catheter materials in vitro. *Infect. Immun.* **1994**, *62* (3), 915–921.
10. Kolter, R.; Losick, R. One for all and all for one. *Science* **1998**, *280* (5361), 226–227.
11. Stewart, P.S. New ways to stop biofilm infections. *Lancet* **2003**, *361* (9352), 97.
12. Davies, D.G.; Parsek, M.R.; Pearson, J.P.; Iglewski, B.H.; Costerton, J.W.; Greenberg, E.P. The involvement of cell-to-cell signals in the development of a bacterial biofilm. *Science* **1998**, *280* (5361), 295–298.
13. Stoodley, P.; Cargo, R.; Rupp, C.J.; Wilson, S.; Klapper, I. Biofilm material properties as related to shear-induced deformation and detachment phenomena. *J. Ind. Microbiol. Biotechnol.* **2002**, *29* (6), 361–367.
14. Thapar, M.K.; Rao, P.S.; Feldman, D.; Linde, L.M. Infective endocarditis: A review. I. Incidence, etiology, pathology and clinical features. *Paediatrician* **1978**, *7* (1–3), 65–84.
15. Lewis, K. Riddle of biofilm resistance. *Antimicrob. Agents Chemother.* **2001**, *45* (4), 999–1007.
16. Quirynen, M.; Bollen, C.M. The influence of surface roughness and surface-free energy on supra- and subgingival plaque formation in man. A review of the literature. *J. Clin. Periodontol.* **1995**, *22* (1), 1–14.
17. Busscher, H.J.; Bos, R.; van der Mei, H.C. Initial microbial adhesion is a determinant for the strength of biofilm adhesion. *FEMS Microbiol. Lett.* **1995**, *128* (3), 229–234.
18. Bandyk, D.F.; Kinney, E.V.; Riefsnyder, T.I.; Kelly, H.; Towne, J.B. Treatment of bacteria-biofilm graft infection by in situ replacement in normal and immune-deficient states. *J. Vasc. Surg.* **1993**, *18* (3), 398–405. Discussion 405–396.
19. Rudnicka, W.; Sadowska, B.; Ljungh, A.; Rozalska, B. Specific immune response to staphylococcal antigens during long-lasting biomaterial implantation. *FEMS Immunol. Med. Microbiol.* **1997**, *19* (1), 7–14.
20. Dobke, M.K.; Svahn, J.K.; Vastine, V.L.; Landon, B.N.; Stein, P.C.; Parsons, C.L. Characterization of microbial presence at the surface of silicone mammary implants. *Ann. Plast. Surg.* **1995**, *34* (6), 563–569. Discussion 570–561.
21. Virden, C.P.; Dobke, M.K.; Stein, P.; Parsons, C.L.; Frank, D.H. Subclinical infection of the silicone breast implant surface as a possible cause of capsular contracture. *Aesthetic. Plast. Surg.* **1992**, *16* (2), 173–179.
22. Lamont, R.J.; Bryers, J.D. Biofilm-induced gene expression and gene transfer. *Methods Enzymol.* **2001**, *336*, 84–94.
23. Padberg, F.T., Jr.; Ruggiero, J.; Blackburn, G.L.; Bistrrian, B.R. Central venous catheterization for parenteral nutrition. *Ann. Surg.* **1981**, *193* (3), 264–270.
24. Blot, F.; Schmidt, E.; Nitenberg, G.; Tancrede, C.; Leclercq, B.; Laplanche, A.; Andremont, A. Earlier positivity of central-venous- versus peripheral-blood cultures is highly predictive of catheter-related sepsis. *J. Clin. Microbiol.* **1998**, *36* (1), 105–109.
25. Welling, M.M.; Lupetti, A.; Balter, H.S.; Lanzeri, S.; Souto, B.; Rey, A.M.; Savio, E.O.; Paulusma-Annema, A.; Pauwels, E.K.; Nibbering, P.H. 99mTc-labeled antimicrobial peptides for detection of bacterial and *Candida albicans* infections. *J. Nucl. Med.* **2001**, *42* (5), 788–794.
26. Welling, M.M.; Paulusma-Annema, A.; Balter, H.S.; Pauwels, E.K.; Nibbering, P.H. Technetium-99m labelled antimicrobial peptides discriminate between bacterial infections and sterile inflammations. *Eur. J. Nucl. Med.* **2000**, *27* (3), 292–301.
27. Lupetti, A.; Welling, M.M.; Mazzi, U.; Nibbering, P.H.; Pauwels, E.K. Technetium-99m labelled fluconazole and antimicrobial peptides for imaging of *Candida albicans* and *Aspergillus fumigatus* infections. *Eur. J. Nucl. Med. Mol. Imaging* **2002**, *29* (5), 674–679.
28. Lupetti, A.; Welling, M.M.; Pauwels, E.K.; Nibbering, P.H. Radiolabelled antimicrobial peptides for infection detection. *Lancet Infect. Dis.* **2003**, *3* (4), 223–229.
29. Vinjamuri, S.; Hall, A.V.; Solanki, K.K.; Bomanji, J.; Siraj, Q.; O’Shaughnessy, E.; Das, S.S.; Britton, K.E. Comparison of 99mTc infecton imaging with radiolabelled white-cell imaging in the evaluation of bacterial infection. *Lancet* **1996**, *347* (8996), 233–235.
30. Burrows, L.L.; Khoury, A.E. Issues surrounding the prevention and management of device-related infections. *World J. Urol.* **1999**, *17* (6), 402–409.
31. Stickler, D.J.; King, J.B.; Winters, C.; Morris, S.L. Blockage of urethral catheters by bacterial biofilms. *J. Infect.* **1993**, *27* (2), 133–135.
32. Cormio, L.; La Forgia, P.; La Forgia, D.; Siitonen, A.; Ruutu, M. Bacterial adhesion to urethral catheters: Role

- of coating materials and immersion in antibiotic solution. *Eur. Urol.* **2001**, *40* (3), 354–358. Discussion 359.
33. Reid, G.; Habash, M.; Vachon, D.; Denstedt, J.; Riddell, J.; Beheshti, M. Oral fluoroquinolone therapy results in drug adsorption on ureteral stents and prevention of biofilm formation. *Int. J. Antimicrob. Agents* **2001**, *17* (4), 317–319. Discussion 319–320.
 34. Levin, A.; Mason, A.J.; Jindal, K.K.; Fong, I.W.; Goldstein, M.B. Prevention of hemodialysis subclavian vein catheter infections by topical povidone-iodine. *Kidney Int.* **1991**, *40* (5), 934–938.
 35. Sesso, R.; Barbosa, D.; Leme, I.L.; Sader, H.; Canziani, M.E.; Manfredi, S.; Draibe, S.; Pignatari, A.C. Staphylococcus aureus prophylaxis in hemodialysis patients using central venous catheter: Effect of mupirocin ointment. *J. Am. Soc. Nephrol.* **1998**, *9* (6), 1085–1092.
 36. Lok, C.E.; Stanley, K.E.; Hux, J.E.; Richardson, R.; Tobe, S.W.; Conly, J. Hemodialysis infection prevention with polysporin ointment. *J. Am. Soc. Nephrol.* **2003**, *14* (1), 169–179.
 37. Sedor, J.; Mulholland, S.G. Hospital-acquired urinary tract infections associated with the indwelling catheter. *Urol. Clin. North Am.* **1999**, *26* (4), 821–828.
 38. Hedlund, H.; Hjelm, K.; Jonsson, O.; Klarskov, P.; Talja, M. Hydrophilic versus non-coated catheters for intermittent catheterization. *Scand J. Urol. Nephrol.* **2001**, *35* (1), 49–53.
 39. Fader, M.; Moore, K.N.; Cottenden, A.M.; Pettersson, L.; Brooks, R.; Malone-Lee, J. Coated catheters for intermittent catheterization: Smooth or sticky? *BJU Int.* **2001**, *88* (4), 373–377.
 40. Johnson, J.R.; Delavari, P.; Azar, M. Activities of a nitrofurazone-containing urinary catheter and a silver hydrogel catheter against multidrug-resistant bacteria characteristic of catheter-associated urinary tract infection. *Antimicrob. Agents Chemother.* **1999**, *43* (12), 2990–2995.
 41. Darouiche, R.O.; Safar, H.; Raad, I.I. In vitro efficacy of antimicrobial-coated bladder catheters in inhibiting bacterial migration along catheter surface. *J. Infect. Dis.* **1997**, *176* (4), 1109–1112.
 42. Liedberg, H.; Lundberg, T. Silver alloy coated catheters reduce catheter-associated bacteriuria. *Br. J. Urol.* **1990**, *65* (4), 379–381.
 43. Stickler, D.; Hewett, P. Activity of antiseptics against biofilms of mixed bacterial species growing on silicone surfaces. *Eur. J. Clin. Microbiol. Infect. Dis.* **1991**, *10* (3), 157–162.
 44. Nickel, J.C.; Downey, J.; Costerton, J.W. Movement of *Pseudomonas aeruginosa* along catheter surfaces. A mechanism in pathogenesis of catheter-associated infection. *Urology* **1992**, *39* (1), 93–98.
 45. Pugach, J.L.; DiTizio, V.; Mittelman, M.W.; Bruce, A.W.; DiCosmo, F.; Houry, A.E. Antibiotic hydrogel coated Foley catheters for prevention of urinary tract infection in a rabbit model. *J. Urol.* **1999**, *162* (3 Pt 1), 883–887.
 46. Niel-Weise, B.S.; Arend, S.M.; van den Broek, P.J. Is there evidence for recommending silver-coated urinary catheters in guidelines? *J. Hosp. Infect.* **2002**, *52* (2), 81–87.
 47. Marcic, K.D.; Veenstra, D.L.; Lipsky, B.A.; Saint, S. Which antimicrobial impregnated central venous catheter should we use? Modeling the costs and outcomes of antimicrobial catheter use. *Am. J. Infect. Control* **2003**, *31* (1), 1–8.
 48. Rediske, A.M.; Roeder, B.L.; Brown, M.K.; Nelson, J.L.; Robison, R.L.; Draper, D.O.; Schaalje, G.B.; Robison, R.A.; Pitt, W.G. Ultrasonic enhancement of antibiotic action on *Escherichia coli* biofilms: An in vivo model. *Antimicrob. Agents Chemother.* **1999**, *43* (5), 1211–1214.
 49. Houry, A.E.; Lam, K.; Ellis, B.; Costerton, J.W. Prevention and control of bacterial infections associated with medical devices. *Asaio. J.* **1992**, *38* (3), M174–M178.
 50. Stewart, P.S.; Wattanakaroon, W.; Goodrum, L.; Fortun, S.M.; McLeod, B.R. Electrolytic generation of oxygen partially explains electrical enhancement of tobramycin efficacy against *Pseudomonas aeruginosa* biofilm. *Antimicrob. Agents Chemother.* **1999**, *43* (2), 292–296.
 51. Woo, G.L.; Yang, M.L.; Yin, H.Q.; Jaffer, F.; Mittelman, M.W.; Santerre, J.P. Biological characterization of a novel biodegradable antimicrobial polymer synthesized with fluoroquinolones. *J. Biomed. Mater. Res.* **2002**, *59* (1), 35–45.
 52. Hentzer, M.; Riedel, K.; Rasmussen, T.B.; Heydorn, A.; Andersen, J.B.; Parsek, M.R.; Rice, S.A.; Eberl, L.; Molin, S.; Hoiby, N.; Kjelleberg, S.; Givskov, M. Inhibition of quorum sensing in *Pseudomonas aeruginosa* biofilm bacteria by a halogenated furanone compound. *Microbiology* **2002**, *148* (Pt 1), 87–102.
 53. Balaban, N.; Goldkorn, T.; Nhan, R.T.; Dang, L.B.; Scott, S.; Ridgley, R.M.; Rasooly, A.; Wright, S.C.; Larrick, J.W.; Rasooly, R.; Carlson, J.R. Autoinducer of virulence as a target for vaccine and therapy against *Staphylococcus aureus*. *Science* **1998**, *280* (5362), 438–440.
 54. Singh, P.K.; Parsek, M.R.; Greenberg, E.P.; Welsh, M.J. A component of innate immunity prevents bacterial biofilm development. *Nature* **2002**, *417* (6888), 552–555.
 55. Lyte, M.; Freestone, P.P.; Neal, C.P.; Olson, B.A.; Haigh, R.D.; Bayston, R.; Williams, P.H. Stimulation of *Staphylococcus epidermidis* growth and biofilm formation by catecholamine inotropes. *Lancet* **2003**, *361* (9352), 130–135.
 56. Nagino, K.; Kobayashi, H. Influence of macrolides on mucoid alginate biosynthetic enzyme from *Pseudomonas aeruginosa*. *Clin. Microbiol. Infect.* **1997**, *3* (4), 432–439.
 57. van der Mei, H.C.; Free, R.H.; Elving, G.J.; Van Weissenbruch, R.; Albers, F.W.; Busscher, H.J. Effect of probiotic bacteria on prevalence of yeasts in oropharyngeal biofilms on silicone rubber voice prostheses in vitro. *J. Med. Microbiol.* **2000**, *49* (8), 713–718.
 58. Free, R.H.; Busscher, H.J.; Elving, G.J.; van der Mei, H.C.; van Weissenbruch, R.; Albers, F.W. Biofilm formation on voice prostheses: In vitro influence of probiotics. *Ann. Otol. Rhinol. Laryngol.* **2001**, *110* (10), 946–951.

Inner Ear Implants

Graeme M. Clark

University of Melbourne, East Melbourne, Victoria, Australia

INTRODUCTION

The cochlear implant is an electronic device that brings useful hearing to severely to profoundly deaf people through multiple-channel electrical stimulation of the auditory nerves in the inner ear. This is required if their inner ears are so badly damaged by injury and disease, or so inadequately developed, that they cannot provide sufficient hearing for communication, even when the sound is amplified with a hearing aid. By stimulating the nerve directly with patterns of electrical pulses, the implant bypasses the normal function of the sense organ of hearing in the inner ear to partially reproduce the coding of sound. It consists of a wearable speech processor that picks up sound with a microphone, analyzes the signal, and then sends it by radio waves to the implanted receiver stimulator, which decodes the message and stimulates the electrode wires inserted into the inner ear (Fig. 1).

NORMAL HEARING

With normal hearing, the external ear or pinna collects sound, which then passes along the external auditory canal to the eardrum (tympanic membrane) (Fig. 2). The eardrum vibrates to the sound, and these vibrations are transmitted through three small articulated bones (ossicles)—(malleus, incus, and stapes)—in the middle ear to the inner ear or cochlea. The ossicles amplify the sound waves, and this overcomes the greater resistance of fluid in the inner ear to the transmission of sound. The transmission to the fluid in the inner ear is via the stapes, whose footplate lies in an opening in the inner ear (oval window). The vibrations in the fluid cause a membrane placed across the cochlea (the basilar membrane), on which lies the sense organ of hearing (organ of Corti), to vibrate (Fig. 3). The basilar membrane acts as a filter for sound frequencies, with the high frequencies, producing a maximal vibration at the basal end (i.e., where the sound enters) and the lower frequencies at the more distant apical region. The hair cells of the organ of Corti convert sound waves into electrical signals. This electrical activity initiates temporal and spatial patterns of action potentials in the auditory nerve and higher brain centers, thus encoding the sound. Deafness occurs if

the structures or pathways are damaged or diseased. Deafness is referred to as conductive if the sound cannot be conducted to the inner ear. Deafness is sensorineural if the cochlea or auditory nerve is damaged.

PRINCIPLES OF THE COCHLEAR IMPLANT

Principles underlying the cochlear implant are firstly the partial reproduction of the coding of frequency and intensity of sound. The coding of frequency is illustrated in Fig. 4. As shown, the temporal coding of frequency is the result of action potentials locked to the same phase of the sine wave and the detection of this information through coincidences in the arrival of the action potentials at central auditory neurons. The place coding of frequency first occurs as a result of filtering by the basilar membrane of the cochlea, with the maximal displacement of the membrane for the high frequencies in the basal turn, moving to the apical turn for the low frequencies. The filtering is due to the mechanical properties of the membrane, which is stiffer at the basal end. This ordering of frequencies on a spatial basis is preserved throughout the auditory system via connections between the cochlea and the higher centers in the central auditory pathways. The spatial or tonotopic organization of the brain is used with a cochlear implant electrode array to stimulate different fibers in the cochlea, conveying frequency information on a place basis. The auditory nerves in the inner ear are splayed out for the transmission of frequency information on a place-coding basis. The amplitude of the sound is converted into current level.

The reproduction of the coding of sound frequency and amplitude with electrical stimulation is imperfect, and thus the interface between the world of sound and the central auditory pathways can be considered an electroneural “bottle-neck.” For this reason, the information from speech in particular must be carefully chosen so that the most relevant data is effectively coded for transmission along the central pathways.

Cochlear implants are effective in providing speech understanding only if they are multichannel, and improved place coding of sound has been responsible for the greatest advances in cochlear implant speech perception to the point where the results are comparable

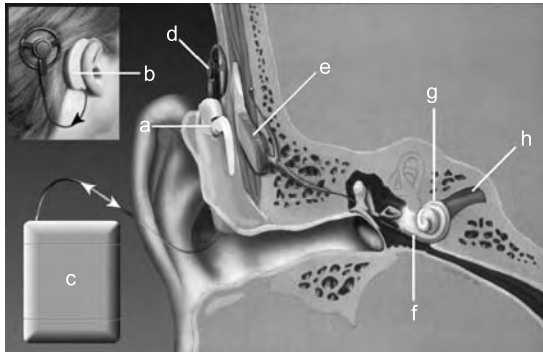


Fig. 1 The cochlear implant. The components are a) microphone; b) behind-the-ear speech processor; c) body-worn speech processor; d) transmitting aerial; e) receiver-stimulator; f) electrode bundle; g) inner ear (cochlea); h) auditory or cochlear nerve. (From Ref. [1].)

to someone with a severe loss of hearing using a hearing aid to good effect.

IMPLEMENTATION OF SPEECH PROCESSOR

The external speech-processing unit of a cochlear implant is now small enough to be worn behind the ear, or a more versatile one for an extended range of advanced strategies can be worn on a belt or in the pocket (Fig. 1). The speech processor has a directional microphone that converts sound into electrical voltages, and these are sent to the speech processor. The speech processor filters this waveform into frequency bands. The outputs of the filters are referred to a map containing the patient's electrical current thresholds and comfortable listening levels for each individual electrode in the inner ear. A code is produced for the

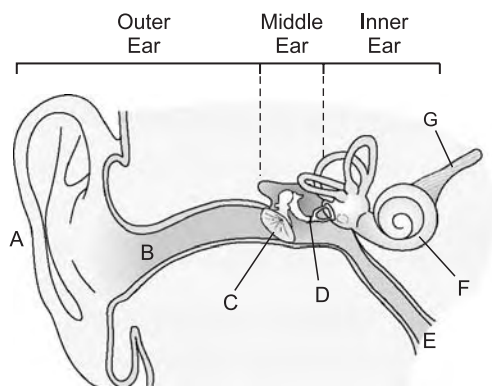


Fig. 2 Diagram of the outer, middle, and inner ears and auditory nerve passing to the central auditory system. A) pinna; B) external auditory canal; C) eardrum; D) ossicles; E) Eustachian tube; F) inner ear; G) auditory (cochlear) nerve. (From Ref. [2].)

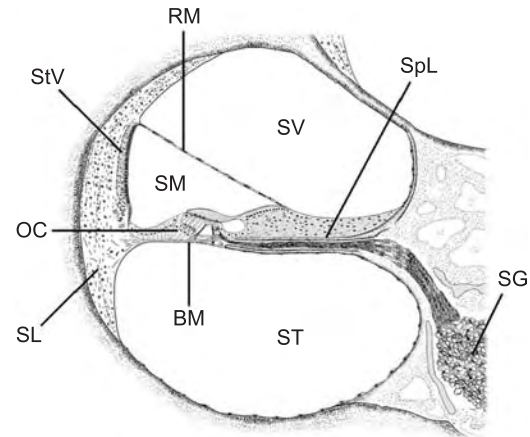


Fig. 3 A cross section of the human cochlea showing the scala tympani (ST), the scala media (SM), and the scala vestibuli (SV), separated by the basilar membrane (BM) and the Reissner's membrane (RM). The basilar membrane is attached externally to the spiral ligament (SL) and internally to the spiral lamina (SpL). The stria vascularis (StV) generates the endolymphatic potential in the scala media. This powers the production of cochlear microphonics in the organ of Corti (OC), and the resulting action potentials pass along the peripheral processes in the spiral lamina to the spiral ganglion cells (SG) in the spiral canal in the modiolus. (From Ref. [2].)

stimulus parameters via electrode site and current level. To represent the speech signal at each instant in time, this code together with power is transmitted by radio waves via a circular aerial through intact skin to the implanted receiver-stimulator section that is embedded in the mastoid bone behind the ear. The receiver-stimulator decodes the signal and produces a pattern of stimulus currents in a bundle of electrodes that are inserted around the first turn of the inner ear so that the auditory nerve fibers, particularly in the speech frequency range, can be stimulated. The electrode bundle lies close to but not attached to the spiral ganglion cells in the inner ear, the first station for processing information.

The sound is converted into electrical signals by a directional microphone that increases in sensitivity from 1500 Hz up to approximately 5000 Hz at 6 dB per octave before falling off at 18 dB per octave. This emphasizes the high frequencies of speech that are important for intelligibility but low in intensity. The signal passes to the speech processor via a cable. The speech processor has an analogue front end to amplify the signal to a level that can be best handled by the speech processor (Fig. 5). The front end also has an automatic gain control (AGC), which is a compression amplifier to keep the variations in the amplitude of the speech signal within an acceptable range. The time taken for the AGC to operate (attack) and cease (release) must be carefully controlled for optimal processing of the speech sounds.

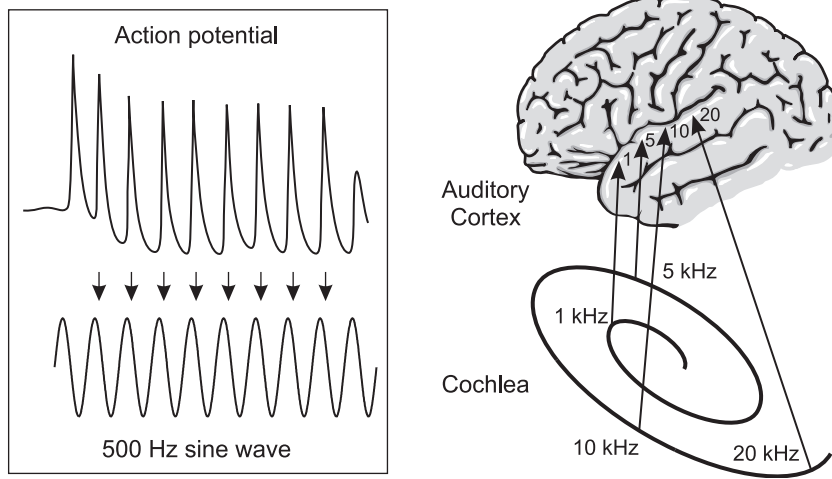


Fig. 4 Left: the temporal showing the phase-locking of the brainstem cell responses to the sine wave. Right: the place coding of frequency showing how the sound is first filtered by the cochlea and then connected to higher brain centers in an orderly way so that a frequency scale is preserved. (From Ref. [3].)

The next important function for the speech processor is to filter the sounds into component frequencies and detect the energy level in each frequency band (Fig. 6). Switched-capacitor filters are used as they have greater flexibility in selecting the center frequencies and the slope of the upper and lower frequency cutoffs. The resulting analogue representation of the sound is digitized with an analogue-to-digital converter (ADC) for further processing. Alternatively, the signal from the front end is first digitized and then filtered using a Fast Fourier Transform (FFT). There are 20 filters for the Nucleus SPEAK and ACE strategies (Fig. 7). The first seven have center frequencies distributed lineally over the range from 140 to 1000 Hz, and the remaining ones are logarithmically spaced up to 6000 Hz, the upper limit for speech frequencies. Although there is a logarithmic

distribution of frequency versus distance along the basilar membrane, the above frequency to electrode allocation gives a better emphasis to the frequencies of importance for speech intelligibility. The fixed filter strategies for presenting the CIS strategy have six to eight filters with center frequencies ranging from 250 to 5500 Hz and distributed on a logarithmic basis.

The correspondence between formant or filter frequency and electrode can be determined from psychophysical studies. The electrode to be stimulated for each frequency band is assigned from a table in the digital memory. The person's threshold and maximum comfortable level for each electrode is also stored in memory as a reference map for setting the stimulus levels to be transmitted to the receiver-stimulator. Microprocessors are used to control the signals for implementing the different strategies.

The data from the speech processor to instruct the receiver-stimulator on the electrical stimulus settings are transferred serially as a digital stream of information. This is more effective than analogue transmission as it provides immunity to noise and component variations. The signal parameters from the microprocessor are fed to the transmitter section. The microprocessor controls the times at which the electrical parameter values are sampled and configures the radio frequency signal for transmission to the implanted receiver-stimulator with alignment ensured through magnets incorporated in the centers of both coils. The transmitted code is made up of a digital data stream representing the stimuli and transmitted by pulsing the radio frequency (RF) carrier. Data for stimulation are transmitted in frames with the Nucleus system. The data in each frame are coded as the number of cycles in bursts of RF. Each burst consists of a number of cells or just discriminable RF cycles. Successive frames are separated by an interframe gap.

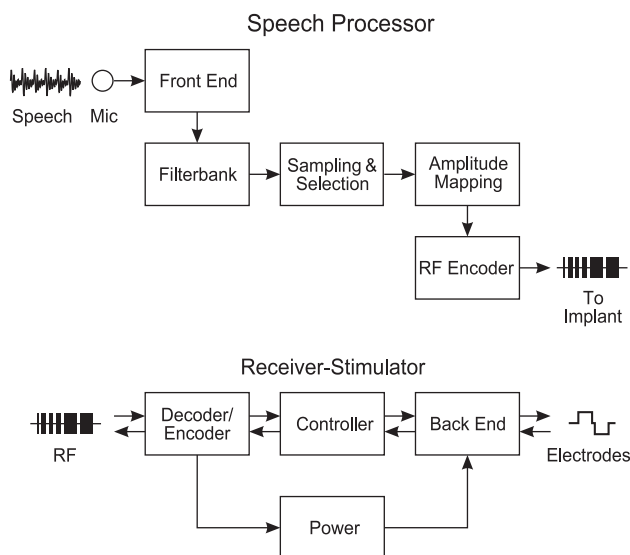


Fig. 5 Top: the principal components of a speech processor. Bottom: the principal components of the receiver-stimulator. (From Ref. [2].)

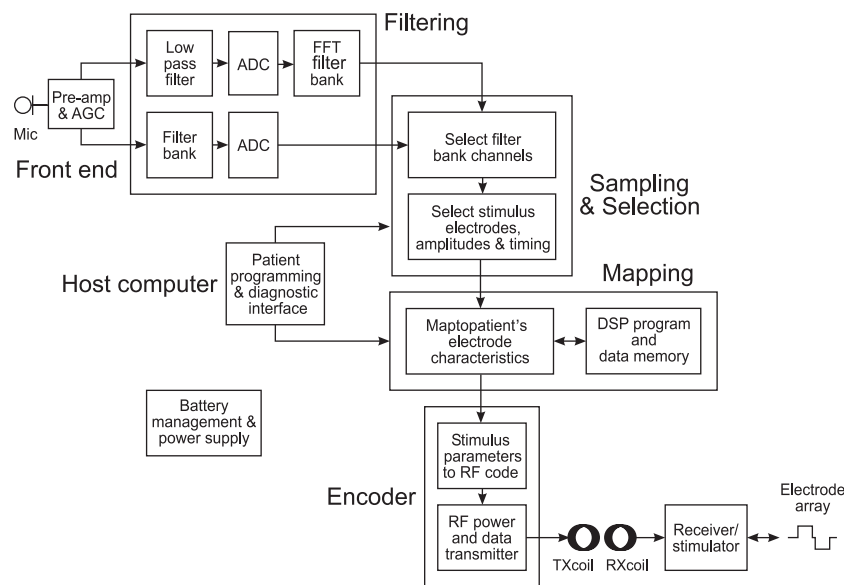


Fig. 6 The Nucleus behind-the-ear and body-worn speech processors are implemented using either a standard filter bank or Fast Fourier transform (FFT) filter bank, respectively. The front end sends the signal to a signal processor chip via a filter bank or to a digital signal processor (DSP) chip, which carries out an FFT and signal processing. The signal processor selects the filter bank channels and the stimulus electrodes and amplitudes and then maps these to the patient's requirements. An encoder section converts the stimulus parameters to a code for transmitting to the receiver-stimulator on a radio-frequency (RF) signal together with power to operate the device. ADC, A-to-D converter; RX, receiver; TX, transmitter. (From Ref. [4].)

IMPLEMENTATION OF RECEIVER-STIMULATOR

The receiver-stimulator decodes the information into instructions for the selection of the electrode mode of stimulation (bipolar, common ground, or monopolar), current level, pulse width, and interpulse gap. Power to operate the receiver-stimulator is also transmitted by the radio frequency carrier. The receiver-stimulator has circuitry for decoding the transmitted signal and also for converting the energy in the radio frequency signal into power for its operation and stimulation of the auditory nerve fibers.

Electrical charge is the important parameter for stimulating nerves, and this will vary if there are alterations in the impedance between the electrode and the tissue. To avoid this problem, constant current stimulation is used. Bipolar stimulation localizes the current

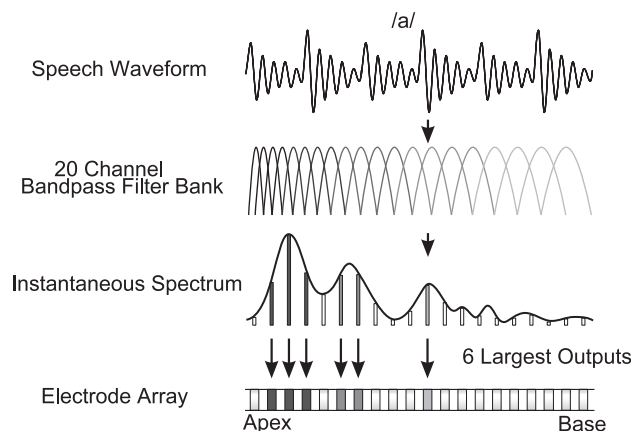


Fig. 7 The Nucleus Spectral Maxima (SPEAK) sound processor. (From Ref. [5].)

to separate groups of nerve fibers. In this case, the electrical current passes between one of the electrodes and a neighboring one usually separated by one intervening electrode. Common-ground stimulation provides localized stimulation, and this occurs with the passage of current from an active electrode to all the other electrodes on the array shorted to form a common ground. Finally, monopolar stimulation provides localized stimulation if the electrode is placed close to the nerve fibers to be stimulated, and it has the added advantage in that a lower threshold is achieved, enabling smaller processors to be used with longer battery life. It is important that the electrical pulses be biphasic so that one cancels out the other to ensure there is no buildup of DC at the electrode-tissue interface as this can be damaging to tissue. In addition, stimulators have a capacitor in series to ensure there is a high impedance for any DC in the circuit. Another important principle in stimulating the nerve fibers in the cochlea is non-simultaneous stimulation; otherwise there will be an interaction of the electrical fields and unpredictable variations in loudness. The number of stimulus channels or electrodes to be used will vary depending on whether the strategy is formant extraction, spectral selection, or fixed filter. Multiple-channel devices are used with channel numbers varying from 6 to 22. The fixed-filter strategies require only up to eight channels. Formant selection or spectral maxima strategies perform better as they use up to 22 channels. Studies with acoustic models and electrical stimulation indicate that up to 22 channels are especially necessary for the best speech perception in the presence of background noise.

The intensity of the stimulus can be varied through the current level or the pulse width. Both current and pulse width alter the charge presented to the nerve

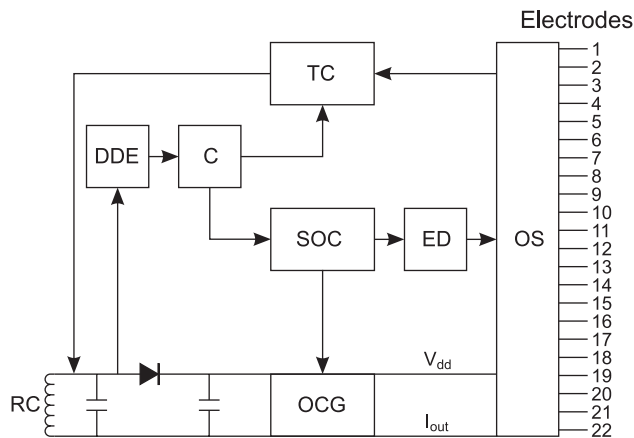


Fig. 8 A block diagram of the sections of the receiver-stimulator circuitry. RC, receiving coil; DDE, data decoder/encoder; C, clock; SOC, stimulus output controller; ED, electrode decoder; OCG, output current generator; S, switches; TC, telemetry controller. (From Ref. [2].)

membrane. There are other parameters such as rate and total charge that also affect neural stimulation and the perception of loudness. It is also important that charge density and charge per phase be kept within certain strict limits as they can cause corrosion of the electrode and gassing if a charge density of $300 \mu\text{C cm}^2$ geometric phase is exceeded. Lower charge density levels may still cause neural damage especially if presented at high rates of stimulation. Consequently, the charge and charge density of the neural stimulators need to be controlled and will depend upon the strategy to be used as well as the geometry of the electrodes. Animal experimental studies are required to confirm the safety and efficacy of the speech processing strategies and electrode designs.

In implementing the instructions from the speech processor, the receiver-stimulator has circuitry that sends the data via a clock unit to control the timing and then to an output controller to determine the instructions required for producing the stimulus parameters (Fig. 8). These instructions are then fed through an electrode decoder to control the mode and duration of stimulation. With bipolar stimulation, the current flows first in one direction when the voltage rail V_{dd} is connected to the V_{ss} and then in a reverse direction to complete the stimulus. The second phase of the pulse is identical in magnitude but opposite in sign to the first. This technique makes it possible to match the current in both phases of the pulse and not produce an imbalance.

In addition, receiver stimulators have telemetry. This allows the voltages on electrodes generated while delivering a stimulus pulse to be transmitted back to the external programming system (Fig. 9). The voltages can help determine the tissue impedance around the electrodes and thus assess any pathological changes

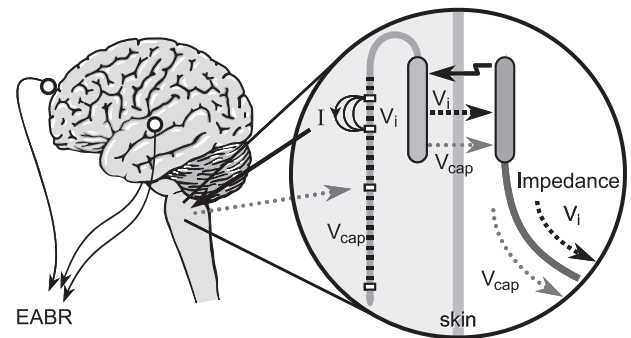


Fig. 9 A diagram of telemetry showing I , the stimulus current; V_i , the voltage between electrodes for the calculation of the impedance; V_{cap} , the compound action potential or voltage from the auditory nerve and brainstem; EABR, the electrically evoked brainstem response and cortical evoked potential. (From Ref. [6].)

occurring in the neighboring tissue. In addition, the telemetry permits the recording of compound nerve action potentials from the auditory nerve. This enables the neural stimulus thresholds and dynamic ranges to be gauged for infants and young children who cannot easily respond and explain the sensations they are receiving. Thus the speech processing maps in young children can be made more efficiently.

ELECTRONEURAL INTERFACE

The electrodes required to stimulate the auditory nerve are insinuated along the basal turn of the cochlear spiral at a surgical operation (Fig. 10). They are introduced through a small opening drilled in the bone overlying the scala tympani, one of the fluid compartments in the inner ear. The electrodes either lie



Fig. 10 A drawing of the banded array passing around the basal turn of the cochlea in the scala tympani. (From Ref. [7].)

along the outer wall or have a precurved spiral so that they come to lie closer to the neural elements in the central column of the cochlea.

The electrodes must be inserted with the right mechanical properties so that they will pass up the tightening spiral far enough to lie close to the fibers conducting the speech frequencies to the brain with minimal trauma. It was found they needed to have graded stiffness with more flexibility at the tip so that the delicate inner ear structures were not injured, which would in turn lead to the loss of the neural elements it was necessary to stimulate. The electrode carrier also needed to be free-fitting to minimize trauma and not stretch or tear vital structures. Furthermore, it had to be smooth without protruding electrodes so that it could be easily taken out and another one inserted should there be any failure or breakages of the wires.

It is essential to make sure that middle ear infection will not extend around the electrode bundle into the inner ear and thence to the meningeal linings of the brain leading to serious meningitis, particularly in young children. Experimental studies at the University of Melbourne showed the importance of using a fibrous tissue seal packed around the entry point to prevent the ingress of infection (Fig. 11). The series of basic experimental studies demonstrated that a sheath extending from middle ear around the single component electrode into the cochlea is important in preventing meningitis. This sheath enables three lines of defense by the body to counter the entry of infection from the middle ear. The first is through a change in the lining cells that extend around the sheath to ones that produce mucus that slows down the spread of bacteria, the hair cells of these lining cells sweep the

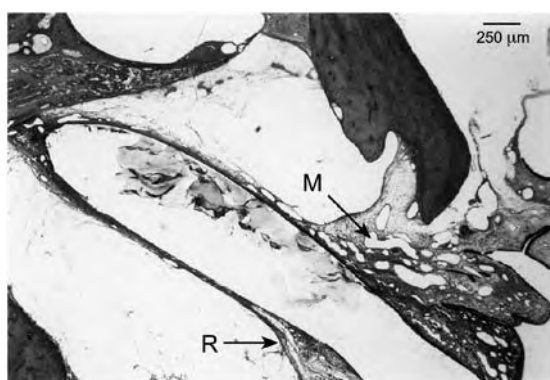


Fig. 11 Photomicrograph showing a cat cochlea and a spontaneous middle ear infection several weeks after implantation. An envelope has formed around the prosthesis inside the scala tympani, the round window membrane is thickened (R), and the round window niche is filled with mucosal folds (M); the organ of Corti and neurons are intact. (From Ref. [8].)

mucus and bacteria away. The second line of defense is through phagocytosis. When the invasive organism causes the blood vessels around the sheath to dilate, the granulocyte white cells enter the tissue and engulf the bacteria. The third line of defense is also facilitated through the sheath around the single component array. The sheath allows the entry of specialized cells; the type B lymphocytes produce antibodies and the type T lymphocytes kill the bacteria directly. Meningitis can, however, be a problem in children who have a congenital deformity of the cochlea, as they often have a large communication between the inner ear and the fluid-filled subarachnoid space around the brain.

CLINICAL ISSUES

Once the receiver-stimulator and the electrode carrier are correctly placed at the operation, the electrode carrier needs to be fixed at a point in the temporal bone that does not move relative to the inner ear in young children. This avoids changes in head growth leading to its slippage.

The cochlear implant and speech processor have been used in tens of thousands of adults and children in over 120 countries. Clinical studies have shown that speech processing strategies that extract speech frequencies and present them to appropriate electrodes around the cochlea on a frequency to place basis as well as presenting voicing as rate of stimulation or amplitude variations produce very significant improvements in speech understanding (Fig. 12). The average deaf adult who had hearing before going deaf can understand running speech without the need to lip-read at a level comparable with that of a person with severe hearing loss who has residual hearing and uses a

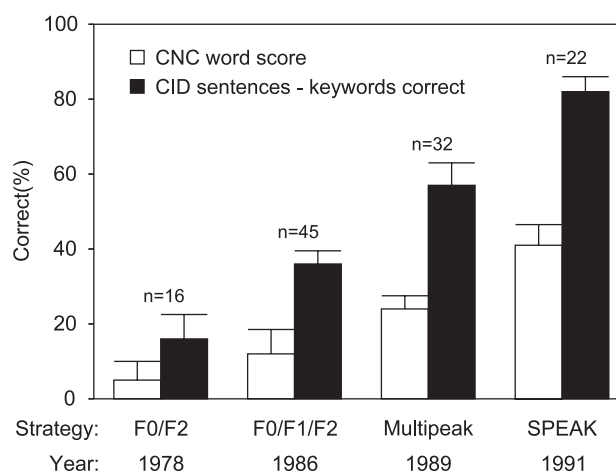


Fig. 12 The open-set CID sentence and CNC word scores for the Nucleus F0/F2, F0/F1/F2, "Multippeak," and SPEAK strategies. (From Ref. [9].)

hearing aid. Similarly, the same applies to children who are born deaf, especially when they are operated on at a young age even as early as 12 months.

With young children it is also important for them to speak correctly and to develop language. There is an interrelationship between their ability to perceive speech and their language development. It has been found, however, that like children with severe hearing loss and a hearing aid, the language levels on average do not always reach those of children with excellent hearing but can do so with extra training.

The level of success can be partially predicted in adults and children. The important factors are age when deafened, age at implantation, duration of deafness, duration of implantation, the presence of residual hearing, speech reading ability, speech processing strategy, the communication strategy before and after surgery, and other factors such as multiple handicaps.

The Nucleus first and second formant and “Multi-peak” speech processing strategies were approved by the U.S. Food and Drug Administration as safe and effective for children older than two years in 1990. Studies on children also showed they obtained better results when they were operated on at a younger age.

Before operating on children under two years, it was essential to evaluate three safety issues. These were firstly ensuring that infections in the middle ear, which are frequent in this age group, would have minimal risk of extending to the inner ear and lead to meningitis; secondly determining that the effect of head growth, which is rapid at this age, would not lead to the extraction of the lead wires from the inner ear; and thirdly establishing that electrical stimulation did not have an adverse effect on the maturing nervous system. Once these issues were found not to present special difficulties for children under two, operations were carried out down to 12 months, especially as it is now possible to diagnose hearing loss accurately with steady-state evoked response audiometry.

To further improve speech processing strategies, it was found desirable to develop an electrode array that could pass around the spiral of the inner ear and lie close to the central axis where the ganglion cells of the hearing nerve lie. This was a complex bioengineering task. The electrode array had to be made as a spiral and then held straight with a stylet. After partial insertion, the stylet is withdrawn and then the electrode advanced to curl around the cochlea. These electrodes achieve lower thresholds of stimulation and more localized stimulation of the nerve elements.

CONCLUSION

The cochlear implant has been a major achievement in the management of profoundly deaf children. It is

the first significant improvement since Sign Language of the Deaf was developed at the Paris Deaf School 200 years ago. It is also the first clinically established interface between the environment and the central nervous system and a major step forward in the corrective management of brain disorders.

ARTICLES OF FURTHER INTEREST

Hearing Mechanisms; Implanted Medical Devices, Power Sources for; Tinnitus Devices

REFERENCES

1. Clark, G.M. *Sounds from Silence*; Allen & Unwin: Sydney, 2000.
2. Clark, G.M. *Cochlear Implants: Fundamentals and Applications*; Springer-Verlag: New York. *in press*.
3. Clark, G.M. Electrical stimulation of the auditory nerve: The coding of frequency, the perception of pitch and the development of cochlear implant speech processing strategies for profoundly deaf people. *Clin. Exp. Pharmacol. Physiol.* **1996**, *23*, 766–776.
4. Clark, G.M. Cochlear Implants. In *Diseases of the Ear*, 1st Ed.; Wright, A., Ludman, H., Eds.; Edward Arnold: London, 1998; 149–163.
5. Clark, G.M. Cochlear implants: Climbing new mountains. The Graham Fraser memorial lecture 2001. *Cochlear Implants Int.* **2001**, *2*, 75–97.
6. Clark, G.M. Learning to Understand Speech with the Cochlear Implant. In *Perceptual Learning*, 1st Ed.; Fahle, M., Poggio, T., Eds.; MIT Press: Cambridge, MA, 2002; 147–160.
7. *Surgical Procedure Manual: Nucleus 22 Channel Cochlear Implant System, Issue 5*; Cochlear Corporation: Englewood, CO, 1987.
8. Shepherd, R.K.; Franz, B.K-H.; Clark, G.M. The Biocompatibility and Safety of Cochlear Prostheses. In *Cochlear Prostheses*, 1st Ed.; Clark, G.M., Tong, Y.C., Patrick, J.F., Eds.; Churchill Livingstone: London, 1990; 69–98.
9. Clark, G.M. The cochlear implant: A search for answers. *Cochlear Implants Int.* **2000**, *1*, 1–17.

FURTHER READING

More detailed information is available in the following texts:

- Clark, G.M. *The University of Melbourne—Nucleus Multi-Electrode Cochlear Implant*; Advances in Oto-Rhino-Laryngology, Karger: Basel, 1987; Vol. 38.
- Cochlear Implantation for Infants and Children—Advances*; Clark, G.M., Cowan, R.S.C., Dowell, R.C., Eds.; Singular Publishing Group Inc.: San Diego, 1997.

Insulin Delivery

Robert S. Parker

Department of Chemical and Petroleum Engineering, University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.

INTRODUCTION

Insulin is an endogenous protein that facilitates the cellular uptake of glucose in the body. After release from the pancreatic β -cells, circulating insulin interacts with cellular receptors that, through a cascade of intracellular events, increases the transport of glucose across the cell membrane. Patients suffering from diabetes have either little or no endogenous insulin production (type I diabetes) or a combination of insulin resistance and insufficient pancreatic insulin release (type II diabetes). Type I and type II diabetes sufferers may require insulin, and a key issue in disease treatment is the delivery of the hormone so as to regulate patient glucose levels within a fairly narrow band (70 to 120 mg/dl). Classical delivery techniques have relied on the administration of insulin into a subcutaneous depot. A variety of insulin formulations have been employed, each characterized by its respective profile of insulin release into the bloodstream. Given that optimal glucose-control performance is limited to a significant extent by the insulin delivery mechanism, advances in this area are necessary.^[1] These advances may be accomplished in any of three primary areas, discussed in greater depth in the following sections: route of delivery, insulin formulation, and novel approaches to insulin delivery.

DELIVERY ROUTES FOR INSULIN

At the time of Banting and Best (1923), early insulin preparations were delivered intramuscularly and at high volume. This resulted in significant pain for the patient, as well as medical complications (e.g., abscess) due to the large required delivery volume. Developments since have focused on reducing the volume of the insulin preparation delivered subcutaneously, increasing patient quality of life by developing less invasive methods for insulin delivery and normalizing the glucose distribution in the patient by using delivery routes other than the subcutaneous (subQ). This last aim is partially driven by the natural pathway for insulin release.

Healthy patients release insulin into the portal vein, such that insulin first passes through the liver and is

then distributed throughout the systemic circulation. The high concentrations of insulin in the liver promote storage of glucose as glycogen, and the systemic insulin levels are decreased by the so-called “first pass effect,” in which the liver extracts a significant fraction of the insulin entering the organ (approximately 40%).^[2] In contrast, the exogenous delivery of insulin is rarely to a space that drains into the portal vein (implantable pumps excepted). The result is a glucose distribution weighted heavily toward the periphery because the lower liver insulin concentrations reduce the amount of glucose stored as glycogen in the liver and the elevated systemic insulin levels promote glucose storage in the muscle and fat cells. Based on the preceding discussion, the choice of delivery route plays a significant role in the quality of treatment, and hence patient quality of life.

Subcutaneous Delivery

The classical insulin delivery method has been injection to the subcutaneous tissue space. This creates a small depot of insulin that diffuses into the circulation over a period of time. A variety of factors impact the effectiveness of this treatment.^[3] The stability of an insulin formulation will establish the maximum theoretical bioavailability of subQ insulin. Practically, the bioavailability may be lower than this maximum due to the concentration-dependent nature of insulin release from subcutaneous depots. Surrounding tissue may metabolize some of the injected insulin; also, after a period of time, there may be an insufficient concentration of insulin in the depot to further drive diffusion to the bloodstream. In addition to the magnitude of insulin that reaches the circulation, rate of appearance in the bloodstream is important. The rate of absorption from the depot affects such important characteristics as peak insulin concentration, time of peak concentration, and duration of insulin action. The rate and magnitude of insulin effect can be altered by changing the formulation and biochemical structure of the injected insulin, as discussed in the section titled “Insulin Formulations.”

As an alternate path to multiple daily injection therapy, there has been an ongoing push to develop insulin pumps.^[4] Typically these pumps contain a

two- to three-day supply of insulin that is administered through a flexible tube; the distal end resides in the (typically abdominal) subcutaneous tissue of the patient. This transcutaneous line is positioned by the patient via a needle, and the location is changed every 2–3 days to maintain viability of the delivery site. The development of patient-friendly pumps and intermediate-duration indwelling tubing has resulted in continuous subcutaneous insulin infusion (CSII) therapy. Here, regular insulin is infused continuously at a low rate to a subcutaneous depot over a period of a few days. This produces a more consistent basal insulin pattern than multiple daily injections of longer-acting insulin preparations. Generally, the aim is to satisfy basal insulin requirements using CSII and to use injection therapy to address meal effects. While this doesn't eliminate the need for injections altogether, it does reduce the necessary number of transcutaneous punctures. Alternatively, some pumps include priming mechanisms that allow them to deliver boluses or (square-)waves of insulin. In these cases, the patient can eliminate injection therapy by instead activating the pump before consuming a meal. An issue with CSII is the possibility that the infusion port will become blocked, thereby stopping insulin delivery to the patient. As long as patients consistently monitor their blood glucose, effective treatment can continue. Recent studies have shown the effectiveness of CSII therapy versus traditional multiple daily injections in reducing HbA_{1c}, a measure of “overall” glucose control performance in diabetic patients.^[5]

Alternate Routes of Administration

The development of non-subQ insulin delivery approaches is generally motivated by the desire(s) 1) to improve the pharmacokinetics or pharmacodynamics of either acute or basal insulin administration, and/or 2) to reduce the lifestyle impact of multiple daily subQ insulin injections. Different routes offer advantages in either or both goals; however, there are typically additional complications associated with these alternate delivery locations that challenge scientists, engineers, and clinical practitioners.

Oral enteric insulin

A primary challenge in developing an oral insulin dosage form that targets intestinal absorption is the need to protect the sensitive protein from denaturation by stomach acids and intestinal enzymes. Even with successful arrival of insulin in the intestinal tract, the bioavailability of the protein is still significantly less than that achieved by injection therapy.^[6,7] Finally, the absorption profile is poorly understood at present, and

the variability of meal content could significantly complicate the establishment of repeatable uptake. The two primary advantages associated with this approach provide sufficient incentive to continue development in spite of the aforementioned concerns. A key advantage of this technique is patient compliance, as it is far more convenient and acceptable (in the personal and societal sense) to take a pill than a self-administered injection. Furthermore, uptake of insulin from the intestinal tract is into the portal vein. This is an important biological advantage, as the insulin available to the body undergoes the “first pass effect” in the liver. This should lead to a normalization of the glucose distribution in the body (liver instead of peripheral storage) and an elimination of the undesirable systemic hyperinsulinemia observed in peripheral delivery.

Ocular, nasal, and buccal mucosa delivery

Ocularly delivered drugs are generally swept away by tearing and blinking to the nasal mucosa, where they are absorbed.^[8] The properties of mucosal absorption are reasonably similar, thereby allowing simultaneous discussion of ocular, nasal, and buccal (or mouth) delivery. All of these approaches avoid the pain associated with injection therapy. Also, transport across mucosal layers is quite fast, and buccal studies have shown insulin appearance in the circulation in about 10 minutes.^[9] However, there are some issues in these delivery pathways that limit the utility of insulin delivery. Ocular and nasal delivery often requires absorption enhancers to improve the uptake of larger (6000 kDa) proteins and to reduce the clearance rate from the nasal cavity. Unfortunately, many of these enhancers are irritants of some form (eye stinging, mucosal irritation, etc.). In addition, insulin bioavailability is less than 10% for nasal absorption, so that significantly greater amounts of insulin will be required. By using an ocular device (gel insert), the duration of insulin effect and bioavailability may be markedly increased; the long-term effects of this approach have yet to be evaluated. Buccal delivery is another option, and a mist-based system has been developed.^[9] Although the bioavailability appears to be markedly reduced, more than 10 times less than subcutaneous injection, the systemic insulin profile has a short time to peak (about 30 minutes) that would aid in rapid reduction of meal-induced hyperglycemia.

Insulin inhalation

Inhalation of insulin is another route that avoids the pain associated with injection therapy. In this case, either liquid droplets or powders with size 1–3 μm are inhaled with slow even breaths so that the drug will be

deposited in the alveoli of the lungs.^[10] The particle size is important as particles outside this size range are deposited in the mouth or on the walls of the bronchial tubes. While the bioavailability is reduced versus subcutaneous injection (10–40% of total inhaled dose absorbed),^[11,12] the concentration versus time profile of inhaled regular insulin is similar to that of insulin lispro delivered subcutaneously.^[12] Along with the apparent advantages, there are several concerns. There have been no studies evaluating long-term safety of inhaled insulin formulations, although the expected levels of particulates are less than the current recommended maximum set by the American Council of Governmental Industrial Hygienists. Inhaled insulin response is reproducible, but any factor influencing lung operation (lung disease, common cold, etc.) could alter the drug absorption. Since 80% of the inhaled dose reaches the alveoli, while only 10–40% reaches the circulation in the healthy patient, overabsorption could lead to potentially lethal hypoglycemia. Finally, the use of inhaled insulin is primarily for acute dosing and meal compensation, as there are no ongoing studies on inhalation therapy using long-acting insulin formulations.

Transcutaneous transport

Delivering drugs across the skin offers some significant advantages versus alternate delivery routes, including no gastrointestinal degradation (oral), steady delivery (acute therapies), and better compliance (injection). However, a major challenge is the transport of large compounds, such as insulin, across the stratum corneum—the upper layer of skin and the generally accepted rate-limiting layer of the epidermis.^[13,14] By changing from hypodermic to microneedles, the depth of penetration can be reduced, thereby reducing the level of pain associated with insulin delivery.^[15] However, the rate of insulin delivery would be reduced due to the significant reduction in needle size. Increasing skin permeability provides an alternative to puncturing the skin with micro- or hypodermic needles. Methods for increasing transcutaneous transport include chemical, electrical, and sonic approaches. Low-frequency ultrasound (~20 kHz) has proven effective for the transcutaneous delivery of insulin, which is similar in size (molecular weight) to insulin.^[13] An added benefit of ultrasound is the pain-free nature of the technology. Further improvements in transport rate result from synergistic effects between ultrasound and chemical or electrical (e.g., electroporation, iontophoresis) means. Synergistic approaches may result in increased skin irritation by chemical transport enhancers or the electric voltages associated with iontophoresis and electroporation.

Intravenous administration

In order to mimic the human body, insulin would have to be administered intravenously (IV) to the patient. While this has significant advantages, such as reducing the rate of progression of diabetic nephropathy,^[16] it has drawbacks of equal or greater importance. Clearly, the bioavailability of a compound delivered IV is 100%. Furthermore, there is no delay from administration to changes in the circulating concentration, so regular insulin can be used and the onset of action would be nearly immediate. Likewise, a stoppage of insulin would immediately eliminate insulin appearance; this would be important in the case of exercise or hypoglycemia compensation. However, this immediate effect of stopped insulin would be problematic if undesired; glucose concentration rise would initiate much sooner than under other forms of therapy. Also, it is impractical given current technology to maintain an indwelling intravenous line for drug delivery. Blockage by immunoreaction may occur, sterility for a transcutaneous port would be a major concern, and stability and rupture of the venous blood vessel are serious safety issues. With the possible exception of implantable pumps, which could deliver to the portal vein, the point of insulin delivery would likely be systemic. Hence the benefits of IV delivery would be lost, to an extent, by not having the liver “first pass effect,” and therefore the misdistribution of stored glucose would occur including elevated peripheral glycogen levels and a reduced capacity of the liver to buffer glucose demand changes. So while IV delivery of insulin to the portal vein is the ideal scenario, safety and technological concerns currently limit this approach to clinical (intensive care) settings and basal mode implantable pumps.

INSULIN FORMULATIONS

Insulin has traditionally been delivered into subQ depots, from which it would diffuse into the circulation. Effective subQ therapy has been dependent upon a variety of issues, including insulin preparation stability, bioavailability, biocompatibility, and rate of release from the depot. By selecting the proper dose and formulation of insulin to accomplish glucose regulation, which typically requires a finite set of insulin preparations and doses, reasonable control of glucose can be achieved.^[17,18]

Two primary issues must be addressed in insulin delivery, and hence these impact the formulation of delivered insulins: basal insulin replacement and acute insulin demand. In healthy patients, the pancreas releases a continuous stream of insulin to satisfy the basal tissue need.^[19] This basal secretion maintains the blood glucose concentration within the normoglycemic

range of approximately 70 to 120 mg/dl.^[17] In addition to this requirement, acute perturbations, e.g., meals, can dramatically upset the physiological glucose concentration. In this case, additional insulin is typically administered to reduce the hyperglycemic (>150 mg/dl) excursion. Of concern is that aggressive acute dosing may lead to an increased risk of hypoglycemia (glucose <60 mg/dl) after the meal. In general, the basal and acute insulin requirements are met using different insulin formulations and schedules.

Acute Insulin Delivery

The action of insulin produced by the pancreas is rapid. This is important, as glucose concentrations often rise in less than 30 minutes after the start of meal consumption. Hence, the body must act quickly to avoid overt hyperglycemia. Likewise, diabetic patients need an insulin formulation that acts quickly after administration. Rapid absorption is also beneficial for hypoglycemia avoidance because short-acting insulin formulations do not have the extended response profiles of subQ-delivered regular or longer-acting insulins. Hence, insulin release from the subcutaneous depot occurs during meal absorption, rather than extending well beyond the duration of meal uptake.

Early regular insulins were bovine or porcine in origin, but advances in recombinant DNA technology and production in *E. coli* made available “regular” human insulin. The pharmacokinetics and pharmacodynamics of regular insulin are identical to natively produced insulin when administered intravenously, but subcutaneous delivery results in a delayed absorption (30 minutes to IV appearance), time of peak action (peak at about 80–230 minutes), and duration of action (4 to 6 hours).^[20] When administered subcutaneously, regular insulin has a tendency to form dimers, hexamers, and aggregates. In order to diffuse out of the subcutaneous depot, insulin must be in monomeric form. Hence, the delays in absorption and time to peak action, as well as the extended absorption profile, are a function of the time required for dissociation of the insulin aggregates, hexamers, and dimers.

Insulin lispro and insulin aspart are regular human insulin derivatives designed to reduce the degree of insulin self-association. Lispro inverts the position of lysine (B29 → B28) and proline (B28 → B29) on the B-chain.^[20] Aspart replaces the B28 proline with aspartic acid.^[20] These two insulins both have rapid action onset (≤15 min), a reduced time to peak (45–60 min), and double the peak concentration of regular insulin, with similar total exposure when delivered subQ.^[20] This time profile is much closer to that of healthy patients, which improves overall glucose control by reducing the postmeal glucose peak magnitude as well as

the frequency of hypoglycemic episodes.^[20] Unfortunately, rapid-acting insulins have some drawbacks. They are more expensive, and the reduced duration of action may result in a more rapid onset of ketoacidosis in the case of infusion pump failure or basal insulin absorption loss. For patients not on intensive insulin therapy, there is also an increase in the number of daily injections required for glucose control. However, given the potential improvements in glucose performance provided by intensive insulin therapy^[17,18] combined with rapid-acting insulin, the case is stronger than ever for patients to switch to more intensive treatment of their disease. In spite of the drawbacks, most patients using injection therapy to treat their disease are currently employing rapid-acting insulin analogues as a part of their therapy. These analogues are also useful for inhalation therapy, as the rapid uptake times improve glucose control for these patients as well.

Basal Insulin Replacement

Basal insulin in the healthy patient is a continuous secretion from the β -cells of the pancreas.^[21] Ideally, the delivery of basal insulin to diabetic patients would match the profile of endogenous healthy-patient production. Key to this approach is the development of long-acting insulins that diffuse out of a subcutaneous depot at rates significantly slower than the regular insulins previously discussed. Early developments of this type include neutral protamine Hagedorn (NPH) and (70% ultra-/30% semi-)lente insulins. Both of these insulins display a peak of action between 4 and 5 hours after subcutaneous injection, followed by a decrease in action after about 6 hours,^[21] with a total duration of action of approximately 12 hours. This peaking and nonflat behavior does not represent native insulin secretion in the healthy patient; furthermore, the peak concentration and subsequent loss of activity for twice-daily injections occurs in the middle of the night and is a significant contributor to overnight hypoglycemia and morning hyperglycemia, respectively.

Like the rapid-acting insulins, amino acid modification of insulin can be used to tune the subcutaneous insulin release profile. Insulin glargine is formed by substituting a glycine for the asparagine on the A-chain at position 21 and adding two arginines to the B-chain after position 30.^[20] The result is an insulin with a modified isoelectric point that precipitates upon administration to subcutaneous tissue. This insulin formulation is peakless and has a duration of action of approximately 24 hours. Hence, a single nightly dose results in an insulin release profile that does not induce overnight hypoglycemia (no insulin peak), counteracts waking hyperglycemia (no loss of action over 12 hours), and approximates the flat healthy-patient

insulin release profile. The use of insulin glargine improves overall glycemic control (as measured by HbA_{1c}) and fasting glucose levels, while reducing the frequency of hypoglycemic episodes experienced by patients.^[20] Two drawbacks are associated with insulin glargine use. First, it cannot be mixed with shorter-acting insulins due to its isoelectric point, resulting in an increase in the number of daily injections.^[20,21] Second, the amount of rapid-acting insulin required at mealtime may be increased.^[20]

INDUSTRIAL CONTRIBUTIONS TO INSULIN DELIVERY

It is estimated that 6% of the United States population suffers from diabetes, and treating this disease requires a collection of products provided by the corporate sector. While it is infeasible to provide a comprehensive review of the industrial contributions to glucose management and insulin delivery, representative corporate contributions are summarized here. In the context of diabetes care, the necessary components include a glucose sensor and an insulin delivery device. For diabetic patients to properly estimate their insulin needs, a measurement of glucose is required. In the current state of clinical practice, this has meant a (finger-)prick to obtain a drop or two of blood, which is analyzed by a glucose-sensing strip.^[22] Strip technology typically relies on an enzymatically catalyzed reaction of glucose plus oxygen converting to gluconic acid and hydrogen peroxide in the presence of glucose oxidase.^[23] More advanced glucose-sensing technologies take advantage of glucose transport through the skin (in reverse iontophoresis^[24]) or subcutaneous glucose sensors (such as the continuous glucose monitoring system^[25]). These automated systems aim to reduce patient discomfort in glucose measurement (versus the finger-prick) while providing significantly more information regarding glucose dynamics during treatment. Where as transcutaneous measurements are taken four to six times daily, these automated devices record glucose concentrations as rapidly as every 10 minutes. Furthermore, measurements are continued during the nighttime hours, such that dangerous overnight hypoglycemia can be diagnosed and treated.

The ability to deliver insulin keeps a diabetic patient alive. Since its discovery, insulin has been delivered via syringe in a variety of preparations;^[26,27] those preparations currently employed for diabetes treatment have already been detailed. For subcutaneous delivery via injection, the needle gauge (diameter) contributes to the level of discomfort and healing time for an injection. There is a continual push for smaller needles, in terms of both width (pain) and length (stability of the needle), as well as prefilled delivery systems (such

as the insulin pen), as patient comfort and convenience are as important here as in glucose sensing.^[28] As an alternative to injection therapy, insulin pumps can be employed.^[5] Continuous infusion pumps are either implanted or worn externally, with the latter having a transcutaneous line for subcutaneous delivery.^[29] These pumps may also be primed to deliver a pulse or profile of insulin when commanded by the patient. In this way, it is possible for extracorporeal pump users to forgo the insulin injection normally taken with a meal in favor of additional insulin delivery from the pump. Ideally, the continuous pump system could be integrated with the glucose sensor in such a way that insulin delivery would automatically be adjusted to the glucose dynamics observed in the patient. This is a mechanical replacement for the natural pancreas, known as the artificial pancreas. A closer look at this insulin replacement technique concludes this article.

NOVEL INSULIN DELIVERY APPROACHES

Physiological Insulin Replacement

A key reason for the development of fast-acting and long-acting insulins, as well as the evaluation of various delivery routes, was the number of issues associated with tissue-based insulin replacement.^[30–32] First among these is the need for transplanted or implanted tissue to be histocompatible with the recipient. Histoincompatible tissue is attacked by the recipient's immune system unless immunosuppression is employed; while beneficial in controlling blood glucose, this approach compromises the immune system of the patient, so it is rarely employed unless the pancreas is cotransplanted with another organ (such as the kidney).^[33] Allogeneic grafts of β -cells have proven successful in the absence of immunosuppression,^[34] although this approach uses a genetically modified cell line that requires long-term antibiotic treatment to keep the growth of the grafted islets under control. Given the difficulties encountered when employing the above approaches, the bulk of research on tissue-oriented insulin delivery has focused on the implantation of either islets of Langerhans or just the insulin-producing β -cells in a manner that partitions foreign tissue from that of the host.

Microencapsulation using a cross-linked polymer gel is one approach proposed for the development of a bioartificial pancreas. These gels may contain islets^[35] or cells that will mature into islets.^[36] The success of this style of approach is dependent upon a collection of factors.^[37] The cells within the gel must remain viable; sufficient oxygen and nutrients must reach the islets to maintain cellular function. Furthermore, the cells must be able to sense their extracellular environment, and

that environment must be representative of the physiological state. Hence, glucose transport rates through the gel must be fast enough that the encapsulated islets can respond to the fluctuations in physiological glucose levels. The diffusion characteristics of the gel must allow insulin produced by the islets to escape into the surrounding tissue and hence the circulation. Finally, the gel coating must protect the encapsulated cells from the immune system of the host. For islet microencapsulation, Weir and coworkers used a high-mannuronic acid alginate cross-linked with BaCl_2 .^[35,36] This membrane provided superior performance in islet protection, and hence glucose control, through its durable mechanical properties and nonreactivity with the host immune system.

Another islet-protection approach is macroencapsulation. This approach uses larger capsules (millimeter size and larger) to hold islets rather than coating small numbers of islets as in microencapsulation. This makes the implanted device(s) easier to remove or to reload with islets. The diffusion limitations of these larger systems are more severe than in microcapsules,^[31] leading to a potential trade-off between islet viability and immunoblocking capability. This drawback can be overcome, to a certain degree, by reducing the size of the macrocapsule.^[38] Given the diffusion characteristics, the physiological location of the implant will play a more dominant role. For devices with reduced diffusion rate, increased vascularization and blood flow is necessary for both device performance and cell survival. Hollow fiber macrocapsules containing islets were implanted in the intraperitoneal cavity of diabetic mice.^[39] Blood glucose was returned to normal, the islets remained viable, and insulin release from the macrocapsule had the characteristic biphasic profile in response to a glucose challenge.

Polymeric Delivery Systems

Polymeric delivery systems have been extensively examined for the purposes of providing long-term reproducible insulin delivery to diabetic patients.^[40] Polymer gels that respond to pH changes are made glucose-sensitive through immobilization of glucose oxidase to the polymer network. In the presence of this enzyme, glucose reacts with oxygen to produce gluconic acid and hydrogen peroxide. The gluconic acid produced alters the pH of the device microenvironment. The pH change leads to an increase in membrane porosity and release of insulin from the polymer device. Obviously, insulin induces physiological glucose removal, causing a decrease in the blood glucose concentration, and a slowing of the aforementioned glucose reaction within and around the gel. This leads to a decrease in porosity and a decrease in insulin release.

In this way, the membrane is a combined glucose sensor and feedback control system, sensing the glucose concentration and adjusting membrane porosity to yield an appropriate insulin flux.

Reservoir systems represent another formulation of implantable insulin delivery devices. As described previously, an environmentally sensitive polymer membrane separates the bloodstream from an insulin reservoir. Horbett and coauthors performed experimental^[41,42] studies on potential membranes for these systems, demonstrating insulin release in the physiological glucose range (50–400 mg/dl glucose) across macroporous and pressurized homogeneous polymer membranes. A theoretical study^[43] examined the oxygen limitations to glucose sensitivity in membrane devices. An alternate reservoir device design was presented by Siegel and Firestone.^[44] A pH-sensitive polymer membrane would respond to glucose in the bloodstream through the same glucose-to-gluconic acid reaction. The expansion of the membrane in response to high glucose concentrations would exert pressure on an intermediate fluid, thereby forcing insulin release through an orifice opposite the “sensor” membrane. Decreased glucose concentrations would induce a deswelling of the polymer membrane and a corresponding cessation of insulin release. Although the study examined direct pH effects (rather than glucose-imparted pH changes), the swelling and deswelling behavior was shown to be repeatable, such that periodic insulin release in response to meal consumption was feasible.

Intelligent (“smart”) polymer systems provide an alternate approach to the pump or reservoir systems. These devices represent an “all-in-one” philosophy, where the environmentally sensitive polymer would also contain insulin, thereby acting as a unified sensor, controller, and reservoir. Goldraich and Kost^[45] examined cross-linked polymethacrylate polymer disks for insulin delivery. Although these systems responded to glucose, the response was not entirely within the physiological range, and insulin had difficulty escaping the polymer disks. Peppas and coauthors^[46,47] utilized different monomers to construct cross-linked disks and microparticles for glucose-induced insulin delivery. Insulin was introduced to these devices through imbibition, and catalase was incorporated to improve oxygen permeability and eliminate hydrogen peroxide. Response to glucose was at low concentrations, but insulin release was significantly greater than in the earlier disk system.

The Artificial Pancreas

Insulin delivery, as it is currently practiced, requires direct patient interaction in the roles of sensor (taking

glucose measurements), controller (determining the proper insulin dose to deliver), and actuator (delivering the insulin). Hence, the quality-of-disease treatment is directly impacted by patient decisions including the timing and rate of glucose measurement and insulin injections, and these decisions may include quality-of-life concerns such as pain, risk assessment, convenience, and others. Automated insulin delivery could lead to significant improvements in glucose control (reduced HbA_{1c}, fewer hypoglycemic episodes, etc.) by removing the patient from the control loop. A schematic of this approach is shown in Fig. 1. One formulation of a device of this type would be implantable and would include a glucose sensor, a dosage adjustment algorithm (controller), and a variable-rate insulin pump with reservoir. Alternatively, the hardware could be an external device with subcutaneous glucose sensing and insulin delivery. In either case, the pump mechanism is currently available, and the sensor is a current topic of intense academic and industrial interest.^[48,49] Key to the success of the artificial pancreas is the performance of the control algorithm that converts a glucose measurement into a safe and effective insulin pump delivery rate.

Mathematical models of the patient can be employed in the development of algorithms for glucose control. This yields a model-based approach to controller design, and model-based control is the current state of the art. However, it is important to note that the performance of a model-based controller is directly related to the quality of the model, such that model quality limits theoretically achievable controller performance.^[50] As a result, there is a significant focus on the control-relevance of models for drug delivery.^[51] The objective is to employ a mathematical model of minimum complexity to represent the behaviors of the system of interest, such as the diabetic patient, as accurately as possible.

Modeling has been a focus of glucose-insulin interaction analysis since the 1960s. A variety of models for

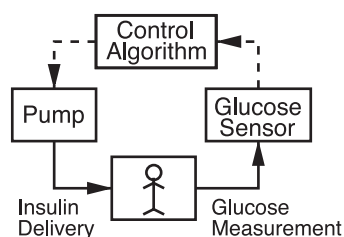


Fig. 1 Schematic of an automated closed-loop insulin delivery system. A glucose sensor collects a measurement of patient glucose concentration, which is sent to a control algorithm. This controller transmits a desired insulin delivery rate to the insulin pump, which delivers the corresponding amount of insulin to the patient.

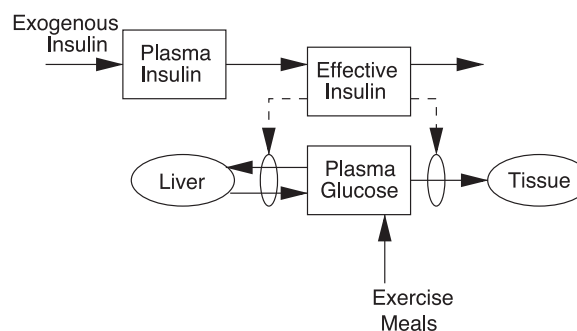


Fig. 2 Schematic representation of Bergman's minimal model.

representing the human glucose-insulin system have been suggested, ranging from the very simple and aphysiological (two linear differential equations^[60]) to the physiologically motivated and complex (tens of coupled nonlinear differential equations). Schematic representations of two control-relevant models are shown in Figs. 2 and 3. Figure 2 depicts the Bergman minimal model,^[52] which was designed to capture glucose and insulin dynamics using a minimum number of equations (three) and only six parameters that need to be fit from data. This empirical model is adequate for representing composite behavior under certain conditions, but it offers little insight into physiological mechanisms that impact insulin delivery needs. Figure 3 shows a physiologically motivated compartmental model representing the pharmacokinetics and pharmacodynamics of glucose, insulin, and glucagon (a counterregulatory hormone).^[53,54] While physiological models are far more time-consuming to construct, specific metabolic effects are easily incorporated within the model structure, as each differential equation corresponds to a physiologically measurable quantity.

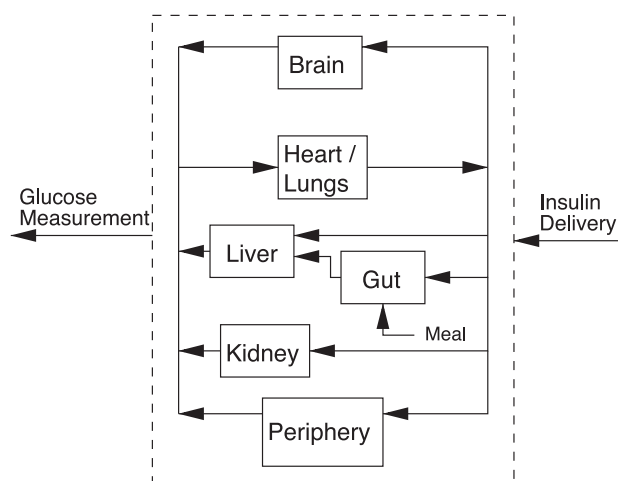


Fig. 3 Schematic representation of a physiologically based pharmacokinetic/pharmacodynamic model.

Furthermore, a controller designed for a patient using a relevant physiologically motivated model would have better performance than one designed for the same patient using the minimal model.

$$\min_{\Delta U(k|k)} \{ \|\Gamma_y(R(k+1|k) - Y(k+1|k))\|_2^2 + \|\Gamma_u \Delta U(k|k)\|_2^2 \} \quad (1)$$

Model predictive control is a model-based control algorithm that has been successfully applied to biomedical problems, such as diabetes,^[55] blood pressure control,^[56] and depth of anesthesia.^[57] As measurements are collected, the model predictive control algorithm solves the preceding optimization problem online. The first term penalizes for deviations of model-based glucose predictions (Y) from the desired glucose reference (R), and the second term penalizes insulin delivery adjustments (measurement noise effects). The relative importance of glucose control performance and pump wear can be adjusted by changing the weighting matrices, Γ_y and Γ_u , respectively. This routine is employed in a “receding horizon” strategy, where the insulin delivery rate is adjusted each time a glucose measurement is taken. The model accuracy determines the quality of the glucose predictions (Y) based on past insulin delivery rate information, and therefore model quality explicitly impacts the controller performance. This algorithm was employed with a linear version of the model shown in Fig. 3 to counteract a 50 g glucose challenge initiated at time $t = 50$ min by delivering insulin intravenously. The results are plotted in Fig. 4. Significant improvements in glucose control are possible by employing this insulin delivery algorithm, including a reduction in meal-induced hyperglycemia and a more rapid return to normoglycemia. The appearance of a hypoglycemic excursion (around $t = 230$ min) is a function of a modest overdelivery of insulin because of the aggressively tuned controller. The magnitude of this episode is dependent on the quality of the model and may increase as the model is less representative of the patient, but the degree of hypoglycemia can be systematically altered by adjusting the controller tuning parameters.

The objective in Eq. 1 is commonly employed in engineering problems and penalizes equally for deviations of glucose above or below the reference. From a medical perspective, however, these deviations are of unequal importance. Hypoglycemia is immediately dangerous, while hyperglycemia is generally associated with the long-term complications of diabetes. Hence, it is more clinically relevant to penalize deviations below the glucose target more heavily than hyperglycemic excursions. Furthermore, the dynamics of insulin action previously discussed are not instantaneous. Many medical metrics of glycemic control relate to glucose concentrations at a point time-distant to meal

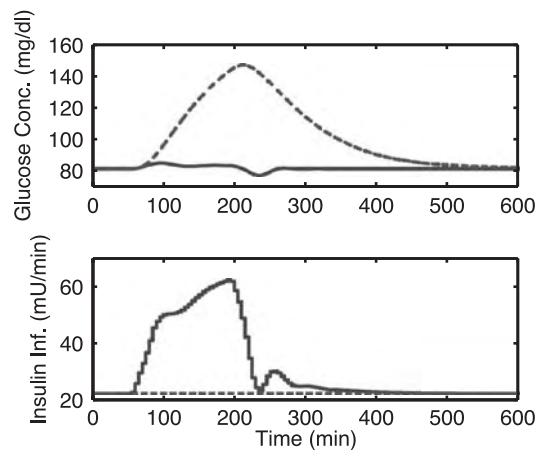


Fig. 4 A comparison of closed-loop glucose control with variable-rate insulin delivery (solid) versus basal insulin delivery (dashed). The 50 g glucose challenge took place at time $t = 50$ min, and insulin delivery was to the intravenous space.

consumption (2 hours after, for example). Finally, glucose control objectives are generally prioritized. An ordered example set might include 1) avoid overt hypoglycemia; 2) keep glucose below an upper threshold; and 3) maintain glucose as near to a reference value as possible. The types of objectives discussed here force a reformulation of the original control problem and employ more complicated optimization methods to solve.^[58,59] The use of a more clinically relevant formulation requires a more powerful computer for online solution, but the more powerful control algorithm can reduce the incidence of hypoglycemia without sacrificing meal compensation performance.^[58]

ARTICLES OF FURTHER INTEREST

Controlled Release; Drug Delivery, Controlled; Glucose Sensors

REFERENCES

1. Kennedy, F.P. Recent developments in insulin delivery techniques, current status and future potential. *Drugs* **1991**, *42*, 213–227.
2. Samols, E.; Ryder, J.A. Studies on tissue uptake of insulin in man using a differential immunoassay for endogenous and exogenous insulin. *J. Clin. Invest.* **1961**, *40*, 2092–2102.
3. Heinemann, L. Variability of insulin absorption and insulin action. *Diabetes Technol. Ther.* **2002**, *4*, 673–682.
4. American Diabetes Association. Continuous subcutaneous insulin infusion (position statement). *Diabetes Care* **2001**, *24* (Suppl. 1), S98.
5. Lenhard, M.J.; Reeves, G.D. Continuous subcutaneous insulin infusion: A comprehensive review of insulin

- pump therapy. *Arch. Intern. Med.* **2001**, *161*, 2293–2300.
6. Still, J.G. Development of oral insulin: Progress and current status. *Diabetes/Metab. Res. Rev.* **2002**, *18* (Suppl. 1), S29–S37.
 7. Lowman, A.M.; Morishita, M.; Kajita, M.; Nagai, T.; Peppas, N.A. Oral delivery of insulin using pH-responsive complexation gels. *J. Pharm. Sci.* **1999**, *88*, 933–937.
 8. Lee, Y.-C.; Simamora, P.; Pinsuwan, S.; Yalkowsky, S.H. Review on the systemic delivery of insulin via the ocular route. *Int. J. Pharm.* **2002**, *233*, 1–18.
 9. Modi, P.; Mihic, M.; Lewin, A. The evolving role of oral insulin in the treatment of diabetes using a novel RapidMist™ system. *Diabetes/Metab. Res. Rev.* **2002**, *18* (Suppl. 1), S38–S42.
 10. Patton, J.S. Deep-lung delivery of therapeutic proteins. *Chemtech* **1997**, *27*, 34–38.
 11. Bindra, S.; Rosenstock, J.; Cefalu, W.T. Inhaled insulin: A novel route for insulin delivery. *Expert Opin. Investig. Drugs* **2002**, *11*, 687–691.
 12. Klonoff, D.C. Inhaled insulin. *Diabetes Technol. Ther.* **1999**, *1*, 307–313.
 13. Kost, J. Ultrasound-assisted insulin delivery and non-invasive glucose sensing. *Diabetes Technol. Ther.* **2002**, *4*, 489–497.
 14. Scheuplein, R.J.; Blank, I.H. Permeability of the skin. *Physiol. Rev.* **1971**, *51*, 702–747.
 15. Prausnitz, M.R. Overcoming skin's barrier: The search for effective and user-friendly drug delivery. *Diabetes Technol. Ther.* **2001**, *3*, 233–236.
 16. Dailey, G.E.; Boden, G.H.; Creech, R.H.; Johnson, D.G.; Gleason, R.E.; Kennedy, F.P.; Weinrauch, L.A.; Weir, M.; D'Elia, J.A. Effects of pulsatile intravenous insulin therapy on the progression of diabetic nephropathy. *Metabolism* **2000**, *49*, 1491–1495.
 17. DCCT—The Diabetes Control and Complications Trial Research Group. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. *N. Engl. J. Med.* **1993**, *329*, 977–986.
 18. DCCT—The Diabetes Control and Complications Trial Research Group. The absence of a glycemic threshold for the development of long-term complications: The perspective of the diabetes control and complications trial. *Diabetes* **1996**, *45*, 1289–1298.
 19. Pfeifer, M.A.; Halter, J.B.; Porte, D., Jr. Insulin secretion in diabetes mellitus. *Am. J. Med.* **1981**, *70*, 579–588.
 20. Gerich, J.E. Novel insulins: Expanding options in diabetes mellitus. *Am. J. Med.* **2002**, *113*, 308–316.
 21. Bolli, G.B. Physiological insulin replacement in type 1 diabetes mellitus. *Exp. Clin. Endocrinol. Diabetes* **2001**, *109* (Suppl. 2), S317–S332.
 22. Bayer Corporation Glucometer® dex diabetes care system **2003**. <http://www.bayerdiag.com/products/pages/self/glucometerdex.html>.
 23. Roche Diagnostics Accu-chek® test strips **2003**. <http://www.accu-chek.com/products/main.cfm?pid=3000>.
 24. Potts, R.O.; Tamada, J.A.; Tierney, M.J. Glucose monitoring by reverse iontophoresis. *Diabetes/Metab. Res. Rev.* **2002**, *1*, S49–S53.
 25. Steil, G.M.; Rebrin, K.; Mastrototaro, J.; Bernaba, B.; Saad, M. Determination of plasma glucose during rapid glucose excursions with a subcutaneous glucose sensor. *Diabetes Technol. Ther.* **2003**, *5*, 27–31.
 26. Eli Lilly and Company Products **2003**. <http://www.lilly.com/products.html>.
 27. Novo Nordisk Insulin **2003**. <http://www.novonordisk-us.com/view.asp?ID=1177>.
 28. Becton, Dickinson, and Company BD pharmaceutical systems **2003**. <http://www.bd.com/pharmaceuticals/>.
 29. Medtronic MiniMed Which pump is for you? **2003**. http://www.minimed.com/patientfam/pf_ipt_paradigm_whichpump.shtml.
 30. Lacy, P. Treating diabetes with transplanted cells. *Sci. Am.* **1995**, *273*, 50–58.
 31. Weir, G.C.; Bonner-Weir, S. Scientific and political impediments to successful islet transplantation. *Diabetes* **1997**, *46*, 1247–1256.
 32. Siebers, U.; Horcher, A.; Bretzel, R.G.; Federlin, K.; Zekorn, T. Alginate-based microcapsules for immunoprotected islet transplantation. *Ann. N.Y. Acad. Sci.* **1997**, *831*, 304–312.
 33. Reach, G. Islet transplantation: A field on the move. *Nephrol. Dial. Transplant.* **2001**, *16*, 893–896.
 34. Pericin, M.; Althage, A.; Freigang, S.; Hengartner, H.; Rolland, E.; Dupraz, P.; Thorens, B.; Aebischer, P.; Zinkernagel, R.M. Allogenic β -islet cells correct diabetes and resist immune rejection. *Proc. Natl. Acad. Sci.* **2002**, *99*, 8203–8206.
 35. Duvivier-Kali, V.F.; Omer, A.; Parent, R.J.; O'Neil, J.J.; Weir, G.C. Complete protection of islets against allojection and autoimmunity by a simple barium-alginate membrane. *Diabetes* **2001**, *50*, 1698–1705.
 36. Omer, A.; Duvivier-Kali, V.F.; Trivedi, N.; Wilmot, K.; Bonner-Weir, S.; Weir, G.C. Survival and maturation of microencapsulated porcine neonatal pancreatic cell clusters transplanted into immunocompetent diabetic mice. *Diabetes* **2003**, *52*, 69–75.
 37. Colton, C.K.; Avgoustiniatos, E.S. Bioengineering in development of the hybrid artificial pancreas. *J. Biomech. Eng.* **1991**, *113*, 152–170.
 38. Trivedi, N.; Keegan, M.; Steil, G.M.; Hollister-Lock, J.; Hasenkamp, W.M.; Colton, C.K.; Bonner-Weir, S.; Weir, G.C. Islets in alginate macrobeads reverse diabetes despite minimal acute insulin secretory responses. *Transplantation* **2001**, *71*, 203–211.
 39. Hegre, O.D.; Lacy, P.E.; Dionne, K.E.; Gentile, F.; Aebischer, P.; Laurance, M.; Fiore, D.; Gardiner, A.; Hazlett, T.; Purzycki, M. Macroencapsulation of islets of Langerhans. *Diabet. Nutr. Metab.* **1992**, *5* (Suppl. 1), 159–162.
 40. Albin, G.; Horbett, T.; Ratner, B. Glucose-Sensitive Membranes for Controlled Release of Insulin. In *Pulsed and Self-Regulated Drug Delivery*; CRC Press, 1990; 160–185.
 41. Albin, G.; Horbett, T.; Ratner, B. Glucose sensitive membranes for controlled delivery of insulin: Insulin transport studies. *J. Control. Release* **1985**, *2*, 153–164.
 42. Cartier, S.; Horbett, T.; Ratner, B. Glucose-sensitive membrane coated porous filters for control of hydraulic

- permeability and insulin delivery from a pressurized reservoir. *J. Membr. Sci.* **1995**, *106*, 17–24.
43. Klumb, L.; Horbett, T. Design of insulin delivery devices based on glucose sensitive membranes. *J. Control. Release* **1992**, *18*, 59–80.
 44. Siegel, R.; Firestone, B. Mechanochemical approaches to self-regulating insulin pump design. *J. Control. Release* **1990**, *11*, 181–192.
 45. Goldraich, M.; Kost, J. Glucose-sensitive polymeric matrices for controlled drug delivery. *Clin. Mater.* **1993**, *13*, 135–142.
 46. Lowman, A.; Peppas, N. Complexation graft copolymers as oral drug delivery systems. *Polym. Prepr.* **1997**, *38*, 566–567.
 47. Hassan, C.; Doyle III, F.J.; Peppas, N. Dynamic behavior of glucose-responsive poly(methacrylic acid-glycol) hydrogels. *Macromolecules* **1997**, *30*, 6166–6173.
 48. Kerner, W. Implantable glucose sensors: Present status and future developments. *Exp. Clin. Endocrinol. Diabetes* **2001**, *109* (Suppl. 2), S341–S346.
 49. Maran, A.; Cerpaldi, C.; Tiengo, A.; Grassi, G.; Vitali, E.; Pagaon, G.; Bistoni, S.; Calabrese, G.; Santeusano, F.; Leonetti, F.; Ribaud, M.; DiMario, U.; Annuzzi, G.; Genovese, S.; Riccardi, G.; Prefiti, M.; Cucinotta, D.; Giorgino, F.; Bellomo, A.; Biorgino, R.; Poscia, A.; Varalli, M. Continuous subcutaneous glucose monitoring in diabetic patients: A multicenter analysis. *Diabetes Care* **2002**, *25*, 347–352.
 50. Morari, M.; Zafiriou, E. *Robust Process Control*; Prentice-Hall: Englewood Cliffs, NJ, 1989.
 51. Parker, R.S.; Doyle III, F.J. Control-relevant modeling in drug delivery. *Adv. Drug Deliv. Rev.* **2001**, *48*, 211–228.
 52. Bergman, R.; Phillips, L.; Cobelli, C. Physiologic evaluation of factors controlling glucose tolerance in man. *J. Clin. Invest.* **1981**, *68*, 1456–1467.
 53. Parker, R.S.; Ward, J.H.; Peppas, N.A.; Doyle III, F.J. Robust H-infinity glucose control in diabetes using a physiological model. *AIChE J.* **2000**, *46*, 2537–2549.
 54. Sorensen, J. *A Physiologic Model of Glucose Metabolism in Man and Its Use to Design and Assess Improved Insulin Therapies for Diabetes*; Ph.D. Thesis-Department of Chemical Engineering MIT, 1985.
 55. Parker, R.S.; Doyle III, F.J.; Peppas, N.A. A model-based algorithm for blood glucose control in type I diabetic patients. *IEEE Trans. Biomed. Eng.* **1999**, *46* (2), 148–157.
 56. Yu, C.L.; Roy, R.J.; Kaufman, H.; Bequette, B.W. Multiple-model adaptive predictive control of mean arterial blood pressure. *IEEE Trans. Biomed. Eng.* **1992**, *39*, 765–778.
 57. Linkens, D.A.; Mahfouf, M. Generalised Predictive Control (GPC) in Clinical Anaesthesia. In *Advances in Model-Based Predictive Control*; Clarke, D., Ed.; Oxford University Press, 1994; 429–445.
 58. Parker, R.S.; Gatzke, E.P.; Doyle III, F.J. Advanced Model Predictive Control [MPC] for Type I Diabetic Patient Blood Glucose Control. In *Proc. American Control Conf.*; , 2000; Vol. 5, 3483–3487.
 59. Tyler, M.L.; Morari, M. Propositional logic in control and monitoring problems. *Automatica* **1999**, *35*, 565–582.
 60. Bolie, V.W. Coefficients of normal blood glucose regulation. *J. Appl. Physiol.* **1961**, *16*, 783–788.

Integrins

David G. Simpson

Department of Anatomy and Neurobiology, Virginia Commonwealth University,
Richmond, Virginia, U.S.A.

INTRODUCTION

The integrins are a family of extracellular matrix (ECM) receptors that bind to a broad variety of ligands. The prototypic binding event mediated by these receptors establishes a continuum that physically anchors the extracellular environment to the intracellular environment. These binding events allow a cell to detect and respond to specific extracellular constituents, essentially allowing the cell to “evaluate” the composition of the surrounding microenvironment. The molecular architecture of any one integrin provides at least two potential routes for the transduction of information across the plasma membrane. First, integrin ligation to a target molecule establishes a physical continuum that mediates the propagation of mechanical perturbations across the plasma membrane. Second, the accessory signal proteins that accumulate in association with the cytoplasmic domains of the integrins provide a site for the initiation of intracellular signal cascades. Classically, signal transduction is considered as a response of a cell to events that originate in the extracellular environment. However, mechanical and biochemical signals that originate within the intracellular space can be transduced by the integrins into events that impact the local microenvironment that surrounds the cell. The identity and profile of integrins present on the cell surface may be modulated in response to developmental events and in the progression of pathophysiological disease states. A shift in integrin identity under these conditions may be a reflection of an adaptive process that allows cells to detect and respond to new environmental conditions.

OVERVIEW OF INTEGRIN STRUCTURE/FUNCTION

The functional integrin receptor is a bidirectional signaling complex that is composed of two transmembrane subunits, an *alpha* (α) and a *beta* (β) chain.^[1,2] The α (120–180 kDa in mass) and β (90–110 kDa) chains are transcribed from separate gene sequences and delivered to the cell surface as an intact heterodimeric receptor. Approximately 18 different α chains and eight distinct β chains exist, forming at least

26 separate, and functional, receptors.^[1] Not all of these receptor isotypes exist in humans. Numerous splice variants exist, further increasing the molecular and structural variability that can be observed in the functional receptor.^[3] Alternative splicing events can occur as a function of tissue type,^[4] the state of tissue differentiation, and in parallel with the onset and progression of disease. The full significance of alternative splicing events remains to be determined; however, in embryonic stem cells, a shift from the β_{1A} to the β_{1D} variant is associated with changes in adhesion, migratory activity, and the state of differentiation.^[5] Specific splice variants also have been correlated with the degree of malignancy in some tumors.^[6] Post-translational modifications, in the guise of tyrosine phosphorylation to the cytoplasmic tail of the β chain, also exist and serve to modulate integrin function during platelet aggregation.^[7]

Integrin Fine Structure

The α and β chains of the integrin receptor each exhibit an intracellular, a transmembrane, and an extracellular domain (Fig. 1). The intracellular domains of the prototypic integrin project a short distance into the cytosol, usually less than 50 amino acids. In the cytoplasmic space, the terminal domains of the β chain are indirectly anchored to elements of the cytoskeleton through interactions mediated by a multimolecular complex of accessory proteins. These domains play a central role in the assembly of the heterodimeric receptor complex.^[8] The transmembrane domains of each integrin chain exhibit a helical structure and pass through the lipid bilayer, anchoring the receptor within the plasma membrane. These domains, in concert with the cytoplasmic domains, function in the clustering and assembly of integrins into the mature cell–matrix adhesion structure called the focal adhesion.^[9,10]

The α and β chains both project from the cell surface and into the extracellular space in one of three, gross conformational configurations.^[11–13] These extracellular domains can be further subdivided into a series of subdomains. Analysis of the crystal structure of the $\alpha_V\beta_3$ integrin^[14] reveals a proximal β tail domain located immediately adjacent to the plasma membrane,

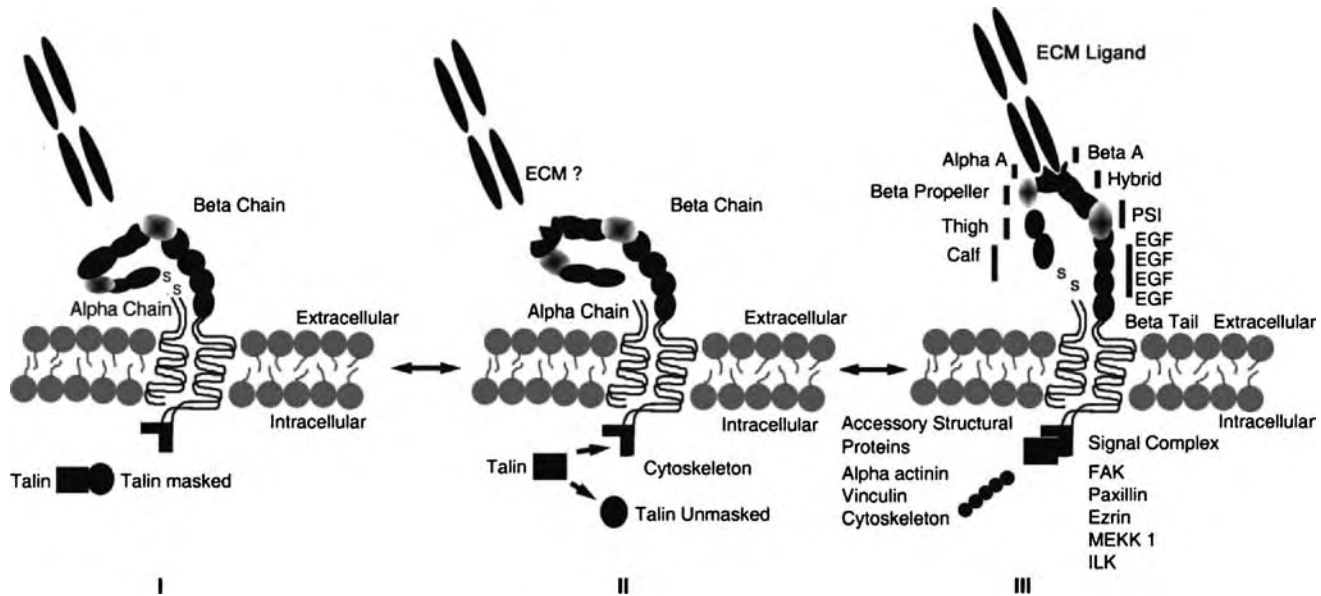


Fig. 1 Integrin structure. The prototypic integrins exist in one of three gross conformational states. This figure depicts each of the three states: (I). folded, closed state; (II) folded, open state; and (III) extended, open state. Talin plays a central role in promoting integrin activation and the transition to high-affinity binding. In the masked state, talin is sequestered from binding to integrins. With unmasking, talin binds to the integrin and promotes activation. Integrin sub domains and an inventory of structural and cell-signaling molecules associated with this receptor are provided in the final panel of the illustration.

where the subunit emerges from the lipid bilayer. This domain is followed by a series of epidermal growth factor (EGF) repeats, a plexin semaphorin integrin (PSI) domain, a hybrid domain, and a distal β A domain. In the α chain, there is a proximal calf domain, a thigh domain, a β propeller, and a distal α A-domain (summarized in Fig. 1). The α and β chains of the functional receptor are typically depicted as a pair of elongated stalks projecting into the extracellular environment. However, the entire receptor can fold into a compact shape that brings the distal ligand-binding domain into close proximity with the plasma membrane. Folding occurs in the transition between the EGF/PSI domains of the β chain, and between the calf/thigh domains of the α chain. The domains subject to folding have been referred to as the knee of the receptor.^[11] There is evidence that the folded conformation functions to limit ligand accessibility; disruption of the folded conformation with antibodies extends the receptor and promotes ligand recognition/binding.^[15]

The head of the prototypic receptor is formed by the distal β A domain of the β chain and the α A-domain of the α chain. These structural features combine to form a cleft that functions as the ligand-binding site. The identity of the α and β chains present in the receptor dictate the conformation of this domain, and serve to define the range of binding interactions that can be mediated by the receptor. Specific amino acid sequences within each head domain and a metal ion-dependent adhesion site (MIDAS) further modulate

the structure and function of the ligand-binding site.^[11] Given the wide variety of potential sites where integrin structure can be modulated, it is not surprising that ligand-binding events are subject to a complex pattern of regulation.

Regulation of Integrin–Ligand Binding

It is useful to conceptualize integrin–ligand binding events as a lock and key interaction. However, this simplified model understates the true complexity of the events that control ligand binding. Integrins may exhibit unique, redundant, and overlapping binding activity to different ECM substrates. For example, $\alpha_7\beta_1$ integrin binds to laminin; both $\alpha_1\beta_1$ and $\alpha_2\beta_1$ integrin bind to collagen and laminin; the $\alpha_5\beta_1$ and $\alpha_v\beta_1$ bind fibronectin and not collagen or laminin, and the $\alpha_3\beta_1$ binds to collagen, laminin, and fibronectin.^[1] Binding affinities, a reflection of binding strength, can be described for each of these interactions and are unique to each integrin–ligand binding partner.

Intrinsic Factors that Regulate Integrin-Binding Events

Identity of the α and β chains

Integrin specificity and ligand binding are regulated, first and foremost by the identity of the α and β chains

that are present within the receptor.^[1,3,16] The primary sequence and tertiary architecture of the α and β subunits serve to define the range of structural and conformational features that are available to the distal β A and α A domains of the receptor. The α A domain encodes the bulk of the information necessary to determine ligand specificity (Table 1). Integrin–ligand binding events, once established, are stabilized by electrostatic (ionic) and hydrophobicity-based interactions.^[11] The molecular, crystallographic, and

functional evidence describing the ligand-binding event compliment one another to a high degree.

Activation state of the integrin receptor

The prototypic integrin complex exists in a low-affinity and a high-affinity state. The actual ligand-binding event is initiated within the intracellular space when talin, a cytoskeletal protein, becomes physically

Table 1 Integrin ligands

Integrin <i>beta</i> chain	Integrin <i>alpha</i> chain	Classical ECM binding partners
<i>Beta</i> ₁		
β_1	α_1	Collagen, ^c laminin
β_1	α_2	Collagen, ^c laminin
β_1	α_3	Laminin, collagen, and fibronectin
β_1	α_4	Lymphocyte trafficking, VCAM-1
β_1	α_5	RGD, classic fibronectin receptor
β_1	α_6	Laminin
β_1	α_7	Laminin
β_1	α_8	RGD peptide
β_1	α_9	Tenascin, non-RGD-dependent binding to fibronectin ^b
β_1	α_{10}	Collagen
β_1	α_{11}	Collagen
<i>Beta</i> ₂		
β_2	α_L	Leukocyte-specific receptor
β_2	α_M	Leukocyte-specific receptor
β_2	α_X	Leukocyte-specific receptor
β_2	α_D	Leukocyte-specific receptor
<i>Beta</i> ₃		
β_3	α_{II_b}	RGD peptide, highly promiscuous ^a
<i>Beta</i> ₄		
β_4	α_4	Plaque proteins, plectin, and others ^c
β_4	α_6	Laminin
<i>Beta</i> ₇		
β_7	α_E	Leukocyte-specific receptor
β_7	α_4	Lymphocyte trafficking, VCAM-1
Other		
$\beta_1, \beta_3, \beta_5, \beta_6, \beta_8$	α_V^d	RGD peptide

^aRGD-dependent binding interactions are often promiscuous in nature. This target sequence is a common motif in many ECM structural proteins. The $\alpha_{II_b}\beta_3$ binds to fibrinogen, vWf, fibronectin, and vitronectin.

^bThe α_9 has several unusual binding/functional properties. In myogenic lines, the binding of ADAM12, a disintegrin metalloproteinase, to the $\alpha_9\beta_1$ integrin mediates cell-to-cell adhesion, leading to the onset of myogenic fusion and differentiation. (From Ref. [17].)

^cConcentrated in hemidesmosomes of epithelial tissue, the β_4 exhibits specialized cytoplasmic domains for docking proteins and cytoskeletal elements of the intermediate filament network.

^dInitial characterization of integrins composed of the α_V integrin subunit suggested the expression of this subunit was an in vitro artifact. Subsequent accumulation of data demonstrates that this subunit, in combination with the listed β chains, forms physiologically relevant receptors. For example, $\alpha_V\beta_1$ plays a role in myofibroblast differentiation, a process selectively associated with fibrotic encapsulation of implanted materials. (From Ref. [18].)

^eCollagen represents the classical ligand for this receptor; extensive and redundant binding activity with laminin is observed.

associated with the cytoplasmic tail of the β chain.^[19,20] This event initiates a transition from the low to the high-affinity ligand-binding state. This transition, commonly referred to as integrin activation, is correlated with unfolding at the knee, conformational deployment of the stalks, and the extension of the receptor away from the plasma membrane.^[15] Determining how talin induces these rather dramatic changes in conformation remains the subject of intense interest. Integrin activation primes the receptor to recognize extracellular ligands and precedes the high-affinity ligand-binding event. Signals that originate in the extracellular space may serve to initiate and/or facilitate the activation process, providing an alternative path leading to high-affinity binding. The integrin-binding site on talin appears to be masked when the integrin is in the low-affinity state. Partial digestion of talin with calpain *in vitro* unmasks the integrin-binding site and promotes talin–integrin interactions.^[21] Talin-mediated activation may or may not represent a generic control checkpoint for all integrins. The physiological signals that regulate talin–integrin interactions remain to be fully defined.

The observation that integrins can bind to extracellular ligands in the low-affinity conformation complicates any discussion describing the molecular events leading to ligand recognition and binding. A comprehensive working model to rectify this complication with existing models of talin-initiated activation of the receptor complex has not been developed. From a functional standpoint, binding in the low-affinity state may underlie events where low-avidity (i.e., “low strength”) adhesion is necessary, for example, in cell migration.^[22,23] In cells undergoing active translocation across a surface, a subset of integrins exists in the low-affinity state. As an increasing proportion of these receptors undergo the transition to the high-affinity state, there is a cessation of motility and the formation of stable cell–substrate contacts. Leukocytes are attached with low affinity and continuously roll along the endothelium without sticking, in an integrin-dependent fashion. Upon activation, in response to chemokine proadhesive signals, this cell type will selectively arrest and undergo firm attachment to the underlying endothelium.^[24]

Closed and open state considerations

Local changes in the conformation of the binding pocket can also modulate integrin ligation. The α A domain exists in two distinct structural states,^[25] a closed and an open conformation. High-affinity ligand binding is correlated with the open state configuration.^[26] Experimentally locking this domain into the open configuration (a process that may be equivalent

to the integrin activation in response to an extracellular event) in lymphocytes is sufficient to initiate the transition from rolling adhesion to firm attachment.^[27] The exact relationships between integrin activation in response to talin, global changes associated with unfolding at the knees, and the conformational rearrangements that can occur in the ligand-binding pocket, all remain active areas of research.

In summary, structural studies suggest that the integrin receptor exists in three distinct conformational states.^[12,13] In turn, these different conformational states appear to correspond to three distinct ligand-binding states. These conformational states include:

- A folded conformation at the knees in which the distal head domains of the ligand-binding site are adjacent to the plasma membrane. The distal ligand-binding pocket exists in a closed state. This conformation correlates with low-affinity binding.
- An extended unfolded conformation associated with talin-initiated activation with the distal ligand-binding site in a closed state. This represents an intermediate conformation; the binding properties are unclear and may be expected to be intermediate between the fully bent state and fully extended conformation in the open state.
- The extended unfolded conformation with the distal ligand-binding site in the open state. This conformation represents the state associated with high-affinity ligand-binding activity.

Extrinsic Factors that Regulate Integrin-Binding Events

Aspartic acid- and glutamic acid-based target sequences

If integrins represent the receptacle in the lock and key analogy of ligand binding, a key to specificity may be envisioned in the composition, primary sequence, and tertiary structure of the target molecules. Aspartic acid- or glutamic acid-based residues located within flexible loops that project from the surface of ECM target molecules are features common to many integrin ligands. The majority of integrins recognize aspartic acid-based structures; these include the amino acid sequence arginyl–glycyl–aspartic acid (RGD) and closely related tripeptides.^[16,28] A small number of integrins bind to glutamic acid-based sequences through events mediated by the α A domain of the integrin α chain. Considering the narrow range of peptide sequences that function as substrates for the integrins, the degree of specificity that members of this class of receptors actually exhibit for target

ligands is surprising. The native ECM is composed of a very complex mixture of materials, yet cells selectively bind and interact with specific constituents of this network.

Ligand accessibility

The molecular identity of a target ligand defines a surface topology; amino acid sequences surrounding the aspartic acid- and glutamic acid-based target sequences dictate the availability of these substrates for binding. The role of these flanking sequences in the presentation of the integrin-binding peptides should not be underestimated. It is only when the contribution of these ancillary characteristics are taken into consideration that it is possible to use just two, three-peptide sequences to accommodate in a specific manner the bulk of all integrin-binding events. These flanking sequences present the tripeptide target sequences to the external environment, in a topology that can be recognized and undergo integrin ligation.

Synergy sites

In addition to the role that flanking sequences play in regulating integrin–ligand binding events, sites that are more distal to the target peptide sequence also have been implicated in the modulating the binding event. Synergy sites within fibronectin act to modify the affinity of the $\alpha_5\beta_1$ integrin to the target RGD site of this glycoprotein in a strain-dependent fashion. In the unstrained state, the synergy site resides approximately 32 Å away from the RGD tripeptide integrin-binding site.^[29] Mechanical perturbations that separate the spatial distance between the RGD sequence and the synergy site reduce the binding affinity of the $\alpha_5\beta_1$ integrin for the target sequence. In effect, the synergy site functions as a mechanosensory switch.^[30] There is also precedence for the masking of integrin recognition sites in ECM molecules. Proteolytic and/or physical perturbations to the structure of the matrix can uncover these cryptic sites and make them available for binding. For example, partial proteolytic digestion alters the conformation of Type IV collagen and disrupts the $\alpha_1\beta_1$ integrin-binding site and uncovers an $\alpha_v\beta_3$ binding site.^[31] Cryptic sites also have been identified in laminin^[32] and can be expected to be present in a variety of other matrix molecules. In perhaps what might be considered a converse pattern of regulation, actin-based contraction induces changes in fibronectin structure that lead to the fibrilogenesis of this matrix glycoprotein.^[33]

Local cation concentration

Divalent cations can modulate the functional state of integrins that exhibit the αA domain. Under normal circumstances, a single cation occupies the MIDAS site of the central β sheet. The ions Ca^{2+} , Mg^{2+} , and Mn^{2+} all support physiological binding at micromolar concentrations; however, at millimolar concentrations, Ca^{2+} appears to inhibit ligand binding.^[16]

The activation of synergy sites and/or the unmasking of cryptic sites can be expected to modulate the identity, density, and/or avidity of integrin-binding events. In turn, these changes may lead to increased and/or decreased migratory activity, altered rates of proliferation, and/or specific changes in gene expression and protein content. These integrin-mediated events may effectively limit the range of materials that can be effectively used in some tissue-engineering applications. Conversely, with the proper fabrication strategies, it may be possible to exploit integrin-mediated events to selectively induce cell migration, regulate cell distribution, and/or modulate cell phenotype (differentiation).

Postadhesion Events

Once ligand recognition and binding occurs, integrin–matrix interactions appear to exist along a continuum of structures, ranging from the focal complex or focal contact to the focal adhesion. The use of a continuum to describe these structures, rather than referring to specific classes of structures, is exploited in this discussion to emphasize the dynamic nature of the adhesion event. The focal complex (contact) exists at one end of the adhesion continuum.^[34] The architecture of this type of adhesion is relatively “rudimentary” in nature and appears to represent the minimal valiancy of an integrin–ligand binding event. High-resolution microscopic mapping of focal complexes in living cells suggests that these structures represent a cluster of three to four integrins.^[35] A binding event mediated by a single cluster of integrins may not represent a site of firm attachment to the surrounding environment. However, when taken in an aggregate as a population of binding events scattered over the entire surface of the plasma membrane, even this type of structure can serve to anchor a cell in place. A single embryonic fibroblast has an estimated $3\text{--}7 \times 10^5$ integrins on the cell surface.^[36] The aggregate number of adhesive events within any given organ system is staggering when the total number of integrins on each individual cell are considered and the total number of cells are tallied. These cell–matrix interactions may have profound effects on the degree of resting tension that is present within the ECM.

The focal adhesion exists at the other end of the adhesion continuum. Focal contacts can “mature” through the sequential recruitment of other structural and cytoskeletal proteins, to form more elaborate, stable, and long lasting focal adhesions.^[37] Most of the ensuing discussion describes cell–matrix adhesions within the context of tissue culture experiments. However, these focal adhesions are not limited to the cultured cell; structures that exhibit homologous molecular identity, architecture, and function exist *in vivo*. Examples include the dense plaques of smooth muscle, the myotendinous junctions of skeletal muscle, the Z-disks of cardiac muscle, and the hemidesmosomes of the epithelium.^[38–41] It should be noted that the molecular identity^[42] and mechanical properties^[43] of any one focal contact or focal adhesion site vary within a cell. This degree of heterogeneity may represent different stages of maturation and/or functional states that exist along the adhesion continuum.

The formation of focal complexes requires an integrin substrate,^[44] and is closely allied with changes in actin polymerization states.^[45] The structural properties of the actin-based cytoskeleton are regulated by the ras homology protein (RHO), a member of the ras superfamily of small guanosine triphosphatases (GTPases). The principal members of this family, ras-related c3 botulinum toxin substrate (RAC), Cdc42, and RHO each occupy a critical site in the signal transduction pathways that modulate distinct architectural aspects of the microfilament network. For example, RAC induces G-actin to undergo polymerization into F-actin in a process associated with the formation of lamellipodia at the cell periphery.^[46] The Cdc42 protein drives actin assembly in the fine filipodial projections that emanate from the leading edge of lamellipodia. The formation of stress fibers^[47] by bundling small-scale actin fibrils into larger order assemblies^[45] is promoted by RHO. The coordinate regulation of these distinctive structures is essential in the promotion of directed cell motility.

Real-time microscopic examination of focal complex assembly reveals that integrins are carried along in the plane of the plasma membrane by underlying actin fibrils that are undergoing polymerization.^[35] The movement of integrins by this process may facilitate initial integrin–ligand interactions and the subsequent formation of the focal complex, the “smallest unit” of adhesion. Prior to incorporation into a focal complex, clusters of integrins appear to be freely diffusible in the lipid bilayer. The integrins in this freely diffusible compartment are presumably in an unactivated, low-affinity, closed state. Once activation and a nucleation process have been initiated, the integrins localized to a nascent focal complex become substantially less mobile in the plasma membrane.

The molecular events that mark the transformation of the focal complex into a more elaborate adhesion site are ill defined. However, under baseline conditions, nominally RHO-dependent events are necessary to induce this transformation in most cell types.^[48] Unlike the actin polymerization-dependent assembly of focal complexes, the assembly of actin stress fibers does not directly induce the formation of focal adhesions.^[49] Rather, it is the mechanical signals associated with the contractile forces generated by these fibrils that promote the clustering and concentration of integrins into specific domains on the plasma membrane. Strain applied across the plasma membrane is sufficient to drive these processes, even when endogenous contractile events are experimentally suppressed.^[50] Nonspecialized domains of the plasma membrane may exhibit 50–400 integrins/ μm^2 ; by comparison, the fully elaborated focal adhesion may have 700–1200 integrins/ μm^2 .^[35]

As noted, the cytoplasmic and transmembrane domains of the receptor complex facilitate the accumulation and clustering of integrins into higher order adhesion sites.^[9,10] In parallel with this clustering process, a multimolecular complex of accessory structural and intracellular signaling proteins accumulate in association with the cytoplasmic domain of the β chains. Elegant studies using live cell imaging have been used to map the accumulation of these accessory proteins in space and time.^[35,37,43,51]

It is unclear as to how the mechanical perturbations are coupled to the processes that regulate the assembly of focal adhesions. The RAC and RHO-induced changes in the organization of the actin-based cytoskeleton can occur independently of any detectable change in the organization of adhesion sites. This observation argues that RAC and RHO are proximal to the signals that ultimately couple actin dynamics to the assembly of adhesion sites.^[45] There is evidence that microtubules “fine tune” the assembly process. Physical perturbations to the cytoplasmic compartment induce global changes in microtubule dynamics, and the disruption of this cytoskeletal network suppresses RAC and RHO deployment in response to mechanical signals.^[52] Perhaps more importantly, the regional assembly of microtubules modulates the degree of strain present within specific domains of the cell;^[53] providing a potential mechanism for the regulation of adhesion site assembly and structure, through subcellular or local changes in the viscoelasticity of the cytosol. A local increase in the stiffness of the cytosol, induced by microtubule assembly, can be expected to increase the amount of contractile force necessary to initiate changes in integrin distribution and the assembly of focal adhesions. Consistent with this model, the pharmacologic disruption of microtubules leads to an increase in the surface area of

existing focal adhesions.^[54] Conversely, when drug induced microtubule depolymerization is reversed during washout experiments, the focal adhesions' surface area is reduced in parallel with the assembly of the recovering microtubules.^[55] These are intriguing results and demonstrate the fundamental role that mechanical forces play in regulating internal cytoarchitecture.^[40,41,56–59]

The intimate correlation between changes in actin organization and the onset of integrin mediated-adhesion indicates that cytoskeletal links are present at the earliest stages of the adhesion event. These nascent interconnections serve to “prime” the processes that promote the subsequent elaboration of focal complexes into focal adhesions. Molecular modeling predicts that RAC, RHO, actin polymerization states, and focal complex assembly/maturation, all function within an elaborate feedback loop that self-regulates the assembly process.^[60] In aggregate, the dispersed focal contacts appear to function as a population of receptors for mechanical perturbations. When a mechanical perturbation exceeds some threshold of force, the assembly process is favored and tipped toward the formation of stable focal adhesions. The mechanical regulation of integrin clustering and the assembly of focal adhesions allow the cell to adapt and respond to mechanical events that originate in the external microenvironment. Simultaneously, this regulatory program allows the cell to selectively transmit contractile forces to the surrounding environment. A relevant example of this type of regulatory pattern is observed in endothelial cells subjected to shear forces. Changes in shear directly modulate the elaboration of focal adhesion structure,^[60] and insure the elaboration of formation of adhesions sufficient to firmly anchor these cells to the underlying basement membrane in a challenging hemodynamic environment.

The subtle balance of internally and externally derived forces that modulate integrin-mediated adhesion, intracellular signaling, and cell phenotype in native tissues is difficult to recapitulate within a tissue-engineering scaffold. This is one reason that many design criteria for this type of engineered material include specifications that allow for the biodegradation and replacement of implanted materials with endogenous ECM materials.^[61,62]

Integrin Deadhesion

An aspect of integrin binding that is often underappreciated is a consideration of how an established integrin–ligand binding interaction might be terminated. The cyclic regulation of cell adhesion must routinely occur during cell division and active motility, and provides the stationary, adhesive cell the capacity

to become mobile. These transitions also occur during the native wound healing process, and underlie the events that allow a cell to interact with implanted biomaterials and devices. The physiological events that mediate deadhesion are not fully defined, and a comprehensive molecular model to fully account for all of the events that have been correlated with the disruption of integrin-mediated binding is just beginning to emerge. Experimental data suggest that deadhesion is not a simple reversal of the assembly process.

Integrin-mediated binding events may be altered through multiple pathways. Because models of disassembly are drawn from many different cell types (normal vs. transformed vs. genetic knockouts vs. functional knockouts vs. cells from different tissues) and models of adhesion (in vitro migration vs. in vitro static cell adhesion vs. dynamic live vs. non-live-cell imaging) using techniques of varying fidelity, a note of caution is in order. Developing molecular models of deadhesion are further complicated by the observation that the composition of adhesion sites varies as a function of the assembly process (i.e., the degree of maturation along the adhesion continuum). Defining how different intracellular signals interact during integrin deadhesion promises to be an evolving area of research.

Intracellular events in the regulation of deadhesion

Prior to the accumulation of a specific protein entity and/or a critical mass of accessory proteins, the integrity of integrin-dependent adhesion sites is likely to be regulated through a dynamic interaction between RAC- and RHO-sensitive events.^[60] At the focal adhesion end of the adhesive continuum, cell culture experiments have identified several candidate signaling events/proteins that are associated with the loss of adhesion. During the elaboration of focal adhesions, the signal peptide, FAK (focal adhesion kinase), accumulates in association with integrin-positive adhesion sites.^[54] The FAK protein plays a dual role in these structures; first, it functions as a scaffolding protein for the assembly of a signaling complex; second, it functions as a signal peptide. FAK has specific binding activity against a number of accessory proteins that localize to integrins, including talin, paxillin, ezrin, mitogen-activated protein kinase kinase kinase 1, and other target substrates.^[63] Over time, additional accessory proteins including ILK (integrin linked kinase), α -actinin, and vinculin and other cytoskeletal elements accumulate in the multimolecular signaling complex, and FAK begins to undergo phosphorylation on tyrosine residues. Early in the assembly process, the concentration of FAK and the relative amount of FAK phosphorylation increase in parallel with the

accumulation of the other accessory proteins. At some, as of yet to be precisely defined, levels of FAK phosphorylation (and/or other accessory proteins), focal adhesion structure begins to deteriorate.

Relative changes in FAK content at adhesion sites exhibit a high correlation with the assembly/disassembly process. Several observations link the functional state and distribution of FAK, microtubule dynamics, and focal adhesion structure into a coherent model of deadhesion. Cells null in FAK develop extensive focal adhesions and display reduced rates of cell migration.^[64,65] Inducing microtubule assembly in the FAK null cell fails to grossly alter focal adhesion structure.^[55] Introducing functional FAK into this phenotypic background rescues the defect, and leads to the disassembly of focal adhesions and the restoration of migratory activity. Microtubule assembly plays a subtle and critical role in these processes. High-resolution microscopy has revealed that a subset of microtubules terminate within molecular distances of adhesion sites.^[66] These microtubules are associated with dynamin, a GTPase that regulates endocytic events.^[67] In working models, dynamin is delivered to the vicinity of the focal adhesion site by these microtubules, and then, just prior to the onset of deadhesion, dynamin coimmunoprecipitates with FAK.^[55] The adaptor protein, Grb2 (growth factor receptor-binding protein 2), which also functions in endocytic pathways,^[68] also becomes associated with FAK at this critical juncture in the process.^[69]

Subsequent to the interactions of dynamin-FAK-Grb2, cell-surface integrins are observed within an intracellular vesicular compartment.^[35] These internalized receptors may ultimately be delivered to a lysosomal compartment and/or are recycled back to the cell surface. Dynamin participates in concert with Grb2 in the internalization of EGF receptors.^[70] However, the nature of the molecular events that link the formation of the dynamin-FAK-Grb2 complex and the delivery of integrins to the intracellular vesicular compartment remains a black box. Not all integrins are processed to the vesicular compartment during deadhesion. A subset of integrins undergo an increase in mobility and diffuse in the plasma membrane away from adhesion sites.^[35]

In tissue culture, the adhesion sites of actively migrating cells are disrupted by a proteolytic event that severs intracellular links between the integrin complex and the cytoskeleton through a calpain-sensitive pathway.^[23] In this process, integrins remain attached to the substrate as the moving cell translocates forward.^[35,71] Little or no detectable cytoskeletal or associated accessory proteins appear to be associated with the cast off receptors, suggesting that proteolysis occurs very close to the cytoplasmic domains of the integrin complex.^[72] This would seem to be a rather

energetically expensive regulatory strategy, and the physiological significance of this type of proteolytic processing *in vivo* is unknown. There is limited evidence that selective ectodomain shedding of integrins occurs during rapid transitions in growth and onset of pathology in the heart,^[73,74] and in cancer cell metastasis.^[75] In fibroblasts, matrix metalloprotease (MMP) inhibitors stabilize focal adhesion structure, suggesting that proteases may play an active role in integrin ectodomain processing.^[76] Disruption of matrix conformation by MMPs and/or other proteases also undoubtedly plays a role in regulating the deadhesion process within intact tissues by perturbing the structure of existing ECM ligands.^[31] The MMP-1 binds directly to the $\alpha_2\beta_1$ integrin, when keratinocytes migrate on type I collagen.^[77]

There is a constantly growing constellation of signal peptides that interact to regulate cell adhesion events. For example, FAK-related nonkinase (FRNK) is a distinct protein species, composed of the C-terminal domain of FAK. Unlike FAK, FRNK inhibits the assembly of focal adhesions, potentially by competing for FAK-binding sites in the integrin-signaling complex.^[78] In selected tissues, the RHO-related peptides, Rnd (“round”), promote focal adhesion breakdown, the loss of cell-substratum adhesion and cell rounding.^[79] The loss of Src-kinase, a proto-oncogene protein, activity suppresses the activation and localization of calpain to focal adhesion sites, stabilizes FAK against proteolysis, increases focal adhesion area, and suppresses cell motility.^[80,81] The loss of Src theoretically impairs the proteolytic processing and shedding of integrins; the adhesion defects are probably further exacerbated because interactions between Grb2 and FAK are Src dependent.^[82] Blocking RAS, another proto-oncogene protein, impairs focal adhesion breakdown and compromises cell motility.^[83] Consistent with this observation, a high percentage of malignant cancers, which commonly exhibit defects in adhesion-dependent growth, have constitutive RAS activity.

Extracellular events in the regulation of deadhesion

Integrin adhesion also can be inhibited by extracellular events. For example, ADAMs (a disintegrin and metalloprotease)^[84] are a family of transmembrane glycoproteins that function as binding partners for integrins.^[85] The ligation of integrins to an ADAM forms a heterodimeric complex that mediates a cell-to-cell binding event.^[86] Heterotypic binding of integrins and cadherin molecules (cadherins normally mediate cell-to-cell interactions in a homotypic fashion) also have been identified, providing yet another potential (and “unorthodox”) regulatory site for the control of

integrin-binding events.^[87] The formation of heterodimeric complexes can be expected to indirectly alter cell–matrix interactions, intracellular strain patterns, actin dynamics, and intracellular signaling.

From a research perspective, there are peptides that can be used to disrupt integrin-mediated adhesion. The integrin target sequence RGD can be prepared as a cyclic peptide. This synthetic molecule competes for adhesion sites, suppresses integrin clustering, and blocks adhesion. Peptides isolated from the venom of pit vipers, called disintegrins, disrupt and block selected integrin-mediated binding events.^[88] This particular class of antiadhesive agents is a potent inhibitor of platelet aggregation.^[89] Antibodies that selectively block specific integrin subtypes have long been a mainstay of adhesion research. These agents may be designed to inhibit adhesion through blocking specific receptor isotypes or through the targeting of binding moieties present in extracellular ligands. Conversely, antisera can be designed to activate cluster integrins in the extracellular space, and initiate downstream intracellular signal transduction events. Microinjection can be used to deliver antisera to the cytosolic compartment. Various adenoviral^[58] and small siRNAs^[90] also can be used to examine integrin-dependent events. A variety of nonspecific and specific inhibitors directed against intracellular signal peptides also have been developed.

Blocking deadhesion processes does not necessarily interrupt all of the events associated with cell motility. For example, inhibition of RAS pathways suppresses the focal adhesion turnover and the directed cell migration in wound assays of motility, but does not appear to impact the formation of lamellipodia.^[83] This result emphasizes the structural and functional compartmentalization that characterizes actin dynamics (e.g., RAC, RHO, and Cdc42 regulated events) and the continuum of adhesion events. This discussion of integrin-associated regulatory peptides that modulate deadhesion is by no means exhaustive. The specific regulatory events that transpire during deadhesion are likely to vary with the degree of maturation of the adhesion site. Signal transduction events that disrupt integrin-mediated adhesion may occur both in a linear (i.e., unique substrates that must interact in a serial fashion, A to B to C) or parallel pattern of regulation (where events do not have to happen in a specific sequence). Events that are redundant and carried out through alternative pathways most certainly exist within the processes that regulate deadhesion.

Functional Considerations: Integrin-Mediated Signaling

The integrin complex is far more than a simple tether that functions to anchor the cell and provide traction

against the surrounding environment. Integrins represent a site for the active transmission of phenotypic information across the plasma membrane. Classically, cell-signaling events describe processes in which information from the extracellular environment is transmitted into the intracellular space. This process is referred to as “outside-in signaling.” In the context of integrin-based signaling, this type of pathway functions in the detection and transduction of phenotypic information that is stored in the composition and tertiary structure of the ECM.^[91] Integrins also mediate the transmission of information from the intracellular space into the extracellular environment, in a process termed “inside-out signaling.” This provides a pathway for the cell to manipulate physical or chemical characteristics of the local microenvironment.^[1]

Intracellular signal transduction pathways are isolated, synergistic and antagonistic, redundant, and unique. It is possible to map many of the specific signal events initiated by integrins with a high degree of certainty. Intracellular signal transduction pathways may function in a “response”- and/or “tissue”-specific pattern of regulation. A response-specific pattern describes events that occur in the majority of cell types through a common intracellular pathway(s). The potential number of these pathways is multiplied by the observation that intracellular signals may be initiated by the ligation of different integrin receptors (e.g., $\alpha_1\beta_1$ in one cell type and $\alpha_2\beta_1$ in another). Tissue-specific patterns refer to events that occur in a particular cell type, in response to an integrin-generated signal. Because of the enormous variability that response- and tissue-specific events exhibit, it is more productive in this type of monograph to survey some specific signal pathways and examples of the downstream, end effects of integrin-based signaling. Outside the laboratory setting, few intracellular signals occur in a vacuum, a characteristic that is important to recognize in any discussion of this topic.

Outside-in signaling

In the outside-in mode of signaling, the integrin receptor represents an antenna that functions to detect extracellular ligands. Integrin ligation can transmit information to the intracellular space through at least two distinct pathways. First, binding events between an integrin and a target ligand establishes a physical continuum that can be traced from constituents of the ECM, across the plasma membrane at the integrin, through elements of the cytoskeleton and into the nuclear matrix. Physical perturbations propagated across this network can impact nuclear structure, DNA organization, and rates of gene transcription.^[92] The degree of specificity that can be encoded by signals

that are physically propagated to the nucleus is open to question. Integrin ligation and mechanical signals also induce the formation of microcompartments that in turn function to induce the accumulation and concentration of mRNA and ribosomes to the vicinity of focal adhesions,^[93] providing the potential for the site-specific synthesis of proteins.

Integrin ligation can detect, encode, and transduce more specific information through the complex of soluble accessory structural and signaling proteins that accumulate in association with fully elaborated adhesion sites. Many of the intracellular signals induced by the integrins act, in concert, with growth factors to modulate cellular phenotype. At least 20 distinct protein species become associated with the cytoplasmic domains of the integrin receptor. Signals propagated through this multimolecular complex can be divided into five broad categories of downstream events; these include signals that: (I) control cell survival; (II) regulate gene transcription; (III) modulate rates of cell proliferation; (IV) regulate cell motility; and (V) modulate cytoskeletal organization (for an excellent overview, see Ref. [1]). Like nearly all the signaling events, there is considerable overlap in many of these processes.

(I) *Integrin Mediated Control of Cell Survival (i.e., Anchorage-Dependent Survival)*. In many cell types, integrin ligation is associated with the accumulation and activation of ILK. This leads to the signaling through phosphatidylinositol 3 kinase and nuclear factor-kappa B (NF-kappaB), and the phosphorylation of protein kinase B (PKB) and a transcription factor (Akt).^[94] The loss of down stream signals propagated through this particular pathway, and specifically the inhibition of PKB/Akt activity, lead to increased rates of apoptosis in many different cell types, including fibroblasts,^[95] keratinocytes,^[96] endothelial cells^[97] embryonic hippocampal neurons,^[98] and a variety of other normal cell types. The loss of signals through ILK pathways also leads to apoptosis in prostate cancer cells^[99] and other transformed cell types. Conversely, overexpression of ILK results in oncogenic transformation; the potential for using this signal molecule as a therapeutic target in the treatment of cancer has been noted by several researchers.

(II) *Regulation of Gene Transcription*. Several specific examples of genes induced by integrin ligation can be readily cited. For example, ligation of the $\alpha_5\beta_1$, a fibronectin receptor, induces collagenase gene expression through an activating protein 1 and/or polyomavirus enhancer activator 3; a member of the ETS family of transcription factors, c-fos (an oncogenic protein)-dependent pathway(s).^[100] In chondrocytes, $\alpha_5\beta_1$ integrin-specific signals induce the expression of several different proinflammatory cytokines through NF-kappaB.^[101] In NIH 3T3 cells, FAK signals

through c-SRC, Grb2, and Ras to activate the mitogen-activated protein (MAP) kinase and extracellular response kinase 2 (ERK2).^[63] The ERK2, and related family members, phosphorylate cytoplasmic proteins, freely enter the nucleus, and activate targets in the nucleus that regulate gene transcription. Activation of the ERK pathways, a relatively promiscuous signal in nature, modulates the expression of many different gene families associated with proliferation and differentiation.

(III) *Modulation of Cell Proliferation*. The MAP kinase pathways are activated in a nearly ubiquitous pattern in response to integrin ligation. This pathway is activated in response to the autophosphorylation of the transmembrane protein, receptor tyrosine kinase (RTK), a class that includes many growth factor receptors. This event establishes docking sites on the cytoplasmic domains of RTK for Grb2 and son of sevenless protein (homolog of the *Drosophila melanogaster* SOS). In turn, SOS interacts with the Raf-1 (a serine threonine kinase), which phosphorylates MEK1 and MEK2 (also known as mitogen activated protein kinase kinase) leading to phosphorylation of Erk1 and Erk2 (MAP kinase family, extracellular response kinase). The Erk peptides, once activated, can translocate to the nucleus and activate transcription factors that regulate the cell cycle. Integrins potentiate signals through this pathway by increasing RTK autophosphorylation and the efficiency of Raf-1, MEK, and Erk signaling.^[102]

Experiments conducted with NIH 3T3, human embryonic kidney, and umbilical vein endothelial cells indicate that anchorage-dependent regulation of the progression into G1 is mediated by integrin- and growth factor-derived signals through c-Jun N-terminal kinase (JNK).^[103] In this particular regulatory path, FAK initiates events through an Src kinase, P130^{CAS} (Crk-associated substrate), Crk (oncogene protein with homology to Src), and ultimately JNK. In endothelial cells, integrin activation leads to a transient tyrosine phosphorylation of janus kinase 2 (JAK2, a family of cytoplasmic signal molecules) and signal transducer and activator of transcription (STAT5). Subsequent to these events, STAT5 localizes to the nucleus and induces c-fos expression.^[104] It should be noted that c-fos represents a member of the early growth response genes; as such, it is upstream of many of the processes that mediate the onset of cell proliferation.

(IV) *Regulation of Cell Motility*. The interplay that exists between the formation of adhesion sites and rates of cell motility has been discussed to some degree in previous sections (see Ref. [105] for a review of RAC, cdc42, and RHO). A variety of signal molecules that are associated directly and tangentially to RAC and RHO have been characterized. For example, RAC

resides within a complex web of signal events. This peptide can activate p21-activated kinase, which, in turn, targets filamin. Filamin is a multifunctional protein that crosslinks F-actin, a function necessary for the formation of lamellipodial ruffles. The RAC protein can simultaneously inhibit myosin activity, thereby suppressing contractile activity, stress fiber formation, and the assembly focal adhesion signaling complexes, processes that suppress the formation of stable cell-matrix adhesions and allow the cell to sustain motility. Integrin-derived signals can indirectly support cell migration by regulating protease activity, allowing the cell to remodel, penetrate, and migrate through the ECM. As already noted, signals propagated by the $\alpha_5\beta_1$ induce collagenase gene expression.^[100] In immortalized keratinocytes, signals propagated by the $\alpha_3\beta_1$ integrin increase the half-life of MMP-9 gelatinase mRNA.^[106] The $\alpha_2\beta_1$ integrin of normal keratinocytes binds MMP-1, localizing protease activity to the immediate vicinity of the cell.^[77]

(V) *Modulation of Cytoskeletal Organization.* The regulation of the actin-based cytoskeleton by integrin-derived events is intimately associated with the formation of cell-matrix adhesion sites and the process of cell migration as discussed. Direct and indirect connections between the integrins, the actin cytoskeleton, and other cytoskeletal elements also can be defined. In the epithelium, intermediate filaments terminate at the plasma membrane in association with hemidesmosomes. These cell-matrix junctions, essentially an *in vivo* correlate to the focal adhesion of cultured cells, anchor epithelial tissues to the basement membrane. Clusters of $\alpha_6\beta_4$ integrins are concentrated in these domains, linking intermediate filaments to the plasma membrane and forming a continuum that anchors this cytoskeletal network to the external environment.^[107] The cytoplasmic tail of β_4 integrin exhibits specialized cytoplasmic structures that function as docking domains for accessory proteins that link intermediate filaments to the hemidesmosome.^[108] Considerable efforts have been directed at identifying the molecular signature of the proteins necessary to form these elaborate structures. The hemidesmosome is subject to regulation, at least as complex as actin-associated focal adhesions. Intermediate filament networks undergo assembly and disassembly in parallel with the progression of the cell cycle. The disruption of hemidesmosomes in the autoimmune disease, *bullous pemphigoid*, causes epithelial blistering and compromises the intermediate filament network.

Inside-out signaling

Inside-out signaling typically refers to the process of integrin activation. Integrin activation and

ligand-binding events impact the local mechanical environment and may lead to signals that modulate the biochemical composition of the immediate surroundings of the cell. Cardiac fibroblasts migrating on a collagen gel apply tension to this type of extracellular network through integrin-mediated processes.^[109] This cell-derived tension represents a mechanical perturbation that can be propagated to adjacent cells. The potential role of this type of signaling has not been investigated to any great extent. However, in the heart (and potentially other organ systems, i.e., the skin), it is possible that the aggregate effect of the tension applied to the ECM by the resident interstitial cells could have an impact on ventricular wall compliance. We can only speculate how the selective pharmacologic disruption of interstitial cell-matrix interactions might effect cardiac performance in the normal and failing heart. Suppression of tensile forces generated by dermal fibroblasts has been used in attempts to manage scar formation during wound healing. The integrin-mediated regulation of MMP expression and distribution can modulate ECM structure, the degree of tension present in this network, and initiate the release of growth factors and cytokines that are bound to constituents of the ECM.^[31,77] Integrins also indirectly modulate the biochemical composition of the ECM by modulating cytokine expression.^[101]

CONCLUSIONS

Integrins function as a critical interface between the intracellular and extracellular environments. These receptors are multifunctional; in the outside-in mode of signaling, the integrins detect and transduce signals that regulate a variety of cellular processes. Simultaneously, in the inside-out mode of signaling, the integrins can modulate the composition of the local microenvironment. Ideally, the bioengineering paradigm will develop strategies to exploit this unique class of receptors to direct tissue reconstruction, modulate the wound healing process, and promote more complete tissue regeneration during recovery from acute and chronic injuries. It may be possible to develop physiologically relevant tissue-engineering scaffolds that selectively direct cell migration, cell survival, and cell phenotype.

ACKNOWLEDGMENTS

This work is supported, in part, by NIH R01EB003087 (Simpson). My thanks to Todd A. Telemeco, MP, PhD for his comments, and my colleague Garrett N. Simpson for figure preparation.

ARTICLES OF FURTHER INTEREST

Biofunctional Polymers; Biomaterials Immune Response

REFERENCES

- Hynes, R.O. Integrins: bidirectional, allosteric signaling machines. *Cell* **2002**, *110*, 673–687.
- Schoenwaelder, S.M.; Burridge, K. Bidirectional signaling between the cytoskeleton and integrins. *Curr. Opin. Cell Biol.* **1999**, *11* (2), 274–286.
- Ross, R.S.; Borg, T.K. Integrins and the myocardium. *Circ. Res.* **2001**, *88* (11), 1112–1119.
- Belkin, A.M.; Zhidkova, N.I.; Balzac, F.; Altruda, F.; Tomatis, D.; Maier, A.; Tarone, G.; Kotliansky, V.E.; Burridge, K. Beta 1D integrin displaces the beta 1A isoform in striated muscles: localization at junctional structures and signaling potential in nonmuscle cells. *J. Cell Biol.* **1996**, *132* (1–2), 211–226.
- Gimond, C.; Baudoin, C.; Sonnenberg, A. Defects in adhesion and migration, but not in proliferation and differentiation, of embryonic stem cells upon replacement of integrin subunit beta1A by beta1D. *Differentiation* **2000**, *66* (2–3), 93–105.
- Tennenbaum, T.; Belanger, A.J.; Glick, A.B.; Tamura, R.; Quaranta, V.; Yuspa, S.H. A splice variant of alpha6 integrin is associated with malignant conversion in mouse skin tumorigenesis. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92* (15), 7041–7045.
- Jenkins, A.L.; Nannizzi-Alaimo, L.; Silver, D.; Sellers, J.R.; Ginsberg, M.H.; Law, D.A.; Phillips, D.R. Tyrosine phosphorylation of the beta3 cytoplasmic domain mediates integrin-cytoskeletal interactions. *J. Biol. Chem.* **1998**, *273* (22), 13,878–13,885.
- Briesewitz, R.; Kern, A.; Marcantonio, E.E. Assembly and function of integrin receptors is dependent on opposing alpha and beta cytoplasmic domains. *Mol. Biol. Cell* **1995**, *6* (8), 997–1010.
- Hantgan, R.R.; Gibbs, W.; Stahle, M.C.; Aster, R.H.; Peterson, J.A. Integrin clustering mechanisms explored with a soluble alphaIIb beta3 ectodomain construct. *Biochim. Biophys. Acta* **2004**, *1* (suppl 1700 (1)), 19–25.
- Li, R.; Gorelik, R.; Nanda, V.; Law, P.B.; Lear, J.D.; DeGrado, W.F.; Bennett, J.S. Dimerization of the transmembrane domain of integrin α_{IIb} subunit in cell membranes. *J. Biol. Chem.* **2004**, *279* (25), 26,666–26,673.
- Humphries, M.J.; McEwan, P.A.; Barton, S.J.; Buckley, P.A.; Bella, J.; Mould, A.P. Integrin structure: heady advances in ligand binding, but activation still makes the knees wobble. *Trends Biochem. Sci.* **2003**, *28* (6), 313–320.
- Takagi, J.; Petre, B.M.; Walz, T.; Springer, T.A. Global conformational rearrangements in integrin extracellular domains in outside-in and inside-out signaling. *Cell* **2002**, *110* (5), 599–611.
- Takagi, J.; Springer, T.A. Integrin activation and structural rearrangement. *Immunol. Rev.* **2002**, *186*, 141–163.
- Xiong, J.P.; Stehle, T.; Diefenbach, B.; Zhang, R.; Dunker, R.; Scott, D.L.; Joachimiak, A.; Goodman, S.L.; Arnaout, M.A. Crystal structure of the extracellular segment of integrin alphaV beta3. *Science*. **2001**, *294* (5541), 339–345.
- Mould, A.P.; Travis, M.A.; Barton, S.J.; Hamilton, J.A.; Askari, J.A.; Craig, S.E.; Macdonald, V.; Kammerer, R.A.; Buckley, P.A.; Humphries, M.J. Evidence that monoclonal antibodies directed against the integrin beta subunit plexin/semaphorin/integrin domain stimulate function by inducing receptor extension. *J. Biol. Chem.* **2005**, *280* (6), 4238–4246.
- Arnaout, M.A.; Goodman, S.L.; Xiong, J.P. Coming to grips with integrin binding to ligands. *Curr. Opin. Cell Biol.* **2002**, *14* (5), 641–651.
- Lafuste, P.; Sonnet, C.; Chazaud, B.; Dreyfus, P.A.; Gherardi, R.K.; Wewer, U.M.; Authier, F.J. ADAM12 and alpha9 beta1 integrin are instrumental in human myogenic cell differentiation. *Mol. Biol. Cell* **2005**, *16* (2), 861–870.
- Lygoe, K.A.; Norman, J.T.; Marshall, J.F.; Lewis, M.P. alphaV integrins play an important role in myofibroblast differentiation. *Wound Repair Regen.* **2004**, *12* (4), 461–470.
- Tadokoro, S.; Shattil, S.J.; Eto, K.; Tai, V.; Liddington, R.C.; de Pereda, J.M.; Ginsberg, M.H.; Calderwood, D.A. Talin binding to integrin beta tails: a final common step in integrin activation. *Science* **2003**, *302* (5642), 103–106.
- Calderwood, D.A. Integrin activation. *J. Cell Sci.* **2004**, *117* (Pt 5), 657–666.
- Yan, B.; Calderwood, D.A.; Yaspan, B.; Ginsberg, M.H. Calpain cleavage promotes talin binding to the beta3 integrin cytoplasmic domain. *J. Biol. Chem.* **2001**, *276* (30), 28,164–28,170.
- Huttenlocher, A.; Ginsberg, M.H.; Horwitz, A.F. Modulation of cell migration by integrin-mediated cytoskeletal linkages and ligand-binding affinity. *J. Cell Biol.* **1996**, *134* (6), 1551–1562.
- Palecek, S.P.; Huttenlocher, A.; Horwitz, A.F.; Laufner, D.A. Physical and biochemical regulation of integrin release during rear detachment of migrating cells. *J. Cell Sci.* **1998**, *111* (Pt 7), 929–940.
- Feigelson, S.W.; Grabovsky, V.; Winter, E.; Chen, L.L.; Pepinsky, R.B.; Yednock, T.; Yablonski, D.; Lobb, R.; Alon, R. The Src kinase p56(lck) up-regulates VLA-4 integrin affinity. Implications for rapid spontaneous and chemokine-triggered T cell adhesion to VCAM-1 and fibronectin. *J. Biol. Chem.* **2001**, *276* (17), 13,891–13,901.
- Lee, J.O.; Bankston, L.A.; Arnaout, M.A.; Liddington, R.C. . Two conformations of the integrin A-domain (I-domain): a pathway for activation? *Structure* **1995**, *3* (12), 1333–1340.
- Lu, C.; Shimaoka, M.; Zang, Q.; Takagi, J.; Springer, T.A. Locking in alternate conformations of the integrin alphaL beta2 I domain with disulfide bonds reveals functional relationships among integrin domains. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98* (5), 2393–2398.
- Salas, A.; Shimaoka, M.; Chen, S.; Carman, C.V.; Springer, T. Transition from rolling to firm adhesion is

- regulated by the conformation of the I domain of the integrin lymphocyte function-associated antigen-1. *J. Biol. Chem.* **2002**, *277* (52), 50,255–50,262.
28. Ruoslahti, E. The RGD story: a personal account. *Matrix Biol.* **2003**, *22* (6), 459–465.
 29. Krammer, A.; Craig, D.; Thomas, W.E.; Schulten, K.; Vogel, V. A structural model for force regulated integrin binding to fibronectin's RGD-synergy site. *Matrix Biol.* **2002**, *21* (2), 139–147.
 30. Krammer, A.; Lu, H.; Isralewitz, B.; Schulten, K.; Vogel, V. Forced unfolding of the fibronectin type III module reveals a tensile molecular recognition switch. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96* (4), 1351–1356.
 31. Xu, J.; Rodriguez, D.; Petitclerc, E.; Kim, J.J.; Hangai, M.; Moon, Y.S.; Davis, G.E.; Brooks, P.C. Proteolytic exposure of a cryptic site within collagen type IV is required for angiogenesis and tumor growth in vivo. *J. Cell Biol.* **2001**, *154* (5), 1069–1079.
 32. Schulze, B.; Mann, K.; Poschl, E.; Yamada, Y.; Timpl, R. Structural and functional analysis of the globular domain IVa of the laminin alpha 1 chain and its impact on an adjacent RGD site. *Biochem. J.* **1996**, *314* (Pt 3), 847–851.
 33. Zhong, C.; Chrzanowska-Wodnicka, M.; Brown, J.; Shaub, A.; Belkin, A.M.; Burrige, K. Rho-mediated contractility exposes a cryptic site in fibronectin and induces fibronectin matrix assembly. *J. Cell Biol.* **1998**, *141* (2), 539–551.
 34. Geiger, B.; Bershadsky, A. Assembly and mechanosensory function of focal contacts. *Curr. Opin. Cell Biol.* **2001**, *13* (5), 584–592.
 35. Wiseman, P.W.; Brown, C.M.; Webb, D.J.; Hebert, B.; Johnson, N.L.; Squier, J.A.; Ellisman, M.H.; Horwitz, A.F. Spatial mapping of integrin interactions and dynamics during cell migration by image correlation microscopy. *J. Cell Sci.* **2004**, *117* (Pt 23), 5521–5534.
 36. Neff, N.T.; Lowrey, C.; Decker, C.; Tovar, A.; Damsky, C.; Buck, C.; Horwitz, A.F. A monoclonal antibody detaches embryonic skeletal muscle from extracellular matrices. *J. Cell Biol.* **1982**, *95*, 654–666.
 37. Zimmerman, B.; Volberg, T.; Geiger, B. Early molecular events in the assembly of the focal adhesion-stress fiber complex during fibroblast spreading. *Cell Motil. Cytoskeleton* **2004**, *58* (3), 143–159.
 38. Terracio, L.; Simpson, D.G.; Hilenski, L.; Carver, W.; Decker, R.S.; Vinson, N.; Borg, T.K. Distribution of vinculin in the Z-disk of striated muscle: analysis by laser scanning confocal microscopy. *J. Cell Physiol.* **1990**, *145* (1), 78–87.
 39. Danowski, B.A.; Imanaka-Yoshida, K.; Sanger, J.M.; Sanger, J.W. Costameres are sites of force transmission to the substratum in adult rat cardiomyocytes. *J. Cell Biol.* **1992**, *118* (6), 1411–1420.
 40. Simpson, D.G.; Decker, M.L.; Clark, W.A.; Decker, R.S. Contractile activity and cell-cell contact regulate myofibrillar organization in cultured cardiac myocytes. *J. Cell Biol.* **1993**, *123* (2), 323–336.
 41. Sharp, W.W.; Simpson, D.G.; Borg, T.K.; Samarel, A.M.; Terracio, L. Contractile activity and external mechanical load affect focal adhesion formation by cultured neonatal rat cardiac myocytes. *Am. J. Physiol.* **1996**, *273*, H546–H556.
 42. Zamir, E.; Katz, B.Z.; Aota, S.; Yamada, K.M.; Geiger, B.; Kam, Z. Molecular diversity of cell-matrix adhesions. *J. Cell Sci.* **1999**, *112* (Pt 11), 1655–1669.
 43. Matthews, B.D.; Overby, D.R.; Alenghat, F.J.; Karavitis, J.; Numaguchi, Y.; Allen, P.G.; Ingber, D.E. Mechanical properties of individual focal adhesions probed with a magnetic microneedle. *Biochem. Biophys. Res. Commun.* **2004**, *313* (3), 758–764.
 44. Hotchin, N.A.; Hall, A. The assembly of integrin adhesion complexes requires both extracellular matrix and intracellular rho/rac GTPases. *J. Cell Biol.* **1995**, *131* (6 Pt 2), 1857–1865.
 45. Machesky, L.M.; Hall, A. Role of actin polymerization and adhesion to extracellular matrix in Rac- and Rho-induced cytoskeletal reorganization. *J. Cell Biol.* **1997**, *138* (4), 913–926.
 46. Nobes, C.D.; Hall, A. Rho, rac, and cdc42 GTPases regulate the assembly of multimolecular focal complexes associated with actin stress fibers, lamellipodia, and filopodia. *Cell* **1995**, *81* (1), 53–62.
 47. Ridley, A.J.; Hall, A. The small GTP-binding protein rho regulates the assembly of focal adhesions and actin stress fibers in response to growth factors. *Cell* **1992**, *70* (3), 389–399.
 48. Rottner, K.; Hall, A.; Small, J.V. Interplay between Rac and Rho in the control of substrate contact dynamics. *Curr. Biol.* **1999**, *9* (12), 640–648.
 49. Chrzanowska-Wodnicka, M.; Burrige, K. Rho-stimulated contractility drives the formation of stress fibers and focal adhesions. *J. Cell Biol.* **1996**, *133* (6), 1403–1415.
 50. Riveline, D.; Zamir, E.; Balaban, N.Q.; Schwarz, U.S.; Ishizaki, T.; Narumiya, S.; Kam, Z.; Geiger, B.; Bershadsky, A.D. Focal contacts as mechanosensors: externally applied local mechanical force induces growth of focal contacts by an mDia1-dependent and ROCK-independent mechanism. *J. Cell Biol.* **2001**, *153* (6), 1175–1186.
 51. Zaidel-Bar, R.; Ballestrem, C.; Kam, Z.; Geiger, B. Early molecular events in the assembly of matrix adhesions at the leading edge of migrating cells. *J. Cell Sci.* **2003**, *116* (22), 4605–4613.
 52. Putnam, A.J.; Cunningham, J.J.; Pillemer, B.B.; Mooney, D.J. External mechanical strain regulates membrane targeting of Rho GTPases by controlling microtubule assembly. *Am. J. Cell Physiol.* **2003**, *284* (3), C627–C639.
 53. Kaverina, I.; Krylyshkina, O.; Small, J.V. Microtubule targeting of substrate contacts promotes their relaxation and dissociation. *J. Cell Biol.* **1999**, *146*, 1033–1044.
 54. Kirchner, J.; Kam, Z.; Tzur, G.; Bershadsky, A.D.; Geiger, B. Live-cell monitoring of tyrosine phosphorylation in focal adhesions following microtubule disruption. *J. Cell Sci.* **2003**, *116*, 975–986.
 55. Ezratty, E.J.; Partridge, M.A.; Gundersen, G.C. Microtubule-induced focal adhesion disassembly is

- mediated by dynamin and focal adhesion kinase. *Nat. Cell Biol.* **2005**, *7*, 581–590.
56. Simpson, D.G.; Carver, W.; Borg, T.K.; Terracio, L. The role of mechanical stimulation in the establishment and maintenance of muscle cell differentiation. *Int. Rev. Cyto.* **1993**, *150*, 69–94.
 57. Simpson, D.G.; Sharp, W.; Terracio, L.; Price, R.L.; Borg, T.K.; Samarel, A.M. Mechanical regulation of cardiac protein turnover and myofibrillar structure. *Am. J. Physiol.* **1996**, *270*, C1075–C1087.
 58. Simpson, D.G.; Reaves, T.A.; Shih, D.; Burgess, W.; Borg, T.K.; Terracio, L. Cardiac integrins: the ties that bind. *Cardiovasc. Pathol.* **1997**, *7*, 135–143.
 59. Simpson, D.G.; Majeski, M.; Borg, T.K.; Terracio, L. Regulation of cardiac protein turnover and myofibrillar structure in vitro by specific directions of stretch. *Circ. Res.* **1999**, *85*, 59E–69E.
 60. Civelekoglu-Scholeya, G.; Orrc, W.; Novak, I.; Meistera, J.; Schwartze, M.A.; Mogilner, A. Model of coupled transient changes of Rac, Rho, adhesions and stress fibers alignment in endothelial cells responding to shear stress. *J. Theor. Biol.* **2005**, *232*, 569–585.
 61. Matthews, J.A.; Wnek, G.E.; Simpson, D.G.; Bowlin, G.L. Electrospinning collagen nanofibers. *J. Macromol.* **2002**, *11*, 232–238.
 62. Telemeco, T.A.; Ayres, C.; Bowlin, G.L.; Wnek, G.; Boland, E.D.; Cohen, N.M.; Baumgarten, C.M.; Mathews, J.; Simpson, D.G. Regulation of cellular infiltration into tissue engineering scaffolds composed of submicron diameter fibrils produced by electrospinning. *Acta Biomaterialia* **2005**, *1*, 376–384.
 63. Schlaepfer, D.D.; Jones, K.C.; Hunter, T. Multiple Grb2-mediated integrin-stimulated signaling pathways to ERK2/mitogen-activated protein kinase: summation of both c-Src and focal adhesion kinase initiated tyrosine phosphorylation events. *Mol. Cell Biol.* **1998**, *18* (5), 2571–2585.
 64. Richardson, A.; Malik, R.K.; Hildebrand, J.D.; Parsons, J.T. Inhibition of cell spreading by expression of the C-terminal domain of focal adhesion kinase (FAK) is rescued by co-expression of Src or catalytically inactive FAK: a role for paxillin tyrosine phosphorylation. *Mol. Cell Biol.* **1997**, *17* (12), 6906–6914.
 65. Hsia, D.A.; Mitra, S.K.; Hauck, C.R.; Streblov, D.N.; Nelson, J.A.; Ilic, D.; Huang, S.; Li, E.; Nemerow, G.R.; Leng, J.; Spencer, K.S.; Cheres, D.A.; Schlaepfer, D.D. Differential regulation of cell motility and invasion by FAK. *J. Cell Biol.* **2003**, *160* (5), 753–767.
 66. Krylyshkina, O.; Anderson, K.I.; Kaverina, I.; Upmann, I.; Manstein, D.J.; Small, V.; Toomre, D.K. Nanometer targeting of microtubules to focal adhesions. *J. Cell Biol.* **2003**, *161*, 853–859.
 67. Marks, B.; Stowell, M.H.; Vallis, Y.; Mills, I.G.; Gibson, A.; Hopkins, C.R.; McMahon, H.T. GTPase activity of dynamin and resulting conformation change are essential for endocytosis. *Nature* **2001**, *410* (6825), 231–235.
 68. Sorkin, A.; McClure, M.; Huang, F.; Carter, R. Interaction of EGF receptor and grb2 in living cells visualized by fluorescence resonance energy transfer (FRET) microscopy. *Curr. Biol.* **2000**, *10* (21), 1395–1398.
 69. Schlaepfer, D.D.; Hanks, S.K.; Hunter, T.; van der Geer, P. Integrin-mediated signal transduction linked to Ras pathway by GRB2 binding to focal adhesion kinase. *Nature* **1994**, *372* (6508), 786–791.
 70. Wang, Z.; Moran, M.F. Requirement for the adapter protein GRB2 in EGF receptor endocytosis. *Science* **1996**, *272* (5270), 1935–1939.
 71. Laukaitis, C.M.; Webb, D.J.; Donais, K.; Horwitz, A.F. Differential dynamics of alpha 5 integrin, paxillin, and alpha-actinin during formation and disassembly of adhesions in migrating cells. *J. Cell Biol.* **2001**, *153* (7), 1427–1440.
 72. Palecek, S.P.; Schmidt, C.E.; Lauffenburger, D.A.; Horwitz, A.F. Integrin dynamics on the tail region of migrating fibroblasts. *J. Cell Sci.* **1996**, *109* (Pt 5), 941–952.
 73. Ding, B.; Price, R.L.; Goldsmith, E.C.; Borg, T.K.; Yan, X.; Douglas, P.S.; Weinberg, E.O.; Bartunek, J.; Thielen, T.; Didenko, V.V.; Lorell, B.H. Left ventricular hypertrophy in ascending aortic stenosis mice: anoikis and the progression to early failure. *Circulation* **2000**, *101* (24), 2854–2862.
 74. Goldsmith, E.C.; Carver, W.; McFadden, A.; Goldsmith, J.G.; Price, R.L.; Sussman, M.; Lorell, B.H.; Cooper, G.; Borg, T.K. Integrin shedding as a mechanism of cellular adaptation during cardiac growth. *Am. J. Physiol. Heart Circ. Physiol.* **2003**, *284* (6), H2227–H2234.
 75. Demetriou, M.C.; Pennington, M.E.; Nagle, R.B.; Cress, A.E. Extracellular alpha 6 integrin cleavage by urokinase-type plasminogen activator in human prostate cancer. *Exp. Cell Res.* **2004**, *294* (2), 550–558.
 76. Ho, A.T.; Voura, E.B.; Soloway, P.D.; Watson, K.L.; Khokha, R. MMP inhibitors augment fibroblast adhesion through stabilization of focal adhesion contacts and up-regulation of cadherin function. *J. Biol. Chem.* **2001**, *276* (43), 40,215–40,224.
 77. Dumin, J.A.; Dickeson, S.K.; Stricker, T.P.; Bhattacharyya-Pakrasi, M.; Roby, J.D.; Santoro, S.A.; Parks, W.C. Pro-collagenase-1 (matrix metalloproteinase-1) binds the alpha(2)beta(1) integrin upon release from keratinocytes migrating on type I collagen. *J. Biol. Chem.* **2001**, *276* (31), 29,368–29,374.
 78. Richardson, A.; Parsons, T. A mechanism for regulation of the adhesion-associated proteintyrosine kinase pp125FAK. *Nature* **1996**, *380* (6574), 538–540.
 79. Nobes, C.D.; Lauritzen, I.; Mattei, M.G.; Paris, S.; Hall, A.; Chardin, P. A new member of the Rho family, Rnd1, promotes disassembly of actin filament structures and loss of cell adhesion. *J. Cell Biol.* **1998**, *141*, 187–197.
 80. Cary, L.A.; Klinghoffer, R.A.; Sachsenmaier, C.; Cooper, J.A. SRC catalytic but not scaffolding function is needed for integrin-regulated tyrosine phosphorylation, cell migration, and cell spreading. *Mol. Cell Biol.* **2002**, *22* (8), 2427–2440.
 81. Westhoff, M.A.; Serrels, B.; Fincham, V.J.; Frame, M.C.; Carragher, N.O. SRC-mediated phosphorylation of focal adhesion kinase couples actin and

- adhesion dynamics to survival signaling. *Mol. Cell Biol.* **2004**, *24* (18), 8113–8133.
82. Schlaepfer, D.D.; Hunter, T. Evidence for in vivo phosphorylation of the Grb2 SH2-domain binding site on focal adhesion kinase by Src-family protein-tyrosine kinases. *Mol. Cell Biol.* **1996**, *16* (12), 7182–7184.
 83. Nobes, C.D.; Hall, A. Rho GTPases control polarity, protrusion, and adhesion during cell movement. *J. Cell Biol.* **1999**, *144* (6), 1235–1244.
 84. Wolfsberg, T.G.; Straight, P.D.; Gerena, R.L.; Huovila, A.P.; Primakoff, P.; Myles, D.G.; White, J.M. ADAM, a widely distributed and developmentally regulated gene family encoding membrane proteins with a disintegrin and metalloprotease domain. *Dev. Biol.* **1995**, *169* (1), 378–383.
 85. Schlondorff, J.; Blobel, C.P. Metalloprotease-disintegrins: modular proteins capable of promoting cell-cell interactions and triggering signals by protein-ectodomain shedding. *J. Cell Sci.* **1999**, *112*, 3603–3617.
 86. White, J.M. ADAMs: modulators of cell-cell and cell-matrix interactions. *Curr. Opin. Cell Biol.* **2003**, *15*, 598–606.
 87. Higgins, J.M.; Mandlebrot, D.A.; Shaw, S.K.; Russell, G.J.; Murphy, E.A.; Chen, Y.T.; Nelson, W.J.; Parker, C.M.; Brenner, M.B. Direct and regulated interaction of integrin alphaE beta7 with E-cadherin. *J. Cell Biol.* **1998**, *140* (1), 197–210.
 88. Scarborough, R.M.; Rose, J.W.; Naughton, M.A.; Phillips, D.R.; Nannizzi, L.; Arfsten, A.; Campbell, A.M.; Charo, I.F. Characterization of the integrin specificities of disintegrins isolated from American pit viper venoms. *J. Biol. Chem.* **1993**, *268* (2), 1058–1065.
 89. Musial, J.; Niewiarowski, S.; Rucinski, B.; Stewart, G.J.; Cook, J.J.; Williams, J.A.; Edmunds, L.H., Jr. Inhibition of platelet adhesion to surfaces of extracorporeal circuits by disintegrins. RGD-containing peptides from viper venoms. *Circulation* **1990**, *82* (1), 261–273.
 90. Tilghman, R.W.; Slack-Davis, J.K.; Sergina, N.; Martin, K.H.; Iwanicki, M.; Hershey, E.D.; Beggs, H.E.; Reichardt, L.F.; Parsons, J.T. Focal adhesion kinase is required for the spatial organization of the leading edge in migrating cells. *J. Cell Sci.* **2005**, *118* (Pt 12), 2613–23.
 91. Simpson, D.G.; Terracio, L.; Terracio, M.; Price, R.L.; Turner, D.C.; Borg, T.K. Modulation of cardiac myocyte phenotype in vitro by the composition and orientation of the extracellular matrix. *J. Cell Physiol.* **1994**, *161*, 89–105.
 92. Maniotis, A.J.; Chen, C.S.; Ingber, D.E. Demonstration of mechanical connections between integrins, cytoskeletal filaments, and nucleoplasm that stabilize nuclear structure. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94* (3), 849–854.
 93. Chicurel, M.E.; Singer, R.H.; Meyer, C.J.; Ingber, D.E. Integrin binding and mechanical tension induce movement of mRNA and ribosomes to focal adhesions. *Nature* **1998**, *392* (6677), 730–733.
 94. Troussard, A.A.; Mawji, N.M.; Ong, C.; Mui, A.; St-Arnaud, R.; Dedhar, S. Conditional knock-out of integrin-linked kinase demonstrates an essential role in protein kinase B/Akt activation. *J. Biol. Chem.* **2003**, *278* (25), 22,374–22,378.
 95. Nho, S.; Xia, H.; Kahm, J.; Kleidon, J.; Diebold, D.; Henke, C.A. Role of integrin-linked kinase in regulating phosphorylation of Akt and fibroblast survival in type I collagen matrices through a beta 1 integrin viability signaling pathway. *J. Biol. Chem.* **2005**, *280* (28), 26,630–26,639.
 96. Wang, X.Q.; Sun, P.; Paller, A.S. Inhibition of integrin-linked kinase/protein kinase B/Akt signaling: mechanism for ganglioside-induced apoptosis. *J. Biol. Chem.* **2001**, *276* (48), 44,504–44,511.
 97. Scatena, M.; Almeida, M.; Chaisson, M.L.; Fausto, N.; Nicosia, R.F.; Giachelli, C.M. NF-kappaB mediates alphavbeta3 integrin-induced endothelial cell survival. *J. Cell Biol.* **1998**, *141* (4), 1083–1093.
 98. Gary, D.S.; Mattson, M.P. Integrin signaling via the PI3-kinase-Akt pathway increases neuronal resistance to glutamate-induced apoptosis. *J. Neurochem.* **2001**, *76* (5), 1485–1496.
 99. Persad, S.; Attwell, S.; Gray, V.; Delcommenne, M.; Troussard, A.; Sanghera, J.; Dedhar, S. Inhibition of integrin-linked kinase (ILK) suppresses activation of protein kinase B/Akt and induces cell cycle arrest and apoptosis of PTEN-mutant prostate cancer cells. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97* (7), 3207–3212.
 100. Tremble, P.; Damsky, C.H.; Werb, Z. Components of the nuclear signaling cascade that regulate collagenase gene expression in response to integrin-derived signals. *J. Cell Biol.* **1995**, *129* (6), 1707–1720.
 101. Pulai, J.I.; Chen, H.; Im, H.J.; Kumar, S.; Hanning, C.; Hegde, P.S.; Loeser, R.F. NF-kappa B mediates the stimulation of cytokine and chemokine expression by human articular chondrocytes in response to fibronectin fragments. *J. Immunol.* **2005**, *174* (9), 5781–5788.
 102. Juliano, R.L.; Reddig, P.; Alahari, S.; Edin, M.; Howe, A.; Aplin, A. Integrin regulation of cell signaling and motility. *Biochem. Soc. Trans.* **2004**, *32* (Pt 3), 443–446.
 103. Oktay, M.; Wary, K.K.; Dans, M.; Birge, R.B.; Giancotti, F.G. Integrin-mediated activation of focal adhesion kinase is required for signaling to Jun NH2-terminal kinase and progression through the G1 phase of the cell cycle. *J. Cell Biol.* **1999**, *145* (7), 1461–1469.
 104. Brizzi, M.F.; Defilippi, P.; Rosso, A.; Venturino, M.; Garbarino, G.; Miyajima, A.; Silengo, L.; Tarone, G.; Pegoraro, L. Integrin-mediated adhesion of endothelial cells induces JAK2 and STAT5A activation: role in the control of c-fos gene expression. *Mol. Biol. Cell* **1999**, *10* (10), 3463–3471.
 105. Burrridge, K.; Wennerberg, K. Rho and Rac take center stage. *Cell* **2004**, *116*, 167–179.
 106. Iyer, V.; Pumiglia, K.; DiPersio, C.M. Alpha3 beta1 integrin regulates MMP-9 mRNA stability in immortalized derationocytes: a novel mechanism of integrin-mediated MMP gene expression. *J. Cell Sci.* **2005**, *118*, 1185–1195.

107. Green, K.J.; Jones, J.C. Desmosomes and hemidesmosomes: structure and function of molecular components. *FASEB J.* **1996**, *10* (8), 871–881.
108. Spinardi, L.; Ren, Y.L.; Sanders, R.; Giancotti, F.G. The beta 4 subunit cytoplasmic domain mediates the interaction of alpha6 beta4 integrin with the cytoskeleton of hemidesmosomes. *Mol. Biol. Cell* **1993**, *4* (9), 871–884.
109. Carver, W.; Molano, I.; Reaves, T.A.; Borg, T.K.; Terracio, L. Role of the alpha 1 beta 1 integrin complex in collagen gel contraction in vitro by fibroblasts. *J. Cell Physiol.* **1995**, *165* (2), 425–437.



Intellectual Property Management

Michael Blakeney

Director, Queen Mary Intellectual Property Research Institute, University of London,
London, U.K.

INTELLECTUAL PROPERTY AND RESEARCH

Intellectual Property is a broad term for the various rights, which the law gives for the protection of economic investment in creative effort. These rights may be conferred by statute or as a consequence of the development of legal principles in the courts. Intellectual property legislation may be obliged as a result of international agreement, such as the World Trade Organization Agreement on Trade-Related aspects of Intellectual Property Rights (TRIPS) or, in the case of U.S.A., as a consequence of its membership of the North America Free Trade Association (NAFTA). The principal categories of intellectual property, which are going to be relevant to research institutes are patents, industrial designs, plant variety rights, trademarks, confidential information, copyright, and database rights.

Patents

A patent is a statutory privilege granted by a government to an inventor and to other persons deriving their rights from the inventor, for a fixed period of years, to exclude other persons from manufacturing, using, or selling a patented product or from using a patented method or process. Patent rights are conferred by statute as a matter of right to the person who is entitled to apply for it and who fulfils the prescribed registration requirements. The protection secured by the registration of a patent is usually limited in time. For example, under the TRIPS Agreement, which sets the basic minimum standard for intellectual property protection, the term of patent protection is 20 yr. In some countries, there may be opportunities for extensions of protection for particular categories of invention such as for pharmaceutical processes. At the end of the period of protection, the patented invention is said to be within the public domain that is available for anyone to exploit. As will be discussed below, this means that patent information in the public domain is a valuable source of data for research institutes.

An invention is usually defined as an idea, which permits the solution of a specific problem in a field of technology. The applicant for the protection of an invention is usually the inventor or his/her successor in

title. For an invention to be protected by a patent it must (i) be new; (ii) involve an inventive step; (iii) be industrially applicable; and (iv) not be a category of excluded invention.

Invention

A distinction is often drawn between inventions and discoveries, as the latter are not considered patentable. Scientific theories without a practical application are not patentable either because as discoveries they are not inventions, or because in their restriction to the province of theory, they cannot easily be characterized as methods of manufacture.^[1] Thus a computer program comprising no more than a mathematical algorithm is not patentable,^[2] whereas the application of an algorithm for some practical purpose, such as for the production of images in computer display graphics, has been characterized as a method of manufacture.^[3] Similarly, a computer program, which permitted a word processor to assemble text in Chinese characters, was held to be patentable.^[4] It is possible for legislation to decree that something considered as a discovery is deemed to be an invention. For example, the European Parliament in its Biotechnology Directive provides that biological material, which is isolated from its natural environment or produced by means of a technical process, is deemed to be an invention even if this material previously occurred in nature. The U.S. Supreme Court took a similar position in *Diamond vs. Chakrabarty*^[5] in ruling that a bacterium genetically engineered to degrade crude oil was an invention.

Novelty

An invention is considered to be new “if it does not form part of the state of the art.” The state of the art is taken to comprise all matter (whether a product, a process, information about either, or anything else), which has been disclosed to the public by prior publication in a tangible form, by oral disclosure, or by use in any way prior to the filing of the patent application. Disclosure may also have occurred through scholarly publication, through a public address or lecture, being placed on the Internet, or through the use of a product

or process embodying the invention. Given that the right to patent an invention can be lost through premature disclosure, it may be incumbent upon an agricultural research institute to institute a system for the management of the publication activities of its researchers.

Inventive step

An invention is said to involve an “inventive step” if, having regard to the prior art, it would not have been obvious to a person having an ordinary skill in the art. In other words, the invention must involve a creative advance on existing knowledge. Researchers from public research institutes will often be requested to provide expert testimony in patents cases as to the current state of the technological art.

Industrial applicability

The requirement that an invention be industrially applicable, or useful, excludes from patent protection purely theoretical inventions, which cannot be carried out in practice. The notion of an applicability, which is “industrial,” connotes a commercial scale of application.

Exclusions from patentability

The TRIPS Agreement envisages the exclusion from patentability of (i) inventions the publication or exploitation of which could broadly be considered to encourage immoral behavior or (ii) “any variety of animal or plant or any essentially biological process for the production of animal or plants, not being a micro-biological process or the product of such process.”

The registration procedure

To obtain a patent, an application is filed with a Patent Office. The application will contain, among other things, a description of the invention, with any drawings referred to in the description and the claims made for the invention. The description must disclose the invention in a manner sufficiently clear and complete for it to be carried out by a person skilled in the art. The disclosure of the invention has to present the invention in the context of the state of the art. To be patentable, the invention must offer a novel solution to a technical problem and the description has to relate the invention to the background art. The function of the claim is to define the scope of the protection, which is sought. A patent application is examined by the Patents Office to ensure that the application meets the

formal registration requirements. The application may then proceed to examination as to substance. For example, the registration authority may institute a search of the patent documents of other nations and of significant technical journals and other publications to ensure that an applicant’s invention has not been previously disclosed. Some countries permit the registration of patents for inventions, which have only been partially disclosed in prior art. Some countries confine relevant prior art to national disclosure, or to prior use and prior oral disclosure.

The application may be published or laid open for public inspection before a patent is granted. An opportunity may be given for third parties to oppose the grant of protection. After the examination of the application as to form and to substance and after the consideration of any opposition, the registration authority will decide on whether to grant a patent. The fact of the granting of the patent will be published in an official gazette.

Patents are national rights. The Patent Cooperation Treaty provides for some procedural efficiencies, for example in having a single application designating a number of countries and having a single state-of-the-art search report available for each national office. In some countries, patent protection is not available for all inventions. For reasons of national interest, some countries withhold patent protection from inventions pertaining to agriculture, food, medical and pharmaceutical products, and nuclear and computer technology. Because patent protection extends only to inventions of a technological nature, protection is withheld from advances relating to methods of business, including financial and accounting techniques, as well as from medical treatments, plant varieties, and animal breeds.

Most countries have compulsory licensing provisions whereby if a patent is not “worked” within a stipulated period of time, the patent holder may be obliged to license the patent right to another person in return for compensation. Compulsory licensing has become a particularly controversial subject. Granting authorities, on one hand, are concerned to prevent patents being used as a means of preventing market entry. Patentees, on the other hand, seek to extend the categories of justification for non-working.

Patent documents as a source of research information

Patent documents are an easily accessible source of research information because they have a fairly uniform structure throughout the world. Patent documents contain both bibliographic and technical information and a search report.

Bibliographic information is presented on the first page of a patent document and includes the following:

1. Dates, names, and addresses of the persons involved in the patent, such as the inventor, the owner of the patent right, and the representative of the owner, or patent agent, as well as particulars of the issuing authority.
2. Classification symbols of the International Patent Classification (IPC), and in some cases, also the national patent classification.
3. Title of the invention, abstract of the description, and a representative drawing or a chemical formula.

Each of the bibliographic data items on the first page of a patent document is identified by a two-digit numeric code, which is universally adopted and facilitates the understanding of the names, dates, addresses, and classification symbols, even without any knowledge of the language in which the patent documentation is published.

The technical information contained in a patent document usually includes:

1. A short description of the state of the art of technology as known to the inventor.
2. The detailed description of the invention in such a manner that a technician skilled in the art is able to work the invention.
3. Drawings or chemical formulae illustrating the invention.
4. The claims, which define the scope of the invention.

The length of a patent document will vary according to the field of technology, but they usually average between 10 and 15 pages.

The search report is established by the patent examiner in the Patent Office. It will contain references to all patent documents, which describe similar or identical technical solutions as the purported invention.

A number of advantages have been identified with the use of patent documents as a source of technological information:

1. Current patent documents contain the latest information. Because of the loss of novelty through publication in academic journals, patent documents will often be the first place where an invention is disclosed, e.g., the Hollerith punch card, Baird television, and Whittle jet engine were disclosed in patent documents a number of years prior to their discussion in the academic literature.

2. Patent documents contain information, which would not otherwise be disclosed, e.g., studies in U.K. and U.S.A. indicate that less than 20% of technical solutions published in patent documents were later published in other scientific journals.
3. Patent documents have a uniform structure, using classification symbols, which simplifies the access to information.
4. Patent documents identify the commercial application of inventions.
5. Patent documents indicate the name and address of the applicant, the patentee, and the inventor. This assists the identification of the source of licensing opportunities.

Typically, adopting the formalities prescribed under the Patents Cooperation Treaty (PCT), an application will contain a request, a description, claims, and an abstract. Occasionally, drawings may be required. The request will identify the applicant and will seek protection for a named invention. The description will contain the disclosure of the invention, identifying the technical field to which the invention relates, a description of the state of the art, together with a bibliography, and the disclosure of the invention with sufficient clarity and completeness to enable it to be understood by a person having ordinary skill in the art. The description will also identify the novelty in the invention by reference to the preexisting art. Finally, the description will indicate the way in which the invention will be industrially applied. The claims, which are made for an invention will define the scope of protection, which is offered. A tension will exist between the width of the claims, which are made by an applicant and the limitation to those claims, which may be considered to be supported by the description and also having account of matters, which are considered already to be part of the state of the art or which are considered to be obvious. Failure to satisfy the disclosure requirements may result from insufficiency of the description or from ambiguity in the claims. Disclosure must be sufficient to permit an informed reader with reasonable skill in the trade to perform the invention. Read with the knowledge of that reader, claims must clearly and succinctly delineate the invention. Finally, the claims must be fairly based on the material revealed in the specification.

Industrial Designs

An industrial design is the ornamental or esthetic aspect of a useful article. Industrial designs are protected through registration. As with patents, most countries require novelty. The standard of novelty varies between universal and national novelty. A difficult issue

in designs protection is the extent to which a design must differ from an earlier design to be considered novel. Minor variations are usually inadequate. A desirable test is whether the design claimed is subjectively new in the sense that it is not an imitation of designs already known to the creator. The critical feature of industrial applicability is that the design is repeatable in commercial quantities. Thus items of artistic craftsmanship are outside the scope of design protection and are more properly protectable under copyright laws.

Industrial designs are protected against unauthorized copying or imitation for 15 yr, dated from the end of the calendar year in which the design was first recorded in a design document or an article was first made to the design, whichever occurred first.

Trademarks

A trademark is usually defined as a sign capable of being represented graphically, which is capable of distinguishing goods or services of one undertaking from those of other undertakings. Originally trademarks were protected for use in relation to goods, but in recent years, the marks used in relation to services have been embraced by this style of protection.

To be registered as a trademark, a sign must be represented in a visible form. In some countries, audible or olfactory signs may be registered if they can be recorded in an appropriate manner. Visible signs typically include names, invented or existing words, letters, numbers, pictures and symbols, or combinations of these signs. Under most trademark laws, the shape of goods or their containers, or the design of their labels or fabrics can be registered as marks. The marks, which are excepted from registration, are those that are not distinctive, or deceptively similar to existing marks and marks, which violate public order or morality.

Most trademark laws allow separate registrations for a mark in respect of each of the 42 categories of goods and services laid down in the International Classification of Goods and Services, which was established in accordance with the Nice Agreement of 1957 and its subsequent revisions. Registration may be permitted with the disclaimer of some elements of a mark. For example, in a word mark, there may be disclaimer of those words, which would be common to the relevant trade.

The application process usually requires an examination by the granting office to ensure compliance with the formal registration requirements, as well as with the substantive requirement of distinctiveness. There also has to be a check as to whether a mark is in conflict with prior rights. After the publication of an

application, most countries provide for an opposition process whereby an interested third party may protest the registration of a mark, usually on the grounds of prior rights or deceptive similarity with another mark.

Upon acceptance of a mark, registration is conferred for a term of between 10 and 20 yr, with a possibility for renewal. A mark will expire if a renewal is not sought. Expungement of a mark may also be sought where its use becomes deceptive or where the mark becomes generic of goods or services. For example, the marks "Vaseline" and "gramophone" are two examples of marks, which became generic descriptions of the type of goods to which they were appended.

A controversial requirement of some trade mark laws is that registration of a trademark is contingent upon its use or a bona fide intention to use upon or in close association with the classes of goods or services in respect of which it is registered. A similar requirement provides for the removal of the registration after a prescribed period of non-use. Protection without registration may be extended to "well-known marks," i.e., those with a significant reputation in a country. Such marks invariably have a substantial international reputation through advertising and use.

Registration of a mark confers protection against emulation by traders using identical or substantially similar marks. Most systems of registration permit assignment or licensure. A system of registered user may be provided to record trademark licenses.

In the event of infringement of a registered mark, a trademark proprietor may seek relief in the form of injunction, compensation orders, and seizure of infringing goods.

The use of a trademark is a warranty of the quality of the goods or services supplied under that mark. The name, acronym, or logo of a research institute is often a warranty of the quality of the services supplied by that institute. Its designation is worthy of protection, particularly because unauthorized traders may falsely represent an affiliation with the institute. Similarly, the products, which are produced by an agricultural research institute may also be worthy of protection, e.g., the IRRI prefix for rice types developed at that institute.

A new form of trademark style protection is registration of an Internet domain name. Imitation of a trademark by an Internet domain name is a trademark infringement and the unauthorized registration of another's name as an Internet domain name is also actionable.

Confidential Information (Including Trade Secrets)

Intellectual property rights, such as patents, trademarks, and industrial designs, which arise as a result of

registration may be contrasted with the rights, which arise under know-how agreements. In the latter case, know-how, which is also described as trade secrets, is protected by the law of confidential information. This is judge-made law (i.e., arising without legislation) in contrast with the various intellectual property rights, which arise as a result of legislation.

The statutory rights, such as patents, trademarks, and copyright are conferred in respect of specific acts of artistic or technological creativity, whereas know-how agreements tend to involve the more general application of technology in a commercial situation. The commercialization of intellectual property through a manufacturing enterprise will typically involve the following: (i) preinvestment studies, including the preparation of a feasibility study and project report; (ii) engineering, including preparation of machinery specifications, plant design, and factory layout; (iii) selection of equipment, plant construction, erection and installation of machinery, and start-up of plant; (iv) acquisition of manufacturing technology; and (v) technical assistance during postinstallation period, including training programs and management assistance.^[6] In other words, intellectual property rights cannot be effectively utilized without the deployment of substantial quantities of know-how. However, unlike for intellectual property rights, there is no system of registration for know-how, and it will be protected only if it is regarded as confidential or if a restriction is placed upon its unauthorized communication.

In the case of employee contracts, most common law systems imply a contractual term obliging employees not to divulge information, which is considered to be the property of an employer. This information will include much of the information relevant to the commercialization of intellectual property listed above such as plant and equipment design, production methods, and customer lists. This contractual term is enforceable only against an employee. To prevent a third party from wrongfully acquiring and using a person's trade secrets, the courts have fashioned rights in equity and tort against third parties. In determining whether information is protectable as a trade secret, the courts have regard to: (i) the extent to which the information is known by employees and by persons outside the relevant business; (ii) the extent of measures made to guard its secrecy; (iii) the value of the information to the business and to competitors; (iv) the amount of effort or money expended in developing the information; and (v) the ease or difficulty with which the information could be properly acquired or duplicated by others.^[7]

The research data and test results assembled by a research institute may form part of the corpus of its confidential information or trade secrets and be protectable under the law.

Biotechnological Rights and Plant Varieties

Biotechnological invention, particularly through the practice of genetic engineering, has become increasingly important for agriculture and for the treatment of disease. Historically, the question of the patentability of "animate" substances proceeded down a separate legal channel to that concerning the patentability of plant varieties. Originally, it was considered that discoveries involving living organisms and material were not inventions for the purposes of most patent statutes. Exceptions to this principle were made for microorganisms used in fermentation and in antibiotics. However, in 1969, the Supreme Court of the Federal Republic of Germany ruled that animal breeding techniques were patentable, provided that the technique was repeatable.^[8] In U.S.A., the courts had consistently rejected claims for the patentability of animate matter until the 1980 decision of the Supreme Court in *Diamond vs. Chakrabarty*.^[9] In that case, the Court ruled that a genetically engineered bacterium capable of breaking down the components of crude oil was patentable. It recognized that the touchstone of patentability was not whether an invention involved living or inanimate subject matter, but whether it involved a man-made invention.

As an alternative to the protection of plant varieties under patent laws, plant varieties may be protected under plant breeder's rights (PBR) laws modeled on the Union for the Protection of New Varieties of Plants (UPOV) Convention International Convention for the Protection of New Varieties of Plants. Generally, under this legislation, the plant breeder is conferred an exclusive right to perform or to license the following acts in relation to propagating material of the variety:

- Produce or reproduce the material.
- Condition the material for the purpose of propagation.
- Offer the material for sale.
- Sell the material.
- Import the material.
- Export the material.
- Stock the material for the purposes described above.

The general duration of PBR under the UPOV Convention is 25 yr in the case of trees and vines and 20 yr for any other variety.

As with patents, PBRs are established after a registration process. A plant variety is considered to be registrable, if it has a breeder, is distinct, uniform, stable, and has not been or has only recently been exploited. A plant variety is considered distinct if it is clearly distinguishable from any other variety whose existence is a matter of common knowledge. It is

uniform if, subject to the variation, which may be expected from the particular features of its propagation, it is uniform in its relevant characteristics on propagation. A plant variety is stable if its relevant characteristics remain unchanged after repeated propagation. A plant variety is taken not to have been exploited if it or propagating material has not been sold to another person by or with the consent of the breeder. The test of no commercial exploitation is easier to satisfy than the test for novelty under patent law, and the choice between the two forms of intellectual property right is a matter to be considered by a research institute.

A similar disclosure of the technology to that required under patent law is required under the Plant Varieties Act 1983. Thus documents, which are published by the various plant variety registries, will provide a valuable source of research data.

Copyright

Copyright law is concerned with the protection and exploitation of the expression of ideas in a tangible form. The central right, which the law confers is to prevent unauthorized persons from copying a work. To be protected as copyright, ideas have to be expressed in an original way; that is they must have their origin in the labor of the creator. Works are protected irrespective of their quality.

Subject matter of protection

Originally, the subject matter of copyright protection was printed in literary artistic and literary works. As reprographic technology has improved, protection has been extended to technical drawings, maps, paintings, three-dimensional works such as sculptures and architectural works, and to photographs and cinematographic works. More recently, copyright protection has been extended to computer programs, Internet web sites, and to databases.

Rights comprised in copyright

The owner of a copyrighted work may exclude others from using it without authorization. The acts, which require the authorization of the copyright owner are usually the following: (i) copying or reproducing the work; (ii) performing the work in public; (iii) making a sound recording of the work; (iv) making a motion picture of the work; (v) broadcasting a work through the electromagnetic spectrum or through cable diffusion; and (vi) translating or adapting the work.

In addition to these rights, certain “moral rights” have been recognized by the Berne Convention for the

Protection of Literary and Artistic Works.^[10] These include the right to claim authorship of a work and the right to object to any distortion, mutilation, or other modification of, or other derogatory action in relation to a work, which would be prejudicial to an author’s honor or reputation. These moral rights usually remain with an author, even after the transfer of the various economic rights mentioned above.

Neighboring rights

Three kinds of rights neighbor upon copyright protection. These are: (i) the rights of performing artists in their performances; (ii) the rights of producers of phonograms; and (iii) the rights of broadcasting organizations in their radio and television programs.

Copyright and research institutes

Obviously protected by copyright law are the publications, which emanate from research institutes. A sometimes-difficult question concerns the ownership of the scholarly articles and papers of researchers within an institute. Strictly speaking, if produced in the course of their employment, the copyright in this scholarship is owned by the research institute. However, as a matter of practice, these ownership rights are conventionally waived, as it is often considered to be antithetical to the research ethos of an institute to insist upon these rights. On the other hand, where this scholarly output forms part of the teaching materials produced by an institute, a clarification of the ownership of this material is counseled. The moral rights of researchers in their scholarly publications will carry across to teaching materials that incorporate those publications and this has to be borne in mind in the way in which editorial work is affected.

Obviously a research institute must monitor the use of the copyright work of others in its own publications and teaching materials. This will include online works, which are incorporated into any online materials or web sites.

Database rights

The European Database Directive, which was implemented in U.K. in the Copyright and Rights in Databases Regulations, 1997, provides for the protection of material contained in databases against unauthorized extraction or reutilization. A “database” is defined as “a collection of independent works, data, or other materials arranged in a systematic or methodical way and individually accessible by electronic or other means.” Much of the data assembled by

teams and researchers in a research institute will fall within this definition. Also protected are “the materials necessary for the operation or consultation of certain databases such as thesaurus and indexation systems.” Thus bibliographic tools maintained by research libraries may fall within this protective regime. The duration of the right to prevent unauthorized extraction is 15 yr. The regulations also provide that if a database is made by an employee in the course of employment, the employer shall be regarded as the maker, subject to any agreement to the contrary. This will be an import issue for research institutes.

In U.S.A., databases are protected, where some effort has been applied for their development.

OWNERSHIP OF INTELLECTUAL PROPERTY BY RESEARCH INSTITUTES

Accountability

Much of the funding received by university research institutes is obtained from public sources such as governments, international organizations, or charitable foundations. Consequently, institutional accountability is needed for the intellectual property emerging from its research; institutes may take the view that their research results are freely available to all. However, an unauthorized person may seek to register a patent arising from that research, thereby preventing its use as a public good. This may have to be preempted by the institute seeking its own intellectual property protection. Patent or plant variety rights, thus obtained, can be licensed to users on a non-exclusive basis.

As owner of the intellectual property, the institute can apply quality assurance procedures to its evaluation, protection, development, and exploitation through administrative procedures dedicated to this task. It can also assist innovators in the commercialization of the intellectual property by concentrating the expertise in an area available to all research teams.

The purpose of intellectual property ownership by a research institute is invariably to promote the fundamental research mission of the institute, to protect its integrity and welfare, to provide a resource for industry, to obtain appropriate return for the use of facilities, resources, and services provided by the institution, to encourage the growth, progress, and success of the research institute through ventures with industry, seeking commercial returns, to provide fair and reasonable reward (and incentive) to staff and students who apply their intellectual activity, to increase the acumen and accountability for management and use of public funds, and to foster

the identity of the research institute and an esprit de corps.

Enhancing Public/Private Cooperation

Much research carried out at university research institutes is conducted with industry collaboration. Industry operates in an environment, which uses intellectual property rights to protect the value of its products and the money expended on R&D. Many corporations will not invest in, or transfer technology to, countries, which do not have an intellectual property infrastructure, so corporations are increasingly beginning to insist that research institutes have an equivalent infrastructure. Where collaborative research is to be conducted, the private collaborator will insist that the fruits of the research, often representing a significant investment by it, are not appropriated by some unauthorized third party. Sometimes the research will involve an elaboration of a piece of intellectual property contributed from the private sector, and the contributor may insist that the research institute has an appropriate intellectual property regime.

Ownership of Intellectual Property

As a general rule, a research institute will own all the intellectual property generated by its staff (and research students). This is, at least, on the basis that the intellectual property has been created in the course of employment or that research institute resources and services have contributed, including, often, preexisting intellectual property belonging, or licensed by the research institute. The question of intellectual property ownership is also clarified through the adoption of an intellectual property statute, particularly one, which provides for the insertion of an intellectual property clause in the contract of employment.

Most institutions deal with the ownership issue in their contracts of employment or staff policies.

MANAGING THE INTELLECTUAL PROPERTY OF A RESEARCH INSTITUTE

Intellectual Property Policy

In the modern world of biomaterials and bio-engineering, it is imperative for a research institute to formulate an intellectual property policy, both to protect the assets created by the institute and to comply with any contractual obligations concerning third party intellectual property used by an institute. Invariably, a considerable amount of investment will underpin the activities of a research institute, and a

failure to secure the ownership of any intellectual property assets, which are generated by the institute can represent a squandering of that investment. The failure to manage third party intellectual property rights may also have financial repercussions, if the institute is held liable for breach of contract in relation to proprietary rights made available for limited purposes.

Compliance by an institute's researchers with an institute's intellectual property policy can be built into the contracts of employment issued by an institute. In any event, an effective intellectual property policy will be one, which is expressed to be in support of an institute's research mission. For example, intellectual property assets, which are generated can be traded for third party intellectual property rights, which are necessary to the research of the institute.

Providing for a central point of expertise to manage intellectual property is one way to assist university research institutes to deal with the complexities of the legal protection of intellectual property rights and commercial arrangements needed for the acquisition and the exploitation of those rights. Additionally, some research may depend upon the acquisition of a third party's intellectual property rights. For example, a private company may wish to contribute a patented gene. This may be on the basis that it is exclusively for the use of the institute or on condition that any improvements or applications are jointly patented. This will also be more conveniently managed centrally. Given that the right to patent an invention can be lost through premature disclosure, it may be incumbent upon an agricultural research institute to institute a system for the management of the publication activities of its researchers.

The research institute may choose between a specially dedicated research liaison office (RLO) and an official within the institute with intellectual property responsibilities. The decision will usually be made on the basis of funds available to pay salaries and to house relevant staff. A critical factor will be the expected revenues, which is expected to be generated from intellectual property by the research institute. Where the research institute is associated with a university, it will usually be able to avail itself of the services of a central university institution. Where a research institute is not able to avail itself of the facilities offered by an associated university or government body, it will probably be advised to deploy one of its officers to perform the intellectual property coordination function.

The RLO may decide not to exploit ownership for certain intellectual property, in which case, the rights will be released to the innovator(s), usually subject to a non-exclusive license to use the intellectual property for research in the research institute. In those cases where the RLO forms the view that the research institute should be involved in the exploitation of the

intellectual property, it will work with the innovator(s) in market evaluation and in finding commercial collaborators to exploit their intellectual property. Each case differs, and an appropriate strategy must be devised in conjunction with the research staff involved. The RLO will be responsible for securing professional assistance under appropriate conditions of secrecy. The ultimate vehicle for commercialization may be the research institute, through the RLO, or through the institute's commercial arm. The RLO will ensure the coordination of commercialization activities and prevent conflicts of interest.

To enable the RLO to evaluate the commercial potential of inventions and discoveries, researchers should be obliged to divulge their research activities in an appropriate form. This form will capture all preliminary information required for initial evaluation. Initial evaluation will involve assessing the invention's patentability (or other intellectual property protection) and commercial potential, and fulfilling any obligations to research sponsors. For example, companies supporting specific research projects may have in the relevant research agreement the first right to negotiate licenses for any inventions resulting from that research. Non-profit research sponsors may also have certain rights under previously negotiated research agreements. These rights and obligations must be fully defined and resolved before the invention can be made available for licensing.

The time needed to complete initial evaluation varies greatly for each invention. For example, efforts to determine commercial interest may need to await receipt of further research data or resolution of outstanding sponsor obligations.

Owing to the high cost of prosecuting and then maintaining patents, the RLO will usually have to delay filing of patent applications until strong commercial potential interest in technologies has been confirmed. In many cases, initial filing is authorized only when companies commit to inventions through letter or option agreements. During the course of the patenting and licensing process, multiple patent applications may be filed to maximize patent coverage for a given technology and cover developments made during the preparation period. Applications also may well be filed in foreign countries. Again, the commercial potential of the invention must be considered when making subsequent decisions to file or maintain patents.

Where commercialization of intellectual property rights is decided upon, the RLO will approach the most appropriate firms from a list prepared in conjunction with the innovators. The RLO will prepare executive summaries of inventions and submits these to relevant publications and databases. Marketing packages can be targeted to potential licensees identified

through discussions with inventors and the use of various internal and external resource lists. The RLO will negotiate agreements in conjunction with the innovators to ensure conformity with the role and mission of the research institute.

Once potential licensees have been identified, negotiations proceed in a variety of ways. Short-term letter agreements may be used to confirm company interest in technologies and establish their commitment to pay patent costs during an evaluation and negotiation period. Option agreements define each party's rights and responsibilities of the two parties and grant licensees the right to utilize the institute's intellectual property.

Depending on the nature of individual technologies and the estimated resources necessary for development and commercialization, the research institute may grant licensees exclusive rights, rights exclusive for a particular field of use, or non-exclusive rights. For example, most inventions in the health-care field require very large investments of time and resources to develop and commercialize. Companies will often only make the necessary investment within the context of an exclusive license agreement. In contrast, plant varieties would have already been tested and are ready for commercialization, thus they are licensed (typically) non-exclusively to multiple companies.

CONCLUSIONS

The principal objective of the intellectual property management policy of a research institute is to further the research mission of the institution. This is achieved by creating a fruitful climate for innovation and invention, and facilitating the establishment of strategic linkages with researchers in other research institutes and with the private sector. The management of

an institute's intellectual property resources may be through a specially established RLO or by utilizing the intellectual property services, which may be offered by an equivalent office within an associated institution such as a university. Through its RLO, a better understanding of the rights, which the law gives for the protection of creative effort, will be communicated to researchers and funding will be secured through collaborative research and development agreements. The RLO will be responsible for protecting and maintaining the intellectual property assets of the research institute and assist in developing an awareness and appreciation of the use of patent documents and registered plant variety data as research resources.

ARTICLE OF FURTHER INTEREST

Biomedical Ethics

REFERENCES

1. Neilson vs. Minister of Public Works (NSW) (1914), 18 CLR 423.
2. Re-application of Honeywell Bull Inc (1991), 22 INTELLECTUAL PROPERTY 463.
3. IBM Corp vs. Commissioner of *Patents* (1991), 22 INTELLECTUAL PROPERTY 417.
4. CCOM Pty Ltd vs. Jiejing Pty Ltd (1994), 28 INTELLECTUAL PROPERTY 481.
5. 447 US 303 (1980).
6. UNIDO. *Guidelines for the Acquisition of Foreign Technology in Developing Countries*; 1973; 4-5.
7. Restatement of the Law of Torts, Art. 757.
8. Rote Taube (Red Dove) (1970), 1 IIC 136.
9. 447 US 303 (1980).
10. See 1.2.1 infra.

Interpenetrating Polymeric Networks

Nicholas A. Peppas

Jing Zhang

Departments of Chemical and Biomedical Engineering and Division of Pharmaceutics,
University of Texas at Austin, Austin, Texas, U.S.A.

INTRODUCTION

Interpenetrating polymer networks (IPNs) are combinations of two or more polymers in network form, with at least one of them polymerized and/or crosslinked in the immediate presence of the other. Usually two networks are involved in an IPN system. The IPNs may have similar characteristics to polymer blends, more specifically a combined property that can be achieved from the interpenetration of the two networks. The interpenetration of two dissimilar networks may result in immiscible phases or phase separation in the IPN system. However, interpenetration of networks does not change the properties of each network. The proportion and properties of each polymer can be varied independently. Usually, network interpenetration increases the number of physical crosslinks in the system to achieve higher apparent crosslinking density and to strengthen the mechanical behavior of the system.

From a topological point of view, IPNs are closely related to polymer blends, block, graft, and crosslinked copolymers.^[1,2] An IPN can be distinguished from other multipolymer systems in three ways: 1) an IPN swells, but does not dissolve in solvents; 2) creep and flow are suppressed because of the interlocked crosslinks in the IPN; and 3) IPNs can exhibit characteristic morphologies because of mutual incompatibility.^[1]

Research on biomedical IPNs has been focused on the studies of IPN synthesis,^[1,2] morphological behavior,^[3-9] properties and industrial applications.^[1,10-12] Sperling,^[1,4,13] Klempner,^[10,14] Frisch,^[12] Lipatov,^[8,15] and their associates have contributed significantly to the area of IPNs.

BACKGROUND

Depending on the preparation method, IPNs can be classified^[1] as follows:

Sequential IPN: In sequential IPNs, the polymer network I is synthesized first. Monomer II plus a crosslinking agent and an initiator are swollen into network I and polymerized in situ.

Simultaneous interpenetrating network: The monomers and/or prepolymers plus a crosslinking agent and

initiators of both networks are mixed, followed by simultaneous polymerization via noninterfering reactions.

Latex IPN: The polymers are made in the form of lattices. Thus, each particle constitutes a micro-IPN. Usually, a core-shell structure develops in latex IPNs.

Gradient IPN: In gradient IPNs, the overall composition or crosslinking density of the material varies from location to location on the macroscopic level. One way of preparing these materials involves partial swelling of polymer network I by monomer II, followed by rapid polymerization before diffusional equilibrium takes place. Thus, a film can be made with polymer network I predominantly on one surface and polymer network II on the other surface, with a gradient composition throughout the interior.

Thermoplastic IPN: These materials utilize physical crosslinks, which may involve chain entanglements and a crystalline region. As thermoplastics, they may flow at elevated temperatures.

PHASE SEPARATION AND MECHANICAL BEHAVIOR OF IPNs

As neither of the two polymers is 100% miscible, IPNs polymer can be phase separated at some specific composition and temperature range. This is because of the low entropy of mixing caused by the low compatibility of the two components, which can lead to a negative Gibbs free energy change, as shown in Eq. (1).

$$\Delta G_{\text{mix}} = \Delta H_{\text{mix}} - T\Delta S_{\text{mix}} \quad (1)$$

Thus, several phase diagrams are possible for polymer/polymer IPN systems.^[1] For the phase separated IPNs, phase continuity, phase size, and shape all contribute to the mechanical behavior of the system.^[1] Glass transition temperatures may be broadened or shifted by greater or lesser mixing. Of course, in the limit of compatibility, only one transition will be observed. The morphology of IPNs is usually investigated by scanning electron microscopy (SEM).

Because in IPNs there is no chemical bonding between the two networks involved in the system, each network should retain its properties. Thus, the IPN

behavior should be the combination of the two networks, and the IPN property can be tuned by simply changing the composition of each component. The molecular interpenetration results in physical entanglements in the IPNs, which increase the apparent crosslinking density and further lead to higher mechanical strength of the material.

IPNs made from conventional polymers have been used as ion-exchange resins, sound and vibration damping materials, sheet molding compounds, and tough plastics^[10] depending on the homopolymers used.

Recently, IPNs made of biocompatible polymers have been proposed for applications as membranes, drug delivery carriers, artificial teeth, and artificial tissues. Environmentally sensitive biocompatible hydrogels have also been used to make IPNs^[16–19] especially for medical applications.

HYDROGELS AS IPNs

IPN hydrogels are crosslinked polymeric networks, which are able to absorb large amounts of water and swell up to equilibrium. These networks may be physically or chemically crosslinked to maintain their structural integrity.

IPN hydrogels are used in a variety of applications especially in the biomedical field. Because of the high water content of these materials, gel IPNs are very similar to natural tissue and often exhibit good biocompatibility.^[20] These characteristics allow hydrogels to be used in drug delivery systems, biosensors, contact lenses, superabsorbents, and artificial tissue. Also, because of their water content, hydrogels are good candidates for membrane applications in which the water allows for solute diffusion through the gels.^[21–26]

IPN hydrogels can be classified as neutral and ionic hydrogels depending on the type of repeating units and side chains on the polymer backbone. For example, a neutral hydrogel used in an IPN, such as poly(vinyl alcohol), does not have ionizable groups on the polymer backbone or side chain. Thus, its swelling ratio is independent of pH or ionic strength. Ionic hydrogels, on the contrary, are hydrogels with ionizable functional groups. They can be further classified as anionic and cationic hydrogels depending on the nature of ionizable moieties on their backbones.

Anionic gels contain acidic groups, which ionize when the pH of the surrounding medium rises above their pK_a . Anionic IPN gels swell at high pH values and collapse at low pH values. In the unionized state, the hydrogen of the carboxylic group might be donated to a nearby electron-rich group to form hydrogen bonding which leads to the formation of polymer complexes.

Poly(acrylic acid) (PAA) and poly(methacrylic acid) (PMAA) are typical examples of anionic hydrogels

utilized in IPNs. Similarly, copolymers of PMAA and PAA with poly(ethylene glycol)^[27,28] and poly(2-hydroxyethyl methacrylate)^[29] have been used in IPNs as drug release carriers. Copolymers of PMAA with poly(*N*-isopropyl acrylamide) (PNIPAAm) show interesting coupling of pH and temperature-sensitive swelling behavior and have been proposed for drug delivery and biomedical applications.^[30,31]

Cationic gels usually contain amino groups which form $-NH_3^+$ when the pH of the surrounding medium decreases below pK_a . Dimethylaminoethyl methacrylate and diethylaminoethyl methacrylate (DEAEM) are the typical cationic monomers used in IPNs. Cationic hydrogels swell at low pH and collapse at high pH. Thus, they can be used in insulin delivery.^[32,33]

ENVIRONMENTALLY SENSITIVE POLYMERS IN HYDROGELS

Environmentally sensitive hydrogels can be used as IPN components. They are gels whose properties, most notably their swelling ratio, change in response to specific environmental stimuli, such as temperature, pH, electric field, solvent quality, light, pressure, and ionic strength.^[34] Among these, pH-sensitive and temperature-sensitive hydrogels have been investigated extensively.^[35–42]

Polymers of methacrylic acid (MAA) display sharp sensitivity to external pH^[43,44] with capability of controlling solute permeation by the corresponding changes in swelling behavior. At high pH values, the carboxylic acid side groups become ionized, creating a repulsive force, which expands crosslinked PMAA networks. In acidic solutions, the carboxylic groups maintain the $-COOH$ structure, and the polymer remains in a collapsed state. Fig. 1 illustrates how the polymer networks response to the environmental pH.

Temperature influence on the swelling ratio in IPNs varies according to the dependence and sensitivity of a given hydrogel. Hydrogels with positive thermosensitivity experience increased swelling with increasing

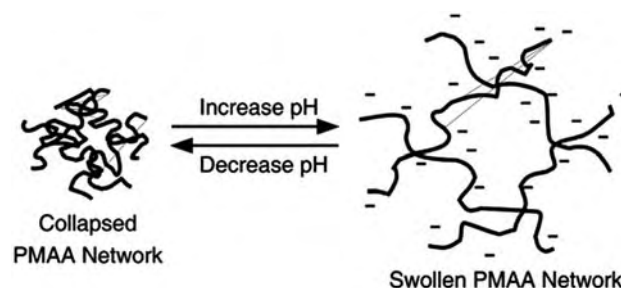


Fig. 1 The effect of pH on pure PMAA hydrogel in aqueous solution.

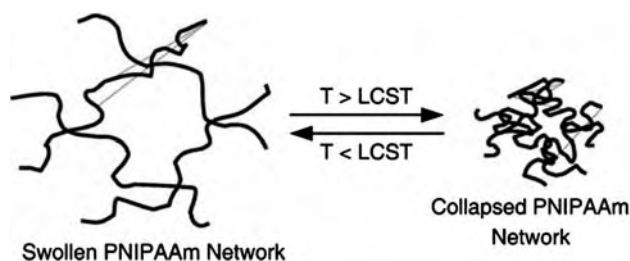


Fig. 2 The effect of temperature on the poly(*N*-isopropyl acrylamide) hydrogel in aqueous solution.

temperature. Those with negative thermosensitivity experience decreased swelling with increasing temperature as shown in Fig. 2. The affinity between solvent and the monomer units in a gel greatly affects the type of transition. Positive thermosensitive gels contain mostly hydrophilic monomers. On the other hand, negative thermosensitive gels are composed of monomers such as *N*-methylacrylamide, *N,N*-dimethylacrylamide, and *N*-isopropyl acrylamide, containing hydrophobic substituents.^[45]

Since the first report by Tanaka,^[46] crosslinked PNIPAAm have been the most studied temperature-sensitive hydrogels, especially in IPN applications. Such a swelling behavior of PNIPAAm is characterized by a lower critical solution temperature (LCST).

The LCST can be defined as the critical temperature at which a polymer solution undergoes phase transition from a soluble to an insoluble state when the temperature is raised. In the case of crosslinked hydrogels, LCST represents the volume-phase transition temperature of swelling/deswelling process. Pure PNIPAAm gels have an LCST of 32°C swelling at temperatures lower than 32°C and collapsing at temperatures higher than 32°C. The LCST of PNIPAAm can be modified by copolymerization with acrylamide, acrylic acid (AA), and alkyl methacrylates.^[47–49] However, when PNIPAAm is modified by grafting or conjugation^[50] or by interpenetrating with a second network,^[49,51–53] the LCST remained almost constant. This is significant for biomedical applications.

Many theories have been proposed to explain the LCST phenomenon in PNIPAAm single chains and gels, as summarized by Schild.^[54] In the discussion of PNIPAAm single chain theory, Heskins and Guillet^[55] noted that the water is often “ice-like” (more ordered) near the hydrophobic groups (the *N*-isopropyl) and that the ordering may be because of the hydrogen bondings between the hydrophilic group (amide) and water. They speculated that both these effects, hydrophobic interaction and hydrogen bonding, might contribute to the LCST transition. There has been much debate on whether the “hydrophobic effects” or “hydrogen bonding effects” are dominant in the

transition process. Otake et al.^[56] proposed that the driving force is entirely by “hydrophobic interactions,” while on the other extreme, Prausnitz and coworkers^[57] attributed the LCST entirely to hydrogen bonding.

The earliest prediction of the PNIPAAm behavior was by Hirotsu et al.^[58] They incorporated the Flory–Huggins mixing term and ideal elasticity into the free energy expression. However, this theory could not predict the discontinuous transitions for nonionic gels.

Prausnitz and coworkers^[57] developed a theory based on hydrogen bonding in a lattice thus overcoming the deficiencies of the Flory–Huggins theory. This theory predicts the discontinuous transition for nonionic gels and fits with Tanaka’s data, but not with Prausnitz group’s own observation. Cussler and coworkers^[59,60] defined a parameter for an interaction energy from the various cohesive energy densities in the system which fitted the experimental results but overpredicted the pressure effect on gel swelling. By assuming the mixing term as a virial expansion and assuming that the entire interaction term is derived from hydrophobic crosslinks, Saito and coworkers^[61] predicted continuous transitions that become discontinuous upon ionization.

With such diverse theories on PNIPAAm single chains and gels, it is clear that the theoretical explanation of the LCST is far from complete, especially when these networks are used in IPNs. In general, it is accepted that there is a delicate hydrophilic/hydrophobic balance (both hydrophobic interaction and hydrogen bonding are being considered) at the LCST of PNIPAAm. At temperatures above the LCST, this balance is broken by the increased mobility of both PNIPAAm chains and solvent, which lead to the collapse of the PNIPAAm chains.

APPLICATIONS OF TEMPERATURE-SENSITIVE IPN HYDROGELS

Because of their temperature responsive characteristics, temperature-sensitive hydrogels used as components in IPNs have been proposed for applications in separation, biosensors, drug delivery systems, and surface modified biomaterials.

PNIPAAm has been used in size-selective solvent extraction.^[62] Dry PNIPAAm gels can be swollen in aqueous solutions with different size molecules as solutes. When the gels swell, small size solutes enter the networks while the larger solutes are excluded from entering the gels because of size limitations of the gel mesh. Therefore, when the gels are removed, the solution with the higher molecular weight solute is concentrated. On the other hand, by raising the temperature higher than LCST, solvent with a higher concentration of the smaller molecule will be extracted from the gel.

Temperature-sensitive IPN gels have also been proposed for use in mechanical valves, where the polymer can control the extent to which a valve opens by its degree of swelling.

There are also numerous reports of using temperature-sensitive hydrogels in drug delivery applications. Okano, Bae, Kim, Hoffman, Peppas, and their coworkers have made significant contributions in this area.^[16,20,48,49,51,53,63–69]

Okano's group has also applied the temperature-sensitive PNIPAAm to prepare surface modified biomaterials and devices. By introducing PNIPAAm grafts onto surfaces with regulated morphology, the surface hydrophilicity can be modulated by external signals. Kikuchi and Okano^[70] used the surface hydrophilic/hydrophobic alteration property to create cell culture systems.

Okano et al.^[66] prepared copolymers of NIPAAm and butylmethacrylate (BMA) and IPNs of PNIPAAm and poly(tetramethylene ether glycol) (PTMEG). The shrinking behavior for disk-shaped gels with increasing temperature was dependent on the gel composition and the fabrication techniques used.

An important observation in these studies was that the initial shrinkage of the gels occurred at the surface of the membranes. The consequence of this phenomenon was that the outer surface formed a denser structure, which regulated water and subsequently the solute transport. For example, the release of indomethacin from these gels was studied, and it was shown that at low temperatures in the range of 20°C, indomethacin followed pseudo-zero-order or first-order release kinetics depending on the networks formed, while at higher temperatures of approximately 30°C, indomethacin could not diffuse out of the gels. This "on-off" release in response to temperature could be explained by the drastic surface response to the temperature rather than the bulk swelling change.

The effect of copolymerization of NIPAAm with BMA and interpenetration of PNIPAAm with PTMEG on the thermosensitivity of swelling showed that the LCST of the copolymer decreased because of the incorporation of hydrophobic BMA, while the LCST of the IPNs was not changed by the incorporation of the hydrophobic PTMEG.

Temperature- and pH-sensitive IPNs containing silicone rubber domains within copolymers of NIPAAm and AA were also synthesized by Dong et al.^[71] These gels contained macropores so that they could swell and collapse much faster than homopolymers of NIPAAm. Biocatalyst immobilization was studied.^[71] Copolymer of NIPAAm and MAA was also studied by Dong and Hoffman.^[47] P(NIPAAm-co-MAA) exhibited both pH and temperature sensitivity, and the gels showed higher swelling ratios in the swollen state because of the hydrophilicity of the MAA group.

In an effort to obtain superior biomedical IPNs, we^[30,72] synthesized copolymers of NIPAAm and MAA. These gels displayed both temperature and pH sensitivity with sharp transition regions that were functions of environmental conditions and polymer composition. The mesh space available for diffusion in the gels changed greatly between the collapsed and swollen states. This change was shown to control streptokinase release in a pulsatile pattern with respect to small changes in external pH and temperature.

When subjected to changes in temperature and pH at the same time, the swelling ratio responded more quickly and to a greater degree than when only one condition was changed. In subsequent work in our laboratory, Vakkalanka et al.^[31] showed that terpolymers of NIPAAm, AA, and 2-hydroxyethyl methacrylate (HEMA) could be prepared and used to add mechanical strength and integrity to the hydrogels. Pulsatile pH- and temperature-dependent swelling studies were performed to determine the extent and rate of gel swelling in response to changing conditions. Results showed that increasing the NIPAAm concentration in the terpolymers resulted in a higher degree of temperature-sensitive swelling. The terpolymers also exhibited an improved temperature sensitivity compared to the copolymers at a relatively low NIPAAm concentration. Poly(NIPAAm-co-BMA-DEAEMA) were also synthesized by Feil et al.^[73] with BMA as a hydrophobic component to increase the mechanical stability and DEAEMA (DEAEM) as the pH-sensitive component.

IPN gels composed of poly(ethylene oxide–dimethyl siloxane–ethylene oxide) triblock copolymers and PNIPAAm were synthesized by Mukae et al.^[52] Transmission electron microscopy and differential scanning calorimetry (DSC) studies of bulk IPNs showed phase separation with partial phase mixing between the two chemically independent networks. Aqueous equilibrium swelling behavior was examined as a function of temperature. DSC thermograms of the swollen IPNs indicated the same swelling transition temperature as that of the PNIPAAm gels. This was in sharp contrast with the random copolymer of NIPAAm in which the transition was affected by the presence of comonomer. The swelling/deswelling kinetics of IPNs was also controlled by gel compositions and applied temperatures.

Temperature-sensitive semi-IPNs of crosslinked PNIPAAm and linear poly[ether(urethane-urea)] (Biomer[®]) were prepared by Gutowska et al.^[51] The results indicated that the semi-IPNs exhibited LCST transitions. The incorporation of Biomer strongly influenced the mechanical strength, equilibrium swelling properties, and the deswelling kinetics of the semi-IPN gels. Compared to the PNIPAAm gels, the mechanical strength increased after the incorporation of hydrophobic Biomer because of the increased apparent

effective crosslinking densities and lower swelling levels. Swelling studies showed that the incorporation of Biomer did not affect the LCST but resulted in decreased thermosensitivity and lower swelling ratios. Semi-IPNs with relatively higher Biomer concentrations exhibited faster deswelling rate because of the presence of Biomer[®] which prevented the formation of the skin-type layer on the gel surface which in turn increased the deswelling rate. Loading and release of heparin was studied as a function of temperature and Biomer content in the semi-IPNs. The partition coefficients of heparin within the networks decreased with increasing temperature and Biomer content. The release of heparin was decreased by the formation of a skin layer in the hydrogels with lower Biomer content.

Lim et al.^[17] synthesized IPNs of PNIPAAm and polyurethane (PU) and IPNs of PNIPAAm and PAA. As expected, the urethane network in PNIPAAm/PU IPNs improved the mechanical strength but reduced the swelling and drug release rates because of the hydrophobic characteristics of PU. Again, the LCST was little affected by the incorporation of PU. The drug releasing ability at 28°C was controlled competitively by both PAA and PNIPAAm components. However, at 37°C, the release was controlled mostly by the PAA because of the collapse of PNIPAAm which inhibited the swelling ability of PAA network.

Research has also been conducted on temperature-sensitive IPNs based on other *N*-alkylacrylamide materials. IPNs composed of PAA and poly(acrylamide-co-butyl methacrylate) and IPNs composed of poly(acrylamide) and PAA were synthesized by Katono et al.^[16] Positive temperature sensitivity was observed in these IPNs which was attributed to the continuity of hydrogen-bonded polymer complexes between PAA and poly(acrylamide), or the so-called chain-chain zipper effect. This phenomenon was not shown in PNIPAAm-based IPNs because of the steric hindrance of the *N*-isopropyl group.

In our laboratory, IPNs composed of the temperature-sensitive PNIPAAm and the pH-sensitive PMAA were prepared by a sequential UV polymerization method.^[74-77] The IPN hydrogels were characterized for their temperature- and pH-responsive behavior by equilibrium swelling studies, oscillatory swelling studies, and DSC. The equilibrium swelling behavior as a function of temperature and pH is shown in Fig. 3 for PMAA/PNIPAAm IPNs containing 70 mol% of PNIPAAm. Fig. 4 shows the significant effect that the ionic strength has on the swelling of the same networks. The oscillatory swelling behavior as a function of time and temperature at pH values of 7.4 and 4.0 (Fig. 5) shows the versatility that these IPNs have in medical applications, as they can act as actuators or valves.

The permeability of these IPNs was investigated under various pH and temperature conditions. The

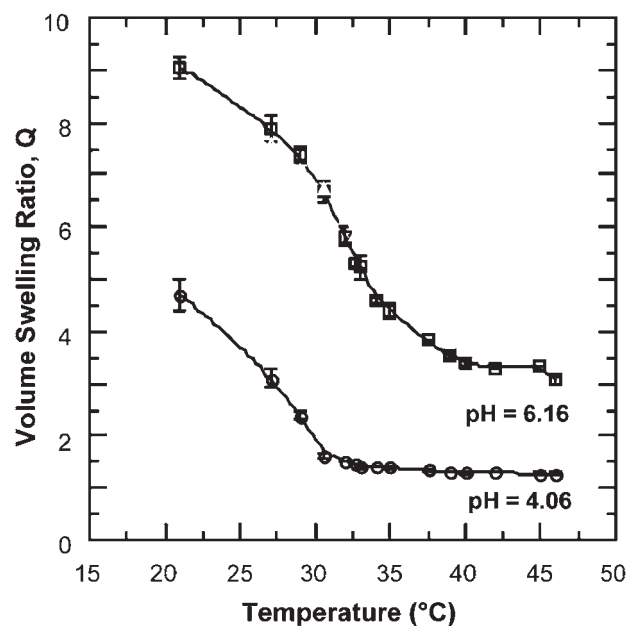


Fig. 3 Equilibrium swelling behavior as a function of temperature at pH 4.06 (○) and pH 6.16 (□) for PMAA/PNIPAAm IPNs containing 70 mol% of PNIPAAm.

results showed that these hydrogels exhibited a combined pH and temperature sensitivity at a temperature range of 31–32°C and a pH value of approximately 5.5. Permeation study results indicated a significant size exclusion behavior while model drugs with different sizes permeate through the IPN membranes. The permeability of the IPN membrane could be significantly affected by varying the pH and temperature conditions.

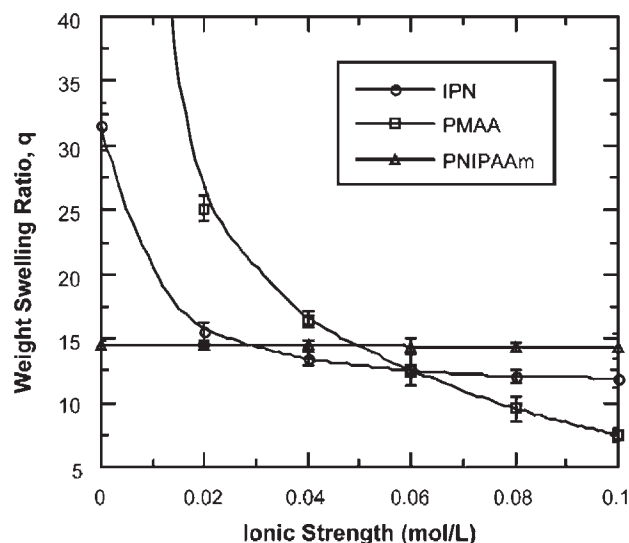


Fig. 4 Equilibrium swelling behavior as a function of ionic strength at pH 7.0, 22°C for PMAA/PNIPAAm IPN containing 70 mol% of PNIPAAm (○), PMAA (□), and PNIPAAm (Δ).

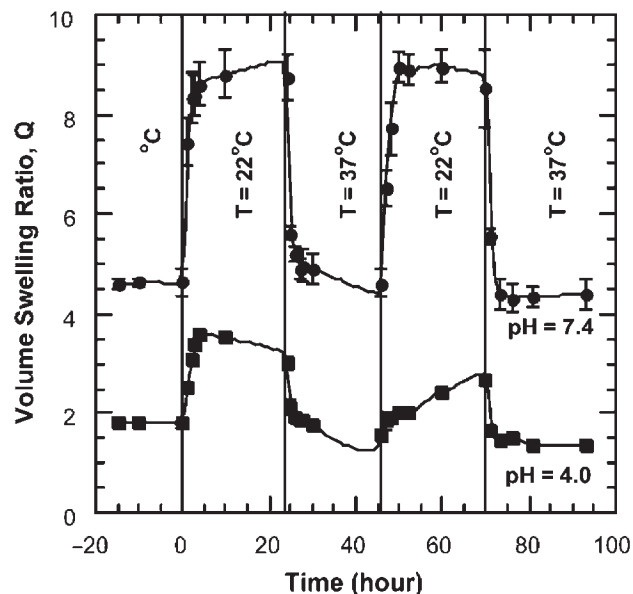


Fig. 5 Oscillatory swelling behavior as a function of time and temperature at pH 7.4 and 4.0 for PMAA/PNIPAAm IPN samples containing 70 mol% of PNIPAAm.

The morphology of these IPNs was investigated by the SEM and atomic force microscopy.^[75,77] The environmental conditions induced morphological changes for these dual sensitive IPN hydrogels which were studied by cryogenic SEM, when the gels were analyzed in their wet state. The results showed that the porous

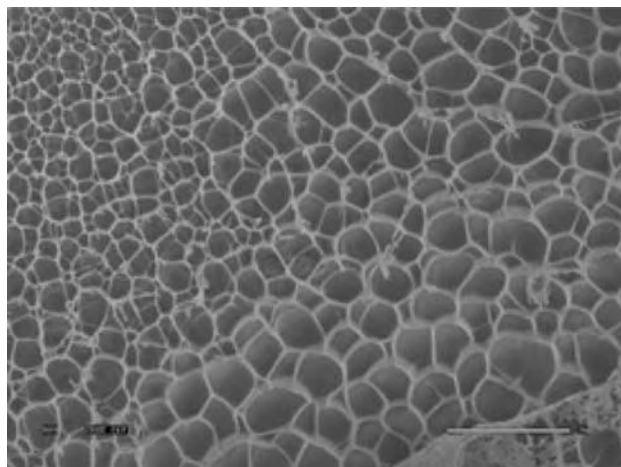


Fig. 6 Scanning electron micrograph of a PMAA/PNIPAAm IPN sample prepared from 40 vol% of MAA methanolic solution (for the primary network) and 0.2 g/ml NIPAAm methanolic solution (for the secondary network), with 1 wt% of initiator, DMPA, and 1 mol% of crosslinking agent, tetraethylene glycol dimethacrylate (TEGDMA), for each solution. Sample was dried first in air for two days and then in vacuum oven for five days and was fractured after liquid nitrogen quenching. The bar indicates 10 μm .

size (see also Fig. 6) in the IPN was strongly influenced by the environmental pH and temperature. Decrease in pH and increase in temperature resulted in significant pore size decrease for the swollen IPNs hydrogels.

CONCLUSIONS

In recent years considerable research attention has been focused on IPN hydrogels that are able to alter their volume and properties in response to environmental stimuli such as pH, temperature, ionic strength, and electric field. Because of their drastic swelling and syneresis in response to environmental stimuli, these polymeric hydrogels have been investigated for many biomedical and pharmaceutical applications, including controlled drug delivery, molecular separation, tissue culture substrates, enzyme activity controlling systems, and materials for improved biocompatibility. By chain interpenetration, we may attain combination of properties from these two polymer networks. As there is no chemical bonding between the two component networks, each network may retain its own property while the proportion of each network can be varied independently. Interpenetration of the two networks may also lead to a much higher mechanical strength with respect to the homopolymer network.

ARTICLE OF FURTHER INTEREST

Biologic and Synthetic Apatites

REFERENCES

1. Sperling, L.H. *Interpenetrating Polymer Networks and Related Materials*; Plenum: New York, NY, 1981.
2. Gupta, N.; Srivastava, A.K. Interpenetrating polymer networks: a review on synthesis and properties. *Polym. Int.* **1994**, *35* (2), 109–118.
3. Millar, J.R. Interpenetrating networks. *J. Chem. Soc.* **1960**, *311*.
4. Sperling, L.H. *Multiphase Macromolecular Systems. Contemporary Topics in Polymer Science*; Culbertson, B.M., Ed.; Plenum: New York, 1989; Vol. 6.
5. Shilov, V.V.; Lipatov, Y.S.; Karabanova, L.; Sergeeva, L. Phase separation in the interpenetrating polymeric network on the basis of polyurethane and polystyrene. *J. Polym. Sci. Polym. Chem.* **1979**, *17*, 3083.
6. Schlick, S.; Harvey, R.D.; Alonso-Amigo, M.; Klempner, D. Study of phase separation in interpenetrating polymer networks using nitroxide spin labels. *Macromolecules* **1989**, *22*, 822–830.
7. Binder, K.; Frisch, H.L. Phase stability of weakly crosslinked interpenetrating polymer networks. *J. Chem. Phys.* **1984**, *81*, 2126–2136.

8. Lipatov, Y.S. *Advances in Interpenetrating Polymer Networks*; Klempler, D., Frisch, K.C., Eds.; Technomic: Lancaster, PA, 1989; 261.
9. Ma, S.; Tang, X. Morphology of acrylic-epoxy bichain simultaneous interpenetrating polymer networks. In *Interpenetrating Polymer Networks. Advances in Chemistry*; Klempler, D., Sperling, L.H., Utracki, L.A., Eds.; American Chemical Society: Washington, DC, 1994.
10. Klempler, D., Frisch, K.C., Eds.; *Advances in Interpenetrating Polymer Networks*; Technomic: Lancaster, PA, 1994.
11. Hourston, D.J.; Huson, M.G.; McCluskey, J.A. Semi- and fully interpenetrating polymer networks based on polyurethane-polyacrylate systems. X. Polyurethane-poly(ethyl acrylate) interpenetrating polymer networks. *J. Appl. Polym. Sci.* **1986**, *32*, 3881.
12. Sophia, D.; Klempler, D.; Sendjarevic, V.; Suthar, B.; Frisch, K.C. Interpenetrating polymer networks as energy-absorbing materials. In *Interpenetrating Polymer Networks*; Klempler, D., Sperling, L.H., Utracki, L.A., Eds.; Advances in Chemistry Series, American Chemical Society: Washington, DC, 1994; 39–75.
13. Kim, S.C.; Sperling, L.H.; Eds. *IPNs Around the World: Science and Engineering*; John Wiley & Sons: New York, NY, 1997.
14. Klempler, D.; Frisch, H.L. Topologically interpenetrating elastomeric networks. *J. Polym. Sci. A-2*, **1970**, *8*, 921–935.
15. Lipatov, Y.S. Equilibrium and non-equilibrium micro-phase structures of interpenetrating polymer networks. In *Interpenetrating Polymer Networks*; Klempler, D., Sperling, L.H., Utracki, L.A., Eds.; Advances in Chemistry Series, American Chemical Society: Washington, DC, 1994; 125–139.
16. Katono, H.; Maruyama, A.; Sanui, K.; Ogata, N.; Okano, T.; Sakurai, Y. Thermo-responsive swelling and drug release switching of interpenetrating polymer networks composed of poly(acrylamide-co-butyl methacrylate) and poly(acrylic acid). *J. Contr. Rel.* **1991**, *16*, 215–228.
17. Lim, Y.H.; Kim, D.; Lee, D.S. Drug releasing characteristics of thermo- and pH- sensitive interpenetrating polymer networks based on poly(*N*-isopropylacrylamide). *J. Appl. Polym. Sci.* **1997**, *64*, 2647–2655.
18. Aoki, T.; Kawashima, M.; Katono, H.; Sanui, K.; Ogata, N.; Okano, T.; Sakurai, Y. Temperature-responsive interpenetrating polymer networks constructed with poly(acrylic acid) and poly(*N,N*-dimethylacrylamide). *Macromolecules* **1994**, *27*, 947–952.
19. Buyanov, A.L.; Revelskaya, L.G.; Kuznetsov, Y.P.; Shestakova, A.S. Cellulose-poly(acrylamide or acrylic acid) interpenetrating polymer network membranes for the pervaporation of water-ethanol mixture. *J. Appl. Polym. Sci.* **1998**, *69*, 761–769.
20. Hoffman, A.S. Conventional and environment-sensitive hydrogels for medical and industrial uses: a review paper. *Polym. Prepr.* **1990**, *31* (1), 220–221.
21. Kost, J.; Langer, R. Equilibrium swollen hydrogels in controlled release applications. In *Hydrogels in Medicine and Pharmacy*; Peppas, N.A., Ed.; CRC Press: Boca Raton, FL, 1986; 95–107.
22. Peppas, N.A. Other biomedical applications of hydrogels. In *Hydrogels in Medicine and Pharmacy*; Peppas, N.A., Ed.; CRC Press: Boca Raton, FL, 1986; 177–194.
23. Ratner, B. Biomedical applications of hydrogels: critical appraisal and review. In *Biocompatibility of Clinical Implant Materials*; Williams, D.F., Ed.; CRC Press: Boca Raton, FL, 1981; 145–153.
24. Peppas, N.A.; Korsmeyer, R.W. Dynamic swollen hydrogels in controlled release applications. In *Hydrogels in Medicine and Pharmacy*; Peppas, N.A., Ed.; CRC Press: Boca Raton, FL, 1986; 109–135.
25. Langer, R.; Hsieh, D.S.T.; Rhine, W.; Folkman, J. Control of release kinetics of macromolecules from polymers. *J. Membr. Sci.* **1980**, *7*, 330–350.
26. Langer, R. Polymeric delivery system for controlled drug release. *Chem. Eng. Comm.* **1980**, *6*, 1–48.
27. Bell, C.L.; Peppas, N.A. Poly(methacrylic acid-g-ethylene glycol) hydrogels as pH-responsive biomedical materials. In *Biomaterials for Drug and Cell Delivery*; Miikos, A.G., Murphy, R., Bemstein, H., Peppas, N.A., Eds.; Pittsburgh, PA, 1994; 199–204.
28. Lowman, A.M.; Peppas, N.A. Analysis of the complexation/decomplexation phenomena in graft copolymer networks. *Macromolecules* **1997**, *30*, 4959–4965.
29. Khare, A.R.; Peppas, N.A. Swelling/deswelling of anionic copolymer gels. *Biomaterials* **1995**, *16*, 559–567.
30. Brazel, C.S.; Peppas, N.A. Synthesis and characterization of thermo- and chemomechanically responsive poly(*N*-isopropylacrylamide-co-methacrylic acid) hydrogels. *Macromolecules* **1995**, *28*, 8016–8020.
31. Vakkalanka, S.K.; Brazel, C.S.; Peppas, N.A. Temperature- and pH- sensitive terpolymers for modulated delivery of streptokinase. *J. Biomater. Sci. Polym. Edn.* **1996**, *8*, 119–129.
32. Seminoff, L.; Olsen, G.; Zheng, J.; Kim, S.W. Self-regulated insulin release. *Proc. Int. Symp. Control. Rel. Bioact. Mater.* **1998**, *15*, 160–161.
33. Shino, D.; Murata, Y.; Kubo, A.; Kataoka, K.; Koyama, Y.; Kikuchi, A.; Yokohama, M.; Sakurai, Y.; Okano, T. Amine containing phenylboronic acid gel for glucose-responsive insulin release under physiological pH. *J. Contr. Rel.* **1995**, *37*, 269–276.
34. Gehrke, S.H. Synthesis, equilibrium swelling, kinetics, permeability and applications of environmentally responsive gels. *Adv. Polym. Sci.* **1993**, *110*, 81–144.
35. Bell, C.; Peppas, N.A. Biomedical membranes from hydrogels and interpolymer complexes. In *Biopolymer II*; Peppas, N.A., Langer, R.S., Eds.; Advances in Polymer Science; Springer-Verlag: New York, 1995; 125–176.
36. Bae, Y.H. Stimuli-sensitive drug delivery. In *Controlled Drug Delivery Challenges and Strategies*; Park, K., Ed.; American Chemical Society: Washington, DC, 1997; 147–162.
37. Kaneko, Y.; Saki, K.; Kikuchi, A.; Sakurai, Y.; Okano, T. Fast swelling/deswelling kinetics of comb-type aragted poly(*N*-isopropyl acrylamide) hydrogels. *Macromol. Symp.* **1996**, *109*, 41–53.
38. Vakkalanka, S.K.; Peppas, N.A. Swelling behavior of temperature- and pH- sensitive block terpolymers for drug delivery. *Polym. Bull.* **1996**, *36*, 221–225.

39. Chen, G.; Hoffman, A.S. Graft copolymer that exhibit temperature-induced phase transitions over a wide range of pH. *Nature* **1995**, *373*, 49–52.
40. Schwarte, L.M.; Podual, K.; Peppas, N.A. Cationic hydrogels for controlled release of proteins and other macromolecules. In *Tailored Polymeric Materials for Controlled Delivery System*; McCulloch, I., Shalaby, S.W., Eds.; American Chemical Society: Washington, DC, 1998; 57–66.
41. Ramkissoon-Ganokar, C.; Liu, F.; Baudys, M.; Kim, S.W. Modulating insulin-release profile from pH/thermosensitive polymeric beads through polymer molecular weight. *J. Contr. Rel.* **1999**, *59*, 87–98.
42. Okano, T. *Biorelated Polymers and Gels: Controlled Release and Applications in Biomedical Engineering*; Academic Press: San Diego, CA, 1998.
43. Weiss, A.M.; Adler, K.P.; Grodzinsky, A.J.; Yarmush, M.L. Variable permeability membranes. Network structure of poly(methacrylic acid) and its relation to diffusive transport. *J. Membr. Sci.* **1991**, *58*, 153–173.
44. Seno, M.; Len, M.L.; Iwamoto, K. Swelling of poly(methacrylic acid) gels and adsorption of L-lysine and its polymer on the gels. *Colloid Polym. Sci.* **1991**, *269*, 873–879.
45. Saito, S.; Konno, M.; Inomata, H. Volume phase transition of N-alkylacrylamide gels. *Adv. Polym. Sci.* **1993**, *109*, 207–232.
46. Tanaka, T. Volume phase transition in poly(N-isopropylacrylamide) gel. *Phys. Rev. Lett.* **1978**, *40*, 820.
47. Hoffman, A.S.; Afrassiabi, A.; Dong, L.C. Thermally reversible hydrogels. II. Delivery and selective removal of substances from aqueous solution. *J. Contr. Rel.* **1986**, *4*, 213–222.
48. Gutowska, A.; Bae, Y.H.; Kim, S.W. Heparin release from thermosensitive hydrogels. *J. Contr. Rel.* **1992**, *22*, 95–104.
49. Bae, Y.H.; Okano, T.; Kim, S.W. “On-off” thermo-control of solute transport. I. Temperature dependence of swelling of N-isopropylacrylamide networks modified with hydrophobic components in water. *Pharm. Res.* **1991**, *8*, 531–537.
50. Chen, J.P.; Yang, H.J.; Hoffman, A.S. Polymer-protein conjugates. I. Effect of protein conjugation on the cloud point of poly(N-isopropylacrylamide). *Biomaterials* **1990**, *11*, 625–630.
51. Gutowska, A.; Bae, Y.H.; Jacobs, H.; Feijen, J.; Kim, S.W. Thermosensitive interpenetrating polymer networks: synthesis, characterization, and macromolecular release. *Macromolecules* **1994**, *27*, 4167–4175.
52. Mukae, K.; Bae, Y.H.; Okano, T.; Kim, S.W. A new thermo-sensitive hydrogel: poly(ethylene oxide-dimethyl siloxane-ethylene oxide)/poly(N-isopropylacrylamide) interpenetrating polymer networks I. Synthesis and characterization. *Polym. J.* **1990**, *22*, 206–217.
53. Mukae, K.; Bae, Y.H.; Okano, T.; Kim, S.W. A new thermosensitive hydrogel: poly(ethylene oxide-dimethyl siloxane-ethylene oxide)/poly(N-isopropylacrylamide) interpenetrating polymer networks. II. On-off regulation of solute release from thermo-sensitive hydrogels. *Polym. J.* **1992**, *2*, 250–265.
54. Schild, H.G. Poly(N-isopropylacrylamide): experiment, theory and application. *Progr. Polym. Sci.* **1992**, *17*, 163–249.
55. Heskins, M.; Guillet, J.E. Solution properties of poly(N-isopropylacrylamide). *J. Macromol. Sci. Chem.* **1968**, *A2*, 1441–1455.
56. Otake, K.; Inomata, H.; Konno, M.; Saito, S. Thermal analysis of the volume phase transition with N-isopropylacrylamide gels. *Macromolecules* **1990**, *23*, 283–289.
57. Prange, M.M.; Hooper, H.H.; Pransnitz, J.M. Thermodynamics of aqueous systems containing hydrophilic polymers or gels. *AIChE J.* **1989**, *35*, 803–813.
58. Hirotsu, S.; Hirokawa, Y.; Tanaka, T. Volume-phase transitions of ionized N-isopropylacrylamide gels. *J. Chem. Phys.* **1987**, *87*, 1392–1395.
59. Marchetti, M.; Prager, S.; Cussler, E.L. Thermodynamic predictions of volume changes in temperature-sensitive gels. 2. Experiments. *Macromolecules* **1990**, *23*, 3445–3450.
60. Lee, K.K.; Cussler, E.L.; Marchetti, M.; McHugh, M.A. Pressure-dependent phase transitions in hydrogels. *Chem. Eng. Sci.* **1990**, *45*, 766–767.
61. Otake, K.; Inomata, H.; Konno, M.; Saito, S. A new model for the thermally induced volume phase transition of gels. *J. Chem. Phys.* **1990**, *91*, 1345–1350.
62. Freitas, R.F.S.; Cussler, E.L. Temperature sensitive gels as extraction solvents. *Chem. Eng. Sci.* **1987**, *42*, 97–103.
63. Katono, H.; Sanui, K.; Ogata, N.; Okano, T.; Sakurai, Y. Drug release OFF behavior and deswelling kinetics of thermo-responsive IPNs composed of poly(acrylamide-co-butyl methacrylate) and poly(acrylic acid). *Polym. J.* **1991**, *23*, 1179–1189.
64. Kaneko, Y.; Nakamura, S.; Sakai, K.; Aoyagi, T.; Kikuchi, A.; Sakurai, Y.; Okano, T. Rapid deswelling response of poly(N-isopropylacrylamide) hydrogels by the formation of water release channels using poly(ethylene oxide) graft chains. *Macromolecules* **1998**, *31*, 6099–6105.
65. Kim, Y.-H.; Bae, Y.H.; Kim, S.W. pH/temperature-sensitive polymers for macromolecular drug loading and release. *J. Contr. Rel.* **1994**, *28*, 143–152.
66. Okano, T.; Bae, Y.H.; Jacobs, H.; Kim, S.W. Thermally on-off switching polymers for drug permeation and release. *J. Contr. Rel.* **1990**, *11*, 255–265.
67. Dong, L.C.; Hoffman, A.S. Synthesis and application of thermally reversible heterogels for drug delivery. *J. Contr. Rel.* **1990**, *13*, 21–31.
68. Yoshida, R.; Sakai, K.; Okano, T.; Sakurai, Y. Pulsatile drug delivery systems using hydrogels. *Adv. Drug Deliv. Rev.* **1993**, *11*, 85–108.
69. Serres, A.; Baudys, M.; Kim, S.W. Temperature and pH-sensitive polymer for human calcitonin delivery. *Pharm. Res.* **1996**, *13*, 196–201.
70. Kikuchi, A.; Okano, T. Temperature-responsive polymers as On-Off switches for intelligent biointerfaces. In *Biorelated Polymers and Gels*; Okano, T., Ed.; Academic Press: San Diego, CA, 1998; 1–28.

71. Dong, L.C.; Yan, Q.; Hoffman, A.S. Controlled release of amylase from a thermal and pH-sensitive, macroporous hydrogel. *J. Contr. Rel.* **1992**, *19*, 171–178.
72. Brazel, C.S.; Peppas, N.A. Pulsatile local delivery of thrombolytic and antithrombotic agents using poly(N-isopropylacrylamide-co-methacrylic acid) hydrogels. *J. Contr. Rel.* **1996**, *39* (57–64).
73. Feil, H.; Bae, Y.H.; Feijen, J.; Kim, S.W. Mutual influence of pH and temperature on the swelling of ionizable and thermosensitive hydrogels. *Macromolecules* **1992**, *25*, 5528–5530.
74. Zhang, J.; Peppas, N.A. Synthesis and characterization of pH- and temperature-sensitive PMAA/PNIPAm interpenetrating polymer networks. *Macromolecules* **2003**, *3*, 102–107.
75. Zhang, J.; Peppas, N.A. Structure and solute size exclusion of poly(methacrylic acid)/poly(N-isopropyl acrylamide) interpenetrating polymeric networks. In *Polymeric Drugs and Drug Delivery Systems*; Ottenbrite, R.M., Kim, S.W., Eds.; Technomic: Lancaster, PA, 2001; 157–168.
76. Zhang, J.; Peppas, N.A. Molecular interactions in poly(methacrylic acid)/poly(N-isopropyl acrylamide) interpenetrating polymeric networks. *J. Appl. Polym. Sci.* **2001**, *82*, 1077–1082.
77. Zhang, J.; Peppas, N.A. Morphology of poly(methacrylic acid)/poly(N-isopropyl acrylamide) interpenetrating polymeric networks. *J. Biomater. Sci. Polym. Ed.* **2002**, *13*, 511–525.



Intravascular Ultrasound

Charles D. Choi

General Dynamics Advanced Information Systems, Fairfax, Virginia, and Department of Biomedical Engineering, University of Michigan, Ann Arbor, Michigan, U.S.A.

Yao Wang

Pfizer, Inc., Ann Arbor, Michigan, U.S.A.

Javier de Ana

Matthew O'Donnell

Department of Biomedical Engineering, University of Michigan, Ann Arbor, Michigan, U.S.A.

INTRODUCTION

Intravascular ultrasound (IVUS) is an established imaging modality used to visualize vascular pathology and morphology in real time. IVUS is a catheter-based imaging system using small ultrasound transducer(s) to image from inside a blood vessel. The two types of IVUS catheters are a 64-element circumferential array transducer and a single-element rotating (element or mirror) transducer. Images produced from IVUS are cross-sections of the artery perpendicular to the flow direction in the vessel. IVUS has several advantages over other imaging modalities. It provides real-time ultrasound images from inside the vessel at very high resolution with large penetration depths. Furthermore, IVUS enables not only anatomical imaging of vessels and plaques but also functional imaging. Currently, the primary disadvantages of IVUS include being minimally invasive (the catheter is placed inside the vasculature via the femoral artery), imaging only one cross-section of the vessel at a time, and requiring additional time and cost in the catheterization lab for the procedure. Overall, IVUS has been shown to change clinical outcomes by aiding clinicians in better diagnosing plaque composition, quantifying the degree of stenosis, determining the appropriate interventional therapy, and evaluating interventional results. Future developments in IVUS could further increase its role in the catheterization laboratory with such applications as molecular imaging, a forward-viewing array, quantitative flow imaging, and an integrated balloon ultrasound catheter guiding interventions in real time.

CLINICAL PROBLEM

Before IVUS technology and applications are discussed, it is important to understand the clinical

problem it was designed to address. Cardiovascular disease remains the primary cause of death in the United States, leading to nearly one million deaths a year. The American Heart Association estimates that nearly 60 million Americans have some form of cardiovascular disease, and the total economic burden of hospitalization, treatment, and lost time from work amounts to \$330 billion per year.^[1]

The most dangerous form of cardiovascular disease is atherosclerotic coronary artery disease, which leads to narrowing and stiffening of arteries that supply blood to the heart. Atherosclerosis results from repeated injury, subsequent inflammation, and repair of the vascular wall. Endothelial cell injury leads to an inflammatory response and lipid accumulation mostly from low-density lipoprotein cholesterol in the bloodstream. Endothelial cells, along with macrophages, promote rapid accumulation of lipids leading to fatty streaks or foam cells. At this stage, atherosclerosis is considered to be reversible with appropriate changes to risk factors (smoking, exercise, healthy diet, etc.). If atherosclerosis continues, a plaque (atheroma) forms around a lipid core. A stable plaque (fibroatheroma) is characterized by a thick fibrous cap composed of smooth muscle cells and a small lipid core. In contrast, a plaque susceptible to rupture is characterized by a large lipid core and a thin fibrous cap. They are also referred to as vulnerable plaques. Note that the lipid core is highly thrombolytic. Upon rupture, lipids interact with the blood to induce thrombosis (blood clot formation), artery blockage, and ultimately myocardial infarction (heart attack). If the atheroma stabilizes, smooth muscle cell proliferation can prevent rupture but still lead to calcification and severe fibrosis. Continued plaque growth can lead to severe stenosis (vessel occlusion) and ischemia (lack of oxygenated blood). In the past, conventional wisdom in cardiovascular research assumed that most myocardial

infarctions and strokes were caused by artery occlusion due to plaque growth.^[2] Recent studies have shown that plaque rupture, not vessel occlusion, is the leading cause of these acute events.^[3]

CLINICAL TREATMENTS

To effectively treat atherosclerosis, several interventions have been developed and are currently used in the clinic. The most common interventional procedure is percutaneous transluminal coronary angioplasty (PTCA). In PTCA, a balloon catheter is inserted into the arterial vasculature via the femoral artery over a smaller guide wire. Once the balloon catheter arrives at the coronary lesion to be dilated, the balloon is pressurized with fluid to expand the arterial lumen. The first balloon angioplasty procedure was performed in 1977 by Dr. Grüntzig in Zurich, Switzerland.^[4] The procedure is most effective for softer plaques. Although angioplasty proved effective in many cases, restenosis (subsequent occlusion of the artery lumen after the procedure) rates can still be 20–30%, based on follow-up studies performed about six months after the procedure.^[5,6]

Due to the limited success of PTCA procedures, stents were developed to be placed inside the vessel to maintain lumen diameter after the balloon intervention. Stents are meshed scaffold tubes delivered over an angioplasty balloon. As the angioplasty balloon is inflated, the stent expands and is implanted inside the vessel lumen. Since gaining FDA approval in 1994, stents have become the primary interventional procedure accounting for a majority of angioplasty procedures in catheterization laboratories.^[7] Stents effectively prop open the occluded vessel by pushing on the plaque and arterial wall to expand the lumen diameter. Clinical complications related to stent deployment are underdeployment and overdeployment. Underdeployment can lead to stent collapse if it is not in complete contact with the vessel lumen. Furthermore, blood can collect in the gap between the arterial tissue and the underdeployed stent leading to thrombosis.

To avoid these complications, stents are usually overdeployed or overpressurized. While preventing stent collapse, overpressurizing the stent can lead to vessel rupture, edge dissections of the vessel, or distal embolism. Furthermore, stent overdeployment leads to restenosis as balloon overpressurization triggers an injury response from the vessel. The inflammatory response due to endothelial cell damage results in neointimal hyperplasia (smooth muscle cell proliferation into the lumen leading to restenosis) in the very vessel the stent was intended to open. Most recently, drug-coated stents have been developed to deliver

chemicals inhibiting neointimal proliferation. The recent study by Sousa et al.^[8] shows promising initial results in human patients with no in-stent, edge restenosis, or major clinical events eight months after a stent procedure. Although more recent studies with a larger number of patients and longer follow-up evaluation have yielded less promising results,^[9] drug-coated stents appear to significantly reduce restenosis rates.

Other common interventional procedures include atherectomy, brachytherapy, and coronary bypass surgery. Atherectomy uses a small rotating blade on the tip of a catheter to cut away the plaque and collect debris with a suction tube. This procedure is usually reserved for calcified plaques resistant to angioplasty or stent procedures,^[10] but it has the risk of vessel rupture or injury. Another interventional treatment for coronary artery disease is brachytherapy. This procedure involves delivering focused radiation at the lesion site using a catheter. By delivering localized radiation to the lesion, the goal is to kill unhealthy plaque cells. In more severe cases, when angioplasty or stent procedures go awry or do not significantly change the lumen diameter of the coronary artery for the long term, coronary bypass surgery is required. Vessels are usually grafted from extremities to provide collateral flow around occluded areas in the coronary vasculature. This surgical procedure generally requires cracking the sternum and opening the rib cage to gain access to the heart.

For all of the therapies mentioned above, pharmaceuticals plays a significant and ever growing role in the treatment of cardiovascular disease. During the procedure, anticoagulants such as heparin are usually delivered immediately after the procedure to help minimize thrombosis in the treated artery. Furthermore, cholesterol-lowering drugs have been shown to be valuable adjuvants to lifestyle changes (stopping smoking, regular exercise, and healthy diet) to help manage coronary artery disease.^[11,12] Research is also being carried out to study the effects of anti-inflammatory drugs to minimize the inflammatory response in the treated region. Also, chemicals that inhibit vascular endothelial growth factor (VEGF) are used in drug-coated stents.^[8,13]

IVUS IMAGING

To help diagnose and treat cardiovascular disease, IVUS imaging was developed to help the clinician visualize arterial pathologies from inside the vessel. IVUS uses small ultrasound transducer(s) integrated into an intraluminal catheter. Brightness mode (B-mode) cross-sectional images of the vessel are acquired and displayed in real time for the clinician, as

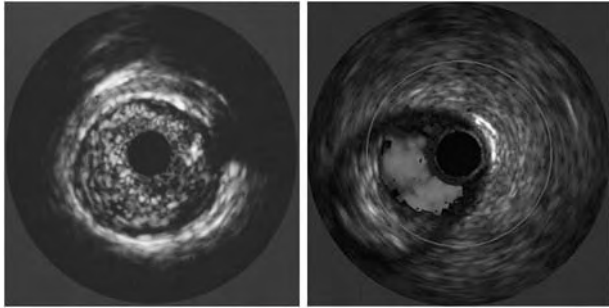


Fig. 1 B-mode IVUS image (left) of an occluded artery. A plaque filling nearly the entire lumen is easily differentiated from the vascular wall. Simultaneous B-mode and flow image highlighting the open lumen within a partially occluded vessel is shown on the right. The flow region is delineated in white (print version) or shown in red (online version). Flow images are produced using a non-Doppler decorrelation method. (Reprinted with permission from Ref. [20]. © 1997 IEEE.)

illustrated in the left panel of Fig. 1. The typical IVUS catheter is only about 1 mm in diameter and is inserted into the patient's vasculature via the femoral artery. The two most common IVUS catheters are a 64-element circumferential array transducer and a single-element rotating (element or mirror) transducer. A drawing illustrating both designs is shown in Fig. 2. The single element catheter operates at a center frequency of 30 MHz at frame rates up to 50 Hz. This catheter can produce image artifacts due to nonuniform rotational distortion, a result of frictional forces when the catheter is inside the vessel.^[10] The solid-state circumferential array transducer operates at a center frequency of 20 MHz using synthetic aperture processing to form images at frame rates of about 30 Hz. The advantages of this transducer are that there are no mechanical parts, elements are controlled electronically, and its smaller footprint allows it to be integrated into combined devices, as discussed subsequently.

Intravascular ultrasound is primarily used as a diagnostic tool to measure lumen diameter and degree of stenosis pre- and postintervention (angioplasty or stent deployment). In addition, IVUS is also used to



Fig. 2 Two types of IVUS catheters. The top drawing is a single rotating ultrasound element transducer and the lower drawing is of a solid-state ultrasound array transducer. (Reprinted with permission from Ref. [20]. © 1997 IEEE.)

measure blood flow to determine if normal physiological flow has been restored after an interventional procedure.

Intravascular ultrasound provides real-time, high-resolution (50–75 μm) images with large penetration depths of 5–8 mm. By imaging from inside the vessel, IVUS allows for both anatomical and functional imaging to help diagnose and treat cardiovascular diseases. The primary disadvantages of IVUS include being minimally invasive and imaging only one cross-section of the vessel at a time. Furthermore, IVUS often increases both the time and cost of interventional procedures.

OTHER IMAGING MODALITIES

To better understand IVUS, it is important to compare it to other imaging modalities used for cardiovascular imaging. X-ray angiography remains the gold standard for coronary arteries. Its main advantages are high spatial resolution (150 μm) and relatively low cost compared to other modalities.^[14] The main drawbacks are the radiation exposure for the patient during an intervention procedure as multiple injections of radioactive contrast agent are delivered. In general, contrast agents can be radioactive (usually for X-ray, computed tomography (CT), and nuclear imaging) or nonradioactive (usually for magnetic resonance imaging (MRI) and ultrasound), and their purpose is to enhance image quality by highlighting certain tissues or features. Another drawback to X-ray angiography is that planar X-ray imaging captures only a projection of the vasculature and can often misdiagnose asymmetric lesions.^[14]

Computed tomography is an extension angiography. X-ray CT is an imaging tool that uses a ring of X-ray detectors opposed to the X-ray transmitter. The main advantage of CT is its noninvasive, cross-sectional imaging of vessels to diagnose vascular occlusions and atherosclerotic calcified lesions. In addition, fast multislice CT allows for three-dimensional (3-D) volume data acquisition and slice images can be reconstructed with an in-plane resolution of 600 $\mu\text{m} \times 600 \mu\text{m}$ and a slice thickness of 1 mm. CT can also be used to image vessels after an interventional procedure and during noninvasive follow-up studies. Unfortunately, CT, like X-ray angiography, exposes the patient to radioactive iodinated contrast agents often used during the procedure.

Another noninvasive imaging modality is MRI. MRI uses a high-field whole-body magnet and radio-frequency coils to manipulate nuclear spins in the body for functional and anatomical imaging.^[15] MRI has the advantage of being able to identify large lipid regions in a plaque, which is useful for identifying vulnerable

plaques. Furthermore, like CT, MRI collects a 3-D volume of data and allows for any slice to be viewed after the data acquisition. The resolution for magnetic resonance angiography images is typically about $1000\ \mu\text{m} \times 1000\ \mu\text{m}$ (or $1\ \text{mm} \times 1\ \text{mm}$) with a frame rate of about 5 Hz. MRI slice thickness can range from 1 to 10 mm, depending on frame rate. In MRI, paramagnetic contrast agents can be used to enhance images. The drawbacks of MRI are the long data acquisition times and the tradeoffs of higher resolution images for lower frame rates.

Nuclear imaging uses radiopharmaceuticals injected into the patient intravenously in small doses to study physiological parameters (as opposed to imaging anatomical features) by collecting images over time. As different regions of the body selectively absorb these radioactive materials, gamma rays are emitted and detected. Although nuclear imaging provides important functional information, both spatial and temporal resolutions are lower than with other modalities, and the patient must be injected with radioactive isotopes to achieve contrast.

Optical coherence tomography (OCT) imaging is a relatively new modality developed by Fujimoto in 1991.^[16] OCT operates much like a single-element ultrasound transducer but uses optics to transmit and receive one-dimensional signals, which are then combined to form an image. Its primary advantage is high resolution (about $10\ \mu\text{m}$) at a frame rate around 10 Hz. This resolution is on the order of histological images and can help characterize plaques and image inside a vessel. The major obstacle to vascular imaging with OCT is that the transmitted light is highly scattered by blood.^[17]

CLINICAL APPLICATIONS

Conventional IVUS images show the vessel cross-section to a depth of 5–8 mm from the catheter. Quantitative lumen measurements are made from these images to diagnose vessel stenosis. With the development of vascular interventional procedures such as angioplasties and stents, comparing pre- and post-procedure lumen measurements is especially useful in evaluating an intervention. Thus, a clinician can immediately determine if further interventions are necessary.

In addition to visualizing lumen morphology, IVUS images can also assess coronary blood flow during diagnostic or posttherapeutic interventions, as illustrated in the right panel of Fig. 1. One of the challenges to measuring flow with IVUS is that blood flows normal to the side-looking imaging plane (vessel cross-section). Therefore, Doppler processing cannot be used to image the flow field. Recent developments in

IVUS technology have allowed for real-time blood speed imaging with IVUS using a filter bank approach.^[18] With quantitative flow measurements, clinicians can evaluate the degree of stenosis before and after an intervention and validate the procedure before pulling back the catheter.

FUTURE DIRECTIONS

Intravascular ultrasound technology is continually being improved to create high-resolution images (through smaller high-frequency ultrasound elements) and smaller catheters. In addition to these advances in the catheter itself, several promising new applications and IVUS designs could lead to a more significant role for IVUS in diagnosing and treating cardiovascular disease.

Due to the increase in cardiovascular diseases, there has been a growing interest in measuring mechanical properties of arterial tissues and pathologies using strain imaging and IVUS.^[19–21] If vulnerable plaques can be identified and their mechanical properties quantitatively measured, they can help guide interventional procedures and plaque rupture could be better understood. IVUS images can be used to produce strain and elasticity images (with deformations applied by a balloon catheter or the cardiac pulse) to differentiate soft (lipid) and hard (fibrotic) regions of a plaque. Figure 3 shows a thrombus located inside a rabbit thoracic aorta that is hard to distinguish from the vessel in the B-mode image on the left but clearly visible in the strain image on the right. Tissue displacements can be accurately measured with speckle-tracking algorithms by imaging from inside a balloon

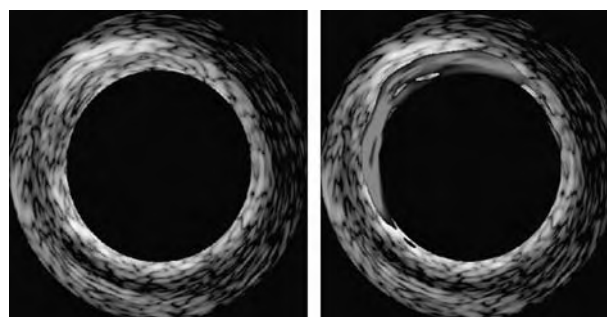


Fig. 3 IVUS images of an excised rabbit thoracic aorta with a thrombus imaged from inside a balloon catheter placed in a water tank. The left image is the B-mode and the right is the strain image. The B-mode image is displayed over a 40-dB log scale and strains are displayed over a truncated range of 10% (black) to 35% (white). The strain image highlights the soft thrombus inside the aorta that is not clearly visible in the B-mode image alone. (Reprinted with permission from Ref. [25]. © 2002 IEEE.)

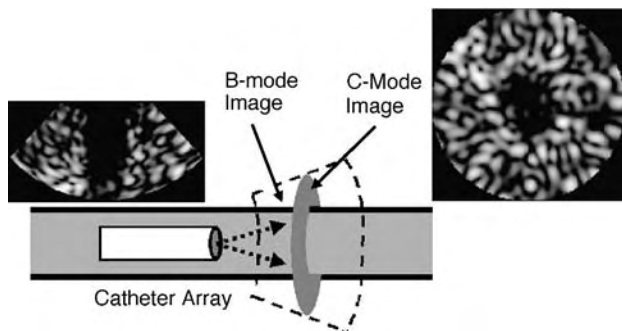


Fig. 4 Forward-viewing IVUS array drawing and simulated images. The forward-viewing IVUS array enables 3-D imaging and imaging of the C-mode plane denoted by the dark gray disk and the simulated image on the right. The B-mode image in this geometry is orthogonal to the C-mode plane, as shown on the left.

catheter as a deformation is applied to the artery or lesion. Displacement data are accumulated and processed to create strain and elasticity images of vascular tissues and pathologies. Clinical evidence has shown that distensibility of atherosclerotic coronary lesions may vary widely, from undilatable hard lesions to soft highly compliant ones.^[22] In particular, the elastic moduli of a normal arterial wall and a coronary plaque can vary by over an order of magnitude.^[23]

Another exciting future direction for IVUS is a forward-viewing catheter to image areas of the vessel in front of the device.^[24] Using a ring-annular array of ultrasound elements, 3-D data can be acquired to image tissue and flow in front of the catheter. Multi-view images, as shown in Fig. 4, would allow the clinician to accurately guide the IVUS catheter to the lesion of interest and view the plaque without disturbing it. The C-mode image in this geometry is similar to the conventional B-mode image in the side-looking IVUS catheter and images the vessel cross-section. The B-mode image in this geometry is orthogonal to this plane and can image along the length of the vessel. Using this forward-viewing B-mode image, Doppler flow measurements would be possible in the region distal to the IVUS array. This forward-viewing array can be fabricated by modifying the current solid-state IVUS catheter to image in both the side-viewing (conventional) and forward-viewing modes to allow for imaging at the location of the catheter as well as the region in front of it.

Conventional IVUS imaging can be extended to provide 3-D images by acquiring a series of real-time cross-sectional images and using image segmentation and rendering to allow the clinician to view the vessel and pathology from any angle. The physician could rotate the 3-D image and also view different imaging planes (or slices) to obtain a better view of the plaque and vessel morphology. Three-dimensional

IVUS images would be particularly useful for unusual vessel or plaque shapes, suspicious plaques, and regions near vessel branches.

Perhaps the greatest clinical limitation in balloon angioplasty and stent deployment procedures is the lack of image guidance and quantitative feedback during the procedure. Even preprocedure estimates of lumen diameter and occlusion can lead to significant errors due to planar projections obtained by X-ray angiography.^[14] Due to the lack of image guidance, stents are often underdeployed, leading to thrombosis or stent collapse, or overdeployed, resulting in greater intimal injury and higher restenosis rates. However, current uses of IVUS are still limited to imaging before and after the procedure but not *during* it. Furthermore, imaging pre- and postprocedure requires the use of additional catheters and lengthens the total intervention time. Both of these factors can significantly add to the total cost for patient treatment.

To address these important clinical problems, an integrated balloon ultrasound catheter has been proposed to combine IVUS imaging and therapy in realtime during an interventional procedure.^[19,25] Figure 5 shows a drawing of the integrated device along with a photo. The clinical opportunities and the significance of an integrated device are numerous. First, imaging from inside a balloon catheter enables strain imaging of arterial pathologies to characterize tissues and guide interventions. By imaging in real time as the balloon probes the tissue, lesions could be characterized before a complete angioplasty procedure is carried out. Even during stent procedures, strain imaging could prevent overdeployment. For balloon procedures such as angioplasty and stent deployment, the integrated balloon ultrasound catheter could

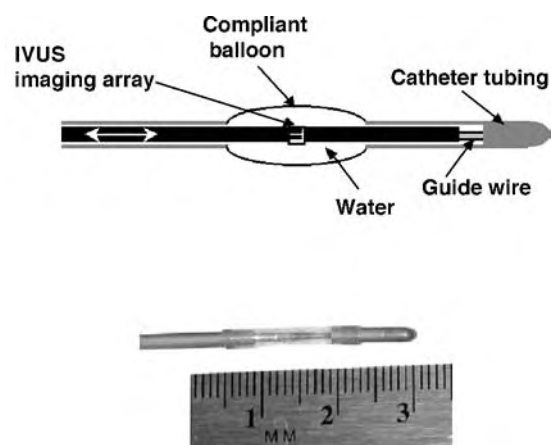


Fig. 5 Integrated balloon ultrasound catheter. The top is a drawing with the IVUS array shown inside the balloon catheter. Below the drawing is a photo of an integrated balloon ultrasound catheter prototype. (Reprinted with permission from Ref. [25]. © 2002 IEEE.)

provide both real-time image guidance and quantitative feedback during the interventional procedures. Both visual and quantitative feedback would help clinicians achieve optimal stent deployment. After a procedure is performed, the balloon can be deflated and moved in pullback mode along the length of the lesion to verify that the stent has been fully deployed. Last, the current limitations of IVUS can be overcome by providing a single catheter device for balloon interventions and reducing total intervention time, at the same time providing significant feedback during the procedure and allowing for a more controlled and reliable intervention.

The ultimate goal of an integrated device is to significantly change clinical outcomes by reducing intimal injury and subsequent restenosis rates. By verifying stent deployment with the same integrated balloon ultrasound catheter after the procedure, complete deployment can be easily validated in pullback mode. If any regions are underdeployed, the balloon can easily be realigned and reinflated using the same metric and real-time image guidance. Ultimately, the integrated balloon ultrasound catheter could transform cardiovascular interventions into a single catheter procedure. By reducing the number of catheters used, the total procedure time would decrease for the patient and lead to increased patient throughput for a catheterization laboratory.

In addition to exploring these potential advancements in IVUS technology, several research groups are currently investigating ultrasound-mediated gene therapy and molecular imaging applications with high-frequency ultrasound. Recent work by one research group has explored ligand-targeted contrast agents using microbubbles for IVUS.^[26,27] The potential benefits of molecular imaging are the ability to detect, monitor, and treat different pathologies. Other groups are studying the effects of IVUS therapy to enhance transgenic expression^[28] and reduce neointimal hyperplasia.

CONCLUSION

Intravascular ultrasound is a valuable imaging modality in the catheterization laboratory used to diagnose cardiovascular diseases. Recent developments and current research show promise for IVUS gaining a larger role in guiding interventional procedures. With the development of a forward-viewing IVUS array and an integrated balloon ultrasound catheter, IVUS can be used as a single interventional device to not only diagnose but also guide therapies and provide real-time feedback to the clinician. In addition, further advances in IVUS catheter fabrication will lead to higher-resolution images and smaller catheters. Last,

emerging applications such as 3-D rendering, strain processing, and molecular imaging will continue to make IVUS an important imaging tool in treating cardiovascular disease.

ARTICLES OF FURTHER INTEREST

Medical Imaging, 3-D; Ultrasound Doppler; Ultrasound and Tissue Interaction; Vascular Grafts; Vascular Grafts: Host Interactions; Vascular Imaging

REFERENCES

1. American Heart Association. *2002 Heart and Stroke Statistical Update*; American Heart Association: Dallas, TX, 2001.
2. Peterson, M.; Dangas, G.; Fuster, V. Atherosclerosis and Thrombosis. In *The Manual of Interventional Cardiology*, 3rd Ed.; Safian, R.D., Freed, M.S., Eds.; Physicians' Press: Royal Oak, MI, 2001; 959–973.
3. Lee, R.T.; Libby, P. The unstable atheroma. *Atheroscler. Thromb. Vasc. Biol.* **1997**, *17*, 1859–1867.
4. Norell, M.S. History, Development and Current Activity. In *Essential Interventional Cardiology*; Norell, M.S., Perrins, E.J., Eds.; Harcourt Publishers Limited: London, England, 2001; 3–12.
5. Schiele, F.; Vuilleminot, A.; Meneveau, N.; Pales-Espinosa, D.; Gupta, S.; Bassand, J-P. Effects of increasing balloon pressure on mechanism and results of balloon angioplasty for treatment of restenosis after Palmaz-Schatz stent implantation: An angiographic and intravascular ultrasound study. *Catheter. Cardiovasc. Interv.* **1999**, *46*, 231–314.
6. Shiran, A.; Mintz, G.S.; Waksman, R.; Mehran, R.; Abizaid, A.; Kent, K.M.; Pichard, A.D.; Satler, L.F.; Popma, J.J.; Leon, M.B. Early lumen loss after treatment of in-stent restenosis: An intravascular ultrasound study. *Circulation* **1998**, *98*, 200–203.
7. Lee, S.F.; Brooks, M.R.; Livingston, D.L.; Squire, J.C. Smart catheter for stent placement. *IEEE Potentials* **2003**, *22*, 19–22.
8. Sousa, J.E.; Costa, M.A.; Abizaid, A.; Abizaid, A.S.; Feres, F.; Pinto, I.M.F.; Seixas, A.C.; Staico, R.; Mattos, L.A.; Sousa, A.G.M.R.; Falotico, R.; Jaeger, J.; Popma, J.J.; Serruys, P.W. Lack of neointimal proliferation after implantation of sirolimus-coated stents in human coronary arteries: A QCA and 3-D IVUS study. *Circulation* **2001**, *103*, 192–195.
9. Liistro, F.; Stankovic, G.; Di Mario, C.; Takagi, T.; Chieffo, A.; Moshiri, S.; Montorfano, M.; Carlino, M.; Briguori, C.; Pangotta, P.; Albiero, R.; Corvaja, N.; Colombo, A. First clinical experience with a Paclitaxel derivative-eluting polymer stent system implantation for in-stent restenosis. *Circulation* **2002**, *105*, 1883–1886.
10. Atar, S.; Siegel, R.J. Intravascular Echocardiography. In *Imaging in Cardiovascular Disease*; Pohost, G.M.,

- O'Rourke, R.M., Eds.; Lippincott Williams and Wilkins: Philadelphia, 2000; 97–113.
11. Dupis, J.; Tardif, J.C.; Cernacek, P.; Theroux, P. Cholesterol reduction rapidly improves endothelial function after acute coronary syndromes. The RECIFE (reduction of cholesterol in ischemia and function of the endothelium) trial. *Circulation* **1999**, *99*, 3227–3233.
 12. Mikhailidi, D.P.; Ganotakis, E.S.; Spyropoulos, K.A.; Jagroop, I.A.; Byrne, D.J.; Winder, A.F. Prothrombotic and lipoprotein variables in patients attending a cardiovascular risk management clinic: Response to ciprofibrate or lifestyle advice. *Int. Angiol.* **1998**, *17*, 225–233.
 13. Grube, E.; Gerckens; Rowold, S.; Yeung, A.C.; Stertz, S.H. Inhibition of in-stent restenosis by a drug-eluting polymer stent: Pilot trial with 18-month follow-up. *Circulation* **2000**, *102*, II-554.
 14. Ohnesorge, B.M.; Becker, C.R.; Flohr, T.G.; Reiser, M.F. *Multi-slice CT in Cardiac Imaging: Technical Principles, Clinical Application and Future Developments*; Springer: New York, 2002.
 15. Nishimura, D.G. *Principles of Magnetic Resonance Imaging*; Nishimura: Stanford, CA, 1996.
 16. Huang, D.; Swanson, E.A.; Lin, C.P.; Schuman, J.S.; Stinson, W.G.; Chang, W.; Hee, M.R.; Flotte, T.; Gregory, K.; Puliafito, C.A.; Fujimoto, J.G. Optical coherence tomography. *Science* **1991**, *254*, 1178–1181.
 17. Brezinski, M.E.; Fujimoto, J.G. Optical coherence tomography: High resolution imaging in non-transparent tissue. *IEEE J. Sel. Top. Quantum Electron.* **1999**, *5*, 1185–1192.
 18. Crowe, J.R.; O'Donnell, M. Quantitative blood speed imaging with intravascular ultrasound. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **2001**, *48*, 477–487.
 19. Shapo, B.M.; Crowe, J.R.; Skovoroda, A.R.; Eberle, M.J.; Cohn, N.A.; O'Donnell, M. Displacement and strain imaging of coronary arteries with intraluminal ultrasound. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **1996**, *43*, 234–246.
 20. O'Donnell, M.; Eberle, M.J.; Stephens, D.N.; Lizza, J.L.; Shapo, B.M.; Crowe, J.R.; Choi, C.D.; Chen, J.J.; Muller, D.W.M.; Kovach, J.A.; Lederman, R.J.; Ziegenbein, R.C.; San Vicente, K.; Bleam, D. In Catheter Arrays: Can Intravascular Ultrasound Make a Difference in Managing Coronary Artery Disease?, Proceedings of the IEEE Ultrasonics Symposium, Toronto, Canada, October, 1997; IEEE UFFC, 1997; 1447–1456.
 21. de Korte, C.L.; van der Steen, A.F.W.; Cespedes, E.I.; Pasterkamp, G. Intravascular ultrasound elastography in human arteries: Initial experience in vivo. *Ultrasound Med. Biol.* **1998**, *24*, 401–408.
 22. Cespedes, E.I.; de Korte, C.L.; van der Steen, A.F.W. Intraluminal ultrasonic palpation: Assessment of local and cross-sectional tissue stiffness. *Ultrasound Med. Biol.* **2000**, *26*, 285–296.
 23. Falk, E. Why do plaques rupture? *Circulation* **1992**, *86* (Suppl. III), III-30–III-42.
 24. Wang, Y.; Stephens, D.N.; O'Donnell, M. Optimizing the beam pattern of a forward-viewing ring-annular ultrasound array for intravascular imaging. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **2002**, *12*, 1652–1664.
 25. Choi, C.D.; Skovoroda, A.R.; Emelianov, S.Y.; O'Donnell, M. An integrated compliant balloon ultrasound catheter for intravascular strain imaging. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **2002**, *11*, 1552–1560.
 26. Lanza, G.M.; Wallace, K.D.; Scott, M.J.; Catheris, W.P.; Kaufmann, R.J.; Gaffney, P.J.; Wickline, S.A. A novel site-targeted ultrasonic contrast agent with broad biomedical application. *Circulation* **1998**, *94*, 3334–3340.
 27. Lanza, G.M.; Trousil, R.L.; Wallace, K.D.; Rose, J.H.; Hall, S.C.; Scott, M.J.; Miller, J.G.; Eisenberg, P.R.; Gaffney, P.J.; Wickline, S.A. In vitro characterization of a novel, tissue-targeted ultrasound contrast system with acoustic microscopy. *J. Acoust. Soc. Am.* **1999**, *104*, 3665–3672.
 28. Lawrie, A.; Brisken, A.F.; Francis, S.E.; Tayler, D.I.; Chamberlain, J.; Crossman, D.C.; Cumberland, D.C.; Newman, C.M. Ultrasound enhances reporter gene expression after transfection of vascular cells in vitro. *Circulation* **1999**, *99*, 2617–2620.

In Vivo Cardiovascular Models

B. H. Walpoth

Department of Cardiovascular Surgery, University Hospital, Geneva, Switzerland

O. M. Hess

Swiss Cardiovascular Center, University Hospital, Bern, Switzerland

INTRODUCTION

Synthetic vascular grafts are increasingly used for replacement of medium and large arteries. However, synthetic grafts have failed as vascular conduits in small arteries (<6 mm) such as coronary and peripheral vessels.^[1,2] Various attempts have been made to overcome this problem by using autologous material (vein, artery, pericardium, umbilical cord, etc.) with satisfactory long-term results.^[3-6] However, in the case of diseased autologous material or previous surgery, homologous or, eventually, heterologous or synthetic material have to be used, but results are less satisfactory.^[7-10] Tissue engineering, as a new promising technique, has gained a lot of interest for improving long-term patency of small synthetic vascular grafts by seeding endothelial and/or smooth muscle cells.^[11-16] Animal models are, therefore, essential for preclinical assessment of novel synthetic or tissue-engineered vascular grafts.^[17-19]

The main goal of these animal models is:

- To test new synthetic or biological materials (non-degradable or degradable).
- To study blood and tissue compatibility.
- To assess rejection and healing processes.
- To evaluate genetically modified auto-/allo- or xenografts.

MODELS

Application of new devices for implantation in humans requires thorough in-vitro testing for physical characteristics (compliance and tensile strength) and tissue reactions in cell culture models. As a next step, in vivo models should be used to test for vascular pathophysiology and several conditions have to be taken into consideration for the testing of vascular prostheses, low and high pressure systems; pulsatile and non-pulsatile flow; conduit or capacitance vessels; presence or absence of endothelium; and normal or arteriosclerotic vessels (inflammation).^[20-24]

Animal Models

Multiple factors must be considered for animal models, such as animal species, circulation models, implantation sites, and concomitant treatment.^[25-40]

Rodents (mice and rats) have several advantages, such as availability and cost effectiveness, but operations and follow-up tests are usually difficult and results are less applicable to humans.^[28,29] The rabbit model is often chosen for its ability to replicate human arteriosclerotic lesions after cholesterol-rich feeding.^[39] Pigs and primates are used for preclinical studies, especially when similar devices as in humans have to be used, such as stents, valve prostheses, etc.^[40] A few questions can be answered by using acute experiments such as feasibility studies for graft suturing or early blood thrombogenicity testing. Mostly, chronic experiments are applied in order to assess changes over time (mainly in small animals) and long-term patency, etc. In larger animals, these changes can be assessed repetitively using newer, sophisticated, noninvasive and invasive methods such as echo-doppler, MRI imaging and flow studies, quantitative angiography, and/or intravascular ultrasound (IVUS).^[41-43]

Site of Implantation

Major differences are found after vascular implantation in a low-pressure (venous side) or a high-pressure (arterial side) system. Furthermore, results will be influenced by the size of the prosthesis, the native flow, and peripheral resistance of the vessel being tested.^[44] In general, coronary arteries show a high flow rate compared to their size and, thus, better results than distal peripheral artery replacements of comparable sizes.

Different surgical techniques can be employed such as end-to-end or end-to-side bypass anastomosis or even a looped vascular graft.^[45-47] Furthermore, animal treatment (such as antithrombotic agents) will vary according to the species used.^[48] As a general rule, vascular replacement of the infrarenal aorta typically shows a good long-term patency without any drug therapy in the rat model. In contrast, carotid artery

replacements in the pig model with 3–4 mm synthetic vascular grafts require antithrombotic long-term therapy to yield a satisfactory patency rate.^[38,49]

Graft Material

During the last 20 years, innumerable tests have been carried out with various synthetic and biological grafts to overcome the bad patency rates in small-caliber vascular prostheses.^[11–16,50–55]

The best clinical results for arterial and venous replacement are obtained with autologous materials such as saphenous vein, internal mammary artery-free transplants, or in situ, or radial, artery.^[3–6] Clinically, an off-the-shelf vascular graft with all diameters and lengths comparable to autologous material would be the surgeon's dream. However, even the best clinical long-term results for small vascular prostheses are unsatisfactory. For large arterial replacements such as aortic aneurysms, Dacron and expanded-polytetrafluoroethylene (ePTFE) prostheses have been used with good long-term results due to the large size and high-flow situation.^[18] For peripheral vascular surgery, synthetic ePTFE grafts have a certain role, but patency is limited. These grafts can be improved by surface coatings such as heparin bonding, bi-layer surface treatment, etc.^[49,56–59] Others have tested biological replacement grafts such as homografts (cryopreserved) or xenografts such as glutaraldehyde-denatured bovine carotid arteries or uvine ureters.^[9,10] More recent studies have explored autologous endothelial cell seeding on synthetic ePTFE grafts with promising results.^[11,14] The technique, however, is time- and cost-intensive and cannot be used for emergency surgery because the grafts have to be prepared with autologous culture-expanded cells.

Vascular tissue engineering using nondegradable or degradable synthetic grafts or grafts produced from recombinant human proteins have been used successfully as scaffolds for cell growth.^[60–62] Potential cells to be used in such scaffolds are autologous endothelial cells, smooth muscle cells, or progenitor stem cells.^[63–65] Experimental and clinical data on autologous cell seeding techniques on nondegradable vascular grafts have shown reasonably good results.^[11,66] However, regardless of the method (direct seeding or culture expansion), the rate of surface coverage after implantation is still incomplete (about 50–75%).^[67] One of the explanations for this observation is the high shear stress in the arterial circulation.^[23] The use of progenitor or even embryonic stem cells has not yet passed the experimental stage. Progenitor cells have to be harvested and enriched before graft seeding. More recently, several groups have been working with the

concept of using blood-borne endothelial progenitor cells to attach to the inner graft surface in vivo (statins and/or growth factors).^[68–70] Ultimately, a three-dimensional biodegradable vascular scaffold made out of synthetic or, better yet, of human recombinant proteins with embryonic or progenitor cell seeding (either in vitro or in vivo), may be the solution for the future.^[71,72]

Functional Assessment

A variety of functional tests are available in animal models. The larger the animal, the easier the use of these techniques, which normally include flow, pressure, and resistance measurements. Most commonly the transit time Doppler flow technique is used (CardioMed, Norway). This technique represents a fast and cost-effective method that yields excellent accuracy and reproducibility.^[73,74] Other techniques include Duplex sonograph or angiographic examinations. These methods combine a morphological image with some information about blood flow and flow distribution. The assessment of compliance is difficult under in-vivo conditions but comparisons with normal vessels may give important information on long-term patency.^[44–46] After explantation of the vessels, vasomotion can be tested with the vascular ring technique. This technique, however, can be applied only to autologous or possibly tissue-engineered neovessels.^[75]

Morphological Assessment

On completion of a study, an in situ pressure-fixation with glutaraldehyde is recommended in order to preserve all anatomical details of the anastomosis and the vascular prosthesis because any fixation will create a shrinkage of the vessel wall. After explantation several parameters can be determined: Macroscopically, the in-growth of the prosthesis into the adjacent tissue is thought to have a beneficial effect on patency (due to adventitial neovessels promoting endothelial coverage) and may represent some protection against possible graft infections.^[76] Certainly such tissue ingrowth may induce fibrosis with a reduction in vascular compliance leading possibly to accelerated intimal hyperplasia reaction and graft occlusion.^[77] Microscopically, patency, luminal surface, intimal hyperplasia, and tissue reaction, as well as prosthesis in-growth, can also be studied. Intimal hyperplasia formation is normally assessed quantitatively with micrometry on vascular longitudinal or cross sections. Normally the intimal hyperplasia thickness as well as the ratio of stenosis to the normal vessel is calculated.^[49] In the case of biological or degradable grafts, degradation rate and

foreign body reaction can be assessed.^[78] Scanning electron microscopy (SEM) is the method of choice to examine the inner surface.^[79] An evaluation of cellular and platelet deposition and microthrombus formation is indicative of blood compatibility.^[80–82] Further studies such as immunohistochemistry, molecular analysis, or gene expression may be used to evaluate molecular adaptation of these vascular prostheses to the surrounding tissue.^[33,83]

Description of Some Currently Used In-Vivo Cardiovascular Models

Blood thrombogenicity in the acute pig model

Early blood thrombogenicity may be tested in a femorofemoral arteriovenous shunt model. A specific canula is placed in the femoral artery and connected to a similar canula in the femoral vein. The vascular grafts (normally pieces of about 3–5 cm length with various diameters) are placed between these two canulae. Blood is then circulated for varying time intervals. Normally morphological and structural changes are assessed after 3, 9, and 30 minutes. These tests are performed under constant flow and pressure. A typical study protocol includes the following two steps: perfusion without heparin followed by perfusion with systemic heparinization of the animal.

Morphological examination of the vascular grafts is done on the explanted pieces after pressure fixation using scanning electron microscopy. Figures 1 and 2

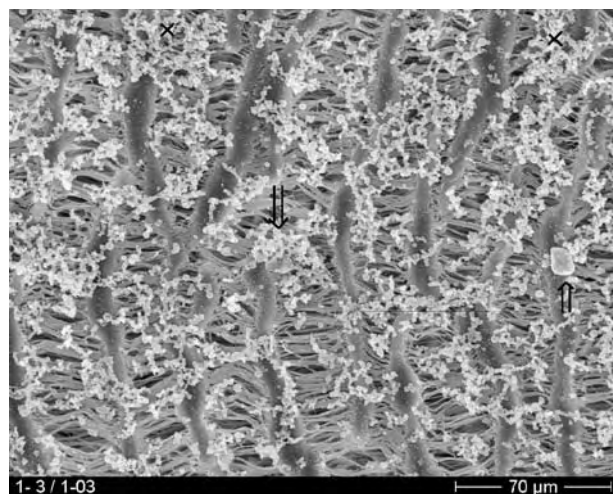


Fig. 1 Scanning electron microscopy (SEM) picture of the inner surface of the ePTFE (control) graft after 9 minutes blood perfusion (arteriovenous shunt model in the pig). Note the moderate cellular (WBC [↓]) deposition and the severe platelet deposition [↑] with some micro-thrombi formation [X].

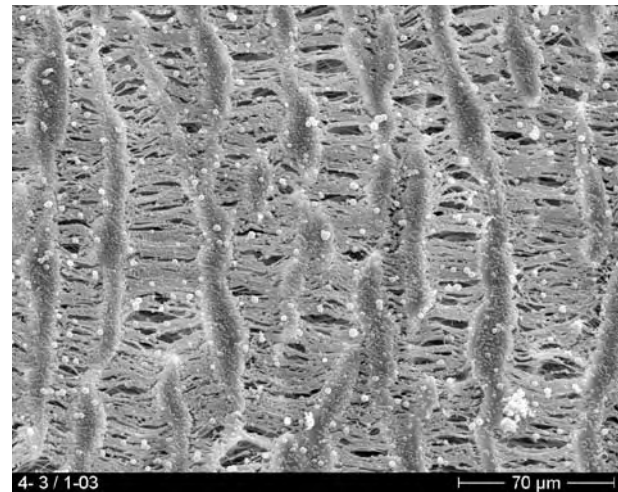


Fig. 2 SEM picture of the inner surface of the ePTFE-coated graft with PPS-PEG dry after 9 minutes of blood perfusion. Note the striking difference from Fig. 1 with only little white blood cell deposition and no platelets or microthrombi.

show the same ePTFE graft (control vs. coating) after 9 minutes of perfusion without heparin. Note the significant cellular and platelet deposition in Fig. 1 (control) as compared to Fig. 2 using polypropylene sulfide and poly-ethylene glycol (PPS-PEG) coating.^[84] Thrombogenicity is normally evaluated by the amount of cellular deposition and especially the deposition of platelets and microthrombi.

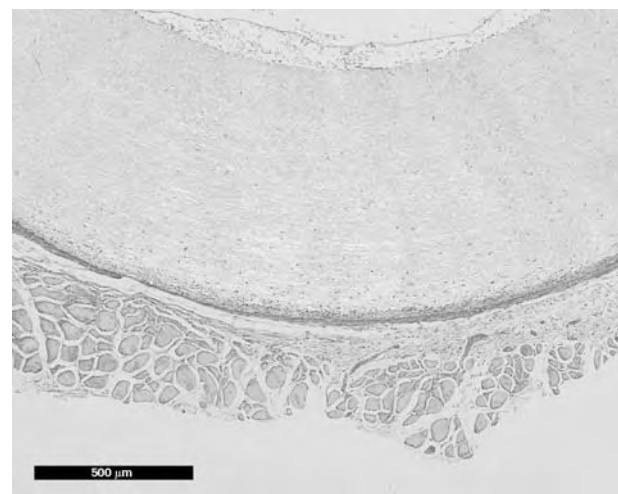


Fig. 3 Histological cross section of a poly(glycolic acid) (PGA) graft manufactured by electrospinning after 2 weeks of subcutaneous implantation. Note the early degradation of the material (loosening of the material) and mild cellular ingrowth as well as tissue reaction of the perivascular tissue.

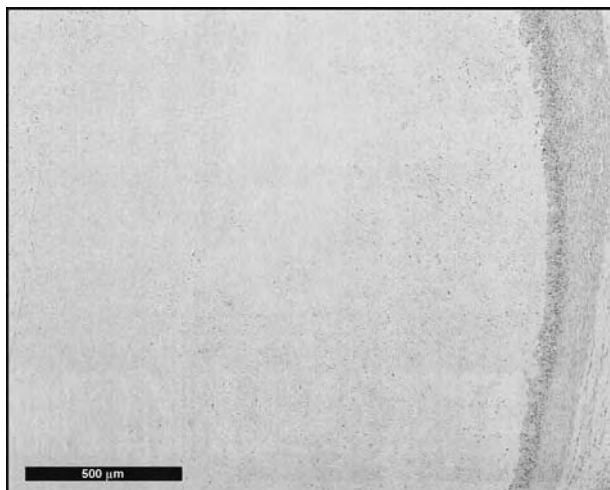


Fig. 4 Histological cross section of a poly(lactic acid) (PLA) graft, manufactured by electrospinning after 2 weeks of subcutaneous implantation. Note the advanced degradation of the graft material with a more pronounced tissue reaction of the perivascular tissue compared to Fig. 3.

Degradation and tissue reaction in the rat model

Subcutaneous implantation of pieces of vascular grafts is used, e.g., for testing new degradable synthetic or biological graft material. Advantages of this technique are the simple surgery and the fact that several materials can be compared in the same animal. In addition, material calcification can be easily studied in this model before using costly long-term survival experience as, for example, when studying new cell-seeded or tissue-engineered valves or grafts. Besides histological staining, a high resolution radiography will clearly

show microcalcification in the material.^[29,85] After subcutaneous implantation animals are sacrificed at 2, 4, 8, or 12 weeks and morphological assessments are performed as mentioned. Figures 3 and 4 are examples of degradable synthetic grafts after two weeks of implantation. Note the slight degradation with mild cellular in-growth and tissue reaction of the perivascular tissue in Fig. 3 poly(glycolic acid) (PGA) as compared to Fig. 4 polylactic acid (PLA) showing an advanced degradation of the graft with more tissue reaction than in Fig. 3.

Vascular graft replacement in the small animal

Following subcutaneous implantation of the grafts, the vascular prostheses with degradation rate ranging between 4 and 8 weeks with little tissue reaction are implanted in a small animal for assessment of graft patency and intimal hyperplasia formation.^[49] This implies microvascular surgery using magnifying loops or an operative microscope with specialized instruments and trained personnel. Due to the small sample size, noninvasive follow-up examinations are difficult and normally animals will be euthanized for morphological assessment in the rat model. Adequate information is normally obtained after 4 weeks of implantation. However, long-term implantation up to 6 or 12 months is advisable for assessing intimal hyperplasia, calcification, and aneurysm formation. Figure 5 shows a polyurethane graft at infrarenal aortic implantation in the rat (left panel) and at sacrifice after 4 weeks of implantation (middle panel). The right panel shows an histological longitudinal section with graft anastomosis.

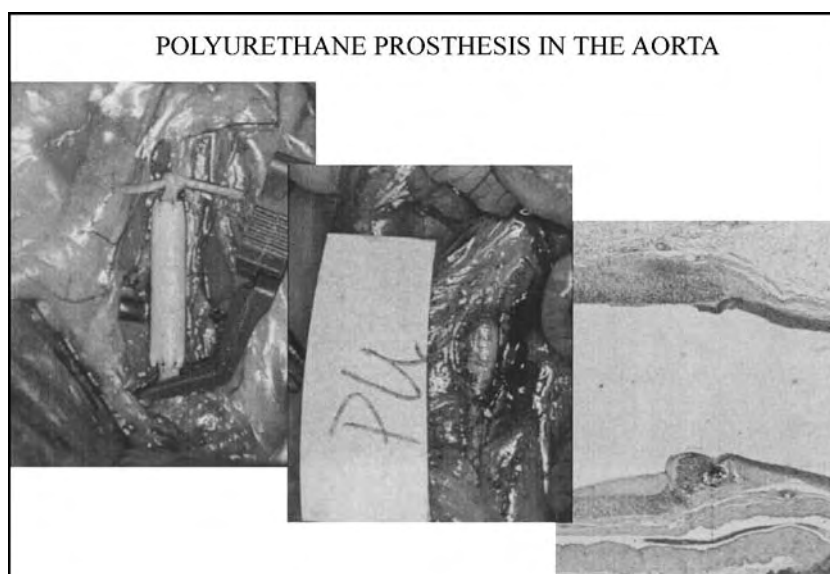


Fig. 5 The polyurethane graft upon implantation at infrarenal aortic position in the rat (left panel) and at sacrifice after 4 weeks of implantation (middle panel). The right panel shows a histological longitudinal section with graft anastomosis. Towards the left, the graft structure is clearly seen covered by about ten layers thickness of neointimal cells. In the middle, some tissue reaction is visible mainly due to the surgical anastomosis, and on the right side, the normal arterial wall structure is recognizable.

Pig carotid bypass

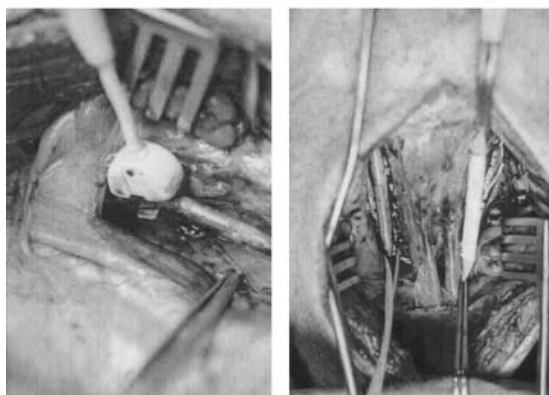


Fig. 6 Carotid artery flow measurement using the transit time technique (left) and an ePTFE carotid artery end-to-end replacement (right).

Vascular graft replacement in the atherosclerotic (cholesterol-rich diet) rabbit

The majority of experimental studies are performed in healthy juvenile vessels. In clinical situations, vascular replacement, however, is performed most often in patients with advanced atherosclerotic disease. Thus, the animal models may not be appropriate to answer clinical questions and requirements.^[39] The high cholesterol-fed rabbit model is well established and induces severe arteromatous vascular lesions that are similar to the early changes seen in human arteriosclerosis.^[39]

These changes appear in the rabbit after 4–6 weeks, but lack calcification and plaque ulcerations. Nevertheless, it is the model of choice for testing new anti-proliferative drugs such as statins or rapamycin or taxol.^[83,86]

Bilateral carotid artery replacement in the pig

This animal model has several advantages compared to studies based on rodents. In the past, such experiments have been performed with mongrel or dogs bred for experimental use with good results. An advantage of the canine model is the large size and good surgical accessibility of not only the carotid vessels but also femoral arteries.^[57,59] However, for ethical reasons, most studies with dogs have been discontinued. In the porcine model, vascular anatomy and morphology of the heart and great vessels are similar to humans. Furthermore, the advantage of using a bilateral carotid artery replacement is that each animal can serve as its own control, reducing interanimal variability.^[31,38,87] Of equal importance is the small size of the carotid artery in the pig; such a vessel diameter (3–4 mm internal diameter), would be appropriate for human coronary bypass vessels. Furthermore, clinically available invasive and noninvasive methods can be used repetitively for assessing long-term patency and intimal hyperplasia formation. Figure 6 shows a carotid artery flow measurement using the transit time

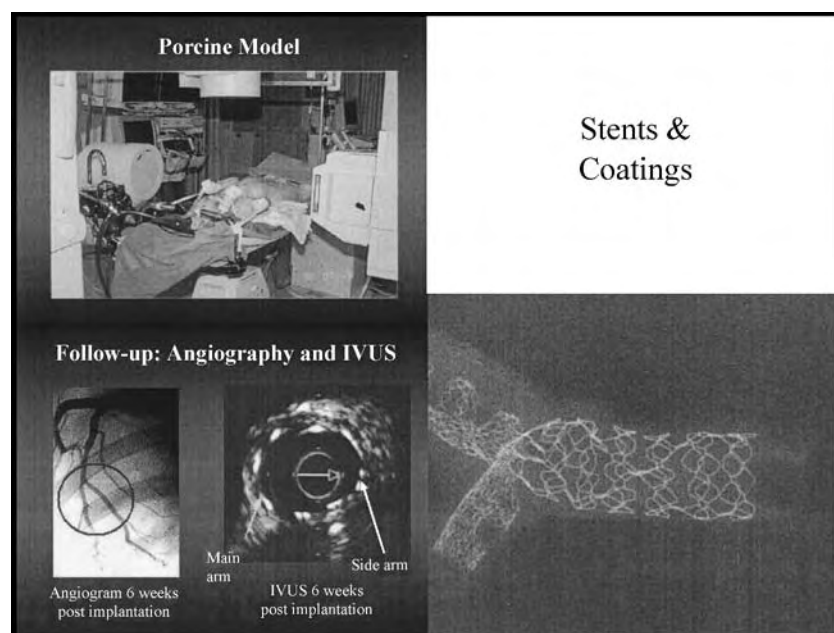


Fig. 7 Animal at follow-up during selective coronary angiography and intravascular Doppler flow measurement 6 weeks after implantation of a bifurcation stent (left). The angiogram shows a widely patent stented bifurcation, which is confirmed by IVUS. The right panel shows an ex-vivo high-resolution radiograph of the stent structure.

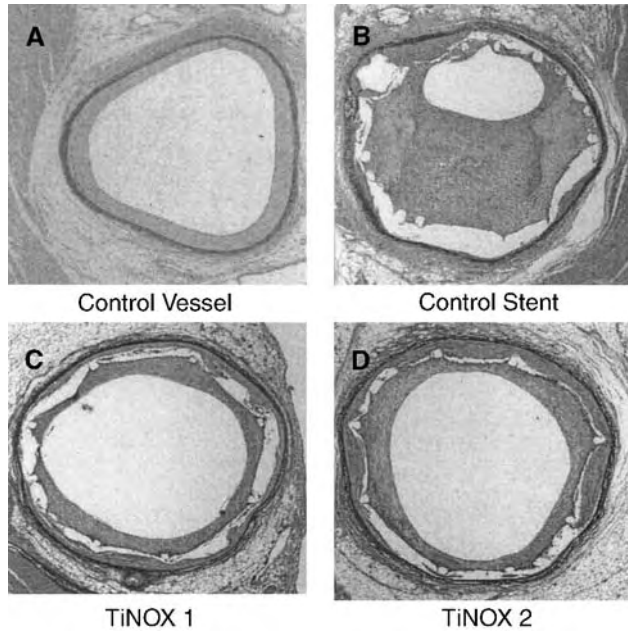


Fig. 8 Cross sections of coronary arteries of the pig, either native (A), or after 6 weeks of stent implantation (B,C,D). Note the severe intimal formation (near occlusion) in the control stents as compared to the stents coated with TiNOX 1 or 2. (Reproduced from Ref. [93] with permission from Circulation.)

technique (left) and an ePTFE carotid artery end-to-end replacement (right).

Porcine coronary artery overstretch model and stent implantation

The porcine coronary anatomy is similar to humans, especially with regard to collateralization.^[88] Therefore, this model is often used for preclinical testing of vascular implants. The Achilles heel of percutaneous

transluminal coronary angioplasty (PTCA) is restenosis due to excessive neointimal hyperplasia with negative vascular remodeling.^[89] Intracoronary stents have reduced restenosis but in-stent restenosis (constrictive vascular remodeling) remains a major problem. Thus, new stent designs with passive or active surface coatings with or without drug elution are often tested in the porcine overstretch model.^[90-92] (See Fig. 7.)

The effect of novel stent coatings on coronary artery remodelling has been studied in 12 pigs, each instrumented with three stainless steel stents (15 mm length, 2.5–3.0 mm diameter), and a control and two TiNOX-coated (titanium-nitrid-oxide) stents (TiNOX 1 = ceramic; TiNOX 2 = metallic coating). After 6 weeks of survival, animals were euthanized and histological specimens of the stented segments were performed by digital morphometry. Largest neointimal hyperplasia (expressed as surface area) was found in uncoated (2.6 mm²) followed by TiNOX 1-coated (1.5 mm²) and TiNOX 2-coated (1.4 mm²) stent (Fig. 8). Thus, restenosis rate was reduced by 40% in TiNOX-coated stents.^[93]

In a similar study, the effect of poly(propylene sulfide)-poly(ethylene glycol) PPS-PEG and gold coating was tested (Fig. 9). The results show a significantly lesser degree of intimal proliferation in the PPS-PEG-coated stents compared to control or gold-coated stents.^[94]

Despite a large variety of in-vitro tests and new diagnostic procedures including molecular and genetic analysis, in-vivo testing of cardiovascular devices and vascular prostheses remains important before clinical implantation. Several models have been described above, each of which having certain advantages. A stepwise progression from acute to small chronic implantations followed by chronic long-term testing in the large animals is recommended.

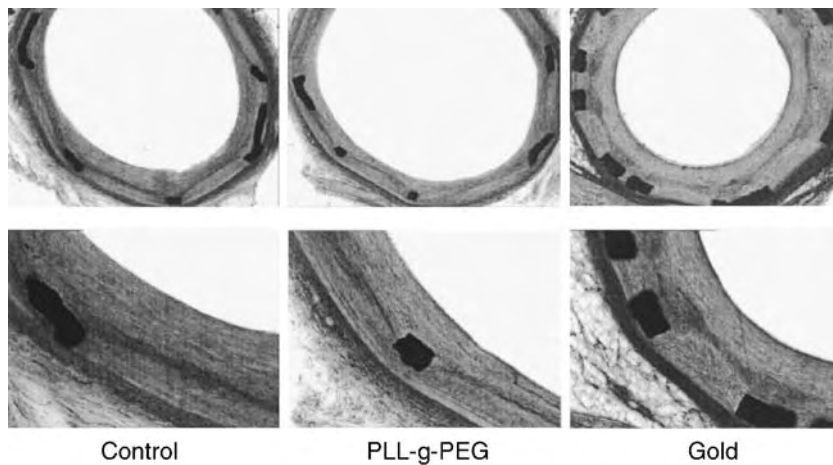


Fig. 9 Side-by-side histological cross sections (low- and high-magnification) after 6 weeks of implantation for (left) control: uncoated stented vessel; (center) PPS-PEG coated stented vessel; and (right) gold-coated stented vessel.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Mustafa Cikirkcioglu for his expert knowledge of the literature in this field. The cited studies have been partly supported by the Swiss National Science Foundation and the Swiss Heart Foundation.

ARTICLES OF FURTHER INTEREST

Artificial Heart Fluid Dynamics: Positive Displacement Pumps; Biocompatibility Testing; Blood-Material Interactions; Blood Vessel Mechanics; Cell Culture Assays; Thrombosis; Tissue Engineering of Blood Vessel; Tissue Engineering of Cardiac Muscle; Vascular Grafts: Development Strategies; Vascular Grafts: Host Interactions

REFERENCES

- Weinstein, P.R.; Reinert, R.L.; Brittain, F. Delayed thrombosis of synthetic microvascular bypass grafts. *Neurosurgery* **1981**, *9* (3), 268–274.
- Madras, P.N.; Johnson, W.R.; Ward, C.A. Enhancement of thrombus formation by vascular prostheses. *Can. J. Surg.* **1980**, *23* (6), 544–546.
- Goldman, S.; Zadina, K.; Krasnicka, B.; Moritz, T.; Sethi, G.; Copeland, J.; Ovitt, T.; Henderson, W.; Department of Veterans Affairs Cooperative Study Group No. 297 Predictors of graft patency 3 years after coronary artery bypass graft surgery. *J. Am. Coll. Cardiol.* **1997**, *29* (7), 1563–1568.
- Plecha, E.J.; Freischlag, J.A.; Seabrook, G.R.; Towne, J.B. Femoropopliteal bypass revisited, an analysis of 138 cases. *Cardiovasc. Surg.* **1996**, *4* (2), 195–199.
- Cooper, G.J.; Underwood, M.J.; Deverall, P.B. Arterial and venous conduits for coronary artery bypass, a current review. *Eur. J. Cardio-Thorac. Surg.* **1996**, *10* (2), 129–140.
- Benedetti-Valentini, F.; Gossetti, B.; Irace, I.; Martinnelli, O.; Gattuso, R. Composite grafts for critical ischemia. *Cardiovasc. Surg.* **1996**, *4* (3), 372–376.
- Donzeau-Gouge, P.; Touati, G.; Vouhe, P.R.; Roy, A.; Farge, C.; Leca, F.; Neveux, J.Y. Coronary artery bypass with bovine internal mammary artery graft. *Arch. Mal. Coeur Vaiss.* **1990**, *83* (12), 1811–1815.
- Iaffaldano, R.A.; Lewis, B.E.; Johnson, S.A.; Piffare, R.; McKiernan, T.L. Patency of cryopreserved saphenous vein grafts as conduits for coronary artery bypass surgery. *Chest* **1995**, *108* (3), 725–729.
- Conklin, B.S.; Richter, E.R.; Kreutziger, K.L.; Zhong, D.S.; Chen, C. Development and evaluation of a novel decellularized vascular xenograft. *Med. Eng. Phys.* **2002**, *24* (3), 173–183.
- Lee, W.K.; Park, K.D.; Han, D.K.; Suh, H.; Park, J.C.; Kim, Y.H. Heparinized bovine pericardium as a novel cardiovascular bioprosthesis. *Biomaterials* **2000**, *21* (22), 2323–2330.
- Zilla, P.; Fasol, R.; Deutch, M.; Fischlein, T.; Minar, E.; Hammerle, A.; Krupicka, O.; Kadletz, M. Endothelial cell seeding of polytetrafluoroethylene vascular grafts in humans. *J. Vasc. Surg.* **1987**, *6* (6), 535–541.
- Massia, S.P.; Hubbell, J.A. Tissue engineering in the vascular graft. *Cytotechnology* **1992**, *10* (3), 189–204.
- Nerem, R.M.; Seliktar, D. Vascular tissue engineering. *Annu. Rev. Biomed. Eng.* **2001**, *3*, 225–243.
- Dunn, P.F.; Newman, K.D.; Jones, M.; Yamada, I.; Shayani, V.; Virmani, R.; Dichek, D.A. Seeding of vascular grafts with genetically modified endothelial cells. *Circulation* **1996**, *93* (7), 1439–1446.
- Teebken, O.E.; Haverich, A. Tissue engineering of small diameter vascular grafts. *Eur. J. Vasc. Endovasc. Surg.* **2002**, *23* (6), 475–485.
- Ratcliffe, A. Tissue engineering of vascular grafts. *Matrix Biol.* **2000**, *19* (4), 353–357.
- Valente, M.; Pettenazzo, E.; Di Filippo, L.; Laborde, F.; Rinaldi, S.; Thiene, G. Biodegradable polymer (D,L-lactide-caprolactone) in aortic vascular prosthesis, morphological evaluation in an animal model. *Int. J. Artif. Organs* **2002**, *25* (8), 777–782.
- Zippel, R.; Wilhelm, L.; Marusch, F.; Koch, A.; Urban, G.; Schlosser, M. Antigenicity of polyester (Dacron) vascular prostheses in an animal model. *Eur. J. Vasc. Endovasc. Surg.* **2001**, *21* (3), 202–207.
- Macneill, B.D.; Pomerantseva, I.; Lowe, H.C.; Oesterle, S.N.; Vacanti, J.P. Toward a new blood vessel. *Vasc. Med.* **2002**, *7* (3), 241–246.
- Merhi, Y.; Bernier, J.; Marois, Y.; Guidoin, R.L. Acute thrombogenicity of arterial prostheses exposed to reduced blood flow in dogs, effects of heparin, aspirin and prostacyclin. *J. Cardiovasc. Pharmacol.* **1995**, *26* (1), 1–5.
- Ziegler, T.; Nerem, R.M. Tissue engineering a blood vessel, regulation of vascular biology by mechanical stresses. *J. Cell. Biochem.* **1994**, *56* (2), 204–209.
- Hess, F.; Steeghs, S.; Jerusalem, R.; Reijnders, O.; Jerusalem, C.; Braun, B.; Grande, P. Patency and morphology of fibrous polyurethane vascular prostheses implanted in the femoral artery of dogs after seeding with subcultivated endothelial cells. *Eur. J. Vasc. Surg.* **1993**, *7* (4), 402–408.
- Fernandez, P.; Bareille, R.; Conrad, V.; Midy, D.; Bordenave, L. Evaluation of an in vitro endothelialized vascular graft under pulsatile shear stress with a novel radiolabeling procedure. *Biomaterials* **2001**, *22* (7), 649–658.
- Nerem, R.M. Role of mechanics in vascular tissue engineering. *Biorheology* **2003**, *40* (1–3), 281–287.
- Abbott, W.M.; Callow, A.; Moore, W.; Rutherford, R.; Veith, F.; Weinberg, S. Evaluation and performance standards for arterial prostheses. *J. Vasc. Surg.* **1993**, *17* (4), 746–756.
- Sparks, S.R.; Tripathy, U.; Broudy, A.; Bergan, J.J.; Kumins, N.H.; Owens, E.L. Small-caliber mesothelial cell-layered polytetrafluoroethylene vascular grafts in New Zealand white rabbits. *Ann. Vasc. Surg.* **2002**, *16* (1), 73–76.

27. Shigematsu, K.; Yasuhara, H.; Shigematsu, H. Topical application of antiangiogenic agent AGM-1470 suppresses anastomotic intimal hyperplasia after e-PTFE grafting in a rabbit model. *Surgery* **2001**, *129* (2), 220–230.
28. Yue, X.; Van der Lei, B.; Schakenraad, J.M.; Van Oene, G.H.; Kuit, J.H.; Feijen, J. Smooth muscle cell seeding in biodegradable grafts in rats, a new method to enhance the process of arterial wall regeneration. *Surgery* **1988**, *103* (2), 206–212.
29. Joshi, R.R.; Underwood, T.; Frautschi, J.R.; Phillips, R.E.; Schoen, F.J.; Levy, R.J. Calcification of polyurethanes implanted subdermally in rats is enhanced by calciphylaxis. *J. Biomed. Mater. Res.* **1996**, *31* (2), 201–207.
30. Bobryshev, Y.V.; Inder, S.J.; Cherian, S.M.; Lord, R.S.; Ao, P.Y.; Hawthorne, W.J.; Fletcher, J.P. Colonisation of prosthetic grafts by immunocompetent cells in a sheep model. *Cardiovasc. Surg.* **2001**, *9* (2), 166–176.
31. Ortenwall, P.; Bylock, A.; Kjellstrom, T.; Risberg, B. Seeding of ePTFE carotid interposition grafts in sheep and dogs, species-dependent results. *Surgery* **1988**, *103* (2), 199–205.
32. Galletti, G.; Ussia, G.; Farruggia, F.; Baccarini, E.; Biagi, G.; Gogolewski, S. Prevention of platelet aggregation by dietary polyunsaturated fatty acids in the biodegradable polyurethane vascular prosthesis, an experimental model in pigs. *Ital. J. Surg. Sci.* **1989**, *19* (2), 121–130.
33. Marois, Y.; Guidoin, R.; Wagner, E.; Roy, R.; Douville, Y. Fresh venous allografts as arterial substitutes in dogs, the importance of donor-recipient tissue compatibility. *J. Invest. Surg.* **1994**, *7* (5), 393–408.
34. Akoum, A.; Guidoin, R.; King, M.W.; Marois, Y.; Sigot, M.; Sigot-Luizard, M.F. A new bioactive molecule for improving vascular graft patency, exploratory trials in dogs. *Clin. Invest. Med.* **1992**, *15* (4), 318–330.
35. Mehran, R.J.; Ricci, M.A.; Graham, A.M.; Carter, K.; Symes, J.F. Porcine model for vascular graft studies. *J. Invest. Surg.* **1991**, *4* (1), 37–44.
36. van der Bas, J.M.; Quax, P.H.; van den Berg, A.C.; van Hinsbergh, V.W.; van Bockel, J.H. Ingrowth of aorta vascular cells into basic fibroblast growth factor-impregnated vascular prosthesis material, a porcine and human in vitro study on blood vessel prosthesis healing. *J. Vasc. Surg.* **2002**, *36* (6), 1237–1247.
37. Tucker, O.P.; Syburra, T.; Augstburger, M.; van Melle, G.; Gebhard, S.; Bosman, F.; von Segesser, L.K. Small intestine without mucosa as a growing vascular conduit, a porcine experimental study. *J. Thorac. Cardiovasc. Surg.* **2002**, *124* (6), 1165–1175.
38. Lin, P.H.; Chen, C.; Surowiec, S.M.; Conklin, B.; Bush, R.L.; Chaikof, E.L.; Lumsden, A.B.; Weiss, V. A porcine model of carotid artery thrombosis for thrombolytic therapy and angioplasty, application of PTFE graft-induced stenosis. *J. Endovasc. Ther.* **2000**, *7* (3), 227–235.
39. Bocan, T.M. Animal models of atherosclerosis and interpretation of drug intervention studies. *Curr. Pharm. Des.* **1998**, *4* (1), 37–52.
40. Kleine, P.; Abdel-Rahman, U.; Kleisus, A.A.; Sherer, M.; Simon, A.; Moritz, A. Comparison of hemodynamic performance of Medtronic Hall 21 mm versus St. Jude Medical 23 mm prostheses in pigs. *J. Heart Valve Dis.* **2002**, *11* (6), 857–863.
41. McLennan, G.; Trerotola, S.O.; Forney, M.; Jellison, B.; Dreesen, R.G.; Tennery, J. Short-term patency and safety of an expanded polytetrafluoroethylene encapsulated endoluminal device at the venous anastomosis of a canine arteriovenous graft model. *J. Vasc. Interv. Radiol.* **2001**, *12* (2), 227–234.
42. Omary, R.A.; Green, J.D.; Schirf, B.E.; Li, Y.; Finn, J.P.; Li, D. Real-time magnetic resonance imaging-guided coronary catheterization in swine. *Circulation* **2003**, *107* (21), 2656–2659.
43. Odenstend, H.; Aneman, A.; Oi, Y.; Svensson, M.; Stenqvist, O.; Lundin, S. Descending aortic blood flow and cardiac output, a clinical and experimental study of continuous oesophageal echo-doppler flowmetry. *Acta Anaesthesiol. Scand.* **2001**, *45* (2), 180–187.
44. Dobrin, P.B.; Mirande, R.; Kang, S.; Dong, Q.S.; Mrkvicka, R. Mechanics of end-to-end artery-to-PTFE graft anastomoses. *Ann. Vasc. Surg.* **1998**, *12* (4), 317–323.
45. Morasch, M.D.; Dobrin, P.B.; Dong, Q.S.; Mrkvicka, R. Mechanics of spatulated end-to-end artery-to-vein anastomoses. *Ann. Vasc. Surg.* **1998**, *12* (1), 55–59.
46. Tozzi, P.; Hayoz, D.; Ruchat, P.; Corno, A.; Oedman, C.; Botta, U.; von Segesser, L.K. Animal model to compare the effects of suture technique on cross-sectional compliance on end-to-side anastomoses. *Eur. J. Cardio-Thorac. Surg.* **1999**, *19* (4), 477–481.
47. Walpoth, B.H.; Ris, H.B.; Amonn, A.; Mettler, D.; Schilt, W.; Hoflin, F.; Felix, R.; Schaffner, T.; Althaus, U. Evaluation of recent vascular prostheses in the arteriovenous shunt. *Helv. Chir. Acta* **1990**, *57* (2), 347–350.
48. Walpoth, B.H.; Amonn, A.; Galdikas, J.; Ris, H.B.; Schaffner, T.; Hoflin, F.; Schilt, W.; Mettler, D.; Nachnbur, B.; Althaus, U. Experimental assessment of thrombogenicity in vascular prostheses before and during prostaglandin E1 treatment. *Eur. J. Vasc. Surg.* **1993**, *7* (5), 493–499.
49. Walpoth, B.H.; Rogulenko, R.; Tikhvinskaia, E.; Gogolewski, S.; Schaffner, T.; Hess, O.M.; Althaus, U. Improvement of patency rate in heparin-coated small synthetic vascular grafts. *Circulation* **1998**, *98* (suppl. 19), II-319–II-324.
50. Edelman, E.R. Vascular tissue engineering, designer arteries. *Circ. Res.* **1999**, *85* (12), 1115–1118.
51. Hoerstrup, S.P.; Zund, G.; Sodian, R.; Schnell, A.M.; Grunfelder, J.; Turina, M.I. Tissue engineering of small caliber vascular grafts. *Eur. J. Cardio-Thorac. Surg.* **2001**, *20* (1), 164–169.
52. Ratcliffe, A. Tissue engineering of vascular grafts. *Matrix Biol.* **2000**, *19* (4), 353–357.
53. Teebken, O.E.; Bader, A.; Steinhoff, G.; Haverich, A. Tissue engineering of vascular grafts. Human cell seeding of decellularised porcine matrix. *Eur. J. Vasc. Endovasc. Surg.* **2000**, *19* (4), 381–386.

54. Teebken, O.E.; Haverich, A. Tissue engineering of small diameter vascular grafts. *Eur. J. Vasc. Endovasc. Surg.* **2002**, *23* (6), 475–485.
55. Matsumura, G.; Hibino, N.; Ikada, Y.; Kurosawa, H.; Shin'oka, T. Successful application of tissue engineered vascular autografts, clinical experience. *Biomaterials* **2003**, *24* (13), 2303–2308.
56. Bujan, J.; Garcia-Honduvilla, N.; Contreras, L.; Gimeno, M.J.; Escudero, C.; Bellon, J.M.; San-Roman, J. Coating PTFE vascular prostheses with fibroblastic matrix improves cell retention when subjected to blood flow. *J. Biomed. Mater. Res.* **1998**, *39* (1), 32–39.
57. Zarge, J.I.; Gosselin, C.; Hauang, P.; Greisler, H.P. Platelet deposition on ePTFE grafts coated with fibrin glue with or without FGF-1 and heparin. *J. Surg. Res.* **1997**, *67* (1), 4–8.
58. Zarge, J.I.; Husak, V.; Huang, P.; Greisler, H.P. Fibrin glue containing fibroblast growth factor type 1 and heparin decreases platelet deposition. *Am. J. Surg.* **1997**, *174* (2), 188–192.
59. Chen, C.; Lumsden, A.B.; Ofenloch, J.C.; Noe, B.; Campbell, E.J.; Stratford, P.W.; Yianni, Y.P.; Taylor, A.S.; Hanson, S.R. Phosphorylcholine coating of ePTFE grafts reduces neointimal hyperplasia in canine model. *Ann. Vasc. Surg.* **1997**, *11* (1), 74–79.
60. Boland, E.D.; Wnek, G.E.; Simpson, D.G. Tailoring tissue engineering scaffolds using electrostatic processing techniques, a study of poly(glycolic acid). *J. Macromol. Sci.* **2001**, *38*, 1231–1243.
61. Niu, S.; Kurumatani, H.; Satoh, S.; Kanda, K.; Oka, T.; Watanabe, K. Small diameter vascular prostheses with incorporated bio-absorbable matrices. A preliminary study. *ASAIO J.* **1993**, *39* (3), 750–753.
62. Izhar, U.; Schwalb, H.; Borman, J.B.; Hellener, G.R.; Hotoveli-Salomon, A.; Marom, G.; Stern, T.; Cohn, D. Novel synthetic selectively degradable vascular prostheses, a preliminary implantation study. *J. Surg. Res.* **2001**, *95* (2), 152–160.
63. Stitzel, J.D.; Pawlowski, K.J.; Wnek, G.E.; Simpson, D.G.; Bowlin, G.L. Arterial smooth muscle cell proliferation on a novel biomimicking biodegradable vascular graft scaffold. *J. Biomater. Appl.* **2001**, *16* (1), 22–33.
64. Levenberg, S.; Golub, J.S.; Amit, M.; Itskovitz-Eldor, J.; Langer, R. Endothelial cells derived from human embryonic stem cells. *PNAS* **2002**, *99* (7), 4391–4396.
65. Kadner, A.; Hoerstrup, S.P.; Zund, G.; Eid, K.; Maurus, C.; Melnitchouk, S.; Grunenfelder, J.; Turina, M.I. A new source for cardiovascular tissue engineering, human bone marrow stromal cells. *Eur. J. Cardio-Thorac. Surg.* **2002**, *21* (6), 1055–1060.
66. Herring, M.B.; LeGrand, D.R. The histology of seeded PTFE grafts in humans. *Ann. Vasc. Surg.* **1989**, *3* (2), 96–103.
67. Sodian, R.; Lemke, T.; Fritsche, C.; Hoerstrup, S.P.; Fu, P.; Potapov, E.V.; Hausmann, H.; Hetzer, R. Tissue-engineering bioreactors: A new combined cell-seeding and perfusion system for vascular tissue engineering. *Tissue Eng.* **2002**, *8* (5), 863–870.
68. Walter, D.H.; Rittig, K.; Bahlmann, F.H.; Kirchmair, R.; Silver, M.; Murayama, T.; Nishimura, H.; Losordo, D.W.; Asahara, T.; Isner, J.M. Statin therapy accelerates endothelialization. A novel effect involving mobilization and incorporation of bone-marrow-derived endothelial progenitor cells. *Circulation* **2002**, *105* (25), 3017–3024.
69. Weatherford, D.A.; Sackman, J.E.; Reddick, T.T.; Freeman, M.B.; Stevens, S.L.; Goldman, M.H. Vascular endothelial growth factor and heparin in a biologic glue promotes human aortic endothelial cell proliferation with aortic smooth muscle cell inhibition. *Surgery* **1996**, *120* (2), 433–439.
70. Shi, Q.; Bhattacharya, V.; Hong-De Wu, M.; Sauvage, L.R. Utilizing granulocyte colony-stimulating factor to enhance vascular graft endothelialization from circulating blood cells. *Ann. Vasc. Surg.* **2002**, *16* (3), 314–320.
71. Bowlin, G.L.; Pawlowski, K.J.; Stitzel, J.D. Electrospinning of Polymer Scaffolds for Tissue Engineering. In *Tissue Engineering and Biodegradable Equivalents, Scientific and Clinical Applications*; Lewandrowsky, K., Ed.; Marcel Dekker: New York, 2002; 65–178.
72. Llavador, J.; Murasawa, S.; Kureishi, Y.; Uchida, S.; Masuda, H.; Kawamoto, A.; Walsh, K.; Isner, J.M.; Asahara, T. HMG-CoA reductase inhibitor mobilizes bone marrow-derived endothelial progenitor cells. *J. Clin. Invest.* **2001**, *108* (3), 399–405.
73. Beldi, G.; Bosshard, A.; Hess, O.M.; Althaus, U.; Walpoth, B.H. Transit time flow measurement, experimental validation and comparison of three different systems. *Ann. Thorac. Surg.* **2000**, *70* (1), 212–217.
74. Walpoth, B.H.; Beldi, G.; Bosshard, A.; Hess, O.M.; Carrel, T. Transit Time Flow Measurements, From Bench to Bedside. In *Intraoperative Graft Patency Verification in Cardiac and Vascular Surgery*; D'Ancona, G., Karamanoukian, H.L., Ricci, M., Salerno, T.A., Bergsland, J., Eds.; Futura Publishing Company: New York, 2001; 191–200.
75. He, G.W.; Yang, C.Q. Comparative study on calcium channel antagonists in the human radial artery, clinical implications. *J. Thorac. Cardiovasc. Surg.* **2000**, *119* (1), 94–100.
76. Contreras, M.A.; Quist, W.C.; Logerfo, F.W. Effect of porosity on small-diameter vascular graft healing. *Microsurgery* **2000**, *20* (1), 15–21.
77. Wang, Z.G.; Zhang, H.; Pu, L.Q.; Du, W.; Li, G.D.; Wu, J.D.; Wang, D.J.; Sha, R.Y. Can endothelial seeding enhance patency and inhibit neointimal hyperplasia? Experimental studies and clinical trial of endothelial seeded venous prostheses. *Int. Angiol.* **2000**, *19* (3), 259–269.
78. Valente, M.; Pettenazzo, E.; Di Filippo, L.; Laborde, F.; Rinaldi, S.; Thiene, G. Biodegradable polymer (D,L-lactide-caprolactone) in aortic vascular prosthesis, morphological evaluation in an animal model. *Int. J. Artif. Organs* **2002**, *25* (8), 777–782.
79. Gogolewski, S.; Walpoth, B.; Rheiner, P. Polyurethane macroporous membranes as pericardial substitutes. *Colloid Polym. Sci.* **1987**, *265*, 971–977.
80. Kambayashi, J.; Watase, M.; Itoh, T.; Kawasaki, T.; Shiba, E.; Sakon, M.; Mori, T. Blood compatibility of

- venous prosthesis made of textile or non-textile material. *Thromb. Res.* **1992**, *66* (4), 365–372.
81. Nojiri, C.; Okano, T.; Jacobs, H.A.; Park, K.D.; Mohammad, S.F.; Olsen, D.B.; Kim, S.W. Blood compatibility of PEO grafted polyurethane and HEMA/styrene block copolymer surfaces. *J. Biomed. Mater. Res.* **1990**, *24* (9), 1151–1171.
 82. Vale, B.H. Current concepts for assessing blood compatibility, small diameter vascular prostheses. *J. Biomater. Appl.* **1987**, *2* (1), 149–159.
 83. Walpoth, B.H.; Pavlicek, M.; Celik, B.; Nicolaus, B.; Schaffner, T.; Althaus, U.; Hess, O.M.; Carrel, T.; Morris, R.E. Prevention of neointimal proliferation by immuno-suppression in synthetic vascular grafts. *Eur. J. Cardio-Thorac. Surg.* **2001**, *19* (4), 487–492.
 84. Beringer, J.P.; Terrettaz, S.; Michel, R.; Tirelli, N.; Vogel, H.; Textor, M.; Hubbell, J.A. Chemisorbed poly (propylene sulphide)-based copolymers resist biomolecular interactions. *Nat. Mater.* **2003**, *2* (4), 259–264.
 85. Chanda, J.; Kuribayashi, R.; Abe, T. Heparin coupling in inhibition of calcification of vascular bioprosthesis. *Biomaterials* **1999**, *20* (19), 1753–1757.
 86. Shigematsu, K.; Zasuhrara, H.; Shigematsu, H. Topical application of antiangiogenic agent AGM-1470 suppresses anastomotic intimal hyperplasia after e-PTFE grafting in a rabbit model. *Surgery* **2001**, *129* (2), 220–230.
 87. Torem, S.; Schneider, P.A.; Paxton, L.D.; Yasuda, H.; Hanson, S.R. Factors influencing acute thrombus formation on carotid artery vascular grafts. *ASAIO Trans.* **1988**, *34* (4), 916–920.
 88. Helisch, A.; Schaper, W. Arteriogenesis, the development and growth of collateral arteries. *Microcirculation* **2003**, *10* (1), 83–97.
 89. Hoffmann, R.; Mintz, G.S.; Dussaillant, G.R.; Popma, J.J.; Pichard, A.D.; Satler, L.F.; Kent, K.M.; Griffin, J.; Leon, M.B. Patterns and mechanisms of in-stent restenosis. A serial intravascular ultrasound study. *Circulation* **1996**, *94* (6), 1247–1254.
 90. Hardhammar, P.A.; van Beusekom, H.M.; Emanuelsson, H.U.; Hofma, S.H.; Albertsson, P.A.; Verdouw, P.D.; Boersma, E.; Serruys, P.W.; van der Giessen, W.J. Reduction in thrombotic events with heparin-coated Palmaz-Schatz stents in normal porcine coronary arteries. *Circulation* **1996**, *93* (3), 423–430.
 91. Sousa, J.E.; Costa, M.A.; Sousa, A.G.; Abizaid, A.C.; Seixas, A.C.; Abizaid, A.S.; Feres, F.; Mattos, L.A.; Falotico, R.; Jaeger, J.; Popma, J.J.; Serruys, P.W. Two-year angiographic and intravascular ultrasound follow-up after implantation of sirolimus-eluting stents in human coronary arteries. *Circulation* **2003**, *107* (3), 381–383.
 92. Kastrati, A.; Schomig, A.; Dirschinger, J.; Mehilli, J.; von Welsch, N.; Pache, J.; Schühlen, H.; Schilling, T.; Schmitt, C.; Neumann, F.J. Increased risk of restenosis after placement of gold-coated stents: Results of a randomized trial comparing gold-coated with uncoated steel stents in patients with coronary artery disease. *Circulation* **2000**, *101* (21), 2478–2483.
 93. Windecker, S.; Mayer, I.; De Pasquale, G.; Maier, W.; Dirsch, O.; De Groot, P.; Wu, Y.P.; Noll, G.; Leskosek, B.; Meier, B.; Hess, O.M.; Working Group on Novel Surface Coating of Biomedical Devices (SCOL). Stent coating with titanium-nitride-oxide for reduction of neointimal hyperplasia. *Circulation* **2001**, *104* (8), 928–933.
 94. Grigoni, K.S.; Windecker, S.; Hubbell, J.A. Stent coating for prevention of in-stent restenosis: Biocompatible polymer versus inert metal. (Submitted).

Kidneys, Artificial

Churn K. Poh

Department of Mechanical Engineering, University of Kentucky, Lexington, Kentucky, U.S.A.

Junfeng Lu

Center for Biomedical Engineering, University of Kentucky, Lexington, Kentucky, U.S.A.

William R. Clark

Clinical Strategy and Therapy Development, Gambro Renal Products (Intensive Care), Indianapolis, Indiana, U.S.A.

Dayong Gao

Department of Mechanical Engineering and Center for Biomedical Engineering, University of Kentucky, Lexington, Kentucky, U.S.A.

INTRODUCTION

Artificial kidneys or hemodialyzers play an important role in saving the life of patients with kidney failures. The reported incident rates of patients with end-stage renal disease (ESRD) worldwide have increased steadily since 1989, and the U.S.A. continues to have the highest incident rates according to the U.S. Renal Data System 2000 Annual Data Report. The number of ESRD patients in the U.S.A. had increased significantly over the past 20 years, and it is projected to grow exponentially over the next 10 years. This number is projected to increase to 661,330 in 2010, an increase of 78% compared to 372,407 in 2000. Among these ESRD patients, those undergoing dialysis therapy are projected to increase to 520,240 in 2010 with an increase of 85% compared to 281,355 in 2000. In addition, the total annual cost of Medicare for ESRD patients in the U.S.A. is projected to be \$28.3 billion in 2010, an increase of 99% compared to \$14.2 billion in 2000.

In this article, we will discuss the fundamentals of artificial kidneys from the engineering perspective. We will first discuss the basic functions of human kidneys, classification of uremic solutes, basic designs and classification of artificial kidneys, property and classification of dialysis membranes, mass transfers in artificial kidneys, parameters used to characterize performance of artificial kidneys, and finally different forms of renal replacement therapies for acute renal failure (ARF). Because it is spatially impossible to discuss in depth the topic of artificial kidneys in solely one article, we encourage the readers to refer to the "Articles of Further Interest," "References," and "Bibliography" cited at the end of this article for further readings.

HUMAN KIDNEYS

Kidneys (Fig. 1) remove uremic solutes and excess plasma water from our blood and excrete them as urine. Healthy kidneys process about 190 L of blood per hour and produce approximately 1–2 L of urine per day.^[1] In addition to excretion of metabolic wastes and water, kidneys also regulate the acid–base balance in our blood and release three important hormones: erythropoietin, renin, and calcitriol. Erythropoietin stimulates our bones to generate red blood cells. Renin regulates our blood pressure. Calcitriol stimulates calcium reabsorption at the tubule and hence helps maintain a proper calcium level in our body.

There are three types of kidney failures: acute, chronic, and end-stage renal failures. End-stage renal failure is commonly called ESRD. Acute renal failure (ARF) is a sudden drop of the renal function caused by a sudden injury to the kidneys such as trauma and drugs or alcohol poisoning. The ARF may be counteracted if the kidneys are not seriously damaged. Chronic renal failure, on the other hand, is a gradual irreversible loss of the renal function. The ESRD is the final stage of kidney failure.

Kidney failure is said to occur when the kidneys lost 80–90% of their renal function. Most people remain healthy after having lost 50% of their renal function.^[1] Kidney failure, if untreated, the accumulation of uremic solutes in the blood can cause uremia and eventually lead to death; therefore, all kidney-failure patients must receive either renal replacement therapy or kidney transplantation to survive. An excellent discussion on the human kidneys and their diseases can be found in Refs. [2–6].

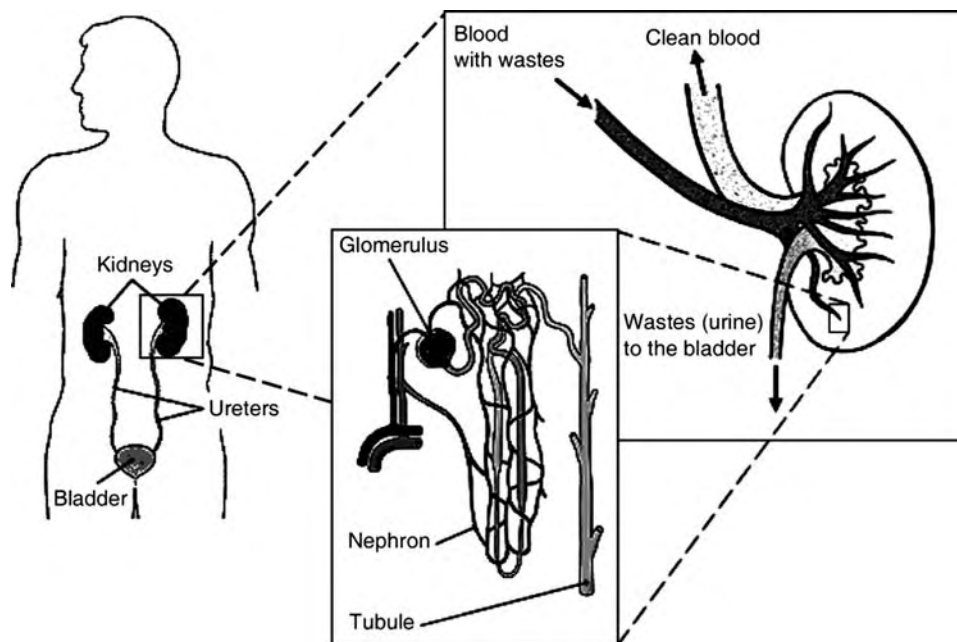


Fig. 1 The structure of human kidneys. Each kidney contains about one million nephrons where blood purification takes place. The excess plasma water and metabolic wastes are filtered from the blood-carrying capillaries in the nephrons and are sent to the bladder as urine through the ureters.

CLASSIFICATION OF UREMIC SOLUTES

Uremic solutes can be divided into four groups^[7] based on their molecular weights. The so-called “small solutes” have molecular weights less than 300 Daltons (Da), “middle molecules” have molecular-weights between 500 and 5000 Da, “low molecular weight protein” have molecular weights between 5000 and 50,000 Da, and “large proteins” have molecular weights greater than 50,000 Da. A Dalton is a unit of molecular weight. Examples of small solutes are urea (60 Da) and creatinine (113 Da), which are byproducts of protein metabolism and are commonly measured to estimate kidney health; low molecular-weight proteins are β_2 -microglobulin (11,818 Da) and parathyroid hormone (9425 Da); and large protein is albumin (66,000 Da). This classification scheme will be used throughout this article.

History Perspective of Uremic Solutes

In the 1960s, the retention of small solutes in the human blood was believed to be the main cause of uremic syndrome. At that time, conventional hemodialysis mediated by diffusive transport of small solutes was claimed to be adequate for treating kidney-failure patients. In addition, some researchers^[8,9] claimed that larger solutes are unimportant to uremic syndrome, and they therefore emphasized on the clearance of small solutes as the measurement of

adequacy of hemodialysis. In contrast, other researchers^[10,11] found discrepancy between urea (small solute) clearance and uremic syndrome.

In the early 1970s, the Middle Molecule hypothesis^[12,13] was proposed, and it resulted in the classification “middle molecule.” The Middle Molecule hypothesis states that uremic syndrome is at least in part caused by the retention of middle molecules in the blood.

In the 1980s, the discrepancy between urea clearance and uremic syndrome, and the Middle Molecule hypothesis, however, were ignored when experimental results from the National Cooperative Dialysis Study (NCDS) showed that the removal of urea was correlated to the morbidity and mortality of hemodialysis patients.^[14,15] Despite the NCDS results, Rockel et al.^[16,17] found elevated concentration of low molecular-weight proteins, which normally undergo catabolism in the kidneys, in chronic hemodialysis patients. Gejyo et al.^[18] and Gorevic et al.^[19,20] discovered that the elevated concentration of β_2 -microglobulin, a middle molecule, leads to systemic amyloidosis in long-time hemodialysis patients. The β_2 -microglobulin is the first middle molecule to be linked to a specific clinical syndrome occurring exclusively in hemodialysis patients.^[21]

In the 1990s, Vanholder et al.^[22] revisited the Middle Molecule hypothesis and confirmed that the uremic syndrome is at least in part caused by the retention of middle molecules in the human blood. In addition, Anderstam et al.^[23] showed evidence that suggests that

the retention of middle molecules in the blood causes uremic appetite suppression, where they found that the middle molecules isolated from the uremic ultrafiltrate inhibit ingestive behavior in rats. All these findings have shown the importance of the middle molecules in uremic syndrome.

Most recently, a more sophisticated classification scheme of uremic solute has been proposed by the European Uremic Toxicity Working Group, which was initiated by the European Society for Artificial Organs in 1999. Details of that classification scheme can be found in Refs. [24,25]. A review on the pathophysiologic effects of uremic solutes has been done by Vanholder and De Smet.^[26]

ARTIFICIAL KIDNEYS

Artificial kidneys or hemodialyzers are used in extracorporeal renal therapies for the removal of uremic solutes and excess plasma water from the blood of patients with kidney failure. The terms *artificial kidney* and *hemodialyzer* will be used interchangeably throughout this article.

Historical Perspective of Artificial Kidneys

The first scientific description of hemodialysis principle was published by Graham^[27] in 1854, and the first prototype hemodialyzer (Fig. 2) was developed by

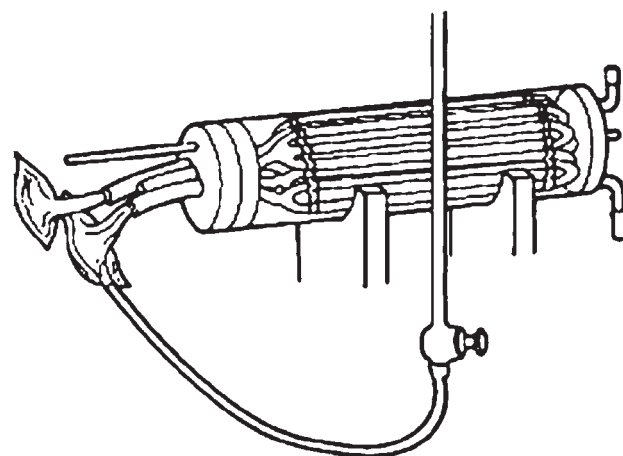


Fig. 2 The first prototype hemodialyzer developed by Abel et al. in 1914. (From Ref. [28].)

Abel, Rowntree, and Turner^[29] in 1913. The hemodialyzer was cylindrical in shape with an inner diameter of 10 cm and a length of 50 cm. It contained 30 collodion tube membranes with an inner diameter of 8 mm, with hirudin as the anticoagulant and saline as the dialysate. Abel et al. performed the first historical *in vivo* hemodialysis on a dog and succeeded in removing sodium salicylate administered to the dog from its body fluid. They referred to the process as *vividiffusion*.

The first hemodialyzer used on human was reported by Haas^[30] in 1923. Haas also used collodion tube membranes. He used hirudin as anticoagulant but later

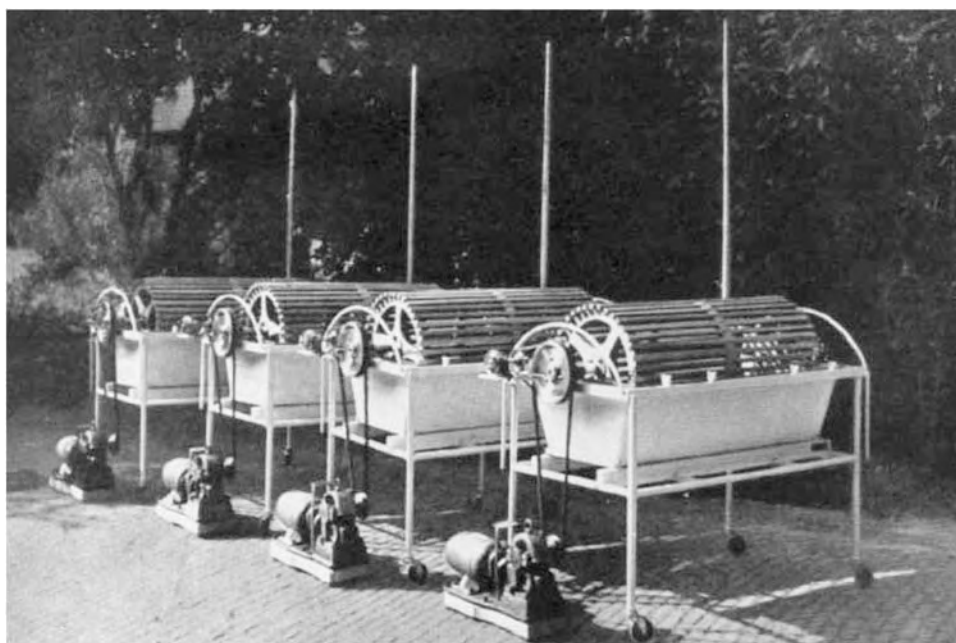


Fig. 3 The rotating-drum artificial kidney developed by William J. Kolff that, for the first time, succeeded in treating a patient with ARF in 1945. (From Ref. [33].)

switched to heparin and built several models and sizes of hemodialyzers. None of Haas' patients, however, survived.

In 1945, Kolff^[31,32] developed the rotating-drum artificial kidney (Fig. 3) and succeeded in treating a patient with ARF for the first time. Kolff spirally wrapped a 30-m cellophane tube with an inner diameter of 25 mm around a wooden drum. The drum was rotated and half immersed in a 100-L stationary dialysate bath. Kolff^[34] later improved his rotating-drum artificial kidney by developing a twin-coil disposable artificial kidney (Fig. 4) in 1956.

In 1960, a plate-and-frame hemodialyzer called the Kiil dialyzer was developed by Kiil.^[35] It was made of multiple flat sheets of cellophanes sandwiched between plastic boards having multiple grooves. The Kiil dialyzer has a total membrane surface area of 1.8 m² and extracorporeal volume of 700 mL.

In 1964, the hollow-fiber hemodialyzer called the capillary kidney was first proposed by Stewart, Cerny, and Mahon.^[36] The capillary kidney was very similar to the prototype hemodialyzer developed by Abel, Rowntree, and Turner^[29] except that it has a much smaller inner diameter of hollow-fiber membrane. Hollow-fiber hemodialyzers are the most widely used hemodialyzers today.

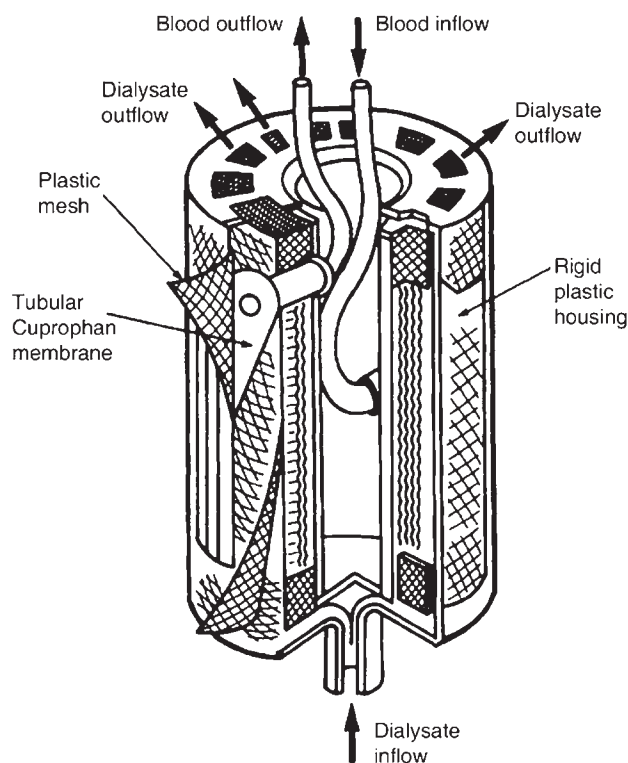


Fig. 4 A twin-coil disposable artificial kidney developed by William J. Kolff in 1956 (From Ref. [28].)

Hollow-Fiber Artificial Kidneys

A typical hollow-fiber hemodialyzer contains between 6000 and 12,000 hollow fibers, depending on the size of the hemodialyzer. These hollow fibers have an inner diameter of about 200 μm and a wall thickness between 15 and 50 μm . These hollow fibers are made of semi-permeable membranes, and they act as a semipermeable barrier for the mass transfer of uremic solutes and excess plasma water from the blood to the dialysate. These hollow fibers are potted with either polyurethane or epoxy at both ends of the hemodialyzer to form tubesheets. Hollow-fiber hemodialyzers are available in different sizes (about 2–5 cm in diameter and 15–30 cm in length), and the choice of the size of hemodialyzer depends upon the size of the dialysis patient, where a large dialysis patient requires a large hemodialyzer.

From an engineering perspective, a hollow-fiber hemodialyzer is a countercurrent mass exchanger. The hollow-fiber design is preferred because it has a higher surface area per unit blood volume, thereby maximizing the overall mass-transfer surface area compared with other designs such as flat sheet and spiral wound. A hollow-fiber hemodialyzer has two compartments: blood and dialysate compartments. Its engineering design (Fig. 5) is very similar to a shell-and-tube heat exchanger. The tube side (inside the lumen of hollow fibers) is called the blood compartment, and the shell side (outside the lumen of hollow fibers) is called the dialysate compartment. The blood and dialysate are introduced countercurrently through the hemodialyzer to optimize the mass transfer of solutes by enhancing their concentration difference across the hollow-fiber membrane. The typical blood flow rates used in the clinical settings are between 200 and 400 mL/min, and the dialysate flow rates are between 500 and 800 mL/min, depending upon the size and physical conditions of the dialysis patient and the type of renal replacement therapy.

The geometrical designs of hemodialyzers play a critical role in regulating the flow distributions in the blood and dialysate compartments, and hence in determining their performance in removing uremic solutes. Ideally, the flow distributions in both the blood and dialysate compartments should be spatially uniform to ensure optimal removal of uremic solutes from the blood to the dialysate. Any mismatch of the blood and dialysate flows or the presence of shunt flow in either compartment can impair the removal efficacy of uremic solutes. In practice, it is very difficult to introduce uniform flow distribution throughout the dialysate compartment because of the presence of thousands of hollow fibers that create a complex pathway for the dialysate. We have studied the flow distribution in hollow-fiber hemodialyzers extensively using magnetic resonance imaging.^[37–41] Fig. 6 shows

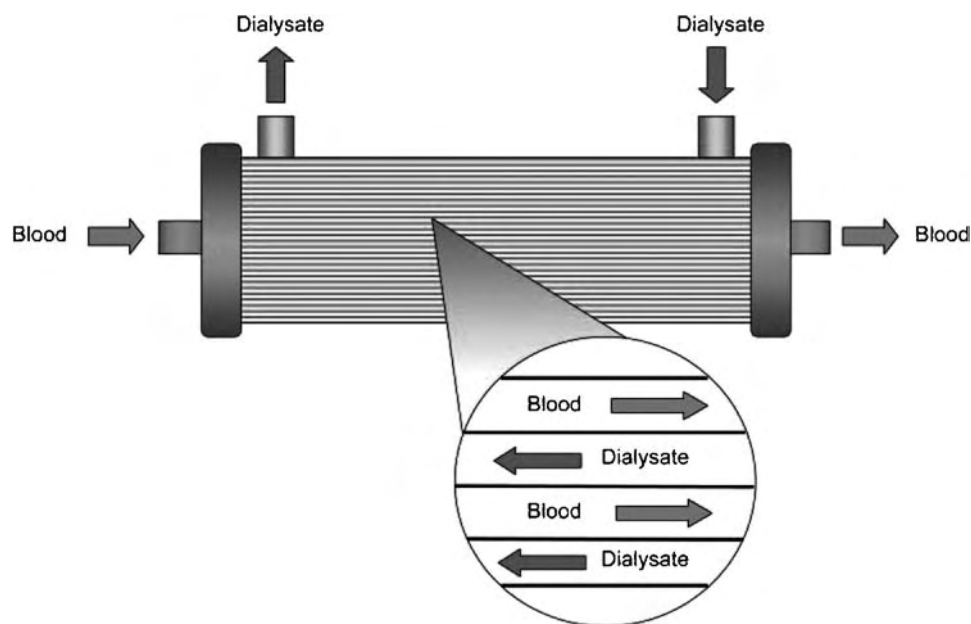


Fig. 5 A general schematic drawing of an artificial kidney. The shape and size of an artificial kidney vary from different manufacturers.

an example of dialysate flow distribution at the middle cross-section of a polyamide hemodialyzer^[39] containing 10,000 hollow fibers, each with an inner diameter of 215 μm and a wall thickness of 50 μm . The velocity was imaged in the axial direction, and the dialysate flow rate was 600 mL/min. The dialysate flow distribution of this hemodialyzer was nonuniform where channeling flows occurred at the periphery. The channeling flows were caused by the uneven spacing of hollow fibers in the hemodialyzer. We have also shown that the design of flow baffle in the dialysate compartment affects significantly the dialysate flow distribution.^[40] Fig. 7 shows the designs of two different

flow baffles, and Fig. 8 shows their corresponding dialysate flow distributions near the inlet and the outlet of dialysate ports at dialysate flow rates between 200 and 1000 mL/min.

CLASSIFICATION OF ARTIFICIAL KIDNEYS

Artificial kidneys can be classified based upon the transport property of their membranes or their uses in dialysis therapy. The classification schema, nevertheless, remain somewhat debatable. Some researchers classify hemodialyzers based on their water

POLYAMIDE HEMODIALYZER Dialysate Compartment

Flow Rate: 600 mL/min

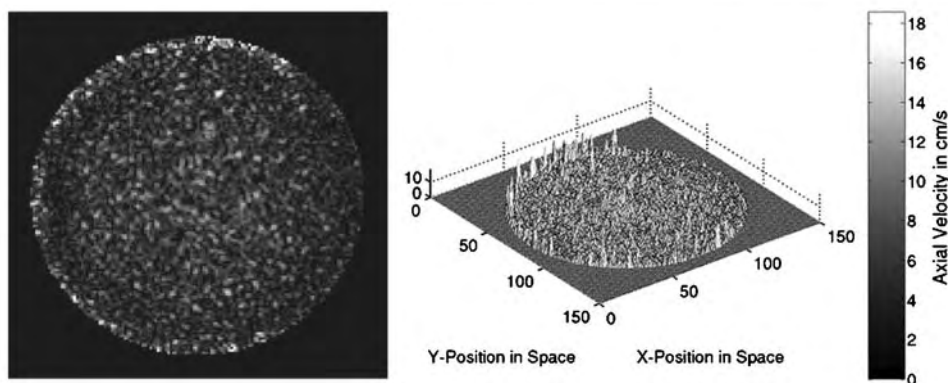


Fig. 6 Dialysate-side velocity distribution at the middle cross-section of a polyamide hemodialyzer containing 10,000 hollow fibers. (From Ref. [39].)

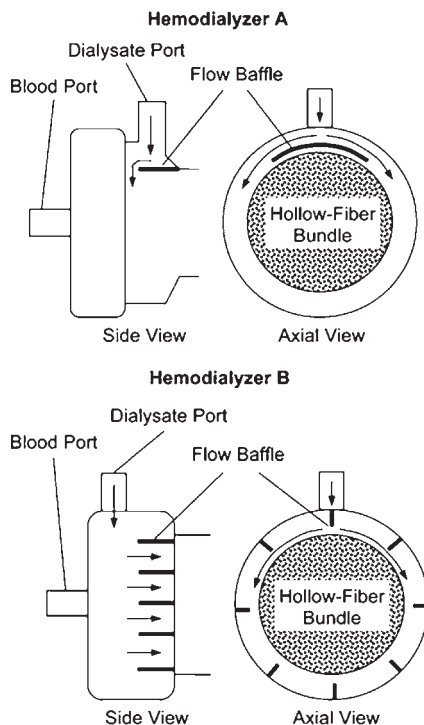


Fig. 7 Pictorial representation of two different flow baffles in hemodialyzers A and B. (From Ref. [40].)

permeabilities, while others classify hemodialyzers based on their solute permeabilities. The terms *low flux*, *high flux*, and *high efficiency* are most widely used in the literature. In the most general sense, we may associate *low flux* with hemodialyzers that remove small solutes, *high flux* with hemodialyzers that remove middle molecules and low molecular-weight proteins, and *high efficiency* with hemodialyzers that either require short dialysis treatment time or have a large membrane area.

Low Flux and High Flux

Some researchers classify hemodialyzers based on their water permeabilities, where low and high fluxes are associated with low and high ultrafiltration coefficients, respectively; whereas, other researchers classify hemodialyzers based on their solute permeabilities, where low and high fluxes are associated with low and high solute permeabilities, respectively. The classification scheme based on ultrafiltration coefficient has also been reported, and it remains debatable^[42] because different sizes of hemodialyzers made from the same membrane can be classified as either low or high flux depending on the total surface area of their membranes; for instance, for a given type of membrane and transmembrane pressure, a hemodialyzer with a large membrane area has a higher ultrafiltration

coefficient and hence a higher ultrafiltration rate compared with a hemodialyzer with a small membrane area. This classification scheme would be more useful in comparing hemodialyzers if we normalized the ultrafiltration coefficient with respect to the membrane area.

Subsequently, in 1991, the Center for Disease Control and Prevention (CDC) defined a high-flux hemodialyzer^[43] as having an ultrafiltration coefficient per unit membrane area equal to or greater than $20 \text{ mL/hr/mmHg/m}^2$. In 1995, the Hemodialysis (HEMO) Study,^[42,44] nonetheless, defined low-flux and high-flux hemodialyzers as having β_2 -microglobulin clearance less than 10 and greater than 20 mL/min , respectively. A classification of high-flux hemodialyzers as having hydrophobic membrane, high mass-transfer-area coefficient k_0A of urea, and ultrafiltration coefficient per unit membrane area greater than $10 \text{ mL/hr/mmHg/m}^2$ has also been reported in the literature.^[45]

Despite the above classifications, the only classification scheme recognized by the U.S. Food and Drug Administration is based on the water permeability,^[46] with low permeability hemodialyzers having ultrafiltration coefficients of less than 8 mL/hr/mmHg and high permeability hemodialyzers having ultrafiltration coefficients of equal to or greater than 8 mL/hr/mmHg .

We would also like to point out that the water permeability of dialysis membrane is not necessarily correlated to its solute permeability.^[46] Consider two different membranes A and B of equal area and porosity. Membrane A has a large number of small pores, and B has a small number of large pores. Although the diffusive transport properties of small solutes through these membranes would be equal, hence equal permeability of small solutes, because of equal porosity; the water permeability in membrane B would be greater than that in A because the ultrafiltrate flow, hence water permeability, is proportional to the fourth power of the membrane's pore radius.^[47]

High Efficiency

The term *high efficiency* is first used by Keshaviah et al.^[48] during the 1980s to describe the type of dialysis therapy rather than the type of hemodialyzer. These researchers showed that a significant reduction in treatment time with a high clearance of urea is possible by using a high blood flow rate and a cellulose-acetate hemodialyzer with a large membrane area. Subsequently, others have started using the term *high efficiency* in the literature with reference to hemodialyzers having a short dialysis treatment time or a large membrane area.

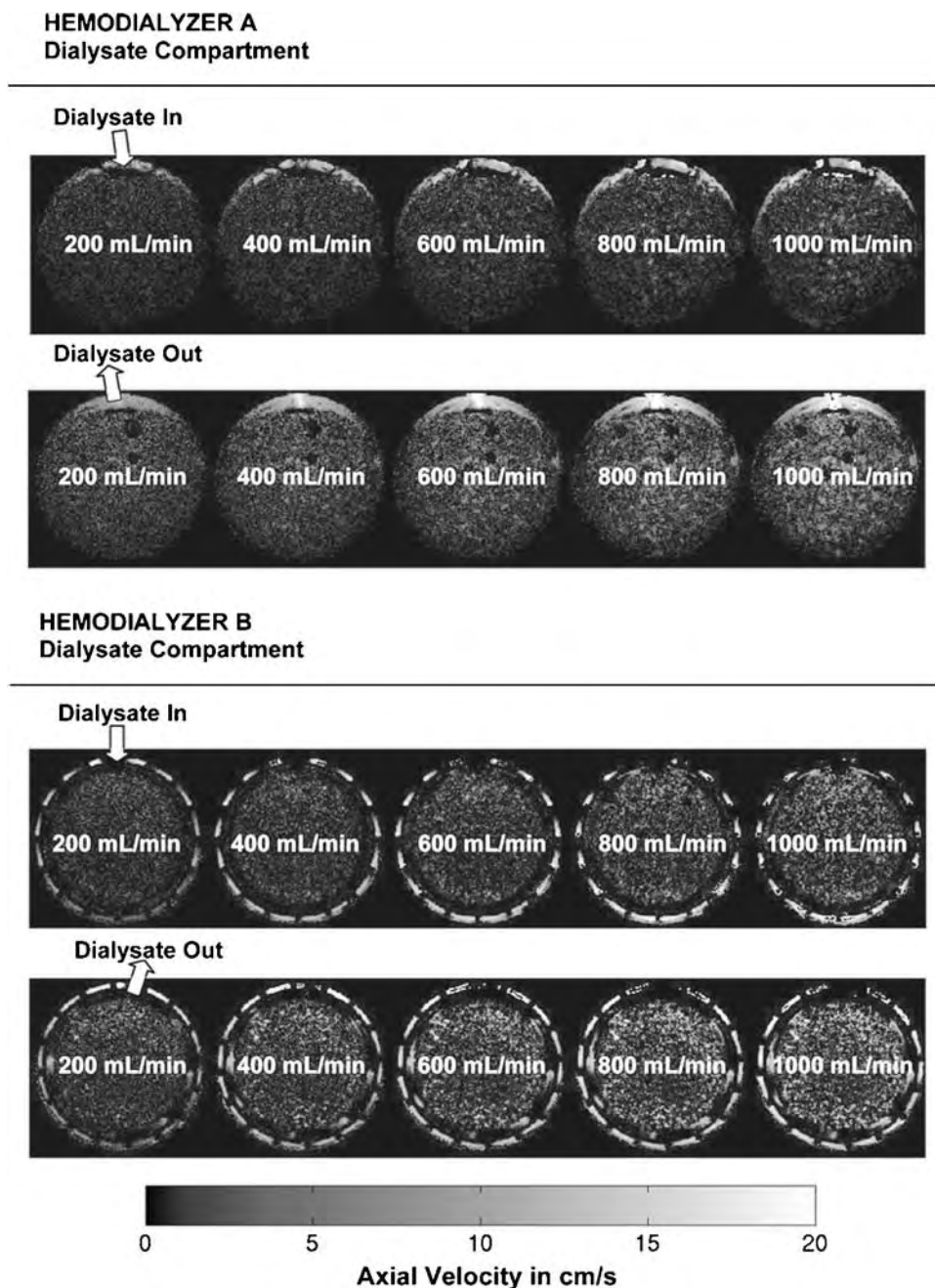


Fig. 8 Dialysate-side velocity distributions at the dialysate ports of hemodialyzers A and B, each with different designs of flow baffles. (From Ref. [40].)

A high-efficiency hemodialyzer^[45] was also referred to as having either an urea clearance greater than 150 mL/min or an urea clearance greater than 200 mL/min and a mass-transfer-area coefficient k_0A of urea greater than 450 mL/min. Other researchers have defined a high-efficiency hemodialyzer based on urea clearance per body weight^[49] and on urea clearance per distribution volume.^[50]

In 1991, the CDC defined a high-efficiency hemodialyzer^[43] as having an ultrafiltration coefficient

between 10 and 19 mL/hr/mmHg. This definition distinguishes a hemodialyzer that needs an ultrafiltration control system from the one that does not.

The term *high efficiency*, however, has subsequently become less apparent in the literature because the current goal of dialysis treatment^[45] has shifted from shortening treatment time to optimizing solute clearances of middle molecules and low molecular-weight proteins. Retrospective data have shown that mortality increases in dialysis patients with treatment time less

than 3.5 hr.^[51] A longer dialysis time is also claimed to improve blood pressure regulation in dialysis patients.^[52]

Another interesting classification scheme, which is notably mentioned, has subsequently been proposed by Akizawa, Kinugasa, and Ideura.^[53] They classify hemodialyzers by performance into five types: standard low flux, high flux, highly permeable, adsorptive, and high performance. They claimed that many other proposals and discussions are required before determining the final classification.

DIALYSIS MEMBRANES

Dialysis membranes act as a semipermeable barrier for the removal of uremic solutes and excess plasma water from the blood to the dialysate. Dialysis membranes vary in morphology and chemical composition. The morphology determines the transport properties of uremic solutes, while the chemical composition determines the biocompatibility of the membrane. An ideal dialysis membrane should be capable of selectively removing a broad range of uremic solutes comparable to those removed by the native kidneys and should not induce any adverse reactions between the membrane and the blood or any administered drugs during dialysis therapy. Equally important, the physicochemical properties of an ideal dialysis membrane should be resistant to change under different operating conditions and during sterilization.

Biocompatibility

The discovery that pulmonary leukostasis may have been resulted from the complement activation by cellulosic membranes in the 1970s^[54] has opened a new dimension to the concept of biocompatibility in dialysis membranes. Complement activation is the sequential activation of serum components C1 through C9, initiated by either an erythrocyte-antibody complex or microbial polysaccharides owing to the presence of microbes or other foreign materials in the blood, resulting in an inflammatory response. The complement system, once activated, fires in a cascade-like fashion in which one component activates the next, and it is an adjunct to the antibody system. The extent of complement activation depends on the membrane material and is used as a parameter for the assessment of membrane biocompatibility,^[55] for instance, the plasma concentration of terminal C5b-9 complement complex can be used for the quantification of complement activation and hence as a marker of biocompatibility of the membrane.^[56,57]

Another important biocompatibility aspect of the dialysis membrane is the adsorption of proteins on the membrane surface, also known as protein fouling. The adsorption of proteins increases the transmembrane resistance to fluid and solute transport, and hence decreases the water and solute permeabilities of the membrane. Protein adsorption is usually seen on the surface of hydrophobic membrane.^[58] Morti and Zydney^[59,60] have shown that the protein layer has a significant effect on the clearance of uremic solutes with molecular weights greater than 10,000 Da in cellulose-triacetate, polyacrylonitrile, and polysulfone hemodialyzers. They have also shown that the thickness of the protein layer increases with increasing dialysis time, ranging from 0.25 μm after 40 min to 0.86 μm after 180 min. Although protein adsorption on the membrane surface has an adverse effect on its water and solute permeabilities, the ability of a membrane to adsorb specific proteins can be advantageous for removing harmful low molecular-weight proteins such as β_2 -microglobulin from the blood, which will be discussed later.

CLASSIFICATION OF DIALYSIS MEMBRANES

We can classify dialysis membrane into three categories based on their polymeric compositions and preparation processes: regenerated cellulosic, modified cellulosic, and synthetic membranes. The effect of membrane composition and structure on the solute removal and biocompatibility in hemodialysis has been discussed by Clark, Hamburger, and Lysaght.^[46]

Regenerated Cellulosic Membranes

Regenerated cellulosic membranes are the first generation of dialysis membranes with low water and solute permeabilities (small pore size) and a relative thin wall thickness (5–15 μm). They are made from cellulose, a naturally occurring substance, and are structurally symmetric. These membranes are highly hydrophilic because of the presence of a large number of free hydroxyl groups on their cellulose monomers. The hydrophilic nature of these membranes promotes complement activation in dialysis patients because of the interaction of complement cascade products with the hydroxyl groups found in the membranes.^[46,61]

Apart from complement activation, regenerated cellulosic membranes offer great permeabilities on small solutes, but poor permeabilities on middle molecules and low molecular-weight proteins. An example of regenerated cellulosic membrane is Cuprophane[®] (Membrana GmbH, Wuppertal, Germany). Cuprophane is still being used in more than 50% of all

hemodialyzers throughout the world.^[62] In the U.S.A., however, the use of Cuprophane in hemodialyzers had declined from approximately 65% in 1990 to less than 20% in 1997.^[63] The popularity of Cuprophane has decreased over the past years because of its lack of biocompatibility and poor clearances on middle molecules and low molecular-weight proteins.

Modified Cellulosic Membranes

In an effort to improve the biocompatibility by decreasing the degree of complement activation in cellulosic membranes, modified cellulosic membranes were introduced in the 1980s. Modified cellulosic membranes are also made from cellulose. They are sometimes called substituted cellulosic membranes because parts or all of the hydroxyl groups on the cellulose monomers have been substituted by chemical groups such as the acetyl and benzyl groups. They have greater permeabilities of water and middle molecules (larger pore size) compared with the regenerated cellulosic membranes, but poor permeabilities of low molecular-weight proteins.

Examples of substituted cellulosic membrane modified by acetylation are cellulose acetate (substituted by one acetyl group), cellulose diacetate (substituted by two acetyl groups), and cellulose triacetate (substituted by three acetyl groups). These substituted cellulosic membranes have symmetric structure and relatively thin wall thickness. As opposed to the hydroxyl group, the acetyl group does not bind avidly to the C3 molecule to initiate activation of the complement cascade,^[64] consequently, complement activation and leukopenic response are attenuated.

Hemophan[®] (Membrana GmbH, Wuppertal, Germany) is another commonly used substituted cellulosic membrane. The diethylaminoethyl (DEAE) groups are used as the substituent on the cellulose monomers, and only a small percentage (less than 5%) of the hydroxyl groups are substituted^[64] in Hemophan. The DEAE groups are bulky and effective in shielding the hydroxyl groups through steric hindrance, hence attenuating the degree of complement activation and leukopenic response.

SMC[®] (synthetically modified cellulose), on the other hand, is made from cellulose where parts of the hydroxyl group on the cellulose monomer are substituted by etherification with the benzyl group. They are marketed under the name SMC (Membrana GmbH, Wuppertal, Germany).

The Hemophan membrane has been shown experimentally to be more biocompatible than the cellulose acetate, followed by the Cuprophane.^[65] The SMC membrane has been shown experimentally to be more biocompatible than the Cuprophane.^[66]

Synthetic Membranes

Synthetic membranes are new generation of membranes that are still under development. These membranes are developed in response to the pronounced complement activation and poor removal of middle molecules and low molecular-weight proteins associated with unmodified cellulosic membranes. Synthetic membranes are made from polymers such as polyacrylonitrile, polyamide, polymethylmethacrylate, polysulfone, and polyethersulfone. The use of polysulfone membranes in hemodialyzers in the U.S.A. had increased significantly from approximately 25% in 1990 to approximately 75% in 1997.^[63]

Synthetic membranes have different morphologies and various pore sizes depending on their manufacturing processes and polymeric compositions. Synthetic membranes have greater wall thickness (at least 20 μm) compared with cellulosic membranes and may be structurally symmetric or asymmetric. Asymmetric synthetic membranes usually have two or more layers with different porosities and pore-size distributions. The thin layer has small nominal pore size and acts as an active filtering layer; the thick layer has large nominal pore size and acts as a porous support layer for the thin layer. In addition, the porous support layer sometimes acts as an adsorption site to remove certain uremic solutes that are not removed by the thin layer.^[67] An example of a cross section of a synthetic hollow fiber^[41] is shown in Fig. 9. This hollow fiber has an asymmetric finger-like cross-sectional structure with three distinct layers combined with a microdomain structure at the inner surface. The outer and inner layers have a thickness of 0.1 and 0.5 μm , respectively.

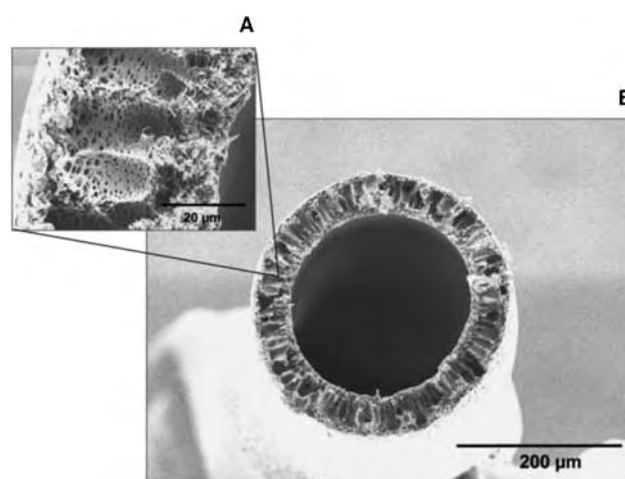


Fig. 9 Scanning electron micrographs of a polyamide hollow fiber at (A) 2000 \times and (B) 200 \times magnifications. These micrographs show the unique morphology of the hollow fiber. (From Ref. [41].)

The intermediate finger-like structure has a thickness of 45–50 μm . The microdomain inner surface is created by inserting hydrophilic areas into the hydrophobic polymer to create a balanced combination of hydrophilicity and hydrophobicity. A balanced hydrophilic–hydrophobic membrane has an excellent biocompatibility for use in hemodialyzers as discussed below.

In general, hydrophilic membranes promote complement activation as discussed earlier, and hydrophobic membranes promote adsorption of proteins on their surfaces. Conventional synthetic membranes made from polyamide, polysulfone, and polyethersulfone are highly hydrophobic, and thereby their solute permeabilities are highly affected by the adsorption of proteins on their membrane surfaces. A hydrophilic additive, polyvinylpyrrolidone, is added during the manufacturing process to create a balance between hydrophilicity, hydrophobicity and hence a balance between complement activation and proteins adsorption in these membranes. A balance between complement activation and proteins adsorption improves the overall biocompatibility of these membranes. The discussion on the biocompatibility advantage of membranes having a balanced combination of hydrophilicity and hydrophobicity can be found in Refs.^[66,68] Hakim et al.^[69] have demonstrated that the relative risk of mortality of patients dialyzed with synthetic or substituted cellulosic membranes was at least 25% less than that of the patients dialyzed with unsubstituted cellulosic membranes. Vitamin E-bonded dialysis membranes with improved biocompatibility have also been reported in the literature.^[70–80]

HOLLOW-FIBER MEMBRANES

Although clear differences exist among different types of membrane, a number of common features are evident in hollow-fiber membranes. These hollow fibers typically have inner diameter between 180 and 220 μm , wall thickness between 15 and 50 μm , and length between 15 and 30 cm. The dimensions of hollow fibers are dictated by the operating conditions required during dialysis therapy and are result of a compromise between opposing forces.^[64] A small inner diameter of hollow fiber, on one hand, is preferred because it provides a short diffusive distance for removal of uremic solutes. For a given blood flow rate, a smaller inner diameter also provides a larger shear rate that results in greater attenuation of the effect of blood-side boundary layer.^[81] A small inner diameter, on the other hand, has an undesirable effect that induces large resistance to the blood flow as discussed below. We can model hollow fibers as nonporous cylindrical tubes, and the blood flow inside them is governed by the

Hagen–Poiseuille equation^[82]

$$Q_B = \frac{N\pi r^4 \Delta P}{8\mu L} \quad (1)$$

where Q_B is the blood flow rate, N is the number of hollow fibers, r is the radius of hollow fibers, ΔP is the axial pressure drop, μ is the blood viscosity, and L is the length of hollow fibers. We can rearrange Eq. (1) into

$$Q_B = \frac{N\Delta P}{R} \quad \text{where } R = \frac{8\mu L}{\pi r^4} \quad (2)$$

R is the resistance to the blood flow. Eq. (2) shows that R is inversely proportional to the fourth power of r ; hence, a small decrease in inner diameter of hollow fibers causes a large increase in resistance to the blood flow. Eq. (2) also shows that an increase either in hollow-fiber length or hematocrit level (increasing hematocrit level increases blood viscosity) increases the resistance to the blood flow. To maintain the blood flow rate with increasing resistance to the blood flow for a given number of hollow fibers, the axial pressure drop increases, as seen from Eq. (2), and causes significant backfiltration of dialysate under normal operating conditions.^[83] Backfiltration increases the likelihood of transfer of cytokine-inducing substances from contaminated dialysate to the blood.^[84]

MEMBRANE PORE AND PORE-SIZE DISTRIBUTION

As an approximation, we can view the membrane as having straight cylindrical pores perpendicular to the membrane surface^[85] as shown in Fig. 10. We assume all pores have the same radius R_p , which is larger than the stoke radius R_s of a hypothetical solute. The ultrafiltrate (convective) flow through these straight cylindrical pores is governed by the Hagen–Poiseuille equation [Eq. (1)] and is proportional to the fourth power of the pore radius; hence, although the ultrafiltrate flow (water permeability) is affected by the number of pores, it is affected strongly by the pore size. The diffusive flow, on the other hand, is also affected strongly by the membrane pore size in addition to the membrane thickness; but not as strong as the ultrafiltrate flow. The diffusive flow is proportional to the second power of the pore radius [Eqs. (6) and (7)].

The actual membranes used in hemodialyzers have distribution of pore sizes, and the pores are non-cylindrical in shape (Fig. 9). Fig. 11 shows the pore-size distributions and the corresponding sieving coefficient profiles of three hypothetical dialysis membranes.^[67] Membrane 1 has a large number of small pores

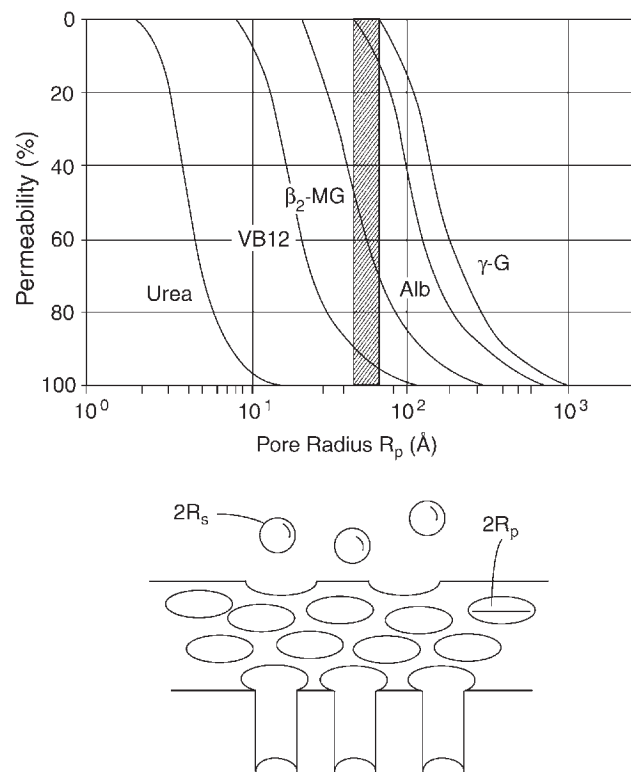


Fig. 10 Pictorial representation of an idealized membrane model with an equal pore radius R_p , which is larger than the stoke radius R_s of a hypothetical solute (lower panel); and the associated relationship between pore radius and solute permeability (upper panel). (From Ref. [85].)

(nominal diameter: 15 Å), and 2 has a large number of large pores (nominal diameter: 30 Å). Because of their narrow pore-size distributions, membranes 1 and 2 have sharp molecular-weight cutoffs of 10 and 60 kDa, respectively, which are favorable in dialysis membranes. Additionally, both membranes have favorable diffusive transport properties because of their large number of pores. Membrane 3, on the other hand, has an unfavorable pore-size distribution. The broad pore-size distribution explains the early drop-off in the sieving coefficient at low molecular-weight and the “tail” effect at high molecular-weight. The “tail” effect is highly undesirable because it may lead to albumin loss during dialysis treatment. Membrane 3 also has poor diffusive transport properties because of its small number of pores.

MASS TRANSFERS IN ARTIFICIAL KIDNEYS

The removal of uremic solutes in artificial kidneys is governed by three basic mass-transfer mechanisms: diffusion, convection, and adsorption. Diffusion and convection depend in part on the molecular size of

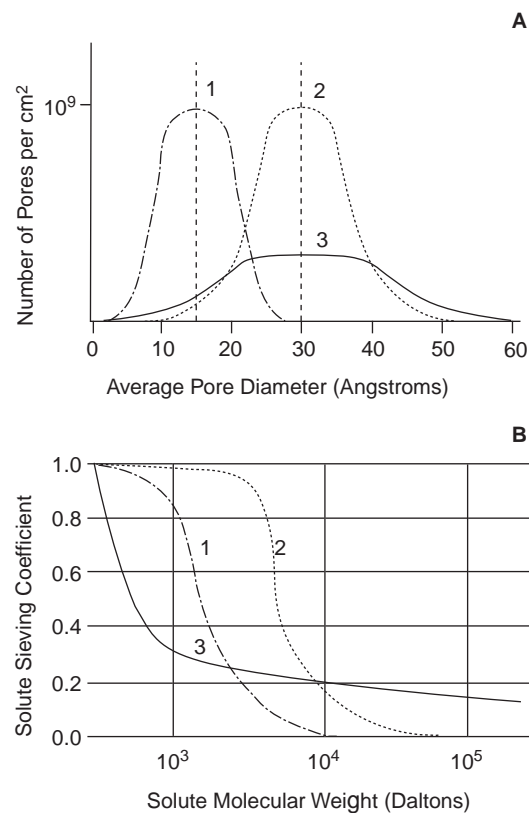


Fig. 11 Pore-size distributions and solute sieving coefficient profiles of three hypothetical dialysis membranes 1, 2, and 3. (A) The number of pores as a function of average pore diameter. (B) The solute sieving coefficient as a function of molecular weight. (From Ref. [67].)

uremic solutes and pore size of dialysis membranes, while adsorption depends on the physicochemical properties of uremic solutes and dialysis membranes (and the electrochemical properties for charged dialysis membranes).

Small solutes are predominantly removed by diffusion, and low molecular-weight proteins are removed by convection and occasionally by adsorption. Middle molecules are generally removed by diffusion and convection in high-flux and low-flux hemodialyzers, respectively, and the removal of middle molecules can be significantly enhanced by convection in high-flux hemodialyzers.

Diffusion

Diffusion is a statistical transport of solutes across a membrane by a difference in solute concentration, where solutes diffuse from the region of high concentration to the region of low concentration. Additionally, diffusion is strongly dependent on the solute molecular weight (large molecular weight corresponds to large molecular size and vice versa) and the

membrane morphology as will be discussed later. Diffusion was the primary transport mechanism of uremic solutes in the earlier development of hemodialyzers.

Diffusion of uremic solutes across the membrane, in the simplest sense, is governed by Fick's first law^[86]

$$J_{D,s} = -D_s \frac{dC}{dx} \quad (3)$$

where $J_{D,s}$ is the diffusive mass flux of a solute (mass removal rate per unit membrane area) across the membrane, D_s is the diffusion coefficient of a solute, and dC/dx is the solute concentration gradient in the x direction across the membrane. D_s is an intrinsic property of the membrane for a given solute. From Fick's first law [Eq. (3)], we can see that for a given concentration gradient across the membrane, the diffusive flux is proportional to the diffusion coefficient of the solute and is inversely proportional to the thickness of membrane. The negative sign on the right-hand side of Eq. (3) indicates that the net solute flux diffuses in the direction of negative concentration gradient, i.e., in the direction of decreasing solute concentration. We would like to point out, however, that Fick's first law is a simple phenomenological model that neglects the physicochemical details of the membrane and uremic solutes, and hence it provides no information on the relationship between the solute diffusive flux and the physicochemical natures of the membrane and uremic solutes.

When experimental data are unavailable, we can estimate the diffusion coefficient of a solute from the Stokes–Einstein equation^[87]

$$D_s = \frac{\kappa T}{6\pi r_s \mu} \quad (4)$$

where κ is Boltzmann's constant, T is the absolute temperature, r_s is the Stokes radius of the solute, and μ is the solution viscosity. We can estimate the Stokes radius of a solute from the following equation

$$r_s = \frac{3M_s}{4\pi\rho N_A} \quad (5)$$

M_s is the molecular weight of the solute, ρ is the hydrated density of the solute, and N_A is the Avogadro's number. Eqs. (4) and (5) show that diffusion is relatively favorable for low Molecular-weight solutes because of the inverse relationship between solute diffusion coefficient and MW. Eq. (4) also shows that, phenomenologically, the diffusive solute transport is dependent on the solution viscosity and temperature in addition to the solute molecular size. Stokes–Einstein equation has been reported to overestimating the

actual solute diffusion coefficient by 10% for solutes with molecular weights greater than several thousands.^[81] This equation is only applicable for laboratory experiments, where aqueous solutions with low solute concentration are used instead of human blood. The real transport of uremic solutes in human blood across the dialysis membrane is more complex than that in aqueous solution because of the coupling effect of blood, dialysate, and uremic solutes.

A common misconception among researchers is to relate the solute diffusion rate to the pore size of the membrane without taking into consideration the porosity of the membrane. Consider a membrane with n cylindrical straight pores with equal radius r perpendicular to the membrane surface, we can express the diffusive mass flux as^[47]

$$J_{D,s} = \frac{D_s \lambda \phi \Delta C}{x} \quad (6)$$

where λ is the solute partition coefficient ($\lambda=1$ for small solutes such as urea and creatinine, and less than one for larger solutes), ϕ is the membrane porosity, ΔC is the solute concentration difference across the membrane, and x is the membrane thickness. The membrane porosity is a function of pore size and pore number, and is given by

$$\phi = \frac{n\pi r^2}{A} \quad (7)$$

where A is the membrane area. Eq. (6) shows that solute diffusion is not only dependent on the membrane pore size, but also on porosity. Larger porosity allows greater access of small solutes to the membrane pores, and hence greater diffusion rate, provided the pore is relatively larger than the solute. Now, if we have two different membranes with equal surface area and thickness that are subjected to the same solute and solute concentration difference; one membrane with large pore and low porosity and another membrane with small pore and high porosity can have equivalent solute diffusion rates provided their pores are larger than the solute of interest.

Convection

Convection is the movement of fluid and solutes across a membrane by a difference in pressure. The uremic solutes are removed by the movement of fluid or solvent drag across the membrane. The pressure in the blood compartment of a hemodialyzer is made higher than that in the dialysate compartment and forces the uremic solutes to travel along with the moving fluid from the blood to the dialysate. Convection is

dependent on both the transport properties and dynamic characteristics of the moving fluid.

Convection was initially used to remove excess plasma water from the blood. With the advent of membrane technology, new high-flux dialysis membranes with improved solute transport properties of middle molecules and low molecular-weight proteins have been developed; and, to date, convection has been used to remove these uremic solutes in high-flux hemodialyzers in addition to the removal of excess plasma water.

It is worthwhile pointing out that even though convection is the main transport mechanism for the removal of middle molecules and low molecular-weight proteins in high-flux hemodialyzers, it does function in low-flux hemodialyzers as ultrafiltration for the removal of excess plasma water from the blood.

The convective volumetric rate of water from the blood to the dialysate, better known as the ultrafiltration rate Q_{UF} , is given by

$$Q_{UF} = K_{UF}\Delta P \quad (8)$$

where K_{UF} is the ultrafiltration coefficient, an intrinsic transport property of the membrane; and ΔP is the transmembrane pressure across the membrane. The transmembrane pressure is limited by the mechanical properties of the membrane. The ultrafiltration rate increases linearly with the applied transmembrane pressure until the maximum transmembrane pressure sustainable by the membrane; further increase in transmembrane pressure across the membrane beyond the maximum transmembrane pressure will cause the membrane to lose its mechanical strength and break.

The convective mass flux of a solute $J_{C,s}$, on the other hand, is given by^[88]

$$J_{C,s} = (1 - \sigma)J_{UF}C_m \quad (9)$$

where σ is the reflection coefficient (a measure of solute rejection by the membrane), J_{UF} is the ultrafiltrate flux (ultrafiltration rate per unit membrane area), and C_m is the mean intramembrane solute concentration (the average of the blood-side and dialysate-side solute concentrations, but the blood-side solute concentration is determined from the submembranous concentration rather than the bulk phase concentration). We can view the quantity $(1-\sigma)$ as the membrane resistance to convective flow, where $\sigma=1$ implies no convective flow, and $\sigma=0$ implies no resistance to the convective flow. Eq. (9) shows that the convective solute flux is dependent on the ultrafiltrate flux and hence on the transmembrane pressure, where it increases with increasing transmembrane pressure. The convective solute flux is also dependent on the solute molecular size and membrane physical properties. For a given

high-flux membrane and transmembrane pressure, the convective solute flux is independent of its molecular size for small solutes and some middle molecules ($\sigma=0$), and is strongly dependent on its molecular size for large molecules such as myoglobin and albumin ($\sigma>0$). The convective solute flux increases with decreasing solute molecular size. For a given solute and transmembrane pressure, on the other hand, the convective solute flux is highly dependent on the solute permeability of the membrane. The convective solute flux increases with increasing solute permeability of the membrane. The convective solute flux is also dependent on the solute charge, if any, and surface charge of the membrane.

Adsorption

Adsorption is the binding of uremic solutes onto the surface of a membrane. Adsorption plays a significant role in removing certain low molecular-weight proteins such as β_2 -microglobulin.^[89-91] Adsorption primarily occurs inside the internal structure of synthetic membranes rather than on the surface; thereby, the adsorption of low molecular-weight proteins greatly depends on the accessibility of these proteins to the membrane's internal structure.^[92] Consequently, adsorption of low molecular-weight proteins inside low-flux membranes is expected to be clinically insignificant compared with that occurs inside high-flux membranes with open structure. The adsorption affinity of low molecular-weight proteins on certain high-flux hydrophobic synthetic membranes is relatively high compared with hydrophilic cellulosic membranes.^[93] Ronco et al.^[94] have shown that adsorption accounts for more than 90% of the β_2 -microglobulin removal in a hydrophobic polysulfone hemodialyzer.

Adsorption also prevents the transfer of cytokine-inducing substances (CIS) from contaminated dialysate to the blood.^[95] Lonnemann and Koch^[96] and Li et al.^[97] have demonstrated that certain high-flux synthetic membranes attenuate either significantly or totally in vitro transfer of CIS from highly contaminated dialysate to the blood. This attenuation is attributed to the adsorptive removal of CIS by the membrane's internal structure via nonspecific hydrophobic interactions.

PERFORMANCE PARAMETERS OF ARTIFICIAL KIDNEYS

Clearance

One of the most frequently used parameters to evaluate the performance of hemodialyzers is clearance.

Clearance describes the ability of a hemodialyzer in removing a given solute in a single pass and is defined as the ratio of the solute mass removal rate to the solute concentration at the blood inlet of a hemodialyzer. From the engineering perspective, nonetheless, clearance provides little information on the understanding of mass transfer in hemodialyzers, but it is of practical importance for comparing the performance of different hemodialyzers and is often used in the medical community.

From the mass balance of a solute in a hemodialyzer with countercurrent blood-dialysate flow, assuming no adsorption occurs, we have

$$Q_{Bi}C_{Bi} - Q_{Bo}C_{Bo} = Q_{Do}C_{Do} - Q_{Di}C_{Di} \quad (10)$$

where Q_B , Q_D , C_B , and C_D are the blood flow rate, dialysate flow rate, blood solute concentration, and dialysate solute concentration, respectively. Subscripts *i* and *o* denote the inlet and the outlet, respectively. Dividing the left-hand side of Eq. (10) by C_{Bi} , we have the blood-side clearance K_B as

$$K_B = \frac{Q_{Bi}C_{Bi} - Q_{Bo}C_{Bo}}{C_{Bi}} \quad (11)$$

Because ultrafiltration rate Q_{UF} is given as

$$Q_{UF} = Q_{Bi} - Q_{Bo} \quad (12)$$

We can rearrange Eq. (11) as

$$K_B = \frac{Q_{Bo}(C_{Bi} - C_{Bo})}{C_{Bi}} + Q_{UF} \quad (13)$$

In the case of zero ultrafiltration, where $Q_{Bi} = Q_{Bo}$, Eq. (13) reduces to Eq. (11).

The blood-side clearance is also known as the whole blood clearance. It substantially overestimates the true solute clearance of a hemodialyzer^[98] because the actual water volumetric rate at which the solute is cleared is less than the blood volumetric rate used in the calculation.

We can reduce the overestimation of the blood-side clearance by measuring the blood water clearance. The blood water flow rate Q_{BW} , which accounts for the hematocrit-dependent effects on the effective solute distribution volume, is given by^[98]

$$Q_{BW} = 0.93Q_B [1 - \text{Hct} + K_{RBC}(1 - e^{-\alpha t})\text{Hct}] \quad (14)$$

where *Hct* is the hematocrit level, K_{RBC} is the partition coefficient between the red blood cell and plasma water for a given solute, α is the transcellular rate constant, and t is the characteristic hemodialyzer residence time. Estimates for these parameters can be found in

Shinaberger, Miller, and Gardner.^[99] We can then determine the blood water clearance K_{BW} by substituting Q_{BW} for Q_B in either Eq. (11) or (13). Eq. (13) shows that the blood-side clearance increases with increasing blood flow rate and ultrafiltration rate.

Similarly, dividing the right-hand side of Eq. (10) by C_{Bi} , we have the dialysate-side clearance K_D as

$$K_D = \frac{Q_{Do}C_{Do} - Q_{Di}C_{Di}}{C_{Bi}} \quad (15)$$

In the laboratory, C_{Di} is equal to zero because the dialysate is usually prepared to be free of the solute of interest; hence, we can further reduce Eq. (15) to

$$K_D = \frac{Q_{Do}C_{Do}}{C_{Bi}} \quad (16)$$

The dialysate-side clearance provides more accurate result and is considered the “gold standard” for evaluating hemodialyzers because there are no hematocrit-dependent effects on the effective solute distribution volume.^[83] The major drawback of the dialysate-side clearance is the difficulty in assaying solutes of very low concentrations in the dialysate. An example of the in-vitro clearances of a Gambro Polyflux 170H hemodialyzer^[100] at different blood flow rates is shown in Fig. 12.

Sieving Coefficient

Another performance parameter commonly seen in the literature is the sieving coefficient. Sieving coefficient is used to evaluate the convective removal capability of a

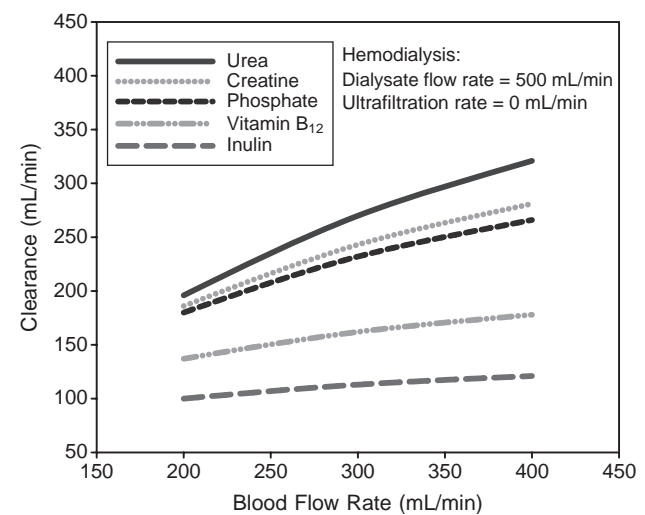


Fig. 12 In-vitro clearances of a Gambro Polyflux 170H hemodialyzer. (Data adapted from Ref. [100].)

hemodialyzer for a given solute. It is a dimensionless parameter that determines the extent of a solute permeating through the membrane. The sieving coefficient of a hemodialyzer can be determined experimentally by operating the hemodialyzer on a hemofiltration mode with a specific ultrafiltration rate, where no dialysate is used; and the inlet of the dialysate compartment is closed. The equation for determining the sieving coefficient SC is

$$SC = \frac{2C_{D_o}}{C_{B_i} + C_{B_o}} \quad (17)$$

where C_{B_i} and C_{B_o} are the solute concentrations at the inlet and outlet of the blood compartment, respectively; and C_{D_o} is the solute concentration at the outlet of the dialysate compartment.

Sieving coefficient depends strongly on the membrane pore size and is a function of ultrafiltration rate. The sieving coefficient depends also on the membrane charge, if any. The sieving coefficient increases with increasing ultrafiltration rate, up to an asymptotic value, and further increasing the ultrafiltration rate will cause the membrane to lose its mechanical strength and break. A sieving coefficient value of one indicates that 100% of the solute is permeating through the membrane, and a value of zero indicates that none of the solute is permeating.

All current hemodialyzers, irrespective of membrane type, have a sieving coefficient value of one for small solutes and less than one for middle molecules and low molecular-weight proteins. We can get an immediate appreciation of the permeability properties of a dialysis membrane by studying its sieving coefficient profile (sieving coefficients of solutes at different molecular weights). The sieving coefficient profiles of three hypothetical dialysis membranes with different pore-size distributions have been shown in Fig. 11. An example of an in-vitro sieving coefficient profile of a Gambro Polyflux R hemodialyzer^[101] is shown in Fig. 13.

RENAL REPLACEMENT THERAPIES IN ARF

The most serious forms of ARF are found with patients in the intensive care unit, which have the highest morbidity and mortality.^[102] There are several renal replacement therapies available for critically ill patients with ARF, and they can be grouped into intermittent renal replacement therapy (IRRT) and continuous renal replacement therapy (CRRT). Theoretically, CRRT is advantageous over IRRT because CRRT, which performs uremic filtration continuously, mimics the kidneys most closely and is hemodynamically more stable than IRRT, which performs uremic filtration

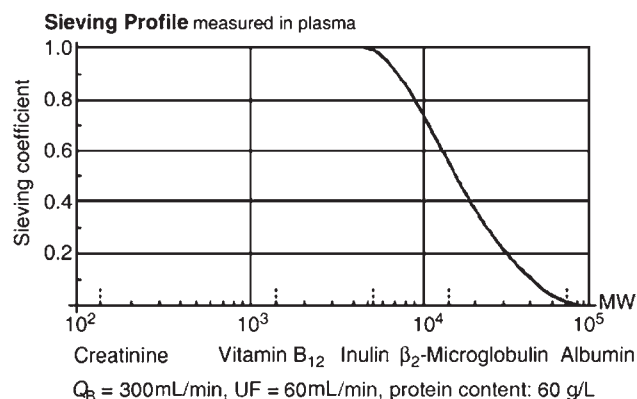


Fig. 13 An in-vitro sieving coefficient profile of a Gambro Polyflux R hemodialyzer. Q_B , Q_{UF} , and MW represent the blood flow rate, ultrafiltration rate, and molecular weight, respectively. (From Ref. [101].)

intermittently. The clinical advantage of CRRT over IRRT, however, remains controversial.^[103]

IRRT

Intermittent hemodialysis (IHD) has been the classic form of IRRT. It is the most used modality to treat ARF patients in the U.S.A.^[104] IHD is usually performed thrice weekly and lasts for 3–4 hr per session. A moderate blood flow rate between 200 and 250 mL/min and a dialysate flow rate of 500 mL/min are used.

Among advantages of IHD are rapid removal of solutes and fluids that results in rapid correction of electrolyte imbalance and removal of drugs or other substances in fatal intoxications.^[105] IHD requires less anticoagulation compared with CRRT because of its faster blood flow rate and shorter therapy duration.

One major drawback of IHD is hemodynamic instability caused by the rapid removal of solutes and fluids. Many patients develop hypotension^[106] during IHD that exacerbates renal injury and delays recovery.^[107] Rapid removal of solutes from the intravascular space can cause cerebral edema and increased intracranial pressure.^[108] ARF patients with head trauma or hepatic encephalopathy are at risk of brain edema and herniation.^[109] The cellulosic membranes commonly used in IHD have also been reported to have adverse effect on renal recovery by inducing activation of various inflammatory mediators and cascades.^[110,111]

Continuous Renal Replacement Therapy

CRRT is performed at a slower rate but continuously compared with IRRT, and hence provides better hemodynamic stability. There are several CRRT

modalities available. The first CRRT was developed by Kramer et al.^[112] called the arteriovenous hemofiltration, now called the continuous arteriovenous hemofiltration (CAVH) to treat diuretic-unresponsive fluid overload in critically ill patients with ARF. CAVH is the simplest form of CRRT. Blood is drawn from the femoral artery, flowed through a hemofilter under the influence of arterial pressure alone, and returned to the patient through the femoral vein. The blood flow rate is between 90 and 150 mL/min.^[113] The transmembrane pressure is generated from the arterial pressure, and CAVH filters between 10 and 15 L of plasma water per day. Substitution fluid is used to replace lost fluid during ultrafiltration.

Since then, various forms of CRRT have evolved that differ in the blood access and mass-transfer mode. Continuous arteriovenous hemodialysis (CAVHD) is very similar to CAVH, except that the uremic solutes are removed by diffusion instead of convection, and dialysate is used instead of substitution fluid compared with CAVH. In continuous arteriovenous hemodiafiltration (CAVHDF), on the other hand, the combination of diffusion and convection is used to remove uremic solutes. Several disadvantages of these arteriovenous therapies include the dependence of blood flow rate and ultrafiltration rate on the arterial pressure; low blood flow rate decreases the efficiency of the hemofilter and encourages blood clotting; and the arterial cannulation causes vascular complications such as ischemia, bleeding, and pseudoaneurysm.^[114]

Subsequently, the continuous arteriovenous therapies have been replaced by continuous venovenous therapies, which include continuous venovenous hemofiltration (CVVH), continuous venovenous hemodialysis (CVVHD), and continuous venovenous hemodiafiltration (CVVHDF). In these therapies, blood is drawn from either the femoral, internal jugular, or subclavian vein through a dual-lumen catheter; pumped through a hemofilter or hemodialyzer; and returned to the patient through the dual-lumen catheter inserted at the vein. Pumps are used to control the blood flow rate and

ultrafiltration rate, and hence to ensure adequate solute clearance and ultrafiltration rate. The substitution fluid is administered either before (predilution) or after (postdilution) the hemofilter in CVVH and CVVHDF. The predilution requires 15% more ultrafiltrate volume and offers enhanced solute removal along with diminished anticoagulation requirements.^[115] Table 1 lists the advantages of CRRT,^[116] Table 2 shows the comparison between different CRRTs,^[113] and Table 3 shows the comparison between the characteristics of IHD and CRRT. Detailed reviews and discussions on the above CRRT can be found in Refs.^[113,118–121]

Another rarely used CRRT in ARF is peritoneal dialysis. Peritoneal dialysis uses the peritoneal membrane to remove uremic solutes and excess plasma water from the blood. It is rarely used in ARF because it provides inadequate solute clearance in critically ill catabolic patients, increases the risk of peritonitis, compromises respiratory function by impeding diaphragmatic excursion, and is contraindicated in patients with recent abdominal surgery or abdominal sepsis.^[122] Detailed discussion on peritoneal dialysis has been presented by Daugirdas, Blake, and Ing.^[123]

Table 1 Advantages of CRRT

Hemodynamically well tolerated; minimal change in plasma osmolality
Better control of azotemia and electrolyte and acid–base balance; corrects abnormalities as they evolve; steady-state chemistries
Highly effective in removing fluid (postsurgery, pulmonary edema, and acute respiratory distress syndrome)
Facilitates administration of parenteral nutrition and obligatory intravenous medications (i.e., pressor, inotropic drugs) by creating unlimited “space” by virtue of continuous ultrafiltration
Recent efforts to make procedure user-friendly

(From Ref. [116].)

Table 2 Comparison between different CRRTs

Modality	Blood pump	Dialysate (D) or substitution fluid (SF)	Urea clearance (L/day)	Urea clearance (mL/min)	Middle molecules clearance	Simplicity
CVVH	Yes	SF	22–24	15–17	+++	++
CAVHD	No	D	24–30	17–21	n/a	+
CVVHD	Yes	D	24–30	17–21	n/a	++
CAVHDF	No	SF + D	36–38	25–26	+++	+++
CVVHDF	Yes	SF + D	36–38	25–26	+++	+++

Symbols: +, most simple; ++, moderate difficult; +++, most difficult. (From Ref. [113].)

Table 3 Comparison between IHD and CRRT

Intermittent hemodialysis	Continuous renal replacement therapy
Mainly diffusive	Mainly convective
Low-flux membrane	High-flux membrane
High dialysate flow	Low dialysate flow
On-line dialysate production	Industrially manufactured substitution fluid
A few hours per day	In theory, continuously
Technically demanding	Technically less demanding
Less labor intensive	Labor intensive

(From Ref. [117].)

Sustained Low-Efficiency Dialysis

Sustained low-efficiency dialysis (SLED)^[124,125] is a fairly new renal replacement therapy that combines the advantages of IHD and CRRT. SLED is performed at low blood and dialysate flow rates of 200 and 100 mL/min, respectively, for 6–12 hr/day. It provides hemodynamic stability, improved correction of hypervolemia, and superior solute removal compared with IHD. It uses the conventional hemodialysis machine and hence costs less than CRRT that requires expensive machine. SLED can be done intermittently based on the need of the patient; hence, the interruption of therapy for various diagnostics and therapeutic procedures can be avoided.

Only few data have been reported on SLED, and the clinical reliability and efficacy of SLED remain to be seen. We have used kinetic modeling to compare the effective doses delivered by CVVH, daily hemodialysis, and SLED.^[126] We used a modified equivalent renal clearance (EKR) to account for the unsteady initial stage of dialysis. We found that small solute EKR in CVVH is 8% and 60% higher than those in SLED and daily hemodialysis, respectively. The differences in EKR are more pronounced for middle and large solutes, where their EKR in CVVH are approximately twofold and fourfold larger than their corresponding values in daily hemodialysis and SLED, respectively. The superior removal of middle and large solutes in CVVH is caused by the powerful combination of convection and continuous operation in CVVH.

CONCLUSIONS

We have scratched the surface of the topic of artificial kidneys from the engineering perspective, where we have discussed the basic functions of human kidneys, classification of uremic solutes, basic designs and

classification of artificial kidneys, property and classification of dialysis membranes, mass transfers in artificial kidneys, parameters used to characterize performance of artificial kidneys, and different forms of renal replacement therapies for ARF.

The topic *artificial kidneys* encompasses a vast area of subjects, and it is spatially infeasible to discuss them all in purely one article; we thereby highly encourage the readers to refer to the “Article of Further Interest”, “References”, and “Bibliography” cited at the end of this article to further satisfy their quest for knowledge of artificial kidneys.

We hope that this fundamental entry provides the necessary momentum to further enhance the readers’ interest in artificial kidneys, and we wish the readers all the best for their endeavors in the research of artificial kidneys and in improving the quality of life of patients with kidney failures.

ARTICLES OF FURTHER INTEREST

Bioreactors; Blood Purification

REFERENCES

1. (NIDDK) National Institute of Diabetes and Digestive and Kidney Diseases. *Your Kidneys and How They Work*, NIH Publication No. 98-4241; National Kidney and Urologic Diseases Information Clearinghouse (NKUDIC): Bethesda, MD, U.S.A, March 1998.
2. Seldin, D.W., Giebisch, G., Eds.; *The Kidney: Physiology and Pathophysiology*, 2nd Ed.; Raven Press: New York, 1992.
3. Stanton, B.A.; Koepfen, B.M. The kidney. In *Physiology*, 3rd Ed.; Berne, R.M., Levy, M.N., Eds.; Mosby-Year Book, Inc.: St. Louis, MO, 1993; 717–809.
4. Jacobs, C., Kjellstrand, C.M., Koch, K.M., Winchester, J.G., Eds.; *Replacement of Renal Function by Dialysis*, Kluwer Academic Publishers, 4th Ed.: Boston, MA, 1996.
5. Schrier, R.W., Ed.; *Atlas of Kidney Diseases*; Blackwell Science: Philadelphia, 1999.
6. Brenner, B.M., Ed.; *Brenner & Rector’s The Kidney*; W.B. Saunders: Philadelphia, PA, 2004.
7. Gao, D.; Kraus, M.A.; Ronco, C.; Clark, W.R. The biology of dialysis. In *Pediatric Dialysis*; Warady, B.A., Schaefer, F.S., Fine, R.N., Alexander, S.R., Eds.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 2004; 13–34.
8. Cambi, V.; Dall’Aglio, P.; Savazzi, G.; Arisi, L.; Rossi, E.; Migone, L. Clinical assessment of haemodialysis patients with reduced small molecules removal. *Proc. Eur. Dial. Transplant. Assoc.* **1972**, *9*, 67–73.
9. Cambi, V. Limits of dialysis technology. *Proceedings of the Sixth International Congress of Nephrology*, 1975. Karger AG: Basel, Switzerland, 1976; 624–628.

10. Jebesen, R.H.; Tenckhoff, H.; Hoult, J.C. Natural history of uremic polyneuropathy and the effects of dialysis. *N. Engl. J. Med.* **1967**, *277*, 327–333.
11. Tenckhoff, H.; Curtis, F.K. Experience with maintenance peritoneal dialysis in the home. *Trans. Am. Soc. Artif. Intern. Organs* **1970**, *16*, 90–95.
12. Babb, A.L.; Popovich, R.P.; Christopher, T.G.; Scribner, B.H. The genesis of the square-meter hour hypothesis. *Trans. Am. Soc. Artif. Organs* **1971**, *17*, 81–91.
13. Babb, A.L.; Ahmad, S.; Bergstrom, J.; Scribner, B.H. The middle molecule hypothesis in perspective. *Am. J. Kidney Dis.* **1981**, *1*, 46–50.
14. Lowrie, E.G.; Laird, N.M.; Parker, T.F.; Sargent, J.A. Effect of the hemodialysis prescription on patient morbidity. *N. Engl. J. Med.* **1980**, *305*, 1176–1181.
15. Vanholder, R.; De Smet, R.; Vogeleere, P.; Ringoir, S. Middle molecules: toxicity and removal by hemodialysis and related strategies. *Artif. Organs* **1995**, *19* (11), 1120–1125.
16. Rockel, A.; Gilge, U.; Liewald, A.; Heidland, A. Elimination of low molecular weight proteins during hemofiltration. *Artif. Organs* **1982**, *6*, 307–311.
17. Rockel, A.; Abdelhamid, S.; Fiegel, P.; Walb, D. Elimination of low molecular weight proteins with high flux membranes. *Contr. Nephrol.* **1985**, *46*, 69–74.
18. Gejyo, F.; Yamada, T.; Odani, S.; Nakagawa, Y.; Arakawa, M.; Kunitomo, T.; Kataoka, H.; Suzuki, M.; Hirasawa, Y.; Shirahama, T.; Cohen, A.; Schmid, K. A new form of amyloid protein associated with chronic hemodialysis was identified as beta 2-microglobulin. *Biochem. Biophys. Res. Commun.* **1985**, *129* (3), 701–706.
19. Gorevic, P.D.; Casey, T.T.; Stone, W.J.; Di Raimondo, C.R.; Prelli, F.C.; Frangione, B. Beta-2 microglobulin is an amyloidogenic protein in man. *J. Clin. Invest.* **1985**, *76*, 2425–2429.
20. Gorevic, P.D.; Munoz, P.C.; Casey, T.T.; Di Raimondo, C.R.; Stone, W.J.; Prelli, F.C.; Rodrigues, M.M.; Poulik, M.D.; Frangione, B. Polymerization of intact β_2 -microglobulin in tissue causes amyloidosis in patients on chronic hemodialysis. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 7908–7912.
21. Colton, C.K. Analysis of membrane processes for blood purification. *Blood Purif.* **1987**, *5*, 202–251.
22. Vanholder, R.; De Smet, R.; Hsu, C.; Vogeleere, P.; Ringoir, S. Uremic toxicity: the middle molecule hypothesis revisited. *Semin. Nephrol.* **1994**, *14* (3), 205–218.
23. Anderstam, B.; Mamoun, A.H.; Sodersten, P.; Bergstrom, J. Middle-sized molecule fractions isolated from uremic ultrafiltrate and normal urine inhibit ingestive behavior in the rat. *J. Am. Soc. Nephrol.* **1996**, *7* (11), 2453–2460.
24. Vanholder, R.; Argiles, A.; Baurmeister, U.; Brunet, P.; Clark, W.; Cohen, G.; De Deyn, P.P.; Deppisch, R.; Descamps-Latscha, B.; Henle, T.; Jorresm, A.; Massy, Z.A.; Rodriguez, M.; Stegmayr, B.; Stenvinkel, P.; Wratten, M.L. Uremic toxicity: present state of the art. *Int. J. Artif. Organs.* **2001**, *24* (10), 695–725.
25. Vanholder, R.; De Smet, R.; Glorieux, G.; Argiles, A.; Baurmeister, U.; Brunet, P.; Clark, W.; Cohen, G.; De Deyn, P.P.; Deppisch, R.; Descamps-Latscha, B.; Henle, T.; Jorres, A.; Lemke, H.D.; Massy, Z.A.; Passlick-Deetjen, J.; Rodriguez, M.; Stegmayr, B.; Stenvinkel, P.; Tetta, C.; Wanner, C.; Zidek, W. Review on uremic toxins: classification, concentration, and interindividual variability. *Kidney Int.* **2003**, *63* (5), 1934–1943.
26. Vanholder, R.; De Smet, R. Pathophysiologic effects of uremic retention solutes. *J. Am. Soc. Nephrol.* **1999**, *10* (8), 1815–1823.
27. Graham, T. Osmotic force. *Philos. Trans. R. Soc. Lond.* **1854**, *144*, 177–228.
28. Sakai, K. Artificial kidney engineering—dialysis membrane and dialyzer for blood purification. *J. Chem. Eng. Jpn.* **1997**, *30* (4), 587–599.
29. Abel, J.J.; Rowntree, L.G.; Turner, B.B. On the removal of diffusible substances from the circulating blood of living animals by means of dialysis. *J. Pharmacol. Exp. Ther.* **1914**, *5*, 275–316.
30. Haas, G. Dialysis of the flowing blood in the patient. *Klin. Wochenschr.* **1923**, *70*, 1888.
31. Kolff, W.J. *New Ways of Treating Uraemia: The Artificial Kidney, Peritoneal Lavage, Intestinal Lavage*; Churchill: London, 1947.
32. Kolff, W.J. The beginning of the artificial kidney. *Artif. Organs* **1993**, *17* (5), 293–299.
33. Kolff, W.J. The artificial kidney and its effect on the development of other artificial organs. *Nat. Med.* **2002**, *8* (10), 1063–1065.
34. Kolff, W.J.; Watschinger, B.; Further, B. Development of a coil kidney. Disposable artificial kidney. *J. Lab. Clin. Med.* **1956**, *47*, 969–977.
35. Kiil, F. Development of a parallel-flow artificial kidney in plastics. *Acta Chir. Scand. Suppl.* **1960**, *253*, 142–150.
36. Stewart, R.D.; Cerny, J.C.; Mahon, H.I. The capillary kidney: preliminary report. *Univ. Mich. Med. Cent. J.* **1964**, *30*, 116.
37. Poh, C.K.; Hardy, P.A.; Clark, W.R.; Gao, D. Characterization of flow in hemodialyzers using MRI. *Proc. Intl. Soc. Mag. Reson. Med.* **2001**, *9*, 1984.
38. Poh, C.K.; Hardy, P.A.; Liao, Z.; Huang, Z.; Gao, D.; Clark, W.R. Characterization of flow distribution in hemodialyzers using MRI (abstract). *ASAIO J.* **2002**, *48* (2), 179.
39. Poh, C.K.; Hardy, P.A.; Liao, Z.; Clark, W.R.; Gao, D. Noninvasive characterization of fluid transport phenomena in hollow-fiber membrane modules using MRI: an innovative experimental approach. In *New Insights into Membrane Science and Technology: Polymeric and Biofunctional Membranes*; Bhattacharyya, D., Butterfield, D.A., Eds.; Elsevier Science B.V.: Amsterdam, The Netherlands, 2003; 89–122.
40. Poh, C.K.; Hardy, P.A.; Liao, Z.; Huang, Z.; Clark, W.R.; Gao, D. Effect of flow baffles on the dialysate flow distribution of hollow-fiber hemodialyzers: a noninvasive experimental study using MRI. *J. Biomech. Eng.* **2003**, *125* (4), 481–489.

41. Poh, C.K.; Hardy, P.A.; Liao, Z.; Huang, Z.; Clark, W.R.; Gao, D. Effect of spacer yarns on the dialysate flow distribution of hemodialyzers: a magnetic resonance imaging study. *ASAIO J.* **2003**, *49* (4), 440–448.
42. Leypoldt, J.K. Dialyzers and dialysis membranes: research directions. *ASAIO J.* **1996**, *42* (3), 133–135.
43. Tokars, J.I.; Alter, M.J.; Favero, M.S.; Moyer, L.A.; Bland, L.A. National surveillance of dialysis associated diseases in the United States, 1991. *ASAIO J.* **1993**, *39* (4), 966–975.
44. Cheung, A.K.; Agodoa, L.Y.; Daugirdas, J.T.; Depner, T.A.; Gotch, F.A.; Greene, T.; Levin, N.W.; Leypoldt, J.K. Hemo Study Group. Effect of hemodialyzer reuse on clearances of urea and β_2 -microglobulin. *J. Am. Soc. Nephrol.* **1999**, *10*, 117–127.
45. Barth, R.H. Pros and cons of short, high efficiency, and high flux dialysis. In *Replacement of Renal Function by Dialysis*, 4th Ed.; Jacobs, C., Kjellstrand, C.M., Koch, K.M., Winchester, J.F., Eds.; Kluwer Academic Publishers: Boston, MA, 1996; 418–453.
46. Clark, W.R.; Hamburger, R.J.; Lysaght, M.J. Effect of membrane composition and structure on solute removal and biocompatibility in hemodialysis. *Kidney Int.* **1999**, *56* (6), 2005–2015.
47. Lysaght, M.J. Hemodialysis membranes in transition. *Contrib. Nephrol.* **1988**, *61*, 1–17.
48. Keshaviah, P.; Luehmann, D.; Ilstrup, K.; Collins, A. Technical requirements for rapid high-efficiency therapies. *Artif. Organs* **1986**, *10* (3), 189–194.
49. Collins, A.J.; Keshaviah, P. High efficiency therapies for clinical dialysis. In *Clinical Dialysis*, 2nd Ed.; Nissenson, A.R., Fine, R.N., Gentile, D.E., Eds.; Appleton & Lange: Norwalk, 1990; 687 pp.
50. Collins, A.J. High-flux, high-efficiency procedures. In *Principles and Practice of Dialysis*; Henrich, W.L., Ed.; Williams & Wilkins: Baltimore, 1994; 76 pp.
51. Held, P.J.; Levin, N.W.; Bovbjerg, R.R.; Pauly, M.V.; Diamond, L.H. Mortality and duration of hemodialysis. *JAMA* **1991**, *265*, 871–875.
52. Pastan, S.; Bailey, J. Dialysis therapy. *N. Engl. J. Med.* **1998**, *338* (20), 1428–1437.
53. Akizawa, T.; Kinugasa, E.; Ideura, T. Classification of dialysis membranes by performance. *Contrib. Nephrol.* **1995**, *113*, 25–31.
54. Craddock, P.R.; Fehr, J.; Dalmasso, A.P.; Bingham, K.L.; Jacob, H.S. Hemodialysis leukopenia: pulmonary vascular leukostasis resulting from complement activation by dialyzer cellophane membranes. *J. Clin. Invest.* **1977**, *59*, 879–888.
55. Girndt, M.; Heisel, O.; Köhler, H. Influence of dialysis with polyamide vs haemophan haemodialysers on monokines and complement activation during a 4-month long-term study. *Nephrol. Dial. Transplant.* **1999**, *14*, 676–682.
56. Deppisch, R.; Schmitt, V.; Bommer, J.; Hansch, G.M.; Ritz, E.; Rauterberg, E.W. Fluid phase generation of terminal complement complex as a novel index of bioincompatibility. *Kidney Int.* **1990**, *37* (2), 696–706.
57. Schaefer, R.M.; Rayterberg, E.W.; Deppisch, R.; Vienken, J. Assembly of terminal SC5b-9 complement complexes: a new index of blood-membrane interaction. *Miner. Electrolyte Metab.* **1990**, *16* (1), 73–76.
58. Valette, P.; Thomas, M.; Déjardin, P. Adsorption of low molecular weight proteins to hemodialysis membranes: experimental results and simulations. *Biomaterials* **1999**, *20*, 1621–1634.
59. Morti, S.M.; Zydney, A.L. Protein-membrane interactions during hemodialysis: effects on solute transport. *ASAIO J.* **1998**, *44* (4), 319–326.
60. Morti, S.M.; Zydney, A.L. Transport characteristics of different hemodialysis membranes (abstract). *ASAIO J.* **2000**, *46* (2), 220.
61. Walton, D.; Cheung, A. Membrane biocompatibility. In *Clinical Dialysis*, 3rd Ed.; Nissenson, A., Fine, R., Gentile, D., Eds.; Appleton & Lange: Norwalk, CT, 1995; 93–120.
62. Cheung, A.K.; Leypoldt, J.K. The hemodialysis membranes: a historical perspective, current state and future prospect. *Semin. Nephrol.* **1997**, *17* (3), 196–213.
63. U.S. Renal Data System. USRDS 2000 Annual Data Report: Atlas of End-Stage Renal Disease in the United States; National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases: Bethesda, MD, June 2000.
64. Clark, W.R.; Gao, D. Properties of membranes used for hemodialysis therapy. *Semin. Dial.* **2002**, *15* (1), 191–195.
65. Lucchi, L.; Bonucchi, D.; Acerbi, M. Improved biocompatibility by modified cellulose membranes: the case of hemophan. *Artif. Organs* **1989**, *13*, 417–421.
66. Bowry, S.K.; Rintelen, T.H. Synthetically modified cellulose: a cellulosic hemodialysis membrane with minimized complement activation. *ASAIO J.* **1998**, *44*, M579–M583.
67. Ronco, C.; Ballestri, M.; Cappelli, G. Dialysis membranes in convective treatments. *Nephrol. Dial. Transplant.* **2000**, *15* (suppl 2), 31–36.
68. Deppisch, R.; Gohl, H.; Smeby, L. Microdomain structure of polymeric surfaces: potential for improving blood treatment procedures. *Nephrol. Dial. Transplant.* **1998**, *13*, 1354–1359.
69. Hakim, R.M.; Held, P.J.; Stannard, D.C.; Wolfe, R.A.; Port, F.K.; Daugirdas, J.T.; Agodoa, L. Effect of the dialysis membrane on mortality of chronic hemodialysis patients. *Kidney Int.* **1996**, *50* (2), 566–570.
70. La Greca, G.; Ronco, C. *Vitamin E-Bonded Membrane: A Further Step in Dialysis Optimization*; Karger AG: Basel, Switzerland, 1999.
71. Sasaki, M.; Hosoya, N.; Saruhashi, M. Vitamin E modified cellulose membrane. *Artif. Organs* **2000**, *24* (10), 779–789.
72. Huraib, S.; Tanimu, D.; Shaheen, F.; Hejaili, F.; Giles, C.; Pagayon, V. Effect of vitamin-E-modified dialysers on dialyser clotting, erythropoietin and heparin dosage: a comparative crossover study. *Am. J. Nephrol.* **2000**, *20* (5), 364–368.
73. Dhondt, A.; Vanholder, R.; Glorieux, G.; Waterloos, M.A.; De Smet, R.; Lesaffer, G.; Lameire, N. Vitamin

- E-bonded cellulose membrane and hemodialysis bioincompatibility: absence of an acute benefit on expression of leukocyte surface molecules. *Am. J. Kidney Dis.* **2000**, *36* (6), 1140–1146.
74. Shimazu, T.; Ominato, M.; Toyama, K.; Yasuda, T.; Sato, T.; Maeba, T.; Owada, S.; Ishida, M. Effects of a vitamin E-modified dialysis membrane on neutrophil superoxide anion radical production. *Kidney Int.* **2001**, *59* (Suppl 78), S137–S143.
 75. Mydlik, M.; Derzsiova, K.; Racz, O.; Sipulova, A.; Lovasova, E.; Petrovicova, J. A modified dialyzer with vitamin E and antioxidant defense parameters. *Kidney Int.* **2001**, *59* (Suppl 78), S144–S147.
 76. MacGinley, R.; Westhuyzen, J.; Saltissi, D.; Morgan, C.; Healy, H.; Thirlwell, G.K.; Disney, A.P. Evaluation of a novel vitamin E coated cellulosic membrane hollow fiber dialyzer. *ASAIO J.* **2001**, *47* (1), 66–73.
 77. Tsuruoka, S.; Kawaguchi, A.; Nishiki, K.; Hayasaka, T.; Fukushima, C.; Sugimoto, K.; Saito, T.; Fujimura, A. Vitamin E-bonded hemodialyzer improves neutrophil function and oxidative stress in patients with end-stage renal failure. *Am. J. Kidney Dis.* **2002**, *39* (1), 127–133.
 78. Nakatan, T.; Takemoto, Y.; Tsuchida, A.K. The effect of vitamin E-bonded dialyzer membrane on red blood cell survival in hemodialyzed patients. *Artif. Organs* **2003**, *27* (3), 214–217.
 79. Kobayashi, S.; Moriya, H.; Aso, K.; Ohtake, T. Vitamin E-bonded hemodialyzer improves atherosclerosis associated with a rheological improvement of circulating red blood cells. *Kidney Int.* **2003**, *63* (5), 1881–1887.
 80. Triolo, L.; Malaguti, M.; Ansali, F.; Comunian, M.C.; Arcangeloni, O.; Coppolino, F.; Marrocco, F.; Sicoli, R.; Biagini, M. Vitamin E-bonded cellulose membrane, lipoperoxidation, and anemia in hemodialysis patients. *Artif. Cells Blood Substit. Immobil. Biotechnol.* **2003**, *31* (3), 185–191.
 81. Colton, C.K.; Lowrie, E.G. Hemodialysis: physical principles and technical considerations. In *The Kidney*; Brenner, B.M., Rector, F.C., Eds.; W.B. Saunders: Philadelphia, PA, 1981; 2425–2489.
 82. Bird, R.B.; Stewart, W.E.; Lightfoot, E.N. Velocity distributions in laminar flow. In *Transport Phenomena*, 1st Ed.; Bird, R.B., Stewart, W.E., Lightfoot, E.N., Eds.; John Wiley & Sons: New York, 1960; 34–70.
 83. Clark, W.R. Quantitative characterization of hemodialyzer solute and water transport. *Semin. Dial.* **2001**, *14* (1), 32–36.
 84. Lonnemann, G. Chronic inflammation in hemodialysis: the role of contaminated dialysate. *Blood Purif.* **2000**, *18* (3), 214–223.
 85. Takeyama, T.; Sakai, Y. Polymethylmethacrylate: one biomaterial for a series of membrane. *Contrib. Nephrol.* **1999**, *125*, 9–24.
 86. Kessler, S.B.; Klein, E. Dialysis: theory. In *Membrane Handbook*; Ho, W.S.W., Sirkar, K.K., Eds.; Van Nostrand Reinhold: New York, 1992; 168 pp.
 87. Welty, J.R.; Wicks, C.E.; Wilson, R.E. *Fundamental of Momentum, Heat, and Mass Transfer*, 3rd Ed.; John Wiley & Sons: New York, 1984; 494 pp.
 88. Werynski, A.; Waniewski, J. Theoretical description of mass transport in medical membrane devices. *Artif. Organs* **1995**, *19* (5), 420–427.
 89. Jorstad, S.; Smeby, L.; Balstad, T.; Wideroe, T. Removal, generation, and adsorption of β_2 -microglobulin during hemofiltration with five different membranes. *Blood Purif.* **1988**, *6* (2), 96–105.
 90. Jindal, K.K.; McDougall, J.; Woods, B.; Nowakowski, L.; Goldstein, M.B. A study of the basic principles determining the performance of several high-flux dialyzers. *Am. J. Kidney Dis.* **1989**, *14* (6), 507–511.
 91. Klinke, B.; Rockel, A.; Abdelhamid, S.; Fiegel, P.; Walb, D. Transmembranous transport and adsorption of beta-2-microglobulin during hemodialysis using polysulfone, polyacrylonitrile, polymethylmethacrylate and cuprammonium rayon membranes. *Int. J. Artif. Organs* **1989**, *12* (11), 697–702.
 92. Clark, W.R.; Macias, W.L.; Molitoris, B.A.; Wang, N.H. Membrane adsorption of β_2 -microglobulin: equilibrium and kinetic characterization. *Kidney Int.* **1994**, *46* (4), 1140–1146.
 93. Clark, W.R.; Macias, W.L.; Molitoris, B.A.; Wang, N.H. Plasma protein adsorption to highly permeable hemodialysis membranes. *Kidney Int.* **1995**, *48* (2), 481–488.
 94. Ronco, C.; Heifetz, A.; Fox, K.; Curtin, C.; Brendolan, A.; Gastaldon, F.; Crepaldi, C.; Fortunato, A.; Pietribasi, G.; Caberlotto, A.; Brunello, A.; Milan Manani, S.; Zanella, M.; La Greca, G. Beta 2-microglobulin removal by synthetic dialysis membranes. Mechanisms and kinetics of the molecule. *Int. J. Artif. Organs* **1997**, *20* (3), 136–143.
 95. Clark, W.R.; Gao, D.; Ronco, C. Membranes for dialysis: composition, structure and function. *Contrib. Nephrol.* **2002**, *137*, 70–77.
 96. Lonnemann, G.; Koch, K.M. β_2 -microglobulin amyloidosis: effects of ultrapure dialysate and type of dialyzer membrane. *J. Am. Soc. Nephrol.* **2002**, *13* (Suppl 1), S72–S77.
 97. Li, B.; Huang, Z.; Poh, C.; Liao, Z.; Gao, D.; Clark, W.R. Prevention of cytokine-inducing substance transport by a new high-flux synthetic dialyzer (abstract). *J. Am. Soc. Nephrol.* **2001**, *12*, 269A.
 98. Lim, V.S.; Flanigan, M.J.; Fangman, J. Effect of hematocrit on solute removal during high efficiency hemodialysis. *Kidney Int.* **1990**, *37* (6), 1557–1562.
 99. Shinaberger, J.; Miller, J.; Gardner, P. Erythropoietin alert: risks of high hematocrit hemodialysis. *Trans. Am. Soc. Artif. Intern. Organs* **1988**, *34*, 179–184.
 100. Gambro, Inc. *Polyflux[®] H Dialyzer Spec Sheet*, 306150084 Rev. A; Gambro Renal Products, Inc.: Lakewood, Colorado, U.S.A., 2004.
 101. Gambro, Inc. *English Operator's Manual POLYFLUX RTM*; N 50 123 rev. 004; Gambro Renal Products, Inc.: Lakewood, Colorado, U.S.A.
 102. Lameire, N.; Van Biesen, W.; Vanholder, R.; Colardijn, F. The place of intermittent hemodialysis in the

- treatment of acute renal failure in the ICU patient. *Kidney Int.* **1998**, *53* (Suppl 66), S110–S119.
103. Bellomo, R.; Ronco, C. Continuous versus intermittent renal replacement therapy in the intensive care unit. *Kidney Int.* **1998**, *53* (Suppl 66), S125–S128.
 104. Mehta, R.; Letteri, J. National Kidney Foundation Council on Dialysis. Current status of renal replacement therapy for acute renal failure: a survey of US nephrologists (abstract). *J. Am. Soc. Nephrol.* **1996**, *7*, 1457.
 105. O'Reilly, P.; Tolwani, A. Renal replacement therapy III: IHD, CRRT, SLED. *Crit. Care. Clin.* **2005**, *21* (2), 367–378.
 106. Zucchelli, P.; Santoro, A. Dialysis-induced hypotension. A fresh look at Pathophysiology. *Blood Purif.* **1993**, *11*, 85–98.
 107. Conger, J.D. Does hemodialysis delay recovery from ARF? *Semin. Dial.* **1990**, *3*, 146–148.
 108. Davenport, A.; Finn, R.; Goldsmith, H.J. Management of patients with renal failure complicated by cerebral oedema. *Blood Purif.* **1989**, *7* (4), 203–209.
 109. Davenport, A.; Will, E.J.; Davidson, A.M. Effect of renal replacement therapy on patients with combined acute renal failure and fulminant hepatic failure. *Kidney Int.* **1993**, *43* (Suppl 41), S245–S251.
 110. Schiff, H.; Lang, S.M.; Konig, A.; Strasser, T.; Haider, M.C.; Held, E. Biocompatible membranes in acute renal failure: prospective case-controlled study. *Lancet* **1994**, *344* (8922), 570–572.
 111. Hakim, R.M.; Wingard, R.L.; Parker, R.A. Effect of the dialysis membrane in the treatment of patients with acute renal failure. *N. Engl. J. Med.* **1994**, *331* (20), 1338–1342.
 112. Kramer, P.; Wigger, W.; Rieger, J.; Matthaei, D.; Scheler, F. Arteriovenous haemofiltration: a new and simple method for treatment of over-hydrated patients resistant to diuretics (in German). *Klin. Wochenschr.* **1977**, *55* (22), 1121–1122.
 113. Manns, M.; Sigler, M.H.; Teehan, B.P. Continuous renal replacement therapies: an update. *Am. J. Kidney Dis.* **1998**, *32* (2), 185–207.
 114. Tominaga, G.T.; Ingegno, M.; Ceraldi, C.; Waxman, K. Vascular complications of continuous arteriovenous hemofiltration in trauma patients. *Trauma* **1993**, *35*, 285–289.
 115. Lebedo, I. Predilution hemofiltration: a new technology applied to an old therapy. *Int. J. Artif. Organs.* **1995**, *18* (11), 735–742.
 116. Sigler, M.H.; Teehan, B.O.; Daugirdas, J.T.; Ing, T.S. Slow continuous therapies. In *Handbook of Dialysis*, 3rd Ed.; Daugirdas, J.T., Blake, P.G., Ing, T.S., Eds.; Lippincott Williams & Wilkins: Philadelphia, PA, 2001; 199–230.
 117. Vanholder, R.; Van Biesen, W.; Lameire, N. What is the renal replacement method of first choice for intensive care patients? *J. Am. Soc. Nephrol.* **2001**, *12* (Suppl 17), S40–S43.
 118. Burchardi, H. History and development of continuous renal replacement techniques. *Kidney Int.* **1998**, *53* (Suppl 66), S120–S124.
 119. Clark, W.R.; Ronco, C. Renal replacement therapy in acute renal failure: solute removal mechanisms and dose quantification. *Kidney Int.* **1998**, *53* (Suppl 66), S133–S137.
 120. Ronco, C.; Bellomo, R. Continuous renal replacement therapy: evolution in technology and current nomenclature. *Kidney Int.* **1998**, *53* (Suppl 66), S160–S164.
 121. Bellomo, R.; Ronco, C. Continuous haemofiltration in the intensive care unit. *Crit. Care* **2000**, *4* (6), 339–345.
 122. Goel, S.; Saran, R.; Nolph, K.D. Indications, contraindications, and complications of peritoneal dialysis in the critically ill. In *Critical Care Nephrology*; Ronco, R., Bellomo, R., Eds.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1998; 1373–1381.
 123. Daugirdas, J.T., Blake, P.G., Ing, T.S., Eds.; *Handbook of Dialysis*; 3rd Ed.; Lippincott Williams & Wilkins: Philadelphia, PA, 2001; 281–410.
 124. Chatoth, D.K.; Shaver, M.J.; Marshall, M.R.; Golper, T.A. Daily 12-hour sustained low-efficiency hemodialysis (SLED) for the treatment of critically ill patients with acute renal failure: initial experience (abstract). *Blood Purif.* **1999**, *17*, 16.
 125. Marshall, M.R.; Golper, T.A.; Shaver, M.J.; Alam, M.G.; Chatoth, D.K. Sustained low-efficiency dialysis for critically ill patients requiring renal replacement therapy. *Kidney Int.* **2001**, *60* (2), 777–785.
 126. Liao, Z.; Zhang, W.; Hardy, P.A.; Poh, C.K.; Huang, Z.; Kraus, M.A.; Clark, W.R.; Gao, D. Kinetic comparison of different acute dialysis therapies. *Artif. Organs.* **2003**, *27* (9), 802–807.

BIBLIOGRAPHY

- Dhondt, A.; Vanholder, R.; Van Biesen, W.; Lameire, N. The removal of uremic toxins. *Kidney Int.* **2000**, *58* (suppl 76), s47–s59.
- Gutch, C.F.; Stoner, M.H.; Corea, A.L. Review of Hemodialysis for Nurses and Dialysis Personnel. 6th. Mosby, Inc.: St. Louis, MO, 1999.
- Kolff, W.J. The Artificial Kidney. J.H. Kok: Kampen, Holland, 1946.
- Vanholder, R. Problems and solutions for artificial kidney. *Technol. Health Care* **2000**, *8* (6), 373–379.
- Van Noordwijk, J. Dialysing for Life: the Development of the Artificial Kidney. Kluwer Academic Publishers: Boston, MA, 2001.

Knee Joint: Overuse Injuries

Ivan Bojanić
Alan Ivković
Marko Pećina

*Department of Orthopaedic Surgery, School of Medicine, University of Zagreb,
Zagreb, Croatia*



INTRODUCTION

Athletic activities and recreational exercises are the phenomenon of modern civilized societies. The Olympic logo “Citius, Altius, Fortius” may serve us better nowadays as “Even faster, even higher, even stronger.” As exercise continues to proliferate, so do the concomitant injuries, especially damage to the locomotor system. Damage of the locomotor system is the result of a series of repetitive microtraumas that overwhelm the tissue’s ability to repair itself. Therefore, many authors view it as one of the microtraumatic illnesses, but etiologically and pathogenetically, it would be better to term it as an overuse injury. The knee is a frequently injured joint in athletic and recreational activities. Overuse injuries are also frequent in the knee joint, the reason for this is that the knee joint participates in all sports activities (running, jumping, kicking, etc.) and that this joint area is characterized by numerous attachment sites for muscles and their tendons as well as by numerous bursae. It is also else to the nature of the specific joint between the patella and the femur (patellofemoral joint) that constitutes a part of the knee joint. These reasons amply explain why as much as 40% of all overuse injuries occurring in runners develop in the knee joint area. Speaking in general terms, all overuse injuries around the knee joint can be differentiated, depending on their anatomical localization into four groups (see Table 1).

As indicated in Table 1, there are numerous possible reasons for anterior knee pain. Note, however, that Table 1 primarily consists of chronic causes that lead to the development of anterior knee pain. Based on the anatomical localization (see Table 1), this entry will provide basic information on epidemiology, diagnostics, and treatment of overuse injuries around the knee joint.

PATELLOFEMORAL PAIN SYNDROME

There are numerous possible reasons for patellofemoral pain syndrome. Irregularities in the course of the patella and malalignment of the extensor system of the knee are

all described when discussing pain in the patellofemoral syndrome. Therefore, we feel it is imperative to clearly define the terms patellar malalignment, chondromalacia of the patella, and anterior knee pain.^[1]

Patellar Malalignment

Perhaps, the easiest way to define patellar malalignment is to say that malalignment of the extensor system of the knee entails changes in the architecture of the patellofemoral joint, which in turn change the relationship between the patellar stabilizers, leading to excessive tightness in some and insufficiency in others. Patellar malalignment also leads to an incorrect distribution of strain in the patellofemoral joint, resulting in increased pressure in certain areas of patellofemoral contact (see Fig. 1).

Patellar malalignment is a translational or rotational deviation of the patella relative to any axis, and it can be a major component of patellar pain. There are two theories to explain the origin of pain: the neural theory and the mechanical theory, but it should be noted that malalignment of the knee extensor system is not always accompanied by pain in the patellofemoral joint. During clinical examination of the patient who exhibits symptomatology characteristic of the patellofemoral joint, the physician should direct attention to defining the underlying causative factors that have led to the development of malalignment of the extensor system. There are several clinical tests to assess patellar mobility but these tests have fair and poor reliability. Radiographs (axial view, CT) and especially MR imaging (kinematic and dynamic axial imaging), as a noninvasive and nonionizing modality, have made significant contribution to the understanding of patellofemoral malalignment (see Fig. 2). The basic principle of therapy consists of correcting the underlying causative factors of patellar malalignment.

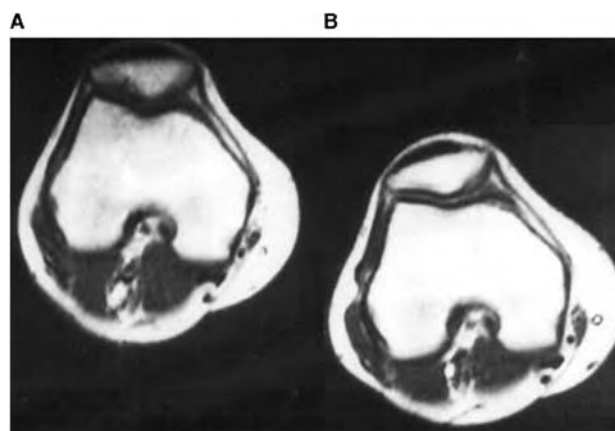
Chondromalacia of the Patella

The term was introduced into the medical literature by Konig in 1924. Chondromalacia, the word that has

Table 1 Overuse injuries of the knee joint

Anterior aspect	Medial aspect
Patellofemoral pain syndrome	Breaststroker's knee
Jumper's knee	Pes anserinus tendinitis/bursitis
Osgood-Schlatter disease	Plica syndrome
Sinding-Larsen-Johansson disease	Semimembranosus tendinitis
Stress fracture of the patella	Medial retinaculitis
Fat pad syndrome	
Posterior aspect	Lateral aspect
Fabellitis	Iliotibial band friction syndrome
Gastrocnemius strain	Popliteal tendinitis
	Bicipital tendinitis

defined the field of patellar pain during the last century, is now mostly a source of confusion and should be abandoned. The presence of soft cartilage (the literal translation of the term chondromalacia) is not strongly correlated with patellar pain. The confusion is further compounded by the use of poorly defined grades (for example, grade III chondromalacia). These types of lesions should be termed chondral or cartilage lesions, and they should be described qualitatively (for example, as blistering or fissuring) and quantitatively (in terms of size, location, and depth). The term chondromalacia patellae should not be used to diagnose patellofemoral pain but, rather, to describe lesions of articular cartilage. The term chondromalacia patellae, although once used as an all-inclusive term for anterior knee pain, is now widely accepted as a term used to describe pathological lesions of the patellar articular cartilage found at arthroscopy. Numerous different categorization systems of chondromalacia of the patella are found in the medical literature, depending on the state and characteristics of the change in the cartilage

**Fig. 1** Schematic representation of the tilt of the patella.**Fig. 2** Magnetic resonance imaging of the patellofemoral joint at (A) 15° and (B) 0° of the flexion.

of the patellofemoral joint and the size and location of the changes. Outerbridge's classification is most frequently used, as well as the arthroscopic categorization by Joensen. Hungerford differentiates among closed chondromalacia, open chondromalacia, chondrosclerosis, surface changes in the cartilage, and development of tufts.

Anterior Knee Pain

The term anterior knee pain is suggested to encompass all pain-related problems, excluding anterior knee pain due to intraarticular pathology. Anterior knee pain is a common complaint in the orthopedic clinic. The differential diagnosis is wide and the principal goal of initial assessment is to detect remediable causes. A clear understanding of the pathophysiology of anterior knee pain is inhibited by the use of imprecise, poorly defined, and often interchanged words, such as malalignment, patellar alignment, maltracking, subluxation, dislocation, and congruence. Although, some physicians compare the diagnosis of anterior knee pain to a "large wastepaper basket," these physicians still agree that it is better to begin treatment, even with this "provisory" diagnosis, and try to determine the precise location of the injury and/or damaged area while monitoring the development of the syndrome. Ireland recognizes mechanical, inflammatory, and mixed causes of pain. According to Fulkerson, use of the term anterior knee pain will underline the need to search further for a specific cause of pain in each patient, remembering that there is not always a correlation between articular lesions and pain.^[2] Anterior knee pain caused by overactivity in an athletic or military population is described by many authors. As we pointed out in the beginning of this entry, a synergistic effect and definite correlation exist among malalignment of the knee extensor system, instability of the

patella, chondromalacia of the patella, and anterior knee pain, and this is the reason why it is better to call all these disturbances according to symptomatology, patellofemoral pain syndrome. The complexity related to the patellofemoral pain syndrome explains the complexity present when treating this syndrome. After a precise diagnosis, the physician should prescribe treatment of the cause of the pain, which can be either nonoperative or operative. If, for example, the primary cause of pain is excessive pressure on the lateral part of the patella caused by malalignment, then the physician can attempt to correct this malalignment with electrostimulation of the vastus medialis muscle and/or by having the patient wear a brace designed for stabilizing the patella. Surgically, treatment consists of full arthroscopic stabilization of the patella.^[3] The same principle is considered for the other causes of the patellofemoral pain syndrome.

PATELLAR TENDINITIS/TENDINOSIS (JUMPER'S KNEE)

Patellar tendinitis/tendinosis is characterized by pathological changes in the quadriceps tendon and its insertion to the proximal pole of the patella, and the patellar tendon (patellar ligament) and its proximal insertion to the apex of the patella or distal insertion to the tibial tubercle. Other terms are also found in the medical literature that are frequently used to describe the same syndrome: patellar apicitis, enthesitis apicis patellae, jumper's knee, volleyball's knee, and basketball's knee. Jumper's knee is a clinical entity most commonly found in athletes who, during their athletic activities, habitually place excessive strain on the extensory system of their knees with numerous jumps or a long period of running. Kujala et al. reported in their retrospective study that during a 5 yr interval in a test sample of 2672 ambulatory patients with various knee injuries, fully 26.4% suffered from complications related to jumper's knee.^[4] Repeated mechanical excessive strain is, without doubt, the basic prerequisite for the development of this syndrome. An analysis of training methods and procedures has indicated that the type of training method used has no significant bearing on the development of jumper's knee. A much more important causative factor was found to be the amount of training (both the amount of time and the amount of mechanical strain placed on the knee) that the athlete habitually carries out. The development of jumper's knee, to a large degree, also depends on the somatic characteristics of the athlete, of which the best known and most intensively studied are the anatomic features of the lower extremities. Although, most of the studies carried out to date have been less interested in the external factors causing jumper's knee, it is quite

certain that excessive strain caused by training and exercising on hard surfaces contributes significantly to the development of this syndrome. With regard to the group of anatomical changes of the lower extremity, the unequal lower extremities and patella alta are causative factors for the development of jumper's knee. We feel confident in concluding that the underlying causes of patellar tendinosis are based on the mechanical characteristics of the tendon and the area of insertion of the tendon into the bone (its elasticity and its ability to stretch). The basic etiopathogenetic course of events leading to the development of overuse syndrome of the quadriceps tendon and the patellar tendon, i.e., jumper's knee, can be defined as the state that arises when mechanical overstrain overcomes the adaptive ability of the excessively strained tissue. Today, it is believed that insertional tendinopathy along with changes in the tendon microstructure (tendinosis) is the mainstay of this condition. We performed an investigation to document pathological changes in bone, the bone–tendon junction, and the patellar ligament in 34 athletes treated surgically.^[1] Pathological changes in the bone were found in one-third of the operated patients, on the bone–tendon junction in two-thirds of patients, and in the patellar ligament pathological changes were found in all operated patients. Pathohistological changes included cellular proliferation within the tendinous tissue accompanied by a prominence of capillary proliferation, chronic inflammatory cell infiltration, hyaline degeneration, mucoid degeneration, microtears of the tendinous tissues, neovascularisation, and tenocyte infiltration, etc. Polarization microscopy showed loss of clear demarcation of collagen bundles accompanied by loss of the normal dense, homogenous polarization pattern. On the basis of our histopathological examinations we support the opinion of the majority of authors that the time has come to change the name of this overuse injury colloquially named, jumper's knee, from patellar tendinitis to patellar tendinosis.

The clinical picture of jumper's knee is characterized by the presence of pain as the basic symptom and a decreased functional ability of the afflicted lower extremity. The pain can appear either in the area of the upper or lower pole of the patella, or in the area of the tibial tuberosity. Pećina and Bojanić reported the appearance of pain on the upper pole of the patella in 10% of their patients, in the area of tibial tuberosity also in 10% of patients, and at the tip of the patella in 80% of athletes suffering from jumper's knee.^[5] The main features of this pain consist of sharpness of varying intensity that gradually evolves and whose appearance is not connected with evident trauma. Some authors have suggested the categorization of jumper's knee into different stages (three or five), based on the prevalence of the main clinical symptoms—pain

Table 2 Classification of jumper's knee according to the main clinical symptoms

Stage 1	Pain only after sports activity
Stage 2	Pain at the beginning of sports activity, disappearing after warm-up and reappearing after activity
Stage 3	Pain at the beginning, during, and after sports activity
Stage 4	Constant pain at rest and during activity Inability to participate in sports
Stage 5	Complete rupture of patellar tendon

and loss of function in the extremity. We prefer categorization or classification in five stages (see Table 2).

Radiographic analysis may reveal bony changes on the poles of the patella and the tibial tuberosity, as well as ossification in the tendon structure. The diagnosis of jumper's knee is facilitated by the use of ultrasound examination of the quadriceps tendon and the patellar ligament.^[6] Some authors have recommended an ultrasound classification of jumper's knee, again in three or five stages (see Fig. 3). To date, there have been a number of reports concerning the success of MR imaging diagnostics of jumper's knee. The authors feel that based on these results they are better able to objectively prescribe the most suitable therapy.

Although jumper's knee is relatively easy to diagnose, its treatment is more difficult. Of the numerous nonoperative treatments available (cryotherapy, anti-inflammatory drugs, stretching exercises and strengthening of the extensor muscle system of the knee, stretching of the hamstring muscles, strengthening of

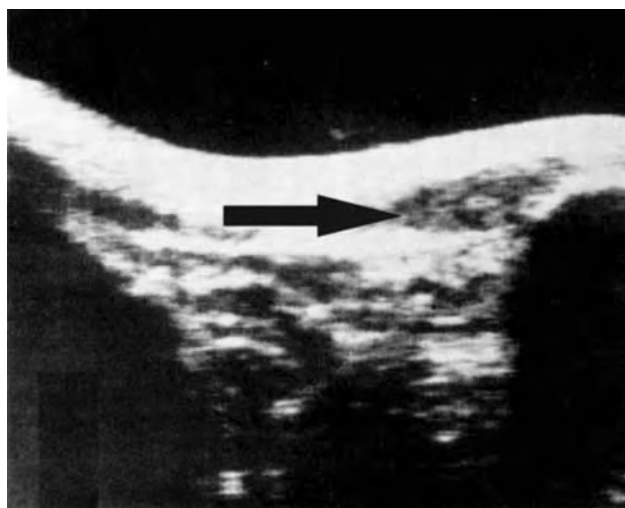


Fig. 3 Longitudinal ultrasound showing focal hypoechoic area (arrow) in the substance of the proximal tendon.

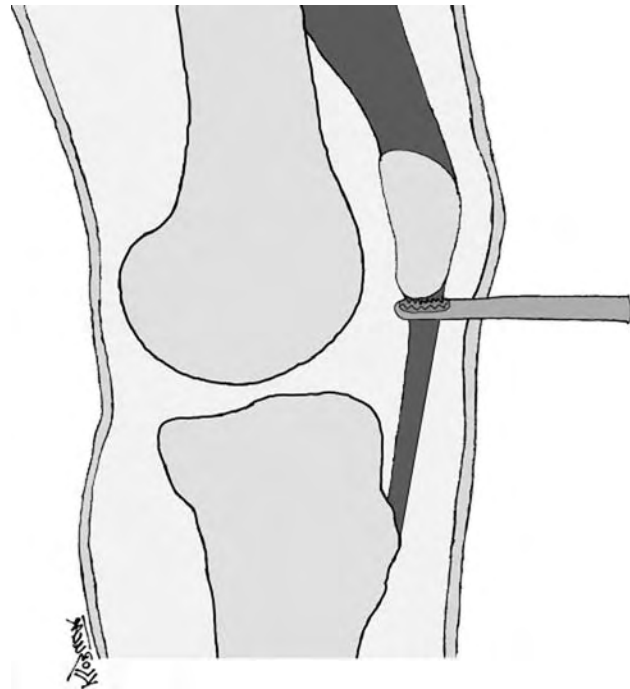


Fig. 4 Schematic representation of arthroscopic patellar apicotomy.

the extensor system of the knee by concentric and eccentric exercises) specific physiotherapy and correction of technical errors are often efficient. Surgical treatment is indicated only if a prolonged and well-supervised conservative treatment program fails. Surgical excision of irreversible lesions, demonstrated on ultrasonography or MR images, are logical procedures, which provide good results.^[7] There are numerous different operative approaches and we prefer the surgical procedure known as arthroscopically assisted approach or fully arthroscopic technique (see Fig. 4), taking into consideration the localization of the pain, knee extensor mechanism alignment, ultrasound, MR imaging, and arthroscopic intraoperative findings.^[1]

ILIOTIBIAL BAND FRICTION SYNDROME

Iliotibial band friction syndrome (ITBFS) results from activity comprising many repetitive flexion and extension movements of the knee, during which rubbing of the band against the lateral femoral epicondyle occurs. This produces irritation and inflammatory reaction within the iliotibial band or the underlying bursa (see Fig. 5).

Iliotibial band friction syndrome is one of the most common overuse injuries in professional runners and also in recreational joggers. The term iliotibial band friction syndrome was introduced by Renne in 1975.^[8] Orava diagnosed ITBFS in 6.4% athletes with overuse

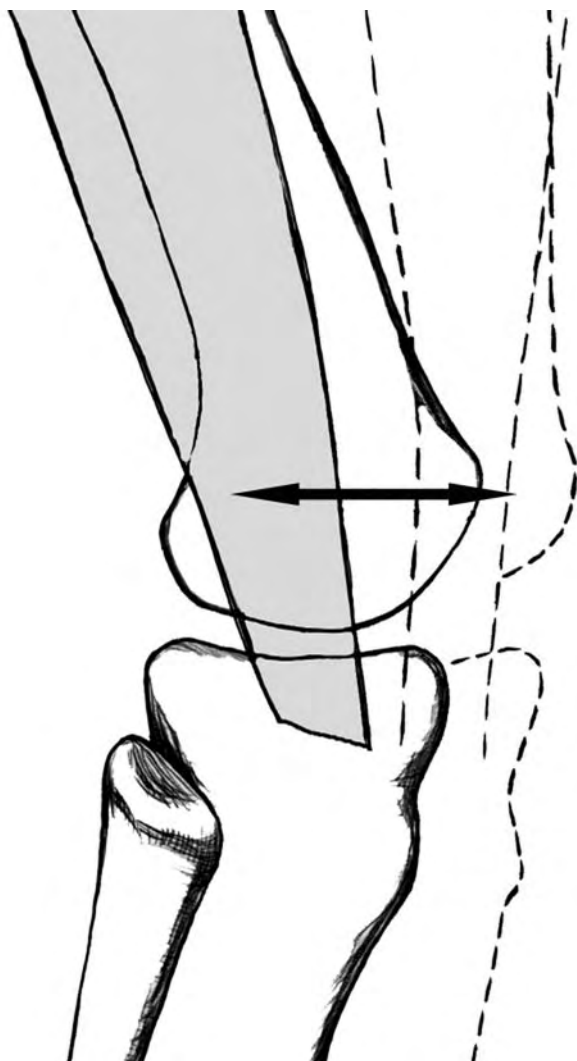


Fig. 5 Mechanism of development of ITBFS.

injuries, and some other authors found a 4.7% frequency of ITBFS in 4173 injured runners.^[1,9] Also it is noted that there is a higher incidence of ITBFS in long-distance runners than in middle-distance runners and sprinters. The iliotibial band (TIB) or tract is a thickened band of fascia extending from the iliac crest down the lateral side of the thigh to its attachment on the lateral tibial condyle (Gerdy's tubercle). With the knee in the extended position, the ITB lies anterior to the lateral femoral epicondyle; in knee flexion of 30° it lies behind the lateral femoral epicondyle. The friction (or impingement) occurs near footstrike, predominantly in the foot contact phase at or slightly below 30° of the flexion, between the posterior edge of the iliotibial band and the underlying lateral femoral epicondyle. The activities entailing many repetitive flexion and extension movements of the knee produce irritation and subsequent inflammatory reactions within the ITB or formation of underlying bursae and secondary inflammation. Most commonly, the ITBFS is a result

of training errors and anatomical malalignment, of the lower extremity. Malalignment conditions of forefoot varus and/or subtalar varus, tibia vara, and genu varum lead to excessive and/or prolonged forefoot pronation during the support phase of the running cycle and this causes ITBFS in more than 50% of cases. Other predisposing factors to ITBFS have been cited: poor ITB flexibility, leg length discrepancies, increased prominence of the lateral femoral epicondyle, hard and/or uneven running surface, and inadequate workout shoes. The patients with ITBFS experience pain at the lateral side of the knee. Pain is stinging in nature, and is located at the lateral femoral condyle 2 cm above the joint line, but it may radiate downward along the iliotibial band, or upward along the lateral side of the thigh. Initially, the pain appears immediately after engagement in activity and usually disappears following a few hours of rest. The next phase is characterized by pain at the very beginning of the activity, its disappearance after warming up, and appearance again on completion of the activity. It is common that the athlete does not seek medical attention in that time and it results in progressive development of the syndrome. In the final stage, the pain impairs even normal walking. The clinical history data are of great importance in the diagnosis of ITBFS, but also various provocative tests for ITBFS have been described. During Noble's test the patient lies in a supine position and the knee is flexed to 90° (see Fig. 6). The examiner holds the ankle of the affected leg in one hand and presses on the lateral femoral epicondyle with the thumb of the other hand. While maintaining constant pressure, the examiner slowly extends the leg. Pain appears when the knee is flexed to 30°, and the patient usually reports that it is the kind of pain experienced when running. Modification of this test according to Pecina consists of putting the lower leg in varus position when performing knee extension. Ober's test is used to diagnose flexibility of the ITB



Fig. 6 Provocation of pain in ITBFS (Noble's test).



Fig. 7 Ober's test.

(see Fig. 7). Ultrasonographic and MR examinations may depict some abnormalities in patients with ITBFS. As a part of differential diagnosis several other conditions should be taken in to consideration: popliteal tendinitis, rupture of the lateral knee joint capsule and/or lateral meniscus, cyst of the lateral meniscus, lesion of the lateral collateral ligament, lateral patellar compression syndrome, but most often the differential diagnosis should be made between the ITBFS and popliteal tendinitis.

The treatment of ITBFS is usually nonoperative: correction of predisposing factors, short-term cessation or modification of athletic activities, ITB stretching exercises, ice massage of the painful area, analgesic/nonsteroidal anti-inflammatory drugs, although infiltration of steroids combined with local anesthetic into the site of the most intense pain is recommended by some authors. When formation of a bursa has been confirmed, the application of steroids injection is also indicated. Surgery is recommended in resistant cases: a transverse cut is made in the 2 cm-long posterior portion of the ITB at the level of the lateral femoral epicondyle and in this way, a V-shaped defect is obtained preventing friction between ITB and lateral femoral epicondyle.

BREASTSTROKER'S KNEE

A high percentage of breaststroke swimmers complained of pain in the medial part of the knee joint. This condition has been termed breaststroker's knee, although further studies showed that breaststroker's knee was diagnosed in swimmers who used the free-style, backstroke, or butterfly stroke. Kennedy and Hawkins believed that a tibial collateral ligament strain on the proximal femoral origin of the ligament, resulting from repeated stretching of the origin during

breaststroke swimming, was the primary disorder leading to breaststroker's knee.^[10] The stretching of the ligament origin is caused by extension of the knee during the "whip kick" phase of breaststroke swimming, accompanied by an excessive valgus stress on the knee joint and the outward rotation of the leg in the final phase of the stroke (see Fig. 8).

In most cases the pain is localized in the lower area of the medial facet of the patella. The inflammation, thickening, and fibrosis of the medial parapatellar plica as potential causes of breaststroker's knee pain are described and arthroscopically confirmed. Arthroscopic findings range from localized medial synovitis of the knee to generalized synovitis and chondromalatic changes on the medial facet of the patella. Slow-motion analysis of the mechanics of the breaststroke showed that the primary cause of breaststroker's knee was the incorrect technique during the "whip kick" phase. Breaststroker's knee symptoms usually appear after 3 yr of competitive swimming. In most cases, both knees are affected and in the beginning the area of the greatest tenderness is localized in the proximal origin of the medial collateral ligament. The later phase is characterized by pain in the lower medial patellar facet area. As with other overuse injuries, the ideal treatment of breaststroker's knee is prevention. Identifying the mechanism of injury and prescribing appropriate management is not easy, unless one has a thorough understanding of proper technique of the four competitive strokes. Proper warm-up exercise, local application of ice, and ultrasound can be helpful. Stretching exercises of the hamstring muscles and strengthening exercises for the medial vastus muscle should be performed as well. Surgical treatment is rarely indicated.



Fig. 8 Biomechanical explanation of the causes leading to the development of breaststroker's knee. Note the stretching of the medial collateral ligament (black arrow) and external rotation of the legs (grey arrows).

OSGOOD–SCHLATTER DISEASE

The most often encountered overuse injuries in children and adolescents are the traction apophysitises, among which the most common and best known is Osgood–Schlatter disease, that is, traction apophysitis of the tibial tubercle. Traction apophysitis of the tibial tubercle develops during the adolescent growth spurt, most often at 11 yr of age in girls and at 13 yr of age in boys. It occurs slightly more often in boys. The clinical picture consists of pain localized to the area of tibial tubercle. The tubercle may be swollen and hypertrophied. Most authors agree that rapid adolescent skeletal growth leads to relative “tightness” of the soft tissues, as skeletal growth is faster than the elongation of the muscle–tendon units. This in turn creates increased tensile forces on the muscles, tendons, and their attachment sites (the apophyses), which cause avulsion fractures of the bony part of the traction apophysis on the tibial tubercle. Physical examination reveals pain during palpation of the tibial tubercle. Resisted extension of the knee from the 90° flexed position will usually reproduce pain, but resisted straight-leg rising is usually painless. The lateral radiograph generally shows the characteristic picture of prominent tibial tubercle with irregularly fragmented ossific nucleus, or a free bony fragment proximal to the tubercle (see Fig. 9).

Nonoperative treatment of Osgood–Schlatter disease, as with other traction apophysites, is based on the



Fig. 9 Osgood–Schlatter disease in a young basketball player.

same principles that apply to all overuse injuries. Today, there is no need for total immobilization or for totally refraining from athletic activities. Corticosteroid injections are not indicated because the condition is not primarily inflammatory and because of the numerous unwanted side effects. Of vital importance is that the physician informs the parents, the coach, and the child athlete of the natural course of Osgood–Schlatter disease. The symptoms of Osgood–Schlatter disease will generally decline and disappear in most patients if nonoperative treatment is carried out long enough, especially after bone growth is terminated. In athletes or physically active young people surgical treatment gives good results in chronic unresolved cases.^[11]

SINDING–LARSEN–JOHANSSON DISEASE

This is a traction apophysitis of the inferior pole of the patella (in rare cases it affects the superior pole as well) and this entity is similar to Osgood–Schlatter disease. Sinding–Larsen–Johansson disease is more common in boys, generally between the ages of 10–15 yr, and is characterized by tenderness of the knee and limping. Clinical examination reveals localized tenderness and swelling over the distal pole of the patella. Lateral radiograph reveals delayed ossification of the inferior pole of the patella. Ultrasound examination gives complete information about the involvement of bone, cartilage, and patellar tendon. The lower pole of the patella appears irregular, fragmented, with swelling of the cartilage and thickening at the insertion of patellar tendon. Treatment is the same as with Osgood–Schlatter disease.

FAT PAD SYNDROME (HOFFA DISEASE)

The primary cause of fat pad syndrome is repeated traumatization of the infrapatellar fatty pad (corpus adiposum Hoffa) during activities that require constant repetition of maximal extension in the knee. It is necessary to distinguish two phases, acute and chronic. Fat pad syndrome is characterized by contact tenderness of the fatty pad, as well as during palpation along the edge of the patellar ligament. Fat pad syndrome is further characterized by the appearance of pain during hyperextension of the knee, the so-called Smillie’s sign (appearance of pain during sudden passive hyperextension of the knee), by the appearance of blockades that are hard to differentiate from meniscal blockades. Magnetic resonance imaging is a useful tool in the study of Hoffa’s fat pad, whose local and systemic involvement is an often ignored cause of anterior knee pain.^[12] In addition, the infrapatellar fat pad may be involved secondary from intrinsic and

extrinsic processes: intracapsular chondroma, localized nodular synovitis, postarthroscopy and postsurgery fibrosis, meniscal and ganglion cyst, Cyclops lesion, etc. In most cases, treatment is nonoperative, but it includes elevation of the heels of the shoes, which reduces, and in some cases completely disables, hyperextension of the knee. In persistent cases arthroscopic resection of the fat pad is performed.

PLICA SYNDROME

Plicae synoviales are normal synovial structures of the knee joint cavity. They are remnants of the mesenchymal tissue that occupies the space between the distal femoral and proximal tibial epiphyses in the 8-week embryo. Incomplete resorption leaves synovial pleats in most of the knee. The infrapatellar plica is the most constant and is also called the mucosum ligament. The suprapatellar plica partially divides the suprapatellar recessus from the knee joint, and the medial plica, which can in some cases be joined to the suprapatellar plica, projects along the anteromedial part of the knee joint to the infrapatellar fatty pad.^[13] The infrapatellar plica was found to be thickened and fibrotic and to have lost in normal elasticity and consequently to be impinged at the intercondylar notch and trochlea, blocking further extension beyond 25° of knee flexion. Arthroscopy allows a very precise assessment of the plica including dynamic examination, in plica syndrome the symptoms are, in most cases, caused by a chronically changed (fibrotic, enlarged, and sometimes calcified) medial (also called mediopatellar) plica. While bending the knee, the affected plica strikes the medial facet of the patella; while extending the knee, it is in contact with the anteromedial part of the femoral condyle (see Fig. 10).

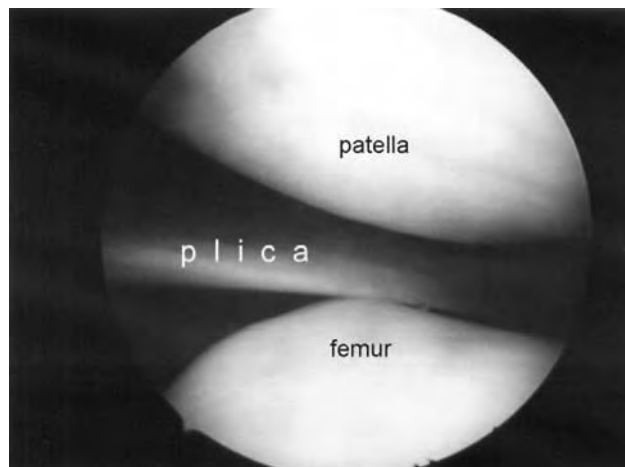


Fig. 10 Arthroscopic view of the symptomatic plica synovialis.

Anamnestic data of repeated audible effects described as explosive snaps and occasional blockades help in diagnosing the plica syndrome. Clinical examination sometimes reveals the plica during palpation, but in most cases the only symptom is tenderness to palpation of the medial facet of the patella and medial femoral condyle. Ultrasound, MR imaging, and arthroscopy are very reliable diagnostic tools. Non-operative treatment includes wearing of special knee braces. Arthroscopic excision of a painful medial plica provides satisfactory relief of symptoms.

SEMIMEMBRANOSUS TENDINITIS

Semimembranosus tendinitis is often an overlooked cause of posteromedial knee pain. The main symptom is pain located at the very posteromedial corner of the knee just below the joint line. During physical examination, pain is produced by applying pressure to the anterior medial tendon of the semimembranosus muscle, immediately below the knee joint, and at the insertion of the posterior medial tendon on the posterior part of the tibial medial condyle. Tenderness is amplified if the knee is in 90° flexion and the lower leg in maximum outward rotation. Differential diagnosis includes ruling out the possibility of medial collateral ligament bursitis, and injury of the medial meniscus. The pes anserinus tendons should also be palpated to rule out the possibility of enthesitis of those tendons and, in some cases, the inflammation of the pes anserinus bursae. In most cases, treatment is non-operative and surgical intervention is rarely indicated consisting of surgical exploration, drilling of the insertion site, and semimembranosus tendon transfer.

PES ANSERINUS TENDINITIS AND BURSITIS

The tendon of the sartorius, semimembranosus, and gracilis muscles insert on the medial plane of the tibia, just below the condyle, forming the pes anserinus (goose foot). Pes anserinus syndrome is frequently found in long-distance runners. Predisposing factors for the development of this syndrome include incorrect training techniques, excessive tightness of the hamstring muscles, valgus alignment of the knee, and excessive rotation of the lower leg in the outward direction.^[14] Bursitis is more frequent and presents fewer problems, both in diagnostics and in treatment. A slight swelling is usually found on the medial side of the knee joint. Ultrasound and MR imaging facilitate correct diagnosis of bursitis. Diagnosing tendinitis (enthesitis) presents more problems, especially regarding differential diagnosis of an internal derangement of the medial compartment of the knee. Magnetic

resonance imaging can be useful in diagnosing pes anserinus tendinitis, and in some cases, scintigraphy is performed to rule out the possibility of stress fracture of the tibia. Treatment is usually nonoperative. Local injection of corticosteroid plus anesthetic is more effective than noninjection therapy.

POPLITEAL TENDINITIS

Popliteal tendinitis is characterized by localized tenderness in the area above the origin of the popliteal muscle on the lateral part of the lateral femoral condyle.^[15] Excessive and/or extended pronation of the foot during running is considered to be a predisposing factor in developing this syndrome. Popliteal tendinitis or popliteal tenosynovitis is one of the many injuries to the knee joint that are characterized by tenderness in the lateral area of the knee. To localize the exact area of greatest tenderness, the affected leg is placed in the so-called figure of 4 position where the popliteus muscle tendon and the lateral collateral ligament are extended and thus accessible to palpation (see Fig. 11). Pain is produced by applying pressure with a finger to the area of the insertion of the popliteus muscle located slightly anteriorly and distally from the origin of the lateral collateral ligament. Treatment is usually nonoperative and carried out according to the principles



Fig. 11 Figure of 4—diagnostic test for popliteal tendinitis.

set down for treating enthesitis. Surgical treatment is rarely performed.

FABELLA SYNDROME (FABELLITIS)

The fabella is a sesamoid bone located on the posterior side of the lateral femoral condyle in the lateral head of the gastrocnemius muscle. It is present in 10–18% of the population. Injuries to the fabella can cause tenderness and sensation of radiating pain in the lower leg. Although fracture and dislocation of the fabella is described in athletes, fabella syndrome is mostly the direct result of long-lasting microtraumas.^[16] Clinical manifestations are characterized by the gradual onset of pain that radiates into the lower leg, specifically the peroneus communis nerve dermatoma. The pain is usually moderate in intensity and associated with knee movements, either because of extension of the muscles during hyperextension of the knee or because of direct pressure on the fabella, e.g., when sitting with legs crossed. The fabella syndrome appears to be unique to late adolescence, that is 15–17 yr of age. Differential diagnosis must rule out the possibility of rupture of the posterior horn of the lateral meniscus and also tendinitis of the lateral head of the gastrocnemius muscle, the biceps femoris, and the popliteus muscle. If the usual nonoperative treatment, including infiltration of anesthetics with corticosteroids, does not produce improvement, surgical treatment, consisting of extirpation of the fabella is indicated.

CONCLUSIONS

Because of its exceptional vulnerability, and its extremely complex anatomic and biomechanical structure, the knee joint, like no other joint in the human body, poses great challenges to a treating physician. The knee joint is a frequently injured joint in athletic and recreational activities, and according to some statistics, one-half, or possibly even more, of all sports-related injuries are incurred in the knee joint. Although most of the injuries of the knee joint are acute ones, such as meniscal or anterior cruciate ligament ruptures, overuse injuries are becoming more and more present in overall knee pathology. The reasons for this are that the knee joint participates in all sports activities (running, jumping, kicking, etc.), and also the joint area is characterized by numerous attachment sites for muscles and their tendons as well as by numerous bursae. It is also because the specific joint between the patella and the femur (articulatio femoropatellaris) constitutes a part of the knee joint. Overuse injuries of the knee joint deserve special attention because, though very common, they are complex in

terms of their detection and differential diagnosis. Thorough knowledge of epidemiology, etiology, pathophysiology and diagnostic modalities is crucial to diagnose these injuries correctly. As with most of the overuse injuries, the treatment is usually conservative and the concept is known as “relative rest.” If the conservative treatment fails, numerous surgical options, including minimally invasive procedures such as arthroscopy and endoscopy, are available to the treating physician.

ARTICLE OF FURTHER INTEREST

Tissue Engineering of Cartilage

REFERENCES

1. Pecina, M.; Bojanic, I. *Overuse Injuries of the Musculoskeletal System*, 2nd Ed.; CRC Press: Boca Raton, 2003; 189–254.
2. Fulkerson, J.P. *Disorders of the Patellofemoral Joint*, 3rd Ed.; Williams and Wilkins: Baltimore, 1997; 137–153.
3. Haspl, M.; Cicak, N.; Klobucar, H.; Pecina, M. Full arthroscopic stabilization of the patella. *Arthroscopy* **2002**, *18* (1), E2 (1–3).
4. Kujala, M.H.; Kvist, M.; Osterman, K. Knee injuries in athletes, review of exertion injuries and retrospective study of outpatient sports clinic material. *Sports Med.* **1986**, *3*, 447–460.
5. Pecina, M.; Bojanic, I. Surgical treatment of jumper’s knee in top level athletes. In *Sports, Medicine and Health*; Hermans, G.P.H., Ed.; Elsevier: Amsterdam, 1990; 299–304.
6. Fritschy, D.; Gautard, R. Jumper’s knee and ultrasonography. *Am. J. Sports Med.* **1988**, *16*, 637–640.
7. Fritschy, D.; Wallensten, R. Surgical treatment of patellar tendinitis. *Knee Surg. Sports Traumatol. Arthrosc.* **1993**, *1*, 131–133.
8. Renne, J.W. The iliotibial band friction syndrome. *J. Bone Joint Surg.* **1975**, *57A*, 1110–1111.
9. Orava, S. Iliotibial tract friction syndrome in athletes—an uncommon exertion syndrome on the lateral side of the knee. *Br. J. Sports Med.* **1978**, *12*, 69–73.
10. Kennedy, J.C.; Hawkins, R.J. Breaststroker’s knee. *Phys. Sportsmed.* **1974**, *2*, 33–38.
11. Orava, S.; Malinen, L.; Karpakka, J.; Kvist, M.; Leppilahti, J.; Rantanen, J.; Kujala, V.M. Results of surgical treatment of unresolved Osgood-Schlatter lesion. *Ann. Chir. Gynecol.* **2000**, *89*, 298–302.
12. Saddik, D.; McNally, E.G.; Richardson, M. MRI of Hoffa’s fat pad. *Skeletal Radiol.* **2004**, *33*, 433–444.
13. Dupont, J.Y. Synovial plicae of the knee. Controversies and review. *Clin. Sports Med.* **1997**, *16*, 87–122.
14. Uson, J.; Aguado, P.; Bernad, M.; Mayordomo, L.; Naredo, E.; Balsa, A.; Martin-Mola, E. Pes anserinus tendino-bursitis: what are we talking about? *Scand. J. Rheumatol.* **2000**, *29*, 184–186.
15. Faletti, C.; De Stefano, N.; Giudice, G.; Larciprete, M. Knee impingement syndromes. *Eur. J. Radiol.* **1998**, *27*, 60–69.
16. Robertson, A.; Jones, S.C.; Paes, R.; Chakrabarty, G. The fabella: a forgotten source of knee pain? *Knee.* **2004**, *11*, 243–245.

Knee Joint Replacement

M. J. Cross

*Australian Institute of Musculo-Skeletal Research, and University of Sydney, Sydney,
New South Wales, Australia*

E. N. Parish

S. Machan

Australian Institute of Musculo-Skeletal Research, Sydney, New South Wales, Australia

K

INTRODUCTION

Total knee replacement (TKR) has been around in some form for over fifty years. The complexities of the knee joint began to be understood only thirty years ago. Because of this, TKR was initially not as successful as Sir John Charnley's artificial hip. However, over the last twenty years, dramatic advancements in the knowledge of knee mechanics have led to modifications of the design of the artificial knee, which appear to be durable. As with most techniques in modern medicine, the envelope is constantly expanding so that increasing numbers of patients are receiving the benefits of knee arthroplasty.

Knee joint replacement is a safe operation with good success rates. To date, cemented fixation has been the gold standard for TKR with excellent long-term results. Medium- and long-term results from cementless TKR have proven to be satisfactory. The effectiveness of bioactive coatings in augmenting TKR fixation has achieved positive early results, and bioactive coatings may be an alternative to cement in achieving long-term durability.

HISTORY

In the 1860s Fergusson reported performing a resection arthroplasty of the knee for arthritis, and Verneuil is thought to have performed the first interposition arthroplasty using joint capsule. Other substances were subsequently tried including skin, muscle, fascia, fat, and pig bladder.

The first artificial implants were tried in the 1940s as molds fitted to the femoral condyles (following similar designs in the hip). In the next decade tibial replacement was also attempted but both designs led to problems with loosening and persistent pain. Combined femoral and tibial articular surface replacements appeared in the 1950s as simple hinges. These implants failed to account for the complexities of knee motion and consequently had high failure rates.

In 1971, Gunston^[1] importantly recognized that the knee does not rotate on a single axis like a hinge.

Gunston showed that the femoral condyles roll and glide on the tibia with multiple instant centers of rotation. His polycentric knee replacement had early success with its improved kinematics over hinged implants, but failed because of inadequate fixation of the prosthesis to bone.

The highly conforming and constrained Geomedic knee arthroplasty,^[2] introduced in 1973 at the Mayo Clinic, ignored Gunston's work and gave rise to a kinematic conflict. Knee prosthesis designs since then have either followed Gunston's principle in attempting to reproduce normal knee kinematics, or have provided a more conforming articulation to govern knee motion.

A variety of condylar prostheses have since been designed. The Total Condylar prosthesis designed by Insall in 1973^[3] was one of these. This prosthesis concentrated on mechanics and did not try to reproduce normal knee motion. Ranawat et al.^[4] reported 94% survivorship at 15-year follow-up, which is the most impressive reported to date; however, range of motion was not recorded. The component was subsequently redesigned to more closely reproduce the normal kinematics of the knee and to improve range of motion of the component. At the same time, a prosthesis with more natural kinematics was developed at the Hospital for Special Surgery, relying on the retained cruciate ligaments to provide knee motion.

The argument as to whether the posterior cruciate ligament should be preserved or sacrificed goes on to this day. Long-term follow-up studies do not show any significant differences, although gait has been shown to be more normal, especially when walking up and down stairs, if ligaments are preserved.^[5] One approach to incorporating normal kinematics and maximal conformity is with mobile tibial bearings. Long-term follow-up of these prostheses has so far shown encouraging results with predictable long-term survival.^[6]

DEFINITION OF PROBLEM

Patients with painful, deformed, and unstable knees secondary to degenerative or inflammatory conditions need a prosthesis that will provide reproducible relief

of pain and improvement in function. Morbidity and complications from the procedure should be minimal. The complexities of a normal knee joint however, are not easily reproduced with modern techniques and patients should understand that the prosthesis will not be a normal knee. The prosthesis should be durable, requiring patients to undergo only one definitive procedure in their lifetime, although in younger patients, this may be unrealistic. Revision procedures should always be contemplated, and bone preservation is mandatory in the initial procedure.

Indications for Surgery

The primary indication for TKR is to relieve pain caused by severe arthritis. Correction of significant deformity is an important indication, but is rarely used as the primary indication for surgery. Before surgery is considered, all conservative treatment measures should have been exhausted.

Knee replacement has a finite expected survival that is adversely affected by activity level. It is generally indicated in older patients with more modest activity levels, but is also clearly indicated in younger patients who have limited function because of systemic arthritis with multiple joint involvement. Young patients requesting knee replacement, especially those with post-traumatic arthritis, are not excluded by age but must be significantly disabled and must understand the inherent longevity of joint replacement. Rarely, severe patellofemoral arthritis may justify arthroplasty because the expected outcome of arthroplasty is superior to patellectomy. Deformity can sometimes become the principal indication for knee replacement in patients with moderate arthritis when flexion contracture or varus or valgus laxity is significant.

Contraindications

Absolute contraindications to TKR include knee sepsis including previous osteomyelitis, a remote source of ongoing infection, extensor mechanism dysfunction, severe vascular disease, recurvatum deformity secondary to muscular weakness, and the presence of a well-functioning knee arthrodesis. Relative contraindications include medical conditions that preclude safe anesthesia, the demands of surgery and rehabilitation, a neuropathic joint, obesity, or skin conditions within the field of surgery such as psoriasis.

RELEVANT ANATOMY

Movement of the knee joint can be classified into six degrees of freedom—three translations (anterior/

posterior, medial/lateral, and inferior/superior) and three rotations (flexion/extension, internal/external, and abduction/adduction). The movements of the knee joint are determined by the shape of the articulating surfaces of the tibia and femur and the orientation of the four major ligaments of the knee joint (the anterior and posterior cruciate ligaments and the medial and lateral collateral ligaments). The two cruciate ligaments create a four-bar linkage system that maintains a fixed femoral and tibial relationship throughout flexion.

Knee flexion and extension involves a combination of rolling and sliding, termed femoral rollback, an ingenious way of allowing increased ranges of flexion. Because of asymmetry between the lateral and medial femoral condyles, the lateral condyle rolls a greater distance than the medial condyle during 20 degrees of knee flexion. This causes a coupled external rotation of the tibia, which is described as the screw-home mechanism that locks the knee into extension.

The primary function of the medial collateral ligament is to restrain valgus rotation of the knee joint, with secondary function being control of external rotation. The lateral collateral ligament restrains against varus rotation and resists internal rotation.

The primary function of the anterior cruciate ligament (ACL) is to resist anterior displacement of the tibia on the femur when the knee is flexed, and to control the screw-home mechanism of the tibia in terminal extension of the knee. A secondary function of the ACL is to resist varus or valgus rotation of the tibia, especially in the absence of the collateral ligaments. The ACL also resists internal rotation of the tibia. The main function of the posterior cruciate ligament (PCL) is to allow femoral rollback in flexion and to resist posterior translation of the tibia relative to the femur. The PCL also controls external rotation of the tibia with increasing knee flexion. Retention of the PCL in TKR has been shown biomechanically to provide a more normal kinematic rollback of the femur on the tibia. This is also important for improving the lever arm of the quadriceps mechanism with flexion of the knee.

The movement of the patellofemoral joint can be characterized as rotating, gliding, and sliding. During flexion of the knee, the patella moves distally on the femur. This movement is governed by its attachments to the quadriceps tendon, ligamentum patellae, and the anterior aspects of the femoral condyles. The muscles and ligaments of the patellofemoral joint are responsible for producing extension of the knee. The patella acts as a pulley in transmitting the force developed by the quadriceps muscles to the femur and the patellar ligament. It also increases the mechanical advantage of the quadriceps muscle relative to the instant center of rotation of the knee.

The mechanical axis of the lower limb is an imaginary line through which the weight of the body passes. It runs from the center of the hip to the center of the ankle through the middle of the knee. This is altered in the presence of deformity and must be reconstituted at surgery. This allows normalization of gait and protects the prosthesis from eccentric loading and early failure.

PROSTHETIC DESIGN FACTORS

A typical total knee replacement has four components (Fig. 1). These include a femoral component (usually cobalt–chromium alloy); a tibial component (either titanium or cobalt–chromium alloy); a meniscal insert (polyethylene), which can be fixed or mobile; and a patella-resurfacing component (polyethylene). The design and material used in each component are important when attempting to recreate the kinematics of a normal knee joint. Cobalt-based alloy (CoCrMo) is said to be a reliable material that meets the prerequisites of biocompatibility and durability that make it suitable for use in joint prostheses of the knee, hip, and shoulder.^[7,8]

Femoral and Tibial Component Design Factors

A functional design that also reproduces the kinematics of a normal knee joint is the goal of all prosthetic designs. In order to do this, the femoral component must recreate the correct anatomical sizes and shapes of the femur. If the PCL is to be retained, adequate replacement of the resected femoral condyles and tibial cuts is required to maintain full flexion and femoral rollback (Fig. 2). Excessive or insufficient thickness may result in loss of flexion and absence of rollback. Andriacchi et al.^[5] compared the kinematic characteristics of five different designs of knee replacement and found that all resulted in abnormalities during gait, although the cruciate-retaining designs produced a more normal range of motion during stair climbing. Abnormalities



Fig. 1 Location of the (A) femoral and (B) tibial beaded component surfaces.



Fig. 2 Flexion obtained in a TKR with a fixed bearing.

during knee motion following TKR are also the result of interactions of surrounding soft tissue structures and efforts to achieve a suitable balance.

The decision to conserve or sacrifice the PCL has important consequences when considering component design and patient function outcome. Advantages of keeping the PCL include preservation of proprioception, maintenance of femoral rollback, prevention of posterior subluxation, and greater reproduction of the kinematics of a normal knee. Disadvantages include joint looseness or tightness and decreased meniscal insert conformity resulting in increased contact stresses due to femoral rolling and sliding. Advantages of sacrificing the PCL include a technically easier operation, increased conformity of the articular surface resulting in increased anterior-posterior stability, and no sliding. Disadvantages are reduced range of motion due to absence of rollback; increased constraint related to conformity of the meniscal insert, which increases stress at the fixation interface on the tibia; and altered kinematics and proprioception of the knee.

Tibial component design is also important, especially as concerns fixation. Tibial loosening is a major cause of failure in total knee arthroplasty due to tibial lift-off and subsequent subsidence. Many factors influence the initial stability of an uncemented tibial tray, including implant design, geometry of the bearing surfaces, quality of the underlying cancellous bone, and surgical technique. In an uncemented knee arthroplasty, stable initial fixation is a prerequisite of successful bone ingrowth and long-term fixation (Fig. 3). To allow successful osseointegration, screw fixation must elicit less than 100 μm initial motion,^[9–11] with best fixation occurring with the use of four screws. Other designs of tibial components incorporate stems and cement in order to further enhance fixation and prevent lift-off and shear. In fact, it has been shown that short or long stems do not enhance fixation with cemented or



Fig. 3 Machine setup for testing tibial plate micromotion with four-screw fixation.

uncemented arthroplasty,^[12] and that use of a stem results in reduced metaphyseal bone density.^[13]

Cortical coverage is important when designing and implanting the tibial component. Subsidence can occur when inadequate cortical coverage is achieved due to an undersized tibial component, especially in the uncemented knee. A broad variety of sizes should be available to allow adequate cortical coverage without affecting the congruence of femoral component articulation.

Meniscal Insert Material and Design Properties

The function of the meniscal insert is to provide effortless articulation between the femur and tibia and to allow for correct anatomical function of the joint. Meniscal insert failure is the primary reason for total knee replacement failure.^[14] The metal femoral component is 100 to 200 times stiffer than the plastic meniscal insert, and hence it acts like a rigid indenter, deforming the plastic.^[15] Consequently, the design of the meniscal insert is critical to the longevity and success of the prosthesis. The choice of material and manufacturing processes used, as well as the geometry of the meniscal insert, are the key design factors in reducing stress in the component, hence increasing its lifetime.

The principle reasons for failure of meniscal inserts are delamination and pitting.^[16] Delamination is caused by tiny cracks forming under the articulating



Fig. 4 Meniscal inserts demonstrating delamination.

surface. These cracks start to propagate, running parallel to the surface and resulting in sheets of the material laminating away.^[17] Pitting is a similar mechanism, where the cracks start at the surface and propagate inward. Both types of failure—especially delamination due to excessive amounts of material breaking away at one time—are detrimental to the function of the joint (Fig. 4). Failure can lead to infection in the joint (because of the biological reaction to debris particles) and loosening of the prosthesis.^[18] To increase the longevity of the meniscal insert, stresses in the component must be reduced and manufacturing processes refined to reduce defects in the material that may predispose it to failure.

The standard material currently used for the meniscal insert is medical-grade, ultra-high, molecular-weight polyethylene (UHMWPE). UHMWPE is a biocompatible material with excellent wear-resistant characteristics and a very low coefficient of friction with cobalt–chromium alloy, commonly used for the femoral component. The coefficient of friction between UHMWPE and cobalt–chromium alloy has been reported at between 0.03 and 0.16 with excellent wear rates.^[19] Other polymeric materials have been tried in the past to increase the strength of the articulating component, but UHMWPE has surpassed other materials with excellent wear rates and fatigue resistance.^[19] More recently, the strengthening of UHMWPE by cross-linking has been studied to achieve a stronger material with higher wear resistance. Although high wear resistance has been achieved, cross-linking produces a more brittle material, detrimental to fatigue resistance properties and hence lowering longevity.^[20]

The manufacturing and sterilization processes used to make meniscal inserts also affect the longevity of the component. UHMWPE meniscal inserts were originally manufactured by heat pressing, a process that heats the material under compression and allows it to flow into all cavities of the mold, where it then cools.

This process produces inconsistent cooling of the material. Significant grain boundaries are consequently formed that are weak points in the material.^[21] These imperfections are key regions where the delamination process can initiate. More successful is the recent use of machining UHMWPE from ram-extruded or compression-molded rods to manufacture the components. This process produces material with minimal evidence of grain boundaries and consequently higher fatigue resistance.^[21] Two main types of sterilization are used for medical purposes. The most common is gamma irradiation. This process has been shown to oxidate the material, which degrades the density of UHMWPE over time.^[22] Because this is detrimental to the fatigue resistance of UHMWPE,^[23] ethylene oxide sterilization is now preferred as an alternative.

The geometry of the meniscal insert is critical to not only the function of the joint, but also to the stresses produced in the component and consequently to the longevity of the prostheses. Two key design factors concerning the geometry of the component are the thickness of the meniscal insert and its conformity with the femoral component. Studies have shown that greater meniscal insert thickness results in lower stresses, with minimum thicknesses below 6 mm causing excessive stresses. The problem with increasing meniscal insert thickness is maintaining anatomically correct joint space, accomplished by sacrificing proximal tibial bone.^[15] A balance must therefore be achieved to maximize meniscal thickness, reduce bone loss, and maintain joint space. Increasing conformity of meniscal insert to femoral component increases contact area and consequently reduces contact stresses.^[18] Because the kinematics of the knee joint is very complex, more conforming designs will likely restrict the natural movement and function of the joint. Again, balance must be attained made between conformity and function of the knee joint.

Mobile-Bearing Insert

The mobile-bearing insert in total knee arthroplasty is designed to move freely in the joint during knee motion. There are generally two types of mobile bearing: the semiconstraining, rotating platform has a central peg that allows rotation but no translations; the other, less constraining, self-centering mobile bearing allows both rotation and translation. Both types sit on a polished tibial surface and are highly conforming to the femoral component (Fig. 5). The potential benefits of mobile meniscal-bearing total knee prostheses are good mobility, low contact stress, and low shear stress on the components, thus reducing polyethylene wear and implant failure.^[24-29]



Fig. 5 Polished meniscal-bearing tibial plates.

The perceived advantages of meniscal-bearing TKRs are the potential for reduced polyethylene wear while allowing unrestrained tibiofemoral movement^[28,30,31] and increased range of movement (Fig. 6) compared to similar designs with fixed bearing. Mobile-bearing TKRs have significantly reduced upper- and lower-surface stresses compared with fixed-bearing components.^[32] The mean rate of polyethylene penetration of congruent meniscal bearings is less than that of fixed bearings and has been estimated in vitro to be as low as 0.05 mm^[33] to 0.01 mm^[28] per year.

By allowing completely unconstrained motion with a mobile-bearing system, as compared to central peg systems and systems where the bearings are constrained by tracks, high-shear stresses in the polyethylene and at the bone/component interface are avoided. However, there is an increased possibility for subluxation of the meniscal insert with these designs, although the incidence of this is very low.^[29] Knees undergoing replacement are diseased and their individual kinematics



Fig. 6 Flexion obtained in a TKR with meniscal bearing. (Note femoral rollback.)

is impossible to forecast. As a result, it is essential that the mobile bearing be allowed to find its own place and range of motion to suit the knee in which it has been implanted. In order to achieve this, the bearing should be self-centering. It is essential to have equal flexion and extension gaps when using mobile bearings.

Femoral Morphology and Component Rotation

Femoral morphology is commonly assumed to be relatively consistent when performing total knee arthroplasty.^[34] However, recent work has demonstrated three distinct shapes of the distal femur: ectomorphic, mesomorphic, and endomorphic.^[35] This variation in morphology affects the choice of femoral components. To avoid undersizing the prosthesis, prosthesis kits in the future may offer narrow as well as wide femoral components, to accommodate the variation in femoral morphology.

Bone cuts are made to prepare the distal femur to accept the femoral component using anatomical landmarks for reference. Correct varus–valgus alignment is usually achieved with either an intramedullary guide against the anatomical axis or an extramedullary guide to demonstrate the mechanical axis. Femoral component rotation is also dependent on the shape of the distal femur.

Accurate rotation of the femoral component in total knee arthroplasty is important for normal patella tracking, symmetrical patellofemoral joint contact, neutral varus–valgus positioning in flexion, correct rotational alignment of the tibia in extension, and avoidance of anterior femoral notching.^[36,37] Internal rotation of the femoral component causes a shift into valgus alignment with flexion and an increase in the quadriceps angle (commonly known as the Q-angle) with harmful effects on patella tracking. Internal rotation of the femoral component also causes differences in the flexion and extension gaps by altering the relative dimensions of the posterior condyles in flexion. Flexion then causes asymmetrical tension across the prosthesis and gapping on the lateral side.^[36]

COMPONENT FIXATION

Cemented and Cementless Fixation in TKR

There has been much debate in recent years regarding the optimal method of fixation in TKR. The consensus generally accepts cement fixation as the ideal, with randomized studies revealing longer survival rates with few complications.^[4,38–43] However, more recent studies comparing cemented and cementless techniques have revealed few differences, if any.^[39,44–47] Unlike with the

hip, early use in TKR of a cementless technique was unsuccessful, with high failure rates due to inadequate fixation and failure of bone ingrowth onto porous surfaces. A poor prosthetic design rationale and use of metal-backed patellae also resulted in the cementless prosthesis being associated with high rates of failure, raising concerns over use of such prostheses.

The long-term survival of TKR fixation using polymethylmethacrylate (PMMA) cement is well-published with excellent results confirming satisfactory functional scores and prosthesis survival greater than ninety % after more than 10 years.^[4,38–43] Many surgeons choose to use a cemented fixation due to its decreased technical difficulty with implantation and reliable long-term results. Cement, however, has been indicated as a major factor in prosthetic loosening due to its limited fatigue life, poor stress properties, and association with osteolysis. The presence of cement disease has been reported with hip replacements where, in younger patients, cement can be implicated as the major cause of aseptic loosening due to its unnatural biomaterial properties.^[48] The question thus arises whether there is a need to use cement or not.

Porous coated prostheses were first used in the 1980s as an alternative to cemented components in an attempt to overcome the problem of aseptic loosening. Ingrowth of the prosthesis occurs when the porous coated surface of the prosthesis is applied to a living bone surface with an initially stable interface with minimal micromotion. The porous surface is penetrated with hematoma and inflammatory cells following implantation. As early as 4 weeks postimplant, the porous surface is replaced with early bone structures, and by 6 months postimplant, mature bone formation occurs.^[34]

Medium- and long-term results from cementless TKR have proven satisfactory.^[39,49,50] Theoretically, cementless fixation has the advantages of less third-body wear, decreased maintenance, better restoration of bone stock, and safer implantation after bacterial arthritis. Long-term studies reveal little difference in survival rates between cemented and cementless designs, leading to an increase in the popularity of cementless fixation.^[26,46,51] Comparable results with cemented and cementless design, theoretically longer and more durable outcomes, and the long-term efficacy of calcium phosphates (i.e., hydroxyapatite) that enhance fixation have led to the continuing increase in the use of cementless knee arthroplasty prostheses.

Hydroxyapatite (HA) in TKR

Over the last decade simple press-fit total knee prostheses have not demonstrated enough intrinsic stability, and ingrowth into porous coated implants has been inconsistent, possibly because of inexact bone

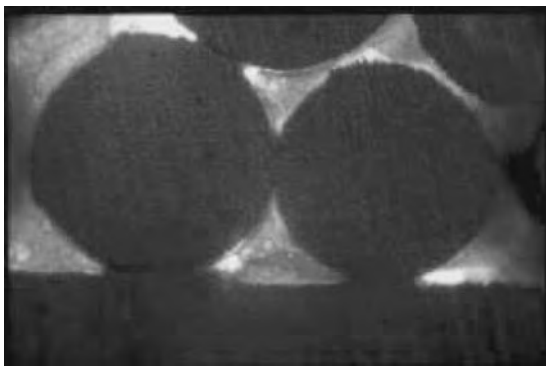


Fig. 7 Evidence of osseointegration achieved with hydroxyapatite (HA) coating.

cuts. The problem with some early cementless designs was that they allowed inadequate tibial cortical coverage; the tibial inserts therefore often sank. Preliminary reports,^[40,52–54] suggested that HA-coated implants could improve stability and fixation of cementless TKR.

The effectiveness of hydroxyapatite (HA) in augmenting TKR fixation has been proven.^[55–57] The presence of HA has been shown to encourage bone growth onto the porous coated prostheses (Fig. 7), as indicated by the reduction of lucent lines when compared to similar uncemented prostheses.^[58] Hydroxyapatite coating has been shown to enhance fixation in total hip replacement,^[59,60] however, its role in TKR has only recently been demonstrated. Early results with bioactive coatings on prosthetic interfaces have been encouraging. Growing concern over cement fixation and so-called cement disease has led to increased use of cementless fixation in TKR.

FUNCTION AFTER TKR

Many knee replacement systems have demonstrated adequate functional results and range of motion at middle- to long-term follow-up. Most follow-up studies' functional scoring systems use pain, ability to walk, ability to ascend and descend stairs, and use of a walking aid as measures of outcome.^[61,62] Kneeling ability is usually not considered, although some scoring systems, derived from patient questionnaires, have included kneeling as a measure of knee function.^[61,63,64]

The kneeling position is important in many activities of daily living and is important in some vocations.^[65,66] The kneeling position has been shown to be a predisposing factor for osteoarthritis of the knee, and so preselects patients who will need to kneel after TKR. Kneeling has also been shown to be an important

intermediate body position by which older adults rise from the floor or garden.^[67] Many patients ask questions about kneeling ability before TKR surgery. It is difficult to answer this inquiry because of the dearth of studies on the subject.^[68,69] One study reveals that 70% of patients were able to kneel with minimal discomfort following surgery.^[69]

A number of reports in the literature indicate that up to 56% of patients return to sports after total hip replacement.^[70,71] There are only a few studies on patients returning to sports after TKR, even though knee replacement rates approach those of hip replacement. One study^[72] reports up to 77% of patients able to return to sports following TKR. With increasing implant survivorship, patient expectation of a good, functional result will also rise.

CONCLUSION

The majority of patients are satisfied with their knee replacements. If relief of pain is the main indication for surgery, then this should indeed be the case. Good knee function is usually restored after TKR, and the majority of patients are able to return to low-impact sporting activity. Long-term studies confirm satisfactory functional scores, and show 91–96% prosthesis survival at 14–15-year follow-up. The ideal prosthesis will eventually be cementless with minimal friction because of highly polished surfaces, mobile bearing, and minimal bone resection, giving rise to a normal range of motion and the kinematics of a normal joint.

ARTICLES OF FURTHER INTEREST

Alumina; Bone Cement; Implant, Total Hip; Tendons and Ligaments; Tissue Engineering of Cartilage; Tissue Engineering of Ligament; Tissue Engineering of Tendons; Ultra-High Molecular Weight Polyethylene (UHMWPE); Zirconia Ceramics

REFERENCES

1. Gunston, F.H. Polycentric knee arthroplasty. *J. Bone Jt. Surg.* **1971**, *53-A*, 272–275.
2. Coventry, M.B.F.G.A.M.; Riley, L.H.; et al. A new geometric knee for total knee arthroplasty. *Clin. Orthop.* **1972**, *94*, 171–184.
3. Insall, J.N.S.W.N.; Ranawat, C.S. The total condylar knee prosthesis. A report of two hundred and twenty cases. *J. Bone Jt. Surg.* **1979**, *61-A*, 173–180.
4. Ranawat, C.S.F.W.F.; Saddler, S.; et al. Long-term results of the total condylar knee arthroplasty: A 15-year survivorship study. *Clin. Orthop.* **1993**, *286*, 94.

5. Andriacchi, T.P.G.J.O.; Fermier, B.S.; Illinois, C. The influence of total knee replacement design on walking and stair climbing. *J. Bone Jt. Surg.* **1982**, *64-A* (9), 1328–1335.
6. Buechel, F.F. Long-term followup after mobile bearing total knee replacement. *Clin. Orthop.* **2002**, *404*, 40–50.
7. Marti, A. Cobalt-base alloys used in bone surgery. *Injury* **2000**, *31* (Suppl. 4), 18–21.
8. Dearnley, P.A. A review of metallic, ceramic and surface-treated metals used for bearing surfaces in human joint replacements. *Proc. Inst. Mech. Eng., H: J. Eng. Med.* **1999**, *213*, 107–135. (Part H).
9. Engh, C.A.O.C.D.; Jatsy, M.; McGovern, T.F.; Boby, J.D.; Harris, W.H. Quantification of implant micromotion, strain shielding, and bone resorption with porous-coated anatomic medullary locking femoral prostheses. *Clin. Orthop.* **1992**, *285*, 13–29.
10. Pilliar, R.M.L.J.M.; Maniopoulos, C. Observations on the effect of movement on bone ingrowth into porous-surfaced implants. *Clin. Orthop.* **1986**, *208*, 108–113.
11. Volz, R.G.N.J.K.; Lee, R.W.; McMurtry, M.G. The mechanical stability of various noncemented tibial components. *Clin. Orthop.* **1988**, *226*, 38–42.
12. Stern, S.H.D.W.R.; Gilbert, J.L. The effect of tibial stem design on component micromotion in knee arthroplasty. *Clin. Orthop.* **1997**, *345*, 44–52.
13. Lonner, J.H.K.M.; Levitz, C.; Lotke, P.A. Changes in bone density after cemented total knee arthroplasty: Influence of stem design. *J. Arthroplast.* **2001**, *16* (1), 107–111.
14. Heim, C.S.; Postak, P.D.; Greenwald, A.S. Chapter 34: Factors Influencing the Longevity of UHMWPE Tibial Components. In *Instructional Course Lectures*; AAOS, 1996; Vol. 45.
15. Bartel, D.; Burstein, A.; Toda, M.; Edwards, D. The effect of conformity and plastic thickness on contact stresses in metal-backed plastic implants. *J. Biomech. Eng.* **1985**, *107*, 193–199.
16. Blunn, G.; Walker, P.S.; Joshi, A.; Hardinge, K. The dominance of cyclic sliding in producing wear in total knee replacements. *Clin. Orthop. Relat. Res.* **1991**, *273*, 253–260.
17. Walker, P.S.; Blunn, G.; Joshi, A.; Sathasivam, S. Modulation of Delamination by Surface Wear in Total Knees. 39th Annual Meeting of the ORS, San Francisco, CA, February 1993; 15–18.
18. Bartel, D.; Rawlinson, J.; Burstein, A.; Ranawat, C.; Flynn, W. Stresses in polyethylene components of contemporary total knee replacements. *Clin. Orthop. Relat. Res.* **1995**, *317*, 76–82.
19. McKellop, H.; Clarke, I.; Markolf, K.; Amstutz, H. Friction and wear properties of polymer, metal, and ceramic prosthetic joint materials evaluated on a multi-channel screening device. *J. Biomed. Mater. Res.* **1981**, *15*, 619–653.
20. Baker, D.; Hastings, R.; Pruitt, L. Study of fatigue resistance of chemical and radiation crosslinked medical grade ultrahigh molecular weight polyethylene. *J. Biomed. Mater. Res.* **1999**, *6* (573–581), 573–581.
21. Farrar, D.F.; Brain, A.A. The microstructure of ultra-high molecular weight polyethylene used in total joint replacements. *Biomaterials* **1997**, *18* (24), 1677–1685.
22. Kurtz, S.; Rimnac, C.M.; Bartel, D. Degradation rate of ultra-high molecular weight polyethylene. *J. Orthop. Res.* **1997**, *15*, 57–61.
23. Pruitt, L.; Bailey, L. Factors effecting near-threshold fatigue crack propagation behaviour of orthopedic grade ultra high molecular weight polyethylene. *Polymer* **1998**, *39* (8–9), 1545–1553.
24. Buechel, F.F.P.M.J. The New Jersey low-contact-stress knee replacement system: Biomechanical rationale and review of the first 123 cemented cases. *Arch. Orthop. Trauma Surg.* **1986**, *105* (4), 197–204.
25. Goodfellow, J.; O'Connor, J. The mechanics of the knee and prosthesis design. *J. Bone Jt. Surg.* **1978**, *60-B* (3), 358–369.
26. Jordan, R.O.J.L.; Voorhorst, P.E. Survivorship analysis of cementless meniscal bearing total knee arthroplasty. *Clin. Orthop.* **1997**, *338*, 119–123.
27. Polyziodes, A.J.D.G.K.; Tsakonas, H. The rotaglide total knee arthroplasty prosthesis design and early results. *J. Arthroplast.* **1996**, *11* (4), 453–459.
28. Psychoyios, V.C.R.W.; O'Connor, J.J.; Murray, D.W. Wear of congruent meniscal bearings in unicompartmental knee arthroplasty. *J. Bone Jt. Surg.* **1998**, *80-B* (6), 976–982.
29. Morgan-Jones, R.L.R.G.J.; Solis, G.; Parish, E.N.; Cross, M.J. Meniscal bearing uncemented total knee arthroplasty: Early clinical results at a minimum 2-year review. *J. Arthroplast.* **2003**, *18* (1), 41–44.
30. Murray, D.W.G.; O'Connor, J.J. The oxford medial unicompartmental arthroplasty. *J. Bone Jt. Surg.* **1998**, *80-B* (6), 983–989.
31. Argenson, J.N.O.C.J.J. Polyethylene wear in the meniscal knee replacement. A one to nine-year retrieval analysis of the oxford knee. *J. Bone Jt. Surg.* **1992**, *74-B* (2), 228–232.
32. Matsuda, S.W.S.E.; Williams, V.G.; McCarthy, D.S.; Whiteside, L. A Contact stress analysis in meniscal bearing total knee arthroplasty. *J. Arthroplast.* **1998**, *13* (6), 699–706.
33. Tsakonas, A.C.P.A.J. Reduction of polyethylene wear in a congruent meniscal knee prosthesis. *Acta Orthop. Scand., Suppl.* **1997**, *275*, 127–131.
34. Fu, F.H.C.D.; Vince, K.G. Total Knee Arthroplasty and Disorders of Articular Cartilage. In *Knee Surgery*; Williams and Wilkins: Baltimore, 1994; Vol. 2.
35. Palmer, S.H.; Machan, S.; Parish, E.N.; Cross, M.J. Morphology of the Distal Femur. 13th Triennial Congress of the Asia Pacific Orthopaedic Association, Adelaide, April 1–6, 2001.
36. Anouchi, Y.S.W.L.; Kaiser, A.; Milliano, M. The effects of axial rotation alignment of the femoral component on knee stability and patella tracking in total knee arthroplasty demonstrated on autopsy specimens. *Clin. Orthop.* **1993**, *321*, 168–172.
37. Whiteside, L.A.J. The anteroposterior axis for femoral rotational alignment based on the anterior-posterior axis in total knee arthroplasty is a valgus knee: A technical note. *J. Bone Jt. Surg.* **1995**, *77-A*, 1331–1334.

38. Duffy, G.P.B.D.J.; Rand, J.A. Cement versus cementless fixation in total knee arthroplasty. *Clin. Orthop.* **1998**, *356*, 66–72.
39. McCaskie, A.D.D.J.; Green, T.P.; et al. Randomised, prospective study comparing cemented and cementless total knee replacement. *J. Bone Jt. Surg.* **1998**, *80-B* (6), 971–975.
40. Onsten, I.N.A.; Carlsson, A.S.; et al. Hydroxyapatite augmentation of the porous coating improves fixation of tibial components: A randomised RSA study in 116 patients. *J. Bone Jt. Surg.* **1998**, *80-B* (3), 417–425.
41. Rorabeck, C.H.B.R.B.; Nott, L. The cemented kinematic-II and the non-cemented porous-coated anatomic prostheses for total knee replacement. *J. Bone Jt. Surg.* **1988**, *70-A* (4), 483–490.
42. Rand, J.A.I.D.M. Survivorship analysis of total knee arthroplasty. *J. Bone Jt. Surg.* **1991**, *73-A*, 397.
43. Ritter, M.A.H.S.A.; Keating, E.M.; et al. Long-term survival analysis of a posterior cruciate-retaining total condylar total knee arthroplasty. *Clin. Orthop.* **1994**, *309*, 136.
44. Collins, D.N.H.S.A.; Nelson, C.L.; Smith, P.S. Porous-coated anatomic total knee arthroplasty. *Clin. Orthop.* **1991**, *267*, 128–136.
45. Dodd, C.A.F.H.D.S.; Krackow, K.A. Total knee arthroplasty fixation. Comparison of the early results of paired cemented versus uncemented porous coated anatomic knee prostheses. *Clin. Orthop.* **1990**, *260*, 66–70.
46. Rosenburg, A.G.B.R.; Galante, J.O. A comparison of cemented and cementless fixation with the miller-galante total knee arthroplasty. *Orthop. Clin. North Am.* **1989**, *20* (1), 97–111.
47. Rand, J.A. Cement or cementless fixation in total knee arthroplasty. *Clin. Orthop.* **1991**, *273*, 52–62.
48. Jones, L.C.H.D.S. Cement disease. *Clin. Orthop.* **1987**, *225*, 192–206.
49. Hoffman, A.A.E.J.D.; Ferguson, R.P.; et al. Ten- to 14-year clinical follow-up of the cementless Natural Knee system. *Clin. Orthop.* **2001**, *388*, 85–94.
50. Parker, D.A.R.C.H.; Bourne, R.B. Long-term follow-up of cementless versus hybrid fixation for total knee arthroplasty. *Clin. Orthop.* **2001**, *388*, 68–76.
51. Nielsen, P.T.H.E.B.; Rechnagel, K. Cementless total knee arthroplasty in unselected cases of osteoarthritis and rheumatoid arthritis. *J. Arthroplast.* **1992**, *7* (2), 137–143.
52. Ryuichi, G.A.S.; Takizawa, T. Fixation of the NexGen HA-TCP-coated cementless, screwless total knee arthroplasty. *J. Arthroplast.* **2002**, *17* (4), 449–456.
53. Toksvig-Larsen, S.J.P.; Ryd, L.; et al. Hydroxyapatite-enhanced tibial prosthetic fixation. *Clin. Orthop.* **2000**, *370*, 192–200.
54. Nielsen, P.T.V.E.R.; Rozing, P.M. The effect of hydroxyapatite on the micromotion of total knee prosthesis: A prospective, randomised, double-blind study. *J. Bone Jt. Surg.* **1998**, *80-A*, 1665–1672.
55. Basset, R.W. Result of 1000 performance knees: Cementless versus cemented fixation. *J. Arthroplast.* **1998**, *13* (4), 409–413.
56. Cross, M.J.R.G.J.; Dimmick, R. The effect of hydroxyapatite on bone-prosthesis interface in total knee arthroplasty. *J. Bone Jt. Surg.* **1998**, *80-B* (2S), 140.
57. Cross, M.J.D.P.; Chitnavis, J.; Parish, E.N. Medium term clinical results of hydroxyapatite-coated total knee replacements. *Key Eng. Mater.* **2003**, *240–242*, 857–858.
58. Cross, M.J.R.G.J.; Holt, M.; Bradbury, N.; Dixon, P.; Parish, E.N. Enhanced fixation of uncemented knee replacement with hydroxyapatite. *Key Eng. Mater.* **2003**, *240–242*, 773–776.
59. Palm, L.J.S.A.; Ivarsson, I. Hydroxyapatite coating improves 8- to 10-year performance of the link RS cementless femoral stem. *J. Arthroplast.* **2002**, *17* (2), 172–175.
60. Oosterbos, C.J.R.A.I.; Tonino, A.J. Hydroxyapatite coated hip prosthesis followed up for 5 years. *Int. Orthop.* **2001**, *25* (1), 17–21.
61. Insall, J.N.R.C.S.; Aglietti, P.; Shine, J.A. A comparison of four models of total knee replacement prosthesis. *J. Bone Jt. Surg.* **1976**, *58-A*, 754–765.
62. Lequesne, M.G.M.C.; Samson, M.; et al. Indexes of severity for osteoarthritis of the hip and knee. Validation-value in comparison with other assessment tests. *Scand. J. Rheumatol., Suppl.* **1987**, *65*, 85–89.
63. Bellamy, N.B.W.W.; Goldsmith, C.H.; et al. Validations of WOMAC: A health status instrument for measuring clinically important patient relevant outcomes to anti-rheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J. Rheumatol.* **1988**, *15* (12), 1833–1840.
64. Dawson, J.F.R.; Murray, D.; et al. Questionnaire on the perceptions of patients about total knee replacement. *J. Bone Jt. Surg.* **1998**, *80-B* (1), 63–69.
65. Cooper, C.M.T.; Coggon, D.; Egger, P.; Dieppe, P. Occupational activity and osteoarthritis of the knee. *Ann. Rheum. Dis.* **1994**, *53* (2), 90–93.
66. Coggon, D.C.P.; Kellingray, S.; Barrett, D.; McLaren, M.; Cooper, C. Occupational physical activities and osteoarthritis of the knee. *Arthritis Rheum.* **2000**, *43* (7), 1443–1449.
67. Ulbrich, J.R.A.; Alexander, N.B. Body position used by healthy and frail older adults to rise from the floor. *J. Am. Geriatr. Soc.* **2000**, *48* (12), 1626–1632.
68. Schai, P.A.G.A.J.; Scott, R.D. Kneeling ability after total knee arthroplasty: Perception and reality. *Clin. Orthop.* **1999**, *367*, 195–200.
69. Palmer, S.H.S.C.T.J.; Maguire, J.; Parish, E.N.; Cross, M.J. Ability to kneel after total knee replacement. *J. Bone Jt. Surg.* **2002**, *84-B*, 220–222.
70. Ritter, M.A.M.J.B. Total hip arthroplasty: Can the patient play sports again? *Orthopaedics* **1987**, *10*, 1447–1452.
71. Dubs, L.G.N.; Munzinger, U. Sport after total hip arthroplasty. *Arch. Orthop. Trauma Surg.* **1983**, *101*, 161–169.
72. Bradbury, N.B.D.; Spoo, G.; Cross, M.J. Participation in sports after total knee replacement. *Am. J. Sports Med.* **1998**, *26* (4), 530–535.

Laboratory Animals: Ethics and Regulations for Care and Use

Kathy E. Laber

M. Michael Swindle

Department of Comparative Medicine, Medical University of South Carolina, Charleston, South Carolina, U.S.A.

INTRODUCTION

This article reviews common ethical arguments used both for and against the use of animals as biomedical research models, thereby providing information which will aid the investigator who either volunteers or is unwittingly engaged in discussing the use of animals in biomedical research. It also provides a summary of the laws and regulations, which govern the use of animals in research in the U.S.A. A working understanding of these laws and regulations is critical for those individuals who choose to utilize animals in their efforts to advance biomedical research. It is important to also recognize that although many of the European countries and Canada share in common the ethical arguments and parallel the animal use principles presented in this article, many of these countries have their own unique sets of laws regulating the use of animals. If animal work is to be conducted in collaboration with foreign countries, further resources specific to that country should be consulted.

ETHICS

The use of animals for biomedical research is a controversial and complex moral issue. Medical professionals and their patients clearly depend upon the advances made through the use of animals in research on a daily basis, but that unavoidable dependency does not eliminate the moral quandary, i.e., “whether it is ‘right’ to take the life of animal in pursuit of scientific knowledge and advancing medical treatment.” The emphasis and interest in this moral dilemma has increased in recent years as our society has continued to move away from its agricultural base. In western society the only interaction that many individuals have with animals is the emotion-based relationship of a pet-owner.^[1]

Moral values are primarily imprinted by religious followings, and many religions support the concept that human life is more valuable than animal life. Most religions support the animal’s position as a food and

work source, but religious teachings generally condemn cruelty toward animals. Religions tend to encourage individuals to act as responsible animal guardians. The major religions such as Christianity, Judaism, and Islam generally support the use of animals in research with the significant caveats being that animal pain and distress is minimized and that the research is likely to lead to the saving of human life or the relief of human suffering.^[2]

Philosophers have purported the concept that humans have greater “moral standing” than animals as humans have vastly superior sensory and cognitive abilities. For example, humans are more aware of pain and pleasure, thereby making the lesser-evolved animal species the more humane choice for research studies. After an evaluation of the level of moral standing an animal possesses and a determination that the animal has lesser moral standing than humans, the issue then debated is whether it is acceptable to utilize that animal *because* it has lesser moral standing than humans. This philosophical concept has created great angst over the use of nonhuman primates, particularly great apes, as they are felt to have a moral standing more closely related to man’s as evidenced by sensory and cognitive abilities that include language and self-consciousness. Other philosophers have argued that there is not a difference in “degrees” of moral standing and the acknowledgement of moral standing comes with the simple possession of life, the ability to feel pain and consciousness, and the possession of beliefs and desires. Clearly, if one adopts these tenets many animals used for science would be considered as having moral standing equal to man; therefore the use of animals in basic research is comparable to the use of humans in basic research. The mass embracement of this philosophical ideal would unmistakably bring biomedical research to a halt.^[3]

The utilitarian approach to the animal use debate is used quite eloquently by both scientists and animal activists, i.e., the greatest good for the greatest number justifies the means. Scientists argue that animal research is justified because it brings benefits to humans that outweigh the harm to animals. Antivivisectionists

argue in turn that the suffering of laboratory animals is marked, and the benefits are too limited to warrant this suffering. Both of these arguments beg the question of “how much does an animal suffer?” In its natural habitat a mouse ranges through wide areas and avoids human contact. Does something as simple as putting a mouse into a cage with frequent human contact cause a level of suffering that outweighs the benefit achieved from that animal? Does that mouse suffer more or less than a rabbit placed into a cage? Do all animals experience pain to the same degree? Clearly, because of the large variety of species used in research in combination with the lack of objective data on what constitutes distress or even pain for each individual species, we cannot cleanly define how much an animal suffers in any given experiment. Therefore the utilitarian approach to the use of animals in research provides logical arguments that support both sides of the debate.^[4]

The advent of transgenic technology has created a new era in animal research with additional ethical issues requiring thoughtful deliberation. Modifying the genetic endowment that defines a living being is felt by some to be intrinsically immoral. The process of inserting a foreign genetic material has the potential to create pain and distress in animals. The insertion of DNA occurs randomly in the genome creating the potential for gross mutations, severe disease, and mortality. Counter to that argument is the fact that transgenic technology has decreased the use of larger, potentially more sentient animals and has allowed for the rapid advancement of medical therapies targeted at the gene level.^[5]

The intense emotional response that is evoked in individuals about the topic of animal use for research has resulted in the organization of extremist groups such as People for the Ethical Treatment of Animals and the Animal Liberation Front. These groups are adept at presenting information within media campaigns to negatively sway public opinion against animals in research. They have also been very effective at presenting these views throughout the primary school systems. Extremists have, with increased frequency, resorted to terrorist tactics to “liberate” animals housed in research facilities, vandalize laboratories, and intimidate investigators and their families. The governments of the U.S.A. and several other nations have worked to develop legislation which prosecutes these activities; however such campaigns have effectively diminished the private business research sector from growing in places such as the U.K. Historically, scientists have not aggressively combated the animal activist movement; although in recent years the National Association for Biomedical Research (www.nabr.org) has lobbied for legislation that protects the research enterprise. The affiliated Foundation

for Biomedical Research (www.fbresearch.org) has worked to educate individuals about the benefits of biomedical research. Of potential future concern is the growing number of law schools that are expanding their curriculum to study the legal rights of animals. This could lead to the legal stance that humans are not “owners” of animals but rather their guardians. The adoption of this philosophy at a grass root level would clearly add monumental barriers to the use of animals in research.

At present, the general populace in the U.S.A. support the idea that research is important, that animals deserve some level of moral standing in our society, and that research on animals can be justified under certain conditions. This support, however, is dependent in part upon the species being used and the type of research being performed. For example, there is more support for the use of mice in cancer research than the use of rabbits for cosmetic testing. The use of pet species such as dogs and cats and nonhuman primates generates more concern from the public than rodent species.^[6] The concerns that the public has about the use of animals in research have led to adoption of a variety of guiding principles and laws which ensure the public that certain ethical considerations are being given to the use of animals in research.

ANIMAL WELFARE ACT

Although animals have been utilized in biomedical research for centuries, the first law in the U.S.A. impacting on biomedical research was the *Laboratory Animal Welfare Act of 1966* [later renamed the *Animal Welfare Act* (AWA) <http://www.nal.usda.gov/awic/pubs/awicdocs.htm>]. The public outcry that leads to the creation of this law was the result of the media publicizing that dogs and cats were being stolen and used for research, or housed by animal dealers in inappropriate ways prior to being sold to research institutions.^[7] The law implemented licensing requirements for animal dealers in addition to detailing transportation, husbandry, and veterinary care requirements for certain species animals used in biomedical research.

The AWA has been amended several times over the years, with each amendment resulting in additional regulations; however, it was the 1985 amendment to the AWA that broke new ground resulting in a very significant impact on the use of animals in biomedical research. One of the major changes pursuant to this amendment was the requirement that institutions appoint an Institutional Animal Care and Use Committee (IACUC). This committee, with its required specified membership inclusive of scientists, was, by law, assigned the responsibility for the animal care and



use program in coordination with the attending veterinarian. Historically, the weight of ensuring compliance with the AWA had been placed primarily on the back of the attending veterinarian. This arrangement often created a state of conflict between the attending veterinarian who was enforcing the AWA and the investigators who were focused on completing their research in the most efficient manner irrespective of the laws and regulations, thereby creating the potential for a dysfunctional research environment. Key to empowerment of the IACUC was its legal ability to suspend the use of animals on a project if the animal use was not being conducted in a manner consistent with the AWA principles or consistent with the scientific plan that the IACUC was to review and approve. The IACUC now became responsible for reviewing facilities where animals were held as well as the program for animal care and use twice yearly with a report of their activities forwarded to an individual who had to have the authority and resources to correct any issues the committee identified. Deficiencies in either the program or the facilities had to be corrected within a specified time frame, or the deficiency then had to be reported to the U.S. Department of Agriculture (USDA). The IACUC was also charged with reviewing the scientific plan for the animals to ensure that: 1) procedures were carried out by an individual adequately trained to conduct them; 2) procedures would avoid or minimize pain and distress, and be conducted with appropriate anesthetics and analgesics; 3) appropriate perioperative surgical care would be provided; 4) multiple survival surgery would be scientifically justified; 5) alternatives to the use of animals had been considered and documented; 6) a qualified veterinarian was consulted in the planning of the project; and 7) euthanasia would be conducted in an acceptable manner. In addition to the husbandry and veterinary care standards that were present in the original law, the amendment required that attention be given to the psychological well-being of nonhuman primates and that dogs received exercise and socialization.^[8]

Compliance with the AWA is monitored by trained veterinarians through unannounced visits employed by the USDA. The visits vary in frequency and length dependent in part on the size of the program, the species utilized, and the history of compliance. Deviations from the law are cited and the institution has the opportunity to either respond with corrections or contest the findings. Uncorrected or repeat deficiencies often result in significant fines being levied on the institution. In addition, information that the USDA collects in its inspection reports is accessible to the public through the *Freedom of Information Act* which has been used by animal rights activists to miscommunicate the findings to news media.

Of particular note is the fact that the AWA specifically *excludes* certain animals in its coverage: farm animals used for agricultural purposes, cold blooded animals, birds, and the majority of animals that are used for research—rats and mice.

HEALTH RESEARCH EXTENSION ACT

Concurrent with the Animal Welfare Amendment of 1985 was the enactment of Public Law 99-158 known as the *Health Research Extension Act*. This Act, in broad brush strokes, defined principles that institutions receiving support from the *Public Health Service* (PHS) were required to follow.^[9] The principles were expanded in the publication *Public Health Service Policy on Humane Care and Use of Laboratory Animals* (PHS Policy), enforced by the Office of Laboratory Animal Welfare—OLAW (<http://grants.nih.gov/grants/olaw/olaw.htm>).^[10] The principles required that each institution which received National Institutes of Health (NIH) funding submit an assurance describing their animal care and use program, which must then be approved for compliance with the Policy by OLAW prior to the release of funds from NIH. The Policy is a less voluminous document when compared to the AWA; however within its pages is the requirement that institutions use the *Guide for the Care and Use of Laboratory Animals* (Guide) as a basis for developing and implementing an institutional program for activities involving animals.^[11] The Guide, in turn, is a complex document which defines the operations of an animal care program in much greater detail (see below). The Policy expands the needed IACUC membership from three, as stated in the AWA, to five. It also increases well beyond the scope of the AWA the circumstances which are reportable to OLAW. The AWA only requires that the IACUC report suspension of activities, whereas the Policy requires that the IACUC report serious or continuing noncompliance with the Policy or serious deviations from the provisions of the Guide to OLAW. The OLAW, in turn, expects that institutions will be able to self correct the deficiencies that are identified and notified OLAW accordingly (Table 1).

Although OLAW, the organization responsible for enforcing the PHS Policy, rarely conducts unannounced inspections, they have investigated institutions for cause with the ramifications of negative finding being much more financially significant than the fines imposed by the USDA for noncompliance. The OLAW has the ability to withhold or restrict PHS animal welfare assurances which will eliminate or restrict funding from PHS agencies such as the NIH or the Food and Drug Administration (FDA). The OLAW has an additional

Table 1 Requirements of PHS policy/NIH guide and Animal Welfare Act

Requirement	PHS policy/NIH guide	Animal Welfare Act
IACUC membership	Five members: one nonscientific community, one scientist, one veterinarian	Three members: one community representative, one scientist, one veterinarian
IACUC duties	Review protocols, semi annual inspection, report suspended activities, report serious deviation from the Guide, report serious or continuing noncompliance	Review protocols, semiannual inspection, report suspension of an activity
Annual report	IACUC through the Institutional Official (IO) to OLAW annually	IO to APHIS prior to December 1 reporting fiscal year (October 1–September 30)
Suspensions/noncompliance	IACUC through the IO to OLAW	IACUC through the IO to APHIS
Species covered	All vertebrates	Includes mammalian species. Excludes rats, mice, birds, farm animals used for food and fiber research and cold blooded animals
Registration	PHS approved assurance required for federal funding. Approved for up to five years	Required registration when using covered species. Updated every three years
Training requirements	All personnel	All personnel
Inspections	Potentially OLAW	USDA at least annually
Shipping regulations	None	Specific for covered species
Exercise and socialization	General guidelines	Specific for dogs
Psychological well being	General guidelines for enrichment	Specific for primates
Husbandry standards	Specific standards	Specific for covered species
Surgical standards	Facility, personnel, anesthesia/analgesia, sterile technique, postoperative care. Multiple survival surgery requires scientific justification	Appropriate anesthesia/analgesia, no paralytic agents without anesthesia. Multiple survival surgery requires scientific justification
Construction standards	Specific	General
Veterinary consultation and care	Required	Required
Alternatives to animals for invasive procedures	Literature search required	Literature search required and use of National Agricultural Library should be encouraged
Sanitation standards	Specific	General

mission of being both an educator and a resource for the biomedical research community. They frequently cohost didactic lectures across the U.S.A. providing information on best practices for IACUC function in addition to helping to fund educational materials on research techniques related to animals.

AAALAC INTERNATIONAL ACCREDITATION

During the post-World War II science boom, leading veterinarians and researchers recognized the need for

an independent organization to encourage high standards for humane animal care and use in science. A committee called the Animal Facilities Accreditation Board was formed, and in 1963 they drafted the first *Guide for Laboratory Animal Facilities and Care*. In 1964, this group conducted 26 site visits of institutions using animals in research using this Guide on a trial basis. At the same time, the committee requested financial assistance from various biomedical associations who were committed to humane treatment of animals in research. The concept was well received by the biomedical research community, and in 1965, a

permanent organization called the American Association for the Accreditation of Laboratory Animal Care (AAALAC; www.aaalac.org) was formed. The organization was established to be operated by a Council of expert veterinarians, scientists, and administrators well versed in complexities of animal research. The organization reported then to a Board of Trustees that represented the scientific community. The mission of AAALAC, as defined by the Board of Trustees, was to enhance life sciences by promoting the responsible treatment of animals used in research, teaching and testing through voluntary accreditation and assessment programs. In 1996, AAALAC changed its name to the Association for the Assessment and Accreditation of Laboratory Animal Care, International to reflect its expansion outside the U.S.A. Today, more than 670 institutions are accredited in 25 different countries. Unlike reviews conducted by either the USDA or the OLAW, AAALAC site visits are completely *voluntary*, and the results are completely *confidential*. The site visits are conducted as peer-review assessments evaluating performance-based outcomes. The guidelines by which programs are evaluated include the "Guide" and certain widely accepted peer-reviewed reference resources such as the *American Veterinary Medical Association (AVMA) Panel on Euthanasia*,^[12] the *Guide for the Care and Use of Agricultural Animals in Agricultural Research and Teaching*,^[13] and the *Guidelines for the Care and Use of Mammals in Neuroscience and Behavioral Research*.^[14] Countries outside the U.S.A. which participate in the AAALAC process are also evaluated by laws and regulations specific to that country such as the European Union's *Council Directive on the Approximation of Laws, Regulations and Administrative Provisions of the Member States Regarding the Protection of Animals Used for Experimental and Other Scientific Purposes (EEC 86/609)*,^[15] or the Council of Europe's *European Convention for the Protection of Vertebrate Animals Used for Experimental and Other Scientific Purposes (Convention ETS 123)*^[16] or the Canadian Council for Animal Care's *Guide to the Care and Use of Experimental Animals*.^[17] A listing of Council approved reference resources that AAALAC utilizes to evaluate programs can be found on their web site. The outcome of the visit decided by a vote of Council can be: 1) full accreditation; 2) deferred accreditation which indicates that some relatively easily correctable deficiencies were present; 3) probation which indicates that some serious deficiencies requiring intensive programmatic adjustments were present; or 4) revocation of accreditation which must be approved by the Board of Trustees.

Site visits occur every three years, but annual reports are required. In addition, deficiencies cited by or self-reported to the USDA or OLAW must be

forwarded to AAALAC. Certain funding agencies require AAALAC accreditation (i.e., American Heart Association), and achieving full accreditation helps tremendously in streamlining the application for a PHS assurance as required to receive NIH funding. In addition, accreditation signals to both the general public and the scientific community that the institution has a commitment to maintaining a high quality animal care and use program.

THE GUIDE FOR CARE AND USE OF LABORATORY ANIMALS

Any discussion of laws and regulations would be remiss without further explanation given to the Guide, the cornerstone of both the PHS policy and the AAALAC accreditation process. As mentioned above, the first version of the Guide was published in 1963 by a group of dedicated professionals who formed a committee within the American Association of Laboratory Animal Science organization. The next revision of the Guide followed rapidly in 1965, published by the National Research Council through the Institute of Laboratory Animal Resources. The AWA and the Guide presented the very first guidelines and laws to impact on the biomedical research community at about the same time. All future revisions of the Guide have occurred through the National Research Council. The last revision of the Guide occurred in 1996, at which time the concept of performance vs. engineering standards was introduced. The Guide now in many areas defines what the outcome should be, but leaves it up to professional judgment to determine how to best achieve that outcome. Previous versions of the Guide were much more prescriptive in their requirements for husbandry standards. The newest version of the Guide also put increased emphasis on behavioral management for all animals. Enrichment should be provided for animals in terms of conspecific contact for social animals, exercise opportunities, and structural cage enrichments such as nesting material for rodents and "play" items for larger species such as pigs and goats.

The Guide is one of the most detailed documents instructing institutions about how to conduct an animal care and use program. Its scope encompasses all vertebrate animals. It has four chapters that discuss institutional policies and responsibilities, the necessary animal environment, housing and animal oversight, the depth of veterinary medical care, and the physical plant requirements. The research leadership at the vast majority of U.S. institutions which use animals in research is intimately familiar with the guidelines and recommendations found within this document, and use

them as the framework on which to base their animal care and use program.

ENDANGERED SPECIES ACT

The Endangered Species Act^[18] was enacted in 1973. Relative to animals, this act was designed to provide a means whereby threatened species may be conserved. It therefore prohibits transport and commerce of endangered species and prohibits the import of primates except for research purposes. The endangered animals that existed within the U.S.A. prior to 1973 are exempt from the act. The regulatory authority for this act is vested with the U.S. Department of the Interior (USDI), and implemented by the USDI's Fish and Wildlife Service.

GOOD LABORATORY PRACTICES ACT

The Good Laboratory Practices (GLP) Act^[19] was enacted in 1978 to provide standards for the conduction of preclinical trials in animals. The act requires that institutions, which are mainly pharmaceutical companies, follow all applicable laws and regulations. In addition, the GLP provide for intensive documentation of all procedures that are performed. It includes complete postmortem examinations of animals undergoing the testing procedures. GLP standards are of such intensity that offices which specifically monitor the procedures and maintain the paperwork are a practical necessity. Thus, most of the GLP trials are conducted by biomedical corporations and contract research laboratories. The law is regulated by the FDA which periodically conducts compliance inspections of the institutions.

OTHER REGULATIONS

In addition to the federal laws there are a number of other regulations that apply to animals in research. Notably the Veterans Administration (VA) has requirements that include mandatory AAALAC accreditation of their research facilities. The VA also has a protocol review process and position statements that are more in depth than the requirements of most academic institutions. The Department of Defense (DOD) also has specific policies, procedures, and responsibilities detailing the use of animals that are supported by DOD funding.

Professional societies also publish guidelines for performing research in specific fields. In particular, the AVMA publishes the *AVMA Guidelines for Euthanasia* which is utilized as the standard for which OLAW, the

USDA, and AAALAC expect institutions to meet. Many other professional societies such as the Society of Neurosciences also publish guidelines for conducting experiments which are within their field of interest.

CONCLUSIONS

The responsible conduct of research utilizing animals requires both a thorough understanding of the laws and regulations applicable to their use as well as an appreciation for the ethical dilemma that many who work with animals experience.

The wise researcher is aware of the negative public perception of animal research that often prevails and realizes that the biomedical research community as a whole could vastly benefit from proactive vs. reactive public education. Ongoing efforts by the researcher to refine and replace the use of animals in research are both a legal requirement and a moral one. Science could not advance without the use of animals, but their use in research must be executed with a solid balance of compassion for both the humans receiving the benefits and the animals whose lives are being sacrificed.

ARTICLE OF FURTHER INTEREST

Biomedical Ethics

REFERENCES

1. Hart, L. The human/animal relationship in the research setting. In *The Human/Research Animal Relationship*; Krulisch, L., Mayer, S., Simmonds, R., Eds.; Scientists Center for Animal Welfare, May 1996.
2. Parker, J.V. The human/animal relationship: jewish, christian and islamic perspectives. In *Bioethics and the Use of Laboratory Animals: Ethics in Theory and Practice*; Kraus, L.A., Renquist, D., Eds.; Gregory Benoit Publishing: Dubuque, 2000; 29–45.
3. Russow, L.A. Bioethics, animal research, and ethical theory. In *Animals in Biomedical Ethics*; ILAR, 1999; Vol. 40 (1), 15–21.
4. Rowan, A. Ethical principles for animal research and the sundowner principles. In *Bioethics and the Use of Laboratory Animals: Ethics in Theory and Practice*; Kraus, L.A., Renquist, D., Eds.; Gregory Benoit Publishing: Dubuque, 2000; 5–29.
5. Gonder, J.; Prentice, E.; Russow, L.M., Eds.; *Genetic Engineering and Animal Welfare: Preparing for the 21st Century*; Scientists Center for Animal Welfare, October 1999.
6. Rowan, A.; de Liedekerck, V. Attitudes to animal research. In *Bioethics and the Use of Laboratory Animals: Ethics in Theory and Practice*; Kraus, L.A., Renquist, D., Eds.; Gregory Benoit Publishing: Dubuque, 2000; 171–184.



7. Sideris, L.; McCarthy, C.; Smith, D.H. Roots of concern with nonhuman animals in biomedical ethics. In *Animals in Biomedical Ethics*; ILAR, 1999; Vol. 40 (1), 3–14.
8. Animal Welfare Act and Animal Welfare Regulations, United States Department of Agriculture, Animal and Plant Health Inspection Service, August 2002.
9. Animals in Research, Health Research Extension Act of 1985, Public Law 99–158; November 20, 1985.
10. *Public Health Service Policy on Humane Care and Use of Laboratory Animals*; OLAW, 2002.
11. National Research Council. *Guide for the Care and Use of Laboratory Animals*; National Academy Press: Washington, D.C., 1996.
12. American Veterinary Medical Association. Report of the AVMA panel on euthanasia. *JAVMA* **2001**, *218* (5), 669–696.
13. *Guide for the Care and Use of Agricultural Animals in Agricultural Research and Testing*; Federation of Animal Sciences, 1999.
14. *Guidelines for the Care and Use of Mammals in Neuroscience and Behavioral Research* (NRC 2003); (<http://www.nap.edu/catalog/10732.html>).
15. Council Directive on the Approximation of Laws, Regulations and Administrative Provisions of the Member States Regarding the Protection of Animals Used for Experimental and Other Scientific Purposes. European Union (Directive 86/609/EEC), 1986. (Adopted May 1999).
16. European Convention for the Protection of Vertebrate Animals Used for Experimental and Other Scientific Purposes. Council of Europe (Convention ETS 123), 1985 (Adopted May 1999); <http://conventions.coe.int/treaty/EN/cadreprincipal.htm>.
17. *Guide to the Care and Use of Experimental Animals*. Canadian Council on Animal Care, 1984, Vol. 2. (Adopted May 1999) Canadian Council on Animal Care: Albert St., Ottawa, Ontario, Canada, K1R 1B1, 315–350; http://www.ccac.ca/english/gui_pol/guides/english/TOC_V2.HTM.
18. Endangered Species Act, PL–101-707; November, 1988.
19. Good Laboratory Practice for Non-Clinical Studies, Title 21 Code of Federal Regulations, Chapter 1: FDA, Department of Health and Human Services, 4-1-97 Ed., Subchapter A, Part 58.

Laser-Tissue Interaction

E. Duco Jansen

Department of Biomedical Engineering, Vanderbilt University, Nashville, Tennessee, U.S.A.



INTRODUCTION

When applying laser light to biological tissue, a variety of complex interactions may occur (Fig. 1). A comprehensive discussion of all aspects of laser-tissue interaction is clearly beyond the scope and page limits of this article, so we will limit our discussion to the most important concepts. Specific tissue characteristics as well as laser parameters contribute to this diversity. Most important among optical tissue properties are the refractive index and the wavelength-dependent coefficients of absorption, and scattering. Together they determine the distribution of light and total transmission of the tissue at a certain wavelength. This initial distribution of light lies at the very heart of all ensuing effects, whether they are therapeutic or diagnostic. On the other hand, the following parameters are given by the laser radiation itself: wavelength, exposure time, laser power, applied energy, spot size, radiant exposure (energy/unit area), and irradiance (power/unit area). Among these, the exposure time is a crucial parameter when selecting the type of interaction, as we will find shortly.

TISSUE OPTICS

In describing the optical properties and light propagation in tissues, light is treated as photons. The primary reason for this approach is that biological tissue is an inhomogeneous mix of compounds, many with unknown properties. Hence, analytical solutions to Maxwell's equations in this medium pose an intractable mathematical problem. The representation of light as photons presents the opportunity to apply probabilistic approaches that lend themselves particularly well to numerical solutions that are tractable in computer simulations. Photons in a turbid medium such as tissue can move randomly in all directions and may be scattered (described by its scattering coefficient μ_s [m^{-1}]) or absorbed (described by its absorption coefficient μ_a [m^{-1}]). These coefficients along with anisotropy (i.e., the direction in which a photon is scattered) and index of refraction are referred to as the

optical properties of a material. If photons impinge on tissue several things can happen:

1. Some photons will reflect off the surface of the material (similar to what happens to glass and other materials—Fresnel's equations hold for tissue as well).
2. The majority of the photons will enter the tissue, upon which the following can happen:
 - a) The photon is absorbed (and can be converted to heat, trigger a chemical reaction, or cause fluorescence emission).
 - b) The photon is scattered (bumps into a particle and changes direction, but continues to exist).
 - c) Some photons can make it through the entire slab without running into anything. They are neither scattered nor absorbed and will emerge on the other side—these are called ballistic photons.

Absorption

We will first consider the case when only light is absorbed. In tissue optics, absorption of photons is a crucial event:

- Absorption is the primary event that allows a laser or other light source to cause a potentially therapeutic (or damaging) effect on a tissue. Without absorption, there is no energy transfer to the tissue and the tissue is left unaffected by the light.
- Absorption of light can provide diagnostic information by providing a clue as to the chemical composition of a tissue, and serve as a mechanism of optical contrast during imaging. Absorption is used in both spectroscopic and imaging applications.

Molecules that absorb light are called chromophores. Some examples of chromophores are water, blood (hemoglobin), melanin, amino acids, proteins, and fatty acids.

- In the UV: Tissue absorption is high due to protein, amino acid, fatty acid, and DNA absorption.
- In the visible: Whole blood (hemoglobin) is a strong absorber for some wavelengths (blue/green), but its

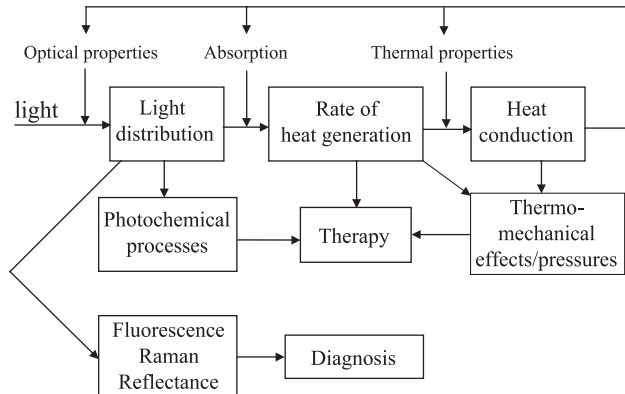


Fig. 1 Schematic of laser-tissue interactions.

effect is limited since the fraction of blood in tissue is small, unless one considers highly perfused tissues or the vasculature itself. Melanin is a strong absorber, again providing a small fraction of tissue on average except when considering areas with strong pigmentation such as the skin or retina.

- In the IR: Tissue absorption is dominated by water absorption.

The absorption of light can be characterized using Beer's Law, which predicts that the light intensity in a material decays exponentially with depth:

$$E(z) = E_0 e^{-c\xi(\lambda)z} \quad (1)$$

where E_0 is the incident irradiance [W/m^2], $E(z)$ is the irradiance through some distance z of the medium, c is the concentration of absorbing molecules in the medium, and $\xi(\lambda)$ is the molecular absorption coefficient (also known as the molar extinction coefficient), which is a function of the wavelength.

Often, c and $\xi(\lambda)$ are not known for tissue. Instead the two are lumped together and defined in a probability approach. We then replace $(c\xi(\lambda))$ with the absorption coefficient ($\mu_a(\lambda)$) and rewrite Beer's Law as:

$$E(z) = E_0 e^{-\mu_a(\lambda)z} \quad (2)$$

The absorption coefficient is defined as follows: When a photon travels over an infinitesimal distance Δz , the probability for absorption is given by $\mu_a \Delta z$ where μ_a is defined as the absorption coefficient (m^{-1}) (i.e., $1/\mu_a$ is the mean free path a photon travels before an absorption event takes place).^[1]

A related and useful parameter is the penetration depth, defined as the depth in the medium at which the energy or irradiance is reduced to $1/e$ times ($\sim 37\%$) the incident irradiance at the surface. Per definition, the penetration depth equals $1/\mu_a$ in cases where there is no scattering.

We have defined the irradiance (sometimes referred to as intensity) as power per unit area [W/m^2]. Thus, the irradiance gives us information about how much light made it to a certain point in the tissue. However, it does not tell us how much of that light is absorbed at that point. We define a new term called the heat source term or rate of heat generation, S , as the number of photons absorbed per unit volume [W/m^3]. Note that number of photons absorbed can be related to amount of heat generated, i.e., heat source. Mathematically, heat source can be written as the product of the irradiance at some point in the tissue and the probability of absorption of that light at that point:

$$S(z) = \mu_a E_0 e^{-\mu_a z} = \mu_a E(z) \quad (3)$$

Scattering

Scattering of light occurs in media that contain spatial fluctuations in the refractive index, n , whether such fluctuations are discrete particles or more continuous variations in n .

In biomedical optics, scattering of photons is an important event:

- Scattering provides feedback during therapy. For example, during laser coagulation of tissues, the onset of scattering is an observable endpoint that signals a desired therapeutic goal. Scattering also strongly affects the dosimetry of light during therapeutic procedures that depend on absorption. The scattering affects where the light goes and thus where absorption will occur.
- Scattering has diagnostic value. Scattering depends on the microstructure of a tissue, e.g., the density of lipid membranes in the cells, the size of nuclei, the presence of collagen fibers, the status of hydration in the tissue, etc. Scattering is used for both spectroscopic and imaging applications.

The scattering coefficient is defined as follows: When a photon travels over an infinitesimal distance Δz , the probability for scattering is given by $\mu_s \Delta z$ where μ_s is defined as the scattering coefficient (m^{-1}). The mean free path of scattering is $1/\mu_s$.^[1] Although the scattering coefficient tells us something about the likelihood that scattering occurs, it says nothing about the direction in which the light is scattered if scattering occurs. This scattering direction is described by the phase function $p(s, s')$, which represents the probability that a photon will scatter from direction s to direction s' . Light scattering in tissue is not isotropic, but strongly forward directed. A measure of the degree of anisotropy in scattering is given by the anisotropy factor, g , which is defined as the average cosine of the

scattering angle, θ . With this, it should be easy to understand that: total forward scattering: $g=1$; isotropic scattering: $g=0$; total backscattering: $g=-1$. For tissues in the visible and near-IR wavelengths, $g \approx 0.7-0.95$. With knowledge about both the probability of scattering and the scattering direction, we can define a few new useful parameters, known as the reduced optical properties:

- Reduced scattering coefficient: $\mu_s' = \mu_s(1-g)$
- Reduced attenuation coefficient: $\mu_t = \mu_a + \mu_s'$
- Effective attenuation coefficient:

$$\mu_{eff} = \sqrt{3\mu_a(\mu_a + \mu_s')}$$
- Penetration depth = mean free path = $1/\mu_t$
- Effective penetration depth: $\delta_{eff} = 1/\mu_{eff}$

Light scattered from the collimated beam undergoes multiple scattering events as it propagates through tissue. The transport theory is used to describe the transfer of energy through tissue. This theory is based on a statistical approximation of photon particle transport in a multiple scattering medium. The light transport equation relates the gradient of radiance at position r in direction s to losses due to absorption and scattering and gain due to scattering from all other directions s' into s .^[1]

$$\frac{dL(r, \hat{s})}{ds} = -\mu_a L(r, \hat{s}) - \mu_s L(r, \hat{s}) + S(r, \hat{s}) + \mu_s \int_{4\pi} p(\hat{s}, \hat{s}') L(r, \hat{s}') d\omega' \quad (4)$$

where dL/ds = change in radiance at r in the direction s , $L(r, s)$ = radiance defined as the propagation of photon power, $\mu_a L(r, \hat{s})$ = losses due to absorption, $\mu_s L(r, \hat{s})$ = losses due to scattering, $\mu_s \int_{4\pi} p(\hat{s}, \hat{s}') L(r, \hat{s}') d\omega'$ - gain due to scattering from all s' into s where $p(s, s')$ is the phase function, and $S(r, s)$ = source of power generated at r in the direction s .

The light transport equation is a complicated differential equation that has no analytical solution except for a few special cases. In most problems involving light scattering, different models can be used to solve the light transport equation. These include the Adding-Doubling method, Delta-Eddington method, Kubelka-Munk or two-flux theory, diffusion approximation method, and Monte Carlo simulations.^[1]

LIGHT-TISSUE INTERACTIONS

During the first decades after invention of the laser by Maiman in 1960, many studies have been conducted, investigating potential interaction effects by using all types of laser systems and tissue targets. Although the

number of possible combinations for the experimental parameters is unlimited, mainly three different types of interaction mechanisms are classified today: photochemical, photothermal, and photomechanical. Each of these interaction mechanisms will be discussed in detail in the following sections. In particular, the physical principles governing these interactions are reviewed.

Before going into detail, an interesting observation deserves to be stated. All these seemingly different interaction types share a common property: the characteristic radiant exposure [J/cm^2] ranges from approximately $1 \text{ J}/\text{cm}^2$ to $1000 \text{ J}/\text{cm}^2$. This is surprising since the irradiance itself varies over more than 15 orders of magnitude. Thus, a single parameter distinguishes and primarily controls these processes: the duration of the laser exposure, which is largely similar to the interaction time itself.^[2]

A log-log plot of irradiance versus pulse duration with the basic interaction mechanisms is shown in Fig. 2. In this figure, the photomechanical mechanisms are separated into three subcategories: photoablation, plasma-induced ablation, and photodisruption. In the graphs, the two diagonals show constant radiant exposures at $1 \text{ J}/\text{cm}^2$ and $1000 \text{ J}/\text{cm}^2$, respectively. According to this graph, the timescale can roughly be divided into three major sections: continuous wave or exposure times $>1 \text{ s}$ for photochemical interactions, 100 s down to $1 \mu\text{s}$ for photothermal interactions, and $1 \mu\text{s}$ and shorter for photomechanical interactions. It should be clear, however, that these boundaries are not strict, and adjacent interaction types cannot always be separated. Thus, overlap in these main regions does exist. For example, in the range of $1 \mu\text{s}$ to several

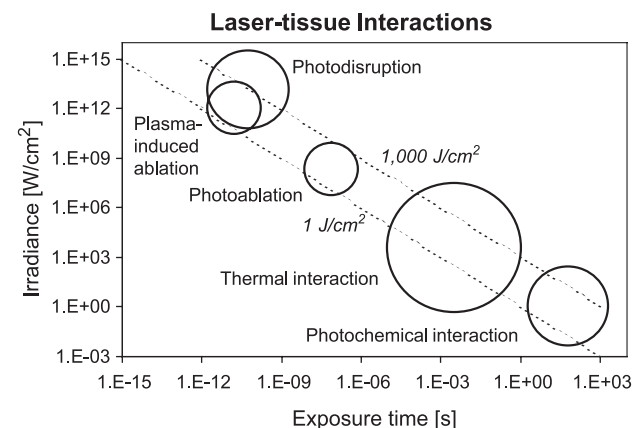


Fig. 2 Laser-tissue interaction regimes. Note that despite the fact that the irradiance varies over more than 15 orders of magnitude, the characteristic radiant exposure [J/cm^2] ranges from approximately $1 \text{ J}/\text{cm}^2$ to $1000 \text{ J}/\text{cm}^2$. (Figure adapted from Ref. [2].)

hundreds of μs , the interaction mechanisms typically have photothermal as well as photomechanical components to them, while many photochemical interactions also exhibit photothermal components. Finally, there is the possibility that multiple pulses of a particular interaction type collectively contribute to another interaction type. For example, even pulses of a duration of 100 ps, each of which has negligible thermal interactions, may add up to a measurable temperature increase if applied at repetition rates of 20 Hz or higher, depending on the laser. We will now consider each of the main interaction mechanisms in some more detail.

PHOTOCHEMICAL INTERACTIONS

We will only briefly discuss photochemical interactions. An excellent overview of mechanisms and some applications can be found in chapter 3 of Niemz's book on laser-tissue interaction.^[2] In brief, the group of photochemical interactions is based on the fact that light can induce chemical effects and reactions within macromolecules or tissues. The most obvious example of this is created by nature itself—photosynthesis. In the field of medical laser applications, photochemical interaction mechanisms play a role during photodynamic therapy (PDT).^[3] Frequently, biostimulation is also attributed to photochemical interactions, although this is not scientifically ascertained.

Photochemical interactions take place at very low irradiances (typically 1 W/cm^2) and long exposure times ranging from seconds to tens of minutes. Careful selection of laser parameters (most notably wavelength) yields a radiation distribution inside the tissue that is primarily determined by scattering. In most cases, wavelengths in the red and near-infrared part of the spectrum (600–900 nm) are used because of their low absorption and resulting deep penetration in tissue.

During PDT, chromophores (called photosensitizers) are injected in the body. Monochromatic light may then trigger selective photochemical reactions. The end result of a significant amount of complicated reactive species chemistry is that the activated photosensitizer, acting as a catalyst, causes highly cytotoxic compounds to be released, causing an irreversible oxidation of essential cell structures. Applications of PDT have traditionally focused on treatment of tumor although recently, excellent results have been reported in nononcological applications such as treatment of age-related macular degeneration (AMD).^[4]

Another type of photochemical laser-tissue interaction is molecular bond breaking. The absorbed photon excites the molecule to a higher electronic, vibrational, or rotational state. If the photon energy is

large enough to excite the molecule to an electronic state that is not bound, the molecule dissociates. Then it may form radicals or it may break. Clearly, the prerequisite for such a mechanism is ample supply of high-energy (short wavelength, UV) photons, because only those photons possess sufficient energy to overcome the molecular bond strength of organic molecules.

Recent and exciting developments that rely on photochemical interactions include experimental applications in the field of gene therapy where light may be used to activate caged DNA.^[5] In this scenario, therapeutic DNA is synthesized and linked to an inactivating group (a caged compound). Upon light exposure, cleavage of the DNA from its cage is achieved, rendering the active, translatable form of the therapeutic DNA only where light exposure is activated. Similarly, light-induced sense or antisense therapy may hold significant promise in this field.

Summary of Photochemical Interactions

- Main idea: Using photosensitizer acting as catalyst (PDT); optical activation/modulation of certain biochemical pathways (Biostimulation).
- Observations: No macroscopic observations.
- Typical lasers: Red dye lasers, diode lasers.
- Typical pulse durations: 1 second to CW.
- Typical irradiances: $0.01\text{--}50 \text{ W/cm}^2$.
- Special applications: PDT, biostimulation, gene therapy (?).

PHOTOTHERMAL INTERACTIONS

Photothermal interactions include a large group of interaction types resulting from the transformation of absorbed light energy to heat, leading to a local temperature increase. Thermal effects can be induced by either continuous wave (CW) or pulsed laser radiation. While photochemical processes are often governed by a specific reaction pathway, photothermal effects generally tend to be nonspecific and are mediated primarily by absorption of optical energy and secondly, governed by fundamental principles of heat transport. Depending on the duration and peak value of the temperature achieved, different effects like coagulation, vaporization, melting, or carbonization may be distinguished. An excellent overview of these interaction regimes can be found.^[6] It is essential to emphasize that thermal interactions in tissue are typically governed by rate processes, i.e., not just the temperature plays a role, but the duration for which the tissue is exposed to a particular temperature is also a parameter of major importance.

Once deposited in tissue and given sufficient time, the traditional mechanism of heat transfer applies to laser-irradiated biological tissues. Heat flows in biological tissue whenever a temperature difference exists. The transfer of thermal energy is governed by the laws of thermodynamics: 1) energy is conserved; 2) heat flows from areas of high temperature to areas of low temperature. The primary mechanisms of heat transfer to consider include: conduction, convection, and radiation.^[7] In addition, the following processes may have important effects on the overall temperature field of the laser-irradiated tissue: metabolic heat, perfusion, and evaporation.

Ablation

The word ablation comes from the latin *ablatus*, which means to remove. There are two meanings of the word ablation. In a physics/engineering sense, it refers to the physical removal of material (tissue), while in the medical/biological sense, it refers to the physiological removal of tissue (i.e., causing the tissue to stop performing the processes needed for survival of the tissue, that is, killing the cells, but not necessarily physically removing them). The disciplines in medicine where laser ablation has received the most attention are:

- Cardiology—angioplasty, arrhythmia treatment, TMR.
- Ophthalmology—corneal reshaping (LASIK), cataract surgery.
- Orthopedic surgery—cartilage, bone ablation (knees, joints, herniated discs).
- Dentistry—drilling cavities, soft tissue (gums).
- Dermatology—burn eschar debridement, skin resurfacing, hair removal.

Generally, below the boiling point of water, thermal interactions are dominated by protein denaturation and tissue coagulation, resulting in many cases in cell apoptosis or necrosis. Several medical applications use this type of interaction to achieve the clinical endpoint. For example, during interstitial hyperthermia, relatively weakly absorbed wavelengths are used, combined with long exposure times (minutes), to slowly coagulate or cook a mass of tissue such as a tumor using fiber-optic delivery of laser light.

Around 100°C, the thermal interactions become dominated by evaporation of water and we typically operate in the regime where physical material removal occurs. In this case, the goal is usually to remove or ablate the target tissue while minimizing the damage to the tissue left behind. Some notable exceptions to this paradigm include those applications where, in addition to cutting with the laser, some coagulation is desirable,

for instance, when the target tissue is highly vascularized. The amount of collateral thermal damage is largely determined by the amount of heat that can diffuse out of the irradiated tissue prior to the ejection of this material. We will now look at the process of ablation in some more detail.

The chronicle of ablation with a continuous wave laser generally follows the following sequence:

1. Light interaction with tissue (governed by tissue optical properties).
2. Conversion of optical energy to thermal energy (heat generation).
3. Tissue temperature rise (heat source and conduction).
4. Denaturation and dehydration of the tissue (this may affect the thermal and optical properties of the tissue).
5. Increase of subsurface pressure.
6. Explosion resulting in the expulsion of heated tissue fragments—this removes energy.
7. Exposure of cooler subsurface layers to light energy (then going back to 1).
8. At some point, all the tissue water has been evaporated and carbonization/pyrolysis of tissue will occur.

During tissue vaporization, the position of the tissue surface, where tissue and air meet, moves deeper into the tissue. The velocity at which the ablation front moves into the tissue is called the ablation velocity. By using a moderate irradiance and measuring surface temperatures with an IR thermal camera, it is possible to document the initial surface heating and explosive ejection of tissue. The onset of ablation occurs above 100°C due to subsurface superheating. As hot tissue is removed, a cooler layer is exposed to the laser irradiation. The temperature of the exposed layer remains at approximately 100°C while continued irradiation dehydrates the tissue. The removal of water decreases the local thermal conductivity, reducing heat conduction to the surrounding area. Continued radiation rapidly increases the tissue temperature until it reaches about 350 to 450°C; the tissue burns and carbonizes.

At, or just above 100°C, water vapor is generated volumetrically, the equilibrium is pushed toward the vapor phase, and, for sufficiently high laser fluence rates (irradiances), more vapor is produced than can escape by simple diffusion. The excess vapor is trapped in the tissue layers, forming vacuoles. The superheated vapor contained in the vacuoles will expand quickly, compressing the surrounding, rapidly drying tissues that form the vacuolar walls. As the vacuoles expand, the wall separating the vapor pockets from each other or the tissue surface become thin. The walls rupture as the force of the increasing pressures of expanding



vapor overcomes the mechanical strength of the tissue. In most applications of lasers as ablative devices, the goal is to remove the target tissue while minimizing collateral damage to adjacent tissue layers. Based on the very nature of the interaction of laser radiation with tissue, as we have seen, the damage mechanisms may be thermal or mechanical.^[6] In brief, mechanical damage is caused by subsurface explosion of heated tissue material (in particular, in cases where the absorption coefficient is relatively small), whereas thermal damage is caused by two sources: 1) thermal diffusion (from heated tissue to surrounding cooler tissue); and 2) direct deposition of heat in layers beyond the point where the energy density was sufficient for ablation, i.e., below threshold. The irony is that reducing the pulse duration will minimize the amount of heat diffusion that can take place before heated material is ejected. However, this same reduction of the duration of the pulse will cause the ablation process to be even more explosive and violent, thus increasing the collateral mechanical damage to the tissue.

It should be obvious that by choosing wavelengths that are highly absorbed (i.e., large absorption coefficient) in tissue, the energy is confined to a smaller volume and thus, the irradiance or radiant exposure necessary to achieve ablation will be less. In addition, the extent of the zone of tissue that is heated to subablative energy densities will be greatly reduced in this case.

Pulsed Laser Ablation

The question we now need to ask ourselves is: When is a laser a pulsed laser? Whereas there are certainly good answers and clear definitions from a laser physics point of view (i.e., how the machine works), we will look at it with a laser-tissue interaction mindset. As we have seen, laser ablation appears to be a trade-off between causing thermal damage with long pulses and mechanical damage with short pulses. First, when considering thermal diffusion, when the laser penetration depth is less than the laser spot radius, the thermal diffusion time, τ_{th} , can be defined as:

$$\tau_{th} = \frac{\delta^2}{4\alpha} \quad (5)$$

where δ is the penetration depth of laser light in tissue with thermal diffusivity, $\alpha = 0.15 \text{ m}^2/\text{s}$ for water at 37°C . For laser pulses shorter than the thermal diffusion time, the distribution of thermal energy is determined by the laser light distribution. This situation is known as thermal confinement. If the laser pulse duration is larger than the thermal diffusion time, the

thermal energy propagates into the tissue during the laser pulse. The criterion for thermal confinement is:

$$\tau_{pulse} < \tau_{th} \quad (6)$$

Next, we need to consider mechanical effects. A major photomechanical interaction during most tissue ablation events is tissue water vaporization, resulting in explosive removal of tissue structures, which has been described as the popcorn effect for CW lasers. Note that in this scheme, the explosiveness of the process may be severe enough to eject liquid or solid tissue fragments. As such, the energy needed to accomplish a certain crater size may be significantly less than expected based on calculations that assume true vaporization of the tissue. This ejection process is driven by the fact that water when converted to steam has a volumetric expansion of approximately 1620 at 1 atm. A superb and comprehensive overview of the physics of pulsed laser ablation of tissue has been published by Vogel and Venugopalan.^[8]

Summary of Photothermal Interactions

- Main idea: Achieving a certain temperature that leads to the desired thermal effect.
- Observations: Either coagulation, vaporization, melting, or carbonization.
- Typical lasers: CO_2 , Nd:YAG, Ho:YAG, Er:YAG, Argon ion, diode lasers.
- Typical pulse durations: 1 μs – minutes.
- Typical irradiances: $10\text{--}10^6 \text{ W/cm}^2$.
- Special applications: Coagulation, vaporization, melting, thermal decomposition, treatment of retinal detachment, Port-wine stain, retinal photocoagulation, interstitial hyperthermia.

PHOTOMECHANICAL INTERACTION

Photomechanical interaction includes mechanisms associated with photoablation, explosive vaporization, plasma-induced ablation, and effects of laser-induced acoustic (shock waves). Note that even though we discuss them separately, photothermal and photomechanical events are closely intertwined. Photomechanical mechanisms can be caused by the thermoelastic expansion of tissue due to extremely rapid heating of the target tissue by the pulsed laser light and recoil caused by ejection of ablated material. These photomechanical interactions cause stress waves, which propagate with the speed of sound or faster. If the laser pulse length is shorter than the time it takes the stress wave to propagate out of the irradiated tissue volume, large peak stresses can be

reached. These stresses may contribute to the ablation mechanism and may inflict damage to adjacent tissue.

Laser-Induced Acoustics

In addition to the explosion and material ejection, a second mechanical effect is the formation of pressure and shock waves. These pressure waves can be caused by the thermoelastic expansion of tissue or due to recoil caused by the ejection of ablation material. Thermoelastic expansion describes the event that due to heating of a material, the material expands (due to an on average larger molecular spacing). If the heating is rapid enough, significant expansion of the material may take place, leading to pressure waves that are formed in the heated volume and will propagate in the tissue outside the heated volume.

The instantaneous pressure rise is proportional to the temperature rise and can be calculated:^[14]

$$\Delta P = \frac{1}{\gamma} \frac{\Delta V}{V} = \frac{1}{\gamma} \beta \Delta T \quad (7)$$

where:

$$\gamma = \frac{1}{\rho \sigma^2} \frac{c_v}{c_p} [Pa^{-1}] \quad (8)$$

γ = coefficient of isothermal compressibility [Pa^{-1}],
 β = volumetric expansion coefficient [K^{-1}], c_v = specific heat at constant volume [J/kg K], c_p = specific heat at constant pressure [J/kg K].

Thus, we can write ΔP as:

$$\Delta P(z) = \beta \rho \sigma^2 \frac{c_v}{c_p} \Delta T(z) = \frac{\beta \sigma^2}{c_p} \mu_a H_0 e^{-\mu_{eff} z} \quad (9)$$

where $\beta \sigma^2 / c_p$ is known as the Gruneisen coefficient, Γ (dimensionless). This simplifies the expression for the pressure rise to:

$$\Delta P(z) = \Gamma \mu_a H_0 e^{-\mu_{eff} z} \quad (10)$$

The Gruneisen coefficient is a temperature-dependent parameter, approximately 0.11 for water at 20°C. It can be seen from Eq. 10 that the pressure as a function of depth of the acoustic wave will be proportional to (and thus exactly follow) the temperature distribution as a function of depth. This in turn means it will exactly mimic the light distribution in tissue. Hence, for an absorption-dominated case where the light distribution follows a Beer's law, exponential distribution, the shape of the pressure wave, will also be an exponential decay with depth.

These photomechanical interactions (thermoelastic expansion, ablative recoil, and explosive vaporization)

cause stress waves, which propagate with the speed of sound, $\sigma = 1500$ m/s, or even faster as in the case of shockwaves. Analogous to heat diffusion, we can define a stress diffusion time:

$$\tau_{str} = \frac{\delta}{\sigma} \quad (11)$$

where τ_{str} is the stress confinement time, δ is the penetration depth of the light in the tissue, and σ is the speed of sound in the medium. If the laser pulse length, τ_p , is shorter than the time it takes the stress wave to propagate out of the irradiated volume, large peak stresses can be reached. These stresses may contribute to the ablation mechanism and may inflict damage to adjacent tissues, even far away from the irradiated zone. The criterion for so-called stress confinement is given by:

$$\tau_{pulse} < \tau_{str} \quad (12)$$

Confinement (and with it the determination if the effect of the laser is considered pulsed or not) depends on the laser pulse length and the optical penetration depth (i.e., absorption coefficient). In Fig. 3, the pulse length versus penetration depth is shown; marked in it are various common lasers with their characteristic pulse durations and optical penetration depths (recall that these depend on the wavelength). For example, from this figure, it follows that the interaction of both the XeCl excimer and the Holmium:YAG laser are thermally confined, but not stress confined.

Explosive Surface Evaporation and Phase Explosion

Regardless of whether conditions of stress confinement are met (and hence thermoelastic expansion-induced pressure waves are created), short laser pulses when

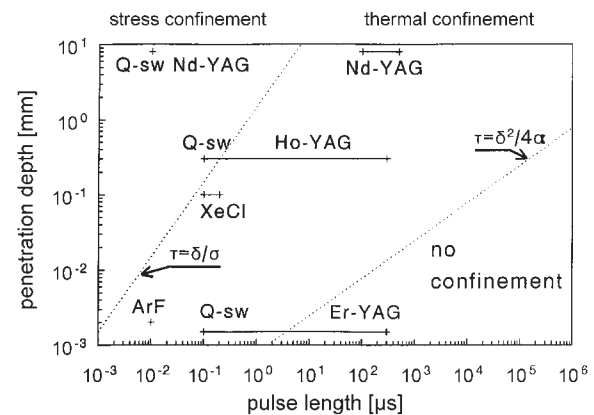


Fig. 3 Confinement regimes for laser-tissue interactions. (Figure adapted from Ref. [16].)

containing sufficient energy can cause the explosive phase change of water. This explosive surface vaporization or phase explosion results in the forceful ejection of debris. In turn, the conservation of momentum dictates that the force of the ejected debris must be compensated by an equal force in the opposite direction. The net result is that if explosive ablation occurs, a second source of pressure wave generation, namely the ablative recoil force, causes forceful compressive stress waves into the tissue. The amplitude of these pressure waves scales with the laser radiant exposure.

Plasma-Induced Ablation

A last regime of laser-tissue interactions involves the use of ultrashort (ps and fs) lasers. Using these type of laser devices, the peak irradiances are so high ($>10^{12}$ W/cm²) that the electric field of the light permeating the medium is capable of ionizing the material, a process known as plasma formation or dielectric breakdown. This can occur even when the linear absorption of the wavelength used is negligible. The result is the generation of a small cloud of free electrons and positively charged ions, known as a plasma. In turn, once formed, the plasma becomes self-absorbing, causing the remaining laser energy to be efficiently coupled in the growing plasma. Prerequisites for the occurrence of a plasma are: 1) a very short laser pulse (in order to increase the power, even when only small pulse energies are used); and 2) a small spot size (the irradiance scales with the inverse of the spot radius squared. The resulting plasma typically involves small volumes (several microns), but may in this microscopic volume reach local temperatures exceeding thousands of degrees. In the moments immediately following the laser pulse, the plasma energy diffuses to the surrounding tissue (water). The result is virtually always the formation of a pressure or shock wave and the rapid vaporization of tissue water, resulting in an explosive ablation process. Notable medical applications of this technology include the use of Q-switched Nd:YAG lasers (few ns, 1064 nm) in cataract surgery and more recently, the use of a fs laser for the creation of bladeless corneal flaps in LASIK surgery.^[9] In the latter case, the near-IR laser is focused inside the corneal stroma at the desired depth and each fs pulse generates a rapidly expanding and collapsing microbubble of approximately 3 μ m in diameter, which separates corneal layers. The laser beam is scanned over the cornea at 20 kHz and adjacent voids together create a clean hinged flap. Other exciting applications of plasma-induced microevents have been reported in cellular manipulation and subcellular microsurgery.^[15] A detailed

discussion of the physics of laser-induced plasma formation and photodisruption is beyond the scope of this article and the reader is referred to.^[2]

BIOLOGICAL EFFECTS

Regardless of the source of damage (thermal, mechanical, or even photochemical), biological tissue typically responds to lethal damage with either a necrotic or apoptotic response. The exact decision of whether a cell will move down the necrotic or apoptotic pathway in response to laser-induced damage is largely unknown. Some evidence exists that oxidative stress, as generated, for example, in photodynamic therapy, shows a marked upregulation of apoptotic markers.^[10] Similar sublethal damage has been shown to upregulate the expression of molecular chaperones such as the heat shock protein family, in particular hsp70.^[11,12] In many lethal thermal damage scenarios, the biological response can be classified as a coagulation necrosis, followed by a normal wound healing response that includes formation of granulation tissue, scar formation, and scar remodeling.^[13] The promise of lasers to substantially increase wound healing speed or cosmetic appearance of scars has thus far remained elusive.

SUMMARY

Long pulses cause collateral thermal damage to tissue adjacent to the ablation crater while short pulses cause collateral mechanical damage. Thermal damage can be the result of thermal diffusion or direct irradiation of tissue at subthreshold levels. Mechanical damage can be the result of explosive vaporization or laser-induced pressure transients. The latter can originate by thermoelastic expansion, ablative recoil, or as a side effect of rapid explosive vaporization. The pulse duration and depth of penetration determine whether the interaction of the laser pulse with tissue is thermally confined and/or mechanically confined.

ARTICLES OF FURTHER INTEREST

Laser Transmyocardial Revascularization; Optics, Biomedical

REFERENCES

1. Welch & Van Gemert. *Optical Thermal Response of Laser-Irradiated Tissue*; Plenum Press, 1995.
2. Niemz, M. *Laser-Tissue Interaction*; Springer-Verlag: Berlin, 1996.

3. Ochsner, M. Photophysical and photobiological processes in the photodynamic therapy of tumours. *J. Photochem. Photobiol.*, **B 2000**, 39 (1), 1–18.
4. Wachtlin, J.; Heimann, H.; Behme, T.; Foerster, M.H. Long-term results after photodynamic therapy with verteporfin for choroidal neovascularizations secondary to inflammatory chorioretinal diseases Graefes. *Arch. Clin. Exp. Ophthalmol.* **2003**, 241 (11), 899–906.
5. Monroe, W.T.; McQuain, M.M.; Chang, M.S.; Alexander, J.S.; Haselton, F.R. Targeting expression with light using caged DNA. *J. Biol. Chem.* **1999**, 274 (30), 20895–20900.
6. Thomsen, S.L. Pathologic analysis of photothermal and photomechanical effects of laser-tissue interactions. *Photochem. Photobiol.* **1991**, 53 (6), 825–835.
7. Incropera & DeWitt. *Introduction to Conduction*; John Wiley & Sons: New York, NY, 1990.
8. Vogel, A.; Venugopalan, V. Mechanisms of pulsed laser ablation of biological tissues. *Chem. Rev.* **2003**, 130 (2), 577–644.
9. Ratkay-Traub, I.; Ferincz, I.E.; Juhasz, T.; Kurtz, R.M.; Krueger, R.R. First clinical results with the femtosecond neodymium-glass laser in refractive surgery. *J. Refract. Surg.* **2003**, 19 (2), 94–103.
10. Oleinick, N.L.; Morris, R.L.; Belichenko, I. The role of apoptosis in response to photodynamic therapy: What, where, why and how. *Photochem. Photobiol. Sci.* **2003**, 1 (1), 1–21.
11. Beckham, T.; Mackanos, M.A.; Crooke, C.; Takahashi, T.; O'Conner-Rodwell, C.E.; Contag, C.H.; Jansen, E.D. Assessment of cellular response to thermal laser injury through bioluminescence imaging of hsp70. *Photochem. Photobiol.* **2004**, 79 (1), 76–85.
12. Desmettre, T.; Maurage, C.A.; Mordon, S. Heat shock protein hyperexpression on chorioretinal layers after transpupillary thermotherapy. *Invest. Ophthalmol. Vis. Sci.* **2001**, 42 (12), 2976–2980.
13. Wu, N.J.; Jansen, E.D.; Davidson, J.M. Comparison of MMP-13 expression in free-electron laser and scalpel incisions during wound healing. *J. Invest. Dermat.* **2003**, 21 (4), 926–932.
14. Esenaliev, R.O.; Oraevsky, A.A.; Letokhov, V.S.; Karabutov, A.A.; Malinsky, T.V. Studies of acoustical and shock waves in the pulsed laser ablation of bio-tissue. *Lasers Surg. Med.* **1993**, 13 (4), 470–484.
15. Venugopalan, V.; Guerra, A., III; Nahen, K.; Vogel, A. Role of laser-induced plasma formation in pulsed cellular microsurgery and micromanipulation. *Phys. Rev. Lett.* **2002**, 88 (7), 078103.
16. Jacques, S.L. Laser-tissue interactions: Photochemical, photothermal and photomechanical. *Surg. Clin. N. Am.* **1992**, 72, 531–558.



Laser Transmyocardial Revascularization

James W. Jones
Nancy A. Crigger
Bruce W. Richman

Baylor College of Medicine, Houston, Texas, U.S.A.

INTRODUCTION

Evidence continues to mount that Transmyocardial Laser Revascularization (TMR) is a safe and effective treatment for angina, as more and more patients have had their angioplasties and coronary artery bypass grafts experience failure over time. Important questions about TMR nevertheless remain unanswered. TMR appears to relieve the symptoms without affecting the underlying atherosclerotic disease, and its method of action has not been established. Like angioplasty and coronary bypass, TMR is in fact a palliative therapy rather than a cure for atherosclerotic heart disease. The most plausible current formulations of TMR's therapeutic mechanism posit a reconfiguration of the microcirculation, with blood shunted from epicardial to endocardial areas. These unresolved issues notwithstanding, TMR benefits patients with end-stage coronary disease and represents a pioneering effort to remodel the microcirculation of patients with arteriosclerotic occlusive disease.

BACKGROUND

Beck attempted the first direct surgical treatment of ischemic heart disease in 1934 when he tried to improve the microcirculation by suturing the left greater pectoral muscle onto the surface of a poorly perfused heart.^[1] Others continued the effort using creative but unscientific methods, and with varying degrees of success.^[1] Vineburg's procedure showed the best results in externally augmenting the microcirculation,^[2] but this was soon rendered obsolete by the advent of coronary artery bypass to treat macro-circulatory obstructions. TMR is a return to direct treatment of the microcirculation. The approach has several theoretical advantages for treating all patients with ischemic heart disease, because suitable target arteries needed for bypass or balloon angioplasty procedures are not required. This will become more important in coming years, because an increasing number of bypassed and angioplastied patients have exhausted their target vessels and are no longer candidates for conventional procedures.

The concept of TMR emerged from an erroneous hypothesis that the human heart contained sinusoids connecting the ventricular chamber with the myocardium. This concept stimulated mechanical attempts to create additional ventricular chamber connections,^[3] but the idea was discredited when it was observed that the new passages degenerated into scar tissue and closed. The advent of surgical lasers suggested the possibility that better channels might remain open and patent as new blood conduits.^[4] In 1984, after 15 years of preparation, Mirhoseini used a relatively low-powered CO₂ laser to create channels through areas he could not bypass during myocardial revascularization procedures.^[2] He had excellent clinical results in 12 patients, and reported channels on postoperative ventriculograms. These early results, coupled with the development of a high-power 850-watt CO₂ laser, encouraged several prestigious surgical groups to begin transmyocardial laser revascularization as an independent beating heart procedure.^[5,6] These pioneering efforts were anchored in the belief that new channels immediately improved myocardial blood flow by allowing seepage of oxygenated blood in the ventricular chamber into the myocardial wall.

Animal studies in dogs, originally intended to compare healing properties of different lasers, showed that scar tissue developed in every instance, not an artificially created sinusoid.^[7,8] The hearts of patients who underwent TMR but succumbed to noncardiac causes were histologically examined with no evidence of patency in any of more than 200 laser channels.^[8] Sophisticated blood flow studies using microspheres showed closure within minutes.^[9] A therapy with an unproven hypothesis had proven effective in reducing angina, but the putative mechanism of action upon which the therapy had been based was completely discredited. TMR was and still is an empirical therapy in search of a therapeutic mechanism.

Current Hypotheses on Mechanism of Action

The most prominent currently favored hypotheses of TMR's mechanism of action include angiogenesis and denervation. Denervation is suggested by the

conspicuous absence of angina immediately upon postoperative emergence from anesthesia.^[5,6,10] Al-Sheikh compared PET scans before and after TMR in eight patients, finding evidence for sympathetic denervation in all.^[11] Reports of both denervation and nondenervation in dog hearts following TMR were confused by other variables in two recently published studies.^[12-14] Denervation would be an explanation consistent with the clinical course, effective in reducing angina, but unsatisfactory as a therapeutic modality. There are less expensive and less invasive means of accomplishing the same goal. The real danger of cardiac denervation in ischemic heart disease is elimination of the early warning function of angina, producing potentially deadly silent ischemia. We reviewed exercise tolerance tests before and after both TMR and percutaneous myocardial revascularization (PMR) and found that silent ischemic changes were no more frequent than in the preoperative studies.^[15] The additional time patients were able to spend on the treadmill was not attributable to the absence of angina; from lack of angina, the percentage of people showing electrocardiographic evidence of ischemia without symptoms was the same pre- and postoperatively.

Angiogenesis is an evident part of the histopathology of healing from the laser injuries in every study.^[16-21] This neovascularization is clearly visible on photomicrographs, with open channels communicating to intramyocardial vessels in humans.^[18] In an ischemic canine model, no differences were seen in blood flow either acutely or two months after TMR in resting animals. When stressed at the two-month interval, however, the animals with TMR averaged 40% greater blood flow than control animals,^[22] indicating that functional connections were allowing blood to reach the endocardial areas. The cardiac microcirculatory anatomy determined by corrosion scanning electron microscopy runs in cable-like bundles surrounding myocyte layers with extensive connections longitudinally, but not from superficial to deep.^[23] Combining these findings, the likely effect of TMR is injury-induced angiogenesis facilitating redistribution of flow to areas of marginal perfusion during periods of greater demand. A recent study used magnetic resonance imaging and found a more modest but important benefit. Both the perfusion and global performance of the heart were unchanged, but the lasered areas were prevented from scarring, as indicated by the absence of fixed perfusion defects in the TMR patients.^[24]

THE PROCEDURE

Between October 1996 and December 2002, we performed 125 transmyocardial revascularization operations

using the Holmium-YAG laser. 115 male patients with an average age of 62 years and an average ejection fraction of 0.37 received TMR laser surgery. At the time of surgery, patients were prescribed an average of five cardiac-related medications and had 4.5 failed conventional revascularization attempts; all had at least one previous coronary bypass procedure. Eighty-seven had one or more serious comorbid conditions, including prior stroke, COPD with steroids, morbid obesity with sleep apnea, renal failure, leukemia, Lupus Erythematosus, or peripheral vascular disease requiring amputation.

Initially, a Holmium-YAG laser (CardioGenesis Corporation, Sunnyvale, California), gated to the electrocardiographic R wave, firing in triple 2-Joule bursts, was used. Later in the series, a non-gated Holmium-YAG laser manufactured by Eclipse was used. The change in lasers occurred following the merger of manufacturer Eclipse with CardioGenesis. The merged companies maintained the CardioGenesis name.

The candidate for transmyocardial revascularization surgery must receive complete disclosure of all potential risks and complications and understand that this procedure has a currently unknown mechanism of action. The patient's emotional status must also be evaluated preoperatively. Many are at the end-stage of their heart disease and may have psychological sequelae to multiple failed operations, catheter interventions, and other therapies. Qualified patients should be provided with appropriate reassurance, sedatives, and/or anxiolytics to avoid increased stress and its deleterious effect on the heart.

Candidates for transmyocardial revascularization surgery are almost always high risk patients with severe heart disease. They nevertheless vary considerably in their mortality risk. Highest-risk patients have severe triple-vessel disease without any directly perfused areas.^[25] Their inflow is severely limited, consistent with the concept of redistributed blood flow, and is compatible with the redistribution of blood flow mechanism concept. Such patients may have operative mortality rates as high as 25%,^[25] an excessive price for symptom relief. Patients with congestive heart failure (CHF) are high risk^[6] because they lack myocardial reserve. Poor early appreciation of how TMR initially injured the heart without creating an immediate new source of blood flow resulted in inappropriate application of TMR therapy with sometimes disastrous results. Uncompensated CHF has since been recognized as a contraindication for TMR. Patients with well-compensated CHF did not have a significantly higher operative mortality, but a greater number died during a six-month follow-up period.^[26] An ejection fraction of at least 25% has now been incorporated into the standard guidelines for



determining patient suitability for TMR therapy. Using TMR or conventional bypass grafting on patients with unstable coronary syndromes increases risk. In one sizable study, patients with unmanageable angina for acute syndromes had a five-fold increase in operative mortality.^[27] Others confirmed the increased mortality in severely unstable patients.^[26] A stabilizing period of several weeks after such episodes has been recommended before TMR is attempted. We have administered TMR in desperation to ten such patients who could not be stabilized with medical therapy; all survived.^[28] This is not to recommend TMR in acutely unstable patients, but to confirm that TMR can achieve good results in carefully selected clinical situations.

Specific operative steps can be taken to minimize the potential complications associated with this complex surgical group. As suggested by Horvath,^[5] insertion of an epidural catheter and the tranquil induction of anesthesia reduces pain and limits coronary artery spasms. Operating time is decreased and the procedure made much easier by the insertion of a double lumen or Univent endotracheal tube collapsing the left lung to provide adequate exposure. Radial and femoral arterial lines, a Swan-Ganz catheter, I.V. lines, and a Foley catheter are all standard. Initially, we placed access lines in the femoral veins in the event of an intraoperative emergency, but have thus far had no occasion to use them and have stopped placing them. The femoral arterial line is started to provide a consistently reliable arterial pressure and ready access should it become necessary to insert an intraaortic balloon pump (IABP) during surgery. We have had to insert the IABP postoperatively in two patients. Five patients were taken to the operating room with IABPs in place.

Transesophageal echocardiography (TEE) is available for intraoperative cardiac monitoring and to observe the appearance of intraventricular bubbles during laser application as an indication that we have entered the ventricular chamber. TEE was very useful as a monitor of cardiac function, providing early detection of transient myocardial ischemia in the early stages of the procedure's development. We have contained operative costs by not using TEE in our last 75 patients, with no discernable adverse effects. When monitoring intraventricular bubbles to determine chamber entry, TEE is operator-dependent and less reliable than the surgeon's ability to visualize pulsatile blood flow from the new channel. Transient intraoperative ischemia can nevertheless be detected early by monitoring ST-segment elevation. A signal from the anesthesiologist gives the surgeon an opportunity to rest the heart. Channeling too quickly can destabilize cardiac homeostasis.

After induction of anesthesia, the patient is placed in the left lateral decubitus position at an angle

thirty degrees from perpendicular to the operating table. Proper positioning presents the anterior area of the chest as a flat surface and provides the most direct access and exposure to the heart. The left side of the thorax, the anterior half of the abdomen, and the left groin area are all prepped in the sterile field. The entire area is circumscribed with towels and a thyroidectomy drape is placed to the anterior midline just below the left nipple and slanted upward toward the axilla.

The skin is infiltrated with 1% lidocaine without epinephrine prior to incision. Anterior, lateral, or muscle-sparing posterolateral thoracotomies all provide equally satisfactory access. It is our impression that patients in whom the anterior approach is used have less postoperative pain and recover more rapidly than those treated with other incisions.

A 4-inch or smaller incision is made over either the fourth or fifth interspace. The fifth interspace can be used to provide direct exposure of the heart's inferior surface when it is the only area to be treated, but may not be adequate for ready access to the base of anterior and lateral areas. An incision in the fourth interspace is recommended for entry to all other divisions of the left ventricle, either alone or in combination with the inferior region. Others have used the fifth and sixth interspaces satisfactorily.

We have arbitrarily divided the left ventricle into three main regions as served by the principal coronary arteries: anterior (anterior descending), lateral (Circumflex), and inferior (right coronary). These zones are further divided into a total of seven sections (anterior, anterior-lateral, lateral-anterior, lateral, lateral-inferior, inferior-lateral, and inferior), facilitating localization of channel placement. These are in the geometric form of roughly 1-centimeter-wide strips oriented longitudinally, the total encompassing the entire area of left ventricular surface. If one does not use a ruler initially for perspective, it is possible to "overchannel." The newly created laser channels are 1-millimeter vaporized tracts surrounded by a 1–2-millimeter zone of necrosis and varying radii of edema. Overchanneling can create a significant area of dead heart muscle.

After skin incision we use two Gelpy retractors to gain exposure and insert a Toufier retractor when the intercostal space is entered. The retractor is opened slowly and division between the intercostal muscles is extended as tension develops. The pectoralis muscles and the latissimus dorsi muscle are preferably undermined rather than divided. In most patients this permits access to the entire left ventricle, although exposure is limited to one region at a time. When this technique is inadequate to permit satisfactory exposure, the upper rib is "shingled" laterally with a rib shear. When wound closure is begun the rib is

reapproximated with a single No. 5 wire suture, as are any unintended rib fractures.

Upon entering the thorax one may find numerous adhesions of the pleura to the anterior cardiac surface. Adhesions containing vascular structures are ligated with either clips or suture and not cauterized. Electrocautery should be limited to structures that are away from the phrenic nerve, and this practice is best instituted before the nerve is visualized, since its course can vary considerably in patients who have had previous cardiac surgery. From the midpoint of its pericardial course the phrenic nerve travels slightly anteriorly. Therefore, the safest dissection either keeps it continuously in sight or, if that becomes difficult, turns decidedly anterior as the apex of the heart is approached.

Once the pericardium is opened, the dissection proceeds parallel to and about 1 centimeter anterior to the phrenic nerve. This approach creates a small pocket inferiorly that maximizes exposure of the inferior and inferolateral areas of the heart, with limited traction on the phrenic nerve. Dissection of the pericardium off the heart is facilitated by placing traction sutures as exposure is developed. As the dissection continues, the anterior pericardium is opened from the center and folded back to expose the area for laser channeling. The separation of the pericardium is not usually difficult, even in patients who have undergone sternectomy and flap placement for sternal osteomyelitis after coronary bypass surgery, and this approach avoids the problem areas.

Upon completion of the dissection, the ventricular apex is directly facing the surgeon standing on the left side of the operating table. Maneuvering the heart within the dissected space can expose all of the left ventricle, one segment at a time. Exposure is carried anteriorly to the anterior descending artery, and inferiorly to the posterior descending artery, depending on the areas to be treated. A 1-inch malleable retractor covered by a 4-by-4 sponge tied in place has been most helpful in providing gently stabilized retraction in the limited available space.

The Holmium-YAG laser probe is placed on the epicardial surface and fired until the probe enters the ventricular chamber. The probe must not be forced, as this movement will allow entry without firing the laser; instead, the probe is advanced after each succeeding laser firing only until the ventricular chamber is reached. Since this procedure is performed on a beating heart, the surgeon must be able to control the probe to avoid inadvertent damage to normal myocardium. The probe can be stabilized either by using vascular forceps to grasp the epicardium, or by firing the laser once on the epicardial surface to create a small surface hollow. The probe itself can be held either directly in the hand or in an instrument designed

for the purpose. The probe holder used in this study is malleable and works well in reaching areas at the base of the anterior and lateral zones and the more apical inferior zone, but it requires too much space to easily position the laser to the base of the heart in the inferior area. That area is best reached by positioning the probe with two vascular forceps.

The conventional intent of laser surgery is formation of channels that enter the ventricular chamber perpendicular to the myocardial wall. Although the concept of primary conduit formation from the ventricular chamber is questionable, we hesitate to alter the empirical method that may or may not have contributed to the technique's clinical success. There are four indicators of laser entry into the ventricular chamber: 1) pulsatile spurting blood; 2) bubble formation as seen on transesophageal echocardiography; 3) a change in the sound of the laser; and 4) feeling the probe enter the ventricular chamber. Because the thickness of the ventricle decreases from the basal areas toward the apex of the heart, the closer one is to the apex the easier it becomes to penetrate to the ventricular chamber. The probe used in our study measured 17 mm in length and may not be long enough to traverse the ventricular wall in basal areas, especially in patients with hypertension-related cardiac hypertrophy.

Pulsatile blood flow is the best indicator of ventricular penetration. Bleeding stops with simple finger pressure in virtually all cases, and once stopped does not resume. This has been true in patients on aspirin, with low platelet counts, and with thinned ventricular walls. In rare cases when the procedure may be prolonged due to bleeding channels, slight pressure on a hemostatic pad placed over the channel is effective (Fig. 1).

The proper amount of myocardium to be treated varies among patients and among treating surgeons.

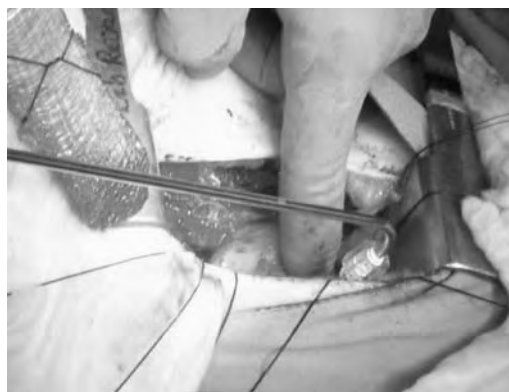


Fig. 1 Applying finger pressure to a Holmium-YAG laser channel in a patient.

L

Areas shown to be ischemic by thallium scan are typically treated with minimal extension onto adjacent normal myocardium. Arguments for directed rather than global treatment include the historically limited success of prophylactic surgical procedures. Most of these patients are quite ill and frail, and the surgical procedures to which they are subjected should impose the least possible physiological stress. The directed treatment strategy has shown good short-term results, and the procedure can be repeated in the event of recurrent angina. It is nevertheless essential that all ischemic areas be treated; directed treatment is unlikely to be successful otherwise.

The areas to be treated are determined preoperatively. We have had excellent correlation between angiographic anatomy and the reversible ischemia shown by thallium scan. Areas with nonreversible perfusion defects were not treated. We did see some discrepancy between the thallium scan reading and the anatomy, in that the distal inferior surface ischemia was often ischemia of the apical segment.

The patient's intraoperative clinical status must be controlled at all times. We have twice seen decreased regional wall motion on TEE associated with dramatic ST-T wave changes on the monitor screen and concomitant substantial rises in pulmonary pressures. In both cases, the deteriorating clinical status responded to suspension of surgical manipulation, removing the chest retractor, and allowing time for the anesthesiologist to intensify vasodilator therapy and increase the depth of anesthesia. Ischemic problems did not recur in these patients and there was no postoperative evidence of myocardial damage.

Perioperative arrhythmias are of considerable concern in this patient population. The operating room should be equipped with external sterile and unsterile defibrillator paddles and easily accessible internal paddles, though none of our patients has required defibrillation or displayed dysrhythmias during laser firing. In four patients, however, the initial positioning of the probe in preparation for firing the laser resulted in considerable ectopy. This was easily controlled by firing the laser in the target area, thus scoring and desensitizing the myocardium before the actual channeling began.

At the end of the procedure, all areas of dissection and laser treatment are inspected with irrigation and any accumulated clot is removed. The pericardium was left open in all cases.

Before reinflation of the left lung, intercostal blocks are placed under direct vision in the intercostal space of the incision and at least two interspaces above and below. The blocks are placed as far posterior as possible. A 3½-inch spinal needle is usually adequate, but in exceptionally large patients a 5-inch needle can be used. A mixture of 60 ml of .25% Marcaine

and 80 milligrams of Depo-Medrol usually provides one to three days of pain relief. A properly placed single 24 F chest tube is adequate and seems to be less painful than the larger tubes commonly used for other thoracic procedures. Intraoperative care to avoid damage to the left lung facilitates early chest tube removal.

Patients are typically extubated in the ICU shortly after arrival. A continuous nitroglycerin infusion is maintained at the maximum level tolerated by the patient's blood pressure. The propofol drip is discontinued and pain is controlled by the epidural. The anesthesiologists are responsible for epidural analgesia and treating the associated side effects of the epidural until its removal. About half of the patients require only oral analgesic supplementation, but for those requiring more pain control, PCA morphine sulfate is the preferred analgesic. When carbon dioxide retention is a short-term problem, intravenous ketorolac tromethamine is helpful.

If the chest tube drains less than 100 ml in the preceding 12 hours, it is removed the evening of surgery or on the first postoperative morning. In all our patients the chest tubes were removed in less than 24 hours. One patient required drainage of a recurrent pleural effusion. The nitroglycerin drips are slowly reduced (5 ml/30 minutes), and a nitroglycerin patch applied when the drip reaches 10 ml/hr. Those who tolerate discontinuation of the drips are transferred to the floor. Alert patients are allowed to administer their own previously prescribed home medications. Oral analgesics are ordered every three hours to avoid delays associated with prn requests.

Ambulation is begun on postoperative day one, with most discharges one to two days later. We have observed that post-TMR patients who are asymptomatic tend to reduce medications too rapidly and without appropriate medical consultations; their discharge instructions should strongly caution patients against medication reduction without medical supervision.

The beneficial results of TMR therapy are often dramatic, provided that perioperative risks can be minimized. Our approach stresses techniques that reduce pain and limit the chance of perioperative myocardial infarction. At 12-month follow-up these patients have shown significant angina class improvement when compared to both control patients and their own preoperative status.^[29] Likewise, average exercise tolerance with a Modified Bruce Protocol has been significantly better among TMR patients than among medication controls. Holmium-YAG laser channeling of the myocardium has been effective in reducing anginal pain and improving quality of life in ischemic heart disease patients who are without alternative therapeutic options.^[26,29-34]

CLINICAL RESULTS

Clinical outcomes determine the usefulness of medical therapies. At this time, TMR is a specialized therapy. Surgical therapy of ischemic heart disease began with attempts to manipulate the microcirculation and progressed to coronary bypass surgery as a macroscopic method of redirecting blood flow around circulatory obstructions. Percutaneous Transarterial Catheter Arterioplasty restored blood flow within the obstructed channels themselves.

Subjective Clinical Results

All of the studies published thus far have reported that angina pectoris is relieved to a significant extent by TMR. In longitudinal uncontrolled trials, the CO₂ laser^[5,30] and the newer Holmium-YAG laser^[33,34] showed clinical dramatic improvement. While this data stimulated further use and development of TMR, it did not fulfill the compelling need for evidence-based data to support wide implementation of a major new invasive treatment method. The proponents of this therapy started three controlled trials in the United States and one controlled trial in England, all of which were published in 1999.^[29-32] The results of all four studies showed highly significant differences in the percentage of TMR and control patients registering improvement of at least two scale angina classes 12 months after operation. Schofield and colleagues^[31] reported similar findings but did not consider the improvements to be clinically important. Other randomized trials showed improvement that doubled or tripled the magnitude of those results.^[29,30,32] When a blinded method of determining the patient's angina class was used, the scores were markedly lower than in the nonblinded ratings, but remained significantly better than those for the medically treated group.^[32] Patients in all four randomized studies reported higher quality of life scores, consistent with improving symptoms. During the 12-month follow-up, TMR patients required fewer hospitalizations for cardiac events.^[29-31] The symptoms, if not necessarily the process, of ischemic heart disease showed positive responses to TMR. The possibility of a surgical placebo effect exists, but is unlikely to account for these results. The treated population was effectively bedridden prior to operation, and most were able to resume normal lifestyles postoperatively. Surgical placebo effects typically persist for two to three months, and affect no more than 30% of patients.^[10]

Objective Clinical Results

Several investigators report evidence of increased perfusion on thallium scan,^[30,35,36] but other equally

well-conducted work reports no significant changes.^[37] One group^[30] showed a significant number (20%) of patients with better-perfused segments after TMR, but 26.4% of these study patients had incomplete follow-up and 19% of the patient evaluations could not be used, weakening the value of the overall report. One possible explanation for the differences in studies and the overall failure to show improvement in perfusion is the variability in thallium scans over time.^[37] In support of the improved perfusion concept, dobutamine stress echocardiography has shown improvement of global contractile reserve in a porcine model of chronic ischemia^[38] as well as in humans.^[30,39]

Exploring this apparent inconsistency, Frazier used PET scanning, the most sensitive test for myocardial perfusion, and reported that patients showed an increase in the relative perfusion of the subendocardial to epicardial areas after TMR.^[6] Most important, these findings were seen in patients who had no significant improvement in thallium scans. If these conflicting reports are accepted as fact, it is conceivable that the thallium scan lacks the necessary sensitivity, or that flow redistributes from areas of adequate perfusion to underperfused myocardium, such as epicardium and subendocardium, through new microvascular connections.

Clarification and enhancement of the mechanism of action of TMR holds promise as a forerunner of future therapies to treat vascular disease through remodeling the microcirculation. At present it illustrates that treatment of illness without treating the disease using empiric therapies remains a valuable staple of medicine.

ARTICLES OF FURTHER INTEREST

Laser-Tissue Interaction; Optics, Biomedical

REFERENCES

1. Jones, J.W.; et al. Surgical myocardial revascularization. *Surg. Clin. North Am.* **1998**, 78 (5), 705-727.
2. Mirhoseini, M.; Cayton, M.M. Transmyocardial laser revascularization: Historical background and future directions. *J. Clin. Laser Med. Surg.* **1997**, 15 (6), 245-253.
3. Mittal, C.M. Profulla Kumar Sen: His contributions to cardiovascular surgery. *Texas Heart Inst. J.* **2002**, 29 (1), 17-25.
4. Chester, M.; Hammond, C.; Leach, A. Long-term benefits of stellate ganglion block in severe chronic refractory angina. *Pain* **2000**, 87 (1), 103-105.
5. Horvath, K.A.; et al. Transmyocardial laser revascularization: Operative techniques and clinical results at



- two years. *J. Thorac. Cardiovasc. Surg.* **1996**, *111* (5), 1047–1053.
6. Frazier, O.H.; et al. Myocardial revascularization with laser. Preliminary findings. *Circulation* **1995**, *92* (Suppl. 9), II58–II65.
 7. Burkhoff, D.; et al. Histologic appearance of transmyocardial laser channels after 4 1/2 weeks. *Ann. Thorac. Surg.* **1996**, *61* (5), 1532–1534. discussion 1534–1535.
 8. Krabatsch, T.; et al. Histological findings after transmyocardial laser revascularization. *J. Card. Surg.* **1996**, *11* (5), 326–331.
 9. Kohmoto, T.; et al. Does blood flow through holmium: YAG transmyocardial laser channels? *Ann. Thorac. Surg.* **1996**, *61* (3), 861–868.
 10. Jones, J.W.; et al. Holmium: YAG laser transmyocardial revascularization relieves angina and improves functional status. *Ann. Thorac. Surg.* **1999**, *67* (6), 1596–1601. discussion 1601–1602.
 11. Al-Sheikh, T.; et al. Cardiac sympathetic denervation after transmyocardial laser revascularization. *Circulation* **1999**, *100* (2), 135–140.
 12. Kwong, K.F.; et al. Transmyocardial laser treatment denervates canine myocardium [see comments]. *J. Thorac. Cardiovasc. Surg.* **1997**, *114* (6), 883–889. Discussion 889–890.
 13. Kwong, K.F.; et al. Nontransmural laser treatment incompletely denervates canine myocardium. *Circulation* **1998**, *98* (Suppl. 19), II67–II71. Discussion II71–II72.
 14. Hirsch, G.M.; et al. Transmyocardial laser revascularization does not denervate the canine heart. *Ann. Thorac. Surg.* **1999**, *68* (2), 460–468. Discussion 468–469.
 15. Myers, J.; et al. Do transmyocardial and percutaneous laser revascularization induce silent ischemia? An assessment by exercise testing. *Am. Heart J.* **2002**, *143* (6), 1052–1057.
 16. Hughes, G.C.; et al. Induction of angiogenesis after TMR: A comparison of holmium: YAG, CO₂, and excimer lasers. *Ann. Thorac. Surg.* **2000**, *70* (2), 504–509.
 17. Kohmoto, T.; et al. Evidence of vascular growth associated with laser treatment of normal canine myocardium [see comments]. *Ann. Thorac. Surg.* **1998**, *65* (5), 1360–1367.
 18. Schweitzer, W.; et al. Transmyocardial laser revascularization. Histopathology of laser channels in 10 postoperatively deceased patients 1 to 18 days after treatment with a CO₂ laser. *Pathologie* **1997**, *18* (5), 374–384.
 19. Gassler, N.; Rastar, F.; Hentz, M.W. Angiogenesis and expression of tenascin after transmural laser revascularization. *Histol. Histopathol.* **1999**, *14* (1), 81–87.
 20. Galli, M.; et al. Percutaneous transmyocardial revascularization with holmium laser in patients with refractory angina: A pilot feasibility study. *G. Ital. Cardiol.* **1999**, *29* (9), 1020–1026.
 21. Fleischer, K.J.; et al. One-month histologic response of transmyocardial laser channels with molecular intervention. *Ann. Thorac. Surg.* **1996**, *62* (4), 1051–1058.
 22. Yamamoto, N.; et al. Angiogenesis is enhanced in ischemic canine myocardium by transmyocardial laser revascularization. *J. Am. Coll. Cardiol.* **1998**, *31* (6), 1426–1433.
 23. Ono, T.; et al. Scanning electron microscopic studies on microvascular architecture of human coronary vessels by corrosion casts: Normal and focal necrosis. *Scan Electron Microsc.* **1986**, (Pt. 1), 263–270.
 24. Holmstrom, M.; et al. Wall motion and perfusion analysis of transmyocardial laser revascularization. *SCJ, Scand. Cardiovasc. J.* **2003**, *37* (2), 91–97.
 25. Burkhoff, D.; et al. Factors correlating with risk of mortality after transmyocardial revascularization. *J. Am. Coll. Cardiol.* **1999**, *34* (1), 55–61.
 26. Guleserian, K.J.; et al. Quality of life and survival after transmyocardial laser revascularization with the holmium: YAG laser. *Ann. Thorac. Surg.* **2003**, *75* (6), 1842–1847. Discussion 1847–1848.
 27. Hattler, B.G.; et al. Transmyocardial laser revascularization in the patient with unmanageable unstable angina. *Ann. Thorac. Surg.* **1999**, *68* (4), 1203–1209.
 28. Jones, J.; Guleserian, K.J.; Camillo, C.J.; DAmiano, R.J.; Moon, M.R. Is transmyocardial revascularization too high-risk for high-risk patients. *Ann. Thorac. Surg.* **2003**, *75*, 1847–1848. Discussion.
 29. Allen, K.B.; et al. Comparison of transmyocardial revascularization with medical therapy in patients with refractory angina [see comments]. *N. Engl. J. Med.* **1999**, *341* (14), 1029–1036.
 30. Frazier, O.H.; March, R.J.; Horvath, K.A. Transmyocardial revascularization with a carbon dioxide laser in patients with end-stage coronary artery disease [see comments]. *N. Engl. J. Med.* **1999**, *341* (14), 1021–1028.
 31. Schofield, P.M.; et al. Transmyocardial laser revascularisation in patients with refractory angina: A randomised controlled trial [see comments] [published erratum appears in *Lancet* 1999 May 15;353(9165):1714]. *Lancet* **1999**, *353* (9152), 519–524.
 32. Burkhoff, D.; et al. Transmyocardial laser revascularisation compared with continued medical therapy for treatment of refractory angina pectoris: A prospective randomised trial. ATLANTIC investigators. Angina treatments-lasers and normal therapies in comparison. *Lancet* **1999**, *354* (9182), 885–890.
 33. Diegeler, A.; et al. Transmyocardial laser revascularization using the Holium-YAG laser for treatment of end stage coronary artery disease. *Eur. J. Cardio-Thorac. Surg.* **1998**, *13* (4), 392–397.
 34. Milano, A.; et al. Transmyocardial revascularization with a holmium laser: Preliminary results. *G. Ital. Cardiol.* **1997**, *27* (10), 1011–1018.
 35. Horvath, K.A.; et al. Transmyocardial laser revascularization: Results of a multicenter trial with transmyocardial laser revascularization used as sole therapy for end-stage coronary artery disease. *J. Thorac. Cardiovasc. Surg.* **1997**, *113* (4), 645–653. Discussion 653–654.
 36. Moosdorf, R.; et al. Transmyocardial laser revascularization in stable and unstable angina pectoris. *Herz* **1997**, *22* (4), 198–204.
 37. Burkhoff, D.; Jones, J.W.; Becker, L.C. Variability of myocardial perfusion defects assessed by thallium-201 scintigraphy in patients with coronary artery disease

- not amenable to angioplasty or bypass surgery. *J. Am. Coll. Cardiol.* **2001**, *38* (4), 1033–1039.
38. Hughes, G.C.; et al. Improved perfusion and contractile reserve after transmyocardial laser revascularization in a model of hibernating myocardium. *Ann. Thorac. Surg.* **1999**, *67* (6), 1714–1720.
39. Donovan, C.L.; et al. Improvement in inducible ischemia during dobutamine stress echocardiography after transmyocardial laser revascularization in patients with refractory angina pectoris. *J. Am. Coll. Cardiol.* **1997**, *30* (3), 607–612.



Ligament Repair: Animal Models

Robert C. Bray

*Department of Surgery, University of Calgary, Calgary, Alberta, Canada, and
McCaig Centre for Joint Injuries and Arthritis Research*

Catherine A. Leonard

Kevin R. Forrester

Paul T. Salo

Department of Surgery, University of Calgary, Calgary, Alberta, Canada

INTRODUCTION

Ligament damage is a frequent sequel to joint trauma, and anyone experiencing pain and disability after a significant ligament injury will attest to the urgent need of finding new treatments to accelerate and promote improved healing. Damage to the anterior cruciate (ACL) and medial collateral (MCL) ligaments of the knee is particularly common. Once thought of as inert structures, ligaments are now seen to participate actively in joint function and homeostasis by means of complex mechanical, physiological, and cellular interactions. Ligament injuries that do not heal therefore result in abnormal loading of other joint structures, leading to degeneration and the development of osteoarthritis.

Joint trauma generally produces complex injuries involving multiple tissues. For research purposes, the use of animal models allows the degree and type of ligament injury to be isolated and controlled, and permits the application of analysis techniques that cannot be used in a clinical setting. This entry gives a brief overview of ligament structure, function, and response to injury, followed by a discussion of specific animal models of ligament injury, healing, and reconstruction. The application of new technology for the assessment of ligament healing and physiology is also discussed.

NORMAL LIGAMENT STRUCTURE AND FUNCTION

Anatomy, Structure, and Composition

Skeletal ligaments are dense, white, hypovascular tissues connected at both ends to bone. Some ligaments, such as the MCL of the knee, are distinct cord-like structures, but many consist of several discrete bands that merge into one another. These distinct bands can provide functional heterogeneity within otherwise homogeneous tissue. At their bony

attachments (insertions), ligaments show unique structural alterations, with varying cell shapes and fiber orientations and a gradual change in the composition and organization of the fibrous tissue that becomes calcified and merges into the bone surface.

Ligament tissue consists of closely spaced collagen fibers aligned predominantly along the long axis of the ligament. Under polarized light microscopy, the fibers show a regular sinusoidal pattern along their length that is termed “crimp.” Fiber crimping is thought to act as a mechanism for load damping. Collagen fibers within ligaments are organized parallel bundles or fascicles delineated by a cellular layer, the *endoligament*.^[1] Ligament cells comingle with the collagen fibers within a fascicle and are organized into rows that run along the long axis of the ligament. The ligament fascicles are encased in a cellular layer, the *epiligament*, which is populated by both nerves and blood vessels. Adjacent rows are interconnected by cytoplasmic processes joining one cell to the next via functional gap junctions. This arrangement produces a complex interconnected array of cells that is thought to extend from one end of the ligament to the other. This array has been termed the *cellular matrix*.

Ligament tissue is relatively hypovascular, with approximately 2% of its volume typically occupied by blood vessels.^[2] The majority of blood vessels are located in the *epiligament*, forming vascular plexuses with a random vessel orientation that show a typical branching pattern into small vascular channels. Deeper ligament tissue is less vascular, and the vessels are more linearly aligned with the long axis. These deep vessels run parallel to the collagen bundles in the ligament, and show infrequent anastomoses in the tissue substance. Blood vessels in ligaments are often associated with nerves showing a similar longitudinal organization.

The innervation of ligaments originates mainly from neighboring muscular, cutaneous, and articular nerve trunks.^[3] In the *epiligament*, nerve fibers and blood vessels form a rich plexus, sending branches that penetrate the ligament. Most nerve fibers do not enter

the ligament, but terminate as nerve endings on the surface. Myelinated sensory, unmyelinated sensory, and sympathetic nerve fiber types have been identified in ligament tissue. In addition to the classical neurotransmitters (acetylcholine, amines, amino acids, and nitric oxide), 10 or more slow-transmitting neuropeptides have recently been identified in ligaments: sensory neuropeptides (substance P and calcitonin gene-related peptide), opioid peptides, and sympathetic peptides (neuropeptide Y and vasoactive intestinal polypeptide).

Functions of Ligaments

Ligament function has been traditionally viewed as purely mechanical, stabilizing and controlling joint articulation to distribute forces and protect other tissues. More recent evidence suggests that ligaments also play a significant role in sensory feedback, and that the mechanical properties of ligament are influenced by neurovascular function.

Ligaments show nonlinear mechanical behavior, being relatively compliant under low-load conditions and showing increasing stiffness with increasing tensile load.^[4] The crimped structure and viscoelastic properties of collagen fibers both contribute to this behavior, which is also determined by age, gender, and loading history of the tissue. Recently, it has been suggested that ligaments function under repeated or sustained low loads in normal daily activity, and respond to this environment by the phenomenon of creep (elongation) rather than relaxation. Excessive ligament creep can result in significant joint laxity and subsequent tissue injury, and is therefore of particular clinical relevance.

Recent experiments showing that ligament viscoelastic properties are affected by tissue water content have brought the importance of understanding physiological processes in connective tissues into a new focus. Tissue water content is determined by a variety of factors under vasomotor control, and ligament creep behavior is dependent on the state of hydration. Increasing hydration leads to increased creep, while displacing tissue water with sucrose leads to decreased creep.^[5] The physiologic regulation of connective tissue water content is therefore crucial to the maintenance of mechanical function within normal parameters.

There is good evidence to suggest that the mechanoreceptors found in ligaments play a substantial role in motor control. Sensory input from ligament receptors appears to be important for normal balancing of muscle forces, inducing reflex joint stiffening in response to applied loads, and increasing tension in both flexor and extensor muscles. This stiffens and stabilizes the joint, thereby protecting the articular surfaces from mechanical damage.

The physiology of normal ligaments is complex with regard to cell-matrix structure, as well as heterogeneous with respect to neural and microvasculature elements; this has considerable implications for the normal regulation of these tissues.

MODELS OF LIGAMENT INJURY AND HEALING

Injury to ligaments occurs as a direct result of trauma, or subsequent to altered joint mechanics in unstable or osteoarthritic joints. The knee joint in particular has been the focus of numerous studies, partly owing to the frequency with which clinical injury occurs, and partly because the internal and external ligaments of the knee demonstrate significant differences in their response to injury: the MCL shows a healing response similar to other soft tissues and can regenerate its structure to closely resemble the original. Conversely, one of the major stabilizing ligaments of the knee, the ACL shows a very minor healing capability, and injury to this ligament generally results in joint instability and osteoarthritis. This has led to the development of models that create an injury to the ligament tissue while retaining mechanical stability of the joint. Ligament injury models fall into four main categories: ligament division (rupture or transection), joint laxity (micro-injury), reconstruction, and in vitro techniques. Models of ligament injury are illustrated in Figs. 1 and 2.

Rupture or Transection

Ligament injury results from tearing of the collagen bundles by varying degrees. Subject to sufficient force, collagen fibers are completely torn, and the entire ligament is severed. Early work on ligament injuries attempted to reproduce clinical trauma by the crude method of subjecting the joint to larger-than-normal forces, resulting in injuries of varying extent and location. Later studies focused more specifically on controlling the site of damage; for example, Frank et al. used a model of MCL injury where a braided steel suture was placed under the ligament and pulled upwards perpendicular to the joint to cause a midsubstance rupture.^[6] This model of ligament rupture is similar to a clinical injury in that the ends of the ligament are frayed, producing a characteristic “mop-end” appearance (Fig. 1A). However, because other structures in the joint and regions of the ligament remote from the actual injury site are also subjected to aberrant forces, interpretation of the results from these experiments is complex.

While rupture models of ligament injury may closely approximate clinical injuries, the degree and location of damage vary widely. To give better control over the

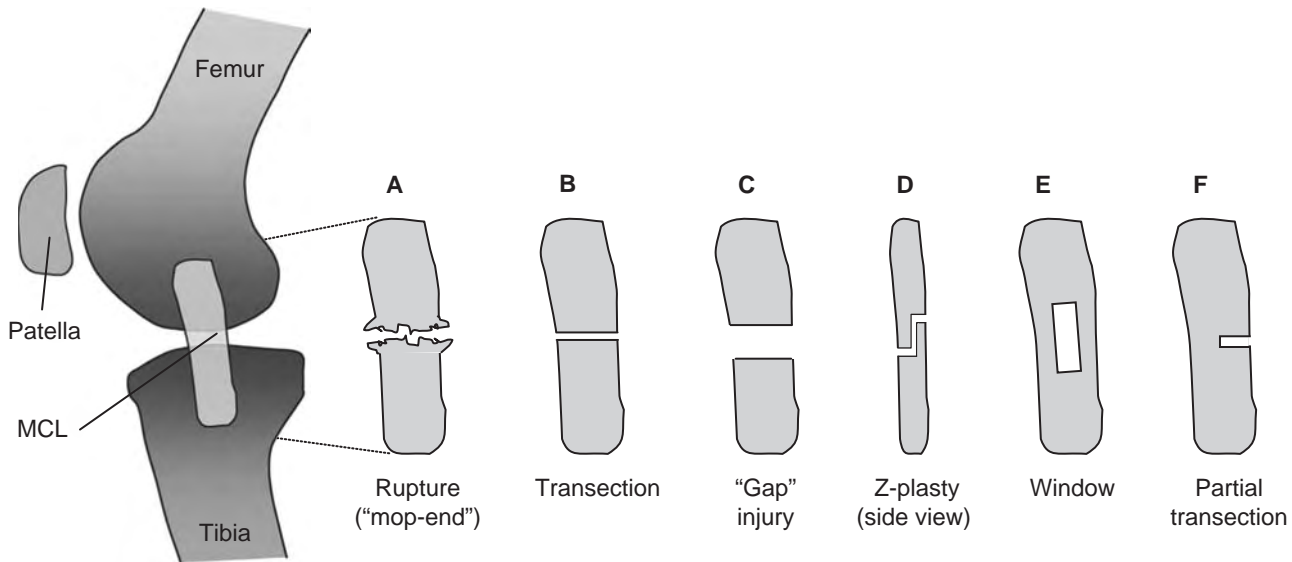


Fig. 1 Schematic showing surgical models of MCL injury and healing.

site and extent of ligament injury, surgical transection models were developed. Complete division of the MCL or ACL is easily achieved under surgical anesthesia using a scalpel, where the ligament is transected at its midpoint, and two uniform blunt ends are produced (Fig. 1B). Initially, frayed ends were thought to be prerequisite for normal MCL healing; however, it was shown that there was no difference in the mechanical or histological properties of a healing ligament whether the ends were initially frayed or cut.^[7]

During healing of the MCL, scar tissue forms that is thought to affect the mechanical properties of the ligament, showing disorganization of the collagen fiber structure and increased vascularity that reduces tensile strength of the tissue (see below for details of ligament healing responses). To assess the contribution of

scar tissue to the mechanical properties of healing ligaments, two further models of transection were developed in the rabbit MCL,^[8] one aimed at increasing the amount of scar tissue by separating the two cut ends of the ligament ("gap" injury, Fig. 1C), and one aimed at reducing the amount of scarring by maximizing the contact area of the cut ends ("Z-plasty," Fig. 1D). This study demonstrated that both gap and contact injuries healed with scar tissue, recovering their biomechanical viscoelastic behaviors to 68–92% of contralateral control ligaments by 14–40 weeks. However, "contact scars remodeled more quickly, recovered laxity more quickly, and were generally closer to contralateral control ligaments than gaps in terms of their structural strength, stiffness, and material behaviors after 40 weeks of healing."^[8] The

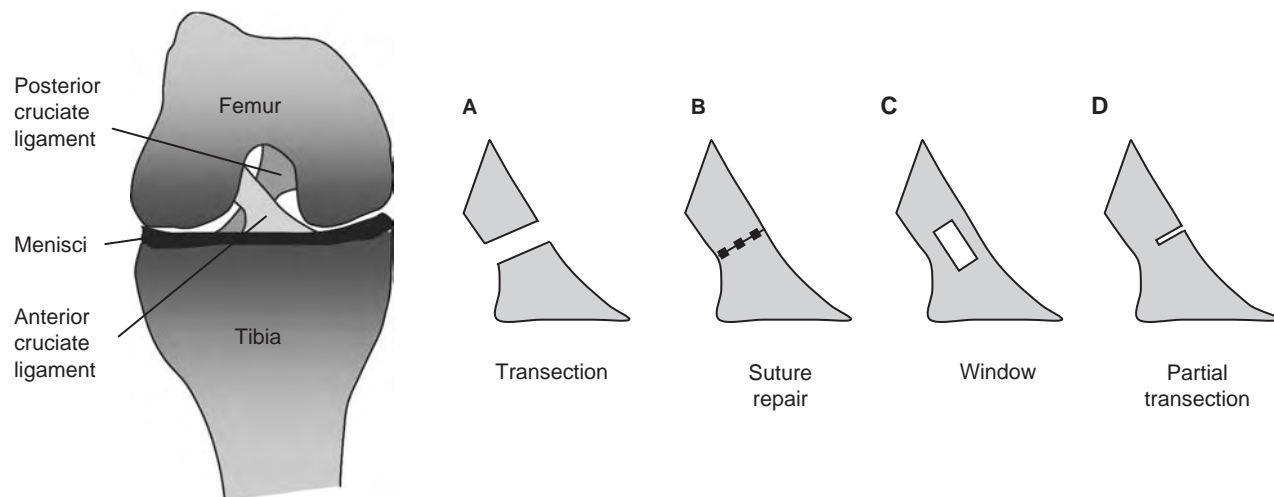


Fig. 2 Schematic showing surgical models of ACL injury, joint laxity, and osteoarthritis.

authors concluded that there were differences in the early rate and possibly in the later quality of ligament healing when cut rabbit MCL ends were in proximity.^[8] Other models of MCL injury such as tissue windows and partial sections have been developed for comparison with ACL injury models detailed below (Figs. 1 and 1F).

In contrast to the MCL, the ACL of the knee generally shows poor healing capacity in both clinical patients and animal models. In fact, it may be stated that for all intents and purposes, the transected ACL never heals.^[9,10]

It has proven quite useful to study the transected ACL in animal models (Fig. 2A) and compare its healing response with an extra-articular ligament such as the MCL that shows much better recovery after injury. Early attempts at severing the ACL used a stab incision to the anterior surface of the joint using a scalpel blade; however, the success of this model was very limited because of inconsistent injuries to the ACL and added damage to the cartilage and other joint structures. A more recently established and much more consistent approach to ACL transection is to open the joint cavity by a lateral capsulotomy, retract the fat pad, isolate the ACL with a hooked probe, and completely transect the ligament using a scalpel blade (Fig. 2A). Completeness of ACL transection is tested by drawing the tibia anteriorly relative to the femur, and the wound is sutured closed. This provides a consistent model of injury, where the ACL is completely transected while minimizing damage to other tissues.

A possible reason for poor healing of the ACL is the intra-articular location of the ligament, where the ACL provides the major mechanical resistance to anterior translation of the tibia. When the ACL is ruptured or cut (Fig. 2A), the biomechanics of the knee are significantly altered, and the motion of the tibia relative to the femur is greatly increased (“joint laxity”). Healing of the ACL may be impossible in this environment because of lack of apposition of the ligament ends. This hypothesis has led to the development of knee injury models that either physically reattach the cut ligament ends (suture repair, Fig. 2B), or attempt to induce the formation of scar tissue while retaining the mechanical integrity of the joint (tissue window and partial transactions, Figs. 2 and 2D). Histological, mechanical, and physiological measures of ACL healing can then be compared with control ACLs, or with the MCL to determine differences in the reaction to injury.

Joint Laxity and Osteoarthritis

In addition to studying the healing responses of the ACL itself, the ACL-deficient model of joint injury

(Fig. 2A) is also useful in two other areas of study: microtrauma to other joint structures as a consequence of joint laxity, and in the longer term, development of osteoarthritis in the joint.

Disruption of the ACL produces increased joint laxity and altered joint mechanics, and forces placed on the joint that would normally be resisted by the ACL are redistributed throughout the remaining connective tissues. The chronically abnormal biomechanical environment in an ACL-deficient knee induces a series of maladaptive physiological and mechanical alterations in secondary joint stabilizing structures that may in turn undergo injury as a result. Specific results of chronic instability are pain and inflammation, hyperemia, and angiogenesis in articular soft tissue, including ligaments. Assessing the function of the MCL in an ACL-deficient joint allows examination of the tissue response to an altered mechanical environment without inducing a direct injury to the MCL. This model has produced interesting results that are discussed below.

Long-term untreated joint laxity induces a cycle of recurring inflammation and progressively alters the tissue mechanics, and ultimately leads to the development of osteoarthritis in the joint. The ACL-deficient knee of several species has been used to study the mechanics, cartilage properties, and bone remodeling in osteoarthritis joints.

Ligament Reconstruction

In light of the poor healing potential of the ACL even when the ends of the injured ligament are apposed, alternative treatments involving replacement or reconstruction of the ligament have been developed and clinical ACL reconstruction has become almost standard practice.^[11] In animal models, replacement of the injured ACL with biologic or synthetic materials has been tested in a variety of species with significant success.

Autografting of the ligament with other tissues such as hamstring tendon, semitendinosus tendon, and patellar tendon from the same animal is common. This precludes graft rejection but causes morbidity of the donor site that may be problematic, especially in smaller species. As an alternative allografting of tissue from donor animals has been used, but there is increased risk of infection and graft rejection. Synthetic materials such as polyester ribbon can also be used to stabilize the joint; however, the viscoelastic properties of such materials are very different from biological tissue and can lead to difficulty in interpreting the results. Nonetheless, ACL reconstruction models, especially autografts, have provided a wealth of information on the relation between



joint laxity and the development of microinjury and osteoarthritis.

In Vitro Models

An alternative approach to in vivo models of ligament injury is the use of tissue explants or cell cultures to study the properties and migration of ligament fibroblasts and cellular responses to a variety of growth factors. While these models have received little attention in the past, recently developed culture techniques have shown differences in the responsiveness of cells from ACL and MCL tissues to cytokines and neuro-peptides.^[12] Medial collateral ligament fibroblasts derived from tissue explants show greater proliferation, migration, and collagen synthesis than ACL cells and are more responsive to externally applied factors such as tumor growth factor-B. These intrinsic differences in the properties of MCL and ACL fibroblasts may go some way to explaining the differential healing potentials of these two ligaments.

ANIMAL SPECIES IN LIGAMENT RESEARCH

In the clinical setting, joint trauma generally results in complex injuries involving several tissues, and it is therefore difficult to evaluate the contribution of the various structures to the injury response. The use of animal models allows specific types of ligament injury to be induced under controlled conditions where the degree of healing can be directly correlated to changes in biomechanics, joint physiology, and homeostasis. The choice of animal species is dependent on a number of variables, some of which are nonscientific factors such as ease of handling and housing costs. In general, larger species are used for studies of biomechanics, where the size and articulation of the joints more closely resembles human joints. Smaller species are more commonly used for histologic assessments of healing and innervation.

Sheep and Goats

These species have relatively large knee joints, and are well suited for studies of the ACL healing and reconstruction.^[13–15] The anatomy and flexion mechanics of sheep and goat knees more closely resemble that of humans than do smaller species, because of the stiffness of the stifle joint. Sheep are somewhat amenable to exercise training on a treadmill; however, the specialized housing needs of these animals may preclude their use at some institutions.

Dogs

Dogs were used in some of the earliest studies of ligament injury,^[16] and are still widely used, because they are easy to handle and can be trained in a variety of exercise protocols. There is also a large existing database on canine ligament healing from veterinary research. The disadvantages of the model are the relatively high costs of animal purchase and housing and restrictions placed on experimental protocols by institutions because of the companion nature of the species.

Rabbits

There is a large database on the biomechanics, histology, and physiology of ligament injuries in the rabbit.^[4,6] The New Zealand White rabbit is a commonly used strain because of its relatively low purchase cost, basic housing requirements, and ease of handling. In addition, adult rabbits weigh between 5 and 7 kg, and so the knee joint of these animals is large enough for meaningful mechanical testing to be done, yet small enough that histological sectioning of the entire joint is still feasible. Models of MCL injury, ACL injury (joint laxity), and ACL reconstruction have all been established in the rabbit model.

Rats

Rats have not traditionally been used for ligament studies, mainly because of the small size of their joints, and the fact that rats were thought to grow continually throughout their life so that bone and soft tissues are being continually remodeled. However, rodent models have recently been used successfully to study the innervation of joints, ligament, and tendon.^[3,17] In addition, we now know that although the epiphyseal plates of rats do not ossify, they do become dormant shortly after sexual maturity; so these animals may be more commonly used in studies of ligament biology because of the potential cost savings and the wealth of biological probes and techniques available for use in rat tissues.

LIGAMENT HEALING RESPONSES

The effectiveness of ligament healing depends upon a variety of factors, including location of the ligament, type and extent of injury, and the animal species used for the study. Ligament anatomy and healing has been assessed using a wide variety of techniques, including biomechanical assessment of tissue viscoelastic properties, histological and immunohistochemical studies

of structure and innervation, and functional indicators of vascular physiology such as perfusion imaging. In this section, a comparison of healing in the MCL and ACL is used to illustrate the application of some of these techniques.

Phases of Ligament Healing

The injury response of ligament tissue follows a similar pattern to other soft tissues, although the time taken for the process to occur is relatively long.^[18] Some ligaments such as the MCL show almost complete healing, while others such as the ACL demonstrate almost no injury response. In ligaments where an injury response occurs, healing can be divided into three phases: hemorrhage and inflammation, proliferation, and remodeling.

Initially, torn ligament ends are bridged by a blood clot that is absorbed as the injury site is infiltrated by cells; damaged material is degraded, and repair is initiated by cellular and vascular factors. Blood flow to the ligament is significantly increased and remains elevated for several weeks.^[2]

In the proliferative phase, infiltration by hypertrophic fibroblast cells produces “scar” tissue, a hypercellular collagen matrix that joins the torn ends of the ligament. This scar tissue is structurally disorganized and mechanically weak, and contains collagen fibers that are randomly oriented and of smaller diameter than normal ligament tissue. With time, collagen fibers and bundles are found to be realigned with the long axis of the ligament. Scar tissue is relatively vascular, and it has been suggested that the perforation of a normally highly organized tissue by myriad small, structurally weak blood vessels contributes significantly to the inferior mechanical properties of healing ligaments.

After a ligament injury, it takes months to years for tissue remodeling. Components of the scar tissue change their composition and orientation, and the cellularity and the vascularity gradually decrease, but the tissue remains mechanically weak. Some matrix components return to near-normal levels, but the remodeled tissue still contains different proteoglycans, collagen fiber types, and diameters that appear to create a new functional baseline. In the long term, the extent of recovery after a ligament injury depends on many factors including the degree of joint movement, apposition of the ligament ends, and neurovascular responses.

Biomechanics

Testing the material properties and mechanical behaviors of bone–ligament–bone complexes has been widely used to assess the recovery of ligament function

after injury. The major function of ligaments is to resist and distribute forces applied to the joint, and so the viscoelastic properties of the tissue have been extensively studied in a variety of injury models. For a detailed review of the biomechanics of knee ligaments see Ref. [4].

In comparing the mechanical properties of the MCL and the ACL, the most commonly used protocol is to leave the ligament attached to the femur and tibia, which are then clamped into the testing device. Using a bone–ligament–bone complex rather than the ligament alone eliminates the problems of securely gripping the tissue and the concern that the presence of clamp–tissue interfaces may affect the measurement of its properties. A bone–ligament–bone complex provides a structural unit to which load can be applied to determine mechanical behaviors such as stress, strain, and creep. The low-load viscoelastic properties of ligaments have recently received much attention, because it is now believed that these properties are most representative of ligament function under normal conditions.

The biomechanical properties of MCL scars are generally inferior to normal ligaments, demonstrating a slow recovery of many properties. During cyclic loading, or a single static load, the stress relaxation of scar tissue is more than normal and therefore relies more heavily on other tissues to compensate. The creep properties of the tissue are reduced, and scarred ligaments show twice as much creep when subjected to cyclic loads far below their failure strength. The length of the ligament is reestablished during later remodeling, but at 14 weeks after a “gap” injury, failure strength of the MCL complex in the rabbit is still only half of sham-operated controls.^[7,8]

Following complete transection of the ACL, there is generally insufficient scar formation to allow mechanical testing; the ligament stumps remain separated and may eventually be absorbed. However, models of ACL repair or replacement such as suture repair, partial injury, or reconstructions have been compared with control ACLs. One study in dogs showed that the failure and tensile strength of ACL reconstructions using patellar tendon autografts was initially only 10% of control values; these properties did improve over time, but never recovered to within normal range.^[11] Factors such as graft fixation, position, composition, and tension must be taken into consideration in the interpretation of such results.

Histology

Ligament structure, cellularity, and innervation have been characterized using a wide range of histological and immunohistochemical techniques. Simple techniques such as hematoxylin and eosin staining can



be used to study the appearance of ligament cells, blood vessels, and organization of the collagen fiber structure, providing a measure of the metabolic and mechanical status of the tissue. Recent ultrastructural studies have revealed ligament cellular organization and communication via gap junctions,^[1] while immunohistochemical staining for nerve fibers and vasoactive peptides has provided new information on the role of vasomotor reflexes and innervation in ligament function and healing.^[3,17,19]

Following injury to the MCL, there is significant inflammation followed by the formation of scar tissue that is highly cellular and vascular and shows a derangement of collagen fiber structure. During the remodeling phase, the fibers gradually realign themselves with the long axis of the ligament, and the cellular organization of the tissue is restored.^[18] In contrast, injuries to the ACL show almost no cellular response, even when the ligament ends are physically apposed. The MCL also shows a significant remodeling response in ACL-deficient knees, presumably in response to increased and aberrant forces in a lax joint. The ligament thickens along most of its length, demonstrating similarities in appearance to scar tissue formed after direct injury.

Immunohistochemistry has been used to detect nerve fibers and vasoactive neuropeptides in ligaments.^[3,19] Fig. 3 shows nerve fibers in the rabbit MCL

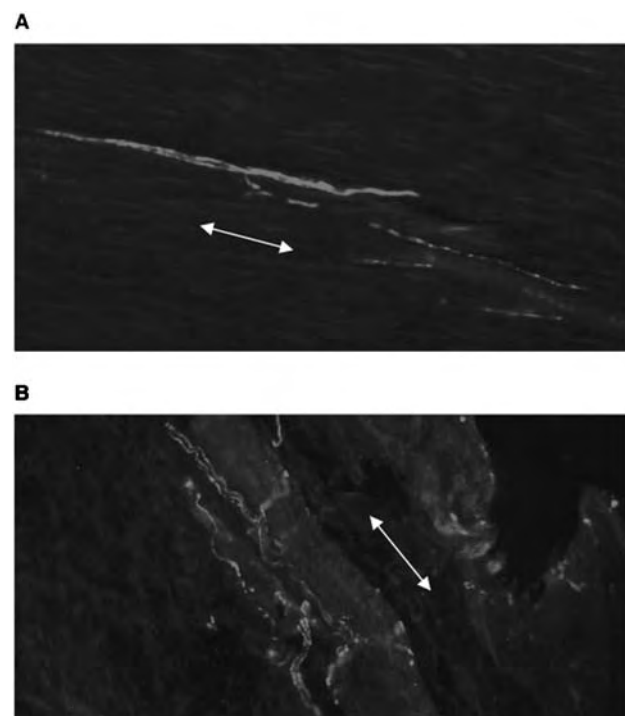


Fig. 3 Immunostaining of rabbit MCL for (A) the neuronal marker PGP9.5 and (B) the vasoactive peptide CGRP. Arrows indicate the longitudinal axis of the ligament, and in (B), is also aligned with an adjacent blood vessel.

immunostained for the neuronal marker protein gene product 9.5 (PGP 9.5) and the neuropeptide CGRP.^[3] Six weeks after gap injury in the MCL, calcitonin gene-related peptide (CGRP) immunoreactivity is increased in the scar tissue. The vasoactive and trophic properties of neuropeptides in injured ligaments likely contribute to their healing potential, affecting angiogenesis, inflammation, and scar tissue formation.

Vascular Physiology

The blood flow, vascularity, and water content of ligaments significantly affect their mechanical properties, and a vascular response after injury is crucial to tissue healing. The vasoactive responses of blood vessels in ligaments have remained largely unknown, because of a lack of accurate measures of normal tissue vascularity and vasomotor activity. A range of techniques have been used to measure ligament blood flow, beginning with vascular casting techniques in the late 1980s and progressing through colored

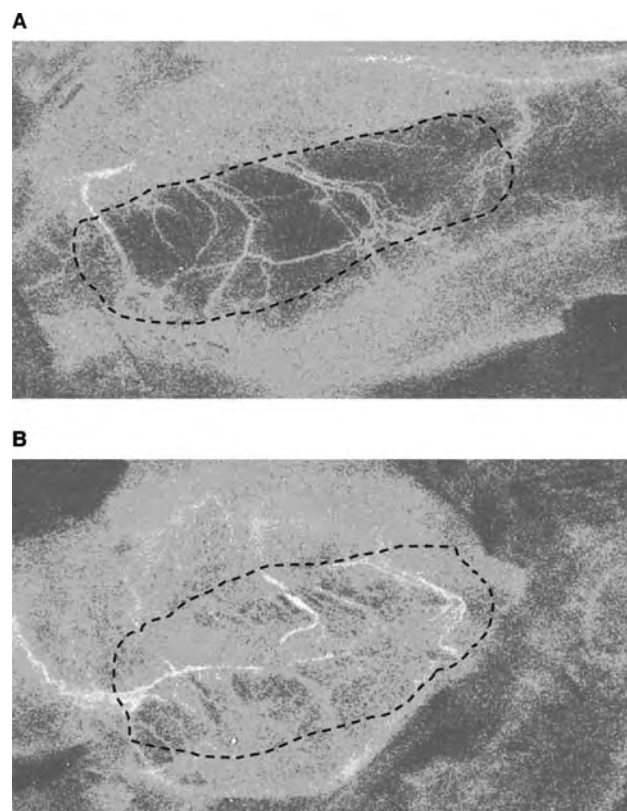


Fig. 4 Laser speckle perfusion images showing blood flow to the rabbit MCL in (A) control and (B) ACL-deficient joints. Areas of low flow are in blue, areas of moderate flow are in green and in yellow, and areas of high flow are in red. (From Ref. [17].)

microsphere infusion and laser Doppler imaging.^[2] Most recently, new technology for the real time assessment of tissue blood flow, laser speckle perfusion imaging (LSPI), allows visualization of ligament blood vessels and microcirculation in living tissue with minimal invasiveness. This enables physiologic studies of vasoregulatory processes under a variety of clinical and research conditions.^[20]

There is a clear difference in the vascular response of the MCL and the ACL after partial transection injuries in the rabbit, which almost certainly contributes to the lack of healing in the ACL. Using colored microsphere infusion, we have shown that the MCL responds to partial transection with a massive increase in blood flow and vascularity, peaking at six weeks after injury and returning toward control levels by 16 weeks.^[9] In contrast, blood flow to the ACL after partial transection was not significantly different from sham-operated animals at any time.

The application of LSPI to ligament injuries allows a more rapid and dynamic assessment of tissue vascular responses.^[21] Fig. 4A shows the blood supply to a normal rabbit MCL measured by LSPI. The perfusion image is color coded so that regions of low blood flow appear blue, regions of moderate flow appear green and yellow, and areas of high flow appear red. In Fig. 4A, individual blood vessels supplying the MCL appear green and yellow on a blue background. The larger epiligamentous vessels appear as a branching network over the surface of the ligament, while the capillary microcirculation is seen as background perfusion without distinct vessels. Fig. 4B shows the injury response of the MCL in a lax joint, six weeks after surgical transection of the ACL. Blood flow and vascularity of the MCL is significantly increased, especially in the areas of greatest tissue thickening.

Application of new techniques such as LSPI allows assessment of connective tissue vascular responses at high speed and high resolution.

CONCLUSIONS

The use of animal models in the study of ligament injury has allowed significant advances in the understanding of structural, mechanical, and physiological changes occurring during connective tissue healing. The use of these models in combination with advances in analysis techniques has the potential to establish new therapeutic regimens following traumatic joint injury.

ACKNOWLEDGMENTS

Dr. Bray is a Senior Scholar of the Alberta Heritage Foundation for Medical Research. Funding was

provided by the Canadian Institutes for Health Research, the Arthritis Society of Canada, and the Whitaker Foundation.

ARTICLES OF FURTHER INTEREST

Tendons and Ligaments, Mechanical Testing of; Tissue Engineering of Ligament; Tissue Engineering of Rotator Cuff Tendons; Tissue Engineering of Tendons

REFERENCES

- Lo, I.K.; Chi, S.; Ivie, T.; Frank, C.B.; Rattner, J.B. The cellular matrix: a feature of tensile bearing dense connective tissues. *Histol. Histopathol.* **2002**, *17*, 523–537.
- Bray, R.C. Blood supply of ligaments. *Orthop. Int.* **1995**, *3*, 39–48.
- Salo, P.T.; Theriault, E. Number, distribution and neuropeptide content of rat knee joint afferents. *J. Anat.* **1997**, *190*, 515–522.
- Woo, S.L.-Y.; Debski, R.E.; Withrow, J.D.; Janaushek, M.A. Biomechanics of knee ligaments. *Am. J. Sports Med.* **1999**, *27*, 533–543.
- Thornton, G.M.; Shrive, N.G.; Frank, C.B. Altering ligament water content affects ligament pre-stress and creep behaviour. *J. Orthop. Res.* **2001**, *19* (5), 845–851.
- Frank, C.B.; Woo, S.L.-Y.; Amiel, D.; Harwood, F.; Gomez, M.; Akeson, W. Medial collateral ligament healing: a multidisciplinary assessment in rabbits. *Am. J. Sports Med.* **1983**, *11*, 379.
- Chimich, D.; Frank, C.; Shrive, N.; Bray, R.C.; King, G.; McDonald, D. No effect of mop-ending on ligament healing: rabbit studies of severed collateral knee ligaments. *Acta Orthop. Scand.* **1993**, *64*, 587.
- Chimich, D.; Frank, C.; Shrive, N.; Dougall, H.; Bray, R. The effects of initial end contact on medial collateral ligament healing: a morphological and biomechanical study in the rabbit. *J. Orthop. Res.* **1991**, *9*, 37.
- Bray, R.C.; Leonard, C.A.; Salo, P.T. Correlation of healing capacity with vascular response in the anterior cruciate and medial collateral ligaments of the rabbit. *J. Orthop. Res.* **2003**, *21* (6), 1118–1123.
- Woo, S.L.-Y.; Young, E.P.; Ohland, K.J.; Marcin, J.P.; Horibe, S.; Lin, H.-C. The effects of transection of the anterior cruciate ligament on healing of the medial collateral ligament. A biomechanical study of the knee in dogs. *J. Bone Joint Surg. Am.* **1990**, *72*, 382–392.
- Newton, P.O.; Horibe, S.; Woo, S.L.-Y. Experimental studies on anterior cruciate ligament autografts and allografts. In *Knee Ligament: Structure, Function, Injury and Repair*; Daniel, D., Akeson, W., O'Connor, J., Eds.; Raven Press: New York, 1990; 389 pp.
- Chun, J.; Tuan, T.L.; Han, B.; Vangsness, C.T.; Nimni, M.E. Cultures of ligament fibroblasts in fibrin matrix gel. *Connect. Tissue Res.* **2003**, *44* (2), 81–87.
- Fabbricani, C.; Mulas, P.D.; Ziranu, F.; Deriu, L.; Zarelli, D.; Milano, G. Mechanical analysis of fixation methods for anterior cruciate ligament reconstruction



- with hamstring tendon graft. An experimental study in sheep knees. *Knee* **2005**, *12* (2), 135–138.
14. Hunt, P.; Scheffler, S.U.; Unterhauser, F.N.; Weiler, A. A model of soft-tissue graft anterior cruciate ligament reconstruction in sheep. *Arch. Orthop. Trauma Surg.* **2005**, *125* (4), 238–248.
 15. Buma, P.; Kok, H.J.; Blankevoort, L.; Kuijpers, W.; Huijskes, R.; Van Kampen, A. Augmentation in anterior cruciate ligament reconstruction—a histological and biomechanical study on goats. *Int. Orthop.* **2004**, *28* (2), 91–96.
 16. Clayton, M.L.; Weir, G.J., Jr. Experimental investigations of ligamentous healing. *J. Bone Joint Surg.* **1969**, *43A*, 1167.
 17. Ackermann, P.W.; Finn, A.; Ahmed, M. Sensory neuro-peptidergic pattern in tendon, ligament and joint capsule. A study in the rat. *Neuroreport* **1999**, *10* (10), 2055–2060.
 18. Lo, I.K.Y.; Thornton, G.; Miniaci, A.; Frank, C.B.; Rattner, J.B.; Bray, R.C. Structure and function of diarthrodial joints. In *Operative Arthroscopy*, 3rd Ed.; McGinty, J.B., Ed.; Lippincot, Williams and Wilkins: Philadelphia, 2003; 41–126.
 19. McDougall, J.J.; Bray, R.C.; Sharkey, K.A. Morphological and immunohistochemical examination of nerves in normal and injured collateral ligaments of rat, rabbit and human knee joints. *Anat. Record* **1997**, *248*, 29–39.
 20. Forrester, K.R.; Tulip, J.; Leonard, C.; Stewart, C.; Bray, R.C. A laser speckle imaging technique for measuring tissue perfusion. *IEEE Trans. Biomed. Eng.* **2004**, *51* (11), 1994–2005.
 21. Miller, D.; Forrester, K.; Leonard, C.; Salo, P.; Bray, R.C. ACL deficiency and incipient osteoarthritis impairs the vasoconstrictive efficacy of neuropeptide Y in articular tissues: a laser speckle perfusion imaging study. *J. Appl. Physiol.* **2005**, *98* (1), 329–333.

Liver, Bio-Artificial

François Berthiaume

Christina Chan

Martin L. Yarmush

Center for Engineering in Medicine, Massachusetts General Hospital, Harvard Medical School, and the Shriners Hospital for Children, Boston, Massachusetts, U.S.A.



INTRODUCTION

The liver is one of the most complex and metabolically active organs in the body and performs many detoxification and protein synthetic functions that are essential to life. Extracorporeal bioartificial liver (BAL) systems consisting of functioning, viable hepatocytes may provide temporary support for patients with acute hepatic failure and save the lives of patients awaiting orthotopic liver transplantation (OLT).

In the past few years, we have seen several clinical studies testing the efficacy of BAL devices. These preliminary studies have provided some promising results, although current-generation liver-assist devices have not yet demonstrated sufficient efficacy and reliability for routine use. Some of the issues that the field of BAL development continues to wrestle with are: 1) how to support a large cell mass without substrate limitations so that cells function with maximum efficiency; 2) how to maintain high levels of stable long-term liver-specific function in an artificial (and potentially inhospitable) environment; and 3) how to minimize the priming volume (or dead space) that must be filled by blood or plasma from the patient. It is becoming clearer every day that a more fundamental understanding of the effect of environmental parameters on hepatocellular function, as well as host-BAL interactions, is necessary before the concept of BAL becomes a reality available at reasonable cost. This article describes the tissue-engineering principles and critical technologies that are relevant to the design of BAL devices, and provides future directions for the development of the next generation of BAL support systems.

CLINICAL RELEVANCE

The Clinical Problem

Approximately 30,000 patients die each year from end-stage liver disease in the United States. About 80% of these patients have decompensated chronic liver disease, typically caused by alcoholism or chronic

hepatitic C infection, and less commonly by a genetic—hepatocellular or anatomic—defect of liver function, or cancer. The other 20% die of acute liver failure (without preexisting chronic liver disease), which has various etiologies, including ischemia-reperfusion injury during liver surgery, acetaminophen poisoning, viral hepatitis, severe sepsis, idiosyncratic drug reactions, etc. Acute liver failure symptoms develop over a period of 6 weeks to 6 months and lead to death in over 80% of the cases, usually from cerebral edema, complications due to coagulopathy, and renal dysfunction. A more severe form of acute liver failure—fulminant hepatic failure—is characterized by a more rapid evolution (2–6 wk).

Orthotopic liver transplantation is the only clinically proven effective treatment for patients with end-stage liver disease. Acute liver patients have the highest priority for donor livers; however, because acute liver failure is an uncommon condition, it only accounts for 12% of OLTs.^[1] The latest data from the United Network for Organ Sharing, the organization that keeps track of organ transplants in the United States, show that the number of liver transplants performed yearly has increased at a constant rate of 240/year since 1992, to reach 4950 in 2000 (Fig. 1). However, the number of patients dying while on the waiting list has increased exponentially at a rate of 16–17% per year. As of August 31, 2002, a staggering 17,436 patients were on the liver transplant waiting list. Currently, 90% of the donated livers are transplanted; thus, any significant increase in the donor supply can only be achieved through dramatic improvements in the methods of organ procurement, as well as expansion of donor criteria.

The majority of donor livers are obtained from brain-dead cadavers that still possess respiratory and circulatory functions at the time of organ retrieval. A large pool that remains largely untapped consists of nonheart-beating cadavers.^[2] In 2001, only 3% of the cadaveric livers were from nonheart-beating donors, although it is estimated that the potential may be on the order of 250,000 annually.^[3] Such livers, however, are prone to primary nonfunction posttransplantation, and much more research is needed to develop

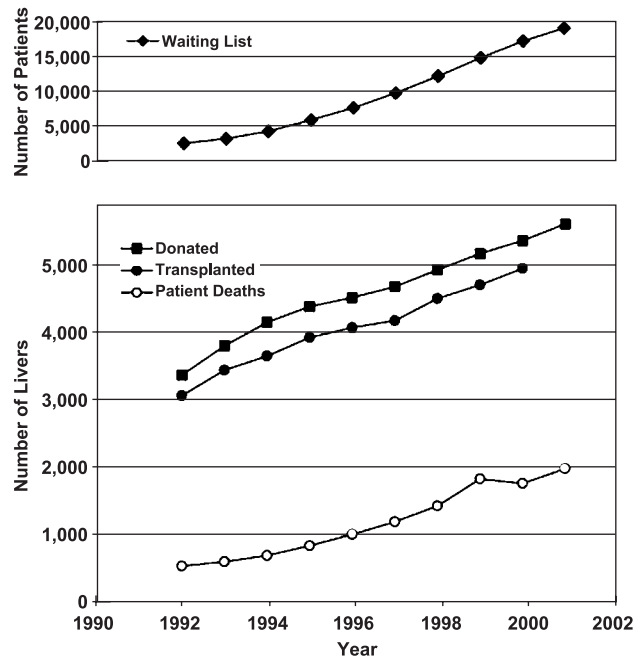


Fig. 1 Evolution of the supply and demand for donor livers in the United States. (Data from Ref. [1].)

techniques that would recondition or repair these organs to make them acceptable donors. Living-related donor transplantation, which involves using a graft obtained by means of a partial hepatectomy from a living donor (usually a parent or relative), is another promising alternative. Nevertheless, split- and living-liver donor techniques represent only about 3% of the total number of transplants performed in the United States,^[4] and are inherently limited because this procedure represents a significant risk for the donor.

All the aforementioned approaches to expand the donor pool are being actively pursued, and have the advantage that they do not require extensive regulatory review, thus shortening the time from concept to practice. Nevertheless, the extent of the problem is such that other alternatives to OLT are desperately needed. Some of the recent and more innovative approaches currently being explored are: 1) xenotransplantation; 2) hepatocyte transplantation; and 3) temporary liver support.

Xeno (Cross-Species Transplantation)

The transplantation of pig or baboon livers into humans has had several failed attempts in the recent past. For this approach to succeed, patients must first overcome hyperacute rejection, a very rapid immune response (~1 hour) that is characteristic of xenotransplantations, and must actively develop

microchimerism, an interaction of two coexisting cell populations leading to the down regulation of both recipient and graft immune systems.^[5] Another major concern is the potential transmission of infectious agents across species.^[6] Recent advances in genetic engineering may allow the modification of the immune response, and sophisticated breeding programs may provide animal breeds devoid of zoonotic diseases. These advances may eventually make xenogeneic livers a viable option for use as a bridge to transplantation, or possibly even as a permanent graft.

Hepatocyte transplantation

Transplantation of liver parenchymal cells (hepatocytes) offers the possibility of treating several patients with one single donor liver. The survival and function of implanted hepatocytes can be improved by incorporating the cells into implantable devices such as microcarriers (beads 200–300 μm in diameter that the cells can attach to), hydrogels, and other polymeric matrices. The most suitable implantation sites appear to be those that provide a microenvironment resembling that of liver, i.e., with a substrate that promotes hepatocyte attachment and a venous blood supply network, such as the splenic pulp and the host liver itself. Furthermore, permanent incorporation may require that the liver have significant preexisting damage and that the transplanted cells have a competitive advantage (e.g., they proliferate faster) v. those in the host's liver.^[7] Studies so far suggest that the efficiency of engraftment is quite low and that a lag time (which may be as long as 48 hours) is necessary before any clinical benefit occurs.^[8] (A 48-hour lag time may be too long in a rapidly deteriorating patient.) On the other hand, this approach offers an attractive prospect for correcting nonemergency conditions such as inherited metabolic defects of the liver.

Other implantable devices, such as capsules and hollow fibers, have been used to protect hepatocytes from the immune system of the host. Some encouraging results have been obtained over a period of a few weeks after implantation, but not beyond, possibly due to deterioration of the biomaterial.^[9,10] A common recurrent problem is the presence of a foreign-body reaction against the capsule material itself, leading to the generation, over a period of days to weeks, of a fibrotic layer around it, compromising metabolite transport between the implanted cells and the surroundings. Besides improvements in the biocompatibility of the material, one avenue that may improve the function of these devices is the use of materials or factors that promote the growth of blood vessels near the surface of the capsule.

Temporary liver support

Temporary liver support is a logical option for patients with acute liver failure where the host's liver has the potential to fully recover, as well as for use as a bridge to transplantation. The concept of temporary liver support is best exemplified by results obtained with auxiliary partial liver transplantation (APLT). Auxiliary partial liver transplantation, currently the only reliable method of temporary liver support, consists of implanting a functioning liver without removing the host's native organ. Auxiliary partial liver transplantation has been used to treat specific inborn liver disorders as well as acute fulminant hepatic failure. In certain cases of hepatic failure, the native liver recovered and the auxiliary liver could be safely removed, hence eliminating the need for lifelong immunosuppression.^[11] Indeed, one of the greatest appeals of temporary liver support is the possibility of spontaneous recovery without the need for transplantation. This would free up donor livers as well as eliminate the need for lifelong immunosuppression that is required by all transplant recipients, translating into significant cost savings for the health care system. Temporary liver support may also be used to maintain the patient's life during the long wait for a donor, as well as prevent liver failure-related complications that may make the patient too ill to tolerate a liver transplantation procedure.

Auxiliary partial liver transplantation depends on the availability of donated livers; there is thus a need to develop an off-the-shelf device akin to the kidney dialysis machine. Early devices for temporary liver support contained no cells (i.e., charcoal adsorption columns), and could not demonstrate therapeutic effectiveness.^[12] Nevertheless, there remains an interest in improving this methodology. Because the liver provides a myriad of detoxifying as well as protein synthetic functions, efforts in this area are mostly dedicated to the development of bioartificial systems incorporating live liver cells.

CELL SOURCE AND THERAPEUTIC CELL MASS

What Is the Therapeutic Cell Mass?

The cell mass required to support a patient during acute liver failure has not been systematically determined. Devices that have undergone clinical testing have used 5×10^9 – 6×10^{10} porcine hepatocytes^[13,14] or 4×10^{10} human hepatoma cells.^[15] In a more recent study on humans with acute liver failure, intrasplenic and intraarterial injections of human hepatocytes ranging from 10^9 – 4×10^{10} per patient (i.e., 1–10% of the total liver mass) transiently improved brain

function and several blood chemistry parameters, but not survival.^[8] Results from animal studies also suggest that relatively few hepatocytes (between 2% and 10% of the host's liver) are needed to effect a therapeutic benefit. This may be due in part to the fact that the exogenously supplied hepatocytes may aid the regeneration of the native liver. Current consensus is that the minimum cell mass necessary to support a human undergoing acute liver failure is about 5–10% of the total liver weight, or about 10^{10} cells.

Cell Source and Replication of Hepatocytes in Culture

Primary (i.e., directly isolated from an organ) human hepatocytes are the natural choice for the BAL. However, they are scarce due to the competing demand for livers for OLT. Human hepatocyte lines have been generated by the spontaneous transformation of long-term cultures of adult hepatocytes, as well as by a reversible transformation strategy based on Cre-Lox recombination of the SV40T viral antigen.^[16] No one has yet achieved the goal of generating a safe and fully-functional yet clonal, immortalized, or genetically-engineered human cell that can be substituted for primary hepatocytes. Besides providing a sufficient number of cells for BALs, proliferating cells would have the advantage of being able to be genetically engineered to overexpress metabolically significant genes (e.g., glutamine synthetase, cytochrome P450s, etc.). Currently, the only way to permanently integrate transgenes into the cellular genome is to infect replicating cells with retroviruses.

Recent data suggest that there are stem cells in the liver that can differentiate into fully mature hepatocytes.^[17] Furthermore, studies in rats, mice, and humans have shown that bone marrow is a major extrahepatic source of stem cells that may contribute to liver regeneration after hepatic injury. Pending more research on the mechanisms of stem cell differentiation into hepatocytes, stem cells remain a potentially significant source of liver cells.

In BAL devices tested so far in the clinic, the only human cell used has been the hepatoma C3A line,^[15,18] which originated from a tumor and may have significantly lower detoxification activities than primary hepatocytes. Most current BAL concepts are based on the use of pig hepatocytes. These present no risk of transmitting malignancies to the patient, but pose the risks of hyperacute rejection,^[5] transmission of zoonoses,^[6] and potential mismatch with human liver functions. Specialized breeding of transgenic animals could address the first two concerns. Very little is known about the metabolic compatibility of liver functions across species. Xenotransplantation studies



may help provide useful information in that respect, once there is reliable success in overcoming hyperacute rejection and early death after clinical xenotransplantation.

Development and Optimization of Hepatocyte and Liver Preservation Techniques

The development of optimal preservation protocols for hepatocytes that enable the storage and ready availability of cells for BALs, has been the subject of several studies. Hepatocytes have been cryopreserved shortly after isolation as well as after culture for several days. Compared to isolated cells, cultured hepatocytes exhibit greater resistance to high concentrations of the cryoprotective agent dimethyl sulfoxide, as evidenced by preservation of cell viability, cytoskeleton, and function. Based on experimental and theoretical studies, cooling rates between 5 and 10°C/min caused no significant decrease in albumin secretion rate compared to control, unfrozen, cultures.^[19,20]

There have been attempts to store hepatocyte cultures in various solutions used for cold storage of whole donor livers. One study showed that cultured hepatocytes maintained at 4°C lose significant viability after a few hours of cold storage, but that addition of polyethylene glycol significantly extends functionality and survival.^[21] Interestingly, the use of the University of Wisconsin (UW) solution, currently the most widely-used solution for cold organ storage, has not performed better than leaving the cells in standard hepatocyte culture medium. It is conceivable that the UW solution mediates its effect by prolonging the survival of non-parenchymal cells. It is hoped that further improvements in preservation solutions will enable the storage of BAL systems, as well as lengthen the useful cold storage time of whole livers for transplantation.

ENGINEERING A BIOARTIFICIAL LIVER FUNCTIONAL UNIT

Hepatocyte Function and the Environment

The basic functional unit in the liver is the hepatic lobule (Fig. 2), where hepatocytes are organized into a radial network of plates that are one cell thick. On each side of the plate is a thin layer of extracellular matrix (ECM) material that is covered by vascular endothelial cells. Other cell types, such as fat-storing cells and macrophages are dispersed throughout the lobule. Epithelial cells form bile ductules that collect the bile secreted by the hepatocytes. In the lobule, blood flows from the periportal outer region toward the central hepatic vein through specialized porous capillaries

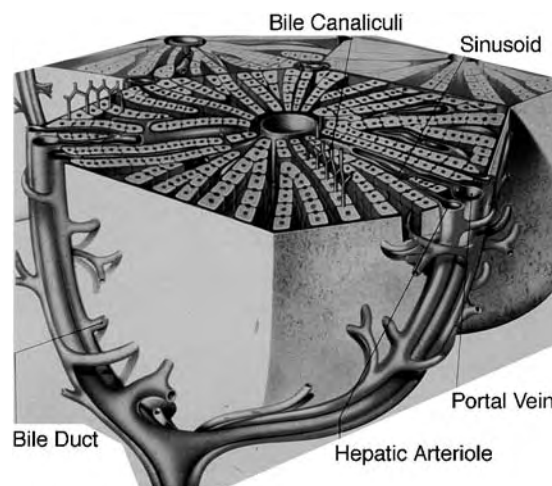


Fig. 2 Structure of the liver lobule. (Reproduced from Ref. [11] by permission of the publisher.)

called sinusoids. Hepatocytes in the periportal, intermediate or centrilobular, and perivenous zones exhibit different morphological and functional characteristics. For example, urea synthesis is a process with high capacity to metabolize ammonia and low affinity for the substrate, which occurs in the periportal and intermediate zones. In the perivenous zone, ammonia is removed by glutamine synthesis, a high-affinity but low-capacity process that removes traces of ammonia that cannot be metabolized by the urea cycle. Replicating the functional heterogeneity of hepatocytes in the lobule may be important to maximize the performance of the BAL.

The maintenance of functional heterogeneity in the liver is dependent on several factors, including gradients of hormones, substrates, oxygen, and ECM composition.^[22] In one study where hepatocytes were chronically exposed to increasing oxygen tensions within the physiological range of about 5 mm Hg (perivenous) to 85 mm Hg (periportal), urea synthesis increased about tenfold, whereas cytochrome P450-dependent detoxifying activity decreased slightly and albumin secretion was unchanged.^[23] These data suggest that by creating environmental conditions that emulate certain parts of the liver sinusoid, it is possible to modulate hepatocyte metabolism in a way that is consistent with in-vivo behavior, and to purposely upregulate or downregulate specific liver metabolic functions.

Biomaterials in Liver Tissue Engineering

Materials used in tissue engineering

Most of the natural ECM materials used for liver cell culture were obtained via solubilization of tissues by chemical processing. These protein solutions can often

be reconstituted into three-dimensional gels of any shape or form, and retain many chemical features of the ECM proteins, including bound growth factors, found *in vivo*. Commonly used reconstituted matrices include type I collagen, which is isolated from rat tail or bovine skin by mild acid treatment. The acid solution of collagen can be induced to form a gel upon restoring a physiological pH of 7.4, which causes the polymerization of collagen molecules into a large network of fibrils. The extent of cross-linking in this collagen is very low in comparison with that of the native tissue, but chemical cross-linking can be induced by either glutaraldehyde or dehydrothermal (vacuum and $\sim 100^\circ\text{C}$) treatment.

Naturally derived matrices provide good substrates for cell adhesion because cells express the adhesion receptors that specifically recognize and bind to ECM molecules that make up these matrices. Nevertheless, there have been considerable advances in the development of synthetic biocompatible polymers, which theoretically have an unparalleled range of physical and chemical properties. In practice, however, most studies have used relatively few synthetic materials, in part due to a reluctance to expend time and money to secure regulatory approval for clinical use of untested biomaterials.

Optimization of surface chemistry

Cells do not usually directly attach to artificial substrates, but rather to ECM proteins that are physically adsorbed (i.e., by virtue of hydrophobic and electrostatic interactions) or chemically (i.e., via covalent bonds) attached to the surface. Surface treatment, such as with ionized gas, is often used to favor protein adsorption. This process only modifies the surface of the material, and thus minimally affects its bulk mechanical properties. Surfaces have been coated with positively charged materials such as poly-L-lysine to promote cell adhesion by attractive electrostatic interactions with the cells. This is because the latter usually display a negative surface charge due to the presence of negative sialic acid residues on their surface glycocalyx.

Because physisorption is notoriously nonselective, covalent modification of substrates or chemisorption is used if necessary to provide more control over the type, density, and distribution of adhesive protein on the surface of the material. The first step involves using a reactive chemical that bonds to the surface and has a free functional group that easily reacts with free thiol, hydroxyl, carboxyl, or amine groups on proteins. This step often requires harsh chemical conditions, whereas the second step (which involves conjugation of the protein) can be accomplished under physiological conditions. This approach is also suitable to graft small

adhesive peptides (e.g., arginine-glycine-aspartic acid (RGD)-peptide) that otherwise would not stably bind to surfaces by physisorption.

Various ECM proteins, such as collagen and laminin, regulate the activities of transcription factors that control the expression of liver-specific genes such as albumin. It has been hypothesized that these factors bind ECM-responsive elements. It is important to realize that no matter what type of surface is used to seed the cells, over a period of several days of culture, cells will have secreted significant quantities of their own ECM onto the substrate, and the initial surface properties of the material will be dramatically altered. For example, cultured hepatocytes continually synthesize collagen, fibronectin, and laminin,^[24] and this secretion is important for the maintenance of long-term liver-specific protein secretion in hepatocytes cultured in collagen gels.

More chemically complex ECM, such as matrix extracted from Engelbroth-Holmes sarcoma (EHS) tumors grown in mice, whose composition more closely resembles basal lamina (it consists predominantly of type IV collagen, laminin, and heparan sulfate proteoglycan), induce the expression of liver-specific functions and phenotypic features typical of hepatocytes *in vivo*.^[25,26] There is evidence that the mechanochemical properties of such surfaces, whether they favor cell spreading, rounding, or aggregation, impact the responsiveness of the hepatocytes to growth and other trophic factors.^[27]

Extracellular matrix and cell polarity

Extracellular matrix geometry (or topology) affects hepatocyte morphology and function. The effect of matrix topology was investigated in a controlled fashion using a culture technique whereby cells cultured on a single surface were overlaid with a second layer of ECM, thereby creating a sandwich configuration around the cells. In the case of hepatocytes, the overlay establishes an ECM configuration that resembles that found in the liver (i.e., where hepatocytes are generally bounded by ECM at each of their opposite basolateral membrane domains). The result is that hepatocytes remain as a monolayer, but exhibit a dramatic increase in the expression of liver-specific functions compared to cultures on a single ECM-coated surface.^[24]

Rat hepatocytes cultured in the collagen sandwich, when compared to a single collagen gel substrate, have a dramatically altered organization and expression of cytoskeletal proteins. The collagen sandwich induces the formation of distinct apical and basolateral membrane domains, each expressing specific surface markers^[24] (Fig. 3). Cytoskeletal actin filaments are



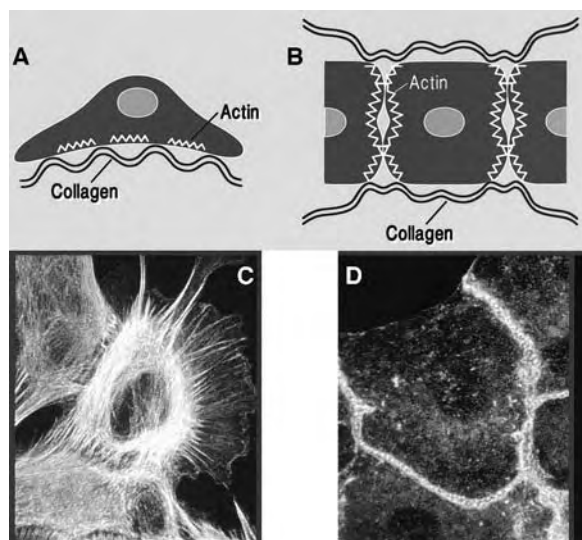


Fig. 3 Morphology and localization of actin fibers in hepatocyte cultures on a single collagen gel (A and C) and in a collagen sandwich (B and D). A and B show the distribution of apical and basolateral membranes in relation to the localization of actin. C and D show actual immunofluorescence staining of actin filaments in rat hepatocytes. (C and D reproduced from Ezzell, R.M., et al., *Exp Cell Res* **1993**, *208*, 442–452, by permission of the publisher.)

concentrated under the plasma membrane in regions of contact with neighboring cells, forming the sheathing of a functional bile canalicular network. In contrast, hepatocytes cultured on a single gel exhibit actin-containing stress fibers on the ventral surface in contact with the substrate and no bile canaliculi, a pattern more typical of fibroblasts.

Controlling cell–cell interactions

In general, when cells are seeded on a substrate, they initially form a monolayer that is held together by both cell–substrate and cell–cell adhesions. If the substrate is only moderately adhesive, cell–cell adhesions may be stronger and the cells may spontaneously reorganize into compact aggregates that interact minimally with the surface. When hepatocytes are induced to form aggregates, they maintain viability and stable expression of liver-specific functions for several weeks. Furthermore, recent data show that hepatocytes in spheroids form functional bile canaliculi, which indicates that the cells exhibit distinct apical and basolateral membrane domains.^[28] Hepatocyte spheroids have been used as the basis for several BALs.^[29]

Approximately 20 years ago, it was discovered that hepatocytes could be cultured on feeder (or supportive) cells to maintain their viability and function.^[30] More recent studies showed that nonhepatic cells, even from

other species, may be used. In these culture systems, cell–cell interactions among hepatocytes and cells of another type (rat liver epithelial cells, liver sinusoidal endothelial cells, or mouse embryonic fibroblasts), or heterotypic interactions, are critical for the expression of hepatocellular functions. The disadvantages of coculture systems include the potential variability in the cell line used and the additional work needed to propagate that cell line and attend to the isolation of hepatocytes.

It may be desirable to optimize heterotypic cell–cell interactions in order to maximize the expression of liver-specific functions of the cocultures. Keeping in mind that cells cultured on surfaces do not usually layer onto each other (except for malignant cancer cell lines), random seeding using a low ratio of parenchymal cells to feeder cells will achieve this goal, but at the expense of using much of the available surface for fibroblasts, which do not provide the desired metabolic activity. On the other hand, micropatterning techniques enable the optimization of the seeding pattern of both cell types to ensure that each hepatocyte is near a feeder cell while minimizing the number of feeder cells.^[31] As a result, metabolic function per area of culture is increased and the ultimate size of a BAL with the required functional capacity is reduced. In prior studies using circular micropatterns, function per hepatocyte increased when the hepatocyte circle diameter decreased, and function per unit area of culture increased when the space occupied by fibroblasts in between the hepatocyte islands decreased (for a constant cell number ratio of the two cell types).

Various methods for patterning the deposition of ECM or other cell attachment factors onto surfaces have been developed.^[32] Photolithography involves spin-coating a surface (typically silicon or glass) with a layer approximately 1 μm thick of photoresist (a photo sensitive chemical resistant to acid), exposing the coated material to ultraviolet light through a mask that contains the pattern of interest, and treating the surface with a developer solution that dissolves only the exposed regions of photoresist (Fig. 4). This process leaves photoresist in previously unexposed areas of the substrate. The exposed areas of substrate can be chemically modified for attaching proteins, etc., or can be treated with hydrofluoric acid to etch the material. The etching time controls the depth of the channels created. Subsequently, the leftover photoresist is removed using an appropriate solvent, leaving a surface patterned with different molecules or grooves. A disadvantage of this method is that it uses chemicals that are toxic to cells, and generally harsh conditions that may denature proteins.

The etched surfaces produced by photolithography can be used to micromold various shapes in a polymer called poly(dimethylsiloxane) (PDMS). The PDMS cast faithfully reproduces the shape of the silicon or

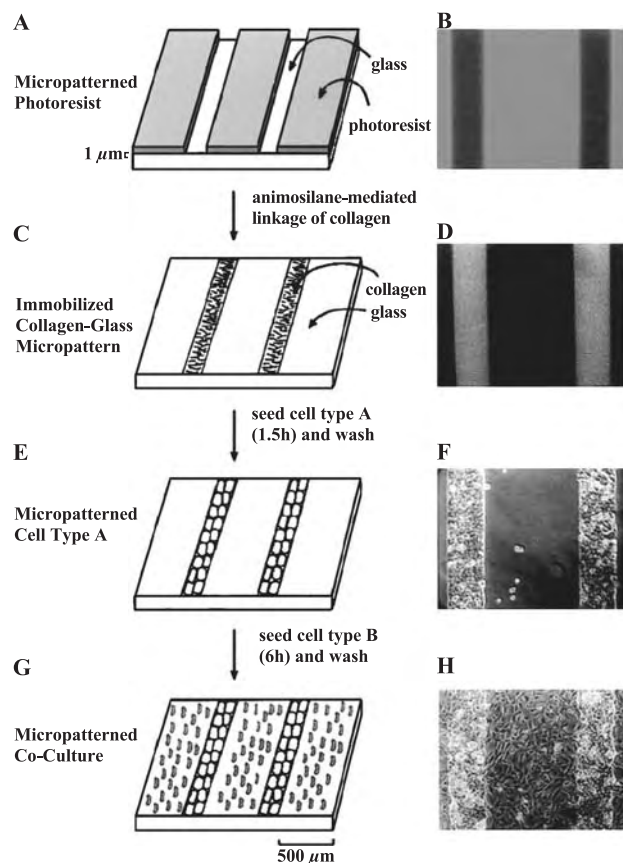


Fig. 4 Photolithographic technique for patterning cells on a glass surface. (Reproduced from Ref. [28] by permission of the publisher.)

glass mold to the μm scale, and can be used in various soft lithography techniques, including microstamping, microfluidic patterning, and stencil patterning. An infinite number of identical PDMS casts can be generated from a single master mold, which makes the technique very inexpensive. Soft lithography methods can be used on virtually any type of surface, including curved surfaces, owing to the flexibility of PDMS. In using these approaches, it is important that the base material be resistant to physisorption, or the selectivity of the adhesive groups may be significantly reduced *in vivo*. A successful approach to prevent adhesion to the base material is via covalent attachment of anti-adhesive factors on the remaining functional groups.

Porous three-dimensional matrices

Porous matrices with pore sizes in the range of 30–300 μm have been used to culture hepatocytes, and have been prepared by salt-leaching or freeze-drying techniques. The first method involves adding water-soluble crystals (e.g., NaCl) of size range similar to the desired pores in the melted base polymer

material. After solidification of the polymer, the salt crystals in the resulting solid are dissolved by exposure to aqueous solutions, leaving a pore in the place of every crystal. An alternative approach is to use supercritical carbon dioxide to create pores by the induction of microbubble formation within the polymer. The freeze-drying technique is based on the general principle that when freezing a solution, the solvent forms pure solid crystals while all solute materials are concentrated in the remaining unfrozen fraction. During the subsequent drying process, the solid crystals evaporate and leave pores. The morphology of the solid crystals is dependent on the physicochemical properties of the solution, the temperature gradient at the liquid–solid interface, and the velocity of that interface. Furthermore, the solid crystals tend to orient in the direction of the temperature gradient, so that the direction of the pores can be controlled as well. For practical applications, however, it is more typical to freeze solutions containing biomaterials in a bulk fashion. It is noteworthy that the rate of freezing and the temperature gradient are difficult to control and to maintain constant throughout the freezing process. Thus, porous materials made by this technique typically exhibit nonuniform pore sizes as one moves from the surface to the center.

Per unit volume, such porous materials have a large surface area; at least theoretically they can support the attachment of a large number of cells. This is in practice not the case because of the lack of efficient oxygen transport in porous matrices that lack a functional vascular system or other means to provide convective (i.e., flow-driven) delivery of nutrients. Pore size may affect whether the cells spread onto the substrate or invade the pores, in which case they tend to form three-dimensional aggregates. Such aggregates have been shown to express high levels of liver-specific function, presumably due to the formation of specialized cell–cell contacts among the hepatocytes.^[33]

Soluble Factors

Hepatocyte culture media

Certain hormonal additives have been found to be essential for the maintenance of long-term protein secretion and viability in cultured hepatocytes. For example, supplementation with corticosteroids such as hydrocortisone and dexamethasone is used to promote the expression of liver-specific functions in various hepatocyte culture configurations. Serum-free cultures do not benefit from the contribution of hormones and other metabolites found in the serum, and have more requirements in this respect. For example,

serum-free cultures of hepatocytes require exogenous insulin and proline.^[24] Insulin is an anabolic hormone, and proline is necessary for endogenous collagen synthesis. It is noteworthy that the expression of liver-specific genes in cultured liver slices is only slightly better in hormonally supplemented medium as opposed to basal medium. Further refinements in the composition of the ECM may thus relax some of the requirements for hormonal and other supplements in the culture medium.

Effect of plasma and blood

Rat hepatocytes that are seeded and maintained in standard hepatocyte culture medium and then exposed to either rat or human plasma become severely fatty within 24 hours, with a concomitant reduction in liver-specific functions. Plasma thus appears to be a rather inhospitable environment for hepatocytes, although it is clear that hepatocytes must be made to tolerate plasma for the concept of bioartificial liver to become reality. Supplementation of human anticoagulated plasma with hormones and amino acids (to bring those metabolites to levels similar to those found in standard hepatocyte culture medium) eliminates intracellular lipid accumulation and restores albumin and urea synthesis as well as P450-dependent detoxification.^[34,35] However, direct supplementation of plasma, especially with respect to the high levels of hormones used, is high in cost and potential health risk to the patient.

Hepatocytes are cultured in an artificial medium for several days prior to their use in a BAL, where they will contact plasma or blood from the patient. Culture media containing supraphysiological levels of hormones—and more specifically, insulin (used at about 10^4 times physiological concentration in standard hepatocyte culture medium)—predispose the hepatocytes to an abnormal response when subsequently exposed to plasma. In a recent study in our laboratory, we found that preconditioning hepatocytes in low insulin-containing medium reduced intracellular lipid accumulation during subsequent plasma exposure.^[36] Direct amino acid supplementation of plasma increased both urea and albumin secretion rates by the hepatocytes. We thus concluded that a combination of both preconditioning and plasma supplementation can be used to upregulate liver-specific functions of hepatocytes during plasma exposure.

BIOREACTORS

Transport Issues in Scale-up of Hepatocyte Cultures

In normal liver, no hepatocyte is farther than a few micrometers from circulating blood; thus, transport by

diffusion has to occur only over very short distances. Although oxygen diffusivity is an order of magnitude greater than that of many other small metabolites (e.g., glucose and amino acids), it has very low solubility in physiological fluids deprived of oxygen carriers. Thus, it is not possible to create large concentrations of gradients, which would provide the driving force for rapid oxygen transport over long distances. This, in addition to the fact that hepatocytes have a relatively high oxygen uptake rate, makes oxygen transport the most constraining parameter in the design of BAL devices.

Oxygen transport and uptake of hepatocytes have been extensively studied in the sandwich culture configuration in order to obtain the essential oxygen uptake parameters needed for the design of bioreactor configurations.^[37] The maximum oxygen uptake rate of cultured rat hepatocytes was measured to be about 13.5 pmol/s/ μg DNA, which is fairly stable after the first day in culture and continuing for up to 2 weeks. Interestingly, the oxygen uptake was about twice on the first day after cell seeding, presumably because of the increased energy required for cell attachment and spreading. Perhaps this should be taken into account when seeding hepatocytes into a BAL. Oxygen uptake was not sensitive to oxygen tension in the vicinity of the hepatocytes up to a lower limit of about 0.5 mm Hg. Below this, oxygen uptake decreased, suggesting that it becomes a limiting substrate for intracellular hepatocyte metabolism. Because oxygen is essential for hepatic adenosine triphosphate (ATP) synthesis, a reasonable design criterion would specify that oxygen tension remain above approximately 0.5 mm Hg.

Based on these parameters, it is possible to estimate oxygen concentration profiles in various bioreactor configurations based on a simple diffusion-reaction model that assumes that the process follows Michaelis–Menten kinetics. One can generally estimate that the maximum thickness of a static layer of aqueous medium on the surface of a confluent single hepatocyte layer is about 400 μm .^[38] Calculations of oxygen transport through hepatocyte aggregates suggest that even a relatively low density of cells (10^7 cells/ cm^3) cannot have a thickness exceeding about 300–500 μm . At cell densities of 10^8 cells/ cm^3 , which is similar to those found in normal liver, that thickness is only 100–200 μm .

Hollow-Fiber Systems

The hollow-fiber system has been the most widely used type of bioreactor in BAL development.^[39,40] The hollow-fiber cartridge consists of a shell traversed by a large number of small-diameter tubes. The cells may be placed within the fibers in the intracapillary space or

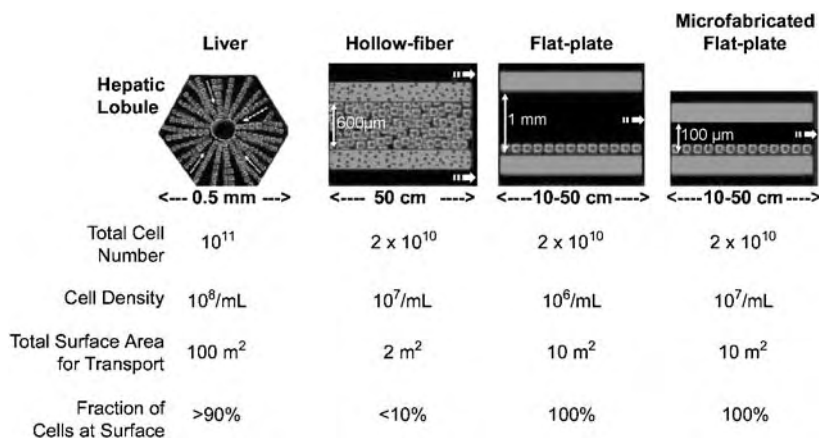


Fig. 5 Comparison among popular hepatocyte bioreactor configurations.

on the shell side in the extracapillary space. The compartment that does not contain the cells is generally perfused with culture medium or with the patient's plasma or blood. The fiber walls may provide the attaching surface for the cells or act as a barrier against the immune system of the host. Microcarriers have also been used as a means to establish an attachment surface for anchorage-dependent cells introduced in the shell side of hollow-fiber devices. Many studies have been conducted to determine optimal fiber dimensions, spacing, and reactor length based on oxygen transport considerations.^[41]

One difficulty with the hollow-fiber configuration is that interfiber distances (and consequently transport properties within the shell space) are not well controlled. Thus, it may be advantageous to place cells in the lumen of small fibers because the diffusional distance between the shell (where the nutrient supply would be) and the cells is essentially equal to the fiber diameter. In one configuration, hepatocytes have been suspended in a collagen solution and injected into the lumen of fibers where the collagen is allowed to gel. Contraction of the collagen lattice by the cells even creates a void in the intraluminal space, which can be perfused with hormonal supplements, etc., to enhance the cells' viability and function while the patient's plasma flows on the shell side. Because of the relatively large diameter of the fibers used as well as transport limitations associated with the fiber wall, these systems have been prone to substrate transport limitations. To improve oxygen delivery, additional fibers that carry oxygen straight into the device have been used.^[29,42]

Parallel Plate Systems

An alternative bioreactor configuration is based on a flat surface geometry^[37] where it is easier to control internal flow distribution and ensure that all cells are adequately perfused. The main drawback to this

configuration is that it is difficult to build a system that contains sufficient cell concentration (Fig. 5). For example, a 1 mm channel height would result in a 10-liter reactor to support 20×10^9 hepatocytes cultured on a surface 10 m^2 in area. For a liver failure patient who is probably hemodynamically unstable, it is generally accepted that the priming volume of the system not exceed 1 liter.

The volume of the device in the flat-plate geometry can be decreased by reducing the channel height (Fig. 5). However, this forces the fluid to move through a smaller gap, which rapidly increases the drag force (shear stress) imparted by the flow to the cells. Recent data suggest that hepatocyte function decreases significantly at shear stresses $>5 \text{ dy/cm}^2$.^[37] To reduce the deleterious effects of high shear, it may be possible to use grooved surfaces, wherein cells lodge and are less exposed to shear stress, allowing for faster flow without causing cell damage.

In an attempt to provide adequate oxygenation to cells and to protect them from shear in perfused bioreactors, gas-permeable membranes and membranes separating cells from plasma have been incorporated into the flat-plate geometry. A flat-plate microchannel bioreactor wherein cells directly contact the circulating medium has been developed.^[37] The channel is closed by a gas-permeable membrane on one surface, which decouples oxygen transport from the flow rate in the device. Comparing this with a similar flat-plate design wherein a nonpermeable glass surface is substituted to the membrane, internal membrane oxygenation removed the oxygen limitations that occur at low volumetric flow rates.

Monitoring the Performance of Bioartificial Livers

The performance of BALs is often assessed based on clinical parameters such as survival, grade of

encephalopathy, and blood and urine chemistry values. Because these parameters can vary from one patient to another depending on the severity of the disease, it is worthwhile to assess BAL function through other methods. Further insight into the function of the BAL can be obtained by measuring the change in metabolite levels across the BAL, including oxygen, CO₂, ammonia, urea, amino acids, and ketone bodies.

An enhanced perspective of metabolism and cellular function can be obtained with a mathematical modeling framework that considers the stoichiometric constraints of the intracellular reaction network of the cells in the BAL. One such methodology that is especially useful for the analysis of the metabolic function of organs and tissues is metabolic flux analysis. Metabolic flux analysis refers to the calculation of intracellular rates of metabolic reactions using steady-state metabolite balance models applied to measured rates of uptake and release of extracellular metabolites. This approach has been extensively used to study and improve strains of microorganisms (bacteria and yeasts) of significance in biotechnology, and has more recently been applied to mammalian cell systems, including perfused liver.^[43] Metabolic flux analysis, once validated for the particular case under study, is potentially very useful as it is noninvasive and cost effective.

CONCLUSION

The extracorporeal bioartificial liver is a promising technology for the treatment of liver failure, but significant technical challenges remain before systems with sufficient processing capacity and of manageable size can be developed. Most efforts to date have focused on device design and construction, and more recently on the development of methods to generate a continuous supply of human hepatocyte cell lines. New designs are not yet able to stably sustain the large cell populations needed for therapeutic purposes, and concerns about the safety of transformed cell lines remain. On the other hand, there have been fewer efforts to improve the specific functional capacity of the hepatocytes used in bioartificial livers. Increasing the efficiency of the cellular component of bioartificial livers would greatly facilitate their design. The functional capacity of cells used in bioartificial livers may be improved on at least three different levels: 1) by altering the internal machinery of individual cells to upregulate critical functions; 2) by judiciously controlling the spatial distribution of multiple cell populations with different specializations to emulate the organization of the liver acinus; and 3) by optimizing the overall treatment protocol, including the interval, duration, and number of treatment sessions.

ACKNOWLEDGMENTS

This work was supported by grants from the Shriner's Hospitals for Children and the National Institutes of Health.

ARTICLES OF FURTHER INTEREST

Bioreactors; Extracellular Matrix Scaffolds; Microelectromechanical Systems (MEMS) Manufacturing; Tissue Engineering of Liver; Tissue Engineering, Microscale; Tissue Engineering Scaffolds; Xenografts

REFERENCES

1. United Network for Organ Sharing, Richmond, VA. *Policy, Organ Distribution: Allocation of Livers (3.6)*; Accessed November 5, 2002.
2. Potts, J.T.; Herdman, R. *Non-Heart-Beating Organ Transplantation: Medical and Ethical Issues in Procurement*; National Academy Press: Washington, DC, 1997.
3. Safar, P. Clinical death symposium. *Crit. Care Med.* **1988**, *16*, 919–920.
4. Sindhi, R.; Rosendale, J.; Mundy, D.; Taranto, S.; Baliga, P.; Reuben, A.; Rajagopalan, P.R.; Hebra, A.; Tagge, E.; Othersen, H.B., Jr. Impact of segmental grafts on pediatric liver transplantation—A review of the UNOS scientific registry data (1990–1996). *J. Pediatr. Surg.* **1999**, *34*, 107–111.
5. Butler, D. Last chance to stop and think on risks of xenotransplantation. *Nature* **1998**, *391*, 320–324.
6. LeTissier, P.; Stoye, J.P.; Takeuchi, Y.; Patience, C.; Weiss, R.A. Two sets of human-tropic pig retrovirus (letter). *Nature* **1997**, *389*, 681–682.
7. Malhi, H.; Gorla, G.R.; Irani, A.N.; Annamaneni, P.; Gupta, S. Cell transplantation after oxidative hepatic preconditioning with radiation and ischemia-reperfusion leads to extensive liver repopulation. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 13114–13119.
8. Bilir, B.M.; Guinette, D.; Karrer, F.; Kumpe, D.A.; Krysl, J.; Stephens, J.; McGavran, L.; Ostrowska, A.; Durham, J. Hepatocyte transplantation in acute liver failure. *Liver Transplant.* **2000**, *6*, 32–40.
9. Gomez, N.; Ballardur, P.; Calmus, Y.; Baudrimont, M.; Honiger, J.; Delelo, R.; Myara, A.; Crema, E.; Trivin, F.; Capeau, J.; Nordlinger, B. Evidence for survival and metabolic activity of encapsulated xenogeneic hepatocytes transplanted without immunosuppression in Gunn rats. *Transplantation* **1997**, *63*, 1718.
10. Yang, M.B.; Vacanti, J.P.; Ingber, D.E. Hollow fibers for hepatocyte encapsulation and transplantation: Studies of survival and function in rats. *Cell Transplantation.* **1994**, *3*, 373–385.
11. Rosenthal, P.; Roberts, J.P.; Ascher, N.L.; Emond, J.C. Auxiliary liver transplant in fulminant failure. *Pediatrics* **1997**, *100*, E10.

12. Yarmush, M.L.; Dunn, J.C.; Tompkins, R.G. Assessment of artificial liver support technology. *Cell Transplant*. **1992**, *1*, 323–341.
13. Gerlach, J.; Trost, T.; Ryan, C.J.; Meissler, M.; Hole, O.; Muller, C.; Neuhaus, P. Hybrid liver support system in a short term application on hepatectomized pigs. *Int. J. Artif. Organs* **1994**, *17*, 549–553.
14. Rozga, J.; Holzman, M.D.; Ro, M.S.; Griffin, D.W.; Neuzil, D.F.; Giorgio, T.; Moscioni, A.D.; Demetriou, A.A. Development of a hybrid bioartificial liver. *Ann. Surg.* **1993**, *217*, 502–509.
15. Sussman, N.L.; Gislason, G.T.; Conlin, C.A.; Kelly, J.H. The Hepatix extracorporeal liver assist device: Initial clinical experience. *Artif. Organs* **1994**, *18*, 390–396.
16. Kobayashi, N.; Fujiwara, T.; Westerman, K.A.; Inoue, Y.; Sakaguchi, M.; Noguchi, H.; Miyazaka, M.; Cai, J.; Tanaka, N.; Fox, I.J.; Leboulch, P. Prevention of acute liver failure in rats with reversibly immortalized human hepatocytes. *Science* **2000**, *287*, 1258–1262.
17. Susick, R.; Moss, N.; Kubota, H.; Lecluyse, E.; Hamilton, G.; Luntz, T.; Ludlow, J.; Fair, J.; Gerber, D.; Bergstrand, K.; White, J.; Bruce, A.; Drury, O.; Gupta, S.; Reid, L.M. Hepatic progenitors and strategies for liver cell therapies. *Ann. N.Y. Acad. Sci.* **2001**, *944*, 398–419.
18. Ellis, A.J.; Hughes, R.D.; Wendon, J.A.; Dunne, J.; Langley, P.G.; Kelly, J.H.; Gislason, G.T.; Sussman, N.L.; Williams, R. Pilot-controlled trial of the extracorporeal assist device in acute liver failure. *Hepatology* **1996**, *24*, 1446–1451.
19. Karlsson, J.O.M.; Cravalho, E.G.; Borel-Rinkes, I.H.M.; Tompkins, R.G.; Yarmush, M.L.; Toner, M. Long-term functional recovery of hepatocytes after cryopreservation in a three-dimensional culture configuration. *Cell Transplant* **1992**, *1*, 281–292.
20. Borel-Rinkes, I.H.M.; Toner, M.; Sheehan, S.J.; Tompkins, R.G.; Yarmush, M.L. Nucleation and growth of ice crystals inside cultured hepatocytes during freezing in the presence of dimethylsulfoxide. *Biophys. J.* **1993**, *65*, 2524–2536.
21. Stefanovich, P.; Toner, M.; Ezzell, R.M.; Sheehan, S.J.; Tompkins, R.G.; Yarmush, M.L. Effects of hypothermia on the function, membrane integrity, and cytoskeletal structure of hepatocytes. *Cryobiology* **1995**, *23*, 389–403.
22. Reid, L.M.; Fiorino, A.S.; Sigal, S.H.; Brill, S.; Holst, P.A. Extracellular matrix gradients in the space of Disse: Relevance to liver biology. *Hepatology* **1992**, *15*, 1198–1203.
23. Bhatia, S.N.; Toner, M.; Foy, B.D.; Rotem, A.; O'Neil, K.M.; Tompkins, R.G.; Yarmush, M.L. Zonal liver cell heterogeneity: Effects of oxygen on metabolic functions of hepatocytes. *Cell. Eng.* **1996**, *1*, 125–135.
24. Berthiaume, F.; Moghe, P.V.; Toner, M.; Yarmush, M.L. Effect of extracellular matrix topology on cell structure, function, and physiological responsiveness: Hepatocytes cultured in a sandwich configuration. *FASEB J.* **1996**, *10*, 1471–1484.
25. Moghe, P.V.; Berthiaume, F.; Ezzell, R.M.; Toner, M.; Tompkins, R.G.; Yarmush, M.L. Role of extracellular matrix composition and configuration in maintenance of hepatocyte polarity and function. *Biomaterials* **1996**, *17*, 373–385.
26. Jauregui, H.O.; McMillan, P.N.; Driscoll, J.; Naik, S. Attachment and long-term survival of adult rat hepatocytes in primary monolayer cultures: Comparison of different substrata and tissue media formulations. *In Vitro Cell Dev. Biol.* **1986**, *22*, 13–22.
27. Semler, E.J.; Moghe, P.V. Engineering hepatocyte functional fate through growth factor dynamics: The role of cell morphologic priming. *Biotechnol. Bioeng.* **2001**, *75*, 510–520.
28. Abu-Absi, S.F.; friend, J.R.; hansen, L.K.; Hu, W.-S. Structural polarity and functional bile canaliculi in rat hepatocyte spheroids. *Exp. Cell Res.* **2002**, *274*, 56–67.
29. Sauer, I.M.; Obermeyer, N.; Kardassis, D.; Theruvath, T.; Gerlach, J.C. Development of a hybrid liver support system. *Ann. N.Y. Acad. Sci.* **2001**, *944*, 308–319.
30. Guguen-Guillouzo, C.; Clement, B.; Baffet, G.; Bearumont, C.; Morel-Chany, E.; Glaise, D.; Guillouzo, A. Maintenance and reversibility of active albumin secretion by adult rat hepatocytes co-cultured with another liver epithelial cell type. *Exp. Cell Res.* **1983**, *143*, 47–53.
31. Bhatia, S.N.; Balis, U.J.; Yarmush, M.L.; Toner, M. Effect of cell–cell interactions in preservation of cellular phenotype: Co-cultivation of hepatocytes and non-parenchymal cells. *FASEB J.* **1999**, *13*, 1883–1900.
32. Folch, A.; Toner, M. Microengineering of cellular interactions. *Annu. Rev. Biomed. Eng.* **2000**, *2*, 227–256.
33. Ranucci, C.S.; Kumar, A.; Batra, S.P.; Moghe, P.V. Control of hepatocyte function on collagen foams: Sizing matrix pores toward selective induction of 2-D and 3-D cellular morphogenesis. *Biomaterials* **2000**, *21*, 783–793.
34. Washizu, J.; Chan, C.; Berthiaume, F.; Tompkins, R.G.; Toner, M.; Yarmush, M.L. Amino acid supplementation improves cell specific functions of rat hepatocytes exposed to human plasma. *Tissue Eng.* **2000**, *6*, 497–504.
35. Washizu, J.; Berthiaume, F.; Mokuno, Y.; Tompkins, R.G.; Toner, M.; Yarmush, M.L. Long-term maintenance of cytochrome P450 activities by rat hepatocyte/3T3 cell co-cultures in heparinized human plasma. *Tissue Eng.* **2001**, *7*, 691–703.
36. Chan, C.; Berthiaume, F.; Washizu, J.; Toner, M.; Yarmush, M.L. Metabolic pre-conditioning of cultured hepatocytes prior to plasma exposure: Effect of insulin. *Biotechnol. Bioeng.* **2002**, *78*, 753–760.
37. Tilles, A.W.; Berthiaume, F.; Yarmush, M.L.; Toner, M. Critical issues in bioartificial liver development. *Technol. Health Care* **2001**, *9*, 1–10.
38. Yarmush, M.L.; Toner, M.; Dunn, J.C.Y.; Rotem, A.; Hubel, A.; Tompkins, R.G. Hepatic tissue engineering. Development of critical technologies. *Ann. N.Y. Acad. Sci.* **1992**, *665*, 238–252.
39. Tzanakakis, E.S.; Hess, D.J.; Sielaff, T.D.; Hu, W.-S. Extracorporeal tissue engineered liver-assist devices. *Ann. Rev. Biomed. Eng.* **2000**, *2*, 607–632.
40. Allen, J.W.; Hassanein, T.; Bhatia, S.N. Advances in bioartificial liver devices. *Hepatology* **2001**, *34*, 447–455.



41. Catapano, G. Mass transfer limitations to the performance of membrane bioartificial liver support devices. *Int. J. Artif. Organs* **1996**, *19*, 18–35.
42. Flendrig, L.M.; la Soe, J.W.; Jorning, G.G.; Steenbeek, A.; Karlsen, O.T.; Bovee, W.M.; Ladiges, N.C.; te Velde, A.A.; Chamuleau, R.A. In vitro evaluation of a novel bioreactor based on an integral oxygenator and a spirally wound nonwoven polyester matrix for hepatocyte culture as small aggregates. *J. Hepatol.* **1997**, *26*, 1379–1392.
43. Lee, K.; Berthiaume, F.; Stephanopoulos, G.N.; Yarmush, M.L. Metabolic flux analysis: A powerful tool for monitoring tissue function. *Tissue Eng.* **1999**, *5*, 347–368.

Lung, Artificial: Basic Principles and Current Applications

William J. Federspiel

*Department of Chemical Engineering, University of Pittsburgh,
Department of Bioengineering, University of Pittsburgh, and McGowan Institute,
University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.*

Kristie A. Henchir

*Department of Bioengineering, University of Pittsburgh, and McGowan Institute,
University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.*



INTRODUCTION

Artificial lungs currently used or in development are medical devices designed to take over or supplement the respiratory function of the lung, which is to oxygenate the blood and remove carbon dioxide. Current artificial lungs are also known as blood oxygenators and are simple passive modules composed of bundles of hollow fiber membranes through which blood is pumped. Blood oxygenators are used routinely in cardiopulmonary bypass for open-heart surgical procedures. Much less prevalent is the use of artificial lungs in treating respiratory insufficiencies or failures, in applications generally known as Extracorporeal Membrane Oxygenation (ECMO) or Extracorporeal Life Support (ECLS). However, the growing incidence of lung disease associated with our aging population, along with advances in biomaterials, has spurred significant recent development work toward next-generation artificial lungs that could be used to successfully treat patients with a variety of respiratory failures.

Acute and chronic diseases of the lung remain major healthcare problems. Each year nearly 350,000 Americans die of some form of lung disease.^[1] Overall, lung disease is America's number-three killer and is responsible for one in seven deaths. Adult respiratory distress syndrome (ARDS), only one form of acute lung failure, afflicts approximately 150,000 patients every year in the United States,^[2] and despite advances in critical care medicine, ARDS mortality remains between 40% and 50%.^[3,4] Most lung disease is chronic, and an estimated 30 million Americans are now living with chronic lung disease. Chronic obstructive pulmonary disease (COPD) includes emphysema and chronic bronchitis and afflicts over 16 million adults in the United States annually.^[5] COPD is the fourth leading cause of death and the second most common disability. Each year nearly 400,000 COPD patients will be hospitalized with acute exacerbations of their chronic condition, a rate that has risen nearly threefold in the last decade.^[6] Respiratory support for these patients is most often

maintained using mechanical ventilation, and although adequate gas exchange may be achieved, ventilatory support can lead to barotrauma, volutrauma, and other iatrogenic injuries, further exacerbating the acute respiratory insufficiency in many patients.^[7,8] Unlike mechanical ventilation, artificial lungs provide respiratory support independent of the lungs and allow a reversibly injured lung to rest and heal, thus offering the promise of improved treatment for acute respiratory insufficiencies or an effective bridge-to-lung transplant for chronic respiratory disease. Artificial lungs for longer-term respiratory support are likely far away, awaiting significant improvements in biomaterials and more efficient gas exchange strategies incorporated into their designs. Currently, lung transplantation remains the best treatment for chronic lung failures.

BACKGROUND

The natural lung represents a remarkable organ for gas exchange, and developing an artificial lung that approaches the gas exchange prowess of the natural lung is a significant engineering challenge. The alveoli of the lung, the tiny gas sacs at the termini of all the branching airways of the lung, offer intimate contact between inspired gas and blood flowing through capillaries in the lung (Fig. 1). The O₂ and CO₂ diffusing capacities of the lungs are proportional to the gas exchange area of the alveolar-capillary membrane and to the inverse of the diffusion distance across the alveolar-capillary membrane into blood.^[9] The substantial gas exchange capacity of the lung stems from an alveolar-capillary area comparable to a tennis court surface, 100–150 m², packaged compactly with a high surface area to blood volume ratio of approximately 300 cm⁻¹^[10] and a diffusion distance between gas and blood phases of no more than about 1–2 μm.^[11] The natural lung can provide gas exchange ranging from resting levels for both O₂ and CO₂ (about 200–250 ml/min in average adults) to 10–20 times that under exercise conditions,^[12] and it does so

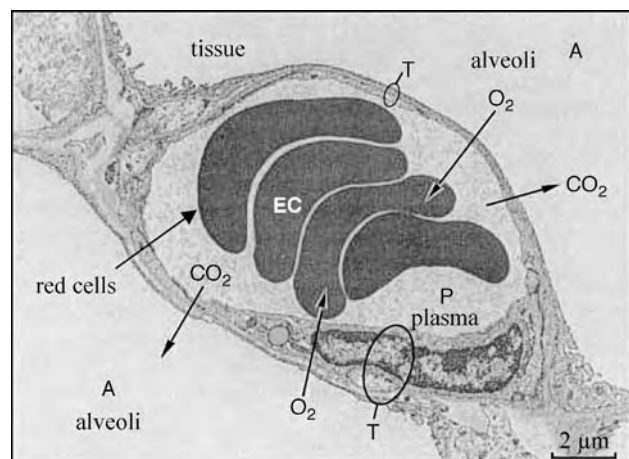


Fig. 1 The alveolar-capillary membrane in the natural lung showing intimate contact between inspired gas and blood in the lung. (Reprinted with written permission from Ref. [11], page 342, by Ewald R. Weibel/Harvard University Press.)

using room air as its oxygen supply gas. In contrast, current hollow fiber blood oxygenators, as used in cardiopulmonary bypass, have membrane areas ranging from 1 to 4 m² that are packaged much less compactly than in the natural lung, with a surface area to blood volume ratio 10 times less than in the natural lung (Fig. 2). The effective distance that gas diffuses between blood and gas flow pathways in artificial lungs is approximately 10–30 μm, an order of magnitude greater than in the natural lung.^[13] Thus, even with using 100% oxygen gas, artificial lungs currently used or under development aim at gas exchange levels that can support *resting* metabolic needs in patients. Artificial lungs that will allow patients any significant level of increased metabolic activity are not on the immediate horizon.

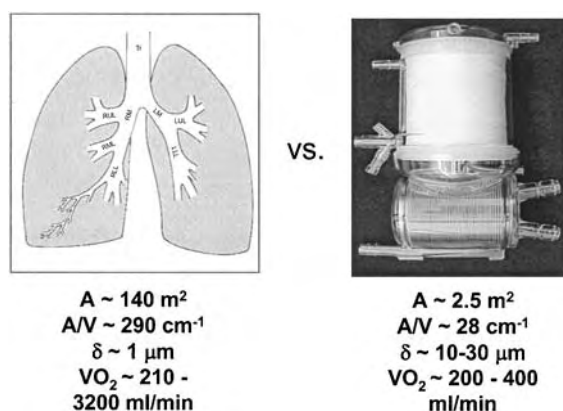


Fig. 2 Comparison of gas exchange parameters of the natural lung and current artificial lungs or blood oxygenators.

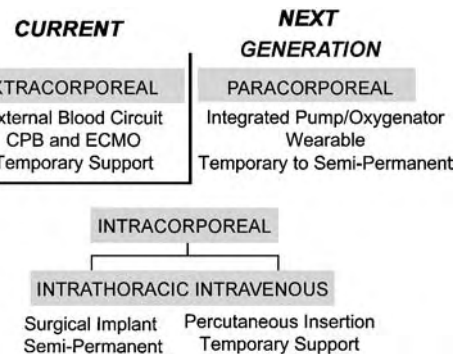


Fig. 3 Classification of artificial lung technology.

This article reviews artificial lung technology in terms of basic operational principles and present-day applications of standard blood oxygenators in cardiopulmonary bypass and ECMO/ECLS. Artificial lungs can be generally classified as extracorporeal, paracorporeal, intravascular, or intrathoracic (Fig. 3). The blood oxygenators used in cardiopulmonary bypass and ECMO/ECLS are extracorporeal artificial lungs incorporated into pump-containing external blood circuits. The companion article to this one (see Lung, Artificial: Current Research and Future Directions) will review research and development work on next-generation paracorporeal, intravascular, and intrathoracic artificial lungs. This review will not cover the interesting and significant work on past artificial lungs, devices which are no longer in use or in development, such as bubble oxygenators, disc-type oxygenators, or even the early work based on hollow fiber membranes. For information pertaining to these earlier efforts in artificial lung development, the interested reader is referred to several previous comprehensive review articles.^[13–15]

BASIC PRINCIPLES OF OPERATION

Hollow fiber membranes form the basic gas exchange units of contemporary artificial lungs and are small polymer tubes with microporous walls of 20 to 50 μm thickness and with outer diameters from 200 to 400 μm (Fig. 4). The wall pores have characteristic sizes typically below about 0.1 μm, and the porosity or volume fraction of the fiber wall can vary from about 30% to 50%.^[16] Hollow fiber membranes for artificial lungs are made from hydrophobic polymers (often polypropylene), so that the membrane wall pores remain gas-filled and respiratory gases can diffuse readily across it. In most artificial lung applications, an oxygen (O₂) “sweep gas” flows through the inside lumens of the hollow fibers, while blood flows outside the hollow fibers through the interstitial spaces in the hollow fiber bundle. Oxygen diffuses down its concentration gradient across

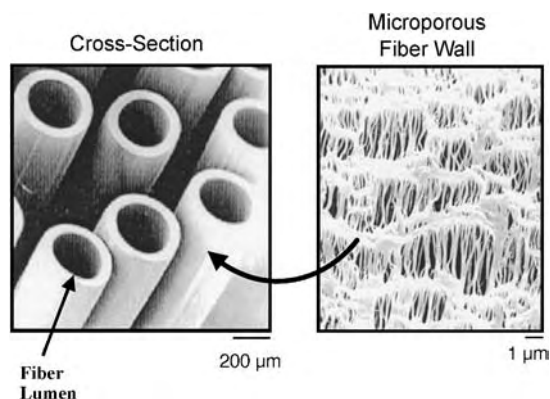


Fig. 4 Scanning electron micrograph of microporous hollow fiber membranes used in artificial lungs. The walls of the fibers (right) contain submicron pores where respiratory gases diffuse.

the fiber wall into blood, while carbon dioxide (CO_2) diffuses down its concentration gradient from the blood into the sweep gas flowing through the fibers and is removed when the sweep gas exits the device.

Determinants of Gas Exchange

The gas exchange permeance, K , of an artificial lung represents an overall mass transfer coefficient for either O_2 or CO_2 exchange.^[9] The overall O_2 exchange rate, \dot{V}_{O_2} , is related to the O_2 permeance according to

$$\dot{V}_{\text{O}_2} = K_{\text{O}_2} A (P_{\text{O}_{2g}} - P_{\text{O}_{2b}}) \quad (1)$$

where $P_{\text{O}_{2g}}$ and $P_{\text{O}_{2b}}$ are the average O_2 partial pressures in the sweep gas and blood phases, respectively, flowing through the artificial lung, and A is the total membrane area of the hollow fiber bundle. The gas exchange permeance for CO_2 removal can be related to the CO_2 exchange rate similarly using

$$\dot{V}_{\text{CO}_2} = K_{\text{CO}_2} A (P_{\text{CO}_{2b}} - P_{\text{CO}_{2g}}) \quad (2)$$

with the driving force for CO_2 exchange being the P_{CO_2} difference between blood and sweep gas. In artificial lungs the overall permeances for O_2 and CO_2 gas exchange are dictated by the diffusional resistances encountered as these gases diffuse between the sweep gas and blood flow pathways. Permeance is the inverse of a diffusional resistance, and the overall transfer resistance, in an artificial lung device has two principal components:

$$\frac{1}{K} = \frac{1}{K_m} + \frac{1}{K_b} \quad (3)$$

where K_m and K_b are the membrane and blood-side permeances for each gas (O_2 and CO_2). The term $1/K_m$

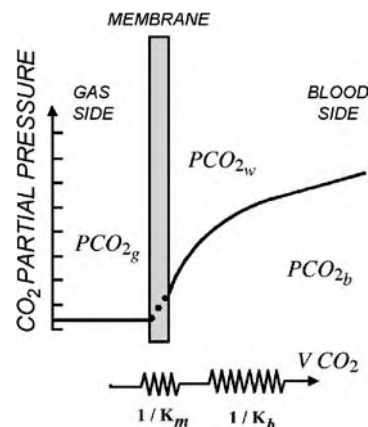


Fig. 5 Illustration of principal determinants of gas exchange in artificial lungs. $P_{\text{CO}_{2w}}$ represents the partial pressure of CO_2 at the membrane wall.

represents a diffusional resistance for the membrane itself, while $1/K_b$ represents a resistance for gas diffusing between the membrane and the flowing blood stream.

Figure 5 illustrates the membrane and blood-side diffusional resistances to gas exchange in artificial lungs by showing the general gradient in CO_2 partial pressure from the sweep gas to the blood pathway. Transfer resistance within the sweep gas pathway is negligible. Most of the diffusional resistance resides within a blood-side diffusional boundary layer, and secondarily within the membrane itself. The blood-side and membrane permeances dictate overall gas exchange in artificial lungs and represent serial transport processes whose resistances add directly to determine overall resistance, as in Eq. 3. As serial “resistors” the smallest permeance or largest resistance controls overall gas exchange in an artificial lung.

Membrane Permeance

Most artificial lungs, including standard blood oxygenators, use microporous hollow fiber membranes. Microporous hollow fibers have fixed submicron pores within the wall that are contiguous from outer to inner lumen, and gas exchange occurs by diffusion through these gas-filled pores. The polymer used does not dictate gas exchange through the membrane as much as the pore characteristics and the fiber wall porosity. In artificial lungs the hydrophobic nature of the polymers (e.g., polypropylene) used to make the fiber membranes prevents intrusion of blood plasma into the fiber pores under normal conditions. Most microporous hollow fiber membranes for artificial lungs are manufactured by Celgard (Charlotte, NC), Membrana (Germany), and Mitsubishi Rayon (Japan). Table 1 summarizes the properties of several commercially

Table 1 Properties of commercially available hollow fiber membranes

Fiber	OD/ID (microns)	Wall thickness (microns)	Pore size (width × length) (microns)	Porosity
Celgard x30–240	300/244	28	0.03 × 0.2	40%
Mitsubishi KPF-190	245/200	22	0.04 (width)	40–55%
Membrana PP 50/280	380/280	50	0.2	N/A
DIC	255/205	25	Nonporous coated	Nonporous coated

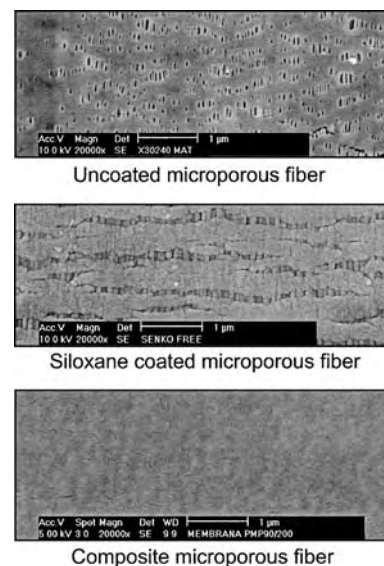
available microporous hollow fibers commonly used in artificial lung devices.

The membrane diffusional resistance of a microporous hollow fiber depends on the permeance, K_m , of the fiber membrane. Membrane permeance is not usually reported by fiber manufacturers because its effect on the overall gas exchange performance of artificial lungs is negligible compared to blood-side permeance. The K_m for microporous hollow fiber membranes can be estimated theoretically using simple diffusion principles, and doing so predicts membrane permeances of approximately 2×10^{-2} ml/cm²/s/cm Hg for O₂ and CO₂ gases in the Celgard X30-240 hollow fiber membrane.^[17] The K_m of microporous hollow fibers can be measured using gas–gas test systems (i.e., with the fibers immersed in a gas rather than a liquid), so that all the transfer resistance is associated with the membrane. Kamo et al.^[18] measured the oxygen permeance of the Mitsubishi KPF fiber as $K_m = 6.7 \times 10^{-2}$ ml/cm²/s/cm Hg. Lund^[16] determined K_m values of 1.72×10^{-2} and 1.47×10^{-2} ml/cm²/s/cm Hg for O₂ and CO₂, respectively, at room temperature for the Celgard X30-240 fiber. A membrane permeance of 10^{-2} ml/cm²/s/cm Hg for microporous hollow fibers represents a very large gas exchange capacity. For example, if membrane permeance dictated overall gas exchange, an artificial lung with 2 m² membrane area perfused with blood at a PCO₂ of 50 mm Hg would remove CO₂ at a theoretical rate of 60 liters per minute! The gas exchange rate of artificial lungs is much smaller than this because overall gas exchange is dictated by diffusional boundary layers that arise on fiber surfaces in the flowing blood stream. In practice, therefore, $K \cong K_b$ unless hollow fibers are coated with nonporous polymers (true membranes) to resist plasma wetting, as will now be described.

Membrane permeance can play an important role when coated or composite hollow fiber membranes are used to prevent plasma wetting in artificial lungs, a process in which blood plasma infiltrates the microporous walls of hollow fibers. Plasma wetting is a common problem when extracorporeal oxygenators are used in extended respiratory support and can lead to device failure within days.^[19,20] Plasma wetting results primarily from phospholipids, lipoproteins

and/or proteins in blood^[19] that adsorb onto the fiber polymer surfaces at the plasma interface, rendering the interface hydrophilic and allowing for wetting of the pores by either partial or complete plasma infiltration. Plasma infiltration markedly diminishes the membrane permeance, K_m , because relatively rapid gas phase diffusion is replaced by diffusion through stagnant plasma within fiber pores. The membrane permeance for a completely wetted hollow fiber is in the range of 10^{-7} ml/cm²/s/cm Hg for O₂, a 100,000-fold decrease compared with K_m for gas-filled pores.^[17] Thus even partial plasma infiltration into fiber membranes can significantly reduce membrane permeance and degrade artificial lung performance.

Composite hollow fibers incorporate a thin nonporous polymer layer as a true membrane or “skin” on the microporous fiber surface (Fig. 6). The true membrane blocks infiltration of plasma into pores and is a key functional requirement of artificial lungs for longer-term respiratory support. Composite hollow fiber membranes^[18,21,22] are made either by coating an existing microporous fiber with a thin nonporous polymer (a true composite hollow fiber) or by

**Fig. 6** SEMs of uncoated and coated hollow fiber membrane surfaces.

modifying the fabrication of the microporous fiber itself to seal off pores at the surface (an asymmetric hollow fiber). The nonporous polymer skin that prevents plasma wetting also diminishes membrane permeance because a nonporous polymer can present an impediment to gas diffusion. Indeed, the membrane permeance of a composite hollow fiber is essentially dominated by the nonporous polymer layer and is given by

$$K_m = \frac{\alpha_p D_p}{\delta} = \frac{P_m}{\delta} \quad (4)$$

where α_p and D_p are the solubility and diffusivity of the gas within the nonporous polymer and δ is the polymer layer thickness. Polymer manufacturers usually report the *product* of polymer solubility and diffusivity, the polymer permeability (P_m) to specific gases. The design of composite hollow fiber membranes for artificial lungs requires a K_m that does not significantly reduce overall gas exchange. As an example, if coated or composite fibers are to exert no more than a 5% reduction in overall gas exchange for a particular artificial lung design, then K_m needs to be greater than 20 times K_b . For this reason, composite hollow fiber membranes for artificial lungs require nonporous polymers with relatively high gas permeabilities (~ 100 Barriers or greater^a) that can be coated in a continuous layer of 1 μm thickness or less on microporous hollow fiber surfaces.^[21]

Diffusional Boundary Layers

The blood-side permeance of an artificial lung, K_b , accounts for gas movement through the diffusional boundary layers that exist adjacent to the fiber surfaces, where fluid velocity is reduced by drag forces. Gas molecules traverse the boundary layer by molecular diffusion before being exposed to sufficient convection by the blood flowing past fiber surfaces. The blood-side permeance can be expressed as

$$K_b = \frac{\alpha_b D_b}{\delta_{bl}} \quad (5)$$

where α_b and D_b are the effective solubility and diffusion coefficient of the diffusing gas in blood and δ_{bl} is an average boundary layer thickness. For O_2 and CO_2 the effective solubility accounts for increased solubility due to hemoglobin binding (for O_2) or carriage as bicarbonate ion (for CO_2).

^aA Barrier is a common unit of polymer permeability and is equal to 10^{-10} ml-cm/cm²/s/cm Hg.

The boundary layer thickness, δ_{bl} , depends on the local interaction between diffusional and velocity fields in the flowing blood phase subjacent to the fiber surfaces of the artificial lung. The nature of these diffusional boundary layers is complex, but the simple boundary layer paradigm of laminar flow past a flat membrane surface can be instructive (Fig. 7). Boundary layer thickness on a flat surface grows with distance along the surface in the direction of flow according to

$$\delta_{bl}(x) \approx \left(\frac{\nu}{D_b}\right)^{1/6} \sqrt{\frac{D_b}{V}} x \quad (6)$$

where ν is the kinematic viscosity, D_b is the species diffusion coefficient, and V is the bulk flow velocity past the surface.^[23] An important concept is that boundary layer thickness can be decreased by increasing the blood flow velocity past the fiber surfaces, and the resulting increase in gas exchange permeance (see Eqs. 5 and 6) varies as the square root of flow velocity. Furthermore, because boundary layers grow along the fiber surface, permeance and gas exchange are less with *longitudinal* flow, parallel to the fiber axes, than with *transverse* or *cross* flow, perpendicular to the fiber axes. The simple boundary layer paradigm predicts that K_b for transverse versus longitudinal flow would be $K_b^{tran}/K_b^{long} \approx \sqrt{L/d}$, where L and d are fiber length and diameter, respectively. Since L/d in hollow fiber bundles can vary from 100 to 1000, an appreciable mass transfer benefit exists for transverse compared to parallel blood flow through hollow fiber bundles.

Mass Transfer Correlations

The blood-side permeability, K_b , for artificial lungs can be estimated from mass transfer correlations obtained for flow through bundles of hollow fiber membranes.

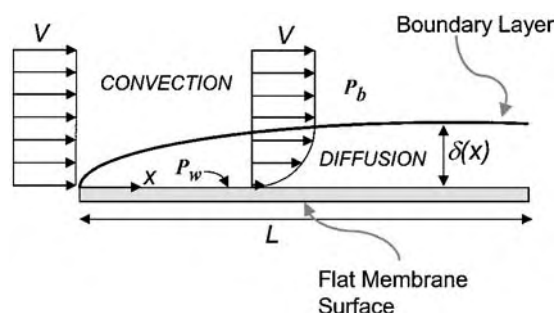


Fig. 7 Classical diffusional boundary layer on a flat surface. P_w represents the species partial pressure at the membrane wall.

Convective mass transfer correlations have the general form^[24]

$$Sh = aRe^bSc^{1/3} \quad (7)$$

where Sh is the Sherwood number, $Sh = K_b d_h / \alpha D$; Re is the Reynolds number, $Re = V d_h / \nu$; and Sc is the Schmidt number, $Sc = \nu / D$. These correlations involve the hydraulic diameter, d_h , of the fiber bed; the average interstitial blood flow velocity, V ; the kinematic viscosity of blood, ν ; and the solubility, α , and diffusivity of the diffusing species in blood, D . Several studies^[25–29] have reported mass transfer correlations in water and blood for steady flow through various module geometries of packed hollow fiber membranes, over a range of bundle porosities and Reynolds numbers. The a and b parameters in these mass transfer correlations depend on the fabricated fiber module, its porosity, and the flow patterns through the bundle. For example, Yang and Cussler^[29] reported the correlation $Sh = 1.38Re^{0.34}Sc^{0.33}$ for flow through a 750-fiber bundle with a porosity of $\epsilon = 0.3$, and the correlation $Sh = 1.9Re^{0.4}Sc^{0.33}$ for cross flow through a 72-fiber bundle with a much greater porosity of $\epsilon = 0.93$. In the design of an implantable artificial lung, Vaslef et al.^[30] used a correlation of $Sh = 0.52Re^{0.29}Sc^{0.33}$ based on measurements in a Sarns Membrane Oxygenator. Hewitt and Federspiel^[24] averaged appropriate cross-flow correlations from the literature and used $Sh = 0.524Re^{0.523}Sc^{0.33}$ for model studies of gas exchange in an intravascular artificial lung. More complex correlations^[31] were proposed for blood flowing through modules of cross-laid fiber mats oriented at various angles to the direction of blood flow.

Vaslef et al.^[27] validated a useful method for using Sh versus Re correlations to relate O_2 exchange for water flow through an oxygenator bundle to that which would arise for blood flow through the same bundle. A dimensionless mass transfer correlation, $Sh = aRe^bSc^{1/3}$, determined for a fiber bundle applies to blood and water alike (with same a and b constants), but the diffusivity in Sc number must appropriately account for hemoglobin binding of O_2 in blood. The Schmidt number involves an effective diffusivity dependent on the slope of the oxyhemoglobin dissociation curve:^[24]

$$D_b^{eff} = \frac{D_b}{1 + C_{Hb} \frac{dS}{dP_{O_2}} \alpha_b^{-1}} \quad (8)$$

where C_{Hb} is hemoglobin concentration and S is hemoglobin saturation, while the Sherwood number involves the ordinary diffusivity of oxygen in blood, D_b . The permeance in blood depends on the blood P_{O_2} within the fiber bundle and can be 2–3 times that in

water near the steep portion of the oxyhemoglobin dissociation curve.^[24] The prediction of O_2 exchange in blood from water measurements worked well for several standard blood oxygenators,^[27] and while a similar approach could be developed for relating CO_2 exchange in blood and water, none has yet to be specifically proposed or validated.

THE BLOOD OXYGENATOR

Artificial lungs that are used currently are membrane blood oxygenators consisting of either microporous polypropylene hollow fiber membranes or, as in one design, silicone sheets. The general anatomy of the oxygenator is similar between the two types of devices despite the differing gas exchange surfaces. Blood enters the oxygenator through an inlet port and flows either along the outside of the hollow fibers or the outside of the silicone sheet. The blood is then collected in a manifolded region, flows through a heat exchanger, and then exits the device through an outlet port. The gas, which can be pure oxygen or a mixture of oxygen and room air, enters the oxygenator through a gas inlet port, flows through the inside of the hollow fibers/silicone sheets, and exits the device via an outlet port (Fig. 8). The key design considerations in blood oxygenators include minimizing the resistance to blood flow, reducing the priming volume, ensuring easy debubbling at setup, and minimizing blood activation and thrombogenicity.

Most current blood oxygenators (Fig. 9) have fiber membranes with outer diameters of 200–400 μm and wall thickness of 20–50 μm , total membrane surface area of 2–4 m^2 , and blood priming volume of 135–340 ml.^[17] The hollow fibers are wound or matted

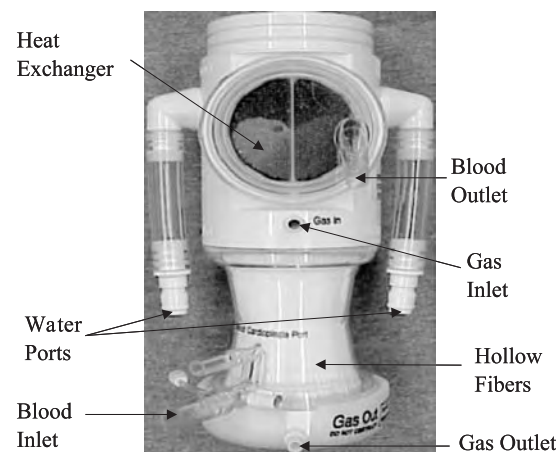


Fig. 8 A modern artificial lung showing the various ports for blood and gas flows. The heat exchanger warms the blood using water at 37°C.

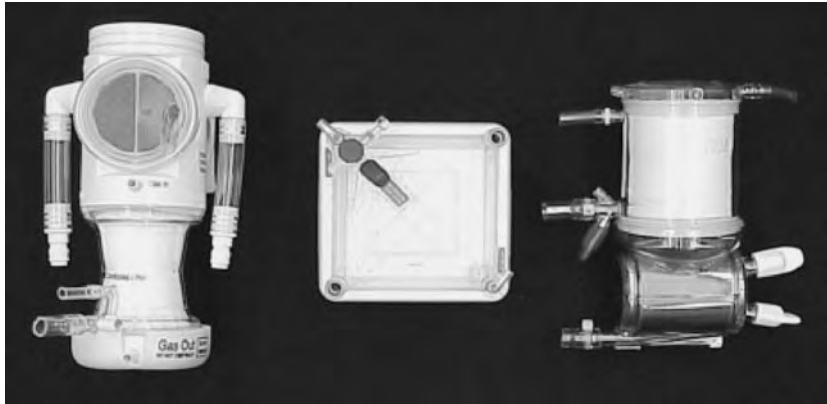


Fig. 9 Picture of currently used oxygenators. From left to right: Capiiox[®] SX from Terumo Cardiovascular Systems, Quadrox[®] from Jostra, and Affinity[®] from Medtronic.

within a hard plastic outer shell to produce fiber packing densities in the bundle of 40–60%, and the arrangement of the fiber bundle and blood flow patterns differ between devices.^[32] For example, fibers are helically wound in the Medtronic Affinity NT oxygenator. Blood enters the device through a central core channel and is then distributed radially through the fiber bundle. Fibers in the Jostra Quadrox oxygenator are aligned so that blood flow is perpendicular to the gas pathways. Hollow fiber oxygenators with intraluminal blood flow have been designed but are rarely used due to a generally unfavorable high resistance to blood flow.

The fabrication or wrapping of the fiber bundle in a blood oxygenator can be important as the geometry obtained impacts diffusional boundary layers, secondary flows, and gas exchange efficiency.^[33] A blood oxygenator is often characterized by its rated flow as a measure of the gas exchange capacity of the device. Rated flow (see Fig. 10) is the flow rate through the oxygenator at which an inlet blood saturation of 70% can be oxygenated to an outlet blood saturation of 95%.^[33] The rated flow can range from 1–1.8 l/min for a neonatal oxygenator and up to 7 l/min for an adult oxygenator (Table 2). A greater rated flow indicates an oxygenator with increased gas exchange capacity.

Silicone membrane oxygenators are often used in extracorporeal membrane oxygenation for respiratory

support since plasma leakage does not occur as it does in microporous hollow fiber oxygenators. Kolobow^[33] is generally credited with developing the first spiral-wound silicone membrane oxygenators in 1963. The oxygenator contains two silicone sheets sealed around the edges, which are wound around a polycarbonate core. Gas flows within the sealed sheets and blood flows countercurrently between the spiral wraps. The surface area of silicone membrane oxygenators ranges from 0.4 to 4.5 m² and the priming volumes range from 90 to 665 ml.^[33] Because diffusion occurs across a nonporous silicone sheet, the thickness of these sheets was reduced to 100–200 μm. Nevertheless, the gas exchange efficiency of silicone oxygenators is substantially below that of hollow fiber oxygenators. The AVECOR 0800 silicone oxygenator (a descendant of the Kolobow silicone oxygenator) has an O₂ transfer efficiency of 88 ml/min/m² compared to 150 ml/min/m² for the Affinity hollow fiber device.^[33] The resistance to blood flow is also higher in silicone sheet oxygenators compared with hollow fiber oxygenators, and debubbling the sheet oxygenators can be more difficult.

Oxygenators in Cardiopulmonary Bypass

Cardiopulmonary bypass (CPB) using an external flow circuit incorporating a blood oxygenator is used in

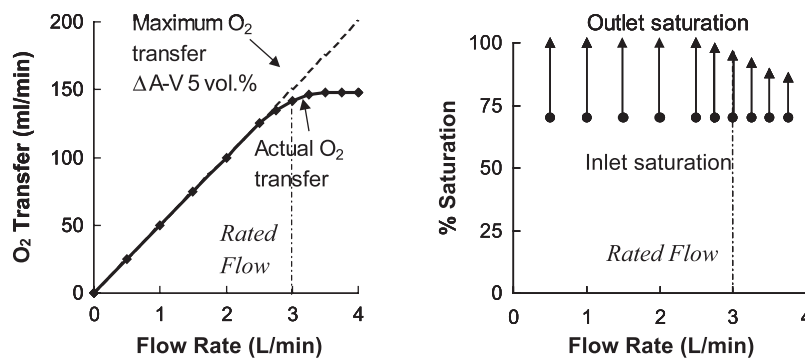


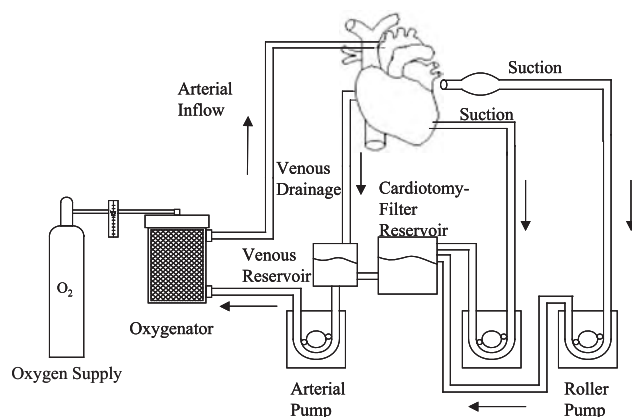
Fig. 10 The rated flow of a device is the maximum flow rate at which blood leaves the oxygenator at least 95% saturated. The rated flow depicted here is 3 l/min. (Adapted from Ref. [44].)

Table 2 Properties of blood oxygenators currently used clinically

	Membrane material ^a	Surface area (m ²)	Priming volume (ml)	Blood flow rate (l/min)	Rated blood flow (l/min)	Oxygen transfer (ml/min)
Terumo: Capiiox SX10	P HFM	1	135	0.5–4.0	–	240
Terumo: Capiiox SX18	P HFM	1.8	270	0.5–7.0	–	260
Terumo: Capiiox SX25	P HFM	2.5	340	0.5–7.0	–	300
Jostra: Quadrox	P HFM	1.8	250	0.5–7.0	–	288
Avecor 0800	SMS	0.8	100	<1.2	1.2	70
Avecor 1500	SMS	1.5	175	1.0–1.8	1.8	113
Affinity	P HFM	2.5	270	1.0–7.0	7.0	374

^aP, polypropylene; HFM, hollow fiber membranes; SMS, silicone membrane sheets.

open-heart procedures to take over the function of both the heart and the lungs. Over 700,000 open-heart surgeries, including valve replacements, coronary artery bypass grafting, and heart transplants, were performed in the United States in 1999.^[34] In cardiopulmonary bypass, blood is drained by gravity from the inferior/superior vena cava or the right atrium into a venous reservoir and is then pumped through the oxygenator by either a roller or centrifugal pump back into the ascending aorta (Fig. 11). Blood flow during CBP is kept low (2–2.4 l/m²/min) to minimize bleeding.^[35] A heat exchanger is required to cool and rewarm the patient and is typically incorporated into the oxygenator. Oxygen, or a mixture of oxygen and carbon dioxide, is fed through flowmeters and blenders into the oxygenator at flow rates of 5–10 l/min, which is 2–3 times the flow rate of blood.^[36] The oxygenator must be capable of transferring up to 250 ml/min of oxygen and 200 ml/min of carbon dioxide during cardiopulmonary bypass in order to meet the

**Fig. 11** Schematic of cardiovascular bypass circuit.

metabolic needs of the patient.^[36] The bypass circuit also includes suction devices that are used to maintain a blood-free surgical field. The suctioned blood is collected and filtered in a cardiotomy reservoir and is then pumped into the venous reservoir. Other components of the bypass circuit include pressure and temperature monitors, sampling ports, filters, tubing, and cannulae.

A significant complication associated with oxygenators in cardiopulmonary bypass is activation of the coagulation cascades and thrombosis. Patients are anticoagulated with heparin to achieve an activated clotting time (ACT) of 480 seconds to prevent thrombosis-related oxygenator failures.^[37] Thrombus formation in the oxygenator can cause an increase in resistance to blood flow and a decrease in gas transfer. Thus, to reduce the risk of clot formation, oxygenators are designed to minimize regions of blood flow stasis, which typically promote thrombus formation. The high level of anticoagulation, however, can lead to an increased risk of bleeding. Oxygenators and the entire bypass circuit are now being coated with heparin in order to prevent clotting in the circuit while reducing the required amount of systemic anticoagulation. There are several different heparin coatings currently available on the market.^[38] The Carmeda Bioactive Surface has been used for more than a decade and utilizes a covalent end-point attachment of heparin to the surface. The BioLine Coating by Jostra first coats the surface with polypeptides and then with a low-molecular-weight heparin, Liquemin. The coating is available in two types, one for CBP and one for long-term use in extracorporeal membrane oxygenation. Other more recent coatings include AOThel by Artificial Organ Technology, Corline by Corline Systems AB, and the Trilium Bio-passive Surface by Avecor. Many studies have been performed on the efficacy of

the heparin coatings and the required level of systemic heparin with the coatings. Aldea et al.^[39] compared noncoated circuits and an ACT of 480 seconds with coated circuits and an ACT of 280 seconds. The heparin coating resulted in a 34% decrease in the need for blood products, 13.8% less bleeding, 43.6% shorter intubation time, 41.7% less time in the intensive care unit, and 17.8% less time in the hospital compared with noncoated circuits.

Extracorporeal Membrane Oxygenation

Extracorporeal membrane oxygenation (ECMO) uses blood oxygenators in pump-driven external circuits to provide respiratory support and lung rest and recovery for prolonged periods of time (1–30 days).^[40] ECMO is used in patients with severe lung failure who fail traditional mechanical ventilation. Similar to CPB circuits, the ECMO circuit contains a pump, a heat exchanger, and an oxygenator, but unlike CPB circuits, a venous reservoir and suctioning equipment is not used. In ECMO, the patient is continuously anticoagulated with heparin to achieve an ACT of 160–240 seconds, much less than that found in CPB.^[41] The required blood flow in ECMO is 120 ml/kg/min for neonates, 75 ml/kg/min for pediatric, and 50 ml/kg/min for adults.^[33] Extracorporeal membrane oxygenation is used to treat neonatal, pediatric, and adult patients with lung failure, and the effectiveness of ECMO differs in each of these groups. ECMO is most commonly used in neonates with a survival rate of 80%.^[42] Indications for neonatal ECMO include meconium aspiration syndrome, respiratory distress syndrome, persistent fetal circulation, persistent pulmonary hypertension, and hyaline membrane disease.^[43] Pediatric and adult patients have lower survival rates of 53% and 41%, respectively.^[42] Indications for ECMO in pediatric or adult patients are viral, bacterial, or aspiration pneumonia and acute respiratory distress syndrome (ARDS), which can be caused by trauma, pneumonia, or sepsis.^[44]

Three different cannulation techniques can be used in ECMO, including venovenous, venoarterial, and arteriovenous, referring to the locations of the blood drainage and return sites. Venovenous ECMO drains and returns to the venous system, venoarterial ECMO drains from the venous system and returns into the arterial system, and arteriovenous ECMO is the opposite. Venovenous (VV) ECMO was established in the 1960s and 1970s and is now the most commonly used cannulation technique.^[33,45] VV ECMO has several different sites for cannulation including the internal jugular, saphenous, or femoral veins or the right atrium. In neonates, VV ECMO can use a single dual-lumen cannula or two cannulae. The single

double-lumen cannula is used in the jugular vein, and the septum offset produces a larger channel for venous inflow into the ECMO circuit. The cannula must be designed to reduce recirculation of returned blood directly back into the ECMO circuit. Cardiac output, pump flow rate, cannula position, and right atrium size are all factors that can affect recirculation. Single double-lumen cannulation cannot be used in pediatric and adult ECMO due to inadequate venous inflow into the circuit and also high levels of hemolysis, recirculation, and pressure with flow rates greater than 600 ml/min.^[45]

Venoarterial (VA) ECMO was the original cannulation technique used in ECMO and is indicated when cardiac support is required in addition to respiratory support. In contrast, VV ECMO provides no cardiac support and is not used in patients with cardiac arrest, arrhythmias, or myocardial failure.^[45] The disadvantages of VA ECMO include cannulation of a major artery, lack of pulmonary perfusion, decreased cardiac output due to a higher afterload, and increased risk of neurological events. VV ECMO has several advantages over VA ECMO including preserving pulsatility and avoiding the cannulation of a major artery. Neurological events can also be reduced since thromboemboli from the circuit travel to the lungs instead of the brain. VV ECMO also prevents ischemic injury to the lungs since the lungs remain perfused with blood, but blood flow must be carefully regulated in order to prevent an imbalance in the central venous system.^[45] Given its advantages compared to VA ECMO, several institutions are now using VV ECMO and comparing results with VA ECMO. Knight et al.^[46] found an increased survival rate of 91% with VV ECMO compared to 80% with VA ECMO in neonates. Zahraa et al.^[47] performed a retrospective study from 1986–1997 comparing VV and VA ECMO in pediatric patients and found a trend for improved survival with VV ECMO with survival rates of 60% and 56%, respectively.

The inflammatory and thrombogenic complications associated with cardiopulmonary bypass are exacerbated in ECMO due to the longer blood exposure to the extracorporeal circuit. As for CPB circuits, ECMO circuits and oxygenators are heparin-coated to help minimize systemic heparinization, decrease inflammatory responses, and prevent thrombosis. One complication of ECMO not seen in CPB is plasma wetting of hollow fiber membranes from the longer-term exposure of the ECMO oxygenator to blood. Plasma wetting decreases gas exchange, can occur quickly and unpredictably, and requires replacement of the oxygenator. Microporous hollow fiber membranes can be coated with thin siloxane layers to prevent plasma wetting and increase the biocompatibility.^[48–52] New polymer coatings are also being developed to resist



plasma leakage while attenuating the inflammatory response. Saito et al.^[52] coated CBP circuits with poly(2-methoxyethylacrylate) (PMEA) and compared the inflammatory response with that caused by uncoated circuits in swine. Protein adsorption was significantly less on the PMEA circuits compared with control ($0.3 \pm 0.03 \mu\text{g}/\text{cm}^2$ versus $3.42 \pm 0.04 \mu\text{g}/\text{cm}^2$). Peek et al.^[53] performed an initial clinical trial with the Medos Hilite 7000LT oxygenator, which uses a polymethyl pentene (PMP) asymmetric hollow fiber membrane, which was also coated with heparin. Additional studies are needed to fully evaluate the effectiveness of these new coated fiber oxygenators.

SUMMARY

Artificial lungs use bundles of microporous hollow fiber membranes made into modules designed to bring blood and gas phases in intimate contact separated by only the thin walls of the hollow fibers. This article reviewed the basic principles of gas exchange in hollow fiber-based artificial lungs. Artificial lungs used today are extracorporeal modules called blood oxygenators primarily employed for pulmonary support during open-heart surgical procedures involving cardiopulmonary bypass. To a much lesser degree, blood oxygenators are also used clinically to provide support to the failing lung in acute respiratory failure. Several research efforts are underway on the development of next-generation artificial lungs designed specifically for respiratory support of the failing lung either acutely or as a bridge to lung transplant in patients requiring chronic respiratory support. These efforts, reviewed in a companion article (see Lung, Artificial: Current Research and Future Directions), involve paracorporeal, intravascular, and intrathoracic devices. Approaches developed to specifically overcome limitations in gas exchange, biocompatibility, and other design factors of the present-day blood oxygenator are described in this article.

ACKNOWLEDGMENTS

Supported by HL70051 from the National Institutes of Health and by the U.S. Army Medical Research, Development, Acquisition, and Logistics Command under Grant No. DAMD17-98-1-8638. The views, opinions, and/or findings contained in this article are those of the authors and should not be construed as an official Department of the Army position, policy, or decision unless so designated by other documentation. KH is a recipient of a graduate fellowship in biomedical engineering from the Whitaker Foundation.

ARTICLES OF FURTHER INTEREST

Lung, Artificial: Current Research and Future Directions; Lung Surfactants

REFERENCES

1. <http://www.lungusa.org/data>. (accessed May 2001).
2. Demling, R.H. The modern version of adult respiratory distress syndrome. *Annu. Rev. Med.* **1995**, *46*, 193–202.
3. Zilberberg, M.D.; Epstein, S.K. Acute lung injury in the medical ICU: Co-morbid conditions, age, etiology, and hospital outcome. *Am. J. Respir. Crit. Care Med.* **1998**, *157* (4 Pt. 1), 1159–1164.
4. Bartlett, R.H.; Roloff, D.W.; Custer, J.R.; Younger, J.G.; Hirschl, R.B. Extracorporeal life support: The University of Michigan experience. *JAMA* **2000**, *283* (7), 904–908.
5. <http://www.lungusa.org/data>. (accessed March 2000).
6. <http://www.lungusa.org/data>. (accessed March 2001).
7. Mortensen, J.; Bezzant, T. Risk and hazard of mechanical ventilation: A collective review of published literature. *Dis.-Mon.* **1994**, *40* (11), 2193–2196.
8. Weinacker, A.B.; Vaszar, L.T. Acute respiratory distress syndrome: Physiology and new management strategies. *Annu. Rev. Med.* **2001**, *52*, 221–237.
9. AAMI Standards and Recommended Practices. Association for the Advancement of Medical Instrumentation. Volume 2.1 Biomedical Equipment, Part 1, Equipment Therapy and Surgery Cardiovascular implants and artificial organs: Blood gas exchangers. *AAMI* **1996**, *7199*, 633–648.
10. Newhouse, M.T. Tennis anyone? The lungs as a new court for systemic therapy. *CMAJ* **1999**, *161* (10), 1287–1288.
11. Weibel, E.R. *The Pathway for Oxygen: Structure and Function in the Mammalian Respiratory System*; Harvard University Press: Cambridge, 1984.
12. Federspiel, W.J.; Sawzik, P.J.; Borovetz, H.S.; Reeder, G.D.; Hattler, B.G. Temporary Support of the Lungs—The Artificial Lung. In *The Transplantation and Replacement of Thoracic Organs*; Cooper, D.K.C., Miller, L.W., Patterson, G.A., Eds.; Kluwer Academic Publishers: Boston, 1996; 717–728.
13. High, K.M.; Snider, M.T.; Bashein, G. Principles of Oxygenator Function: Gas Exchange, Heat Transfer, and Blood-Artificial Surface Interaction. In *Cardiopulmonary Bypass Principles and Practice*; Gravlee, G.P., Davis, R.F., Utley, J.R., Eds.; Williams & Wilkins: Philadelphia, 1993; 28–54.
14. LaPierre, R.A., Jr.; Howe, R.J.; Haw, M.P.; Elliott, M. Oxygenators for Pediatric Cardiac Surgery. In *Cardiopulmonary Bypass in Neonates, Infants and Young Children*; Jonas, R.A., Elliott, M.J., Eds.; Butterworth-Heinemann: Boston, 1994; 173–185.
15. Beckley, P.D.; Holt, D.W.; Tallman, R.D., Jr. Oxygenators for Extracorporeal Circulation. In *Cardiopulmonary Bypass: Principles and Techniques of Extracorporeal Circulation*; Mora, C.T., Ed.; Springer-Verlag: New York, 1995; 199–219.

16. Lund, L. Measurement of Hollow Fiber Membrane Permeance in a Gas-Liquid System Ph.D.; University of Pittsburgh, 2000.
17. Hattler, B.G.; Federspiel, W.J. Gas Exchange in the Venous System: Support for the Failing Lung. In *The Artificial Lung*; Vaslef, S.N., Anderson, R.W., Eds.; Landes Bioscience, 2002; 133–174.
18. Kamo, J.; Uchida, M.; Hirai, T.; Yosida, H.; Kamada, K.; Takemura, T. A new multilayered composite hollow fiber membrane for artificial lung. *Artif. Organs* **1990**, *14* (5), 369–372.
19. Montoya, J.P.; Shanley, C.J.; Merz, S.I.; Bartlett, R.H. Plasma leakage through microporous membranes. Role of phospholipids. *ASAIO J.* **1992**, *38* (3), M399–M405.
20. Mottaghy, K.; Oedekoven, B.; Starmans, H.; Muller, B.; Kashefi, A.; Hoffmann, B.; Bohm, S. Technical aspects of plasma leakage prevention in microporous capillary membrane oxygenators. *ASAIO Trans.* **1989**, *35* (3), 640–643.
21. Mulder, M. Preparation of Synthetic Membranes. In *Basic Principles of Membrane Technology*; Kluwer Academic Publishers: Boston, 1996; 71–156.
22. Niimi, Y.; Ueyama, K.; Yamaji, K.; Yamane, S.; Tayama, E.; Sueoka, A.; Kuwana, K.; Tahara, K.; Nose, Y. Effects of ultrathin silicone coating of porous membrane on gas transfer and hemolytic performance. *Artif. Organs* **1997**, *21* (10), 1082–1086.
23. Probstein, R.F. Solutions of Uncharged Molecules. In *Physicochemical Hydrodynamics*; John Wiley & Sons, Inc.: New York, 1994; 53–107.
24. Hewitt, T.J.; Hattler, B.G.; Federspiel, W.J. A mathematical model of gas exchange in an intravenous membrane oxygenator. *Ann. Biomed. Eng.* **1998**, *26* (1), 166–178.
25. Mockros, L.F.; Leonard, R.J. Compact cross-flow tubular oxygenators. *ASAIO Trans.* **1985**, *31*, 628–632.
26. Vaslef, S.N.; Cook, K.E.; Leonard, R.J.; Mockros, L.F.; Anderson, R.W. Design and evaluation of a new, low pressure loss, implantable artificial lung. *ASAIO J.* **1994**, *40* (3), M522–M526.
27. Vaslef, S.N.; Mockros, L.F.; Anderson, R.W.; Leonard, R.J. Use of a mathematical model to predict oxygen transfer rates in hollow fiber membrane oxygenators. *ASAIO J.* **1994**, *40* (4), 990–996.
28. Wickramasinghe, S.R.; Semmens, M.J.; Cussler, E.L. Mass transfer in various hollow fiber geometries. *J. Membr. Sci.* **1992**, *69*, 235–250.
29. Yang, M.C.; Cussler, E.L. Designing hollow fiber contactors. *AIChE J.* **1986**, *32*, 1910–1916.
30. Vaslef, S.N.; Mockros, L.F.; Cook, K.E.; Leonard, R.J.; Sung, J.C.; Anderson, R.W. Computer-assisted design of an implantable, intrathoracic artificial lung. *Artif. Organs* **1994**, *18* (11), 813–817.
31. Capatano, G.; Papenfuss, H.D.; Wodetzki, A.; Baurmeister, U. Mass and momentum transport in extra-luminal flow (ELF) membrane devices for blood oxygenation. *J. Membr. Sci.* **2001**, *184* (1), 123–135.
32. Dierickx, P.W.; De Wachter, D.S.; De Somer, F.; Van Nooten, G.; Verdonck, P.R. Mass transfer characteristics of artificial lungs. *ASAIO J.* **2001**, *47*, 628–633.
33. Hirschl, R. Devices. In *ECMO Extracorporeal Cardiopulmonary Support in Critical Care*; Zwischenberger, J.B., Bartlett, R.H., Eds.; Extracorporeal Life Support Organization: Ann Arbor, MI, 1995; 159–190.
34. <http://www.americanheart.org/presenter.jhtml?identifier=4674>. (accessed November 2002).
35. Bartlett, R.H. Physiology of Extracorporeal Life Support. In *ECMO Extracorporeal Cardiopulmonary Support in Critical Care*; Zwischenberger, J.B., Bartlett, R.H., Eds.; Extracorporeal Life Support Organization: Ann Arbor, MI, 1995; 27–52.
36. Galletti, P.M.; Colton, C.K. Artificial Lungs and Blood-Gas Exchange Devices. In *The Biomedical Engineering Handbook*; Bronzino, J.D., Ed.; CRC Press LLC: Boca Raton, 2000; 1–19.
37. Gravlee, G.P. Anticoagulation for Cardiopulmonary Bypass. In *Cardiopulmonary Bypass Principles and Practice*; Gravlee, G.P., Davis, R.F., Utley, J.R., Eds.; Williams & Wilkins: Philadelphia, 1993; 340–380.
38. Wendel, H.P.; Ziemer, G. Coating-techniques to improve the hemocompatibility of artificial devices used for extracorporeal circulation. *Eur. J. Cardio-Thorac. Surg.* **1999**, *16* (3), 342–350.
39. Aldea, G.S.; Doursounian, M.; O’Gara, P.; Treanor, P.; Shapira, O.M.; Lazar, H.L.; Shemin, R.J. Heparin-bonded circuits with a reduced anticoagulation protocol in primary CABG: A prospective, randomized study. *Ann. Thorac. Surg.* **1996**, *62*, 410–418.
40. Zwischenberger, J.B.; Bartlett, R.H. An Introduction to Extracorporeal Life Support. In *ECMO Extracorporeal Cardiopulmonary Support in Critical Care*; Zwischenberger, J.B., Bartlett, R.H., Eds.; Extracorporeal Life Support Organization: Ann Arbor, MI, 1995; 11–14.
41. Dark, J.H. Extracorporeal Respiratory Support. In *Techniques in Extracorporeal Circulation*; Kay, P.H., Ed.; Butterworth-Heinemann Ltd.: Boston, 1992; 309–320.
42. Tracy, T., Jr.; DeLosh, T.; Stolar, C. The Registry of the Extracorporeal Life Support Organization. In *ECMO Extracorporeal Cardiopulmonary Support in Critical Care*; Zwischenberger, J.B., Bartlett, R.H., Eds.; Extracorporeal Life Support Organization: Ann Arbor, MI, 1995; 251–260.
43. Ichiba, S.; Bartlett, R.H. Current status of extracorporeal membrane oxygenation for severe respiratory failure. *Artif. Organs* **1996**, *20* (2), 120–123.
44. Mols, G.; Loop, T.; Geiger, K.; Farthmann, E.; Benzing, A. Extracorporeal membrane oxygenation: A ten-year experience. *Am. J. Surg.* **2000**, *180* (2), 144–154.
45. Cornish, J.; Clark, R. Principles and Practice of Venovenous Extracorporeal Membrane Oxygenation. In *ECMO Extracorporeal Cardiopulmonary Support in Critical Care*; Zwischenberger, J.B., Bartlett, R.H., Eds.; Extracorporeal Life Support Organization: Ann Arbor, MI, 1995; 87–110.
46. Knight, G.R.; Dudell, G.G.; Evans, M.L.; Grimm, P.S. A comparison of venovenous and venoarterial extracorporeal membrane oxygenation in the treatment of neonatal respiratory failure. *Crit. Care Med.* **1996**, *24* (10), 1678–1683.
47. Zahraa, J.N.; Moler, F.W.; Annich, G.M.; Maxvold, N.J.; Bartlett, R.H.; Custer, J.R. Venovenous versus



- venoarterial extracorporeal life support for pediatric respiratory failure: Are there differences in survival and acute complications? *Crit. Care Med.* **2000**, *28* (2), 521–525.
48. Shimono, T.; Shomura, Y.; Tahara, K.; Hioki, I.; Tenpaku, H.; Maze, Y.; Hirano, R.; Shimpo, H.; Shionoya, Y.; Yokoyama, A.; Morikan, T.; Yada, I. Experimental evaluation of a newly developed ultrathin silicone layer coated hollow fiber oxygenator. *ASAIO J.* **1996**, *42* (5), M451–M454.
 49. Shimono, T.; Shomura, Y.; Hioki, I.; Shimamoto, A.; Tenpaku, H.; Maze, Y.; Onoda, K.; Takao, M.; Shimpo, H.; Yada, I. Silicone-coated polypropylene hollow-fiber oxygenator: Experimental evaluation and preliminary clinical use. *Ann. Thorac. Surg.* **1997**, *63* (6), 1730–1736.
 50. Watanabe, H.; Hayashi, J.; Ohzeki, H.; Moro, H.; Sugawara, M.; Eguchi, S. Biocompatibility of a silicone-coated polypropylene hollow fiber oxygenator in an in vitro model. *Ann. Thorac. Surg.* **1999**, *67* (5), 1315–1319.
 51. Shimamoto, A.; Kanemitsu, S.; Fujinaga, K.; Takao, M.; Onoda, K.; Shimono, T.; Tanaka, K.; Shimpo, H.; Yada, I. Biocompatibility of silicone-coated oxygenator in cardiopulmonary bypass. *Ann. Thorac. Surg.* **2000**, *69* (1), 115–120.
 52. Saito, N.; Motoyama, S.; Sawamoto, J. Effects of new polymer-coated extracorporeal circuits on biocompatibility during cardiopulmonary bypass. *Artif. Organs* **2000**, *24* (7), 547–554.
 53. Peek, G.J.; Killer, H.M.; Reeves, R.; Sosnowski, A.W.; Firmin, R.K. Early experience with a polymethyl pentene oxygenator for adult extracorporeal life support. *ASAIO J.* **2002**, *48* (5), 480–482.

Lung, Artificial: Current Research and Future Directions

William J. Federspiel

Department of Chemical Engineering, University of Pittsburgh,
Department of Bioengineering, University of Pittsburgh, and Artificial Lung Laboratory,
McGowan Institute of Regenerative Medicine, University of Pittsburgh, Pittsburgh,
Pennsylvania, U.S.A.

Robert G. Svitek

Department of Chemical Engineering, University of Pittsburgh, and Artificial Lung
Laboratory, McGowan Institute of Regenerative Medicine, University of Pittsburgh,
Pittsburgh, Pennsylvania, U.S.A.



INTRODUCTION

Artificial lungs are medical devices designed to take over or supplement the respiratory function of the lung: oxygenating the blood and removing carbon dioxide. They are generally classified as extracorporeal, paracorporeal, intravascular, or intrathoracic devices (Fig. 1). The artificial lungs used clinically today are extracorporeal blood oxygenators, primarily used in operations requiring cardiopulmonary bypass, but also used less frequently for support of patients with respiratory failure. The growing incidence of lung disease associated with the aging population has spurred recent work toward next-generation artificial lungs that may be used to successfully treat patients with a variety of respiratory failures. Next-generation artificial lungs include 1) paracorporeal approaches (wearable devices that will be attached directly to patients); 2) intravascular approaches (respiratory catheters placed within the vena cava through a peripheral vein); and 3) intrathoracic approaches (devices that will be placed internally within the thoracic or abdominal cavities). Intravascular artificial lungs are aimed primarily at treatment of acute respiratory failures, whereas paracorporeal and intrathoracic artificial lungs are aimed primarily at bridge-to-transplant respiratory support because of the more invasive procedures required to implement them. This article reviews the current state of research on next-generation artificial lungs for support of patients with failing lungs. The basic operational principles of artificial lungs and the current clinical applications of standard blood oxygenators are reviewed in a companion article (*Lung, Artificial: Basic Principles and Current Applications*).

ARTERIOVENOUS CARBON DIOXIDE REMOVAL

Arteriovenous carbon dioxide removal (AVCO₂R) is more the development of a new respiratory support technique than a next-generation artificial lung device. AVCO₂R represents a simpler means of extracorporeal

respiratory support and involves a blood oxygenator (either a standard blood oxygenator or one developed specifically for the application) connected directly from the arterial to venous circulations in a paracorporeal approach, thus obviating the need for a blood pump and associated tubing in the extracorporeal circuit. The goal of AVCO₂R is to provide lung rest to patients suffering from acute respiratory distress syndrome (ARDS) or other acute lung failures^[1] by reducing the tidal volume, minute ventilation, and pressures associated with mechanical ventilation.^[2,3] The specific technique used in AVCO₂R consists of diverting blood flow through a femoral arterial cannula to a commercial blood oxygenator (as the artificial lung) and back to the patient through a femoral venous cannula (Fig. 2). The AVCO₂R circuit principally removes CO₂ before returning the blood back to the patient, but some oxygenation of the blood occurs as well. The flowrate is much lower than in typical extracorporeal membrane oxygenation (ECMO) circuits and is dictated by the arterial-to-venous pressure difference of the patient and the hydraulic resistance of the artificial lung. Although AVCO₂R eliminates the need for a mechanical blood pump, it ideally requires an oxygenator/artificial lung with a sufficiently low hydraulic resistance to ensure adequate blood flow from the arteriovenous pressure difference to remove a sufficient amount of CO₂.

Recent animal studies using healthy adult sheep have characterized the flowrate and gas exchange capabilities of an AVCO₂R system.^[4] The artificial lung used in the AVCO₂R application was a 2.5 m² Affinity™ blood oxygenator connected to the left carotid artery via an 18-Fr cannula and the left jugular vein via a 22-Fr cannula. Blood flowrates through the AVCO₂R circuit were 25–29% of cardiac output during the 6-hour study, with a pressure drop across the oxygenator of less than 10 mm Hg. Maximum CO₂ removal was approximately 100 ml/min, or 96% of total CO₂ production by the animal. Importantly, the study showed a decrease in minute

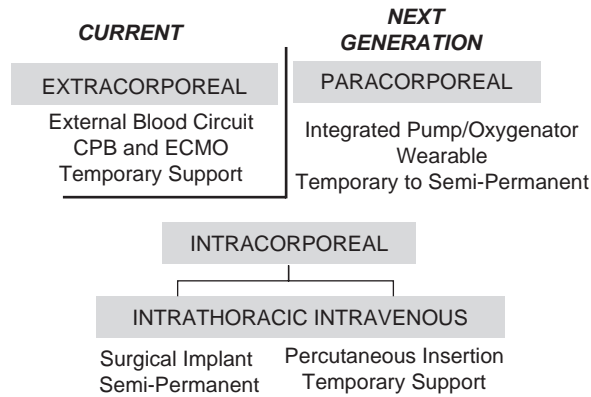


Fig. 1 Classification of artificial lung technology.

ventilation from approximately 10 l/min before AVCO₂R to 0.5 l/min, and peak inspiratory pressure also decreased from 41 cmH₂O to 20 cmH₂O in subsequent hours after initiation of the therapy. The effect of AVCO₂R blood flow on organ perfusion was investigated using colored microspheres in a conscious ovine model with the carotid-to-jugular arteriovenous circuit using the Affinity™ 2.5 m² oxygenator.^[5] AVCO₂R flow

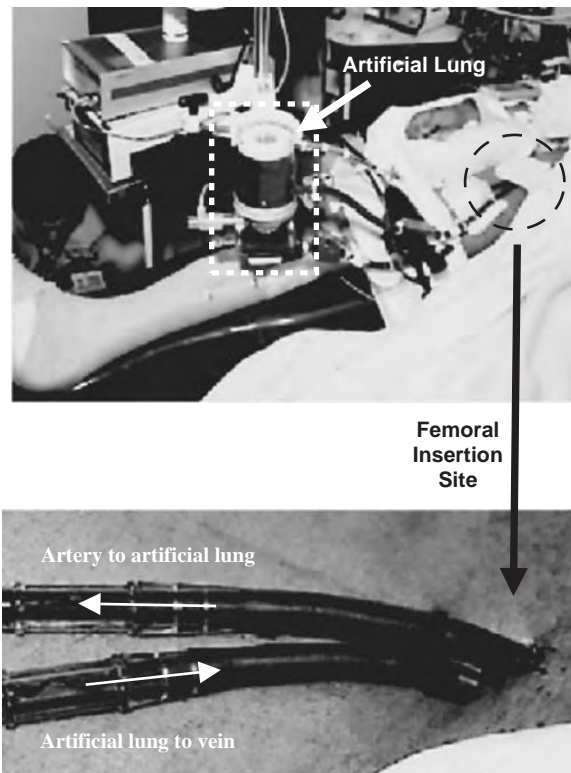


Fig. 2 This simple circuit consists of the inlet cannula, the artificial lung, and the return cannula to the patient. The insertion site in this case is the femoral artery and vein in the patient's right leg. (Courtesy of Joseph B. Zwischenberger, University of Texas Medical Branch, Galveston, TX.)

varied from 5–25% of baseline cardiac output and organ flow was measured using microsphere techniques. A 10–20% decrease in organ perfusion accompanied a 5% shunt, but organ perfusion did not decrease further with increased shunt flow through the AVCO₂R device. Hemodynamic parameters such as heart rate, mean arterial, and pulmonary arterial pressures remained unchanged throughout the study, indicating that blood shunting through the AVCO₂R system was well-tolerated by the animals.^[5]

A 7-day study of AVCO₂R in sheep using a smoke-inhalation injury model of severe respiratory failure showed gas exchange capabilities similar to those in previous AVCO₂R studies using the same Affinity™ blood oxygenator with no statistically significant changes in heart rate, cardiac output, mean arterial pressure, or pulmonary artery pressure.^[6] Increased survival among sheep managed with AVCO₂R versus ventilator-managed sheep has also been shown for this ARDS model.^[7] In the latter study, 18 sheep received an LD₅₀ severe smoke inhalation and a 40% third-degree burn on their external flanks. Animals were randomized to either a ventilator-managed sham group ($n=9$) or an AVCO₂R therapy group ($n=9$) and treated for 7 days. AVCO₂R allowed significant ventilator reductions in tidal volume, peak inspiratory pressure, minute ventilation, respiratory rate, and FIO₂ (Fig. 3). All AVCO₂R-managed animals survived the study, whereas only three of the ventilator-managed sheep survived the duration of the study.

Respiratory support with AVCO₂R has recently expanded into human clinical trials. In a pilot study of feasibility and safety, percutaneous AVCO₂R was

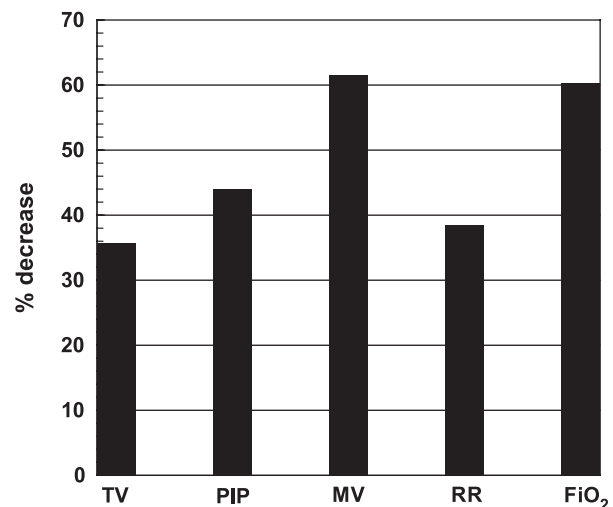


Fig. 3 Decreased ventilator parameters after instituting AVCO₂R in sheep with burn and smoke-inhalation-induced ARDS. This study lasted 7 days. TV: tidal volume; PIP: peak inspiratory pressure; MV: minute ventilation; RR: respiratory rate; FIO₂: fraction of inspired O₂.

studied in five patients suffering from severe ARDS and CO₂ retention who were unresponsive to standard mechanical ventilation protocols.^[8] AVCO₂R was initiated for 72 hours using the 2.5 m² Affinity™ blood oxygenator with 12- to 15-Fr venous and 10- to 12-Fr arterial cannulae inserted percutaneously into the femoral vein and artery, respectively. Average blood flowrates through the artificial lung ranged from approximately 600–1100 ml/min with maximum CO₂ removal rate of 208 ml/min. The minute ventilation was significantly decreased from 7.2 ± 2.3 l/min before initiation of AVCO₂R to 3.4 ± 0.8 l/min after 24 hours of AVCO₂R therapy. A follow-up phase I clinical study lasting 72 hours in eight patients with acute respiratory failure and hypercapnia was also conducted.^[9] Blood gas measurements, ventilator parameters, and hemodynamics were measured before and after initiation of AVCO₂R. The arterial CO₂ concentration decreased from 90.8 ± 7.5 mm Hg to 51.8 ± 3.1 mm Hg, and minute ventilation was reduced from 6.92 ± 1.65 l/min to 3.00 ± 0.53 l/min after 2 hours of AVCO₂R. Hemodynamic changes were not significant, nor were any major complications associated with AVCO₂R reported during the study. Zwischenberger and colleagues from the University of Texas Medical Branch continue work in this area, focused on developing AVCO₂R into a standard therapy for respiratory support during severe respiratory failure.

INTRAVASCULAR RESPIRATORY CATHETERS

Intravascular artificial lungs have been studied and are being developed as a less expensive, less personnel-intensive alternative to respiratory support with extracorporeal artificial lungs. Anatomical and physiological constraints of device placement in major blood vessels of the human body impose significant challenges in developing intravascular artificial lungs. Most of the intravascular devices that have been developed are intended for insertion through a peripheral vein (femoral or jugular) and placement in the vena cava, the largest blood vessel in the body through which blood returns to the heart (Fig. 4). The adult human inferior vena cava ranges on average from 2.2 cm to 3.3 cm in diameter and the superior vena cava ranges from 1.5 cm to 2.2 cm.^[10] Intravascular artificial lungs must be compact for insertion, yet possess sufficient membrane area to achieve adequate respiratory support. The primary objective of intravascular artificial lungs is to supplement the gas exchange of a failing lung, but not completely replace it. Respiratory support at 40–60% of the body's resting metabolic needs has generally been considered an appropriate target for intravascular artificial lungs.^[11]

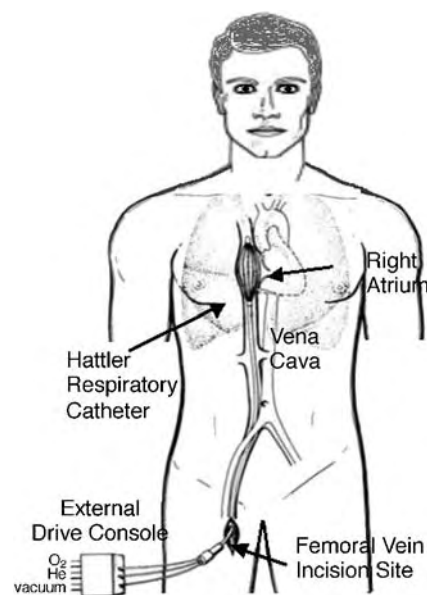


Fig. 4 Schematic of intravascular artificial lung or respiratory support catheter placed within the human vena cava.

Four principal development efforts have tackled the challenge of intravascular artificial lungs since the 1980s. Mortensen and colleagues at CardioPulmonics, Inc. (Salt Lake City, UT) developed the IVOX, the only intravascular artificial lung that has undergone human clinical trials.^[12–14] The IVOX consisted of a bundle of crimped hollow fiber membranes joined at the distal end to the inner lumen of a dual-lumen gas conduit, and at the proximal end to the outer lumen of the gas conduit, which led outside the body to a console for providing sweep gas flow through the fibers. The crimped fibers (Fig. 5) of the IVOX helped minimize fiber clumping in the vena cava and also helped disturb blood flow and diffusional boundary layers on fiber surfaces to improve overall gas exchange permeance.^[15,16] A total of 160 patients with severe acute respiratory distress were studied in applications that

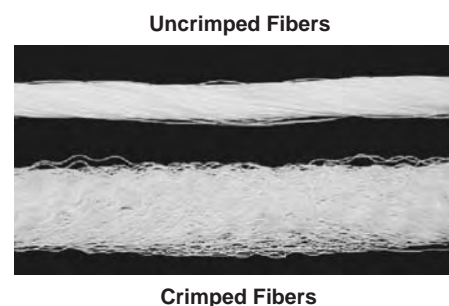


Fig. 5 Crimped fibers were used in the IVOX device to minimize fiber clumping and disrupt blood flow.

lasted up to 28 days of support.^[12] The clinically tested IVOX ranged from 0.21 m² to 0.51 m² in membrane area, and the average rates of O₂ and CO₂ transfer accomplished in the trials ranged from 40 to 70 ml/min, or about 20–30% of baseline metabolic needs. The IVOX demonstrated that intravascular artificial lungs can be implanted within the vena cava and perform for extended periods without significant complications *in situ* (for example, from blood thrombosis). About 30% of the patients showed improvement in blood gases that allowed reduction in mechanical ventilation, but overall the level of respiratory support provided was considered marginal.^[12] The choice of a clinical trial without a random prospective control arm hampered efforts at gaining Food and Drug Administration (FDA) approval, and CardioPulmonics eventually discontinued its attempt at FDA approval and halted further device development.

Several innovative intravascular artificial lung prototypes were developed at Northwestern University as the ILAD, for intravascular lung assist device.^[17–19] To enhance gas exchange, an active or dynamic D-ILAD device was developed, consisting of sheets of short microporous fibers stacked, folded, and twisted around a central shaft to create a helical or screw-like arrangement of fiber surfaces.^[20] Rotation of the entire D-ILAD device increased blood convection across the fiber surfaces, thus improving gas transfer performance while also providing some pumping motion to blood flow. In pilot experiments, oscillatory motion of the D-ILAD bundle showed increased gas exchange compared to steady rotational motion. Prototype D-ILAD devices with membrane areas from 0.1 to 0.5 m² accomplished O₂ and CO₂ transfer rates of 208 ml/min/m² and 310 ml/min/m², respectively, in bench tests involving bovine blood flow through a mock vena cava at 2 l/min.^[21] The D-ILAD only achieved adequate gas exchange with significant fiber rotation, which may be difficult to implement in the vena cava; the efforts of the Northwestern group have shifted to developing a total artificial lung device.^[22,23]

An intravascular artificial lung designed to be placed through the right ventricle into the pulmonary artery (PA) was developed at Penn State University as the PENSIL, for Penn State Intravascular Lung.^[24,25] Placement within the PA allows the entire cardiac output to flow past the fiber surfaces to help maximize gas exchange. The PENSIL consisted of a central catheter with spaced manifold sleeves around it from which blind-ended hollow fibers emanated. The central catheter served as the spine of the PENSIL, and the fibers emanating from the manifold resembled long bristles of a bottle brush. The use of blind-ended fibers meant that sweep gas must diffuse into fibers rather than being convected through them, as in most

artificial lungs. Oscillation of the pressure in the central catheter was studied as a method to provide some oscillatory convection of sweep gas in and out of the blind-ended fibers.^[26] PENSIL devices were fabricated with membrane areas ranging from 0.04 to 0.38 m² and, compared to the IVOX, achieved reasonable O₂ exchange efficiencies at 140 ml/min/m². CO₂ exchange (corrected to a blood PCO₂ driving force of 50 mm Hg) was low, however, at only 25 ml/min/m². Catheter vibration was recently considered as a means of improving gas exchange in the PENSIL,^[14] but other than this effort, the development of the PENSIL appears dormant.

The only currently active development program on intravascular artificial lungs is the effort at the McGowan Institute for Regenerative Medicine at the University of Pittsburgh. This intravascular artificial lung or respiratory support catheter (also known as the Hattler Catheter, and formally referred to as the intravenous membrane oxygenator (IMO)) uses a centrally positioned pulsating balloon within a fiber bundle made from fiber fabric wrapped around the central balloon (Fig. 6). Balloon pulsation pumps blood around the hollow fiber surfaces at greater blood flow velocities than would otherwise flow in the vena cava, and as a result enhances gas exchange by decreasing diffusional boundary layers. Development efforts over the past seven years have led to improvements in balloon pulsation (faster possible pulsation rates) and balloon-bundle design to further increase the gas exchange of the respiratory catheter. The design of the respiratory catheter has evolved substantially from earlier prototypes^[27–31] and has been tested extensively in animals, including *ex-vivo* tests^[29] used to make comparisons with the IVOX, acute animal implants,^[32] and some limited chronic implant tests.^[33] Initially intended for surgical insertion in the femoral vein through a cut-down procedure, the respiratory catheter is now undergoing further

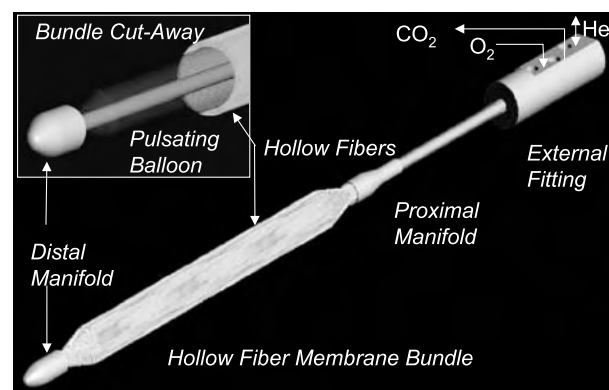


Fig. 6 Respiratory support catheter.

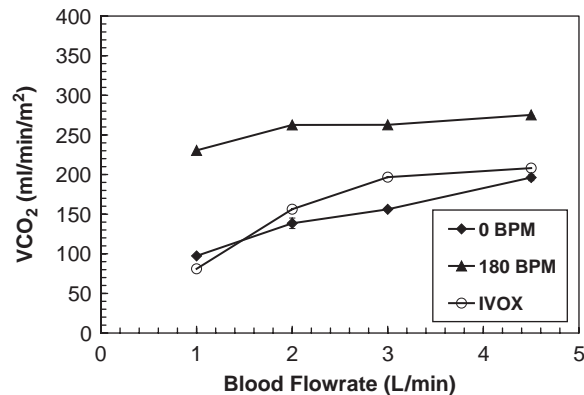


Fig. 7 Ex-vivo results for pulsating respiratory catheter (at 0 and 180 beats per minute) versus IVOX as function of blood flowrate.

development to reduce catheter insertion size and allow for percutaneous insertion.

The pulsating balloon has been a key component of our respiratory support catheter that allows for greater gas exchange than a passive device like the IVOX. Recent ex-vivo tests^[29] were conducted by placing the catheter in a mock vena cava vessel (1-inch tube) connected in an extracorporeal circuit to a calf, a test and setup analogous to those used to study the clinically tested IVOX.^[11] Figure 7 shows the CO₂ removal rate achieved with the respiratory (Hattler) catheter (HC) tested ex vivo in comparison to the IVOX over a range of relevant blood flowrates through the mock vena cava. Balloon pulsation increased gas exchange 200–300% at the lowest blood flowrate, compared to 50–100% enhancement at the highest blood flowrate. In this manner balloon pulsation eliminated much of the dependence of the gas exchange rate on blood flowrate as seen in passive oxygenators, suggesting that the respiratory catheter may exhibit less gas exchange variability related to cardiac output variability in patients. In the ex-vivo studies, the CO₂ and O₂ gas exchange rates of the catheter at maximal balloon pulsation varied from approximately 250 to 300 ml/min/m² and, depending on the blood flowrate, were from 50–500% greater than IVOX gas exchange in these tests.^[34] In a series of acute implants of the respiratory catheter in calves,^[7] the CO₂ transfer rate was 320 ml/min/m², comparable to that achieved in the ex-vivo tests.

TOTAL ARTIFICIAL LUNGS

Several research groups are developing total artificial lung (TAL) devices for treating chronic respiratory failure, primarily as bridge-to-lung transplant devices. Whereas implantable total artificial lungs that can be

placed in the thoracic or abdominal cavities may be an ultimate goal, the initial implementation and testing of total artificial lungs appears to favor paracorporeal applications, with the TAL external to but immediately attached to the patient. The attachment mode of the TAL is an important design consideration, and in-series, in-parallel, and hybrid configurations have been studied.^[35] The in-series configuration connects the artificial lung to the proximal pulmonary artery, diverting all the cardiac output through the device and returning it to the distal pulmonary artery immediately upstream of the natural lungs. Although this mode enables the natural lung to be an effective embolic filter, the mechanical load on the right heart increases because it must provide the pumping energy for blood flow through both the natural and the artificial lung. An estimated 0.11 W/(l/min) is required by the right ventricle with the TAL in-series configuration.^[36] The in-parallel configuration attaches the artificial lung between the pulmonary artery and the left atrium so that only a fraction of the blood flow diverts through the artificial lung. The fraction of blood flow through the artificial lung depends on its impedance relative to that of the natural lung. The in-parallel configuration has the clear advantage that the right heart workload is reduced, but only a fraction of total cardiac output receives respiratory support from the artificial lung and that fraction is not exposed to the metabolic and filtering functions of the natural lung. The power requirement of the in-parallel configuration (assuming 2/3 of total flow through the TAL and 1/3 through the natural lung) is roughly half that for the in-series configuration.^[36] The hybrid configuration attaches the inlet of the artificial lung to the proximal pulmonary artery, and uses a split return to the distal pulmonary artery (and natural lung) and to the left atrium. The hybrid configuration allows all the cardiac output to flow through the artificial lung with less resistance than the in-series configuration, and also allows greater flow through the natural lung than the in-parallel configuration. The power requirement of the right heart for this configuration depends on the amount of blood flow through the artificial lung relative to the natural lung, and has been estimated at between approximately 0.08 W/(l/min) and 0.10 W/(l/min).^[36] Patients with a weak or failing right ventricle would require either the in-parallel or hybrid configurations because of the reduced power required for adequate perfusion of the artificial lung and natural lung.

The implantable Intrathoracic Artificial Lung (ITAL) under development at Northwestern University focuses on resting the lung in acute respiratory failure and as a bridge-to-lung transplantation in chronic lung failure.^[37] Mathematical models were developed to estimate the required surface area for 200 ml/min of oxygen transfer at a blood flowrate of

Table 1 Specifications of a first ITAL prototype

Specification	Design target
O ₂ transfer rate	> 200 ml/min
Blood-side pressure drop	< 15 mm Hg
Blood flowrate	3–5 l/min
Fiber outer diameter	0.038 cm
Surface area	2.2 m ²
Priming volume	240 ml

5 l/min with a pressure drop of less than 15 mm Hg.^[38] Based on these analyses a prototype implantable artificial lung with the design specifications shown in Table 1 was fabricated and tested.^[39] In-vitro gas exchange rates in bovine blood were 150–200 ml/min of O₂ and over 200 ml/min of CO₂ at 4 l/min blood flow. The device was implanted in two Yorkshire pigs for less than 6 hours, and gas exchange rates were approximately 99 ml/min of O₂ and 86–144 ml/min of CO₂. The pressure drop across the device was approximately 10 mm Hg. The ITAL was later modified to reduce the resistance to flow of the device. These changes included increasing the device porosity—or void fraction—from 0.53 to 0.74, reducing the surface area from 2.2 m² to 1.83 m², increasing the inlet and outlet grafts to the device from $\frac{1}{4}$ " to nearly $\frac{3}{4}$ ", and adding an inlet compliance chamber made of thin-walled latex tubing 1 inch in inner diameter.^[37] The device, tested on three Yorkshire pigs, had average gas exchange ranging from 156–204 ml/min of O₂ and 187–242 ml/min of CO₂. The power requirements of the right ventricle were determined to be approximately 0.06 W/(l/min) for full blood flow through the ITAL. The effects of the device housing and inlet graft compliance on the right heart power requirements have also been studied using a computational model.^[36]

The BioLung™ total artificial lung under development at MC3, Inc. (Ann Arbor, MI) and the University of Michigan is intended for complete respiratory support as a bridge to transplant for 1–6 months (Fig. 8). Initial BioLung prototypes were evaluated paracorporeally and connected in series with the natural lung.^[22] Five of 7 sheep survived with 75–100% of cardiac output diverted through the device during a 24-hour study with oxygen transfer rates of approximately 150–240 ml/min. Two of the animals did not tolerate complete occlusion of the pulmonary artery; therefore total cardiac output was not diverted through the artificial lung in these animals. A 168-hour study found 50% of the sheep died of right heart failure when the device was placed in series with the natural lung due to increased flow impedance caused by the device.^[23] Several changes to the TAL were made to reduce its flow resistance (or impedance),



Fig. 8 BioLung™ artificial lung prototype from MC3, Inc. The center port is the blood inlet, and the two side ports are blood exit ports. (Courtesy of MC3, Ann Arbor, MI.)

including addition of an inflow cannula compliance chamber, and inlet blood flow separator and modification of the outlet geometry. A 72-hour healthy sheep survival model using the modified device showed 6 of the 7 sheep exhibited good cardiac function with gas exchange averages of 220 ml/min of O₂ and 166 ml/min of CO₂.^[40] The most recent study has compared volume-controlled mechanical ventilation to the modified BioLung TAL in a 5-day smoke inhalation/burn ARDS sheep study. Six of 8 animals survived when treated with the artificial lung, and only 1 of 6 animals survived when treated with mechanical ventilation.^[41] Average gas transfer rates for the artificial lung were 218.6 ± 17.7 ml/min O₂ and 183.0 ± 27.8 ml/min CO₂. Latest developments on the BioLung device have focused on using computational fluid dynamics (CFD) to optimize the device.^[42] A flow separator has been placed in the inner core of the fiber bundle and the axial length of the fiber bundle has been shortened, reducing the surface area from 2.1 m² to 1.6 m². A second outlet has also been added, which decreased the pressure drop across the device 10–40%. The newer design exhibited higher O₂ transfer efficiency (almost 300 ml/min/m² versus less than 250 ml/min/m² in the old design) in a 6-hour study of 6 adult male sheep at a flowrate of 6 l/min.^[42]

A paracorporeal total artificial lung for chronic respiratory support (Chronic Artificial Lung, or CAL) is under development at the University of Maryland as a continuation of earlier work at the University of Pittsburgh. The CAL is intended as a bridge-to-transplant device with the goal of 21-day support of basal metabolic needs using a device less than 0.5 m² in fiber membrane area.^[43] The CAL uses active mixing from a rapidly rotating disc made of microporous hollow fiber membranes that enhance gas exchange by increasing blood flow velocity past fiber surfaces

and reducing diffusional boundary layers. The disc rotates within a housing and the centrifugal motion imparted to blood enables the CAL to pump blood (which may reduce the impact of the CAL on the right heart in its intended in-series attachment mode). The motor controller directing disc rotation can generate pulsatile or nonpulsatile flow. The CAL generated 5 l/min flow against a 100 mm Hg pressure head at 1600 rpm during steady flow in bench tests using bovine blood, but adding pulsatility to the flow decreased pumping. Published data are lacking but the gas exchange efficiency of the CAL appears promising, with 550 ml/min/m² and 450 ml/min/m² reported for O₂ and CO₂ exchange efficiencies, respectively, in scaled-down prototypes.^[43]

OTHER EFFORTS IN ARTIFICIAL LUNGS

Several other artificial lung technologies are in various stages of development. The Integrated Heart–Lung Assist Device (IHLAD) combines the functions of a blood pump and artificial lung into one device with the goal of functioning as emergency percutaneous cardiopulmonary support.^[44] The IHLAD incorporates a central six-vane impeller driven by a magnetic coupling to an outer motor. The impeller is surrounded by gas-exchanging hollow fiber membranes. The blood-contacting surfaces are treated with covalently-bonded heparin to reduce thrombogenicity and allow pre-primed storage for emergency use. This IHLAD has a surface area of 0.85 m² and was evaluated in an in-vivo goat study with the device attached paracorporeally in a venoarterial bypass. Transfer rates of 180 ml/min O₂ and 110 ml/min CO₂ were achieved at 5 l/min of blood flow with sweep gas flows of three to five times the blood flowrate. The impeller enables pumping performance of 9 l/min output at 400 mm Hg pressure head at 3000 rpm. After 3 months of preprimed storage, the device was tested for gas exchange rates in an ex-vivo goat experiment. Whereas O₂ transfer rates were unaffected by storage, the CO₂ removal decreased approximately 30% after 3 months.

One artificial lung concept involves rotating an annular fiber bundle to promote active mixing and increased gas exchange, in addition to generating a pumping force.^[45] The annular fiber bundle is potted around a stationary central flow diffuser for distributing blood flow throughout the length of the rotating bundle. This is similar in concept to the CAL, in that the rotation of the fibers increases the relative velocity of blood flow past fiber surfaces to increase the efficiency of gas transfer, so less fiber membrane area is required for gas exchange, helping to reduce priming volume and adverse blood–material interactions. A prototype with a priming volume of 1250–2500 ml and

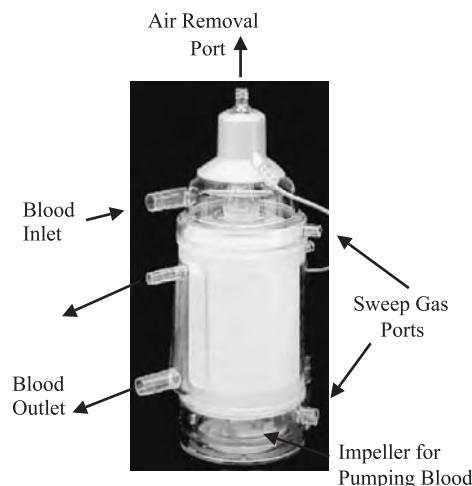


Fig. 9 The CORx™ integrated pump-oxygenator from cardiovention, Inc.

0.2 m² of membrane area exhibited a transfer rate of approximately 60 ml/min of O₂ at a rotational speed of 600 rpm.

The CORx™ System developed by Cardiovention in California consists of a hollow fiber membrane bundle similar to those used in typical commercial artificial lungs attached directly to an impeller pump (Fig. 9).^[46] The pump generates enough pressure to perfuse the artificial lung and return oxygenated and decarbonated blood back to the patient, thus achieving the primary gas exchange requirements and blood flowrates of standard cardiopulmonary bypass (CPB). The device also includes a special technology for removing air that inadvertently enters the circuit during the surgical procedure. This compact unit has a total blood-contacting surface area of less than 1.4 m², which reduces the blood surface contact area and minimizes hemodilution of the patient. The device has recently been approved by the FDA and is currently used for cardiac procedures at hospitals across the United States.^[46]

A novel photolytic artificial lung is under development that uses photosynthesis to convert CO₂ to O₂, obviating the need for O₂ to supply gas (as in membrane artificial lungs).^[47] The device consists of a blood pump, photolytic module of stacked cells, and a light source. The photolytic cells are created with vacuum-deposited titanium metal, a coating of photolytically sensitive TiO₂, and a layer of MnO₂ adhered to the surface of a glass substrate. In the photolytic modules, bicarbonate is protonated to form carbonic acid, which is converted to water and CO₂ in the presence of the carbonic anhydrase. The resultant water is converted to active oxygen at the TiO₂ catalyst when exposed to light, and the active oxygen is converted to dissolved oxygen at the MnO₂ catalyst.^[47] Many such photolytic

units will be integrated into a functional device to achieve relevant rates of O₂ production and CO₂ removal.

CONCLUSION

The next-generation artificial lungs described here derive directly in a conceptual sense from the hollow fiber membrane and membrane module technology used in standard clinical blood oxygenators. Many of these artificial lung devices will hopefully achieve some clinical success in the next five to ten years. On a more distant horizon, the future of artificial lungs may depart significantly from the use of purely synthetic devices based on existing hollow fiber membranes fabricated into modules resembling those used today. The natural lung is a remarkable gas exchange organ because of the microvascular scale of blood pathways combined with the intimate micron-scale juxtaposition of its blood and gas pathways (see the article titled “Lung, Artificial: Basic Principles and Current Applications” in this encyclopedia). Future artificial lung technologies are likely to be biohybrid artificial lungs, combining both synthetic and natural tissue components, into gas exchange units aimed at mimicking the scale and function of the alveolar-capillary units of the natural lung. Biohybrid artificial lungs of the future (Fig. 10) may involve micron-scale blood channels intimately contacting gas channels to provide not only for short diffusion paths and efficient gas exchange but to also create compact devices with surface area to blood volume ratios approaching those of the natural lung. The challenge of biocompatibility inherent in making these microvascular-scale blood channels, with an extensive blood contact area, non-thrombogenic and noninflammatory may require the

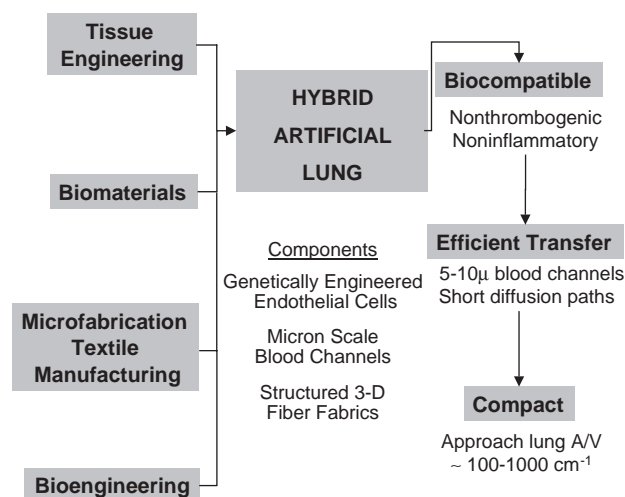


Fig. 10 The future of artificial lungs.

use of endothelial cells, perhaps genetically engineered for enhanced performance or for the robustness required in this application. Significant advances in tissue engineering, biomaterials, microfabrication, and bioengineering will all need to be harnessed for the technological development of these future artificial lungs. At the same time, the need for artificial lungs in the distant future may be eclipsed by significant advances in regenerative medicine that enable tissue repair and regeneration of the failing lung.

ACKNOWLEDGMENTS

Supported by HL70051 from The National Institutes of Health and by the U.S. Army Medical Research, Development, Acquisition and Logistics command under Grant No. DAMD17-98-1-8638. The views, opinions, and/or findings contained in this report are those of the authors and should not be construed as an official Department of the Army position, policy, or decision unless so designated by other documentation.

ARTICLES OF FURTHER INTEREST

Lung, Artificial: Basic Principles and Current Applications; Lung Surfactants

REFERENCES

- Deslauriers, J.; Awad, J.A. Is extracorporeal CO₂ removal an option in the treatment of adult respiratory distress syndrome? *Ann. Thorac. Surg.* **1997**, *64* (6), 1581–1582.
- Brower, R.G.; Matthay, M.A.; Morris, A.; Schoenfeld, D.; Thompson, B.T.; Wheeler, A. Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndrome. *New Engl. J. Med.* **2000**, *342* (18), 1301–1308.
- Brunston, R.L., Jr.; Zwischenberger, J.B.; Tao, W.; Cardenas, V.J., Jr.; Traber, D.L.; Bidani, A. Total arteriovenous CO₂ removal: Simplifying extracorporeal support for respiratory failure. *Ann. Thorac. Surg.* **1997**, *64* (6), 1599–1604.
- Tao, W.; Brunston, R.L., Jr.; Bidani, A.; Pirtle, P.; Dy, J.; Cardenas, V.J., Jr.; Traber, D.L.; Zwischenberger, J.B. Significant reduction in minute ventilation and peak inspiratory pressures with arteriovenous CO₂ removal during severe respiratory failure. *Crit Care Med.* **1997**, *25* (4), 689–695.
- Brunston, R.L., Jr.; Tao, W.; Bidani, A.; Traber, D.L.; Zwischenberger, J.B. Organ blood flow during arteriovenous carbon dioxide removal. *ASAIO J.* **1997**, *43* (5), M821–M824.

6. Brunston, R.L., Jr.; Tao, W.; Bidani, A.; Alpard, S.K.; Traber, D.L.; Zwischenberger, J.B. Prolonged hemodynamic stability during arteriovenous carbon dioxide removal for severe respiratory failure. *J. Thorac. Cardiovasc. Surg.* **1997**, *114* (6), 1107–1114.
7. Zwischenberger, J.B.; Alpard, S.K.; Tao, W.; Deyo, D.J.; Bidani, A. Percutaneous extracorporeal arteriovenous carbon dioxide removal improves survival in respiratory distress syndrome: A prospective randomized outcomes study in adult sheep. *J. Thorac. Cardiovasc. Surg.* **2001**, *121* (3), 542–551.
8. Zwischenberger, J.B.; Conrad, S.A.; Alpard, S.K.; Grier, L.R.; Bidani, A. Percutaneous extracorporeal arteriovenous CO₂ removal for severe respiratory failure. *Ann. Thorac. Surg.* **1999**, *68* (1), 181–187.
9. Conrad, S.A.; Zwischenberger, J.B.; Grier, L.R.; Alpard, S.K.; Bidani, A. Total extracorporeal arteriovenous carbon dioxide removal in acute respiratory failure: A phase I clinical study. *Intensive Care Med.* **2001**, *27* (8), 1340–1351.
10. Luzsa. *X-Ray Anatomy of the Vascular System*; J.B. Lippincott Company, 1974.
11. Tao, W.; Zwischenberger, J.B.; Nguyen, T.T.; Tzouanakis, A.E.; Matheis, E.J.; Traber, D.L.; Bidani, A. Performance of an intravenous gas exchanger (IVOX) in a venovenous bypass circuit. *Ann. Thorac. Surg.* **1994**, *57* (6), 1484–1490.
12. Conrad, S.A.; Bagley, A.; Bagley, B.; Schaap, R.N. Major findings from the clinical trials of the intravascular oxygenator. *Artif. Organs* **1994**, *18* (11), 846–863.
13. Conrad, S.A.; Zwischenberger, J.B.; Eggerstedt, J.M.; Bidani, A. In vivo gas transfer performance of the intravascular oxygenator in acute respiratory failure. *Artif. Organs* **1994**, *18* (11), 840–845.
14. Conrad, S.A.; Eggerstedt, J.M.; Morris, V.F.; Romero, M.D. Prolonged intracorporeal support of gas exchange with an intravenacaval oxygenator. *Chest* **1993**, *103* (1), 158–161.
15. Mortensen, J.D.; Berry, G. Conceptual and design features of a practical, clinically effective intravenous mechanical blood oxygen/carbon dioxide exchange device (IVOX). *Int. J. Artif. Organs* **1989**, *12* (6), 384–389.
16. Mortensen, J.D. Intravascular oxygenator: A new alternative method for augmenting blood gas transfer in patients with acute respiratory failure. *Artif. Organs* **1992**, *16* (1), 75–82.
17. Makarewicz, A.J.; Mockros, L.F.; Anderson, R.W. A pumping intravascular artificial lung with active mixing. *ASAIO J.* **1993**, *39* (3), M466–M469.
18. Vaslef, S.N.; Mockros, L.F.; Anderson, R.W. Intravascular Lung Assist Device and Method. US Patent 5,037,383, 8-6-1991.
19. Vaslef, S.N.; Mockros, L.F.; Anderson, R.W. Development of an intravascular lung assist device. *Trans. ASAIO* **1989**, *35*, 660–664.
20. Makarewicz, A.J.; Mockros, L.F.; Anderson, R.W. A dynamic intravascular artificial lung. *ASAIO J.* **1994**, *40* (3), M747–M750.
21. Aldea, G.S.; Doursounian, M.; O’Gara, P.; Treanor, P.; Shapira, O.M.; Lazar, H.L.; Shemin, R.J. Heparin-bonded circuits with a reduced anticoagulation protocol in primary CABG: A prospective, randomized study. *Ann. Thorac. Surg.* **1996**, *62*, 410–418.
22. Lynch, W.R.; Montoya, J.P.; Brant, D.O.; Schreiner, R.J.; Iannettoni, M.D.; Bartlett, R.H. Hemodynamic effect of a low-resistance artificial lung in series with the native lungs of sheep. *Ann. Thorac. Surg.* **2000**, *69* (2), 351–356.
23. Lick, S.D.; Zwischenberger, J.B.; Alpard, S.K.; Witt, S.A.; Deyo, D.M.; Merz, S.I. Development of an ambulatory artificial lung in an ovine survival model. *ASAIO J.* **2001**, *47* (5), 486–491.
24. Snider, M.T.; High, K.M.; Richard, R.B.; Panol, G.; Campbell, E.A.; Service, C.V.; Stene, J.K.; Ultman, J.S. Small intrapulmonary artery lung prototypes: Design, construction, and in vitro water testing. *ASAIO J.* **1994**, *40* (3), M533–M539.
25. High, K.M.; Nicholson, T.; Richard, R.B.; Panol, G.R.; Shelley, K.; Snider, M.T. Effects of blood phase oscillation on gas transfer in a microporous intravascular lung. *ASAIO J.* **1994**, *40* (3), M735–M739.
26. Nodelman, V.; Baskaran, H.; Ultman, J.S. Enhancement of O₂ and CO₂ transfer through microporous hollow fibers by pressure cycling. *Ann. Biomed. Eng.* **1998**, *26* (6), 1044–1054.
27. Body, S.C.; FitzGerald, D.; Voorhees, C.; Hansen, E.; Crowley, C.; Voorhees, M.E.; Sherman, S.K. Effect of nitric oxide upon gas transfer and structural integrity of a polypropylene membrane oxygenator. *ASAIO J.* **1999**, *45* (6), 550–554.
28. Conrad, S.A.; Brown, E.G.; Grier, L.R.; Baier, J.; Blount, J.; Heming, T.; Zwischenberger, J.B.; Bidani, A. Arteriovenous extracorporeal carbon dioxide removal: A mathematical model and experimental evaluation. *ASAIO J.* **1998**, *44* (4), 267–277.
29. De Somer, F.; Van Belleghem, Y.; Foubert, L.; Caes, F.; Francois, K.; Dubrulle, F.; Van Nooten, G. Feasibility of a pumpless extracorporeal respiratory assist device. *J. Heart Lung Transplant.* **1999**, *18* (10), 1014–1017.
30. Gross, B.D.; Sacristan, E.; Peura, R.A.; Shahnarian, A.; Devereaux, D.; Wang, H.L.; Fiddian-Green, R. Supplemental systemic oxygen support using an intestinal intraluminal membrane oxygenator. *Artif. Organs* **2000**, *24* (11), 864–869.
31. Keszler, M.; Moront, M.G.; Cox, C.; Milewski, M.; Visner, M.S. Oxygen delivery with a single cannula tidal flow venovenous system for extracorporeal membrane oxygenation. *ASAIO J.* **1995**, *41* (4), 850–854.
32. Fiore, G.B.; Costantino, M.L.; Fumero, R.; Montevicchi, F.M. The pumping oxygenator: Design criteria and first in vitro results. *Artif. Organs* **2000**, *24* (10), 797–807.
33. Watanabe, H.; Hayashi, J.; Ohzeki, H.; Moro, H.; Sugawara, M.; Eguchi, S. Biocompatibility of a silicone-coated polypropylene hollow fiber oxygenator in an in vitro model. *Ann. Thorac. Surg.* **1999**, *67* (5), 1315–1319.
34. Lund, L. Measurement of Hollow Fiber Membrane Permeance in a Gas–Liquid System Ph.D.; University of Pittsburgh, 2000.



35. Zwischenberger, J.B.; Anderson, C.M.; Cook, K.E.; Lick, S.D.; Mockros, L.F.; Bartlett, R.H. Development of an implantable artificial lung: Challenges and progress. *ASAIO J.* **2001**, *47* (4), 316–320.
36. Boschetti, F.; Perlman, C.E.; Cook, K.E.; Mockros, L.F. Hemodynamic effects of attachment modes and device design of a thoracic artificial lung. *ASAIO J.* **2000**, *46* (1), 42–48.
37. Cook, K.E.; Makarewicz, A.J.; Backer, C.L.; Mockros, L.F.; Przybylo, H.J.; Crawford, S.E.; Hernandez, J.M.; Leonard, R.J.; Mavroudis, C. Testing of an intrathoracic artificial lung in a pig model. *ASAIO J.* **1996**, *42* (5), M604–M609.
38. Vaslef, S.N.; Mockros, L.F.; Cook, K.E.; Leonard, R.J.; Sung, J.C.; Anderson, R.W. Computer-assisted design of an implantable, intrathoracic artificial lung. *Artif. Organs* **1994**, *18* (11), 813–817.
39. Vaslef, S.N.; Cook, K.E.; Leonard, R.J.; Mockros, L.F.; Anderson, R.W. Design and evaluation of a new, low pressure loss, implantable artificial lung. *ASAIO J.* **1994**, *40* (3), M522–M526.
40. Lick, S.D.; Zwischenberger, J.B.; Wang, D.; Deyo, D.J.; Alpard, S.K.; Chambers, S.D. Improved right heart function with a compliant inflow artificial lung in series with the pulmonary circulation. *Ann. Thorac. Surg.* **2001**, *72* (3), 899–904.
41. Zwischenberger, J.B.; Wang, D.; Lick, S.D.; Deyo, D.J.; Alpard, S.K.; Chambers, S.D. The paracorporeal artificial lung improves 5-day outcomes from lethal smoke/burn-induced acute respiratory distress syndrome in sheep. *Ann. Thorac. Surg.* **2002**, *74* (4), 1011–1016.
42. Chambers, S.D.; Merz, S.I.; McGillicuddy, J.W.; Bartlett, R.H. Development of the MC3 Biolung. In *IEEE, Proceedings of the Second Joint EMBS/BMES Conference*, Annual Fall Meeting of the Biomedical Engineering Society, Houston, TX, October 2002; Omnipress, 2002; 1581–1582.
43. Gartner, M.; Litwak, P.; Borovetz, H.; Griffith, B. Development of a Pumping Artificial Lung. In *IEEE, Proceedings of the Second Joint EMBS/BMES Conference*; Annual Fall Meeting of the Biomedical Engineering Society, Houston, TX, October 2002; Omnipress, 2002; 1589–1590.
44. Tatsumi, E.; Takano, H.; Taenaka, Y.; Nishimura, T.; Kakuta, Y.; Nakata, M.; Tsukiya, T.; Nishinaka, T. Development of an ultracompact integrated heart–lung assist device. *Artif. Organs* **1999**, *23* (6), 518–523.
45. Maloney, J., Jr.; Buckberg, G. Mass and Thermal Transfer Means for Use in Heart Lung Machines, Dialyzers, and Other Applications. US Patent 5,900,142, 5-4-1999.
46. <http://www.cardiovention.com> (accessed March 2003).
47. Dasse, K.A.; Monzyk, B.F.; Gilbert, R.J. Development of a Photolytic Artificial Lung (PAL). In *ASAIO Journal Abstracts New York, June 12, 2002*; Lippincott Williams & Wilkins: Baltimore, MD, 2002; 139.

Lung Surfactants

Robert H. Notter

*Department of Pediatrics, University of Rochester School of Medicine,
Department of Environmental Medicine, University of Rochester School of Medicine, and
Department of Biomedical Engineering, University of Rochester School of Medicine,
Rochester, New York, U.S.A.*

Zhengdong Wang

*Department of Pediatrics, University of Rochester School of Medicine, Rochester,
New York, U.S.A.*



INTRODUCTION

This article reviews the composition, activity, dysfunction, and clinical significance of endogenous and exogenous lung surfactants. Pulmonary surfactant, a mixture of lipids and proteins produced in the alveolar epithelial lining, is required for normal respiration. Surfactant reduces the work of breathing, stabilizes alveolar inflation and deflation to optimize gas exchange, and reduces the driving force for pulmonary edema. Surfactant acts on pulmonary pressure-volume mechanics through specific surface properties generated by molecular interactions between its lipid and protein components. Biophysically important surfactant components include dipalmitoyl phosphatidylcholine (DPPC), a mix of other saturated and unsaturated phospholipids, and three surfactant proteins (SP)-A, -B, and -C. Severe disease results if lung surfactant is deficient or becomes dysfunctional. Surfactant deficiency is the cause of respiratory distress syndrome (RDS) in premature infants, and surfactant dysfunction is an important contributor to clinical acute lung injury (ALI) and the acute respiratory distress syndrome (ARDS), which can affect patients of all ages. One approach to treating diseases of surfactant deficiency and dysfunction is to administer exogenous surfactants by airway instillation or other methodology. Understanding the biochemistry, biophysics, and physiology of lung surfactant materials is crucial for optimizing surfactant-based therapies for RDS and ALI/ARDS.

ENDOGENOUS SURFACTANT AND ITS NORMAL ACTIVITY IN LUNGS

Although surface tension is a small force in many macroscopic systems, it is highly important in the lungs. The majority of the extensive alveolar surface in mammals ($\sim 1 \text{ m}^2/\text{kg}$ body weight) is covered by a thin aqueous layer or hypophase. This hypophase is

contacted by inspired air, resulting in a large liquid–air interface with associated surface tension forces. The significant contribution of surface tension to the mechanics of breathing was demonstrated by von Neergaard in 1929,^[1] who showed that higher pressures were needed to inflate lungs or maintain them at fixed volume with air than with liquid (normal saline). This difference is due to surface tension forces, which are present along with tissue forces in air-filled lungs but not in liquid-filled lungs. Although von Neergaard's work documented the importance of pulmonary surface tension forces, the understanding that normal respiration depends on surface active agents (surfactants) to moderate these forces came later. The existence of lung surfactant and the linkage between its deficiency and respiratory distress syndrome (RDS or hyaline membrane disease) in premature infants were demonstrated in the 1950s.^[2–5]

Lung surfactant is synthesized, packaged, stored, secreted, and recycled by type II cells in the alveolar epithelial lining.^{[6–8]a} Secreted surfactant in the alveolar hypophase exists as a heterogeneous population of phospholipid-rich aggregates with incorporated apoproteins. These aggregates vary from tens of nanometers to several microns in size, with larger aggregates typically having the greatest surface activity and highest apoprotein content. Surfactant adsorbs from the hypophase to form a highly active interfacial film that lowers and varies alveolar surface tension during breathing. By lowering surface tension, surfactant reduces the nonflow component of respiratory work (increases lung compliance, $\Delta V/\Delta P$). The extent of

^aType II pneumocytes not only produce and recycle pulmonary surfactant, but also are stem cells for the alveolar epithelium and participate extensively in intercellular signaling during lung development, growth, inflammatory injury, and repair. Some lung surfactant components, particularly apoproteins, are additionally found in airway cells such as nonciliated bronchiolar epithelial cells (Clara cells). The present article focuses on the composition and activity of whole surfactant synthesized by alveolar type II pneumocytes.

surface tension lowering in any given alveolus is proportional to the concentration of surfactant molecules in the film. Surface tension is lowered to the greatest extent when the film is compressed as alveolar radius decreases during expiration, and becomes higher when alveolar radius increases during inspiration. This moderates pressure differences within and between different-sized alveoli based on the Laplace equation,^b stabilizing small airsacs against collapse (reducing atelectasis) and promoting more uniform inflation. Alveolar stability in vivo is further increased by specialized connective-tissue support fibers and by an interconnected network structure that allows airsacs sharing common septa to help each other resist collapse. However, active alveolar surfactant is crucial for normal respiration, and the consequences of its deficiency are strikingly apparent in premature infants with RDS who exhibit alveolar collapse and overdistension, decreased lung volumes and compliance, intrapulmonary shunting with impaired arterial oxygenation, and diffuse pulmonary edema. The importance of active surfactant is also apparent in patients with surfactant dysfunction who develop acute respiratory failure in association with ALI/ARDS.

Alveolar surfactant is a complex mixture containing approximately 85–90% phospholipids, 7–10% protein, and 4–7% neutral lipid by weight^[6,9,10] (Table 1). Phosphatidylcholine (PC) is by far the most common phospholipid class, composing about 80% of total surfactant phospholipid. DPPC (C16:0, C16:0) is the most prevalent molecular species, accounting for 40–50% of the PC fraction. Other disaturated molecules make up about 10–15% of PC, including isomers of palmitoyl-myristoyl PC (C16:0, C14:0) and a small amount of palmitoyl-stearoyl PC (C16:0, C18:0). Fluid unsaturated compounds (one or more double bonds) account for about 35–45% of PC, with the most abundant being palmitoyl-palmitoleoyl PC (C16:0, C16:1) and palmitoyl-oleoyl PC (C16:0, C18:1). Anionic phospholipids (phosphatidylglycerol, phosphatidylinositol, and phosphatidylserine) together compose about 15% of total surfactant phospholipid, and sphingomyelin and phosphatidylethanolamine make up the remaining 5%. As with the PC fraction, each of these phospholipid classes contains a mix of disaturated and unsaturated compounds. The protein content of lung surfactant is much smaller than its phospholipid content, but is crucial to surface

^bThe Laplace equation is given by $\Delta P = 2\sigma/R$, where ΔP is the pressure drop across a spherical interface of radius R and surface tension σ . A similar stability argument can be made for nonspherical alveoli from the more general Law of Young and Laplace, which incorporates the two principal radii of curvature R_1 and R_2 defining an interface of arbitrary shape: $\Delta P = \sigma(1/R_1 + 1/R_2)$.

Table 1 Approximate composition of endogenous pulmonary surfactant

<u>85–90% Phospholipids</u>	
80%	Phosphatidylcholine (PC)
40–50%	DPPC
10–15%	Other disaturated PCs
35–45%	Unsaturated PCs
15%	Anionic phospholipids (PG, PI, PS)
5%	Other phospholipid classes (PE, Sph)
<u>7–10% Apoproteins</u>	
	SP-A
	SP-B
	SP-C
	SP-D (not involved in biophysical function)
<u>4–7% Neutral lipids</u>	
	Cholesterol
	Cholesterol esters, glycerides

Values are representative averages in weight percent for surfactant lavaged (washed) from the lungs of normal animals of different species and ages. *Abbreviations:* PC: phosphatidylcholine; PG: phosphatidylglycerol; PI: phosphatidylinositol; PS: phosphatidylserine; PE: phosphatidylethanolamine; Sph: sphingomyelin; SP: surfactant protein. Adapted from Ref. [6].

active function. Three biophysically active proteins, SP-A, SP-B, and SP-C, are found in lung surfactant (Table 2). SP-D, a fourth protein not involved in biophysical function but important in host defense, is also present. Lung surfactant also contains about 4–7% neutral lipids, primarily cholesterol plus small amounts of cholesterol esters, diglycerides and triglycerides.

The physiological actions of lung surfactant depend on its ability to generate specific surface properties (Table 3). Secreted surfactant in the alveolar hypophase must be able to adsorb readily into a film at the air–water interface. In order to reduce the work of breathing and stabilize alveoli, this film must reduce surface tension to low values during dynamic compression and it must vary surface tension with interfacial area during cycling. Theory does not predict a unique level of surface tension lowering required for active lung surfactant. However, multiple studies have shown that films and dispersions of lavaged lung surfactant and related organic solvent extracts can lower surface tension to extremely low values <1 mN/m under dynamic compression at physiological rates at 37°C.^{[6,9]c} Lung

^cMinimum surface tension values of “0 mN/m” are sometimes reported for lung surfactant in the literature. However, since surface tension is necessarily finite at a stable air–water interface, a lower-limit minimum value of <1 mN/m is utilized here.

Table 2 Molecular characteristics of lung surfactant proteins

Surfactant protein (SP)	Selected structural and functional characteristics
SP-A	<p>MW 26–38 kD (monomer), 228 AA in length in humans.</p> <p>Most abundant surfactant apoprotein, relatively hydrophilic.</p> <p>Acidic glycoprotein with multiple post-translational isoforms.</p> <p>C-type lectin and member of the collectin family of host-defense proteins.</p> <p>Oligomerizes to an active octadecamer (six trimer subunits).</p> <p>Aggregates and orders lipids.</p> <p>Necessary for tubular myelin formation (along with SP-B, Ca⁺⁺).</p> <p>Helps regulate surfactant reuptake/recycling in type II cells.</p>
SP-B	<p>MW 8.5–9 kD (monomer), 79 AA in length in humans (active peptide).</p> <p>Overall hydrophobic in nature, with 2–3 amphipathic helical regions.</p> <p>Sequence includes multiple charged AA (e.g., 10 positive Arg/Lys residues).</p> <p>Forms dimers of presumptive functional significance in humans.</p> <p>Interacts biophysically with both the headgroups and chains of phospholipids.</p> <p>Necessary for tubular myelin formation (along with SP-A, Ca⁺⁺).</p> <p>Disrupts and fuses lipid bilayers and promotes lipid insertion into surface films.</p> <p>Enhances adsorption, film spreading, and dynamic surface activity of lipids.</p> <p>Most active SP in increasing adsorption and overall dynamic surface activity.</p>
SP-C	<p>MW 4.2 kD (monomer), 35 AA in length in humans (active peptide).</p> <p>Most hydrophobic surfactant protein.</p> <p>Forms dimers and other oligomers of uncertain functional significance.</p> <p>Human form has two palmitoylated cysteine residues.</p> <p>Contains only two charged amino acids at neutral pH (2 Arg/Lys residues).</p> <p>Primarily α-helical in structure, with a length that spans a lipid bilayer.</p> <p>Interacts biophysically primarily with phospholipid fatty chains.</p> <p>Disrupts and fuses lipid bilayers.</p> <p>Enhances adsorption, film spreading, and surface activity of lipids.</p>
SP-D	<p>MW 39–46 kD (monomer), 355 AA in length in humans.</p> <p>Has significant structural similarity to SP-A.</p> <p>Oligomerizes to a dodecamer (four trimer subunits).</p> <p>C-type lectin and member of collectin family of host-defense proteins.</p> <p>Not implicated in lung surfactant biophysics.</p> <p>May participate in surfactant metabolism in addition to host-defense.</p>

surfactant films must also respread effectively at the air–water interface during repetitive cycling, i.e., molecules ejected from the surface film during compression must reintegrate with remaining interfacial components during expansion. Respreading from film-associated collapse phases or related structures in the interfacial region, along with adsorption from the subphase, ensure that sufficient surfactant remains present in the film to lower surface tension effectively during successive breathing cycles.

METHODS FOR MEASURING THE SURFACE ACTIVITY OF LUNG SURFACTANTS

Details about interfacial and molecular biophysical methods used in studying lung surfactant films and bilayers are given in the text by Notter.^[6] Two widely used instruments for measuring the overall surface activity of lung surfactants are the pulsating bubble surfactometer^[11] and the captive bubble surfactometer.^[12] Measurements of surface tension on the

Table 3 Physiological Actions and Surface Properties of Functional Lung SurfactantPhysiological actions of functional lung surfactant

- Reduces the work of breathing (increases lung compliance).
- Increases alveolar stability against collapse during expiration (reduces atelectasis).
- Improves the uniformity of alveolar inflation during inspiration.
- Reduces the hydrostatic driving force for pulmonary edema.

Biophysical properties of functional lung surfactant

- Adsorbs rapidly to the air–water interface.
- Reaches very low minimum surface tensions during dynamic film compression.
- Varies surface tension with surface area during dynamic cycling.
- Respreads effectively at the interface during successive cycles.

pulsating bubble surfactometer reflect a physiologically relevant combination of adsorption, dynamic film compression, and film resreading behavior at 37°C. A tiny air bubble in a surfactant dispersion is pulsated at 20 cycles/min between maximum and minimum radii of 0.55 and 0.40 mm (50% area compression for truncated spheres), and surface tension is calculated from the Laplace equation and the measured interfacial pressure drop.^[11] The sample chamber of the pulsating bubble apparatus has a volume of only 40 microliters so that very small amounts of surfactant can be studied. Pulsating bubble data are typically expressed as minimum surface tension (surface tension at minimum radius) as a function of time of pulsation. Maximum surface tension (surface tension at maximum radius) is also sometimes reported. Adsorption can also be defined with this instrument in the static mode, although a diffusion resistance is present. The captive bubble surfactometer determines surface tension from its theoretical relation to pressure, volume, and dimensional parameters in a bubble captured against an agar layer and compressed or expanded by an external piston.^[12] This design rules out film migration effects that can influence data during initial cycles in the pulsating bubble surfactometer. The captive bubble surfactometer provides data on surface tension as a function of surface area, and adsorption can also be measured with subphase stirring to minimize diffusion resistance. However, the apparatus requires more complex data analysis based on computer solutions to interfacial equations in conjunction with bubble shape measurements from recorded video images.

The Langmuir–Wilhelmy surface balance is another important experimental technique used to study molecular interactions and resreading in cycled films of lung surfactant and its components.^[6] Films are most commonly spread directly at the air–water interface in the balance trough, which is made of Teflon[®] or some other hydrophobic material. A transverse dam or continuous ribbon barrier allows the film to be compressed and expanded while surface tension is measured from the force on a finely roughened Wilhelmy slide (e.g., sandblasted platinum) dipped into the interface. Data are typically reported as surface pressure (the amount that surface tension is lowered below that of the pure subphase) as a function of surface area or film concentration in square Angstroms per surfactant molecule at fixed temperature, cycling rate, and compression ratio. Surface pressure–area isotherms from the Wilhelmy balance directly assess resreading and molecular interactions in compressed surfactant films, although cycling rates for this instrument are much smaller than in the lungs. Adsorption can also be assessed on the Wilhelmy balance, but is much more commonly studied in a small Teflon[®] dish containing a subphase stirred with a magnetic bar to minimize the effects of diffusion. Surfactant is injected into the stirred subphase at time zero, and adsorption surface tension (or surface pressure) is measured as a function of time from the force on a hanging Wilhelmy slide dipped in the surface.^[6]

ROLES OF SPECIFIC COMPONENTS IN LUNG SURFACTANT ACTIVITY

Lung surfactant is extraordinary in its ability to adsorb at the air–water interface and to lower surface tension in a dynamically compressed film (see Refs. [6,9] for detailed review and literature citations). Surfactant lavaged from normal animal lungs rapidly adsorbs to equilibrium surface tensions of 22–23 mN/m at 37°C (equilibrium surface pressures of 47–48 mN/m) at the air–water interface. Similar rapid adsorption into a surface film is also exhibited by chloroform:methanol extracts of lavaged surfactant dispersed in the aqueous phase. As noted earlier, when such films of whole or extracted lung surfactant are compressed at rapid physiological rates, extremely low minimum surface tensions <1 mN/m are generated. These surface active behaviors result from molecular interactions between the lipid and protein components of lung surfactant. As an overview, DPPC and other rigid long-chain disaturated phospholipids are major contributors to the low minimum surface tensions found in dynamically compressed lung surfactant films. DPPC and other disaturated phospholipids are also essential for varying surface tension with area during cycling.

Fluid lung surfactant phospholipids greatly enhance film respreading during cycling, and improve adsorption relative to DPPC. Neutral lipids such as cholesterol also help to improve film respreading and adsorption relative to DPPC, but are detrimental to dynamic surface tension lowering if present in excess amounts. The most crucial components for increasing lung surfactant adsorption are SP-A, SP-B, and SP-C. Surfactant proteins are also highly important in refining the composition of the surface film during cycling to maximize dynamic surface tension lowering and respreading.

The molecular biophysical behavior of lung surfactant phospholipids and proteins has received extensive research study.^[6,9,10] Due to their high gel to liquid crystal transition temperatures and tight chain-packing characteristics, disaturated phospholipids such as DPPC form rigid interfacial films that are highly effective in lowering and varying surface tension under rapid compression at 37°C. However, rigid long-chain disaturated phospholipids do not respread effectively in cycled films, nor do they adsorb readily at the air–water interface. Unsaturated phospholipids pack less tightly than their disaturated counterparts in films and bilayers, and are in the liquid-crystal phase at 37°C. They do not lower surface tension as well as DPPC in cycled films, but exhibit much better respreading. Unsaturated phospholipids also adsorb more rapidly than DPPC (although not nearly as well as whole or extracted lung surfactant that contains apoproteins). SP-A, SP-B, and SP-C all have extensive molecular biophysical interactions with phospholipids that contribute to surface activity (Table 2). SP-A functions biophysically as a large octadecamer containing six trimeric subunits. SP-A in the presence of calcium increases the molecular order and aggregation of phospholipids. It is cooperative with SP-B in forming tubular myelin, which along with other large aggregate forms has been associated with rapid lung surfactant adsorption. SP-B and SP-C can disrupt and fuse phospholipid bilayers, and SP-B also facilitates the insertion of phospholipids into surface films. Multiple studies have demonstrated that SP-B is the most active of all the surfactant proteins in increasing the adsorption and overall dynamic surface activity of phospholipids (e.g., Refs. [13–17]). Its amphipathic structure allows strong interactions with both the headgroups and chains of phospholipid molecules in films and bilayers, making it a particularly important functional component of endogenous and exogenous lung surfactants.

DEFICIENCY AND DYSFUNCTION OF ENDOGENOUS LUNG SURFACTANT

The term “lung surfactant deficiency” implies a decreased amount of surfactant in the alveoli, whereas

the term “surfactant dysfunction” (also called surfactant inhibition or inactivation) generally means a decrease in surface activity. Lung surfactant deficiency and dysfunction are not independent, because surface activity is decreased if surfactant concentration is decreased. However, surfactant deficiency is usually associated with abnormalities in type II cell function, either from incomplete development as in premature infants or from cellular injury or alteration. Surfactant dysfunction, on the other hand, can arise from any biochemical or biophysical process that compromises the surface activity of alveolar surfactant, and associated abnormalities in type II cells may or may not be present. Both endogenous and exogenous lung surfactants are potentially subject to surfactant dysfunction or inactivation.

INHIBITOR-INDUCED LUNG SURFACTANT DYSFUNCTION

Multiple compounds and processes can detrimentally affect the surface activity of lung surfactants (see Refs. [6,9,18,19] for detailed review and literature citations). Inhibitors that act biophysically to inhibit lung surfactant activity include plasma and blood proteins such as albumin, fibrinogen, fibrin monomer, or hemoglobin; cellular lipids and lysophospholipids; fluid free fatty acids; meconium; and antibodies to surfactant proteins. Plasma and blood proteins reduce surface activity primarily by competitive adsorption and interfacial shielding. These large molecules contain polar and nonpolar amino acids and have an intrinsic surface activity. When they adsorb, they form a relatively impenetrable film that occupies the interface and limits the entry of lung surfactant components. Smaller molecules such as fluid free fatty acids and unsaturated cell membrane lipids, on the other hand, act primarily by mixing with lung surfactant components within the interfacial film to compromise surface tension lowering during dynamic compression. In addition, free fatty acids and lysophospholipids can also damage the cellular integrity of the alveolocapillary membrane and generate pulmonary edema containing plasma proteins and other inhibitors.

The ability of albumin, hemoglobin, or red blood cell membrane lipids to reduce the dynamic surface tension lowering ability of lavaged lung surfactant is illustrated in Table 4.^[20,21] The degree of inhibition is strongly dependent on surfactant concentration. At a low surfactant phospholipid concentration of 0.5 mM, minimum surface tension values above 20 mN/m are reached in the presence of the inhibitors after 10 minutes of cycling on a pulsating bubble surfactometer. However, doubling surfactant concentration to 1 mM phospholipid results in minimum surface



Table 4 Inhibitory effects of plasma proteins and red blood cell membrane lipids on the dynamic surface activity of endogenous lung surfactant

Mixtures	Phospholipid concentration (mM)	Inhibitor concentration (mg/ml)	Minimum surface tension (mN/m) at time (min)		
			0	5	10
Albumin	–	2–200	45	45	45
Hemoglobin	–	2–200	36	35	35
Lung Surfactant (LS)	0.5	–	20	6	<1
LS + albumin	0.5	10	45	44	29
LS + hemoglobin	0.5	25	34	29	26
LS + membrane lipids	0.5	0.4	24	22	21
LS	1	–	19	<1	<1
LS + albumin	1	100	44	3	<1
LS + hemoglobin	1	100	36	25	<1
LS + membrane lipids	1	0.4	22	20	19
LS + membrane lipids	2	0.4	20	–	<1

LS was obtained by bronchoalveolar lavage of intact calf lungs. Minimum surface tension values at 0, 5, and 10 min are on a pulsating bubble surfactometer (20 cycles/min, 50% area compression, 37°C) with a subphase of 0.15 M NaCl + 1.4 mM CaCl₂. Time 0 is < 15 sec from the start of bubble pulsation. Data from Refs. [20,21].

tensions of <1 mN/m even if blood protein concentrations are increased. When surfactant concentration is raised to 2 m M phospholipid, the inhibitory effects of red blood cell membrane lipids on dynamic surface activity are also abolished. Concentration-dependent inhibition patterns conceptually similar to those in Table 4 for dynamic surface activity are also found for lung surfactant adsorption in the presence of blood proteins.^[20,22,23] In addition to their effects as individuals, some inhibitory substances exhibit additivity in reducing surface activity. For example, mixtures of albumin or hemoglobin with lyso-PC or red blood cell membrane lipids can impair the surface tension lowering of lung surfactant more severely than the individual inhibitors alone.^[24] Additive inhibition, if present, is again most apparent at low surfactant concentrations and is mitigated as surfactant concentration is raised.^[24] The magnitude of additivity also depends on the mechanisms of inhibitor action involved. The effects of some inhibitors can actually be lessened in mixtures, as occurs when free fatty acids are bound by albumin when the two substances are present together.^[24]

Lung surfactant activity can also be impaired by additional pathways during pulmonary injury and inflammation. This includes exposure to chemically acting inhibitors such as phospholipases, proteases, and reactive oxygen/nitrogen species that are present in the alveoli and interstitium during inflammatory injury. Enzymes such as phospholipases and proteases

not only reduce the concentration of functional surfactant components by chemical degradation, but also produce reaction products that can cause further inhibition (e.g., phospholipases A₁ and A₂ generate inhibitory free fatty acids and lysophospholipids when they chemically degrade phospholipids). Reactive oxygen and nitrogen species that can chemically alter lung surfactant lipids and proteins during inflammatory injury include nitric oxide (\bullet NO), superoxide anion (O₂ \bullet^-), hydrogen peroxide (H₂O₂), hydroxyl radical (\bullet OH), peroxynitrite (ONOO⁻), nitrogen dioxide (NO₂), dinitrogen trioxide (N₂O₃), and S-nitrosothiols. In addition, another important mechanism by which surfactant activity can be impaired during lung injury involves the selective depletion or alteration of large surfactant aggregates or subtypes, which as noted earlier normally have better surface activity and higher apoprotein contents than small aggregates. Aggregate-related changes in surfactant activity can result from direct interactions with inhibitory substances, impaired intraalveolar aggregate processing, injury to type II cells, or a combination of these factors.

PHYSIOLOGICAL CORRELATES OF LUNG SURFACTANT ACTIVITY AND DYSFUNCTION

Decades of research have shown the importance of correlating data on lung surfactant composition and

Table 5 Animal models of lung surfactant deficiency or dysfunction**I. ANIMAL MODELS OF SURFACTANT DEFICIENCY****A. Models emphasizing pressure–volume (P–V) mechanics**

- Excised adult rat lungs depleted in surfactant by multiple lavage.
- Small premature fetal animals (e.g., 27-day gestation rabbits).

B. Models emphasizing pulmonary function as well as mechanics in vivo

- Premature large animals (e.g., lambs, baboons, monkeys).
- Adult animals such as rats or rabbits depleted in surfactant by in-vivo lung lavage (also used as a model of acute lung injury).

II. ANIMAL MODELS OF SURFACTANT DYSFUNCTION

Multiple models of acute lung injury in adult or newborn animals. Initiators of lung injury include: hyperoxia, instillation or injection of bacteria or bacterial endotoxin, aspiration of acid or meconium, bilateral vagotomy, oleic acid infusion, viral infection, infusion of antibodies to lung surfactant or lung tissue, administration of toxic chemicals such as N-Nitroso-N-methylurethane (NNNMU), pulmonary radiation, inhalation of toxic dusts or particles, and in-vivo lung lavage.

For reviews of the tabulated animal models, and their use in surfactant replacement applications, see Refs. [6,9,19].

surface activity with physiological effects in intact lungs. Animal models of surfactant deficiency and dysfunction are reviewed in detail elsewhere.^[6,9,19] As an overview, the most meaningful assessments of any lung surfactant material include studies in animal models and correlation of physiological findings with data on composition, surface activity, and inhibition resistance. Animal models can be categorized as primarily involving surfactant deficiency or surfactant dysfunction (Table 5). Although there is overlap, models of surfactant deficiency generally relate most directly to RDS, whereas models of surfactant dysfunction and lung injury are applicable for ARDS. Within each category, animal models can be subdivided based on their detailed characteristics. For example, some models of surfactant deficiency emphasize measurements of pressure–volume (P–V) mechanics in small premature animals, whereas others emphasize functional assessments of respiratory physiology (Table 5). These latter models generally require animals of sufficient age and/or size to allow detailed lung functional measurements during mechanical ventilation over a timescale of hours. Models of acute lung injury and surfactant dysfunction

are generally categorized by the mechanism of injury-induction (Table 5). Adult animals are most commonly used in lung injury models, although newborn animals can also be studied. Research in animal models of lung surfactant deficiency, along with interfacial biophysical studies, were essential in laying the basis for successful clinical exogenous surfactant therapy in premature infants with RDS. Research in animal models of acute lung injury with surfactant dysfunction is similarly important to extend surfactant therapy to ARDS-related lung injury. The consensus of current animal research indicates that lung surfactant materials with the greatest overall surface tension lowering ability and resistance to inhibitor-induced dysfunction have the highest physiological activity in reversing surfactant deficiency and dysfunction in intact lungs.

DISEASES OF LUNG SURFACTANT DEFICIENCY AND DYSFUNCTION

The major disease of lung surfactant deficiency is RDS or hyaline membrane disease in premature infants (for clinical details see standard texts of pathology and pediatric medicine). The rate of occurrence of RDS is very low in infants ≥ 35 weeks gestation (term = 40 weeks; prematurity = < 37 weeks), and increases as gestational age and birth weight decrease. The incidence of RDS is most significant in infants < 32 weeks gestation. Approximately 77,500 infants < 32 weeks, and 219,000 infants 32–35 weeks gestation, were born in the United States in the year 2000.^[25] There were over 24,000 reported cases of RDS, with an overall incidence of 6 per 1000 live births.^[25] Aside from the degree of prematurity, the risk of RDS is increased by maternal diabetes, male sex, caucasian race, and perinatal factors such as caesarian section. RDS is caused by surfactant deficiency due to type II cell immaturity, but superimposed lung injury and surfactant dysfunction can also enter its clinical course as the result of mechanical ventilation, hyperoxia, and diverse complications of prematurity. However, surfactant dysfunction is more prominent in the acute respiratory failure associated with clinical ALI/ARDS.

ARDS was originally described primarily in adults,^[26,27] and was commonly referred to as the “adult” respiratory distress syndrome. The now standard nomenclature of the “acute” respiratory distress syndrome reflects the pathophysiological origins of this syndrome as a severe acute pulmonary injury that can manifest in patients of any age. The American–European Consensus Committee in 1994 defined clinical ARDS based on acute onset, bilateral infiltrates demonstrated on frontal chest radiograph, a



ratio of arterial oxygen partial pressure to inspired fraction of oxygen ($\text{PaO}_2/\text{FiO}_2$ ratio) of ≤ 200 mmHg, and a pulmonary capillary wedge pressure of ≤ 18 mmHg (if measured) or no evidence of left atrial hypertension.^[28] The Consensus Committee defined clinical ALI identically, except for a less severe $\text{PaO}_2/\text{FiO}_2$ ratio of ≤ 300 mmHg.^[28] Thus, by definition, all patients with ARDS have ALI. Clinical ARDS affects 50,000 to 150,000 patients in the United States each year, and has a substantial mortality of 30–50% despite sophisticated mechanical ventilatory support and intensive medical care. Surfactant abnormalities have been well-documented in bronchoalveolar lavage from patients with ALI/ARDS in multiple studies.^[18,29]

An additional group of surfactant-related diseases involves a genetic deficiency in a functional surfactant component (particularly an apoprotein). Such conditions are much more rare than RDS or ARDS, affecting only one in hundreds of thousands of infants. The best known is hereditary SP-B deficiency, the major form of which is a homozygous mutation in exon 4 of the SP-B gene at codon 121, referred to as 121ins2.^[30,31] Heterozygous infants do not have overt respiratory pathology. Because of the crucial functionality of SP-B, the homozygous 121ins2 mutation is associated with decreased surfactant activity and lethal perinatal respiratory failure (an early alveolar proteinosis is also common).^[30,31] Reduced levels of surfactant phosphatidylglycerol and abnormalities in SP-C processing with accumulation of abnormal proSP-C are also present in homozygous infants.^[31–33] SP-B knockout mice similarly have lethal acute respiratory failure and compromised surfactant activity.^[34] There is genetic heterogeneity in human SP-B deficiency, and some variants allow survival into childhood.^[35] Congenital SP-B deficiency is not effectively treated with exogenous surfactant^[32] or extracorporeal membrane oxygenation.^[36] Lung transplantation is the only current option for infants with severe forms of this condition,^[37] and gene therapy is under investigation. In contrast to SP-B deficiency, a congenital deficiency of SP-C is not associated with severe acute respiratory disease in humans^[38] or transgenic mice.^[39]

CLINICAL EXOGENOUS SURFACTANT THERAPY

Clinical surfactant therapy is straightforward in concept: if endogenous surfactant is deficient or dysfunctional, then it can be supplemented by delivering an active exogenous surfactant to the alveoli. This therapy is by necessity acute and requires that the lungs ultimately develop or recover the ability to

produce active endogenous surfactant. Although simple in concept, developing effective exogenous surfactant therapy for premature infants with RDS required a substantial time. Initial treatment of premature infants with aerosolized DPPC in the 1960s was ineffective, both because DPPC alone is not an adequate lung surfactant substitute and because pulmonary delivery by aerosolization was inefficient. Decades of subsequent biophysical and animal research were required to clarify the molecular basis of lung surfactant function and define clinically effective exogenous surfactants. The utility of surfactant therapy to prevent or treat RDS in premature infants is now well-documented.^[6,18,40] Exogenous surfactants are dispersed in saline and delivered by intratracheal instillation through an endotracheal tube at doses of ~ 100 mg/kg body weight (divided dosing by tracheal or bronchoscopic instillation is typically used in older patients). Despite having theoretical advantages, aerosol methods for delivering exogenous surfactants are still not perfected.

Surfactant therapy for clinical ALI/ARDS is currently not as developed as is the case for RDS in premature infants. As described earlier, basic research has shown that many forms of surfactant dysfunction can be reversed or mitigated by raising surfactant concentration even in the continued presence of inhibitor substances (e.g., Table 4). Multiple ARDS-related animal models have also been shown to respond favorably to exogenous surfactants, including acute lung injury from hyperoxia, meconium aspiration, anti-lung serum administration, bacterial or endotoxin administration, vagotomy, in-vivo lavage, N-nitroso-N-methylurethane administration, and viral pneumonia.^[6,9,19] In addition, clinical studies have shown that exogenous surfactant therapy is efficacious in term infants with meconium aspiration or pneumonia, and improves lung function in children with ARDS-related acute respiratory failure (for citations see Refs. [6,18]). Surfactant therapy in adults with ARDS has had less success. By far the largest published trial of surfactant therapy in adults with ARDS reported no clinical benefits following treatment with aerosolized Exosurf.^[41] However, Exosurf is known to have low activity compared to animal-derived exogenous surfactants (see below), and aerosol delivery methods are not as effective as instillation. Many of the most active exogenous surfactant drugs have not been studied in detail in adults with ARDS. Moreover, the complex lung injury pathophysiology of ALI/ARDS includes inflammation, vascular dysfunction, and multiorgan disease in addition to surfactant dysfunction. Exogenous surfactant may thus need to be used in a combined-modality approach with other therapies that target multiple aspects of pathology in ARDS.^[6,18]

CLINICAL EXOGENOUS SURFACTANTS

Clinical exogenous surfactants currently used to treat surfactant-related lung disease are listed in Table 6. These preparations are divided into three functionally relevant groups: 1) organic solvent extracts of lavaged endogenous lung surfactant from animals; 2) organic solvent extracts of processed animal lung tissue with or without additional synthetic additives; and 3) synthetic preparations not containing surfactant material from animal lungs. Organic solvent extracts of lavaged alveolar surfactant in Category I in principle contain all of the hydrophobic lipid and protein components of endogenous surfactant, although compositional details vary somewhat depending on solvents and processing. Extracts of minced or homogenized lung tissue in Category II are noted separately because they contain a higher content of tissue-derived material and require more extensive processing that can alter the content of surfactant components. The synthetic clinical surfactants in Category III that have been most widely studied to date are Exosurf and ALEC (artificial lung expanding compound). Exosurf is a mixture of DPPC:hexadecanol:tyloxapol (1:0.11:0.075 by weight) and ALEC is a mixture of 7:3 DPPC:egg PG. Synthetic KL4 and recombinant SP-C surfactant are currently under clinical evaluation.

Table 6 Clinical exogenous surfactant drugs used to treat surfactant-related lung disease

I. Organic solvent extracts of lavaged animal lung surfactant

Infasurf (CLSE)

bLES

Alveofact

II. Supplemented or unsupplemented organic solvent extracts of processed animal lung tissue

Survanta

Surfactant-TA

Curosurf

III. Synthetic exogenous lung surfactants

Exosurf

ALEC

KL4

Recombinant SP-C surfactant

Infasurf (ONY, Inc. and Forest Laboratories), Survanta (Abbott/Ross Laboratories), and Curosurf (Chesi Farmaceutici and Dey Laboratories) are currently FDA-approved in the United States, and KL4 is under clinical evaluation. Exosurf (Glaxo-Wellcome) is also FDA-approved, but is no longer used. Additional synthetic surfactants are under development as described in the text. Table adapted from Ref. [6].

RELATIVE ACTIVITY AND INHIBITION RESISTANCE OF EXOGENOUS SURFACTANT DRUGS

The relative activity and efficacy of exogenous surfactant drugs are crucial for evaluating and optimizing therapy. Differences in efficacy between clinical surfactants have been demonstrated in comparison trials in premature infants and in retrospective meta analyses.^[6] Such studies indicate that natural surfactants from animal lungs (Categories I and II, Table 6) have significantly greater efficacy than the protein-free synthetic surfactants Exosurf and ALEC. The hydrophobic surfactant proteins are highly active, and substituting for them in synthetic surfactants is a challenging task. The surface and physiological activities of Exosurf, for example, are significantly increased by the addition of purified bovine SP-B/SP-C, demonstrating that its synthetic components do not adequately replace these active endogenous apoproteins.^[42] Animal-derived clinical surfactants also differ markedly in their surface activity and ability to resist inhibitor-induced dysfunction. Biophysical and animal research indicates that the surface activity, inhibition resistance, and physiological effects of extracts of lavaged animal surfactant (Category I, Table 6) are greater than those of other available clinical surfactants. For example, the surface and physiological activities of Infasurf are substantially greater than Survanta in basic research.^[17,42-44] These activity differences correlate directly with the content of SP-B in the two preparations.^[17,32,43] Survanta has a very low SP-B content by ELISA of only 0.044% by weight relative to phospholipid, while Infasurf has an SP-B content of 0.9% by weight that approaches that of lavaged whole surfactant.^[17] As described earlier, SP-B is the most active of the hydrophobic surfactant proteins in enhancing the adsorption and overall dynamic surface activity of phospholipids.^[13-17] The addition of SP-B or synthetic SP-B peptides to Survanta significantly improves its activity towards that of Infasurf and whole surfactant.^[17,43,45]

FUTURE EXOGENOUS SURFACTANT DEVELOPMENT

The high activity, low toxicity, and close analogy to alveolar surfactant of surfactants from animal lung lavage in Category I of Table 6 provide significant challenges for new drug development. However, synthetic surfactants manufactured under laboratory conditions in vitro have significant potential advantages in manufacturing, formulation, quality control, and availability over animal-derived preparations. In addition, the risk of transmitting infectious agents



such as prions is negligible for synthetic surfactants. Examples of conceptual approaches being followed for new exogenous surfactants include: 1) Combining human sequence recombinant apoproteins or related synthetic peptides with synthetic glycerophospholipids or other components; 2) modifying current clinical exogenous surfactants with added recombinant apoproteins or synthetic peptides to increase activity; and 3) developing new exogenous surfactants for use in inflammatory lung injury that have added beneficial properties based on novel synthetic components. KL4 and recombinant SP-C surfactant in Table 6 are examples of the first of these approaches. Examples of the second approach include adding human sequence SP-A or SP-A peptides to enhance inhibition resistance in exogenous surfactants, or adding hydrophobic surfactant apoproteins or peptides to Exosurf or Survanta as described above.^[17,42,43,45] An example of the third approach is the use of synthetic phospholipase-resistant diether phospholipids that have dynamic surface-tension lowering ability comparable to DPPC, but greatly improved adsorption and respreading properties (e.g., Ref. [46]). Surfactant preparations containing such phospholipids combined with purified surfactant proteins or synthetic peptides have a potential for very high overall surface activity plus the ability to resist degradation by phospholipases in ARDS-related lung injury. For more detailed review and discussion of on-going basic research on exogenous surfactants see Refs. [6,9].

SUMMARY AND CONCLUSIONS

This article has detailed the composition, activity, dysfunction, and replacement of pulmonary surfactant. Lung surfactant is a physiologically essential mixture of lipids and proteins required for normal respiration. Its actions include reducing the work of breathing (increasing lung compliance), stabilizing alveoli against collapse and overdistension, and decreasing the hydrostatic driving force for pulmonary edema. Lung surfactant achieves its physiological effects through specific surface active properties that result from the molecular biophysical interactions of its components. Biophysically important components of endogenous surfactant include DPPC and related long-chain disaturated phospholipids, a mix of fluid unsaturated phospholipids, and three specific surfactant proteins SP-A, SP-B, and SP-C. SP-D, a fourth surfactant protein, does not participate in surfactant biophysics but is active in host defense along with SP-A and potentially other surfactant components.

Potentially lethal respiratory disease occurs when lung surfactant is deficient or dysfunctional. Surfactant

deficiency is the cause of RDS in premature infants, and surfactant dysfunction is an important contributor to the pathophysiology of clinical lung injury (ALI and ARDS). Clinical surfactant therapy involves the delivery of exogenous surface active material to the alveoli to supplement deficient or dysfunctional endogenous surfactant. Life-saving exogenous surfactant therapy is now available to prevent and treat RDS in premature infants. Surfactant therapy is currently being optimized in premature infants and is being extended to patients of all ages with ALI and ARDS. Biophysical and animal research indicates that detriments to surface activity in lung-injured patients can be reversed or mitigated if active exogenous surfactants able to resist inhibition are delivered effectively to the alveoli. Exogenous surfactant therapy has been shown to have significant benefits in term infants with lung injury from meconium aspiration and infection, and also to improve respiration in children with ARDS. Current experience with surfactant therapy in adults with ARDS is mixed, but many of the most active clinical exogenous surfactants have not been extensively evaluated in these patients. Combination therapies using exogenous surfactant in conjunction with other interventions against inflammatory lung injury have also not received detailed study. Optimizing surfactant therapy depends on a rational integration of mechanism-based basic biophysical and animal research with focused, prospective, controlled clinical investigations.

ACKNOWLEDGMENT

The support of grant HL-56176 from the National Institutes of Health is gratefully acknowledged.

ARTICLES OF FURTHER INTEREST

Lung, Artificial: Basic Principles and Current Applications; Lung, Artificial: Current Research and Future Directions

REFERENCES

1. von Neergaard, K. Neue auffassungen uber einen grundbegriff der atemmechanik. Dieretraktionskraft der lunge, abhangig von der oberflachenspannung in den alveolen. *Z. Gesamte Exp. Med.* **1929**, *66*, 373–394.
2. Pattle, R.E. Properties, function, and origin of the alveolar lining layer. *Nature* **1955**, *175*, 1125–1126.
3. Clements, J.A. Surface tension of lung extracts. *Proc. Soc. Exp. Biol. Med.* **1957**, *95*, 170–172.

4. Gruenwald, P. The significance of pulmonary hyaline membranes in newborn infants. *J. Am. Med. Assoc.* **1958**, *166*, 621–623.
5. Avery, M.E.; Mead, J. Surface properties in relation to atelectasis and hyaline membrane disease. *Am. J. Dis. Child.* **1959**, *97*, 517–523.
6. Notter, R.H. *Lung Surfactants: Basic Science and Clinical Applications*; Marcel Dekker, Inc.: New York, 2000.
7. Batenburg, J.J. Surfactant phospholipids: Synthesis and storage. *Am. J. Physiol.* **1992**, *262*, L367–L385.
8. Rooney, S.A.; Young, S.L.; Mendelson, C.R. Molecular and cellular processing of lung surfactant. *FASEB J.* **1994**, *8*, 957–967.
9. Notter, R.H.; Wang, Z. Pulmonary surfactant: Physical chemistry, physiology and replacement. *Rev. Chem. Eng.* **1997**, *13*, 1–118.
10. Creuwels, L.A.J.M.; van Golde, L.M.G.; Haagsman, H.P. The pulmonary surfactant system: Biochemical and clinical aspects. *Lung* **1997**, *175*, 1–39.
11. Enhorning, G. Pulsating bubble technique for evaluation of pulmonary surfactant. *J. Appl. Physiol.* **1977**, *43*, 198–203.
12. Schurch, S.; Bachofen, H.; Goerke, J.; Possmayer, F. A captive bubble method reproduces the in situ behavior of lung surfactant monolayers. *J. Appl. Physiol.* **1989**, *67*, 2389–2396.
13. Wang, Z.; Gurel, O.; Baatz, J.E.; Notter, R.H. Differential activity and lack of synergy of lung surfactant proteins SP-B and SP-C in surface-active interactions with phospholipids. *J. Lipid Res.* **1996**, *37*, 1749–1760.
14. Yu, S.H.; Possmayer, F. Comparative studies on the biophysical activities of the low-molecular-weight hydrophobic proteins purified from bovine pulmonary surfactant. *Biochim. Biophys. Acta* **1988**, *961*, 337–350.
15. Oosterlaken-Dijksterhuis, M.A.; van Eijk, M.; van Golde, L.M.G.; Haagsman, H.P. Lipid mixing is mediated by the hydrophobic surfactant protein SP-B but not by SP-C. *Biochim. Biophys. Acta* **1992**, *1110*, 45–50.
16. Seeger, W.; Günther, A.; Thede, C. Differential sensitivity to fibrinogen inhibition of SP-C- vs. SP-B-based surfactants. *Am. J. Physiol.* **1992**, *261*, L286–L291.
17. Notter, R.H.; Wang, Z.; Egan, E.A.; Holm, B.A. Component-specific surface and physiological activity in bovine-derived lung surfactants. *Chem. Phys. Lipids* **2002**, *114*, 21–34.
18. Notter, R.H.; Apostolakos, M.; Holm, B.A.; Willson, D.; Wang, Z.; Finkelstein, J.N.; Hyde, R.W. Surfactant therapy and its potential use with other agents in term infants, children and adults with acute lung injury. *Perspect. Neonatol.* **2000**, *1* (4), 4–20.
19. Lewis, J.F.; Jobe, A.H. Surfactant and the adult respiratory distress syndrome. *Am. Rev. Respir. Dis.* **1993**, *147*, 218–233.
20. Holm, B.A.; Notter, R.H.; Finkelstein, J.H. Surface property changes from interactions of albumin with natural lung surfactant and extracted lung lipids. *Chem. Phys. Lipids* **1985**, *38*, 287–298.
21. Holm, B.A.; Notter, R.H. Effects of hemoglobin and cell membrane lipids on pulmonary surfactant activity. *J. Appl. Physiol.* **1987**, *63*, 1434–1442.
22. Holm, B.A.; Enhorning, G.; Notter, R.H. A biophysical mechanism by which plasma proteins inhibit lung surfactant activity. *Chem. Phys. Lipids* **1988**, *49*, 49–55.
23. Holm, B.A.; Wang, Z.; Notter, R.H. Multiple mechanisms of lung surfactant inhibition. *Pediatr. Res.* **1999**, *46*, 85–93.
24. Wang, Z.; Notter, R.H. Additivity of protein and non-protein inhibitors of lung surfactant activity. *Am. J. Respir. Crit. Care Med.* **1998**, *158*, 28–35.
25. Martin, J.A.; Hamilton, B.E.; Ventura, S.J.; Menacker, F.; Park, M.M. Births: Final data for 2000. *Natl. Vital Stat. Rep.* **2002**, *50*, 18. Tables 43 and 48.
26. Ashbaugh, D.G.; Bigelow, D.B.; Petty, T.L.; Levine, B.E. Acute respiratory distress in adults. *Lancet* **1967**, *2*, 319–323.
27. Petty, T.L.; Ashbaugh, D.G. The adult respiratory distress syndrome. Clinical features, factors influencing prognosis and principles of management. *Chest* **1971**, *60*, 233–239.
28. Bernard, G.R.; Artigas, A.; Brigham, K.L.; Carlet, J.; Falke, K.; Hudson, L.; Lamy, M.; Legall, J.R.; Morris, A.; Spragg, R. The American-European Consensus Conference on ARDS: Definitions, mechanisms, relevant outcomes, and clinical trial coordination. *Am. J. Respir. Crit. Care Med.* **1994**, *149*, 818–824.
29. Griese, M. Pulmonary surfactant in health and human lung diseases: State of the art. *Eur. Respir. J.* **1999**, *13*, 1455–1476.
30. Noguee, L.M.; Wert, S.E.; Proffitt, S.A.; Whitsett, J.A. Allelic heterogeneity in hereditary surfactant protein B (SP-B) deficiency. *Am. J. Respir. Crit. Care Med.* **2000**, *161*, 973–981.
31. Noguee, L.M.; Garnier, G.; Dietz, H.C.; Singer, L.; Murphy, A.M.; deMello, D.E.; Colten, H.R. A mutation in the surfactant protein B gene responsible for fatal neonatal respiratory disease in multiple kindreds. *J. Clin. Invest.* **1994**, *93*, 1860–1863.
32. Hamvas, A.; Cole, F.S.; deMello, D.E.; Moxley, M.; Whitsett, J.A.; Colten, H.R.; Noguee, L.M. Surfactant protein B deficiency: Antenatal diagnosis and prospective treatment with surfactant replacement. *J. Pediatr.* **1994**, *125*, 356–361.
33. Vorbroker, D.K.; Proffitt, S.A.; Noguee, L.M.; Whitsett, J.A. Aberrant processing of surfactant protein C in hereditary SP-B deficiency. *Am. J. Physiol.* **1995**, *268*, L647–L656.
34. Clark, J.C.; Wert, S.B.; Bachurski, C.J.; Stahlman, M.T.; Stripp, B.R.; Weaver, T.E.; Whitsett, J.A. Targeted disruption of the surfactant protein B gene disrupts surfactant homeostasis, causing respiratory failure in newborn mice. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 7794–7798.
35. Dunbar, A.E.; Wert, S.E.; Ikegami, M.; Whitsett, J.A.; Hamvas, A.; White, F.V.; Piedboeuf, B.; Jobin, C.; Guttentag, S.; Noguee, L.M. Prolonged survival in hereditary surfactant protein B (SP-B) deficiency associated with a novel splicing mutation. *Pediatr. Res.* **2000**, *48*, 275–282.
36. Moulton, S.L.; Krous, H.F.; Merritt, T.A.; Odell, R.M.; Gangitano, E.; Cornish, J.D. Congenital pulmonary alveolar proteinosis: Failure of treatment with



- extracorporeal life support. *J. Pediatr.* **1992**, *120*, 297–302.
37. Hamvas, A.; Noguee, L.M.; Mallory, G.B.; Spray, T.L.; Huddleston, C.B.; August, A.; Dehner, L.P.; deMello, D.E.; Moxley, M.; Nelson, R.; Cole, F.S.; Colten, H.R. Lung transplantation for treatment of infants with surfactant protein B deficiency. *J. Pediatr.* **1997**, *130*, 231–239.
 38. Noguee, L.M.; Dunbar, A.E.; Wert, S.E.; Askin, F.; Hamvas, A.; Whitsett, J.A. A mutation in the surfactant protein C gene associated with familial interstitial lung disease. *N. Engl. J. Med.* **2001**, *344*, 573–579.
 39. Glasser, S.W.; Burhans, M.S.; Korfhagen, T.R.; Na, C.-L.; Sly, P.D.; Ross, G.F.; Ikegami, M.; Whitsett, J.A. Altered stability of pulmonary surfactant in SP-C deficient mice. *Proc. Natl. Acad. Sci.* **2001**, *98*, 6366–6371.
 40. Jobe, A.H. Pulmonary surfactant therapy. *N. Engl. J. Med.* **1993**, *328*, 861–868.
 41. Anzueto, A.; Baughman, R.P.; Guntupalli, K.K.; Weg, J.G.; Wiedemann, H.P.; Raventos, A.A.; Lemaire, F.; Long, W.; Zaccardelli, D.S.; Pattishall, E.N.; Exosurf ARDS Sepsis Study Group Aerosolized surfactant in adults with sepsis-induced acute respiratory distress syndrome. *N. Engl. J. Med.* **1996**, *334*, 1417–1421.
 42. Hall, S.B.; Venkitaraman, A.R.; Whitsett, J.A.; Holm, B.A.; Notter, R.H. Importance of hydrophobic apoproteins as constituents of clinical exogenous surfactants. *Am. Rev. Respir. Dis.* **1992**, *145*, 24–30.
 43. Mizuno, K.; Ikegami, M.; Chen, C.-M.; Ueda, T.; Jobe, A.H. Surfactant protein-B supplementation improves in vivo function of a modified natural surfactant. *Pediatr. Res.* **1995**, *37*, 271–276.
 44. Seeger, W.; Grube, C.; Günther, A.; Schmidt, R. Surfactant inhibition by plasma proteins: Differential sensitivity of various surfactant preparations. *Eur. Respir. J.* **1993**, *6*, 971–977.
 45. Walther, F.J.; Hernandez-Juviel, J.; Bruni, R.; Waring, A. Spiking Survanta with synthetic surfactant peptides improves oxygenation in surfactant-deficient rats. *Am. J. Respir. Crit. Care Med.* **1997**, *156*, 855–861.
 46. Turcotte, J.G.; Lin, W.H.; Pivarnik, P.E.; Sacco, A.M.; Bermel, M.S.; Lu, Z.; Notter, R.H. Chemical synthesis and surface activity of lung surfactant phospholipid analogs. II. Racemic N-substituted diether phospholipids. *Biochim. Biophys. Acta* **1991**, *1084*, 1–12.

Magnetic and Electrophoretic Cell Separation

S. Chethana

Ganapathi Patil

K. S. M. S. Raghavarao

Department of Food Engineering, Central Food Technological Research Institute, Mysore, Karnataka, India



INTRODUCTION

Improved techniques for separating cells are increasingly important to biomedical science and biotechnology because separation is frequently a limiting factor in downstream processing. Other downstream processing contributes a large portion of the product cost.^[1] Magnetic and electrophoretic cell sorting is finding wide application due to multidisciplinary interactive research in areas such as physics, material science, physical and polymer chemistry, biology, medicine and engineering (electrical, electronics, and biomedical).

Many cell separation/sorting methods have been reviewed in the literature.^[2] Magnetic and electrophoretic techniques have the potential to be alternative or often complimentary to other methods currently used in biotechnology. In this article, new and improved separation methods are discussed with brief mention of existing methods.

Multistage methods, namely ADSEP (ADvanced SEPARATOR), MAGSEP (MAGnetic SEPARATOR) and ELECSEP (ELECTrophoretic SEPARATOR), for quantitatively separating cells are given a special emphasis. These methods are expected to cost much less in comparison with the competing technologies at similar scale and precision and are built on the counter-current distribution technology of classical chemical engineering separations. The theory behind field-assisted separations provide the opportunity to predict the performance of these extractors and also to improve their design. Hence, mathematical models are stressed in this chapter. We also present a few suggestions for future work while analyzing the scale-up and economic aspects of these separation methods.

MAGNETIC CELL SEPARATION

Magnetic separation methods are used in a variety of scientific applications, such as endocrinology, growth regulation, virology, carcinogenesis, infectious diseases, neurology, and nutrition. Magnetic separation

based on magnetic sorbents, carriers, and modifiers has been employed for isolation, modification, detection, immobilization, and removal of a variety of cells. The developments that aroused fresh technical interest in magnetic separation of cells are high magnetic field gradients, superconducting magnets, and newer magnetic labeling techniques.

Existing Methods

Magnetic separation methods are mainly classified into two types: 1) magnetic separation of separands that are intrinsically magnetic and 2) rendering one or more nonmagnetic components of a mixture magnetic by the attachment of a magnetically responsive entity, which can be manipulated using an external magnetic field. The cells exhibit magnetic properties, namely paramagnetism, ferromagnetism, and diamagnetism. The magnetoactive aquatic bacteria that are separated by high gradient magnetic separations (HGMSs) contain small single-domain crystals of magnetite, and the ability to be oriented in a magnetic field remains even when the organism dies. Most of the applications of HGMS of paramagnetic material is pertains to red blood cells. There are other ways to induce paramagnetism. For instance, large positive susceptibilities can be introduced into biological particles by their adsorption of magnetic cations. Ferromagnetic properties can be imparted to biological matter supplementing the solution with small magnetite particles. An improvement of this technique by employing ferrofluids is reviewed by Rosensweig.^[3] Diamagnetic materials can be separated by using a strongly paramagnetic background fluid, which increases the contrast. White blood cells and platelets have been extracted by this method although in general the range of concentration of the carrier medium is limited by the osmotic pressure effects on the cell membrane.

Two possible modes, direct and indirect, are available for the isolation of target cells. In the direct mode, the magnetic particles with immobilized affinity ligands are directly suspended in the solution or suspension. The target cells bind with the immobilized ligand

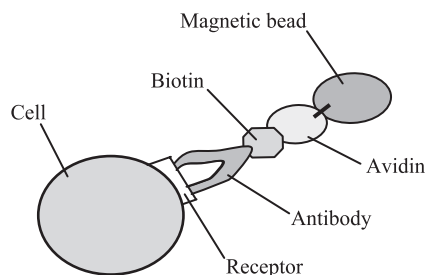


Fig. 1 Magnetic bead attached to cell receptor.

during the course of incubation. The whole complex can then be separated using a magnet. In the indirect mode, the target cells interact with the affinity ligand, usually a primary antibody that has first been added to the solution or suspension in the free form, which forms a complex during incubation. Magnetic particles with immobilized secondary antibodies are then added. Finally the magnetic complex formed is separated in a magnetic separator. This technique has been used for the isolation of various microbial cells from culture media^[4] and also used for the isolation of prokaryotic and eukaryotic cells.^[5-7]

The method of cell separation using a magnetic field has been implemented as a binary separation between cells that have and have not bound magnetic microspheres on the basis of a cell receptor (a specific type of surface ligand), as shown in Fig. 1. Many other applications in different biorelated areas are discussed by Safarik and Safarikova.^[8]

Different methods available for separating cells other than magnetic separation are sedimentation, flow cytometry, electrophoresis, and differential adsorption.^[9] From close observation of all the cell separation methods, magnetic separation is easy to implement and can be made quantitative based on the amount of the microspheres bound.

Multistage Magnetic Method

The development of user-friendly devices that are capable of separating particles according to quantity of ligand seems to be the area of greatest need in improving magnetically assisted separation devices. The

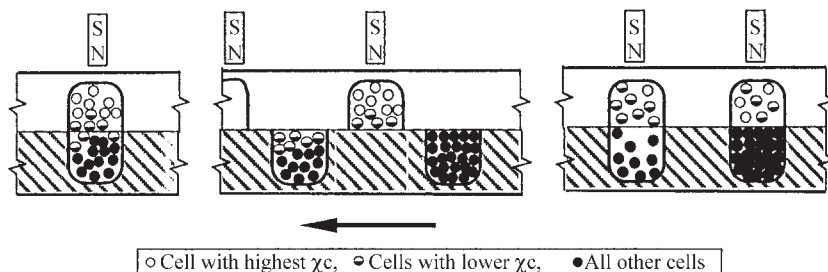


Fig. 2 One transfer in the multistage electromagnetic separator process.

magnetic separation industry has made considerable progress in this regard but the commercial technology to date has been limited to binary separation methods. Combination of magnetic extraction with counter-current extraction has made the magnetic separation of cells possible, as both quantitative and qualitative technique. It is quantitative with respect to the amount of ligand bound to cells of the same type and qualitative with respect to the amount of ligand bound to each kind of cell. The innovation presented here is magnetic separation (MAGSEP), a reliable method for differential magnetic separation of cells on the basis of small differences in surface composition.

In MAGSEP, capture could be isocratic or gradient mode. In the gradient mode (as shown in Fig. 2), the first stage has no magnet and no upper cavity and serves the purpose of homogenizing the cell mixture by stirring just before the beginning of transfers. The second stage also has no magnet and serves the purpose of adding magnetic particles to the cell suspension, mixing them together and allowing them to react. The third stage has a very weak magnet in the upper cavity and attracts only the highly magnetized cells, namely those with the most receptors for bonding with the magnetic microspheres. The fourth stage has a stronger magnet than the third stage in its upper compartment and attracts more weakly magnetized cells, and so on until, at the last-but-one stage, the strongest magnet of all captures the cells with the fewest receptors. The final stage has no magnet and will contain remaining unmagnetized cells after the final transfer. In the presence of gravity, uncaptured cells settle into the lower cavities by gravitational sedimentation if the transfer times are made sufficiently long. In the absence of gravity, uncaptured cells would remain in both the upper and lower cavities at each transfer. However, continued mixing with each transfer would have the effect of removing the uncaptured cells in each cavity. This method can separate both particulate (cells) and soluble (proteins) separands.

A comprehensive mass-balance model of multistage separation has been developed. Figure 3 is a schematic representation of a multistage electromagnetic separator showing comparison with a hypothetical magnetic chromatography column. The MAGSEP device

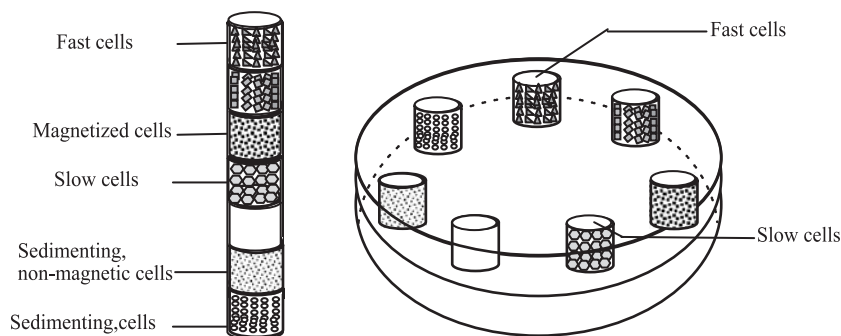


Fig. 3 Multistage electromagnetic separator compared with a magnetic chromatography column.

utilizes a step-wise rotary distribution and containment system, which selects, isolates, and stores particles of different magnetophoretic mobilities. Thus, cells are separated according to the quantity of ligand on their surface.^[10]

Mathematical Models

The range of magnetic field required for cell separation is obviously decided by the magnetic and mechanical properties of the cells. A review of the mathematical formulae that are relevant to cell extraction and their illustration by selected examples is presented well by Zboroswki.^[11]

The magnetic field is a space domain where an electrical charge, e , moving with velocity v_e experiences a magnetic force, F_m , given by

$$F_m = IlB \quad (1)$$

where I is the current intensity (A), l is length of element (m), B is the magnetic field intensity (kg A⁻¹s⁻² or Tesla, T). The presence of matter in the magnetic field modifies the field fluxes at given constant field strengths. The magnetic properties of matter are defined by induced polarization M to account for the variation in magnetic flux. For isotropic media

$$M = \chi_m H \quad (2)$$

where χ_m is magnetic susceptibility and H is the magnetic field strength (Am⁻¹). In the uniform magnetic field the magnetic particle undergoes rotation until Maxwell stress tensor becomes zero and then remains stationary with respect to the medium. In the nonuniform field, differences in the Maxwell stresses result in a net force, F_m , acting on the magnetic particle given as

$$F_m = V_p(M \cdot \Delta)B \quad \text{or} \quad F_m = V_p \chi_m H \frac{dB}{dz} \quad (3)$$

where V_p is the volume of the particle. In homogenous media, magnetic field B is parallel and proportional to

the magnetic field strength H . Magnetic force lines are defined as curves that are tangent to the direction of vector F_m . In general, due to nonlinear relationship between the magnetic force F_m and the magnetic field H , the lines of magnetic force do not follow the lines of magnetic field B . This is drastically different from the static electric field, in which electric force lines are synonymous with electric field lines.

Now the work required to bring all the components of the magnetic system from infinity to their given spatial position is defined as magnetostatic potential energy, U_m , which can be given as

$$U_m = V_p \chi_m u_m = -\frac{V_p \chi_m B^2}{2\mu_0} \quad (4)$$

where ' u_m ' is magnetic energy density, given by $-B^2/2\mu_0$ and μ_0 is the magnetic permeability of free space. The relationship between the force and the potential energy leads to the following expression of the magnetic force:

$$F_m = -\nabla U_m = -V_p \chi_m \nabla \mu_m = \frac{V_p \chi_m \nabla B^2}{2\mu_0} \quad (5)$$

The direction of the magnetic force F_m relative to the energy density gradient ∇B^2 depends on the sign of χ_m . For paramagnetic substances, $\chi_m > 0$ and the force vector points toward the direction of the maximum increase in magnetic field energy density, termed magnetic attraction. For diamagnetic substances, $\chi_m < 0$ and the force vector points in the opposite direction to the maximum increase in the field energy density, termed as magnetic repulsion. When the particles are suspended in a medium of magnetic susceptibility χ_m , then in the above expression χ_m is to be replaced by $\Delta\chi_m$ (magnetic susceptibility difference). Consequently the above discussion applies to $\Delta\chi_m$ rather than χ_m . Thus the very basis for all magnetic cell extractors is the observation that forces acting on small magnetic particles follow the lines of gradient of the magnetic field energy density.



Model for Viscous Medium

A simple mathematical model is developed based on the theory discussed so far in this article, for the motion of magnetically labeled cells in a viscous medium in the presence of magnetic field.^[11] The following assumptions are involved: 1) cells are small compared to the characteristic magnetic field and fluid flow dimension; 2) cells are treated as magnetized rigid spheres while calculating magnetic and fluid drag forces; 3) specific density of the cell is similar to that of the medium; and 4) inertial forces are too small to be considered when compared with the viscous ones. With these assumptions the equation of motion [involving the forces magnetic (F_m), drag (F_d), and buoyancy (F_b)] is written as

$$F_m + F_d + F_b = 0 \quad (6)$$

where

$$F_m = \frac{(\Delta\chi_m)V_{\text{cell}}\nabla B^2}{2\mu_0} \quad (7)$$

$$F_d = 6\pi\eta a_{\text{cell}}(v_{\text{cell}}) \quad (8)$$

$$F_b = V_{\text{cell}}(\Delta\rho)g \quad (9)$$

when the circulation in the medium is neglected (v_{medium}) due to high viscosity. It may be noted that $V_{\text{cell}} = 4\pi a_{\text{cell}}^3/3$ (volume of cell, m^3) where a_{cell} is cell radius (m) and η is viscosity (kg ms^{-1}), $\Delta\chi_m = (\chi_{\text{cell}} - \chi_{\text{medium}})$ with magnetic susceptibility is replaced by volume average cell susceptibility, and $\Delta\rho = (\rho_{\text{cell}} - \rho_{\text{medium}})$ where χ_{cell} is magnetic susceptibility of cell, χ_{medium} is magnetic susceptibility of medium, ρ_{cell} is density of the cell (kg m^{-3}), and ρ_{medium} is density of the medium (kg m^{-3}). On solving for v_{cell} (velocity of the cell, m/s)

$$v_{\text{cell}} = -\frac{(\Delta\chi_m)a_{\text{cell}}^2\nabla B^2}{9\mu_0\eta} - \frac{2(\Delta\rho)a_{\text{cell}}^2g}{9\eta} \quad (10)$$

In order to obtain the flux J , both sides are multiplied by concentration of cells C :

$$J = -\frac{C(\Delta\chi_m)a_{\text{cell}}^2\nabla B^2}{9\mu_0\eta} = C v_{\text{cell}} + \frac{2(\Delta\rho)a_{\text{cell}}^2g}{9\eta} C \quad (11)$$

This equation states that the flux of magnetically labeled cells relative to the medium follows the lines of magnetic energy density gradient. Here the mass flux of small magnetic cells/particles is coupled to the driving force of the magnetic energy density gradient. As the

magnetic particles are assumed to be rigid bodies of finite volume such that convective, gravity, and buoyancy forces can be neglected, the magnetic particles are expected to accumulate, forming layers around the magnetic surfaces. Further, the surfaces of the magnetic layers follow the surfaces of constant magnetic energy density or B^2 is constant.

This model of Zborowski^[11] could be easily adapted for MAGSEP. Additional assumptions made at this point include: 1) All particle motion is vertical; 2) the drag force is negligible except for in the vertical direction; 3) the axial magnetic field is constant over the radius of the cylindrical cavity of the MAGSEP; 4) creeping flow conditions (low Reynolds number); and 5) particle velocity is constant and is equal to the mean velocity of migration ($d^2z/dt^2 = 0$).

Magnets with large cross-sections (relative to the cavity cross-section) are assumed and, as a result, the magnetic field was considered to be constant over the radial cross-section ($\partial B_z/\partial r = 0$). By taking the dot product of the magnetic field and the gradient of the magnetic field, the following differential relationship is determined:

$$B \cdot \nabla B = B_z \cdot \frac{\partial B_z}{\partial r} + B_r \cdot \frac{\partial B_r}{\partial z} \quad (12)$$

where B_r is magnetic field component in r direction ($\text{kg A}^{-1}\text{s}^{-2}$ or Tesla, T), B_z is magnetic field component in z direction ($\text{kg A}^{-1}\text{s}^{-2}$ or Tesla, T), $\partial B_z/\partial r$ is rate of change of the magnetic field along the vertical axis r and $\partial B_r/\partial z$ is rate of change of the magnetic field along the vertical axis z .

This relationship, combined with the above assumptions to simplify the magnetic force equation, yields the scalar magnetic force relationship similar to that given in Eq. 7. If all variable parameters for the MAGSEP experiment in this equation (B and particle size radius, a_p) are set, then the susceptibility difference, $\Delta\chi_m$, is the controlling variable governing the magnetic force. Based on this calculation, the separation of the cells as it relates to the susceptibility difference can be governed by controlling the magnetic field, B , the vertical height of separation, Δz , and the time interval required for separation, Δt .

The volumetric susceptibility difference $\Delta\chi_m$ can be measured by sampling a distribution of particle positions. If the particles are placed in a thin flat layer across the bottom of a cavity and covered with a solution to a depth Δz , the particles can then be aligned below a shallow upper cavity with an activated magnet. The magnetic field of the upper magnet can then attract particles for a time Δt , and when a sufficient amount of particles have migrated into the shallow upper cavity, the cavities can be half-stepped to separate the particles as shown in Fig. 2. The volumetric

susceptibility can then be calculated by the following relationship:

$$\Delta\chi_m = \frac{9v_p\eta/a_p^2 + 2g(\Delta\rho)}{B \cdot \frac{\partial B}{\partial z}} \quad (13)$$

where v_p is the velocity (m/s) and can be assumed to be equal to $\Delta z/\Delta t$. The number of particles transferred represents a distribution of the total particles and is also a representation of the number of particles within a volumetric susceptibility range. If enough separation in these stages takes place, then this volumetric susceptibility range can be assumed to be an absolute volumetric susceptibility. Continuous flow magnetic cell sorting using soluble immunomagnetic label and the corresponding theory was presented by Zborowski et al.^[12]

ELECTROPHORETIC SEPARATION

Electrophoresis is a leading method for resolving mixtures of charged cells/macromolecules. Electrophoretic separations are influenced by many factors including the size (or molecular weight), shape, secondary structure, and charge. These features can influence electrophoretic properties either separately or jointly. In order to achieve the required scale-up, scientists and engineers resort to flowing methods.^[13] Scale-up of electrophoresis is hindered by Ohmic heating, which causes free convection/mixing within the system. Excess heat can also denature the labile cells.

Existing Methods

Free electrophoresis could enjoy more widespread use, because it is a high-resolution separation method that does not require adsorption to solid media and the subsequent solids handling. Furthermore, it can handle the cells/particles as well as solutes/macromolecules alike. Some of the applications of free electrophoresis include the separation of different cells of peripheral blood and bone marrow in hematological and immunological research with therapeutic applications. In spite of its positive applications, free electrophoresis has never gained popularity as a preparative or as an industrial separation method owing to gravity-dependent characteristics such as thermal convection, electroosmosis, particle sedimentation, particle aggregation, and electrohydrodynamic zone distortion. None of these principal gravity-dependent and gravity-independent processes have been investigated in multistage electrophoresis, which is designed to minimize the impact of these processes on separation quality.

Another type of mixing problem encountered in free electrophoresis, although less critical compared to Ohmic heating, is the mixing caused by gas release at the electrodes. This problem is addressed by employing nongassing electrodes or membrane-separated electrodes.^[14]

A score of methods have been developed to effect free electrophoresis.^[9] These methods can be broadly divided into static and flowing methods, neither of which has satisfactory capacity for application as a manufacturing tool. Batch and continuous methods have also been developed. In almost all cases, maximum sample input rates have been of the order of a few millilitre per hour. In one important case, Tulp et al.^[15] designed a reorienting free electrophoresis device consisting of a flat disk-shaped container with thin sample bands and a short migration distance. The top and bottom electrode fluids served as coolant and the total height of the separation column was 1–2 cm, and its diameter was greater than 15 cm. The distance between unrelated separands was about 1–2 mm, and this distance was increased during fractionation after electrophoresis by reorienting the disk.

Based on research experience an attempt has been made by using a multistage electrophoretic method to overcome the major bottlenecks of free electrophoresis in various fields of bioprocessing and analysis such as virology, endocrinology, enzymology, hematology, and mammalian cell culture.

Multistage Electrophoretic Method

The multistage electrophoretic method has been developed by combining free electrophoresis and multistage extraction and is explored as an improved alternate method for the purification and concentration of cells or macromolecules. The present discussion is restricted to cells. The design is based on a derivative of the thin-layer multistage extractor design of Albertsson^[16] and Treffrey et al.^[17] and called ADSEP, designed by SHOT, Inc.,^[18] which is shown in Fig. 4. The extractor consists of a 22-cavity, multistage, thin-layer extraction system.^[17] Half-cavities oppose each other in disks that are sealed together and rotate with respect to each other. The half cavities are disk shaped, and top cavities have flat tops while bottom cavities have flat bottoms. Both contain palladium (Pd) metal electrodes that produce an electric field when the two cavities are in phase with each other. This extractor, like the Tulp device,^[15] takes advantage of short column height to maintain isothermal conditions. Each cavity is only a few mm in height so that the fluid within it remains isothermal during the application of an electric field that transfers the separand particles or cells from the bottom to the

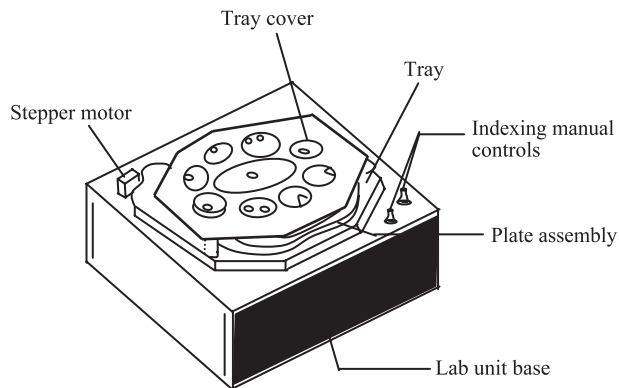


Fig. 4 ADvanced SEPARation (ADSEP) apparatus.

top cavity. As each separand is transferred to a new cavity, it is either drawn into the upper cavity by the electric field or left in the lower cavity, depending on its electrophoretic mobility. ADSEP is driven by an independent power supply (Lamda, Model # LP-532-FM) for rotating the plates to bring the chambers into interfacial contact with each other. This ADSEP is modified to use as ELECSEP by replacing the chamber bottoms with metal cover plates. Electrodes are kept over these cover plates with gaskets in between them.

A combination of free electrophoresis and multistage extraction is explored as an improved and alternative method for the extraction of cells. A simple mathematical model is presented to describe the mass and heat transfer during the electrophoretic separation process in an electrophoretic countercurrent extractor (ECCE). In order to validate the model and also to test the extractor, the extraction of fixed human red blood cells and latex particles of different sizes suspended in phosphate buffer is performed at different electric field

strengths. The experimental results are found to agree reasonably well with those predicted by the model.^[19]

It is observed that the cell transport velocity increases proportionally to the applied electric field strength under otherwise similar conditions (Eq. 14), and is also noted that when the field strength is doubled the cell extraction is completed in much fewer transfers.^[19]

Mathematical Models

In an ECCE apparatus, the separation of bioparticles having different electrophoretic mobilities is achieved by contacting the buffer solutions of top chambers with bottom chambers containing the bioparticles in buffer or fermentation broth, and then applying an electric field (E) at regular, predetermined intervals.

Mass transfer

Electrophoretic extraction of cells is a rate process (not an equilibrium process) and the particles having higher electrophoretic mobility are separated ahead of those having relatively lower mobility. The physical description of the multistage extraction is shown in Fig. 5. An ECCE apparatus has n extraction stages. Let us consider a situation in which all particles of electrophoretic mobility, μ_E , are initially in one (first) bottom chamber. In the first extraction step, the top and bottom chambers of stage 1 are brought into interfacial contact with each other. Then the field of preselected intensity (E) is switched on. Cells having negative mobility move to the top chambers (having an anode). After applying the field for a predetermined time period (τ), the field is switched off. The bottom cavity is moved to a position to be in contact with a

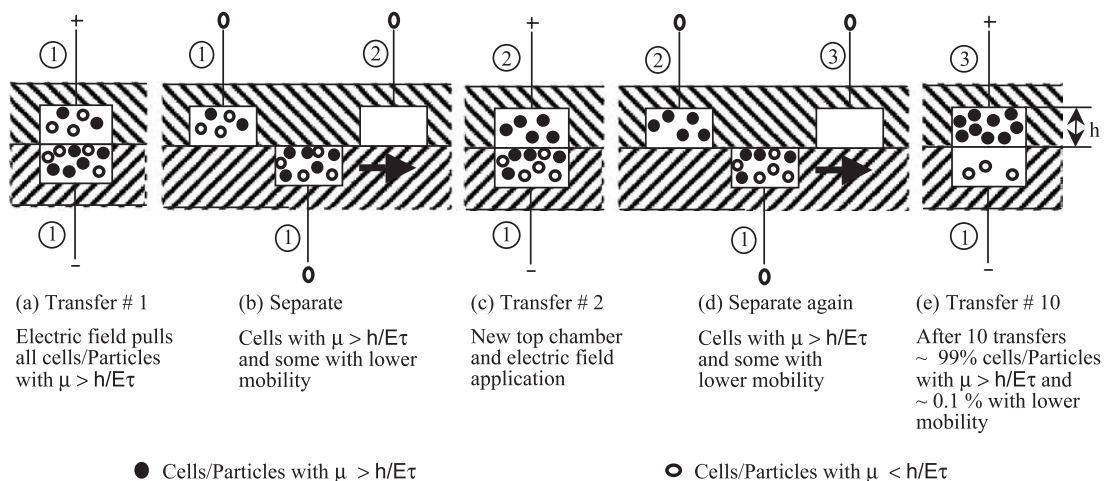


Fig. 5 Physical description of multistage extraction of cells/particles.

new top cavity having buffer. This process is repeated as many times as necessary to achieve the desired separation. Even when one type of particle is used, the particles usually have an approximately normal electrophoretic mobility distribution and not a single value. However, the average mobility can be estimated under actual conditions without the drawbacks of the conventional methods.

The quantity of bioparticles initially ($t=0$) present in the first bottom chamber is denoted by N . Now let us consider one chamber, whose total depth is H_t and radius is R with the bioparticles suspended in buffer solution filling the chamber. When a vertical electric field is applied, the bioparticles move upward as a slug due to their electrophoretic mobility. Their velocity will be proportional to the applied field. In other words

$$dy/dt \propto E \quad \text{or} \quad dy/dt = \mu_E E \quad (14)$$

where the proportionality constant, μ_E , is electrophoretic mobility and the characteristic of the bioparticle and its magnitude is decided by the surface charge of the bioparticles.

Integrating the above equation between the limits $y=0$ to h and $t=0$ to τ where τ is the time of application of the electric field results in

$$h = \mu_E E \tau \quad (15)$$

so that a bioparticle of mobility μ_E moves a distance h in an electric field of intensity E applied for a time of τ . It is obvious that if E is increased, τ can be decreased to achieve the same distance of migration and vice-versa, under otherwise similar conditions.

The bioparticles are randomly distributed in space over the volume of the bottom chamber. However, the time required for an individual particle to move into the top chamber increases as its distance from the top surface of the bottom chamber increases, under a given set of experimental conditions of E and τ ; that is, the ratio of these heights (h/H_t) gives the relative number of bioparticles that migrate to the top chamber at any given set of E and τ . To calculate the absolute number of particles transferred to the top chamber, the ratio of the distance to the top of the chamber to the total height H_t has to be multiplied by the concentration of the particles (i.e., number of particles per unit volume), since the ratio of the heights is nothing but the ratio of the volumes of the chamber corresponding to the location considered.

Therefore, if it were desired to capture the particles with mobility μ_E located at distance h from the top surface of bottom chamber, the number of bioparticles that would migrate during a single step is

$$m = (h/H_t)(N) = (\mu_E E \tau / H_t)(N) \quad (16)$$

where $N = (c)(\pi R^2 H_t)$, c is the cell concentration (cells/ml), $\pi R^2 H_t$ is the volume of the chamber, and $h = \mu_E E \tau$ (from Eq. 15). The particles move as a slug and in each transfer step a slice of slug of height h (containing m number of cells/particles) is extracted into the top chamber. In each transfer the slug of cells/particles move by the same distance and hence the same number of cells/particles is extracted in each transfer, until cells are exhausted.

The change in the number of particles in a stage n during step r will be equal to the number of bioparticles that migrated to the top chamber during step r . Thus the general equation can be written for this situation, by material balance, for number of particles extracted in each transfer step.

$$-(N)_{n,r} + (N)_{n,r-1} = (m)_{n,r} \quad (17)$$

where $(m)_{n,r}$ is the number of particles with mobility μ_E that migrated from stage n during step r to the top chamber N is the total number of bioparticles present in any of the n bottom chambers at $t=0$.

Now using Eq. 16, the above Eq. 17 can be written as

$$(N)_{n,r} = -(\mu_E E \tau / H_t)(N)_{n,r-1} + (N)_{n,r-1} \quad (18)$$

This is a general equation that enables us to estimate the number of bioparticles having mobility μ_E in any stage n , provided their concentration is known in the previous stage. Assuming that the bioparticles are uniformly distributed in the chamber and they move in a plug flow under the influence of an applied electric field and hence the same number of particles migrate to the top chamber in each transfer step (which are contained by the slug of height h of Eq. 16), the following material balance equation can be written as

$$(N)_{n,r} = -(r)(\mu_E E \tau / H_t)(N) + (N) \quad (19)$$

where r is the step number and N is the initial concentration of the bioparticles of mobility μ_E . From this generalized equation, the number of bioparticles in a given stage can be calculated directly from the number of transfers, r , instead of step-by-step sequence, especially at low field strengths.

Heat transfer

A major problem in scale-up of electrokinetic processes is known to be heating, which in turn causes mixing. However, the major advantage of the multistage process is the speed at which it performs separations. Hence the design is modified (such as provision for proper heat transfer/removal) in order to reduce the



adverse effects of heating, while scaling up the process based on the basic laws of heat generation and transmission. Electrical energy is dissipated as heat according to the following equation, which obeys Ohm's law

$$W = I^2/A^2K_E \quad (20)$$

where W is the power density (watts/cm³), I is the current (amperes), A is the area (cm²) over which the field is applied, and K_E is the electrical conductivity of the medium (mho/cm or Siemen/cm). It may be noted that the heat generation increases as the square of the current passed. For this reason, nearly all electrokinetic applications are in the most resistant media compatible with the unit operation, meaning that low-conductance solutions must be used to have low current in order to operate for longer periods of time.

In an adiabatic stagnant system, the temperature gradually increases uniformly with time as

$$\Delta T = \frac{IE_\tau}{C_p\rho} = \frac{I^2\tau}{K_E A^2 C_p \rho} \quad (21)$$

where τ is the time of application of electric field, C_p is the specific heat of the media carrying the current, and ρ is the density of the system (gm/cm³). The above equation represents the maximum (adiabatic) temperature increase in the electrophoretic system.

SCALE-UP AND ECONOMIC ASPECTS

Recent developments such as new methods of generating high magnetic field gradients and availability of reasonably priced efficient and superconducting magnets enable the use of magnetic extraction technology even on a large scale.^[1] Batch HGMS machines are in commercial use and continuous separators are also developed.^[1] Some of the following are already commercialized or have a very high possibility of being commercialized in the near future. Commercially introduced by Dynal A/S (Oslo, Norway) and Miltenyi Biotech (Bergish Gladbach, Germany), immunomagnetic cell separation has become an established method for cell analysis in clinical diagnostics. Its low price makes it an alternative to flow cytometry and very handy for rare cellular events.

The main advantage of magnetic separation here is to eliminate the need for centrifugation with its many associated disadvantages.^[19] Further magnetic separations are cheaper, faster, allow many samples to be handled simultaneously, and help to lend the process towards automation. Ease of conjugation of magnetic immunomicrospheres having the required size with

antibodies paved the way for large-scale immunological cell sorting.^[1] Performance data in simple and complex matrices for the process-scale cell separation with the help of bioreceptor ferrofluids and HGMS were reported.^[20]

On analyzing the available information, it appears that electroextraction may pose more scale-up problems when compared to magnetic extraction for cells. However, the actual situation varies depending on the application and each method has its own advantages and disadvantages. For instance, the advantages of using the magnetic rather than electric field are no need for medium modifications, no biological effects in the range of practical static magnetic field intensities, and no joule heating and therefore no flow distortion or need for a complex cooling system.^[21] On the other hand, the disadvantages are its complex interaction with paramagnetic and ferromagnetic labels, the additional cost of research and development, and its dependency on chemical reagents, antibodies, and magnetic colloids, while all other methods are based on cellular immunoreactivity.^[12]

Practically no reports are available in the literature to date regarding economic aspects of magnetic and electroextractions. However, we are optimistic about the prospects for the electrophoretic methods in processing of cells and macromolecules due to two main factors. As a source of energy due to low currents involved (only mA), electric fields are not expensive. As a force, electric fields are tunable to a very fine degree and involve no moving parts. Small-scale applications are already making a major impact on analytical and preparative separations.

OTHER APPLICATIONS

Magnetic extraction must be thoroughly explored for its application in other areas since it exhibits high specificity and easy scale-up. Therefore, it finds applications in various fields such as food processing, wastewater treatment, medicine parasitology, and clinical and environmental microbiology. Some possible applications are cited here, which will prompt researchers to look for similar applications.

Casein hydrolysis was carried out using chymotrypsin immobilized on magnetic supports.^[22] Magnetic separation is used for the detection of pathogenic microorganisms such as *Staphylococcus aureus*, *Salmonella enteritidis*, *Listeria monocytogenes*, and *Escherichia coli*.^[23]

Separation, immunoassy, magnetic resonance imaging (MRI), drug delivery, and hyperthermia are enhanced by the use of magnetic particles in medical applications and could be used as carriers to introduce DNA into cells. They can also be used as heat

mediators for cancer therapy. Hence, the use of magnetic particles with their unique feature will further improve medical techniques.^[24] Another potential application of magnetic particle separation technology is the use of magnetically delivered therapeutics. Carrier holding the drug can be concentrated at the desired site *in vivo* by a magnetic field.

Electrophoretic extraction can be explored in aqueous two-phase systems, because unlike the magnetic extraction method, it has influence on the partition behavior of cells, as well as on phase demixing rates.^[21] Similarly, electrophoretic methods are finding increasing applications in other areas. For instance, electrophoretic processes are established techniques for the dewatering of fine clays and latexes and also for enhancing liquid–liquid extraction. Adaptation of this method for the extraction of biochemicals was reported.^[25] Greater clarification and faster floc formation were claimed using electrocoagulation in comparison with conventional chemical methods.

SUGGESTIONS FOR FUTURE WORK

Only a few attempts at proportional magnetic particle separation have been made. The technical shortfall is related to the distance over which field gradients must be maintained. Using the preexisting ADSEP technology, these distances can be kept short, the apparatus is manageable, and the separation can be of adequate resolution. Although the technical feasibility of MAGSEP has been established, more research must be carried out to deliver the economically feasible design of multistage magnetic extractor.

There is a need for magnetic extraction of cells on the basis of receptor density. The relatively small sample of recent findings clearly indicates that tools for studying cells with modified receptor densities would be welcome. Identification of applications is one of the major aspects that need attention in the immediate future. A need for proportional (vs. binary) magnetic separation has to be identified in the areas of 1) immunological research, 2) pharmaceutical delivery, and 3) biomedical applications.

Our research work in electrophoretic extraction has confirmed the technical feasibility of multistage extraction of cells using electric fields.^[19] However, a good amount of research needs to be done before bringing ADSEP to the market as ELECSEP. For example, immediate study is required to examine the efficacy of the unit for the fractionation of a mixture of cells. Preliminary experiments have indicated that the resolution is hampered by the depth of the chamber. Thus in future designs, the depth of the chamber should be reduced while the diameter/cross-sectional area of the chambers should be increased.

CONCLUSIONS

Some of the new separation technologies that have considerable promise in the near future for the effective extraction of cells/particles are presented in this article. It is not difficult to appreciate the fact that separation technology is one of the most complex and important areas of biotechnology. Some successful applications of magnetic and electrophoretic extractions are commercialized.^[5] Magnetic and electrophoretic separation methods are undoubtedly complementary to other methods that are in use at present in the area of biotechnology and have the potential even to be alternatives in specific situations where other conventional methods normally tend to fail. Further, these advanced separation methods often result in relatively faster and more selective separation of the target cells. It may be noted that the scope of possible applications of magnetic and electrophoretic separation methods is very wide and obviously not limited to only those discussed here in this article. More and more new applications will be identified and studied in the near future.

An interdisciplinary effort involving a combination of physical, chemical, biomedical, and engineering aspects is essential for developing effective field-assisted separation methods. Some of these have received enough attention, but, the engineering aspects have been largely neglected. Mathematical modeling, which is highlighted in this article, will be of immense use in predicting the amount of cells extracted without the measurement of an inordinately large number of parameters and, hence, due importance should be given to this aspect. Some of the important aspects on which future research efforts need to be focused are suggested in this article. Even if some of these aspects are addressed in greater depth by future researchers, the objective of this article can be considered fulfilled in view of the recognized scientific and industrial potential of magnetic and electrophoretic separation methods.

ACKNOWLEDGMENTS

The authors are thankful to Dr. V. Prakash, Director CFTRI Mysore, India, for the encouragement and keen interest in field-assisted separations. One of the authors, S. Chethana, acknowledges CSIR Govt. of India for a Senior Research Fellowship.

ARTICLES OF FURTHER INTEREST

Flow Cytometry; Magnetic and Electrophoretic Cell Separation; Magnetic Nanoparticles: Structure and



Bioapplications; Micro-Particle Image Velocimetry in Biomedical Applications

REFERENCES

1. Setchell, C.H. Magnetic separations in biotechnology. *J. Chem. Technol. Biotechnol.* **1985**, *35B*, 175.
2. Todd, P.; Pretlow, T.G. The Separation of Living Cells. In *Cell Separation Science and Technology*; Kompala, D.S., Todd, P., Eds.; ACS Symposium Series, American Chemical Society: Washington, 1991; 1.
3. Rosensweig, R.E. Magnetic fluids. *Sci. Am.* **1982**, *247*, 124.
4. Patchett, R.A.; Kelly, A.F.; Kroll, R.G. The absorption of bacteria to immobilized lectins. *J. Appl. Bacteriol.* **1991**, *71*, 271.
5. Olsvik, O.; Popovic, T.; Skjerve, E. Magnetic separation techniques in diagnostic microbiology. *Clin. Microbiol. Rev.* **1994**, *7*, 43.
6. Safarik, I.; Safarikova, M.; Forsythe, S.J. The application of magnetic separation in applied microbiology. *J. Appl. Bacteriol.* **1995**, *78*, 575.
7. Larsson, P.O.; Mosbach, K. Alcohol production by magnetically immobilized yeast. *Biotechnol. Lett.* **1979**, *12* (1), 501.
8. Safarik, I.; Safarikova, M. Scientific and Clinical Applications of Magnetic Carriers. In *Overview of Magnetic Separations Used in Biochemical and Biotechnological Applications*; Hafeli, U., Schutt, W., Teller, J., Zborowski, M., Eds.; Plenum Press: New York, 1997; 323.
9. Todd, P. The Separation of Living Cells. In *Cell Separation Science and Technology*; Kompala, D.S., Todd, P., Eds.; ACS Symposium Series, American Chemical Society: Washington, 1991; 216.
10. Todd, P.; Cooper, R.P.; Doyle, J.F.; Dunn, S.; Vellinger, J.; Deuser, M.S. Multistage magnetic particle separator. *J. Magn. Magn. Mater.* **2001**, *225*, 294.
11. Zborowski, M. Physics of Magnetic Cell Sorting. In *Scientific and Clinical Applications of Magnetic Carriers*; Hafeli, U., Schutt, W., Teller, J., Zborowski, M., Eds.; Plenum Press: New York, 1997; 205.
12. Zborowski, M.; Moore, L.R.; Sun, L.; Chalamers, J.J. Continuous Flow Magnetic Cell Sorting Using Soluble Immunomagnetic Label. In *Scientific and Clinical Applications of Magnetic Carriers*; Hafeli, U., Schutt, W., Teller, J., Zborowski, M., Eds.; Plenum Press: New York, 1997; 247.
13. Ivory, C.F. The prospects for large-scale electrophoresis. *Sep. Sci. Technol.* **1988**, *23*, 875.
14. Cole, K.D.; Todd, P.; Srinivasan, K.; Dutta, B.K. Free-resolution electrophoresis of proteins in an improved density gradient column and by capillary electrophoresis. *J. Chromatogr., A* **1995**, *A707*, 77.
15. Tulp, A.; Timmerman, A.; Bornhorn, M.G. *Electrophoresis*; Stathakos, D., Ed.; W. deGruuyter and Co.: Berlin, 1983; 317.
16. Albertsson, P.A. Multistage Procedure. In *Partition of Cells Particles and Macromolecules*, 2nd Ed.; Albertsson, P.A., Ed.; Wiley intersciences: New York, 1986; 161.
17. Treffey, T.E.; Sharp, P.T.; Walter, H.; Brooks, D.E. Thin Layer Counter Current Distribution. In *Partitioning in Aqueous Two Phase Systems*; Walter, H., Brooks, D.E., Fisher, D., Eds.; Academic Press: Orlando, FL, 1985; 132.
18. Deuser, M.S.; Vellinger, J.C.; Naumman, R.J.; Guinn, M.R.; Todd, P. Apparatus for Aqueous Two-Phase Partitioning for Terrestrial and Space Application, AIAA Life Sciences and Biomedical Conference, Houston, TX, April 1995, 65. Book of Abstracts 95-LS-42.
19. Todd, P.; Raghavarao, K.S.M.S.; Sengupta, S.; Doyle, J.F.; Vellinger, J.; Deuser, M.S. Multistage electrophoresis system for the separation of cells, particles and solutes. *Electrophoresis* **2000**, *21*, 318.
20. Liberti, P.A.; Feeley, B.P. Analytical Process-Scale Separation with Bioreceptor Ferrofluids and High Gradient Magnetic Separation. In *Cell Separation Science and Technology*; Kompala, D.S., Todd, P., Eds.; ACS Symposium Series: Washington, 1991; 268.
21. Raghavarao, K.S.M.S.; Srinivas, N.D.; Chethana, S.; Todd, P. Field-assisted extraction of cells, particles and macromolecules. *Trends. Biotechnol.* **2002**, *20*, 72.
22. Munro, P.A.; Dunhill, P.; Lilly, M.D. Casein Hydrolysis in stirred tank reactors using chymotrypsin immobilized on magnetic supports. *Biotechnol. Bioengg.* **1981**, *23*, 677.
23. Safarika, I.; Safarikova, M. Use of magnetic techniques for the isolation of cells. *J. Chromatogr. B.* **1999**, *722*, 33.
24. Shinkai, M. Functional magnetic particles for medical application. *J. Biosci.* **2002**, *94*, 606.
25. Weatherley, L.R.; Campbell, I.; Slaughter, J.C.; Sutherland, K.M. Electrically Enhanced Solvent Extraction of Biochemicals. In *Separations for Biotechnology*; Verrall, M.S., Ed.; 1987; 341.

Magnetic Nanoparticles: Structure and Bioapplications

Yudhisthira Sahoo

Edward P. Furlani

Earl J. Bergey

*Institute for Lasers, Photonics and Biophotonics, State University of New York at Buffalo,
Buffalo, New York, U.S.A.*



INTRODUCTION

Nanometer-sized particles of nearly all materials exhibit novel properties because their state is intermediate between that of atoms or molecules and that of bulk, solid materials.^[1,2] Particles in the size range of 1–100 nm can exhibit quantum size effects, depending on their bulk Bohr radius. They may show size-dependent optical absorption and emission, coulomb staircase behavior in electrical transport, quantum tunneling of magnetic moment in magnetism, and other properties unique to materials in this size range. Nanoparticles of magnetic substances—metals, alloys, and oxides—acquire noteworthy new qualities and their magnetic characteristics are significantly altered relative to those of the corresponding bulk materials. Magnetic nanoparticles are active components of ferrofluids, recording tape, and flexible disk recording media. There are exciting future applications of magnetic nanoparticles in spintronics—a new paradigm of electronics that utilizes both the intrinsic charge and intrinsic spin of electrons for ultra-high-density data storage and quantum computing. Magnetic nanoparticles are also used in important biomedical applications. These include magnetic bioseparation of labeled cells and biological entities, therapeutic drugs, gene and radionuclide delivery, radiofrequency-induced destruction of cells and tumors via hyperthermia, and contrast-enhancement agents for magnetic resonance imaging (MRI). Magnetic nanoparticles are well suited for these applications for several reasons. First, they can be selectively fabricated in sizes that range from several to hundreds of nanometers making them smaller than or comparable to important biological entities such as cells (10–100 μm), viruses (20–250 nm), proteins (3–50 nm), and genes (10 nm wide and 10–100 nm long). Second, they can be surface functionalized with biological molecules that enable selective interaction with various target bioentities. Third, their motion can be manipulated and controlled by an external magnetic field.^[3,4] Fourth, they can absorb energy from a time-varying magnetic field and dissipate it at the cellular level in the form of heat.^[5]

Studies of magnetic nanostructures encompass a broad range of synthetic and investigative techniques involving tools from and knowledge of chemistry, physics, and engineering. There are two basic approaches to nanoparticle preparation and assembly: “bottom up and top down.” The bottom-up approach takes molecules or a cluster of molecules (nanoparticles) and assembles them up into a pretailored architecture. This approach relies on the energetics of the assembly process to guide it. Typical examples are templated film growth, self- and directed-assembly of colloidal particles, and spinodal wetting/dewetting. The top-down approach, on the other hand, relies on micromachining materials to the desired sizes and patterns, and is generally subtractive in nature. Typical examples of the top-down approach include photolithography, mechanical machining/polishing, laser beam and electron beam processing, and electrochemical removal. In this entry, we are particularly focused on the preparation by the bottom up approach and bioapplications of magnetic nanoparticles. The particle of particular interest is Fe_3O_4 and/or Fe_2O_3 in 10–100 nm size ranges that are built as an aqueous ferrofluid.

SUPERPARAMAGNETISM

Magnetic characteristics of ultra-small particles are very different from that of their bulk counterparts. Magnetic particles in the size range of less than ~ 20 nm consist of single magnetic domains unlike in a bulk ferromagnet where there are several domains. In a single-domain ferromagnetic particle, below the characteristic Curie temperature of the material, the individual spins are ferromagnetically coupled giving rise to a large magnetic moment, typically 10^3 to 10^5 μB , and thus the particle is spontaneously magnetized. Even as the ferromagnetic order within the particle remains intact, the magnetization vector of the particle as a whole points in random directions and thus, the particle as a whole acts as a giant paramagnet, hence the term superparamagnet. Thus, in the presence of an external

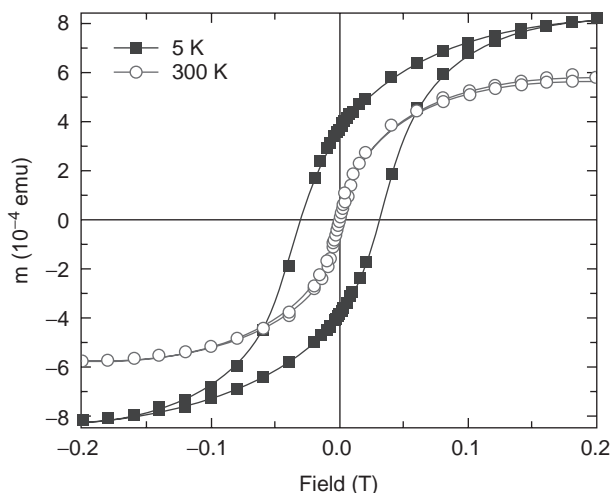


Fig. 1 Superparamagnetic nature of magnetic particles. There is no hysteresis at room temperature. Hysteresis develops owing to blocked states at low temperatures.

magnetic field, a collection of superparamagnetic particles will all point their magnetization vectors along the field direction. But as the field is withdrawn, they relax back to their paramagnetic state and the spins orient randomly. Therefore, the superparamagnetic particles do not show any magnetic hysteresis, except below their blocking temperature (Fig. 1). The super-spins are frozen in definite directions. This property is amply useful in bioapplications.

Another issue to have a magnetic ferrofluid ready for any bioapplication is to have the magnetic particles be colloiddally stable when dispersed in water. It is a technological challenge to control the dispersibility of nanoparticles in desired solvents. The nanoparticles have large surface to volume ratio and therefore high surface energies. Consequently, they tend to aggregate so as to minimize the surface energy. In addition to a high surface energy, magnetic nanoparticles experience interparticle dipolar attractions that tend to further destabilize a colloidal dispersion of them. A suitable surface functionalization and choice of solvent are crucial to achieving sufficient repulsive interactions to prevent aggregation so as to obtain a stable colloidal solution of magnetic nanoparticles—a ferrofluid.

NANOMAGNETIC PARTICLE SYNTHESIS TECHNIQUES

Homogeneous Precipitation

This is a method of precipitation where a short single burst of nuclei are formed when the concentration of the reagents crosses a critical limit.^[6] These nuclei so

formed are then allowed to grow by diffusion of solute within the solution that goes and adds to the surface of the nuclei. The nucleation and growth steps are temporally separated. At the end of the reaction, depending on the system, nearly uniform crop of particles are formed. Often in this process, there is a sharpening of the particle sizes by a process called Ostwald ripening where owing to energy consideration, smaller particles dissolve into the solution and the resulting solute gets added to the large particles surface to make them even bigger. The process continues to make the entire crop of particles nearly monosized.

Introduction of a suitable surfactant during or after the synthesis prevents aggregation of the nanoparticles, and makes a stable colloidal dispersion possible. Most surfactants adhere to surfaces in a substrate-specific manner. For example, thiols bind to gold and silver surfaces, carboxylic acids bind to oxide surfaces, and nitriles and amines bind to platinum. However, the mere presence of a functional group does not guarantee that a ligand will bind to the surface of a nanoparticle, or that a stable dispersion will form. Often, the chain length and solvent system have significant influence on the formation of a truly stable dispersion. In the case of magnetic ferrite particles, the carboxylate ends of long-chain carboxylic acids such as oleic acid and stearic acid bind well to the particles so that the aliphatic chains extending out from the surface of the particle make them effectively hydrophobic and dispersible in nonpolar solvents. Alkyl phosphonates and phosphates also bind to magnetite particles well, and render their surfaces hydrophobic.^[7] Intermediate length carboxylic acids from C₉–C₁₂ can form a bilayered structure and make the surface hydrophilic, enabling formation of a stable aqueous dispersion.^[8] Further if a different carboxylic acid ligand such as citric acid or dimercaptosuccinic acid is introduced onto the particle surface, adsorption takes place, but also the surface of the particle is terminated with a polar group.^[9,10] This results in particles that are soluble in polar groups such as water. In fact, aqueous solutions of magnetite particles with such ligands have been found to be stable for several months without any sedimentation. Fe₃O₄ particles that are surface terminated with carboxylic groups (citric acid) can further bind to a dye molecule and be fluorescently labeled, and cellular uptake can be imaged using fluorescent microscopy (Fig. 2).

Coprecipitation

Coprecipitation is an example of homogeneous precipitation to prepare a magnetite system. Magnetite (Fe₃O₄) particles are prepared by the Massart's coprecipitation of water-soluble iron salts in a basic

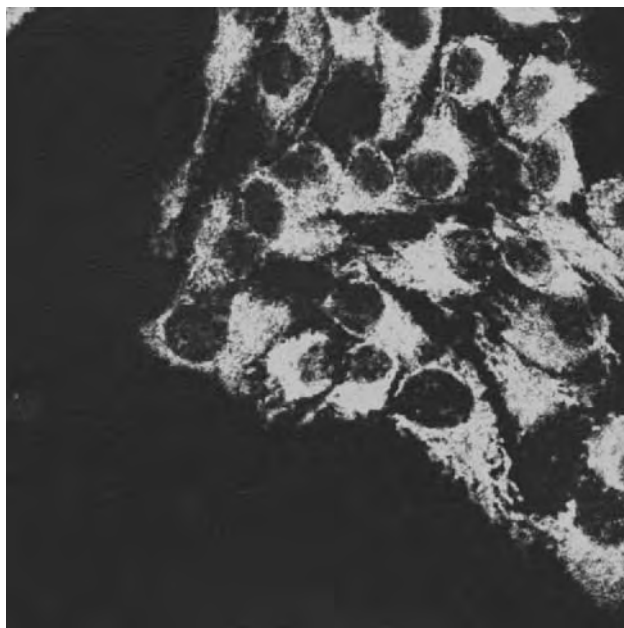


Fig. 2 Imaging of fluorescent magnetic nanoparticle uptake in cultured cells (KB—an oral adenocarcinoma cell line).

medium.^[11] A mixture of FeCl_2 and FeCl_3 (1:2 molar ratio) is taken as a concentrated solution ($\sim 10\text{ M}$) and heated to $70\text{--}80^\circ\text{C}$ in inert environment, and concentrated NH_4OH in excess of three molar equivalents is injected into the reaction vessel. This yields a black precipitate consisting of Fe_3O_4 nanoparticles that remain partially suspended in the reaction mixture at a high pH. At this time, a long-chain carboxylic acid such as oleic acid or stearic acid may be adsorbed on the surface of the particles by the carboxylate end, rendering the particle surfaces alkyl terminated and hence hydrophobic. These particle surfaces can also strongly adsorb phosphonic acids through the phosphonate end. The particles thus surface-modified are soluble in organic nonpolar solvents such as hexane, toluene, etc. Size-selective precipitation may be employed to narrow the size distribution. This method relies on the fact that surface-passivated particles of different sizes, while remaining in a dispersed state, selectively precipitate according to their sizes by differential centrifugation. Thus, a series of different fractions of particles of a different average size can be produced.

Nonaqueous Colloidal Precipitation

This is an example of homogeneous precipitation carried out in an organic medium. Metal complexes such as carbonyls can be decomposed at moderately high temperatures to form metal clusters. The growth of the

metal clusters can be arrested by concomitant surface passivation. Long-chain carboxylic acids, amines, and thiols are known to bind to the surfaces of metals and oxides. For instance, cobalt nanoparticles have been produced by the use of oleic acid as the surfactant and octyl ether as the solvent. The particles were synthesized by reduction of cobalt salt with superhydride at a high temperature. High temperature allows the formation of multiple nuclei, and a long heating time often allows crystallization, so that each nanoparticle is a single crystal. Ferrite particles such as Fe_3O_4 have been prepared by a high-temperature reduction called polyol process where usually a long-chain diol is employed to obtain these particles from the organic precursors.^[12] Organically soluble iron acetylacetonate decomposes at a high temperature ($\sim 240^\circ\text{C}$) in the presence of hexadecanediol to give Fe_3O_4 nanoparticles of uniform sizes (Fig. 3). In this process, the nucleation and growth takes place in the presence of the surface-binding ligand such as oleic acid or oleyl amine, or their mixture. Appropriate mixture of Co and Mn precursors gives rise to CoFe_2O_4 and MnFe_2O_4 nanoparticles, respectively. Because of the uniform sizes of particles already obtained by this method, there is no necessity of size-selective precipitation.

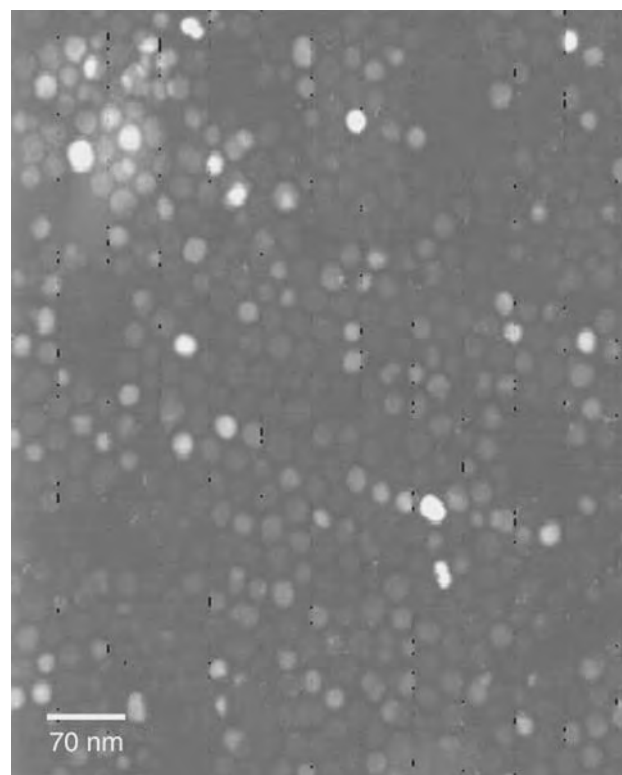


Fig. 3 Transmission electron micrograph of monodispersed Fe_3O_4 particles prepared by organic route.

Microemulsion Formation

Microemulsion is a homogeneous phase obtained from a mixture of water, oil, and surfactant. The surfactant is responsible to form a micelle or a reverse micelle depending on the ratio of water to oil. In oil-in-water microemulsions, small oil globules are formed called micelles, and in reverse stoichiometry, i.e., water-in-oil microemulsions, small droplets of water are formed in the oil phase, called reverse micelles. The micellar and reverse micellar droplets, can both be used as micro/nano reactors to cause a precipitation reaction, and often the product formed is regulated by the sizes of the reactor. Magnetic particles can be formed by appropriate reaction within these reactors. Magnetite particles of ~ 4 nm average diameters have been reported by controlled hydrolysis of ammonium hydroxide of a mixture of FeCl_2 and FeCl_3 in a reverse micelle reactor. Further, Fe_2O_3 nanoparticles of 4–6 nm diameters have been prepared by reaction of $\text{Fe}(\text{AOT})_2$ (AOT: Aerosol OT) with methyl amine in a micellar system.^[13]

Pyrolysis

There are important magnetic materials, such as metallic nickel, for which colloidal synthesis of nanoparticles is difficult or impractical. However, these particles can be readily prepared in the vapor phase by decomposition of suitable precursors. This is accomplished using a laser-driven aerosol synthesis reactor in which an infrared laser rapidly (in less than 1 msec) heats a dilute mixture of precursor vapors to decompose the precursor and initiate particle nucleation.^[14] The laser heating is followed by rapid cooling of the freshly nucleated particles by mixing with unheated gas. This quenches their growth at controllable sizes, typically from 5 to 50 nm diameter. Online aerosol instrumentations allows for direct measurement of the size distribution of particles larger than 10 nm in diameter. The particles can be collected directly into appropriate combinations of solvents and surfactants without exposure to the atmosphere, which may result in unwanted oxidation of the particles. If necessary, an additional size-selection procedure is adopted to narrow the size distribution of the particles produced. This is accomplished by adding a suitable nonsolvent (e.g., ethanol into a hexane dispersion) and centrifuging the mixture at various speeds and collecting fractions with different median sizes of the particles. With addition of small amounts of oxygen to the reactant stream, the same method can be used to prepare metal oxide nanoparticles.

For some materials, volatile precursors are readily available, such as iron carbonyl for preparation of iron

or iron oxide particles. For nickel, the corresponding precursor (nickel carbonyl) is highly toxic. However, for vapor-phase synthesis, nickel carbonyl can be generated in situ from activated nickel and carbon monoxide, and then immediately react it to form nanoparticles and decompose any residual precursor. Appropriate precursors for many materials can be readily identified from the vast literature on chemical vapor deposition of thin films on solid surfaces from vapor phase precursors.

BIOMEDICAL APPLICATIONS

Disease Therapy

Since 1950, magnetic probes or particles have been investigated as a potential alternative treatment for cancer. Studies have demonstrated that the hyperthermia generated by magnetic particles coupled to a high-frequency AC magnetic field could be used as an alternate or an adjuvant to therapy for cancer treatment.^[15] The hyperthermic activity (heat produced by the relaxation of magnetic energy of the magnetic material) was shown to effectively destroy tumor tissue surrounding the probes or particles. The development of ultra-small magnetic particles (ferrofluids) with high crystallinity provided the next step in magnetically induced hyperthermic therapy. This treatment resulted in reduction of tumor size when the particles were directly injected into the tissue, demonstrating the ability of magnetic nanoparticles to damage the protein synthesizing system in mouse liver, *in vivo*. The problem with this approach, like any thermal treatment, is that the heating cannot remain localized, and therefore there is collateral damage of neighboring healthy tissue.

Another mechanism for targeted disease therapy and imaging has been reported. The “nanoclinic” is a multifunctional magnetic nanoparticle having a magnetic core surrounded by a functionalized silica shell to which biotargeting agents can be attached (Fig. 4).^[16,17] These particles are fabricated by first making a nanometers sized magnetic core have a two photon fluorophore attached. This is followed by the encapsulation the surface with a layer of silica. The silica layer is surface terminated with carboxylate group to which a biological agent can be attached to provide targeting ability for specific cell types. This is prototypic magnetic nanoclinic that can selectively destroy specific cancer cells using a DC magnetic field. The significance of this work is that magnetocytolysis (cell lysis) with these magnetic-nanoclinics was achieved using a DC magnetic field. This type of field forms the foundation for MRI, which is commonly used in medical diagnostics. Previous hyperthermic

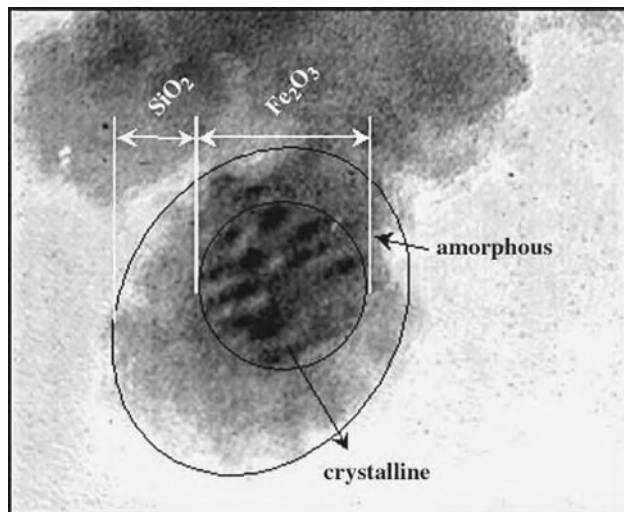


Fig. 4 Transmission electron micrograph of a nanoclinic showing the core/shell structure (with corresponding structured and unstructured pattern).

applications require special instrumentation for generation of AC field, which are not available in clinical settings. Our research has demonstrated that nanoclinics can be synthesized to selectively effect specific cell types with a controllable efficiency. We have fabricated nanoclinics coated with leuteinizing hormone-releasing hormone (LH-RH). The surface coupling of the peptide hormone analogy enables the selective targeting of cancer cells having LH-RH receptors on their surface. Magnetocytolytic activity (effected using a DC magnetic field) was effective only in those cells capable of interacting with these nanoparticles. This effect requires direct contact with and/or internalization into the target cell. These nanoclinics may provide a new noninvasive therapeutic approach to the treatment of selected cancers and could be functionalized to selectively target other cancer cell types.

MRI

In addition to the therapeutic properties that may be associated with magnetic ferrofluids, there has been a significant interest in the use of these nanoparticles as contrast agents in MRI. The MRI measures the changes in the magnetization of hydrogen protons in water molecules resident in the tissues when a pulse of radiofrequency (RF) hits them. Protons of different tissues have different rates of relaxation and mapping their relaxation times essentially maps the tissue system. Magnetic nanoparticles themselves generate a magnetic field and influence the local area around them. The MRI offers the advantage of high spatial resolution of contrast differences between tissues. Because of the relatively long T_1 and T_2 relaxation

times inherent for most tissues, advances in contrast agent-enhanced MRI have strongly favored the development of T_1 -weighted agents. T_1 -weighted agents allow protons to “relax” back to their equilibrium values in a shorter time, resulting in reduced RF saturation effects and an increased number of averages per unit time, yielding increased S/N per unit time or increased resolution per unit time. A desirable T_1 -weighted agent, however, should not significantly affect the bulk magnetic susceptibility of the tissue compartment in which it is localized to minimize any magnetic field inhomogeneities that could result in image artifacts and/or decreased signal intensity. This feature can be exploited in the MRI to enhance the contrast within the tissues of the body. Superparamagnetic iron oxide nanoparticles have been used for image enhancement in the brain of rats to study the morphological and physiological changes in neurological diseases.

Bioseparation

In molecular biology, the ability to separate biomaterials such as cells, enzymes, antigens, or DNA from their native environment is fundamental to the detection and analysis of such entities.^[18–20] Broadly speaking, the primary goal in bioseparation is to separate a low concentration of target biomaterial from a solution, and then release it in sufficiently high concentration to enable a desired analysis. Magnetic bioseparation is a versatile and well-established method to achieve this. It entails the use of magnetic microparticles/nanoparticles with surface treatments (immobilized affinity ligand) that are designed to bind with a target biomaterial. Magnetic particles are commercially available with a wide range of surface functionalizations, custom-tailored for various bioapplications.

Magnetic separation is typically implemented using either a direct or an indirect approach. In the more common direct approach, the surface-treated particles are first mixed with a solution containing the target biomaterial. The mixture is allowed to incubate so that a sufficient amount of biomaterial binds to the particles (bound biomaterial is said to be magnetically tagged or labeled). The labeled biomaterial is magnetically separated from the solution and then released in a higher concentration for further processing. In the indirect approach, the target biomaterial is first incubated in a solution with an affinity ligand (primary antibody), which is added in free form. After a sufficient amount of biomaterial binds to the primary antibody, magnetic particles with surface-bound secondary antibodies (antibodies against the primary antibodies) are introduced into the solution. The mixture is allowed to incubate until a sufficient amount of

the target biomaterial becomes magnetically tagged. This material is then separated from the solution and released in higher concentration for further analysis.

The use of magnetic separation in molecular biology has enjoyed a resurgence of interest over the last decade, and has several advantages over other techniques.^[21] Specifically, it is significantly faster than other methods, and enables the target biomaterial to be isolated directly from crude samples such as blood, bone marrow, tissue homogenates, cultivation media, food, water, etc. Furthermore, the relatively low permeability of the aqueous medium enables efficient coupling between the applied field and the magnetically labeled material. In addition, because biomaterials have a relatively low intrinsic magnetic susceptibility, there is substantial contrast between tagged and untagged material, which enables a high degree of selectivity.

In conventional magnetic separation systems, rare-earth magnets or electromagnets are used to produce a nonuniform field distribution throughout the separation region. When magnetic particles enter this region, they experience a force that moves them toward areas of high field gradient where they are ultimately captured. The particles have a high susceptibility and acquire a dipole moment in an external field, but quickly relax back to an unmagnetized state once the field is removed. Thus, when the field is removed, the separated particles redisperse, thereby enabling the primary function of the separator—isolation of a biomaterial and enhancement of its concentration.

A variety of conventional magnetic separation systems are commercially available and effective, but are less than ideal in that they tend to be relatively large, and they can also be costly and complex requiring significant energy to operate. Accurate manipulation of small particles in small sample volumes is awkward and time consuming in such systems, and the ability to precisely monitor the separation process is limited. However, recent advances in microsystem technology have led to the development of novel magnetic bio separation microdevices that overcome these limitations. Specifically, microbioseparators are energy efficient, and well suited for the analysis and monitoring of small samples (potentially picoliters). Moreover, such devices can be integrated into comprehensive “lab-on-a-chip” systems that comprise several independent stations, each with a different functionality (e.g., biomaterial is isolated in the bioseparation station, and then passed to other biochemical characterization stations).

Several bioseparation microdevices have recently been fabricated and characterized.^[22–24] The performance of such devices can be predicted and optimized from first principles.^[25] Fundamental to such analysis is the ability to predict the behavior of a magnetic

colloid in the presence of an applied magnetic field. The motion of a suspended magnetic particle is governed by several factors including 1) the magnetic force owing to all field sources; 2) viscous drag; 3) inertia; 4) gravity; 5) buoyancy; 6) thermal diffusion; and 7) interparticle effects including: i) magnetic dipole interactions; ii) electric double-layer interactions; and iii) van der Waals force. The relative importance of these factors depends primarily on particle size, and therefore the approach to analyzing particle capture can be broadly divided on this basis. Specifically, for applications involving micron-sized particles, classical Newtonian dynamics suffices. However, for nanoparticles particles, Brownian motion is important and one solves an advection-diffusion equation for the particle concentration, rather than a Newtonian equation for the trajectory of an individual particle. It is difficult to specify the particle radius at which the thermal kinetics become significant, because this depends on the magnetic properties of the particle and the magnitude of the local magnetic field gradient. However, it is generally accepted that when the particle radius is above 50 nm, the classical Newtonian approach is predictive.^[25]

For applications involving micron-sized magnetic particles in low flow regimes (low Reynolds number), the magnetic and viscous drag forces dominate, and the motion of the particle can be predicted using Newton’s second law

$$m_p \frac{dv_p}{dt} = F_m + F_f \quad (1)$$

where m_p and v_p are the mass and velocity of the particle, and F_m and F_f are the magnetic and fluidic forces, respectively. The magnetic force is given by

$$F_m = \mu_0 V_p (M_p \cdot \nabla) H_a \quad (2)$$

where V_p and M_p are the volume and magnetization of the particle, and H_a is the applied magnetic field intensity at the center of the particle. The fluidic force is obtained using Stokes’ law for the drag on a sphere

$$F_f = -6\pi\eta R_p (v_p - v_f) \quad (3)$$

where η and v_f are the viscosity and the velocity of the fluid, respectively. In Eq. (2), it can be assumed that M_p is a linear function of the field intensity up to saturation, at which point it remains constant at a value M_s . Thus, below saturation

$$M_p = \chi_p H_{in} \quad (4)$$

where $\chi_p = \mu_p/\mu_0 - 1$ is the susceptibility of the particle, μ_p is its permeability, and $H_{in} = H_a - H_{demag}$. $H_{demag} = M/3$

is the self-demagnetization field that accounts for the magnetization of the particle.^[26] If the particle is suspended in a magnetically linear fluid of permeability μ_f , the force on it is

$$F_m = \mu_f V_p \frac{3(\chi_p - \chi_f)}{[(\chi_p - \chi_f) + 3(\chi_f + 1)]} H_a \cdot \nabla H_a \quad (5)$$

The Newtonian equation [Eq. (1)] does not account for Brownian motion, which can substantially influence particle capture when the particle radius is much smaller than 25 nm. For particle motion at the nanoscale, one solves the following advection-diffusion equation for the particle concentration $c(r,t)$

$$\frac{\partial c(r,t)}{\partial t} + \nabla \cdot J = 0 \quad (6)$$

where $J = J_D + J_F$ is the total flux of particles, which includes a contribution $J_D = -D\nabla n(r, t)$ owing to diffusion, and a contribution $J_F = v c(r, t)$ owing to the action of all external forces. The diffusion coefficient D is given by the Nernst–Einstein relation $D = \mu kT$ where $\mu = 1/(6\pi\eta R_p)$ (R_p replaces R_F) is the mobility of a particle of radius R_p in a fluid with viscosity η (Stokes' approximation). The drift velocity also depends of the mobility $v(r) = \mu F(r)$, where $F(r) = F_m(r) + F_f(r) + \dots$, is the total force on the particle.

In summary, advances in both microsystem technology and the synthesis of multifunctional magnetic microparticles/nanoparticles for bioapplications have spawned a renewed interest in magnetic bioseparation. Various microbioseparators have been fabricated, and have proven to have significant advantages over conventional devices. Perhaps most importantly, these novel devices will enable a high degree of integration in which the isolation and multiphase analysis of a biomaterial can be performed rapidly and efficiently in a “lab-on-a-chip” environment.

ACKNOWLEDGMENTS

This study was supported by the Directorate of Chemistry and Life Sciences of the Air Force Office of Scientific Research through Defense University Research (grants #F496200110358 and FA95500410158), the National Science Foundation (grants# IBN-9728923, DMR-0318211, HL-49376, NS43621-01), the National Institute of Health (grant# CA104492-01A1), the Oshei Foundation, and the University at Buffalo Interdisciplinary Research and Creative Activities Fund.

ARTICLES OF FURTHER INTEREST

Biological Effects of Radiofrequency Electromagnetic Field; Drug Delivery, Controlled; Micro-Particle Image Velocimetry in Biomedical Applications; Nanoscale Neurosurgery

REFERENCES

1. Prasad, P.N. *Nanophotonics*, Hoboken; John Wiley & Sons, Inc, 2004.
2. Alivisatos, A.P. Perspectives on the physical chemistry of semiconductor nanocrystals. *J. Phys. Chem.* **1996**, *100* (31), 13226–13239.
3. Karumanchi, R.S.M.S.; Doddamane, S.N.; Sampangi, C. et al. Field-assisted extraction of cells, particles and macromolecules. *Trends Biotechnol.* **2002**, *20* (2), 72–78.
4. Mirowski, E.; Moreland, J.; Russek, S.E.; Donahue, M.J. Integrated microfluidic isolation platform for magnetic particle manipulation in biological systems. *Appl. Phys. Lett.* **2004**, *84* (10), 1786–1788.
5. Reynolds, A.R.; Moghimi, S.M.; Hodivala-Dilke, K. Nanoparticle-mediated gene delivery to tumour neovasculature. *Trends Mol. Med.* **2003**, *9* (1), 2–4.
6. Matijevic, E. Uniform inorganic colloid dispersions—achievements and challenges. *Langmuir* **1994**, *10* (1), 8–16.
7. Sahoo, Y.; Pizem, H.; Fried, T.; Golodnitsky, D.; Burstein, L.; Sukenik, C.N.; Markovich, G. Alkyl phosphonate/phosphate coating on magnetite nanoparticles: a comparison with fatty acids. *Langmuir* **2001**, *17* (25), 7907–7911.
8. Shen, L.F.; Laibinis, P.E.; Hatton, T.A. Bilayer surfactant stabilized magnetic fluids: synthesis and interactions at interfaces. *Langmuir* **1999**, *15* (2), 447–453.
9. Sahoo, Y.; Goodarzi, A.; Swihart, M.T.; Ohulchanskyy, T.Y.; Kaur, N.; Furlani, E.P.; Prasad, P.N. Aqueous ferrofluid of magnetite nanoparticles: fluorescence labeling and magnetophoretic control. *J. Phys. Chem. B* **2005**, *109* (9), 3879–3885.
10. Fauconnier, N.; Pons, J.N.; Roger, J.; Bee, A. Thiolation of maghemite nanoparticles by dimercaptosuccinic acid. *J. Coll. Interf. Sci.* **1997**, *194* (2), 427–433.
11. Massart, R. Preparation of aqueous magnetic liquids in alkaline and acidic media. *IEEE Trans. Magnet.* **1981**, *17* (2), 1247–1248.
12. Sun, S.H.; Zeng, H. Size-controlled synthesis of magnetite nanoparticles. *J. Am. Chem. Soc.* **2002**, *124* (28), 8204–8205.
13. Feltn, N.; Pileni, M.P. New technique for synthesizing iron ferrite magnetic nanosized particles. *Langmuir* **1997**, *13* (15), 3927–3933.
14. Sahoo, Y.; He, Y.; Swihart, M.T. et al. An aerosol mediated magnetic colloid: study of nickel nanoparticles. *J. Appl. Phys.* **2005**, *98* (5), 054308.
15. Wada, S.; Tazawa, K.; Furuta, I.; Nagae, H. Antitumor effect of new local hyperthermia using dextran



- magnetite complex in hamster tongue carcinoma. *Oral Dis.* **2003**, *9* (4), 218–223.
16. Bergey, E.J.; Levy, L.; Wang, X.P.; Krebs, L.J.; Lal, M.; Kim, K.S.; Pakatchi, S.; Liebow, C.; Prasad, P.N. DC magnetic field induced magnetocytolysis of cancer cells targeted by LH-RH magnetic nanoparticles in vitro. *Biomed. Microdevices* **2002**, *4* (4), 293–299.
 17. Levy, L.; Sahoo, Y.; Kim, K.S.; Bergey, E.J.; Prasad, P.N. Nanochemistry: synthesis and characterization of multifunctional nanoclinics for biological applications. *Chem. Mat.* **2002**, *14* (9), 3715–3721.
 18. Pankhurst, Q.A.; Connolly, J.; Jones, S.K.; Dobson, J. Applications of magnetic nanoparticles in biomedicine. *J. Phys. D Appl. Phys.* **2003**, *36* (13), R167–R181.
 19. Safarik, I.; Safarikova, M. Magnetic nanoparticles and biosciences. *Monatshefte Fur Chemie* **2002**, *133* (6), 737–759.
 20. Safarýk, I.; Safarýkova, M. *Scientific and Clinical Applications of Magnetic Carriers*; Hafeli, U., et al., Ed.; Plenum Press: London, 1997.
 21. Gijs, M.A.M. Magnetic bead handling on-chip: new opportunities for analytical applications. *Microfluidics Nanofluidics* **2004**, *1*, 22–40.
 22. Smistrup, K.; Hansen, O.; Bruus, H.; Hansen, M.F. Magnetic separation in microfluidic systems using microfabricated electromagnets-experiments and simulations. *J. Magnetism Magnetic Mat.* **2005**, *293* (1), 597–604.
 23. Choi, J.W.; Ahn, C.H.; Bhansali, S.; Henderson, H.T. A new magnetic bead-based, filterless bio-separator with planar electromagnet surfaces for integrated bio-detection systems. *Sens. Actuators B Chem.* **2000**, *68* (1-3), 34–39.
 24. Ahn, C.H.; Allen, M.G.; Trimmer, W.; Yun, Y.N.; Erramilli, S. A fully integrated micromachined magnetic particle separator. *J. Microelectromech. Syst.* **1996**, *5* (3), 151–158.
 25. Fletcher, D. Fine particle high-gradient magnetic entrapment. *IEEE Trans. Magnet.* **1991**, *27* (4), 3655–3677.
 26. Furlani, E.P. *Permanent Magnet and Electromechanical Devices; Materials, Analysis and Applications*; Academic Press: New York, 2001.

Magnetic Resonance Imaging (MRI)

Simon J. Doran

Department of Physics, University of Surrey, Guildford, Surrey, and CRUK clinical Magnetic Resonance Research Group, Institute of Cancer Research, Sutton, Surrey, U.K.

INTRODUCTION

Since its introduction in 1973, magnetic resonance imaging (MRI) has become arguably the preeminent modality of medical imaging and an indispensable tool in hospitals. Should any confirmation of its status be needed, it comes emphatically from the award, in 2003, of the Nobel Prize for Physiology and Medicine to Paul Lauterbur and Sir Peter Mansfield for their seminal contributions to the development of MRI. They joined a growing number of other nuclear magnetic resonance (NMR) laureates: Stern (1943), Rabi (1944), Bloch and Purcell (1952), Ernst (1991), and Wüthrich (2002). It is safe to assume that none of these scientists could, at the time of their initial discoveries, have imagined the enormous scope for applications of NMR and MRI by the early years of the 21st century.

This article will outline: 1) the early history of the subject; 2) the key physical principles; 3) MRI equipment and technology; 4) the sources of contrast in MR images; and 5) the applications of MRI. Whilst the majority of these occur in clinical practice and biomedical research, a significant number of problems in the nonbiological arena may be addressed using MRI, and some of these have great industrial relevance. Given the breadth of the subject, this article does not claim to list exhaustively even all the generic areas of study, let alone the many hundreds of interesting subtopics that have evolved over the last 30 years.

HISTORY

Although the history of MRI is generally traced to the discovery in 1945 of the NMR phenomenon independently by Bloch^[1] and Purcell, Torrey, and Pound,^[2] the foundations of the subject go back at least as far as the discovery of spin in the 1920s. Important experimental work during the 1930s by Stern and Gerlach, Gorter, and then Rabi showed how spin orientation could be measured via an inhomogeneous magnetic field and manipulated using a radiofrequency (RF) field.

During the decades after the second World War, huge strides were made in the understanding of NMR as a *spectroscopic* (i.e., not spatially resolved)

technique. Relaxation was described theoretically by Bloembergen, Purcell, and Pound in 1948. Chemical shift was discovered by Proctor and Yu and by Dickinson, and the spin echo by Hahn, in 1950. The effect of diffusion was analyzed by Carr and Purcell in 1954 and pulsed NMR was first introduced by Ernst and Anderson in 1966. Although all these would later prove invaluable in clinical applications, the focus during this period was primarily on the capability of NMR as a tool for chemical analysis.

Although there had been a number of studies on animal and human tissues throughout the 1950s and 1960s, the undoubted impetus for the use of NMR in medicine came from the 1971 work of Damadian, which showed a systematic difference in the T_1 relaxation properties of cancerous and normal tissues. Damadian filed a patent in 1972 for an “Apparatus and Method for Detecting Cancer in Tissue” and he became a significant player in the subsequent development of MRI. Whilst the originality of the idea proposed in his patent (granted in 1974) was evident, his method for spatial localization of the NMR signal is disputed. No mention is made of magnetic field gradients, now accepted to be key for the formation of an MR image. To this day, he remains a controversial figure. It is generally accepted that the first demonstration of MRI occurred in a Nature paper of March 1973 by Lauterbur,^[3] which showed an image of two tubes of water. Lauterbur’s “projection reconstruction” technique had strong parallels with the x-ray computed tomography (CT) scanning method demonstrated by Hounsfield in the previous year. In September 1973, Mansfield and Grannell,^[4] with a different starting point—their background was in solid state NMR—described theoretically what they called “NMR diffraction.” The graphs that they produced might now be termed one-dimensional (1-D) images, whilst one of the formulae in the paper is instantly recognizable with hindsight as the basis for what is now known as Fourier imaging, a method introduced formally in 1975 by Ernst’s team in Zurich.

The remainder of the 1970s saw a succession of achievements: the first image of a laboratory animal by Hutchison et al. in 1974; the first image of human anatomy in vivo—a finger—and the first MR abdominal images showing the body’s internal organs by



Mansfield and Maudsley in 1976–1977; a cross-sectional image of the chest obtained by Damadian et al. in 1977; and the first brain images from Clow and Young in 1978.

By the early 1980s, clinical applications of MRI were being pioneered in a handful of institutions, notably at the Hammersmith Hospital in London by Young and Bydder. However, at this stage, examinations were slow and the results often of limited clinical utility, hampered by the available technology and a lack of knowledge of some of the fundamental NMR properties of human tissue. Pioneering work by Haase and colleagues in Göttingen, Hennig in Freiburg, and Mansfield's group in Nottingham led to the development of much faster imaging sequences.

With rapid commercialization of MRI in the 1990s, scanners moved from being expensive and speculative research tools to items of routine hospital equipment. Initially, expertise in the field had lain with the manufacturers of spectrometers for chemical research, such as Bruker (Europe), Varian (U.S.A.), and JEOL (Japan). These were joined by a raft of other companies, both big and small. However, by the turn of the century, mergers, acquisitions, regulatory demands,

and other commercial forces had left the majority of production of MR scanners in the hands of three multinationals: General Electric, Philips, and Siemens. Enormous investment continues to be made in technological development by these companies and, with the increased availability of hardware, the number of MRI researchers, both in hospitals and universities, has mushroomed.

PHYSICAL PRINCIPLES^[5]

Signals and Magnetization

Many subatomic particles possess a property called spin (represented by the vector \mathbf{I} in Fig. 1A). Whilst a number of the characteristics of spin may be understood by analogy with the classical angular momentum of a rotating macroscopic body such as a top, spin along a given direction is quantized in half-integer multiples of the basic unit \hbar (Planck's constant/ 2π). Quantum mechanics tells us that it is impossible to measure simultaneously the x -, y -, and z -components

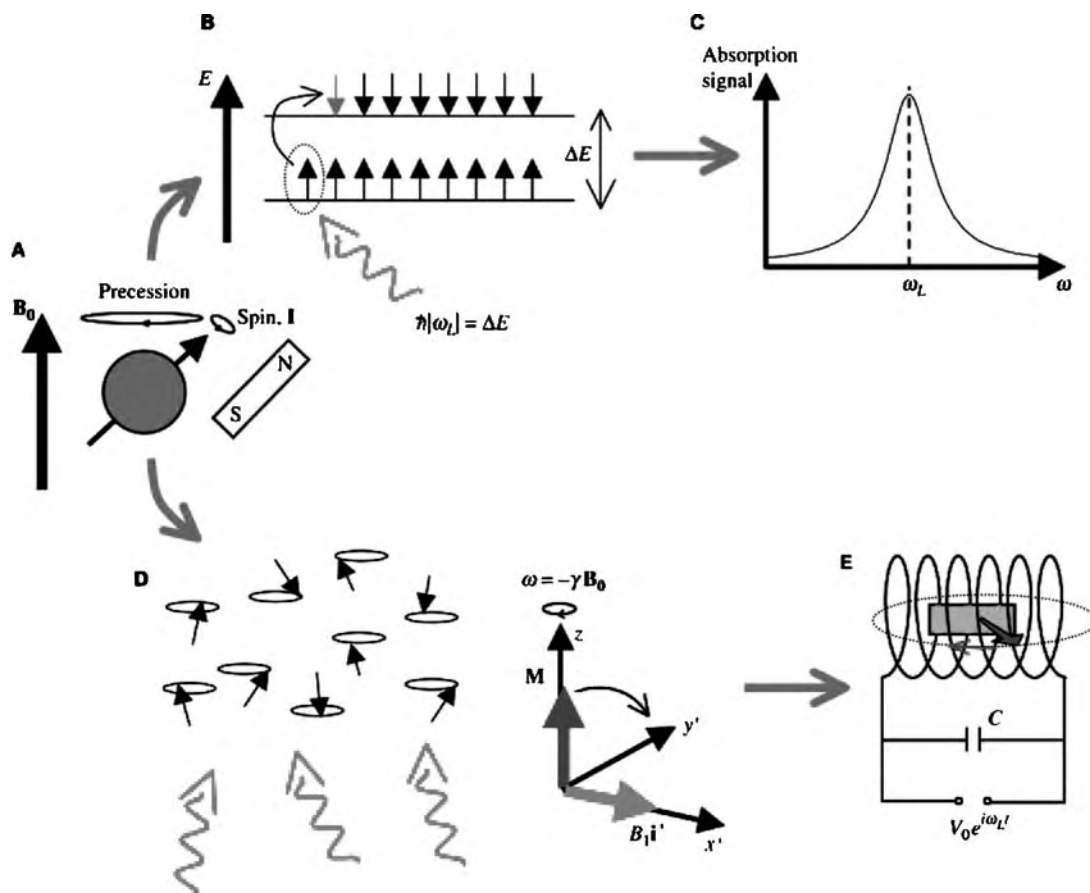


Fig. 1 Physical principles of MRI.

of the spin vector, i.e., to ascertain “the direction” in which a single nuclear spin is pointing. The equations that follow imply that such a definite direction does exist, specifying as they do \mathbf{I} and the associated magnetic moment $\boldsymbol{\mu}$ for an individual nucleus, but the semiclassical approach in this article is justified by a more rigorous quantum statistical mechanical argument.

The nuclei most commonly imaged in medical MRI are those of the hydrogen that is present in the body’s water and fat. For these, quantum mechanics imposes the further restriction that $|\mathbf{I}| = \sqrt{3}\hbar/2$. Suppose one applies a large and constant magnetic field B_0 in the z -direction. Each spin makes an angle $\theta = \cos^{-1}(I_z/|\mathbf{I}|) = \cos^{-1}([\pm\hbar/2]/[\sqrt{3}\hbar/2]) \approx 55^\circ$ or 125° with the z -axis. As $\boldsymbol{\mu} = \gamma\mathbf{I}$, where γ is a constant known as the *gyromagnetic ratio*, the potential energy in the magnetic field is different for spins whose z -components point with and against the field (Zeeman Effect).

$$\mu_z = \gamma I_z = \pm\gamma\hbar/2 \Rightarrow E = -\mu_z B_0 = \mp\gamma\hbar B_0/2 \quad (1)$$

The magnetic field exerts a torque $\mathbf{T} = \boldsymbol{\mu} \times \mathbf{B}_0$, which causes the direction of the spin vector \mathbf{I} to change. The extension of Newton’s Second Law of Motion to this rotating system yields,

$$\mathbf{T} = \frac{d\mathbf{I}}{dt} \Leftrightarrow \boldsymbol{\mu} \times \mathbf{B}_0 = \frac{d(\boldsymbol{\mu}/\gamma)}{dt} \Leftrightarrow \frac{d\boldsymbol{\mu}}{dt} = \gamma\boldsymbol{\mu} \times \mathbf{B}_0 \quad (2)$$

which represents a *precession*, at the so-called Larmor angular frequency $\omega_L = -\gamma B_0$. The spin vector traces out a cone (Figs. 1A and D).

Even the smallest sample interrogated by MRI contains a large enough number of spins for the distribution between the two energy levels (Fig. 1B) to follow accurately Boltzmann’s Law, $n_\downarrow/n_\uparrow = \exp(-\Delta E/kT)$, where n_\uparrow is the number of spins per unit volume in the energetically favorable “up” state ($I_z = +\hbar/2$), n_\downarrow is the number with $I_z = -\hbar/2$ and the energy gap between the levels is $\Delta E = \gamma\hbar B_0 = \hbar|\omega_L|$, k is the Boltzmann constant, and T the absolute temperature. For a sample consisting of an ensemble of many spins, we define the *magnetization* (Fig. 1D) as the magnetic moment per unit volume.

$$\mathbf{M} = \sum_i \boldsymbol{\mu}_i / V \quad (3)$$

It follows that, when the sample is in thermal equilibrium at any finite temperature, there is a population excess in the lower level, leading to a net positive z -magnetization. Because there is no preferred azimuthal angle for the spin vectors, the random x - and y -components of the individual nuclear magnetic moments cancel and $\mathbf{M} = M_0\mathbf{k}$ (Fig. 1D).

RF Excitation and Signal Detection

Two complementary pictures have historically been used to understand the origin of signals in NMR. In Fig. 1B, a quantum of RF energy is absorbed by a nucleus, promoting it from the low to the high energy level. This interpretation, attributable to Purcell, is useful in explaining how resonant absorption (Fig. 1C) of energy occurs in a continuous wave NMR spectrometer. In terms of classical fields, this RF energy is input to the sample in the form of a magnetic field \mathbf{B}_1 , rotating in the transverse (xy -) plane at the Larmor frequency. The equation of *forced precession* is

$$\frac{d\mathbf{M}}{dt} = \gamma\mathbf{M} \times (\mathbf{B}_0 + \mathbf{B}_1) \quad (4)$$

It is very common in the theoretical treatment of MRI to consider a frame of reference rotating about the z -axis at the Larmor frequency (Fig. 1D). In this rotating frame, the spiral motion of \mathbf{M} described by Eq. (4) is replaced by a simple precession about the (now static) vector $\mathbf{B}_1 = B_1\mathbf{i}'$, where \mathbf{i}' is the unit vector in the x' -direction of the rotating frame. In the absence of relaxation (see section “Relaxation”), application of the \mathbf{B}_1 field for a length of time t leads to a rotation of the magnetization through angle $\alpha = -\gamma B_1 t$ about the x' -axis. α is known as the RF pulse (flip) angle and plays an important role in MRI.

Once the RF excitation has been turned off, Eq. (4) applies in the laboratory frame with $\mathbf{B}_1 = \mathbf{0}$. The component of magnetization excited by the pulse into the transverse plane undergoes *free precession* about the longitudinal (z -) axis. The rotating magnetic field thus created can induce an electromotive force (EMF) in a receiver coil that is part of a tuned circuit and this is the way that the NMR signal is detected in most modern spectrometers (Fig. 1E). This interpretation, due to Bloch, turns out to be surprisingly difficult to reconcile with the quantum picture painted by Purcell. A common misconception is that, by analogy with optical spectroscopy, the origin of the signal is the downward transitions in Fig. 1B, whose RF photons are then intercepted by the probe. This is not true:^[6] in fact, there are two terms in the classical expression for the magnetic field of the rotating magnetization, the near- and far-fields. The EMF induced in the probe corresponds primarily to the former (which, in a quantum picture, is mediated by *virtual* photons), whilst the RF photons described above are associated with the (radiative) far-field.

Relaxation

The spins’ environment is a complex of rapidly fluctuating fields created by the precession of neighboring



nuclei, all of which are also in random translational and tumbling motion. Analysis of the effect of these fields is not easy but, happily, the results are entirely consistent with the phenomenological Bloch equations, which are experimentally derived

$$\frac{dM_z}{dt} = \frac{M_0 - M_z}{T_1}, \quad \frac{dM_{xy}}{dt} = -\frac{M_{xy}}{T_2} \quad (5)$$

Upon excitation, the longitudinal component of magnetization M_z returns toward its thermal equilibrium value M_0 in an exponential fashion, with time constant T_1 , whilst the transverse component M_{xy} created by the RF pulse decays exponentially to zero with time constant T_2 . The detail of the quantum mechanics is hidden in the two constants T_1 and T_2 and, for the majority of applications, these relaxation times may be treated as experimentally measured properties of the system. T_1 and T_2 are often referred to by the alternative names *spin-lattice relaxation* and *spin-spin relaxation* time.

Gradients and Imaging

Nuclear magnetic resonance spectra have long been used to obtain *chemical* information about samples: The precession frequency of a nucleus is modified by its electronic environment (chemical shift) as $B_{\text{nucleus}} = (1-\sigma)B_0$, where σ is known as the shielding parameter. Lauterbur and Mansfield's key insight was that one could create a corresponding spectrum in which Larmor frequency was related to *spatial position* by applying a magnetic field gradient. This process is known as *frequency encoding* and the Fourier transform of the signal thus obtained yields a 1-D projection of the sample.

To obtain the full 3-D distribution of spins, one needs to employ additional techniques. *Slice selection* uses a time-varying \mathbf{B}_1 field, in the presence of a gradient, to excite selectively only a thin slice of the sample. To obtain spatial resolution for the second "in-plane" dimension, Lauterbur rotated the direction of the frequency-encoding or "read" gradient, such that he obtained projections over a range of angles. These were reconstructed using the filtered back-projection algorithm. However, the majority of applications today use *phase encoding*, in which a third gradient, orthogonal to both read and slice, is applied during a preparation period before the frequency encoding. By acquiring the 1-D projections with a range of values of this gradient (phase-encoding steps), one obtains a dataset, which is the 2-D Fourier transform of the desired image.

Rapid Imaging

There exist many ways—referred to by an array of acronyms, often vendor specific and baffling to the neophyte—of acquiring an MR image. These are distinguished by the order and duration of the various gradient and RF pulses. The description of a particular MRI experiment is generally called a *pulse sequence* and often presented in diagrammatic form. In the gradient-echo pulse sequence of Fig. 2, the image intensity is given by

$$I(x, y) \propto \rho(x, y) \frac{1 - e^{-T_R/T_1(x, y)}}{1 - \cos \alpha e^{-T_R/T_1(x, y)}} \sin \alpha e^{-T_E/T_2^*(x, y)} \quad (6)$$

Here ρ is the proton density and T_2^* is a modification of the T_2 relaxation time, which takes account of the local magnetic field inhomogeneity, something that may result from the design of the magnet itself, or, more commonly, from spatial variations in the sample's magnetic properties. By an appropriate choice of T_R and T_E (the sequence repetition time and the echo time) and the RF flip angle α , which are all controlled by the user, we may obtain images whose contrast reveals differences in the relaxation times T_1 and T_2^* . By using a so-called "spin-echo," rather than a gradient-echo sequence, T_2 weighting is available.

Whereas a traditional gradient-echo MR image might take a few minutes to acquire, a key goal of medical MRI development is the design of methods to

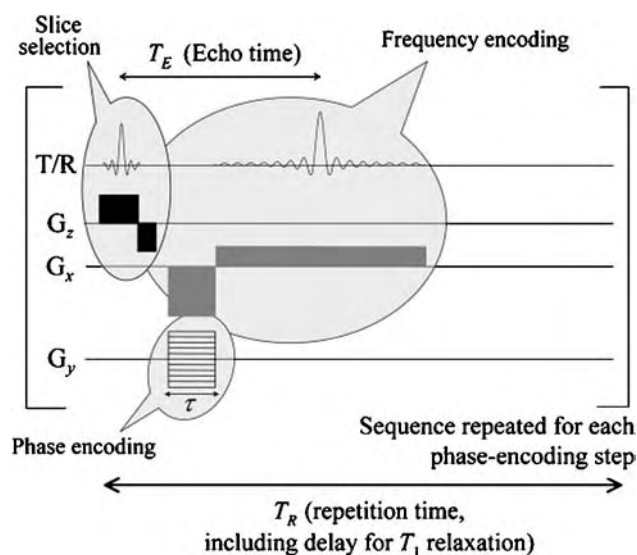


Fig. 2 Pulse sequence for a simple gradient-echo experiment. The various elements of the diagram describe the operation of the RF and gradient systems of the MR scanner.

obtain the necessary information as fast as possible. This is generally achieved either by using short T_R values together with a low flip angle, or by a single excitation followed by a “train” of refocused echoes. If gradient echoes are acquired, then the technique is generally a variant of echo-planar imaging (EPI), whilst, in the technique originally known as rapid acquisition with relaxation enhancement (RARE) and now often referred to as fast spin-echo, spin echoes are refocused by additional RF pulses. The choice between the techniques will be dependent, among other factors, on the T_2^* of the sample. Single-slice imaging is possible within tens of milliseconds and it is now routine for entire 3-D volumes to be acquired within a single breath-hold—this avoids image degradation by respiratory motion.

Ultimately, imaging speed is limited by the image signal-to-noise ratio (SNR). The noise power P_N entering into the NMR measurement is given by the Nyquist/Johnson equation $P_N = 4kT\Delta f$, where the signal bandwidth Δf is inversely related to the gap between data sampling points. Until recently, engineering issues such as the switching speed of the gradients provided other acquisition bottlenecks. However, we now operate close to the *physiological limit* where faster gradient pulses would cause nerve stimulation or even pain, by the induction of EMFs in the body, and thermal deposition of power in the patient from more rapidly repeated RF pulses would exceed regulatory limits.

MRI TECHNOLOGY

Fig. 3 illustrates schematically the apparatus required for a clinical MRI scanner. The main magnet, which produces the B_0 field, represents the largest capital cost. Huge improvements in both the available field strength (currently up to 8 T for human and over 20 T for microscopy systems) and its homogeneity (which is related to T_2^*) have been made over recent years.

The gradient set, together with its associated controllers and current amplifiers, produces the controlled variation in magnetic field required for spatial localization. Here, the key performance parameters are ramp time (i.e., the speed at which the gradient may be changed from one value to another), maximum gradient (currently around 40 mT/m for a typical clinical system), and uniformity of the gradient throughout the imaging region. It was the introduction in the 1980s of *active gradient screening* that revolutionized the performance of gradient systems and made possible all of the rapid imaging we take for granted today.

The part of the clinical scanner that has seen the largest change in recent years has been the RF system. Until the 1990s, this consisted of a single “RF probe”

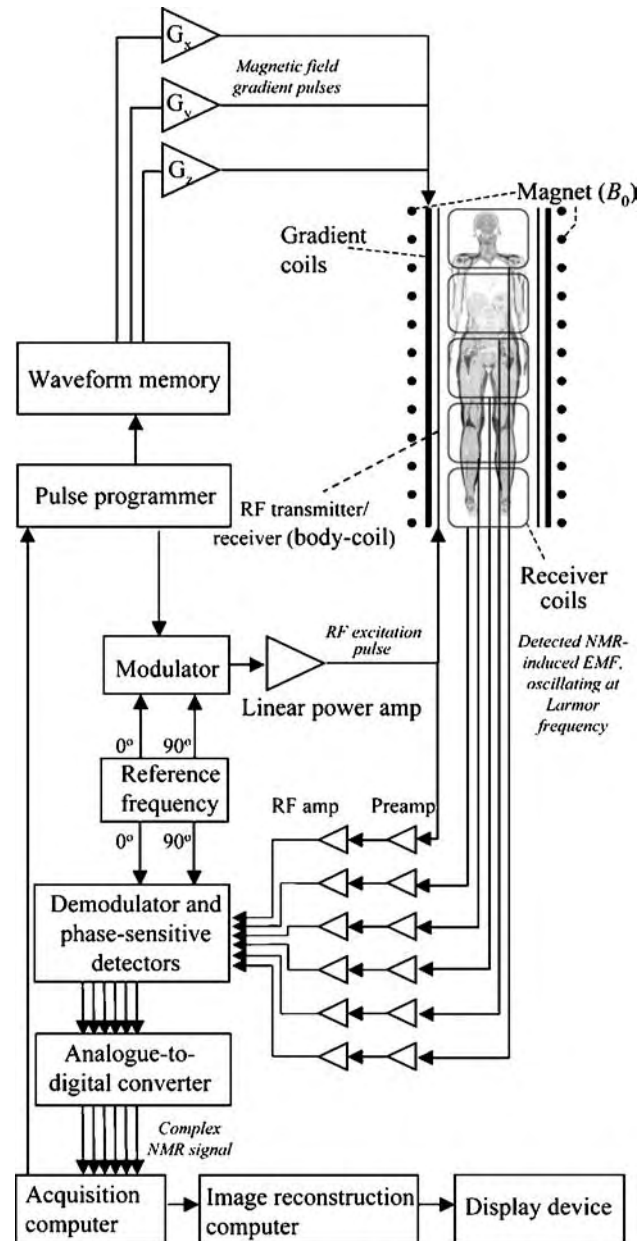


Fig. 3 Schematic diagram of a clinical whole-body MRI scanner.

or antenna, which performed the dual task of transmitting the excitation pulse and receiving the NMR signal. The introduction of *phased array* coils allowed higher SNR with greater coverage and paved the way for the introduction of *partially parallel imaging*. This technique, whose principal methods are known as SMASH and SENSE, uses multiple receiver coils to dispense with the need for a large fraction of the phase-encoding steps. Speedup factors of, typically, 2-6 are obtained without needing to change the RF pulse or gradient pattern, although at the cost of some degradation in SNR. Partially parallel imaging may use



over 30 simultaneously acquired receiver channels, each attached to its own independent RF probe (now restyled “coil element”). Further developments in RF technology allow the patient to be covered in receiver coils, allowing whole-body 3-D coverage via a moving-table technique.

SOURCES OF IMAGE CONTRAST IN MRI^[5]

Image Weighting

Magnetic resonance imaging is possibly the most flexible of all medical imaging modalities, because the image acquisition may be *tailored* to highlight a particular sample property. This is sometimes known as “weighting” the image and leads to the terms “proton-density-weighted,” “ T_1 -weighted,” “ T_2 -weighted,” etc. Compared with the alternative imaging modalities of x-ray CT and positron emission tomography, MRI is sensitive to many more aspects of the sample, as detailed below.

Proton Density

Image intensity is always proportional to M_0 and hence the proton density ρ , but we say that an image is “proton density weighted” when the sequence is specifically desensitized to all the other factors that might affect image intensity. ρ is similar for many types of biological tissue and so this mode of contrast is relatively rare clinically. However, there are numerous industrial applications where quantitative determination of ρ can be very important, for example in the imaging of oil-bearing rock cores. Calibrating such experiments is difficult, but recent developments in multiple-quantum imaging have allowed M_0 to be measured directly in A/m.

Relaxation Times (T_1 , T_2 , T_2^* and $T_{1\rho}$)

Relaxation times are very sensitive probes of structure and environment. Nonspatially resolved NMR uses relaxation times to determine molecular distances and investigate protein–protein interactions, polymer dynamics, and chemical exchange, to name but a few applications. Relaxation times provide the dominant contrast mechanism in clinical imaging, particularly in conjunction with the injection of contrast agents that lower relaxation times as a (reciprocal) function of their concentration. Uptake and dispersion profiles in so-called dynamic contrast studies yield parameters that can be correlated with the blood supply to tumors. Agents with tailored pharmacokinetic properties and the ability to bind to different targets allow image

contrast to be increased for selected organs or pathologies. Analysis of multiple relaxation time components allows one to identify more than one tissue type within a single voxel and thus infer information at a higher spatial resolution than can be imaged directly. T_1 is also temperature sensitive, although latterly measurements based on changes in chemical shift have been used for temperature mapping. In cancer applications, gel dosimetry uses a radiation-induced variation in T_2 to map the distribution of absorbed dose. The dynamic variation of T_2^* around blood vessels is the basis of the blood oxygenation level dependent (BOLD) effect [see section Functional MRI (fMRI)] and the relationship between T_2 and T_2^* may be used to infer the size distribution of blood vessels in some tissues.

Magnetic resonance imaging forms images using signals obtained primarily from hydrogen nuclei that are “mobile” or “weakly bound,” such as those in tissue water. This is because the T_2 values for strongly bound protons are very short (microseconds, rather than the milliseconds) and there is not enough time for imaging pulse sequences to run before the signal decays. Recent hardware developments have dramatically pushed back the frontiers, with the advent of so-called ultrashort T_E sequences, which provide previously unseen contrast in cortical bone, cartilage, tendons, and ligaments. Contrast may also be generated by the mechanism of *magnetization transfer* from bound to mobile protons in close proximity. This occurs though a combination of direct chemical exchange and transfer of the spin state via magnetic interactions, leading to images of mobile protons that are influenced by the “unseen” bound component.

A further relaxation time $T_{1\rho}$ describes the process of relaxation in the rotating frame of off-resonance irradiation. $T_{1\rho}$ is rarely encountered in biomedical imaging, but has an important role in materials MRI.

Motion-Induced Contrast

Macroscopic motion may usefully be divided into two classes: *coherent* within a voxel (e.g., spins in blood flowing along large vessels) and *diffusive*. Perfusion may be considered as a hybrid of these two cases: molecules flow coherently, but along a myriad of randomly oriented capillaries within a single imaging voxel, mimicking diffusive motion, but on a different length scale.

Consider the effect on the NMR signal of spin motion through a gradient applied along x . The Larmor frequency is $\omega_L = -\gamma(B_0 + G_x x)$ and, in the rotating frame of reference, the phase evolution during

the interval $[t_1, t_2]$ is

$$\begin{aligned}\Delta\phi &= -\gamma \int_{t_1}^{t_2} G_x x(t) dt \\ &= -\gamma \int_{t_1}^{t_2} G_x [x_0 + vt + \frac{1}{2}at^2 + \dots] dt\end{aligned}$$

where v and a are the x -components of velocity and acceleration of the spins initially at x_0 .

Random diffusive motion leads to a distribution of phases $\Delta\phi_i$. Summing over all spins, the net magnetization decays exponentially. It is common to write the image intensity as $I(x, y) = I_0(x, y)\exp[-bD(x, y)]$, where the so-called b -value incorporates the influence of the pulse sequence parameters. The gradient G_x initially arose from the imperfect B_0 field of early magnets. However, it soon became convenient to generate controlled field variations, and hence different b -values, via pulses of the imaging gradients of specified amplitude and duration. If two or more such images are available, a quantitative calculation of $D(x, y)$ in m^2/s is possible. Given the need to detect such small displacements, the influence of gross sample motion (both involuntary patient movement and physiological motion associated with respiration and the cardiac cycle) is very destructive in an imaging context, where the measurement takes place over an extended period. The introduction of so-called *navigator echo* techniques was a major advance in this area.

Tissue perfusion may be measured either in terms of a “pseudo-diffusion” coefficient or in terms of a blood flow in ml/min per 100 g of tissue via the relaxation-based techniques of arterial spin labelling or dynamic contrast enhancement.

Where the motion is coherent, all the spins in a given voxel have a similar value of $\Delta\phi$ and the value of v may be extracted from Eq. (7) on the assumption that a and higher order derivatives are zero. By applying gradients along different axes, we are able to map the 3-D velocity field with high spatial resolution. Important clinical applications of this mode of contrast are flow mapping in the heart and angiography (discussed in section “MRI Methodology and Applications in the Clinic”).

Two other types of MRI experiment that exploit motion are worthy of note. In *spin-tagging* experiments, a specialized RF pulse scheme is used to excite tissue whilst leaving thin and regularly repeating “stripes” of unexcited material. As the stripes are associated with the tissue rather than the corresponding image location, they will move if this part of the tissue moves. Such a technique is often used to measure strain in images of the heart. *Magnetic Resonance Elastography* is used to create images whose contrast depends on the local amplitude of an elastic wave

passing through the sample. By synchronizing the timing of a motion-sensitive MRI sequence with the forced mechanical vibration of the object being examined, displacement amplitudes as small as $1\ \mu\text{m}$ may be measured. With various assumptions, the elastic properties of tissue may be mapped. This provides a quantitative assessment of many breast, thyroid, prostate, and abdominal pathologies, which have historically been identified by palpation.

Current Density Imaging

The imaging modality magnetoencephalography uses detectors positioned outside the head to measure electric currents deep inside the brain by making exquisitely sensitive measurements of the tiny ($\sim 10^{-14}$ T) magnetic fields they generate. It has been suggested that direct 3-D imaging of these currents, of major importance in the location of epileptic seizures, might be possible using MRI because of the local changes in Larmor frequency they create. A number of artificial studies have been successful but detection of the effect in vivo remains elusive and it may remain forever below the sensitivity limits of the MRI experiment.

Spectral Information

A modern imaging scanner offers several means to access information on the chemical environment of hydrogen in the body. (Other nuclei of medical interest do give an NMR signal, for example ^{31}P , ^{13}C , and ^{23}Na , but the signal available at physiological concentrations and “natural abundance,” i.e., without isotopic enhancement, is generally insufficient for imaging.) Single voxel spectroscopy allows one to excite a small 3-D region of the sample and apply sophisticated chemical analysis techniques. However, it does not produce an image and if information is required from more than one part of the sample, the procedure must be repeated, which is time intensive. An alternative is to acquire a spectroscopic image, which is a dataset containing an extra dimension, such that each voxel has an entire NMR spectrum associated with it. If information is required about only a single metabolite, then a *spectrally selective* RF pulse may be used, creating a single image whose contrast depends on the concentration of spins within the selected chemical shift range.

The compounds one wishes to detect are generally present in much lower concentrations than water and fat. Hence, an efficient *suppression scheme* is required to avoid swamping the detector with unwanted signals. The weak metabolite signal means spectroscopic images have low SNR and poor resolution. To improve SNR, one may average the signal from multiple



acquisitions, thus increasing the imaging time, or move to higher B_0 fields.

Polarization Transfer^[7]

The cause of the low SNR in MRI is the very small difference between n_{\uparrow} and n_{\downarrow} . The *nuclear polarization* $P = (n_{\uparrow} - n_{\downarrow}) / (n_{\uparrow} + n_{\downarrow})$ for hydrogen at room temperature and 1 T is $\sim 3 \times 10^{-6}$. For certain applications, one may image species whose polarization can be artificially enhanced to a nonthermal state. The most common examples in biomedical MRI are the noble gases, ^3He and ^{129}Xe , which are *hyperpolarized* to $P > 0.1$ by spin-exchange or other techniques with an optically pumped alkali metal vapor. Helium is widely used for imaging the air spaces in the lungs, whilst the enormous chemical shift range of xenon in different environments and its solubility make it an extremely sensitive probe.

Electron spin resonance (ESR) is a well-established method of detecting free radicals, but one not well adapted for high resolution imaging. The proton-electron double-resonance imaging (PEDRI) technique combines ESR with an MRI readout. Electron spins in the free-radical target are irradiated in a low magnetic field, which is then rapidly increased to one suitable for MRI. Spin angular momentum is transferred from the electron to the nuclear spins, thus leading to contrast in the NMR image that reflects the density of the free radicals. Electron spin resonance may also be used for measuring the oxygenation of tissue.

MRI METHODOLOGY AND APPLICATIONS IN THE CLINIC

Cancer^[8,9]

Magnetic resonance imaging plays a major role in all stages of our treatment of cancer, from research on the fundamental biochemistry, through diagnosis (which involves the identification of a lesion, characterization of tumors as benign or malignant, and determination of their extent), to treatment planning, image guidance during surgery and monitoring of the outcome of therapy, assessment of residual disease, and finally detecting relapse in treated or distant tissues.

Key imaging targets are the pelvic organs, breast, and liver (Fig. 4D) and brain, reflecting the prevalence of cancers of these organs. However, MRI is used routinely to diagnose cancers in all parts of the body and also in experimental animal models of disease.

Early imaging studies were based on intrinsic differences in relaxation times between tumors and healthy tissue. However, over the last decade, better

understanding of the tumor microenvironment and improved hardware have dramatically increased the range of techniques available. Dynamic contrast-enhanced (DCE)-MRI measures the leakage of exogenous contrast agent from the capillary vasculature into the extravascular extracellular space, via a rapid succession of T_1 - and T_2^* -weighted images. It gives quantitative values of vessel permeability (elevated in the “new” blood vessels created by tumors) and tissue perfusion. This makes the method ideal for monitoring the treatment efficacy of *antiangiogenic* drugs, which have recently been developed specifically to target the tumor microvasculature. Other techniques used in cancer are the following: targeted contrast agents (see sections “Molecular Imaging” and “Cellular Imaging”); spectroscopy (^1H , ^{19}F , ^{23}Na , ^{31}P) for monitoring tumor and drug metabolism; diffusion-weighted imaging to probe the response to treatment; diffusion tensor imaging to determine tumor infiltration into white matter tracts; BOLD imaging (section “functional MRI”) for both tumor oxygen utilization and functional MRI (fMRI) in surgical planning; correlation of histopathology with micro-MRI of biopsy samples; elastography; monitoring of thermotherapy treatments; and surgical guidance in biopsy.

In many cases, cancer diagnosis requires the development of *biomarkers*, i.e., properties measurable via MRI, which can be correlated with changes taking place at a microscopic level. It is believed that properties such as diffusion or the identification of drug metabolites may be earlier indicators of the success or failure of treatment than gross morphological changes in the tumor.

Neurology

White matter diseases^[11]

One of the diseases most readily detected via MRI, but possibly least understood is multiple sclerosis (MS). Lesions are conspicuous on standard T_2 -weighted scans and change over time, but their presence correlates very poorly with clinical symptoms of the disease. The three major characteristics of MS—*inflammation*, *destruction of the myelin sheaths surrounding axons*, and *axonal injury*—may all be detected via MRI, using a combination of DCE-MRI, cellular imaging of macrophages (see section “Cellular Imaging”), whole brain volumetry, magnetization transfer, diffusion imaging, and spectroscopic imaging.

Some half a dozen or so categories of other white matter disease exist. Standard MRI techniques are not able to distinguish between all of them and research work is therefore ongoing to develop more specific tests.

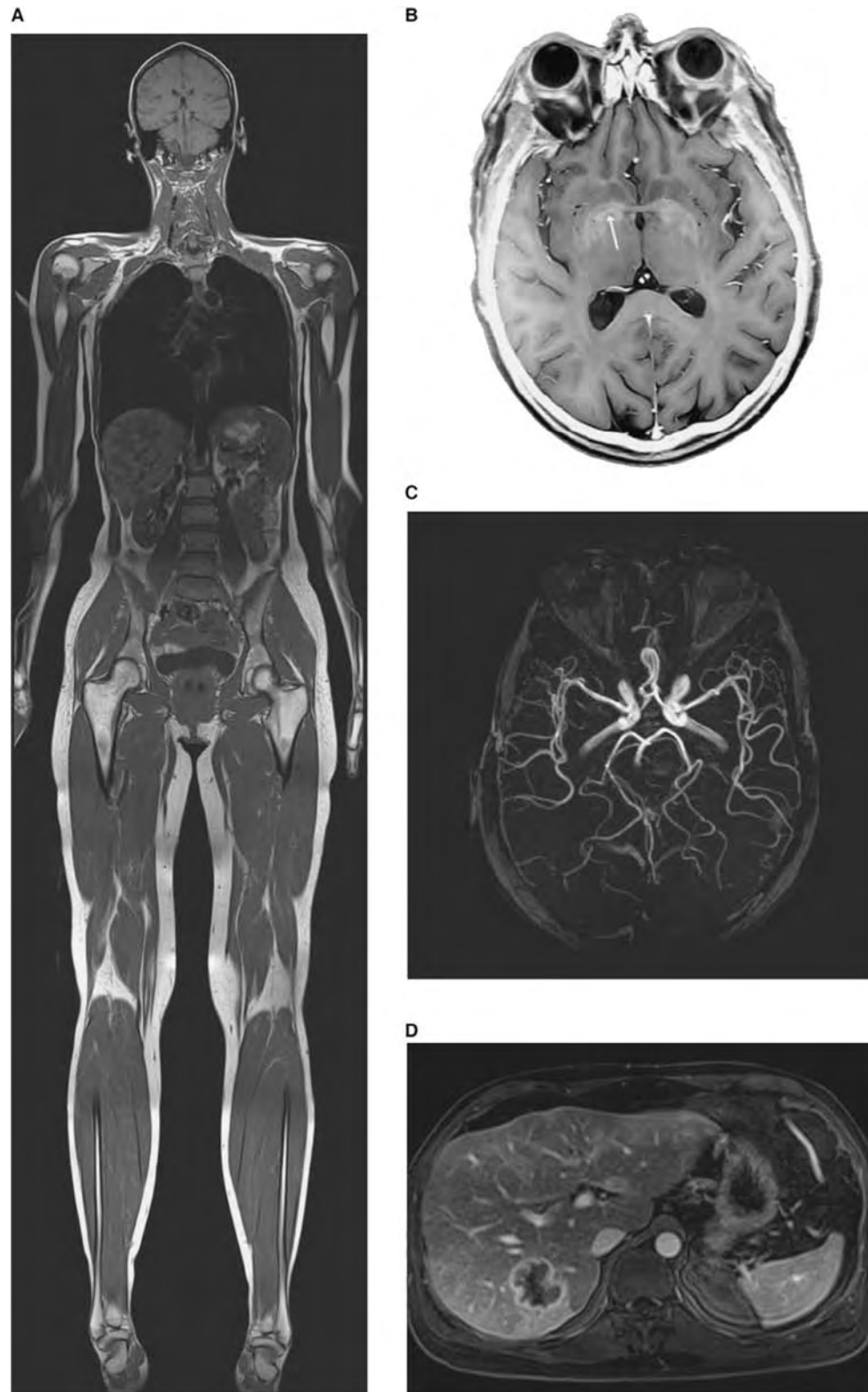


Fig. 4 (A) Whole-body coronal MR image; (B) high-resolution anatomical image of a transverse slice of the brain; (C) MR angiography image of cranial blood vessels; and (D) clinical image showing metastases of a pancreatic carcinoma. (A, C, D: Data courtesy of Siemens Medical Systems; B: from Ref. [10].)

Degenerative diseases^[12]

Although Alzheimer's disease is the most widely known neurodegenerative disease, a number of others are similarly characterized by the loss of nerve cells in the brain. It is often difficult to diagnose these diseases radiologically, because of the lack of distinctive lesions or obvious changes in brain morphology and also because of the significant overlap with the symptoms of normal aging. Magnetic Resonance Imaging in these conditions is important primarily to exclude other conditions such as tumor or stroke. Nevertheless, at a research level some progress is being made in using imaging (e.g., whole-brain volumetry) and spectroscopy (in particular, levels of the compound *N*-acetylaspartate) to investigate the mechanisms of brain damage in these diseases and develop biomarkers markers to assess treatment outcomes.

Diffusion tensor imaging and tractography^[13]

In biological tissues, MRI diffusion imaging (section "Motion Induced Contrast") measures not D , the free diffusion coefficient of water given by the Stokes–Einstein equation $D = kT/6\pi\eta r$, but a quantity known in the literature as the *apparent* diffusion coefficient (ADC). This is usually a factor of 2–5 smaller than D and reflects *restriction* of the diffusive motion of water molecules by the cellular microstructure. It is well known that many tissues (e.g., cardiac muscle tissue and white matter tracts) consist of arrangements of elongated cells with a preferred alignment. Because water diffuses more readily along the cells than across the membranes, the quantitative values of ADC measured in diffusion-weighted imaging (DWI) are strongly dependent on the orientation of the diffusion-encoding gradient. This is known as *anisotropic diffusion* and is studied systematically by repeating the DWI experiment with a number of different diffusion

gradient directions. The output of such measurements is a spatial mapping of the so-called *diffusion tensor*.

Clinical applications of diffusion tensor imaging (DTI) in the heart and brain have developed very rapidly over recent years, supported by technical improvements in rapid imaging capabilities. A typical acquisition with 30 gradient orientations can now be achieved in as little as six minutes. By considering the principal eigenvector of the tensor to be representative of the local fiber orientation and following it from one pixel to another, fiber tracking is now capable of producing extremely high quality atlases (Fig. 5). These classify different regions of white matter in the brain, allow the visualization of the relation of the different tracts with each other and with the gray matter that they connect, and monitor their disruption by pathological processes.

Functional MRI^[14]

The magnetic properties of blood vary depending on its oxygen content. Oxygenated hemoglobin is diamagnetic, while deoxygenated hemoglobin is paramagnetic and has a shorter T_2^* . A sequence such as EPI that gives T_2^* weighting will thus produce images with BOLD contrast. The difference in image intensity between the two states is in the region of 3–5% and depends on B_0 , being larger at higher field strengths. It was found in the early 1990s that brain images obtained in the presence of a stimulus (e.g., a flashing light) differed from those obtained under control conditions (darkness in this case). Furthermore, by appropriate processing of the two datasets—most simply, by subtracting one from the other, although other more sophisticated methods are now used—the regions of the brain *activated* by the stimulus can be determined.

What is observed in fMRI is not the direct activation of neurons, but rather the increased blood flow that accompanies it. Nerve impulses travel from

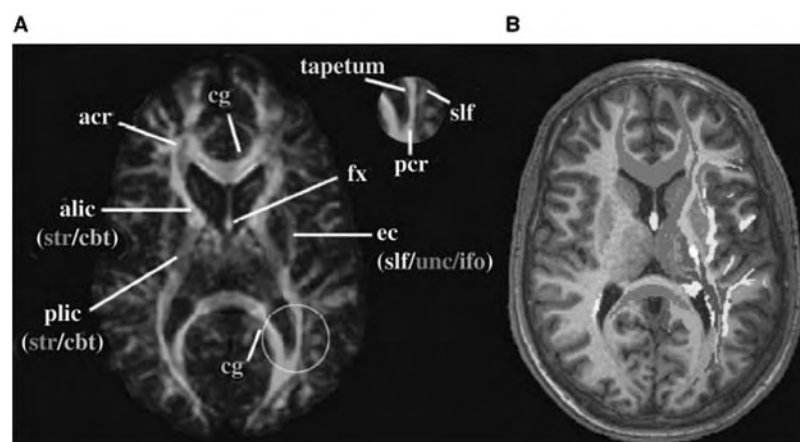


Fig. 5 (A) Transverse DTI color map. Red, green, and blue components of each color represent fibers running along right–left, anterior–posterior, and superior–inferior axes, respectively. Various anatomic landmarks visible are annotated on the basis of existing anatomic knowledge and (B) white matter “parcellation” map with locations of white matter tracts assigned. (From Ref. [13].)

neuron to neuron via the release of neurotransmitters at synapses and the process of recycling these neurotransmitters back into the cell requires energy. This is provided, with a latency of some 4–6 s after the onset of the task, by the hemodynamic response and it is this that the BOLD effect detects.

This discovery has revolutionized the way in which the brain is studied. Because MRI involves no ionizing radiation, studies may be performed repeatedly on healthy volunteers as well as patients. By designing appropriate stimuli, neurophysiologists have been able to elucidate the regions of the brain associated with perception, recognition, interpretation, movement, memory, learning, pain, and a whole host of other functions in normal subjects. Diseases of many types have been studied, including psychiatric conditions such as schizophrenia and compulsive disorders for which objective measurement was hitherto unimaginable. Functional magnetic resonance imaging is now used routinely prior to surgery to map out the position of eloquent structures. Monitoring the rehabilitation after acute stroke has demonstrated the plasticity of the brain, with new areas being recruited to take over functions from damaged regions. It has even been shown that parts of the brain may change in size in response to high usage in the same way that muscles enlarge in response to exercise.

Although fMRI is primarily associated with the brain, the BOLD effect also has applications in other types of tissue (e.g., for cancer studies).

Stroke^[15]

Strokes can be divided into two categories: *ischemic*, caused by a blockage in one of the arteries supplying blood to the brain, and *hemorrhagic*, in which a blood vessel has ruptured. Although the latter can be imaged using susceptibility-weighted MRI, x-ray CT is more commonly used, and it is for the detection of ischemic stroke that MRI finds its greatest application.

When blood flow is interrupted, a cascade of events is set in train, involving altered metabolism and changes in the distribution of neurotransmitters and electrolytes. Cells start to swell as the “pump” maintaining homeostasis fails and there is an influx of water (cytotoxic edema). An accompanying decrease occurs in the water ADC and, within minutes of arterial occlusion, changes in signal intensity can be obvious on diffusion-weighted images.

The extent of neurological deficit can be minimized if a thrombolytic agent is administered while the tissue is still viable (within approximately three hours of the onset of symptoms). However, this treatment is not without risk itself—there is a 10–15% chance of hemorrhage—and so early assessment via MRI is vital.

Lesion size, as measured using DWI in the acute phase, has been shown to correlate well with standard neurological scores and represents *infarcted* tissue. The DWI data may be compared with a map of the corresponding cerebral blood flow (CBF), as measured using perfusion-weighted MR imaging (PWI). This usually indicates a larger abnormal region and so the PWI–DWI mismatch is assumed to represent tissue whose blood supply is compromised but that is still potentially salvageable. The competing modality for the measurement of CBF is currently $H_2^{15}O$ positron emission tomography.

Changes in the MR appearance of stroke continue over about 10 days and the final state is well visualized using T_2 -weighted imaging. Fluid-attenuated inversion-recovery sequences are particularly beneficial in this respect as the strong confounding signal from cerebrospinal fluid is removed.

Angiography^[16]

Angiography is the imaging of the body’s blood vessels (Fig. 5C). This is done for a number of reasons, but primarily because vessels become blocked as a result of atherosclerosis, the deposition of fat in the vessel walls. This leads to ischemic symptoms such as pain, loss of function, severe disability, or even death. Early detection of the plaques that form is routinely obtained with high specificity and sensitivity (>90%) via MR angiography (MRA). Typical targets are the intracranial, carotid, renal, abdominal, pelvic, and peripheral arteries. Further reasons for performing MRA are the visualization of aneurysms and arteriovenous malformations, preoperative planning, the tracking of intravascular catheters as they are moved through the body during surgical procedures, and the positioning of intravascular devices such as stents and filters.

For many years, the “gold standard” for angiographic imaging has been x-ray digital subtraction angiography. However, in more recent times, this method has been superseded by rotational x-ray angiography (RA), which can give true 3-D information. However, in both cases, patients are exposed to ionizing radiation, particularly during the monitoring of interventional procedures, when many such images are acquired. Magnetic resonance angiography provides a radiation-free alternative, with the potential to acquire images of comparable quality. Recent research indicates very high correlations in anatomical depiction between RA and MRA.

Three different classes of MRA technique exist, using different physical properties of the blood to distinguish it from background tissue. *Time-of-flight* measurements exploit the relaxation and flow properties of blood to enhance its signal relative to the

background tissue, whilst *phase contrast* angiography measures quantitative values of flow velocity using the effects described in Section “Motion-Induced Contrast.” In recent times, both methods have been largely superseded by studies in which a large dose (bolus) of contrast agent (CA) is injected and ultrarapid 3-D images are acquired during the first passage of the agent through the vasculature. The CA lowers dramatically the T_1 of blood such that, even at short T_R values for which background tissue is completely saturated, a strong blood signal is still obtained. It is likely that even this technique will be superseded by angiography performed after the administration of macromolecular contrast media that are retained within the vasculature. Other, more experimental, methods for labelling the blood vessels involve the injection of hyperpolarized contrast agents.

Cardiology^[17]

Capturing sharp 3-D images of a moving target with the time-resolution necessary to see individual phases of the heart cycle is a daunting task. Cardiac MRI is, arguably, the field of diagnosis that has seen the greatest impact over recent years of the phenomenal increase in acquisition speed of MR scanners. This has been brought about via the introduction of fast, high-contrast sequences such as fast imaging with steady-state free precession (FISP), gradient modulation strategies (e.g., spiral imaging), and parallel imaging techniques and improvements in gradient performance.

Respiratory motion creates significant difficulties and when cardiac imaging was first introduced, it was necessary to *gate* (i.e., synchronize) the acquisition with both the cardiac and respiratory cycles. This led to extremely long acquisition times. 3-D images may now be acquired within a single breath-hold and so the motion artifacts introduced by respiration are removed. Nevertheless cardiac patients are precisely the group that find holding their breath difficult and so free-breathing strategies are sometimes used (usually in conjunction with navigator techniques) despite the difficulties this may lead to in coregistering slices acquired at different times.

The applications of MRI in cardiology make use of many of the principles described in section “Sources of Image Contrast in MRI.” Standard morphological imaging and cine imaging, employing various mechanisms of contrast to distinguish between different tissues, enable measurement of ventricular volumes and visualization of valve operation, wall thickening, congenital abnormalities, and aortic or pulmonary anomalies. Flow imaging may be used to quantify blood velocities in the chambers of the heart, allowing the ejection fraction (a key diagnostic variable) to be

calculated by integrating over the whole cycle. Velocity-sensitive MRI has the potential to replace color Doppler ultrasound in diagnosing problems such as mitral valve regurgitation. In conjunction with computational fluid dynamics simulations, it may predict pressure gradients across stenoses and regions of turbulence. Angiography techniques are employed for 3-D visualization of the coronary arteries and the evaluation of bypass grafts. Cardiac tagging studies allow one to quantify wall motion, strain, and twisting in the heart muscle. The orientation of the muscle fibers themselves, despite their subvoxel size, can be measured via diffusion tensor mapping. Perfusion imaging via dynamic susceptibility contrast is used to measure regional variations in blood supply to the myocardium and thus to assess the areas of tissue compromised by ischemic heart disease, with a vastly better spatial resolution than the previous technique of ^{99m}Tc single-photon emission computed tomography (SPECT). ^{23}Na MRI can be used to gain an insight into sodium homeostasis in myocytes and thus may allow discrimination between viable and nonviable tissue after a heart attack, whilst specially developed Gd contrast agents may be used to similar effect with ^1H imaging.

In short, MRI is now routinely described as a “one-stop shop” for cardiac diagnosis, allowing visualization of both acute and chronic cardiac problems, including hypertensive and ischemic disease, cardiomyopathies, valve problems, and congenital disorders.

Therapeutic Monitoring^[18]

The use of MRI for monitoring surgical procedures is highly attractive but the challenges to be solved for this to become reality have been considerable. The most fundamental of these is the geometry of the main magnet. For imaging *during* surgery, two designs of “open-access” magnets are currently available. The 0.5 T “double doughnut” consists of two coils, separated by about 50 cm. It gives the surgeon vertical access to the homogeneous region of magnetic field where imaging occurs, allowing operations to be carried out with real-time feedback and with the patient in a seated position. Neurosurgery, in particular, benefits from this arrangement. By contrast, the C-arm approach involves the use of permanent magnet pole pieces above and below the patient. It allows a much wider access from the side of the patient but until recently has only been available as a low-field option (around 0.25 T). A major disadvantage of using many open-access magnets is poor image quality, resulting from the low B_0 field and poor field homogeneity. It is thus often desirable to superimpose the perioperative data onto previously acquired images from a high-field

scanner; this raises important questions of image registration and accuracy. However, newly developed techniques of frameless stereotaxy, involving stereo views from multiple cameras, may improve this situation, as will the recent introduction of higher field systems. An alternative is to use an “in-out” protocol: The patient is slid repeatedly into a standard scanner at one end of the operating theatre, while the interventional procedure itself is performed outside the magnet.

The manufacture of MR-compatible surgical tools represents a second technological challenge. Although nonmagnetic stainless steel is “safe” and not attracted into the magnet bore, it still perturbs the local homogeneity of the magnetic field, causing massive disruption to the resulting images. Titanium and some plastics are better in this regard, but cannot be made as sharp or stiff and may not be so easy to sterilize.

Applications of interventional MRI include *biopsy*, where one wishes both to localize accurately the tissue to be sampled and to visualize structures to be avoided en route to the target; *surgery of brain and skull base tumors*, where access is limited and where there are many critical areas that must not be damaged; and *endovascular treatments*, in which the progress of a catheter or device must be monitored as it passes from the point of insertion to the site of action or placement. Here, MRI has the potential to replace x-ray fluoroscopy, which provides a dynamically updated view of the surgery, but at the expense of a large patient dose of ionizing radiation.

Unique to MRI is the ability to make images of *temperature distribution* within a sample. This allows us to monitor the delivery of clinical hyperthermia treatments, for example, the ablation of liver tumors by laser or thermal catheters or a high-intensity focused ultrasound beam. By a feedback-controlled coupling of the scanner to the delivery system, very precise manipulation of the tissue temperature may be achieved.

Other Medical Applications

Numerous other specialist application areas of MRI exist, the description of which lies beyond the scope of this article. These include: MRI of the musculoskeletal system (with important applications in surgical planning for the lumbar spine, for trauma patients, and for the knee and other joints; the study of the ligaments, tendons, and cartilage and of arthritis); imaging of the gastrointestinal tract; MRI of the kidneys and urinary system; and imaging of the fetus and placenta in utero. Magnetic resonance imaging is also becoming increasingly important for drug development. Whilst this

report has focussed on the significant role of quantitative imaging, the importance of exquisitely detailed morphological information, such as that shown in Fig. 4, must never be underestimated.

BIOMEDICAL APPLICATIONS

Molecular Imaging^[9]

In an era when the human genome has been sequenced and in which we are rapidly learning the biological function of the individual genes that make it up, it is predicted that medical therapy will move rapidly toward a model where drugs are purpose-designed to act on specific genetic/molecular targets. One might imagine that it would be impossible to image such phenomena, as they occur at concentrations and at length scales far below the resolution limits of MRI. However, localization of the point of action of a drug to a value better than the 0.1 mm available on today's MRI scanners is clinically unnecessary. To visualize the effects of a molecularly directed therapy, we simply (!) need a probe with: 1) a high affinity and specificity for the target molecule, enzyme, or receptor; 2) a delivery route to the tissue of interest; and, crucially, 3) the ability to *amplify* nanomolar changes in concentration to signals detectable at the level of an imaging voxel. The most commonly used amplification mechanism in MRI is relaxation: Small quantities of ferromagnetic or paramagnetic materials have an effect on the NMR signal over a large region. Thus, the signal we detect comes not from the relatively few target molecules, but from the much larger pool of surrounding water molecules.

Probes may be divided into two classes: agents that influence water relaxation merely by an increase in concentration, and “smart” agents that change their conformation on interaction with the target. An example of the former is a so-called “polymerized vesicle” containing the relaxation agent Gd^{3+} , which is conjugated to an antibody targeting the compound $\alpha_v\beta_3$ integrin found in many tumors. By contrast, the “Egad” complex contains a β -galactose subunit that prevents access of water molecules to the paramagnetic Gd^{3+} ion trapped inside it. When the enzyme β -galactosidase is present, the ring is cleaved and the relaxation efficiency of the Gd^{3+} is vastly increased. This mechanism has a potentially wide application, because expression of β -galactosidase is widely used as a marker to report successful gene transfection. A ballpark figure of the number of Gd complex units required per cell to obtain successful visualization on an image is $\sim 4-5 \times 10^7$ and initial studies appear to show that this will be achievable in vivo.



Cellular Imaging^[19]

The distinction between molecular and cellular imaging is somewhat muddy. The terms are used loosely and would perhaps be better replaced by “molecular *detection*” and “cellular *detection*,” because (in the common usage) neither is forming images of individual molecules or cells. Broadly speaking, cellular imaging aims to pinpoint the location of specific cells, rather than macromolecules, although it may achieve this via the types of molecular imaging technique described above.

The relaxivity of paramagnetic Gd decreases on internalization (uptake into the cell). This and its poor biocompatibility—the metal is toxic if it becomes dechelated—have led to a preference for iron, in the form of ultrasmall superparamagnetic iron oxide particles as a cellular labelling agent.

Early clinical applications have imaged the distribution of *macrophages* (e.g., visualizing lesions in the liver caused by a deficiency of Kupffer cells, or microscopic metastases in normally sized lymph nodes), but the majority of other studies have been proofs of principle, often in animal models. Nevertheless, the potential applications in humans are already becoming clear. One of the most exciting is the ability to follow labelled *stem cells* as they migrate from the site of injection to their target. The detectable limit for a $100\ \mu\text{m} \times 100\ \mu\text{m} \times 200\ \mu\text{m}$ voxel—the distinction between MRI and MR microscopy is not well defined—at 1.5T has been estimated at $1.4\text{--}3.0 \times 10^{-12}$ g of iron per cell and, given that it appears possible to internalize up to 25×10^{-12} g per stem cell, this opens the possibility of finding the destinations of the first four generations of daughter cells.

True microscopy studies at ultrahigh B_0 fields have demonstrated the feasibility of acquiring images of individual large cells in vitro.

Genetic Phenotyping^[9]

It is known that humans share a very large fraction of the genome with other mammals and, thus, there is a strong likelihood that human gene function may be inferred from studies on transgenic mice. To this end, a systematic program is underway to produce heritable mutations in every gene and study the phenotype of each resulting mouse.

High-resolution ($\sim 100\ \mu\text{m}^3$) 3-D MRI is a valuable screening approach. It has a number of advantages compared with traditional histological methods: 1) it is nondestructive, allowing serial studies; 2) the data do not exhibit the geometric distortion that can be caused by stresses or dehydration during histological sectioning; 3) small organs can be isolated and imaged at

higher resolution, then coregistered with the main dataset; 4) the data acquisition procedure is easier to automate than physical sectioning of the animal—important, as it has been estimated that some 20,000 mouse strains will need to be examined just to characterize the effects of modifying dominant genes; and 5) MRI is inherently 3-D and digital. Not only does this allow the researcher to slice the dataset in any arbitrary orientation to view features of interest, but also via image registration and pattern-matching software, the data for a transgenic phenotype might be compared automatically with a control.

Over recent years, this type of research has led to a growing database of phenotype images with rapid dissemination of data via web-based archives.

INDUSTRIAL AND NONBIOMEDICAL APPLICATIONS^[20]

The role of MRI outside medicine is already huge. However, the high capital cost of equipment and the need, in an industrial setting, to demonstrate profitability have tended to limit the user base to small but active communities in each area. The potential for further growth is considerable as more disciplines discover the capabilities of MRI. All of the techniques hitherto described may be brought to bear on non-medical samples, along with others (for example, stray field imaging, or high-speed magic angle spinning) that cannot be used in vivo.

Materials imaging is a major industrial application, with the compounds investigated including emulsions, paints and coatings, washing powders and other domestic products, plastics, rubbers, composites, dental resins, concretes and cements, textiles, and explosives. Ingress of solvents into polymers and the dissolution of solids (e.g., drug capsules) can be studied, as can be the penetration and efficacy of skin care products. Magnetic resonance imaging has demonstrated its potential in the food industry, for example tracking the migration of fat and water as products are cooked, or measuring the water content of breakfast cereals (and hence determining their crunchiness!). There are applications of imaging in the timber industry (from the mechanism of fungal infection in growing trees to the drying of wood and the manufacture of paper); in the oil industry for studying rock cores and other porous media; and in chemical engineering, where flow through reactors can be monitored and the rheology of complex fluids studied.

Magnetic resonance imaging has many applications in basic science, for example, in following the progress of chemical reactions (particularly those in reaction-diffusion systems, which exhibit spatial heterogeneity), in elucidating mechanisms of water transport in plants,

and in the development of biomaterials, such as artificial tissues.

CONCLUSIONS

This article has provided an introduction to the physics, technology, and applications of MRI. It is clear that the scope of MRI is enormous and spans the physical and medical sciences. Here, we have merely scratched the surface of the subject, providing a starting point for further study.

ARTICLES OF FURTHER INTEREST

Functional MRI: Applications; Magnetic Resonance Imaging in Temperature Measurement; Magnetic Resonance Microscopy; Medical Imaging, 3-D

REFERENCES

1. Bloch, F. Nuclear induction. *Phys. Rev.* **1946**, *70*, 460–461.
2. Purcell, E.M.; Torrey, H.C.; Pound, R.V. Resonance absorption by nuclear magnetic moments in a solid. *Phys. Rev.* **1946**, *69*, 37.
3. Lauterbur, P.C. Image formation by induced local interactions: examples employing nuclear magnetic resonance. *Nature* **1973**, *242*, 190.
4. Mansfield, P.; Grannell, P.K. NMR diffraction in solids. *J. Phys. C Sol. State Phys.* **1973**, *6* (22), L422–L426.
5. Haacke, E.M.; Brown, R.W.; Thompson, M.R.; Venkatesan, R. *Magnetic Resonance Imaging: Physical Principles and Sequence Design*; John Wiley & Sons Inc., 1999; 914 pp.
6. Hoult, D.I.; Ginsberg, N.S. The quantum origins of the free induction decay signal and spin noise. *J. Magn. Reson.* **2001**, *148* (1), 182–199.
7. Goodson, B.M. Nuclear magnetic resonance of laser-polarized noble gases in molecules, materials, and organisms. *J. Magn. Reson.* **2002**, *155* (2), 157–216.
8. Padhani, A., Ed.; *Angiogenesis Imaging*; Boehringer Ingelheim, 2003; Vol. 76.
9. King, C.P.L. Molecular imaging. *J. Magn. Reson. Imag.* **2002**, *16* (4), *special issue*.
10. Thomas, et al. High-resolution fast spin echo imaging of the human brain at 4.7 T: implementation and sequence characteristics. *Magn. Reson. Med.* **2004**, *51* (6), 1254–1264.
11. Filippi, M.; Falini, A.; Arnold, D.L.; Fazekas, F.; Gonen, O.; Simon, J.H.; Dousset, V.; Savoirdo, M.; Wolinsky, J.S. Magnetic resonance techniques for the in vivo assessment of multiple sclerosis pathology: consensus report of the white matter study group. *J. Magn. Reson. Imag.* **2005**, *21* (6), 669–675.
12. Petrella, J.R. Neuroimaging and early diagnosis of Alzheimer disease: a look to the future. *Radiology* **2003**, *226* (2), 315–336.
13. Wakana, S.; Jiang, H.Y.; Nagae-Poetscher, L.M.; van Zijl, P.C.M.; Mori, S. Fiber tract-based atlas of human white matter anatomy. *Radiology* **2004**, *230* (1), 77–87.
14. Jezzard, P.; Matthew, P.M.; Smith, S.M. *Functional Magnetic Resonance Imaging: An Introduction to Methods*; Oxford University Press, 2003.
15. Xavier, A.R.; Qureshi, A.I.; Kirmani, J.F.; Yahia, A.M.; Bakshi, R. Neuroimaging of stroke: a review. *Southern Med. J.* **2003**, *96* (4), 367–379.
16. Schneider, G., Prince, M.R., Meaney, J.F., Vincent, B., Eds.; *Magnetic Resonance Angiography: Techniques, Indications and Practical Applications*; Springer-Verlag Italia Srl, 2005.
17. Woodard, P.K.; Stillman, A.E. Cardiac MR and CT. *J. Magn. Reson. Imag.* **2004**, *19* (6), *special issue*.
18. Schulz, T.; Puccini, S.; Schneider, J.P.; Kahn, T. Interventional and intraoperative MR: review and update of techniques and clinical experience. *Eur. Radiol.* **2004**, *14* (12), 2212–2227.
19. Bulte, J.W.M.; Kraitchman, D.L. Iron oxide MR contrast agents for molecular and cellular imaging. *NMR Biomed.* **2004**, *17* (7), 484–499.
20. Blümmler, P., Blümich, P., Botto, R., Fukushima, E., Eds.; *Spatially Resolved Magnetic Resonance: Methods, Materials, Biology, Rheology, Geology, Ecology, Hardware*; Wiley-VCH: Weinheim, New York, Chichester, Brisbane, Singapore, Toronto, 1998.



Magnetic Resonance Imaging in Temperature Measurement

Nadine Barrie Smith

Department of Bioengineering, Pennsylvania State University, University Park, Pennsylvania, U.S.A.

INTRODUCTION

Magnetic resonance imaging (MRI) techniques offer attractive methods for determining noninvasive temperature changes during thermal therapies for treatment of diseased tissue. For pathological tissues such as tumors, noninvasive or minimally invasive procedures that use heating or cooling are increasingly being employed in clinical applications. Heat is either applied in the form of irreversible tissue ablation, where the temperature of the tissue is increased above 50–60°C for time periods of a few seconds, or hyperthermia in which the tissue temperature is increased to 42–45°C for 30–60 minutes. Cryotherapy/cryosurgery destroys tissue by freezing a local region of tissue over a period of a few minutes. In each case, the treatment protocol aims to attain the required temperature distribution as quickly as possible and to affect only the diseased tissue without damaging surrounding healthy tissue. Noninvasive MRI can be used to locate the position, with respect to the therapy device, of the desired anatomical area for thermal treatment, determine noninvasively the temperature distribution within the target, and also monitor the effects of the treatment after the treatment protocol has finished.^[1–5]

OVERVIEW

Much of the research in clinical MR thermometry has resulted from the desire to monitor temperature changes from new interventional therapeutic treatments using, for example, lasers, radiofrequency (RF) and microwave antennas, cryotherapy probes, or ultrasound transducers.^[3] As a minimally invasive device, laser energy (known as laser interstitial thermotherapy, or LITT) is delivered via a thin, flexible optical fiber to deep-lying lesions or tissues that are usually surgically inaccessible.^[6] In RF ablation, high frequency alternating current is used to thermocoagulate the target tissue through a 1 mm diameter active electrode. Typical RF lesions form an ellipsoidal pattern and are created in 30–60 seconds with temperatures up to 80°C.^[7] For minimally invasive or noninvasive delivery, microwave energy can be delivered to the target through individual antennas or arrays of antennas for either hyperthermia or ablation.^[8] Ultrasound, commonly

used for both hyperthermia and ablation, has been used either as a minimally invasive or noninvasive thermal treatment. One advantage of ultrasound is the ability to treat noninvasively deep tissues and the ability to shape the energy deposition patterns using various transducer designs.^[9] As an alternative to heating tissue, cryosurgery involves the in situ ablation of a target tissue by application of extremely cold temperatures, down to –60°C, using liquid nitrogen.^[10,11]

Determination of the exact location of energy deposition, and associated tissue temperature, in three-dimensional space has been a recurring problem with thermal treatment protocols. Traditionally, invasive thermometric techniques have been used to measure temperatures in the target region. However, these techniques provide only temperature measurements at certain specific points in the target volume. Additionally, devices such as thermocouples can potentially bypass the regions of greatest temperature elevation. MRI methods provide good tissue contrast for detection of pathologies, and excellent visualization of coagulated tissue posttreatment, while MR thermometry provides good spatial and temporal localization of the temperature changes during therapy.

A large number of parameters can be used to measure temperature via magnetic resonance techniques.^[5] For imaging applications, the majority of studies have used temperature-dependent changes in the molecular diffusion coefficient (D) of water, the proton spin-lattice relaxation time (T_1), or the proton resonance frequency (PRF) as a noninvasive thermometer. This article presents the physical basis for the temperature dependence of each of these parameters as well as interventional MRI examples used in conjunction with the clinical thermal therapy techniques described earlier. The reader is also directed to several recent articles and books for further reading on this topic.^[1–5]

TEMPERATURE DEPENDENCE OF MR PARAMETERS

Molecular Diffusion Coefficient

The temperature dependence of the molecular diffusion coefficient of water is described by the Stokes–Einstein

equation:

$$D = D_{\text{inf}} \exp\left(-\frac{E_a}{kT}\right) \quad (1)$$

where D_{inf} is the diffusion coefficient at infinite temperature, k is the Boltzmann constant, T is the temperature measured in degrees Kelvin, and E_a is the activation energy associated with the breaking of hydrogen bonds between water molecules, and is assumed to vary little between different tissue types in medical applications.^[12,13] For temperature imaging, a series of diffusion-weighted images are acquired using a spin-echo sequence, or a fast-imaging sequence preceded by a spin-echo based preparation module. In a particular image, the signal intensity, S , in a voxel is given by:

$$S \propto \exp\left[-\gamma^2 G^2 \delta^2 D \left(\Delta - \frac{\delta}{3}\right)\right] \quad (2)$$

where γ is the gyromagnetic ratio, G is the diffusion encoding gradient strength, δ is the gradient duration, and Δ is the time between the two diffusion-encoding gradient pulses. The pixel-by-pixel value of D is calculated from a series of images, each acquired with a different value of G . Temperature changes from either heating or cooling protocols can be determined based on the acquisition of images at an initial known baseline temperature (T_i) and corresponding diffusion coefficient D_i . The temperature, T_f , at any time during the protocol is given by:

$$T_f = \left(\frac{1}{T_i} - \frac{k}{E_a} \ln \frac{D_f}{D_i}\right)^{-1} \quad (3)$$

where D_f is the corresponding measured diffusion coefficient.

The first imaging study using the temperature dependence of D measured thermal changes in a polyacrylamide gel phantom at 0.5 Tesla using a simple spin-echo sequence.^[13] Using a whole body MR system, temperature images were obtained within the gel, which was heated by an MR-compatible phased array RF applicator. To verify the accuracy of the diffusion technique, a thermocouple was placed adjacent to the heated region. The results indicated that the MR-derived temperature inside a 0.8 cm^2 image region agreed within 0.5°C to that measured by the thermocouple.

One of the drawbacks with conventional diffusion-weighted spin-echo MRI is that the measurement is very sensitive to patient motion. This motion sensitivity can be reduced by using rapid diffusion-weighted echo-planar imaging (EPI), a technique which has been tested on acrylamide gel phantoms and canine brain

in vivo.^[14] Heating of the gel was performed using a surface microwave applicator at 433 MHz and resulted in a temperature sensitivity of the value of D of $3.04 \pm 0.03\%/^\circ\text{C}$ with a temperature resolution of 0.55°C in a 1.8 cm^3 volume. For in-vivo experiments, canine brain was heated with a whole body RF hyperthermia using an annular phased array (130 MHz) and was found to have a temperature sensitivity of $1.9 \pm 0.1\%/^\circ\text{C}$ with a temperature resolution of 0.55°C in a 0.9 cm^3 volume. Although EPI minimizes the effects of patient motion since the data can be collected extremely rapidly, the imaging technique does suffer from a number of disadvantages, including an intrinsically broad point spread function and sensitivity to image artifacts and distortion from both chemical shift and magnetic susceptibility.

Another approach to reducing motion artifacts and data acquisition times, while avoiding the image artifacts associated with EPI, is to use a diffusion-sensitive single-shot rapid acquisition with relaxation enhancement sequence. Using this technique, temperature mapping has been employed for observing temperature changes with in-vivo human calf muscle.^[15] The muscle was cooled using ice, and temperature drops of 18 and 4°C were recorded using a temperature dependence of the diffusion coefficient of $2.4\%/^\circ\text{C}$ in a 2 Tesla clinical magnet. Using both gels and in-vitro muscle samples, temperature changes of $\pm 0.5^\circ\text{C}$ and $\pm 2^\circ\text{C}$, respectively, were detected with a $2 \times 2 \times 8 \text{ mm}$ spatial resolution.

With respect to clinical studies, diffusion-weighted imaging is rarely used in interventional thermal therapies.^[3] In addition to the problems of motion sensitivity and long acquisition times, the temperature dependence of the diffusion coefficient can, in fact, vary with tissue type. Changes in D can also arise from other sources, such as the water content of the tissue, which changes during a heating protocol. Even though there are continuing attempts to improve temperature mapping with diffusion-weighted imaging, T_1 and PRF based techniques are generally the preferred clinical temperature monitoring methods.

Spin-Lattice Relaxation Time

The temperature dependence of the proton T_1 in water (tissue) can be approximated by:

$$T_1 \propto \exp\left(-\frac{E_a}{kT}\right) \quad (4)$$

where E_a is the activation energy for the T_1 process. Although this equation shows an exponential temperature dependence, T_1 is generally linear with T^{-1} over the temperature ranges used for clinical heating



therapies. However, the temperature sensitivity of T_1 varies between tissue types and thus, for each tissue, the T_1 dependence must be known or measured in order to determine accurately the in-vivo tissue temperature. If the tissue is sufficiently heterogeneous, the temperature dependence will be multiexponential, with no corresponding simple linear approximation.

For clinical applications, a pixel-by-pixel estimation of the change in the T_1 value is determined. This can be achieved using a number of different sequences, the most simple being a spoiled gradient-echo sequence in which the image signal intensity (S) is given by:

$$S \propto M_o \frac{\sin \alpha (1 - E_1)}{1 - \cos \alpha E_1} E_2^* \quad (5)$$

where M_o is the net equilibrium magnetization, α is the tip angle, TR is the repetition time between each phase-encoded step, $E_1 = \exp(-TR/T_1)$, T_2^* is the inhomogeneous spin-spin relaxation time, TE is the echo time, and $E_2^* = \exp(-TE/T_2^*)$. Since the temperature dependence of T_2^* is considered negligible, only the temperature dependence of M_o and T_1 contributes to the change in the signal intensity. Although the change in signal intensity is dominated by the temperature dependence of T_1 , the variation in M_o must also be taken into account. Temperature changes are determined by initially acquiring a pretreatment baseline image and then acquiring subsequent images during thermal treatment.

Use of T_1 for temperature imaging was first demonstrated for monitoring local temperature variations in blood samples.^[16] Over a range of 0–40°C, the T_1 was found to vary linearly with T^{-1} with a temperature sensitivity of roughly 1.4%/°C. At higher temperatures, a hysteresis effect was found between the MR signal intensity and temperature. In another study using a neodymium:yttrium–aluminium–garnet (Nd:YAG) laser to heat minced rabbit brain, reversible decreases in M_o and reversible increases in T_1 were found for temperatures up to 45°C.^[17] Above this temperature, up to 60°C, both reversible and irreversible increases in T_1 were found, which were attributed to transformations in cellular metabolism and water distribution within the cell. Above 60°C, only irreversible changes in T_1 were found, corresponding to tissue coagulation and protein denaturation. More recently, a study has been performed to understand the changes in T_1 and other MR parameters (T_2 , diffusion, and magnetization transfer) of tissues due to thermal coagulation using temperatures from 37–90°C.^[18] Using animal and human ex-vivo tissue such as skeletal muscle, cardiac muscle, cerebral hemisphere, white matter, liver, adipose tissue, prostate, and blood, the T_1 value decreased for increasing temperatures except for fat and prostate.

Clinical Applications of T_1 Measurements

While many of the clinical applications of T_1 -based thermometry methods have qualitatively evaluated temperature changes, the application of quantitative temperature imaging using T_1 is increasing. For example, quantitative T_1 thermometry has been used clinically for MR guidance of LITT for inoperable renal tumors.^[19] Delivering the laser therapy via a Nd:YAG source, thermal lesions within kidney were monitored in real time. As the tissue temperature rose, the reduction in signal intensity due to increases in tissue T_1 was determined using continuous-subtraction colorization software, which monitored the development of necrosed tissue. Clinical results for patients with renal carcinoma indicate the feasibility of MR guided laser ablation for inoperable renal tumors.

One important application of quantitative T_1 -based thermometry is for monitoring temperature in breast tissue. As it will be shown in the next section, the PRF technique is limited in its application to adipose tissue since the proton chemical shift in fatty tissue is essentially temperature independent. A recent clinical study investigated the use of focused ultrasound to non-invasively ablate deep malignant breast tumors combined with MR image guidance and on-line MR temperature control.^[20] For treatment of human breast tumors, a focused ultrasound surgery (FUS) therapy unit was designed to operate in a clinical 1.5 Tesla MR scanner. Prior to FUS treatment, T_1 -weighted mammography was performed on the biopsy-confirmed breast cancer. One image is shown in Fig. 1A after injection of a gadolinium-based paramagnetic contrast agent. Strong signal enhancement can be seen in the tumor, revealing a well-circumscribed mass that was located in the central part of the right breast (Fig. 1A). For the FUS treatment, the patient was positioned in the MR scanner and T_1 -weighted images of the breast were acquired perpendicular to the plane of ultrasound propagation. In treatment planning, the target volume was outlined as the MRI visible tumor boundaries to facilitate the pathological analysis of the effects of therapeutic ultrasound on cellular proliferation. Temperature monitoring during the ultrasound sonications used a T_1 -weighted saturation–recovery–turboFLASH sequence. MRI thermometry during FUS with color-coded temperature display for a typical sonication is shown in Fig. 1B. The maximum temperature detected in the breast was approximately $70 \pm 5^\circ\text{C}$; thermometry images were reported to have a temperature resolution of $\pm 2^\circ\text{C}$ and spatial accuracy better than 2 mm. At the end of the sonications, the effect of the ultrasound on the tumor vasculature was evaluated with injected gadolinium contrast agent. Fig. 2B shows a sharply demarcated area with an absence of contrast agent in the treated area, which suggests a complete

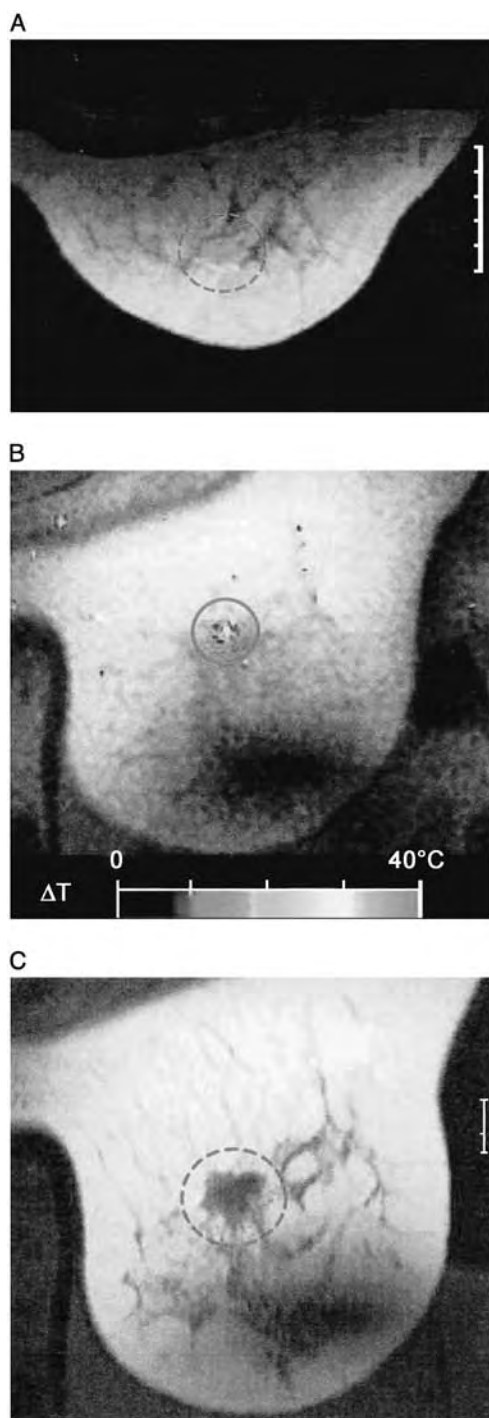


Fig. 1 (A) Before ultrasound treatment, T_1 -weighted images were obtained of the tumor located in the central part of the breast. (B) A temperature map of the spatial heating in the tumor from the focused ultrasound based on the change in the T_1 signal. (C) After FUS treatment and with contrast, a T_2 -weighted image of the breast was acquired exhibiting a complete lack of contrast enhancement after ultrasound therapy. (Reprinted with permission, Copyright 2001, American Association for Cancer Research.)

interruption of blood perfusion, and the area correlates well with the region outlined on the pretherapy MR images. Overall, the results show that clinical non-invasive MRI-guided therapy of breast cancer is both feasible and effective.

One of the drawbacks of using T_1 -based thermometry is the tissue-specific nature of the thermal sensitivity and the complex nature of relaxation in tissue. Above ablation temperatures of 65°C , at which point protein denaturation occurs, changes in T_1 are probably related more to structural changes in the tissue than to the actual temperature.

Proton Resonance Frequency Shift

Temperature information from MR images using the PRF shift method exploits the temperature dependence of the Larmor frequency of protons in water.^[21] The resonant frequency (ω) is given by:

$$\omega(T) = \gamma B_{\text{muc}}(T) = \gamma B_0 \left(1 - \sigma(T) - \frac{2}{3} \chi_v(T) \right) \quad (6)$$

where γ is the proton gyromagnetic ratio, B_{muc} is the local magnetic field experienced by the protons, B_0 is the strength of the external magnetic field, $\sigma(T)$ is the temperature-dependent shielding constant, and χ_v is the temperature-dependent magnetic susceptibility of water. In terms of temperature dependence, the dominant term is the shielding constant, which has a sensitivity of -0.01 parts per million (ppm)/ $^\circ\text{C}$.^[22] The physical mechanism involves the increase in the hydrogen bond length with temperature, which leads to a reduction in the deshielding effect of the oxygen atom.

Use of the PRF for MRI temperature mapping was first introduced by Ishihara et al. in 1992.^[23] Gradient-echo images are used to determine temperature changes, from an initial baseline level, during thermal treatment. To determine the temperature change, phase-sensitive images are processed to estimate the PRF shift. The spatially-dependent temperature change, $\Delta T(x, y, z)$, in a gradient-echo MR image with an echo time, TE , is given by:

$$\Delta T(x, y, z) = \frac{\Delta\phi(x, y, z)}{\gamma \cdot B_0 \cdot \alpha \cdot TE} \quad (7)$$

where $\Delta\phi(x, y, z)$ is the pixel-by-pixel phase change between the baseline image and an image acquired during thermal therapy, and α is a constant (≈ -0.01 ppm/ $^\circ\text{C}$). In practice, the PRF shift is best determined by a complex-subtraction method, instead of a direct phase subtraction, since this process avoids problems of phase wrapping.^[24] Long scan times also require correction of system instabilities such as drift in the

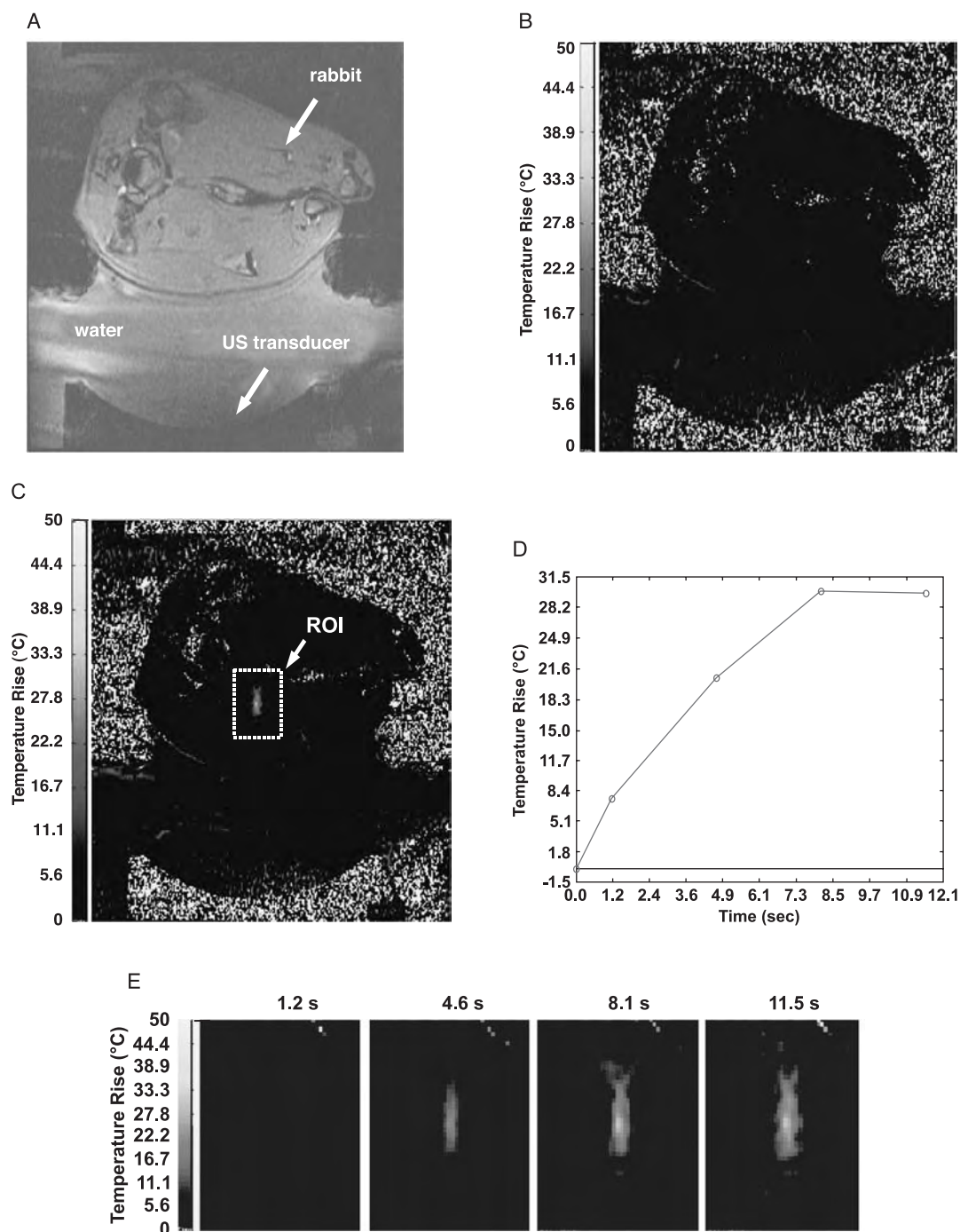


Fig. 2 Magnitude image (A) showing the experimental setup for the delivery of focused ultrasound in rabbit thigh muscle in vivo. Initially, a phase-difference image is obtained before the thermal therapy (B) as a baseline or reference. During the ultrasound sonication deep in the rabbit thigh, the change in phase (C) within the image is used to quantify the temperature change within the ROI. Phase images across the focal point illustrate the temperature elevation during a sonication. The four images (D, left to right) were made at approximately 1.2, 4.6, 8.1, and 11.5 seconds from the beginning of the sonication. As a function of time (E), the temperature rise determined within the ROI is also graphed. (Figures courtesy of Drs. Hynynen and McDannold, Brigham and Women's Hospital, Boston, MA.)

main magnetic field.^[25] First-order-phase-drift correction algorithms use reference phantoms made of oil, which have a very low temperature dependence, $\alpha(T) = -4.59 \times 10^{-5}$ ppm/°C.^[26] Since the drift typically

varies linearly over the imaging plane, spatial interpolation using several phantoms placed around the target can be used to correct for the phase drift in the heated region.

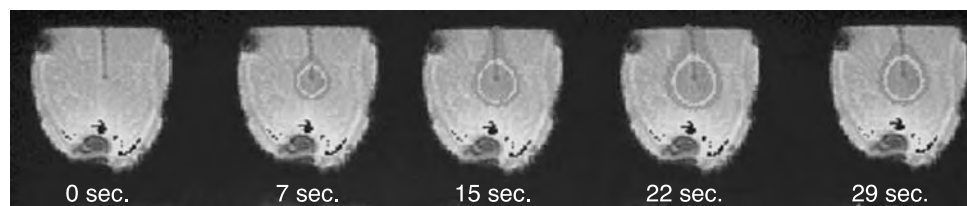


Fig. 3 Spatial and temporal change in the MR temperature mapping during continuous irradiation from a bare end fiber laser interstitial thermotherapy (LITT) for intracranial tumors tested in rabbit brain. (Reprinted with permission, Copyright 2001, John Wiley & Sons, Inc.)

As an example of an in-vivo ultrasound heating procedure monitored using PRF thermometry, Fig. 2A shows a magnitude image of an axial view of a rabbit above a focused ultrasound transducer. The transducer, which was designed to ablate tissue deep within the rabbit's thigh muscle, consisting of a single focused crystal, which had a diameter, radius of curvature, and resonant frequency of 100 mm, 80 mm, and 1.5 MHz, respectively. For this application, the phase shifts were measured using a fast spoiled gradient-echo (FSPGR) sequence. Spherical transducers are known to produce ellipsoidal-shaped heating patterns or lesions,^[27] and the temperature maps shown in Fig. 2 show this type of behavior. The maximum temperature rise within the region of interest (ROI) in Fig. 2C is plotted as function of time in Fig. 2D. Spatial changes in the temperature are plotted as a function of time in Fig. 2E, with both the size of the heated region and the maximum measured temperature increasing with time.

The PRF method has also been used for in-vivo laser heating studies. Lasers have been used to perform thermal ablation in areas such as the brain, head, neck, tongue, liver, kidney, and prostate. Laser designs include variations that can focus energy directly out of the tip (bare end), though a lens, out from a beveled side, or in a diffuse projection pattern. PRF methods have been used to quantify the different temperature profiles produced by each design in rabbit brain.^[28] The results indicated that all of the designs mentioned produced a spherical distribution of elevated temperatures, except the diffuse projection laser, which produced a cylindrical temperature distribution. An example of the spatial heating over time from a bare end fiber design in rabbit brain is shown in Fig. 3. Compared to a fiber optic thermometer also located in the brain, the accuracy of the MR temperature data was reported to be 0.88°C.

With the robustness and rapid data acquisition associated with the PRF technique, evaluation of therapy procedures can be performed close to real time.^[29] One particular application is the use of the MR temperature measurements to control the heating procedure via closed-loop feedback.^[30–32] An example of an in-vivo ultrasound hyperthermia procedure using MR temperature data to control the temperature rise within

tissue is shown in Fig. 4A.^[31] The experiment used a 16-channel ultrasound intracavitary array for localized hyperthermia treatment of prostate tumors. Oil phantoms were placed around the rabbit to correct for drift in the main magnetic field while the ultrasound array was used to heat the thigh muscle of the rabbit. A fiber optic temperature probe was also placed near the ROI to monitor the precision and accuracy of the measurements. MRI-derived temperature measurements within the ROI were used to adjust the power to the transducer through a proportional-plus-integral controller. With the controller programmed to aim for a target temperature of 44°C, the measured steady-state temperatures were within $\pm 1^\circ\text{C}$ of their targets (Fig. 4B). These results demonstrate the feasibility of MR feedback control of hyperthermia treatments.

The advantage of using the PRF method over using the spin-lattice relaxation time is that the temperature coefficient is independent of the type of tissue, unlike the tissue-specific T_1 value. Also, the spatial and temporal resolution of the PRF is superior to that of T_1 thermometry.^[2] A drawback similar to T_1 -based methods is that the PRF technique is susceptible to motion. If the therapy procedure is short, then usually the patient can remain motionless during the treatment and images can be acquired to determine the in-vivo temperatures. However, if the procedure is long, like a hyperthermia treatment, then motion artifacts often arise. One method proposed to reduce the detrimental effects of motion is to acquire multiple echoes rather than a single echo to determine the frequency shift.^[33] The trade-off is that although the sequence improved the temperature measurements in the presence of motion, the temperature sensitivity was reduced. Another problem is that the temperature-dependent magnetic susceptibility is not negligible. Physiologically induced volume magnetic susceptibility changes can arise from blood flow or an altered oxyhemoglobin to deoxyhemoglobin ratio during a hyperthermia treatment. Finally, the mechanism that makes the PRF method attractive for temperature monitoring in muscle, brain, and kidney, for example, cannot be applied to tissue with a high component of fat. One approach to address this problem is to use the PRF technique with a binomial lipid-suppression sequence to remove the

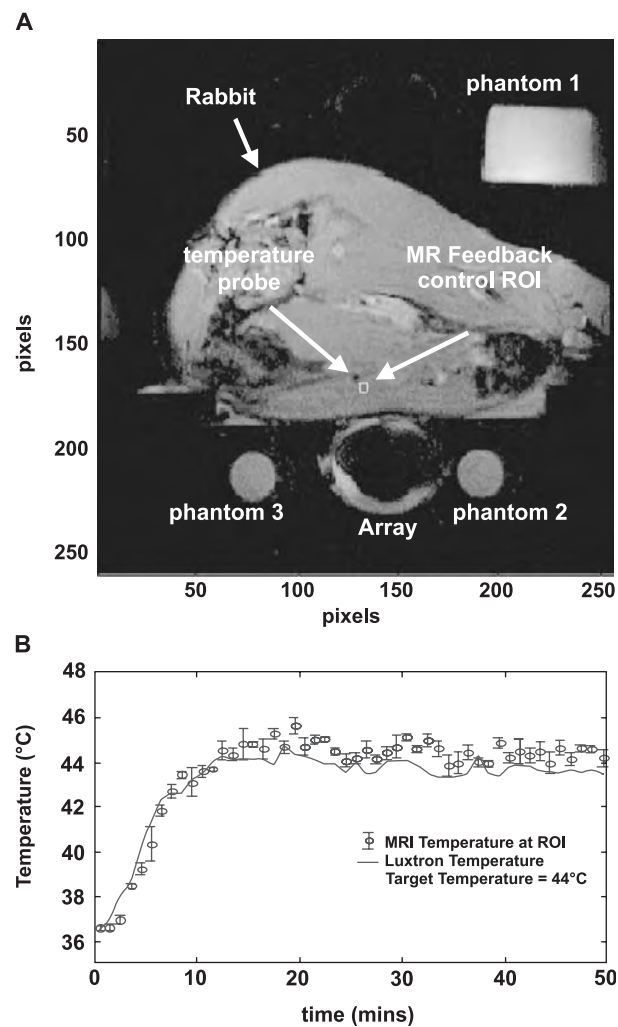


Fig. 4 (A) For in-vivo test of the controller, axial view for a rabbit, and the array along with the three reference oil phantoms used for correcting for the magnetic field drift. (B) The feedback system was set to control the temperatures within the ROI. Temperatures measured by the fiber optic were compared against the MR temperatures determined using the PRF method (mean \pm S.D.) in the ROI. (Reprinted with permission, Copyright 2001, Taylor & Francis Ltd., www.tandf.co.uk.)

effects of adipose tissue from the temperature map,^[34] but considerable care is still needed in interpreting the results.

Clinical Applications of the PRF Method

PRF shift methods have found applications in both animals' experiments and human clinical interventional MR. With MR guidance and temperature control, percutaneous interstitial microwave thermoablation of locally recurrent prostate carcinomas have been performed.^[35] Microwave heating with an

interstitial therapeutic system (Urowave, Dornier MedTech), which consisted of a 915 MHz microwave generator, has been used to increase temperature in the prostate. Each applicator consisted of a 1.2-mm-diameter element housed in a sterile water cooling jacket to prevent carbonization of tissue immediately adjacent to the applicator. The goal of the procedure was to ablate as much as possible of the prostate tissue in the patient by reaching a temperature of at least 50°C, but not exceeding 80°C, for 20 minutes and to preserve the urethra and rectum by not exceeding 45°C in these locations. On-line control monitoring of thermal therapy was performed with a standard fast spoiled gradient-echo sequence modified to support real-time imaging. MR temperature results correlated well with fiber optic temperature probes located on the applicators.

MRI temperature imaging has also been used in conjunction with RF hyperthermia treatments of human high-grade sarcoma of the lower leg.^[36] For the hyperthermia treatment, an RF phased array (130 MHz, 28 cm diameter), which was magnet compatible was designed to fit inside the head coil of a 1.5T whole body scanner. As with much electrical therapy equipment used in conjunction with an MRI scanner, the RF equipment was isolated from the scanner by means of high-order RF filters. Standard gradient echo sequences were used to acquire phase-difference images during the 65-minute-long treatment. Drift in the main magnetic field, B_0 , was corrected using regions of interest within the bone marrow.

One advantage of the PRF shift method over the other thermometry techniques is that the temperature dependence had been shown to be linear above the coagulation or ablation threshold. Thus, this method was used for the first clinical trial of noninvasive MR imaging-guided focused ultrasound surgery of benign fibroadenomas in the breast.^[37] Initially, the tumor was outlined on T_2 -weighted images for multiple sonication treatment planning. Phase-difference imaging was performed during each ultrasound exposure to monitor the change in tissue temperature and focus localization. Results from the temperature mapping indicated that fibroadenoma temperature elevation ranged from 12.8°–49.9°C for the separate sonications. Moreover, the feasibility of this treatment method was demonstrated by the complete or partial lack of contrast material uptake on posttherapy T_1 -weighted images in eight of the 11 lesions treated. Unfortunately, many of the temperature measurements were reported to be unreliable due to patient motion during the sonications. Future treatments may employ such techniques as multiple gradient-echo imaging sequences, mentioned earlier, to overcome motion artifacts.^[33]

Other Clinical Applications and MR Thermometry Methods

One area in which the thermometry methods mentioned previously become problematic is the application to cryotherapy. As the tissue water transitions from the liquid- to solid-state, the signal intensity becomes extremely small in the MR images. This is because the T_2 relaxation of solid ice crystals is so short that, under normal imaging conditions, the ablated area appears as a signal void. In fact, this signal disappearance can be useful since the sharp contrast between the frozen and unfrozen tissue enables a clear estimation of the cryolesion volume.^[10,38] Post cryosurgical treatments have shown good correlation between the MR image lesion volume and the excised necrosed tissue volume. However, temperature measurements within the iceball formed during cryosurgery are not possible unless imaging sequences with very short echo times can be used. Recent successful efforts have quantified the relationship between temperature with ice formation by mapping the R_2^* ($R_2^* = 1/T_2^*$) of bovine liver in vitro.^[11] As the tissue freezes, some extracellular fluid remains in a dehydrated yet unfrozen state from which signal can be measured. With a cryosurgical probe, the temperature was reduced from 37°C to -100°C. Using sequences incorporating a Gaussian half-pulse excitation and a short spiral readout for a very short echo time (0.2 msec) for the first image, while the second image was obtained with a TE of 1.2 msec, allowed determination of the R_2^* . The relationship was found to be linear with temperature down to -25°C.

Besides proton, other heteronuclei such as ^{13}C , ^{15}N , ^{19}F , ^{23}Na , ^{59}Co have been proposed as internal MRI spectroscopic thermometers by measuring temperature-dependent chemical shifts. For example, although not used clinically, tris(ethylenediamine) cobalt(III) trichloride has been used as a temperature sensor due to the large temperature dependence of its ^{59}Co nuclear magnetic resonance chemical shift. Encapsulation of this complex within liposomes was used to target the agent to the reticuloendothelial system, and temperature changes of the order of 0.1°C have been measured in vivo on rats from ultrasound hyperthermia.^[39] An agent that contains two resonances with temperature-dependent chemical shifts, such that the differences in the chemical shifts is also temperature-dependent, can potentially be used to address issues of motion and magnetic susceptibility changes. One example used two fluorine peaks of perfluorooctylbromide to measure the absolute temperature and temperature changes in the liver induced by ultrasound exposure. The chemical was encapsulated in protein microspheres, injected into rats, and subsequently taken up by the macrophages of the liver and spleen.^[40] Though these methods have

been for thermometry with in-vitro and in-vivo animals' experiments, they have yet to be applied for clinical interventional thermal therapies.

CONCLUSIONS

The use of MRI techniques to map temperature during thermal therapies has increased significantly over the past decade. The PRF method provides a robust and easily implemented tool for both measuring and controlling temperature in specific areas for predetermined treatment times. The use of noninvasive or minimally invasive thermal therapies used in combination with MR planning, thermal monitoring, and follow-up has already been shown to be effective in a number of applications.^[3] Interventional MR techniques translate clinically into reduced infections, scar formation, reduced recovery time, and, ultimately, reduced health care costs.

ACKNOWLEDGMENTS

The author would like to thank Dr. Andrew Webb (University of Illinois, Urbana Champaign) for his suggestions.

ARTICLES OF FURTHER INTEREST

Functional MRI: Applications; Magnetic Resonance Imaging (MRI); Magnetic Resonance Microscopy; Medical Imaging, 3-D

REFERENCES

1. Jolesz, F.A.; Young, I.R. *Interventional MR: Techniques and Clinical Experience*; Martin Dunitz Ltd.: London, 1998.
2. Quesson, B.; De Zwart, J.A.; Moonen, C.T. Magnetic resonance temperature imaging for guidance of thermotherapy. *J. Magn. Reson. Imaging* **2000**, *12* (4), 525–533.
3. McDannold, N.J.; Jolesz, F.A. Magnetic resonance image-guided thermal ablations. *Top. Magn. Reson. Imaging* **2000**, *11* (3), 191–202.
4. Germain, D.; Chevallier, P.; Laurent, A.; Saint-Jalmes, H. MR monitoring of tumour thermal therapy. *Magma* **2001**, *13* (1), 47–59.
5. Webb, A.G. Temperature Measurements Using Nuclear Magnetic Resonance. In *Annual Reports on NMR Spectroscopy*, 45 Ed.; Webb, G., Ed.; Academic Press, Elsevier Science Ltd.: London, 2002; 1–67.
6. Kettenbach, J.; Kacher, D.F.; Koskinen, S.K.; Silverman, S.G.; Nabavi, A.; Gering, D.; Tempny, C.M.; Schwartz, R.B.; Kikinis, R.; Black, P.M.; Jolesz, F.A.



- Interventional and intraoperative magnetic resonance imaging. *Annu. Rev. Biomed. Eng.* **2000**, *2*, 661–690.
7. Farahani, K.; Black, K.L.; DeSalles, A.A.; Anzai, Y.; Lufkin, R.B. Radiofrequency Ablation. In *Interventional MR*; Joles, F.A., Young, I.R., Eds.; Martin Dunitz Ltd.: London, 1998; 153–157.
 8. Hoffmann, A.L.; de la Rosette, J.J.; Wijkstra, H. Intraprostatic temperature monitoring during transurethral microwave thermotherapy: Status and future developments. *J. Endourol.* **2000**, *14* (8), 637–642.
 9. Diederich, C.J.; Hynynen, K.H. Ultrasound technology for hyperthermia. *Ultrasound Med. Biol.* **1999**, *25* (6), 871–887.
 10. Tacke, J. Thermal therapies in interventional MR imaging. Cryotherapy. *Neuroimaging Clin. N. Am.* **2001**, *11* (4), 759–765.
 11. Butts, K.; Sinclair, J.; Daniel, B.L.; Wansapura, J.; Pauly, J.M. Temperature quantitation and mapping of frozen tissue. *J. Magn. Reson. Imaging* **2001**, *13* (1), 99–104.
 12. Chang, D.C.; Rorschach, H.E.; Nichols, B.L.; Hazlewood, C.F. Implications of diffusion coefficient measurements for the structure of cellular water. *Ann. N.Y. Acad. Sci.* **1973**, *204*, 434–443.
 13. Le Bihan, D.; Delannoy, J.; Levin, R.L. Temperature mapping with MR imaging of molecular diffusion: Application to hyperthermia. *Radiology* **1989**, *171* (3), 853–857.
 14. Macfall, J.; Prescott, D.M.; Fullar, E.; Samulski, T.V. Temperature dependence of canine brain tissue diffusion coefficient measured in vivo with magnetic resonance echo-planar imaging. *Int. J. Hypertherm.* **1995**, *11* (1), 73–86.
 15. Il'yasov, K.A.; Hennig, J. Single-shot diffusion-weighted RARE sequence: Application for temperature monitoring during hyperthermia session. *J. Magn. Reson. Imaging* **1998**, *8* (6), 1296–1305.
 16. Parker, D.L.; Smith, V.; Sheldon, P.; Crooks, L.E.; Fussell, L. Temperature distribution measurements in two-dimensional NMR imaging. *Med. Phys.* **1983**, *10* (3), 321–325.
 17. Jolesz, F.A.; Bleier, A.R.; Jakab, P.; Ruenzel, P.W.; Huttl, K.; Jako, G.J. MR imaging of laser-tissue interactions. *Radiology* **1988**, *168* (1), 249–253.
 18. Graham, S.J.; Stanis, G.J.; Kecojevic, A.; Bronskill, M.J.; Henkelman, R.M. Analysis of changes in MR properties of tissues after heat treatment. *Magn. Reson. Med.* **1999**, *42* (6), 1061–1071.
 19. de Jode, M.G.; Vale, J.A.; Gedroyc, W.M. MR-guided laser thermoablation of inoperable renal tumors in an open-configuration interventional MR scanner: Preliminary clinical experience in three cases. *J. Magn. Reson. Imaging* **1999**, *10* (4), 545–549.
 20. Huber, P.E.; Jenne, J.W.; Rastert, R.; Simiantonakis, I.; Sinn, H.P.; Strittmatter, H.J.; von Fournier, D.; Wannemacher, M.F.; Debus, J. A new noninvasive approach in breast cancer therapy using magnetic resonance imaging-guided focused ultrasound surgery. *Cancer Res.* **2001**, *61* (23), 8441–8447.
 21. Hindman, J.C. Proton resonance shift of water in the gas and liquid states. *J. Chem. Phys.* **1966**, *44*, 4582–4592.
 22. Kuroda, K.; Tsutsumi, S. Temperature Imaging by Proton Chemical Shift. In *Non-invasive Thermometry*; Miyakawa, M., Bolomey, J.C., Eds.; CRC: Boca Raton, 1994; 64–102.
 23. Ishihara, Y.; Calderon, A.; Watanabe, H.; Mori, K.; Okamoto, K.; Suzuki, T.; Sato, K.; Kuroda, K.; Nakagawa, N.; Tsutsumi, S. A precise and fast temperature mapping using water proton chemical shift. *Proc. SMRM* **1992**, 4803.
 24. Chung, A.H.; Hynynen, K.H.; Cline, H.E.; Colucci, V.; Oshio, K.; Jolesz, F.A. Optimization of spoiled gradient-echo phase imaging for in vivo localization of a focused ultrasound beam. *Magn. Reson. Med.* **1996**, *36* (5), 745–752.
 25. De Poorter, J.; DeWagter, C.; De Deene, Y.; Thomsen, C.; Stahlberg, F.; Achten, E. The proton resonance frequency shift method compared with molecular diffusion for quantitative measurements of two dimensional time dependant temperature distributions in a phantom. *J. Magn. Reson., Ser. B* **1994**, *103*, 234–241.
 26. Kuroda, K.; Suzuki, Y.; Ishihara, Y.; Okamoto, K.; Suzuki, Y. Temperature mapping using water proton chemical shift obtained with 3D-MRSI: Feasibility in vivo. *Magn. Reson. Med.* **1996**, *35*, 20–29.
 27. Fry, F.J. Intense focused ultrasound in medicine. *Eur. Urol.* **1993**, *23* (Suppl. 1), 2–7.
 28. Atsumi, H.; Matsumae, M.; Kaneda, M.; Muro, I.; Mamata, Y.; Komiya, T.; Tsugu, A.; Tsugane, R. Novel laser system and laser irradiation method reduced the risk of carbonization during laser interstitial thermotherapy: Assessed by MR temperature measurement. *Lasers Surg. Med.* **2001**, *29* (2), 108–117.
 29. Quesson, B.; Vimeux, F.; Salomir, R.; De Zwart, J.A.; Moonen, C.T. Automatic control of hyperthermic therapy based on real-time Fourier analysis of MR temperature maps. *Magn. Reson. Med.* **2002**, *47* (6), 1065–1072.
 30. Vimeux, F.C.; De Zwart, J.A.; Palussiere, J.; Fawaz, R.; Delalande, C.; Canioni, P.; Grenier, N.; Moonen, C.T. Real-time control of focused ultrasound heating based on rapid MR thermometry. *Invest. Radiol.* **1999**, *34* (3), 190–193.
 31. Smith, N.B.; Merrlees, N.K.; Dahleh, M.; Hynynen, K.H. Control system for an MRI compatible intracavitary ultrasound array for thermal treatment of prostate disease. *Int. J. Hypertherm.* **2001**, *17* (3), 271–282.
 32. Behnia, B.; Suthar, M.; Webb, A.G. Closed-loop feedback control of phased-array microwave heating using thermal measurements from magnetic resonance imaging concepts in magnetic resonance. *Magn. Reson. Eng.* **2002**, *15* (1), 101–110.
 33. Mulkern, R.V.; Panych, L.P.; McDannold, N.J.; Jolesz, F.A.; Hynynen, K. Tissue temperature monitoring with multiple gradient-echo imaging sequences. *J. Magn. Reson. Imaging* **1998**, *8* (2), 493–502.
 34. Kuroda, K.; Mulkern, R.V.; Oshio, K.; Panych, L.P.; Nakai, T.; Moriya, T.; Okuda, S.; Hynynen, K.; Jolesz, F.A.; Joles, F.A. Temperature mapping using the water proton chemical shift: Self-referenced method with echo-planar spectroscopic imaging. *Magn. Reson. Med.* **2000**, *43* (2), 220–225.

35. Chen, J.C.; Moriarty, J.A.; Derbyshire, J.A.; Peters, R.D.; Trachtenberg, J.; Bell, S.D.; Doyle, J.; Arrelano, R.; Wright, G.A.; Henkelman, R.M.; Hinks, R.S.; Lok, S.Y.; Toi, A.; Kucharzyk, W. Prostate cancer: MR imaging and thermometry during microwave thermal ablation-initial experience. *Radiology* **2000**, *214* (1), 290–297.
36. Carter, D.L.; MacFall, J.R.; Clegg, S.T.; Wan, X.; Prescott, D.M.; Charles, H.C.; Samulski, T.V. Magnetic resonance thermometry during hyperthermia for human high-grade sarcoma. *Int. J. Radiat. Oncol. Biol. Phys.* **1998**, *40* (4), 815–822.
37. Hynnen, K.; Pomeroy, O.; Smith, D.N.; Huber, P.E.; McDannold, N.J.; Kettenbach, J.; Baum, J.; Singer, S.; Jolesz, F.A. MR imaging-guided focused ultrasound surgery of fibroadenomas in the breast: A feasibility study. *Radiology* **2001**, *219* (1), 176–185.
38. Gilbert, J.C.; Onik, G.M.; Rubinsky, B. MRI-Guided Tissue Ablation Using Cryosurgery. In *Interventional MR*; Joles, F.A., Young, I.R., Eds.; Martin Dunitz Ltd.: London, 1998; 159–167.
39. Smith, N.B.; Webb, A.G.; Ellis, D.S.; Wilmes, L.J.; O'Brien, W.D., Jr. Experimental verification of the theoretical in vivo ultrasound heating using cobalt detected magnetic resonance. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **1995**, *42* (4), 489–491.
40. Webb, A.G.; Smith, N.B.; Ellis, D.S.; O'Brien, W.D., Jr. Non-invasive in vivo temperature mapping of ultrasound heating using fluorine-based magnetic resonance imaging agent **1995**, *1*, 1609–1612.



Magnetic Resonance Microscopy

Egidijus Uzgiris

Rensselaer Polytechnic Institute, Troy, New York, U.S.A.

INTRODUCTION

What do we mean by the term magnetic resonance (MR) microscopy? MR imaging is typically done with an in-plane resolution of about 1 mm^2 and slice thickness of 1–3 mm (this resolution can be routinely increased to $300 \times 300\ \mu\text{m}^2$ in-plane resolution with 1-mm slice thickness) at magnetic field strengths of 1.5 T (or, more frequently now, at 3 T) in imaging times of the order of 1 to 10 min or so, depending on the precise nature of the imaging sequence. The field of view in MR images is typically $20 \times 20\ \text{cm}^2$, or larger in the case of whole body imaging with parallel coil signal receive arrangements. The magnetic field gradients used to generate the spatial resolution are on the order of 10–20 mT/m. At this clinical scale of resolution, the soft tissue contrast is superb and superior to all other imaging modalities. Higher resolution is desirable, however, in some clinical applications—in evaluating the microvasculature of tumors, in detecting amyloid plaque in Alzheimer’s patients, in detecting atherosclerotic plaque in arteries, and in detecting small tumorous lesions. These structures and lesions range from 10 to 100 μm in the first case, from 40 to 200 μm for amyloid plaques, and up to 500 μm for small tumor lesions. Typically, the highest resolution in a clinical scanner may be on the order of $300 \times 300 \times 1000\ \mu\text{m}^3$, which fails to meet the above clinical needs (except for tumor detection at the half-millimeter size range). Thus, MR microscopy may be defined as MR imaging at yet a higher resolution than the best clinical case or, specifically, a resolution of at least 100 μm or better. This partitioning into two classes of imaging occurs in practice because of the need for specialized gear to generate the necessary high magnetic field gradients and specialized receive coils for capturing the weak signals from each imaging voxel, as well as the use of very high magnetic fields, 9–14 T.

As we shall see in this review, it is a daunting task to generate MR images with such high resolution and the commensurate small imaging voxels. In fact, we shall demonstrate that *in vivo* MR microscopy will be practical only in rodent studies or in embryology studies for which high-field, small-bore magnets and high-field gradient coils exist. *Ex vivo* imaging may push the resolution limits of MR microscopy to the cellular level, but only with rather difficult

implementations of the limiting geometries of the necessary coil hardware and still very long imaging times: not minutes, but many hours. The fundamental limits of detecting weak radiofrequency signals generated by an ever smaller number of precessing nuclear spins per imaging voxel are not, in the end, to be circumvented.

Can we expect any movement toward meaningful clinical applications such as those listed above? Perhaps only in one or two of the applications will progress with adapted clinical MR systems yield possible resolution (application of gradient coils and specialized receive coils in clinical MR systems magnets) high enough to meet a clinical need. Certainly in preclinical studies of rodents we may expect MR microscopy to develop much valuable information (such as the phenotyping of mice or mouse embryos). There may be a role in developmental studies as well. In *ex vivo* studies, MR microscopy may provide useful delineations of tissue lesions and abnormalities that could complement histology (in terms of organization of cells or cellular structures). Finally, techniques such as labeling with suitable contrast agents may allow detection of a few cells at time and their migration and localization in organs of interest using standard or adapted clinical MR imaging (MRI) systems.

How then can we judge progress in this field? Are efforts that push the resolution limits by the use of ultrasmall receive coils and very strong gradient coil assemblies relevant to important biological or clinical applications? Such efforts are technically admirable but are very distant from the main applications listed above. High-resolution mouse phenotyping by MR microscopy was one important achievement. We must determine whether atherosclerotic plaque can be adequately visualized—with or without contrast agent. Can tumor microvasculature be visualized directly by MR microscopy? Will contrast agent approaches allow the visualization of cell trafficking? Will *in vivo* amyloid plaque imaging be at all possible and clinically relevant? I examine these questions below.

SIGNAL TO NOISE IN MR IMAGING

The physics of MR imaging is discussed in any number of monographs^[1,2] in great detail. Briefly, the MR signal is derived from precessing nuclear spins (spins of

protons of water) that are prepared by a suitable radiofrequency pulse excitation in a magnetic field. The time-varying magnetic fields produced by this precession are sensed by a suitable receive coil (placed around or near the sample) as an induced voltage at the precession frequency. Because the precession rate depends on the field strength, the spatial position of a particular precession frequency is encoded by applying a magnetic field gradient in a suitable fashion in all three spatial directions. Thus, for high spatial delineation, it is necessary to apply high magnetic field gradients. This becomes a hardware problem that was solved for small-bore systems. But the spin signals are weak and noise from the receiver coil (the Johnson noise) and equivalent resistive losses in the sample under study are limiting factors in generating adequate image quality.

What are the barriers to high resolution with MR? Fundamentally, the principal barrier is the loss of signal in direct proportion to the voxel volume—the loss of the number of participating proton nuclear spins that generate the MR signal in that voxel, in direct proportion to the voxel volume. To illustrate this barrier, consider a voxel of $400\ \mu\text{m}^3$ (the limiting resolution for typical clinical imaging) shrinking to $100\ \mu\text{m}^3$ for the case of MR microscopy (in our classification criteria). The signal loss for the latter case is by a factor of $4^3 = 64$. To alleviate this loss with increased averaging time (signal increases linearly with averaging time, whereas the noise increases as the square root of averaging time) requires $64^2 = 4096$ longer imaging times to obtain the same signal-to-noise ratio (SNR). Thus, an imaging time of some 10 min in the clinical case becomes 68 hr in the microscopy case. Clearly, without some other helping factors, that would be the end of it. Imaging at such a high resolution would not be at all practical. But one such factor is the magnetic field, and in MR microscopy the fields used are typically 7 to 14 T, increasing detected signal up to 50 over equivalent voxel size at 1.5 T. Another helping factor is the receive coil configuration for the smaller volume microscopy samples. Together these conspire to give measurement times of about 14 hr for 110- μm voxel resolution at 2 T for whole-mouse imaging (ex vivo and fixed in formalin).^[3] Increasing the magnetic field increases the resolution further for the same measurement time: at 7.1 T a resolution of 50 μm was achieved. Thus, it is clear that MR microscopy is technically hard and even with high fields the necessary measurement times for acceptable SNR quickly explode to impracticality as resolution approaches the biological cell size. Let us keep in mind that the example above is for a formalin-fixed animal. Imaging in vivo has further issues of the effects of breathing motion on the signals and the challenge of keeping animals under anesthesia for

prolonged periods of time. Progress on both of these fronts has been made by gating strategies and animal monitoring in real time. However, for these and other reasons, resolution much below 100 μm in living animals is unlikely.

Another approach may allow visualization of groups of cells even with voxel sizes that are large compared to individual cells: the use of contrast agents to label the cells. A signal amplification results, which allows detection of the cells in those voxels containing the cells. Without resolving the cells individually, one can still detect them and follow their migration in situ in tissues and organs of interest. This is likely to be an important application of MR imaging, but, strictly speaking, it is not MR microscopy by the criteria of resolution limits used in such studies (clinical resolution may be adequate).

The reader may refer to any number of monographs and research articles for a detailed consideration of MR signal to noise.^[2,4,5] As always in MRI, the detected signals are weak and the most challenging barrier for faster imaging or for higher resolution is that of signal to noise. An image must have some minimum value of SNR in order to detect interesting features. For example, an SNR of less than 20 is inadequate for detecting amyloid plaques in brain sections.^[6] The scaling of SNR with the principal experimental variables can be expressed straightforwardly. To this end, and including explicitly only those variables of significance for scaling, the signal to noise of an MRI experiment may be expressed as

$$\text{SNR} \propto A(\text{seq}) \times f(B_0) \times \frac{V}{s} \sqrt{T_{\text{img}}}, \quad (1)$$

where $A(\text{seq})$ is a factor involving pulse sequence details (which may be optimized experimentally), V is the voxel volume, s is the imaging scale size, i.e., the diameter of the receive solenoid or receive coil, T_{img} is the imaging time, and $f(B_0)$ is the magnetic field dependence, which we take to be

$$f(B_0) = B_0^{\frac{7}{4}}. \quad (2)$$

This field dependence assumes that the dominant noise source is the Johnson noise of the receive coil and not the resistive losses of the sample. This assumption seems to hold in many high resolution cases compared below. The important factor of receiver coil scale size, s , can be understood in that the induced current in a coil for a given rate of magnetic field change within the coil is inversely proportional to coil diameter: the smaller the coil diameter, the larger the induced voltage in the coil for a given rate of magnetic field change (see Ref. [4] for detailed discussion of size scaling).



Given fixed values for the first three terms, the coil size, the field, and the pulse sequence, the resolution size and imaging time are the only remaining variables. For some necessary SNR level to adequately distinguish image features, the necessary imaging time rapidly blows up to

$$T_{\text{img}} \propto \frac{1}{A^2 f^2 d^6}, \quad (3)$$

where d is the voxel size. Even with improvements in coil sensitivity and the use of high fields, it is apparent that the last term causes T_{img} to blow up quickly as d becomes less than $100 \mu\text{m}$. As a result, in vivo MR microscopy remains extremely challenging, even for the case of rodent brain studies. Not only must the imaging time utilized in ex vivo studies of 10–30 hr be reduced to something like 1–2 hr, compatible with animal anesthesia exposure, but also the effects of breathing and cardiac motion present yet another set of severe technical impediments.

For clarity, consider a further example. In the imaging of mouse brains,^[6] the voxel sizes used were $50 \times 50 \times 500 \mu\text{m}$ in 2–3 hr for an SNR of about 20, the lower limit for acceptable SNR. To go to the next step, the imaging of a majority of amyloid plaques in rodent brains, for example, one would require voxels sizes not larger than $50 \times 50 \times 50 \mu\text{m}^3$. The imaging time would increase to 200 to 300 hr, which is clearly not practical in live animals. Where can the necessary factor of 100–300 come from? The field is already high in this example and no large improvement can be gained at any further practical field level. The coil sensitivity is as high as the geometry of mouse brain allows. Are we at an impasse for in vivo studies? Not entirely. There are several gambits to pursue in the near term and perhaps more exotic detection schemes in the future.

What can, perhaps, be done is to use high-temperature superconducting (HTS) coils to help reduce

the noise further.^[7,8] It was estimated previously that SNR is improved by a factor of 5–6 over copper receive coils in some experimental situations. These situations seem to require limiting the electromagnetic loading of the sample on the HTS coil. This is hard to do when attempting to image a mouse brain, but perhaps achievable when imaging smaller structures such as embryos. HTS coils are implemented in nuclear MR spectroscopy to good effect, but implementation of such coils for MR imaging has not been widely implemented, as discussed below.

The remaining gambit, after all other optimization knobs have been turned, is to apply contrast agents to ex vivo samples or to live animals. The contrast agents have the effect of reducing tissue longitudinal relaxation times, T_1 , and lead directly to larger signals with certain common pulse sequences for imaging (although such gains are only of the order of factors of 2 or so). This approach has been applied ex vivo to visualize amyloid plaques in 1–2 hr^[9] and it may be that such contrast agents can lead to adequate signal amplification in vivo as well. More exotic approaches have been suggested for MR signal detection and we shall consider these below.

What are the experimental realities? We have considered signal scaling relationships and determined they have some serious limitations for MR microscopy. Have experimenters been able to circumvent these limitations by clever approaches to SNR enhancement? We consider several applications below and see that, indeed, progress is being made and clever solutions exist to the problem of obtaining adequate signal to noise in imaging at high resolution.

For projecting practical resolution limits in MR imaging, it is instructive to compare key experimental data points of SNR thus far achieved with MR microscopy in Table 1.

This comparison of achieved SNR spans many orders of magnitude in voxel size, and the predicted

Table 1 SNR for various experimental conditions in microscopy and the clinic case

Reference	Field (T)	Volume, voxel (mm ³)	Scale size (mm)	Time (hr)	SNR	SNR ^a scaled	Pulse sequence	Sample
10	14.1	2×10^{-7}	0.5	1	8	4	SE	Leaf stem
11	7	1.8×10^{-4}	20	1	70	30	SE	Insect
12	14.1	1.2×10^{-5}	15	18	27	32	T_2^*	Brain section
6	7	2.4×10^{-4}	15	5.6	96	110	T_2^*	Brain section
6	7	1.2×10^{-4}	15	5.6	42	56	T_2^*	Brain section
6	7	5.9×10^{-5}	15	10.9	27	38	T_2^*	Brain section
Clinic ^b	1.5	1	300	.008	20	60	SE	Head

^aSNR scaled in accordance with Eq. 1 using the results of Lee et al.,^[10] but reduced by a factor of 2 to account for the T_1 shortening they invoked by applying CuSO_4 to the sample.

^bStandard head scan; SNR is the estimated range.

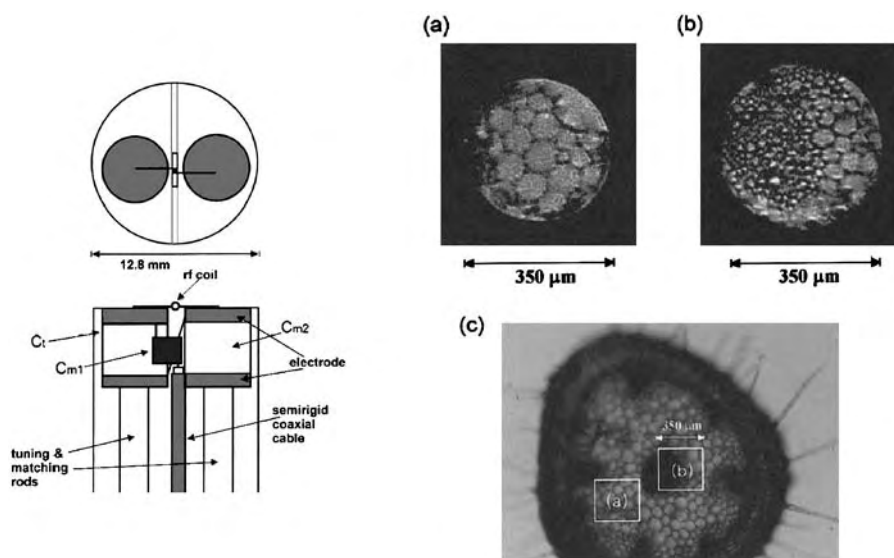


Fig. 1 One-micrometer in-plane resolution experimental setup at 14 T (left). Note the sample is placed inside the radiofrequency coil of about 0.5-mm diameter. (a and b) MR images of geranium leaf stem: voxel size, $2 \times 2 \times 50 \mu\text{m}^3$, imaging time, 56 min. (c) Optical micrograph of leaf stem. Reproduced by permission from Lee et al.,^[10] Figs. 1 and 5.

scaling behavior of the SNR by Eq. 1 is within a factor of ~ 2 of the observed value in the worst case and better than that in the remaining comparisons. Eq. 1 works quite well in predicting SNR behavior, especially if the extrapolations are for optimum pulse sequence conditions. In Table 1, the different pulse sequences were not factored in the comparisons. The best SNR conditions in detail were achieved by Wecker et al.^[11] in their study of insects. In Table 1, Eq. 1 was used to extrapolate the necessary FOV/receive coil size for achieving a given SNR and the resulting resolution limit using a field of 14 T and an imaging time of 1 hr.

For biological systems, cellular resolution (i.e., voxel size less than $10 \mu\text{m}$ on one side) is achieved with acceptable SNR and reasonable imaging time, as illustrated by the work of Lee et al.^[10] at 14 T and a “microscopic” receive coil of 0.5-mm diameter. This rather specialized hardware, as well as the high-resolution image of a leaf stem, is illustrated in Fig. 1. At a more general size scale we are constrained above the cellular resolution limit, even for a low SNR of 10. Nevertheless, MR imaging at a resolution somewhat larger than the biological cell scale size would be valuable in the study of insect morphology and embryology development. Furthermore, future staining techniques may bring out more contrast in the images and increase the available SNR.

HTS RECEIVE COILS

In cases where the sample resistive losses are small, HTS coils in place of copper coils provide a way of increasing the SNR by a factor of ~ 6 .^[7,8] For a voxel size of $3.8 \times 10^{-5} \text{ mm}$, or $17 \times 17 \times 134 \mu\text{m}^3$, Hurlston et al.^[8] achieved an SNR of 29 at 9.4 T for a kidney

sample in 1 hr and a rectangular coil size of $10 \times 15 \text{ mm}$. When this result is scaled in accordance with Eq. 1, the expected SNR for the condition of Wecker et al.^[11] is 61, not very different from the SNR of 70 reported by Wecker et al. for a copper coil of $\sim 20\text{-mm}$ diameter. The samples are not the same and the sequences are similar, but not the same. Nevertheless, this rough comparison suggests that the HTS coil geometries are not as ideal as the configurations that can be fabricated from copper. The evident advantage of HTS in exact comparisons with copper is perhaps reduced in comparison with more optimum copper coil configurations of the same size. There is also the question of signal homogeneity in the HTS coil configuration (see, for example, Fig. 2 in Ref. [8]). This approach is promising in principle, but in practice there are challenges and further work to be done to understand the limitations of the HTS coil systems.

PHENOTYPING OF THE MOUSE

Perhaps the first useful application of MR microscopy was demonstrated in the high-resolution imaging of intact fixed mice for the purpose of phenotyping.^[3] The relationship of genome and structure and function is of vital importance for biological understanding of the many mouse phenotypes generated by genetic alterations, such as in discerning the function of knockout genes. The whole mouse was imaged *ex vivo* at 2 T at a resolution of $110 \mu\text{m}$ isotropically in 14.5 hr. The issue of signal sensitivity was addressed with efficient signal encoding and efficient array of receive coils, but these in themselves were insufficient to yield adequate SNR. The key step in this study was the use of “active” staining with a T_1 contrast agent, together with



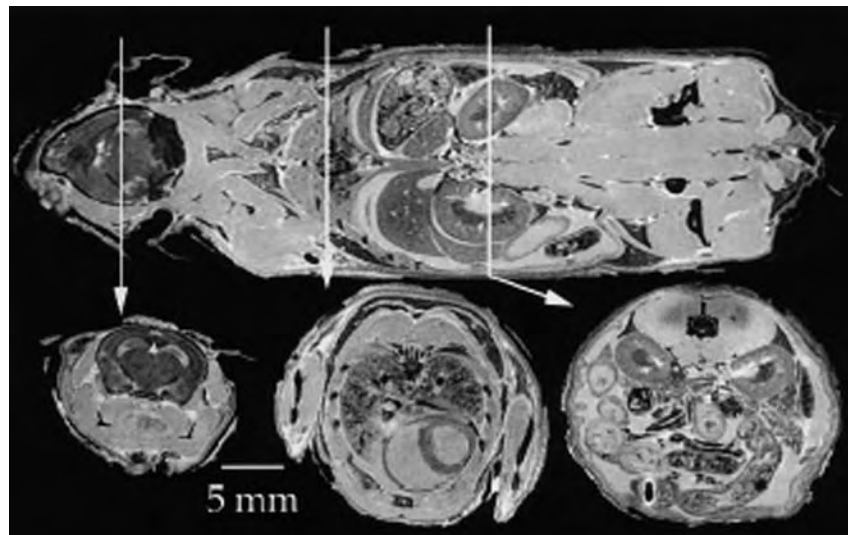


Fig. 2 MR image of whole fixed mouse at 2 T. Voxel size is 110 μm isotropically; imaging time is 14.5 hr. Cross-sections of 110- μm thickness are indicated by arrows. Reproduced by permission from Johnson et al.,^[3] Fig. 3.

formalin fixation by perfusion of the mouse through the jugular vein. The reduction in T_1 by almost a factor of 6 in most tissues gave a useful high-resolution image of the mouse and a data set that could be manipulated in many ways to give desired views of the mouse anatomy, as illustrated in Fig. 2. This study required more than turning “resolution knobs,” as the authors state. The large data sets required separated reconstruction methods and further visualization methods to render rather low SNR images of thin slices into more useful thick-slice rendered images, as illustrated in Fig. 3. Thus, the effective resolution of anatomical structures even with staining is much larger than the voxel size used in data acquisition.

Interpretation of structures at a level of 100 μm from such images is challenging. The effects of staining on the tissues are not known in detail and the visualized anatomical structures may require further studies for correct interpretation. The simplest effect of such staining is the delineation of blood vessels. The question remains, At what level of microvasculature structure can MR microscopy be applied reliably? The

images in Fig. 3 indicate that microvasculature below some 100- μm size is not yet effectively visualized and, thus, application to tumor biology and angiogenesis imaging may require different staining procedures (such as the use of intravasculature staining only) and significant amplification of SNR—which is already pushed to the limit, even in fixed and stained tissue samples.

At high fields, one is generally limited to small bore sizes and studies at fields beyond about 7 T will be used for tissue imaging and perhaps for ex vivo organ imaging, as done here. The severity of the SNR problem in MR microscopy is quite evident in these ex vivo images and the extension of this microscopy to in vivo studies is daunting.

MR MICROSCOPIC HISTOLOGY AND AMYLOID PLAQUES

A less formidable application for MR microscopy is that of histology. The samples are small, suitable for

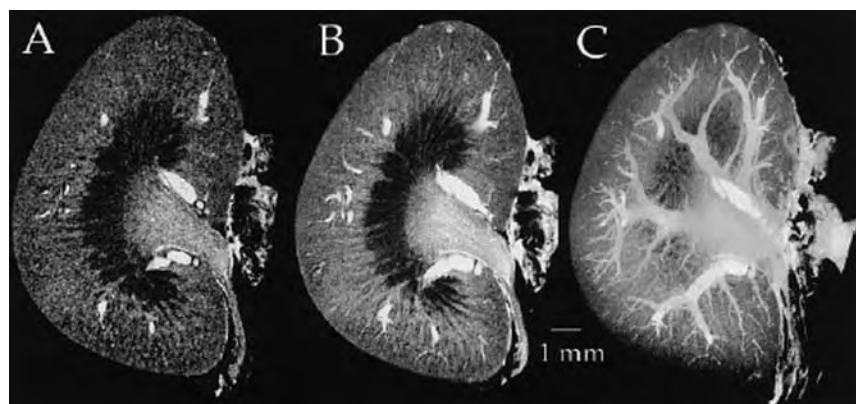


Fig. 3 MR images of Gd-perfused mouse kidney at 9 T with an array size of $256 \times 512 \times 512$, yielding a slice thickness of 25 μm (A). A rendering technique was used to produce (B) a 250- μm -thick section and (C) a 2500- μm -thick section. Imaging time, 14.5 hr. Reproduced by permission from Johnson et al.,^[3] Fig. 5.

very high field systems and for efficient receive coils of small diameter. We shall now examine an important area of research, determining plaque burden in transgenic mice models of Alzheimer's disease. Can amyloid plaques be visualized reliably and give a useful measure of disease progressions in these animal models? If yes, then this would indeed be a significant contribution to the study of this disease. Alternative histological methods exist, and it is still important to demonstrate that the MR technique brings new information or information in a new form (digital encoding, for example, and non-destructive imaging) over the standard optical methods.

It has now been confirmed that an earlier study of human amyloid plaques in the brain^[6] did indeed detect at least a fraction of the amyloid plaques in the *ex vivo* brain sections on T_2 imaging at high spatial resolution. The source of the contrast is thought to be iron deposition in plaques (through an unspecified mechanism). This deposition is age dependent, present in older mice and not in younger mice, and may occur only for a subset of larger plaques.

Benveniste et al.^[6] in a careful study of postmortem Alzheimer's brain sections, delineated the resolution and SNR limits for detection of plaques by MR microscopy. They used T_2^* weighting to visualize hypointense spots in Alzheimer's brain sections but not in control, disease-free samples. The identification was through coregistration of MRI sections with histology sections that were stained for amyloid. The coregistration of histology sections with MR image sections is technically challenging, as demonstrated in this and all subsequent studies of plaque visualization. Errors in registration present real issues of image interpretation in detail. However, no hypointense spots were observed in disease-free sections, giving credence to the interpretation that the spots are associated with amyloid plaques and the hypointensity is most likely caused by associated iron deposits in the plaques leading to susceptibility effects that manifest themselves as loss of signal in T_2^* -weighted imaging. The images were taken at a field of 7.1 T and required 14 hr of acquisition time to achieve an SNR ≈ 30 , a necessary limit for plaque visualization. Importantly, the resolution also had to be of the order of $40 \times 40 \times 40 \mu\text{m}^3$ for the hypointense spots to be reliably discerned. At coarser resolution, the spots became indistinct owing to the partial volume effect associated with voxels larger than the objects of interest.

This early result was encouraging and an impetus for numerous groups to join in this effort. However, this is still a long way from *in vivo* imaging. For example, in 1–2 hr acceptable SNR can be achieved for $50 \times 50 \times 500 \mu\text{m}^3$ at 7.1 T for live mouse brain imaging. To achieve the extra factor of 10 in voxel size reduction and maintain an acceptable SNR, one would

need 200-hr imaging times, clearly untenable for live animals. Because this comparison is already for a high field, the remaining “knobs” are in part more efficient receive coils, more efficient pulse sequences, and the use of contrast agents. We shall examine all of these approaches below. Furthermore, the effects of motion as a result of breathing and cardiac output must also be addressed in order to obtain the necessary high resolution without loss of SNR caused by blurring and other motion artifacts.

Wadghiri et al.^[9] used specific contrast agents to bring out plaque conspicuity at a relatively coarse resolution of $78 \times 78 \times 500 \mu\text{m}^3$ in a ~ 1 -hr scan time at 7 T for transgenic mice. In their study either a Gd or Mion agent was linked to A β 1–40 peptides, which bind strongly to amyloid plaques. These constructs were injected intraarterially, together with mannitol, and the brain scans commenced 6 hr later. In the absence of agents, no spots were observed in the brains of the transgenic mice *in vivo* and *ex vivo*, and no spots were observed in wild-type mice *in vivo* and *ex vivo*. *Ex vivo* MR images were registered with histologic slices and amyloid staining to isolate a few coregistered spots as an indication of specificity for amyloid plaques.

This work was followed by that of Jack et al.^[13] Remarkably, they extended *in vivo* imaging of mouse brains for amyloid plaques without the use of contrast agents using a field of 9.4 T with a pair of small surface coils and a reduced FOV. An optimized spin-echo sequence was used with very careful procedures for mitigating breathing and cardiac motion effects. In transgenic mice, APP/PS1, they were able to visualize numerous hypointense spots at a resolution of $60 \times 60 \times 120 \mu\text{m}^3$ that they could coregister with *ex vivo* images and *ex vivo* histologic slides stained specifically for amyloid with thioflavin S, as illustrated in Fig. 4. Importantly, no spots were observed for control animals *in vivo*, *ex vivo*, or in histologic stains.

How is it that these workers were able to surmount the SNR barriers first indicated by Benveniste et al.^[6] With careful gating and attention to movement artifacts they may have been able to accept lower SNR and still visualize plaques. More efficient pulse sequences also contribute to reduced imaging time. The use of small surface coils and a limited FOV are other significant differences from the earlier work. All in all, a factor of about 10 or so in reduced imaging time was obtained with somewhat coarser resolution.

In all of these studies of amyloid plaques, there is a characteristic that remains troubling. Some spots coregister with stains for amyloid, but most do not. Thus, the specificity for amyloid appears low. Second, many more spots exist in histology stains than appear in the *ex vivo* MR images (see Fig. 4), so the sensitivity is also limited. This would suggest that the MR microscopy images may have limited use in longitudinal studies of



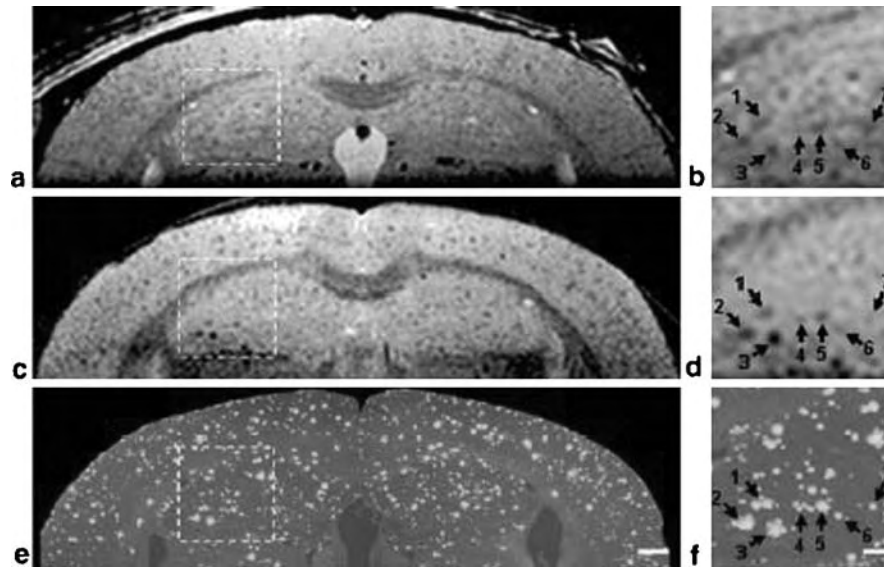


Fig. 4 Three-way correlation of AD mouse brain. (a, c, and e) Full FOV images; (b, d, and f) magnified subregions. (a and b) In vivo images; (c and d) ex vivo images; and (e and f) histological section. Voxel size, $60 \times 60 \times 120 \mu\text{m}^3$, imaging time, ~ 1 hr. Hypointense spots are correlated with histological stain spots (arrows). Reproduced by permission from Jack et al.,^[13] Fig. 5.

pathology or in following therapy response. However, such spots are not observed in control animals. The link to the amyloid gene function in the transgenic animals seems to exist. Perhaps some associated structures in the brains of these animals cause the hypointense spots to appear—these could be microhemorrhages, microvessels, or some sort of filamentary structures associated with this phenotype.

In one study at 11.7T using T_2 imaging and very high resolution of $\sim 23 \mu\text{m}$ isotropically, hypointense spots could not be associated with iron staining (not all that surprising because not all plaques are associated with iron uptake) in ex vivo samples of Alzheimer patients' brains.^[12] They associated hypointensity in all but one case, a rather large 100- μm plaque, with blood vessels and associated blood cell aggregates. This kind of phenomenon may explain the somewhat poor sensitivity and specificity in MR microscopy studies of Alzheimer's brains.

A more convincing demonstration of plaque imaging was obtained by staining in a passive protocol, i.e., soaking ex vivo samples in a Gd-based contrast agent.^[14] The idea in this study was to increase signal to noise by lowering the T_1 of tissue. Because plaques would not take up the agent, a susceptibility difference would develop between the plaques and the surrounding tissue that could be detected more readily and even for plaques smaller than the imaging voxel size (by virtue of the fact that the susceptibility effect has a considerable extension in space). Hypointense spots were delineated only in APP/PS1 mouse strains, with spot density increasing with age and in expected brain areas. The hypointense spots could be coregistered with Congo Red staining (a specific marker of amyloid plaques). The imaging parameters were $\sim 60 \times 60 \times 60 \mu\text{m}^3$ voxel sizes, with 3-D gradient echo

sequence and imaging time of 7–9 hr. Without the passive staining the SNR and plaque contrast was insufficient to reveal specific hypointense spots, as might be expected from the earlier work of Benveniste et al.^[6] and the considerations of SNR and resolution.

CELLS AND GENE EXPRESSION

Visualization of biological cells directly has not been performed, except in the case of the large cylindrical cells found in the leaf stem.^[10] The majority of biological cells of interest are of the order of $10 \mu\text{m}$ in size and there is, as yet, no practical way to get sufficient SNR to visualize such structures. There is also the issue of contrast. What distinguishes a cell from its surrounding medium are differences in terms of MR contrast parameters: T_1 , T_2 , diffusion, proton density. These parameters may not differ much from the interior to the exterior of cells and thus may not offer much contrast, especially at high resolution and low SNR. One way to ameliorate these difficulties is by the use of contrast agents. It was suggested that T_2 agents, superparamagnetic iron oxide particles (SPIO), if internalized by the cell, would, by the previously mentioned susceptibility effect, spread out their magnetic field influence many times the diameter of a single cell. A single cell may thus be visualized at a coarser resolution than the cell dimension. The cell would be detected but not resolved, an acceptable objective for many applications, in particular in studies of cell trafficking and cell lineage development in embryology. Ex vivo imaging of single cells labeled in this manner was successful at visualizing increasingly more dilute solutions of labeled cells. As demonstrated above, in the absence of motion and with small

imaging volumes, adequate SNR can be achieved in reasonable imaging times. However, *in vivo* detection of single cells remains problematic. Nevertheless, groups of cells can be followed in studies of cell migration in living animals. For example, the fate of injected stem cells is of great importance in current clinical investigations. In a recent study, stem cell migration was followed *in vivo* over a 3-week period in a mouse model of stroke using iron oxide particles.^[15] Visualization of the migrating cells was performed at 7 T with a $78 \times 78 \times 78 \mu\text{m}^3$ voxel volume in 1 hr with a SNR of 8–10, sufficient to pick up hypointense voxels in the migrating path. With an efficient cell incorporation technique, these investigators were able to saturate the MRI susceptibility effect so that they could follow cells with no loss of apparent contrast over 10 cell divisions after the initial incorporation. They estimate that for *in vivo* imaging they were able to detect clusters of cells as small as 100 cells, with the rather short imaging time given above. They were able to do so in part by using a specialized small receive coil and in part by the large contrast of the cell clusters relative to background enabled by the high uptake of the dextran-coated iron oxide particles.

An extension of the above approach is to use very large, micrometer-size iron oxide polymer particles to get larger susceptibility effects, which could, in principle, darken a large voxel with just a single cell within.^[16] This contrast agent was able to efficiently label hematopoietic cells without apparent alteration of cell function. An important extension of this approach is to visualize stem cells and their fate after injection in blood or tissue and possibly to do so in a practical manner, with clinical scanners and readily adapted hardware. This would be an important clinical application advance derived from microscopic imaging studies of this kind. The sensitivity of SPIO labeling, in principle, can be quite high, approaching a detectability in the femtomolar range, that of positron emission tomography or single photon emission computed tomography sensitivities, but with submillimeter resolution.^[17]

Louie et al. conducted an elegant study of gene expression by MR microscopy.^[18] They constructed a novel contrast agent responsive to β -galactoside (β -gal) and injected the contrast agent at the two-cell development stage in *Xenopus laevis* and mRNA for β -gal in just one of the two cells. Subsequently, the descendants of the cells with β -gal became bright in T_1 -weighted images (11.7 T, $27 \times 16 \times 16 \mu\text{m}^3$, 3–4 hr image time) in the developed embryo.

TUMOR MICROVESSELS

Visualization of tumor angiogenesis, the development of new neovasculature, is an important clinical need

for prognosis and therapy monitoring. The microvasculature is composed of vessels as small as $10 \mu\text{m}$ in diameter and therefore these structures, most numerous in a developing tumor, are to date not directly visualized by clinical MR or MR microscopy. Their delineation in tumor biology is important in understanding how antiangiogenesis therapy proceeds. This need is an important challenge for MR microscopy that could be addressed in preclinical animal tumor models. In a recent study by Kiessling et al.,^[19] a correlation was demonstrated between the very fine detailed rendition of tumor vascular density by 3-D histological reconstruction and MR angiography data with an intravascular agent ($160 \times 200 \times 320 \mu\text{m}^3$ at 1.5 T) and MR vascular volume parameter data at a much coarser resolution, as illustrated in Fig. 5. Understanding the very fine structure and how it might respond to therapy is of principle importance. Can it be achieved with MR microscopy for a rodent tumor model?

As illustrated in Fig. 5, the resolution required to visualize individual vessels must be much higher than that of the angiography data set indicated. Again, we may apply high-field, small FOV and small surface coils and contrast agents to this problem and see how far the visualization can be pushed. From the work of Wecker et al.,^[11] we may extrapolate to a 7-T system, suitable for animal imaging, and a 1.5-mm size scale, again suitable for tumor imaging in rodents and contrast agent signal amplification. With 1-hr imaging, the delineation of tumor microvessels in great detail should be possible: With an acceptable SNR of 10 achieved through the use of a long circulating contrast agent with high signal amplification of the vessels, it should be possible to achieve $\sim 25\text{-}\mu\text{m}$ isotropic resolution in an imaging time of ~ 1 hr. Thus, with further rendering techniques for 3-D representation of thick sections, the microvessel structures of rodent tumors could be substantially visualized with available hardware and imaging techniques.

VISUALIZING BIOLOGICAL CELLS

Assuming some contrast is available for distinguishing a cell from its surroundings, such as the extracellular matrix, the resolution requirement of $10 \times 10 \times 10 \mu\text{m}^3$ for the delineation of single cells possesses severe limits on required imaging time and available FOV. Table 2 illustrates that for some reasonable imaging time of the order of 1 hr, only a very high field with a microscopic coil arrangement such as that used by Lee et al.^[10] can approach this resolution. The use of contrast agent to help the SNR may alleviate the size constraints, perhaps by factors of 2 or so. Here, the attempt is to resolve the cells and not merely detect them in a tissue



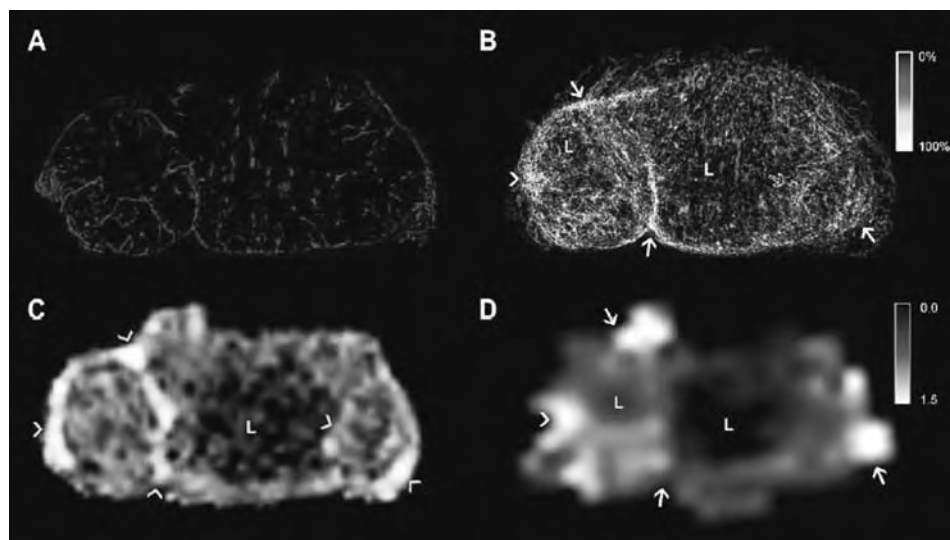


Fig. 5 Images of mouse tumor, HaCaT-ras-A5RT3. (A) Immunofluorescence images of vasculature at 5- μ m thickness. (B) Reconstructed, color-coded vessel volume density (VVD) for a 2-mm-thick section. (C) MR angiography image at 2 mm thickness and (D) an MR parameter map of vascularization. Large vessel bundles in the VVD map are picked up in the angiography image. Arrowhead in D corresponds to the vessel bundle in B (arrowhead). The fine vascular structure in the VVD map is not resolved in the angiography image or the MR parameter map. Reproduced by permission from Kiessling et al.,^[19] Fig. 4.

Table 2 Isotropic resolution in micrometers for a magnetic field of 14 T and 1-hr imaging as a function of imaging scale size and SNR

Scale size (mm)	SNR = 10	SNR = 35
0.5	5.8	9.7
5	10	21
10	12	26
15	13.7	30

milieu, for which the use of susceptibility agents can render cells visible in a coarse voxel resolution, as demonstrated previously.

MOLECULAR IMAGING

Molecular imaging is a term, much in vogue currently,^[20] used to describe *in vivo* imaging of molecular targets in tissues or of molecular processes, such as gene expression in specific tissues. In these efforts, high resolution is not strictly necessary because in many cases the targets are distributed over a reasonably large volume, such as for a tumor of >1 mm in diameter. However, in other cases, such as for the molecular imaging of atherosclerotic plaque, visualization requires if not microscopic resolution, at least sub-millimeter resolution. This is a challenge for *in vivo*

imaging with clinical scanners with limited imaging times in the face of motion artifacts.

The principle challenge for MR in molecular imaging is one of sensitivity. MR contrast agents, small T_1 or T_2 agents, must be in the 10s of micromolar to even the millimolar range (per active ion, Gd for a T_1 agent, or per Fe in an iron oxide Mion T_2 agent) to give detectable MR signal changes, whereas many *in vivo* targets are in the nanomolar to femtomolar range. This mismatch is not addressed by higher resolution, but rather with amplification schemes (many agent molecules per ligand binder or by increase of agent relaxivity, although there are physiochemical limits to such improvements). We described one scheme of large iron oxide particles to give sufficient amplification to detect single cells in *ex vivo* imaging. In some cases the detectability may approach the femtomolar range in Fe.^[17] However, the particles in this case are large, ~ 100 nm or so, and do not readily distribute in the extravascular space to reach targets beyond those of the blood circulation.

ATHEROSCLEROTIC PLAQUES

The imaging of atherosclerotic plaques is a big and challenging enterprise that is beyond the scope of this review. I mention this topic only to emphasize a very great clinical need that conventional MR imaging meets only partially. Atherosclerotic plaques may be of the order of approximately millimeters in size and are,

therefore, not essentially microscopic. But their classification into vulnerable and non-vulnerable classes may push the limits of available SNR and available contrast in clinical protocols. Thus, the techniques of MR microscopy and molecular imaging may contribute to this clinical problem. In small animal models of atherosclerosis, MR microscopy techniques will be essential.

CONCLUSIONS

MR microscopy has demonstrated much potential for clinical and fundamental biological studies. The truly high magnification range in optical microscopy is exceedingly difficult to achieve with MR because of the inherent low signals associated with proton nuclear spins. At present, only one study resolved cells in a biological sample^[10] and it required extremely high fields and very specialized, miniaturized gradient and signal receive coils, not to mention the choice of large cell structures in leaf stems for visualization. Can the SNR barrier be somehow surmounted to yield more practical visualization of biological cells? This problem requires a solution for MR microscopy to be truly microscopic in a more general way. Perhaps it is insurmountable as viewed by current proton staining techniques and a truly clever advance is required. For example, imaging schemes involving hyperpolarized xenon are being investigated for the promise of large sensitivity gains.^[21] Nevertheless, even at a lower resolution than that of the cell, MR microscopy has been demonstrated to be valuable in the visualization of Alzheimer's amyloid plaques in brain, in rodent brain studies, in following cell migration in vivo, and in rodent and embryo phenotyping, to name just the most prominent studies.

ARTICLES OF FURTHER INTEREST

Functional MRI: Applications; Magnetic Resonance Imaging (MRI); Magnetic Resonance Imaging in Temperature Measurement

REFERENCES

1. Kuperman, V. *Magnetic Resonance Imaging, Physical Principles and Applications*; Academic Press: New York, 2000.
2. Vlaardingbroek, M.T.; den Boer, J.A. *Magnetic Resonance Imaging*, 2nd Ed.; Springer: Berlin, 1999.
3. Johnson, G.A.; Cofer, G.P.; Gewalt, S.L.; Hedlund, L.W. Morphologic phenotyping with MR microscopy: The visible mouse. *Radiology* **2002**, *222*, 789–793.
4. Cho, Z.H.; Ahn, C.B.; Juh, S.C.; Lee, H.K.; Jacobs, R.E.; Lee, S.; Yi, J.H.; Jo, J.M. Nuclear magnetic resonance microscopy with 4- μm resolution: Theoretical study and experimental results. *Med. Phys.* **1988**, *15*, 815–824.
5. Callaghan, P.T. *Principles of Nuclear Magnetic Microscopy*; Oxford University Press: Oxford, UK, 1991.
6. Benveniste, H.; Einstein, G.; Kim, K.R.; Hulette, C.; Johnson, G.A. Detection of neuritic plaques in Alzheimer's disease by magnetic resonance microscopy. *Proc. Natl. Acad. Sci. USA* **1999**, *96* (24), 14079–14084.
7. Black, R.D.; Early, T.A.; Roemer, P.B.; Mueller, O.M.; Mogro-Camparo, A.; Turner, L.G.; Johnson, G.A. A high-temperature superconducting receiver for nuclear magnetic resonance microscopy. *Science* **1993**, *259*, 793–795.
8. Hurlston, S.E.; Brey, W.W.; Suddarth, S.A.; Johnson, G.A. A high temperature superconducting helmholtz probe for microscopy at 9.4T. *Magn. Reson. Med.* **1999**, *41*, 1032–1038.
9. Wadghiri, Y.Z.; Sigurdsson, E.M.; Sadowski, M.; Elliot, J.I.; Li, Y.; Scholtzova, H.; Tang, C.Y.; Aguinaldo, G.; Papolla, M.; Duff, K.; Wisniewski, T.; Turnbull, D.H. Detection of Alzheimer's amyloid in transgenic mice using magnetic resonance microscopy. *Magn. Reson. Med.* **2003**, *50*, 293–302.
10. Lee, S.-C.; Kim, K.; Kim, J.; Lee, S.; Yi, J.H.; Kim, S.W.; Ha, K.-S.; Cheong, C. One micron resolution NMR microscopy. *J. Magn. Reson.* **2001**, *150*, 207–213.
11. Wecker, S.; Hornschemeyer, T.; Hoehn, M. Investigation of insect morphology by MRI: Assessment of spatial and temporal resolution. *Magn. Reson. Imaging* **2002**, *20*, 105–111.
12. Dhenain, M.; Prival, N.; Duyckearts, C.; Jacobs, R.E. Senile plaques do not induce susceptibility effects in T_2 weighted MR microscopic images. *NMR Biomed.* **2002**, *15*, 197–203.
13. Jack, C.R., Jr.; Garwood, M.; Wengenack, T.M.; Borowski, B.; Curran, G.L.; Lin, J.; Adrany, G.; Grohn, O.H.; Grimm, R.; Poduslo, J.F. In vivo visualization of Alzheimer's amyloid plaques by magnetic resonance imaging in transgenic mice without a contrast agent. *Magn. Reson. Med.* **2004**, *52*, 1263–1271.
14. Dhenain, M.; Delatour, B.; Walczak, C.; Volk, A. Passive staining: A novel ex vivo MRI protocol to detect amyloid deposits in mouse models of Alzheimer's disease. *Magn. Reson. Med.* **2006**, *55* (3), 687–693.
15. Hoehn, M.; Kustermann, E.; Blunk, J.; Wiedermann, D.; Trapp, T.; Focking, M.; Arnold, H.; Hescheler, J.; Fleischmann, B.K.; Schwandt, W.; Buhrle, C. Monitoring of implanted stem cell migration in vivo: A highly resolved in vivo magnetic resonance imaging investigation of experimental stroke in rat. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 16267–16272.
16. Hinds, K.A.; Hill, J.M.; Shapiro, E.M.; Laukkanen, M.O.; Silva, A.C.; Combs, C.A.; Varney, T.R.; Balaban, R.S.; Koretsky, A.P.; Dunbar, C.E. Highly efficient endosomal labeling of progenitor and stem cells with large magnetic particles allows magnetic resonance imaging of single cells. *Blood* **2003**, *102*, 867–972.



17. Heyn, C.; Bowen, C.V.; Rutt, B.; Foster, P.J. Detection threshold of single SPIO-labeled cells with FIESTA. *Magn. Reson. Med.* **2005**, *53*, 312–320.
18. Louie, A.Y.; Huber, M.M.; Ahrens, E.T.; Rothbacher, U.; Moats, R.; Jacobs, R.E.; Fraser, S.E.; Meade, T.J. In vivo visualization of gene expression using magnetic resonance imaging. *Nat. Biotech.* **2000**, *15*, 321–325.
19. Kiessling, F.; Le-Huu, M.; Kunert, T.; Thorn, M.; Vosseler, S.; Schmidt, K.; Hoffend, J.; Meinzer, H.P.; Fusenig, N.E.; Semmler, W. Improved correlation of histological data with DCE MRI parameter maps by 3D reconstruction, reslicing and parameterization of the histological images. *Eur. Radiol.* **2005**, *15*, 1079–1086.
20. Weissleder, R. Molecular imaging in cancer. *Science* **2006**, *312*, 1168–1171.
21. Schroder, L.; Lowery, T.J.; Hilty, C.; Wemmer, D.E.; Pines, A. Molecular imaging using a targeted magnetic resonance hyperpolarized biosensor. *Science* **2006**, *314*, 446–449.

Mass Transfer in Tissue Function: Roles

E. N. Lightfoot

Karen A. Duca

Virginia Bioinformatics Institute, Virginia State University, Blacksburg, Virginia, U.S.A.

INTRODUCTION

From a philosophical standpoint living organisms can be considered transport–reaction systems that have as one of their primary aims the preservation and reproduction of particular genomes. However, all true organisms exhibit a markedly hierarchical structure, and each level in this structure has its own transport–reaction characteristics. Genome-based reactions appear only in the last stages of this hierarchy. Biochemical reactions, almost always mediated by enzymes, can be surprisingly fast and normally do not limit metabolic rates. Mass transport mechanisms are, however, few, and two are dominant in all but the smallest systems: convection, or flow induced transport, and various forms of concentration diffusion. These are very difficult to augment—or to slow down. As a result the structure and function of living organisms is determined to a very large degree by mass transfer considerations.

However, almost all physiological processes are “designed” by evolutionary processes with a safety factor on the order of 10, and as a result mass transfer limitations only manifest themselves under conditions of extreme stress in the healthy organism. The result is a structural hierarchy paralleled by a time constant hierarchy notable for marked time constant separation. This latter characteristic greatly simplifies physiological analysis. The situation is far different under pathological conditions, and this lack of orderly behavior can complicate the treatment of disease. Moreover, the slow progression of many chronic diseases is hidden by the safety factor of physiological design until functional capacities are much reduced.

While the study of physiological mass transfer began centuries ago, major new insights are still emerging. Early macroscopic descriptions are still useful and will be included here. However, based on technological advances, interest has progressed toward even smaller length scales, and our discussion will follow this same path. Because of space limitations we must satisfy ourselves with short overviews of the topics covered and with references limited largely to major reviews and monographs. A particularly useful summary that covers many unique biological processes, e.g., facilitated or hindered diffusion, excitatory phenomena, and inter-cellular communication, is provided by Keener and

Sneyd.^[1] A more detailed introduction to mass transfer is available in Bird, Stewart, Lightfoot.^[2]

SPATIAL AND TEMPORAL HIERARCHIES

Finding one’s way requires a hierarchical organization and here we use the pronounced spatial hierarchy of living systems. There are six well-defined levels, representing different length scales, which we will explore: whole organisms (m), organ systems (dm), functional units (10–100 μm), cells (1–10 μm), formed elements (0.1–1 μm), and molecular clusters (nm).

Major Organ Systems

Near the top of the hierarchy are the major organs interacting directly with the external environment via a wide variety of mass transport processes: the gut, kidneys, lungs, and skin. The gut absorbs nutrients through the small intestine and discharges solid wastes through the large intestine or colon. The kidneys discharge liquid wastes while the lungs perform a dual role. They provide for oxygen input from the air and discharge of a number of gases, primarily carbon dioxide and water vapor. The skin plays a relatively minor role in mass transfer, but sweating is an important means for release of metabolic heat. The descriptions of organ architecture at the macroscopic level are the province of anatomy, a subject that goes back to Leonardo da Vinci.^[3]

These and other major organs are connected by the blood circulation in an organ network suggested in the pharmacokinetic diagram given in Fig. 1. In addition to the above external organs we will be particularly interested in the brain and the liver that along with the kidneys are the most active metabolically. Examples of metabolic activity are shown in Fig. 2, and more fully described in Schmidt-Nielsen (Ref. [4], p. 94) and Krebs^[5] with units extracted from examination of the original source: ml O₂ (sTP)/100 g dry mass (min). These are tissue slice data that tend to be smaller than in vivo data, but they are sufficient for our present purposes. In particular, they show that metabolic rates



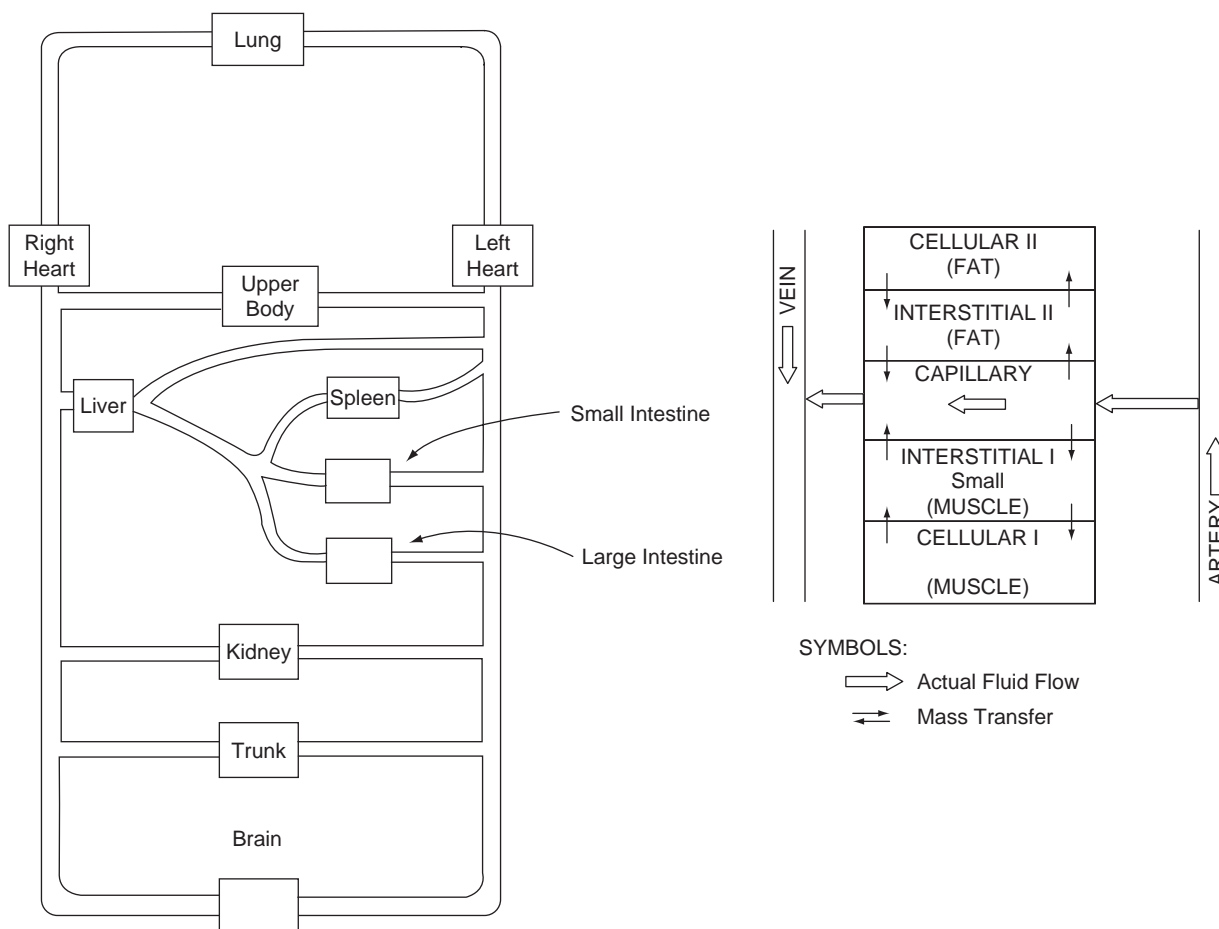


Fig. 1 Pharmacokinetic approximations.

per unit mass tend to decrease with increase in animal size.

The sizes of these organs vary, surprisingly systematically, with body size, but they are always large enough to see with the unaided eye, and the distances between them are large.^[6] It is common practice

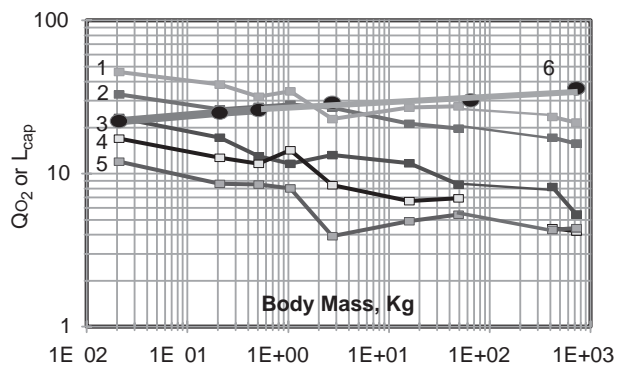


Fig. 2 Specific oxygen consumption and capillary spacing: (1) brain, (2) kidney, (3) liver, (4) spleen, (5) lung [ml O₂ (sTP)/100 g dry mass (min)], and (6) intercapillary spacing for cerebral gray matter (μm). (From Refs. [4,18].)

to express macroscopic body properties, P_i , relative to body mass, M , by an allometric equation of the form:

$$P_i = a_i M^{b_i} \quad (1)$$

and some widely used examples of interest are shown in Table 1. Mass is expressed in kilograms. A wide variety of similar information is available in specialized references and in physiology texts.^[4,6,7,10,11] There have been many attempts to explain the mechanistic basis of the allometric equations in terms of underlying physiology, some quite recent, but none has been completely convincing. They should be considered as useful empiricisms, and it should be recognized that some are based on inadequate sampling.

An instructive example is provided by the first entry of Table 1, fasting whole-body metabolic activity. The first form of this correlation, for a fasting basal state, and prominently reported in a number of sources, is almost certainly badly skewed by misuse of large-animal data (Ref. [10], p. 195). A more thorough study by White and Seymour strongly suggests the second form, adjusted to a temperature of 36.2°C.^[7] The latter

Table 1 Commonly accepted mammalian allometric relations

Item	Allometry
Oxygen consumption [L(STP)/hr]	$0.676M^{0.75}$ (commonly accepted) $0.444M^{0.67}$ (Ref. [7]) ^a
Cardiac output (L/min)	$0.187M^{0.81}$
Blood volume (L)	$0.055M^{0.99}$
Kidney mass (kg)	$0.021M^{0.85}$
Liver mass (kg)	$0.082M^{0.87}$
Lung volume (L)	$0.063M^{1.02}$

^aSee Refs. [8,9] for details of the conditions used.

investigators also discuss the problems of comparing animals of different normal body temperatures and the ambiguities of determining a common reference state. Both correlations are an attempt to reduce a large mass of data taken under a variety of conditions to a single allometric equation, and one is well advised to read the White and Seymour paper, and the literature they cite, carefully before using either. Similar remarks are appropriate for the other table entries.

Functional Units of Organs

We shall find in the next section that diffusion is far too slow to supply nutrients or to eliminate wastes at the size scales of whole organs. As a result, organs are subdivided into large numbers of parallel structures or functional units where the metabolic reactions of the organism actually take place. These form the parenchymal tissues, and they are supported by structural elements, branching blood vessel structures acting as manifolds and the like. The functional units take a wide variety of forms, and describing them is the province of histology.^[12] A classic case, of parallel capillaries, is shown in Fig. 3: grey matter in the brain.^[13] The circular cylinders are approximations to the actual hexagonal prisms and are widely used in transport modeling. For a more detailed cross section see Kessel (Ref. [12], fig. 12.18, p. 269). The spacing between capillaries, which defines the sizes of the corresponding functional units, is quite regular. Moreover, as suggested in line 6 of Fig. 2, it varies little with body size. This spacing results from a compromise between diffusional access and structural constraints.

Thus, we shall find shortly that in typical functional units rates of blood perfusion, lateral diffusion to the parenchymal tissue and chemical reaction in this tissue are in balance with each other. Moreover, capillary density tends to be limited by as yet unidentified material constraints. They represent an intermediate

situation between the organ networks, where diffusion is too slow to be significant, and cell interiors where it typically needs to be suppressed.

Whole Cells

Cells are one of the defining features of true living organisms, and cellular transport and reaction processes are among the most active of research areas today.^[14,15] They come in almost infinite variety: Alberts et al. estimate that there are more than 10 million living species on the earth, and most of them are unicellular. However, the cells share a number of common features, and most are of about the same size: 10 μm to perhaps 50 μm in diameter for animal cells and a few micrometers for bacteria. All are surrounded by a cell membrane with highly selective permeability, and their interiors, or cytoplasm are the sites of most metabolic reactions. As a result the cytoplasm is not only crowded, but it must provide for segregation of this myriad of reactions. Diffusion can now often be considered inconveniently fast, and cells have developed a wide variety of interior structures, surrounded by additional selective membranes, and other means to channel individual reaction sequences spatially. Even the simple prokaryotic cell interiors of bacteria are very crowded.^[16] In bacteria there is, however, minimal organized structure above the level of protein or DNA molecules.

We shall shortly find this crowding to lower diffusion rates substantially, especially for larger entities such as protein molecules. Cell membranes also have elaborate structures based on bilayer lipid membranes to limit transport between the cell interior or cytoplasm and the exterior. Penetrating these bilayers are a wide variety of protein complexes to permit, or actively transport, specific entities into or out of the cell.

Intracellular Organelles

Eukaryotic cells, characteristic of all higher organisms, from yeasts and amebae to fungi, plants, and animals, contain large numbers of substructures, e.g., peroxisomes, endoplasmic reticulum, Golgi apparatus, endosomes, and lysosomes, as suggested in Fig. 4. The mitochondria are particularly important in producing the high-energy phosphate bonds that power most metabolic reactions (Ref. [14], chap. 14). Microtubules permit transport of large entities over distances comparable to cell dimensions. New transport mechanisms come into our story at this point, and we shall lay the groundwork for describing them shortly.



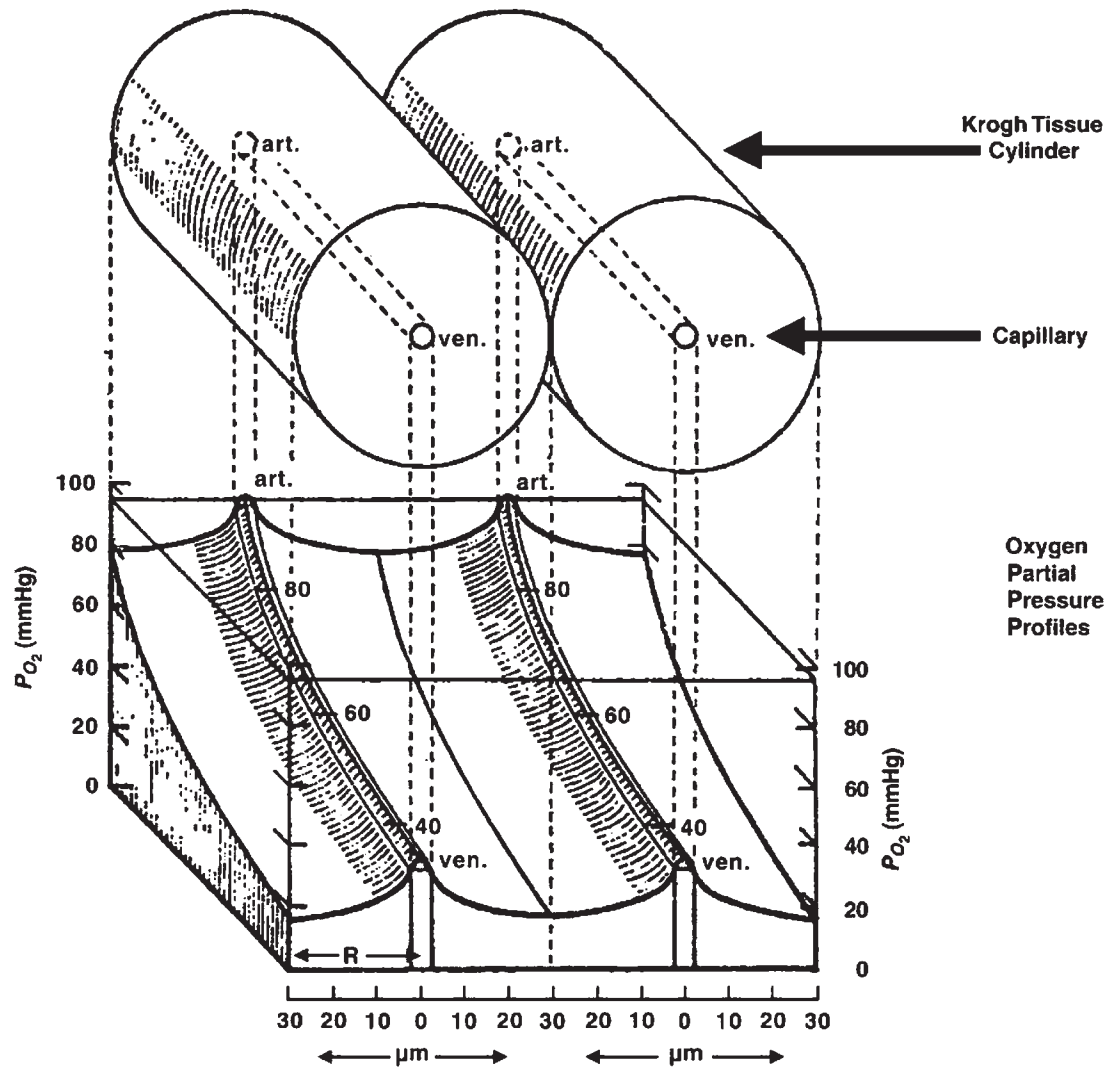


Fig. 3 Cerebral functional unit. (From Ref. [13].)

Molecular Complexes

There are also a very large number of molecular complexes, individual enzymes, enzyme complexes, ion pumps, and other entities of mass transport interest. Among the simpler are the enzyme complexes of Fig. 5 and ion channels (see http://anusf.anu.edu.au/Vizlab/viz_showcase/shinho_chung/) (Ref. [14], p. 99). Much more complex is the electron transport chain of Fig. 6 (Ref. [14], p. 787). This structure, imbedded in the inner mitochondrial membranes, includes simpler enzyme complexes as elements. It would appear from recent research that even this picture can prove to be oversimplified.^[17-19]

Microtubules, as shown in Fig. 4, and more schematically in Fig. 7, in addition to providing structural support, are of particular transport interest as they serve to transport vesicles and small organelles around

the cell. The roles of these various complexes in mass transfer will be discussed below.

FLOW DOMINATED SYSTEMS

Owing to both the above-mentioned safety factor and timescale separation many important biological processes can be described successfully without any significant consideration of diffusion. We shall consider two approaches that are particularly in wide use: application of the method of temporal moments for analysis of transient tracer pulses (Stewart and Spalding in Ref. [2], section 23.6-3, p. 756) and the use of compartment models (Ref. [20], especially Bischoff and Dedrick referenced on p. 3). Both were pioneered by biologists and chemical engineers long

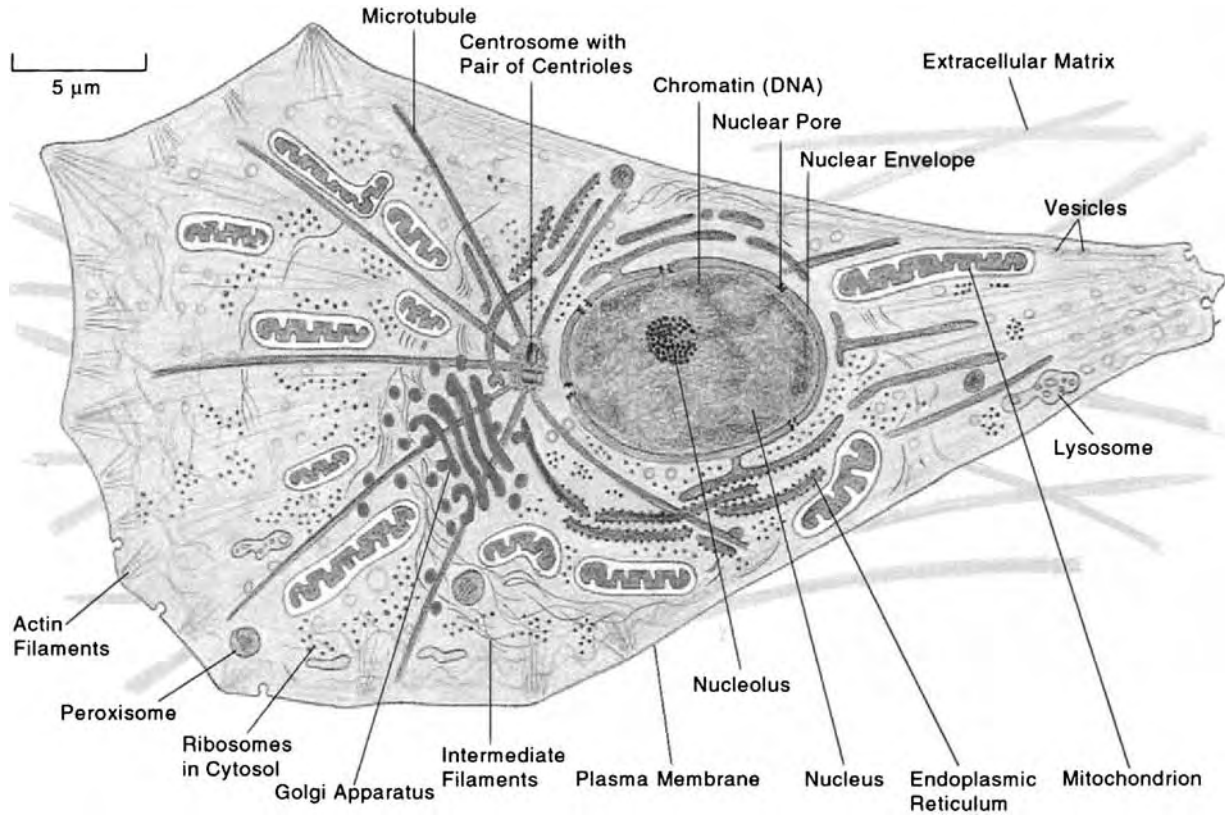


Fig. 4 The internal structures of eukaryotic cells. (From Ref. [14].)

ago and have since been used extensively in both conventional process engineering and environmental studies.

Temporal Moments

Temporal moments, defined immediately below, are restricted in principle to investigation of a simple closed flow system as illustrated in Fig. 8. A system of constant volume V with impermeable walls except for one inlet and one outlet is shown in the figure. Fluid, for example, blood, flows through this system at a

constant volumetric flow rate Q , and a tracer solute “ i ” is transported by the fluid under such circumstances that solute diffusion across the inlet and outlet by diffusion relative to the flow is negligible. The most important temporal moments of the tracer are defined by:

$$M_0 = \int_{-\infty}^{\infty} c_{i0} dt \tag{2}$$

$$M_1 = \int_{-\infty}^{\infty} t c_{i0} dt \tag{3}$$

$$m_2 = \int_{-\infty}^{\infty} (t - \bar{t})^2 c_{i0} dt \tag{4}$$

$$m_3 = \int_{-\infty}^{\infty} (t - \bar{t})^3 dt \tag{5}$$

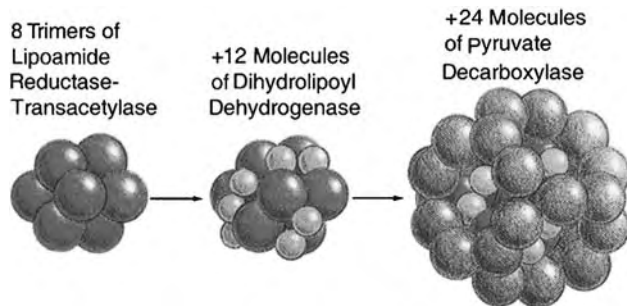


Fig. 5 The pyruvate hydrogenase complex. (From Ref. [14].)

These are known as the zeroth and first absolute moments, and the second and third moments relative to the mean residence time, respectively. Here, c_{i0} is the instant outlet tracer concentration, and t is time. For closed systems in series the values of M_1 , m_2 , and m_3

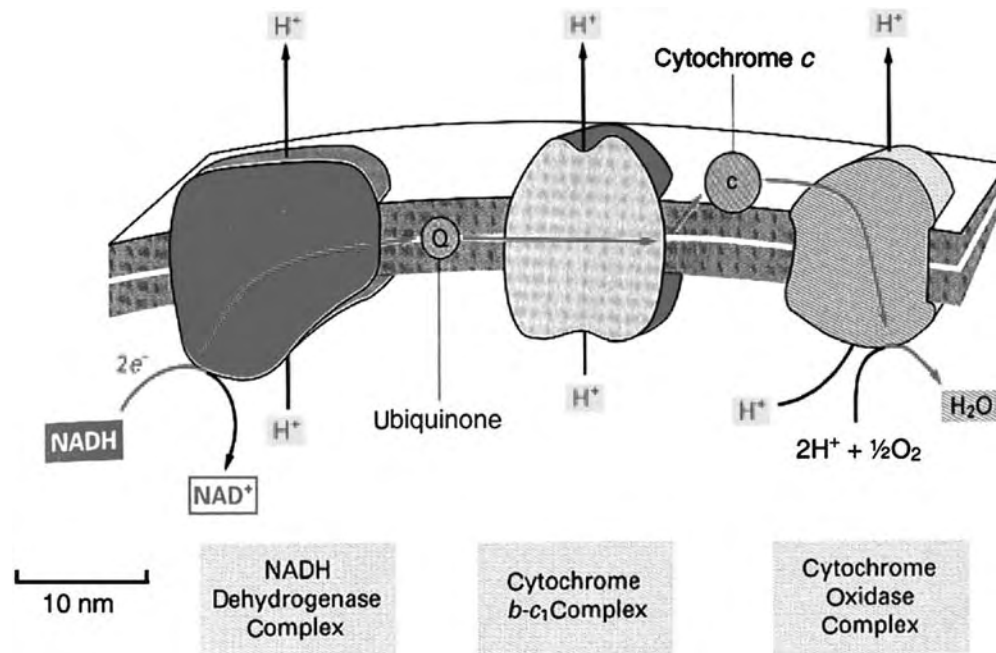


Fig. 6 The mitochondrial electron transport chain. (From Ref. [14].)

are additive, a property that permits subtracting out the effects of tubing and sensors. However, one must be careful to eliminate convective dispersion [Ref. [2], p. 643].

In practice, all of these requirements can often be relaxed, for example, by using time average flows. Moreover, in most cases it is possible to start the

integrals at time zero. However, the validity of these approximations must be determined in each new application. For medical applications recirculation of tracer is a serious problem (see Ref. [21]). Tracer responses for two simple flow situations are shown in Fig. 8.

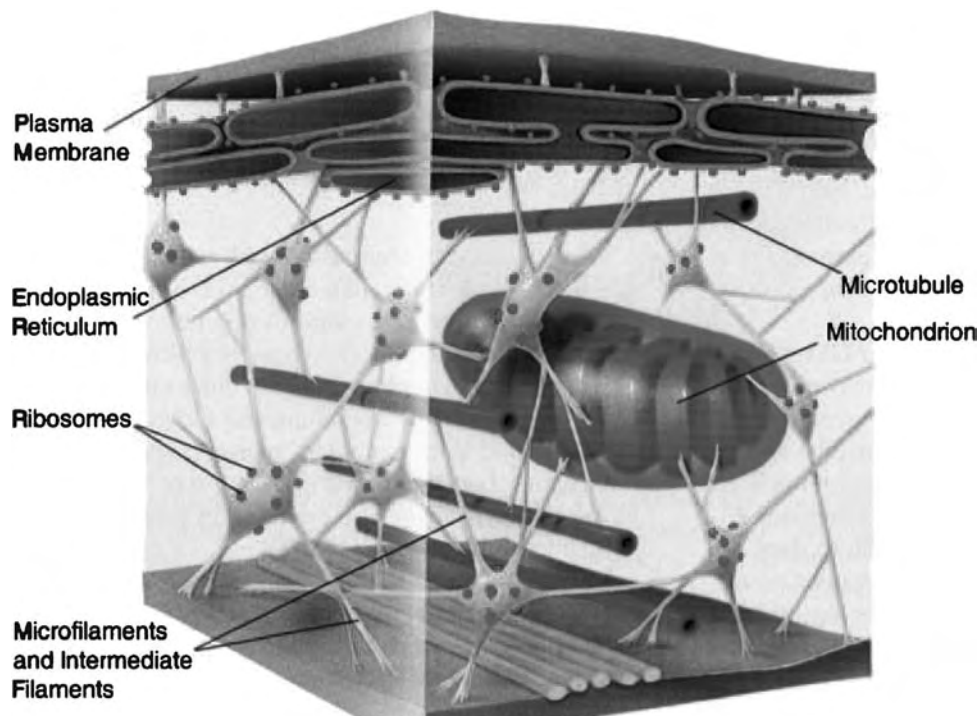


Fig. 7 Microtubules, mitochondria, and ribosomes.

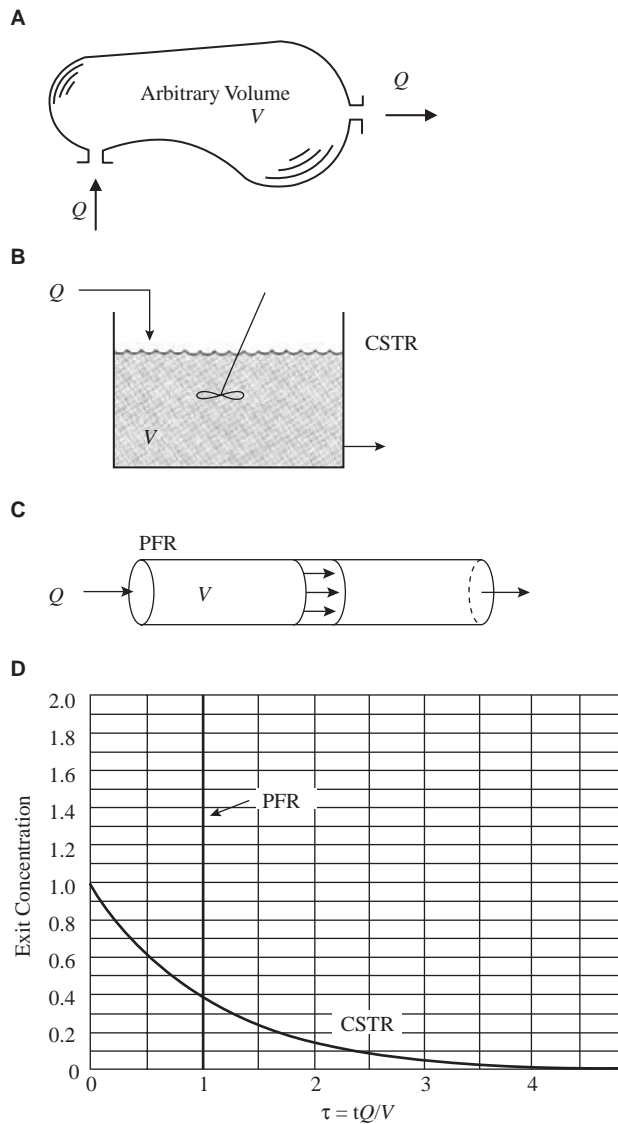


Fig. 8 Closed flow systems.

Eqs. (2) and (3) are particularly useful in providing the flow rate Q and effective volume V without any detailed knowledge of the system. Thus, from a simple mass balance:

$$Q = \frac{M}{\int_{-\infty}^{\infty} c_{i0} dt} = M/M_0 \quad (6)$$

Here, M is the mass of tracer fed. The mean solute residence time \bar{t} is defined as

$$\bar{t} = M_1/M_0 = V/Q \quad (7)$$

If the system is heterogenous with regions of varying solubility for the tracer, V must be replaced by effective volume:

$$V_{\text{eff}} = \sum_{j=1}^n V_j r_j \quad (8)$$

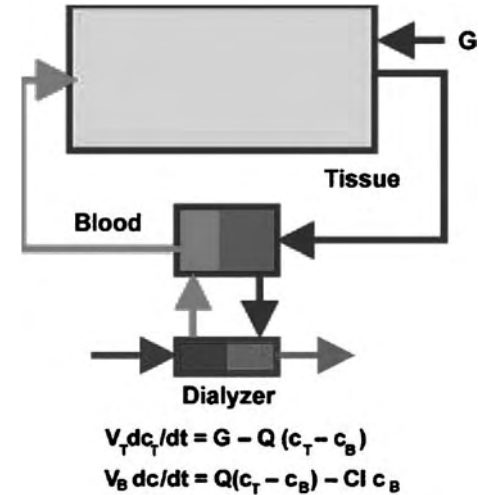


Fig. 9 Hemodialysis dynamics. (From Ref. [14].)

where r_j is the ratio of solubility in region j relative to that of the permeating fluid.

The second and third central moments m_2 and m_3 describe the dispersion and skew of tracer relative to \bar{t} , respectively. These two higher moments are particularly useful for evaluating parameters in structural models of the system. It is normally impossible to obtain accurate values of moments higher than 3, but one can use more than one tracer, a process often referred to as multiple indicator dilution. The use of tracers has become very sophisticated, and there is now a large literature available.^[22,23]

Pharmacokinetic Approximations

One can also construct many physiological models independent of diffusional considerations, and these are perhaps best known as the basis for calculating transient drug concentrations, hence the name of pharmacokinetics. They are often obtained by simplification of Fig. 1, but many are empirical.^[20,24,25] They depend on lumping substantial portions of the body into homogenous compartments and are also known as compartmental models.

The hemodialysis model is a classic example that has been thoroughly tested (Fig. 9) (Ref. [2], p. 733). Here, all body fluids except blood are lumped into a single mixing tank (denoted as tissue), and the blood is represented by another. Toxin buildup is modeled as a constant, continuous input, the dialyzer is described by a single parameter known as clearance. Predictions from this model are typically indistinguishable from actual data (see Fig. 10). Neither the body fluids nor blood act like mixers, and the success of this simple model results from the excellent

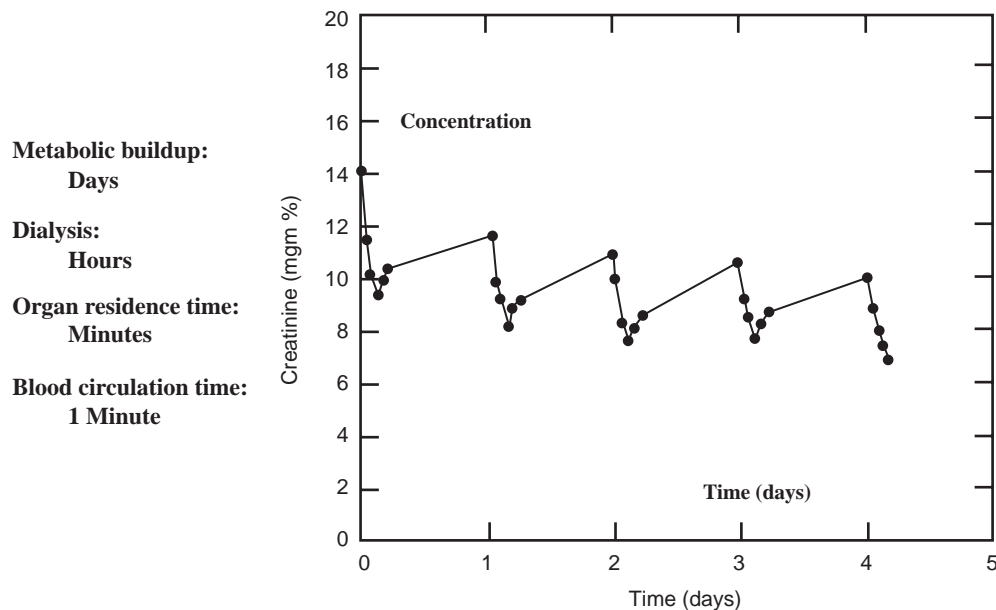


Fig. 10 Test of hemodialysis model. (From Ref. [2].)

separations of timescales for the different processes taking place. Under such circumstances, system transient responses are quite insensitive to the actual dynamics of individual components, and it is possible to use very simple approximations (Ref. [2], fig. 23.1–8, p. 737).

The corresponding model of an alveolus of the lung, the small air sacs at the ends of airway branching, is rather different (Fig. 11). The air sacs are surrounded by a net of capillaries where entering venous blood is aerated. We examine the situation of a small amount of a volatile tracer at concentration c_V entering with the venous blood. Part of it evaporates and enters a previously tracer-free alveolar air, which then leaves with a tracer partial pressure p_{Alv} (see Fig. 12). The arterial blood then has a reduced tracer concentration c_A . This resulting arterial blood is very nearly in equilibrium with the alveolar gas, so that we may assume:

$$p_A = Hc_A = p_{Alv} \quad (9)$$

where H is Henry's law constant. By a simple mass balance we may write:

$$Q(c_A - c_V) = \dot{V}p_{Alv}/RT \quad (10)$$

We may then write:

$$p_V - p_A = p_{Alv}(\dot{V}/Q\lambda) \quad (11)$$

where R is the international gas constant, T is absolute temperature, and λ is just H/RT . The volumetric flow rates of blood and air are Q and \dot{V} , respectively, following the usual convention.^[26]

We may now define fractional retention R , and an extraction E by

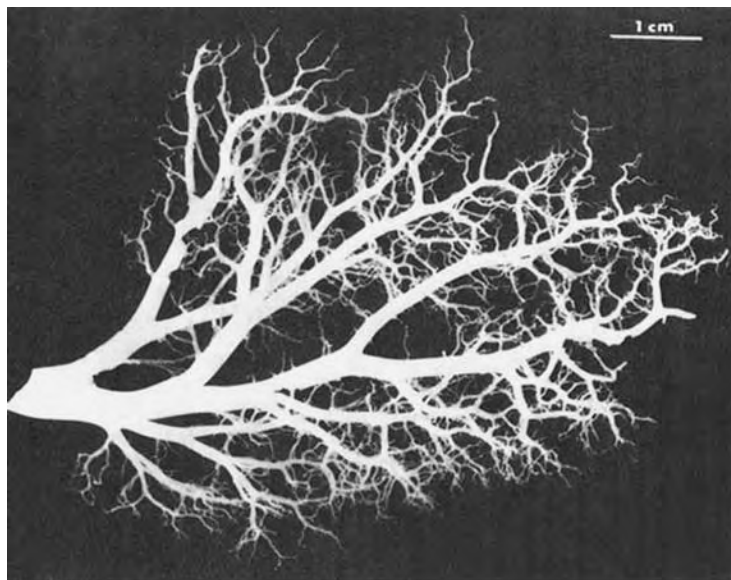
$$R = p_A/p_V = \frac{\lambda}{\lambda + \dot{V}/Q} \quad (12)$$

$$E = p_{Alv}/p_V = \frac{\lambda}{\lambda + \dot{V}/Q} \quad (13)$$

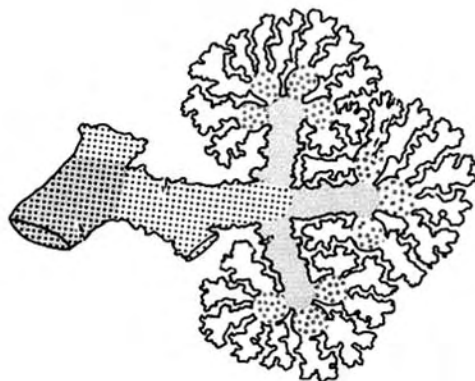
For any one alveolus, these quantities are the same, and they are functions only of λ and \dot{V}/Q . However, if these local partial pressures, p_{Alv} and p_A , are replaced by average values for the lung as a whole, the global retention and extraction differ. These differences can be used to estimate the heterogeneity of the lung as shown in Fig. 13. A dead space is defined as one with no gas flow at all, and such a region can provide no extraction. A shunt is a region where the blood sees no gas, and it must have a retention of unity. Thus, in a real lung the global E is always less than unity, and the global R is always greater than zero. Such heterogeneities reduce the oxygen content of arterial blood, and they tend to become more severe with age and lung damage. Similar heterogeneities occur throughout the circulatory system.^[21]

CONCENTRATION DIFFUSION

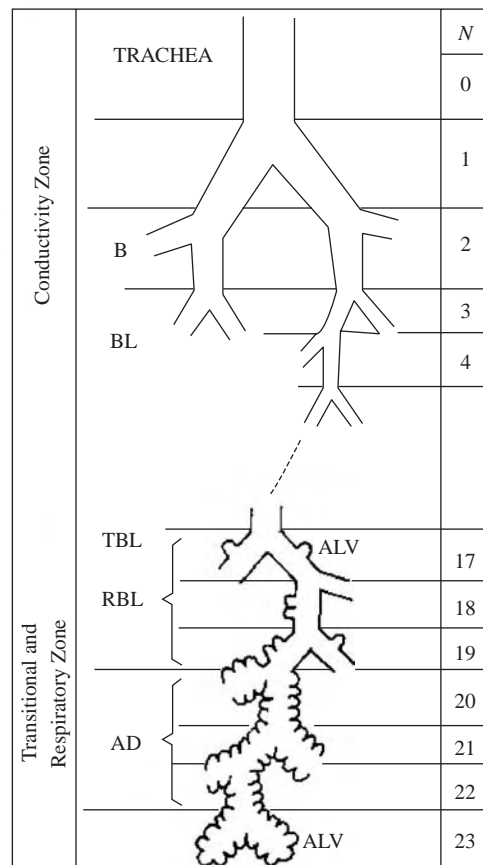
There are many situations where diffusion must be taken into account explicitly, and we will consider a few of them here. Perhaps, the most familiar is the oxygenation of tissue via the lungs and circulating blood. However, there are a large number of less familiar



Cat Pulmonary Tree



Alveoli



Pulmonary Organization



Fig. 11 Pulmonary branching. (From Ref. [9].)

situations requiring detailed analysis of concentration diffusion, and we shall touch on a few of these as well.

The lungs exhibit a branching structure with pronounced fractal characteristics as suggested in Fig. 11.^[9,26] The upper airways in the schematic diagram act as manifolds to distribute air to the alveolar sacs, which in turn are surrounded by a capillary net. It is primarily at this last region that transport of oxygen to the blood occurs. However, particulates of many kinds can enter the lungs and be deposited in the upper airways, and it is important to understand these deposition processes also (Ref. [26], p. 78, 28).

The circulation also forms two branching networks, pulmonary and systemic, and again transport, whether from air or to tissue, takes place via the smallest vessels.^[28] One example, brain capillaries, is shown in Fig. 3. Understanding these transport processes requires an introduction to the laws of diffusion, and we begin with a short overview of those.

Time Constants and Fick's Laws of Diffusion (Ref. [2], chap. 17)

Essentially, all physiological processes depend on individual metabolites moving relative to their surroundings, and the most common mechanism is random, or Brownian, motion. This undirected process, which is primarily responsible for concentration diffusion, tends to move individual molecules of species "i" away from their initial position, and the mean square distance traveled in a quiescent medium "m" over a time interval "t" is:

$$r^2 = x^2 + y^2 + z^2 = 6D_{im}t \tag{14}$$

$$x^2 = 2D_{im}t \tag{15}$$

Here r is distance from the origin, and its components in the direction of the three rectangular coordinates are

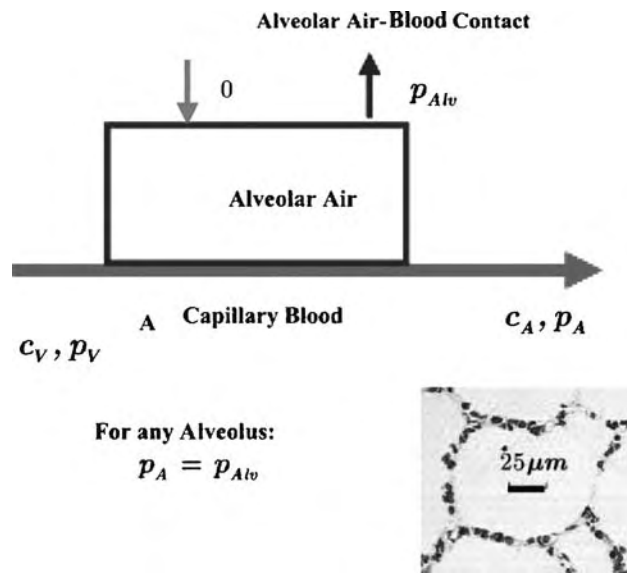


Fig. 12 Alveolar transport.

$x, y,$ and z . The quantity D_{im} is the Brownian diffusivity of species “i” through medium “m.” Eqs. (14) and (15) can be used to define diffusional time constants of the form:

$$T_{dif,R} = R^2/6D_{im} \tag{16}$$

$$T_{dif,L} = L^2/2D_{im} \tag{17}$$

That is, transient diffusion in spheres of radius R or slabs of half-width L will be “essentially” complete in times $T_{dif,R}$ or $T_{dif,L}$, respectively. They are particularly

useful for order-of-magnitude analysis, and we shall illustrate this shortly.

For our present purposes D_{im} may be considered equal to the diffusivity appearing in Fick’s law, which we will approximate here as (Ref. [2], section 17.1, p. 514):

$$N_{ix} = c_i v_{ix} = -D_{im} \partial c_i / \partial x + c_i v_x \tag{18}$$

Here, c_i is the molar concentration of species “i”, v_{ix} is the observable velocity of species “i” in the x direction, and v_x is the corresponding velocity of the surrounding medium.

Eq. (16) can be put into a mass balance to obtain:

$$\frac{\partial c_i}{\partial t} = D_{im} \left(\frac{\partial^2 c_i}{\partial x^2} + \frac{\partial^2 c_i}{\partial y^2} + \frac{\partial^2 c_i}{\partial z^2} \right) + \frac{\partial(c_i v_x)}{\partial x} + \frac{\partial(c_i v_y)}{\partial y} + \frac{\partial(c_i v_z)}{\partial z} + R_i \tag{19}$$

For the common case of quiescent media Eq. (17) takes the form

$$\frac{\partial c_i}{\partial t} = D_{im} \left(\frac{\partial^2 c_i}{\partial x^2} + \frac{\partial^2 c_i}{\partial y^2} + \frac{\partial^2 c_i}{\partial z^2} \right) + R_i \tag{20}$$

Here, R_i is the rate of formation of species “i” per unit volume. Eqs. (14), (17), and (18) suffice for this stage of our discussion.

Thus, Eqs. (16) and (17) are used to define diffusional response times in three-dimensional and one-dimensional systems, respectively. Eq. (19) serves to describe the combined effects of diffusion, reaction, and flow, and Eq. (17) can be used to describe transient

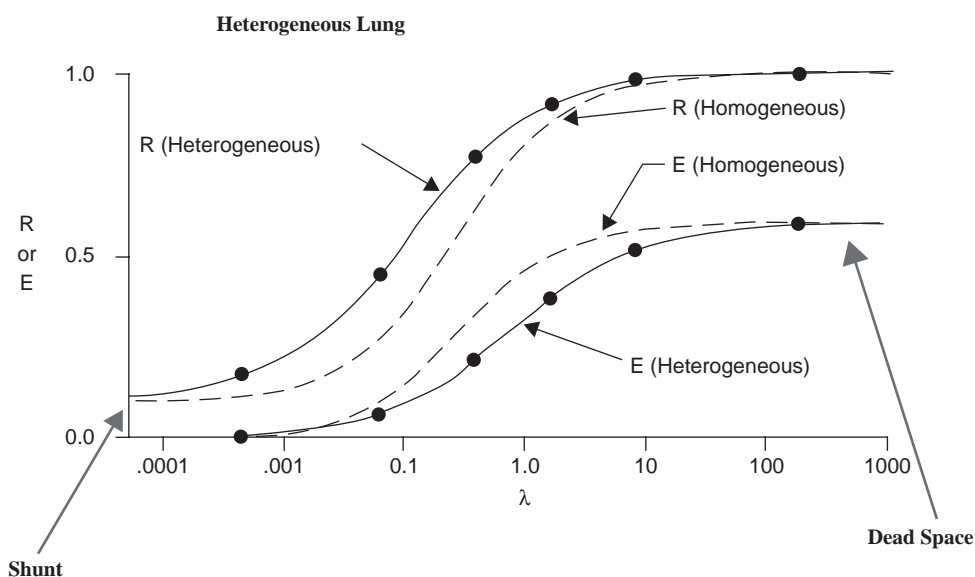


Fig. 13 Global pulmonary retention and extraction. (From Ref. [26].)

processes in quiescent media. We are now ready to look at some characteristic examples.

Orders of Magnitude

We begin by asking ourselves if there is any significant diffusional resistance in the alveolar sacs shown in Fig. 11 under physiological conditions. To do this we need to know that the diffusivity of oxygen in air at 37°C is about 0.21 cm²/sec, and that the normal respiration rate is about 12 breaths/min, or a cycle time of 5 sec (Ref. [2], table 17.1, p. 517; eq. 17.3–12, p. 526). For this system then, the diffusional time constant is about:

$$T_{\text{dif},R} = \frac{(5 \times 10^{-3} \text{ cm})^2}{6 \times 0.21 \text{ cm}^2/\text{sec}} = 1.98 \times 10^{-5} \text{ sec} \quad (21)$$

This is extremely small relative to 5 sec so we may assume that the gas phase concentration is always quite uniform: the general experience is that the transient process can be considered complete in about three time constants.

We next try to identify those blood vessels able to supply oxygen to tissue, and all we have available is the summary of data in Fig. 14. The key to success is to recognize that the last column heading can be rewritten as (Ref. [2], table 22.2–1, p. 676):

$$L/DReSc = (L/\langle v \rangle)/(D^2/D_{O_2B}) \quad (22)$$

where $\langle v \rangle$ is flow average velocity and D_{O_2B} is the effective diffusivity of oxygen through blood. Now in two-dimensional (radial) diffusion the diffusional time constant:

$$T_{\text{dif}} = R^2/4D_{O_2B} \quad (23)$$

while the blood holdup or residence time:

$$\bar{t} = L/\langle v \rangle \quad (24)$$

These two characteristic times are equal when

$$L/DReSc = 1/16 = 0.0625 \quad (25)$$

This is intermediate between the entries for terminal arteries and arterioles in the figure.

Thus, none of the arteries provides enough time for dissolved oxygen to reach the arterial wall, let alone into the surrounding tissue. These are convective transport vessels. The arterioles and venules can provide some oxygen, but it is clearly the capillaries that are the major suppliers. Note also how quickly the characteristics change from “convector” to “diffuser”! Body “design” is truly impressive. Note also how much information we have obtained without any real effort. The very large values of $L/DReSc$ reflect the need to have sufficient transport capacity to permit tissue oxygenation. We return to this point below.

Next, we show how one can often estimate the time course of rather complex diffusional processes, and for this we use bacterial gene expression. The critical process here is reaction of a promoter molecule with the gene to be activated, and this can be envisioned as taking place in three steps.^[29]

First, the promoter diffuses at random in the cell until the probability of its being in all positions is the same. This occurs over the diffusional time scale:

$$T_{\text{dif}} = R_{\text{cell}}^2/6D_{\text{pc}} \quad (26)$$

where R_{cell} is the effective cell radius, and D_{pc} is the effective protein diffusivity in the cytoplasm. It is reasonable to assume the cell radius to be about 1 μm.



Mass Transfer Effectiveness of Blood Vessels 13-kg Dog				
Vessel	Radius (cm)	Length (cm)	$\langle v \rangle$ (cm/s)	L/DReSc
Aorta	0.5	40	50	0.000032
Large Arteries	0.15	20	13.4	0.000033
Main Art. Branches	0.05	10	8	0.005
Terminal Arteries	0.03	1.0	6.0	0.00185
Arterioles	0.001	0.2	0.32	6.25
Capillaries	0.0004	0.1	0.07	89.4
Venules	0.0015	0.2	0.07	12.7
Terminal Veins	0.075	1.0	1.3	0.0014
Main Venous Branches	0.12	10	1.48	0.0047
Large Veins	0.3	20	3.6	0.00064
Venae Cavae	0.625	40	33.4	0.000031

Fig. 14 Classification of blood vessels. (From Ref. [28].)

However, protein diffusion, much slowed down by the crowded nature of the cytoplasm, is about 71 times slower than in water or about 10^{-8} cm²/sec (Ref. [29], Table 115.5, p. 115.7). This randomization process then takes place in about:

$$\begin{aligned} T_{\text{ran}} = T_{\text{dif}} &= (10^{-8} \text{ cm}^2)/(6 \times 10^{-8} \text{ cm}^2/\text{sec}) \\ &= (1/6) \text{ sec} \end{aligned} \quad (27)$$

We shall shortly see that this is very fast, and we are left with the classic problem of diffusion to a spheroid from large surroundings (e.g., ex. 2.2.2, p. 205) with fast irreversible reaction at the surface. It can therefore be shown that the probability of the promoter finding the gene is:

$$P = 1 - \exp[-t/T_{\text{rxn}}] \quad (28)$$

with

$$T_{\text{rxn}} = 3D_{\text{PC}}[(R_P + R_{\text{eff}})/R_{\text{cell}}^3] \quad (29)$$

where R_P is the effective radius of the protein and R_{eff} that of the gene. If one uses just the actual size of a typical protein, 5 nm, and neglects the size of the gene proper:

$$T_{\text{rxn}} = \left[\frac{3 \times 5 \times 10^{-7}}{(10^{-4})^3} \times 10^{-8} \right]^{-1} = 67 \text{ sec} \quad (30)$$

This is in good agreement with extensive numerical calculations (see Berg).^[29] However, actual kinetics are substantially faster because the promoter molecule can sample substantial lengths of DNA on each encounter.

Diffusion and Reaction

The primary purpose of diffusion is to provide feeds for chemical reaction, i.e., organ metabolism. This always requires a concentration gradient as described by Fick's law, Eq. (18), and here we consider diffusion in the absence of flow. We take as a specific example a slab of tissue as described in Fig. 15, but the behavior of cylinders or spheres is almost identical using the reaction modulus:

$$\Phi = \frac{\langle R_i \rangle (V/S)^2}{c_{i0} D_{iT}} \quad (31)$$

shown as an abscissa in the figure (Ref. [2], p. 566). Here, $\langle R_i \rangle$ is the average rate of consumption of species "i," in moles per unit volume and time over the whole tissue volume, V . The surface area exposed to reactant "i" is S , the molar concentration of "i" at the

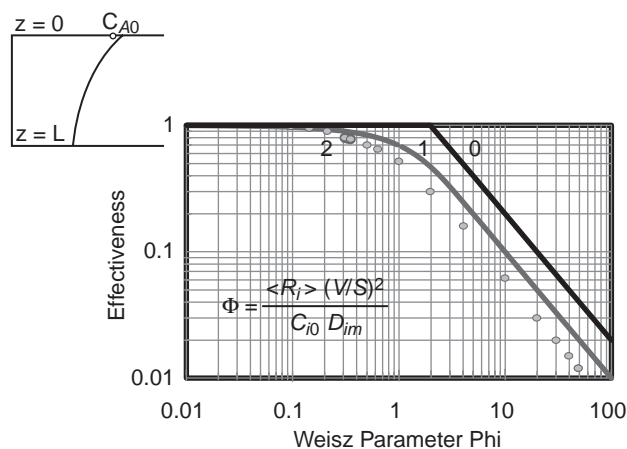


Fig. 15 Effectiveness factors.

surface is c_{i0} , and its diffusivity through the tissue is D_{iT} . The figure is drawn for "power-law" kinetics where local values of consumption rate are given by:

$$R_i = k_n c_i^n \quad (32)$$

with values of reaction order n equal to 0, 1, and 2 as shown. The effectiveness taken as origin is the ratio of observed consumption rates to those at the exposed surface. First, it should be noted that effectiveness is not greatly dependent on n .

More important for our discussion, the results for all biological kinetics, for example, the familiar Michaelis–Menten model, lie between the zero- and first-order lines (Ref. [14], p. 164). It follows that we can obtain a fairly good estimate of diffusional effects just from examination of the tissue involved. We begin by showing the demand for circulating blood in a normal adult human of 70 kg. We need only orders of magnitude here so we can make the following approximations:

$$V = 70 \text{ L} = 70,000 \text{ cm}^3$$

$$S = 10 V^{2/3} = 17,000 \text{ cm}^2; (V/S) = 4 \text{ cm}$$

$$\langle R_i \rangle [0.25 \text{ L(sTP)/min}]/70,000 \text{ cm}^3$$

$$= [0.25/(22.4 \text{ L/gmol} \times 60 \text{ sec/min})]/70,000 \text{ cm}^3$$

$$= 2.66 \times 10^{-9} \text{ gmol/sec/cm}^3$$

$$c_{i0} = 0.21 \times 1.0 \times 10^{-6} \text{ gmol/cm}^3$$

$$= 2.1 \times 10^{-7} \text{ gmol/cm}^3$$

$$D_{O_2T} = 10^{-5} \text{ cm}^2/\text{sec}$$

It follows that the value of the modulus is on the order of 10^4 and the effectiveness factor is on the order

of 0.0001. The bottom line then is that all vertebrate animals need a blood circulation to supply oxygen.

It is perhaps more useful to note that almost all enzymes operate with a modulus between 1/3 and unity (see Ref. [15], section 12.3.2 and in particular the Ref. [31]). There are at least two reasons for the existence of enzyme clusters such as that shown in Fig. 5. First, if an intermediate is present at very low concentration one can still obtain a high effectiveness if the diffusion distance to the next is sufficiently small. Second, a cluster may act to keep an intermediate from diffusing away into the cytosol.

Diffusion, Reaction, and Convection

In most functional units diffusion, reaction, and convection are in balance, and we take as our example oxygenation of Krogh cylinders in the brain (Fig. 3). The brain must obtain all of its metabolic energy from aerobic oxidation of glucose, and the combination of low oxygen diffusivity and solubility is the primary constraint on design.^[21] There have been many descriptions of these processes over the years, but the results shown in Purves are still reasonable (Fig. 3). The diffusion–reaction modulus Φ at the distal (venous) end of cerebral capillaries is usually estimated to be about 1/4. This is a little lower than the range claimed by Weisz, but this is an extraordinary case (Ref. [15], ref. 375). It is critically important that cerebral metabolism be maintained very close to normal values even while the organism is heavily stressed.

MORE COMPLEX TRANSPORT MECHANISMS

Intracellular transport is far more complicated than the above-described situations, and so are the morphologies of the systems involved. One must now consider additional motive forces, for example, coupled processes such as passive and active carrier transport through membranes vesicular traffic along filamentous networks (microtubules), chemotaxis, or single-protein transport using dedicated motors (Ref. [14], chap. 11).^[1,30,32,33] These are too complex for extensive coverage here, but they are too important to ignore.

Many of the diffusing species are electrically charged, and their motion is therefore heavily influenced by electric fields (Ref. [2], eqs. 24.2–8, p. 769, ex. 24.4-3, p. 780). A more limited but simpler picture is provided by Alberts et al. (Ref. [14], p. 634). There are potential differences of some tens of millivolts across most biological membranes, and these are used to control concentrations of charged species c_i , notably

small ions. This action can be explained to a fair approximation by the Nernst equation (Ref. [14], p. 634):

$$V_2 - V_1 = \frac{RT}{z_i F} \ln \left(\frac{c_{i1}}{c_{i2}} \right) \quad (33)$$

Here, V is the electric potential in volts for sides 2 and 1 of the membrane, R is the international gas constant, 1.987 cal/mol $^\circ$ K, T is absolute temperature, and F is Faraday's constant, 2.308×10^4 cal/V(g/eqt.). The ionic charge z_i is, for example, +1 eqt./mol for sodium ion, and –2 eqt./mol for sulfate. Thus, cation concentrations will be higher on the low-voltage side. Membrane potentials are produced by ion pumps (Ref. [1], chap. 3, p. 74; Ref. [14], p. 635). One of the most complex biological mass transfer situations is the transmission of nerve impulses by transient disturbance of membrane potentials (Ref. [1], chap. 4, p. 116).

CONCLUSIONS

It is hoped that even this cursory discussion of biological mass transfer will prove useful and, more specifically, that it will make the large mass of information in the references accessible. The accumulated literature is enormous and the subject matter exceedingly complex. For this reason it is important that the goals of investigators be very carefully posed and that only needed information be sought. Fortunately, the hierarchical design of living systems frequently makes surprisingly simple models useful, and much of the above discussion has been based on such simplified descriptions. At the same time the difficulty of obtaining reliable data, and the tendency of living systems to oppose attempts to probe them, can severely limit the numbers of model parameters that can be determined. It is most important to read the literature critically, and in particular, to examine all basic data carefully, before accepting any generalizations drawn from them. That warning applies also to those provided here. A good point to make!

ARTICLES OF FURTHER INTEREST

Microcirculatory Oxygen Transport; Protein and Cell Signaling with Biomaterials: Interfacial Transport; Tissue Engineering; Vascular Fluid Dynamics

REFERENCES

1. Keener, J.; Sneyd, S. *Mathematical Physiology*; Springer, 1998.
2. Bird, R.B.; Stewart, W.E.; Lightfoot, E.N. *Transport Phenomena*, 2nd Ed.; Wiley, 2002.

3. Netter, F.H. *Atlas of Human Anatomy*, 2nd Ed.; Novartis, 1997.
4. Schmidt-Nielsen, K. *Scaling*; Cambridge, 1984.
5. Krebs, H.A. Body size and tissue respiration. *Biochim. Biophys. Acta* **1950**, *4*, 249–269.
6. Calder, W.A.I. *Size, Function and Life History*; Dover Edition; 1996.
7. White, C.R.; Seymour, R.S. Mammalian basal metabolic rate is proportional to body mass to the 2/3 power. *Proc. Natl. Acad. Sci.* **2003**, *100* (7), 4046–4049.
8. Audesirk, T.; Audesirk, G. *Biology: Life on Earth*, 4th Ed.; Prentice Hall, 1996.
9. Bassingthwaighte, J.B.; Liebovitch, L.S.; West, B.J. *Fractal Physiology*; Oxford, 1994.
10. Schmidt-Nielsen, K. *Animal Physiology*, 5th Ed.; Cambridge, 1997.
11. Johnson, L.R. *Essential Medical Physiology*; Lippincott-Raven, 1998.
12. Kessel, R.G. *Basic Medical Histology*; Oxford, 1998.
13. Purves, M.K. *Physiology of the Circulation*; Cambridge, 1972.
14. Alberts, B.; Johnson, A.; Lewis, J.; Reff, H.; Roberts, K.; Walter, P. *Molecular Biology of the Cell*, 4th Ed.; Garland, 2002.
15. Palsson, B.O.; Bhatia, S.N. *Tissue Engineering*; Pearson Prentice Hall, 2004.
16. Goodsell, D. *The Machinery of Life*; Springer, 1998.
17. Dudkina, N.V.; Eubel, H.; Keegstra, W. Structure of a mitochondrial super complex formed by respiration-chain complexes I and III. *Proc. Natl. Acad. Sci.* **2005**, *102* (9), 3225–3229.
18. Lagerholm, B.C.; Weinreb, G.E.; Jacobsen, G.E.; Thompson, N.L. Detecting microdomains in intact cell membranes. *Annu. Rev. Phys. Chem.* **2005**, *56*, 309–336.
19. Raman, R.; Sasikharan, V.; Sasikharan, R. Structural insights into biological roles of protein-glycosaminoglycan interactions. *Chem. Biol.* **2005**, *12* (3), 267–277.
20. Welling, P.G. *Pharmacokinetics*; 1997.
21. Beard, D.A.; Bassingthwaighte, J.B. Power-law kinetics of tracer washout from physiological systems. *Ann. Biomed. Eng.* **1998**, *26*, 775–779.
22. Bassingthwaighte, J.B., Goresky, C.A., Linehan, J.H., Eds. *Whole Organ Approaches to Cellular Metabolism*; Springer, 1998.
23. Kroll, K.; Wilke, N.; Jerosch-Herold, M. Modeling regional myocardial flows from residue functions of an intravascular indicator. *Am. J. Physiol.* **1996**, *271* (4), H1643–H1655.
24. Alcorn, J.; McNamara, P.J. Pharmacokinetics in the newborn. *Adv. Drug Deliv. Rev.* **2003**, *55* (5), 667–686.
25. Perelson, A.S. Modeling viral and immune system dynamics. *Nature Rev. Immunol.* **2002**, *2* (Jan 1), 29–35.
26. Hlastala, M.P.; Berger, A.J. *Physiology of Respiration*; Oxford, 1996.
27. Schlesinger, F.B. Comparative deposition of inhaled aerosols in experimental animals and humans. *J. Toxicol. Environ. Health* **1985**, *13*, 197–214.
28. Fung, Y.C. *Biomechanics Circulation*, 2nd Ed.; Springer, 1997.
29. Lightfoot, E.N.; Duca, K.A. The roles of mass transfer in tissue function. In *The Biomedical Engineering Handbook*, 2nd Ed.; Bronzino, J.D., CRC Press: 2000; II, 115–1–115-15.
30. Okamoto, G.H.; Lightfoot, E.N. Energy cost of intracellular organization. *Ind. Drug. Chem. Res.* **1992**, *31* (3), 732–735.
31. Weisz, P.B. Diffusion and Chemical Transformation Science. **1973**, *179*, 433–440.
32. de Duce, C. *Blueprint for a Cell*; Patterson, 1991.
33. Beard, D.A.; Bassingthwaighte, J.B. Modeling advection and diffusion of oxygen in complex vascular networks. *Ann. Biomed. Eng.* **2001**, *29*.

Matrix Metalloproteinases

Todd A. Telemeco

Division of Physical Therapy, Shenandoah University, Winchester, Virginia, U.S.A.

James R. Bowman, III

David G. Simpson

Department of Anatomy and Neurobiology, Virginia Commonwealth University, Richmond, Virginia, U.S.A.



INTRODUCTION

The development of clinically relevant tissue-engineered products is highly dependent upon the identification of suitable scaffoldings that can be used to mimic the native extracellular matrix (ECM), a daunting barrier. In vivo, the native ECM defines the architectural organization and the passive mechanical properties of nearly every solid human tissue. This complex network of fibrous proteins also functions as a repository for soluble peptides that can regulate a variety of physiological processes in the surrounding cell population. Classically, this network is often viewed as a static structural scaffolding. However, the composition and tertiary structure of the ECM are subject to regulation throughout life. This natural scaffolding can undergo profound changes in composition in development, at the onset of disease, and during the progression of normal wound repair. Even during homeostasis, constituents of the extracellular matrix are routinely broken down by proteolytic events and replaced by nascent proteins.

In the bioengineering paradigm, scaffoldings play a central role in the fabrication and delivery of tissue-engineered organs to a host site. The mechanical and biochemical perturbations that inevitably occur to the local microenvironment when a tissue-engineered organ is implanted can markedly alter homeostasis and stimulate the remodeling of the surrounding ECM. Fundamental to these remodeling processes is the proteolytic breakdown of pre-existing proteins in the ECM of the host and, potentially, the scaffolding elements of the implanted tissue-engineered organ. As a class, enzymes of the matrix metalloproteinase (MMPs) family exhibit broad specificity against the principal structural proteins of native extracellular matrix (ECM). The implantation of nearly any tissue-engineered organ can be expected to induce the expression, synthesis, or activation of these proteases at the host-implant interface.

There is a dynamic interaction between the ECM and the cellular elements of an organ. Changes in the composition of the ECM can impact the physiological state of the surrounding cells. Conversely, changes in the physiological state of the cellular compartment can

impact the synthesis, deposition, and turnover of matrix constituents. Our laboratories believe that this dynamic interaction must be recognized and incorporated into the design of tissue-engineered organs. For example, when a tissue-engineered organ is implanted, a complex set of events must occur to promote the structural and functional integration of the construct into the host site. At implantation, interstitial fibroblasts from the surrounding tissue are stimulated to migrate into the tissue-engineered organ. If these cells cannot freely penetrate the borders of the implanted tissue, the construct may initiate fibrosis and inflammation, potentially leading to the delamination of the construct from the host site. Notably, as an integral part of the integration process, microvascular endothelial cells must rapidly penetrate and migrate throughout a tissue-engineered organ to establish capillary beds and cardiovascular support. Matrix degradation and remodeling are integral to the processes of cell migration and the formation of these nascent blood vessels.

BACKGROUND

The interrelationships between matrix proteolysis and cell migration place constraints on the design of tissue-engineering scaffoldings. A scaffold composed solely of synthetic materials is not typically subject to enzymatic attack. In this type of matrix, cell migration will be limited by physical constraints imposed by the porosity of the scaffolding and the rate at which the material dissolves. Strategies designed to facilitate scaffolding remodeling include the fabrication of synthetic scaffoldings with MMP-sensitive cleavage sites and biologically active peptides.^[1] The efficacy of this approach in a tissue-engineering application remains to be determined; the synthetic components of these structures still represent a potential nexus for inflammation. Scaffoldings composed of native materials that are subject to proteolysis can be expected to support more rapid and extensive cellular migration, facilitating the integration process. One caveat: if a natural tissue-engineering scaffolding lacks sufficient structural coherence, premature or

excessive rates of proteolysis can lead to the catastrophic failure of the implant before the construct can be fully integrated into the host site. Matrix proteolysis is mediated by a variety of different endopeptidases acting on compositionally distinct ECM proteins. Members of the matrixin subfamily of metalloproteinases, the MMPs, have been implicated in the proteolytic destruction of nearly all the protein constituents of the ECM.^[2,3]

Members of the MMP family are ubiquitously distributed throughout mammalian tissues. These enzymes are expressed, deposited into the extracellular environment, and activated in response to a host of environmental events. The differentiated resident cells and interstitial fibroblasts that populate the organ systems can synthesize and deposit both active and inactive MMPs into the surrounding ECM. Circulating cells of the cardiovascular system that mediate inflammatory processes also produce these proteases. As a result, MMP activity within an implant site can originate from nascent enzymes synthesized by the resident cells of the surrounding tissue, pre-existing enzymes present within the ECM that are activated in response to an environmental signal, or from inflammatory cells that respond to an implant construct.

Jerome Gross and Charles Lapiere (1962)^[4] published the first study describing an enzyme displaying proteolytic activity against ECM constituents in the vertebrate system during morphogenesis. Since that report, members of the MMP subfamily have been implicated in numerous physiological events including developmental processes,^[5,6] cell migration,^[7] normal homeostatic maintenance,^[8] reproduction, wound healing, and pathological processes such as metastatic cancer,^[7] fibrotic disease,^[9] arthritis,^[10,11] ulcerations,^[12,13] and other matrix-weakening diseases.^[3] It is clear from the wide distribution and diversity of functions ascribed to these enzymes that nearly any tissue-engineering application has the potential to invoke the expression and/or activation of these proteases. The MMP subfamily has been divided into five groups: the 1) collagenases, 2) gelatinases, 3) stromelysins, 4) membrane-type MMPs (MT-MMP), and 5) other MMPs^[2,3] (Table 1). These groups are distinguished from one another on structural considerations and substrate specificity.

STRUCTURE AND REGULATION: AN OVERVIEW OF MMPs

Members of the MMP family exhibit a broad range of enzymatic activity against matrix constituents. This robust proteolytic activity is subject to regulation from the transcriptional to the posttranslational levels of control. It is not at all surprising that the unregulated expression and activation of MMP proteolytic activity

can have profound effects on the local microenvironment and the overall structural integrity of a tissue. For example, the consequences of MMP dysregulation are evident in the chronic, nonhealing ulcer.^[20] The local microenvironment of the nonhealing ulcer exhibits elevated levels of MMPs derived from the neutrophils that infiltrate the interstitium surrounding these injuries.^[20,21] The protease activity associated with these enzymes exacerbates the inflammatory process and directly contributes to the pathogenesis of the lesion.^[21] In addition to the increased levels of MMPs found in nonhealing wounds, these sites also often exhibit decreased concentrations of the endogenous inhibitors of these proteases, the tissue inhibitors of matrix metalloproteinases (TIMPs).^[20] Therapeutic agents targeted at suppressing MMP activity represent a fundamental shift in the clinical treatment of chronic ulcers.

At the transcriptional level of control, growth factors and cytokines, including IL-1, PDGF, and TNF- α , hormones, and oncogenic factors have all been demonstrated to induce the expression of MMPs.^[22-24] In contrast, TGF- β , retinoic acid, IL-4, heparin, and corticosteroids suppress the expression of selected MMPs. The regulatory effects of these environmental factors are not uniform across the MMP family, suggesting that a degree of variation exists in the signal transduction pathways that control the expression and activation of these enzymes in different tissues. For example, TGF- β accelerates the synthesis of selected MMPs in the gelatinase subfamily in some tissues, yet the expression of the collagenases and the stromelysins can be inhibited by this growth factor in other tissues.^[25]

At the posttranscriptional level of control, the proteolytic activity of the MMPs is regulated by the tertiary structure of the molecule. The prototypic MMP is expressed and deposited into the extracellular environment as an inactive proenzyme. Activation of enzymatic activity subsequently occurs in the extracellular space in response to specific environmental signals. In the propeptide MMP isoform the catalytic site is sequestered and rendered inactive by a conformational feature in the molecule that is maintained by a cysteine switch. An unpaired cysteine sulfhydryl group located near the C-terminal end of the protein interacts with a regulatory zinc ion that is present within the catalytic domain.^[2,3] Disruption of this cysteine switch by enzymatic cleavage of the propeptide sequence leads to a conformational change and the activation of the enzyme.

In addition to the zinc-dependent cysteine switch, the majority of MMPs contain other common structural features (Fig. 1). Each MMP exhibits a signal sequence. This signal sequence, typically 17–20 amino acid residues in length, directs the nascent polypeptide chain to the endoplasmic reticulum where it is packaged and prepared for deposition into the extracellular compartment. The propeptide domain is approximately

Table 1 MMP subfamilies

Group	MMP-No.	Other names	ECM substrate/proteins digested
Collagenases	MMP-1	Collagenase 1/ fibroblast collagenase	Type I, II, III, VII, VIII, X, and XI collagens, Gelatin, Fibronectin, Vitronectin, Laminin, Tenascin, Entactin, Aggrecan, Link Protein, Myelin basic protein (MBP)
	MMP-8	Collagenase 2/ neutrophil collagenase	Types I, II and III collagens, Aggrecan
	MMP-13	Collagenase 3/ rat interstitial collagenase	Type I, II, III, VI, IX, X, and XIV collagens, Fibronectin, SPARC, and Aggrecan
	MMP-18	Collagenase 4	Unknown
Gelatinases	MMP-2	Gelatinase A	Type I, III, IV, V, VII, X, and XI collagens, Elastin, Fibronectin, Vitronectin, Laminin, Tenascin, Entactin, SPARC, Aggrecan, Decorin, Link Protein, and MBP
	MMP-9	Gelatinase B	Type IV, V, XI, and XIV collagens, Gelatin, Elastin, Vitronectin, Laminin, SPARC Aggrecan, Decorin, Link Protein and MBP
Stromelysins	MMP-3	Stromelysin 1	Type III, IV, V, VII, IX, X, and XI collagens, Gelatin, Elastin, Fibronectin, Vitronectin, Laminin, Tenascin, Entactin, SPARC, Aggrecan, Decorin, Link Protein, and MBP
	MMP-10	Stromelysin 2	Types III, IV, and V collagens, Gelatin, Elastin, Fibronectin, Aggrecan, and Link Protein
	MMP-11	Stromelysin 3	Does not degrade classic extracellular matrix components, ^[14,15]
Membrane-type MMPs	MMP-14	MT1-MMP	Type I, II, and III collagens, Gelatin, Fibronectin, Vitronectin, Laminin, Aggrecan, Nidogen, ^[16] and Perlecan, ^[16]
	MMP-15	MT2-MMP	Fibronectin, Laminin, Tenascin, Aggrecan, Nidogen, ^[16] and Perlecan, ^[16]
	MMP-16	MT3-MMP	Type I and III collagens, Gelatin Fibronectin, Vitronectin and Laminin ^[17]
	MMP-17	MT4-MMP	Unknown
Other MMPs	MMP-24 & 25	MT5-MMP & MT6-MMP	Unknown
	MMP-7	Matrilysin	Type I and IV collagens, Gelatin, Elastin, Fibronectin, Vitronectin, Laminin, Tenascin, Entactin, SPARC Aggrecan, Decorin, Link Protein, and MBP
	MMP-12	Metalloelastase/ macrophage elastase	Type I and IV collagens, Gelatin, Elastin, Fibronectin, Vitronectin, Laminin, Entactin, Aggrecan, and MBP
	MMP-19	No trivial name	Aggrecan and cartilage oligomeric matrix protein (COMP) ^[18]
	MMP-20	Enamelysin	Amelogenin, ^[19] Aggrecan, and COMP ^[18]

Substrate degradation as determined by in-vitro experimentation. Modified from Refs. [2] and [3].

80 amino acids in length and highly conserved with a PRCXXPD sequence containing the cysteine residue that forms the cysteine switch. The catalytic domain is 160–170 amino acid residues and includes calcium-binding sites, a second zinc atom that is integral to the structure of the peptide, and a binding site for the catalytic zinc atom.^[2]

Depending on the specific identity and subtype of MMP under consideration, other domains that may be present include a furin-cleavage site, hemopexinlike

domain, a hinge region, a variable number of fibronectinlike repeats, and/or a membrane insertion extension. Enzymatic cleavage at Furin-sensitive sites occurs within the trans-golgi network and, in some instances, appears to down regulate proteolytic activity of the mature enzyme.^[27] The hemopexinlike domain has a 200-amino-acid sequence that resembles the extracellular proteins hemopexin and vitronectin. This domain is believed to contribute to substrate specificity and represents a binding site for the natural inhibitors of the



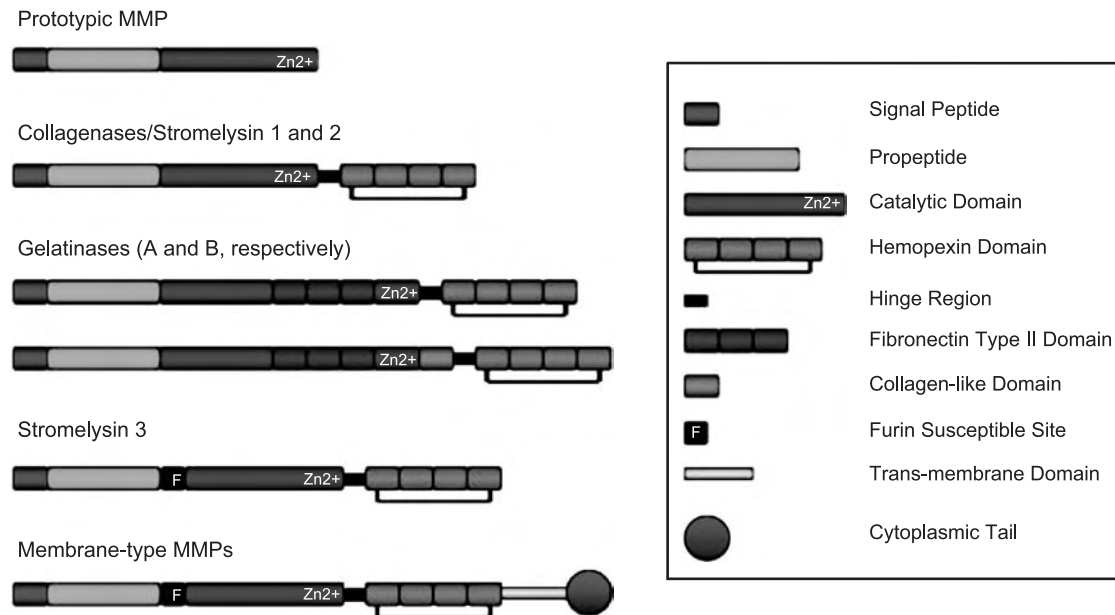


Fig. 1 Domain structure of representative MMPs. Signal peptide; propeptide with the cysteine switch; catalytic domain containing the Zn^{2+} binding site; hemopexin domain that purportedly contributes to MMP substrate specificity; fibronectin type II domains for gelatin binding, collagen-like domain with an unknown function; furin susceptible site, transmembrane domain and the cytoplasmic tail. (Modified from Ref. [26].)

soluble MMPs, the TIMPs. The hinge region acts as a linker that connects the catalytic domain to the hemopexinlike domain. No specific functional characteristics have been ascribed to this structural feature. The fibronectinlike repeats consist of three fibronectin type II-like domains arrayed sequentially in the primary structure of the peptide. These repeats appear to mediate the adhesion of gelatinases to denatured collagen. The membrane insertion domain is a hallmark of the membrane-bound MMPs (MT-MMP); this domain traverses the cell membrane and anchors the MT-MMP protein to the plasma membrane.^[3] We will consider the specific characteristics of representative MMP enzymes from each of these subgroups in the following discussion.

COLLAGENASES

In humans, members of this MMP subfamily include MMP-1 (Collagenase 1), MMP-8 (Collagenase 2) and MMP-13 (Collagenase 3).^[2] These enzymes exhibit activity against fibrillar types I, II, and III collagen. These collagen isotypes represent the most abundant extracellular structural proteins in man and play a role in the development, maintenance, repair, and regeneration of almost every tissue type. Type I collagen is present in high concentrations in bone, skin, and the connective tissue of muscle. A blend of type I and type III collagen is observed in the myocardium and in blood

vessels. Type II collagen predominates in cartilage. Considerable effort has been directed at processing proteins of the collagen family into tissue-engineering scaffolds that mimic the structure of native tissue.^[28,29]

In clinical practice, we believe that MMP-1 and MMP-8 will have the most relevance to a bioengineering paradigm. Specific roles for the collagenases in developmental and pathological processes have been defined in transgenic mice models. For example, mutations at or around the collagenase cleavage site on the alpha 1 (I) chain of type I collagen render it resistant to collagenase digestion *in vitro*. Mice carrying a collagenase-resistant Colla-1 transgene die in late embryogenesis. When the same mutation is introduced into the endogenous Colla-1 sequence by gene targeting, the host undergoes normal development into early adulthood. However, with aging the animals develop marked fibrosis in the dermis in a pattern similar to that observed in human scleroderma.^[30]

MMP-1 (Interstitial Collagenase)

MMP-1 is expressed by interstitial fibroblasts in a variety of tissue types and macrophages.^[31] The nearly ubiquitous distribution of MMP-1 virtually ensures that this protease will be present and activated within any tissue-engineering site. The enzyme contains a minimal set of MMP-specific domains. There is a signal peptide, a propeptide, a hinge region, a catalytic domain, and a hemopexinlike domain. Molecular weight

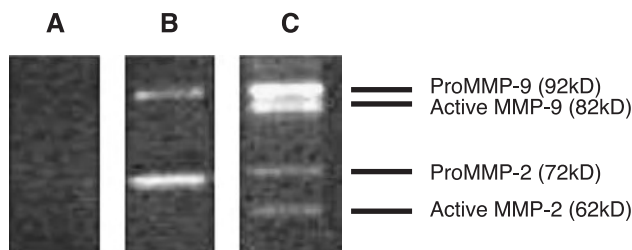


Fig. 2 Representative gelatin zymography. Equal concentration concentrated media proteins were electrophoresed at a constant voltage of 120 V for 1½ hrs. Lane A, No evidence of gelatinolytic activity in serum-free media (DMEM/F-12 supplemented with 5 ml ITS liquid media supplement (100 ×), 1.2% bicarbonate, 1.2% antimycotic-antibiotic). Lane B, Clear bands at 92 kDa and 72 kDa present in 7.5% serum-supplemented media representing the inactive zymogens MMP-9 and MMP-2 respectively. Lane 3, neonatal cardiac fibroblast plated onto a bed of fibronectin [10 µg/ml] in serum-free media. Clear bands observed for both the inactive and active forms of the gelatinases.

of the proisoform is 52 kDa and 41 kDa for the active enzyme. MMP-1 is synthesized and deposited into the extracellular environment as an inactive zymogen (proMMP-1) in response to environmental signals.

The activation cascade of MMP-1 is a multistep process, a feature that represents a common regulatory pattern for the enzymes of this subclass. In the extracellular space the propeptide is partially cleaved from the enzyme to form an intermediate, inactive zymogen species^[2,3] (Fig. 2). This initial cleavage event can be mediated by several different extracellular proteases, including plasmin,^[2] mast cell factor,^[32] soluble proteases present in rheumatoid synovial fluid,^[33] and cartilage procollagenase-activating protein.^[2] The partial cleavage of the propeptide sequence results in a conformational change that initiates an autocatalytic event and the removal of the remaining amino acid residues of the propeptide sequence, forming an intermediate 42 kDa isoform.^[3] This partially activated isoform is relatively short-lived and is further processed to a mature 41-kDa protein that exhibits full collagenase activity. The final proteolytic events that lead to the maturation of the enzyme appear to be mediated predominately by MMP-2 and MMP-3.^[2] The mature isoform of MMP-1 exhibits a degree of substrate specificity and cleaves type III collagen 100 times more rapidly than type II collagen and 10 times more rapidly than type I collagen.^[3]

MMP-8 (Neutrophil Collagenase)

MMP-8 is similar in molecular size to MMP-1 and shares several subdomains with this peptide. Despite this structural homology, MMP-8 exhibits very different

patterns of regulation. MMP-8 was originally believed to be specific to neutrophils and polymorphonuclear leukocytes (PMNs). Further study has demonstrated that normal chondrocytes, synovial fibroblasts, and endothelial cells all express this protease. Notably, the expression of MMP-8 is accelerated in chondrocytes during the onset of osteoarthritis.^[34] This is a degenerative condition associated with the inflammation of the articular cartilage and the surrounding synovial lining. The elevated levels of MMP-8 observed in osteoarthritis originate from intrinsic sources (i.e., the chondrocytes)^[34] and from extrinsic sources (i.e., the PMNs that accumulate in response to the inflammatory process). It remains to be determined if the activation of MMP-8 represents a cause or a consequence of this disease process. However, collagen degradation is recognized as an irreversible step in the destruction of cartilage, ultimately leading to a failure in joint function.^[35]

The molecular weight of the proisoform of MMP-8 is 60 kDa and the mature active enzyme has a molecular weight of 40 kDa. Two different regulatory cascades have been described for this protease. In the PMN cell, MMP-8 is stored within specific intracellular granules.^[36] During the progression of an inflammatory response, macrophages release cytokines and promote the accumulation of PMNs. Once activated, the PMNs release MMP-8 as a bolus into the local microenvironment, accelerating matrix degradation.^[36,37] The release of MMP-8 in this fashion results in a highly localized and concentrated response.

In other cell types, for example, the chondrocyte, MMP-8 is expressed and deposited into the ECM as an inactive proenzyme. Within the extracellular space the enzyme can undergo activation by an autocatalytic cascade similar to that described for MMP-1. Activation can also occur through proteolytic events mediated by a variety of other soluble extracellular proteases.^[3] For enzymes stored in the extracellular matrix, the rate of activation can be modulated by the relative amount of oxygen present within the interstitium. This regulatory pathway can have direct implications in the tissue-engineering paradigm where implanted tissue may be subject to hypoxia. Under aerobic conditions during an inflammatory response, oxidizing substances formed by neutrophils can rapidly initiate the autocatalytic cascade.^[38] Subsequent to this initial processing event the enzyme appears to be cleaved by cathepsin G, leading to a fully activated isoenzyme.^[38] On the surface the activation of this protease through this pathway appears to be completely detrimental to the healing process. However, the matrix remodeling that transpires under these conditions appears to be associated with the re-expression of a more natively like ECM architecture. In contrast, under anaerobic conditions a slower, more progressive proteolytic activation that is mediated cathepsin G, chymases, and/or tryptases



predominates.^[31] Activation through this cascade, while slower, may lead to aberrant ECM remodeling and subsequent fibrosis. In the absence of circulatory support, the accumulation of cellular debris and remnants of provisional matrix constituents results in the build-up of protein within the interstitial fluid.

The complex interplay of regulatory events that control the deposition and activation of MMP-8 are clear demonstrations of how difficult it can be to manage protease activity associated with the MMP family. In osteoarthritis, MMP-8 originates from intrinsic and extrinsic sources and is deposited into the extracellular environment by two completely different processes. Once deposited into the extracellular environment the regulatory cascades can be further modulated by the local concentration of proteases, oxidizing agents, and oxygen tension.

GELATINASES

Members of the gelatinase family include Gelatinase A (MMP-2, 72 kDa) and Gelatinase B (MMP-9, 92 kDa). Collagen is not degraded solely by any one proteolytic enzyme and, in effect, the enzymes of the gelatinase subfamily mediate the relatively late stages of collagen degradation. The complete degradation of this natural polymer is mediated by a battery of proteases. During initial processing, a MMP collagenase or other protease may degrade fibrillar collagen into a smaller fragment. Subsequently, the fragments are further digested by enzymes of the gelatinase family into yet smaller entities. Gelatinases are expressed in nearly every tissue type^[39] and have been described in a variety of interstitial fibroblasts. Expression has also been detected in fully differentiated cell types including, skeletal, cardiac, and smooth muscle cells and endothelial cells. Proteolytic activity associated with these enzymes is critical during angiogenesis,^[40,41] a process essential to the bioengineering paradigm. MMP-2 is usually expressed constitutively at low levels; MMP-9 is expressed in response to environmental signals.

The gelatinases share extensive structural homology with the other members of the MMP family. However, the fibronectin type II repeats within these proteases represent a distinguishing, and diagnostic, feature. This motif contains a 58-amino-acid repeat that adds considerable mass to the gelatinase MMP, making these proteases among the largest of the MMP family. The fibronectin II repeats mediate the adhesion of the MMP gelatinase to elastin and denatured, or partially degraded, collagen (gelatins). This physical interaction appears necessary for the proteolytic activity of the enzyme against these structural proteins.^[42,43] Subtle differences in the sequence of the fibronectinlike repeats

appear to confer a degree of substrate specificity to the enzymes of this family.^[44,45]

Gelatinase activation is classically studied *in vitro* by zymography. Zymography involves electrophoretic separation of denatured but nonreduced proteins in a gelatin-infused poly-acrylamide gel. At the conclusion of the separation, proteins are renatured by exchanging the SDS with a nonionic detergent such as 2.5% (v/v) of Triton X-100 in ddH₂O or proprietary renaturation buffer (BioRad Laboratories, Hercules, CA). The renatured proteins in the gel are incubated in development buffer: 0.05 M Tris-HCl pH 8.8, 5 mM CaCl₂, 0.02% NaN₃. Development buffers are also available commercially. Gels are incubated for 12–24 hours at 37°C revealing clear bands, indicating proteolysis of the gelatin substrate. After the appropriate incubation period, developed gels are fixed in destaining solution (20% methanol, 5% acetic acid, and 75% distilled H₂O) and stained with 0.5% Coomassie brilliant blue R-250 prepared in 45% methanol, 10% acetic acid, and 45% dH₂O. Stained gels are destained in destaining solution until clear bands against a blue background can be visualized (Fig. 2). Clear bands can be analyzed by using densitometry software. For molecular weight comparison, gelatinase zymography standards are available commercially through Chemicon (Temecula, CA). For a complete and more detailed description of zymography refer to *Methods in Molecular Biology*, edited by Ian M. Clark.^[46]

The gelatinases are expressed at elevated levels during tissue remodeling and accelerated gelatinase activity is associated with trophoblast implantation,^[47] bone resorption,^[8] and the process of angiogenesis.^[40,41] In the cardiovascular system, increased gelatinase expression is correlated with morphogenic events during early cardiogenesis, including the remodeling of the pharyngeal arches, heart tube looping, cardiac cushion formation, and the migration of neural crest cells.^[6] In disease states, gelatinase activity is augmented during wound healing and in the progression of rheumatoid and osteogenic arthritis, arthrosclerosis,^[48] dilatory cardiovascular disease,^[49] and the abnormal resorption of periodontal ligaments.^[50] The metastatic potential of some cancers has been correlated with increased gelatinase expression; presumably, this proteolytic activity facilitates the migration of the cells through the basal lamina and into the interstitial compartments.^[51,52]

MMP-2 (Gelatinase A)

MMP-2, a prototypic enzyme of the gelatinase family, is constitutively expressed at low levels by many different cell types. It is secreted into the ECM as a 72-kDa latent proenzyme (zymogen). A series of parallel and overlapping activation cascades regulate the catalytic

activity of the enzyme. The activated isoform of the protein exhibits proteolytic activity against gelatin (denatured triple-helical type I and type III collagen), type IV, V, VII, and IX collagen, fibronectin, laminin, elastin, a variety of proteoglycans, and entactin.^[2] Over prolonged periods of time, MMP-2 also has been shown to bind and cleave soluble type I collagen.^[3]

In the extracellular space the cleavage of the propeptide from the parent proMMP-2 isoform can be achieved through several different pathways. In one well-characterized pathway, proMMP-2 is activated through the formation of a multimolecular complex. In the inactive state the proMMP-2 isoform binds to Tissue Inhibitor Metalloproteinase-2 (TIMP-2), a peptide inhibitor that suppresses the conversion of 72-kDa proMMP-2 to the active 68-kDa isoform.^[2,3] There are two separate binding sites on the proMMP-2 molecule for TIMP-2. One site is adjacent to the active site of the enzyme and the other appears to represent a novel stabilization site. At equimolar concentrations of TIMP-2 and proMMP-2, the conversion of proMMP-2 to the active isoform is inhibited. The binding of TIMP-2 to the stabilization site appears to maintain the tertiary structure of the enzyme and preserve the proisoform by preventing the progression of an autocatalytic activation cascade.^[3] As the concentration of TIMP-2 is increased, the inhibitor binds to a membrane-bound protease of the MMP family, MT1-MMP.^[53,54] The interaction of a TIMP-2 : proMMP-2 complex to MT1-MMP forms a ternary complex at the cell surface. This complex positions the bound proMMP-2 in proximity to an adjacent and free MT1-MMP, leading to the proteolytic processing and subsequent activation of the MMP-2 molecule. At very high concentrations TIMP-2 binds proMMP-2 and saturates the MT1-MMP compartment, neutralizing the activation cascade.^[55] MMP-2 may also be activated by an autoactivation pathway. At high local concentrations, proMMP-2 can be processed by a proteolytic event to an active enzyme, without the addition of exogenous activation factors. It is unclear if the autocatalysis represents a physiologically relevant event.

MMP-9 (Gelatinase B)

MMP-9 is the largest (92 kDa) of the known MMPs. This enzyme exhibits a unique collagen type V-like domain that is located between the catalytic and a hemopexinlike domain. *In vitro*, activated MMP-9 degrades type IV and type V collagen, aggrecan, link protein, myelin basic protein, and amyloid B peptide.^[3] The substrates degraded *in vivo* have not been fully characterized. However, MMP-9-deficient mice exhibit a specific defect in endochondral bone formation that is associated with the accumulation of hypertrophic

cartilage at the growth plates. These animals also exhibit diminished angiogenic activity.^[56]

Unlike MMP-2, MMP-9 is not expressed constitutively.^[57] This isotype is expressed in response to specific environmental signals including, growth factors, cytokines, cell-cell interactions, cell-matrix interactions, and external events that alter cell shape.^[3] The proteolytic activation cascade that regulates the processing of MMP-9 is subject to TIMP-dependent regulation in a pattern similar to MMP-2. TIMP-1 and TIMP-2 can both bind and suppress the conversion of proMMP-9 to the active state. TIMP-2 binds to MMP-9 with a higher affinity than TIMP-1 and is more effective at suppressing activation cascade; however, the TIMP-1-dependent regulation of MMP-9 appears to predominate *in vivo*.^[58]

Like other latent MMPs, the activation of proMMP-9 is mediated by the proteolytic disruption of the cysteine/ Zn^{2+} switch (coordination bond). During processing, the (92-kDa) proMMP-9 species is converted to an (82-kDa) active isoform. The specific pathway to full activation is dependent upon whether MMP-9 is bound in a complex with TIMP-1 or in a free state.^[3] In most cell types, proMMP-9 is deposited into the extracellular environment in complex with TIMP-1. Under these conditions the proMMP-9 : TIMP-1 complex is resistant to cleavage by MMP-3, a soluble protease of the stromelysin MMP subfamily.^[59] As the relative concentration of active MMP-3 increases with respect to the local TIMP concentration, MMP-3 is believed to cleave proMMP-9 to an 82-kDa isoform. This isoenzyme exhibits an intermediate degree of proteolytic activity against ECM substrates.^[59] In the absence of TIMP-1, the intermediate 82-kDa MMP-9 is processed further by MMP-3 to a fully active 67-kDa isoform^[3] (Fig. 3). A plasmin-dependent, TIMP-independent activation pathway for MMP-9^[57] also has been described. This pathway may provide a substantial contribution to the activation of MMP-9.^[60] Plasmin is derived from plasminogen, a peptide secreted by hepatocytes that is a nearly ubiquitous component of plasma.^[61]

STROMELYSINS

Members of the stromelysin subfamily include MMP-3 (Stromelysin-1) and MMP-10 (Stromelysin-2). MMP-3 and MMP-10 both contain typical MMP subdomains, including the signal peptide, the propeptide, and the catalytic domain. These enzymes have broad substrate specificity and can degrade basement membrane collagens, laminin, fibronectin, vitronectin, proteoglycans, and matrix glycoproteins. The stromelysins do not degrade fibrillar collagen;^[62] however, the enzymatic activity of this class of proteases plays a central role in the full activation of the collagenases and gelatinases



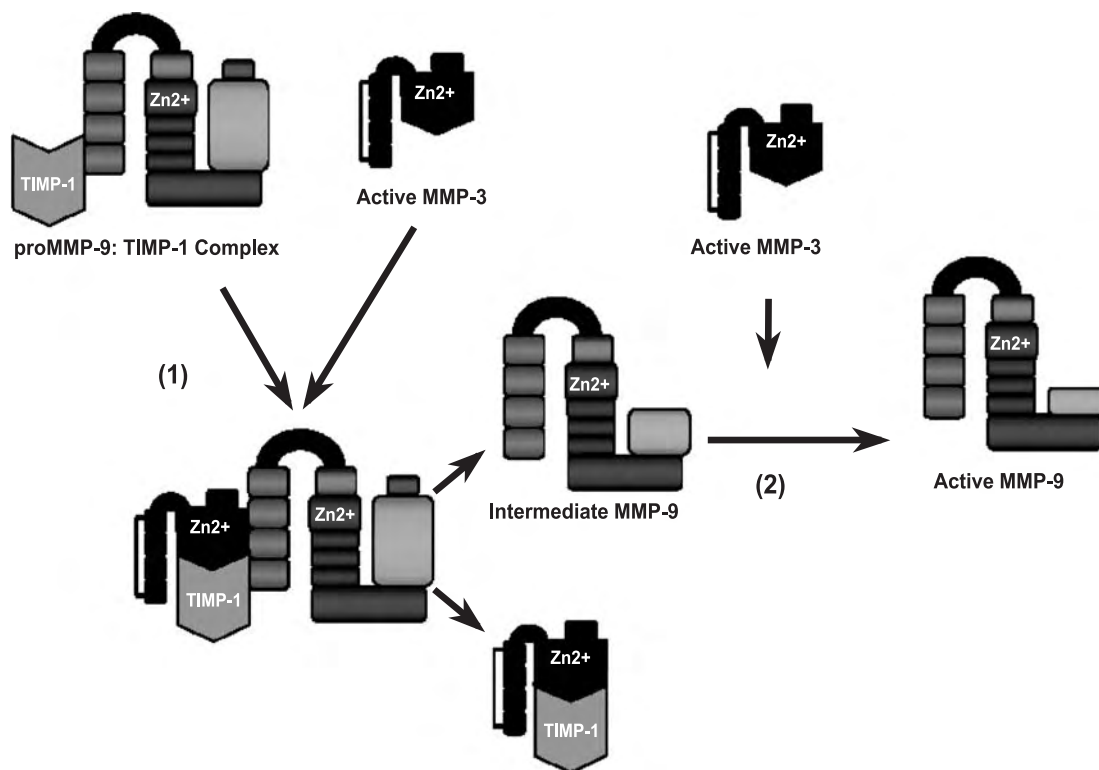


Fig. 3 Schematic representation for the activation of ProMMP-9. Step (1), the relative concentration of active MMP-3 increases with respect to the local TIMP concentration. This interaction is believed to cleave proMMP-9 to an 82-kDa isoform (intermediate level of activity). Step (2), now with the removal of TIMP-1, the intermediate 82-kDa MMP-9 is processed further by MMP-3 (stromelysin) to a fully active 67-kDa isoform. (Modified from Ref. [59].)

(refer to Fig. 3). Given this regulatory role and the broad substrate specificity of the stromelysins against matrix constituents, it is not surprising that these enzymes have been implicated in a wide variety of biological processes. During development there is coordinate expression of the stromelysins, the collagenases, and the gelatinases during blastocyst outgrowth and the early stages of implantation. Stromelysins are concentrated in the growth plates of developing and maturing bone. Tissue culture experiments suggest that osteoblasts express the stromelysins in response to bone resorptive agents and mediate the degradation of non-mineralized collagen.^[63]

Specific patterns of MMP-3 and MMP-10 expression have been observed in keratinocytes in the immediate vicinity of epithelial injuries.^[64] During the onset of proliferation, keratinocytes that are in contact with the basement membrane express MMP-3 but not MMP-10. The enzymatic activity associated with MMP-3 is correlated with the contraction of the wound bed. In MMP-3-null mice there is a delay in the onset of incisional wound closure,^[65] and dermal fibroblasts from these animals exhibit decreased capacity to contract collagen gels *in vitro*.^[66] MMP-10, but not MMP-3, is expressed in cells within the migrating front that forms the

epithelial tongue at the margin of a wound. The specific function of this enzyme in the migration of cells across the provisional matrix of a wound site remains to be fully defined.

MMP-3 (Stromelysin-1)

MMP-3 is expressed by many cell types; including interstitial fibroblasts, vascular smooth muscle cells, chondrocytes, keratinocytes, macrophages, and endothelial cells in response to extrinsic signals. RGD-dependent binding events, mechanical perturbations, cytokines, and growth factors can all invoke the expression and synthesis of this enzyme. The proMMP-3 molecule is a 52-kDa protein that has a signal sequence, a propeptide, a catalytic domain, and a hemopexinlike domain. The hemopexinlike domain, although not essential for catalytic activity, appears to define substrate specificity. This domain contains an amino acid repeat that represents a high-affinity binding site for the extracellular matrix proteins hemopexin and vitronectin.

The activation of proMMP-3 (52 kDa) to its 45-kDa active isoform follows a pattern similar to that of proMMP-1. The initial cleavage of proMMP-3 is

mediated by a large number of proteinases that attack the first α -helix of the peptide. This appears to destabilize the Cysteine/Zn⁺² switch, resulting in an activated intermediate isoform. The final activation steps of MMP-3 do not require other members of the MMP family and a diverse variety of proteases can mediate this event.^[3]

MEMBRANE-TYPE MATRIX METALLOPROTEINASES

The membrane-type matrix metalloproteinases (MT-MMP) are tethered to the cell membrane. The prototypic MT-MMP contains a C-terminal hydrophobic transmembrane domain that passes through the plasma membrane terminating in a short cytoplasmic tail. The characteristic location of these enzymes establishes focused sites of proteolytic activity. Common to all MT-MMPs is the insertion of a potential furin cleavage site adjacent to the propeptide domain.^[3] The function of this furin site remains controversial. In some cell types the cleavage of this site appears to down regulate the proteolytic capacity of the mature enzyme.^[27] In other cell types this site appears to mediate the association of the MT-MMPs with TIMPs during the initial stages of the activation cascades that regulate the maturation of the soluble MMPs.^[67]

MT1-MMP

Of the six known MT-MMPs, MT1-MMP is perhaps the most well characterized. The molecular weight of the pro-MT1-MMP is 64 kDa, whereas the fully active form has a molecular weight of 54 kDa. The pro-MT-MMP can be activated during intracellular transport to the cell surface by a furinlike proteinase and/or at the cell surface by furin or plasmin-mediated processing events. Activation can also be mediated at the cell surface by other proteases, but these events have not yet been fully characterized.^[3] As discussed, MT1-MMP is best recognized for its essential role in the activation of proMMP-2 (Gelatinase A), but this protease can directly degrade gelatin (denatured collagen), aggrecans, fibronectin, vitronectin, and laminin.

MMPs IN TISSUE ENGINEERING

Enzymes of the MMP family are widely distributed in the mammalian system and exhibit robust proteolytic activity against a wide variety of matrix constituents. A broad spectrum of regulatory factors controls the synthesis, expression, and activation of these proteases. Within in any given local domain protease activity

associated with the MMPs may arise from the resident cells of an organ system or from cells that originate from the circulatory system. As a consequence of the ubiquitous distribution of these enzymes nearly any surgical procedure designed to implant a bioengineered organ can be expected to initiate MMP protease activity.

In the design of tissue-engineering scaffolds a balance must be achieved between structural and functional considerations. Scaffolds that can withstand manual manipulation and simultaneously support cell infiltration and the integration of the implant into the host site must be developed. The support of cellular infiltration is critical to the formation of nascent blood vessels. Ideally, a tissue-engineering scaffold should define the initial structural properties of a bioengineered construct until the cells of that construct can elaborate a native extracellular matrix. To reduce the risks of inflammation and fibrosis, the ideal fabrication scaffolding also should be biodegradable and be replaced by matrix materials expressed by the implant.

Prospects for developing synthetic materials that meet the enumerated design considerations seem unlikely. Identifying synthetic materials that exhibit biological, regulatory, and commercial efficacy is not a trivial task. Synthetics that function well in the research setting may face intense scrutiny from regulatory agencies (e.g., The Food and Drug Administration). The tripartite hurdle of demonstrating biological efficacy, meeting regulatory approval, and exhibiting commercial viability is often overlooked and represents an unappreciated constraint that can limit the movement of materials from the laboratory setting to clinical practice.

Our laboratories believe that novel methods must be developed to process native or true analogs of native materials into surgical biomaterials suitable for use as tissue-engineering scaffolds.^[68,69] This approach, coupled with tissue fabrication and surgical strategies that limit the remodeling response, represent an avenue to manage the potentially destructive protease activity associated with the MMP family. We have approached this issue by developing scaffolds that more closely mimic the chemical composition and material properties of the host site of an implant.^[70] MMP-mediated protease activity is critical to a diverse variety of homeostatic processes and it seems unlikely that it will be feasible to suppress the activity of these enzymes completely, even within the local microenvironment of a tissue-engineered implant. Some degree of targeted disruption, perhaps directed at key regulatory steps in the MMP activation cascades, may be tolerated and have efficacy over short intervals of time.^[71] However, with nonspecific or prolonged suppression of MMP-mediated protease activity, cell infiltration and capillary formation will be primarily defined by the porosity of the scaffolding material and density of the implant. This



places additional design constraints on the scaffold and the bioengineered organ. Until a functional capillary network can develop, most bioengineered constructs must be supported by the passive diffusion of nutrients from the surrounding interstitium. Nonspecifically inhibiting MMP-mediated protease activity can be expected to slow healing^[49] and disrupt the formation of nascent capillary networks, placing additional limitations on the mass of tissue that can be successfully implanted.

ACKNOWLEDGMENT

The Jeffress Grant Foundation (Grant# J-680) is gratefully acknowledged for their financial support of our work on MMPs.

ARTICLES OF FURTHER INTEREST

Angiogenesis Inhibitors; Collagen; Elastin; Extracellular Matrix Scaffolds

REFERENCES

- Zisch, A.H.; Lutolf, M.P.; Hubbell, J.A. Biopolymeric delivery matrices for angiogenic growth factors. *Cardiovasc. Pathol.* **2003**, *12* (6), 295–310.
- Woessner, J.F.; Nagase, H. *Matrix Metalloproteinases and TIMPs*; Oxford University Press: New York, 2000; 1.
- Woessner, J.F., Jr. The Matrix Metalloproteinase Family. In *Matrix Metalloproteinases*; Parks, W.C., Mecham, R.P., Eds.; Academic Press: San Diego, 1998; 1–13.
- Gross, J.; Lapiere, C.M. Collagenolytic activity in amphibian tissues: A tissue culture assay. *Proc. Natl. Acad. Sci.* **1962**, *48*, 1014–1022.
- Werb, Z.; Chin, J.R. Extracellular matrix remodeling during morphogenesis. *Ann. N. Y. Acad. Sci.* **1998**, *857*, 110–118.
- Cai, D.H.; Vollberg, T.M., Sr.; Hahn-Dantona, E.; Quigley, J.P.; Brauer, P.R. MMP-2 expression during early avian cardiac and neural crest morphogenesis. *Anat. Rec.* **2000**, *259* (2), 168–179.
- Itoh, Y.; Nagase, H. Matrix metalloproteinases in cancer. *Essays Biochem.* **2002**, *38*, 21–36.
- Uchida, M.; Shima, M.; Shimoaka, T.; Fujieda, A.; Obara, K.; Suzuki, H.; Nagai, Y.; Ikeda, T.; Yamato, H.; Kawaguchi, H. Regulation of matrix metalloproteinases (MMPs) and tissue inhibitors of metalloproteinases (TIMPs) by bone resorptive factors in osteoblastic cells. *J. Cell. Physiol.* **2000**, *185* (2), 207–214.
- Henry, M.T.; McMahon, K.; Mackarel, A.J.; Prikk, K.; Sorsa, T.; Maisi, P.; Sepper, R.; Fitzgerald, M.X.; O'Connor, C.M. Matrix metalloproteinases and tissue inhibitor of metalloproteinase-1 in sarcoidosis and IPF. *Eur. Respir. J.* **2002**, *20* (5), 1220–1227.
- Sun, H.B.; Yokota, H. Messenger-RNA expression of matrix metalloproteinases, tissue inhibitors of metalloproteinases, and transcription factors in rheumatic synovial cells under mechanical stimuli. *Bone* **2001**, *28* (3), 303–309.
- Cawston, T. Matrix metalloproteinases and TIMPs: Properties and implications for the rheumatic diseases. *Mol. Med. Today* **1998**, *4* (3), 130–137.
- Ladwig, G.P.; Robson, M.C.; Liu, R.; Kuhn, M.A.; Muir, D.F.; Schultz, G.S. Ratios of activated matrix metalloproteinase-9 to tissue inhibitor of matrix metalloproteinase-1 in wound fluids are inversely correlated with healing of pressure ulcers. *Wound Repair Regen.* **2002**, *10* (1), 26–37.
- Mirastschijski, U.; Impola, U.; Jahkola, T.; Karlsmark, T.; MS, A.G.; Saarialho-Kere, U. Ectopic localization of matrix metalloproteinase-9 in chronic cutaneous wounds. *Human Pathol.* **2002**, *33* (3), 355–364.
- Fromigue, O.; Louis, K.; Wu, E.; Belhacene, N.; Loubat, A.; Shipp, M.; Auburger, P.; Mari, B. Active stromelysin-3 (MMP-11) increases MCF-7 survival in three-dimensional Matrigel culture via activation of p42/p44 MAP-kinase. *Int. J. Cancer* **2003**, *106* (3), 355–363.
- Pan, W.; Arnone, M.; Kendall, M.; Grafstrom, R.H.; Seitz, S.P.; Wasserman, Z.R.; Albright, C.F. Identification of peptide substrates for human MMP-11 (stromelysin-3) using phage display. *J. Biol. Chem.* **2003**, *278* (30), 27820–27827.
- d'Ortho, M.P.; Will, H.; Atkinson, S.; Butler, G.; Messent, A.; Gavrilovic, J.; Smith, B.; Timpl, R.; Zardi, L.; Murphy, G. Membrane-type matrix metalloproteinases 1 and 2 exhibit broad-spectrum proteolytic capacities comparable to many matrix metalloproteinases. *Eur. J. Biochem.* **1997**, *250* (3), 751–757.
- Shimada, T.; Nakamura, H.; Ohuchi, E.; Fujii, Y.; Murakami, Y.; Sato, H.; Seiki, M.; Okada, Y. Characterization of a truncated recombinant form of human membrane type 3 matrix metalloproteinase. *Eur. J. Biochem.* **1999**, *262* (3), 907–914.
- Stracke, J.O.; Fosang, A.J.; Last, K.; Mercuri, F.A.; Pendas, A.M.; Llano, E.; Perris, R.; Di Cesare, P.E.; Murphy, G.; Knauper, V. Matrix metalloproteinases 19 and 20 cleave aggrecan and cartilage oligomeric matrix protein (COMP). *FEBS Lett.* **2000**, *478* (1–2), 52–56.
- Bartlett, J.D.; Ryu, O.H.; Xue, J.; Simmer, J.P.; Margolis, H.C. Enamelysin mRNA displays a developmentally defined pattern of expression and encodes a protein which degrades amelogenin. *Connect. Tissue Res.* **1998**, *39* (1–3), 101–109. Discussion 141–149.
- Armstrong, D.G.; Jude, E.B. The role of matrix metalloproteinases in wound healing. *J. Am. Podiatr. Med. Assoc.* **2002**, *92* (1), 12–18.
- Nwomeh, B.C.; Liang, H.X.; Cohen, I.K.; Yager, D.R. MMP-8 is the predominant collagenase in healing wounds and nonhealing ulcers. *J. Surg. Res.* **1999**, *87* (2), 189–195.
- Dollery, C.M.; McEwan, J.R.; Henney, A.M. Matrix metalloproteinases and cardiovascular disease. *Circ. Res.* **1995**, *77* (5), 863–868.
- Robbins, J.R.; McGuire, P.G.; Wehrle-Haller, B.; Rogers, S.L. Diminished matrix metalloproteinase 2

- (MMP-2) in ectomesenchyme-derived tissues of the Patch mutant mouse: Regulation of MMP-2 by PDGF and effects on mesenchymal cell migration. *Dev. Biol.* **1999**, *212* (2), 255–263.
24. Johansson, N.; Ahonen, M.; Kahari, V.M. Matrix metalloproteinases in tumor invasion. *Cell. Mol. Life Sci.* **2000**, *57* (1), 5–15.
25. Cleutjens, J.P. The role of matrix metalloproteinases in heart disease. *Cardiovasc. Res.* **1996**, *32* (5), 816–821.
26. Sternlicht, M.D.; Werb, Z. How matrix metalloproteinases regulate cell behavior. *Annu. Rev. Cell Dev. Biol.* **2001**, *17*, 463–516.
27. Wang, X.; Pei, D. Shedding of membrane type matrix metalloproteinase 5 by a furin-type convertase: A potential mechanism for down-regulation. *J. Biol. Chem.* **2001**, *276* (38), 35953–35960.
28. Kato, Y.P.; Silver, F.H. Formation of continuous collagen fibres: Evaluation of biocompatibility and mechanical properties. *Biomaterials* **1990**, *11* (3), 169–175.
29. Suzuki, S.; Matsuda, K.; Isshiki, N.; Tamada, Y.; Yoshioka, K.; Ikada, Y. Clinical evaluation of a new bilayer “artificial skin” composed of collagen sponge and silicone layer. *Br. J. Plast. Surg.* **1990**, *43* (1), 47–54.
30. Liu, X.; Wu, H.; Byrne, M.; Jeffrey, J.; Krane, S.; Jaenisch, R. A targeted mutation at the known collagenase cleavage site in mouse type I collagen impairs tissue remodeling. *J. Cell Biol.* **1995**, *130* (1), 227–237.
31. Sorsa, T.; Konttinen, Y.T.; Lindy, O.; Ritchlin, C.; Saari, H.; Suomalainen, K.; Eklund, K.K.; Santavirta, S. Collagenase in synovitis of rheumatoid arthritis. *Semin. Arthritis Rheum.* **1992**, *22* (1), 44–53.
32. Simpson, J.W.; Taylor, A.C. Regulation of gingival collagenase: A possible role for a mast cell factor. *Proc. Soc. Exp. Biol. Med.* **1974**, *145* (1), 42–47.
33. Kruze, D.; Wojtecka, E. Activation of leucocyte collagenase proenzyme by rheumatoid synovial fluid. *Biochim. Biophys. Acta* **1972**, *285* (2), 436–446.
34. Tetlow, L.C.; Adlam, D.J.; Woolley, D.E. Matrix metalloproteinase and proinflammatory cytokine production by chondrocytes of human osteoarthritic cartilage: Associations with degenerative changes. *Arthritis Rheum.* **2001**, *44* (3), 585–594.
35. Shingleton, W.D.; Hodges, D.J.; Brick, P.; Cawston, T.E. Collagenase: A key enzyme in collagen turnover. *Biochem. Cell. Biol.* **1996**, *74* (6), 759–775.
36. Hasty, K.A.; Reife, R.A.; Kang, A.H.; Stuart, J.M. The role of stromelysin in the cartilage destruction that accompanies inflammatory arthritis. *Arthritis Rheum.* **1990**, *33* (3), 388–397.
37. Reynolds, J.J. Collagenases and tissue inhibitors of metalloproteinases: A functional balance in tissue degradation. *Oral Dis.* **1996**, *2* (1), 70–76.
38. Claesson, R.; Johansson, A.; Belibasakis, G.; Hanstrom, L.; Kalfas, S. Release and activation of matrix metalloproteinase 8 from human neutrophils triggered by the leukotoxin of *Actinobacillus actinomycetemcomitans*. *J. Periodontal Res.* **2002**, *37* (5), 353–359.
39. Bailey, A.J. Perspective article: The fate of collagen implants in tissue defects. *Wound Repair Regen.* **2000**, *8* (1), 5–12.
40. Luttun, A.; Dewerchin, M.; Collen, D.; Carmeliet, P. The role of proteinases in angiogenesis, heart development, restenosis, atherosclerosis, myocardial ischemia, and stroke: Insights from genetic studies. *Curr. Atheroscler. Rep.* **2000**, *2* (5), 407–416.
41. Nguyen, M.; Arkell, J.; Jackson, C.J. Human endothelial gelatinases and angiogenesis. *Int. J. Biochem. Cell Biol.* **2001**, *33* (10), 960–970.
42. Murphy, G.; Nguyen, Q.; Cockett, M.I.; Atkinson, S.J.; Allan, J.A.; Knight, C.G.; Willenbrock, F.; Docherty, A.J. Assessment of the role of the fibronectin-like domain of gelatinase A by analysis of a deletion mutant. *J. Biol. Chem.* **1994**, *269* (9), 6632–6636.
43. Shipley, J.M.; Doyle, G.A.; Fliszar, C.J.; Ye, Q.Z.; Johnson, L.L.; Shapiro, S.D.; Welgus, H.G.; Senior, R.M. The structural basis for the elastolytic activity of the 92-kDa and 72-kDa gelatinases. Role of the fibronectin type II-like repeats. *J. Biol. Chem.* **1996**, *271* (8), 4335–4341.
44. O’Farrell, T.J.; Pourmotabbed, T. The fibronectin-like domain is required for the type V and XI collagenolytic activity of gelatinase B. *Arch. Biochem. Biophys.* **1998**, *354* (1), 24–30.
45. O’Farrell, T.J.; Pourmotabbed, T. Identification of structural elements important for matrix metalloproteinase type V collagenolytic activity as revealed by chimeric enzymes. Role of fibronectin-like domain and active site of gelatinase B. *J. Biol. Chem.* **2000**, *275* (36), 27964–27972.
46. Hawkes, S.P.; Hongxia, L.; Taniguchi, G.T. *Matrix Metalloproteinase Protocols, in Methods in Molecular Biology*; Clark, I.M., Ed.; Humana Press: Totowa, NJ, 2001.
47. Bischof, P.; Meisser, A.; Campana, A. Control of MMP-9 expression at the maternal-fetal interface. *J. Reprod. Immunol.* **2002**, *55* (1–2), 3–10.
48. Lijnen, H.R. Extracellular proteolysis in the development and progression of atherosclerosis. *Biochem. Soc. Trans.* **2002**, *30* (2), 163–167.
49. Galis, Z.S.; Khatri, J.J. Matrix metalloproteinases in vascular remodeling and atherogenesis: The good, the bad, and the ugly. *Circ. Res.* **2002**, *90* (3), 251–262.
50. Chang, Y.C.; Yang, S.F.; Lai, C.C.; Liu, J.Y.; Hsieh, Y.S. Regulation of matrix metalloproteinase production by cytokines, pharmacological agents and periodontal pathogens in human periodontal ligament fibroblast cultures. *J. Periodontal Res.* **2002**, *37* (3), 196–203.
51. Opdenakker, G.; Van Damme, J. Chemokines and proteinases in autoimmune diseases and cancer. *Verh. K. Acad. Geneesk. Belg.* **2002**, *64* (2), 105–136.
52. Giannelli, G.; Antonaci, S. Gelatinases and their inhibitors in tumor metastasis: From biological research to medical applications. *Histol. Histopathol.* **2002**, *17* (1), 339–345.
53. Strongin, A.Y.; Collier, I.; Bannikov, G.; Marmer, B.L.; Grant, G.A.; Goldberg, G.I. Mechanism of cell surface activation of 72-kDa type IV collagenase. Isolation of the activated form of the membrane metalloprotease. *J. Biol. Chem.* **1995**, *270* (10), 5331–5338.



54. Sato, H.; Takino, T.; Okada, Y.; Cao, J.; Shinagawa, A.; Yamamoto, E.; Seiki, M. A matrix metalloproteinase expressed on the surface of invasive tumour cells. *Nature* **1994**, *370* (6484), 61–65.
55. Li, H.; Simon, H.; Bocan, T.M.; Peterson, J.T. MMP/TIMP expression in spontaneously hypertensive heart failure rats: The effect of ACE-and MMP-inhibition. *Cardiovasc. Res.* **2000**, *46* (2), 298–306.
56. Vu, T.H.; Shipley, J.M.; Bergers, G.; Berger, J.E.; Helms, J.A.; Hanahan, D.; Shapiro, S.D.; Senior, R.M.; Werb, Z. MMP-9/gelatinase B is a key regulator of growth plate angiogenesis and apoptosis of hypertrophic chondrocytes. *Cell* **1998**, *93* (3), 411–422.
57. Ramos-DeSimone, N.; Hahn-Dantona, E.; Siple, J.; Nagase, H.; French, D.L.; Quigley, J.P. Activation of matrix metalloproteinase-9 (MMP-9) via a converging plasmin/stromelysin-1 cascade enhances tumor cell invasion. *J. Biol. Chem.* **1999**, *274* (19), 13066–13076.
58. Howard, E.W.; Bullen, E.C.; Banda, M.J. Preferential inhibition of 72-and 92-kDa gelatinases by tissue inhibitor of metalloproteinases-2. *J. Biol. Chem.* **1991**, *266* (20), 13070–13075.
59. Ogata, Y.; Itoh, Y.; Nagase, H. Steps involved in activation of the pro-matrix metalloproteinase 9 (progelatinase B)-tissue inhibitor of metalloproteinases-1 complex by 4-aminophenylmercuric acetate and proteinases. *J. Biol. Chem.* **1995**, *270* (31), 18506–18511.
60. Devy, L.; Noel, A.; Baramova, E.; Bajou, K.; Tentesaux, C.; Jardillier, J.C.; Foidart, J.M.; Jeannesson, P. Production and activation of matrix metalloproteinase-9 (MMP-9) by HL-60 promyelocytic leukemia cells. *Biochem. Biophys. Res. Commun.* **1997**, *238* (3), 842–846.
61. Irigoyen, J.P.; Munoz-Canoves, P.; Montero, L.; Koziczak, M.; Nagamine, Y. The plasminogen activator system: Biology and regulation. *Cell. Mol. Life Sci.* **1999**, *56* (1–2), 104–132.
62. Woessner, J.F., Jr. Matrix metalloproteinases and their inhibitors in connective tissue remodeling. *FASEB J.* **1991**, *5* (8), 2145–2154.
63. Meikle, M.C.; Bord, S.; Hembry, R.M.; Compston, J.; Croucher, P.I.; Reynolds, J.J. Human osteoblasts in culture synthesize collagenase and other matrix metalloproteinases in response to osteotropic hormones and cytokines. *J. Cell. Sci.* **1992**, *103* (Pt 4), 1093–1099.
64. Saarialho-Kere, U.K.; Pentland, A.P.; Birkedal-Hansen, H.; Parks, W.C.; Welgus, H.G. Distinct populations of basal keratinocytes express stromelysin-1 and stromelysin-2 in chronic wounds. *J. Clin. Invest.* **1994**, *94* (1), 79–88.
65. Bullard, K.M.; Lund, L.; Mudgett, J.S.; Mellin, T.N.; Hunt, T.K.; Murphy, B.; Ronan, J.; Werb, Z.; Banda, M.J. Impaired wound contraction in stromelysin-1-deficient mice. *Ann. Surg.* **1999**, *230* (2), 260–265.
66. Bullard, K.M.; Mudgett, J.; Scheuenstuhl, H.; Hunt, T.K.; Banda, M.J. Stromelysin-1-deficient fibroblasts display impaired contraction in vitro. *J. Surg. Res.* **1999**, *84* (1), 31–34.
67. Polette, M.; Birembaut, P. Membrane-type metalloproteinases in tumor invasion. *Int. J. Biochem. Cell Biol.* **1998**, *30* (11), 1195–1202.
68. Matthews, J.A.; Wnek, G.E.; Simpson, D.G.; Bowlin, G.L. Electrospinning of collagen nanofibers. *Biomacromolecules* **2002**, *3* (2), 232–238.
69. Matthews, J.A.; Boland, E.D.; Wnek, G.E.; Simpson, D.G.; Bowlin, G.L. Electrospinning of collagen type II: A feasibility study. *J. Bioact. Compat. Polym.* **2003**, *18*, 125–134.
70. Telemeco, T.; Bowlin, G.L.; Vaida, M.; Cohen, N.; Simpson, D.G. Collagen as a surgical biomaterial: Fabrication of a skeletal muscle prosthetic. *Nat. Biotechnol.* **2004**, *in review*.
71. Sawa, M.; Kiyoi, T.; Kurokawa, K.; Kumihara, H.; Yamamoto, M.; Miyasaka, T.; Ito, Y.; Hirayama, R.; Inoue, T.; Kirii, Y.; Nishiwaki, E.; Ohmoto, H.; Maeda, Y.; Ishibushi, E.; Inoue, Y.; Yoshino, K.; Kondo, H. New type of metalloproteinase inhibitor: Design and synthesis of new phosphonamide-based hydroxamic acids. *J. Med. Chem.* **2002**, *45* (4), 919–929.

Mechanical Circulatory Support

Eric J. Okum

Vigneshwar Kasirajan

Department of Cardiothoracic Surgery, Medical College of Virginia, Richmond, Virginia, U.S.A.

Robert S. D. Higgins

Department of Cardiovascular and Thoracic Surgery, Rush-Presbyterian-St. Luke's Medical Center, Chicago, Illinois, U.S.A.



INTRODUCTION

Cardiovascular disease remains the number one cause of death in the United States. Congestive heart failure (CHF) continues to increase in frequency and prevalence, even as the mortality from coronary artery disease and valvular disease decreases. In the United States alone, more than two million outpatient CHF clinic visits occur per year, at a cost of more than \$10 billion annually. Failure to optimize outpatient management leads to added inpatient care, burdening the healthcare system and the individual patient with additional demands. Total annual direct costs for managing and treating heart failure approach 60 billion dollars, a significant part of national healthcare expenditures. The median survival of patients with new-onset congestive heart failure (CHF) is five years. CHF affects some 4 to 5 million Americans, with approximately 500,000 new cases diagnosed each year. Coupled with the aging of the American population, CHF continues to present clinicians and policymakers with vexing clinical and financial complexities. Heart failure is a particularly important problem in the elderly population, where its prevalence is approaching 10%. In fact, more than 75% of patients with heart failure are more than 65 years old. With the continued aging of the population, both the incidence and the prevalence of heart failure are expected to increase. Heart failure is also the most frequent cause of hospitalization in those older than 65.

OVERVIEW

CHF is a complex clinical disorder. Both systolic and diastolic heart failure are associated with high rates of morbidity, mortality, and cost. New York Heart Association (NYHA) functional class assignment is widely recognized as a classification schema for severity of CHF. At any given time, 30% of patients are at

NYHA Class III or IV, indicating severe impairment in functional ability. Despite recent advances in medical therapy, the one- to two-year mortality rate is 40% to 50% for advanced heart failure, and the four- to five-year mortality rate ranges from 15% to 40% for those with asymptomatic left ventricular (LV) dysfunction or only mild to moderate symptoms. Pharmacologic therapy can ameliorate CHF symptoms and slow disease progression, but there are limitations. These include patient compliance with complex medical regimens and access issues relating to the cost of newer therapies (including natriuretic peptides). For severe end-stage disease, cardiac transplantation provides improvements in quality of life and intermediate long-term (10 year) survival approaching 50%. It is the therapy of choice for patients with refractory failure, for whom medical therapy has failed. The downstream effect of this is seen in the number of patients awaiting cardiac transplantation each year, a number limited by the critical shortage of donors. The lack of donor organs has provoked intense interest in mechanical devices to support the failing heart.

Research into cardiac assist devices gained momentum in the 1970s, when the National Heart, Lung, and Blood Institute (NHLBI) began funding research that resulted in the development of intra-abdominal left ventricular assist devices (LVADs). These devices evolved as a means to support the left ventricle (LV) in an era when the total artificial heart had not been fully developed. While clinical outcomes were initially discouraging, progress was made. The use of such devices as a temporizing measure for patients awaiting transplantation gained support and ultimately received FDA approval in 1997.

Several iterations of these assist devices have been developed. The Randomized Evaluation of Mechanical Assistance for the Treatment of Congestive Heart Failure (REMATCH) was conducted from 1998 through July 2001 in 20 cardiac transplant centers in the United States. The study enrolled 129 patients with

end-stage heart failure who had not responded to medical management and were ineligible for cardiac transplantation. This randomized trial showed significant extension of life and improvement in quality of life for a subset of heart failure patients.^[1] This has added evidence to support potential clinical destination applications where assist devices are utilized as permanent therapy, instead of the transplantation option. Also of interest is the finding by some clinicians that CHF patients supported by mechanical circulatory assist devices apparently recover and actually improve myocardial function (which allows removal of the assist device), achieving adequate cardiac function that sustains them for a period of time.

This article will focus on three FDA-approved ambulatory LVADs: the TCI HeartMate Vented Electric (VE), the TCI HeartMate Implantable Pneumatic (IP) (Thoratec Corporation, Pleasanton, CA), and the Novacor LVAS (WorldHeart Corporation, Ottawa, Canada). We will also provide some insights on the status of the total artificial heart, specifically highlighting the AbioCor Artificial Heart (Abiomed Corporation, Danvers, MA).

Each of these devices allows the patient to be ambulatory, and some patients can even be at home while waiting for transplant. But each device has its merits and potential problems, as will be seen when these

devices are described in detail. Complications including thromboembolism, bleeding, infection, and device failure will be reviewed, and future directions will be addressed.

PUMP PHYSIOLOGY

The left ventricular assist devices support the left side of the heart and hence require an adequately functioning right side and low pulmonary vascular resistance to allow blood flow through the lungs (Fig. 1). The inflow cannula drains the left ventricle passively, so the entire left ventricle becomes a passive conduit. The pump body then ejects the blood through the outflow graft, unidirectional bloodflow being maintained by the inflow and outflow valves. The blood is ejected into the ascending aorta, so the aortic valve remains closed. Also, due to the drainage of the LV the interventricular (IV) septum is moved toward the LV. This may in fact reduce the efficiency of the right ventricular function, but in most instances the reduction in left ventricular pressure leads to improved afterload reduction to the right ventricle. This is more than enough to improve the RV function to compensate for the IV septum shift.

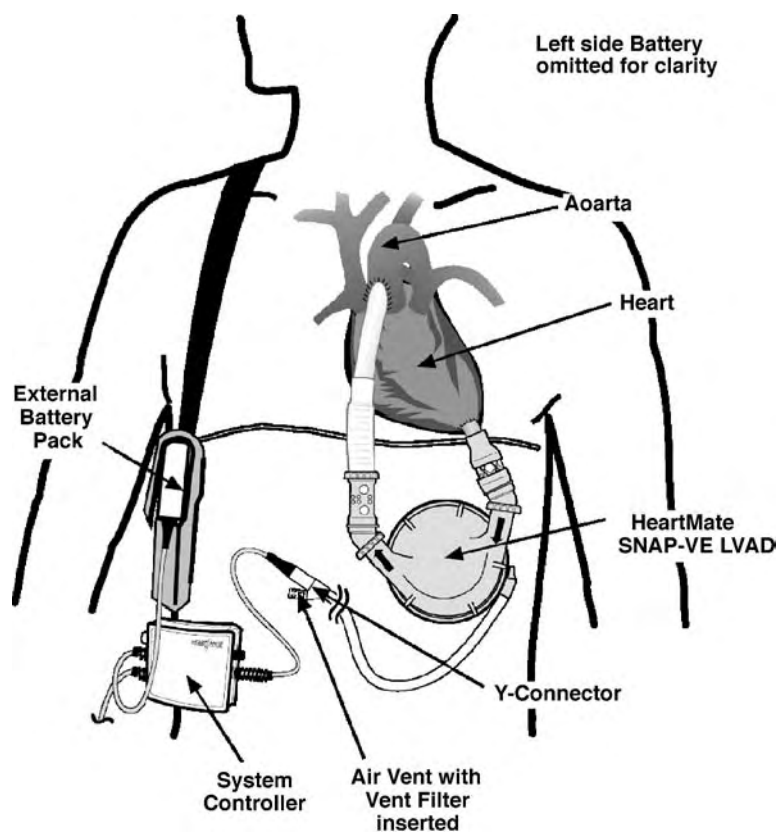


Fig. 1. LVAD pump physiology.. (Reprinted with permission from Thoratec Corporation.)

AMBULATORY IMPLANTABLE LEFT VENTRICULAR ASSIST DEVICES

HeartMate Vented Electric (VE)

The HeartMate VE is an implantable pulsatile pump made of titanium (Fig. 2). It has a polyurethane diaphragm that is interposed between two pusher plates. The power source is an internal electric motor supplied by a battery pack that the patient can wear. The inflow cannula is attached to the left ventricular apex after removing a circular core of heart muscle. Blood is passively drained from the left ventricle and ejected into an outflow cannula that is attached to the ascending aorta.

Porcine valves, each 25 mm, are attached to both the inflow and the outflow cannula. This allows for unidirectional flow. Blood is pumped through the device into a 20 mm woven Dacron graft.

The blood-pumping surfaces, composed of titanium microspheres sintered onto a titanium surface, are textured. The polyurethane diaphragm is also textured. This environment provides a nidus for growth of a thin layer of intima comprising both acellular and cellular components that form a biocompatible layer. The



Fig. 2 The HeartMate vented electric pump. (Reprinted with permission from Thoratec Corporation.)

exact mechanism by which this layer forms is unknown but has been studied.^[2] Thrombus formation is unusual, so patients only need to be maintained on low-dose antiplatelet therapy (aspirin 81 mg).

The polyurethane diaphragm separates the titanium housing from the blood. Between this membrane and the housing lies the rotor that moves the diaphragm and propels the blood forward. A long driveline enters the skin above the umbilicus and is tunneled percutaneously to attach to the left side of the device. The driveline attaches to an external power pack for venting.

The HeartMate VE is 11 cm in diameter and 4 cm thick, and occupies a volume of 460 cc. Two batteries and an external controller supply the external power pack, and their sum weight is approximately 1.5 kg. The batteries supply 4 to 7 hours of function, with an emergency 24-hour battery and an external power source available if battery failure should occur.^[3]

The HeartMate VE can eject blood at a maximum stroke volume of 83 ml. Volumes should be adjusted by changing the heart rate or preload from the patient's intravascular volume status, so that stroke volumes stay between 70–80 ml. Anything less risks thromboembolism due to an overall decreased flow through the pump. Heart rates can be adjusted to between 50 and 120 beats per minute (bpm), which makes cardiac output 4 to 10 L/min. Decreased flows also reduce the patient's exercise tolerance. Hence, flows can be optimized via the patient's intravascular volume and the pump's stroke volume and rate. The HeartMate VE ejects independently of the patient's rhythm, but that has not been problematic clinically.

The HeartMate VE can be implanted within the rectus sheath, preperitoneally, or within the abdomen. Factors that contribute to the choice of site are potential length of LVAD implantation, previous abdominal surgery, abdominal location of cardiodefibrillators, and the surgeon's preference.

HeartMate Implantable Pneumatic (IP)

The HeartMate IP was the original version, followed by the HeartMate VE. The IP unit is very similar to the VE, containing a textured titanium housing with a polyurethane diaphragm. An air chamber lies between the rigid titanium housing and the flexible membrane. Air is pumped through the external driveline and displaces blood in the diaphragm, thereby ejecting blood into the outflow conduit.

The HeartMate IP is powered by one of three sources: an external power cart, a portable driver, or a small hand pump for emergencies. The cart weighs 73 pounds and the portable driver weighs 20 pounds, thereby making home use difficult.

Pumping and flow characteristics of the VE and IP are similar. Both devices have an 83-ml maximum stroke volume and pump rates between 50 and 120 bpm, and both function independently of intrinsic heart function and rate. Neither device can be placed into a patient with a body surface area less than 1.5 square meters, because the large size of the devices impedes the function of intra-abdominal organs.

Novacor LVAS

The development of the Novacor device (Fig. 3) was initiated in 1969, and in 1984 it was first successfully used as a bridge to transplant.^[4] This device is a fully implantable, electrically powered system attached to a percutaneous driveline that exits the anterior abdominal wall in the right lower quadrant.

The Novacor device consists of a seamless, smooth polyurethane sac that is compressed between two solenoid pusher plates. The housing is composed of fiberglass and epoxy. Inflow and outflow conduits are composed of a 25-mm gelatin-sealed polyester graft (Sulzer Vascutek Ltd., Renfrewshire, Scotland, UK) attached to porcine valves that ensure unidirectional flow.

The Novacor device is implanted in the left posterior rectus abdominus sheath. Inflow and outflow conduits are tunneled through the diaphragm on the left and right side, respectively. The inflow cannula is attached to the left ventricular apex, whereas the outflow cannula is attached to the right side of the



Fig. 3 The Novacor LVAS. (Reprinted with permission from WorldHeart Corporation.)

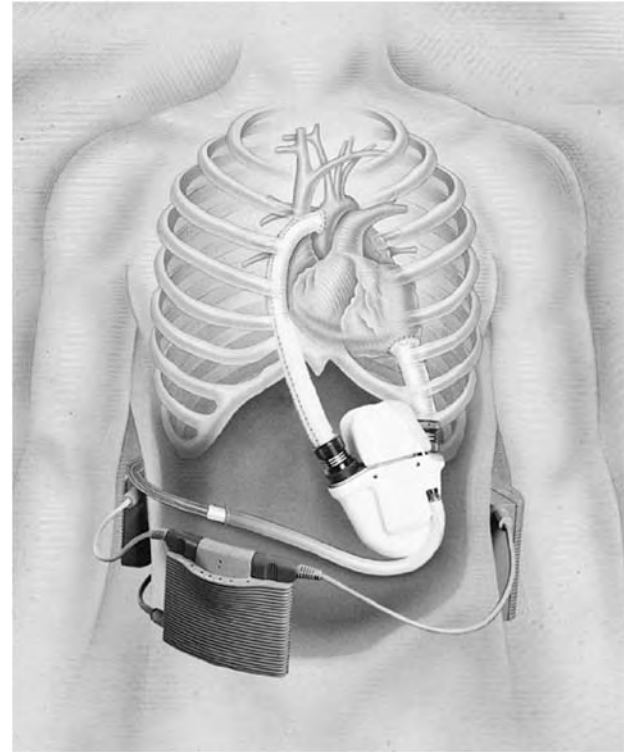


Fig. 4 Anatomical placement of Novacor LVAS. (Reprinted with permission from WorldHeart Corporation.)

ascending aorta. The external driveline is tunneled subcutaneously and exits through the abdominal wall to the right of the umbilicus.

The system controller is worn on the body and is connected to the external driveline that acts as an electrical lead and a vent (Fig. 4). The control system provides electrical energy to the solenoid pusher plates, which are coupled to a spring mechanism. Overall, the energy conversion process is approximately 65% efficient.^[5]

Unlike the HeartMate devices, the Novacor can function both asynchronously or synchronously with the patient's heart rate, but it is usually programmed in the latter mode. The pump fills passively during ventricular systole and ejects during diastole. The device has a stroke volume of 65 ml and can generate outputs of 7 liters per minute.^[6]

COMPLICATIONS

Thromboembolism

Clinically significant neurological events, defined as any change in neurological status detected by physical exam, can often be attributed to thromboembolism in

LVAD patients. LVADs activate the coagulation cascade, resulting in clot formation on the devices. Clot on the devices, in combination with turbulent blood flow across the surfaces, leads to thromboembolic events such as stroke, kidney failure, and distal extremity ischemia. Even with anticoagulation therapy, these events occur in up to 30% of LVAD patients.^[7] Thromboembolic events can occur with both HeartMate and Novacor devices, but they are much more prevalent with the latter device.

The HeartMate systems have a textured interior surface on which an adherent pseudointima may form.^[8] This may become a layer between the device surface and blood flow. The exact nature and cell types involved in this layer remain controversial. However, it is likely that soon after implantation, platelet deposition on the blood-contacting surfaces of the LVAD leads to rapid deposition of a fibrin matrix, which acts as the initial scaffolding for the attachment of a variety of hematopoietic cells, including monocytes and pluripotent cells. This process then leads to myofibroblast attachment and proliferation—possibly as a result of chemotactic factors released by monocytes—which leads to the formation of a pseudointima of collagenous matrix. This can later lead to the attachment of host endothelial cells.^[9] In one series of 223 patients, a clinically significant thromboembolic rate of 2.7% was observed.^[10] Other design characteristics that minimize clot formation are a short, rigid inflow cannula and a cornerless pumping chamber. The diameter of the inflow and outflow grafts is 20 mm in the HeartMate device, compared to 25 mm in the Novacor.

The incidence of neurologic complications is more frequent with the Novacor device. Thus, all patients with this device need systemic anticoagulation with warfarin. Studies report up to 47% incidence of neurologic complications.^[11] In most patients with the Novacor device, a poorly adherent pannus was seen in the inflow conduit. The Novacor inflow conduit is longer than the HeartMate's and also more flexible, which allows for separation and embolization of this pannus material. Portner et al. demonstrated, in a nonrandomized, retrospective review, that the incidence of embolic strokes was significantly reduced from 21% to 12% by changing conduit materials from a woven, crimped polyester graft to a gelatin-sealed, knitted graft with wall reinforcement. These authors also found a more stable neointima on the latter grafts, and computer modeling of both grafts showed lower mural shear rates, with less variation, in the gelatin noncrimped grafts.^[12] Further reduction in length of the inflow and change to polytetrafluoroethylene (PTFE) with more rigidity will likely reduce the evidence of embolic strokes.

Infection

Currently, the risk of infection from the LVADs remains the single most important factor precluding long-term support. The presence of an external power source requiring the percutaneous driveline, ill patients with multiple vascular access lines, and the poor nutritional status of patients may all contribute to infection. Infection, in turn, directly increases the risk of device failure, thromboembolic events, and bleeding. The Cleveland Clinic reviewed more than 200 patients receiving LVADs between 1992 and 2000, and 49% of them had a bloodstream infection. The most common pathogens were coagulase-negative *Staphylococcus* (23%), *Staphylococcus aureus* (13.5%), *Candida* (13.5%), and *Pseudomonas*. Bloodstream infection was associated with a higher mortality. Fungal infections had a hazard ratio of 10.9 for death.^[13]

In a nonrandomized prospective trial comparing the Novacor and HeartMate devices, there was a significantly higher incidence of infections in the HeartMate group (55% vs. 20%). Although the exact reason for this is unknown, it is suggested that the rotary movement of the HeartMate pump increases the reaction between the pump and surrounding tissues, thereby increasing infection.^[14] Studies have also shown a state of aberrant T cell activation, progressive defects in cell-mediated immunity, and increased risk of opportunistic infections in patients supported with LVADs.^[15]

Bleeding

Bleeding remains a significant problem after LVAD insertion. Most of these patients are quite ill before surgery, with significant hepatic dysfunction and low levels of clotting factors synthesized in the liver. The use of vitamin K parenterally prior to surgery and aprotinin—a serine protease inhibitor—during surgery is recommended to minimize bleeding and the need for blood products. Patients with HeartMate LVADs may have significant activation of coagulation with secondary fibrinolysis, despite both clinical stability and “normal” screening values of routine hemostatic parameters such as platelet count, prothrombin time, and activated partial thromboplastin time.^[16] Similar studies of the Novacor device are not available. Since most studies advocate aggressive early anticoagulation within a few hours of implanting the Novacor device, the amount of blood loss was higher when compared with the HeartMate. El-Banayosy demonstrated that although blood loss was greater in the Novacor group, there was no difference between groups in the frequency of reoperation to control bleeding.^[14]



Device Failure

Multiple device failures plagued the HeartMate devices prior to initial FDA approval as a bridge to transplantation. McCarthy noted that 12 patients out of 100 in the Cleveland Clinic series had device failure, which included driveline fractures, leaking from inflow conduits, and failure of the electrical system. Half of the patients died.^[17] Since then, mechanical failure of the device has been extremely uncommon in support durations of less than one year. Data from the manufacturer indicate an average exchange rate of 2.4% for device failure from 1993 to 1999.^[3] These numbers may not truly reflect deaths caused by device failures.

One of the modes of device dysfunction relates to failure of the HeartMate inflow valve. This is manifested as an increase in pump rates and calculated flows, with decreased exercise tolerance in the patient.^[18] In the recent REMATCH trial, the incidence of device failure was 17% during long-term support (more than one year).

Mechanical failure of the Novacor device has been uncommon. The main mode of late failure appears to be wear in the bearing of the actuator. This seems to occur after three to five years and can be detected by analysis of the pump's electrical signals. This information can then be used to electively exchange the pump.

THE ABIOCOR IMPLANTABLE ARTIFICIAL HEART

The modern left ventricular assist devices have been improved and developed in extraordinary ways over the past 20 years to provide excellent hemodynamic support for patients with left ventricular failure. Despite the many potential complications in the perioperative and postoperative periods, including bleeding, infection, and thromboembolic events, the major limitation of modern-day left ventricular assist devices is that these pumps need an external power source. Obviously, the ultimate goal of any investigative and research effort in the field of mechanical circulatory support is the development of a fully implantable replacement heart that allows the patient to be fully ambulatory without tethering to external power sources. The AbioCor™ implantable replacement heart is an advanced medical system that has been designed to fully sustain the body's circulation. These devices have been developed over the past 30 years in collaboration with and support from the NHLBI. Currently, the AbioCor is being tested in patients with irreversible, end-stage, biventricular heart failure with no other therapeutic options and with a life expectancy of less than 30 days.

The Federal Drug Administration is testing the AbioCor and has authorized up to fifteen implants in groups of five, with review after each group. The primary end point is mortality from any cause, evaluated 60 days following implantation. The secondary end points include quality-of-life measurements and adverse events determinations involving complications. The device requires the utilization of three decision support tools: 1) the Abio-score, a prognostic tool comprising laboratory and clinical data significant to 30-day mortality in advanced heart failure patients; 2) the AMI-shock index, a second prognostic tool for patients with acute myocardial infarction and cardiogenic shock; and 3) the Abio-fit screening tool, used prior to surgery to test the anatomical compatibility between potential subjects and the AbioCor thoracic unit. The Abio-fit accepts a range of digitalized thoracic imagery, including computed tomography, magnetic resonance imaging, or conventional X-rays. Prior to surgery, it creates a three-dimensional image of the AbioCor superimposed on the imagery of the patient's thoracic structures to help surgeons gauge the feasibility of implanting the replacement heart. This ensures that the AbioCor fits within the borders of the left chest cavity, with good alignment of the inflows to the two atria and of the outflow track to the great vessels, while not interfering with the pulmonary veins or the left bronchus.

The AbioCor consists of four implanted subsystem components. The implanted controller drives the thoracic unit and is powered by the implanted battery or the internal transcutaneous energy transmission secondary coil (TET), and it interfaces with the AbioCor console via radio frequency link (Fig. 5). The portable console provides primary power to the system transcutaneously, via the external TET primary coil that is placed on the skin area overlying the internal TET coil. It also provides two-way radio frequency communication with the implanted controller through an RF communication module, for monitoring alarm

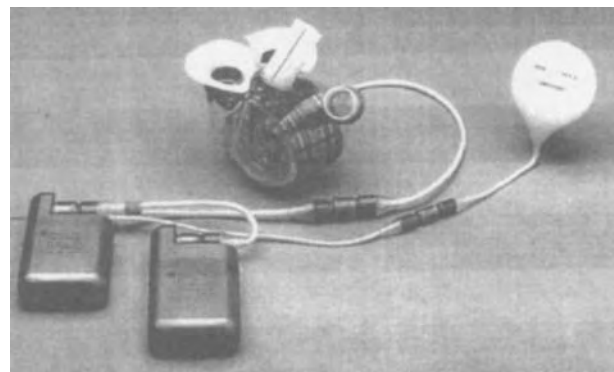


Fig. 5 The AbioCor totally implantable artificial heart. (Reprinted with permission from Abiomed Corporation.)

and system control if needed. The implanted TET converts high-frequency power received from the external coil to a DC output suitable for powering the internal components.

The thoracic unit has three primary functional subunits: 1) a balance chamber that manages the right and left blood flow by maintaining physiologic left-atrial pressures; 2) the self-contained hydraulic unit, separated by pairs of flexing membranes situated between the two blood pumps, which serves as an undulating pseudo-septum actuating the left and right blood pump alternately; and 3) the blood-contacting components, including tri-leaflet valves and blood pumps, that are fabricated from a nontoxic, non-mutagenic, polyurethane called Angioflex^R. The thoracic unit, weighing approximately 1000 grams and the size of a grapefruit, includes two artificial ventricles with corresponding valves and a motor-driven hydraulic pumping system. This pump performs the back-and-forth movement of an internal ball within a sphere, supplying blood to all the vital organs. Between the two chambers, the small ball hits the soft walls with jets of fluid compressing the ventricles in a quick pulsation, forcing their contents into the arteries. The AbioCor operates on internal and external lithium batteries. The internal batteries are considered to be for emergency purposes lasting one-half hour only, but they allow the patient to move around without the external power source and the external battery pack. The external battery packs can be used for many hours, depending on the number of batteries that are carried.

To implant the device, the surgeon removes most of the patient's diseased heart, leaving only the left and right atrial cuffs. The device is then attached to the atria and to the aorta. The AbioCor can pump up to 8 liters per minute from physiologic left filling pressure into a mean arterial pressure range of from 60–160 mm of mercury, and from physiologic right filling pressure into a mean pulmonary pressure range of from 15–120 mm of mercury. This allows it to support patients with chronic obstructive pulmonary disease and pulmonary hypertension. The AbioCor has durability and reliability consistent with one year of trouble-free operation, with potential for longer operational life. Durability and reliability tests of valves and blood pumps have exceeded 20,200 and 100,000,000 cycles, respectively, under physiologic load conditions.

Since July of 2001, the AbioCor has been implanted in six patients. The first patient had a remarkable physiologic recovery and survived 151 days, but died of a stroke-related complication. Other patients have now survived with the device in place for longer than a year without significant complication. They are fully rehabilitated and functional, with an excellent quality of life. Significant obstacles to widespread utilization

of this technology continue to be biomechanical and materials problems; incompatibility with foreign blood elements; and the need for an external power source and the size of the pump, which prohibit its use in women and adolescents weighing less than 90 kg.

FUTURE DIRECTIONS IN THE USE OF IMPLANTABLE PULSATILE DEVICES

As the numbers of patients with advanced heart failure increase, the use of LVADs as long-term therapy will become more important. The landmark study—Randomized Evaluation of Mechanical Assistance for the Treatment of Congestive Heart Failure (REMATCH)—was conducted from 1998 to 2001 in 20 centers in the United States. Eligible patients were adults with chronic endstage heart failure and with contraindications to heart failure. One type of vented electric device—the HeartMate LVAD—was selected for the study. This device contains textured internal surfaces, and patients receiving it do not require anticoagulation. Study patients were randomized in a 1:1 ratio to the vented electric system or to optimal medical therapy. The primary endpoint was death from any cause. The trial was designed to enroll 140 patients and to continue until 92 deaths occurred. Secondary endpoints included the incidence of serious adverse events, the number of days of hospitalization, quality-of-life indices, symptoms of depression, and functional status. A total of 129 patients were enrolled, 68 receiving left ventricular assist devices and 61 assigned to undergo optimal medical management. There was a reduction of 48% in the risk of death from any cause in the group that received the left ventricular devices, compared to the medical therapy group ($P=0.001$). The Kaplan–Meier estimates of survival at one year were 52% in the device group and 25% in the medical therapy group ($P=0.002$). At two years, these estimates had dropped to 23% and 8%, respectively ($P=0.09$). Terminal heart failure was the cause of most of the deaths in the medical therapy group, whereas most deaths in the device group were caused by sepsis (41%) and failure of the device (17%). Clearly, in patients ineligible for transplantation, the left ventricular devices offer a superior survival outlook compared to optimal medical therapy.

Unfortunately, despite substantial survival benefit, there was considerable mortality and morbidity associated with the device. Infection was a major limiting factor. Predisposing factors for infection included the status of the percutaneous driveline, malnutrition, and severe preoperative debilitation due to heart failure. Mechanical failure of the device, inflow valve insufficiency, and neurological events were the other major limiting factors. It is apparent that these problems will



have to be addressed before further generalized use of these devices in patients with CHF.

With the AbioCor total artificial heart system, novel power sources, transcutaneous delivery of power with no percutaneous lines, acceptable improvements in device reliability, and a deeper understanding of patient–pump interactions have been advanced. The newer pumps also are expected to be smaller in order to accommodate a wider range of patients, including women and children.

ARTICLES OF FURTHER INTEREST

Blood Vessel Mechanics; Mass Transfer in Tissue Function: Roles; Vascular Fluid Dynamics

REFERENCES

- Rose, E.A.; for the REMATCH Study Group. Long term use of a left ventricular assist device for endstage heart failure. *NEJM* **2001**, *345*, 1435–1443.
- Spanier, T.B.; Chen, J.M.; Oz, M.C.; Stern, D.M.; Rose, E.A.; Schmidt, A.M. Time-dependent cellular population of textured-surface left ventricular assist devices contributes to the development of a biphasic systemic procoagulant response. *J. Thorac. Cardiovasc. Surg.* **1999**, *118*, 404–413.
- Long, J.W. Advanced mechanical circulatory support with the HeartMate ventricular assist device in the year 2000. *Ann. Thorac. Surg.* **2001**, *71*, S176–S182.
- Portner, P.M.; Oyer, P.E.; McGregor, C.G.A. First human use of an electrically powered implantable ventricular assist system. *Artif. Organs* **1985**, *9*, 36.
- Robbins, R.C.; Oyer, P.E. Bridge to transplant with the Novacor left ventricular assist system. *Ann. Thorac. Surg.* **1999**, *68*, 695–697.
- Kasirajan, V.; McCarthy, P.M.; Hoercher, K.J.; Starling, R.C.; Young, J.B.; Banbury, M.K.; Smedira, N.G. Clinical experience with long term use of implantable left ventricular assist devices: Indications, implantation, and outcomes. *Semin. Thorac. Cardiovasc. Surg.* **2000**, *12*, 229–237.
- Wagner, W.R.; Johnson, P.C.; Kormos, R.L.; Griffith, P. Evaluation of bioprosthetic valve-associated thrombus in ventricular assist device patients. *Circulation* **1993**, *88*, 2023–2029.
- Rose, E.A.; Levin, H.; Oz, M.C.; Frazier, O.H.; Mcmanus, Q.; Burton, N.A.; Lefrak, E.A. Artificial circulatory support with textured interior surfaces: A counterintuitive approach to minimizing thromboembolism. *Circulation* **1994**, *90*, II-87–II-91.
- Scott-Burden, T.; Frazier, O.H. Cellular linings of ventricular assist devices. *Ann. Thorac. Surg.* **1995**, *60*, 1561–1562.
- Slater, J.P.; Rose, E.A.; Levin, H.R.; Frazier, O.H.; Roberts, J.K.; Weinberg, A.D.; Oz, M.C. Low thromboembolic risk without anticoagulation using advanced-design left ventricular assist devices. *Ann. Thorac. Surg.* **1996**, *62*, 1321–1327.
- Thomas, C.E.; Jichici, D.; Petrucci, R.; Urrutia, V.C.; Schwartzman, R.J. Neurologic complications of the Novacor left ventricular assist device. *Ann. Thorac. Surg.* **2001**, *72*, 1311–1315.
- Portner, P.M.; Jansen, P.G.M.; Oyer, P.E.; Wheeldon, D.R.; Ramasamy, N. Improved outcomes with an implantable left ventricular assist system: A multicenter study. *Ann. Thorac. Surg.* **2001**, *71*, 205–209.
- Gordon, S.M.; Schmitt, S.K.; Jacobs, M.; Smedira, N.M.; Goormastic, M.; Banbury, M.K.; Yeager, M.; Serkey, J.; Hoercher, K.; McCarthy, P.M. Nosocomial infections in patients with left ventricular assist devices. *Ann. Thorac. Surg.* **2001**, *72*, 725–730.
- El-banayasy, A.; Arusoglu, L.; Kizner, L.; Tenderich, G.; Minami, K.; Inoue, K.; Korfer, K. Novacor left ventricular assist system versus heartmate vented electric left ventricular assist system as a long term mechanical circulatory support device in bridging patients: A prospective study. *J. Thorac. Cardiovasc. Surg.* **2000**, *119*.
- Ankersmit, H.A.; Tugulea, S.; Spanier, T.; Weinberg, A.D. Activation-induced T cell death and immune dysfunction after implantation of left ventricular assist device. *Lancet* **1999**, *354*, 550–555.
- Spanier, T.; Oz, M.C.; Levin, H.; Weinberg, A.; Stamatis, K.; Stern, D.; Rose, E.A.; Schmidt, M. Activation of coagulation and fibrinolytic pathways in patients with left ventricular assist devices. *J. Thorac. Cardiovasc. Surg.* **1996**, *112*, 1090–1097.
- McCarthy, P.M.; Smedira, N.O.; Vargo, R.L.; Goormastic, M.; Hobbs, R.E.; Starling, R.C.; Young, J.B. One hundred patients with the HeartMate left ventricular assist device: Evolving concepts and technology. *J. Thorac. Cardiovasc. Surg.* **1998**, *115*.
- Poirier, V.L. Inflow valve incompetence. *J. Congest. Heart Fail. Circ. Support* **2001**, *1* (2), 23–25.

Mechano-Regulation of Fibroblast Function

James H.-C. Wang
Bhavani P. Thampatty

MechanoBiology Laboratory, Departments of Orthopaedic Surgery, Bioengineering and Mechanical Engineering, University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.



INTRODUCTION

The human body is subjected to various types of mechanical forces. For example, a state of tension exists naturally in the skin, and mechanical forces generated during motion exert on the tendons and ligaments of joints. The cardiovascular and respiratory systems are organ systems that are constantly subjected to mechanical forces; varying levels of tensile, compressive, and shearing stresses because of blood flow in the cardiovascular system and air flow in the respiratory system constantly act on blood vessels and the lining of the lungs. As a result of their dynamic environment, cells in the skin, tendons, ligaments, blood vessels, and lung respond to mechanical forces by altering gene expression and the levels of protein secreted, which lead to changes in the extracellular matrix (ECM) and eventually in the tissue structure and function. These cells are referred to as mechano-responsive cells.

A major type of mechano-responsive cells in the human body is the fibroblast, which is a mesenchymal cell that plays an essential role in maintaining, repairing, and remodeling cardiac and lung tissues, as well as other connective tissues including the skin, tendons, and ligaments. This entry will review the mechano-regulation of fibroblast function, with a focus on the effects of mechanical forces on ECM gene expression and protein synthesis. In addition, the interactions of mechanical loads with growth factors and cytokines in terms of their effects on the ECM and tissue engineering constructs will be presented. This will be followed by a discussion of cellular mechano-transduction mechanisms by which mechanical forces are transduced into a cascade of cellular and molecular events, and the discussion will conclude with the authors' views on future research directions in the area of fibroblast mechano-biology.

MECHANO-REGULATION OF GENE EXPRESSION IN FIBROBLASTS

Fibroblasts in cardiac, lung, and connective tissues are responsible for the production of ECM proteins, including different types of collagens, elastin, fibronectin,

and proteoglycans. All these matrix components provide for the structural and functional integrity of the tissue. In addition, fibroblasts secrete various soluble factors, including hormones [e.g., insulin-like growth factors I and II (IGF-I and IGF-II)], growth factors [e.g., transforming growth factor- β (TGF- β), platelet-derived growth factor (PDGF), fibroblast growth factor (FGF), endothelial growth factor (EGF), and connective tissue growth factor (CTGF)], and cytokines [e.g., interleukin-1 (IL-1) and tumor necrosis factor- α (TNF- α)].^[17,18] These factors modulate many cellular functions such as DNA synthesis, protein secretion, cell migration, and differentiation.^[19] They also play an essential role in tissue wound healing.^[20]

Fibroblasts are subjected to various forms of mechanical loading that regulate their functions. To study the mechano-regulation of fibroblast function in a controlled environment, several cell loading systems have been developed.^[21,22] These devices make use of flexible substrates with smooth culture surfaces, such as silicone elastomer coated with ECM proteins to which the cells are attached. Cyclic biaxial strains are applied by stretching the substrate in two orthogonal directions. When equibiaxial stretches are applied, the substrate strains transmitted to the cell are the same in all directions; however, when non-equibiaxial stretches are applied, the substrate strains change with the cell's orientation with respect to the stretching direction. As a result, the strains that cells experience depend on their orientations (Fig. 1A). On the other hand, cyclic uniaxial strains can also be applied by stretching the substrate in one direction, and as a result, the substrate is lengthened in the stretching direction whereas it is shortened in its perpendicular direction (Fig. 1B). Again, the strains that cells experience are cell orientation dependent. In both cases of stretching, the strain level ranges from 1% to 30%, and the strain rates commonly vary from 0.1 to 10 Hz.^[23]

Besides stretching fibroblasts on the substrates with smooth surfaces, a microgrooved silicone substrate coated with ProNectin-F, an engineered fibronectin-like protein, has also been used.^[24,25] When tendon fibroblasts in the microgrooved substrate are subjected to uniaxial stretching, they assume an elongated shape and are oriented in the direction of stretching,

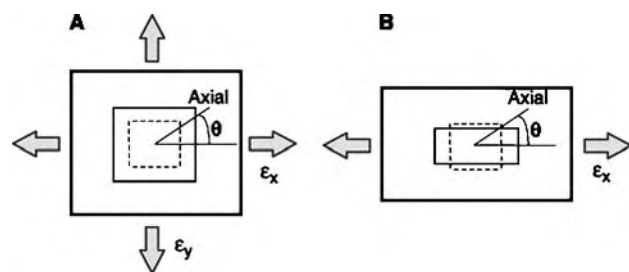


Fig. 1 The application of biaxial stretching and uniaxial stretching to fibroblasts. (A) Biaxial stretching. There are two cases of biaxial stretching: One is the equibiaxial stretching conditions, where $\epsilon_x = \epsilon_y$; the other is the non-equibiaxial stretching, where $\epsilon_x \neq \epsilon_y$. Under the equibiaxial stretching, the surface strains do not change with the orientation, θ . Therefore, the cells on the surface are subjected to the same strains, regardless of their orientation with respect to stretching direction. However, under non-equibiaxial stretching, surface strains change with the orientation, θ . In other words, the cells on the surface are subjected to different levels of surface strains, depending on their orientation. (B) Uniaxial stretching. Under uniaxial stretching, surface strains change with the orientation, θ . This means that cells with different orientation are subjected to different levels of surface strains.

mimicking the alignment and loading conditions of tendon fibroblasts in vivo (Fig. 2).

Fibroblast-populated collagen gels (FPCGs) are also suitable in vitro cell culture models to study the response of cells to mechanical forces.^[26,27] Three-dimensional collagen gels provide improved maintenance of cell morphology and phenotype, which create an environment for studying fibroblast response to mechanical loading that more closely mimics the in vivo environment. The FPCG is maintained in a culture dish with growth medium, and fibroblasts in it generate contraction (Fig.3). As a result, the FPCG is under tension if its bottom is attached to the dish, or if its two ends are held in place. However, when the

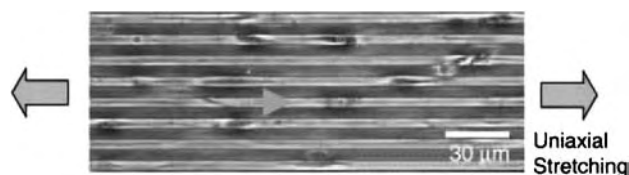


Fig. 2 The application of cyclic uniaxial stretching to tendon fibroblasts in a microgrooved substrate. It is seen that tendon fibroblasts in the microgrooves are elongated and aligned in the direction of the microgrooves, which is also the stretching direction. The advantage of this type of stretching is that one can control the alignment and mechanical stretching conditions of tendon fibroblasts to mimic their in vivo organization and loading conditions.

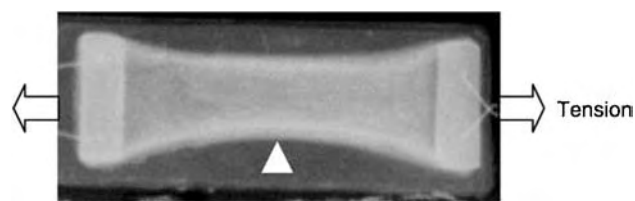


Fig. 3 The tension in the FPCG. Fibroblasts in the gel generate contraction, so that the gel changes from rectangular to a parabolic shape (arrow). Note that tension is maintained by holding two ends of the gel in place. Once the ends are released, the tension disappears, such as in the case of a floating gel in medium. When this occurs, the cells in the gel are in the stress-free state.

FPCG is detached, it floats in the medium, and all tension is released. Besides internal tension generated in FPCGs, external loading can also be applied to the FPCGs.^[28-30]

In response to mechanical loads, fibroblasts alter DNA synthesis, the expression of ECM genes and proteins, and activation of signaling molecules (Table 1). For example, cyclic and static uniaxial stretching of cardiac fibroblasts increased both collagen III mRNA expression and the ratio of collagen III to collagen I.^[31] Cyclic mechanical stretching also significantly elevated the procollagen I mRNA levels of these cells^[32] as well as collagen I mRNA expression and protein secretion.^[2] These results are consistent with the fact that cardiac hypertrophy through mechanical loading is accompanied by an increase in the deposition of ECM proteins.^[33]

Also, mechanical stretching in cardiac fibroblasts was found to regulate collagen type III and fibronectin mRNA levels and TGF- β 1 activity in a stretching magnitude-dependent manner.^[1] A 10% stretching magnitude significantly increased mRNA expression levels of collagen type III and fibronectin while collagen I mRNA only increased slightly. However, a 20% stretching magnitude decreased collagen type III and fibronectin mRNA levels while collagen I did not change significantly. Similarly, equibiaxial strain also caused up- or down-regulation of specific ECM genes, where the genes affected depended on the magnitude of stretching applied. For instance, 3% equibiaxial strain increased both collagen III and fibronectin mRNA levels, whereas a 6% strain decreased collagen III mRNA levels but did not create a significant change in fibronectin levels compared to the unstretched control. In addition, both low and high levels of compressive equibiaxial strains decreased fibronectin mRNA levels,^[1] suggesting that cardiac fibroblasts show a differential response to tensile and compressive strains in terms of their ECM gene expression and protein synthesis.

Table 1 Mechano-regulation of fibroblast function

Cells	Response	Type of load	Substrate	Significance	References
Cardiac fibroblasts	Increase in collagen type III and fibronectin mRNA levels and TGF- β 1 activity	Cyclic uniaxial and biaxial stretch, 10% and 3%, respectively	Collagen type I coated	Differential response to different stretching magnitudes and types of loading (stretching vs. compression)	[1]
	Decrease in collagen III and fibronectin mRNA	Cyclic uniaxial stretch, 20%			
	Decrease in collagen III mRNA, no change in fibronectin mRNA	Cyclic biaxial stretch, 6%			
	Decrease in fibronectin mRNA, no significant change in collagen III mRNA	Compressive equibiaxial strain, 3–6%			
Cardiac fibroblasts	Increase in collagen III mRNA	Cyclic uniaxial stretch, 5%	Laminin coated	Substrate specific response	[2]
	Mitogen-activated protein kinase (MAPK) activation, increase in collagen-I mRNA	Cyclic uniaxial stretch, 5%	Randomly organized collagen and uncharged silastic membrane		
Cardiac fibroblasts	Increase in procollagen mRNA	Cyclic equibiaxial stretch, 20%	Elastin coated	Synergistic stimulation by load and growth factors	[3]
Lung fibroblasts	Increase in procollagen mRNA and protein	Cyclic biaxial stretch, 20%	Elastin and laminin coated	Matrix protein specific response	[4]
	Decrease in procollagen mRNA	Cyclic biaxial stretch, 20%	Fibronectin coated		
Dermal fibroblasts	Increase in α (I) collagen mRNA and matrix metalloproteinase-1 (MMP-1) mRNA	Cyclic biaxial and static stretch, 20%	Collagen type I coated vs. plastic dish	Substrate and strain specific gene regulation	[5]
Dermal fibroblasts	Increase in tenascin-C mRNA and protein	Cyclic equibiaxial stretch, 10%	Fibronectin coated	Additive effect with growth factors; requirement of Rho-ROCK signaling for strain-induced effects	[6]
Dermal fibroblasts	Cell alignment, orientation toward stretching direction and increase in β 1-integrin and collagen I mRNA	Cyclic uniaxial stretch, 10%	Smooth surface vs. microgrooved surface	Surface topography based response	[7]
Tendon fibroblasts	DNA synthesis stimulation	Cyclic biaxial stretch, 5%	Collagen type I coated	Synergistic action of growth factors and mechanical load	[8,9]
Tendon fibroblasts	Stimulation of collagen type I and TGF- β 1 gene expression	Cyclic uniaxial stretch, 8%	ProNectin-F coated and microgrooved	TGF- β 1 mediated response to load	[10]

(Continued)



Table 1 Mechano-regulation of fibroblast function (*Continued*)

Cells	Response	Type of load	Substrate	Significance	References
Tendon fibroblasts	Increase in MMP-1 and MMP-3 mRNA	Cyclic biaxial stretch, 5%	Collagen type I coated	Matrix degradation initiation by synergistic effect of cytokine and mechanical load	[11]
Ligament [anterior cruciate ligament (ACL)] fibroblasts	Increase in collagen I and III mRNA, and TGF- β 1 protein	Cyclic uniaxial stretch, 10%	Collagen type I coated	TGF- β 1 mediated response	[12]
Ligament [ACL and medial collateral ligament (MCL)] fibroblasts	Increase in collagen I mRNA with no change in collagen I (ACL), increase in collagen III with no change in collagen I mRNA	Cyclic equibiaxial stretch, 5%	Collagen type I coated	Differential response to different cell source	[13]
Ligament (ACL) fibroblasts	Increase in collagen synthesis, cell alignment in the direction of stretch	Cyclic uniaxial stretch, 5%	Aligned nanofiber sheets	Surface material and topography specific response	[14]
Ligament (periodontal) fibroblasts	Increase in collagen I and fibronectin synthesis, decrease in tropoelastin synthesis	Cyclic equibiaxial stretch, 5%	Collagen type I coated	Differential ECM protein synthesis in response to specific magnitudes of stress	[15]
Ligament (periodontal) fibroblasts	No change in collagen I synthesis, increase in fibronectin synthesis, decrease in tropoelastin synthesis	Cyclic equibiaxial stretch, 10%			
Ligament (periodontal) fibroblasts	Decrease in collagen I and fibronectin protein, increase in MMP-2 and no change in tissue inhibitor of metalloproteinase-2 (TIMP-2) mRNA	Cyclic compressive force, 10%	Collagen type I coated	Differential ECM synthesis and degradation in response to different types of mechanical load	[16]
Ligament (periodontal) fibroblasts	Increase in collagen I protein, no significant increase in fibronectin synthesis, increase in MMP-2 and TIMP-2 mRNA	Cyclic equibiaxial stretch, 10%			

Furthermore, synergistic stimulation of procollagen gene expression by mechanical loading and growth factors was observed. Tensile mechanical loading of cardiac fibroblasts in combination with IGF-I was shown to stimulate procollagen mRNA levels three fold compared with that of IGF-I treatment alone. Also, a combination of tensile mechanical loading and TGF- β increased procollagen α_1 (I) (*COL1A1*) mRNA levels 4.3-fold compared with that of TGF- β treatment alone.^[3] The strain-induced *COL1A1* promoter activation in cardiac fibroblasts is TGF- β dependent, suggesting a novel TGF- β response element in the *COL1A1* gene.^[34]

The response to mechanical load also depends on the composition of the ECM in cardiac fibroblasts. In vitro studies have shown that different ECM components (underlying ECM substrates) modify cardiac fibroblast activity differently during mechanical stretch. For example, stretched cardiac fibroblasts cultured on silastic membrane coated with different ECM substrates (i.e., aligned collagen, randomly organized collagen, fibronectin, laminin, and charged membrane) responded differently in terms of cell morphology, proliferation, collagen production, and cell signaling.^[17] In cells stretched on randomly organized collagen and uncoated charged membrane, mitogen-activated protein (MAP) kinase was most rapidly activated, and collagen I mRNA expression was most abundant. Decrease in cell proliferation was observed in stretched cells regardless of the nature of the ECM substrate; however, the decrease was most marked in cells stretched on randomly organized collagen. Finally, cells stretched on all ECM substrates increased their surface area, but the most significant increase was noted in cells adherent to aligned collagen, randomly organized collagen, and uncoated charged membrane compared to the membrane coated with fibronectin or laminin.

Similar to cardiac fibroblasts, lung fibroblasts are also mechano-responsive cells that alter procollagen gene expression in response to changes in mechanical tension.^[4] The mechano-response of lung fibroblasts appears to depend on ECM proteins because cells cultured on elastin or laminin, but not on fibronectin, increased procollagen gene expression.^[4] Furthermore, mechanical stress in FPCGs regulates gene expression of collagen α_2 (I) and collagen α_1 (III).^[35] Also, mechanical stress in conjunction with TGF- β induces a high level of CTGF, which is known to influence the synthesis of ECM proteins.^[35] Hence, it is likely that this growth factor is responsible for the increased synthesis of collagen and other ECM proteins in stressed fibroblasts. Besides CTGF, PDGF released by mechanically loaded lung fibroblasts also stimulates procollagen production.^[36]

Mechanical forces as regulators of ECM synthesis have been the subject of various studies in dermal

fibroblast monolayer as well as in FPCGs. When subjected to biaxial stretching in the presence of serum or TGF- β 1, these cells increased procollagen α_1 (I) mRNA expression.^[37] Also, both static and cyclic stretching of dermal fibroblasts cultured on silicone substrate increased α_1 (I) collagen mRNA expression^[5] and tenascin-C mRNA and protein levels.^[6] In addition, cyclic stretching of cells on silicone membranes at low frequencies increased the MMP-1 mRNA level whereas expression remained low in cells cultured on plastic dishes or stressed collagen gels.^[5] The results suggest that the type of stress (cyclic vs. static) as well as the substrate properties can influence the gene expression of dermal fibroblasts.

Furthermore, mechano-responsive genes encoding ECM proteins [collagen α_1 (I) and tenascin-C], growth factors, protease inhibitors, components of focal adhesions, and the cytoskeleton in dermal fibroblasts in mechanically stressed collagen gels have been identified by microarray technology and confirmed by Northern blot hybridization.^[5] This study demonstrated that cells adapt to mechanical stress by modulating the transcription of various genes and reorganizing their intracellular structures. The application of tension leads to an increase in the transcription of genes that are involved in cell proliferation.

The combined effect of mechanical stretching and soluble factors on the ECM synthesis has also been studied. Cyclic tensile strain and growth factors, such as TGF- β 1 and PDGF, increased tenascin-C mRNA expression in an additive manner,^[6] and a mechano-responsive regulatory sequence in the promoter of tenascin-C gene has been identified.^[38] This mechano-responsive sequence enables a mode for tightly regulating the cell's response to mechanical forces.

Interestingly, neonatal and adult dermal fibroblasts displayed a differential response to varying strain patterns in terms of cell proliferation and collagen synthesis.^[29,30] Neonatal cells were significantly stimulated in response to uniaxial strain but were unaffected by non-uniform and biaxial strains. Similarly, uniaxial strain stimulated adult fibroblast proliferation, but the other two strain patterns inhibited cell proliferation. Collagen synthesis in neonatal cells also increased slightly by uniaxial strain while synthesis in adult cells was significantly inhibited.

Additionally, in dermal fibroblasts, the topography of a substrate surface has been shown to influence cell shape, spreading, and orientation under mechanical stretching.^[29] Cells stretched on microgrooved surfaces were better aligned along the groove direction while cells on smooth surfaces were oriented randomly.

Fibroblasts are the dominant cell type in connective tissues, where they respond to mechanical forces by changing their expression of several ECM proteins including collagen type I.^[39] Mechanical stretching also



modulated the secretion of growth factors such as TGF- β 1, PDGF, and bFGF in human tendon fibroblasts, which may stimulate cell proliferation, differentiation, and matrix formation.^[40] In the presence of PDGF and IGF-I, mechanical loading stimulated DNA synthesis in tendon fibroblasts.^[8,9] Moreover, cyclic mechanical stretching significantly stimulated collagen type I and TGF- β 1 gene expression and protein secretion in the absence of serum. In addition, anti-TGF- β antibody eliminated the elevated collagen type I production, which suggests that TGF- β 1 mediates stretch-induced collagen type I production in tendon fibroblasts.^[10] Furthermore, using micro-grooved deformable substrates, the fibroblast response to mechanical stretching was shown to depend on cell orientation with respect to stretching direction because they expressed higher levels of smooth muscle actin (α -SMA) when they were stretched parallel to their long axes compared to stretching perpendicular to the long-axis direction.^[41] In addition, mechanical stretching of tendon cells in the presence of IL-1 β induced the gene expression of matrix-degrading enzymes such as MMP-1 and MMP-3, suggesting that when an inflammatory mediator such as IL-1 β is present, mechanical loading has catabolic effects on cells, which may eventually lead to tendon matrix degeneration.^[11] A recent study further showed that 4% cyclic uniaxial stretching of human tendon fibroblasts decreased the cyclooxygenase-2 (COX-2) and MMP-1 gene expression and prostaglandin E₂ (PGE₂) production that were stimulated by IL-1 β (10 pM); however, a 8% stretching further increased these gene expression and PGE₂ production in addition to the effects of IL-1 β stimulation. Therefore, the study suggests that repetitive, small-magnitude stretching is anti-inflammatory, whereas large-magnitude stretching is proinflammatory.^[42]

Similar to tendon fibroblasts, ligament fibroblasts also respond to mechanical loading by altering ECM gene expression. Mechanical stretching increased collagen types I and III mRNA expression in human ACL fibroblasts^[12] and increased the secretion of TGF- β 1 protein. This increase was ablated by anti-TGF- β 1 antibody, which suggests that the stretch-induced collagen expression is mediated via an autocrine mechanism of TGF- β 1.^[12] In addition, the ACL and MCL fibroblasts exhibited a differential response in the collagen mRNA induction to equibiaxial strain.^[13] In response to cyclic strains, ACL fibroblasts increased the expression levels of collagen type I mRNA with no apparent changes in type III mRNA. MCL fibroblasts, however, exhibited a significant increase in collagen type III but not in collagen type I mRNA. The effect of mechanical loading on gene expression has also been studied in an organ culture system, where an increase in collagen I mRNA expression was found when the

ACL was subjected to mechanical loading.^[43] Similar to the differential responses in collagen gene expression, differential responses in terms of MMP-2 protein expression due to biaxial stretching were also observed in ACL and MCL fibroblasts.^[44] Although both ACL and MCL fibroblasts induced MMP-2 within 72 hr after stretching, ACL fibroblasts produced significantly more pro-MMP-2 protein than MCL fibroblasts. Thus, mechano-regulation of gene and protein expression appears to depend on the source of fibroblasts.

Furthermore, mechanical stretching of ACL fibroblasts aligned on nanofibers in the direction of mechanical strain affected the ECM production.^[14] More collagen was produced in cells on aligned nanofiber sheets than on randomly oriented sheets when subjected to mechanical stretching. Also, an *in vitro* bioreactor was used to test the effect of stretching and alignment of MCL fibroblasts on the intracellular calcium concentration response to mechanical stretching.^[45] The cells that were both aligned and subjected to mechanical stretching propagated a calcium wave better than cells that were aligned, but not subjected to stretching.

Besides ACL and MCL fibroblasts, the effect of mechanical loading on expression patterns of ECM proteins and matrix-degrading enzymes has also been noted in periodontal ligament (PDL) fibroblasts. These cells exhibited differential expression patterns of ECM proteins in response to varying mechanical loads. For instance, PDL fibroblasts subjected to 5% biaxial strain exhibited a significant increase in the synthesis of collagen type I and fibronectin and a decrease in tropoelastin.^[15] Similarly, a 10% strain also increased the fibronectin synthesis and decreased the tropoelastin synthesis. However, at this strain level, collagen type I protein levels did not differ from that secreted by control cells. The differential synthesis of collagen I observed at 5% and 10% strain may be because of the differential effect of mechanical loading on genes encoding collagens and collagenases. Furthermore, these cells exhibited differential synthesis of ECM proteins such as collagen I and fibronectin and MMP-2 and TIMP-2 mRNA when subjected to different loading conditions: tensile and compressive.^[16] In addition, mechanical unloading of PDL fibroblasts in the free floating gel model was shown to decrease the collagen content and increase the activity of MMP-2 and MMP-9.^[46]

MECHANO-REGULATION IN TISSUE ENGINEERED CONSTRUCTS

As in native tissues such as tendons and ligaments, mechanical forces also regulate the functionality of

tissue engineered constructs.^[47] In cardiac cells in gelatin-matrix scaffolds, for instance, the application of cyclic stretching stimulated the formation and organization of the ECM.^[48] Also, the application of unidirectional mechanical stretching generated highly differentiated cardiac tissue constructs from neonatal rat cells.^[49,50]

Furthermore, in collagen gels seeded with dermal fibroblasts, the application of cyclic strain, superimposed on a preload, can have both anabolic and catabolic effects on cells.^[29,30] Cyclic strain superimposed on preloads of 2 and 10 mN increased cell proliferation compared with that of the just preload alone. However, enhanced collagen synthesis in response to cyclic strain was observed only in collagen gel constructs preloaded at 2 mN, implying that the preload level is an important factor to be considered in engineering functional tissue constructs. In addition, the cyclic strain on the cell-seeded constructs induced both active and latent MMPs on a 10 mN preload, but not on 2 mN. The data demonstrate that cyclic strain can shift fibroblasts to a more active catabolic state, depending on the level of the initial preload.

Tendons and ligaments are two other prime candidates for the development of engineered tissue constructs. A novel system for engineering bioartificial tendons (BATs) via the application of mechanical loading on cells cultured on collagen gels has been developed.^[51] The application of mechanical strain greatly enhanced the mechanical properties of BATs by increasing the deposition of ECM components such as collagens I, III, and XII, as well as aggrecan, fibronectin, and tenascin. Furthermore, cells in BATs aligned with the principal strain direction assumed the morphology of fascicles of the whole tendon. In addition, mechanically loaded BATs gained an ultimate tensile strength that was three fold greater than that of unloaded BATs. Similarly, using a bioreactor system to engineer ligaments from mesenchymal stem cells,^[52] mechanical loading for two weeks was found to induce cell alignment in the direction of the force and increase the expression of ligament specific markers such as collagens I, III, and tenascin-C. Taken together, these studies show that mechanical loading regulates specific phenotype expression of cells in tissue engineered constructs, and that dynamic forces may be essential to the engineering of tissue constructs with mechanical properties suitable for implantation.

CELLULAR MECHANO-TRANSDUCTION

The effects of mechanical forces on fibroblasts are now well recognized; however, the cellular mechano-transduction mechanisms (Fig. 4) by which cells sense and respond to mechanical forces to undergo a cascade

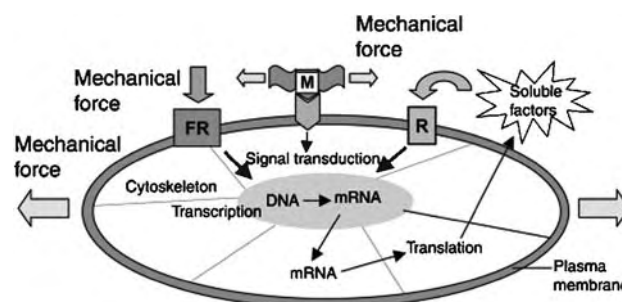


Fig. 4 A conceptual illustration of cellular mechano-transduction mechanisms. Mechanical forces can induce signal transduction by directly transmitting forces to the cytoskeleton and nucleus. As a result, transcription and translation occur. Also, mechanical forces can unfold a domain of the extracellular protein (M) and expose a cryptic site that may serve as an activating ligand for a cell surface receptor, which results in a series of signaling events. Also, mechanical force applies to a “force receptor” (FR), such as integrins and G proteins, and then initiate signal transduction, which results in transcription, followed by translation. This may result in the secretion of soluble factors into ECM, and a soluble factor in turn acts on the receptor (R) and then initiates a cascade of signaling events.

of cellular and molecular events (e.g., ECM synthesis and cell differentiation) are only beginning to be understood. One model of cellular mechano-transduction mechanisms is “hard wired.”^[53,54] In this model, mechanical forces applied to the cell membrane are directly and immediately transmitted to the nucleus through the interconnected cytoskeleton composed of actin filaments, microtubules, and intermediate filaments. Previous studies with various cell types have shown that the application of a mechanical stress to integrins can alter the cytoskeleton and activate signal transduction and gene expression in a stress-dependent manner.^[55–57] Therefore, integrins are likely the force receptors and transducers.^[58] Cellular mechano-transduction, however, may also involve intracellular second messengers that are similar to those from chemical stimuli.^[59] Mechanical stresses, for instance, elevate cytoplasmic Ca^{2+} concentration,^[60,61] which is known to activate Ca^{2+} -responsive proteins, such as protein kinase C, CAM kinase (calcium/calmodulin-dependent kinase), and their downstream transcription factors [e.g., cyclic adenosine monophosphate response element-binding protein (CREB)].^[62]

While many intracellular molecules such as Ca^{2+} and protein kinases^[63] for cellular mechano-transduction have been identified, how mechanical force is initially sensed is still not clear. Candidates for the “force sensor” may include extracellular proteins. In response to mechanical stress, a domain of the extracellular protein might unfold and expose a cryptic site that may serve as an activating ligand for an adjacent

receptor.^[59] This possibility is supported by the fact that a smaller force (80 pN) and a larger force (200 pN) unfold the weakest domain and the most stable domain of fibronectin, respectively.^[64] Also, the direct activation of a transmembrane protein such as an integrin by force can occur, where the transmembrane protein may undergo a conformational change when the force is transmitted from the ECM to the protein. A mechanical force can then be transmitted via cytoskeleton to the nucleus, inducing the transcription and synthesis of a nuclear factor that transactivates a specific ECM gene.^[65] For example, the transcription factor early growth response-1 (EGR-1) and growth factor PDGF-B are regulatory genes that can be directly activated by mechanical stress.^[66,67] All ECM genes with EGR-1 binding sites in their promoters are targets for activation.

Mechanical forces, however, may also indirectly induce synthesis and secretion of growth factors (e.g., TGF- β), which in turn bind to a cell surface receptor, ultimately regulating ECM gene and protein expression, just as chemical stimuli act on a cell. Cyclically stretched fibroblasts, for instance, secrete active TGF- β , which then induces the expression of the procollagen α_1 (I) gene.^[34] Similarly, tensile stress regulates CTGF transcription and secretion.^[35] Other examples of autocrine feedback loops in response to mechanical stress include PDGF,^[68] IL-4,^[69] and adenosine triphosphate.^[70] All these factors may exert their effects on the cell through ligand-receptor binding and result in multimolecular events in an autocrine fashion. Thus, cellular mechano-transduction likely involves a mechano-electrochemical sensory system,^[32,33] including integrin-cytoskeleton machinery, load-conformation sensitive receptors, and mechanically sensitive ion channels.

Note that the description given here is meant to provide a general discussion of cellular mechano-transduction. For a detailed, more in-depth discussion of mechano-transduction in fibroblasts, interested readers should consult existing excellent reviews.^[60,71,72]

CONCLUSIONS

Fibroblasts in cardiac, lung, and connective tissues such as skin, tendons, and ligaments are mechano-responsive cells. In response to mechanical loads, these cells change gene expression and secrete proteins into the ECM, thereby maintaining, repairing, and remodeling these tissues. Determining the mechano-effects of fibroblasts is essential for a better understanding of the physiology and pathology of the various tissues, not to mention their wound healing mechanisms. In addition, the combined effects of

mechanical forces with soluble factors on fibroblasts should be further investigated, because the findings may provide clues to positive and negative cellular responses that are physiological under one set of mechanical loading conditions, but pathophysiological under another. It should be noted that while this review emphasizes the common mechano-responses of fibroblasts, these cells are actually a heterogeneous group,^[73] and their response to mechanical loads likely depend on species, tissue type, and anatomic location.^[74]

More importantly, however, it is necessary to further our understanding of cellular mechano-transduction mechanisms by which these cells sense and respond to mechanical forces. To this end, researchers in the area of fibroblast mechano-biology will face the challenge of developing novel experimental techniques and theoretical models to determine the type and magnitude of forces experienced at the cellular and subcellular levels. It is equally challenging to identify specific “force receptors” or “force sensors,” of which specific proteins at the membrane-cytoskeletal interface (e.g., integrins^[53,58] and G proteins^[75,76]) are good candidates. Ultimately, additional research in the area of fibroblast mechano-biology will aid in designing new approaches to engineer tissue constructs as well as to develop new therapeutic strategies for enhanced tissue repair.

ACKNOWLEDGMENTS

This work was supported in part by the Arthritis Investigator Award and NIH grant AR049921 (JHW).

ARTICLE OF FURTHER INTEREST

Blood Vessel Mechanics

REFERENCES

1. Lee, A.A.; Delhaas, T.; McCulloch, A.D.; Villarreal, F.J. Differential responses of adult cardiac fibroblasts to in vitro biaxial strain patterns. *J. Mol. Cell. Cardiol.* **1999**, *31*, 1833–1843.
2. Atance, J.; Yost, M.J.; Carver, W. Influence of the extracellular matrix on the regulation of cardiac fibroblast behavior by mechanical stretch. *J. Cell Physiol.* **2004**, *200*, 377–386.
3. Butt, R.P.; Bishop, J.E. Mechanical load enhances the stimulatory effect of serum growth factors on cardiac fibroblast procollagen synthesis. *J. Mol. Cell. Cardiol.* **1997**, *29*, 1141–1151.
4. Breen, E.C. Mechanical strain increases type I collagen expression in pulmonary fibroblasts in vitro. *J. Appl. Physiol.* **2000**, *88*, 203–209.

5. Kessler, D.; Dethlefsen, S.; Haase, I.; Plomann, M.; Hirche, F.; Krieg, T.; Eckes, B. Fibroblasts in mechanically stressed collagen lattices assume a “synthetic” phenotype. *J. Biol. Chem.* **2001**, *276*, 36,575–36,585.
6. Chiquet, M.; Sarasa-Renedo, A.; Tunc-Civelek, V. Induction of tenascin-C by cyclic tensile strain versus growth factors: distinct contributions by Rho/ROCK and MAPK signaling pathways. *Biochim. Biophys. Acta* **2004**, *1693*, 193–204.
7. Loesberg, W.A.; Walboomers, X.F.; van Loon, J.J.; Jansen, J.A. The effect of combined cyclic mechanical stretching and microgrooved surface topography on the behavior of fibroblasts. *J. Biomed. Mater. Res. A* **2005**, *75*, 723–732.
8. Banes, A.J.; Tsuzaki, M.; Hu, P.; Brigman, B.; Brown, T.; Almekinders, L.; Lawrence, W.T.; Fischer, T. PDGF-BB, IGF-I and mechanical load stimulate DNA synthesis in avian tendon fibroblasts in vitro. *J. Biomech.* **1995**, *28*, 1505–1513.
9. Banes, A.J.; Tsuzaki, M.; Yamamoto, J.; Fischer, T.; Brigman, B.; Brown, T.; Miller, L. Mechanoreception at the cellular level: the detection, interpretation, and diversity of responses to mechanical signals. *Biochem. Cell Biol.* **1995**, *73*, 349–365.
10. Yang, G.; Crawford, R.C.; Wang, J.H. Proliferation and collagen production of human patellar tendon fibroblasts in response to cyclic uniaxial stretching in serum-free conditions. *J. Biomech.* **2004**, *37*, 1543–1550.
11. Archambault, J.; Tsuzaki, M.; Herzog, W.; Banes, A.J. Stretch and interleukin-1beta induce matrix metalloproteinases in rabbit tendon cells in vitro. *J. Orthop. Res.* **2002**, *20*, 36–39.
12. Kim, S.G.; Akaike, T.; Sasagaw, T.; Atomi, Y.; Kurosawa, H. Gene expression of type I and type III collagen by mechanical stretch in anterior cruciate ligament cells. *Cell Struct. Funct.* **2002**, *27*, 139–144.
13. Hsieh, A.H.; Tsai, C.M.; Ma, Q.J.; Lin, T.; Banes, A.J.; Villarreal, F.J.; Akeson, W.H.; Sung, K.L. Time-dependent increases in type-III collagen gene expression in medical collateral ligament fibroblasts under cyclic strains. *J. Orthop. Res.* **2000**, *18*, 220–227.
14. Lee, C.H.; Shin, H.J.; Cho, I.H.; Kang, Y.M.; Kim, I.A.; Park, K.D.; Shin, J.W. Nanofiber alignment and direction of mechanical strain affect the ECM production of human ACL fibroblast. *Biomaterials* **2005**, *26*, 1261–1270.
15. Howard, P.S.; Kucich, U.; Taliwal, R.; Korostoff, J.M. Mechanical forces alter extracellular matrix synthesis by human periodontal ligament fibroblasts. *J. Periodontal Res.* **1998**, *33*, 500–508.
16. He, Y.; Macarak, E.J.; Korostoff, J.M.; Howard, P.S. Compression and tension: differential effects on matrix accumulation by periodontal ligament fibroblasts in vitro. *Connect. Tissue Res.* **2004**, *45*, 28–39.
17. Hackam, D.J.; Ford, H.R. Cellular, biochemical, and clinical aspects of wound healing. *Surg. Infect. (Larchmt)* **2002**, *3* (suppl. 1), S23–S35.
18. Sacco, O.; Silvestri, M.; Sabatini, F.; Sale, R.; Defilippi, A.C.; Rossi, G.A. Epithelial cells and fibroblasts: structural repair and remodelling in the airways. *Paediatr. Respir. Rev.* **2004**, *5* (suppl. A), S35–S40.
19. Jones, S.M.; Kazlauskas, A. Connecting signaling and cell cycle progression in growth factor-stimulated cells. *Oncogene* **2000**, *19*, 5558–5567.
20. Grazul-Bilska, A.T.; Johnson, M.L.; Bilski, J.J.; Redmer, D.A.; Reynolds, L.P.; Abdullah, A.; Abdullah, K.M. Wound healing: the role of growth factors. *Drugs Today (Barc.)* **2003**, *39*, 787–800.
21. Banes, A.J.; Gilbert, J.; Taylor, D.; Monbureau, O. A new vacuum-operated stress-providing instrument that applies static or variable duration cyclic tension or compression to cells in vitro. *J. Cell Sci.* **1985**, *75*, 35–42.
22. Lee, A.A.; Delhaas, T.; Waldman, L.K.; MacKenna, D.A.; Villarreal, F.J.; McCulloch, A.D. An equibiaxial strain system for cultured cells. *Am. J. Physiol.* **1996**, *271*, C1400–C1408.
23. Huang, H.; Kamm, R.D.; Lee, R.T. Cell mechanics and mechanotransduction: pathways, probes, and physiology. *Am. J. Physiol. Cell Physiol.* **2004**, *287*, C1–C11.
24. Wang, J.H.; Jia, F.; Yang, G.; Yang, S.; Campbell, B.H.; Stone, D.; Woo, S.L. Cyclic mechanical stretching of human tendon fibroblasts increases the production of prostaglandin E2 and levels of cyclooxygenase expression: a novel in vitro model study. *Connect. Tissue Res.* **2003**, *44*, 128–133.
25. Li, Z.; Yang, G.; Khan, M.; Stone, D.; Woo, S.L.; Wang, J.H. Inflammatory response of human tendon fibroblasts to cyclic mechanical stretching. *Am. J. Sports Med.* **2004**, *32*, 435–440.
26. Brown, R.A.; Prajapati, R.; McGrouther, D.A.; Yannas, I.V.; Eastwood, M. Tensional homeostasis in dermal fibroblasts: mechanical responses to mechanical loading in three-dimensional substrates. *J. Cell Physiol.* **1998**, *175*, 323–332.
27. Grinnell, F. Fibroblast-collagen-matrix contraction: growth-factor signalling and mechanical loading. *Trends Cell Biol.* **2000**, *10*, 362–365.
28. Prajapati, R.T.; Chavally-Mis, B.; Herbage, D.; Eastwood, M.; Brown, R.A. Mechanical loading regulates protease production by fibroblasts in three-dimensional collagen substrates. *Wound Repair Regen.* **2000**, *8*, 226–237.
29. Berry, C.C.; Cacou, C.; Lee, D.A.; Bader, D.L.; Shelton, J.C. Dermal fibroblasts respond to mechanical conditioning in a strain profile dependent manner. *Biorheology* **2003**, *40*, 337–345.
30. Berry, C.C.; Shelton, J.C.; Bader, D.L.; Lee, D.A. Influence of external uniaxial cyclic strain on oriented fibroblast-seeded collagen gels. *Tissue Eng.* **2003**, *9*, 613–624.
31. Carver, W.; Nagpal, M.L.; Nachtigal, M.; Borg, T.K.; Terracio, L. Collagen expression in mechanically stimulated cardiac fibroblasts. *Circ. Res.* **1991**, *69*, 116–122.
32. Papakrivopoulou, J.; Lindahl, G.E.; Bishop, J.E.; Laurent, G.J. Differential roles of extracellular signal-regulated kinase 1/2 and p38MAPK in mechanical load-induced procollagen alpha1(I) gene expression in cardiac fibroblasts. *Cardiovasc. Res.* **2004**, *61*, 736–744.



33. Bishop, J.E.; Lindahl, G. Regulation of cardiovascular collagen synthesis by mechanical load. *Cardiovasc. Res.* **1999**, *42*, 27–44.
34. Lindahl, G.E.; Chambers, R.C.; Papakrivopoulou, J.; Dawson, S.J.; Jacobsen, M.C.; Bishop, J.E.; Laurent, G.J. Activation of fibroblast procollagen alpha 1(I) transcription by mechanical strain is transforming growth factor-beta-dependent and involves increased binding of CCAAT-binding factor (CBF/NF-Y) at the proximal promoter. *J. Biol. Chem.* **2002**, *277*, 6153–6161.
35. Schild, C.; Trueb, B. Mechanical stress is required for high-level expression of connective tissue growth factor. *Exp. Cell Res.* **2002**, *274*, 83–91.
36. Bishop, J.E.; Butt, R.; Dawes, K.; Laurent, G. Mechanical load enhances the stimulatory effect of PDGF on pulmonary artery fibroblast procollagen synthesis. *Chest* **1998**, *114*, 25S.
37. Parsons, M.; Kessler, E.; Laurent, G.J.; Brown, R.A.; Bishop, J.E. Mechanical load enhances procollagen processing in dermal fibroblasts by regulating levels of procollagen C-proteinase. *Exp. Cell Res.* **1999**, *252*, 319–331.
38. Chiquet-Ehrismann, R.; Tannheimer, M.; Koch, M.; Brunner, A.; Spring, J.; Martin, D.; Baumgartner, S.; Chiquet, M. Tenascin-C expression by fibroblasts is elevated in stressed collagen gels. *J. Cell Biol.* **1994**, *127*, 2093–2101.
39. Banes, A.J.; Horesovsky, G.; Larson, C.; Tsuzaki, M.; Judex, S.; Archambault, J.; Zernicke, R.; Herzog, W.; Kelley, S.; Miller, L. Mechanical load stimulates expression of novel genes in vivo and in vitro in avian flexor tendon cells. *Osteoarthr. Cartil.* **1999**, *7*, 141–153.
40. Skutek, M.; van Griensven, M.; Zeichen, J.; Brauer, N.; Bosch, U. Cyclic mechanical stretching modulates secretion pattern of growth factors in human tendon fibroblasts. *Eur. J. Appl. Physiol.* **2001**, *86*, 48–52.
41. Wang, J.H.; Yang, G.; Li, Z. Controlling cell responses to cyclic mechanical stretching. *Ann. Biomed. Eng.* **2005**, *33*, 337–342.
42. Yang, G.; Im, H.J.; Wang, J.H. Repetitive mechanical stretching modulates IL-1beta induced COX-2, MMP-1 expression, and PGE(2) production in human patellar tendon fibroblasts. *Gene* **2005**.
43. Hsieh, A.H.; Sah, R.L.; Paul Sung, K.L. Biomechanical regulation of type I collagen gene expression in ACLs in organ culture. *J. Orthop. Res.* **2002**, *20*, 325–331.
44. Zhou, D.; Lee, H.S.; Villarreal, F.; Teng, A.; Lu, E.; Reynolds, S.; Qin, C.; Smith, J.; Sung, K.L. Differential MMP-2 activity of ligament cells under mechanical stretch injury: an in vitro study on human ACL and MCL fibroblasts. *J. Orthop. Res.* **2005**, *23*, 949–957.
45. Jones, B.F.; Wall, M.E.; Carroll, R.L.; Washburn, S.; Banes, A.J. Ligament cells stretch-adapted on a microgrooved substrate increase intercellular communication in response to a mechanical stimulus. *J. Biomech.* **2005**, *38*, 1653–1664.
46. Von den Hoff, J.W. Effects of mechanical tension on matrix degradation by human periodontal ligament cells cultured in collagen gels. *J. Periodontal Res.* **2003**, *38*, 449–457.
47. Butler, D.L.; Goldstein, S.A.; Guilak, F. Functional tissue engineering: the role of biomechanics. *J. Biomech. Eng.* **2000**, *122*, 570–575.
48. Akhyari, P.; Fedak, P.W.; Weisel, R.D.; Lee, T.Y.; Verma, S.; Mickle, D.A.; Li, R.K. Mechanical stretch regimen enhances the formation of bioengineered autologous cardiac muscle grafts. *Circulation* **2002**, *106*, I137–I142.
49. Zimmermann, W.H.; Schneiderbanger, K.; Schubert, P.; Didie, M.; Munzel, F.; Heubach, J.F.; Kostin, S.; Neuhuber, W.L.; Eschenhagen, T. Tissue engineering of a differentiated cardiac muscle construct. *Circ. Res.* **2002**, *90*, 223–230.
50. Zimmermann, W.H.; Melnychenko, I.; Eschenhagen, T. Engineered heart tissue for regeneration of diseased hearts. *Biomaterials* **2004**, *25*, 1639–1647.
51. Garvin, J.; Qi, J.; Maloney, M.; Banes, A.J. Novel system for engineering bioartificial tendons and application of mechanical load. *Tissue Eng.* **2003**, *9*, 967–979.
52. Vunjak-Novakovic, G.; Altman, G.; Horan, R.; Kaplan, D.L. Tissue engineering of ligaments. *Annu. Rev. Biomed. Eng.* **2004**, *6*, 131–156.
53. Ingber, D.E.; Dike, L.; Hansen, L.; Karp, S.; Liley, H.; Maniotis, A.; McNamee, H.; Mooney, D.; Plopper, G.; Sims, J. Cellular tensegrity: exploring how mechanical changes in the cytoskeleton regulate cell growth, migration, and tissue pattern during morphogenesis. *Int. Rev. Cytol.* **1994**, *150*, 173–224.
54. Ingber, D.E. Tensegrity II. How structural networks influence cellular information processing networks. *J. Cell Sci.* **2003**, *116*, 1397–408.
55. Schmidt, C.E.; Horwitz, A.F.; Lauffenburger, D.A.; Sheetz, M.P. Integrin-cytoskeletal interactions in migrating fibroblasts are dynamic, asymmetric, and regulated. *J. Cell Biol.* **1993**, *123*, 977–991.
56. Wang, N.; Butler, J.P.; Ingber, D.E. Mechanotransduction across the cell surface and through the cytoskeleton [see comment]. *Science* **1993**, *260*, 1124–1127.
57. Urbich, C.; Dernbach, E.; Reissner, A.; Vasa, M.; Zeiher, A.M.; Dimmeler, S. Shear stress-induced endothelial cell migration involves integrin signaling via the fibronectin receptor subunits alpha(5) and beta(1). *Arterioscler. Thromb. Vasc. Biol.* **2002**, *22*, 69–75.
58. Ingber, D. Integrins as mechanochemical transducers. *Curr. Opin. Cell Biol.* **1991**, *3*, 841–848.
59. Janmey, P.A.; Weitz, D.A. Dealing with mechanics: mechanisms of force transduction in cells. *Trends Biochem. Sci.* **2004**, *29*, 364–370.
60. Silver, F.H.; Siperko, L.M. Mechanosensing and mechanochemical transduction: how is mechanical energy sensed and converted into chemical energy in an extracellular matrix? *Crit. Rev. Biomed. Eng.* **2003**, *31*, 255–331.
61. Munevar, S.; Wang, Y.L.; Dembo, M. Regulation of mechanical interactions between fibroblasts and the substratum by stretch-activated Ca²⁺ entry. *J. Cell Sci.* **2004**, *117*, 85–92.
62. Iqbal, J.; Zaidi, M. Molecular regulation of mechanotransduction. *Biochem. Biophys. Res. Commun.* **2005**, *328*, 751–755.

63. Ullrich, A.; Schlessinger, J. Signal transduction by receptors with tyrosine kinase activity. *Cell* **1990**, *61*, 203–212.
64. Oberhauser, A.F.; Badilla-Fernandez, C.; Carrion-Vazquez, M.; Fernandez, J.M. The mechanical hierarchies of fibronectin observed with single-molecule AFM. *J. Mol. Biol.* **2002**, *319*, 433–447.
65. Liu, C.; Yao, J.; Mercola, D.; Adamson, E. The transcription factor EGR-1 directly transactivates the fibronectin gene and enhances attachment of human glioblastoma cell line U251. *J. Biol. Chem.* **2000**, *275*, 20,315–20,323.
66. Schwachtgen, J.L.; Houston, P.; Campbell, C.; Sukhatme, V.; Braddock, M. Fluid shear stress activation of egr-1 transcription in cultured human endothelial and epithelial cells is mediated via the extracellular signal-related kinase 1/2 mitogen-activated protein kinase pathway. *J. Clin. Invest.* **1998**, *101*, 2540–2549.
67. Li, C.; Xu, Q. Mechanical stress-initiated signal transductions in vascular smooth muscle cells. *Cell Signal.* **2000**, *12*, 435–445.
68. Wilson, E.; Mai, Q.; Sudhir, K.; Weiss, R.H.; Ives, H.E. Mechanical strain induces growth of vascular smooth muscle cells via autocrine action of PDGF. *J. Cell Biol.* **1993**, *123*, 741–747.
69. Millward-Sadler, S.J.; Wright, M.O.; Lee, H.; Nishida, K.; Caldwell, H.; Nuki, G.; Salter, D.M. Integrin-regulated secretion of interleukin 4: A novel pathway of mechanotransduction in human articular chondrocytes. *J. Cell Biol.* **1999**, *145*, 183–189.
70. Tsuzaki, M.; Bynum, D.; Almekinders, L.; Yang, X.; Faber, J.; Banes, A.J. ATP modulates load-inducible IL-1beta, COX 2, and MMP-3 gene expression in human tendon cells. *J. Cell. Biochem.* **2003**, *89*, 556–562.
71. MacKenna, D.; Summerour, S.R.; Villarreal, F.J. Role of mechanical factors in modulating cardiac fibroblast function and extracellular matrix synthesis. *Cardiovasc. Res.* **2000**, *46*, 257–263.
72. Chiquet, M.; Renedo, A.S.; Huber, F.; Fluck, M. How do fibroblasts translate mechanical signals into changes in extracellular matrix production? *Matrix Biol.* **2003**, *22*, 73–80.
73. Eastwood, M.; McGrouther, D.A.; Brown, R.A. Fibroblast responses to mechanical forces. *Proc. Inst. Mech. Eng. [H]* **1998**, *212*, 85–92.
74. Wall, M.E.; Banes, A.J. Early responses to mechanical load in tendon: role for calcium signaling, gap junctions and intercellular communication. *J. Musculoskel. Neuron. Interact.* **2005**, *5*, 70–84.
75. Gudi, S.R.; Lee, A.A.; Clark, C.B.; Frangos, J.A. Equibiaxial strain and strain rate stimulate early activation of G proteins in cardiac fibroblasts. *Am. J. Physiol.* **1998**, *274*, C1424–C1428.
76. Clark, C.B.; McKnight, N.L.; Frangos, J.A. Strain and strain rate activation of G proteins in human endothelial cells. *Biochem. Biophys. Res. Commun.* **2002**, *299*, 258–262.

Medical Imaging, 3-D

Oskar Skrinjar
Paul J. Benkeser

Wallace H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology
and Emory University, Atlanta, Georgia, U.S.A.

INTRODUCTION

Medical imaging technology has come a long way since its beginnings in 1895 with Röntgen's discovery of X rays and their ability to image structures in two dimensions within the body. Many modern imaging technologies are capable of acquiring, displaying, and manipulating imagery in three dimensions to aid in the diagnosis and treatment of disease and injuries. This article describes the instrumentation of a number of three-dimensional (3-D) imaging technologies commonly used in the practice of medicine, together with a description of the image processing and analysis techniques that aid in the interpretation of the imagery produced by these technologies.

MAGNETIC RESONANCE IMAGING

Magnetic resonance imaging (MRI) is an imaging technique used to produce images of structures inside the body. It is based on the principles of nuclear magnetic resonance (NMR), a standard spectroscopic technique used to obtain chemical and physical information about molecules.^[1] The nucleus of atoms with an odd atomic number or an odd number of neutrons can be detected using NMR techniques. There are many such atoms in the body, but for the purposes of MRI, the hydrogen atom (which consists of a single proton) is of primary interest due to its large magnetic moment. Largely inspired by Lauterbur's development in 1972^[2] of the spatial encoding principles used to form images from NMR data, Kumar, Welti, and Ernst devised a two-dimensional (2-D) Fourier transform approach using field gradient pulses and phase encoding to form the modern foundation for MRI instrumentation.^[3] Such instrumentation is used to acquire multidimensional image data sets in which the contrast is a function of the density, distribution, and relaxational properties of hydrogen protons within the region of the body being imaged. MRI is used to image both structures and function within the body.

Magnetic resonance imaging systems in clinical use today consist of the following components: 1) a magnet

and shim coils to generate a static homogeneous magnetic field with a field strength typically in the range of 0.2 to 3 T; 2) a coil to transmit radiofrequency (RF) signals into the body; 3) a receiver coil to detect the RF signals transmitted by the hydrogen nuclei; 4) gradient coils to provide spatial localization of the RF signals; 5) a computer to control the system timing and reconstruct the RF signals into the image; and 6) a video display system. An example of such a system is illustrated in Fig. 1.

Magnetic resonance imaging systems have a number of attractive features for medical imaging applications. They can generate 2-D sectional images at a variety of orientations, 3-D volumetric images, or four-dimensional (4-D) images representing spatial-spectral distributions, all without any mechanical movement of the instrumentation.^[4] Unlike single photon emission computed tomography (SPECT) and positron emission tomography (PET), MRI does not require injection of radioisotopes. In addition, the imaging process involves the use of RF energy that lies in the nonionizing region of the electromagnetic spectrum.

X-RAY COMPUTED TOMOGRAPHY

The introduction of X-ray computed tomography (CT) in the early 1970s was one of the most profound developments in the field of medical imaging since the time of Röntgen.^[5,6] X-ray CT involves generating X-ray photons, exposing the body to them, detecting and digitizing those photons that are transmitted through the body, and tomographically reconstructing the digitized data to form an image. The function being imaged is the distribution of the linear attenuation coefficient within the body. The logarithm of the ratio of intensities of the generated and transmitted photons is proportional to the integral of the attenuation coefficient along the path. Contrast in X-ray CT is thus dependent on the spatial differences in the attenuation coefficient, which is strongly correlated to tissue density. X-ray CT is primarily used to image structures within the body. However, with high-speed systems and the introduction of contrast agents, it may also be



Fig. 1 Varian 3T magnetic resonance imaging system.

used to image functional changes in organs, such as blood flow rate.

X-ray CT systems in clinical use today typically consist of the following components: 1) a gantry containing a fan-beam X-ray source and a circular array of 600 to 4800 (depending on the manufacturer) independent X-ray detectors that surrounds the patient; 2) a table on which the patient lies that may be translated through the hole in the center of the gantry; 3) a computer for controlling the image acquisition parameters and for reconstructing the image; and 4) a video display system. In fourth-generation systems, the X-ray source is an X-ray tube that is rotated 360 degrees around the patient to acquire multiple projections needed to reconstruct the image. In fifth-generation systems, a high-energy electron beam is electronically scanned along a tungsten strip anode to produce the fan beam of X-rays without mechanical rotation of any component. With these fifth-generation systems, projection data can be acquired fast enough to image the beating heart without significant motion artifacts. Spiral scanning systems, consisting of fourth-generation systems that acquire projection data while the patient is translated through the gantry in a continuous motion, also can be used to image moving structures. An example of an X-ray CT system is shown in Fig. 2.

NUCLEAR TOMOGRAPHIC IMAGING

This is a class of nuclear medicine imaging techniques that utilize tomographic reconstruction methods to image the nuclear radiation emitted from radiopharmaceuticals that have been ingested, inhaled, or



Fig. 2 GE CT/i™ x-ray computed tomography system.

injected into the body. This radiation is in the form of either high-energy gamma rays or positrons, which travel very short distances before they annihilate to produce two 511 KeV photons. This differs from X-ray CT in that the emission source, not the attenuation, is measured. There are two different types of nuclear tomographic imaging techniques: single photon emission computed tomography (SPECT) and positron emission tomography (PET). Both of these techniques are used to image function, not structure. For example, in cardiology applications, with the appropriate radiopharmaceutical (e.g., ^{99m}Tc and FDG) these techniques can be used to measure perfusion in the myocardium.

Single Photon Emission Computed Tomography (SPECT)

The SPECT method was initially investigated by Kuhl and Edwards in 1963 prior to the development of X-ray CT, PET, or MRI.^[7] In a manner similar to X-ray CT, SPECT requires the rotation of a gamma-ray photon detector array around the body to acquire projection data from multiple angles. SPECT systems in clinical use today typically consist of the following components: 1) single or dual photon detector arrays (aka gamma cameras), each with its own collimator; 2) gantry for rotating the camera(s); 3) table on which the patient lies that is positioned in the center of the gantry; 4) computer for controlling the image acquisition parameters and for reconstructing the image; and 5) video display system.

The image reconstruction task for SPECT is far more challenging than for X-ray CT because the emission sources are at unknown positions and densities inside the body. The problem is compounded by the need to use a collimator to acquire projection data, which rejects all photons that are not at a particular

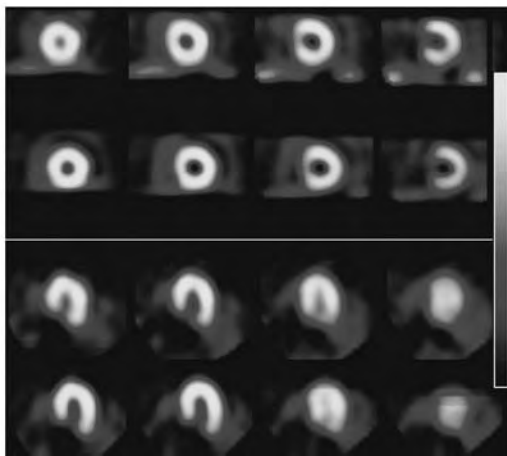


Fig. 3 (top) ADAC Vertex™ dual-head SPECT imaging system; (bottom) normal SPECT myocardial perfusion images.

trajectory angle, thus resulting in a low detection efficiency. Losses due to attenuation of the photons in the body further reduce the obtainable resolution. The latter losses can be partially compensated for during the reconstruction process.^[8] The typical resolution achievable by SPECT systems is on the order of 7 to 15 mm. An example of a SPECT system and typical images are shown in Fig. 3.

Positron Emission Tomography (PET)

The development of computed tomography concepts in the early 1970s spurred the development of the early PET systems.^[9] PET has several inherent advantages over SPECT. First, no collimator is required because the pair of detected 511 KeV gamma-ray photons produced when the positron is annihilated travel along the same line and thus can be traced back to their origin. Second, higher energy photons experience less attenuation, and the attenuation that does occur can

be more easily corrected for compared to SPECT. Third, PET does not require the mechanical rotation of gamma cameras around the patient. Instead, modern PET systems encircle the patient with photon detectors that are electronically coupled in such a manner as to discriminate the coincident detection of the two photons from individual annihilation events. The resolution achievable by this detector system is limited primarily by the uncertainty associated with the path length of the positron prior to annihilation. These advantages are the primary reasons why PET is capable of achieving resolutions three to four times better than SPECT. Although SPECT imaging resolution is inferior to that of PET, the cost of a typical SPECT imaging procedure is currently a factor of two to three times lower than the cost of similar PET procedures. In addition to its higher cost, PET requires the use of radiopharmaceuticals with extremely short half-lives, requiring a close proximity to a particle accelerator. An example of a PET system and typical images are shown in Fig. 4.

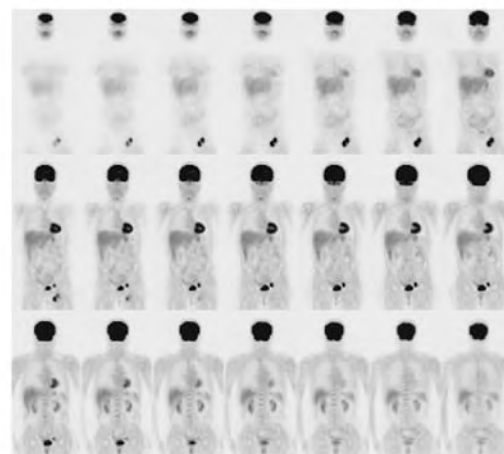


Fig. 4 (top) GE Medical Systems Discovery LS™ with PET and CT imaging capabilities; (bottom) PET whole-body images.



Fig. 5 Philips Integris™ 3-D rotational angiography system.

OTHER CLINICAL 3-D IMAGING TECHNOLOGIES

Although MRI, PET, SPECT, and X-ray CT are the most pervasive 3-D imaging technologies used in clinical settings today, a number of others are worth noting. Three-dimensional imaging can be performed with ultrasound using either mechanical scanning of one-dimensional (1-D) array transducers or through electronic scanning of 2-D array transducers.^[10] Electrical impedance tomography has received considerable attention in the research community, but has found only limited clinical utility to date.^[11] Angiography systems have been developed with solid-state digital X-ray detectors mounted on rotating C-arms, as shown in Fig. 5, to provide 3-D insight into vascular pathologies.^[12]

MEDICAL IMAGE PROCESSING AND ANALYSIS

The development of medical image acquisition techniques has provided some of the most effective diagnostic tools in medicine. These systems are also used for surgical planning and navigation, as well as for imaging in biology. Two-, three-, and higher-dimensional images contain vast amounts of data that need to be interpreted in a timely, accurate, and meaningful manner to benefit clinical and research applications. To assist in interpretation of biomedical images, a number of automated and semiautomated image processing and analysis techniques have been developed.^[13] This overview presents the most standard methods used in interpretation of 3-D medical images: segmentation, registration, and visualization. In addition, a brief summary of other methods is given at the end.

SEGMENTATION

Image segmentation is the partitioning of an image into regions that are homogeneous with respect to one or more characteristics or features.^[13,14] In medical applications, image segmentation is important for feature extraction, shape analysis, image measurements, and image display. There is no one standard segmentation method that can be successfully used in all imaging applications. Rather, a variety of segmentation techniques have been proposed, many of which are application-specific.^[17-19] A recent overview of medical image segmentation methods has been published.^[33] An example of 3-D medical image segmentation is illustrated in Fig. 6.

Image segmentation methods can be classified in several ways:^[13]

- Manual, semiautomatic, and automatic.
- Pixel-based (local methods) and region-based (global methods).
- Manual delineation, low-level segmentation, and model-based segmentation.
- Classical (thresholding, edge-based, and region-based techniques), statistical, fuzzy, and neural network techniques.

Furthermore, most segmentation techniques can be classified as region-based (techniques that look for regions satisfying a given homogeneity criterion) or boundary-based (techniques that look for edges between regions with different characteristics). Most image segmentation methods use one image modality (MRI, CT, ultrasound, PET, SPECT, etc.). Their performance can be enhanced by combining images of different modalities (multispectral segmentation^[26,27]) or images acquired over time (dynamic, motion, or temporal segmentation^[20,28,29]).

Among region-based segmentation methods, the most common ones are thresholding, clustering, region splitting and merging, and region growing.^[13,20] Thresholding is the simplest region-based segmentation method.^[19-21] In this technique a threshold is selected and an image is divided into two groups of pixels: one having values less than the threshold and the other with values greater or equal to the threshold. The threshold can be set globally for the entire image, or can be locally adjusted in an adaptive manner. In addition, methods have been proposed for automatic setting of an optimal threshold value.^[20] Segmentation approaches based on thresholding can produce acceptable results only in simple situations. One of the rare cases in which thresholding is successfully used in medical applications is the segmentation of bony structures from CT images (see Fig. 7 for an example). Thresholding approaches can be enhanced with



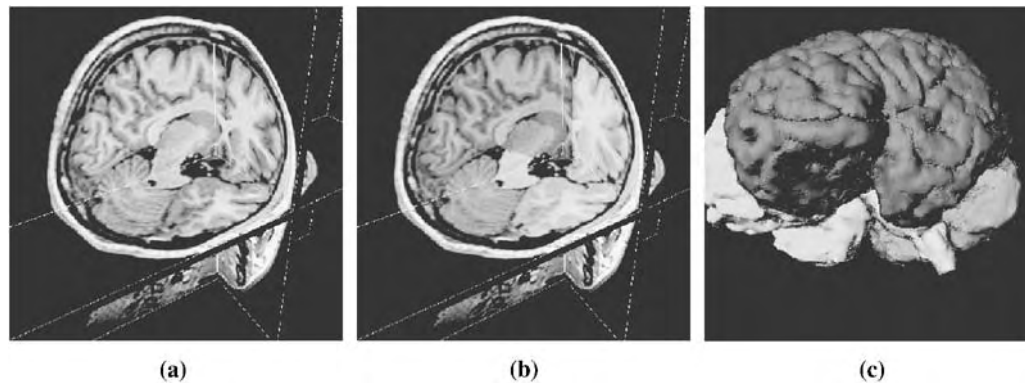


Fig. 6 Example of 3-D medical image segmentation showing a) three orthogonal sections through a 3-D MR image of the human head; b) the same image segmented into the left hemisphere (red), right hemisphere (green), cerebellum (blue), and brain stem (yellow); and c) the 3-D surface models generated for each of the four segmented regions.

mathematical morphology.^[20] Such algorithms, although usually simple, can provide effective and computationally efficient ways to segment complex medical structures. An example of the output of such an algorithm^[32] is illustrated in Fig. 8. Clustering segmentation methods partition the image into clusters of pixels that have strong similarity in the feature space.^[13,22-24] Each pixel is examined and assigned to the cluster that best represents the value of its characteristic vector of features of interest. Region splitting and merging is another region-based segmentation method.^[20] The image is subdivided initially into a set of arbitrary disjointed regions that are then merged and split in an attempt to satisfy certain conditions. Region growing starts from a set of “seed” points and assigns adjacent pixels or regions to the same group if their image values are similar enough, according to some predefined criterion.^[13,20] A typical representative of region growing methods is watershed segmentation.^[20] The idea behind boundary-based segmentation methods is to extract object boundaries and to segment regions enclosed by the

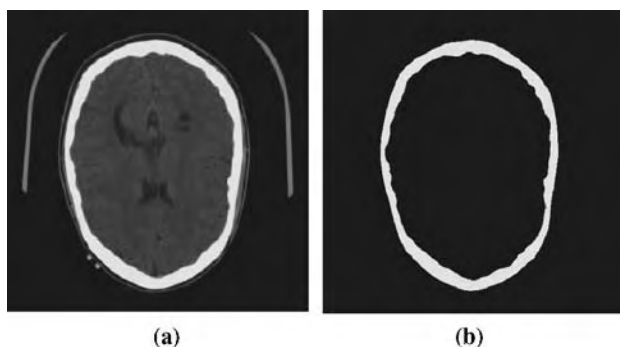


Fig. 7 Skull segmentation b) from a CT image a) of the head by global thresholding (an axial slice through the 3-D volume is shown).

boundaries.^[13,20,25] These algorithms typically rely on an edge detector, such as Sobel, Prewitt, Roberts, or Canny.^[20] One of the most successful classes of boundary-based segmentation strategies are methods based on active boundaries. This line of work was inspired by the pioneer work on active contours (also known as snakes),^[34] which is a 2-D version of the approach. The idea is that the object contours are modeled as deformable curves that automatically deform and attempt to align with image edges, at the same time preserving certain predefined properties (e.g., smoothness, proximity to certain landmarks, sharp corners, etc). The work was extended to 3-D and applied to medical images.^[35-37]

Often, the complexity of medical structures requires the design of application-specific segmentation methods that embody concepts from both region-based and boundary-based approaches. For example, an algorithm for the segmentation of the cortex is described^[38] that, in addition to using image information, restricts the cortical thickness to be within a predefined range, which greatly improves its performance. A more general way to incorporate prior knowledge into segmentation algorithms is by statistical approaches. The idea is to segment a number of objects from a certain class, learn the statistics of the shape variability of each object, and then use each to constrain the segmentation of a new object from the same class. Such an approach has been used to segment the corpus callosum from 3-D MR images of the head.^[39]

REGISTRATION

Image registration is the process of aligning images so that corresponding features can be easily related.^[15] The inputs to an image registration algorithm are the two images to be registered; the output is the geometric

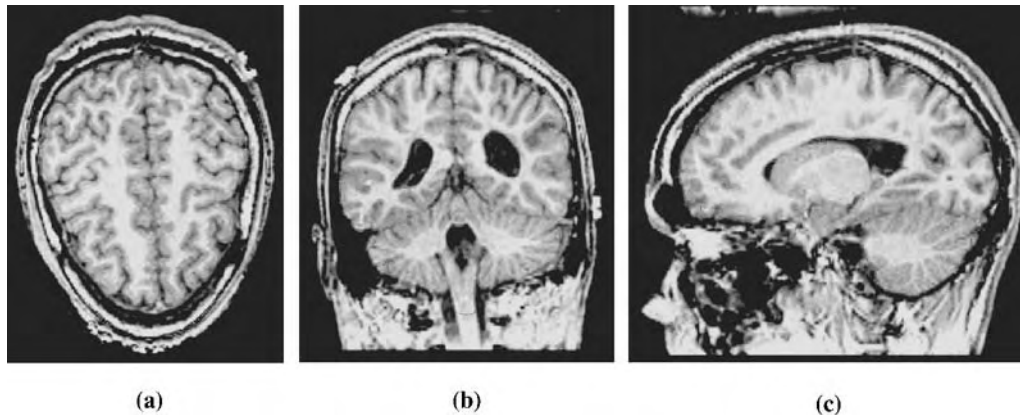


Fig. 8 3-D brain segmentation from MR images using thresholding and mathematical morphology. Contours of the segmented brain are shown in axial a), coronal b), and sagittal c) sections through the 3-D MR image of the head.

transformation that maps points in one image to points in the other.^[14] Many image registration methods have been proposed, and they may be classified in a number of ways.^[30,31] Medical image registration has a wide range of applications that include, but are not limited to: combining information from multiple imaging modalities (e.g., combining images that show function and images that show anatomical structures); monitoring changes in size, shape, location, or image intensity over time; relating preoperative images to the data acquired intraoperatively or during therapy; and relating an individual's anatomy to a standardized atlas.^[15]

Medical image registration methods can be classified based on the following classification criteria: image dimensionality, registration basis, geometric transformation, degree of interaction, optimization procedure, modality, subject, and object.^[14] Each of the criteria is briefly discussed in the remainder of the section.

Medical image registration typically involves 3-D images (3-D to 3-D), but sometimes 2-D images (2-D to 2-D) or images of different dimensions (3-D to 2-D) need to be registered. Figures 9 and 10 illustrate the 3-D to 3-D case, in which functional and anatomical 3-D images are registered and fused.^[40] An

example of the 2-D to 2-D case is the registration of temporal mammograms, which are 2-D images.^[41] Registration of 3-D computed tomography images and 2-D portal images,^[42] which is an example of the 3-D to 2-D case, is often needed in prostate radiation therapy. Portal images are the X-ray images obtained with a radiation treatment beam during the prostate radiation therapy.

Registration basis refers to the underlying mechanics of the registration. By this criterion, image registration algorithms are classified as feature-based, intensity-based, or hybrid. Feature-based algorithms use various features (usually points, curves, and surfaces that are manually or automatically extracted from the two images) to determine the geometric transformation necessary to register the two images.^[43-45] Image-based algorithms use image intensity as the only information to register the images,^[46] and hybrid methods combine both approaches.^[47]

Geometric transformation is the characteristic of image registration algorithms that determines the mathematical form of the geometric mapping used to align the two images. Geometric transformations used in practice range from rigid^[48] to affine^[49] to

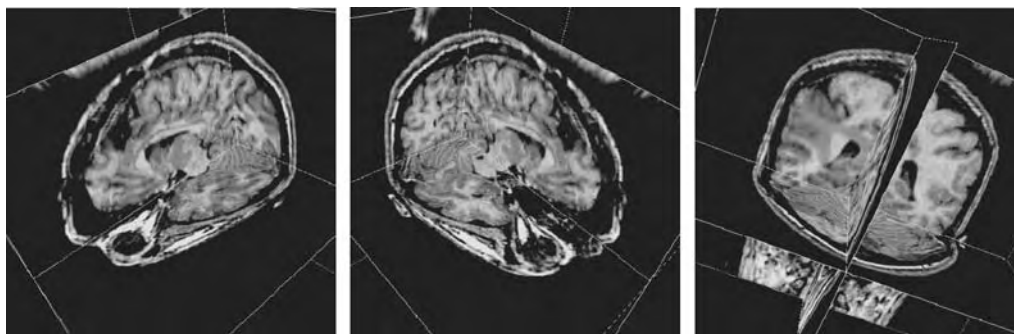


Fig. 9 Rigidly registered and fused SPECT and anatomical MR 3-D images.

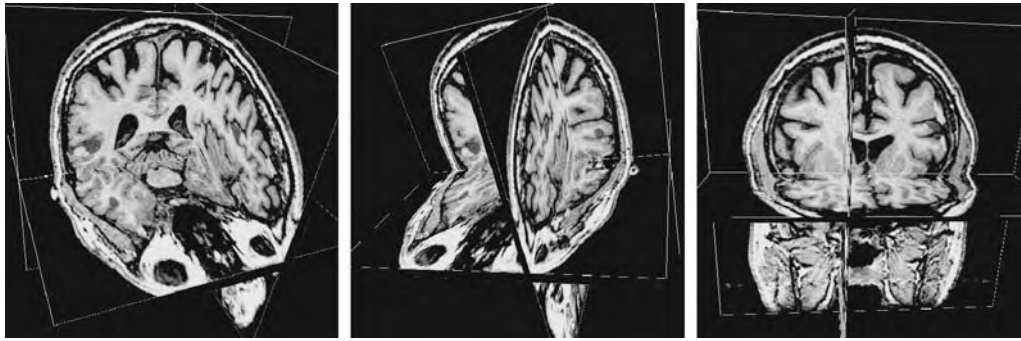


Fig. 10 Rigidly registered and fused functional MR and anatomical MR 3-D images.

nonrigid.^[46] Rigid transformation models (composed of translation and rotation) are used in applications where there is no deformation between the two images except for a change in the location and orientation. For example, in the case of registration of 3-D images of the head of the same patient, the skull, falx, and tentorium provide almost perfectly rigid support for the soft brain tissues, and one can assume a rigid transformation between the two (see Figs. 9 and 10). Affine transformation, in addition to translation and rotation, can account for global scale and shear between the two images. Nonrigid transformation models are needed in cases where the deformation between two images is more complex than affine, e.g., in cases with soft-tissue deformation (cardiac images, pre- and postoperative images, etc.) or in aligning a patient's anatomy to that in an atlas.

Degree of interaction refers to the amount of manual interaction needed by a human operator in performing image registration. The manual interaction may range from no interaction in a fully automated algorithm, which is the ideal situation, to simple initialization of certain parameters or to manual adjustments needed throughout the registration process.

Optimization procedure refers to the approach employed to determine the geometric transformation, i.e., its parameters. Usually, a function of the two images and the mapping between them is used to measure the quality of the registration, and the goal is to optimize (minimize or maximize) the function with respect to the parameters of the transformation. Ideally, the algorithm is based on a closed-form solution that is guaranteed to produce the global extremum of the function. The global extremum can also be determined by an exhaustive search, which in most cases is computationally too expensive, and in some cases it can be determined by means of dynamic programming. In cases in which it is computationally prohibitive to compute the global extremum, iterative optimization schemes are used. Their main disadvantage is that they might find a local extremum that is not the global one,

but they are computationally less expensive and for this reason widely used in practice.

Modality refers to the type of images to be registered. When the two images are of the same modality (e.g., MRI–MRI or CT–CT), the registration is called intramodality or single-modality or monomodality. When the two images are of different modalities (e.g., MRI–SPECT as in Fig. 9 or MRI–fMRI as in Fig. 10), the registration is called intermodality or multimodality.

Medical image registration algorithms based on the subject imaged are classified as: inpatient (the two images are of the same patient), interpatient (the two images are of different patients), and patient-to-atlas (an image of a patient needs to be registered to an atlas image, which itself is typically based on one or more images of different patients).

VISUALIZATION

Three-dimensional visualization generally refers to transformation and display of 3-D objects so as to effectively represent the 3-D nature of the objects. Although an extensive discussion of various visualization techniques for 3-D medical imaging has been published,^[16] here we briefly present the most commonly used ways to display data from 3-D medical images.

Sections through 3-D image volume can be displayed as individual 2-D images (e.g., Fig. 7a, Fig. 8a–c). These 2-D image sections can be displayed together in a 3-D coordinate system as texture-mapped rectangular plane sections (e.g., Fig. 6a–b, Fig. 9, Fig. 10). Such displays provide a better insight about spatial relationships between image structures compared to the case of individually displayed 2-D sections. Usually, the displayed 2-D sections are orthogonal to the three main axes, but sometimes oblique 2-D sections are used.

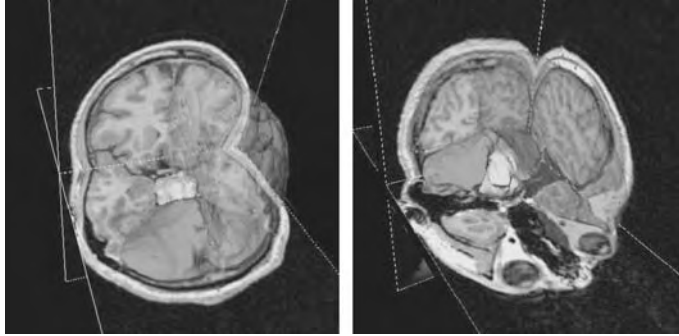


Fig. 11 Examples of surface-rendered models displayed together with orthogonal sections through an MR 3-D image of the human head. Models of the right (green) and left (red) cerebral hemisphere are displayed as transparent surfaces, and models of the cerebellum (blue) and brain stem (yellow) are displayed as opaque surfaces.

In order to provide even better appreciation for the spatial relationships between objects in the 3-D image, the objects of interest are segmented and a surface model is generated around each object, which tessellates the surface with polygons allowing for rapid computer manipulation. These models are then surface-rendered, often combined with other visualization techniques. One can control the color, transparency, visibility, smoothness, complexity, and other properties of these models. Examples of surface-rendered objects are given in Figs. 6c, 11, and 12b–c.

Another way to display 3-D images is volume rendering, during which the entire image is displayed and transparency values are assigned to each voxel in the data set. Ray tracing techniques then give the sense of depth perception without the need for a priori segmentation. Typically only a part of the image volume is shown by using clip planes. Examples are shown in Fig. 12, where Fig. 12b–c illustrates a way to combine surface and volume rendering techniques for visualization of medical 3-D image data.

Note that more complex displays usually require a number of image processing methods to prepare the data for visualization. For example, in order to generate the displays in Fig. 12b–c, the head was segmented, its surface model (which represents the skin)

was generated, and the anatomical MR and functional MR images were (rigidly) registered.

OTHER TECHNIQUES AND VALIDATION

Although segmentation, registration, and visualization are the most widely used techniques for interpretation of 3-D medical image data, a number of other methods have been developed. These methods include image enhancement and restoration, feature extraction, statistical analysis, biomechanical modeling, and quantification.^[13,14,16]

A critical component of every medical image interpretation technique is its validation, because its clinical application affects the health and sometimes the life of the patient. Abnormalities, pathological cases, variability of normal human structures, limited image quality, and image artifacts are among the factors that make automated medical image interpretation difficult. On the other hand, even small errors in interpretation of medical images can negatively and sometimes fatally affect the patient's health. For this reason, any medical image interpretation technique that is intended to be used clinically needs to be validated on a large number of subjects under various conditions. An ideal

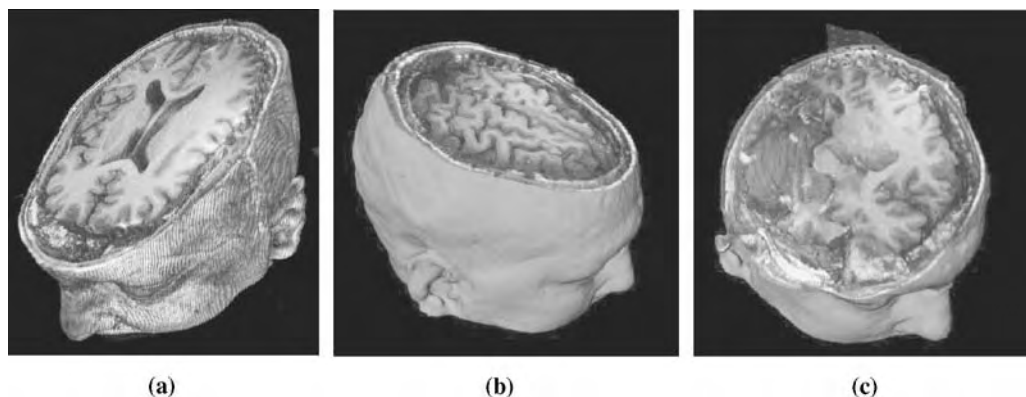


Fig. 12 Examples of visualization of 3-D medical image data. In a), a purely volume-rendered MR 3-D image of a human head is shown cut with an oblique clip plane; in b) and c), volume-rendered fused anatomical MR and functional MR images are displayed together with a surface-rendered skin model.

validation study involves comparison of the output of the medical image interpretation method against ground truth. However, ground truth is rarely available and other nonideal validation approaches need to be used. These include simulations, phantom studies, and qualitative and quantitative comparisons of expert-generated output to the output produced by the medical image interpretation method.

CONCLUSION

Modern imaging technologies are capable of acquiring, displaying, and manipulating imagery in three dimensions to aid in the diagnosis and treatment of disease and injuries. There is no single imaging technology available to satisfy the needs of the entire spectrum of clinical diagnostic procedures employed. Rather, a collection of technologies is available for clinicians to select from in order to achieve the best possible contrast and resolution for a particular imaging procedure.

Image processing and analysis techniques, ranging from simple thresholdbased segmentation to more sophisticated methods, including deformable models, statistical methods, and multimodality image registration, are available for use with the imagery produced with these technologies. The goal of these techniques is to enhance medical procedures by improving their accuracy and reducing the length and cost. The complexity of medical problems often requires application-specific solutions, which need to be extensively validated in order to be acceptable for clinical use. Finally, the 3-D nature of these problems requires advanced visualization techniques in order to properly display the data and allow for an easier interpretation.

ARTICLES OF FURTHER INTEREST

Electrical Impedance Imaging; Functional MRI: Applications; Magnetic Resonance Imaging in Temperature Measurement; Microcomputed Tomography and Its Applications; Optical Coherence Tomography; Positron Emission Tomography (PET); Real Time Tomographic Reflection; Ultrasound Doppler

REFERENCES

- Röntgen, W.K. On a new kind of rays. *Nature* **1896**, 53, 274–276.
- Hornak, J.P. *The Basics of MRI*. <http://wwwcis.rit.edu/htbooks/mri/> (accessed June 2003).
- Lauterbur, P.C. Image formation by induced local interactions: Examples employing nuclear magnetic resonance. *Nature* **1973**, 242, 190–191.
- Kumar, A.; Welte, D.; Ernst, R. NMR Fourier zeugmatography. *J. Magn. Reson.* **1975**, 18, 69.
- Liang, Z.P.; Lauterbur, P.C. Introduction. In *Principles of Magnetic Resonance Imaging*; Akay, M., Ed.; IEEE Press: New York, 2000; 1–12.
- Cho, Z.H.; Jones, J.P.; Singh, M. X-ray Computerized Tomography. In *Foundations of Medical Imaging*; John Wiley & Sons: New York, 1993; 148–164.
- Kuhl, D.E.; Edwards, R.Q. Image separation radioisotope scanning. *Radiology* **1963**, 80, 653–661.
- Zaidi, H.; Hasegawa, B. Determination of the attenuation map in emission tomography. *J. Nucl. Med.* **2003**, 44 (2), 291–315.
- Cho, Z.H.; Chan, J.K.; Eriksson, L. Circular ring transverse axial position camera for 3-D reconstruction of radionuclide distribution. *IEEE Trans. Nucl. Sci.* **1976**, 23, 613–622.
- Fenster, A.; Downey, D.B.; Cardinal, H.N. Three-dimensional ultrasound imaging. *Phys. Med. Biol.* **2001**, 46 (5), R67–R99.
- Zou, Y.; Guo, Z. A review of electrical impedance techniques for breast cancer detection. *Med. Eng. Phys.* **2003**, 25 (2), 79–90.
- Hagen, G.; Wadstrom, J.; Magnusson, A. 3D rotational angiography of transplanted kidneys. *Acta Radiol.* **2003**, 44 (2), 193–198.
- Handbook of Medical Imaging. Processing and Analysis*; Bankman, I.N., Ed.; Academic Press: San Diego, 2000.
- Medical Image Processing and Analysis. In *Handbook of Medical Imaging*; Sonka, M., Fitzpatrick, J.M., Eds.; Spie Press: Bellingham, WA, 2000; Vol. 2.
- Medical Image Registration*; Hajnal J.V., Hill, D.L.G., Hawkes, D.J., Eds.; Biomedical Engineering Series, CRC Press: Boca Raton, FL, 2001.
- Robb, R.A. *Biomedical Imaging, Visualization, and Analysis*; Wiley-Liss, John Wiley & Sons: New York, 2000.
- Bezdek, J.C.; Hall, L.O.; Clarke, L.P. Review of MR image segmentation techniques using pattern recognition. *Med. Phys.* **1993**, 20 (4), 1033–1048.
- Clarke, L.P.; Velthuizen, R.P.; Camacho, M.A.; Heine, J.J.; Vaidyanathan, M.; Hall, L.O.; Thatcher, R.W.; Silbiger, M.L. MRI segmentation: Methods and applications. *Magn. Reson. Imaging* **1995**, 13 (3), 343–368.
- Pal, N.R.; Pal, S.K. A review of image segmentation techniques. *Pattern Recogn.* **1993**, 26 (9), 1227–1249.
- Gonzalez, R.C.; Woods, R.E. *Digital Image Processing*, 2nd Ed.; Prentice Hall: Upper Saddle River, NJ, 2002.
- Sahoo, P.K.; Soltani, S.; Wond, A.K.; Chen, Y.C. A survey of thresholding techniques. *Comput. Vis. Graph. Image Process.* **1988**, 41, 233–260.
- Bramdt, M.E.; Bohan, T.P.; Kramer, L.A.; Fletcher, J.M. Estimation of CSF, white and gray matter volumes in hydrocephalic children using fuzzy clustering of MR imaging. *Comput. Med. Imaging Graph.* **1994**, 18 (1), 25–34.
- Duda, R.O.; Hart, P.E. *Pattern Recognition and Scene Analysis*; Wiley: New York, 1973.

24. Jain, A.K.; Flynn, P.J. Image Segmentation Using Clustering. In *Advances in Image Understanding*; Bowyer, K., Ahuja, N., Eds.; IEEE Computer Society Press: Los Alamitos, CA, 1996.
25. Haralick, R.M.; Shapiro, L.G. Survey: Image segmentation techniques. *Comput. Vis. Graph. Image Process.* **1985**, *29*, 100–132.
26. Fletcher, L.M.; Marsotti, J.B.; Hornak, J.P. A multispectral analysis of brain tissues. *Magn. Reson. Med.* **1993**, *29*, 623–630.
27. Reddick, W.E.; Glass, J.O.; Cook, E.N.; Elkin, T.D.; Deaton, R.J. Automated segmentation and classification of multispectral magnetic resonance images of brain using artificial neural networks. *IEEE Trans. Med. Imag.* **1997**, *16* (6), 911–918.
28. Lucas-Quesada, F.A.; Sinha, U.; Sinha, S. Segmentation strategies for breast tumors from dynamic MR images. *J. Magn. Reson. Imaging* **1996**, *6*, 753–763.
29. Rogowska, J.; Preston, K., Jr.; Hunter, G.J.; Hamberg, L.M.; Kwong, K.K.; Salonen, O.; Wolf, G.L. Applications of similarity mapping in dynamic MRI. *IEEE Trans. Med. Imag.* **1995**, *14* (3), 480–486.
30. Maintz, J.B.A.; Viergever, M.A. A survey of medical image registration. *Med. Image Anal.* **1998**, *2*, 1–36.
31. Maurer, C.D.; Fitzpatrick, J.M. A Review of Medical Image Registration. In *Interactive Image-Guided Neurosurgery*; Maciunas, R.J., Ed.; American Association of Neurological Surgeons: Park Ridge, IL, 1993; 17–44.
32. Stokking, R. Integrated Visualization of functional and Anatomical Brain Images. PhD Thesis; Utrecht University, 1998.
33. Duncan, J.S.; Ayache, N. Medical image analysis: Progress over two decades and the challenges ahead. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *22* (1), 85–106.
34. Kass, M.; Witkin, A.; Terzopoulos, D. Snakes: Active contour models. *Int. J. Comput. Vis.* **1988**, *1* (4), 321–331.
35. Miller, J.V.; Breen, D.E.; Lorensen, W.E.; O'Bara, R.M.; Wozny, M.J. Geometrically deformed models: A method for extracting closed geometric models from volume data. *Comput. Graph.* **1991**, *25* (4), 217–226.
36. McInerney, T.; Terzopoulos, D. A dynamic finite element surface model for segmentation and tracking in multidimensional medical images with application to cardiac 4D image analysis. *Comput. Med. Imaging Graph.* **1995**, *19* (1), 69–83.
37. Cohen, I.; Cohen, L.D.; Ayache, N. Using deformable surfaces to segment 3D images and infer differential structures. *CVGIP, Image Underst.* **1992**, *56* (2), 242–263.
38. Zeng, X.; Staib, L.H.; Schultz, R.T.; Duncan, J.S. Segmentation and measurement of the cortex from 3-D MR images using coupled-surfaces propagation. *IEEE Trans. Med. Imag.* **1999**, *18* (10), 927–937.
39. Leventon, M.E.; Grimson, W.E.L.; Faugeras, O. Statistical shape influence in geodesic active contours. *IEEE Conf. Comput. Vis. Pattern Recognit.* **2000**, *1*, 316–323.
40. Studholme, C.; Hill, D.L.G.; Hawkes, D.J. An overlap invariant entropy measure for 3d medical image alignment. *Pattern Recogn.* **1999**, *32* (1), 71–86.
41. Richard, F.J.P.; Cohen, L.D. A new image registration technique with free boundary constraints: Application to mammography. *Comput. Vis. Image Underst.* **2003**, *89*, 166–196.
42. Bansal, R.; Staib, J.H.; Zhe, C.; Rangarajan, A.; Knisely, J.; Nath, R.; Duncan, J.S. Entropy-based dual-portal-to-3-DCT registration incorporating pixel correlation. *IEEE Trans. Med. Imag.* **2003**, *22* (1), 29–49.
43. Haili, C.; Rangarajan, A. A Feature Registration Framework Using Mixture Models. In *IEEE Workshop on Mathematical Methods in Biomedical Image Analysis, Proceedings*; IEEE Computer Society: Los Alamitos, CA, 2000; 190–197.
44. Pelizzari, C.A.; Chen, G.T.Y.; Spelbring, D.R.; Weichselbaum, R.R.; Chen, C.T. Accurate three-dimensional registration of CT, PET, and/or MR images of brain. *J. Comput. Assist. Tomogr.* **1989**, *13*, 20–26.
45. Cuchet, E.; Knoploch, J.; Dormont, D.; Marsault, C. Registration in neurosurgery and neuroradiotherapy applications. *J. Image Guid. Surg.* **1995**, *1*, 198–207.
46. Rueckert, D.; Sonoda, L.I.; Haynes, C.; Hill, D.L.G.; Leach, M.O.; Hawkes, D.J. Non-rigid registration using free-form deformations: Application to breast MR images. *IEEE Trans. Med. Imag.* **1999**, *18* (8), 712–721.
47. Johnson, H.J.; Christensen, G.E. Consistent landmark and intensity-based image registration. *IEEE Trans. Med. Imag.* **2002**, *21* (5), 450–461.
48. Holden, M.; Hill, D.L.G.; Denton, E.R.E.; Jarosz, J.M.; Cox, T.C.S.; Rohlfing, T.; Goodey, J.; Hawkes, D.J. Voxel similarity measures for 3D serial MR brain image registration. *IEEE Trans. Med. Imag.* **2002**, *19* (2), 94–102.
49. Maes, F.; Collignon, A.; Vandermeulen, D.; Marchal, G.; Suetens, P. Multimodality image registration by maximization of mutual information. *IEEE Trans. Med. Imag.* **1997**, *16* (2), 187–198.

Medical Robotics

John E. Speich

Mechanical Engineering, Virginia Commonwealth University, Richmond, Virginia, U.S.A.

Jacob Rosen

Electrical Engineering, University of Washington, Seattle, Washington, U.S.A.

INTRODUCTION

Today, robotic devices are used to replace missing limbs, perform delicate surgical procedures, deliver neuro-rehabilitation therapy to stroke patients, teach children with learning disabilities, and perform a growing number of other health related tasks. According to the Robot Institute of America, a robot is “a reprogrammable, multifunctional manipulator designed to move material, parts, tools, or specialized devices through various programmed motions for the performance of a variety of tasks” (1979). Given this definition, medical robotics includes a number of devices used for surgery, medical training, rehabilitation therapy, prosthetics, and assisting people with disabilities.

REHABILITATION ROBOTICS

The most extensive use of robotic technology for medical applications has been in rehabilitation robotics, which traditionally includes assistive robots, prosthetics, orthotics, and therapeutic robots. Assistive robots provide greater independence to people with disabilities by helping them perform activities of daily living. For example, robot manipulators can assist individuals who have impaired arm or hand function with basic tasks such as eating and drinking, or with vocational tasks such as opening a filing cabinet. Assistive robotics also includes mobility aides such as wheelchairs and walkers with intelligent navigation and control systems, for individuals with impaired lower-limb function. Robotic prosthetics and orthotics have been developed to replace lost arms, hands, and legs and to provide assistance to weak or impaired limbs. Therapeutic robots are valuable tools for delivering neuro-rehabilitation to the limbs of individuals with disabilities following stroke. An insightful summary of rehabilitation robotics from the 1960s to 2003, can be found in the commentary by Hillman.^[1]

Assistive Robots

A number of robotic systems for assisting individuals with severe disabilities are commercially available. The most widely used is the Handy 1 (Rehab Robotics Limited, UK), which was developed by Topping in 1987.^[2,3] This device enables people with little or no hand function to independently complete everyday functions such as eating, drinking, washing, shaving, and teeth cleaning. MANUS (Exact Dynamics, Netherlands) is a wheelchair-mounted, general-purpose manipulator with six degrees of freedom (DOF) and a two-fingered gripper. It was also designed to assist people with disabilities in completing tasks of daily living.^[4] More than 100 people have used MANUS in their homes in the Netherlands, France, and other countries. The Raptor (Applied Resources Corporation, U.S.A.) is a 4-DOF wheelchair-mounted robot that allows individuals with disabilities to feed themselves and reach objects on the floor, on a table, or above their heads.^[5]

Mobility assistance devices

Robotic technology can be used to equip mobility aides such as wheelchairs and walkers with intelligent navigation and control systems. Such mobility aides are commonly used by the elderly and people with impaired lower limb function or impaired vision. For example, Wasson et al. at the University of Virginia Medical Automation Research Center have developed an intelligent wheeled walker that can assist the user with obstacle avoidance and drop-off detection, and provide minor corrections to the user's steering input.^[6,7] Prassler et al. at the University of Ulm (Germany) have designed a robotic wheelchair called MAid (Mobility Aid for Elderly and Disabled People) with an intelligent navigation and control system for people with limited motor skills.^[8,9]

Vocational assistance devices

Recent studies have shown that robotic technology can greatly benefit motion-impaired individuals during the

performance of vocational tasks. In one study, nine people with manipulation disabilities used a robotic workstation to perform manipulation tasks that they would have been unable to perform otherwise.^[10] In another study, impaired individuals used a force-reflecting PHANToM (SensAble Technologies, Inc., U.S.A.) haptic interface to control a robot manipulator and to perform occupational tasks used in manual dexterity tests.^[11] The results showed that the assistance provided by the force-feedback device improved task performance and decreased task completion time. These studies show that robotic technology has the potential to provide people with disabilities with much greater access to vocational opportunities. ProVAR (Professional Vocational Assistant Robot) is a 7-DOF desktop robot system, currently being developed by Van der Loos et al. at Stanford University and the Veterans Affairs Palo Alto Health Care System, that will be used in vocational environments by individuals with high-level spinal cord injuries.^[12]

Prosthetics

A prosthetic is a mechanical device that substitutes for a missing part of the human body. These devices are often used to provide mobility or manipulation abilities when a limb is lost. The Utah Arm (Motion Control, Inc., U.S.A.) is a computer-controlled, above-the-elbow prosthesis developed by Jacobsen at the University of Utah in the 1980s.^[13] This commercially available arm is controlled using feedback from electromyography (EMG) sensors that measure the response of a muscle to nervous stimulation (electrical activity within muscle fibers). Motion Control, Inc. also makes a two-fingered prosthetic hand that is controlled using myoelectric signals from the remnant limb. Another prosthetic hand is currently being developed at the Scuola Superiore Sant'Anna, Italy,^[14] and Rutgers University^[15] is creating a robotic prosthetic hand with five fingers and twenty DOF using shape-memory alloys as artificial muscles (Fig. 1).

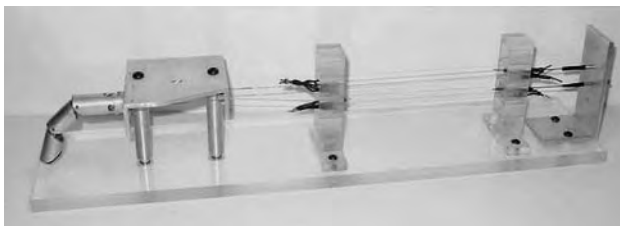


Fig. 1 Photograph of a robotic prosthetic hand under development at Rutgers University. (Photo courtesy of Kathryn De Laurentis and Constantinos Mavroidis, Rutgers University.)

Robotic prosthetics can also be used to replace lower limbs. The MIT LegLab is testing an intelligent prosthetic knee that enables above-the-knee amputees to walk and climb stairs more naturally by adapting the swing rate of the knee accordingly.^[16]

One challenging area of prosthetics research is determining the intended action of the human so that the prosthetic device can be properly controlled. Mussa-Ivaldi at Northwestern University has developed a fish-machine interface that allows a robot to be controlled by the brain of a fish.^[17] Nicolelis et al. at Duke University Medical Center have developed a system that uses implanted electrodes to measure the brain signals in an owl monkey and enables the monkey to control a robot arm to reach for a piece of food.^[18] This research may eventually lead to brain-machine interfaces that can control prosthetic limbs.

Orthotics

An orthotic is a mechanism used to assist or support a weak or ineffective joint, muscle, or limb. Many orthotics utilize robotic technology, and they often take the form of an exoskeleton—a powered anthropomorphic suit that is worn by the patient. Exoskeletons have links and joints that correspond to those of the human and actuators that assist the patient with moving his or her limb or lifting external loads. For example, the Wrist-Hand Orthosis (WHO) uses shape memory alloy actuators to provide a grasping function for quadriplegic patients.^[19] In ongoing research, Rosen et al. at the University of Washington are developing the exoskeleton shown in Fig. 2, which can be controlled by myosignals from the wearer's arm.^[20]

Robot-Assisted Rehabilitation Therapy

Robots have the potential to be valuable tools for rehabilitation therapy. They may enhance traditional treatment techniques by enabling more precise and consistent therapy, especially in therapies that involve highly repetitive movement training. New therapy techniques may be developed using robotic devices that can actively assist and/or resist the motion of the patient. Therapeutic robots can also continuously collect data that can be used to quantitatively measure the patient's progress throughout the recovery process, enabling therapists to optimize treatment techniques. In addition, robot-assisted therapy systems have the potential to provide extended periods of unsupervised therapy, which could increase efficiency and reduce cost by decreasing the amount of one-on-one time that a therapist must spend with a patient.



Fig. 2 Photograph of an exoskeleton for assisting arm movement under development at the University of Washington.

Upper-limb devices

Preliminary research indicates that robotic devices have the potential to greatly enhance the neuro-rehabilitation therapy of stroke patients.^[21–27] Burgar, Lum et al.,^[21,22] Krebs et al.,^[23–25] and Reinkensmeyer et al.^[26,27] have demonstrated that the use of robot-aided therapy can yield positive results in the rehabilitation of forearm movement in stroke patients.

Burgar, Lum et al. at Stanford University and the Veterans Affairs Palo Alto Health Care System have conducted clinical trials using the Mirror-Image Motion Enabler (MIME) robot system shown in Fig. 3, which uses a 6-DOF PUMA 560 robot to interact with the impaired arm.^[21,22] This system can operate in three unilateral modes and one bilateral mode. The unilateral modes are passive, in which the patient remains passive while the robot moves the arm along a preprogrammed path; active-assisted, in which the patient initiates movement and the robot assists and guides the motion along the desired path; and active-constrained, in which the robot resists motion along the path and provides a restoring force in all other directions. This system was used in clinical trials to compare robot-assisted therapy to traditional therapy in stroke patients, and the results

showed greater improvements in the robot group than in the control group.^[21,22]

Krebs et al. at MIT and the Burke Medical Research Institute have conducted clinical trials with the MIT-MANUS, a backdrivable robotic system for delivering upper-extremity neuro-rehabilitation therapy.^[23–25] This system can move and/or guide a patient's arm within a horizontal planar workspace, while recording the motion and applied forces. Clinical trials conducted with the MIT-MANUS have found who patients who underwent robot-aided rehabilitation improved more than patients in the control group.^[24]

In other research, Reinkensmeyer et al. at the University of California at Irvine developed the ARM Guide (Assisted Rehabilitation and Measurement Guide) to evaluate and treat arm impairment following stroke, using linear reaching movements.^[27] The ARM Guide has a single actuator, and the motion of the patient's arm is constrained to a linear path that can be oriented within the horizontal and vertical planes. Initial results with the ARM Guide show that the system produces quantifiable benefits in the neuro-rehabilitation of stroke patients.^[27]

Robotic therapy devices have also been developed for rehabilitating the hand and fingers. Jack et al. have performed preliminary tests indicating that rehabilitation of the hand and fingers may be enhanced using the Rutgers Master II-ND (RMII) force-feedback glove.^[28] This glove has four pneumatic actuators, located in its palm, which interact independently with the index, middle, and ring fingers and the thumb of the right hand. This system provides force feedback and allows the user to interact with a virtual environment. In a pilot clinical trial, three stroke patients used the system daily for two weeks and showed improved hand parameters at the end of the study.



Fig. 3 Photograph of the MIME robotic system for delivering rehabilitation therapy to patients with arm impairment following stroke. (Photo courtesy of Peter Lum, Virginia Commonwealth University.)

Lower-limb devices

The NASA Jet Propulsion Laboratory and UCLA are developing a robotic stepper for lower-limb rehabilitation.^[29] This device uses a pair of robotic arms that resemble knee braces to guide the patient's legs while they move on a treadmill. The system uses a harness to support the patient's weight and several sensors to measure the patient's force, speed, acceleration, and resistance. Astronauts may eventually use this system in microgravity as they exercise to help maintain normal locomotion skills, muscle mass, and calcium levels in bones. In other research, Colombo et al. at the University Hospital Balgrist (Switzerland) have implemented a robotic orthosis to move the legs of spinal cord injury patients during rehabilitation training on a treadmill.^[30] Reinkensmeyer et al. at the University of California at Irvine have also developed a robotic device for measuring and manipulating stepping on a treadmill.^[31] Van Der Loos et al. at the Veterans Affairs Palo Alto Health Care System are studying lower-limb biomechanics using a servomotor-controlled bicycle that can provide both assistance and resistance independently to each leg.^[32]

SURGICAL ROBOTICS

In the last decade, surgery and robotics have reached a maturity that has allowed them to be safely assimilated to create a new kind of operating room. This new environment includes robots for local surgery and telesurgery, audiovisual telecommunication for telemedicine and teleconsultation, robotic systems with integrated imaging for computer-enhanced surgery, and virtual reality (VR) simulators enhanced with haptic feedback, for surgical training. According to Satava, "the operating room of the future will be a sophisticated mix of stereo imaging systems, microbots, robotic manipulators, virtual reality/telepresence workstations, and computer integrated surgery."^[33]

Human–Machine Interfaces in Surgery

Performing a surgical task involves three primary entities: the surgeon, the medium, and the patient. The medium is the means through which the surgeon sees, interacts, and communicates with the patient. It may include standard surgical instruments, an endoscopic camera system, laparoscopic instruments, a robotic surgery system, and/or various other technologies. Figure 4 schematically depicts the human–machine interfaces for various surgical setups. As the physician is moved farther away from the patient, the medium introduced between the surgeon and the patient

becomes more complex and places more constraints into the audio, visual, and physical communication/interaction channels between the surgeon and patient. Nevertheless, in some setups this complex medium may introduce valuable information by providing force feedback, enhancing vision, or enhancing the surgeon's kinematic capabilities by scaling down motion and filtering out hand tremor.

In conventional open surgery, the surgeon interacts with the internal tissues through a relatively large open incision, using direct hand contact or surgical instruments (Fig. 4a). There is no mediator in any of the communication channels: audio, visual, motion, or haptic (force feedback). In the absence of constraints on the surgical tool, the surgeon can translate and orient it anywhere in the surgical scene, using six DOF. In a minimally invasive surgery (MIS) setup, the tools and endoscopic camera are inserted through ports into the body's cavity (Fig. 4b). The port/tool and port/camera interfaces introduce a fulcrum, while decreasing the number of available DOF from six in open surgery to four in MIS, allowing only one (in/out) tool translation. The MIS setup requires at least two operators: the surgeon who is controlling the endoscopic tools and an assistant who is manipulating and positioning the endoscopic camera. The human assistant can anticipate the surgeon's intentions and reposition or track the surgical tools with the endoscopic camera, using minimal directions from the surgeon. However, the assistant is subject to fatigue from holding the camera in one position for long time segments. The assistant can be replaced by AESOP (Computer Motion Inc., U.S.A.), a voice-activated 7-DOF robotic arm that automates the critical task of endoscopic camera positioning and provides the surgeon with direct control over a smooth, precise, and stable view of the internal surgical field.^[34] Broderick, Merrell, et al. have demonstrated that AESOP can also provide improved visibility during open surgery.^[35]

In the United States, robotic surgery can now be performed by the two commercially available systems shown in Fig. 5: ZEUS by Computer Motion^[34] and da Vinci by Intuitive Surgical,^[36,37] which are FDA (Food and Drug Administration) approved for specific cardiac and thoracic surgical procedures. As of July 2003, these two companies have merged into a single company known as Intuitive Surgical. The typical surgical robot architecture follows a classical master/slave teleoperation setup (Fig. 4c,d). This setup consists of two modules: the surgeon console (master) and the robot (slave). The surgeon console includes a set of handles, a vision system, and in some cases voice command components. The robotic system interacting with the patient includes at least three robotic arms: two to manipulate the surgical instruments and a third to control the endoscopic camera. The surgeon controls

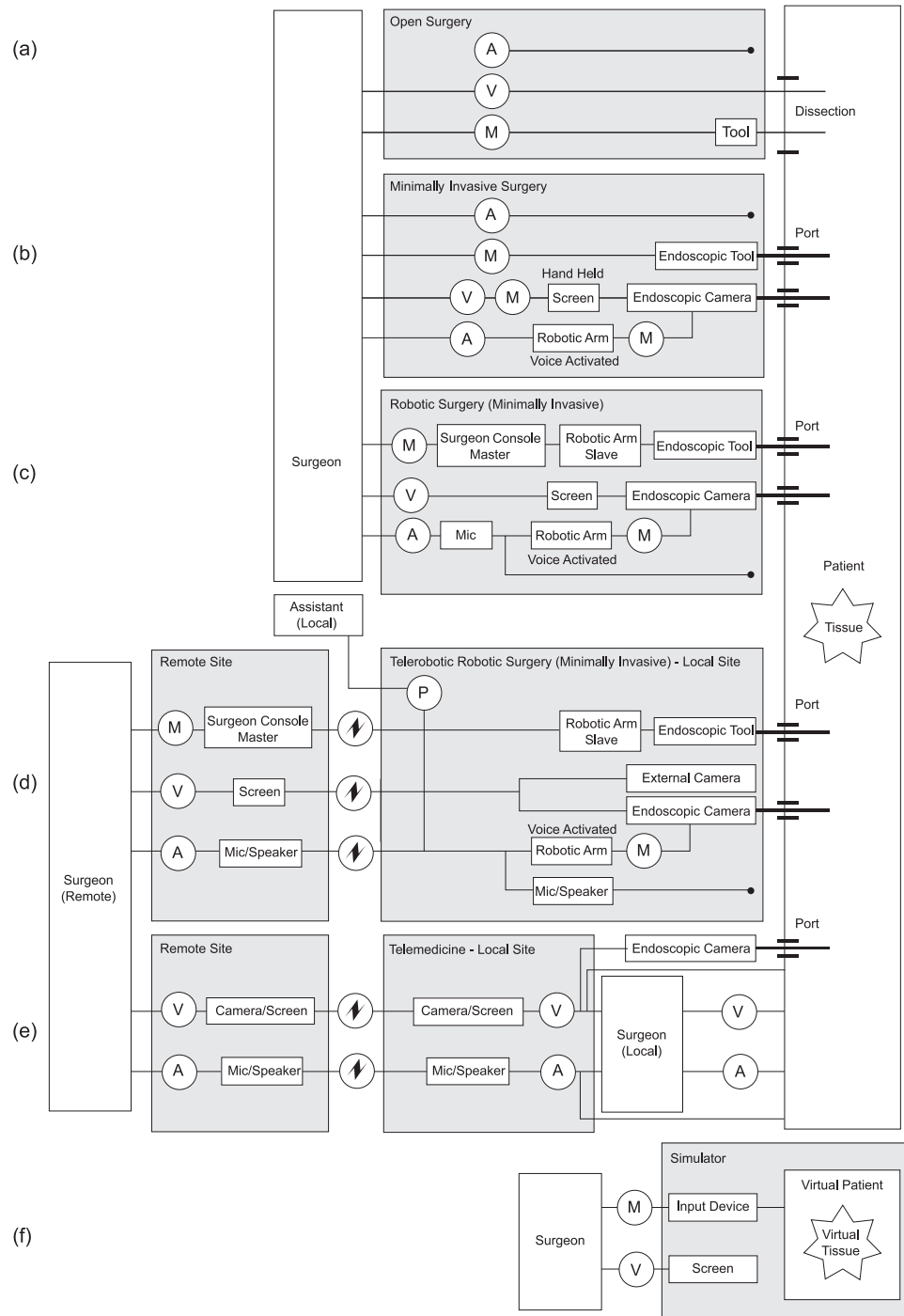


Fig. 4 Modalities used in different configurations for performing surgery: (a) open surgery; (b) minimally invasive surgery; (c) robotic surgery; (d) telerobotic surgery; (e) telemedicine or teleconsultation during surgery; (f) surgical simulation. The type of information being transferred is denoted by (A)–Audio; (M)–Motion, Haptics or Force Feedback; (V)–Vision; and (P)–Positioning.

the position of the robot arms by manipulating the two handles at the console. The endoscopic camera arm is controlled by voice commands from the surgeon, and the view is transmitted back to the surgeon console. None of the currently available surgical systems incorporate force feedback, but the Black Falcon,^[38]

which was used in part as the foundation for the da Vinci system, and other surgical systems^[39] have experimentally tested force feedback. This feedback allows the surgeon to feel the forces generated as the surgical tools interact with the tissue, using a bilateral (position and force) teleoperation mode. Currently, the

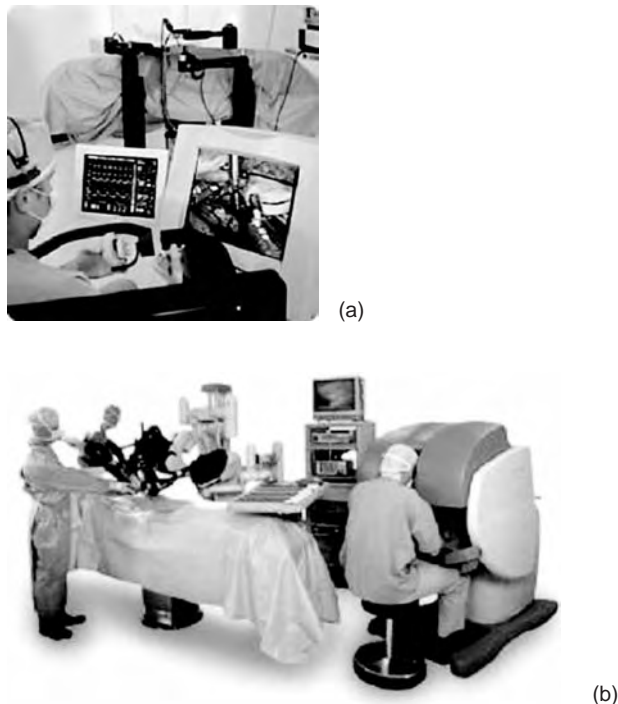


Fig. 5 Commercially available surgical robots: (a) ZEUS by Computer Motion (www.computermotion.com) and (b) da Vinci by Intuitive Surgical (www.intuitivesurgical.com).

FDA has approved only robot-assisted surgical procedures in which the entire robotic system (master/slave) is located in the operating room. However, the same robotic system has been used to perform a surgery telerobotically across the Atlantic Ocean.^[40]

In telemedicine and teleconsultation (Fig. 4e), the remote physician communicates with either the local physician or the patient through audiovisual telecommunication channels. Using systems like SOC-RATES and HERMES by Computer Motion, the remote surgeon can control an external camera, share the view from the endoscopic camera with the local surgeon, and use it as a whiteboard to draw graphics on the image seen by both surgeons.^[34]

For training purposes, the patient, tissue, instruments, and robotic arms can be replaced using computerized simulations (Fig. 4f). The surgeon can practice specific surgical tasks or full clinical procedures by interacting with virtual tissue. Haptic technology can be used to allow the physician to feel the forces generated as a result of the interaction between the virtual tools and the virtual tissue.

One element that all the modalities in Fig. 4 have in common is a human-machine interface, in which visual, kinematic, dynamic, and haptic information are shared between the surgeon and the various modalities. This interface, rich with multidimensional data, is a valuable source of information that can be used to

objectively assess technical surgical skills. Algorithms that are developed for objective skill assessment are independent from the modality being used; therefore, the same algorithms can be incorporated into any of these technologies.^[41–44]

Surgical Robots

The recent evolution of surgical robotics is the result of profound research in the field of robotics and tele-robotics over the past four decades.^[45] By examining the list of strengths and weaknesses of humans and robots in Table 1, it is apparent that combining them into a single system may benefit the level of health care delivered during surgery. The combined system allows the human to provide high-level strategic thinking and decision making while allowing the robot to deliver the actual tool/tissue interaction, using its high precision and accuracy. Because of these characteristics, robots have emerged in the field of surgery and other fields of medicine almost naturally. A number of authors have written reviews of the state of the art in surgical robotics, including Green et al.,^[46] Taylor et al.,^[47,48] Howe,^[49] Buess et al.,^[50] Cleary et al.,^[51] Ballantyne,^[52] and Li et al.^[53]

Surgical robots can be classified into three categories: class i)—semi-autonomous systems; class ii)—guided systems; and class iii)—teleoperation systems. Special robotic arms have been designed in one or more of these categories to meet the requirements of various surgical specialties, including neurological, orthopedic, urological, maxillofacial, ophthalmological, cardiac, and general surgery. Each discipline in surgery has a special set of requirements, dictated by the anatomical structure and the surgical procedure, that necessitate special design and configuration of the robotic system. However, some robotic arm configurations (e.g., ZEUS and da Vinci) are equipped with specially designed sets of tools and may be used for various types of procedures across different surgical disciplines.

Similar to industrial robotics, the tool path of a surgical robot operating in a semi-autonomous mode (class i) is predefined based on a visual representation of the anatomy acquired by an imaging device (e.g., CT, MRI) and preoperative planning. Once the path is defined, the relative locations of the anatomical structure and the robot are registered, and the robot executes the task using position commands without any further intervention on behalf of the surgeon. For obvious safety reasons, the surgeon can stop the action, but altering the path requires replanning. Semi-autonomous robotic systems are suitable for orthopedic or neurological surgical procedures with well-constrained anatomical structures such as hard

Table 1 Characteristics of human and robotic systems

Characteristic	Human	Rank	Robot	Rank
Coordination	Visual/Motor—limited	–	Geometry—Highly accurate	+
Dexterity	High within the range of sensor information	+	Limited by the number and types of sensors—Range exceeds human perception	+
Info. integration	High level—High capacity	+	High level—Limited by AI algorithms	–
	Low level—Limited (info. overload)	–	Low level—High capacity	+
Adaptability	High	+	Limited by design	–
Stable performance	Degrades fast as a function of time	–	Degrades slowly as a function of time	+
Scalability	Inherently limited	–	Limited by design	+
Sterilization	Acceptable	+	Acceptable	+
Accuracy	Inherently limited	–	Designed to exceed human capacity	+
Space occupation	Limited to the human body space	+	Currently exceeds the volume needed to replace the human operator (surgeon)	–
Exposure	Susceptible to radiation and infection	–	Unsusceptible to environmental hazards	+
Specialty	Generic	+	Specialized	–

tissues and bones or with soft tissue such as the brain, confined by the skull.

Surgical robots can be used as guided systems (class ii) in cases where high precision is required, such as in microsurgery, microvascular reconstruction, ophthalmology, or urology. The surgeon interacts directly with the robotic arm and moves the tool in space. The surgical arm provides stable, steady, and precise tool movements, using an impedance control. Forces and torques applied on the system by the surgeon's hand are sensed by a force/torque sensor and translated into a velocity command to the robot.

The architecture of a teleoperated surgical robot (class iii), as previously explained, consists of three fundamental components: the surgical console, the robotic arms, and the vision system. Using the bilateral (motion and force) mode of operation depicted in Fig. 6, the surgeon generates position commands to the robot by moving the input devices (the master) located at the surgeon's console. The position commands are

transferred through a controller to the surgical robotic arms (the slave), which have actuators that move the arms and the surgical tools to the proper positions. In some systems, force feedback may be generated by actuators attached to the master input device, enabling the surgeon to feel the forces between the tool and the tissue. The force-feedback command in a bilateral mode of operation is defined as the difference between the position command generated by the operator and the actual position achieved by the robot.^[37,38] As the difference between the position command and the actual position increases, the force feedback to the operator increases proportionally. Although the bilateral mode of operation does not require additional sensors for generating the force feedback, the high level of friction due to the use of non-direct-drive actuators and the high inertia due to large robotic arms impair the quality of the force-feedback signal using this algorithm. A different approach for incorporating force feedback requires the use of force/torque sensors

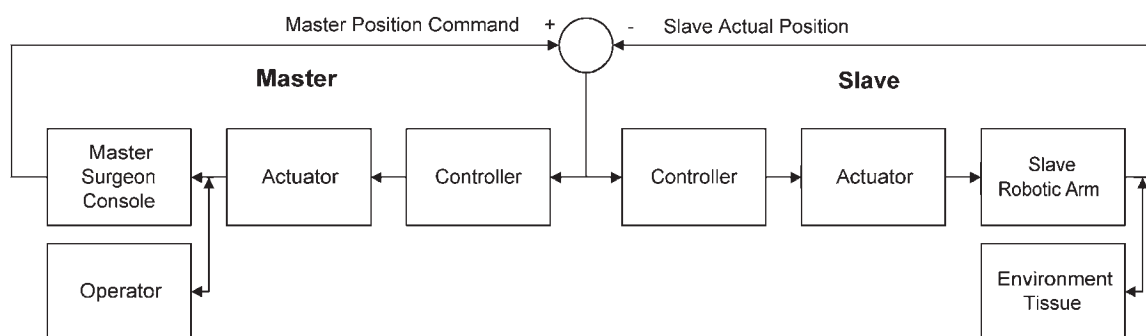


Fig. 6 A block diagram of a typical bilateral teleoperation system used in class iii robotic systems. The actuators and controllers on the master console are eliminated if force feedback is not incorporated into the system.

as close as possible to the end-effector to diminish the mechanical and dynamic interference. Given the harsh environments associated with the tool sterilization and operation, attaching force sensors to the tool and protecting them is still a technological challenge. For MIS, placing force sensors at the distal end of the tool is further limited by the 5–10 mm diameter of the port.

Surgical Training Simulators and Haptics

Training surgical residents adds substantial cost to medical care, including costs associated with inefficient use of operating room time and equipment. Residents are currently trained on a variety of modalities, from using plastic models to operating on live animals and human patients. A resident is more likely to make a mistake than an expert surgeon, and these mistakes can have dire economic, legal, and societal impacts. Ever-increasing costs and louder demands for efficiency have brought surgical simulators to the forefront of training options as a cost-effective and efficacious methodology. Medical simulators are inspired by the aviation simulators used by airline and military pilots to train in virtual reality. Realistic virtual reality surgical simulators allow more comprehensive training without endangering patients' lives. Residents can train for difficult scenarios or anatomy, and they can practice repairing mistakes. In addition, simulators also reduce the need to use animals and cadavers, with obvious ethical and financial benefits.

A typical virtual reality surgical simulator includes both hardware and software. Some input devices include only position sensors and thus provide only the positions of the tools as inputs to the simulator. More advanced input devices, called haptic interfaces, incorporate actuators in addition to position sensors. These actuators generate the appropriate force feedback as the tools are interacting with the virtual medium. Due to the wide variety of surgical procedures, there is no generic input device. Specific input devices are usually developed to match the actual human–machine interface associated with a specific medical procedure as realistically as possible. However, some generic input devices do exist for MIS, including the Laparoscopic Surgical Workstation by Immersion^[54] or a modified version of the PHANTOM by SensAble Technologies.^[55] For an extensive review of medical input devices, see the reference by Chen.^[56] In addition to input devices that are specifically designed for surgical simulators, any surgeon console (master) of a robotic system can be used as an input device, while the patient is replaced by a virtual model. Connecting two or more consoles together may allow two surgeons to share both the visual view and the haptic sensation of

the surgical scene, either in real surgery or in a VR simulation. This will allow the surgeons to regain the collaborative capability that exists in open surgery, but is somewhat lost in MIS. The ability to collaborate may enable a local surgeon to assist an expert surgeon operating from a remote location. In addition, this mode of collaboration may be used during training sessions, where the same tool is controlled by both a trainee and a senior surgeon with overruling authority.

A variety of simulators for surgical training and preoperative planning have been developed by the National Capital Area Medical Simulation Center at the Uniformed Services University of the Health Sciences,^[57] and by the National Center for Biocomputation, a collaboration between Stanford University and the NASA Ames Research Center.^[58] At the heart of a simulator is its computational engine, which accepts the tool positions as inputs and is responsible for presenting the graphical representation of the surgical scene along with generating force feedback as outputs based on a model of the virtual material. Both mass-spring and finite element models are in use for simulating soft tissues. These models are considered to be an oversimplification of real soft tissue, which exhibits nonlinear, heterogeneous, viscoelastic behavior. Measuring the biomechanical characteristics of soft tissue *in vivo* is the subject of active research.^[39,59] Most of the data currently available were acquired *in situ* or under postmortem conditions, which alters the fundamental characteristics of the tissue.

Developing objective methodologies for surgical competence and performance is of paramount importance to superior surgical training. The methodology for assessing surgical skill as a subset of surgical ability has gradually shifted from subjective scoring, based on expert and certainly biased opinion using fuzzy criteria, toward more objective, quantitative analysis. This shift has been enabled by the incorporation of surgical simulators and robots into the surgical training curriculum, in addition to using these tools for demonstrating continued competency among practicing surgeons. The kinematics and dynamics of the surgical tools are fundamental sources of data for objective assessment of surgical skill, regardless of the modality being used. Simple measures like completion time, tool tip path, forces, and torques are currently used as objective criteria, but they fail to provide an integrated approach for analyzing surgery as a multi-dimensional process. Markov modeling can be used to decompose the surgical task and analyze its internal hierarchy, using the kinematics and dynamics of the tools. This technique holds the promise of providing an integrated approach and objective means for quantifying training and skills acquisition prior to clinical implementation.^[43,44]

The Future of Robot-Assisted Surgery

Analysis of the surgical robot's role in the currently available operating room (OR) setup demonstrates that the surgeon can be safely removed from the immediate surgical scene and still maintain interaction with the patient in a teleoperational mode. Although this revolutionary mode of operation may have benefits for the patient, it is far from being efficient because of the lack of supporting technologies. The increased setup and operational time of the current robotic systems is due to lack of automation and the presence of sophisticated interfaces. As a result, the simple act of changing tools or readjusting the robot's position produces inefficient interactions between the clinical staff and the technology. These examples demonstrate the incomplete integration of surgical robotic systems into the OR.

The OR of the future has been envisioned as an integrated information system.^[60] Figure 7 shows a futuristic rendering of some of the subsystems that may be combined within the OR of the future. Much of the medical staff may be removed from the OR and replaced during surgery, in part by hardware in the form of supportive electromechanical devices and in part by software for documenting, assisting, and assessing the operation. The patient may be scanned by an imaging device, which will allow the surgeon to practice critical steps of the operation using the robotic console within a virtual reality environment based on patient-specific data. Then, the operation will be conducted by the surgeon utilizing the robot, tool changer, and equipment dispenser in an OR similar to a class 100 clean room. Smart tags will be incorporated into tools and



Fig. 7 A futuristic rendering of some of the subsystems that might be incorporated into the operating room of the future, including a modular operating table, surgical robotic arms, a tools changer, an equipment dispenser, and an imaging device.

equipment, and once they are used, the billing process and the inventory updates will be executed immediately. Surgical performance will be monitored in the background, and critical decisions may be made through consultation with an expert system incorporated into the system. Much of the core technology for materializing this vision already exists, but whether this vision will become common practice in the next few decades is still an open question.

OTHER MEDICAL ROBOTICS APPLICATIONS

Training

Robotic mannequins have been developed for simulated medical training. The commercially available Medsim-Eagle Patient Simulator developed at Stanford University and the Veterans Affairs Palo Alto Health Care System has several computer-controlled electromechanical features, including eyes that open and close, arms that move, arms and legs that swell, and lungs embedded in the chest that breathe spontaneously.^[61]

Tele-echography

A French consortium has developed a telerobotic echography system consisting of a slave robot, with a real probe as its end-effector, and a master interface with a virtual probe.^[62] This system transmits motion and force information bidirectionally, allowing an expert interacting with the master interface to perform an examination at a remote location, using the slave robot.

Robots for Special Education

AnthroTronix has developed JesterBot™ and CosmoBot™ for the rehabilitation and special education of children.^[63] These robots combine therapy, education, and recreation and can be controlled using body movements, voice commands, or an interactive control station.

Robots for the Deaf and Blind

Dexter, a robotic hand communication aid for people who are both deaf and blind, uses fingerspelling to communicate information typed on a keyboard, stored in a computer, or received from a special telephone.^[64,65]

CONCLUSION

Robotic technology has successfully produced valuable tools for rehabilitation, surgery, and medical training, as well as new and improved prosthetics and assistive devices for people with disabilities. Future applications of robotic technology will continue to provide advances in these and other areas of medicine. The most significant role of medical robots will most likely be to perform tasks that are otherwise impossible, such as enabling new microsurgery procedures by providing high-dexterity access to small anatomical structures, integrating imaging modalities into the OR, providing functional replacements for lost limbs, and enabling new human-machine interfaces and techniques for delivering neuro-rehabilitation therapy.

ACKNOWLEDGMENTS

The authors greatly appreciate the contributions made by Lynette Penn, Pamela Speich, and Jie Li in the preparation of this article. The authors also thank Blake Hannaford, who reviewed the manuscript and contributed his insight and vast knowledge to help shape this article.

ARTICLES OF FURTHER INTEREST

Ergonomics; Eye Tracking: Characteristics and Methods; Eye Tracking: Research Areas and Applications; Gait Analysis; Microelectromechanical Systems (MEMS) Manufacturing; Telemedicine

REFERENCES

- Hillman, M. Rehabilitation Robotics from Past to Present. Proceedings of the Eighth International Conference on Rehabilitation Robotics, April 2003.
- Topping, M. An overview of the development of Handy 1, a rehabilitation robot to assist the severely disabled. *J. Intell. Robot. Syst.: Theory Appl.* **2002**, *34* (3), 253–263.
- <http://www.rehabrobotics.com> (accessed June 2003).
- <http://www.exactdynamics.nl> (accessed June 2003).
- <http://www.appliedresource.com/RTD/Products/Raptor> (accessed June 2003).
- http://marc.med.virginia.edu/projects_elderarerob.html (accessed July 2003).
- Wasson, G.; Gunderson, J.; Graves, S.; Felder, R. An Assistive Robotic Agent for Pedestrian Mobility. International Conference on Autonomous Agents, 2001; 169–173.
- Prassler, E.; Scholz, J.; Fiorini, P. A robotic wheelchair for crowded public environments. *IEEE Robot. Autom. Mag.* **2001**, *8* (1), 38–45.
- http://www.helfenderoboter.de/Produktblaetter/Produktblatt_Rollstuhl_Maid.pdf (accessed June 2003).
- Schuyler, J.; Mahoney, R. Assessing human-robotic performance for vocational placement. *IEEE Trans. Rehabil. Eng.* **2000**, *8* (3), 394–404.
- Pernalet, N.; Yu, W.; Dubey, R.; Moreno, W. Development of a robotic haptic interface to assist the performance of vocational tasks by people with disabilities. *Proc. IEEE Int. Conf. Robot. Autom.* **2002**, *2*, 1269–1274.
- <http://guide.stanford.edu/Projects/02projects/vdl2.html> (accessed June 2003).
- <http://www.utaharm.com> (accessed June 2003).
- Massa, B.; Roccella, S.; Carrozza, M.C.; Dario, P. Design and Development of an Underactuated Prosthetic Hand. 2002 IEEE International Conference on Robotics and Automation, Washington, DC, May 11–15, 2002.
- DeLaurentis, K.; Mavroidis, C. Mechanical design of a shape memory alloy actuated prosthetic hand. *Technol. Health Care* **2002**, *10* (2), 91–106.
- http://www.ai.mit.edu/projects/leglab/mpeg_vcd (accessed June 2003).
- <http://news.bbc.co.uk/1/hi/sci/tech/1043001.stm> (accessed June 2003).
- <http://www.globaltechnoscan.com/22Nov-28Nov/robot.htm> (accessed June 2003).
- Makaran, J.; Dittmer, D.; Buchal, R.; MacArthur, D. The SMART(R) wrist-hand orthosis (WHO) for quadriplegic patients. *J. Prosthet. Orthot.* **1993**, *5* (3), 73–76.
- http://brl.ee.washington.edu/Research_Active/Exoskeleton/Device_03/Exoskeleton_03.html (accessed July 2003).
- Burgar, C.; Lum, P.; Shor, P.; Van der Loos, M. Development of robots for rehabilitation therapy: The Palo Alto VA/Stanford experience. *J. Rehabil. Res. Dev.* **2000**, *37* (6).
- Lum, P.; Burgar, C.; Shor, P.; Majmundar, M.; Van der Loos, M. Robot-assisted movement training compared with conventional therapy techniques for the rehabilitation of upper-limb motor function after stroke. *Arch. Phys. Med. Rehabil.* **2002**, *83* (7), 952–959.
- Krebs, H.; Volpe, B.; Aisen, M.; Hogan, N. Increasing productivity and quality of care: Robot-aided neuro-rehabilitation. *J. Rehabil. Res. Dev.* **2000**, *37* (6).
- Krebs, H.; Hogan, N.; Aisen, M.; Volpe, B. Robot-aided neuro-rehabilitation. *IEEE Trans. Rehabil. Eng.* **1998**, *6* (1), 75–87.
- Fasoli, S.; Krebs, H.; Stein, J.; Frontera, W.; Hogan, N. Effects of robotic therapy on motor impairment and recovery in chronic stroke. *Arch. Phys. Med. Rehabil.* **2003**, *84* (4), 477–482.
- Reinkensmeyer, D.; Dewald, J.; Rymer, W. Guidance-based quantification of arm impairment following brain injury: A pilot study. *IEEE Trans. Rehabil. Eng.* **March 1999**, *7* (1), 1–11.
- Reinkensmeyer, D.; Kahn, L.; Averbuch, M.; McKenna-Cole, A.; Schmit, B.; Rymer, W. Understanding and treating arm movement impairment after

- chronic brain injury: Progress with the ARM guide. *J. Rehabil. Res. Dev.* **2000**, *37* (6).
28. Jack, D.; Boian, R.; Merians, A.; Tremaine, M.; Burdea, G.; Adamovich, S.; Recce, M.; Poizner, H. Virtual reality-enhanced stroke rehabilitation. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2001**, *9* (3), 308–318.
 29. <http://www.jpl.nasa.gov/releases/2000/stepper.html> (accessed June 2003).
 30. Colombo, G.; Joerg, M.; Schreier, R.; Dietz, V. Treadmill training of paraplegic patients using a robotic orthosis. *J. Rehabil. Res. Dev.* **2000**, *37* (6), 693–700.
 31. Reinkensmeyer, D.; Wynne, J.; Harkema, S. A Robotic Tool for Studying Locomotor Adaptation and Rehabilitation. Proceedings in the 2002 IEEE Engineering in Medicine and Biology 24th Annual Conference and the 2002 Fall Meeting of the Biomedical Engineering Society (BMES/EMBS), Houston, TX, October 2002.
 32. Van der Loos, M.; Kautz, S.; Schwandt, D.; Anderson, J.; Chen, G.; Bevly, D. A Split-Crank, Servomotor-Controlled Bicycle Ergometer Design for Studies in Human Biomechanics. 2002 IEEE/RSJ International Conference on Intelligent Robots and Systems, Lausanne, Switzerland, September 2002.
 33. Satava, R. *Cybersurgery: Advanced Technologies for Surgical Practice*; John Wiley & Sons, Inc.: New York, 1997.
 34. <http://www.computermotion.com/>, (accessed August 2003).
 35. Broderick, T.; Russell, K.; Doarn, C.; Merrell, R. A novel method for visualizing the open surgical field. *J. Laparoendosc. Adv. Surg. Tech.* **2002**, *12* (4), 297–302.
 36. <http://www.intuitivesurgical.com/>, (accessed August 2003).
 37. Guthart, G.; Salisbury, K. The Intuitive™ Telesurgery System: Overview and Application. In Proceedings of 2000 IEEE International Conference on Robotics and Automation, **2000**, 618–621.
 38. Madhani, A.; Niemeyer, G.; Salisbury, K. The Black Falcon: A Teleoperated Surgical Instrument for Minimally Invasive Surgery. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS), Victoria B.C., Canada, October 1998.
 39. Rosen, J.; Hannaford, B.; MacFarlane, M.; Sinanan, M. Force controlled and teleoperated endoscopic grasper for minimally invasive surgery—Experimental performance evaluation. *IEEE Trans. Biomed. Eng.* **October 1999**, *46* (10), 1212–1221.
 40. Marescaux, J.; Leroy, J.; Gagner, M.; Rubino, F.; Mutter, D.; Vix, M.; Butner, S.; Smith, M.K. Transatlantic robot-assisted telesurgery. *Nat. Mag.* **2001**, *413*, 379–380.
 41. Rosen, J.; Hannaford, B.; Richards, C.; Sinanan, M. Markov modeling of minimally invasive surgery based on tool/tissue interaction and force/torque signatures for evaluating surgical skills. *IEEE Trans. Biomed. Eng.* **May 2001**, *48* (5), 579–591.
 42. Rosen, J.; Brown, J.; Chang, L.; Barreca, M.; Sinanan, M.; Hannaford, B. The Blue DRAGON—A System for Measuring the Kinematics and the Dynamics of Minimally Invasive Surgical Tools in-vivo. Proceedings of the 2002 IEEE International Conference on Robotics & Automation, Washington, DC, USA, May 11–15, 2002.
 43. Rosen, J.; Solazzo, M.; Hannaford, B.; Sinanan, M. Objective evaluation of laparoscopic skills based on haptic information and tool/tissue interactions. *Comput. Aided Surg.* **July 2002**, *7* (1), 49–61.
 44. Rosen, J.; Chang, L.; Brown, J.D.; Hannaford, B.; Sinanan, M.; Satava, R. *Minimally Invasive Surgery Task Decomposition—Etymology of Endoscopic Suturing, Studies in Health Technology and Informatics—Medicine Meets Virtual Reality*; IOS Press: January 2003; Vol. 94, 295–301.
 45. Hannaford, B. Feeling is Believing: Haptics and Telerobotics Technology. *The Robot in the Garden, Telerobotics and Telepistemology on the Internet*, Cambridge, MA, 1999; Goldberg, J.K., Ed.; MIT Press: Cambridge, MA, 1999.
 46. Green, P.; Hill, J.; Jensen, J.; Shah, A. Telepresence surgery. *IEEE Eng. Med. Biol.* **1995**, *14*, 324–329.
 47. Taylor, R.; Lavallee, S.; Burdea, G.; Mosges, R. *Computer-Integrated Surgery*; MIT Press: Cambridge, MA, 1996.
 48. Taylor, R. Medical Robotics. In *Handbook of Industrial Robotics*, 2nd Ed.; Nof, S.Y., Ed.; Wiley: New York, 1999; 1213–1230.
 49. Howe, R.; Matsuoka, Y. Robotics for surgery. *Annu. Rev. Biomed. Eng.* **1999**, *1*, 211–240.
 50. Buess, G.; Schurr, M.; Fischer, S.; Sabine, C. Robotics and allied technologies in endoscopic surgery. *Arch. Surg.* **2000**, *135*, 229–235.
 51. Cleary, K.; Nguyen, C. State of the art in surgical robotics: Clinical applications and technology challenges. *Comput. Aided Surg.* **2001**, *6*, 312–328.
 52. Ballantyne, G.H. The pitfalls of laparoscopic surgery: Challenges for robotics and telerobotic surgery, special issue on surgical robotics. *Surgical Laparoscopy, Endoscopy and Percutaneous Techniques* **2002**, *12* (1).
 53. Li, Q.; Zamorano, L.; Pandya, A.; Perez, R.; Gong, J.; Diaz, F. The application accuracy of the NeuroMate robot—A quantitative comparison with frameless and frame-based surgical localization systems. *Comput. Aided Surg.* **2002**, *7*, 90–98.
 54. <http://www.immersion.com/>, (accessed August 2003).
 55. <http://www.sensable.com/>, (accessed August 2003).
 56. Chen, E.; Marcus, B. Force feedback for surgical simulation. *Proc. I.E.E.E.* **March 1998**, *86* (3), 524–530.
 57. <http://www.simcen.org/surgery/> (accessed January 2004).
 58. <http://www-biocomp.stanford.edu/> (accessed January 2004).
 59. Brown, J. In-vivo and Postmortem Biomechanics of Abdominal Organs Under Compressive Loads: Experimental Approach in a Laparoscopic Surgery Setup. Ph.D. Dissertation; University of Washington: Seattle, WA, 2003.
 60. Satava, R. Disruptive visions: The operating room of the future. *Surg. Endosc.* **2003**, *17* (1), 104–107.

61. <http://anesthesia.stanford.edu/VASimulator/sim.htm> (accessed June 2003).
62. <http://www.laas.fr/iarp-france/status-reports/2001/medical.html> (accessed July 2003).
63. <http://www.anthrotronix.com> (accessed June 2003).
64. Gilden, D.; Jaffe, D. Dexter—A robotic hand communication aid for the deaf-blind. *Int. J. Rehabil. Res.* **1988**, *II* (2), 198–199.
65. <http://guide.stanford.edu/TTran/jrrd.html> (accessed June 2003).

Melt Spinning

Eija Marjut Pirhonen

Inion Oy, Tampere, Finland

Ville Ellä

Tampere University of Technology, Institute of Biomaterials, Tampere, Finland

INTRODUCTION

Several spinning techniques are used in the production of man-made fibers, including solution spinning (wet or dry), melt spinning, gel spinning, and emulsion spinning, the latter two of which are variations of solution spinning. Of these techniques, melt spinning is the oldest and most widely used. It is an economic and safe method because no toxic solvents are involved and the production rates are high. However, melt spinning by its nature limits the usable polymers to thermoplasts with suitable molecular conformations.

In order to form fibers from glass by melt spinning, glass must have wide working range properties, i.e., the viscosity changes should occur over a large temperature range to allow accurate viscosity control of the melt. Thus, glasses that devitrify easily are very difficult to melt spin into fibers. Some other ceramics and metals have been converted into fibers by melt spinning, but only the phenomena and techniques to convert polymers and glasses into fibers are discussed here.

This chapter gives an overview of the basic principles involved in melt spinning of polymer and glass fibers, as well as application areas of melt-spun biomaterial fibers in biomedical engineering. Melt spinning of polymers involves extrusion, quenching, and drawing. Within these processes, the molecular structure of the polymer changes from randomly oriented bulk melt to highly oriented molecular chains, which give fibers the significantly higher strength properties compared with bulk polymers. With glasses, the structure of the glass network does not significantly differ in bulk and in fibers, but the melt spinning causes stress distribution across the fiber, which, together with surface quality, affects the strength properties of melt-spun glass fibers.

Sutures comprise the best known application area for melt-spun biomaterials, but in recent years interest in the use of fibers in biomedical applications has widely expanded. Fibers are bringing new advantageous properties for scaffolds, especially in tissue engineering, because in many cases fibers resemble biological materials.

MELT SPINNING OF POLYMERS

Characteristics of Polymers Suitable for Melt Spinning

The polymer suitable for melt spinning must meet several requirements:^[1,2]

- It must be a thermoplast, i.e., it must be formable using an elevated temperature.
- Molecular chains of the polymer should be linear, long, and flexible; side groups should be simple, small, or polar.
- Polymer chains should be capable of being oriented and crystallized.
- The polymer should not degrade to a high extent at melting temperature.

Molecular orientation is best reached with polymers that have long, linear, and flexible chains. For practical purposes, the melting temperature of the polymer should be high enough to allow the use of fibers in moderate or high temperatures, such as fibers applied for cloth and reinforcements. Polymers that decompose at high spinning temperatures are usually not suitable for melt spinning, although for some biodegradable polymers the loss of molecular weight during melt spinning can be taken into account in processing. The glass transition temperature should not be too close to the melting range to ensure efficient orientation temperature. However, it should be high enough to retain the characteristics of fibers in their working environment.

The temperature used to melt the polymer has an effect on the rheological properties of the polymer melt. Rheological properties influence the elongational flow of the melt through the spinneret. On the other hand, the elongational flow has an effect on the molecular orientation. The molecular orientation has an effect on forming crystallization because it affects the kinetics of solidification during cooling. Solidification, which is the transformation from a molten state to a solid state, accompanied with the possibility of uniaxial orientation from the melt, is one of the most

important factors in the production of fibers by melt spinning.^[2-4]

The most important structural characteristics of polymers influencing spinnability are degree of crystallinity, crystalline orientation, and the size of the crystallites. These factors govern the spinnability of polymers and, therefore, the spinning procedure itself. Very-high-molecular-weight crystalline polymers with high melt viscosity will induce higher shear stress in the extruder compared with lower-molecular-weight polymers with lower viscosities. For the best spinning results, a wide molecular weight distribution is optimal. Spinning results can be affected, to some extent, by modifications to the machinery or by controlling the chemistry of the polymer.^[1,2,4]

The variables of melt spinning can be divided into three categories. Primary variables include material properties, melt spinning temperature, type of spinneret and orifices, mass outflow, length of spinning path, take-up velocity, and quenching parameters. Secondary variables include extrusion speed, diameter of preoriented fibers, and spin-draw ratio. Resulting variables are take-up tension, take-up stress, fiber temperature, texture, and physical properties of preoriented fibers. Figure 1 presents the scheme of a simplified melt spinning machinery.^[1,2]

Extrusion of Polymer

In the melt spinning process the polymer is melted in the extruder. In the extruder the rotating screw compresses and homogenizes the polymer, which is then farther transferred under pressure directly or via gear pump into the spinneret. Possible additives such as pigments are already mixed into the master batch or added via a separate mixing and feeding unit along the extrusion line, either to the screw or to a metering pump at the end of the extrusion line. The design of the screw and the spinning temperatures vary according to the polymers. Polymer melt is pressed through the spinneret and the filaments are formed. Spinneret type and designs vary according to the end use specifications of the fibers. It is very important that the design of the spinneret enables constant and uniform flow of the melt for every spinneret orifice with equal residence times. Not only circular cross-sections but also different kinds of symmetric and asymmetric patterns are manufactured. The number of holes in a spinneret can vary from a single hole (monofilament spinneret) to hundreds or even thousands of holes (multifilament spinneret).^[3,5]

Quenching of Fibers

After exiting the spinneret, the individual filaments are quenched (i.e., rapidly cooled down) to a solid state so

that they can be hauled off. During quenching, the speed of cooling down and the uniform controlled cooling affects the quality of the filaments as well as the drawing parameters. Orientation already occurred in the melt when the polymer exits the extruder. It should be noted that the speed of the melt exiting the spinneret is lower than the haul of the speed of the filaments, resulting in a preoriented structure. The rate of cooling has an effect on polymer crystallization. With rapid cooling the crystalline structure can be minimized if needed, and by controlled cooling the size of crystallites can be varied. For most polymers the distance from the spinneret to the beginning of the quenching area is 50 to 150 mm. Some polymers do not need quenching. One of these is poly(hydroxybutyrate) (PHB), where quenching below T_g could result in low-quality fibers, whereas drawing from the melt without quenching resulted in preoriented filaments that could be oriented to good-quality fibers.^[6] After quenching, the filaments are oriented via drawing units.^[2-5]

The formed crystalline structure of the undrawn fibers prior to drawing plays a major role in melt spinning. The optimal molecular structure of the undrawn fibers is the foundation for the optimal final structure and for the best physical properties of the drawn fibers.^[2,4]

Drawing of Fibers

In drawing units the fibers are further oriented from the preoriented state either partly or fully. Drawing of melt-spun filaments can be divided into two categories, hot-drawing and high-speed melt spinning with a spin-draw machine. In hot drawing the produced filaments are drawn through ovens, where the filaments are heated up and drawn to a higher degree of orientation. The drawing procedure can be independent from the melt extrusion process, whereas in high-speed melt spinning and the spin-draw machine the filaments are continuously taken from the spinneret through quenching onto the heated godets, where the polymer is heated and drawn. Fig. 1 illustrates how drawing happens between the godets by spinning the D2 godet faster than the D1 godet. Different drawing and relaxation procedures can be obtained by the godet placing, temperature, and speed variations. After godets the fiber is collected on the spools. For online non-woven production the winder is replaced by a staple fiber-cutting machine with possible extensions. For textile use a spin finish (oil or emulsions of oil and water) is applied to the filaments to eliminate electrostatic charge and friction prior to the winding procedure.^[2-5]

The optimum temperature for drawing is between the glass transition temperature and the melting range.



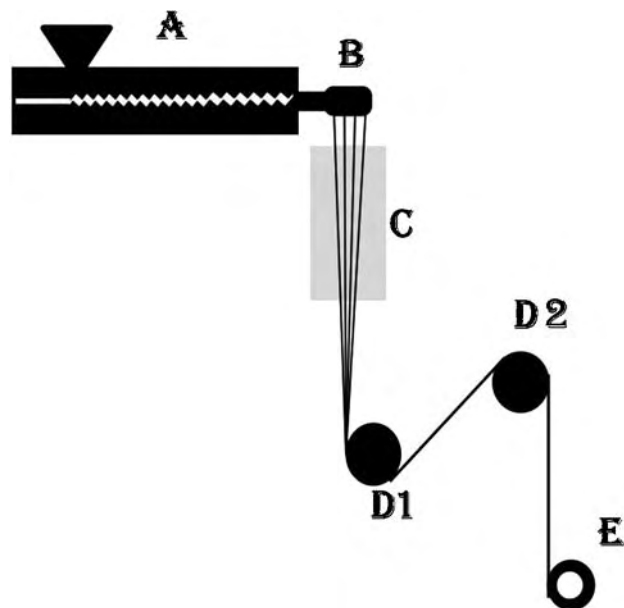


Fig. 1 Scheme of a melt-spinning machinery: (A) extruder, (B) spinneret, (C) quenching area, (D1) first drawing unit or godet, (D2) second drawing unit or godet, and (E) winding unit.

The specific drawing temperature is dependent on the polymer and the requirements of the fiber. As a result of drawing, the final orientation is achieved. Molecular orientation is the parallelization and extension of molecular chains and structural units along the fiber axis. For flexible polymers, the uniaxial orientation is carried out by unfolding the folded molecular chains. When the degree of orientation is increased by drawing, the degree of crystallinity, Young's modulus, and ultimate tensile strength are increased but the elongation (i.e., stretching capacity) is decreased. As the drawing progresses, the movement of the polymer chains becomes more difficult as a result of even closer packing and alignment of the chains. Electrical and optical behavior, as well as absorption properties of solvents, are also affected by the change in orientation. Partially drawn fibers possess the capability of further orientation. They also usually need heat setting after the drawing prior to use in the textile industry as a result of the stretching instability. If heat setting is carried out under stress, the orientation factor can still increase, but when heat setting is carried out free of load, the fibers are reheated and left to shrink longitudinally. Fully drawing the fibers avoids this instability.^[1,2,4]

Properties of Melt-Spun Polymer Fibers

The melt spinning process results in fibers that can be used for different applications. Examples are presented

in Fig. 2. The properties of these oriented fibers differ from each other, depending on the polymer used and the spinning conditions. A proper orientation of molecules gives fibers the optimal properties in the direction of the fiber axis.^[1,2] The difference between a polymer and an oriented fiber can be summarized with polypropylene. As a billet form, its longitudinal yield strength is 25 MPa and elongation at break is around 800%, whereas in a die-drawn form (deformation ratio 20) it has a fiber axis tensile strength of 312 MPa and elongation at break of only 11%.^[7] After melt spinning tensile strength of absorbable polylactide multifilament fibers is generally around 350 to 500 MPa, elastic modulus of 6 to 8 GPa and strain at yield of 25 to 30%.^[8] In comparison, the commercial synthetic absorbable sutures have a tensile breaking strength of 450 to 920 MPa and non-absorbable sutures from 370 to 1060 MPa.^[9]

MELT SPINNING OF BIOACTIVE GLASS

Manufacturing of glass fibers by melt spinning is an important branch of the technical glass industry and technical glass fibers are used, for example, for insulation, as reinforcement of polymers, in special textiles, and as optical fibers. Bioactive glasses were introduced by Hench and colleagues in the late 1960s and since then bioactive glasses have created a great deal of interest for the biomedical society as well as clinicians. Bioactive glass is an amorphous solid that is capable of forming a cohesive bond with both hard and soft tissue when exposed to appropriate *in vivo* or *in vitro* environments, such as simulated body fluid, by developing a surface layer of hydroxycarbonate apatite by release of ionic species from the bulk material. However, in the biomedical field only results from research experiments utilizing bioactive glass fibers have been reported. As bioactive glass fibers present new possibilities to the biomedical materials field they will likely be utilized much more widely within the coming years in the biomedical field as well.

Compositional Dependence of Bioactive Glass Fiber Spinning

The ability of a substance to form glass does not depend upon any particular chemical or physical property. Any liquid, in principle, can be transformed into glass if cooled sufficiently quickly and brought below the crystal transformation range. A good glass-forming material is one for which the rate of crystallization is very slow in relation to the rate of cooling. B_2O_3 , SiO_2 , GeO_2 , and P_2O_5 all readily form glasses on their own and are commonly known as glass

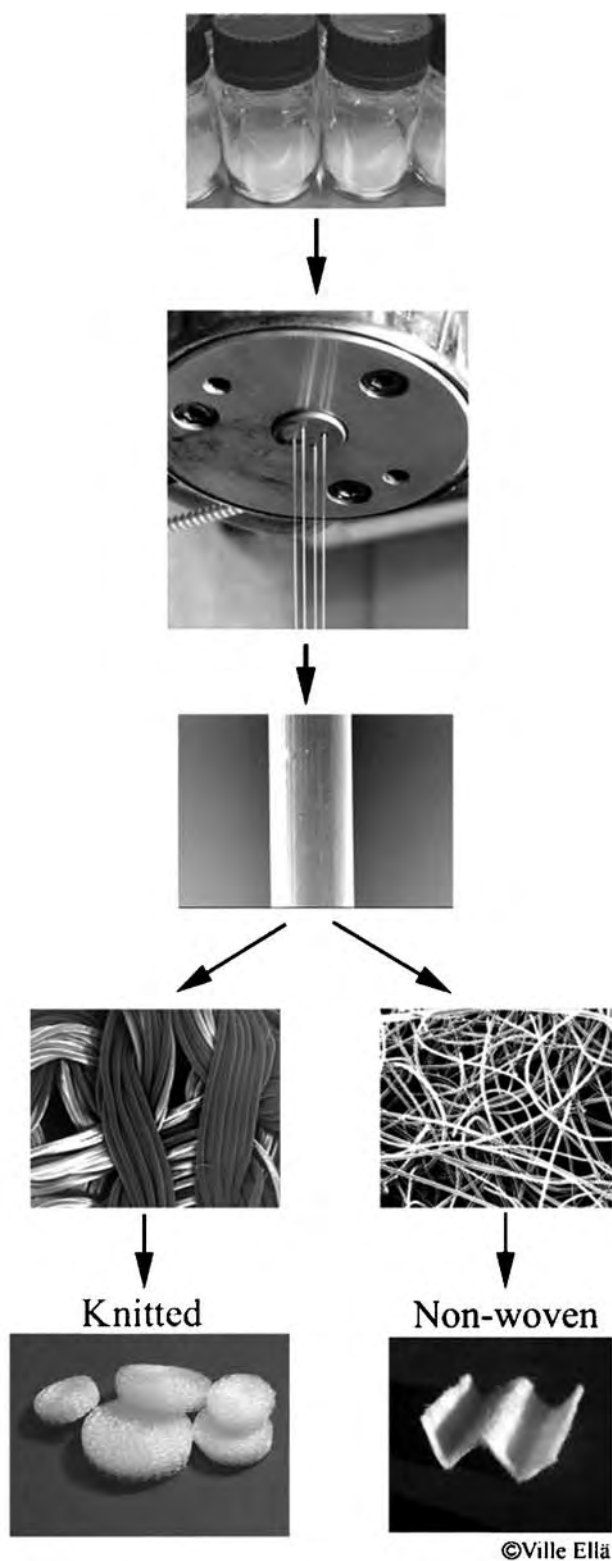


Fig. 2 Melt-spinning steps from a raw material to a product.

formers or network formers because they provide the backbone in other mixed-oxide glasses. With so-called network modifiers (Na^+ , Ca^{2+} , K^+ , Mg^{2+} , etc.), glass properties like degradation rate, bioactivity,

viscosity–temperature dependence, and crystallization behavior can be modified and optimized.

The form of bioactive glasses in clinical use has been limited mainly to the use of glass in the form of crushed particulates. The ability of most of the early bioactive glasses to crystallize when heated renders further heat treatments with these bioactive glasses difficult. Another important feature to enable fiber formation from glass is the viscosity–temperature dependence, or the “working range” of glass. A glass has a narrow working range if the viscosity is significantly changed within a relatively narrow temperature region and it has a large working range if the viscosity changes over a large temperature region. Because in melt spinning viscosity control of the glass melt is a prerequisite for successful controlled manufacture, glass compositions with a large working range are significantly easier to manufacture by melt spinning compared with glasses with a narrow working range. Currently both phosphate- and silicate-based bioactive glasses can be melt spun into fibers.

The most commonly used phosphate glasses are composed of phosphate and calcium, but other network modifiers such as sodium can be added. The first proposals regarding the use of phosphate-based glass fibers for biomedical purposes are from the mid-1980s. One of the limitations of the use of phosphate-based glasses is their rapid degradation. This property can be controlled to some extent by modifying the composition and structure of the glasses.

The first report of the manufacture and osteoconductivity of melt-spun silicate-based bioactive glass fiber was by Pazzaglia and coworkers in 1989.^[10] The composition of the used glass was as follows: 46–53 wt% SiO_2 ; 4–8 wt% P_2O_5 ; 9–20 wt% CaO ; 7–24 wt% Na_2O ; 2.8 wt% K_2O ; 1.11 wt% MgO ; and 0.12 wt% Al_2O_3 . Vita Vinzi Zalman et al.^[7] reported that bioactive glass can be formed into fiber without devitrification when it has the following composition: 40–55 wt% SiO_2 ; 5–8 wt% P_2O_5 ; 20–40 wt% CaO (MgO); 20–30 wt% Na_2O (K_2O); 2–9 wt% K_2O ; 0.5–2.5 wt% Al_2O_3 .

Brink^[11] further studied different glasses in a system containing boron, sodium, potassium, magnesium, calcium, phosphorus oxides, and silica. Brink carried out an extensive study of the crystallization, viscosity, and biological activity of different glass compositions in this system. Bioactivity occurred when the glass contained 14–30 mol % of alkali, as well as alkaline earth oxides, and <59 mol % of silica. The forming properties of bioactive glasses were improved when the content of silica in the glass was increased to 53–56 mol %. Brink also reported that the amount of alkali oxides should be decreased and the amount of alkaline earth oxides increased in glass compared with

the early bioactive glass compositions (e.g., Bioglass[®]) to achieve a large working range.

The Physics of Glass Fiber Drawing

The viscosity of molten glass is very important for manufacturing success. During drawing the viscosity should be relatively low. If it is too high the fiber will break during drawing; however, if it is too low the glass will form droplets rather than be drawn out into a fiber. The optimal viscosity range for fiber spinning is typically between 500 and 1000 dPa sec. The critical time (t) in which a very viscous (η = viscosity) circular jet of radius a leaving an orifice would break up is

$$t = 5\eta a / \sigma$$

if σ is the surface tension. However, fiber is drawn from the nozzle at such speed that its radius is much decreased immediately after the orifice and its radius is so small that it cools and becomes rigid in less than the critical time.

Under the conditions used in making glass fibers it may be assumed that the flow through the nozzle is governed by Poiseuille's law, so that the mass rate of flow M through an orifice of radius a is

$$M = \rho \pi \Delta P a^4 / (8\eta L),$$

where ΔP is the pressure difference causing flow (the hydrostatic head), ρ is the density of the glass, and L is the length of the nozzle. Drawing the fiber more quickly reduces its final diameter without increasing the mass flow rate.^[12]

Glass Fiber Melt Spinning Machinery

In the direct melt process, which is used mainly in technical glass fiber manufacturing, the molten glass in the furnace goes right to the bushing after glass formation. The temperature of the bushing plate is controlled to keep the glass at the correct temperature for fiber formation. The bushing plate normally contains 200 to 4000 nozzles with a diameter of approximately 0.5 mm and is made of platinum alloyed with rhodium for durability. Platinum is used because the glass melt has a natural affinity for wetting it.

Laboratory scale melt spinning can be carried out from the previously manufactured glass block using a simple platinum crucible. The glass block is heated to the desired temperature in a platinum crucible, which has holes as nozzles at the bottom of the crucible. Because the hydrostatic pressure resulting from the glass melt is lower in the small-scale method than in the direct melt process, the required nozzle diameter is larger (approximately 4 mm).^[13] Fig. 3 presents a

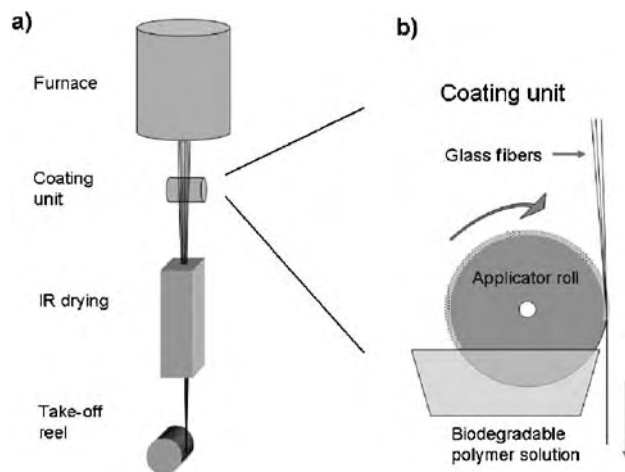


Fig. 3 Scheme of laboratory-scale machinery for bioactive glass melt spinning, including a coating unit.

schematic diagram on laboratory-scale melt spinning machinery, including a coating unit.

As glass flows through the nozzle it forms a drop, which is suspended from the end. As it falls, it leaves a thread attached by the meniscus to the nozzle as long as the viscosity is in the correct range for fiber formation. The formed threads are drawn and attached, e.g., to the rotating coil in order to collect the formed fibers. With controlled drawing speed the diameter of the formed fiber can be adjusted to the desired level.

In technical-scale glass fiber melt spinning the fibers are rapidly cooled right after formation by spraying water or water together with a sizing agent onto the formed fibers before they are wound onto a roll. Sizing agents act as a surface lubricant to minimize damage caused by glass-to-glass abrasion during drawing and winding but their use as a keying agent is also common when the fiber is to be used in some kind of polymer composite.

Structure and Strength of Melt-Spun Glass Fibers

Glass and melt-spun glass fibers are identical with regard to their chemical composition, but have structural differences caused by the manufacturing process and the larger surface area of the fiber. The rapid cooling of glass fiber in melt spinning causes stress distribution to fibers, resulting in compression stress on the surface parts and tensile stress in the inner parts. This toughening phenomenon is caused by the stress distribution, which is a result of the density differences in the different layers of the fiber caused by cooling. Thus, glasses bear viscoelastic behavior, which can be explained by the molecular relaxations of material, so structural changes and stress relaxations can be presumed to occur in glass fibers over long time periods.

The theoretical strength, which can be calculated from the bond energies between different ions and molecules, is about $15,000 \text{ N/mm}^2$ for technical “E” glass fiber. In a laboratory setting strength values close to the theoretical value have been achieved with very thin, faultless fibers. In practice, obtained strength values are much lower depending on the manufacturing method, the fiber diameter, and mostly the fiber surface quality. The structural defects, such as cracks, voids, and impurities on the fiber surface, significantly decrease fiber strength. Because thinner fibers have a smaller surface, there is less likelihood of flaws, which leads to higher strength values with thinner fibers. The Griffith theory is also supported by the dependency of the fiber length on the fiber strength. Because even slight abrasion of the fiber surface drastically reduces the strength of bioactive glass fibers, it is crucial to apply a sizing or coating to the fiber surfaces within fiber melt spinning to maintain the initial high strength of the formed fibers.^[13]

BIOMEDICAL APPLICATIONS OF MELT-SPUN FIBERS

Polymer Fibers

Textiles have a long history in medicine dating back to 2000 BC. The most widely used applications are hospital textiles, such as beddings and sheets, and the largest selling product group are wound dressings. In these applications the material is temporarily and is in contact with the human body for only a short period of time; thus, materials do not need to be biocompatible (i.e., accepted by human tissues).

The most used surgical products are sutures, threads, and yarns used to close wounds. They exist in both biostable (i.e., material remains in the body unchanged) and bioabsorbable (material intentionally and at a predictable rate disappears from the body) forms. The first man-made synthetic bioabsorbable suture materials, Dexon[®] and Vicryl[®], were introduced in the 1970s when the fibers emerged as an interesting aspect in scaffold manufacturing. Bioabsorbable sutures have been extensively used in medicine for decades and new materials have also been developed. Along with new materials, new techniques and modifications of old textile manufacturing systems have emerged.^[9,14]

The main purpose of fibers in the field of biomedical applications has been wound closure operations. Much knowledge has been gained in this area of expertise. Biomaterials used for this purpose are biostable-like silk, polypropylene, nylon 6, nylon 66, polyethylene terephthalate, polybutylene terephthalate, poly(tetramethylene ether) terephthalate-co-tetramethylene terephthalate, and bioabsorbable-like regenerated

collagen, polyglycolic acid, polyglycolide-co-lactide, polyglycolide-co- ϵ -caprolactone, polyglycolide-co-trimethylene carbonate, and poly-*p*-dioxanone. These bioabsorbable sutures must have a certain degree of crystallinity for mechanical strength and in vivo retention time. The development of biomaterial fiber production also gave rise to a whole new set of products and applications. In the field of biomedical engineering the main focus at the moment is filling voids, creating new tissue by creating a support for the tissue, and manufacturing 3-D scaffolds as support structures for living organisms such as bone, ligaments, tendons, veins, and body walls. By manufacturing the scaffold from fibers it is possible to create open and porous networks that still have suitable strength and endurance capabilities. Different processing methods can be utilized with fibers depending on their application. The most commonly used methods with polymeric fibers are weaving, knitting, and braiding.^[9,14]

Wovens and non-wovens

Fabrics weaved from fibers are called woven fabrics. These products are generally used as cloths and drapes in hospitals. In an atmosphere where sterility is a key issue, some fabrics are meant for single use. For these applications non-woven fabrics are the most economic solutions because of their lower manufacturing costs and prices. Several methods for producing non-wovens from staple fibers exist, such as air laying, needle punching, thermal bonding, and wet laying. A non-woven production method for continuous fibers is a spun-laying technique. Manufacturing methods can change the porosity of the fabric. High-porosity fabrics, where the porosity is around 90%, can be manufactured, although increased porosity results in decreased mechanical strength and fabric stability. When comparing the non-woven method with other methods of production, the structure of the non-woven fabric is always random because of the method used, whereas manufacture by weaving, knitting, and braiding results in a well-organized structure. Non-wovens are used for applications such as non-degradable condensed poly(tetrafluoroethylene) PTFE hernia meshes and surgical drapes. Non-wovens produced from melt spun polylactide fibers for cell carrier scaffolds (Fig. 2) have been studied by Sittinger et al.^[15] and Ellä et al.^[8] and some products made from biodegradable non-wovens are already on the market.

Knitting

Knitting is a manufacturing method that includes styles like weft knitting and warp knitting. In knitting



there is continuous feed of the fiber to the fabric for every layer. Both techniques have the capability to produce flat or tubular structures, of which weft knitted are more deformable. Using warp knitting it is possible to manufacture tricot and raschel knits, among other structures. These structures can be used as core structures in engineering composites and implantable device manufacturing or as scaffolds for tissue engineering purposes. The needles used in the machines make loops in the knit and, depending on the fiber and the loop geometry, the loop is the weakest point, thus lowering the maximum possible strength of the knit.^[16] Knitting has been used to manufacture stents and vascular grafts for use in the human body.^[14] As a more engineered implant, Honkanen et al.^[17] studied a knitted circular implant made from polylactic acid fibers as a joint scaffold arthroplasty in metacarpophalangeal joints (Fig. 2).

Braiding

As a manufacturing method, braiding differs from knitting because the fibers are fed from every spool while the spools are spinning around each other, making a criss-cross pattern. This method is used for everyday applications such as shoelaces and cords. Some commercial surgical sutures are manufactured using braiding. Compared with a monofilament, multifilament braiding increases knot break strength, permeability, and flexibility. Knot break strength is the most important factor for a suture material: if a knot breaks during suturing, it cannot be fastened. Braiding is also used to manufacture stents and vascular conduits. The suitability of the method was also studied for use in reinforcement for implantable products and as artificial ligaments. The main interest in braided products lies in their good pulling strength along the braiding axis compared with other manufacturing methods.^[9,14]

Bioactive Glass Fibers

The medical field, and especially the field of tissue engineering, continues to set new challenges for biomaterials engineering. Despite dramatic progress in the field, there is still a need to further enhance the properties of these materials. One fruitful area of research in biomaterials is the use of melt-spun bioactive glass fibers in raw form or as part of biomaterial composites, because bioactive glass fibers are well tolerated by living tissues. Advantageous properties can be achieved, for example, when strong bioactive glass fibers are used as reinforcement in biomedical polymer matrixes.

Sintered porous scaffolds

Bioactive glass is proposed as a good option to achieve repair of bone tissue because of its resorbability and its osteostimulating effect. With the sintering method, in which chopped bioactive glass fibers are connected to each other using heat to form a structure, a scaffold with porous network is obtained. The network has highly interconnected pores, as illustrated in Fig. 4. The porosity, pore size, scaffold architecture, mechanical properties, and degradation rate of the material of the scaffold can all be optimized using this method. The preliminary in vivo and in vitro results indicate that this material bears advantageous properties in use as synthetic bone grafting material and bone tissue engineering applications.^[18]

Reinforcement for bioabsorbable composite structures

Using strong bioactive glass fibers as part of a polymer-based composite, the properties of the plain polymer can be enhanced: 1) bioactive glass fibers reinforce the polymer and the mechanical properties of the obtained composite may reach the level of hard bone tissue; and 2) bioactive glass imparts osteoconductive characteristics to the composite.

The advantage of fiber-reinforced composite materials is the ability to tailor the mechanical or physical properties of the material. For example, the modulus of a unidirectional fiber composite correlates closely with a rule of mixtures,

$$E_{\text{composite}} = fE_{\text{fiber}} + (1 - f)E_{\text{matrix}},$$

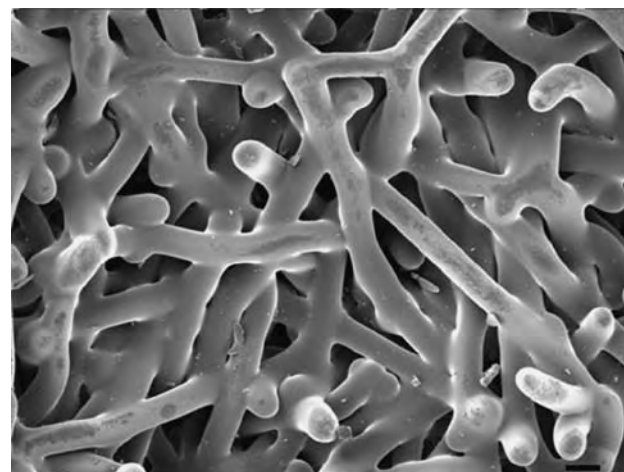


Fig. 4 Scanning electron microscopy image of a porous bioactive glass scaffold manufactured by sintering the melt-spun bioactive glass fibers. Scale bar = 100 μm .

which states that the modulus of a unidirectional fiber composite is proportional to the volume fractions (f) of the materials in the composite. The rule of mixtures can also be applied to composite strength and composite density. Beside increased mechanical performance, bioactive glass as a part of a composite may enhance the biocompatibility of the formed composite when the appropriate amount of glass is free for surrounding tissue fluids. Bioactive glass is reported to enhance osteoblast growth, proliferation, and differentiation and recent findings indicate bioactive glass as a part of tissue engineered scaffolds positively affects the formation of vascularization to the tissue engineered construct.^[13]

CONCLUSIONS

Melt spinning is an old manufacturing technique that can be used with materials that can be melted and reformed. It is mainly utilized to manufacture polymer and glass fibers. Within past decades, the introduction of new biomaterials and further developments in melt spinning and fiber processing techniques, together with the use of more accurate and comprehensive testing, all significantly influenced advancing fibers and fabrics for medical applications. Especially in the field of tissue engineering, ever-new challenges are presented for biomaterials. Many biological structures have fibers in some scale and, thus, it is likely that fibers will play an even greater role in the future in the fields of biomaterials and biomedical engineering. Further information on the subject can be found in the references.

ACKNOWLEDGMENTS

Prof. Minna Kellomäki, Tampere University of Technology, is highly appreciated for valuable comments on the text. Mikko Tukiainen, Tampere University of Technology, is acknowledged for supplying Fig. 3 in this article and Inion Oy, Tampere, Finland, is acknowledged for supplying Fig. 4.

ARTICLES OF FURTHER INTEREST

Electrospinning; Fabrics; Polymers; Vascular Grafts

REFERENCES

1. Nakjima, T. *Advanced Fiber Spinning Technology*; Woodhead: Cambridge, UK, 1994.
2. Ziabicki, A. *Fundamentals of Fiber Formation: The Science of Fiber Spinning and Drawing*; John Wiley & Sons: London, UK, 1976.
3. Billmeyer, F.W., Jr. *Textbook of Polymer Science*, 3rd Ed.; John Wiley & Sons: New York, 1984.
4. Ward, I.M. *Structure and Properties of Oriented Polymers*, 2nd Ed.; Chapman & Hall: London, UK, 1997.
5. Braun, S.; Stausberg, G. Monofilament Extrusion. In *Fiber Extrusion in Plastic Extrusion Technology*, 2nd Ed.; Mayer, M., Ed.; Hanser: Munich, Germany, 1997.
6. Chodak, I.; Blackburn, R.S. Poly(hydroxyalkanoates) and polycaprolactone. In *Biodegradable and Sustainable Fibers*; Blackburn, R.S., Ed.; Woodhead: Cambridge, UK, 2005.
7. Vita Vinzi Zalman, E.; Locardi, B.; Gabbi, C.; Tranquilli Leale, P. Bioactive vitreous composition for bone implants, filaments made therefrom and method. WO 91/12032. August 22, 1991.
8. Ellä, V.E.; Gomes, M.E.; Reis, R.L.; Törmälä, P.; Kellomäki, M. Studies of P(L/D)LA 96/4 non-woven scaffolds and fibres; properties, wettability and cell spreading before and after intrusive treatment methods. *Journal of Materials Science: Materials in Medicine* **2007**, *18*, 1253–1261.
9. Chu, C.C. Textile-based biomaterials for surgical applications. In *Polymeric Biomaterials Second Edition Revised & Expanded*; Dumitriu, S., Ed.; Marcel Dekker: New York, 2002.
10. Pazzaglia, U.P.; Gabbi, C.; Locardi, B.; Di Nucci, A.; Zatti, G.; Cherubino, P. Study of the osteoconductive properties of bioactive glass fibers. *J Biomed Mater Res* **1989**, *23*, 1289–1297.
11. Brink, M. *Bioactive Glasses with a Large Working Range*; Åbo Akademi University: Turku, Finland, 1997; Doctoral Thesis.
12. Cable, M. Classical glass technology. In *Glasses and Amorphous Materials*; Zarzycki, J., Ed.; John Wiley & Sons: Weinheim, 1991; 9.
13. Pirhonen, E. *Fibers and Composites for Potential Biomaterials Applications*; Tampere University of Technology, Finland, 2006; Doctoral Thesis.
14. Gupta, B.S. Medical textile structures: An overview. *Medical Plastics and Biomaterials* **Jan. 1998**, 16–31.
15. Sittinger, M.; Reitzel, D.; Dauner, M.; Hierlemann, H.; Hammer, C.; Kastenbauer, E.; Plack, H.; Burmester, G.R.; Bujia, J., J. *Biomed. Mater. Res. Part B: Appl. Biomater.* **1996**, *33*, 57–63.
16. Leong, K.H.; Ramakrishna, S.; Huang, Z.M.; Bibo, G.A. The potential of knitting for engineering composites—A review. *Composites A* **2000**, *31*, 197–220.
17. Honkanen, P.B.; Kellomäki, M.; Lehtimäki, M.Y.; Törmälä, P.; Mäkelä, S.; Lehto, M.U.K. Bio-reconstructive joint scaffold implant arthroplasty in metacarpophalangeal joints: Short-term results of a new treatment concept in rheumatoid arthritis patients. *Tissue Engineering* **2003**, *9* (5), 957–966.
18. Moimas, L.; Biosotto, M.; Di Lenarda, R.; Olivo, A.; Schmid, C. Rabbit pilot study on the resorbability of three-dimensional bioactive glass fibre scaffolds. *Acta Biomater.* **2006**, *2*, 191–199.



Metallic Orthopedic Devices: Failure Analysis

Lyle D. Zardiackas

Department of Biomedical Materials Science, University of Mississippi Medical Center, Jackson, Mississippi, U.S.A.

INTRODUCTION

Failure, as defined by Webster,^[1] is paraphrased as the omission or inability to perform a function. As the function of implanted metallic orthopedic and oral maxillofacial devices may or may not be to serve as definitive substitutes for bone or other hard tissues, this definition provides an interesting dilemma in relation to failed implants. Assessment of a defect is dependent on application rather than on fracture or deterioration. It is, therefore, imperative to carefully consider the function of the device when assessing success or failure.

Prediction of the physiological loads and the biological changes, which may occur, is not possible, and success or failure of the implant is often dependent on the patient. When assessing the adequacy of a device, especially a fracture fixation device, it is important, therefore, to correlate the "failure" of the device with patient record information. In establishing intended function, it is necessary to point out that these metallic implants may be separated into two categories: (i) Fracture fixation devices are used to anatomically align the bone to promote healing and (ii) Metallic prostheses are designed for long-term replacement of bone.

BIOLOGY, STRUCTURE, AND PROPERTIES OF BONE

Although it is true that all systems change with environment and time (expand, contract, corrode, change phase, etc.), none changes with such dynamics as does a living, growing system such as bone. In addition to the bone itself, the presence and interaction of soft tissues (muscles, ligaments, tendons, blood vessels, etc.) guide and often limit the implanted device in dimension, composition, and properties, and it is the patient in whom the device is implanted who often controls success and limits the restoration of function. Failure to consider these factors, can and will, in many cases, lead to patient problems such as infection, rejection, and delayed union or non-union owing to a variety of factors including stress shielding. These factors are far more profound considerations than fracture of the implant.

Bone is elastic, anisotropic, and hard and tough in compression and tension.^[2] Outer cortical bone is dense and compact and is often compared to reinforced concrete. The inner cancellous bone consists of trabeculae and lamellae, which are tiny slender spicules. Both cortical and cancellous bones are porous; however, the former is less porous. The elastic properties of bone can be defined by Hooke's Law (elastic behavior). A thorough discussion of bone on the cellular level may be found in many texts.^[3-5]

Many biological and physiological mechanisms are activated by bone fracture, but the actual nature of the stimulus for new bone formation following trauma is unknown. However, when a bone is broken or sectioned, blood from ruptured medullary, periosteal, and adjacent soft tissue vessels extravasates so that a hematoma or clot is formed around the fractured or sectioned location, periosteal and intraosseous osteoblasts are activated, and bone stem cells form new osteoblasts. Within a short period of time, a callus, which includes cartilage and fibrocartilage, is formed between the two ends. In the case of a fracture or osteotomy, the osteoblastic activity is accelerated by mechanical internal fixation, which causes compression and tension of cortical bone.^[6] If bone receives an insufficient amount of stress or strain (stress or strain shielding), bone remodeling is hindered.^[7-8] Mineralization and demineralization, which are controlled by nutrition, activity, hormones, patient physiology, and rest, affect bone growth.

IMPLANTS FOR FRACTURE FIXATION AND JOINT REPLACEMENT

As previously stated, implants for fixation are designed to provide stabilization during healing. As the sites and patient needs are highly variable, great versatility and adaptability of these implants are required. Various sizes and types of plates, nails, screws, wires, etc., which may be further customized during surgery to meet specific anatomical conditions, have been developed; but this array of devices cannot be infinite.

Metallic prostheses are most often designed to be replacements for joints and are often thought of as permanent. They should not, however, be considered

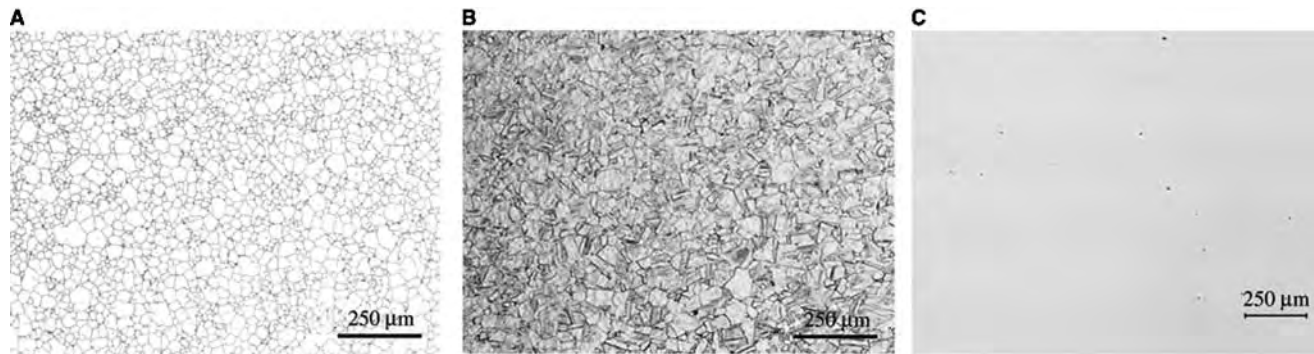


Fig. 1 (A) Microstructure of annealed 316 L stainless steel (100x), (B) microstructure of cold worked 316 L stainless steel (100x), and (C) globular oxide (Type D) and sulfide stringer (Type A) inclusions in 316 L stainless steel (100x).

permanent but rather as definitive and will deteriorate over time because they are subject to the same physiological environment, which caused failure of the original biological joint. It is recognized that implant degradation begins once it is placed or, at least, as soon as it is loaded. Various sizes, types, and material combinations are used in the manufacture of prostheses such as hips, knees, shoulder, and temporomandibular joints. Unlike the situation with fracture fixation devices, there has been success in providing custom prostheses primarily because joint replacement is anticipated before surgery and is patient specific.

Complications arising from any one or more of the following factors may contribute to or cause fracture prior to healing of the bone:

1. Loosening of the implant owing to bone resorption.
2. Delayed union or non-union owing to biological considerations.
3. Secondary trauma at or near the original fracture site.
4. Overloading of the implant by the patient.
5. Inadequate reduction or alignment.
6. Improper implant handling during placement.

7. Implant degradation owing to biological considerations.
8. Implant wear.

Great care has been taken here to elucidate these concepts because of the use or misuse of the term “implant failure.” This terminology is used by those who are knowledgeable in the field to describe the failure of an implant to achieve its intended function owing to inadequacy of materials, design, or manufacturing defects that render the device unsuitable. Material defect levels must be compared with values defined by the appropriate specifications for the specific material by which the implant is made. The existence of manufacturing defects must be determined based upon limitations imposed by the host and surgical limitations. Additionally, the appropriate manufacturing documents for the specific device, as well as upon ASTM International^[2] and International Standards Organization (ISO) specifications, where appropriate must be examined.

Design defects are much more difficult to determine because the device must be examined based upon research including in vitro evaluation, and, most importantly, use by patients. Input from experts in

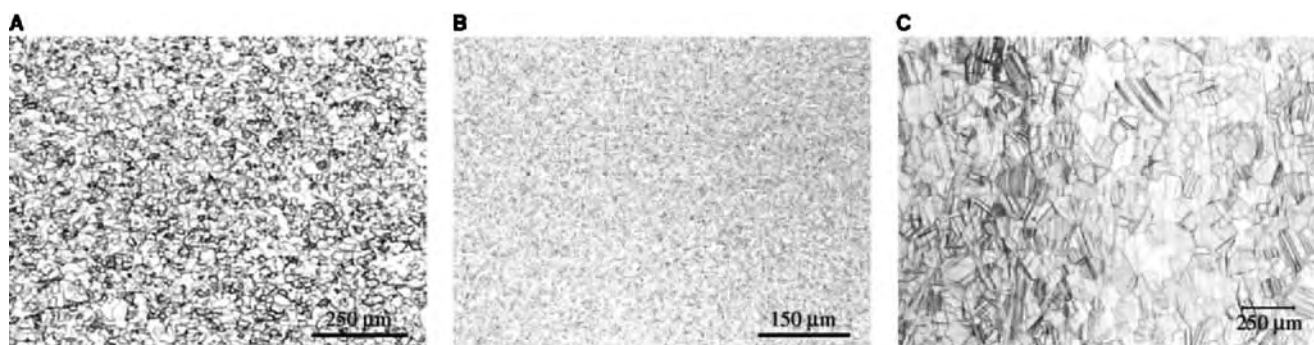


Fig. 2 (A) Grain structure of annealed Grade 4 CP Ti (100x); (B) grain structure of annealed Ti-6Al-4V alloy (200x); and (C) grain structure of annealed β -Ti15Mo alloy (100x).

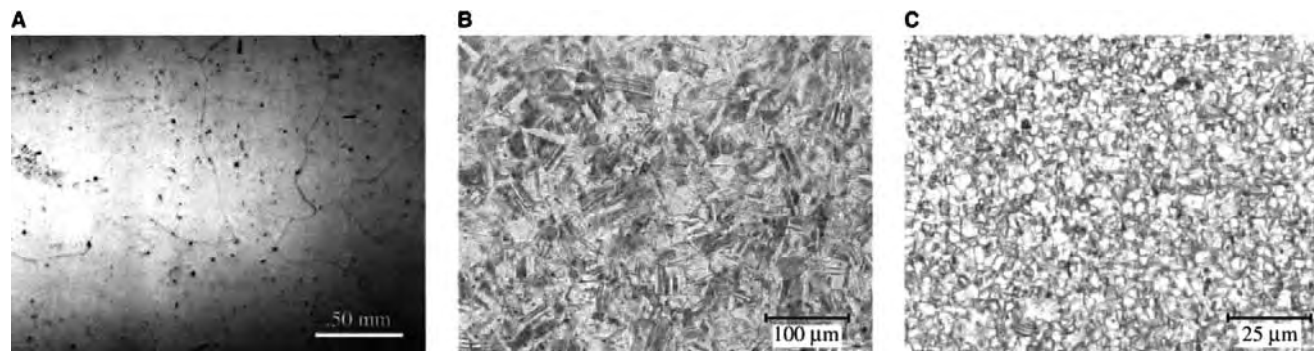


Fig. 3 (A) Grain structure of cast Co–Cr ASTM F75 alloy (50x); (B) grain structure of wrought Co–Cr ASTM F90 alloy (100x); and (C) grain structure of wrought Co–Cr ASTM F1537 alloy (500x).

biomaterials, biomechanics, and certainly from surgeons is required to evaluate the suitability of a device for its intended application. It must always be recognized that fracture of a device does not necessarily mean failure or even inadequate design, faulty material, or surgical error, and it is always a race between implant failure and bone healing. More often than not, repeated mechanical stresses on the implant beyond its design capabilities result in fracture.

The reasons for delayed union and non-union of bone are varied and numerous, and are generally not in the control of the surgeon and implant or device manufacturer; and patient non-compliance is often not easily discernible. In the case of failed prostheses, bone resorption, cement fracture, and fracture and/or wear of polymeric materials, which constitute one side of a bearing surface are often responsible for, or at least contribute to, failure of the implant.

Metals and Alloys

A number of metallic materials are currently accepted for use in implants and these are considered in depth elsewhere in this text. These materials have been found to be appropriate based upon their physical,

mechanical, electrochemical, and biological properties. Appropriate determination of such characteristics as the composition, grain structure and size, inclusions, foreign phases, mechanical properties, and corrosion susceptibility is key to the evaluation of the failure of an implant

The stainless steel used for fracture fixation devices is implant quality austenitic stainless steel supplied in the annealed or work-hardened condition (Fig. 1A and B). Examination of the microstructure for grain boundary precipitates, grain structure, and grain size is necessary, as these parameters are indicative of the acceptability of the material. Inclusion levels are rated in the longitudinal direction owing to elongation during cold working and the must be below certain concentration levels (Fig. 1C). A relatively new low nickel/high nitrogen-containing alloy has a number of interesting properties including higher strength, greater ductility, and longer fatigue life in the physiological environment.

Titanium (Ti) and Ti alloys constitute the second major metallic material used for this application. α -Ti exists at equilibrium at room temperature and the structure is transformed at 883°C from hexagonal close packed (hcp) to body-centered cubic (bcc) β -Ti. Commercially pure (CP) Ti (α) is supplied in four grades. As the grade number increases from 1 to 4, the

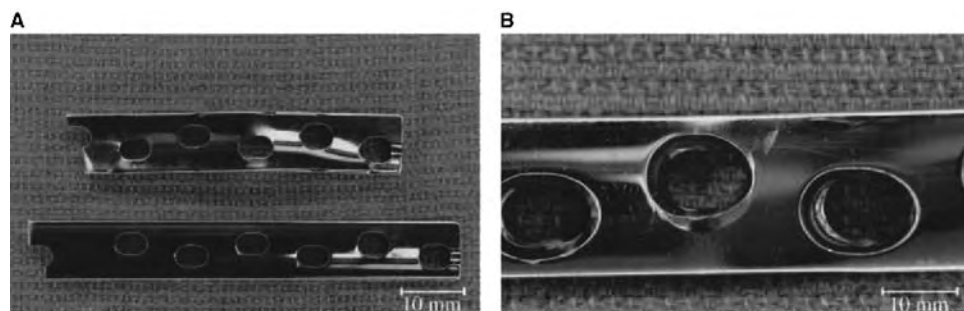


Fig. 4 (A) Photograph of DC plate after retrieval and (B) photograph showing fretting in screw holes of DC plate.

Table 1 Quantitative ICP analysis of dynamic compression plate

Concentration (dilution calculation)											
No.	Cr %	Mn %	Fe [cps]	Mo %	Ar [cps]	Ni %	Si %	S %	P %	Cu %	C %
1	17.82	1.759	32,470	2.928	3991	14.99	0.3576	0.0041	0.0135	0.0828	0.0127
2	17.51	1.757	36,289	2.770	4334	14.92	0.3414	0.0033	0.0136	0.0681	0.0111
3	17.64	1.719	34,804	2.636	4381	14.82	0.3571	0.0037	0.0138	0.0698	0.0109
(n = 3)											
x	17.65	1.745	34,521	2.778	4236	14.91	0.3520	0.0037	0.0136	0.0736	0.0115
s	0.1530	0.0223	1925	0.1464	212.9	0.0838	0.0092	0.0004	0.0002	0.0081	0.0009
sr	0.867	1.277	5.577	5.270	5.028	0.562	2.615	9.952	1.343	10.975	8.132

oxygen concentration increases, and, therefore, the strength increases. The strength of this material may also be appreciably increased by work hardening. The grain structure of CP Ti in the annealed condition is equiaxed (Fig. 2A). The α/β Ti alloys exhibit a two-phase microstructure and have increased strength as compared to CP Ti (Fig. 2B). Over the last several years, several β -Ti alloys have been introduced (Fig. 2C). These alloys have good strength and corrosion resistance with increase ductility as compared to the α/β Ti alloys.

Metals with Co and Cr as the primary alloying elements constitute the final major metallic system. These alloys are not used for fixations devices but are used for prostheses owing to their high elastic modulus, high strength (tensile and yield), high rate of work hardening with a commensurate reduction in ductility, as well as other factors. Of primary importance in their use for a fixation device is an increased probability of stress shielding of bone resulting in delayed or non-union. These alloys also form single-phase microstructures. The grains in the as-cast condition are very large, and only a few grains may be visible in a transverse section of a

casting (Fig. 3A). The structure of wrought alloys is generally equiaxed with much smaller grains than either the cast alloy or the other wrought alloy (Fig. 3B). Over the last several years, alloys with higher strength and hardness have been developed by conventional and powder metallurgical techniques. The microstructure of these alloys may have a small grain size and very fine, segregated carbides (Fig. 3C).

In addition to compositional analysis and metallurgical analysis to assure compliance with the appropriate specifications, evaluation of compliance with manufacturing documents is required. Evaluation of dimensional specifications is often impossible as the retrieved implant may have areas of corrosion, wear, or deformation caused during implantation, retrieval, handling, and/or while in the patient.

FAILURE ANALYSIS OF FRACTURE FIXATION DEVICES

It is necessary to ensure that retrieved implants are properly sterilized prior to handling, and it is often

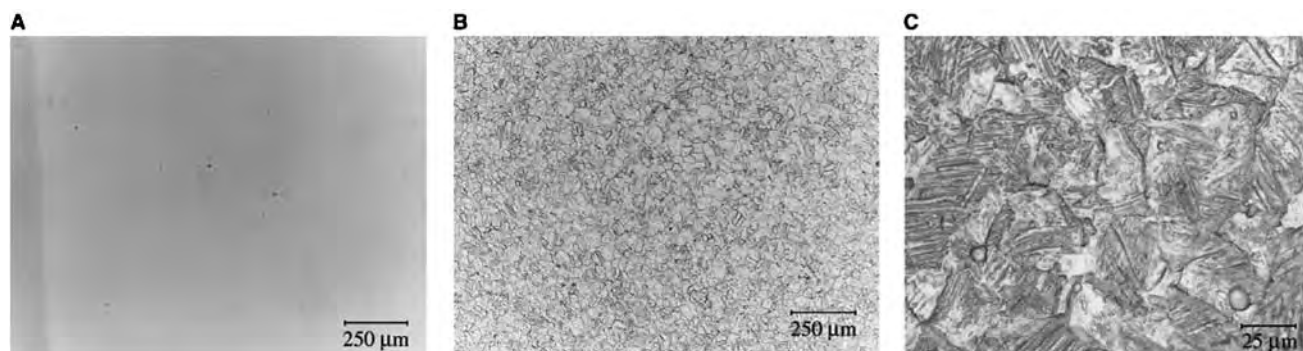


Fig. 5 (A) Light micrograph showing longitudinal inclusions in DC plate of 316L stainless steel (100x); (B) light micrograph showing grain structure and ASTM grain size of 8.0 (100x); and (C) light micrograph showing etch pits and discontinuous grain boundary ditching in 316L stainless steel after screening test according to ASTM A262.



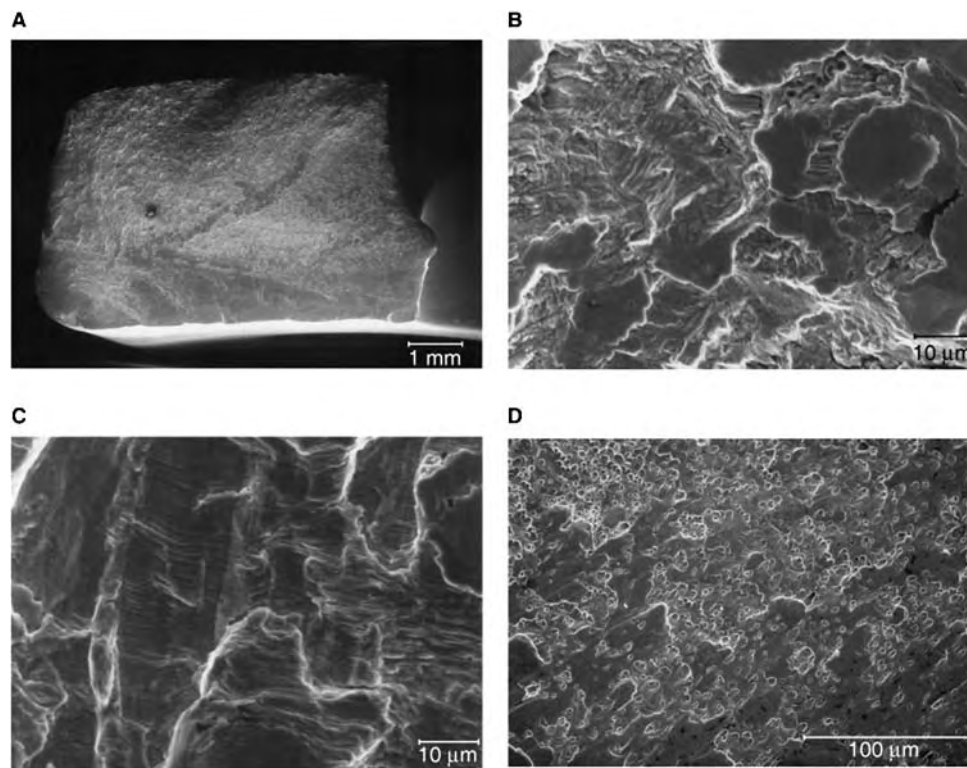


Fig. 6 (A) SEM micrograph of fracture surface of DC plate on left side of screw hole—crack initiation at upper right; (B) SEM micrograph showing wear, secondary cracks, and striations in area of crack initiation; (C) SEM micrograph showing fine striations and secondary fatigue cracks; and (D) SEM micrograph showing dimples in area of final overload fracture (Stage III fatigue).

necessary to perform a thorough cleaning of the fracture surfaces to permit complete examination. Cleaning is generally accomplished using alcohol, acetone, or mild detergent solution in an ultrasonic cleaner and/or by the use of an enzyme if the amount of tissue is excessive.

While light microscopy is useful for the examination of these implants, it is limited by magnification and depth of field. Additionally, light microscopy is most often used for metallurgical analysis for evaluation of microstructure both prior to and after etching or

staining. Scanning electron microscopy (SEM) is used to examine areas such as fracture surfaces owing to increased depth of field at higher magnification. Often, only one of two opposing fracture surfaces needs to be examined because they are mirror images; however, if the morphological characteristics have been obscured or obliterated, examination of the opposing fracture surface is required.

In addition to examination of the implant itself, manufacturing documents, internal and appropriate

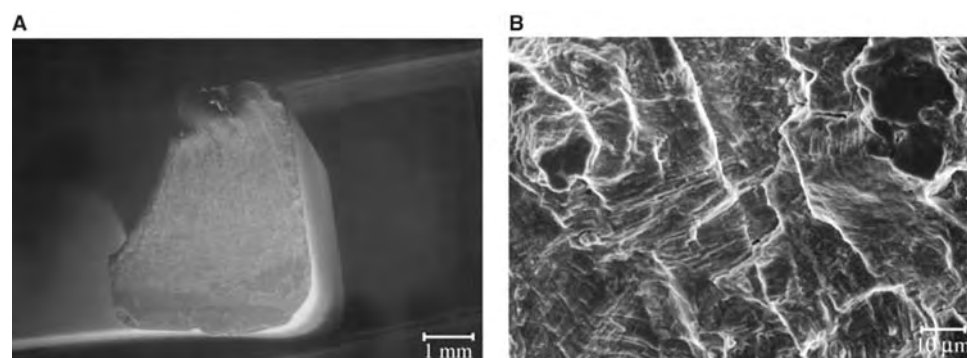


Fig. 7 (A) SEM micrograph of fracture surface of DC plate on right side of screw hole—crack initiation at upper right and (B) SEM micrograph showing striations over fracture surface in Fig. 7A.

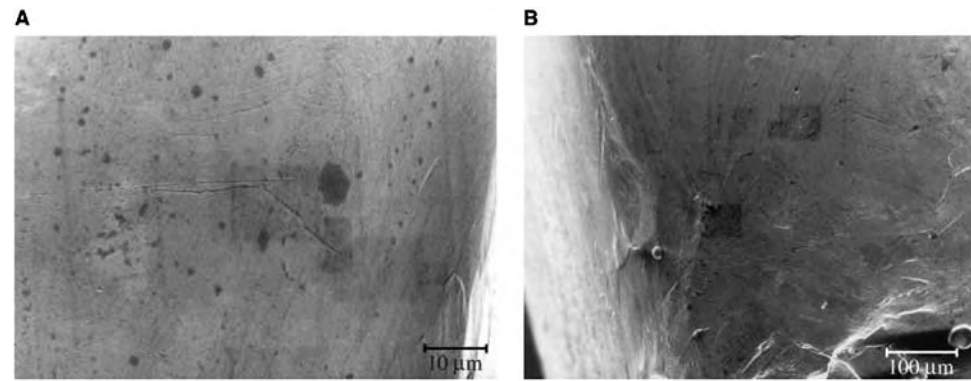


Fig. 8 (A) SEM micrograph showing cracks on surface near area of crack initiation and (B) SEM micrograph showing multiple slip bands near screw hole.

standard specifications, patient records, and X-rays should be examined to correlate with the results of direct observations of the implant. There may be situations where all records are not available, but the structure and properties of the metal as well as the morphology of the fracture surface are often sufficient to determine the cause of failure.

Failure Analysis of Dynamic Compression Plate

Visual examination of the plate revealed that fracture had occurred transversely through the sixth screw hole

from one end of the plate (Fig. 4A). Visual and light microscopic examination of the pieces of the dynamic compression (DC) plate revealed no abnormal markings other than those that appear to have been caused by use. Examination of the screw holes showed minor-to-moderate fretting and/or fretting corrosion typical for a retrieved fixation device of this type (Fig. 4B).

To fully examine the fracture surfaces in the SEM and to provide a piece for metallurgical analysis, the plate was sectioned using an aluminum oxide blade on a high-speed metallurgical saw. The pieces were

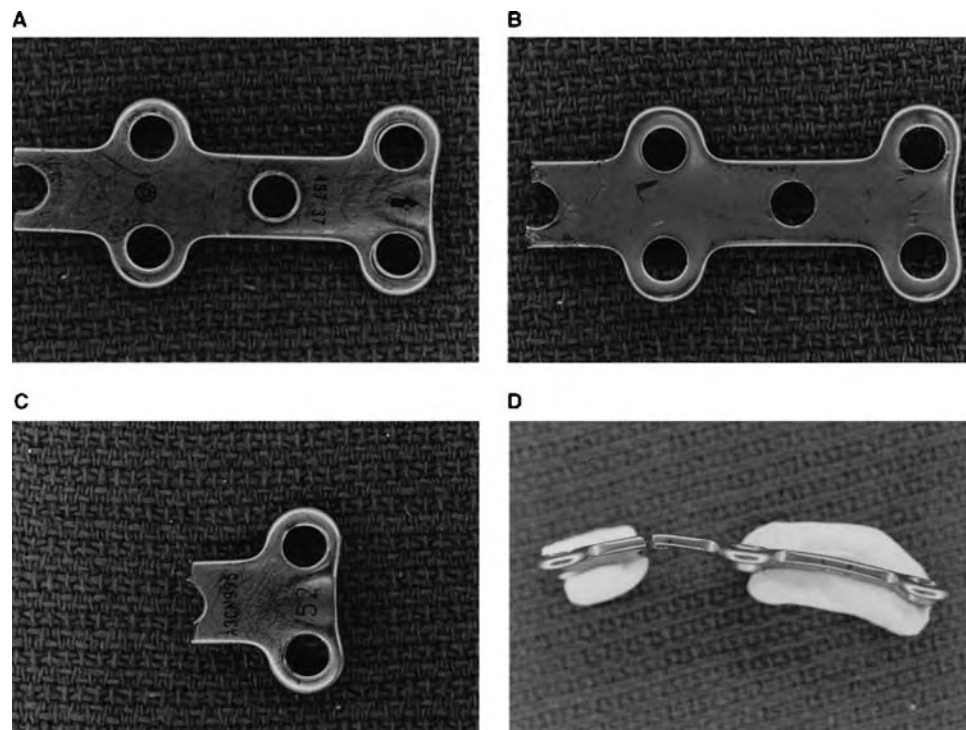


Fig. 9 (A) Photograph of anterior cervical spine plate; (B) photograph of cephalad end anterior cervical spine plate showing wear near fracture surfaces; (C) photograph of caudal end anterior cervical spine plate showing wear near fracture surfaces; and (D) photograph of anterior cervical spine plate showing contour in plate.



Fig. 10 X-ray of anterior cervical spine plate after placement.

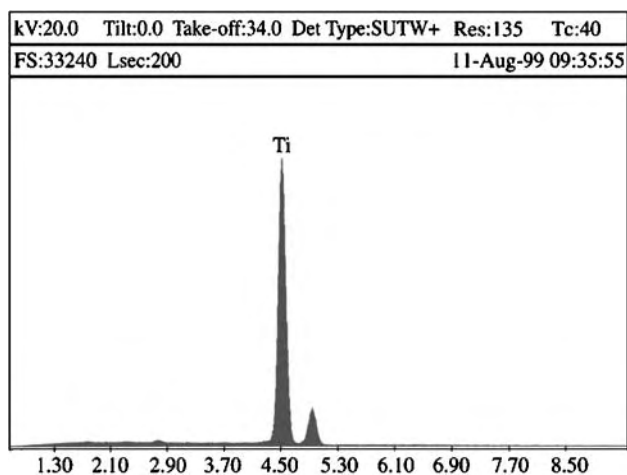


Fig. 11 Qualitative energy-dispersive spectroscopy of cervical anterior spine plate.

cleaned in an ultrasonic cleaner for approximately 15 min using a solution of Alconox in warm water, rinsed in water and ethanol, and dried with forced air. After instrument calibration, the small-sectioned piece was placed in the spark stand of the inductively coupled plasma (ICP) spectrometer for quantitative analysis. Three samplings were taken to determine the concentrations of the critical elements and they were found to be in compliance with ASTM and ISO specifications (Table 1).

After compositional analysis, the piece was mounted longitudinally in heat-cure transparent acrylic; the surface was polished to a final finish using 0.05 μm alumina and examined for non-metallic inclusions. Only globular oxide (Type D) inclusions were noted and rated according to ASTM E45 at a level of 1.0 thin/0.0 heavy (Fig. 5A) The sample was etched with a solution of CuCl_2 in HCl and ethanol and examined for the presence of δ -ferrite. No δ -ferrite was observed at 100x magnification.

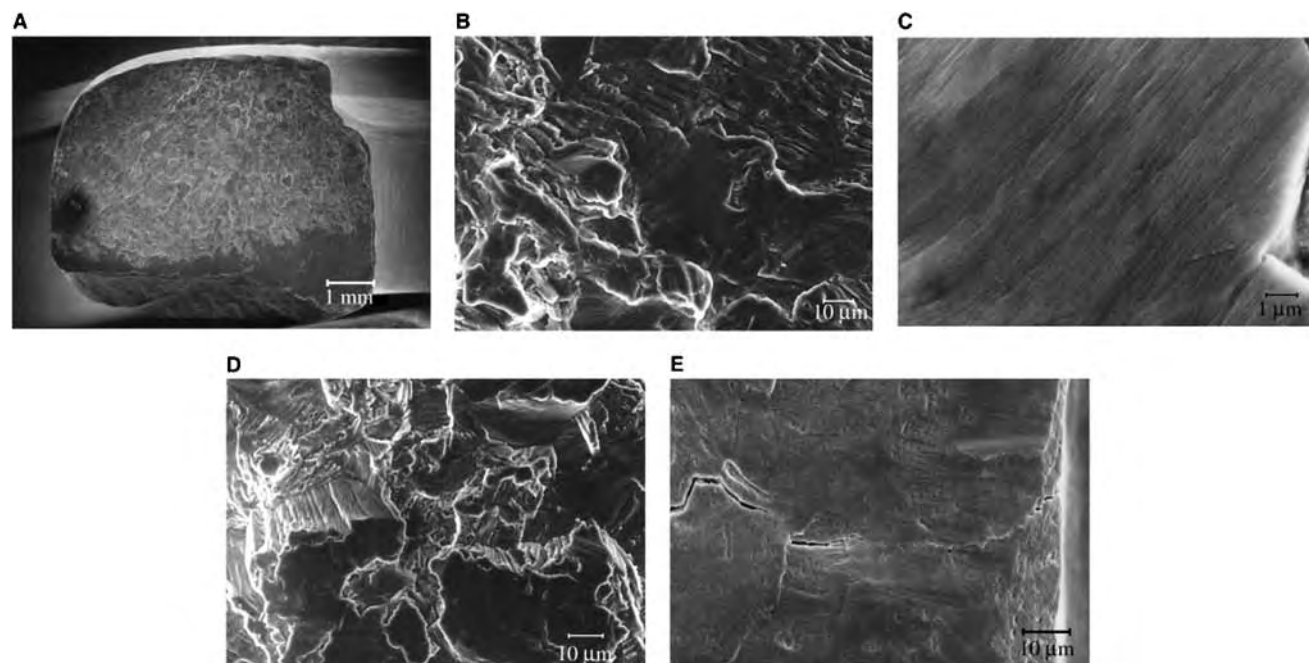


Fig. 12 (A) SEM micrograph showing fracture surface left of screw hole on caudal section in Fig. 9C; (B) SEM micrograph of fracture surface in area of crack initiation at upper/right in Fig. 12A; (C) higher magnification of fracture surface in area of crack initiation showing fine striations; (D) higher magnification of fracture surface showing striations, wear, flutes and terraces; and (E) SEM micrograph showing cracking on anterior surface near fracture surface and screw hole.

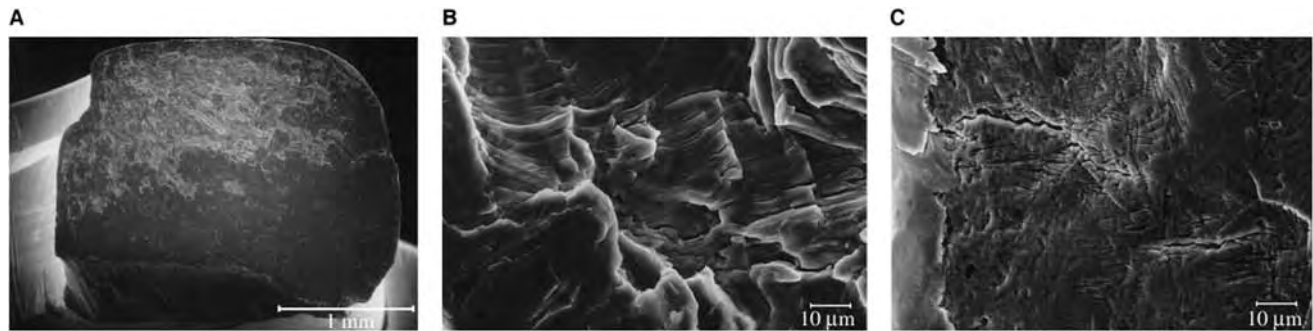


Fig. 13 (A) SEM micrograph of fracture surface right of screw hole on caudal section; (B) higher magnification SEM micrograph of area of crack initiation at upper/left in Fig. 13A; and (C) SEM micrograph showing slip bands and cracks on anterior surface near fracture surface in Fig. 13A.

After the evaluation for δ -ferrite, the sample remounted transversely and polished to a final finish with $0.05\ \mu\text{m}$ alumina. Five^[5] Vickers microhardness (H_v) measurements were made under a 500 g load for 12 sec along the approximate midline toward each edge, at the approximate center, and at the approximate midpoints between the center and edges. These measurements resulted in a mean H_v value of 331.6, which is within the range expected for work-hardened 316 L stainless steel. The sample was etched using a solution of FeCl_3 in HCl and HNO_3 , and the ASTM grain size was determined according to ASTM E-112. The grain structure was free of carbide precipitates at the grain boundaries, the structure was twinned owing to cold working, and the ASTM grain size was determined to be 8.0 (Fig. 5B), which is within ASTM and ISO specifications. Since examination of the surface, the screw holes and the microstructure showed no signs of excess corrosion or a suspect microstructure, evaluation of intergranular corrosion was deemed unnecessary; however, should corrosion be suspect, the test for intergranular corrosion (ASTM A262) should be performed. An example of an acceptable microstructure for 316 L stainless steel after the screening test for intergranular corrosion as described in ASTM A262 is seen in Fig. 5C.

Fig. 6A is an SEM micrograph of the fracture surface on the left side of the screw hole. The chamfer for the screw head is oriented toward the top/right of the micrograph and chevron marks and clamshell marks are visible. The orientation of these marks shows that the main crack front, which resulted in fracture on this side of the screw hole, was initiated at the upper/right near the intersection of the screw hole and lateral side of the plate and the crack propagated toward the lower/left. The morphology near the site of crack initiation, seen at higher magnification in Fig. 6B, is typical of that seen over most of the fracture surface. Worn and smeared metal, secondary cracking, as well as striations, indicative of crack propagation

owing to Stage II fatigue, were noted. Many of these striations were very closely spaced ($<0.5\ \mu\text{m}$) indicating high-cycle fatigue, but the varied striation spacing indicates a variable load spectrum with both high- and low-cycle Stage II fatigue crack propagation resulting (Fig. 6C). Shear dimples, indicative of Stage III fatigue (overload), were observed at the lower left (Fig. 6D). This morphology indicates that final overload fracture on this side of the screw hole occurred in this area.

The fracture surface on the right side of the screw hole shows orientation of the chevron marks, and clamshell marks, which indicate crack initiation at the upper/left and propagation toward the lower/right (Fig. 7A). Fig. 7B shows an area at the approximate center of the fracture surface where the morphology is also indicative of Stage II fatigue under a variable load spectrum. Fig. 8A and B show the presence of additional cracks and the morphology of the crack interfaces as well as the presence of visible slip bands at and near the screw hole near the fracture surface. The morphology indicates that these cracks had formed after manufacturing and while in the patient.

Examination of the X-rays showed the position of the fracture in the bone, the condition of the bone



Fig. 14 Photograph of fractured screws.

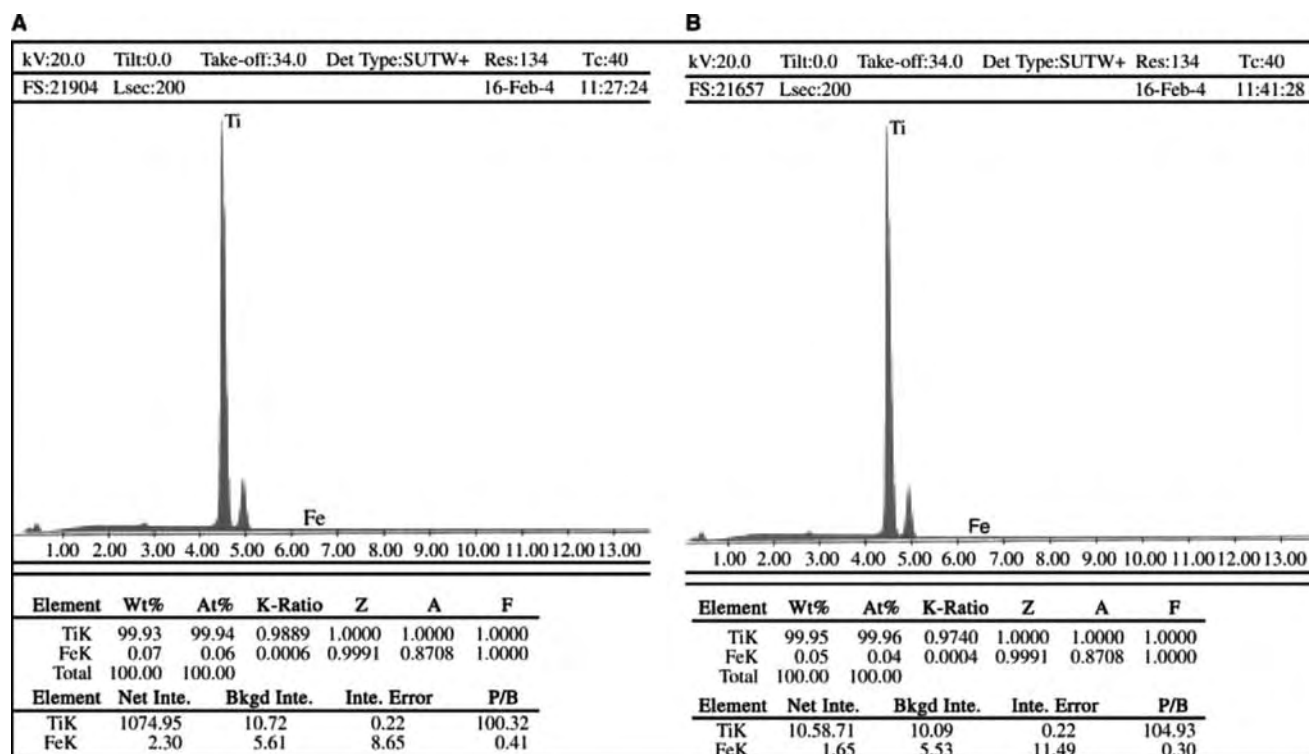


Fig. 15 (A) Quantitative energy-dispersive spectroscopy of screw A and (B) quantitative energy-dispersive spectroscopy of screw B.

around the fracture site, and the relative position of the fracture of the plate in relation to the fracture of the bone. As further examination of X-rays and the patient record indicated a non-union, the implant eventually failed. Additionally, based upon the preceding tests and evaluations, it is concluded that the plate was produced in accordance with the appropriate metallurgical requirements and the appropriate manufacturing documents. It is concluded that the primary reason for fracture of this plate was fatigue owing to repetitive cyclical stressing by the patient, and that fracture was not related to deficiencies in the plate.

Failure Analysis of Anterior Cervical Spine Plate

Visual and light microscopic examination at up to 40x magnification revealed multiple nicks and scratches through the anodized surfaces, no generalized surface or fretting corrosion indicating occurrence after manufacturing (Fig. 9A). Wear of the posterior surface on the cephalad piece and the anterior surface on the caudal piece indicated that the cephalad section had become displaced over the caudal section after fracture (Fig. 9B and C). Additional examination of the plate and X-rays indicated that it had been contoured during

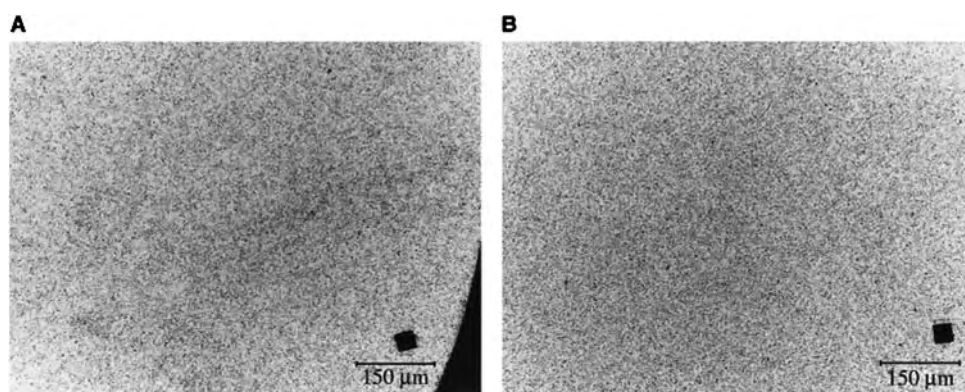


Fig. 16 (A) Microstructure of screw A and (B) microstructure of screw B.

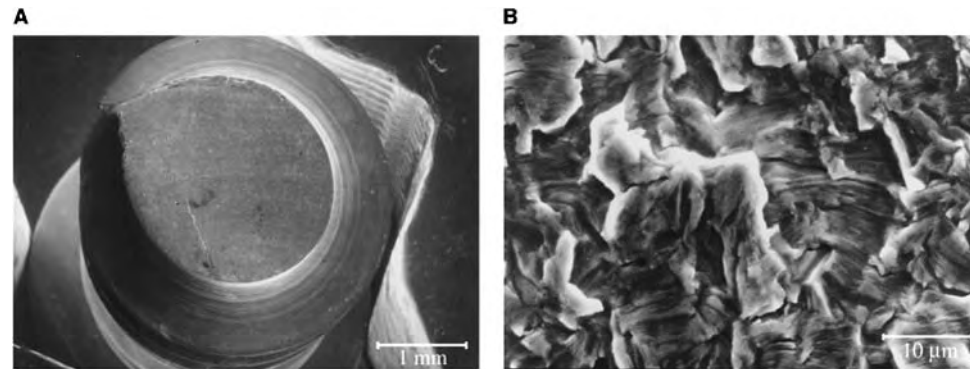


Fig. 17 (A) SEM micrograph showing fracture surface of screw A and orientation of clamshell marks and shear lip and (B) higher magnification SEM micrograph showing striations secondary cracking and wear.

placement (Figs. 9 and 10). Prior to examination by SEM, the pieces of the plate were ultrasonically cleaned in a mild detergent solution to remove biological debris, rinsed in distilled water and ethanol, and air-dried. The anterior aspect of the cephalad section was qualitatively analyzed for major elemental composition by energy-dispersive spectrometry. The spectrum, seen in Fig. 11, shows that the only major element detected was Ti indicating that the material is CP Ti.

Fig. 12A shows the fracture surface to the left of the screw hole, the cephalad piece. Much of the fracture surface toward the posterior aspect of the plate as well as the posterior surface itself was highly worn. Wear indicates that the opposing fracture surfaces and the anterior/posterior aspects of the plate rubbed against each other during and/or after the crack had formed. The orientation of the chevron marks indicates that crack initiation was at the intersection of the screw hole and the anterior surface of the plate (upper/right Fig. 12A). The area of crack initiation is seen at higher magnification in Fig. 12B. Fine and coarse striations and wear were noted. The very fine nature of the striations ($<1\mu\text{m}$) seen at higher magnification in

Fig. 12C is indicative of high-cycle Stage II fatigue crack propagation, while coarser striations indicate higher loading during crack propagation. The area at the bottom/left in Fig. 12A showed striations, wear, and indications of tear structures often seen in fatigue of Ti (Fig. 12D). Areas near the fracture surface (Fig. 12E) show the presence of slip bands and the presence of additional cracking near the intersection of the fracture surface and both the anterior and posterior aspect.

The fracture surface to the right of the screw hole is seen in Fig. 13A. The orientation of the chevron marks, while somewhat obscured, indicates that fracture is initiated at the upper/left near the intersection of the screw hole and anterior surface of the plate. The area at the top/left of Fig. 13A, near the site of crack initiation, is seen at higher magnification in Fig. 13B. This area, as well as all other areas of this fracture surface showed striations indicative of crack propagation owing to Stage II fatigue. As with the anterior aspect on the left side of the screw hole, the presence of slip bands, some of which were opening to form additional cracks, indicates stressing of the plate after manufacturing (Fig. 13C).

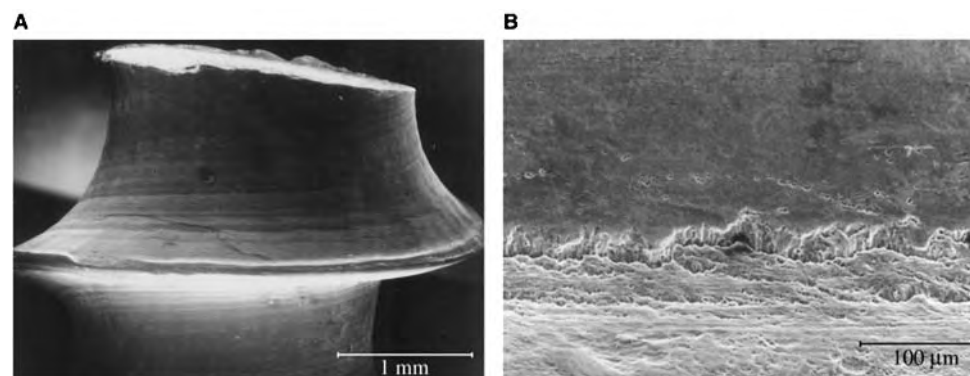


Fig. 18 (A) Cracking in screw thread below fracture surface and (B) higher magnification of crack showing crack interface.

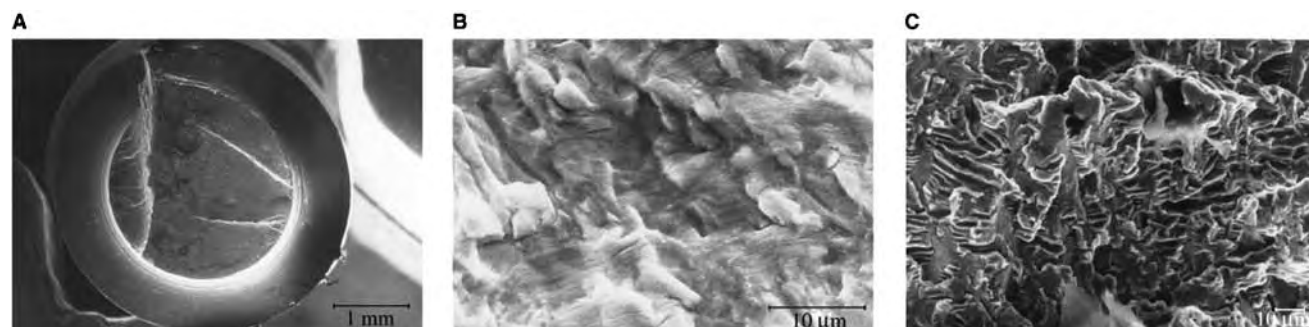


Fig. 19 (A) SEM micrograph showing fracture surface of screw B—note chevron marks; (B) higher magnification SEM micrograph showing typical fracture surface morphology; and (C) higher magnification SEM micrograph showing flutes and terraces with feather marks.

Failure Analysis of Ti Screws

The two broken screws, which were labeled A and B for identification purposes, are seen in Fig. 14. Light microscopic examination at up to 40x magnification revealed no generalized surface or fretting corrosion, and no manufacturing defects. After photographing the implants, they were rinsed in ethanol and blown dry with forced air. The heads of each screw were removed using a diamond wafering blade, and the cut surfaces of the head sections were polished to a 9 μm diamond finish and quantitatively analyzed for major alloying elements using energy-dispersive spectrometer, calibrated using a standard of CP Ti. The results as seen in Fig. 15A and B indicate that the primary element is Ti with a minor amount of iron (Fe), which is within ASTM specifications for Grade 4 CP Ti.

After compositional analysis, the heads of the screws were mounted, and polished to a final finish with colloidal silica. H_v measurements were made in three areas giving a mean value of 283.2 for screw A

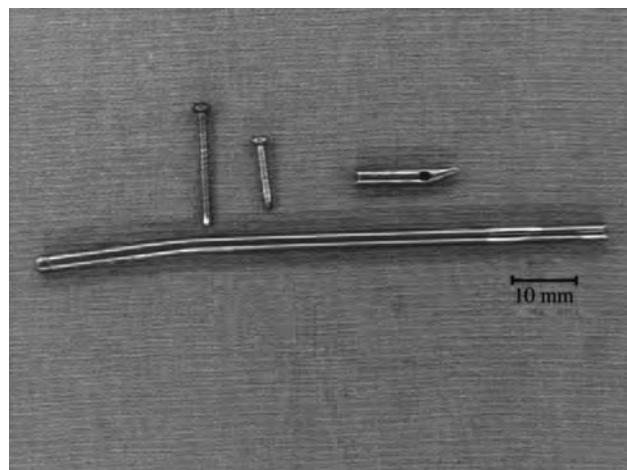


Fig. 20 Photograph of fractured intramedullary (IM) nail.

and a mean value of 287.2 for screw B. These values are within the range expected for cold-worked Grade 4 CPTi. After hardness testing, the samples were etched revealing a single-phase α -microstructure with an ASTM grain size for screw A of 9.5 and an ASTM grain size between 9.5 and 10.0 for screw B (Fig. 16A and B).

The fracture surfaces and areas near the fracture surfaces of both screws were examined using SEM. Fig. 17A shows the orientation of the clamshell marks on screw A, which radiate from the area of

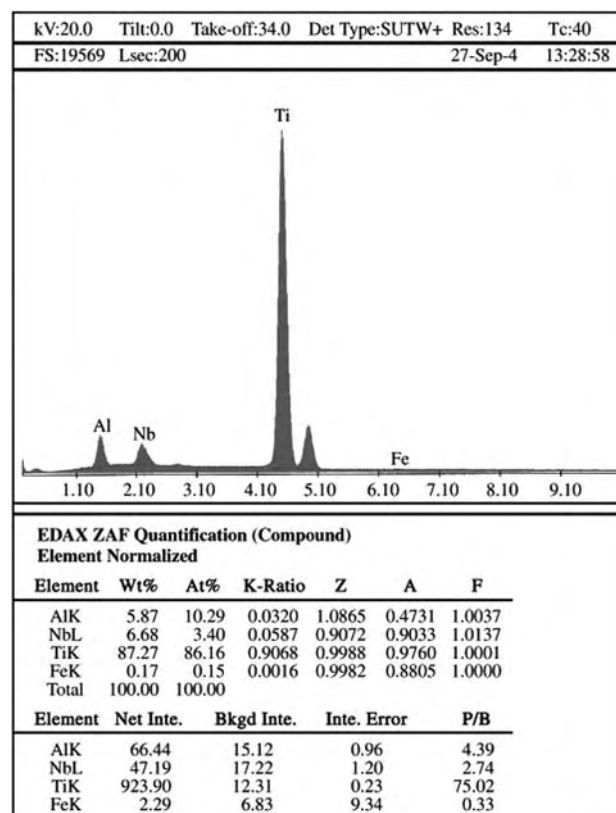


Fig. 21 Quantitative energy-dispersive spectroscopy of IM nail.

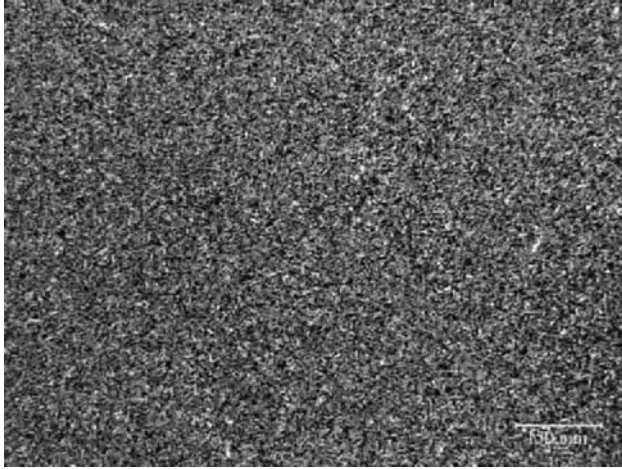


Fig. 22 Light micrograph showing α/β grain structure of IM nail.

initiation of the major crack front as marked with an arrow. Essentially all areas of the fracture surface, left and right of the ridge in Fig. 17A, as well as at the approximate center and at the top of the micrograph showed fine striations, secondary cracking and wear, typically seen in high-cycle Stage II fatigue crack propagation caused by patient motion (Fig. 17). Fig. 17A also shows a shear lip at the upper periphery, which is indicative of final fracture owing to overload.

The morphology of the crack in the thread immediately below the fracture surface in the region of crack initiation (Fig. 18A) as well as the crack near the head of the screw (Fig. 18B) indicates the application of significant loads after manufacturing, which led to the formation of additional cracks.

The orientation of the chevron marks, radiating from both the left and right sides of the fracture surface of screw B (Fig. 19A), indicates that two cracks were propagating and met at the ledge running from top to bottom (between the arrows). Examination of the areas at both the left and the right showed

striations, secondary cracking, and wear typical in crack propagation owing to high-cycle Stage II fatigue (Fig. 19B). Fig. 19C is typical of the morphology seen on the ledge joining the two fracture surfaces with flutes, and terraces with feather marks. This morphology is typical in fatigue fracture under a high stress intensity and/or shear of Ti. Results of examination of areas near the fracture surface and in the threads revealed no additional cracking of this screw.

These results show that the material is CP Ti with an Fe concentration well within the limits of ASTM F67, and hardness values are within the range expected for implant quality cold-worked Grade 4 CP Ti. The grain structure and grain size are within the range expected for this material. As expected, crack initiation was at or near the root of the threads and propagation was primarily high-cycle fatigue for both of the screws.

M

Failure Analysis of Intramedullary Nail

Initial visual and light microscopic examination was performed and the implant pieces were photographed (Fig. 20). A piece containing the fracture surfaces was cut from the proximal section and an additional piece was cut adjacent to the piece containing these fracture surfaces. There was significant amount of wear, which had removed the anodized layer from much of the surface of the nail and there were additional marks and gouges through the anodized surface indicating that these marks occurred after manufacturing, and no surface corrosion was observed. Examination of the fracture surfaces using light microscopy showed that the fracture surfaces were highly worn owing to rubbing against one another during and/or after fracture of the nail.

After light microscopy, fracture surfaces were ultrasonically cleaned, using a detergent in warm water, rinsed in water and ethanol, and blown dry with forced air. After polishing one of the cut surfaces with

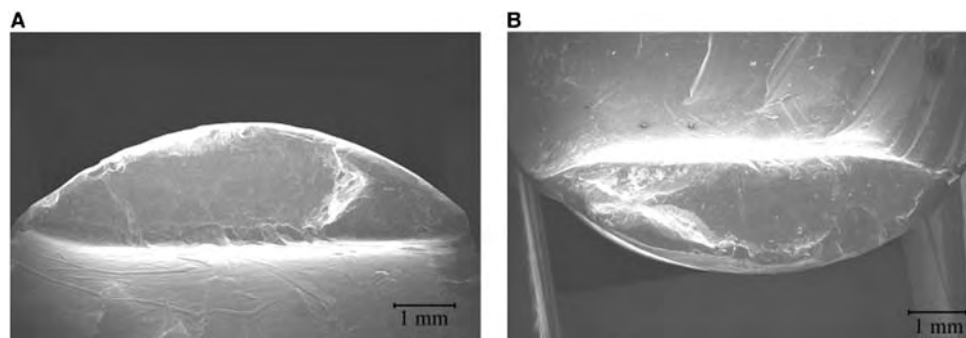


Fig. 23 (A) SEM micrograph of distal fracture surface on one side of screw hole and (B) SEM micrograph of distal fracture on side of screw hole opposite Fig. 23A.

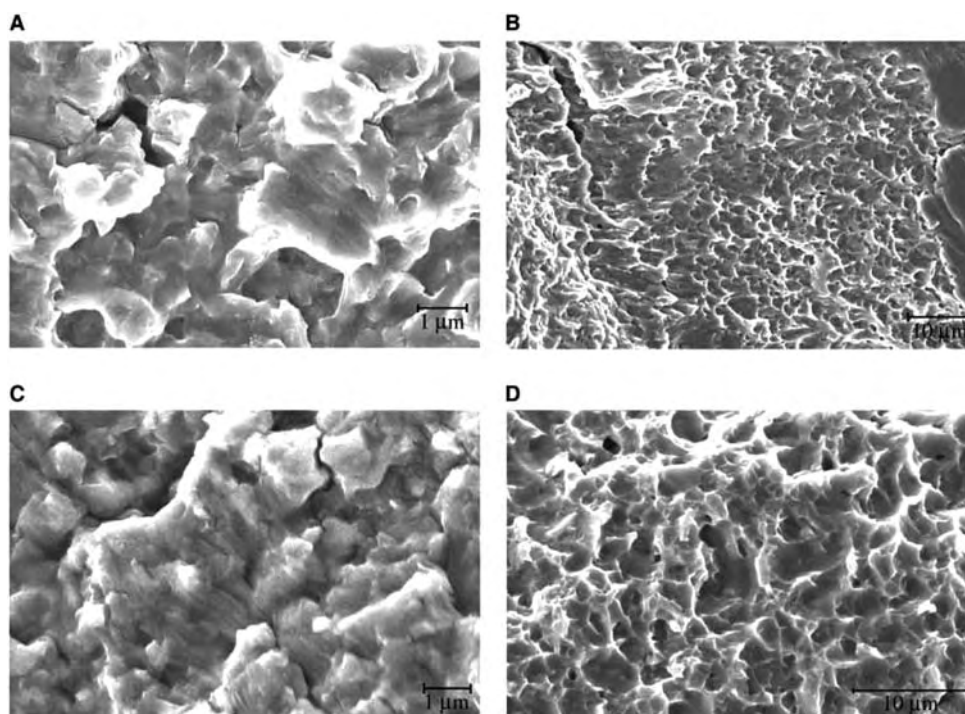


Fig. 24 (A) SEM micrograph showing fine fatigue striations and secondary cracks; (B) SEM micrograph of area between crack front in Fig. 24A showing orientation of shear dimples; (C) SEM micrograph showing fatigue striations and secondary cracks on proximal fracture surface; and (D) SEM micrograph showing dimples indicative of ductile tensile overload fracture.

9 μm diamond, quantitative analysis of the major alloying elements was performed using a Ti–6Al–7Nb standard. The spectrum and results, as seen in Fig. 21, show that the major elements were Ti, Al, and Nb within the levels required by ASTM and ISO specifications.

After compositional analysis, the sample was mounted and polished to a final finish with colloidal silica. Five H_v measurements were made along the diameter of the sample resulting in a mean value of 340.0, which is within the range expected for this alloy. The polished surface was etched with Kroll's reagent to reveal an α/β microstructure, with no α -case in compliance with ASTM F1295 (Fig. 22).

The fracture surfaces on both sides of the screw hole on both the distal and the proximal sections were

examined because they were all extremely worn owing to rubbing after fracture. Fig. 23A and B show fracture surfaces on both sides of the screw hole. Both clamshell marks and chevron marks indicate two different cracks initiating at the intersection of the screw hole and nail surface indicating reverse loading. Fine striations ($<1\ \mu\text{m}$) and secondary cracking indicative of crack propagation owing to high-cycle Stage II fatigue were observed (Fig. 24A). This morphology is typical of that seen over most of this fracture surface, except where the surface is highly worn and the area of overload fracture as seen in Fig. 24B. These obliterated dimples are indicative of shearing stress between the two primary crack fronts.

No further information was gleaned from examination of the fracture surface on the other side of the screw hole owing to wear; however, the opposing

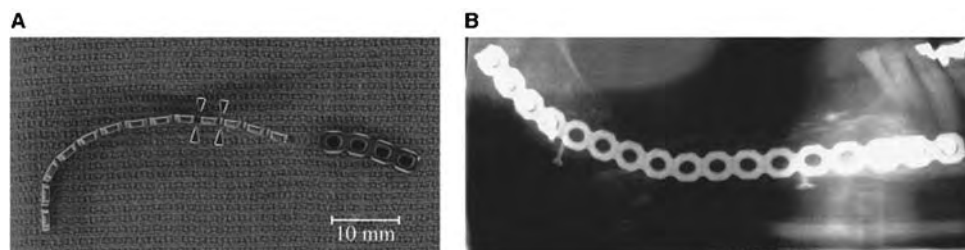


Fig. 25 (A) Photograph of fractured mandibular plate and (B) X-ray showing plate in situ.

proximal fracture surface was not as significantly worn, fine fatigue striations ($<1\ \mu\text{m}$) and secondary cracking were observed (Fig. 24C). As with the fracture surface on the other side of the screw hole, final fracture was owing to ductile overload (Fig. 24D).

Review of the patient record summary, device history record, material certification documents, and X-rays indicated that the bone fracture was highly comminuted and had occurred in the central to distal third of the humerus. Fracture of the nail was seen nearly 11 mo after insertion in the area of greatest stress to the nail.

All observations indicate compliance with ASTM and manufacturer's documents and crack initiation and propagation were consistent with cyclical loading by the patient, causing fatigue fracture of the device in the area of lowest section modulus and highest stress. These findings were consistent with the patient record and the nature of the fracture of the bone.

Failure Analysis of Mandibular Reconstruction Plate

Visual and light microscopic examination of the locking reconstruction plate seen in Fig. 25A as well as examination of the patient X-rays (Fig. 25B) showed fracture between the fourth and fifth screw hole from

the posterior end of the plate in the scalloped area. Multiple nicks and scratches through the anodized surface were observed indicating that they occurred after manufacturing. The plate had been contoured through the area of fracture as well as through the two adjacent anterior and posterior scalloped areas and screw holes to achieve anatomical alignment of the mandibular arch. Fracture of the plate occurred in the area of mandibular bone resection between the ends of the posterior and anterior bone segments where the plate was unsupported and underwent maximum stress. Additionally, the X-rays indicate that the reduced cross-section region (scalloped area) through which final fracture occurred coincided with the end of the posterior mandibular bone section.

After photographing the implants, they were rinsed in ethanol and blown dry with forced air. A small piece of the plate was removed using a diamond wafering blade, and a cut surface was polished to a $9\ \mu\text{m}$ diamond finish and quantitatively analyzed for the major elements using an energy-dispersive spectrometer, which had been calibrated using a standard of CP Ti. The results as seen in Fig. 26 indicates that the primary element is Ti with a minor amount of Fe, which is within ASTM specifications for Grade 4 CP Ti.

After compositional analysis, the piece was mounted, and polished to a final finish with colloidal silica. H_v measurements were made in three areas giving a mean value of 279.8, which is within the range expected for cold-worked Grade 4 CP Ti. After hardness testing, the sample was etched revealing a single-phase α -microstructure with an ASTM grain size of 9.5 (Fig. 27).

The fracture surfaces were ultrasonically cleaned in a detergent solution, rinsed in water and ethanol, and then blown dry with forced air. Fig. 28A shows the

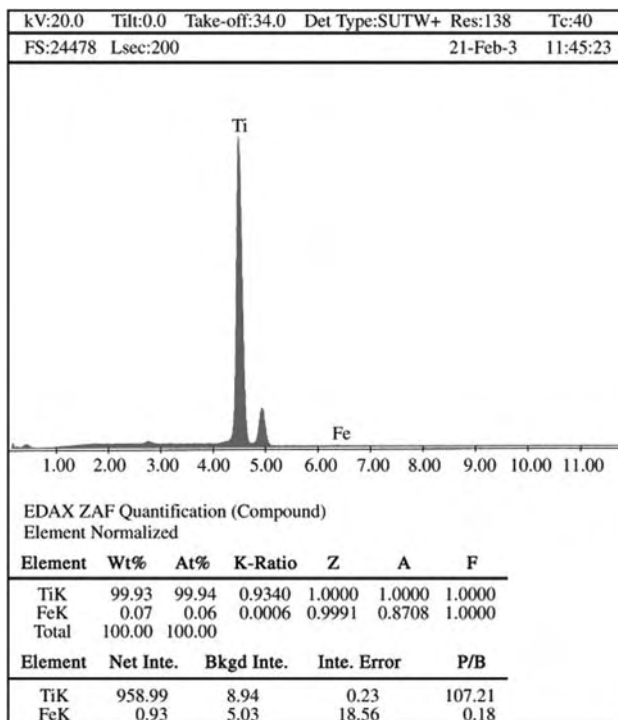


Fig. 26 Quantitative X-ray energy-dispersive spectroscopy of mandibular plate.

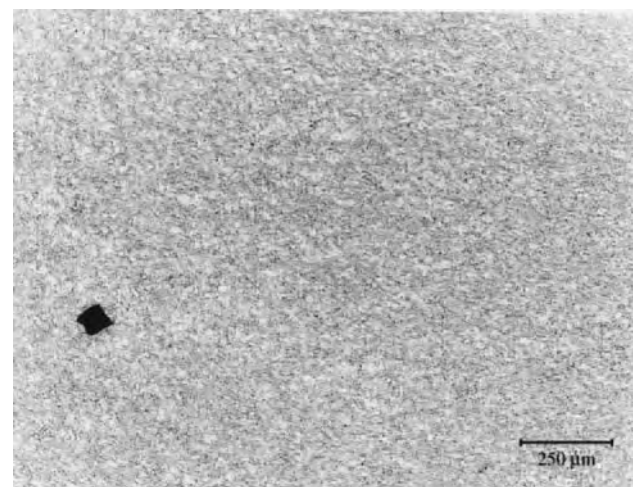


Fig. 27 Light micrograph of grain structure.



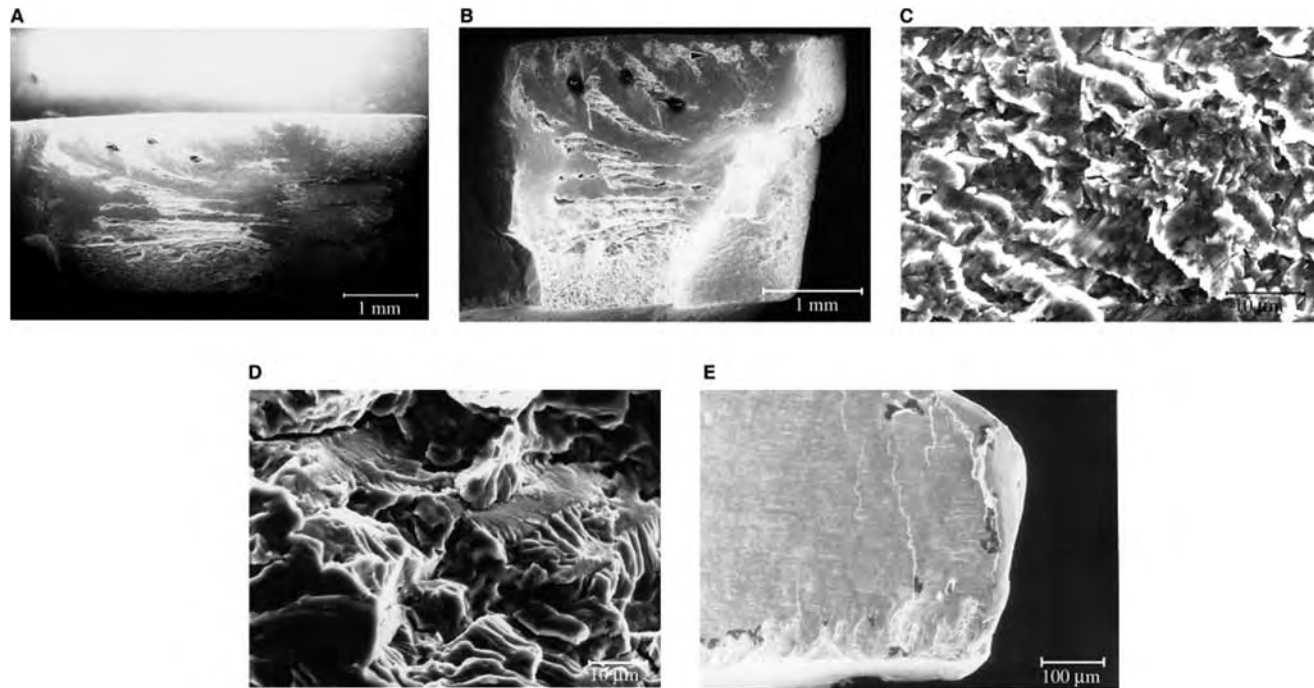


Fig. 28 (A) SEM micrograph showing distal fracture surface of mandibular reconstruction plate; (B) SEM micrograph showing oblique view of fracture surface; (C) SEM micrograph of central area of fracture surface showing fatigue striations; (D) SEM micrograph of area of fracture surface showing flutes and tear structures; and (E) SEM micrograph showing slip bands and additional cracks near fracture surface.

distal fracture surface oriented with the convex surface (buccal aspect) and the concave surface (lingual aspect) at the bottom. The nature of the fracture surface morphology indicates that the crack is initiated at the superior aspect in the area of reduced cross section between the screw holes (scallop). Fig. 28B shows the same fracture surface slightly tilted to better show the scalloped area and the orientation of the chevron marks emanating from the superior surface indicating the direction of crack propagation. While much of this fracture surface was worn, owing to repeated contact during and after fracture, Fig. 28C shows striations and wear typical of the predominant morphology seen

over most of this fracture surface. Fine striations ($< 1.0 \mu\text{m}$), secondary cracks, and wear are indicative of crack propagation owing to high-cycle Stage II fatigue. Fig. 28D shows not only striations and

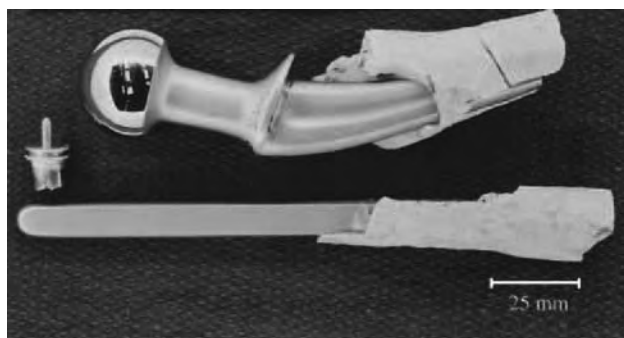


Fig. 29 Photograph showing pieces of Co-Cr hip prosthesis.

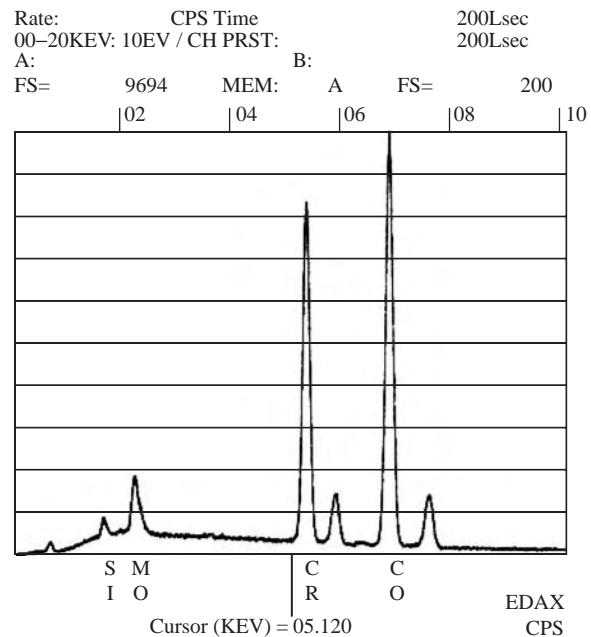


Fig. 30 X-ray fluorescence analysis of hip prosthesis.

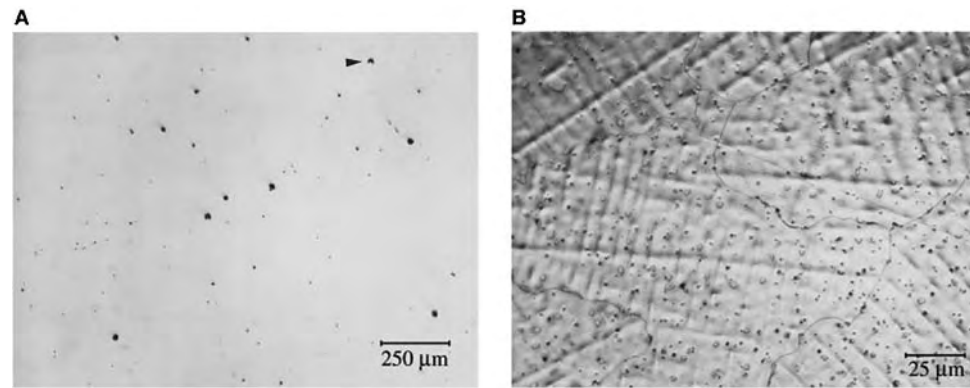


Fig. 31 (A) Light micrograph showing carbide distribution in ASTM F75 Co–Cr–Mo hip stem (100x) and (B) light micrograph showing large grain microstructure and dendritic grain growth in ASTM F75 Co–Cr–Mo hip stem.

secondary cracking but also fluting and terraces with feather marks indicative of a complex state of stress, often seen in fatigue fracture of hexagonal metals such as Ti. Examination of the areas near the fracture surface and the two superior scalloped areas posterior to the fracture surface showed additional cracks and multiple slip bands indicative of relatively high stresses being placed on the plate by the patient (Fig. 28E). The presence of the slip bands and the morphology of the additional cracks initiating at or near the superior surface of the plate indicates stressing of the plate by the patient and not during manufacturing.

In conclusion, compositional analysis confirms that the material is CP Ti and the concentration of Fe is within the limits defined by ASTM F67. Further, the hardness values, grain structure and grain size are within the range expected for this material, and no manufacturing defects were detected. Crack initiation and propagation was varied and consistent with the use as identified by the patient record and X-rays.

FAILURE ANALYSIS OF PROSTHESES

The primary difference between metals used for implanted metal prostheses are that they may be cast

rather than wrought and the prosthesis is designed for long-term use. Because of the prolonged time, it may be difficult to obtain patient records and device history records. As previously stated, there are a number of factors, which may have contributed to the failure of the prosthesis other than fracture of the metal components. Another significant difference is that because the prostheses are definitive devices designed to replace bone, fracture of the metallic components has become increasingly rare.

Failure Analysis of Co–Cr Hip Stem

Fracture of the hip prosthesis discussed in the following paragraphs had occurred transversely through the stem approximately two-thirds of the distance from the distal end (Fig. 29). Prior to microstructural evaluation, a piece was cut from the stem and polished with 1 μm diamond, qualitatively analyzed using X-ray fluorescence analysis as seen in Fig. 30. The analysis indicates that the material composition is in agreement with those to be expected for implant quality casting alloy conforming to ASTM F75. This same polished sample was evaluated for Vickers hardness (H_v) and a mean value of 347.1 was



Fig. 32 (A) SEM micrograph of fracture surface of Co–Cr–Mo hip stem; (B) SEM micrograph showing saw tooth pattern on fracture surface indicative of Stage I fatigue crack growth; and (C) SEM micrograph showing fatigue striations and secondary cracking on fracture surface.

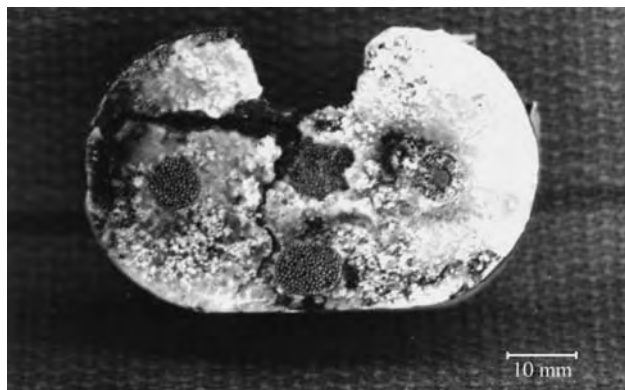


Fig. 33 Photograph of fractured tibial tray.

recorded. This value is in the range expected for cast Co–Cr–Mo alloy (ASTM F75). The inclusion content and grain structure were examined on the sample used for hardness testing. Evaluation of inclusions showed typical carbide distribution (Fig. 31A) and pores typical of castings. The sample was etched to reveal a typical cast Co–Cr–Mo alloys with large grains and dendritic grain growth (Fig. 31B).

After cleaning ultrasonically in mild detergent solution, the fracture surface was viewed from the lateral toward the medial (Fig. 32A). In addition to the fatigue striations noted in many areas of the fracture surface (Fig. 32B), of particular interest is the multifaceted appearance of much of the fracture surface. This type of morphology is consistent with planar slip in this cobalt alloy and is indicative of fatigue crack at low values of ΔK while striations normally occur at higher ΔK values. Material separation has taken place along these preferred glide planes which is typical in this alloy. Also, the small voids normally present in this type of casting should be noted. During Stage II fatigue, striations are often contained on plateaus at different levels and are joined by ledges containing striations or by tear ridges. Examination at several orientations revealed that the area above the secondary crack is not in the same plane as the area below the crack. Another view of the fracture surface, but with significant tilting can be seen in Fig. 32C, which clearly shows that the area above the crack is at a significant angle to the area below which accounts for the apparent lack of fatigue striation alignment across the crack. As the two planes are in different crystallographic directions, they are joined together by a ledge or wall containing fatigue striations.

As neither the X-rays nor the patient record were available, discussion with the patient indicated that the prosthesis had been in place for in excess of 15 yr. In addition, the patient record indicated significant fracture and deterioration of the cement as well as bone resorption. This information, along with the failure analysis, the age, and the weight of the patient

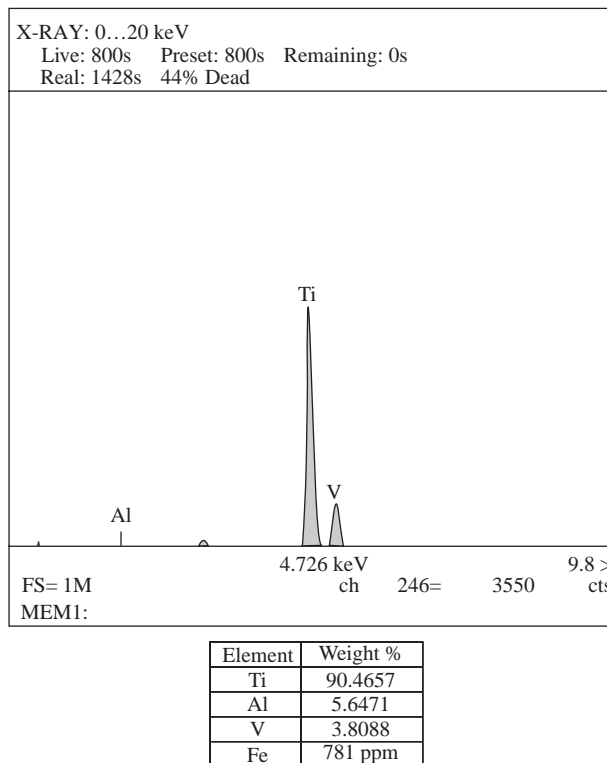


Fig. 34 Quantitative X-ray energy-dispersive spectra of Ti–6Al–4V fractured tibial tray.

indicated that fatigue of the hip stem was related to factors other than the prosthesis itself.

Failure Analysis of Tibial Tray

The fractured tibial tray, which is the subject of this fracture analysis, is seen in Fig. 33. After removal of a

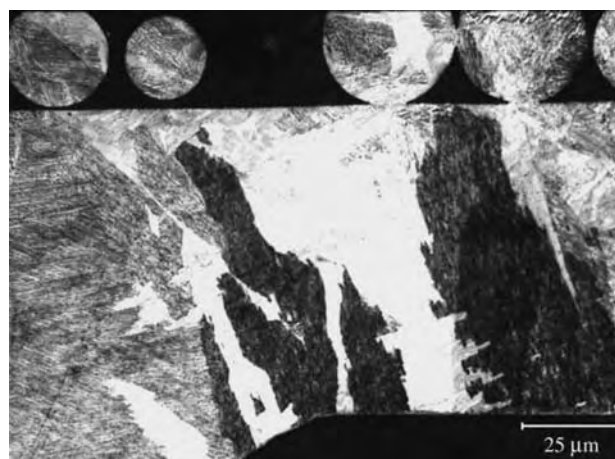


Fig. 35 Light micrograph showing large grained microstructure of Ti–6Al–4V tibial tray—note microcrack (35x) at left.

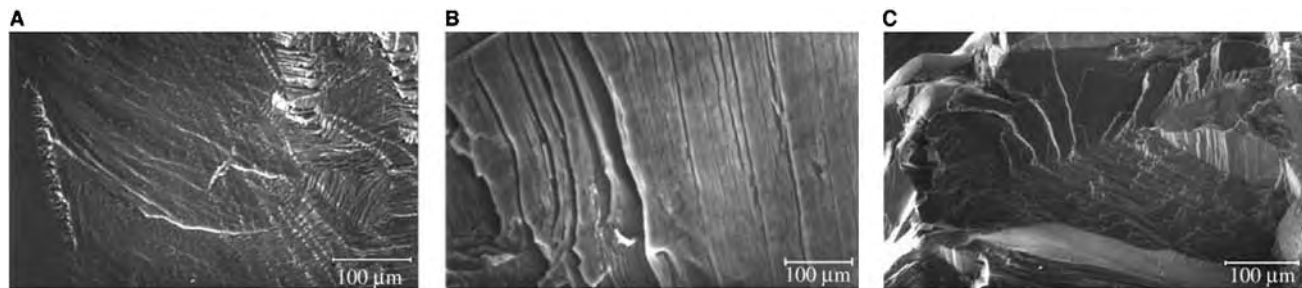


Fig. 36 (A) SEM micrograph of typical fracture surface morphology; (B) higher magnification SEM micrograph showing variable striation spacing; and (C) higher magnification SEM micrograph showing striations, fluting, and terrace.

small piece of the implant, compositional analysis of major alloying elements using X-ray fluorescence analysis showed that the material was Ti-6Al-4V (Fig. 34). This piece was mounted and polished to a final finish with colloidal silica, and three H_v measurements were made resulting in a mean value of 313. The surface was then etched revealing a large grained microstructure and additional microcracks in the section (Fig. 35). This large grain microstructure indicates a cast rather than wrought tibial tray.

SEM examination of the fracture surface revealed a mixed morphology, which is typical of fatigue fracture of this alloy having a large grained microstructure. Striations, fluting, cleavage, tongues, and a saw tooth structure typical in fatigue fracture of hcp metals such as Ti-6Al-4V alloys under a variable load spectrum was observed (Fig. 36A through C).

Examination of the patient record and X-rays showed significant bone resorption below the tibial tray. No materials defects were observed and this resorption, leaving an unsupported tibial component, accounts for excessively high stresses leading to fatigue fracture.

CONCLUSIONS

It is extremely important that in none of the devices evaluated was a defect in materials, design, or manufacturing noted. This observation is most often the case. Great advances in materials and design have been made over the last several decades and while defects can occur, they are extremely rare. Analysis of any failed device requires careful attention to detail. Precise information concerning the history of the specific device is often needed to render an accurate opinion. Because implants have a direct and very personal impact on people, failure analysis information is especially important as it may lead to changes in materials, design, and manufacturing. However in most cases, failure of the device can be traced to problems with

patient biology, physiology, or failure to comply with the prescribed course of treatment. Therefore, failure analysis should not be thought of as a method to prove or disprove the adequacy of a device. Rather, the scientist should embark on the analysis with the express intent of determining: (i) If the device is faulty and (ii) What contributing factors could have led to failure if no materials, manufacturing, and/or design faults are noted.

ARTICLES OF FURTHER INTEREST

Shape Memory Metals; Titanium and Its Alloys

REFERENCES

1. *Webster's Ninth New Collegiate Dictionary*; Merriam-Webster Inc.: Springfield, MA, 1991; 445.
2. Schmid-Schönbein, G.W.; Woo, S.L-Y.; Zweifach, B.W. *Frontiers in Biomechanics*; Springer-Verlag: New York, 1986; 197-201.
3. Mears, D.C. Materials and orthopaedic surgery. In *Tissues of the Musculo-Skeletal System*; Williams & Wilkins Company: Baltimore; 1, 142-150.
4. Anthony, C.P.; Kolthoff, N.J. *Textbook of Anatomy and Physiology*; C.V. Mosby Company: St. Louis, 1971.
5. Hole, J.W. *Essentials of Human Anatomy and Physiology*, 2nd Ed.; W.C. Brown: Dubuque, Iowa, 1986.
6. Perren, S.M.; Matter, P.; Ruedi, R.; Allgöwer, M. Biomechanics of fracture healing after internal fixation. In *Surgery Annual*; Nyhus, L.M., Ed.; Appleton-Century-Crofts: New York, 1975; 361.
7. Wainwright, S.A.; Biggs, W.D.; Currey, J.D.; Gosline, J.M. *Mechanical Design in Organisms*; 1976; 169.
8. Engh, C.A.; O'Connor, D.; Jasty, M.; McGovern, T.F.; Bobyn, J.D.; Harris, W.H. Quantification of implant micromotion, strain shielding, and bone resorption with porous-coated anatomic medullary locking femoral prostheses. *Clin. Orth. Rel. Res.* **1992**, 285, 13.

Microcirculatory Oxygen Transport

Roland N. Pittman

Department of Physiology, Virginia Commonwealth University, and Department of Biomedical Engineering, Emergency Medicine, Virginia Commonwealth University, Richmond, Virginia, U.S.A.

INTRODUCTION

Mammalian cells require an adequate supply of energy, stored in the form of the high-energy phosphate bond of adenosine triphosphate (ATP), to carry out their various functions. Consequently, cells need a continuous supply of oxygen, as they prefer to generate ATP by the biochemical process of oxidative phosphorylation in the mitochondria. Oxygen is carried in the blood, and almost all of it (about 98%) is reversibly bound to hemoglobin in the red blood cells (RBCs). The pulmonary circulation oxygenates blood from the pulmonary artery, which has returned to the right side of the heart after giving up part of its oxygen in the peripheral tissues. The oxygenated blood from the pulmonary vein now flows into the left side of the heart, where it is pumped into the systemic circulation by the left ventricle. The oxygenated blood finally reaches the peripheral tissues through the intricate microvascular networks of the circulation, where oxygen exchange takes place between the blood and tissues across the walls of the arterioles, capillaries, and venules. Microcirculatory oxygen transport forms the subject of this article.

August Krogh, the founder of the field of microcirculatory oxygen transport, thought that most oxygen exchange occurred across the walls of the smallest microvessels, the capillaries, owing to their high surface area-to-volume ratio and the low velocity of RBCs flowing through them—conditions that maximized the opportunity for diffusive exchange. However, research over the last 30 years has indicated otherwise, because arterioles can be a major site of oxygen diffusion. There has been an increased pace of activity in this research area during the past 10 years, spurred in large part by the maturation of new methods, development of artificial oxygen carriers, and the realization of the important role played by oxygen (or rather its lack) in a number of pathophysiological situations. Based on his histological and *in situ* observations of capillaries in skeletal muscle, Krogh^[1] formulated a simple conceptual model consisting of a capillary surrounded by a concentric cylinder of oxygen-consuming tissue with uniform consumption and oxygen diffusion properties. This mathematical model has served as the basis

for most modeling in this area, although many of the simplifying assumptions of the model are not strictly valid under the most general conditions. Duling and Berne^[2] made the first systematic measurements of oxygen levels in the microcirculation and unexpectedly found that a substantial amount of oxygen was lost from the arteriolar network, prior to the loss from the capillaries.

The rapid evolution of the field of microcirculatory oxygen transport over the past 30 years has been closely linked to the development of new methods for intravital microscopy. These methods allow the direct visualization of the microcirculation and measurement of the following key oxygen transport variables in individual vessels of microvascular networks: geometric (diameter and segment length); hemodynamic (velocity, hematocrit, blood flow, and RBC flow); and oxygenation (oxygen tension and oxygen saturation). These measurements can be combined to quantify the convective flow of oxygen in microvessels, and application of mass balance for oxygen can be used to calculate the rate at which oxygen diffuses from the lumen of microvessels. This article focuses on the study of oxygen transport from the point of view of the microcirculation. The fundamental conceptual basis for the subject, the Krogh cylinder model and principles of mass transport, are emphasized. A description of the basic methods applied to measure oxygen transport variables in microvessels is presented. Results of measurements in arterioles, capillaries, and venules are summarized, along with unresolved issues and controversies associated with the measurements. Finally, future research directions involving artificial oxygen carriers, nitric oxide (NO), and pathophysiological circumstances resulting in cellular hypoxia are presented.

NEED FOR OXYGEN: CELLULAR METABOLISM

Mammalian cells need a continuous and adequate supply of oxygen to carry out their various functions. The cardiovascular system serves this need by providing blood vessels in intimate contact with the surface of most cells: the flowing blood serves a dual purpose by providing needed oxygen and by carrying

away metabolic waste products. All cells have the biochemical machinery to produce ATP. The most efficient means for producing ATP is through the process of oxidative phosphorylation. ATP is produced at several steps in the Krebs cycle and along the electron transport chain in the mitochondria. Molecular oxygen reacts with the terminal enzyme in the chain—cytochrome oxidase. As long as the partial pressure of oxygen or oxygen tension (PO_2) inside the cell is higher than about 1 mmHg, the cellular and mitochondrial oxygen consumption will not be limited by the availability of oxygen. For oxygen to reach the intracellular space and the mitochondria, it must diffuse from the RBC through several barriers imposed by the RBC membrane and plasma in the blood, the microvessel wall, the interstitial fluid, the parenchymal cell membrane, the cytoplasm, and finally the mitochondrial membrane. The PO_2 progressively decreases from the RBC to the mitochondrion, because oxygen moves by diffusion and net diffusion requires a gradient in PO_2 as its driving force. In some cells (e.g., skeletal muscle cells with oxidative fiber type), the heme protein myoglobin is present and can serve as a temporary storage depot for oxygen, and possibly as a facilitator of oxygen diffusion through carrier-mediated transport.

CARRIAGE OF OXYGEN

Oxygen is carried in two forms in the blood: 1) physically dissolved in plasma and RBC cytoplasm and 2) reversibly bound to hemoglobin molecules in the RBC. The concentration of oxygen in the dissolved form, $[O_2]_D$ in ml O_2 /ml, accounts for only about 2% of the total oxygen in blood at normal arterial oxygen saturation, and it is given by Henry's law

$$[O_2]_D = \alpha PO_2 \quad (1)$$

where α is the solubility coefficient of oxygen in ml O_2 /ml/mmHg. Most of the oxygen (about 98%) is bound to the hemoglobin inside the RBCs and the concentration of oxygen in the bound form, $[O_2]_B$, is given by

$$[O_2]_B = C_{Hb}SO_2[Hb] \quad (2)$$

where C_{Hb} is the oxygen-binding capacity of hemoglobin (1.34 ml O_2 /g/Hb), SO_2 is the oxygen saturation (or fraction of heme-binding sites occupied by oxygen), and $[Hb]$ is the hemoglobin concentration in blood (g/ml).

As RBCs flow through the circulation, the hemoglobin binds or releases oxygen according to the local PO_2 gradient. Thus, in the pulmonary circulation, there is a net uptake of oxygen by the RBCs, while in

the systemic circulation there is a net release of oxygen from the RBCs. The equilibrium relationship between SO_2 and PO_2 is the oxygen dissociation curve. Its shape is described as sigmoid and has to do with the cooperative nature of oxygen binding to hemoglobin. Hemoglobin is a tetrameric molecule composed of four polypeptide chains, two α and two β chains, each with a heme group where ligands such as oxygen, carbon monoxide, and NO can bind. The heme group consists of a porphyrin ring surrounding an iron (Fe) atom. Ferrous iron (Fe^{2+}) can reversibly bind oxygen, whereas ferric iron (Fe^{3+}) cannot. The latter oxidized form of hemoglobin is called methemoglobin, and it is not capable of binding oxygen. Hemoglobin exhibits positive cooperativity in which successive oxygen molecules bind more tightly than the previous one. A simple quantitative description of oxygen binding by hemoglobin is given by Hill's equation

$$SO_2 = (PO_2/P_{50})^n / [1 + (PO_2/P_{50})^n] \quad (3)$$

where P_{50} is the PO_2 at which the hemoglobin is 50% saturated with oxygen and n is Hill's coefficient, which describes the degree of cooperativity of binding. As P_{50} decreases, the affinity of hemoglobin for oxygen increases, making it harder for hemoglobin to release oxygen. The converse is also true: as P_{50} increases, the affinity of hemoglobin for oxygen decreases, making it easier for hemoglobin to release oxygen. P_{50} varies with temperature, pH, and partial pressure of CO_2 (PCO_2). Increases in temperature, $[H^+]$ and PCO_2 , separately or in combination, will lead to a rightward shift of the oxygen dissociation curve (i.e., to higher values of P_{50}); decreases in any of these factors will produce a leftward shift. Shifts in the curve caused by changes in $[H^+]$ and/or PCO_2 are called the Bohr effect.

CARDIOPULMONARY SYSTEM

Overall management of oxygen transport in an organism is under the coordinated control of the lungs, the heart, and the blood vessels that connect them to the systemic organs, the cardiopulmonary system. Deoxygenated blood from the peripheral tissues is directed to the right atrium, from where it passes to the right ventricle and then is pumped to the pulmonary artery, and on through the small vessels of the lung. While flowing through the pulmonary capillaries, the blood is in intimate contact with the alveolar gas spaces, where it releases carbon dioxide and picks up oxygen, both by passive diffusion. Before the blood leaves the pulmonary capillaries, it comes into equilibrium with the gas in the alveoli (i.e., both PO_2 and PCO_2 in the blood are equal to their respective partial

pressures in the alveolar gas).^[3] The oxygenated blood passes through the pulmonary vein into the left atrium and then into the left ventricle, where it is ejected into the aorta. The blood is then conducted through the arterial network to the various tissues of the body, where it flows through the microvascular networks of the individual tissues. Exchange of oxygen takes place in the microcirculation of each tissue and the deoxygenated blood is collected in the venous network and conducted back to the right atrium.

MICROCIRCULATION

Roughly speaking, the microcirculation is the collection of blood vessels that are small enough so as to require a magnified image to visualize clearly. The microcirculatory network is composed of series and parallel connections of arterioles, capillaries, and venules. The primary function of the arterioles is to control the magnitude and distribution of blood flow within the network; their walls are composed of an inner layer of endothelial cells, surrounded by vascular smooth muscle. The primary function of the capillaries is the exchange of water and small solutes; their walls are composed of a single layer of endothelial cells. The primary function of the venules is collection of the blood flowing from the capillaries; their walls are composed of an inner layer of endothelial cells surrounded by vascular smooth muscle. In regard to the exchange of oxygen, all these vessels can participate, because their walls are permeable to oxygen and PO_2 differences exist between their lumens and the interstitial fluid.^[4]

APPROACHES TO STUDYING OXYGEN TRANSPORT

Oxygen transport can be studied at various levels of organization, from the intact organism to isolated organs to isolated cells to subcellular components (e.g., mitochondria and enzymes of the electron transport chain). The combination of a small microcirculatory network (i.e., feeding arteriole, capillaries, and collecting venules) and the tissue it serves represents the smallest complete functional unit that exhibits all the features of the intact tissue. Thus, studies at the level of the microcirculation can provide a rich source of information that can be linked to behavior on a larger scale.

Mathematical modeling of microcirculatory oxygen transport can be traced back to the seminal work of Krogh,^[1] referred to earlier. Although the Krogh cylinder model has formed the basis of most modeling in this area, recent advances in computational hardware

(i.e., processor speed, memory, storage space, and parallel processing) and software design, coupled with the availability of more accurate and complete experimental data, have made it possible to create models that incorporate realistic network geometry, hemodynamic schemes, and metabolic pathways.^[5,6]

Most experimental approaches have appealed to the principles of mass balance for oxygen, in which measurements of geometry, hemodynamics, and oxygenation in unbranched segments of a network are used to compute the convective flow of oxygen (QO_2^{CONV}) into and out of the segment (Fig. 1). The difference between the rate of inflow and outflow of oxygen in the segment is equal to the rate at which oxygen diffuses (QO_2^{DIFF}) into the wall of the vessel from the lumen

$$QO_2^{DIFF} = QO_2^{CONV}(In) - QO_2^{CONV}(Out) \quad (4)$$

The diffusive flux of oxygen out of the lumen (JO_2) is the diffusive flow divided by the luminal surface area

$$JO_2 = QO_2^{DIFF} / 2\pi RL \quad (5)$$

where R is the luminal radius and L is the length of the segment. Assuming that the vessel wall does not have oxygen consumption and oxygen permeability properties different from those of most other tissues, the decrement in PO_2 across the wall (i.e., transmural PO_2) should be relatively small. This expectation is consonant with results reported by Duling and Berne.^[2] The application of this approach using the methods described in the next section yields the following expression for QO_2^{CONV} (Fig. 1):

$$QO_2^{CONV} = \pi R^2 v[Hb] SO_2 C_{Hb} \quad (6)$$

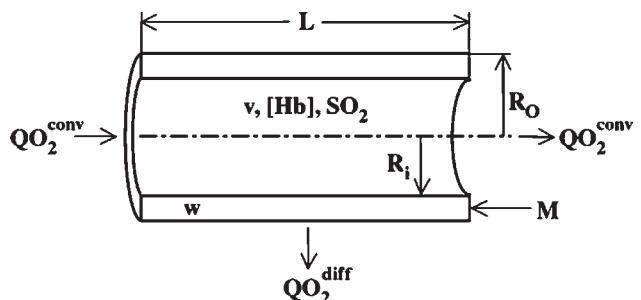


Fig. 1 Calculation of diffusive oxygen flow (QO_2^{DIFF}) or flux ($QO_2^{DIFF}/\text{luminal surface area}$) from oxygen mass balance in an unbranched arteriolar segment of internal radius, R_i , and length, L . Variables used to compute QO_2^{CONV} are R_i , v , $[Hb]$, and SO_2 . Luminal surface area is $2\pi R_i L$. Wall thickness, w , is $R_o - R_i$. Oxygen consumption of the wall of this segment is $\pi(R_o^2 - R_i^2) L M$, where M is the volume-specific oxygen uptake of the wall. (From Ref. [7].)

where v is the average blood velocity and the other symbols have been defined previously. $[Hb]$ and SO_2 are the appropriate flow-weighted average values across the vessel lumen.^[8] However, because the radial dependencies of these variables are generally unknown, measured values are averages of $[Hb]$ and SO_2 along a vertical line through the center of the vessel. Results from this approach are described below.

DEVELOPMENT OF EXPERIMENTAL METHODS

The study of microcirculatory oxygen transport requires the measurement of several geometric, hemodynamic, and oxygenation variables in the living microcirculation. The availability of an instrument to perform intravital microscopy is common to all the measurements. A research-grade microscope equipped with illumination sources for transillumination of thin tissues and epifluorescence of thin or thick tissues is needed. Microscopic images are typically viewed with a closed-circuit video camera, either analog or digital, and video monitor, with additional capability to record desired images for later observation and analysis. Recordings of video images can be made with a videocassette recorder or sequences of digital images can be stored in computer memory or on a hard drive. Geometric measurements such as vessel diameter or segment length can be made directly from the image on a video monitor or by using standard image analysis software. Hemodynamic measurements such as particle velocity and hematocrit require specialized additions to the intravital microscope. Two approaches to measuring the velocity of fluorescent particles (e.g., labeled RBCs and microspheres) using video technology include multiple flash illumination^[9] and electronic shuttering of the video camera.^[10] Both these methods provide multiple images of the moving particles. By measuring the distance moved between successive particle images of known temporal spacing, the velocity can be calculated. An alternative procedure requires the use of a dual photosensor arrangement to detect the modulation of transmitted light by the passage of RBCs past the detection region.^[11] The photometric signals at the closely spaced upstream and downstream measurement points are cross-correlated and the velocity is computed from the ratio of photosensor separation and relative time shift of the signals that produces maximum cross-correlation (so-called “dual sensor” method). The measurement of hematocrit is based on the spectrophotometric determination of $[Hb]$ in a microvessel. The optical density of blood at a wavelength that is insensitive to the oxygenation state of hemoglobin (i.e., an isosbestic wavelength) is measured and used to compute $[Hb]$ from Beer’s law of light absorption. Because RBCs scatter light, in addition to that absorbed by the hemoglobin in the RBC, it is usually

necessary to measure the optical density at a second isosbestic wavelength to correct for the contribution of light scattering to the optical density. This photometric measure of $[Hb]$ is, in general, proportional to neither tube hematocrit (volume-weighted concentration) nor discharge hematocrit (flow-weighted concentration) as discussed by Ellsworth and Pittman,^[12] owing to the nonuniform luminal distribution of RBCs. Rather, it gives the $[Hb]$ averaged over a vertical line through the center of the vessel. The measurement of SO_2 is also a spectrophotometric method requiring video images of a microvessel at two (for capillaries^[13]) or three wavelengths (for arterioles and venules^[14,15]), one of which must be a wavelength at which there is a large difference between light absorption by oxy- and deoxyhemoglobin. The measurement of PO_2 in tissues or microvessels can be carried out with polarographic microelectrodes. The recessed-tip electrodes are most appropriate for microcirculatory measurements, because they have been shown to create the least disturbance of the existing PO_2 field. Detailed descriptions of all these methods, as well as the surgical preparation of various tissues for microscopic viewing, are contained in the monograph *Microcirculatory Technology*.^[16] A more recently introduced approach to measure PO_2 relies on the ability of oxygen to quench the phosphorescence of molecules excited by a brief flash of light (phosphorescence-quenching method). The lifetime of the phosphorescence decay is inversely related to the PO_2 (Stern–Volmer relation). This method has been used to measure PO_2 in microvessels and interstitial fluid.

FINDINGS IN DIFFERENT TISSUES

The methods and experimental approaches described above have been used to investigate the transport of oxygen in the microcirculation of several different tissues in a variety of species. These studies have led to discoveries of a number of interesting and unexpected results.

Results in Arterioles

Duling and Berne^[2] reported the first systematic study of oxygen levels in the arteriolar network, using oxygen microelectrodes to measure perivascular PO_2 . Contrary to the expectation from Krogh’s work, that oxygen exchange between the blood and tissues would be the exclusive domain of the capillaries, they found a substantial and progressive loss of oxygen in the precapillary arteriolar network—a longitudinal gradient in PO_2 . Following the development of the spectrophotometric method to measure SO_2 ,^[14] subsequent work showed that the fall in perivascular PO_2 was



accompanied by a parallel fall in luminal SO_2 predictable from the oxygen dissociation curve. The simultaneous measurement of diameter, velocity, hematocrit, and SO_2 at upstream and downstream sites in unbranched arterioles allowed the calculation of convective oxygen flow at the two sites, the difference of which, by mass balance for oxygen, was the flow of oxygen from the lumen by diffusion across the arteriolar wall.^[7] It was found in resting skeletal muscle that most of the oxygen, about two-thirds, diffused from the arteriolar network and only about one-third diffused from the capillaries,^[17] corroborating the previous results of Duling and Berne.^[2] Careful comparison of the experimental results with the predictions of computational models of oxygen transport led to the finding that the apparent rate at which oxygen diffused from the arterioles was an order of magnitude higher than expected.^[18] Attempts to explain this large discrepancy have followed two different paths: 1) possible misinterpretation of incomplete data on luminal SO_2 ^[8] and 2) proposed high oxygen consumption by the arteriolar wall.^[19] There is no doubt that oxygen is lost from the arteriolar network, leading to the conclusion that there must be radial gradients in SO_2 and PO_2 .^[7] Because the measurements of SO_2 referred to above were made above the centerline of the vessel, each represented an average value along a line through the center of the arteriole. Using empirical shapes for the luminal SO_2 profiles, one can calculate the relationship between measured centerline SO_2 and the parameters used to characterize the SO_2 profile.^[8] It can be shown that reasonable values of the parameters are compatible both with the measured SO_2 and with the expected rate of oxygen diffusion from the arteriole, thereby removing the order of magnitude disagreement between experimental and computational studies. While this explanation works in principle, definitive experimental evidence in the form of actual SO_2 profiles remains to be presented. The alternative hypothesis, based on the measurement by phosphorescence-quenching of a large transmural gradient in PO_2 ,^[20] proposes that the large efflux of oxygen from the arteriolar lumen is accounted for by high oxygen consumption of the wall. This possibility is more difficult to understand, because, to account for the apparent large oxygen flux from the arterioles, the oxygen consumption of the cells in the wall must be about two orders of magnitude higher than previously reported values for vascular tissue.^[21] This circumstance appears to be unlikely and it is not clear how to otherwise reconcile the report of the large transmural PO_2 gradient with previous and widely accepted results on oxygen transport in the microcirculation. The eventual resolution of these disparate results should aid our understanding of oxygen transport in microvascular networks.

Results in Capillaries

The diameter of a capillary is typically small enough so that RBCs are constrained to flow through it in single file. The first systematic measurements of oxygen transport in capillaries were reported by Ellsworth, Popel, and Pittman.^[22] A local longitudinal gradient in SO_2 was observed, which was consistent with tissue oxygen consumption and diffusion properties of skeletal muscle, in contrast to the order of magnitude discrepancy between predicted and observed oxygen flux from the arterioles. An unexpected finding was that there was not a monotonic decrease in SO_2 from the entrance to the exit of the capillary. Rather, after an initial fall in SO_2 , there was a rise and then fall in SO_2 associated with the passage of an arteriole across the path of the capillary.^[23] This observation was later validated by computational modeling of the diffusive uptake of oxygen, from a nearby arteriole, by RBCs flowing through the capillary.^[24] A recent finding is the occurrence of fluctuations in PO_2 as RBCs and plasma gaps alternately pass an observation point along a capillary.^[25] This phenomenon of erythrocyte-associated transients in PO_2 was predicted by theoretical models from the late 1970s,^[26,27] but was only verified after the spatial (1 μm) and temporal (10 msec) resolution of the phosphorescence-quenching technique was improved to this level.

Results in Venules

Venules collect blood that has flowed through the capillaries, so that venular blood generally has a lower level of oxygenation than blood from other sites. However, in some cases the oxygenation of venular blood gradually increases from the smallest to the largest venules.^[28] Thus, it is of interest to consider potential sources of this oxygen. Because of the heterogeneity of oxygen release from the capillaries, the blood that flows into venules from different capillaries also has a similar degree of heterogeneity in respect to oxygen, and it is possible that the blood added to the venules at each successive branching order would result in an increase in blood oxygenation. In addition to this convective explanation for the increase in venular oxygenation, two potential diffusive sources of oxygen exist: 1) paired arterioles that run parallel to the venules and 2) neighboring regions of interstitial fluid with higher PO_2 than that of the venular blood.^[28] Although the case of diffusive shunting of oxygen directly between adjacent arterioles and venules in close proximity is appealing, calculations have shown that the contribution to venular oxygenation by this mechanism is negligible, except for instances where blood flow is lower than normal.^[5] The PO_2 of the

interstitial fluid that bathes all the microvessels and parenchymal cells appears to lie between the values of nearby arterioles and venules, so that it can be a source of oxygen for the venules. Because the walls of all microvessels appear to be permeable to oxygen, net diffusion of oxygen only requires for there to be a PO_2 difference between microvascular blood and a neighboring region.

EMERGING AREAS OF RESEARCH

Several research areas with a close relationship to microcirculatory oxygen transport are ripe for investigation. Three of these areas are highlighted below in this section.

Artificial Oxygen Carriers

The addition of artificial oxygen carriers or blood substitutes to the circulation can be of potential benefit when the normal complement of oxygen-carrying RBCs has been reduced, as in intentional normovolemic hemodilution or hemorrhage followed by resuscitation with crystalloids or colloids. Two major types of blood substitutes have been pursued: 1) hemoglobin-based oxygen carriers (HBOCs) that carry oxygen in the bound form and 2) perfluorocarbon emulsions (PFCs) that carry oxygen in the dissolved form.^[29]

The amount of oxygen that can be carried by a HBOC is proportional to the concentration of oxygen-binding sites in the molecule. As the name implies, HBOCs are derived from hemoglobin, typically modified by chemically cross-linking the two α or β chains of the tetramer together, preventing the tetramer from dissociating into lower molecular weight dimers. Further modifications are often made, two of the more common being polymerization of the cross-linked tetramers and decoration of the tetramer with molecules such as polyethylene glycol; these changes are made to increase the size of the HBOC.^[29] The reason for increasing the size of the HBOC is that the infusion of tetrameric hemoglobin into the circulation leads to a rapid rise in arterial blood pressure, and this pressor response can be mitigated by using larger HBOCs. The mechanism for the elevated blood pressure is under investigation, but the leading candidate is the NO-scavenging action of free hemoglobin. NO (see below) is a vascular smooth muscle relaxant produced in the endothelial cells and other places. When its concentration in the arteriolar wall decreases, a net constriction occurs, raising total peripheral resistance and arterial blood pressure. NO binds tightly to hemoglobin, so that introduction of a HBOC into the

circulation causes a rapid reduction in NO concentration, both in the lumen of microvessels and their wall, leading to vasoconstriction and a rise in blood pressure. Increasing the size of the HBOC reduces the pressor effect, presumably by making it more difficult for the HBOC to move close to the endothelial surface and reducing the ability of the HBOC to extravasate from the microcirculation and influence NO levels from the interstitial side of arterioles.

PFCs consist of submicron particles about $0.2\ \mu\text{m}$ in diameter. The amount of oxygen that can be carried by the PFC is proportional to the perfluorocarbon concentration and to the ambient PO_2 , according to Henry's law. Because PFCs do not scavenge NO, they do not exhibit the pressor response of the HBOCs. For substantial amounts of oxygen to be carried by PFCs, the ambient PO_2 must be considerably higher than that of air at sea level, and 100% oxygen is often used with PFCs to increase the amount of oxygen they can deliver.

NO and Oxygen Transport

NO is a small, diatomic, gaseous molecule that is produced enzymatically from L-arginine in the presence of oxygen. The enzyme responsible for NO synthesis is nitric oxide synthase (NOS) of which there are several isoforms, depending on the source of NO: endothelial or eNOS, neuronal or nNOS, inducible or iNOS, and mitochondrial or mtNOS. NO can influence oxygen transport in two distinct ways. First, because NO is a smooth muscle relaxant and is formed by the endothelial cells, any change in the local NO concentration will alter the diameter of the nearby resistance vessels, and hence the local blood flow responsible for the convective supply of oxygen. NO levels can increase by the action of NO-dependent vasoactive agents (e.g., acetylcholine and bradykinin) and by increases in shear stress on the endothelium. Second, because NO interacts with cytochrome oxidase and reduces mitochondrial oxygen consumption, it can also modify the demand for oxygen. Thus, NO is positioned to modulate the balance between oxygen supply and demand at the level of the microcirculation.

Pathophysiology of Oxygen Transport

As having an adequate supply of oxygen is essential to the normal functioning of all tissues, any pathological or disease state in which the oxygen supply relative to oxygen demand is reduced places a tissue in functional jeopardy. Examples of such states are cancer, diabetes, environmental hypoxia, hemodilution, hemorrhage, hypertension, inflammation, ischemia/reperfusion, sepsis, and sickle cell anemia. In most

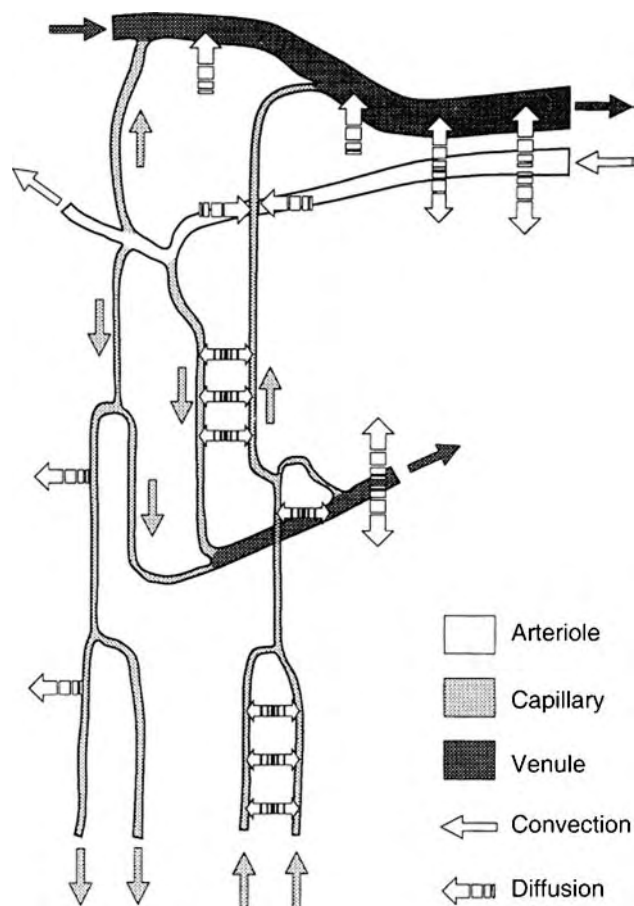


Fig. 2 Oxygen exchange among microvessels. Current understanding of oxygen supply to tissue requires consideration of diffusive interactions among all neighboring microvessels. A microvascular network is shown, containing arterioles, venules, and capillaries, with diffusion indicated by interrupted arrows and convection of oxygen in arterioles, capillaries, and venules, indicated by appropriately shaded solid arrows. The key to understanding oxygen supply to tissue requires a consideration of the complex nature of diffusive interactions among various microvessels. (From Ref. [4].)

cases, the supply of oxygen is the limiting factor and this can be compensated for by increasing blood flow to the affected tissue or by a transfusion of RBCs or HBOC. In cases where the demand for oxygen exceeds the supply, oxygen consumption can be reduced by lowering the temperature of the affected tissue or partially inhibiting the reaction of oxygen with the electron transport chain of the mitochondria.

CONCLUSIONS

An adequate supply of oxygen to cells is critical, as almost all cells require oxygen to produce the energy that supports their function. The exchange of oxygen between the blood and tissue occurs by diffusion at

the level of microvascular networks, in particular in the arterioles and capillaries. The various possibilities for oxygen exchange are summarized in Fig. 2. Typically, oxygen is lost from the arterioles and capillaries and is acquired by the venules. The primary site of oxygen transport in a specific tissue depends on its blood flow and metabolic rate: arterioles appear to be the main site in tissues with low flow and metabolism (e.g., resting skeletal muscle), whereas the capillaries appear to be preferred in tissues with high flow and metabolism (e.g., brain and contracting skeletal muscle^[30]). Quantitative assessment of oxygen transport in arterioles has revealed a large, and as yet unresolved, discrepancy between observed oxygen flux from arterioles and that predicted by computational modeling, whereas in capillaries, there is good agreement between experimental findings and predictions. Further and more complete experimental measurements are needed to resolve the findings in arterioles.

Based on recent trends in the field, in the near future one can expect increased efforts aimed at investigating the microcirculatory mechanisms by which artificial oxygen carriers, both HBOCs and PFCs, aid in the transport of oxygen to tissues. For HBOCs, this includes optimizing the oxygen-binding properties (P_{50} and Hill coefficient) of the modified hemoglobin and understanding the mechanism responsible for the pressor effect of some HBOCs. In recognition of the roles played by NO in microcirculatory oxygen transport, there will be more research focused on direct measurement of NO with improved spatial and temporal resolution, so that its involvement in oxygen supply and demand can be delineated more clearly. There will be increased emphasis on studying microcirculatory oxygen transport in animal models of disease and other pathological conditions, so that detailed mechanisms can be elucidated for future applications to clinical medicine.

ACKNOWLEDGMENTS

The research of the author, which is included in this article, was supported in part by grants HL18292 and HL79087 from the National Heart, Lung, and Blood Institute of the National Institutes of Health. The author thanks Dr. Aleksander S. Popel and Dr. Aleksander S. Golub for discussions that have helped to clarify many of the areas covered in this article.

ARTICLES OF FURTHER INTEREST

Mass Transfer in Tissue Function: Roles; Tissue Engineering of Microvascular Networks; Vascular Fluid Dynamics

REFERENCES

1. Krogh, A. The number and distribution of capillaries in muscles with calculations of the oxygen pressure head necessary for supplying the tissue. *J. Physiol.* **1919**, *52*, 409–415.
2. Duling, B.R.; Berne, R.M. Longitudinal gradients in periarteriolar oxygen tension. A possible mechanism for the participation of oxygen in local regulation of blood flow. *Circ. Res.* **1970**, *27* (5), 669–678.
3. Boron, W.F.; Boulpaep, E.L. *Medical Physiology*; Saunders: Philadelphia, 2003.
4. Ellsworth, M.; Ellis, C.; Popel, A.; Pittman, R.N. Role of microvessels in oxygen supply to tissue. *News Physiol. Sci.* **1994**, *9*, 119–123.
5. Popel, A.S. Theory of oxygen transport to tissue. *Crit. Rev. Biomed. Eng.* **1989**, *17* (3), 257–321.
6. Secomb, T.W.; Hsu, R.; Dewhirst, M.W. Synergistic effects of hyperoxic gas breathing and reduced oxygen consumption on tumor oxygenation: a theoretical model. *Int. J. Radiat. Oncol. Biol. Phys.* **2004**, *59* (2), 572–578.
7. Pittman, R.N. Oxygen transport and exchange in the microcirculation. *Microcirculation* **2005**, *12* (1), 59–70.
8. Pittman, R.N. Influence of microvascular architecture on oxygen exchange in skeletal muscle. *Microcirculation* **1995**, *2* (1), 1–18.
9. Rovainen, C.M.; Wang, D.B.; Woolsey, T.A. Strobe epillumination of fluorescent beads indicates similar velocities and wall shear rates in brain arterioles of newborn and adult mice. *Microvasc. Res.* **1992**, *43* (2), 235–239.
10. Parthasarathi, A.A.; Japee, S.A.; Pittman, R.N. Determination of red blood cell velocity by video shuttering and image analysis. *Ann. Biomed. Eng.* **1999**, *27* (3), 313–325.
11. Wayland, H.; Johnson, P.C. Erythrocyte velocity measurement in microvessels by a two-slit photometric method. *J. Appl. Physiol.* **1967**, *22* (2), 333–337.
12. Ellsworth, M.L.; Pittman, R.N. Evaluation of photometric methods for quantifying convective mass transport in microvessels. *Am. J. Physiol.* **1986**, *251* (4 Pt 2), H869–H879.
13. Ellis, C.G.; Ellsworth, M.L.; Pittman, R.N. Determination of red blood cell oxygenation in vivo by dual video densitometric image analysis. *Am. J. Physiol.* **1990**, *258* (4 Pt 2), H1216–H1223.
14. Pittman, R.N.; Duling, B.R. Measurement of percent oxyhemoglobin in the microvasculature. *J. Appl. Physiol.* **1975**, *38* (2), 321–327.
15. Pittman, R.N.; Duling, B.R. A new method for the measurement of percent oxyhemoglobin. *J. Appl. Physiol.* **1975**, *38* (2), 315–320.
16. Baker, C.; Nastuk, W. *Microcirculatory Technology*; Academic Press: Orlando, 1986.
17. Swain, D.P.; Pittman, R.N. Oxygen exchange in the microcirculation of hamster retractor muscle. *Am. J. Physiol.* **1989**, *256* (1 Pt 2), H247–H255.
18. Popel, A.S.; Pittman, R.N.; Ellsworth, M.L. Rate of oxygen loss from arterioles is an order of magnitude higher than expected. *Am. J. Physiol.* **1989**, *256* (3 Pt 2), H921–H924.
19. Tsai, A.G.; Friesenecker, B.; Mazzoni, M.C.; Kerger, H.; Buerk, D.G.; Johnson, P.C.; Intaglietta, M. Microvascular and tissue oxygen gradients in the rat mesentery. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95* (12), 6590–6595.
20. Tsai, A.G.; Johnson, P.C.; Intaglietta, M. Oxygen gradients in the microcirculation. *Physiol. Rev.* **2003**, *83* (3), 933–963.
21. Vadapalli, A.; Pittman, R.N.; Popel, A.S. Estimating oxygen transport resistance of the microvascular wall. *Am. J. Physiol. Heart Circ. Physiol.* **2000**, *279* (2), H657–H671.
22. Ellsworth, M.L.; Popel, A.S.; Pittman, R.N. Assessment and impact of heterogeneities of convective oxygen transport parameters in capillaries of striated muscle: experimental and theoretical. *Microvasc. Res.* **1988**, *35* (3), 341–362.
23. Ellsworth, M.L.; Pittman, R.N. Arterioles supply oxygen to capillaries by diffusion as well as by convection. *Am. J. Physiol.* **1990**, *258* (4 Pt 2), H1240–H1243.
24. Secomb, T.W.; Hsu, R. Simulation of O₂ transport in skeletal muscle: diffusive exchange between arterioles and capillaries. *Am. J. Physiol.* **1994**, *267* (3 Pt 2), H1214–H1221.
25. Golub, A.S.; Pittman, R.N. Erythrocyte-associated transients in PO₂ revealed in capillaries of rat mesentery. *Am. J. Physiol. Heart Circ. Physiol.* **2005**, *288* (6), H2735–H2743.
26. Federspiel, W.J.; Popel, A.S. A theoretical analysis of the effect of the particulate nature of blood on oxygen release in capillaries. *Microvasc. Res.* **1986**, *32* (2), 164–189.
27. Federspiel, W.J.; Sarelius, I.H. An examination of the contribution of red cell spacing to the uniformity of oxygen flux at the capillary wall. *Microvasc. Res.* **1984**, *27* (3), 273–285.
28. Stein, J.C.; Ellis, C.G.; Ellsworth, M.L. Relationship between capillary and systemic venous PO₂ during nonhypoxic and hypoxic ventilation. *Am. J. Physiol.* **1993**, *265* (2 Pt 2), H537–H542.
29. Riess, J.G. Oxygen carriers (“blood substitutes”)—raison d’être, chemistry, and some physiology. *Chem. Rev.* **2001**, *101* (9), 2797–2920.
30. Pittman, R.N. Oxygen supply to contracting skeletal muscle at the microcirculatory level: diffusion vs. convection. *Acta Physiol. Scand.* **2000**, *168* (4), 593–602.



Microcomputed Tomography and Its Applications

Thomas Dufresne
Paula Chmielewski
Babul Borah

Procter & Gamble Pharmaceuticals, Mason, Ohio, U.S.A.

Andres Laib

Scanco Medical AG, Bassersdorf, Switzerland

INTRODUCTION

Microcomputed tomography, commonly referred to as μ CT, is a relatively new field in the area of non-destructive imaging. Like conventional computed tomography, μ CT is based on the collection of the projections of X rays through a specimen and the application of tomographic principles to reconstruct the 3-D structure of the specimen. The term micro denotes a scanning system that is much higher in resolution than conventional clinical scanners. Clinical tomographic scanners may have resolutions on the order of a millimeter or less. However, high-resolution μ CT scanners may have resolutions below five microns. This difference in resolution is what separates μ CT from conventional CT scanners. In this article, micro-computed tomography will refer to tomographic scanners that can produce images with resolutions higher than 100 microns.

With the large increase in system resolution in microcomputed tomography, there is also a concomitant decrease in the field of view that can be imaged. Historically, this has limited the use of these scanners to research areas for examination of various materials, biopsied biological tissue, or the direct examination of small animal models. Also, there are differences in the source-detector gantry. In μ CT systems where a non-living specimen is being examined, it is more convenient to rotate the specimen and keep the source and detector in a fixed position. Because there is less chance of the source and detector becoming misaligned, more precise measurements are possible.

Over the past two decades, a large variety of μ CT scanners have been designed for various uses. A primary application for μ CT is the examination of biopsied human samples or small animal models for medical and pharmaceutical research. Other applications include the examination of geological specimens for impurities and the creation of physical models through rapid prototyping methods. Most of these systems use conventional X-ray tubes to produce polychromatic radiation. However, synchrotron radiation is used for

very high-resolution imaging. These instruments produce monochromatic radiation using a large particle accelerator and are limited to a handful of laboratories around the world.

BACKGROUND

The mathematical foundation for using projection data to reconstruct an image stems back to 1917, when Dr. Johann Radon developed the Radon transform. This transform provided the mathematical framework for the reconstruction of 3-D images from a set of X-ray projection data. The first practical application of image reconstruction based on X-ray projections was attributed to Godfrey Hounsfield and Allen Cormack, who shared the Noble Prize in 1979 for their independent discovery.^[1,2] Shortly after this time, researchers began to explore the idea that computerized tomography could be realized at high resolution, allowing the unique examination of both human biopsies and small animal models. In 1980, William Oldendorf outlined some of the possible applications of high-resolution computerized tomography.^[3] The idea of using a fan beam collimator with a fixed source and detector was first introduced, which allowed the specimen to rotate in order to examine tissue in situ. This concept is now widely used in μ CT for the examination of bone biopsies for structural changes in the cancellous bone.

In 1984, Kujoory et al. created a unique system for high-resolution CT of a rat nephrogram.^[4] In their system, a horizontal fan beam of X rays was illuminated through the specimen, then through a slit collimator onto an X-ray film cassette that was continuously translating the film. The specimen was located on a turntable that had two fixed lead markers, allowing the center of the projection to be easily defined. The film cassette was then digitized offline. To avoid lengthy reconstruction times, several digitized points were averaged together, giving an effective detector resolution of 133 microns with a beam height of approximately 1 mm. Overall, this resulted in their

ability to obtain 1-mm-thick coronal sections through the rat with a lateral resolution of 150 μm . In the same year, Seguin et al.^[5] published results showing that they had achieved 50 μm resolution for a single CT slice. The advantage in their system was their use of a digital detector using an integrated array of silicon photo-diodes placed behind a scintillator.

The next improvement in μCT scanners occurred in 1989, when Feldkamp et al. published their work on the measurement of trabecular bone via 3-D computerized tomography.^[6] This scanner used cone beam tomography based on a practical algorithm developed by Feldkamp several years earlier.^[7] The system resolution was 50 microns, but this was achievable in all three dimensions (isotropic resolution). An image intensifier converted the signal to visible light, followed by a vidicon camera and frame grabber to convert the signal to a digital array. This configuration allowed the first utilization of direct 3-D measures to quantify aspects of the bone architecture derived from 3-D μCT data. Notably, Odgaard et al. had previously applied some of these architecture measures from 3-D bone samples obtained through serial sectioning.^[8]

In 1995, Kinney et al. showed the power of using synchrotron radiation for the examination of in vivo bone samples from a rat hind limb.^[9] In this study, they performed longitudinal imaging of the tibias of rats. The high-quality images clearly show the loss of trabecular bone after ovariectomy. The high photon count and the use of monochromatic radiation have led to much wider use of this technique for the examination of bone.^[10,11] Most notably, Ritman et al. have done extensive work in the examination of vascular structure through the use of synchrotron imaging.^[12] Nuzzo et al. showed a high correlation between synchrotron imaging and other modalities (backscattered SEM, micro-radiography) for assessing mineralization.^[13]

In 1997, Scanco Medical (Bassersdorf, Switzerland) produced the first commercial system for non-destructive evaluation of bone architecture.^[14] Unlike Feldkamp's system, the Scanco $\mu\text{CT}20$ used only a 1-D detector and created single-slice reconstructions and then stacked these slices to obtain a 3-D data set. Similar to Feldkamp's system, the generated data was isotropic by matching the slice separation to the in-plane resolution. The Scanco $\mu\text{CT}20$ system included extensive analysis software for a broad suite of 3-D imaging measurements, which were automatically calculated and stored in a database. Hildebrand and R uegsegger developed direct measures of trabecular thickness, trabecular number, and trabecular separation.^[15] Hildebrand and R uegsegger also developed a novel approach for calculating the ratio of plates to rods using a method called structure model index (SMI).^[16] This measurement provides a quantitative measure of how plate- or rodlike a structure is, yielding

a value of zero for a fully platelike structure to a value of three for a complete rodlike structure.

As technology continues to improve, μCT is becoming more common in both academia and industry. Today there are several commercially available systems including in vivo instruments for examining small animal models.^[17,18] Most scanners today are of the cone beam type, allowing for much faster scan times and lower dose to the animal under investigation. The improvement of 2-D detector arrays and faster electronics has allowed high-resolution scans exceeding 1 gigabyte of data to be acquired in only a few minutes.

BASIC X-RAY AND COMPUTERIZED TOMOGRAPHY THEORY

Since μCT is very similar to conventional CT, a short theoretical framework is presented for computerized tomography. This provides an overview of the generation of X rays and their attenuation in material, as well as the theory behind image reconstruction. More thorough discussion of tomographic theory and the physics of X-ray generation can be found in several excellent books.^[19,20]

In general, most μCT systems have a similar design. Typically, they generate X rays with a microfocal X-ray tube. The X rays are usually filtered to remove very low-energy X rays. The X rays pass through and are attenuated by the specimen, which is mounted on a spindle and rotated to collect different projections. The X rays pass through a scintillator, which converts the X rays into visible light, which is then fed into a solid-state detector (e.g., CCD array). The data are allowed to collect in the detector for a certain amount of time, referred to as the integration time, and then read out and stored on the computer. A fine stepper motor is used to rotate the sample to collect the projections at precise positions. The set of projections is then processed on a computer workstation to produce the cross-sectional image data. The spindle is translated and the projection data are collected for the next stack of data. The whole set of cross-sectional images then makes up the complete 3-D image array. The value stored in a pixel of a CT image is ideally only determined by the local attenuation coefficient of the material at that pixel's location. The components of computerized tomography are described in greater detail later in this article.

GENERATION OF X RAYS

A typical X-ray tube consists of a negatively charged cathode and a positively charged anode. A heated



filament at the cathode and a large voltage potential established between these two ends results in the high-speed flow of electrons from the cathode to the anode. As the electrons strike the anode surface, they are deflected in different ways and their kinetic energy is converted to radiation through the emission of photons. Most of this emitted radiation is in the form of heat (~ 99%). The remaining energy is transmitted as X rays, which will have a wide variation in wavelength (energy) because of the variation in random interactions with the electron beam and the anode surface. A focusing cup is used at the cathode to keep the electron beam focused on a small area in the anode target. The effective focal spot size of the X-ray tube is a function of the focusing cup, the angle of the anode surface with the cathode, and the number of electrons striking the surface of the anode. Modern μ CT scanners can obtain a focal spot size of five microns.

Generated X rays are of two types. The most common type is called Bremsstrahlung X rays (also called general radiation). This type of radiation is a result of the deflection of an electron as it approaches the nucleus of an atom on the anode. The coulomb field of the electron interacts with the electromagnetic field of the atom, and part of the energy of the electron is converted to Bremsstrahlung and emitted as an X ray. Therefore, the resulting energy is, on average, less than the maximum keV of the system, resulting in a broad spectrum of all emitted X rays. The second type of radiation produced is called characteristic radiation. This radiation is produced when the electron bombardment causes an electron to be ejected from the inner orbits of the target atoms. Depending on the binding energy of that electron shell for that type of atom, a photon of a very specific energy is released as the electron space is refilled. Therefore, the energy is considered characteristic for an atom's electron shell. It is common for μ CT systems to use aluminum filters in front of the source to suppress the low-energy X rays where many characteristic X-ray spikes tend to appear in the transmitted X-ray energy spectrum.

ATTENUATION OF X RAYS

Although there are several ways that X rays can be attenuated, the two most common in medical imaging are: 1) photoelectric effect; and 2) Compton scattering. In the photoelectric effect, the X ray striking an atom within the specimen is attenuated by the transference of energy to eject a tightly bound electron from orbit. The photoelectric effect is more dominant when the impending X ray is of an energy just slightly greater than the binding energy of the electron. Also, it is more likely to occur when the atomic number of the element in the specimen is high. The photoelectric effect

provides good contrast in tissues having elements of different atomic numbers (e.g., compact bone vs. marrow).

Compton scattering is similar to the photoelectric effect in that it also results from an X-ray photon ejecting an electron from its orbit. The difference is that in Compton scattering, the electrons are loosely bound so they need virtually no energy to be ejected. However, the kinetic energy of this recoil electron causes an attenuation of the original photon. Also, the direction of the scattered photon is changed from its original forward direction. The higher the energy of the photon, the less the photon will scatter. Since the angles of the photons are small and much of the energy is retained by the photon, it is possible that the X-ray photon could pass into the wrong detector element, causing noise. This effect can usually be mitigated by the presence of an X-ray absorbing collimator in front of the detector to block stray photons. In bone, Compton scattering is the dominant attenuator at 60 keV and higher, while the photoelectric effect is dominant between 20 and 60 keV.

For a given thickness of known material, a certain percentage of photons will be absorbed. Therefore, the number of X rays reaching the detector after passing through a material of thickness x with an attenuation coefficient of μ can be described by the exponential equation:

$$I = I_0 e^{-\mu x} \quad (1)$$

I_0 is usually measured through air during the scan, since μ (air) is (virtually) 0. This, coupled with the dark current measure (detector signal with the X-ray shutter closed), allows accurate measurement of the relative X-ray attenuation of a given projection. This attenuation is a function of not only material composition and material density, but also of X-ray beam energy.

When polychromatic radiation is used, an artifact called beam hardening is introduced into the imaging reconstruction. This is due to the preferential absorption of low-energy (soft) X-ray photons as the X rays pass through the sample, since attenuation is higher for lower X-ray energy, given the energy range and materials used in μ CT. As the X rays reach the center portion of the specimen, many of the low-energy X rays have been attenuated, and the beam has a mean keV that is higher, resulting in less attenuation of photons. After reconstruction, this artifact appears as a cupping effect in the center of the image. Beam-hardening is more pronounced when there is substantial attenuating material on the periphery of the image (such as the skull surrounding the brain). Synchrotron avoids this problem by using monochromatic energy. This allows the quantification of small changes in density using this type of instrument.

COLLECTION OF PROJECTION DATA

In CT imaging, the projection data are collected over many angles at a single point along the length of the specimen. The specimen (or the gantry) is translated to a new position and the collection of projection data is repeated. After translating the entire specimen length, the reconstructed data are stacked together to yield the final 3-D image. The projection data can be collected as parallel beams, fan beams, or cone beams. Parallel beam was the first technique developed and it is the simplest to reconstruct. A collimated beam is passed through the specimen in a given direction to create a set of projection points representing a single projection. The source detector is then rotated a small increment (or the specimen is rotated) and another projection is collected. This is repeated for a 180-degree arc about the specimen (slightly greater for fan- or cone-beam projection). In many μ CT systems, the specimen is rotated a full 360 degrees to obtain a field of view that is twice the size of the detector. This is achieved by placing the specimen's axis of rotation along one edge of the detector. Therefore, at any one time, a projection of one-half of the object is created. After 360 degrees of rotations, both halves of the specimen have been subjected to projections over 180 degrees, allowing the reconstruction of the entire specimen.^[21] In theory, the number of projections used should be greater than the number of sample points for every projection by a factor of $\pi/2$ ^[22] (e.g., for a 512×512 image, one should use approximately 800 projections). However, due to mechanical limits of the stepper motor, it is common practice in μ CT to have the projections and the number of sample points per projection be roughly equivalent, with little image degradation.

Using fan-beam geometry, a line of projection data can be collected without translating the source. Unlike parallel projection, the X rays for an entire projection are emanated from a single point (depending on the focal spot size of the X-ray tube, the width of this point will change). A one-dimensional array of detectors is used to collect the data. Collimation is used at the source to limit the slice thickness, and at the detector end to limit scatter. One advantage of using a fan of X rays is that the image can be magnified by moving it closer to the source. If a 2-D detector is used in this collection scheme, it becomes cone-beam geometry. This technique, because of its rapid ability to acquire 3-D data, is generally used in most modern μ CT scanners. Cone angle is limited to insure that the X rays from every projection pass through each point in the 3-D volume of interest.^[23] Although collimation and reconstruction become more difficult for cone-beam geometries, the acquisition time is decreased by a magnitude, while maintaining high image quality.

IMAGE RECONSTRUCTION

Given a 2-D image, I , there is known relationship between the 1-D Fourier transform of a projection through I , at angle α , ($P_\alpha(I)$) and the corresponding 2-D Fourier transform of I , $F\{I\}$. $P_\alpha(I)$ will form a line, passing through the origin, at angle α , in $F\{I\}$. This is referred to as the Fourier Slice Theorem and can be easily derived mathematically.^[24] By rotating the projection angle, we can begin to fill in spatial frequency information in all directions. This becomes a series of radial lines passing through the origin of fourier space. At low frequencies, where all the radial lines are converging, we have an abundance of information. At high frequency, we have a deficit of information. If the entire fourier space was completely filled in, we could transform back to image space and have the reconstructed image, with no loss of information. To fill in the fourier space, we must apply a weighting scheme that emphasizes the high-frequency information, so as to approximate an interpolation of this region of fourier space. This interpolation process can be visualized as a replacement of the lines in the fourier space with pie wedges, so that they neatly fill the fourier space.^[24]

The most common way that the fourier space is interpolated and transformed back into an image is through a reconstruction method referred to as convolution-backprojection. Since a multiplication in fourier space corresponds to a convolution in real space, the weighting is done by convolution of the projection data in real space. The final step is called backprojection, which sums the weighted projections together by smearing of the data along the direction of the projection. Many modern μ CT units utilize some form of convolution-backprojection for image reconstruction. However, many sophisticated improvements have been made to the basic algorithm to handle parallel, fan-beam, and cone-beam reconstruction.

PRECISION AND ACCURACY OF MICRO-CT

To determine the accuracy and precision of the micro CT instrumentation, Chmielewski et al. developed two phantoms and methods to perform a series of repeated measurements on a Scanco μ CT20 scanner.^[25] The first phantom consisted of a single 128.13-micron-diameter aluminum wire. This wire was measured by both μ CT and light microscopy, which was calibrated with a NIST traceable standard. The mean diameter measured on the μ CT system was 125.5 microns, with a standard deviation of 0.7 microns, resulting in only a 2.1% error from the calibrated microscopy measurement. The length of the wire measured by μ CT had only a 0.3% error from the calibrated microscopy



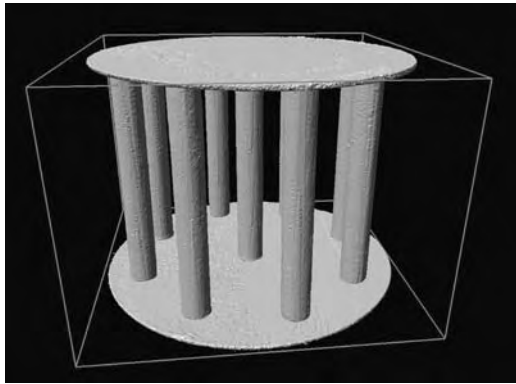


Fig. 1 Aluminum tube phantom used to validate the μ CT scanner accuracy.

measurement. These results confirm a high degree of accuracy of this instrument for measuring changes in structure. To measure the precision of the instrument, a second aluminum phantom was used. This phantom consisted of eight concentric aluminum tubes (250-micron-diameter) with a ninth tube directly in the center, as shown in Fig. 1. Repeat measures showed a very high reliability in the precision as well. The total volume measured across the tubes had a coefficient of variation of approximately 2% with repositioning of the same specimen over multiple scans. The total surface area measurement of the tubes had a coefficient of variation of less than 0.5%. The physical phantom measurements provided a means of monitoring the precision of the image data across entire studies and showed there was little drift over the measurements over an entire year.

Muller et al. reported high correlations between micro CT measurements with conventional histomorphometric measures.^[26] Bone volume/total volume, bone surface/total volume, and trabecular separation all exhibited regression coefficients above 0.9. The mean percentage differences in absolute differences between the μ CT and histology measurements ranged between 2.5 and 6.1%. The combination of high regression and low error between conventional histomorphometry and μ CT provide further evidence of the accuracy and precision of this modality.

3-D MEASUREMENTS WITH μ CT

With the advent of 3-D μ CT, parameters such as trabecular thickness and separation can be estimated directly from the 3-D data.^[15] Using only 2-D slices, these important parameters of bone change could only be estimated based on a simple model that all components of the measured trabeculae form simple rods

or plates.^[27] Unfortunately, this model is somewhat weak because bone fits neither of these simple descriptions. Furthermore, if the morphometry of bone is shifting between a rodlike structure to a more platelike structure, it could lead to a misrepresentation that these trabecular measurements are changing when in fact they are not. Beyond these measurements, the 3-D data afforded by μ CT technology allow additional measurements such as connectivity density (# of connections/mm³), degree of anisotropy (the orientation of the trabecular architecture), and cortical porosity (% of holes in the cortical bone) and cortical thickness.

Another advantage of 3-D μ CT is that it can serve as input to a finite element model. A finite element model allows virtual mechanical testing of a sample based on its 3-D data set. Micro CT provides both density information as well as high-resolution structural data, making ideal input for such a system. The voxels in the 3-D data set are replaced by brick elements. These brick elements have a modulus of elasticity and a Poisson ratio associated with each element. The modulus of elasticity can often be derived from the CT values and represents the stiffness of the brick element in each direction. In the finite element model of bone, one side of the model is generally constrained and an opposing side is displaced a predetermined distance. Based on the material properties of the brick elements, this introduces displacements in each of the elements. The displacement of each element is governed by a set of equilibrium equations. The simultaneous solution of these equations produces a strain distribution across the specimen. The resulting distribution can then be compared against known maximum strains for the material in question, allowing prediction of regions susceptible to fracture.^[28]

APPLICATIONS OF MICRO-CT TECHNOLOGY

Microcomputed Tomography in Osteoporosis Research

Osteoporosis is a widespread disease affecting mainly the elderly, where the bone becomes fragile and prone to fracture. By the definition provided by the World Health Organization at the Consensus Development Conference in 1991, osteoporosis is a skeletal disease characterized by low bone mass and microarchitectural deterioration of bone tissue with a consequent increase in bone fragility and susceptibility to fracture.^[29] The idea that the microarchitecture of the bone plays an important role in the likelihood of fracture has led to great interest in quantifying changes in this complex structure using μ CT technology.

To better illustrate the changes in cancellous bone that occur with disease progression, contrast the change in trabecular structure between a normal 52-year-old woman and an osteoporotic 84-year-old shown in Fig. 2. These images are computer renderings of data collected on a benchtop μ CT scanner (μ CT20, Scanco Medical, Bassersdorf, Switzerland). It is evident from these images that there is not only a change in overall bone mass, but also a change in the architecture of the bone. The osteoporotic bone has much thinner rods with more perforations. Also, there is a preferential loss of horizontal trabeculae, which run perpendicular to the load direction in the vertebra. It is believed that the absence of these horizontal trabeculae leads to buckling effect of the vertical trabeculae during normal loading, leading to fracture.^[30]

In recent years, there has been an increasing number of compelling results demonstrating the importance of architectural examination of the bone through the use of μ CT technology to provide insights into osteoporosis treatment. In the ovariectomized minipig vertebrae, we showed that there was a preservation of architectural parameters in risedronate-treated animals, independent of bone density.^[31] Ulrich et al. showed the relative contribution of trabecular architecture to mechanical strength through the examination of iliac crest biopsies taken from a wide

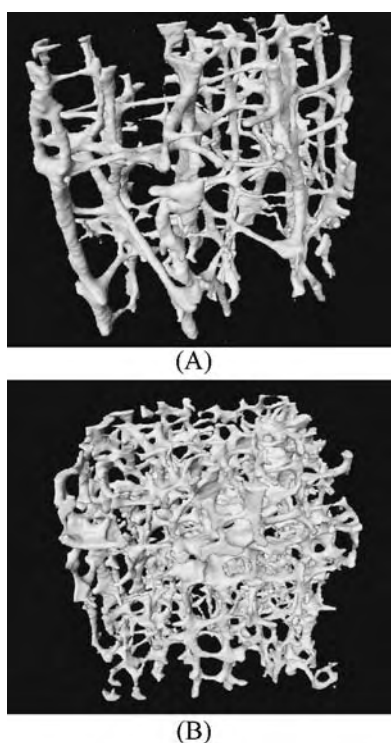


Fig. 2 μ CT rendering of female vertebral cores from (A) osteoporotic 84-year-old and (B) normal 52-year-old.

age range of patients.^[32] Using paired iliac crest biopsies, we have also used μ CT technology to show that risedronate preserved trabecular bone in both an early postmenopausal group of women as well as an older osteoporotic group of women^[33,34] Micro-CT studies of iliac crest biopsies before and after treatment with PTH^[1-32] a bone anabolic agent, showed an improvement in bone architecture through an increase in trabecular connectivity and cortical thickness.^[36] These important μ CT studies had illustrated differences between antiresorptives like risedronate, and bone anabolics like PTH^[1-32] in their effect on bone microarchitecture and their mechanism of action for reducing the probability of fracture.^[35]

TRABECULAR BONE ARCHITECTURE ASSESSMENT AS A NOVEL SCREEN IN ANABOLIC RESEARCH

Due to the high resolution of the images acquired with μ CT, it is particularly well-suited for the examination of trabecular architecture in small animal models. We compared treatment effects of two bone anabolics, hPTH^[1-32] and Fluprostenol (a prostaglandin FP receptor agonist) in a severely osteopenic rat model.^[37,38] As shown in Fig. 3, three-dimensional μ CT showed that both agents significantly increased trabecular bone volume in the lumbar vertebra compared to the ovariectomized group (Fluprostenol: 45.3%, PTH: 55.6%, Final Ovx: 24.4%). However, the therapy-induced changes in trabecular architecture were greatly different for the two therapies, as can be visualized in Fig. 3. While PTH caused extensive trabecular thickening (Fig. 3D), the formation of new trabecular bone is visually evident with Fluprostenol (Fig. 3C) and resulted in trabecular bone similar to the normal intact group (Fig. 3A). This is best reflected by the changes in connectivity density measured between the two anabolic agents. In the high-dose Fluprostenol group, the connectivity density was more than twice that found in the PTH group (Fluprostenol: 84.5 conn./mm³, PTH: 37.4 conn./mm³). This study underscores the importance of being able to measure beyond just bone mass and examine the complete 3-D trabecular architecture of the bone.

3-D IMAGING OF VESSELS FOR THE UNDERSTANDING OF ANGIOGENESIS

The visualization and quantification of new blood vessel formation is important in pharmaceutical drug development of angiogenic drugs as well as understanding the natural occurrence of angiogenesis



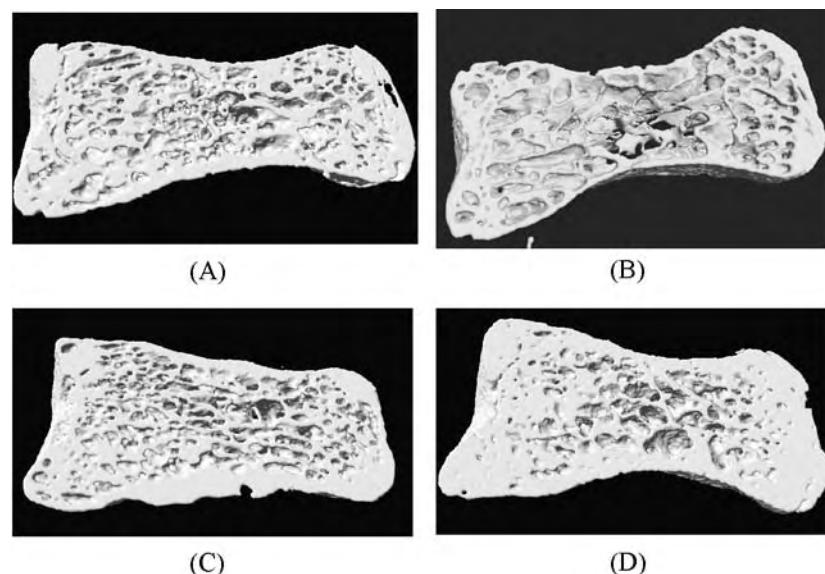


Fig. 3 μ CT rendering of rat vertebra cross section of (A) normal, (B) ovariectomized, (C) fluprostenol, and (D) PTH-treated vertebrae.

in the body. In pharmaceutical research, the rat hind limb ischemia model is commonly used to measure the efficacy of proangiogenic compounds.^[39] In this model, the femoral artery of the rat is ligated, which causes ischemic regions to develop distal to the ligation. Angiogenic compounds are administered to the animal for a fixed number of days. Prior to euthanasia, the rat is injected with microspheres at a fixed rate. The tissue of the ischemic region is then examined and the count of the microspheres embedded in the tissue provides an accurate assessment of the flow of blood in the ischemic region. Changes in blood flow rates are then the endpoints for efficacy of the compound. However, although this is an important endpoint, it does not address whether the increased blood flow is a result of new collateral growth around the ligation, or a remodeling of existing vessels to increase their flow.

Micro CT technology provides an excellent method of visualizing and quantifying vascular changes in small animal models.^[12] The natural X-ray attenuation of the vessels is very low, therefore, a contrast agent needs to be injected into the vessels prior to imaging with μ CT. For the experiments in our lab, we have chosen the use of Microfil[®] as a contrast agent. The contrast agent is injected into the aorta, postmortem, while maintaining physiological blood pressure. Once the Microfil[®] has hardened overnight, tissues are harvested and fixed for scanning. Figure 4A shows a normal rat hind limb injected with Microfil[®]. The loss of active blood vessels is quite apparent immediately after the ligation has occurred (Fig. 4B). After two weeks of treatment with a proangiogenic compound, many of the vessels are restored (Fig. 4C). In addition to examining collateral vascular structure, it can be

applied to a variety of organs where the 3-D vascular structure needs to be elucidated.

MEASURING COATING THICKNESS IN 3-D ON ENTERIC COATED TABLETS

Ulcerative colitis is characterized by an inflammation of the lining of the rectum and colon. Therefore, in order to treat this condition, it is important that the drug be delivered in proximity to the site of the ulceration. Enteric coating allows the tablet to pass through the stomach and into the colon where it will dissolve at a pH of seven or greater. Since the stomach has greater acidity, it will pass through to the colon intact. Equally important is that the enteric coating has a uniform thickness and little deviation between thicknesses of tablets within a lot.

In our study, we examined 400 mg Asacol[®] tablets. These tablets are prescribed for ulcerative colitis and contain an iron oxide-based enteric coating, making the coating sufficiently attenuating for 3-D micro-computed tomography. The goal of this study was to use μ CT imaging as a calibration standard for near infrared analysis and conventional scanning electron microscopy.^[40] This study examined tablets that were being coated with a newly designed system. Custom designed software was developed so that the distinct regions of the upper and lower dome as well the girdle could be independently measured. To get a broad range of sample thicknesses, the tablets were coated for 45, 70, and 90 minutes (finished form). Figure 5 shows the coating thickness for two different time points. This application has been used repeatedly, not only

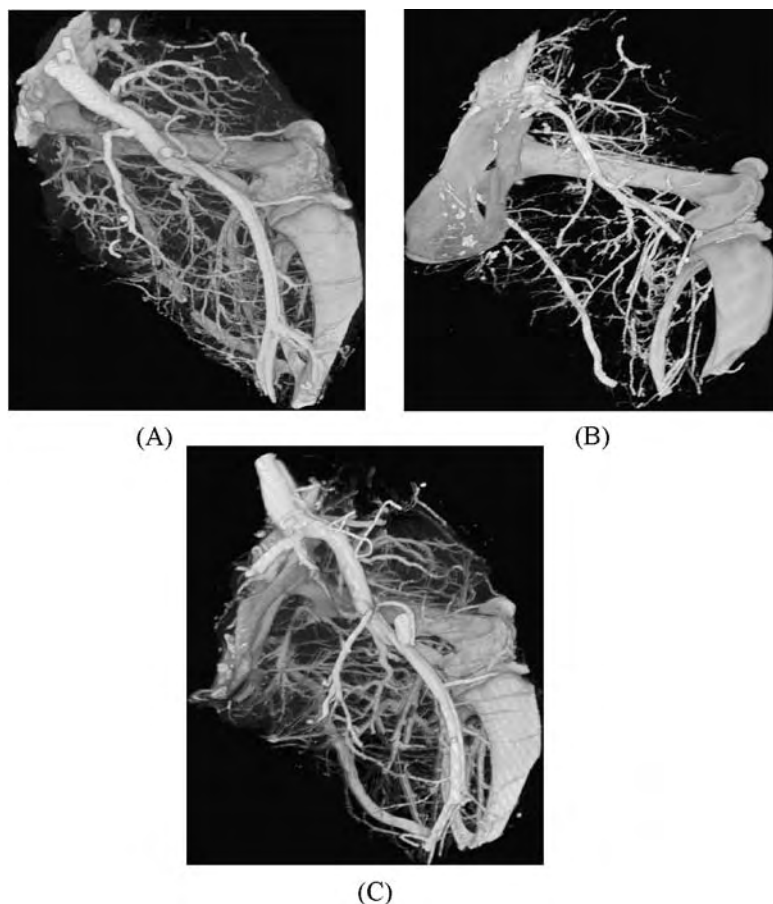


Fig. 4 μ CT rendering of vessels in rat hind limb ischemic rat model; (A) baseline, (B) one day after ligation of femoral artery, and (C) two weeks of treatment after ligation.

calibrating other modalities, but also troubleshooting issues with coating distribution.

CONCLUSION

Micro CT technology is a continually evolving field that is providing new insight into many fields of re-

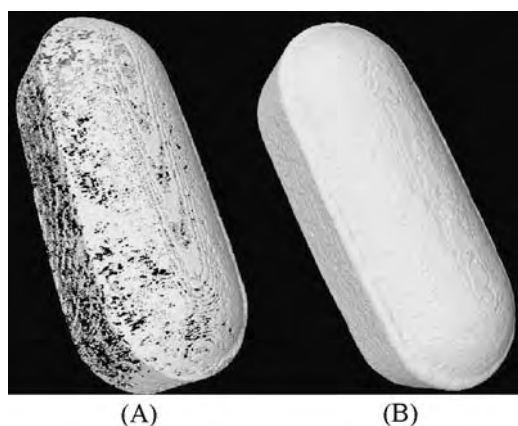


Fig. 5 μ CT analysis of enteric coated tablets; (A) 3-D rendering of coating after 45 minutes in coater, and (B) 3-D rendering of coating after 90 minutes in coater.

search. As detector performance continues to increase with improved resolution and better signal to noise, the image quality of the reconstructed images will continue to improve. Reconstruction times, which are largely a function of computer CPU speed, will continue to speed up, allowing for large reconstructed images to be created concurrent with data collection. In the area of bone research, many new in vivo scanners are being produced, which allow excellent delineation of trabecular bone in longitudinal studies of small animal models. This, in turn, is allowing sophisticated modeling of strength changes in bone using finite element analysis.

ACKNOWLEDGMENTS

I would like to thank Mses. Cindy Blanton and Elaine Eby-Wilkens for their work with using Microfil[®] for vessel identification, Dr. Mark Lundy and Ms. Katie Combs for their work in anabolic screening, Dr. Tom Cambron and Mr. Van Bullman for their efforts in near infrared and microscopy of the enteric coated tablets, and Dr. Sheng Liu for her assistance with the editing of this manuscript.

ARTICLES OF FURTHER INTEREST

Functional MRI: Applications; Medical Imaging, 3-D; Magnetic Resonance Imaging; Optical Coherence Tomography; Positron Emission Tomography (PET); Real Time Tomographic Reflection

REFERENCES

- Hounsfield, G.N. A method of and apparatus for examination of a body by radiation such as X-ray or gamma radiation. *Pat. Specif.* **1972**, 1283915.
- Cormack, A.M. Representation of a function by its line integrals, with some radiological applications. *J. Appl. Phys.* **1963**, *34*, 2722–2727.
- Oldendorf, W.H. Some possible applications of computerized tomography in pathology. *J. Comput. Assist. Tomogr.* **1980**, *4* (2), 141–144.
- Kujoory, M.A.; Hillman, B.J.; Barrett, H.H. High-resolution computed tomography of the normal rat nephrogram. *Invest. Radiol.* **1980**, *15*, 148–154.
- Seguin, F.H.; Burstein, P.; Bjorkholm, P.J.; Homburger, F.; Adams, R.A. X-ray computed tomography with 50- μm resolution. *Appl. Opt.* **1985**, *24*, 4117–4123.
- Feldkamp, L.A.; Goldstein, S.A.; Parfitt, M.A.; Jesion, G.; Kleerekoper, M. The direct examination of three-dimensional bone architecture in vitro by computed tomography. *J. Bone Miner. Res.* **1989**, *4* (1), 3–11.
- Feldkamp, L.A.; Davis, L.C. Practical cone-beam algorithm. *J. Opt. Soc. Am., A* **1984**, *1*, 612–619.
- Odgaard, A.; Andersen, K.; Ullerup, R.; Frich, L.H.; Melsen, F. Three-dimensional reconstruction of entire vertebral bodies. *Bone* **1994**, *15* (3), 335–342.
- Kinney, J.H.; Lane, N.E.; Haupt, D.L. In vivo, three-dimensional microscopy of trabecular bone. *J. Bone Miner. Res.* **1995**, *10* (2), 264–270.
- Bonse, U.; Busch, F.; Gunnewig, O.; Beckman, F.; Pahl, R.; Delling, G.; Graeff, HahnW. 3D computed X-ray tomography of human cancellous bone at 8 μm spatial and 10^{-4} energy resolution. *Bone Miner.* **1994**, *25*, 25–38.
- Peyrin, F.; Salome, M.; Cloetens, P.; Laval-Jeantet, A.M.; Ritman, E.; Ruegsegger, P. Micro-CT examinations of trabecular bone samples at different resolutions: 14,7,2 micron level. *Technol. Health Care* **1998**, *6*, 391–401.
- Ritman, E.L.; Bolander, M.E.; Fitzpatrick, L.A.; Turner, R.T. Micro-CT imaging of structure-to-function relationship of bone microstructure and associated vascular involvement. *Technol. Health Care* **1998**, *6*, 403–412.
- Nuzzo, R.; Lafage-Proust, M.H.; Martin-Badosa, E.; Boivin, G.; Thomas, T.; Alexandre, C.; Peyrin, F. Synchrotron radiation microtomography allows the analysis of three-dimensional microarchitecture and degree of mineralization of human iliac crest biopsy specimens: Effects of etidronate treatment. *J. Bone Miner. Res.* **2002**, *17*, 1372–1382.
- Ruegsegger, P.; Koller, B.; Muller, R. A microtomographic system for the nondestructive evaluation of bone architecture. *Calcif. Tissue Int.* **1996**, *58*, 24–29.
- Hildebrand, T.; Ruegsegger, P. A new method for the model-independent assessment of thickness in three-dimensional images. *J. Microsc.* **1996**, *185*, 67–75.
- Hildebrand, T.; Ruegsegger, P. Quantification of bone microarchitecture with the structure model index. *Comput. Methods Biomech. Biomed. Eng.* **1997**, *1*, 15–23.
- Salmon, P.L.; Buelens, E.; Sasov, A. Radiation Dosimetry, Anaesthesia and Welfare of Rodents in Micro-CT In Vivo Scanning. Skyscan, Aartselaar, Belgium, Proceedings of the IBMS-JSBMR 2003, Osaka, Japan, June 3–7, 2003. Poster 130W.
- David, V.; Laroche, N.; Boudignon, B.; Lafage-Proust, M.H.; Alexandre, C.; Ruegsegger, P.; Vico, L. Noninvasive in vivo monitoring of bone architecture alterations in hindlimb-unloaded female rats using novel three-dimensional microcomputed tomography. *J. Bone Miner. Res.* **2003**, *18* (9), 1622–1631.
- Curry, T.S.; Dowdey, J.E.; Murry, R.C. *Christensen's Physics of Diagnostic Radiology*, 4th Ed.; Lea & Febiger: Philadelphia, PA, 1990.
- Kak, A.C. *Principles of Computerized Tomographic Imaging*; IEEE Press: New York, NY, 1988.
- Ritman, E.L.; Borah, B.; Dufresne, T.E.; Phipps, R.J.; Sacha, J.P.; Jøsgansen S.M.R.T. 3-D Synchrotron Micro-CT allows unique insight of changes in bone quality. *JBMR* **2002**, *17* (Suppl 1), S415.
- Kak, A.C. Aliasing Artifacts and Noise in CT Images. In *Principles of Computerized Tomographic Imaging*; IEEE Press: New York, NY, 1988; 183–186.
- Smith, B.D. Image reconstruction from cone-beam projections: Necessary and sufficient conditions and reconstruction methods. *IEEE Trans. Med. Imag.* **1985**, *MI-4*, 14–25.
- Kak, A.C. Algorithms for Reconstructing with Non-diffracting Sources. In *Principles of Computerized Tomographic Imaging*; IEEE Press: New York, NY, 1988; 56–62.
- Chmielewski, P.A.; Dufresne, T.E.; Borah, B.; Combs, K.H.; Lundy, M.W. Validation of a Micro-CT System for the 3-D Measurement of Bone. *JBMR* **2001**, *16* (Suppl 1), p. S461.
- Muller, R.; Van Campenhout, H.; Van Damme, B.; Van Der Perre, G.; Dequeker, J.; Hildebrand, T.; Ruegsegger, P. Morphometric analysis of human bone biopsies: A quantitative structural comparison of histological sections and micro-computed tomography. *Bone* **1998**, *23* (1), 59–66.
- Parfitt, A.M. The Stereologic Basis of Bone Histomorphometry. Theory of Quantitative Microscopy and Reconstruction of the Third Dimension. In *Bone Histomorphometry. Techniques and Interpretations*; CRC Press: Boca Raton, FL, 1983; 53–87.
- Ulrich, D.; Hildebrand, T.; Van Rietbergen, B. The Quality of Trabecular Bone Evaluated with Micro-CT, FEA and Mechanical Testing. In *Bone Research in Biomechanics*; IOS Press: Amsterdam; 1997; 97–112.

29. Consensus Development Conference. Prophylaxis and treatment of osteoporosis. *Am. J. Med.* **1991**, *90* (1), 107–110.
30. Mosekilde, L. Vertebral structure and strength in vivo and in vitro. *Calcif. Tissue Int.* **1993**, *53* (suppl), S121–S126.
31. Borah, B.; Dufresne, T.E.; Chmielewski, P.A.; Gross, G.J.; Prenger, M.C.; Phipps, R.J. Risedronate preserves trabecular architecture and increases bone strength in vertebra of ovariectomized minipigs as measured by 3-D micro-computed tomography. *J. Bone Miner. Res.* **2002**, *17*, 1139–1147.
32. Ulrich, D.; van Rietbergen, B.; Laib, A.; R uegsegger, P. The ability of 3D structural indices to reflect mechanical aspects of trabecular bone. *Bone* **1999**, *25* (1), 55–60.
33. Dufresne, T.E.; Chmielewski, P.A.; Manhart, M.A.; Johnson, T.D.; Borah, B. *Risedronate Preserves Bone Architecture in Early Postmenopausal Women in 1 Year as Measured by Three-Dimensional Microcomputed Tomography*; Calcified Tissue International, Springer-Verlag: New York. *Calc. Tissue Int.* **2003**, *73*, 423–432.
34. Borah, B.; Dufresne, T.E.; Chmielewski, P.A.; Prenger, M.C.; Eriksen, E.F. Risedronate (Ris) preserves bone architecture in osteoporotic postmenopausal women as measured by 3-D microcomputed tomography. *J. Bone Miner. Res.* **2001**, *16* (suppl 1), S218.
35. Borah, B.; Gross, G.J.; Dufresne, T.E.; Smith, T.S.; Cockman, M.D.; Chmielewski, P.A.; Lundy, M.W.; Hartke, J.R.; Sod, E.W. Three-dimensional microimaging (MrmI and μ CT), finite element modeling and rapid prototyping provide unique insights into bone architecture in osteoporosis. *Anat. Rec.* **2001**, *265*, 101–110.
36. Dempster, D.W.; Cosman, F.; Kurland, E.S.; Zhou, H.; Nieves, J.; Woelfert, L.; Shane, E.; Plavetic, K.; Muller, R.; Bilezikian, J.; Lindsay, R. Effects of daily treatment with parathyroid hormone on bone architecture and turnover in patients with osteoporosis: A paired biopsy study. *J. Bone Miner. Res.* **2001**, *16*, 1846–1853.
37. Borah, B.; Dufresne, T.E.; Chmielewski, P.A.; Pierce, S.A.; Hartke, J.R.; Lundy, M.W. Fluprostenol, a Prostaglandin fp Agonist, Improves 3-Dimensional Trabecular Architecture in Rat Vertebrae as Measured by Micro-computed Tomography. Proceedings of 21st ASBMR, St. Louis, Missouri, 1999; S401. *J. Bone Miner Res.*
38. Dufresne, T.E. Segmentation techniques for analysis of bone by three-dimensional computed-tomographic imaging. *Technol. Health Care* **1998**, *6*, 351–359.
39. Herzog, S.; Sager, H.; Khmelevski, E.; Deylig, A.; Ito, W.D. Collateral arteries grow from preexisting anastomoses in the rat hindlimb. *Am. J. Physiol, Heart Circ. Physiol.* **2002**, *283* (5), H2012–2020.
40. Dufresne, T.E.; Borah, B.; Chmielewski, P.A.; Combs, K.H.; Loudy, D.E.; Lundy, M.W.; Samuelsson, S.J. 3-D Micro-Computed Tomography in Pharmaceutical Research. *J. Scan. Microsc.* **1999**, *21* (2), 129–130.



Microelectromechanical Systems (MEMS) Manufacturing

Marc Madou

University of California, Irvine, California, and Department of Mechanical and Aerospace Engineering and Biomedical Engineering Department, Integrated Nanosystems Research Facility (INRF), U.S.A.

INTRODUCTION

Humans are still best at building structures they can grasp with their hands and see with their eyes, and have only relatively recently started to craft objects at the micro- and nanoscale level. The manufacture of miniature devices is still principally based on top-down manufacturing methods; starting from bigger building blocks—say a whole Si wafer—and reducing them into smaller and smaller pieces by cutting, etching, and slicing to make, for example, an integrated circuit chip. Bottom-up manufacturing methods (also nanochemistry) are considered by many as the next natural step in human manufacture in which small particles such as atoms, molecules, and atom clusters are added or removed for the construction of bigger functional structures. Bottom-up methods are nature's way of growing things and in biomimetics one studies how, building bottom-up, that is, atom by atom, nature, through eons of time, developed manufacturing methods, materials, structures, and intelligence.

Given the fundamental differences between top-down and bottom-up manufacturing approaches, we have treated these manufacturing approaches in separate sections. As mentioned earlier, bottom-up manufacturing approaches are often nature-inspired, and we dedicate yet another section to biomimetics. Besides the aforementioned three sections, we start this short review on miniaturization methods with a historical perspective on human manufacturing techniques and we conclude it with our vision on manufacturing approaches for the next 10 to 15 years.

Miniaturization science (MEMS) in the year 2004 is still predominantly practiced by electrical and mechanical engineers, but as applications are becoming more and more biological in nature, biologists, materials scientists, chemists, and physicists must work together with electrical and mechanical engineers to develop new miniaturization solutions including bottom-up methods. We hope that this short article will further stimulate such collaborations and help bridge the gap between dry engineering, in air and with rigid materials, and wet engineering, with flexible materials in watery solutions.

HISTORICAL PERSPECTIVE

Toolmaking for manufacturing an object began perhaps when humans learned to walk erect and had their hands free to grasp objects of wood and stone and to shape them. Although tool-making hominids probably emerged millions of years earlier, we can only trace toolmaking history back two million years to the oldest surviving and recognizable flint hand axes.^[1]

The birth of mechanical, more precise machining dates back to the invention of the lathe, attributed to Anarcharis the Scythian (300 B.C.).^[1] Precision machining with a hand-operated lathe got a boost in the mid-1600s from demands for more precision in building better clocks.^[2] Ruling engines for the control of diffraction gratings have driven state-of-the-art precision machine tool design since the late 1700s. More than a century ago, Rowland at John Hopkins University built ruling engines with part-per-million (ppm) accuracy.^[3,4]

Tools driven by mechanical power brought on the machine age in the late 1700s.^[5] By 1920, the Lord's Prayer had been engraved with a diamond tool into an area measuring 100 μm by 40 μm . Cutting with very small diamond tools (e.g., 22 μm diameter) was developed in Japan in the 1980s (see Fig. 1 for a current example).

Today, the mechanical technology upon which photolithographic alignment systems are based is still grounded in the ruling engine design described earlier. Highly precise instruments such as servomotors, feedback devices, and computers had been implemented in manufacturing tools by 1977 and many types of machine tools are now equipped for computerized numerical control, commonly called CNC.

Processes more sophisticated than mechanical machining can be performed, often more effectively, by using nonmechanical forms of energy, such as chemical, electrochemical, and electrothermal. In electrothermal removing and forming processes, thermal energy, provided by a heat source, melts and/or vaporizes the volume of the material to be removed or deposited. Among electrothermal methods, electro discharge machining, or EDM, using electrical sparks for cutting, is one of the oldest. In the United States during the 1960s, EDM machining was used to create features in the 75- μm size

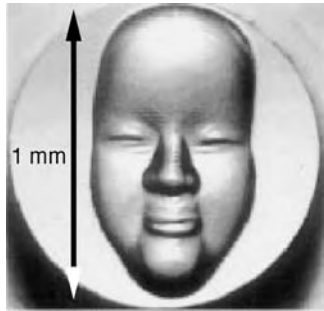


Fig. 1 Noh mask made by mechanical precision machining on Fanuc's ROBOnano UiA. (From Fanuc.)

range, and even more precise EDM, employing rotating spindles, was pioneered in Japan in the late 1980s (5- μm size range). Electron-beam machining (EBM) and laser-beam machining (LBM), two other examples of thermal techniques, were developed in the 1960s and 1980s, respectively. In EBM, high-velocity electrons are used instead of sparks, and in LBM, intense photon pulses are employed.

Chemical and photochemical etching underlies what we know today as integrated circuit (IC) manufacturing and microelectromechanical systems, or MEMS. In 1822, Niépce developed the photoprocess and five years later, Lemaître performed the subsequent chemical milling of a Niépce photoplate—marking the first deliberate use of photolithography in combination with chemical milling.^[6] The IC revolution in the 1950s

employed these same photolithography techniques and, on its coattails, brought along micromachining or MEMS in the late 1960s with the fabrication of the first Si strain gauges. The first Si micromachined devices were isotropically etched. The discovery of anisotropic etching of Si and etch-stop techniques catapulted MEMS firmly onto the commercial scene.

The use of X-ray lithography in combination with electroplating and molding, introduced by Ehrfeld and his colleagues in 1982, demonstrated to the world that lithography may be merged with more traditional manufacturing processes to make master molds of unprecedented aspect ratios and tolerances in order to replicate microstructures in ceramics, plastics, and metals^[7] (see Fig. 2A for an example of a LIGA product, and Fig. 2B for the LIGA process steps).

In soft lithography, a Si master mold is made based on “traditional” lithography and an elastomeric stamp (e.g., polydimethylsiloxane, or PDMS) is molded from this master (Fig. 3A). As illustrated in Fig. 3B, this method enables the manufacture of nonplanar 3D objects. This technique also combines top-down and bottom-up machining: Indeed, the PDMS stamp in Fig. 3A can be inked with a self-assembled monolayer (SAM) to print (stamp) substrates with micro- or nano-sized patterns. This method was introduced by Whitesides et al. in 1997.^[8]

In the coming decades, bottom-up manufacturing methods, in which small particles such as atoms, proteins, and atom clusters are added or removed for the construction of bigger functional structures, will become

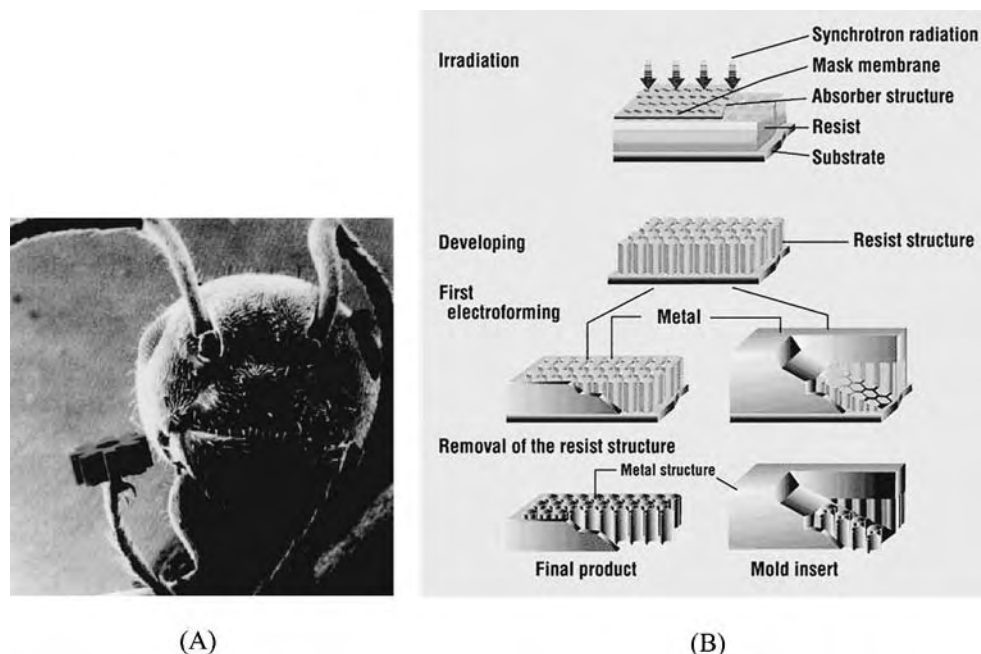


Fig. 2 LIGA (A) Example product: Ant holds a Ni gear produced with the LIGA technique. (B) Basic LIGA process steps. (From Forschungszentrum, Program Microsystem Technologies, with permission.)

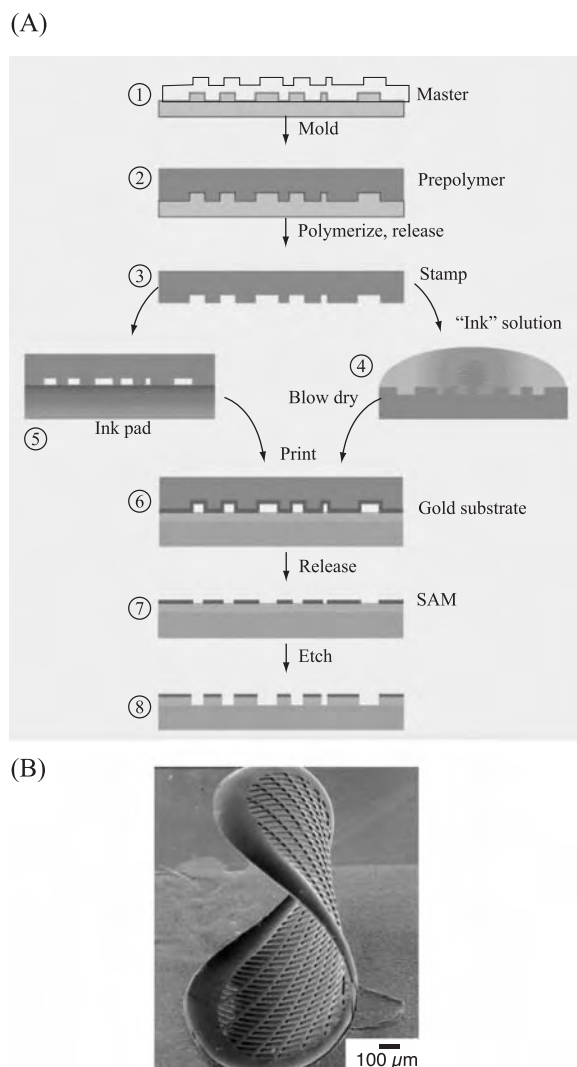


Fig. 3 Soft lithography: (A) Process flow: A prepolymer (2) covering the master (1) is cured by heat or light, and demolded to form an elastomeric stamp (3). The stamp is inked by immersion (4) or contacted with an inkpad (5), and printed onto a substrate (6), forming a self-assembled monolayer (SAM). The ink pattern (7) is then transferred into the substrate by a selective etch (8). (B) Curved glassy carbon structure made from photoresist patterned with an elastomeric stamp. (Courtesy of Dr. G. M. Whitesides, Harvard University.)

the most important research domain for the future of manufacturing. Bottom-up methods draw from work in supramolecular chemistry, protein engineering, molecular self-assembly, proximal probes such as scanning tunneling microscopes (STMs) and atomic force microscopes (AFMs), photon manipulation of individual atoms, carbon nanotubes, etc. An example of bottom-up manufacturing can be found in Eigler et al.'s work at IBM: Using an AFM, this team herds atoms together in any desired patterns (see Fig. 4).



Fig. 4 Kanji character for "atom," by Lutz and Eigler, IBM, Almaden. An example of manipulation/mechanosynthesis of iron atoms on a Cu surface with an STM. (Courtesy of D. Eigler.)

For the next two decades, we expect the size of objects that can be built with top-down and bottom-up approaches to overlap more and more as we learn new ways to combine ever smaller components manufactured through nanofabrication with the products of nature (e.g., rotor proteins). As a result, hybrid top-down and bottom-up constructs can be expected to provide guidance for future manufacturing strategies (see Fig. 5 for a recent example).

TOP-DOWN MANUFACTURING

Introduction

From the myriad of top-down manufacturing techniques (see the preceding discussion), we will only

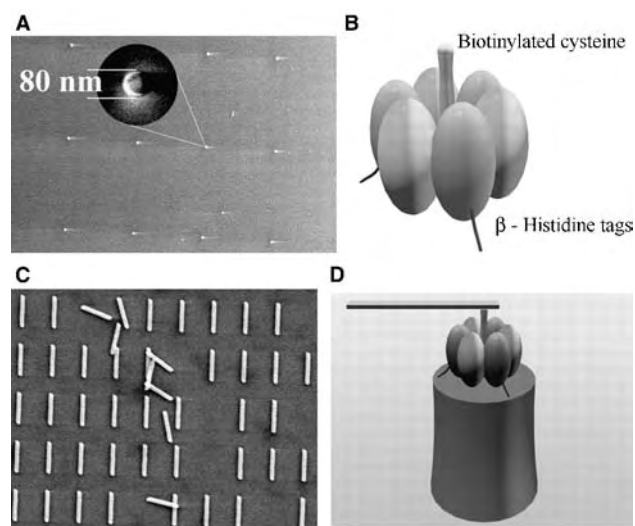


Fig. 5 A hybrid nanodevice with a Ni micromachined propeller attached to a biomolecular motor (a rotor protein). The propeller is powered by adenosine triphosphate (ATP). (Courtesy of Dr. C. Montemagno, UCLA.)

consider those that result in small devices ($<1\text{ cm}^3$) incorporating yet smaller features ($<1\text{ mm}$). More specifically, we will compare precision mechanical machining with Si micromachining or MEMS (micro-electromechanical systems).

Comparing Mechanical Precision Machining with MEMS

3D vs. 2D

Shaping device contours in top-down mechanical manufacturing results in truly 3D shapes (Fig. 6), whereas with MEMS lithography techniques, one can only obtain projected 2D shapes (Fig. 2).

Batch vs. Serial

MEMS and IC fabrication techniques lend themselves to economies of scale (large batches), which are unavailable with mechanical serial machining techniques. Although micromachining is often compared to IC microelectronics, it lacks the high degree of generality characteristic of microelectronics. ICs can be grouped in a limited number of classes, within which design and production follow well-defined and common steps. As a consequence, the price/performance ratio allows industry to make profits on complex ICs, and it makes ICs suited for mass production. MEMS sensors and actuators, on the other hand, are more specific and economy of scale does not apply as commonly as it does for ICs. Compared to mechanically machined devices, MEMS devices still do have a huge economy of scale advantage though.

Relative tolerance

For over 100 years, precise, rather large machines have been designed and built with mechanical removal techniques achieving part-per-million relative tolerances; recently, this type of machining has entered the realm of 0.1-ppm relative tolerances.^[4] In comparing Si

micromachines (MEMS) with larger, mechanically machined structures, Slocum^[4] notes that, whereas MEMS micromachines are impressive for their small size, their relative accuracy is two orders of magnitude worse than is typically achieved in large machines, which, moreover, are much more complex. The surface roughness of micromachines looks high compared to the specular finishes of bearing surfaces. In reality, the absolute roughness is about the same, but in micromachines, surface forces such as friction are relatively more important. Thermal errors, on the other hand, in a micromachine are more relaxed; the small scale and fast thermal equilibrium of the smaller structures is their saving grace. Another key issue Slocum brings up is that of position measurement systems; verification of fabricated geometry and tolerances are much more difficult for micromachines since measurement of a displacement of one part in 10,000 will typically get one down to the nanometer level. It is thus important to realize that whereas MEMS can achieve excellent absolute tolerances, relative tolerances are rather poor compared to those achieved by more traditional techniques. The decrease in manufacturing accuracy with decreasing size is striking; in traditional machining, relative tolerances of 10^{-6} (ppm) are becoming standard, whereas in the integrated circuit (IC) industry, a 10^{-2} relative tolerance is considered good. The definition of precision machining, with relative tolerances of 10^{-4} , actually excludes micromachining! To describe micromachining with the same terminology, it must be expanded to cover all machining methods where the relative accuracies are at least 10^{-4} or where absolute size in one or more dimensions is in the micrometer range. Precision machining and Si micromachining are complementary; both strive to improve absolute and relative tolerances, with Si-based technology better at obtaining smaller features (absolute tolerance, see the following section) and traditional methods better at obtaining tighter relative tolerances.

Absolute tolerances

In mechanical machining, various factors such as deformation of the workpiece and tool, vibration, thermal

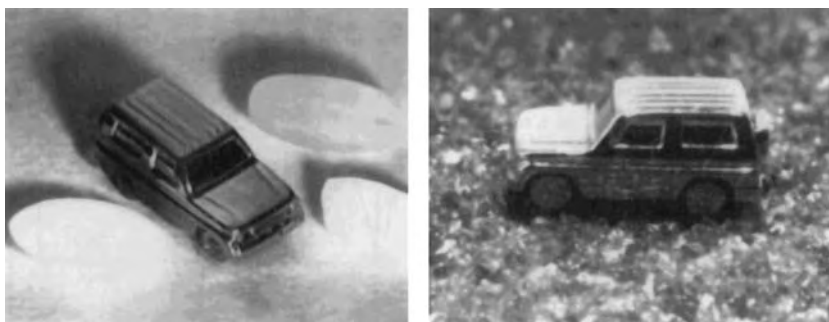


Fig. 6 Microcar and rice grains. The car is 7 mm long, 2.3 mm wide, and 3 mm high and is made with traditional mechanical machining methods. (This figure is from A. Teshigahara, M. Watanabe, N. Kawahara, Y. Ohtsuka, and T. Hattori, *J. Microelectromech. Syst.*, 76–80, 1995.)

deformation, inaccuracies of machine tools, etc., affect the machining accuracy. Ultrahigh precision machines with sharp single-crystal diamond tools have made sub-micrometer precision machining possible (see Fig. 1). The first computer numerical controlled (CNC) machines became available in 1977. The smallest movement the machines could reproducibly make was $0.05\ \mu\text{m}$ by 1993, and today there is equipment available featuring $0.01\ \mu\text{m}$ and even nanometer step resolution.^[9] This progress is sketched in Fig. 7, where the Taniguchi curves show a machining accuracy for ultraprecision machining of $1\ \text{nm}$ for the year 2000. Because of these types of precisions, ultraprecision manufacturing is still the commercially preferred technique for the production of computer hard discs, mirrors for X-ray applications, photocopier drums, heads and camcorder viewfinders, high-definition television (HDTV) projection lenses, and VCR scanning heads, etc.

When looking at the progress in absolute tolerances achieved in ICs and MEMS as presented by Moore's law (Fig. 7), lithography methods only reached critical dimensions (CDs) of $0.18\ \mu\text{m}$ in 2000, as contrasted with the $1\ \text{nm}$ already achieved with mechanical machining at that time. It must be pointed out however, that the accuracy axis for traditional machining

describes accuracy in the height or z direction, whereas for lithography-based methods, it is the accuracy of carving out a feature in the x,y plane. In the x,y plane, mechanical machining is considerably less accurate than lithography-based methods because the writing tool in mechanical machining is crude compared to writing with light or electrons—the smallest diamond styluses are still about $22\ \mu\text{m}$ in diameter! In mechanical machining, for example the polishing of a mirror, the size of the tool does not influence the resulting accuracy in the z -direction, the edge of the tool does.

An intertwined future

Progress in absolute tolerances in top-down mechanical machining and lithography is, in part, based on continued improvement in positioning stages and that progress has slowed down for both machining methods.^[3] In lithography, the reasons for the slowdown go beyond the positioning stage though. The frequent alignment steps in lithography may be checked by laser-interferometer-controlled stages to an absolute positioning accuracy of $<10\ \text{nm}$, thereby allowing the registration of successive layers of a semiconductor device to a similar

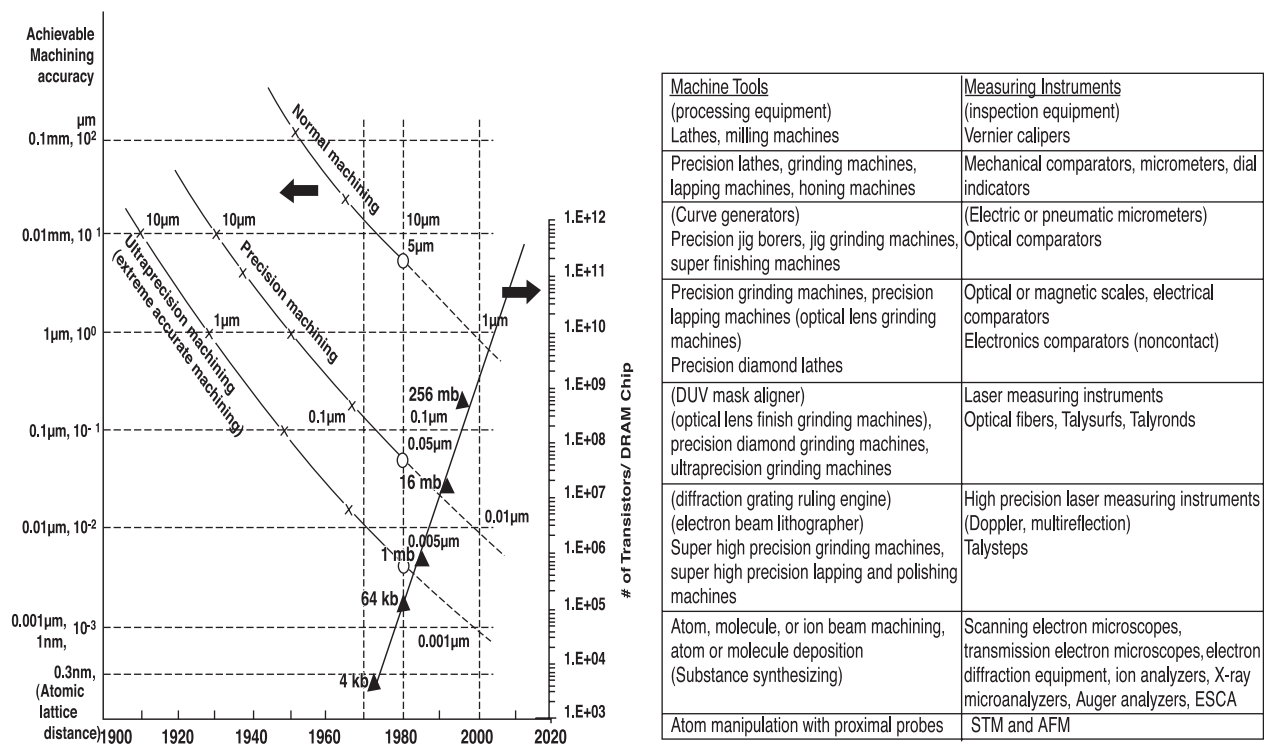


Fig. 7 Progress of accuracy in machining. The development of achievable machining accuracy over the last 70 years is shown in this figure under the generalized classification of normal machining, precision machining, and ultraprecision machining. Left side ordinate: increase of manufacturing over time (based on Taniguchi). Right side ordinate: increase in transistor density over time according to Moore's law. The figure also lists the machines, process equipment, and metrology means by which the indicated resolution can be achieved. By ultraprecision machining, Taniguchi means those processes/machines by which the highest possible dimensional accuracy is or has been archived at a given point in time.

precision. Since interferometric schemes enable alignment to less than 10 nm, it is feasible to develop aligners compatible with line widths of 0.1 μm or even 50 nm.^[10] It is obviously not the metrology equipment that hampers further accuracy improvement in lithography techniques. The reason for the slowdown in electronic device miniaturization hinges indeed on securing more mechanical stability in the positioning stages, but more so on manufacturing cost and the fact that transistors are becoming too leaky below 0.1 μm where quantum effects start playing a role. To further extend manufacturing accuracy in electronics, one has to look beyond current manufacturing methodologies and consider methods such as molecular engineering and novel nanofabrication techniques. Micromachining of special lithography tools will play a pivotal role in those new machining approaches. For example, the scanning tunneling microscope (STM), based on micromachining ultra sharp tips on Si cantilevers, creates machining opportunities with accuracies on the atomic level,^[11] well beyond the predictions of the Taniguchi and Moore curves (see Fig. 4).

For mechanical structures, operation in the micro- and nanodomain imposes barriers other than those described for electronic devices. New design philosophies that are more error-insensitive with simpler mechanisms, smaller numbers of parts, and, wherever possible, flexible members rather than rigid ones must be adopted.

BIOMIMETICS

Introduction

Biomimetics is the study of how nature, building atom by atom, through eons of time, developed manufacturing methods, materials, structures, and intelligence. These studies are inspiring engineering and design of manmade miniature objects. As nanochemistry is currently inspiring a new wave of biomimetic efforts, caution is in order since, as the legend of Icarus so pointedly reminds us, most previous attempts at biomimetics have resulted in failure. In macroengineering, as S. Vogel points out in *Cats' Paws and Catapults* (W. W. Norton, New York, 1998),^[12] human and natural technologies form a separate, well-integrated entity, operating in an internally coherent context. Nature, for example, does not use steel, nor does it favor the production of flat surfaces and sharp corners—all very useful in human manufacturing. Nature builds with proteins and produces mostly curved surfaces and rounded corners, resulting in such masterfully engineered objects as biological cells. In large-scale engineering projects, both natural and human manufacturing approaches have their merits within their own proper frame of reference. It is in the nanoworld that

nature is way ahead of human engineering as it has learned to work with much smaller, more versatile building blocks and master the self-assembly of those building blocks. Because it uses relatively large building blocks, human manufacturing is rapid and the expectation is that nature, because it uses much smaller building blocks, must be very slow. To offset the time it takes to work with small, basic building blocks, nature, in growing an organism, relies on an additive process featuring massive parallelism and self-assembly. We believe that, just as in the macroworld, blindly mimicking nature will be unproductive, but studying nature and taking hints as to how to build nanostructures from the bottom-up (nanochemistry) certainly will be productive.

Hints from Nature

A few of the guiding principles involved in the construction of living organisms are:^[13]

- The use of composites (rather than monolithic materials in human engineering): Nanocomposite materials, for example, are widespread in biological systems as in bone, cell walls, soft body tissue, etc. The nacre (mother-of-pearl) of abalone shell, an oriented coating composed of alternating layers of aragonite (CaCO_3) and a biopolymer, is an often studied example (nacre is twice as hard and 1000 times as tough as its constituent phases).
- Almost all natural manufacturing processes occur in a liquid phase.
- Carbon is the prevalent element in building living things, not silicon.
- Massive parallelism: Multitude of cells all work in parallel.
- Several successive levels of organization (hierarchy): Proteins, bone, and tissue fibers have a hierarchical organization much more complex than materials of the nonliving world (e.g., crystals). This organization bestows many benefits on living organisms such as a high strength-to-density ratio and the ability to serve several functions simultaneously.
- Self-repair. Biological materials might have a narrower operating window than inorganics, but they are constantly repaired when damaged.
- Soft flexible materials: Skin, flesh, muscle, etc.
- Self-assembly and self-replication, all through weak bonds: Self-assembly is the lock- and-key assembly of complex molecules such as in the case of neurotransmitters assembling with molecular recognition sites on proteins that elicit chemical responses, for example, acetylcholine recognition at nerve-muscle junctions resulting in muscle contractions.^[14] In self-replication, molecules such as DNA encode the genetic information for cells and the resulting



organisms, thus allowing the manufacturing information to be passed on from cell to cell and from generation to generation.^[15]

- The use of templates: For example, the reading of genes by messenger RNA (m-RNA) to build proteins in the cells' ribosomes.
- Compartmentalization: Tissue is compartmentalized into cells and within cells there are many organelles.

These principles provide inspiration for the implementation of human nanochemistry. Nanochemistry holds the promise of the versatility of design offered by nature itself, and molecular self-assembly and replication add to the tremendous appeal of this type of nanotechnology.

BOTTOM-UP MANUFACTURING APPROACHES

Introduction

At the beginning of the twenty-first century, the performance criteria for top-down manufacturing techniques, that is, Moore's law, which projects the progress in achievable transistor density on a chip, and Taniguchi's curves, which predict improvements in accuracy for mechanical machining, have started to exhibit signs of a slowdown (see Fig. 7). Further miniaturization progress will most likely be based on methods originating in

nanochemistry, that is, bottom-up methodologies as well as combinations of top-down nanofabrication using traditional IC methods and nanochemistry. Over the next 10 years, the size overlap of objects fabricated with either approach will become more significant. Consequently, molecular engineers and supramolecular chemists who are manipulating, with rapidly improving dexterity, nature's building blocks such as atoms, molecules, proteins, and DNA, will, jointly with MEMS specialists who are now reaching the 0.1- μm lithography limits, find new ways to combine top-down nanofabricated components with bottom-up natural products (see Fig. 5).

Multiparallelism Micro- and Nanoassembly

Today, IC and MEMS hybrid assembly is performed mostly by humans with tweezers and microscopes or with pick-and-place robots.^[16] As long as we cannot master self-assembly, the challenge to assemble smaller and smaller components from different technologies is staggering. The existing microassembly approaches, divided into serial and parallel processes, are summarized in Table 1. In serial microassembly, parts are picked up and placed together on a substrate or workpiece one by one. Table 1 covers many methods enabling serial manipulation of very small particles, including individual atoms for nanoassembly, but unless one can

Table 1 Microassembly approaches

Serial microassembly		Parallel microassembly	
Manual assembly with tweezers and microscopes		<i>Deterministic</i>	<i>Stochastic driving forces</i>
Visually-based and teleoperated microassembly	[17,18]	Flip-chip wafer-to-wafer	[19,20] Fluidic agitation and mating part shapes [21–23]
High-precision macroscopic robots with submicron resolution	[24,25]	Microgripper arrays for massively parallel pick-and-place	[26] Vibratory agitation and electrostatic force fields [27–31]
Microgrippers (with gripper sizes 100 μm or less)	[26,32–34]		Vibratory agitation and mating part shapes [35]
Optical trapping	[36]		Mating patterns of self-assembling monolayers [37]
Microcapillary/pipette-based positioning	[37]		Magnetic field-mediated separation [38]
AFM/STM atom manipulation	[11]		Dielectrophoretic trapping [39]
			Ultrasonic levitation and transportation in standing acoustic waves [40]
			Hydrodynamic/electro-osmotic object manipulation in microchannels [41,42]

(From Ref. [16].)

manipulate many, many particles all in parallel, most of these techniques will remain research tools rather than manufacturing options. In parallel microassembly, where parts are assembled simultaneously, we distinguish between deterministic and stochastic microassembly.^[16] In the former, the relationship between the part and its destination on a substrate is known in advance. In the latter, this relationship is unknown or random; parts self-assemble during a stochastic process driven by some motive force.

Stochastic self-assembly is an attractive option to integrate dissimilar process technologies. To achieve self-assembly, there must be bonding forces present, the bonding must occur selectively, and the assembling parts must be able to move so that they can come together. Stochastic microassembly often encompasses vibration of the parts in combination with electrostatic, fluidic, and other forces that operate on singulated parts in various media (fluids, air, or vacuum). After reviewing scaling issues of the assembly process, examples of micro- and nanoassembly will be presented. The smaller the building blocks, the higher the need for stochastic self-assembly as handling the individual components becomes more and more of a mechanical challenge and the process would be too slow. Nature, building on the nanoscale, does rely heavily on self-assembly, and DNA-assisted self-assembly (see subsequent discussion) is an early attempt at coopting nature for human manufacturing endeavors.

Scaling of the Assembly Process

The newest robots with submicrometer resolution must be used to assemble nanoparticles as conventional robots that pick, transport, and place small parts having a control error of 100 μm at best, which would translate in relative errors of over 100%. Moreover, small parts (<1 mm and masses less than 10^{-6}g) are difficult to release as adhesion forces between gripper and object are significant compared to gravitational forces at that scale.^[16] In the microdomain, surface forces such as adhesion dominate over volume forces such as gravity or inertia. Adhesive forces involved in pick and place include surface tension due to moisture on the part and gripper, van der Waals forces (significant for gaps below 100 nm), and electrostatic attractions (e.g., induced by triboelectrification). To avoid some of these problems, the smallest components are often manipulated in a liquid medium, for example, with laser trapping or dielectrophoresis; the fluid eliminates electrostatic and surface tension effects. On a larger scale, vibration is used extensively in industrial part feeders which singulate, position, and orient parts before they are transferred to an assembly station.

Microassembly Examples

Serial microassembly

Many of the serial manipulation methods listed in Table 1 are used for manipulation of components such as biological cells, subcellular organelles, and even atoms.^[43] So the same tools may be used for micro- and bottom-up nanoassembly.

A small enough object that is relatively transparent to laser light of a particular frequency refracts an incident laser beam, bending the light. As a result of this refraction, momentum is transferred from the light to the particle. When the geometry of light beams and particle is correct, the momentum imparted pulls the particle in the direction of the incident beam. The beam (or multiple beams) can then hold the particle in place, or by moving the beam(s), the laser operator can move the particle from place to place. In other words, light has momentum and can be used to catch and manipulate objects in a size range from nm to μm in so-called optical tweezers. The laser wavelengths in optical tweezers are typically between 0.7 and 1.06 μm , and 25 to 500 milliwatts of power is generated in a focal spot between 0.5 and 1.0 mm in diameter. The first work with optical tweezers was done by Ashkin et al. at AT&T Bell Laboratories in pioneering works in 1985 and 1986, successfully demonstrating that small particles could be trapped by focused low-power laser light (under one watt).^[44,45] The first work on trapping of living biological objects was also done by Ashkin et al. in 1987.^[36] The authors not only trapped and manipulated single cells, they demonstrated that the trap was harmless to living biological species. Steven Chu, a 1997 Nobel laureate in physics, at Stanford University showed that laser tweezers can be used to stretch molecules such as DNA: with colloidal spheres attached to the DNA strands, the spheres were trapped and the DNA was stretched by moving in fluid flows.^[46] The most common form of optical tweezers in biology is the single-beam gradient trap, immobilizing transparent dielectric particles near the focal point of a tightly focused continuous-wave laser with a force proportional to the electric field gradient.^[43,47,48] Such a beam generates forces in the piconewton range, more than sufficient for trapping cells and moving organelles inside and outside the cells.

Two different serial manipulation approaches can be used together; for example, cages for biological cells have been made by combining optical trapping and dielectrophoresis.^[49] Such a combination enables manipulation of two bioparticles in close proximity, which would be impossible using two laser beams or two electrode fields, as those fields interfere. Dielectrophoresis is an AC electrokinetic force and it is a powerful tool for separating particles of the size of biological cells, but is less applicable for particles below 1 μm .



In mechanosynthesis, an STM is used to manipulate individual atoms, as demonstrated in Fig. 4 with the Kanji character for the word atom. Such serial manufacturing processes that use too small building blocks are not viable manufacturing candidates unless an army of them can work together on the same construction project.

Parallel microassembly

Many stochastic approaches for microassembly are under investigation. Yeh et al.^[22,23] trapped semiconductor LEDs, originally suspended in a liquid, in micro-machined wells on a wafer by solvent-surface forces. Carrier fluid containing the GaAs parts is dispensed over the host Si wafer with etched holes to assemble the GaAs blocks. Because of the trapezoidal design, the blocks fit preferentially into the holes in the design orientation. Random mechanical vibration of the microstructures enables large numbers of microstructures to be positioned into precise registration with the sites.

Cohn et al.^[28] demonstrated an interesting variation on the work of Yeh et al. by using alignment capabilities over electrostatic traps. A critical problem in attempting to vibrate microstructures into specific sites from a liquid is that when structures stick to the substrate or to each other, their progress toward optimal, precise registration is halted. Cohn et al. addressed this problem by levitating the microstructures a short distance above the target electrostatic trap site before letting them settle. As a result, random sticking is prevented and critical positioning and orientation of the microstructure with respect to the trap become possible with a relatively low applied field ($10\text{ V}/\mu\text{m}$). The levitation over the trap site is accomplished by creating a short-range repulsive force between the microstructure and the target site. This prevents contact between the site and the suspended microstructure. Critical positioning and orientation become possible, in the manner of a compass needle resting on a low-friction bearing.^[28] Böhringer et al. have subsequently also demonstrated the ability to break surface stiction forces using ultra low-amplitude vibration in vacuum, which holds promise as an extremely sensitive technique for positioning parts, as well as discriminating part orientation, shape, and other physical properties.^[29]

Several MEMS research groups have built actuator arrays for micromanipulation; these arrays typically consist of a regular grid of motion pixels (mechanical members such as cilia).^[50–57]

At Nanogen, electric fields are used to transport and control the placement of proteins, RNA, and DNA as well as semiconductor components on a substrate equipped with an array of electrodes in a low-conductivity

solution.^[31,58] In Fig. 8, we illustrate the use of electric fields for precision placement of microLEDs ($20\text{-}\mu\text{m}$ diameter) onto a Si substrate using electroosmosis. The InGaAs microLEDs used all have contacts on one side of the device, as shown in the cross-section schematic in Fig. 8A. The corresponding substrate has two sets of electrodes designed for the mounting process. As seen in Fig. 8B, a first pair of larger horseshoe electrodes brackets an inner set of electrodes. The inner set of

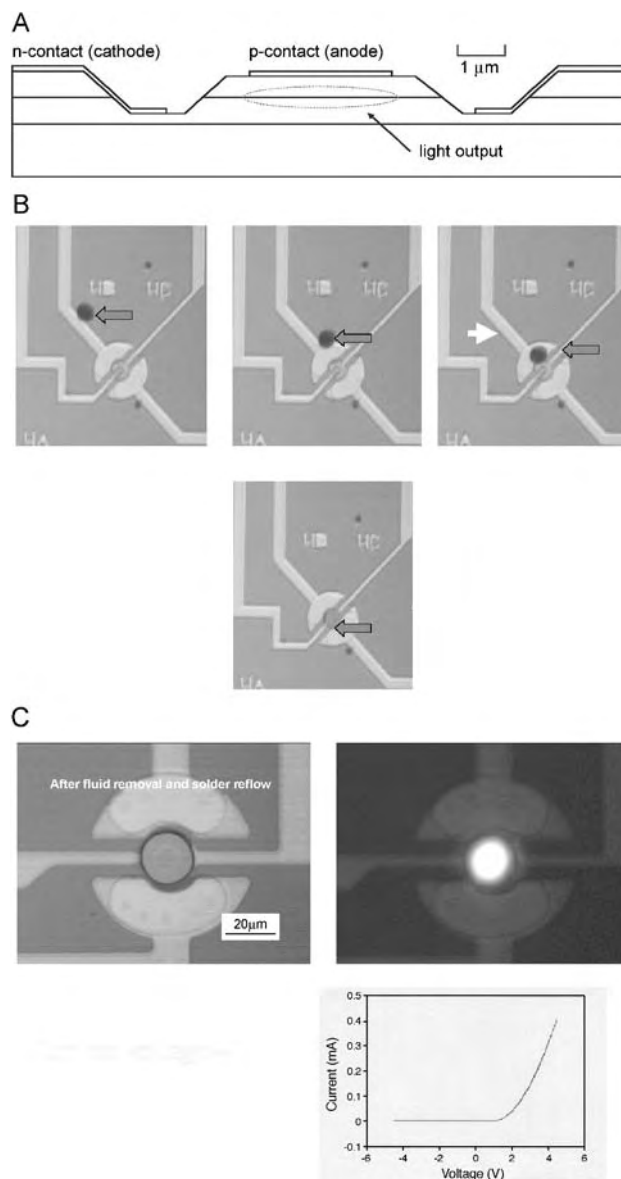


Fig. 8 (A) Schematic InGaAs microLEDs have contacts on one side of the device, as shown in the cross section. (Courtesy of Dr. M. Heller, Nanogen.) (B) Electric field directed placement of LED. (Courtesy of Dr. M. Heller, Nanogen.) (C) Circuit integrity and LED activation demonstrated by forward biasing the LED. (Courtesy of Dr. M. Heller, Nanogen.)

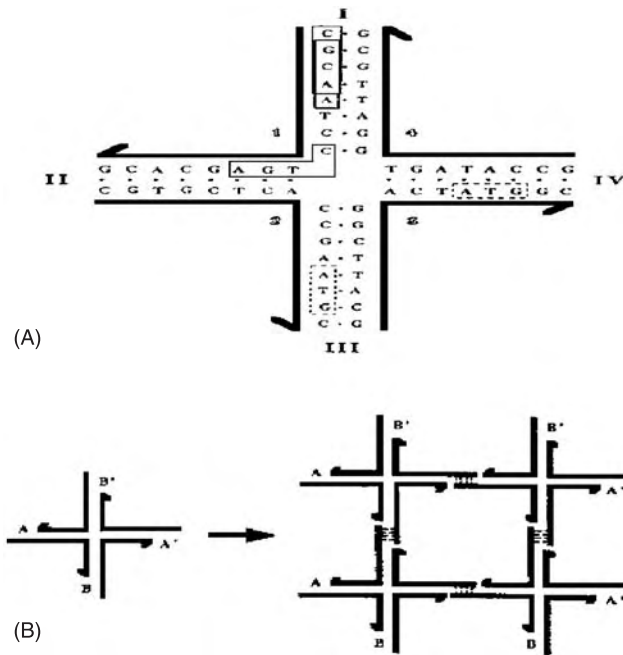


Fig. 9 (A) Four-armed stable branched junction made from DNA molecules; (B) use of branched junctions to form periodic crystals. (From Ref. [60].)

electrodes is designed to mate and bond to the LED and is standard tin/lead electroplated to a height 50 nm above the surrounding surface. The linear structures connect traces, which are insulated to avoid contact with the overlaying aqueous solution. Activation of the outer electrodes with respect to a counterelectrode (not shown in the figure) causes an electroosmotic flow along the surface of the chip, and the LED placed in the solution moves with the bulk fluid flow. Once the LED is on the electrode, the electroosmotic flow is briefly reversed and the reversal of the fluid flow drives the LED onto the awaiting bonding points with microscale precision. After the fluid is removed, the solder is reflowed under forming gas at 250°C. Circuit integrity and LED activation is demonstrated by forward biasing the LED using the silicon circuit's contacts, as demonstrated in Fig. 8C.

The authors point out that the same process could be used to assemble a large variety of components simultaneously and selectively on a circuit by the parallel activation of sets of electrodes while maintaining a number of components present in the low-conductivity solution. They speculate that the process might be aided by the use of a self-sorting mechanism by attaching short sections of complementary DNA on components and target sites on the circuit (see next section).^[58]

DNA-mediated assembly

Many researchers have recognized that the exquisite molecular recognition of various complementary

biological molecules may be used to assist in the assembly of micro- or nanocomponents into complex networks. In this approach, DNA nanostructures and DNA-mediated assembly of nano- and microstructures have received the most attention in the literature. In the DNA nanostructures approach, pioneered by N. C. Seeman, the 2-nm-wide DNA molecules themselves are used to form 2D and 3D structures.^[59–63] As shown in Fig. 9A, DNA four-armed branched junctions are made by properly choosing the sequence of complementary strands. These DNA scaffolds can be linked together to make periodic crystals, as shown in Fig. 9B.

To use DNA for assembling nano- or microparts, one first needs to attach DNA molecules to these parts. The most widely used method is the covalent bond between sulfur and gold, that is, a metal thiolate (~44 kcal/mol). DNA molecules can be functionalized with a thiol (S–H) or disulfide (S–S) group and, upon immersion of microparts with clean gold surfaces in solutions of thiol-derivatized oligonucleotides, the sulfur adsorbs on the gold surfaces, forming a single layer of molecules. Using this methodology, Mirkin et al. described a method of assembling colloidal gold nanoparticles into macroscopic aggregates using DNA as the linking element.^[64] Noncomplementary DNA oligonucleotides were attached to the surfaces of two batches of 13-nm gold particles. An oligonucleotide duplex with ends complementary to the grafted sequences was then introduced into the particle mixture and the nanoparticles self-assembled into aggregates. The process is illustrated in Fig. 10. It can be reversed by increasing the temperature, thereby melting the double-stranded DNA linkers. Closely-packed assemblies of aggregates with uniform particle spacings of about 60 Å were demonstrated this way.

The concept of DNA-mediated self-assembly has been extended to metallic nanowires/rods.^[65] The idea behind this is to fabricate gold and/or platinum wires, functionalize them with single stranded DNA (ssDNA), and assemble a wiring network on a substrate, which has the complementary ssDNA molecules attached at specific sites. Thus, self-assembly of interconnects and wires is rendered possible. DNA has also been used as a template for the fabrication of nanowires. Braun et al. formed a DNA bridge between two gold electrodes, again using thiol attachment. Once a DNA bridge is formed between the 12–16 μm spacing of the electrodes, a chemical deposition process is used to deposit silver ions along the DNA through Ag⁺/Na⁺ ion exchange and reduction of the Ag complexes with hydroquinone. The result is a silver nanowire that is formed using DNA as a template or skeleton. The most ambitious goal is the DNA-assisted assembly (molecular velcro) of semiconductor active devices such as transistors, laser diodes, etc., proposed by Edman et al. at Nanogen.^[58]

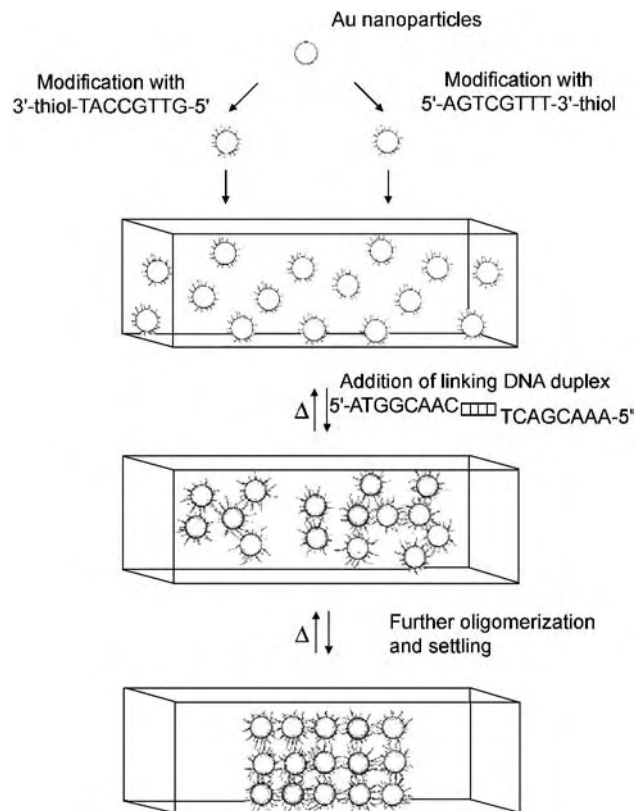


Fig. 10 Fabrication process for the aggregated assembly of DNA-conjugated gold nanoparticles. (From Ref. [64].)

A vision for the future of manufacturing

IC techniques for batch manufacturing are applied today to a wide variety of non-IC products and are also combined with traditional precision manufacturing methods. Undoubtedly, the merging of traditional precision engineering and IC manufacturing methodology will lead to many more interesting new manufacturing processes. In the next 15 years, Moore's law will continue to run its course, but in parallel, we will start adapting nanochemistry principles and combine them with more traditional IC fabrication techniques.

There is little doubt that biotechnology analysis tools will keep on improving at an increasingly faster rate. Desktop DNA sequencers and three-dimensional protein readers will be a reality in the not-so-distant future. Genetic engineering will have the most profound impact though on how mankind looks at manufacturing. Since it is already possible to synthesize a virus bottom-up, given the sequence of the bases in its genes, it seems quite likely that we will be able to manufacture synthetic viruses designed to enter a cell and carry out diagnostic and therapeutic tasks. Along this line, Venter has proposed building bottom-up a minimal gene setup required to sustain life in a test tube. Many other applications will become evident; photovoltaic solar cells today have a high conversion efficiency of about 10 to

15%, but are expensive to deploy and maintain. Crops grown for energy are also expensive, involve harvesting, and are only 1% efficient. It is conceivable that genetic engineering will enable the production of energy crops that convert sunlight into fuel at a 10% efficiency. Through nanochemistry, the current digital information technology (IT) revolution might well be followed by a new analogue manufacturing revolution. Today, computers let us shape our digital environment, but by giving computers the means to manipulate the analogue world of atoms as easily as they manipulate bits, the same kind of personalization may be brought to our physical three-dimensional environment. In this context, Gershenfeld from the MIT Media Laboratory envisions a personal fabricator (PF) akin to the PC.

A human society based on nanomachining will be a much more balanced one with a manufacturing approach based on how the species itself is made. Products will be based on a fundamental understanding of the assembly of their ultimate components, atoms, molecules, and proteins, and on how to induce self-assembly into useful objects. Materials will be degradable, flexible, and fully reusable. The smaller building blocks used in manufacture will enable products of more variety and intelligence. There will be less emphasis on the traditional engineering materials such as steel, wood, stone, composites, and carbon, and proteins will become much more important.

The transition toward a nanosociety will require a major shift in workforce skill level as manipulation of data and applying knowledge of bioengineering will be part of a manufacturing worker's daily duties. In academia, less hyperspecialization and better grounding in all the sciences and engineering will be a must as the traditional dry engineering and sciences, such as electrical engineering, and mechanical engineering, will merge with the wet sciences and engineering, such as biology and bioengineering.

ARTICLES OF FURTHER INTEREST

Biosensors; Liver, Bio-Artificial; Medical Robotics; Neuroprostheses; Tissue Engineering of Liver; Tissue Engineering, Microscale; Tissue Engineering of Microvascular Networks; Photopolymerization

REFERENCES

1. Hodges, H. *Technology in the Ancient World*; Barnes & Noble: New York, 1970.
2. Zuurveen, F. Precisie-Technologie: Kennis en Kunde, Inzicht en Uitzicht. In *Precisie-Technologie-Jaarboek 1994*; NVFT: Eindhoven, 1994; 99-118.
3. Slocum, A.H. *Precision Machine Design*; Prentice Hall: Englewood Cliffs, NJ, 1992.

4. Slocum, A.H. Precision Machine Design: Macro-machine Design Philosophy and Its Applicability to the Design of Micromachines. In *Proceedings: IEEE Micro Electro Mechanical Systems (MEMS '92)*; Travemunde, Germany, 1992; 37–42.
5. Compton *Compton's Interactive Encyclopedia (Interactive Multimedia): Computer Data and Program*; Compton's New Media: Carlsbad, 1994.
6. Harris, T.W. *Chemical Milling*; Clarendon Press: Oxford, 1976.
7. Becker, E.W.; Ehrfeld, W.; Munchmeyer, D.; Betz, H.; Heuberger, A.; Pongratz, S.; Glashauser, W.; Michel, H.J.; Siemens, V.R. Production of separation nozzle systems for uranium enrichment by a combination of x-ray lithography and galvanoplastics. *Naturwissenschaften* **1982**, *69*, 520–523.
8. Xia, Y.; Whitesides, G.M. Soft lithography. *Angew. Chem., Int. Ed. Engl.* **1998**, *37*, 550–575.
9. Szepesi, D. Sensoren en Actuatoren in Ultraprecisie Draaibanken. In *Sensoren en Actuatoren in de Werktuigbouw/Machinebouw, Centrum voor Micro-Electronica*; The Hague, Netherlands, 1993; 99–107.
10. Smith, H.I.; Schattenburg, M.L. Why bother with x-ray lithography? *SPIE* **1992**, *1671*, 282–298.
11. Stroschio, J.A.; Eigler, D.M. Atomic and molecular manipulation with the scanning tunneling microscope. *Science* **1991**, *254*, 1319–1326.
12. Vogel, S. *Cats' Paws and Catapults*; W.W. Norton: New York, 1998.
13. Ball, P. *Made to Measure*; Princeton University Press: Princeton, NJ, 1997.
14. Whitesides, G.M.; Mathias, J.P.; Seto, C.T. Molecular self-assembly and nanochemistry: A chemical strategy for the synthesis of nanostructures. *Science* **1991**, *254*, 1312–1318.
15. Achilles, T.; Von Kiedrowski, G. A self-replicating system from three starting materials. *Angew. Chem., Int. Ed. Engl.* **1993**, *32*, 1198–1201.
16. Böhringer, K.F.; Fearing, R.S.; Goldberg, K.Y. Microassembly. In *Handbook of Industrial Robotics*; Nof, S., Ed.; Wiley: New York, 1998.
17. Nelson, B.; Vikramaditya, B. Visually Guided Micro-assembly Using Optical Microscopes and Active Vision Techniques. In *Proceedings: IEEE Int. Conf. on Robotics and Automation (ICRA)*; 1997. Albuquerque, N.M.
18. Feddema, J.T.; Simon, R.W. CAD-Driven Micro-assembly and Visual Servoing. In *Proceedings: IEEE Int. Conf. on Robotics and Automation (ICRA)*; 1998. Leuven, Belgium.
19. Cohn, M.B.; Howe, R.T. Wafer-to-Wafer Transfer of Microstructures Using Breakaway Tethers. In *U.S. Patent Application*; 1997.
20. Singh, A.; Horsley, D.A.; Cohn, M.B.; Pisano, A.P.; Howe, R.T. Batch Transfer of Microstructures Using Flip-Chip Bump Bonding. In *Digest: Int. Conf. on Solid State Sensors and Actuators*; Transducers Research Foundation: Chicago, 1997.
21. Yeh, H.-J.; Smith, J.S. Fluidic self-assembly for the integration of GaAs light-emitting diodes on Si substrates. *IEEE Photonics Technol. Lett.* **1994**, *6*, 706–708.
22. Yeh, H.-J.; Smith, J.S. Fluidic Self-Assembly of Microstructures and Its Application to the Integration of GaAs on Si. In *Proc. IEEE International Workshop on Micro Electro Mechanical Systems (MEMS '94)*; 1994; 279–284. Oiso, Japan.
23. Yeh, H.-J.; Smith, J.S. Integration of GaAs vertical-cavity surface-emitting laser on Si by substrate removal. *Appl. Phys. Lett.* **1994**, *64*, 1466–1468.
24. Quaid, A.E.; Hollis, R.L. Cooperative 2-Dof Robots for Precision Assembly. In *Proceedings: IEEE Int. Conf. on Robotics and Automation (ICRA)*; 1996. Minneapolis, Minn.
25. Zesch, W. Multi-Degree-of-Freedom Micropositioning Using Stepping Principles. In *Ph.D. Thesis*; Swiss Federal Institute of Technology: Zurich, Switzerland, 1997.
26. Keller, C.G.; Howe, R.T. Hexsil Tweezers for Teleoperated Micro-Assembly. In *Proceedings: IEEE Workshop on Micro Electro Mechanical Systems (MEMS)*; 1997; Nagoya, Japan.
27. Cohn, M.B. Self-Assembly of Microfabricated Devices. U.S. Patent 5,355,577, 1992.
28. Cohn, M.B.; Howe, R.T.; Pisano, A.P. Self-Assembly of Microsystems Using Non-Contact Electrostatic Traps. In *Proceedings of the ASME International Congress and Exposition, Symposium on Micromechanical Systems (IC '95)*; 1995; 893–900. San Francisco.
29. Böhringer, K.-F.; Cohn, M.B.; Goldberg, K.; Howe, R.; Pisano, A. Electrostatic Self-Assembly Aided by Ultrasonic Vibration. AVS 44th National Symposium; 1997. San Jose, Calif.
30. Böhringer, K.-F.; Goldberg, K.; Cohn, M.B.; Howe, R.; Pisano, A. Parallel Microassembly with Electrostatic Force Fields. In *Proceedings: IEEE Int. Conf. on Robotics and Automation (ICRA)*; 1998; Leuven, Belgium.
31. Edman, C.F.; Gurtner, C.; Formosa, R.E.; Coleman, J.J.; Heller, M.J. Electric-Field-Directed Pick-and-Place Assembly. In *HDI*; October, 2000; 30–35.
32. Kim, C.-J.; Pisano, A.P.; Muller, R.S. Silicon-Processed Overhanging Microgripper. *J. Microelectromech. Syst.* **1992**, *1*, 31–36.
33. Pister, K.S.J.; Judy, M.W.; Burgett, S.R.; Fearing, R.S. Microfabricated hinges. *Sensors Actuators, A* **1992**, *33*, 249–256.
34. Keller, C.; Howe, R.T. Nickel-Filled Hexsil Thermally Actuated Tweezers. In *Digest: Int. Conf. on Solid-State Sensors and Actuators*; Transducers Research Foundation: Stockholm, Sweden, 1995.
35. Hosokawa, K.; Shimoyama, I.; Miura, H. Dynamics of self-assembling systems: Analogy with chemical kinetics. *Artif. Life* **1995**, *1*.
36. Ashkin, A.; Dziedzic, J.M.; Yamane, T. Optical trapping and manipulation of single cells using infrared laser beams. *Nature* **1987**, *330*, 769–771.
37. Brown, K.T.; Flaming, D.G. Advanced Micropipette Techniques for Cell Physiology. In *IBRO Handbook: Methods in Neurosciences*; Chichester, Ed.; Wiley: New York, 1992.
38. Miltenyi, S.; Mueller, W.; Weichel, W.; Radbruch, A. Cytometry **1990**, *11*, 231–238.



39. Schnelle, T.; Hagedorn, R.; Fuhr, G.; Fiedler, S.; Müller, T. . *Biochim. Biophys. Acta* **1993**, *1157*, 127–140.
40. Coackley, W.T. . *TIBTECH* **1997**, *5*, 506–511.
41. Giddings, J.C. . *Science* **1993**, *260*, 1456–1464.
42. Fu, A.; Spence, C.; Scherer, A.; Arnold, F.H.; Quake, S.R. . *Nat. Biotechnol.* **1999**, *17*, 1109–1111.
43. Wheeler, A.R.; Morishima, K.; Arnold, D.W.; Rossi, A.B.; Zare, R.N. Single Organelle Analysis with Integrated Chip Electrophoresis and Optical Tweezers. In *Micro Total Analysis Systems 2000*; van den Berg, A., Olthuis, W., Bergveld, P., Eds.; Kluwer Academic Publishers: Enschede, The Netherlands, 2000; 25–28.
44. Ashkin, A.; Dziedzic, J.M. Observation of radiation-pressure trapping of particles by alternating light beams. *Phys. Rev. Lett.* **1985**, *54*, 1245–1248.
45. Ashkin, A.; Dziedzic, J.M.; Bjorkholm, J.E.; Chu, S. Observation of a single-beam gradient force optical trap for dielectrical particles. *Opt. Lett.* **1986**, *11*, 228–290.
46. Perkins, T.T.; Smith, D.E.; Larson, R.G.; Chu, S. Stretching of a single tethered polymer in a uniform flow. *Science* **1995**, *268*, 83–87.
47. Ashkin, A.; Dziedzic, J.M. Optical trapping and manipulation of viruses and bacteria. *Science* **1987**, *235*, 1517–1520.
48. Svoboda, K.; Block, S.M. Biological applications of optical forces. *Annu. Rev. Biophys. Biomol. Struct.* **1994**, *23*, 247–285.
49. Fuhr, G.R.; Reichle, C. Living Cells in Opto-Electrical Cages: Characterisation, Manipulation and Force Measurements. In *Micro Total Analysis System*; van den Berg, A., Olthuis, W., Bergveld, P., Eds.; Kluwer Academic Publishers: Enschede, The Netherlands, 2000; 261–264.
50. Darling, R.B.; Suh, J.W.; Kovacs, G.T. Ciliary micro-actuator array for scanning electron microscope positioning stage. *J. Vac. Sci. Technol., A* **1998**, *16*, 1998–2002.
51. Pister, K.S.J.; Fearing, R.; Howe, R. A Planar Air Levitated Electrostatic Actuator System. In *Proceedings: IEEE Workshop on Micro Electro Mechanical Systems (MEMS)*; 1990; 67–71. Napa Valley, Calif.
52. Fujita, H. Group Work of Microactuators. In *International Advanced Robot Program Workshop on Micromachine Technologies and Systems*; 1993; 24–31. Tokyo, Japan.
53. Böhringer, K.-F.; Donald, B.R.; Mihailovich, R.; MacDonald, N.C. A Theory of Manipulation and Control for Microfabricated Actuator Arrays. In *Proceedings: IEEE Workshop on Micro Electro Mechanical Systems (MEMS)*; 1994. Oiso, Japan.
54. Storment, C.W.; Borkholder, D.A.; Westerlind, V.; Suh, J.W.; Maluf, N.I.; Kovacs, G.T.A. Flexible, dry-released process for aluminum electrostatic actuators. *J. Microelectromech. Syst.* **1994**, *3*, 90–96.
55. Suh, J.W.; Glander, S.F.; Darling, R.B.; Storment, C.W.; Kovacs, G.T.A. Combined Organic Thermal and Electrostatic Omnidirectional Ciliary Microactuator Array for Object Positioning and Inspection. In *Proceedings: Solid-State Sensor and Actuator Workshop*; 1996. Hilton Head, S.C.
56. Liu, C.; Tsao, T.; Will, P.; Tai, Y.; Liu, W. A Micro-Machined Magnetic Actuator Array for Micro-Robotics Assembly Systems. In *Digest: Int. Conf. on Solid-State Sensors and Actuators*; Transducers Research Foundation: Stockholm, Sweden, 1995.
57. Liu, W.; Will, P. Parts Manipulation on an Intelligent Motion Surface. In *IEEE/RSJ Int. Workshop on Intelligent Robots & Systems (IROS)*; 1995; Pittsburgh, Penn.
58. Edman, C.F.; Swint, R.B.; Gurtner, C.; Formosa, R.E.; Roh, S.D.; Lee, K.E.; Swanson, P.D.; Ackley, D.E.; Coleman, J.J.; Heller, M.J. Electric field directed assembly of an ingaas led onto silicon circuitry. *IEEE Photonics Technol. Lett.* **2000**, *12*, 1198–1200.
59. Seeman, N.C. *J. Theor. Biol.* **1982**, *99*, 237.
60. Seeman, N.C. *Nanotechnology* **1991**, 149.
61. Seeman, N.C.; Zhang, Y.; Chen, J. *J. Vac. Sci. Technol.* **1993**, *A12*, 1895.
62. Seeman, N.C. *Annu. Rev. Biophys. Biomol. Struct.* **1998**, *27*, 225.
63. Winfree, E.; Liu, F.; Wenzler, L.; Seeman, N.C. *Nature* **1998**, *394*, 539.
64. Mirkin, C.A.; Letsinger, R.L.; Mucic, R.C.; Storhoff, J.J. *Nature* **1996**, *382*, 607.
65. Martin, B.R.; Dermody, D.J.; Reiss, B.D.; Fang, M.; Lyon, L.A.; Natan, M.J.; Mallouk, T.E. *Adv. Mater.* **1999**, *11*, 1021.

Micro-Particle Image Velocimetry in Biomedical Applications

Jinhua Cao

Steve T. Wereley

Department of Mechanical Engineering, Purdue University, West Lafayette, Indiana, U.S.A.

INTRODUCTION

In recent years, the biological and biomedical applications of micro- and nanotechnology [commonly referred to as biomedical or biological micro-electro-mechanical system (BioMEMS)] have become increasingly prevalent and have found widespread use in a variety of applications such as diagnostics, therapeutics, and tissue engineering.

Fluid flow at the microscale exhibits unique non-intuitive behavior that can be leveraged to fabricate devices and components capable of performing functions useful for biological studies. Microfluidic systems for drug discovery in the pharmaceutical industry,^[1] molecular recognition in clinical diagnostics,^[2] and cell culture and manipulation for cellular and tissue engineering^[3] are changing the way in which biological information is processed. BioMEMS are devices that, much like microfluidic MEMS, are designed to interact specifically with biological entities. Some benefits of microscale fluidic technology for medical applications include reduced fluid consumption, reduced waste products, decreased reaction times, and improved parallel chemical processing capability.

The i-STAT device (i-STAT, Inc.) is one of the first microfabricated fluid devices that have seen routine use in the medical community for blood analysis. Other examples of microfluidic devices for biomedical research include microscale flow cytometers for cancer cell detection,^[4] micromachined electrophoretic channels for DNA fractionation, and polymerase chain reaction (PCR) chambers for DNA amplification.^[5]

BACKGROUND

As the field of BioMEMS and the development of microfluidic devices continue to expand, so too does the physical and geometric complexity of these devices. The details of the fluid motion through small channels, coupled with nonlinear interactions between macromolecules, cells, and the surface-dominated physics of the channels create very complicated phenomena, which can be difficult to simulate numerically.

Developing cutting-edge technologies requires cutting-edge tools. Various experimental techniques have

been developed for microfluidic research, such as laser Doppler velocimetry, the dual-slit method, optical Doppler tomography, scalar image velocimetry, molecular tagging velocimetry, and so on.^[6] Measurements with spatial resolutions smaller than 10 μm are necessary for making measurements in many MEMS applications. Only microparticle image velocimetry (μPIV) has demonstrated this high spatial resolution. The μPIV is a nonintrusive method that can be used to obtain high-resolution two-dimensional (2-D) full velocity fields instantaneously. By using a combination of advanced imaging and processing techniques, spatial resolutions on the order of single microns (and even smaller) can be achieved. μPIV also forms the basis for several applications in BioMEMS devices besides just velocity measurement. This article focuses on the applications of μPIV in biomedical research. We will give an overview of μPIV techniques first and then review its biomedical applications.

MICROPARTICLE IMAGE VELOCIMETRY

The experimental setup of a PIV system typically consists of several subsystems. Tracer particles are added to the flow in most applications, and these particles must be illuminated at least twice within a short time interval. The light scattered by the particles must be recorded either on a single camera frame or on a sequence of frames. The displacement of the particle images between the light pulses is determined through evaluation of the PIV recordings.

For evaluation, the particle images are divided into small interrogation windows. The local displacement vector for each window is determined by correlation-based statistical methods.^[7] Correlation-based PIV determines the average motion of small groups of particles contained within small interrogation windows, instead of the displacement of individual particles. The determination of the average particle displacement is accomplished by computing the spatial autocorrelation of a doubly exposed single image, or preferably the spatial cross-correlation of a pair of singly exposed images of the particle images over many interrogation windows. One displacement vector is calculated per interrogation window. For a multiple-exposure image,



the average displacement of the recorded particle image pairs is determined by computing the autocorrelation of the interrogation window. If the double frame/single exposure recording technique is employed, the cross-correlation function is used. An important advantage of the cross-correlation method is that this method removes the so-called directional ambiguity, which is inherent in the autocorrelation method.

Generally, after having acquired two sequential images of the flow, cross-correlation is employed to extract the velocity vectors from the images via an $N \times N$ pixel interrogation window, which is shown schematically in Fig. 1. The area of interest in the first frame is A and outside of this window is labeled A_0 . Their counterparts in the second frame are B and B_0 , respectively. An interrogation window $f(i, j)$ centered at a particular position (x, y) is sampled within A . Its counterpart is $g(i, j)$ within B . $f(i, j)$ and $g(i, j)$ are the gray value distributions of the two interrogation windows. The cross-correlation coefficient $R_{fg}(k, l)$ between the patterns $f(i, j)$ and $g(i, j)$ can be estimated by the convolution integral^[8]

$$R_{fg}(k, l) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} f(i, j)g(i+k, j+l) \quad (1)$$

A high cross-correlation value is observed when many particle images in the first image match their corresponding spatially shifted partners in the second image. Small cross-correlation peaks may be observed when individual particle images match with other particle images. The highest correlation peak is considered to represent the best match of particle images between the windows $f(i, j)$ and $g(i, j)$. The offset $(\Delta x, \Delta y)$ of the peak, which are measured from the origin of the cross-correlation coefficient map, represent the average spatial shift of particles from image A to image B at position (x, y) . Typical 2-D correlation function is shown in Fig. 8.

Micro Particle Image Velocimetry refers to the application of PIV to measure velocity fields of fluid motion

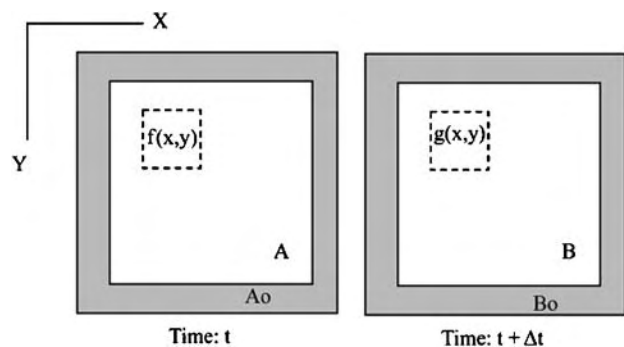


Fig. 1 Schematic interrogation windows of particle image velocimetry.

with length scales of order $100 \mu\text{m}$, and with spatial resolution of individual velocity measurements of order $1\text{--}10 \mu\text{m}$. In one of the predecessors to μPIV , flow-tracing particles were used estimate flow velocities in microscale geometries.^[9] A mercury arc lamp continuously illuminated $0.9 \mu\text{m}$ diameter fluorescent particles in an epi-fluorescent microscope. The length of the particle streaks in long-exposure images was used to estimate the local velocity. The resolution of the measurements was on the order of $10 \mu\text{m}$. The technique was extended by Santiago et al.^[10] using a mercury arc lamp to continuously illuminate 300 nm fluorescent particles. Discrete particle images were recorded by an intensified charge-coupled device (CCD) camera, using the intensifier as the shuttering mechanism. Correlation analysis was then applied to the particle-image field. This produced regularly spaced velocity data in contrast to Brody's work,^[9] where the measurements are spaced irregularly. To record particle images with continuous light sources, tens of milliseconds are required for exposure. This limits the magnitude of velocity that can be practically measured by μPIV using submicron flow tracing particles of the order $100 \mu\text{m}/\text{sec}$. Meinhart, Wereley, and Santiago^[11] used a pulsed Nd:YAG laser to illuminate 200 nm polystyrene beads. The pulsed illumination source creates effective exposure times equal to the pulse duration of the laser, $\sim 5 \text{ nsec}$. The pulsed illumination allows for much higher velocities to be measured ($\sim 10 \text{ m}/\text{sec}$) and smaller particles to be recorded. Meinhart, Wereley, and Santiago^[11] achieved a spatial resolution of $0.9 \mu\text{m}$. The first μPIV measurement in circular tubes was obtained by Koutsiaris, Mathioslakis, and Tsangaris.^[12] They used a 20 W Halogen lamp to illuminate a glycerol suspension of $10 \mu\text{m}$ diameter glass spheres flowing through $\sim 200 \mu\text{m}$ diameter glass capillaries. The $10 \mu\text{m}$ particle size allowed for relatively short exposure times with continuous illumination. The spatial resolution of these measurements was approximately $26 \mu\text{m}$. The μPIV technique has been applied to a variety of flow regimes ranging from fundamental studies of fluid motion to practical applications such as commercial inkjet print-heads, and BioMEMS devices. Fig. 2 is a schematic of a typical μPIV system, such as that originally described by Meinhart, Wereley, and Santiago.^[11] It is common to use a pulsed monochromatic light source, such as a pulsed Nd:YAG laser system. The laser system is specifically designed for PIV applications, and consists of two Nd:YAG laser cavities, beam combining optics, and a frequency doubling crystal. The laser emits two pulses of light at $\lambda = 532 \text{ nm}$. The time delay between light pulses can vary from hundreds of nanoseconds to a few seconds. The illumination is delivered to the microscope through beam-forming optics, which can consist of a variety of optical elements that will sufficiently modify the light so that it will fill the back of the

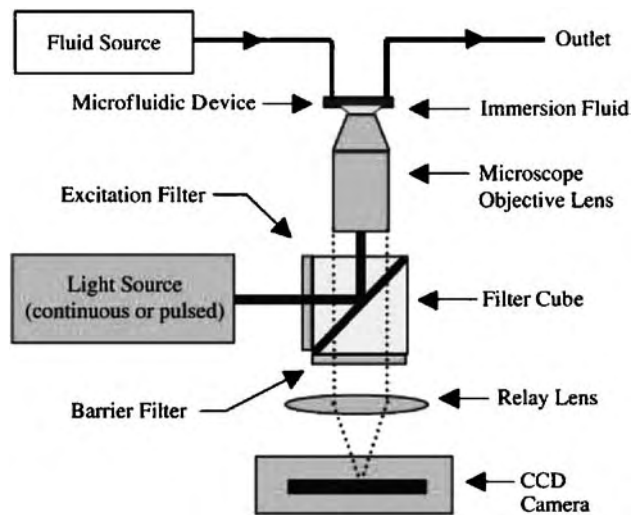


Fig. 2 Schematic of a μ PIV system using either a continuous or pulsed illumination light source. (From Ref. [13].)

objective lens, and thereby broadly illuminate the microfluidic device. The beam-forming optics can consist of a liquid-filled optical fiber to direct the light into the microscope or a simple beam expander. The illumination is reflected upward toward the objective lens by a dichroic mirror (designed to reflect 532 nm and transmit 612 nm).

Three fundamental differences of μ PIV from conventional macroscopic PIV are that the particles are small compared to the wavelength of the illuminating light; the particles are small enough that the effects of Brownian motion must be addressed; and the illumination source is typically not a light sheet, but an illuminated volume of the flow.^[11]

Because the diameter of the flow-tracing particles must be small enough that the particles not disturb the flow being measured, they can frequently be on the order of 50–100 nm. The diameter (d_p) is then 1/10 to 1/5 the wavelength (λ) of green light, and therefore approaching the Rayleigh scattering criteria where the amount of light scattered by a particle varies as d^6 . This makes it extremely difficult to record particle images. One solution to the imaging problem is to use epi-fluorescence imaging to record light emitted from fluorescently labeled particles, using an optical filter to remove the background light.

When the seed particle size becomes small, the collective effect of collisions between the particles and a moderate number of fluid molecules is unbalanced, resulting in a random particle motion known as Brownian motion, which prevents the particle from following the flow to some degree, causing a measurement error. It is an unbiased error that can be reduced by averaging over groups of particles.

The third significant difference between μ PIV and conventional PIV is the illumination mode. Conventional PIV uses light sheet illumination while μ PIV uses volume illumination because forming a light sheet inside a microscale device and aligning the light sheet with the focal plane would almost be impossible. Because the entire depth of the fluid is illuminated, the tracer particles within the volume are illuminated regardless of whether these particles are within the depth of field of the optical imaging system. Particle image size, intensity, and shape vary with distance from the focal plane. In conventional PIV, all illuminated particles are in focus because the depth of field is larger than the thickness of the light sheet. However, in μ PIV, the depth of field is determined by the optical system and camera characteristics.^[13]

The spatial resolution can be determined in terms of the depth of correlation. The depth of correlation is defined as the axial distance, Z_{corr} , from the object plane in which a particle becomes sufficiently out of focus so that it no longer contributes significantly to the signal peak in the particle-image correlation function. The expression of Z_{corr} , given by Wereley and Meinhart,^[13] is strongly dependent on numerical aperture, NA, and particle size, d_p , and is weakly dependent on magnification.

The resolution of μ PIV measurements can be increased further by applying a number of advanced evaluation techniques, which also work well for standard PIV systems. Measurement uncertainty of PIV data includes both bias and precision error. One of the most effective methods for reducing the bias error of PIV in complex flows is the central difference interrogation (CDI) method.^[14] Adaptive window offsetting is widely used with the fast fourier transform (FFT)-based correlation algorithm for reducing the evaluation error and with the image pattern tracking algorithms for increasing the spatial resolution. When using CDI, the first and second interrogation windows are shifted backwards and forwards, respectively, each by half of the expected particle image displacement. By applying CDI, the accuracy is improved to order Δt^2 (Δt is the time interval between the two images) instead of order Δt for conventional forward difference interrogation, by which only the first interrogation window is shifted forward.

To further reduce the measurement uncertainty, the random error must also be reduced. In a complex flow with both translation and rotation velocity, even when the flow is ideally seeded and the PIV recordings are made without any noise, evaluation errors may result from the deformation of the measured flow. To account for the deformation of the PIV image pattern, image correction techniques have been developed. Central difference image correction (CDIC), an efficient image correction algorithm, combines ideas of central

difference interrogation, adaptive continuous window shifting, and image correction, and enables a very reliable and accurate evaluation of digital PIV recordings.^[15] In the CDIC procedure, the particle image displacements at the four corners of each interrogation window are calculated based on previous iterations and used to deform the image patterns in the interrogation area for both exposures of the PIV recording pair using one of several interpolation schemes, so that the image patterns have an optimal match despite spatial velocity gradients at the particle image displacement.

The volume illumination used in μ PIV limits the number density of particles that can be used to trace the flow. If the particle density is too high, then background noise from out-of-focus particles will dominate the image and reduce the visibility of in-focus particles. The correlation function for a single-exposed PIV image pair has a peak at the position of the mean particle image displacement in the interrogation window, which should be the highest among all the peaks of $R_{fg}(k, l)$. Smaller subpeaks, which result from noise or mismatch of particle images, are usually obviously lower than the main peak, i.e., the peak of the particle displacement. However, when the interrogation window does not contain enough particle images or the noise level is too high, the main peak will become weak and may be lower than some of the “sub” peaks, causing an erroneous velocity vector. Because many of the low Reynolds number flows in microfluidic devices are either steady or periodic, it is sufficient to measure the time-averaged or phase-averaged velocity field. In steady or periodic flows, the main peak of $R_{fg}(k, l)$ is always at the same position for PIV recording pairs taken at different times while the subpeaks appear with random intensities and positions in different recording pairs. Therefore, when averaging R_{fg} for a sufficiently large number (N) of PIV recording pairs, the main peak will occur at the position of the particle image displacement with the same intensity level, but the subpeaks, occurring randomly, will disappear as N becomes large. By averaging the correlation function, the effective particle concentration is increased, so that the errors resulting from the low image density can be avoided.^[16]

FLOW FIELD MEASUREMENT

One straightforward application of μ PIV is to measure the velocity field inside a BioMEMS microfluidic device. In the present example, particle dynamics in response to both flow and electrical forces is studied using μ PIV.

DEP Particle Trap Device

It is important to determine the flow field at the micro-scale. Many biological processes involve chemical

reactions and the transport of macromolecules or cells through complicated channel networks. The details of the fluid motion in these small channels, coupled with interactions between macromolecules, cells, and the surface-dominated physics of the channels create complex phenomena, which can be difficult to simulate numerically. Understanding the flow patterns inside the micro channels by high resolution velocity measurement helps to understand the physics detail of the micro scale flow, characterize and quantify the performance of the BioMEMS device, and optimize the operating parameters.

Dielectrophoresis (DEP) is becoming a popular method for manipulating cells and other biological entities, such as virus, proteins, and DNA molecules. Dielectrophoresis has been used to separate blood cells and to capture DNA molecules.^[17] Despite the attention it is receiving, there is presently an incomplete understanding of the phenomenon, especially as it interacts with the hydrodynamic force on cells. The μ PIV technique was applied in a DEP device to characterize the flow field by Whitacre and Wereley.^[18]

The device was designed and manufactured at Purdue University.^[19] It has been designed to trap, separate, and concentrate biological components carried in solution. It consists of a microchannel with a depth of 11.6 μm , width of 350 μm , and length of 3.3 mm. The channel was anisotropically etched in silicon to produce a trapezoidal cross section and covered by a piece of anodically bonded glass. A schematic view and photo of the device are shown in Fig. 3. Bright regions represent platinum electrodes and the dark regions represent the electrode gaps. The electrodes are arranged in interdigitated pairs so that the first and third electrodes from Fig. 3 are always at the same potential. The second and fourth electrodes are also at the same potential, but can be at a different potential than the first and third electrodes. An alternating electric potential is applied to the interdigitated electrodes to create an electromagnetic field with steep spatial gradients. Particle motion through the resulting electric field gradients causes polarization of the suspended components, resulting in a body force that repels particle motion into increasing field gradients.

In the experiments, sample solutions were injected into the chamber using a syringe pump. The flow rate could be adjusted. An HP 33120A arbitrary waveform generator was used as the AC signal source to produce sinusoidal signal with frequency specified at 1 MHz. Images of particles were acquired using a standard μ PIV system, as shown in Fig. 2. A mercury lamp is used to illuminate the 0.7 μm polystyrene latex (PSL) microspheres (Duke Scientific) that are suspended in deionized water in concentrations of about 0.1% by volume. The particles are coated with a red fluorescing dye ($\lambda_{\text{abs}} = 542 \text{ nm}$, $\lambda_{\text{emit}} = 612 \text{ nm}$). The images were

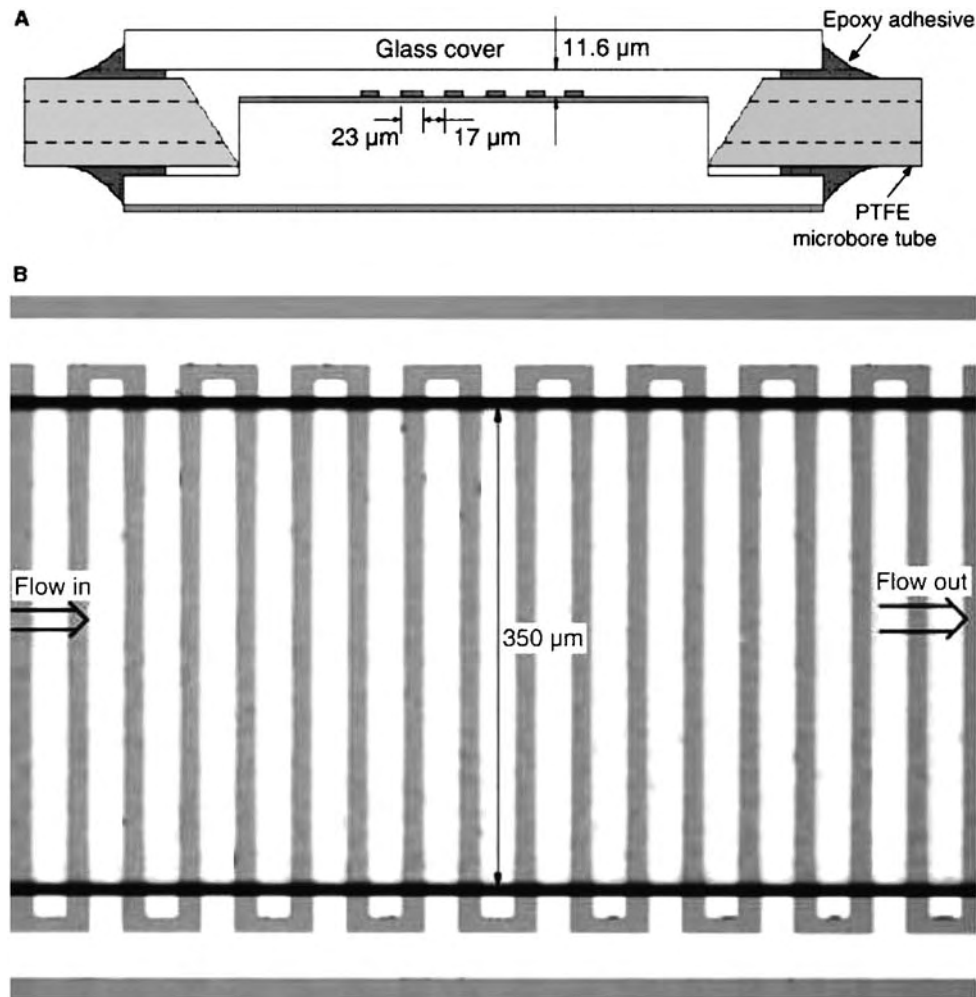


Fig. 3 (A) Schematic view of experimental apparatus and (B) photo of apparatus. (From Ref. [18].)

acquired using a Photometric CoolSNAP HQ interline transfer monochrome camera (Roper Scientific).

The experiments used six sets of 800 images each to analyze the effect of DEP on particle motion in the test device. These images are high quality with low readout noise, as can be seen in the example fluorescent image of Fig. 4. Fig. 4A demonstrates the many different particle intensity distributions, which are typically present in a μ PIV image in which the particles are distributed randomly within the focal plane. Fig. 4B shows how, as the result of the DEP force, the particles migrate to the top of the channel and all have nearly identical images. Fig. 4B also shows that when a significant DEP force exists, the particles are trapped at the electrode locations by the increase in DEP force there.

The PIV results for experimental cases ranging from 0.5 to 4.0 V are summarized in Fig. 5, which is a plot of the average axial velocities for the six electrode voltage cases. An interrogation region of 32 by 32 pixels and a grid spacing of 8 by 8 pixels were used. While this grid

spacing over samples the data, it provides a good means of interpolating between statistically independent measurements. The advanced interrogation options included CDI, continuous window shifting, ensemble correlation averaging, and a multiple iteration scheme set to two passes. The velocity curves show that particles move fastest over the centers of the electrodes and slowest on the leading edge of the electrodes (electrodes are located at 20–40, 60–80, 100–120 μ m, etc., from the inlet of the channel), which testifies to the feasibility of DEP particle trapping. The three lowest voltages show a trend of decreasing particle velocity in the downstream direction, which may be explained by the impediment by each electrode a particle encounters. Another interesting result observed from Fig. 5 is that the average particle velocity increases as the voltage increases from 0.5 to 2.0 V, but decreases as the voltage increases from 2.0 to 4.0 V. This phenomenon is explained by particles being displaced from the channel bottom into faster area of the fluid flow, which biases the velocity

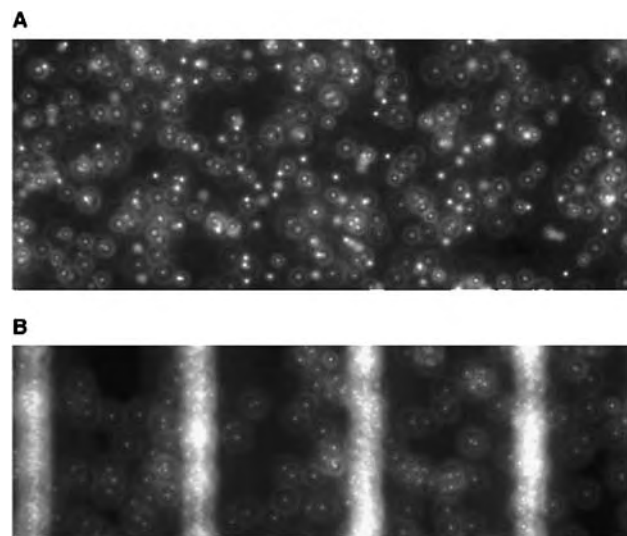


Fig. 4 Photo of PSL particles with 0.0 V (A) and 4.0 V (B). (From Ref. [18].)

distribution toward higher velocities, altering the shape of the cross-correlation peak to favor higher velocities even though the flow is constant. For higher voltages, the effect of particles being hindered by axial electric field gradients is compounded by particles being forced beyond the high-speed center of the flow profile by the DEP force.

Microviscometry

Based on the velocity measurement, many other related flow properties can be derived, such as shear stress, viscosity, pressure gradient, and even the wall location

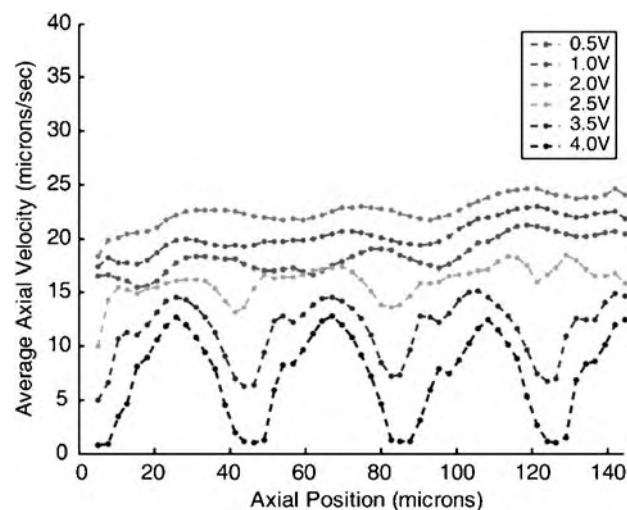


Fig. 5 Average axial velocity from PIV results for all electrode voltage cases (From Ref. [18].)

in microvessels. Applications of μ PIV to microvascular hemodynamics help to get many dynamic, kinematic, and rheological properties of the blood flow, which have been difficult or impossible to assess in microvessels in vivo.

A microviscometry method based on μ PIV is presented by Long et al.^[20] The central tenet of this approach is that distributions in the local viscosity, $\mu(r)$, as a function of the radial position, r , over the cross section of microvessels, can be determined analytically from the cross-sectional axial velocity distribution, $v_z(r)$, where $v_z(r)$ can be extracted from particle tracers in the flow by using μ PIV. In addition to the viscosity distribution, various flow parameters in microvessels in vivo can also be predicted by this method, including axial pressure gradient, dp/dz ; volume flow rate, Q ; and the relative apparent blood viscosity, η_{rel} , defined as the ratio of steady volume-flow rates per unit pressure drop of blood plasma relative to whole blood. Quantitative validation of microviscometric methods is given and a decrease in local and apparent blood viscosity in individual microvessels as a direct result of reducing systemic hematocrit is shown. They also detect the presence of the endothelial-cell surface layer and estimate its hydrodynamically relevant thickness by least square fitting to the μ PIV data. This idea is first proposed by Stone et al.,^[20a] named “microfluidic nanoscope” for measuring the shape of internal cavities and flow channels with a spatial resolution of tens of nanometers. This technique uses μ PIV to measure the motion of fluid near a cavity wall. This information, combined with a model for the fluid’s boundary condition, typically the no-slip boundary condition, can be used to determine the wall’s location and topology.

Typically, rheological data are extracted from force measurements, which do not provide the detailed spatial distributions that this approach offers. The idea of using a velocity field to extract rheological data has applications beyond microvascular hemodynamics to a variety of generalized linearly viscous fluids and particle-suspension flows, including polymer melts, emulsions, extrusion, and polymeric drag reduction.

The small fluorescent particles ($<0.1 \mu\text{m}$) used in this method increase the spatial resolution in rheology studies and increase the size of particle fluctuations ($<1/r$), and therefore the random error. Fluorescent particles allow viscoelasticity to be probed in complex optical objects, e.g., allowing them to be easily located within living cells. However, fluorescence studies can encounter challenges with respect to the chemistry of the fluorescent tags. Photobleaching gives a finite time limit for the experiments. It is important to quantify the degree of noise inherent in PIV experiments, because this defines the largest moduli (smallest amplitude of particle fluctuations) that can be measured. Superior calibration of the PIV experiments can be achieved with

respect to the viscosity of a range of standard Newtonian fluids, e.g., glycerol/water mixtures at a range of concentrations. The dependence of the microrheological measurements on the fluorescent particle size and on its chemistry need to be carefully considered. Thus, a microrheologist must study the interactions of their probe particle with the complex fluids under examination. Ideally, the particle chemistry should have a negligible influence on the result of a rheological measurement.

Similar to that above, μ PIV can also be applied to measure the red blood velocity field,^[21] to elucidate the effects of disturbed flow on monocyte adhesion to endothelial cells by characterizing the disturbed flow in a vertical-step flow chamber,^[22] to study the deformation of DNA molecules under homogeneous extensional flow,^[23] and so on. A similar particle-tracking technique^[24] is also used in microrheology. It has been used to probe the dynamics of membrane systems^[25] and to measure the local mechanical properties of cells and cytoplasm.^[26] A novel method for measuring the microrheology of soft viscoelastic media, based on cross-correlating the thermal motion of pairs of tracking particles is demonstrated by Crocker et al.^[27] This method does not depend on the exact nature of the coupling between the tracers and the medium, and yields accurate rheological data for highly inhomogeneous materials.

PARTICLE DETECTION

Using the μ PIV technique, we can obtain not only velocimetry data by conventional (correlation-based) PIV analysis, but also other useful information by different analysis methods. For example, the peak location of particles in the PIV images can be found by Gaussian peak fitting. The point response function of the microscope objective lens can be approximated well by the Gaussian function.^[28] The particle image-intensity distribution can be derived by convolving the point-spread function with the particle geometric-image distribution. Assuming the intensity distribution of the particle geometric image is also a Gaussian function, which is the most common approximation used in PIV analysis, a Gaussian function would be a good approximation to the intensity distribution of a particle image. By using the Gaussian function to fit the particle image intensity function, several parameters of a single particle can be extracted: the peak intensity of the particle image, peak location, and particle diameter.^[29]

Concentration Measurement

This technique is used to study the Brownian particle migration in a microchannel tube flow.^[29] A rigid

spherical particle translating at a small tube Reynolds number in a shearing flow experiences lateral migration owing to small inertia and wall effect. In a suspension flow of sufficiently small particles, Brownian motion is another factor. The Peclet number is a dimensionless parameter to describe the relative importance of Brownian motion with respect to shear-induced migration, which is conceptually defined as

$$Pe = \frac{6\pi\mu\gamma a^3}{KT} \quad (2)$$

where μ is the absolute viscosity of the fluid; γ is the shear rate; a is the particle radius; K is the Boltzmann's constant; and T is the temperature. Large Peclet numbers imply that the shear forces dominate Brownian diffusion while the case is opposite for small Peclet numbers.

In this work, a dilute suspension flow (particle volume fractions $\ll 0.01$) of Brownian particles is studied using a μ PIV system to directly image particle concentrations and velocity. The suspensions consist of fluorescent PSL particles and water. The mean particle diameter d_p are 0.69, 1.0, and 3.0 μm . Seventy-five micrometer diameter capillary tubing made of fluorinated ethylene propylene (FEP) Teflon is used as the flow channel. As the refractive index of FEP is very close to that of water, the tubing is immersed into water to match the refractive index. The flow inside the tubing is driven by a pressure head. The flow velocity is varied, which results in the bulk Peclet number ranging over four orders of magnitude.

To study the spatial particle distribution, the locations of all the particles in each image are extracted by fitting the Gaussian function to the particle image intensity function. When a local maximum in an image is detected, the intensity matrix $I(X, Y)$ around that maximum is extracted, then a least squares curve fit is applied to get the parameters that describe the particle characteristics: the peak intensity of the particle image, peak location, and particle image diameter. After each particle image is detected, it is removed from the original image and the next particle is found. From the locations of the detected particles, the concentration distribution can be calculated and expressed as a probability density function.

The velocity is calculated using the software package EDPIV, developed by Dr. Lichuan Gui.^[15] Two hundred image pairs of images are used to ensemble-average velocity fields in correlation space to reduce random errors owing to image noise and tracer particle Brownian motion.

The relative concentration profiles of the 3.0 μm particles are shown in Fig. 6. The maximum tube velocity is in the range of 0.009–0.21 m/sec, corresponding to Peclet numbers between 3820 and 89700 and particle



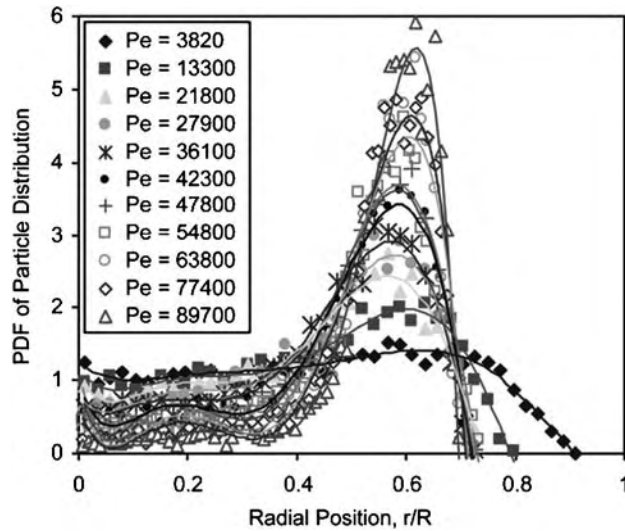


Fig. 6 PDF (relative concentration) of $3.0\ \mu\text{m}$ particles. (From Ref. [29].)

Reynolds numbers between 0.013 and 0.32. It has been found that when Pe is small, particles migrate away from the channel wall, and that the particle concentration in regions remote from the wall is nearly uniform. The dominant mechanism is Brownian motion; when Pe increases, particles move away both from the wall and the center, toward a preferred position (r^*) with the migration effect becoming progressively stronger as Pe increases. The dominant mechanism is the shear-induced migration. We can also see that there is a particle-free zone near the channel wall and this zone expands as the velocity increases. A quantitative analysis based on the behavior of three different sizes of particles shows that the transition between these behaviors happens between Pe 1000 and 100,000.^[29] The radial velocity data agrees with this migration behavior by showing that particles move toward the equilibrium position r^* .

Many BioMEMS devices work with “particles” (cells, bacteria, DNA, etc.) suspended in fluids. The performance of some of those devices strongly depends on particle distribution as they pass through the sensor channels, such as the multifrequency particle impedance characterization system developed by Fuller et al.^[30] The work described above gives a good understanding of the migration of particles small enough to be susceptible to Brownian motion. It is very helpful to characterize and improve the performance of those BioMEMS devices.

It is also found that the equilibrium position r^* has a dependence on the particle size, it decreases as the particle-to-channel size ratio increases as is shown in Fig. 7. Based on this knowledge, a shear-activated cell sorter can be proposed. The advantage of this sorter is

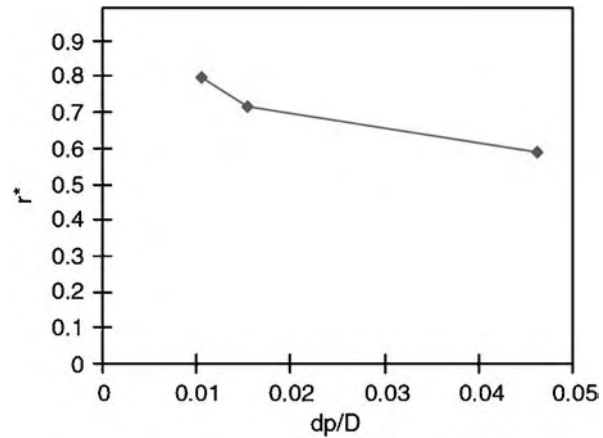


Fig. 7 r^* vs. d_p/D (r^* decreases as the particle-channel size ratio increases). (From Ref. [29].)

the simple working mechanism—only hydrodynamic forces are needed.

Diffusion Measurement

This particle-detection method can also be used to measure the diffusion coefficients. Particle image diffusiometry is proposed by Gorti et al. Virus identification with particle image diffusiometry (submitted for publication) to directly measure the diffusion coefficients of submicron particles as a means of pathogen detection.

The mean square displacement traveled by a particle with diffusivity D owing to diffusion over a time period Δt is given by

$$\langle s^2 \rangle = 2nD\Delta t \quad (3)$$

where n is the number of translational degrees of freedom. They obtained 700 images with 300 msec between frames for each set. The particles in the images are detected by Gaussian fitting discussed earlier in this paper, and the center location of each particle is extracted. Because the volume fraction of the solution used here is extremely low (the particles images are separated by at least 10 particle diameters), the particle images in one image can be easily matched with those in the next image so that individual particles can be tracked over long times. From the trajectory of each particle, a large amount data of distance traveled by the particle over Δt (300 msec) can be obtained. This displacement data formed a Gaussian function and can be used to estimate $\langle s^2 \rangle$ by statistical means. In the 2-D flow field, the diffusion coefficient is one-fourth the slope of $\langle s^2 \rangle$ vs. t curve [Eq. (3)].

At a constant temperature, the diffusion coefficient is a function of the drag on the particle. If the particles

are functionalized with antibodies against a specific analyte and introduced into a sample containing that analyte, binding of the analyte with the particles will increase its hydrodynamic force, which results in a decrease in diffusion. The reduction in diffusion is correlated with the amount of analyte present, so that the amount of analyte can be sensed by measuring the diffusion coefficient by this particle image diffusometry technique. It is shown (Fig. 5, Ref. Gorti et al. Virus identification with particle image diffusometry, submitted for publication.) that analytes as small as 1 nm can cause measurable changes in the diffusion coefficients of appropriate particles, and therefore can be detected.

Typically, affinity biosensors use “sandwich” assays for detection and measurement. One drawback of sandwich assays is the difficulty in producing two antibodies that each bind to different sites of an antigen, and the probability of two binding events on a single antigen is significantly less than that of a single binding event. This method avoids the sandwich assay, as only one binding event is needed, and offers several advantages: high resolution, multianalyte detection capability, avoidance of sandwich assays, less preparation time, and quicker results. It is a more direct approach to biosensing compared to many other methods, such as immunoassays,^[31] enzyme-catalyzed reactions,^[32] nucleic acid-based reactions, use of biomimetic materials, and membrane protein-based detection.

CHARACTERISTICS OF CORRELATION FUNCTION

Correlation-based PIV analysis detects the peak location of the correlation function to obtain the displacement of the interrogation window. However, the correlation function carries more information than just the peak location, for example, the peak width.

It is noted earlier in this paper that for small tracer particles and low speed flows, Brownian particle motion is measurable and could be significant enough to introduce considerable error into the velocity measurements. Because it was viewed as an undesirable effect on velocity data, the effect of Brownian motion on the PIV correlation function—namely a width-wise spreading of the correlation peak, which added to the uncertainty of locating the peak center—was often intentionally and substantially reduced through ensemble averaging over multiple images. However, Olsen and Adrian^[33] developed the theoretical relationships with which such spreading of the correlation peak could be used to deduce fluid temperature. Furthermore, because peak location (which yields velocity information) and peak width are independent parameters, it is conceivable that

PIV may be used to simultaneously measure both temperature and velocity.

The diffusivity D of particles of diameter d_p immersed in a liquid of temperature, T , and dynamic viscosity, μ , is given by expression

$$D = KT/3\pi\mu d_p \quad (4)$$

Combining Eqs. (3) and (4), it is observed that an increase in fluid temperature, with all other factors held constant, will result in a greater expected particle displacement, $\sqrt{\langle s^2 \rangle}$. However, absolute viscosity, μ , is a strong function of temperature and $\langle s^2 \rangle \propto T/\mu$. The effect of this relationship depends on the phase of the fluid. For a liquid, increasing temperature decreases absolute viscosity—so the overall effect on the ratio T/μ follows the change in T . For a gas, however, increasing temperature increases absolute viscosity, meaning that the effect on T/μ , and hence $\langle s^2 \rangle$, would need to be determined from fluid-specific properties.

Brownian motion necessitates that, in a given elapsed time window, all particles immersed in a fluid undergo slight random variations in displacement, resulting in a correlation function with a predictable and measurable peak-width increase over the case of no Brownian motion. The relationship between a particle’s actual cross-sectional area and that of its image is governed by the characteristics of the imaging optics. Olsen and Adrian^[33] developed an expression for the correlation peak width based on particle diameter, imaging optics characteristics, and Brownian motion. Hohreiter et al.^[34] extended this expression to allow the temperature to be extracted from the correlation peak width (Fig. 8) shows the possible variations in peak width owing to Brownian motion for a pair of experimental PIV images.

This microparticle image thermometry idea is elucidated and applied experimentally in a zero-velocity particle-seeded water cell to measure temperature over the range 20–50°C.^[34] When compared with thermocouple-measured temperatures, results agreed within a range of $\pm 3^\circ\text{C}$.

Because volume illumination is employed in μPIV , the out-of-focus particles negatively contribute to the image by effectively varying the particle image size and luminous intensity, reducing the signal-to-noise ratio, and ultimately decreasing the strength of the correlation.^[35,36] However, these effects are present in both the auto-correlation and cross-correlation peak shapes, and hence can be removed from consideration by using the difference of the cross-correlation and autocorrelation peak widths to infer temperature. The time delay, Δt , must satisfy two criteria. It needs to be: 1) sufficiently long to allow measurable particle motion to take place, yet and 2) short enough that a significant fraction of the particles in the first image also be present in the second. The former criterion is imposed by Brownian particle

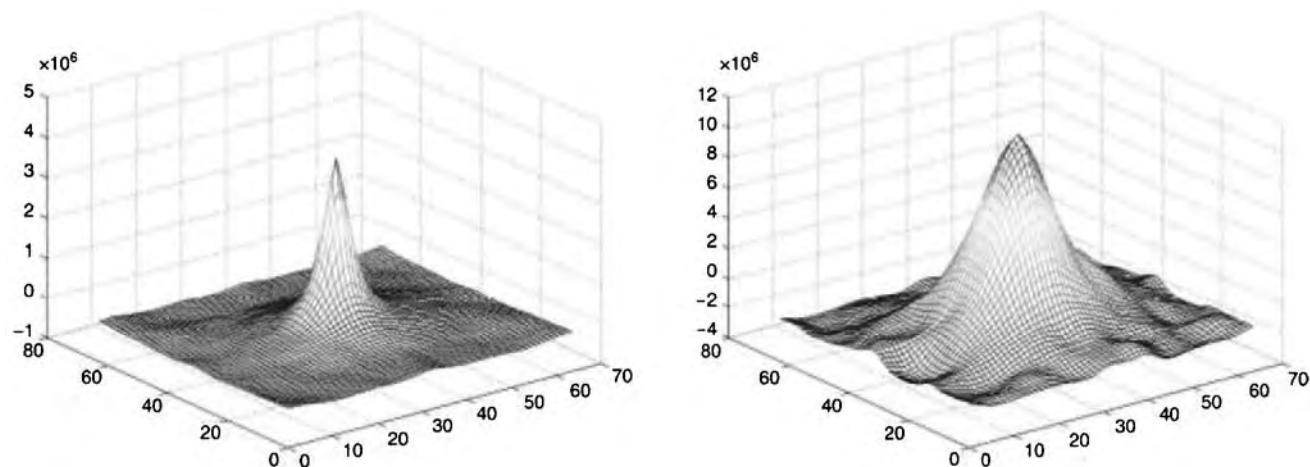


Fig. 8 A pair of correlation functions demonstrates possible variations in peak width. (From Ref. [34].)

motion and the latter by bulk fluid motion. Because the time interval between images cannot be too short, the sensitivity to dynamic change in temperature is not high. Experimental error for PIV thermometry is tied closely to the optical resolution of the PIV system being used—the higher the optical resolution, the more accurately fluid temperatures can be measured using this technique.

Thermal management is an important function required of emerging lab-on-chip systems as well as heat-producing integrated circuits. A variety of functions, such as DNA amplification and sensing, often require precise temperature control to guarantee successful and repeatable performance. For instance, PCR devices require operation at two or three different temperatures with a precision of $\pm 1^\circ\text{C}$.^[37,38] To monitor the temperature inside microfluidic channels, many methods were applied. A technique using the temperature dependence of the fluorescein dye intensity was first successfully employed by Ross, Gaitan, and Locascio to measure longitudinal temperature distributions in microchannels;^[39] another technique was developed using the temperature dependency of the conformational state of *N*-alkyl-substituted acrylamides polymer resulting in solubility below and precipitation above distinct (critical) temperatures.^[40] Including the microparticle image thermometry stated above, all three techniques allow visualizing thermal patterns in microfluidic systems under operating conditions. They can all be recommended for temperature measurements in optically accessible systems. The polymer method is ideal for detecting isotherms in steep temperature gradients close to heat sources. The fluorescein-based thermometry and the microparticle image thermometry have a comparable precision of about $\pm 3^\circ\text{C}$. The microparticle image thermometry has higher sensitivity over larger temperature ranges while the fluorescein-based thermometry has

higher sensitivity at a lower temperature range. The microparticle image thermometry technique can be introduced to monitor and control the temperature inside the flow systems. The advantages are that it is a nonintrusive method and it can measure both the temperature and the velocity simultaneously.

Because the cross-correlation peak area change is directly related to the expected particle displacement $\sqrt{\langle s^2 \rangle}$ and the temperature, T , through the diffusivity, D , it is conceivable that the diffusivity can be measured by this method too. It is another promising diffusometry method.

CONCLUSIONS

μPIV is a powerful, high-resolution measurement tool in the microfluidics field. The applications of μPIV are reviewed in this article. They can be categorized into three categories by the different analysis method that is used.

1. μPIV can be used to measure the flow field inside BioMEMS devices, microcirculation, and other biological environments. Both velocimetry data and other flow-related properties can be obtained, such as viscosity, shear stress, and pressure gradient. The velocimetry data can also be used to locate the wall position inside the microvessel.
2. Single particle images in a μPIV image can be detected by a Gaussian fitting. Particle location and particle size can be extracted. The motion and concentration of particles can be obtained. Microparticle image diffusometry, based on the Gaussian particle fitting, can be used to measure the diffusion coefficient and then detect the pathogen.

- Variation of diffusivity can result in variation of the PIV correlation function. By measuring the peak area change of the correlation function, we can measure diffusion-related parameters such as temperature and diffusion coefficient.

ARTICLES OF FURTHER INTEREST

Blood Vessel Mechanics; Flow Cytometry; Vascular Fluid Dynamics

REFERENCES

- Cheng, S.B.; Skinner, C.D.; Taylor, J.; Attiya, S.; Lee, W.E.; Picelli, G.; Harrison, D.J. Development of a multichannel microfluidic analysis system employing affinity capillary electrophoresis for immunoassay. *Anal. Chem.* **2001**, *3* (7), 1472–1479.
- Webster, J.R.; Burke, D.T.; Burns, M.A.; Mastrangelo, C.H. Electrophoresis system with integrated on-chip fluorescence detection. *Proc. IEEE Micro Electro Mech. Syst.* **2000**, 306–310.
- Bhatia, S.N.; Balis, U.J.; Yarmush, M.L.; Toner, M. Probing heterotypic cell interactions: hepatocyte function in microfabricated co-cultures. *J. Biomater. Sci. Polym.* **1996**, *9* (11), 1137–1160.
- Cruz, I.; Ciudad, J.; Cruz, J.J.; Ramos, M.; Gomez-Alonso, A.; Adansa, J.C.; Rodriguez, C.; Orfao, A. Evaluation of multiparameter flow cytometry for the detection of breast cancer tumor cells in blood samples. *Am. J. Clin. Pathol.* **2005**, *123* (1), 66–74.
- Northup, M.A.; Hills, R.F.; Landre, R.; Lehew, H.D.; Watson, R.A. A MEMS-based DNA analysis system. *Transducers'95, Eighth International Conference on Solid State Sensors and Actuators*, Stockholm, Sweden, June, 1995; 764–767.
- Wereley, S.T.; Gui, L.; Meinhart, C.D. Advanced algorithms for microscale particle image velocimetry. *AIAA J.* **2002**, *40* (6), 1047–1055.
- Raffel, M.; Willert, C.E.; Kompenhans, J. *Particle Image Velocimetry; A Practical Guide*, 3rd Ed.; Springer-Verlag: Berlin Heidelberg, Germany, 1998.
- Keane, R.; Adrian, R.J. Theory of cross-correlation analysis of PIV images. *Appl. Scientific Res.* **1992**, *49* (3), 191–215.
- Brody, J.P.; Yager, P.; Goldstein, R.E.; Austin, R.H. Biotechnology at low Reynolds numbers. *Biophys. J.* **1996**, *71*, 3430–3441.
- Santiago, J.G.; Wereley, S.T.; Meinhart, C.D.; Beebe, D.J.; Adrain, R.J. A micro particle image velocimetry system. *Exp. Fluids.* **1998**, *25* (4), 316–319.
- Meinhart, S.D.; Wereley, S.T.; Santiago, J.G. PIV measurements of a microchannel flow. *Exp. Fluids* **1999**, *27*, 414–419.
- Koutsiaris, A.G.; Mathioslakis, D.S.; Tsangaris, S. Microscope PIV for velocity-field measurement of particle suspensions flowing inside glass capillaries. *Meas. Sci. Technol.* **1999**, *10*, 1037–1046.
- Wereley, S.T.; Meinhart, C.D. Micro-resolution particle image velocimetry. In *Micro- and Nano-Scale Diagnostic Techniques*; Breuer, K., Ed.; Springer-Verlag: New York, 2003.
- Wereley, S.T.; Meinhart, C.D. Second-order accuracy particle image velocimetry. *Exp. Fluids* **2001**, *31*, 258–268.
- Wereley, S.T.; Gui, L. PIV Measurement in a four-roll-mill flow with a central difference image correction (CDIC) method, 4th International Symposium on Particle Image Velocimetry, Göttingen, Germany, Sept 2001.
- Meinhart, C.D.; Wereley, S.T.; Santiago, J.G. A PIV algorithm for estimating time-averaged velocity fields. *J. Fluids Eng.* **2000**, *122*, 285–289.
- Miles, R.; Belgrader, P.; Bettencourt, K.; Hamilton, J. Nasarabadi. Dielectrophoretic manipulation of particles for use in microfluidic devices. In *MEMS-Vol. 1, Micro-electromechanical Systems (MEMS)*, Proceedings of the ASME International Mechanical Engineering Congress and Exposition, Nashville, TN, Nov 14–19, 1999.
- Whitacre, I.; Wereley, S.T. Particle dynamics in a dielectrophoretic microdevice. In *BioMEMS and Biomedical Nanotechnology; Biomolecular Sensing, Processing and Analysis*; Ferrari, M., Bashir, R., Wereley, S.T., Eds.; Kluwer: Boston, Vol. 4, Part III *in press*.
- Li, H.B.; Zheng, Y.; Akin, D.; Bashir, R. Characterization and modeling of a microfluidic dielectrophoresis filter for biological species. *J. Microelectromechanical Sys.* **2005**, *14* (1), 103–112.
- Long, D.S.; Smith, M.L.; Pries, A.R.; Ley, K.; Damiano, E.R. Microviscometry reveals reduced blood viscosity and altered shear rate and shear stress profiles in microvessels after hemodilution. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 10060–10065.
- Stone, S.W.; Meinhart, C.D.; Wereley, S.T. A microfluidic-based nanoscope. *Exp. Fluids.* **2002**, *33* (5), 613–619.
- Sugii, Y.; Nishio, S.; Okamoto, K. *In vivo* PIV measurement of red blood cell velocity field in microvessels considering mesentery motion. *Physiol. Meas.* **2002**, *23*, 403–416.
- Chiu, J.J.; Chen, C.N.; Lee, P.L.; Yang, C.T.; Chuang, H.S.; Chien, S.; Usami, S. Analysis of the effect of disturbed flow on monocytic adhesion to endothelial cells. *J. Biomech.* **2003**, *36*, 1883–1895.
- Wong, P.K.; Lee, Y.K.; Ho, C.M. Deformation of DNA molecules by hydrodynamic focusing. *J. Fluid Mech.* **2003**, *497*, 55–65.
- Tseng, Y.; An, K.M.; Wirtz, D. Microheterogeneity controls the rate of gelation of actin filament networks. *J. Biol. Chem.* **2002**, *277*, 18143–18150.
- Saxton, M.J.; Jacobson, K. Single-particle tracking: applications to membrane dynamics. *Biomol. Struct.* **1997**, *26*, 373–399.
- Valentine, M.; Weitz, D.A. Local mechanical properties of cells and cytoplasm; <http://www.deas.harvard.edu/projects/weitzlab/research/cells.html> (accessed August 2005).
- Crocker, J.C.; Valentine, M.T.; Weeks, E.R.; Gisler, T.; Kaplan, P.D.; Yodh, A.G.; Weitz, D.A. Two-point microrheology of inhomogeneous soft materials. *Phys. Rev. Lett.* **2000**, *85*, 888–891.



28. Adrian, R.J.; Yao, C.S. Pulsed laser technique application to liquid and gaseous flows and the scattering power of seed materials. *Appl. Optics* **1985**, *24*, 44–52.
29. Cao, J.; Wereley, S.T. *Brownian Particle Distribution in Tube Flows*, Proceedings of the ASME International Mechanical Engineering Congress and Exposition, Anaheim, CA, Nov 13–19, 2004.
30. Fuller, C.K.; Hamilton, J.; Ackler, H.; Krulevitch, P.; Boser, B.; Eldredge, A.; Becker, F.; Yang, J.; Gascoyne, P. Microfabricated multi-frequency particle impedance characterization system. *Micro Total Anal. Syst.* **2000**, *11*, 265–268.
31. Anderson, G.P.; King, K.D.; Gaffney, L.H.; Johnson, L.H. Multi-analyte interrogation using the fiber optic biosensor. *Biosens. Bioelectronics* **2000**, *14*, 771–777.
32. Eu, J.Y.; Wang, C.Y.; Andrade, J. Homogeneous bioluminescence assay for Galactosuria: interference and kinetic analysis. *Anal. Biochem.* **1999**, *271* (2), 168–176.
33. Olsen, M.G.; Adrian, R.J. Brownian motion and correlation in particle image velocimetry. *Opt. Laser Technol.* **2000**, *32*, 621–627.
34. Hohreiter, V.; Wereley, S.T.; Olsen, M.G.; Chung, J.N. Cross-correlation analysis for temperature measurement. *Meas. Sci. Technol.* **2002**, *13*, 1072–1078.
35. Meinhart, C.D.; Wereley, S.T.; Gray, M.H.B. Volume illumination for two-dimensional particle image velocimetry. *Meas. Sci. Technol.* **2000**, *11*, 809–814.
36. Olsen, M.G.; Adrian, R.J. Out-of-focus effects on particle visibility and correlation in microscopic particle image velocimetry. *Exp. Fluids* **2000**, *29*, 166–174.
37. Wittwer, C.T.; Fillmore, G.C.; Garling, D.J. Minimizing the time required for DNA amplification by efficient heat transfer to small samples. *Anal. Biochem.* **1990**, *186*, 328–331.
38. Chiou, J.; Matsudaira, P.; Sonin, A.; Ehrlich, D. A closed-cycle capillary polymerase chain reaction machine. *Anal. Chem.* **2001**, *79* (3), 2018–2021.
39. Ross, D.; Gaitan, M.; Locascio, L.E. Temperature measurement in microfluidic systems using a temperature-dependent fluorescent dye. *Anal. Chem.* **2001**, *73*, 4117–4123.
40. Seger-Sauli, U.; Panayiotou, M.; Schnydrig, S.; Jordan, M.; Renaud, P. Temperature measurements in microfluidic systems: heat dissipation of negative dielectrophoresis barriers. *Electrophoresis* **2005**, *26* (11), 2239–2246.

Microporous Materials

Julian R. Jones
Larry L. Hench

Department of Materials, Imperial College London, London, United Kingdom



INTRODUCTION

When a significant loss of tissue occurs as a result of trauma or by the excision of diseased or cancerous tissue, total healing occurs only with the aid of graft implants. Therapies employing graft material retrieved from a different site in a patient (autograft), from another human donor (homograft), or from other living or nonliving species (heterografts or xenografts) have been restricted by limited material availability, complicated multistage surgery at the detriment of the harvest site, and the risk of disease transmission. These factors create a great demand for synthetic substitutes specially designed and manufactured to act as a scaffold for tissue engineering and regeneration. One of the main criteria for a scaffold is that it is composed of an interconnected network of pores with diameters in excess of 100 μm (microporous materials) to allow tissue ingrowth, vascularization, and nutrient delivery. This entry begins by describing all the properties of an ideal scaffold and reviews processes that have been used in an attempt to create microporous ceramics and glasses for various applications, from fuel filters to biomedical scaffolds, and compares them to the criteria of an ideal scaffold. The entry continues by describing the latest advances in microporous materials that have the potential to be used to regenerate damaged tissues to their natural state and function and ends with an assessment of the attainability of the ideal scaffold.

TISSUE ENGINEERING AND REGENERATION

At present, the gold standard for surgeons repairing damaged bone is to restore it to its natural form and function by harvesting the patient's tissue from a donor site and transplanting it to a host site, at times even maintaining blood supply. The future goal is to be able to eliminate the need for transplants, but still repair tissue to its natural form.^[1,2] Tissue regeneration techniques involve the use of a scaffold that can be implanted into a defect to guide and stimulate tissue regrowth in situ. The scaffold should resorb (dissolve) as the tissue grows, leaving no trace of damage. In tissue engineering applications, the scaffolds are seeded

with cells in vitro to produce the basis of a tissue before implantation;^[2,3] cells could be extracted from a patient, seeded on a scaffold of the desired architecture, and the replacement tissue grown in the laboratory, ready for implantation.^[4]

AN IDEAL SCAFFOLD

An ideal scaffold is one that mimics the extracellular matrix of the tissue that is to be replaced so that it can act as a template in three dimensions onto which cells attach, multiply, migrate, and function. The criteria for an ideal scaffold for bone regeneration are that it:

1. Is made from a material that is biocompatible, i.e., not cytotoxic.
2. Acts as template for tissue growth in three dimensions.
3. Has an interconnected microporous network containing pores with diameters in excess of 100 μm for cell penetration, tissue ingrowth and vascularization, and nutrient delivery to the center of the regenerating tissue on implantation.^[5,6]
4. Bonds to the host tissue without the formation of scar tissue, i.e., is made from a bioactive and osteoconductive material.
5. Exhibits a surface texture that promotes cell adhesion, adsorption of biological metabolites.^[7]
6. Influences the genes in the bone-generating cells to enable efficient cell differentiation and proliferation
7. Resorbs at the same rate as the tissue is repaired, with degradation products that are nontoxic and that can be easily excreted by the body, for example, via the respiratory or urinary systems.
8. Is made from a processing technique that can produce irregular shapes to match that of the defect in the bone of the patient.
9. Exhibits mechanical properties sufficient to be able to regenerate tissue in the particular application such as bone in load-bearing sites.
10. Has the potential to be commercially producible to the required ISO (International Standards Organization) or FDA (Food and Drug Administration) standards.

11. Can be sterilized and maintained as a sterile product to the patient.
12. Can be produced economically to be covered by national and/or private healthcare insurances.

For bone regeneration, where the scaffold is implanted directly into a bone defect, mechanical properties of the scaffold are critical and the modulus and elastic strength of the microporous material should match that of the natural bone. If a scaffold with a modulus much lower than the host bone is implanted into a load-bearing site, the scaffold will fracture. If the modulus of the scaffold is much higher than bone, the load will be transmitted through the scaffold instead of the bone (stress shielding), causing bone resorption rather than bone regeneration. Stress shielding causes bone resorption because the body is continually remodeling bone according to the stress to which the bone is subjected. Osteogenic cells lay down new bone matrix to strengthen bone under high loads. When bone is not loaded, other cells called osteoclasts resorb it. This is one of the most common reasons for total hip replacements to fail.^[3]

For tissue engineering applications, only the mechanical properties of the final tissue engineered construct is critical.

BIOACTIVE MATERIALS

When bioinert materials are placed in the body, scar (fibrous) tissue surrounds the implant. Scar tissue will hinder repair and affect functionality of the tissue. Bioactive materials bond to tissue on implantation. The most commonly used bioactive materials are ceramics and glasses. Class A bioactive materials bond to both hard and soft tissue, whereas class B bioactive materials bond only to bone.^[3,7] Bioactive glasses are class A bioactive materials that are biocompatible, bioactive, resorbable, osteoconductive (allows bone to grow along its surface), and even osteoinductive (encourages bone growth away from the original bone surface).^[7]

Fig. 1 shows a scanning electron (SEM) micrograph of rabbit trabecular bone. Trabecular bone is a natural composite of collagen fibers and microporous bone mineral (55–70% interconnected porosity), which is similar to synthetic hydroxyapatite (HA) ceramic, $\text{Ca}_{10}(\text{PO}_4)_6(\text{OH})_2$.^[1,8] Therefore, it is desirable to produce a scaffold that mimics the morphology of the bone mineral and then can regenerate it. Synthetic HA and tri-calcium phosphate (TCP) have gained great attention as fillers, spacers, and bone graft substitutes in orthopedic and maxillofacial applications.^[7] They are class B bioactive materials that are osteoconductive. Synthetic stoichiometric HA dissolves very slowly

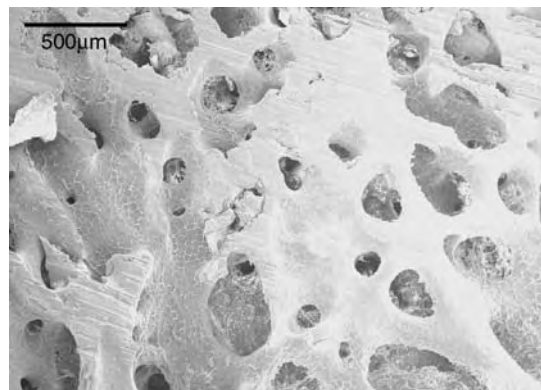


Fig. 1 SEM micrograph of the microporous structure of rabbit trabecular bone.

in the body, so slowly that it is usually considered as not resorbable. In contrast, TCP resorbs at a very high rate, sometimes too fast for bone regeneration applications.

HA, TCP, and bioactive glasses have been used successfully in the clinic as bone filler materials in powder form,^[7,8] however, the challenge is to develop them into 3-D microporous scaffolds with the properties listed earlier. The following sections will focus on that challenge.

HA AND TCP CERAMICS

Sintering

The simplest way to generate porous scaffolds from ceramics such as HA or TCP is to sinter particles. Particles are usually mixed with a wetting solution, such as poly(vinyl alcohol), to aid compaction. The powders are compacted by cold isostatic pressing the particles into cylinders at approximately 200 MPa^[9] to form a green body. The gaps between the particles are the pores and the green body will have low mechanical properties. Mechanical properties of the materials are improved by sintering at high temperatures, i.e., the material is raised to high temperature (e.g., 1200°C) in a furnace. As sintering temperature increases, pore diameter decreases and mechanical properties increase as the packing of the spheres increases. Porosity can be increased by adding fillers such as sucrose to the powder and the wetting solution, which burnout on sintering.^[10] Komlev et al.^[11] produced porous HA scaffolds with interconnected interparticle pore diameters of approximately 100 μm, and a tensile strength of approximately 0.9 MPa, using 500 μm spheres. Mechanical properties can be increased further by hot isostatic pressing (HIPing), which will also decrease the pore diameter.^[12]

Combustibles

A foreign combustible organic material, such as an amino acid derivative, is added to a ceramic powder and burned away during sintering, leaving voids in the matrix.^[13,14] Total porosities of 60–90% can be attained with pores in the range of 100–250 μm ,^[14] but there tend to be large distances between the pores and, therefore, the interconnectivity of the pores is generally low.

SHS and Combustion Synthesis (Volatile Agents)

Porous ceramics can be made by self-propagating high-temperature synthesis (SHS), otherwise known as combustion synthesis. Cylindrical compacts of stoichiometric powder mixture and volatile agents are mixed and pressed at 30 MPa. The SHS reaction is ignited in a nitrogen atmosphere. Without the inclusion of volatile agents, such a process creates high-purity products, but porosity is low ($\sim 1 \mu\text{m}$ in diameter) and generally closed. However, volatile agents can be added to increase the total porosity and interconnectivity to produce ceramics with an open porosity of up to 83% and pores in excess of 100 μm .^[15] Disadvantages of this procedure are that the pore distribution is difficult to control and the technique is complicated with the risk of toxic residues being left in the scaffolds.

Freeze-Dry Process

In this environmentally friendly technique, a ceramic slurry is frozen while the growth direction of ice is controlled.^[16] The frozen slurry is then dried under reduced pressure to cause sublimation of the ice, creating pores macroscopically aligned in the direction of freezing. The green bodies were then sintered to provide strength. Although porosity is open and the pore structure can be controlled by the concentration of the slurry, freezing temperature, and the sintering parameters, the maximum pore size obtained was not much greater than 10 μm and mechanical strength was low. Such a process is therefore more applicable in the field of catalyst supports and filters than the field of biomaterials.

Polymer Foam Replication

Open-celled polyurethane foams can be immersed in slurries of HA or β -TCP under vacuum to allow the slurry to penetrate into the pores of the foam. The foams are then heated at 250°C to burn out the organic

components (pyrolysis) and sintered at 1350°C for three hours, producing a scaffold with 300 μm interconnected pore diameters.^[17] Glass frit can be incorporated into the slurry to reinforce the material.^[18]

In a similar process, porous ceramics can be produced in certain compositions, such as silicon oxycarbide, by the foaming and pyrolysis of preceramic polymers. Preceramic polymers, generally containing silicon, are polymers that undergo polymer-ceramic transition when heated to temperatures above 800°C (pyrolysis).^[19,20] However, there are no preceramic polymers for calcium phosphate ceramics.

Rapid Prototyping

Rapid prototyping is a generic term for a processing technique that produces materials in a shape determined by CAD (computer aided design) software on a computer. Such materials are usually built up layer-by-layer using a liquid phase or slurry of the material that cures or sets on contact with a substrate.

Rapid prototyping of HA scaffolds has been developed as a way of controlling the pore architecture of the scaffold.^[21] Stereo lithography (SL) apparatus is used to create an epoxy resin negative image of the virtual scaffold that is developed using the CAD software. A thermal-curable HA/acrylate suspension is poured into the resin mold and cured at 85°C. The scaffolds are then removed from the mold and pyrolyzed in a furnace and then sintered at 1350°C. The scaffolds produced are cylinders with diameter of 8 mm and height of 6 mm containing channels with diameters of 340 μm , a porosity of 35%, and a compressive strength of the scaffolds is approximately 30 MPa.

Direct Foaming

Highly porous ceramics can be produced by foaming particulate suspensions or colloidal sols^[22] to obtain scaffolds with pore diameters from 20 μm to 1–2 mm. Foaming techniques have proven very successful in producing microporous ceramics and glasses with potential for tissue engineering applications. The different methods of direct foaming will be discussed later.

SOL-GEL-DERIVED BIOACTIVE GLASSES

There are two methods to produce bioactive glasses 1) the traditional method of melting oxides at high temperatures in a furnace and casting into solid shapes or quenching into a liquid to form powders called frits; and 2) via the sol-gel process.

The sol-gel process involves the production of a colloidal solution (sol) of Si–O groups by the

hydrolysis of alcoxide precursors, such as tetraethyl orthosilicate (TEOS) in excess water under acidic catalysis. Simultaneous polycondensation of Si–O groups will continue after hydrolysis is complete, beginning the formation of the silicate network. As the network connectivity increases, viscosity increases and eventually a gel will form. The gel is then subjected to carefully controlled thermal processes of ageing (60°C) to strengthen the gel, drying (130°C) to remove the liquid by-product of the polycondensation reaction and thermal stabilization/sintering to remove organics from the surface of the gel (500–800°C). A chemically stable glass is then produced. An advantage of gel-derived bioactive glasses over the melt-derived is that they exhibit a mesoporous (pores with diameters in the range 2–50 nm) texture. This texture enhances bioactivity and resorbability of the glasses and creates anchorage sites for cells such as osteoblasts.^[23,24] Microporous sol-gel-derived bioactive glasses have been developed in several ways, each exhibiting a hierarchical pore structure; micropores for tissue ingrowth and textural porosity in the nanometer range for enhanced bioactivity, cell seeding, and controlled resorbability.

Sintering

Vallet-Regi and coworkers^[25] produced five gel-derived glasses, each with different silica content, by milling the gels after the drying stage, pressing them under isostatic conditions, and sintering at 700°C. The glasses exhibited micropores and mesopores that decreased in pore volume and pore diameter as silica content increased. The modal pore diameter, from mercury intrusion porosimetry, for a 58S (58 wt% SiO₂, 36 wt% CaO, 6 wt% P₂O₅) glass was 2.01 μm decreasing to 0.52 μm for an 80S (80 wt% SiO₂, 16 wt% CaO, 4 wt% P₂O₅) glass. Such small pore diameters are unsuitable for tissue ingrowth and vascularization.

Phase Separation

Phase separation associated with the hydrolysis and polycondensation of alkoxysilanes (e.g., TEOS), in the presence of water-soluble polymers such as poly(ethylene oxide), can be used to fabricate films or fibers with controlled microporosity and mesoporosity.^[26] The solution separates into gel (silicate-rich) and liquid (solvent-rich) areas before the sol-gel transition. After the transition, the solvent evaporates, leaving micropores, but the modal interconnected pore diameter of a very narrow pore size distribution (mercury porosimetry) was 1 μm, which again is too small for tissue repair applications, but ideal for high pressure liquid chromatography.

Freeze-Gelation

In a similar process to freeze drying, freeze-gelation has been used to control the textural porosity of sol-gel-derived materials and prevent shrinkage during thermal processing. Freezing of many silica sols initiates gelation, due to high concentrations of colloidal silica particles between ice interstices. Once gelation is complete, the ice crystals are melted, leaving pores.^[27] It is very difficult to generate large micropores with this method.

Templating

Many authors have cast silica-based sol-gel glasses around removable templates such as surfactants and latex or polystyrene spheres in processes similar to those described for HA and TCP ceramics.^[28–30] Once the sol has gelled and templates are removed, voids are left in the gel. Yan and coworkers^[30] used close-packed arrays of monodisperse spheres, such as polystyrene (PS) and poly(methyl methacrylate) (PMMA), as templates to produce ordered microporous (3DOM) gel-derived bioactive glasses. The close-packed beads were added to the sol after hydrolysis was completed. After soaking, the impregnated arrays were removed for gelation, ageing, and drying. The organics were burned out when thermal stabilization was carried out at 600–800°C. The pore network was very homogeneous and interconnected, due to the close packing of the spheres. However, the mean pore diameter was only 410 nm and mean interconnected pore diameter was approximately 100 nm. The pore size is limited by the difficulty in producing organic spheres with diameters of the order of 500 μm required to produce an ideal scaffold for tissue engineering applications.

FOAMING OF CERAMICS AND GLASSES

In this entry, foaming is defined as the incorporation of a gas (often air) into a ceramic to produce a porous material, with pores either interconnected or closed. The incorporation of bubbles is achieved either by injection of gases into the fluid medium, mechanical agitation, blowing agents, evaporation of compounds, or by evolution of gas by in-situ chemical reaction.^[31] A surfactant is generally used to stabilize any bubbles formed in the liquid phase by reducing the surface tension of the gas-liquid interfaces.^[32] Surfactants are macromolecules composed of two parts, one hydrophobic and one hydrophilic. Owing to this configuration, surfactants tend to adsorb onto gas-liquid interfaces with the hydrophobic part being expelled

from the solvent and a hydrophilic part remaining in contact with the liquid. This behavior lowers the surface tension of the gas–liquid interfaces, making the foam films thermodynamically stable, which would otherwise collapse in the absence of surfactant.

A permanent method of stabilization is required to solidify the foam and maintain the foam structure during solidification, providing strength. There are therefore three main stages in the process for foaming ceramics:

1. Foam generation from the ceramic suspension, with the aid of a surfactant to stabilize the bubbles in the liquid phase.
2. Foam solidification to provide permanent stabilization.
3. Sintering to provide strength.

In the time period between foam generation and solidification, several transformations occur in the bubble structure, such as bubble coalescence and bubble coarsening. During coarsening, large bubbles expand at the expense of smaller bubbles, which is caused by the diffusion of gas that is driven by a pressure gradient from smaller to larger bubbles.^[33]

Bubble coalescence is a consequence of the displacement of liquid around neighboring bubbles to the point that the films become so thin that they rupture. The thinning of the cell walls is caused by mechanisms such as:

1. Drainage due to gravity.
2. Drainage due to capillary action.
3. Local depression of surface tension due to hot spots.
4. Van der Waals attraction between surfaces of thin films.^[33]

Other processes also oppose cell wall thinning, such as:

1. The ability of films to repair differential concentrations of surfactant, which may result from disruptions (elasticity).
2. The relatively high viscosity of liquid in the bulk.
3. The electrostatic repulsion between the two sides of a very thin film.
4. Cohesion between the adsorbed surfactant chains around the film.

These mechanisms affect the final foam structure and properties such as permeability and strength. For foams to exhibit high pore interconnectivity, the films between cells must rupture to form an open cell structure.

The methods for the setting of fluid suspensions that have been commonly used are:

1. Drying or freeze drying the solvent.
2. Compositions that can form a gel intrinsically (sol-gel process).
3. Incorporation of gelling substances, such as cellulose derivatives or alginates.
4. The in situ polymerization of monomers (gel casting).

Different types of foaming methods, using combinations of the aforementioned processes will now be discussed.

Gel-Casting

Sepulveda et al.^[31,34] used the gel-casting method to produce microporous bioactive hydroxyapatite (HA) ceramics. Gel-casting employs the foaming of suspensions, which contain HA particles, water, and dispersing agents, and organic monomers (6 wt% acrylate/diene). The organic monomers must be water-soluble and retain a high reactivity. Once the slurry has foamed, in-situ polymerization of the monomers is initiated and cross-linking occurs, forming a 3-D polymeric network (gel), which produces strong green bodies. Foaming was achieved by agitation at 900 rpm with the addition of Tergitol TMN10 surfactant (poly(ethylene glycol trimethylnonyl ether)) under a nitrogen atmosphere. The polymerization process was initiated using ammonium persulphate and a catalyst (TEMED, N,N,N',N'-Tetramethyl ethylene diamine) before casting. Porous samples were then sintered to provide mechanical strength and to burn out the organic solvents. Foam volume (and hence, porosity) could be controlled by the surfactant concentration in the slurry, but only low foam volumes could be generated, even at high surfactant concentrations. This was thought to be due to the high viscosity of the HA slurry, but it could also have been due to the choice of surfactant used. The materials produced exhibited pores of maximum diameter of 100–200 μm , which is ideal for tissue engineering applications. Unfortunately, data on the interconnected pore size distribution were unavailable.

Interconnected pore size distribution data were available for gel-cast alumina foams.^[31] Foams with a final bulk density of 0.33 gcm^{-3} exhibited a wide interconnected pore size distribution with a mode at approximately 130 μm and showing pores present up to 1500 μm . The flexural strength of these foams, which were 12% of the theoretical density of alumina, was found to be approximately 3MPa. Higher density foams (30% theoretical density) exhibited flexural

strengths up to 24 MPa. Although alumina is bioinert, it is not bioactive or resorbable. Therefore, applications of alumina foams would be filters or weight-saving structures.

Theoretically, the gel-casting process could be applied to melt-derived bioactive glass powders. However, such glasses undergo surface reactions on contact with solutions to produce an HCA surface layer, and it is desirable to have control over the reaction before a scaffold is ready for clinical use.

Foaming Agents

Herein, foaming agents are defined as reagents that create gas bubbles on reaction in solution. Gun et al.^[35] added 6–10 vol% hydrogen peroxide to the base-catalyzed sol-gel process. Decomposition of the hydrogen peroxide forms bubbles, which remain throughout polymerization. Crack-free large monoliths (rods 20 cm × 0.5 cm) were prepared that did not shrink during aging, but there were few pores greater than 10 μm in diameter.

Wu et al.^[36] applied polymer foaming techniques to a sol-gel system and utilized viscosity control to stabilize bubbles created by freon (CCl₃F) droplets dispersed in a sol to create cellular silica with porosities of 55–90% and corresponding mean cell diameters of 30–1000 μm. Porous silica can be resorbable if the specific surface area is high enough. The viscosity control was obtained by adjustment of pH with H₂SO₄. The foam structure was stabilized by:

1. A rapid polymerization reaction, which caused a rapid viscosity increase in the sol during the sol-to-gel transition and stabilized the foam structure.
2. The addition of an anionic surfactant (sodium dodecyl sulphate, SDS) that lowered the bubble-sol interfacial tension.
3. The addition of a cosurfactant (1-Dodecanol) that stabilized the bubbles.
4. The addition of a cosolvent (methanol) that dispersed the freon in the aqueous sol.

Bubbles formed because freon has a boiling point of 23.8°C, and incubation was carried out at temperatures above this boiling temperature. This step was controlled to occur simultaneously with gelation. An increase in freon concentration and incubation temperature caused porosity and cell size interconnectivity to increase. The bending strength (three point bend) of the silica foams ranged from 2 MPa at a porosity of 86% to approximately 8 MPa at a porosity of 66%. Although pore diameters were high and pores were spherical, the pore size distribution appeared broad and interconnectivity appeared to be low from SEM observations.

Sol-Gel-Derived Bioactive Glass Foams—An Ideal Scaffold?

Sol-gel derived bioactive glasses have been directly foamed by vigorous agitation, with the aid of surfactants, to produce scaffolds that fulfill many of the criteria of the ideal scaffold.^[22] Fig. 2 shows a SEM micrograph of a typical foam of the 70S30C composition (70 mol% SiO₂, 30 mol% CaO). The scaffolds can be produced with interconnected micropores, with porosities in the range of 60–90% and a modal interconnected pore diameter of up to 150 μm^[37] with many interconnections in excess of 300 μm in diameter. Comparing Fig. 2 to Fig. 1 shows that the pore structure closely mimics that of trabecular bone.

The sol-gel-derived bioactive glass contributes high bioactivity, controlled resorbability, and the potential for the ionic dissolution products (Si and Ca) to stimulate the genes in bone cells to enhance bone regeneration.^[38]

Due to the nature of the sol-gel process, the scaffolds can be produced in many shapes, which are determined simply by the shape of the casting mold. The scaffolds can be produced from various compositions of gel-derived glasses. All foam compositions can be easily cut to a required shape. Fig. 3 shows foams produced in various shapes.

The only criterion not addressed is the matching of mechanical properties of the scaffolds to bone for in-situ bone regeneration applications. The compressive strength (~1 MPa)^[37] of the foams is much less than that of trabecular bone (~10 MPa). However, the mechanical properties of these foams should be sufficient for tissue engineering applications, where bone would be grown on a scaffold in the laboratory before implantation. Work on improving the mechanical properties is ongoing.

The surface of the gel-derived foams has also been modified with organic groups and proteins to create scaffolds that have the potential for lung tissue engineering.^[39,40]

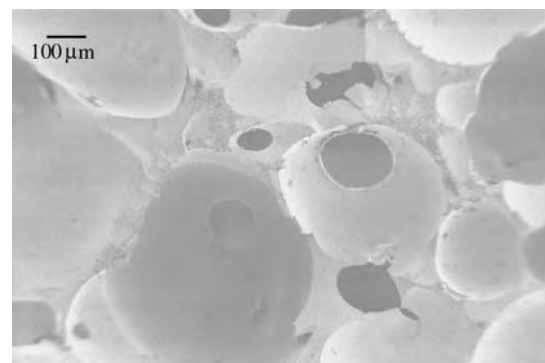


Fig. 2 SEM micrograph of a sol-gel-derived microporous foam of the 70S30C composition (70 mol% SiO₂, 30 mol% CaO).



Fig. 3 Typical sol-gel-derived microporous scaffolds made in various shapes. (Picture courtesy of Dr. Pilar Sepulveda.)

Fig. 4 shows a flow diagram depicting the foaming process. The scaffolds are produced by using a catalyst (hydrofluoric acid) to cause rapid gelation of the sol. As viscosity increases, the sol is foamed by vigorous agitation. Air is entrained into the sol, forming bubbles that are stabilized using a surfactant. As gelation occurs, the bubbles are permanently stabilized and become pores in the scaffold. The porous gel is then subjected to the standard thermal processes used in the sol-gel process.^[22] The properties of the pore network can be controlled at each stage of the process depicted in Fig. 4, e.g., by the type and concentration of catalyst and surfactant, the glass composition, and the process temperature.^[41]

POROUS MELT-DERIVED BIOACTIVE GLASSES

Melt-derived bioactive glasses of the 45S5 composition (Bioglass[®], 46.1 mol% SiO₂, 24.4 mol% Na₂O, 26.9 mol% CaO, and 2.6 mol% P₂O₅) have been used

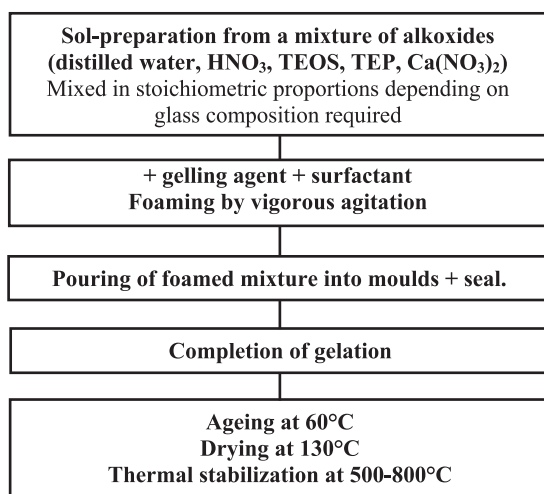


Fig. 4 Flow diagram of the sol-gel foaming process.

in the clinic in a powder form for over a decade.^[7] Therefore, producing porous Bioglass[®] scaffolds would provide a quicker route to clinical use than the sol-gel derived bioactive glass foams. However, porous melt-derived glasses have only recently been produced and it has proven difficult to produce materials with high porosity.

Livingston et al.^[42] produced a simple sintered scaffold by mixing 45S5 melt-derived bioactive glass (Bioglass[®]) powders, with particle size range of 38–75 μm, with 20.2 wt% camphor (C₁₀H₁₆O) particles, with particle size range of 210–350 μm. The mixture was dry-pressed at 350 MPa and heat treated at 640°C for 30 minutes. The camphor decomposed to leave porous Bioglass[®] blocks. Micropores were in the region of 200–300 μm in diameter, however, total porosity was just 21%, as there were large distances between pores, therefore interconnectivity was low.

Yuan et al.^[43] produced similar scaffolds by foaming Bioglass[®] 45S5 powder with a dilute H₂O₂ solution and sintered at 1000°C for two hours to produce a porous glass-ceramic. The pores were irregular in shape and relatively few in number, implying that interconnectivity was poor, but pore diameters were in the range 100–600 μm, ideal for tissue ingrowth. The pores appeared to be more like orientated channels running through the glass, rather than a pore network. The samples were implanted into the muscle tissue of dogs and were found to be osteoinductive, containing osteoblasts and osteocytes, in soft tissue. Pathological calcification (not containing osteocytes) was also observed in pores near the surface of the material. Crystal layers rich in Ca and O formed in pores away from the outer surface implant and were thought to be unrelated to the Ca-P rich layer that forms on the surface of bioactive glasses in vivo. Bone was formed directly on the solid surface and on the surface of crystal layers that formed in the inner pores and in pores containing pathological calcification. Osteogenic cells were observed to aggregate near the material surface and secrete bone matrix, which then calcified to form bone. However, although the implants had a porosity of about 30%, only 3% bone was formed. The authors suggested that any osteoconductive calcium phosphate-based biomaterial can be osteoinductive if it is made to the right geometry.

COMPOSITES

The obvious way to try to improve the mechanical properties of bioactive ceramics and glasses is to increase toughness by combining them with polymers, forming a composite. This is an extensive field of research.^[44,45] Microporous composites containing a



bioactive phase and a resorbable polymeric phase show promise for tissue engineering applications including soft tissue regeneration,^[46] however, a drawback to resorbable composites for in-situ applications is that the polymers do not resorb at the same rate as the ceramic or glass. The mechanical properties of the scaffolds therefore decrease rapidly with time and glass particles may migrate from the implantation site. Research to control the resorption rates of both phases is very active.

SUMMARY

Many techniques have been employed to produce microporous bioactive ceramics. Techniques such as the incorporation of organic phases that are eliminated during firing and the partial sintering of ceramics produce pore sizes up to 100 nm in diameter, which are not suitable for tissue engineering applications. The replication of polymer foams by impregnation of slurries produces pores in excess of 300 μm , however, mechanical properties were generally low. The most promising techniques for producing microporous ceramics and glasses with pore sizes suitable for tissue engineering seem to be direct foaming and gel-casting.

The possibility of an ideal microporous scaffold rests in the further development of present scaffolds by improving mechanical properties or by producing a microporous bioactive composite, where the resorbability of the polymer can be closely controlled.

ACKNOWLEDGMENTS

The authors wish to thank Lloyds Tercentenary Foundation and EPSRC (UK) for support.

ARTICLES OF FURTHER INTEREST

Bioactive Glass; Extracellular Matrix Scaffolds; Polymer Foams; Supercritical Fluid Processing

REFERENCES

- Davies, J.E. *Bone Engineering*; EM² Incorporated: Toronto, 2000.
- Langer, R.; Vacanti, J.P. Tissue engineering. *Science* **1993**, *260* (5110), 920–926.
- Jones, J.R.; Hench, L.L. Biomedical materials for the new millennium: A perspective on the future. *J. Mater. Sci. Technol.* **2001**, *17*, 891–900.
- Ohgushi, H.; Caplan, A.I. Stem cell technology and bioceramics: From cell to gene engineering. *J. Biomed. Mater. Res.* **1999**, *48*, 913–927.
- Lu, J.X.; Flautre, B.; Anselme, K.; Hardoïun, P.; Gallur, A.; Deschamps, M.; Thierry, B. Role of interconnections in porous bioceramics on bone recolonisation in vitro and in vivo. *J. Mater. Sci., Mater. Med.* **1999**, *10*, 111–120.
- Okii, N.; Nishimura, S.; Kurisu, K.; Takeshima, Y.; Uozumi, T. In vivo histological changes occurring in hydroxyapatite cranial reconstruction—Case report. *Neurol. Med.* **2001**, *41* (2), 100–104.
- Hench, L.L. Bioceramics: From concept to clinic. *J. Am. Ceram. Soc.* **1991**, *74* (7), 1487–1510.
- Gibson, I.R.; Bonfield, W. Novel synthesis and characterisation of an AB-type carbonate-substituted hydroxyapatite. *J. Biomed. Mater. Res.* **2002**, *59*, 697–708.
- Ota, Y.; Kasuga, T.; Abe, Y. Preparation and compressive strength behaviour of porous ceramics with β -Ca(PO₃)₂ fiber skeletons. *J. Am. Ceram. Soc.* **1997**, *80* (1), 225–231.
- Andrade, J.C.T.; Camilli, J.A.; Kawachi, E.Y.; Bertran, C.A. Behaviour of dense and porous hydroxyapatite implants and tissue response in rat femoral defects. *J. Biomed. Mater. Res.* **2002**, *62*, 30–36.
- Komlev, V.S.; Barimov, S.M. Porous hydroxyapatite ceramics of bi-modal pore size distribution. *J. Mater. Sci., Mater. Med.* **2002**, *13*, 295–299.
- Yin, S.; Uchida, S.; Fujishiro, Y.; Ohmori, M.; Sato, T. Preparation of porous ceria doped tetragonal zirconia ceramics by capsule free hot isostatic pressing. *Br. Ceram.* **1999**, *98* (1), 19–23.
- Fabbri, M.; Celotti, G.C.; Ravaglioli, A. Hydroxyapatite-based porous aggregates: Physical-chemical nature, structure, texture and architecture. *Biomaterials* **1995**, *16*, 225–228.
- Engin, N.O.; Tas, A.C. Manufacture of macroporous calcium hydroxyapatite bioceramics. *J. Eur. Ceram. Soc.* **1999**, *19*, 2569–2572.
- Zhou, W.; Hu, W.B.; Zhang, D. Combustion synthesis of highly porous ceramics: The TiC–Al₂O₃ system. *J. Mater. Sci.* **1999**, *34*, 4469–4473.
- Fukasawa, T.; Deng, Z.-Y.; Ando, M.; Ohji, T.; Goto, Y. Pore structure of porous ceramics synthesised from water-based slurry by freeze-dry process. *J. Mater. Sci.* **2001**, *36*, 2523–2527.
- Zhang, Y.; Zhang, M. Three-dimensional macroporous calcium phosphate bioceramics with nested chitosan sponges for load-bearing bone implants. *J. Biomed. Mater. Res.* **2002**, *61* (1), 1–8.
- Callcut, S.; Knowles, J.C. Correlation between structure and compressive strength in a reticulated glass-reinforced hydroxyapatite foam. *J. Mater. Sci., Mater. Med.* **2002**, *13*, 485–489.
- Columbo, P.; Griffoni, M.; Modesti, M. Ceramic foams from a preceramic polymer and polyurethanes: Preparation and morphological investigations. *J. Sol-Gel Sci.* **1998**, *13*, 195–199.
- Takahashi, T.; Munstedt, H.; Colombo, P.; Modesti, M. Thermal evolution of a silicon resin/polyurethane

- blend from preceramic to ceramic foam. *J. Mater. Sci.* **2001**, *36*, 1627–1639.
21. Chu, G.T.-M.; Orton, D.G.; Hollister, S.J.; Feinberg, S.E.; Halloran, J.W. Mechanical and in vivo performance of hydroxyapatite implants with controlled architectures. *Biomaterials* **2002**, *23*, 1283–1293.
 22. Sepulveda, P.; Jones, J.R.; Hench, L.L. Bioactive sol-gel foams for tissue repair. *J. Biomed. Mater. Res.* **2002**, *59* (2), 340–348.
 23. Hench, L.L.; West, J.K. The sol-gel process. *Chem. Rev.* **1990**, *90*, 33–72.
 24. Li, P.; Zhang, F. The electrochemistry of a glass-surface and its application to bioactive glass in solution. *J. Non-Cryst.* **1990**, *119*, 112–116.
 25. Balas, F.; Arcos, D.; Perez-Periente, J.; Vallet-Regi, M. Textural properties of $\text{SiO}_2\text{-CaO-P}_2\text{O}_5$ glasses prepared by the sol-gel method. *J. Mater. Res.* **2001**, *16* (5), 1345–1348.
 26. Nakanishi, K. Porous gels made by phase separation: Recent progress and future directions. *J. Sol-Gel Sci. Technol.* **2000**, *19*, 65–70.
 27. Statham, M.J.; Hammett, F.; Harris, B.; Cooke, R.G.; Jordan, R.M.; Roche, A.J. Net-shape manufacture of low-cost ceramic shapes by freeze-gelation. *J. Sol-Gel Sci.* **1998**, *12*, 171–175.
 28. Lebeau, B.; Fowler, C.E.; Mann, S.; Farcet, C.; Charleux, B.; Sanchez, C. Synthesis of hierarchically ordered dye-functionalised mesoporous silica with macroporous architecture by dual templating. *J. Mater. Chem.* **2000**, *10*, 2105–2108.
 29. Khramov, A.N.; Collinson, M.M. Sol-gel preparation of macroporous silica films by templating with polystyrene microspheres. *Chem. Commun.* **2001**, *8*, 767–768.
 30. Yan, H.; Zhang, K.; Blandford, C.F.; Francis, L.F.; Stein, A. In vitro hydroxycarbonate apatite mineralisation of CaO-SiO_2 sol-gel glasses with a three-dimensionally ordered macroporous structure. *Chem. Mater.* **2001**, *13*, 1374–1382.
 31. Sepulveda, P.; Binner, J.G.P. Processing of cellular ceramics by foaming and in situ polymerisation of organic monomers. *J. Eur. Ceram. Soc.* **1999**, *19*, 2059–2066.
 32. Rosen, M.J. *Surfactants and Interfacial Phenomena*, 2nd Ed.; John Wiley & Sons: New York, 1989; 277–303.
 33. Bikerman, J.J. *Foams*; Springer Verlag: New York, 1973.
 34. Sepulveda, P.; Binner, J.G.P.; Rogero, S.O.; Higa, O.Z.; Bressiani, J.C. Production of porous hydroxyapatite by the gel-casting of foams and cytotoxic evaluation. *J. Biomed. Mater. Res.* **2000**, *50*, 27–34.
 35. Gun, J.; Lev, O.; Regev, O.; Pevzner, S.; Kucernak, A. Sol-gel formation of reticular methyl-silicate materials by hydrogen peroxide decomposition. *J. Sol-Gel Sci.* **1998**, *13*, 189–193.
 36. Wu, M.; Fujii, T.; Messing, G.L. Synthesis of cellular inorganic materials by foaming sol-gels. *J. Non-Cryst.* **1990**, *121*, 407–412.
 37. Jones, J.R.; Hench, L.L. Effect of porosity on the mechanical properties of bioactive foam scaffolds. *Key Eng. Mater.* **2003**, *240–242*, 209–212.
 38. Xynos, I.D.; Edgar, A.J.; Buttery, L.D.K.; Hench, L.L.; Polak, J.M. Ionic products of bioactive glass dissolution increase proliferation of human osteoblasts and induce insulin-like growth factor II mRNA expression and protein synthesis. *Biochem. Biophys. Res. Commun.* **2000**, *276*, 461–465.
 39. Lenza, R.F.S.; Jones, J.R.; Vasconcelos, W.L.; Hench, L.L. Surface-modified 3D scaffolds for tissue engineering. *J. Mater. Sci., Mater. Med.* **2002**, *13*, 837–842.
 40. Tan, A.; Romanska, H.M.; Lenza, R.; Jones, J.; Hench, L.L.; Polak, J.M.; Bishop, A.E. The effect of 58S bioactive glass sol-gel derived foams on the growth of murine lung epithelial cells. *Key Eng. Mater.* **2003**, *240–242*, 209–212.
 41. Jones, J.R.; Hench, L.L. The effect of surfactant concentration and glass composition on the structure and properties of bioactive foam scaffolds. *J. Mater. Sci.* **2003**, *38*, 3783–3790.
 42. Livingston, T.; Ducheyne, P.; Garino, J. In vivo evaluation of a bioactive scaffold for bone tissue engineering. *J. Biomed. Mater. Res.* **2002**, *62*, 1–13.
 43. Yuan, H.; de Bruijn, J.D.; Zhang, X.; van Blitterswijk, C.A.; de Groot, K. Bone induction by porous glass ceramic made from Bioglass[®] (45S5). *J. Biomed. Mater. Res.* **2001**, *58* (3), 270–276.
 44. Bonfield, W.; Grynpas, M.D.; Tully, A.E.; Bowmann, J.; Abram, J. Biomechanics of bone. *Biomaterials* **1981**, *2* (3), 185–188.
 45. Roether, J.A.; Gough, J.E.; Boccacini, A.R.; Hench, L.L.; Maquet, V.; Jerome, R. Novel bioresorbable and bioactive composites based on bioactive glass and polylactide foams for bone tissue engineering. *J. Mater. Sci., Mater. Med.* **2002**, *13* (12), 1207–1214.
 46. Day, R.; Boccacini, A.R.; Roether, J.A.; Surey, S.; Forbes, A.; Hench, L.L.; Gabe, S. The effect of Bioglass[®] on epithelial cell and fibroblast proliferation and incorporation into a PGA matrix. *Gastroenterology* **2002**, *122* (4), T875. Suppl. 1.

Modeling, Biomedical

Andrew J. Pullan

Nicolas P. Smith

Department of Engineering Science, University of Auckland, Auckland, New Zealand

INTRODUCTION

Mathematical modeling is the process of combining quantitative data with a qualitative understanding to produce an explanatory and predictive tool. The level of advancement from qualitative to quantitative prediction via mathematical modeling is one measure that marks the maturity of a scientific discipline.

In contrast to the areas of physics and chemistry, where many major modeling-related advances were made in the 18th and 19th centuries, respectively, the medical and life sciences are only now undergoing this transition, albeit very rapidly. The advancement of modeling in the biomedical field has been delayed because of the inherent complexity in characterizing biological systems. This complexity manifests itself in two forms. The first is in the number of variables required to define the state of a system; the second is in the interactions between these variables that determine function.

Experimental techniques from the fields of molecular biology, structural biochemistry, physiology, and medical imaging are now able to provide enormous amounts of data relevant to the huge range of spatial-temporal scales needed to define the state of many biological systems. This large quantity of data means the link between measurement and function outside a specific scale is beyond rationalization using the human mind alone. However, recent increases in computational power are now providing the ability to rapidly process and database these large data sets.

Thus to take full advantage of these data sets and to apply basic science in a clinical or biological context, the development of formal mathematical frameworks is being accelerated. In the simplest sense the purpose is one of biological bookkeeping although through iteration with experiment it has the power to elucidate mechanisms, predict function, and provide information not accessible via measurement.

In this article we expand further on the place of mathematical modeling in the biomedical field, and where this field is heading. A brief history is given first, together with our view on the philosophy of modeling. We present some of the major difficulties associated with biomedical modeling, not the least of which is the very large range of time and space scales that must be

addressed in an integrated biophysically based model of an organ. Despite these difficulties, increasingly sophisticated whole-organ models are currently being developed. By way of illustration, we examine one such model. Finally, we provide a glimpse into the future of biomedical modeling when we describe the Physiome project and its grand ambitions to integrate all information from genomic level to clinical medicine using mathematical frameworks.

A BRIEF HISTORY

Only a brief glimpse of the history of biomedical modeling is given here to illustrate the rich and varied past. One of the first uses of mathematics in the biomedical field was made by William Harvey (1578–1657), a doctor of King Charles I of England. In 1628, Harvey published his famous *Anatomical Treatise on the Motion of the Heart and Blood in Animals*,^[1] in which he explained how blood was pumped from the heart around the body and then returned to the heart for recirculation. This was a very controversial idea that cost Harvey several patients—the common view at that time was that blood did not circulate, but sloshed back and forth, an idea that, by using simple arithmetic, was shown by Harvey to be absurd.

Galileo Galilei (1564–1642) was also publishing controversial ideas around the same time, although in a different field. He was ultimately imprisoned for his writings on astronomy, which were at odds with the Catholic Church at that time, and forbidden to publish anything more. However, his book *Discourses on Two New Sciences*^[2] was smuggled out of Italy and published in Leiden (in the Netherlands) in 1638. In this work Galileo presented a study of the relations between length, volume, weight, and strength and gave applications of this work to the shape and strength of animals. Giovanni Borelli (1608–1679) was also interested in this subject, and in 1680 presented a study of bones and muscles and the forces between them in his book *De Motu Animalium*.^[3]

Edmund Halley (1656–1742), most famously known for the comet that bears his name, published a set of mortality tables for the city of Breslau that were designed to aid in calculating annuities and pensions.^[4]

This was probably the first time data of this kind had been published and it was highly influential in the future production of actuarial tables for life insurance. Daniel Bernoulli (1700–1782) also spent some time studying the mortality of man. He is known to have analyzed the effectiveness of inoculation against smallpox, reaching the conclusion that the number of people who died as a result of the treatment was in fact smaller than the number who would otherwise have died. Bernoulli's greatest achievements, however, were in the field of hydrodynamics, and one of the most famous results in this field bears his name (Bernoulli's law). Together with Leonard Euler (1707–1783) (who is regarded as the greatest Swiss mathematician of all time), Bernoulli focussed significant attention on the relationship between the speed at which the blood flows and the pressure. His work ultimately led to a device to record blood pressure, albeit rather painfully, and soon physicians all over Europe were sticking point-ended glass tubes directly into patient's arteries.

Thomas Young (1773–1829) was an English physician and physicist who, through his studies on the behavior of light, was the initiator (with Herman Helmholtz (1821–1894) of the three-color theory of perception. The basis of this theory is that there are three different groups of cones in the retina, and each group is particularly sensitive to only one of the colors red, green, and blue. Only in recent years has conclusive evidence been obtained that shows that this theory is indeed accurate. Young also did work in elasticity, and a measure of the rigidity of materials, Young's modulus, is named after him. As well as the three-color theory of perception, Helmholtz made many other famous contributions in the biomedical field. He studied, among other things, the conservation of energy, the velocity of nerve impulses, and the heat output of muscles and through his work unified the fields of medicine, physiology, anatomy, and physics. Helmholtz was a superb exponent of combining mathematical modeling with experimentation and through his studies he brought about major advances in experimental measurement techniques and equipment.

Developments grew apace in the 20th century and we cannot hope to touch on them all here. However, there are at least three major works that must be commented on, even in this briefest of reviews. The first involves the work of Willem Einthoven (1860–1927). Einthoven measured the electrical activity of the human heart using external recording devices and produced a simple dipole model of the electrical activation of the human heart.^[5] This work, and his simple model, gave birth to the most widely used diagnostic technique in clinical cardiology, namely the electrocardiogram. Einthoven was duly recognized for such an important contribution with the Nobel Prize in Physiology/Medicine in 1924.

The second famous 20th-century use of biomedical modeling is the work of Alan Hodgkin (1914–1998) and Andrew Huxley (1917–) (for which they also won the Nobel Prize in Physiology/Medicine in 1963). Hodgkin and Huxley performed a series of studies and developed a mathematical model of a nerve cell based on the results of these experiments. This model was then used elegantly to test a number of fundamental hypotheses not directly measurable by experiment.^[6,7] This work still forms the basis of our current understanding of electrophysiology and ion regulation of cells in a multitude of tissue types today. Remarkably, Andrew Huxley was also a key figure in the parallel field of muscle mechanics, developing the sliding filament hypothesis. Huxley used mathematical modeling to predict the tissue response from a molecular model of how contractile proteins bind and unbind to generate tension.^[8] Only recently have sophisticated imaging techniques provided sufficient resolution to verify this model successfully and fully.^[9]

While the application of mathematics to study aspects of biomedical processes dates back more than three centuries, the complexity in the models has increased significantly in the latter half of the 20th century. This increase is largely due to parallel advances in the related fields of experimental biology and computer science.

In biology techniques such as DNA sequencing, protein structure measurement, and cell patch clamping, magnetic and nuclear magnetic resonance imaging have rapidly been developed and refined. These procedures provide rich data sets that, in combination, can provide the full gambit of information relevant to effective biological modeling. For instance, individual model parameters can be directly measured and related to whole system behavior, a process that is essential for model verification.

The second significant advance has been the increase in computational power. There is an oft-quoted mantra in the computer industry, encapsulated in Moore's law,^[10] which states that the computational power available for a given price will double every 18 months. This rate of advance shows no sign of abating in the near future and we now have readily available powerful computer systems that can be used to solve the increasingly sophisticated models. Before the computer revolution, models had to be solved by hand, which by necessity limited the amount and type of detail that could be included. Contemporary biomedical modeling is characterized by a constant interplay between the level of detail needed to characterize a system effectively (which is provided by experimental techniques) and the ability to process this information (which is provided by computational technologies). The increased information generated by the advances in experimental and clinical technologies

has demanded increased model sophistication as limitations of existing models have been revealed. However, without the power to solve such advanced models, the increased detail would have been little more than academic.

MODELING PHILOSOPHY

The goal or intended use of a mathematical model is an important influence not only on its level of detail but also its form. At the most basic level, a model for interpolating experimental data can be a phenomenologically fitted function whose form has no reference to the physical system. However, to predict function outside of a set of experimental conditions, it becomes important to include an anatomical and/or biophysical structure in the model. When such a model has been extensively verified, it then becomes possible to predict function powerfully, including, for instance, the effect of interventions such as medical treatments.

Central to the philosophy and process behind mathematical modeling in all fields is the distinction of key elements and parameters from those which can be simplified. To model everything that is known about a system is often both prohibitively complex and ultimately adds nothing to the understanding of the system. At the beginning of model development, insight is needed to decide what is important and what can be simplified. These decisions often depend not only on the system being studied but also the questions being asked of the model e.g., fast process variables that change on a slow time-scale can be assumed to be constant.

Biomedical modeling, particularly that involving whole-organ function, can be extremely complicated. It typically involves many subsystems that can span several orders of magnitude in the relevant spatial and temporal scales. Spatial scales can range from the 1-nm pore diameter of a membrane protein to the 1-m scale of the human body, a range of 9 orders of magnitude. Temporal scales range from 1 microsecond for Brownian motion to 70 years for a human life span, a range of 15 orders of magnitude. There are also typically multiple types of physical problem at the same spatial scale. In the heart, for example, myocardial contraction governed by the laws of mechanics is initiated by an electrical wave whose propagation is governed by conservation of electrical current, and supplied with energy via a blood stream also obeying the equations of mechanics but with entirely different constitutive laws. It is challenging in itself to understand these subsystems. To understand how these systems interact to produce the observed large-scale behavior is significantly more difficult. Due to the increasing sophistication in experimental techniques

we are continuously increasing our knowledge about the structure and function of the different subsystems. An obvious example of this is the genome project. From a sequenced gene the structure of the corresponding protein can be found, a prerequisite for understanding how it functions. This type of research has led to an explosion in the amount of information that is available on systems at the smallest scales.

Modeling provides a framework within which the building blocks coming from such research and experiments can be placed. This is a common approach in the traditional areas of engineering and physics where many researchers routinely design and analyze complex structures through computer models. The fact that this is a common approach in these areas is partly due to the solid base of information on material properties and interactions that is available for many traditional building materials. Biological materials and systems are far more complex than human-engineered systems and the complex interactions of these systems are correspondingly harder to predict. For this reason it is important to verify model results with all possible experimental data. Such verification determines the correctness of the model and its scope of use. Furthermore, a biophysically based model that does not fit experimental data points to a flaw in the underlying assumptions, which is often an interesting and relevant result in itself.

The complexity in biology often means systems are grouped in submodels that contain extra detail at a finer spatial or temporal scale than that of the main model, but interface to the main model using a few variables. An example of this technique is the spatial embedding of cell models in a tissue model. At the cellular level, a large number of chemical concentrations are tracked through time but output only one parameter, such as transmembrane potential or muscle tension, to the tissue model.

THE USE AND NEED OF MODELS

In all forms of modeling it is important to challenge the use or scope of a model and the need it satisfies or its ability to elucidate a given hypothesis. With the additional complexity that accompanies many biomedical models and the associated need for extensive development and computational resource, these questions have even more relevance and thus require consideration.

One argument that is raised against the mathematical models in the biological field is that they cannot be predictive, meaning that they are only capable of reproducing the experimental results that were used to create them. The sliding filament theory of Huxley mentioned previously shows just how wrong that

generalization can be. There are many other such counter-examples that show how computer-based models can allow insights into underlying mechanisms in a complex, integrated setting. For instance, in a recent paper,^[11] mutations of the cardiac sodium channel were investigated. A Markovian model of the channel with both normal (wild-type) and mutated channels was placed within a cardiac cellular model. This allowed the mechanism of early after depolarizations (a clinically important, but not well understood phenomenon) to be elucidated through the inactivation and subsequent reactivation of the L-type calcium channel. This would not have been possible using experiments on isolated sodium channels and illustrates the importance of modeling to gain an understanding of the overall picture.

The above example on the cardiac channels also reinforces the fact that behavior of the whole cannot be understood by simply understanding the behavior of each component. The process is highly nonlinear with very complicated interactions between each of the components that cannot be understood without modeling. Only by performing computer simulations is it possible to aggregate the pieces into a whole and see what consequences an altered component has on the overall system. An example of this is to investigate the effect of a drug designed to block the passage of chemically charged ions through a channel in the cellular membrane. The immediate effect is a reduction of the ion flow through that particular channel, but the resulting voltage action potential is not found by simply removing this current. The reason for this is that other currents are affected by the altered current and thus the overall result can be counter-intuitive.

These examples, and many others like it, provide strong evidence that models are already being used in a predictive manner and not just to reproduce experimental data. Bassingthwaight comments on the fact that even though our knowledge of these complex biological systems is very limited, computer-based models seem to be the only conceivable alternative to follow up on animal experimentation.^[12]

However, the issue still remains that the range of spatial and temporal scales is too large to provide an exact model of the physical situation. It is not practical to adopt a traditional top-down or bottom-up approach to these problems. Instead, the modeling may be undertaken using a middle-out approach, where the desired spatial and temporal scales are specified and information is gleaned from the scales that lie immediately below those of interest. While this approach may lead to some simplifications, groups of researchers at all levels may then benefit from the advances made at the next level down. Another approach could be to optimize the model for a particular medical problem based on knowledge of the important biological

determinants. The fact that under certain circumstances all models can be proved wrong can put modeling of this nature into a negative light. However, that does not mean one should not model, but that one should recognize that all models are to some degree approximations of reality, and one should be aware of the assumptions and approximations underlying any interpretation of the outcome of a model.

THE HEART AS AN EXAMPLE

To demonstrate many of the modeling techniques and philosophies outlined above, it is instructive to review an example. The methods demonstrated in the following example have also been successfully applied to model the gastrointestinal, pulmonary, and musculoskeletal systems. However, because of its relevance to mortality and mobility, modeling of the heart is arguably the most advanced example of integrative biomedical modeling that currently exists.

The heart is a variable-speed biochemical pump. Electrical impulses originating in specialized cells and caused by the flow of ions across the membrane of the cell give rise to a self-perpetuating wave of activation that spreads through the entire myocardium. This activating wave initiates and, via the nature of its spreading, coordinates the contraction of the cardiac cells, pumping blood to the body and to the heart itself. This simple description belies the complexity and many subprocesses that exist at a large number of scales within the beating of a heart. The modeling approaches associated with each spatial scale are described below along with an outline of how each scale is integrated into a whole-heart modeling framework that can be used to determine function in health and disease.

Molecular-Level Modeling

At the molecular level, protein chains interact to enable a cell to contract or regulate its internal or external ion concentrations. Actin and myosin are long protein chains that progressively bind and unbind in a ratcheting motion to contract a cell along their length. Their function, as discussed previously, was first elucidated by Andrew Huxley using a mathematical model of two coupled partial differential equations that account for the spatial strain and the rates of attachment and detachment.^[8] This approach has since been extended in a number of models to include a larger number of binding states between the two proteins and a related increase in the number of coupled equations.

The interaction of proteins also controls the flux of ions into or out of a cell by lining pores in the cellular



membrane. There are a large number of models that use a single or system of ordinary differential equations to quantify how the overall conductance of channel changes over time or is affected by the concentration of ions that may bind to sites on the channel. The structure and function of the pore proteins is determined by their encoded gene sequence. A mutated gene can affect ion motility across the cell membrane, affecting the activation and recovery of the cell, which, in the worst-case scenario, can lead to cardiac arrhythmia and death. Detailed imaging data is now becoming available that shows individual channel structure. In the near future we expect the emerging link between DNA sequencing data and channel protein structure to rapidly produce continuum-based mathematical models that link genomics to channel structure and channel function.

Cell-Level Modeling

Cellular models that describe the chemical reactions within the cell are used to determine functions such as excitability, tension generation, and metabolic energy production. Figure 1 shows a schematic diagram of the biochemical reactions encapsulated in a recently developed cell model.^[13] Each of these reactions is expressed as a series of coupled rate equations whose parameters include such properties as temperature, ion concentrations, and dimensions of the cell.

Well-known examples of such models include N98^[14] and LRH^[15,16] for the electrical activity, HMT^[17] for mechanics, and the Salem model^[18] for cardiac metabolism. Each of these models contains

three different components: 1) model parameters such as physical constants and fixed reaction rates or channel conductances, 2) state variables that define the state of the cellular system at any given point in time, for instance, ion concentrations or channel-gating states, and 3) derived variables such as an ion current through the membrane. Typically, functions of the derived variables and model parameters are used to construct expressions that determine the rate of change of each state variable with time. The system of equations composing the cellular model is then integrated through time to determine transient cellular function. One difficulty of coupling distinct cellular functions together are the very different time scales of each system. For example, at the cellular level excitation occurs on the scale of milliseconds, whereas metabolic processes can take many minutes to develop fully. This problem is often overcome using mathematical techniques that solve a model of a fast system using a fine time-scale and a model of a slow system using a much courser time-scale while still storing the values of state variable at all times. Recent coupled cellular models encompassing excitation, contraction, and metabolism contain up to 50 state variables, 100 parameters, and 80 derived variables.^[13]

Tissue-Level Modeling

Cardiac structure plays a pivotal role in the contraction of the heart—cells individually contract along their length, but this cell shortening in itself is not sufficient to generate the volume changes seen in a whole heart beat. Significant sliding of groups of cells

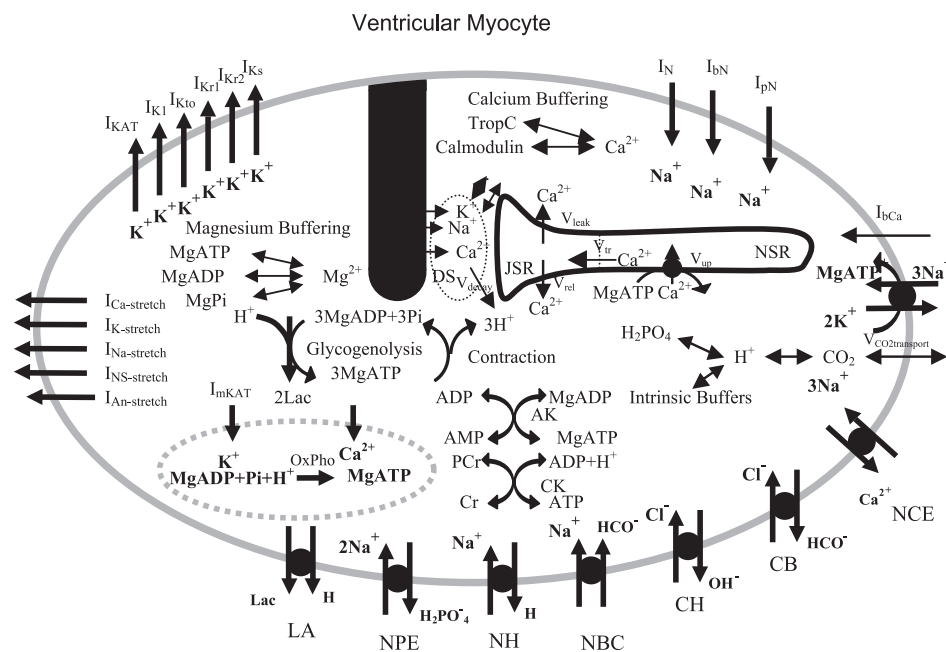


Fig. 1 Schematic illustration of some of the major processes involved with the electrical, contractile, and metabolic activity of the cell. (Based on the model from Ref. [13].)

over other groups also occurs, the directions of which are determined by the arrangement of the cells within the heart, and the distribution of collagen. Conduction speeds through tissue are also critically dependent on tissue structure, being typically tenfold faster in the fiber direction.

To model both health and disease, these effects of tissue structure are clearly beyond the scope of single-cell models. Thus, to determine tissue and ultimately whole-organ function, a modeling framework to which structural information can be added to cellular function is required.

The cell models outlined above are used to provide sources of current, mechanical tension, and oxygen tension to tissue-level models. Using a finite element mesh to spatially represent the tissue sample, a cell model can be embedded at a number of points whose locations within the mesh can be defined at whatever spatial resolution is required. Each point is treated as a black box, whose source/sink characteristics are determined by the complex ion kinetics of the underlying cellular model. The spread of current is then modeled by numerically solving the well-known bidomain equations underlying myocardial activation and extracellular current flow, which are formulated with a diffusion tensor based on the structural model. These equations are typically solved on a spatial scale much finer than that used to represent tissue geometry, meaning many cell model embedding points are defined within each finite element. Within the tissue model, conductivity and

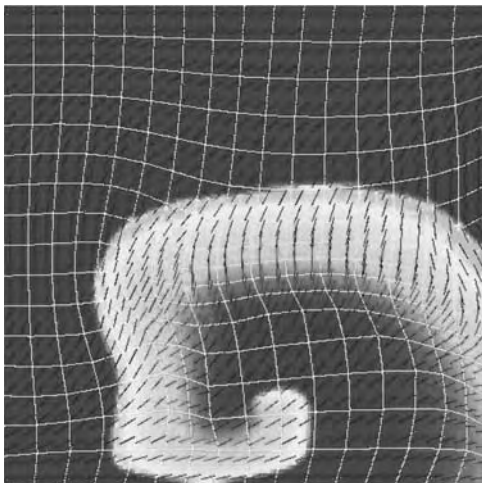


Fig. 2 A simulation on a two-dimensional block of cardiac tissue demonstrating the propagation of a spiral wave and the contraction of the tissue. The white lines represent the elements of the finite element mesh. In the resting state each element is square—distortions from this are the consequence of the contraction that is caused by the electrical wave moving through the tissue. The direction of the cardiac fibers is shown with the black lines.

capacitance determine the electrical coupling between approximation points and ultimately the spread of activation throughout the tissue. The anisotropic nature of cardiac tissue means this conductive coupling is represented by a tensor with values that are determined from the tissue microstructure.

In addition to the active tension generated by the embedded cellular models, the mechanical behavior of tissue models is dependent on the passive stress strain relationships. This passive relationship is quantified using a nonlinear constitutive law that determines the tissue stiffness locally about the axes defined by the tissue organization. The parameters of this constitutive law are estimated from biaxial tension, compression, and shear tests on small samples of ventricular tissue. Using these material properties, tissue deformation is determined using the finite element method to solve the governing equations of finite deformation elasticity. An illustration of the results of such a model is given in Fig. 2.

Organ-Level Modeling

The step up in scale from tissue to whole-organ modeling is initially one of adding geometry and characterizing the regional variation in structure. Careful dissections of hearts^[19] together with detailed imaging^[20] have gathered significant information on the unique organization of cardiac myocytes within the heart. These detailed measurements have been encapsulated within whole-organ continuum models using nonlinear optimization fitting techniques and piecewise parametric modeling. Figure 3 shows a finite

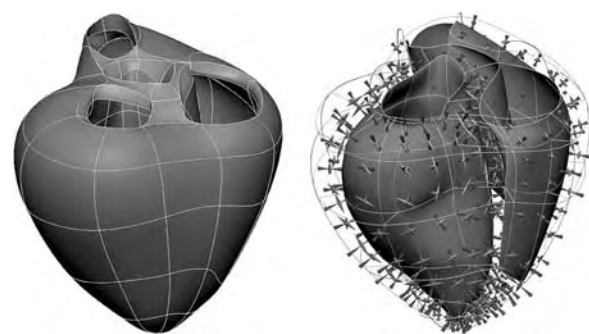


Fig. 3 A finite element mesh of the ventricles of a heart that has been constructed using detailed information from a pig heart (left). The regional variation in principle stresses and their direction at end-diastole at midwall points of the finite element mesh (right). Outward pointing cones represent tension while compression is represented by inwardly pointing cones. (Unpublished work from the Ph.D. thesis of Dr. Carey Stevens, Auckland Bioengineering Institute.)

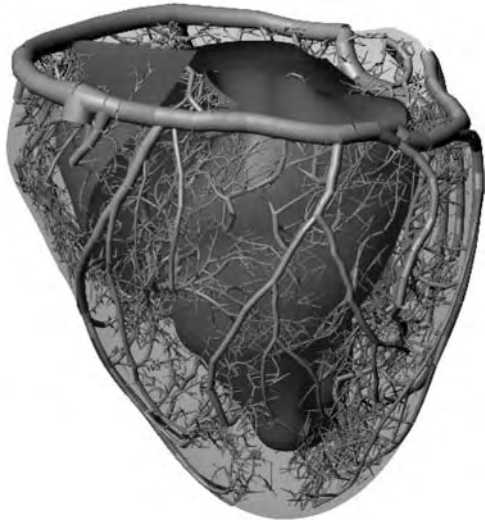


Fig. 4 An anatomically based model of the largest six generations of the coronary arterial network of blood vessels embedded in cardiac geometry. (Original figure generated using the techniques described in Ref. [21].)

element mesh of the ventricles of a heart that has been constructed by using detailed information from a pig heart.

Applying the tissue modeling techniques described previously to this anatomical model, the dynamics of heart excitation and contraction can be determined at the whole-organ level. Figure 3 (right) shows the directions and magnitudes of the principal stresses in the heart at the end of contraction, illustrating the ability of a model to calculate and display data that is extremely difficult to measure experimentally.

To understand heart function in health and disease more fully, a model of coronary blood flow needs to be added that can predict the regional distribution of blood flow through the heart wall. Figure 4 shows an anatomically based finite element model of the coronary blood vessels embedded in the ventricular geometry. The anatomy of the epicardial coronary vasculature is obtained from direct measurements. To represent smaller vessels that branch into the heart wall a nonlinear network-growth algorithm is used to add spatial information to detailed topological measurements made from silicon casts of the coronary vessel tree.^[22] A prediction of the regional and temporal variation in coronary blood flow is achieved by applying computational fluid dynamic methods to the domain of the anatomically generated model of the large coronary vessels.

Body-Level Modeling

Physically the cardiac electrical activity is connected to current flow in the torso via the extracellular potential



Fig. 5 The body surface potentials calculated at 45 ms after the onset of activation using an anatomically based torso model. Voltage is scaled between -1 and 1 mV and each band represents a 0.25 mV increment. (Original figure generated using the techniques described in Ref. [23].)

field. Cutaneous electrodes are routinely used in electrocardiology clinics all around the world to record such activity. Mathematical models of human torsos have been developed that can incorporate the information from the cell- and organ-level models described previously, and thus simulate electrocardiographic (ECG) signals. While the individual details of a single cell's electrical activity cannot be uniquely distinguished at the body surface, to simulate an ECG accurately still requires much of the detail of the cellular-level models described previously. An example of the torso potentials generated from one such model is illustrated in Fig. 5.

WHERE MODELING IS HEADING—THE PHYSIOME PROJECT

We are on the verge of a model-based revolution in the biomedical field. Much publicity has been afforded to the completion of the first draft of the human genome sequence.^[24] However, rather than signal the end of an event, this achievement, remarkable as it is, really signifies the beginning of something much bigger. The authors themselves allude to this when they state “In principle, the string of genetic bits holds long-sought secrets of human development, physiology, and medicine. In practice, our ability to transform such information into understanding remains woefully inadequate.” The challenge, therefore, is clear—to integrate this incredible wealth of information to allow

the determination of structure and function at all levels of biological organization.

This challenge has been recognized by the International Union of Physiological Sciences (IUPS), and confronted face-on by the establishment of what has been termed the “Physiome project.” This concept was presented in a report from the Commission on Bioengineering in Physiology to the IUPS at the 32nd World Congress in Glasgow in 1993. The name comes from “physio-” (life) and “-ome” (as a whole), and is intended to provide a quantitative description of physiological dynamics and functional behaviour of the intact organism. A satellite workshop “On designing the Physiome Project,” organized and chaired by the Chair of the IUPS Commission on Bioengineering in Physiology (Prof. James Bassingthwaite), was held in Petrodvoretz, Russia, following the 33rd World Congress in St. Petersburg in 1997. A satellite meeting on the Physiome project was held at the 34th World Congress of IUPS in Christchurch, New Zealand, in August 2001, and the Physiome Project was designated by the IUPS executive committee as a major focus for IUPS during the next decade.

The long-range goal of the Physiome project is to understand and describe the human organism, its physiology and pathophysiology, and to use this understanding to improve human health. A major aim is to develop mathematical and computer models to integrate the observations from many laboratories into quantitative, self-consistent and comprehensive descriptions. Integrated, biophysically based modeling is thus set for a revolution. Glimpses of this can already be seen. The pharmaceutical industry is beginning to appreciate the importance of mathematical modeling in reducing the massive cost of bringing a new drug to market: “Computer modelling will even provide the tools with which to perform in silico clinical trials, based on whole organ body models that test for everything.”^[25] Several companies have recently been formed to exploit biological modeling for drug discovery (Entelos, Pharsight, Physiome Sciences, Simulations Plus). The U.S. Food and Drug Administration (FDA) has indicated clearly that it would like submissions to include computer simulation and this in fact has been used by Roche to gain approval from the FDA for its cardiac calcium-channel blocker Posicor.^[26]

CONCLUSION

The use of modeling in the biomedical field lags behind that of many other fields. This is perhaps not surprising given the complexity of biological systems compared with many constructed by humanity.

However, the stage is set for a revolution in the biological sciences driven by the rapid development and application of biomedical modeling. This is being driven in part by the significant increases in computational power, enabling ever more complicated models to be constructed and solved.

Accelerated model development is also being driven by the massive amounts of data now being generated, courtesy of such technologies as gene sequencing and advances in medical imaging. One has very little choice but to use models to help cope with this explosion in complexity. This has been recognized by a number of companies and organizations, and has led to the development of the so-called Physiome project (<http://www.physiome.org/>). A major aim of this project is to develop mathematical and computer models to integrate the observations from many laboratories into quantitative, self-consistent, and comprehensive descriptions and to use this understanding to improve human health. Although very new, momentum is gathering, and one should not be surprised to find in the not too distant future the use of modeling in the biomedical field to be as pervasive as all other forms of experimentation.



ACKNOWLEDGMENTS

The authors kindly acknowledge the help Mr. Garry Tee, provided for the historical references. The work presented in Figs. 2–5 are the results of several past and present members of the Auckland Bioengineering Institute. In particular, the following people contributed to the work that led to these figures: Dr. Greg Sands and Dr. Martyn Nash (Fig. 2), Dr. Carey Stevens (Fig. 3) and Dr. Chris Bradley and Dr. Martin Buist (Fig. 5).

ARTICLES OF FURTHER INTEREST

Bone Remodeling; Finite Element Analysis; Spine Biomechanics

REFERENCES

1. Harvey, W. *The Works (1628)*; Sydenham Society: London, 1847. Translated by Willis R.
2. Galilei, G. *Discourses Concerning Two New Sciences (1638)*; McGraw-Hill: New York, 1963. (With an Introduction by A. Favaro), translated by Henry Crew and Alfonso de Salvio.
3. Borelli, G.A. *De Motu Animalium (1680)*; Springer-Verlag: Berlin, 1989. On the movement of animals (translated to English by Paul Maquet).

4. Halley, E. An estimate of the degrees of mortality of mankind, drawn from curious tables of the births and funerals at the city of Breslaw; With an attempt to ascertain the price of annuities on lives. *Philos. Trans. R. Soc. Lond.* **1693**, *17*, 596–610.
5. Einthoven, W. Die galvanometrische registrering des menschlichen elektrokardiogramms, zugleich eine beurtheilung der anwendung des capialr-elektrometers in der physiologie. *Pfluegers Arch.* **1903**, *99*, 472–480.
6. Hodgkin, A.L.; Huxley, A.F. Currents carried by sodium and potassium ions through the membrane of the giant axon of Logio. *J. Physiol.* **1952**, *116*, 448–472.
7. Hodgkin, A.L.; Huxley, A.F. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.* **1952**, *117*, 500–544.
8. Huxley, A.F. Muscle structure and theories of contraction. *Prog. Biophys.* **1957**, *7*, 255–318.
9. Craig, R.; Lehman, W. Crossbridge and tropomyosin positions observed in native, interacting thick and thin filaments. *J. Mol. Biol.* **2001**, *31* (311(5)), 1027–1036.
10. Moore, G.E. Cramming more components onto integrated circuits. *Electronics* **1965**, *38* (8), 114–117.
11. Rudy, Y. From Genome to Physiome: Integrative Models of Cardiac Excitation. *Ann. Biomed. Eng.* **2000**, *28*, 945–950.
12. Bassingthwaighte, J. *The Physiome Project: Why Now? and How?*, Proceedings of the Physiological Society of New Zealand, Aug 2001; Vol 20. Supplement 1 (Proceedings of the XXXIV International Congress of Physiological Sciences).
13. Mulquiney, P.J.; Smith, N.P.; Clark, K.; Hunter, P.J. Mathematical modelling of the ischaemic heart. *Non-linear Anal.* **2001**, *47*, 235–244.
14. Noble, D.; Varghese, A.; Kohl, P.; Noble, P. Improved guinea-pig ventricular cell model incorporating a diadic space, IKr and IKs, and length- and tension-dependent processes. *Can. J. Cardiol.* **1998**, *14* (1), 123–134.
15. Luo, C.; Rudy, Y. A dynamic model of the cardiac ventricular action potential. I. Simulations of ionic currents and concentration changes. *Circ. Res.* **1994**, *74* (6), 1071–1096.
16. Luo, C.; Rudy, Y. A dynamic model of the cardiac ventricular action potential. II. Afterdepolarizations, triggered activity, and potentiation. *Circ. Res.* **1994**, *74* (6), 1097–1113.
17. Hunter, P.J.; McCulloch, A.D.; ter Keurs, H.E. Modelling the mechanical properties of cardiac muscle. *Prog. Biophys. Mol. Biol.* **1998**, *69* (2–3), 289–331.
18. Salem, J.E.; Saei, G.M.; Stanley, W.C.; Cabrera, M.E. Mechanistic model of myocardial energy metabolism under normal and ischemic conditions. *Ann. Biomed. Eng.* **2002**, *30* (2), 202–216.
19. Legrice, I.J.; Hunter, P.J.; Smaill, B.H. Laminar structure of the heart: A mathematical model. *Am. J. Physiol.* **1997**, *272* (5 Pt 2), H2466–H2476.
20. Hooks, D.A.; Tomlinson, K.A.; Marsden, S.G.; LeGrice, I.J.; Smaill, B.H.; Pullan, A.J.; Hunter, P.J. Cardiac microstructure: Implications for electrical propagation and defibrillation in the heart. *Circ. Res.* **2002**, *91* (4), 331–338.
21. Smith, N.P.; Pullan, A.J.; Hunter, P.J. Generation of an anatomically based geometric coronary model. *Ann. Biomed. Eng.* **2000**, *28* (1), 14–25.
22. Kassab, G.S.; Rider, C.A.; Tang, N.J.; Fung, Y.C. Morphometry of pig coronary arterial trees. *Am. J. Physiol.* **1993**, *265* (1 Pt 2), H350–H365.
23. Bradley, C.P.; Pullan, A.J.; Hunter, P.J. Effects of material properties and geometry on electrocardiographic forward simulations. *Ann. Biomed. Eng.* **2000**, *28* (7), 721–741.
24. International Human Genome Mapping Consortium. A physical map of the human genome. *Nature* **2001**, *409* (6822), 934–941.
25. PricewaterhouseCoopers. An industrial revolution in R&D. Pharma **2005**, 1998. http://www.pwcglobal.com/gx/eng/about/ind/pharma/industrial_revolution.pdf.
26. Noble, D.; Colatsky, T.J. A return to rational drug discovery: Computer-based models of cells, organs and systems in drug target identification. *Emerg. Ther. Targets* **2000**, *4*, 39–49.

Nanoscale Neurosurgery

William C. Broaddus

Peter J. Haar

Department of Neurosurgery, Virginia Commonwealth University, Richmond, Virginia, U.S.A.

George T. Gillies

*Department of Neurosurgery, Virginia Commonwealth University, Richmond, and
Department of Biomedical Engineering, University of Virginia, Charlottesville, Virginia, U.S.A.*



INTRODUCTION

The fields of nanoscience and nanotechnology are creating new possibilities for investigating and dealing with difficult neurosurgical problems such as brain tumors, neurodegenerative disorders, and spinal cord injury. Glioblastoma multiforme (GBM), a common highly malignant primary tumor of the central nervous system (CNS), is characterized by a tumor mass surrounded by normal brain containing invading tumor cells. In this disease, many of the tumor cells invade the healthy regions of brain, but systemically delivered therapeutic agents are largely ineffective against them because of the screening effects of the blood-brain barrier (BBB). However, standard tools of nanotechnology, such as atomic force microscopy, are now being used to study the structures of individual GBM cells in ways that are beginning to shed light on how the invasion process works.

Moreover, the technique of positive pressure infusion is presently being deployed as a means of driving the flow of therapeutic agents directly through the nanoscale interstitial space between the cells within the CNS, thus providing a means of circumventing the screening effects of the BBB. The knowledge gained from laboratory studies of GBM cells and clinical studies of infusion therapies will carry over to other neurosurgical applications, too. For example, nanostructured materials that might be used as postresection surgical implants to deliver targeted therapies might be modified to serve as in situ scaffolds for guiding axonal growth cones in restorative neurosurgical procedures aimed at bridging regions of glial scar tissue in the treatment of spinal cord injury. In what follows, the scientific basis, clinical status, and potential applications for nanoscale neurosurgical modalities are described and discussed.

THE STRUCTURE OF THE BRAIN AT THE SUBMICRON SCALE

With roughly 1 trillion cells, 100 trillion synapses, and 1 million kilometers of neural pathways,^[1] the human

brain is often considered the most structurally complex object in the known universe. The brain's cellular and subcellular organizations have been studied for over a century and, far from being completely characterized, continue to amply furnish fresh subjects of inquiry and debate. Almost all knowledge of brain microstructure has been drawn from two-dimensional planar measurements; yet novel techniques, such as diffusion tensor imaging of the brain and computerized neuronal forest reconstruction, promise to greatly advance the micro- and nanostructural understanding of it. The brain is composed of many different cell types, among which are the most morphologically variable in the body.

Neurons are the CNS cells specialized for sending and receiving signals, and therefore they have highly extended cytoarchitectures to allow multiple intercellular contacts locally and at great distances. The nucleus is located in the cell body, also called the soma or perikaryon. Axons conduct signals away from neurons; most neurons have only a single axon, which can vary greatly in length in different neuron types. Dendrites are branching processes that receive impulses. Most CNS neurons are multipolar, meaning that they have at least several dendrites. Axons and dendrites are collectively termed neurites. Neurons are structurally diverse. Neuronal cell bodies measure anywhere from 5 to 135 μm across. Pyramidal cells in the parietal cortex in humans generally range from 8 to 36 μm (Ref. [2], as cited in Ref. [3]), but the giant pyramidal cells of Betz can reach 92 μm in length and 71 μm in width. The volume of human pyramidal cells ranges from 297 μm^3 to 3200 μm^3 depending on the area and layer of cortex (Ref. [4], as cited in Ref. [3]). The average volume of cortical neurons in man is 1600 μm^3 (the range is from 300 to 20,000). Neuron cell body densities have been thoroughly characterized in the different layers and areas of the adult human cortex. These densities range from 30 to 106 neuron cell bodies per 0.001 mm^3 (Ref. [4], as cited in Ref. [3]). In another study, cortex cell body densities were measured to be between 15 and 165 per 0.001 mm^3 .^[5] The human mean has been stated to be 48 neuronal cell bodies per 0.001 mm^3 . The neuronal

cell body cytoplasm is rich in structural elements such as microtubules, which are 25 nm in diameter; and neurofilaments, which are similar to intermediate filaments of other cell types and are 7.5 to 10 nm in thickness. In the neuronal perikaryon, microfilaments and neurotubules seem to occupy most of the space not preempted by the larger and more prominent organelles. In the neuronal cell body, these elements are of relatively short length and run parallel to one another in loose bundles. Microfilaments, which are 4 nm in diameter and composed of the contractile protein actin, are particularly abundant in the subplasmalemmal region of the cell.^[6]

Axons are specialized for conduction and transmission of neuronal output. The majority of CNS neurons have a single axon that can deliver signals to numerous other neurons. Axons are cylindrical, with a uniform diameter and a round or oval cross section. Within most cell types, axon diameter is proportional to axon length. Human axon diameters range from 100 nm to

20 μm .^[7] Collateral branches may be given off at right angles, and tend to be distal to the cell body. Axon branches form at approximately right angles and tend to have diameters similar to the parent stem. Terminal branches called telodendria end in synaptic terminals also known as boutons terminaux. Axons of the CNS may be quite long. For example, axons of certain pyramidal neurons course several centimeters in the subcortical white matter to furnish cortical-thalamic and cortical-cortical connections, often giving off recurrent collaterals or radial collateral branches that travel up to 8 mm. Because the lengths of these axons are many orders of magnitude greater than the micro- and nano-mechanical scales of length, it is more meaningful to consider the idea of axon length-density. Axon length-density indicates the total length of axons that are packed into a small volume of brain tissue. Using three different approaches, it has been estimated^[8] that the axonal length-density in the mammalian cortex is between 1.5 and 4.1 m per 0.001 mm^3 . Table 1 contains a

Table 1 Volume percentage, length-density, and Young's modulus measurements^a

Cortical tissue component	Percentage of total cortex volume (%)	Length-density (0.001 mm^{-3})	Effective Young's modulus (N m^{-2})
<i>Gray matter</i>			
Neuronal cell bodies	7.8	—	—
Axons	25.2	≈ 2.0 m	$0.7\text{--}4.0 \times 10^3$
Dendrites	25.9	~ 1.0 m	$0.7\text{--}4.0 \times 10^3$
Dendritic spines	10.4	—	—
Protoplasmic astrocytes	5.1	≈ 26 mm	$1.0\text{--}20 \times 10^3$
Oligodendrocytes	3.0	—	—
Extracellular space	17.0	—	—
Small arterioles	0.6	0.004 mm	1.0×10^5
Capillaries	3.2	1.00 mm	1.1×10^5
Small venules	1.8	0.012 mm	1.1×10^5
<i>White matter</i>			
Neuronal cell bodies	≈ 0	—	—
Axons (total)	31.9	338 mm	(see below)
Myelinated nerve fibers (5 μm)	76.8	39.2 mm	5.5×10^4
Small axons (0.3 μm)	2.1	299 mm	$0.7\text{--}4.0 \times 10^3$
Oligodendrocytes	46.4	39.2 mm	(see above)
Dendrites	≈ 0	≈ 0 mm	—
Fibrous astrocytes	1.8	43.8 mm	$1.0\text{--}20 \times 10^3$
Extracellular space	18.0	—	—
Small arterioles	0.2	0.002 mm	1.0×10^5
Capillaries	1.1	0.300 mm	1.1×10^5
Small venules	0.6	0.004 mm	1.1×10^5

^aDetails of the citations the many sources for the values reported in this table are contained in the document "Brain Tissue Structure and Micromechanics," available at the website <http://www.vcu.edu/mdphd/CurrentStudents/StudentPages/Haar/haar.html>.

list of a number of length-density characteristics for various cellular components of the CNS that have been compiled from the literature.

The cytoplasm of an axon, called the axoplasm, contains many structural support elements, including neurofilaments, microtubules, and actin microfilaments, among other sparsely scattered organelles, such as mitochondria and fragments of smooth endoplasmic reticulum. Microtubules are involved in rapid internal transport of protein and small molecules. Neurofilaments usually have side arms that limit their packing density. It has been shown that axon caliber is a function of neurofilament and microtubule numbers with neurofilaments predominating in large axons. Actin microfilaments are particularly abundant in presynaptic terminals and in the tips of growing axons, termed growth cones. It is believed that microfilaments play a key role in responding to external changes and in positioning the various receptors and ion channels at specific locations on the neuronal surface, by virtue of direct and/or indirect interactions with the intracellular portions of transmembrane components associated with the receptors and ion channels. Axons in the human CNS greater than $2\ \mu\text{m}$ are usually myelinated. Most cortical pyramidal cell axons are encased in a myelin sheath, which begins near the origin of the axon and ends short of its terminal branching. The myelin is laid down by oligodendocytes. The length of each myelin sheath, known as the internodal distance, is proportional to fiber diameter for each cell type and typically ranges from 150 to $1500\ \mu\text{m}$. The thickness of the myelin sheath and the internodal distance are directly proportional to the axon's diameter and length. The bare portions of axon between the internodes are termed nodes of Ranvier. The axon and its myelin sheath are collectively termed a nerve fiber. Dendrite architecture is organized to integrate synaptic responses from many other neurons. A typical CNS neuron receives about 1000 to 10,000 synaptic inputs, but some may receive an estimated 1 million inputs. Dendrites are continuous with the perikaryal cytoplasm and a transition point cannot be distinguished. Dendrites taper from the cell body and may branch immediately. Dendrites narrow along their length and therefore have irregular cross sections. Like axons, dendrites also have abundant structural support. The dendritic cytoskeleton appears less organized than that of axons. Microtubules dominate in large dendrites, and actin microfilaments are particularly abundant in dendritic gemmules. Microtubules are not well aligned with each other, and are less regular in their spacing than they are in axons.

Neuroglia—literally nerve glue cells—include all of the cells of the brain that are not neurons or cells of the vascular stroma. Neuroglia are not primarily involved with excitation, inhibition, and propagation of the nerve impulse. Neuroglia and neurons are in close contact

with each other and are highly interdependent metabolically. Although neuroglia function to support and maintain neuronal physiology, their overall functions are not completely understood in most cases. In the entire brain, the number of neuroglial cells has been estimated to be roughly 10 times greater than the total number of neurons; however, in the cortex, neuroglia outnumber neurons by only five times. On average, a $0.001\ \text{mm}^3$ volume of human brain contains 80 to 100 neuroglial cells (Ref. [9], as cited in Ref. [3]). In humans, the ratio between oligodendrocytes, astrocytes, and microgliaocytes in motor cortex was found to be 54:38:8.^[10] Astrocytes are the largest of the neuroglial cells and in the adult brain they perform many functions, both demonstrated and hypothesized. These functions include structural support, reaction to injury, neuronal insulation, and extracellular environment regulation. Astrocyte cell bodies have an average diameter of $40\text{--}50\ \mu\text{m}$, and astrocytes are believed to account for roughly 20–50% of the volume of most brain areas. Within the CNS, astrocytes are divided into two general categories. Protoplasmic astrocytes, also called velate astrocytes, are found in gray matter. They have numerous, thick, short, flattened, and highly branched processes which form delicate lamellae around terminal branches of axons, dendrites, and synapses. Fibrous astrocytes occur in the white matter. They have long processes with coarse bundles of gliofilaments. Each fibrous astrocyte has approximately 10–30 processes that are long, thin, smooth, and poorly branched. When grown in vitro with neurons, fibrous astrocytes exhibit fine complex processes hundreds of microns long. Both protoplasmic and fibrous astrocytes exhibit a high degree of morphological plasticity and can undergo marked structural changes.

Astrocytes project processes to neuron somas, ensheath synapses and dendrites, and form cuffs around nodes of Ranvier. Astrocytes are connected to each other by gap junctions, forming a syncytium that allows transfer of ions and small molecules. Astrocytes are also joined by desmosomes that create mechanical coupling. Astrocytes do not form the blood-brain barrier, but induce and maintain the tight junctions in endothelial cells that form the barrier. Astrocytic cytoplasm contains abundant gliofilaments, which are slightly finer than neurofilaments and are gathered into bundles. These intermediate filaments lack side arms, which allows for tighter packing. Gliofilaments are made of glial fibrillary acidic protein. Several observations seem to indicate that gliofilament bundles play a role in brain structural support, much like the connective tissue in other organs. Astrocytes often contain a large number of filaments arranged in parallel bundles, which are attached to the relatively rigid vascular tree and are held together by intercellular junctions. Protoplasmic astrocytes are more readily damaged than fibrous ones

because they contain a smaller number of fibrils. Following intense and prolonged mechanical stresses, astrocytes undergo an increase in size and show a marked increase in the amount of their fibrils. However, evidence indicates that the blood vessels, not the neuroglia, are the primary support structure of the brain.

Astrocytes have very close functional and structural relationships with neurons and the extracellular space (ECS). It has been demonstrated that neurotransmitters such as norepinephrine and glutamate can alternately stimulate and reverse glial stellation, respectively. Changes in astrocytic morphology would in turn affect the volume of the ECS and consequently neuronal activity. If accurate, this model suggests that the morphology of astrocytes and the volume of the ECS may be constantly changing under the influence of the activity of different sets of neurons. A quantitative analysis of the elasticity of astrocytes was performed using atomic force microscopy.^[11,12] The Young's modulus of the surface layer of cultured astrocytes ranged from 1 to $20 \times 10^3 \text{ N m}^{-2}$. Over the nucleus, the cell was softest, having a Young's modulus of only $2\text{--}3 \times 10^3 \text{ N m}^{-2}$; where the cell membrane was dense with F-actin, the cell was stiffest, having a Young's modulus of $10\text{--}20 \times 10^3 \text{ N m}^{-2}$. It is important to note that these studies measured elastic moduli of the astrocyte cell body, not the astrocyte processes. Given the fibrous structure of astrocyte processes, they are hypothesized to behave viscoelastically like the neuron processes.

Oligodendrocytes form the myelin sheaths that surround axons. Microglia mediate immune responses in nervous tissue by acting as macrophages. Resting microglial cells compose about 5% of the total neuroglial population. Ependyma is the simple cuboidal or columnar epithelium that lines the ventricular system. Ependymocytes comprise the majority of ependymal cells. Tanycytes are primarily found in the floor of the third ventricle. These cells have long basal processes that end on the pia mater and on the blood vessels in the median eminence of the hypothalamus. Tanycytes may respond to hormone levels in the CSF and help control the endocrine system. Choroidal epithelial cells cover the surfaces of the choroid plexus and have microvilli at their apical surfaces. They are characterized as having tight junctions. The vasculature allows a continuous blood perfusion to brain tissue, but a difference between brain endothelial cells and those of the systemic circulation is the presence of interendothelial tight junctions, also known as the zonula occludens. These tight junctions are responsible for the BBB that limits communication between the intravascular space and the brain ECS.

The complex geometries of blood vessels, neurons, and neuroglial cells are in close apposition to each other, and the ECS between them is thin, highly convoluted, and characterized geometrically by a parameter

referred to as the tortuosity. Spaces of only 10 to 20 nm are observed by electron microscopy between the cells and their processes.^[13] The ECS widens only at specific sites—for example, the nodes of Ranvier. As a result, estimations for the volume of this system of spaces have varied greatly. Early researchers reported that the ECS occupied 35% of the total volume in gray matter. Other researchers arrived at ECS percentages in gray matter of only 5–7%, and only 15% in white matter (Ref. [14], as cited in Ref. [3]). More recent determinations of the volume percentage of the ECS, using several different experimental methods, indicate that it occupies 17–20% of the brain.^[15,16] It is believed that standard fixation procedures greatly reduce measurements of this space. The ECS has a matrix of proteoglycans and glycosaminoglycans. The matrix composition varies with the cytoarchitectonics, but little is known about its density. Astrocytes are a major source of ECM proteins and adhesion molecules in the CNS; these proteins include nerve cell-nerve cell adhesion molecules, laminin, fibronectin, cytactin, and the J-1 family members janusin and tenascin.

DELIVERY OF THERAPEUTIC AGENTS THROUGH THE NANOSCALE STRUCTURE OF THE ECS

The complex process of glioma cell invasion takes place within the submicron scale milieu of cells, cellular processes, and vasculature described earlier. The fluid-filled ECS provides a natural conduit for the delivery of therapeutic agents to glioma cells via positive pressure infusion, also referred to as high flow-rate microinfusion and intraparenchymal or intracerebral clysis. In this technique, one or more catheters are surgically placed in the brain using appropriate methods of image-guidance, and the selected agent is then pumped through them; it subsequently infuses into the ECS. Pioneering studies of this technique and related transport mechanisms have been conducted by a number of workers—for example, Rech and Domino,^[17] Nicholson and Phillips,^[18] and Bobo et al.,^[19] who over the past several years have investigated the mechanisms by which macromolecules and other species are transported within the ECS. Various analyses of such data and parallel efforts to provide useful mathematical models of the fluid transport process have led to the elucidation of parameters such as the tortuosity factor, the volume fraction, and the porosity parameter that help to provide insight into the complex path for interstitial fluid movement. Experimental validation of the theoretical transport models has provided the foundation for the design and implementation of treatment planning systems for infusion therapy protocols.

To develop an understanding of the clinical basis for image-guided intraparenchymal drug delivery, one must first identify the biophysical mechanisms underlying the transport process. For simplicity of the current discussion, we consider the brain tissues to be isothermal and we ignore the slow physiological perfusion of the interstitial fluid that occurs continuously within the brain. Under these assumptions, there are then two principal mechanisms by which transport of externally introduced fluids can take place within the parenchymal tissues of the CNS; the first is flow along a concentration gradient (diffusive flow) and the second is flow along a pressure gradient (infusive or bulk flow). Drug delivery via the second mechanism is synonymously referred to as convection-enhanced delivery. The starting point for the modeling of the diffusion process is most typically Fick's laws of diffusion, modified to take into account the effects of the tortuosity and volume fraction of the tissues. Expressions can be developed that relate the flux to the concentration gradient. This then permits one to estimate the diffusion time constants and location-dependent concentration profiles, all as a function of the molecular weight of the diffusate.^[20] In contrast, the infusion process is most often analyzed in terms of Darcy's law for the flow through a poroelastic medium,^[21] which takes into account the swelling of the tissues due to the presence and spread of the initial bolus of infusate through the ECS, and the subsequent change in the permeability or hydraulic conductivity of the tissues due to the swelling. Such modeling of fluid behavior enables predictions of the volumes of distribution of the infusate as a function of the fluid's time-dependent pressure and velocity profiles. The pressure profiles can be measured by inline monitoring during the treatment protocol, and the time-dependent volume of distribution (and thus the direction-dependent velocity of the fluid through the ECS) can be mapped by magnetic resonance (MR) imaging techniques.

Because the flow of therapeutic or diagnostic agents through the ECS consists of both diffusive and infusive components, the determination of the transport ratio for these flow components requires an understanding of the dynamics of both processes. An important point to note is that during an infusion therapy session that lasts for a few hours, the transport distance traversed by a macromolecule through the nanoscale ECS due to diffusion will be significantly less than that due to the advective flow. Therefore, the volume of distribution that would be due only to diffusion is negligible over that period and we can assume that the advective component of the flow is what will be driving the spread of the infusate.^[22] This works to simplify the development of any associated treatment-planning algorithm that the clinician might wish to use, as there is only one physical process to model for the early stage part of the

protocol. The final concentration of the agent in the tissues will result from a balance between the source and sink terms in the model, which represent the influx of the agent into the region of tissue being treated, the vascular clearance, the rate and directionality of interstitial fluid flow, and the metabolic uptake or degradation of the agent. Techniques for MR-based confirmation of the volumes of distribution and of the concentrations are being developed by several workers—for example, Kucharczyk and Moseley.^[23] Such techniques will be critical to the ultimate clinical success of the method.

A useful consequence of the infusion of therapeutic agents directly into the nanoscale ECS arises because this delivery occurs to the brain side of the BBB; agents selected for their ability to penetrate the BBB would also tend to diffuse into the brain microvasculature after infusion and would thus be more readily cleared. Therefore, drug candidates that have been selected for delivery to the CNS by way of parenteral routes may actually be those that will be more rapidly cleared after direct intraparenchymal infusion. As a result, listing of possible therapeutic compounds in order of their lipid solubility would have to be inverted in this case to determine which candidates would be more likely to persist within the CNS parenchyma postinfusion. Broaddus et al.^[24,25] have discussed the status of clinical applications for positive pressure infusion.

Last, we note that a useful tool for optimizing the performance characteristics of neurocatheters designed for infusion therapies is positive pressure infusion into agarose gel surrogates of CNS tissues. In particular, 0.6% agarose gel has a nanoscale porosity that mimics that of the *in vivo* human ECS. This makes it a very attractive test bed for evaluating the volumes of distribution and the pressure profiles produced by catheters that are candidates for clinical applications. Figure 1 shows an infusion of a low-molecular-weight dye (bromphenol blue, MW = 693) into a Petri dish of 0.6% agarose gel. The gel was poured into a dish that contained glass fibers 8 μm in diameter in order to introduce a controlled anisotropy into this otherwise homogeneous and isotropic medium. The principal advantages offered by such an experimental arrangement are the simplicity, low cost, and robustness of the model, which make it possible to do critical preclinical testing on many types of neurocatheter designs before any level of *in vivo* testing becomes necessary. Gillies and colleagues^[26,27] have reported the results of various experimental studies of nanoscale flows in brain tissue surrogates of these kinds. In addition to the valuable model for brain anisotropy just described, the availability of inert size-delimited nanoparticles is expected to provide great utility in probing the behavior of the interstitial space of animal brain and *in vitro* model systems.^[25]



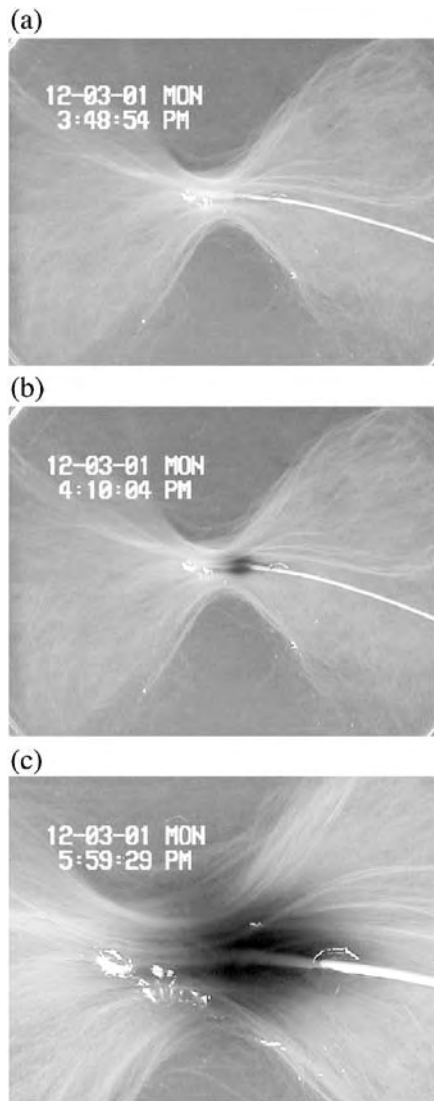


Fig. 1 A sequence of photographs showing the spread of a low-molecular-weight blue dye through the ≈ 200 nm interstitial pores of a 1-cm thick layer of 0.6% agarose gel into which 8- μ m diameter glass fibers have been added. The fibers introduce a simple type of controlled anisotropy that is meant to mimic the gross features of nerve fiber bundles and axons in the CNS.

NANOTECHNOLOGICAL TOOLS FOR INVESTIGATING TUMOR CELL INVASION AND AXONAL EXTENSION

Cell motility and movement is a well-developed area of study, and the invasion of tumor cells through the brain tissues is a subspecialty of that field that receives a considerable amount of attention. Standard methods of monitoring cell invasion mechanics involve the use of videomicroscopy, e.g., in experimental studies of tumor cell motility-based measures of malignancy. In

terms of monitoring invasion via cellular markers, a variety of different invasion assays are under development, several based on measurements of expression of matrix metalloproteinase (MMP). However, a new line of work now being undertaken in various laboratories centers on using atomic force microscopy (AFM) to join these other lines of effort. Scanning probe systems of this kind can be used to obtain images of glioma cell structure that are based on morphological maps of the cell surface obtained while the cells are plated on nanostructured substrates. A touch-point between this type of work and the invasion assays is that AFM can be used in principle to image directly whether or not different types and levels of MMPs do indeed influence the cell's motility; studies to explore this possibility are now being planned.

Recent AFM studies of glioma cells include those of Fillmore et al.,^[28,29] which focused on obtaining scanning probe images of the invadopodia extended by glioma cells into their surroundings. These slender cellular processes can have length-to-diameter ratios in excess of 100:1 and are involved in the sensing of and movement through the environment local to the cell; they play an important role in glioma invasion. For instance, AFM observations of T98 human glioma cells have revealed that some invadopodia produced by cells plated on collagen substrates can be approximately 100 microns long, with diameters of 600 to 700 nm. In one case,^[28] the invadopodia maintained a linearity in their path of extension of better than 1%, with interesting implications for the nanodynamics of the invadopodia growth process and cell-to-cell signaling mechanisms. In a related study^[29] observations were made on glioma cells that produced multiple invadopodia.

In terms of the axonal mechanics, many different methods are under study for the repair of paralysis-producing spinal cord injuries; Gillies et al.^[27] have published a succinct review of these methods. Central to these efforts is the need to regrow or otherwise bridge axonal bundles through or across the zone of damage that consists of traumatized and disrupted axonal processes, reactive glial and inflammatory cells, and developing glial scar tissue. Several strategies have been proposed for initiating such a process. These include the surgical placement of nerve grafts and/or scaffolding implants and a variety of other approaches in which axons from either existing or stem cell-grown neurons can then bridge the glial scar via chemotactic and cytomechanical action. Mechanical tension in particular has long been known to be an effective means of achieving axonal extension. Smith et al.^[30] have recently shown that axons under continuous mechanical tension can grow successfully when drawn at elongation rates of $\approx 12 \text{ nm s}^{-1}$ but will not work at a rate twice this fast. The question of how one would carry out such a procedure in situ in the patient is a key one, and several

strategies have been discussed. One interesting possibility proposed is the magnetic stretching of the axons, and various groups are exploring this approach. In principle, particles of magnetic material would be either attached to or taken up by neurons in vitro, and their axons would then be stretched by application of an external magnetic field. To gauge the magnitude of the magnetic force that would be needed, another standard tool of nanotechnology—the optical tweezer—is being proposed for use. In such an approach the cell body would be fixed to a substrate and a dielectric particle would be attached to the growth cone. The radiation pressure field of an appropriately shaped beam of laser light can then be used to capture the dielectric particle and place the axon under tension. Pre- and post-experiment calibrations allow one to determine the size of the force to be applied.

CONCLUSION

The development of scientific tools for use in measuring subnanonewton forces and for observing and controlling motions in the nanometer range of distances is making it possible to envision new clinical techniques for attacking difficult neurosurgical problems such as glioblastoma multiforme and spinal cord injury. These approaches constitute a movement toward nanoscale neurosurgery, in which new modalities for treating CNS disorders are expected to arise from an improved understanding of the nanomechanics of glioma cell invasion, extracellular matrix, CNS interstitial space, and neurite process extension. In this article we have presented a survey of the relevant neuroanatomy at the submicron scale of distances, described how positive pressure infusion makes use of the nanoscale interstices of the CNS to deliver therapeutic agents to the CNS parenchyma and invading tumor cells, and discussed how tools of modern nanotechnology are being put into use to develop improved understanding of the dispersal of primitive neural cell types and the processes extended by them. For further background information on these topics, please see the references to the cited literature.

ACKNOWLEDGMENTS

The authors thank H. L. Fillmore, I. Chasiotis, Z.-J. Chen, J. A. C. Humphrey, K. L. Holloway, G. E. Wnek, and several other colleagues for useful discussions and advice. We also thank University of Virginia undergraduate student Tim Wilhelm for the laboratory work leading to Fig. 1 and Prof. J. A. C. Humphrey for suggesting the use of glass wool filaments to introduce the controlled anisotropy in the gels shown in that

figure. We thank the Kopf Family Foundation, Inc., for financial support of the work at the University of Virginia. We thank the Hord and Cullather funds of the MCV Foundation for support of the work at Virginia Commonwealth University.

ARTICLES OF FURTHER INTEREST

Cardiac Catheters; Fibrin Sealants; Magnetic Nanoparticles: Structure and Bioapplications; Nerve Guides

REFERENCES

1. Swanson, L.W.; Lufkin, T.; Colman, D.R. Organization of the Nervous System. In *Fundamental Neuroscience*; Zigmond, M.J., Bloom, F.R., Landis, S.C., Roberts, J.L., Squire, L.R., Eds.; Academic Press: San Diego, 1999; 9–37.
2. Stankevich, I.A.; Shevchenko, Y.G. Variability of the structure of the cerebral cortex. *Trans. Inst. Mozg.* **1935**, *1*, 119–172.
3. Blinkov, S.M.; Glezer, I.I. *The Human Brain in Figures and Tables: A Quantitative Handbook*; Plenum Press: New York, 1968.
4. Glezer, I.I. Quantitative Characteristics of Certain Stages of Development of the Cortex of the Frontal Lobe in Human Postnatal Ontogenesis. In *Proceedings of the Third Conference on Age Physiology, Morphology and Biochemistry*; Nauka: Moscow, 1959; 387–394.
5. Scholte, W. Zur Gliarchitektonik der menschlichen Grosshirnrinde. *Arch. Psychiatr. Nervenkr.* **1959**, *199*, 573–595.
6. Peters, A.; Palay, S.L.; Webster, H.D. *The Fine Structure of the Nervous System: Neurons and Their Supporting Cells*, 3rd Ed.; Oxford University Press: Oxford, 1991.
7. Cricenti, A.; De Stasio, G.; Generosi, R.; Perfetti, P.; Ciotti, M.T.; Mercanti, D. Atomic force microscopy of neuron networks. *Scanning Microsc.* **1995**, *9*, 695–699.
8. Braitenburg, V.; Schüz, A. *Cortex: Statistics and Geometry of Neuronal Connectivity*, 2nd Ed.; Springer-Verlag: Berlin, 1998.
9. Blinkov, S.M.; Ivanitskii, G.R. The number of neuroglial cells in the human brain. *Biofizika* **1965**, *10*, 817–825.
10. Brownson, R.H. Perineuronal satellite cells in the motor cortex of aging brains. *J. Neuropathol. Exp. Neurol.* **1956**, *14*, 424–432.
11. Yamane, Y.; Shiga, H.; Haga, H.; Kawabata, K.; Abe, K.; Ito, E. Quantitative analyses of topography and elasticity of living and fixed astrocytes. *J. Electron Microsc.* **2000**, *49*, 463–471.
12. Yamane, Y.; Hatakeyama, D.; Tojim, T.; Kawabata, K.; Usiki, T.; Ogura, S.; Abe, K.; Ito, E. Fine surface images that reflect cytoskeletal structures in cultured glial cells by atomic force microscopy. *Jap. J. Appl. Phys.* **1998**, *37* (6B), 3849–3854.



13. Dempsey, E.F.; Luse, S. Fine Structure of the Neuropil in Relation to the Neuroglial Cells. In *Biology of Neuroglia*; Windle, W., Ed.; Charles C. Thomas: Springfield, IL, 1958; 99.
14. Horstmann, E. Was wissen wir über den intercellulären Raum in ZNS? *World Neurol.* **1962**, *3*, 112–117.
15. Nicholson, C.; Sykova, E. Extracellular space structure revealed by diffusion analysis. *Trends Neurosci.* **1998**, *21*, 207–215.
16. Nicholson, C. Diffusion and related transport mechanisms in brain tissue. *Rep. Prog. Phys.* **2001**, *64*, 815–854.
17. Rech, R.H.; Domino, E.F. Observations on injections of drugs into the brain substance. *Arch. Int. Pharmacodyn.* **1959**, *121*, 429–442.
18. Nicholson, C.; Phillips, J.M. Ion diffusion modified by tortuosity and volume fraction in the extracellular microenvironment of the rat cerebellum. *J. Physiol.* **1981**, *321*, 225–257.
19. Bobo, R.H.; Laske, D.W.; Akbasak, A.; Morrison, P.F.; Dedrick, R.L.; Oldfield, E.H. Convection-enhanced delivery of macromolecules in the brain. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 2076–2080.
20. Nicholson, C. Diffusion from an injected volume of a substance in brain tissue with arbitrary volume fraction and tortuosity. *Brain Res.* **1985**, *333*, 325–329.
21. Basser, P.J. Interstitial pressure, volume, and flow during infusion into brain tissue. *Microvasc. Res.* **1992**, *44*, 143–165.
22. Morrison, P.F.; Laske, D.W.; Bobo, R.H.; Oldfield, E.H.; Dedrick, R.L. High-flow microinfusion: Tissue penetration and pharmacodynamics. *Am. J. Physiol.* **1994**, *266*, R292–R305.
23. Kucharczyk, J.; Moseley, M.E. Method and Apparatus for Use with MR Imaging. U.S. Patent 6,061,587, May 9, 2000.
24. Broaddus, W.C.; Gillies, G.T.; Kucharczyk, J. Advances in image-guided delivery of drug and cell therapies into the central nervous system. *Neuroimaging Clin. N. Am.* **2001**, *11*, 727–735.
25. Broaddus, W.C.; Gillies, G.T.; Kucharczyk, J. Image-Guided Intraparenchymal Drug and Cell Delivery. In *Diagnostic and Therapeutic Imaging of the Nervous System*; Latchaw, R.E., Kucharczyk, J., Moseley, M.E., Yarborough, B.V., Eds.; Mosby: Philadelphia, 2003.
26. Gillies, G.T.; Allison, S.W.; Tissue, B.M. Positive pressure infusion of fluorescent nanoparticles as a probe of the structure of brain phantom gellatins. *Nanotechnology* **2002**, *13*, 484–486.
27. Gillies, G.T.; Wilhelm, T.D.; Humphrey, J.A.C.; Fillmore, H.L.; Holloway, K.L.; Broaddus, W.C. Spinal cord surrogate with nanoscale porosity for in vitro simulations of restorative neurosurgical techniques. *Nanotechnology* **2002**, *13*, 587–591.
28. Fillmore, H.L.; Chasiotis, I.; Cho, S.W.; Gillies, G.T. Atomic force microscopy observations of tumour cell invadopodia: Novel cellular nanomorphologies on collagen substrates. *Nanotechnology* **2003**, *14*, 73–76.
29. Chasiotis, I.; Fillmore, H.L.; Gillies, G.T. Atomic force microscopy measurement of cytostructural elements involved in the nanodynamics of tumour cell invasion. *Nanotechnology* **2003**, *14*, 557–561.
30. Smith, D.H.; Wolf, J.A.; Meaney, D.F. A new strategy to produce sustained growth of central nervous system axons: Continuous mechanical tension. *Tissue Eng.* **2001**, *7*, 131–139.

Nerve Guides

Burkhard Schlosshauer
Martin Lietz

*NMI Naturwissenschaftliches und Medizinisches Institut an der Universität Tübingen,
Reutlingen, Germany*

INTRODUCTION

The human nervous system has the potential to regenerate, although a number of biological restrictions are evident. Healing of lesions must occur via true regeneration, because simple scar formation (i.e., filling of tissue gaps by nonneuronal tissue) will reestablish electrical connectivity neither within the neuronal network nor with target cells, including muscle fibers. Neuronal regeneration is confined to regrowth of severed nerve fibers and proliferation of such accessory cells as myelinating glial cells. With the exception of quantitatively negligible stem cell proliferation, no new neurons can be generated in the adult nervous system. In the broadest sense, the aims of neural tissue engineering are: 1) to restore aspects of the three-dimensional architecture using biological or synthetic scaffolds (e.g., nerve guides in lesion sites); 2) to supplement biological functions, as by the implantation of neurotransmitter-producing cells (e.g., dopamine-producing cells in Parkinson's disease); or 3) to counteract defective metabolism, inflammation, or cell death using genetic engineering with viral vectors (e.g., apoptosis in amyotrophic lateral sclerosis).

This review provides a survey of nerve guides used as implants to bridge gaps of lesioned nerves. Through observations of biological phenomena during de- and regeneration, the applications for and potential uses and limitations of nerve-guide concepts have been elucidated. Different nerve guides have been developed—involving biological, synthetic, inert, resorbable, purely material-based, and biohybrid approaches—and have proven to be promising methods to enhance regeneration after implantation. Nevertheless, further tissue-engineering improvements are needed to fully exploit the complete regenerative capacity of the nervous system. In the future, neural tissue engineering will greatly profit from more detailed understanding of cellular mechanisms, allowing for more rational designs of implants.

PERIPHERAL NERVOUS SYSTEM (PNS)

If lesions have compromised the integrity of neural tissue, the damage may be repaired using nerve guide implants. Understanding the functional architecture of

nerves will clearly facilitate prosthesis development. In most peripheral nerves, several hundreds or thousands of individual nerve fibers (nonmyelinated axons and axons myelinated by Schwann cells) are grouped into units called fascicles (Fig. 1). An individual nerve contains a limited number of fascicles, with each fascicle surrounded by a perineurium composed of flattened cells and basement membranes, which together form a mechanically stabilizing envelope. All fascicles of a nerve are embedded in loose connective tissue termed epineurium. The intrafascicular connective tissue is called endoneurium. Nutrients are supplied by a microvascular system organized in a segmental fashion. Extrinsic vessels often approach the nerve trunk in a loosely coiled manner, then typically divide into branches running longitudinally in the epineurium (Fig. 2). Collaterals are formed in the perineurial sheet, continuing mainly as capillaries in the endoneurium. Endothelial cells of these endoneurial capillaries display specific barrier and carrier functions, collectively summarized as a blood-nerve barrier that is not evident in perineurial vessels. Vascular plexuses are found at all nerve layers. Because no lymph vessels are present in the endoneurial space,^[4] following nerve trauma, edema may form in the epineurium, but will not penetrate into the fascicle due to the protective ensheathment of the perineurium.

The key components of peripheral nerves are axons, which are long cell processes of neurons situated outside the nerve in functional relay stations, as in the spinal cord or in peripheral ganglia. The axons propagate neuronal information from relay stations to peripheral targets, including muscle fibers and endocrine glands, or convey information in the opposite direction (e.g., from the skin to more central relay stations). To ensure proper function of the electrically conductive axons of the PNS, axons must be ensheathed by Schwann cells, be embedded in a connective tissue scaffold, and receive nutrients via a vascular network as outlined in the foregoing. Schwann cells provide some of the most vital structural and biochemical support for PNS axons. The great importance of these support cells and nonneuronal tissue elements must be appreciated, especially in light of the fact that axons



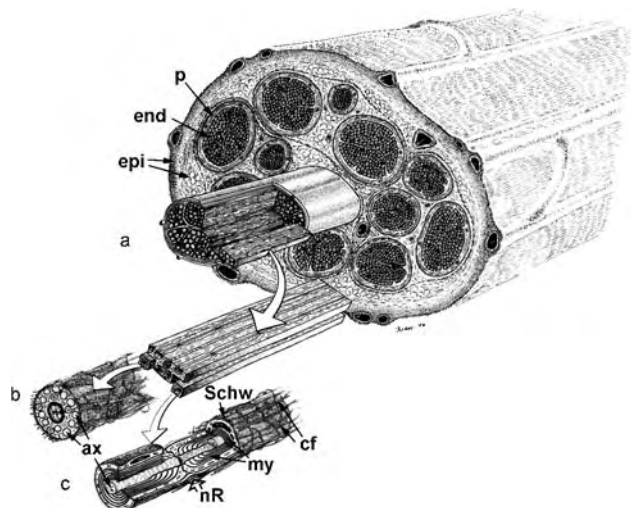


Fig. 1 Microanatomy of the peripheral nerve trunk. a, Bundles of nerve fibers are surrounded by the perineurium (p), thereby creating fascicles that are embedded in loose connective tissue—the epineurium (epi). The intrafascicular connective tissue is called the endoneurium (end). The appearance of nonmyelinated and myelinated nerve fibers is illustrated in b and c, respectively. Schw—Schwann cell; my—myelin sheat; ax—axon; nR—node of Ranvier. (From Lundborg, G.: *Nerve Injury and Repair*; used with permission of Churchill and Livingstone, Edinburgh, 1988.)

represent cell extensions of enormous length. These lengths of one meter or more mean that the distance between the distal axon tip and its nurturing neuronal cell body can be 10,000 times the diameter of the neuron itself.

During movements, nerve trunks are protected by a conjunctive adventitia on their exterior surface, which enables gliding of the nerve trunk. The mobility of fascicles in deeper nerve layers is ensured because the fascicles can slide against each other. The plexus brachialis shows a total excursion of about 50 mm during abduction-adduction of the shoulder. During elbow flexion-extension, median and ulnar nerves glide some 8 mm proximal to the elbow and some 14 mm near the carpal tunnel. If these gliding mechanisms are hampered by extra- or intraneuronal fibrosis after nerve lesions, chronic nerve irritation or loss of function may occur. Partial or complete compression and transient or permanent disruption of one or more of these neuronal or nonneuronal structures may severely affect the function of nerve conduction.

BIOLOGICAL ASPECTS OF DEGENERATION AND REGENERATION

According to the nomenclature of Seddon, nerve lesions are classified into three categories: 1) neurapraxia;

2) axonotmesis; and 3) neurotmesis.^[5] Neurapraxia describes traumata originating from severe compression without significant structural damage to the nerve. Transient motor paralysis is a typical consequence, whereas less vulnerable sensory and sympathetic functions of unmyelinated axons remain largely unaffected. Spontaneous recovery without neurosurgical intervention is observed within 3 months. Axonotmesis describes a severe crush or traction injury. Axons within otherwise preserved endoneurial tubes are ruptured. Due to the fact that truncated axons can regenerate within endoneurial tubes, no surgery is performed. Neurotmesis describes lesions with loss of continuity of some or all connective tissue components (endo-, peri-, or epineurium). If the internal organization of the nerve is disrupted or the nerve is completely transected, surgical repair becomes obligatory. Of the three types of nerve lesions, only neurotmesis injuries are therapeutically addressed using nerve guide implants.

Because the adult nervous system is essentially unable to generate new neurons, neuronal regeneration is restricted to the regrowth of new cell processes (axons and dendrites). If nerve lesions do not occur too close to the neuronal cell bodies, the neurons survive, pass through a series of degenerative states, and initiate regeneration in concert with support cells of the severed nerve. After injury, neurons swell, decrease metabolism associated with neurotransmission, and increase the production of cytoskeleton and membrane components necessary for axonal regrowth. During an initial response, the tip of the proximal nerve stump still in continuity with the neural cell body retracts for several millimeters. After several days, axons within the proximal nerve stump form growth cones that

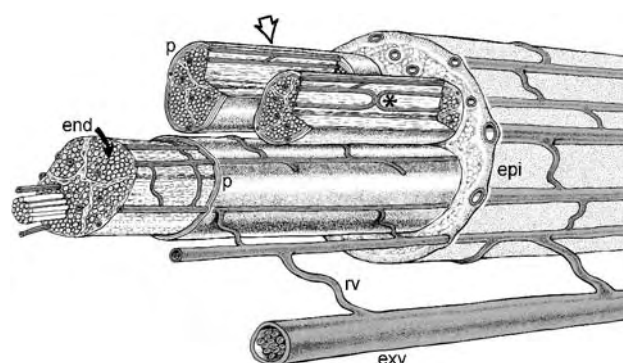


Fig. 2 Microvascular system of the nerve trunk. Extrinsic vessels (exv) support the regional feeding vessels (rv). The vascular plexuses are located in the epineurium (epi), perineurium (p), and endoneurium (end). The vessels pierce through the perineurium obliquely (open arrowhead). Note the intrafascicular double-loop formation (*). (From Lundborg, G.: *Nerve Injury and Repair*; used with permission of Churchill and Livingstone, Edinburgh, 1988.)

represent the leading tip of regrowing axons. The highly motile growth cone is the sensimotoric unit that directs axonal regrowth by exploring the microenvironment. The growth cone must adhere to permissive substrata such as laminin-containing basal lamina via integrin receptors, or will retract after encountering repulsive extracellular matrix (ECM) components or membrane constituents, including proteoglycans or myelin-associated proteins.^[6] Growth cones extend numerous exploratory filopodia until the first of these contacts a permissive terrain. Subsequently, the other filopodia die back in a pruning process. The leading filopodium becomes stabilized and continuously elongates via insertion of membrane components at the growth cone base, after which new filopodia again begin to explore the microenvironment. Any neuroprosthetic concept must take into account these fundamental growth characteristics, which are based on selective and transient cell and ECM adhesion.

Wallerian degeneration begins in the distal nerve stump within hours, characterized by complete deterioration of the distal axon segments that have been separated from their cell bodies. Within the next few days, the myelin sheets formed by Schwann cells around axons become degraded. Macrophages invade affected nerve segments and remove cell debris via phagocytosis. Resident and migrating cells release growth- and differentiation factors in addition to many types of interleukins, which act in concert to regulate the degeneration and subsequent regeneration processes. During this period the basal lamina components remain surprisingly resistant to degradation, which is important because its ECM components provide a directional scaffold for proliferating Schwann cells and axonal regrowth. Additionally, in the absence of Schwann cells, the ECM alone can function as a directional cue for axonal extension, even weeks after the nerve lesions were induced. Proliferating Schwann cells organize themselves in longitudinally oriented bands of Büngner. If axons can successfully navigate across the lesion site between the proximal and distal nerve stump, they are likely to encounter these Schwann cell columns, which then provide a perfect environment for axonal elongation. Several nerve-guide concepts attempt to imitate the positive influence of permissive matrix molecules and Schwann cells (discussed subsequently). In humans, the axonal elongation rate is on average 1 mm per day until the axons have traversed the length of the distal nerve and contacted synaptic target cells. Misrouting during muscular and sensory reinnervation is a rather frequently observed phenomenon, occurring mainly at the site of injury. Attempts to reduce mismatched connections have focussed on reconnecting individual fascicles rather than nerves as a whole. However, due in part to the

nonlinear architecture along the longitudinal axis of nerves, it is notoriously difficult to correctly rejoin proximal and distal fascicle stumps. Axonal subtypes (motor, sensory, sympathetic) display synaptic specificity, which may result in the elimination of incorrectly rewired axons. Despite this tissue specificity, the specificity for distinct nerve fascicles has remained controversial.^[7] Due to the functional plasticity of the adult nervous system, relearning based on rewiring of central neural networks is possible. This rewiring allows for some adaptation to changes in the microanatomy after aberrant synaptogenesis within distinct populations (e.g., motoneurons). Note that loss of innervation leads to progressive atrophy of such synaptic targets as muscles. Therefore, a primary clinical goal is the acceleration of regeneration to prevent muscle wasting and loss of sensory units. In addition, complete failure to reestablish any connectivity may lead to the formation of a neuroma at the lesion site. A neuroma is a pathological aggregate with dense, irregular tangles of axons devoid of peripheral synaptic targets. Because neuromas often cause painful sensations, nerve-guide implants may be used to direct axons to (even inappropriate) synaptic targets, as to neighboring muscles. A major cause of unsuccessful regeneration is fibrosis, a connective tissue scar formed by infiltrating fibroblasts. These secrete collagen type IV, which, although a permissive substrate, may function additionally as a molecular scaffold for inhibitory extracellular matrix components (including chondroitin sulfate proteoglycan).^[8] In addition to neuroma therapy, nerve guides are indicated where nerve gaps cannot be bridged by natural means.

NERVE SUTURES AND TRANSPLANTATIONS

In humans, essentially no regeneration is observed across severed nerves with a gap or extensive scars between both nerve ends. Therefore, therapeutic approaches always aim to reestablish some sort of continuous growth line. If the ends of severed nerves can be joined without tension, an end-to-end repair is possible. Fine monofilament nylon sutures are employed, passing through either the epineural connective tissue (perineural repair) or the perineural tissue (fascicular repair) to realign nerve ends or fascicles, respectively. Fibrin glue—based on the polymerization of fibrinogen after enzymatic activation by thrombin—is another means by which truncated nerve ends can be stably interconnected.

In instances where true nerve gaps are present, some sort of bridging becomes essential. Currently, autologous nerve graft transplants continue to represent the gold standard. Sensory nerves are employed in most cases, the removal of which result in only minor

functional deficits. Most frequently, sural nerves from the right and left lower leg are transplanted. Obvious drawbacks occur at the secondary surgical site, with concomitant risks of infection and inflammation, morbidity at the donor site, and potential formation of painful neuromas. To circumvent these problems, heterologous transplants have been employed in animal models. Additionally, decalcified bone channels or inside-out veins and arteries have been used, which expose the collagen-rich adventitia to regrowing axons. In the case of heterologous grafts, transplantation has been followed by immunosuppression to prevent graft rejection. To remove the most highly immunogenic components, alternating freezing-and-thawing and detergent extraction have been employed.^[9-11] Common to all these approaches is preservation of the fairly durable basal lamina, which provides a template for axonal regrowth. Evaluation of these acellular grafts indicates that regeneration is delayed compared to regeneration resulting from use of fresh grafts *in vivo*. This indicates the importance of such functioning cells as Schwann cells. As outlined subsequently, integration of Schwann cells into nerve-guide conduits improves regeneration substantially. To date, none of the heterologous approaches has fulfilled its promise. These approaches consequently are not in clinical use.

NERVE-GUIDE PRINCIPLES

Nerve guides are compact or porous scaffolds or tubular conduits of natural or synthetic polymers destined to replace biological transplants in order to bridge the gaps between injured nerve stumps. Besides rerouting regenerating axons to synaptic targets, nerve guides may serve to enhance angiogenesis, prevent infiltration of fibrous tissue, and limit dilution of growth factors secreted by accessory cells, while still allowing for diffusion of vital nutrients and gases. The material used for nerve-guide production should neither provoke inflammation nor enhance scar formation, and ought to be generally biocompatible without release of toxic degradation products, such as an excessively rapid shift in pH due to acidic hydrolysis. In addition to these biological aspects, nerve guides should be pliable but able to resist collapse during body movement, should maintain their shape during regeneration, and should undergo resorption after regeneration is complete. Polymer hydrolysis or degradation times of resorbable nerve guides should be temporally adapted to the regeneration period required to bridge nerve gaps. This aspect is of vital importance, because myelination of axons follows regrowth, increasing the diameter of individual fibers and leading in turn to compression if inert conduit walls are used.

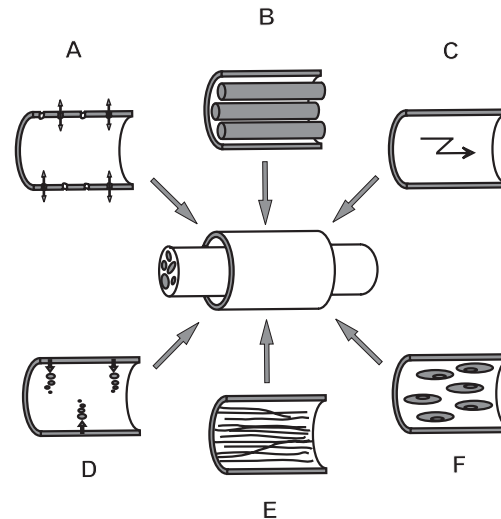


Fig. 3 Nerve-guide concepts. Aspects of nerve-guide designs that promise to aid medical performance are: (A) Biodegradability and porosity to control nutrient exchange and exclude scar-forming cells; (B) intraluminal channels to mimic the fascicular structure of nerves; (C) electrical microdevices, piezoelectric elements, and electrets to stimulate growth or externally control axonal activity; (D) controlled release of regenerative factors; (E) longitudinally-oriented matrix to guide regenerating axons; and (F) incorporation of support cells. (Adapted from Hudson, T.W., Evans, G.R., and Schmidt, C.E., *Engineering strategies for peripheral nerve repair*, 1999, *Tissue Engineering*, **26**, 617-628.)

In addition, nerve guides may be fashioned to actively stimulate regeneration and play a structural role during axonal outgrowth. Surface texture, longitudinally-oriented microchannels or polymer fibers, porosity, and electrical stimulation devices are adjuncts that may improve the regenerative performance of guidance channels. The impregnation, inclusion, or covalent linkage of biomolecules to conduit polymers (smart materials) represent means to regulate neuronal survival via retrograde transport, axonal elongation, proliferation of Schwann cells, chemotaxis of blood vessel-forming endothelial cells, or repulsion of scar-forming fibroblasts. In addition, Schwann cells or genetically-engineered fibroblasts that act as self-sustained biofactories for growth factors and ECM proteins may be introduced into nerve guides. Several conduit concepts are summarized in Fig. 3. A broad range of synthetic polymers for nerve-guide production has been reviewed.^[12]

THE SILICONE TUBE MODEL/ NONRESORBABLE NERVE GUIDES

The chronological sequence of regeneration in nerve guides was evaluated in simple inert silicone tubes used

to bridge 10-mm gaps of rat sciatic nerve.^[13] Although this model has been shown to be relevant only for very short regeneration distances, it elucidates four principle steps: 1) matrix formation; 2) cell migration; 3) axon migration; and 4) myelination. 1) After the proximal and distal nerve stumps have been coadapted to an empty silicone tube, the conduit fills with wound fluid within hours. Because the fluid contains the clotting factor fibrinogen, a fibrin network forms within the first week. If the tube distance does not exceed 10 mm, this network may be oriented in a longitudinal fashion. 2) Within the second week (or earlier), fibroblasts, Schwann cells, endothelial cells, and macrophages migrate from both nerve stumps into the tube. 3) Axons appear to be the last cellular constituents, despite the fact that axonal migration is in principle nearly one order of magnitude faster than the migration of Schwann cells.^[14] The likely explanation for this temporally reversed regeneration sequence in conduits is the fact that Schwann cells must biologically pave the way for axons. Based on histological sections, the forefront of regenerating axons does not exceed the front of migrating Schwann cells. Axons sprouting from the proximal nerve end traverse the silicone tube within about 4 weeks. 4) During this period, myelination will have begun, leading to nerve-fiber conduction velocities determined to be 60% of control values.^[15] This paradigm clearly highlights the basic regeneration potential of peripheral nerves, including removal of debris after degeneration, axonal regrowth, myelination, angiogenesis, and successful synaptogenesis. Nevertheless, it has been acknowledged that inert, nonpermeable conduits do not provide a therapeutic solution for nerve gaps longer than 5–10 mm.^[16]

Because of their inert and elastic properties, silicone-based nerve guides were among the first and most frequently used implants, including applications in humans (discussed subsequently). Entubulation of regenerating nerves, however, suffered drawbacks due to fibrosis and chronic nerve compression. In a number of cases, pronounced initial axonal regrowth and functional recovery were observed using nerve-guide tubes. Unfortunately, a decline in sensory and motor performance became evident after longer recovery periods. Histological analysis revealed that progressive myelination led to deleterious compression of axons within the limited volume of the conduit. These observations led to the conclusions that: 1) nerve guides should be resorbable; and 2) disintegration of nerve guides should begin soon after axons have traversed the conduit. We suggest that conduits are ideally resorbed in a graded fashion, with fast degradation near the proximal nerve end and slower resorption at the distal end. This concept may be especially important in the case of nerve-guide implants traversing long distances.

Electrical stimulation via leads integrated into silicone tubes and used as cuffs surrounding lesioned nerves was shown to increase regeneration in the rat model.^[17] Biocompatible electrets such as polytetrafluorethylene (PTFE) are dielectric polymers that provide electrically stable, monopolar charges without the need for an external power supply. Polyvinylidene fluoride (PVDF) is a piezoelectric material that displays transient charges depending on the degree of physical deformation (i.e., dipole displacement). Although the cellular and molecular mechanisms remain unclear, nerve guides made of positively-poled PVDF and PTFE appear to facilitate in-vivo regeneration to a higher degree than negatively-poled tubes.^[18,19] To our knowledge, the feasibility of clinical applications for this scientifically interesting approach implementing nonresorbable PTFE and PVDF implants has not yet been examined.

RESORBABLE ARTIFICIAL NERVE GUIDES

Numerous biological and resorbable synthetic materials have been investigated for neural-tissue engineering applications (Table 1). Tubes of poly- β -hydroxybutyrate (PHB), purified from energy storage granules of microorganisms, have been employed in the rabbit common peroneal-nerve injury model. Nerve regeneration across 2–4 cm gaps was observed 1–2 months postoperative, similar to the recovery seen using autologous nerve transplants.^[20] Nerve guides have also been synthesized from chemically modified collagen type I. Instead of tubes, 2000 collagen filaments 20 μ m in diameter were stabilized with a cross-linking epoxy reagent. Experimental defects in rat sciatic nerves were bridged using these 20-mm implants. Eight weeks postoperative the implants displayed the same amount of regeneration as autologous nerve transplants (about 5000 myelinated axons, 2–3 μ m thick).^[21] One of the earliest grafts was made from a Vicryl mesh (polyglactin or PLGA 10:90, noted subsequently). This graft resulted in some nerve regeneration, but provoked minor inflammatory responses and yielded nerve fibers with a morphology slightly different from that of normal nerves.^[22,23] Another family of degradable polymers includes polyglycolic acid (PGA),^[24,25] poly(L-lactide-co-caprolactone),^[26] and PLGA, which contains copolymers of lactic acid and glycolic acid in various ratios. These polymers were especially attractive because of their early availability, ease of processing, and approval by the U.S. Food and Drug Administration (FDA). For a series of studies, 75:25 PLGA was formed into tubes with 1.6 mm inner diameter and 3.2 mm outer diameter. These tubular conduits were interposed in rat sciatic



Table 1 Nerve-guide materials^a

Polymeric materials
Nonresorbable
Nonporous
Ethylene–vinyl acetate copolymer (EVA)
Polytetrafluoroethylene (PTFE)
Polyethylene (PE)
Silicone elastomers (SE)
Polyvinyl chloride (PVC)
Microporous
Expanded polytetrafluoroethylene (ePTFE)
Millipore (cellulose filter)
Semipermeable
Polyacrylonitrile (PAN)
Polyacrylonitrile/poly vinyl chloride (PAN/PVC)
Polysulfone (PS)
Piezoelectric
Poly vinylidene fluoride (PVOF)
PTFE-electret
Release of trophic factors
Ethylene vinyl acetate (EVA)
Bioresorbable
Polyglycolide (PGA)
Polylactide (PLLA)
PGA/PLLA blends
Polycaprolactone (PCL)
Polyhydroxybutyrate (PHB)
Polyhydroxyethyl methacrylate
Metals
Stainless steel
Tantalum
Biological Materials
Artery/vein
Decalcified bone
Collagen/gelatine
Hyaluronic acid derivatives
Alginate

^aA selection of classified synthetic and biological components used for generation of nerve guides is listed.

(Adapted from Valentini, R.F. and Aebischer, P., Strategies for the Engineering of Peripheral Nervous Tissue Regeneration. In R.P. Lanza, R. Langer, and W.L. Chick (Eds.), Principles of Tissue Engineering, Academic Press, R.G. Landes Company, Austin, 1997, pp. 671–684.)

nerve defects. Walking track analysis revealed restoration of muscle innervation and sciatic functional indexes approaching control values at 12 weeks.^[27] However, histological data demonstrated an

inappropriately early partial collapse of the conduits. As an alternative, poly(L-lactic acid) (PLLA) was used in subsequent trials to generate more stable geometries. In-vivo evaluation suggested the PLLA conduits to be superior to PLGA conduits of the same design, because there was no evidence of conduit collapse or breakage upon ambulation.^[28] However, unsatisfactory results were obtained with poly(L-lactic acid) blended with 10% poly(L-lactic acid) oligomer conduits in the rat spinal-cord lesion model. Despite obvious axonal regrowth and myelination over the first 2 months, the rate of regeneration declined thereafter, possibly due to compression and hazardous fragmentation of the polymer walls.^[29] In another approach, random copolymers of trimethylene carbonate and ϵ -caprolactone (50:50) showed good biocompatibility in vitro (Schwann-cell survival, adhesion, metabolism, and proliferation) and only a mild soft tissue reaction with no adverse chronic inflammatory reaction in vivo.^[30] Poly(L-lactide-co-caprolactone) (PLA) employed in vivo as 12-mm long semipermeable tubes with walls 1.3-mm thick yielded good regeneration over a period of 3–10 weeks.^[31] As later discussed in greater detail, these examples elucidate the delicate task of balancing the needs for stability, flexibility, and such structural features as porosity for metabolic diffusion with the requirement that implants degrade spontaneously (representing a loss of the aforementioned attributes).

In addition to polyesters, resorbable polyurethanes^[32] and poly(organo)phosphazine^[33] were employed as guide-channel materials. Although the number of available conduit materials has continuously increased over the years, the need for further development is evident. To our knowledge to date, no composition has been found that provides all required mechanical properties and excellent biological performance without causing complications like inflammation or toxicity.

BIOHYBRID NERVE GUIDES

This review considers biohybrid nerve guides composed of both synthetic and biological components (proteins, cells).

ECM-Derived Guides

Because growth cones (the leading tip of regrowing axons) require adhesion to appropriate substrata for proper growth, various ECM components including laminin, collagen, and fibronectin or mixtures of components (Matrigel[®]) have been introduced into polymer conduits. Alternatively, short oligopeptides of 3–10

amino acids (GRGDS, YIGSR, SIKVAV, etc.) that are recognized by axonal integrin receptors have successfully been employed as substitutes for large ECM molecules. These short-recognition sequences are less likely to elicit immune responses and more durable than large molecules. Three-dimensional hydrogels chemically modified with laminin-derived oligopeptides promoted axonal outgrowth *in vitro* and *in vivo*.^[34] Another innovative approach was incorporation of peptides derived from laminin and the cell-adhesion molecule N-cadherin into a fibrin matrix. With the enzymatic aid of factor XIIIa, lysine residues of neuroactive peptides were cross-linked to glutamine residues of fibrin. The resulting matrix resulted in enhanced regeneration when used as a nerve guide in lesioned rats.^[35]

Hydrogels, Alginate

Other bioresorbable conduit materials used in a biohybrid composition are hydrogels such as poly(2-hydroxyethyl methacrylate)^[36] or alginate, a polysaccharide composed of 1,4-linked β -D-mannuronate and 1,4-linked α -L-glucuronate blocks. Alginate has been used to form nerve-guide implants as both a sponge alone and as spongy filling for a polyglycolic acid mesh (PGA).^[37] Gaps in cat sciatic nerves 50 mm in length were successfully bridged using these implants, with locomotor function recovered within 8 months in both approaches. By this time, the myelinated nerve displayed a fascicular organization and alginate had been completely resorbed. Even longer nerve gaps were investigated in dogs. Polyglycolic acid-collagen tubes 80 mm in length and filled with either laminin-soaked sponges or collagen fibers were implanted into severed peroneal nerves. Motor and sensory nerves had partially regenerated within 12 months, with reestablishment of functional synapses (as deduced from essentially normal dorsiflexion of ankle joints and based on histological data).^[38]

Alginate has also been used to coat PHB fibers together with fibronectin. The coated PHB fibers were implanted into lesioned rat spinal cord. Eight weeks postoperative, 50% of rubrospinal neurons could be rescued, although they are most affected in this paradigm.^[39] Although no enhanced axonal regrowth onto these fibers was evident, the neuroprotective effect observed may have been the result of coated PHB fibers serving as a molecular scaffold for diffusible growth factors, consequently increasing the local concentration or half-life of growth factors.

Growth Factors

Factors that support neuronal survival (neurotrophic) or attract axons (neuritotropic) include nerve growth

factor (NGF), brain-derived neurotrophic factor (BDNF), neurotrophin-3 and -4/5 (NT-3, NT-4/5), platelet-derived growth factor (PDGF), glial cell line-derived neurotrophic factor (GDNF), acidic and basic fibroblast growth factor (aFGF, bFGF), ciliary neurotrophic factor (CNTF), and insulin-like growth factors (IGF), to name just the most prominent ones. Various factors address neurons, whereas, e.g., NGF also promotes Schwann-cell proliferation and migration in an autocrine fashion.^[40] Nerve injury leads to a biphasic synthesis of NGF, with a first peak 6 hours and a second peak 3 days postinjury. Blood-borne macrophages are attracted during the first phase, whereas during the second phase, Schwann-cell responses are triggered.^[41] A number of studies have demonstrated that coapplication of these factors enhances regeneration.^[39,42-44] Critical aspects of applying macromolecules *in vivo* can be rapid inactivation (conformational changes, enzymatic modifications) or degradation *in vivo*. Therefore, it will be extremely helpful if drug-delivery and controlled-release techniques, and strategies for molecular stabilization of macromolecules for extended regenerative periods (i.e., weeks or months) can be improved.

Support Cells

The most promising alternative to the application of recombinant factors and synthetic peptides is the integration of self-sustained cells that produce many of these factors and may even respond to microenvironmental signals in self-regulatory feedback loops. Fibroblasts engineered to express the factors bFGF and NGF constitutively or in a drug-controlled fashion do not display all of these features but have been shown to enhance regeneration (e.g., after spinal cord lesions).^[45] Schwann cells appear to be more potent than transfected/transduced cells as additives in nerve guides.^[29] These cells are not only self-regulatory sources of growth factors NGF, BDNF, IGF-1, and CNTF, but also represent an excellent substratum for growing axons *in vitro* and *in vivo*.^[46] The cells produce a growth-permissive ECM and interact with blood vessel-forming endothelial cells (blood-nerve barrier) and connective tissue cells. Although Schwann cells are able to invade nerve conduits more than a limited distance from the distal and proximal nerve stumps, it appears to be advantageous to insert Schwann cells into conduits before implantation. *In vitro* axons migrate 8 times faster than Schwann cells^[14] (Fig. 4), while axons *in vivo* do not migrate beyond the forefront of Schwann cells, suggesting that the slow proliferation of Schwann cells is the limiting factor during regeneration through nerve conduits.^[47] Procedures have recently been established to purify



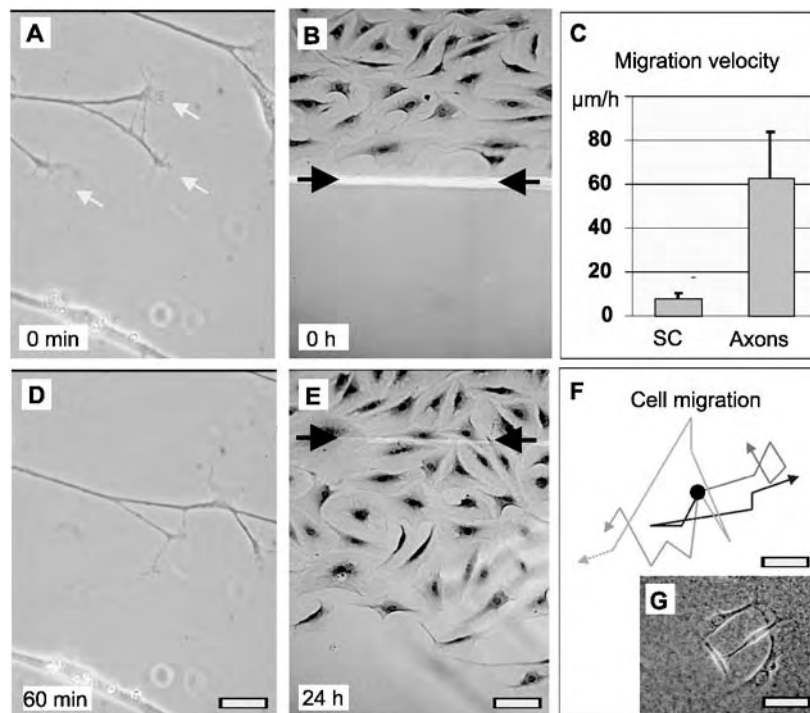


Fig. 4 Migration and growth velocities of Schwann cells and axons. Dorsal root ganglia were explanted in vitro and axonal outgrowth was monitored via time-lapse video recording, shown (A) at time point 0 and (D) after 60 min. Schwann cells were synchronized in a two-field migration assay at a start line (the line between black arrows in part B). (E) After 24 h, cultures were fixed and Coomassie-stained, and the migration distance was determined. (C) Quantitative comparison revealed that axons migrate about 8 times faster than Schwann cells. (F) Time-lapse video recording of low-density cultures was also used to analyze the directionality of Schwann-cell migration. The enlarged color-coded trajectories of four Schwann cells were plotted from a virtual starting point (black dot in F). Cell positions were determined every 30 minutes. (G) Cell migration was nondirected. Cells were evaluated for (F) at time point 0. All figures represent views from the top. Scale bars (D for A and D): 25 µm; (E for B and E): 50 µm; (G): 65 µm; (F): 60 µm. (From Ref. [14] with permission.)

human and rat adult Schwann cells and accelerate their growth via the combinatorial addition of glial growth factor, forskolin, and isobutylmethylxanthine, whereby the proliferation of fibroblasts becomes inhibited.^[48] A combination of methods may further improve the biological performance. Regeneration in a spinal-cord lesion model may be significantly enhanced by implantation of Schwann-cell seeded polyacrylonitrile-polyvinylchloride copolymer conduits and coapplication of BDNF and NT-3 into neighboring spinal-cord tissue with a miniosmotic pump.^[49]

PHYSICAL PROPERTIES OF NERVE GUIDES

Physical and physicochemical properties of nerve guides are of pivotal importance for successful regeneration. This survey focuses on nerve guides in the form of tubular conduits, although guides composed of meshes or a collection of filaments have been introduced.^[21,39] Features such as conduit dimension, resorption time, degradation by-products, wall porosity,

surface texture and hydrophilicity, flexibility, and torsional stability must be considered. Length, cross-sectional area, and wall thickness are important parameters, because nerve guides should be adapted to the native dimensions of the severed nerve, which in the clinical setting may range from 1–10 mm in diameter and from 0.5 to about 50 cm in length. In the rat model, consistently good regeneration has been achieved for lengths up to 1 cm, whereas longer distances could often be bridged only upon coapplication of growth factors or growth factor-producing cells. A comparative study suggested that for a distinct nerve-lesion model, a conduit diameter of 1.8 mm was superior to smaller (1.2 mm) and larger (3.1 mm) diameters. These restrictions may reflect counteracting phenomena—excessively small diameters are likely to cause damage upon insertion of nerve stumps into conduits, or may later result in compression during myelination of axons. Conversely, an overly large opening may permit infiltration of scar-forming fibroblasts and outward diffusion of vital neurotrophic factors produced by migrating Schwann cells.

Furthermore, sufficient diffusion of gases and metabolites across the conduit diameter and through the conduit wall is a prerequisite for cell survival and differentiation—especially during the initial phase of regeneration, when blood vessels have not yet been formed. Semipermeable conduit walls are clearly preferable to nonpermeable walls, because these allow in- and outward diffusion of nutrients and intoxicating cellular waste products but hamper the infiltration of scarring fibroblasts. Again, counteracting aspects must be balanced, because longer diffusion distances and lower wall porosity will lead to oxygen concentrations lower than 2 µg/ml, which is considered the minimal requirement for axonal elongation.^[1] Calculations have suggested that the diffusion distance should not exceed 1 mm. In addition, it is noteworthy that an early increase in oxygen levels in permselective tubes may minimize connective tissue formation, due to hindered fibroblast migration. On the other hand, limited anoxia represents a natural stimulus for endothelial cell invasion and subsequent angiogenesis, which are indispensable.

Pore size may be a means to not only regulate the in- and outward movement of cells versus molecules, but also to differentiate between molecules of different sizes.^[50] Some growth-inhibiting, high-molecular-weight ECM components (e.g., chondroitin sulfate proteoglycan) produced by the tissue that fills the nerve guide may possibly be excluded. For resorbable conduits, the ability to act as a molecular sieve typically depends exclusively on inner and outer surface layers just microns thick. The temporal stability of these thin layers must be considered, because the pores will enlarge with progressive degradation. All of these requirements must be met while still providing compression resistance and flexibility, as discussed earlier. As a basis for nerve-guide generation, it is noteworthy that in animal experiments, compression of 25 mm Hg results in reduced blood flow in epineural venules. Long-term pressure on endoneurial capillaries increases permeability, because oxygen depletion leads to edema in the endoneurium. Extraneural pressure of 80 mm Hg results in total intraneural ischemia.^[51] Notably, nonmyelinated axons (and possibly early axons) are less vulnerable to the effects of compression than are myelinated axons.

The selection of distinct synthetic nerve-guide materials will determine the polymer's crystallinity, and consequently affects not only the rigidity and flexibility of the polymer, but also its rates of hydrolyzation and degradation. Polymers such as PGA and PLA are resorbed after acidic hydrolysis, which acidifies the microenvironment. (This process accelerates with decreasing pH.) The total amount of polymer used will determine whether cells are capable of buffering the increasing hydrogen ion concentration in order to survive. This leads back to the question of how best to

design nerve guides—conduit walls, in particular—that are only transiently mechanically stable. (This is because the less polymer used, the lower the degree of acidification upon degradation.)

Hydrophilicity and surface texture are additional features to consider. PLA, PGA, alginate, and others are essentially nonpermissive for outgrowing axons, although Schwann-cell attachment is partially evident. Increasing the hydrophilicity via oxygen-plasma treatment improves cell attachment but simultaneously decreases the stability of nerve-guide walls.^[14] Surface functionalization via coating with polycationic polymers such as degradation-stabilized poly-D-lysine improves cell attachment.^[46] Additionally, accelerated regeneration is likely to occur if meandering of axons in the nerve guide can be prevented. As mentioned earlier, some approaches use collagen fibers as a scaffold,^[21,38,52] although microstructured compact polymer filaments with microgrooves are favored. These microgrooves induce longitudinal alignment of Schwann cells. Cocultures with nervous tissue *in vitro* have shown that axons orient themselves along the Schwann-cell chains very much as they align along the bands of Büngner during regeneration *in vivo* (Lietz, Tschöcke, Doser, and Schlosshauer, unpublished). In this manner, the direct passage of regrowing axons through nerve guides will likely result in faster synaptogenesis and thus stop the ongoing atrophy of previously denervated muscle fibers. Surface texture in the form of an oriented microstructure plays the central role in this regeneration concept. Surface texture also appears to be fundamental for the stability of the ECM–nerve conduit linkage, based on comparison of regeneration through smooth- and rough-walled conduits.^[53]

Microelectrodes represent an additional means to improve nerve implants, by allowing for communication with regenerating axons and external modulation of muscle function. Outgrowing axons may be demonstrated to penetrate polyimide-embedded sieve electrodes.^[54] These electrode arrays were integrated into short nerve guides and implanted into lesioned rat sciatic nerves. Stimulation of the regenerated axons evoked nerve and muscle responses.^[55]

HUMAN APPLICATIONS

The clinical application of nerve guides has been limited in the past, but usage is expected to increase with the recent advent of commercially available implants. Polytetrafluorethylene (ePTFE, Gore-Tex) conduits were implanted in the upper extremities for clinical repair of 15–60-cm long median and ulnar nerve defects in 43 patients. The recovery of motor and sensory function was nearly 80% for repairs up to 40 mm in length, although the rate of recovery declined to about



10% if the implants were longer.^[56] Also using non-resorbable Gore-Tex tubing implants, only minor regeneration was observed in 7 patients (16–56 years of age, 2–15 mm defects) who suffered from lesions of the inferior alveolar and lingual nerves. In another early study, median nerves were successfully reconstructed via silicone entubulation in 2 male patients with nerve gaps of 3–5 mm. At 36 months, motor recovery was about 80%.^[57] When the same authors compared microsurgical and tubular repair (silicone conduits) in a subsequent randomized clinical prospective study, follow-up examinations after 4–5 years revealed comparable recovery of sensory and motor function in both groups.^[58] In some cases where median and ulnar nerves had been severed, gaps of up to 30 mm were bridged with silicone-based conduits.^[59]

Clinical nerve reconstruction with poly(glycolic acid) tubes was first performed in 15 patients with 5- to 30-mm gaps in digital nerves.^[24] Sensibility was evaluated within a postoperative interval of 11 to 32 months, with excellent to good recovery observed in 86% of all cases. Absence of pain at the site of reconstruction was noted by 73% of patients. Additionally, a randomized multiclinical prospective study of nerve guides made from poly(glycolic acid) was performed.^[60] This study enrolled 98 subjects with a total of 136 nerve transections in the hand. Treatment involved implantation of autologous nerve grafts, poly(glycolic acid) conduits, or end-to-end repair. Two-point discrimination was quantified at 3, 6, 9, and 12 months postsurgery. After receiving short implants, about 75% of all patients in all groups exhibited good to excellent recovery. In patients with implants of both types longer than 8 mm, PGA conduits provided better regeneration. Seven of 17 PGA patients showed excellent recovery, but none in the group with autologous nerve transplants displayed this level of improvement.

Besides bridging lesion gaps, nerve-guide implantation promises to be a helpful approach to prevent aberrant axonal sprouting and formation of painful neuromas. A case report described secondary digital reconstruction in the foot after a posttraumatic neuroma of the medial plantar nerve had formed. Introduction of the nerve trunk into a bioresorbable PGA nerve conduit after complete neuroma resection eliminated neuroma pain in a male child.^[61] Silicone tubes 20 mm long were also used to treat neuromas after accidental transection of radial nerves at wrist levels, with patients describing different degrees of pain relief.^[62]

THE ECONOMIC IMPACT

More than 8 million surgical interventions are performed in the United States each year to treat organ and tissue deficiencies. The annual costs associated with these

medical treatments exceed \$400 billion.^[63,64] Because tissue engineering has considerable potential for improving quality of life and reducing health care costs, operating expenses allocated for tissue engineering exceeded \$450 million in 1997. The amount of tissue research increases 20–25% per annum, making spin-off benefits for neural-tissue engineering more than likely. Future savings for the health care field due to these activities has been estimated to be in the range of \$80 billion, in addition to the value of their positive impact on quality-of-life. By 1998, the capital value of start-up companies with a focus on tissue engineering was \$1.7 billion.^[65]

The biotechnological progress made in various fields of tissue engineering will also foster therapeutic developments in neuropathology. Some 400,000 patients suffer neurotraumas of the peripheral nervous system annually, subjecting them to neurosurgery (www.polygarnics.com). Additionally, the many patients suffering from idiopathic damage, compression syndromes, traumata, iatrogenic injuries, and systemic diseases must be considered. In Germany, about 1% of the population of 80 million suffers a sports-related injury annually, with about 5% of all peripheral-nerve lesions arising from these injuries.^[66] With regard to the central nervous system, spinal-cord injuries have an annual incidence of 30,000 cases worldwide.^[67] These lesions could be treated to some extent using nerve-guide implants. The market for nerve-guide products in the peripheral nervous system has been calculated to be in the range of \$50 million, plus applications in the central nervous system. To our knowledge, by the summer of 2003, three SME-biotech-companies had received approval for their nerve-guide implants from the U.S. FDA. There is clearly a medical need for nerve-guide implants. The benefits gained by the health care field will likely compensate for expenses arising from research spending, especially when advances arising from neurological tissue engineering research^[1–3] are combined with advances from nonneural research.

CONCLUSION

It is obvious that future developments will benefit from more detailed insights into the intercellular and molecular regulatory mechanisms of regeneration, yielding intelligent nerve-guide designs. Axonal regeneration recapitulates embryonic development to a considerable extent, as deduced from gene expression studies.^[68] Integration of knowledge about growth-promoting factors and, perhaps more important, inhibitory and repulsive factors in the CNS and PNS will foster implant development. Major classes of these factors are ephrins, semaphorins, netrins, slits, NOGOS, and proteoglycans. As for the latter factors, it has been shown that their enforced degradation greatly enhances regeneration

after spinal-cord lesions.^[8] Integration of counteractive therapeutic means such as pharmacologically active antagonists of inflammatory factors, regeneration-suppressing molecules, or their second-messenger cascades will lend new qualities to nerve guides.^[69] Such approaches will also allow more stringent treatment of CNS lesions, where current conduit implantations merely serve to bridge fluid-filled cysts after spinal-cord lesions. The suggested biohybrid concepts are likely to successfully manipulate the inhibitory microenvironment in order to induce reentry and synaptogenesis of regrowing axons into the spinal cord. Besides the focus on repulsive or inhibitory growth mechanisms (fibrosis, myelin-mediated inhibition, etc.), emphasis should be also put on aspects of angiogenesis. Recent efforts concentrate principally on axonal regrowth, to the extent that support cells and nonneural tissue domains have been essentially neglected. For treatment of neuropathic pain (nerve compression without transection), which spares nerve continuity and blood vessels, regeneration performs at its best. This situation should probably be duplicated in neurotmesis (complete transection) by establishing a capillary network before regeneration of long axons is attempted. In this way an optimal nutrient supply may be established, helping to overcome the current restriction; namely, that nerve guides often fail to facilitate regeneration over distances longer than about 4 cm. The medical need for neural-tissue engineering products and the existing market for medical products of this type demand quick remedies for the present shortcomings of nerve-regeneration treatments. Thanks to advances in materials sciences as well as in life sciences, nerve-guide development is predicted to accelerate in the near future, providing promising solutions.

ACKNOWLEDGMENTS

We are grateful to Drs. Carl (U.S.A.), Fricke (Reutlingen), Lundborg (Stockholm), Stoop (Reutlingen), and Schaller (Tuebingen) for critical discussions, figures, and correction of the manuscript.

ARTICLES OF FURTHER INTEREST

Nanoscale Neurosurgery; Neuroprostheses; Tissue Engineering of Peripheral Nerve

REFERENCES

1. Rutkowski, G.E.; Heath, C.A. Development of a bioartificial nerve graft. I. Design based on a reaction-diffusion model. *Biotechnol. Prog.* **2002**, *18*, 362–372.
2. Dunnett, S.B.; Björklund, A. Prospects for new restorative and neuroprotective treatments in Parkinson's disease. *Nature* **1999**, *399*, A32–A39.
3. Hottinger, A.F.; Azzouz, M.; Déglon, N.; Aebischer, P.; Zurn, A.D. Complete and long-term rescue of lesioned adult motoneurons by lentiviral-mediated expression of glial cell line-derived neurotrophic factor in the facial nucleus. *J. Neurosci.* **2000**, *20* (15), 5587–5593.
4. Lundborg, G.; Dahlin, L.B. Anatomy, function, and pathophysiology of peripheral nerves and nerve compression. *Hand Clin.* **1996**, *12* (2), 185–193.
5. Seddon, H. *Surgical Disorders of the Peripheral Nerves*, 2nd Ed.; Churchill and Livingstone: Edinburgh, 1972.
6. Horner, P.J.; Gage, F.H. Regenerating the damaged central nervous system. *Nature* **2000**, *407*, 963–970.
7. Doolabh, V.B.; Hertl, M.C.; Mackinnon, S.E. The role of conduits in nerve repair: A review. *Rev. Neurosci.* **1996**, *7*, 47–84.
8. Bradbury, E.J.; Moon, L.D.; Popat, R.J.; King, V.R.; Bennett, G.S.; Patel, P.N.; Fawcett, J.W.; McMahon, S.B. Chondroitinase ABC promotes functional recovery after spinal cord injury. *Nature* **2002**, *416* (6881), 636–640.
9. Gulati, A.K.; Cole, G.P. Immunogenicity and regenerative potential of acellular nerve allografts to repair peripheral nerve in rats and rabbits. *Acta Neurochir. (Wien)* **1994**, *126* (2–4), 158–164.
10. Sondell, M.; Lundborg, G.; Kanje, M. Regeneration of the rat sciatic nerve into allografts made acellular through chemical extraction. *Brain Res.* **1998**, *795* (1–2), 44–54.
11. Johnson, P.C.; Duhamel, R.C.; Meezan, E.; Brendel, K. Preparation of cell-free extracellular matrix from human peripheral nerve. *Muscle Nerve* **1982**, *5* (4), 335–344.
12. Pachence, J.M.; Kohn, J. *Biodegradable Polymers for Tissue Engineering*; Lanza, R.P., Langer, R., Chick, W.L., Eds.; Academic Press: San Diego, CA, 1997; 273–293.
13. Williams, L.R.; Longo, F.M.; Powell, H.C.; Lundborg, G.; Varon, S. Spatial-temporal progress of peripheral nerve regeneration within a silicone chamber: Parameters for a bioassay. *J. Comp. Neurol.* **1983**, *218* (4), 460–470.
14. Schlosshauer, B.; Muller, E.; Schroder, B.; Planck, H.; Muller, H.W. Rat Schwann cells in bioresorbable nerve guides to promote and accelerate axonal regeneration. *Brain Res.* **2003**, *963* (1–2), 321–326.
15. Fields, R.D.; Ellisman, M.H. Axons regenerated through silicone tube splices. I. Conduction properties. *Exp. Neurol.* **1986**, *92* (1), 48–60.
16. Lundborg, G.; Dahlin, L.B.; Danielsen, N. Nerve regeneration in silicone chambers: Influence of gap length and distal stump components. *Exp. Neurol.* **1982**, *76*, 361–375.
17. Kerns, J.M.; Fakhouri, A.J.; Weinrib, H.P.; Freeman, J.A. Electrical stimulation of nerve regeneration in the rat: The early effects evaluated by a vibrating probe and electron microscopy. *Neuroscience* **1991**, *40* (1), 93–107.
18. Aebischer, P.; Valentini, R.F.; Dario, P.; Domenici, C.; Galletti, P.M. Piezoelectric guidance channels enhance



- regeneration in the mouse sciatic nerve after axotomy. *Brain Res.* **1987**, *436* (1), 165–168.
19. Valentini, R.F.; Sabatini, A.M.; Dario, P.; Aebischer, P. Polymer electret guidance channels enhance peripheral nerve regeneration in mice. *Brain Res.* **1989**, *480* (1–2), 300–304.
 20. Young, R.C.; Wiberg, M.; Terenghi, G. Poly-3-hydroxybutyrate (PHB): A resorbable conduit for long gap repair in peripheral nerves. *Br. J. Plast. Surg.* **2002**, *55*, 235–240.
 21. Yoshii, S.; Oka, M. Collagen filaments as a scaffold for nerve regeneration. *J. Biomed. Mater. Res.* **2001**, *56* (3), 400–405.
 22. Molander, H.; Olsson, Y.; Engquist, O.; Bowald, S.; Eriksson, I. Regeneration of peripheral nerve through a polyglactin tube. *Muscle Nerve* **1982**, *5*, 54.
 23. Gibson, K.L.; Remson, L.; Smith, A.; Satterlee, N.; Strain, G.M.; Daniloff, J.K. Comparison of nerve regeneration through different types of neural prostheses. *Microsurgery* **1991**, *12* (2), 80–85.
 24. Mackinnon, S.E.; Dellon, A.L. Clinical nerve reconstruction with a bioabsorbable polyglycolic acid tube. *Plast. Reconstr. Surg.* **1990**, *85* (3), 419–424.
 25. Keeley, R.D.; Nguyen, K.D.; Stephanides, M.J.; Padilla, J.; Rosen, J.M. The artificial nerve graft: A comparison of blended elastomer-hydrogel with polyglycolic acid conduits. *J. Reconstr. Microsurg.* **1991**, *7* (2), 93–100.
 26. den Dunnen, W.F.A.; Schakenraad, J.M.; Zondervan, G.J.; Pennings, A.J.; van der Lei, B.; Robinson, P.H. A new PLLA/PCL copolymer for nerve regeneration. *J. Mater. Sci., Mater. Med.* **1993**, *4*, 521–525.
 27. Hare, G.M.; Evans, P.J.; Mackinnon, S.E.; Best, T.J.; Bain, J.R.; Szalai, J.P.; Hunter, D.A. Walking track analysis: A long-term assessment of peripheral nerve recovery. *Plast. Reconstr. Surg.* **1992**, *89* (2), 251–258.
 28. Evans, G.R.D.; Brandt, K.; Widmer, M. In vivo evaluation of poly(L-Lactic acid) porous conduits for peripheral nerve regeneration. *Biomaterials* **1999**, *20*, 1109–1115.
 29. Oudega, M.; Gautier, S.E.; Chapon, P.; Fragoso, M.; Bates, M.L.; Parel, J.M.; Bunge, M.B. Axonal regeneration into Schwann cell grafts within resorbable poly(alpha-hydroxyacid) guidance channels in the adult rat spinal cord. *Biomaterials* **2001**, *22*, 1125–1136.
 30. Fabre, T.; Schappacher, M.; Bareille, R.; Dupuy, B.; Soum, A.; Bertrand-Barat, J.; Baquey, C. Study of a (trimethylenecarbonate-co-ε-caprolactone) polymer—Part 2: In vitro cytocompatibility analysis and in vivo ED1 cell response of a new nerve guide. *Biomaterials* **2001**, *22*, 2951–2958.
 31. den Dunnen, W.F.A.; Stokroos, I.; Blaauw, E.H.; Holwerda, A.; Pennings, A.J.; Robinson, P.H.; Schakenraad, J.M. Light-microscopic and electron-microscopic evaluation of short-term nerve regeneration using a biodegradable poly(DL-lactide-epsilon-caprolacton) nerve guide. *J. Biomed. Mater. Res.* **1996**, *31*, 105–115.
 32. Robinson, P.H.; van der Lei, B.; Hoppen, H.J.; Leenslag, J.W.; Pennings, A.J.; Nieuwenhuis, P. Nerve regeneration through a two-ply biodegradable nerve guide in the rat and the influence of ACTH4–9 nerve growth factor. *Microsurgery* **1991**, *12* (6), 412–419.
 33. Langone, F.; Lora, S.; Veronese, F.M.; Caliceti, P.; Parnigotto, P.P.; Valenti, F.; Palma, G. Peripheral nerve repair using a poly(organo)phosphazene tubular prosthesis. *Biomaterials* **1995**, *16* (5), 347–353.
 34. Bellamkonda, R.; Ranieri, J.P.; Aebischer, P. Laminin oligopeptide derivatized agarose gels allow three-dimensional neurite extension in vitro. *J. Neurosci. Res.* **1995**, *41*, 501–509.
 35. Schense, J.C.; Bloch, J.; Aebischer, P.; Hubbell, J.A. Enzymatic incorporation of bioactive peptides into fibrin matrices enhances neurite extension. *Nat. Biotechnol.* **2000**, *18*, 415–419.
 36. Woerly, S. *Microreconstruction Surgery of the Central Nervous System by Means of Engineered Polymer Constructs*; Sames, K., Ed.; Ecomed: Landsberg-Lech, 2000; 1–14.
 37. Sufan, W.; Suzuki, Y.; Tanihara, M.; Ohnishi, K.; Suzuki, K.; Endo, K.; Nishimura, Y. Sciatic nerve regeneration through alginate with tubulation or non-tubulation repair in cat. *J. Neurotrauma* **2001**, *18* (3), 329–338.
 38. Toba, T.; Nakamura, T.; Shimizu, Y.; Matsumoto, K.; Ohnishi, K.; Fukuda, S.; Yoshitani, M.; Ueda, H.; Hori, Y.; Endo, K. Regeneration of canine peroneal nerve with the use of a polyglycolic acid-collagen tube filled with laminin-soaked collagen sponge: A comparative study of collagen sponge and collagen fibers as filling materials of nerve conduits. *J. Biomed. Mater. Res.* **2001**, *58*, 622–630.
 39. Novikov, L.N.; Novikova, L.N.; Mosahebi, A.; Wiberg, M.; Terenghi, G.; Kellerth, J.-O. A novel biodegradable implant for neuronal rescue and regeneration after spinal cord injury. *Biomaterials* **2002**, *23*, 3369–3376.
 40. Anton, E.S.; Weskamp, G.; Reichardt, L.F.; Matthew, W.D. Nerve growth factor and its low-affinity receptor promote Schwann cell migration. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91* (7), 2795–2799.
 41. Fu, S.Y.; Gordon, T. The cellular and molecular basis of peripheral nerve regeneration. *Mol. Neurobiol.* **1997**, *14* (1–2), 67–116.
 42. Barras, F.M.; Pasche, P.; Bouche, N.; Aebischer, P. Glial cell line-derived neurotrophic factor released by synthetic guidance channels promotes facial nerve regeneration in the rat. *J. Neurosci. Res.* **2002**, *70*, 746–755.
 43. Fine, E.G.; Decosterd, I.; Papaloizos, M.; Zurn, A.D.; Aebischer, P. GDNF and NGF released by synthetic guidance channels support sciatic nerve regeneration across a long gap. *Eur. J. Neurosci.* **2002**, *15*, 589–601.
 44. Bloch, J.; Fine, E.G.; Bouche, N.; Zurn, A.D.; Aebischer, P. Nerve growth factor- and neurotrophin-3-releasing guidance channels promote regeneration of the rat dorsal root. *Exp. Neurol.* **2001**, *172*, 425–432.
 45. Nakahara, Y.; Gage, F.H.; Tuszynski, M.H. Grafts of fibroblasts genetically modified to secrete NGF, BDNF, NT-3, or basic FGF elicit differential responses in the adult spinal cord. *Cell Transplant* **1996**, *5*, 191–204.

46. Steuer, H.; Fadale, R.; Müller, E.; Müller, H.W.; Planck, H.; Schlosshauer, B. Biohybrid nerve guide for regeneration: Degradable polylactide fibers coated with rat Schwann cells. *Neurosci. Lett.* **1999**, *277*, 165–168.
47. Hobson, M.I.; Green, C.J.; Terenghi, G. VEGF enhances intraneural angiogenesis and improves nerve regeneration after axotomy. *J. Anat.* **2000**, *197*, 591–605.
48. Rutkowski, J.L.; Kirk, C.J.; Lerner, M.A.; Tennekoon, G.I. Purification and expansion of human Schwann cells in vitro. *Nat. Med.* **1995**, *1* (1), 80–83.
49. Bamber, N.I.; Li, H.; Lu, X.; Oudega, M.; Aebischer, P.; Xu, X.M. Neurotrophins BDNF and NT-3 promote axonal re-entry into the distal host spinal cord through Schwann cell-seeded mini-channels. *Eur. J. Neurosci.* **2001**, *13*, 257–268.
50. Aebischer, P.; Guenard, V.; Winn, S.R.; Valentini, R.F.; Galletti, P.M. Blind-ended semipermeable guidance channels support peripheral nerve regeneration in the absence of a distal nerve stump. *Brain Res.* **1988**, *454*, 179–187.
51. Lundborg, G.; Dahlin, L.B.; Danielsen, N.; Nachemson, A.K. Tissue specificity in nerve regeneration. *Scand. J. Plast. Reconstr. Surg.* **1986**, *20* (3), 279–283.
52. Tong, X.J.; Hirai, K.; Shimada, H.; Mizutani, Y.; Izumi, T.; Toda, N.; Yu, P. Sciatic nerve regeneration navigated by laminin-fibronectin double coated biodegradable collagen grafts in rats. *Brain Res.* **1994**, *663* (1), 155–162.
53. Aebischer, P.; Guenard, V.; Brace, S. Peripheral nerve regeneration through blind-ended semipermeable guidance channels: Effect of the molecular weight cutoff. *J. Neurosci.* **1989**, *9* (10), 3590–3595.
54. Schlosshauer, B.; Brinker, T.; Müller, H.-W.; Meyer, J.-U. Towards micro electrode implants: In vitro guidance of rat spinal cord neurites through polyimide sieves by Schwann cells. *Brain Res.* **2001**.
55. Navarro, X.; Calvet, S.; Rodríguez, F.J.; Stieglitz, T.; Blau, C.; Butí, M.; Valderrama, E.; Meyer, J.U. Stimulation and recording from regenerated peripheral nerves through polyimide sieve electrodes. *J. Peripher. Nerv. Syst.* **1998**, *3* (2), 91–101.
56. Stanec, S.; Stanec, Z. Reconstruction of upper-extremity peripheral-nerve injuries with ePTFE conduits. *J. Reconstr. Microsurg.* **1998**, *14* (4), 227–232.
57. Lundborg, G.; Rosen, B.; Abrahamson, S.O.; Dahlin, L.; Danielsen, N. Tubular repair of the median nerve in the human forearm. Preliminary findings. *J. Hand Surg., Br.* **1994**, *19* (3), 273–276.
58. Lundborg, G.; Rosen, B.; Dahlin, L.; Danielsen, N.; Holmberg, J. Tubular versus conventional repair of median and ulnar nerves in the human forearm: Early results from a prospective, randomized, clinical study. *J. Hand Surg., Am.* **1997**, *22* (1), 99–106.
59. Braga-Silva, J. The use of silicone tubing in the late repair of the median and ulnar nerves in the forearm. *J. Hand Surg., Br.* **1999**, *24* (6), 703–706.
60. Weber, R.A.; Breidenbach, W.C.; Brown, R.E.; Jabaley, M.E.; Mass, D.P. A randomized prospective study of polyglycolic acid conduits for digital nerve reconstruction in humans. *Plast. Reconstr. Surg.* **2000**, *106* (5), 1036–1045.
61. Kim, J.; Dellon, A.L. Reconstruction of a painful post-traumatic medial plantar neuroma with a bioresorbable nerve conduit: A case report. *J. Foot Ankle Surg.* **2001**, *40* (5), 318–323.
62. Dahlin, L.B.; Lundborg, G. Use of tubes in peripheral nerve repair. *Neurosurg. Clin. N. Am.* **2001**, *12*, 341–352.
63. Patrick, C.W.J.; Mikos, A.G.; McIntire, L.V. *Prospectus in Tissue Engineering*; Pergamon: Oxford, 1998; 3–11.
64. Langer, R.; Vacanti, J.P. Tissue engineering. *Science* **1993**, *260*, 920–926.
65. Lysaght, M.J.; Nguy, N.A.; Sullivan, K. An economic survey of the emerging tissue engineering industry. *Tissue Eng.* **1998**, *4* (3), 231–238.
66. Lang, C.; Stefan, H. Sports injuries of the nervous system. *Fortschr. Neurol. Psychiatr.* **1999**, *67* (8), 373–384.
67. Marino, R.J.; Ditunno, J.F.; Donovan, W.H.; Maynard, F.J. Neurologic recovery after traumatic spinal cord injury: Data from the model spinal cord injury systems. *Arch. Phys. Med. Rehabil.* **1999**, *80* (11), 1391–1396.
68. Bonilla, I.E.; Tanabe, K.; Strittmatter, S.M. Small proline-rich repeat protein 1A is expressed by axotomized neurons and promotes axonal outgrowth. *J. Neurosci.* **2002**, *22* (4), 1303–1315.
69. Fournier, A.E.; Takizawa, B.T.; Strittmatter, S.M. Rho kinase inhibition enhances axonal regeneration in the injured CNS. *J. Neurosci.* **2003**, *23* (4), 1416.



Neuroprostheses

Milos R. Popovic

Inst. Biomaterials and Biomedical Eng., University of Toronto, Toronto, Ontario, Canada

T. Adam Thrasher

Lyndhurst Centre—Research, Toronto Rehabilitation Institute, Toronto, Ontario, Canada

INTRODUCTION

A neuroprosthesis, sometimes called a neural prosthesis, is a device that provides short bursts of electrical impulses to the central or peripheral nervous system to produce sensory and/or motor functions. Over the past four decades, neuroprostheses have been developed for a wide variety of applications. Some have achieved great success, such as the cochlear implants and bladder management stimulators that are produced in large volume worldwide. Other neuroprostheses, such as those for walking and grasping, have not yet matured to a level that creates a significant consumer demand. There are far too many neuroprostheses on the market and in development to list comprehensively, so this review involves the general features, principles, and functions of some of the most notable devices past and present.

PHYSIOLOGICAL OVERVIEW

In nerve cells, information is coded and transmitted as a series of electrical impulses called action potentials, which represent a brief change in cell electric potential of approximately 80 mV. Nerve signals are frequency modulated; that is, the number of action potentials that occur in a unit of time is proportional to the intensity of the transmitted signal. Typical action potential frequency is between 4 and 12 Hz. An action potential can be elicited artificially by changing the electric potential of a nerve cell or a nerve axon by inducing electrical charge into the cell (Fig. 1). This process, when used to produce action potentials in motor neurons to generate body function, is termed functional electrical stimulation (FES).

Where sufficient electrical charge is provided to a nerve cell, a localized depolarization of the cell wall occurs, resulting in an action potential that propagates toward the end of the axon (orthodromic propagation). Concurrently, an action potential will propagate backward towards the cell body (antidromic propagation). Typically, FES is concerned with orthodromic impulses, using them to generate muscle contractions by stimulating motor nerve axons that

can produce desirable body functions. Until recently, antidromic impulses were considered a useless side effect of FES, but there is new interest in the potential role of antidromic impulses in neural rehabilitation.^[1]

Since generation of action potentials and their propagation occur in the axons, the motor nerves of the stimulated muscles must be intact. If peripheral axons are missing (if they have been cut or have degenerated, for example), the muscle becomes denervated and therefore highly resistant to electrical stimulation. However, contractions can be elicited from denervated muscles by applying extremely intense electrical fields across the muscle fibers, as demonstrated by researchers at the University of Vienna.^[2]

Nerves can be stimulated using monophasic or biphasic current or voltage pulses. The monophasic pulses are seldom used because they lead to unbalanced charge delivery to the tissues, potentially causing damage due to galvanic processes. Most modern FES systems implement biphasic current or voltage pulses, or so-called monophasic compensated pulse shapes.

Another way to activate muscles is to stimulate ascending axons of sensory neurons that trigger reflex arcs. The case where electrical stimulation is used to stimulate sensory neurons and thus alter reflexes or central nervous system functions is commonly described by the term neuromodulation.

TECHNOLOGY

Neuroprostheses come in many different shapes and sizes and serve many different purposes. The common components in all neuroprostheses are: 1) a power source; 2) a stimulus generator; 3) a user-control interface; and 4) electrodes. Most modern neuroprostheses use batteries, disposable or rechargeable, as a power source. Some still use external AC power. Stimulus generators have been miniaturized dramatically over the years. Nowadays, commercial and laboratory-class stimulators tend to be lightweight (less than 1 kg) and handheld. User-control interfaces usually consist of a simple control panel with standard manual controls such as switches, buttons, dials, and

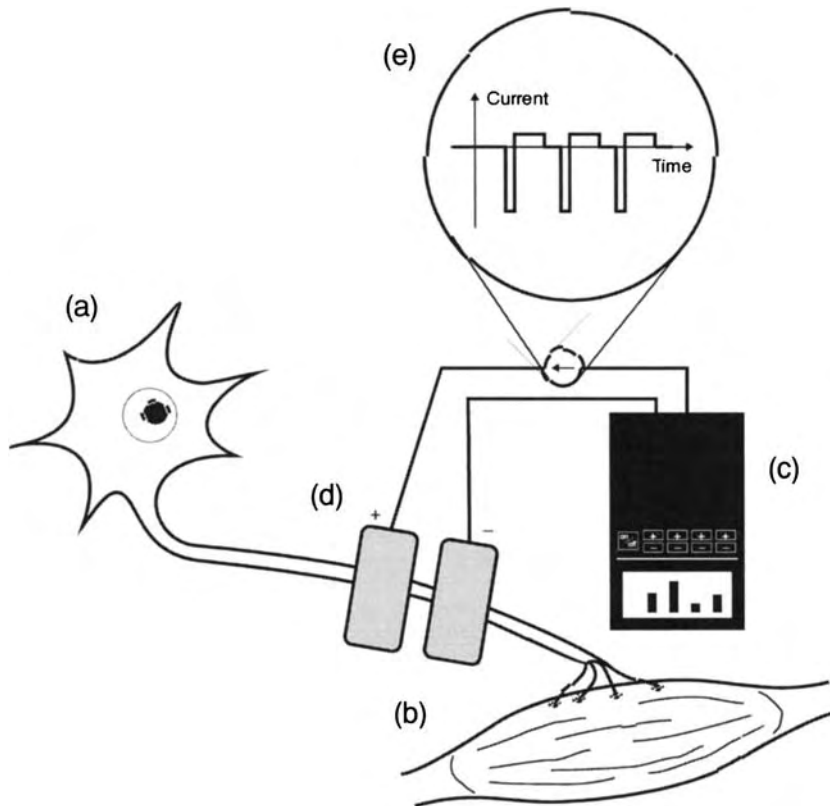


Fig. 1 Illustration of direct stimulation of a motor neuron. The cell body (a) is responsible for synthesizing input from dendrites and deciding whether or not to generate signals, which are transmitted to the corresponding muscle fibers (b). Following a stroke or spinal cord injury, muscles are impaired because motor neurons no longer receive sufficient input from the central nervous system. A neuroprosthesis (c) injects electrical current into the cell axon (d). A train of negative pulses (e) produces a series of action potentials. Depolarization occurs where negative current enters the axon at the active electrode indicated.

sliders, plus some kind of visual output such as light emitting diodes or, on more sophisticated models, a liquid crystal display. In addition, input devices are often mounted on the user's assistive devices, such as a pushbutton attached to a cane or walker. In some cases, input devices are attached to the user's clothing or body, such as an inclinometer on the shank of the leg or pressure sensors in the insole of a shoe. The most sophisticated neuroprostheses use real-time feedback control, which requires sensors such as goniometers, accelerometers, or gyroscopes to provide continuous-state feedback.

Nerves can be stimulated using either surface (transcutaneous), percutaneous, or implanted electrodes. Surface electrodes contact the skin (Fig. 2). They are noninvasive, easy to apply, and generally inexpensive. However, due to the impedance of the skin and the dispersion of current, much higher-intensity signals are required than with subcutaneous electrodes. Current amplitude typically ranges from 10–150 mA in surface stimulation. A major limitation is that some nerves, for example, those innervating the hip flexors, are too profound to be stimulated by surface electrodes. Percutaneous electrodes consist of thin wires that are inserted through the skin and into muscular tissue, remaining in place for a temporary period of time. In percutaneous stimulation, current amplitude is rarely higher than 25 mA. The third class of electrodes is implanted electrodes, which are

permanently implanted. Compared to surface electrodes, implanted and percutaneous electrodes potentially have higher stimulation selectivity with much less electrical charge applied, both of which are desired characteristics of FES systems. The drawbacks are that implants require a lengthy, invasive surgical process to install and that percutaneous electrodes can be used only temporarily and may cause infection at the penetration site.

There is a brand of miniature electrode, the BION™, that can be implanted via hypodermic needle.^[3] They are cylindrical in shape, with a diameter of 2 mm and a length of 15 mm. Once implanted, they are powered

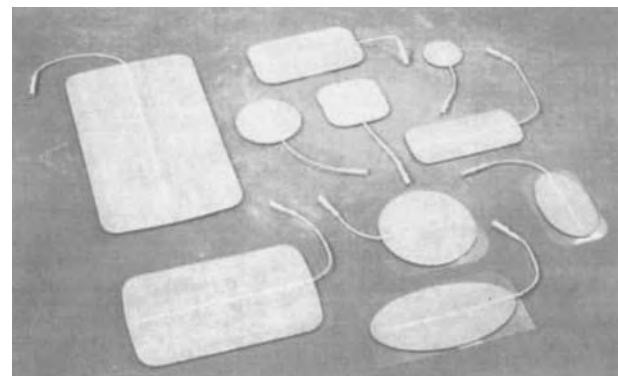


Fig. 2 Reusable, self-adhesive electrodes for surface stimulation come in a variety of shapes and sizes.

and controlled via radio waves from an external controller that can be worn by the patient.

HISTORICAL NOTES

In 46 A.D., Scribonius Largus described what may be considered the first neuroprosthesis: the torpedo ray, which is capable of generating an electric potential of 25 to 30 V.^[4] For centuries, torpedoes were prescribed for all sorts of ailments, including headaches, hemorrhoids, and even mental illness.

Following the invention of the electrostatic generator in the late 17th century, electrical discharges were found to excite animal muscles. Stronger discharges, and hence stronger biological responses, were made possible when the first capacitor was invented in 1745. Physicians began treating a wide range of diseases by applying electrical discharges to their patients. Benjamin Franklin pioneered some of these techniques.

In 1791, Luigi Galvani published his discovery that dissected frog legs could be stimulated by touching a bimetallic rod to nerve and muscle. Michael Faraday built the first electric generator in 1831. It introduced the possibility of applying a series of high-frequency electrical pulses to nerves, which is the basis for all modern electrical stimulation. G. B. Duchenne utilized Faradism extensively in the latter part of the nineteenth century to treat various neurological disorders. Duchenne developed electrodes for localizing currents, and he produced a set of maps of the body indicating locations called motor points, where electrodes can be positioned to excite specific muscles.

NEUROPROSTHESES FOR WALKING

There are many neuroprostheses that address lower-extremity movement. As early as 1960, Kantrowitz demonstrated paraplegic standing by applying continuous surface FES to the quadriceps and gluteus maximus muscles of a patient with complete spinal cord injury.^[5] Around the same time, Liberson and colleagues developed a simple neuroprosthesis to correct drop foot. This common symptom in hemiplegia is characterized by a lack of dorsiflexion during the swing phase of gait, resulting in short, shuffling strides.^[6] Liberson's device, which has the distinction of being the first neuroprosthesis to receive a patent, consisted of a power supply worn on a belt, two surface electrodes positioned for stimulation of the common peroneal nerve, and a heel switch. The stimulation was activated whenever the heel lost contact with the ground, and was deactivated when the heel regained contact with the ground.

Stimulation of the common peroneal nerve causes contraction of the muscles responsible for dorsiflexion (i.e., tibialis anterior and extensor hallucis longus, among others). It can also trigger the flexor withdrawal reflex, which may not be desirable. The flexor withdrawal reflex occurs naturally when a sudden, painful sensation is applied to the sole of the foot. It results in flexion of the hip, knee, and ankle of the affected leg and extension of the contralateral leg in order to get the foot away from the painful stimulus as quickly as possible. To prevent this from happening during FES-assisted ambulation, Vodovnik proposed using a low-pass filter to slow the onset of stimulation current.^[7]

Following Liberson's invention and Vodovnik's revisions, a number of drop foot stimulators were developed. Some were commercialized, for example the MikroFES (Josef Stefan Institute, Ljubljana, Slovenia) and the Odstock Dropped Foot Stimulator.^[8] The latter was shown to significantly increase walking speed and efficiency, and a carryover effect was observed in stroke patients; that is, their walking speed and efficiency without the stimulator were improved.^[9] Similar studies have reported no carryover effect.^[10] Users of the Odstock device were generally satisfied with it, but almost all of them identified the surface electrodes as problematic, and two thirds would consider an implanted system instead.^[11]

The first commercially available implanted drop foot stimulator was developed by Rancho Los Amigos Medical Centre and Medtronic Inc.^[12] The surgically implanted compounds were a radio-frequency (RF) receiver, a pulse train generator, and one bipolar electrode implanted adjacent to the peroneal nerve. An external unit worn on the belt delivered power via the RF coil and received input commands from a wireless foot switch. Despite some problems with electrode migration and infection, the device was considered successful. Since then, more reliable and easier to implant systems such as the IPPO^[13] and the Aalborg University implanted stimulator^[14] have been devised, but they are not commercially available. The latter uses input from an implanted cuff electrode around the sural nerve, which is the nerve innervating the skin sensors on the sole of the foot. This system is unique in that it requires no external sensors.

Most modern drop foot stimulators continue to use a heel switch for active input. Burrige et al. tried using the foot switch on the nonaffected leg, but found it was not preferable unless the patient was unable to reliably achieve heel contact on the affected leg.^[15] Vodovnik was one of the first to experiment with manual push-buttons and EMG sensors.^[7] Other alternatives to the heel switch include a heel/toe switch,^[16] an array of four single-axis accelerometers positioned on the shank,^[17] a tilt sensor positioned on the shank,^[18]

electroneurography,^[14] a knee goniometer,^[19] and the Gait Phase Detection System.^[20]

The earliest neuroprostheses for paraplegic gait provided continuous stimulation to the quadriceps to produce a mode of gait similar to long leg-brace walking. Later systems used alternating bilateral quad/glut stimulation (during stance phase) out of phase with peroneal nerve stimulation (during swing phase). One such system was a six-channel stimulator developed at the University of Ljubljana in Slovenia.^[16] Later at the same institution, Kralj and colleagues described a technique for paraplegic gait using surface stimulation, which remains the most popular method today.^[21] According to Kralj's technique, four channels of stimulation are used. Electrodes are placed over the quadriceps muscles and peroneal nerves bilaterally. The user controls the neuroprosthesis with two pushbuttons attached to the left and right handles of a walking frame, or on canes or crutches (Fig. 3). When the neuroprosthesis is turned on, both quadriceps are stimulated. The left button initiates swing phase in the left leg by briefly stopping stimulation of the left quadriceps and stimulating the peroneal nerve. This stimulation is applied suddenly so as to trigger the flexor withdrawal reflex, resulting in simultaneous hip and knee flexion as well as dorsiflexion. After a fixed period of time, peroneal nerve stimulation is stopped and quadriceps stimulation is resumed. Similarly, the right button initiates swing phase in the right leg. Kralj and colleagues successfully applied this system to more than 50 subjects with spinal cord injury.

Many neuroprostheses for walking have employed the basic technique described in this section. As

microprocessor technology developed, neuroprostheses became more portable and flexible. The Parastep system uses Kralj's technique.^[22,23] It is the only neuroprosthesis for walking to receive approval from the United States Food and Drug Administration (FDA) and the first neuroprosthesis of any kind to receive FDA approval. It includes an ankle-foot orthosis to bolster ankle stiffness. The Parastep is commercially available, and more than 600 people have used it successfully.

A major limitation of neuroprostheses for walking that are based on surface stimulation is that the gait is slow, awkward, and unnatural looking. Perhaps a major reason for this is that the hip flexors cannot be stimulated directly. Therefore, hip flexion during walking must come from voluntary effort, which is often absent in paraplegia, or from the flexor withdrawal reflex (initiated by peroneal nerve stimulation). Implanted systems have the advantage of being able to stimulate the hip flexors. They also provide better muscle selectivity and more natural gait patterns. Two such systems are the Praxis24 and the system proposed by Kobetic, which use 24 and 32 electrodes respectively.^[24,25] The Praxis24 system also enables bladder voiding.

NEUROPROSTHESES FOR REACHING AND GRASPING

A number of neuroprostheses have been developed and used to assist stroke and spinal cord-injured subjects to improve their grasping function. The best-known grasping neuroprostheses are the Freehand system,^[26]

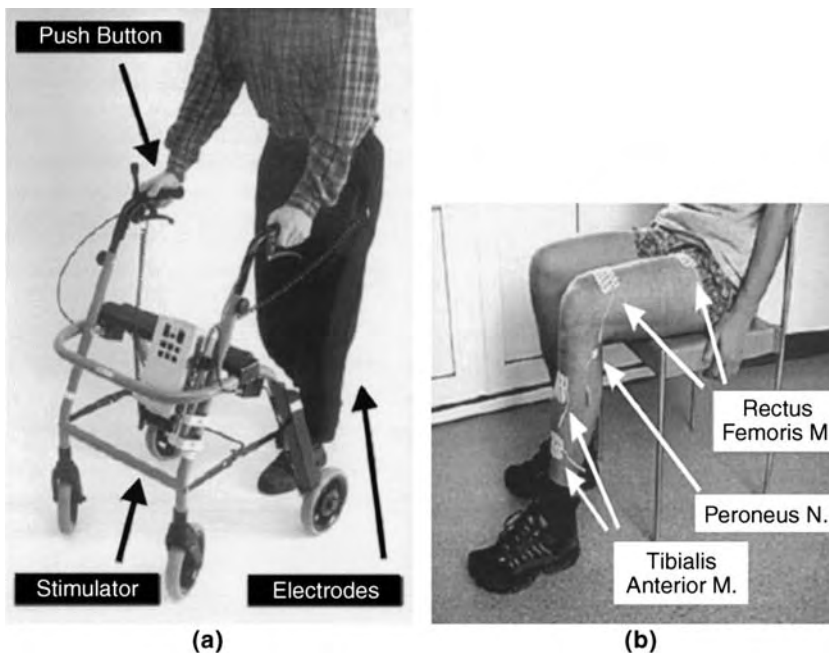


Fig. 3 The ETHZ-ParaCare walking neuroprosthesis for hemiplegic subjects and subjects with unilateral paraplegia, developed in Zurich, Switzerland. (From Ref. [38].) (a) Stimulator and push button attached to walker; (b) surface electrodes attached to legs underneath clothing.

the Handmaster NMS-1,^[27] the Bionic Glove,^[28] the NEC FESMate system,^[29] the Compex Motion neuroprosthesis for grasping,^[30,31] and the systems developed by Rebersek and Vodovnik^[32] and Popovic et al.^[33] With the exception of the Freehand and NEC FESMate systems, all use surface stimulation.

The key element for achieving the synergistic activity of muscles that results in reaching and grasping is the appropriate sequencing of electrical pulses. The available neuroprostheses for grasping can restore the two most frequently used grasping styles: the palmar and the lateral grasp. The palmar grasp is used to hold bigger and heavier objects such as cans and bottles, and the lateral grasp is used to hold smaller and thinner objects such as keys, paper, and compact discs. The lateral grasp is generated by first flexing the fingers to provide opposition, which is followed by the thumb flexion. The palmar grasp is generated by first forming opposition between the thumb and the palm, which is followed by simultaneous flexion of both the thumb and the fingers.

The Freehand system, manufactured and distributed by NeuroControl Co., U.S.A.,^[34] consists of eight implanted epimysial stimulation electrodes that stimulate flexion and extension of the fingers and the thumb in order to provide lateral and palmar grasp. Commands are given by an external position sensor that is placed on the shoulder of the subject's opposite arm. An additional external switch allows the user to choose between palmar and lateral grasp. This sequence is then sent via a radio frequency coil to the implanted unit, which generates the stimulation sequences for each channel.

The electrode leads are tunneled subcutaneously to the implanted stimulator located in the pectoral region. Surgical procedures to enhance both voluntary and stimulated hand functions are often performed in conjunction with the stimulator implantation. More than 200 quadriplegic subjects have received the Freehand neuroprosthesis at more than a dozen sites around the world. The subjects have demonstrated the ability to grasp and release objects and to perform activities of daily living more independently when using the neuroprosthesis. The Freehand system is the first neuroprosthesis for grasping approved by the FDA. The main advantage of the Freehand system is that it is implanted, and the time needed to don and doff the system is shorter compared to most of the surface FES systems.

In the 1980s, the group led by Handa at Sendai University, Japan, developed a microcomputer-controlled neuroprosthesis for grasping. Soon after that, Handa's team proposed a system with 16 percutaneous intramuscular stimulation electrodes that is both portable and programmable.^[29] This system consists of a NEC PC-98LT personal computer and an

external microcontroller-based stimulator. The stimulator applies trapezoidal stimulation patterns to generate muscle contractions. The stimulation patterns were "cloned" from the muscle activity recorded during voluntary grasping movements of able-bodied subjects. Stimulation sequences were triggered with a pushbutton or a pneumatic pressure sensor. This system demonstrated that spinal cord injured subjects with complete C4 to C6 spinal cord lesions could reach and grasp. In collaboration with NEC Inc., the Sendai team developed a fully implantable 16-channel electric stimulator called the NEC FESMate. Although 200 of these stimulators have been manufactured,^[35] the NEC FESMate is not available outside of Japan.

The neuroprosthesis developed by Rebersek and Vodovnik was one of the first FES systems for grasping.^[32] This neuroprosthesis has three stimulation channels, which are used to generate grasping by stimulating the finger flexors and extensors and the thumb flexors. Although this device was developed almost three decades ago, it is one of the rare systems that allows a subject to control the stimulator via different sensory interfaces such as an EMG sensor, a sliding resistor, or a pressure sensor. This option is important because it allows the neuroprosthesis to be tailored to the subject. The main disadvantages of this neuroprosthesis are that donning and doffing times are long and that selectivity of stimulation is quite limited. Merelitti et al. modified this system and used it for stroke subjects.^[36] They applied two channels to augment the elbow and fingers/wrist extension. They concluded that FES contributed greatly to recovery of hand and elbow movements in five stroke subjects, but in the remaining three the improvement was significant only at the elbow joint.

The Handmaster (Fig. 4) is a neuroprosthesis for grasping that is manufactured and distributed by NESS Ltd., located in Israel.^[27] It consists of an orthosis that has built-in flexibility to enhance and control freedom of movement within the forearm and hand, while supporting the wrist joint at a functional angle of extension. The Handmaster multiplexes a single channel of stimulation through a selected combination of surface electrodes on the inner surface of the orthosis, which effectively transforms the device into a six-channel neuroprosthesis. One stimulation channel is used to stimulate the extensor digitorum communis at the dorsal side of the forearm. The second channel stimulates the flexor digitorum superficialis. Electrodes are positioned over the muscles of the forearm and hand intrinsics during an initial setup session with a clinician. The setup position of the electrodes depends on the device user's specific needs. The Handmaster is controlled with an array of push buttons allowing the subject to select the operating mode and to trigger programmed movement

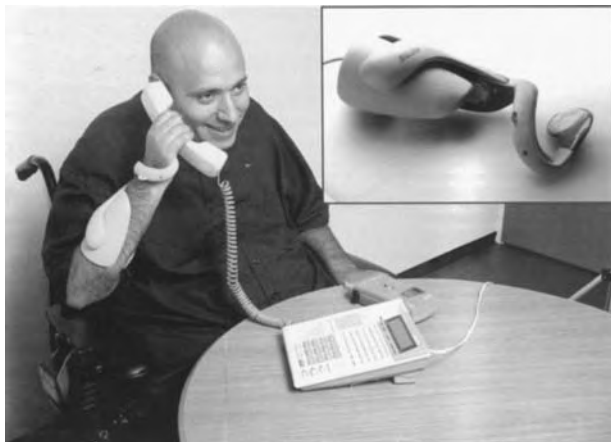


Fig. 4 The Handmaster system manufactured and distributed by NESS Ltd., Israel. (From Ref. [27].) The flexible, adjustable unit is worn over the forearm, wrist, and palm.

sequences. Using the buttons, the subject can also control stimulation intensity and thumb posture, thereby adjusting the grasp to the size and the shape of the target object. Originally, the Handmaster was envisioned as a permanent orthotic system, but it is also used as an exercise and rehabilitation tool. One of the major advantages of the Handmaster is that it is easy to don and doff. It is exceptionally well designed and is one of the best neuroprostheses for grasping on the market. There are currently more than 2000 in use.

The Bionic Glove is a neuroprosthesis designed to enhance the tenodesis grasp in subjects who have good voluntary control over wrist flexion and extension.^[28] By extending their wrist, users can cause passive finger flexion due to the limited length of the finger flexors. The Bionic Glove stimulates finger flexors and extensors during tenodesis grasp, significantly enhancing the strength of the grasp. Four self-adhesive surface stimulation electrodes provide stimulation, and the stimulator and a wrist position sensor are located on the forearm of the glove. An easy-to-use interface with three push buttons on the stimulator is used to set the stimulation parameters, and the optional audio feedback facilitates faster learning. Clinical evaluation of the Bionic Glove has indicated that it is generally beneficial to quadriplegic subjects, but only about 30% of potential users accepted it for long-term use.^[37] It is available only for clinical evaluation from the University of Alberta, Canada, and it is presently being modified into a new system called the Tetron that will provide several grasping patterns and strategies.

The Belgrade Grasping-Reaching System proposed by Popovic et al. is a neuroprosthesis for grasping that also provides reaching function.^[33] It consists of four stimulation channels, three of which are used to generate grasping function. The fourth channel is used

to stimulate the triceps brachii muscle augmenting elbow extension. Reaching is achieved by measuring the subject's shoulder velocity with a goniometer and by generating a synergistic elbow motion by stimulating the triceps brachii muscle. This neuroprosthesis, similar to the system proposed by Rebersek and Vodovnik, requires more time to don than the Handmaster and is not yet commercially available.

The Complex Motion neuroprosthesis is a very flexible device designed to improve grasping and walking functions in both spinal cord injury and stroke patients.^[38] This multichannel surface stimulation system is programmable and can be interfaced with any sensor system. As a four-channel neuroprosthesis for grasping, it provides both palmar and lateral grasps. It can be controlled with proportional EMG, discrete EMG, pushbuttons, or sliding resistor control strategies. Thus far, more than 50 stroke and spinal cord-injured patients have used the neuroprosthesis in a clinical setting or at home in activities of daily living. One of the main disadvantages of this system is that it requires about eight-minutes to put on or take off.

NEUROPROSTHESES FOR BLADDER MANAGEMENT

Neuroprostheses have been very successful in treating lower urinary tract dysfunctions commonly associated with spinal cord injury, such as urge incontinence and urinary retention. The first attempts to electrically stimulate the bladder were made in the 1950s, when researchers sought ways to induce bladder emptying. At that time, a bladder wall stimulator was developed and implanted in three humans,^[39] and animal studies of pelvic nerve stimulation were carried out.^[40] Later, it was found that electrical stimulation of the sacral anterior roots produces excellent results, and this led to the development of the Finetech-Brindley stimulator, which is the most widely used neuroprosthesis for bladder management today.^[41]

Attempts to manage incontinence using electrical stimulation began in the 1960s.^[42] It was found that urethral resistance could be increased by stimulating the muscles of the pelvic floor, vagina, and rectum using external electrodes.^[43] Eventually, fully implanted systems were developed to suppress the detrusor muscle, thus preventing reflex incontinence and increasing bladder volume.^[44] Most spinal cord injuries result in reflex incontinence. Typically, detrusor-sphincter dyssynergia develops, in which the detrusor and urethral sphincter contract simultaneously rather than reciprocally. The detrusor also becomes hyper-reflexic, and the bladder becomes overactive. The standard treatments are anticholinergic

medication, which blocks the neuromuscular junctions, and sensory rhizotomy (surgical transection of the posterior sacral roots). Neuroprostheses for bladder management serve as a practical alternative to these treatments. They can also augment sensory rhizotomy.

The Finetech-Brindley stimulator has been implanted in more than 2000 patients, usually those who have had a rhizotomy.^[45] The electrodes are positioned on the second, third, and fourth sacral roots, bilaterally and extradurally. If a rhizotomy has not been performed, the electrodes must be implanted inside the dura to prevent crossover stimulation of the sensor neurons, which will trigger the detrusor reflex. A portable external controller transmits power to the implant via radio frequency coil, and the user initiates bladder voiding by pushing buttons on the external unit. Micturition is usually achieved with residual volumes of less than 50 mL, contributing to a dramatic reduction in urinary tract infections.^[46] The Finetech-Brindley stimulator has proven to be extremely robust, with only one failure expected every 80 implant-years.^[47]

The Medtronic Interstim stimulator is a sacral root implant for incontinence, using neuromodulation to correct the inappropriate reflex behaviour.^[48] It consists of fine wire electrodes inserted into the sacral foramina. When active, these electrodes inhibit the detrusor, but the mechanism of this inhibition is not yet properly understood. Thorough testing must be done using a temporary implant before permanent implantation is recommended. The stimulation parameters commonly used are a pulsewidth of 60–270 μ s and a frequency of 10–15 Hz, with the stimulation on for 10 s, then off for 2 s. Current amplitude is twice the sensory threshold. The clinical success rate of this device is about 50%. Bladder emptying has to be achieved either voluntarily or by means of intermittent catheterization.

COCHLEAR IMPLANTS

Cochlear implants are neuroprostheses for the hearing impaired who have severe (70 to 90 dB) or profound (>90 dB) hearing loss. A long wire electrode is implanted directly into the cochlear duct, and electrical stimulation is applied to the residual spiral ganglion cells of the cochlear nerve. These devices were first developed in France in 1957. Since then, cochlear implants have been refined and miniaturized, and now they have received widespread acceptance, more so than any other class of neuroprostheses. More than 75,000 patients have received cochlear implants worldwide. Originally, few hearing-impaired people were eligible for cochlear implantation, but as the technology has improved, the selection criteria have

expanded greatly to include a wide range of hearing impairments.^[49]

Due to the success and popularity of cochlear implants, there are many different brands on the market. Most brands, however, are essentially similar. Differences between the currently available cochlear implants mainly involve the number of electrode channels (12 to 22), speech coding strategies, and the mode of electrode stimulation.^[50] A recent study carried out at the University of Toronto, for example, concluded that the Clarion CI (Advanced Bionics, Symlar, CA) and the Nucleus 22 (Cochlear Corp., Sydney, Australia) cochlear implants were totally comparable in function and performance.^[51] Both devices succeeded in reducing tinnitus, thereby increasing word and sentence recognition, but there was no significant reduction in vestibular function. Among the implantees, 76% reported that they were satisfied with their implants, and 96% reported an overall positive impact on quality of life. Some other brands of cochlear implant are the COMBO 40+ system (MED-EL, Durham, NC), Digisonic (MXM, France), and the SOUNDTEC direct system (SOUNDTEC, Palo Alto, CA), most of which are FDA approved.

Cochlear implants generally consists of the following: 1) an external earpiece; 2) a speech processor; and 3) an internally implanted unit (Fig. 5). The earpiece, usually very small and lightweight, is worn comfortably behind the ear, much like a hearing aid. It

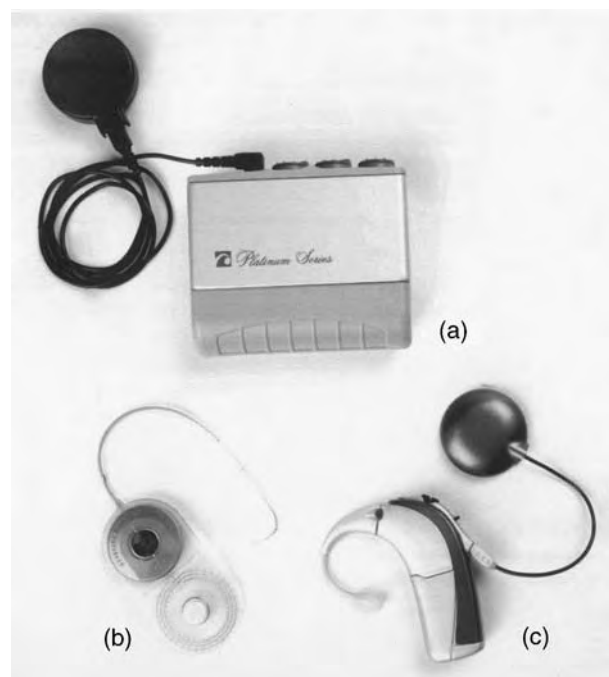


Fig. 5 The Clarion CII Bionic Ear system by Advanced Bionics (Symlar, California). (a) Sound processing unit; (b) internal, implantable unit; (c) external earpiece.

contains an ear-level or in-ear microphone and a radio frequency coil to transmit signals to the implanted components. The speech processor, can be in the form of a small box worn somewhere on the body or, in some models, it is contained in the external unit worn behind the ear. The internally implanted unit consists of a receiver coil located underneath soft tissue in a cavity drilled in the temporal squama, and a 20–24 mm insulated wire ending in a multichannel electrode array, which is inserted into the cochlear duct. Sound waves are received by the external microphone and converted into electrical signals that are input to the external speech processor. There, the signals are digitally encoded and transmitted to the internal unit via radio frequency coil. The internal unit decodes the radio signals back into elementary electrical signals to stimulate each channel of the electrode array. Therefore, the multichannel device provides a complex sound analysis similar to the physiological analysis of sound in normal patients.

CONCLUSIONS

The first modern FES devices were developed over 40 years ago, and since then there has been a great deal of innovation resulting in scores of new neuroprostheses. The most successful of these technologies, in terms of consumer acceptance, are cochlear implants—more than 75,000 units have been implanted worldwide. Bladder management stimulators have also achieved wide acceptance, but not quite to the same degree.

Despite many advances and positive reports over the decades, neuroprostheses for walking and grasping have not achieved widespread approval. Most are used only for research purposes, and few have been used regularly by patients for activities of daily living. Some have been put to clinical use, and those are usually abandoned after a short period of time. The general perception among clinicians is that the neuroprostheses for grasping and walking are not fully matured and their application is often labor-intensive, whereas a favorable outcome cannot be guaranteed. Nevertheless, recent studies indicate that these tools have great potential in the rehabilitation of stroke and spinal cord-injured subjects.^[30,3152,53] In particular, these studies indicate that a significant number of patients who were trained with these systems recover voluntary reaching, grasping, and walking functions due to intensive and repetitive training with these systems. Current efforts are focused on understanding the mechanism of short- and long-term improvements and recoveries observed in these patients.

Neuroprostheses are a new and emerging technology that has significant potential. However, implementation

of this technology to its full potential presents numerous challenges that have yet to be addressed. We believe that the 21st century will be the century when most of these technical and implementation issues are resolved and neuroprostheses are established as one of the important classes of rehabilitation systems available to patients with disabilities ranging from spinal cord injury to blindness.

ARTICLES OF FURTHER INTEREST

Artificial Muscles; Excitable Tissue, Electrical Stimulation of; Nerve Guides; Tissue Engineering of Peripheral Nerve

REFERENCES

1. Rushton, D.N. Functional electrical stimulation and rehabilitation—An hypothesis. *Med. Eng. Phys.* **2003**, *25* (1), 75–78.
2. Reichel, M.; Breyer, T.; Mayr, W.; Rattay, F. Simulation of the three-dimensional electrical field in the course of functional electrical stimulation. *Artif. Organs* **2002**, *26* (3), 252–255.
3. Loeb, G.E. Presentation highlights: Bionic neurons (BIONs™). *J. Rehabil. Res. Dev.* **2002**, *39* (Supplement 3), 5–6.
4. Schechter, D.S. Origins of electrotherapy. *N. Y. State J. Med.* **1971**, *71*, 997–1008.
5. Kantrowitz, A. A Report of the Maimonides Hospital. In *Electronic Physiologic Aids*; Brooklyn: New York, 1960.
6. Liberson, W.T.; Holmquest, H.J.; Scot, D., Dow, M. Functional electrotherapy: Stimulation of the peroneal nerve synchronized with the swing phase of the gait of hemiplegic patients. *Arch. Phys. Med. Rehabil.* **1961**, *42*, 101–105.
7. Vodovnik, L.; Long, C.; Reswick, J.B.; Lippay, A.; Starbuck, D. Myo-electric control of paralyzed muscles. *IEEE Trans. Biomed. Eng.* **1965**, *12* (3), 169–172.
8. Taylor, P.N.; Burridge, J.H.; Dunkerley, A.L.; Wood, D.E.; Norton, J.A.; Singleton, C.; Swain, I.D. Clinical use of the Odstock Dropped Foot Stimulator: Its effect on speed and effort of walking. *Arch. Phys. Med. Rehabil.* **1999**, *80* (12), 1577–1583.
9. Taylor, P.N.; Burridge, J.H.; Dunkerley, A.L.; Lamb, A.; Wood, D.E.; Norton, J.A.; Swain, I.D. Patients' perceptions of the Odstock Dropped Foot Stimulator. *Clin. Rehabil.* **1999**, *13* (5), 439–446.
10. Granat, M.H.; Maxwell, D.J.; Ferguson, A.C.; Lees, K.R.; Barbenel, J.C. Peroneal stimulator; evaluation for the correction of spastic drop foot in hemiplegia. *Arch. Phys. Med. Rehabil.* **1996**, *77* (1), 19–24.
11. Taylor, P.N.; Mann, G.; Wood, D.E.; Hobby, J. Pilot Study to Evaluate the Safety and Efficacy of an Implanted Dropped Foot Stimulator, From Technology to Market: Bridging the Gap Proceedings of the International Functional Electrical Stimulation Society, Maroochydore, Australia, July 1–5, 2003.



12. Waters, R.L.; McNeal, D.; Perry, J. Experimental correction of foot-drop by electrical stimulation of the peroneal nerve. *J. Bone Jt. Surg., Am.* **1975**, *57-A*, 1047–1054.
13. Strojnik, P.; Acimovic, R.; Vavken, E.; Simic, V.; Stanic, U. Treatment of drop foot using an implantable peroneal underknee stimulator. *Scand. J. Rehab. Med.* **1987**, *19*, 37–43.
14. Haugland, M.K.; Sinkjaar, T. Cutaneous whole nerve recordings used for correction for footdrop in hemiplegic man. *IEEE Trans. Biomed. Eng.* **1995**, *3*, 307–317.
15. Burrige, J.H.; Taylor, P.N.; Hagan, S.A.; Wood, D.E.; Swain, I.D. The effects of common peroneal stimulation on the effort and speed of walking: A randomized controlled trial with chronic hemiplegic patients. *Clin. Rehabil.* **1997**, *11* (3), 201–210.
16. Strojnik, P.; Kralj, A.; Ursic, I. Programmed six-channel electrical stimulator for complex stimulation of leg muscles during walking. *IEEE Trans. Biomed. Eng.* **1979**, *26*, 112–116.
17. Willemsen, A.T.M.; Bloemhof, F.; Boom, H.B.K. Automatic stance-swing phase detection from accelerometer data for peroneal nerve stimulation. *IEEE Trans. Biomed. Eng.* **1990**, *37*, 1201–1208.
18. Dai, R.; Stein, R.B.; Andrews, B.J.; James, K.B.; Wieler, M. Application of tilt sensors in functional electrical stimulation. *IEEE Trans. Rehabil. Eng.* **1996**, *4* (2), 63–72.
19. Sweeney, P.C.; Lyons, G.M. Fuzzy gait event detection in a finite state controlled FES drop foot correction system. *J. Bone Jt. Surg., Br.* **1999**, *81-B* (Suppl. I), 93.
20. Pappas, I.P.; Popovic, M.R.; Keller, T.; Dietz, V.; Morari, M. A reliable gait phase detection system. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2001**, *9* (2), 113–125.
21. Kralj, A.; Bajd, T.; Turk, R. Enhancement of gait restoration in spinal injured patients by functional electrical stimulation. *Clin. Orthop. Relat. Res.* **1988**, *223*, 34–43.
22. Graupe, D.; Davis, R.; Kordylewski, H.; Kohn, K.H. Ambulation by traumatic T4-12 paraplegics using functional neuromuscular stimulation. *Crit. Rev. Neurosurg.* **1998**, *8*, 221–231.
23. Graupe, D.; Kohn, K.H. Functional neuromuscular stimulator for short-distance ambulation by certain thoracic-level spinal-cord-injured paraplegics. *Surg. Neurol.* **1998**, *50*, 202–207.
24. Davis, R.; Houdayer, T.; Andrews, B.J.; Barriskill, A. In *Paraplegia: Implanted Praxis24±FES System for Multi-Functional Restoration*, Proceedings of the International Functional Electrical Stimulation Society, Sendai, Japan, 1999.
25. Kobetic, R.; Triolo, R.; Morsolais, E. Muscle selection and walking performance of multichannel FES systems for ambulation in paraplegia. *IEEE Trans. Rehabil. Eng.* **1997**, *5*, 23–28.
26. Smith, B.; Peckham, P.H.; Keith, M.; Roscoe, D. An externally powered, multichannel, implantable stimulator for versatile control of paralyzed muscle. *IEEE Trans. Biomech. Eng.* **1987**, *34* (7), 499–508.
27. Ijzerman, M.; Stoffers, T.; 't Groen, F.; Klatte, M.; Snoek, G.; Vorsteveld, J.; Nathan, R.; Hermens, H. The NESS handmaster orthosis: Restoration of hand function in C5 and stroke patients by means of electrical stimulation. *J. Rehabil. Sci.* **1996**, *9* (3), 86–89.
28. Prochazka, A.; Gauthier, M.; Wieler, M.; Kenwell, Z. The Bionic Glove: An electrical stimulator garment that provides controlled grasp and hand opening in quadriplegia. *Arch. Phys. Med. Rehabil.* **1997**, *78*, 1–7.
29. Hoshimiya, N.; Handa, Y. A master-slave type multi-channel functional electrical stimulation (FES) system for the control of the paralyzed upper extremities. *Automedica* **1989**, *11*, 209–220.
30. Popovic, M.R.; Hajek, V.; Takaki, J.; Bulsen, A.; Zivanovic, V. In *Restoration of Reaching and Grasping Functions in Hemiplegic Patients with Severe Arm Paralysis*, From Technology to Market: Bridging the Gap, Proceedings of the International Functional Electrical Stimulation Society, Maroochydore, Australia, July 1–5, 2003.
31. Adams, M.; Takes, V.; Popovic, M.R.; Bulsen, A.; Zivanovic, V. Restoration of Grasping Functions in Patients with Quadriplegia. From Technology to Market: Bridging the Gap Proceedings of the International Functional Electrical Stimulation Society, Maroochydore, Australia, July 1–5, 2003.
32. Rebersek, S.; Vodovnik, L. Proportionally controlled functional electrical stimulation of hand. *Arch. Phys. Med. Rehabil.* **1973**, *54*, 168–172.
33. Popovic, D.; Popovic, M.; Stojanovic, A.; Pjanovic, A.; Radosavljevic, S.; Vulovic, D. In *Clinical Evaluation of the Belgrade Grasping System*, Proceedings of the 6th Vienna International Workshop of FES, Vienna, Austria, 1998.
34. Smith, B.; Peckham, P.; Keith, M.; Roscoe, D. An externally powered, multichannel, implantable stimulator for versatile control of paralyzed muscle. *IEEE Trans. Biomech. Eng.* **1987**, *34*, 499–508.
35. Takahashi, K.; Hoshimiya, N.; Matsuki, H.; Handa, Y. Externally powered implantable FES system. *Jpn. J. Med. Electron. Biol. Eng.* **1999**, *37*, 43–51.
36. Merletti, R.; Acimović, R.; Grobelnik, S.; Cvilak, G. Electrophysiological orthosis for the upper extremity in hemiplegia: Feasibility study. *Arch. Phys. Med. Rehabil.* **1975**, *56* (12), 507–513.
37. Popovic, D.; Stojanovic, A.; Pjanovic, A.; Radosavljevic, S.; Popovic, M.; Jovic, S.; Vulovic, D. Clinical evaluation of the Bionic Glove. *Arch. Phys. Med. Rehabil.* **1999**, *80* (3), 299–304.
38. Popovic, M.R.; Keller, T.; Pappas, I.P.I.; Dietz, V.; Morari, M. Surface-stimulation technology for grasping and walking neuroprostheses. *IEEE Trans. Eng. Med. Biol. Mag.* **2001**, *20*, 82–93.
39. Boyce, W.H.; Lathem, J.E.; Hunt, L.D. Research related to the development of an artificial electrical stimulator for the paralyzed human bladder: A review. *J. Urol.* **1964**, *91*, 41–51.
40. Bors, E.; Comarr, E. *Neurological Urology: Physiology of Micturition, Its Neurological Disorders and Sequelae*; Karger Verlag, 1971.

41. Brindley, G.S.; Polkey, C.E.; Rushton, D.N.; Cardozo, L. Sacral anterior root stimulators for bladder control in paraplegia: The first 50 cases. *J. Neurol. Neurosurg. Psychiatry* **1986**, *49*, 1104–1114.
42. Caldwell, K.P.S. The electrical control of sphincter incontinence (preliminary communication). *Lancet* **1963**, *2*, 174–175.
43. Alexander, S. Research and clinical experience in the treatment of the neurogenic bladder by electronic implant and prosthesis. *Paraplegia* **1968**, *6* (3), 183–193.
44. Rijkhoff, N.J.M.; Wijkstra, H.; van Kerrebroeck, P.E.V.; Debruyne, F.M.J. Urinary bladder control by electrical stimulation: Review of electrical stimulation techniques in spinal cord injury. *NeuroUrol. Urodyn.* **1997**, *16*, 39–53.
45. Schurch, B.; Rodic, B.; Jeanmonod, D. Posterior sacral rhizotomy and intradural anterior sacral root stimulation for treatment of the spastic bladder in spinal cord injured patients. *J. Urol.* **1997**, *157*, 610–614.
46. Creasey, G.H. Restoration of bladder, bowel, and sexual function. *Top Spinal Cord Inj., Rehabil.* **1999**, *5* (1), 21–32.
47. Vastenholt, J.M.; Ijzerman, M.J.; Buschman, H.P.J.; Snoek, G.J.; van der Aa, H.E. In *Seven Year Follow-Up of Sacral Anterior Root Stimulation and Sacral Posterior Root Rhizotomy for Bladder Control in Patients with a Spinal Cord Injury, Complications and Quality of Life*, Proceedings of the International Functional Electrical Stimulation Society, Cleveland, Ohio, 2001.
48. Siegel, S.W. Management of voiding dysfunction with an implantable neuroprosthesis. *Urol. Clin. North Am.* **1992**, *19* (1), 163–170.
49. Lenarz, T. Cochlear implants, selection criteria and shifting borders. *Acta-Rhino-Laryngol. Belg.* **1998**, *52* (3), 183–199.
50. Gstoettner, W.; Adunka, O.; Hamzavi, J.; Baumgartner, W.D. Rehabilitation of hearing-impaired patients with cochlear implants, a review. *Wien. Klin. Wochenschr.* **2000**, *112* (11), 464–472.
51. Higgins, K.M.; Chen, J.M.; Nedzelski, J.M.; Shipp, D.B.; McIlmoyl, L.D. A matched-pair comparison of two cochlear implant systems. *J. Otolaryngol.* **2002**, *31* (2), 97–105.
52. Popovic, M.B.; Popovic, D.B.; Sinkjaer, T.; Stefanovic, A.; Schwirtlich, L. Restitution of reaching and grasping promoted by functional electrical therapy. *Artif. Organs* **2002**, *26* (3), 271–275.
53. Thrasher, T.A.; Popovic, M.R. In *FES-Assisted Walking for Rehabilitation of Incomplete Spinal Cord Injury, From Technology to Market: Bridging the Gap*, Proceedings of the International Functional Electrical Stimulation Society, Maroochydore, Australia, July 1–5, 2003.



Neutrophil-Biomaterial Interactions

Vibeke Videm

Department of Laboratory Medicine, Children's and Women's Health, The Norwegian University of Science and Technology, and Department of Immunology and Transfusion Medicine, Trondheim University Hospital, Trondheim, Norway

INTRODUCTION

This chapter first gives a brief overview of normal neutrophil functions that parallel many of the reactions to biomaterials. Differences that influence neutrophil reactions in various settings employing biomaterials are mentioned, and mechanisms for neutrophil adhesion to biomaterials are discussed. The various consequences of neutrophil adhesion to biomaterials are described, including the oxidative burst, degranulation, neutrophil functional changes, and changes in neutrophil lifespan. The last part of the chapter provides a brief overview of some methods commonly employed to study neutrophil–biomaterial interactions.

Neutrophil Granulocytes—Part of the First-Line Defense

Neutrophil granulocytes (polymorphonuclear cells) constitute approximately 40–75% of the leukocytes in blood. These short-lived, highly mobile cells are important in the body's first-line defense, and their reactions to biomaterials are parallel in many respects to their reactions against other foreign substances such as microorganisms. A brief description of these normal reactions is therefore given.^[1]

In the resting state, neutrophils patrol the microcirculation. If they encounter activating stimuli that are liberated as a result of tissue damage, microorganisms, etc., they first become primed and develop a lower threshold for activation and subsequently may become more and more strongly activated. Activated neutrophils express adhesion molecules that bind to complementary receptors on activated endothelium in the threatened area. If the activating stimuli are strong enough to cause firm adhesion, the neutrophils traverse the endothelium and migrate into the tissues, guided by chemoattractants that lead them to the damaged or infected area. Here, neutrophils phagocytose (ingest) microbes and kill them by a combination of mediators such as proteinases from their abundant cytoplasmic granules and highly reactive oxygen species generated in reactions initiated by the enzyme complex NADPH

oxidase. This increased production of reactive oxygen metabolites is denoted the respiratory burst. During activation and degranulation (liberation of mediators from the cytoplasmic granules), there is always some external release of substances that may harm the surrounding tissue. Faced with foreign material that is too big to be phagocytosed, such external degranulation is augmented, a process sometimes called frustrated phagocytosis. Adhesion of slightly activated neutrophils to endothelium may be reversible, permitting primed cells with a reduced activation threshold to reenter the circulation. Resting neutrophils undergo apoptosis after a few days, whereas activated neutrophils live somewhat longer in the tissues until they either undergo apoptotic or non-apoptotic cell death.

Neutrophil–Biomaterial Interactions Depend on the Setting

Neutrophils readily identify biomaterials as foreign and react by becoming primed or activated. The interactions depend on the specific biomaterial, but also on the specific application. Important modifying factors are whether the biomaterial contact is time limited (e.g., during cardiopulmonary bypass) or permanent (e.g., in a vascular graft) and whether the exposure is repeated (e.g., hemodialysis). For biomaterials with direct blood contact, blood flow rates and the level of shear stress will also influence the interactions. Neutrophils normally react more strongly if there are other activating stimuli present as well, such as complement activation products or proinflammatory cytokines. The presence or absence of other cells such as monocytes or platelets influences the reactions, as well as plasma proteins including complement. Finally, there are large individual variations in the intensity of the reactions among different persons. Metabolic derangements, e.g., as in uremic patients, may also modify the responses. Thus, even though some general principles may form the basis for anticipation of the responses, specific testing of a biomaterial in the intended application is mandatory to fully characterize its interactions with neutrophils.

A number of studies on neutrophil–biomaterial interactions involve the use of isolated neutrophils. Isolation procedures should be carried out at 4°C to avoid in vitro priming or activation by biomaterials in pipette tips, tubes, etc., and polypropylene tubes should be used instead of polystyrene.

NEUTROPHIL INTERACTIONS ARE MAINLY WITH ADHERENT PROTEINS

In vivo, neutrophils react directly with the biomaterial surface only to a small extent, even if they do so during in vitro tests employing protein-free buffers. This is because biomaterials in contact with blood immediately become coated with plasma proteins and later platelets as well. Although less well characterized, protein adsorption will also take place on implants in tissues. They may become coated with plasma proteins as a result of bleeding during implantation and also absorb proteins from the extracellular fluid. The nature of the adsorbed protein coat is assumed to be decisive for the subsequent interactions with neutrophils. The composition of the protein layer, in turn, is highly dependent on the biomaterial, e.g., the chemical structure, charge and whether it is hydrophobic or hydrophilic, whether active molecules such as heparin or anti-inflammatory agents have been incorporated, etc. The adsorbed proteins may become more or less denatured, and their orientation subsequently exposing or hiding reactive groups is of importance. Some biomaterials activate the complement system and contribute to formation of complement activation products that may react with neutrophils, either in the circulation or locally if the biomaterial also binds the activation products.

MECHANISMS FOR NEUTROPHIL ADHESION TO BIOMATERIALS

A number of proteins and chemical groups on biomaterials may react with neutrophil granulocytes. For some molecules, the receptor or receptors on neutrophils are identified, but receptors yet unidentified are clearly involved as well^[2] and unknown groups on the biomaterials also engage in the interactions. Table 1 summarizes neutrophil receptors for various proteins that are commonly adsorbed onto biomaterials in vivo. In settings where platelets adhere to the adsorbed proteins, adhesion of neutrophils to these platelets may also take place. Table 2 summarizes neutrophil receptors to the most important ligands on biomaterial-adherent platelets. It is probable that neutrophil adhesion to platelets may also involve bridging molecules such as fibrinogen or von Willebrand factor.

Table 1 Neutrophil receptors for proteins on biomaterials

Surface-adsorbed protein	Neutrophil receptor
Albumin	CD43 (leukosialin)
Coagulation factor X	CD11b/CD18 (CR3, Mo1, Mac-1)
Complement	
C1q	Various C1q-receptors
C3 breakdown products	CD35 (CR1), CD11b/CD18 (CR3, Mo1, Mac-1)
C5a	CD88 (C5a receptor)
Fibrinogen/fibrin	CD11b/CD18 (CR3, Mo1, Mac-1), CD11c/CD18 (p150,95)
Fibronectin	CD49e/CD29 (VLA-5), CD51/CD29, integrin α V β 5
IgG	CD16 (Fc γ RIII), CD32 (Fc γ RII)
Kininogens	CD11b/CD18 (CR3, Mo1, Mac-1), CD49e/CD29 (VLA-5)
von Willebrand factor	CD162 (PSGL-1), CD11b/CD18 (CR3, Mo1, Mac-1), CD11c/CD18 (p150,95)
Hydroxyl groups (–OH)	Carbohydrates?
Unknown	Proteoglycans?

Abbreviations: CR3, complement receptor-3; VLA-5, very late antigen-5; PSGL-1, P-selectin glycoprotein ligand-1.

Because the detailed composition of the adsorbed proteins on different biomaterials in use in vivo is mostly unknown, we can only point at possible interactions with neutrophils. In addition, the most important mechanisms may not necessarily involve the most abundant proteins.

Table 2 Neutrophil receptors for platelets on biomaterials

Platelet receptor	Neutrophil receptor
CD62P (P-selectin)	CD162 (PSGL-1)
CD42b–V-IX complex	CD11b/CD18 (CR3, Mo1, Mac-1)
JAM-C	CD11b/CD18 (CR3, Mo1, Mac-1)?
CD41/CD61 (GPIIb/IIIa)	CD11c/CD18 (p150,95), fibrinogen

Abbreviations: CR3, complement receptor-3; JAM-C, junctional adhesion molecule-C; PSGL-1, P-selectin glycoprotein ligand-1.



The dominating mechanisms for neutrophil adhesion to hydrophilic and hydrophobic surfaces are probably different, but generalization is difficult because the specific chemical groups and charged molecules of the biomaterial also influence the reactions. The protein layer adsorbed to hydrophobic surfaces contains substantial amounts of fibrinogen, and the complement system is also strongly activated by many hydrophobic surfaces. Neutrophil adhesion may take place directly to fibrin/fibrinogen, to complement activation products, and to platelets. Similar mechanisms are involved on hydrophilic surfaces, but the adsorbed fibrinogen is subsequently partly exchanged with coagulation factor XII (Hageman factor) and kininogens. Factor XII induces coagulation and adsorption of thrombin. Complement activation is less intense with many hydrophilic surfaces, and neutrophil adhesion to adsorbed IgG may therefore become more dominating. Surfaces with negatively charged groups bind kininogens, but the site that reacts with neutrophil receptors is thereby hidden. Adsorbed kininogens therefore tend to reduce the material's reactions with neutrophils. A similar "pacifying" effect is seen with adsorbed albumin. Some data indicate that increased hydroxylation and expression of other negative groups on the surface result in more neutrophil adhesion,^[3,4] but positive charges may also increase the reactions.^[5]

FACTORS THAT MODIFY NEUTROPHIL-BIOMATERIAL ADHESION

A number of other factors than biomaterial composition influence neutrophil adhesion. Smooth surfaces generally lead to less adhesion than rough ones,^[6,7] but the mechanisms for this effect are not well understood. The number of adherent neutrophils increases gradually with shear stress up to approximately 3–4 dyn/cm² and then decreases.^[8,9] The use of different anticoagulants during extracorporeal circulation may also be of importance: Citrate reduces the available Ca²⁺, which may inhibit neutrophil functions, whereas heparin in clinical doses primes neutrophils and reduces their activation threshold.^[10]

CONSEQUENCES OF NEUTROPHIL BIOMATERIAL ADHESION

Neutrophil adhesion is coupled to other aspects of activation. The β 2-integrin CD11b/CD18 (CR3, Mo1, Mac-1) is involved in adhesion to many substrates (Table 1), and both this and other adhesive receptors interact with the cell's cytoskeleton. As they adhere, the neutrophils change shape, becoming flat and

outstretched and acquiring pseudopods and the ability for amoeboid movement (diapedesis). Most often, these shape changes parallel the degree of activation and are related to how firmly the cells have adhered, but sticking and spreading should be regarded as partly separate phenomena. The presence of pillars, holes, ridges, etc., greatly influences neutrophil motility on implanted biomaterials.^[11]

Two other important downstream events are assembly of the NADPH oxidase, responsible for production of reactive oxygen metabolites, and degranulation, which are further described in the following paragraphs. Neutrophil adhesion to biomaterials may therefore lead to liberation of the blood or tissues of substances that contribute to local damage and sometimes even systemic activation of inflammation. The more intense activation the biomaterial elicits, the higher the concentrations and wider the selections of mediators are released, and the more harm may be caused.

Oxidative Burst

The NADPH oxidase consists of several protein components from the cytosol as well as the membrane-bound cytochrome *b*₅₅₈ and other membrane-bound proteins. Probably as a safeguard against erroneous production of highly reactive oxygen metabolites, these components are not physically assembled before an activating stimulus such as adhesion to a biomaterial is received. Activation of the oxidase results in the so-called oxidative burst, i.e., the production of large quantities of superoxide anions and liberation of hydrogen peroxide (H₂O₂) that is used for the production of long-lived reactive species, including hypochlorous acid (HOCl), catalyzed by the enzyme myeloperoxidase. These reactive oxygen species may damage a variety of molecules, including lipids in cell membranes, but they also lead to activation of proenzymes to form tissue-destructing proteinases such as collagenase and gelatinase, as well as impairment of the antioxidant defenses by inactivation of α 1-antitrypsin. Ideally, these highly reactive substances should mainly be released into phagosomes containing internalized microorganisms to be killed, but in the setting of biomaterial adhesion, they are often released extracellularly in substantial amounts and may cause damage.

Degranulation

Neutrophils contain three types of granules in addition to secretory vesicles, forming reservoirs of molecules necessary to rapidly change phenotype and perform new functions upon activation. Table 3 indicates a few

Table 3 Some constituents of neutrophil granules and vesicles^[12]

Granule/vesicle type ^a	
Secretory vesicles	Membrane-bound Adhesion molecules (CD11b/CD18) Activation receptors (complement receptor 1, C1q-receptor, FMLP receptor, CD14, CD16)
Gelatinase granules	Membrane-bound Adhesion molecules (CD11b/CD18) Activation receptors (FMLP receptor) Cytochrome <i>b</i> ₅₅₈ Matrix Enzymes (gelatinase) Antimicrobial substances (lysozyme)
Specific (secondary) granules	Membrane-bound Adhesion molecules (CD11b/CD18, CD66, laminin receptor, vitronectin receptor) Cytochrome <i>b</i> ₅₅₈ Matrix Enzymes (collagenase, gelatinase, sialidase) Antimicrobial substances (lactoferrin, lysozyme)
Azurophilic (primary) granules	Matrix Enzymes (myeloperoxidase, proteinase-3, sialidase, elastase) Antimicrobial substances (lysozyme, cathepsins, defensins, BPI)

Abbreviations: BPI, bactericidal permeability increasing protein; FMLP, formyl-methionyl-leucyl-phenylalanine.

^aThe distinction into different granules is less absolute in vivo than indicated here, with partly overlapping species.

important constituents of these granules and vesicles, but is far from complete.^[12] Upon activation, the secretory vesicles are mobilized first and their membranes are fused to the cell membrane, immediately providing the cell with more adhesion molecules and receptors to receive further activating stimuli. Next, gelatinase granules are mobilized and then the specific granules, and upon even stronger stimulation the azurophilic granules. The gelatinase and specific granules contain enzymes that can degrade basement membranes and adhesion molecules, facilitating migration into the tissues. The further down in the degranulation hierarchy, the more molecules involved in microbial killing—including cytochrome *b*₅₅₈ and myeloperoxidase—become available.

Functional Changes

Neutrophil–biomaterial contact may also result in functional changes. In this context, it is important to note that adhesion may not necessarily be permanent, especially in settings with biomaterial–blood contact such as hemodialysis or cardiopulmonary bypass.

Shear stress and changes in receptor binding strength over time may permit cells to reenter the bloodstream. Such neutrophils may be either primed or functionally less responsive (exhausted). The changes may be complex, as indicated in an in vitro model study of cardiopulmonary bypass, where neutrophil adhesion to endothelial cells was unchanged after recirculation but reduced upon subsequent stimulation,^[13] but where adhesion to gelatin representative of collagen in the extracellular matrix was increased. Both primed and exhausted neutrophils may pose risks to the patient, especially because there also is extensive fluid-phase liberation of proinflammatory mediators during extracorporeal circulation. Primed neutrophils released after biomaterial contact may be “overactive” in this proinflammatory milieu, adhering to activated endothelium or exposed subendothelial tissues in different organs and contributing to the “post-pump syndrome” following cardiopulmonary bypass or respiratory symptoms in hemodialysis. At the same time, extracorporeal circulation may increase the risk of infections because neutrophils that are exhausted after biomaterial contact are incapable of performing their normal antimicrobial functions.



In the setting of implants, functional changes in adhered neutrophils may pose similar risks. Depending on the nature of the adsorbed proteins on a biomaterial surface, bacterial phagocytosis and killing may be impaired, perhaps as the result of an excess of cytoskeletal organization induced upon adhesion.^[14] The nature of the biomaterial and topographic details also influences mobility of adherent neutrophils,^[11,15] which may be of importance for their ability to move toward microbes in order to phagocytose and kill them. Finally, products released from adherent neutrophils may induce hyporesponsiveness in fresh neutrophils,^[16] a phenomenon that may contribute to explaining implant-associated infections.

Biomaterial-Associated Changes in Neutrophil Lifespan

Biomaterial contact may change neutrophil lifespan, resulting in either increased or decreased apoptosis or non-apoptotic cell death. Various inflammatory mediators including reactive oxygen species, complement activation products, and proinflammatory cytokines such as IL-6 and IL-8 downregulate neutrophil apoptosis. This may be seen as a useful adaptation when combating infections. In the context of cardiopulmonary bypass, however, delayed apoptosis^[17] may contribute to prolonged inflammatory response. Dialysis patients demonstrate greater neutrophil apoptosis than control subjects, both as a consequence of biomaterial contact and through the interaction with monocytes.^[18] This may contribute to the increased propensity to infection in patients on hemodialysis. Adhesion to certain biomaterials may also induce non-apoptotic cell death, and this effect is dependent on surface roughness. Neutrophils adherent to a roughened surface produce more reactive oxygen species than cells adherent to a smooth surface of the same biomaterial, and, in parallel, activation of a src kinase family-mediated signaling pathway triggers non-apoptotic cell death.^[7] This may represent an alternative form to apoptosis of activation-induced cell death, perhaps as a regulatory mechanism to avoid too-strong inflammatory responses. However, in non-apoptotic cell death released cellular contents such as enzymes may damage the surrounding tissue.

Factors That Modify Systemic Neutrophil Activation upon Biomaterial Contact

In extracorporeal circulation, the foreign surface area relative to the blood volume is important for the degree of systemic neutrophil activation. With a relatively larger area, activation of non-adherent neutrophils is greater.^[19] Conceivably, this effect is of

importance both in hemodialysis and in cardiopulmonary bypass. However, because substantial differences in activation are seen among the various membrane biomaterials employed in hemodialysis, the choice of membrane may be much more important for neutrophil activation than the surface area in this setting.

Biomaterial-induced neutrophil activation increases with lower hematocrit levels, although hemodilution alone does not induce neutrophil activation.^[20] This is of importance in cardiopulmonary bypass, especially in children. The relationship between surface area and degree of hemodilution with the magnitude of the inflammatory responses is an important reason for the trend toward smaller cardiopulmonary bypass circuit surface areas.

METHODS FOR EVALUATION OF NEUTROPHIL-BIOMATERIAL INTERACTIONS

As outlined above, neutrophil-biomaterial interactions are diverse, including adhesion, functional changes of adherent and non-adherent cells, and changes in lifespan, and may be influenced by several other factors than the chemical composition of the biomaterial. This complexity renders evaluation of such interactions difficult, and there is no general consensus regarding which methods should be preferred. Two paramount requirements should be considered. First, more than one parameter should be evaluated, because biomaterial contact induces a number of changes that are not necessarily correlated. Second, evaluation should be performed during the intended application or in a model setting that mimics this application. Some standard methods to evaluate various parameters are given in Table 4 and are further discussed in the following paragraphs.

Biomaterial-Related Adhesion

Many of these methods are based on different forms of microscopy, with the advantage that not only the number of adherent neutrophils can be quantified, but also cellular morphology. The degree of spreading often provides an indication of the level of activation. Methods based on microscopy are labor intensive and, with the exception of light microscopy that may even be performed on unstained cells, often costly. Cell enzyme immunoassay is a versatile, inexpensive method that is useful for large series of experiments. The cell number on a surface is quantified by employing an enzymatic reaction dependent on binding of a primary antibody to the adherent cells. Primary antibodies against myeloperoxidase are often used

Table 4 Commonly used methods to evaluate neutrophil-biomaterial interactions^a

Test parameter	Method
Adhesion to biomaterial	Scanning electron microscopy
	Fluorescence or confocal microscopy
	Light microscopy
	Cell enzyme immunoassay (cell ELISA)
	Quantification of constituent after solubilization
Adhesion molecule expression	Counts in incubation medium or blood sample
	Cell enzyme immunoassay on adherent cells
	Flow cytometry on adherent cells after detachment
Degranulation	Flow cytometry on circulating cells
	Enzyme immunoassay of supernatant
Phagocytosis	Enzymatic assay in supernatant
	Flow cytometry
Oxidative burst	Methods based on radiolabeled bacteria
	Flow cytometry
Bacterial killing	Chemiluminescence
Motility	Bacterial colony counts
	Chemotaxis assay under agarose, in Boyden chamber, or in tissue culture wells
Viability	Chemotaxis assay on biomaterial surface
	Dye uptake or exclusion tests

Abbreviation: ELISA, enzyme-linked immunosorbent assay.

^aSee text for details.

because neutrophils contain abundant quantities, but antibodies against any granule constituent that is relatively specific for neutrophils may be employed. No information on morphology is available, and results are given on a relative scale if not calibrated against a method based on some form of counting. Adherent neutrophils may also be solubilized to allow granule constituents to be released to the medium, in which quantification, for example of myeloperoxidase, is performed. The advantages and disadvantages are as for cell enzyme immunoassay, but results often demonstrate somewhat larger variation as a result of the extra steps involved. A very simple method is to quantify the number of neutrophils in the incubation medium, assuming that any differences from the

starting counts are a result of removal by adhesion to the biomaterial. This works reasonably well in vitro even if cell death may be a source of error, either because cells disintegrate or because they become apoptotic and divide into a number of apoptotic particles. In vivo, changes in margination and liberation of neutrophils from the bone marrow as part of the inflammatory response render such counts fairly unreliable.

In some experiments, circulating neutrophils after biomaterial contact are tested for their subsequent ability to adhere to specific substrates, such as gelatin, vitronectin, or laminin, or to cultured endothelial cells in order to evaluate priming or exhaustion. The mentioned methods may also be employed in such studies.

Adhesion Molecule Expression

Cell enzyme immunoassay on adherent cells may also be used to quantify changes in adhesion molecule expression. With this method, it is important that the cells are not permeabilized, because this will allow the primary antibody to bind to intracellular stores of adhesion molecules, as well as those expressed on the cell surface. An alternative method is to stain the cells with antibodies against the relevant adhesion molecules after detachment from the biomaterial and analyze them using flow cytometry. For neutrophils, this method has some caveats, rendering thorough evaluation of the detachment conditions necessary. If detachment is performed using trypsin, the adhesion molecules one wishes to quantify may be partly or completely removed, even with low trypsin concentrations. Detachment using ethylenediaminetetraacetic acid is not always sufficiently effective for the most activated cells, thereby creating a selection bias. Removal using a cell scraper may be useful. Flow cytometry is labor intensive, but has a higher sensitivity than cell enzyme immunoassay and permits simultaneous assessment of many adhesion molecules. Flow cytometry is an excellent method for measurement of adhesion molecule expression on circulating neutrophils after in vivo or in vitro extracorporeal circulation, providing information specific to each sampling time point. However, the most activated neutrophils may be adherent to the biomaterials or activated endothelium and thus are unavailable for assessment. The β 2-integrins (CD11/CD18, especially CD11b) are often used as activation markers. It is important to note that conformational changes in integrins may activate them to increased binding without changes in receptor numbers. L-selectin, which is shed during activation, may also be a useful marker.

Degranulation

Even if degranulation is required for changes in adhesion molecule expression, the two parameters are not always correlated. One explanation may be that adhesion molecule expression gives a single-moment picture, whereas degranulation products become accumulated over time and may stem from both adherent and non-adherent neutrophils. Enzyme immunoassay of supernatants or plasma after neutrophil–biomaterial contact is an easy, inexpensive method with high sensitivity that is well suited for large series. By combining analysis for constituents from different granules, a more detailed picture of events may be obtained. Lactoferrin is often used as a marker for specific granules and myeloperoxidase or elastase may be used as markers for azurophilic granules. Because the granules contain a number of enzymes, quantification of enzymatic activity in supernatants or plasma may also serve as an indicator of degranulation. Sensitivity is often lower than for enzyme immunoassay, but these methods are often rapid, inexpensive, and well suited for large series. Enzyme immunoassay or tests for enzymatic activity may also be performed in neutrophil lysates after biomaterial contact in order to quantify the remaining amount of some granule constituent, often in comparison with lysate from precontact samples. Measurement of degranulation products released from neutrophils after biomaterial contact and subsequent stimulation, e.g., with phorbol ester or formyl-methionyl-leucyl-phenylalanine (FMLP), gives an indication of the “remaining function,” i.e., if the cells have become primed or exhausted.

Phagocytosis

The capacity of neutrophils for phagocytosis after biomaterial contact is often measured by flow cytometry after an incubation period with opsonized fluorochrome-labeled bacteria. This method is simple and sensitive and imposes no risk of infection. Other methods are based on phagocytosis of radiolabeled bacteria with subsequent counting of neutrophil-related radioactivity. This is also a sensitive method, but requires adequate protection of the researchers and routines for disposal of the radioactive waste.

Oxidative Burst

The simplest and most sensitive methods are based on flow cytometry. The neutrophils are incubated with a non-fluorescent substance such as dihydrorhodamine, which is oxidized to a fluorescent compound by hydrogen peroxide or superoxide anions. Both spontaneous

production of reactive oxygen metabolites and production after stimulation may be quantified, and the method is easily combined with measurements of adhesion molecule expression. Other methods are based on chemiluminescence, which is usually dependent on stimulation of the neutrophils and thus indicates only the “remaining function,” not the spontaneous level of reactive oxygen metabolite production. Using luminol or lucigenin as probes, intracellular or secreted oxidants, respectively, may be measured with high sensitivity. The cytochrome *c* reduction assay is an alternative method to measure the oxidative burst. It is specific for measurement of secreted superoxide anions, but will not detect intracellular oxidant production. Both the cytochrome *c* reduction assay and chemiluminescence-based methods are less used after the introduction of the much simpler flow cytometric assays. All methods for measurement of oxidative burst are relatively labor intensive.

Bacterial Killing

The most standard methods are based on incubation of bacteria with neutrophils that have been in contact with biomaterials, followed by plating of the bacteria and subsequent counting of bacterial colonies.

Motility

Many motility assays are dependent on transferring the neutrophils that have been in contact with biomaterials to the test setup, e.g., a petri dish with wells cut out in agarose, a Boyden chamber, or a tissue culture well with a porous filter insert. The cells are then induced to move in one direction by a chemoattractant such as FMLP, and the results are evaluated by staining and counting. In the tissue culture well setup, the laborious counting step may be substituted if the neutrophils are preincubated with a fluorescent or radioactive label. The cells that have passed through the pores are solubilized and quantified using a fluorometer or gamma/beta counter. All these methods are unsuitable for evaluating the mobility of cells adherent to the biomaterial. In order to overcome this disadvantage, assays have been developed in which the biomaterial is fixed on the bottom of a tissue culture well under a layer of agarose. Neutrophils are added to a well in the agarose so that they move on the biomaterial surface in the direction of a chemoattractant. The counting step may also be substituted with image analysis techniques, permitting a much more detailed analysis of movement, including repeated observation of single neutrophils.

Viability

The simplest tests are based on differences in dye exclusion or uptake between live and dead neutrophils. For example, trypan blue, ethidium bromide, and propidium iodide are excluded as long as the cells are alive, whereas acridine orange permeates the membranes of live cells. Stained and unstained cells are evaluated either by light microscopy (trypan blue) or by fluorescence microscopy/flow cytometry (e.g., ethidium bromide, propidium iodide, acridine orange). Detection of apoptosis may be performed using a number of methods that are beyond the scope of this text. One alternative is to employ flow cytometry with fluorochrome-conjugated annexin V, which binds to phosphatidylserine that is exposed on the plasma membrane in apoptotic cells.

CONCLUSION

Upon biomaterial contact, neutrophil granulocyte priming, adhesion, or activation may lead to clinical consequences such as increased inflammation with subsequent organ dysfunction, as well as an increased risk of infections. Therefore, biomaterials intended for clinical applications should be thoroughly screened for potential neutrophil-related adverse effects by examining various aspects of neutrophil activation and functions in a relevant setting.

ARTICLES OF FURTHER INTEREST

Biofunctional Polymers; Biomaterials Immune Response; Biomaterials: Protein-Surface Interactions

REFERENCES

1. Edwards, S.W. *Biochemistry and Physiology of the Neutrophil*; Cambridge University Press: Cambridge, 1994.
2. Videm, V. Endpoint-attached heparin blocks neutrophil sticking and spreading. *Biomaterials* **2004**, *25*, 43–51.
3. Sundaram, S.; Lim, F.; Cooper, S.L.; Colman, R.W. Role of leucocytes in coagulation induced by artificial surfaces: investigation of expression of Mac-1, granulocyte elastase release and leucocyte adhesion on modified polyurethanes. *Biomaterials* **1996**, *17*, 1041–1047.
4. Tomczok, J.; Sliwa-Tomczok, W.; Klein, C.L.; van Kooten, T.G.; Kirkpatrick, C.J. Biomaterial-induced alterations of human neutrophils under fluid shear stress: Scanning electron microscopical study in vitro. *Biomaterials* **1996**, *18*, 1359–1367.
5. Nimeri, G.; Öhman, L.; Elwing, H.; Wetterö, J.; Bengtsson, T. The influence of plasma proteins and platelets on oxygen radical production and F-actin distribution in neutrophils adhering to polymer surfaces. *Biomaterials* **2002**, *23*, 1785–1795.
6. Eriksson, C.; Lausmaa, J.; Nygren, H. Interactions between human whole blood and modified TiO₂-surfaces: Influence of surface topography and oxide thickness on leukocyte adhesion and activation. *Biomaterials* **2001**, *22*, 1987–1996.
7. Chang, S.; Popowich, Y.; Greco, R.S.; Haimovich, B. Neutrophil survival on biomaterials is determined by surface topography. *J. Vasc. Surg.* **2003**, *37*, 1082–1090.
8. Kao, W.J. Evaluation of leukocyte adhesion on polyurethanes: The effects of shear stress and blood proteins. *Biomaterials* **2000**, *21*, 2295–2303.
9. Shive, M.S.; Hasan, S.M.; Anderson, J.M. Shear stress effects on bacterial adhesion, leukocyte adhesion, and leukocyte oxidative capacity on a polyurethane. *J. Biomed. Mater. Res.* **1999**, *46*, 511–519.
10. Videm, V. Heparin in clinical doses “primes” granulocytes to subsequent activation as measured by myeloperoxidase release. *Scand. J. Immunol.* **1996**, *43*, 385–390.
11. Tan, J.; Saltzman, W.M. Topographical control of human neutrophil motility on micropatterned materials with various surface chemistry. *Biomaterials* **2002**, *23*, 3215–3225.
12. Borregaard, N.; Cowland, J.B. Granules of the human neutrophilic polymorphonuclear leukocyte. *Blood* **1997**, *89*, 3503–3521.
13. Åsberg, A.E.; Videm, V. Neutrophil dysfunction after biomaterial contact in an in vitro model of cardiopulmonary bypass. *Eur. J. Cardiothorac. Surg.* **2006**, *30*, 744–748.
14. De La Cruz, C.; Haimovich, B.; Greco, R.S. Immobilized IgG and fibrinogen differentially affect the cytoskeletal organization and bactericidal function of adherent neutrophils. *J. Surg. Res.* **1998**, *80*, 28–34.
15. Zhou, Y.; Doerschuk, C.M.; Anderson, J.M.; Marchant, R.E. Biomaterial surface-dependent neutrophil mobility. *J. Biomed. Mater. Res.* **2004**, *69A*, 611–620.
16. Kaplan, S.S.; Basford, R.E.; Jeong, M.H.; Simmons, R.L. Biomaterial–neutrophil interactions: Dysregulation of oxidative functions of fresh neutrophils induced by prior neutrophil–biomaterial interaction. *J. Biomed. Mater. Res.* **1996**, *30*, 67–75.
17. Chello, M.; Mastroberto, P.; Aurino, A.; Cuda, G.; Peticone, F.; Cirillo, F.; Covino, E. Inhibition of neutrophil apoptosis after coronary bypass operation with cardiopulmonary bypass. *Ann. Thorac. Surg.* **2002**, *73*, 123–130.
18. Nahar, N.; Shan, H.; Siu, J.; Colvin, R.; Bhaskaran, M.; Ranjan, R.; Wagner, J.D.; Singhal, P.C. Dialysis membrane-induced neutrophil apoptosis is mediated through free radicals. *Clin. Nephrol.* **2001**, *56*, 52–59.
19. Gourlay, T.; Stefanou, D.C.; Asimakopulos, G.; Taylor, K.M. The effect of circuit surface area on CD11b(mac-1) expression in a rat recirculation model. *Artif. Org.* **2001**, *25*, 475–479.
20. Gourlay, T.; Samartzis, I.; Taylor, K.M. The effect of haemodilution on blood–biomaterial contact-mediated CD11b expression on neutrophils: Ex vivo studies. *Perfusion* **2003**, *18*, 87–93.



Nickel and Nickel Alloys

Rachel L. Williams

Department of Clinical Engineering, University of Liverpool, Liverpool, United Kingdom

INTRODUCTION

Nickel alloys play an important role in dentistry, orthopedics, and cardiovascular applications. In these applications, Ni is alloyed with another metal that forms a stable oxide layer—for example, chromium or titanium. These control the dissolution of Ni into the biological environment, which is important because of potential adverse biological responses to the nickel ions.

NICKEL CHEMISTRY

Nickel is a transition element with atomic number 28.^[1] It has a melting point of 1455°C and a boiling point of 2730°C. The prevalent oxidation states of nickel are 0 (as in nickel metal and its alloys) and 2+ (as in most of its inorganic compounds). The main nickel compounds are its oxides, sulfides, and salts. As a pure metal it is silvery-white in color, has ferromagnetic properties, and is malleable and ductile. Ni²⁺ forms complexes in aqueous solutions with negative groups or neutral polar molecules. Nickel is widely distributed in nature, making up about 0.008% of the earth's crust. Its main ores are pentlandite (a nickel/iron sulfide) and garnierite (a nickel/magnesium silicate). Of the many hundred thousand tons of nickel produced annually, 70% is used in the production of alloys. These alloys are used in a wide range of industrial processes. Alloys of particular interest to biomedical applications are the Ni–Cr dental casting alloys, austenitic stainless steels, and Ni–Ti shape-memory alloys.

NICKEL ALLOYS

Ni–Cr Alloys

Cast Ni–Cr alloys have been extensively used in dentistry since the 1960s for partial denture frameworks, crowns and bridges, and porcelain-fused-to-metal restorations.^[2–6] More recently, Co–Cr alloys have been predominantly used for framework structures, but Ni–Cr alloys continue to be used for crowns, bridgework, and endodontic posts^[3] (Figs. 1 and 2).

The important properties for these applications are castability, sufficient mechanical properties, and corrosion resistance. The fulfilment of these requirements derives from the relationship between composition, structure, processing, and material properties.

A typical composition for a cast Ni–Cr dental alloy is shown in Table 1. The major alloying element Cr is added to provide the alloy's corrosion resistance by forming a passivating oxide layer. The minor alloying elements, in fact, have a greater influence on the physical and mechanical properties of the alloy. Mo and Al both contribute to increasing the strength of the alloy. Aluminium forms a Ni₃Al compound that significantly increases the yield and ultimate tensile strength of the alloy. Carbon plays a significant role in increasing the hardness of the alloy due to the formation of metal carbides with other alloying elements—for example, Cr, Si, Mo, and Ni. The actual percentage of C must be carefully controlled: If too much C is added (i.e., more than 0.2%), the alloy becomes too hard and brittle for use in dentistry; if too little is added, the yield and ultimate tensile strengths become too low for the alloy to be effective. Beryllium (Be) is added as a solid-solution strengthener. It plays an important role in increasing the castability of the alloy by lowering the melting temperature by about 100°C, and has been demonstrated to improve the surface oxide and enhance adhesion of porcelains to the metal alloy. The Al content can help in this regard as well as Fe, which also reduces shrinkage following casting. Si and Mn are added to increase the fluidity and castability of the alloys.

Typical ranges for mechanical properties are given in Table 2. The percentage elongation is critically dependent on the microporosity of the casting. This can be controlled by the melting and casting variables (e.g., the alloy should not be overheated during melting). Control of the C and Mo content can also increase ductility without reducing strength. A higher elastic modulus can be an advantage because the greater rigidity of the structure allows the design of restorations with reduced dimensions. It is also argued that a well-designed rigid appliance provides a better distribution of forces to the supporting tissues during use.^[2]

The corrosion resistance of Ni–Cr dental alloys has dominated research on these materials. The oral

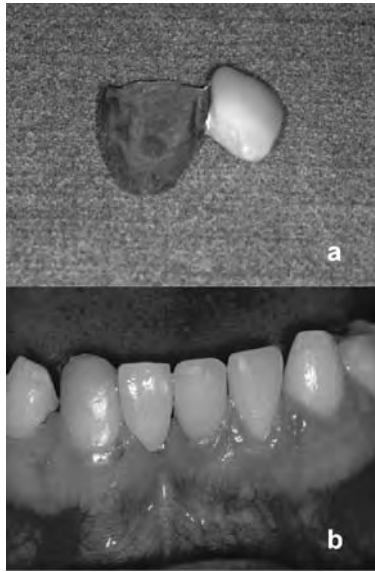


Fig. 1 Use of Ni–Cr dental casting alloy in crown and bridgework. (a) The bridge and (b) the crown are positioned with the bridge stuck the back of the neighboring tooth.

environment is aggressive to metal surfaces, and degradation of the material surface occurs due to corrosion in the saliva, wear and erosion caused by food and chewing, and bacterial activity.^[7] These all result in the release of metal ions into the environment, leading to the formation of various compounds (e.g., chlorides, sulfides, and oxides). The mechanism to minimize ion release is by control of the formation of a strong passive layer, usually an oxide, on the surface of the alloy. In Ni–Cr alloys this is achieved with Cr, but it must be remembered that all metal will release ions through the passivating oxide layer, although the rate may be very low. The biological significance of the metal ion release is discussed later.

Stainless Steel

Stainless steels are Fe–C alloys with the addition of Cr to provide a passivating oxide layer of Cr_2O_3 . There are many stainless steels with different properties depending on their composition and microstructure that, as a consequence, have a role in many different biomedical applications. The 300 series (American Iron and Steel Institute, AISI) of stainless steels^[8] also contains Ni and has applications in medical devices, including orthodontic wires in dentistry, wires, pins, plates and screws in orthopedics (Fig. 3), and stents in cardiovascular applications. The following discussion concentrates on this group of alloys.

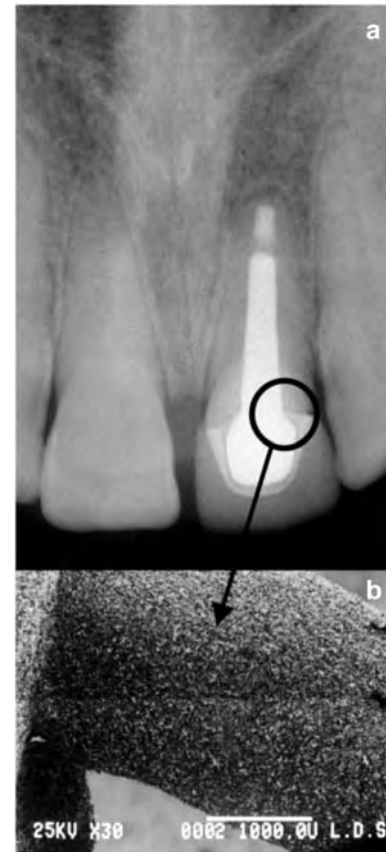


Fig. 2 Ni–Cr endodontic post. (a) A radiograph showing positioning of the post; (b) SEM of etched surface of the post. (Reproduced with kind permission of Dr. Nicolas Jedynakiewicz.)

Typical compositions of the 300 series of stainless steels are shown in Table 3. The main constituents are Cr, Ni, and Mo with low levels of other elements. (C content in particular is carefully controlled.) Chromium is added to provide a stable oxide, which

Table 1 Typical composition of Ni–Cr dental casting alloy

Element	Weight %
Ni	71–78
Cr	12–16
Mo	3.0–5.5
Fe	0–1.5
Si	0–3.0
Al	0–2.2
Be	0.4–1.9
Mn	Trace
C	0.1

(Adapted from Ref. [9].)

Table 2 Typical mechanical properties of cast Ni–Cr dental alloys

Mechanical properties of typical cast Ni–Cr alloys	
Yield strength at $E=0.002$ (MPa)	296–717
UTS (MPa)	> 800
Elongation %	Up to 18
Elastic modulus (GPa)	158–212
Vickers hardness	167–309

(Adapted from Refs. [2,4].)

maintains the corrosion resistance of the alloy. There must be at least 12% chromium to impart sufficient corrosion resistance for implantation into the body; Mo is also added to increase corrosion resistance. The amounts of Cr and Mo are controlled so that $\%Cr + 3.3 \times \%Mo \geq 26$. This is known as the pitting resistance equivalent (PRE).^[9] Nickel is added to the composition to maintain a completely austenitic microstructure at room temperature. It is important to have a completely austenitic microstructure to maintain the corrosion resistance and nonmagnetic properties, even though this restricts the maximum strength of the alloy that can be obtained. The carbon content is very carefully controlled. In implantable 316L stainless steel, the carbon content is maintained below 0.03%. It has been observed that in higher-carbon content 316 stainless steel ($C=0.08\%$) after prolonged heating above 400°C, chromium carbides can form at the grain boundaries. This has a significant influence on the alloy's corrosion resistance owing to the depletion of Cr in the vicinity of the carbides, and therefore reduction in the stability of the passive oxide layer in these areas. The sulfur content is kept low to restrict the formation of sulfides (e.g., MnS) that can impair surface quality. The Si content is also kept low to decrease silicate-type inclusions, thus providing greater stability of the austenitic structure. A low phosphorus content maintains improved ductility of the alloy, especially in implant devices that require cold-working to provide sufficient strength.

The microstructure of implant-quality stainless steel has single-phase austenitic structure and fine grain size. This allows the production of an alloy with a good combination of tensile strength and fatigue resistance. The alloy cannot be heat-treated to increase mechanical properties, but can be cold-worked. Typical values of the mechanical properties are presented in Table 4. The annealed alloy is particularly used for wires and fracture fixation plates where the strength is sufficient and the higher ductility an advantage to allow the surgeon to conform the plate to the shape of the bone during surgery. The cold-worked alloy is used in the manufacture of bone screws, bone plates, and intramedullary nails where increased strength is required. It is also reported^[9] that the torsional properties of 316L stainless steel allow the surgeon to feel when a bone screw has reached its yield torque, thus reducing the potential to overtighten screws, causing failure. Cold-drawn material is used for pins and wires where high resistance to bending is required (e.g., in Kirschner wires and Schanz screws). The cold-worked alloy also has increased fatigue resistance compared to annealed material, but at the expense of corrosion resistance (due to increased dislocation density).

In orthodontics, 316 stainless steel is used in the manufacture of archwires that apply force to straighten teeth. The wires are produced in wrought form by rolling and drawing. After being formed into appliance-specific correct shape, they are usually heat-treated to relieve stresses induced by fabrication. This produces wires with good corrosion resistance and the excellent mechanical properties of yield strength, elastic modulus, and spring rate.

Coronary artery stents are used to maintain patency of blood vessels following interventional cardiology and have rapidly developed into the principal method of percutaneous coronary revascularization.^[10] Ideal stent characteristics include flexibility, low unconstrained profile, high radial strength, high corrosion resistance, good biocompatibility, and low thrombogenicity. Most commercially available stents are manufactured from 316L stainless steel. The actual properties of each design depend on the configuration

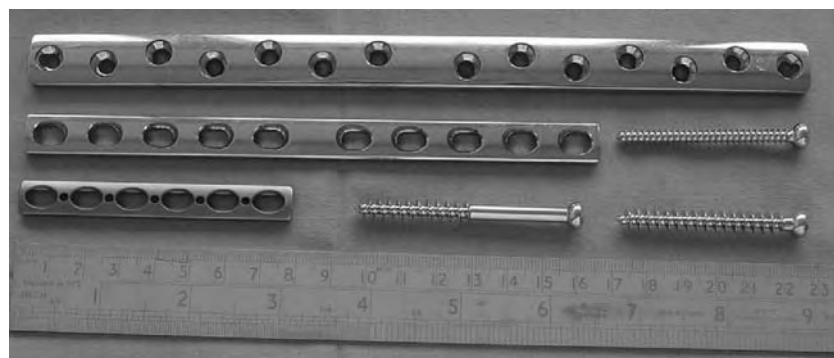
**Fig. 3** 316L stainless steel fracture fixation plates and screws.

Table 3 Typical composition of the 300 series austenitic stainless steels

	302	303	304	305	316	316L	317
Cr	17–19	17–19	18–20	17–19	16–18	16–18	18–20
Ni	8–10	8–10	8–10.5	10.5–13	10–14	10–14	11–15
C	0.15	0.15	0.08	0.12	0.08	0.03	0.08
Mn	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Mo		0.6 (optional)			2–3	2–3	3–4
P	0.045	0.20	0.045	0.45	0.045	0.045	0.045
S	0.03	≤0.15	0.03	0.3	0.03	0.03	0.03
Si	1.0	1.0	1.0	1.0	1.0	1.0	1.0

(Adapted from Refs. [2,4].)

of the stent. Concerns over the potential release of Ni from 316L stainless steel and its influence on restenosis^[11] have led to the development of coating technologies with other metals, metal oxides, or polymers containing antithrombotic or antiproliferative drugs.

The corrosion resistance of stainless steel is afforded by the chromium oxide that forms on the surface, protecting the underlying alloy from corrosive attack. The stability of the passive layer is critical to the corrosion resistance of the alloy. Many conditions can influence the stability of the oxide. In anodic polarization studies of medical-grade stainless steel, it has been shown that the difference between the corrosion potential and the protection potential is less than in other metallic biomaterials (e.g., Ti alloys and Co–Cr alloys),^[12] and it is therefore relatively easy to breach the passive layer and allow active corrosion to proceed. This can be influenced by geometrical factors (e.g., the existence of crevices) or relative motion between components (e.g., fretting). Areas of local high stress—perhaps resulting from dislocations or impurities—may lead to weakening of oxide stability and result in pitting corrosion.

Ni–Ti

Ni–Ti is an alloy with exceptional properties owing to its crystal structure. These properties include shape-memory behavior and pseudoelasticity (sometimes known as superelasticity). Ni–Ti alloys possess shape memory and pseudoelastic properties owing to the temperature- and stress-dependent transformation of their crystal structure from a parent phase, B2 (often called austenite), to a martensitic phase, B19'. (Find details of these transformation processes elsewhere in this encyclopedia.) Ni–Ti shape-memory alloys are based on an equiatomic composition of nickel and titanium. They can be manufactured by conventional

vacuum-induction melting^[13] or powder metallurgy^[14] techniques. The microstructure consists primarily of a single phase with small amounts of other phases dispersed in the matrix. If the composition changes from the equiatomic, the amount of the second phase can increase, affecting the properties of the alloy. The temperatures associated with the phase transformation (M_s , the martensite start temperature; M_f , the martensite finish temperature; A_s , the austenite start temperature; and A_f , the austenite finish temperature) are critically dependent on the composition. It is reported that, particularly on the Ni-rich side of the equiatomic, control of the composition must be on the order of 0.01%.^[13] A 0.1% difference in compositions can cause a 10°C change in M_s .^[14] The transformation characteristics may also be influenced by the inclusion of impurities such as C, O, H, and N during the manufacturing process due to Ti's high affinity for these elements.^[15]

The thermomechanical treatment of Ni–Ti alloys is critical to their performance and must therefore be carefully controlled. Fully annealed equiatomic Ni–Ti alloys have very good pseudoelastic and shape-memory properties. Thermomechanical treatment, however, is often required to increase the strength of the alloys. Cold-working alone will increase the yield strength of the alloy but destroy the pseudoelastic and shape-memory properties. Annealing postcold-working will

Table 4 Typical mechanical properties of implant quality 316L stainless steel

Condition	UTS (MPa)	Yield strength (MPa)	Elongation (%)
Annealed	490–690	> 190	> 40
Cold-worked	860–1100	> 690	> 12
Cold-drawn	1350–1600	–	–

(Adapted from Ref. [8].)



restore the memory effect but decrease the yield strength. By controlling the annealing conditions and degree of cold-working, it is possible to optimize the mechanical properties of the alloy. It is believed^[13] that cold-working alone introduces a high density of random dislocations and that annealing rearranges the dislocations into cells of relatively dislocation-free zones within which the martensite twins can be mobile but are surrounded by dislocation networks that restrict further mobility.

The introduction of dislocations also has a significant influence on transformation temperatures. A sample with a high dislocation density will have a lower M_s and a higher A_f than the fully annealed sample. The dislocations stabilize the B2 (austenite) phase such that on cooling, the transformation from B2 to B19' is hindered by the dislocations.^[13,16-18] There will also be a similar change in the critical stress required to initiate stress-induced martensite. Samples with low dislocation density and coarse grain structure (i.e., high annealing temperature) will have a lower critical stress than samples with high dislocation density and fine grain structure (i.e., low annealing temperature). The dislocations present in the matrix act as obstacles to the martensite growth during stress-induced martensite formation.

The shape-memory effect relies on the transformation of the phase structure from a B2 structure at high temperatures to a B19' structure at lower temperatures. In fully annealed alloys, the transformation occurs directly from B2 to B19' on cooling. For a cold-worked sample, the transformation occurs via the R-phase (i.e., B2 to R to B19').^[15,16] The mechanical properties of a Ni-Ti alloy are critically dependent on the amounts of the three phases present.^[17] The R-phase is a rhombohedral distortion of the B2 parent lattice structure. It can nucleate on a single dislocation, whereas the martensitic lattice structure can not, due to the large transformation shape change compared to the strain around a single dislocation.^[15] The R-phase appears on cooling prior to martensite under certain conditions and can be associated with a temperature hysteresis as small as 1.5°C.^[19] It shows pseudoelastic and shape-memory characteristics, and can occur more easily than the transformation to the B19' phase, as it requires smaller temperature changes or smaller loads. Thermomechanical treatments can be used to optimize the properties of Ni-Ti alloys for use in medical devices. The treatments can enhance the mechanical properties and be used to tailor the phase transformation temperatures to values suitable for in-vivo applications.

As with most implant alloys, the corrosion-resistance behavior in Ni-Ti alloys relies on a passive oxide layer. Many studies have examined the corrosion behavior of Ni-Ti alloys. Ni-Ti alloys process a stable,

passive oxide layer that will maintain its passivity when exposed to potentiodynamic polarization over a potential range similar to other Ti-based alloys. However, if for any reason the passive layer is breached, then repassivation may be slow and difficult. Any configuration or prior treatments that influence the stability of the passive oxide can therefore have a very significant effect on the corrosion resistance of the alloys. This includes scratching and wear,^[20,21] crevice geometry,^[22] roughened surfaces,^[23,24] stress concentrations,^[25] intermetallic inclusions,^[26] acid environments,^[22] protein adsorption,^[27] and pretreatments that modify the chemistry of the oxide.^[28,29]

The passivity of the oxide layer will depend on its composition. Some studies have suggested that the oxide is entirely Ti-oxide-based without any detection of Ni in the surface.^[27] Other studies have demonstrated the presence of metallic Ni and Ni oxides throughout the Ti-oxide layer.^[28-30] Depth profiling through the oxide has shown that its composition is critically dependent on prior treatment.^[31-33] However, it would appear that in many situations the outermost layer of the oxide is composed of TiO₂. It may be that this is how the wide passive region is conferred to the alloys that has been measured in some studies. However, if the outermost oxide is removed by wear or a scratch, then the underlying mixed oxide of Ti oxide and Ni oxide may be less stable in the aggressive chloride environment, leading to a slow and less stable repassivation.

A crevice geometry or roughened surface may exploit weaknesses in the surface protection and, as discussed in the foregoing, this may be due to the inclusion of Ni and Ni oxides in the predominantly Ti-oxide surface layer. Many practical Ni-Ti alloys have been pretreated with thermomechanical processes to optimize their mechanical performance. The presence of the dislocations formed could be considered to produce preferential sites for passive film breakdown owing to internal stress concentrations.^[25] Intermetallic phases (NiTi₂) have also been shown to act as preferential sites for pit initiation in Ni-Ti alloys.^[26] The intermetallic phase has been shown to form with only a very slight excess of Ti.^[13]

It has been reported^[34] that corrosion in an aggressive chloride environment has the ability to modify the microstructure of the alloy, even though the chemical composition remains unchanged. For heavily corroded sections of the sample, the phase transformation with change in temperature had almost been eliminated. Further into the corroded sample, the ability of the phase transformation to take place had been greatly reduced and the temperatures of the transformation had been lowered. It is also reported^[26] that the microstructure of the alloy can influence its corrosion resistance, and that a greater corrosion

resistance was observed on an alloy with an austenitic phase structure than a martensitic phase structure. Therefore, if corrosion causes a shift in the phase transformation temperature, the phase composition of the alloy at body temperature may change, thereby influencing its corrosion resistance as well as its mechanical properties.

Most of the studies on the corrosion behavior of Ni–Ti alloys have been performed using inorganic electrolytes to represent the *in vivo* or artificial saliva environments. It is well known, however, that proteins in the *in vivo* environment can have a significant influence on the corrosion behavior of implant metals and alloys. It has been suggested that the protein molecules in serum have a strong influence on the passivity of Ni–Ti alloy.^[27] Protein was shown to adsorb onto the oxide surface and cause a decrease in the breakdown potential and an increase in the passive corrosion current density.

Pretreating the surface of NiTi in some oxidizing environments (e.g., dilute HNO₃) can increase the passivity of the surface, resulting in a wider passive region and reducing the likelihood of localized corrosion due to pitting.^[22,24,28] Shabalovskaya^[29] reports on a study of the effect on surface properties of a Ni–Ti alloy of several procedures used for sterilization. This study reports that the pretreatments have a highly significant influence on the composition of the surface oxide, and that the composition of the oxide correlates with the elemental release into cell culture media from these surfaces. In particular, it was found that surfaces exposed to hydrogen peroxide possess a high Ni:Ti ratio in the surface oxide, and consequently a large release of Ni from the surface into the cell culture medium was measured. The oxide film on specimens that had been autoclaved was ten times more effective at preventing Ni release, compared with specimens exposed to hydrogen peroxide. Autoclaving or boiling the specimens in water resulted in very low Ni content in the surface, but high Ni concentration was measured in the water following treatment, indicating that the Ni diffused to the surface and was released under these conditions. When the specimens were initially autoclaved in steam, the Ni concentration in the surface was higher than for those specimens autoclaved in water. (This was assumed to be due to diffusion of Ni to the surface without a medium for it to be released into.) If the samples were steam autoclaved for longer periods, however, the Ni concentration in the surface was reduced. It was suggested that this is due to the tendency of the treatment to increase the thickness of the Ti-oxide layer in preference to Ni-containing oxides.

The corrosion resistance of Ni–Ti alloys can be deliberately influenced by modifying the surface oxide. The effect of electropolishing, heat treatment, and

nitric acid passivation on the surface oxide of a Ni–Ti alloy was examined by Trepanier.^[28] It was demonstrated that the uniformity of the oxide layer was the predominant factor in protecting the NiTi stent from corrosion. Nonuniform oxide layers grown on the plastically deformed surface of mechanically polished specimens were found to be less corrosion-resistant than more homogeneous and uniform surface oxides grown on electropolished specimens. The thickness of the oxide layer was shown to have less influence on the corrosion resistance than the uniformity. In fact, heat-treated specimens tended to have a higher Ni:Ti ratio in the surface oxide (attributed to diffusion of the Ni into the oxide through a defective Ti-oxide layer).

Medical devices have been manufactured from Ni–Ti alloys with considerable success. Two examples are shown in Fig. 4—vascular stents and bone staples. Vascular stents and bone staples take advantage of the shape-memory properties of the alloy. Vascular stents can be cooled and deformed to reduce their diameter, thus allowing them to be introduced into the vascular system via a small-diameter catheter. Once in position, the stent will be released from the inside of a catheter and will warm to body temperature and transform into its austenitic shape. Bone staples are used to increase compressive forces across a fracture, again taking advantage of the shape-memory effect. In its low-temperature condition, the staple is U-shaped. When it is heated, the legs of the staple bend toward each other, pulling the fracture together.

Other devices successfully manufactured from Ni–Ti alloys include orthodontic archwires and springs, which take advantage of the alloy's

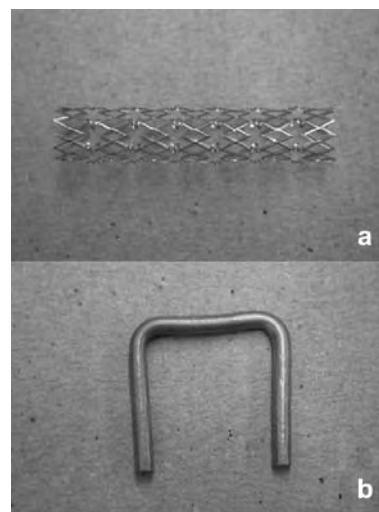


Fig. 4 Ni–Ti medical devices. (a) Vascular stent and (b) bone staple.

pseudoelastic behavior to move teeth more effectively. When a stainless steel wire or spring is used to move teeth, high force is initially applied, but as the teeth move in response to the force, the force applied will be reduced. When using a Ni–Ti wire on the other hand, the force applied to the teeth will be maintained over a greater range of stain, allowing lower force to be initially applied and maintained for a longer period, resulting in improved clinical response.

BIOLOGICAL CONSEQUENCES

From the preceding discussion, it is clear that Ni ions can be released into the biological environment from any of the Ni alloys used in medical applications. Although the amount of Ni ions released may be quite low, it is important to consider the biological consequences of Ni²⁺ accumulation in the body. Many studies have been conducted on the biological response to nickel-containing materials.

It is well known that Ni²⁺ can induce a hypersensitivity reaction. It is reported^[35] that Ni²⁺ is the most common contact allergen, with up to 10% of the Caucasian population having a positive skin test to Ni²⁺. The resulting Ni²⁺-induced allergic contact dermatitis is caused by a delayed-type hypersensitivity. When Ni ions are released into the biological environment they are incomplete antigens owing to their small size. However, they bind readily to the peptides and form complexes to become full antigens. The Ni²⁺ associates in a trimolecular complex of MHC II molecules, peptide, and T-cell receptor to induce a specific T-cell response. Ni²⁺-specific T cells express CD4 and CD8 surface molecules and produce T-helper cytokines (Th1:IL-2 and INF- γ or Th2: IL-4 and IL-5). The presence of Ni²⁺-reactive T cells in peripheral blood or the skin of individuals demonstrating a Ni allergy suggests that these cells are effector cells in Ni²⁺-induced contact sensitivity. The route of administration of the Ni ion is important to the body load of the ion. It has been shown that skin administration leads to a greater uptake of the metal ion and a slower excretion rate through the urine than intramuscular administration. The longer retention of Ni²⁺ in the skin therefore leads to a prolonged period for antigen processing and recognition in dermal tissue.

A link has been investigated between the hypersensitivity reaction induced by Ni²⁺ and the restenosis rate following implantation of Ni-alloy stents. Restenosis following stent implantation results from an excessive fibroproliferative and inflammatory response. It is reported^[11] that immunocompetent cells have been detected (particularly around stent struts), suggestive of a local immune and inflammatory response. It was found that all patients with delayed-type

hypersensitivity to Ni required clinical intervention for in-stent restenosis. Corrosion products from 316L stainless steel wires^[36] have been shown to be toxic to primary smooth muscle cells, and the inhibition of cell growth was correlated with the concentration of nickel ions above 11.7 ppm.

It is known that Ni ions are released into the oral cavity when Ni alloys are used in dental applications. They can be absorbed in the gastrointestinal tract, the oral mucosa and skin, or in the respiratory tract.^[7] The resultant hypersensitivity reaction can manifest as an oral contact allergy demonstrated by redness and swelling of the oral tissues, mucosal erosion, and lichenoid reactions. A series of studies on the release of metal ions from dental materials and their effect on various components of the biological response have been reported.^[37–40] They show that metal ions continue to be released for periods up to 10 months, that a longer-term low-level exposure to monocytes of Ni²⁺ alters cells' metabolism and proliferation, that the Ni ion is taken up rapidly by macrophages, and that it accumulates in the nucleus, possibly influencing its function. A further study investigated whether the mechanism of the recruitment of macrophages and lymphocytes following implantation of Ni-based alloys occurs via a change in the expression of CAMs on endothelial cells. It was found that the influence of Ni ions on the expression of ICAM-1 depends on its concentration; at high concentrations ICAM-1 expression may be promoted, and at low concentrations the expression of ICAM-1 on stimulated cells may be suppressed.

Many studies reported in the literature have investigated the potential toxicity of Ni–Ti alloy in either in vitro cell culture or in vivo^[41–46] environments. Attempts have been made to correlate the biological response to the amount of Ni or Ti released from the specimen over a certain period of time. Some studies relate the cytotoxicity of the NiTi to the surface oxide composition.^[28,29] These demonstrate that pretreatment of the surfaces can have a very significant influence on their cytotoxicity. In particular, surface oxides that contain significant amounts of Ni^[29] appear to release it into the biological environment and cause significant cytotoxicity. Careful pretreatment of Ni–Ti alloys using techniques that increase their passivation can increase their biocompatibility.^[47] Trepanier et al.^[47] evaluated the in vitro fibroblast response in terms of proliferation and the in vivo inflammatory response to heat-treated, electropolished, air-aged, and acid-passivated Ni–Ti stents. It was concluded that in general the in vitro response is very similar to the control group and that after longer implantation, the fibrous capsules surrounding the different implants reduce in thickness to a similar low value.

CONCLUSION

Nickel alloys play a significant role in the production of medical devices, although there is concern over the potential release of Ni ions into the body and the consequent adverse biological response to these ions. Optimization of the material properties in terms of their corrosion resistance allows these alloys to be used with minimal release of nickel. Some experimentation in the design of nickel-free austenitic stainless steels^[9] may overcome this problem for fracture fixation devices; careful pretreatment of Ni–Ti devices can stabilize the Ti-oxide surface and minimize Ni release from these alloys.^[33]

ARTICLES OF FURTHER INTEREST

Corrosion; Orthodontic Wires; Shape Memory Metals

REFERENCES

- Schaller, K.-H.; Raithel, H.-J.; Angerer, J. Nickel. In *Handbook on Metals in Clinical and Analytical Chemistry*; Seiler, H.G., Sigel, A., Sigel, H., Eds.; Marcel Dekker, Inc.: New York, 1994; 505–518.
- Craig, R.G. *Restorative Dental Materials*, 10th Ed.; Mosby-Year Book, Inc.: St. Louis, 1997.
- van Noort, R. *Introduction to Dental Materials*, 2nd Ed.; Mosby: Edinburgh, 2002.
- Test, J.A.; Waterstrat, R.M. Base Metal Alloys for Dental Use. In *Concise Encyclopedia of Medical and Dental Materials*; Williams, D., Ed.; Pergamon Press plc: Oxford, 1990; 41–45.
- Ferracane, J.L. *Materials in Dentistry; Principles and Applications*; J.B. Lippincott Co.: Philadelphia, 1995.
- Breme, J.; Biehl, V. Metallic Biomaterials. In *Handbook of Biomaterial Properties*; Black, J., Hastings, G., Eds.; Chapman & Hall: London, 1998; 135–214.
- Lygre, H. Prosthodontic biomaterials and adverse reactions: A critical review of the clinical and research literature. *Acta Odontol. Scand.* **2002**, *60*, 1–9.
- Sutow, E.J. Iron-Based Alloys. In *Concise Encyclopedia of Medical and Dental Materials*; Williams, D., Ed.; Pergamon Press plc: Oxford, 1990; 232–240.
- Disegi, J.A.; Eschbach, L. Stainless steel in bone surgery. *Injury* **2000**, *31* (Suppl. 4), 2–6.
- Kandzari, D.E.; Tchong, J.E.; Zidar, J.P. Coronary artery stents: Evaluating new designs for contemporary percutaneous intervention. *Catheter. Cardiovasc. Interv.* **2002**, *56*, 562–576.
- Köster, R.; Vieluf, D.; Kiehn, M.; Sommerauer, M.; Kähler, J.; Baldus, S.; Meinertz, T.; Hamm, C. Nickel and molybdenum contact allergies in patients with coronary in-stent restenosis. *Lancet* **2000**, *356*, 1895–1897.
- Gurappa, I. Characterisation of different materials for corrosion resistance under simulated body fluid conditions. *Mater. Charact.* **2002**, *49*, 73–79.
- Melton, K.N. NiTi Based Shape Memory Alloys. In *Engineering Aspects of Shape Memory Alloys*; Duerig, T.W., Melton, K.N., Stöckel, D., Wayman, C.M., Eds.; Butterworth-Heinemann: London, 1990; 21–35.
- Matsumoto, M. Shape memory alloys. *Artif. Organs Today* **1992**, *2*, 1–7.
- Beyer, J. Recent advances in the martensitic transformation of Ti–Ni alloys. *J. Phys., IV* **1995**, *5* (C2), 433–442.
- Hornbogen, E. Lattice Defects and Martensitic Transformation. In *Micromechanics of Advanced Materials: A Symposium in Honor of Professor James Li's 70th birthday*; Chu, S.N.G., Liaw, P.K., Arsenault, R.J., Sadananda, K., Chan, K.S., Gerberich, W.W., Chau, C.C., Kung, T.M., Eds.; The Minerals, Metals & Materials Society, 1995; 307–315.
- Filip, P.; Musialek, J.; Lorethova, H.; Nieslanik, J.; Mazanec, K. TiNi shape memory clamps with optimized structure parameters. *J. Mater. Sci., Mater. Med.* **1996**, *7*, 657–663.
- Filip, P.; Pech, J.; Mazanec, K. Physical metallurgy of TiNi shape memory alloys applied for dynamic splints. *BHM* **1994**, *139*, 174–178.
- Otsuka, K. Introduction to the R-Phase Transition. In *Engineering Aspects of Shape Memory Alloys*; Duerig, T.W., Melton, K.N., Stöckel, D., Wayman, C.M., Eds.; Butterworth-Heinemann: London, 1990; 36–45.
- Rondelli, G.; Vicentini, B.; Cigada, A. The corrosion behaviour of nickel titanium shape memory alloys. *Corros. Sci.* **1990**, *30*, 805–812.
- Rondelli, G. Corrosion resistance tests on NiTi shape memory alloy. *Biomaterials* **1996**, *17*, 2003–2008.
- Melton, K.N.; Harrison, J.D. Corrosion of NiTi Based Shape Memory Alloys. In *Proceedings of the First International Conference on Shape Memory and Superelastic Technologies*; Pelton, A.R., Hodgson, D., Duerig, T., Eds.; Pacific Grove: California, 1994; 187–196.
- Abiko, Y.; Sachdeva, R.; Endo, K.; Araki, Y.; Kaku, T.; Ohno, H. Corrosion Resistance and Biological Evaluation of NiTi Alloys with Varied Surface Textures. In *Proceedings of the First International Conference on Shape Memory and Superelastic Technologies*; Pelton, A.R., Hodgson, D., Duerig, T., Eds.; Pacific Grove: California, 1994; 203–209.
- Rondelli, G.; Vicentini, B. Localized corrosion behaviour in simulated human body fluids of commercial NiTi orthodontic wires. *Biomaterials* **1999**, *20*, 785–792.
- Montero-Ocampo, C.; Lopez, H.; Salinas Rodriguez, A. Effect of compressive straining on corrosion resistance of a shape memory NiTi alloy in ringer's solution. *J. Biomed. Mater. Res.* **1996**, *32*, 583–591.
- Dutta, R.S.; Madangopal, K.; Gadiyar, H.S.; Banerjee, S. Biocompatibility of NiTi shape memory alloy. *Br. Corros. J.* **1993**, *28*, 217–221.
- Endo, K.; Sachdeva, R.; Araki, Y.; Ohno, H. Corrosion Behaviour of NiTi Shape Memory Alloy in a Cell Culture Medium. In *Proceedings of the First International Conference on Shape Memory and Superelastic*



- Technologies*; Pelton, A.R., Hodgson, D., Duerig, T., Eds.; Pacific Grove: California, 1994; 197–201.
28. Trepanier, C.; Tabrizian, M.; Yahia, L'H.; Bilodeau, L.; Piron, D. Effect of modification of oxide layer on NiTi stent corrosion resistance. *J. Biomed. Mater. Res.* **1998**, *43*, 433–440.
 29. Shabalovskaya, S.A. On the nature of the biocompatibility and on medical applications of NiTi shape memory and superelastic alloys. *Biomed. Mater. Eng.* **1996**, *6*, 267–289.
 30. Wever, D.J.; Veldhuizen, A.G.; de Vries, J.; Busscher, H.J.; Uges, D.R.A.; van Horn, J.R. Electrochemical and surface characterisation of nickel–titanium alloy. *Biomaterials* **1998**, *19*, 761–769.
 31. Shih, C.-C.; Lin, S.-J.; Chung, K.-h.; Chen, Y.-L.; Su, Y.-Y. Increased corrosion resistance of stent materials by converting current surface film of polycrystalline oxide into amorphous oxide. *J. Biomed. Mater. Res.* **2000**, *52*, 323–332.
 32. Thierry, B.; Tabrizian, M.; Savadogo, O.; Yahia, L'H. Effects of sterilization processes on NiTi alloy: Surface characterisation. *J. Biomed. Mater. Res.* **2000**, *49*, 88–98.
 33. Thierry, B.; Tabrizian, M.; Trepanier, C.; Savadogo, O.; Yahia, L'H. Effect of surface treatment and sterilization processes on the corrosion behaviour of NiTi shape memory alloy. *J. Biomed. Mater. Res.* **2000**, *51*, 685–693.
 34. Li, C.; Wu, K.H. Corrosion Behaviour of NiTi Shape Memory Alloy in Artificial Seawater. In *Proceedings of the First International Conference on Shape Memory and Superelastic Technologies*; Pelton, A.R., Hodgson, D., Duerig, T., Eds.; Pacific Grove: California, 1994; 227–232.
 35. Büdinger, L.; Hertl, M. Immunological mechanisms in hypersensitivity reactions to metal ions: An overview. *Allergy* **2000**, *55*, 108–115.
 36. Shih, C.-C.; Shih, C.-M.; Chen, Y.-L.; Su, Y.-Y.; Shih, J.-S.; Kwok, C.-F.; Lin, S.-J. Growth inhibition of cultured smooth muscle cells by corrosion products of 316L stainless steel wire. *J. Biomed. Mater. Res.* **2001**, *57*, 200–207.
 37. Edwards, D.L.; Wataha, J.C.; Hanks, C.T. Uptake and reversibility of uptake of nickel by human macrophages. *J. Oral Rehab.* **1998**, *25*, 2–7.
 38. Wataha, J.C.; Lockwood, P.E.; Schedle, A.; Noda, M.; Bouillaguet, S. Ag, Cu, Hg and Ni ions alter the metabolism of human monocytes during extended low-dose exposures. *J. Oral Rehab.* **2002**, *29*, 133–139.
 39. Wataha, J.C.; Lockwood, P.E. Release of elements from dental casting alloys into cell-culture medium over 10 months. *Dent. Mater.* **1998**, *14*, 158–163.
 40. Wataha, J.C.; Sun, Z.L.; Hanks, C.T.; Fang, D.N. Effect of Ni ions on expression of intercellular adhesion molecule 1 by endothelial cell. *J. Biomed. Mater. Res.* **1997**, *36*, 145–151.
 41. Takeshita, F.; Takata, H.; Ayukawa, Y.; Suetsugu, T. Histomorphometric analysis of the response of rat tibiae to shape memory alloy (nitinol). *Biomaterials* **1997**, *18*, 21–25.
 42. Ryhänen, J.; Niemi, E.; Serlo, W.; Niemelä, E.; Sandvik, P.; Pernu, H.; Salo, T. Biocompatibility of nickel–titanium shape memory metal and its corrosion behaviour in human cell cultures. *J. Biomed. Mater. Res.* **1997**, *35*, 451–457.
 43. Wever, D.J.; Veldhuizen, A.G.; Sanders, M.M.; Schakenraad, J.M.; van Horn, J.R. Cytotoxic, allergic and genotoxic activity of a nickel–titanium alloy. *Biomaterials* **1997**, *18*, 1115–1120.
 44. Assad, M.; Yahia, L.H.; Rivard, C.H.; Lemieux, N. In vitro biocompatibility assessment of a nickel–titanium alloy using electron microscopy in situ. *J. Biomed. Mater. Res.* **1998**, *41*, 154–161.
 45. Ryhänen, J.; Kallioinen, M.; Tuukkanen, J.; Junila, J.; Niemelä, E.; Sandvik, P.; Serlo, W. In vivo biocompatibility evaluation of nickel–titanium shape memory metal alloy: Muscle and perineural tissue responses and capsule membrane thickness. *J. Biomed. Mater. Res.* **1998**, *41*, 481–488.
 46. Ryhänen, J.; Kallioinen, M.; Tuukkanen, J.; Lehenkari, P.; Junila, J.; Niemelä, E.; Sandvik, P.; Serlo, W. Bone modeling and cell–material interface responses induced by nickel–titanium shape memory alloy after periosteal implantation. *Biomaterials* **1999**, *20*, 1309–1317.
 47. Trepanier, C.; Leung, T.K.; Tabrizian, M.; Yahia, L'H.; Bienvenu, J.-G.; Tanguay, J.-F.; Piron, D.L.; Bildeau, L. Preliminary investigation of the effects of surface treatments on biological response to shape memory NiTi stents. *J. Biomed. Mater. Res. (Appl. Biomater.)* **1999**, *48*, 165–171.

Nitric Oxide

Matthew J. Bizzarro
Vineet Bhandari

*Department of Pediatrics, Yale University School of Medicine and Children's Hospital,
New Haven, Connecticut, U.S.A.*

INTRODUCTION

The first synthesis of nitric oxide (NO) in the laboratory has been credited to Jan Baptista van Belmont in 1610. In 1772, Joseph Priestly studied some of the molecule's actions and called it "nitrous air." In 1987, NO was discovered to be identical to endothelial-derived relaxing factor (EDRF). The role of NO as a unique signalling molecule in the vascular system is well known. It was also soon appreciated that NO had a diverse range of physiological and pathological actions. From then on, research on NO has been exponential in nature and has expanded our knowledge of this molecule as exemplified by over 50,500 citations in the literature.

SYNTHESIS AND CHEMISTRY OF NITRIC OXIDE

The realization that NO was identical to EDRF^[1-3] has spurred the understanding of the production, release, and effects of NO. L-arginine is the substrate for the three known NO synthases (NOS), which result in the production of L-citrulline and the generation of NO. NOS1 was the first enzyme to be purified and, as a result of its extraction from neural tissue, was dubbed neuronal NOS (nNOS). NOS2, also called inducible NOS (iNOS), is a calcium-independent isoform present in leukocytes. NOS3 is the isoform first derived from endothelial cells—hence endothelial NOS (eNOS). NOS1 and NOS3 are termed constitutive since they are expressed continuously in neurons and endothelial cells, respectively.^[4] Unlike NOS2, NOS1 and NOS3 are dependent on a rise in tissue calcium concentration for activity.^[4] However all the enzymes require nicotinamide adenine dinucleotide phosphate (NADPH), flavin mononucleotide (FMN), flavin adenine dinucleotide (FAD), and tetrahydrobiopterin (H₄B) as cofactors.^[5]

Induction of NOS2 results in continuous production of NO because of tight calmodulin binding, which allows electrons to pass from the heme group to the flavin (FMN, FAD) and nicotinamide (NADPH) electron-accepting cofactors.^[4] In contrast, induction of NOS1 and NOS3 results in production of lower, transient concentrations of NO.^[4]

The arrangement of one atom of nitrogen and one of oxygen leaves an unpaired electron, which makes NO a highly reactive free radical.^[4] NO readily reacts with superoxide, oxygen, and thiol groups to produce a number of other products, such as nitrosothiols, peroxynitrite, and nitrites.^[4] The interaction between NO and superoxide ion leads to the generation of peroxynitrite, which is a powerful oxidant. Peroxynitrite is very toxic through nitrosation and/or nitration of amino acids such as tyrosine and cysteine on various proteins.^[6-8] Such modifications alter protein function and disrupt cellular activity.^[6-8] Peroxynitrite is particularly toxic to DNA and enzymes involved in DNA repair.^[9] Comprehensive reviews of NO chemistry, biogenesis, and regulation have been published.^[9,10]

MECHANISM OF ACTION OF NITRIC OXIDE

NO diffuses freely from its site of production to mediate biologic actions in adjacent target tissues and has a half-life in the order of seconds. For example (Fig. 1), NO produced by endothelial cells diffuses into surrounding smooth muscle where it induces relaxation of vascular smooth muscle and vasodilation.^[11] NO also diffuses into the flowing blood and is inactivated by its reaction with the iron-containing heme groups of oxyhemoglobin, which produces methemoglobin and nitrate ions.^[11] Despite this deactivation of NO, enough bioactivity remains to exert control over vascular tone.^[11] The rapid destruction of NO by hemoglobin raises the issue of the molecule having only local effects (paracrine) or the possibility that NO can disseminate throughout the body. If NO acts like the latter, as a hormone, then it could have potential for significant pharmacologic use.^[11]

It has been suggested by some investigators that S-nitrosohemoglobin is a source of bioactive NO.^[12] S-nitrosohemoglobin is formed when NO reacts with the cysteine at position 93 of the beta chain of hemoglobin.^[11] It has been proposed that S-nitrosohemoglobin forms in tissues at high oxygen concentrations, whereas at low oxygen concentrations, it decomposes and releases NO.^[11] However, other investigators have questioned this mode of action.^[13,14]



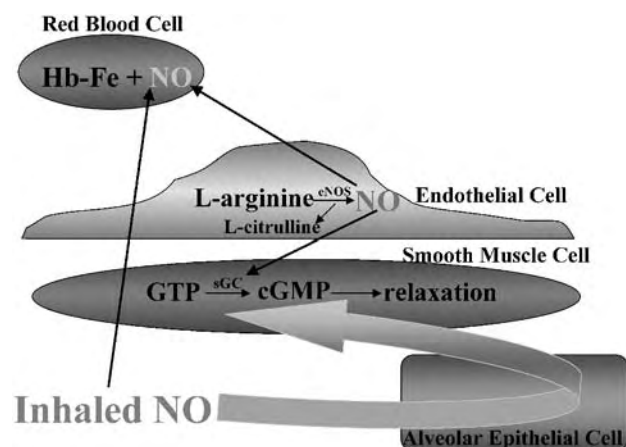


Fig. 1 Synthesis of endogenous NO and its interaction with iNO and pulmonary vascular smooth muscle. This figure depicts the synthesis of endogenous nitric oxide within the vascular endothelial cell. L-arginine and oxygen are converted to nitric oxide (NO) and L-citrulline by the enzyme endothelial nitric oxide synthase (eNOS). NO then diffuses out of the endothelial cell and into the vascular smooth muscle. Here, NO binds with the enzyme soluble guanylate cyclase (sGC) to convert guanosine triphosphate (GTP) to cyclic guanosine 3',5'-monophosphate (cGMP). cGMP is then responsible for mediating relaxation of the smooth muscle. NO can also diffuse out of endothelial cell and into the blood stream, where it can combine with the iron in hemoglobin (Hb-Fe) to form methemoglobin and eventually nitrite and nitrate. The figure also depicts the interaction of iNO with vascular smooth muscle. iNO diffuses through the alveolar epithelial cell and into the smooth muscle cell, where it exerts the same action as endogenous NO. The molecule then diffuses into the blood stream, where it interacts with the red blood cells and is deactivated via the same pathway described previously. The net effect of both exogenous and endogenous NO is vascular smooth muscle relaxation and pulmonary vasodilation.

Other models include the possibility that NO remains active in the plasma longer than previously assumed or it can react with plasma proteins in ways that stabilize and protect it.^[15] Moreover, iron-nitrosyl hemoglobin, nitrite ions, and the erythrocyte membrane are also potential sources of NO under certain conditions.^[11]

The physiologic actions of NO can be divided into two main groups. The first occurs as a result of the molecule's interaction with the heme group in the enzyme guanylate cyclase, which results in the activation and generation of the second messenger cyclic guanosine monophosphate (cGMP). The physiologic effects of this pathway include regulation of vascular tone and neuromodulation of the central nervous system. The second grouping of actions are cGMP-independent.^[4] These play a pivotal role in angiogenesis especially through interactions with vascular endothelial growth factor (VEGF).^[4,16] (Table 1).

Table 1 Effects of endogenous and exogenous nitric oxide

Endogenous NO produced via neurons; neuronal, Type 1 NOS:

Acts as a neurotransmitter through regulation of:

Nonadrenergic, noncholinergic function resulting in reduction of smooth muscle tone in:

- Respiratory tract
- Gastrointestinal tract
- Genitourinary tract

Information storage and memory

Pain perception

Cortical and retinal development

Catecholamine release in the cardiovascular system

Endogenous NO produced via leukocytes; inducible, Type 2 NOS:

Inhibition of neutrophil adhesion and release of inflammatory mediators

Activation of cytokines

Promotion of leukocyte demargination

Promotion of neovascularization

Endogenous NO produced via endothelial cells; endothelial, Type 3 NOS:

Reduction of pulmonary and systemic vascular tone

Inhibition of cellular proliferation

Inhibition of platelet adhesion and activation

As research has uncovered more and more physiologic effects of endogenous NO, clinicians and investigators have sought the discovery and production of an exogenous form to be implemented in the treatment of various disease states. This eventually came in the form of inhaled NO (iNO). This review will highlight the major physiological effects of iNO and focus on its use and potential for use in different disease conditions as well as in the field of biomedical engineering.

PHYSIOLOGIC EFFECTS OF INHALED NITRIC OXIDE

The use of iNO has been implemented in many diverse clinical scenarios. To fully comprehend the therapeutic rationale behind the use of iNO in medicine, its most beneficial physiologic effects must first be reviewed (Table 2).

PULMONARY VASODILATION

Like endogenous NO, exogenous iNO is a potent, but selective, pulmonary vasodilator. Within the endothelial

Table 2 Substantiated and proposed^a effects of exogenous, inhaled NO

Selective pulmonary vasodilation
Bronchodilation ^a
Inhibition of neutrophil aggregation ^a
Inactivation of inflammatory mediators ^a
Increased oxygen affinity of hemoglobin S ^a

cells of the pulmonary vascular bed, NO is synthesized, diffuses into, and reacts with the cells of the vascular smooth muscle, which results in a reduction in muscular tone and a subsequent vasodilation.^[17] Once iNO diffuses out of the pulmonary vascular smooth muscle, it is inactivated and thus cannot exert its effects on the systemic vascular resistance.^[17]

PROTECTION AGAINST CELLULAR TOXICITY

NO may also play a role in preventing cellular damage from oxygen-free radicals and cellular-toxic oxidants.^[18] Although, in high concentrations iNO has been shown to cause cellular damage, some animal models have demonstrated that low-dose iNO may instead protect the lung from injury.^[19] This is mediated through NO's reaction with—and inactivation of—H₂O₂, O₂⁻ and alkyl peroxides, all of which are directly injurious to pulmonary tissue.^[20,21] Local injury can trigger a systemic inflammatory response, which results in further damage to the lung. iNO may diminish the damage by inhibiting neutrophil migration, adhesion, and release of damaging oxidants into the alveoli.^[22,23]

THE USE OF INHALED NITRIC OXIDE IN THE NEONATE

Although iNO has several clinical uses in the treatment of newborn infants, its most significant impact has been made in the treatment of persistent pulmonary hypertension of the newborn (PPHN). PPHN is a life-threatening condition that occurs in about 1 in 1000 live births^[24] and is caused by a sustained increase in pulmonary vascular resistance due to any of a number of diverse etiologic possibilities. Conventional therapies in the treatment of PPHN are aimed at promoting systemic to pulmonary blood flow. Despite these maneuvers, a significant number of infants fail conventional treatment and require the use of extracorporeal membrane oxygenation (ECMO), a procedure with its own associated morbidities.

The discovery of iNO's action as a selective pulmonary vasodilator seemed the perfect therapeutic

intervention for PPHN and fueled several large, randomized controlled trials.^[25–28] The results of these trials revealed that, whereas use of iNO in term neonates with PPHN did not improve their overall rate of survival, it did result in a significant improvement in arterial oxygenation. Its use also effectively reduced the need for ECMO. As a result, the implementation of iNO in newborns with PPHN has become part of standard therapeutic intervention and remains the only FDA-approved application of the inhaled medication to date.

Another area in which iNO therapy plays a significant role is in the treatment of neonatal congenital heart disease (CHD). It is estimated that approximately 0.8 to 1% of all newborns have some form of CHD.^[29] Several forms of CHD may result in cyanosis and hypoxemia in the neonate, which result in a secondary pulmonary hypertension. It would seem logical that iNO would therefore be an effective therapeutic intervention. However, decreasing pulmonary vascular resistance in the setting of CHD associated with poor left ventricular function, i.e., hypoplastic left heart, can have drastic consequences and expedite cardiac failure.^[30] Alternatively, iNO can have beneficial therapeutic effects for patients with right ventricular outflow tract obstruction and severe right to left shunting through the improvement of pulmonary blood flow.^[17,30] Therefore, it is essential that proper anatomic definition of the underlying structural defect be made before implementation of iNO is considered.

When used in the correct setting, iNO can have significant therapeutic benefit to the CHD patient. Cardiac surgery often necessitates the use of cardiopulmonary bypass (CPB). CPB can result in a secondary pulmonary hypertension in the newborn as a result of microemboli, excess thromboxane and endothelin production, platelet aggregation, and hypoxic pulmonary vasoconstriction.^[30] This complication, combined with baseline elevated pulmonary vascular resistance as a result of the underlying cardiac anomaly can result in life-threatening pulmonary hypertensive crises in the postoperative period. iNO may therefore be beneficial in treating this process as described in several randomized controlled trials.^[31–33] In all, the use of postoperative iNO has been shown to facilitate postoperative management and reduce the need for postoperative ECMO such that more patients may be eligible for corrective surgical repair.^[34,35]

In addition to PPHN and CHD, much research has focused recently on the use of iNO in premature infants. In particular, there is some discussion involving the possible role of iNO in the treatment of respiratory distress syndrome (RDS). iNO may enhance pulmonary blood flow and improve the ventilation-perfusion mismatching that is the hallmark of this disease process. Studies on both animal models and premature



neonates have yielded similar results showing significant increases in arterial oxygenation without added morbidity and yet also without significant reduction in mortality.^[36-38] Despite these findings, no current recommendations regarding the use of iNO in premature infants have been made.

THE USE OF INHALED NITRIC OXIDE IN CHILDREN

ARDS, or acute respiratory distress syndrome, is an inflammatory-mediated process triggered by some localized pulmonary or systemic insult. Despite optimal conventional management, the overall mortality rate for ARDS still remains about 30–50%.^[39,40] iNO, as the result of its documented action as a pulmonary vasodilator, its potential as an inhibitor of neutrophil activation and cellular injury, and its possible role in reducing pulmonary edema through decreasing pulmonary capillary pressure and the transvascular movement of albumin,^[41,42] would seem a likely therapeutic agent for use in ARDS. Clinical trials in the pediatric population revealed that, although iNO was somewhat effective in reducing pulmonary arterial pressure and improving oxygenation, there was no significant improvement in overall mortality.^[17,40,43]

The use of iNO in the pediatric population is also seen in the presence of CHD. Its use as a pulmonary vasodilator is utilized in the postoperative period for reasons similar to that described in the neonate.^[31-33] However, in the pediatric population, iNO also plays a significant diagnostic role in evaluating patients as potential surgical candidates.

In children with pulmonary overcirculation, progressive increased pulmonary vascular resistance may result in irreversible pulmonary hypertension.^[17] As a result, these infants may not be candidates for certain corrective surgical procedures.^[30] iNO can be used in the preoperative period to assess pulmonary vasoreactivity and thus separate patients who may benefit from surgery from those in whom surgery would be contraindicated.

There are other proposed clinical uses of iNO in the pediatric population that are still being investigated. One is in association with sickle cell disease (SCD). SCD results in production of defective hemoglobin, termed hemoglobin S (HbS), which has a low affinity for oxygen. It is the resulting deoxygenation of HbS that causes the erythrocyte to take on an abnormal shape, which results in hemolysis and vasoocclusion. iNO has been investigated as a potential therapy in SCD based on data that has shown that iNO may increase the oxygen affinity of HbS and thus decrease the potential for SCD crises.^[44]

Another area of particular interest in the neonatal and pediatric populations is the use of iNO in patients with septic shock. Animal studies have yielded data that supports the use of iNO in this scenario through its actions in limiting the effects of leukocyte sequestration and subsequent inflammation in addition to inhibiting platelet-activating factor induced septic shock.^[45-47] Alternatively, results of a prospective study in 15 neonates and infants being treated with iNO showed that the use of iNO decreased neutrophil production of superoxide anion, a substance necessary for host defense against microorganisms.^[48] Therefore, iNO may put the pediatric patient with an already compromised immune system at an increased risk for developing infection. As a result of this conflicting data, no recommendations have been made for or against the use of iNO in patients with or at risk for sepsis.

THE USE OF INHALED NITRIC OXIDE IN ADULTS

ARDS is a disease process that also affects adults. One large multicenter trial enrolled and randomized 177 patients to receive either placebo or iNO. Patients in the iNO group, as a whole, achieved significant improvement in arterial oxygen saturations and achieved lower mean pulmonary arterial pressures as compared with the placebo group. However, no significant difference in mortality was seen.^[39]

iNO also plays a role in adults in management of patients undergoing cardiac surgery. Like in children, adults often experience a transient increase in pulmonary arterial pressure believed secondary to damage induced by CPB.^[18] iNO has been shown to significantly reduce pulmonary vascular tone in this setting and decrease the need for ECMO.^[49] iNO may also play a role in reducing pulmonary arterial pressure in patients with long-standing left atrial pressure from valvular disease.^[18]

iNO has been implemented in the adult population in an attempt to treat asthma and COPD. It is believed that endogenous NO also plays a role as a neurotransmitter in controlling smooth muscle tone in the bronchial airways.^[50] Some animal models revealed that iNO may act as an effective bronchodilator, comparable to β_2 -agonists, in bronchospastic disease.^[51] However, when used in humans, the results were less encouraging. In fact, iNO (80 ppm) appeared to have no bronchodilatory effect on patients with COPD and only a minimal effect on relaxation of airway smooth muscle in patients with asthma.^[52] However, it is believed that higher doses of iNO may play some role in bronchodilation, but the data is currently unavailable.^[53]

THE USE OF NO IN BIOMEDICAL ENGINEERING

In addition to its use in clinical medicine, the use of NO has been extensively investigated and implemented in the field of biomedical engineering. NO-generated biomaterials and NO-modified tissue production have utilized the action of the molecule as an inhibitor of platelet-activating factor and inflammation to prevent operative and postoperative complications.

USE OF NO GAS INFUSION

ECMO procedures have been shown to induce complement and leukocyte activation and release of endotoxin, inflammatory mediators, and platelet-activating factors. Heparin-coated bypass circuits have been implemented in an attempt to improve the biocompatibility of ECMO.^[54,55] In an in-vitro system to study the effect of NO on platelet deposition on biomaterials, gaseous NO was delivered via a semi-permeable membrane to a radio-labeled platelet suspension.^[56] Platelet inhibition was essentially complete at 0.1 ppm gaseous NO exposure.^[56] NO gas infusion to the oxygenator circuit enhanced the biocompatibility of heparin-coated ECMO circuits.^[55] In a study comparing the effects of heparin-coated bypass circuit and NO gas infusion, NO (100 ppm) was equivalent to the heparin-coated circuit for attenuating bypass-induced blood activation, and a combination of the two significantly surpassed the results of either modification alone.^[57]

USE OF NO-GENERATING BIOMATERIALS

The success of many intravascular medical devices has depended on favorable biocompatibility of the blood-contacting surfaces of such implants.^[58] A great deal of research has been devoted to developing polymeric materials with improved thromboresistance.^[58] As mentioned above, NO is useful in prevention of platelet adhesion and activation. Thus, one attractive option would be the use of slow NO-releasing hydrophobic polymer materials.^[58] Different approaches to preparing polyurethane and polyvinyl chloride polymer films containing NO-releasing diazeniumdiolate have been described.^[58] These films release NO for variable periods (10–72 hours) and differ in the mechanism, location, and amount of NO released.^[58] In one study, use of the NO-releasing diazeniumdiolate, spermine/NO, produced a marked localized inhibition of neointimal proliferation in balloon-injured arteries in a rat model.^[59] In contrast, tantalum wire coil coronary stents coated with polycaprolactone impregnated with the diethylenetriamine adduct of

nitric oxide, (Z)-1-[N-(2-aminoethyl)-N-(2-aminoethyl)amino]diazene-1-ium-1,2-diolate did not prevent restenosis in a pig balloon injury model.^[60] The investigators speculated that these findings could have been due to a lack of efficacy of NO or a profuse and overwhelming stimulatory effect of the polymer in the coated stents.^[60] However, incorporation of the NO-releasing polymer [NO donor, (Z)-1-[N-Methyl-N-[6-(N-methylammoniohexyl)amino]]diazene-1-ium-1,2-diolate (MAHMA/N₂O₂)] coated onto the inner surface of ECMO circuits reduced platelet consumption and eliminated the need for systemic heparinization in a rabbit model.^[61]

Another approach utilized three NO donors with different half-lives by covalent incorporation into photopolymerized polyethylene glycol hydrogels.^[62] Under physiological conditions, NO was produced by these hydrogels over periods ranging from hours to months, depending upon the polymer formulation.^[62] NO production was inhibited at acidic pH, which may be useful for storage of these materials.^[62] The NO releasing materials successfully inhibited smooth muscle cell proliferation in culture and prevented platelet adhesion to collagen-coated surfaces.^[62]

NO-releasing silicone rubbers have been developed that have been shown to exhibit improved thromboresistance in vivo in an ECMO circuit in a rabbit model.^[63] Use of amperometric, oxygen-sensing catheters coated with such NO-releasing silicone rubber polymer (DACA/N₂O₂ SR) in vivo (in a swine model) have shown to have accurate measurements of pO₂ levels with the sensors exhibiting no significant platelet adhesion or thrombus formation.^[64] In comparison, non-NO-coated sensors had a high propensity for cell adhesion and bulk clot formation.^[64]

Other investigators have utilized L-arginine by immobilizing it into a supramolecular-structured polyrotaxane.^[65] L-arginine-immobilized polyrotaxane exhibited sustained NO-generation for a period of 250 hours.^[65] Consequently, after the generation of NO and the release of L-arginine from the L-arginine-immobilized polyrotaxane, the residual component would be a polyrotaxane with superior biocompatibility and mechanical properties.^[65]

USE OF NO-GENERATING TISSUE VIA BIOMEDICAL ENGINEERING

The seeding of the blood-contacting surfaces of cardiovascular prostheses with autologous endothelial cells to improve their biocompatibility had little success as the cells subsequently sloughed off under high-flow conditions. On the other hand, left ventricular assist devices (LVADs) have shown some efficacy. Use of genetically engineered smooth muscle



cells (GE-SMC) transduced with NOS3 gene have resulted in viable cells that produce NO at concentrations that reduce platelet deposition and SMC proliferation when tested in vitro.^[66,67] An in-vitro circulatory loop was designed to expose cell-lined LVADs to in-vivo operating conditions.^[67] Cumulative cell loss was less than 10% after 24 hours of flow, and, based on these results, the first implantation of an LVAD containing GE-SMC lining was successfully implemented in a bovine model.^[67]

Using adenoviral vectors, gene transfer of NOS3 has been achieved.^[68-70] Using a different (nonviral) approach, local intramural delivery of NOS2 lipoplexes have been shown to exert therapeutic effects in inhibiting stent-induced neointimal lesion formation.^[71] Other potential uses include that of a hybrid artificial lung with interleukin-10 and NOS3 gene-transfected endothelial cells inhibiting interleukin-8 release and increased NO production,^[72] NOS3 gene transfection of saphenous vein grafts with naked DNA or liposomal vector,^[73] and cell-based NOS3 gene transfer to pulmonary artery SMCs to combat pulmonary hypertension.^[74] NOS gene transfer has also shown promise as a potential treatment of erectile dysfunction.^[75-77]

CONCLUSION

Endogenous NO has many physiologic effects on various organ systems in the human body. As a result, investigators have sought to develop and incorporate the medical implementation of exogenous NO in the newborn, pediatric, and adult populations. iNO has been sought as one such therapy due to its selective effect on reducing pulmonary arterial pressure. To date, however, FDA-approved use of inhaled nitric oxide remains limited to PPHN despite some data to support its potential use in other clinical scenarios. Further research is needed though in the form of large, randomized controlled trials to evaluate its efficacy in these and other, possibly yet undiscovered, areas of medicine.

In biomedical engineering, the usefulness of NO gas in inhibition of platelet activation and inflammation in ECMO circuits has sparked the development of NO-generating biomaterials to improve the biocompatibility of intravascular devices. Another active area of research is in the use of NOS gene transfer via different mechanisms to endothelial cells and/or SMCs to achieve the same purpose. The potential for using the NOS gene transfer approach in certain cardiovascular diseases and erectile dysfunction is actively being pursued.

It is evident that discovery of the physiologic actions of endogenous NO has initiated extensive research in

an attempt to apply this physiology to the treatment of disease processes. Modification and utilization of endogenous and exogenous NO has resulted in a great deal of progress both in clinical medicine and biomedical engineering. While more information to the potential uses of NO is being uncovered, we eagerly await the possibility of further therapeutic implementations.

ARTICLES OF FURTHER INTEREST

Angiogenesis Inhibitors; Biofunctional Polymers; Bone Remodeling; Integrins

REFERENCES

1. Palmer, R.M.; Ferrige, A.G.; Moncada, S. Nitric oxide release accounts for the biological activity of endothelial-derived relaxing factor. *Nature* **1987**, *327* (6122), 524-526.
2. Ignarro, L.J.; Buga, G.M.; Wood, K.S.; Byrns, R.E.; Chaudhuri, G. Endothelium-derived relaxing factor produced and released from artery and vein is nitric oxide. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84* (24), 9265-9269.
3. Ignarro, L.J. Nitric oxide as a unique signaling molecule in the vascular system: A historical overview. *J. Physiol. Pharmacol.* **2002**, *53* (4 part 1), 503-514.
4. Brennan, P.A.; Thomas, P.J.; Langdon, J.D. The role of nitric oxide in oral diseases. *Arch. Oral Biol.* **2003**, *48* (2), 93-100.
5. Schwentker, A.; Billiar, T.R. Inducible nitric oxide synthase: From cloning to therapeutic applications. *World J. Surg.* **2002**, *26* (7), 772-778.
6. Hobbs, A.J.; Higgs, A.; Moncada, S. Inhibition of nitric oxide synthase as a potential therapeutic target. *Annu. Rev. Pharmacol. Toxicol.* **1999**, *39*, 191-220.
7. Bhandari, V. The role of nitric oxide in hyperoxia-induced injury to the developing lung. *Front. Biosci.* **2003**, *8*, e361-369.
8. Bhandari, V.; Johnson, L.; Kirwin-Smith, S.; Vigliotta, G.; Funanage, V.; Chandar, A. Hyperoxia and nitric oxide reduce surfactant components (DSPC and surfactant proteins) and increase apoptosis in adult and fetal rat type II pneumocytes (TIIP). *Lung* **2002**, *180* (6), 301-317.
9. Wink, D.A.; Vodovotz, Y.; Laval, D.; Laval, F.; Dewhirst, D.W.; Mitchell, J.B. The multifaceted roles of nitric oxide in cancer. *Carcinogenesis* **1998**, *19* (5), 711-721.
10. Bian, K.; Murad, F. Nitric oxide (NO)- biogenesis, regulation, and relevance to human diseases. *Front. Biosci.* **2003**, *8*, D264-D278.
11. Schechter, A.N.; Gladwin, M.T. Hemoglobin and the paracrine and endocrine functions of nitric oxide. *N. Engl. J. Med.* **2003**, *348* (15), 1483-1485.
12. McMahon, T.J.; Moon, R.E.; Lusching, B.P.; Carraway, M.S.; Stone, A.E.; Stolp, B.W.; Gow, A.J.;

- Pawloski, J.R.; Watke, P.; Singel, D.J.; Piantadosi, C.A.; Stamler, J.S. Nitric oxide in the human respiratory cycle. *Nat. Med.* **2002**, *8* (7), 711–717.
13. Joshi, M.S.; Ferguson, T.B., Jr.; Han, T.H.; Hyduke, D.R.; Liao, J.C.; Rassaf, T.; Bryan, N.; Feelisch, M.; Lancaster, J.R., Jr. Nitric oxide is consumed, rather than conserved, by reaction with oxyhemoglobin under physiologic conditions. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (16), 10341–10346.
 14. Gladwin, M.T.; Wang, X.; Reiter, C.D.; Yang, B.K.; Vivas, E.X.; Bonaventura, C.; Schechter, A.N. S-nitrosohemoglobin is unstable in the reductive erythrocyte environment and lacks O₂/NO-linked allosteric function. *J. Biol. Chem.* **2002**, *277* (31), 27818–27828.
 15. Rassaf, T.; Preik, M.; Kleinbongard, P.; Lauer, T.; Heiss, C.; Strauer, B.E.; Feelisch, M.; Kelm, M. Evidence for in vivo transport of bioactive NO in human plasma. *J. Clin. Invest.* **2002**, *109* (9), 1241–1248.
 16. Bhandari, V.; Link, H.; Lee, C.; Tang, C.; Baluk, P.; McDonald, D.; Homer, R.; Elias, J. Transgenic VEGF causes hemorrhage, pulmonary edema, and mucus metaplasia and enhances survival in hyperoxia via a nitric oxide-dependent pathway(s). *Am. J. Resp. Crit. Care Med.* **2003**, *167*, A563.
 17. Nelin, L.D.; Hoffman, G.M. The use of inhaled nitric oxide in a wide variety of clinical problems. *Pediatr. Clin. North. Am.* **1998**, *45* (3), 531–548.
 18. Steudel, W.; Hurford, W.E.; Zapol, W.M. Inhaled nitric oxide. Basic biology and clinical applications. *Anet* **1999**, *91* (4), 1090–1121.
 19. Wink, D.A.; Cook, J.A.; Pacelli, R.; Liebmann, J.; Krishna, M.C.; Mitchell, J.B. Nitric oxide (NO) protects against cellular damage by reactive oxygen species. *Toxicol Lett.* **1995**, *82–83*, 221–226.
 20. Poss, W.B.; Timmons, O.D.; Farrukh, I.S.; Hoidal, J.R.; Michael, J.R. Inhaled nitric oxide prevents the increase in pulmonary vascular permeability caused by hydrogen peroxide. *J. Appl. Physiol.* **1995**, *79* (3), 886–891.
 21. Kavanagh, B.P.; Mouchawar, A.; Goldsmith, J.; Pearl, R.G. Effects of inhaled NO and inhibition of endogenous NO synthesis in oxidant-induced acute lung injury. *J. Appl. Physiol.* **1994**, *76* (3), 1324–1329.
 22. Guidot, G.M.; Repine, M.J.; Hybertson, B.M.; Repine, J.E. Inhaled nitric oxide prevents neutrophil-mediated, oxygen radical-dependent leak in isolated rat lungs. *Am. J. Physiol.* **1995**, *13* (1), L2–L5.
 23. Bloomfield, G.L.; Holloway, S.; Ridings, P.C.; Fisher, B.J.; Blocher, C.R.; Sholley, M.; Bunch, M.; Sugerma, H.J.; Fowler, A.A. Pretreatment with inhaled nitric oxide inhibits neutrophil migration and oxidative activity resulting in attenuated sepsis-induced acute lung injury. *Crit. Care Med.* **1997**, *25* (4), 584–593.
 24. Reece, E.A.; Moya, F.; Yazigi, R.; Holford, T.; Duncan, C.; Ehrenkranz, R.A. Persistent pulmonary hypertension: Assessment of perinatal risk factors. *Obstet. Gynecol.* **1987**, *70* (5), 696–700.
 25. Roberts, J.D.; Polaner, D.M.; Lang, P.; Zapol, W.M. Inhaled nitric oxide in persistent pulmonary hypertension of the newborn. *Lancet* **1992**, *340*, 818–819.
 26. Kinsella, J.P.; Neish, S.R.; Shaffer, E.; Abman, S.H. Low-dose inhaled nitric oxide in persistent pulmonary hypertension of the newborn. *Lancet* **1992**, *340*, 819–820.
 27. Neonatal Inhaled Nitric Oxide Study Group. Inhaled nitric oxide in full-term and nearly full-term infants with hypoxic respiratory failure. *NEJM* **1997**, *336* (9), 597–604.
 28. Clark, R.H.; Kueser, T.J.; Walker, M.W.; Southgate, W.M.; Huckaby, J.L.; Perez, J.A.; Roy, B.J.; Kesler, M.; Kinsella, J.P. Low-dose nitric oxide therapy for persistent pulmonary hypertension of the newborn. *NEJM* **2000**, *342* (7), 469–474.
 29. Schamberger, M.S. Cardiac emergencies in children. *Pediatr. Ann.* **1996**, *25* (6), 339–344.
 30. Atz, A.M.; Wessel, D.L. Inhaled nitric oxide in the neonate with cardiac disease. *Semin. Perinatol.* **1997**, *21* (5), 441–455.
 31. Miller, O.I.; Tang, S.F.; Keech, A.; Pignott, N.B.; Beller, E.; Celermajer, D.S. Inhaled nitric oxide and prevention of pulmonary hypertension after congenital heart surgery: A randomized double-blind study. *Lancet* **2000**, *356*, 1464–1469.
 32. Morris, K.; Beghetti, M.; Petros, A.; Adatia, I.; Bohn, D. Comparison of hyperventilation and inhaled nitric oxide for pulmonary hypertension after repair of congenital heart disease. *Crit. Care Med.* **2000**, *28* (8), 2974–2978.
 33. Russell, I.A.; Zwass, M.S.; Fineman, J.R.; Balea, M.; Rouine-Rapp, K.; Brook, M.; Hanley, F.L.; Silverman, N.H.; Cahalan, M.K. The effects of inhaled nitric oxide on postoperative pulmonary hypertension in infants and children undergoing surgical repair of congenital heart disease. *Anesth. Analg.* **1998**, *87* (1), 46–51.
 34. Wernovsky, G.; Wypij, D.; Jonas, R.A.; Mayer, J.E., Jr.; Hanley, F.L.; Hickey, P.R.; Walsh, A.Z.; Chang, A.C.; Castaneda, A.R.; Newburger, J.W. Postoperative course and hemodynamic profile after the arterial switch operation in neonates and infants: A comparison of low-flow cardiopulmonary bypass versus circulatory arrest. *Circulation* **1995**, *92* (8), 2226–2235.
 35. Goldman, A.P.; Delius, R.E.; Deanfield, J.E.; de Leval, M.R.; Sigston, P.E.; Macrae, D.J. Nitric oxide might reduce the need for extracorporeal support in children with critical postoperative pulmonary hypertension. *Ann. Thorac. Surg.* **1996**, *62* (3), 750–755.
 36. Kinsella, J.P.; Ivy, D.D.; Abman, S.H. Inhaled nitric oxide improves gas exchange and lowers pulmonary vascular resistance in severe experimental hyaline membrane disease. *Pediatr. Res.* **1994**, *36* (3), 402–408.
 37. Skimming, J.W.; Bender, K.A.; Hutchinson, A.A.; Drummond, W.H. Nitric oxide inhalation in infants with respiratory distress syndrome. *J. Pediatr.* **1997**, *130* (2), 225–230.
 38. Kinsella, J.P.; Walsh, W.F.; Bose, C.; Gerstmann, D.R.; Labella, J.J.; Sardesai, S. Randomized controlled trial of inhaled nitric oxide in premature neonates with severe hypoxemic respiratory failure. *Lancet* **1999**, *354*, 1061–1065.
 39. Dellinger, R.P.; Zimmerman, J.L.; Taylor, R.W.; Straube, R.C.; Hauser, D.L.; Criner, G.J.; Davis, K.J.;



- Hyers, T.M.; Papadakos, P. Effects of inhaled nitric oxide in patients with acute respiratory distress syndrome: Results of a randomized phase II trial. *Crit. Care Med.* **1998**, *26* (1), 15–23.
40. Abman, S.H.; Griebel, J.L.; Parker, D.K.; Schmidt, J.M.; Swanton, D.; Kinsella, J.P. Acute effects of inhaled nitric oxide in children with severe hypoxemic respiratory failure. *J. Pediatr.* **1994**, *124* (6), 881–888.
 41. Benzing, A.; Geiger, K. Inhaled nitric oxide lowers pulmonary capillary pressure and changes longitudinal distribution of pulmonary vascular resistance in patients with acute lung injury. *Acta Anaesthesiol. Scand.* **1994**, *38* (7), 640–645.
 42. Benzing, A.; Bräutigam, P.; Geiger, K.; Loop, T.; Beyer, U.; Moser, E. Inhaled nitric oxide reduces pulmonary transvascular albumin flux in patients with acute lung injury. *Anesthesiology* **1995**, *83* (6), 1153–1161.
 43. Day, R.W.; Guarin, M.; Lynch, J.M.; Vernon, D.D.; Dean, J.M. Inhaled nitric oxide in children with severe lung disease: Results of acute and prolonged therapy with two concentrations. *Crit. Care Med.* **1996**, *24* (2), 215–221.
 44. Head, C.A.; Brugnara, C.; Martinez-Ruiz, R.; Kacmarek, R.M.; Bridges, K.R.; Kuter, D.; Bloch, K.D.; Zapol, W.M. Low concentrations of nitric oxide increase oxygen affinity of sickle erythrocytes in vitro and in vivo. *J. Clin. Invest.* **1997**, *100* (5), 1193–1198.
 45. Berger, J.I.; Gibson, R.L.; Redding, G.L.; Standaert, T.A.; Clarke, W.R.; Truog, W.E. Effects of inhaled nitric oxide during group B streptococcal sepsis in piglets. *Am. Rev. Respir. Dis.* **1993**, *147* (5), 1080–1086.
 46. Dahm, P.L.; Blomquist, S.; DeRobertis, E.; Jonson, B.; Myhre, E.; Svantesson, C.; Thorne, J. Effects of NO inhalation on pulmonary leukocyte sequestration and blood volume in porcine endotoxaemia. *Intensive Care Med.* **2000**, *26* (3), 336–343.
 47. Emil, S.; Berkeland, J.; Kosi, M.; Atkinson, J. Inhaled nitric oxide prevents experimental platelet activating factor-induced shock. *Arch. Surg.* **1996**, *131* (8), 855–859.
 48. Gessler, P.; Nebe, T.; Birle, A.; Mueller, W.; Kachel, W. A new side effect of inhaled nitric oxide in neonates and infants with pulmonary hypertension: Functional impairment of neutrophil respiratory burst. *Intensive Care Med.* **1996**, *22* (3), 252–258.
 49. Shimp, H.; Mitani, Y.; Tanaka, J.; Mizumoto, T.; Onoda, K.; Tani, K.; Yuasa, H.; Yada, I.; Maruyama, K. Inhaled low-dose nitric oxide for post-operative care in patients with congenital heart defects. *Artif. Organs* **1997**, *21*, 10–13.
 50. Belvisi, M.G.; Ward, J.K.; Mitchell, J.A.; Barnes, P.J. Nitric oxide as a neurotransmitter in human airways. *Ach. Int. Pharmacodyn. Ther.* **1995**, *329* (1), 97–110.
 51. Dupuy, P.M.; Shore, S.A.; Drazen, J.M.; Frostell, C.; Hill, W.A.; Zapol, W.M. Bronchodilator effect of inhaled nitric oxide in guinea pigs. *J. Clin. Invest.* **1992**, *90* (2), 421–428.
 52. Högman, M.; Frostell, C.G.; Hedenström, H.; Hedenstierna, G. Inhalation of nitric oxide modulates adult human bronchial tone. *Am. Rev. Respir. Dis.* **1993**, *148* (6), 1474–1478.
 53. Troncy, E.; Francoeur, M.; Blaise, G. Inhaled nitric oxide: Clinical applications, indications, and toxicology. *Can. J. Anaesth.* **1997**, *44* (9), 973–988.
 54. Wendel, H.P.; Ziemer, G. Coating-techniques to improve the hemocompatibility of artificial devices used for extracorporeal circulation. *Eur. J. Cardiothorac. Surg.* **1999**, *16* (3), 342–350.
 55. Hayashi, Y.; Sawa, Y.; Nishimura, M.; Chang, J.C.; Amemiya, A.; Kagisaki, K.; Taketani, S.; Yamaguchi, T.; Hirata, N.; Ohtake, S.; Matsuda, H. Nitric oxide gas infusion to the oxygenator enhances the biocompatibility of heparin coated extracorporeal bypass circuits. *ASAIO J.* **1998**, *44* (5), M456–461.
 56. Ramamurthi, A.; Lewis, R.S. Nitric oxide inhibition of platelet deposition on biomaterials. *Biomed. Sci. Instrum.* **1999**, *35*, 333–338.
 57. Hayashi, Y.; Sawa, Y.; Hirata, N.; Nishimura, M.; Ueda, H.; Naka, Y.; Yamaguchi, T.; Ohtake, S.; Matsuda, H. Improvement of bypass circuit biocompatibility: Comparison and combination of heparin-coated circuit and nitric oxide gas infusion. *J. Card. Surg.* **2002**, *17* (6), 477–484.
 58. Mowery, K.A.; Schoenfisch, M.H.; Saavedra, J.E.; Keefer, L.K.; Meyerhoff, M.E. Preparation and characterization of hydrophobic polymeric films that are thromboresistant via nitric oxide release. *Biomaterials* **2000**, *21* (1), 9–21.
 59. Kaul, S.; Cercek, B.; Rengstrom, J.; Xu, X.P.; Molloy, M.D.; Dimayuga, P.; Parikh, A.K.; Fishbein, M.C.; Nilsson, J.; Rajavashisth, T.B.; Shah, P.K. Polymeric-based perivascular delivery of a nitric oxide donor inhibits intimal thickening after balloon denudation arterial injury: Role of nuclear factor-kappaB. *J. Am. Coll. Cardiol.* **2000**, *35* (2), 493–501.
 60. Buegler, J.M.; Tio, F.O.; Schulz, D.G.; Khan, M.M.; Mazur, W.; French, B.A.; Raizner, A.E.; Ali, N.M. Use of nitric-oxide-eluting polymer-coated coronary stents for prevention of restenosis in pigs. *Coron. Artery Dis.* **2000**, *11* (4), 351–357.
 61. Annich, G.M.; Meinhardt, J.P.; Mowery, K.A.; Ashton, B.A.; Merz, S.I.; Hirschl, R.B.; Meyerhoff, M.E.; Bartlett, R.H. Reduced platelet activation and thrombosis in extracorporeal circuits coated with nitric oxide release polymers. *Crit. Care Med.* **2000**, *28* (4), 915–920.
 62. Bohl, K.S.; West, J.L. Nitric oxide-generating polymers reduce platelet adhesion and smooth muscle cell proliferation. *Biomaterials* **2000**, *21* (22), 2273–2278.
 63. Zhang, H.; Annich, G.M.; Miskulin, J.; Osterholzer, K.; Merz, S.I.; Bartlett, R.H.; Meyerhoff, M.E. Nitric oxide releasing silicone rubbers with improved blood compatibility: Preparation, characterization, and in vivo evaluation. *Biomaterials* **2002**, *23*, 1485–1494.
 64. Frost, M.C.; Rudich, S.M.; Zhang, H.; Maraschio, M.A.; Meyerhoff, M.E. In vivo biocompatibility and analytical performance of intravascular amperometric oxygen sensors prepared with improved nitric oxide-releasing silicone rubber coating. *Anal. Chem.* **2002**, *74* (23), 5942–5947.
 65. Lee, W.K.; Kobayashi, J.; Ooya, T.; Park, K.D.; Yui, N. Synthesis and characterization of nitric oxide

- generative polyrotaxane. *J. Biomater. Sci., Polym. Ed.* **2002**, *13* (10), 1153–1161.
66. Scott-Burden, T.; Tock, C.L.; Schwarz, J.J.; Casscells, S.W.; Engler, D.A. Genetically engineered smooth muscle cells as linings to improve the biocompatibility of cardiovascular prostheses. *Circulation* **1996**, *94* (9 Suppl), II235–II238.
67. Tock, C.L.; Bosley, J.P.; Parnis, S.M.; Clubb, F.J., Jr.; Macris, M.P.; Frazier, O.H.; Scott-Burden, T. A genetically engineered, nonthrombogenic cellular lining for LVADs: In vitro preconditioning before in vivo implantation. *ASAIO J.* **1999**, *45* (3), 172–177.
68. Dulak, J.; Jozkowicz, A.; Guevara, I.; Dembinska-Kiec, A. Gene transfer of vascular endothelial growth factor and endothelial nitric oxide synthase—Implications for gene therapy in cardiovascular diseases. *Pol. J. Pharmacol.* **1999**, *51* (3), 233–241.
69. Frey, A.; Schneider-Rasp, S.; Marienfeld, U.; Yu, J.C.; Paul, M.; Poller, W.; Schmidt, H.H. Biochemical and functional characterization of nitric oxide synthase III gene transfer using a replication-deficient adenoviral vector. *Biochem. Pharmacol.* **1999**, *58* (7), 1155–1166.
70. Tanner, F.C.; Meier, P.; Greutert, H.; Champion, C.; Nabel, E.G.; Luscher, T.F. Nitric oxide modulates expression of cell cycle regulatory proteins: A cytostatic strategy for inhibition of human vascular smooth muscle cell proliferation. *Circulation* **2000**, *101* (16), 1982–1989.
71. Muhs, A.; Heublein, B.; Schletter, J.; Herrmann, A.; Rudiger, M.; Sturm, M.; Grust, A.; Malms, J.; Schrader, J.; von der Leyen, H.E. Preclinical evaluation of inducible nitric oxide synthase lipoplex gene therapy for inhibition of stent—Induced vascular neointimal lesion formation. *Hum. Gene Ther.* **2003**, *14* (4), 375–383.
72. Ohata, T.; Sawa, Y.; Takagi, M.; Inoue, T.; Yoshida, T.; Kogaki, S.; Matsuda, H. hybrid artificial lung with interleukin-10 and endothelial constitutive nitric oxide synthase gene-transfected endothelial cells attenuates inflammatory reactions induced by cardiopulmonary bypass. *Circulation* **1998**, *98* (10 suppl), II269–II274.
73. Kalra, M.; Jost, C.J.; Severson, S.R.; Miller, V.M. Adventitial versus intimal liposome-mediated ex vivo transfection of canine saphenous vein grafts with endothelial nitric oxide synthase gene. *J. Vasc. Surg.* **2000**, *32* (6), 1190–1200.
74. Campbell, A.I.; Kuliszewski, M.A.; Stewart, D.J. Cell-based gene transfer to the pulmonary vasculature: Endothelial nitric oxide synthase overexpression inhibits monocrotaline-induced pulmonary hypertension. *Am. J. Respir. Cell Mol. Biol.* **1999**, *21* (5), 567–575.
75. Tirney, S.; Mattes, C.E.; Yoshimura, N.; Yokayama, T.; Ozawa, H.; Tzeng, E.; Birder, L.A.; Kanai, A.J.; Huard, J.; de Groat, W.C.; Chancellor, M.B. Nitric oxide synthase gene therapy for erectile dysfunction: Comparison of plasmid, adenovirus, and adenovirus-transduced myoblast vectors. *Mol. Urol.* **2001**, *5* (1), 37–43.
76. Chancellor, M.B.; Tirney, S.; Mattes, C.E.; Tzeng, E.; Birder, L.A.; Kanai, A.J.; de Groat, W.C.; Huard, J.; Yoshimura, N. Nitric oxide synthase gene transfer for erectile dysfunction in a rat model. *BJU Int.* **2003**, *91* (7), 691–696.
77. Bivalacqua, T.J.; Usta, M.F.; Champion, H.C.; Adams, D.; Namara, D.B.; Abdel-Mageed, A.B.; Kadowitz, P.J.; Hellstrom, W.J. Gene transfer of endothelial nitric oxide synthase partially restores nitric oxide synthesis and erectile function in streptozotocin diabetic rats. *J. Urol.* **2003**, *169* (5), 1911–1917.

Nuclear Magnetic Resonance Spectroscopy

Mari A. Smith
Peter B. Barker

Department of Radiology, Johns Hopkins University School of Medicine, Baltimore, Maryland, U.S.A.

INTRODUCTION

Nuclear magnetic resonance (NMR) spectroscopy is based upon the measurement of the absorption of electromagnetic radiation in the radiofrequency region by atomic nuclei. For absorption to occur, the degeneracy of the nuclei's energy states is removed by placement in an intense magnetic field.

HISTORICAL PERSPECTIVES

The concept of NMR spectroscopy was first postulated in 1924 by W. Pauli's suggestion that exposure to a magnetic field would cause the energy levels of atomic nuclei possessing spin and angular momentum properties to separate.^[1] However, experimental verification of the absorption of electromagnetic radiation in bulk matter by nuclei did not occur until 1946, when both F. Bloch and E. Purcell independently demonstrated that a strong magnetic field induced splitting of the energy levels, and thus they detected the resonance phenomena.^[2,3] The method was originally of interest only to physicists for the measurement of gyromagnetic ratios (γ) of different nuclei, a constant specific to a particular nucleus, but applications of NMR to chemistry became apparent after the discovery of chemical shift and spin-spin coupling effects in 1950 and 1951, respectively.^[4,5] The spectra of high-resolution liquid NMR contain fine structure information because the absorption frequency (nuclear resonance) is influenced by neighboring nuclei and the chemical environment; hence NMR spectroscopy has rapidly become an important and widely used technique for chemical analysis and structure elucidation of chemical and biological compounds.

Major technical advances in the 1960s included the introduction of superconducting magnets in 1965, which were very stable and allowed higher field strengths to be attained than was possible with conventional electromagnets; and the use of the Fourier transform (FT) in 1966 for signal processing. In nearly all contemporary spectrometers, the sample is subjected to periodic radiofrequency pulses directed perpendicularly to the external field; the signal is

Fourier transformed to give a spectrum in the frequency domain. FT NMR provides increased sensitivity compared to previous techniques, and has led to the development of a huge variety of pulsed NMR methods, including multidimensional NMR techniques.

Biological and medical applications of magnetic resonance were developed in the early 1970s with the introduction of magnetic resonance imaging (MRI) and magnetic resonance spectroscopy (MRS) of biological tissue. In vivo MRS of humans became possible in the early 1980s with the advent of whole body magnets with sufficiently high field strength and homogeneity.^[6] Magnetic resonance imaging studies in humans are based on the proton resonances of tissue water, in which the water content and the water relaxation times T_1 and T_2 are used to generate tissue contrast. T_1 and T_2 are the longitudinal and transverse lifetimes and describe the decay of the parallel and perpendicular components of magnetization, respectively, relative to the static magnetic Zeeman field. In addition to anatomical imaging, there has been a recent surge of interest in functional imaging in which dynamic properties such as water diffusion, brain-blood flow and blood volume, oxygenation, and macromolecule-water proton interactions can be used as image contrast.^[7,8] Although clinical applications of MRI were immediately apparent because of its anatomical imaging capabilities, MRS has until recently been predominantly a tool for biomedical research. For example, in vivo MRS and MR spectroscopic imaging (MRSI) of the brain is used to measure equilibrium levels of brain metabolites and to study active brain transport and metabolism. Early MRS studies predominantly used the phosphorus-31 (^{31}P) nucleus, but more recently the proton (^1H) nucleus, which is the most sensitive and fully abundant naturally occurring nucleus, has been used more extensively, particularly for studies of the brain.

BASIC THEORY

The NMR experiment involves the observation of transitions between nuclear spin states that are

perturbed by a Zeeman interaction with an external magnetic field. If the magnetic field is described by B_0 , then the energy of the nuclear spin state is given by

$$E = -\mu \cdot B_0 \quad (1)$$

where μ is the nuclear magnetic moment. The magnetic moment is related to the spin angular momentum P by the gyromagnetic ratio γ ,

$$\mu = \gamma P = \gamma(I\hbar/2\pi) \quad (2)$$

where γ is a characteristic constant for each nucleus, I is the spin quantum number, and \hbar is Planck's constant. By definition, the direction of B_0 is taken to specify the Z axis, so Eq. 1 reduces to

$$E = -\mu_z B_0 \quad (3)$$

where μ_z is the component of μ along the Z axis. For a nucleus with spin quantum number I , there are $(2I + 1)$ different possible orientations of μ in the field, each with a component m_I in the Z direction. For example, for a spin $1/2$ nucleus ($I = 1/2$), m_I can take the values $+1/2$ and $-1/2$ (Fig. 1).

Applying the magnetic dipole-allowed selection rule $\Delta m_I = \pm 1$, the resonance frequency is given by

$$\nu_0 = \Delta E/\hbar = \gamma h B_0 / (2\pi\hbar) = \gamma B_0 / 2\pi \quad (4)$$

which can be expressed in angular frequency units as

$$\omega_0 = \gamma B_0 \quad (5)$$

ω_0 is called the Larmor frequency. With typical magnetic field strengths currently available (≈ 0.2 to 20 Tesla), most magnetic nuclei resonate in the high-frequency to very-high-frequency regions of the electromagnetic spectrum (10 to 900 MHz for the proton). Systems for human spectroscopy studies typically

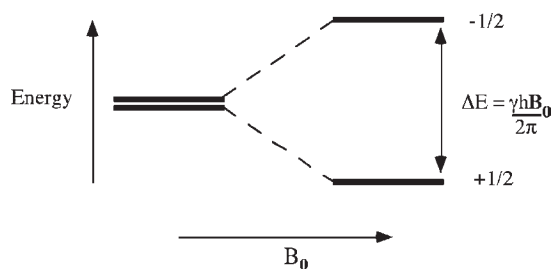


Fig. 1 Energy levels for a nucleus with spin quantum number $I = \pm 1/2$. In the absence of an externally applied field, B_0 , the two energy levels are degenerate. Nuclei with spin $1/2$ have two energy levels corresponding to the two discrete values of I . The spin with $I = +1/2$ is aligned with the external magnetic field and thus is lower in energy.

operate between 0.5 to 4.0 Tesla (1.5 T being the most common); smaller horizontal-bore systems for animal studies are available up to about 9 Tesla (384 MHz). Vertical small-bore superconducting magnet systems for high-resolution NMR spectroscopy are commercially available to over 20 Tesla (900 MHz).

Signal and Signal-to-Noise Ratio

In NMR, the difference between energy levels is very small; hence it is an intrinsically insensitive technique. The excess population of the lower level compared to the upper level can be calculated from the Boltzmann factor

$$\frac{n(\text{upper})}{n(\text{lower})} = \exp(-\Delta E/kT) \quad (6)$$

where k is the Boltzmann constant and T is the absolute temperature (for protons at room temperature at 100 MHz, $n(\text{upper})/n(\text{lower}) \approx 0.999984$). The excess population of the lower level creates the macroscopic nuclear magnetization that is observed in the NMR experiment. The net nuclear magnetization M_0 is given by

$$M_0 = NB_0\gamma^2\hbar^2I(I+1)/3kT \quad (7)$$

where N is the total number of spins in the sample and \hbar is Planck's constant divided by 2π . The signal (S) detected in the receiver coil is proportional to the magnetization times the resonant frequency

$$S \propto M_0\omega_0 \quad (8)$$

giving a final dependence of the signal on B_0^2 for a given nucleus (γ fixed) or alternatively a γ^3 dependence for different nuclei at a fixed B_0 . In a typical experiment, the NMR signal induced in the receiver coil is of the order of a few microvolts. For nonconducting samples, the noise in the system is predominantly from the random (thermal) motion of electrons in the receiver coil (Johnson noise). The noise power is independent of frequency; however, the noise voltage increases with approximately the square root of frequency, because the resistance of the coil increases with frequency (the skin depth effect). Therefore, for a given nucleus, the signal-to-noise ratio increases as $B_0^{3/2}$, whereas for a given field strength the S/N ratio increases as $\gamma^{5/2}$. For instance, the ratio of γ for proton and carbon-13 is about 4:1; therefore, the S/N ratio is 32 times greater on ^1H than on ^{13}C for equal numbers of nuclei. For biological samples, the noise is often dominated by the sample's resistance, which increases linearly with frequency. In this case, the S/N is only linearly proportional to B_0 for fixed γ , or to γ^2 for fixed B_0 .



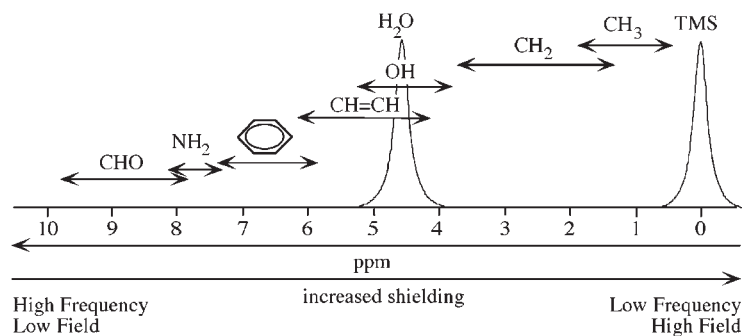


Fig. 2 Typical ^1H chemical shift range (10 ppm) for various functional groups. The resonances of water and the standard reference compound TMS are shown at 4.7 and 0.0 ppm, respectively.

THE NMR SPECTRUM

The Chemical Shift

Early in the development of NMR, it was discovered that nuclei in different molecular environments resonate at slightly different frequencies.^[4] The origin of this effect lies in the response of the molecule's electrons to the applied magnetic field; in simple terms, a rotating current is induced that generates a small magnetic field that opposes the external field (a diamagnetic effect). Hence, the nucleus experiences a smaller net field than that which is actually applied. The effective field at the nucleus, B , can be expressed in terms of a shielding parameter, σ ,

$$B = B_0(1 - \sigma) \quad (9)$$

so that the nuclear resonance frequency becomes

$$\nu = \gamma B / 2\pi = \nu_0(1 - \sigma) \quad (10)$$

This effect is known as the chemical shift; it is of prime importance because nuclei in different chemical environments can be distinguished on the basis of their resonant frequencies. The shielding parameter is usually defined in parts per million (ppm) of the resonance frequency measured relative to a reference compound. Although the shielding parameter is a constant, the chemical shift (measured in Hz) increases linearly with field strength (Eq. 7). Thus the resolution of the NMR spectrum increases linearly with increasing field strength. To compare chemical shifts measured at different field strengths, it is standard to report chemical shift values (δ) in ppm relative to a standard reference compound; for ^1H and ^{13}C spectra (in vitro), tetramethylsilane (TMS) is often used as the reference,

$$\delta = (\nu - \nu_{\text{TMS}}) * 10^6 / \nu_{\text{TMS}} \quad (11)$$

where ν and ν_{TMS} are the frequencies of the signal of interest and the reference signal, respectively. Figure 2 illustrates the typical ^1H chemical shift range for various functional groups. Whereas the typical

chemical shift range for ^1H is quite small (≈ 10 ppm), much larger chemical shift ranges (e.g., up to several 100 ppm) exist for ^{19}F , ^{31}P , ^{13}C , and ^{15}N .

Spin-Spin Couplings

In addition to chemical shift effects, it was discovered that spectra from liquid samples exhibit further fine structure in the forms of splittings (or multiplets).^[5] These arise from the electron-coupled spin-spin interaction, also known as J- or scalar-coupling. Spin-spin coupling results from nuclei experiencing the magnetic fields of their neighboring nuclei through polarization of the electrons in the molecular bonds between them. The effective magnetic field experienced by one nucleus depends on the spin state of a neighboring coupled nucleus. The nomenclature to describe spin systems assigns letters to each individual spin, which are close together in the alphabet if the spins are strongly coupled or far from one another in the alphabet if the spins are weakly coupled ($|\delta_A - \delta_X| \gg J_{AX}$). The simplest multiplet pattern that can be observed is in a two-spin AX system (Fig. 3).

Spin-spin couplings have the following properties:

1. They act through the bonding electrons and are therefore intramolecular.

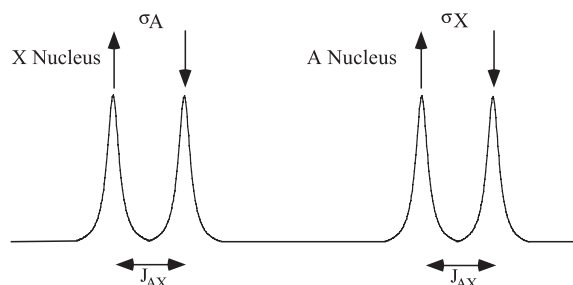


Fig. 3 Multiplet pattern of a weakly coupled two-spin (AX) system. The splitting is equal to J_{AX} . σ represents the shielding constant of the A or X spin. The arrows indicate alignment with (↓) or against (↑) the static magnetic field.

2. They are independent of the strength of the applied magnetic field.
3. Spin multiplet structure reflects the states of neighbor nuclei.
4. The interaction is reciprocal; if A splits X, then X splits A.
5. Splitting cannot be observed due to coupling between equivalent nuclei.
6. Coupling can be homonuclear or heteronuclear.

Typical coupling patterns

For spins that have widely different chemical shifts, the multiplet pattern is symmetrical around the chemical shift frequency, with relative intensities given by the binomial coefficients $(1+x)^N$ for $I=1/2$ nuclei.

Strong coupling

The simple treatment of multiplet patterns is valid only if $|\delta_A - \delta_X| \gg J_{AX}$, which is known as the weak coupling, or first-order approximation. When this is not true, the spectra are strongly-coupled or second-order, the symmetry of the multiplet is destroyed, and the splittings are no longer equal to the coupling constant. New lines may appear (combination lines), and lines are no longer assignable to a single nucleus. In general, the spectrum can be assigned only through the use of computer simulation, although some simple spin-coupling features may remain that can be assigned by factoring the spin system into subsystems and introducing an additional spin to calculate how it affects the subsystem (the X approximation).^[9,10] This is justified as long as the new spin is only weakly coupled to the spins under consideration.

Coupling with $I > 1/2$ nuclei

Nuclei with $I > 1/2$ produce more complicated multiplet patterns. The deuterium nucleus ($I=1$) produces a triplet in neighboring atoms with an intensity ratio 1:1:1, corresponding to $m_I = -1, 0,$ and $+1$, respectively. However, some quadrupolar nuclei relax too rapidly to generate observable splittings. It should also be noted that spin-spin splitting can be modified or eliminated by chemical exchange or double resonance experiments (decoupling).

PULSED FOURIER TRANSFORM (FT) SPECTROSCOPY

The production of transverse magnetization through perturbation by radiofrequency (RF) pulses is required

to record a spectrum. In an FT NMR experiment,^[11] an RF pulse is applied and the receiver is then turned on to record the time varying signal (free induction decay, or FID) from the resulting magnetization as it precesses in a static magnetic field B_0 . The signal from the entire spectrum is collected in a single scan (provided the RF pulse has sufficient bandwidth, i.e., it is a hard pulse). Therefore, FT NMR has an enormous sensitivity advantage over continuous wave (CW) NMR where each resonance is observed in turn. The signal-to-noise ratio can be improved through time-averaging, i.e., signals are digitized and summed in a computer memory; the signal S (which is coherent) accumulates linearly ($S \propto n$, where n is the number of scans), whereas the noise \sqrt{N} (which is incoherent) accumulates in memory as n ; therefore the signal-to-noise ratio S/N improves as a function of \sqrt{n} .



The Fourier Transform

The FID consists of the summation of all of the resonant frequencies in the spectrum:

$$f(t) = \sum_{j=1}^k \exp(i\Omega_j t) \exp(-\lambda_j t) \quad (12)$$

Ω_j and λ_j are the resonant frequencies and decay rates, respectively, of the j th out of k resonances in the spectrum. In a perfectly uniform (homogeneous) magnetic field, λ is the inverse of the T_2 relaxation time ($\lambda = 1/T_2$). In an inhomogeneous magnetic field, the rate of decay of the magnetization is faster because of destructive interference of the signals precessing at different frequencies, and the time constant is called T_2^* ("tee-two-star," $\lambda = 1/T_2^*$). The spectrum is calculated from the FID by Fourier transformation:

$$f(\omega) = \text{F.T.}[f(t)] = \int_{-\infty}^{+\infty} \exp(-i\omega t) f(t) dt \quad (13)$$

Substituting Eq. 12 into Eq. 13 and evaluating the integral gives (for the case of a single resonance, $k=1$)

$$\begin{aligned} f(\omega) &= \frac{\lambda}{\lambda^2 + (\Omega - \omega)^2} + \frac{i(\Omega - \omega)}{\lambda^2 + (\Omega - \omega)^2} \\ &= A(\omega) + iD(\omega) \end{aligned} \quad (14)$$

$A(\omega)$ and $D(\omega)$ are Lorentzian functions. The real part of this expression is called the "absorption" lineshape (A); the imaginary part is called the "dispersion" lineshape (D) (see Fig. 4). The absorption lineshape is the preferred lineshape for spectral display because it has no negative intensity and tails off more rapidly than the dispersion mode. In practice, the continuous

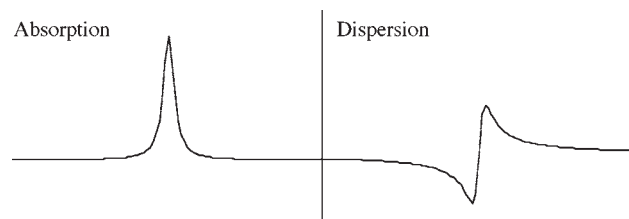


Fig. 4 Real (absorptive) and imaginary (dispersive) components of the complex frequency domain NMR spectrum obtained by Fourier transformation of the FID.

time domain signal $f(t)$ is sampled discretely on an equidistant set of time points, Δt , within the Nyquist frequency, the maximum frequency for a given t ,

$$f_n = \frac{1}{2\Delta t} \quad (15)$$

Therefore, the frequency spectrum is conventionally generated from a finite discrete Fourier transformation (DFT) of the time domain signal. The DFT closely approximates the continuous FT.

SIMPLE PULSED NMR EXPERIMENTS: THE ROTATING FRAME

The theoretical description of NMR experiments can be done at various different levels of complexity. Although spin angular momentum is inherently a quantum-mechanical effect, the classical (vector) description of the interaction between the net macroscopic nuclear magnetization and the applied magnetic fields is the simplest to understand and most widely used. For NMR experiments on weakly coupled spin systems, the product operator formalism is often used,^[12] but for strongly coupled spin systems a full quantum-mechanical density matrix approach is required.^[13]

For most experiments in MRI and localized spectroscopy, the classical vector approach is sufficient. Visualization of the motion of the magnetization vector during a pulse sequence is most easily appreciated in the rotating frame of reference, a coordinate system that rotates about the Z axis at the Larmor frequency. In this frame, the precessional frequency of the spins around B_0 and the oscillations of the B_1 field are both removed, so the only effective field during an RF pulse is the (now static) B_1 field, taken to define the X axis (Fig. 5A). If the initial magnetization is aligned along the Z axis, application of the B_1 field causes a rotation of the magnetization around the X axis with an angular frequency $\omega_1 = \gamma B_1$. Therefore, if the RF pulse is applied for a period τ , the flip angle α (in radians) is given by

$$\alpha = \gamma B_1 \tau \quad (16)$$

Because the NMR spectrometer detects signal in the transverse plane, maximum signal is excited (on a single scan, ignoring relaxation) when $\alpha = \pi/2$ (i.e., a 90° pulse) (Fig. 5A). In general, it is not possible for all spins to be exactly on-resonance at the same time in the rotating frame because of the inherent frequency distribution in the spectrum. Therefore, during an RF pulse, spins that are off-resonance precess around an effective field in the rotating frame that has both a Z component ($= (\omega - \omega_0)/\gamma$, where ω is the frequency of the spin and ω_0 is the frequency of the rotating frame) and an X component (γB_1) (Fig. 5B). As the spins become farther off-resonance, the effective field becomes progressively larger and tilted toward the Z axis; once the field is tilted more than 45° , it is impossible to rotate the magnetization fully into the transverse plane using simple radiofrequency pulses. Selective excitation and slice-selective processes are based on the use of this off-resonance effect.

Measurement of T_1 Relaxation Times

Progressive saturation—optimization of sensitivity

The simplest pulsed NMR experiment uses a single RF pulse followed by detection of the FID (Fig. 6A). When a pulse-and-collect NMR experiment is repeated several (n) times to improve S/N, the longitudinal magnetization reaches a steady state after a few repetitions. The steady-state magnetization depends on the flip angle used (α) and the spin-lattice (T_1) relaxation time, whose inverse describes the rate at which longitudinal magnetization recovers after it has been displaced from equilibrium by an RF pulse. For a 90° pulse ($\alpha = \pi/2$), the steady-state signal is given by

$$S(TR) = S_0(1 - \exp(-TR/T_1)) \quad (17)$$

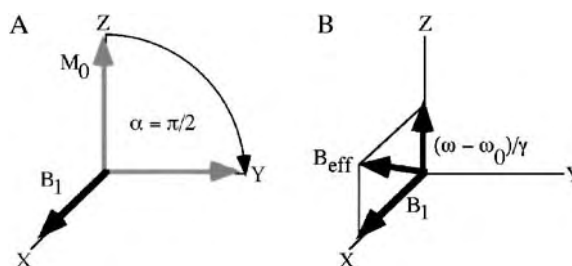


Fig. 5 (A) Effect of an applied RF field, B_1 , on initial magnetization, M_0 , in the rotating reference frame. The application of a 90° pulse with phase x creates transverse magnetization, shown here to be aligned along the Y axis; (B) orientation of the effective field, B_{eff} , experienced by spins that are off-resonance.

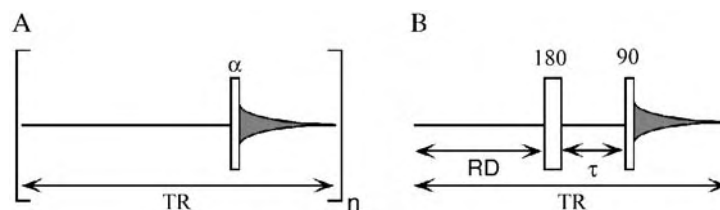


Fig. 6 (A) The simplest pulsed experiment consists of a single RF pulse of flip angle, α , followed by detection of the FID, and can be repeated n times to improve the signal-to-noise ratio. For a 90° pulse ($\alpha=90$) with a repetition time (TR) of zero, the detected signal will also equal zero, indicating that the spins are saturated; (B) diagram of the inversion recovery experiment that can be used to determine the T_1 relaxation time by varying the delay, τ , provided the recovery delay, RD, is $\sim 5 > T_1$.

where TR is the repetition time of the experiment. In the limit $TR=0$, the signal will also be zero (the spins are saturated). In a progressive saturation experiment, T_1 can be calculated by recording the signal as a function of several different TR values. At any given TR, there will be an optimum flip angle for maximum signal known as the Ernst angle:

$$\cos \alpha = \exp(-TR/T_1) \quad (18)$$

Inversion recovery

One of the earliest pulsed FT NMR experiments was the inversion recovery experiment, which can also be used to measure T_1 relaxation times. Inversion recovery consists of a 180° pulse to invert the Z magnetization, a delay (τ) during which T_1 relaxation occurs, and a 90° read pulse to detect the NMR signal (Fig. 6B). The experiment is repeated with several different values of τ ; the T_1 relaxation time can be obtained from fitting the detected signal to the expression

$$S(\tau) = S_0(1 - 2 \exp(-\tau/T_1)) \quad (19)$$

The repetition delay RD between experiments must be sufficiently long to allow complete recovery of Z magnetization before the next pulse is applied (typically, $RD > 5T_1$).

Spin Echoes: Measurement of T_2 Relaxation Times

The rate of decay of transverse magnetization is described by the T_2 , or spin-spin, relaxation time constant. In addition to true T_2 relaxation, transverse magnetization is also dephased as a result of B_0 magnetic field inhomogeneities. The effects of field inhomogeneity can be reversed by applying a 180° refocusing pulse that generates a spin echo, a reappearance of the signal a finite time after disappearance of the initial FID^[14] (Fig. 7).

The amplitude of the spin-echo signal is given by

$$S(TE) = S_0 \exp(-TE/T_2) \quad (20)$$

T_2 can be calculated by recording the signal as a function of the echo time TE. If the 180° pulse is imperfect, or if spins diffuse into different regions of magnetic field strength before and after the 180° , echo rephasing will be imperfect, leading to a reduction in signal intensity and an underestimate of the T_2 relaxation time. These problems can be partially overcome by using multiple 180° pulses (the Carr-Purcell Method B^[15]) with appropriate phase-shifted 180° pulses (the Carr-Purcell-Meiboom-Gill (CPMG) method^[16]) (Fig. 8A).

Pulses of flip angle other than 180° can produce echoes (Hahn's original spin-echo paper was based on the 90° - 90° sequence^[14]), and if more than two pulses are

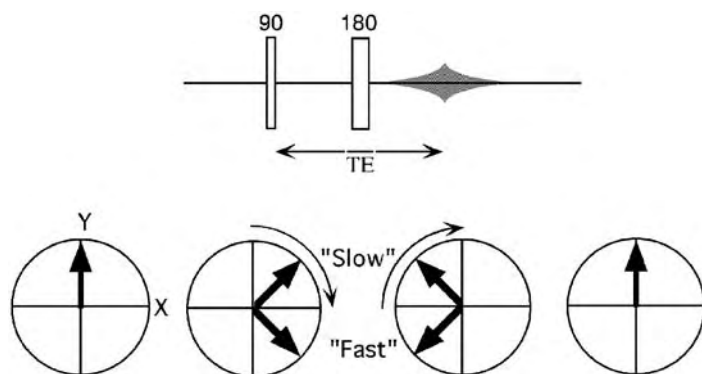


Fig. 7 Illustration of a spin echo. After application of a 90° pulse with x phase, the magnetization is aligned along the Y axis. Immediately prior to application of the 180° pulse the magnetization has experienced a delay ($TE/2$) that causes the magnetization to dephase. Application of the 180° pulse causes the magnetization to be flipped 180° , and after an equivalent delay ($TE/2$), the magnetization will refocus along the Y axis as a spin echo, represented by the shaded diamond icon.

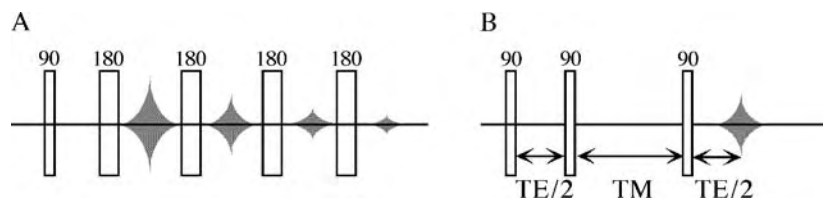


Fig. 8 (A) Carr-Purcell-Meiboom-Gill (CPMG) method. Multiple 180° pulses are employed to counteract imperfect 180° pulses—those that cannot flip all spins within a desired bandwidth, and/or spin diffusion to produce an echo of the signal; (B) similarly, a (third) stimulated signal echo can be observed at a time $TE/2$ after the application of the third 90° pulse in a train of three 90° pulses.

applied, more than one echo can be observed. In particular, for a train of three 90° pulses, a third stimulated echo can be observed at a time $TE/2$ after the third pulse (Fig. 8B). Spin and stimulated echoes are used extensively both in MRI imaging and in localized NMR spectroscopy.

MAGNETIC NUCLEI OFTEN USED FOR BIOLOGICAL STUDIES

Many elements have isotopes with nonzero nuclear spin. However, to generate a strong high-resolution NMR signal, the following factors are required: 1) spin $1/2$ (or small quadrupole moment if $I > 1/2$); 2) large chemical shift range; 3) high γ ; and 4) high natural abundance (unless isotopic enrichment is used). ^1H and ^{31}P best fulfill these criteria for studies of biological tissue.

An example ^{31}P spectrum from the human brain is shown in Fig. 9. In vivo, the phosphocreatine resonance is often used as a chemical shift reference (set to zero ppm). The ATP resonances exhibit homonuclear ^{31}P spin-spin couplings ($J \approx 20$ Hz); the

β -ATP is a triplet because it is coupled to both the α and the γ spins, whereas the α and γ resonances are doublets because they are coupled only to the β nucleus.

Generally, magnetic resonance studies are associated with anatomical imaging generated from the tissue contrast due to water content and relaxation times T_1 and T_2 (see Fig. 10). Imaging, however, is a relatively recent application of nuclear magnetic resonance spectroscopy. NMR spectroscopy is routinely used in the analysis of liquid and solid state materials. For example, solid state nuclear magnetic resonance (SSNMR) and imaging have been used to study both natural and synthetic biomaterials such as bone mineral, calcification (atherosclerosis), and polymer, ceramic, and composite implants.^[17,18] SSNMR is useful for structural studies of molecules that cannot be easily crystallized or dissolved, including surface-adsorbed biopolymers and membrane-associated proteins, and for large and complex molecules such as proteins and ribonucleic acids (RNA).^[19,20] SSNMR defines molecular structure through the measurement of internuclear distances and torsion angles. Structural measurements

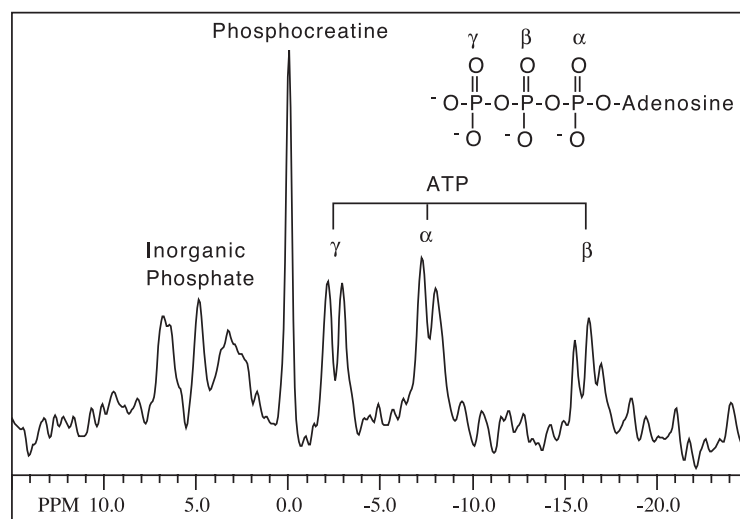


Fig. 9 ^{31}P spectrum from the human brain. (ATP = adenosine triphosphate.)

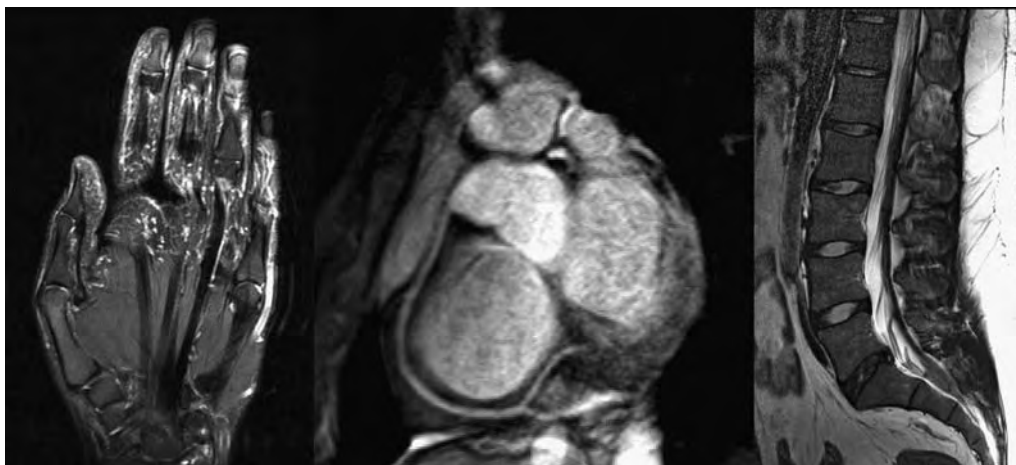


Fig. 10 Examples of magnetic resonance imaging of representative human anatomies. Shown is a T₂-weighted image of a hand (left), a coronary image displaying the right coronary artery (middle), and a sagittal T₂-weighted image of the lumbar spine (right). (Images courtesy of Terri Brawner, F.M. Kirby Research Center, Kennedy Krieger Institute.)

are accomplished through the application of magic angle sample spinning to remove anisotropic magnetic interactions, in conjunction with a pulse sequence that selectively recouples the dipolar interaction, from which direct structural information can be obtained.

FUNDAMENTALS OF HUMAN (BRAIN) SPECTROSCOPY

In vivo MR spectroscopy is possible in many human organ systems and with many nuclei: ³¹P spectroscopy of muscle metabolism, ¹³C spectroscopy of the brain and liver, and ¹H spectroscopy of the brain, breast,^[67] and prostate^[21] are a few examples. Early studies focused on the phosphorus nucleus, because this was the most technically feasible at that time. Methods were developed for spatially localized ³¹P MRS,^[22] and studies of major neuropathologies (such as stroke or brain tumors) were performed.^[23–25] A significant problem with ³¹P MRS, however, is its low sensitivity, mainly because of the relatively low γ of ³¹P and low concentrations of phosphorus-containing compounds. Because the spatial resolution in in vivo spectroscopy is largely limited by the signal-to-noise ratio, the typical voxel size for ³¹P spectroscopy of the human brain is usually at least 30 cm³ using conventional techniques and 1.5 Tesla magnets. ³¹P spectroscopy can therefore usually be applied only to either very large lesions or diffuse and/or global diseases.

In recent years, there has been more interest in ¹H spectroscopy, particularly after it was demonstrated that it was possible to obtain high-resolution spectra from small well-defined regions in reasonably

short scan times.^[26] The higher sensitivity of the ¹H is due to several factors, including higher γ , higher metabolite concentrations, and more favorable relaxation times compared to ³¹P or other nuclei. Also, because the same nucleus is measured as in MRI, no additional hardware is required. Although ¹H spectroscopy has been demonstrated in a number of organ systems, the overwhelming number of applications have been in the brain because of the absence of free lipid signals in normal cerebrum, ease of shimming, and lack of motion artifacts. Human brain pathology has now been extensively studied by ¹H spectroscopy.^[27]

Spatial Localization Techniques for In Vivo Spectroscopy

Generally, two different approaches are used for ¹H spectroscopy of the brain: 1) single voxel methods based on the STEAM (STimulated Echo Acquisition Mode: 90°-TE/2-90°-TM-90°-TE/2-Aquire, where TM is the mixing time)^[26] or PRESS (Point RESolved Spectroscopy: 90°-TE₁-180°-TE₁₊₂-180°-TE₂-Aquire)^[28] pulse sequences for localization; or 2) spectroscopic imaging (SI, also known as chemical shift imaging (CSI)) studies usually done in two dimensions using a variety of different pulse sequences (spin-echo, PRESS).^[29–31] Usually, but not exclusively, single voxel scans are recorded at short echo times (TE < 30 ms), whereas SI studies are done at long echo times (e.g., TE > 135 ms). Short echo-time spectra contain signals from more compounds and have better signal-to-noise ratios, but also have worse water and lipid suppression factors. Long echo-time spectra have lower signal-to-noise ratios, fewer detectable

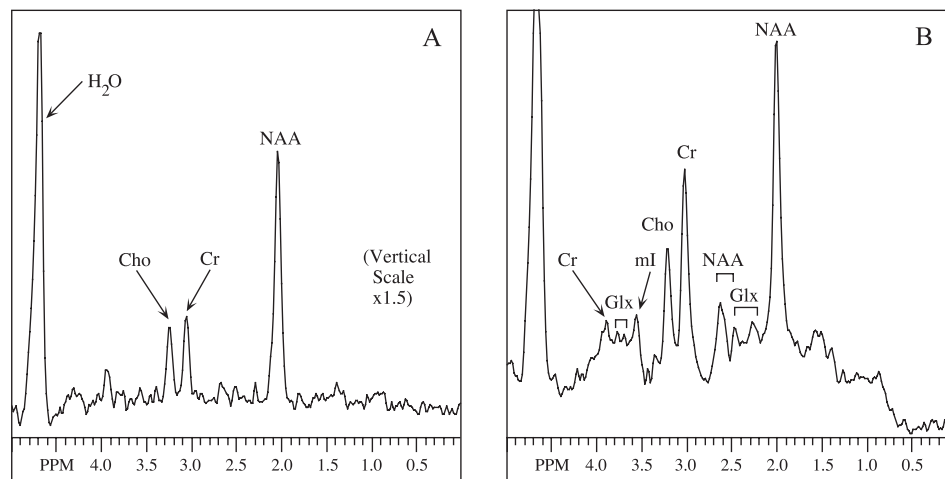


Fig. 11 (A) Long (TE 272 ms) and (B) short (TE 30 ms) ^1H MR spectra from normal human brain recorded using the STEAM pulse sequence. H_2O = residual water, Cho = choline, Cr = creatine, NAA = N-acetyl aspartate, mI = *myo*-inositol, Glx = glutamine and glutamate.

compounds, and a variable amount of T_2 -weighting, but are usually better resolved spectra with flatter baselines. Lactate is usually best detected at long echo times (TE = 140 or 280 ms so that the J-modulation is rephased) to distinguish it from lipid signals. Examples of short and long TE SV spectra of the human brain are shown in Fig. 11. Whereas single-voxel methods allow the optimized detection of individual brain regions at short echo times, they also have a number of disadvantages; voxel sizes are rather large, and in many conditions choice of voxel location is difficult. For diagnostic purposes, one would in principle like to study the complete brain. The combination of MRS with imaging (MRSI) can alleviate these problems. An example of a spectroscopic imaging study is given in Fig. 12, along with a T_1 -weighted image for reference. The metabolite maps display elevated levels of choline, the presence of lactate, and the absence of N-acetyl aspartate (NAA) in the tumor region, as does the spectrum selected from a voxel within the tumor region. Figure 13 demonstrates a conventional proton MRSI in normal human breast. The figure shows a T_1 localizer image (middle of 3 slices); reconstructed

images of residual water, residual lipid (both after suppression), and choline; and selected spectra from individual regions of interest in adipose and glandular tissue. The data show that no choline signal was detectable above the noise floor of the spectroscopic image, and that residual water and lipid signals after suppression were reasonably small. Water signals are predominantly observed only in glandular tissue and chest wall, whereas lipid signals are predominantly seen in the adipose tissue, indicating sufficient resolution for proton MRSI measurements of choline in breast cancer patients.

The choice of method depends (in addition to availability) on the information required by the particular medical or research application. For instance, if spectroscopy were being used to search for the location of a stroke or a seizure focus, spectroscopic imaging would be preferable because this generates maps of metabolite levels that can be screened for abnormalities in different locations. Alternatively, if the issue is to observe changes in compounds such as glutamine/glutamate or myoinositol, which can only be detected in short echo-time spectra, in global or diffuse

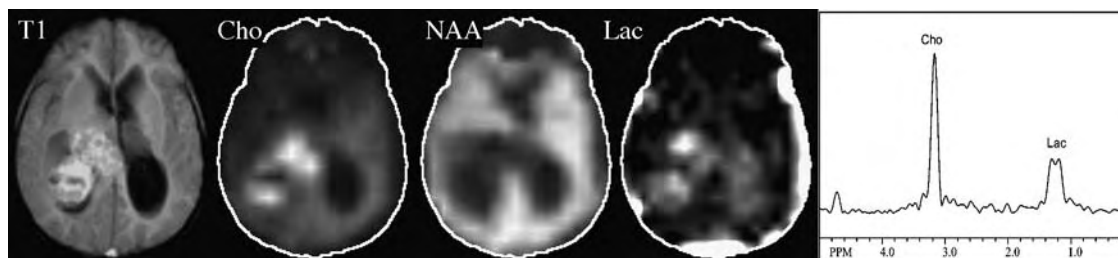


Fig. 12 Example of an MRSI in a 4-yr-old boy with a choroid plexus carcinoma. Shown are metabolic maps of Cho, NAA, and Lac along with a T_1 -weighted MRI (left) and spectra from the tumor region (right) for comparison. Both the tumor spectra and spectroscopic images show elevation of choline and lactate with complete absence of N-acetyl aspartate and creatine (see spectra).

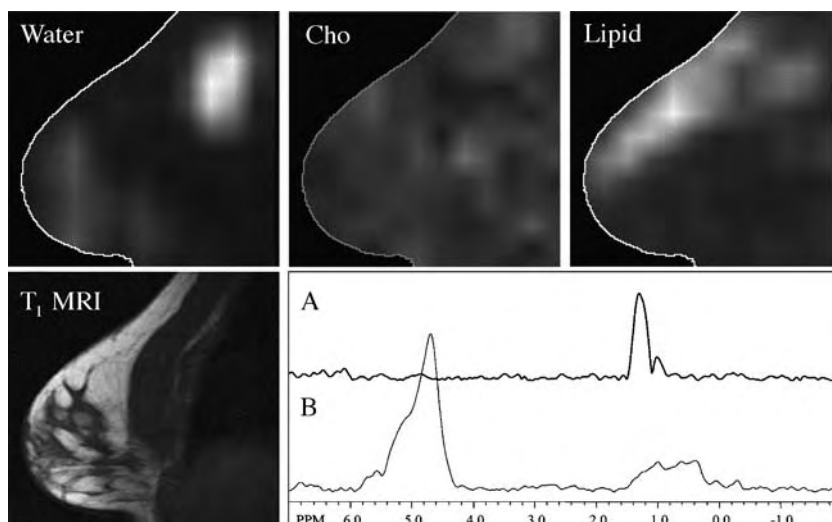


Fig. 13 Spectroscopic images of residual water, choline (integration range 3.3 to 3.1 ppm), and residual lipid from one slice in a normal human breast. Selected spectra are shown from individual 1.27-cm^3 voxels in (A) adipose tissue showing methyl and methylene lipid signals; and (B) chest wall showing predominantly residual water signal. Note that the broad water signal is the residual water signal in the chest wall where field homogeneity is worse than in the breast.

diseases such as hepatic encephalopathy, then short echo-time single-voxel spectroscopy is the method of choice.

Water and Lipid Suppression

Brain metabolite levels are on the order of 10 mM or less, whereas brain water is 80 M and lipids in pericranial fat are also present in very high concentrations. Therefore, water and lipid suppression techniques are essential in ^1H spectroscopy in order to reliably observe the much smaller metabolite signals. Numerous methods for solvent (water) suppression have been developed in high-resolution NMR spectroscopy, and some of these methods have been applied to in vivo spectroscopy. The most common approach is to presaturate the water signal using frequency-selective 90° pulses (CHESS pulses^[32]) prior to the localization pulse sequence. By using more than one pulse, and with correct choice of flip angles,^[33,34] very good suppression factors can be attained (>1000).

Lipid suppression can be performed in several different ways. One approach is to avoid exciting the lipid signal, using STEAM^[26] or PRESS^[28] localization, confined to brain regions, and avoiding the scalp. Alternatively, outer-volume suppression pulses can be used to presaturate the lipid signal.^[31] An inversion pulse can also be used for lipid suppression, making use of the different T_1 values of lipid (typically 300 ms) versus metabolites (typically 1000–2000 ms).^[35] Choice of a short inversion time of approximately 200 ms ($= T_1 \cdot \ln(2)$) (derived from Eq. 19 with $S(\tau)=0$) will selectively null the lipid signal, and most of the metabolite signal will remain inverted. Because both water and lipid resonances have shorter T_2 relaxation times

than many metabolites, suppression factors are also usually better in long TE compared to short TE spectra.

Information Content of Proton MR Spectra of the Brain

Figure 11 shows examples of proton spectra recorded at long ($TE = 272$ ms, Fig. 11A) and short ($TE = 30$ ms, Fig. 11B) echo times. The largest metabolite signal, resonating at 2.02 ppm, occurs from the N-acetyl group of N-acetyl aspartate (NAA), with perhaps a small contribution from N-acetyl aspartyl glutamate (NAAG).^[36] An important property of NAA is that hardly any is found outside the nervous system.^[37,38] The evidence for NAA as a neuronal marker is widespread,^[27,39] for instance, NAA reductions measured by ^1H MRS were in excellent agreement with neuronal enzyme activity assays in an animal model of neuronal damage induced by striatal injection of kainic acid.^[40] In a similar model of excitotoxic neuronal injury,^[41] NAA levels were also shown to decrease in regions with neuronal loss. In addition, immunocytochemical staining of rat brain sections using monoclonal antibodies specific for NAA showed immunoreactivity in neurons and axons only.^[42] Decreased levels of NAA have also been observed in numerous human pathologies associated with neuronal and axonal loss, such as brain tumors (of glial origin), multiple sclerosis, stroke, or epilepsy.^[43–49] Although there is some evidence that NAA is present in glial cells (isolated oligodendrocyte preparations),^[68] on balance it appears that most of the NAA signal in mature brain is of neuronal-axonal origin, and that reductions in NAA may be a reasonably good surrogate marker for neuronal-axonal injury in many neuropathologies.^[69]

The choline signal (Cho, 3.24 ppm) arises from the $N(\text{CH}_3)_3$ groups of glycerophosphocholine (GPC), phosphocholine (PC), and a small amount of free choline, compounds that are involved in membrane synthesis and degradation. Both increases and decreases in Cho have been reported in pathological conditions; processes leading to elevation of choline signal include active demyelination,^[45] resulting from the degradation of myelin phospholipids primarily to GPC, or increased numbers of glial cells.^[50,51] Low Cho has been observed in hepatic encephalopathy,^[52] and there is also some evidence to suggest that dietary intake of choline can modulate cerebral Cho levels.^[53]

The creatine signal (Cr, 3.02 ppm) is a composite peak consisting of creatine and phosphocreatine (PCr), compounds involved in energy metabolism via the creatine kinase reaction generating ATP. Because creatine is synthesized in the liver, chronic liver disease leads to lower cerebral creatine concentration.^[27,42] In vitro, glial cells contain a two- to fourfold higher concentration of creatine compared to neurons.^[54] In the past, it was suggested that the sum of these Cr plus PCr may be relatively constant, both in different brain regions and in pathological processes. For this reason, Cr is sometimes used as a reference peak. However, it should also be recognized that there are quite strong regional variations in Cr^[70] and many conditions in which Cr will change (e.g., reduced due to general cell loss (as might occur, for example, in cerebral infarction)), so it is unwise to assume that Cr is invariant in most cases.

In normal human brain, lactate (1.33 ppm) is below (or at the limit of) detection in most studies. Any detectable increase in lactate can therefore be considered abnormal. Increased lactate is usually the result of deranged energy metabolism, and has been observed in ischemia (both acute (highest) and chronic^[46,47]), brain tumors,^[43] mitochondrial diseases,^[55] and other conditions. Small elevations of lactate have also been reported in the visual cortex during photic stimulation,^[56] believed to be due to increased nonoxidative glycolysis, but this effect does not appear to be particularly reproducible.^[57]

At short echo times, additional compounds are detected that are not visible at long echo times, because of either short T_2 relaxation times or the dephasing effects of J-coupling (Fig. 11B). One of the largest signals occurs from *myo*-inositol (mI) at 3.56 ppm. The function of mI is largely unknown, but it appears to be altered in many pathological conditions. It has been suggested that mI is found only in glial cells, and could therefore be a glial marker.^[58] mI has also been found to be reduced in hepatic encephalopathy^[59] and increased in Alzheimer's dementia^[60] and demyelinating diseases.^[61]

Glutamate (Glu) and glutamine (Gln) are difficult to separate in proton spectra at 1.5T and are often labeled as a composite peak Glx. At higher fields (such as 3 or 4T), the C4 resonances of Glu and Gln start to become resolved. Increased cerebral glutamine has been found in patients with liver failure (hepatic encephalopathy,^[59] Reye's syndrome^[52]) as the result of increased blood ammonia levels, which increase glutamine synthesis.

Under certain conditions (e.g., the use of editing sequences, high field strengths, or very short echo times), other signals can also be detected in proton brain spectra, such as GABA,^[62] glucose,^[63] *scyllo*-inositol,^[64] purine nucleotides,^[65] lipids, and macromolecules.^[66] Finally, occasionally exogenous substances that cross the blood-brain barrier can be detected by MRS; these include propandiol, mannitol, ethanol, and methylsulphonylmethane.

CONCLUSIONS

Many methods are presently available to perform magnetic resonance spectroscopy in vivo, all of which have been applied successfully to study cells, both animal and human; metabolic images of equilibrium levels of metabolites can be obtained on standard clinical equipment. In vivo MR spectroscopy of humans is possible in many organ systems and with many nuclei. Although ^1H spectroscopy has been demonstrated in a number of organ systems, the overwhelming number of applications have been in the brain, and human brain pathology has now been extensively studied by ^1H spectroscopy.

ACKNOWLEDGMENT

Supported in part by NIH grant P41 RR15241.

ARTICLES OF FURTHER INTEREST

Magnetic Resonance Imaging in Temperature Measurement; Medical Imaging, 3-D; Microcomputed Tomography and Its Applications

REFERENCES

1. Pauli, W. The theoretical significance of the satellites of some spectrum lines and the effect on them of magnetic fields. *Naturwissenschaften* **1924**, *12*, 741.
2. Purcell, E.U.; Torrey, H.C.; Pound, R.V. Resonance absorption by nuclear magnetic moments in a solid. *Phys. Rev.* **1946**, *69*, 37-38.

3. Bloch, F. Nuclear induction. *Phys. Rev.* **1946**, *70*, 460–474.
4. Proctor, W.G.; Yu, F.C. *Phys. Rev.* **1950**, *77*, 717.
5. Gutowsky, H.S.; McCall, D.W.; Slichter, C.P. Coupling among nuclear magnetic dipoles in molecules. *Phys. Rev.* **1951**, *84*, 589.
6. Radda, G.K. The use of NMR spectroscopy for the understanding of disease. *Science* **1986**, *233*, 640–645.
7. Le Bihan, D. *Diffusion and Perfusion Magnetic Resonance Imaging. Applications to Functional MRI*; Raven Press: New York, NY, 1995.
8. Moonen, C.T.W.; van Zijl, P.C.M.; Frank, J.A.; Le Bihan, D.; Becker, E.D. Functional magnetic resonance imaging in medicine and physiology. *Science* **1990**, *250*, 53–61.
9. Pople, J.A.; Schaefer, T. The analysis of complex nuclear magnetic resonance spectra. 1. Systems with one pair of strongly coupled nuclei. *Mol. Phys.* **1960**, *3*, 547–556.
10. Diehl, P.; Pople, J.A. The analysis of complex nuclear magnetic resonance spectra. 2. Some further systems with one strong coupling constant. *Mol. Phys.* **1960**, *3*, 557–561.
11. Ernst, R.R.; Anderson, W.A. Application of Fourier transform spectroscopy to magnetic resonance. *Rev. Sci. Instrum.* **1966**, *37*, 93–102.
12. Sorensen, O.W.; Eich, G.W.; Levitt, M.H.; Ernst, R.R. Product operator-formalism for the description of NMR pulse experiments. *Prog. NMR Spectr.* **1983**, *16*, 163–192.
13. Ernst, R.R.; Bodenhausen, G.; Wokaun, A. Principles of Nuclear Magnetic Resonance in One and Two Dimensions. In *International Series of Monographs on Chemistry 14*; Oxford University Press: Oxford, 1987; 1–610.
14. Hahn, E.L. Spin echoes. *Phys. Rev.* **1950**, *80*, 580–594.
15. Carr, H.Y.; Purcell, E.M. Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Phys. Rev.* **1954**, *94*, 630–638.
16. Meiboom, S.; Gill, D. Modified spin-echo method for measuring nuclear relaxation times. *Rev. Sci. Instrum.* **1958**, *29*, 688–691.
17. Wu, Y.; Ackerman, J.L.; Chesler, D.A.; Graham, L.; Wang, Y.; Glimcher, M.J. Density of organic matrix of native mineralized bone measured by water- and fat-suppressed proton projection MRI. *Magn. Reson. Med.* **2003**, *50*, 59–68.
18. Long, J.R.; Oyler, N.; Drobny, G.P.; Stayton, P.S. Assembly of alpha-helical peptide coatings on hydrophobic surfaces. *J. Am. Chem. Soc.* **2002**, *124*, 6297–6303.
19. Long, J.R.; Shaw, W.J.; Stayton, P.S.; Drobny, G.P. Structure and dynamics of hydrated statherin on hydroxyapatite as determined by solid-state NMR. *Biochemistry* **2001**, *40*, 15451–15455.
20. Park, S.H.; Mrse, A.A.; Nevzorov, A.A.; Mesleh, M.F.; Oblatt-Montal, M.; Montal, M.; Opella, S.J. Three-dimensional structure of the channel-forming trans-membrane domain of virus protein “u” (Vpu) from HIV-1. *J. Mol. Biol.* **2003**, *333*, 409–424.
21. Kurhanewicz, J.; Swanson, M.G.; Nelson, S.J.; Vigneron, D.B. Combined magnetic resonance imaging and spectroscopic imaging approach to molecular imaging of prostate cancer. *J. Magn. Reson. Imaging* **2002**, *16*, 451–463.
22. Luyten, P.R.; Groen, J.P.; Vermeulen, J.W.; den, H.J. Experimental approaches to image localized human 31P NMR spectroscopy. *Magn. Reson. Med.* **1989**, *11*, 1–21.
23. Arnold, D.L.; Shoubridge, E.A.; Emrich, J.; Feindel, W.; Villemure, J.G. Early metabolic changes following chemotherapy of human gliomas in vivo demonstrated by phosphorous magnetic resonance spectroscopy. *Invest. Radiol.* **1989**, *24*, 958–961.
24. Cadoux, H.T.; Blackledge, M.J.; Rajagopalan, B.; Taylor, D.; Radda, G.K. Human primary brain tumour metabolism in vivo: A phosphorus magnetic resonance spectroscopy study. *Br. J. Cancer* **1989**, *60*, 430–436.
25. Levine, S.R.; Helpert, J.A.; Welch, K.M.; Vande Linde, A.M.; Sawaya, K.L.; Brown, E.E.; Ramadan, N.M.; Deveshwar, R.K.; Ordidge, R.J. Human focal cerebral ischemia: Evaluation of brain pH and energy metabolism with P-31 NMR spectroscopy. *Radiology* **1992**, *185*, 537–544.
26. Frahm, J.; Bruhn, H.; Gyngell, M.L.; Merboldt, K.D.; Hanicke, W.; Sauter, R. Localized high-resolution proton NMR spectroscopy using stimulated echoes: Initial applications to human brain in vivo. *Magn. Reson. Med.* **1989**, *9*, 79–93.
27. Ross, B.D.; Michaelis, T. Clinical applications of magnetic resonance spectroscopy. *Magn. Reson. Q.* **1994**, *10*, 191–247.
28. Bottomley, P. *Selective Volume Method for Performing Localized NMR Spectroscopy*; U.S.A., 1984. Inventor.
29. Luyten, P.R.; Marien, A.J.; Heindel, W.; van Gerwen, P.H.; Herholz, K.; Hollander, den J.A.; Friedmann, G.; Heiss, W.D. Metabolic imaging of patients with intracranial tumors: H-1 MR spectroscopic imaging and PET. *Radiology* **1990**, *176*, 791–799.
30. Brown, T.R.; Kincaid, B.M.; Ugurbil, K. NMR chemical shift imaging in three dimensions. *Proc. Natl. Acad. Sci. U. S. A.* **1982**, *79*, 3523–3526.
31. Duyn, J.H.; Gillen, J.; Sobering, G.; van Zijl, P.C.M.; Moonen, C.T.W. Multislice proton MR spectroscopic imaging of the brain. *Radiology* **1993**, *188*, 277–282.
32. Haase, A.; Frahm, J.; Hanicke, W.; Matthaei, D. 1H NMR chemical shift selective (CHESS) imaging. *Phys. Med. Biol.* **1985**, *30*, 341–344.
33. Ogg, R.J.; Kingsley, P.B.; Taylor, J.S. WET, a T1- and B1-insensitive water-suppression method for in vivo localized 1H NMR spectroscopy. *J. Magn. Reson. B.* **1994**, *104*, 1–10.
34. Moonen, C.; van Zijl, P. Highly efficient water suppression for in vivo proton NMR spectroscopy. *J. Magn. Reson.* **1990**, *88*, 28–41.
35. Spielman, D.M.; Pauly, J.M.; Macovski, A.; Glover, G.H.; Enzmann, D.R. Lipid-suppressed single- and multisection proton spectroscopic imaging of the human brain. *JMRI* **1992**, *2*, 253–262.



36. Frahm, J.; Michaelis, T.; Merboldt, K.D.; Hanicke, W.; Gyngell, M.L.; Bruhn, H. On the N-acetyl methyl resonance in localized ^1H NMR spectra of the human brain in vivo. *NMR Biomed.* **1991**, *4*, 201–204.
37. Birken, D.L.; Oldendorf, W.H. N-acetyl-L-aspartic acid: A literature review of a compound prominent in ^1H NMR spectroscopic studies of brain. *Neurosci. Behav. Rev.* **1989**, *13*, 23–31.
38. Tallan, H.H. Studies on the distribution of N-acetyl-L-aspartic acid in brain. *J. Biol. Chem.* **1957**, *224*, 41–45.
39. Howe, F.A.; Maxwell, R.J.; Saunders, D.E.; Brown, M.M.; Griffiths, J.R. Proton spectroscopy in vivo. *Magn. Reson. Q.* **1993**, *9*, 31–59.
40. Guimaraes, A.R.; Schwartz, P.; Prakash, M.R.; Carr, C.A.; Berger, U.V.; Jenkins, B.G.; Coyle, J.T.; Gonzalez, R.G. Quantitative in vivo ^1H nuclear magnetic resonance spectroscopic imaging of neuronal loss in rat brain. *Neuroscience* **1995**, *69*, 1095–1101.
41. Strauss, I.; Williamson, J.M.; Bertram, E.H.; Lothman, E.W.; Fernandez, E.J. Histological and ^1H magnetic resonance spectroscopic imaging analysis of quinolinic acid-induced damage to the rat striatum. *Magn. Reson. Med.* **1997**, *37*, 24–33.
42. Simmons, M.L.; Frondoza, C.G.; Coyle, J.T. Immunocytochemical localization of N-acetyl-aspartate with monoclonal antibodies. *Neuroscience* **1991**, *45*, 37–45.
43. Alger, J.R.; Frank, J.A.; Bizzi, A.; Fulham, M.J.; DeSouza, B.X.; Duhaney, M.O.; Inscoc, S.W.; Black, J.L.; van Zijl, P.C.M.; Moonen, C.T.W.; Dichiro, G. Metabolism of human gliomas: Assessment with H-1 MR spectroscopy and F-18 fluorodeoxyglucose PET (see comments). *Radiology* **1990**, *177*, 633–641.
44. Arnold, D.L.; Riess, G.T.; Matthews, P.M.; Francis, G.S.; Collins, D.L.; Wolfson, C.; Antel, J.P. Use of proton magnetic resonance spectroscopy for monitoring disease progression in multiple sclerosis. *Ann. Neurol.* **1994**, *36*, 76–82.
45. Davie, C.A.; Hawkins, C.P.; Barker, G.J.; Brennan, A.; Tofts, P.S.; Miller, D.H.; McDonald, W.I. Detection of myelin breakdown products by proton magnetic resonance spectroscopy. *Lancet* **1993**, *341*, 630–631.
46. Barker, P.B.; Gillard, J.H.; van Zijl, P.C.; Soher, B.J.; Hanley, D.F.; Agildere, A.M.; Oppenheimer, S.M.; Bryan, R.N. Acute stroke: Evaluation with serial proton MR spectroscopic imaging. *Radiology* **1994**, *192*, 723–732.
47. Petroff, O.A.; Graham, G.D.; Blamire, A.M.; al-Rayess, M.; Rothman, D.L.; Fayad, P.B.; Brass, L.M.; Shulman, R.G.; Prichard, J.W. Spectroscopic imaging of stroke in humans: Histopathology correlates of spectral changes. *Neurology* **1992**, *42*, 1349–1354.
48. Cendes, F.; Andermann, F.; Preul, M.C.; Arnold, D.L. Lateralization of temporal lobe epilepsy based on regional metabolic abnormalities in proton magnetic resonance spectroscopic images. *Ann. Neurol.* **1994**, *35*, 211–216.
49. Hugg, J.W.; Laxer, K.D.; Matson, G.B.; Maudsley, A.A.; Weiner, M.W. Neuron loss localizes human temporal lobe epilepsy by in vivo proton magnetic resonance spectroscopic imaging. *Ann. Neurol.* **1993**, *34*, 788–794.
50. Gill, S.S.; Small, R.K.; Thomas, D.G.; Patel, P.; Porteous, R.; Van Bruggen, N.; Gadian, D.G.; Kauppinen, R.A.; Williams, S.R. Brain metabolites as ^1H NMR markers of neuronal and glial disorders. *NMR Biomed.* **1989**, *2*, 196–200.
51. Gill, S.S.; Thomas, D.G.T.; Van Bruggen, N.; Gadian, D.G.; Peden, C.J.; Bell, J.D.; Cox, I.J.; Menon, D.K.; Iles, R.A.; Bryant, D.J.; Coutts, G.A. Proton MR spectroscopy of intracranial tumours: In vivo and in vitro studies. *J. Comput. Assist. Tomogr.* **1990**, *14*, 497–504.
52. Kreis, R.; Ross, B.D.; Farrow, N.A.; Ackerman, Z. Metabolic disorders of the brain in chronic hepatic encephalopathy detected with H-1 MR spectroscopy. *Radiology* **1992**, *182*, 19–27.
53. Stoll, A.L.; Renshaw, P.F.; De Micheli, E.; Wurtman, R.; Pillay, S.S.; Cohen, B.M. Choline ingestion increases the resonance of choline-containing compounds in human brain: An in vivo proton magnetic resonance study. *Biol. Psychiatry* **1995**, *37*, 170–174.
54. Urenjak, J.; Williams, S.R.; Gadian, D.G.; Noble, M. Proton nuclear magnetic resonance spectroscopy unambiguously identifies different neural cell types. *J. Neurosci.* **1993**, *13*, 981–989.
55. Mathews, P.M.; Andermann, F.; Silver, K.; Karpati, G.; Arnold, D.L. Proton MR spectroscopic characterization of differences in regional brain metabolic abnormalities in mitochondrial encephalomyopathies. *Neurology* **1993**, *43*, 2484–2490.
56. Prichard, J.; Rothman, D.; Novotny, E.; Petroff, O.; Kuwabara, T.; Avison, M.; Howseman, A.; Hanstock, C.; Shulman, R. Lactate rise detected by ^1H NMR in human visual cortex during physiologic stimulation. *Proc. Natl. Acad. Sci. U. S. A.* **1991**, *88*, 5829–5831.
57. Merboldt, K.D.; Bruhn, H.; Hanicke, W.; Michaelis, T.; Frahm, J. Decrease of glucose in the human visual cortex during photic stimulation. *Magn. Reson. Med.* **1992**, *25*, 187–194.
58. Flogel, U.; Willker, W.; Leibfritz, D. Regulation of intracellular pH in neuronal and glial tumour cells, studied by multinuclear NMR spectroscopy. *NMR Biomed.* **1994**, *7*, 157–166.
59. Ross, B.D.; Jacobson, S.; Villamil, F.; Korula, J.; Kreis, R.; Ernst, T.; Shonk, T.; Moats, R.A. Subclinical hepatic encephalopathy: Proton MR spectroscopic abnormalities. *Radiology* **1994**, *193*, 457–463.
60. Shonk, T.K.; Moats, R.A.; Gifford, P.; Michaelis, T.; Mandigo, J.C.; Izumi, J.; Ross, B.D. Probable Alzheimer disease: Diagnosis with proton MR spectroscopy. *Radiology* **1995**, *195*, 65–72.
61. Kruse, B.; Hanefeld, F.; Christen, H.J.; Bruhn, H.; Michaelis, T.; Hanicke, W.; Frahm, J. Alterations of brain metabolites in metachromatic leukodystrophy as detected by localized proton magnetic resonance spectroscopy in vivo. *J. Neurol.* **1993**, *241*, 68–74.
62. Rothman, D.L.; Petroff, O.A.; Behar, K.L.; Mattson, R.H. Localized ^1H NMR measurements of gamma-aminobutyric acid in human brain in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 5662–5666.

63. Gruetter, R.; Garwood, M.; Ugurbil, K.; Seaquist, E.R. Observation of resolved glucose signals in ^1H NMR spectra of the human brain at 4 Tesla. *Magn. Reson. Med.* **1996**, *36*, 1–6.
64. Michaelis, T.; Helms, G.; Merboldt, K.D.; Hanicke, W.; Bruhn, H.; Frahm, J. Identification of Scyllo-inositol in proton NMR spectra of human brain in vivo. *NMR Biomed.* **1993**, *6*, 105–109.
65. van Zijl, P.C.; Moonen, C.T. In situ changes in purine nucleotide and N-acetyl concentrations upon inducing global ischemia in cat brain. *Magn. Reson. Med.* **1993**, *29*, 381–385.
66. Hwang, J.H.; Graham, G.D.; Behar, K.L.; Alger, J.R.; Prichard, J.W.; Rothman, D.L. Short echo time proton magnetic resonance spectroscopic imaging of macromolecule and metabolite signal intensities in the human brain. *Magn. Reson. Med.* **1996**, *35*, 633–639.
67. Kvistad, K.A.; Bakken, I.J.; Gribbestad, I.S.; Ehrnholm, B.; Lundgren, S.; Fjosne, H.E.; Haraldseth, O. Characterization of neoplastic and normal breast tissues with ^1H MR spectroscopy. *JMRI* **1999**, *10*, 159–164.
68. Bhakoo, K.K.; Pearce, D. In vitro expression of N-acetyl aspartate by oligodendrocytes: Implications for proton magnetic resonance spectroscopy signal in vivo. *J. Neurochem.* **2000**, *74* (1), 254–262.
69. Barker, P.B. N-acetyl aspartate—A neuronal marker? *Ann. Neurol.* **2001**, *49*, 423–424.
70. Jacobs, M.A.; Horska, A.; van Zijl, P.C.M.; Barker, P.B. Quantitative proton MR spectroscopic imaging of normal human cerebellum and brain stem. *Magn. Reson. Med.* **2001**, *46*, 699–705.

Ocular Implants

Andrew W. Lloyd

Susan Sandeman

Richard G. A. Faragher

Biomedical Materials Research Group, University of Brighton, Brighton, United Kingdom

Stephen P. Denyer

Biomedical Materials Research Group, University of Brighton, Brighton, and Welsh School of Pharmacy, Cardiff University, Cardiff, United Kingdom

INTRODUCTION

The human eye is a complex organ of vital importance for everyday life and a major target for a range of implants and accessory biomedical devices. Biomaterials are used to fabricate ocular devices to correct functional deficiencies caused by disease, age, and ocular trauma. This article examines the various types of ocular implants, with particular reference to the approaches adopted for improving ocular compatibility and ensuring the longevity of indwelling ocular devices.

ORBITAL IMPLANTS

In contrast to many indwelling devices, orbital implants are employed with mainly cosmetic intent.^[1] A wide range of problems can lead to the need to eviscerate or enucleate the eye including retinoblastoma, trauma, uveitis, and rubeotic glaucoma. Removal of the globe or its internal contents requires the insertion of a solid implant to maintain the orbital volume; device motility is also desirable to optimize cosmesis. A range of materials have been used to manufacture orbital implants including glass, cork, ivory, and aluminium.^[1] It is now recognized that cellular invasion and vascularization is a desirable characteristic of orbital implants because it discourages bacterial colonization of the surface and permits treatment of low-grade ocular infection with systemic antibiotics.^[2] For this reason, modern implants are frequently based upon hydroxyapatite or porous polyethylene.^[3,4] Evisceration involves the removal of the contents of the globe, introduction of the implant into the scleral shell, and conjunctival closure. Enucleations involve direct suturing of the extraocular muscles to the implant, or, more commonly, to the implant wrapped in donor or synthetic material, and the closure of Tenon's capsule followed by separate conjunctival closure. Porous implants have been shown to have an overall

failure rate of approximately 10% with higher failure rate when used in eviscerations than enucleations. Failed implants that have been histologically characterized reveal a failure of fibroblast ingrowth and a chronic inflammatory response. It has been suggested that such a response can be partially avoided by wrapping the implant in autologous donor tissue.^[5]

KERATOPROSTHESES AND INTRACORNEAL IMPLANTS

On occasions, direct intervention surgery is necessary to replace a diseased cornea or alter its refractive state; such surgery can involve the implantation of permanent indwelling devices such as keratoprotheses or intracorneal implants. The former devices involve full-thickness penetration of the cornea and result in the full substitution of corneal function, whereas the latter are implanted within the cornea to improve its refractive performance.

The intracorneal lens is implanted in the central corneal stroma to correct conditions such as myopia and hyperopia. Modern intracorneal lenses are most commonly fashioned from PMMA, which combines a range of good optical and mechanical properties with acceptable biocompatibility.^[6] An intrinsic problem with solid PMMA lenses is the disruption of nutrient transport across the cornea; mechanical irritation also renders such implants less than ideal. Epithelial thinning over the implant, reduced keratocyte density anterior to the lens, and stromal opacity have all been reported as side effects of such implants.^[7]

To address the problem of nutrient flow, hydrogel intracorneal lenses have been evaluated. Primate trials indicate that these implants are well tolerated for at least five years.^[16] Clinical studies indicated a reasonable initial correction of hyperopia although this had returned to preoperative values within one year. Lens migration and epithelial cell invasion were reported as significant side effects in approximately 10% of

patients. A more generic issue is the relatively low refractive index of currently available hydrogels relative to PMMA.^[6] These implants exert their refractive effect principally through the alteration of corneal curvature.

In contrast to hydrogels, polysulfone lenses have extremely high refractive indices and can thus be made extremely thin, simplifying the surgery needed to implant them. Polysulfones, however, are impermeable to aqueous solutions and thus disrupt nutrient flow in an analogous manner to PMMA; an acceptable compromise appears to be the fenestration of polysulfone lenses with a network of 10- μ m diameter holes to allow fluid movement.^[8] Such holes allow the retention of adequate optical quality and appear to be too small to permit cellular invasion of the implant.

An alternative method for the correction of myopia involves the implantation of an intrastromal ring designed to alter the anterior corneal curvature.^[9] Such rings are inserted into a surgically prepared enclosed tunnel within the cornea and flatten the central corneal curvature. Varying ring thickness alters the degree of flattening produced. Intrastromal rings are generally manufactured from PMMA and are tolerated in both rabbit and human eyes.

In addition to intrinsic functional deficits, a wide range of external hazards can permanently damage the healthy cornea. Penetrating injury, alkali burn (particularly exposure to lime in the building industry), viral infection, and ulceration can damage and opacify the cornea. The most common clinical response in this situation is to perform a penetrating keratoplasty (a corneal graft) using a donor cornea. Unfortunately, supplies of donor corneas are limited and graft failure is common, with the risk increasing over time. In particular, patients with severe chemical burns, chronic ocular surface inflammation, herpes simplex viral infection, and Stevens-Johnson syndrome have a poor prognosis for grafting. A variety of solutions have been offered including corneal repopulation using cell lines^[10] and tissue grafts based on hyaline cartilage. In the immediate term, however, polymer-based tissue replacements, keratoprotheses, appear to offer the greatest potential for success. A range of materials and designs of these implants are now available.

In contrast to intracorneal implants, keratoprotheses are penetrating replacements of the cornea. A general problem with keratoprotheses is thus the maintenance of a clear visual window while allowing sufficient cellular invasion to fix the implant firmly in place. Extrusion of the keratoprosthesis as a result of internal globe pressure is a common problem resulting from a failure to promote wound healing at the implant margins. The original keratoprotheses designed by Cardona and coworkers in the 1960s had a failure rate of over 30% as a result of tissue erosion followed

by extrusion of the implant.^[11] Other problems with such keratoprotheses result from a fixed implant size and a fixed surgical position. In the original Cardona-style keratoprosthesis a nut-and-bolt manufactured from high-quality PMMA^[12] was employed. Following surgical removal of the epithelial layer and extraction of the underlying corneal stroma and endothelium, the implant was secured into the extracted tissue by means of a trephined bolt hole and the tissue containing the keratoprosthesis was then sutured back into the eye. Overtightening of the implant had a regrettable tendency to trigger aseptic necrosis of the tissue.

Girard keratoprotheses are modifications of the Cardona type.^[13] Doane et al. have manufactured twin-plate keratoprotheses from PMMA but with design modifications that permit their use in very dry eyes.^[14] These have been used in the clinical situation. Attempts have also been made to improve the retention of such solid keratoprotheses by surface modification, but, in general, PMMA keratoprotheses have not achieved large-scale usage or great success. It has been suggested that an ideal keratoprosthesis should be 1) tightly retained in the cornea to prevent extrusion, 2) easily and completely colonized by corneal epithelial cells on the external face, and 3) should suppress downgrowth of such cells into the implant bed. Keratoprotheses have also been fashioned from a range of materials including melt-blown polyolefins and vitreous carbon. Studies using the commercial polytetrafluoroethylene (PTFE) Proplast demonstrated over 50% implant extrusion after four months.^[15] Comparative studies using Impra, Gore-tex, and Proplast PTFE demonstrated that a significant invasive response was promoted only by Impra.^[16]

As water impermeable materials generally appear to be poorly tolerated by the residual corneal tissue, considerable interest is being shown in the exclusive use of pHEMA-based hydrogels for the fabrication of new keratoprotheses.^[17] However, the results of clinical trials using such materials have shown considerable patient variability.^[18]

An alternative device, the osteo-odonto-keratoprosthesis (OOKP), was developed in Italy by Strampelli in 1963.^[19] This device uses a lamina prepared from a single-root tooth extracted from the patient. A PMMA optic is cemented into the center of the lamina and the device is implanted in a subcutaneous pocket below the eye in the patient's cheek. After three months the device is removed from the pocket, the optic is placed through a hole trephined in the cornea into the anterior chamber, and the device is sutured over the cornea below a covering of buccal mucosa. Although complex, this approach probably has had the greatest clinical success to date. Further details on these various technologies may be found in the article entitled *Corneal Implants*.

SCLERAL BUCKLES

Retinal detachment complicates a small percentage of natural age-related posterior vitreous detachment cases and is routinely repaired using a scleral implant known colloquially as a buckle. Buckles indent the sclera bringing the separated layers of the neural retina and the underlying pigment epithelium together. Since the introduction of this implant technique in the 1930s a range of materials have been used to manufacture the actual buckle.^[20] A broad distinction can be made between absorbable buckles, which appear to have limited advantages and many problems, and non-absorbable implants.

General problems associated with the use of scleral buckles include implant extrusion, chronic irritation, alterations in the intraocular pressure of the buckled globe, decreased extraocular motility, and infection. Absorbable materials address the problem of extrusion by being broken down and replaced by scar tissue in the normal course of tissue turnover. The simplest absorbable materials consist of autogenous tendon or fascia lata from the patient.^[20] Such grafts have the advantage that rejection is eliminated but the patient is subject to extra surgery. Cadavers are occasionally used as sources of donor material, in which case lyophilized or fixed samples of dura mater or sclera are implanted. Cadaver material that is generally stored in ethanol or lyophilized is generally more rigid than autologous graft material; this increases the risk of implant erosion. Host rejection and transmission of a novel pathogen are also associated problems with the use of such material. A generic problem with grafts is their lack of elasticity, which limits the ability to produce scleral buckles of exact dimensions.

A variety of other degradable materials have been used for buckles including stabilized fibrin, absorbable gut, and collagen. However, the most widely used protein in this respect remains pigskin gelatin. Gelatin has the advantage that the implant can be precisely shaped and impregnated with antibiotics to control potential infections. Reabsorption proceeds smoothly over a three- to six-month period, however it has a regrettable tendency to trigger low-grade inflammation of the eye that can escalate to include vitreous haze and macular edema. All absorbable buckles suffer from the problem that reabsorption of the implant reduces the buckling effect and the removal of a partially absorbed implant is a difficult surgical technique. Temporary buckles based on injection of plasma into the suprachoroidal space have also been attempted; this produces a buckling effect for approximately one week. Plasma injection, and related techniques such as the use of hyaluronan, can, however, cause choroidal perforation and substantial hemorrhaging. In addition, control of buckle shape and size by injection is also very poor.

Nonabsorbable devices have undergone extensive design work in recent years and now offer distinct advantages over the absorbable materials. A range of polymers have been used to produce nonabsorbable implants, particularly silicones and hydrogels. Polyethylene tubing has been used in the production of encircling scleral implants since the 1950s. The main advantage of this material was the facility to introduce a suture into the lumen of the tubing that could then be tensioned to regulate buckle height. An incidental advantage of the use of such tubing was the opportunity to fill the lumen with antibiotics to reduce the risk of infection. However, the buckles produced using polyethylene tubing tended to be narrow, infections still occurred, and the tubes themselves had a tendency to erode the underlying sclera and choroid over time. To address these problems, implants based on softer silicones and silicone sponges were introduced. These materials remain flexible when implanted and thus do not cause erosion. A further advantage of the use of silicones is the formation of a tough fibrous capsule around the implant that both strengthens the sclera and allows the easy removal of the implant should reintervention be required. Soft silicones tend to be superior to sponges in this respect due to their greater surface smoothness but sponges do produce a more even buckling effect and can also be impregnated with antibiotics to limit bacterial colonization of the implant. Unfortunately, silicone implants show a post-operative infection rate of 2–5%.^[21] An elastic material with a higher endogenous biocompatibility might offer advantages in this regard.

Poly(glyceryl methacrylate) (PGMA), poly(2-hydroxyethyl acrylate) (PHEA) and co-poly(methyl acrylate-2-hydroxyethyl acrylate) (MAI) have all been used as base materials for scleral buckles. The major advantages of the hydrogels are softness, defined swelling on hydration, and the potential to act as drug delivery devices for antibiotics. Of the three hydrogels commonly used, PGMA suffers from a lack of tensile strength when swollen and PHEA has a tendency to fragment after swelling. MAI has good bulk characteristics and promotes the formation of a strong surrounding capsule. As a result of these properties it is the most commonly used hydrogel and is sold under the commercial name Miragel. Although these various technologies have been developed, the clinical preference remains for the silicone tire or sponge.

GLAUCOMA FILTRATION IMPLANTS (GFIs)

Glaucoma is a common cause of severe visual handicap in which raised intraocular fluid pressure (IOP) causes progressive optic nerve damage and visual field loss. The principal surgical treatment, glaucoma

filtration surgery (GFS), involves the formation of a drainage passage through the sclera/cornea connecting the anterior chamber of the eye to the subconjunctival space, bypassing the normal drainage routes. Although GFS is generally highly successful, in a number of cases an artificial drainage implant is required. Early approaches included the use of various types of seton made from a range of materials including horse hair, silk, metals (such as gold, platinum, and magnesium), and polymers such as nylon. These early devices had high rates of filtration failure primarily as a consequence of the chronic foreign body inflammatory response caused by the devices, which led to subconjunctival fibrosis and fistula closure. In 1969, Molteno introduced a two-piece drainage implant that comprised a drainage tube attached to a plate device.^[22] The plate, which is sutured to the sclera under a quadrant limbal conjunctival flap, was designed to maintain the patency of the subconjunctival filtration reservoir during subconjunctival fibrosis. Since then a range of different tube and plate devices have been developed.^[23] The early Molteno and Baerveldt devices had no resistance elements and suffered from problems with early-stage hypotony. Molteno attempted to address this by developing a two-stage surgical technique in which the device was implanted and allowed to encapsulate before insertion of the tube element into the anterior chamber in a second operation. Other approaches that have been adopted to overcome this problem include various ligation procedures and temporary occlusion of the tube. Later embodiments of the Molteno and Baerveldt GFIs incorporated plate-based resistance elements that depend on tissue fibrosis over the plate to restrict the flow. Other workers have developed various alternative systems, including the use of microtubules and various types of valve, to restrict fluid flow and avoid postoperative hypotony. Recent in-vitro studies of these devices have indicated considerable variation in resistance values obtained from devices from the same manufacturer, indicating deficiencies in manufacturers' quality control and likely variability between implanted devices.^[24] The recommended technique for insertion of the device involves the formation of a paracentesis track using a 22- or 23-gauge hypodermic needle. As a consequence, uncontrolled extrinsic leakage caused by a poor tube to paracentesis fit is believed to contribute to the problems of hypotony following device implantation. Other common complications include chronic inflammation, internal occlusion of the tube lumen, infection, and corneal endothelial damage from tube contact. It has been suggested that these problems are related to both the micromovement of the device following implantation causing tissue irritation and inappropriate selection of device material for optimal ocular compatibility.

INTRAOCULAR LENSES

Cataract is the most common treatable form of blindness throughout the world. Cataract extraction and intraocular lens implantation is the most frequent form of ophthalmic surgical procedure, with an excess of 1.6 million operations being performed per annum in the United States alone.

The surgical method most frequently employed in the developed world for the treatment of this condition involves removal of the cataractous material while leaving most of the lens capsule in place and insertion of an intraocular lens (IOL) to compensate for the loss of the natural crystalline lens. Of increasing popularity is the technique of phacoemulsification, in which the hardened nucleus of the crystalline lens is emulsified and removed through a 3.5-mm incision into the eye, as opposed to a 14-mm incision for a standard extracapsular cataract extraction. The advantages of this technique lie in the smaller, generally sutureless wound, causing quicker, safer rehabilitation and less surgically induced astigmatism.

During cataract extraction the anterior lens capsule is opened and the contents of the capsular bag removed. Following this, it is necessary to correct the loss of refracting power of the lens by the use of either a spectacle correction, which results in magnification and distortion of the image seen by the patient, a contact lens with all the complications associated with maintenance, or an IOL. Three types of IOL have been used: an anterior chamber IOL that sits in front of the iris but behind the cornea, an iris clip lens that straddles the pupil, and finally a posterior chamber IOL that sits behind the iris within or on the capsular bag. This latter type of lens is now the lens of choice.

PMMA was the first IOL material developed for the purpose by Harold Ridley in 1949. These IOLs are generally lathe-cut from PMMA rods or buttons and comprise a central optic that is supported by haptics, projections from the main body of the lens. The haptics are usually constructed from PMMA or the base material of the optic, although other polymers may be used. Current lenses usually employ a 7-mm-diameter optic, a 6.5-mm oval-shaped optic, or a 5-mm round optic. These posterior chamber lenses may thus be inserted wholly within the capsular bag or with the optic supported within the capsular bag remnants and the haptics lodged in the ciliary sulcus, the anatomical groove between the iris and the ciliary body. The lenses require a large incision and are now less frequently used, being unsuitable for small incision phacoemulsification surgery.

IOLs reside within the eye as foreign bodies, but were until recently largely regarded as inert. Indeed, PMMA was selected as the material of choice because of its low weight and biocompatibility. Although this

material has been used for over 40 years, major problems still occur as a consequence of its relatively low surface energy, which may result in both corneal endothelial damage on insertion and postoperative adhesion of inflammatory cells to the IOL. Uveal contact with the IOL surface has also been shown to cause increased and prolonged postoperative intraocular inflammation, which may lead to iris adhesion to the IOL, uveitis, breakdown of the blood retinal barrier causing cystoid macular edema, or loss of vision.^[25]

Attempts to improve the biocompatibility of IOLs have included: process modifications to produce a highly polished surface, the generation of both soft, high-energy surfaces using NVP and HEMA and hard low-energy surfaces using perfluoropropane, and the binding of heparin and hyaluronic acid to the outer surface of the lens. More recently, the use of phosphorylcholine-based polymeric coatings has been reported.^[26] These latter coatings were shown to reduce protein adsorption, cellular adhesion, and neutrophil activation by the PMMA surface and were shown in vivo to reduce cellular deposition onto the lens. Such materials have the propensity to reduce material-associated postimplantation opacification, which appears to occur as a consequence of an ongoing inflammatory response to the implanted material.

Phacoemulsification, with its small incision, has encouraged the development of foldable IOLs for implantation. The need to insert these devices through a 3.5-mm incision has also encouraged companies to investigate other design modalities including the concept of plate design IOLs. These foldable lenses have been manufactured from silicone elastomers, collagen copolymers, PHEMA hydrogels, and acrylic polymers. In developing these materials, particular emphasis has been placed on the handling, foldability, and unfolding characteristics of the lenses since they must be easy to insert, unfold slowly, and leave no crease mark. In this respect, it has been reported that the acrylic lens unfolds more slowly and in a more controlled fashion than the silicone lenses; the higher refractive index of the acrylic material gives rise to a thinner IOL. Evidence suggests that the hydrophilic materials are less damaging to the corneal endothelium and produce lower levels of inflammatory response in terms of cellular adhesion and foreign body response when compared to PMMA IOLs. Studies have also suggested that phosphorylcholine-based acrylate polymers may have application in the development of novel biocompatible foldable IOLs.^[27]

Recent emphasis has focused on the development of materials that will reduce posterior capsule opacification (PCO). This phenomenon is attributed to the migration of lens epithelial cells across the posterior capsule bag. Although this secondary cataract can be

treated using a neodymium:YAG laser to rupture the capsular bag, poor focussing can result in dislocation, decentralization, fragmentation, and pitting of PMMA lenses. Acrylic, silicone, and PHEMA IOLs have been shown to be more tolerant to the YAG laser than PMMA. Furthermore, recent in-vivo studies have suggested that PCO occurs less often following implantation of acrylic IOLs than other materials. This may be attributed to the physical inhibition of lens cell migration by adhesion of the posterior capsule to the IOL following implantation. Other approaches that have been investigated to reduce PCO include the modification of the IOL surface with antimetabolites. Although these approaches appear to be effective in in-vitro models, there is some concern over the effects of these antimetabolites on other ocular tissues.

Unlike the natural crystalline lens, standard IOLs provide no means of accommodation for near and far vision. In most cases the visual adjustment is achieved through the wearing of spectacles. Various approaches have been investigated for the fabrication of lenses that utilize the natural muscular action exerted on the capsular bag as a means of accommodation.^[28] A two-piece lens system has been patented in which the distance between the two lenses is controlled by the pressure exerted by the ciliary muscle on a U-shaped flange connecting the periphery of the two lenses. Relaxation of the ciliary muscle causes the lens to flatten, thereby changing the focal distance of the lens. Other approaches rely on the postoperative healing following surgery, during which the anterior capsule remnant fuses to the posterior capsule causing fibrosis around the lens haptic on the implanted lens, which leads to a rearward deflection of the lens against the posterior capsule. Accommodation is achieved through contraction and relaxation of the ciliary muscle that relaxes and stretches the fused remnant causing the lens to move forward and backward relative to the retina, thereby adjusting the focal distance. These approaches have recently led to the clinical evaluation of a number of accommodating lens designs.^[29] Other approaches include the use of viscoelastic IOL implants that may be injected into the capsular bag following removal of the natural crystalline lens. The viscoelastic gel's dimensions and refractive properties are modified as the ciliary muscle contracts and relaxes causing the capsular bag to either stretch or bulge, respectively.^[30]

Recent concern over the potential damage to the posterior segment of the eye arising from UV radiation following removal of the natural crystalline lens has encouraged companies to include UV absorbing chromophores in IOL materials. These chromophores are generally derivatives of benzotriazole that absorb UV radiation below 400 nm and that are either physically blended or chemically incorporated into the IOL material.^[31]

SILICONE OIL

Various agents have been used in conjunction with scleral buckles in the treatment of retinal detachment. These agents are injected into the eye to provide an intravitreal tamponade. Silicone oil is the most commonly used material despite various problems including poor biocompatibility and emulsification within the eye.^[32] In addition to the use of silicone oil as a tamponade, various workers have investigated the potential for the use of silicone oil for the delivery of antiproliferative agents for the treatment of proliferative vitreoretinopathy.^[33] This is a field of ophthalmology that could potentially benefit significantly from the development of novel high-density liquid biocompatible materials.

VISCOELASTIC AGENTS

Viscoelastic agents play an important role in ophthalmic surgery for both protection of the corneal endothelium during cataract extraction and in the treatment of retinal detachment. Corneal endothelial cell damage, particularly following implantation of anterior chamber IOLs, was until recently a common occurrence. This problem was overcome to some extent with the increased use of posterior IOLs that are placed in the capsular bag rather than in front of the iris. However, the insertion and removal of IOLs and surgical instruments was still found to cause some corneal endothelial cell damage. To prevent this, viscoelastic agents have been introduced into the anterior chamber to protect the corneal endothelium while allowing the surgeon to insert and remove optical instruments and implants. These agents have been prepared from a variety of materials including hydroxypropylmethylcellulose (HPMC), hyaluronic acid, chondroitin sulphate, polyacrylamide, collagen, and various mixtures of these materials.^[34] 1% hyaluronic acid has also been successfully employed as an internal tamponade in the treatment of retinal detachment but is generally less effective in patients with proliferative vitreoretinopathy than other agents.^[35]

DRUG DELIVERY

Various biomaterials have been employed for drug delivery to the eye including in-situ forming gels, nanoparticulates, and microparticulates; these have been recently reviewed by Ding (1998).^[36] Various implantable devices are in commercial use including Ocusert and Vitrasert with several other systems under development including osmotic minipumps, biodegradable scleral plugs, and biodegradable

poly(caprolactone) matrices. The use of biomaterials in ocular drug delivery will inevitably increase in future years and there will be a need to improve biocompatibility and patient comfort in these implanted systems.

CONCLUSION

This article illustrates the extensive range of biomaterials that have found use in the development of indwelling ocular implants. In many cases, these materials have been selected primarily on the basis of their mechanical properties and biofunctionality without full consideration of their ocular compatibility. With an improved understanding of the factors that affect material-related ocular compatibility has come the realization of the importance of material surface characteristics. Companies will need to consider both surface characterization and the assessment of ocular compatibility in the development of new biomaterials for ocular applications.

ACKNOWLEDGMENTS

This article has been reproduced in part from Biomaterials, Vol 22, Lloyd et al: 'Ocular Biomaterials and Implants,' pp. 769–785, Copyright (2001) with permission from Elsevier.

ARTICLES OF FURTHER INTEREST

Ocular Implants for Drug Delivery; Tissue Engineering of Cornea; Contact Lenses: Silicone Hydrogels

REFERENCES

1. Perry, A.C. Advances in enucleation. *Ophthalmol. Clin. North Am.* **1991**, *4*, 173–182.
2. Rubin, P.A.D.; Popham, J.K.; Bilyk, J.R.; Shore, J.W. Comparison of fibrovascular ingrowth into hydroxyapatite and porous polyethylene orbital implants. *Ophthalmic Plastic Reconstr. Surg.* **1994**, *10*, 96–110.
3. Dutton, J.J. Coralline hydroxyapatite as an ocular implant. *Ophthalmology* **1991**, *98*, 370–377.
4. Bilyk, J.R.; Rubin, P.A.D.; Shore, J.W. Correction of enophthalmos with porous polyethylene implants. *Int. Ophthalmol. Clin.* **1992**, *32*, 151–156.
5. Remulla, H.D.; Rubin, P.A.D.; Shore, J.W.; Sutula, F.C.; Townsend, D.J.; Woog, J.J.; Jahrling, K.V. Complications of porous spherical orbital implants. *Ophthalmology* **1995**, *102*, 586–593.
6. Flowers, C.W.; McDonnell, P.J. Mechanical methods in refractive corneal surgery. *Curr. Opin. Ophthalmol.* **1994**, *5*, 81–89.



7. Werblin, T.P.; Peiffer, R.L.; Binder, P.S.; McCarey, B.E.; Patel, A.S. Eight years experience with permalens intracorneal lenses in nonhuman primates. *Refract. Corneal Surg.* **1992**, *8*, 12–22.
8. Lane, S.S.; Lindstrom, R.L. Polysulfone intracorneal lenses. *Int. Ophthalmol. Clin.* **1991**, *31*, 38–46.
9. Hoffman, F.; Kruse, H.; Schuler, A. Mechanical methods in refractive corneal surgery. *Curr. Opin. Ophthalmol.* **1993**, *4*, 84–90.
10. Lehrner, M.S.; Sun, T.T.; Lavker, R.M. Strategies of epithelial repair: Modulation of stem cell and transit amplifying cell proliferation. *J. Cell. Sci.* **1998**, *111*, 2867–2875.
11. Catroveijo, R.; Cardona, H.; DeVoe, A.G. Present state of prosthokeratoplasty. *Am. J. Ophthalmol.* **1969**, *68*, 613–625.
12. Cardona, H. Mushroom transcorneal keratoprosthesis. *Am. J. Ophthalmol.* **1969**, *68*, 604–612.
13. Girard, L.J. Keratoprosthesis. *Cornea* **1983**, *2*, 207–224.
14. Doane, M.G.; Dohlman, C.H.; Bearnse, G. Fabrication of a keratoprosthesis. *Cornea* **1996**, *15*, 179–184.
15. Barber, J.C.; Feaster, F.T.; Prior, D.J. The acceptance of a vitreous carbon alloplastic material, Proplast, into the rabbit eye. *Investig. Ophthalmol. Vis. Sci.* **1980**, *19*, 182–190.
16. Legeais, J.M.; Rossi, C.; Renard, G.; Savoldelli, M.; D'hermies, F.; Pouliquen, Y. A new fluorocarbon for keratoprosthesis. *Cornea* **1992**, *11*, 538–545.
17. Chirila, T.V. Artificial cornea with a porous polymeric skirt. *Trends Polym. Sci.* **1998**, *5*, 346–348.
18. Chirila, T.V. An overview of the development of artificial corneas with porous skirts and the use of PHEMA for such an application. *Biomaterials* **2001**, *22*, 3311–3317.
19. Strampelli, B. Osteo-odonto-keratoprosthesis. *Ann. Ophthalmol. Clin. Ocul.* **1963**, *89*, 1039–1044.
20. Schepens, C.L.; Acosta, F. Scleral implants: An historical perspective. *Surv. Ophthalmol.* **1991**, *35*, 447–453.
21. Hahn, Y.S.; Lincoff, A.; Lincoff, H.; Kreissig, I. Infection after sponge implantation for scleral bucking. *Am. J. Ophthalmol.* **1979**, *87*, 180–185.
22. Molteno, A.C.B. New implant for drainage in glaucoma. Animal trial. *Br. J. Ophthalmol.* **1969**, *53*, 161–168.
23. Lim, K.S.; Allan, B.D.S.; Lloyd, A.W.; Muir, A.; Khaw, P.T. Glaucoma filtration implants: Past, present and future. *Br. J. Ophthalmol.* **1998**, *82*, 1083–1089.
24. Lim, K.S.; Allan, B.; Khaw, P.T.; Willis, S.; Lloyd, A.W.; Muir, A.; Gard, P.; Faragher, R.G.A.; Olliff, C.J.; Hanlon, G.W.; Wong, L.; Reed, S.; Denyer, S. Experimental flow studies in glaucoma drainage device development. *Br. J. Ophthalmol.* **2001**, *85*, 1231–1236.
25. Obstbaum, A.S. Biologic relationship between poly-(methyl methacrylate) intraocular lenses and uveal tissue. *J. Cataract Refract. Surg.* **1992**, *18*, 219–231.
26. Lloyd, A.W.; Dropcova, S.; Faragher, R.G.A.; Gard, P.R.; Hanlon, G.W.; Mikhailovsky, S.V.; Olliff, C.J.; Denyer, S.P. The development of in vitro biocompatibility tests for the evaluation of intraocular biomaterials. *J. Mater. Sci., Mater. Med.* **1999**, *10*, 621–627.
27. Lloyd, A.W.; Bowers, R.W.J.; Dropcova, S.; Denyer, S.P.; Faragher, R.G.A.; Gard, P.R.; Hall, B.; Hanlon, G.W.; Jones, S.A.; Muir, A.; Olliff, C.J.; Rosen, P.H.; Riding, M. In vitro evaluation of novel biomimetic intraocular lens materials. *Investig. Ophthalmol. Vis. Sci.* **1997**, *38*, S178.
28. Findl, O.; Kiss, B.; Petternel, V.; Menapace, R.; Georgopoulos, M.; Rainer, G.; Drexler, W. Intraocular lens movement caused by ciliary muscle contraction. *J. Cataract Refract. Surg.* **2003**, *29*, 669–676.
29. Legeais, J.M.; Werner, L.; Werner, L.; Abenham, A.; Renard, G. Pseudoaccommodation: BioComFold versus a foldable silicone intraocular lens. *J. Cataract Refract. Surg.* **1999**, *25*, 262–267.
30. Rosen, E.S. Intraocular lenses. *Curr. Opin. Ophthalmol.* **1993**, *4*, 44–53.
31. Ravi, N.; Chuck, R.S. Development of an injectable intraocular lens for preservation of accommodation. *Investig. Ophthalmol. Vis. Sci.* **1996**, *37*, 3539.
32. Capone, A., Jr.; Aaberg, T.M. Silicone oil in vitreoretinal surgery. *Curr. Opin. Ophthalmol.* **1995**, *6*, 33–37.
33. Arroyo, M.H.; Refojo, M.F.; Ariaz, J.J.; Tolentino, F.I.; Cajita, V.N.; Elner, V.M. Silicone oil as a delivery vehicle for BCNU in rabbit proliferative vitreoretinopathy. *Retina* **1993**, *13*, 245–250.
34. Liesegang, T.J. Viscoelastic substances in ophthalmology. *Surv. Ophthalmol.* **1990**, *34*, 268–293.
35. Goa, K.L.; Benfield, P. Hyaluronic acid—A review of its pharmacology and use as a surgical aid in ophthalmology, and its therapeutic potential in joint disease and wound-healing. *Drugs* **1994**, *47*, 536–566.
36. Ding, S. Recent developments in ophthalmic drug delivery. *Pharm. Sci. Technol. Today* **1998**, *8*, 328–335.

Ocular Implants for Drug Delivery

Susan S. Lee

National Eye Institute (NEI), National Institutes of Health (NIH), Bethesda, and Biomedical Engineering Department, Johns Hopkins University, Baltimore, Maryland, U.S.A.

Peng Yuan

Pharmacy Department, National Institutes of Health Clinical Center, Bethesda, Maryland, U.S.A.

Michael R. Robinson

National Eye Institute (NEI), National Institutes of Health (NIH), Bethesda, Maryland, U.S.A.

INTRODUCTION

Advances in materials science, biomedical engineering, and surgical techniques have encouraged the development of a variety of intraocular therapeutic devices. One class of intraocular devices includes implants for drug delivery, designed to improve the ability to deliver therapeutic drug levels into the eye. Most sight-threatening diseases occur in the back of the eye and frequently involve the retinal tissue. Diabetic retinopathy and age-related macular degeneration (AMD) are the most common causes of blindness in the United States. Local drug delivery with implants can overcome the many barriers that inhibit effective treatment of retinal diseases with traditional mechanisms of drug delivery, such as with eye drops or medications given systemically (i.e., by oral or intravenous route). The modified mucosal surface that covers the cornea is a significant barrier to topically applied eye drop medications (Fig. 1). The tight junctions between the basal epithelial cells of the cornea and the tear fluid drainage causes less than 5% of topically applied drug to permeate the cornea and reach intraocular tissues.^[1,2] As a result, drugs delivered with eye drop formulations are generally limited to treating tissues in the front of the eye. Examples include topical antimicrobials used to treat bacterial conjunctivitis or ocular antihypertensive drops that function by reducing the aqueous humor production at the level of the ciliary body (located in the front of the eye) to treat glaucoma. An attempt to deliver drugs to the posterior structures of the eye with systemic medications is impeded by the blood-retinal-barrier, which consists of the tight junctional complexes of the retinal pigment epithelium and retinal capillaries.^[1] This barrier reduces the penetration of most drugs given systemically to the posterior segment of the eye and hinders the treatment of many retinal diseases. Although high doses of systemically administered

medications may be able to penetrate the blood-retinal-barrier and provide therapeutic drug levels to the back of the eye, serious systemic side effects may occur.^[3] Delivery of drugs directly into the eye by intravitreal injection through the pars plana is another approach to achieve therapeutic levels of drug to treat retinal diseases (Fig. 2).^[4-7] However, given the short half life of most drugs injected into the vitreous, frequent injections 1–3 times per week are required to maintain therapeutic drug levels. As a result, intravitreal drug injections are very useful in treating acute bacterial infections of the eye (i.e., endophthalmitis)^[8] that may require one or two injections, but are not well tolerated by patients that may require months or years of injections to treat chronic eye diseases such as AMD or cytomegalovirus (CMV) retinitis (Fig. 3). Furthermore, frequent injections increase the risk of retinal detachment, vitreous hemorrhage, endophthalmitis, and cataracts.^[9-11]

Intraocular implants for drug delivery are capable of bypassing the barriers to intraocular drug absorption and avoid frequent procedures required using intravitreal injection therapy. Drug release rates of implants can be controlled so that therapeutic drug levels are maintained in the eye, while avoiding toxic or subtherapeutic levels. Moreover, higher intraocular drug levels can be achieved using implants compared with systemic or topical administration, making the drug more effective for treating a variety of retinal diseases.^[12] Since the dose of drug released by intraocular implants is in the range of 0.5 to 10 micrograms per day, these small doses significantly reduce the risk of systemic drug toxicity. Finally, ocular implants are more convenient for patients and reduce patient compliance issues with taking frequent eye drops or systemic medications.^[13] This article reviews the development of intraocular drug delivery implants and the potential for treating eye diseases.

O

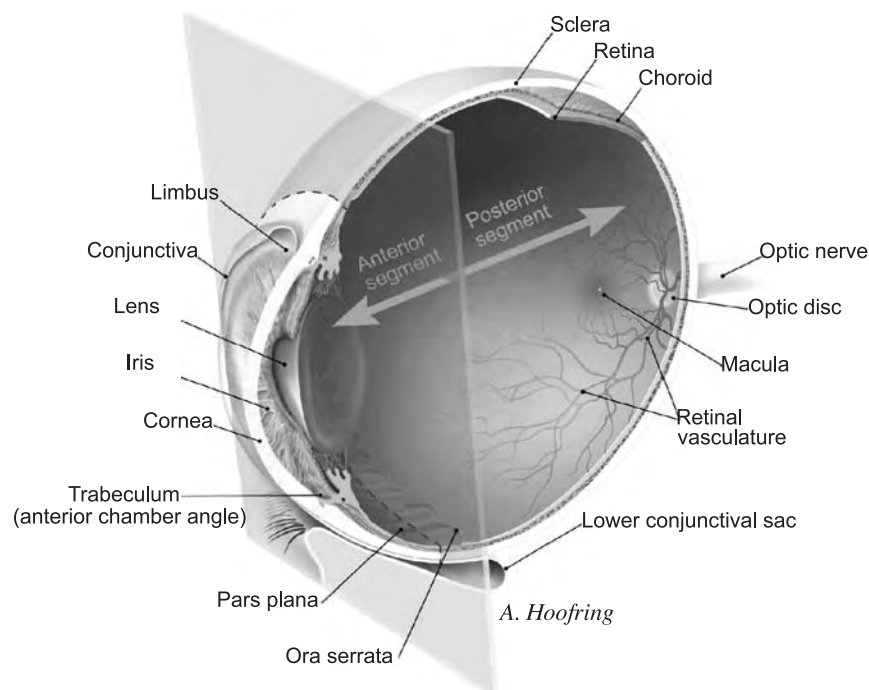


Fig. 1 Anatomy of the eye.

DRUG DELIVERY IMPLANTS: BASIC CONCEPTS

Ocular implants are engineered for optimal therapeutic efficacy. However, ocular implants should be biocompatible and properly sterilized to minimize the potential for an inflammatory reaction in the eye. In addition, implants should release drug at predictable and constant release rates to achieve therapeutic levels of drug to the target tissue with minimal toxicity to adjacent tissues. It is necessary for implants to be comfortable, preferably visible for inspection in the eye, while avoiding blockage of the visual axis. Implants should be easily inserted and removed, if not bioerodible^[14,15] and must be cost-effective when compared with systemic administration of the same drug. Drug stability in the implant is important and needs to be evaluated early during the implant development. From an ophthalmologist and patient's viewpoint, when considering drug delivery with an implant, the severity of the disease being treated should warrant the risks of implant surgery and should also have the potential for greater efficacy and less toxicity compared with the equivalent drug given by systemic administration.

Drug delivery implants can be designed to deliver drug to different parts of the eye (Fig. 2). The choice of implant design and location in the eye depends on the disease the ophthalmologist is treating. Since most of the current implants available clinically are diffusion-based, the closer the implant is located to the target tissue, the more drug will be delivered to that site. For example, an implant designed for the subconjunctival

space will deliver higher drug concentrations to best treat diseases that affect the conjunctiva or sclera (Fig. 2C). Implants designed to directly deliver drug into the vitreous cavity (Fig. 2B,D,E) are used to treat retinal diseases. The decision to use one implant design over another may also be influenced by the surgical risks of implanting a particular device. For example, implants that require an incision leading to the vitreous cavity significantly increase the risk of vitreous hemorrhage, infection, and retinal detachment. Implants inserted in the subconjunctival space are associated with few complications, especially if they are biodegradable and do not require removal. However, implants placed in the subconjunctival space may have some difficulty in treating retinal diseases because the sclera and choroid may pose a formidable obstacle to drug diffusion. Some subconjunctival implants using drugs with high scleral permeability have a reasonable potential to deliver drugs into the eye to treat retinal diseases. Passive solute diffusion through an aqueous pathway is the primary mechanism of drug permeation across the sclera.^[2] In-vitro and in-vivo flux studies have shown sclera to be quite permeable to drugs, including monoclonal antibodies, with a wide range of molecular weights up to 150 kilo Daltons.^[16-18] Trans-scleral permeability is inversely related to the drug's molecular weight; therefore, small molecule drugs are beneficial when delivering drugs from the subconjunctival space into the eye.

The solubility characteristics of a particular drug are also a major factor in determining the ability of a drug to be delivered by a particular device. For example,

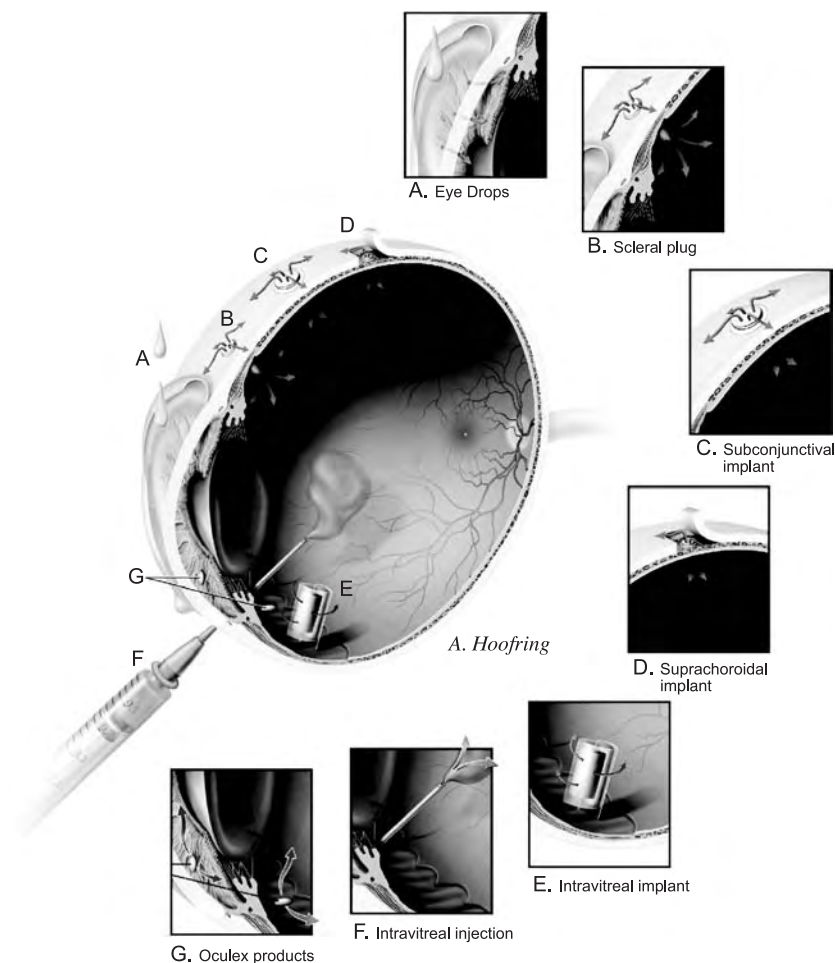


Fig. 2 Principal methods of local drug delivery to the eye.

aminoglycoside antibiotics, used in the past for gram negative bacteria endophthalmitis, are highly water-soluble and when delivered into the vitreous, have difficulty penetrating the relatively lipophilic retina leading to significant toxicity including tissue infarction and necrosis.^[10,19] In contrast, highly lipophilic drugs, like cyclosporine and corticosteroids, are cleared from the eye via a trans-retinal route and are more appropriate to use when treating diseases of the retina and choroid.^[20–22] The two main types of diffusion-based implants used in ocular drug delivery, matrix and reservoir systems, will be discussed separately.

MATRIX IMPLANTS: OVERVIEW

The matrix implant design for ocular implants typically consists of uniformly distributed drug in a nonreactive, bioerodible polymer.^[7,23] The most common polymer used is poly(lactic-glycolic acid), or PLGA, which is a copolymer of poly(glycolic acid) and poly(lactic acid). PLGA hydrolyzes into organic monomers of lactic acid and glycolic acid and is the same material used for manufacturing absorbable sutures. During the

manufacturing process of matrix implants, some of the drug dissolves in the polymeric solution, but the majority remains in solid phase (Fig. 4).^[23] The relative concentrations of the polymers, such as the lactide:glycolide ratio in PLGA, as well as polymer weight ratios can be altered to adjust the drug release rate.^[1,24,25] Furthermore, an additional polymer coating and holes in the implants can influence release rates.^[26] The polymer coating slows release rates, and larger hole diameters increase the surface area of release, which correlates with higher release rates. In addition, adjustments of drug release rate are dependent on drug properties. The use of insoluble drugs results in drug release that more closely relates to the dissolution rate of the matrix. Furthermore, the hydrophilicity of the polymer affects the release rate, as more controlled release is exhibited from hydrophobic polymers.^[27] Generally, the release of drug from bulk-eroding matrix implants follows first-order kinetics, i.e., the initial release is robust followed by a rapid decrease in release as the drug in the implant declines.^[7,28] Therefore, matrix style implants are most appropriate for eye conditions with acute onset, such as infectious endophthalmitis or postoperative inflammation, because these conditions are best treated with

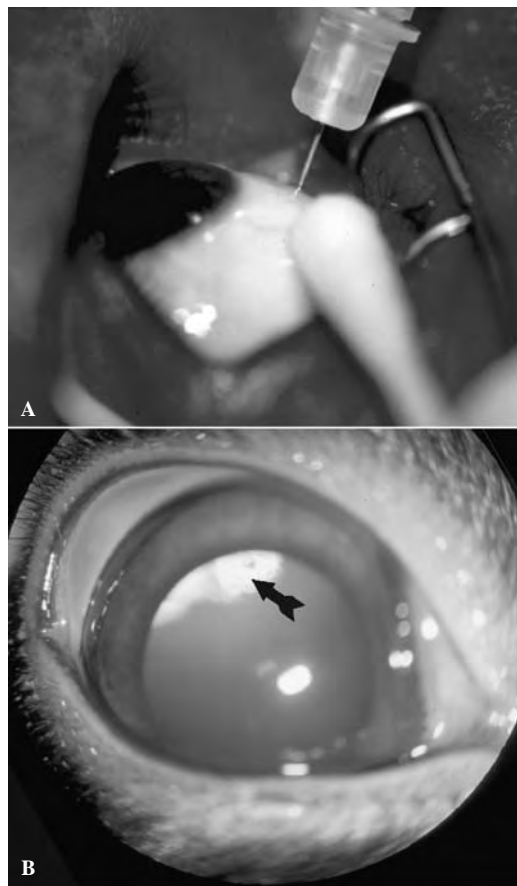


Fig. 3 (A) External photograph of the eye showing injection of an antiviral medication into the vitreous cavity to treat CMV retinitis in a patient with advanced HIV-infection. These injections are repeated twice per week. (B) External photograph of rabbit eye showing front of the eye with pupil widely dilated. A corticosteroid medication (arrow) had been injected a week before and positioned in the vitreous cavity behind the lens.

an initial loading dose of antimicrobials or corticosteroids followed by a gradual tapering of the medications over a defined period of time, generally ranging from one day to three months.

MATRIX IMPLANTS: OCULAR INSERTS

Examples of matrix implants that are placed on the outside surface of the eye are ocular inserts. The ocular inserts are bioerodible, solid matrix implants that are composed of hydrophilic polymers, including poly(vinyl alcohol) (PVA) and hydroxypropyl methylcellulose (HPMC), and drug. Inserts are placed beneath the lower eyelid (i.e., lower conjunctival sac), mix with the patient's tears, and dissolve within 12 to 24 hours. Drugs used in inserts include atropine, pilocarpine, and antibiotics and are used to apply drug to the surface of the eye to treat glaucoma or a variety of other eye diseases that affect the anterior segment of the eye. In addition, the polymer matrix alone can offer therapy to patients with dry eye. An example is the manufacturing of rod-shaped hydroxypropylcellulose (HPC) inserts that are placed in the lower conjunctival sac and undergo gradual dissolution, coating the surface of the eye with a tear/HPC lubricant continuously for approximately 24 hours.^[29] To further prolong the release of a drug from an ocular insert matrix system, use of nonbiodegradable polymers such as silicone can be employed.^[30] In addition, blending cellulose polymers with carbomer and creating the Bioadhesive Ophthalmic Drug Insert (BODI) can be done to increase release of drugs in the inferior conjunctival sac for >24 hours.^[31] Aside from delivering drugs to treat eye diseases, ocular inserts made of an absorbable gelatin foam have been employed in the inferior conjunctival sac to deliver insulin systemically to treat diabetes mellitus.^[14]

MATRIX IMPLANTS: SCLERAL PLUGS

Scleral plugs are an example of a matrix implant made for intraocular insertion and delivery of drug into the vitreous cavity (Fig. 2B). Scleral plugs are biodegradable, release drug, and are placed through a sclerotomy site without suture material at the completion of a pars

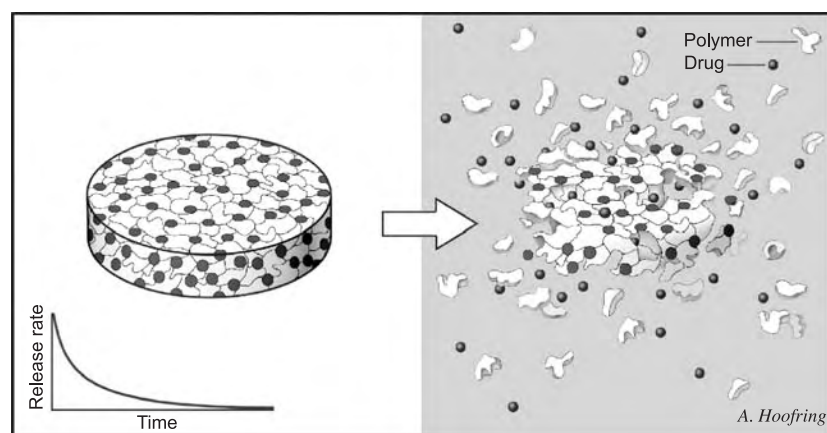


Fig. 4 Illustration showing a matrix implant (left) with drug (red spheres) dispersed in the polymer. When the implant is placed in an aqueous environment (right), the polymeric structure dissolves, releasing the drug. The release rate (lower-left inset) follows first-order kinetics, with an initial burst of drug followed by a rapid reduction in release.

plana vitrectomy procedure or as a separate procedure.^[32–35] These implants are composed of a matrix of PLGA and drug, and are constructed by a compression-molding technique to 1 mm in diameter. In addition, scleral plugs are heated to 80°C, thus all the drugs used must be thermally stable. Rabbit models showed that the drug diffusion into the vitreous does not impair the clarity of the ocular media. The scleral plugs may cause a mild foreign body reaction at one month after implantation, and the retina, lens, and other tissues remain normal for the entire duration of the implant. At four months, the devices are completely degraded and the sclerotomy site is replaced by loose connective tissue.^[32–34,36–38] Drug is released by diffusion from the plug during hydrolytic degradation, where the device swells and then gradually decreases in size. The drug diffusion from the polymer matrix is generally of first-order kinetics and triphasic, with an initial burst, a diffusive phase, and a second burst.^[34] Scleral plugs with drugs of smaller molecular weights release drug faster than plugs with drugs of larger molecular weights.^[36] Some controllable factors of the release profile of drugs are composition of polymers, ratio of polymers, molecular weight, shape of plug, and amount of drug.^[33,38] In addition to the release rate of the drug from the device, the diffusion rate of the drug in the vitreous and the elimination rate of the drug from the vitreous influence the concentration of drug in the vitreous.^[36] Altering the molecular weight ratio and type of polymers has been demonstrated to change the release rate of drug by producing biphasic scleral plugs. Biphasic implants produce only an initial burst and diffusive phase, and not the undesirable final burst, which may result in ocular toxicity. These pseudo-zero-order kinetics implants eliminate the adverse burst in the last phase of release. Scleral plugs contain poly (lactic acid) at two different molecular weights, PLA-70000 and PLA-5000 at a blending ratio of 80/20. These implants maintained an effective level of drug for six months without a significant burst or significant retinal toxicity.^[25,34] The biphasic release of the blended implant was largely due to ratio of molecular weights.^[25,32] Since PLA-5000 degraded faster due to its smaller molecular weight, water channels were formed connecting the surface to the inside of the implant. The drug present inside the implant was released constantly by diffusion throughout the water channels. Because water could move freely through the implant, degradation of the implant proceeded homogeneously, avoiding a final burst of drug inherent to matrix implants without blended PLGA designs.

A number of drugs have been tested in scleral plugs. PLGA scleral plugs with 25% ganciclovir have proven that ganciclovir release lasts longer in the vitreous with the implant compared with intravitreal injections.^[34,38] Ganciclovir is an antiviral medication used for the

treatment of CMV retinitis in patients with advanced HIV-infection. The scleral plug implants can release antiproliferative drugs such as doxorubicin^[32,33] that may be useful for treating proliferative vitreoretinopathy (PVR), a disease that can occur following surgery to repair retinal detachments. PVR results from a proliferation of fibroblasts in the eye that leads to scar formation on the retina and retinal detachment can recur. It is desirable to have approximately one to two months of an antiproliferative drug in the vitreous cavity to help reduce the incidence of postoperative PVR. Scleral plugs with doxorubicin maintained a therapeutic level for one month, as 25% of the drug in the plug was released.^[32] To treat fungal infections inside the eye, scleral plugs with fluconazole, a potent antifungal medication, demonstrated a maximum concentration in the vitreous to be 2.3 times higher than after intravenous injection of fluconazole, and maintained an effective level in the rabbit vitreous for three weeks.^[37]

MATRIX IMPLANTS: OCULEX DESIGNS

A biodegradable matrix implant using PLGA and dexamethasone is currently being investigated by Oculex Pharmaceuticals, Inc. (Sunnyvale, CA). The implants are rod-shaped and approximately 2 mm in length and diameter. They can be placed in the anterior chamber of the eye to treat inflammation in the front of the eye, and also in the posterior segment, to treat retinal diseases (Fig. 2G). Currently, Oculex has completed a phase II clinical trial evaluating the company's lead product, Posurdex, for the treatment of persistent macular edema related to diabetic retinopathy, retinal vein occlusion, and uveitis. The Posurdex implant releases 0.7 mg of dexamethasone over a four to six-week period and can be injected into the vitreous through the pars plana. The Oculex product, called Surodex, is placed in the anterior segment at the conclusion of cataract surgery to prevent postoperative inflammation and has been studied extensively in the United States and in Singapore where many of the clinical trials were conducted. The Surodex implant is approved for use in many countries including China where access to topical steroid drops for control of postoperative cataract surgery inflammation is limited. The Surodex implant is placed in the anterior chamber following the completion of the cataract surgery procedure and dexamethasone is released over 7–10 days.^[39,40] The 60 microgram of dexamethasone in one Surodex implant is approximately the same amount of drug in one drop of topical 0.1% dexamethasone eye drops. When compared to topical steroid eye drops in the setting of a clinical trial, the Surodex implant was shown to be at least as effective as topical steroids in reducing anterior chamber inflammation during the first month after surgery. Patient studies showed that

Surodex does not migrate in the ocular tissues and does not need to be secured using suture materials.^[40] No significant adverse complications resulted from usage of Surodex, and there was no occurrence of rebound inflammation after depletion of drug in the implant.^[39–41]

Others have investigated biodegradable implants similar to the Oculex design to deliver corticosteroids in the anterior chamber. Morita et al.^[42] have created a rod-shaped implant releasing the corticosteroid, fluorometholone. The biocompatible polymers, poly(lactic acid) and poly(vinyl pyrrolidone), were used to create a first-order pharmacokinetic release for 40 days and the release rates could be adjusted by changing the poly(vinyl pyrrolidone) content. To date, the fluorometholone implant has only been described for use in animal models.

MATRIX IMPLANTS: MISCELLANEOUS DESIGNS

Rubsamen et al.^[43] have studied a cylindrical matrix implant of PLGA and 1 mg of fluorouracil inserted at the vitreous base for the treatment of PVR. Fluorouracil, a common chemotherapy agent used in cancer trials, is a potent inhibitor of fibroblast proliferation. The fluorouracil implant swelled slightly, and then slowly degraded, with residual matrix left at 28 days. The implant prevented tractional retinal detachment in experimental PVR models, without introducing any toxicity to the adjacent tissues. The mean fluorouracil concentrations were maintained between 2.2 and 6.7 microgram/ml in the vitreous cavity for 14 days and above 0.3 microgram/ml for 3 weeks. There was minimal variation in performance among individual implants. Sustained release fluorouracil from a polycaprolactone (PCL) matrix has also been reported to be effective in preventing tractional retinal detachment in experimental PVR, without significant complications.^[44]

Investigators have employed matrix systems to deliver antifibrotic medications to improve the success of glaucoma filtering surgery. In glaucoma surgery, a communication is surgically made to redirect aqueous humor flow to the subconjunctival space near the limbus to lower the intraocular pressure. The surgical channel can eventually scar down with fibrosis occurring in the subconjunctival space and the intraocular pressure increases. Morales et al.^[45] investigated the antiproliferative agent daunomycin in a matrix system that delivered drug to the subconjunctival space to improve the outcome of glaucoma-filtering surgery in experimental models. In the Morales system, daunomycin was conjugated to hyaluronic acid by an acid labile bond. Hyaluronic acid, a high molecular-weight carrier for daunomycin, was chosen to restrict the movement of the conjugated drug from the surgical

site. It is a naturally occurring substance in the anterior segment with affinity for extracellular matrix. Daunomycin was slowly released from the hyaluronic acid with an in-vitro half life of 51 hours, and in-vivo release extending beyond 24 hours. The pharmacologically effective concentrations of daunomycin were restricted to the surgical site, making this device attractive for the delivery of daunomycin to improve the long-term success of filtering procedures. Other matrix systems that have been employed in experimental glaucoma-filtering models use polyanhydrides as the matrix polymer. Matrix disks of a polyanhydride 1,3-bis(p-carboxyphenoxy) propane and sebacic acid, with the drug etoposide, were placed in the subconjunctival region and delivered a sufficient concentration to reduce fibroblast proliferation for at least 12 days following glaucoma-filtration surgery.^[46] When considering a drug delivery system to improve the success rate of glaucoma-filtering surgery, one has to weigh the potential risks of having potent antiproliferative drugs in the eye with the potential benefits since some of these drugs have been associated with significant local toxicities.^[47]

Another novel matrix system releases daunomycin^[48,49] into the eye and was developed for use in experimental models of PVR. Poly capillary fiber (PCF) is a nonerodible, biocompatible, anisotropic ultrafiltration fiber with several desirable features for slow-release drug delivery purposes. Properties that make PCF suited for prolonged delivery include the deep macrovoids in the outer membranes that create more surface area for drug diffusion. This slow release of daunomycin reduces the potential for local toxicity without reducing its effectiveness in inhibiting fibroblast proliferation.

Applications of matrix style implants to the eye also extend to immunosuppressants, such as cyclosporine, as well as various types of polymer materials other than PLGA. A subconjunctival, PLGA matrix implant containing cyclosporine was designed for sustained immunosuppressant therapy to the anterior segment following corneal transplant therapy to prevent graft rejection.^[50] The PLGA disk is 5 mm in diameter and 0.5 mm thick, with a 1:1 polymer to drug ratio. In-vivo experiments were successful in preventing corneal rejection in animal models.

Presoaked matrices, including soft contact lenses and gels, are also simple sustained-release devices. Drug release from these devices is very rapid in the beginning and proceeds at a rate that declines exponentially with time. The basic concept of the presoaked matrix is to extend the time that the dissolved drug will continue to be effective compared with an eye drop. Soft contact lenses soaked in 2% pilocarpine can reduce intraocular pressure for almost 24 hours.^[51] Matrix implants have also been extended to oligodeoxynucleotide (ODN) therapy for the treatment of virus-related retinitis disease.^[52] These rod-shaped implants made from PLGA

release ODN following pseudo zero-order kinetics for more than 20 days in phosphate buffer solution.

Finally, injectable matrices are also effective in short periods of drug release. Fibrin sealant was found to provide a more controlled and localized release of carboplatin to treat retinoblastoma from a subconjunctival injection.^[53] The fibrin sealant, FDA-approved for facilitating blood clotting, appeared to be well-tolerated and has the potential to extend the release of multiple drugs in the subconjunctival space. Other injectable matrices that have been studied include a semisolid poly(ortho ester) that delivered fluorouracil in an experimental glaucoma-filtering model to prevent conjunctival fibrosis.^[54] The poly(ortho ester), a bio-erodible hydrophobic polymer, undergoes surface erosion confined to the polymer-water interface and can improve the duration of drug release compared with more soluble polymers like fibrin sealant.^[15]

RESERVOIR IMPLANTS: OVERVIEW

Reservoir implants deliver continuous amounts of drug over months to years, which is desirable for treatment of chronic eye diseases. These devices typically have a central core of drug, coated with a nonreactive polymer.^[23] Although polymers used for reservoir implants are generally nonerodible, there are exceptions such as a bio-erodible reservoir ciprofloxacin implant for the treatment of intraocular bacterial infections.^[7] Non-erodible silicone, ethylene vinyl acetate (EVA), and poly(vinyl alcohol) (PVA) are commonly used in reservoir devices. The primary polymers in the first extraocular implant (placed in the inferior conjunctival sac) used for drug delivery were EVA and PVA, developed by Alza Corporation (Mountain View, California)^[55,56] and marketed as the Ocusert[®] implant. The Ocusert released pilocarpine, a medication used in the treatment of glaucoma, over a seven-day period and was available in two different doses. The Ocusert improved patient compliance since pilocarpine eye drops were required

every four hours to maintain adequate intraocular pressure control.

The application of different polymers to reservoir implants is determined primarily by the permeability of a given drug through the polymer. Silicone and PVA are freely permeable to a number of drugs while providing enough support to hold the drug reservoir together. Silicone, used for the past three decades in the manufacture of a five-year releasing contraceptive implant (Norplant[®], Wyeth Pharmaceuticals, Collegeville, Penn.), has been very attractive for use in eye implants because it is biocompatible, inexpensive, easy to handle, and has a long track record of safety in the medical field. PVA, a commonly used polymer in commercial eye drops, can be used in high concentrations (5 to 50%) to coat drug pellets. PVA can be heated up to 200°C to cross-link the polymer, rendering it less soluble to drugs. The duration and temperature of the heat treatment can be varied to control the diffusion of drugs through the PVA. EVA is impermeable to most medications and used as a coating around the drug pellet to provide more controlled diffusion of the drug from the implant. The drug delivery depends on the water that diffuses through the polymer coating and dissolves part of the pellet, forming a saturated drug solution (Fig. 5). The drug then diffuses back out of the device and as long as the solution inside the device is saturated, the release rate is constant.^[57] This mechanism allows the device to deliver a small burst of drug, followed by a steady-state release rate that persists until greater than 90% of the drug pellet has diffused through the polymer. The movement of drug through the polymer in reservoir implants is governed by Fick's Law of Diffusion, where the release rate is directly proportional to both the surface area of the implant, the diffusivity (i.e., ability of a particular drug to transit through the polymer), and indirectly proportional to the thickness of the polymeric coating.^[23] The most important determinant of the longevity of a reservoir implant is the solubility of the drug in the pellet in the surrounding media. Typically, relatively insoluble drugs will give

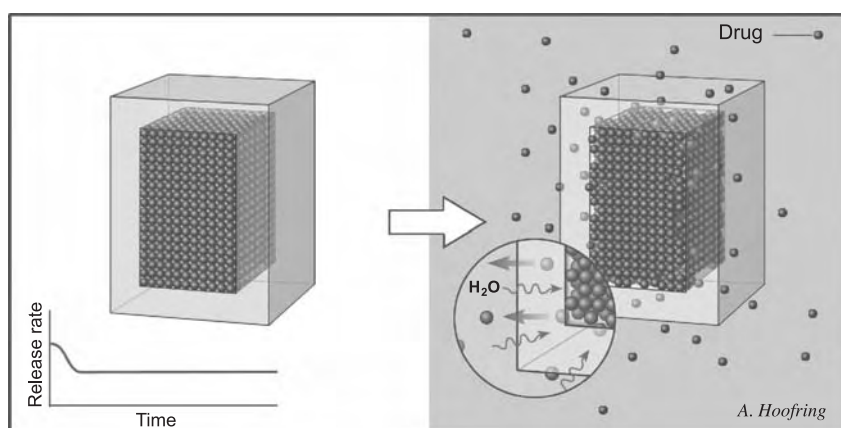


Fig. 5 Illustration showing reservoir implant (left) with compressed drug pellet (red spheres) surrounded by a polymer. When the implant is placed in an aqueous environment (right), water diffuses through the polymer wall and dissolves the periphery of the drug pellet. The solubilized drug diffuses out of the implant through the permeable polymer. The release rate (lower-left inset) shows an initial small burst of drug followed by sustained release following zero-order kinetics.

reservoir implants the longest release rates lasting from six months to five years. Chronic eye diseases such as AMD and diabetic retinopathy are optimally treated with a constant level of drug release, or zero-order kinetics, delivered using reservoir implants.

RESERVOIR IMPLANTS: VITRASERT (GANCICLOVIR IMPLANT)

The advent of the acquired immunodeficiency syndrome in the 1980s was a challenging time for ophthalmologists. Patients with profound immunodeficiency developed opportunistic infections including a blinding condition known as CMV retinitis. Although there were no effective treatments for CMV disease, pharmaceutical companies began investigating specific anti-CMV medications such as ganciclovir. Ganciclovir was effective when administered intravenously; however, ocular penetration was poor, leading to the development of resistant CMV isolates and rapid progression of the retinitis.^[58] Further problems with intravenous ganciclovir included systemic complications including neutropenia, anemia, and renal failure. Since a central intravenous catheter was required to administer the required daily infusions, complications from the indwelling catheter in these fragile patients included sepsis and death.^[59] Investigators started investigating the use of a ganciclovir implant that would be inserted into the vitreous cavity and attached to the inside of the sclera at the pars plana using suture material (Fig. 2E). The pars plana is an ideal area in the eye to access the vitreous cavity since it is anterior to the insertion of the retina and posterior to the lens. As a result, implants inserted through the pars plana will not damage the lens, causing a cataract, and will not damage the retina, which may lead to a retinal detachment. The ganciclovir implant was the first intraocular drug delivery implant available for humans and had a profound impact on treatment of CMV retinitis.^[60] Clinical trials using the implant were started in 1990 and approval from the FDA occurred in 1996. Marketed under the trade name of Vitrasert[®] (Bausch and Lomb, Rochester, NY), the implant was constructed by coating a compressed ganciclovir pellet with PVA, which is permeable to the drug, and then partially coated with EVA, which is impermeable to the drug, to reduce the release rate (Fig. 6). The pellet complex was attached to a PVA suture stub, heated to partially cross-link the PVA and adjust the in-vitro release rate to 1.4 microgram/hr. In vivo, the intravitreal drug levels were 4.1 microgram/ml, considerably higher than those levels achieved following intravenous therapy (0.93 microgram/ml).^[61,62] The device was capable of delivering drug for approximately eight months. Clinical studies showed that the median time to progression of retinitis

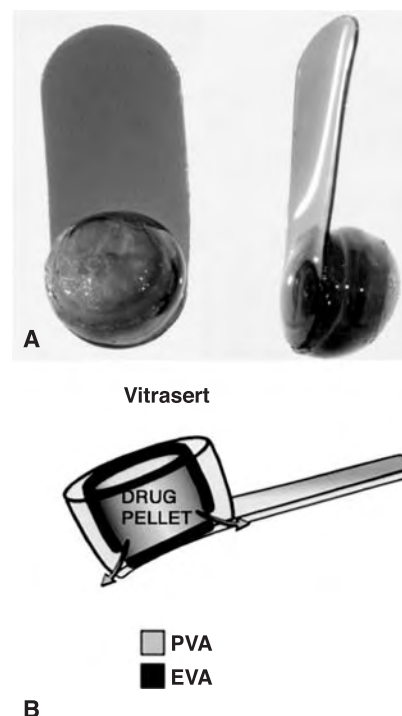


Fig. 6 (A) Vitrasert implant showing front and side views. The drug pellet contains ganciclovir, a potent antiviral agent to treat CMV retinitis. The long suture stub is trimmed and used to fixate the implant to the sclera at the pars plana, allowing the drug reservoir to release drug into the vitreous cavity. (B) Schematic drawing of Vitrasert implant showing the drug pellet surrounded by two polymers, PVA and EVA. Ganciclovir is freely permeable to PVA; however, to reduce the release rate, EVA (not permeable to ganciclovir) coats approximately 90% of the pellet and drug is released (arrows) at the base of the implant.

in untreated patients (controls) was 15 days, compared with 226 days for patients assigned to immediate treatment with the implant. This compared favorably with reported times to progression in patients treated with intravenous ganciclovir, which was 71 days.^[12,63] Vitrasert implant has emerged as the most efficacious method of treating CMV retinitis and the patient acceptance is high, given the alternatives of frequent biweekly intravitreal injections or daily drug infusions (Fig. 7). Because the device eventually runs out of drug and must be replaced, appropriate implant exchange practices have been adopted. The optimal time for exchanging the ganciclovir implant is dictated by the predictability of the duration of release of each device, the location of retinitis, and the overall health of the patient. The duration of release of each implant is dependent on the original quantity of drug, the rate of drug release from the device, and the rate of tissue clearance of the drug.^[64] The prototype ganciclovir

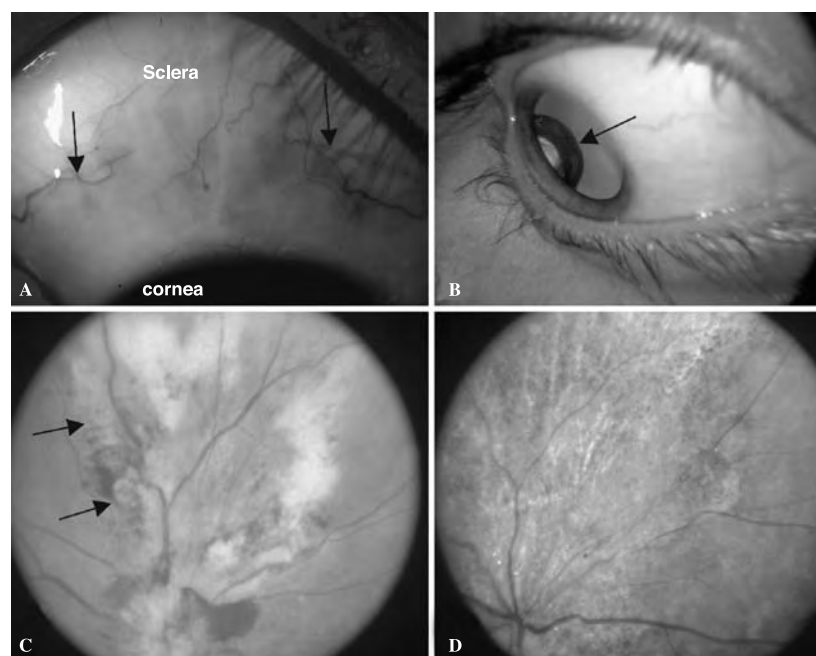


Fig. 7 (A) External photograph of the eye showing the sclera above and cornea below. This patient has two Vitrasert implants placed in the eye. The fixating suture material (arrows), using nonresorbing prolene material, is visible in the subconjunctival area. (B) External photograph of right eye showing the patient looking down to the right. The pupil is widely dilated and the top edge of a Vitrasert implant (arrow) is visible inside the vitreous cavity. (C) Photograph of the back of the eye of a patient with CMV retinitis. The retina is visible with areas of hemorrhage and necrosis (arrows) at the location of the CMV lesion. (D) Photograph of the CMV lesion in Fig. 7C following the insertion of a Vitrasert implant. Note the areas of hemorrhage and necrosis have resolved, and a large retinal scar is present in the area of previous infection.

implants had drug reservoirs of 5 mg and in-vitro release rates of either 1 microgram/hr or 2 microgram/hr. Therefore, expected duration of release from each prototype was 208 days and 104 days, respectively, assuming the in-vivo release rates were similar to the in vitro. Multiple reentries into the vitreous cavity through the same wound may be associated with an increased risk for vitreous hemorrhage and retinal detachment; thus, an implantation site in the eye is used no more than twice. There are acceptable ocular complications that occur in patients receiving the Vitrasert implant including a temporary decrease in visual acuity lasting a few weeks after surgery.^[63] Also, the intraocular surgery entails risks such as endophthalmitis, vitreous hemorrhage, retinal detachment, extrusion of the strut, and improper placement of the implant in the suprachoroidal space that occur in <5% of cases.^[12,60,65] Endophthalmitis can possibly be prevented by using a modified surgical technique with a scleral flap covering the incision.^[66] Patients treated with implants also may be at higher risk for developing extraocular CMV disease; thus, the combination of implantation and systemic administration of ganciclovir using the oral prodrug formulation is most effective in combating CMV infection.

RESERVOIR IMPLANTS: CORTICOSTEROID IMPLANTS

Corticosteroids are important in the management of ocular inflammatory disease that involves inflammation and/or fibrovascular proliferation such as PVR and uveitis. Experiments with dexamethasone showed that an intraocular injection of dexamethasone can successfully inhibit fibrous proliferation in the vitreous cavity, so dexamethasone was one of the first corticosteroids to be placed in a sustained release reservoir device.^[67] The dexamethasone implant contained a 5 mg drug pellet, and released at 1.5 microgram/hr in the vitreous, maintaining a vitreous concentration of approximately 2–3 microgram/ml.^[68] The implant was well tolerated, with no signs of ocular toxicity; however, since the implant would last for only five months, corticosteroids with less solubility were studied in reservoir implants. A reservoir implant releasing fluocinolone, capable of delivering drug for three years in the vitreous cavity, rapidly moved from preclinical testing^[69] to clinical human trials (Fig. 8A and B). A pilot study using fluocinolone acetonide reservoir implants in patients with refractory uveitis proved to be effective in reducing ocular inflammation and significantly improving

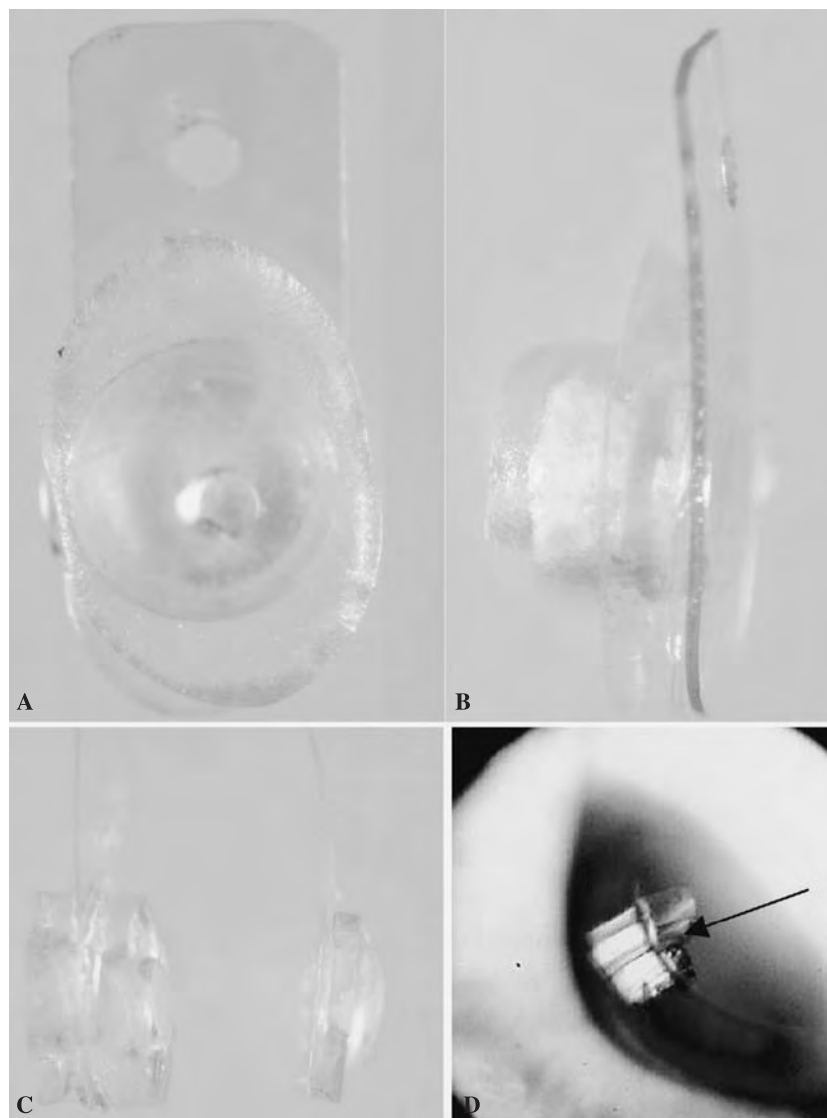


Fig. 8 (A & B) A fluocinolone implant showing front and side views. (C) Cyclosporine implant showing side view of a double-pelleted implant (left) and a single-pelleted implant (right). The implant with two pellets will deliver approximately twice the release rate of a single-pellet implant. Two implant doses are standard for clinical trials examining the efficacy of an implant for treating an eye disease. (D) External photograph of right eye showing the patient looking down to the right. The pupil is widely dilated and a double-pelleted implant (arrow) is visible inside the vitreous cavity.

vision.^[22] Complications included glaucoma and some patients required topical antiocular hypertension drops to lower the eye pressure. Larger multicenter trials using the fluocinolone implant are in progress internationally for uveitis, diabetic macular edema, and choroidal neovascularization associated with AMD.

RESERVOIR IMPLANTS: CYCLOSPORINE IMPLANTS

Oral cyclosporine, a potent immunosuppressant used to prevent allograft rejection, is an effective treatment for chronic uveitis.^[70] However, side effects including renal insufficiency and hypertension often limit its long-term use. Cyclosporine reservoir implants have been designed to release drug into the eye for a number of years to treat chronic uveitis.^[21] The implants were constructed using a core of 5 mg cyclosporine,

measuring 2.5 mm in diameter, coated with medical grade silicone, cured at 104°C for one hour, then attached to a PVA suture tab for placement in the vitreous cavity through the pars plana.^[20] Experimental studies have shown that devices releasing cyclosporine at a rate of 1.3 microgram/day can achieve intravitreal levels over a six month period of 500 ng/ml, which is five times the therapeutic level needed to suppress T-cell activation. These implants can theoretically release for three–five years, which is appropriate for the disease that can be active for decades. In both equine^[71] and rabbit^[21] models of uveitis, the cyclosporine implant showed reduced inflammation compared to the control eyes. A pilot study using the CsA implant for refractory cases of uveitis in humans is currently underway at the National Eye Institute in Bethesda, MD (Fig. 8). Reservoir implants for both intravitreal and subconjunctival (Fig. 9) placement have also been used to deliver angiostatic agents including 2-methoxyestradiol

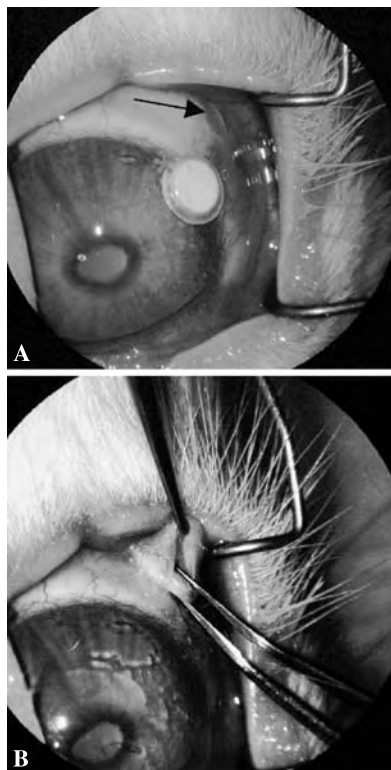


Fig. 9 (A) Reservoir implant (red asterisk) containing investigational drug to treat retinal disease will be placed through previous incision made through conjunctiva (arrow) of rabbit eye. (B) Implant is placed through conjunctival incision and will be sutured closed to prevent implant extrusion.

and thalidomide in models choroidal neovascularization associated with age-related macular degeneration.^[23,28,72] Long-term safety studies on animals are mandated by the FDA and are a prerequisite to human trials to minimize the risk to the patient (Fig. 10).

RESERVOIR IMPLANTS: CODRUG IMPLANTS

The combination of the anti-inflammatory effects of corticosteroids and antiproliferative properties of an antimetabolite, such as fluorouracil, is the basis of codrug reservoir implants that can target several processes of a disease process such as PVR. Delivering multiple drugs simultaneously requires the conjugation of one of more drugs with a covalent carbonate bond linkage and compressing the codrugs into a pellet to be placed in a reservoir device. The drug conjugation makes the codrug complex less soluble, increasing the half life of each drug moiety.^[67] Codrug release into solution occurs in two steps. First is the codrug dissolution from the pellet, which is determined by the solubility of the compound, and is the principal rate-limiting step. The second step is

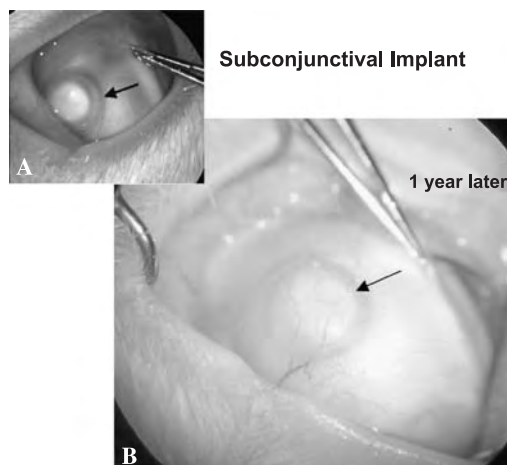


Fig. 10 (A) Reservoir implant (arrow) in subconjunctival space of rabbit eye three months after implantation. Tissues surrounding implant are intact and there are no signs of toxicity. (B) Same reservoir implant (arrow) seen in Fig. 10A after one year showing reduced drug reservoir and no long-term toxicity visible on the surface of the eye.

codrug hydrolysis, which occurs rapidly, and releases the drug moieties into equimolar quantities. It had been demonstrated that sustained release of triamcinolone and fluorouracil codrug was effective in the inhibition of a rabbit model of PVR without toxic effects evident on clinical or histopathologic examination.^[73] Codrug reservoirs that have been developed also include a CsA-dexamethasone implant to treat ocular inflammation,^[74] and a dexamethasone-fluorouracil codrug implant that was effective in a model of PVR.^[67] Other approaches have been described to release multiple drugs simultaneously without the use of codrug technology. Zhou et al. reported a multiple drug delivery implant prototype that had been designed and prepared to test a local controlled-release strategy and multiple-drug regimen approach for intravitreal treatment of PVR.^[75] The device was composed of PLGA, in three contiguous cylindrical segments, each containing 5-fluorouridine, triamcinolone, and human recombinant tissue plasminogen activator (a thrombolytic agent). The first two drugs were used as antiproliferative agents to inhibit fibroblast proliferation and the third drug dissolves hemorrhage that may have occurred in the vitreous during the surgery.

RESERVOIR IMPLANTS: LIVE CELL IMPLANTS

Reservoir implants for ocular delivery have generally used small molecular weight drugs that are stable under the curing conditions of the polymer and the biologic effects of the drug are not altered during implant

construction. Large molecular weight compounds such as cytokines, proteins, and monoclonal antibodies have not been successfully incorporated in reservoir implants for the eye. To overcome these obstacles, Neurotech USA (Lincoln, Rhode Island) developed a novel approach to release proteins into the vitreous using encapsulated cell therapy implants. The implants contain cells that are genetically engineered to secrete therapeutic factors for long-term delivery into the vitreous. Neurotech constructed an implant to release ciliary neurotrophic factor (CNTF), a protein that may be a potential treatment for neurodegenerative diseases of the eye such as retinitis pigmentosa. The titanium implants (10 mm length and 1.0 mm diameter) are loaded with mammalian cells that are genetically engineered to secrete CNTF. Genetically modified cells are packaged in a hollow-fiber, semipermeable membrane, mostly of polyether sulfone. The hollow-fiber prevents immune molecules and host immune cells from entering the device, but allows nutrients and therapeutic molecules to diffuse freely across the membrane. The encapsulated cells are in a matrix material and derive nourishment from the host milieu. The device is anchored to the sclera at the pars plana by a small titanium wire loop. In one study, CNTF was delivered through encapsulated cells directly into the vitreous of the eye in a canine model of retinitis pigmentosa.^[76] The CNTF implant had a protective effect as demonstrated by a higher photoreceptor density in eyes treated with CNTF compared with controls. No adverse effects were observed on the retina in the treated eyes. Protection of photoreceptors appeared to be dose-dependent, with minimal protection at CNTF doses of 0.2–1.0 ng/day and greater protection at higher doses. The implants were well tolerated and the cells remained viable for the seven-week implantation interval. Clinical studies in humans are currently in development to use the CNTF implant in patients with advanced retinitis pigmentosa.

RESERVOIR IMPLANTS: OSMOTIC PUMP SYSTEMS

Osmotic pumps have been investigated for the last three decades at Alza Corporation (Mountain View, California) principally for systemic drug administration.^[77] Alza Corp. gained FDA approval in 2000 for the Viadur[®] implant that releases leuprolide to treat patients with advanced prostate cancer. Small-size osmotic systems are also an effective drug delivery mechanism for eye disease.^[1,27,78] However, these pumps are too large to be placed in the eye, but attempts at fixating the drug reservoir under the skin near the eye or on the external eyelid, with a tube leading from the reservoir into eye tissue, have been performed in experimental models. In a model of endophthalmitis, an osmotic

pump was placed subcutaneously, which infused gentamicin sulfate into the vitreous over a 4.5-day period, releasing at approximately 0.1 mg/hr, and proved to be effective in treating the infection. Ambati and colleagues^[18] showed a similar pump could deliver antivascular endothelial growth factor antibodies through the sclera into the choroid and subretinal space to potentially treat choroidal neovascularization in AMD.

CONCLUSION

Drug delivery implants have had a profound impact in clinical ophthalmology, especially in the management of retinal diseases that can lead to blindness. Ocular implants have been engineered to avoid the barriers that often impede the delivery of drug via traditional methods. Matrix implants, with drug release generally lasting less than three months, are being investigated to treat acute eye diseases, while reservoir implants, with drug release lasting up to five years, are used for chronic eye diseases. The investigation and clinical use of ocular drug implants requires a multidisciplinary approach including biomedical, materials, and chemical engineers interacting with ophthalmologists. This team approach will facilitate the development of ocular implants for drug release, thus providing more treatment options to reduce the severity of sight-threatening diseases.

ARTICLES OF FURTHER INTEREST

Contact Lenses: Silicone Hydrogels; Controlled Release; Corneal Implants; Drug Delivery, Controlled; Ocular Implants

REFERENCES

1. Metrikina, D.C.; Anand, R. Intravitreal drug administration with depot devices. *Curr. Opin. Ophthalmol.* **1994**, *5* (3), 21–29.
2. Geroski, D.H.; Edelhauser, H.F. Drug delivery for posterior segment eye disease. *Invest. Ophthalmol. Vis. Sci.* **2000**, *41* (5), 961–964.
3. Fraunfelder, F.T. National registry of drug-induced ocular side effects. *Am. J. Ophthalmol.* **1990**, *110* (4), 426–427.
4. Mochizuki, K.; Torisaki, M.; Kawasaki, K.; Shirao, Y.; Yamashita, Y.; Kitano, K.; Yonemura, D. Retinal toxicity of antibiotics: Evaluation by electroretinogram. *Doc. Ophthalmol.* **1988**, *69* (2), 195–202.
5. Peyman, G.A.; Schulman, J.A. Intravitreal drug therapy. *Jpn. J. Ophthalmol.* **1989**, *33* (4), 392–404.
6. Young, S.H.; Morlet, N.; Heery, S.; Hollows, F.C.; Coroneo, M.T. High dose intravitreal ganciclovir in the

- treatment of cytomegalovirus retinitis. *Med. J. Aust.* **1992**, *157* (6), 370–373.
7. Hainsworth, D.P.; Conklin, J.D.; Bierly, J.R.; Ax, D.; Ashton, P. Intravitreal delivery of ciprofloxacin. *J. Ocular Pharmacol. Ther.* **1996**, *12* (2), 183–191.
 8. Mamalis, N.; Kearsley, L.; Brinton, E. Postoperative endophthalmitis. *Curr. Opin. Ophthalmol.* **2002**, *13* (1), 14–18.
 9. Heinemann, M.H. *Staphylococcus epidermidis* endophthalmitis complicating intravitreal antiviral therapy of cytomegalovirus retinitis. Case report. *Arch. Ophthalmol.* **1989**, *107* (5), 643–644.
 10. Brown, G.C.; Eagle, R.C.; Shakin, E.P.; Gruber, M.; Arbizio, V.V. Retinal toxicity of intravitreal gentamicin. *Arch. Ophthalmol.* **1990**, *108* (12), 1740–1744.
 11. Seawright, A.A.; Bourke, R.D.; Cooling, R.J. Macula toxicity after intravitreal amikacin. *Aust. N. Z. J. Ophthalmol.* **1996**, *24* (2), 143–146.
 12. Musch, D.C.; Martin, D.F.; Gordon, J.F.; Davis, M.D.; Kuppermann, B.D. Treatment of cytomegalovirus retinitis with a sustained-release ganciclovir implant. The Ganciclovir Implant Study Group. *N. Engl. J. Med.* **1997**, *337* (2), 83–90.
 13. Taylor, S.A.; Galbraith, S.M.; Mills, R.P. Causes of non-compliance with drug regimens in glaucoma patients: A qualitative study. *J. Ocular Pharmacol. Ther.* **2002**, *18* (5), 401–409.
 14. Lee, Y.C.; Simamora, P.; Yalkowsky, S.H. Effect of Brij-78 on systemic delivery of insulin from an ocular device. *J. Pharm. Sci.* **1997**, *86* (4), 430–433.
 15. Bernatchez, S.F.; Merkli, A.; Minh, T.L.; Tabatabay, C.; Anderson, J.M.; Gurny, R. Biocompatibility of a new semisolid bioerodible poly(ortho ester) intended for the ocular delivery of 5-fluorouracil. *J. Biomed. Mater. Res.* **1994**, *28* (9), 1037–1046.
 16. Olsen, T.W.; Edelhauer, H.F.; Lim, J.I.; Geroski, D.H. Human scleral permeability. Effects of age, cryotherapy, transscleral diode laser, and surgical thinning. *Invest. Ophthalmol. Vis. Sci.* **1995**, *36* (9), 1893–1903.
 17. Ambati, J.; Canakis, C.S.; Miller, J.W.; Gragoudas, E.S.; Edwards, A.; Weissgold, D.J.; Kim, I.; Delori, F.C.; Adamis, A.P. Diffusion of high molecular weight compounds through sclera. *Invest. Ophthalmol. Vis. Sci.* **2000**, *41* (5), 1181–1185.
 18. Ambati, J.; Gragoudas, E.S.; Miller, J.W.; You, T.T.; Miyamoto, K.; Delori, F.C.; Adamis, A.P. Transscleral delivery of bioactive protein to the choroid and retina. *Invest. Ophthalmol. Vis. Sci.* **2000**, *41* (5), 1186–1191.
 19. Verma, L.; Arora, R.; Sachdev, M.S. Macular infarction after intravitreal injection of amikacin. *Can. J. Ophthalmol.* **1993**, *28* (5), 241–243.
 20. Pearson, P.A.; Jaffe, G.J.; Martin, D.F.; Cordahi, G.J.; Grossniklaus, H.; Schmeisser, E.T.; Ashton, P. Evaluation of a delivery system providing long-term release of cyclosporine. *Arch. Ophthalmol.* **1996**, *114* (3), 311–317.
 21. Jaffe, G.J.; Yang, C.S.; Wang, X.C.; Cousins, S.W.; Gallemore, R.P.; Ashton, P. Intravitreal sustained-release cyclosporine in the treatment of experimental uveitis. *Ophthalmology* **1998**, *105* (1), 46–56.
 22. Jaffe, G.J.; Ben-Nun, J.; Guo, H.; Dunn, J.P.; Ashton, P. Fluocinolone acetonide sustained drug delivery device to treat severe uveitis. *Ophthalmology* **2000**, *107* (11), 2024–2033.
 23. Ross, M.L.; Yuan, P.; Robinson, M.R. Intraocular drug delivery implants. *Rev. Ophthalmol.* **2000**, *7*, 95–99.
 24. Morita, Y.; Ohtori, A.; Kimura, M.; Tojo, K. Intravitreal delivery of dexamethasone sodium m-sulfobenzoate from poly(DL-lactic acid) implants. *Biol. Pharm. Bull.* **1998**, *21* (2), 188–190.
 25. Kunou, N.; Ogura, Y.; Yasukawa, T.; Kimura, H.; Miyamoto, H.; Honda, Y.; Ikada, Y. Long-term sustained release of ganciclovir from biodegradable scleral implant for the treatment of cytomegalovirus retinitis. *J. Control. Release* **2000**, *68* (2), 263–271.
 26. Wang, G.; Tucker, I.G.; Roberts, M.S.; Hirst, L.W. In vitro and in vivo evaluation in rabbits of a controlled release 5-fluorouracil subconjunctival implant based on poly(D,L-lactide-co-glycolide). *Pharm. Res.* **1996**, *13* (7), 1059–1064.
 27. Shell, J.W. Ophthalmic drug delivery systems. *Surv. Ophthalmol.* **1984**, *29* (2), 117–128.
 28. Robinson, M.R.; Baffi, J.; Yuan, P.; Sung, C.; Byrnes, G.; Cox, T.A.; Csaky, K.G. Safety and pharmacokinetics of intravitreal 2-methoxyestradiol implants in normal rabbit and pharmacodynamics in a rat model of choroidal neovascularization. *Exp. Eye Res.* **2002**, *74* (2), 309–317.
 29. Cordonnier, M.; Pereleux, A.; Herode, A.; Zanen, A. The treatment of dry eye with Lacrisert. *Bull. Soc. Belge Ophthalmol.* **1984**, *212*, 65–69.
 30. Chetoni, P.; Di Colo, G.; Grandi, M.; Morelli, M.; Saettone, M.F.; Darougar, S. Silicone rubber/hydrogel composite ophthalmic inserts: Preparation and preliminary in vitro/in vivo evaluation. *Eur. J. Pharm. Biopharm.* **1998**, *46* (1), 125–132.
 31. Baeyens, V.; Kaltsatos, V.; Boisrame, B.; Fathi, M.; Gurny, R. Evaluation of soluble Bioadhesive Ophthalmic Drug Inserts (BODI) for prolonged release of gentamicin: Lachrymal pharmacokinetics and ocular tolerance. *J. Ocular Pharmacol. Ther.* **1998**, *14* (3), 263–272.
 32. Hashizoe, M.; Ogura, Y.; Kimura, H.; Moritera, T.; Honda, Y.; Kyo, M.; Hyon, S.H.; Ikada, Y. Scleral plug of biodegradable polymers for controlled drug release in the vitreous. *Arch. Ophthalmol.* **1994**, *112* (10), 1380–1384.
 33. Hashizoe, M.; Ogura, Y.; Takanashi, T.; Kunou, N.; Honda, Y.; Ikada, Y. Implantable biodegradable polymeric device in the treatment of experimental proliferative vitreoretinopathy. *Curr. Eye Res.* **1995**, *14* (6), 473–477.
 34. Yasukawa, T.; Kimura, H.; Kunou, N.; Miyamoto, H.; Honda, Y.; Ogura, Y.; Ikada, Y. Biodegradable scleral implant for intravitreal controlled release of ganciclovir. *Graefes Arch. Clin. Exp. Ophthalmol.* **2000**, *238* (2), 186–190.
 35. Kurz, D.; Ciulla, T.A. Novel approaches for retinal drug delivery. *Ophthalmol. Clin. North Am.* **2002**, *15* (3), 405–410.
 36. Kimura, H.; Ogura, Y.; Hashizoe, M.; Nishiwaki, H.; Honda, Y.; Ikada, Y. A new vitreal drug delivery system using an implantable biodegradable polymeric device. *Invest. Ophthalmol. Vis. Sci.* **1994**, *35* (6), 2815–2819.
 37. Miyamoto, H.; Ogura, Y.; Hashizoe, M.; Kunou, N.; Honda, Y.; Ikada, Y. Biodegradable scleral implant for



- intravitreal controlled release of fluconazole. *Curr. Eye Res.* **1997**, *16* (9), 930–935.
38. Hashizoe, M.; Ogura, Y.; Takanashi, T.; Kunou, N.; Honda, Y.; Ikada, Y. Biodegradable polymeric device for sustained intravitreal release of ganciclovir in rabbits. *Curr. Eye Res.* **1997**, *16* (7), 633–639.
 39. Chang, D.F.; Wong, V. Two clinical trials of an intraocular steroid delivery system for cataract surgery. *Trans. Am. Ophthalmol. Soc.* **1999**, *97*, 261–274. discussion 274–269.
 40. Tan, D.T.; Chee, S.P.; Lim, L.; Lim, A.S. Randomized clinical trial of a new dexamethasone delivery system (Surodex) for treatment of post-cataract surgery inflammation. *Ophthalmology* **1999**, *106* (2), 223–231.
 41. Tan, D.T.; Chee, S.P.; Lim, L.; Theng, J.; Van Ede, M. Randomized clinical trial of Surodex steroid drug delivery system for cataract surgery: Anterior versus posterior placement of two Surodex in the eye. *Ophthalmology* **2001**, *108* (12), 2172–2181.
 42. Morita, Y.; Saino, H.; Tojo, K. Polymer blend implant for ocular delivery of fluorometholone. *Biol. Pharm. Bull.* **1998**, *21* (1), 72–75.
 43. Rubsamén, P.E.; Davis, P.A.; Hernandez, E.; O'Grady, G.E.; Cousins, S.W. Prevention of experimental proliferative vitreoretinopathy with a biodegradable intravitreal implant for the sustained release of fluorouracil. *Arch. Ophthalmol.* **1994**, *112* (3), 407–413.
 44. Borhani, H.; Peyman, G.A.; Rahimy, M.H.; Thompson, H. Suppression of experimental proliferative vitreoretinopathy by sustained intraocular delivery of 5-FU. *Int. Ophthalmol.* **1995**, *19* (1), 43–49.
 45. Morales, J.; Kelleher, P.J.; Campbell, D.; Crosson, C.E. Effects of daunomycin implants on filtering surgery outcomes in rabbits. *Curr. Eye Res.* **1998**, *17* (8), 844–850.
 46. Uppal, P.; Jampel, H.D.; Quigley, H.A.; Leong, K.W. Pharmacokinetics of etoposide delivery by a bioerodible drug carrier implanted at glaucoma surgery. *J. Ocul. Pharmacol.* **1994**, *10* (2), 471–479.
 47. Hardten, D.R.; Samuelson, T.W. Ocular toxicity of mitomycin-C. *Int. Ophthalmol. Clin.* **1999**, *39* (2), 79–90.
 48. Rahimy, M.H.; Peyman, G.A.; Fernandes, M.L.; el-Sayed, S.H.; Luo, Q.; Borhani, H. Effects of an intravitreal daunomycin implant on experimental proliferative vitreoretinopathy: Simultaneous pharmacokinetic and pharmacodynamic evaluations. *J. Ocul. Pharmacol.* **1994**, *10* (3), 561–570.
 49. Rahimy, M.H.; Peyman, G.A.; Chin, S.Y.; Golshani, R.; Aras, C.; Borhani, H.; Thompson, H. Polysulfone capillary fiber for intraocular drug delivery: In vitro and in vivo evaluations. *J. Drug Target.* **1994**, *2* (4), 289–298.
 50. Apel, A.; Oh, C.; Chiu, R.; Saville, B.; Cheng, Y.L.; Rootman, D. A subconjunctival degradable implant for cyclosporine delivery in corneal transplant therapy. *Curr. Eye Res.* **1995**, *14* (8), 659–667.
 51. Hillman, J.S.; Marsters, J.B.; Broad, A. Pilocarpine delivery by hydrophilic lens in the management of acute glaucoma. *Trans. Ophthalmol. Soc. U. K.* **1975**, *95* (1), 79–84.
 52. Yamakawa, I.; Ishida, M.; Kato, T.; Ando, H.; Asakawa, N. Release behavior of poly(lactic acid-co-glycolic acid) implants containing phosphorothioate oligodeoxynucleotide. *Biol. Pharm. Bull.* **1997**, *20* (4), 455–459.
 53. Simpson, A.E.; Gilbert, J.A.; Rudnick, D.E.; Geroski, D.H.; Aaberg, T.M.; Edlhauser, H.F., Jr. Transscleral diffusion of carboplatin: An in vitro and in vivo study. *Arch. Ophthalmol.* **2002**, *120* (8), 1069–1074.
 54. Einmahl, S.; Zignani, M.; Varesio, E.; Heller, J.; Veuthey, J.L.; Tabatabay, C.; Gurny, R. Concomitant and controlled release of dexamethasone and 5-fluorouracil from poly(ortho ester). *Int. J. Pharm.* **1999**, *185* (2), 189–198.
 55. Quigley, H.A.; Pollack, I.P.; Harbin, T.S., Jr. Pilocarpine ocuserts. Long-term clinical trials and selected pharmacodynamics. *Arch. Ophthalmol.* **1975**, *93* (9), 771–775.
 56. Hitchings, R.A.; Smith, R.J. Experience with pilocarpine Ocuserts. *Trans. Ophthalmol. Soc. U. K.* **1977**, *97* (1), 202–205.
 57. Ashton, P.; Blandford, D.L.; Pearson, P.A.; Jaffe, G.J.; Martin, D.F.; Nussenblatt, R.B. Review: Implants. *J. Ocul. Pharmacol.* **1994**, *10* (4), 691–701.
 58. Peters, B.S.; Beck, E.J.; Anderson, S.; Coleman, D.; Coker, R.; Main, J.; Migdal, C.; Harris, J.R.; Pinching, A.J. Cytomegalovirus infection in AIDS. Patterns of disease, response to therapy and trends in survival. *J. Infect.* **1991**, *23* (2), 129–137.
 59. McGavin, J.K.; Goa, K.L. Ganciclovir: An update of its use in the prevention of cytomegalovirus infection and disease in transplant recipients. *Drugs* **2001**, *61* (8), 1153–1183.
 60. Anand, R.; Nightingale, S.D.; Fish, R.H.; Smith, T.J.; Ashton, P. Control of cytomegalovirus retinitis using sustained release of intraocular ganciclovir. *Arch. Ophthalmol.* **1993**, *111* (2), 223–227.
 61. Ashton, P.; Brown, J.D.; Pearson, P.A.; Blandford, D.L.; Smith, T.J.; Anand, R.; Nightingale, S.D.; Sanborn, G.E. Intravitreal ganciclovir pharmacokinetics in rabbits and man. *J. Ocul. Pharmacol.* **1992**, *8* (4), 343–347.
 62. Smith, T.J.; Pearson, P.A.; Blandford, D.L.; Brown, J.D.; Goins, K.A.; Hollins, J.L.; Schmeisser, E.T.; Glavinos, P.; Baldwin, L.B.; Ashton, P. Intravitreal sustained-release ganciclovir. *Arch. Ophthalmol.* **1992**, *110* (2), 255–258.
 63. Martin, D.F.; Parks, D.J.; Mellow, S.D.; Ferris, F.L.; Walton, R.C.; Remaley, N.A.; Chew, E.Y.; Ashton, P.; Davis, M.D.; Nussenblatt, R.B. Treatment of cytomegalovirus retinitis with an intraocular sustained-release ganciclovir implant. A randomized controlled clinical trial. *Arch. Ophthalmol.* **1994**, *112* (12), 1531–1539.
 64. Martin, D.F.; Ferris, F.L.; Parks, D.J.; Walton, R.C.; Mellow, S.D.; Gibbs, D.; Remaley, N.A.; Ashton, P.; Davis, M.D.; Chan, C.C.; Nussenblatt, R.B. Ganciclovir implant exchange. Timing, surgical procedure, and complications. *Arch. Ophthalmol.* **1997**, *115* (11), 1389–1394.
 65. Charles, N.C.; Freisberg, L. Endophthalmitis associated with extrusion of a ganciclovir implant. *Am. J. Ophthalmol.* **2002**, *133* (2), 273–275.
 66. Guembel, H.O.; Krieglsteiner, S.; Rosenkranz, C.; Hattenbach, L.O.; Koch, F.H.; Ohrloff, C. Complications after implantation of intraocular devices in patients

- with cytomegalovirus retinitis. *Graefe Arch. Clin. Exp. Ophthalmol.* **1999**, *237* (10), 824–829.
67. Berger, A.S.; Cheng, C.K.; Pearson, P.A.; Ashton, P.; Crooks, P.A.; Cynkowski, T.; Cynkowska, G.; Jaffe, G.J. Intravitreal sustained release corticosteroid-5-fluorouracil conjugate in the treatment of experimental proliferative vitreoretinopathy. *Invest. Ophthalmol. Vis. Sci.* **1996**, *37* (11), 2318–2325.
68. Hainsworth, D.P.; Pearson, P.A.; Conklin, J.D.; Ashton, P. Sustained release intravitreal dexamethasone. *J. Ocular Pharmacol. Ther.* **1996**, *12* (1), 57–63.
69. Jaffe, G.J.; Yang, C.H.; Guo, H.; Denny, J.P.; Lima, C.; Ashton, P. Safety and pharmacokinetics of an intraocular fluocinolone acetonide sustained delivery device. *Invest. Ophthalmol. Vis. Sci.* **2000**, *41* (11), 3569–3575.
70. Masuda, K.; Nakajima, A.; Urayama, A.; Nakae, K.; Kogure, M.; Inaba, G. Double-masked trial of cyclosporin versus colchicine and long-term open study of cyclosporin in Behcet's disease. *Lancet* **1989**, *1* (8647), 1093–1096.
71. Gilger, B.C.; Wilkie, D.A.; Davidson, M.G.; Allen, J.B. Use of an intravitreal sustained-release cyclosporine delivery device for treatment of equine recurrent uveitis. *Am. J. Vet. Res.* **2001**, *62* (12), 1892–1896.
72. Velez, G.; Whitcup, S.M. New developments in sustained release drug delivery for the treatment of intraocular disease. *Br. J. Ophthalmol.* **1999**, *83* (11), 1225–1229.
73. Yang, C.S.; Khawly, J.A.; Hainsworth, D.P.; Chen, S.N.; Ashton, P.; Guo, H.; Jaffe, G.J. An intravitreal sustained-release triamcinolone and 5-fluorouracil codrug in the treatment of experimental proliferative vitreoretinopathy. *Arch. Ophthalmol.* **1998**, *116* (1), 69–77.
74. Enyedi, L.B.; Pearson, P.A.; Ashton, P.; Jaffe, G.J. An intravitreal device providing sustained release of cyclosporine and dexamethasone. *Curr. Eye Res.* **1996**, *15* (5), 549–557.
75. Zhou, T.; Lewis, H.; Foster, R.E.; Schwendeman, S.P. Development of a multiple-drug delivery implant for intraocular management of proliferative vitreoretinopathy. *J. Control Release.* **1998**, *55* (2–3), 281–295.
76. Tao, W.; Wen, R.; Goddard, M.B.; Sherman, S.D.; O'Rourke, P.J.; Stabila, P.F.; Bell, W.J.; Dean, B.J.; Kauper, K.A.; Budz, V.A.; Tsiaras, W.G.; Acland, G.M.; Pearce-Kelling, S.; Laties, A.M.; Aguirre, G.D. Encapsulated cell-based delivery of CNTF reduces photoreceptor degeneration in animal models of retinitis pigmentosa. *Invest. Ophthalmol. Vis. Sci.* **2002**, *43* (10), 3292–3298.
77. Zaffaroni, A. Overview and evolution of therapeutic systems. *Ann. N. Y. Acad. Sci.* **1991**, *618*, 405–421.
78. Michelson, J.B.; Nozik, R.A. Experimental endophthalmitis treated with an implantable osmotic minipump. *Arch. Ophthalmol.* **1979**, *97* (7), 1345–1346.



Optical Coherence Tomography

Xingde Li

Department of Bioengineering, University of Washington, Seattle, Washington, U.S.A.

James G. Fujimoto

Massachusetts Institute of Technology, Cambridge, Massachusetts, U.S.A.

INTRODUCTION

Optical coherence tomography (OCT) is an emerging biomedical imaging technology, which is capable of performing cross-sectional imaging of internal microstructures of biological tissues *in vivo* and in real time.^[1–4] OCT is analogous to ultrasound B-mode imaging except that OCT uses near infrared light rather than sound. The optical echo delay time (related to depth) and intensity of light backscattered from biological tissues are measured using a technique called low-coherence interferometry. The backscattered light intensity versus the echo delay time (or depth) and the transverse position forms a cross-sectional image that can be displayed on a gray or false-color scale. An axial resolution of 1–10 μm has been achieved using advanced low-coherence light sources. Different from other medical imaging modalities such as X-ray computed tomography (CT), positron emission tomography (PET), magnetic resonance imaging (MRI), and ultrasound, OCT image contrast arises from optical scattering and absorption properties of biological tissues. OCT imaging depth is governed by optical attenuation (dominated by optical scattering), and imaging depths of 1–2 mm can be achieved in most highly scattering tissues using a near infrared light source in the 800–1300 nm wavelength range. Therefore, OCT is ideal for imaging the skin, and the epithelial/endothelial linings of internal luminal organs. OCT is very attractive in biomedicine because imaging can be performed in a noninvasive, non-contact fashion with a resolution approaching that of histology, but without the need for tissue removal or processing. OCT is a fiber-optically-based technology, making the OCT system very compact and potentially cost-effective. Flexible fiber-optic OCT imaging probes permit imaging of internal organs and integration with many medical instruments. In addition to structural imaging, OCT can also perform high-resolution, high-sensitivity imaging of blood flow and tissue birefringence.^[5–9] Studies have demonstrated OCT can assess tissue structural morphologies and functional processes in a variety of organs and/or systems including the eye, skin, gastrointestinal tract, oral cavity,

and cardiovascular system etc.^[10–19] OCT is becoming a clinically viable imaging modality for early detection of diseases, precision guidance of surgical interventions, and repeated monitoring of treatment. In the following sections, we will discuss the operating principles technological aspects and advances of OCT and representative clinical and biomedical applications.

PRINCIPLE OF OCT OPERATION

OCT is analogous to ultrasound except that near infrared low-coherence light is utilized instead of acoustic waves. A schematic of a typical OCT system is shown in Fig. 1A. The OCT system is based on a fiber-optically implemented Michelson interferometer. The interferometer contains a fiber-optic coupler (e.g., a beam splitter/combiner), which splits the low-coherence light (e.g., light with a broad spectral bandwidth) into two parts. One part is delivered to the tissue to be imaged through one of the two interferometer arms called the sample arm; the other part is sent to the reference arm, where the optical pathlength is scanned. The light backscattered from the tissue will be recombined by the coupler with light from the reference arm with a known delay, and directed to a photo detector. The echo delay time and backscattered light intensity are detected by measuring the interference signal at the detector. The interference signal only occurs when the distances traveled by the light from the sample and reference arms match to within the coherence length of the light source (see Fig. 1B). The frequency of the interference signal (or the beat frequency) is determined by the Doppler frequency shift induced by optical pathlength scanning in the reference arm. OCT essentially performs heterodyne optical detection and therefore extremely high sensitivity exceeding 100 dB can be achieved, meaning OCT can detect the backscattered light from the tissue at a level of one part in 10 billion of the incident light intensity. The echo delay time and intensity of light backscattered at different depths in the tissue can be measured by scanning the optical

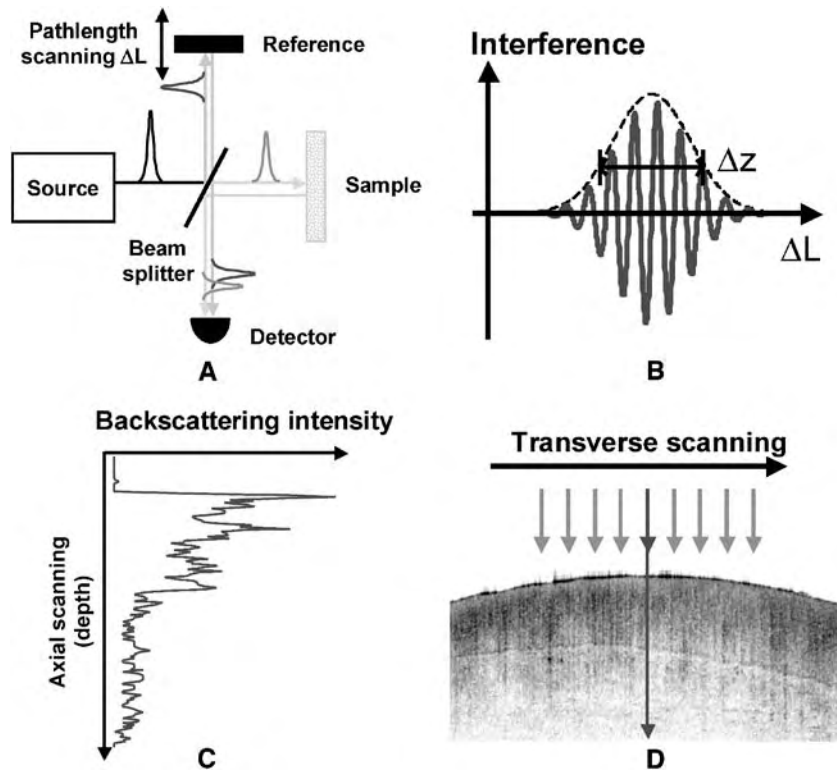


Fig. 1 (A) Schematic of a fiber-optic OCT system. (B) Localized interference occurs when the distances traveled by the light in the OCT reference and sample arms match to within the coherence length of the low-coherence light source (e.g., the two pulses returned from the reference and sample arms overlap at the detector). (C) Intensity versus the optical echo delay time (or depth) forms an axial scan. (D) A cross-sectional tomographic image is formed by multiple adjacent axial scans acquired at different transverse positions.

pathlength in the reference arm. The echo delay time is translated into imaging depth, and the backscattering intensity versus depth forms an axial scan (see Fig. 1C). A two- or three-dimensional data set consisting of multiple adjacent axial scans at different transverse positions can be acquired by laterally scanning the optical beam across the tissue (see Fig. 1D). The cross-sectional image can be displayed using a false-color or gray-scale in a Cartesian or polar coordinate system depending upon the application.

In order to analyze the OCT signal, we assume that the optical pathlength (or depth) scanning is performed in the reference arm by translating a mirror with a scanning velocity of v_s . For a given color of light with an optical frequency, f , the Doppler frequency shift in the reference arm is $f_d = (2v_s/c)f$, where c is the speed of light in a vacuum. The input light intensity and electrical field to the Michelson interferometer are $I_o(f)$ and $E_o(f)$, respectively. The E-field reflection coefficients (which are in general complex numbers) for the scanning mirror in the reference arm and the tissue in the sample arm are denoted by $r(f)$ and $s(f)$, respectively. For simplicity, we assume the 50/50 beam splitter has a wavelength-independent splitting ratio. At the detector, the E-field of the light returned from the reference arm is given by $E_r(f) = r(f)E_o(f)/2$, and the E-field of the light backscattered from the tissue is $E_s(f) = s(f)E_o(f)/2$. The interference signal measured at

the detector is thus

$$\begin{aligned}
 I_f(t) &= \langle |E_r(f)e^{i2\pi ft} \\
 &\quad + E_s(f)e^{i2\pi(f+f_d)t}|^2 \rangle_{\text{Time Average over } 1/f_d \gg t \gg 1/f} \\
 &= \frac{I_o(f)}{4} \left[|r(f)|^2 + |s(f)|^2 \right. \\
 &\quad \left. + 2|r(f)s(f)| \cos(2\pi f_d t + \phi_o) \right] \quad (1)
 \end{aligned}$$

where $t = z/v_s$ is the scanning time related to the scanning depth z and the scanning velocity v_s , and ϕ_o is the phase difference between the two beams returned from the reference and sample arms. In general, the reflectivity in the reference arm is flat and independent of the wavelength within the range of interest (e.g., when using a perfect retroreflecting mirror). Therefore, $|r(f)| = r$ will be a constant and its value represents the optical attenuation in the reference arm. In many cases, we are only interested in the AC component of the interference signal, $I_f^{AC}(t) = [I_o(f)/2]r|s(f)|\cos(2\pi f_d t + \phi_o)$, which is the beating term of two beams at the difference frequency. The constant reflectivity, r , of the reference arm functions as a heterodyne gain for the AC component. Hereinafter, without further specification, the interference signal means the AC component of the interference signal.

Since OCT uses a low-coherence light, the total interference signal will be an integral of Eq. 1 over the entire spectrum of the light source, and the OCT interference signal is then given by:

$$\begin{aligned} I_{\text{All}f}^{AC}(t) &= r \int_f \frac{I_o(f)}{2} |s(f)| \cos(2\pi f_d t + \phi_o) df \\ &= \frac{r}{2} \int_f I_o(f) |s(f)| \cos\left(\frac{4\pi v_s t}{c} f + \phi_o\right) df \end{aligned} \quad (2)$$

Assuming the phase difference, ϕ_o , is wavelength-independent, the previous equation can be rewritten as:

$$\begin{aligned} I_{\text{All}f}^{AC}\left(t' = \frac{2v_s t}{c}\right) &= \text{Re}\left\{\frac{r}{2} \int_f I_o(f) |s(f)| \exp[i(2\pi f t' + \phi_o)] df\right\} \\ &= \frac{r}{2} \text{Re}\{\exp(i\phi_o) FT[I_o(f) |s(f)|]\} \end{aligned} \quad (3)$$

Equation 3 indicates the OCT interference signal is related to the Fourier transform (FT) of the product of the source power spectrum $I_o(f)$ and the tissue reflectivity $|s(f)|$. For a majority of OCT signal processing, Eq. 3 is the starting point for analyzing the tomographic envelope, Doppler flow, and spatially resolved spectroscopy.

OCT TECHNOLOGY

In this section, we discuss some of the key technological aspects of OCT, which include imaging resolution, speed, beam delivery, and collection probes, and data processing for obtaining spectroscopic and functional information. As in any imaging technology, resolution is one of the most important parameters and therefore we consider this first.

OCT Imaging Resolution

Unlike confocal microscopy, the transverse and axial resolutions of OCT are determined independently by different physical mechanisms. The transverse resolution ΔX is governed by the focusing optics and determined by the transverse focused spot size $\Delta X = (2\lambda/\pi)NA$, where $NA = d/2f$, d is the beam spot size on the objective lens, and f is the focal length. The axial resolution ΔZ of OCT, however, is given by the coherence length of the light source, which can be derived from Eq. 3. To simplify the derivation, we assume the backscattered spectrum from the tissue does not change and the spectrum $|s(f)| = s$ is thus wavelength-independent. Under this condition, Eq. 3 indicates

the OCT interference signal is simply the Fourier transform of the power spectrum of the light source, which is the field temporal correlation function according to the Wiener-Khinchin theorem.^[20] Consider a low-coherence light source with a Gaussian spectrum, e.g.,

$$I_o(f) = I_o \frac{2\sqrt{\ln 2/\pi}}{\Delta f} \exp\left[-4 \ln 2 \left(\frac{f - f_o}{\Delta f}\right)^2\right] \quad (4)$$

where Δf is the full-width-at-the-half-maximum (FWHM) of the source power spectrum, f_o is the center frequency of the light source, and I_o is a constant (e.g., the total power of the light source over the entire spectrum). The Fourier transform of a Gaussian spectrum can be easily calculated and Eq. 3 then becomes:

$$\begin{aligned} I_{\text{OCT}}(t') &= \frac{rs}{2} \text{Re}\{\exp(i\phi_o) FT[I_o(f)]\} \\ &= \frac{rs}{2} \exp\left[-\frac{(\pi\Delta f t')^2}{4 \ln 2}\right] \cos(2\pi f_o t' + \phi_o) \end{aligned} \quad (5)$$

After substituting the variable t' with time t (e.g., $t' = 2v_s t/c$), we have the following time-dependent OCT interference signal:

$$I_{\text{OCT}}(t) = \frac{rs}{2} \exp\left[-\frac{(2\pi\Delta f v_s t/c)^2}{4 \ln 2}\right] \cos\left(\frac{4\pi v_s f_o t}{c} + \phi_o\right) \quad (6)$$

As can be seen from Eq. 6, the OCT interference signal is amplitude-modulated. The exponential part forms the modulation envelope, and the free-space axial resolution is then given by the FWHM of the envelope multiplied by the reference arm scanning velocity v_s , i.e.:

$$\Delta Z = \Delta t \cdot v_s = \frac{2 \ln 2 c}{\pi \Delta f} = \frac{2 \ln 2}{\pi} \frac{\lambda_o^2}{\Delta \lambda} \quad (7)$$

where λ_o is the center wavelength of the light source and $\Delta \lambda$ is the spectrum bandwidth. It should be mentioned that the axial resolution in tissue is the free-space axial resolution divided by the tissue index of refraction. From Eq. 7, we see that the OCT axial resolution is inversely proportional to the spectrum bandwidth of the light source. A light source with a broader spectrum bandwidth produces a finer resolution. White light, which has a very broad spectrum, would be a good source for achieving high axial resolution.^[21,22] However, conventional white light sources such as an arc lamp have a poor spatial coherence, and the coupling efficiency of white light into a single-mode optical fiber is extremely low, making it difficult to use a white light source in a fiber-optic OCT system. Therefore, an ideal light source for high-resolution fiber-optic OCT imaging should have a

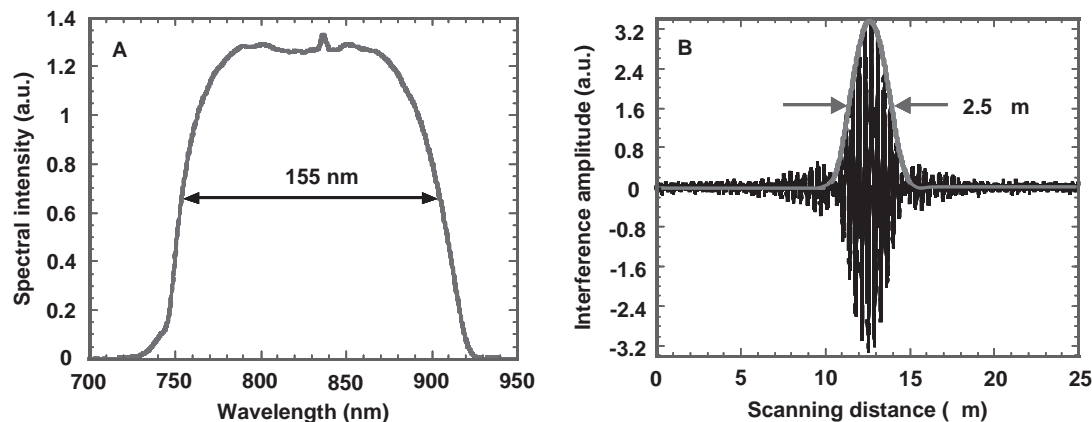


Fig. 2 (A) Spectrum of a short-pulse Ti:Sapphire laser (7–8 fs, transform limited, plus ~ 5 –10% spectrum broadening when launched into a single-mode fiber). (B) The OCT interference signal, of which the FWHM of the envelope determines the OCT axial resolution.

short temporal coherence length, and be spatially single mode. Other requirements for an OCT light source include emission in the near-infrared range (e.g., 800–1500 nm for deeper imaging penetration in highly scattering tissues), high irradiance power (e.g., 5–15 mW for a better signal-to-noise ratio), and a nearly Gaussian spectrum shape (e.g., for minimizing side lobes on the interference signal). Super luminescence diodes (SLDs) are commonly used for OCT imaging. SLDs with a center wavelength at 800 nm or 1300 nm are commercially available and can achieve an axial resolution of about 5–20 μm . The SLDs are compact, making the fiber-optic OCT system very portable.

Recently, advanced low-coherence light sources have been developed including short-pulse femto-second (fs) solid-state lasers and continuum generation in nonlinear photonic crystals.^[23–26] Representative fs lasers are 5–9 fs Ti:Sapphire lasers operating at a center wavelength of approximately 800 nm and a 14–20 fs Cr:Forsterite laser operating at a center wavelength of approximately 1400 nm.^[23,24,27] The resultant axial resolution is about 1.5–5 μm .^[28,29] Most of the high-performance short-pulse lasers rely on specially designed and expensive chirped or double-chirped mirrors for intracavity dispersion management.^[23,27,30] Recently, our group at the University of Washington has built a cost-effective, 7–8 fs transform-limited Ti:Sapphire laser with a moderate footprint (75L \times 28W \times 18H cm) using off-the-shelf optical components. As shown in Fig. 2, the laser has a broad spectrum of more than 150 nm smooth providing a free-space axial resolution of 2.5 μm .

Another type of advanced broadband light sources for OCT is continuum generation in microstructured or tapered nonlinear photonic fibers. A super-broad optical spectrum can be generated, ranging from the visible to the near-infrared (500–1600 nm).^[25,26]

In addition to the broad spectrum, the continuum light source provides high spatial coherence, high power, and can be potentially made very compact. The challenges of using the continuum include the requirement of using a short-pulse pump laser, careful polarization management, and spectrum shaping. It has been demonstrated that a 2- μm axial resolution can be achieved using the continuum with a center wavelength of approximately 1200–1300 nm.^[31,32] Figure 3 illustrates an OCT image of in vivo hamster cheek pouch using the continuum light source centered at 1300 nm.^[31] Microstructures such as the epithelium, muscle layer, and a blood vessel can be clearly identified.

When using a broad spectrum for OCT imaging, the dispersion between the sample and the reference arms

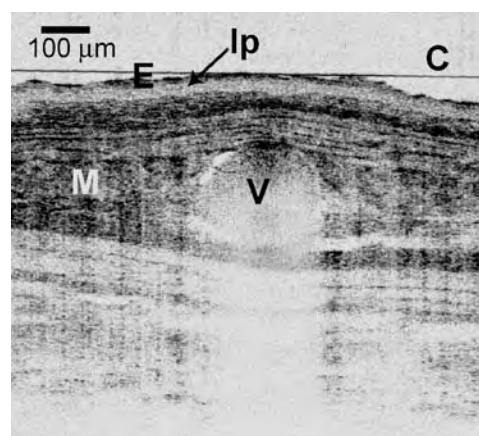


Fig. 3 In vivo OCT image of a hamster cheek pouch acquired with an ultrabroad continuum generated in a microstructure crystal fiber. Structures such as the epithelium (E), lamina propria (Ip), muscle layer (M), and a blood vessel (V), can be clearly identified. (Adapted from Ref. [31].)

needs to be carefully matched in order to achieve the optimal resolution provided by the light source.^[33,34] Unmatched dispersion will cause broadening of the temporal envelope of the interference signal in time domain, or cause out-of-phase summation of the interference signals at different wavelengths in the frequency domain, which will degrade the resolution offered by the light source.

OCT Imaging Speed

OCT imaging speed depends on how fast the optical pathlength in the reference arm can be scanned except when spectral domain techniques are employed.^[35–39] The typical scanning range required is about a few millimeters. Many OCT systems perform optical pathlength scanning by mechanically translating a retroreflector in the reference arm. This is one of the simplest ways to scan optical pathlength. As shown in Fig. 4A, the retroreflector can be translated reciprocally by a galvanometer to achieve a few-millimeter scan with a micron precision and a constant scanning velocity. A constant scanning velocity is advantageous because the Doppler frequency shift at a given wavelength is constant. Thus, heterodyne detection can be performed at a fixed center frequency with a fixed frequency bandwidth, resulting in a better signal-to-noise ratio than those where the center frequency changes with the scanning depth. This simple, mechanically translated scanner can be very compact and perform optical pathlength scanning at a velocity up to 100 mm/s with a repetition rate up to 100 Hz. An image of 500 axial scans can be acquired in about five seconds using this scanner. The most attractive feature of the retroreflecting scanner is that it does not induce extra dispersion during pathlength scanning, thus making it suitable for scanning an optical beam with an ultra broad spectrum without degrading the resolution.

Real-time OCT imaging became possible when the rapid-scanning optical phase delay line (RSOD) was developed.^[40–42] As shown in Fig. 4B, the delay line is based on a grating-lens pair, similar to those used for

ultrafast pulse shaping.^[43,44] The working principle of RSOD is elegant, and it relates to the time-shift theorem in the Fourier transform.^[45] The grating-lens pair first performs an optical Fourier transform; the scanning mirror adds a phase shift linear to each optical frequency (or color). The phase-shifted light is then sent back to the lens-grating pair, and a group time delay is generated after the inverse Fourier transform. As the mirror scans, the slope of the linear phase shift varies, resulting in a varying group delay or an equivalent pathlength scanning. With a RSOD, real-time OCT imaging at 2000 axial scans per second over a 2–3 mm depth is feasible. This corresponds to an imaging speed of four frames per second for an image size of 500 axial scans. Imaging at a higher repetition rate of 4000 scans/s is possible by using a resonant scanner.^[42] However, in this case the scanning velocity has a sinusoidal time-dependence, requiring a broader detection bandwidth. The RSOD has become the most commonly used pathlength scanner for real-time OCT imaging. The center Doppler frequency is tunable by choosing the offset of the beam incident on the scanning mirror relative to the scanning axis, which can be extremely convenient, depending on the applications and detection hardware.

Other high-speed optical pathlength scanners include a rotating glass slab,^[46] and a rotating cube driven by an air turbine.^[47] The fastest scanning velocity achieved to date is about 175 m/s with a repetition rate at 28,600 scans/s. The trade-off of fast scanning is the increased complexity of detection electronics and the reduced signal-to-noise ratio.

One challenge still remains with the aforementioned fast scanners. The RSOD and rotating cube scanners are dispersive. Extra dispersion is introduced in the reference arm during depth scanning, which causes dispersion mismatch between the sample and the reference arms. Due to the high scanning speed, dynamic dispersion compensation is extremely difficult, if not impossible. To overcome this difficulty, a rapid scanning all-reflective optical delay line (RSAD) has been recently developed^[48] which does not introduce scanning-depth-dependent dispersion. As shown in the schematic (Fig. 5), the RSAD only uses three reflective

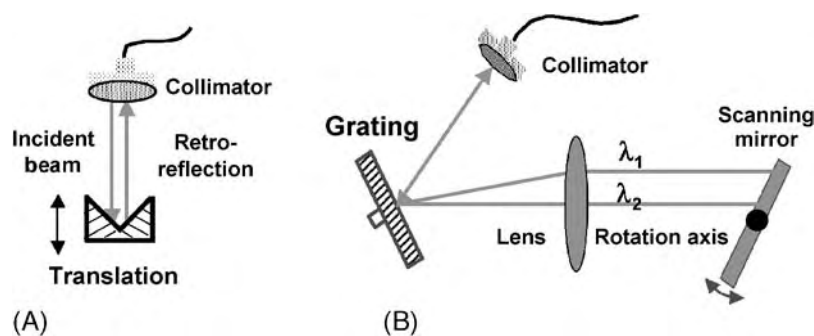


Fig. 4 (A) Schematic of a retroreflecting optical pathlength scanning delay line. The scanning repetition rate is up to approximately 100 Hz. (B) Rapid-scanning optical phase-controlled delay line (RSOD) based on a grating-lens pair and a scanning mirror. The scanning repetition rate is in the kHz range.

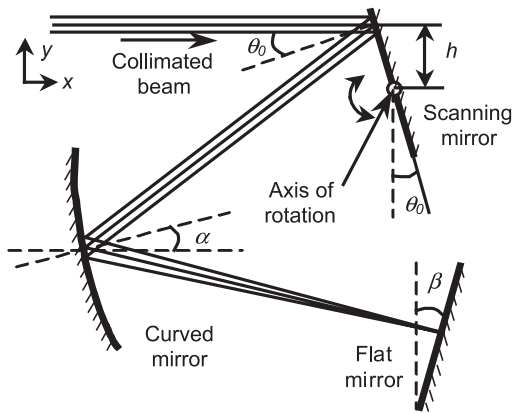


Fig. 5 Schematic of a rapid-scanning all-reflective delay line (RASD), which is nondispersive and suitable for real-time ultrahigh resolution OCT imaging when using an ultrabroad spectrum. Optical pathlength scanning is in the kHz range. (See Ref. [48] for details.)

mirrors, one curved and two flat mirrors. The collimated beam is incident on the scanning mirror with an offset relative to the scanning axis. Compared to another elegant all-reflective pathlength scanner based on a reflective rotary mirror array,^[49] the RSAD is extremely simple and easy to implement. Compared to the multiple-pass Herriott-cell-based scanner,^[50] RSAD is easy to align and has a high and flat coupling efficiency during the depth scanning. The overall size of the RAD is about 2.5×2.5 inches, and it can perform 1–1.2 mm depth scanning at a 4000-Hz repetition rate. The fluctuation of the light intensity in the reference arm is less than $\pm 5\%$, which keeps the heterodyne gain relatively flat over the imaging depth. A larger scanning range is possible by using a double-pass configuration or by using a large scanning mirror.

OCT Imaging Probes

A variety of OCT applications have been made possible by the development of novel OCT imaging probes. The OCT imaging probes deliver and focus a single-mode optical beam to the sample, and collect the beam backscattered from the sample. These imaging probes include a microscope-type device for imaging the skin and the eye,^[51–53] a scanning handheld probe for open-field imaging or surgical guidance,^[54,55] a flexible fiber-optic catheter/endoscope for intraluminal imaging of hollow organs,^[56–59] a Doppler catheter for imaging blood flow and a rigid microneedle for interstitial imaging of soft solid tissues.^[60,61] The flexible OCT catheter/endoscope was first reported in 1996,^[56] which greatly expanded OCT applicability for imaging internal organs including the gastrointestinal (GI) tract, bladder, female reproductive tract, and cardiovascular system.^[18,62–67] Figure 6 shows the schematic and a photo of the distal end of an OCT catheter/endoscope. It consists of a single-mode optical fiber, a graded-index (GRIN) lens to focus the beam, and a miniature prism to deflect the beam by 90° . The entire catheter is encased within a speedometer cable and has an overall diameter of about 1 mm (3 French). The catheter can be rotated by a motor at its proximal end in order to perform circumferential (or radial) imaging; it can also be longitudinally translated in order to perform imaging in a rectangular plane. The OCT catheter/endoscope has no active components and can be used directly to image hollow organs. Forward-imaging endoscopes have also been developed based on magnetically actuated fibers, MEMS scanners and piezo-driven fiber-optic scanners.^[59,68,69] The forward-imaging catheter/endoscope can be used in a wide range of applications and potentially attractive for guiding surgery or excisional biopsy since it permits

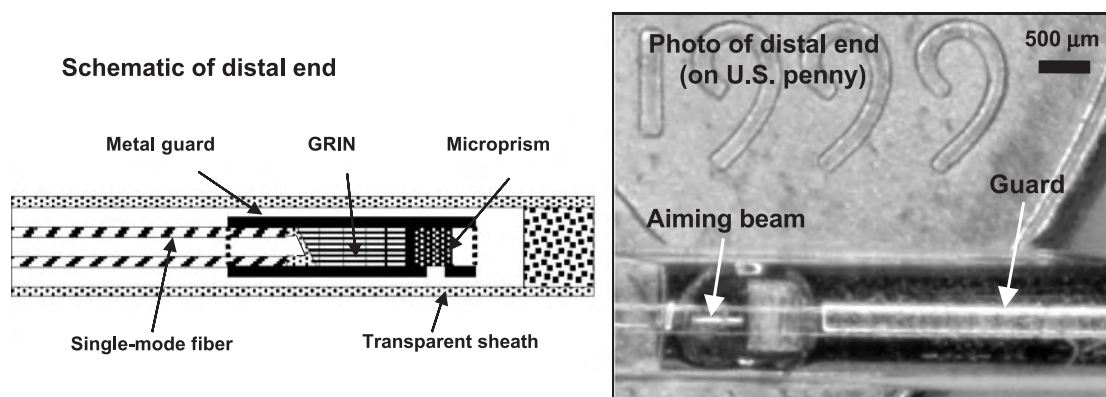


Fig. 6 Schematic and photograph of an OCT catheter/endoscope. The catheter scans the beam in a circumferential or longitudinal direction to cross-sectionally image hollow organs or vessels. A flexible speedometer cable encases a single-mode optical fiber, distal lens and a microprism. The beam is emitted perpendicularly to the catheter axis. (From Ref. [65].)

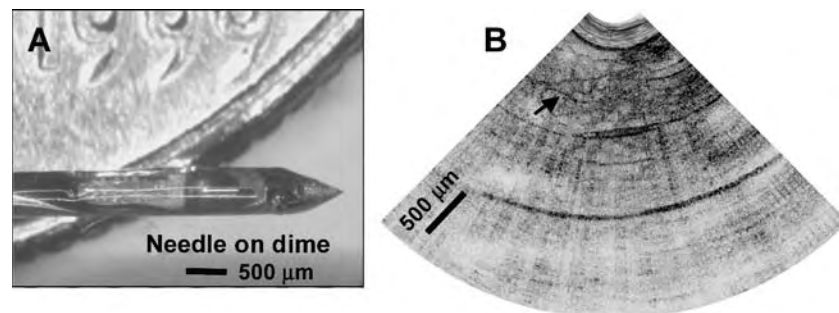


Fig. 7 (A) Photo of a 27-gauge OCT needle for interstitial imaging of internal tissue microstructures. (B) In vivo OCT image of a hamster leg muscle acquired with the OCT needle. Muscle fiber bundles can be identified as indicated by the arrow (F). (From Ref. [60])

imaging of tissues ahead of the device. Development of novel OCT endoscopes/catheters is still an area of active research. The small diameter of these imaging devices permits interfacing the catheter/endoscope with a wide range of medical instruments such as standard endoscopes by introducing the OCT device through the accessory channel. To prevent the polarization state of light from changing due to the torque on the fiber during circumferential imaging, a polarization-insensitive OCT endoscope with a built-in Faraday rotator at the distal end has been developed.^[58]

Due to the strong optical attenuation, OCT imaging depth is limited to 1–2 mm from the surface. Internal imaging of solid tissues was not possible prior to the development of a 27-gauge OCT imaging needle.^[60] The schematic of an OCT needle is similar to the OCT catheter and Fig. 7A shows a photo of a prototype OCT imaging needle. The overall diameter is about 410 μm (27-gauge) and devices with an even smaller diameter are possible. The relative distances between the optical components are set to achieve desired beam parameters including the focused spot size and the working distance. The distal tip of the needle is sharpened to facilitate interstitial delivery. Figure 7B illustrates an in vivo OCT needle image of the hamster leg muscle, from which the muscle fiber bundles can be clearly identified. The needle diameter is smaller than most excisional fine aspiration needles, so that tissues can be imaged with minimal trauma. Compared to a core biopsy needle, which generally samples the tissues with a diameter of 1 mm, the OCT imaging needle can have a sampling diameter of 2–6 mm (which is twice the imaging depth). This represents a 4- to 36-fold increase in the sampling volume. In addition, OCT needle imaging does not require tissue removal. Applications include fine needle optical biopsy and image-guided excisional biopsy.

Spectroscopic, Doppler, and Polarization OCT

In addition to structural imaging, OCT can also perform spectroscopic imaging, Doppler flow imaging,

and tissue birefringence imaging with a micron scale spatial resolution. As can be seen in Eq. 3, the wavelength-dependent E-field reflection coefficient, $s(f)$, contains information on tissue optical scattering and absorption. The backscattered optical spectra at different depths can be obtained by taking a short-time Fourier transform or wavelet transform of the interference signal followed by normalization with the source spectrum.^[70,71] Unlike conventional spectroscopy, OCT provides a spatially resolved reflection spectrum with a micron spatial resolution. Wavelength-dependent tissue optical attenuation due to scattering and/or absorption at different depths can be obtained by using the spectroscopic OCT. The spectral information can enhance optical contrast similar to staining employed in histology. Some pathologic changes involve cellular level variations such as nucleus enlargement in premalignant or early cancer, which result in changes in tissue scattering properties. Therefore, spectroscopic OCT could potentially reveal early pathologic changes.

Doppler OCT (or DOCT) is similar to Doppler ultrasound. It utilizes coherence gating to localize and resolve the flow volume where the backscattered light experiences a Doppler frequency shift (in addition to the known Doppler shift induced in the reference arm). Different from clinically viable laser Doppler velocimetry—which has been applied to many clinical scenarios for more than two decades^[72–74]—Doppler OCT provides depth-resolved quantitative flow information.^[5,6] In principle, the signal processing technique for DOCT is very similar to the spectroscopic OCT. The flow in the imaged tissue volume alters the heterodyne frequency, f_d , which will be the sum of the projected Doppler blood flow velocity along the optical beam and the pathlength scanning velocity in the reference arm, i.e., $f_d = 2fv_s/c - 2nfv_b \cos\theta/c$, where n is the index of refraction of the blood (or fluid in general), θ is the angle between the flow and the optical beam, and v_b is the blood flow velocity. Many methods have been developed to extract the depth-resolved flow velocity including the Short-time Fourier transform, Hilbert transform or phase-resolved method.^[6,75,76] Doppler OCT has been demonstrated for noninvasive

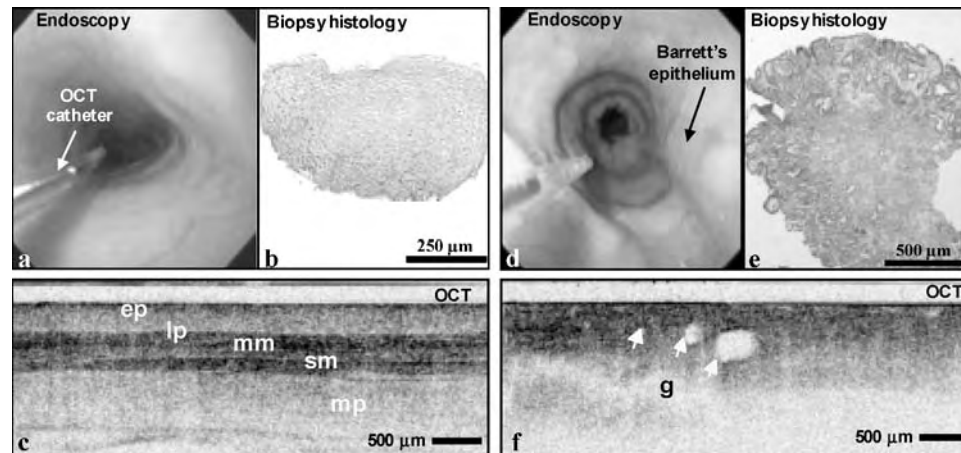


Fig. 8 Endoscopic and OCT images with corresponding pinch histology of normal (a–c) versus Barrett’s esophagus (d–f) of a patient. The OCT images were acquired at four frames per second using the longitudinal scanning scheme with $11 \times 30 \mu\text{m}$ (axial \times trans) resolution at a 1300 nm center wavelength. ep: epithelium; lp: lamina propria; mm: muscularis mucosa; sm: submucosa; mp: muscularis mucosa; g: gland. (From Ref. [65].)

assessment of blood flow on human skin and small animal skin-flap, and for evaluating the effect of photodynamic therapy and hemostatic therapy in rats.^[77,78] One area of particular interest is to quantify the blood vessel density using Doppler OCT, which might quantitatively reveal the angiogenesis process associated with tumor progression.

Many biological tissues such as muscle, tendon, cartilage, and tooth that contain organized collagen and elastin fibers exhibit birefringence, and the polarization state of light propagating within these tissues can be altered. Because the OCT interference signal represents the cross-correlation of two vector E-fields, it will be sensitive to the polarization states of light backscattered from the tissue and returned from the reference arm. Polarization-sensitive OCT (or PSOCT) was developed to quantitatively evaluate the depth-dependent tissue birefringence by measuring the birefringence effect on the interference signal.^[7–9,79] By changing the incident light polarization state and measuring the backscattered light polarization state, PSOCT can fully quantify the depth-dependent Jones or Muller matrix of biological tissues. Local changes of the polarization state of the light backscattered from thermally damaged skin have been revealed by PSOCT.^[80] It has also been demonstrated that PSOCT can enhance the imaging contrast compared to the standard envelope OCT images.^[81]

OCT APPLICATIONS

In a broad sense, OCT can be applied for diagnosis of diseases, guidance of surgical interventions, and monitoring of therapeutic effects. Optical coherence tomography was first developed to image the relatively

transparent tissues in the eye.^[1,10–12,51] Diagnostic ophthalmic OCT systems have been commercially available since 1996.^[82] With technology advances in the past half-decade, OCT has been successfully applied for imaging highly scattering tissues such as the skin, GI tract, and cardiovascular system. Although strong optical attenuation limits OCT imaging depth to 1–2 millimeters, this scale is ideal to cover the epidermis and the dermis of the skin and the epithelial/endothelial linings of the internal organs, from which many diseases originate such as neoplasia,^[13,14,57,63–65,83] and high-risk atherosclerotic plaques.^[17,18] In addition, the 1–2 mm imaging depth is comparable to (if not larger than) excisional biopsy depth, with the advantages that OCT provides image information in real time without the need to remove and process the specimen.

Clinical OCT studies in several organs and systems including (but not limited to) the skin, GI tract, female reproductive tract, joint, and cardiovascular system have just begun.^[13,14,18,57,63–65,83] It is beyond the scope of this article to describe the detailed findings for each disease that has been studied with OCT. In addition to the eye, the GI tract is the second organ that has been most extensively studied among all clinical OCT applications. Figure 8 illustrates representative OCT images acquired with an OCT imaging endoscope showing the normal and Barrett’s esophagus.^[65] During a routine GI endoscopy procedure, a 3.2-french OCT imaging probe was passed down the accessory port of a standard GI endoscope to the patient’s esophagus. Linear OCT images were taken at four frames per second by longitudinally scanning. As shown in Fig. 8, layered structures in the normal esophagus including the epithelium (ep), lamina propria (lp), muscularis mucosa (mm), submucosa (sm),

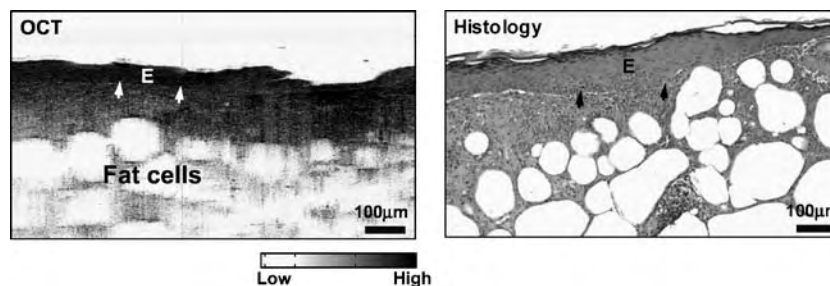


Fig. 9 OCT image and corresponding H&E histology micrograph of wounded skin of a diabetic mouse at day 10 post full-thickness skin punch biopsy. The thick, regenerated epidermis (E) bridges the wounded area and the dermis-epidermis junction can be clearly identified on the OCT image (as indicated by arrows). The OCT image correlates well with the corresponding histology (Details can be found in Ref. [95]).

and muscularis propria (mp) can be clearly distinguished from the alternating low and high back-scattering. The tissue within the epithelial layer is uniform, correlating very well with the histology of the pinch biopsy specimen. In contrast, normal layered structures were disrupted in the Barrett's esophagus and glandular structures became dominant. Again, the glandular metaplastic characteristics observed on the OCT image correlate well with the histologic findings. Recently, OCT imaging, in conjunction with quantitative tissue scattering analysis, has shown that OCT can differentiate normal tissues, metaplasia, dysplasia, and early cancer in the GI tract, and the findings are statistically significant.^[84] In the area of cancer diagnosis with OCT, efforts are being focused on detecting neoplasia at early stages when cancer is still curable. Promising approaches for detecting early neoplasia include the aforementioned quantitative method, spectroscopic OCT,^[84] the combination of angular-dependent light scattering with OCT,^[85,86] and integration of OCT with confocal microscopy.

Preliminary clinical studies have shown that OCT is capable of assessing structural changes in hard tissues such as the joint cartilage and teeth.^[16,55,87,88] Potential applications include detecting osteoarthritic cartilage degeneration, dental caries, and evaluating dental restoration. The micron resolution also makes OCT very promising for identifying high-risk atherosclerotic plaques in the cardiovascular system,^[18,89] in particular for those with a cap thickness of approximately 60 μm or less.^[90] The cardiovascular OCT application would require catheterization, similar to the intravascular ultrasound (IVUS) procedure, but provides a resolution of almost an order of magnitude higher than IVUS.^[18,91]

In addition to clinical applications, OCT has many applications as a basic research tool as well. Optical coherence tomography permits repeated assessment of the same animal *in vivo* and in real time. This can be particularly valuable for developmental biology.^[92-94] OCT can also be a useful tool for monitoring disease

progression and/or treatment outcome on an animal model over the entire time course, avoiding interanimal variance when using a large colony of animals. Recently, one example to use OCT for assessing impaired cutaneous wound healing on a diabetic mouse model.^[95] After a full-thickness cutaneous wound was created by skin punch biopsy, the wounded skin was repeatedly monitored by OCT imaging. Optical coherence tomography images at different time points during the healing process depict the formation of granulation tissue, the bridging of the regenerated epidermis over the wound, and the thickening of the regenerated epidermis. Figure 9 shows a representative OCT image of the cutaneous wound 10 days post-wound induction. The regenerated epidermis that bridges over the wounded area, the thickening of the new epidermis, and the dermis-epidermis (DE) junction can be clearly identified on the OCT image. Optical coherence tomography findings correlate very well with corresponding histology. Longitudinal OCT imaging of the same animal also reveals that the diabetic mouse exhibits a retarded wound healing process compared to the normal mouse (littermate). Preliminary results strongly suggest that OCT can be a valuable noninvasive tool for objectively evaluating the wound healing process. This research can be closely related to biomaterials and tissue engineering^[96,97] The effects of various growth factors on wound healing and the control mechanisms used for releasing the growth factors can be systematically studied using OCT on an animal model *in vivo*.

CONCLUSION

Optical coherence tomography is emerging as an attractive, noninvasive imaging modality, which is viable for clinical diagnosis and basic biomedical research. The past decade has witnessed rapid advances of the OCT technology. At present, OCT is well positioned for laboratory-based research. Except in

ophthalmology, widespread clinical OCT applications in other areas will be boosted by further technology development for improving resolution, beam delivery/collection, and imaging speed. Functional integration of structural, Doppler flow, birefringence, and spectroscopy would potentially enhance OCT image contrast and specificity for diagnosis of diseases at early stages. Development of image processing algorithms and firm understanding of the theoretical foundation of OCT—in particular about resolution degradation due to light-tissue interaction, chromatic dispersion, and the nature of OCT contrast—could lead to another level of image quality.

ACKNOWLEDGMENTS

The contributions of Dr. Yuchuan Chen, Mr. Michael J. Cobb, Dr. Michael B. Kimmey, Dr. Xiumei Liu, Dr. John E. Olerud, Robert Underwood, and Marcia Usui to the unpublished data/images are gratefully acknowledged. For other useful discussions over many years, the authors are also grateful to many former and present colleagues and collaborators, in particular, Dr. Mark E. Brezinski, Dr. Stephen Boppart, Spencer Bowen, Dr. Christian Chudoba, Dr. Wolfgang Drexler, Dr. Ingmar Hartl, Dr. Thomas Hatsukami, Dr. Joo Ha Hwang, Dr. Christopher Kemp, Dr. Hiroshi Mashimo, Dr. Patrick Stayton, Dr. Melissa Upton, Dr. Jacques Van Dam, and Dr. Chun Yuan. Financial supports from the National Institutes of Health, the Whitaker Foundation, and the American Heart Association are acknowledged.

ARTICLES OF FURTHER INTEREST

Electrical Impedance Imaging; Medical Imaging, 3-D; Microcomputed Tomography and Its Applications

REFERENCES

- Huang, D.; Swanson, E.A.; Lin, C.P.; Schuman, J.S.; Stinson, W.G.; Chang, W.; Hee, M.R.; Flotte, T.; Gregory, K.; Puliafito, C.A.; Fujimoto, J.G. Optical coherence tomography. *Science* **1991**, *254* (5035), 1178–1181.
- Schmitt, J.M. Optical coherence tomography (OCT): A review. *IEEE J. Sel. Top. Quantum Electron.* **1999**, *5* (4), 1205–1215.
- Handbook of Optical Coherence Tomography*; Bouma, B.E., Tearney, G.J., Eds.; Marcel Dekker, Inc.: New York, 2001.
- Fujimoto, J.G. Optical coherence tomography for ultrahigh resolution in vivo imaging. *Nat. Biotechnol.* **2003**, *21* (11), 1361–1367.
- Chen, Z.P.; Milner, T.E.; Srinivas, S.; Wang, X.J.; Malekafzali, A.; vanGemert, M.J.C.; Nelson, J.S. Noninvasive imaging of in vivo blood flow velocity using optical Doppler tomography. *Opt. Lett.* **1997**, *22* (14), 1119–1121.
- Izatt, J.A.; Kulkarni, M.D.; Yazdanfar, S.; Barton, J.K.; Welch, A.J. In vivo bidirectional color Doppler flow imaging of picoliter blood volumes using optical coherence tomography. *Opt. Lett.* **1997**, *22* (18), 1439–1441.
- de Boer, J.F.; Milner, T.E.; vanGemert, M.J.C.; Nelson, J.S. Two-dimensional birefringence imaging in biological tissue by polarization-sensitive optical coherence tomography. *Opt. Lett.* **1997**, *22* (12), 934–936.
- Everett, M.J.; Schoenenberger, K.; Colston, B.W.; Da Silva, L.B. Birefringence characterization of biological tissue by use of optical coherence tomography. *Opt. Lett.* **1998**, *23* (3), 228–230.
- Yao, G.; Wang, L.V. Two-dimensional depth-resolved Mueller matrix characterization of biological tissue by optical coherence tomography. *Opt. Lett.* **1999**, *24* (8), 537–539.
- Hee, M.R.; Izatt, J.A.; Swanson, E.A.; Huang, D.; Schuman, J.S.; Lin, C.P.; Puliafito, C.A.; Fujimoto, J.G. Optical coherence tomography of the human retina. *Arch. Ophthalmol.* **1995**, *113* (3), 325–332.
- Puliafito, C.A.; Hee, M.R.; Lin, C.P.; Reichel, E.; Schuman, J.S.; Duker, J.S.; Izatt, J.A.; Swanson, E.A.; Fujimoto, J.G. Imaging of macular diseases with optical coherence tomography. *Ophthalmology* **1995**, *102* (2), 217–229.
- Drexler, W.; Morgner, U.; Ghanta, R.K.; Kartner, F.X.; Schuman, J.S.; Fujimoto, J.G. Ultrahigh-resolution ophthalmic optical coherence tomography. *Nat. Med.* **2001**, *7* (4), 502–507.
- Gladkova, N.D.; Petrova, G.A.; Nikulin, N.K.; Radenska-Lopovok, S.G.; Snopova, L.B.; Chumakov, Y.P.; Nasonova, V.A.; Gelikonov, V.M.; Gelikonov, G.V.; Kuranov, R.V.; Sergeev, A.M.; Feldchtein, F.I. In vivo optical coherence tomography imaging of human skin: Norm and pathology. *Skin Res. Technol.* **2000**, *6* (1), 6–16.
- Welzel, J. Optical coherence tomography in dermatology: A review. *Skin Res. Technol.* **2001**, *7* (1), 1–9.
- Brand, S.; Poneros, J.M.; Bouma, B.E.; Tearney, G.J.; Compton, C.C.; Nishioka, N.S. Optical coherence tomography in the gastrointestinal tract. *Endoscopy* **2000**, *32* (10), 796–803.
- Feldchtein, F.I.; Gelikonov, G.V.; Gelikonov, V.M.; Iksanov, R.R.; Kuranov, R.V.; Sergeev, A.M.; Gladkova, N.D.; Ourutina, M.N.; Warren, J.A.; Reitze, D.H. In vivo OCT imaging of hard and soft tissue of the oral cavity. *Opt. Express* **1998**, *3* (6), 239–250.
- Brezinski, M.E.; Tearney, G.J.; Bouma, B.E.; Izatt, J.A.; Hee, M.R.; Swanson, E.A.; Southern, J.F.; Fujimoto, J.G. Optical coherence tomography for optical biopsy—Properties and demonstration of vascular pathology. *Circulation* **1996**, *93* (6), 1206–1213.
- Jang, I.K.; Bouma, B.E.; Kang, D.H.; Park, S.J.; Park, S.W.; Seung, K.B.; Choi, K.B.; Shishkov, M.; Schlendorf, K.; Pomerantsev, E.; Houser, S.L.; Aretz,



- H.T.; Tearney, G.J. Visualization of coronary atherosclerotic plaques in patients using optical coherence tomography: Comparison with intravascular ultrasound. *J. Am. Coll. Cardiol.* **2002**, *39* (4), 604–609.
19. Jang, I.K.; MacNeill, B.D.; Yabushita, H.; DeJoseph, D.; Kauffman, C.R.; Tearney, G.J.; Bouma, B.E. In-vivo characterization of coronary plaques in patients with ST elevation acute myocardial infarction using optical coherence tomography (OCT). *Circulation* **2002**, *106* (19), 698.
 20. Saleh, B.E.A.; Teich, M.C. *Fundamentals of Photonics*; John Wiley & Sons, Inc.: New York, 1991.
 21. Matsumoto, H.; Hirai, A. A white-light interferometer using a lamp source and heterodyne detection with acousto-optic modulators. *Opt. Commun.* **1999**, *170* (4–6), 217–220.
 22. Vabre, L.; Dubois, A.; Boccarda, A.C. Thermal-light full-field optical coherence tomography. *Opt. Lett.* **2002**, *27* (7), 530–532.
 23. Morgner, U.; Kartner, F.X.; Cho, S.H.; Chen, Y.; Haus, H.A.; Fujimoto, J.G.; Ippen, E.P.; Scheuer, V.; Angelow, G.; Tschudi, T. Sub-two-cycle pulses from a Kerr-lens mode-locked Ti: Sapphire laser. *Opt. Lett.* **1999**, *24* (6), 411–413.
 24. Chudoba, C.; Fujimoto, J.G.; Ippen, E.P.; Haus, H.A.; Morgner, U.; Kartner, F.X.; Scheuer, V.; Angelow, G.; Tschudi, T. All-solid-state Cr: Forsterite laser generating 14-fs pulses at 1.3 μm . *Opt. Lett.* **2001**, *26* (5), 292–294.
 25. Birks, T.A.; Wadsworth, W.J.; Russell, P.S. Super-continuum generation in tapered fibers. *Opt. Lett.* **2000**, *25* (19), 1415–1417.
 26. Ranka, J.K.; Windeler, R.S.; Stentz, A.J. Visible continuum generation in air–silica microstructure optical fibers with anomalous dispersion at 800 nm. *Opt. Lett.* **2000**, *25* (1), 25–27.
 27. Unterhuber, A.; Povazay, B.; Hermann, B.; Sattmann, H.; Drexler, W.; Yakovlev, V.; Tempea, G.; Schubert, C.; Anger, E.M.; Ahnelt, P.K.; Stur, M.; Morgan, J.E.; Cowey, A.; Jung, G.; Le, T.; Stingl, A. Compact, low-cost Ti: Al₂O₃ laser for in vivo ultrahigh-resolution optical coherence tomography. *Opt. Lett.* **2003**, *28* (11), 905–907.
 28. Drexler, W.; Morgner, U.; Kartner, F.X.; Pitris, C.; Boppart, S.A.; Li, X.D.; Ippen, E.P.; Fujimoto, J.G. In vivo ultrahigh-resolution optical coherence tomography. *Opt. Lett.* **1999**, *24* (17), 1221–1223.
 29. Bouma, B.E.; Tearney, G.J.; Bilinsky, I.P.; Golubovic, B.; Fujimoto, J.G. Self-phase-modulated Kerr-lens mode-locked Cr: Forsterite laser source for optical coherence tomography. *Opt. Lett.* **1996**, *21* (22), 1839–1841.
 30. Kowalevicz, A.M.; Schibli, T.R.; Kartner, F.X.; Fujimoto, J.G. Ultralow-threshold Kerr-lens mode-locked Ti: Al₂O₃ laser. *Opt. Lett.* **2002**, *27* (22), 2037–2039.
 31. Hartl, I.; Li, X.D.; Chudoba, C.; Ghanta, R.K.; Ko, T.H.; Fujimoto, J.G.; Ranka, J.K.; Windeler, R.S. Ultrahigh-resolution optical coherence tomography using continuum generation in an air–silica microstructure optical fiber. *Opt. Lett.* **2001**, *26* (9), 608–610.
 32. Wang, Y.M.; Zhao, Y.H.; Nelson, J.S.; Chen, Z.P.; Windeler, R.S. Ultrahigh-resolution optical coherence tomography by broadband continuum generation from a photonic crystal fiber. *Opt. Lett.* **2003**, *28* (3), 182–184.
 33. Hitzenberger, C.K.; Baumgartner, A.; Drexler, W.; Fercher, A.F. Dispersion effects in partial coherence interferometry: Implications for intraocular ranging. *J. Biomed. Opt.* **1999**, *4* (1), 144–151.
 34. Niblack, W.K.; Schenk, J.O.; Liu, B.; Brezinski, M.E. Dispersion in a grating-based optical delay line for optical coherence tomography. *Appl. Opt.* **2003**, *42* (19), 4115–4118.
 35. Fercher, A.F.; Hitzenberger, C.K.; Kamp, G.; Elzaiat, S.Y. Measurement of intraocular distances by back-scattering spectral interferometry. *Opt. Commun.* **1995**, *117* (1–2), 43–48.
 36. Hausler, G.; Lindner, M.W. Coherent radar and spectral radar—New tools for dermatological diagnosis. *J. Biomed. Opt.* **1998**, *3* (1), 21–31.
 37. Wojtkowski, M.; Leitgeb, R.; Kowalczyk, A.; Bajraszewski, T.; Fercher, A.F. In vivo human retinal imaging by Fourier domain optical coherence tomography. *J. Biomed. Opt.* **2002**, *7* (3), 457–463.
 38. Yun, S.H.; Tearney, G.J.; de Boer, J.F.; Itimtia, N.; Bouma, B.E. High-speed optical frequency-domain imaging. *Opt. Express* **2003**, *11* (22), 2953–2963.
 39. Nassif, N.; Cense, B.; Park, B.H.; Yun, S.H.; Chen, T.C.; Bouma, B.E.; Tearney, G.J.; de Boer, J.F. In vivo human retinal imaging by ultrahigh-speed spectral domain optical coherence tomography. *Opt. Lett.* **2004**, *29* (5), 480–482.
 40. Kwong, K.F.; Yankelevich, D.; Chu, K.C.; Heritage, J.P.; Dienes, A. 400-Hz Mechanical scanning optical delay-line. *Opt. Lett.* **1993**, *18* (7), 558–560.
 41. Tearney, G.J.; Bouma, B.E.; Fujimoto, J.G. High-speed phase- and group-delay scanning with a grating-based phase control delay line. *Opt. Lett.* **1997**, *22* (23), 1811–1813.
 42. Rollins, A.M.; Kulkarni, M.D.; Yazdanfar, S.; Ungarunyawee, R.; Izatt, J.A. In vivo video rate optical coherence tomography. *Opt. Express* **1998**, *3* (6), 219–229.
 43. Treacy, E.B. Optical pulse compression with diffraction gratings. *IEEE J. Quantum Electron.* **1969**, *QE-5* (9), 454–458.
 44. Heritage, J.P.; Weiner, A.M.; Thurston, R.N. Picosecond pulse shape manipulation in a grating pulse—Compressor. *J. Opt. Soc. Am., A, Opt. Image Sci. Vision* **1985**, *2* (13), P43.
 45. Oppenheim, A.V.; Willsky, A.S.; Nawab, S.H. *Signals & Systems*; Prentice Hall: Upper Saddle River, NJ, 1997.
 46. Su, C.B. Achieving variation of the optical path length by a few millimeters at millisecond rates for imaging of turbid media and optical interferometry: A new technique. *Opt. Lett.* **1997**, *22* (10), 665–667.
 47. Szydlow, J.; Delachenal, N.; Giannotti, R.; Walti, R.; Bleuler, H.; Salathe, R.P. Air-turbine driven optical low-coherence reflectometry at 28.6-kHz scan repetition rate. *Opt. Commun.* **1998**, *154* (1–3), 1–4.

48. Liu, X.M.; Cobb, M.J.; Li, X.D. Rapid scanning all-reflective optical delay line for real-time optical coherence tomography. *Opt. Lett.* **2004**, *29* (1), 80–82.
49. Chen, N.G.; Zhu, Q. Rotary mirror array for high speed optical coherence tomography. *Opt. Lett.* **2002**, *27* (8), 607–609.
50. Hsiung, P.L.; Li, X.D.; Chudoba, C.; Hartl, I.; Ko, T.H.; Fujimoto, J.G. High-speed path-length scanning with a multiple-pass cavity delay line. *Appl. Opt.* **2003**, *42* (4), 640–648.
51. Swanson, E.A.; Izatt, J.A.; Hee, M.R.; Huang, D.; Lin, C.P.; Schuman, J.S.; Puliafito, C.A.; Fujimoto, J.G. In-vivo retinal imaging by optical coherence tomography. *Opt. Lett.* **1993**, *18* (21), 1864–1866.
52. Schmitt, J.M.; Yadlowsky, M.J.; Bonner, R.F. Sub-surface imaging of living skin with optical coherence microscopy. *Dermatology* **1995**, *191* (2), 93–98.
53. Pan, Y.; Farkas, D.L. Noninvasive imaging of living human skin with dual-wavelength optical coherence tomography in two and three dimensions. *J. Biomed. Opt.* **1998**, *3* (4), 446–455.
54. Boppart, S.A.; Bouma, B.E.; Pitris, C.; Tearney, G.J.; Fujimoto, J.G.; Brezinski, M.E. Forward-imaging instruments for optical coherence tomography. *Opt. Lett.* **1997**, *22* (21), 1618–1620.
55. Li, X.D.; Drexler, W.; Pitris, C.; Ghanta, R.; Jesser, C.; Hermann, J.; Stamper, D.; Golden, D.; Martin, S.; Fujimoto, J.G.; Brezinski, M.E. Imaging of Osteoarthritic Cartilage With Optical Coherence Tomography: Micronstructure and Polarization Sensitivity. In *Conference on Lasers and Electro-Optics*; Optical Society of America and IEEE: Baltimore, MD, 1999; 339.
56. Tearney, G.J.; Boppart, S.A.; Bouma, B.E.; Brezinski, M.E.; Weissman, N.J.; Southern, J.F.; Fujimoto, J.G. Scanning single-mode fiber optic catheter-endoscope for optical coherence tomography. *Opt. Lett.* **1996**, *21* (7), 543–545.
57. Sergeev, A.M.; Gelikonov, V.M.; Gelikonov, G.V.; Feldchtein, F.I.; Kuranov, R.V.; Gladkova, N.D.; Shakhova, N.M.; Suopova, L.B.; Shakhov, A.V.; Kuznetzova, I.A.; Denisenko, A.N.; Pochinko, V.V.; Chumakov Yu, P.; Streltsova, O.S. In vivo endoscopic OCT imaging of precancer and cancer states of human mucosa. *Opt. Express* **1997**, *1* (13).
58. Rollins, A.M.; Ung-arunyawee, R.; Chak, A.; Wong, R.C.K.; Kobayashi, K.; Sivak, M.V.; Izatt, J.A. Real-time in vivo imaging of human gastrointestinal ultrastructure by use of endoscopic optical coherence tomography with a novel efficient interferometer design. *Opt. Lett.* **1999**, *24* (19), 1358–1360.
59. Pan, Y.T.; Xie, H.K.; Fedder, G.K. Endoscopic optical coherence tomography based on a microelectromechanical mirror. *Opt. Lett.* **2001**, *26* (24), 1966–1968.
60. Li, X.D.; Chudoba, C.; Ko, T.; Pitris, C.; Fujimoto, J.G. Imaging needle for optical coherence tomography. *Opt. Lett.* **2000**, *25* (20), 1520–1522.
61. Li, X.D.; Ko, T.H.; Fujimoto, J.G. Intraluminal fiber-optic Doppler imaging catheter for structural and functional coherence tomography. *Opt. Lett.* **2001**, *26* (23), 1906–1908.
62. Tearney, G.J.; Brezinski, M.E.; Bouma, B.E.; Boppart, S.A.; Pitris, C.; Southern, J.F.; Fujimoto, J.G. In vivo endoscopic optical biopsy with optical coherence tomography. *Science* **1997**, *276* (5321), 2037–2039.
63. Bouma, B.E.; Tearney, G.J.; Compton, C.C.; Nishioka, N.S. High-resolution imaging of the human esophagus and stomach in vivo using optical coherence tomography. *Gastrointest. Endosc.* **2000**, *51* (4 Pt 1), 467–474.
64. Sivak, M.V., Jr.; Kobayashi, K.; Izatt, J.A.; Rollins, A.M.; Ung-Runyawee, R.; Chak, A.; Wong, R.C.; Isenberg, G.A.; Willis, J. High-resolution endoscopic imaging of the GI tract using optical coherence tomography. *Gastrointest. Endosc.* **2000**, *51* (4 Pt 1), 474–479.
65. Li, X.D.; Boppart, S.A.; Van Dam, J.; Mashimo, H.; Mutinga, M.; Drexler, W.; Klein, M.; Pitris, C.; Krinsky, M.L.; Brezinski, M.E.; Fujimoto, J.G. Optical coherence tomography: Advanced technology for the endoscopic imaging of Barrett's esophagus. *Endoscopy* **2000**, *32* (12), 921–930.
66. Feldchtein, F.I.; Gelikonov, G.V.; Gelikonov, V.M.; Kuranov, R.V.; Sergeev, A.M.; Gladkova, N.D.; Shakhov, A.V.; Shakhova, N.M.; Snopova, L.B.; Terent'eva, A.B.; Zagainova, E.V.; Chumakov Yu, P.; Kuznetzova, I.A. Endoscopic applications of optical coherence tomography. *Opt. Express* **1998**, *3* (6).
67. Fujimoto, J.G.; Boppart, S.A.; Tearney, G.J.; Bouma, B.E.; Pitris, C.; Brezinski, M.E. High resolution in vivo intra-arterial imaging with optical coherence tomography. *Heart* **1999**, *82* (2), 128–133.
68. Feldchtein, F.I.; Gelikonov, V.M.; Gelikonov, G.V. *Handbook of Optical Coherence Tomography*; Bouma, B.E.; Tearney, G.J., Eds.; Marcel Dekker, Inc.: New York, 2001; 125–142.
69. Liu, X.M.; Cobb, M.J.; Chen, Y.C.; Li, X.D. Miniature Lateral Priority Scanning Endoscope for Real-Time Forward-Imaging Optical Coherence Tomography. *Opt. Lett.* *in press*.
70. Kulkarni, M.D.; Izatt, J.A. Spectroscopic Optical Coherence Tomography. In *Conference on Lasers and Electro Optics (CLEO'96)*; Optical Society of America: Anaheim, CA, 1996; 59–60.
71. Morgner, U.; Drexler, W.; Kaertner, F.X.; Li, X.D.; Pitris, C.; Ippen, E.P.; Fujimoto, J.G. Spectroscopic optical coherence tomography. *Opt. Lett.* **2000**, *25* (2), 111–113.
72. Tanaka, T.; Riva, C.; Ben-Sira, B. Blood velocity measurements in human retinal vessels. *Science* **1974**, *186* (4166), 830–831.
73. Stern, M.D. In vivo evaluation of microcirculation by coherent light scattering. *Nature* **1975**, *254* (5495), 56–58.
74. Bonner, R.F.; Nossal, R. *Laser Doppler Blood Flowmetry*; Shepherd, A.P., Odberg, P.A., Eds.; Kluwer Academic: Boston, 1990; 17–45.
75. Zhao, Y.H.; Chen, Z.P.; Saxer, C.; Xiang, S.H.; de Boer, J.F.; Nelson, J.S. Phase-resolved optical coherence tomography and optical Doppler tomography for imaging blood flow in human skin with fast scanning speed and high velocity sensitivity. *Opt. Lett.* **2000**, *25* (2), 114–116.

76. Yang, V.X.D.; Gordon, M.L.; Mok, A.; Zhao, Y.H.; Chen, Z.P.; Cobbold, R.S.C.; Wilson, B.C.; Vitkin, I.A. Improved phase-resolved optical Doppler tomography using the Kasai velocity estimator and histogram segmentation. *Optics Commun.* **2002**, *208* (4–6), 209–214.
77. Major, A.; Kimel, S.; Mee, S.; Milner, T.E.; Smithies, D.J.; Srinivas, S.M.; Chen, Z.P.; Nelson, J.S. Microvascular photodynamic effects determined in vivo using optical Doppler tomography. *IEEE J. Sel. Top. Quant. Electron.* **1999**, *5* (4), 1168–1175.
78. Wong, R.C.K.; Yazdanfar, S.; Izatt, J.A.; Kulkarni, M.D.; Barton, J.K.; Welch, A.J.; Willis, J.; Sivak, M.V. Visualization of subsurface blood vessels by color Doppler optical coherence tomography in rats: Before and after hemostatic therapy. *Gastrointest. Endosc.* **2002**, *55* (1), 88–95.
79. Hee, M.R.; Huang, D.; Swanson, E.A.; Fujimoto, J.G. Polarization-sensitive low-coherence reflectometer for birefringence characterization and ranging. *J. Opt. Soc. Am., B* **1992**, *9*, 903–908.
80. de Boer, J.F.; Srinivas, S.M.; Malekafzali, A.; Chen, Z.P.; Nelson, J.S. Imaging thermally damaged tissue by polarization sensitive optical coherence tomography. *Opt. Express* **1998**, *3* (6), 212–218.
81. Jiao, S.L.; Wang, L.H.V. Two-dimensional depth-resolved Mueller matrix of biological tissue measured with double-beam polarization-sensitive optical coherence tomography. *Opt. Lett.* **2002**, *27* (2), 101–103.
82. Humphrey-Website. (Information on commercial ophthalmic optical coherence tomography is available on <http://www.meditec.zeiss.com/>).
83. Jackle, S.; Gladkova, N.; Feldchtein, F.; Terentjeva, A.; Brand, B.; Gelikonov, G.; Gelikonov, V.; Sergeev, A.; Fritscher-Ravens, A.; Freund, J.; Seitz, U.; Schroder, S.; Soehendra, N. In vivo endoscopic optical coherence tomography of the human gastrointestinal tract—Toward optical biopsy. *Endoscopy* **2000**, *32* (10), 743–749.
84. Pfau, P.R.; Sivak, M.V.; Chak, A.; Kinnard, M.; Wong, R.C.K.; Isenberg, G.A.; Izatt, J.A.; Rollins, A.; Westphal, V. Criteria for the diagnosis of dysplasia by endoscopic optical coherence tomography. *Gastrointest. Endosc.* **2003**, *58* (2), 196–202.
85. Wax, A.; Yang, C.H.; Dasari, R.R.; Feld, M.S. Measurement of angular distributions by use of low-coherence interferometry for light-scattering spectroscopy. *Opt. Lett.* **2001**, *26* (6), 322–324.
86. Wax, A.; Yang, C.H.; Backman, V.; Badizadegan, K.; Boone, C.W.; Dasari, R.R.; Feld, M.S. Cellular organization and substructure measured using angle—Resolved low-coherence interferometry. *Biophys. J.* **2002**, *82* (4), 2256–2264.
87. Colston, B.W.; Sathyam, U.S.; DaSilva, L.B.; Everett, M.J.; Stroeve, P.; Otis, L.L. Dental OCT. *Opt. Express* **1998**, *3* (6), 230–238.
88. Otis, L.L.; Everett, M.J.; Sathyam, U.S.; Colston, B.W., Jr. Optical coherence tomography: A new imaging technology for dentistry. *J. Am. Dent. Assoc.* **2000**, *131* (4), 511–514.
89. Tearney, G.J.; Yabushita, H.; Houser, S.L.; Aretz, H.T.; Jang, I.K.; Schlendorf, K.H.; Kauffman, C.R.; Shishkov, M.; Halpern, E.F.; Bouma, B.E. Quantification of macrophage content in atherosclerotic plaques by optical coherence tomography. *Circulation* **2003**, *107* (1), 113–119.
90. Virmani, R.; Kolodgie, F.D.; Burke, A.P.; Farb, A.; Schwartz, S.M. Lessons from sudden coronary death: A comprehensive morphological classification scheme for atherosclerotic lesions. *Arterioscler. Thromb. Vasc. Biol.* **2000**, *20* (5), 1262–1275.
91. Patwari, P.; Weissman, N.J.; Boppart, S.A.; Jesser, C.; Stamper, D.; Fujimoto, J.G.; Brezinski, M.E. Assessment of coronary plaque with optical coherence tomography and high-frequency ultrasound. *Am. J. Cardiol.* **2000**, *85* (5), 641–644.
92. Boppart, S.A.; Brezinski, M.E.; Bouma, B.E.; Tearney, G.J.; Fujimoto, J.G. Investigation of developing embryonic morphology using optical coherence tomography. *Dev. Biol.* **1996**, *177* (1), 54–63.
93. Boppart, S.A.; Tearney, G.J.; Bouma, B.E.; Southern, J.F.; Brezinski, M.E.; Fujimoto, J.G. Noninvasive assessment of the developing xenopus cardiovascular system using optical coherence tomography. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94* (9), 4256–4261.
94. Yelbuz, T.M.; Choma, M.A.; Thrane, L.; Kirby, M.L.; Izatt, J.A. Optical coherence tomography—A new high-resolution imaging technology to study cardiac development in chick embryos. *Circulation* **2002**, *106* (22), 2771–2774.
95. Cobb, M.J.; Chen, Y.C.; Underwood, R.A.; Thariani, R.A.; Usui, M.L.; Olerud, J.E.; Li, X.D. Assessment of Cutaneous Wound Healing Using Optical Coherence Tomography. In *OSA Biomedical Optics Topical Meetings*; Optical Society of America: Miami Beach, FL, 2004; FH22.
96. Patterson, J.; Li, X.D.; Herring, S.; Stayton, P.S. Regenerative Matrices for Oriented Craniofacial Bone Growth. In *The 2003 Annual BMES Fall Meeting*; Session 2.6: Tissue Engineering of Bone and Cartilage; Biomedical Engineering Society: Nashville, TN, 2003.
97. Dunkers, J.P.; Cicerone, M.T.; Washburn, N.R. Collinear optical coherence and confocal fluorescence microscopies for tissue engineering. *Opt. Express* **2003**, *11* (23), 3074–3079.

Optical Detection of Cancers

Xin-Hua Hu
Jun Qing Lu

Department of Physics, East Carolina University, Greenville, North Carolina, U.S.A.

INTRODUCTION

The American Cancer Society estimates that approximately 570,000 cancer deaths will occur in 2005 in United States.^[1] The key to achieving the desired patient prognosis is to detect cancer in the early stages. Among various diagnosis modalities, optical detection stands out for its potential of safety to patients and health professionals with nonionizing radiation, non-invasive nature of detection, high spatial resolution, and modest cost of equipment. Optical detection of cancers is a vast subject that has attracted intensive research efforts over the last few decades. It is therefore impossible to survey the field in a relative short essay like the current one. Instead, we will focus on the in vivo results that may lead to the noninvasive staging of dysplasia and cancers. Some in vitro data will be included here to reflect the historical development of a method or when in vivo results are not available.

It has been widely accepted that cancer originates from the accumulated mutation of DNA molecules within a cell or cells that tips the balance of cell cycle toward proliferation.^[2] Cancerous cells multiply without regulation to form a morphologically distinctive mass of tissue consisting of their clones in a local organ, often termed dysplasia, before the mass grows in size and eventually metastasizes into other organs to become a malignant tumor. Nearly all cancers develop in a staged fashion and the gold standard of cancer diagnosis and staging is the histopathologic analysis of biopsied tissues.^[3] Even with this very invasive and time-consuming technique, the criteria for staging of the dysplasia or the precancer stages are still not well defined and diagnosis varies significantly among pathologists. As a result, optical diagnosis of dysplasia lacks some clear standards to be validated against albeit greatly desired by physicians.

Optical detection of lesions was called “optical biopsy” in various publications in reference to the hope that a noninvasive replacement of biopsy and histopathology could be achieved optically. To reach this goal, one must identify the optical signatures of dysplasia and subsequent stages of cancer development, which may be correlated to three different types of variations: the mutation at the molecular level, the morphological variations at the cellular level, and

variations in tissue and vasculature (blood volume and oxygen saturation) structures at the tissue level. Optical detection of these changes at the three levels has been pursued with different regimes of measurements to acquire light signals from the ultraviolet to the infrared regions. These measurement regimes are sketched in Fig. 1 according to the wavelength relations between the excitation and the measured light signals. The following sections are divided on the basis of these relations.

THE RADIATIVE TRANSFER METHOD

Response of a biological system such as cells or tissues to an incident light field can be categorized into different types, as shown in Fig. 1. In the spectral region from about 300 to 2500 nm wavelength, elastic scattering of incident light by various components of tissues including the cells is often the dominant form of interaction. The measurement of optical signals as a result of elastic scattering and absorption is denoted here as the radiative transfer method because of the modeling methodology to be described below.

The light absorption and scattering can be understood fundamentally through the theory of classical electrodynamics in terms of the local fluctuation of the dielectric coefficient $\epsilon(\mathbf{r}, \omega)$ or the related index of refraction $n(\mathbf{r}, \lambda_h) = [\epsilon(\mathbf{r}, \omega)]^{1/2}$ at the scales of intracellular organisms, where \mathbf{r} is the position vector and $\lambda_h = 2\pi v/\omega$ is the wavelength, v and ω are the speed and angular frequency of incident light in the host medium, respectively. To demonstrate this point quantitatively, consider the scattering of an incident light field ($\mathbf{E}_i, \mathbf{B}_i$) with a wave vector \mathbf{k}_i by a biological cell. The scattering configuration and the intracellular structure, given by $\epsilon(\mathbf{r}, \lambda_h)$, are schematically presented in Fig. 2. Using Maxwell's equations, the scattered light field ($\mathbf{E}_s, \mathbf{B}_s$) along the direction of the wave vector \mathbf{k} in the far-field region can be defined by an amplitude function $\mathbf{F}(\mathbf{k}, \mathbf{k}_0)$ as^[4]

$$\mathbf{E}_s(\mathbf{r}, \lambda_h) = \mathbf{F}(\mathbf{k}, \mathbf{k}_i) \frac{e^{i\mathbf{k}\cdot\mathbf{r}}}{r} \quad (1)$$

where \mathbf{k} and \mathbf{k}_i are related to the wavelength as $k = |\mathbf{k}| = |\mathbf{k}_i| = 2\pi/\lambda_h$. The amplitude function $\mathbf{F}(\mathbf{k}, \mathbf{k}_i)$

O

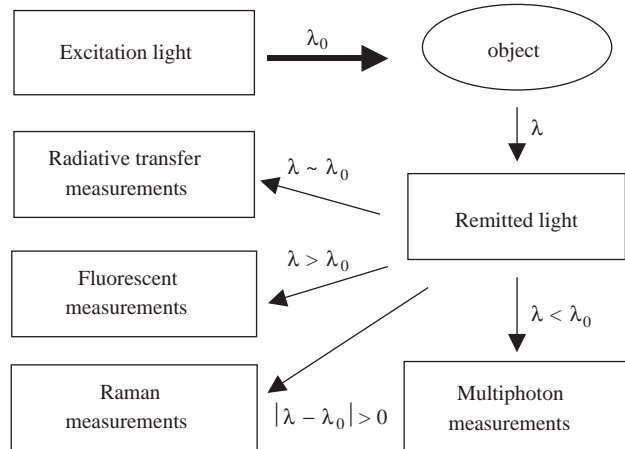


Fig. 1 Different approaches of optical measurements.

and the total field \mathbf{E} , the incident plus the scattered, are to be solved from the following equation:

$$\mathbf{F}(\mathbf{k}, \mathbf{k}_i) = \frac{1}{4\pi} \int_V (k^2 + \nabla \nabla \cdot) \times [\varepsilon(\mathbf{r}', \omega) - \varepsilon_h] \mathbf{E}(\mathbf{r}', \omega) e^{-j\mathbf{k} \cdot \mathbf{r}'} d^3r' \quad (2)$$

where V is the cellular volume in which $\varepsilon - \varepsilon_h$ is a spatial function of \mathbf{r} . These relations show that light scattering is caused by the local inhomogeneity in terms of the deviation of the dielectric coefficient $\varepsilon(\mathbf{r}', \lambda)$ from the background value ε_h of the host medium. Further, light absorption can be accounted for by using a complex dielectric coefficient or complex refractive index $n = n_r + jn_i$, where $j = \sqrt{-1}$ and n_i causes attenuation of the incident wave in addition to scattering. Although analytical solutions can be obtained for scatterer of regular geometric shapes and structures, numerical solutions are the only feasible methods for scatterers of irregular but realistic shapes and inhomogeneous internal structures.

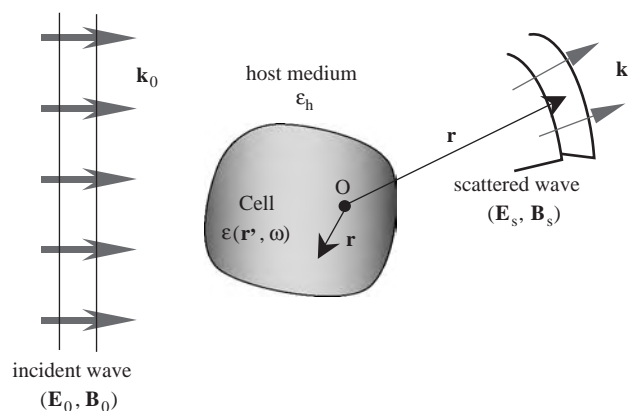


Fig. 2 Light scattering by a biological cell.

Biological tissues consist of many types of cells and extracellular components and usually possess no long-range orders or symmetry. The structural complexity of tissues renders the first-principle approach demonstrated in Eqs. (1) and (2) impractical. Even numerical solutions of tissue optics problems are impossible at the present time owing to the prohibitive computing cost. For modeling tissue optics problems, radiative transfer theory is a much preferred approach in which only the distribution of light radiance, $I(\mathbf{r}, \hat{s}, t)$, instead of the electromagnetic fields, is to be solved. The light radiance $I(\mathbf{r}, \hat{s}, t)$ is defined as the electromagnetic energy propagating along the \hat{s} direction per unit solid angle and per unit area at location \mathbf{r} and time t . The distribution of the radiance I can be obtained from a radiative transfer equation based on energy conservation:^[5]

$$\frac{1}{v} \frac{\partial I(\mathbf{r}, \hat{s}, t)}{\partial t} + \hat{s} \cdot \nabla I = -(\mu_s + \mu_a) I(\mathbf{r}, \hat{s}, t) + \mu_s \int_{4\pi} p(\hat{s}, \hat{s}') I(\mathbf{r}, \hat{s}', t) d\Omega' + S(\mathbf{r}, \hat{s}, t) \quad (3)$$

where \hat{s} is a unit vector of direction, μ_s and μ_a are the scattering and absorption coefficient, respectively, $p(\hat{s}, \hat{s}')$ is the scattering phase function, and S is the source term. In principle the optical parameters of tissue, μ_s , μ_a , and $p(\hat{s}, \hat{s}')$, can be derived from the scattering theory presented in Eq. (2) if the details of the optical inhomogeneity ε or n are completely known.^[6] It has recently been proved experimentally that the complex refractive index of microspheres, relative to the known refractive index of water as the host medium, can be determined accurately from optical measurements of the microsphere suspensions based on Eq. (3) and the Mie theory of light scattering by spheres.^[7] For complex turbid systems such as the biological tissues, it is nearly impossible to solve analytically the practical boundary problems described by Eq. (3). Instead, numerical solutions have to be resorted to and a statistical method of Monte Carlo simulation has been widely adopted for its algorithm simplicity and ease of implementation of the parallel computing technique.^[8,9]

If scattered light can be neglected or removed from the measured data using techniques such as spatial filtering, Eq. (3) can be solved for a source-free medium by simple integration to yield the modified Beer-Lambert law:

$$\frac{I(\mathbf{r}, \hat{s})}{I(\mathbf{r}_0, \hat{s})} = \exp \left\{ - \int_{\mathbf{r}_0}^{\mathbf{r}} \mu_t(\mathbf{r}') dr' \right\} \quad (4)$$

where we define the attenuation coefficient $\mu_t = \mu_s + \mu_a$, neglect the second term on the right-hand side of Eq.

(3) because it produced scattered light only, and consider only the steady-state solutions to drop the time variable. Eq. (4) yields a “shadowgram” principle to relate the projection image data in the transmission mode to the distribution of μ_t in the imaged object and forms the foundation for computed tomography (CT).^[10]

At the other extreme of approximation, the radiative transfer equation can be simplified to a diffusion equation of a “diffuse photon density wave,” represented by $\rho_d(r, t) = \int_{4\pi} I_d(\mathbf{r}, \hat{s}, t) d\Omega$, by considering only the portion of radiance that has been multiply scattered, i.e., I_d . Once the propagation directions of these photons are randomized by multiple scattering, the distribution of ρ_d is no longer related to the detailed form of the scattering process and depends only on the absorption coefficient μ_a and a reduced scattering coefficient $\mu_s' = \mu_s(1-g)$, where $g = \int_{4\pi} (\hat{s} \cdot \hat{s}') p(\hat{s}, -\hat{s}') d\Omega'$ is the first moment of the scattering phase function. The diffusion approximation breaks down in two cases. The first case arises when μ_a becomes appreciable in comparison with μ_s because the multiply scattered portion I_d is significantly reduced by absorption and thus may not dominate in the light distribution I . The second case is in the region near a light source where most incident photons do not yet experience multiple scattering.

Continuous-Wave Imaging Studies

One of the earliest attempts to detect pathological conditions of tissues has been reported by Cutler, who observed projection images of breasts with visible light in the transmission mode in a dark room.^[11] Projection imaging of the breast or diaphanography has been actively studied in the 1970s and 1980s by using both visible and near-infrared light sources with films or video cameras as the recording devices.^[12] These studies employed the shadowgram principle for image interpretation and often were compared with x-ray mammography.^[13] Unlike the case of x-rays in which the photons that undergo volumetric trajectories due to scattering can be effectively removed from the projection data, light scattering dominates over absorption in interaction of the visible or near-infrared photons with the breast tissues. Consequently, blurring is inherent in the optical projection images. Both contrast and spatial resolution of these images, for discriminating different types of breast tissues and lesions, have been shown to be poor in comparison with x-ray mammography and research interest has diminished.

For organs other than breast, transillumination imaging is not practical and imaging in the reflection mode is the only way to measure the optical response

from tissues noninvasively. Several low-magnification microscopes or endoscopes have been routinely used in clinics by physicians as the cancer screening tools to enhance their ability of visual inspection. These include dermoscopy in detection of skin cancers, colposcopy in cervical cancers, colonoscopy in colon cancers, and bronchoscopy in lung and airway cancers. Recently, multispectral imaging has been investigated for early diagnosis of dysplastic nevi and melanoma.^[14] Quantitative modeling of the projection data in the reflection mode to obtain tomographic images, however, is extremely difficult because these signals originate from the refractive index mismatch at the tissue surface and scattered photons in the bulk tissue. Researchers have followed two directions: 1) to extract patterns and statistical features from image data and/or spectroscopy data to correlate with the pathology for cancer detection and 2) to determine the local tissue optical parameters from multipoint measurements of the reflectance with a tissue optics model. These will be discussed in this and the following sections. Elbaum et al. acquired in vivo images from 246 pigmented skin lesions, ranging from benign nevi to dysplasia to melanoma, with 10 wavebands between 430 and 950 nm using a dermoscope.^[15] The multispectral images were processed with different algorithms of patterns analysis to extract about 25 image parameters in two groups. Comparison of the two-parameter groups as classifiers with the histopathology analysis demonstrated good sensitivity (100% or 95%) and relatively poor specificity (84% or 68%). A similar in vivo method for melanoma diagnosis has been investigated on 237 pigmented skin lesions and 80% sensitivity and 51% specificity (with lesion dimension considered) were reported.^[16]

Spectroscopy Measurements

For noninvasive in vivo detection of cancer, reflectance spectroscopy with a continuous-wave (cw) white light source has been the preferred mode of measurements by researchers for its instrumental simplicity. A technique of spatially resolved diffuse reflectance has been introduced to determine μ_a and μ_s' from the spatial dependence of reflectance signals. The spatially resolved signals are acquired with multiple fibers and fitted to calculated curves based on a diffusion model of a homogenous tissue.^[17] The wavelength dependence of the parameters μ_a and μ_s' can be obtained from the spectroscopy data to analyze the chromophores in the tissue. But clinical application of this spatially resolved technique is limited because the homogeneity should not be assumed in tissues with cancerous regions. As a result, published in vivo spectroscopy studies up to this time have combined only the



single-point detection with pattern analysis of data based on various statistical methods. Improvement in modeling accuracy with Monte Carlo simulations has been reported.^[18]

Mirabal et al. investigated reflectance spectroscopy in the regions from 390 to 590 nm and from 350 to 650 nm to evaluate its potential for diagnosis of cervical tissue conditions in patients with abnormal Pap smears.^[19,20] They analyzed the spectral data statistically to extract classifiers for correlating with pathological diagnosis and compared the correlation with that determined with 355-nm-excited fluorescence spectroscopy measurements. It was concluded that fluorescence spectroscopy provided better specificity and sensitivity in the detection of cervical intraepithelial neoplasia.^[19] In vivo measurements of the reflectance spectral data and statistical analysis have been performed in an attempt to detect breast cancers, lesions in ovarian tissues, skin lesions including dysplastic nevi and basal cell carcinomas, and adenomatous colon polyps.^[21–24] Recently, Johnson et al. reported the study of spectral reflectance data, between 340 and 900 nm, acquired from 139 cancerous and normal sentinel nodes removed from 68 patients with breast cancers.^[25] Statistical analysis of the spectral data yielded scores for discriminating different types of nodes with sensitivity and specificity comparable to those of frozen section histology.

In the above studies, the measured light signals included those photons that are scattered only once or a few times in the backward directions and others that are scattered multiple times. Detection of singly backscattered photons can be realized experimentally with a single fiber of small diameter of about 200 μm for both launch and detection of light. Because multiply scattered photons in the remitted light are most likely to migrate away from the fiber, the detected signal is dominated by singly backscattered photons from the superficial layer of the tissue.^[26] Georgakoudi et al. compared three types of in vivo spectral data to evaluate their potentials for detection of dysplasia in 16 patients with Barrett's esophagus: singly backscattered light; diffuse reflectance; and intrinsic fluorescence of multiple excitation wavelengths. It was found that singly backscattered light spectroscopy data provided the best overall sensitivity and specificity among the three techniques in separating different degrees of dysplasia.^[27]

Frequency-Domain Measurements

Intensity modulation of a cw laser beam at a megahertz frequency adds "signatures" to the beam in terms of amplitude demodulation M and phase shift ϕ . When such a beam is launched into tissue, the local optical

properties of the tissue cause variation in M and ϕ in the detected signals relative to the source. With a diffusion model, M and ϕ can be inverted to obtain tissue's optical parameters μ_a and μ_s' at selected wavelengths if multiple laser sources are used.^[28] Further, several fibers can be employed to launch and measure at different locations to reconstruct tomographic images with contrast built on M , ϕ , their spatial derivatives, or μ_a and μ_s' . An early study of one breast cancer patient using the frequency-domain method at 690 and 825 nm was reported by Fantini et al., who also determined μ_a and μ_s' for normal and tumor tissues.^[29] Their very limited in vivo result suggested that μ_a and μ_s' of tumor tissues are both larger than those of the normal tissues at both wavelengths. A similar method has been used to determine μ_a and μ_s' at the sites of lesions and normal tissues on one patient with lung carcinoma and subcutaneous tumors and 10 patients with cervical carcinoma in situ at four wavelengths from 674 to 956 nm.^[30,31] The results, however, exhibited significant overlap in the values of these parameters between the normal tissue sites and lesion sites at the four wavelengths. The frequency-domain method can be further improved to enhance the image contrast of the blood oxygenation by incorporating a spatial derivative algorithm.^[32] This technique was used for an in vivo study conducted on four patients with ductal carcinomas and fibroadenoma in their breasts. Transmission images were obtained at four wavelengths by scanning the source and detection fibers in tandem and image data yielded evidence of a correlation between the lesion locations, identified by x-ray mammography and histopathology, and those of abnormality in the oxygenation index images.^[32]

Time-Resolved Measurements

The significant blurring in the cw imaging data due to the dominant presence of multiply scattered light causes substantial deterioration in the contrast of images interpreted through the shadowgram principle. As ultrashort pulsed solid-state lasers and fast photodetectors became routinely available, research interest in shadowgram imaging for cancer detection has been revived since the early 1990s. Various time-gating methods have been devised to extract the early-arrival photons, which are either not scattered (ballistic photons) or only weakly scattered (snake photons).^[33] Optical parameters μ_a and μ_s' can be obtained with a diffusion model or Monte Carlo method from the time-resolved signals of reflectance or transmittance at the wavelength of excitation.^[34,35] An in vitro study of 20 excised breast samples consisting of either fat (lipid) or fibrous tissues with a streak camera of 2 psec time

resolution has shown that the mean value of μ_s' is about 0.5 mm^{-1} for lipid tissue and 0.9 mm^{-1} for fibrous tissue at the wavelength of 800 nm.^[36] Patients with breast cysts or tumors have been studied with the time-resolved technique to obtain transmission images by scanning the source fibers connected to a diode laser generating 400 psec pulses and a photomultiplier detector on the opposite side of the compressed breast.^[37,38] Even though lesions or tumors of a few centimeters in size can be identified from the images, with the prior knowledge from MRI and x-ray mammography, no significant reduction in blurring can be observed in the time-resolved images when compared with the cw images, which may be attributed to the low time resolution of the above optical systems.^[12]

Diffuse Optical Tomography

Unlike its x-ray counterpart, optical tomography remains an open problem despite intensive research efforts. The complexity of the problem arises from the fact that the measured light signals from a turbid object are dominated by photons with volumetric trajectories in the imaged region as a result of light scattering. Inverse algorithms for reconstructing 2D images of optical parameters from the projections have been investigated in phantoms with simple-shaped structures embedded. The existence of unique solutions in practical cases, however, is unknown.^[39] Most published studies on optical tomography employed the method of fiber-based multipoint detection in the frequency domain to obtain measured data, instead of the cw or time-resolved techniques, for instrumental simplicity and large dynamic range. In diffuse optical tomography (DOT), the diffusion model is used for forward calculations of the measured data from trial values of optical parameters based on simplified models of tissue heterogeneity. Even though DOT possesses the ability to exhibit the distribution of blood oxygen saturation in tissues, fast and robust inverse algorithms for image reconstruction remain to be developed. Existing inverse algorithms for DOT are iterative processes of least-squares type or of Bayesian type using an objective function and their solutions are practical only if the number of unknown optical parameters is kept small by either using a priori structural information as constraints or reducing the number of voxels (or spatial resolution).^[40] Consequently, recent research efforts have been concentrated on the development of DOT as an auxiliary modality to other high-resolution imaging modalities such as x-ray CT or MRI to enhance the functional imaging capacity.^[41] A preliminary in vivo study of a breast tumor with DOT demonstrated the overlapping of the abnormal

locations in the images of μ_a and μ_s' at multiple wavelengths with the biopsy identified tumor site.^[42]

Optical Coherence Tomography

Reconstruction of tomographic images of an object in optical coherence tomography (OCT) is accomplished in reflection mode by combining the single-point measurement using a converging light beam of broad spectral bandwidth with 2D or 3D scanning. The OCT signal is based on the light backscattered from the focal point and is separated from the background noise by interference of the signal beam with a reference beam. High spatial resolution can be achieved in the longitudinal (depth) direction, ranging from 2 to 20 μm , using a light source of short coherent length.^[43] The resolutions in the transverse directions relate to both the imaging optics and the size of the scanning step. The OCT image contrast depends on two factors: the attenuation of the incident beam along its trajectory toward the focal point and the local gradient of the refractive index at the focal point. As a result, OCT signals are attenuated exponentially as $\exp\{-2 \int_0^d \mu_t dr\}$ within the interrogated tissue, where d is the pathlength from tissue surface to the focal point.

Optical coherence tomography has been employed successfully for detection of intraocular tumors because μ_t for most ocular tissues is much less than 1 mm^{-1} in the visible and near-infrared regions.^[44] For turbid biological tissues such as the human skin and mucosa, μ_t typically ranges from 1 to 10 mm^{-1} and the penetration depth of OCT in these tissues is limited to about 1 mm or less.^[45] The small penetration depth in turbid tissues is the major factor limiting the application of OCT for cancer detection. In a recent study, two patients with diagnosis of invasive basal cell carcinoma in their skin have been imaged with a polarization sensitive OCT method to identify structural changes in the superficial layers of the sites.^[46] The OCT images showed variations in both tumor sites, which were attributed to the local changes in the refractive index due to the loss of normal tissue structure by the invasive growth of dense tumor stroma. An in vivo study of OCT in the detection of oral tumors has been conducted in an animal model of 35 hamsters with squamous cell carcinoma and it was found that in 80% of cases the OCT-based diagnosis correlates with those based on histopathology.^[47]

Optical coherence tomography can also be implemented in the endoscopic form for in vivo imaging of luminal cancers by passing a catheter-based fiber probe through the biopsy channel of an endoscope.^[48,49] Examination of upper gastrointestinal tracts has been performed with endoscopic OCT in 32 patients suspected of Barrett's esophagus and esophageal



adenocarcinoma, followed by the endoscopic mucosal biopsies.^[50] The OCT images were compared to the histopathology analysis of the biopsied tissues and it has been found that OCT images provide sufficient contrast to differentiate among the normal esophagus, Barrett's esophagus, and esophageal adenocarcinoma. A similar *in vivo* study of 19 patients with Barrett's esophagus and esophageal adenocarcinoma confirmed that the abnormal morphologic changes in the mucosa tissues can be identified from the OCT images.^[51]

THE FLUORESCENCE METHOD

Fluorescence signals from biological macromolecules depend on the structure of electronically excited states of these molecules. This relation allows the possibility to discriminate molecules in cancerous cells or tissues from those in the normal counterparts. The strong signals, assuming the existence of either intrinsic or extrinsic fluorophores, and the ability for molecular imaging or detection have made the fluorescence method a popular approach to probe biological systems. These advantages, however, can be significantly reduced owing to the overlapping of spectral features in the broadband fluorescence spectra from different biological fluorophores and the dependence of the spectral shape on the tissue optical parameters resulting from the transport of fluorescence light in the interrogated tissues. This typically leads to poor specificity as the barrier for cancer detection in clinical settings. Future research efforts have to be invested for translation of the fluorescence method into clinical applications with much improved signal-to-noise ratios. We review here only the measurements of fluorescence signals emitted by intrinsic fluorophores, which is also called autofluorescence.

Spectroscopy Measurements

Measurements of fluorescence signals are usually carried out with cw excitation at wavelengths λ_0 in the ultraviolet and blue regions because of large absorption cross sections and often, high-fluorescence quantum yields. The distinctive features of cw fluorescence signals are given by the excitation and emission spectra and associated quantum yields or signal intensities.

Alfano et al. compared the difference in fluorescence spectra by intrinsic fluorophores of normal and tumor tissues of animals and humans.^[52,53] The fluorescence spectra were obtained with an argon laser beam at 488 nm and power of up to 100 mW focused on a spot of about 0.1 mm in diameter. These fluorescence spectra were shown to be dominated by a broadly

peaked curve in the wavelength range from 500 to 650 nm with no significant structure. It has been suggested that the signal differences at selected wavelengths in the emission spectra between tissues pathologically classified as normal and cancerous may provide the means for cancer detection.^[52,53] These early spectral measurements did not provide sufficient evidences, however, for cancer detection because the contrast or signal-to-noise ratio was small and sample-to-sample fluctuation was not considered. Additional studies of fluorescence with ultraviolet excitation from 300 to 460 nm on breast tissues provided statistically significant results, which demonstrated that spectral features are very similar among the normal and tumor tissues but the fluorophore concentration and/or fluorescence efficiency may differ in certain spectral regions.^[54,55] Even though these researchers reported that features of emission spectra can be utilized in the form of ratios of fluorescence signals at two wavelengths to differentiate cancerous tissues from the normal breast tissues, the spectra presented in these results vary from each other with no explanation provided. The technique of fluorescence emission ratio has been used for *in vivo* investigation of bronchial tissues with excitation between 350 and 495 nm. After correcting the effects of probe-tissue distance and intensity variation of excitation, it has been found that the normal bronchial tissue exhibits slightly stronger emission than the dysplasia tissue with nearly identical spectral shapes, while those of the cancerous ones were much weaker than both tissues.^[56] Similar *in vivo* study of emission ratio differences has also been reported for distinguishing nasopharyngeal carcinoma and normal tissue.^[57]

The similarity in the shapes of emission spectra between the normal and cancerous tissues significantly reduces the reliability of this technique for cancer detection. Several methods have been investigated to address this problem by either statistical analysis of spectral data to decrease the noise or employment of multiwavelength excitation to increase the signal content. Multivariate analysis methods based on principal component analysis (PCA) and support vector machines were introduced to identify spectral features that can be correlated to the pathology condition of tissues.^[58,59] A method of excitation–emission matrix (EEM) with multiple excitation wavelengths was applied to colonic mucosa tissue analysis.^[60] For over 10 yr, a research group led by Richards-Kortum has studied the fluorescence detection of cervical cancer with methods of both EEM and PCA.^[61,62] Brewer et al. collected the *in vivo* emission spectra of ovarian tissues with 18 excitation wavelengths from 330 to 500 nm from patients suspected to have ovarian cancers.^[63] On the basis of EEM analysis, it has been suggested that certain spectral patterns in EEM plots

may be correlated to the pathological conditions of the ovarian tissues for cancer detection.

Time-Resolved Measurements

One valuable method to differentiate molecules with overlapping spectra and/or reduce the effect of tissue optical parameters on the spectral shapes of fluorescence signals is to determine the relaxation times of fluorescence signals with a time-resolved technique. Time-resolved measurements probe directly the relaxation of electronically excited fluorophores back to their ground state in the time domain. Typical lifetimes of transitions between the singlet excited and ground states are 1–10 nsec. Hence, subnanosecond light pulses from pulsed lasers such as nitrogen lasers are required for excitation and detectors with similar time resolution must be used. Single- or multi-component fitting with exponential decay curves can be used to extract relaxation times from the fluorescence data. An *in vitro* study of 18 normal, benign, and malignant human breast tissues has been performed with excitation at 310 nm and emission at 340 nm.^[33] By fitting the measured data with two-component exponential curves, a large difference was found in the ratio of the two decay components among the samples: the malignant tumor tissues have the largest portion of the slow component (with relaxation time about 2 nsec) among the three types. A similar assessment of fixed but unstained breast tissues from six patients with benign or malignancy-associated collagenous stroma has been recently reported with a fluorescence lifetime imaging system of 2 nsec time resolution and excitation at 415 nm.^[64] A seemingly opposite trend, however, has been observed in the above study: pixels of benign stroma show larger relaxation times than those of malignancy-associated stroma. Further studies are obviously needed.

OTHER METHODS

Owing to the limitation of space, we have to exclude certain optical methods from discussion, which nevertheless have the potential for cancer detection. Among these, three are briefly presented here: confocal microscopy, Raman spectroscopy, and multiphoton techniques. Confocal microscopes have become a standard imaging tool in biology laboratories since the early 1990s for reconstruction of 3D images of cells and other microscopic agents. Its optical principle is similar to that of OCT except for the collection of backscattered light through spatial filtering with a microscopic objective of large numerical aperture. Confocal images of turbid tissues typically exhibit contrasts superior to the OCT images.^[65] *In vivo* applications of confocal

imaging in clinics have been investigated mostly in dermatology for the detection of skin structures and lesions.^[66] In comparison with the very active research on radiative transfer and fluorescence methods, the study of Raman spectroscopy and multiphoton techniques received relatively less attention for *in vivo* optical detection of cancers in tissues. The major challenge relates to the detection of weak signals, which requires intense light sources and sensitive detectors. Lasers are necessary to meet the requirement of high irradiance for nonlinear measurements and both high irradiance and narrow frequency bandwidth for Raman measurements. These methods have been studied for cancer detection with Raman spectroscopy and multiphoton microscopy and for analysis of the skin dermis by the second-harmonic signals.^[67–69]

CONCLUSIONS

Compared with established diagnosis modalities, such as x-ray, CT, and MRI, noninvasive optical detection of cancers has been regarded to have considerable advantages, which include patient safety, high spatial resolution, and low equipment cost. These advantages have attracted intensive research efforts over the last two decades and substantial strides have been made toward modeling of light–tissue interactions and instrumentation development for data acquisition. Despite this progress, optical detection of cancers has not yet achieved the status of the established imaging modalities. The fundamental barrier impeding the translation of optical technology from the laboratory to wide applications in clinics, in our view, relates to the lack of efficient modeling tools and algorithms to invert the measured data from turbid human tissues to the spatial distribution of optical parameters that is inherently heterogenous at the scale of millimeters or lower. It is, therefore, hoped that this barrier can be significantly reduced in the near future with interdisciplinary efforts contributed by researchers from multiple fields including the biomedical, computational, engineering, and physical sciences.

ACKNOWLEDGMENT

We acknowledge the partial support from the National Institutes of Health through a grant (1R15GM70798-01).

ARTICLES OF FURTHER INTEREST

Laser-Tissue Interaction; Optical Coherence Tomography; Optical Mapping; Optics, Biomedical



REFERENCES

1. ACS. Cancer Facts & Figures 2005. American Cancer Society, 2005.
2. Lodish, H.F.; Berk, A.; Matsudaira, P.; Kaiser, C.A.; Krieger, M.; Scott, M.P.; Zipursky, S.L.; Darnell, J. *Molecular Cell Biology*; 5th Ed.; W.H. Freeman and Company: New York, 2003.
3. Cotran, R.S.; Kumar, V.; Collins, T.; Robbins, S.L. *Robbins' Pathologic Basis of Disease*; 6th Ed.; Saunders: Philadelphia, 1999.
4. Lu, J.Q.; Yang, P.; Hu, X.H. Simulations of light scattering from a biconcave red blood cell using the FDTD method. *J. Biomed. Opt.* **2005**, *10* (2) 024022.
5. Chandrasekhar, S. *Radiative Transfer*; Dover Publications: New York, 1960.
6. van de Hulst, H.C. *Light Scattering by Small Particles*; Wiley: New York, 1957.
7. Ma, X.; Lu, J.Q.; Brock, R.S.; Jacobs, K.M.; Yang, P.; Hu, X.H. Determination of complex refractive index of polystyrene microspheres from 370 to 1610 nm. *Phys. Med. Biol.* **2003**, *48*, 4165–4172.
8. Wilson, B.C.; Adam, G. A Monte Carlo model for the absorption and flux distributions of light in tissue. *Med. Phys.* **1983**, *10* (6), 824–830.
9. Li, K.; Lu, J.Q.; Brock, R.S.; Yang, B.; Hu, X.H. Quantitative modeling of skin images using parallel Monte Carlo methods. *SPIE Proc.* **2005**, *5693*, 82–87.
10. Kak, A.C.; Slaney, M. *Principles of Computerized Tomographic Imaging*; IEEE Press: New York, 1988.
11. Cutler, M. Transillumination as an aid in the diagnosis of breast lesions. *Surg. Gynecol. Obstet.* **1929**, *48* (6), 721–729.
12. D'Orsi, C.J.; Bartrum, R.J.; Moskowitz, M.M. Light scanning of the breast. In *Mammography, Thermography, and Ultrasound in Breast Cancer Detection*; 2nd Ed.; Gold, R.H., Ed.; Grune & Stratton: New York, 1987; 169–177.
13. Profio, A.E.; Navarro, G.A.; Sartorius, O.W. Scientific basis of breast diaphanography. *Med. Phys.* **1989**, *16* (1), 60–65.
14. Moncrieff, M.; Cotton, S.; Claridge, E.; Hall, P. Spectrophotometric intracutaneous analysis: a new technique for imaging pigmented skin lesions. *Br. J. Dermatol.* **2002**, *146* (3), 448–457.
15. Elbaum, M.; Kopf, A.W.; Rabinovitz, H.S.; Langley, R.G.; Kamino, H.; Mihm, M.C., Jr.; Sober, A.J.; Peck, G.L.; Bogdan, A.; Gutkowitz-Krusin, D.; Greenbaum, M.; Keem, S.; Oliviero, M.; Wang, S. Automatic differentiation of melanoma from melanocytic nevi with multispectral digital dermoscopy: a feasibility study. *J. Am. Acad. Dermatol.* **2001**, *44* (2), 207–218.
16. Farina, B.; Bartoli, C.; Bono, A.; Colombo, A.; Lualdi, M.; Tragni, G.; Marchesini, R. Multispectral imaging approach in the diagnosis of cutaneous melanoma: potentiality and limits. *Phys. Med.* **2000**, *45* (5), 1243–1254.
17. Kienle, A.; Lilge, L.; Patterson, M.S.; Hibst, R.; Steiner, R.; Wilson, B.C. Spatially resolved absolute diffuse reflectance measurements for noninvasive determination of the optical scattering and absorption coefficients of biological tissue. *Appl. Opt.* **1996**, *35*, 2304–2314.
18. Pfefer, T.J.; Matchette, L.S.; Bennett, C.L.; Gall, J.A.; Wilke, J.N.; Durkin, A.J.; Ediger, M.N. Reflectance-based determination of optical properties in highly attenuating tissue. *J. Biomed. Opt.* **2003**, *8* (2), 206–215.
19. Nordstrom, R.J.; Burke, L.; Niloff, J.M.; Myrtle, J.F. Identification of cervical intraepithelial neoplasia (CIN) using UV-excited fluorescence and diffuse-reflectance tissue spectroscopy. *Lasers Surg. Med.* **2001**, *29* (2), 118–127.
20. Mirabal, Y.N.; Chang, S.K.; Atkinson, E.N.; Malpica, A.; Follen, M.; Richards-Kortum, R. Reflectance spectroscopy for in vivo detection of cervical precancer. *J. Biomed. Opt.* **2002**, *7* (4), 587–594.
21. Bigio, I.J.; Bown, S.G.; Briggs, G.; Kelley, C.; Lakhani, S.; Pickard, D.; Ripley, P.M.; Rose, I.G.; Saunders, C. Diagnosis of breast cancer using elastic-scattering spectroscopy: preliminary clinical results. *J. Biomed. Opt.* **2000**, *5* (2), 221–228.
22. Utzinger, U.; Brewer, M.; Silva, E.; Gershenson, D.; Blast, R.C., Jr.; Follen, M.; Richards-Kortum, R. Reflectance spectroscopy for in vivo characterization of ovarian tissue. *Lasers Surg. Med.* **2001**, *28* (1), 56–66.
23. Garcia-Urbe, A.; Kehtarnavaz, N.; Marquez, G.; Prieto, V.; Duvic, M.; Wang, L.V. Skin cancer detection by spectroscopic oblique-incidence reflectometry: classification and physiological origins. *Appl. Opt.* **2004**, *43*, 2643–2650.
24. Zonios, G.; Perelman, L.T.; Backman, V.; Manoharan, R.; Fitzmaurice, M.; Van Dam, J.; Feld, M.S. Diffuse reflectance spectroscopy of human adenomatous colon polyps in vivo. *Appl. Opt.* **1999**, *38* (31), 6628–6637.
25. Johnson, K.S.; Chicken, D.W.; Pickard, D.C.; Lee, A.C.; Briggs, G.; Falzon, M.; Bigio, I.J.; Keshtgar, M.R.; Bown, S.G. Elastic scattering spectroscopy for intraoperative determination of sentinel lymph node status in the breast. *J. Biomed. Opt.* **2004**, *9* (6), 1122–1128.
26. Canpolat, M.; Mourant, J.R. Particle size analysis of turbid media with a single optical fiber in contact with the medium to deliver and detect white light. *Appl. Opt.* **2001**, *40* (22), 3792–3799.
27. Georgakoudi, I.; Jacobson, B.C.; Van Dam, J.; Backman, V.; Wallace, M.B.; Muller, M.G.; Zhang, Q.; Badizadegan, K.; Sun, D.; Thomas, G.A.; Perelman, L.T.; Feld, M.S. Fluorescence, reflectance, and light-scattering spectroscopy for evaluating dysplasia in patients with Barrett's esophagus. *Gastroenterology* **2001**, *120* (7), 1620–1629.
28. Madsen, S.J.; Anderson, E.R.; Haskell, R.C.; Tromberg, B. Portable, high-bandwidth frequency-domain photon migration instrument for tissue spectroscopy. *Opt. Lett.* **1994**, *19* (23), 1934–1936.
29. Fantini, S.; Walker, S.A.; Franceschini, M.A.; Kaschke, M.; Schlag, P.M.; Moesta, K.T. Assessment of the size, position, and optical properties of breast tumors in vivo by noninvasive optical methods. *Appl. Opt.* **1998**, *37* (10), 1982–1989.

30. Fishkin, J.B.; Coquoz, O.; Anderson, E.R.; Brenner, M.; Tromberg, B.J. Frequency-domain photon migration measurements of normal and malignant tissue optical properties in a human subject. *Appl. Opt.* **1997**, *36* (31), 10–20.
31. Hornung, R.; Pham, T.H.; Keefe, K.A.; Berns, M.W.; Tadir, Y.; Tromberg, B.J. Quantitative near-infrared spectroscopy of cervical dysplasia in vivo. *Hum. Reprod.* **1999**, *14* (11), 2908–2916.
32. Heffer, E.; Pera, V.; Schutz, O.; Siebold, H.; Fantini, S. Near-infrared imaging of the human breast: complementing hemoglobin concentration maps with oxygenation images. *J. Biomed. Opt.* **2004**, *9* (6), 1152–1160.
33. Das, B.B.; Liu, F.; Alfano, R.R. Time-resolved fluorescence and photon migration studies in biomedical and model random media. *Rep. Prog. Phys.* **1997**, *60*, 227–292.
34. Patterson, M.S.; Chance, B.; Wilson, B.C. Time resolved reflectance and transmittance for the noninvasive measurement of tissue optical properties. *Appl. Opt.* **1989**, *28* (12), 2331–2336.
35. Kienle, A.; Patterson, M.S. Determination of the optical properties of turbid media from a single Monte Carlo simulation. *Phys. Med. Biol.* **1996**, *41*, 2221–2227.
36. Zacharakis, G.; Zolindaki, A.; Sakkalis, V.; Filippidis, G.; Papazoglou, T.G.; Tsiftsis, D.D.; Koumantakis, E. In vitro optical characterization and discrimination of female breast tissue during near infrared femtosecond laser pulses propagation. *J. Biomed. Opt.* **2001**, *6* (4), 446–449.
37. Torricelli, A.; Spinelli, L.; Pifferi, A.; Taroni, P.; Cubeddu, R. Use of a nonlinear perturbation approach for in vivo breast lesion characterization by multi-wavelength time-resolved optical mammography. *Opt. Exp.* **2003**, *11* (8), 853–867.
38. Grosenick, D.; Wabnitz, H.; Moesta, K.T.; Mucke, J.; Moller, M.; Stroszczyński, C.; Stossel, J.; Wassermann, B.; Schlag, P.M.; Rinneberg, H. Concentration and oxygen saturation of haemoglobin of 50 breast tumours determined by time-domain optical mammography. *Phys. Med. Biol.* **2004**, *49* (7), 1165–1181.
39. Arridge, S.R. Optical tomography in medical imaging. *Inverse Problems* **1999**, *15*, R41–R93.
40. Arridge, S.R.; Hebden, J.C. Optical imaging in medicine: II. Modelling and reconstruction. *Phys. Med. Biol.* **1997**, *42* (5), 841–853.
41. Li, A.; Miller, E.L.; Kilmer, M.E.; Brukilacchio, T.J.; Chaves, T.; Stott, J.; Zhang, Q.; Wu, T.; Chorlton, M.; Moore, R.H.; Kopans, D.B.; Boas, D.A. Tomographic optical breast imaging guided by three-dimensional mammography. *Appl. Opt.* **2003**, *42* (25), 5181–5190.
42. Dehghani, H.; Pogue, B.W.; Poplack, S.P.; Paulsen, K.D. Multiwavelength three-dimensional near-infrared tomography of the breast: initial simulation, phantom, and clinical results. *Appl. Opt.* **2003**, *42* (1), 135–145.
43. Huang, D.; Swanson, E.A.; Lin, C.P.; Schuman, J.S.; Stinson, W.G.; Chang, W.; Hee, M.R.; Flotte, T.; Gregory, K.; Puliafito, C.A., et al. Optical coherence tomography. *Science* **1991**, *254* (5035), 1178–1181.
44. Shields, C.L.; Mashayekhi, A.; Luo, C.K.; Materin, M.A.; Shields, J.A. Optical coherence tomography in children: analysis of 44 eyes with intraocular tumors and simulating conditions. *J. Pediatr. Ophthalmol. Strabismus* **2004**, *41* (6), 338–344.
45. Knuttel, A.; Boehlau-Godau, M. Spatially confined and temporally resolved refractive index and scattering evaluation in human skin performed with optical coherence tomography. *J. Biomed. Opt.* **2000**, *5* (1), 83–92.
46. Strasswimmer, J.; Pierce, M.C.; Park, B.H.; Neel, V.; de Boer, J.F. Polarization-sensitive optical coherence tomography of invasive basal cell carcinoma. *J. Biomed. Opt.* **2004**, *9* (2), 292–298.
47. Wilder-Smith, P.; Jung, W.G.; Brenner, M.; Osann, K.; Beydoun, H.; Messadi, D.; Chen, Z. In vivo optical coherence tomography for the diagnosis of oral malignancy. *Lasers Surg. Med.* **2004**, *35* (4), 269–275.
48. Sergeev, A.M.; Gelikonov, V.M.; Gelikonov, G.V.; Feldchtein, F.I.; Kuranov, R.V.; Gladkova, N.D.; Shakhova, N.M.; Snopova, L.B.; Shakhov, A.V.; Kuznetzova, I.A.; Denisenko, A.N.; Pochinko, V.V.; Chumakov, Y.P.; Streltsova, O.S. In vivo endoscopic OCT imaging of precancer and cancer states of human mucosa. *Opt. Exp.* **1997**, *1* (13), 432–440.
49. Poneros, J.M.; Nishioka, N.S. Diagnosis of Barrett's esophagus using optical coherence tomography. *Gastrointest. Endosc. Clin. N. Am.* **2003**, *13* (2), 309–323.
50. Bouma, B.E.; Tearney, G.J.; Compton, C.C.; Nishioka, N.S. High-resolution imaging of the human esophagus and stomach in vivo using optical coherence tomography. *Gastrointest. Endosc.* **2000**, *51* (4 pt 1), 467–474.
51. Jackle, S.; Gladkova, N.; Feldchtein, F.; Terentjeva, A.; Brand, B.; Gelikonov, G.; Gelikonov, V.; Sergeev, A.; Fritscher-Ravens, A.; Freund, J.; Seitz, U.; Schroder, S.; Soehendra, N. In vivo endoscopic optical coherence tomography of esophagitis, Barrett's esophagus, and adenocarcinoma of the esophagus. *Endoscopy* **2000**, *32* (10), 750–755.
52. Alfano, R.R.; Tata, D.B.; Cordero, J.; Tomashefsky, P.; Longo, F.W.; Alfano, M.A. Laser induced fluorescence spectroscopy from native cancerous and normal tissue. *IEEE J. Quantum Electron.* **1984**, *20* (12), 1507–1511.
53. Alfano, R.R.; Tang, G.C.; Pradhan, A.; Lam, W.; Choy, D.S.J.; Opher, E. Fluorescence spectra from cancerous and normal human breast and lung tissues. *IEEE J. Quantum Electron.* **1987**, *23* (10), 1806–1811.
54. Gupta, P.K.; Majumder, S.K.; Uppal, A. Breast cancer diagnosis using N2 laser excited autofluorescence spectroscopy. *Lasers Surg. Med.* **1997**, *21* (5), 417–422.
55. Hage, R.; de la Riviere, A.B.; Seldenrijk, C.A.; van den Bosch, J.M. Update in pulmonary carcinoid tumors: a review article. *Ann. Surg. Oncol.* **2003**, *10* (6), 697–704.
56. Zellweger, M.; Goujon, D.; Conde, R.; Forrer, M.; van den Bergh, H.; Wagnieres, G. Absolute autofluorescence spectra of human healthy, metaplastic, and early cancerous bronchial tissue in vivo. *Appl. Opt.* **2001**, *40* (22), 3784–3791.
57. Qu, J.Y.; Wing, P.; Huang, Z.; Kwong, D.; Sham, J.; Lee, S.L.; Ho, W.K.; Wei, W.I. Preliminary study of in vivo autofluorescence of nasopharyngeal carcinoma



- and normal tissue. *Lasers Surg. Med.* **2000**, *26* (5), 432–440.
58. Ramanujam, N.; Mitchell, M.F.; Mahadevan, A.; Thomsen, S.; Malpica, A.; Wright, T.; Atkinson, N.; Richards-Kortum, R. Development of a multivariate statistical algorithm to analyze human cervical tissue fluorescence spectra acquired in vivo. *Lasers Surg. Med.* **1996**, *19* (1), 46–62.
 59. Lin, W.; Yuan, X.; Yuen, P.; Wei, W.I.; Sham, J.; Shi, P.; Qu, J. Classification of in vivo autofluorescence spectra using support vector machines. *J. Biomed. Opt.* **2004**, *9* (1), 180–186.
 60. Zangaro, R.A.; Silveira, L.J.; Manoharan, R.; Zonios, G.; Itzkan, I.; Dasari, R.R.; Van Dam, J.; Feld, M.S. Rapid multiexcitation fluorescence spectroscopy system for in vivo tissue diagnosis. *Appl. Opt.* **1996**, *35* (25), 5211–5219.
 61. Ramanujam, N.; Mitchell, M.F.; Mahadevan, A.; Warren, S.; Thomsen, S.; Silva, E.; Richards-Kortum, R. In vivo diagnosis of cervical intraepithelial neoplasia using 337-nm-excited laser-induced fluorescence. *Proc. Natl. Acad. Sci. USA* **1994**, *91* (21), 10193–10197.
 62. Nath, A.; Rivoire, K.; Chang, S.; West, L.; Cantor, S.B.; Basen-Engquist, K.; Adler-Storthz, K.; Cox, D.D.; Atkinson, E.N.; Staerckel, G.; MacAulay, C.; Richards-Kortum, R.; Follen, M. A pilot study for a screening trial of cervical fluorescence spectroscopy. *Int. J. Gynecol. Cancer* **2004**, *14* (6), 1097–1107.
 63. Brewer, M.; Utzinger, U.; Silva, E.; Gershenson, D.; Bast, R.C., Jr.; Follen, M.; Richards-Kortum, R. Fluorescence spectroscopy for in vivo characterization of ovarian tissue. *Lasers Surg. Med.* **2001**, *29* (2), 128–135.
 64. Tadrous, P.J.; Siegel, J.; French, P.M.; Shousha, S.; Lalani, el N.; Stamp, G.W. Fluorescence lifetime imaging of unstained tissues: early results in human breast cancer. *J. Pathol.* **2003**, *199* (3), 309–317.
 65. Neerken, S.; Lucassen, G.W.; Bisschop, M.A.; Lenderink, E.; Nuijs, T.A. Characterization of age-related effects in human skin: a comparative study that applies confocal laser scanning microscopy and optical coherence tomography. *J. Biomed. Opt.* **2004**, *9* (2), 274–281.
 66. Rajadhyaksha, M.; Gonzalez, S.; Zavislan, J.M.; Anderson, R.R.; Webb, R.H. In vivo confocal scanning laser microscopy of human skin II: advances in instrumentation and comparison with histology. *J. Invest. Dermatol.* **1999**, *113* (3), 293–303.
 67. Hanlon, E.B.; Manoharan, R.; Koo, T.W.; Shafer, K.E.; Motz, J.T.; Fitzmaurice, M.; Kramer, J.R.; Itzkan, I.; Dasari, R.R.; Feld, M.S. Prospects for in vivo Raman spectroscopy. *Phys. Med. Biol.* **2000**, *45* (2), R1–59.
 68. Skala, M.C.; Squirrell, J.M.; Vrotsos, K.M.; Eickhoff, J.C.; Gendron-Fitzpatrick, A.; Eliceiri, K.W.; Ramanujam, N. Multiphoton microscopy of endogenous fluorescence differentiates normal, precancerous, and cancerous squamous epithelial tissues. *Cancer Res.* **2005**, *65* (4), 1180–1186.
 69. Yasui, T.; Tohno, Y.; Araki, T. Characterization of collagen orientation in human dermis by two-dimensional second-harmonic-generation polarimetry. *J. Biomed. Opt.* **2004**, *9* (2), 259–264.

Optical Mapping

Igor R. Efimov
Vladimir P. Nikolski

Case Western Reserve University, Cleveland, Ohio, U.S.A.

Gil Bub

Physiology Department, McGill University, Montreal, Quebec, Canada

INTRODUCTION

Optical mapping of electrical activity with the voltage-sensitive dye-fluorescence technique in the excitable tissues such as heart and brain has become a common research tool. These techniques are especially powerful in studies of spatiotemporal organization of electrophysiological activity in complex multicellular preparations, cell cultures, and organs.

Optical recordings of transmembrane potentials can be performed in a wide range of spatial resolutions from the subcellular level to the whole heart. The response time of fast voltage-sensitive dyes lies in the microsecond range, and the temporal resolution of the technique can potentially exceed that of conventional microelectrode recordings. Spatial resolution is defined by optical magnification and ranges from micrometer to millimeter range. Progress in modern computer technology permits simultaneous optical recordings from multiple sites, providing high-resolution spatiotemporal maps of electrical activity, unachievable by electrode-based techniques. One of the pioneers of fluorescent methods in neurophysiology, B.M. Salzberg, wrote many years ago that voltage-sensitive dyes “could, we believe, provide a powerful new technique for measuring membrane potential in systems where, for reasons of scale, topology, or complexity, the use of electrodes is inconvenient or impossible”. Based on our current experience in cardiac electrophysiology Salzberg’s list needs to be extended to recordings of action potentials in the presence of external electric fields during stimulation and defibrillation, which were impossible with extracellular and intracellular electrodes.

While intracellular microelectrode technique still represents the gold standard for recording single-cell transmembrane action potentials, the chief disadvantage of this technique lies in the impossibility of maintaining stable recordings over longer periods of time from more than two or three sites, especially if the preparation is moving. Similarly, monophasic action potential recordings can be maintained only at a few sites for short periods of time. Thus, spatiotemporal mapping by either method requires the use of a roving probe, limiting the application of either method to the

study of periodic activation patterns. Extracellular contact and non-contact multielectrode arrays are excellent for recording activation maps from the heart *in vitro*, and they also remain the only techniques for *in vivo* mapping. Problems do, however, exist concerning the precise interpretation of the electrogram data, the determination of repolarization times being particularly unreliable. Thus, optical mapping is the only technique capable of recording high-resolution maps of cardiac repolarization and uninterrupted artifact-free registration of the transmembrane potentials during electrical stimulation and defibrillation.

BASICS OF FLUORESCENCE

The general mechanism underlying fluorescence is the absorption of photons of certain energy by a fluorescent compound molecule, which creates an excited unstable energy-rich state, $S1^*$. Then the energy of $S1^*$ is partially dissipated, resulting in relaxed excited state $S1$ from which fluorescence emission originates. When the molecule falls back to the ground state $S0$, it emits a photon of a lower energy than the exciting photon. The wavelength of a photon being a function of its energy, the emitted light always has a longer wavelength than the exciting light, the so-called Stokes’ shift.

So called fast voltage-sensitive dyes that are used in cardiac electrophysiology have response time in the range of microseconds. The mechanisms of voltage-sensitivity for fast dyes are not fully understood. According to the electrochromic theory, the spectral shift is voltage dependent if the photon excitation of the chromophore molecule is accompanied by a large shift in electronic charge distribution in the direction of the electric field. An alternative hypothesis is the solvatochromic mechanism, which is related to electric-field-induced reorientation of the dye molecule.

INSTRUMENTATION DESIGN

Fluorophore

Designers of voltage-sensitive dyes had to solve several problems: 1) to find a chromophore that is capable of



producing the largest movements of charges during quantum transition from the ground to the excited state; 2) to ensure the possibility of delivering lipophilic dye molecules to the cellular membrane; 3) to ensure the proper orientation of the dye molecule perpendicular to the membrane; 4) to maximize duration of stay of the dye molecules in the desired position; 5) to minimize photobleaching of the dye; and 6) to minimize side effects of the dye on the preparation in the presence and absence of light.

Among several types of fluorophore, styryl dyes represent the most popular family of fast dyes, RH-421, di-4-ANEPPS, and di-8-ANEPPS being the most important members of this family. The spectroscopic properties of these dyes have been shown to change linearly with membrane potential changes in normal physiological range of transmembrane voltages.^[1] The orientation of the molecules of these dyes in the cell membrane is ensured by the presence of lipophilic and hydrophilic groups at opposite ends of the molecule.

The stability of the position of dye molecule in the membrane can be improved by increasing the length of the hydrocarbon tails, as was done in the ANEPPS family (di-4-ANEPPS, di-8-ANEPPS, di-12-ANEPPS, nomenclature described by Loew).^[2] In single-layer cell-culture preparations, Rohr et al. demonstrated a significantly improved stability of optical action potentials in preparations stained with di-8-ANEPPS as compared with di-4-ANEPPS on continuous illumination for 60 seconds, presumably on the basis of the slower translocation of di-8-ANEPPS into the cell.^[3] However, di-8-ANEPPS is less water soluble; this necessitates the use of surfactants, such as pluronic acid, which could have toxic side effects. Due to difficulties of staining we failed to confirm superiority of di-8-ANEPPS in Langendorff-perfused and superfused heart preparations as compared to di-4-ANEPPS. In our opinion, di-4-ANEPPS remains the dye of choice for whole-heart and tissue preparations.

The signal-to-noise ratio in stained cardiac preparations is dependent not only on the dye itself, but also on its mode of delivery. For example, one needs to consider slow diffusion of lipophilic dyes through the layers of the heart during staining by superfusion.^[4] Staining by coronary perfusion provides more uniform dye delivery. However, the depth of optical recordings in preparations stained by perfusion is still debatable, ranging from 300 μm to 1–2 mm.

Light Sources

Four types of excitation light sources are used in optical recordings: tungsten lamps, arc lamps, laser, and light-emitting diodes (LED). The choice of a light

source depends on the required spatiotemporal resolution. Arc lamps and lasers are typically used in micrometer scale measurements; tungsten lamps are most commonly used in macroscopic preparations, where a resolution of hundreds of microns is sufficient. Emerging LED-based-light-source technology is used in microscopic cell culture studies, in whole heart experiments and as components of a fiber optical “optrode” used in microscopic and macroscopic settings.

Lasers can provide intense illumination, which can be easily and rapidly delivered to a small spot using computer-controlled scanner. However, lasers have significant instability of beam intensity, which is comparable to the average fractional signal intensity recorded from most voltage-sensitive dyes.^[5] Ratio-calculation feedback signal processing techniques have been applied during recordings to eliminate signal noise related to laser light intensity variability.^[5] In addition, lasers are more expensive as compared to alternative sources.

While the intensities of arc lamps have been reported to exceed 50–100 times that of tungsten lamps, this large difference in intensity cannot be directly translated into better signal-to-noise ratios, as arc lamps have a significant intensity only at distinct narrow lines of the spectrum. Unfortunately, the excitation frequency of the most efficient dyes, which produce the largest fractional change in fluorescence, such as di-4-ANEPPS and RH421, do not overlap with spectral lines of arc lamps. Since the absorption peaks of most of the voltage-sensitive dyes are relatively wide, the lower light output of tungsten lamps at a given spectral line can be compensated for by the use of wider band-pass excitation filters.

Tungsten lamps remains the most popular light source used for optical recordings. They provide a very stable output of light over a very wide range of the spectrum without the sharp peaks observed in the arc lamp’s spectra. Cost of tungsten lamps is lower compared with laser and arc lamp sources.

Light-emitting diodes were recently introduced for excitation of the voltage-sensitive dye by Kodama et al.^[6] Advances in technology brought to the market a variety of ultra bright LEDs for different wavelengths. The most suitable set from CREE has 460-nm 4–10-mW, 470-nm 10-mW, 505-nm 8-mW, and 527-nm 7-mW peak intensity wavelengths and power. This technology appears to be very promising.

Light Collection

Four engineering solutions have been used over the last two decades in brain and heart electrophysiology for optical multiple-channel recording systems: photodiode arrays (PDA),^[7] laser scanner systems,^[8] charged

coupled device (CCD) cameras,^[9] and optrode technology^[10] based on embedding illumination and detection into a single fiber, which allows custom assembling of photodetection arrays in an arbitrary form. These methods differ not only by their way of collecting the fluorescent output of the specimen but also by the way the excitation light is delivered.

CCD cameras were introduced into optical mapping to achieve higher spatial resolutions and avoid the complications of setting up PDA systems. However the accuracy of mapping data with CCD cameras is limited by a number of factors: 1) The signal-to-noise ratio in signals from CCD cameras is usually poor and requires spatial averaging, defeating the advantage of high pixel resolution. 2) The time resolution of fast CCD cameras is approaching 1 ms, but, it is accompanied by the need for sophisticated CCD-cooling systems and still yields inferior signal quality. 3) Many cameras do not allow subtraction of the background fluorescence, and if they do, it is a uniform offset potential subtracted from all channels. 4) Amplitude resolution is usually not more than 12 bits. In general, CCD camera-based optical imaging is a popular solution due to ease of setting up and low cost.

Laser scanning systems represent an entirely different approach.^[14] The output of a single laser is acoustico-optically deflected to scan some 100 sites of the whole preparation. The fluorescence emitted by each site at the time of illumination is collected by a single photodetector, which thus sequentially records the optical signals from all sites scanned by the laser beam. The advantage of laser scanning systems over PDAs is that they can cover a wider area and that flatness of the preparation is not an optical requirement. However, the time resolution is limited to some 1 ms per 100 scanning sites and considerable photobleaching by intense light beam results in a significant beat-to-beat decrease of the fluorescence level.

PDA have been used for optical mapping studies in neurophysiology and cardiology since 1981.^[7] With a quantum efficiency above 0.8, photodiodes are the most sensitive sensors for medium-to-large light intensities, their main drawback being the size of their dark current, which may limit their usefulness at very low light intensities, as in neurophysiological applications. Photodiodes are packaged in arrays of 100, 144, 256, or more. Each recording channel can be independently signal-conditioned before analog-to-digital conversion.

Novel approach to emission detection with fiber-optic custom geometry arrays^[3] in contrast to the commercially available photodiode array detectors with predefined layout has the following advantages: 1) the spatial resolution is not limited by the detector geometry; 2) the array can be set to a specific preparation shape; 3) the light can be collected from several

optical ports, enabling multiemission measurements; and 4) the type of sensing photodiode could be selected to suit the particular emission characteristics.

ARTIFACTS AND PROBLEMS

Pharmacological Effects of Voltage-Sensitive Dyes

The application of voltage-sensitive dyes has been demonstrated to cause side effects on preparations. Phototoxic effects have been documented under intense illumination, with alterations of electric activity both in neurons and in isolated cardiac myocytes. In the isolated cardiac myocytes, 1 min. of illumination with 1 W/cm² will first cause a gradual depolarization of the membrane resting potential and a decay of the action potential amplitude. This is followed by after-depolarizations, occasionally triggered activity, and finally cell death. Recently, we have shown that voltage-sensitive dye RH421 increases the contractility of isolated rat cardiac cells and Langendorff-perfused hearts. Similar effects were observed during staining with di-4-ANEPPS and di-8-ANEPPS. Vasoconstriction is another pharmacological side effect of potential-sensitive styryl dyes. This effect was described for the dyes RH-414 and di-4-ANEPPS.

Mechanical Motion Artifact

Faithful optical recordings of cardiac action potentials require immobilization of the preparation. Movement artifacts in optical action potentials tend to be most pronounced during the action potential plateau and during the action potential downstroke, whereas action potential upstrokes are generally well preserved. Movement artifacts can occur due to physical movement of the tissue from the field of view of one detector element into the field of view of a neighboring element leading to artifactual recordings of transmembrane potentials from different cells in different phases of the action potential or with different degree of staining. This effect is most pronounced at the edges of the preparation. Action potentials can also be distorted due to the modulation of light scattering by mechanical contraction, a phenomenon observed in monolayer cell cultures even in the absence of gross cell movement across the photodetector array.

There are three basic approaches to reduce motion artifacts: 1) mechanical immobilization, 2) use of motion-artifact-insensitive signal analysis algorithms, and 3) pharmacological immobilization.

Preparations can be mechanically immobilized by restricting the movement of the muscle between the

walls of the tissue chamber and/or pressure pads or by stretching the tissue. While mechanical immobilization is free of pharmacological side effects, this method can be used only in preparations with moderate amplitude of contractions; otherwise, the pressure needed to eliminate movements can cause ischemia. This technique has been successfully applied to guinea pig heart preparations.^[11] Yet, mechanical immobilization does not fully eliminate motion artifacts in all channels even in a guinea pig heart and special signal processing techniques^[11] should be used to measure action potential duration.

Motion artifacts can be reduced by another signal processing technique that is based on ratiometry—simultaneous recordings of optical signals at two different excitation or emission wavelengths with subsequent division of the two signals.^[12]

Pharmacological methods of producing mechanically quiescent preparations may significantly affect electrical activity. Several methods are being employed: perfusion with low calcium,^[13] Ca-channel blockers,^[14] 2,3-butanedione monoxime (BDM),^[9] and, most recently, cytochalasin D (cyto D).^[15] At the present time, low calcium solutions and calcium-channel blockers are rarely used in optical mapping because of their effect on crucial calcium-dependent cellular processes. BDM has effects on a variety of channels and gap junctions affecting the action potential duration. A decrease of action potential duration in rabbit, sheep, guinea pig, and dog hearts and an increase in the rat hearts were observed. Cyto D, an actin filament disrupter, has recently been shown to block contraction without affecting action potential shape or duration in a number of species.^[15] Thus, Cyto D is a promising new excitation-contraction uncoupling agent for optical mapping studies of cardiac repolarization. However, high cost of cyto D is a concern.

Noise

There are three major sources of noise present in optical recordings: shot noise, dark noise, and extraneous noise.

Shot noise is the natural fluctuation in the number of photons detected by a photodetector caused by the quantum-statistical nature of photon emission and detection. Thus, shot noise occurs even in the presence of an ideal noise-free light source and ideal noise-free detector and cannot be eliminated. Shot noise is estimated by the root mean square deviation of the number of photons hitting a photodetector per unit time and is equal to the square root of their number. The number of photons detected by the photodetector will finally be reduced by the quantum efficiency of the photodetector, which is defined as the number of photoelectrons per photon. Of the two types of

photodetectors used for optical recordings, photodiodes and photocathodes, photodiodes have quantum efficiencies of 0.8–1.0 and the photocathodes of only 0.15. Thus, photodiodes have a significant advantage over photocathodes, potentially yielding a signal-to-noise ratio that is nearly six times higher.

Dark noise is the noise signal emitted by a photodetector in the absence of light. Photodiodes tend to have much higher dark noise than photocathodes, which, despite their much higher quantum efficiency, limits their utility at low light intensities such as in single-cell measurements, in which case the dark noise can be comparable to or even higher than the shot noise. Thus, photocathodes may yield higher signal-to-noise ratios than photodiodes at very low light levels, and a direct experimental comparison between the two detectors is required in each case.

Extraneous noise is caused by noise sources in the laboratory environment. The following measures serve to cut extraneous noise down to acceptable levels: using a dark or DC-light-illuminated room to eliminate stray light from noisy sources, an anti-vibration table to isolate the setup from mechanical vibration generated in the building and the rest of the experimental setup, a Faraday cage to reduce radiofrequency noise, grounding equipment to a common isolated ground to eliminate 50- or 60-Hz noise picked up by ground loops, using low-noise light source, and isolating power supplies and amplifiers from computing equipment.

PHOTODIODE ARRAY OPTICAL SYSTEM DESIGN

The core of the optical mapping system (Fig. 1) is the Hamamatsu C4675 detector, which combines a 16 × 16-element photodiode array (chip size 17.45 × 17.45 mm² with 256 square 0.95 × 0.95-mm² photodiodes spaced 0.15 mm apart) with 256 current-to-voltage converters in a single compact (136 × 136 × 154 mm) enclosure. The optical system is built around a central beamsplitter cube (Oriol Corp.) bearing the illumination system, a bellows apparatus with the imaging lens, and a ground glass screen and reticule for focusing the imaging optics. The bellows, beamsplitter cube, and photodetector are all mounted on an optical rail. The excitation light is produced by a 250-W quartz tungsten halogen lamp (Oriol Corp.) powered by a low-noise direct current (DC) power supply (Oriol Corp.). After filtering out the infrared part of the spectrum, the light is controlled by an electronic shutter (Oriol Corp.), which opens for only a few seconds during each scan. The light beam is made quasimonochromatic by passing it through a 520 ± 45-nm interference filter (Omega

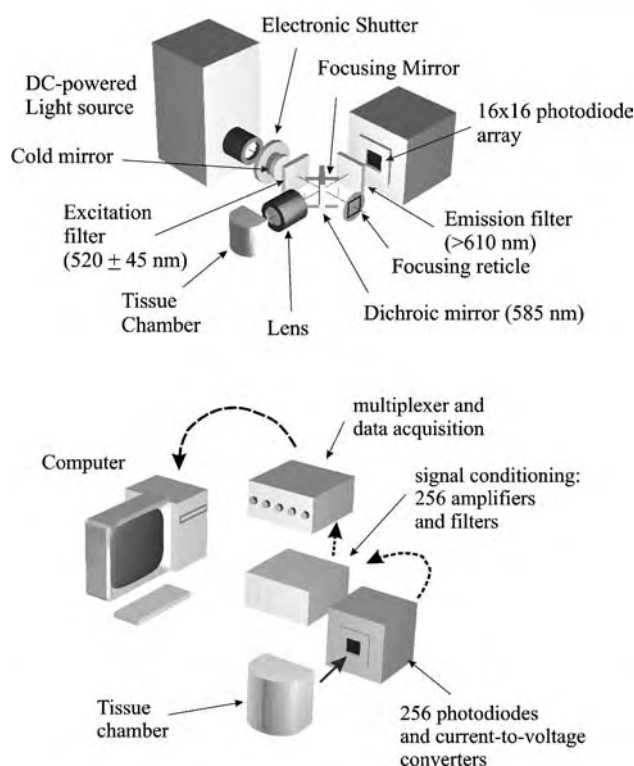


Fig. 1 Experimental setup. See text for details.

Optical), which corresponds to excitation wavelength of the dye di-4-ANEPPS. A 585-nm dichroic mirror (Omega Optical) held in the beamsplitter cube deflects the excitation light into the Nikon photographic imaging lens (50 mm f/1.2 Nikkor), which then focuses the excitation light onto the preparation. Then, the fluorescent light emitted from the preparation passes the dichroic mirror without reflection and is filtered by a long-pass filter (> 610 nm, Schott) before hitting the sensing area of the 16×16 photodiode array.

The photocurrent produced by each photodiode is first-stage amplified by its own low-noise operational amplifier inside the compact housing of the Hamamatsu C4675 detector (feedback resistors 10–100 M Ω , resulting in a gain of 10^7 – 10^8 V/A). Increasing the first stage amplification improves signal-to-noise ratios in optical signals. However, increasing the feedback resistors in the C4675 camera reduced its frequency response from 15 kHz to 1.5 kHz, which may be undesirable in some applications. The outputs of the first stage amplifiers were connected to 256 second-stage amplifiers, four 64-channel cards available from Innovative Technology, Inc., which offer computer-controlled DC coupling and AC coupling with several time constants (short time constant to remove the DC offset of the optical signals caused by background fluorescence, time constant of 30 s during data acquisition). Signals were filtered by Bessel filters with a cutoff frequency of 500–2000 Hz, depending

on the sampling rate of the data acquisition system in order to avoid aliasing errors. The signals were fed into four A/D converter boards (PCI 6033E, National Instruments). Each frame included 256 optical channels and eight instrumentation channels including surface electrogram, stimulation and defibrillation triggers, and aortic pressure, which were stored on a hard disk for off-line analysis.

An example of a raw optical signal, recorded with the described photodiode array system, is shown in Fig. 2A. Cellular depolarization during an action potential causes a reduction in fluorescence of 1 to 10% of the total fluorescence signal. Background fluorescence is caused by accumulation of dye in non-excitable cells, in the inner layer of the lipid membrane, or the lipid membranes of intracellular organelles across which there is no potential change during excitation. To reconstruct the intracellular action potential it is therefore necessary: 1) to subtract the background fluorescence, 2) to invert the signal and 3) to normalize the signals to uniform amplitudes assuming a homogeneity of action potential amplitudes in the imaged

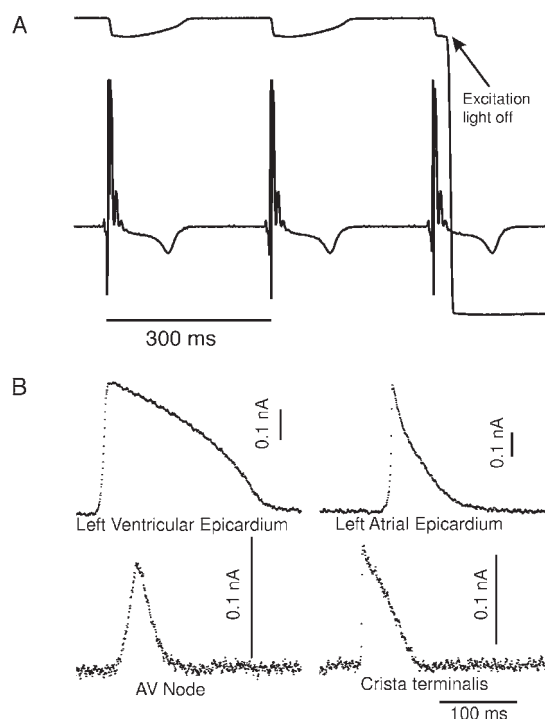


Fig. 2 Optical recordings of action potentials. A: An epicardial fluorescence recording from a rabbit Langendorff-perfused heart stained with di-4-ANEPPS (top trace) is juxtaposed with a simultaneous bipolar electrogram from the apex of the same heart. The light was turned off before the end of the recording to demonstrate the amount of background fluorescence in fluorescence signals without DC-offset subtraction. B: Photodiode optical current traces induced by the fluorescence from different heart tissues.

area of the preparation. In Fig. 2B one can see the examples of inverted optical signals detected from different areas of the heart.

EXPERIMENTAL PREPARATIONS

Experiments can be performed *in vitro* on Langendorff-perfused whole heart preparations, and isolated superfused or coronary-perfused atrial and ventricular preparations in rabbits, dogs, guinea pigs, rats, and humans, as well as in isolated cells and tissue cell cultures.

For staining, a stock solution of 5 mg di-4-ANEPPS (Molecular Probes) is prepared in 4 ml Dimethyl Sulfoxide (DMSO, Fisher Scientific) and is stored frozen at +4°C. After re-warming immediately before the experiment, it is added to the perfusate at final concentration of 0.5 $\mu\text{mol/L}$ for 5 min. during perfusion or for 40 min. during superperfusion. The method of gradual hand-injection of dye into the injection port in the bubble-trap is less cumbersome but less reproducible. Recordings started 5–15 minutes following the staining procedure.

Levels of optical signals and signal-to-noises ratios (SNR) decrease over time, presumably due to the translocation of the voltage-sensitive dye molecules into the cell interior with a half-life time about $\tau = 100 \pm 30$ min. Re-staining could be done once or twice during experiment, if needed.

Organ Mapping: Two-Dimensional Submillisecond Mapping of Transmembrane Polarization During Defibrillation Shocks in Intact Heart

Optical recordings of electrical activity are the only methods that resolve questions related to the

interaction of strong electric fields with excitable cells, conventional electrode recordings being distorted by large-amplitude artifacts.^[16,17] Using fluorescent mapping techniques, we have been able to record changes in transmembrane voltage optically with a high signal-to-noise ratio from all 256 channels. This has permitted us to acquire two-dimensional patterns of shock-induced polarization with submillisecond resolution.

Fig. 3D presents an example of optically recorded changes in transmembrane potential before, during, and after monophasic shock, which was applied at the plateau phase of a propagated action potential in rabbit heart. Fig. 3A shows the preparation photograph with a field of view of the optical mapping system. Analysis of data from all 256 channels allows reconstructing the surface map of transmembrane polarization at the end of the shock (Fig. 3B), as well as the activation map of the subsequent arrhythmia genesis (Fig. 3C). Only optical experimental techniques allow one to link together several important parameters that are responsible for arrhythmogenesis during shock-induced arrhythmia or failed defibrillation shocks: shock induced dispersion of transmembrane polarization, followed by dispersion of repolarization, and new wavefront generation, followed by formation of phase singularities and reentry.^[18]

Isolated Tissues: Conduction from the SA Node to and Through the AV Node

We recently demonstrated a nonradial spread of activation from the SA node to the AV node in isolated preparations of the rabbit right atrium.^[19] The structural complexity of the right atrial preparation, especially in the nodal areas, makes it difficult to

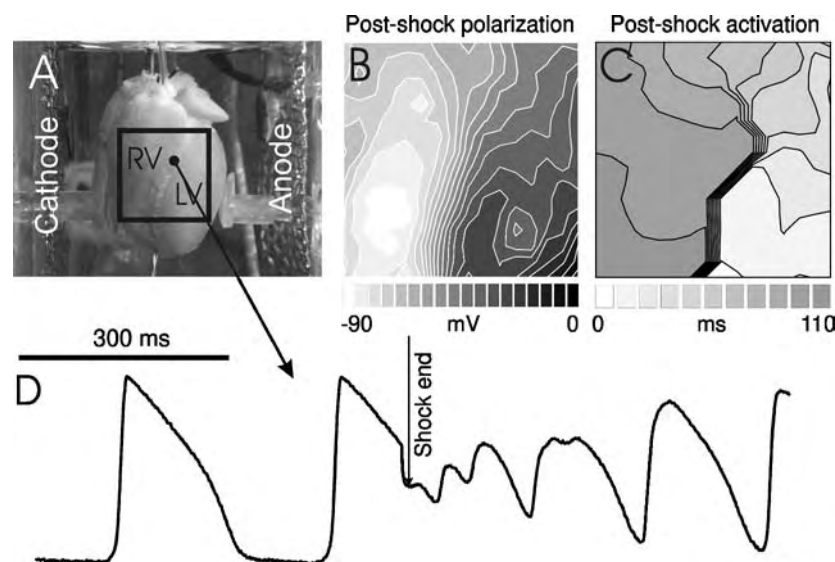


Fig. 3 A: Langendorff-perfused rabbit heart preparation. An optical mapping field of view is marked with a black square. B: map of the epicardial transmembrane polarization detected at the end of 8-ms, 100-V monophasic electric shock applied during plateau phase. C: Activation map of the first shock-induced reentrant beat constructed from the optical recordings. D: example of optical recording at a single optical channel (see marked position at the preparation photograph). It contains optical action potential upstroke of a propagated response during epicardial apex pacing and an arrhythmogenic response induced during the next action potential.

apply conventional electrode mapping techniques. While in unipolar electrograms the distinction between local and distant events is often impossible, bipolar electrograms represent complex spatial derivatives of the underlying spatiotemporal substrate of electrical activity.^[20] Since the precise field of view of both types of electrodes is unknown, interpretation of the electrograms remains challenging and requires verification by means of simultaneous microelectrode recordings.

Like electrograms, optical recordings represent the integrated electrical response of cells from a volume of tissue, the two-dimensional extent of which is precisely defined, even though the exact depth of the recording is less certain. We have previously suggested that double-component optical action potentials observed in proximal AV node and area of the bundle of His may be due to spatial averaging of electrical activity from several layers.^[19] Using microelectrode recordings we were able to demonstrate that optical recordings from the proximal part of the AV nodal area carry inscriptions of both superficial layers of transitional cells and a deeper layer of compact nodal and NH regions.^[21] Fig. 4 illustrates this finding by juxtaposition of two optical traces with conventional bipolar surface electrograms recorded from the interatrial septum and the

bundle of His during two basic and a single premature beats. Activation rapidly spread from the stimulation site on the interatrial septum toward the AV nodal area and finally conducted at the bundle of His. The optical trace recorded from the distal node shows multicomponent responses. Whereas the first component (T) in every beat represents activation of superficial transitional cells, in which conduction velocity is fast, the second component (NH) denotes activation of intranodal structures of the NH-region (NH). The increased delay between the first and the second components during conduction of the premature beat represents slowing of the intranodal conduction, which is also indicated by the His electrogram. It should be mentioned that relative amplitudes of the two components in the optical trace do not represent differences in action potential amplitudes, but rather the relative contribution of different amounts of tissues from transitional and NH regions, and different relative amplitudes of the two components can be seen in different areas of the proximal node.^[21]

Mapping Conduction in Cardiac Cell Culture

Heart cell monolayers are thin layers of tissue, grown in culture dishes from embryonic or neonatal cardiac cells, that have the capacity to form gap junctional connections with neighboring cells in culture and support propagating waves of excitation over long distances. Cardiac monolayers were popular twenty years ago as models of two-dimensional conduction. Optical potential mapping techniques have renewed interest in cultured monolayers, allowing the study of intercellular and cellular electrophysiology on microscopic and macroscopic scales with several advantages when compared to intact tissue. Signals from individual cells can be identified and mapped with sub-cellular resolution in cell culture while optical signals from intact tissue are compound signals from cells at the surface and deeper intramural layers. The geometry of cell cultures can be precisely controlled,^[22] allowing conduction patterns to be observed independently of the influence of micro- and macro-anatomy. Cardiac monolayers can be grown in conditions that cause cells to have similar anisotropy, connectivity, and longitudinality to transverse cell diameter ratios to cells in intact tissue. As a result of these advantages, several fundamental experiments have been performed on cell culture systems that could not easily be addressed in whole-tissue studies.

Propagation in cardiac muscle depends on both excitable and passive electrical properties of the coupled cell network. The passive properties of the tissue are influenced by factors such as discontinuities

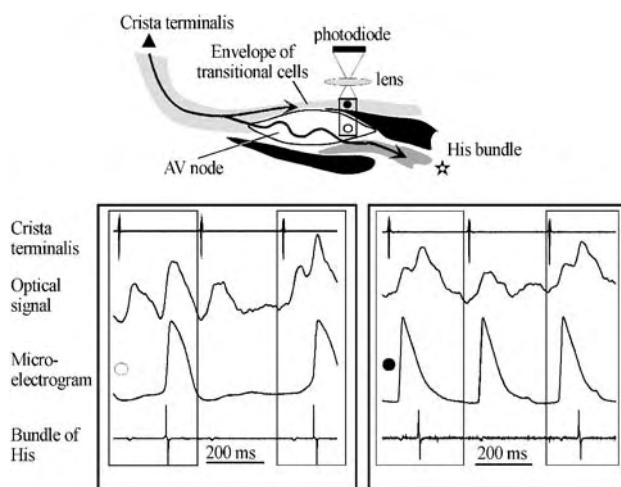


Fig. 4 Three-dimensional (multilayer) imaging of the conduction through the atrioventricular (AV) node. The right atrium was opened through the tricuspid valve, right atrial appendage, and superior vena cava to expose crista terminalis (CrT), intra-atrial septum (IAS), AV node; and the bundle of His. The preparation was paced at IAS with a cycle length S1S1 = 280 ms, (left panel) and 274 ms (right panel), which caused partial AV conduction block. Bipolar electrograms were recorded from CrT (triangle) and bundle of His (star). Optical traces represent the superimposition of signals from the septum and the distal AV nodal areas (see rectangle on the preparation schematic). Places of microelectrode recordings are marked with black and white dots.



introduced by capillaries and connective tissue, variations in cell connectivity, and tissue anisotropy. The local structure dictates whether or not waves propagate, block, or form reentrant circuits. Control of cell orientation by brushing the collagen substrate, and cell position using photolithographic techniques, allows studying the effects of geometry on propagation.

Optical mapping of propagation in monolayers plated as strands of different widths connected to a large tissue area demonstrated that impulses propagating into the site of expansion are blocked if the strand diameter is less than 200 μm .^[23] The instability at the expansion site may be modulated by electrophysiological characteristics: rapid pacing, which reduces excitability by lowering $I(\text{Na})$, results in conduction-block in wider strands, while gap junction-block, which lowers the number of cells directly driven by cells at the expansion, permits propagation through narrower strands. Tissue geometry may also act to stabilize propagation. Propagation through a culture plated as a main strand with many side branches is slowed due to the current load imposed by the branches and is stable at speeds down to 1 cm/sec,^[23] explaining how diseased myocardium, which is a complex patchwork of healthy and dead tissue, supports stable reentry at small space scales.

Theoretical studies have demonstrated that depolarizing and hyperpolarizing regions develop at resistive barriers in the presence of a large extracellular defibrillation field, initiating waves that terminate the arrhythmia. Whether such secondary sources occur at each cell-cell boundary or at larger anatomical structures is still debatable. Optical mapping experiments in cell culture indicate that secondary sources do not form at cell boundaries in densely packed strands of cultured myocytes, but are observed at the boundaries of cell strands and at larger intracellular clefts^[24] increasing as a function of cleft length and electrical field gradient. Additional studies with monolayers without clefts but with bifurcations, abrupt curves, and sudden expansions reveal that secondary sources form at the site of geometric nonlinearities. These results suggest that the target of defibrillatory extracellular shocks is heterogeneous or anatomically complex myocardium.

Monolayers of cardiac cells monitored at macroscopic scales (~ 1 cm) over long time periods (> 10 s) have been used to monitor activation patterns associated with arrhythmia. Depending on culturing conditions, monolayers of rhythmically beating cells display periodic target waves, stable spiral waves, or spiral waves that spontaneously initiate and terminate^[25] (Fig. 5), providing a model of stable and paroxysmal reentrant tachycardias in diseased tissue. A model of reentry in quiescent, excitable tissue was developed where waves propagating around an anatomical obstacle were terminated using extracellular

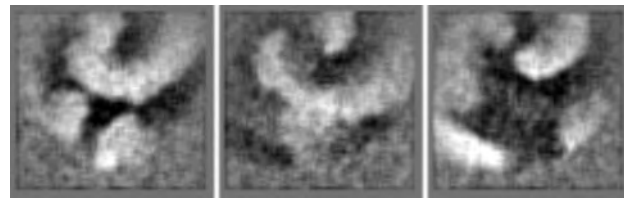


Fig. 5 Reentrant activity in a monolayer culture of chick cardiomyocytes. The image was obtained using calcium-sensitive dye (Calcium Green) and a cooled CCD camera (Roper Scientific, EEV 576 \times 384-frame transfer CCD). Each panel shows calcium fluorescence from a 1-square centimeter region of tissue. Panels are separated in time by 150 ms.

shocks.^[26] Ectopic arrhythmias were investigated in a system where a small area of the monolayer was transiently exposed to a solution that induced ischemialike response. Washout of ischemia inducing solution resulted in arrhythmic episodes lasting up to 200 s originating from ectopic sources located at the border of the ischemic zone.^[27]

CONCLUSION

Fluorescent imaging methods developed and refined in many laboratories have proven to be valuable tools extending the frontiers in cardiac electrophysiology.^[28] We believe that recent technological advances will make these experimental techniques available for a larger number of research institutions. However, future development of the fast fluorescent mapping will have to solve several limitations and challenges.

A major limitation of current techniques of potential imaging is the lack of absolute calibration. Unlike many ratio-metric fluorescent probes for calcium imaging, voltage-sensitive dyes can provide only relative information about changes in transmembrane voltage. Although the changes in the absolute amount of fluorescence excited at one wavelength linearly depends on transmembrane voltage of the viewed cells, accurate calibration has so far been impossible because the number of cells contributing to the signal is unknown. Monatana et al. demonstrated that measurements of fluorescence ratios excited at two wavelengths can provide such information.^[29] They have shown that the ratio between di-4-ANEPPS fluorescence levels excited at 440 and 505 nm linearly depends upon transmembrane potential in a lipid vesicle in the range from -125 mV to $+125$ mV. Similarly, the ratio of di-8-ANEPPS fluorescence from N1E-115 neuroblastoma cell excited at 450 and 530 nm linearly depends on transmembrane voltage. Similar efforts are now being made to measure absolute changes in membrane potentials in the heart.

Another area of application is the simultaneous recording of voltage and other intra- and extra-cellular characteristics using combined staining with voltage-sensitive dyes and other fluorescent probes. We demonstrated that voltage and cytosolic calcium can be measured from the same Langendorff-perfused guinea pig heart stained with the voltage-sensitive dye RH421 and the $[Ca^{2+}]$ -sensitive dye Rhod-2, which have close excitation frequencies, but different emission spectra.^[30] Similar imaging studies of multiple fluorescent probes have been presented by several laboratories, involving different voltage- and calcium-sensitive dyes. In the future, such multispectral measurements can be done with a novel liquid crystal or standing-wave, fast-switching filter technology.

ARTICLES OF FURTHER INTEREST

Cardiac Bioelectricity; Optics, Biomedical

REFERENCES

- Ross, W.N.; Salzberg, B.M.; Cohen, L.B.; Grinvald, A.; Davila, H.V.; Waggoner, A.S.; Wang, C.H. Changes in absorption, fluorescence, dichroism, and birefringence in stained giant axons: Optical measurement of membrane potential. *J. Membr. Biol.* **1977**, *33* (1–2), 141–183.
- Loew, L.M. Voltage-sensitive dyes: Measurement of membrane potentials induced by DC and AC electric fields. [Review]. *Bioelectromagn. Suppl.* **1992**, *1*, 179–189.
- Rohr, S.; Salzberg, B.M. Multiple site optical recording of transmembrane voltage (MSORTV) in patterned growth heart cell cultures: Assessing electrical behavior, with microsecond resolution, on a cellular and subcellular scale. *Biophys. J.* **1994**, *67* (3), 1301–1315.
- Nikolski, V.; Efimov, I. Fluorescent imaging of a dual-pathway atrioventricular-nodal conduction system. *Circ. Res.* **2001**, *88* (3), E23–E30.
- Bullen, A.; Patel, S.S.; Saggau, P. High-speed, random-access fluorescence microscopy: I. High-resolution optical recording with voltage-sensitive dyes and ion indicators. *Biophys. J.* **1997**, *73* (1), 477–491.
- Kodama, I.; Sakuma, I.; Shibata, N.; Honjo, H.; Toyama, J. Arrhythmogenic changes in action potential configuration in the ventricle induced by DC shocks. *J. Electrocardiol.* **1999**, *32* (Suppl), 92–99.
- Grinvald, A.; Cohen, L.B.; Leshner, S.; Boyle, M.B. Simultaneous optical monitoring of activity of many neurons in invertebrate ganglia using a 124-element photodiode array. *J. Neurophysiol.* **1981**, *45* (5), 829–840.
- Dillon, S.; Morad, M. A new laser scanning system for measuring action potential propagation in the heart. *Science* **1981**, *214* (4519), 453–456.
- Davidenko, J.M.; Pertsov, A.V.; Salomonsz, R.; Baxter, W.; Jalife, J. Stationary and drifting spiral waves of excitation in isolated cardiac muscle. *Nature* **1992**, *355* (6358), 349–351.
- Neunlist, M.; Zou, S.Z.; Tung, L. Design and use of an “optrode” for optical recordings of cardiac action potentials. *Pflugers Archiv. Eur. J. Phys.* **1992**, *420* (5–6), 611–617.
- Efimov, I.R.; Huang, D.T.; Rendt, J.M.; Salama, G. Optical mapping of repolarization and refractoriness from intact hearts. *Circulation* **1994**, *90* (3), 1469–1480.
- Knisley, S.B.; Justice, R.K.; Kong, W.; Johnson, P.L. Ratiometry of transmembrane voltage-sensitive fluorescent dye emission in hearts. *Am. J. Physiol.* **2000**, *279*, H1421–H1433.
- Salama, G.; Morad, M. Merocyanine 540 as an optical probe of transmembrane electrical activity in the heart. *Science* **1976**, *191* (4226), 485–487.
- Dillon, S.M. Optical recordings in the rabbit heart show that defibrillation strength shocks prolong the duration of depolarization and the refractory period. *Circ. Res.* **1991**, *69* (3), 842–856.
- Biermann, M.; Rubart, M.; Wu, J.; Moreno, A.; Josiah-Durant, A.; Zipes, D.P. Effects of cytochalasin D and 2,3-butanedione monoxime on isometric twitch force and transmembrane action potentials in isolated canine right ventricular trabecular fibers. *J. Cardiovasc. Electrophysiol.* **1998**, *9* (12), 1348–1357.
- Franz, M.R. Method and theory of monophasic action potential recording. *Prog. Cardiovasc. Dis.* **1991**, *33* (6), 347–368.
- Berbari, E.J.; Lander, P.; Geselowitz, D.B.; Scherlag, B.J.; Lazzara, R. The Methodology of Cardiac Mapping. In *Cardiac Mapping*; Shenasa, M., Borggreffe, M., Breithardt, G., Eds.; Futura Publishing: Mount Kisco, NY, 1993.
- Efimov, I.R.; Cheng, Y.; Van Wagoner, D.R.; Mazgalev, T.; Tchou, P.J. Virtual electrode-induced phase singularity: A basic mechanism of failure to defibrillate. *Circ. Res.* **1998**, *82*, 918–925.
- Efimov, I.R.; Fahy, G.J.; Cheng, Y.N.; Van Wagoner, D.R.; Tchou, P.J.; Mazgalev, T.N. High resolution fluorescent imaging of rabbit heart does not reveal a distinct atrioventricular nodal anterior input channel (fast pathway) during sinus rhythm. *J. Cardiovasc. Electrophysiol.* **1997**, *8*, 295–306.
- Steinhaus, B.M. Estimating cardiac transmembrane activation and recovery times from unipolar and bipolar extracellular electrograms: A simulation study. *Circ. Res.* **1989**, *64* (3), 449–462.
- Efimov, I.R.; Mazgalev, T.N. High-resolution three-dimensional fluorescent imaging reveals multilayer conduction pattern in the atrioventricular node. *Circulation* **1998**, *98*, 54–57.
- Rohr, S.; Scholly, D.M.; Kleber, A.G. Patterned growth of neonatal rat heart cells in culture. Morphological and electrophysiological characterization. *Circ. Res.* **1991**, *68* (1), 114–130.
- Rohr, S.; Kleber, A.G.; Kucera, J.P. Optical recording of impulse propagation in designer cultures. *Cardiac*

- tissue architectures inducing ultra-slow conduction. *Trends Cardiovasc. Med.* **1999**, *9* (7), 173–179.
24. Fast, V.G.; Rohr, S.; Gillis, A.M.; Kleber, A.G. Activation of cardiac tissue by extracellular electrical shocks: Formation of 'secondary sources' at intercellular clefts in monolayers of cultured myocytes. *Circ. Res.* **1998**, *82* (3), 375–385.
 25. Bub, G.; Glass, L.; Publicover, N.G.; Shrier, A. Bursting calcium rotors in cultured cardiac myocyte monolayers. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95* (17), 10283–10287.
 26. Entcheva, E.; Lu, S.N.; Troppman, R.H.; Sharma, V.; Tung, L. Contact fluorescence imaging of reentry in monolayers of cultured neonatal rat ventricular myocytes. *J. Cardiovasc. Electrophysiol.* **2000**, *11* (6), 665–676.
 27. Arutunyan, A.; Swift, L.M.; Sarvazyan, N. Initiation and propagation of ectopic waves: Insights from an in vitro model of ischemia-reperfusion injury. *Am. J. Physiol, Heart Circ. Physiol.* **2002**, *283* (2), H741–H749.
 28. Cohen, L.B.; Lesher, S. Optical Monitoring of Membrane Potential: Methods of Multisite Optical Measurement. In *Optical Methods in Cell Physiology*; De Weer, P., Salzberg, B.M., Eds.; Wiley-Interscience, 1986.
 29. Montana, V.; Farkas, D.L.; Loew, L.M. Dual-wavelength ratiometric fluorescence measurements of membrane potential. *Biochemistry* **1989**, *28* (11), 4536–4539.
 30. Efimov, I.R.; Rendt, J.M.; Salama, G. Optical maps of intracellular $[Ca^{2+}]_i$ transients and action potentials from the surface of perfused guinea pig hearts. *Circulation* **1994**, *90* (II), 632.

Optics, Biomedical

Sergio Fantini

Department of Biomedical Engineering, Tufts University, Medford, Massachusetts, U.S.A.

INTRODUCTION

Biomedical optics is a multidisciplinary field that embraces the studies of light propagation in biological samples, processes of light/tissue interaction, and applications of optical techniques in biology and medicine. Application areas of biomedical optics range from medical therapy, monitoring, and diagnostics, to microscopy studies of biological samples. Here we describe the fundamental properties of light, the optical properties of tissue, the basic mechanisms of light-tissue interaction (and their role in therapeutic applications of light), and a physical model that describes light propagation in tissue. We also describe a number of optical spectroscopy and imaging approaches that find applications in medical monitoring and diagnostics and in the study of biological samples at a tissue, cellular, and subcellular level.

NATURE OF LIGHT

Wave Description of Light

Light is one form of electromagnetic radiation, which is specified in terms of two vectors, the electric field \mathbf{E} and the magnetic field \mathbf{H} . Both vectors obey the wave equation:

$$\nabla^2 \mathbf{E} = \frac{\mu_r \varepsilon_r}{c_0^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} \quad (1)$$

and

$$\nabla^2 \mathbf{H} = \frac{\mu_r \varepsilon_r}{c_0^2} \frac{\partial^2 \mathbf{H}}{\partial t^2} \quad (2)$$

where ∇^2 is the Laplacian operator, μ_r and ε_r are the permeability and permittivity of the medium, respectively, relative to those of vacuum (so that μ_r and ε_r of vacuum are 1, by definition), and c_0 is the speed of light in vacuum (2.998×10^8 m/s). Eqs. 1 and 2 imply that light propagates with a speed $c = c_0 / \sqrt{\mu_r \varepsilon_r}$ in a medium characterized by relative permeability μ_r and relative permittivity ε_r . For simplicity let's consider a one-dimensional case, where the fields \mathbf{E} and \mathbf{H} vary in

only one direction, say z . Then, it can be shown that the following functions of z (spatial coordinate) and t (time):

$$E_x = E_{x0} \cos \left[2\pi \left(\frac{z}{\lambda} - ft \right) + \phi_x \right] \quad (3)$$

and

$$E_y = E_{y0} \cos \left[2\pi \left(\frac{z}{\lambda} - ft \right) + \phi_y \right] \quad (4)$$

are solutions to the wave equation for the x and y components of \mathbf{E} , respectively, and similar cosinusoidal functions are solutions for the H_x and H_y components of the magnetic field. Here, E_{x0} (E_{y0}) is the amplitude of the E_x (E_y) oscillations, λ is the wavelength, f is the frequency, and ϕ_x and ϕ_y are arbitrary phase shifts. The solutions to the wave equation provided by Eqs. 3 and 4 represent the particular case of a plane harmonic wave, which is important in optics and contains the key parameters used to describe the wave nature of light. The wavelength and the frequency are not independent, as their product is equal to the speed of light propagation in the medium ($\lambda f = c$). Therefore, the wavelength and the frequency can be used interchangeably to describe the color of light.

In biomedical optics the choice of wavelength depends on the particular application. A wide range of wavelengths is used, extending over the ultraviolet (UVC: 100–290 nm; UVB: 290–320 nm; UVA: 320–400 nm), visible (400–700 nm), near-infrared (IR) (0.7 to ~ 5 μm), mid-IR (~ 5 to ~ 40 μm), and far-IR (~ 40 to ~ 300 μm) spectral regions. For example, the fluorescence excitation and emission of chromophores found in proteins and nucleic acids (tryptophan, tyrosine, phenylalanine, adenine, guanine, cytosine, uracil, etc.) are typically in the UV region of the spectrum; proteins containing prosthetic groups such as hemes, flavins, etc., present absorption bands (associated with electronic transitions) from the near-UV to the near-IR, and vibrational spectroscopy of molecules is performed at wavelengths in the mid- to far-IR.

In the wave description of light, the energy density associated with the electric and magnetic fields are $0.5 \varepsilon_r \varepsilon_0 \mathbf{E}^2$ and $0.5 \mu_r \mu_0 \mathbf{H}^2$, respectively, where $\varepsilon_0 = 8.854 \times 10^{-12}$ F/m is the permittivity of vacuum and $\mu_0 = 4\pi \times 10^{-7}$ H/m is the permeability of vacuum.

The intensity of the wave (I), i.e., its energy per unit time and unit area, is given by $c\epsilon_r\epsilon_0\mathbf{E}^2$ or, equivalently, $c\mu_r\mu_0\mathbf{H}^2$. The wave description of light lends itself to the treatment of interference and diffraction phenomena. In fact, both phenomena can be described in terms of the superposition of individual waves among which there is a well-defined phase relationship.

Polarization is also typically explained using the wave description of light. The phase relationship between the x and y components of the electric field (i.e., the difference between ϕ_x in Eq. 3 and ϕ_y in Eq. 4) determines the time evolution of the amplitude and direction of the electric field. If $\phi_x - \phi_y$ is zero or 180° , then the light is linearly polarized. If $\phi_x - \phi_y = \pm 90^\circ$ and $E_{x0} = E_{y0}$, the light is said to be circularly polarized, whereas if $\phi_x - \phi_y = \pm 90^\circ$ and $E_{x0} \neq E_{y0}$, the light is said to be elliptically polarized. For values of $\phi_x - \phi_y$ that are not 0 , $\pm 90^\circ$, or 180° , light is also elliptically polarized.

Particle Description of Light

Quantum theory has led to the description of light in terms of quanta of energy, or photons, which behave as particles with a given energy and momentum. The relevance of this representation is in the fact that the exchange of energy between light and matter can occur only through discrete amounts, the photons. The energy associated with each photon depends on the frequency of the light and is equal to hf , where $h = 6.626 \times 10^{-34}$ J s is Planck's constant, and the momentum of a photon (along the direction of propagation) is h/λ . Using the particle description, light propagation in tissue can be represented in terms of a collection of individual photon trajectories. In this picture, the energy density is given by the number of photons per unit volume U multiplied by the energy per photon hf .

The use of intensity-modulated light sources in the optical study of tissue induces time oscillations in the photon density, and the propagation of these temporal oscillations of energy density can be described in terms of photon-density waves.^[1] Photon-density waves differ significantly from electromagnetic radiation waves in that they are strongly damped over a fraction of a wavelength, so that they are intrinsically confined to the near field. Nevertheless, photon-density waves in tissue exhibit refraction,^[2] interference,^[3] and diffraction^[4] phenomena similar to those of optical waves in the far field.

The momentum of a photon changes direction every time the photon is scattered or is refracted at an interface between two media. This change in momentum is associated with a force that is used to trap small particles or cells by laser tweezers.

Radiation Parameters

Using the particle description of light, the following parameters describing the radiation field as a function of position (\mathbf{r}), direction of propagation ($\hat{\Omega}$), and time (t) are commonly used in biomedical optics. These definitions can be extended to describe radiant energy (instead of photon number) by replacing the word "photon" with "energy," and by introducing a factor hf in all definitions (where hf is the energy per photon, of which h is Planck's constant and f is light frequency). A complete nomenclature for quantities used in biomedical optics can be found in Ref. [5].

- Angular photon density is represented by $u(\mathbf{r}, \hat{\Omega}, t)$, with units of $\text{m}^{-3}\text{sr}^{-1}$. $u(\mathbf{r}, \hat{\Omega}, t)d\mathbf{r}d\hat{\Omega}$ represents the number of photons in $d\mathbf{r}$ that travel in a direction within $d\hat{\Omega}$ around $\hat{\Omega}$.
- Photon radiance is represented by $L(\mathbf{r}, \hat{\Omega}, t) = \nu u(\mathbf{r}, \hat{\Omega}, t)$, with units of $\text{s}^{-1}\text{m}^{-2}\text{sr}^{-1}$. $L(\mathbf{r}, \hat{\Omega}, t)d\hat{\Omega}$ represents the number of photons traveling per unit time per unit area (perpendicular to $\hat{\Omega}$) in a range of directions within $d\hat{\Omega}$ around $\hat{\Omega}$.
- Photon density is represented by $U(\mathbf{r}, t) = \int_{4\pi} u(\mathbf{r}, \hat{\Omega}, t)d\hat{\Omega}$, with units of m^{-3} . The photon density is the number of photons per unit volume.
- Photon fluence rate is represented by $E_0(\mathbf{r}, t) = \nu U(\mathbf{r}, t) = \int_{4\pi} L(\mathbf{r}, \hat{\Omega}, t)d\hat{\Omega}$, with units of $\text{s}^{-1}\text{m}^{-2}$. The photon fluence rate is defined as the number of photons traveling per unit time per unit area (perpendicular to the direction of propagation) over all directions.
- Photon current density, or photon flux, is represented by $\mathbf{J}(\mathbf{r}, t) = \int_{4\pi} L(\mathbf{r}, \hat{\Omega}, t)\hat{\Omega}d\hat{\Omega}$, with units of $\text{s}^{-1}\text{m}^{-2}$. The photon flux is a vector that represents the net flow of photons. Its direction points in the direction of the net flux, and its amplitude gives the net number of photons transmitted per unit time per unit area in that direction.

OPTICAL PROPERTIES OF BIOLOGICAL TISSUE

Refractive Index

The refractive index of a medium (n) is defined as $n = \sqrt{\mu_r\epsilon_r}$, so that the refractive index of vacuum is 1. As a result of this definition, the speed of light in a medium of refractive index n is given by $c = c_0/n$. Usually, the refractive index depends on the frequency of light, typically increasing with frequency. Consequently, red light travels faster than blue light. In isotropic media, there are no preferential orientations and the refractive index is independent of the orientation of the electric field. However, anisotropic tissues

may present different indices of refraction for different directions of the electric field. This phenomenon is called birefringence. In birefringent media, the refractive index, and therefore the speed of light propagation, is different for different directions of polarization of linear polarized light.

Absorption Coefficient

Absorption processes consist of the transfer of energy from the light beam to the medium. In the UV, visible, and near-infrared portions of the spectrum, this typically involves electronic transition in tissue macromolecules and pigments, whereas in the mid- and far-IR, vibrational states are involved. Absorption results in an attenuation of the light intensity according to the Beer-Lambert law:

$$I = I_0 e^{-\mu_a L} \quad (5)$$

where I_0 is the incident intensity, μ_a is the absorption coefficient, and L is the distance traveled by the light in the medium. Using the particle description of light, the absorption coefficient μ_a can be defined as the inverse of the average distance traveled by a photon in the medium before being absorbed. The absorption coefficient for a given chromophore is given by the product of its concentration times its molar extinction coefficient. The absorption spectra of three important absorbers in tissue, namely, oxyhemoglobin, deoxyhemoglobin, and water, are shown in Fig. 1. These spectra were obtained from compiled absorption data for water^[6] and hemoglobin.^[7]

Scattering Coefficient

When light travels through tissue, it can be scattered elastically (i.e., without a change in wavelength) or inelastically (i.e., with a change in wavelength, such as in Brillouin or Raman scattering). The definition of the scattering coefficient refers to elastic scattering. Elastic scattering causes a change in the direction of propagation of light and thus an attenuation in the light intensity along the optical axis. Such attenuation is described by the following equation, which is formally identical to the absorption Beer-Lambert law of Eq. 5:

$$I = I_0 e^{-\mu_s L} \quad (6)$$

where μ_s is the scattering coefficient. Similar to the absorption coefficient, μ_s can be defined as the inverse of the average distance traveled by the photons before being scattered.

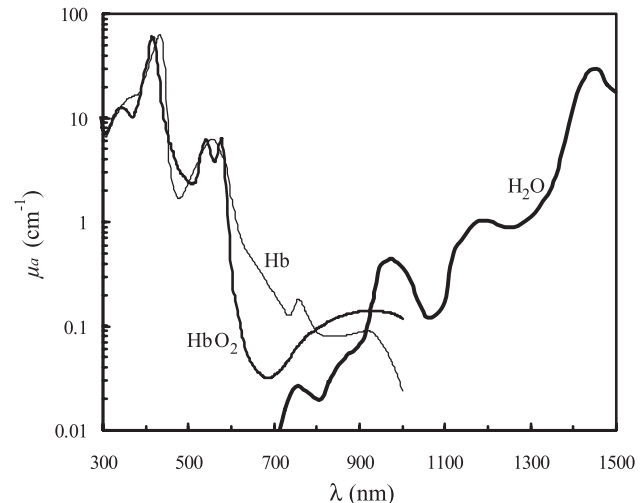


Fig. 1 Spectra of the absorption coefficients (μ_a) vs. wavelength (λ) for three chromophores commonly found in tissue, namely, oxyhemoglobin (HbO_2), deoxyhemoglobin (Hb), and water (H_2O). The concentrations of Hb and HbO_2 are assumed to be $50 \mu\text{M}$. These spectra were obtained from compiled absorption data for water^[6] and hemoglobin.^[7]

Reduced Scattering Coefficient

In most biological tissues, individual scattering events are not isotropic. In other words, the angular distribution of scattered photons is not uniform. The angular distribution of scattered light is described by the probability density $p(\theta)$ that a photon is scattered by an angle θ , which is defined as the angle between the direction of propagation of the scattered photon and that of the incident photon. The anisotropy factor g is defined as the average value of the cosine of the scattering angle, i.e., $g = \langle \cos\theta \rangle$. The value of g in tissue typically ranges from 0.7 to 0.99, indicating forward scattering at angles between 8° and 45° . A g factor of 1 indicates purely forward scattering, a g factor of 0 indicates isotropic scattering, and a g factor of -1 indicates purely backward scattering. In tissue optics the angular distribution of scattering is commonly described with the Henyey–Greenstein phase function.^[8]

$$p(\theta) = \frac{1 - g^2}{2(1 + g^2 - 2g \cos \theta)^{3/2}} \quad (7)$$

The reduced scattering coefficient μ_s' , which is defined as $\mu_s' = \mu_s(1 - g)$, can be interpreted as the inverse of the distance over which the photon loses memory of its initial direction of propagation, so that $1/\mu_s'$ is a measurement of the distance between effectively isotropic scattering events.

LIGHT-TISSUE INTERACTIONS AND RELATED APPLICATIONS

The interaction of light with tissue depends on many parameters, including the optical power and energy density, the optical properties of tissue, and the thermal properties of tissue such as heat conductivity and heat capacity.^[9] Different mechanisms of light-tissue interaction form the basis for a variety of applications of biomedical optics.

Fluorescence: Detection of Metabolic Changes

There are a number of fluorophores in tissue—both in cells (nicotinamide adenine dinucleotide (NADH); flavins; and aromatic amino acids such as tryptophan, tyrosine, and phenylalanine) and in the extracellular matrix (collagen, elastin), that account for endogenous fluorescence, also called autofluorescence. The spectral properties and relaxation times of autofluorescence may be different in normal and cancerous tissue,^[10] thus providing indications toward the detection and diagnosis of cancer. NADH is highly fluorescent (excitation maximum: 340 nm; emission maximum: 450 nm), but its deprotonated form NAD^+ is not. Consequently, fluorescence can be used to measure the NADH/NAD^+ ratio, which depends on the cellular oxidative energy metabolism.

Photochemical Interactions: Photodynamic Therapy

There are several light-induced processes that involve a chemical reaction. Some of these processes are photoaddition, photofragmentation, photo-oxidation, photohydration, *cis-trans* isomerization, and photo-rearrangement.^[11] These processes typically occur at relatively low optical power densities of the order of $1 \text{ W}/\text{cm}^2$ and long exposure times of seconds or more^[9] (Fig. 2). In photodynamic therapy, a photosensitizer (for example, hematoporphyrin derivatives or 5-aminolevulinic acid) is introduced into the bloodstream and distributes over most soft tissues. Its clearance in healthy tissue is faster than in cancerous tissue, so that after 3–7 days most of the photosensitizer is localized at the cancer. At this point, irradiation with light of the proper wavelength (usually in the red spectral region) excites the photosensitizer and triggers a series of processes that result in the generation of cytotoxic singlet oxygen, which induces necrosis of malignant cells. Another biomedical application of low-power (mW) light irradiation, called biostimulation, is believed to be based on photochemical processes. Biostimulation refers to a broad class of benefic effects (wound healing, anti-inflammatory effects, pain relief,

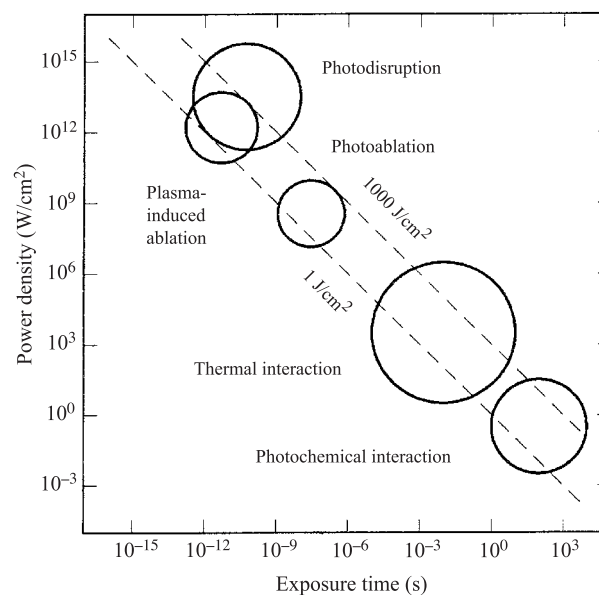


Fig. 2 Different mechanisms of laser-tissue interactions map onto different regions in the power-density/exposure-time space. These regions are schematically represented here. (Reproduced from Ref. [9] with permission.)

etc.) induced by low-level optical irradiation. Biostimulation remains controversial because of a lack of systematic clinical protocols that prove its efficacy and a lack of studies aimed at understanding the specific photochemical processes responsible for the therapeutic effects.

Thermal Interactions: Laser Tissue Welding

Thermal effects refer to the conversion of optical energy absorbed by the tissue into thermal energy as a result of nonradiative relaxation processes. The local heating of tissue depends on the properties of light irradiation as well as the optical and thermal properties of tissue. Depending on the local increase in temperature, four effects can be produced: coagulation, vaporization, carbonization, and melting.^[9] Coagulation requires a local temperature in excess of 60°C and is associated with necrosis. At temperatures above 100°C , the vaporization of water induces thermo-mechanical effects that are responsible for thermal tissue ablation. Carbonization, i.e., the release of carbon, occurs when the tissue temperature exceeds 150°C . At even higher temperatures, tissue melts as a result of the denaturation of tissue proteins (for example, collagen). The latter effect is used in laser tissue welding, in which proteinoous or dye-enhanced solder may also be used. Thermal interactions are typically induced with optical power densities of $10\text{--}10^6 \text{ W}/\text{cm}^2$ and pulse durations of $10^{-6}\text{--}10 \text{ s}$ ^[9] (Fig. 2).

Photoablation: Laser In Situ Keratomileusis (LASIK)

Photoablation refers to the photochemical decomposition of cellular and extracellular components by direct breaking of molecular bonds by UV photons. Photoablation is confined to the irradiated area so that it results in a well-defined area of tissue removal. Excimer lasers (for example, ArF, XeCl, XeF, etc.) emitting at UV wavelengths of 193–351 nm are typically used for photoablation. Power densities are relatively high, in the range 10^7 – 10^{10} W/cm², and pulse durations range from 10 to 100 ns (Fig. 2). LASIK is a routine laser-based method that uses photoablation to flatten the cornea to correct for myopia.

Plasma-Induced Ablation and Photodisruption: Lithotripsy

At high optical power densities of 10^{11} – 10^{13} W/cm², which are typically delivered over relatively short times (10^{-13} – 10^{-10} s), the electrons in the tissue are subject to an electric field that induces forces higher than the atomic attractive forces, thus causing breakdown and the formation of ionized plasma. On the one hand, this plasma of free electrons strongly absorbs optical radiation leading directly to ablation (plasma-induced ablation). On the other hand, at even higher pulse energies, shock waves and cavitation (formation of microscopic gas bubbles) may also take place leading to photodisruption. A key difference between plasma-induced ablation and photodisruption is that the former is spatially confined to the illuminated spot, whereas the latter affects adjacent tissue as well because of the propagation of shock waves. Photodisruption finds applications in the rupture of urinary calculi (lithotripsy).

LIGHT PROPAGATION IN TISSUE

Transport Theory

The description of light propagation in tissue is usually based on the particle description of light in terms of photons. Transport theory treats the flow of particles that experience scattering and absorption in the media where they propagate. The radiative transport equation for the angular photon density u (defined earlier in the section on Radiation Parameters) as a function of position (\mathbf{r}), direction of propagation ($\hat{\Omega}$), and time (t)

is written as follows:^[12]

$$\begin{aligned} \frac{\partial u(\mathbf{r}, \hat{\Omega}, t)}{\partial t} = & -c\hat{\Omega} \cdot \nabla u(\mathbf{r}, \hat{\Omega}, t) \\ & -v(\mu_a + \mu_s)u(\mathbf{r}, \hat{\Omega}, t) \\ & + c\mu_s \int_{4\pi} u(\mathbf{r}, \hat{\Omega}', t)p(\hat{\Omega}', \hat{\Omega})d\hat{\Omega}' \\ & + q(\mathbf{r}, \hat{\Omega}, t) \end{aligned} \quad (8)$$

where $p(\hat{\Omega}', \hat{\Omega})$ is the phase function describing the probability density of scattering from direction $\hat{\Omega}'$ to $\hat{\Omega}$, and $q(\mathbf{r}, \hat{\Omega}, t)$ is the source term, representing the number of photons injected by the light source per unit volume, per unit time, per unit solid angle at position \mathbf{r} , time t , and direction $\hat{\Omega}$. The first term on the right-hand side of Eq. 8 represents the net gain of photons at position \mathbf{r} and direction $\hat{\Omega}$ due to the flow of photons. The second term represents the loss of photons at \mathbf{r} and $\hat{\Omega}$ as a result of absorption and scattering. The third term represents the gain of photons at \mathbf{r} and $\hat{\Omega}$ due to scattering. Finally, the fourth term represents the gain of photons due to the light sources. The transport equation forms the basis for a general treatment of light propagation in absorbing/scattering media. However, it cannot be solved in analytical form even for the simplest geometries and homogeneous media, so that simplified approaches are commonly in use.

Diffusion Approximation

In biomedical optics, a common simplified approach to describe light propagation in tissue is the diffusion approximation, which requires a strongly scattering medium (i.e., $\mu_a \ll \mu_s'$), time variations of the photon density that are slow with respect to the average photon interaction rate $c(\mu_a + \mu_s')$, and distances from sources and boundaries that are much greater than $1/\mu_s'$. Under these conditions the transport equation reduces to the diffusion equation:

$$\frac{\partial U(\mathbf{r}, t)}{\partial t} = D\nabla^2 U(\mathbf{r}, t) - c\mu_a U(\mathbf{r}, t) + S_0(\mathbf{r}, t) \quad (9)$$

where $U(\mathbf{r}, t)$ is the photon density, D is the diffusion coefficient defined as $D = c/[3(\mu_s' + \mu_a)]$, and $S_0(\mathbf{r}, t)$ is the spherically symmetric component in the expansion of the source term $q(\mathbf{r}, \hat{\Omega}, t)$ into spherical harmonics.

Optical studies of tissue are performed using continuous-wave (CW), time-domain (TD), or frequency-domain (FD) techniques. CW studies employ a light source whose emission properties are time-independent, TD techniques employ short (compared to the photon time-of-flight in tissue) pulses, and FD

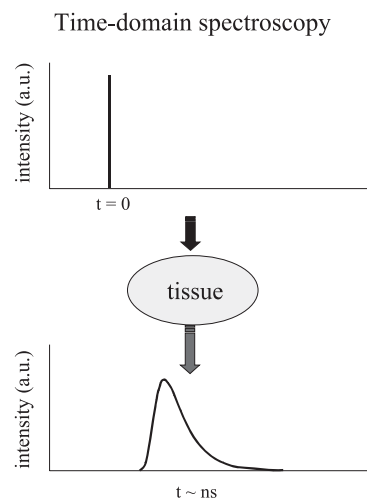


Fig. 3 In time-domain spectroscopy, the tissue is interrogated with a short pulse of light (at time $t=0$), and the temporal distribution (over a timescale of nanoseconds) of the optical pulse that has propagated through tissue can be examined.

techniques employ intensity-modulated light sources, with a modulation period that is comparable to the photon time-of-flight in tissue. The TD and FD techniques are pictorially illustrated in Figs. 3 and 4, respectively. The CW, TD, and FD solutions to the diffusion equation for a point source at $r=0$ in an infinite medium are the following:

$$U_{CW}(r) = \frac{P_0}{4\pi D} \frac{e^{-r\sqrt{3\mu_a\mu_s'}}}{r} \quad (10)$$

$$U_{TD}(r, t) = \frac{N_0}{(4\pi Dt)^{3/2}} e^{-\frac{3\mu_s'^2 r^2}{4ct} - \mu_a ct} \quad (11)$$

and

$$U_{FD}(r, \omega) = \frac{P(\omega)}{4\pi D} \frac{e^{ikr}}{r} \quad (12)$$

where P_0 and $P(\omega)$ are the CW and modulated source powers, respectively, N_0 is the number of photons emitted in the pulse for time-domain studies, and $k = \sqrt{(i\omega - c\mu_a)/D}$ (where the imaginary part of k is positive, so that $^{[12]}$ describes a wave that attenuates rather than amplifies as a function of the distance from the source r). Solutions to the diffusion equation for semi-infinite, slab, cylindrical, and spherical geometries are available in analytical form,^[12] and solutions for irregular geometries and inhomogeneous media are usually obtained using numerical methods.

Diffusion theory is often used to model light propagation in tissue, and its solutions in CW, TD, and FD (Eqs. 10–12) are used to translate optical

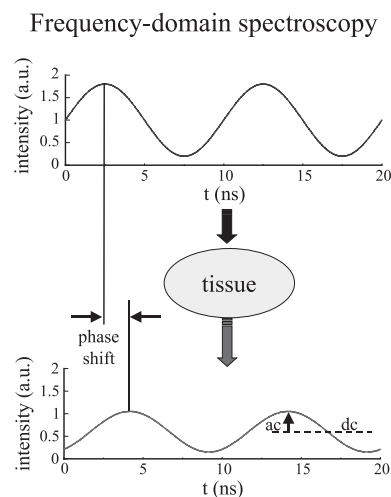


Fig. 4 In frequency-domain spectroscopy, intensity-modulated light is used to illuminate the tissue. The measured parameters are the average value (dc), the amplitude (ac), and the phase of the modulated signal.

measurements into measurements of the absorption and scattering properties of tissue.

BIOMEDICAL OPTICAL SPECTROSCOPY

An optical spectroscopy experiment consists of illuminating a sample with light of a known wavelength (or wavelength range) and measuring some properties of the light that emerges from the sample after having interacted with it. Absorption spectroscopy and elastic scattering techniques measure the fraction of the optical radiation dissipated or scattered, respectively, by the sample; fluorescence, phosphorescence, and inelastic light scattering measure the emitted intensity at wavelengths that are different from the wavelength used for illumination; some spectroscopic techniques involve the measurement of light polarization, for instance, optical rotatory dispersion, circular dichroism, and fluorescence polarization.

Absorption Spectroscopy

The measurement of absorption of UV, visible, and near-IR light, which is associated with transitions between electronic molecular energy states, is a common spectroscopic technique in biomedical optics. IR spectroscopy, which is associated with vibrational transitions, is less common because of the lower absorption of tissue constituents in the IR compared with the UV and visible spectral ranges, and the strong IR absorption of water. Absorption spectroscopy can give indications on the concentrations of specific

chromophores or provide structural information. A specific chromophore is associated with an absorption coefficient per unit concentration (ϵ), also called molar extinction coefficient, so that a measurement of the absorption coefficient (μ_a) yields the chromophore concentration $C = \mu_a/\epsilon$. The absorption properties of a protein or nucleic acid in a polymer are affected by the conformation (three-dimensional arrangement) of other residues in the polymer; thus absorption spectroscopy can provide structural information. A number of features of the absorption spectra of many chromophores (intensity, band shape, wavelength of absorption peak) are also sensitive to local environment (solvent).

Light-Scattering Spectroscopy

The wavelength dependence and the angular distribution of scattered light in tissue allows for the determination of the population density and size distribution of epithelial cell nuclei.^[14] This technique—called light-scattering spectroscopy—provides histological information about tissue *in vivo* and has shown the potential to diagnose early cancerous and precancerous conditions on the basis of the enlargement of dysplastic or malignant nuclei.

Measurements of the optical reduced scattering coefficient (μ_s') of tissue have been correlated with the concentration of glucose in the blood.^[15,16] An increase in the glucose concentration decreases the index of refraction mismatch between the extracellular fluid and the cellular membranes, resulting in a smaller value of μ_s' . It has also been shown that the scattering properties of brain tissue are affected by cell swelling^[17] and by events occurring at the neuronal membranes as a result of brain activity.^[18]

Fluorescence Spectroscopy

The radiative decay from an optically excited electronic state is referred to as luminescence. Fluorescence is the emission associated with a transition from a singlet excited state, and it occurs on a typical time scale of tens of nanoseconds. Because the excitation energy provided by the absorbed photons is partially dissipated through nonradiative relaxation processes, the emitted photons have a longer wavelength (i.e., are red-shifted) with respect to the excitation photons. Fluorescence measurements provide information on the structure, interactions, and dynamics of biological samples at the molecular and cellular level, and are commonly used for the study of biological specimens. In addition to the measurements of excitation and emission spectra, fluorescence techniques also involve measurements of the lifetime of the excited state, the

quantum yield (i.e., the ratio of the number of photons emitted to the number of photons absorbed), and the depolarization (i.e., the loss of polarization of fluorescence induced by dynamical processes of the fluorescent molecule). The fluorescence lifetime can be measured in the time domain by directly measuring the fluorescence decay in response to a pulsed excitation, or in the frequency domain by measuring the phase difference between an intensity-modulated excitation and the fluorescence signal, which is also modulated at the frequency of excitation.

Raman Spectroscopy

Raman spectroscopy is an inelastic scattering technique that probes molecular vibrational energy levels, as in IR spectroscopy. In the case of centrosymmetric molecules, there are vibrational modes that are either Raman-active or IR-active, so that Raman spectroscopy and IR spectroscopy provide complementary information. Raman spectroscopy is typically less sensitive than IR spectroscopy, and Raman signals may be overwhelmed by fluorescence signals in the case of samples that display autofluorescence. However, Raman spectroscopy presents a number of advantages in that it is applicable to aqueous media (because water contributes a small Raman signal, whereas it strongly absorbs IR light), it does not require special sample preparation, and it can be used to obtain spectra on microscopic areas and even on single cells. Raman spectroscopy has been applied to intact tissues for the identification of cancers and precancers.^[19]

BIOMEDICAL OPTICAL IMAGING

Fluorescence Microscopy

Fluorescence microscopy is a powerful bioimaging technique. It is based on measuring the spatial distribution of endogenous fluorophores or exogenous fluorescent molecules that label a specific tissue, cell, or cellular organelle. Fluorescence microscopy is typically performed in epi-illumination, i.e., using the same objective lens to focus the excitation light on the sample and to collect the fluorescence emission from the sample. The excitation light and the emission light are separated using a dichroic mirror, which reflects excitation light at a shorter wavelengths and transmits fluorescence emission light at longer wavelengths. The problem associated with the signal generated at out-of-focus regions in the sample is addressed by confocal microscopy.

In confocal microscopy,^[20] a pinhole is used at the focal plane on the detector side (see the left side of



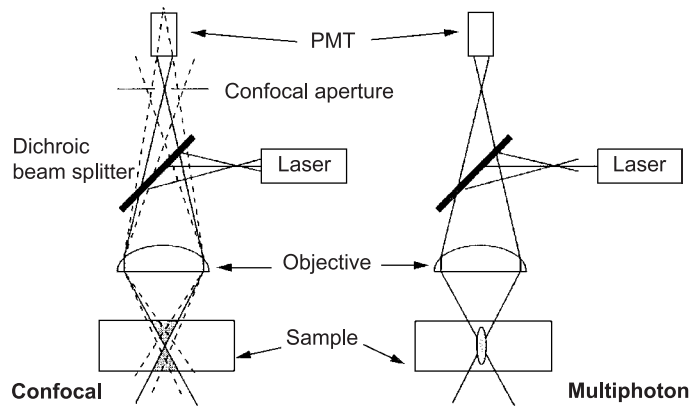


Fig. 5 Schematic diagrams of confocal and multiphoton microscopy. Confocal microscopy (left) achieves z -sectioning by suppressing light generated out of the focal plane by means of a confocal aperture. Multiphoton microscopy (right) achieves z -sectioning by exciting fluorescence only at the focal plane. PMT: photomultiplier tube. (Reproduced from Ref. [11] with permission of John Wiley & Sons, Inc.)

Fig. 5). Objects on the sample at out-of-focus regions are equivalent to extended sources on the illumination focal plane, so that their off-axis emission is rejected by the limited aperture of the pinhole at the detector. Depending on the wavelength used and the size of the pinhole, the size of the optical section in the z direction is on the order of $0.4\text{--}0.8\ \mu\text{m}$. By scanning the focused spot of illumination, confocal microscopy provides three-dimensional optical imaging. Some potential limitations of confocal fluorescence microscopy are the relatively high excitation power required to compensate the signal suppression of the confocal pinhole and the limited optical penetration in tissue of UV/blue light used to excite most biologically relevant fluorophores. In particular, the high power level and the fluorophore excitation over the whole cone of illumination may lead to photobleaching (i.e., the light-induced chemical degradation of a fluorophore) over a relatively large sample volume. These potential problems are overcome by two-photon microscopy.

In two-photon microscopy (right side of Fig. 5) the fluorophore is excited by the simultaneous absorption of two photons, whose individual energy (wavelength) is half (twice) that of photons used for one-photon excitation.^[21] As a result, excitation light for two-photon excitation is typically in the red/near-IR region of the spectrum (where penetration depth in tissue is relatively high) as opposed to the UV/visible spectral range commonly used for one-photon excitation. Furthermore, two-photon excitation is confined to the focal spot of illumination because of the relatively small cross-section for two-photon excitation and the quadratic dependence of the fluorescence intensity on the incident photon flux. First, this fact results in a z -sectioning effect similar to that achieved by confocal microscopy, but without using any emission pinholes. Second, by limiting the fluorophore excitation to the focal spot, two-photon microscopy achieves a spatial confinement of photobleaching and photodamage to the sample.

Fluorescence lifetime imaging microscopy (FLIM) measures the spatial distribution of the fluorescence lifetime within the sample.^[22] Fluorescence lifetime is a sensitive indicator of the environment of the fluorophore and is independent of the excitation intensity and fluorophore concentration. As a result, FLIM is a powerful tool for bioimaging and it has been implemented by using both one-photon and two-photon excitation.

Optical Coherence Tomography (OCT)

Optical coherence tomography is an interferometric, backscattering imaging technique that provides cross-sectional structural images of tissue.^[23] OCT uses low-coherence interferometry to achieve depth discrimination. The basic idea is to split the light emission from the source into a sample beam and a reference beam. The sample beam is focused onto the tissue, whereas the reference beam is reflected by a scanning mirror. The light backscattered by the tissue on the sample arm of the interferometer and the light reflected by the reference mirror on the reference arm are mixed at the optical detector. The two signals produce a well-defined interference pattern only if the difference between the sample and reference pathlengths is within the coherence length of the light source. A short coherence length results in a high level of depth discrimination by scanning the reference mirror so that low-coherence sources are used in OCT. Because the coherence length is inversely proportional to the bandwidth of the light source, broadband light sources such as superluminescent diodes or mode-locked Ti:Sapphire lasers are suitable light sources for OCT. The source of contrast in OCT images is the high-frequency spatial variation (over a few wavelengths) of the coherent backscattering cross-section,^[24] as determined, for example, by discontinuities in the index of refraction. However, additional sources of contrast can be investigated with OCT by combining polarization,

Doppler shifts, or absorption information with the depth-discrimination feature of OCT.

Reflectance Optical Imaging

The spatially-resolved collection of light that is back-scattered/backreflected from biological tissue under pencilbeam or broadbeam illumination conditions is referred to as reflectance imaging. In contrast with the cases of confocal microscopy, two-photon microscopy, and optical coherence tomography, reflectance imaging provides limited capabilities of depth discrimination. However, there are applications that aim at imaging the collective optical features of superficial tissue layers, where depth discrimination is not a requirement. The basic approach is to illuminate the sample either directly with a laser beam or with light guides coupled to a variety of light sources. The light collection is usually performed with charge coupled device cameras. The detection of cutaneous melanoma^[25] and functional imaging of the exposed brain cortex^[26] are two examples of applications of optical reflectance imaging. By combining the spatially-resolved information of optical imaging with wavelength-resolved information, one obtains reflectance spectra at each image pixels. This spectrally- and spatially-resolved technique, which is called spectral imaging, requires tunable bandpass filters (for example, liquid crystal tunable filters or acousto-optic tunable filters) to sequentially select specific spectral regions (as narrow as 1 nm) over the visible/near-IR spectral range of 400–1000 nm.^[27]

Diffuse Optical Imaging

Most biological tissues are optically turbid media, and light propagates inside them in a diffusive regime. This diffusive mode of light propagation (also referred to as photon migration) results in the fact that light loses its polarization, coherence, and memory of the initial direction of propagation over distances on the order of millimeters. Nevertheless, optical imaging can be performed over optically thick tissues by detecting multiply scattered light; this technique is called diffuse optical imaging or diffuse optical tomography.^[28,29] The choice of the spectral region for diffuse optical imaging is dictated by the requirement of achieving an optical penetration depth on the order of centimeters, which in turn requires a relatively small absorption coefficient of tissue. Figure 1 shows that the absorption of hemoglobin and water, two of the dominant light-absorbing substances in tissue, is minimal in the near-IR band of 700–900 nm, which is the spectral range of choice in diffuse optical imaging. Continuous-wave, frequency-domain, and time-domain techniques are all used for diffuse optical imaging, according to a

reflection geometry or transmission geometry. The diffusive nature of light propagation inside tissue poses an intrinsic limit to the spatial resolution that can be achieved with this imaging modality. However, the high optical contrast provided by the spatial distribution of the concentration and oxygen saturation of hemoglobin (see Fig. 1) accounts for the effectiveness of diffuse optical imaging in areas such as tissue oximetry, breast cancer detection, and functional imaging of the brain.

CONCLUSIONS

Biomedical optics, also called biophotonics, is a very active field of research that integrates basic sciences, engineering, and clinical practice for the development of new and more effective tools in medicine and biology. Recent developments point to applications of biophotonics on a nanoscale level, for which the term bionanophotonics has been used. For example, nanoparticles such as quantum dots (fluorescent nanocrystals of semiconductors whose emission wavelengths depend on their size) and metallic nanoparticles are used for optical bioimaging and biosensing, respectively. Even BioMEMS (micro-electro-mechanical system designed for biomedical applications) is moving to the nanoscale. Biomedical optics thus covers applications that range from the macroscopic study of whole organs to the microscopic study at the cellular, subcellular, and molecular level, and to the new developments in the nanoscale world involving nanotechnology and nanophotonics.



ACKNOWLEDGMENTS

I acknowledge support from the National Institutes of Health (Grant No. DA14178) and from The National Science Foundation (Award No. BES-93840).

ARTICLES OF FURTHER INTEREST

Flow Cytometry; Laser-Tissue Interaction; Laser Transmyocardial Revascularization; Optical Detection of Cancers; Optical Mapping

REFERENCES

1. Fishkin, J.B.; Gratton, E. Propagation of photon-density waves in strongly scattering media containing an absorbing semi-infinite plane bounded by a straight edge. *J. Opt. Soc. Am. A* **1993**, *10*, 127–140.

2. O'Leary, M.A.; Boas, D.A.; Chance, B.; Yodh, A.G. Refraction of diffuse photon density waves. *Phys. Rev. Lett.* **1992**, *69*, 2658–2661.
3. Chance, B.; Kang, K.; He, L.; Weng, J.; Sevick, E. Highly sensitive object location in tissue models with linear in-phase and anti-phase multi-element optical arrays in one and two dimensions. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 3423–3427.
4. Boas, D.A.; O'Leary, M.A.; Chance, B.; Yodh, A.G. Scattering and wavelength transduction of diffuse photon density waves. *Phys. Rev., E* **1993**, *47*, R2999–R3002.
5. Hetzel, F.; Patterson, M.; Preuss, L.; Wilson, B. *Recommended Nomenclature for Physical Quantities in Medical Applications of Light*; AAPM Report No. 57, Ed.; American Institute of Physics: Woodbury, NY, 1996; 1–6.
6. Hale, G.M.; Querry, M.R. Optical constants of water in the 200 nm to 200 m wavelength region. *Appl. Opt.* **1973**, *12*, 555–563.
7. S. Prahl (of the Oregon Medical Laser Center, Portland, OR) has tabulated the molar extinction coefficients for oxy-hemoglobin and deoxy-hemoglobin using data from W.B. Gratzer and N. Kollias. These tabulated data are available at <http://omlc.ogi.edu/spectra/hemoglobin/summary.html>.
8. Jacques, S.L.; Wang, L. Monte Carlo modeling of light transport in tissues. In *Optical-Thermal Response of Laser-Irradiated Tissue*; Welch, A.J., van Gemert, M.J.C., Eds.; Plenum Press: New York, NY, 1995; 73–100.
9. Niemz, M.H. *Laser-Tissue Interactions: Fundamentals and Applications*; Springer-Verlag: Berlin, Germany, 1996.
10. Alfano, R.R.; Tata, D.B.; Cordero, J.J.; Tomashefsky, P.; Longo, F.W.; Alfano, M.A. Laser induced fluorescence spectroscopy from native cancerous and normal tissues. *IEEE J. Quantum Electron.* **1984**, *QE-20*, 1507–1511.
11. Prasad, P.N. *Introduction to Biophotonics*; John Wiley and Sons: Hoboken, NJ, 2003.
12. Duderstadt, J.J.; Hamilton, L.J. *Nuclear Reactor Analysis*; Wiley: New York, NY, 1976; 113.
13. Arridge, S.R.; Cope, M.; Delpy, D.T. The theoretical basis for the determination of optical pathlengths in tissue: Temporal and frequency analysis. *Phys. Med. Biol.* **1992**, *37*, 1531–1560.
14. Perelman, L.T.; Backman, V.; Wallace, M.; Zonios, G.; Manohaam, R.; Nusrat, A.; Shields, S.; Seiler, M.; Lima, C.; Hamamo, T.; Itzkan, I.; Van Dam, J.; Crawford, J.M.; Feld, M.S. Observation of periodic fine structure in reflectance from biological tissue: A new technique for measuring nuclear size distribution. *Phys. Rev. Lett.* **1998**, *80*, 627–630.
15. Maier, J.S.; Walker, S.A.; Fantini, S.; Franceschini, M.A.; Gratton, E. Possible correlation between blood glucose concentration and the reduced scattering coefficient of tissues in the near infrared. *Opt. Lett.* **1994**, *19*, 2062–2064.
16. Bruulsema, J.T.; Hayward, J.E.; Farrel, T.J.; Patterson, M.S.; Heinemann, L.; Berger, M.; Koschinsky, T.; Sandahl-Christiansen, J.; Orskov, H.; Essenpreis, M.; Schmelzeisen-Redeker, G.; Böcker, D. Correlation between blood glucose concentration in diabetics and noninvasively measured tissue optical scattering coefficient. *Opt. Lett.* **1997**, *22*, 190–192.
17. MacVicar, B.A.; Hochman, D. Imaging of synaptically evoked intrinsic optical signals in hippocampal slices. *J. Neurosci.* **1991**, *11*, 1458–1469.
18. Stepnoski, R.A.; LaPorta, A.; Raccuia-Behling, F.; Blonder, G.E.; Slusher, R.E.; Kleinfeld, D.J. Non-invasive detection of changes in membrane potential in cultured neurons by light scattering. *Proc. Natl. Acad. Sci. U. S. A.* **1991**, *88*, 9382–9386.
19. Mahadevan-Jansen, A.; Richards-Kortum, R. Raman spectroscopy for the detection of cancers and precancers. *J. Biomed. Opt.* **1996**, *1*, 31–70.
20. Wilson, T.; Sheppard, C. *Theory and Practice of Scanning Optical Microscopy*; Academic Press: New York, NY, 1984.
21. Denk, W.; Strickler, J.H.; Webb, W.W. Two-photon laser scanning fluorescence microscopy. *Science* **1990**, *248*, 73–76.
22. Tadrous, P.J. Methods for imaging the structure and function of living tissues and cells: 2. Fluorescence lifetime imaging. *J. Pathol.* **2000**, *191*, 229–234.
23. Schmitt, J.M. Optical coherence tomography (OCT): A review. *IEEE J. Sel. Top. Quantum Electron.* **1999**, *5*, 1205–1215.
24. Fercher, A.F. Optical coherence tomography. *J. Biomed. Opt.* **1996**, *1*, 157–173.
25. Marchesini, R.; Bono, A.; Bartoli, C.; Lualdi, M.; Tomatis, S.; Cascinelli, N. Optical imaging and automated melanoma detection: Questions and answers. *Melanoma Res.* **2002**, *12*, 279–286.
26. Frostig, R.D.; Lieke, E.E.; Ts'o, D.Y.; Grinvald, A. Cortical functional architecture and local coupling between neuronal activity and the microcirculation revealed by in vivo high-resolution imaging of intrinsic signals. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *87*, 6082–6086.
27. Farkas, D.L.; Becker, D. Applications of spectral imaging: Detection and analysis of human melanoma and its precursors. *Pigment Cell Res.* **2001**, *14*, 2–8.
28. Hebden, J.C.; Arridge, S.R.; Delpy, D.T. Optical imaging in medicine: I. Experimental techniques. *Phys. Med. Biol.* **1997**, *42*, 825–840.
29. Arridge, S.R.; Hebden, J.C. Optical imaging in medicine: II. Modelling and reconstruction. *Phys. Med. Biol.* **1997**, *42*, 841–853.

Organic and Inorganic Matrices

Adele L. Boskey

Musculoskeletal Integrity Program, Hospital for Special Surgery, New York, New York, U.S.A.

INTRODUCTION

Mineralized tissues such as bone, tooth, and shell are composites, consisting of one or more inorganic mineral phases deposited in an oriented fashion upon an organic matrix. The mechanical properties of the composites are markedly different from those of the individual components. To understand the mechanical performance of these tissues, as well as to design tissue engineered replacements for them, it is crucial to be aware of the properties of these tissues, how they are formed, and how they function. This article focuses on bone organic and inorganic matrices, but provides illustrations and examples from teeth and shells.

Definitions

Most mineralized tissues, specifically those mineral-containing extracellular matrices formed in biologic systems, are composites consisting of an organic matrix and an inorganic one. These matrices are produced under cellular control, and together provide their component tissues with mechanical and compressive strength as well as elasticity that far exceeds that of their individual components.^[1-3] By definition, the organic matrix refers to those components made by the cells that consist predominately of proteins, sugars, and lipids, while the inorganic matrix refers to those components that are not made of hydrocarbons. A common feature of the mineralized tissues is their highly anionic organic matrices; these matrices control the deposition of biominerals upon an oriented structural support. Both soluble and insoluble matrices do this, in part, by acting as “templates.” Templating in biomineralization is generally provided by an insoluble matrix, soluble macromolecules, and precipitating ions. The insoluble template provides supramolecular architecture, while the soluble macromolecules associate with the precipitating ions, facilitating crystal nucleation on the template surface. The same soluble macromolecules, having the ability to bind to and stabilize the precipitate, can also regulate crystal growth (size, shape, and aggregation) and the extent to which these crystals proliferate. Copying these naturally occurring tissues and the processes by which they

form (biomimetics) is a fundamental technique in tissue engineering.

In bones, teeth, and a limited number of shells, the inorganic material is an analogue of the naturally occurring mineral, hydroxyapatite (HA). The chemical formula for pure HA is $\text{Ca}_5(\text{PO}_4)_3\text{OH}$, but in biologic tissues, the calcium-to-phosphate molar ratio is often quite different from the stoichiometric 1.67:1 predicted from this formula.^[4] Examples of the inorganic materials found in different tissues within the same species and in different species are listed in Table 1.

Even when the same mineral phase is present, the properties of the mineral may differ with tissue type, animal and tissue age, animal health and disease, and tissue function. The mineral in shells is most frequently a calcium carbonate, but the crystalline phase present may be aragonite, calcite, or vaterite, and often is a mixture of these phases.^[5] Some shells contain non-crystalline (amorphous) calcium carbonates.^[6] Here we review these inorganic components, their organization in the different mineralized tissues, the ways in which they can be analyzed, how the organic matrix on which they are deposited influences their formation, and how this knowledge is being used to design tissue-engineered constructs for bone and tooth repair.

BONE

Structure and Function

The bones of the body are categorized based on shape (long and flat), structural arrangement (cortical and trabecular/lamellar, woven, and compact), or by their mechanism of development (endochondral and intramembranous ossification). Some of these classifications are illustrated in Fig. 1. Flat bones (with flat as distinct from rounded cross-sections) are typified by those in the skull or ribs. They consist of a layer of spongy/lamellar bone between two thin layers of compact/cortical bone. Flat bones do not have a bone marrow cavity, but they do contain marrow. Long bones such as the femur or tibia (Fig. 1) have a shaft and two cartilage covered ends. They are dense bones providing strength, structure, and mobility. Covered by the cortex (compact bone), the long bones have a marrow



Table 1 Inorganic matrix constituents

Mineral phase and chemical formula	Found in biology
Calcium phosphates HA $\text{Ca}_5(\text{PO}_4)_3(\text{OH})$	Bones, dentin, cementum, enamel Phosphatidic shells
Brushite $\text{CaHPO}_4 \cdot 2\text{H}_2\text{O}$	Salivary and kidney stones; ? embryonic bone
Octacalcium phosphate $\text{Ca}_8\text{H}_2(\text{PO}_4)_3$	Enamel
Tricalcium phosphate $\text{Ca}_3(\text{PO}_4)_2$?
Calcium pyrophosphate $\text{Ca}_2\text{P}_2\text{O}_7 \cdot x\text{H}_2\text{O}$	Pathologic deposits
Calcium carbonate Aragonite Calcite Vaterite	Shells, pancreatic stones
Iron oxide	Magnetobacteria
Silicates	Diatoms

cavity and contain both yellow and red marrow. (The red marrow produces blood cells.) The cortical bone consists of unique structural units (osteons or Haversian systems). Each osteon has a central vascular channel

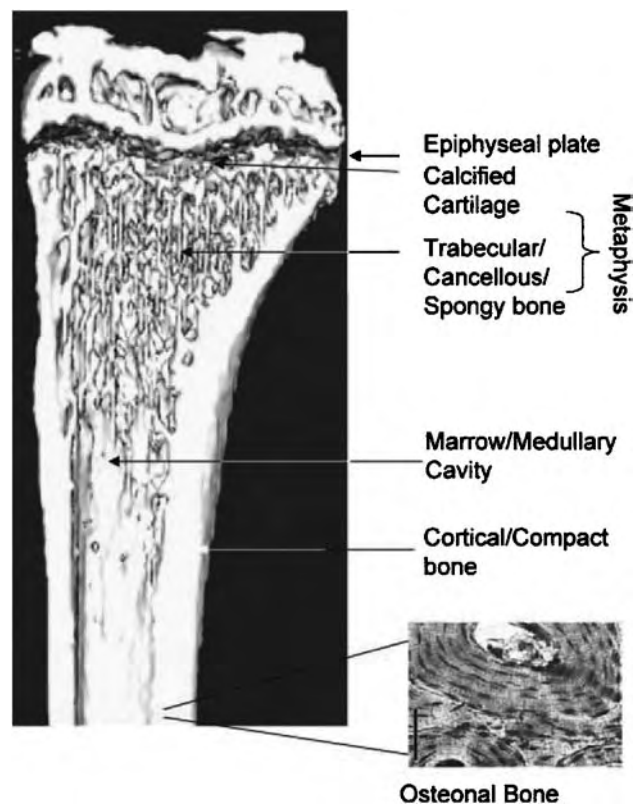


Fig. 1 The structure of a long bone. The anatomic areas are identified on a microcomputed image of a mouse tibia. The pull out shows a cross-section of compact bone, demonstrating the presence of concentric circles of mineral around a vascular cavity (Haversian canal).

(Haversian canal) surrounded by concentric layers of mineralized bone (lamellae). Distinct from cortical/osteonal bone is the trabecular/cancellous bone. It is found in metaphyseal and epiphyseal regions of the long bones. Much of the bone surface contains this less dense, more elastic, high turnover trabecular bone. Its arrangement is mesh-like (Fig. 2). Trabecular bone forms the interior portions of the bones, and is rigid but “spongy,” consisting of short parallel strands of bone, which fuse together.

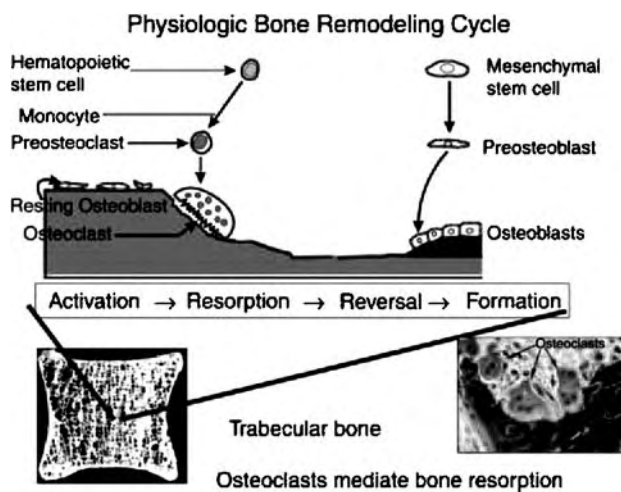


Fig. 2 The process of remodeling in bone is characterized by the coupling of osteoblasts (derived from mesenchymal stem cells), and osteoclasts (derived from hematopoietic stem cells). The pull out shows a segment of a vertebral body showing the individual trabeculae, and a stained image of osteoclasts (multinucleated giant cells) on a bone surface. (Courtesy of Dr. Steven Goldring, Harvard University School of Medicine.)

The entire bone structure provides strength and rigidity, enabling mobility while protecting internal organs. Bone is also the storehouse for calcium and phosphate (and other less abundant) ions required for the metabolism. Bone is constantly in a dynamic state, being formed by the bone-forming cells (osteoblasts) and removed by the macrophage-derived bone-dissolving cells (osteoclasts). These processes are coupled, as illustrated in Fig. 2. When the coupling is defective, i.e., when there is an imbalance between bone remodeling and bone formation, the mechanical properties of the bone composite are impaired. Examples of such uncoupling associated with increased risk of bone fracture include osteoporosis (too much remodeling resulting in loss of bone tissue), osteopetrosis (too little remodeling resulting in rock-like bones), or osteomalacia (too little mineralization).

Composition—The Mineral

The extracellular matrix of the bone composite in the adult animal consists in decreasing order of HA mineral, collagen, non-collagenous proteins, lipids, and water. The amount of mineral is usually determined by measuring the ash content (determining the gravimetric yield after the organic matrix is removed from dry-weighed bone by heating to 600°C). Radiographic methods such as backscatter electron imaging,^[7] quantitative microradiography,^[8] and dual photon absorptiometry^[9] provide two-dimensional estimates of ash content, while quantitative micro-computerized tomography can be used to calculate “mineral tissue density,” a three-dimensional parameter.^[10] While each of these methods provides insight into the amount of mineral that is present and how it is distributed in the tissue, none of them provides information on the nature of the inorganic mineral present.

The nature of the mineral phase present is generally determined by X-ray powder diffraction analysis (XRD).^[4] For this technique, the material in question is ground to a homogeneous powder and exposed to collimated X-rays that are diffracted (reflected) off the surface of the sample. The peaks that are detected (by a radiation detector or classically by a film) occur at positions determined by the spacing between the atomic planes, and are described by Bragg's Law

$$n\lambda = 2d \sin \theta$$

where n is the order of the reflection, λ , the wavelength of the radiation, d , the spacing between the planes, and θ , the angle of incidence (or reflection). The diffraction pattern provides a “fingerprint” of the phase present or if a mixture of the relative amounts of components

based on values published in the powder diffraction files.^[11] Fig. 3A presents a typical XRD pattern of large crystals of HA, such as are found in sintered apatite implant materials. The c -axis (002) reflection used to estimate particle size with the Scherrer equation (below) is indicated. Other spectroscopic techniques such as infrared, Raman, and nuclear magnetic resonance can also provide insight into the environment of molecular constituents, and hence can also provide “fingerprints” of structures;^[4,12] however XRD is the “gold standard.”

The XRD pattern not only allows phase identification, but also provides information on the relative crystal size and perfection of the crystals being examined. This is because the Bragg's equation presented above assumes that the crystals are perfect, all the planes are aligned, and the resultant diffraction peaks are sharp. But for biominerals such as HA, the peaks are broadened. The Scherrer equation can be used

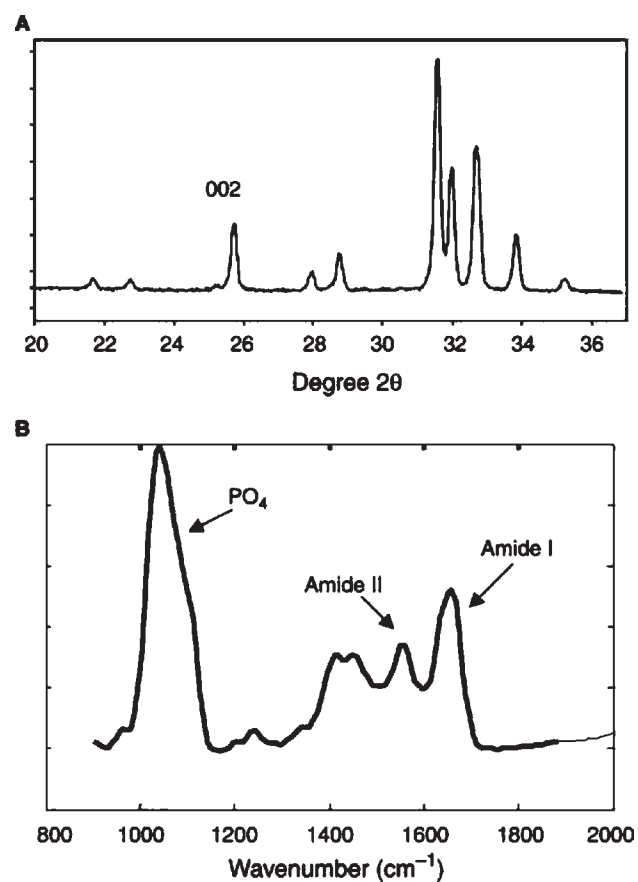


Fig. 3 Illustration of techniques for identification of the inorganic phase. (A) X-ray diffraction of a synthetic HA. The c -axis 002 peak used to estimate crystallite size in biologic tissues is indicated. (B) Infrared spectrum of adult human bone, the absorbances corresponding to vibrations discussed in the text are indicated.

to calculate the average “size and perfection” of component crystals.^[13]

$$t = K\lambda/\beta_{1/2} \cos \theta$$

Where t is the thickness of the particle along the plane being examined, K a constant (usually set at 0.9, $\beta_{1/2}$ is the line width at half maximum for the peak at θ degrees).

Crystallite size can be determined directly by electron microscopy^[14] or atomic force microscopy (AFM).^[15] AFM measurements and electron microscopy provide crystal sizes for individual crystals, rather than the averages as determined by X-ray or electron diffraction. However, for bone, the values determined by all methods are in relative agreement, with the smallest dimension of the crystals being 11–30 nm, and the longest (c -axis length) dimension being 10–60 nm.

The chemical composition of the mineral can also provide insight into the nature of the phase(s) present. However, for calcium phosphates, the data is difficult to interpret not only because there are a variety of calcium phosphates in mineralized tissues (Table 1), but also because the small size of the physiologic HA crystals (and their high surface/volume area) leads to both the inclusion and the surface adsorption of foreign ions. However, at the electron microscopic level using quantitative microprobe analyses, localized compositional changes can be determined, and this can provide unique information on variations in composition within the tissue.^[16]

Infrared microscopy and infrared imaging microscopy can provide spatial information not only on the amount of mineral present (mineral/matrix ratio) but also on the crystallinity, carbonate content, and acid phosphate contents based on comparisons with X-ray diffraction data.^[17] A typical infrared spectrum of adult human bone is shown in Fig. 3B. In infrared microscopy, an infrared spectrometer is coupled to a light microscope, and thin sections of tissue examined with a spatial resolution of as little as 7 μm . Raman microscopy, which has a spatial resolution of as little as 1 μm , can provide complimentary information.^[18] Both these techniques require sections of tissues to be provided for analyses, but while the Raman analyses can be performed on hydrated tissues of variable thickness, the infrared analyses are affected by the presence of water, and require thin (2–4 μm) sections to allow light to be transmitted. Both imaging techniques provide information on the spatial distribution of mineral (and matrix) properties.

Fig. 4 illustrates the application of Fourier transform infrared imaging (FTIRI) to the study of biopsies obtained from human iliac crest bone. The images show the mineral/matrix ratio in (A) trabecular and

(C) cortical bone, along with the pixel distributions corresponding to these images. Using the same spectral data, other parameters can be determined, and Fig. 4B and 4D show the crystallinity distributions for the same trabecular and cortical bone. These data were from a biopsy of a premenopausal woman, with no evidence of bone disease as described elsewhere.^[19] As shown here, in normal healthy bone, there is heterogeneity in the distribution of the mineral properties, with the most recently formed bone having the lowest mineral content, the highest acid phosphate content, the lowest crystallinity, and the lowest carbonate content.^[18,20,21] With age, in the absence of disease, these parameters are altered, such that older bone has higher mineral content, lesser carbonate and acid phosphate contents, and increased crystallinity. However, there normally is a broad distribution of these parameters. In diseases such as osteoporosis, the mineral/matrix ratio is decreased (pixel histograms shift to the left), and the crystallinity of the residual bone is shifted to the right.^[21–24]

Composition—The Organic Matrix

The organic matrix of bone consists predominately of type I collagen (the structured template upon which mineral crystals are deposited in an oriented fashion). Collagen is a unique triple helical molecule that provides the tissue with strength and flexibility.^[25] There are more than 27 types of collagen; however, type I collagen is the most abundant, and is the major constituent of the organic matrix of bone, tendons, lung, dentin, ligaments, and skin.^[26] The collagen gene is one of the more complex in the body; more than 50 intervening sequences have to be spliced out for each of the three triple helical chains and the mRNA respliced, thus although the structure of the protein is quite simple (Glycine-X-Y) $_n$, with X and Y being proline and hydroxyproline most frequently processing of the collagen protein is extremely complex. There are numerous posttranslational modifications (synthesis of hydroxyproline and hydroxylysine, glycosylation of hydroxylysine and collagen propeptides, assembly of the triple helix, and removal of non-helical domains). Once the collagen fibrils are secreted from the cells, they are stabilized by the formation of cross-links, a process that is partially dependent on enzymes and partially non-enzymatic.^[25] This complexity leaves room for errors, and there are various diseases associated with bone fragility, which are either owing to, or associated with, mutations in the collagen gene,^[27–30] improper processing,^[31,32] or impaired collagen degradation.^[33]

The next most abundant components of the organic matrix of bone are a group of proteins that can be extracted without removing the collagen, the so-called

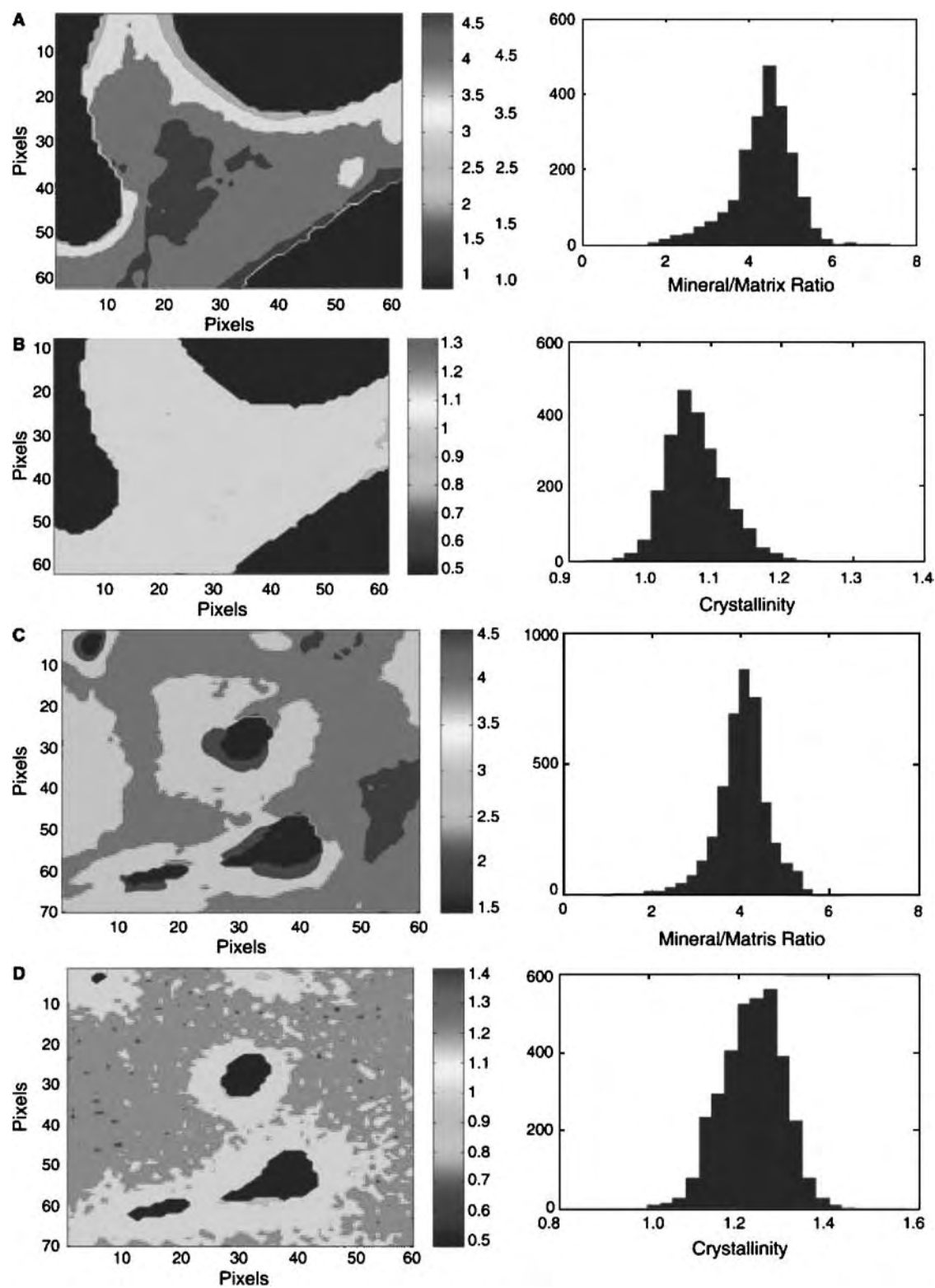


Fig. 4 Infrared imaging of bone in a human iliac biopsy. (A) Mineral/matrix ratio in trabecular bone; (B) crystallinity in this same section; (C) mineral/matrix ratio in cortical bone; and (D) crystallinity in the same cortical section. Pixel histograms are presented on the right of each image describing the distribution of parameters in the image shown.

non-collagenous proteins.^[34] These proteins are highly anionic, interact with collagen, and while these only account for about 5% of the total tissue weight, are involved in regulating the properties of the matrix, and controlling both the initial deposition of the mineral and the growth of the mineral crystals.^[35] Many of these proteins (Table 2) associate with collagen fibrils, are cell binding, and can act as growth factors to stimulate cellular responses. The properties of these proteins have been reviewed extensively elsewhere,^[33,34] and the major components of the bone matrix, along with the effect of these proteins on mineralization in solution, in culture, and/or in animals in which the proteins have been genetically ablated are summarized in Table 2. In addition to the non-collagenous matrix proteins, there are also cell-bound enzymes and intracellular enzymes that are crucial for the function, remodeling, and turnover of the tissue. These enzymes are listed in Table 3. Also in the matrix of bone, in addition to proteins and glycoproteins, are lipids derived from cell membranes, some of which associate with collagen,^[44] and some are components of other extracellular organelles. These lipids serve to regulate cell function and cell shape, and also play roles in mineralization.^[44]

Collagen provides the oriented template upon which the mineral crystals deposit. This deposition is regulated by the cells, which control the flux of calcium and phosphate ions, and by the non-collagenous proteins, which serve as nucleators, and facilitate the deposition of the apatite mineral crystals onto the collagen.

Bone Functional Properties

The most important functional property of bone from an engineering (or health) point of view is its ability to support loads. This load-bearing capacity is dependent on the composite structure^[1] and on the heterogeneity of these structures. Bone strength is determined by a combination of geometric (architecture, number of struts, and wall thickness), and by material properties including the presence of microcracks, mineral content, and crystallite size.^[45] Neither the decalcified organic matrix nor the deorganified mineral matrix has the ability to withstand loads comparable to the composite tissue. It is for this reason that bone tissue engineering seeks to develop scaffolds that can mimic the properties of the healthy bone composite. To make tissue-engineered constructs (scaffold and cells), one requires an understanding of the bone formation process.

Table 2 Organic matrix components^a in vertebrate bone, dentin, and cementum

Protein	Proposed function	Knockout phenotype
Collagen type I	Structural stability/mechanical strength	Brittle bones
Bone sialoprotein	Nucleation of HA	?
Osteopontin	Inhibition of mineralization/ osteoclast recruitment	Impaired remodeling
Dentin matrix protein-1	HA nucleator/cell signaling	Osteomalacia
MEPE	Regulator of mineralization; cell signaling	?
Osteonectin	Regulates collagen fibrillogenesis	Thinned bones, decreased bone density, cataracts, impaired wound healing
Phosphophoryn ^b (PP)	Initiation of dentin mineralization/ cell signaling	?
Dentin sialoprotein ^a (expressed as <i>dspp</i>)	Cell signaling, HA nucleation	Impaired enamel formation; impaired dentinogenesis
Osteocalcin	Recruitment of osteoclasts; inhibitor of mineralization	Increased bone density
Matrix gla-protein	Mineralization inhibitor	Extensive vascular and smooth muscle calcification
Biglycan	Regulates collagen fibrillogenesis; HA nucleator	Thin bones
Decorin	Regulator of collagen fibrillogenesis; inhibits HA formation	?

^aThis list is not comprehensive and only includes those proteins for which functions have been described.

^bFound in dentin only.
(From Ref. [4].)

Table 3 Examples of enzymes that regulate the matrix composition of bones and teeth

Enzyme	Substrate	Function
Alkaline phosphatase ^[36]	Organic phosphate esters	Stimulate mineralization
ANK progressive ankylosis ^[37]	Membrane transport	Transport pyrophosphate
PC-I ^[38]	Pyrophosphate	Hydrolyze inhibitor of mineralization
Casein kinase II ^[39]	Serine rich peptides	Phosphorylate matrix proteins
Matrix metalloproteases ^[40]	Collagen, proteoglycans, matrix proteins	Tissue remodeling
Cathepsin K ^[41]	Extracellular matrix	Osteoclast enzyme that degrades collagenous matrix
Acid phosphatase ^[42,43]	Matrix phosphoproteins	Matrix degradation by osteoclasts and chondroclasts
Carbonic anhydrase ^[43]	Carbon dioxide	Transports protons from cells—important for mineral degradation



Bone Formation

Bones growth occurs by two distinct processes, endochondral ossification and intramembranous ossification. In the first pathway, mesenchymal cells differentiate into chondrocytes and develop a cartilage matrix that is then mineralized and eventually replaced by the bone. These processes occur under the influence of a variety of hormones and growth factors reviewed elsewhere.^[46,47] The cells deposit a matrix consisting, in order of appearance as best we know today, of fibronectin, type I collagen, and non-collagenous proteins. The calcified cartilage that is formed is removed by specialized cells (osteoclasts); bone-forming osteoblasts then deposit a matrix of type I collagen and non-collagenous proteins (osteoid). Key to this process is the invasion of blood vessels, which is why vascular endothelial growth factor (VEGF) is viewed as crucial in the scheme of bone development.^[48,49] Intramembranous ossification or appositional bone growth occurs on an already formed bone matrix, and includes the coupling of resorption and bone formation. Many of the same factors that regulate endochondral ossification also play a role in intramembranous ossification, but the controlling factors are not identical.^[50] In both cases, the final bone composite is similar, although, as noted above, there are slight variations in bone mineral crystal size and composition, depending on the tissue location and age, and on the mechanism of formation.

The first step in bone formation is the formation of a matrix (osteoid in bone and cartilage in the growth plate). This is followed by maturation of both cells and matrix, including modification of the matrix to prepare it for mineral deposition. Because the properties of the mineral crystals and the mechanical properties of the tissues in which they deposit differ, there must be

distinct factors that regulate mineralization in different tissues. However, the general principles are similar. Once an appropriate matrix (which includes the template upon which the crystals will be deposited) is present, a variety of factors lead to the formation of the first insoluble mineral particles (nuclei). Nucleation may occur homogeneously, a very slow process requiring a great deal of energy to arrange the component ions in a structure that will persist as a stable “critical nucleus,” or may occur heterogeneously, on surfaces that resemble the structure of the crystals to be formed. In bone, within which crystals deposit at relatively the same time within multiple sites throughout the entire collagenous matrix, it is thought that the process must be heterogeneous. The collagen, as stated above, provides the template. Non-collagenous matrix proteins associated with the collagen (either on its surface or within its fibrils) can stabilize clusters of ions as they organize to form HA, or, because of their anionic nature, can actually serve as a substrate upon which the ions deposit. Once the HA nuclei are present, when the ion concentration in the solution remains above the solubility of the crystal (i.e., the solution is supersaturated), additional ions can add to the surface of the nuclei, and the crystals can grow in all dimensions, can form branches that support further growth, and can clump together (agglomerate). The same matrix proteins that stabilized the critical nuclei, or served as surfaces for initial crystal deposition, also are involved in the regulation of the shape, proliferation, and agglomeration of the crystals.

Tissue Engineering of Bone

While bone repairs itself by recapitulating the process of endochondral ossification^[51] or intramembranous

ossification, there are often situations where this process fails. Failure of the bone to repair may be because of the presence of neoplasms, extensive trauma, or loss of the vascular supply. In these cases, tissue engineers^[52] working with orthopaedic surgeons seek to replace the bones with scaffolds made of inorganic matrix materials, organic matrix materials, or composites that combine one or more organic matrix component with an inorganic component (Table 4).

The concept is that the scaffold will serve as a substitute template and allow the ordered deposition of mineral. However, in many cases, the inorganic materials do not mimic the poorly crystalline HA of bone, but contain larger crystals. The matrix may consist of a single component or multiple components. Composites with collagen and HA mineral are also being evaluated^[70–73] as substrates for growing bone cells. The inorganic matrix materials used include sintered HA, ceramics, mixtures of HA with bioglass, and mixtures containing other calcium phosphates. None of these has the crystallite properties of bone mineral and none to date has been reported to produce a structure with the mechanical strength of healthy bone. According to Mistry and Mikos^[74] “an ideal

biomaterial scaffold will provide mechanical support to an injured site and also deliver growth factors and cells into a defect to encourage tissue growth” while degrading in a controlled manner “without causing a significant inflammatory response.” Understanding the properties of the healthy bone mineral and matrix should help provide such a material.

Many of these scaffolds are combined with growth factors and cytokines to stimulate the influx of cells. In some cases, the growth factors are adsorbed to the scaffold matrix, in others, cells engineered to express the growth factor have been added. These factors have included, among others, bone morphogenetic proteins, parathyroid hormone, VEGF, and transforming growth factor β (TGF β).^[75,76] The challenges to using these approaches for engineering bone have been reviewed in detail elsewhere,^[77] and include limitations on cell source, problems with developing a scaffold or implant with sufficient porosity that lacks an immunogenic response, the need for a material whose mechanical properties will eventually match those of the tissue to which they are inserted, and the necessity of having a vascular network in the tissue. While most of these materials work well in cell culture experiments,

Table 4 Examples of templates being used for the development of biominerals

Template	Soluble components	Phase formed
Natural Mineral		
HA ^[53]	Simulated body fluid	HA
Polycaprolactone-HA composite ^[54]	Culture media	HA
Amino acid-coated HA nanoparticles in dextran sponge ^[55]		
HA–amelogenin composite ^[56]	Culture media	HA
Calcium carbonate ^[57]	Mollusk proteins	Aragonite (CaCO ₃)
Alumina ^[58]	Osteoblast culture	HA
Collagen		
Patterned Films ^[59]	Fibrinogen and cell culture media	HA
DMP-1 on Collagen ^[60]	Calcium and phosphate ions	HA
Polymers		
PolyHEMA–carboxylate surface ^[61]	Calcium and phosphate ions	HA
Poly(β -hydroxybutyrate-co- β -hydroxyvalerate) ^[62]	Culture media	HA
1-vinyl-2-pyrrolidinone– <i>n</i> -butyl methacrylate co-polymer ^[63]	In vivo	HA
Nanofibrils of L-lactic acid ^[64]	Culture media	HA
Oligolactone macromers ^[65]	Culture media	HA
Porphyrins on glass ^[66]	Calcium and bicarbonate ions	Calcite (CaCO ₃)
Carbon dioxide ^[67]	Carboxymethyl chitosan	Vaterite (CaCO ₃) in flower shaped clusters
Chitin ^[68]	Amino acids in vivo	HA
Chitosan–gelatin composite (\pm HA) ^[69]	Culture medium	HA

their effects in animal models and humans are very variable, and few have been evaluated *in vivo*.

Based on knowledge of the organic components involved in mineralized bone formation, recently there have been some successful attempts to produce biomimetic scaffolds that act like the collagen template, and yield a relative oriented mineralized construct *in vitro*. One of the most promising is the use of scaffolds made of self-assembling molecules by Stupp's group.^[64,78,79] Taking advantage of the need to have a construct with the adsorbed protein-providing sites for cell anchorage, nutrition, signaling to the cells, as well as regulation of mineralization of the scaffold, Stupp's group forms self-assembled biodegradable poly-L-lactic acid fibers bonded by mechanical compression along with a variety of peptides and hydrophobic components. One type of fibrils containing a long alkyl tail, a flexible region, and a single phosphoserine residue was able to orient HA deposition along their long axes *in vitro*.^[64] These materials have not yet been evaluated in animal models, but clearly demonstrate the advantage of providing a mimetic substrate that can serve as a template for HA deposition. Several of the polymers have a cholesteryl component, L-lactic acid, and a flexible protein to which other functional groups can be linked. Another group has used a similar approach with poly (DL-lactic acid) tethered with cholesterol to promote osteoblast attachment and growth *in vitro*.^[80] Song, Malathong, and Bertozzi^[61] similarly used a polymeric hydrogel (polyHema) functionalized with glutamate, aspartate, or phosphoserine. Each of these anionic groups was equally effective in promoting *in vitro* HA deposition.^[61] Table 4 lists additional materials that have been used as templates to form mineral composites and unique structures that have the potential for use in repair of mineralized tissues or induction of new tissue

formation, along with other applications far from biomedical.

A quite different approach to the use of scaffolds was recently reported,^[81] in which the problems of immunogenicity, lack of a vascular supply, and shortage of cell supplies were circumvented by using a space within the host bone as an "in vivo bioreactor." The bioreactor is formed within a space between the long bone and the periosteum that covers it. Using a calcium-rich alginate gel implanted in a space created below the periosteum of rabbits, compact bone formation (via intramembranous ossification) was reproducibly observed. The size of the new bone was dependent on the space into which the alginate was implanted. The implant material functioned equally well with or without growth factors, yielding within six weeks compact bone containing osteons surrounded by lamellar bone. More importantly, the new bone could easily be retrieved and implanted in other sites in the rabbit where it healed defects successfully.

Although many of these tissue-engineered constructs are still in the development stage, there are some already available commercially (Table 5) for repair of bone defects. These are subject to the same limitations as the experimental constructs described above, cell source, immunogenicity, vascularization, and porosity, and thus have restricted usages.

TEETH

Dentin—Composition

Dentin is similar but not identical in composition to bone.^[83] The mineral of dentin is HA, the crystals tend to be slightly larger, and the relative amount of mineral

Table 5 Examples of products for bone repair made from organic and inorganic matrices

Material	What is in it	Company and web location for further detail
Biocoral bone graft substitute	97% calcium aragonite	Biocoral (www.biocoral.com)
Bio Glass	$\text{Na}_2\text{O}-\text{CaO}-\text{CaF}_2-\text{P}_2\text{O}_5-\text{SiO}_2$	NovaBone products (www.novabone.com)
Bio-Oss	Inorganic bone matrix	Geistlich Biomaterials (www.geistlich.de)
Bio-gide	Demineralized bone matrix	
Grafton	Banked bone	Osteotech (www.osteotech.com)
Dynagraft	Demineralized bone matrix + bone morphogenetic protein ^[82]	Citagenix (www.citagenix.com)
Vitoss	Beta-tricalcium phosphate	Orthovita (www.orthovita.com)
Chronos		Synthes (www.stratec.com)
Healos	Cross-linked type I collagen microfibers coated with HA particles	DePuy (www.acumed.com)
Osteoset	Calcium sulfate	Wright Medical (www.wmt.com)



tends to be slightly greater than that found in bone. Dentin has most of the same organic matrix constituents as bone (Table 2), but their relative distribution is different, and dentin has some unique components reviewed elsewhere.^[84] The most unique are the phosphorylated dentin proteins, and, in particular, PP. PP and dentin sialoprotein, another dentin-specific protein, are expressed by a single gene (*dspp*), but the gene products are not colocalized.^[85] Unlike bone, dentin is formed directly by dentin-forming cells (odontoblasts), and in most circumstances, dentin is not remodeled. Dentin, like bone, is important for mechanical strength, and the disruption of its components (for example, the knockout of the *dspp* gene) results in abnormal dentition with impaired mechanical properties, in part attributable to the loss of the enamel coating the dentin.^[86]

Dentin Formation

Dentin, like bone, is of mesenchymal origin. The dentin-forming cells, odontoblasts, deposit their matrix around the epithelial tissue in the jaw, forming the tooth.^[87] As in bone, the mesenchymal cells differentiate, and then synthesize the organic matrix (predentin). There is progressive expression of dentin-specific phosphoproteins, and the deposition of initial mineral crystals, followed by crystal growth and proliferation.^[88] While there is little remodeling, there is an exchange of information between the epithelially derived ameloblasts (which make the enamel) and the odontoblasts. While few dentin-specific matrix components have been used for tissue engineering of other tissues, both PP^[89] and dentin matrix protein 1^[90] have promise for bone repair. A composite sponge consisting of PP and type I collagen was more effective in inducing bone formation in a rat model than collagen alone,^[89] and dentin matrix protein 1, which acts as a HA nucleator, can self-assemble, offering potential for nanoparticle development. A composite of an enamel protein (amelogenin) and HA has also been shown in culture to induce bone-like HA formation.^[56]

Dentin Tissue Engineering

Tissue Engineering of dentin to date has had limited success.^[87] While cells seeded on appropriate synthetic or natural organic, inorganic, or composite matrices can form mineralized structures, developing a structure that has the both shape and properties of an intact tooth has not yet been achieved.^[91] Recent reports indicate that cells cultured in unique dentin environments do express the *dspp* gene, unique to odontoblasts^[92] and express enamel markers. Various

components of the tooth (periodontal ligament^[93] and cementum,^[94,95] etc.) are being developed in vitro.

Enamel—Composition

Covering the surface of the dentin, the enamel composite is a much harder material.^[96] It consists mainly of HA, while the organic matrix is distinct from that of bone and dentin, containing amelogenins, enamelines, and tuftelins but no collagen.^[97,98] Adult enamel, unlike bone, has less than 3% organic matrix. Amelogenins are believed to be the initial mineral nucleators, and enamelines and tuftelins are thought to regulate the growth and orientation of the crystals.^[99] While X-ray diffraction shows that enamel is mainly HA, there is some suggestion that octacalcium phosphate may be the first phase formed in enamel.^[97] The bulk of the mineral is apatitic, and AFM has demonstrated that the crystals form long ribbon-like structures about 1 nm in diameter.^[99] These crystals are much larger crystallite size than that of dentin and bone, and have a different type of carbonate substitution.^[100] The dentin-enamel composite provides the tooth with enhanced mechanical strength—the ability to bite, grind, and chew.

Enamel—Formation

Formation of enamel (amelogenesis) is guided by specialized cells (ameloblasts).^[100] These cells, of epithelial origin, secrete amelogenin nanospheres,^[101] which control the formation of the initial HA crystals, and amelogenin and tuftelin, which help orient the crystals and regulate their growth.^[98] The details of amelogenesis have been reviewed elsewhere,^[102] but are a process quite distinct from the formation of bone and dentin. In a recent in vitro study,^[103] both full-length and a shortened (without the hydrophilic C-terminus) amelogenin inhibit crystal growth in directions other than length (the *c*-axis). In these studies, the crystals organized into parallel arrays only in the presence of the full-length amelogenin in monomeric rather than nanoparticle form. In contrast, preassembled amelogenins had no effect on organization of the crystals, suggesting that the hydrophobic portion of amelogenin regulates crystal growth while the C-terminal domain is essential for the alignment of crystals into parallel arrays. Amelogenin nanospheres thus appear to be acting both as a template and as a regulating soluble component that helps determine the ribbon-like structure of the enamel crystals. This information will be very important for designing materials for repair and regeneration of enamel.

Enamel Tissue Engineering

Enamel repair is a common process in the dental office, and is done by replacing the damaged components, either with other calcium phosphates or with HA itself.^[87,104] However to tissue engineer a tooth, the construct will have to allow for both enamel formation and dentin and cementum formation, thus there are numerous efforts to recapitulate amelogenesis.^[87] The challenge is to engineer a structure that is so distinct from those formed in other mineralized matrices, thus success has been fairly limited. Recent data indicate that enamel-like crystals can be prepared using HA scaffolds.^[105]

SHELLS

Composition

The inorganic matrix of most shells consist of calcium carbonates,^[5,106–108] while their organic matrices may include the polysaccharide chitin,^[109] a series of anionic proteins,^[108] and, in rare cases, collagen-like proteins.^[110] Four different phases of calcium carbonate may deposit, including an amorphous (non-crystalline form), calcite, aragonite, and vaterite.^[6,107,108] Some shells have mixtures of these phases.

These inorganic phases deposit on organic matrices that are highly anionic, and it is the anionic matrix that determines the site of formation of the mineral crystals, the phase formed, and their orientation.^[107,108,111–114] The nature of the phase deposited is determined by the composition of the organic matrix, and extracts of the different proteins from mollusks have been shown to determine when and if calcite or aragonite deposits.^[113,114] The previously described concept that anionic molecules provide the “template” for deposition of mineral ions was first introduced from studies of shells by Weiner and Hood.^[115] Studies of shell proteins and how they regulate mineral nucleation and growth have provided insight into the role of acidic macromolecules in biomineralization in general, and, as discussed in previous sections, have also led to a new industry where biomimetic processes are being used to engineer bone.

Tissue Engineering with Shell Components

Chitin is an abundant component of certain calcium carbonate-containing shells^[109,116] and as it is an inert polysaccharide it has been used for a variety of tissue-engineering applications including drug delivery,^[117] cartilage repair,^[42] and most recently bone repair.^[118] While coral and other inorganic matrices of shells have

been used as a bone graft material,^[119,120] few organic matrix components of shells have been tested as tissue-engineering constructs. Other components of the organic matrices of shells have not yet been tested for tissue repair and regeneration, but as suggested elsewhere,^[120] the potential for using such matrices remains to be developed.

CONCLUSIONS

The inorganic and organic matrices of mineralized tissues in vertebrates and invertebrates share several common features. The inorganic matrices often are similar, and their formation and expansion are controlled by anionic components of the organic matrices. These anionic components associate with or act themselves as templates for oriented mineral deposition. Both the inorganic and the organic matrix have potential for use in tissue repair and regeneration.

ACKNOWLEDGMENTS

Dr Boskey’s research as described in this article was supported by NIH grants DE04141, AR037661, and AR043125.

ARTICLES OF FURTHER INTEREST

Alumina; Biologic and Synthetic Apatites; Composites; Polymers; Zirconia Ceramics

REFERENCES

1. Burr, D.B. The contribution of the organic matrix to bone’s material properties. *Bone* **2002**, *31* (1), 8–11.
2. Hamm, C.E. The evolution of advanced mechanical defenses and potential technological applications of diatom shells. *J. Nanosci. Nanotechnol.* **2005**, *5* (1), 108–119.
3. Landis, W.J. An overview of vertebrate mineralization with emphasis on collagen-mineral interaction. *Gravit. Space. Biol. Bull.* **1999**, *12* (2), 15–26.
4. Boskey, A.L. Bone mineralization. In *Bone Biomechanics*, 3rd Ed.; Cowin, S.C., Ed.; CRC Press: Boca Raton, FL, 2001; 5.1–5.34.
5. Lowenstam, H.A., Weiner, S., Eds.; *On Biomineralization*; Oxford University Press: Oxford, 1989.
6. Weiss, I.M.; Tuross, N.; Addadi, L.; Weiner, S. Mollusc larval shell formation: amorphous calcium carbonate is a precursor phase for aragonite. *J. Exp. Zool.* **2002**, *293* (5), 478–491.
7. Roschger, P.; Plenk, H., Jr.; Klaushofer, K.; Eschberger, J. A new scanning electron microscopy approach to



- the quantification of bone mineral distribution: back-scattered electron image grey-levels correlated to calcium K alpha-line intensities. *Scanning Microsc.* **1995**, *9* (1), 75–86.
8. Boivin, G.; Meunier, P.J. The degree of mineralization of bone tissue measured by computerized quantitative contact microradiography. *Calcif. Tissue Int.* **2002**, *70* (6), 503–511.
 9. Cummings, S.R.; Bates, D.; Black, D.M. Clinical use of bone densitometry: scientific review. *JAMA* **2002**, *288* (15), 1889–1897.
 10. Ding, M.; Day, J.S.; Burr, D.B.; Mashiba, T.; Hirano, T.; Weinans, H.; Sumner, D.R.; Hvid, I. Canine cancellous bone microarchitecture after one year of high-dose bisphosphonates. *Calcif. Tissue Int.* **2003**, *72* (6), 737–44.
 11. JCPDS. Joint Committee on Powder Diffraction Standards. Annual Book of ASTM Standards.
 12. Boskey, A.L. Mineral Analysis provides insights into the mechanism of biomineralization. *Calcif. Tissue Int.* **2003**, *72* (5), 533–536.
 13. Boskey, A.; Spevak, L.; Tan, M.; Doty, S.B.; Butler, W.T. Dentin sialoprotein (DSP) has limited effects on in vitro apatite formation and growth. *Calcif. Tissue Int.* **2000**, *67* (6), 472–478.
 14. Frank, R.M.; Klewansky, P.; Hemmerle, J.; Tenenbaum, H. Ultrastructural demonstration of the importance of crystal size of bioceramic powders implanted into human periodontal lesions. *J. Clin. Periodontol.* **1991**, *18* (9), 669–680.
 15. Tong, W.; Glimcher, M.J.; Katz, J.L.; Kuhn, L.; Eppell, S.J. Size and shape of mineralites in young bovine bone measured by atomic force microscopy. *Calcif Tissue Int.* **2003**, *72* (5), 592–598.
 16. Zhang, Y.; Cheng, F.; Li, D.; Wang, Y.; Zhang, G.; Liao, W.; Tang, T.; Huang, Y.; He, W. Investigation of elemental content distribution in femoral head slice with osteoporosis by SRXRF microprobe. *Biol. Trace Elem. Res.* **2005**, *103* (2), 177–185.
 17. Boskey, A.L.; Mendelsohn, R. Infrared analysis of bone in health and disease. *J. Biomed. Optics* **2005**, *10*(031102-1 - -9).
 18. Tarnowski, C.P.; Ignelzi, M.A., Jr.; Morris, M.D. Mineralization of developing mouse calvaria as revealed by Raman microspectroscopy. *J. Bone Miner. Res.* **2002**, *17* (6), 1118–1126.
 19. Faibish, D.; Gomes, A.; Boivin, G.; Binderman, I.; Boskey, A. Infrared imaging of calcified tissue in bone biopsies from adults with osteomalacia. *Bone*. **2005**, *36* (1), 6–12.
 20. Paschalis, E.P.; Betts, F.; DiCarlo, E.; Mendelsohn, R.; Boskey, A.L. FTIR microspectroscopic analysis of normal human cortical and trabecular bone. *Calcif. Tissue Int.* **1997**, *61* (6), 480–486.
 21. Gadeleta, S.J.; Boskey, A.L.; Paschalis, E.; Carlson, C.; Menschik, F.; Baldini, T.; Peterson, M.; Rinnac, C.M. A physical, chemical, and mechanical study of lumbar vertebrae from normal, ovariectomized, and nandrolone decanoate-treated cynomolgus monkeys (*Macaca fascicularis*) *Bone* **2000**, *27* (4), 541–550.
 22. Paschalis, E.P.; Betts, F.; DiCarlo, E.; Mendelsohn, R.; Boskey, A.L. FTIR microspectroscopic analysis of human iliac crest biopsies from untreated osteoporotic bone. *Calcif. Tissue Int.* **1997**, *61* (6), 487–492.
 23. Paschalis, E.P.; Boskey, A.L.; Kassem, M.; Eriksen, E.F. Effect of hormone replacement therapy on bone quality in early postmenopausal women. *J. Bone Miner. Res.* **2003**, *18* (6), 955–959.
 24. Boskey, A.L.; DiCarlo, E.; Paschalis, E.; West, P.; Mendelsohn, R. Comparison of mineral quality and quantity in Iliac crest biopsies from high-turnover and low-turnover osteoporosis: an FT-IR Microspectroscopic Invest. *Osteopor. Int.* **2005**, Aug 9, ePub ahead of print.
 25. Yamauchi, M. Collagen biochemistry: an overview. *Advances in tissue banking* **2002**, *6*, 445–500.
 26. Eyre, D.R. Collagens and cartilage matrix homeostasis. *Clin. Orthop. Relat. Res.* **2004**, *427*, Suppl S118–S122.
 27. Dalgleish, R. The human type I collagen mutation database. *Nucleic Acids Res.* **1997**, *25* (1), 181–187.
 28. Byers, P.H. Molecular genetics of chondrodysplasias, including clues to development, structure, and function. *Curr. Opin. Rheumatol.* **1994**, *6* (3), 345–350.
 29. Prockop, D.J.; Kuivaniemi, H.; Tromp, G. Molecular basis of osteogenesis imperfecta and related disorders of bone. *Clin. Plast. Surg.* **1994**, *21* (3), 407–413.
 30. Niyibizi, C.; Wang, S.; Mi, Z.; Robbins, P.D. Gene therapy approaches for osteogenesis imperfecta. *Gene Ther.* **2004**, *11* (4), 408–416.
 31. Germain, D. Ehlers-Danlos syndromes. Clinical, genetic and molecular aspects. *Ann. Dermatol. Venereol.* **1995**, *122* (4), 187–204.
 32. Bailey, A.J.; Wotton, S.F.; Sims, T.J.; Thompson, P.W. Post-translational modifications in the collagen of human osteoporotic femoral head. *Biochem. Biophys. Res. Commun.* **1992**, *185* (3), 801–805.
 33. Krane, S.M. Elucidation of the potential roles of matrix metalloproteinases in skeletal biology. *Arthritis Res. Ther.* **2003**, *5* (1), 2–4.
 34. Fisher, L.W.; Termine, J.D. Noncollagenous proteins influencing the local mechanisms of calcification. *Clin. Orthop. Relat. Res.* **1985**, *200*, 362–385.
 35. Gokhale, J.; Robey, P.G.; Boskey, A.L. The biochemistry of bone. In *Osteoporosis*, 2nd Ed.; San Diego: Robert Marcus, David Feldman, Jennifer Kelsey, Eds.; Academic Press, 2001; Vol. 1, 107–189.
 36. Bonucci, E.; Nanci, A. Alkaline phosphatase and tartrate-resistant acid phosphatase in osteoblasts of normal and pathologic bone. *Ital. J. Anat. Embryol.* **2001**, *106* (2 Suppl. 1), 129–33.
 37. Ho, A.M.; Johnson, M.D.; Kingsley, D.M. Role of the mouse ank gene in control of tissue calcification and arthritis. *Science*. **2000**, *289* (5477), 265–270.
 38. Johnson, K.; Terkeltaub, R. Upregulated ank expression in osteoarthritis can promote both chondrocytes MMP-13 expression and calcification via chondrocyte extracellular PPI excess. *Osteoarthritis Cartil.* **2004**, *12* (4), 321–335.
 39. Veis, A.; Sfeir, C.; Wu, C.B. Phosphorylation of the proteins of the extracellular matrix of mineralized

- tissues by casein kinase-like activity. *Crit. Rev. Oral Biol. Med.* **1997**, *8* (4), 360–379.
40. Mott, J.D.; Werb, Z. Regulation of matrix biology by matrix metalloproteinases. *Curr. Opin. Cell Biol.* **2004**, *16* (5), 558–564.
 41. Drake, F.H.; Dodds, R.A.; James, I.E.; Connor, J.R.; Debouck, C.; Richardson, S.; Lee-Rykaczewski, E.; Coleman, L.; Rieman, D.; Barthlow, R.; Hastings, G.; Gowen, M. Cathepsin K, but not cathepsins B, L, or S, is abundantly expressed in human osteoclasts. *J. Biol. Chem.* **1996**, *271* (21), 12,511–12,516.
 42. Frenkel, S.R.; Bradica, G.; Brekke, J.H.; Goldman, S.M.; Ieska, K.; Issack, P.; Bong, M.R.; Tian, H.; Gokhale, J.; Coutts, R.D.; Kronengold, R.T. Regeneration of articular cartilage—evaluation of osteochondral defect repair in the rabbit using multiphasic implants. *Osteoarthritis Cartil.* **2005**, *13* (9), 798–807.
 43. Blair, H.C.; Athanasou, N.A. Recent advances in osteoclast biology and pathological bone resorption. *Histol. Histopathol.* **2004**, *19* (1), 189–99.
 44. Goldberg, M.; Boskey, A.L. Lipids and biomineralizations. *Prog. Histochem. Cytochem.* **1996**, *31* (2), 1–187.
 45. van der Meulen, M.C.; Jepsen, K.J.; Mikic, B. Understanding bone strength: size isn't everything. *Bone* **2001**, *29* (2), 101–104.
 46. Ballock, R.T.; O'Keefe, R.J. Physiology and pathophysiology of the growth plate. *Birth Defects Res. C Embryo Today* **2003**, *69* (2), 123–143.
 47. Tuan, R.S. Biology of developmental and regenerative skeletogenesis. *Clin. Orthop. Relat. Res.* **2004**, *427*, Suppl. S105–S17.
 48. Colnot, C.I.; Helms, J.A. A molecular analysis of matrix remodeling and angiogenesis during long bone development. *Mech. Dev.* **2001**, *100* (2), 245–50.
 49. Zelzer, E.; Olsen, B.R. Multiple roles of vascular endothelial growth factor (VEGF) in skeletal development, growth, and repair. *Curr. Top. Dev. Biol.* **2005**, *65*, 169–87.
 50. Xue, Y.; Karaplis, A.C.; Hendy, G.N.; Goltzman, D.; Miao, D. Genetic models show that parathyroid hormone and 1,25-dihydroxyvitamin D₃ play distinct and synergistic roles in postnatal mineral ion homeostasis and skeletal development. *Hum Mol Genet.* **2005**, *14* (11), 1515–1528.
 51. Ferguson, C.M.; Miclau, T.; Hu, D.; Alpern, E.; Helms, J.A. Common molecular pathways in skeletal morphogenesis and repair. *Ann. N Y Acad. Sci.* **1998**, *23* (857), 33–42.
 52. Crane, G.M.; Ishaug, S.L.; Mikos, A.G. Bone tissue engineering. *Nat. Med.* **1995**, *1* (12), 1322–1324.
 53. Ni, S.; Chang, J.; Chou, L. A novel bioactive porous CaSiO₃ scaffold for bone tissue engineering. *J. Biomed. Mater. Res. A.* **2005**, Epub ahead of print.
 54. Causa, F.; Netti, P.A.; Ambrosio, L.; Ciapetti, G.; Baldini, N.; Pagani, S.; Martini, D.; Giunti, A. Poly-epsilon-caprolactone/hydroxyapatite composites for bone regeneration: in vitro characterization and human osteoblast response. *J. Biomed. Mater. Res. A.* **2005**, Epub ahead of print.
 55. Gonzalez-McQuire, R.; Green, D.; Walsh, D.; Hall, S.; Chane-Ching, J.Y.; Oreffo, R.O.; Mann, S. Fabrication of hydroxyapatite sponges by dextran sulphate/amino acid templating. *Biomaterials* **2005**, *26* (33), 6652–6656.
 56. Du, C.; Schneider, G.B.; Zaharias, R.; Abbott, C.; Seabold, D.; Stanford, C.; Moradian-Oldak, J. Apatite/Amelogenin coating on titanium promotes osteogenic gene expression. *J. Dent. Res.* **2005**, *84* (11), 1070–1074.
 57. Michenfelder, M.; Fu, G.; Lawrence, C.; Weaver, J.C.; Wustman, B.A.; Taranto, L.; Evans, J.S.; Morse, D.E. Characterization of two molluscan crystal-modulating biomineralization proteins and identification of putative mineral binding domains. *Biopolymers* **2003**, *70* (4), 522–533.
 58. Swan, E.E.; Popat, K.C.; Grimes, C.A.; Desai, T.A. Fabrication and evaluation of nanoporous alumina membrane for osteoblast culture. *J. Biomed. Mater. Res. A* **2005**, *72* (3), 288–295.
 59. Ber, S.; Torun Kose, G.; Hasirci, V. Bone tissue engineering on patterned collagen films: an in vitro study. *Biomaterials* **2005**, *26* (14), 1977–1986.
 60. He, G.; George, A. Dentin matrix protein 1 immobilized on type I collagen fibrils facilitates apatite deposition in vitro. *J. Biol. Chem.* **2004**, *279* (12), 11,649–11,656.
 61. Song, J.; Malathong, V.; Bertozzi, C.R. Mineralization of synthetic polymer scaffolds: a bottom-up approach for the development of artificial bone. *J. Am. Chem. Soc.* **2005**, *127* (10), 3366–3372.
 62. Kumarasuriyar, A.; Jackson, R.A.; Grondahl, L.; Trau, M.; Nurcombe, V.; Cool, S.M. Poly(beta-hydroxybutyrate-co-beta-hydroxyvalerate) supports in vitro osteogenesis. *Tissue Eng.* **2005**, *11* (7–8), 1281–1295.
 63. Jansen, E.J.; Sladek, R.E.; Bahar, H.; Yaffe, A.; Gijbels, M.J.; Kuijjer, R.; Bulstra, S.K.; Guldemond, N.A.; Binderman, I.; Koole, L.H. Hydrophobicity as a design criterion for polymer scaffolds in bone tissue engineering. *Biomaterials* **2005**, *26* (21), 4423–4431.
 64. Hargerink, J.D.; Beniash, E.; Stupp, S.I. Self-assembly and mineralization of peptide-amphiphile nanofibers. *Science* **2001**, *294* (5547), 1684–1688.
 65. Vogt, S.; Berger, S.; Wilke, I.; Larcher, Y.; Weisser, J.; Schnabelrauch, M. Design of oligolactone-based scaffolds for bone tissue engineering. *Bio-Medical Materials and Engineering* **2005**, *15* (1/2), 73–85.
 66. Lahiri, J.; Xu, G.; Dabbs, D.M.; Yao, N.; Aksay, I.A.; Groves, J.T. Protoporphyrin amphiphiles as templates for the nucleation of calcium carbonate. *J. Am. Chem. Soc.* **1997**, *119*, 5449–5450.
 67. Liang, P.; Shen, Q.; Zhao, Y.; Zhou, Y.; Wei, H.; Lieberwirth, I.; Huang, Y.; Wang, D.; Xu, D. Petunia-shaped superstructures of CaCO₃ aggregates modulated by modified chitosan. *Langmuir* **2004**, *20* (24), 10,444–10,448.
 68. Mukherjee, D.P.; Tunkle, A.S.; Roberts, R.A.; Clavenna, A.; Rogers, S.; Smith, D. An animal evaluation of a paste of chitosan glutamate and hydroxyapatite as a synthetic bone graft material. *J. Biomed. Mater. Res B Appl Biomater.* **2003**, *67* (1), 603–609.
 69. Zhao, F.; Grayson, W.L.; Ma, T.; Bunnell, B.; Lu, W.W. Effects of hydroxyapatite in 3-D chitosan-gelatin



- polymer network on human mesenchymal stem cell construct development. *Biomaterials* **2005**, epub ahead of print.
70. Bigi, A.; Torricelli, P.; Fini, M.; Bracci, B.; Panzavolta, S.; Sturba, L.; Giardino, R. A biomimetic gelatin-calcium phosphate bone cement. *Int. J. Artif. Organs* **2004**, *27* (8), 664–673.
 71. Fan, Y.; Duan, K.; Wang, R. A composite coating by electrolysis-induced collagen self-assembly and calcium phosphate mineralization. *Biomaterials* **2005**, *26* (14), 1623–1632.
 72. Hempel, U.; Reinstorf, A.; Poppe, M.; Fischer, U.; Gelinsky, M.; Pompe, W.; Wenzel, K.W. Proliferation and differentiation of osteoblasts on Biocement D modified with collagen type I and citric acid. *J. Biomed. Mater. Res. B. Appl. Biomater.* **2004**, *71* (1), 130–143.
 73. Yamauchi, K.; Goda, T.; Takeuchi, N.; Einaga, H.; Tanabe, T. Preparation of collagen/calcium phosphate multilayer sheet using enzymatic mineralization. *Biomaterials* **2004**, *25* (24), 5481–5489.
 74. Mistry, A.S.; Mikos, A.G. Tissue engineering strategies for bone regeneration. *Adv. Biochem. Eng. Biotechnol.* **2005**, *94*, 1–22.
 75. Schek, R.M.; Taboas, J.M.; Segvich, S.J.; Hollister, S.J.; Kresbach, P.H. Engineered osteochondral grafts using biphasic composite solid free-form fabricated scaffolds. *Tissue Eng.* **2004**, *10* (9–10), 1376–1385.
 76. Holland, T.A.; Bodde, E.W.; Baggett, L.S.; Tabata, Y.; Mikos, A.G.; Jansen, J.A. Osteochondral repair in the rabbit model utilizing bilayered, degradable oligo(poly(ethylene glycol) fumarate) hydrogel scaffolds. *J. Biomed. Mater. Res. A.* **2005**, *75* (1), 156–167.
 77. Logeart-Avramoglou, D.; Anagnostou, F.; Bizios, R.; Petite, H. Engineering bone: challenges and obstacles. *J. Cell. Mol. Med.* **2005**, *9* (1), 72–84.
 78. Hwang, J.J.; Iyer, S.N.; Li, L.-S.; Claussen, R.; Harrington, D.A.; Stupp, S.I. Self-assembling biomaterials: liquid crystal phases of cholesteryl oligo(L-lactic acid) and their interactions with cells. *PNAS* **2002**, *99* (15), 9662–9667.
 79. Stendahl, J.C.; Li, L.; Claussen, R.C.; Stupp, S.I. Modification of fibrous poly(L-lactic acid) scaffolds with self-assembling triblock molecules. *Biomaterials* **2004**, *25* (27), 5847–5856.
 80. Yu, G.; Ji, J.; Shen, J. Cholesterol tethered poly(DL-lactic acid) for promoting osteoblast attachment and growth. *J. Bioact. Compat. Polym.* **2005**, *20* (6), 527–540.
 81. Stevens, M.M.; Marini, R.P.; Schaefer, D.; Aronson, J.; Langer, R.; Shastri, V.P. In vivo engineering of organs: the bone bioreactor. *PNAS USA.* **2005**, *102* (32), 11450–11455.
 82. Cheng, H.; Jiang, W.; Phillips, F.M.; Haydon, R.C.; Peng, Y.; Zhou, L.; Luu, H.H.; An, N.; Breyer, B.; Vanichakarn, P.; Szatkowski, J.P.; Park, J.Y.; He, T.C. Osteogenic activity of the fourteen types of human bone morphogenetic proteins (BMPs). *J. Bone Joint Surg. Am.* **2003**, *85-A* (8), 1544–1552.
 83. Butler, W.T.; Brunn, J.C.; Qin, C. Dentin extracellular matrix (ECM) proteins: comparison to bone ECM and contribution to dynamics of dentinogenesis. *Connect. Tissue Res.* **2003**, *44* (Suppl 1), 171–178.
 84. Butler, W.T.; Ritchie, H. The nature and functional significance of dentin extracellular matrix proteins. *Int. J. Dev. Biol.* **1995**, *39* (1), 169–179.
 85. Butler, W.T.; Brunn, J.C.; Qin, C.; McKee, M.D. Extracellular matrix proteins and the dynamics of dentin formation. *Connect. Tissue Res.* **2002**, *43* (2–3), 301–317.
 86. Sreenath, T.; Thyagarajan, T.; Hall, B.; Longenecker, G.; D'Souza, R.; Hong, S.; Wright, J.T.; MacDougall, M.; Sauk, J.; Kulkarni, A.B. Dentin sialoprophosphoprotein knockout mouse teeth display widened predentin zone and develop defective dentin mineralization similar to human dentinogenesis imperfecta type III. *J. Biol. Chem.* **2003**, *278* (27), 24,874–24,880.
 87. Sharpe, P.T.; Young, C.S. Test-tube teeth. *Scientific American* **2005**, *293* (2), 34–41.
 88. Verdelis, K.; Crenshaw, M.A.; Paschalis, E.P.; Doty, S.; Atti, E.; Boskey, A.L. Spectroscopic imaging of mineral maturation in bovine dentin. *J. Dent. Res.* **2003**, *82*, 697–702.
 89. Iejima, D.; Saito, T.; Uemura, T. A collagen-phosphorylamine sponge as a scaffold for bone tissue engineering. *J. Biomater. Sci. Polym. Ed.* **2003**, *14* (10), 1097–1103.
 90. He, G.; Dahl, T.; Veis, A.; George, A. Nucleation of apatite crystals in vitro by self-assembled dentin matrix protein 1. *Nat. Mater.* **2003**, *2* (8), 552–558.
 91. Young, C.S.; Kim, S.W.; Qin, C.; Baba, O.; Butler, W.T.; Taylor, R.R.; Bartlett, J.D.; Vacanti, J.P.; Yelick, P.C. Developmental analysis and computer modelling of bioengineered teeth. *Arch. Oral. Biol.* **2005**, *50* (2), 259–265.
 92. Deng, M.; Shi, J.; Smith, A.J.; Jin, Y. Effects of transforming growth factor beta1 (TGFbeta-1) and dentin non-collagenous proteins (DNCP) on human embryonic ectomesenchymal cells in a three-dimensional culture system. *Arch. Oral. Biol.* **2005**, *50* (11), 937–945.
 93. Giannobile, W.V.; Somerman, M.J. Growth and amelogenin-like factors in periodontal wound healing. A systematic review. *Ann. Periodontol.* **2003**, *8* (1), 193–204.
 94. Zhao, M.; Jin, Q.; Berry, J.E.; Nociti, F.H., Jr.; Giannobile, W.V.; Somerman, M.J. Cementoblast delivery for periodontal tissue engineering. *J. Periodontol.* **2004**, *75* (1), 154–61.
 95. Bosshardt, D.D. Are cementoblasts a subpopulation of osteoblasts or a unique phenotype?. *J. Dent. Res.* **2005**, *84* (5), 390–406.
 96. Weatherell, J.A.; Robinson, C.; Hallsworth, A.S. Variations in the chemical composition of human enamel. *J. Dent. Res.* **1974**, *53* (2), 180–192.
 97. Aoba, T.; Komatsu, H.; Shimazu, Y.; Yagishita, H.; Taya, Y. Enamel mineralization and an initial crystalline phase. *Connect Tissue Res.* **1998**, *38* (1–4), 129–137.
 98. Deutsch, D.; Dafni, L.; Palmon, A.; Hekmati, M.; Young, M.F.; Fisher, L.W. Tuftelin: enamel mineralization and amelogenesis imperfecta. *Ciba Found Symp.* **1997**, *205*, 135–47.
 99. Kirkham, J.; Brookes, S.J.; Shore, R.C.; Bonass, W.A.; Smith, D.A.; Wallwork, M.L.; Robinson, C. Atomic

- force microscopy studies of crystal surface topology during enamel development. *Connect. Tissue Res.* **1998**, *38* (1–4), 91–100.
100. Robinson, C.; Kirkham, J.; Brookes, S.J.; Bonass, W.A.; Shore, R.C. The chemistry of enamel development. *Int. J. Dev. Biol.* **1995**, *39* (1), 145–52.
 101. Du, C.; Falini, G.; Fermani, S.; Abbott, C.; Moradian-Oldak, J. Supramolecular assembly of amelogenin nanospheres into birefringent microribbons. *Science* **2005**, *307* (5714), 1450–1454.
 102. Sire, J.Y.; Delgado, S.; Fromentin, D.; Girondot, M. Amelogenin: lessons from evolution. *Arch. Oral Biol.* **2005**, *50* (2), 205–212.
 103. Beniash, E.; Simmer, J.P.; Margolis, H.C. The effect of recombinant mouse amelogenins on the formation and organization of hydroxyapatite crystals in vitro. *J. Struct. Biol.* **2005**, *149* (2), 182–190.
 104. Devecioglu, D.; Tozum, T.F.; Sengun, D.; Nohutcu, R.M. Biomaterials in periodontal regenerative surgery: effects of cryopreserved bone, commercially available coral, demineralized freeze-dried dentin, and cementum on periodontal ligament fibroblasts and osteoblasts. *J. Biomater. Appl.* **2004**, *19* (2), 107–120.
 105. Chen, H.; Clarkson, B.H.; Sun, K.; Mansfield, J.F. Self-assembly of synthetic hydroxyapatite nanorods into an enamel prism-like structure. *J. Colloid Interface Sci.* **2005**, *288* (1), 97–103.
 106. Wilt, F.H. Developmental biology meets materials science: morphogenesis of biomineralized structures. *Dev. Biol.* **2005**, *280* (1), 15–25.
 107. Krampitz, G.; Witt, W. Biochemical aspects of biomineralization. *Top. Curr. Chem.* **1979**, *78*, 57–144.
 108. Albeck, S.; Addadi, I.; Weiner, S. Regulation of calcite crystal morphology by intracrystalline acidic proteins and glycoproteins. *Connect. Tissue Res.* **1996**, *35* (1–4), 365–370.
 109. Falini, G.; Fermani, S. Chitin mineralization. *Tissue Eng.* **2004**, *10* (1–2), 1–6.
 110. Wright, G.M.; Keeley, F.W.; Robson, P. The unusual cartilaginous tissues of jawless craniates, cephalochordates and invertebrates. *Cell Tissue Res.* **2001**, *04* (2), 165–174.
 111. Wustman, B.A.; Morse, D.E.; Evans, J.S. Structural characterization of the N-terminal mineral modification domains from the molluscan crystal-modulating biomineralization proteins, AP7 and AP24. *Biopolymers* **2004**, *74* (5), 363–376.
 112. Wustman, B.A.; Weaver, J.C.; Morse, D.E.; Evans, J.S. Structure-function studies of the lustrin A polyelectrolyte domains, RKSX and D4. *Connect Tissue Res.* **2003**, *44* (Suppl. 1), 10–15.
 113. Thompson, J.B.; Palocz, G.T.; Kindt, J.H.; Michenfelder, M.; Smith, B.L.; Stucky, G.; Morse, D.E.; Hansma, P.K. Direct observation of the transition from calcite to aragonite growth as induced by abalone shell proteins. *Biophys. J.* **2000**, *79* (6), 3307–3312.
 114. Walters, D.A.; Smith, B.L.; Belcher, A.M.; Palocz, G.T.; Stucky, G.D.; Morse, D.E.; Hansma, P.K. Modification of calcite crystal growth by abalone shell proteins: an atomic force microscope study. *Biophys. J.* **1997**, *72* (3), 1425–1433.
 115. Weiner, S.; Hood, L. Soluble protein of the organic matrix of mollusk shells: a potential template for shell formation. *Science* **1975**, (4218), 987–989.
 116. Chandy, T.; Sharma, C.P. Chitosan—as a biomaterial. *Biomater. Artif. Cells. Artif. Organs.* **1990**, *18* (1), 1–24.
 117. Aimin, C.; Chunlin, H.; Juliang, B.; Tinyin, Z.; Zhichao, D. Antibiotic loaded chitosan bar. An in vitro, in vivo study of a possible treatment for osteomyelitis. *Clin. Orthop. Relat. Res.* **1999**, *366*, 239–247.
 118. Chen, F.; Chen, S.; Tao, K.; Feng, X.; Liu, Y.; Lei, D.; Mao, T. Marrow-derived osteoblasts seeded into porous natural coral to prefabricate a vascularised bone graft in the shape of a human mandibular ramus: experimental study in rabbits. *Br. J. Oral Maxillofac. Surg.* **2004**, *42* (6), 532–537.
 119. Martina, M.; Subramanyam, G.; Weaver, J.C.; Huttmacher, D.W.; Morse, D.E.; Valiyaveetil, S. Developing macroporous bicontinuous materials as scaffolds for tissue engineering. *Biomaterials.* **2005**, *26* (28), 5609–5616.
 120. Green, D.; Walsh, D.; Mann, S.; Oreffo, R.O. The potential of biomimesis in bone tissue engineering: lessons from the design and synthesis of invertebrate skeletons. *Bone* **2002**, *30* (6), 810–815.



Orthodontic Wires

Eser Tüfekçi

Steven J. Lindauer

Department of Orthodontics, Virginia Commonwealth University, Richmond, Virginia, U.S.A.

INTRODUCTION

During orthodontic treatment, arch wires play an important role because they provide the force systems necessary for tooth movement. Once a wire is engaged into the bracket/band, it exerts forces on the tooth. The forces transmitted to a tooth by an arch wire depend on several parameters of the arch wire used and the relationship between the brackets in which the wire is engaged.

The mechanical force system produced depends on the appliance design and the mechanical and physical properties of the wire used. A variety of factors should be considered in the selection of a wire during a particular phase of the treatment, including the amount of force desired, the elastic range, formability, and ease of soldering or welding.

OVERVIEW

During early stages of treatment, when leveling and alignment are the main goals, a wire capable of delivering light and continuous forces is desirable.^[1] Therefore, initial arch wires are flexible and/or small in diameter, so that maximal wire deflection with light forces may be achieved. Figure 1 shows a highly flexible wire demonstrating high deflection when engaged into the brackets of malpositioned teeth. As the tooth alignment improves, wire diameter is increased by the orthodontist to provide more precise control of tooth positioning. During this stage, stiffer wires provide greater control of tooth movement—essential for the stability of the intra-arch units—while enabling inter-arch corrections (Fig. 2).

As discussed earlier, the force system required for tooth movement is transmitted by the arch wire to the teeth via brackets during orthodontic treatment. The bracket is designed to accommodate the arch wire, and ligation keeps the arch wire engaged in the bracket slot. The faciolingual and occlusogingival dimensions of the bracket slot allow the use of wires with different cross-sectional shapes and sizes. The magnitude of the forces generated in the faciolingual (first-order) and

occlusogingival (second-order) direction is in part dependent on the bracket slot size.

Forces in the faciolingual and occlusogingival directions are produced by direct contact between the wire and the bracket within the slot. Wires with either a rectangular/square or a round cross-section are able to generate couples. Besides the mesiodistal bracket width, the clearance between arch wire and slot also has an effect on the magnitude of these force systems. Wires with larger cross-sections are able to generate higher forces by more closely filling the bracket slot than wires with smaller cross-sections. As the cross-section of a wire is increased, the stiffness is also increased.

A third-order couple is another force system commonly used during orthodontic treatment to make changes in the buccolingual inclination (torque) of a tooth. It is important to note that in orthodontics, the term torque is commonly used to describe tooth inclination. A rectangular arch wire with a diagonal cross-sectional dimension greater than the occlusogingival width of the bracket slot may generate torque through two-edge contact with the slot. Therefore, large rectangular wires may be used to bring teeth to proper inclinations.

Many clinicians use the wire sequence just discussed. However, as they move from smaller, flexible wires to larger, stiffer wires, orthodontists must be aware that wire selection depends on their own treatment philosophy and the needs dictated by the presentation of individual patients.

The mechanical behavior of a ductile orthodontic wire, such as stainless steel, in tensile loading may be analyzed in a stress-strain plot (Fig. 3). Stress is the internal response of a wire to the application of external forces, defined as force (load) per cross-sectional area ($\sigma = F/A$). Strain is the deformation of the arch wire as a consequence of the stress, and is defined as the dimensional change divided by the original dimension ($\epsilon = \Delta d/d$).

The initial linear portion of the stress-strain curve represents the elastic range. The slope of this region of the plot corresponds to the modulus of elasticity (E) and expresses the stiffness of the wire material. It is a measure of the force that can be stored in a wire that

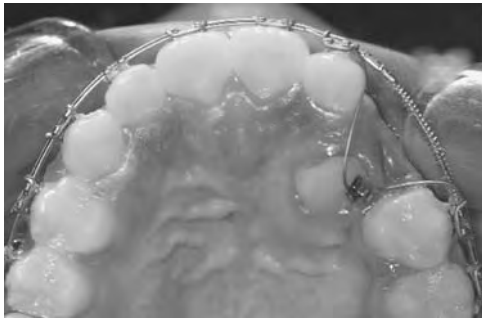


Fig. 1 A nickel-titanium wire tied to an orthodontic attachment showing high deflection.

has been deformed a certain distance. The stress at the end of the linear portion is called the proportional limit. The extreme point on the linear portion of the plot is the elastic limit, from which unloading occurs with no permanent deformation. In crystalline materials, the elastic limit is only slightly beyond the proportional limit and for this reason the two terms are often used interchangeably. If stress is removed from the wire while it is in this portion of the curve, a complete recovery of the initial shape of the wire will occur.

The working range is another important property of an orthodontic wire. It is the maximum deformation (deflection) of an orthodontic wire within its elastic range. In other words, it is the limit of how far a tooth can be moved by one adjustment of a particular wire configuration. Beyond the proportional limit, permanent deformation will take place and the wire will no longer return to its original form when the force is released. Ultimate tensile strength (UTS) is the total load that a wire can withstand before failure. Strength, stiffness, and range are the major properties of clinical importance in the choice of wires to be used. One other important property is the formability of the wire. It describes the amount of permanent deformation that a wire can withstand before it breaks.^[2]



Fig. 2 Stainless steel wire with T-loop bends incorporated to close spaces.

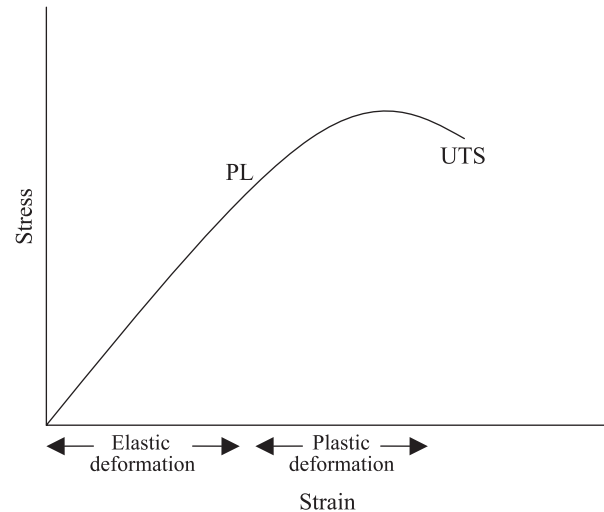


Fig. 3 Stress-strain curve of a ductile material in tensile loading.

The properties of an ideal orthodontic wire are as follows: 1) high strength; 2) low stiffness; 3) high working range; and 4) high formability.^[3] These important characteristics of wires are determined by a variety of factors. Alloy composition and crystal structure of the metal are important variables that are affected by the manufacturing process of the wire, including heat treatment and cold-working. Other factors that influence the mechanical behavior of a wire include the length of the wire across a certain span, or between orthodontic brackets, and the shape and cross-sectional size of the wire.

The effects of the length and cross-section of a beam on strength, stiffness, and range values in bending are shown in Table 1. Based on the engineering bending equation, for example, stiffness of a wire is inversely proportional to the third power of its length (L). Doubling the length of a wire decreases its stiffness to one-eighth.^[4] At the same time, working range is increased four times and strength is halved. The stiffness of a round wire is directly proportional to the fourth power of its cross-sectional diameter (D). Doubling the diameter of a round wire increases

Table 1 Effects of the length and cross-section on the strength, stiffness, and range of orthodontic wires

Strength	Stiffness	Range
<i>Effects of length</i>		
1/L	1/L ³	L ²
<i>Effects of cross-section</i>		
D ³	D ⁴	1/D



its stiffness by 16 times.^[4] At the same time, strength of the wire is increased by eight times, while the working range is decreased by one-half.^[4] It is important for the clinician to be familiar with the mechanical properties of the wires used so that necessary modifications during treatment can be made intelligently. For example, in order to decrease the stiffness and to increase the working range, the length of a span of wire may be controlled by the interbracket distance, which is a function of the distance between the teeth, the position of the brackets on the teeth, and the width of the bracket itself. Incorporating loops and helices into the wire can also increase the effective length of a wire between two brackets, consequently decreasing the stiffness and increasing the working range.

Orthodontic wires are manufactured by a series of proprietary steps.^[5] Usually, more than one company is involved in the manufacturing process. First, the wire alloy is cast in the form of an ingot, which is subjected to several mechanical procedures until the cross-section is small enough for the subsequent wire-drawing process. Drawing of round wires consists of several steps until the desired diameter is obtained. Rectangular wires are manufactured from the round wires by a rolling process, using a Turk's head having two sets of rollers positioned at right angles to each other. Heat treatments are a part of the manufacturing process performed to eliminate the work hardening that takes place during the wire drawing. Detailed information on how the wires are manufactured is proprietary.

In contemporary orthodontic practice, the most commonly used wires include stainless steel, cobalt-chromium-nickel, nickel-titanium, and beta-titanium (Table 2).^[6,7] Precious-metal wires such as gold were commonly used until the 1950s. However, after the introduction of stainless steel into the orthodontic profession, the use of gold wires was abandoned, primarily due to cost factors.^[5,8]

STAINLESS STEEL WIRES

In the 1950s, stainless steel alloy wires became popular because of their good mechanical properties, excellent corrosion resistance, and low cost. Today, stainless steel arch wire is the mainstay of orthodontic practice. The typical composition of the stainless steel alloys used in orthodontics is 73.75% iron, 18% chromium, 8% nickel, and 0.25% cobalt. These wires are in the austenitic phase, and they are usually referred to as 18-8 stainless steel because of the respective percentages of chromium and nickel. Chromium gives stainless steel its corrosion resistance. This element forms a thin, adherent oxide layer on the surface that prevents the reaction of oxygen with the metal below the surface.^[9-11]

The properties of steel wires can be manipulated by heat treatment.^[12] Annealing softens the steel wire, whereas cold-working hardens it. Therefore, annealed steel is very soft and formable. On the other hand, cold-worked wires are very brittle and may break when sharp bends are incorporated into the wire. During the manufacturing process, residual stresses arise which, if not eliminated, may cause breakage of the wire during clinical applications. Therefore, these wires are subjected to heat treatment after manufacturing.

Stainless steel wires may be soldered and welded if solder is used to reinforce the welded parts. This property enables the clinician to fabricate appliances more easily. Both soldering and welding may cause changes in the wire properties if the wire is overheated or underheated. Therefore, extreme caution needs to be taken during these procedures to prevent any decrease in strength of the wire.^[5] When compared to nickel-titanium and beta-titanium arch wires, stainless steel wires exhibit the highest elastic modulus and lowest springback.^[6,13] Although stiffer individually, stainless steel wires that are layered or braided exhibit greater range and lower stiffness than individual strands. That is why these wires were commonly used

Table 2 Important properties of the orthodontic wires

Property	Stainless steel	Cobalt-chromium-nickel	β -titanium	Nickel-titanium
Force levels	High	High	Intermediate	Light
Elastic range	Low	Low	Intermediate	High
Formability	Good	Good	Good	Poor
Ease of joining	Can be soldered; welding only with reinforced solder	Can be soldered; welding only with reinforced solder	Can be soldered and welded	Cannot be soldered nor welded
Friction between the bracket and archwire	Low	Low	Intermediate	High
Cost	Low	Low	High	High

during leveling and alignment before the nickel-titanium wires were available.

COBALT-CHROMIUM ALLOYS

Cobalt-chromium orthodontic wires are very similar to stainless steel in appearance and mechanical properties. Elgiloy (Rocky Mountain Orthodontics), one of the most widely known alloys, has a composition of approximately 40% cobalt, 20% chromium, 15% nickel, 15.8% Fe, 7% Mo 2% Mn 0.15% C, and 0.04 Be.^[6] Due to concerns about the toxicity of beryllium, beryllium-free Cobalt-Chromium alloys (Remaloy, Dentaaurum Inc.) have been introduced to the orthodontic profession.^[5] Cobalt-Chromium wires are available in different tempers (levels of resilience).^[14,15] The soft-temper (Blue Elgiloy) wires are popular among clinicians because they are easily formable and subsequently can be subjected to heat treatment to increase their strength and resilience. These alloys have a modulus of elasticity similar to that of stainless steel,^[6] but they feel softer (prior to the heat treatment) than stainless steel alloys because of their lower yield strength and elastic range.^[6,8]

BETA-TITANIUM WIRES

In 1980, Burstone and Goldberg introduced a beta-titanium alloy to orthodontics under the commercial name of TMA (Ormco, Glendora, CA).^[6] TMA is an acronym for *titanium-molybdenum alloy*. This alloy has approximately 78% titanium and 11% molybdenum in its composition. TMA has a modulus of elasticity approximately 40% that of stainless steel and twice that of Nitinol (nickel-titanium).^[6] This wire has an excellent resilience that provides a wide working range. The wire also exhibits high formability, which allows the clinician to bend the wire and thus incorporate stops and loops into it if desired.^[6,16] One other advantage is that it can also be spot-welded. The properties of TMA make it a wire of choice for fabrication of auxiliary springs and for intermediate and finishing arch wires.

One drawback of this wire is its high coefficient of friction.^[17] This is attributed to the high surface roughness of the wire, confirmed by scanning electron microscopic examinations.^[5,18] Surface roughness should be taken into consideration during the space closure stage of orthodontic treatment, because high friction between the arch wire and the bracket has been shown to slow down the sliding mechanism in vitro. Ion implantation has recently been used to change the surface characteristics of beta-titanium wire improving its ability to enable the use of sliding mechanics.^[19]

TMA wires are the most expensive orthodontic wires available, but their advantages make them a popular wire in current orthodontic practice. Because these wires do not contain nickel in their composition, they may be safely used in patients with a known allergy to nickel.

NICKEL-TITANIUM WIRES

Nitinol was originally developed by Buehler for the U.S. Navy in the early 1960s.^[20] The name *nitinol* was derived from its *nickel-titanium* composition with the suffix *-nol*, which stands for *Naval Ordnance Laboratory*. It was Andreassen whose vision and individual effort led to the introduction of nickel-titanium wires to orthodontics in 1971.^[21] With his collaboration, the first commercial nickel-titanium wire, Nitinol, was marketed by Unitek (3M Unitek, Monrovia, CA). This wire exhibited very important characteristics: 1) a very low modulus of elasticity (E), about one-fifth that of the stainless steel and half that of beta-titanium wires of the same length and cross-sectional dimensions; and 2) an extremely wide working range.^[22] This very first nickel-titanium wire was in a stabilized martensitic form and did not exhibit phase-transition effects. Nevertheless, because of its low modulus of elasticity and high springback, as well as its resistance to deformation, this wire offered a great alternative to braided stainless steel wires during the initial stages of orthodontic treatment.^[23]

In the mid to late 1980s, two new superelastic nickel-titanium alloys—Chinese nickel-titanium (marketed as Ni-Ti by Ormco) and Japanese nickel-titanium (marketed as Sentalloy by GAC International)—were introduced.^[5] The two remarkable properties of the nickel-titanium wires are superelasticity and shape memory.^[24] Nickel-titanium alloy is an equiatomic intermetallic compound that can exist in more than one phase. At low temperatures and high stresses, the alloy is in a martensitic phase with a monoclinic, triclinic, or hexagonal structure. At high temperatures and low stresses, a body-centered cubic (bcc) austenitic phase is present. Both superelasticity and shape memory are related to the transformation within the alloy from the martensitic to the austenitic phase.^[24–26]

Superelasticity is related to the ability of nickel-titanium wires to transform to and from a martensitic phase when subjected to stress.^[24–26] The term pseudoelasticity is also used to describe the wires' superelastic behavior.^[8]

On the stress-strain plot, stress-induced transformation is seen as a plateau where there is not much change in stress with increasing strain (Fig. 4). The austenitic-to-martensitic phase transformation occurs with increasing applied force. The initial linear portion

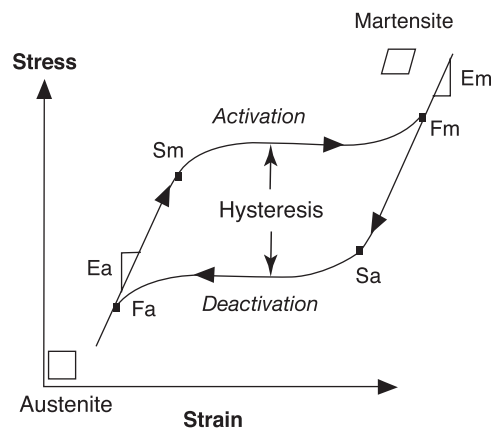


Fig. 4 Stress-strain diagram of an alloy with superelastic behavior. (From Ref. [26].)

of the stress-strain plot represents elastic deformation of the austenitic phase. The slope of this region of the plot is equal to the austenitic modulus of elasticity (E_a). Point S_m is the minimum stress at which martensitic transformation begins. At point F_m , transformation to martensite is complete. Upon unloading, there is a linear relationship between stress and strain. Point S_a is the maximum stress during unloading at which the martensitic structure can still exist, and at this point austenitic transformation begins. With further unloading, at point F_a , the reverse transformation to austenite is complete. While in this region of the curve, these wires exert relatively constant forces independent of activation within a wide range. The force subsequently dissipates once the reverse transformation takes place. The transition temperature is very important, because it determines the force levels that a wire can generate while in the plateau.^[25] By controlling temperature ranges during the manufacturing process, the transition temperature at which the martensitic transformation takes place can be manipulated.^[24-26]

Shape memory is the ability that nickel-titanium wires exhibit to undergo temperature-induced martensitic transformation (Fig. 5).^[24-26] This phenomenon is defined as thermoelasticity.^[8] Thermoelasticity allows a certain shape to be given to a wire at a temperature higher than the martensitic-austenitic transition temperature. When the wire is cooled to a temperature lower than the transition temperature it may be plastically deformed, but when it is reheated it will return to its original shape.

On the stress-strain plot, superelasticity is characterized by a loading curve differing from the unloading curve due to a loss of energy (hysteresis) (Fig. 4). This is a remarkable property, because the force that the wire delivers (unloading portion of the curve) is not the same as the force applied to engage it (loading portion of the curve). In Fig. 4, the second plateau region of

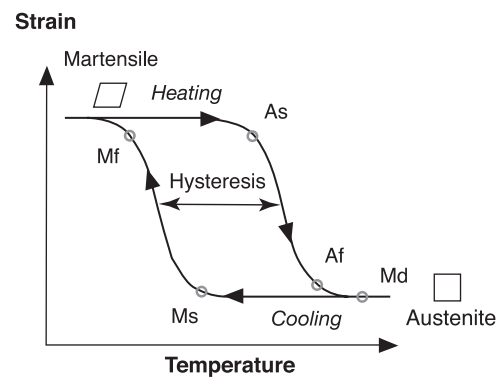


Fig. 5 Stress-strain diagram of an alloy with shape-memory behavior. (From Ref. [26].)

the unloading (deactivation) curve demonstrates how these wires are able to exert low, nearly constant forces. The ideal biological force levels exerted by the superelastic nickel-titanium alloys will move teeth with little discomfort to the patient.

One other striking characteristic of the superelastic nickel-titanium wires is that the unloading curve is affected by the degree of activation. Unloading stiffness is dependent on the level of activation; it decreases as the deflection increases. This phenomenon was first reported by Burstone et al.^[8] In their study, they showed that following an 80° activation, if a tooth moved to the 40° position, the nickel-titanium wire would exert a 380 gm-mm bending moment. However, if the wire was untied and retied, a 700 gm-mm bending moment was produced. Because of this unique behavior, therefore, the force exerted by an activated wire can be increased by simply untying and retying it.

In the mid-1990s, Ormco made a new alloy, copper Ni-Ti, available to the orthodontic profession. This new orthodontic wire is marketed in three different martensite-austenite transformation temperatures of 27°C, 35°C, and 40°C. According to the manufacturer, because shape memory behavior is obtained precisely at those specific temperatures, this wire allows the clinician to select wires on a case-specific basis.^[27] The excellent springback (working range) of nickel-titanium wires make them very useful when large deflections are necessary to engage a wire into the brackets of malpositioned teeth. The wires will generate low forces despite large deflections, resulting in less patient discomfort. These wires display springback approximately 1.4 times that of nitinol wire and 4.6 times that of stainless steel (Fig. 6).

Nickel-titanium wires have a higher coefficient of friction than stainless steel wires as a result of their higher surface roughness. This can be a major disadvantage during space closure, when sliding mechanics along a wire may be slowed down due to the high friction between the wire and the bracket.^[5,17] Nickel

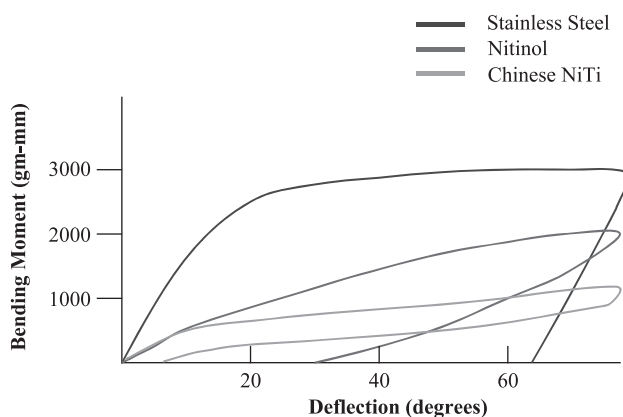


Fig. 6 Activation and deactivation curves of orthodontic wires. (Adapted from Ref. [24].)

content in the composition of nickel-titanium wires is another concern regarding the biocompatibility of this alloy. Clinicians should be careful about using nickel-titanium wires when treating patients with known nickel allergies.

Although these wires offer many desirable properties, they are relatively expensive when compared to stainless steel, exhibit poor formability, and cannot be soldered or welded.^[8]

IS THERE AN IDEAL WIRE?

From this discussion, it is evident that there is not one ideal wire that would suit the needs of all phases of orthodontic treatment. This is somewhat expected, because different stages of treatment will require wires of different stiffness and range.

Nickel-titanium wires provide light forces, wide elastic range, and excellent springback, but they are expensive, offer little formability, and cannot be soldered or welded. Also, their friction is higher than that of beta-titanium and stainless steel wires.^[5,6,8] They are most often suitable for early stages of treatment, where initial alignment of malpositioned teeth is the usual goal.

Stainless steel wires are the least expensive wires and display excellent formability, weldability and soldering properties.^[5,6,8] However, they have the highest force delivery and lowest elastic range and springback when compared to beta-titanium and nickel-titanium alloys.^[5,6,8] Their high stiffness makes them suitable for the midtreatment stages to stabilize aligned teeth, while moving other teeth as required. Their low coefficient of friction allows brackets to slide along the wire—during space closure, for example—more easily than other wire compositions.

Beta-titanium alloys have force delivery, elastic range, and springback values higher than those of stainless steel, but lower than nickel-titanium alloys.^[5,6,8] They can be formed and welded.^[5,6,8] They are expensive wires and have coefficients of friction higher than stainless steel and lower than nickel-titanium alloys.^[5,8] Because of their higher springback characteristics compared to stainless steel and better formability than nickel-titanium, beta-titanium wires are well suited for fabrication of active components of orthodontic appliances. For example, springs made of beta-titanium can be formed to move individual teeth or segments of teeth, as desired. Beta-titanium wires are also often used during the finishing stages of treatment, where small bends in the wire are required to allow for precise tooth positioning.

In summary, there is not one ideal wire best for all stages of orthodontic treatment. Arch wire selection should be based on the context of its intended use.



ARTICLES OF FURTHER INTEREST

Dental Biomechanics; Dental Implants; Nickel and Nickel Alloys; Shape Memory Metals

REFERENCES

1. Reitan, K.; Rygh, P. Biomechanical Principles and Reactions. In *Orthodontics: Current Principles and Techniques*, 3rd Ed.; Graber, T.M., Vanarsdall, R.L., Jr., Eds.; Mosby: St. Louis, 2000; 137–150.
2. Craig, R.G. Mechanical Properties. In *Restorative Dental Materials*, 8th Ed.; Mosby: St. Louis, 1989; 75–91.
3. Proffit, W.R.; Fields, H.W., Jr. Mechanical Principles in Orthodontic Force Control. In *Contemporary Orthodontics*, 3rd Ed.; Mosby: St. Louis, 2000; 326–328.
4. Nikolai, R.J. Behavior of Orthodontic Wire in Bending. In *Bioengineering Analysis of Orthodontic Mechanics*; Lea & Febiger: Philadelphia, 1985; 250–257.
5. Brantley, W.A. Orthodontic Wires. In *Orthodontic Materials Scientific and Clinical Aspects*; Brantley, W.A., Eliades, T., Eds.; Georg Thieme Verlag: Stuttgart, Germany, 2001; 78–104.
6. Burstone, C.J.; Goldberg, A.J. Beta titanium: A new orthodontic alloy. *Am. J. Orthod.* **1980**, *77*, 121–132.
7. Kapila, S.; Sachdeva, R. Mechanical properties and clinical applications of orthodontic wires. *Am. J. Orthod. Dentofac. Orthop.* **1989**, *96*, 100–109.
8. Kusy, R.P. A review of contemporary archwires: Their properties and characteristics. *Angle Orthod.* **1997**, *67*, 197–208.
9. ASM Committee on Wrought Stainless Steels. Wrought Stainless Steels. In *Metals Handbook*, 8th Ed.; American Society for Metals: Metals Park, OH, 1961; Vol. 1.

10. Flinn, R.A.; Trojan, P.K. *Engineering Materials and Their Applications*, 3rd Ed.; Houghton Mifflin: Boston, 1986; 530–531.
11. Budinski, K.G. *Engineering Materials: Properties and Selection*, 5th Ed.; Prentice Hall: Englewood Cliffs, NJ, 1996; 418–420.
12. Backofen, W.A.; Gales, G.F. Heat treating stainless steel wire for orthodontics. *Am. J. Orthod.* **1952**, *38*, 755–765.
13. Goldberg, A.J.; Burstone, C.J. An evaluation of beta titanium alloys for use in orthodontic appliances. *J. Dent. Res.* **1979**, *58*, 593–600.
14. Fillmore, G.M.; Tomlinson, J.L. Heat treatment of Cobalt-Chromium wire. *Angle Orthod.* **1976**, *46*, 187–195.
15. Fillmore, G.M.; Tomlinson, J.L. Heat treatment of Cobalt-Chromium alloys of various tempers. *Angle Orthod.* **1979**, *49*, 126–130.
16. Nelson, K.R.; Burstone, C.J.; Goldberg, A.J. Optimal welding of β titanium orthodontic wires. *Am. J. Orthod. Dentofac. Orthop.* **1987**, *92*, 213–219.
17. Kusy, R.P.; Whitley, J.Q. Effects of sliding velocity on the coefficients of friction in a model orthodontic system. *Dent. Mater.* **1989**, *5*, 235–240.
18. Bradley, T.G.; Mitchell, J.C. Brantley, Surface Composition and Microtopography of nickel-titanium Orthodontic Wires. In *Special Issue—J Dent Res*, 74th General Session and Exhibition of The International Association for Dental Research, San Francisco, CA, March 13–17, Herzberg, M.C., Ed.; American Association for Dental Research: Alexandria, VA, 1996; 61.
19. Kula, K.; Phillips, C.; Gibilaro, A.; Proffit, W.R. Effect of ion implantation of TMA archwires on the rate of orthodontic sliding space closure. *Am. J. Orthod. Dentofac. Orthop.* **1998**, *114*, 577–580.
20. Buehler, W.J.; Wiley, R.C. TiNi-ductile intermetallic compound. *Trans. Aktueller Softw.-Markt* **1962**, *55*, 269–276.
21. Andreasen, G.F.; Hilleman, T.B. An evaluation of 55 cobalt substituted nitinol wire for use in orthodontics. *J. Am. Dent. Assoc.* **1971**, *82*, 1373–1375.
22. Goldberg, A.J.; Morton, J.; Burstone, C.J. The flexure modulus of elasticity of orthodontic wires. *J. Dent. Res.* **1983**, *62*, 856–858.
23. Andreasen, G.F.; Morrow, R.E. Laboratory and clinical analyses of nitinol wire. *Am. J. Orthod.* **1978**, *73*, 142–151.
24. Burstone, C.J.; Qin, B.; Morton, J.Y. Chinese NiTi wire: A new orthodontic alloy. *Am. J. Orthod.* **1985**, *87*, 445–452.
25. Miura, F.; Mogi, M.; Ohura, Y.; Hamanaka, H. The super-elastic property of Japanese NiTi alloy wire for use in orthodontics. *Am. J. Orthod. Dentofac. Orthop.* **1986**, *90*, 1–10.
26. Meling, T.R.; Ødegaard, J. The effect of short-term temperature changes on superelastic nickel-titanium archwires activated in orthodontic bending. *Am. J. Orthod. Dentofac. Orthop.* **2001**, *119*, 263–273.
27. *Copper Ni-Ti Wire to Wire Consistency*; Ormco: Orange, CA, 1998. Promotional literature.

Orthopedic Biomaterials

J. Lawrence Katz

Department of Oral Biology, University of Missouri–Kansas City School of Dentistry, Kansas City, Missouri, and Department of Orthopaedics, University of Texas–Houston School of Medicine, Houston, Texas, U.S.A.

Catherine G. Ambrose

Department of Orthopaedics, University of Texas–Houston School of Medicine, Houston, Texas, U.S.A.

Carl McMillin

Synthetic Body Parts, Brecksville, Ohio, U.S.A.

Paulette Spencer

Department of Oral Biology, University of Missouri–Kansas City School of Dentistry, Kansas City, Missouri, U.S.A.



INTRODUCTION

A large variety of materials is presently used in orthopedic procedures. These include natural tissues such as bone allografts, autogenous bone, demineralized bone and matrix, and bone morphogenic proteins; and synthetic materials such as metals, ceramics, polymers, and composites. They fill a variety of uses such as joint replacement, reconstruction or replacement of skeletal defects, augmentation of fracture repair, strengthening of arthrodeses, filling of defects, and for spinal fusion. This article describes the properties and uses of these orthopedic biomaterials.

In the biomedical industry it can be difficult to compare metals because they are usually referred to by their trademarked names (such as Zimaloy, Vitallium, Stellite, Tivanium, TiOstallo, Protasul). Often the same alloy has multiple trademark names, making it necessary to investigate the actual composition. We have presented the metals using the appropriate American Society for Testing and Materials (ASTM) standard^[2] to simplify the comparison of the commonly used metals.

METALS USED IN ORTHOPEDICS

Metals were among the first orthopedic biomaterials, and are commonly used to this day. Iron wire was reported to have been used as early as 1775 for fracture fixation.^[1] Currently, most orthopedic implants are made from either stainless steel, titanium or one of its alloys, or a cobalt-chrome alloy, although tantalum and Nitinol metals have also been used.

When selecting a biomaterial for orthopedic applications, biocompatibility and corrosion resistance are certainly important, but structural properties are more critical than in other applications. Total hip replacements, total knee replacements, and bone plates must be able to withstand loads several times the body weight of the individual. Consideration of fatigue resistance, stiffness, and surface characteristics are all critical components of the material selection process. For this reason, metals are by far the most commonly used biomaterial in orthopedics.

STAINLESS STEELS

Nickel is added to iron-chromium-carbon alloys to increase the size of the austenite field. If the carbon content is low the steel is virtually all austenite at room temperature. Austenitic stainless steels have excellent ductility, formability, and corrosion resistance (compared to other steels) and are not ferromagnetic. Reducing the carbon content increases corrosion resistance.

In 1948, Fink and Smatko described American Iron and Steel Institute (AISI) Type 302 stainless steel (UNS S30200) as the ideal biomaterial for orthopedic applications.^[3] Today, most orthopedic implants made from steel are manufactured of AISI 316 or its low-carbon form, AISI 316L (ASTM F138 and F139). These stainless steel alloys contain a smaller percentage of carbon than type 302 as well as a small percentage of molybdenum, which improves its corrosion resistance in biological fluids. The compositions of these stainless steels are given in Table 1 and the mechanical properties of the commonly used stainless steels are given in Table 2.

In a study of retrieved 316L stainless steel implants, Weinstein and coauthors found that all implants

Table 1 Compositions of metals used for orthopedic implants

	ASTM F138,139	ASTM F75	ASTM F90	ASTM F67	ASTM F136	ASTM F2063
Aluminum	–	0.10 max	–	–	5.5–6.50	–
Boron	–	0.010 max	–	–	–	–
Carbon	0.030 max	0.35 max	0.05–0.15	0.08 max	0.08 max	0.070 max
Chromium	17.00–19.00	27.00–30.00	19.00–21.00	–	–	0.010 max
Cobalt	–	Balance	Balance	–	–	0.050 max
Copper	0.50 max	–	–	–	–	0.010 max
Hydrogen	–	–	–	0.015 max	0.012 max	0.005 max
Iron	Balance	0.75 max	3.00 max	0.20 max	0.25 max	0.050 max
Manganese	2.00 max	1.00 max	1.00–2.00	–	–	–
Molybdenum	2.25–3.00	5.00–7.00	–	–	–	–
Nickel	13.00–15.00	0.50 max	9.00–11.00	–	–	54.5–57.0
Nitrogen	0.10 max	0.25 max	–	0.03 max	0.05 max	–
Niobium	–	–	–	–	–	0.025 max
Oxygen	–	–	–	0.18 max	0.13 max	0.050 max
Phosphorous	0.025 max	0.020 max	0.040 max	–	–	–
Silicon	0.75 max	1.00 max	0.40 max	–	–	–
Sulfur	0.010 max	0.010 max	0.030 max	–	–	–
Titanium	–	0.10 max	–	Balance	Balance	Balance
Tungsten	–	0.20 max	14.00–16.00	–	–	–
Vanadium	–	–	–	–	3.5–4.5	–

displayed some type of corrosion (either crevice, pitting, or fretting).^[4] The severity of the corrosion had no apparent correlation to the amount of time in situ. There have been many other studies demonstrating that these implants are prone to corrosion in biological environments.

COBALT-CHROME ALLOYS

Cobalt alloys have better corrosion resistance than steels and better abrasive wear resistance. Alloying

elements produce high strength, resistance to creep, and corrosion resistance. In 1924, Zierold reported that a cobalt-chrome alloy (Stellite) was the best-tolerated metal in a study of biocompatibility using dogs.^[5] A derivative of Stellite is Vitallium (ASTM F75), which continues in use today for orthopedic and dental implants. There are basically two type of cobalt-chrome alloys used in medical implants: the castable alloy (ASTM F75) and the alloy that is usually wrought by forging the hot metal (ASTM F90). The compositions of these alloys are listed in Table 1 and the mechanical properties are listed in Table 2.

Table 2 Mechanical properties of metals used for implants

	Ultimate Tensile Strength (MPa)	Yield Strength (0.2% offset) (MPa)	Elongation (%)	Density (g/cc)
AISI 316L Stainless Steel (ASTM F138)	490–1350	190–690	12–40	7.95
Cast Cobalt-Chromium (ASTM F75)	655–1310	450–930	8	8.29
Wrought Cobalt-Chromium (ASTM F90)	860–900	310–830	30	
Unalloyed Titanium-Grade 1 (ASTM F67)	240–300	170–200	24	4.51
Ti-6Al-4V ELI (ASTM F136)	825–860	760–795	8–10	4.43
Nitinol (ASTM F2063)	551–1450	200–350	10	6.45
Unalloyed Tantalum (ASTM F560)	172–517	138–345	1–30	16.6

TITANIUM AND ITS ALLOYS

Titanium is a relatively lightweight metal that provides excellent corrosion resistance, high strength-to-weight ratio, and ductility. Its ultimate strength is lower than stainless steel or CoCr alloys and it has poor shear strength. Titanium is the ninth most abundant element, making it readily available. Cobalt and chrome are critical alloying elements in the biomedical industry but are in lower supply; the availability and price of these elements are highly variable. In the early 1950s, Leventhal reported on the use of titanium for bone screws and plates.^[6] Although commercially pure titanium (ASTM F67) is still used today, titanium alloys—especially those made from titanium, aluminum, and vanadium (ASTM F1108 and F136)—are more commonly used. The compositions of these alloys are given in Table 1 and the material properties in Table 2.

TITANIUM-NICKEL ALLOYS (SHAPE-MEMORY ALLOYS)

This special class of titanium alloy is generally about 50% titanium and 50% nickel. Titanium-nickel alloys (Nitinol, ASTM F2063) demonstrate a shape-memory effect. That is, after deformation the metal can return to its original shape upon heating. This property makes the metal interesting for many orthopedic applications. If the material is worked into a new shape while cold, the implant will regain its original shape upon heating to body temperature. This property has allowed the material to be used for suture anchors (which expand upon heating) and fracture compression staples and wires (which contract upon heating), as well as other medical applications, including cardiovascular stents. The composition and material properties of nitinol are given in Tables 1 and 2.

TANTALUM

Tantalum (ASTM F560) is a biocompatible metal but its mechanical properties make it unsuitable for most orthopedic applications where strength is required. However, it has good biocompatibility and is corrosion-resistant. Recently, this metal has been manufactured as a porous material that allows bony ingrowth. The material is formed by depositing tantalum onto a spongelike carbon structure using vapor deposition techniques. The resulting structure looks like trabecular bone and has mechanical properties similar to trabecular bone, while allowing either soft- or hard-tissue ingrowth.^[7]

POROUS COATING

Metallic implants for joint replacement applications can either be cemented into the body or placed by press-fit or interference-fit techniques. In the case of uncemented implants, the surface of the metal is often treated to allow for bony ingrowth or gripping of the implant. Often, coatings of a ceramic (ASTM F1609) or metal beads or fibers are deposited onto the implant. The metal coatings can be made from either titanium, titanium alloys (ASTM F1580), or cobalt-chrome alloys (ASTM F1377). The coatings can be placed on implants made from either titanium alloys or cobalt-chrome alloys. The coatings can be applied using sintering or plasma spraying techniques.

POLYMERS

Biodegradable Polymers

Most of the biodegradable polymeric products on the market are made from only a few polymers, many of which were first used in sutures. The most common suture materials are the polylactic and glycolic acid polymers and copolymers, the trimethylene carbonate copolymers, and polydioxanone. The advantages of the biodegradable polymeric products include the following: They disappear, so long-term stress shielding is not a concern; there are no long-term device or materials problems; and no second operation is required for removal. The biodegradable polymeric products can be used for drug delivery. There are, however, several disadvantages associated with these materials. Most resorbable polymers degrade too fast. As they degrade, the degradation products accumulate and often result in an inflammatory reaction and sometimes seromas. This appears to be due mostly to the acidic degradation products of the most common biodegradable polymers. The polymers are not high-strength, making load-bearing implants rare. A good review of the mechanical and degradation properties of biodegradable polymers has been published by Daniels et al.^[8]

The rate of polymer degradation depends on a variety of factors. These factors may be broadly subdivided into the following categories: implant features, material characteristics, mechanism of degradation, and recipient characteristics such as age and metabolic condition. Implant features include the size and shape of the implant, material composition of the implant, stress on the implant, and implantation site (e.g., subcutaneous tissue or bone). Material characteristics include the phase of the material, (i.e., amorphous as opposed to crystalline), the racemic mixture of the polymer, and the presence of additives or impurities.

The mechanism of degradation refers to enzymatic cleavage versus hydrolysis.^[9] In order to be biocompatible, the resorbable polymers must be not only free from potentially toxic or carcinogenic residual monomers, stabilizers, polymerization initiators, emulsifiers, etc., but the leachates, degradation products, and subsequent metabolites must also be biocompatible.

Synthetic Biodegradable Polymers

Poly(lactic acid) (PLA)

Degradation rates for PLA vary widely. Lactic acid exists as two optimal isomers, L and D, with the L isomer occurring in nature. Poly(L-lactic acid) (PLLA) is more crystalline and degrades much more slowly than poly(D-lactic acid) or racemic D,L mixtures. PLLA can retain 70% of its strength at 10 months and have significant polymer remaining after 4 years in vivo, while in a prospective study, poly(D,L-lactic acid) cylinders were reported to have resorbed within 1 month.^[10]

Fibers of PLA can be melt spun and various amounts of draw can be incorporated into the process to produce stronger, stiffer, and more crystalline fibers. The tensile strength increases when PLLA is converted from bulk (30–50 MPa) to melt spun fibers (10–70 MPa), high-speed spun and drawn fibers (400–1000 MPa), or chloroform/toluene gel drawn linear, chain extended fibers (1100–2100 MPa). The melt-spinning and hot-drawing process reduced the molecular weight of the PLLA from 480,000 to 70,000.

Self-reinforced (SR) composites are made when a fiber of a material is used to reinforce a matrix of the same material. Nonreinforced bulk PLLA has a flexural strength of 57–72 MPa and a flexural modulus of 2.7 GPa, while that of a SR-PLLA composite rod utilizing drawn fibers has a flexural strength of 180–300 MPa and a flexural modulus of 7–10 GPa.^[11]

Biomaterials composed of polylactic acid, polyglycolic acid, and their copolymers degrade in vivo by hydrolysis into lactic acid and glycolic acid, which are incorporated into the tricarboxylic acid cycle and excreted.^[14] In a comprehensive review of PLA-PGA biomaterials, Athanasiou and others^[14] report that these materials are generally biocompatible, although several studies have reported inflammatory reactions occurring 7–20 weeks after surgical placement of polylactide or polyglycolide implants.

Polyglycolic acid

Polyglycolic acid (PGA) was synthesized by Bischoff and Walden in 1893 and developed by American Cyanamid as the first absorbable synthetic suture in

1962. It has been available commercially since 1970.^[12] This material retains 20% of its initial strength after 3 weeks, but has no strength after 4 weeks. PGA screws (4.5 mm diameter with 3.2 mm core) were found to have an initial shear strength of 180 MPa, bending strength of 320 MPa, torsional strength of 0.6 N-m, and modulus of elasticity of 12 GPa.^[11] Self-reinforced PGA made by drawing has a bending modulus of 18 GPa and a bending strength of 415 MPa, whereas by sintering, the bending modulus is 12 GPa and the bending strength is 360 MPa.^[12] Polylactide, polyglycolide and poly(glycolide-co-lactide) are bioresorbable polymers that belong to the group of poly(α -hydroxy acids).^[15]

Seventeen of 67 patients treated with polyglycolide rods (3.2 or 4.5 mm diameter) between 1985 and 1987 were found to have a clinical course complicated by sudden painful swelling at the ankle, discharging spontaneously through the skin, and releasing sterile liquid with remnants of the degrading implant.^[16] The discharge lasted an average of 4 weeks. Of the 67 patients, 34 had radiographs that showed ovoid osteolytic areas in the implant channels.

Copolymers of lactic and glycolic acids

In vitro, cylinders made of poly(D,L-lactic-co-glycolic acid) 50:50 (PLA₂₅GA₅₀) started to degrade at 7 days, compared with 28 days for cylinders made of poly(D,L-lactic acid) (PLA₅₀) (see U.S. Patents 4,118,470, 3,636,956, and 4,137,921).^[10] Degradation times of 6+ years have been reported for PLLA.^[17] Because of the slow absorption of this material, clinical manifestations of inflammation, if they occur, are typically noted at 15 months to 3 years after implantation.^[17]

The degree of crystallinity of the polymer and the molecular characteristics of the R-groups prescribe the hydrophobicity. The rate of hydrolysis is affected because of the impact of the hydrophobic/hydrophilic character on diffusion of water into the material. Amorphous and hydrophilic regions of the polymer are more accessible to water. For example, in comparing the amorphous D-isomer of PLA (PDLA) and the more crystalline L-isomer (PLLA), PLLA is more biostable. However, copolymers of PLA-PGA or PDLA-PLLA are less hydrophobic and less crystalline, and tend to experience faster rates of hydrolysis.^[9]

Selected examples

Poly(ϵ -caprolactone): a semicrystalline polymer with melting point of 59–64°C,^[17] it degrades more slowly than PLA.^[18] Drug delivery devices can remain active for over a year.

Copolymer of lactic acid and caprolactone: an amorphous copolymer that has been evaluated for use in biodegradable nerve guides by the University of Groningen, The Netherlands.^[19] Two-year rat carcinogenicity tests have been conducted on a Japanese custom-synthesized copolymer of ϵ -caprolactone and L-lactic acid.^[20]

Copolymer of ϵ -caprolactone and glycolide, poly(glycolic co- ϵ -caprolactone) are reported to have a fast resorption rate, almost completely absorbed at 4–8 weeks when used as a matrix for artificial skin.^[20] The sutures made from glycolide and ϵ -caprolactone take 91–119 days to become completely absorbed.^[20]

Polydioxanone and substituted poly(1,4-dioxane-2,5diones)

Polydioxanone (PDS) $T_m = 107\text{--}112^\circ\text{C}$; $T_g = -10^\circ\text{C}$; poly-p-dioxanone: a bioadsorbable polymer that can be made into flexible monofilament surgical sutures (U.S. Patent 4,052,988, U.S. Patent 3,960,152).^[15] Degradation products and rate: hydroxyethoxy acetic acid excreted, resorbed in about 6 months with minimal tissue reaction.^[21]

Poly(ortho esters)

Poly(ortho esters) are not as strong or stiff as crystalline PLA or PGA but slower to degrade. Characteristics: Diketene acetal poly(ortho esters) are amorphous polymers made from diketene acetals.^[22–24] Intermittent cyclic loading (10.5 mm/day @ 8 N) in saline-reduced flexural strength and modulus by 75% after 40 days versus 29% and 20% for static loading at 3.6 N.^[23,24] A polyorthoester mixed with demineralized bone can be molded and contoured.^[25]

Polyanhydrides and poly(anhydride-co-imides)

Poly[bis(p-carboxyphenoxy propane)anhydride and sebacic acid anhydride]. Characteristics: Hydrolysis

results in two dicarboxylic acids.^[21] The incorporation of the imide component into the backbone of the polyanhydrides increases the mechanical strength. Imides have good mechanical properties because of their rigid units, but they are not degradable. In contrast to poly α -hydroxy esters, these polymers degrade by surface erosion.^[26] Poly(anhydride-co-imides) have been evaluated by personnel from the Massachusetts Institute of Technology, Medical College of Pennsylvania, and Drexel University for orthopedic applications. They were found to have low cytotoxicity and excellent mechanical properties.^[27]

Examples of natural biodegradable polymers are provided in Table 3.

Hydrogel Polymers

Hydrogel polymers are cross linked hydrophilic polymers that are significantly swollen in water. Typical hydrogels are hydroxy alkyl methacrylates, poly(vinyl alcohol), and copolymers of N-isopropyl acrylamide. Several hydrogel monomers have been copolymerized with stable monomers. Depending on the ratio of the monomers, this has been done in order to control the water content and mechanical properties of the resulting hydrogel polymer (high hydrogel content) or to increase the swelling and biological interaction of the resulting stable polymer (low hydrogel content).

Clear UHMWPE

Clear acetabular cups have been made from ultrahigh molecular-weight polyethylene (UHMWPE) using a high-pressure (e.g., 15,000 psi) solid-state forming operation by Polyteco Inc., Hayward, CA. This is said to produce an oriented and highly consolidated material with improved wear and creep properties.^[31] UHMWPE (molecular weight $>2 \times 10^6$ g/mol) has been used for load-bearing applications such as acetabular cup of total hip, tibial plateau, and patellar surfaces of knee joints.^[15]

Table 3 Examples of natural biodegradable polymers

Material	Characteristics	References
Collagen cross-linked with hexamethylene diisocyanate fibers	Braided into ACL ligaments	[28]
Chitin, chitosan, and chitin derivatives	Chitosan: biodegradable and partially deacetylated product of natural polysaccharide chitin, based on free and acetylated glucosamine units	U.S. Patent 4,074,366. ^[29]
Alginates	Family of copolymers based on β -D-manuronic and α -L-gluronic acid residues extracted from seaweed	
Cellulose	Partially oxidized	[30]
Fibrin	Physiologically occurring and modified proteins	

Stable Polymers for Implants

In addition to poly(aryl ether ketones) (PAEKs), several other polymers may be suitable for use in implants. One set of these polymers is the liquid crystal polymers such as Vectra. Their biostability looks second-best only to the PAEKs, based on temperature stability, resistance to solvents, and ability to be autoclaved. One study presents rabbit data suggesting that Vectra/carbon fiber composite intramedullary rods are biocompatible.^[32] One other potential criterion is the ability to bond with carbon fiber in a biostable mode, in which the PAEKs are also exceptionally good. One potential drawback of the liquid crystal polymers is their complex molecular structure, which makes it difficult to determine their potential degradation products, and to get approval from the Food and Drug Administration (FDA) for their use in implants.

Aromatic poly(ether ketones) are generally prepared by the nucleophilic displacement of activated aromatic dihalides with alkali-metal bisphenolates in polar aprotic solvents, and to a lesser extent by Friedel-Crafts acylation reactions.^[33] As early as 1967, condensation polymerizations of alkali bisphenates with activated (negatively substituted) aromatic dihalides were reported by Johnson et al. to form ether bonds via nucleophilic substitution of the aromatic dihalides—often fluorides—in tetramethyl sulfone or dimethyl sulfoxide at about 160°C.^[34,35] $\text{KO-Ar-OK} + \text{F-Ar-C=O-AI-F} \rightarrow \text{PEEK}$, but at low molecular weight. Hexafluoroisopropylidene groups have also been incorporated into the polymer backbone to give better polymer solubility.^[33] There are at least 26 synthesis-related patents for poly(aryl ether ketones) identified between August, 1978 and July, 1996. The initial Imperial Chemical Industries (ICI) patents on PEEK are as follows: U.S. Patent No. 4,105,635, U.S. Patent No. 4,113,698, and U.S. Patent No. 4,113,699.

U.S. Patent 5,294,391 (Method of Making a Fiber-Reinforced Composite Structure Including Randomizing of Reinforcing Fibers) was granted to Carl R. McMillin on March 15, 1994. Many PAEK/carbon-fiber spine cages made using this patent and U.S. Patent 5,429,863 (Fiber Reinforced Composite Structure, granted to Carl R. McMillin on July 4, 1995), have been implanted in patients.

CERAMICS AND GLASSES

Both ceramics and glass ceramics are available in various forms for use as orthopedic biomaterials. Ceramic biomaterials are essentially inert within the body. This is a result of improved resistance to corrosion due to ceramic biomaterials' resistance to further oxidation. Although exhibiting more brittle behavior

than metals, ceramics of small grain size and techniques providing full density have prompted increased use of ceramics in orthopedics.

Biostable Ceramics

Biostable ceramics neither resorb nor induce osteoblastic apposition on their surfaces within the body. Aluminum oxide (alumina, ASTM F-603) and a zirconium oxide (zirconia) compound (ASTM F-1873) are the two most common biostable ceramics. Commercial submicron zirconia and alumina powders are used to make highly concentrated aqueous colloidal suspensions that can be cast. Without fiber reinforcements, strengths in three-point bending of 900 to 1200 MPa (115,000 to 175,000 psi) are reported. Advantages are that both aluminum oxide and zirconium oxide are strong and stable (so there is no need to follow degradation products); disadvantages are having weak interface with bone or tissue, low shock resistance, high modulus, and the potential for catastrophic failures. Alumina has been used clinically for many years in Europe and Canada as matched pairs of femoral head and acetabular components in total hip arthroplasty.

Mineralized bone growth into porous calcium aluminate skeletal implants requires a minimum interconnective pore size of 100 μ . The ingrowth of osteoid tissue is said to require a minimum pore size between 40 and 100 μ and the ingrowth of fibrous tissue requires a minimum pore size between 5 and 11 μ .^[36]

Bioactive Ceramics

Certain ceramics and glass ceramics also provide osteophilic surfaces so that bone-forming osteoblasts lay down bone in direct apposition on these surfaces. These are known as bioactive materials. These ceramics (ASTM F-1538) bond to bone and soft tissue. The most common examples are bioactive glasses (Bioglass), bioactive glass-ceramics (Ceravitals A-W Ceramic). Their main advantage is good bonding to tissue and bone; their disadvantage is that they are not as strong as biostable ceramics. A-W Ceramics are being used as space fillers in the iliac crest, as vertebral body replacements in the shape of rectangular bars, and as posterior lumbar interbody fusion (PLIF) cages (used in 13 cases) similar to the Brantigan posterior lumbar interbody fusion (PLIF) cage with an open area for bone growth from top to bottom.

Flexible Ceramics

There are several approaches now available to make ceramics more flexible, including nanophase

technology.^[37] If the brittle nature of ceramics is the only drawback in the design of a new product, these technologies could be used to make tougher, more flexible implants.

Bioresorbable Ceramics

A wide variety of bioresorbable ceramics have been proposed and developed with differing rates of degradation and dissolution. The order of dissolution rates for various apatites and other calcium-containing bio-ceramics is: fluorapatite (FAp) < magnesium whitlockite (MW) < hydroxyapatite (HAp) < calcium-deficient hydroxyapatite (CDHA) < OHA < β -tricalcium phosphate (β -TCP) = α -TCP { < calcium carbonate}.^[38-40] Mixed rates of degradation have been reported. In one case, β -TCP (ASTM F-1088) particles (Biobase) showed 60% degradation in minipigs while α -TCP particles (Cerasorb) showed 30% degradation.^[41] After 20 weeks, β -TCP particles showed 70% degradation while α -TCP (ASTM F-1) particles showed 40% degradation. The Cerasorb particles showed inflammatory cellular reaction over this period. After 68 weeks, 80–90% of the Cerasorb and 90–95% of the Biobase had been transformed into bone. At 86 weeks, 5% Cerasorb and 3% Biobase were still found.^[41]

Biphase calcium phosphate made from β -TCP and HAp (ASTM F-1185) is replaced by bone with the speed of dissolution of the ceramic being related to the ratio of the components.^[42] TCP is said to be osteoconductive but not osteoinductive.^[43] Biphase calcium phosphate (60% HAp:40% β -TCP) has been used instead of allograft or autograft to assist in scoliosis fusions.^[44] In the presence of osteoporosis, calcium phosphate-coated implants have no advantage over porous coatings in the mechanical fixation achieved.^[23,24] TCP dowels compressed and became displaced in 70% of anterior cervical discectomies in 20 dogs within 6 weeks. In another study, porous HAp had similar problems because of its brittle nature.^[45] Biphase ceramic also produced good posterolateral fusions in sheep.^[46]

Plaster of paris has been used in patients as a bone substitute since 1892, and HAp (in cats and monkeys) since 1952;^[45] medical-grade calcium sulfate granules have also been used to substitute for bone. Rapidly resorbable, glassy crystalline materials based on calcium alkali orthophosphates have been studied.^[47]

Apatites and Other Calcium Phosphates

Apatites and other calcium phosphate ceramics are the most intensely studied and clinically used ceramics in orthopedics. HAp, HAp granules, and HAp-sprayed

titanium are commonly used in dental and endosteal implant applications as well as on hip and knee prostheses (more commonly in Europe than in the United States). A HAp plasma flame-spray coating has been developed by many groups. Both HAp and HAp-coated titanium have been used as iliac bone crest spacers.^[48] HAp has been successfully coated onto polymers, including the Brantigan I/F PLIF Cage. COLHAP, a mixture of TCP and HAp (Ca/P ratio 1.6–1.67), is said to have mechanical properties superior to those of HAp. HAp has been mixed with cancellous bone for alveolar ridge augmentation by many users for years in Europe and Canada, and has been added to other bone cements.^[49]

When implanted in soft tissue, calcium phosphate ceramics generate a stable fibrous encapsulation. In bone, dense or microporous CaP materials become incorporated by directly bonding to the bone without an intervening fibrous layer within 3–5 months. In bone, macroporous implants are initially surrounded by fibrous connective tissue that extends into the peripheral pores. This tissue infiltration becomes denser and more organized and extends deeper into the implant until all pores and tunnels are filled, and is accompanied by osteoblastic cellular differentiation and sprouting capillaries. From one week to a month after implantation, osteoid and then new bone are deposited along the implant-fibrous tissue interface. By 4–8 months, complete infiltration is accomplished, with mature new bone bonded directly to the implant material with mature osteons, haversian canals, and mature bone marrow and vessels, without any intervening fibrous tissue. A pore size of 200 μ is needed for mature osteons to form.^[45]

Fluoroapatite (FAp) does not appear to show the same rate of degradation as HAp;^[50] however, its bone-bonding capability is slower and may not be as strong as HAp.^[51] A bioceramic has been developed from high-molecular weight calcium phosphate material with 150 to 250 micron pores converted to HAp that is said to have strength suitable for load-bearing applications.^[52]

Solid HAp and HAp coatings (or fluorapatite) may be good where nonprecision placement of component is anticipated. HAp dissolves slowly at the surface, which enhances bone growth and replacement.^[53] However, at the HAp-metal interface, this produces a weak interface over long periods of time that may cause problems with some implants.^[54] It has been found that thick coatings of HAp do not perform well.

Ceramic Bone Cements

Ceramic bone cements are another active area of research and clinical use. Several different approaches have been taken in the development of a variety of ceramic-based bone cements. Those that consist of

mixtures of ceramics and polymers or other organic compounds are discussed in the Composites section of this article. Examples of bone cements based on ceramics are discussed in the following paragraphs.

Zinc calcium phosphorous oxide ceramic cements can be made either with aluminum calcium phosphorous oxide (ALCAP) or tricalcium phosphate (TCP) for resorbable composites, or with hydroxyapatites for relatively nonresorbable composites; antibiotics such as penicillin or erythromycin or growth factors can be incorporated into the TCP amino acid composite.

A calcium phosphate ceramic that cures in situ in 4 minutes (attaining 90% strength in 4 hrs; cured in 24 hrs) to form a carbonated hydroxyapatite with a yield stress of 25 MPa has been evaluated as an alternative to PMMA for augmentation of transpedicular screws^[54] and in a canine metaphyseal defect.^[55]

A calcium phosphate/hydroxyapatite-based SuperBone family of cements using brushite, which hardens in vivo in about 10 minutes, is under development. SuperBone is also said to be a fast-setting carbonated HAp in preclinical testing.^[56] Under the name Skeletal Repair System (SRS), the carbonated apatite starts to set in 5 minutes, is 50% cured in one hour, and is said to have a compressive strength of 55 MPa after 12 hours.^[57] SRS obtained regulatory approval in the U.S.A. in 1998. It is said to be a dry mixture of monocalcium phosphate, monohydrate, tricalcium phosphate, and calcium carbonate, and to be made into a curable paste with sodium phosphate.

A HAp cement (BoneSource) can initially be formed (like paste) into a shape that becomes hard in 20–30 minutes and converts to HAp in 4 hours. It is sculptable and resorbable because it is a nonceramic HAp, unlike other HAp bone substitutes. It is used for maxillofacial surgery.

Ceramic Foam

Ceramics can also be made into foam to allow tissue ingrowth and to provide a large surface area. Porous hydroxyapatite and porous tricalcium phosphate ceramics have been made.^[58] A fiber-based ceramic foam (made from biostable ceramics) has recently been developed. With resorbable ceramics, a large surface area makes for rapid resorption. The foam can be filled with many different materials for differing applications.

COMPOSITES

A composite is a material whose material properties are different from those of the individual starting components. Bone is a classic example of a biological

composite, in terms of both materials (collagen and apatite) and structure (i.e., the microstructural organization of osteons).^[59] Synthetic composite systems have therefore found several areas of use in orthopedics. Composite analysis is a very active area in mechanics. However, a general understanding of the behavior of composites can be illustrated by several simple algebraic relationships.

Some of the composites presently found in clinical or research orthopedics are carbon fiber-reinforced UHMWPE, PLA (matrix)/PGA (fiber) fracture plates, and carbon fiber-reinforced carbon (CFRC) fracture plates. Carbon fibers have also been used as reinforcements with either epoxy or polysulfone matrices; metal fiber-reinforced polyacetal resin plates are additional composites found in orthopedic practice and study. Apatite-collagen composites are in use as a biomimetic approach to bone replacement. Fiber composites are also under study and in use for tendon replacements.

Ceramic-Collagen Bone Replacement

Ceramic materials are combined with collagen to form an artificial bone that resorbs providing scaffolding on which the patient's bone cells will grow. Two ceramic materials are mixed with bovine collagen and the patient's bone marrow and used as bone cement. A good review of the use of ceramics and other bone graft substitutes has been published by Damien and Parsons.^[60]

A commercial collagen-HAp composite (Collagraft) has been developed and studied. This collagen-HAp composite with the addition of bone marrow, has been used in rat studies.^[61] Collagraft is reported to consist of purified fibrillar collagen, hydroxyapatite, and tricalcium phosphate, and to resorb in a canine femoral shaft defect similar to autogenous bone, with healing not yet complete after one year.^[62]

It has been reported that biphasic granules (40% TCP/60% HAp) mixed with collagen were no better than no graft in a canine spinal fusion model.^[63] However, when combined with a mixture of bone-morphogenic proteins, it was about half as good as autogenous cancellous bone.

A collagen-tetracalcium phosphate bone cement has been studied in femoral condyles of rabbits in Japan.^[48] Organoapatites made from tricalcium phosphate and polyamino acids has been reported.^[64] A soft collagen-based bone defect filler (Osteobit) is being sold in Europe.

Hydroxyapatite-Polymer Composites

A hydroxyapatite-polyethylene composite has been developed for use in orthopedic implants. The material knits together with bone, maintains good mechanical

properties, and can be shaped or trimmed during surgery using a scalpel. The material was developed at Queen Mary and Westfield College and Brunel University in the United Kingdom.^[65]

A bone-joining material for orthopedic plates has been developed in Japan combining fine particles of HAp and PLA at a weight ratio of 30:70 to 60:40. The material is said to be stronger than natural bones and to maintain its strength over 4–6 months, while becoming completely absorbed into natural bone within a few years.^[66]

PMMA-Ceramic Systems

A HAp/acrylic monomer has been prepared, composed of 80% HAp granules coated with silane-coupling agent bonded with bis-phenol-A glycidyl methacrylate, 2,2'-bis(4-methacryloxyethoxy-phenyl) propane, and triethyleneglycol dimethacrylate (35:35:30 w/w), which has a setting time of 5 minutes, a peak curing temperature of 46°C, and a tensile strength of 39.6 MPa (versus 28.6 MPa for a commercial PMMA cement).

A PMMA-HAp cement has been used in total hip replacement surgeries in Japan since 1984.^[48]

Glass-Ionomer or Polyalkenote Cements

When mixtures of boric, zinc, and aluminum oxides; zinc fluoride; tartaric acid; and polyacrylic acid were evaluated 8 of the 20 mixtures had suitable setting characteristics, and 3 produced hard cements.^[67,68] Mechanical properties of polyalkenoate cement have been presented by Canadian researchers.^[69]

Bioresorbable Composites

Bioresorbable composites of various compositions have a 15-year history of clinical use as fixation devices^[70] (e.g., as screws and plate).^[71] In an extensive study of more than 1000 operations, it was found that these materials provided a greater degree of osteogenesis than could be achieved with metallic implants.^[72] There are several potential advantages to the use of bioresorbable fixation devices. They can be impregnated with BMPs or antibacterial drugs. Osteoporosis associated with stress shielding due to the significantly higher modulus of elasticity associated with metallic devices will be reduced or completely avoided. Bioresorbable devices also avoid the need for further surgery (i.e., for their removal). However, there are also potential concerns. Metallic implants are stiffer and more rigid, thus providing better mechanical support under weight-bearing conditions. As early as 1994, research was directed at the improvement of the

mechanical properties, tensile strength, and rigidity of bioresorbable materials.^[73,74]

Other Developments

Japanese researchers have presented a chitosan, hydroxyapatite or β -tricalcium phosphate, malic acid self-hardening paste that generates little heat but is not as strong as bone cement.^[75,76] CaO, MgO, and ZnO are used to modify its properties. An α -TCP with citric acid, malonic acid, and collagen is a cement that is stronger than PMMA cement that has been prepared in Japan.^[77]

ARTICLES OF FURTHER INTEREST

Biologic and Synthetic Apatites; Blood-Material Interactions; Bone Cement; Bone-Implant Interface; Bone Plates and Screws, Bone Remodeling; Diamond and Diamond-Like Carbons; Hydrogels; Implant, Total Hip; Knee Joint Replacement; Orthopedic Fixation Devices; Poly(Glycolic Acid); Titanium and Its Alloys; Zirconia Ceramics

REFERENCES

1. Icart. Journal de Medecine Chirurgie, Pharmacie, Etc. 1775, 44, 169. In *Annual book of CJJ ASTM Standards*; American Society for Testing and Materials: Philadelphia, 2002; Vol 13.
2. Fink, C.G.; Smatko, J.S. J. Electrochem. Soc. **1948**, 94, 271.
3. Weinstein, A.M.; Spires, W.P.; Klawitter, J.J.; Clemow, A.J.T.; Edmunds, J.O. Orthopedic Implant Retrieval and Analysis Study. In *Corrosion and Degradation of Implant Materials*; Syrett, B.C., Acharya, A., Eds.; American Society for Testing and Materials: Philadelphia, 1979; 212–228.
4. Zierold, A.A. Arch. Surg. **1924**, 9, 365.
5. Leventhal, G.S. J. Bone Jt. Surg. **1951**, 33-A (2), 473.
6. Boby, J.D.; Tackpool, G.J.; Hacking, S.A.; Tanzer, M.; Krygier, J.J. Characteristics of bone ingrowth and interface mechanics of a new porous tantalum biomaterial. J. Bone Jt. Surg. **1999**, 81 -B (5), 907–914.
7. Taylor, M.S.; Daniels, A.U.; Andriano, K.P.; Heller, J. Six bioabsorbable polymers: In vitro acute toxicity of accumulated degradation products. J. Appl. Biomater. **1994**, 5 (2), 151–157.
8. An, Y.H.; Woolf, S.K.; Friedman, R.J. Pre-clinical in vivo evaluation of orthopedic bioabsorbable devices. Biomaterials **2000**, 21, 2635–2652.
9. Gautier, S.E.; et al. Poly(α -hydroxyacids) for application in the spinal cord: Resorbability and biocompatibility with adult rat schwann cells and spinal cord. J. Biomed. Mater. Res. **1998**, 42 (4), 642–654.



10. Mayer, M.H.; Hollinger, J.O. Biodegradable Bone Fixation Devices. In *Biomedical Applications of Synthetic Biodegradable Polymers*; Hollinger, J., Ed.; CRC Press: Boca Raton, 1995; 173–195.
11. *Self-Reinforced Bioabsorbable Polymeric Composites in Surgery*; Rokkamen, P., Törmälä, P., Eds. Tampereen, Pikakapio: Tampere, Finland, 1995.
12. Incardona, S.D.; Fambri, L.; Migliaresi, C.; University of Trento, Italy. Poly-L-Lactic acid braided fibres produced by melt spinning: Characterization and in vitro degradation. *J. Mater. Sci., Mater. Med.* **1996**, *7* (7), 387–391.
13. Athanasiou, K.A.; Niederauer, G.G.; Agrawal, C.M. Sterilization, toxicity, biocompatibility and clinical applications of polylactic acid/polyglycolic acid copolymers. *Biomaterials* **1996**, *17*, 93–102.
14. Bronzino, J.D.; Joseph, D. *The Biomedical Engineering Handbook*, 2nd Ed.; CRC Press: Boca Raton, FL, 2000; Vol. 1, 39.
15. Bostman, O.M. Osteolytic changes accompanying degradation of absorbable fracture fixation implants. *J. Bone Jt. Surg., Br.* **1991**, *73-B* (4), 679–682. David L. Helfet—Advances in Orthopedic Surgery, 16 (5), 335–337, March/April 1993.
16. Pietrzak, W.S.; Eppley, B.L. Resorbable polymer fixation for craniomaxillofacial surgery: Development and engineering paradigms. *J. Craniofac. Surg.* **2000**, *11*, 575–585.
17. Ratner, B.D.; et al. *Biomaterials Science an Introduction to Materials in Medicine*; Academic Press; 1996.
18. den Dunnen, W.F.A.; et al. Biological performance of a degradable poly(lactic acid- ϵ -caprolactone) nerve guide: Influence of tube dimensions. *J. Biomed. Mater. Res.* **1995**, *29* (6), 757–776.
19. Nakamura, T.; et al. Biodegradation and tumorigenicity of implanted plates made from a copolymer of ϵ -caprolactone and L-lactide. *J. Biomed. Mater. Res.* **1998**, *42* (4), 475–484.
20. Barrows, T.H. Synthetic Bioabsorbable Polymers. In *High Performance Biomaterials*; Szycher, M., Ed.; Technomic Publishing Co.: Lancaster, 1991; 243–257.
21. Wenger, K.H.; Univ. Utah School of Med. Changes in Mechanical Properties of Three Bioabsorbable Fracture Fixation Pin Types Due to Canine Implantation, 40th Annual Meeting, Orthopedic Research Society, New Orleans, Feb. 21–24, 1994; 538.
22. Daniels, A.U.; Andriano, K.P.; Smutz, W.P.; Chang, M.K.O.; Heller, J. Evaluation of absorbable poly(ortho Esters) for use in surgical implants. *J. Appl. Biomater.* **1994**, *5* (1), 51–64.
23. Rawlings, C.E., III. Modern bone substitutes with emphasis on calcium phosphate ceramics and osteoinductors. *Neurosurgery* **1993**, *33* (5), 935–938.
24. Rawlings, C.E., III. Modern bone substitutes with emphasis on calcium phosphate ceramics and osteoinductors. *J. Neurosurg.* **1992**, *76*, 275–279.
25. Behravesch, E.; et al. Synthetic biodegradable polymers for orthopedic applications. *Clin. Orthop. Relat. Res.* **1999**, *367s*, s118–s125.
26. Attawia, M.A.; Uhrich, K.E.; Botchwey, E.; Fan, M.; Langer, R.; Laurencin, C.T. Cytotoxicity testing of poly(anhydride-co-imides) for orthopedic applications. *J. Biomed. Mater. Res.* **1995**, *29* (10), 1222–1240.
27. Gurav, N.; Downes, S. A qualitative in vitro evaluation of the degradable materials poly(caprolactone), poly(hydroxybutyrate) and a poly(hydroxybutyrate)-(hydroxyvalerate) copolymer. *J. Mater. Sci., Mater. Med.* **1994**, *5* (11), 784–787.
28. Lee, K.Y.; Ha, W.S.; Park, W.H. Blood compatibility and biodegradability of partially N-acetylated chitosan derivatives. *Biomaterials* **1995**, *16* (16), 1210–1216.
29. Radder, A.M.; Leenders, H.; van Blitterswijk, C.A. Interface reactions to PEO/PBT copolymers (Polyactive^{reg}) after implantation in cortical bone. *J. Biomed. Mater. Res.* **1994**, *28* (2), 141–151. (Also similar paper pp. 269–277).
30. Laurencin, C.T.; Norman, M.E.; Elgendy, H.M.; El-Amin, S.F.; Allcock, H.R.; Rucher, S.R.; Ambrosio, A.A. Use of polyphosphazenes for skeletal tissue regeneration. *J. Biomed. Mater. Res.* **1993**, *27* (7), 963–973.
31. Kettunen, J.; Makela, A.; Miettinen, H.; Nevalainen, T.; Heikkila, M.; Torma, P.; Rokkanen, P. The effect of an intramedullary carbon-fiber-reinforced liquid crystalline polymer impant on bone: An experimental study on rabbits. *J. Biomed. Mater. Res.* **1998**, *42* (3), 407–411. [Finland].
32. Tulles, G.L.; Cassidy, P.E.; St. Clair, A.K. Polymers derived from hexafluoroacetone: 12F-poly(ether ketone). *Macromolecules* **1991**, *24*, 6059–6064.
33. Nguyen, H.X.; Ishida, H. Poly(aryl-ether-ether-ketone) and its advanced composites: A review. *Polym. Compos.* **1987**, *8* (2), 57–73.
34. Johnson, R.N.; Farnham, A.G.; Clendinning, R.A.; Hale, W.F.; Merriam, C.N. *J. Polym. Sci., Polym. Chem. Ed.* **1967**, *15*, 2375–2398.
35. Klawitter, J.J.; Hulbert, S.F. Tissue reaction to three ceramics of porous and nonporous structures. *J. Biomed. Res.* **1972**, *6* (5), 347–374.
36. Haher, T. *New and Transfer Technology, Materials; Spinal Instrumentation from Harrington to the Future, Continuing Education Course by AAOS and SRS, Kansas City, Kansas, September 23, 1992.*
37. Dhert, W.J.A.; Klein, C.P.A.T.; Jansen, J.A.; van der Velde, E.A.; Vriesde, R.C.; Rozing, P.M.; de Groot, K. A histological and histomorphometrical investigation of fluorapatite, magnesium whitlockite, and hydroxylapatite plasma-sprayed coatings in goats. *J. Biomed. Mater. Res.* **1993**, *27* (1), 127–138.
38. Radin, S.R.; Ducheyne, P. The effect of varying the calcium phosphate composition and structure on dissolution and precipitation reactions in simulated interstitial fluid. *Trans. Soc. Biomater.* **May 1–5, 1991**, *XIV*, 193.
39. Denissen, H.W.; et al. Stable loaded HA coatings in comparison with experimental plasma sprayed alternatives. *Trans. Soc. Biomater.* **May 1–5, 1991**, *XIV*, 189.
40. Wiltfang, J.; et al. Degradation characteristics of α and β tri-calcium-phosphate (TCP) in minipigs. *J. Biomed. Mater. Res. (Appl. Biomater.)* **2002**, *63* (2), 115–121. (Germany).
41. Daculsi, G.; LeGerros, R.Z.; Nery, E.; Lynch, K.; Kerebel, B. Transformation of biphasic calcium

- phosphate ceramics in vivo: Ultrastructural and physiochemical characterization. *J. Biomed. Mater. Res.* **1989**, *23* (8), 883–894.
42. Nicholas, R.W.; Lange, T.A.; St. Paul-Ramsey Medical Center, St. Paul, MN. Granular tricalcium phosphate grafting of nonstructural cavitary bone lesions in humans. *Trans. Soc. Biomater.* **May 1–5, 1991**, *XIV*, 8.
 43. Passuti, N.; Daculsi, G.; Martin, S.; Deudon, C. Macroporous Polycrystalline Calcium Phosphate Implant for Spinal Fusion in Man and Dogs. In *Handbook of Bioactive Ceramics, Volume II—Calcium Phosphate and Hydroxylapatite Ceramics*; Yamamuro, T., Hench, L., Wilson, J., Eds.; CRC Press: Boca Raton, 1990; 345–354.
 44. Chuckler, J.M.; Ducheyne, P.; Nangle, A.E. *The Effect of Osteoporosis on Skeletal Fixation of Porous and Calcium Phosphate Coated Implants*, Transactions of Fourth World Biomaterials Congress, Berlin, Germany, April 24–28, 1992; 97.
 45. Guigui, P.; Plais, P.Y.; Flautre, B.; Viguier, E.; Blary, M.C.; Chopin, D.; Lavaste, F.; Hardouin, P. Experimental model of posterolateral spinal arthrodesis in sheep: Part 2. Application of the model: Evaluation of vertebral fusion obtained with coral (porites) or with a biphasic ceramic (triosite). *Spine* **1994**, *19* (24), 2798–2803. (France).
 46. Berger, G.; Gildehaar, R.; Ploska, U. Rapid resorbable glassy crystalline materials on the basis of calcium alkali orthophosphates. *Biomaterials* **1995**, *16* (16), 1241–1248.
 47. Oonishi, H. Bioceramics in Orthopedic Surgery—Our Clinical Experiences. In *Bioceramics*; Hulbert, S., Ed.; Rose-Hulman Institute of Technology, 1992; Vol. 3, 31–42.
 48. Ishihara, K.; Arai, H.; Nakabayashi, N. Adhesive bone cement containing hydroxyapatite particle as bone compatible filler. *J. Biomed. Mater. Res.* **1992**, *26*, 937–945.
 49. Dhert, W.J.A.; de Groot, K.; et al. Biomaterials Research Group, Departments of Biomaterials and Orthopaedics, School of Medicine, University of Leiden, The Netherlands The response of trabecular bone to fluorapatite, magnesium whitlockite, and hydroxylapatite plasma-spray coated implants, an experimental study in goats. *Trans. Soc. Biomater.* **May 1–5, 1991**, *XIV*, 3.
 50. Kangasniemi, I.; Verheyen, C.C.P.M.; Wolke, J.G.C.; de Groot, K. Tensile Test of the Bone Bonding Strength to Fluor- and Hydroxylapatite and Titanium, Transactions of the Fourth World Biomaterials Congress, Berlin, Germany, April 24–28, 1992; 99.
 51. Hotter, D.S. Band-aids for broken bones. *Mach. Des.* **April 4, 1996**, 39–44.
 52. Neo, M.; Nakamura, T.; Ohtsuki, C.; Kokubo, T.; Yamamuro, T. Apatite formation on three kinds of bioactive material at an early stage in vivo: A comparative study by transmission electron microscopy. *J. Biomed. Mater. Res.* **1993**, *27* (8), 999–1006.
 53. David, A.; Eitenmuller, J.; Hangst, K.; Bar, H.F.; Muhr, G. Bone-Defect Substitution in Load Bearing Skeletal Segments by Titanium and Hydroxyapatite-Coated Titanium-Implants, Transactions of the Fourth World Biomaterials Congress, Berlin, Germany, April 24–28, 1992; 93.
 54. Sanden, B.; Olerud, C.; Johansson, C.; Larsson, S. Improved extraction torque of hydroxyapatite-coated pedicle screws. *E. Spine J.* **2000**, *9*, pp. 6, 534–537.
 55. Graziano, G.P.; More, D.C.; Maitra, R.S.; Farjo, L.A.; Goldstein, S.A. Calcium Phosphate Ceramic vs. PMMA for the Augmentation of Transpedicular Screws, Scoliosis Research Society, 28th Annual Meeting, Dublin, Ireland, September 18–23, 1993, 92. Exhibit #43.
 56. Frankenburg, E.P.; Goldstein, S.A.; Harris, S.A.; Toomajian, L.R.; Pellizzon, G.G.; Balke, D.A.; Bauer, T.W.; Orthop. Res. Lab. U. Mich. and CCF, Cleveland Evaluation of an in situ-setting calcium phosphate grout in an in vivo canine metaphyseal defect model. *Orthop. Trans.* **1994**, *18* (2), 413.
 57. Caddes, D.E.; Norian Corporation. Novel bone repair biomaterials. *Orthop. Trans.* **1994**, *18* (2), 3C–22C.
 58. *Orthop. Today* **1995**, *15* (5), 10.
 59. Yoshikawa, T.; et al. Biochemical Sequence of Bone Formation in Porous Ceramics. In *Bioceramics*; Hulbert, S., Ed.; Rose-Hulman Institute of Technology, 1992; Vol. 3; 157–169.
 60. Katz, J.L. On the anisotropy of Young's modulus of bone. *Nature* **1980**, *283*, 106–107.
 61. Damien, C.J.; Parsons, J.R. Bone graft and bone graft substitutes: A review of current technology and applications. *J. Appl. Biomater.* **1991**, *2* (3), 187–208.
 62. Nathan, R.; et al. Osteogenesis in a Rat Femoral Defect with a Collagen-Ceramic Composite Containing Bone Marrow, Transactions of the 37th Annual Meeting, Orthopedic Research Society, Anaheim, CA, March 4–7, 1991; 406.
 63. St. John, K.R. Histological and electron microscopic analysis of tissue response to synthetic composite bone graft in the canine. *J. Appl. Biomater.* **1995**, *6* (2), 89–97.
 64. Muschler, G.F.; Huber, B.; Ullman, T.; Barth, R.; Easley, K.; Otis, J.O.; Lane, J.M. Evaluation of bone-grafting materials in a new canine segmental spinal fusion model. *J. Orthop. Res.* **1993**, *11* (4), 514–524.
 65. Stupp, S.I. Organoapatites: Materials for artificial bone. II. Hardening reactions and properties. *J. Biomed. Mater. Res.* **1993**, *27* (3), 289–299.
 66. Smith & nephew licenses composite from BTG. *Clinica* **April 3, 1995**, *648*, 14. Additional reference on page 17.
 67. Witschger, P. New bone joint stronger than natural bone. *Jpn. Chem. Week* **Oct. 3, 1996**, 3.
 68. Witschger, P. New bone joint stronger than natural bone. *Comline Biotechnol. Med.* **Oct. 9, 1996**.
 69. Neve, A.D.; Piddock, V.; Combe, E.C. Development of novel dental cements. I & II. *Clin. Mater.* **1992**, *9*, 7–20.
 70. Jones, D.W.; Rizkalla, A.S.; Sutow, E.J.; Hall, G.C. Influence of Composition on Mechanical Properties of Polyalkenoate Cements, Transactions of the Fourth World Biomaterials Congress, Berlin, Germany, April 24–28, 1992; 91.
 71. Rokkanen, P.; Bostman, O.; Vainionpaa, S.; et al. Biodegradable implants in fracture fixation: Early results of treatment of fractures of the ankle. *Lancet* **1985**, *1*, 1422–1424.



72. Rokkanen, P.U. Bioabsorbable fixation devices in orthopaedics and traumatology. *Ann. Chir. Gynaecol.* **1998**, *87*, 13–20.
73. Bostman, O.; Vainionpaa, S.; Hirvensalo, E.; et al. Biodegradable internal fixation for malleolar fractures. A prospective randomized trial. *J. Bone Jt. Surg., Br.* **1987**, *69*, 516–619.
74. Peltoniemi, H.H.; Tulamo, R.M.; Toivonen, T.; et al. Biodegradable semirigid plate and miniscrew fixation compared with rigid titanium fixation in experimental calvarial osteotomy. *J. Neurosurg.* **1999**, *90*, 910–917.
75. Yamamuro, T.; Matsusue, Y.; Uchida, A.; et al. Bioabsorbable osteosynthetic implants of ultra high strength poly-L-lactide. A clinical study. *Int. Orthop.* **1994**, *18*, 332–340.
76. Maruyama, M.; Terayama, K.; Itou, M.; Yamagishi, T. Experimental Development of a Chitosan-Bonded Hydroxyapatite Self-Hardening Paste, Transactions of the Fourth World Biomaterials Congress, Berlin, Germany, April 24–28, 1992; 93.
77. Ito, M.; Miyazaki, A.; Yamagishi, T.; Yagasaki, H.; Matsumoto Dental College, Nagano Japan Experimental development of a chitosan-bonded β -tricalcium phosphate bone filling paste. *Bio-Med. Mater. Eng.* **1994**, *4* (6), 439–449.
78. Daniel, A.U.; Chang, M.K.O.; Andriano, K.P. Mechanical properties of biodegradable polymers and composites proposed for internal fixation of bone. *J. Appl. Biomater.* **1990**, *1* (1), 57–78.
79. Gogolewski, S.; Jovanovic, M.; Perren, S.M.; Dillon, J.G.; Hughes, M.K. Tissue response and in vivo degradation of selected polyhydroxyacids: Polylactides (PLA), poly(3-hydroxybutyrate) (PHB), and poly(3-hydroxybutyrate-co-3-hydroxyvalerate) (PHB/VA). *J. Biomed. Mater. Res.* **1993**, *27* (9), 1135–1148.
80. Barrows, T.H. Semiabsorbable Bone Plate Spacer. US Patent 5,013,315, filed 12 January 1985, assigned to Minnesota Mining and Manufacturing Company, issued 7 May 1991.
81. Yoda, R.; Komatsuzaki, S.; Hayashi, T. Surface properties and biocompatibility of A-B-A type block copolymer membranes consisting of poly(γ -benzyl-L-glutamate) as the A Component and Polyisoprene as the B Component. *Biomaterials* **1995**, *16* (16), 1203–1210.
82. Ertel, S.I.; Joachim, K.Z.; Zimmerman, M.C.; Parsons, J.R. Evaluation of poly(DTH carbonate), a tyrosine-derived degradable polymer, for orthopedic applications. *J. Biomed. Mater. Res.* **1995**, *29* (11), 1337–1348.
83. Storey, R.F.; Wiggins, J.S. Design and fabrication of polyester-fiber and matrix composites for totally absorbable biomaterials. *Med. Plast. Biomater.* **1996**, *3* (2), 41–44.
84. Kadiyala, S.; Lo, H.; Ponticello, M.S.; Leong, K.W. Poly(phosphoesters): Synthesis, Physico-chemical Characterization and Biological Response. In *Biomedical Applications of Synthetic Biodegradable Polymers*; Hollinger, J.O., Ed.; CRC Press, 1995; 33–57. Chapter 3.
85. Chvapil, M.; et al. Collagen fibers as a temporary scaffold for replacement of ACL in goats. *J. Biomed. Mater. Res.* **1993**, *27* (3), 313–325.
86. Bulstra, S.; et al. *Analysis of the Osteoconductive Properties of Polyactive in the Rabbit Femur by Dual Energy X-Ray Absorptiometric Analysis and Quantitative Histological Analysis*, 40th Annual Meeting, Orthopedic Research Society, New Orleans, Feb. 21–24, 1994; 515.
87. Whitecloud, T.S.; Cook, S.D.; Dalton, J.E.; Prewett, A. *Reduction in Perineural Scar Formation Following Laminectomy with Polyactive Membrane Sheets*, 40th Annual Meeting, Orthopedic Research Society, New Orleans, Feb. 21–24, 1994; 711.
88. Cook, S.D.; Prewett, A.B.; Dalton, J.E.; Whitecloud, T.S., III. Reduction in perineural scar formation after laminectomy with polyactive membrane sheets. *Spine* **1994**, *19* (16), 1815–1825.
89. Conversation of Tom Sander of Polymedics (partnership of U.S. Surgical and Biomet), Warsaw, IN [219 267-6639x1663] with CRM on 6 September 1994.
90. Damien, C.J. A Comparative Analysis of Various Bone Graft Substitutes. Ph.D. Thesis; Rutgers The State University of New Jersey: New Brunswick, 1990. U.M.D.N.J.
91. Cifkova, I.; Lopour, P.; Vondracek, P.; Jelinek, F. Silicone rubber-hydrogel composites as polymeric biomaterials. I. Biological properties of the silicone rubber-p(HEMA) composite. *Biomaterials* **1990**, *11* (6), 393–396.
92. Lopour, P.; Vondracek, P.; Janatova, V.; Sulc, J.; Vacik, J. Silicone rubber-hydrogel composites as polymeric biomaterials. II. Hydrophilicity and permeability to water-soluble low-molecular-weight compounds. *Biomaterials* **1990**, *11* (6), 397–402.
93. Reissis, N.; Downes, S.; Bentley, G.; IRC Biomedical Materials, Middlesex UK. Cartilage Repair Using a New Synthetic Polymer, 40th Annual Meeting, Orthopedic Research Society, New Orleans, Feb. 21–24, 1994; 482.
94. Hydrophilic-hydrophobic binary systems of poly(2-hydroxyethyl methacrylate) and polycaprolactone. Part I: Synthesis and characterization. *J. Bioact. Compat. Polym.* **January 1994**, *9*. Also “Part II: Degradation” in April 1994.
95. Zachariades, A.E. *A New Class of UHMWPE Orthopedic Prosthetic Devices with Enhanced Mechanical Properties*, Transactions of Fourth World Biomaterials Congress, Berlin, April 24–28, 1992; 623.
96. Kerns, L. Clinical trials to begin for injectable fracture-healing bone mineral substitute. *Orthop. Today* **1995**, *15* (2), 23–24.

Orthopedic Fixation Devices

Jill S. Kawalec

Ohio College of Podiatric Medicine, Cleveland, Ohio, U.S.A.

INTRODUCTION

Broken bones, whether accidental (fracture) or intentional (osteotomy) must be fixed in place in order for proper healing to occur. Torn ligaments and tendons must be reattached to bone, so that proper functioning of a joint or limb may occur. Orthopedic fixation devices provide for such needs. There are two categories of orthopedic fixation: internal fixation and external fixation. Within each category are several types of devices with which to fix a fracture or reattach soft tissue. The devices may be made from metals or absorbable polymers. This section will introduce several orthopedic devices currently used by surgeons.

INTERNAL FIXATION DEVICES

Absorbable Fixation

Within the past 15 years or so, the use of absorbable polymers for internal fixation has become very popular. Absorbable polymers currently in use include polydioxanone (PDO), polyglyconate, pure poly(glycolic acid) (PGA), pure poly (L-lactic acid) (PLLA), and PGA-PLLA copolymers. Of these materials, pure PGA has the greatest strength, while PDO and polyglyconate have the least strength.^[1] The majority of studies with absorbable polymers have used pure PGA, pure PLLA, or copolymers of the two.

Complications have been observed with absorbable polymers, however, these are primarily observed with the homopolymers, as opposed to copolymers.^[1] For instance, pronounced fibrous encapsulation was observed with PLLA,^[1,2] whereas sterile sinus formation^[1,3] and bone osteolysis^[1,4] were observed with PGA. The most likely cause of these complications is the degradation characteristics of the material. PLLA absorbs slowly in the body, and thus, the formation of a fibrous capsule around the implant can occur. In addition, the amorphous regions of PLLA degrade faster than the crystalline regions, resulting in the presence of crystals at the site of the implant. This may result in the formation of a fibrous capsule around the device. Conversely, PGA absorbs very quickly, with nearly all strength lost after one month and all mass lost after six to 12 months. Thus, an abundance

of monomer in the area may lower the pH and result in osteolysis or sinus formation.

Copolymers of PGA and PLLA combine the properties of both materials to produce a device that can be used in orthopedic fixation. By varying the ratio of PGA to PLLA, the properties of the material can be varied. For instance, a study that examined the behavior of a copolymer consisting of 10% PLLA and 90% PGA found that, after three to four months, sinus formation occurred in a small portion of patients who received bone pins in the ankle.^[1,5] In contrast, a study that examined the behavior of a copolymer consisting of 82% PLLA and 18% PGA demonstrated that after eight weeks in physiological buffered saline, the peak flexural strength of the copolymer was greater than that of a 1.5-mm titanium plate.^[1] This was most likely due to the slower absorption rate of PLLA.

The ideal absorbable implant can be made with sufficient initial strength for the purpose it will serve. The rate of degradation must be predictable so that adequate strength of the device remains while healing occurs. The particles and chemicals that are leached from the implant must not have a harmful effect on the body. Finally, once the implant is completely absorbed by the body, there should be no evidence that it ever existed.^[1]

Internal fixation devices that are made from absorbable polymers currently include pins, screws, and plates for fracture fixation. Additional absorbable devices include interference screws for repair of the anterior cruciate ligaments,^[6] staples for meniscal repair, and suture anchors, tacks, and staples to aid in soft tissue fixation to bone.^[7]

Fracture Fixation

Over the past 40 years, interfragmental compression has been the most successful means of immobilization of bone fragments, whether of a fracture or an osteotomy. This can be obtained by internal fixation. The principle of AO internal fixation is “early active pain-free mobilization of the muscles and joints of the injured limbs in patients with multiple fractures.”^[8] Four conditions must be satisfied to obtain the best internal fixation and healing of the surrounding tissue: 1) anatomic reduction of the bone fragments; 2) stable

O

internal fixation with a device that is appropriate for the location and type of fracture; 3) early, active, and pain-free mobilization of the soft tissue and joints surrounding the fracture or osteotomies; and 4) preservation of the blood supply to both the bone fragments and the surrounding tissue.^[8]

There are several different types of internal fixation devices that can be used, depending on type and location of the fracture or osteotomy. The concepts behind orthopedic internal fixation devices, particularly those used for fracture or osteotomy fixation, are stress-sharing and stress-shielding.^[9] A stress-sharing device allows for partial transmission of the load across the fracture site. This allows for micromotion at the site of the fracture, and, as a result, micromotion can occur at the site. Due to this micromotion, secondary healing occurs with the formation of a callus.^[9]

A stress-shielding device protects the fracture or osteotomy from functional loads by transferring the stress from the bone to the device itself. Such devices hold the ends of the bone together in compression, allowing negligible motion at the fracture site. As a result, healing is accomplished by primary healing without callus formation.^[9] One concern that must be addressed with a stress-shielding device is the phenomena of stress protection. When a rigid device is present for an extended period of time, bone mass around the device decreases, as a result of Wolff's law. This theory states that the amount of bone present is dependent upon the need for it, and, in turn, the need for bone is based upon the amount of stress applied to the area. When a rigid device is placed adjacent to bone, the stress is carried almost entirely by the device, and the bone is protected from the stress. As a result, the bone may undergo atrophy or loss of bone mass. Another consequence of stress protection is that the neutral axis of the bone, which is normally located in the medullary canal, shifts toward the device.

Rods, pins, and wires

The simplest form of internal fixation is a single rod, pin, or wire placed across a fracture or osteotomy, resulting in partial immobilization at the site. All three devices are stress-sharing devices, and thus, more rapid secondary healing with callus formation typically occurs. Rods, pins, and wires are not recommended for fracture fixation at weight-bearing locations. If the devices are inserted in such a location, immediate weight bearing is not recommended. Often, a second type of fixation is used, such as a cast, to increase the stability of the fixation. Rods, pins and wires are commonly used in fractures of the ankle, patella, metacarpals, and olecranon.^[9] Rods and pins may be

made from metals or from absorbable polymers, such as PLLA, or composites, such as self-reinforced PLLA. Wires are typically made from stainless steel. If the device is made from metal, it is common to remove it once the fracture has healed.^[9]

Wires also are used in tension band fixation. Every bone experiences bending stresses during functional loading. As the bone bends, the convex portion of the bone experiences tension, while the concave side of the bone undergoes compression. The principle behind the tension band is that the implant absorbs the tension during functional loading, while the bone absorbs the compressive stresses.^[10,11] When a fracture occurs, a gap usually forms between fragments on the tensile side of the bone, and tension band fixation aids in compressing the bone fragments. In order for the bone to function normally, a band must be placed on the tensile side of the fracture to assist in absorbing the stress. Tension band wiring is primarily indicated for avulsion fractures, fixation of arthrodeses, and for the attachment of corticocancellous grafts.^[11]

Tension band wires can be used in conjunction with Kirschner wires, or K-wires.^[10,11] This is the most common method of fixation for a fracture of the olecranon.^[10,12] With this method, two K-wires are placed axially parallel to one another. The addition of the K-wires serves two purposes: they add to the rotational stability and also serve as an anchor for the tension band wire. It is important that the K-wires are inserted parallel to one another. If not, the bone fragments may not achieve the desired compression, and the rotational stability may be compromised.^[10,11]

K-wires can also be used alone for internal fixation. K-wires are thin, smooth metal rods.^[9] They are available in diameters of 0.6, 0.8, 1.0, 1.25, 1.6, and 2.0 mm.^[13] The K-wire can be used by passing the wire across the fracture and then cutting the wire to size. K-wires are commonly used in fractures or osteotomies of the foot. For fractures of the lesser phalanx where alignment of the fragments cannot be achieved by closed methods, a K-wire is placed across the fracture for approximately two to three weeks, during which time the foot is placed in a small cast. After the required time, the wire is removed.^[14] K-wires are used in a similar manner for fractures of the great toe phalanx, and the second, third, and fourth metatarsals. In the midfoot, K-wires are used for Lisfranc fractures when the injuries are primarily to the bone, since less time is required for healing.^[15] The patient is placed in a nonweight-bearing cast and the wires are removed after six weeks.

K-wires are also often used in the joints of the hand. This includes unstable fractures of the metacarpal shafts, and the majority of intraarticular fractures of the first and fifth metacarpals. The K-wires can be placed through the medullary canal, or through the

fracture and into the adjacent metacarpals. Fixation can also be obtained by crossing two K-wires. Based on the surgeon's preference, the K-wires can be cut off beneath the skin or may protrude through the skin. The wires are typically removed after three to four weeks.^[16]

Fixation screws

Internal fixation screws are stress-sharing devices, and thus, healing occurs by callus formation.^[9] A secondary form of fixation, such as a cast, is often used in conjunction with a single screw across a fracture or osteotomy. Once healing occurs, the screw is often removed.

There are two major types of screws: cancellous screws and cortical screws (see Fig. 1).^[10] The cancellous screw is typically wider in the thread area, in comparison to the shaft of the screw. It may be fully or partially threaded. These screws are normally used in the epiphyseal and metaphyseal region of bone. If placed in cortical bone, the cancellous screw cannot be removed. As it heals, the cortical bone grows right up to the shaft. In reverse, the threads are unable to cut through the cortical bone and the screws often break.^[10] Cortical screws are fully threaded, and the thread size is not as large as the cancellous screw.

The lag screw is one of the best ways to achieve interfragmental compression.^[9,10] The lag effect occurs when the screw gains purchase in the far cortex, but passes freely through the cortex near the screw head and the fracture or osteotomy. Partially threaded cancellous screws can inherently be lag screws, as long as the smooth shaft of the screw crosses the fracture or osteotomy. A cortical screw can only be used as a lag screw when the cortex adjacent to the head of the screw is overdrilled, such that purchase is gained in the far cortex alone.^[10]



Fig. 1 Cancellous (left) and cortical (right) fixation screws.

Fixation Plates

Fixation plates are typically metal plates with a curved surface that can contour to the shape of the bone. Because of the rigidity of the metal plate, it is considered a stress-shielding device, and thus, healing is achieved by the much slower method of primary healing without callus formation. Since the load is essentially completely absorbed by the plate, failure of the plate must be considered in selecting the proper material for the plate, and bone resorption must be taken into consideration. Often, a secondary fixation method, such as a cast, is used in conjunction with the plate.^[9]

Since the metal fixation plate is a stress-shielding device, loss of bone mass beneath the device is a concern. The weakened bone may be susceptible to fracture once the plate is removed. In addition, the screws and the edges of the plate are points of discontinuity, which may result in increased bone formation. Thus, less rigid fixation is preferred. This may be accomplished by reducing the stiffness of the metal plate by making the device thinner. Another option is to manufacture the device from a material with a modulus of elasticity similar to that of bone. These materials include polymers, such as polyetheretherketone (PEEK), or composites, such as carbon fiber reinforced PEEK.^[17] These materials, however, have not been used extensively in clinical applications.

According to the AO group, fixations serve one or more of four different functions: 1) static compression; 2) dynamic compression; 3) neutralization; and 4) buttressing.^[10] With static compression, the plate is prestressed in tension, and, when fixed into place, results in axial compression of the fracture. This is primarily indicated for fractures of the arm, since one plate without the use of a lag screw inserted obliquely across the fracture should never be used in the tibia or femur.^[9,10]

With dynamic compression, the plate serves as a tension band, and absorbs all of the tensile stresses while the osteotomy or fracture is exposed only to compressive stresses.^[9,10] This technique is not recommended for new fractures, as the location of tensile stresses cannot immediately be determined.

Neutralization is the most common function of fixation plates.^[10] Interfragmental stability is first achieved by the application of lag screws. They may be inserted by themselves or passed through the plate. Once this fixation has been achieved, a protection plate is fixed over the fracture site. The function is termed "neutralization" because the plate serves to protect the lag screw fixation from torsional, bending, and shear forces. The primary application for neutralization is fracture of the tibia, excluding transverse fractures. It should be noted that all short oblique fractures should

first be fixed with lag screws prior to application of the plate, in order to achieve stable fixation.^[10]

A buttress plate is typically a thin metal plate that serves to protect the cortex. It is a stress-sharing device, and thus, healing occurs by callus formation. Weight bearing is normally delayed.^[9] The most common function of buttressing is in conjunction with a lag screw on the proximal tibia after fracture of the tibial plateau has occurred.^[9] Buttressing can also be used to prevent a cancellous bone graft from collapsing.^[10]

Intramedullary nails

Intramedullary nails are stress-sharing rods that are most commonly used for compound fractures of the tibia and femur, and occasionally, the humerus.^[9,18,19] The devices are placed centrally in the medullary canal, at the mechanical axis of the shaft (see Fig. 2).^[18] At this location, they are not susceptible to functional bending loads that may be introduced. Intramedullary nails allow for movement of the joints above and below the fracture while preventing motion at the fracture site.^[9]



Fig. 2 Schematic of a cross section of tibia, showing placement of an intramedullary nail in the medullary canal.

There are two types of intramedullary nails: reamed and unreamed. Reamed nails have a larger transverse diameter, and thus, are a stronger device. They are commonly used in the shaft of the femur and the tibia, where functional loading is high. The reaming process, however, may result in further destruction of the blood supply to the endosteal bone, thus causing delayed healing. Unreamed nails are much thinner and have less strength than reamed nails. With unreamed nails, the bone shaft is not reamed prior to insertion of the nail, and therefore, the endosteal blood supply is maintained.^[9,18]

Both reamed and unreamed nails can be locked into place by screws that are passed transversely through the nail and two cortices both distally and proximally.^[9,18,20] Interlocking screws are primarily used when the fracture is comminuted or segmented and the fragments are not in contact.^[20] The locked intramedullary nail has become the accepted method of fixation of fractures in the femur.^[19] In this instance, the nail is locked into place to prevent rotation and sliding, and acts more as a stress-shielding device until the fracture has healed sufficiently to begin to bear weight. Once callus formation begins, the screws located either proximally or distally can be removed to create compression at the fracture site to further stimulate healing.^[9,18,20]

The screw holes in the locking intramedullary nail introduce areas of high stress concentration to the device. Thus, material selection for the intramedullary nail is very important. The nail must be made of a material that is biocompatible, with good ductility, high fatigue resistance, and high mechanical strength. The metals most commonly used are titanium alloyed with 6% aluminum and 4% vanadium, cobalt–chromium alloy, and stainless steel.^[21] Stainless steel nails have been shown to be more susceptible to failure at the screw hole compared to the other materials.^[22,23] Nails made from the titanium alloy have an increased strength and flexibility and are more biocompatible than the other metals.^[19]

Soft Tissue Attachment to Bone

Suture anchors

Suture anchors are fixation devices with an integrated suture. The suture is passed through an eyelet in the anchor and is used to secure the soft tissue in place. There are three design types for suture anchors: harpoon, wedge, and screw. The harpoon-style anchors have a smooth surface but possess arcs, or prongs, to hold the anchor in bone. There may be two or four arcs, depending on the brand of anchor. Failure of this type of device usually occurs at the

suture knot.^[24] The wedge-type suture anchor is V-shaped. When inserted, the wedge closes and then reopens when in bone, thus preventing the anchor from backing out. The screw-style suture anchors possess threads that allow the device to be screwed into position.

Suture anchors for fixation of soft tissue to bone may be made from metal or absorbable polymers. Metal suture anchors are typically titanium alloy or stainless steel. Absorbable polymers are becoming more popular for use as suture anchors. Such materials include PLLA, PGA, and PLLA-PGA copolymers. Applications for suture anchors include repair of rotator cuffs, Bankart procedures, treatment of acromioclavicular joint separation,^[25] and reattachment of the Achilles tendon.^[26] Mini suture anchors have been used for thumb ulnar collateral ligament reconstruction^[27] and correction of chronic recurrent dislocation of the mandibular condyle.^[28]

Soft tissue staples

Staples are another type of device used to attach soft tissue to bone. These devices are typically U-shaped and consist of at least two legs and a bridge between the legs. The legs are parallel or nearly parallel to one another, and serve to penetrate the bone and hold the tissue in place. The legs may possess barbs, or serrations, to enhance fixation in the bone. The cross section of the legs may be round, square, or polygonal. The bridge of the staple may be either smooth or possess barbs on the underside to enhance fixation into the soft tissue.

The effectiveness of the staple depends largely on the size and quality of the tissue being fixed to bone, differences in the contour of the bone, and the degree of osteopenia.^[29] Several complications may occur with staples, including failure of the barbed staple as a result of the spikes tearing through the soft tissue,^[29] loosening of the staple, and pain that requires removal of the staple.^[24] Impingement of the staple bridge may occur, resulting in an overlying bursa.^[29]

Staples for the fixation of soft tissue to bone may be made from metals, including cobalt-chromium alloy. Metal implants can lead to several of the complications listed above. Thus, staples are also often made from absorbable materials, including PGA-PLLA copolymers and self-reinforced PLLA (SR-PLLA). Applications of staples include repair of the meniscus in the knee and reattachment of the anterior glenohumeral ligament in the shoulder joint.^[24]

Soft tissue tacks

Tacks are also used for fixation of soft tissue to bone. They are passed through the soft tissue and into the

bone. Some designs consist of spikes on the underside of the head, or have spiked washers to enhance stability in the soft tissue. Certain designs do not require drilling or tapping prior to insertion of the tack.

Tacks are typically made from absorbable polymers. Such polymers include PLLA, PGA, and a combination of both polymers. Applications include rotator cuff repair,^[30] repair of acromioclavicular separation, Bankart procedures, superior labrum anterior-posterior (SLAP) lesion repairs,^[31] and craniofacial surgery.^[32]

Interference screws

The interference screw is another device used for the fixation of soft tissue to bone. They are available with a round or flat head, and may be fully or partially threaded. They are most commonly used for repair or reconstruction of the anterior cruciate ligament (ACL) in the knee. In fact, it has become the “gold standard” for bone-patellar tendon-bone grafts in ACL surgery.^[6,33] Other applications include bone-tendon grafts and reattachments. The screws are placed with an interference fit at the site of insertion of tendon into bone. Interference screws can be made of metal or absorbable polymers. Metals include titanium alloy, while absorbable polymers include PLLA, PLLA/PGA copolymers, and a composite of tricalcium phosphate (TCP) and PLLA.^[6]

EXTERNAL FIXATION DEVICES

Casts

Casts are the most commonly used method of stabilizing a fracture or osteotomy. A cast is a stress-sharing device, which allows for more rapid secondary healing.^[9] Typically made from plaster of Paris, they are situated such that the joints above and below the fracture are immobilized to prevent excessive motion at the site of fracture. In most cases, casts allow for early weight bearing on the fractured bone.^[9]

External Fixators

The external fixator for the stabilization of bone fragments dates back to 1902.^[34] Its function is to maintain alignment of a fracture while allowing the patient to be mobile. It is a device that allows for temporary stress-shielding to protect and stabilize the site of a fracture or an osteotomy. It can also be used as a permanent stress-sharing device, where the load is distributed both through the bone and the fixator.^[35]



There are several types of external fixators. The ring fixator was developed by Ilizarov in the 1950s. It is based on the theory that bone union can be stimulated by pulling apart, or distracting, the bone fragments, as well as by compression.^[36] It is designed to hold the bone fragments in place while allowing axial dynamization at the fracture site.^[37] The ring fixator works by passing K-wires under tension or stiff full or one-half pins through the bone segment.^[37,38] The wires or pins are then secured to a full- or semi-circular ring. The ring is then fixed by threaded rods. Hinges, telescoping rods, and spacers may also be added to the construct.^[37] With such a device, adjustments can be made to the length and angulation after initial fixation.^[39] Ring fixators are commonly used for tibial osteotomies,^[40] tibial nonunions^[41,42] and compound fractures,^[43] and angular and rotational deformities of the lower extremity.^[44]

Two other types of external fixators consist of unilateral and bilateral frames.^[45] Unilateral frames include one external rod and a minimum of two transfixing pins, which are placed above and below the fracture. The pins are inserted through the skin and into the bone, and are held in place by the external rod. Bilateral frames consist of two external rods, one on each side of the bone, and at least two transfixing pins. As with the unilateral frame, the pins are fixed on one end by one of the external rods and the pins are inserted transcutaneously into the bone. In the case of bilateral fixation, however, the pins pass completely through the bone and are attached to the second external rod.^[45]

Bilateral frames are intrinsically more stable than unilateral frames. During functional loading, the pins in a bilateral frame may bend, while the stress through the external rods is distributed evenly, thus preventing significant bending. In contrast, when the unilateral frame is loaded, the transfixing pins undergo considerable bending, resulting in observable bending of the external rod.^[45]

External fixators are most commonly used for the stabilization of complex fractures, usually in the tibia.^[35] They are very rarely used for fixation of the femur or of the upper extremities, with the exception of the bones around the wrist. An external fixator is only used under extreme conditions, where there is severe injury and contamination to the surrounding soft tissue. It is also used when there has been vascular injury and prompt stabilization is required in order to rapidly restore circulation to the limb. An external fixator can also be used to repair compound or closed fractures of the lower extremities when the patient is too ill or injured to undergo extensive surgery, and for nonunions and osteotomies.^[35]

External fixators are also used in pediatrics for staged limb lengthening. When used for lengthening

purposes, an osteotomy is created and the lengthening frame is inserted into place. As the fragments are separated, a defect is created in the bone. Once the desired length is achieved, a second surgery is needed to pack the defect with cancellous bone.^[34] Other applications of external fixators are as compression frames to hold bone fragments together after an arthrodesis or when a nonunion occurs, or as soft-tissue frames to stretch soft tissues in order to correct contractures, such as an equinus deformity about the ankle.^[34]

CONCLUSION

There are many orthopedic fixation devices available, whether to repair a fracture or fix an osteotomy, or to fix soft tissue to bone. The devices include both internal and external fixation devices. The device can be made from metals, or more recently, absorbable polymers. Research continues to find means by which to improve the performance and biocompatibility of orthopedic fixation devices.

ARTICLES OF FURTHER INTEREST

Bone Cement; Bone Fracture Fixation; Bone Plates and Screws; Bone Remodeling; Orthopedic Biomaterials; Poly(Glycolic Acid); Poly(lactic acid)s; Sutures; Tendons and Ligaments, Mechanical Testing Of; Titanium and Its Alloys

REFERENCES

1. Pietrzak, W.S.; Sarver, D.R.; Verstynen, M.L. Bio-absorbable polymer science for the practicing surgeon. *J. Craniofac. Surg.* **1997**, *8*, 87–91.
2. Bergsma, J.E.; de Bruijn, W.C.; Rozema, F.R.; Bos, R.R.M.; Boering, G. Late degradation tissue response to poly (L-lactide) bone plates and screws. *Biomaterials* **1995**, *16*, 25–31.
3. Partio, E.K.; Bostman, O.; Hirvensalo, E.; Patiala, H.; Vainionpaa, S.; Vihtonen, K.; Tormala, P.; Rokkanen, P. The indication for the fixation of fractures with totally absorbable SR-PGA screws. *Acta Orthop. Scand.* **1990**, *61* (Suppl. 237), 43.
4. Bostman, O.M. Osteolytic changes accompanying degradation of absorbable fracture fixation implants. *J. Bone Jt. Surg., Br.* **1991**, *73* (4), 679–682.
5. Bostman, O.; Vainionpaa, S.; Hirvensalo, E.; Makela, A.; Vihtonen, K.; Tormala, P.; Rokkanen, P. Biodegradable internal fixation for malleolar fractures: A prospective randomized trial. *J. Bone Jt. Surg., Br.* **1987**, *69* (4), 615–619.

6. Kousa, P.; Jarvinen, T.L.N.; Kannus, P.; Jarvinen, M. Initial fixation strength of bioabsorbable and titanium interference screws in anterior cruciate ligament reconstruction: Biomechanical evaluation by single cycle and cyclic loading. *Am. J. Sports Med.* **2001**, *29*, 420–425.
7. Pietrzak, W.S. Critical concepts of absorbable internal fixation. *J. Craniofac. Surg.* **2000**, *11*, 335–341.
8. Muller, M.E.; Allgower, M.; Schneider, R.; Willenegger, H. *Manual of Internal Fixation: Techniques Recommended by the AO-Group. Second Edition, Expanded and Revised*; Springer-Verlag: Berlin, 1979; 3–27.
9. Hoppenfeld, S. Biomechanical Principles of Fixation Devices. In *Treatment and Rehabilitation of Fractures*; Hoppenfeld, S., Murthy, V.L., Eds.; Lippincott Williams and Wilkins: Philadelphia, 2000; 11–18.
10. Muller, M.E.; Allgower, M.; Schneider, R.; Willenegger, H. *Manual of Internal Fixation: Techniques Recommended by the AO-Group. Second Edition, Expanded and Revised*; Springer-Verlag: Berlin, 1979; 27–173.
11. Heim, U.; Pfeiffer, K.M. Implants and Instruments. In *Internal Fixation of Small Fractures: Technique Recommended by the AO-ASIF Group. Third Edition of the Small Fragment Set Manual*; Springer-Verlag: Berlin, 1988; 34–35.
12. Gaudinez, R.F.; Murthy, V.L.; Hoppenfeld, S. Olecranon Fractures. In *Treatment and Rehabilitation of Fractures*; Hoppenfeld, S., Murthy, V.L., Eds.; Lippincott Williams and Wilkins: Philadelphia, 2000; 141–153.
13. Heim, U.; Pfeiffer, K.M. Implants and Instruments. In *Internal Fixation of Small Fractures: Technique Recommended by the AO-ASIF Group. Third Edition of the Small Fragment Set Manual*; Springer-Verlag: Berlin, 1988; 5–30.
14. McCormack, A.P.; Hoppenfeld, S. Ankle fractures. In *Treatment and Rehabilitation of Fractures*; Hoppenfeld, S., Murthy, V.L., Eds.; Lippincott Williams and Wilkins: Philadelphia, 2000; 483–511.
15. McCormack, A.P. Midfoot Fractures. In *Treatment and Rehabilitation of Fractures*; Hoppenfeld, S., Murthy, V.L., Eds.; Lippincott Williams and Wilkins: Philadelphia, 2000; 461–482.
16. Kram, D.A.; Murthy, V.L. Metacarpal Fractures. In *Treatment and Rehabilitation of Fractures*; Hoppenfeld, S., Murthy, V.L., Eds.; Lippincott Williams and Wilkins: Philadelphia, 2000; 223–240.
17. Jockisch, K.A.; Brown, S.A.; Bauer, T.W.; Merritt, K. Biological response to chopped-carbon-fiber-reinforced PEEK. *J. Biomed. Mater. Res.* **1992**, *26* (2), 133–146.
18. Taffet, R. Management and Classification of Compound Fractures. In *Treatment and Rehabilitation of Fractures*; Hoppenfeld, S., Murthy, V.L., Eds.; Lippincott Williams and Wilkins: Philadelphia, 2000; 65–72.
19. Im, G.-I.; Shin, S.-R. Treatment of femoral shaft fractures with a titanium intramedullary nail. *Clin. Orthop.* **2002**, *401*, 223–229.
20. Klemm, K.W. Interlocking Nailing of Complex Fractures. In *Concepts in Intramedullary Nailing*; Seligson, D., Ed.; Grune & Stratton, Inc.: Orlando, 1985; 293–317.
21. Gleis, G.E.; von Fraunhofer, J.A.; Seligson, D. Material Considerations in Medullary Nailing. In *Concepts in Intramedullary Nailing*; Seligson, D., Ed.; Grune & Stratton, Inc.: Orlando, 1985; 51–68.
22. Bucholz, R.W.; Ross, S.E.; Lawrence, K.L. Fatigue failure of the interlocking nail in the treatment of fracture of the distal part of the femoral shaft. *J. Bone Jt. Surg.* **1987**, *69A*, 1791–1799.
23. Franklin, J.L.; Winqvist, R.A.; Benirschke, S.K.; Hansen, S.T. Broken intramedullary nails. *J. Bone Jt. Surg.* **1988**, *70A*, 1463–1471.
24. Shall, L.M.; Cawley, P.W. Soft tissue reconstruction in the shoulder: Comparison of suture anchors, absorbable staples and absorbable tacks. *Am. J. Sports Med.* **1994**, *22* (5), 715–718.
25. Breslow, M.J.; Jazrawi, L.M.; Bernstein, A.D.; Kummer, F.J.; Rokito, A.S. Treatment of acromioclavicular joint separation: Suture or suture anchors? *J. Shoulder Elbow Surg.* **2002**, *11* (3), 225–229.
26. Janis, L.; Lam, A.T.; Espiritu, T.; Ploot, E.; Husain, S. A comparison of soft-tissue anchors in tendo achilles reattachment. *J. Foot Ankle Surg.* **2001**, *40* (4), 195–207.
27. Weiland, A.J.; Burner, S.H.; Hotchkiss, R.N.; McCormack, R.R., Jr.; Gerwin, M. Repair of acute ulnar collateral ligament injuries of the thumb metacarpophalangeal joint with an intraosseous suture anchor. *J. Hand Surg.* **1997**, *22A*, 585–591.
28. Wolford, L.M.; Pitta, M.C.; Mehra, P. Mitek anchors for treatment of chronic mandibular dislocation. *Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endo.* **2001**, *92*, 495–498.
29. Robertson, D.B.; Daniel, D.M.; Biden, E. Soft tissue fixation to bone. *Am. J. Sports Med.* **1986**, *14* (5), 398–403.
30. Goradia, V.K.; Mullen, D.J.; Boucher, H.R.; Parks, B.G.; O'Donnell, J.B. Cyclic loading of rotator cuff repairs: A comparison of bioabsorbable tacks with metal suture anchors and transosseous sutures. *Arthroscopy* **2001**, *17* (4), 360–364.
31. Trusler, M.L.; Bryan, W.J.; Ilahi, O.A. Anatomic and radiographic analysis of arthroscopic tack placement into the superior glenoid. *Arthroscopy* **2002**, *18* (4), 366–371.
32. Cohen, S.R.; Holmes, R.E.; Amis, P.; Fitchner, H.; Shusterman, E.M. Tacks: A new technique for craniofacial fixation. *J. Craniofac. Surg.* **2001**, *12* (6), 596–602.
33. Costi, J.J.; Kelly, A.J.; Hearn, T.C.; Martin, D.K. Comparison of torsional strengths of bioabsorbable screws for anterior cruciate ligament reconstruction. *Am. J. Sports Med.* **2001**, *29* (5), 575–580.
34. Weber, B.G. General Part. In *External Fixator*; Weber, B.G., Magerl, F., Eds.; Springer-Verlag: Berlin, 1985; 4–26.
35. Taffet, R. Management and Classification of Compound Fractures. In *Treatment and Rehabilitation of Fractures*; Hoppenfeld, S., Murthy, V.L., Eds.; Lippincott Williams and Wilkins: Philadelphia, 2000; 65–72.
36. Martinez, R.L.; Shampo, M.A.; Kyle, R.A.; Gavriel, A. Ilizarov—Russian surgeon develops the external fixator. *Mayo Clin. Proc.* **2002**, *7*, 304.



37. Fleming, B.; Paley, D.; Kristiansen, T.; Pope, M. A biomechanical analysis of the Ilizarov external fixator. *Clin. Orthop.* **1989**, *241*, 95–105.
38. Behrens, F. General theory and principles of external fixation. *Clin. Orthop.* **1989**, *241*, 15–23.
39. Behrens, F. A primer of fixator devices and configurations. *Clin. Orthop.* **1989**, *241*, 5–14.
40. Duda, G.N.; Sollmann, M.; Sporrer, S.; Hoffman, J.E.; Kassi, J.-P.; Khodadadyan, C.; Raschke, M. Interfragmentary motion in tibial osteotomies stabilized with ring fixators. *Clin. Orthop.* **2002**, *396*, 163–172.
41. Paley, D.; Catagni, M.A.; Argnani, F.; Villa, A.; Benedetti, G.B.; Cattaneo, R. Ilizarov treatment of tibial nonunions with bone loss. *Clin. Orthop.* **1989**, *241*, 146–165.
42. Schwartzman, V.; Choi, S.H.; Schwartzman, R. Tibial nonunions: Treatment tactics with the Ilizarov method. *Orthop. Clin.* **1990**, *21* (4), 639–653.
43. Dagher, F.; Roukoz, S. Compound tibial fractures with bone loss treated by the Ilizarov technique. *J. Bone Joint Surg., Br* **1991**, *73-B*, 316–321.
44. Shtarker, H.; Volpin, G.; Stolerio, J.; Kaushansky, A.; Samchukov, M. Correction of combined angular and rotational deformities by the Ilizarov method. *Clin. Orthop.* **2002**, *402*, 184–195.
45. Weber, B.G. On the Biomechanics of External Fixation. In *The External Fixator*; Weber, B.G., Magerl, F., Eds.; Springer-Verlag: Berlin, 1985; 27–53.

Osteoclastic Bone Resorption

Samuel C. Ramage

Department of Biochemistry, Richmond, Virginia, U.S.A.

Mengnai Li

Department of Orthopaedic Surgery, Richmond, Virginia, U.S.A.

Matthew J. Beckman

Department of Orthopaedic Surgery & Biochemistry, Richmond, Virginia, U.S.A.

INTRODUCTION

Osteoclasts are myeloid-derived multi-nuclear cells responsible for bone degradation in the body. Bone is composed of a highly collagenous extracellular matrix made by mesenchymally-derived osteoblasts. The matrix is mineralized by osteoblasts when they generate and release matrix vesicles that initiate crystallization of hydroxyapatite on the bone matrix. Hydroxyapatite, $[\text{Ca}_3(\text{PO}_4)_2]_3\text{Ca}(\text{OH})_2$, adds strength to the bone and serves as the main store of calcium in the body. The activity of osteoclasts is highly regulated to maintain both the structural integrity of the bone and calcium homeostasis.

Hematopoietic stem cells in the bone marrow give rise to circulating mononuclear osteoclast precursors. The resorption cycle characterizes the events in the life cycle of the osteoclast. First, the circulating osteoclast precursors are recruited to the site of resorption, where the osteoclast undergoes activation following adherence to the bone matrix. Activation is evident in morphological changes and polarization of the cell that prepare the osteoclast to degrade bone. The activated osteoclast then dissolves the hydroxyapatite crystals of the bone. Following demineralization, the organic matrix is degraded by secreted proteins from the osteoclast. The resulting enzymatically-cleaved products are removed by the cell through a transcytotic process. The final step in the cycle is the termination of osteolytic bone resorption. The processes involved in the resorption cycle are largely cytokine mediated, and many of the signaling pathways involve both stromal and myeloid cells. The signaling pathway involving the receptor activator of NF- κ B ligand (RANKL), receptor activator of NF- κ B (RANK), and osteoprotegerin (OPG) has been implicated as an important regulator of osteoclast differentiation. Other factors, such as the proinflammatory cytokines tumor necrosis factor- α (TNF α), interleukin-1 (IL-1), interleukin-6 (IL-6), are important regulators of osteoclastogenesis and osteoclast activity.

Healthy bone continuously undergoes the process of bone remodeling, which balances bone formation with

resorption. The loss of balance between osteoblastic formation and osteoclastic degradation appears to be a key feature of bone resorptive pathologies, such as rheumatoid arthritis, periprosthetic osteolysis, osteolytic malignant cancers, and others. The signaling involved in the resorption cycle of osteoclasts offers promising therapeutic targets for these diseases.

OSTEOCLAST MORPHOLOGY

Mature, multinucleated osteoclasts are formed by the fusion of mono-nuclear osteoclast precursors. Activation occurs as the nonresorbing or inactive osteoclast reaches and adheres to the bone surface. Morphological changes allow the creation of a contained compartment between the osteoclast and the bone, called the resorption lacuna, that is the site of resorption. The degraded bone fragments are then removed and secreted away from the bone surface. The specialized morphology of the osteoclast makes it extremely efficient at resorbing bone.

Sealing Zone

Active osteoclasts are highly polarized cells with specific features characterizing the surface adjacent to the bone. After recruitment, receptors on the plasma membrane interact with the bone matrix to form the sealing zone, a tight interaction that creates a sealed extracellular compartment between the osteoclast and bone. Signaling initiated by the adherence of the osteoclast to the bone is accompanied by cytoskeletal rearrangements, such as the formation of an intracellular ring of actin filaments that aligns with the sealing zone and is important for the adherence of the cell.^[1] The adherence of the osteoclast to the bone matrix is crucial for resorption lacuna formation, as well as signaling that activates the osteoclast.

Integrins are transmembrane receptors that have important function in both signaling and adhesion.



$\alpha_v\beta_3$, $\alpha_2\beta_1$, and $\alpha_v\beta_1$ integrins are expressed in mammalian osteoclasts^[2] and have been implicated in the adherence of osteoclasts to bone. $\alpha_2\beta_1$ has been shown to bind to fibronectin^[3] in avian osteoclasts, however, work with human osteoclasts has also shown a capacity of β_1 , and in particular $\alpha_2\beta_1$ integrins to bind native type I collagen.^[4] A large amount of research has focused on $\alpha_v\beta_3$, the most highly expressed integrin in osteoclasts, because antibodies to this integrin as well as soluble ligand analogs^[5,6] prevent attachment of osteoclasts to bone and inhibit bone resorption. The $\alpha_v\beta_3$ integrin, although called the vitronectin receptor,^[7] has been shown to bind to many proteins besides vitronectin. The $\alpha_v\beta_3$ integrin binds to Arg-Gly-Lys (RGD) containing proteins.^[8] Of particular interest to bone biology is the binding of the $\alpha_v\beta_3$ integrin to osteopontin,^[3] an RGD-containing protein present in bone. Despite an apparent crucial role in osteoclast adherence, studies have been unable to show the presence of $\alpha_v\beta_3$ in the sealing zone.^[9] Instead, the $\alpha_v\beta_3$ integrin has been found only in the basolateral and ruffled border membranes adjacent to the sealing zone. Although inhibition of the $\alpha_v\beta_3$ integrin has been shown to have a large effect on bone resorption, the mechanism is not totally understood.

Integrins have been shown to act not only as mediators of adhesion, but also as signaling receptors in the osteoclasts. Integrin signaling affects cellular metabolism, cytoskeletal organization, and protein phosphorylation.^[10] The $\alpha_v\beta_3$ integrin has been shown to affect cellular changes through a variety of intracellular signaling mediators, including the cytoplasmic tyrosine receptor kinase c-Src, the adhesion-dependent kinase PYK2, and phosphatidylinositol-3 kinase.^[11] Integrin signaling in other cell types is responsible for changes in gene regulation, cytoskeletal rearrangement, and cell metabolism of proteins and lipids, suggesting that integrins are at least in part responsible for the activation of osteoclasts. Integrin binding to bone appears to be a key step in the activation of osteoclasts, although the mechanism of activation is not fully characterized.

Ruffled Border

The portion of the membrane, called the ruffled border, encircled by the sealing zone undergoes changes resulting in significant endosomal character and finger-like projections of the membrane towards the bone surface. Activation of the osteoclast triggers intracellular vesicles to fuse with the membrane contained within the sealing zone. Fusion releases HCl and other vesicular contents into the resorption lacuna.^[12,13] Fusion also increases the amount of membrane encompassed by the sealing zone, leading to the development of the

folds and extensions of the membrane into the resorption lacuna. The ruffled border is significantly different from the membrane of the rest of the cell. Experiments that utilized virally infected osteoclasts to examine protein targeting showed a lack of any viral membrane-targeted proteins in the ruffled border in actively resorbing osteoclasts.^[14] The viral proteins were found in the apical and basolateral portions of the osteoclast cell membrane. Within the ruffled border, a high concentration of vacuolar ATPase were found.^[15] Osteoclasts use the many vacuolar ATPases to pump H^+ into and further acidify the resorption lacuna. Carbonic anhydrase II is up-regulated in osteoclasts and found in high levels in the cytoplasm.^[16] The carbonic anhydrase II, in conjunction with HCO_3^-/Cl^- exchangers, work to maintain a steady level of intracellular protons, in addition to supplying them for the ATPases. Palogonkas et al. found further evidence of endosomal character in the localization of other endosomal markers, including RAB7 and Igp110 in the ruffled border.^[17] Vesicle fusion and the activity of vacuolar ATPase are key functions of the ruffled border that create an environment conducive to bone resorption within the resorption lacuna.

Functional Secretory Domain

Products of the degradation of the organic matrix of the bone are endocytosed by the osteoclast and trafficked to the apical region of the cell where they are secreted into the extracellular space. The apical region is called the functional secretory domain of the osteoclast. Salo et al.^[14] showed that there are differences between the basolateral membrane and the apical membrane by utilizing viral proteins trafficked to either the basolateral or the apical membrane of infected endothelial cells. Apically directed proteins in endothelial cells were found in the functional secretory domain of the infected osteoclasts while basolaterally directed proteins were localized in the basolateral portion of osteoclasts. While degradation products are being trafficked from the resorption lacuna, they may be further degraded by enzymes within the vesicles before they are secreted to the extracellular space. The functional secretory domain again illustrates the polarization in activated osteoclasts, as well as specialized functional domains of the cell.

MECHANISM OF BONE DEGRADATION

Mineral Crystal Dissolution

Following osteoclast activation, the resorption lacuna becomes highly acidic as two mechanisms work to

decrease the pH. The first is the fusion of intracellular acidic vesicles with the ruffled border releasing acid into the area of active resorption. The second is vacuolar ATPases, which pump protons across the membrane into the resorption lacuna. The low pH achieved is necessary to dissolve the hydroxyapatite crystals surrounding the organic matrix of the bone. Mineral crystals in bone must be dissolved for the degradative enzymes secreted by the osteoclast to access to the organic matrix. Hydroxyapatite crystals in the bone store a majority of the calcium in the body and osteoclastic dissolution of the mineral component of bone is an important mechanism for maintaining serum levels of calcium and other ions contained in bone such as phosphate. Accordingly, the activity of osteoclasts is regulated in response to conditions, such as hypocalcemia to maintain suitable concentration ranges of those ions.

Organic Matrix Degradation

Following demineralization, the osteoclast secretes enzymes that further degrade the bone. The extracellular matrix made by osteoblasts and osteocytes is principally composed of collagen, but contains other extracellular matrix proteins as well. The matrix is degraded by two types of proteases, cathepsins and metalloproteinases (MMPs). Fig. 1 illustrates the regulation of osteoclastogenesis and osteoclastic bone resorption.

Several cathepsins are present in osteoclasts and are responsible for much of the degradation of the organic matrix of bone. Cathepsins are lysosomal proteases active at acidic pH. Their activity at low pH shows the importance of the osteoclast machinery responsible for lowering the pH of the resorption lacuna, in providing a suitably acidic environment for cathepsins and other acidic proteases, in addition to dissolving the hydroxyapatite. Osteoclasts produce cathepsin B, C, D, E, G, L, and K (D and E are aspartic proteases, while the others are cysteine proteases).^[18]

Cathepsin K is a lysosomal cysteine protease that has been implicated as an important bone resorptive enzyme. Immunolocalization studies have shown cathepsin K localized to the ruffled border and lysosomal vacuoles in the osteoclast and the extracellular space surrounding the cell.^[19] Cathepsin K has high enzymatic activity towards type I collagen, the main component of bone, as well as type II collagen.^[20,21] Cathepsin K also has the ability to act on non-collagenous substrates, such as osteonectin,^[22] a component of the protein matrix of the bone. Cathepsin K has also been hypothesized to cleave the tartrate-resistant acid phosphatase (TRAP) leading to the generation of a more active enzyme.^[23] Studies utilizing

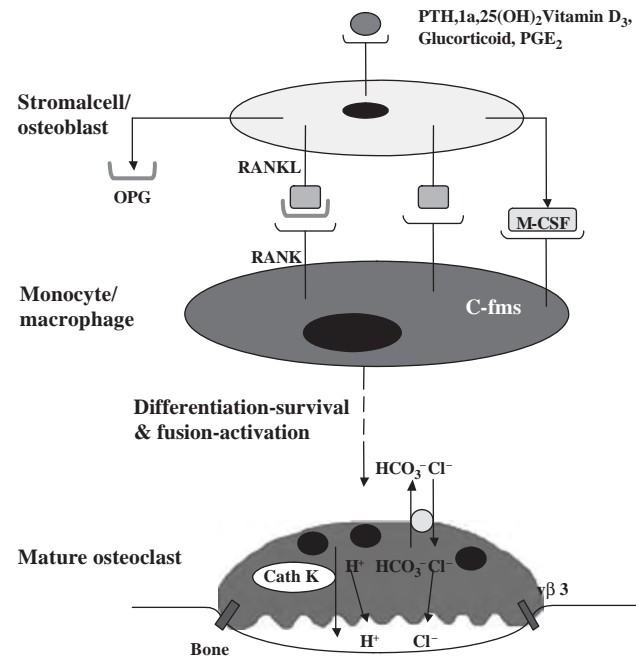


Fig. 1 Extensive research has showed that the accessory cells (osteoblasts and stromal cells) express two key molecules that are essential and sufficient to promote osteoclastogenesis: M-CSF and RANKL. Osteoclasts are multinucleated giant cells deriving from hematopoietic progenitors of the monocyte–macrophage lineage. Stromal cell/osteoblast cell express RANKL, OPG, and M-CSF and their expression is regulated by factors such as 1,25(OH)₂D₃, PTH, PGE₂, and glucocorticoids. RANKL and M-CSF will interact with the cell surface receptors on mono/macrophage cells. These will trigger the osteoclastogenesis process. These postmitotic committed precursors fuse to form multinucleated osteoclasts, which are then activated to resorb bone. Osteoblasts or bone marrow stromal cells are required as supporting cells for the in vitro differentiation of osteoclasts from their progenitor cells. Source: Adapted from American Association for the Advancement of Science (see Ref. [63]); Springer (see Ref. [64]); International and American Association for Dental Research (see Ref. [65]) and American Association for the Advancement of Science (see Ref. [66]).

cathepsin K knockout mice have shown that these animals exhibit osteopetrosis due to decrease matrix degradation, while demineralization still occurs normally.^[24] Cystatin C, a secreted protein produced by osteoclasts and osteoblasts, as well as cells in the liver and kidney,^[25] is an inhibitor of cysteine proteases. The addition of cystatin C decreased bone resorption in in vitro assays.^[26] The importance of cathepsin K is illustrated by the skeletal defects present in pycnodysostosis. In pycnodysostosis, cathepsin K deficiency leaves the ability of osteoclasts to demineralize bone unchanged, but there is no degradation of the matrix, resulting in shortened stature and skeletal deformities

in these individuals.^[27] Similarly, cathepsin K knockout mice show osteopetrosis^[24] due to their inability to degrade bone. The potent resorptive activity of cathepsin K is crucial to bone degradation and is regulated by other factors, such as cystatin, to achieve the desired level of matrix degradation.

Cathepsins, other than cathepsin K, play important roles in osteoclast function. They are present in vesicles and at the resorption border. Many, but not all, of the cathepsins made in the osteoclast are present in the resorption lacuna. The cathepsins not found in the resorption lacuna are suspected to play roles in the degradation of resorbed collagen and other matrix materials, as well as in the cleavage of inactive enzyme precursors that help generate an active enzyme.

MMPs are zinc-dependent proteases produced by a wide variety of cells in the body. Many of the MMPs present in bone show activity against collagen. Several MMPs are made by osteoclasts, including MMP-2, 9, 12, 13, and 14. The role of MMPs is not entirely clear. However, they do appear to function in bone matrix degradation, as well as being important in allowing osteoclast migration. Osteoblasts, osteocytes, and cells adjacent to and associated with osteoclasts at the bone surface, are all capable of producing MMPs in addition to osteoclasts. These MMPs, particularly those made by adjacent cells, have been implicated in bone resorption and have obscured the role of the osteoclastic MMPs. However, many experiments have shown that MMPs are active in bone degradation. MMP-13 is an essential MMP, as knockouts have shown decreased bone resorption, while other MMPs do not show any change in resorption rate.^[28] The unique degradation products of MMPs, called ICTP fragments, are present at elevated levels in the serum in mice with cancer-induced osteolysis. Similar to the regulation of cathepsin activity by cystatin C, MMPs are regulated by tissue inhibitors of metalloproteinases (TIMPs). MMPs have proven to be an important pharmacological target for treating bone resorptive pathologies, as bisphosphonates are used to inhibit and downregulate MMPs in bone resorption and other diseases, such as the metastases of cancer.^[29]

Removal of Degradation Products

Degraded collagen and other matrix proteins are removed from the resorption lacuna, transported through the osteoclast, and exocytosed to the extracellular space. The degraded collagen and other tissue from the extracellular matrix of the bone are removed by vesicles that form at the ruffled border and move the degraded bone products through the osteoclast to the functional secretory domain, where they are

secreted. Immunohistology has shown the presence of TRAP in the transcytotic vesicles, along with degraded collagen. TRAP, because of its presence in the vesicles and its ability to generate reactive oxygen species, has been suggested as an enzyme that further degrades bone matrix products after they are taken into the transcytotic vesicles. Staining techniques to identify osteoclasts utilize the presence of TRAP activity in osteoclasts. The release of matrix fragments through the functional secretory domain is an important function that, as mentioned previously, allows serum detection of fragments that result from the osteoclastic specific bone resorption. A detailed description of the myeloid osteoclast lineage is illustrated in Fig. 2, which shows the molecular sequence of preosteoclast recruitment, proliferation, differentiation, cell fusion, and activation. The progression of cell surface markers and the active factors at each cell stage in osteoclastogenesis is described in the figure legend.

Osteoclast Inactivation/Apoptosis

Stopping bone resorption is necessary to avoid excessive resorption and bone loss. The mechanism of this action in the resorption cycle is currently not well understood. Both inactivation and apoptosis of the osteoclast have been proposed as mechanisms for limiting osteoclastic bone resorption. The prevalence of numerous diseases with increased bone resorption has led to great interest in understanding the mechanism of osteoclast inactivation and apoptosis as possible therapeutic targets. Apoptosis of osteoclasts has been linked to the Fas apoptotic signaling pathway.^[30] Bisphosphonates have been implicated in apoptosis of the giant cell tumors in bone that are similar osteoclasts and have been used in treating diseases that show increased bone resorption.^[31]

SIGNALING IN OSTEOCLAST DIFFERENTIATION AND ACTIVATION

RANKL/RANK/OPG

RANKL, also known as TRANCE, ODF, and OPGL, is a membrane-bound protein that is a potent activator of osteoclast differentiation.^[32] RANKL is a protein in the TNF family that is synthesized by bone marrow stromal cells, activated T cells, fibroblasts, and osteoblasts.^[33] Expression of RANKL is increased in the presence of such factors as 1,25-dihydroxyvitamin D₃ (1,25-(OH)₂-D₃), Parathyroid Hormone (PTH), Prostaglandin E₂ (PGE₂), and glucocorticoids. RANKL, in the presence of another cytokine, macrophage colony stimulating factor (M-CSF), causes the differentiation

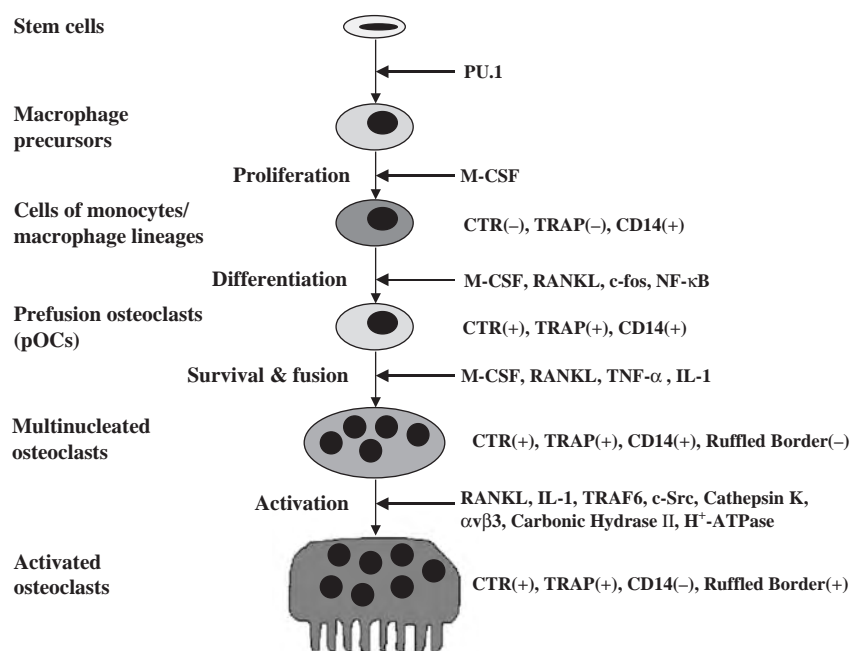


Fig. 2 The molecular mechanisms underlying osteoclast differentiation and function have been thoroughly studied by an osteopetrotic mouse model. It was found that PU.1, a transcription factor present in B-lymphoid cells and macrophages, which binds to a purine-rich sequence that contains a central core with the sequence 5'-GGAA-3' deficiency was lethal since the mice had no early macrophage precursors. Mice lacking M-CSF generated immature macrophages though the number was reduced. This suggests that M-CSF exerts its role at a later stage than does PU.1. In c-Fos, NF- κ B and RANKL mutants, normal macrophage development was evident, but with impaired osteoclast differentiation. Mutations in $\alpha_v\beta_3$ -integrin, TNF receptor associated factor (TRAF)-6 or c-Src mutants were normal in terms of osteoclast numbers, however osteoclasts in these models did not polarize and did not have ruffled membrane close to the bone side. In contrast, cathepsin K, carbonic anhydrase II, and H⁺-ATPase mouse mutants showed normal polarized osteoclasts but these cells failed to resorb bone. This was due to the fact that carbonic anhydrase II, and H⁺-ATPase are required to acidify the microenvironment between the bone and osteoclast, while cathepsin K is necessary for degradation of the organic matrix of the bone. Source: Adapted from Endocrine Society (see Ref. [67]); American Association for the Advancement of Science (see Ref. [63]) and From Nature Publishing Group (see Ref. [68]).

of macrophages/monocytes, and dendritic cells to osteoclasts.^[34,35] RANKL can be cleaved by TNF α -cleavage enzyme (TACE) to generate a soluble form of the protein. RANKL appears to induce osteoclast differentiation through the activation of the NF- κ B and c-Jun signaling pathways in osteoclast precursors.^[36]

RANKL expressed on stromal cells leads to osteoclast differentiation through RANK, a TNF receptor family member. RANK is a transmembrane receptor expressed primarily in cells of the monocyte/macrophage lineage, including B and T cells, dendritic cells, and osteoclast precursors.^[37] RANK is also expressed by differentiated osteoclasts,^[38,39] where RANKL binding is capable of activating the osteoclast.

The third player in the RANKL intercellular signaling pathway is OPG, also called osteoclastogenesis inhibitory factor (OCIF). OPG is a soluble, secreted protein that binds RANKL^[32,40] acting as a decoy receptor. By binding RANKL, OPG blocks RANKL-RANK

binding and the signaling that leads to osteoclastogenesis. By preventing osteoclast differentiation, OPG has potent bone protective effects. In several bone resorptive pathologies, an increase in the RANKL/OPG ratio has highlighted the importance of RANK/RANKL/OPG triad in regulating bone resorption.^[41,42] The protective effects of OPG have made it an attractive target for treating bone resorptive diseases, such as arthritis and metastatic cancers.

In addition to OPG, other aspects of the RANKL/RANK/OPG signaling system are being investigated as therapeutic targets. By using synthetic analogs of OPG and antagonists, or inhibitors of RANKL signaling, investigators are searching for ways to decrease the RANKL/OPG ratio. Another avenue being investigated is the ability to block RANK signaling to inhibit osteoclastogenesis.^[43] The discovery of the RANKL/RANK/OPG signaling triad has led to a better understanding of the mechanisms involved in bone remodeling.

Macrophage Colony Stimulating Factor

M-CSF is a growth factor/cytokine that is essential for osteoclast differentiation and survival. Kodama et al.^[44] demonstrated the necessary role of M-CSF in osteoclast differentiation. It is now known that M-CSF also has a role in promoting survival of both osteoclasts and osteoclast precursors.^[45] Osteoblasts and other stromal cells produce both membrane-bound and secreted forms of M-CSF, although only a small portion is membrane bound. M-CSF works to promote survival and differentiation of osteoclast and osteoclast precursors through the c-Fms receptor, a member of the tyrosine kinase growth factor receptor family. Signaling through the c-Fms receptor activates several pathways that promote osteoclast survival, including the mitogen-activated protein kinase (MAPK) cascade that turns on AP-1 and I κ B-kinase (IKK), leading to nuclear translocation of NF- κ B.

Interleukin-1

Interleukin-1, prior to the discovery of RANKL, was thought to be the main cytokine mediator involved in bone resorption and is still recognized today as an important signal in bone resorption. Interleukin-1 signaling is the result of three proteins of the same gene family: Interleukin-1 α (IL-1 α), Interleukin-1 β (IL-1 β), and Interleukin-1 Receptor agonist (IL-1Ra). IL-1 α and IL-1 β have largely the same effects in bone, but they are not the same in all regards. IL-1 α is more frequently found intracellularly and is released in significant amounts following cell death. Accordingly, serum concentrations of IL-1 α are very low. IL-1 β is synthesized, like IL-1 α , as a precursor that needs enzymatic cleavage for activation. Removal of a small peptide fragment by IL-1 β cleavage enzyme (ICE) from the propeptide generates the functional IL-1 β that is secreted by the cell. IL-1Ra is similar in structure to IL-1 β , but is not capable of eliciting a response upon binding to the IL-1 receptors. There are two IL-1 receptors: IL-1 receptor type I (IL-1RI) and IL-1 receptor type II (IL-1RII). IL-1RI is capable of binding and signaling, whereas IL-1RII is capable of binding but unable to transduce intracellular signaling.^[46] Since the two receptors are concomitantly expressed, there is competition for the binding of IL-1, however only IL-1RI leads to signal transduction. IL-1RI activation is accompanied by dimerization of IL-1RI with the interleukin-1 receptor accessory protein (IL-1RAcP) that is necessary for signaling. Structural differences in IL-1 β and IL-1Ra prevent IL-1RAcP from forming the heterodimer necessary for signaling.

IL-1 acts to stimulate osteoclastogenesis indirectly by affecting stromal cell expression of RANKL. Wei

et al.^[47] have shown that IL-1 is an intermediate in TNF α -induced RANKL expression. TNF α increases both IL-1 production, as well as the expression of its functional receptor, IL-1RI in stromal cells.^[47] The presence of IL-1Ra, or the use of IL-1RI deficient mice, had the effect of decreasing TNF α -induced RANKL expression. IL-1 stimulation increased osteoclastogenesis in macrophages exposed to RANKL and M-CSF.^[48]

IL-1 also has direct effects on osteoclasts and their precursors. Through IL-1RI, IL-1 is capable of activating the transcription factors c-Jun and NF- κ B. Activation of NF- κ B seems to be particularly important as RANKL–RANK signaling also works to activate this transcription factor. To examine IL-1 signaling, Li et al.^[49] used the presence of IL-1 receptor associated kinase-M (IRAK-M), a dominant negative IL-1RI associated signaling protein expressed in monocyte-derived cells that acts to decrease signaling. In the absence of IRAK-M, marked increases in osteoclastogenesis, osteoclast survival, as well as activation of NF- κ B, were observed. IL-1 can also affect osteoclastic bone degradative machinery, such as the expression of cathepsin K.^[50] The multiple effects resulting from both direct and indirect signaling of IL-1 make it a potent osteoclastogenic cytokine.

Tumor Necrosis Factor- α

TNF α , like RANKL, is a member of the TNF superfamily and has pro-inflammatory and bone resorptive effects. TNF α is made primarily by activated macrophages as a membrane-bound protein though TACE, the same protease responsible for the cleavage of RANKL, can act to cleave and release soluble TNF α . There are two receptors for TNF α : TNF receptor I (TNF-R1) and TNF Receptor II (TNF-R2). Signaling through TNF-R1 leads to the activation of two transcription factors: c-Jun and NF- κ B. TNF α has been shown in high concentrations to promote limited osteoclastogenesis in the absence of RANKL.^[51] Zou et al.^[52] have shown a role for TNF α , not only as a stimulator of NF- κ B signaling, but also as an autocrine factor that stimulates NF- κ B signaling in osteoclasts. There is also evidence of TNF α stimulating osteoblasts to produce RANKL, promoting osteoclastogenesis in an indirect manner.^[47] Data on the role of TNF α in bone resorption point to both a direct action in conjunction with RANKL–RANK signaling in osteoclast precursors as well as an the ability to stimulate production of resorptive signals from stromal cells in bone.

Interleukin-6

IL-6 is a cytokine expressed largely by stromal cells in bone^[53] that works to promote osteoclast survival.

IL-6 promotes osteoclast differentiation through receptors present on osteoclast precursors, however, like many of the other bone resorptive cytokines, there is an indirect mechanism that works to promote interaction of stromal cells and osteoclast precursors. IL-6 binds to the IL-6 receptor, which is composed of two subunits, the gp130 and the IL-6R α . Many of the effects of IL-6 are achieved through increases in IL-1 production.^[54] The JAK/STAT pathway is a downstream effector of IL-6 signaling that results in the eventual activation of MAPK and the transcription factor AP-1. In osteoclasts this signaling leads to differentiation and survival, while in stromal cells, IL-6 signaling results in increased expression of stromal factors, namely RANKL, that induce osteoclastogenesis through interaction with myeloid cells.^[55] IL-6 family members LIF and IL-11 are also osteoclastogenic cytokines that act through receptors utilizing the gp130 subunit for signaling. In response to resorptive signaling, the IL-6 family members act to stimulate osteoclastogenesis.

TRAF Signal Transduction

RANKL and other osteogenic cytokines activate a variety of signaling pathways, but they show a convergence of signaling to a few key transcription factors. Tumor necrosis factor receptor associated factors (TRAF) bind to TNFR family members and the IL-1R, initiating signaling cascades following receptor activation. TRAF-2, 5, and 6 have been shown to play key roles in osteoclastic signals. IL-1R, TNFR, and RANK all signal through TRAFs. IL-1R and TNFR both bind to TRAFs through adaptor proteins, while

RANK has 3 TRAF binding domains. TRAF signaling through I- κ B kinase (IKK) and the MAPK activates NF- κ B and AP-1, respectively. The TRAF binding site on RANK adjacent to the membrane is specific for TRAF-6, while the other two sites can bind either TRAF-2 or 5. In addition to interacting with RANK, TRAF-2 interacts with adaptor proteins associated with the TNFR; TRAF-6 interacts with an IRAK adaptor protein on the IL-1R. TRAF-6 knockout mice display severe osteopetrosis that shows the crucial role of TRAF-6^[56] in osteoclastogenesis. TRAF-6 works to activate the soluble tyrosine receptor kinase Src, which in turn activates PI₃ kinase and its downstream signaling partners. Signaling from the three TRAFs described above, and through IL-6, facilitates activation of MAPK and IKK, leading to further downstream activation/translocation of the two major transcription factors in osteoclastogenesis, AP-1 and NF- κ B, respectively. Fig. 3 depicts the divergent receptors and their convergence on functional differentiation and activation of the mature osteoclast.

Recruitment

In normal bone remodeling, bone formation and degradation are balanced, and a crucial part of this balance is in the recruitment, differentiation, and survival of osteoclasts. This appears to be accomplished mainly through osteoblastic signals, thereby maintaining the balance. In cases of excess resorption there is greater recruitment and survival of osteoclasts due to signaling by proinflammatory cytokines involved in osteolysis.

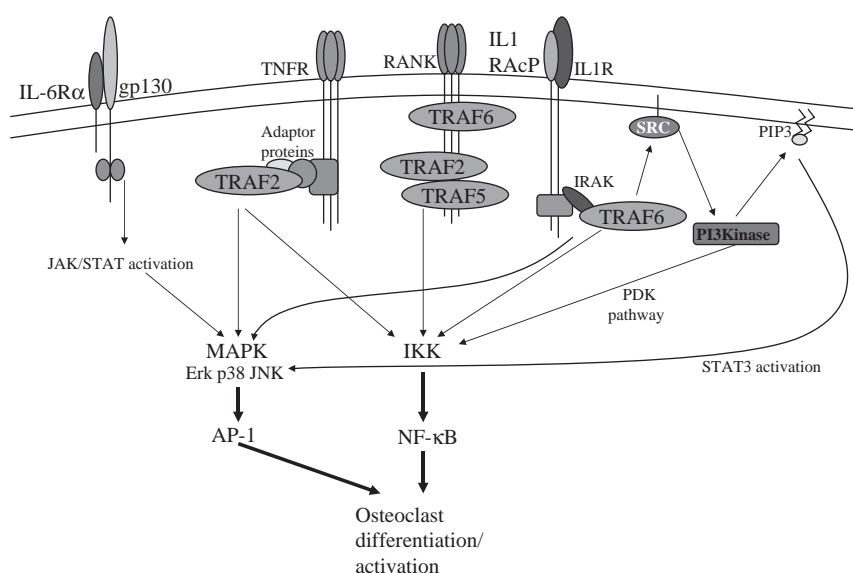


Fig. 3 Cytokines play a major role in regulating osteoclast differentiation and survival. Receptor mediated transduction of osteoclastogenic cytokine signaling leads to the activation of intracellular signaling cascades. TRAFs play an important role in TNFR family signaling and can activate numerous pathways to affect cellular signaling. IL-6R works through heterodimerization with gp130 to activate the MAPK cascade through JAK/STAT proteins. Signaling through the MAPK and IKK systems lead to activation of key transcription factors AP-1 and NF- κ B. Regulation of gene expression through these signaling mechanisms is responsible for the differentiation, activation and survival of the osteoclast.

BONE REMODELING

In healthy bone the process of formation and degradation is constant and balanced. The involvement of numerous cytokines from various sources in osteoclast differentiation and activation suggest that this is a highly coordinated process. Despite the activity of a myriad of cytokines and other signals, recent research has illustrated the major signaling mechanism involved in osteoclastogenesis, the RANKL/RANK/OPG molecular triad. The importance of this signaling system has been shown by the number of pathologies that exhibit some disturbance of this system. Osteoporosis is a condition that affects many people, especially women with diminished levels of estrogen. RANKL has been shown to be a major factor in periprosthetic osteolysis induced by particle wear debris surrounding joint replacements.^[57] A disturbance of the RANKL/OPG ratio has been shown in a mouse model of arthritis.^[58] Individuals with Paget's disease have been shown to express increased levels of RANKL, in part due to IL-6 signaling, as well as greater response of myeloid cells to RANKL signaling. These effects work to increase osteoclast activity in localized areas of resorption that characterize Paget's disease.^[59] One emerging area of research is the role of the RANKL and OPG in the calcification of arterial tissue in atherosclerosis.^[60] In addition, RANKL/RANK/OPG signaling has been implicated in both primary tumors of bone^[61] as well as metastases.^[62]

CONCLUSION

The regulation of osteoclast activity has become an area of significant interest as an increasing number of diseases have been characterized involving an imbalance between bone formation and resorption in remodeling. An understanding of the complex and often combinatorial effects of the cytokines that regulate osteoclast function is necessary to understand these pathologies. Elucidation of the RANKL/RANK/OPG system has provided a greater understanding of osteoclastogenesis as well as a myriad of new therapeutic targets.

ARTICLES OF FURTHER INTEREST

Bone Fracture Fixation; Bone-Implant Interface; Bone Remodeling; Tissue Engineering of Bone

REFERENCES

- Lakkakorpi, P.T.; Vaananen, H.K. Cytoskeletal changes in osteoclasts during the resorption cycle. *Microsc. Res. Tech.* **1996**, *33*, 171–181.
- Nesbitt, S.; Nesbit, A.; Helfrich, M.; Horton, M.; et al. Biochemical characterization of human osteoclast integrins. Osteoclasts express alpha v beta 3, alpha 2 beta 1, and alpha v beta 1 integrins. *J. Biol. Chem.* **1993**, *268*, 16737–16745.
- Ross, F.P.; Chappel, J.; Alvarez, J.I.; et al. Interactions between the bone matrix proteins osteopontin and bone sialoprotein and the osteoclast integrin alpha v beta 3 potentiate bone resorption. *J. Biol. Chem.* **1993**, *268*, 9901–9907.
- Helfrich, M.H.; Nesbitt, S.A.; Lakkakorpi, P.T.; et al. Beta 1 integrins and osteoclast function: involvement in collagen recognition and bone resorption. *Bone* **1996**, *19*, 317–328.
- Fisher, J.E.; Caulfield, M.P.; Sato, M.; et al. Inhibition of osteoclastic bone resorption in vivo by echistatin, an "arginyl-glycyl-aspartyl" (RGD)-containing protein. *Endocrinology* **1993**, *132*, 1411–1413.
- Horton, M.A.; Taylor, M.L.; Arnett, T.R.; Helfrich, M.H. Arg-Gly-Asp (RGD) peptides and the anti-vitronectin receptor antibody 23C6 inhibit dentine resorption and cell spreading by osteoclasts. *Exp. Cell. Res.* **1991**, *195*, 368–375.
- Felding-Habermann, B.; Cheresch, D.A. Vitronectin and its receptors. *Curr. Opin. Cell Biol.* **1993**, *5*, 864–868.
- Helfrich, M.H.; Nesbitt, S.A.; Dorey, E.L.; Horton, M.A. Rat osteoclasts adhere to a wide range of RGD (Arg-Gly-Asp) peptide-containing proteins, including the bone sialoproteins and fibronectin, via a beta 3 integrin. *J. Bone Miner. Res.* **1992**, *7*, 335–343.
- Lakkakorpi, P.T.; Horton, M.A.; Helfrich, M.H.; Karhukorpi, E.-K.; Väänänen, H.K. Vitronectin receptor has a role in bone resorption but does not mediate tight sealing zone attachment of osteoclasts to the bone surface. *J. Cell Biol.* **1991**, *115*, 1179–1186.
- Meredith, J.E., Jr; Winitz, S.; Lewis, J.M.; et al. The regulation of growth and intracellular signaling by integrins. *Endocr. Rev.* **1996**, *17*, 207–220.
- Duong, L.T.; Lakkakorpi, P.; Nakamura, I.; Rodan, G.A. Integrins and signaling in osteoclast function. *Matrix Biol.* **2000**, *19*, 97–105.
- Blair, H.C.; Teitelbaum, S.L.; Ghiselli, R.; Gluck, S. Osteoclastic bone resorption by a polarized vacuolar proton pump. *Science* **1989**, *245*, 855–857.
- Vaananen, H.K.; Karhukorpi, E.K.; Sundquist, K.; et al. Evidence for the presence of a proton pump of the vacuolar H(+)-ATPase type in the ruffled borders of osteoclasts. *J. Cell Biol.* **1990**, *111*, 1305–1311.
- Salo, J.; Metsikko, K.; Palokangas, H.; Lehenkari, P.; Vaananen, H. Bone-resorbing osteoclasts reveal a dynamic division of basal plasma membrane into two different domains. *J. Cell Sci.* **1996**, *109*, 301–307.
- Vaananen, H.K.; Zhao, H.; Mulari, M.; Halleen, J.M. The cell biology of osteoclast function. *J. Cell Sci.* **2000**, *113*, 377–381.
- Gay, C.V.; Mueller, W.J. Carbonic anhydrase and osteoclasts: localization by labeled inhibitor autoradiography. *Science* **1974**, *183*, 432–434.
- Palokangas, H.; Mulari, M.; Vaananen, H.K. Endocytic pathway from the basal plasma membrane to the

- ruffled border membrane in bone-resorbing osteoclasts. *J. Cell Sci.* **1997**, *110*, 1767–1780.
18. Goto, T.; Yamaza, T.; Tanaka, T. Cathepsins in the osteoclast. *J. Electron Microsc. (Tokyo)* **2003**, *52*, 551–558.
 19. Littlewood-Evans, A.; Kokubo, T.; Ishibashi, O.; et al. Localization of cathepsin K in human osteoclasts by in situ hybridization and immunohistochemistry. *Bone* **1997**, *20*, 81–86.
 20. Kafienah, W.; Bromme, D.; Buttle, D.J.; Croucher, L.J.; Hollander, A.P. Human cathepsin K cleaves native type I and II collagens at the N-terminal end of the triple helix. *Biochem. J.* **1998**, *331*, 727–732.
 21. Bromme, D.; Okamoto, K.; Wang, B.B.; Biroc, S. Human cathepsin O2, a matrix protein-degrading cysteine protease expressed in osteoclasts. Functional expression of human cathepsin O₂ in *Spodoptera frugiperda* and characterization of the enzyme. *J. Biol. Chem.* **1996**, *271*, 2126–2132.
 22. Bossard, M.J.; Tomaszek, T.A.; Thompson, S.K.; et al. Proteolytic activity of human osteoclast cathepsin K. Expression, purification, activation, and substrate identification. *J. Biol. Chem.* **1996**, *271*, 12517–12524.
 23. Ljusberg, J.; Wang, Y.; Lang, P.; et al. Proteolytic excision of a repressive loop domain in tartrate-resistant acid phosphatase by cathepsin K in osteoclasts. *J. Biol. Chem.* **2005**, *280* (31), 28370–28381.
 24. Saftig, P.; Hunziker, E.; Wehmeyer, O.; et al. Impaired osteoclastic bone resorption leads to osteopetrosis in cathepsin-K-deficient mice. *Proc. Natl Acad. Sci. USA.* **1998**, *95*, 13453–13458.
 25. Yamaza, T.; Tsuji, Y.; Goto, T.; et al. Comparison in localization between cystatin C and cathepsin K in osteoclasts and other cells in mouse tibia epiphysis by immunolight and immunoelectron microscopy. *Bone* **2001**, *29*, 42–53.
 26. Lerner, U.H.; Johansson, L.; Ranjso, M.; Rosenquist, J.B.; Reinholt, F.P.; Grubb, A. Cystatin C, and inhibitor of bone resorption produced by osteoblasts. *Acta Physiol. Scand.* **1997**, *161*, 81–92.
 27. Motyckova, G.; Fisher, D.E. Pycnodysostosis: role and regulation of cathepsin K in osteoclast function and human disease. *Curr. Mol. Med.* **2002**, *2*, 407–421.
 28. Delaisse, J.M.; Andersen, T.L.; Engsig, M.T.; Henriksen, K.; Troen, T.; Blavier, L. Matrix metalloproteinases (MMP) and cathepsin K contribute differently to osteoclastic activities. *Microsc. Res. Tech.* **2003**, *61*, 504–513.
 29. Teronen, O.; Heikkila, P.; Konttinen, Y.T.; et al. MMP inhibition and downregulation by bisphosphonates. *Ann. N.Y. Acad. Sci.* **1999**, *878*, 453–565.
 30. Wu, X.; McKenna, M.; Feng, X.; Nagy, T.R.; McDonald, J.M. Osteoclast apoptosis: the role of Fas in vivo and in vitro. *Endocrinology* **2003**, *144*, 5545–5555.
 31. Chang, S.S.; Suratwala, S.J.; Jung, K.M.; et al. Bisphosphonates may reduce recurrence in giant cell tumor by inducing apoptosis. *Clin. Orthop. Relat. Res.* **2004**, 103–109.
 32. Lacey, D.L.; Timms, E.; Tan, H.-L.; et al. Osteoprotegerin ligand is a cytokine that regulates osteoclast differentiation and activation. *Cell* **1998**, *93*, 165–176.
 33. Wang, R.; Zhang, L.Y.; Zhang, X.R.; et al. Regulation of activation-induced receptor activator of NF-kappaB ligand (RANKL) expression in T cells. *Eur. J. Immunol.* **2002**, *32*, 1090–1098.
 34. van't Hof, R.J.; Tuinenburg-Bol Raap, A.C.; Nijweide, P.J. Induction of osteoclast characteristics in cultured avian blood monocytes; modulation by osteoblasts and 1,25-(OH)₂ vitamin D₃. *Int. J. Exp. Pathol.* **1995**, *76*, 205–214.
 35. Rivollier, A.; Mazzorana, M.; Tebib, J.; et al. Immature dendritic cell transdifferentiation into osteoclasts: a novel pathway sustained by the rheumatoid arthritis microenvironment. *Blood* **2004**, *104*, 4029–4037.
 36. Darnay, B.G.; Haridas, V.; Ni, J.; Moore, P.A.; Aggarwal, B.B. Characterization of the intracellular domain of receptor activator of NF-kappaB (RANK). Interaction with tumor necrosis factor receptor-associated factors and activation of NF-kappab and c-Jun N-terminal kinase. *J. Biol. Chem.* **1998**, *273*, 20551–20555.
 37. Khosla, S. Minireview: the OPG/RANKL/RANK system. *Endocrinology* **2001**, *142*, 5050–5055.
 38. Hsu, H.; Lacey, D.L.; Dunstan, C.R.; et al. Tumor necrosis factor receptor family member RANK mediates osteoclast differentiation and activation induced by osteoprotegerin ligand. *Proc. Natl Acad. Sci. USA.* **1999**, *96*, 3540–3545.
 39. Myers, D.E.; Collier, F.M.; Minkin, C.; et al. Expression of functional RANK on mature rat and human osteoclasts. *FEBS Lett.* **1999**, *463*, 295–300.
 40. Yasuda, H.; Shima, N.; Nakagawa, N.; et al. Osteoclast differentiation factor is a ligand for osteoprotegerin/osteoclastogenesis-inhibitory factor and is identical to TRANCE/RANKL. *Proc. Natl Acad. Sci. USA.* **1998**, *95*, 3597–3602.
 41. Skoumal, M.; Kolarz, G.; Haberhauer, G.; Woloszczuk, W.; Hawa, G.; Klingler, A. Osteoprotegerin and the receptor activator of NF-kappa B ligand in the serum and synovial fluid. A comparison of patients with longstanding rheumatoid arthritis and osteoarthritis. *Rheumatol. Int.* **2005**, *25*, 373–378.
 42. Grimaud, E.; Soubigou, L.; Couillaud, S.; et al. Receptor activator of nuclear factor kappaB ligand (RANKL)/osteoprotegerin (OPG) ratio is increased in severe osteolysis. *Am. J. Pathol.* **2003**, *163*, 2021–2031.
 43. Goranov, S.E.; Goranova-Marinova, V.S. Bone lesions in multiple myeloma—the OPG/RANKL-ligand system. *Folia Med. (Plovdiv)* **2004**, *46*, 5–11.
 44. Kodama, H.; Yamasaki, A.; Nose, M.; et al. Congenital osteoclast deficiency in osteopetrotic (op/op) mice is cured by injections of macrophage colony-stimulating factor. *J. Exp. Med.* **1991**, *173*, 269–272.
 45. Tanaka, S.; Takahashi, N.; Udagawa, N.; et al. Macrophage colony-stimulating factor is indispensable for both proliferation and differentiation of osteoclast progenitors. *J. Clin. Invest.* **1993**, *91*, 257–263.
 46. Dinarello, C.A. Interleukin-1. *Cytokine Growth Factor Rev.* **1997**, *8*, 253–265.
 47. Wei, S.; Kitaura, H.; Zhou, P.; Ross, F.P.; Teitelbaum, S.L. IL-1 mediates TNF-induced osteoclastogenesis. *J. Clin. Invest.* **2005**, *115*, 282–290.



48. Ma, T.; Miyanishi, K.; Suen, A.; Epstein, N.J.; Tomita, T.; Smith, R.L. Human interleukin-1-induced murine osteoclastogenesis is dependent on RANKL, but independent of TNF- α . *Cytokine* **2004**, *26*, 138–144.
49. Li, H.; Cuartas, E.; Cui, W.; et al. IL-1 receptor-associated kinase M is a central regulator of osteoclast differentiation and activation. *J. Exp. Med.* **2005**, *201*, 1169–1177.
50. Kamolmatyakul, S.; Chen, W.; Yang, S.; et al. IL-1 α stimulates cathepsin K expression in osteoclasts via the tyrosine kinase-NF- κ B pathway. *J. Dent. Res.* **2004**, *83*, 791–796.
51. Kobayashi, K.; Takahashi, N.; Jimi, E.; et al. Tumor necrosis factor α stimulates osteoclast differentiation by a mechanism independent of the ODF/RANKL–RANK interaction. *J. Exp. Med.* **2000**, *191*, 275–286.
52. Zou, W.; Hakim, I.; Tschoep, K.; Endres, S.; Bar-Shavit, Z. Tumor necrosis factor- α mediates RANK ligand stimulation of osteoclast differentiation by an autocrine mechanism. *J. Cell Biochem.* **2001**, *83*, 70–83.
53. Holt, I.; Davie, M.W.; Marshall, M.J. Osteoclasts are not the major source of interleukin-6 in mouse parietal bones. *Bone* **1996**, *18*, 221–226.
54. Kurihara, N.; Bertolini, D.; Suda, T.; Akiyama, Y.; Roodman, G.D. IL-6 stimulates osteoclast-like multinucleated cell formation in long term human marrow cultures by inducing IL-1 release. *J. Immunol.* **1990**, *144*, 4226–4230.
55. O'Brien, C.A.; Lin, S.C.; Bellido, T.; Manolagas, S.C. Expression levels of gp130 in bone marrow stromal cells determine the magnitude of osteoclastogenic signals generated by IL-6-type cytokines. *J. Cell Biochem.* **2000**, *79*, 532–541.
56. Naito, A.; Azuma, S.; Tanaka, S.; et al. Severe osteopetrosis, defective interleukin-1 signalling and lymph node organogenesis in TRAF6-deficient mice. *Genes Cells* **1999**, *4*, 353–362.
57. Clohisy, J.C.; Frazier, E.; Hirayama, T.; Abu-Amer, Y. RANKL is an essential cytokine mediator of polymethylmethacrylate particle-induced osteoclastogenesis. *J. Orthop. Res.* **2003**, *21*, 202–212.
58. Lubberts, E.; Schwarzenberger, P.; Huang, W.; et al. IL-17 promotes bone erosion in murine collagen-induced arthritis through loss of the receptor activator of NF- κ B ligand/osteoprotegerin balance. *J. Immunol.* **2003**, *170*, 2655–2662.
59. Mena, C.; Reddy, S.V.; Kurihara, N.; et al. Enhanced RANK ligand expression and responsiveness of bone marrow cells in Paget's disease of bone. *J. Clin. Invest.* **2000**, *105*, 1833–1838.
60. Collin-Osdoby, P. Regulation of vascular calcification by osteoclast regulatory factors RANKL and osteoprotegerin. *Circ. Res.* **2004**, *95*, 1046–1057.
61. Good, C.R.; O'Keefe, R.J.; Puzas, J.E.; Schwarz, E.M.; Rosier, R.N. Immunohistochemical study of receptor activator of nuclear factor κ -B ligand (RANK-L) in human osteolytic bone tumors. *J. Surg. Oncol.* **2002**, *79*, 174–179.
62. Kitazawa, S.; Kitazawa, R. RANK ligand is a prerequisite for cancer-associated osteolytic lesions. *J. Pathol.* **2002**, *198*, 228–236.
63. Teitelbaum, S.L. Bone resorption by osteoclasts. *Science* **2000**, *289*, 1504–1508.
64. Quinn, J.M.; Neale, S.; Fujikawa, Y.; McGee, J.O.; Athanasou, N.A. Human osteoclast formation from blood monocytes, peritoneal macrophages, and bone marrow cells. *Calcif. Tissue Int.* **1998**, *62*, 527–531.
65. Reddy, S.V.; Roodman, G.D. Control of osteoclast differentiation. *Crit. Rev. Eukaryot. Gene Expr.* **1998**, *8*, 1–17.
66. Teitelbaum, S.L. Bone resorption by osteoclasts. *Science* **2000**, *289*, 1504–1508.
67. Takahashi, N.; Akatsu, T.; Udagawa, N.; et al. Osteoblastic cells are involved in osteoclast formation. *Endocrinology* **1988**, *123*, 2600–2602.
68. Tondravi, M.M.; McKercher, S.R.; Anderson, K.; et al. Osteopetrosis in mice lacking haematopoietic transcription factor PU.1. *Nature* **1997**, *386*, 81–84.

Osteogenic Progenitor Cells of Bone

Richard O.C. Oreffo

University Orthopaedics, University of Southampton, Southampton, United Kingdom

INTRODUCTION

The osteoblast, the cell that is responsible for bone formation, is derived from a pluripotent marrow stromal fibroblastic stem cell, or mesenchymal stem cell, that can give rise to cells of the adipogenic, reticular, osteoblastic, myoblastic, and fibroblastic lineages. These stem cells generate progenitors committed to one or more cell lines with an apparent degree of plasticity or interconversion. However, the isolation and characterization of such stem cells has been hampered by the lack of specifically reactive antibodies and the very low incidence of these cells. A wealth of evidence has accumulated detailing the factors that regulate and modulate osteoprogenitor differentiation into the mature osteoblast phenotype using a variety of in-vitro models with different species, which has also served to provide conflicting observations and interpretations as to the fate, function, and physiology of the mesenchymal stem cell and osteogenic progenitor populations. Indeed, to date, we still know relatively little about the phenotypic characteristics, whether from a morphological or biochemical standpoint. In addition, direct in-vivo confirmation of the lineage potential and plasticity or interconversion potential that exists of mesenchymal stem cells and osteogenic progenitor cells remains unclear.

With increasing age, skeletal remodeling is altered with a resultant net bone loss as a consequence of increased trabecular and cortical bone resorption and decreased osteoblast number and bone formation. Bone formation strategies, although attractive, have yet to yield functional and mechanically competent tissues and clinical imperatives to augment and facilitate tissue lost as a consequence of trauma or degeneration, have brought mesenchymal and osteoprogenitor cell biology to center stage. This is due largely to the observation that bone has a tremendous capacity to regenerate from cells with stem cell characteristics. Thus, current approaches have centered on the isolation and expansion of mesenchymal stem cells and osteogenic progenitor cells and the modulation of their phenotype for skeletal repair using cell, gene, and tissue engineering.

This brief article outlines the concepts, properties, and current understanding in mesenchymal stem cell and osteogenic progenitor cell biology and their

potential for clinical application in tissue reconstruction, which could improve the quality of life for many as a result of advances in cell, gene, and tissue engineering.

MESENCHYMAL STEM CELL LINEAGES

Bone marrow serves as a reservoir for a variety of stem cells including the nonadherent, circulating hemopoietic stem cells and, in close contact and supporting the hemopoietic compartment, marrow stromal cells. The marrow stromal stem cell or mesenchymal stem cell exhibits multilineage differentiation capacity and has been shown to generate cartilage, bone, myelo-supportive stroma, adipocytes and fibrous connective tissue (Fig. 1).^[1–3] The capacity of bone to regenerate has long been recognized, and yet, to date, there is only limited evidence for the existence of a pluripotent mesenchymal or marrow stem cell. This is, in large part, due to their low incidence, indeterminate morphology, and, despite nearly three decades of research, undefined biochemical phenotype.

A number of criteria are required before a cell can be termed a stem cell, namely, self renewal, pluripotentiality, and the capacity for cell division maintained throughout life—terms developed from the hemopoietic system. Thus, a stem cell can be defined as a cell that maintains its own numbers (self renewal) in spite of physiological or artificial removal of cells from the population^[4] and that can differentiate into all cell types (pluripotent) or more than one differentiated cell type (multipotent).^[5] In contrast, it is important to note that (osteogenic) progenitor cells may be multipotential but cannot self-renew and thus are intermediate between a stem cell and differentiated progeny (i.e., osteoblast). Apart from a lack of morphological and defined biochemical phenotype, a certain degree of confusion has been generated by the interchangeable use of names for the pluripotent stromal cell that can give rise to osteogenic, adipogenic, fibroblastic, and reticular lineages. These have included osteogenic stem cells, stromal fibroblastic cells,^[6] stromal stem cells coined by Owen and Friedenstein,^[7] mesenchymal stem cells initiated by Caplan and coworkers,^[8] and, more recently, skeletal stem cells.^[9] Currently, mesenchymal stem cells



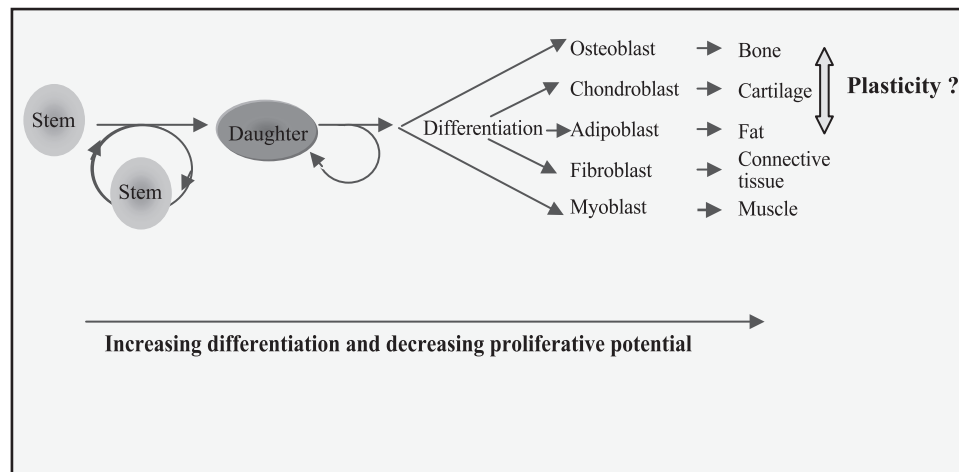


Fig. 1 Differentiation of mesenchymal stem cells from self-renewing stem cells and daughter cells to generate all mesenchymal cell lineages.

(undifferentiated pluripotent cells of the mesenchyme) appear to be the favored term adopted by many workers within the field.

Evidence for an osteoblast stem cell was provided in the early 1960s in seminal studies by Alexander Friedenstein and coworkers. Initial studies showed that bone marrow transplanted under the renal capsule of mice or within an enclosed environment (diffusion chamber) generated osteogenic tissue indicating that marrow contained osteogenic precursors.^[3,10] The mesenchymal stem cell gives rise to a hierarchy of bone cell populations artificially divided into a number of developmental stages including mesenchymal stem cell (MSC), determined osteoprogenitor cell (DOPC), preosteoblast, osteoblast, and, ultimately, osteocyte. However, it is important to note that clear and distinct boundaries between the populations do not exist and these cells exist, in essence, as a developmental continuum.

AGING AND MESENCHYMAL STEM CELLS

Aging is associated with decreased fracture repair and reduced skeletal bone mass as a consequence of a net reduction in bone formation. This can result in osteoporosis with devastating socioeconomic consequences. As indicated, bone formation depends on mesenchymal stem cells present within bone marrow that give rise to colony-forming units—fibroblastic (CFU—F), which differentiate into the osteogenic, adipogenic, fibroblastic, and reticular cell lineages.^[3,10–12] Studies on human osteoprogenitor number and age have been limited and contradictory. Nishida et al.^[13] found that alkaline phosphatase-positive colony-forming

units—fibroblastic (AP+CFU—F) were significantly reduced between 10 and 20 years of age and then gradually after the age of 20. Similarly, D'Ippolito et al.^[14] found a significant decrease with age in AP-positive CFU-F in bone progenitors isolated from human vertebrae. Other workers have recorded decreases in CFU-F colonies or AP+CFU—F in human,^[15] rat,^[16] and mouse^[17] marrow. However, Oreffo et al. in a large study of 99 patients who were osteoarthritic, osteoporotic, or showed no evidence of metabolic bone disease, found no differences in CFU—F or AP+CFU—F with age, disease state, or gender, although a significant decrease in CFU—F colony size was observed.^[18,19] Similarly, Stenderup et al. found the number and proliferative capacity of osteogenic stem cells was maintained during aging and in patients with osteoporosis.^[20] The discrepancies in these various reports may be attributable to different sample population sizes, populations selected, and laboratory protocols used. Nevertheless, the bone loss associated with aging may reflect altered proliferative capacity of progenitor cells or altered responsiveness of CFU—F to systemic or locally released growth factors^[21,22] leading to alteration in subsequent differentiation.

PHENOTYPIC CHARACTERISTICS

To date, there is a paucity of developmental stage-specific markers for mesenchymal stem cells and osteogenic progenitor cells. It is generally accepted that osteoblast is derived from a pluripotent mesenchymal stem cell.^[3,10,11,23,24] A few antibodies do exist and the isolation and expansion of these cells, while retaining

their differentiation ability, has been performed using the following monoclonal antibodies SB-10,^[25] STRO-1,^[26] SH-2,^[27] and HOP-26.^[28] Gronthos and coworkers found the heterogeneity of the stromal cell population could be reduced by isolation using the monoclonal antibody STRO-1, which recognizes a trypsin-resistant cell surface antigen present on a subpopulation of bone marrow cells.^[29] The cells expressing STRO-1 have been found to include predominantly the adherent, high growth potential, colony-forming units—fibroblastic (CFU—F).^[29] However, the epitope for STRO-1 remains to be characterized although the antibody does not bind to myeloid cells, megakaryocytes, or macrophages.^[29] The STRO-1 selected populations have been shown to give rise to osteogenic, fibroblastic, adipogenic, and smooth muscle cells, as well as cells of the osteoblastic lineage. The cells of the osteoblastic lineage, under osteogenic conditions, become alkaline phosphatase-positive, respond to $1,25(\text{OH})_2\text{D}_3$, form mineralized tissue in vitro, and, in vivo, form bone.^[29–31] Beresford and coworkers have shown, using cell sorting, that there is an inverse association between the expression of STRO-1 and alkaline phosphatase.^[30] Pre-osteoblasts or progenitor populations were observed to be STRO-1-positive and alkaline phosphatase-negative, whereas osteoblasts were observed to be STRO-1-negative and alkaline phosphatase-positive.^[30] Epitopes for a number of monoclonal antibodies shown to be reactive with early osteoprogenitor populations have been identified,^[31]—thus SB-10, which has been shown to react with osteoprogenitors at the early stages of differentiation, has been found to react with activated leukocyte-cell adhesion molecule 99 (ALCAM)^[32] and SH-2 with CD105 (Endoglin).^[33] The epitope for HOP-26, derived using human bone-marrow fibroblast cultures, has been shown to be the cell-surface and lysosomal enzyme CD63 or melanoma-associated antigen, a member of the tetraspan glycoproteins.^[34]

Studies from Verfaillie's group have shown that pluripotent mesenchymal stem cells, termed multipotent adult progenitor cells, could be derived from adult marrow by cell sorting, and the phenotype of the human mesenchymal stem cells after extended culture was observed to be CD34, CD44, CD45, c-Kit, and major histocompatibility complex (MHC) class I and II-negative.^[11] These cells have been shown to differentiate not only into mesenchymal cells, but also into cells with visceral mesoderm neuroectoderm and endoderm characteristics in vitro.^[11] Critically, it appeared multipotent adult progenitor cells engrafted in vivo when transplanted into nondamaged recipients. To date, however, issues of cell fusion, nuclear reprogramming and derivation of unique phenotypes in extended culture remain to be resolved in these and

other studies and the phenotype of the human mesenchymal stem cell in vivo is unknown.

PLASTICITY AND INTERCONVERSION POTENTIALS

The multipotentiality of mesenchymal stem cells to form bone cartilage, fat, tendon, and muscle is widely accepted as a key criterion in the definition of these cells. Studies from a number of groups have shown that bone marrow cells can be plated onto tissue culture plastic, and the initial adherent bone marrow-derived stromal colonies are derived from a single mesenchymal stem cell.^[10] These colonies are multipotent and can be induced to form bone, cartilage, and fat by simple manipulation of culture conditions.^[24,35,36] As detailed above, committed progenitors (i.e., restricted in lineage development such as bone under defined conditions) can be identified by functional in-vitro CFU—F assays.^[37] Thus, these stromal colonies retain their multipotentiality in vitro after expansion. Culture in the presence of dexamethasone, methyl isobutylxanthine, insulin, and indomethacin has been found to favor adipogenesis of human bone marrow cells, whereas culture in the absence or presence of serum with TGF β favors chondrogenesis.^[24,35] In contrast, it is well-established that marrow cells cultured in serum with dexamethasone and ascorbic acid favor osteoprogenitor differentiation with enhanced alkaline phosphatase expression, matrix production, nodule formation, and deposition of calcium, confirming the presence of osteoprogenitor cells.^[12,35] With the understanding of defined conditions required to modulate cell phenotype, the potential for cells to switch or differentiate between different phenotypes after considerable differentiation (for example, along the osteogenic lineage to give a mature osteoblast population) has led to the concept of plasticity of phenotype. It has long been known that an association between an increase in marrow adipose tissue and osteopaenia with increasing age and in a variety of experimental and pathological conditions such as disuse osteoporosis and glucocorticoid-induced osteoporosis exists,^[38] suggesting plasticity or interconversion potential between the lineages and that the adipocytic and osteogenic cells share a common lineage. Park et al.^[35] have shown the ability to isolate clonal adipogenic cells and to redifferentiate the cells into osteogenic and adipogenic lineages in vitro, while Nuttall and coworkers^[39] have shown the potential of human bone cells to transdifferentiate along the adipogenic lineage. All of these manipulations in-vitro remain to be conclusively demonstrated in vivo and the extent and physiological relevance of cell plasticity remains unclear although undeniably

clinically attractive from a tissue-engineering and -repair perspective.

DIFFERENTIATION AND PROLIFERATION—GROWTH FACTORS

A number of growth factors have been shown to affect mesenchymal stem cell and osteoprogenitor cell activity, although it is important to note culture conditions (e.g., serum-free vs. serum) and species studies (human marrow vs. mouse marrow or rat calvarial-derived populations) play a significant part in the response/interpretation observed. It is beyond the scope of this short article to cover all the regulatory factors implicated in control of mesenchymal stem cell and osteoprogenitor differentiation, but several key factors emerge in the differentiation and proliferation of mesenchymal stem cells and osteoprogenitor cells. In general, the information that follows will concentrate on data from human systems.

Transforming growth factor beta (TGF β) and other members of the TGF β superfamily, including the bone morphogenetic proteins, play a significant role in the induction and differentiation and modulation of osteoprogenitor cells.^[40,41] TGF β has complex biphasic actions on progenitors and on their differentiation. This matrix protein has been shown to act on committed or determined osteogenic precursors to stimulate proliferation and chemotaxis to generate populations of committed osteoblast cells.^[42] Locklin and coworkers showed TGF β in the absence of serum-promoted cell growth and alkaline phosphatase activity while in the presence of serum-matrix proteoglycan synthesis, and collagen synthesis was increased in human bone-marrow stromal fibroblasts.^[43] In addition, in the same model system, TGF β inhibited dexamethasone-induced adipogenesis. More recently, Walsh and coworkers found TGF β 1 limited the expansion of the osteoprogenitor fraction of human bone-marrow cells with significantly reduced alkaline phosphatase expression.^[44]

The BMPs, which like the TGF β s are sequestered within bone matrix, play a pivotal role in the initiation of mesenchymal stem cells or inducible osteogenic precursors into the osteoblast lineage (Fig. 2).

BMP-2, BMP-4, and BMP-7 have been shown to induce the mouse pluripotent fibroblast cell line C3H10T1/2 and mouse 3T3 cells to differentiate into osteoblasts, chondroblasts, and adipocytes.^[45] Similarly, Rickard et al.^[46] have shown that BMP-2 can induce osteoblast differentiation in rat bone-marrow cell cultures. In addition, differential roles for BMP receptor type IB and IA in the differentiation and specification of mesenchymal progenitor cells to either osteoblasts or adipocytes have been demonstrated by Chen and coworkers.^[41]

Gronthos and Simmons examined the growth factor requirements of selected (STRO-1) human bone-marrow stromal precursors in serum-deprived conditions in vitro and found from 25 purified recombinant growth factors examined that platelet-derived growth factor (PDGF) and epidermal growth factor (EGF) were able to initiate and support clonogenic growth of CFU—F while IL-4 and INF α inhibited colony formation.^[47] In a later study, Kuznetsov and colleagues found the initial proliferation of human and mouse marrow stromal fibroblast colony formation required at least four growth factors: PDGF, bFGF, TGF β , and EGF with differing requirements between mouse and human marrow stromal fibroblasts.^[48] More recently, Oreffo and coworkers showed INF α inhibited human osteoprogenitor cell proliferation, CFU—F formation, alkaline phosphatase expression, and HOP-26 expression, as well as modulating BMP-2 expression.^[49]

A number of studies indicate basic FGF and the glucocorticoids are potent mitogens for marrow stromal fibroblasts. FGF-2 was found to increase colony size and colony number of human stromal cells markedly although colony formation was unaffected.^[50] Furthermore, it appears FGF-2 acts on early osteogenic progenitors, possibly maintaining cells in an immature state, as in combination with the synthetic glucocorticoid dexamethasone, and further cell

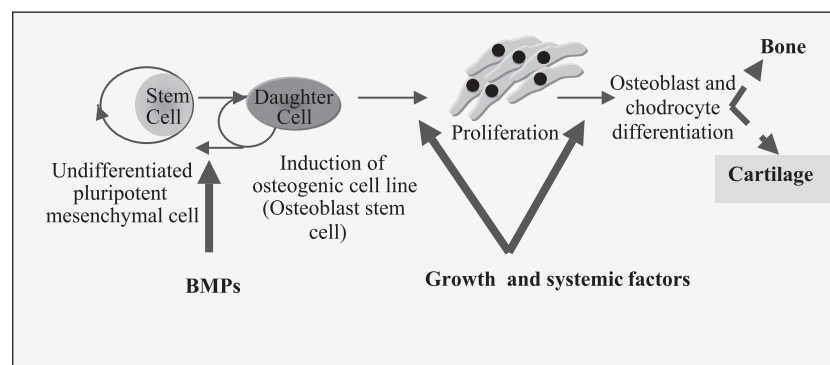


Fig. 2 Differentiation of mesenchymal stem cells and daughter cells to generate bone and cartilage and site of interaction of crucial factors such as BMPs in initiating osteogenesis. The particular stages of differentiation and sites of factor interaction are imprecisely known.

maturation was observed.^[51] In vitro, dexamethasone has been shown in rat, mouse, rabbit, and human marrow stromal cultures to stimulate osteoprogenitor cell differentiation,^[43,52–57] although the known effects of glucocorticoid-induced osteoporosis and impaired fracture repair remain confounding issues.

Other factors of note include 1) PGE₂, which may act to promote osteogenesis by the recruitment of osteoprogenitor cells from the nonadherent mesenchymal precursor cells in bone marrow acting possibly via EP2 receptors,^[58–61] 2) Pleiotrophin (PTN), an extracellular matrix-associated protein, which has been found to induce human osteoprogenitor chemotaxis, proliferation, differentiation, and bone formation,^[62,63] and 3) osteoclastotropic/bone remodeling factors Il-1, Il-6, GM-CSF, and M-CSF, which have all been implicated in the modulation of osteoprogenitor proliferation and differentiation.^[64–66]

Ultimately, delineation of the growth factors and hormones required for human mesenchymal stem cells and osteoprogenitors will be needed to generate protocols for the rapid expansion, ex-vivo, of these cells (using autologous serum) while retaining their differentiation potential for tissue-engineering/regeneration applications.

STEM CELLS, GENES, AND TISSUE ENGINEERING—BONE REGENERATION AND RESTORING AGING BONES

Developments in cell and gene engineering indicate it may be possible to use osteoprogenitor cells for therapeutic use in skeletal and hemopoietic diseases. Tissue loss as a result of injury or disease, in an increasing, aging population, provides reduced quality of life for many at significant socioeconomic cost.^[67] Healthcare costs in treating patients with bone fractures, in the U.K. alone, have been estimated at £1.7 billion.^[68] With an increasing, aging population, these health costs are set to rise. To date, artificial prostheses, which are subjected to wear due to lack of integration, ultimately fail and while joint replacement surgery in rheumatoid arthritis, osteoarthritis, and osteoporosis has significantly improved the quality of life for many, biomaterial failure and incompatibility and problems with detachment of the stem coating and/or fracture of the cement–stem or cement–bone interface^[69] have led to intense research activity to exploit the potential of mesenchymal stem cells and osteoprogenitors in bone repair and regeneration.^[9,10,70] It should be noted that at present, regimes that encourage bone formation, which hold the promise of significantly increasing bone density, have yet to reach the clinic.

Each year in the U.K. there are some 150,000 fractures (wrist, vertebral, and hip) due to osteoporosis. In

particular, hip fracture is associated with high morbidity and mortality with fewer than half the patients returning home after surgery. Moreover, 30–50% of these hip operations will require subsequent revision surgery and, in a significant proportion, bone augmentation will be necessary. The lack of techniques and approaches in reconstructive surgery emphasize the number of clinical applications that would benefit from the use of mesenchymal stem cells and osteoprogenitors to augment bone formation.^[9,70] The field of tissue engineering seeks to resolve these issues through a combination of stem or progenitor cells with appropriate growth factors and tailored scaffolds. Thus, tissue engineering has been defined as the application of scientific principles to the design, construction, modification, and growth of living tissues.^[71] With regard to a source of progenitor cells, it has long been known that bone has a vast capacity for regeneration from cells with stem cell characteristics. Kadiyala and coworkers have shown that culture-expanded bone marrow cells can heal a segmental bone defect following reimplantation.^[72] A number of studies indicate mesenchymal stem cells and osteoprogenitor populations from a variety of species, including human mesenchymal stem cells, can give rise to osteogenic tissue within diffusion chambers.^[73–75] As detailed previously, human bone marrow osteoprogenitors can be isolated and enriched using selective markers, such as STRO-1, from a CD34+ fraction^[76,77] and readily expanded, indicating their potential for marrow repopulation.^[10,28,78,79] Furthermore, the lack of immunogenicity of mesenchymal stem cells^[80] has opened up the potential of these cells in cartilage and bone repair. Although true engraftment of human mesenchymal stem cells, long-term biological effects on the stem cells at the implant site, and issues of cell plasticity remain unknown, preliminary studies from Horwitz and colleagues^[81] using human bone marrow-derived osteoprogenitors transplanted into children with osteogenesis imperfecta suggest some therapeutic effect.^[82] The authors reported that donor-derived mesenchymal stem cells (albeit 2%) were capable of homing to the bone marrow and differentiating into osteoblasts. In follow-up studies an increase in bone mineral content was observed compared to age-matched controls although the precise contribution of the donor cells to the clinical improvements recorded remains to be determined.^[83] Other workers, Connolly^[84] and recently Quarto and coworkers^[85] have indicated the efficacy of autologous bone marrow stromal cells in the treatment of large bone defects. Although mesenchymal stem cell/osteoprogenitor delivery, whether by injection or with a matrix, is attractive and carries a reduced risk of morbidity, for large skeletal defects, these cells will need an appropriate, designed or smart vehicle/scaffold/matrix tailored to the shape and size of

the defect that will provide mechanical competence. To aid this process, advances in scaffold technology and gene delivery offer the possibility of genetic modification of isolated and expanded cells in constructs to produce populations of progenitor cells over-expressing selected signaling molecules for bone regeneration.^[9,83,86-88] Lieberman and coworkers have shown regional cell and gene therapy using BMP-2-producing bone marrow cells on the repair of segmental bone defects in rats.^[89] Similarly, Breitbart and colleagues have cultured periosteal cells retrovirally transduced with BMP-7 in a polyglycolic-acid scaffold in a critical-sized calvarial defect model in rabbits.^[90] Several studies indicate human bone marrow stromal cells expressing BMP-2 by adenoviral infection may prove efficacious in bone regeneration (Fig. 3).^[31,91-93]

Musgrave and coworkers^[94] reported on the use of direct adenoviral-mediated gene therapy to deliver active BMP-2 and produce bone in skeletal muscle, thus circumventing problems associated with delivery of BMP-2 and the requirements for cell expansion. However, limitations associated with immunogenicity *in vivo*, the fate of adenoviral cells, long-term safety, and the requirement for cell expansion in culture prior to viral infection and reimplantation remain significant hurdles before clinical use.

A variety of materials have been used for bone regeneration with mesenchymal stem cells and osteoprogenitors including ceramics or materials based on hydroxyapatite, ceramic forms of β -tricalcium phosphate, and composites of both hydroxyapatite and β -tricalcium phosphate.^[88] However, these materials are not osteoinductive and are resorbed relatively slowly *in vivo*. In addition, developments in recent years have sought to exploit and mimic biological cues that are necessary to mimic the cell-bone matrix interactions with the generation of biomimetic scaffolds.^[95] The development and use of synthetic materials such as poly(lactic acid) (PLA) poly(lactic-co-glycolic acid)

(PLGA), polyglycolic acid (PGA), polyanhydrides, polyimides, and polyphosphazenes, which can be surface-modified with biological agents, has made the development of such biomimetic scaffolds a realistic objective (e.g., Fig. 4).^[95,96] These synthetic materials, approved and widely used in the formation of resorbable sutures, scaffolds, and drug delivery devices, are 1) biocompatible, 2) processable into a three-dimensional structure and, 3) eventually degradable. Healy and coworkers^[97] and Dee and coworkers^[98] have reported on the ability to modulate the adhesion of osteoblasts selectively and significantly by immobilized -Arg-Gly-Asp- (-RGD-) and -Phe-His-Arg-Arg-Ile-Lys-Ala- (-FRRRIKA-) and Lys-Arg-Ser-Arg peptides (KRSR). Recently, Yang et al. have shown the potential to promote human osteoprogenitor differentiation on RGD-coupled biodegradable scaffolds.^[99] Continued advances in gene delivery combined with development of smart designer scaffolds and materials in combination with selection of mesenchymal stem cell and osteoprogenitor populations are needed to deliver on the promise of tissue engineering for skeletal repair.

FUTURE PERSPECTIVES AND SUMMARY

Given the demographic challenge of an aging population, developments in our understanding of mesenchymal stem cell and osteoprogenitor biology are paramount. From a patient's perspective, the ultimate goal will remain the utilization of mesenchymal stem cells and osteoprogenitors in simple, safe, and reproducible strategies for the augmentation of bone, cartilage, and tendon repair to address the pressing orthopedic clinical imperatives of many. Tissue engineering could provide suitable, efficacious, alternative therapies for orthopedic applications and is attractive on a number of fronts: 1) The ability to engineer tissue *in vitro* for transplantation would reduce

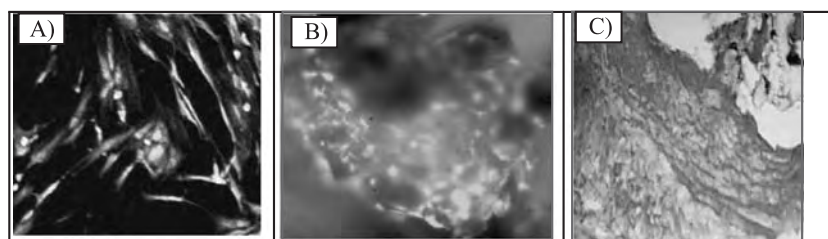


Fig. 3 Bone formation using selected and genetically modified osteoprogenitors. A) Cultures of adherent STRO-1+ cells immunostained using FITC conjugate and ethidium bromide nuclear stain ($\times 50$). B) Growth of STRO-1+ osteoprogenitors transduced with adenoviral BMP-2 and grown on PLA scaffolds demonstrating extensive penetration, colonization, and growth ($\times 25$). C) Alcian blue/sirius red-stained sections of STRO-1-selected cultures transduced with adenoviral BMP-2 and implanted *in vivo* demonstrating extensive new bone formation and organized collagen formation ($\times 25$).

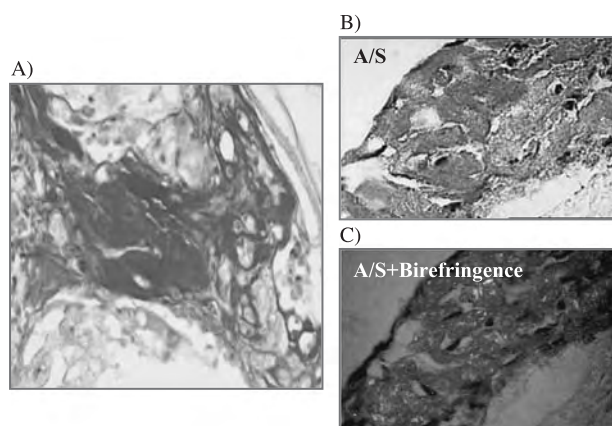


Fig. 4 Bone formation in vivo (A) and in vitro (B, C) using human osteoprogenitor cells on PLA scaffolds. A) Bone formation observed following subcutaneous implantation of human osteoprogenitors with pleiotrophin on PLA scaffolds in mF1 Nu/Nu mice after 6 weeks. B) and C) In-vitro mineralization on PLA scaffolds using human osteoprogenitor cells and pleiotrophin on PLA scaffolds stained using alcian blue and sirius red with birefringence microscopy ($\times 100$).

the requirement for donor tissue, as the number of skeletal cells required could be expanded ex vivo prior to implantation, 2) treatment of diseased tissue at an early stage with mesenchymal stem cells could alleviate or even cure the disease, reducing the need for life-long treatment and improving the quality of life of the patient. Thus, cell delivery vehicles, composite tissue engineering protocols, and smart materials that can orchestrate the bone formation cascade will all undoubtedly have their place. However, it is clear that many strategies are unlikely to reach clinical evaluation for the most basic of structural, biological, or safety issues and that this discipline is still in its infancy. Nevertheless, the ability to select, expand, and differentiate these cells and the potential for ex vivo gene transfer suggests the potential future use of human osteoprogenitors for gene therapy of metabolic bone diseases. Clearly, harnessing the homing potentials of such cells and channeling their development into the required human tissues for regeneration is a significant challenge, and it will be necessary to gain a greater understanding of the continuum of cell development of mesenchymal stem cells to osteoprogenitor to preosteoblast and ultimately osteocyte and the role of the matrix microenvironment and factors that control their fate and plasticity. While development pathways are clearly delineated from in vitro observations, their relevance in vivo and the extent of physiological population modulations (adipocyte to osteoblast, for example) are unclear. Ultimately, the need to define the phenotype of the mesenchymal stem cell, to generate a phenotypic finger print, will enable

isolation and expansion of populations for clinical use, which ultimately may improve the quality of life for many as a result of strategies to augment skeletal regeneration.

ACKNOWLEDGMENTS

Work in the author's laboratory is supported by grants from the BBSRC, EPSRC, and Research into Ageing. The work presented and many useful discussions are derived from all at University Orthopaedics: Daniel Howard, Rahul Tare, Xuebin Yang, David Green, Kris Partridge, Nicholas Clarke, and Trudy Roach, as well as fruitful collaborations with Professors Shakesheff and Howdle (University of Nottingham), Professor Mann (University of Bristol) and Dr. Julian Chaudhuri (University of Bath). The author apologizes to those authors whose valuable work it was not possible to cite due to restrictions of space.



ARTICLES OF FURTHER INTEREST

Biologic and Synthetic Apatites; Bone-Implant Interface; Bone Remodeling; Dental Implants; Histogenesis; Osteoinductive Substances and Materials; Poly(Glycolic Acid); Poly(lactic acid)s; Stem Cells; Tissue Engineering of Bone; Tissue Engineering of Bone Marrow

REFERENCES

1. Friedenstein, A.J.; Petrakova, K.V.; Kurolesova, A.I.; Frolova, G.P. Heterotopic of bone marrow. Analysis of precursor cells for osteogenic and hematopoietic tissues. *Transplantation* **1968**, *6* (2), 230–247.
2. Friedenstein, A.J.; Ivanov-Smolenski, A.A.; Chajlakjan, R.K.; Gorskaya, U.F.; Kuralsova, A.I.; Latzinik, N.W.; Gerasimow, U.W. Origin of bone marrow stromal mechanocytes in radiochimeras and heterotopic transplants. *Exp. Hematol.* **1978**, *6* (5), 440–444.
3. Friedenstein, A.J. Precursor cells of mechanocytes. *Int. Rev. Cyt.* **1976**, *47*, 327–359.
4. Lajtha, L.G. Stem cell concepts. *Differentiation* **1979**, *14* (1–2), 23–34.
5. Lovell-Badge, R. The future for stem cell research. *Nature* **2001**, *414* (6859), 88–91.
6. Triffitt, J.T. Stem cells and the philosopher's stone. *J. Cell. Biochem., Suppl.* **2002**, *38*, 13–19.
7. Owen, M.; Friedenstein, A.J. Stromal stem cells: Marrow-derived osteogenic precursors. *Ciba Found. Symp.* **1988**, *136*, 42–60.
8. Goshima, J.; Goldberg, V.M.; Caplan, A.I. The osteogenic potential of culture-expanded rat marrow mesenchymal cells assayed in vivo in calcium phosphate ceramic blocks. *Clin. Orthop.* **1991**, *262*, 298–311.

9. Bianco, P.; Robey, P.G. Stem cells in tissue engineering. *Nature* **2001**, *414* (6859), 118–121.
10. Bianco, P.; Riminucci, M.; Gronthos, S.; Robey, P.G. Bone marrow stromal stem cells: Nature, biology, and potential applications. *Stem Cells* **2001**, *19* (3), 180–192.
11. Jiang, Y.; Jahagirdar, B.N.; Reinhardt, R.L.; Schwartz, R.E.; Keene, C.D.; Ortiz-Gonzalez, X.R.; Reyes, M.; Lenvik, T.; Lund, T.; Blackstad, M.; Du, J.; Aldrich, S.; Lisberg, A.; Low, W.C.; Largaespada, D.A.; Verfaillie, C.M. Pluripotency of mesenchymal stem cells derived from adult marrow. *Nature* **2002**, *418* (6893), 41–49.
12. Triffitt, J.T.; Oreffo, R.O.C. *Osteoblast Lineage*, 1st Ed.; JAI Press, Inc., 1998; 475–498.
13. Nishida, S.; Endo, N.; Yamagiwa, H.; Tanizawa, T.; Takahashi, H.E. Number of osteoprogenitor cells in human bone marrow markedly decreases after skeletal maturation. *J. Bone Miner. Metab.* **1999**, *17* (3), 171–177.
14. D'Ippolito, G.; Schiller, P.C.; Ricordi, C.; Roos, B.A.; Howard, G.A. Age-related osteogenic potential of mesenchymal stromal stem cells from human vertebral bone marrow. *J. Bone Miner. Res.* **1999**, *14* (7), 1115–1122.
15. Majors, A.K.; Boehm, C.A.; Nitto, H.; Midura, R.J.; Muschler, G.F. Characterization of human bone marrow stromal cells with respect to osteoblastic differentiation. *J. Orthop. Res.* **1997**, *15* (4), 546–557.
16. Egrise, D.; Martin, D.; Vienne, A.; Neve, P.; Schoutens, A. The number of fibroblastic colonies formed from bone marrow is decreased and the in vitro proliferation rate of trabecular bone cells increased in aged rats. *Bone* **1992**, *13* (5), 355–361.
17. Bergman, R.J.; Gazit, D.; Kahn, A.J.; Gruber, H.; McDougall, S.; Hahn, T.J. Age-related changes in osteogenic stem cells in mice. *J. Bone Miner. Res.* **1996**, *11* (5), 568–577.
18. Oreffo, R.O.C.; Bennett, A.; Carr, A.J.; Triffitt, J.T. Patients with primary osteoarthritis show no change with ageing in the number of osteogenic precursors. *Scand. J. Rheumatol.* **1998**, *27* (6), 415–424.
19. Oreffo, R.O.C.; Bord, S.; Triffitt, J.T. Skeletal progenitor cells and ageing human populations. *Clin. Sci. (Colch.)* **1998**, *94* (5), 549–555.
20. Stenderup, K.; Justesen, J.; Eriksen, E.F.; Rattan, S.I.; Kassem, M. Number and proliferative capacity of osteogenic stem cells are maintained during aging and in patients with osteoporosis. *J. Bone Miner. Res.* **2001**, *16* (6), 1120–1129.
21. Pfeilschifter, J.; Diel, I.; Pilz, U.; Brunotte, K.; Naumann, A.; Ziegler, R. Mitogenic responsiveness of human bone cells in vitro to hormones and growth factors decreases with age. *J. Bone Miner. Res.* **1993**, *8* (6), 707–717.
22. Erdmann, J.; Kogler, C.; Diel, I.; Ziegler, R.; Pfeilschifter, J. Age-associated changes in the stimulatory effect of transforming growth factor beta on human osteogenic colony formation. *Mech. Ageing Dev.* **1999**, *110* (1–2), 73–85.
23. Verfaillie, C.M. Adult stem cells: Assessing the case for pluripotency. *Trends Cell Biol.* **2002**, *12* (11), 502–508.
24. Pittenger, M.F.; Mackay, A.M.; Beck, S.C.; Jaiswal, R.K.; Douglas, R.; Mosca, J.D.; Moorman, M.A.; Simonetti, D.W.; Craig, S.; Marshak, D.R. Multi-lineage potential of adult human mesenchymal stem cells. *Science* **1999**, *284* (5411), 143–147.
25. Bruder, S.P.; Horowitz, M.C.; Mosca, J.D.; Haynesworth, S.E. Monoclonal antibodies reactive with human osteogenic cell surface antigens. *Bone* **1997**, *21* (3), 225–235.
26. Simmons, P.J.; Torok-Storb, B. CD34 expression by stromal precursors in normal human adult bone marrow. *Blood* **1991**, *78* (11), 2848–2853.
27. Haynesworth, S.E.; Baber, M.A.; Caplan, A.I. Cell surface antigens on human marrow-derived mesenchymal cells are detected by monoclonal antibodies. *Bone* **1992**, *13* (1), 69–80.
28. Joyner, C.J.; Bennett, A.; Triffitt, J.T. Identification and enrichment of human osteoprogenitor cells by using differentiation stage-specific monoclonal antibodies. *Bone* **1997**, *21* (1), 1–6.
29. Gronthos, S.; Graves, S.E.; Ohta, S.; Simmons, P.J. The STRO-1 + fraction of adult human bone marrow contains the osteogenic precursors. *Blood* **1994**, *84* (12), 4164–4173.
30. Stewart, K.; Walsh, S.; Screen, J.; Jefferiss, C.M.; Chainey, J.; Jordan, G.R.; Beresford, J.N. Further characterization of cells expressing STRO-1 in cultures of adult human bone marrow stromal cells. *J. Bone Miner. Res.* **1999**, *14* (8), 1345–1356.
31. Howard, D.; Partridge, K.; Yang, X.; Clarke, N.M.; Okubo, Y.; Bessho, K.; Howdle, S.M.; Shakesheff, K.M.; Oreffo, R.O.C. Immunoselection and adenoviral genetic modulation of human osteoprogenitors: In vivo bone formation on PLA scaffold. *Biochem. Biophys. Res. Commun.* **2002**, *299* (2), 208–215.
32. Bruder, S.P.; Ricalton, N.S.; Boynton, R.E.; Connolly, T.J.; Jaiswal, N.; Zaia, J.; Barry, F.P. Mesenchymal stem cell surface antigen SB-10 corresponds to activated leukocyte cell adhesion molecule and is involved in osteogenic differentiation. *J. Bone Miner. Res.* **1998**, *13* (4), 655–663.
33. Barry, F.P.; Boynton, R.E.; Haynesworth, S.; Murphy, J.M.; Zaia, J. The monoclonal antibody SH-2, raised against human mesenchymal stem cells, recognizes an epitope on endoglin (CD105). *Biochem. Biophys. Res. Commun.* **1999**, *265* (1), 134–139.
34. Triffitt, J.T.; Oreffo, R.O.C.; Viridi, A.S.; Xia, Z. Osteogenic stem-cell characterization and development: Potentials for cytotераpy. *Cytotherapy* **2001**, *3* (5), 413–416.
35. Park, S.R.; Oreffo, R.O.C.; Triffitt, J.T. Interconversion potential of cloned human marrow adipocytes in vitro. *Bone* **1999**, *24* (6), 549–554.
36. Kuznetsov, S.A.; Mankani, M.H.; Gronthos, S.; Sato-mura, K.; Bianco, P.; Robey, P.G. Circulating skeletal stem cells. *J. Cell Biol.* **2001**, *153* (5), 1133–1140.
37. Friedenstein, A.J.; Deriglasova, U.F.; Kulagina, N.N.; Panasuk, A.F.; Rudakowa, S.F.; Luria, E.A.; Ruadkowi, I.A. Precursors for fibroblasts in different populations of hematopoietic cells as detected by the in vitro colony assay method. *Exp. Hematol.* **1974**, *2* (2), 83–92.

38. Nuttall, M.E.; Gimble, J.M. Is there a therapeutic opportunity to either prevent or treat osteopenic disorders by inhibiting marrow adipogenesis? *Bone* **2000**, *27* (2), 177–184.
39. Nuttall, M.E.; Patton, A.J.; Olivera, D.L.; Nadeau, D.P.; Gowen, M. Human trabecular bone cells are able to express both osteoblastic and adipocytic phenotype: Implications for osteopenic disorders. *J. Bone Miner. Res.* **1998**, *13* (3), 371–382.
40. Wozney, J.M.; Rosen, V. Bone morphogenetic protein and bone morphogenetic protein gene family in bone formation and repair. *Clin. Orthop.* **1998**, *346*, 26–37.
41. Chen, D.; Ji, X.; Harris, M.A.; Feng, J.Q.; Karsenty, G.; Celeste, A.J.; Rosen, V.; Mundy, G.R.; Harris, S.E. Differential roles for bone morphogenetic protein (BMP) receptor type IB and IA in differentiation and specification of mesenchymal precursor cells to osteoblast and adipocyte lineages. *J. Cell Biol.* **1998**, *142* (1), 295–305.
42. Pfeilschifter, J.; Wolf, O.; Naumann, A.; Minne, H.W.; Mundy, G.R.; Ziegler, R. Chemotactic response of osteoblastlike cells to transforming growth factor beta. *J. Bone Miner. Res.* **1990**, *5* (8), 825–830.
43. Locklin, R.M.; Oreffo, R.O.C.; Triffitt, J.T. Effects of TGFbeta and bFGF on the differentiation of human bone marrow stromal fibroblasts. *Cell Biol. Int.* **1999**, *23* (3), 185–194.
44. Walsh, S.; Jefferiss, C.; Stewart, K.; Beresford, J.N. TGFbeta1 limits the expansion of the osteoprogenitor fraction in cultures of human bone marrow stromal cells. *Cell Tissue Res.* **2003**, *311* (2), 187–198.
45. Wang, E.A.; Israel, D.I.; Kelly, S.; Luxenberg, D.P. Bone morphogenetic protein-2 causes commitment and differentiation in C3H10T1/2 and 3T3 cells. *Growth Factors* **1993**, *9* (1), 57–71.
46. Rickard, D.J.; Sullivan, T.A.; Shenker, B.J.; Leboy, P.S.; Kazhdan, I. Induction of rapid osteoblast differentiation in rat bone marrow stromal cell cultures by dexamethasone and BMP-2. *Dev. Biol.* **1994**, *161* (1), 218–228.
47. Gronthos, S.; Simmons, P.J. The growth factor requirements of STRO-1-positive human bone marrow stromal precursors under serum-deprived conditions in vitro. *Blood* **1995**, *85* (4), 929–940.
48. Kuznetsov, S.A.; Friedenstein, A.J.; Robey, P.G. Factors required for bone marrow stromal fibroblast colony formation in vitro. *Br. J. Haematol.* **1997**, *97* (3), 561–570.
49. Oreffo, R.O.C.; Romberg, S.; Viridi, A.S.; Joyner, C.J.; Berven, S.; Triffitt, J.T. Effects of interferon alpha on human osteoprogenitor cell growth and differentiation in vitro. *J. Cell. Biochem.* **1999**, *74* (3), 372–385.
50. Walsh, S.; Jefferiss, C.; Stewart, K.; Jordan, G.R.; Screen, J.; Beresford, J.N. Expression of the developmental markers STRO-1 and alkaline phosphatase in cultures of human marrow stromal cells: Regulation by fibroblast growth factor (FGF)-2 and relationship to the expression of FGF receptors 1–4. *Bone* **2000**, *27* (2), 185–195.
51. Martin, I.; Muraglia, A.; Campanile, G.; Cancedda, R.; Quarto, R. Fibroblast growth factor-2 supports ex vivo expansion and maintenance of osteogenic precursors from human bone marrow. *Endocrinology* **1997**, *138* (10), 4456–4462.
52. Locklin, R.M.; Williamson, M.C.; Beresford, J.N.; Triffitt, J.T.; Owen, M.E. In vitro effects of growth factors and dexamethasone on rat marrow stromal cells. *Clin. Orthop.* **1995**, *313*, 27–35.
53. Oreffo, R.O.C.; Viridi, A.S.; Triffitt, J.T. Modulation of osteogenesis and adipogenesis by human serum in human bone marrow cultures. *Eur. J. Cell Biol.* **1997**, *74* (3), 251–261.
54. Cheng, S.L.; Yang, J.W.; Rifas, L.; Zhang, S.F.; Avioli, L.V. Differentiation of human bone marrow osteogenic stromal cells in vitro: Induction of the osteoblast phenotype by dexamethasone. *Endocrinology* **1994**, *134* (1), 277–286.
55. Kimoto, S.; Cheng, S.L.; Zhang, S.F.; Avioli, L.V. The effect of glucocorticoid on the synthesis of biglycan and decorin in human osteoblasts and bone marrow stromal cells. *Endocrinology* **1994**, *135* (6), 2423–2431.
56. Walsh, S.; Jordan, G.R.; Jefferiss, C.; Stewart, K.; Beresford, J.N. High concentrations of dexamethasone suppress the proliferation but not the differentiation or further maturation of human osteoblast precursors in vitro: Relevance to glucocorticoid-induced osteoporosis. *Rheumatol. (Oxf.)* **2001**, *40* (1), 74–83.
57. Beresford, J.N.; Joyner, C.J.; Devlin, C.; Triffitt, J.T. The effects of dexamethasone and 1,25-dihydroxyvitamin D3 on osteogenic differentiation of human marrow stromal cells in vitro. *Arch. Oral Biol.* **1994**, *39* (11), 941–947.
58. Scutt, A.; Bertram, P. Bone marrow cells are targets for the anabolic actions of prostaglandin E2 on bone: Induction of a transition from nonadherent to adherent osteoblast precursors. *J. Bone Miner. Res.* **1995**, *10* (3), 474–487.
59. Still, K.; Scutt, A. Stimulation of CFU—F formation by prostaglandin E2 is mediated in part by its degradation product, prostaglandin A2. *Prostaglandins Other Lipid Mediat.* **2001**, *65* (1), 21–31.
60. Kasugai, S.; Oida, S.; Iimura, T.; Arai, N.; Takeda, K.; Ohya, K.; Sasaki, S. Expression of prostaglandin E receptor subtypes in bone: Expression of EP2 in bone development. *Bone* **1995**, *17* (1), 1–4.
61. Oreffo, R.O.C.; Wells, N.; Johnstone, D. The presence of prostaglandin receptors on osteoblasts and osteoclasts. *J. Bone Miner. Res.* **1991**, *6* (Suppl. 1), S208.
62. Yang, X.; Tare, R.S.; Partridge, K.A.; Roach, H.I.; Clarke, N.M.; Howdle, S.M.; Shakesheff, K.M.; Oreffo, R.O.C. Induction of human osteoprogenitor chemotaxis, proliferation, differentiation, and bone formation by osteoblast stimulating factor-1/pleiotrophin: Osteoconductive biomimetic scaffolds for tissue engineering. *J. Bone Miner. Res.* **2003**, *18* (1), 47–57.
63. Tare, R.S.; Oreffo, R.O.C.; Clarke, N.M.; Roach, H.I. Pleiotrophin/osteoblast-stimulating factor 1: Dissecting its diverse functions in bone formation. *J. Bone Miner. Res.* **2002**, *17* (11), 2009–2020.
64. Jilka, R.L. Cytokines, bone remodeling, and estrogen deficiency: A 1998 update. *Bone* **1998**, *23* (2), 75–81.
65. Manolagas, S.C.; Jilka, R.L. Bone marrow, cytokines, and bone remodeling. Emerging insights into the

- pathophysiology of osteoporosis. *N. Engl. J. Med.* **1995**, *332* (5), 305–311.
66. Boyce, B.F.; Aufdemorte, T.B.; Garrett, I.R.; Yates, A.J.; Mundy, G.R. Effects of interleukin-1 on bone turnover in normal mice. *Endocrinology* **1989**, *125* (3), 1142–1150.
 67. Chapekar, M.S. Tissue engineering: Challenges and opportunities. *J. Biomed. Mater. Res.* **2000**, *53* (6), 617–620.
 68. Jordan, K.M.; Cooper, C. Epidemiology of osteoporosis. *Best. Pract. Res. Clin. Rheumatol.* **2002**, *16* (5), 795–806.
 69. Spector, M. Biomaterial failure. *Orthop. Clin. North Am.* **1992**, *23* (2), 211–217.
 70. Rose, F.R.; Oreffo, R.O.C. Bone tissue engineering: Hope vs. hype. *Biochem. Biophys. Res. Commun.* **2002**, *292* (1), 1–7.
 71. Langer, R.; Vacanti, J.P. Tissue engineering. *Science* **1993**, *260* (5110), 920–926.
 72. Kadiyala, S.; Young, R.G.; Thiede, M.A.; Bruder, S.P. Culture expanded canine mesenchymal stem cells possess osteochondrogenic potential in vivo and in vitro. *Cell Transplant.* **1997**, *6* (2), 125–134.
 73. Gundle, R.; Joyner, C.J.; Triffitt, J.T. Human bone tissue formation in diffusion chamber culture in vivo by bone-derived cells and marrow stromal fibroblastic cells. *Bone* **1995**, *16* (6), 597–601.
 74. Nakahara, H.; Bruder, S.P.; Haynesworth, S.E.; Holecek, J.J.; Baber, M.A.; Goldberg, V.M.; Caplan, A.I. Bone and cartilage formation in diffusion chambers by subcultured cells derived from the periosteum. *Bone* **1990**, *11* (3), 181–188.
 75. Ohgushi, H.; Goldberg, V.M.; Caplan, A.I. Heterotopic osteogenesis in porous ceramics induced by marrow cells. *J. Orthop. Res.* **1989**, *7* (4), 568–578.
 76. Simmons, P.J.; Storb, B. Torok-identification of stromal cell precursors in human bone marrow by a novel monoclonal antibody, STRO-1. *Blood* **1991**, *78* (1), 55–62.
 77. Stewart, K.; Walsh, S.; Screen, J.; Jefferiss, C.M.; Chainey, J.; Jordan, G.R.; Beresford, J.N. Further characterization of cells expressing STRO-1 in cultures of adult human bone marrow stromal cells. *J. Bone Miner. Res.* **1999**, *14* (8), 1345–1356.
 78. Bruder, S.P.; Jaiswal, N.; Haynesworth, S.E. Growth kinetics, self-renewal, and the osteogenic potential of purified human mesenchymal stem cells during extensive subcultivation and following cryopreservation. *J. Cell. Biochem.* **1997**, *64* (2), 278–294.
 79. Jaiswal, N.; Haynesworth, S.E.; Caplan, A.I.; Bruder, S.P. Osteogenic differentiation of purified, culture-expanded human mesenchymal stem cells in vitro. *J. Cell. Biochem.* **1997**, *64* (2), 295–312.
 80. Kadiyala, S.; Young, R.G.; Thiede, M.A.; Bruder, S.P. Culture expanded canine mesenchymal stem cells possess osteochondrogenic potential in vivo and in vitro. *Cell Transplant* **1997**, *6* (2), 125–134.
 81. Horwitz, E.M.; Prockop, D.J.; Fitzpatrick, L.A.; Koo, W.W.; Gordon, P.L.; Neel, M.; Sussman, M.; Orchard, P.; Marx, J.C.; Pyeritz, R.E.; Brenner, M.K. Transplantability and therapeutic effects of bone marrow-derived mesenchymal cells in children with osteogenesis imperfecta. *Nat. Med.* **1999**, *5* (3), 309–313.
 82. Horwitz, E.M.; Prockop, D.J.; Gordon, P.L.; Koo, W.W.; Fitzpatrick, L.A.; Neel, M.D.; McCarville, M.E.; Orchard, P.J.; Pyeritz, R.E.; Brenner, M.K. Clinical responses to bone marrow transplantation in children with severe osteogenesis imperfecta. *Blood* **2001**, *97* (5), 1227–1231.
 83. Dominici, M.; Hofmann, T.J.; Horwitz, E.M. Bone marrow mesenchymal cells: Biological properties and clinical applications. *J. Biol. Regul. Homeost. Agents* **2001**, *15* (1), 28–37.
 84. Connolly, J.F. Injectable bone marrow preparations to stimulate osteogenic repair. *Clin. Orthop.* **1995**, *313*, 8–18.
 85. Quarto, R.; Mastrogiacomo, M.; Cancedda, R.; Kutepov, S.M.; Mukhachev, V.; Lavroukov, A.; Kon, E.; Marcacci, M. Repair of large bone defects with the use of autologous bone marrow stromal cells. *N. Engl. J. Med.* **2001**, *344* (5), 385–386.
 86. Perry, C.R. Bone repair techniques, bone graft, and bone graft substitutes. *Clin. Orthop.* **1999**, *360*, 71–86.
 87. Boden, S.D. Bioactive factors for bone tissue engineering. *Clin. Orthop.* **1999**, *367* (Suppl. 1), 84–94.
 88. Oreffo, R.O.C.; Triffitt, J.T. Future potentials for using osteogenic stem cells and biomaterials in orthopedics. *Bone* **1999**, *25* (2 Suppl.), 5S–9S.
 89. Lieberman, J.R.; Daluiski, A.; Stevenson, S.; Wu, L.; McAllister, P.; Lee, Y.P.; Kabo, J.M.; Finerman, G.A.; Berk, A.J.; Witte, O.N. The effect of regional gene therapy with bone morphogenetic protein-2-producing bone-marrow cells on the repair of segmental femoral defects in rats. *J. Bone Jt. Surg. Am.* **1999**, *81* (7), 905–917.
 90. Breitbart, A.S.; Grande, D.A.; Mason, J.M.; Barcia, M.; James, T.; Grant, R.T. Gene-enhanced tissue engineering: Applications for bone healing using cultured periosteal cells transduced retrovirally with the BMP-7 gene. *Ann. Plast. Surg.* **1999**, *42* (5), 488–495.
 91. Olmsted, E.A.; Blum, J.S.; Rill, D.; Yotnda, P.; Gugala, Z.; Lindsey, R.W.; Davis, A.R. Adenovirus-mediated BMP2 expression in human bone marrow stromal cells. *J. Cell. Biochem.* **2001**, *82* (1), 11–21.
 92. Turgeman, G.; Pittman, D.D.; Muller, R.; Kurkalli, B.G.; Zhou, S.; Pelled, G.; Peyser, A.; Zilberman, Y.; Moutsatsos, I.K.; Gazit, D. Engineered human mesenchymal stem cells: A novel platform for skeletal cell mediated gene therapy. *J. Gene Med.* **2001**, *3* (3), 240–251.
 93. Partridge, K.; Yang, X.; Clarke, N.M.; Okubo, Y.; Bessho, K.; Sebald, W.; Howdle, S.M.; Shakesheff, K.M.; Oreffo, R.O.C. Adenoviral BMP-2 gene transfer in mesenchymal stem cells: In vitro and in vivo bone formation on biodegradable polymer scaffolds. *Biochem. Biophys. Res. Commun.* **2002**, *292* (1), 144–152.
 94. Musgrave, D.S.; Bosch, P.; Ghivizzani, S.; Robbins, P.D.; Evans, C.H.; Huard, J. Adenovirus-mediated direct gene therapy with bone morphogenetic protein-2 produces bone. *Bone* **1999**, *24* (6), 541–547.
 95. Patel, N.; Padera, R.; Sanders, G.H.; Cannizzaro, S.M.; Davies, M.C.; Langer, R.; Roberts, C.J.; Tandler, S.J.;

- Williams, P.M.; Shakesheff, K.M. Spatially controlled cell engineering on biodegradable polymer surfaces. *FASEB J.* **1998**, *12* (14), 1447–1454.
96. Pearson, R.G.; Bhandari, R.; Quirk, R.A.; Shakesheff, K.M. Recent advances in tissue engineering: An invited review. *J. Long-Term Eff. Med. Implants* **2002**, *12* (1), 1–33.
97. Reznia, A.; Healy, K.E. Integrin subunits responsible for adhesion of human osteoblast-like cells to biomimetic peptide surfaces. *J. Orthop. Res.* **1999**, *17* (4), 615–623.
98. Dee, K.C.; Anderson, T.T.; Bizios, R. Osteoblast population migration characteristics on substrates modified with immobilized adhesive peptides. *Biomaterials* **1999**, *20* (3), 221–227.
99. Yang, X.B.; Roach, H.I.; Clarke, N.M.; Howdle, S.M.; Quirk, R.; Shakesheff, K.M.; Oreffo, R.O.C. Human osteoprogenitor growth and differentiation on synthetic biodegradable structures after surface modification. *Bone* **2001**, *29* (6), 523–531.

Osteoinductive Substances and Materials

Toshimasa Uemura

Hiroko Kojima

Age Dimension Research Center, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Ibaraki, Japan

Takashi Saito

Department of Operative Dentistry and Endodontology, Health Sciences University of Hokkaido, Ishikari-gun, Hokkaido, Japan

INTRODUCTION

The meaning of scientific technical terms changes from time to time, depending on progress in the field. Osteoinduction in the field of biomaterials is one example. In dictionaries of biomaterials, osteoinduction is defined as the “act or process of stimulating osteogenesis.” In this article, the authors define osteoinductive substances and materials as biological molecules such as proteins, cytokines, etc., as well as materials that induce bone formation. This does not include osteoconductive substances and materials, such as hydroxyapatite. In this article, representative osteoinductive substances and materials are described.

OSTEOINDUCTIVE SUBSTANCES

BMP (Bone Morphogenetic Protein)

BMPs are multifunctional, low-molecular weight proteins belonging to the extended transforming growth factor superfamily.^[1] In 1965, Urist isolated an osteoinductive extract from adult bone, which can induce new endochondral bone formation at ectopic sites in rodents.^[2] Currently, several BMPs are known to belong to the TGF- β (transforming growth factor β) superfamily. Members of the BMP subfamily show a pattern of seven conserved sequences in the mature protein. Mature BMP proteins are synthesized as precursor molecules, which are processed to approximately 30 kDa dimers before their secretion from the cells. The BMP family express their function by binding to transmembrane serine/threonine kinases that are named type I and type II receptors. BMPs can bind to type I and type II receptors, but the affinity is low. However, when both receptors exist, BMPs bind strongly. The binding of BMPs to the receptors is regulated by extracellular antagonists that bind to BMP and prevent BMP from binding to the receptor. Representative antagonists are noggin, GDF5, GDF6, chordin, and follistatin. Osteogenic activities of BMPs

have been demonstrated using ectopic and orthotopic implantations, which indicates BMP-2, 4, 5, 6, 7, and 9 are osteogenic molecules.^[3-6] A basic understanding of BMP functions has been achieved.^[7,8] In this section, the clinical utility of BMPs is focused on.

First of all, osteoinductive BMPs are expected to be clinically useful for orthopedic surgery. This is because only small amounts of osteoinductive BMPs are necessary for producing substantial amounts of bone. Osteoinductive BMPs such as BMP-2 and BMP-7 were found to heal large defects in rats, rabbits, dogs, and monkeys using the large, segmental long bone defect model.^[9] Results confirmed the potent osteoinductive ability and demonstrated that BMP induced the formation of bone intimately associated with preexisting surrounding bone and remodeled depending on its location. The spinal fusion procedure utilizes a large amount of autogeneous bone graft and provides a case for the application of BMP as a bone graft substitute. RhBMP-2 (recombinant human bone morphogenetic protein-2) and RhBMP-7 have been successfully examined in preclinical animal models of the intertransverse process and interbody fusion.^[10-13] Clinical studies utilizing spinal fusion applications are underway.

Osteoinductive BMPs are potential mediators of fracture repair when sufficient repair cannot be expected. Studies have demonstrated that BMP-2 and BMP-7 accelerated fracture repair in rabbit, goat, and dog models.^[14-17] Clinical studies are also in progress.

Several studies have used demineralized dentin powder, which is like demineralized bone powder and will be described later in this article. Demineralized dentin implanted into dental pulp induced mineralized tissue formation.^[18-21] Although it is known that BMPs induce dentinogenesis, the mode of action and molecular understanding of dentin BMP is not very well understood. However, recombinant BMPs are expected to induce dentinogenesis. Recently, Nakashima reported that recombinant BMP-2 and BMP-4, but not TGF- β , induced reparative dentinogenesis in dogs when placed on partially amputated dental

pulps.^[22] Similar results were obtained using recombinant human BMP-7.^[23,24] For utilizing osteogenic BMPs, bioresorbable collagen matrices are usually chosen as a delivery substance. However, a biodegradable, synthetic polymer delivery system may prove to be a more useful tool.

Recently, BMP gene therapy has attracted much attention in orthopedics and the tissue engineering field. The BMP gene is introduced directly into the body or into cells and is then followed by *ex vivo* infection and reintroduction into the body. Adenoviral vectors carrying BMP-2, BMP-7, and BMP-9 have been used directly in animal models for enhancing fractures and promoting spinal fusion.^[25–29]

Bone Proteins

Collagen type I, osteopontin (OPN), BSP (bone sialoprotein), osteocalcin, and osteonectin are typical bone proteins in bone matrix, which regulates bone metabolism.^[30] Among them, OPN is not only a differentiation marker of osteoblasts but also a regulating molecule of osteoblasts. OPN is a highly acidic secreted phosphoprotein that binds to cells via an RGD cell adhesion sequence that recognizes $\alpha_v\beta_3$ integrin. The signal transduction pathway of OPN-integrin binding in osteoblasts was studied by our group regarding osteoblast differentiation.^[31–36] The production and autophosphorylation of FAK (focal adhesion kinase), the expression of alkaline phosphatase (ALP) at the transcription level, and their kinetics were examined through OPN-integrin binding compared with collagen type I and integrin binding. Results suggested that the OPN acts as a trigger in the early differentiation of osteoblasts. In this regard, OPN is categorized as an osteoinductive substance. We examined the effect of a coating of OPN on the hydroxyapatite (HA) scaffold for a cultured bone transplantation model.^[37] According to the method of Yoshikawa et al., we subcutaneously implanted OPN-coated HA/bone marrow-derived osteoblast constructs and noncoated constructs (control) into Fischer rats. At 18, 32, and 46 days after implantation, the marrow cell/HA constructs were harvested, and their ALP activities were measured. The ALP activity of the marrow cells/OPN-coated HA construct showed a similar maximum at around 18 days with 40% higher activity than the control group (marrow cells/noncoated-HA). This indicated that the OPN coating resulted in greater bone formation with the marrow cell/HA construct *in vivo*. This result means that OPN activates osteoblast differentiation for bone formation. On the other hand, secreted OPN on HA inhibits further growth of HA.^[38] These properties of OPN might regulate the HA formation *in vivo* in a complex way.

Phosphophoryn

The process of mineralization in vertebrates involves a sequential and localized series of events that leads to the formation and growth of carbonate containing apatite mineral within an organic matrix. The matrix network into which the mineral of bone, dentin, and cementum is deposited is primarily type I collagen.^[39] Acidic, noncollagenous components of the matrix directly associated with the collagen network are thought to regulate collagen mineralization.^[40] This regulation includes control of the nucleation and the postnucleation growth of the crystals and the structural organization of the tissue. Several of the noncollagenous components might have more than one function in the regulation of mineralization.^[41,42] Notably, dentin phosphoprotein (phosphophoryn), a synthetic product of the odontoblast, is accumulated at the mineralization front of the dentin matrix and is strongly associated with the mineral phase.^[43] It has high affinity to calcium ion at physiological pH.^[44] Moreover, it binds about 1.5 additional calcium ions in the presence of phosphate ion.^[45]

The pattern of mineral deposition in dentin is correlated with the structure of the collagen fibrils.^[46] Therefore, one would expect that a mineral nucleator would have a strong association with the collagen fibrils of dentin and have a fixed distribution on the fibrils. Stetler-Stevenson and Veis^[47] found that phosphophoryn binds to collagen molecules and the surface of type I collagen fibrils from skin but not to heat-denatured collagen. The phosphoprotein from egg yolk, phosvitin, competed with phosphophoryn to bind to collagen, but other phosphoproteins such as bovine serum albumin, fibronectin, and osteonectin did not. The bound phosphophoryn enhanced the binding of calcium to collagen fibrils and was considered to have the potential to nucleate apatite crystals oriented with respect to the collagen fibrils. Electron microscopic observations showed that phosphophoryn is adsorbed by collagen fibrils near the gap region and might regulate mineral deposition in the gap region of collagen fibrils.^[48] Fujisawa and Kuboki^[49] confirmed that phosphophoryn bound to type I collagen fibrils from skin, but they found that the phosphoprotein also bound to heat-denatured collagen and that calcium sharply reduced the amount of bound phosphophoryn.

Recently, we established an *in-vitro* mineralization system.^[50–52] First, bovine phosphophoryn was immobilized on reconstituted skin type I collagen fibrils or agarose beads by adsorption or covalent cross-linking. In some samples, the ester phosphate was removed from the covalently cross-linked phosphophoryn by treatment with acid phosphatase. All samples were incubated at 37°C in metastable solutions that do not spontaneously precipitate. Reconstituted

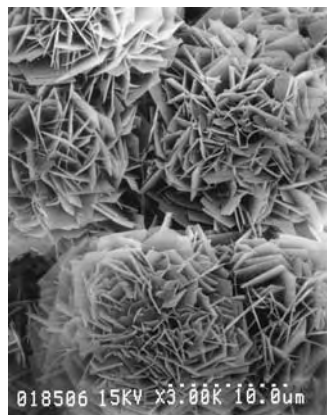


Fig. 1 Mineral formation on phosphophoryn cross-linked to agarose beads.

collagen fibrils alone did not induce mineral formation. The phosphophoryn adsorbed by the agarose beads or the collagen fibrils was desorbed when the mineralization medium was added, and mineral was not induced. The mineral induced by the cross-linked phosphophoryn was confirmed as apatite from the X-ray diffraction pattern, and the crystals were confined to the surface of the agarose beads (Fig. 1) or the collagen fibrils (Fig. 2).^[50,51] With decreasing medium saturation, the time required for mineral induction increased.^[51] The interfacial tensions calculated for apatite formation by immobilized phosphophoryn was about the same as that found for phosphatidic acid liposomes. This similarity in values for the interfacial tension might indicate that the nucleation potential of each of these highly phosphorylated organic surfaces is the same.^[50,51] We concluded that phosphophoryn must be cross-linked to collagen for the mineralization of the collagen fibrillar network in solutions that do not spontaneously precipitate. The phosphate esters of phosphophoryn are required for mineral induction, and the carboxylate groups of phosphophoryn are not sufficient.

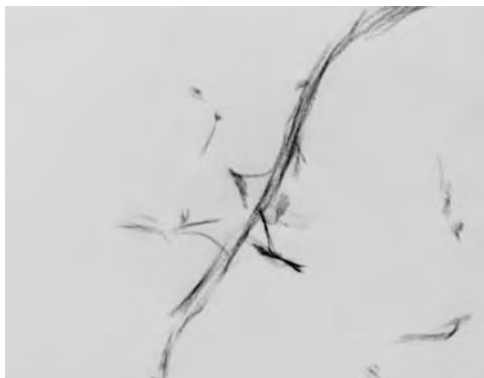


Fig. 2 Mineral formation on phosphophoryn cross-linked to collagen fibrils. Bar = 0.5 μm .

On the basis of the above study, we expected the phosphophoryn/collagen composites to be a good scaffold for bone tissue engineering and examined the formation of bone in orthotopically transplanted phosphophoryn/collagen and bone marrow osteoblast composites in Fischer rats.

Osteoblastic primary cells were obtained from the bone shaft of femorae of 344 seven-week-old male Fischer rats, according to the method of Maniatopoulos et al.^[53] The cell suspension of marrow cells was distributed through a T-75 flask with 15 ml of standard culture medium and incubated at 37°C in a humidified atmosphere containing 95% air and 5% CO₂. When culture dishes were near to confluent, after 10 days, they were treated with 0.1% trypsin and then concentrated by centrifuge to 10⁶ cells/ml and sub-cultured onto the synthesized phosphophoryn/collagen sponge and collagen sponge (control). After 14 days in culture with standard medium supplemented with 10 nM Dexamethasone, 10 mM β -glycerophosphate, and 50 $\mu\text{g/ml}$ ascorbic acid, the composites of phosphophoryn/collagen and osteoblastic cells and control composites were transplanted into holes that were 2 mm in diameter in the bone of Fischer rats. Then, the wounds were sutured. The composites were harvested at one, two, six, and eight weeks after implantation, fixed with 4% paraformaldehyde and 0.25% glutaraldehyde, decalcified, embedded in paraffin, and stained with hematoxylin and eosin.

Results clearly showed that more bone formed in the phosphophoryn/collagen sponge and osteoblasts composites than control composites from one week to eight weeks,^[54] which suggests that the phosphophoryn/collagen sponge is a good candidate as a scaffold for bone tissue engineering.

OSTEOINDUCTIVE MATERIALS

Bioglass

Bioglass in the Na₂O–CaO–SiO₂–P₂O₅ system is a bioactive material that has been clinically used as an important bone substitute.^[55] Bioglass and bioglass-type glasses have gained over a million successes in periodontal bone repair.^[56] As the clinical application of bioglass progresses, its limited mechanical properties are becoming apparent. Currently, several studies are under way to improve this mechanical weakness (e.g., Kokubo et al.^[57]).

Demineralized Bone

Osteogenesis can be induced in experimental animals using demineralized matrix of bone or dentin.^[2,58,59]

Mulliken and Growacki showed that demineralized human bone induced osteogenesis in the rat.^[60] When implanted into subcutaneous sites, the demineralized bone matrix causes connective tissues to produce endochondral bone.^[58,61] Furthermore, demineralized bone powder can bridge large defects in long bones in rats.^[62] Demineralized bone implants were examined for craniofacial,^[63] maxillofacial,^[64] periodontal,^[65] and phalangeal^[66,67] defects. Thus, demineralized bone has applications for several bone defects.

Biodegradable Ceramics

Calcium phosphates, which are used clinically in orthopedic surgery, have attracted attention as bone substitutes due to their good biocompatibility and osteointegrative properties.^[68–77] The application of hydroxyapatite (HA) for this purpose has the advantage of stability and maintenance of compressive strength, especially when block-shaped implants are used.^[69] However, the use of a rigid nonresorbable HA for skeletal reconstructions is associated with potential long-term interference with mechanical stress and strain in load-bearing areas. Furthermore, HA cannot be replaced by new bone. Therefore, a fracture might occur either in the HA itself or at the interface with the host bone. The clinical application of HA has thus been restricted due to its poor absorbability and brittleness. Consequently, a degradable material that can eventually be replaced by bone tissue is needed for bone repairs. A suitable macroporous structure is important to obtain good implant incorporation through rapid vascularization and bone ingrowth.^[78–80]

Recently, porous tricalcium phosphate (TCP) has been attracting attention due to its good biodegradability and biocompatibility.^[71,75,77,81] However, while some reports on biodegradable porous TCP alone implanted at extraskeletal sites suggest that degradation occurred quickly with no formation of bone,^[70,72] at least one study on porous TCP implanted at skeletal sites found that the bone that formed during the initial stages was resorbed later on so that bone repair after one year was not significantly enhanced.^[82] Moreover, some solid β -TCPs have been reported to be completely resorbable at skeletal sites; although they were not porous but granular.^[76,83]

Owing to the rapid degradation and weak mechanical properties of TCP, much research has focused on mixed calcium phosphates (mixtures of TCP and HA or TCP and polymer).^[73,77,79,84–88] However, the purity of β -TCP also influences bone formation and biocompatibility.^[89]

To try to overcome these difficulties, a highly pure porous β -TCP was manufactured as an implantable material, Osferion. This biodegradable β -TCP was

combined with bone marrow-derived osteoprogenitor cells (BMO) and cultured in osteogenic medium in vitro. We hypothesize that this composite of β -TCP and BMO could improve bone formation and slow down the degradation of β -TCP. The results are shown in the following section.

Application of Porous β -TCP for Bone Tissue Engineering

In this section, our application of porous β -TCP for bone tissue engineering is introduced. A schematic diagram of a cultured bone transplantation model is shown in Fig. 3. The cultured bone transplantation is clinically important; however, several problems remain to be solved before any clinical applications. First of all, appropriate materials are necessary for cell scaffolding. Our study focuses on the improvement of

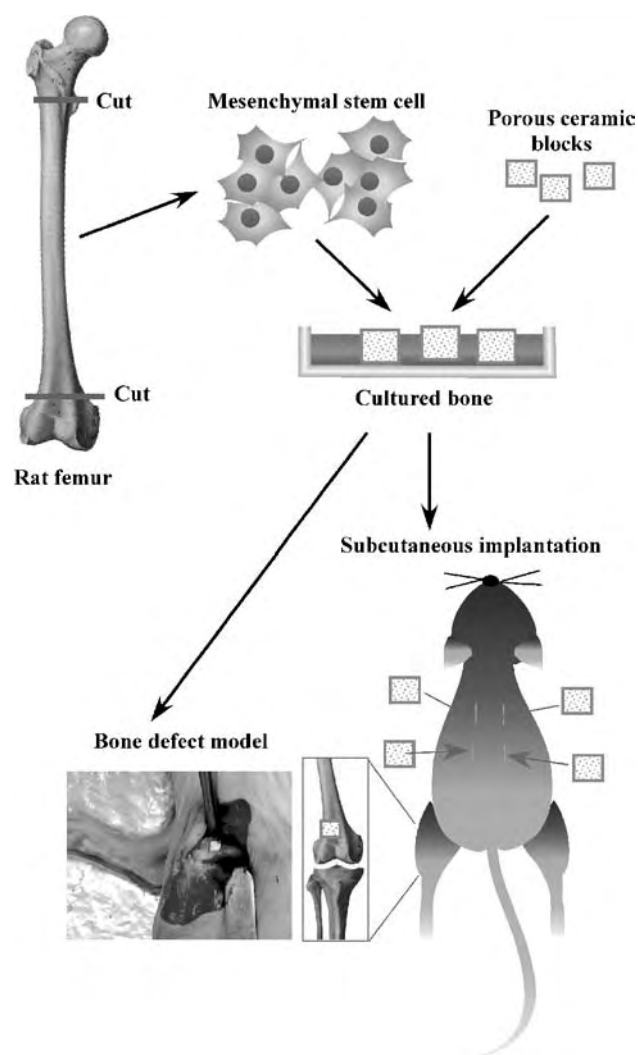


Fig. 3 Schematic diagram of cultured bone transplantation model.

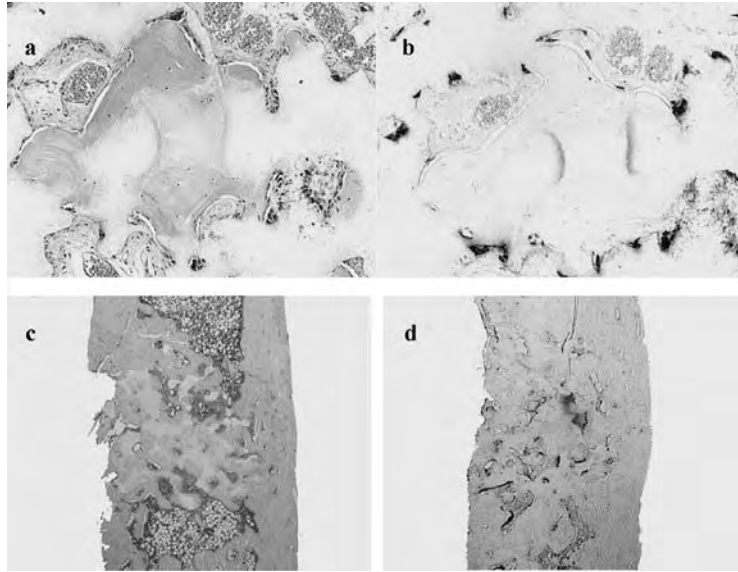


Fig. 4 HE (a,c) and TRAP (b,d) staining of subcutaneously implanted (a,b) and orthotopically implanted (c,d) β -TCP and BMO composites at eight weeks postimplantation.

scaffolds and the in-vitro culture method. Several kinds of porous ceramic scaffolds have been examined for this transplantation. For clinical usage, the ideal scaffold would be a porous ceramic material not only with good biocompatibility and osteointegrative properties but also with high mechanical strength and high biodegradability. Some ceramics have the former two properties; however, no porous scaffolds satisfy both of the latter two properties. We combined blocks of porous β -TCP with rat BMO to obtain TCP/BMO composites. The composites were cultured in osteogenic medium for two weeks and then implanted at subcutaneous sites in syngeneic rats. After implantation, the composites were harvested and prepared for biochemical, histological, and biomechanical analysis. We examined the methodology, the ability of bone formation, absorbability, and biomechanical properties of the TCP/BMO composite. We also evaluated this β -TCP porous material as a scaffold for bone tissue engineering.

Alkaline phosphatase activity and osteocalcin content were maintained at high levels for at least 16 weeks. In the light microscopic analysis, decalcified sections stained with HE showed immature bone had formed in some pores of the implant at two weeks postimplantation. At four weeks, areas of mature bone existed in the pores. At eight weeks, the amount of mature bone had increased in the pores. Fig. 4 shows the HE-stained and TRAP-staining of subcutaneously implanted composites and orthotopically implanted composites at eight weeks postimplantation. Blood vessel formation was observed, which supported bone formation. Rich bone formation was observed in the pores. At the same time, TRAP-positive osteoclasts were observed, which suggested a rapid remodeling is under way. These results also suggest that the β -TCP

porous materials are good candidates for scaffolds in bone tissue engineering. The most serious problem for clinical application is the mechanical weakness of β -TCP porous materials with 2 MPa of compressive strength. Many orthopedic surgeons have doubts over clinical applications due to the rapid degradation of β -TCP in vivo. In some reports, it is suggested that bone formation cannot follow the degradation of β -TCP. However, this difficulty proved to be overcome by tissue engineering treatment of the β -TCP materials. The compressive strength of β -TCP/BMO before implantation was weak at 2.25 MPa; however, it had improved to 4.92 MPa at 12 weeks postimplantation. Moreover, it reached 11.2 MPa at 24 weeks postimplantation. These results suggested that tissue engineering treatment using porous β -TCP increased the mechanical strength of the composites in vivo.^[90]

ARTICLES OF FURTHER INTEREST

Bioactive Glass; Biphasic Calcium Phosphate (BCP) Bioceramics: Chemical, Physical, and Biological Properties; Bone-Implant Interface; Bone Remodeling; Dental Implants; Osteogenic Progenitor Cells of Bone; Tissue Engineering of Bone

REFERENCES

1. Williams, D.F. *The Williams Dictionary of Biomaterials*; Liverpool University Press, 1999; 234.
2. Urist, M.R. Bone formation by autoinduction. *Science* **1965**, *150*, 893-899.
3. Celeste, A.J.; Song, J.J.; Cox, K.A.; Rosen, V.; Wozney, J.M. Bone morphogenetic protein-9, a new member of

- TGF-beta superfamily. *J. Bone Miner. Res.* **1994**, *9*, S137.
4. D'Alessandro, I.S.; Cox, K.A.; Israel, D.I.; LaPan, P.; Moutsatsos, I.K.; Nove, J.; Rosen, V.; Ryan, M.C.; Wozney, J.M.; Wang, E.A. Purification, characterization and activities of recombinant BMP-5. *J. Bone Miner. Res.* **1991**, *6*, S153.
 5. Gitelman, S.E.; Kobrin, M.S.; Ye, J.Q.; Lopez, A.R.; Lee, A.; Derynck, R. Recombinant Vgr-1/BMP-6-expressing tumors induce fibrosis and endochondral bone formation in vivo. *J. Cell Biol.* **1994**, *126* (6), 1595–1609.
 6. Wozney, J.M.; Rosen, V.; Celeste, A.J.; Miotsock, L.M.; Whitters, M.J.; Kriz, R.W.; Hewick, R.M.; Wang, E.A. Novel regulators of bone formation: Molecular clones and activities. *Science* **1988**, *242* (4885), 1528–1534. 16.
 7. Rosen, V.; Wozney, J.M. Bone Morphogenetic Protein. In *Principles of Bone Biology*, 2nd Ed.; Bilezikian, J.P., Raisz, L.G., Rodan, G.A., Eds.; Academic Press: USA, 2002; 919–928.
 8. Miyazono, K. Bone Morphogenetic Protein Receptors and Actions. In *Principles of Bone Biology*, 2nd Ed.; Bilezikian, J.P., Raisz, L.G., Rodan, G.A., Eds.; Academic Press: USA, 2002; 929–942.
 9. Rosen, V.; Wozney, J.M. Bone Morphogenetic Protein and the Adult Skeleton. In *Skeletal Growth Factors*; Canalis, E., Ed.; Lippincott Williams and Wilkins: Philadelphia, 2000; 299–310.
 10. Boden, S.D.; Moskovitz, P.A.; Morone, M.A.; Toribitake, Y. Video assisted lateral intertransverse process arthrodesis. Validation of a new minimally invasive lumbar spinal fusion technique in the rabbit and nonhuman primate models. *Spine* **1996**, *21*, 2689–2697.
 11. David, S.M.; Gruber, H.E.; Meyer, R.A.I.; Murakami, T.; Tabor, O.B.; Howard, B.A.; Wozney, J.M.; Hanley, E.N.J. Lambur spinal fusion using recombinant human bone morphogenetic protein in the canine. A comparison of three dosages and the carriers. *Spine* **1999**, *24*, 1971–1979.
 12. Sandhu, H.S.; Kanim, L.E.; Kabo, J.M.; Toth, J.M.; Zeegan, E.N.; Liu, D.; Seeger, L.L.; Dawson, E.G. Evaluation of rhBMP-2 with an OPLA carrier in a canine posterolateral (transverse process) spinal fusion model. *Spine* **1995**, *20* (24), 2669–2682. 15.
 13. Sandhu, H.S.; Kanim, L.E.; Kabo, J.M.; Toth, J.M.; Zeegen, E.N.; Liu, D.; Delamarter, R.B.; Dawson, E.G. Effective doses of recombinant human bone morphogenetic protein-2 in experimental spinal fusion. *Spine* **1996**, *21* (18), 2115–2122. 15.
 14. Bouxsein, M.L.; Turek, T.I.; Blake, C.D.; D'Augusta, D.; Seeherman, H.; Wozney, J. rhBMP-2 accelerate healing in a rabbit ulnar osteotomy model. *Trans. Orthop. Res. Soc.* **1999**, *24*, 138.
 15. Popich, L.S.; Salkeld, S.L.; Rueger, D.C.; Tucker, M.; Cook, S.D. Critical and non-critical size defect healing with OP-1. *Trans. Orthop. Res. Soc.* **1997**, *22*, 600.
 16. Turek, T.J.; Bostrom, M.P.G.; Camacho, N.; Blake, C.A.; Palmer, R.; Seeherman, H.J.; Wozney, J.M. Acceleration of bone healing in a rabbit ulnar osteotomy model with rhBMP-2. *Trans. Orthop. Res. Soc.* **1997**, *22*, 526.
 17. Welch, R.D.; Jones, A.L.; Bucholz, R.W.; Reinert, C.M.; Tjia, J.S.; Pierce, W.A.; Wozney, J.M.; Li, X.J. Effect of recombinant human bone morphogenetic protein-2 on fracture healing in a goat tibial fracture model. *J. Bone Miner. Res.* **1998**, *13* (9), 1483–1490.
 18. Anneroth, G.; Bang, G. The effect of allogenic demineralized dentin as a pulp capping agent in Java monkeys. *Odontol. Revy* **1972**, *23*, 315–328.
 19. Nakashima, M. Dentin induction by implants of autolyzed antigen-extracted allogenic (AAA) dentin on amputated dental pulps of dogs. *Endod. Dent. Traumatol.* **1972**, *5*, 279–286.
 20. Tziafas, D.; Kolokuris, I. Inductive influences of demineralized dentin and bone matrix on pulp cells. *Dent. Res.* **1990**, *69*, 75–81.
 21. Tziafas, D.; Kolokuris, I.; Alvanou, A.; Kaidoglou, K. Short term dentinogenic response of dog dental pulp tissue after its induction by demineralized or native dentine or pre-dentine. *Arch. Oral Biol.* **1992**, *37*, 119–128.
 22. Nakashima, M. Induction of dentin formation on canine amputated pulp by recombinant human bone morphogenetic protein (BMP)-2 and -4. *J. Dent. Res.* **1994**, *73*, 1515–1522.
 23. Rutherford, R.B.; Wahle, J.; Tucker, M.; Rueger, D.; Charette, M. Induction of reparative dentin formation in monkeys by recombinant human osteogenic protein-1. *Arch. Oral Biol.* **1993**, *38*, 571–576.
 24. Rutherford, R.B.; Spangberg, L.; Tucker, M.; Rueger, D.; Charette, M. Time course of the induction of reparative dentin formation in monkeys by recombinant human osteogenic protein-1. *Arch. Oral Biol.* **1994**, *39*, 833–838.
 25. Alden, T.D.; Pittman, D.D.; Beres, E.J.; Hankins, G.R.; Kallmes, D.F.; Wisotsky, B.M.; Kerns, K.M.; Helm, G.A. Percutaneous spinal fusion using bone morphogenetic protein-2 gene therapy. *J. Neurosurg.* **1999**, *90* (1 Suppl), 109–114.
 26. Baltzer, A.W.; Lattermann, C.; Whalen, J.D.; Wooley, P.; Weiss, K.; Grimm, M.; Ghivizzani, S.C.; Robbins, P.D.; Evans, C.H. Genetic enhancement of fracture repair: Healing of an experimental segmental defect by adenoviral transfer of the BMP-2 gene. *Gene Ther.* **2000**, *7* (9), 734–739.
 27. Franceschi, R.T.; Wang, D.; Krebsbach, P.H.; Rutherford, R.B. Gene therapy for bone formation: In vitro and in vivo osteogenic activity of an adenovirus expressing BMP7. *J. Cell. Biochem.* **2000**, *78* (3), 476–486. 6.
 28. Helm, G.A.; Alden, T.D.; Beres, E.J.; Hudson, S.B.; Das, S.; Engh, J.A.; Pittman, D.D.; Kerns, K.M.; Kallmes, D.F. Use of bone morphogenetic protein-9 gene therapy to induce spinal arthrodesis in the rodent. *J. Neurosurg.* **2000**, *92* (2 Suppl), 191–196.
 29. Musgrave, D.S.; Bosch, P.; Ghivizzani, S.; Robbins, P.D.; Evans, C.H.; Huard, J. Adenovirus-mediated direct gene therapy with bone morphogenetic protein-2 produces bone. *Bone* **1999**, *24* (6), 541–547.
 30. Sodek, J.; Zhang, Q.; Goldberg, H.A.; Domenicucci, C.; Kasugai, S.; Wrana, J.L.; Shapiro, H.; Chen, J.



- Non-collagenous Bone Proteins and Their Role in Substrate-Induced Bioactivity. In *The Bone-Biomaterial Interface*; Davies, J.E., Ed.; University of Toronto Press, 1991; 97–110.
31. Yabe, T.; Nemoto, A.; Uemura, T. Recognition of osteopontin by rat bone marrow derived osteoblastic primary cells. *Biosci. Biotechnol. Biochem.* **1997**, *61*, 754–756.
 32. Uemura, T.; Liu, Y.-K.; Feng, Y.; Nemoto, A.; Yabe, T.; Ushida, T.; Miyamoto, H.; Tateishi, T. The role of sialoprotein in recognition of bone surface by osteoblasts via integrin. *Mater. Sci. Eng.* **1997**, *C4*, 303–309.
 33. Nemoto, A.; Uemura, T. Integrin signal transduction via FAK in osteoblasts. *Seikagaku* **1999**, *71*, 140–145.
 34. Uemura, T.; Nemoto, A.; Deng, H.-Y. Differentiation and function of bone cells induced by bone matrix proteins. *Tanpakushitsu Kakusan Koso* **1999**, *44*, 143–148.
 35. Liu, Y.-K.; Uemura, T.; Nemoto, A.; Yabe, T.; Fujii, N.; Ushida, T.; Tateishi, T. Osteopontin involvement in integrin-mediated cell signaling and regulation of expression of alkaline phosphatase during early differentiation of UMR cells. *FEBS Lett.* **1997**, *420*, 112–116.
 36. Uemura, T.; Nemoto, A.; Liu, Y.-K.; Kojima, H.; Dong, J.; Yabe, T.; Yoshikawa, T.; Ohgushi, H.; Ushida, T.; Tateishi, T. Osteopontin involvement in bone remodeling and its effects on in vivo osteogenic potential of bone marrow-derived osteoblasts/porous hydroxyapatite constructs. *Mater. Sci. Eng.* **2001**, *C17*, 33–36.
 37. Uemura, T.; Nemoto, A.; An, J.; Liu, Y.-K.; Yoshikawa, T.; Ohgushi, H.; Kuboki, Y.; Ushida, T.; Tateishi, T. Osteopontin involvement in osteoblast differentiation and its effect on in vivo osteogenic potential of bone marrow derived osteoblasts. *Bio-ceramics* **1999**, *6*, 127–129.
 38. Buttler, W.T.; Ridall, A.L.; McKee, M.D. Osteopontin. In *Principles of Bone Biology*; Bilezikian, J.P., Raisz, L.G., Rodan, G.A., Eds.; Academic Press, 1996; 167–181.
 39. Butler, W.T. Dentin matrix proteins and dentinogenesis. *Connect. Tissue Res.* **1995**, *33*, 59–65.
 40. Boskey, A.L. Mineral-matrix interactions in bone and cartilage. *Clin. Orthop.* **1992**, *281*, 244–274.
 41. Boskey, A.L.; Maresca, M.; Doty, S.; Sabsay, B.; Veis, A. Concentration-dependent effects of dentin phosphophoryn in the regulation of in vitro hydroxyapatite formation and growth. *Bone Miner.* **1990**, *11*, 55–65.
 42. Campbell, A.A.; Ebrahimpour, A.; Perez, L.; Smesko, S.A.; Nancollas, G.H. The dual role of polyelectrolytes and proteins as mineralization promoters and inhibitors of calcium oxalate monohydrate. *Calcif. Tissue Int.* **1989**, *45*, 122–128.
 43. Weinstock, M.; Leblond, C.P. Radioautographic visualization of the deposition of a phosphoprotein at the mineralization front in the dentin of rat incisor. *J. Cell Biol.* **1973**, *56*, 838–845.
 44. Kuboki, Y.; Fujisawa, R.; Aoyama, K.; Sasaki, S. Calcium-specific precipitation of dentin phosphoprotein: A new method of purification and the significance for the mechanism of calcification. *J. Dent. Res.* **1979**, *58*, 1926–1932.
 45. Marsh, M.E. Self-association of calcium and magnesium complexes of dentin phosphophoryn. *Biochemistry* **1989**, *28*, 339–345.
 46. Butler, W.T. Dentin Collagen: Chemical Structure and Role in Mineralization. In *Dentin and Dentinogenesis*; Linde, A., Ed.; CRC Press: Boca Raton, 1984; 37–54.
 47. Stetler-Stevenson, W.G.; Veis, A. Type I collagen shows a specific binding affinity for bovine dentin phosphophoryn. *Calcif. Tissue Int.* **1986**, *38*, 135–141.
 48. Traub, W.; Jodaikin, A.; Arad, T.; Veis, A.; Sabsay, B. Dentin phosphophoryn binding to collagen fibrils. *Matrix* **1992**, *12*, 197–201.
 49. Fujisawa, R.; Kuboki, Y. Affinity of bone sialoprotein and several other bone and dentin acidic proteins to collagen fibrils. *Calcif. Tissue Int.* **1992**, *51*, 438–442.
 50. Saito, T.; Arsenault, A.L.; Yamauchi, M.; Kuboki, Y.; Crenshaw, M.A. Mineral induction by immobilized phosphoproteins. *Bone* **1997**, *21*, 305–311.
 51. Saito, T.; Yamauchi, M.; Crenshaw, M.A. Apatite induction by insoluble dentin collagen. *J. Bone Miner. Res.* **1998**, *13*, 265–270.
 52. Saito, T.; Yamauchi, M.; Abiko, Y.; Matsuda, K.; Crenshaw, M.A. In vitro apatite induction by phosphophoryn immobilized to collagen fibrils. *J. Bone Miner. Res.* **2000**, *15*, 1615–1619.
 53. Maniopoulos, C.; Sodek, J.; Melcher, A.H. Bone formation in vitro by stromal cells obtained from marrow of young adult rats. *Cell Tissue Res.* **1988**, *254*, 317–330.
 54. Uemura, T.; Dong, J.; Wang, Y.C.; Kojima, H.; Saito, T.; Iejima, D.; Kikuchi, M.; Tanaka, J.; Tateishi, T. Transplantation of cultured bone cells using combinations of scaffolds and culture techniques. *Biomaterials* **2003**, *24*, 2277–2286.
 55. Hench, L.L.; Splinter, R.J.; Allen, W.C.; Greenlee, T.K. Bonding mechanisms at the interface of ceramic prosthetic materials. *J. Biomed. Mater. Res. Symp.* **1971**, *2*, 117–141.
 56. Wilson, J.; Yli-Urpo, A.; Risto-Pekka, H. Bioactive Glasses: Clinical Applications. In *An Introduction to Bioceramics*; Hench, L.L., Wilson, J., Eds.; World Scientific: Singapore, 1993; 63–74.
 57. Kokubo, T.; Kim, H.M.; Kawashita, M. Novel bioactive materials with different mechanical properties. *Biomaterials* **2003**, *24* (13), 2161–2175.
 58. Glowacki, J.; Altobelli, D.; Mulliken, J.B. Fate of mineralized and demineralized osseous implants in cranial defects. *Calcif. Tissue Int.* **1981**, *33* (1), 71–76.
 59. Ray, R.D.; Holloway, J.A. Bone implants. Preliminary results of an experimental study. *J. Bone Jt. Surg.* **1957**, *39A*, 1119–1128.
 60. Mulliken, J.B.; Glowacki, J. Induced osteogenesis for repair and construction in the craniofacial region. *Plast. Reconstr. Surg.* **1980**, *65*, 553–559.
 61. Reddi, A.H.; Huggins, C. Biochemical sequences in the transformation of normal fibroblasts in adolescent rats. *Proc. Natl. Acad. Sci. U. S. A.* **Jun. 1972**, *69* (6), 1601–1605.
 62. Gepstein, R.; Weiss, R.E.; Hallel, T. Bridging large defects in bone by demineralized bone matrix in the

- form of a powder. A radiographic, histological, and radioisotope-uptake study in rats. *J. Bone Jt. Surg., Am.* **1987**, *69* (7), 984–992.
63. Glowacki, J.; Kaban, L.B.; Murray, J.E.; Folkman, J.; Mulliken, J.B. Application of the biological principle of induced osteogenesis for craniofacial defects. *Lancet* **1981**, *1* (8227), 959–962. 2.
 64. Kaban, L.B.; Mulliken, J.B.; Glowacki, J. Treatment of jaw defects with demineralized bone implants. *J. Oral Maxillofac. Surg.* **1982**, *40* (10), 623–626.
 65. Sonis, S.T.; Kaban, L.B.; Glowacki, J. Clinical trial of demineralized bone powder in the treatment of periodontal defects. *J. Oral Med.* **1983**, *38* (3), 117–122.
 66. Upton, J.; Boyajian, M.; Mulliken, J.B.; Glowacki, J. The use of demineralized xenogeneic bone implants to correct phalangeal defects: A case report. *J. Hand Surg., Am.* **1984**, *9* (3), 388–391.
 67. Upton, J.; Glowacki, J. Hand reconstruction with allograft demineralized bone: Twenty-six implants in twelve patients. *J. Hand Surg., Am.* **1992**, *17* (4), 704–713.
 68. Neo, M.; Kotani, S.; Fujita, Y.; Nakamura, T.; Yamamuro, T.; Bando, Y.; Ohtsuki, C.; Kokubo, T. Differences in ceramic-bone interface between surface-active ceramics and resorbable ceramics: A study by scanning and transmission electron microscopy. *J. Biomed. Mater. Res.* **1992**, *26*, 255–267.
 69. Kitsugi, T.; Yamamuro, T.; Nakamura, T.; Kotani, S.; Kokubo, T.; Takeuchi, H. Four calcium phosphate ceramics as bone substitutes for non-weight-bearing. *Biomaterials* **1993**, *14*, 216–224.
 70. Yang, Z.; Yuan, H.; Tong, W.; Zou, P.; Chen, W.; Zhang, X. Osteogenesis in extraskelally implanted porous calcium phosphate ceramics: Variability among different kinds of animals. *Biomaterials* **1996**, *17*, 2131–2137.
 71. Ohura, K.; Bohner, M.; Hardouin, P.; Lemaître, J.; Pasquier, G.; Flautre, B. Resorption of, and bone formation from, new β -tricalcium phosphate-monocalcium phosphate cements: An in vivo study. *J. Biomed. Mater. Res.* **1996**, *30*, 193–200.
 72. Ayers, R.A.; Simske, S.J.; Nunes, C.R.; Wolford, L.M. Long-term bone ingrowth and residual microhardness of porous block hydroxyapatite implants in humans. *J. Oral Maxillofac. Surg.* **1998**, *56*, 1297–1301.
 73. Tancred, D.C.; McCormack, B.A.; Carr, A.J. A synthetic bone implant macroscopically identical to cancellous bone. *Biomaterials* **1998**, *19*, 2303–2311.
 74. Furusawa, T.; Mizunuma, K.; Yamashita, S.; Takahashi, T. Investigation of early bone formation using resorbable bioactive glass in the rat mandible. *Int. J. Oral Maxillofac. Implants* **1998**, *13*, 672–676.
 75. Ohsawa, K.; Neo, M.; Matsuoka, H.; Akiyama, H.; Ito, H.; Kohno, H.; Nakamura, T. The expression of bone matrix protein mRNAs around beta-TCP particles implanted into bone. *J. Biomed. Mater. Res.* **2000**, *52*, 460–466.
 76. Merten, H.A.; Wiltfang, J.; Grohmann, U.; Hoenig, J.F. Intraindividual comparative animal study of alpha- and β -tricalcium phosphate degradation in conjunction with simultaneous insertion of dental implants. *J. Craniofac. Surg.* **2001**, *12*, 59–68.
 77. Kurashina, K.; Kurita, H.; Wu, Q.; Ohtsuka, A.; Kobayashi, H. Ectopic osteogenesis with biphasic ceramics of hydroxyapatite and tricalcium phosphate in rabbits. *Biomaterial* **2002**, *2* (23), 407–412.
 78. Grenga, T.E.; Zins, J.E.; Bauer, T.W. The rate of vascularization of coralline hydroxyapatite. *Plast. Reconstr. Surg.* **1989**, *84*, 245–249.
 79. Schliephake, H.; Neukam, F.W.; Klosa, D. Influence of pore dimensions on bone ingrowth into porous hydroxyapatite blocks used as bone graft substitutes. An isometric study. *Int. J. Oral Maxillofac. Surg.* **1991**, *20*, 53–58.
 80. Tancred, D.C.; McCormack, B.A.; Carr, A.J. A synthetic bone implant macroscopically identical to cancellous bone. *Biomaterials* **1998**, *19*, 2303–2311.
 81. Saito, M.; Shimizu, H.; Beppu, M.; Takagi, M. The role of β -tricalcium phosphate in vascularized periosteum. *J. Orthop. Sci.* **2000**, *5*, 275–282.
 82. Schliephake, H.; Kage, T. Enhancement of bone regeneration using resorbable ceramics and a polymer-ceramic composite material. *J. Biomed. Mater. Res.* **2001**, *56*, 128–136.
 83. Nicholas, R.W.; Lange, T.A. Granular tricalcium phosphate grafting of cavitary lesions in human bone. *Clin. Orthop.* **1994**, *306*, 197–203.
 84. Nery, E.B.; LeGeros, R.Z.; Lynch, K.L.; Lee, K. Tissue response to biphasic calcium phosphate ceramic with different ratios of HA/ β TCP in periodontal osseous defects. *J. Periodontol.* **1992**, *63*, 729–735.
 85. Johnson, K.D.; Frierson, K.E.; Keller, T.S.; Cook, C.; Scheinberg, R.; Zerwekh, J.; Meyers, L.; Sciadini, M.F. Porous ceramics as bone graft substitutes in long bone defects: A biomechanical, histological, and radiographic analysis. *J. Orthop. Res.* **1996**, *14*, 351–369.
 86. Krebsbach, P.H.; Kuznetsov, S.A.; Satomura, K.; Emmons, R.V.; Rowe, D.W.; Robey, P.G. Bone formation in vivo: Comparison of osteogenesis by transplanted mouse and human marrow stromal fibroblasts. *Transplantation* **1997**, *63*, 1059–1069.
 87. Wang, J.; Chen, W.; Li, Y.; Fan, S.; Weng, J.; Zhang, X. Biological evaluation of biphasic calcium phosphate ceramic vertebral laminae. *Biomaterials* **1998**, *19*, 1387–1392.
 88. Alam, I.; Asahina, I.; Ohmamiuda, K.; Enomoto, S. Comparative study of biphasic calcium phosphate ceramics impregnated with rhBMP-2 as bone substitutes. *J. Biomed. Mater. Res.* **2001**, *54*, 129–138.
 89. Yokoyama, N.; Minabe, M.; Sugaya, A.; Hori, T.; Otaki, K.; Hasegawa, A.; Irie, H.; Hakamatsuka, Y. Application of β -tricalcium phosphate to periodontal therapy. Part 1. Fundamental studies and biological tests of β -TCP prepared by mechanochemical method. *Nippon Shishubyo Gakkai Kaishi* **1989**, *31*, 213–223.
 90. Dong, J.; Uemura, T.; Shirasaki, Y.; Tateishi, T. Promotion of bone formation using highly pure porous beta-TCP combined with bone marrow-derived osteoprogenitor cells. *Biomaterials* **2002**, *23* (23), 4493–4502.



Particulate Matter and Host Reactions

Edward M. Greenfield

Department of Orthopedics, Case Western Reserve University, Cleveland, Ohio, U.S.A.

INTRODUCTION

Wear particles from orthopaedic implants stimulate a local inflammatory process that induces proresorptive cytokine production, osteoclast differentiation, bone resorption, osteolysis, implant loosening, pain, loss of implant function, and, ultimately, the need for revision surgery. This review will focus on four aspects of the host reaction to the particles: first, the role of particles themselves compared with other factors, including implant motion, mechanical forces, fluid pressure, and bacterial endotoxins; second, the ability of macrophages and other cell types to detect the particles; third, the actions of osteoclasts and osteoblasts as effector cells; and fourth, future directions for research in this area. The 12 major conclusions of this review are that: (i) wear particles are the primary inducers of aseptic loosening; (ii) the effects of the wear particles are increased by other factors, including implant motion, mechanical forces, fluid pressure, and bacterial endotoxins; (iii) the most important type of wear particle in an individual patient is likely to be the type that is the most abundant; (iv) the macrophage is the primary cell type that detects and responds to the wear particles; (v) the wear particles activate multiple signal transduction pathways in the macrophages; (vi) the wear particles induce macrophages to produce proinflammatory cytokines [tumor necrosis factor- α (TNF- α), interleukin (IL)-1, IL-6, etc] that synergistically induce osteoclast differentiation and osteolysis; (vii) the primary role of fibroblasts and other mesenchymal cells in aseptic loosening is to produce proresorptive cytokines [RANKL and macrophage colony stimulating factor (M-CSF)] in response to the proinflammatory cytokines produced by the macrophages; (viii) lymphocytes likely do not play a major role in aseptic loosening except, perhaps, in a subset of patients with metal sensitivity; (ix) increased osteoclast differentiation is the primary cause of osteolysis; (x) decreased bone formation plays an important, but secondary, role in aseptic loosening; (xi) a better understanding of the genetic regulation of aseptic loosening may lead to better methods for identifying higher risk patients; and (xii) approaches to accelerate osseointegration or to inhibit bacterial growth, wear particle migration, or osteoclast differentiation are promising future directions that deserve further study.

BACKGROUND

Wear particles are the primary cause of aseptic loosening of orthopaedic implants.^[1,2] Aseptic loosening, a major problem in clinical orthopaedics, is defined as loosening, in the absence of clinical or microbiological evidence of infection, owing to osteolysis or local bone loss. The overall process is summarized in Fig. 1. Motion between the implant components generates metallic and polymeric wear particles. The most abundant type of particle in most patients is ultrahigh molecular weight polyethylene (UHMWPE) derived from the bearing surfaces. Although the rate of generation of these particles will likely be reduced by the recent development of cross-linked UHMWPE, aseptic loosening will likely remain a major problem because cross-linking does not eliminate the production of wear particles^[3] and, in fact, leads to the production of smaller, more biologically active, particles.^[4] Moreover, many other sources of wear particles exist, including third-body wear, impingement, backside wear, polymethylmethacrylate bone cement, hydroxyapatite coatings, corrosion, fretting, misalignment, and Morris tapers.^[1]

The wear particles stimulate a local inflammatory process that induces proresorptive cytokine production, osteoclast differentiation, bone resorption, osteolysis, implant loosening, pain loss of implant function, and, ultimately, the need for revision surgery (Fig. 1). This review will focus on four aspects of the host reaction to the particles: first, the role of particles themselves compared with other factors, including implant motion, mechanical forces, fluid pressure, and bacterial endotoxins; second, the ability of macrophages and other cell types to detect the particles; third, the actions of osteoclasts and osteoblasts as effector cells; and fourth, future directions. For details on other aspects of the process, readers are referred to other recent reviews.^[1,2,5,6]

THE ROLE OF PARTICLES AND OTHER FACTORS

Many types of studies have led to the consensus that wear particles that are phagocytosable (i.e., $<10\mu\text{m}$ in size) initiate aseptic loosening of orthopaedic

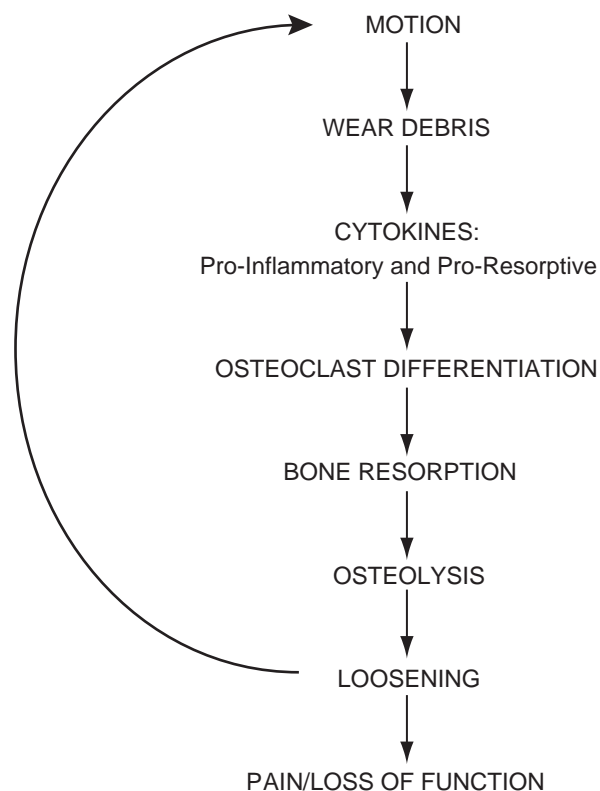


Fig. 1 Overview of the aseptic loosening process.

implants.^[1,2] Moreover, the rate of UHMWPE wear correlates with the risk of aseptic loosening.^[7,8] Despite the importance of wear particles in stimulating implant loosening, it is likely that other factors also contribute to the process.^[6] For example, a number of studies have shown that implant motion, mechanical forces, fluid pressure, and bacterial endotoxins can increase the biological responses induced by wear particles.^[6] We have been especially interested in the potential role of bacterial endotoxins in aseptic loosening.^[6] Although this might seem to be a contradiction in terms, aseptic loosening is defined as osteolysis in the absence of clinical or microbiological evidence of infection, which does not exclude a role for subclinical levels of infection or for molecules derived from bacteria located in other sites in the body. The three most likely sources of bacterial endotoxins during aseptic loosening are the implants themselves; systemic endotoxin derived from gut flora, minor infections, or dental procedures; and the bacterial biofilm found on many implants retrieved from patients with aseptic loosening. Of these three possibilities, the bacterial biofilm is likely to be quantitatively most important.^[6]

Although the most abundant type of particle in most patients with aseptic loosening is UHMWPE, numerous particles can also be generated from the polymethylmethacrylate bone cement and the metallic components of the implants (Fig. 2). Multiple studies

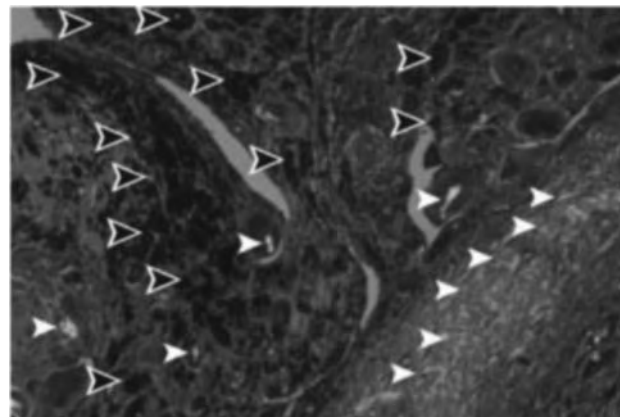


Fig. 2 Hematoxylin and eosin-stained periprosthetic tissue from a loose knee implant. *Black arrows* indicate metal wear particles; *white arrows* indicate UHMWPE wear particles visualized using polarized light microscopy. (Courtesy of Dr. James Anderson, Department of Pathology, Case Western Reserve University.)

have revealed conflicting results on the relative potency of the particles of the different materials (Ref. [9] and references therein) and we have recently shown that titanium and UHMWPE particles induce osteolysis by indistinguishable mechanisms.^[10] Important but hard to control factors such as particle size, shape, and the adherence of biologically active molecules make these studies difficult to interpret. Nonetheless, taken together, the studies suggest that the most important type of wear particle in an individual patient is likely to be the type that is the most abundant in that patient. Consistent with this, the most important effect of oxidation of UHMWPE appears to be the stimulation of particle generation rather than the stimulation of biological activity of the particles (Ref. [10] and references therein).

MACROPHAGES AND OTHER CELL TYPES AS PARTICLE DETECTORS

Abundant evidence indicates that macrophages are the primary cell type responsible for detecting the wear particles, and that the initial response to phagocytosis of the particles by macrophages is the increased production of proinflammatory cytokines.^[1,5,6] One active area of research examines the roles of the individual proinflammatory cytokines produced by macrophages in response to wear particles. As reviewed in Ref. [11], TNF- α is the proinflammatory cytokine with the best evidence for a role in aseptic loosening. In support of this conclusion, pharmacological or genetic inhibition of TNF- α signaling reduces wear particle-induced osteolysis in cell culture and animal models.^[11]

Moreover, a polymorphism in the TNF- α gene promoter is associated with an increased rate of aseptic loosening.^[12] IL-1 also likely plays a role as a polymorphism in the IL-1 receptor antagonist also is also associated with an increased rate of aseptic loosening.^[13] Moreover, we have found that neutralizing antibodies directed against either IL-1 α or IL-1 β inhibit wear particle-induced osteoclast differentiation by ~75% (Fig. 3A). In the same series of experiments, neutralizing antibodies directed against either TNF- α or IL-6 inhibited osteoclast differentiation by ~65% (Fig. 3A). Because the sum of the inhibitions induced by the four antibodies (75% + 75% + 65% + 65%) substantially exceeds 100%, it is most likely that the four cytokines act synergistically to induce osteoclast differentiation in response to wear particles. In agreement with this conclusion, we have also found that IL-1 β , IL-6, and TNF- α can synergistically stimulate osteoclast differentiation in the absence of wear particles (Fig. 3B).

Another area of research examines the macrophage intracellular signal transduction pathways that are activated by wear particles. This is undoubtedly an extremely complex and only partially understood response. Our current working model of signaling pathways activated by wear particles is depicted in Fig. 4. This working model indicates that wear particles bind to a complex of receptor molecules that includes toll-like receptors.^[6] The receptor complex, in turn, strongly activates many signal transduction pathways. The nuclear factor- κ B (NF- κ B), ERK1/2, and PI3K/Akt pathways all contribute to the production of TNF- α in response to wear particles.^[11,14] TNF- α , in turn, is one of the proinflammatory cytokines that stimulate osteoclast differentiation and the resultant osteolysis (reviewed in Ref. [11]). The p38 and c-Jun amino-terminal kinase (JNK) pathways are also activated by wear particles.^[11] Although these pathways do not appear to contribute to TNF- α production, they may be responsible for the production of other proinflammatory cytokines.^[11] All of these results are based on our studies of titanium particles with adherent endotoxin; “endotoxin-free” titanium particles are much less potent at inducing the downstream events such as the activation of all examined signal transduction pathways, the production of proinflammatory cytokines, osteoclast differentiation, and osteolysis.^[6,11,14] “Endotoxin-free” particles likely do not activate toll-like receptors because they are the primary receptors used by the host immune system to detect bacteria and other pathogens.

Cells besides macrophages may also play a secondary role in responding to wear particles. For example, recent studies have shown that mesenchymal cells at various stages of differentiation along the osteogenic pathway (fibroblasts, stromal cells, preosteoblasts, and

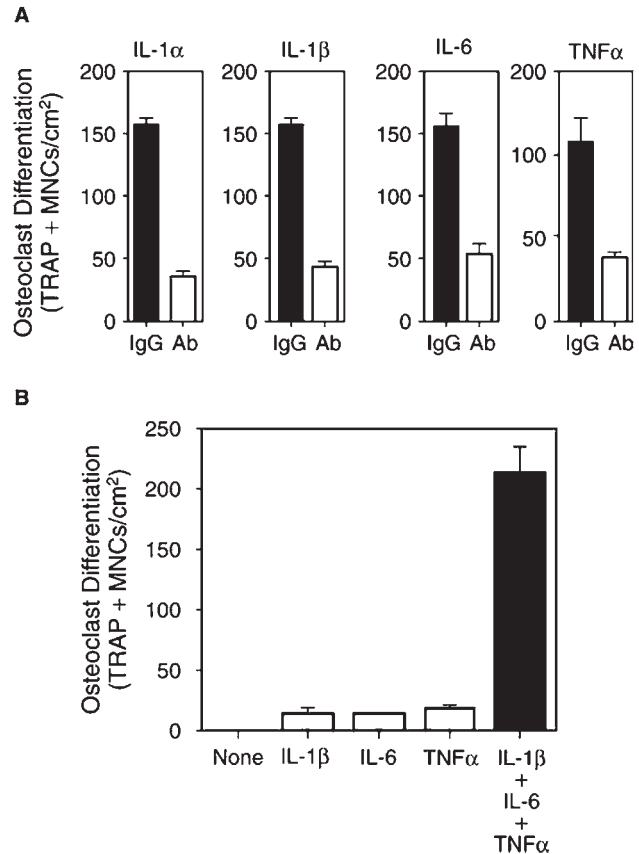


Fig. 3 Proinflammatory cytokines synergistically stimulate osteoclast differentiation in response to titanium particles (panel A) and when added directly to the cell cultures (panel B). In panel A, maximally effective concentrations of neutralizing antibodies to the indicated cytokines were added to osteoclast differentiation assays containing conditioned media from macrophages incubated with titanium particles (open bars). Control groups received equal amounts of isotype-matched IgG (filled bars). In panel B, submaximal amounts of IL-1 β , IL-6, and TNF- α were added either alone (open bars) or in combination (filled bars) to osteoclast differentiation assays in the absence of macrophage conditioned medium. Panel A is modified from Lowe, R.; Nalepka, J.; Van De Motter, R.; Goldberg, V.; Bi, Y.; Greenfield, E. Stimulation of osteoclast differentiation by wear debris depends on synergistic interactions between IL-1 α , IL-1 β , IL-6, and TNF- α . *Trans. Orthop. Res. Soc.* **2001**, *26*, 959. Panel B is modified from Ragab, A.A.; Nalepka, J.L.; Bi, Y.; Greenfield, E.M. Cytokines synergistically induce osteoclast differentiation: support by immortalized or normal calvarial cells. *Am. J. Physiol. Cell Physiol.* **2002**, *283*, C679–C687.

osteoblasts) can produce proinflammatory and pro-resorptive cytokines in response to wear particles (Ref. [15] and references therein). In most patients, however, it is likely that the production of resorptive cytokines in response to inflammatory cytokines produced by macrophages is quantitatively more important (Fig. 5).

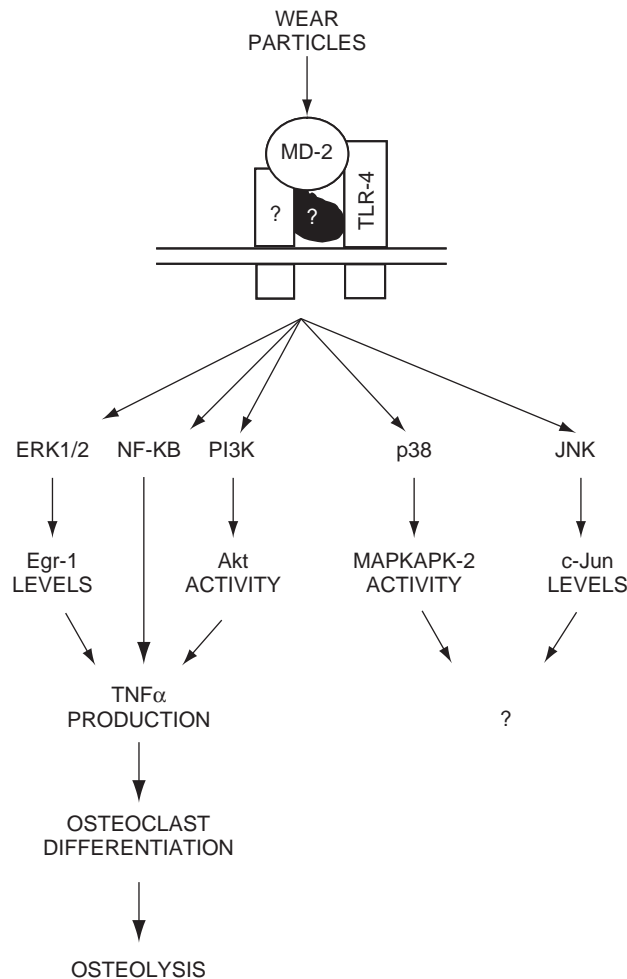


Fig. 4 Signal transduction pathways activated in macrophages by orthopaedic wear particles. (Modified from Ref. [11].)

This conclusion is based on two types of data. First, macrophages are the most abundant cell type surrounding most loose implants and, second, macrophages produce much higher levels of cytokines per cell than are produced by mesenchymal cells. Moreover, a histological, in situ hybridization, and immunohistochemical study by Jiranek et al.^[16] of the periprosthetic tissue surrounding aseptically loose implants showed that macrophages and foreign body giant cells were the only cell types that phagocytosed the wear debris. In addition, these monocyte-derived cells were the only cells that produced IL-1 β mRNA. In contrast, immunohistochemistry showed IL-1 β protein bound to both macrophages and fibroblasts, indicating that IL-1 β protein binds to receptors on both cell types but that macrophages are the primary producers of the proinflammatory cytokine.^[16] The primary role of fibroblasts and other mesenchymal cells would then be to produce proresorptive cytokines, such as RANKL and M-CSF, in response to the proinflammatory cytokines

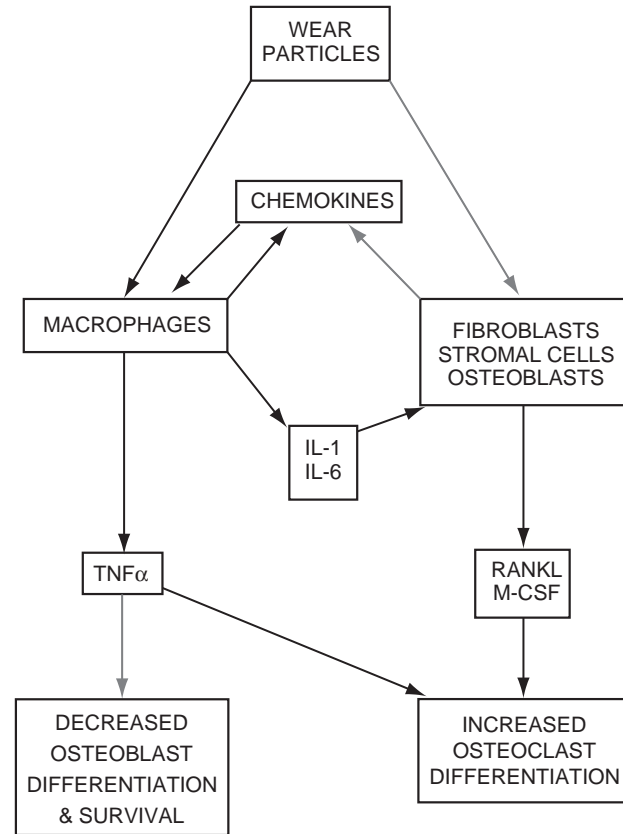


Fig. 5 Cellular responses to orthopaedic wear particles. *Black arrows* indicate primary responses; *gray arrows* indicate secondary responses.

produced by macrophages (Fig. 5). It should also be noted that TNF- α is an exception in that it primarily acts directly on osteoclast precursors to induce their differentiation into mature bone resorptive cells (Fig. 5).^[6] Fibroblasts may also be important in the production of chemokines that recruit macrophages to the periprosthetic environment (Fig. 5).^[2] As wear particles induce macrophages to produce abundant levels of chemokines that recruit additional macrophages (Fig. 5),^[5] the fibroblast-derived chemokines are likely to be quantitatively important only during the initial stage of host responses to wear particles, prior to substantial macrophage recruitment.

The lymphocyte is another cell type suggested to contribute to the responses to orthopaedic wear particles.^[17] In support of this possibility, lymphocytes from a subset of patients with aseptic loosening exhibit sensitization to metals.^[17] However, this presents the classical chicken-or-egg question: do the increased metal levels in patients with loosening contribute to the increased sensitization and/or does sensitization contribute to aseptic loosening? Only the latter possibility would indicate that lymphocytes contribute to aseptic loosening. Evidence against a role for lymphocytes was obtained in animal and cell culture studies that found



that lymphocytes do not detectably contribute to cytokine production or to osteolysis induced by wear particles (Ref. [10] and references therein). Moreover, the study by Jiranek et al.^[16] discussed above failed to show the phagocytosis of wear particles by lymphocytes in periprosthetic tissue. However, three studies have recently shown that the periprosthetic tissue from patients with aseptic loosening of metal-on-metal joint replacements contains high levels of T and B lymphocytes.^[18–20]

OSTEOCLASTS AND OSTEOBLASTS AS EFFECTOR CELLS

Osteoclasts are the cells that resorb bone in response to the local inflammatory process initiated by wear particles.^[5] So, osteoclasts are the primary effector cell type responsible for wear particle-derived osteolysis. Wear particles have little effect, however, on the activity or the survival of mature osteoclasts.^[5] Instead, the recruitment of osteoclast precursors to the periprosthetic environment and their subsequent differentiation into mature bone resorbing cells are both potently upregulated by the inflammatory response initiated by wear particles (Fig. 5 and left side of Fig. 6).^[5] The recruitment and differentiation of osteoclast precursors thus represent attractive targets for the pharmacological treatment of aseptic loosening.

Osteoblasts are the cells that form bone, and decreases in osteoblast number or activity induced by wear particles likely play a secondary role in implant loosening (right side of Fig. 6).^[6] The inhibition of osteoblast number and activity by wear particles appears to be much more potent *in vivo* than *in vitro*,^[6] indicating that this inhibition is likely owing to the inflammatory process induced by the wear particles rather than to direct effects on the osteoblasts themselves. TNF- α may be especially important in this process because it inhibits differentiation of osteoblasts and stimulates their apoptosis (Fig. 5).^[21] Although the decreased bone formation plays a secondary role in the pathogenesis of implant loosening, patients with high rates of bone formation have lower rates of aseptic loosening.^[7] In this regard, we have found that the high rates of bone formation in young mice lead to the rapid repair of osteolytic areas.^[6] Moreover, increased bone formation is responsible for the ability of osteogenic agents, such as growth factors and parathyroid hormone, to increase the osseointegration of orthopaedic implants.^[6,22] Osteogenic agents represent an especially attractive pharmacological approach, because accelerated osseointegration would allow patients to return more rapidly to activities of daily living and would lead to a lower incidence of aseptic loosening in subsequent years.^[6]

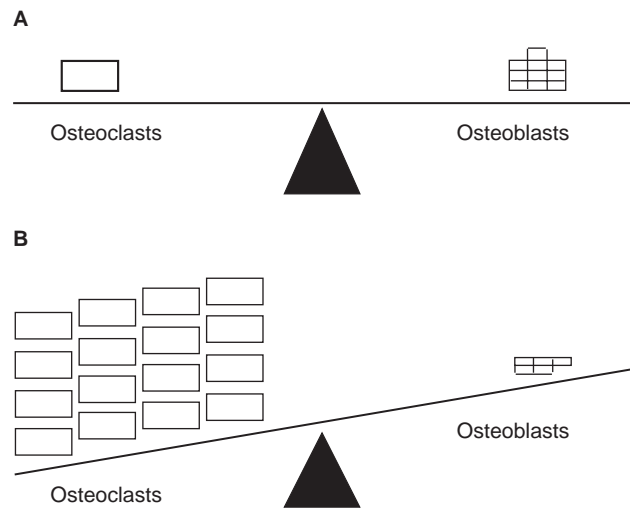


Fig. 6 Osteoclasts and osteoblasts as effector cells responsible for implant loosening. During normal homeostasis (A), the action of the two cell types are balanced with ~10 osteoblasts required to form new bone equivalent to that resorbed by one osteoclast. During implant loosening (B), bone resorption exceeds formation primarily because the number of osteoclasts increases 10–20-fold. The ~50% decrease in the number of osteoblasts is a smaller, secondary response. (From Ref. [5] and Kadoya, Y.; Revell, P.A.; Al-Saffer, N.; Kobayashi, A.; Scott, G.; Freeman, M.A.R. Bone formation and bone resorption in failed total joint arthroplasties: histomorphometric analysis with histochemical and immunohistochemical technique. *J. Orthop. Res.* **1996**, *14*, 473–482.)

FUTURE DIRECTIONS

The inhibition of the inflammatory process, either pharmacologically or by gene therapy, would likely reduce aseptic loosening. However, the long-term systemic inhibition of inflammation may lead to increased infections and other serious complications.^[23] As discussed above, the inhibition of osteoclast differentiation or the acceleration of osseointegration represents more attractive approaches, which would not necessarily require the local delivery of the pharmaceutical or gene construct. Moreover, these approaches would potentially have the additional benefit of reducing osteoporosis in the treated patients. An example is the potential use of bisphosphonates in patients with orthopaedic implants. By inhibiting osteoclastic bone resorption, bisphosphonates can increase the osseointegration of implants in animal models when administered either systemically^[24] or bound to the implants.^[25] Moreover, in patients, bisphosphonates can reduce the increase in local bone turnover and the resultant local bone loss that

occurs in the first few years after total hip arthroplasty.^[26,27]

As a more detailed understanding of the host response to orthopaedic wear particles becomes available, it may be possible to identify those patients who have high risk for aseptic loosening. As discussed above, patients with high rates of bone formation have lower rates of aseptic loosening.^[7] In addition, specific femur geometries correlate with altered rates of aseptic loosening.^[7] Although the genes involved are unknown at this time, it is likely that genetic polymorphisms contribute to both the high rates of bone formation and the femur geometries. Genetic polymorphisms also likely control the magnitude of each patient's inflammatory and resorptive responses. For example, as discussed above (in the section entitled Macrophages and Other Cell Types as Particle Detectors), polymorphisms in both the TNF- α and the IL-1 receptor antagonist genes correlate with altered rates of aseptic loosening.^[12,13] One way to appreciate the likely importance of genetics in controlling the susceptibility to aseptic loosening is to consider the role of genetics in osteoporosis: twin and family studies have shown that 60–80% of the variation in bone density between individuals is genetically determined, and many genes appear to contribute to this effect.^[28]

To the extent that bacterial molecules contribute to aseptic loosening,^[6] antibacterial agents may decrease the rate of aseptic loosening, in addition to decreasing implant infection. Data from the Norwegian Arthroplasty Registry show that both the intravenous administration of antibiotics for one to three days following implant surgery and the inclusion of antibiotics in the polymethylmethacrylate bone cement contribute to decreased aseptic loosening.^[29] The prolonged effects of these therapies likely result from favoring patient-derived cells in their “race for the implant surface” with bacteria.^[6] Covalently linking antibiotics to metal implant surfaces may provide an even longer-lived effect.^[30]

Finally, guided bone regeneration might reduce aseptic loosening.^[31] In this approach, particle migration from the joint space to the bone–implant interface is blocked with a membrane of expanded polytetrafluoroethylene. This membrane also guides bone formation to reinforce the seal and to continue to block particle migration. This approach was tested using a goat total hip replacement model in which the cobalt–chromium bearing surfaces were roughened prior to implantation to accelerate the wear of the UHMWPE.^[30] The polytetrafluoroethylene membrane substantially reduced aseptic loosening assessed radiographically and histologically 12 months after implantation.^[30] The long-term usefulness of this approach remains to be investigated.

CONCLUSIONS

The 12 major conclusions of this review are that: (i) wear particles are the primary inducers of aseptic loosening; (ii) the effects of the wear particles are increased by other factors, including implant motion, mechanical forces, fluid pressure, and bacterial endotoxins; (iii) the most important type of wear particle in an individual patient is likely to be the type that is the most abundant; (iv) the macrophage is the primary cell type that detects and responds to the wear particles; (v) the wear particles activate multiple signal transduction pathways in the macrophages; (vi) the wear particles induce macrophages to produce proinflammatory cytokines (TNF- α , IL-1, IL-6, etc) that synergistically induce osteoclast differentiation and osteolysis; (vii) the primary role of fibroblasts and other mesenchymal cells in aseptic loosening is to produce proresorptive cytokines [RANKL and macrophage colony stimulating factor (M-CSF)] in response to the proinflammatory cytokines produced by the macrophages; (viii) lymphocytes likely do not play a major role in aseptic loosening except, perhaps, in a subset of patients with metal sensitivity; (ix) increased osteoclast differentiation is the primary cause of osteolysis; (x) decreased bone formation plays an important, but secondary, role in aseptic loosening; (xi) a better understanding of the genetic regulation of aseptic loosening may lead to better methods for identifying higher-risk patients; and (xii) approaches to accelerate osseointegration or to inhibit bacterial growth, wear particle migration, or osteoclast differentiation are promising future directions that deserve further study.



ARTICLES OF FURTHER INTEREST

Biocompatibility Testing; Titanium and Its Alloys; Ultra-High Molecular Weight Polyethylene (UHMWPE); Wear Debris; Wear Debris from Joint Arthroplasties, Biological Effects Of; Wear Debris, Bone Resorption Animal Models

ACKNOWLEDGMENTS

I would like to thank all of the past and current members of my laboratory who have contributed to the studies reviewed in this article. Our work related to the subject of this review has been supported by NIH RO1 AR043769. I would also like to thank Valerie Schmedlen for her careful editing of the manuscript.

REFERENCES

1. Wright, T.M.; Goodman, S.B. Implant wear in total joint replacement. Rosemount, Ill: J. Am. Acad. Orthop. Surg. **2001**, 224.
2. Bauer, T. Particles and periimplant bone resorption. Clin. Orthop. Rel. Res. **2002**, 405, 138–143.
3. Dorr, L.D.; Wan, Z.; Shahrदार, C.; Sirianni, L.; Boutary, M.; Yun, A. Clinical performance of a Durasul highly cross-linked polyethylene acetabular liner for total hip arthroplasty at five years. J. Bone Joint Surg. **2005**, 87A, 1816–1821.
4. Ingram, J.; Stone, M.; Fisher, J.; Ingham, E. The influence of molecular weight, crosslinking and counterface roughness on TNF- α production by macrophages in response to ultra high molecular weight polyethylene particles. Biomaterials **2004**, 25, 3511–3522.
5. Greenfield, E.M.; Bi, Y.; Ragab, A.A.; Goldberg, V.M.; Van De Motter, R.R. The role of osteoclast differentiation in aseptic loosening. J. Orthop. Res. **2002**, 20, 1–8.
6. Greenfield, E.M.; Bi, Y.; Ragab, A.A.; Goldberg, V.M.; Nalepka, J.L.; Seabold, J.M. Does endotoxin contribute to aseptic loosening of orthopaedic implants? J. Biomed. Mater. Res. B. Appl. Biomater. **2005**, 72B, 179–185.
7. Kobayashi, S.; Takaoka, K.; Saito, N.; Hisa, K. Factors affecting aseptic failure of fixation after primary Charnley total hip arthroplasty. J. Bone Joint Surg. **1997**, 79A, 1618–1627.
8. Wilkinson, J.M.; Hamer, A.J.; Stockley, I.; Eastell, R. Polyethylene wear rate and osteolysis: critical threshold versus continuous dose-response relationship. J. Orthop. Res. **2005**, 23, 520–525.
9. Haynes, D.R.; Boyle, S.J.; Rogers, S.D.; Howie, D.W.; Vernon-Roberts, B. Variation in cytokines induced by particles from different prosthetic materials. Clin. Orthop. Rel. Res. **1998**, 352, 223–230.
10. Taki, N.; Tatro, J.M.; Nalepka, J.L.; Togawa, D.; Goldberg, V.M.; Rinnac, C.M.; Greenfield, E.M. Polyethylene and titanium particles induce osteolysis by similar, lymphocyte-independent, mechanisms. J. Orthop. Res. **2005**, 23, 376–383.
11. Beidelschies, M.; Islam, A.; Smith, M.; Goldberg, V.; Chen, X.; Greenfield, E. ERK1/2 activation by orthopaedic wear particles is strongly stimulated by adherent endotoxin and is required for stimulation of TNF α production. **2005**, submitted.
12. Wilkinson, J.M.; Wilson, A.; Stockley, I.; Scott, I.; Macdonald, D.; Hamer, A.; Duff, G.; Eastell, R. Variation in the TNF gene promoter and risk of osteolysis after total hip arthroplasty. J. Bone Miner. Res. **2003**, 18, 1995–2001.
13. Gordon, A.; Wilkinson, J.M.; Wilson, A.G.; Stockley, I.; MacDonald, D.; Eastell, R. Polymorphisms in the interleukin-one gene cluster and the risk of aseptic loosening after total hip arthroplasty. J. Bone Miner. Res. **2003**, 18 (S2), 326.
14. Smith, M.; Innocenti, M.; Lee, M.; Rohrer, J.; Sobieraj, M.; Goldberg, V.; Greenfield, E. The PI3K/Akt pathway mediates the biological activity of titanium particles. Trans. Orthop. Res. Soc. **2005**, 30, 1619.
15. Wei, X.; Zhang, X.; Zuscik, M.; Drissi, M.; Schwarz, E.; O'keefe, R. Fibroblasts express RANKL and support osteoclastogenesis in a COX-2-dependent manner after stimulation with titanium particles. J. Bone Miner. Res. **2005**, 20, 1136–1148.
16. Jiranek, W.A.; Machado, M.; Jasty, M.; Jevsevar, D.; Wofe, H.J.; Goldring, S.R.; Goldberg, M.J.; Harris, W.H. Production of cytokines around loosened cemented acetabular components. J. Bone Joint Surg. **1994**, 75A, 863–879.
17. Hallab, N.; Merritt, K.; Jacobs, J. Metal sensitivity in patients with orthopaedic implants. J. Bone Joint Surg. **2001**, 83A, 428–436.
18. Davies, A.; Willert, H.; Cambell, P.; Learmouth, I.; Case, C. An unusual lymphocytic perivascular infiltration in tissue around contemporary metal-on-metal joint replacements. J. Bone Joint Surg. **2005**, 87A, 18–27.
19. Park, Y.; Moon, Y.; Lim, S.; Yang, J.; Ahn, G.; Choi, Y. Early osteolysis following second-generation metal-on-metal hip replacement. J. Bone Joint Surg. **2005**, 87A, 1515–1521.
20. Willert, H.; Buchhorn, G.; Fayyazi, A.; Flury, R.; Windler, M.; Koster, G.; Lohmann, C. Metal-on-metal bearings and hypersensitivity in patients with artificial joints. J. Bone Joint Surg. **2005**, 87A, 28–36.
21. Nanes, M.; Pacifici, R. Inflammatory cytokines. In *Bone Resorption*; Bronner, F., Farach-Carson, M., Rubin, J., Eds.; London: Springer, 2005; 67–90.
22. Skripitz, R.; Bohling, S.; Ruther, W.; Aspenberg, P. Stimulation of implant fixation by parathyroid hormone (1-34)-A histomorphometric comparison of PMMA cement and stainless steel. J. Orthop. Res. **2005**, 23, 1266–1270.
23. Fleischmann, R.; Iqbal, I.; Stern, R. Considerations with the use of biological therapy in the treatment of rheumatoid arthritis. Expert Opin. Drug Saf. **2004**, 3, 391–403.
24. Kurth, A.; Eberhardt, C.; Muller, S.; Steinacker, M.; Schwarz, M.; Bauss, F. The bisphosphonate ibandronate improves implant integration in osteopenic ovariectomized rats. Bone **2005**, 37, 204–210.
25. Peter, B.; Pioletti, D.; Laib, S.; Bujoli, B.; Pilet, P.; Janvier, P.; Guicheux, J.; Zambelli, P.; Bouler, J.; Gauthier, O. Calcium phosphate drug delivery system: influence of local zoledronate release on bone implant osteointegration. Bone **2005**, 36, 52–60.
26. Bhandari, M.; Bajammal, S.; Guyatt, G.; Griffith, L.; Busse, J.; Scunemann, H.; Einhorn, T. Effect of bisphosphonates on periprosthetic bone mineral density after total joint arthroplasty. A meta-analysis. J. Bone Joint Surg. **2005**, 87A, 293–301.
27. Wilkinson, J.; Eagleton, A.; Stockley, I.; Peel, N.; Hamer, A.; Eastell, R. Effect of pamidronate on bone turnover and implant migration after total hip arthroplasty: a randomized trial. J. Orthop. Res. **2005**, 23, 1–8.
28. Albagha, O.; Ralston, S. Genetic determinants of susceptibility to osteoporosis. Endocrinol. Metab. Clin. North Am. **2003**, 32, 65–81.

29. Engesaeter, L.B.; Lie, S.A.; Espehaug, B.; Furnes, O.; Vollset, S.E.; Havelin, L.I. Antibiotic prophylaxis in total hip arthroplasty. *Acta. Orthop. Scand.* **2003**, *74*, 644–651.
30. Parvizi, J.; Wickstrom, E.; Zeiger, A.; Adams, C.; Shapiro, I.; Purtill, J.; Sharkey, P.; Hozack, W.; Rothman, R.; Hickok, N. Titanium surface with biologic activity against infection. *Clin. Orthop. Rel. Res.* **2004**, *429*, 33–38.
31. Bhumbra, R.; Walker, P.; Berman, A.; Emmanuel, J.; Barrett, D.; Blunn, G. Prevention of loosening in total hip replacements using guided bone regeneration. *Clin. Orthop. Rel. Res.* **2000**, *372*, 192–204.

Peripheral Nerve Repair and Regeneration: Historical Perspective

E.E. Sabelman

Department of Neurology, Kaiser-Permanente Medical Center, Redwood City, California, U.S.A.

W. Chang

S. Nguyen

J.M. Rosen

Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire, U.S.A.

INTRODUCTION

Nerve injuries can lead to clinical deficits that range from temporary loss of sensation to quadriplegia. Our understanding of nerve injury and repair has increased dramatically since the time of Galen, when nerves were first distinguished from tendons; however, the clinical application of scientific knowledge continues to be subdued, with functional results of nerve repair largely unchanged during the past 30 years. Direct approximation of proximal and distal ends of cut nerve remains the gold standard whenever possible, despite results that are less than 100% effective. When a significant nerve gap is incurred that cannot be repaired by end-to-end suturing, the current repair method involves implantation of an autograft as a conduit for axonal regeneration. However, surgical nerve grafting using a donor nerve is limited by availability and donor site morbidity, such as loss of sensation and cosmesis. Research in biomaterials and tissue engineering was recently enlisted to overcome these limitations. This article reviews historical and current understanding of surgical nerve repair and introduces alternate techniques using novel tissue engineering constructs for the repair and regeneration of peripheral nerve.

Nerve injuries complicate successful rehabilitation more than any other form of trauma. Peripheral nerve damage occurs in approximately 30% of extremity trauma.^[1] Self-care and other tasks of daily life can be severely impaired because of a lack of tactile discrimination and motor dysfunction.^[2] Pain from neuromas, sympathetic dystrophy, and causalgia often follow nerve injury, often necessitating further surgery.^[3] Causalgia has an estimated occurrence of 1–15% in sharp injuries. Causalgia and sympathetic dystrophy are commonly associated with crush injuries, whereas neuromas commonly result from transected nerve injuries.^[4,5] Improved guidance of regenerating nerves would, in theory, reduce the severity of these disabling sequelae.

RELEVANT ANATOMY

Peripheral nerves carry within them motor, sensory, and sympathetic nerve fibers (axons). Nerves comprise a neural component and a connective tissue component. (See also Sanghvi et al., “Tissue Engineering of Peripheral Nerve.”) The neural component consists of individual axons that can be myelinated or unmyelinated (Fig. 1). The non-neural component is made up of vascular and connective tissue cells and collagenous matrix organized into endoneurium, perineurium, and epineurium layers. The endoneurium is the basal membrane layer that envelops each individual axon; it is secreted by Schwann cells, which are the glial cells of the peripheral nervous system normally ensheathing myelinated axons, except at Ranvier’s nodes. The axons themselves are organized into fascicles, which are surrounded by a layer of perineurium, composed of collagen fibers and concentric layers of epithelial cells united by tight junctions. The fascicles are bundled together and surrounded by an outermost layer of epineurium composed of collagen and elastic fibers; it usually represents 50% of the cross-section of the nerve. The nerve receives its blood supply from vessels from the mesoneurium that is contiguous with the epineurium (Fig. 2).^[6]

Human peripheral nerves are subdivided into monofascicular, oligofascicular, and polyfascicular types. Monofascicular nerves (e.g., terminal branches of digital nerves) contain only one large bundle of several thousand axons with either a pure sensory or a pure motor function. Oligofascicular nerves (e.g., the radial nerve at the wrist) are composed of both sensory and motor elements. Polyfascicular nerves are composed of many small fascicles containing axons that serve a variety of functions.

Cellular Changes Associated with Injury

The most common etiology necessitating peripheral nerve repair is upper limb trauma.^[7] Nerve injuries are

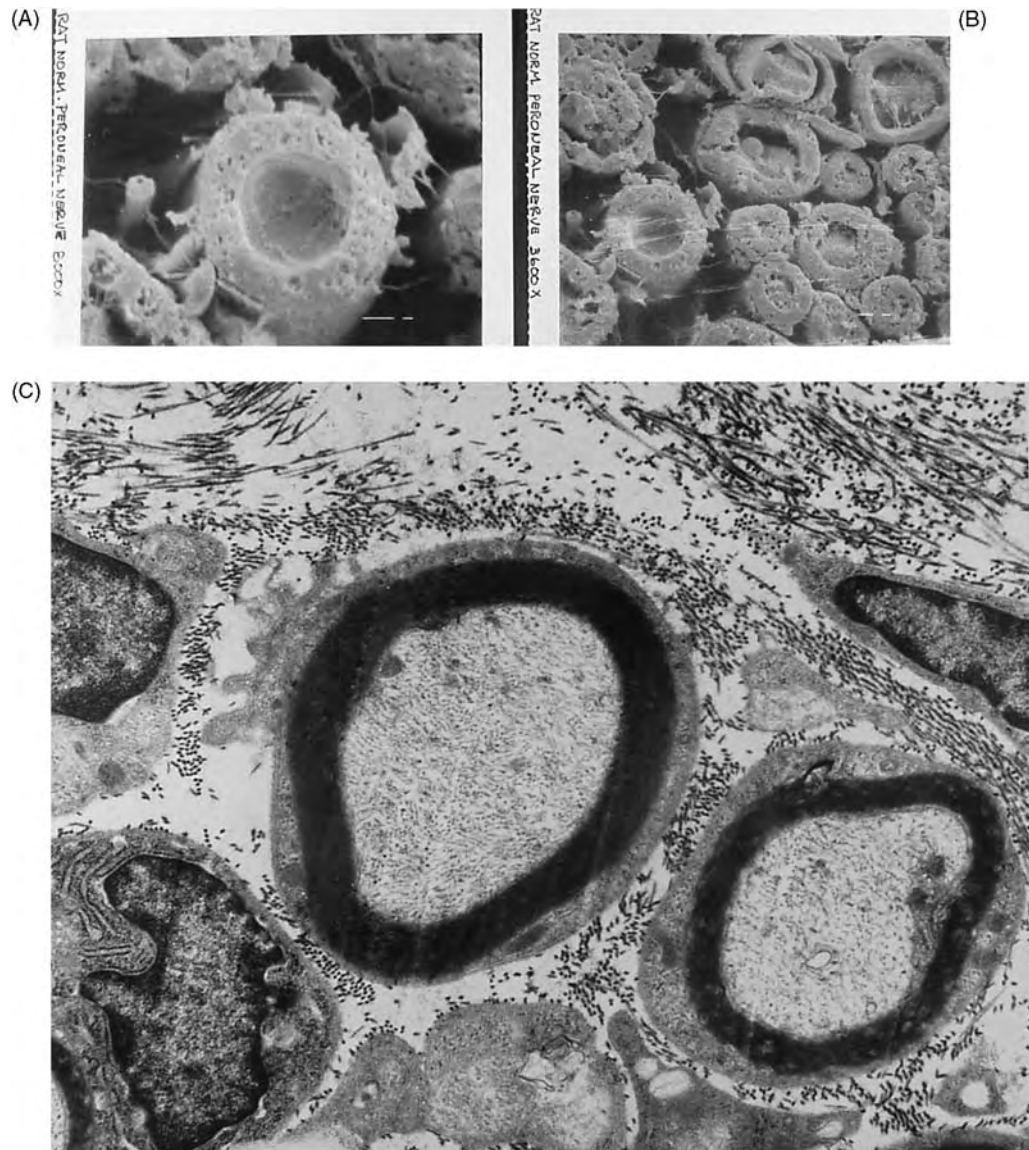


Fig. 1 Rat peroneal nerve cross-section, showing myelinated axons. A, B: Scanning electron micrograph, C. Transmission electron micrograph. Diameter of myelin sheath is typically 20 micrometers. (J. Rosen, unpublished data.)

commonly associated with transection, laceration, or avulsion of the nerve trunk from surrounding tissues. Trauma causes nerves to undergo a succession of changes, not only at the site of injury but also in the brain, which tries to restructure its functional organization as a consequence of the peripheral nerve damage.

The proximal end of the transected axon generally retracts for a distance equal to several Ranvier's nodes, with the injured tip sealing off immediately and forming a proximal stump (Fig. 3). The nucleus and its associated apparatus enlarge to prepare for the regenerative process.^[8] However, in some cell bodies, transection causes cell death through an apoptotic cascade characterized by DNA fragmentation, shrinking of

the cell, chromatin condensation, and formation of apoptotic subcellular bodies.^[6] Thus, the population of neurons capable of axonal regeneration diminishes over time, although apparently never declining to zero.

Distally, transection causes the axons to disintegrate and undergo wallerian degeneration, a sequence of proteolytic processes that disintegrates the axoplasm, leaving remnant debris to be digested by invading macrophages and proliferating Schwann cells. Wallerian degeneration affects the entire distal length of the nerve; because all regenerating axons must traverse this distance, healing of trauma to nerve trunks near the body axis may take years.^[9] (See also Sanghvi et al., "Tissue Engineering of Peripheral Nerve.")

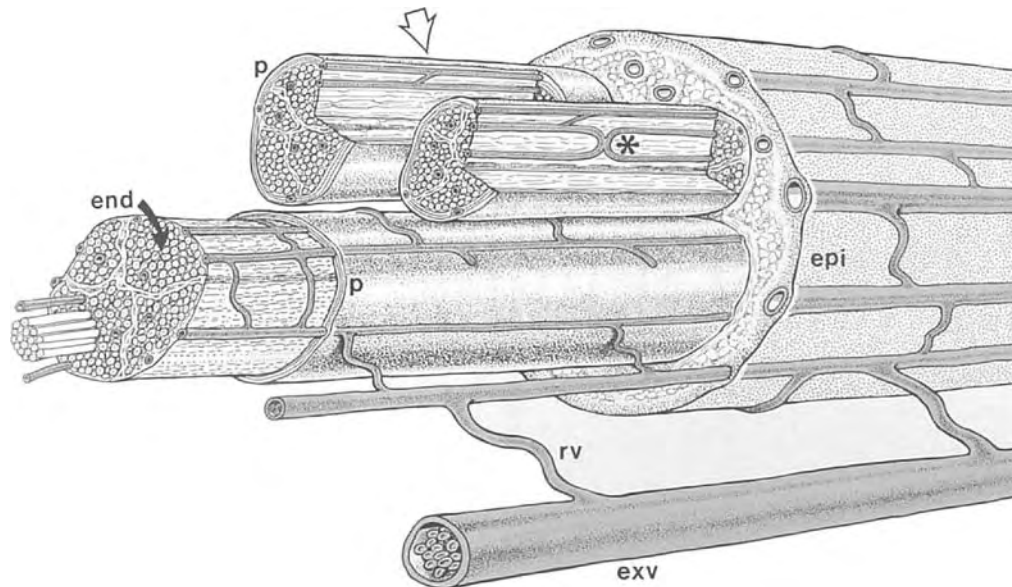


Fig. 2 Vessels from the mesoneurium, which is contiguous with the epineurium. (From Ref. [66].)

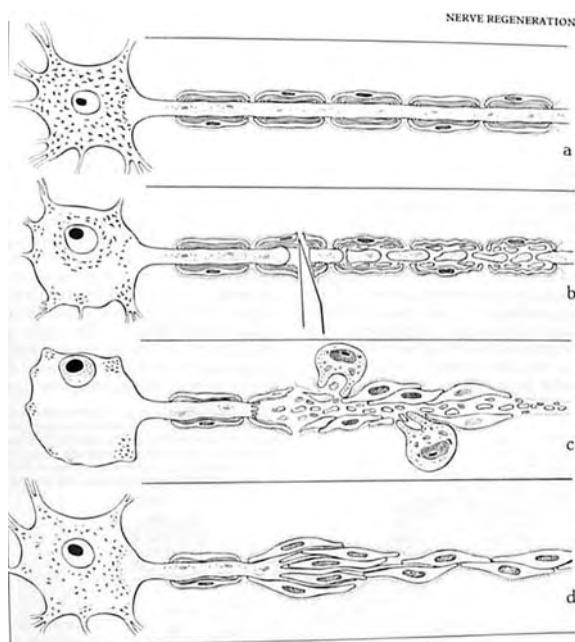


Fig. 3 The proximal end of the transected axon generally retracts for a distance equal to several nodes of Ranvier, with the injured tip sealing off immediately and forming a proximal stump. (From Ref. [66].)

The Role of Schwann Cells

Schwann cells make a crucial contribution to peripheral neural development, maintenance, and regeneration.^[10] In response to injury, Schwann cells divide rapidly, migrate from the two ends of a

transected nerve into the repair zone, phagocytose myelin debris, and create endoneurial “bands of Büngner” to guide axonal elongation through the distal nerve stump (Fig. 4).^[11,12] Schwann cells actively create the correct microenvironment for axonal regeneration by secretion of diffusible molecules such as nerve growth factor,^[13,14] ciliary neurotrophic factor,^[15] dopaminergic neurotrophic factor,^[16] and brain-derived neurotrophic factor.^[17] In vitro studies demonstrated that Schwann cell-secreted extracellular matrix containing Type IV collagen, laminin, and fibronectin^[18] potentiate regeneration.^[13,19–21] Schwann cells also secrete cell adhesion molecules (laminin, L1, N-cadherin, γ -1 integrins, and neural cell adhesion molecule)^[22] and a plasminogen activator for extracellular proteolysis.^[23]

Peripheral grafts without viable Schwann cells fail to support regeneration across peripheral nerve gaps much over 10 mm^[24] or in the central nervous system (CNS).^[25,26] If their multiple functions are not replicated artificially, Schwann cells will be needed throughout the healing nerve to remyelinate axons; if not distributed in the graft at the time of implantation, they will have to migrate from the nerve stumps, slowing the rate of maturation.^[27]

THE COURSE OF NERVE INJURY

Classification

Sir Herbert Seddon in 1943 was the first to describe three discrete types of nerve injury—neuropraxia,

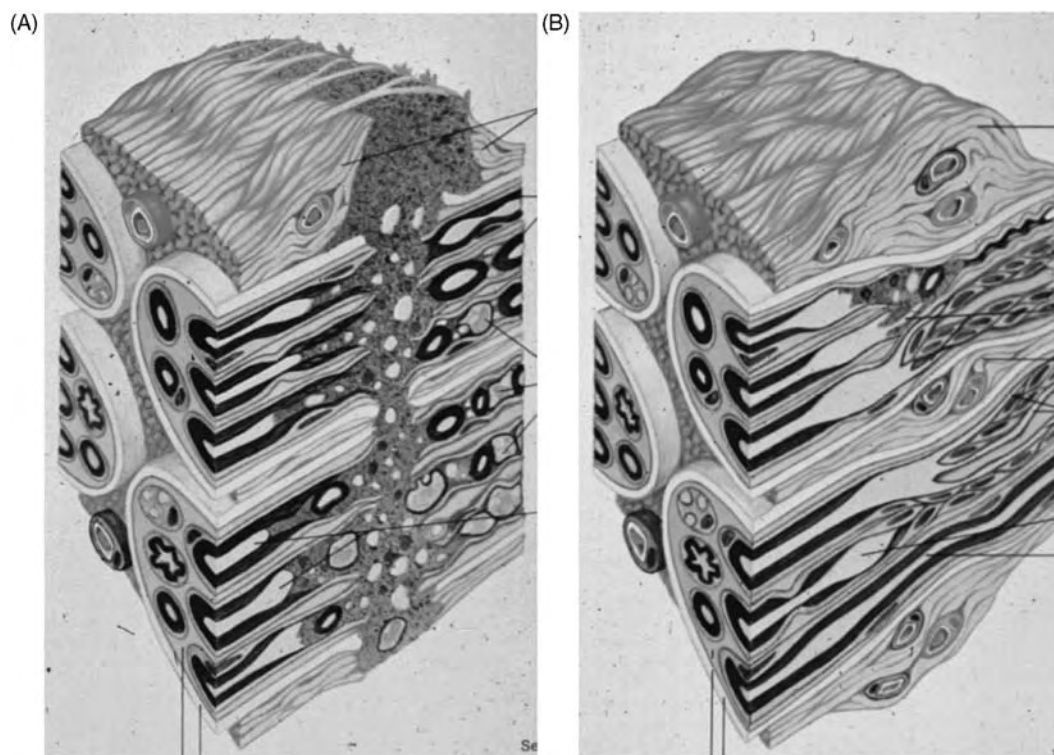


Fig. 4 The Schwann cells de-differentiate at the site of injury to form “Bands of Büngner.” A: Acute nerve injury: axons retract toward cell body, fragment distally. B. Healing nerve: Schwann cells prepare for axons regenerating across injury site. (From Bryant, W.M. “Wound Healing.” CIBA Clinical Symposia 1977, 29 (3), plates 14 and 16. Novartis Foundation.)

axonotmesis, and neurotmesis—as a basis for evaluating and predicting clinical outcome and formulating treatment plans.^[28] In 1951, Sunderland modified the classification system, adding two additional interim steps.^[29] Dellon and Mckinnon later described a six-degree rating scale of injury that dealt with clinical challenges associated with neuroma in situ (Table 1).^[30]

Clinical Assessment of Nerve Repair Outcome

Return of sensation or motor function can be assessed beginning 2 to 4 months postsurgery and continuing 6 to 18 months or longer at monthly intervals, depending on the location of the repair site and cast removal, using routine clinical tests for muscle strength, coordination, and touch. Non-invasive methods for evaluating human motor and sensory recovery include manual and automated strength testing, active range of motion, pain and touch thresholds, sudomotor (sweat gland) response, and surface electrophysiology.^[31] Tinel’s sign is a sensation of tingling felt in the distal extremity of a limb when percussion is performed over the site of an injured nerve, indicating a partial lesion or early regeneration in the nerve.^[32] Specialized equipment for clinical electrophysiology^[33] may be employed. New non-invasive methods for

assessing cross-sectional anatomy of the repaired nerve are available (i.e., high-resolution nuclear magnetic resonance^[34]).

Sensory assessment

Clinical sensory tests of hand function include touch, tactile gnosis (spatial discrimination), pain, and control of sweat glands.^[35] Touch threshold is measured with Semmes–Weinstein monofilaments and tactile gnosis using two-point discrimination. Patients’ self-report of recovery or persistent pain distal to a repair or donor nerve biopsy is significant.^[36] Temperature and vibration thresholds^[37,38] may be tested, but tend to be non-specific or duplicate results of other methods (E. Moberg, as cited by Dellon^[35]).

Recovery of sudomotor response^[39] is believed to parallel return of pain and temperature sensation, although a recent paper^[40] cautions against over-reliance on this assumption. In the ninhydrin test, ninhydrin is applied to the patient’s hand, who then places his or her hand on special paper; color will only develop in areas of innervated sweating skin.^[35] A newer method employs topical application of a known dose of capsaicin, which induces local hyperemia (evinced as reddening) only if arteriolar

Table 1 Classification of peripheral nerve injuries

Type I injury (neuropraxia) represents the mildest type of injury and is associated with full recovery of function. It involves a temporary loss of nerve conduction from ischemia-induced focal demyelination. There is no damage to the nerve itself, and clinically there is no Tinel's sign (see below) at the site of the lesion. A common example of this type of injury is "Saturday night palsy," where an intoxicated patient falls asleep with his or her arm hanging over a chair for an extended period of time, resulting in compression of the radial nerve. Experimental sham surgery, where the nerve is exposed to air without intentional trauma, can produce similar results.

Type II injury (axonotmesis) is an injury where wallerian degeneration occurs in the distal segment. Typically a crush injury, the endoneurium and perineurium remain intact; hence, scar formation is not seen. Recovery will be complete unless the distance of the injury from the motor end plate results in prolonged denervation of the receptor, causing motor recovery to be adversely affected. Nerve regeneration proceeds at 1.0–1.5 mm/day and can be followed clinically with advancement of Tinel's sign. Such injuries do not require surgical intervention.

Type III injury refers to distal wallerian degeneration with the addition of fibrosis of the endoneurium, although the perineurium remains intact. The differentiation between a Type II and Type III injury is based on the degree of recovery. Recovery will be incomplete as a result of scarring of the injury site, which blocks nerve regeneration. There may be a mismatch between regenerating fibers and appropriate end organ receptors. Regeneration occurs at a rate similar to axonotmesis. Surgery is usually not required, but in cases where there is delayed recovery, surgery may be necessary to remove the scar. Clinical recovery will be incomplete.

Type IV injury describes a nerve physically in continuity, but not functioning because of a complete blockage of nerve regeneration by scar tissue. Tinel's sign is seen at the level of injury, but there is no progression of Tinel's sign over time, indicating a lack of regeneration. Wallerian degeneration occurs distal to the injury. Surgical intervention is required for any hope of recovery. This is usually accomplished with primary nerve repair, nerve grafting, or interposition of a nerve conduit to bridge the gap left by scar excision.

Type V injury (neurotmesis) is the result of complete transection of the nerve. Surgery is required before any regeneration can occur. It is usually straightforward to diagnose these types of injuries because almost all complete nerve transactions are open injuries. Treatment is with early surgical exploration and primary repair of the nerve.

Type VI (neuroma in continuity) injury represents a combination of any of the above five levels of injury. It involves a mixed pattern of injury to the individual fascicles of the peripheral nerve, ranging from no injury to Type V injuries. Surgical intervention is usually necessary, with variable clinical recovery. Because only Type IV and Type V injuries are improved by surgery, the surgeon should make a detailed preoperative clinical assessment of the function of each fascicle or group of fascicles so as not to damage uninjured Type I–III injuries during surgery.

(From Ref. [6].)

sphincters are innervated.^[41] Noninvasive measurement of sensory nerve compound action potential is possible, although difficult to interpret.^[42–44] Sural nerve donor site evaluation should be considered in autograft patients.^[45]

Motor assessment

Standard manual anatomic evaluations consist of using a goniometer to measure joint-selected range of motion.^[46] Strengths of individual muscles are assessed non-invasively by the isometric manual motor test—preventing motion of muscles or tendons while the subject attempts motion against the therapist's restraint.^[47] Grades are assigned based on the muscle's ability to contract, total force the muscle can produce, and the amount of resistance, including gravity, against which a muscle can contract. Functional evaluations such as the Jebsen hand function test^[48] and postural assessment^[49] can provide an appraisal of the loss of function as it relates to activities of daily living. Automated strength (grip, palmar pinch, key

pinch, and tip pinch)^[50] and range of motion measurements in the hand can be made using the Greenleaf EVAL™ system.^[51]

Factors Affecting Success of Nerve Repair

Clinical results following nerve repair are influenced by many parameters such as the nature, location, and extent of the injury; the level and timing of the repair; and surgical techniques, as well as factors related to the patient's health (e.g., smoking, compliance with physical therapy).^[52] Probably most important is the prompt timing of surgery,^[6] with primary (i.e., without later secondary reentry of the surgical site) repair being preferable, if possible. Early repair circumvents major problems such as neuronal apoptosis, muscle atrophy, and scar formation, all of which increase with time. Chronic denervation (more than 3 years) leads to atrophy of sensory and motor end organs; therefore, recovery after grafting is likely to be delayed and of lesser extent than if the target tissues are intact.^[53] However, there have been occasional reports of

Table 2 Sensory and motor results after grafting

Gap grafted (cm)	No. cases	Age (yr)	% Children	Follow-up (yr)	Motor Recovery (%)				Sensory Recovery (%)					
					M2	M3	M4	M5	S2	S2+	S3	S3+	S4	
3–15	32			1–15										3
5–15	11	Adult		2–3	0	36	18	0	10	18	45	18	0	0
> 7	33			> 5	0	69	0	0	21	0	69	0	0	0
2–20	38	8–62	3	5–11	7	21	14	46	3	0	60	34	3	3
5–10	8	17–38	0	1.5–2.5	12	0	50	0	13	0	12	63	12	12
4–15	6			1.5–2.5		33	16	33						33
> 2	8	13–57	0	1–5		0	0	0	0	37	38	25	0	0
4–6	5	5–71	2–6		Average M2+		0			Average S2				0

(From Ref. [6].)

successful reinnervation many years after interruption of nerve transmission.^[2,54] Success in restoring motor function to preinjury levels is claimed in 33% to 46% of some series of patients, but more often achieves only a level of 3 on a 0-to-5 scale.^[30] Sensory recovery rarely (3% to 12%) is complete, with two-point discrimination only to 7 mm at best (Table 2).

Factors intrinsic to the injury itself are a significant harbinger of successful outcome. Shorter defects and uncontaminated injuries^[55] are more likely to have better functional outcomes. Meticulous surgical technique is critical to a successful functional outcome. Stretching of the distal and proximal ends of a nerve must be avoided. Coaptation must not be performed under tension because postoperative edema will cause the nerve ends to swell, resulting in poor signal conduction, increased scar formation, and reduced intraneural microvascular flow.^[56] Finally, age and other health factors are important to successful outcome. In general, young children have greater regenerative capability and elderly patients have the poorest results for nerve repair or grafting.

HISTORY OF SURGICAL NERVE REPAIR

Pre-1950s

Our understanding of nerve injury and repair dates back to the time of Ferrara in 1608, when the primary repair of a transected nerve was first documented.^[6] However, it was not until the 1800s that nerve injury and repair were seriously studied and repair techniques were reported in the literature. Before this time, peripheral nerve injuries were usually left untreated or amputated, perhaps because of a lack of antiseptic

measures or antibiotics. In 1850, Waller extensively studied the processes associated with nerve injury, carefully denoting the cellular changes to axons after transection of the hypoglossal nerve in frogs and demonstrating the reproducible pathologic changes following injury that now bear his name.^[57]

The American Civil War (1860–1864) and the resulting large number of peripheral nerve injuries provided greater opportunities for increased understanding. From examination of acute and chronic nerve injury during the Civil War, Mitchell (Fig. 5) and colleagues, in 1864, first demonstrated that partial nerve injury could lead to causalgia, in which the fingers are tapered, painful, and hairless because of a lack of innervation.^[58] From the period of the Civil War until World War II, numerous surgeons began experimenting with nerve repair and grafting. In 1873, Heuter achieved end-to-end coaptation of nerve stumps by epineural suturing, which continues to be the standard used today.^[57] Although ultimately unsuccessful, in 1878 Albert first attempted clinical nerve grafting using a donor nerve from another patient to bridge the median nerve after tumor resection.^[6] In 1898, Tornau published the largest series of nerve repair of his time, comprising 212 primary and secondary repair cases.^[6] Fascicular repair was first advocated by Hashimoto and Langley in 1917,^[59] but achieved limited use until the advent of the surgical microscope.^[60]

Post-1950s

The period following World War II saw renewed interest and experimentation in nerve repair, in part because of the explosion of war injuries now treatable with modern medicine. Woodhall and Beebe (1956)





Fig. 5 Portrait of S. Weir Mitchell from the College of Physicians of Philadelphia. From examination of acute and chronic nerve injury during the Civil War, Mitchell and colleagues first demonstrated, in 1864, that partial nerve injury could lead to causalgia.

reported on 3656 peripheral nerve repairs following war injuries, with relatively poor functional outcomes.^[61] Such poor results colored the perception and expectations of nerve repair for surgeons until a major breakthrough occurred with the work by Millesi in the late 1960s, advocating microsurgical techniques, limited tension, and the use of autografts whenever direct repair was not feasible.^[56] Since Millesi, our understanding of nerve pathophysiology has increased substantially, but there has been little translation of knowledge from experimental models into clinical practice. Millesi's techniques continue to be standard practice today for repair of peripheral injuries.

Prior to development of modern biomaterials, there were nonetheless many attempts to bridge nerve gaps with non-neural materials. Weiss (Fig. 6) cites 53 publications using nerve conduits from the late 1800s to the 1940s.^[62] For example, Forssman in 1898 observed that nerve regeneration could be guided using straw tubes as a bridge. He attributed this phenomenon to neurotropism, the theory that regenerating axons have an affinity for certain substances. This was the view posited by Ramon y Cajal, who believed chemical factors guided the proximal sprout to the distal nerve. Many years later, this concept was called into question by Weiss and Taylor in a series of experiments using Y-shape arteries. Weiss demonstrated that nerve regeneration occurred independently of the distal target, supporting the concept of nerve growth



Fig. 6 Portrait of Paul A. Weiss from the National Academy of Sciences, *Biographical Memoirs*, Vol. 72 (1997). Weiss conducted a review of the many attempts to bridge nerve gaps with non-neural materials, prior to development of modern biomaterials, in the period between the late 1800s and the 1940s.

by contact guidance.^[63] Recent work demonstrated conclusively that both neurotropism and contact guidance both play a critical role in regeneration.^[64] Other studies demonstrated further end-organ specificities, as well as fascicular-specific growth regulation.^[65] These experiments laid the foundation for the concepts of nerve guides and tissue engineering to replace the autograft.

Current Nerve Repair Techniques

The aim of nerve repair is to guide the regenerating axons successfully into the distal nerve with minimal blockage or escape. Nerve repair methods are divided into direct repair (neurorrhaphy) and graft techniques. Direct repairs can be further subclassified into epineural repair and fascicular repair. Graft repairs are performed using harvested nerve segments from autogenous sensory nerves such as the sural or antebrachial cutaneous nerve^[66] (Table 3). The surgical method selected is dependent on the nerve type and tension at the repair site, which is influenced by the gap length, timing of repair, and nature of the injury. In complicated wounds, primary treatment may consist of debridement, reestablishing a vascular supply, and skin closure, with secondary nerve and tendon repair delayed some weeks until inflammation subsides.^[67] Donor nerve excision can be performed by an open

Table 3 Common donor nerve graphs

Nerve graft	Abbreviation	Length (cm)	Morbidity
Sural		30–50	+
Medial antebrachial cutaneous	MACN (above elbow)	10–12	+
Medial antebrachial cutaneous	MACN (below elbow)	8–10	+
Lateral antebrachial cutaneous	LACN (below elbow)	10–12	++
Superficial branch radian	SBRN	20–30	++
Dorsal cutaneous branch ulnar	DCBUN	4–6	++
Lateral femoral cutaneous	LFCN thigh	15–20	+++

(From Ref. [66].)

approach or using endoscopic surgery.^[68,69] Post-surgical rehabilitation includes immediate splinting, followed by successively higher intensity exercise.

Epineural vs. Fascicular Repair

Clinical and experiment studies compared epineural repair to fascicular repair in the treatment of peripheral nerve injuries^[70] (Table 4). Theoretically, fascicular repair should be superior because of better alignment of proximal axons with the bands of Büngner, leading to their preinjury destinations; however, epineural repair is easier, faster, and demands less manipulation, thereby minimizing the disruption to the fine internal structure and critical blood supply. In the acute setting, epineural repair is sufficient where there is a sharp, clean, and clearly delineated lesion. Millesi reports advantages of microsurgical repair under favorable conditions, but the increase in surgical operative time and scar formation should be considered.^[56] If it were possible to completely align the fascicles with minimal trauma and disruption, fascicular repair would likely produce better clinical results than epineural repair.

Nerve Graft Surgical Procedure

When the direct repair of a nerve injury cannot be performed without tension, nerve grafting is recommended.^[71] The graft is a conduit for axonal progression from the proximal to the distal stump, most often being an autologous non-vascularized nerve. The surgical technique of nerve grafting is similar to direct neurorrhaphy, requiring the same microsurgical, atraumatic technique to ensure best results. When

Table 4 Epineural versus fascicular repairs

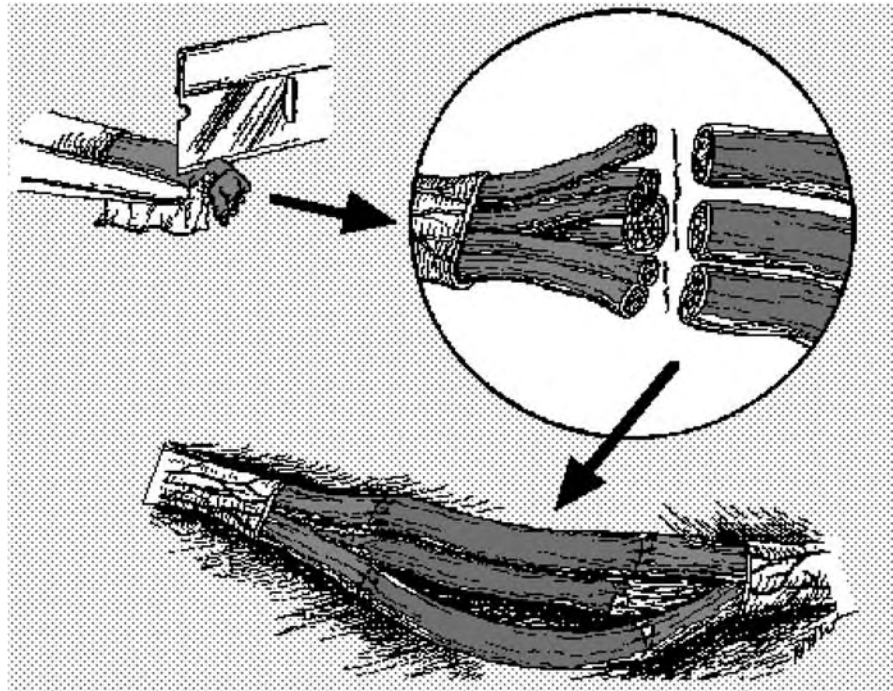
Epineural repair
Heuter's end-to-end coaptation of nerve stumps is the most traditional and commonly used method of repair for transected nerves. Sutures are passed through the epineural sheath using 8-O to 10-O non-absorbable nylon sutures under magnification. It is important to identify the fascicular pattern of the nerve so as to obtain proper anatomical alignment. Care must be taken not to disturb other intraneural contents so as to minimize subsequent fibrosis and scar formation. The epineural suturing technique is relatively simple because only low magnification is required; however, one major disadvantage is that individual fascicles may be misaligned, which in turn may result in defective reinnervation.
Fascicular repair
Fascicular repair involves the coaptation of corresponding fascicles in the proximal and distal nerve stumps to try to achieve anatomic alignment. Such manipulation became possible with the introduction of the operating microscope having over 20 × magnification and long working focus distance. Reconnection of isolated fascicles requires the suturing of corresponding fascicles in the distal and proximal stumps using 10-O and 11-O sutures. Meticulous attention must be paid to surgical technique in order to avoid tension and to prevent injury to the interfascicular content. An advantage associated with fascicular compared with epineural repair is the decreased risk of axonal misdirection; however, the risk of damage to the fascicles may increase because of microherniation of endoneural content. In addition, interfascicular sutures increase intraneural foreign body reactions and, therefore, the likelihood of intraneural fibrosis.

(From Ref. [70].)

performing nerve grafting, the surgeon tries to match the sensory and motor components of the individual fascicles at the proximal and distal repair sites as accurately as possible (Fig. 7). This can be difficult when grafting in the proximal extremity because the fascicular patterns are frequently mixed. Intraoperative identification is often necessary to differentiate the fascicular components. A variety of techniques have been used to aid identification, including electrostimulation, acetylcholinesterase staining, and carbonic anhydrase histochemistry.^[72] Most commonly, surgeons use anatomic cues such as the pattern of epineural blood vessels to properly align the nerve stumps (Fig. 2).

The standard practice of microsurgical nerve reanastomosis has been to use 9-O or 10-O nylon sutures to connect the perineurium of the nerve segments.^[70,73–75] Fibrin tissue adhesive (e.g., Tisseel[®], Baxter-Hyland) was approved for clinical use in 1998, affording a faster, less error-prone method of anastomosis. Nerve autografts and vein-to-nerve grafts





(A)



(B)

Fig. 7 A. Epineural repair to fascicular repair. (From Ref. [74].) B. Epineural repair to fascicular repair. (From Kim, D.H.; Murovic, J.A.; Tiel, R.L.; Kline, D.G. Penetrating injuries due to gunshot wounds involving the brachial plexus. *Neurosurg. Focus* **2004**, *16* (5).) Courtesy of D. Kim.

are now routinely affixed using a bead of fibrin surrounding the junction and extending several millimeters to either side. (See Carr et al., "Fibrin Sealants.")

Donor Site Morbidity

Donor nerves excised for nerve grafting are ideally expendable sensory nerves whose harvest results in

minor disability. Donor nerves include the sural, lateral antebrachial, anterior division of the medial antebrachial cutaneous, intercostal,^[76] or supraclavicular nerves^[77] (Table 2). In selecting a donor graft, it is important to consider the cross-sectional area of the nerve to be repaired, the length of the nerve gap, and donor site morbidity. Donor site morbidity involves loss of donor nerve function, hyperesthesia adjacent to the region of sensory loss, and the possibility of a

painful neuroma. Neuromas are nodular outgrowths from a damaged nerve that occur when axons escape past the remaining perineurium^[78] or encounter an obstacle such as a blood clot.^[5] Risk of neuroma is minimized if the proximal stump is tunneled into innervated muscle, which inhibits axonal sprouting. One study on harvesting of the sural nerve reported that 44% of patients complained of sensory defect around the lateral aspect of the foot and ankle, 42% of patients complained of calf tenderness, and intolerable pain associated with neuroma occurred in 16% of patients.^[6] Sensory loss in the leg as a result of excising the sural nerve can adversely affect gait and balance.^[79]

ALTERNATIVES TO NERVE AUTOGRAFTS

The greatest need for donor autograft nerve at a single injury site is for repair of the brachial plexus—the complex network of cervical nerves proximal to the upper limb.^[80] Problems with autograft repair include a second surgical site, donor site morbidity, and limited supply.^[81–83] Experimental alternatives to autografts include allo- or homografts, heterografts, and artificial materials in tubular form.^[84] An artificial nerve graft (ANG) potentially would solve autograft-associated problems and promote functional recovery equal or superior to an autograft repair.^[30]

Nerve xenografts (cross-species) and allografts (same species) entail rejection or health risks associated with systemic immunosuppression to allow graft acceptance.^[70,85,86] Recent experiments demonstrated that immunogenic compounds can be extracted from allogeneic donor nerve.^[87,88] Other grafts of biologic origin include autogenous veins^[89,90] and pseudosynovial or mesothelial tubes,^[91] experimental results were deemed acceptable but far from ideal. Chiu et al.^[89] reported clinical trials using veins to bridge nerve gaps of 15 mm and longer. Beyond 15 mm, they noted that nerve regeneration ceased and suggested that this may be because of a lack of factors needed for regeneration.^[56]

Guidance Tubes

Although autogenous nerve grafting continues to be the gold standard, one possible alternative is the design of artificial constructs to replace the elements necessary for axonal regeneration. The simplest such artificial nerve graft replaces only the perineurium with a tubular conduit; a complete tissue-engineered ANG would emulate the internal geometry, cellular population, and neurotrophic environment of the natural autograft. The ideal result of implantation of an artificial nerve graft is return of sensory and muscle function as good

or better than that of an autograft^[81] (typically 80% of the intact uninjured nerve; see Table 2). Because no substantial amount of donor nerve is required, the risk of functional loss, scarring, or neuroma at a donor site is minimized.^[92] (See also Schlosshauer and Lietz, “Nerve Guides.”)

There has been a resurgence in interest in the potential for interventions to restore lost neural cellular population, structure, and function.^[93,94] Early human trials are being conducted using relatively simple microgeometry, cell manipulation, and surgical techniques. Notable among these trials is that of Integra Life Sciences, which obtained license to the Yannas patents for collagen/glycosaminoglycan composite implants.^[95] Implantation of unfilled 2- to 5-cm collagen tubular grafts in Danish patients reportedly began in September 1998.^[96] This is a large-scale product-oriented second-stage trial, compared with one-time trials of synthetic biomaterials in relatively few patients by Mackinnon and Dellon in the 1980s.^[97]

Biomaterials for ANGs

Biocompatible materials are required for both the tubular outer conduit and the optional internal scaffold through which the axonal fibers regenerate.^[98] Natural biopolymers, insoluble synthetics, biodegradable synthetics, and electrically active materials have been used for this purpose. The limitations of natural biomaterials include immunoreactivity with the patient and non-uniformity in the graft material. It may be more feasible to alter the chemical and engineering properties of synthetic biomaterials so as to achieve the desired geometric configuration, biocompatibility, porosity, degradation, and mechanical strength.

Silicone is the most frequently employed synthetic biomaterial,^[99,100] but, like other non-resorbable materials,^[101] results in declining function after approximately 6 months as a result of constriction of the enlarging nerve.^[102] The common finding that regenerating axons detach from the wall and form a thin bundle or cable in the center of the silicone tube illustrates cellular non-adhesion; texturing the inner surface can result in a much different, looser structure.^[66] (See also Wynne and Lambert, “Silicones.”)

Experiments with synthetic resorbable materials have had some success.^[103,104] As in the case of vein grafts, clinical trials of bioresorbable tubes fabricated from polyglycolic acid (PGA) without matrix or cells in 15-mm gaps yielded results comparable to that of sutured autografts, but performed poorly in gaps of 30 mm or more.^[97,105] (See also Boland et al., “Poly(Glycolic Acid).”) Newer bioresorbable materials such as glycolide trimethylene

carbonate (GTMC, Maxon™),^[106] polydioxanone, and poly(lactic-ε-caprolactone)^[107] are less brittle and longer lived in vivo. These materials have varying time courses of inflammation and induced fibroblast proliferation.^[108] As with PGA, as long as particles of GTMC or polydioxanone remain (3–6 months), inflammation, mononuclear macrophages, multinuclear giant cells^[97] and compression of the nerve persist.^[107] (See also West, “Biofunctional Polymers.”)

The predominant natural biomaterial for nerve conduit fabrication has been Type I collagen. Type I collagen is the principal extracellular protein of connective tissue and is a polymer of triple-helical subunits.^[109] (See also Beckman et al., “Collagen.”) In human femoral nerve, Type I is present in the perineurium,^[110] whereas collagen Type III is predominantly located in the endoneurium.^[111] Dissolved collagen can be repolymerized to form gels, coatings, and microfibrillar structures,^[112] which can be strengthened by molecular cross-links.^[113]

Experiments demonstrated that silicone tubes filled with Type I collagen result in successful bridging of short nerve gaps.^[114–116] Empty tubes made of collagen support nerve regeneration equal to other biomaterials.^[117,118] This finding was recently confirmed in sites adjacent to the spinal cord as well as distally.^[119] Saline-filled bovine collagen tubes across short (4-mm) gaps in peripheral nerves in primates demonstrate slower initiation of regeneration but equal long-term results compared with autografts.^[120] The combination of collagen with other natural biopolymers such as fibrin may yield better control over resorption rate and inflammation.^[121] Some researchers concluded that regeneration in collagen is impeded,^[122] but these investigators did not compare the same collagen preparations. Parallel orientation of collagen fibrils at the molecular level promoted unidirectional axon elongation.^[123,124]

Tissue Engineering Concepts

Tissue engineering, defined as the combination of living cells with a biomaterial scaffold, holds promise for replacing tissues lost to injury or disease.^[125] (See also Auger and Germain, “Tissue Engineering.”) Whereas other approaches to tissue engineering for human implantation may use donor cells that can be rejected, our approach is based on using a patient’s own (“autologous”) Schwann cells.^[126,127] One of the critical assumptions we made in formulating the ANG is that incorporation of Schwann cells will enhance axonal regeneration by de novo synthesis of growth factors and neuron-specific adhesion factors if they are placed in a biocompatible matrix with suitable micrometer-scale geometry. Researchers working on

peripheral nerve at the Miami Project to Cure Paralysis^[128] made the same assumption. Relying on the patient’s own Schwann cells to provide neurotrophic factors entails the burden of obtaining and isolating the cells and custom-fabricating grafts for each individual patient,^[127] as well as the risk that the cells will not survive or secrete the factors in sufficient quantity (autologous Schwann cells genetically altered to hyperproduce growth factors are eventually possible^[129]).

Schwann Cell-Seeded Grafts

Much has been learned about isolation and culture of adult human Schwann cells.^[130,131] It is now known that growth factors such as heregulin preferentially stimulate proliferation of Schwann cells.^[128,131–133] Previously, we purified Schwann cells using a complement-lysis method to eliminate fibroblasts; this method has a low yield of viable cells.^[134] We since modified the successive-migration explant culture method used by the Bunge group^[130] to minimize the initial non Schwann cell population by microdissection, removing blood vessels as well as perineurium before enzymatically releasing Schwann cells; residual fibroblasts are <5% of the total cell number. More efficient methods of isolating and culturing large numbers of Schwann cells are based on immunoaffinity methods employing antibodies to Schwann cell-specific surface antigens (e.g., S-100)^[135] and on selective growth media.^[136]

Example Artificial Nerve Graft Concept

Ideally, artificial nerve grafts are composed of viable Schwann cells, a linearly oriented matrix, and an outer bioresorbable polymer conduit, emulating the native nerve after injury. The matrix distributes and maintains the donor Schwann cells inside the graft, and the conduit facilitates surgical emplacement in the host site and reduces scar-forming inflammatory ingrowth. (See also Sanghvi et al., “Tissue Engineering of Peripheral Nerve.”) A typical graft of 1- to 3-mm diameter would replace a single nerve fascicle; for large nerves, fascicles in proximal and distal stumps would be reconnected with individual parallel grafts. Preparation of the graft essentially consists of repolymerization of solubilized Type I collagen with added cultured Schwann cells and insertion into a GTMC or cross-linked collagen conduit. This ANG concept evolved from work by Drs. Hentz and Rosen on a sutureless method (fascicular tubulization) for nerve repair.^[137,138]

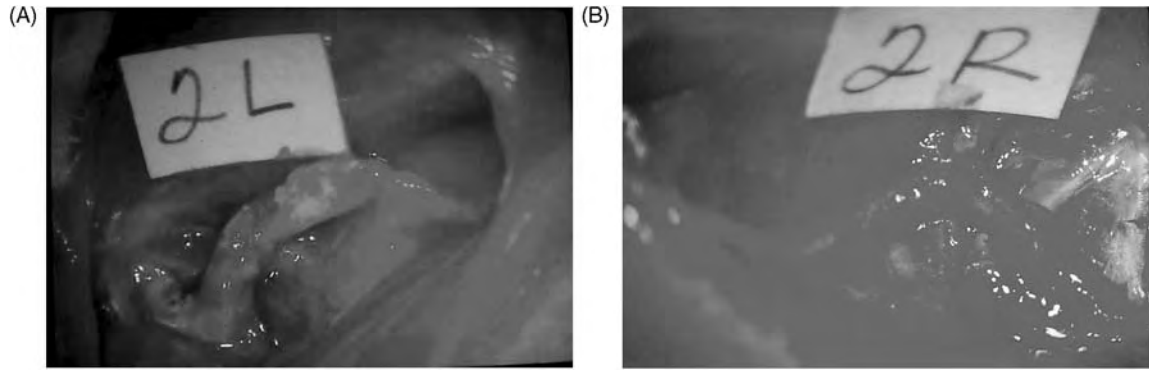


Fig. 8 Rat peroneal nerve implants 28 days post-surgery; artificial nerve grafts were tested in vivo by implantation into 10–14 mm peroneal nerve gaps in the rat. A. Cell-free artificial nerve graft at explantation. B. Sutured autograft at explantation. (J. Rosen, unpublished data.)

ANG Results

ANGs are initially tested in vivo by implantation into 10- to 14-mm peroneal nerve gaps in the rat (Fig. 8). Recovery of motor function is assessed by walking track analysis^[139] at weekly intervals. (See also Sanghvi et al., “Tissue Engineering of Peripheral Nerve.”) Repeated walking track analysis demonstrated that Schwann cell-containing ANGs are superior to collagen-only ANGs and ANGs containing dead Schwann cells or mixed fibroblast/Schwann cells in collagen, which in turn are superior to empty tubes (Fig. 9). Electrophysiology and histological analysis also demonstrate that neonatal Schwann cells favorably influence nerve regeneration if present in adequate numbers.^[140]

If nerve fibers do not reconnect across the gap, the cause is most often because one end came loose from the nerve stump; it may be reconnected by means of another surgical procedure. Another cause is interposition of sutures, blood clots, or other non-nerve tissue, thus increasing the likelihood of scar formation. Both causes also occur in autografts at about the same frequency as expected in artificial grafts.^[52]

Another cause of failure that occurs rarely in autografts is premature resorption of the graft, so the guiding sheath disappears before axons reach their destinations; this possibility is minimized by increasing conduit wall thickness and cross-linking in proportion to the length of the graft. Both in our primate study (unpublished data) and in clinical trials by Mackinnon and Dellon,^[97] failure of axons to reach the distal

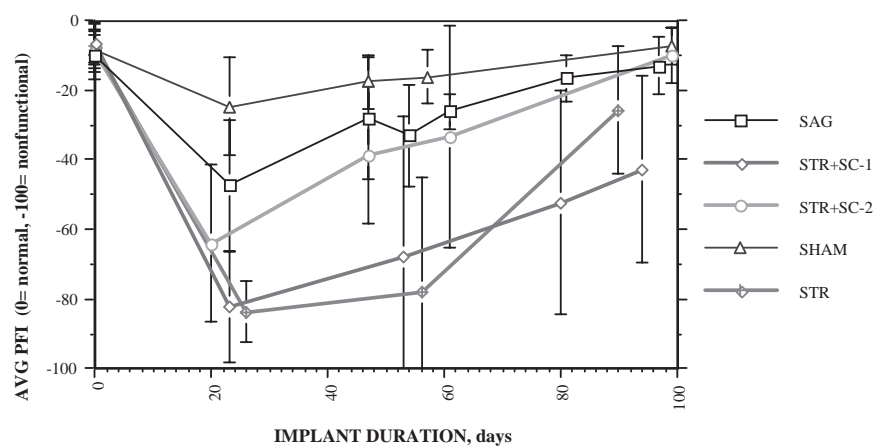


Fig. 9 Repeated walking track analysis has shown that Schwann cell-containing artificial nerve grafts (ANGs) are superior to collagen-only ANGs, ANGs containing dead Schwann cells, or mixed fibroblast/Schwann cells in collagen, which in turn are superior to empty tubes. (From Keeley, R.D.; Sabelman, E.; Kadlcik, P.; Agras, J.; Eng, L.; Wiedman, T-W.; Nguyen, K.; Sudekum, A.; Rosen, J. Synthetic nerve graft containing collagen and syngeneic Schwann cells improves functional, electrophysiological, and histological parameters of peripheral nerve regeneration. *Restorative Neurol. and Neurosci.* **1993**, *5*, 353–366.)

nerve stump was probably a result of premature lysis of the sheath.

According to published reports^[141] and our experience, synthetic polymers do not dissolve uniformly, but by cracking into fragments that can interfere with regeneration. Degradation of collagen is by cell-mediated enzymolysis^[142] and its persistence is a function of cell metabolism and degree of enzyme-resistant cross-linking.^[143] Concerns about flexion-induced fragmentation or pinching of the ANG^[81] where it crosses a joint appear to be minimized by routine postsurgical immobilization for 6–8 weeks. A collagen-based ANG approximates the properties of natural veins used as nerve grafts, which are reported to perform better when subjected to flexing than synthetic polymers.^[144]

CONCLUSIONS

Despite greater understanding of the process of nerve regeneration, surgical practice and restoration of function after peripheral nerve injury have changed little for decades.^[145] The general concept of tissue engineering, or facilitated regeneration through the use of composite artificial/cellular tissues, has broad applications beyond peripheral nerve grafts. In the near future, ANGs would be most frequently used in distal nerve loss in the limbs,^[74] larger diameter multi-fascicular nerve trunks (including branching nerves as in the brachial plexus),^[146,147] facial nerve palsy^[148,149] and other craniofacial sites,^[77] and iatrogenic damage such as occurs in pharyngeal and prostate cancer surgery.

There is a high probability that a Schwann cell-seeded equivalent to a peripheral nerve autograft would find application in the CNS, particularly for bridging trauma in the spinal cord.^[25,150] Other CNS applications are based on using a matrix carrier for Schwann cells or CNS cells implanted into the brain.^[151,152] Clinical trials of peripheral nerve autografts into the CNS are already under way.^[153,154] Another therapeutic application following spinal cord injury would be to use ANGs to reroute some fascicles of nerves with intact CNS communication to denervated muscles^[155,156] as an alternative to tendon transfer to restore hand function.^[1]

ACKNOWLEDGMENTS

We acknowledge VA Rehabilitation R&D Merit Review Grants B588-1, -2, -3RA “Artificial Nerve Graft: Union of Cellular and Non-Cellular Components,” and “Clinical Trial of Artificial Nerve Graft” to J. M. Rosen, M.D., and E.E. Sabelman,

Ph.D. V.R. Hentz, M.D., provided long-term support and inspiration. We thank H. Pham and F. Zhang for animal surgery, J. Kipp, P. Kadlcik, and M. Hu for histology, R. Keeley for expert rat handling, and S. Long and R. Masher for editorial assistance.

ARTICLES OF FURTHER INTEREST

Biofunctional Polymers; Collagen; Fibrin Sealants; Nanoscale Neurosurgery; Nerve Guides; Neuroprostheses; Poly (Glycolic Acid); Silicones; Tissue Engineering; Tissue Engineering of Peripheral Nerve

REFERENCES

1. Seddon, H.J. *Surgical Disorders of the Peripheral Nerves*, 2nd Ed.; Churchill-Livingstone: Edinburgh, 1975.
2. Omer, G.G., Spinner, M., Eds.; *Management of Peripheral Nerve Problems*; W.B. Saunders: Philadelphia, 1980; 487–653.
3. Yüksel, F.; Kislaoğlu, E.; Durak, N.; Uçar, C.; Karacaoğlu, E. Prevention of painful neuromas by epineural ligatures, flaps and grafts. *Br. J. Plast. Surg.* **Apr. 1997**, *50* (3), 182–185.
4. Castaldo, J.E.; Ochoa, J.L. Mechanical injury of peripheral nerves: Fine structure and dysfunction. *Clin. Plast. Surg.* **1984**, *11* (9).
5. Kline, D.G.; Nulsen, F.E. The neuroma in continuity; its pre-operative and operative management. *Surg. Clin. N. Am.* **1972**, *52* (1189).
6. Mackinnon, S.E.; Dellon, A.L. *Surgery of the Peripheral Nerve*; Thieme Medical Publishers, Inc.: New York, 1988; 40–52, 551–574.
7. Beasley, R.W.; Brown, H.G.; Meyer, V.E.; Rabischong, P. *Atlas of Surgical Anatomy of the Hand and Forearm* (pamphlet); Foundation for Hand Research: New York, 1981.
8. Kelly, J. Reactions of Neurons to Injury. In *Principles of Neural Science*; Kandel, E.R., Schwartz, J.H., Eds.; Elsevier: New York, 1985; 25–35, 187–194.
9. Sunderland, S. Rate of regeneration in human peripheral nerve. *Arch. Neurol. Psych.* **1947**, *58*, 251–295.
10. Hall, S.M. The Schwann Cell: A reappraisal of its role in the peripheral nervous system. *Neuropathol. Appl. Neurobio.* **1978**, *4*, 165–176.
11. Torigoe, K.; Tanaka, H.F.; Takahashi, A.; Awaya, A.; Hashimoto, K. Basic behavior of migratory Schwann cells in peripheral nerve regeneration. *Exp. Neurol.* **1996**, *137*, 301–308.
12. Le Beau, J.; LaCorbiere, M.; Powell, H.; Ellisman, M.; Schubert, M. Extracellular fluid conditioned during peripheral nerve regeneration stimulates Schwann cell adhesion, migration and proliferation. *Brain Res.* **1988**, *459*, 93–104.

13. Bunge, M.B.; Jeffrey, J.; Wood, P. Different contributions of Schwann cells and fibroblasts to the collagenous components of peripheral nerves. *J. Cell Biol.* **1977**, *75*, 161a.
14. Distefano, P.; Johnson, E. Nerve growth factor receptors on cultured rat Schwann cells. *J. Neurosci.* **1988**, *8* (1), 231–241.
15. Rende, M.; Hagg, T.; Magal, E.; Varon, S.; Manthorpe, M. Ciliary Neurotrophic Factor (CNTF) immunoreactivity in the rat sciatic nerve. *Soc. Neurosci. Abstr.* **1990**, *16*.
16. Springer, J.; Jacovitti, B.; Maguire, B.; Collier, T. A dopaminergic neurotrophic factor (DTNF) is secreted by Schwann cells grown in culture. *Soc. Neurosci. Abstr.* **1990**, *16*, 822.
17. Barker, P.; Acheson, A.; Pareek, S.; Miller, F.; Murphy, R. Detection of brain-derived neurotrophic factor-like biological activity and mRNA in epidermal fibroblasts and Schwann cells. *Soc. Neurosci. Abstr.* **1990**, *16*, 219.
18. Bixby, L.; Lilien, J.; Reichardt, L. Identification of the major proteins that promote neuronal process outgrowth on Schwann cells in vitro. *J. Cell Bio.* **1988**, *107*, 353–361.
19. Carbonetto, S.; Evans, D.; Cochard, P. Nerve fiber growth in culture on tissue substrata from central and peripheral nervous systems. *J. Neurosci.* **1987**, *7* (2), 610–620.
20. Nath, R.K.; Mackinnon, S.E.; Jensen, J.N.; Parks, W.C. Spatial pattern of type I collagen expression in injured peripheral nerve. *J. Neurosurg.* **May 1997**, *86* (5), 866–870.
21. Nathaniel, E.J.; Pease, D.C. Regenerative changes in rat dorsal roots following Wallerian degeneration. *J. Ultrastructure Res.* **1963**, *9*, 533–549.
22. Mirsky, R. The neurobiology of Schwann cells. *Brain Path.* **1999**, *9*, 293–311.
23. Kalderon, N. Plasminogen activator and Schwann cell plasticity in peripheral nerve regeneration. *Soc. Neurosci. Abstr.* **1990**, *16*, 484.
24. Gulati, A. K. Evaluation of acellular and cellular nerve grafts in repair of rat peripheral nerve. *J. Neurosurg.* **1988**, *68*, 117–123.
25. Guest, J.D.; Bunge, R.P. Human Schwann cells can enhance axonal regeneration and myelination in the nude rat spinal cord. *Soc. Neurosci. Abstr.* **1994**, *20*, 458.9.
26. Smith, G.; Stevenson, J. Peripheral nerve grafts lacking viable Schwann cells fail to support central nervous system axonal regeneration. *Exp. Brain Res.* **1988**, *69*, 299–306.
27. Trumble, T.E.; Parvin, D. Physiology of peripheral nerve graft incorporation. *J. Hand Surg.* **1994**, *19A*, 420–427.
28. Seddon, H.J. A review of work on peripheral nerve injuries in Great-Britain during World War II. *Nervous and Mental Diseases* **1948**, *108*.
29. Sunderland, S. A classification of peripheral nerve injury producing loss of function. *Brain* **1958**, *74*, 491.
30. Mackinnon, S.E.; Dellon, A.L. Future Directions in Peripheral Nerve Surgery. In *Surgery of the Peripheral Nerve*; Mackinnon, S.E., Dellon, A.L., Eds.; Thieme Medical Publishers, Inc.: New York, 1988; 40–52, 551–574.
31. Rosen, J.M.; Jewett, D.L. Physiological Methods of Evaluating Experimental Nerve Repair. In *Nerve Repair and Regeneration, Its Clinical and Experimental Basis*; Jewett, D.L., McCarroll, H.R., Jr. Eds.; C.V. Mosby: St. Louis, 1980; 150–162.
32. Henderson, W.R. Clinical assessment of peripheral nerve injuries: Tinel's test. *Lancet* **1948**, *2*, 801.
33. Husstedt, I.W.; Evers, S.; Grotemeyer, K.H. Reproducibility of different nerve conduction velocity measurements in healthy test subjects and patients suffering from diabetic polyneuropathy. *Electromyogr. Clin. Neurophysiol.* **1997**, *37* (6), 359–363.
34. Shinar, H.; Seo, Y.; Navon, G. Discrimination between the different compartments in sciatic nerve by 2H double-quantum-filtered NMR. *J. Magn. Reson.* **November 1997**, *129* (1), 98–104.
35. Dellon, A.L. *Evaluation of Sensibility and Re-Education of Sensation in the Hand*; Williams & Wilkins: Baltimore, 1981.
36. Dahlin, L.B.; Eriksson, K.F.; Sundkvist, G. Persistent postoperative complaints after whole sural nerve biopsies in diabetic and non-diabetic subjects. *Diabet. Med.* **May 1997**, *14* (5), 353–356.
37. Werner, R.A.; Franzblau, A.; Johnston, E. Quantitative vibrometry and electrophysiological assessment in screening for carpal tunnel syndrome among industrial workers: a comparison. *Arch. Phys. Med. Rehabil.* **1994**, *75*, 1228–1232.
38. Bril, V.; Kojic, J.; Ngo, M.; Clark, K. Comparison of a neurothesiometer and vibration in measuring vibration perception thresholds and relationship to nerve conduction studies. *Diabetes Care* **September 1997**, *20* (9), 1360–1362.
39. Kihara, M.; Opfer-Gehrking, T.L.; Low, P.A. Comparison of directly stimulated with axon-reflex-mediated sudomotor responses in human subjects and in patients with diabetes. *Muscle & Nerve* **1993**, *16*, 655–660.
40. Healy, C.; LeQuesne, P.M.; Lynn, B. Collateral sprouting of cutaneous nerves in man. *Brain* **December 1996**, *119* (Pt 6), 2063–2072.
41. Larkin, S.W.; Williams, T.J. Evidence for sensory nerve involvement in cutaneous reactive hyperemia in humans. *Circul. Res.* **1993**, *73*, 147–154.
42. Kuypers, P.D.L.; Gielen, F.L.H.; Wai, R.T.J.; Hovius, S.E.R.; Godschalk, M.; van Egeraat, J.M. A comparison of electric and magnetic compound action signals as quantitative assays of peripheral nerve regeneration. *Muscle & Nerve* **1993**, *16*, 634–641.
43. Raynor, E.M.; Preston, D.C.; Logigian, E.L. Influence of surface recording electrode placement on nerve action potentials. *Muscle & Nerve* **March 1997**, *20* (3), 361–363.
44. Xiao, S.J.; McGill, K.C.; Hentz, V.R. Inverse volume-conductor method for noninvasive assessment of nerve injuries. *Proc. IEEE Engineering in Medicine & Biology Society 11th Annual International Conference* **1989**, *11*, 994–995.



45. Theriault, M.; Dort, J.; Sutherland, G.; Zochodne, D.W. A prospective quantitative study of sensory deficits after whole sural nerve biopsies in diabetic and nondiabetic patients. Surgical approach and the role of collateral sprouting. *Neurol.* **February 1998**, *50* (2), 480–484.
46. LaStayo, P.C.; Wheeler, D.L. Reliability of passive wrist flexion and extension goniometric measurements: A multicenter study. *Physical Therapy* **1994**, *74* (2), 162–176.
47. Daniels, L.; Worthingham, C. *Muscle Testing; Techniques of Manual Examination*, 5th Ed.; W.B. Saunders: Philadelphia, 1986.
48. Jepsen, R.H.; Taylor, N.; Trieschmann, R.B.; Trotter, M.J.; Howard, L.A. An objective and standardized test of hand function. *Arch. Phys. Med. Rehabil.* **1969**, *56*, 311–319.
49. Lin, F.M.; Pan, Y.C.; Hom, C.; Sabbahi, M.; Shenaq, S. Ankle stance angle: a functional index for the evaluation of sciatic nerve recovery after complete transection. *J. Reconstr. Microsurg.* **April 1996**, *12* (3), 173–177.
50. Jones, E.; Hanly, J.G.; Mooney, R.; Rand, L.L.; Spurway, P.M.; Eastwood, B.J.; Jones, J.V. Strength and function in the normal and rheumatoid hand. *J. Rheumatol.* **1991**, *18*, 1313–1318.
51. Greenleaf EVALtm system. <http://www.steeval.com/>.
52. Doyle, J.R. Factors Affecting Clinical Results of Nerve Suture. In *Nerve Repair and Regeneration, Its Clinical and Experimental Basis*; Jewett, D.L., McCarroll, H.R., Jr. Eds.; C.V. Mosby: St. Louis, 1980; 263–266.
53. Houle, J.D.; Ye, J.H. Changes occur in the ability to promote axonal regeneration as the post injury period increases. *Neuro. Report* **1997**, *8*, 751–755.
54. Dellon, A.L. Wound healing in nerve. *Clinics in Plastic Surgery* **1990**, *17*, 545–570.
55. Cairns, H.; Young, J.Z. Treatment of gunshot wounds of peripheral nerves. *Lancet.* **3 August 1940**, 123–130.
56. Millesi, H. The interfascicular nerve grafting of the median and ulnar nerves. *J. Bone Joint Surg. Am.* **1972**, *54*, 727–748.
57. Waller, A.V. Experiments on the glossopharyngeal and hypoglossal nerves of the frog and observations produced thereby in the structure of their primitive fibres. *Phil. Trans. R. Soc. Lond.* **1850**, *140*, 423.
58. Mitchell, S.W.; Morehouse, G.R.; Keen, W.W. *Gunshot Wounds and Other Injuries of Nerves*; J.B. Lippincott: Philadelphia, 1864.
59. Langley, J.N.; Hashimoto, M. On the suture of separate nerve bundles in a nerve trunk and on internal nerve plexus. *J. Phys. (Lond.)* **1917**, *51*.
60. Smith, J.W. Microsurgery of peripheral nerves. *Plast. Reconstr. Surg.* **1964**, *33*, 317.
61. Woodhall, B.; Beebe, W.G. In *Peripheral Nerve Regeneration: A Follow-up Study of 3,656 World War II Injuries*, VA Medical Monograph; US Government Printing Office: Washington, DC, 1956.
62. Weiss, P.A. *From Cell Research to Nerve Repair*, 2nd Ed.; Futura Publishing: Mt. Kisco, New York, 1976; 316.
63. Weiss, P.; Taylor, A.C. Further experimental evidence against “neurotropism” in nerve regeneration. *J. Exp. Zool.* **1944**, *95*, 289–300.
64. Seckel, B.R.; Ryan, S.E.; Gagne, R.G.; Chiu, T.H.; Watkins, E., Jr. Target-specific nerve regeneration through a nerve guide in the rat. *Plast. Reconstr. Surg.* **1986**, *78* (6), 793–800.
65. Brushart, T.M.; Seiler, W.A. Selective reinnervation of distal motor stumps by peripheral motor axons. *Exp. Neurol.* **1987**, *97*, 289–300.
66. Lundborg, G. *Nerve Injury and Repair*; Longman Group UK, Ltd.: New York, 1988; 64–70, 158–178, 196–199.
67. Kallio, P.K.; Vastamaki, M.; Kolonen, K.A. The results of secondary microsurgical repair of radial nerve in 33 patients. *J. Hand Surg.* **1993**, *18B*, 320–322.
68. Capek, L.; Clarke, H.M.; Zuker, R.M. Endoscopic sural nerve harvest in the pediatric patient. *Plast. Reconstr. Surg.* **October 1996**, *98* (5), 884–888.
69. Jaroszynski, G.; Johnston, G.H. Harvesting of the sural nerve with a tendon stripper. *Microsurg.* **1996**, *17* (4), 217–220.
70. Terzis, J.K. Principles and Techniques of Peripheral Nerve Surgery. In *Reconstructive Microsurgery*; Daniel, R.K., Terzis, J.K., Eds.; Little Brown: Boston, 1977; 387–390.
71. Hentz, V.R.; Rosen, J.M.; Xiao, S.; McGill, K.C.; Abraham, G.A. The nerve gap dilemma. A study comparing nerves repaired end-to-end under tension to tension-free nerve graft techniques in a clinically relevant primate model. *J. Hand Surg.* **1993**, *18*, 417–425.
72. Staniforth, P.; Fisher, T.R. The effects of sural nerve excision in autogenous nerve grafting. *The Hand* **1978**, *10*, 187–190.
73. Stancic, M.F.; Micovic, V.; Bobinac, D.; Starcevic, G.; Fuzinac, A.; Tomljanovic, Z. Electromyographic evaluation of experimental nerve grafts suggests better recovery with microscope assistance. *Pflugers Arch.* **1996**, *431* (6 Suppl. 2), R285–R286.
74. Chase, R.A. *Atlas of Hand Surgery*, Vol. 2; W.B. Saunders: Philadelphia, 1984; 78–91.
75. Shaw-Wilgis, E.F. Nerve Repair and Grafting. In *Operative Hand Surgery*, 2nd Ed., Vol. 2; Green, D.P., Ed.; Churchill Livingstone: New York, 1988; 1373–1403.
76. Zhao, S.; Beurman, R.W.; Kline, D.G. Neurotization of motor nerves innervating the lower extremity by utilizing the lower intercostal nerves. *J. Reconstr. Microsurg.* **January 1997**, *13* (1), 39–45.
77. Mucci, S.; Dellon, A.L. Restoration of lower lip sensation: Neurotization of the mental nerve with the supraclavicular nerve. *J. Reconstr. Microsurg.* **1997**, *13*, 151–155.
78. London, P.S. Neuroma and neurosis in the hand. *Handchirurgie* **1972**, *4*, 103–104.
79. Zehr, E.P.; Stein, R.B.; Komiyama, T. Function of sural nerve reflexes during human walking. *J. Physiol. (Lond.)* **February 1998**, *507* (Pt. 1), 305–314.
80. Allieu, Y.; Chammas, M.; Picot, M.C. Paralysis of the brachial plexus caused by supraclavicular injuries in

- the adult. Long-term comparative results of nerve grafts and transfers. *Rev. Chir. Orthop. Reparatrice Appar. Mot.* **1997**, *83* (1), 51–59.
81. Madison, R.D.; Archibald, S.J.; Krarup, C. Peripheral Nerve Injury. In *Wound Healing. Biochemical and Clinical Aspects*; Cohen, I.K., Diegelmann, R.F., Lindblad, W.J., Eds.; W.B. Saunders: Philadelphia, 1992; 450–487.
 82. Rosen, J.M. Concepts of peripheral nerve repair. *Ann. Plast. Surg.* **1981**, *7*, 165.
 83. Sunderland, S. *Nerves and Nerve Injuries*, 2nd Ed.; Churchill Livingstone: New York, 1978; 423–428.
 84. Fields, R.D.; Le Beau, J.M.; Longo, F.M.; Ellisman, M.H. Nerve regeneration through artificial tubular implants. *Prog. Neurobiol.* **1989**, *33*, 87.
 85. Katsube, K.; Doi, K.; Fukumoto, T.; Fujikura, Y.; Shigetomi, M.; Kawai, S. Nerve regeneration and origin of Schwann cells in peripheral nerve allografts in immunologically pretreated rats. *Transplantation* **15 December 1996**, *62* (11), 1643–1649.
 86. Midha, R.; Munro, C.A.; Mackinnon, S.E.; Ang, L.C. Motor and sensory specificity of host nerve axons influence nerve allograft rejection. *J. Neuropathol. Exp. Neurol.* **April 1997**, *56* (4), 421–434.
 87. Dumont, C.E.; Hentz, V.R. Enhancement of axon growth by detergent-extracted nerve grafts. *Transplantation* **15 May 1997**, *63* (9), 1210–1215.
 88. Sondell, M.; Lundborg, G.; Kanje, M. Regeneration of the rat sciatic nerve into allografts made acellular through chemical extraction. *Brain Res.* **June 1998**, *795* (1–2), 44–54.
 89. Chiu, D.T.W.; Janecka, I.; Krizek, T.J.; Wolff, M.; Lovelace, R.E. Autogenous vein graft as a conduit for nerve regeneration. *Surg.* **1982**, *91*, 226–233.
 90. Strauch, B.; Ferder, M.; Lovelle-Allen, S.; Moore, K.; Kim, D.J.; Llena, J. Determining the maximal length of a vein conduit used as an interposition graft for nerve regeneration. *J. Reconstr. Microsurg.* **November 1996**, *12* (8), 521–527.
 91. Lundborg, G.; Dahlin, L.B.; Danielsen, N.P.; et al. Reorganization and orientation of regenerating nerve fibres, perineurium, and epineurium in preformed mesothelial tubes. An experimental study on the sciatic nerve of rats. *J. Neurosci. Res.* **1981**, *6*, 265.
 92. Herndon, J.H. Neuromas. In *Operative Hand Surgery*, 2nd Ed.; Green, D.P., Ed.; Churchill Livingstone: New York, 1988; 1405–1422.
 93. Lundborg, G.; Kanje, M. Bioartificial nerve grafts. A prototype. *Scand. J. Plast. Reconstr. Surg. Hand Surg.* **June 1996**, *30* (2), 105–110.
 94. Strauch, B. Use of nerve conduits in peripheral nerve repair. *Hand Clin.* **February 2000**, *16* (1), 123–130.
 95. Ellis, D.L.; Yannas, I.V. Recent advances in tissue synthesis in vivo by use of collagen-glycosaminoglycan copolymers. *Biomaterials* **February 1996**, *17* (3), 291–299.
 96. Integra. <http://www.integra-ls.com/products/?product=88>.
 97. Mackinnon, S.E.; Dellon, A.L. A study of nerve regeneration across synthetic (Maxon) and biological (collagen) nerve conduits for nerve gaps up to 5 cm in the primate. *J. Reconstr. Microsurg.* **1990**, *6*, 117–121.
 98. Woerly, S.; Plant, G.W.; Harvey, A.R. Neural tissue engineering: From polymer to biohybrid organs. *Biomaterials* **February 1996**, *17* (3), 301–310.
 99. Jenq, C.-B.; Coggeshall, R.E. Nerve regeneration through holey silicone tubes. *Brain Res.* **1985**, *361*, 233–241.
 100. Rich, K.M.; Alexander, T.D.; Pryor, J.C.; Hollowell, J.P. Nerve growth factor enhances regeneration through silicone chambers. *Exper. Neurol.* **1989**, *105*, 162–170.
 101. Borkenhagen, B.; Stoll, R.C.; Neuenschwander, P.; Suter, U.W.; Aebischer, P. In vivo performance of a new biodegradable polyesterurethane system used as a nerve guidance channel. *Biomaterials* **1998**, *19* (23), 2154–2166.
 102. Merle, M.; Dellon, A.; Campbell, J.; Chang, P. Complications from silicon-polymer intubulation of nerves. *Microsurg.* **1989**, *10*, 130–133.
 103. Archibald, S.J.; Krarup, C.; Shefner, J.; Li, S.-T.; Madison, R.D. A collagen-based nerve guide conduit for peripheral nerve repair: An electrophysiological study of nerve regeneration in rodents and nonhuman primates. *J. Compar. Neurol.* **1991**, *306*, 685–696.
 104. Kline, D.G.; Hayes, G.J. The use of a resorbable wrapper for peripheral nerve repair. *J. Neurosurg.* **1964**, *21*, 737.
 105. Crawley, W.A.; Dellon, A.L. Inferior alveolar nerve reconstruction with a polyglycolic acid bioabsorbable nerve conduit. *Plast. Reconstr. Surg.* **1992**, *90*, 300–302.
 106. Rosen, J.M.; Padilla, J.A.; Nguyen, K.D.; Siedman, J.; Pham, H.N. Artificial nerve graft using glycolide trimethylene carbonate as a nerve conduit filled with collagen compared to sutured autograft in a rat model. *J. Rehabil. Res. Devel.* **1992**, *29* (2), 1–12.
 107. den Dunnen, W.F.A.; van der Lei, B.; Robinson, P.H.; Holwerda, A.; Pennings, A.J.; Schakenraad, J.M. Biological performance of a degradable poly(lactic acid-ε-caprolactone) nerve guide: Influence of tube dimensions. *J. Biomed. Mater. Res.* **1995**, *29*, 757–766.
 108. Satou, T.; Nishida, S.; Hiruma, S.; Tanji, K.; Takahashi, M.; Fujita, S.; Mizuhara, Y.; Akai, F.; Hashimoto, S. A morphological study on the effects of collagen gel matrix on regeneration of severed rat sciatic nerve in silicone tubes. *Acta Pathol. Jpn.* **1986**, *36* (2), 199–208.
 109. Miller, E.J.; Gay, E.S. The Collagens: an overview and update. *Methods in Enzymol.* **1987**, *144*, 4–40.
 110. Seyer, J.M.; Kang, A.H.; Whitaker, J.N. The characterization of type I and type III of collagens from human peripheral nerve. *Biochim. Biophys. Acta* **1977**, *492*, 415–425.
 111. Junqueira, L.C.U.; Montes, G.S.; Bezerra, M.S.F. Do Schwann cells produce collagen type III? *Experientia* **1978**, *35*, 114.
 112. Sabelman, E.E. Biology, Biotechnology and Biocompatibility of Collagen. In *Biocompatibility of Tissue Analogs*, Vol. 1; Williams, D.F., Ed.; CRC Press: Boca Raton, Florida, 1985; 27–66.



113. Reiser, K.; McCormick, R.J.; Rucker, R.B. Enzymatic and nonenzymatic cross-linking of collagen and elastin. *FASEB J.* **1992**, *6* (7), 2439–2449.
114. Madison, R.; Da Silva, C.; Dikkes, P. Entubulation repair with protein additives increases the maximum nerve gap distance successfully bridged with tubular prostheses. *Brain Res.* **1988**, *447*, 325–334.
115. Rosen, J.M.; Padilla, J.A.; Nguyen, K.D.; Padilla, M.A.; Sabelman, E.E.; Pham, H.N. Artificial Nerve graft using collagen as an extracellular matrix for nerve repair compared with sutured autograft in a rat model. *Annals of Plast. Surg.* **1990**, *25* (5), 375–387.
116. Satou, T.; Nishida, S.; Hiruma, S.; Tanji, K.; Takahashi, M.; Fujita, S.; Mizuhara, Y.; Akai, F.; Hashimoto, S. A morphological study on the effects of collagen gel matrix on regeneration of severed rat sciatic nerve in silicone tubes. *Acta Pathol. Jpn.* **1986**, *36* (2), 199–208.
117. Colin, W.; Donoff, R.B. Nerve regeneration through collagen tubes. *J. Dental Res.* **1984**, *63*, 987.
118. Takahashi, M.; Satou, T.; Hashimoto, S. Experimental in vivo regeneration of peripheral nerve axons and perineurium guided by resorbable collagen film. *Acta Pathol. Jpn.* **1988**, *38*, 489–502.
119. Liu, S.; Bodjarian, N.; Langlois, O.; Bonnard, A.S.; Boisset, N.; Peulvè, P.; Saôd, G.; Tadiè, M. Axonal regrowth through a collagen guidance channel bridging spinal cord to the avulsed C6 roots: functional recovery in primates with brachial plexus injury. *J. Neurosci. Res.* **March 1998**, *51* (6), 723–734.
120. Aguayo, A.J.; Bray, G.M.; Perkins, C.S.; Duncan, I.D. Axon-sheath cell interactions in peripheral and central nervous system transplants. *Soc. Neurosci. Symp.* **1979**, *4*, 361–383.
121. Griffiths, R.; Horch, K.; Stensaas, L. A collagen and fibrin tube for nerve repair. *Rest. Neuro. and Neurosci.* **1990**, *1*, 339–346.
122. Eccleston, P.A.; Mirsky, R.; Jessen, K.R. Type I collagen preparations inhibit DNA synthesis in glial cells of the peripheral nervous system. *Exp. Cell Res.* **1989**, *182*, 173–185.
123. Duby, N.; Letourneau, P.C.; Tranquillo, R.T. Enhanced neurite and Schwann cell invasion into magnetically-aligned collagen gel rods for peripheral nerve entubulation repair. *Trans. Soc. for Biomaterials—Symposium: Engineering Neural Tissue*, San Diego, California, **April 22–25, 1998**.
124. Rizvi, A.H.; Wasserman, A.J.; Zazanis, G.; Silver, F.H. Evaluation of peripheral nerve regeneration in the presence of longitudinally aligned collagen fibers. *Cells & Materials* **1991**, *1*, 279–289.
125. Hudson, R.; Evans, G.R.D.; Schmidt, C. Engineering strategies for peripheral nerve repair. In *Clinics Plastic Surgery*, Vol. 31, No. 3; Edington, H., Ed.; W. B. Saunders: Philadelphia, 2000; 485–497. Reprinted in *Orthopaedic Clinics of North America*.
126. Anselin, A.D.; Fink, T.; Davey, D.F. Peripheral nerve regeneration through nerve guides seeded with adult Schwann cells. *Neuropathol. Appl. Neurobiol.* **October 1997**, *23* (5), 387–398.
127. Sittinger, M.; Bujia, J.; Rotter, N.; Reitzel, D.; Minuth, W.W.; Burmester, G.R. Tissue engineering and autologous transplant formation: practical approaches with resorbable biomaterials and new cell culture techniques. *Biomaterials* **February 1996**, *17* (3), 237–242.
128. Levi, A.D.O.; Guénard, V.; Aebischer, P.; Bunge, R. The functional characteristics of Schwann cells cultured from the human peripheral nerve after transplantation into a gap within the rat sciatic nerve. *J. Neurosci.* **1994**, *14*, 1309–1319.
129. Sørensen, J.; Haase, G.; Krarup, C.; Gilgenkrantz, H.; Kahn, A.; Schmalbruch, H. Gene transfer to Schwann cells after peripheral nerve injury: a delivery system for therapeutic agents. *Ann. Neurol.* **February 1998**, *43* (2), 205–211.
130. Casella, G.T.; Bunge, R.P.; Wood, P.M. Improved method for harvesting human Schwann cells from mature peripheral nerve and expansion in vitro. *Glia* **1996**, *17*, 327.
131. Moretto, G.; Kim, S.U.; Shin, D.H.; Pleasure, D.E.; Rizzuro, N. Long-term cultures of human adult Schwann cells isolated from autopsy materials. *Acta Neuropathol. (Berlin)* **1984**, *64*, 15–21.
132. Livesey, F.J.; O'Brien, J.A.; Li, M.; Smith, A.G.; Murphy, L.J.; Hunt, S.P. A Schwann cell mitogen accompanying regeneration of motor neurons. *Nature* **December 1997**, *390*, 614–618.
133. Sondell, M.; Fex-Svenningsen, A.; Kanje, M. The insulin-like growth factors I and II stimulate proliferation of different types of Schwann cells. *Neuroreport* **September 1997**, *8* (13), 2871–2876.
134. Brockes, J.P. Studies of cultured rat Schwann cells. Establishment of purified populations from cultures of peripheral nerve. *Brain Res.* **1979**, *165*, 105–118.
135. Askanas, V.; Engel, W.K.; Dalakas, M.C.; Lawrence, J.V.; Carter, L.S. Human Schwann cells in tissue culture. Histochemical and ultrastructural studies. *Arch. Neurol.* **1980**, *37*, 329–337.
136. Hardman, P.; Spooner, B.S.; Klement, B.J. A technique for promoting Schwann cell growth from fresh and frozen biopsy nerve utilizing D-valine medium (letter). *In vitro Cell. & Dev. Biol.* **1990**, *12* (26), 1116–1117.
137. Rosen, J.M.; Kaplan, E.N.; Jewett, D.L. Suture and Sutureless Methods of Repairing Experimental Nerve Injuries. In *Nerve Repair and Regeneration, Its Clinical and Experimental Basis*; Jewett, D.L., McCarrroll, H.R., Jr. Eds.; C.V. Mosby: St. Louis, 1980; 235–242.
138. Rosen, J.M.; Pham, H.N.; Hentz, V.R. Fascicular tubulization: A comparison of experimental nerve repair techniques in the cat. *Ann. Plas. Surg.* **1989**, *22* (6), 467–468.
139. Bain, J.R.; Mackinnon, S.E.; Hunter, D.A. Functional evaluation of complete sciatic peroneal and posterior tibial nerve lesions in the rat. *Plast. and Reconstr. Surg.* **1989**, *3* (1), 129–136.
140. Keeley, R.D.; Sabelman, E.; Kadlcik, P.; Agras, J.; Eng, L.; Wiedman, T-W.; Nguyen, K.; Sudekum, A.; Rosen, J. Synthetic nerve graft containing collagen and syngeneic Schwann cells improves functional,

- electrophysiological and histological parameters of peripheral nerve regeneration. *Restorative Neurol. and Neurosci.* **1993**, *5*, 353–366.
141. Eilert, J.B.; Binder, P.; McKinney, P.W.; et al. Polyglycolic acid synthetic absorbable sutures. *Am. J. Surg.* **1971**, *121*, 561.
142. Jeffrey, J. Collagen Degradation. In *Wound Healing: Biochemical and Clinical Aspects*; Cohen, I.K., Diegelmann, R.F., Lindblad, W.J., Eds.; W. B. Saunders: Philadelphia, 1992; 177–194.
143. Roe, S.C.; Milthorpe, B.K.; Schindhelm, K. Collagen cross-linking and resorption: effect of glutaraldehyde concentration. *Artificial Organs* **1990**, *14*, 443–448.
144. Chiu, D.T.W.; Lovelace, R.E.; Yu, L.I.; Wolff, M.; Stengel, S.; Middleton, L.; Janecka, I.P.; Krizek, T.J. Comparative electrophysiologic evaluation of nerve grafts and vein grafts as nerve conduits. *Surg. Forum* **1983**, *34*, 601.
145. Lundborg, G. A 25-year perspective of peripheral nerve surgery: evolving neuroscientific concepts and clinical significance. *J. Hand Surg.* **2000**, *25A*, 391–414.
146. Ohara, K.; Doi, K.; Sakai, K.; Kuwata, N.; Kawai, S. Restoration of sensibility in the hand after complete brachial plexus injury. *J. Hand Surg. (Am.)* **May 1996**, *21* (3), 381–386.
147. Bertelli, J.A.; Taleb, M.; Mira, J.C.; dos Santos, A.R.; Calixto, J.B.; Kassir, L. Selective restoration of sensation by peripheral nerve grafts directly implanted into the contralateral C7 dorsal root ganglion: an experimental study in rat brachial plexus. *Neurosurg.* **January 1998**, *42* (1), 125–129.
148. Jacobs, J.M.; Laing, J.H.; Harrison, D.H. Regeneration through a long nerve graft used in the correction of facial palsy. A qualitative and quantitative study. *Brain* **February 1996**, *119* (Pt. 1), 271–279.
149. Kozek, J.; Voska, P.; Tiché, M. Contemporary state of surgical treatment of facial nerve paresis. Preliminary experience with new procedures. *Acta Chir. Plast.* **1997**, *39* (4), 125–131.
150. Moissonnier, P.; Reviron, T.; Ye, J.H.; Horvat, J.C. Motoneurons of the injured spinal cord of the adult dog can grow lengthy axons into an autologous peripheral nerve graft. A retrograde axonal tracing study. *Spinal Cord* **1996**, *34* (6), 320–325.
151. Gage, F.H. Method of Grafting Genetically Modified Cells to Treat Defects, Disease or Damage of the Central Nervous System. US Patent 5,650,148, July 22, 1997.
152. Stichel, C.C.; Lips, K.; Wunderlich, G.; Müller, H.W. Reconstruction of transected postcommissural fornix in adult rat by Schwann cell suspension grafts. *Exp. Neurol.* **July 1996**, *140* (1), 21–36.
153. Date, I.; Imaoka, T.; Miyoshi, Y.; Ono, T.; Asari, S.; Ohmoto, T. Chromaffin cell survival and host dopaminergic fiber recovery in a patient with Parkinson's disease treated by cogafts of adrenal medulla and pretransected peripheral nerve. Case report. *J. Neurosurg.* **April 1996**, *84* (4), 685–689.
154. Wirth, E.D.; Fessler, R.G.; Reier, P.J.; Thompson, F.J.; Uthman, B.; Behrman, A.; Remson, E.J.; Anderson, D.K. Neural tissue transplantation in syringomyelia patients: feasibility and safety. *Amer. Soc. for Neural Transplantation*, 5th Annual Conference, 1998.
155. Romano, V.M.; Blair, S.J.; Wurster, R.D. Rerouting peripheral nerves for spinal cord lesions. *J. Orthop. Res.* **1991**, *9*, 54–60.
156. Eberhard, D.; Millesi, H. Split nerve grafting. *J. Reconstr. Microsurg.* **February 1996**, *12* (2), 71–76.

Phosphorylcholine (PC) Technology

Andrew L. Lewis

Drug Delivery Division, Biocompatibles UK Limited, Farnham, Surrey, United Kingdom

INTRODUCTION

Terms such as biomimicry and bio-inspired have been used to describe systems that have been designed specifically by the application of principles taken from observing nature. PC Technology™ is one example of such a system. It encompasses a range of biomaterials, the chemical compositions of which are founded upon the structure of the cell membrane.

OVERVIEW

Although the lipid bilayer assembly of plasma membranes has been known for some time, it was not appreciated until the early 1970s that the lipid components composing the outer and inner halves of the bilayer are very different. It was Zwaal and coworkers who first reported in *Nature* the fact that if the bilayer is separated into its inner and outer leaflets and then individually exposed to blood, the outer layer does not cause the blood to clot, whereas the inner layer is thrombogenic.^[1] This is explained in terms of the predominant phospholipids in each layer, the outer leaflet composed mainly of those with zwitterionic phosphorylcholine (PC) headgroups (Fig. 1), as opposed to the larger proportion of negatively-charged phosphorylserine-based phospholipids found on the inside. This observation led to the seminal work of the late Professor Dennis Chapman, who demonstrated that surfaces functionalized with the PC headgroup could be rendered non-thrombogenic.^[2-6] This led Chapman to found a company in 1984 called Biocompatibles in order to develop and commercialize a range of materials and surface treatments based on PC Technology that would bring significant benefit to devices used in medicine by improving their tolerance within the body.

PC POLYMERS—EVOLUTION AND SYNTHETIC PERSPECTIVE

Early methods of functionalizing surfaces with PC included direct coupling of reactive small-molecule PC species to the desired substrate. Although the biocompatibility of the surfaces was improved, these techniques

generally required complex pretreatment steps, were inconsistent, and often were not practically viable for medical device manufacture. The first polymer-based approach was achieved using an analogue of the naturally occurring lipid dipalmitoylphosphatidylcholine (DPPC, Fig. 1) that possessed acetylenic bonds within the alkyl chains.^[7] This diacetylenic phosphatidylcholine (DAPC) could be polymerized onto a surface by use of UV or γ -irradiation, which induced cross-linking by addition across the triple bonds. The resultant coating was stable and biocompatible, but again the method of application was fraught with handling difficulties. Other polymer systems investigated included PC-based polyesters and polyurethanes made by the direct polycondensation of the PC diol glycerophosphorylcholine with appropriate diacid chlorides or diisocyanates. Although the biological properties of such materials were enhanced, the presence of the PC group resulted in a significant detrimental effect on the mechanical properties of these types of polymer.

The development in Japan of the PC methacrylate monomer, 2-methacryloyloxyethyl phosphorylcholine (MPC, Fig. 2), opened the door to a whole new area of flexible chemistry that utilizes free-radical polymerization. An improved route to the monomer synthesis allowed multigram preparations to be achieved and hence provided potential commercial viability. The monomer is prepared by the coupling of 2-hydroxyethyl methacrylate, a commonly used material in the medical industry, with 2-chloro-2-oxo-1,3,2-dioxaphospholane, followed by a ring opening reaction of the intermediate alkoxyphospholane with trimethylamine. A large family of polymers could now be prepared via a variety of free-radical techniques including:

- Conventional one-shot or monomer-starved solution methods for the preparation of linear and cross-linkable coating polymers.^[8,9]
- Specialized techniques such as atom transfer radical polymerization (ATRP) to prepare controlled architectures such as block copolymers, narrow polydispersities, and macromonomer synthesis.^[10-12]
- Emulsion methods for water-based PC latexes and suspension polymerization for PC-based microsphere fabrication.
- Bulk polymerizations for molded articles such as contact lenses.^[13,14]

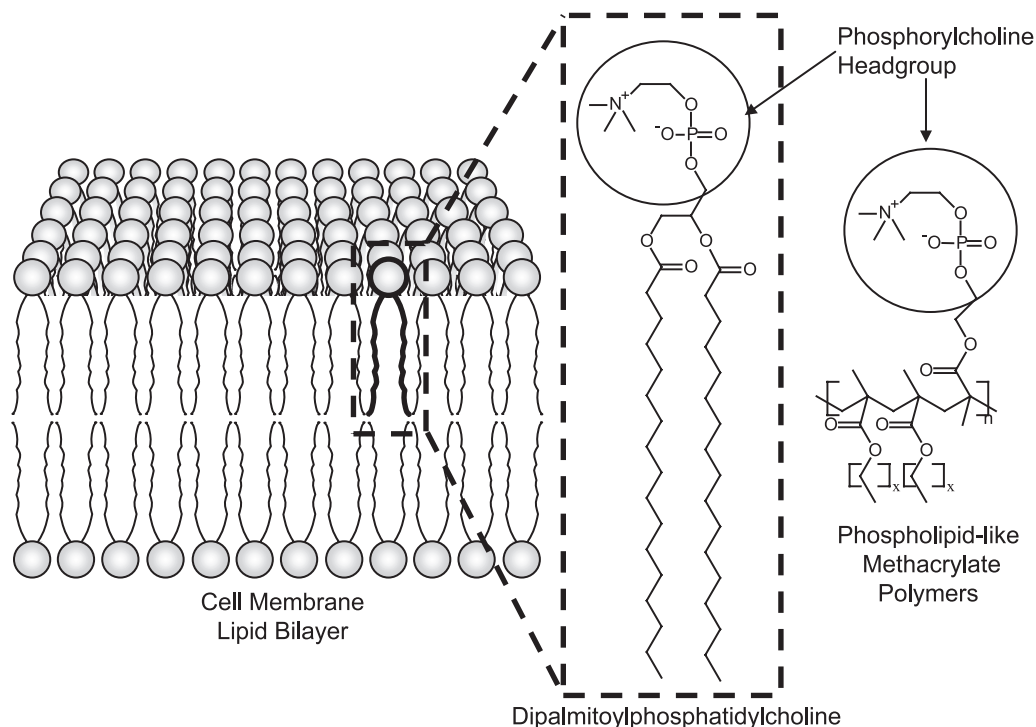


Fig. 1 Cell membrane, structure of a typical phospholipid, and PC polymer analogues.

One of the first PC polymers to be synthesized using these methods was a copolymer with *n*-dodecylmethacrylate (MPC-*co*-DMA (1:2), Fig. 2) which, with two long-chain alkyl monomers for every one MPC unit, is a close structural analogue of DPPC (Fig. 1). This material is amphiphilic in nature due to the highly polar PC headgroups and the hydrophobic alkyl components, and requires the use of a subtle mixture of solvents during synthesis via a solution method due to the differences in monomer solubilities. The hydrophobic portion allows stable coatings to be formed on various substrates such as glass, metals, and a variety of plastics, by virtue of physi-adsorption through hydrophobic interaction. The PC component attracts a high water content (~60%) into the coating to form a hydrogel with remarkable hemocompatibility.^[15]

For some applications, particularly in which a change in shape or dimension of the device may ensue (for instance, in the case of an expandable stent), it is desirable to have a more mechanically robust coating. A cross-linkable PC polymer was developed for this purpose (XL-MPC-*co*-DMA, or PC1036); the polymer is synthesized and isolated from solution as a linear polymer, but with silyl groups pendent from the polymer backbone that provide a method of cross-linking after the coating has been applied.^[9] The extent of curing can be controlled by thermal treatment of the coating and further enhanced by use of a γ -irradiation cycle typical of that employed in the sterilization of

medical devices. Variations on this polymer have also been developed; they contain a cationic component that allows the coating to interact with species that carry a net negative charge, such as heparin and DNA (PC1062 & PC2028, Fig. 2). These coatings can be applied to devices where antithrombogenicity (as opposed to non-thrombogenicity) is required, or in which the ability to deliver locally large biomolecules, such as genetic material, is required.^[16]

In the case of molded articles such as soft contact lenses, adhesion to a substrate is not a requirement and the hydrophobic component of the polymer is replaced by a suitable comonomer that acts as a solvent for the MPC during the bulk polymerization.^[13,14] In the case of the Proclear[®] contact lens (PC1015, Fig. 2), the comonomer is 2-hydroxyethyl methacrylate which, together with a small amount of cross-linker, produces a 59% water content hydrogel (omafilcon A) that has excellent transparency and oxygen permeability necessary for contact lens fabrication.^[17] The lens can be produced from polymer molding buttons which are individually lathed into the various lens powers, or more commonly nowadays by direct cast molding in which the monomer formulation is polymerized directly within a mold that creates the desired shape of the lens. For applications where only a surface treatment is necessary, in-mold coating (IMC) techniques allow for the application of suitably functionalized PC polymers to the inside surfaces of a mold, after which the mold is filled with an

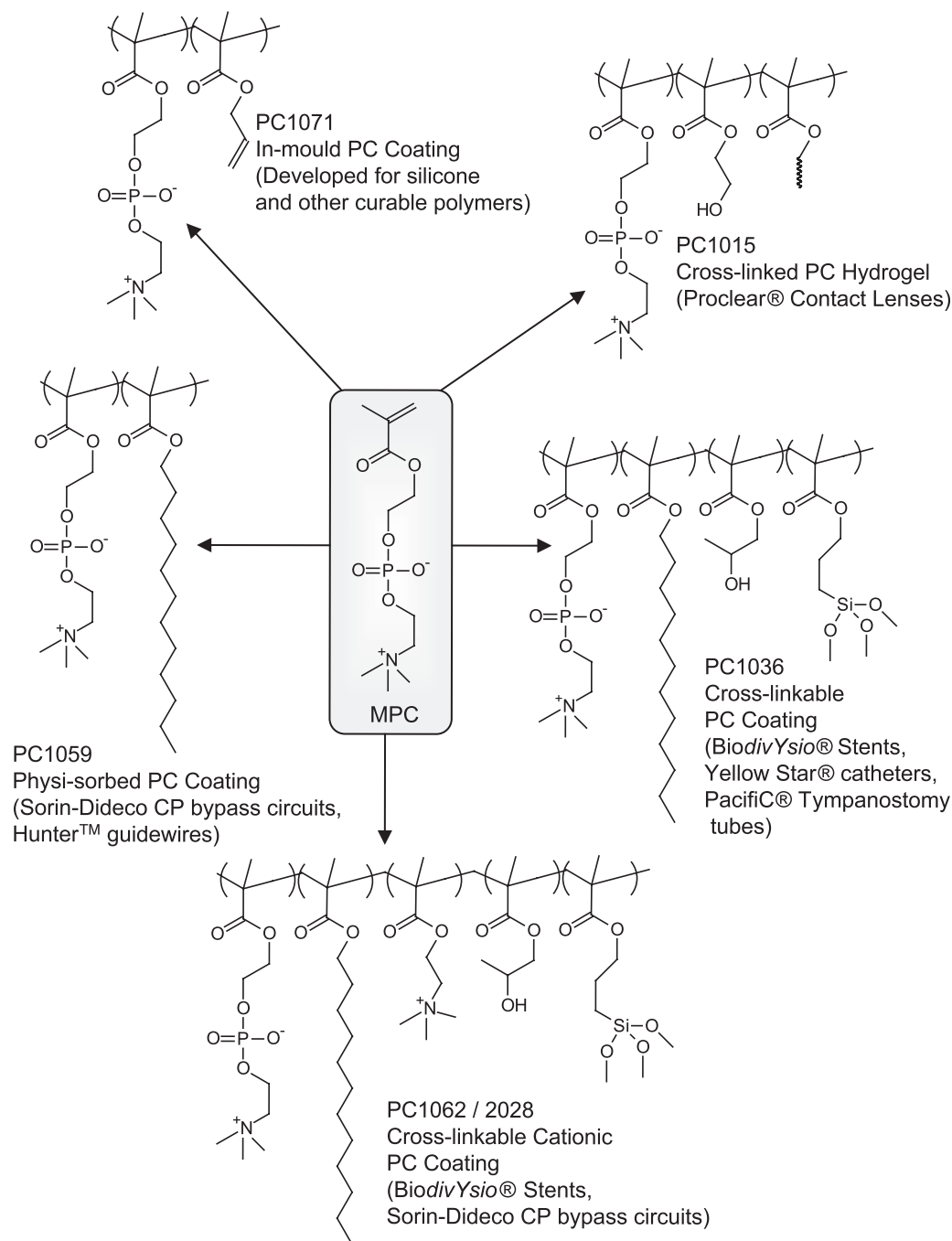


Fig. 2 Structure of the MPC monomer building block and various polymers that have been developed.

appropriate polymer formulation and cured.^[18] As the functional groups within the mold react, the reactive functionalities of the PC polymer become grafted into the surface of the forming article, producing a highly stable, wettable and biocompatible layer. The use of the allylic functionality is particularly useful (PC1071, Fig. 2) and can dramatically alter the surface characteristics of materials such as silicone that are classically very hydrophobic in the absence of any surface modification.

These few examples serve only to demonstrate the flexibility of the polymerization technology and by no means form an exhaustive list of the range of polymers that fall under the PC Technology umbrella. With the advent of new radical polymerization techniques such as ATRP and RAFT that allow more control over polymer architectures, there will be further advances in novel PC polymer synthesis to report in the near future.

BIOPHYSICAL CHARACTERIZATION

Surface Mobility and Swelling Properties

PC coating polymers that are based on lipidlike structures are amphiphilic due to the combination of the polar PC moieties and hydrophobic alkyl chain components. Solutions of these materials in suitable solvents can be easily applied to surfaces via any of the conventional techniques such as dip, spin, or spray coating. As the solvent evaporates from the forming film, the polymer will orientate its hydrophobic groups at the air interface in order to minimize the interfacial free energy. This has been demonstrated by the use of TOF-SIMS which shows that the PC groups are buried within the film when exposed to air or high-vacuum conditions.^[8,19] Once exposed to an aqueous environment, the coating rapidly rearranges so that the PC groups are expressed at the surface, a phenomenon demonstrated in a number of studies using contact angle measurement. These properties are important during consideration of the intended application of the coating, as a spontaneously wettable surface may be desired which requires re-engineering of the PC polymer to avoid the hysteresis effect.

Studies using spectroscopic ellipsometry and neutron diffraction demonstrate that the PC films swell very rapidly when exposed to water (some 50% of water is absorbed within the first minute).^[20,21] The kinetics of swelling and distribution of water within the coating is dependent upon whether the coating is the noncross-linked or cross-linked type. For the latter case, the swelling follows a two-stage mechanism involving a rapid, essentially diffusion-controlled ingress of water, followed by a much slower phase controlled by relaxation of the polymer chains. The hydrophilic:hydrophobic group balance and the extent of cross-linking of the polymer network dominate the swelling properties. Various estimates have placed the number of water molecules directly associated with each PC headgroup in the region of 12–18. These molecules are bound tightly to the PC; it is this attribute that provides the PC hydrogels with a level of dehydration resistance that makes them particularly suitable as contact lens materials, where on-eye dehydration can be associated with discomfort to the wearer.

Coating Stability

The stability of the PC coatings is obviously an important feature of their properties. For the simple physisorbed copolymer materials, studies in which part of the alkyl chain of the hydrophobic portion was radiolabeled with tritium showed that, once coated onto a substrate, the film was stable and did not leach.^[7] For the

cross-linkable polymers, atomic force microscopy techniques were used to demonstrate the increased force required to remove the coating from a substrate compared to the noncross-linked systems, and to show how the force of removal varied with curing and sterilization regimes. An adapted nanoindentation method also showed that the low level of cross-linking gave the coating some slightly elastic property; this ability to allow the coating to recover following deformation is particularly useful with respect to expandable devices such as stents.^[9] The stability of the coating on the *BioDivYsio* coronary stent was studied in vivo both in a porcine coronary model and on a stent removed from a human subject. Both studies showed the coating was still present, required the same force to remove, and had the same thickness, even after six months' implantation.^[23,24]

Drug Delivery Capability

The cross-linked networks formed in polymers such as PC1036 (see Fig. 2) lend themselves perfectly to drug delivery.^[25,26] Although the polymer coating is not macroporous in nature, the interstitial spaces between polymer chains are filled with imbibed water and provide a medium for drug loading, and ultimately a route for release of the therapeutic agent from the coating. Measurements made from fully cured membranes cast from this polymer indicate that the polymer has an effective molecular weight cut-off around 1200 daltons, a figure confirmed by loading and release studies of different molecular weight compounds from coronary stent coatings of this polymer.^[16,26] Furthermore, the coating has been investigated using fluorescent probe analysis to confirm the presence of hydrophobic domains within the hydrated polymer. Pyrene has a fused ring structure that gives rise to a characteristic fluorescent spectrum in which particular bands change in intensity when the molecule is in different environments. A very low level of pyrene in water gives a particular spectrum indicative of a polar environment, compared to that when the pyrene is allowed to partition from the water into the coating, during which the spectrum changes to one that is characteristic of a hydrophobic surrounding. This showed that there are hydrophobic domains within the hydrogel with which hydrophobic drugs could potentially interact and hence provide a method of controlled delivery from the coating. In point of fact, studies involving the loading and release of a range of different therapeutic agents from the PC-coated stent showed that there was a relationship between hydrophobicity and rate of release that was not explained by solubilities of the drugs alone. A weak correlation between the partition coefficient of the drug between octanol and water ($\log P_{(o/w)}$)



and the time for 90% release of the agent to be achieved was found to exist, supporting the hypothesis that the more hydrophobic compounds (higher log $P_{(o/w)}$) interact with the hydrophobic domains in the coating and hence retard their release.

As the PC1036-based coatings had a molecular weight limitation, it was necessary to develop other coatings that would allow the delivery of higher-molecular-weight species. Of particular interest for coronary stent purposes are large biomolecules such as heparin and DNA fragments. Such species carry a net negative charge, and so cationic PC coatings such as PC1062 and PC2028 were developed that are still highly biocompatible but capable of interacting with the biomacromolecules by surface ionic attraction.^[16]

In Vitro Bioevaluation

The extent of in vitro biological evaluation of various PC polymers is substantial^[27] (see Tables 1 and 2).

Many studies have focused on the first steps of any biofouling process, namely, the adsorption of proteins to the polymer surfaces. Any medical device that is placed on or within the body will illicit a response from the host, and the first test for foreign articles in this process of nonself-recognition is their interaction with proteins. When presented with a foreign surface, the proteins may change their three-dimensional shape in order to adopt a more energetically favorable conformation. This in turn exposes chemical moieties within the protein structure that are not normally accessible, and the growing protein layer takes on a higher level of biological activity which can ultimately result in adverse immunological and inflammatory responses. When proteins contact PC-based surfaces they are met by a layer of bound water molecules which act as a barrier to the permanent adhesion of the molecule. The use of circular dichroism spectroscopy has confirmed that the conformation of the protein remains unchanged, there being no thermodynamic drive for the structure to unfold, and consequently PC surfaces are largely protein-resistant.^[28] By addressing this

Table 1 Selection of protein adsorption studies on PC materials and their methods of evaluation

Study	Techniques used	Polymer system	References
<i>Protein interactions</i>			
Fibrinogen (Fg)	Enzyme immunoassay (EIA)	DAPC	[7]
Fg	EIA	DAPC	[29]
Human serum albumin (HSA)			
Immunoglobulin G (IgG)			
von Willebrand Factor (vWF)			
Fg	EIA	MPC-co-DMA	[15]
Hageman factor (FXII)			
BSA	Ellipsometry	MPC-co-DMA	[30,31]
Lysozyme	Neutron reflection		
Fg	Surface plasmon resonance	MPC-co-DMA	[19]
Fg HSA	Surface plasmon resonance	MPC-co-DMA blends	[32,33]
Fg	Ellipsometry	MPC-co-DMA	[34]
BSA	Neutron reflection	XL-MPC-co-DMA	
Lysozyme			
Fg	EIA	XL-MPC-co-DMA	[9]
Total protein	micro-BCA protein assay		
Fg	EIA	MPC-HEMA hydrogel	[35,36]
Fg	Bright field microscopy + image analysis	MPC-HEMA hydrogel	[37]
Total tear protein (in vitro and in vivo)	Spectrofluorimetry	Proclear™ (MPC-HEMA hydrogel)	[38,39]
Lysozyme	UV methods		
Total tear protein	Spectrofluorimetry	PC-IMC-lens	[18]

Table 2 Selection of cell and bacterial adhesion studies on PC materials and their methods of evaluation

Study	Techniques used	Polymer system	References
<i>Cells & bacteria</i>			
Platelets	Adhesion by ATP assay, activation by P selectin immunoassay	DAPC	[7]
Platelets	Adhesion by ATP & GPIb immunoassay, activation P selectin immunoassay	MPC-co-DMA	[15]
Whole blood	Blood contact SEM	MPC-co-DMA	[8]
Whole blood	Blood contact SEM	XL-MPC-co-DMA	[9]
3T3 fibroblasts, Rabbit lens epithelia, Primary human macrophage, Human granulocytes, Murine fibroblasts	ATP luciferase bioluminescence assay	MPC-co-DMA	[40–43]
3T3 fibroblasts, Rabbit lens epithelia, Primary human macrophage, Giant cells	ATP luciferase bioluminescence assay	MPC-co-HEMA hydrogel	[35,36]
Diffuse cells	Rabbit eye implantation in vivo	MPC-co-DMA	[44]
3T3 fibroblasts Rabbit lens epithelia Primary human macrophage Human granulocytes	ATP luciferase bioluminescence assay	Proclear™ (MPC-HEMA hydrogel)	[45]
Rabbit corneal epithelia 3T3 fibroblasts	DAPI stain + fluorescent microscopy	XL-MPC-co-DMA	[46]
Bovine vascular endothelia			
<i>Staphylococcus epidermidis</i>	ATP luciferase bioluminescence assay	MPC-co-DMA	[40–43,47]
<i>Escherichia coli</i>	EIA	XL-MPC-co-DMA	[9]
<i>S. epidermidis</i>	ATP luciferase bioluminescence assay	MPC-co-HEMA hydrogel	[36,43]
<i>S. epidermidis</i> ,	ATP luciferase bioluminescence assay	Proclear™	[48]
<i>Pseudomonas aeruginosa</i> ,	Phase contrast microscopy + image analysis	(MPC-HEMA hydrogel)	
<i>Serratia marcescens</i>			

fundamental step in the cascade of reactions designed to rid the body of foreign bodies, PC can be used to improve the biotolerance of medical devices. Table 1 details work carried out on a range of proteins, using a wide variety of different techniques, in the analysis of their interactions with different PC-based polymer systems.^[7–9,15,18,19,29–39] In all of these studies without exception, it is shown that protein adhesion to a wide range of substrates is significantly decreased by modification of the surface with a PC polymer.

The interaction of cells with surfaces is largely controlled in the first instance by nonspecific protein adsorption. Receptors on cell surfaces rely upon the interaction with particular chemical sequences within certain proteins, which only become available when they alter their conformation. Once the cells are able to

anchor to the surface they can spread, divide, and generally orchestrate an attack against the foreign surface. By virtue of their ability to resist protein fouling, PC-based surfaces have been shown to reduce cellular adhesion of a number of different cell types in vitro^[27] (Table 2). Infection by certain bacteria is another process that may be controlled by reducing the extent of protein adhesion to a surface. For many types of bacteria, it is necessary to adhere to a surface in order to reproduce. In order to facilitate this, the organisms produce various excretions; surfaces that can resist the adhesion of these, thus preventing the bacteria from sticking, will have improved infection resistance. Table 2 outlines a number of studies that have reported successful reduction of bacterial adhesion to PC surfaces.^[7–9,15,29–48]

APPLICATION OF PC TECHNOLOGY

In the past, materials for use within the human body were selected mainly on the basis of the physical function they were to perform. With the aged population rising, there is an increasing emphasis on the duration over which these medical devices are required to serve, and thus greater focus on improving their tolerance by the recipient. This requires an understanding of the biological processes involved, with a focus on the interface between tissue and device. Over the past few decades, PC Technology has offered a proven strategy for the management of this interface in a number of different application areas.

Blood Contact

The bloodstream is a very unforgiving environment for any prospective device that must function within it. The hemostatic mechanism is responsible for minimizing the blood loss from any damaged blood vessel, but the process can educe undesirable events when the blood is presented with the artificial surface. There is a complex sequence of interdependent reactions that

begins with the adsorption of plasma proteins on the surface, and involves platelet adhesion and activation together with coagulation initiated from surface-mediated and extrinsic pathways, culminating in the formation of thrombus by cellular entrapment in an insoluble fibrin gel. In order to overcome these events, prior to any procedure involving blood-contacting devices, the anticoagulant heparin is administered, more commonly nowadays together with an anti-platelet regime. In addition to the extra cost involved, there are often problems associated with lowered platelet counts and bleeding. Devices that are therefore naturally resistant to the formation of thrombus have obvious advantages.

The *in vitro* studies described in the preceding section (see also Fig. 3) demonstrate that by mimicry of the red blood cell membrane, PC Technology is able to dramatically reduce the processes leading to thrombus formation. This exceptional hemocompatibility has led to its evaluation on a wide variety of devices used in cardiovascular applications. Initial studies of the technology were made using a variety of *ex vivo* models in which flowing blood was passed through various PC-coated tubing, with significant reduction in thrombus and platelet adhesion.^[49–52] Based on these results,

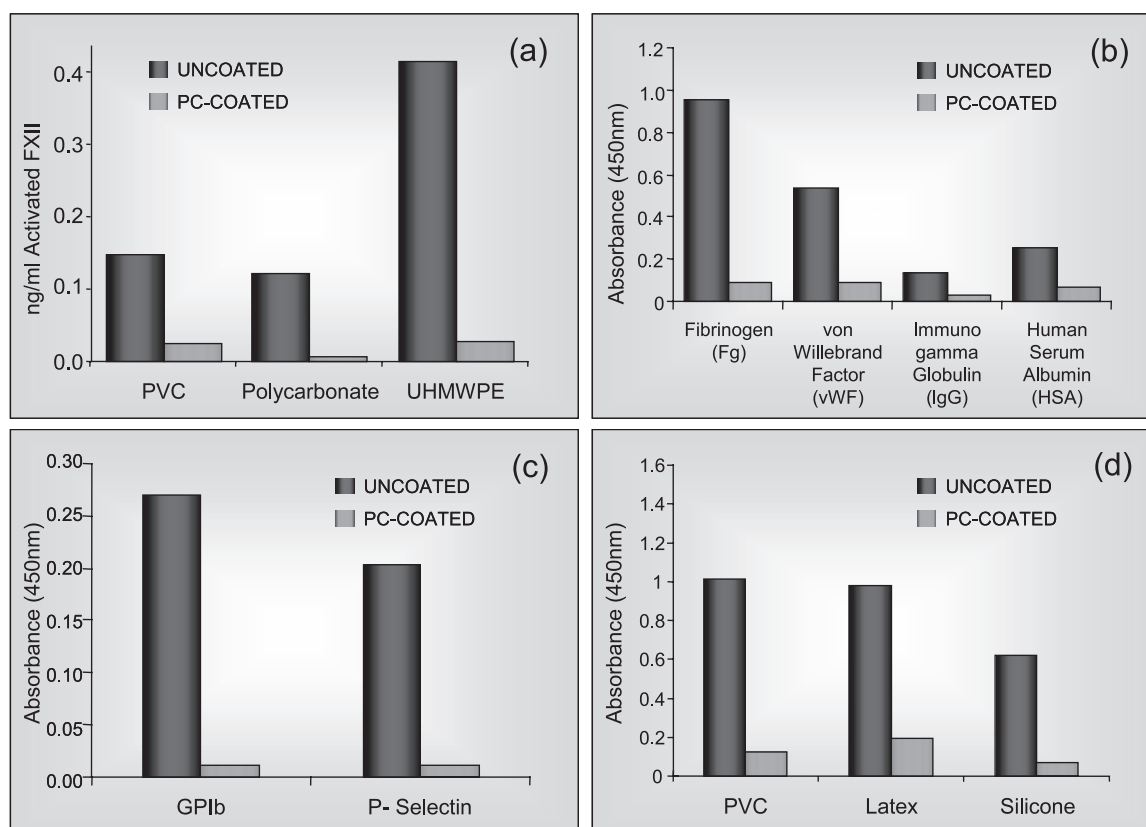


Fig. 3 In vitro performance of PC coatings: (a) reduction of contact activation; (b) reduced protein adhesion; (c) reduced platelet adhesion and activation; (d) reduced bacterial adhesion (*E. coli*). (From Refs. [9] and [15].)

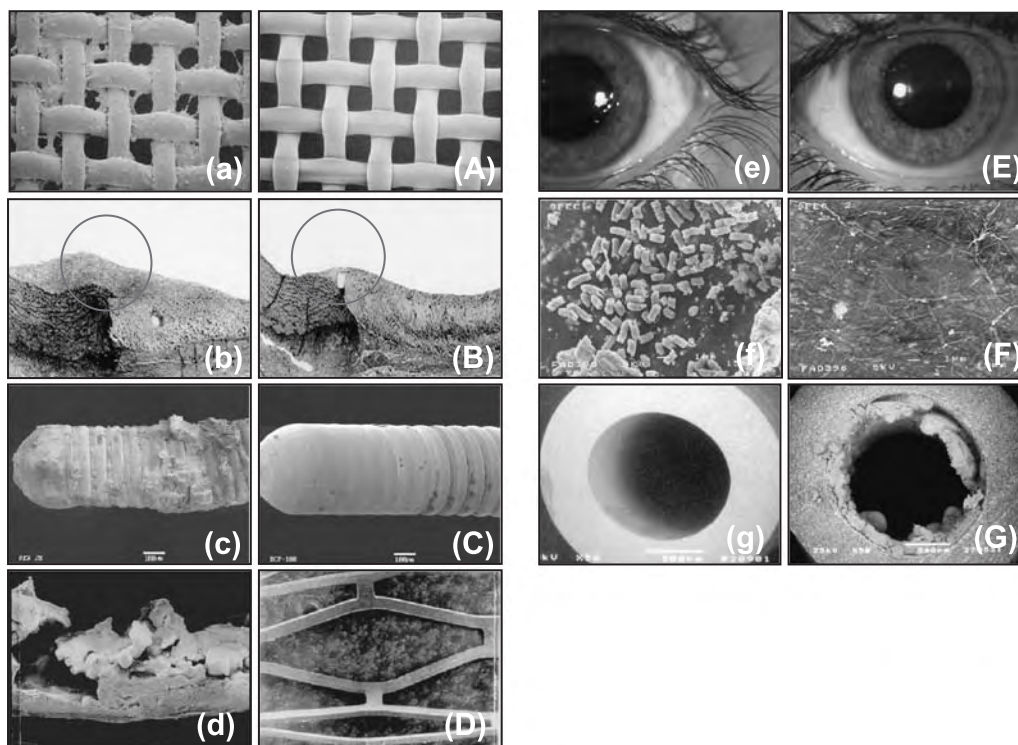


Fig. 4 The problems with medical devices. Blood contacting applications: uncoated extracorporeal circuit blood filter after use (a), compared to a PC-coated filter (A); the anastomoses of a PTFE graft and the related intimal hyperplasia (circled) for uncoated (b) and PC-coated grafts (B) in a canine model; tip of an uncoated PTCA guidewire after a 28-minute procedure (c) and the PC-coated guidewire after a 108-minute procedure (C); uncoated coronary stent from an ex vivo shunt study covered in thrombus (d), compared to the PC-coated stent in the same study (D). Elsewhere in the body: a poorly wetting soft contact lens (e) compared to a PC-containing soft contact lens (E); bacterial adhesion to an uncoated latex surface (*E. coli*) (f), compared to the PC-coated latex surface (F); a PC-coated ureteric catheter after use in a patient (g), compared to an uncoated catheter from the same patient (G).

PC-coated thoracic drain catheters were the first of the various temporary in-dwelling devices to show a significant clinical advantage.^[53] In the years since these early successes, the technology has been applied to ever-more-complex devices, and is currently licensed to Sorin-Dideco for use on the entire circuitry of their cardiopulmonary bypass systems, including components such as the oxygenators, heat exchangers, tubing, filters (see Fig. 4a,A), and reservoirs. This serves to illustrate the flexibility of the coating, as these components are made from a variety of different materials including stainless steel, polypropylene, polycarbonate, PVC, and epoxy resin. Clinical data from the use of such devices show that up to 30% more blood is preserved, and many of the critical blood factors are significantly less activated, when a PC-coated circuit is employed.^[54] Moreover, the use of fully-coated circuits allows for reduced systemic heparinization of the patient, resulting in improved platelet count and lower post-operative bleeding.^[55] Future developments in this area include the use of a second-generation PC coating that has the capability of binding heparin to its surface

(see Fig. 2) in order to address areas of extreme hemodynamic flow which induce thrombosis regardless of surface contact.

Vascular grafts are another device in which potential benefit has been demonstrated by application of PC Technology. These are essentially synthetic blood vessels, commonly made of expanded poly(tetrafluoroethylene) (ex-PTFE) or Dacron (poly(ethylene terephthalate), PET), that are used to create a bypass around diseased arteries or replace damaged vessels. There are currently no suitable materials for use in the manufacture of narrow-bore vascular grafts (less than 3 mm), as conventional materials thrombose and block. Use of PC in combination with such devices is shown to be of benefit, not only from the nonthrombogenicity bestowed upon the device,^[27] but also from the fact that studies in a canine model showed reduced intimal hyperplasia at the anastomoses where synthetic graft is sutured to the tissue^[56] (see Fig. 4b,B).

Percutaneous coronary angioplasty (PTCA) is a minimal access procedure that is used to unblock the arteries of the heart without the need for bypass surgery.

It involves making a small incision in the leg or arm and, with the aid of low-intensity X-rays, advancing a steerable guidewire through the vasculature and into the affected coronary artery. The guidewire is passed through the narrowed region (known as the stenosis) and then a balloon catheter is advanced into position over the guidewire. Once in place within the stenosis, the balloon is filled with saline to pressures up to 20 atmospheres so that it inflates and opens the blockage. Upon deflation, the blood flow is reestablished and the catheter and guidewire removed prior to closure of the access incision in the artery. The guidewires are made from metals such as stainless steel and can be present within the bloodstream from minutes to hours. Even with heparinization, thrombus can form on the guidewire, which could be of potential danger to the patient if it were to dislodge and travel around the body. An *in vivo* study on guidewires coated with a simple PC copolymer (Fig. 2) showed that the coating significantly reduced platelet adhesion in a porcine model.^[57] Recently, Plante et al. reported on a clinical study using a variety of different guidewires with various coatings, and demonstrated by SEM that only the PC-coated guidewire had no sign of any thrombus present after clinical use, all other wires showing significant coverage with thrombi^[58] (Fig. 4c,C).

Clinical trials have shown that long-term outcomes after PTCA are significantly improved if a device known as a stent is placed within the reopened section of the vessel to maintain its patency. The stent can be of several types, but most popular are metal scaffolds that are crimped onto the balloon, and are then expanded into place when the balloon is inflated. Stents are also commonly made of stainless steel and, once deployed, remain in contact with the blood until such a time that the vessel heals and the stent becomes covered with a new endothelial layer. This process can take months and patients are required to remain on antiplatelet therapies over this period to prevent the risk of stent-related thrombosis blocking the treated vessel completely. The PC-coated stent was developed with the objective of reducing the risk of subacute thrombosis by masking the underlying thrombogenicity of the substrate. Initial studies of the coated stent were made *ex vivo* using an arteriovenous shunt made of silastic tubing in which the device was deployed; the shunt was then attached between a suitable artery and vein, some of the flowing blood being redirected through the shunt and over the stent. Platelets were radioactively-labeled with indium and their adhesion monitored using a gamma camera focused over the shunt. SEM was used to reveal the extent of occlusion of the shunt as thrombus formed (Fig. 4d,D). PC-coated stents displayed the expected resistance to platelet adhesion and thrombus formation,^[27] prompting more challenging studies *in vivo* in a range of more relevant preclinical models. Results from these assessments further support

the excellent hemocompatibility of the PC coating. Moreover, histological examination of the implants revealed that the tissue had no adverse reaction to the coating,^[59,60] unlike many other coatings previously studied that induced severe inflammatory reactions.^[61] Data such as these supported the clinical evaluation of the coated stent in a number of different registries and other trials, culminating in a randomized study used for U.S. approval of the device called DISTINCT (*BiodivYsio™ stent in controlled clinical trial*).^[62] This remains the only randomized trial with a coated stent to yield a 0% subacute thrombosis rate.

Although serious when it occurs, the incidence of stent-induced subacute thrombosis in the real world is reported to be around 3%, but lower in most clinical trials. This is due primarily to the introduction of more effective antiplatelet agents. A much more serious problem has been the occurrence of restenosis, in which the stented segment reblocks over time, necessitating a repeat procedure to unblock the artery again. The additional procedure not only incurs extra cost, but is further anguish to the patient and is made more difficult due to the presence of the initial implanted device. Restenosis is in essence an overactive healing response to the injury caused by balloon expansion and stent placement. The molecular biology of the process is complex, not well understood, and beyond the scope of this article. The rates of restenosis with PC-coated stents are comparable to market-leading uncoated devices, as the PC is bioinert and cannot address the injury caused during the procedure. Many approaches have been suggested to try and control the overproduction of new intimal tissue within the stented segment of the vessel (or neointimal hyperplasia, as it is known), but most recently, the delivery of drugs locally from the stent itself is proving to be the most efficacious. The PC coating has been shown to be a proven vehicle for the delivery of a range of therapeutic agents,^[25,26] and thus PC drug delivery stents were developed. The *BiodivYsio* Matrix stent systems were developed in order for physicians to load a drug of their choice onto the device by simple immersion of the stent in a suitable solution of the drug for a few minutes just prior to stent implantation.^[63] Upon removal from the solution, the stent is allowed to dry in air and then immediately deployed. Matrix LO is based on PC1036 and allows loading of drugs up to 1.2 kilodaltons. It has been used to load a range of drugs including angiotensin,^[64] the anti-inflammatory drug dexamethasone,^[65] and the healing promoter oestradiol,^[66] the latter two drugs showed promise in preclinical studies and have progressed to clinical trials (*STRIDE* (*study of anti-restenosis with the biodivYsio dexamethasone-eluting stent*) and *EASTER* (*estrogen and stents to eliminate restenosis*) for dexamethasone and oestradiol, respectively). Matrix HI is based on a cationically

modified PC coating (PC2028), and has been used in a number of preclinical studies including the study of Oliglow (radioactive oligonucleotide), DNA plasmid that codes for vascular endothelial growth factor-2 (VEGF2), and an antisense codon that codes for *c-myc*. There is currently no clinical activity based on this device, although there is a growing interest in the use of gene-based therapy to address restenosis.

Elsewhere in the Body

Outside of the bloodstream, the ocular environment is the second area of major application for PC Technology. The tear film is a complex structure composed of an aqueous mixture of proteins and lipids that bathe the eye, providing nutrients and protection against bacterial infection. Moreover, the cornea has no blood vessels and the tissue relies upon the direct absorption of oxygen from the atmosphere into the tear film where it can then diffuse into the cells. Any device that is required to function in contact with the cornea will significantly impact upon the tear film and will inevitably become a barrier toward the diffusion of oxygen. A PC-based soft contact lens material known as omafilcon A, has been developed; it falls within the soft contact classification scheme as a high-water-content nonionic material. The Proclear[®] range of soft contact lenses provides for the maintained health of the cornea on a number of different levels. Studies have shown that both protein and lipid deposition on the lens are markedly reduced when PC is present;^[38,39] this has the benefit of more comfort for the wearer throughout the day, as the deposition of biomolecules on the lens is essentially an immune response that can cause irritation to the cornea.

Once in contact with the cornea, the lens must allow for adequate permeability toward oxygen, ions, and other nutrients. The water within the hydrogel lens provides the medium through which these may be transported to the tissue. The amount of PC in the lens has been engineered to produce a bulk water content of around 59%, more than sufficient to provide more than adequate oxygen permeability, at the same time maintaining sufficient mechanical strength to allow handling. The PC also dominates the surface properties of the lens. The highly hydrophilic nature of the material ensures that the surface of the lens is easily wetted by the tear film, as demonstrated in Fig. 4e,E, in which the highlights on each lens clearly show the difference between the poor wetting characteristics of an ordinary lens (e) compared to that of a PC-based lens (E). The tear film is shown to break up more rapidly, and the lens surface wettability deteriorates, as proteins and lipids in the tears begin to adhere and spoil the lens surface. With the enhanced resistance to the spoliation processes, the PC-based lens material maintains its

excellent wetting properties—and hence a stable tear film—for much longer periods.

Furthermore, the PC helps to reduce evaporative loss of water from the lens. On-eye studies have shown that the PC-based lens dehydrates significantly less over a day's wear than other soft lenses made from conventional materials.^[67,68] In a double-blind clinical study, two lens-user groups compared the PC-based lens with an ordinary market-leading lens over a six-week period. One group wore Proclear for six weeks and then switched to the other lens, and vice versa for the other group. Each was asked to rank the lens on the basis of comfort at the end of each week, and in both groups the PC-based material was ranked significantly higher.^[69] On the basis of such data, the product has a label claim allowed by the FDA based upon potential benefits provided by the lens for dry-eye sufferers.

Elsewhere in the eye, PC Technology has been investigated on such devices as glaucoma shunts, punctum plugs, and intraocular lenses and for extended-wear contact lenses. Although there are no commercial products in these areas as yet, the benefits of PC are realized in feasibility. Its use in intraocular lenses (IOLs) in particular, as both a coating material for existing PMMA-based IOLs and as a wholly new foldable bulk IOL material, has shown great promise. Significant reductions in the adhesion of proteins such as fibrinogen and lysozyme, various cell lines such as corneal epithelia, fibroblasts, and macrophages, and bacteria such as *Staphylococcus epidermidis* and *Pseudomonas aeruginosa* have been reported (see Table 2 and references cited therein).

Infection is a common problem encountered with implants of all types, and can often result in additional trauma associated with the need to remove the device. The ability of PC Technology to resist colonization by a variety of strains of bacteria has therefore driven its application on devices used in other regions around the body. Studies in vitro have shown that the PC coating is more effective than antimicrobial silver-based treatments in reducing both *S. aureus* and *P. aeruginosa* biofilm formation on fluoroplastic tympanostomy tubes^[70] (licensed by Gyrus for the PacifiC[®] tympanostomy tubes). Other in vitro studies have shown reductions in adhesion of *E. coli*, *S. epidermidis*, *Proteus mirabilis* and *P. aeruginosa* on various coated alloplasts compared to uncoated controls (see Fig. 4f,F^[71]). Furthermore, the ability of the coating to resist the adhesion of proteins, in particular fibrinogen, is significant because of the implication of this protein in the formation of conditioning films necessary to encourage *S. aureus* attachment to surfaces. Urological devices are especially renowned for complications induced by infection due to the relative ease by which bacteria can gain access to the device surfaces. Certain of these bacterial species are known to increase the possibility of mineralization or



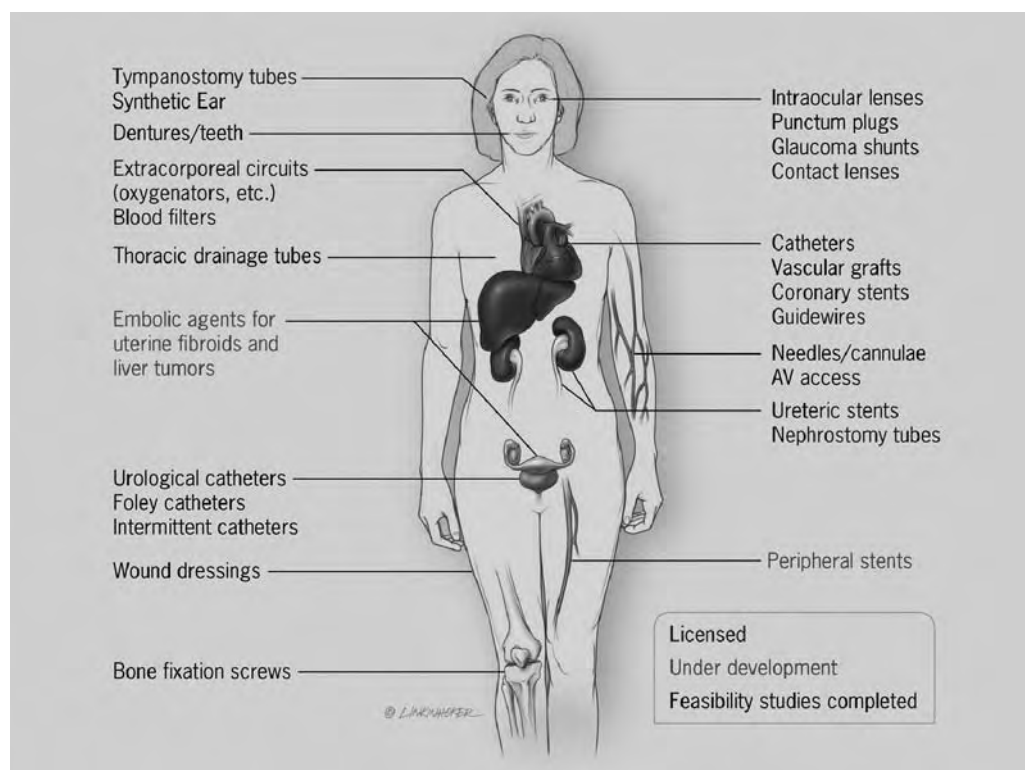


Fig. 5 The use of PC coatings and materials around the body from feasibility to commercial reality.

encrustation of the devices. Urease-secreting organisms such as *P. mirabilis* convert urea to ammonia, causing an increase in pH and initiation of crystallization of calcium and magnesium salts. These minerals encrust the lumens of the device causing blockage. Laboratory methods that mimic this encrustation process using *P. mirabilis* in artificial urine have shown that PC coatings produce large reductions in magnesium on latex and silicone substrates.^[27,71] The same effect has also been observed in vivo in a study on urological catheters, in which uncoated and PC-coated catheters were used consecutively in the same patients and the extent of encrustation assessed in each case (Fig. 4g, G^[27]). Moreover, a clinical study was performed using PC-coated polyurethane ureteral stents which were implanted for three months and then removed and analyzed for the presence of biofilm. This was compared to standard stents removed previously from the same patients and analyzed using the same techniques; the percentage of stents with biofilm visible by scanning electron microscopy was shown to be reduced from 70 to 30% for the PC-coated devices.^[71]

CONCLUSIONS AND FUTURE DIRECTIONS

PC Technology is an important addition to the field of biomaterials that has translated into significant clinical

benefit and commercial investment in a number of medical device applications. It has been studied in a vast array of indications around the body, as illustrated in Fig. 5. The current focus for the technology is on the development of improved embolic agents, materials that are introduced into blood vessels with the objective of occluding the blood flow (for instance, in order to stem flow to a tumor or block a vascular malformation). This type of minimally invasive procedure is on the increase and requires more advanced materials with improved long-term biocompatibility and the potential for drug delivery.

Medical devices have been transformed over recent years with the emphasis shifting from a purely mechanical standpoint to include consideration of the material-tissue interface. More recently, there has been a desire to introduce some form of bioactivity that will enhance the ability of the tissue to heal in the presence of the device. PC Technology is playing an important role in this biomaterials revolution, proving to be an extremely flexible platform for combining exceptional biocompatibility with the ability to deliver therapeutics in a controlled fashion. The continued evolution of the technology will be somewhat dependent upon the development of new techniques for polymer synthesis in order to allow novel architectures to be created, and subsequent device construction to convert these into useful biomaterials. Some initial advances have been

made using novel polymerization methods, such as ATRP, that have allowed well-defined, low polydispersity block copolymers to be made, which have the potential to form biocompatible stimuli-responsive structures that could find use in a multitude of drug delivery applications in the future.

ARTICLES OF FURTHER INTEREST

Biocompatibility Testing; Blood-Material Interactions; Drug Delivery, Controlled; Host Reactions; Protein and Cell Signaling with Biomaterials: Influence of Surfactants; Surface Coatings; Surface Modification; Thrombosis

REFERENCES

- Zwaal, R.F.A.; Comfurius, P.; van Deenen, L.L.M. Membrane asymmetry and blood coagulation. *Nature* **1977**, *268*, 358–360.
- Hayward, J.A.; Chapman, D. Biomembrane surfaces as models for polymer design: The potential for haemocompatibility. *Biomaterials* **1984**, *5*, 135–142.
- Durrani, A.A.; Hayward, J.A.; Chapman, D. Biomembranes as models for polymer surfaces II. The syntheses of reactive species for covalent coupling of phosphorylcholine to polymer surfaces. *Biomaterials* **1986**, *7*, 121–125.
- Hayward, J.A.; Durrani, A.A.; Shelton, C.J.; Lee, D.C.; Chapman, D. Biomembranes as models for polymer surfaces III. Characterisation of a phosphorylcholine surface covalently bound to glass. *Biomaterials* **1986**, *7*, 126–131.
- Hayward, J.A.; Durrani, A.A.; Lu, Y.; Clayton, C.R.; Chapman, D. Biomembranes as models for polymer surfaces IV. ESCA analyses of a phosphorylcholine surface covalently bound to hydroxylated substrates. *Biomaterials* **1986**, *7*, 252–258.
- Hall, B.; Bird, R.; le, R.; Kojima, M.; Chapman, D. Biomembranes as models for polymer surfaces V. Thromboelastographic studies of polymeric lipids and polyesters. *Biomaterials* **1989**, *10*, 219–224.
- Chapman, D.; Charles, S.A. A coat of many lipids—In the clinic. *Chem. Br.* **1992**, *28*, 253–256.
- Lewis, A.L.; Hughes, P.D.; Kirkwood, L.C.; Leppard, S.W.; Tolhurst, L.A.; Stratford, P.W. Synthesis and characterisation of phosphorylcholine-based polymers useful for coating blood filtration devices. *Biomaterials* **2000**, *21* (18), 1847–1859.
- Lewis, A.L.; Cumming, Z.C.; Goreish, H.H.; Kirkwood, L.C.; Tolhurst, L.A.; Stratford, P.W. Crosslinkable coatings from phosphorylcholine-based polymers. *Biomaterials* **2001**, *22*, 99–111.
- Lobb, E.J.; Ma, I.; Billingham, N.C.; Armes, S.P.; Lewis, A.L. Facile synthesis of well-defined, biocompatible phosphorylcholine-based methacrylate copolymers via atom transfer radical polymerisation at 20°C. *J. Am. Chem. Soc.* **2001**, *123*, 7913–7914.
- Ma, I.; Lobb, E.J.; Billingham, N.C.; Armes, S.P.; Lewis, A.L.; Lloyd, A.W.; Salvage, J. Synthesis of biocompatible polymers I. Homopolymerisation of 2-methacryloyloxyethyl phosphorylcholine via ATRP in protic solvents: An optimisation study. *Macromolecules* **2002**, *35* (25), 9306–9314.
- Ma, Y.; Tang, Y.; Billingham, N.C.; Armes, S.P.; Lewis, A.L.; Lloyd, A.W.; Salvage, J.P. Well-defined biocompatible block copolymers via atom transfer radical polymerisation of 2-methacryloyloxyethyl phosphorylcholine in protic media. *Macromolecules* **2003**, *36*, 3475–3484.
- Bowers, R.W.J.; Stratford, P.W.; Jones, S.A. Contact Lens Material. Biocompatibles Ltd. US 6,420,453, 1991.
- Bowers, R.W.J.; Stratford, P.W.; Jones, S.A. Contact Lens Material. Biocompatibles Ltd. US 6,423,761, 1995.
- Campbell, E.J.; O’Byrne, V.; Stratford, P.W.; Quirk, I.; Vick, T.A.; Wiles, M.C.; Yianni, Y.P. Biocompatible surfaces using methacryloylphosphorylcholine laurylmethacrylate copolymers. *ASAIO J.* **1994**, *40* (3), 853–857.
- Lewis, A.L.; Vick, T.A. Site-specific drug delivery from coronary stents. *Drug Deliv. Syst. Sci.* **2001**, *1* (3), 65–71.
- Allsopp, G. Performance review of a biocompatible monthly disposable lens. *Optician* **1997**, *213*, No. 5595.
- Willis, S.L.; Court, J.L.; Redman, R.P.; Wang, J.-H.; Leppard, S.W.; O’Byrne, V.J.; Small, S.A.; Lewis, A.L.; Jones, S.A.; Stratford, P.W. A novel phosphorylcholine-coated contact lens for extended wear use. *Biomaterials* **2001**, *22* (24), 3261–3272.
- Clarke, S.; Davies, M.C.; Roberts, C.J.; Tendler, S.J.B.; Williams, P.M.; O’Byrne, V.; Lewis, A.L.; Russell, J.C. Surface mobility of 2-methacryloyloxyethyl phosphorylcholine-co-lauryl methacrylate polymers. *Langmuir* **2000**, *16*, 5116–5122.
- Tang, Y.; Lu, J.R.; Lewis, A.L.; Vick, T.A.; Stratford, P.W. Swelling of zwitterionic polymer films characterised by spectroscopic ellipsometry. *Macromolecules* **2001**, *34* (25), 8768–8776.
- Tang, Y.; Lu, J.R.; Lewis, A.L.; Vick, T.A.; Stratford, P.W. The structural effects on swelling of thin phosphorylcholine polymer films. *Macromolecules* **2002**, *35* (10), 3955–3964.
- Russell, J.C.; Jones, J.R.; Vick, T.A.; Stratford, P.W. The preparation of tritium labelled biocompatible polymers. *J. Labelled Compd Radiopharm.* **1993**, *33*, 957–964.
- Lewis, A.L.; Tolhurst, L.A.; Stratford, P.W. Analysis of a phosphorylcholine-based polymer coating on a coronary stent pre- and post-implantation. *Biomaterials* **2002**, *23* (7), 1697–1706.
- Lewis, A.L.; Furze, J.D.; Small, S.; Robertson, J.D.; Higgins, B.J.; Taylor, A.S.; Ricci, D.R. Long-term stability of a coronary stent coating post-implantation. *J. Biomed. Mater. Res., Appl. Biomater.* **2002**, *63* (6), 699–705.
- McNair, A.M. Using hydrogel polymers for drug delivery. *Med. Device Technol.* **1996**, *7* (10), 16–20.
- Lewis, A.L.; Vick, T.A.; Collias, A.M.C.; Hughes, L.G.; Palmer, R.R.; Leppard, S.W.; Furze, J.D.; Taylor, A.S.; Stratford, P.W. Phosphorylcholine-based polymers for



- stent drug delivery. *J. Mater. Sci., Mater. Med.* **2001**, *12* (10–12), 865–870.
27. Lewis, A.L. Phosphorylcholine-based polymers and their use in the prevention of biofouling. *Colloids Surf., B Biointerfaces* **2000**, *18*, 261–275, and references therein.
 28. Ishihara, K.; Nomura, H.; Mihara, T.; Kurita, K.; Iwasaki, Y.; Nakabayashi, N. Why do phospholipid polymers reduce protein adsorption? *J. Biomed. Mater. Res.* **1998**, *39*, 323–330.
 29. Campbell, E.J.; Hall, B.; Parmer, A.R.; Sullivan, A.M. Improved haemocompatibility of artificial surfaces by treatment using phosphorylcholine derivatives. *Thromb. Haemost.* **1993**, *69*, 935.
 30. Murphy, E.F.; Lu, J.R.; Brewer, J.; Russell, J.C.; Penfold, J. The reduced adsorption of proteins at the phosphoryl choline incorporated polymer-water interface. *Langmuir* **1999**, *15*, 1313–1322.
 31. Murphy, E.F.; Keddie, J.L.; Lu, J.R.; Brewer, J.; Russell, J.C. The reduced adsorption of lysozyme at the phosphorylcholine incorporated polymer/aqueous solution interface studied by spectroscopic ellipsometry. *Biomaterials* **1999**, *20* (16), 1501–1511.
 32. Clarke, S.; Davies, M.C.; Lewis, A.L.; O'Byrne, V.; Roberts, C.J.; Tendler, S.J.B.; Williams, P.M. Novel biomimetic polymers incorporating phosphorylcholine. *J. Pharm. Pharmacol.* **2000**, *52* (Suppl.), 114.
 33. Clarke, S.; Davies, M.C.; Roberts, C.J.; Tendler, S.J.B.; Williams, P.M.; Lewis, A.L.; O'Byrne, V. Atomic force microscope and surface plasmon resonance investigation of polymer blends of poly([2-(methacryloyloxy)ethyl] phosphorylcholine-co-lauryl methacrylate) and poly (lauryl methacrylate). *Macromolecules* **2001**, *34*, 4166–4172.
 34. Murphy, E.F.; Lu, J.R.; Lewis, A.L.; Brewer, J.; Russell, J.C.; Stratford, P.W. Characterisation of protein adsorption at the phosphorylcholine incorporated polymer-water interface. *Macromolecules* **2000**, *33* (12), 4545–4554.
 35. Faragher, R.G.A.; Bowers, R.W.J.; Denyer, S.P.; Dropcova, S.; Gard, P.R.; Hall, B.; Hanlon, G.W.; Jones, S.A.; Lloyd, A.W.; Mikhailovsky, S.V.; Olliff, C.J.; Riding, M.; Rosen, P.H. In-vitro evaluation of biomimetic hydrogels: Potential new base materials for intraocular lenses. *J. Pharm. Pharmacol.* **1995**, *47* (Suppl.), 1068.
 36. Lloyd, A.W.; Dropcova, S.; Mikhailovsky, S.V.; Denyer, S.P.; Gard, P.R.; Hanlon, G.W.; Olliff, C.J.; Faragher, R.G.A.; Bowers, R.W.J.; Hall, B.; Muir, A.; Jones, S.A.; Rosen, P.; Riding, M. Biocompatible hydrogels for IOLs. *Vis. Res.* **1996**, *36*, S113.
 37. Andrews, C.S.; Lloyd, A.W.; Sharma, D.P.; Hall, B.; Hanlon, G.W.; Denyer, S.P. In-vitro evaluation of biomimetic hydrogels. *Vis. Res.* **1996**, *36*, S200.
 38. Young, G.; Bowers, R.W.J.; Hall, B.; Port, M. Clinical comparison of omafilcon A with four control materials. *CLAO J.* **1997**, *23* (4), 249.
 39. Bowers, R.W.J.; Port, M.; Young, G. *Optician* **1995**, *210*, 23.
 40. Lloyd, A.W.; Dropcova, S.; Mikhailovsky, S.V.; Denyer, S.P.; Gard, P.R.; Hanlon, G.W.; Olliff, C.J.; Faragher, R.G.A.; Bowers, R.W.J.; Hall, B.; Muir, A.; Jones, S.A.; Rosen, P.; Riding, M. A biocompatible coating for IOLs. *Vis. Res.* **1996**, *36*, S166.
 41. Faragher, R.G.A.; Bowers, R.W.J.; Denyer, S.P.; Dropcova, S.; Gard, P.R.; Hall, B.; Hanlon, G.W.; Jones, S.A.; Lloyd, A.W.; Mikhailovsky, S.V.; Muir, A.; Olliff, C.J.; Riding, M.; Risen, P.H. Novel base materials for intraocular lenses. *Eur. J. Pharm. Sci.* **1996**, *4*, S176.
 42. Hanlon, G.W.; Bowers, R.W.J.; Dropcova, S.; Denyer, S.P.; Faragher, R.G.A.; Gard, P.R.; Hall, B.; Jones, S.A.; Lloyd, A.W.; Muir, A.; Olliff, C.J.; Rosen, P.; Riding, M. In-vitro evaluation of a novel biocompatible intraocular lens coating. *Investig. Ophthalmol. Vis. Sci.* **1997**, *38*, S179.
 43. Lloyd, A.W.; Bowers, R.W.J.; Dropcova, S.; Denyer, S.P.; Faragher, R.G.A.; Gard, P.R.; Hall, B.; Hanlon, G.W.; Jones, S.A.; Muir, A.; Olliff, C.J.; Rosen, P.H.; Riding, H. In-vitro evaluation of novel biomimetic intraocular lenses. *Investig. Ophthalmol. Vis. Sci.* **1997**, *38*, S178.
 44. Long, S.F.; Lewis, A.L.; Hanlon, G.W.; Lloyd, A.W. Differential cell adhesion to phosphorylcholine polymers with varying cationic charge. *Eur. Cells Mater.* **2001**, *2* (Suppl. 1), 50–51.
 45. Filipec, M.; Bowers, R.W.J.; Dropcova, S.; Denyer, S.P.; Faragher, R.G.A.; Gard, P.R.; Hall, B.; Hanlon, G.W.; Jones, S.A.; Lloyd, A.W.; Muir, A.; Olliff, C.J.; Rosen, P.H.; Riding, M. In-vivo evaluation of novel biomimetic intraocular lenses. *Investig. Ophthalmol. Vis. Sci.* **1997**, *38*, S178.
 46. Lloyd, A.W.; Faragher, R.G.A.; Reed, S.; Wassall, M.; Wong, L. Proceedings of the 9th Symposium on the Material Science and Chemistry of Contact Lenses, New Orleans, **1997**.
 47. Dropcova, S.; Bowers, R.W.J.; Denyer, S.P.; Faragher, R.G.A.; Gard, P.R.; Hall, B.; Hanlon, G.W.; Jones, S.A.; Lloyd, A.W.; Mikhailovsky, S.V.; Olliff, C.J.; Riding, M.; Rosen, P.H. The effects of adsorbed phosphorylcholine based biomimetic copolymers on bacterial adhesion to PMMA intraocular lens material. *J. Pharm. Pharmacol.* **1995**, *47* (Suppl.), 1072.
 48. Andrew, C.S.; Denyer, S.P.; Hall, B.; Hanlon, G.W.; Lloyd, A.W. A comparison of the use of an ATP-based bioluminescence assay and image analysis for the assessment of bacterial adhesion to standard HEMA and biomimetic soft contact lenses. *Biomaterials* **2001**, *22*, 3225–3233.
 49. Chronos, N.A.F.; Robinson, K.A.; Kelly, A.B. Thromboresistant phosphorylcholine coating for coronary stents. *Circulation* **1995**, *92* (8), pl-685, Abst 3290.
 50. von Segesser, L.K.; Olah, A.; Leskosek, B.; Turina, M. Coagulation patterns in bovine left heart bypass with phospholipid versus heparin surface coating. *ASAIO Trans.* **1993**, *39* (1), 43–46.
 51. von Segesser, L.K.; Tonz, M.; Leskosek, B.; Turina, M.J. Evaluation of phospholipidic surface coatings ex-vivo. *Int. J. Artif. Organs* **1994**, *17* (5), 294–299.
 52. Goodall, A.H.; Campbell, E.J.; Wilson, D.J.; Taylor, A.S.; Chronos, N.A.F.; Nwose, O.D.; Yianni, Y.P. Improved haemocompatibility of artificial surfaces after coating with phosphorylcholine derivatives: An ex-vivo human model using native flowing blood. *Blood Coagul. Fibrinolysis* **1995**, *6*, 151. Abst.

53. Hunter, S.; Angelini, G.D. Phosphatidylcholine-coated chest tubes improve drainage after open heart operation. *Ann. Thorac. Surg.* **1993**, *56*, 1339–1342.
54. De Somer, F.; Francois, K. Phosphorylcholine coating of extracorporeal circuits provides natural protection against blood activation by the material surface. *Eur. J. Cardio-Thorac. Surg.* **2000**, *1*, 602–606.
55. De Somer, F.; van Belleghem, Y.; Caes, F.; Francois, K.; Arnout, J.; Bossuyt, X.; Taeymans, Y.; van Nooten, G. Phosphorylcholine coating offers natural platelet preservation during cardiopulmonary bypass. *Perfusion* **2002**, *1*, 39–44.
56. Chen, C.; Lumsden, A.B.; Ofenloch, J.; Beverly, N.; Campbell, E.J.; Stratford, P.W.; Yianni, Y.P.; Taylor, A.S.; Hanson, S.R. Phosphorylcholine coating of ePTFE grafts reduced neointimal hyperplasia in canine model. *Ann. Vasc. Surg.* **1997**, *11* (1), 74–79.
57. Veil, K.R.; Chronos, N.A.F.; Palmer, S.J.; Yianni, Y.P.; Campbell, E.J.; Cippola, G.D.; Robinson, K.A. Phosphorylcholine: A biocompatible coating for coronary angioplasty devices. *Circulation* **1995**, *92* (8), pl-685, Abst 2337.
58. Gobeil, F.; Juneau, C.; Plante, S. Thrombus formation on guide wires during routine PTCA procedures. A scanning electron microscopic evaluation. *Can. J. Cardiol.* **2002**, *18*, 263–269.
59. Whelan, D.M.; van der Giessen, W.J.; Krabbendam, S.C.; van Vliet, E.A.; Verdouw, O.D.; Serruys, P.W.; van Beusekom, H.M.M. Biocompatibility of phosphorylcholine coated stents in normal porcine coronary arteries. *Heart* **2000**, *83* (3), 338–345.
60. Malik, N.; Gunn, J.; Shepherd, L.; Crossman, D.C.; Cumberland, D.C.; Holt, C.M. Phosphorylcholine-coated stents in porcine coronary arteries: In-vivo assessment of biocompatibility. *J. Invasive Cardiol.* **2001**, *13*, 193–201.
61. van der Giessen, W.J.; Lincoff, A.M.; Schwartz, R.S.; van Beusekom, M.M.; Serruys, P.W.; Holmes, D.R.; Ellis, S.G.; Topol, E.J. Marked inflammatory sequelae to implantation of biodegradable and nonbiodegradable polymers in porcine coronary arteries. *Circulation* **1996**, *94* (7), 1690–1697.
62. Moses, J.W.; Aluka, A.O.; Barbeau, G.R.; et al. The first clinical trial comparing a coated *versus* a non-coated coronary stent: The biocompatibles *BiodivYsio*TM stent in randomised control trial (distinct). *Circulation* **2000**, *102* (Suppl. 18).
63. Rothman, M.L.; Lewis, A.L. Chapter 36, Biocompatibles Drug Eluting Stents. In *Handbook of Coronary Stents*, 4th Ed.; Serruys, P.W., Rensing, B.J., Eds.; Martin Dunitz: London, 2001; 299–310.
64. Armstrong, J.; Gunn, J.; Arnold, N.; Malik, N.; Chan, K.H.; Vick, T.A.; Stratford, P.W.; Cumberland, D.C.; Holt, C.M. Angiopeptin-eluting stents: Observations in human vessels and pig coronary arteries. *J. Invasive Cardiol.* **2002**, *14*, 230–238.
65. DeScheerder, I. Chapter 2, Anti-Inflammatory Approach To Restenosis. In *The Restenosis Matrix. Restenosis: Multiple Strategies for Stent Drug Delivery*; Rothman, M.T., Ed.; Remedica: London, 2001; Vol. 1, 12–31.
66. New, G.; Moses, J.W.; Roubin, G.S.; Leon, M.B.; Colombo, A.; Iyer, S.S.; Tio, F.O.; Mehran, R.; Kipshidze, N. Estrogen-eluting, phosphorylcholine-coated stent implantation is associated with reduced neointimal formation but no delay in vascular repair in a porcine coronary model. *Catheter. Cardiovasc. Interv.* **2002**, *57* (2), 266–271.
67. Young, G.; Bowers, R.J.W.; Hall, B.; Port, M. Six month clinical evaluation of a biomimetic hydrogel contact lens. *CLAO J.* **1997**, *23* (4), 226–236.
68. Lebow, K.; Bridgewater, B. A three month comparative daily-wear study of two high-water content soft lenses. *ICLC* **1997**, *24*, 198–205.
69. Lemp, M.A.; Caffery, B.; Lebow, K.; Lembach, R.; Park, J.; Foulks, G.; Hall, B.; Bowers, R.J.W.; McGarvey, S.; Young, G. Omaficon A (Proclear) soft contact lenses in dry eye population. *CLAO J.* **1999**, *25* (1), 40–47.
70. Berry, J.A.; Biedlingmaier, J.F.; Whelan, P.J. In-vitro resistance to bacterial biofilm formation on coated fluoroplastic tympanostomy tubes. *Otolaryngol. Head Neck Surg.* **2000**, *123*, 246–251.
71. Russell, J.C. Bacteria, biofilms and devices: The possible protective role of phosphorylcholine materials. *J. Endourol.* **2000**, *1*, 39–42.

FURTHER READING

This entry provides an overview of the characterization, properties and application of a commercial technology based on a family of proprietary PC-containing polymers. The author recognizes that there are numerous publications in the scientific literature on other PC-based polymers and there was no intent to provide a review of this whole area. Interested readers may wish to consider the following references as further reading on similar systems:

- Ishihara, K.; Ueda, T.; Nakabayashi, N. Preparation of phospholipid polymers and their properties as polymer hydrogel membranes. *Polym. J.* **1990**, *22*, 355–360.
- Ishihara, K.; Aragaki, R.; Ueda, T.; Watanabe, A.; Nakabayashi, N. Reduced thrombogenicity of polymers having phospholipid polar group. *J. Biomed. Mater. Res.* **1990**, *24*, 1069–1077.
- Ishihara, K.; Oshida, H.; Nakabayashi, N. Hemocompatibility of human whole blood on polymers with a phospholipid polar group and its mechanism. *J. Biomed. Mater. Res.* **1992**, *26*, 1543–1552.
- Sugiyama, K.; Shiraiishi, K.; Aoki, H. Surface modified polymer microspheres obtained by emulsion copolymerisation of 2-methacryloyloxyethyl phosphorylcholine with various monomers. *Polym. J.* **1994**, *26*, 561–569.
- Ueda, T.; Oshida, H.; Kurita, K.; Ishihara, K.; Nakabayashi, N. Preparation of 2-methacryloyloxyethyl phosphorylcholine copolymers with alkyl methacrylates and their blood compatibility. *Polym. J.* **1992**, *24* (11), 1259–1269.



Photopolymerization

Dong-an Wang

Jennifer H. Elisseff

Department of Biomedical Engineering, Johns Hopkins University, Baltimore, Maryland, U.S.A.

INTRODUCTION

Photopolymerization is defined as the reaction of monomers or macromers to produce polymeric structures by light-induced initiation (excitation of a photoinitiator) and subsequent polymerization. Fundamental research and practical applications of photopolymerization have evolved significantly over the past several decades leading to a better understanding of the mechanisms of the reaction and to numerous industrial and medical applications. The general concept of photopolymerization is to conjugate and solidify the fluid reactants into a desired stable structure using light. The procedure is usually irreversible. Complex patterns can be created using photopolymerization to print integrated circuits, wiring boards, or microfabricated biomaterial scaffolds. Photopolymerization technologies have also been applied in medicine for applications ranging from dental to lung sealants.

For the preparation of polymeric biomaterials, it is necessary to distinguish the concepts of photopolymerization, photocrosslinking and photografting. Photopolymerization refers to the polymerization of small molecular monomers or oligomers polymerizing via photoreactive moieties to generate a linear macromolecular. Poly(N-hydroxyethyl methacrylate) (pHEMA) is an example of a linear polymer created by photopolymerization. Photocrosslinking defines a process by which functional groups on linear or branched macromer chains crosslink to form a three-dimensional network. Poly(ethylene glycol diacrylate) (PEGDA) is a popular macromer used to generate hydrogels by photocrosslinking. In general terms, photopolymerization describes both linear photopolymerization and photocrosslinking. The third concept, photografting, describes the procedure of covalently incorporating functional additives to a polymeric matrix or surface using a light-induced mechanism. Photografting is an important modality for biomaterial surface modification. This article will review the mechanism of photoinitiation and photopolymerization, as well as novel applications in biomaterials and tissue engineering.

Photopolymerization causes significant changes to material physical properties, such as phase, solubility,

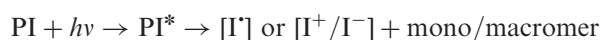
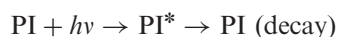
adhesion, color, optical index, and electrical conductivity. This technique for material synthesis has been applied to microfabrication, coating, and surface modification of devices and materials, including adhesives and sealants; printing materials; optical and electro-optical materials; and electronic materials.^[1] More recently, the benefits of photopolymerization have been utilized for numerous biomedical applications. The majority of photopolymerized biomaterials are based on acrylate systems. The prepolymers are usually water-soluble unsaturated acrylate or methacrylate-containing oligomers or macromers that can be covalently cross-linked via addition polymerization with the excitation of radical-generating photoinitiators. For dental or orthopedic applications, photopolymerization has several advantages including rapid gelation and the ability to create a moldable crosslinked structure in situ via minimally invasive injection. Furthermore, photopolymerization, enables encapsulation of cells for in-situ tissue engineering.

MECHANISMS OF PHOTOPOLYMERIZATION

Photopolymerization requires three essential elements: electromagnetic radiation ($h\nu$), chromophore-containing photoinitiators (PI), and mono/macromers with unsaturated moieties such as a vinyl group and its derivatives including styrene, acrylate, or methacrylate groups. Ultraviolet (UV: 190~400 nm) and visible light (Vis: 400~800 nm) can be harnessed for photopolymerization. Accordingly, the maximal absorption wavelength (λ_{\max}) of the PI chromophores (the carbon-carbon or carbon-oxygen π -electron conjugates in PI molecular structure) is concentrated in the relatively high energetic UV or near UV visible range.

Successful photopolymerization relies on effective initiation and propagation. The irradiated photons with proper frequency ($\nu = c/\lambda = E/h$, where c = light traveling speed, λ = wavelength, E = photon energy, and h = Planck's constant) are absorbed by the π -electron conjugates of PI chromophores that are simultaneously promoted to a higher energy orbital—PI*. This excited quantum state has a half-life of less than a μsec . It

rapidly decays back to the original state inducing photon emission or molecular lattice vibration. Numerous PI^* may also chemically decompose into reactive fragments, such as radicals $[I^*]$ or ions $[I^+/I^-]$. As reactive intermediates, these fragments are more stable and serve as direct initiators for radical polymerization or anionic/cationic polymerization (Eq. 1).



Eq. 1 Excitation of photoinitiators and process of photopolymerization.

Most of the practical biomedical applications of photopolymerization use a free radical mechanism to initiate addition polymerization of vinyl group-containing mono/macromers (Eqs. 2 and 3).

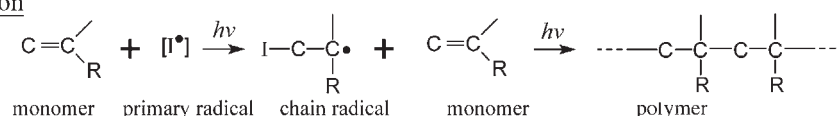
For traditional thermal radical polymerization, chain initiation and propagation usually require a temperature higher than 40°C . However, since electromagnetic radiation provides highly efficient energy for PI excitation, the photoinitiation and the subsequent polymer chain propagation can proceed at lower temperatures ($10\sim 40^\circ\text{C}$), which is compatible with physiological environments (37°C). Furthermore, chain transfer processes that terminate reactions in thermal radical polymerization are minimized. The termination of chain propagation usually proceeds by

standard radical coupling and unsaturation exchange through β -hydrogen migration involving ambient oxygen.^[2]

As described above, the primary radicals for polymerization initiation usually originate from the PI (PI^*). However, for some applications the exogenous PI is unnecessary. A typical example is the photografting strategy for the surface modification of polyurethane biomaterials. This mechanism is based on the in situ generation of primary radicals from the polymer matrix by photooxidation with a peroxide reagent and UV irradiation. The polyether moieties in polyurethane matrix can be oxidized into peroxide radicals that are capable of initiating subsequent chain propagation (Eq. 4).^[3] Another novel example of photo-oxidation is the generation of tyrosyl radicals on the tyrosine residues of collagen fibers in the superficial layer of cartilage tissue. By initiating these domestic radicals, the (meth)acrylated macromers are covalently grafted onto the tissue (Eqs. 4 and 5).^[4,5]

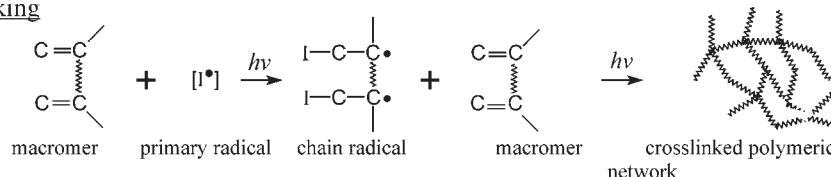
To study non-free radical systems, cationic photopolymerization of phenyl glycidyl ether was recently performed using two kinds of iodonium PI: diaryliodonium hexafluoroantimonate (IHA) and (tolcumyl) iodonium tetrakis (pentafluorophenyl) borate (IPB). Whereas the free radical system depends on light for both initiation and propagation, cationic photopolymerization requires light only for the initiation and creation of the cationic active center.^[6] This study demonstrates feasibilities for the dark-curing photoinitiated polymerization. However, the PI toxicity

Photopolymerization

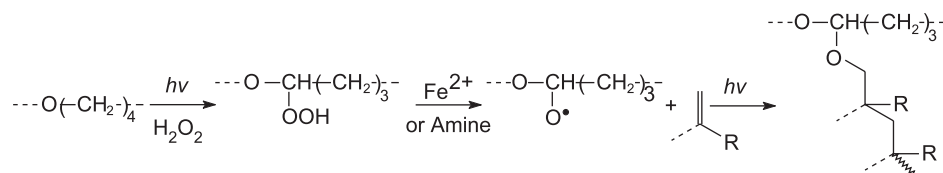


Eq. 2 Photopolymerization of vinyl monomers.

Photocrosslinking



Eq. 3 Photocrosslinking of multifunctional vinyl group-containing macromers.



Eq. 4 Photo-oxidation of polyether and photografting-polymerization of vinyl mono/macromers by initiation of peroxide radicals. (From Ref. [3].)

may limit biomedical applications of this novel approach.

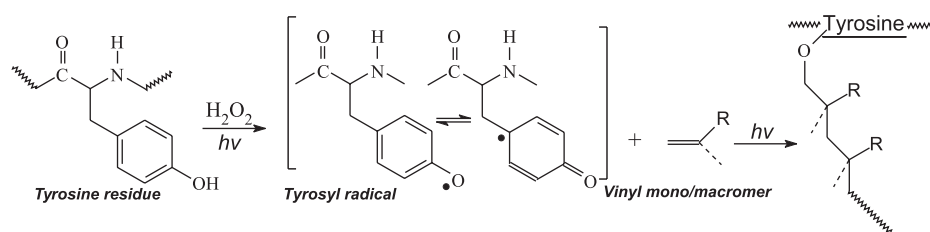
BIOCOMPATIBLE PHOTOINITIATORS

The role of photoinitiators (PI) in photopolymerization is to generate reactive intermediates (free radicals) to initiate the polymerization of unsaturated (vinyl) mono/macromers with the energy from electromagnetic radiation (i.e., UV or near UV-range visible light). Photo-initiation involves photon absorption, PI molecular excitation, and finally reactive fragment formation after PI decomposition. The function of photon absorption is executed by the π -electron-delocalized chromophores in the PI molecular structure and the successive light-induced PI cleavage also occurs at the boundary of the π -electron-conjugate moieties. Each chromophore has a specific maximal absorption wavelength (λ_{\max}) and a corresponding molar extinction coefficient (ϵ , $\text{mol}^{-1} \text{cm}^{-1}$), which is an experimental measure of photon absorption probability at that wavelength. The intensity of absorbance (A) can be calculated using the equation $A = \epsilon \cdot C \cdot l$, where C represents the concentration of chromophores and l represents the length of light-path. The typical units of chromophores are: $C=C$ (alkene, $\lambda_{\max} = 195 \text{ nm}$, $\epsilon = 10,000$), $C \equiv C$ (alkyne, $\lambda_{\max} = 195 \text{ nm}$, $\epsilon = 2000$), $C=C-C=C$ (diene, $\lambda_{\max} = 215 \text{ nm}$, $\epsilon = 20,000$), Benzene (benzoyl or phenyl, $\lambda_{\max} = 185 \text{ nm}$, $\epsilon = 60,000$; $\lambda_{\max} = 200 \text{ nm}$, $\epsilon = 8000$; $\lambda_{\max} = 255 \text{ nm}$, $\epsilon = 200$), RCOR (Ketone, $\lambda_{\max} = 190 \text{ nm}$, $\epsilon = 900$; $\lambda_{\max} = 280 \text{ nm}$, $\epsilon = 15$), RCOOH (Carboxyl, $\lambda_{\max} = 205 \text{ nm}$, $\epsilon = 60$), $\text{N}=\text{N}$ (Azo, $\lambda_{\max} = 345 \text{ nm}$, $\epsilon = 10$), NO_2 (Nitryl, $\lambda_{\max} = 270 \text{ nm}$, $\epsilon = 18.6$), $\text{S}=\text{O}$ ($\lambda_{\max} = 210 \text{ nm}$,

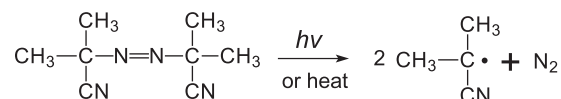
$\epsilon = 1500$), ROOR (Peroxide, $\lambda_{\max} = 200 \sim 300 \text{ nm}$).^[1,2] However, the practical value of ϵ relies on the individual chemical environments, such as the properties of adjacent groups, substituents, and local conjugates.

Commonly used thermal radical initiators, azo and peroxide initiators, are also capable of photoinitiation. The typical species are respectively 2,2'-azobisisobutyronitrile (AIBN, $\lambda_{\max} = 350 \text{ nm}$) and benzoyl peroxide (BPO, $\lambda_{\max} = 230$ and 270 nm). Light-induced radical formation is illustrated schematically in Eqs. 6 and 7. Unfortunately, these compounds suffer from low photoinitiation efficiency and cell toxicity, and are not water soluble, thus limiting biological applications. Limited photoinitiator water-solubility may lead to low initiation efficiency and toxicity. Solubility in water may be improved by dissolving the initiator in an initiator stock solution of 70% ethanol.

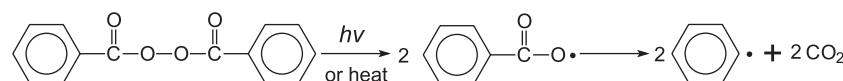
Toxicity of the photopolymerization system may be caused by the electromagnetic irradiation, chemical properties of the reagents, or initiator reactive intermediates. The frequency and intensity of light can be manipulated to optimize photopolymerization efficiency and reduce toxicity. A higher intensity and/or shorter wavelength of light enhances initiation efficiency at the cost of increased cell and tissue injury. On the other hand, a lower irradiation intensity will require longer exposure times for polymerization, which can also be harmful to the viability of living cells or tissues. One potential solution is to apply a biocompatible range of ultraviolet ($\lambda \geq 365 \text{ nm}$)^[7] or visible blue light ($\lambda = 420 \sim 500 \text{ nm}$). A class of photoinitiators containing benzoyl or phenyl-ketone π -electron-delocalized conjugates respond to these wavelengths of light. The typical products of UV-range PI (Ciba-Geigy) are Irgacure



Eq. 5 Generation of tyrosyl radicals by photo-oxidation^[4] and photografting-polymerization of vinyl mono/macromers by initiation of tyrosyl radicals. (From Ref. [5].)



Eq. 6 Reactive radical generation by AIBN decomposition.

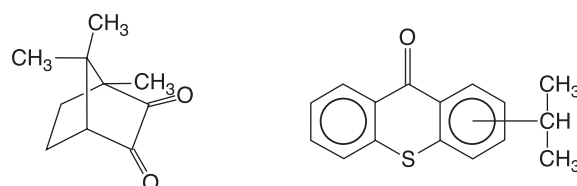


Eq. 7 Reactive radical generation by BPO decomposition.

family of 2,2-dimethoxy-2-phenylacetophenone (I 651, $\lambda_{\max}=365$ nm, $\epsilon=94.6$), 1-hydroxycyclohexyl phenyl ketone (I 184, $\lambda_{\max}=365$ nm, $\epsilon=7.7$), 2-methyl-1-[4-(methylthio) phenyl]-2-(4-morpholinyl)-1-propanone (I 907, $\lambda_{\max}=365$ nm, $\epsilon=120.8$), and Darocur family of 2-hydroxy-1-[4-(hydroxyethoxy) phenyl]-2-methyl-1-propanone (D 2959, $\lambda_{\max}=365$ nm, $\epsilon=6.7$). The prototypical visible-light initiating system is camphorquinone (CQ, $\lambda_{\max}=470$ nm, $\epsilon=24.8$, Aldrich) (Eq. 8a) with ethyl 4-*N,N*-dimethylaminobenzoate (4EDMAB), triethyl amine (TEA), and a popular photosensitizer, isopropyl thioxanthone (Quantacure ITX, $\lambda_{\max}=365$ nm, Great Lakes Fine Chemicals) (Eq. 8b). The mechanism of radical generation is demonstrated in Eqs. [9–12].^[7]

The potential toxicity mediated by the combined factors of chemicals and irradiation has been investigated using various cells. These experiments were performed by suspending the cells in PI solutions and

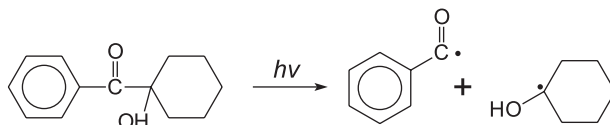
directly exposing them to light irradiation. As reported by Bryant and Anseth,^[7] with ~ 6 mW/cm² (intensity) of 365-nm UV or ~ 60 mW/cm² of 470–490-nm visible light, respectively, the UV-PI D 2959 at concentration of $\leq 0.05\%$ (w/w) or the visible-PI CQ at concentration of $\leq 0.01\%$ (w/w) is cytocompatible for NIH/3T3 fibroblast cells and also capable of initiating the photopolymerization of macromers. For chondrocytes, the use of 0.05% (w/w) D 2959 is safe with ~ 8 mW/cm² of 365 nm UV exposure for 10 minutes. According to the data reported by Williams, Malik, and Elisseff,^[8] the use of 0.05% (w/w) D 2959 with ~ 8 mW/cm², 365 nm-UV exposure for 5 minutes is also safe for goat and human mesenchymal stem cells (gMSC and hMSC). For corneal epithelial cells (SIRC cell-line), the cell viability is around 50%; for human embryonic germ cells (LVEC cell-line) and human fetal osteoblast (hFOB), the cell viability is lower than 50%.



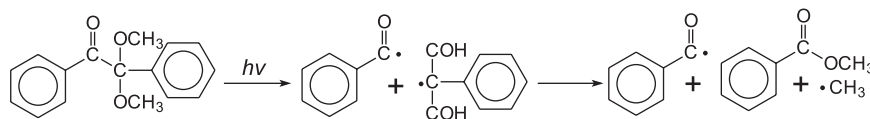
(a) Camphorquinone (CQ)

(b) Isopropyl Thioxanthone (Quantacure ITX)

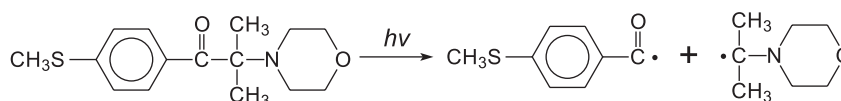
Eq. 8 Chemical structure of (a) CQ and (b) photosensitizer ITX.



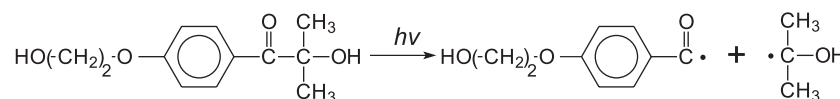
Eq. 9 Reactive radical generation by I 184 photo-decomposition.



Eq. 10 Reactive radical generation by I 651 photo-decomposition.



Eq. 11 Reactive radical generation by I 907 photo-decomposition.



Eq. 12 Reactive radical generation by D 2959 photo-decomposition.

PHOTOPOLYMERIZED BIOMATERIALS

The concept of biomaterials covers all natural and synthetic nonmedicinal matters or blends that, in whole or partially, fulfill the purposes of physically reconstituting, improving, substituting, or reconstructing defective human tissues or organs. Biomaterials include metal, ceramic, synthetic organic materials, and natural macromolecular materials, of which functional polymers are one of the largest families. For biomaterial fabrication, the process of photopolymerization has advantages of short reaction time, mild reaction conditions, in situ gelation, and biocompatibility. Hydrogels are one class of photopolymerized biomaterials. A photopolymerization system comprises mono/macromers, photoinitiators, processing additives, and functional additives in aqueous reaction medium. For tissue engineering applications, living cells are also incorporated. Mono/macromers are usually hydrophilic, with the essential unsaturated structures needed for polymerization. Degradable moieties and functional groups are frequently incorporated into the hydrogels. Members of the polyester family; poly(orthoester), poly(propylene fumarate) (PPF), polyphosphoesters, poly(lactic acid) (PLA), poly(glycolic acid) (PGA), and polyanhydride, as well as natural macromolecules, are commonly employed as the degradation sites on the polymeric scaffold of hydrogels. Biological signals, such as proteins, peptides, sugars, enzymes, and growth factors, are often incorporated as either functional pendent groups on polymeric skeleton or functional additives physically mixed in the system. Hyaluronic acid (HA) is an example of processing additive for controlling the systemic viscosity and fluidity.

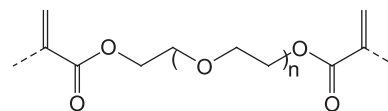
Photopolymerized Hydrogels

Photopolymerized hydrogels created from small molecular weight monomers, such as poly(2-hydroxyethyl methacrylate) (pHEMA), have unique properties. These hydrogels are composed of small-molecular-weight repeating units, and correspondingly have a small pore size and low swelling ratio (water content). High-density pHEMA hydrogels have a high degree of polymerization, which results in a higher bulk strength and adhesive capacity. As biomaterials, they are primarily used as injectable hard-tissue substitutes in the spine or as permanent dental fillers. Due to the physically condensed composition of these hydrogels, living cells may not be encapsulated.

Photocrosslinked Hydrogels

Photocrosslinked hydrogels for use as tissue engineering scaffolds are formed by larger molecular weight

monomers, macromers. The macromers are usually water-soluble polymers with photocrosslinkable moieties. Poly(ethylene glycol)-based polymers, including PEG-diacrylate (PEGDA, Eq. 13), are frequently used for tissue-engineering applications and provide a simple model for illustrating their practicality.



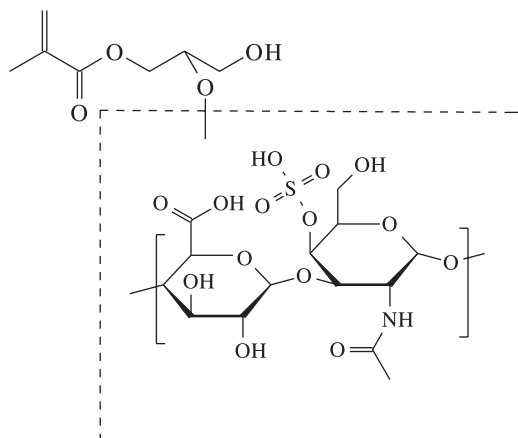
Eq. 13 Chemical Structure of PEGDA.

In PEGDA structure, the macromolecular poly(ethylene glycol) (PEG) is the backbone of the macromer and the diacrylate is the endgroup. After the formation of a PEG hydrogel by photocrosslinking, the three-dimensional crosslinked PEG network forms the scaffold with the acrylate endgroups saturated into chemical crosslinking points. Two or more photocrosslinkable moieties are present on one macromer chain to create a water insoluble network. Crosslinking moieties are generally acrylate derivatives, including methacrylate, methacrylic anhydride, acrylamide and propylene fumarate. The candidates for the macromer backbone are hydrophilic macromers with the molecular weight mostly varying from 10^3 to 10^5 Dalton. Examples of hydrophilic macromers include water-soluble polyethers (such as PEG), polyesters, polyols (such as PVA), polysaccharides and polyester-ether dendrimers.^[9]

Two classes of photocrosslinked hydrogels are PEG gels and polysaccharide (polyol) gels. PEG (or polyethylene oxide, PEO) is a widely used biomaterial due to its superior properties of biocompatibility, hydrophilicity, and mechanical flexibility. These attributes also make it an ideal candidate as a hydrogel scaffold. PEG macromers can be roughly divided into two categories: PEG-backbone and PEG-branched. The PEG-backbone class comprises versions of PEGDA with a template similar to a block copolymer, (meth)acrylates-*Fg*-PEG-*Fg*-(meth)acrylate, where *Fg* represents functional groups for degradation or biological signaling. The PEG-branching class generally encompasses comblike grafting (coupling) of PEG to create co-polymers. The PEG side chains with (meth)acrylate endgroups are pendent branching chains or chains with a comb attached to a backbone. Functional groups can be hybridized either on the ends of PEG chains or on the macromer backbone.

Preparing polysaccharide macromers requires the conjugation of (meth)acrylate moieties onto polysaccharide chains. A large number of sugar hydroxyl groups (OH) or other reactive functional groups such as amine (NH_2), carboxyl (COOH), or sulfonic/sulfate (SO_3/SO_4) groups are utilized as the conjugating sites. Glycidyl methacrylate and methacrylic anhydride may

be employed as the methacrylation reagents. The polysaccharide species used for photocrosslinked hydrogels include alginate, hyaluronan,^[10] dextran,^[11] chitosan,^[12] heparin,^[13] and chondroitin sulfate.^[14] As natural macromolecules, polysaccharides have the advantages of biocompatibility, biodegradation, and specific biological/biophysical functions. For example, chondroitin sulfate, a major component of native cartilage, can be methacrylated to form a hydrogel (Eq. 14, an example of a methacrylated polysaccharide).



Eq. 14 Chemical structure of chondroitin sulfate-glycidyl methacrylate.

Crosslinking

The pore size of hydrogels depends on the molecular dimensions and hydrophilicity of the macromer as well as the actual crosslinking density. The crosslinking density is determined by the (meth)acrylate density and crosslinking efficiency. In general, a highly crosslinked and/or hydrophobic polymer scaffold tends to form hydrogels with a smaller pore size. The hydrogel crosslinking efficiency, C_e , can be calculated by Eq. 15,^[15] which includes the addition polymerization efficiency of the macromer's di-acrylate endgroups.

$$C_e = \left(\frac{\overline{M}_n}{\overline{M}_c} \right) / f_n \quad (15)$$

Eq. 15 Calculation of crosslinking efficiency, C_e .

where f_n is the number of the macromer's crosslinking functionality. For example, each PEGDA macromer molecule (with two acrylate endgroups) may maximally crosslink with four other macromer molecules with acrylate endgroups. Thus f_n is 4, \overline{M}_c represents the average molecular weight shared by each crosslinking functionality of the macromer in the crosslinked network, and \overline{M}_n represents the number

average molecular weight of the macromer. \overline{M}_c can be calculated using the Peppas–Merrill model,^[16] which originates from the Flory–Rehner model and is applicable to systems with the solvent present during crosslinking. The Peppas–Merrill model is expressed by Eq. 16.

$$\frac{1}{\overline{M}_c} = \frac{2}{\overline{M}_n} - \frac{\left(\frac{\overline{v}_0}{\overline{V}_1} \right) [\ln(1 - v_s) + v_s + \chi_1 v_s^2]}{v_r \cdot \left[\left(\frac{v_s}{v_4} \right)^{\frac{1}{3}} - \frac{1}{2} \cdot \left(\frac{v_s}{v_r} \right) \right]} \quad (16)$$

Eq. 16 Calculation of average molecular weight per crosslinking functionality, \overline{M}_c .

where v_0 is the specific volume of the polymer/macromer, \overline{V}_1 is the molar volume of the swelling agent (water), v_s is the polymer volume fraction in the equilibrium-swollen hydrogel, χ_1 is the Flory polymer–solvent interaction parameter, and v_r is the polymer volume fraction in the hydrogel immediately after crosslinking but before swelling, which is equivalent to the volume concentration of the solution where crosslinking occurs. The values of v_s can be calculated by Eq. 17:

$$v_s = \frac{Q_v - W_a / \rho_{\text{water}}}{Q_v} \quad (17)$$

Eq. 17 Calculation of polymer volume fraction in the equilibrium-swollen hydrogel, v_s .

where ρ_{water} is the density of water, Q_v is the volume-swelling ratio of the gels, and W_a is the equilibrium water content (weight) in the gels. The expression $C_e \cdot f_n$ is the average number of crosslinking that occurs on each macromer molecule. If the polymer/macromer's f_n value is not greater than 2, or the system's average C_e value is not greater than 2/4, gelation will not occur. This is also demonstrated by the Peppas–Merrill model for water-soluble macromers ($\chi_1 < 0.5$) in crosslinking-swelling system ($v_s < v_r$)—the \overline{M}_c value is always smaller than $\overline{M}_n/2$.^[15]

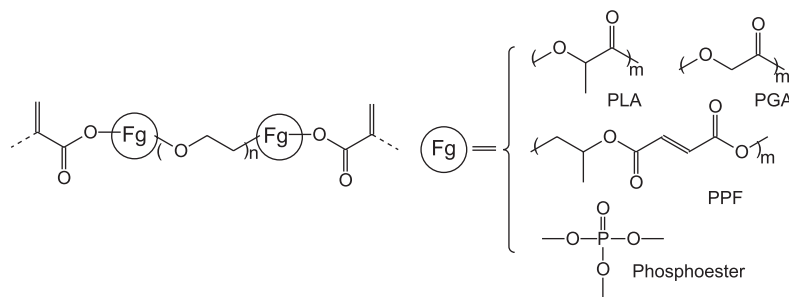
Degradation

The in-vitro degradation of PEGDA hydrogels is based on the hydrolytic cleavage of the degradable groups into macromers (at the F_g site of the template “(meth)acrylates- F_g -PEG- F_g -(meth)acrylate”). Degradable groups that have been attached as F_g include oligo PLA/PGA,^[17] PPF,^[18] and phosphoesters.^[15] The corresponding degradation products are the water-soluble macromers and bioresorbable oligo (meth)acrylate,

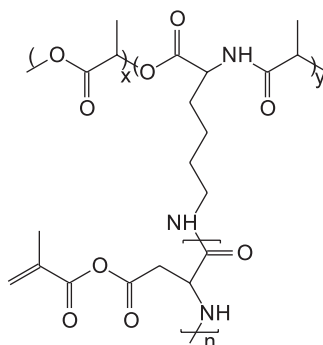
PEG (when molecular weight <10,000 Da), as well as nontoxic small molecules (Eq. 18).

A second form of hydrogel degradation relies on the backbone-degradation of polyesters, polypeptides, polysaccharides, and their combinations. Polyester backbones are usually composed of PLA/PGA, poly(ϵ -caprolactone) (PCL), poly phosphoester, and PPF. Polysaccharide backbones include alginate, hyaluronan, chitosan, and chondroitin sulfate (Eq. 14) and protein backbones comprise collagen,^[19] gelatin^[13,20] and fibrinogen.^[21] The combination of polyester backbone with poly(aspartic acid)-grafted combs with methacrylate groups (Eq. 19),^[22] and the combination of polyphosphoester backbone with PEG-grafting co-polymer methacrylated with methacryloyl chloride (Eq. 20) have been reported.^[23] An example of a combination degradable hydrogel is a (meth)acrylated linear polyester-polyether (PEG) hybridized dendrimer (Eq. 21).^[9]

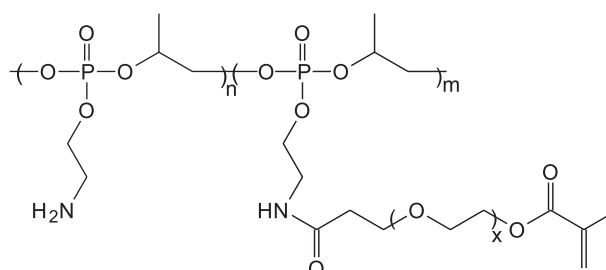
A traditional hydrolytic mechanism of polymer degradation occurs through cleavage of ester groups. Variation of environment temperature and pH can significantly affect the degradation kinetics of the ester groups. For example, alkaline conditions dramatically accelerate poly(phosphoester) degradation by neutralizing the acidic degradation products.^[24] When the environmental temperature and pH are stable, as in the physiological conditions, hydrogel degradation is impacted by three factors: the chemical (hydrolysis) property, the density of degradable groups, and the pore size of the gel which controls the water content and the efficiency of water-polymer interaction. These three factors can be manipulated to control degradation. For example, hydrolysis may be accelerated by increasing the density of degradable sites or the pore size. However, hydrolytic degradation can be difficult to control due to lack of specificity.



Eq. 18 Chemical structure of functionalized PEGDA-type macromers. (From Refs. [13,15,16].)



Eq. 19 Chemical structure of polyester-g-poly (aspartic acid) macromer. (From Ref. [22].)



Eq. 20 Chemical structure of polyphosphoester-g-PEG macromer. (From Ref. [23].)

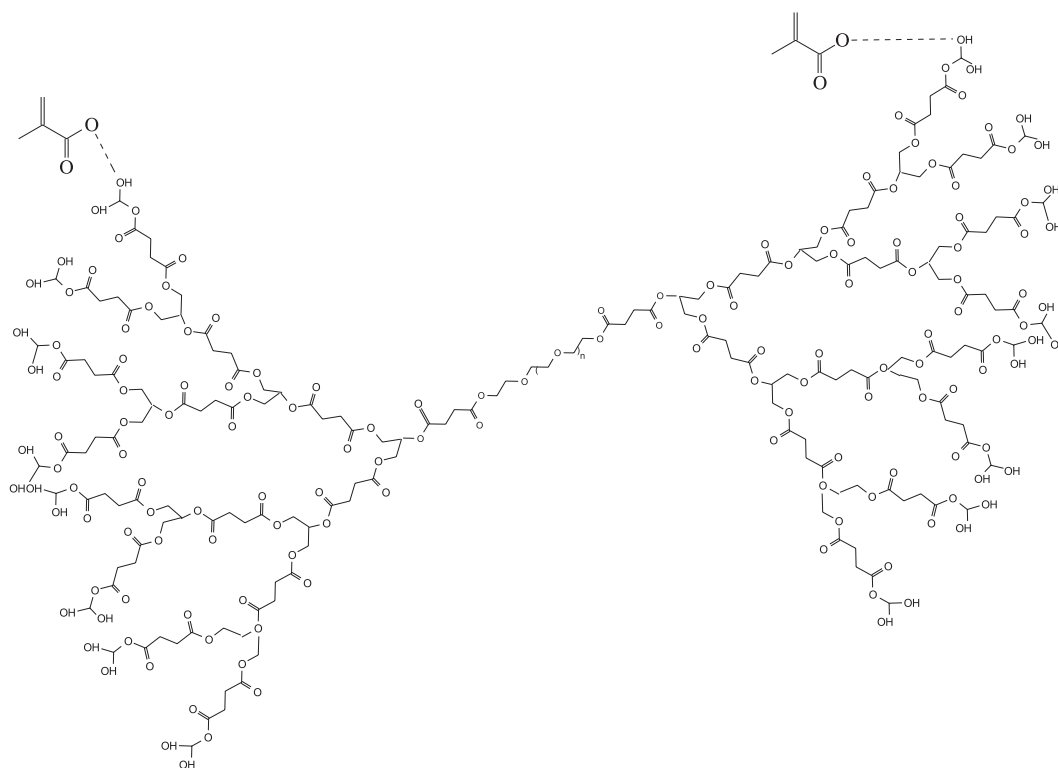
The synchronization of biomaterial scaffold degradation and tissue development is important for optimizing tissue engineering systems. Biological strategies may be investigated to create a material that can specifically respond to its environment. The concept of *biodegradation* is presented as a method to enzymatically cleave polymeric scaffolds and hydrogels. Bio-responsive degradation of scaffolds is triggered by specific changes in the environment. In recent investigations of cartilage–bone tissue engineering scaffolds, chondroitinase ABC that is produced by mature chondrocytes and alkaline phosphatase that is secreted by osteocytes were respectively used for the biodegradation of chondroitin sulfate gels and phosphoester gels.^[24] Results indicate the bioresponsive relationships between the relevant tissue-produced enzymes and the degradation of the corresponding biomaterials.

Functionalization

Photopolymerized biomaterials generally play the roles of carriers or fillers for biological systems. In addition to the physical properties and functions, biomaterials are also expected to fulfill biological functions, such as promoting cell proliferation, inducing cell differentiation, or enhancing the growth of surrounding tissues for defect

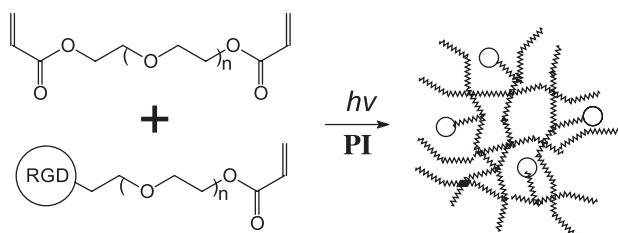
regeneration. A straightforward method to incorporate function is to create directly photocrosslinkable hydrogels made up of natural macromolecules or macromolecular blends. To promote cartilage and bone healing, (meth)acrylated polysaccharide gels and blends of PEOA-type gels encapsulating extracellular matrix extracts “Matrigel”^[25] or “Cartrigel”^[26] have been developed. Hyaluronic acid (HA) was also utilized in a blend with PEOA gels with the dual roles as a process additive for the viscosity control and a functional additive for tissue concrescence (Genzyme, Inc.).

For tissue-engineering purposes, a number of enzymatic and peptide biological signals are employed as functional additives in hydrogels to improve development of engineered tissues. For hard tissue regeneration, transforming growth factor- β (TGF- β)^[27] and bone morphogenetic protein-2 (BMP-2) have been hybridized with photopolymerized hydrogels. The arginine–glycine–aspartic acid tripeptide sequence (RGD) is widely used to improve cell adhesion on biomaterial surfaces and in hydrogels.^[28] Biological signals such as growth factors can be physically encapsulated in the hydrogels with cells. TGF- β and BMP-2 have been combined into mesenchymal stem cells (MSC)—PEGDA photopolymerizing gel systems *in vitro* and *in vivo*, respectively, for chondrogenesis and osteogenesis. To prevent the loss of these additives



Eq. 21 Chemical structure of methacrylated linear polyester–polyether PEG) hybridized dendrimer. (From Ref. [9].)

from the hydrogels via water leaching or extraction, they can be covalently tethered to the hydrogel. Tethering sites can be on the reactive groups of the polymer backbone such as the double bond in propylene fumarate of PPF (Eq. 18), or at the ends of the pendent chain residues, as in the free amino groups (NH_2) of the polyphosphoester-*g*-PEG macromers shown in Eq. 20. Another choice is to synthesize the functionalized and (meth)acrylated mono/macromers separately, and then mix with the standard mono/macromers (such as PEGDA) for co-photogelation (co-photopolymerization/crosslinking). The strategy is schematically illustrated in Eq. 22 using an example of RGD-PEGDA tethering system. Methods for covalent tethering include commonly used bioconjugate techniques, such as the sulfonyl (acyl) chloride, carbonyl diimine (CDI), *N*-hydroxy succinimide (NHS), and divinyl sulfone routes.



Eq. 22 RGD-PEGDA gel tethering system by co-photogelation.

Photografting

Photografting is a technique used for biomaterial surface modification. The basic strategy is: 1) surface photooxidation of a polyether/ester matrix by UV irradiation and hydrogen peroxide to generate reactive (peroxide) radicals; 2) the surface radicals initiate in situ photografting and successive photopolymerization (usually low polymerization degree) of vinyl (acrylate) monomer/macromers with UV irradiation but without photoinitiators; 3) functional moieties (groups, compounds, or polymers) are conjugated for material surface functionalization based on the reactive pendent groups of the photografted poly/oligomers (Eq. 4). This photografting technique has been applied to the surface modification of man-made blood vessel materials and hard-tissue substitution materials. Polyurethane elastomers, including poly(ether/ester urethane)s, are employed as the matrix material for blood-contact devices for cardiovascular tissue engineering. Reactive surface modifying additives, 2-hydroxyethyl (meth)acrylate (HEMA or HEA), acrylamide (reactive amine group containing), and methacrylic diethyldithiocarbamate^[29] can be photografted using the strategy described above. Then with the reactive groups ($-\text{OH}$ or $-\text{NH}_2$),

hydrophilic polymer chains and/or biological ligands are conjugated onto the surface. PEG chains, because of their high hydrophilicity, chain-mobility, and steric exclusion, are often individually grafted as the surface inhibitors for platelet adhesion. PEG chains are also used as spacers for biological ligands, playing the roles of physical supporter and mobility provider at the water interface of the materials. Ligands include surface stearyl groups or Cibacron Blue that is capable of specifically binding blood albumins to make a surface inert and nonthrombogenic. Surface-activating signals like the RGD sequence are also used to promote the adhesion and proliferation of endothelial cells on the modified surface achieving endothelialization.^[30] For cartilage-tissue engineering, degradable PLA/PGA and PCL are often employed as scaffold matrix materials. Using similar photooxidation and photografting strategies, cell(chondrocytes)-growth-promoting ligands are covalently introduced onto the surface of the degradable polymeric scaffolds to improve cartilage tissue regeneration.

PHOTOPOLYMERIZATION FOR TISSUE-ENGINEERING APPLICATIONS

Tissue engineering is the application of the principles and methods of engineering and the life sciences toward the fundamental understanding of structure–function relationships in normal and pathological tissue and the development of biological substitutes to restore, maintain, or improve functions. Photopolymerized biomaterials have been applied to tissue-engineering applications with numerous roles including non-cell involved fillers, or dressings; cell growth guiders or promoters; and matrixes for two- and three-dimensional cell culture. Acellular photopolymerized biomaterials comprise dental fillers and defect dressings for hard tissues. Practical systems include PEGDA-HA blends for bone and cartilage dressings (Genzyme, Inc.) and hernia fixing. PolyHEMA or chondrotin sulfate gels are being used as the fillers for bone repair. Photopolymerized hydrogels are also believed to have the talent of inducing and guiding cells to regenerate tissues^[15,20] Sharma et al. are developing strategies to organize chondrocytes and hydrogels to promote chondrocyte proliferation and three-dimensional tissue production, impossible in two-dimensional systems. For applications in two-dimensional cell culture systems, photopolymerization techniques are generally utilized for surface modification. Biological signals or ligands are immobilized onto the biomaterial surfaces using photografting approaches, by which the cell–material interfaces are optimized for cell adhesion and proliferation. For example, as described in photografting section, RGD was conjugated onto polyurethane elastomers via

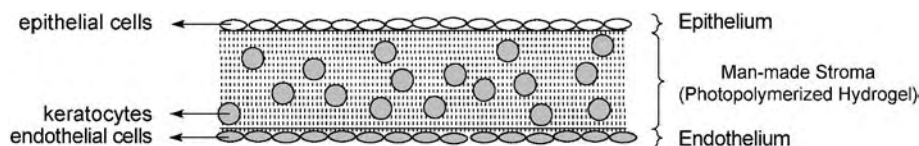


Fig. 1 Schematic illustration of man-made cornea using photopolymerized hydrogels. (From Ref. [33].)

photografting to construct endothelial cell-engineered surfaces for cardiovascular tissue engineering. Unlike endothelial cells, which naturally grow on two-dimensional surfaces, the native environment for most musculoskeletal cells is in three dimensions. Common methods to create a three-dimensional environment are 1) to encapsulate cells in hydrogels or 2) to place cells on porous biomaterial scaffolds. Photopolymerization strategies have been applied to the latter method as demonstrated by photografting of biological signals to optimize the material-cell interfaces. BMP or RGD-conjugated porous materials of biodegradable PLA, PGA, and PCL have been used to enhance scaffolds for cartilage repair.^[31] A unique approach for three-dimensional culture is cell encapsulation by photogelation of injectable macromer-cell-photoinitiator systems. This has been applied to hard tissue regeneration, liver engineering (with hepatocytes),^[32] and bioartificial corneas.^[33]

Applying hydrogels to tissue engineering has numerous advantages. Hydrogels are composed primarily of water, similar to the composition of biological systems. Mechanical properties and water can be manipulated to suit a particular application. The crosslinked hydrophilic polymer network is capable of suspending and immobilizing cells, whereas it also allows for the free penetration of water, nutrients, and metabolic waste, even allowing limited cell migration. A unique tissue-engineering application of hydrogels is engineered corneas. The biomimetic structures may be constructed into a multilayered structure with a superficial epithelial monolayer, deep endothelial monolayer, and a photopolymerized hydrogel encapsulating keratocytes, as a transparent substitute of cornea stroma, in the middle (Fig. 1).^[33]

Hard Tissue Repair

Among the current tissue-engineering applications using photopolymerized biomaterials, entries for hard tissue

repair are the most numerous. A significant number represent dental substitutions (fillings) and musculoskeletal repair templates. The fundamental questions lie in the construction of implants and their integration between the implants and the surrounding tissues. The formulation of the photopolymerization precursors is always composed of mono/macromers, photoinitiators and diluents. The mono/macromers are usually di(meth)acrylate compounds. For dental materials, small molecules or stiff polymers, such as the aromatic ring-containing 2,2-bis[4-(2'-hydroxy-3'-methacryloyl-oxypropoxy) phenyl] propane (bis-GMA) or 1,6-bis(2'-methacryloyl-oxyethoxycarbonylamino)2,4,4-trimethylhexane (UDMA), are applied with relatively high solid density. The system is complemented with inorganic fillers to improve compatibility and strength. The photoinitiator CQ (Eq. 8) is widely used with visible blue light. The photopolymerized (crosslinked) system has a low shape/size distortion and low water absorption, as well as good adhesion, making it ideal as a dental filler.

For the musculoskeletal dressings or implants typically used on bones or cartilage, the diluents are usually water or aqueous fluids and cells are encapsulated in the implants. The in situ injection protocol for (knee) articular cartilage repair is demonstrated in Fig. 2.^[34]

The procedure demonstrates in situ injection of the precursors for polymerization and then in situ photogelation (cell encapsulation) with UV irradiation. The in vivo trials have been performed in nude mice by transdermal injection (Fig. 3)^[35,36] and also in defects created on the articular surface of a goat model (Fig. 4).^[34]

Chondrocyts and osteoblasts are respectively encapsulated for cartilage and bone regeneration. Mesenchymal stem cells (MSC) may also be employed for three-dimensional cultivation in hydrogels. The in-vitro and in-vivo studies of chondrogenesis and osteogenesis were performed using the corresponding growth factors of TGF- β and BMP-2 to promote stem cell differentiation. Furthermore, a novel technique of separately encapsulating different cell lines or cell types within a multilayer hydrogel was developed to mimic the zonal

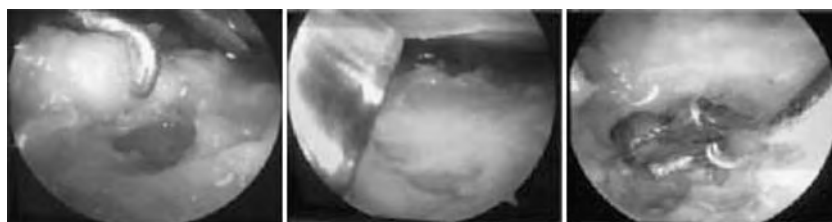


Fig. 2 In situ injection and photogelation of hydrogels for human (knee) articular cartilage repair. (From Ref. [34].)

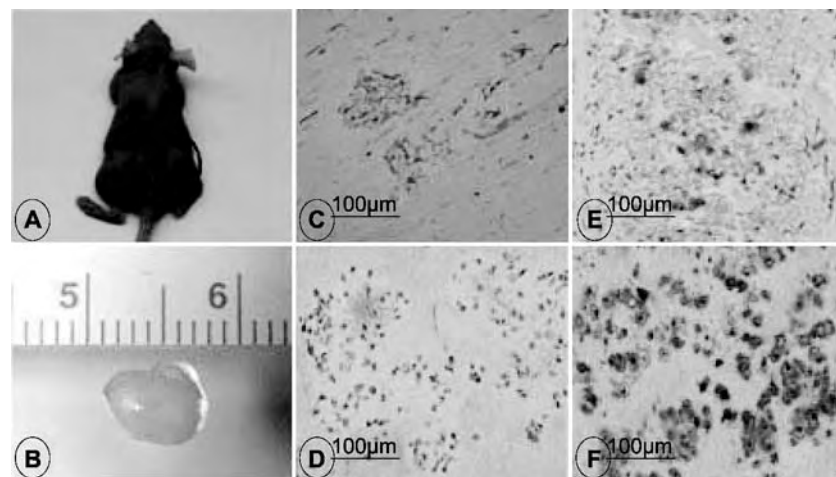


Fig. 3 A) Transdermal implants made with PEODA + HA (left) and PEODA + HA + TGF (right) 3 weeks post implantation. Circle highlights one construct. B) Construct containing PEODA + HA + MSCs harvested after 3 weeks (~8 mm diam.). Safranin-O staining of C) PEODA + HA + MSCs, D) PEODA + MSCs + TGF, E) PEODA + HA + MSCs and F) PEODA + HA + MSCs + TGF. (From Refs. [35] and [36].)

organization and behavior of chondrocytes respectively in the superficial, middle, and deep levels of native cartilage.^[37]

Integration of biomaterials with the body has been a longstanding problem in medicine. Lack of proper biomaterial implant integration with the body sacrifices longevity and function. Integration of biomaterials with hard tissues such as cartilage, which lacks the ability to self-repair, is particularly challenging due to the nature of the dense extracellular matrix and strong mechanical

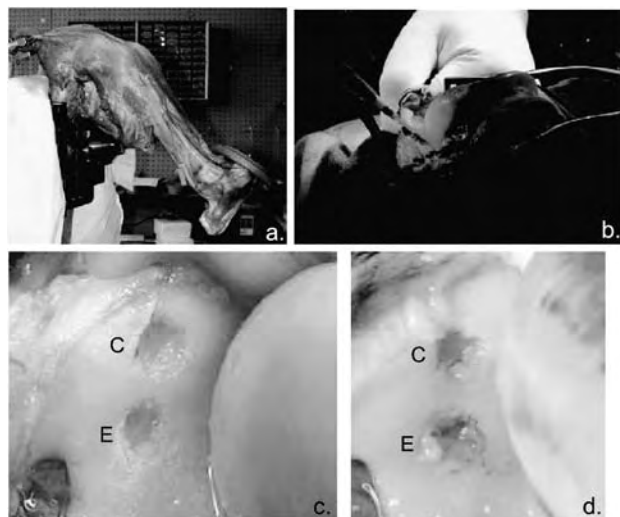


Fig. 4 a) In-vitro bovine leg undergoing mechanical simulator studies to determine hydrogel integrity in chondral and osteochondral defects. b) In-situ photopolymerization of a MSC-hydrogel construct in a medial tibial defect (5 mm diam) in a goat in vivo. c) Chondral defects in goat before hydrogel placement. d) The same defects as in c after one month. E, experimental defect with photopolymerized MSC-hydrogel and C, control defect with no hydrogel. (From Ref. [34].)

forces that the tissue must withstand. Traditional strategies rely on adhesive primers for acellular dressings. Wang and Elisseff developed a novel strategy of tissue-initiated photopolymerization for the integration of the (PEGDA) hydrogels and the native cartilage tissue.^[5] The technique is based on: 1) enzymatic clearance of the proteoglycans in cartilage to expose the collagen network; 2) in situ generation of tyrosyl radicals by photo-oxidation of tyrosine residues on collagen with H_2O_2 under low intensity UV irradiation; and 3) introduction of a macromer solution and in situ photogelation via tyrosyl radical initiation and UV excitation. Since the reaction is initiated by components of the tissue, the macromers are covalently grafted onto the tissue and crosslinked (gelled) to form a hydrogel without exogenous photoinitiators. A cartilage-hydrogel implant created by the tissue-initiated polymerization is pictured in Fig. 5.^[5]

Step 3 has been improved for cell encapsulation by controlling the degree of photopolymerization. Instead of achieving sufficient photogelation, which needs relatively tough reaction conditions, Step 3 is terminated at the termination of photografting (also the photoinitiation) of diacrylate macromer but prior to their further propagation or crosslinking. The free acrylate groups covalently introduced onto the cartilage surface (with PEG spacers) act as the primers for integration. Then living cells can be implanted through the standard protocol with PEGDA macromers and photoinitiators. The diacrylate macromers in the implant will be gelled via self-photocrosslinking and also simultaneously crosslinked with the primer acrylate groups. Therefore, cell encapsulation and the integration with cartilage are both fulfilled. This integrating technique, besides using the principles of basic protein biochemistry (the generation of tyrosyl radicals), is due to a combined strategy of photografting and photopolymerization.

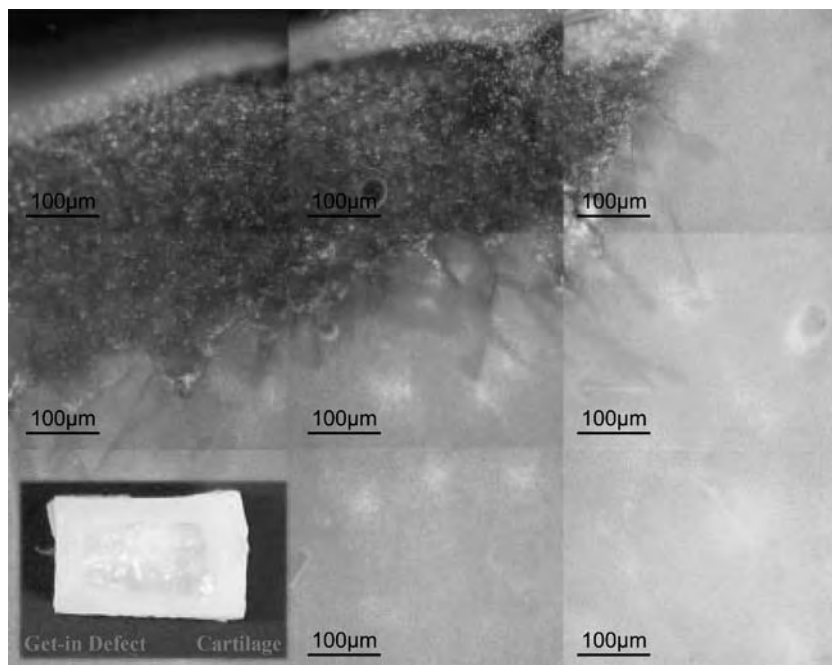


Fig. 5 Photographic illustrations and fluorescent Life/Dead assay (the living and dead cells respectively visualize in green and red by the excitation of blue and green fluorescence) for hydrogel–cartilage integration. (From Ref. [5].)

P

ARTICLES OF FURTHER INTEREST

Biofunctional Polymers; Cartilage Regeneration; Glues; Hydrogels; Microelectromechanical Systems (MEMS) Manufacturing; Rapid Prototyping; Surface Modification; Tissue Engineering of Bone

REFERENCES

1. Peiffer, R.W. Applications of Photopolymer Technology. In *Photopolymerization Fundamentals and Applications*; Scranton, A.B., Bowman, C.N., Peiffer, R.W., Eds.; ACS Symposium Series, American Chemical Society: Washington, DC, 1996; Vol. 673, 1–15.
2. Belfield, K.D.; Crivello, J.V. *Photoinitiated Polymerization*; American Chemical Society: Washington, DC, 1996.
3. Feng, X.D.; Voong, S.T.; Sun, Y.H.; Qiu, K.Y. Selective grafting of hydrogels onto multiphase block copolymers. *Makromol. Chem., Macromol. Chem. Phys.* **1985**, *186* (8), 1533–1541.
4. Barr, D.P.; Gnther, M.R.; Deterding, L.J.; Tomer, K.B.; Mason, R.P. ESR spin-trapping of a protein-derived tryosyl radical from the reaction of cytochrome c with hydrogen peroxide. *J. Biol. Chem.* **1996**, *271* (26), 15498–15503.
5. Wang, D.A.; Williams, C.G.; Elisseeff, J.H. Enhancing the tissue-biomaterial interface: Tissue-initiated integration of biomaterials. *Adv. Funct. Mater.* **2003**. Submitted.
6. Sipani, V.; Scranton, A.B. Dark-cure studies of cationic photopolymerizations of epoxides: Characterization of the active center lifetime and kinetic rate constants. *J. Polym. Sci. Polym. Chem.* **2003**, *41* (13), 2064–2072.
7. Bryant, S.J.; Nuttelman, C.R.; Anseth, K.S. Cytocompatibility of UV and visible light photoinitiating system on culture NIH/3T3 fibroblast in vitro. *J. Biomater. Sci. Polym. Ed.* **2000**, *11* (5), 439–457.
8. Williams, C.G.; Malik, A.N.; Kim, T.K.; Manson, P.N.; Elisseeff, J.H. Variable cytocompatibility of six cell lines with photoinitiators used for polymerizing hydrogels and cell encapsulation. *Biomaterials* **2003**. Submitted.
9. Carnahan, M.A.; Middleton, C.; Kim, J.; Kim, T.; Grinstaff, M.W. Hybrid dendritic–linear polyester–ethers for in situ photopolymerization. *J. Am. Chem. Soc.* **2002**, *124* (19), 5291–5293.
10. Semds, K.A.; Grinstaff, M.W. Photocrosslinkable polysaccharides for in situ hydrogel formation. *J. Biomed. Mater. Res.* **2001**, *54* (1), 115–121.
11. van Dijk-Wolthuis, W.N.E.; Franssen, O.; Talsma, H.; van Steenberg, M.J.; Kettenes-van den Bosch, J.J.; Hennink, W.E. Synthesis, characterization, and polymerization of glycidyl methacrylate derivatized dextran. *Macromolecules* **1995**, *28* (18), 6317–6322.
12. Francis Suh, J.-K.; Matthew, H.W.T. Application of chitosan-based polysaccharide biomaterials in cartilage tissue engineering: A review. *Biomaterials* **2000**, *21* (24), 2589–2598.
13. Matsuda, T.; Magoshi, T. Preparation of vinylated polysaccharides and photofabrication of tubular scaffolds as potential use in tissue engineering. *Biomacromolecules* **2002**, *3* (5), 942–950.
14. Li, Q.; Wang, D.A.; Elisseeff, J.H. Heterogeneous-phase reaction of glycidyl methacrylate and chondroitin sulfate: Mechanism of ring-opening-transesterification competition. *Macromolecules* **2003**, *36* (7), 2556–2562.
15. Wang, D.A.; Williams, C.G.; Li, Q.; Sharma, B.; Elisseeff, J.H. Synthesis, photopolymerization, and characterization of a novel biodegradable phosphate-containing hydrogel. *Biomaterials* **2003**, *24* (22), 3969–3980.

16. Peppas, N.A.; Merrill, E.W. Polyvinyl-alcohol hydrogels—Reinforcement of radiation—Crosslinked networks by crystallization. *J. Polym. Sci. Polym. Chem.* **1976**, *14* (2), 441–457.
17. Elbert, D.L.; Hubbell, J.A. Conjugate addition reactions combined with free-radical cross-linking for the design of materials for tissue engineering. *Biomacromolecules* **2001**, *2* (2), 430–441.
18. Timmer, M.D.; Ambrose, C.G.; Mikos, A.G. In vitro degradation of polymeric network of poly(propylene fumarate) and the crosslinking macromer poly(propylene fumarate)-dicacrylate. *Biomaterials* **2003**, *24* (4), 571–577.
19. Poshusta, A.K.; Anseth, K.S. Photopolymerized biomaterials for application in the temporomandibular joint. *Cells Tissue Organs* **2001**, *169*, 272–278.
20. Nakayama, Y.; Kim, J.Y.; Nishi, S.; Ueno, H.; Matsuda, T. Development of high-performance stent: Gelations photogel-coated stent that permits drug delivery and gene transfer. *J. Biomed. Mater. Res.* **2001**, *57* (4), 559–566.
21. Halstenberg, S.; Panitch, A.; Rizzi, S.; Hall, H.; Hubbell, J.A. Biologically engineered protein-graft-poly(ethylene glycol) hydrogels: A cell adhesive and plasmin-degradable biosynthetic material for tissue repair. *Biomacromolecules* **2002**, *3* (4), 710–723.
22. Elisseeff, J.H.; Anseth, K.; Langer, R.; Hrkach, J.S. Synthesis and characterization of photo-cross-linked polymers based on poly(L-lactic acid-co-aspartic acid). *Macromolecules* **1997**, *30* (7), 2182–2184.
23. Li, Q.; Wang, J.; Sun, D.D.N.; Shahani, S.; Sharma, B.; Elisseeff, J.H.; Leong, K.W. Biodegradable and photo-crosslinkable polyphosphoester hydrogel. *Biomaterials* **2004**. Submitted.
24. Wang, D.A.; Williams, C.G.; Cher, N.; Lee, S.; Elisseeff, J.H. Bioresponsive phosphoester hydrogels for bone tissue. *Adv. Funct. Mater.* **2003**. Submitted.
25. Sefton, M.V.; May, M.H.; Lahooti, S.; Babensee, J.E. Making microencapsulation work: Conformal coating, immobilization gels and in vivo performance. *J. Control. Release* **2000**, *65* (1–2), 173–186.
26. Lum, L.Y.B.; Cher, N.L.C.; Williams, C.G.; Elisseeff, J.H. An extracellular matrix extract for tissue-engineered cartilage. *IEEE Eng. Med. Biol.* **2003**, *22* (5), 71–76.
27. Mann, B.K.; Schmedlen, R.H.; West, J.L. Tethered-TGF- β increases extracellular matrix production of vascular smooth muscle cells. *Biomaterials* **2001**, *22* (5), 439–444.
28. Burdick, J.A.; Anseth, K.S. Photoencapsulation of osteoblasts in injectable RGD-modified PEG hydrogels for bone tissue engineering. *Biomaterials* **2002**, *23* (22), 4315–4323.
29. Luo, N.; Hutchison, J.B.; Anseth, K.S.; Bowman, C.N. Surface-initiated photopolymerization of poly(ethylene glycol) methyl ether methacrylate on a diethyldithiocarbamate-mediated polymer substrate. *Macromolecules* **2002**, *35* (7), 2487–2493.
30. Guan, J.J.; Gao, C.Y.; Feng, L.X.; Sheng, J.C. Surface photo-grafting of polyurethane with 2-hydroxyethyl acrylate for promotion of human endothelial cell adhesion and growth. *J. Biomater. Sci. Polym. Ed.* **2000**, *11* (5), 523–536.
31. Cao, Y.L.; Vacanti, J.P.; Paige, K.T.; Upton, J.; Vacanti, C.A. Transplantation of chondrocytes utilizing a polymer–cell construct to produce tissue-engineered cartilage in the shape of a human ear. *Plast. Reconstr. Struct.* **1997**, *100* (2), 297–302.
32. Liu, V.A.; Bhatia, S.N. Three-dimensional photopatterning of hydrogels containing living cells. *Biomed. Microdevices* **2002**, *4* (4), 257–266.
33. Griffith, M.; Hakim, M.; Shimmura, S.; Watsky, M.A.; Li, F.; Carlsson, D.; Doillon, C.J.; Nakamura, M.; Suuronen, E.; Shinozaki, N.; Nakata, K.; Sheardown, H. Artificial human corneas scaffolds for transplantation and host regeneration. *Cornea* **2002**, *21* (Suppl. 2), S54–S61.
34. Williams, C.G.; Anseth, K.; Gurkan, I.; Sharma, B.; Parks, B.; Wang, D.A.; Meyers, J.D.; Wenz, J.; Elisseeff, J.H. Cartilage repair using a photocurable-mesenchymal stem cell hydrogel in large animal model. ORS'50, The (50th Annual Meeting of the Orthopedic Research Society, San Francisco, CA, U.S.A., March 7–10, 2004.
35. Elisseeff, J.; Anseth, K.; Sims, D.; McIntosh, W.; Randolph, M.; Langer, R. Transdermal photopolymerization for minimally invasive implantation. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (6), 3104–3107.
36. Williams, C.G.; Kim, T.K.; Taboas, A.; Malik, A.; Manson, P.; Elisseeff, J.H. In vitro chondrogenesis of bone marrow-derived mesenchymal stem cells in a photopolymerizing hydrogel. *Tissue Eng.* **2003**, *9* (4), 679–688.
37. Kim, T.K.; Sharma, B.; Williams, C.G.; Ruffner, M.A.; Malik, A.; McFarland, B.G.; Elisseeff, J.H. Experimental model for cartilage tissue engineering to regenerate the zonal organization of articular cartilage. *Osteoarthritis Cartilage* **2003**, *11* (9), 653–664.

Piezoelectric Materials

Zoubeida Ounaies

Virginia Commonwealth University, Richmond, Virginia, U.S.A.

INTRODUCTION

This article provides a brief history and definition of piezoelectricity, followed by an introduction to a variety of piezoelectric materials, with a focus on processing, characterization, and properties. We will illustrate the impact of piezoelectricity on biomaterials and biomedical applications through a description of applications and functions.

PIEZOELECTRICITY

The piezoelectric effect is a linear effect closely related to the microscopic structure of the material. Piezoelectricity stems from the Greek term *piezo* (for pressure); it follows that piezoelectricity is the generation of electrical polarization in a material in response to a pressure or mechanical stress. This phenomenon is known as the direct effect. Piezoelectric materials always display the converse effect in which a mechanical deformation ensues upon application of an electrical charge or signal.

Piezoelectricity is a property of many noncentrosymmetric (lacking a center of symmetry) ceramics, polymers, and other biological systems. The moment-generating nature of this phenomenon makes it ideal for sensor and actuator applications, underlining its similarity to biological-type behavior in which a system senses a change in its environment and reacts to that change through a proportional response.

A subset of piezoelectricity is pyroelectricity. A pyroelectric material is a material that exhibits a change in polarization in response to a uniform temperature variation. Some pyroelectric materials are also ferroelectric; a ferroelectric material possesses a spontaneous polarization that can be reversed by application of a realizable electric field. The application of an AC field large enough to yield polarization-switching results in a hysteresis loop. A typical field-polarization hysteresis loop is illustrated in Fig. 1. One key feature of the loop is the remanent polarization P_r , indicative of the potential piezoelectric performance of the material. At moderate to large electric field levels, inherent nonlinearities (as demonstrated by the hysteresis loop) are observed. At any drive level, some hysteretic behavior is always present. However, at relatively low electric

field levels, the electromechanical response can be regarded as a linear phenomenon. It is then governed by the linear constitutive equations shown in upcoming sections.

HISTORY OF PIEZOELECTRICITY

Pierre and Jacques Curie discovered the piezoelectric effect in 1880 during their study of the effects of pressure on the generation of electrical charge by tourmaline crystals. They subsequently showed that quartz, cane sugar, and Rochelle salt also exhibited the piezoelectric effect. By building upon their work on the symmetry of crystals and their earlier observations on pyroelectricity, the Curie brothers were able to predict the classes of crystals and the conditions under which piezoelectricity would be observed.^[1] In 1881, the term piezoelectricity was first suggested by W. Hankel, and the converse effect was deduced by Lipmann from thermodynamics principles the same year.^[2] The Curie brothers immediately confirmed the converse effect experimentally and proved the complete reversibility of the electrical and mechanical responses.

In the next three decades, collaborations within the European scientific community established the field of piezoelectricity; by 1910, Voigt's *Lerbuch der Kristallphysic* was published and became a standard reference work detailing the complex electromechanical relationships in piezoelectric crystals.^[3] However, the complexity of the science of piezoelectricity made it difficult for it to mature to an application phase until a few years later. In 1916, Paul Langevin developed a piezoelectric ultrasonic detector using the first quartz-sandwich-transducer: quartz crystal sandwiched between two metal plates; as the quartz resonates, the plates emit sound waves through the water. Another quartz device measures the rebounding waves, and the time between emitting and rebounding waves is determined and used to calculate the distance to a possible reflecting source.^[4] This was clearly the forerunner to the sonar. His success opened up opportunities for piezoelectric materials in underwater applications as well as a host of other applications such as ultrasonic transducers, microphones, accelerometers, etc. In 1935, Busch and Scherrer discovered piezoelectricity in potassium dihydrogen phosphate (KDP).^[5] The KDP family was the first major



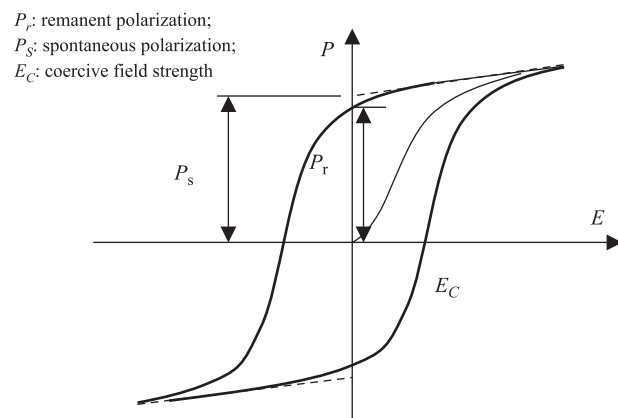


Fig. 1 Typical electric field-polarization hysteresis loop.

family of piezoelectric materials to be discovered. During World War II, research in piezoelectric materials expanded to the United States, the Soviet Union, and Japan. Until then, limited performance by these materials inhibited commercialization, but that changed when a major breakthrough came with the discovery of barium titanate (BT) and lead zirconate titanate (PZT) in the 1940s and 1950s, respectively. These families of materials exhibited very high dielectric and piezoelectric properties. Furthermore, they offered the possibility of tailoring their behavior to specific responses and applications by the use of dopants (i.e., a foreign substance added to a material in a relatively small concentration to alter its properties). To date, PZT is one of the most widely used piezoelectric materials. Most commercially available ceramics (such as KDP, BT, and PZT) are based on the perovskite structure (Fig. 2). The perovskite structure (ABO_3) is the simplest arrangement in which the corner-sharing oxygen octahedra are linked together in a regular cubic array with smaller cations (Ti, Zr, Sn, Nb, etc.) occupying the central octahedral B-site, and larger cations (Pb, Ba, Sr, Ca, Na, etc.) filling the interstices between octahedra in the larger A-site. Compounds such as $BaTiO_3$, $PbTiO_3$, $PbZrO_3$,

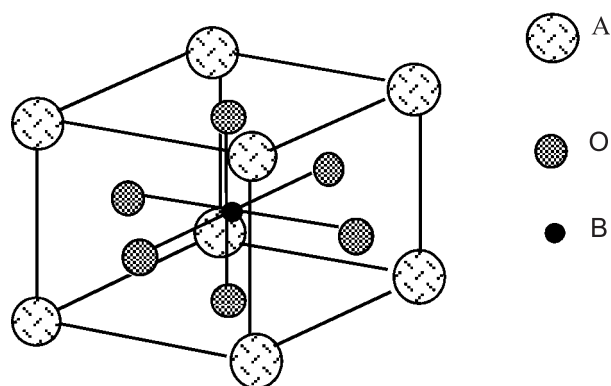


Fig. 2 Perovskite structure (chemical structure ABO_3).

$NaNbO_3$, and $KNbO_3$ have been studied at length, and their high-temperature ferroelectric and antiferroelectric phases have been extensively exploited. This structure also allows for multiple substitutions on the A- and B-sites resulting in a number of useful, though more complex, compounds such as $(Ba,Sr)TiO_3$, $(Pb,Sr)(Zr,Ti)O_3$, $Pb(Fe,Ta)O_3$, $(K,Bi)TiO_3$, etc. It follows that tailoring of the properties is made possible due to the flexibility the perovskite structure affords through judicious selection of dopants and substitutions.

Starting around 1965, several Japanese companies focused on developing new processes and applications, and opening new commercial markets for piezoelectric devices. The success of the Japanese effort attracted other nations, and today the needs and uses of piezoelectric materials extend from medical applications to the communications field, military applications, automotive field, and the consumer industry. Prior to that, pioneering work by Fukada et al. paved the way to extending the field of piezoelectricity to organic materials. As early as the 1950s, Fukada and coworkers investigated the piezoelectric behavior of collagen and other biological polymers, culminating in their utilization of tendon and bone as gramophone pickups in 1959.^[6] However, due to the low piezoelectric coefficient of biological polymers, the applications, and hence the interest, were very limited. This was soon changed by one significant discovery: that of piezoelectricity in stretched and poled poly(vinylidene fluoride) (PVDF) by Kawai in 1969.^[7] He showed that thin films of PVDF that had been poled exhibited a very large piezoelectric coefficient, 6–7 pC/N, a value that is about ten times larger than had been observed in any other polymer.

For more on the early history of piezoelectricity, the reader is referred to an excellent summary focusing on the 19th- and early 20th-century developments by Trainer.^[8] Two other notable works are a text by W. G. Cady^[9] and the 1971 Jaffe, Cook, and Jaffe book *Piezoelectric Ceramics*,^[10] still one of the most referenced works on piezoelectricity.

PIEZOELECTRIC CLASSES

The unit cell contains the smallest grouping of atoms representing the chemical formula of the crystalline material. The specific symmetry of a unit cell determines the possibility of piezoelectricity in the crystal under consideration. For example, if the unit cell has no center of inversion, an external stress causes a displacement of charges, and the charge distribution is no longer symmetric. A net polarization develops and results in an internal electric field. In general, the shapes of the unit cells fall into seven general crystal systems: cubic, tetragonal, trigonal, monoclinic, triclinic, orthorhombic, and

hexagonal. Crystallographers further classify crystal systems based on point groups. All crystals can be divided into 32 classes or point groups, emerging from three types of symmetry elements: pure rotation axes of symmetry (one-, two-, three-, four-, and sixfold), planes of reflection (mirror) symmetry, and axes of rotary inversion (rotation followed by inversion through a center). The point groups determine the form and symmetries of the elastic, piezoelectric, and dielectric matrices that appear in the constitutive relations (more on that in subsequent sections). Table 1 shows the seven crystal systems and the 32 point groups. As illustrated in the table, of the 32 point groups, 20 are both non-centrosymmetric and piezoelectric. The lack of center of symmetry means that a net movement of the positive and negative ions with respect to each other as a result of stress produces an electric dipole. The ceramic, being composed of the random orientation of these piezoelectric crystallites, is inactive, i.e., the effects from the individual crystals cancel each other and no discernable piezoelectricity is present. Regions of equally oriented polarization vectors are known as domains. “Poling” is a commonly used method to orient the domains by polarizing the ceramic through the application of a static electric field. The electrodes are applied to the ceramic (through painting, spraycoating, evaporating,

screenprinting, sputtering, etc.) and a sufficiently high electric field is applied such that the domains rotate and switch in the direction of the electric field. The result is never a full orientation of all domains; however the polycrystalline ceramic exhibits a large piezoelectric effect. During this process, there is a small expansion of the material along the poling axis and a contraction in both directions perpendicular to it (see Fig. 3).

PIEZOELECTRIC CONSTITUTIVE EQUATIONS

When the constitutive equation for a piezoelectric material is written, account must be taken of changes of strain and electrical displacement in three orthogonal directions caused by cross-coupling effects due to applied electrical and mechanical stresses. Tensor notation is first adopted; the reference axes are shown in Fig. 4. The state of strain is described by a second-rank tensor S_{ij} , and the state of stress is also described by a second-rank tensor T_{kl} . The relationships relating the stress tensor to the strain tensor, compliance s_{ijkl} and stiffness c_{ijkl} , are then fourth-rank tensors. The relationship between the electric field E_j (first-rank tensor) and the electric displacement D_i (also a first-rank tensor) is the permittivity ϵ_{ij} , which is a

Table 1 Crystallographic point groups

CRYSTAL SYSTEM	CENTRIC POINT GROUPS		ACENTRIC POINT GROUPS					OPTIC AXES
			POLAR		NONPOLAR			
TRICLINIC	$\bar{1}$		1		NONE			B I A X I A L
MONOCLINIC	2/m		2	m	NONE			
ORTHORHOMBIC	mmm		mm2		222			
TETRAGONAL	4/m	4/mmm	4	4mm	$\bar{4}$	$\bar{4}2m$	422	U N I A X I A L
TRIGONAL	$\bar{3}$	$\bar{3}m$	3	3m	32			
HEXAGONAL	6/m	6/mmm	6	6mm	$\bar{6}$	$\bar{6}m2$	622	
CUBIC	m3	m3m	NONE		23	$\bar{4}3m$	432	ISOTROPIC
	11 GROUPS		10 GROUPS		11 GROUPS			

Piezoelectric^a

^aThe point group 432 (cubic), although lacking a center of symmetry, does not permit piezoelectricity due to a combination of the other symmetries. (Source: Ref. [11].)

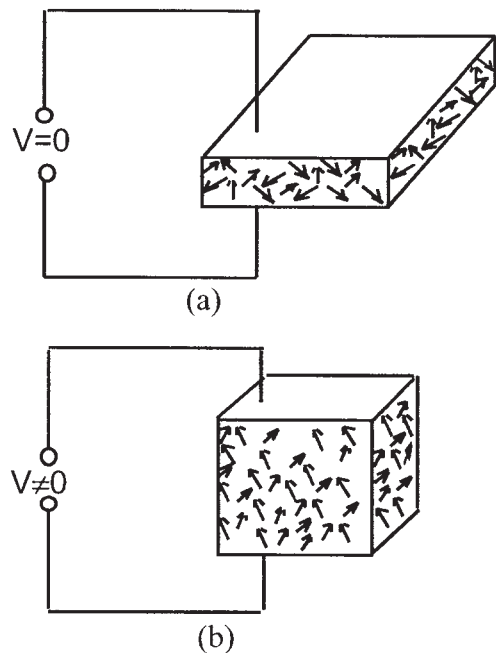


Fig. 3 Poling of a polycrystalline material, (a) unpoled, (b) poled.

second-rank tensor. The piezoelectric equations are therefore written as:

$$D_i = \varepsilon_{ij}^T E_j + d_{ijk} T_{jk} \quad (1)$$

$$S_{ij} = \underline{d}_{ijk} E_k + s_{ijkl}^E T_{kl} \quad (2)$$

where d_{ijk} and \underline{d}_{ijk} are the piezoelectric constants (third-rank tensor). Superscripts T and E denote that the dielectric constant ε_{ij} and the elastic constant s_{ijkl} are measured under conditions of constant stress and constant electric field, respectively. In general, a first-rank tensor has three components, a second-rank tensor has nine components, a third-rank tensor has 27 components, and a fourth-rank tensor has 81 components. Not all the tensor components are independent. Both of these relations (i.e., Eqs. 1 and 2) are direction-dependent; they describe a set of equations that relate these properties in different orientations of the material. The crystal symmetry and the choice of reference axes reduce the number of independent components. A convenient way of describing them is by using axis directions as given by Fig. 4. The convention is to define the poling direction as the 3-axis; the shear planes are indicated by the subscripts 4, 5, and 6, and are perpendicular to directions 1, 2, and 3, respectively. This simplifies the notations introduced above, where a 3-subscript tensor notation ($i,j,k=1,2,3$) is replaced by a 2-subscript matrix notation ($i=1,2,3$ and $j=1,2,3,4,5,6$), and a 2-subscript tensor notation ($i,j=1,2,3$) is replaced by a 1-subscript matrix notation ($i=1,2,3,4,5,6$).

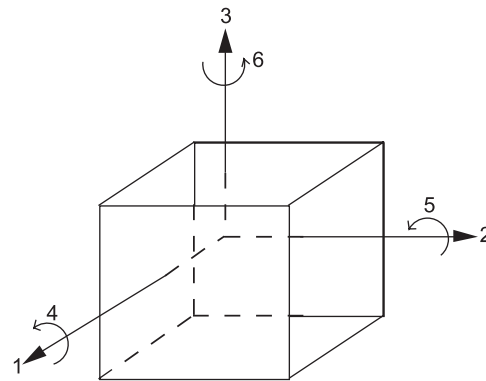


Fig. 4 Reference axes.

A shear strain such as S_4 is a measure of the change of angle between the two initially orthogonal axes in the plane perpendicular to axis 1. The first subscript of the d -constant gives the “electrical” direction (field or dielectric displacement) and the second gives the component of mechanical deformation or stress.

PIEZOELECTRIC CERAMICS

Barium Titanate

Barium titanate is a perovskite with the chemical formula BaTiO_3 (see Fig. 2, where $A = \text{Ba}$, $B = \text{Ti}$). BT has a cubic structure above 120°C , which transforms into a tetragonal structure upon cooling below 120°C . Another transition is found at 0°C , going from tetragonal to orthorhombic, and a third one at -90°C , from orthorhombic to rhombohedral. The most important transition in the piezoelectricity of BT is the one that takes place at 120°C . This temperature is known as the Curie temperature T_c , and the phase transition takes BT from the paraelectric cubic structure (symmetric structure) to the ferroelectric tetragonal structure (non-centrosymmetric). In BT, the spontaneous polarization that gives rise to piezoelectricity and ferroelectricity is a consequence of the positioning of the Ba^{2+} , Ti^{4+} , and O^{2-} ions within the unit cell. The Ba^{2+} ions are located at the corners of the unit cell, the O^{2-} ions are located near, but slightly below, the centers of each of the six faces, and the Ti^{4+} ion is displaced upward from the unit cell center. The resulting dipole moment is a consequence of the relative displacements of the Ti^{4+} and O^{2-} ions from their symmetrical positions (see Fig. 5). BT was the first piezoelectric ceramic to be developed and exploited commercially. However, since the discovery of PZT a decade later, BT has been largely replaced by PZT, which has piezoelectric coefficients. Since then, use of BT has shifted to the capacitor industry, specifically the multilayer capacitor for its high dielectric constant.

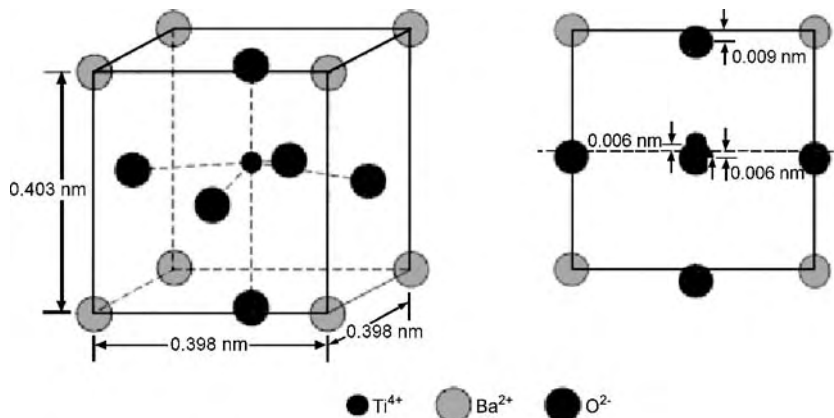


Fig. 5 Tetragonal barium titanate crystal structure.

Lead Zirconate Titanate

Lead zirconate titanate has the chemical formula $\text{Pb}(\text{Ti}, \text{Zr})\text{O}_3$. Similar to BT, it has a perovskite crystalline structure. It is the most widely used piezoelectric ceramic due to the combination of properties it affords. Figure 6 shows the phase diagram for PZT.^[10] As can be ascertained from the figure, T_c (transition from paraelectric to ferroelectric) is anywhere between 240 and 480°C depending on the ratio of Ti to Zr in the solid solution. The implication is that piezoelectricity can now be exploited over a wide temperature range, which is not possible for the BT system. The phase boundary transition between the tetragonal structure and the rhombohedral structure, referred to as the morphotropic phase boundary (MPB), is nearly independent of temperature. Both the dielectric constant and the piezoelectric response increase dramatically at the MPB. Table 2 lists the in-plane coefficients of PZT. To further increase the performance of PZT, and improve its resistance to fracture, researchers have developed actuators based on the ceramic configured

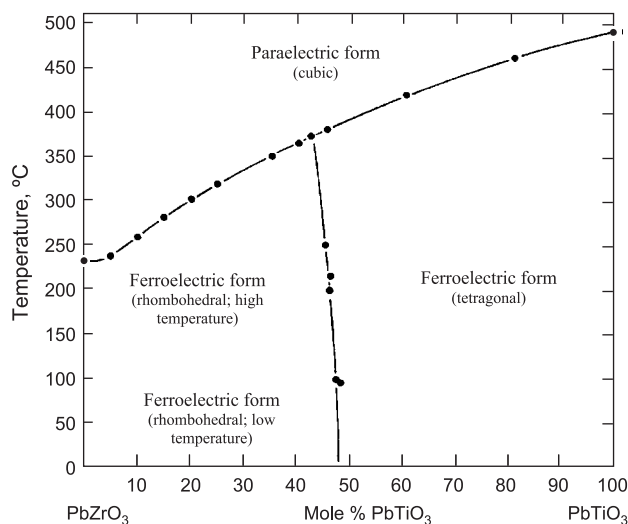


Fig. 6 Phase diagram of PZT. (From Ref. [10].)

with a metal backing or other similar layers.^[12,13] Among these devices are Unimorphs[®], Bimorphs[®], RAINBOWs[™], and THUNDERS[™], to name a few. The result is a more durable and flexible actuator, in which the ceramic is often present in a state of pre-stress.

P

PIEZOELECTRIC POLYMERS

The properties of polymers are different in comparison to those of inorganics (Table 2) in that they are uniquely qualified to fill niche areas where single crystals and ceramics are incapable of performing as effectively. As noted in Table 2, the piezoelectric strain constant (d_{31}) for the polymer is lower than that of the ceramic. However, piezoelectric polymers have much higher piezoelectric stress constants (g_{31}) indicating that they are much better sensors than ceramics. The fourth column of Table 2 reveals that the electromechanical coupling coefficient k_{31} for the ceramic is almost three times that of the polymer. The electromechanical coupling coefficient describes the conversion of energy from mechanical to electrical and, vice versa, electrical to mechanical. As such, it is a measure of the piezoelectric strength in materials.^[10] One would then conclude that the piezoelectric ceramic shown has a higher piezoelectric performance than the polymer. Piezoelectric polymeric sensors and actuators offer the advantage of processing flexibility because they are lightweight, tough, and readily manufactured into large areas, and can be cut and formed into complex shapes. Polymers also exhibit high strength and high impact resistance.^[14] Other notable features of polymers are low dielectric constant, low elastic stiffness, and low density, which result in a high voltage sensitivity (excellent sensor characteristic) and low acoustic and mechanical impedance (crucial for medical and underwater applications). Polymers also typically possess a high dielectric breakdown and high operating field strength, which means that they can withstand much higher driving fields than ceramics. Polymers offer the ability

Table 2 Comparison of piezoelectric polymers and piezoelectric ceramics^a

	d_{31} (pm/V)	g_{31} (mV-m/N)	k_{31}	Features
Polyvinylidene fluoride (PVDF)	28	240	0.12	Better sensor, flexible, lightweight, low acoustic and mechanical impedance
Lead zirconium titanate (PZT)	175	11	0.34	Better actuator, brittle, heavy, load-carrying capabilities

^a d_{31} = piezoelectric strain constant; g_{31} = piezoelectric stress constant; k_{31} = electromechanical coupling coefficient.

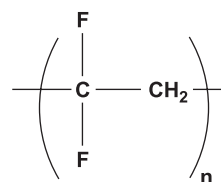
to pattern electrodes on the film surface, and pole only selected regions. Based on these features, piezoelectric polymers possess their own established area for technical applications and useful device configurations.

Synthetic Piezoelectric Polymers

PVDF

Polyvinylidene fluoride (PVDF) is a fluoropolymer that, due to a combination of processability, mechanical strength, and good chemical resistance, has been used as an inert lining or pipework material in chemical plants. However, thanks to Kawai's discovery in 1969,^[7] it became better known for its piezoelectric properties and has been extensively investigated since. PVDF is inexpensive, flexible, and lightweight, and accounts for virtually all of the commercially significant piezoelectric polymer application.

PVDF is a semicrystalline polymer with typically 50 to 60% crystalline content, depending on thermal and processing history. Generally, PVDF is synthesized by the free-radical polymerization of 1,1-difluoroethylene, and the repeating unit is $-\text{CH}_2-\text{CF}_2-$ (Fig. 7). The amorphous phase in PVDF has a glass transition that is well below room temperature (-50°C); hence the material is quite flexible and readily strained at room temperature. There are two conditions that must be fulfilled in order to render PVDF piezoelectric, namely, the crystal structure should possess a nonzero net dipole moment, i.e., it must be polar, and the dipoles should have the same preferred orientation. PVDF has four polymorphic forms: α , β , γ , and δ phases. The α -phase is the most common and stable form of PVDF. Normally, the α -phase is obtained by melt extrusion. It is nonpolar as it produces an inversion of the dipole moments in

**Fig. 7** Structure of PVDF.

alternate chains, which leads to a cancellation of the overall dipole moment. The β -phase is the most important polymorph of PVDF because it is polar and is predominantly responsible for the piezoelectric properties. Usually, the β -phase is obtained by mechanical stretching of the α -phase film. Generally, stretching is either uniaxial or biaxial, which yield different properties. The γ -phase is also a polar unit cell. Generally, the γ -phase is obtained by solution crystallization and readily transforms into the β -phase upon mechanical deformation. The δ -phase is the polar form of α -phase. It is obtained by applying a high electric field to an α -phase film. As a result, the unit-cell dimensions and the configuration are the same as in the α -phase. As mentioned earlier, the β -phase is the form most commonly used for piezoelectricity because it results in the higher piezoelectric coefficients. After mechanical stretching to produce the polar β -phase, the poling process begins with heating of the film and application of a high electric field. During the poling process, the dipoles are aligned relative to the direction of the poling field. The polarization achieved depends on the applied field and on poling temperature and poling time. When the piezoelectric film is operating in an electromechanical or motor mode, the film elongates and contracts as the polarity of the alternating field changes. When operating in a mechano-electrical or generator mode, external forces are applied that produce compressive and tensile strain. These deformations cause a change in the surface charge density, resulting in a voltage between the electrodes, the polarity of which changes as the direction of the force is reversed.

PVDF copolymers

Copolymers of polyvinylidene fluoride with trifluoroethylene (TrFE) and tetrafluoroethylene (TFE) have also been shown to exhibit strong piezoelectric, pyroelectric, and ferroelectric effects. Here, these polymers are discussed together because they behave similarly when copolymerized with PVDF. An attractive morphological feature of the comonomers is that they force the polymer into an all-*trans* conformation that has a polar crystalline phase, thus eliminating the need for mechanical stretching to yield a polar phase.

P(VDF-TrFE) crystallizes (up to 90%) to a much greater extent than PVDF yielding a higher remanent polarization, lower coercive field, and much sharper hysteresis loops. TrFE also extends the use temperature by about 20°, to close to 100°C. Conversely, copolymers with TFE have been shown to exhibit a lower degree of crystallinity and a suppressed melting temperature as compared to the PVDF homopolymer. Although the piezoelectric constants for the copolymers are not as large as for the homopolymer, the advantages of P(VDF-TrFE) associated with processability, enhanced crystallinity, and higher use temperature make it favorable for applications.

Recently researchers have reported that highly-ordered lamellar crystals of P(VDF-TrFE) can be made by annealing the material at temperatures between the Curie temperature and the melting point. They refer to this material as a single crystalline film. A relatively large single crystal P(VDF-TrFE) (75/25 mol%) copolymer has been grown that exhibits a room temperature $d_{33} = -38$ pm/V and a coupling factor $k_{33} = 0.33$.^[15]

Nylons

Kawai et al. first reported piezoelectricity in nylons (polyamides) in 1970.^[16] In 1980, Scheinbeim reported improved levels of piezoelectricity in odd-numbered nylons.^[17] One remaining limitation is that the presence of the hydrogen bonding in the nylons renders them hydrophilic; studies have shown that water absorption can have a deleterious effect on their dielectric and piezoelectric properties.^[18] Water absorption in nylon-11 and nylon-7 can reach as high as 4.5% (by weight) and more than 12% for nylon-5,^[19] whereas it is less than 0.02% for PVDF and its copolymers.

The monomer unit of odd nylons consists of even numbers of methylene groups and one amide group with a dipole moment of 3.7 D. Nylons crystallize in all-*trans* conformations and are packed so as to maximize hydrogen bonding between adjacent amine and carbonyl groups for an even- and an odd-numbered polyamide. For the odd-numbered monomer, the dipoles result in a net dipole moment. The amide dipole cancels for an even-numbered nylon, although remanent polarizations have been measured for some even-numbered nylons.^[20] The unit dipole density is dependent on the number of methylene groups present, and the polarization increases with decreasing number of methylene groups from 58 mC/m² for nylon-11 to 125 mC/m² for nylon-5.^[17,18,20]

At room temperature, odd-numbered nylons have lower piezoelectric constants than PVDF; however, annealing of the films enhances their polarization stability through a denser packing of the hydrogen-bonded sheet structure in the crystalline regions. This results in a reduction of free volume for rotation and hence a hindrance of the dipole switching.^[18,20]

Typical values of the piezoelectric constants for some of the semicrystalline polymers discussed above are given in Table 3.

Amorphous polymers

The piezoelectricity in amorphous polymers differs from that in semicrystalline polymers and inorganic crystals in that the polarization is not in a state of thermal equilibrium, but rather a quasistable state due to the freezing-in of molecular dipoles. The result is a piezoelectriclike effect. One of the most important properties of an amorphous piezoelectric polymer is its glass transition temperature (T_g , temperature below which

Table 3 Comparison of piezoelectric properties of some semicrystalline polymeric materials^a

Polymer	Structure	T_g (°C)	T_m (°C)	Max use temp (°C)	d_{31} (pC/N)
PVDF	$\left(\begin{array}{c} \text{F} \\ \\ \text{---C---CH}_2 \\ \\ \text{F} \end{array} \right)_n$	-35	175	80	20–28
PTrFE	$\left(\begin{array}{c} \text{F} \quad \text{F} \\ \quad \\ \text{---C---CH} \\ \\ \text{F} \end{array} \right)_n$	32	150	90–100	12
Nylon-11	$\left(\begin{array}{c} \text{O} \\ \\ \text{---NH---(CH}_2\text{)}_{10}\text{---C} \\ \\ \text{---} \end{array} \right)_n$	68	195	185	3 @ 25°C 14 @ 107°C

^a T_g = glass transition temperature; T_m = melting temperature; d_{31} = piezoelectric strain constant.

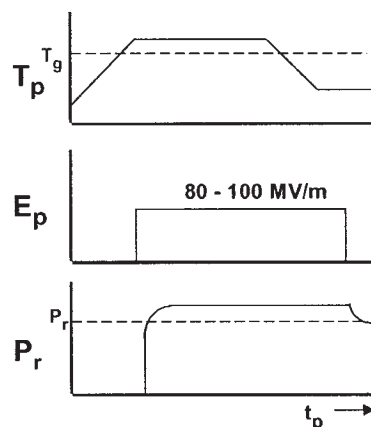


Fig. 8 Poling profile of an amorphous polymer, where T_p is the poling temperature, E_p is the poling electric field, P_r is the remanent polarization, and t_p is the poling time.

the material exhibits glasslike characteristics, and above which it has rubberlike properties) as it dictates temperature for use and defines the poling process conditions. Orientation polarization of molecular dipoles is responsible for piezoelectricity in amorphous polymers. It is induced, as shown in Fig. 8, by applying an electric field E_p at an elevated temperature ($T_p \geq T_g$) where the molecular chains are sufficiently mobile and allow dipole alignment with the electric field. Partial retention of this orientation is achieved by lowering the temperature below T_g in the presence of E_p , resulting in a piezoelectriclike effect. The remanent polarization P_r is directly proportional to E_p and the piezoelectric response. It is a measure of the degree of piezoelectricity in the polymer. The following linear equation for the remanent polarization results when the Clausius Mossotti equation is used to relate the dielectric constant to the dipole moment.^[21]

$$P_r = \Delta\epsilon \epsilon_0 E_p \quad (3)$$

It can be concluded that the remanent polarization P_r and hence the piezoelectric response of a material is determined by $\Delta\epsilon$, the change in the dielectric constant below and above T_g , making it a practical criterion to use when designing piezoelectric amorphous polymers. The dielectric relaxation strength $\Delta\epsilon$ may be the result of either free or cooperative dipole motion.

The literature on amorphous piezoelectric polymers is much more limited than that for semicrystalline systems.^[20] This is due in part to the fact that no amorphous piezoelectric polymers have exhibited responses high enough to attract commercial interest. Much of the previous work on amorphous piezoelectric polymers resides in the area of nitrile-substituted polymers including polyacrylonitrile (PAN),^[22–24] poly

(vinylidene cyanide vinylacetate) (PVDCN/VAc),^[25–28] polyphenylethynitrile (PPEN),^[29,30] and poly(1-bicyclobutanecarbonitrile).^[31] Weak piezoelectric activity in polyvinyl chloride (PVC) and polyvinyl acetate (PVAc) has also been found.^[32–34] The most promising of these materials are the vinylidene cyanide copolymers that exhibit large dielectric relaxation strengths and strong piezoelectricity. Table 4 shows molecular structures of the most commonly encountered amorphous piezoelectric polymers.

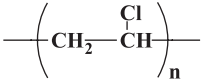
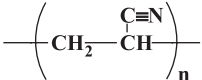
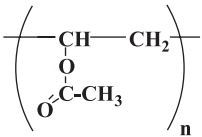
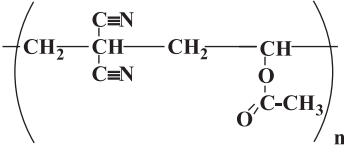
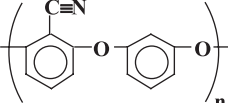
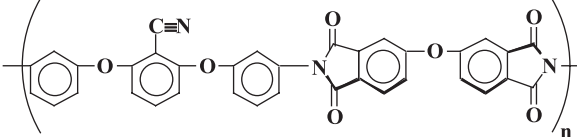
Biological Piezoelectric Polymers

Piezoelectricity of biopolymers was first reported for keratin in 1941.^[35] Consequently, other biological materials have been found to be piezoelectric, including tendon, dentin, ivory, aorta, trachea, intestine, wood, and collagen;^[36,37] polypeptides such as poly- γ -methylglutamate and poly- γ -benzyl-L-glutamate;^[38,39] oriented films of DNA,^[40] polylactic acid,^[41] and chitin,^[42] and the nucleic acids. Bone, however, has been the most frequently studied tissue. Fukada and his group investigated the piezoelectricity in bone in depth starting as early as 1957.^[43] Other researchers subsequently confirmed their findings and further contributed to this area.^[44–46] The piezoelectric coefficient d_{14} , which relates a shear stress applied along the long axis of a bone to the electric displacement that appears on a surface perpendicular to the axis (Fig. 4), is the critical coefficient in this case. The piezoelectricity of bone is of great interest to many researchers because it could shed light on bone physiology. For example, it is known that electric polarization in bone influences bone growth^[46] and, therefore, one could hypothesize that the piezoelectric signal in bone could regulate bone growth.

Because most natural biopolymers possess D_∞ symmetry, they exhibit shear piezoelectricity. As mentioned earlier, a shear stress in the plane of polarization produces electric displacement perpendicular to the plane of the applied stress, resulting in a $-d_{14} = d_{25}$ piezoelectric constant. The piezoelectric constants of biopolymers are small relative to those of synthetic polymers, ranging in value from 0.01 pC/N for DNA to 2.5 pC/N for collagen. However, investigating the piezoelectric mechanism in these biopolymers may help us develop synthetic polymers with higher piezoelectric shear coefficients, uniquely suited to some applications such as flow sensing, for example.

At present, piezoelectricity in biopolymers is an undisputable, albeit still a phenomenological, fact. The physiological significance of piezoelectricity, however, is not yet fully understood, and that opens up a whole potential of exciting and challenging studies.

Table 4 Structure, polarization, and T_g for piezoelectric amorphous polymers^a

Polymer	Structure	T_g (°C)	d_{31} (pC/N)	P_r (mCm ²)	Reference
PVC		80	5	16	[33]
PAN		90	2	25	[24]
PVAc		30	–	5	[34]
P(VDCN-VAc)		170	10	50	[26]
PPEN		145	–	12	[29]
(β-CN) APB/ODPA		220	5 @ 150°C	20	[34]

^a T_g = glass transition temperature; d_{31} = piezoelectric strain constant; P_r = remanent polarization.

APPLICATIONS

The application potential for piezoelectric and other electroactive polymers is immense. To date, ferroelectric polymers have been incorporated into numerous sensing and actuation devices for a wide array of applications. Typical applications include devices in medical instrumentation, robotics, optics, computers, and ultrasonic, underwater, and electroacoustic transducers. One important emerging application area for electroactive polymers is in the biomedical field in which polymers are being explored as potential artificial muscle actuators, as invasive medical robots for diagnostics and microsurgery, as actuator implants to stimulate tissue and bone growth, and as sensors to monitor vascular grafts and to prevent blockages.^[47,48] Such applications are ideal for polymers as they can be made to be biocompatible and they have excellent conformability and impedance matching to body fluids and human tissue. The intent of this manuscript is not to detail specific applications, but the interested reader may refer to excellent sources on applications of piezoelectric and ferroelectric polymers.^[49–51]

CONCLUSION AND FUTURE DIRECTIONS

It has been more than 120 years since the discovery of piezoelectricity. Great strides have been accomplished both on the material and on the application sides. Both piezoelectric ceramics and piezoelectric polymers have become viable for commercial purposes. Early on in the history of piezoelectricity, emphasis was on investigating and mastering the fundamental structure-property relationship in the crystals. This approach has enabled the development of numerous synthetic piezoelectric materials with properties far exceeding that of natural or biological materials. This made possible a number of applications, which further generated more interest in piezoelectricity.

As discussed previously, piezoelectricity was first discovered in biological materials; then, in the next few decades, it centered on new synthetic families of crystals with increasing piezoelectric performance. It is interesting to note that researchers have now come full circle, rediscovering biological polymers and other naturally occurring materials and working on unraveling the piezoelectric phenomenon in them and the physiological

role it plays in the function of the material. Researchers continue to diligently uncover novel ways to characterize complex electromechanical properties and, as they do so, new processing methods and applications are revealed.

ACKNOWLEDGMENT

The author expresses her most sincere appreciation to Dr. J. S. Harrison (NASA Langley Research Center) for providing technical insight in the area of piezoelectric polymers.

ARTICLES OF FURTHER INTEREST

Artificial Muscles; Biological Effects of Radiofrequency Electromagnetic Field; Electroactive Polymeric Materials; Tissue Engineering of Bone; Vascular Grafts: Development Strategies

REFERENCES

- Curie, J.; Curie, P. Développement, par pression, de l'électricité polaire dans les cristaux hémihédres à faces inclinées. *C. R. Acad. Sci.* **1880**, *91*, 294–295.
- Poynting, J.H.; Thomson, J.J. *Textbook of Physics, Electricity and Magnetism*; Griffen: London, 1914; 148–163.
- Voigt, V.W. *Lehrbuch der Kristallphysik*; B. G. Teubner: Leipzig-Berlin, 1910.
- Mason, W.P. Piezoelectricity, its history and applications. *J. Acoust. Soc. Am.* **1981**, *70*, 1561–1566.
- Busch, G.; Scherrer, P. A new seignettelectric substance. *Naturwissenschaften* **1935**, *23*, 737.
- Fukada, E. The piezoelectric effect in fibrous proteins. *Rep. Prog. Polym. Phys. Jpn.* **1959**, *2*, 168–170.
- Kawai, H. The piezoelectricity of poly(vinylidene fluoride). *Jpn. J. Appl. Phys.* **1969**, *8*, 975.
- Trainer, M. Kelvin and piezoelectricity. *Eur. J. Phys.* **2003**, *24*, 535–542.
- Cady, W.G. *Piezoelectricity; An Introduction to the Theory and Applications of Electromechanical Phenomena in Crystals*; Dover Publications: New York, 1964.
- Jaffe, B.; Cook, W.R., Jr.; Jaffe, H. *Piezoelectric Ceramics*; Academic Press: London, 1971.
- Ballato, A. Piezoelectricity: History and New Thrusts, IEEE Ultrasonics Symposium, 1996; 575–583.
- Mossi, K.; Selby, G.; Bryant, R. Thin-layer composite unimorph ferroelectric driver and sensor properties. *Mater. Lett.* **1998**, *35*, 39–49.
- Furman, E.; Li, G.; Haertling, G. An investigation of the resonance properties of RAINBOW devices. *Ferroelectrics* **1994**, *160*, 357–369.
- Davis, G.T. Piezoelectric and Pyroelectric Polymers. In *Polymers for Electronic and Photonic Applications*; Academic Press, Inc.: Boston, MA, 1993; 435.
- Omote, K.; Ohigashi, H.; Koga, K. Temperature dependence of elastic, dielectric and piezoelectric properties of 'single crystalline' films of vinylidene fluoride trifluoroethylene copolymer. *J. Appl. Phys.* **1997**, *81* (6), 2760.
- Kawai, H.; Henji, I. Electrostriction and piezoelectricity of elongated polymer films. *Oyo Butsuri* **1970**, *39*, 413.
- Newman, B.A.; Chen, P.; Pae, K.D.; Scheinbeim, J.I. Piezoelectricity in nylon-11. *J. Appl. Phys.* **1980**, *51* (10), 5161–5164.
- Mei, B.Z.; Schienbeim, J.I.; Newman, B.A. The ferroelectric behavior of odd-numbered nylons. *Ferroelectrics* **1993**, *144*, 51–60.
- Newman, B.A.; Kim, K.G.; Scheinbeim, J.I. The effects of water content on the piezoelectric properties of nylon-11 and nylon-7. *J. Mater. Sci.* **1990**, *25*, 1779–1783.
- Harrison, J.; Ounaies, Z. Piezoelectric Polymers. In *Encyclopedia of Smart Materials*; John Wiley and Sons: New York, 2002.
- Hilzner, B.; Malecki, J. *Electrets: Studies in Electrical and Electronic Engineering*; Elsevier: New York, NY, 1986; Vol. 14; 19.
- Comstock, R.J.; Stupp, S.I.; Carr, S.H. Thermally stimulated discharge currents from polyacrylonitrile. *J. Macromol. Sci., Phys.* **1977**, *B13* (1), 101–115.
- Ueda, H.; Carr, S.H. Piezoelectricity in polyacrylonitrile. *Polym. J.* **1984**, *16*, 661–667.
- von Berlepsch, H.; Kunstler, W.; Wedel, A.; Danz, R.; Geib, D. Piezoelectric activity in a copolymer of acrylonitrile and methylacrylate. *IEEE Trans. Electr. Insul.* **1989**, *24*, 357–362.
- Miyata, S.; Yoshikawa, M.; Tasaka, S.; Ko, M. Piezoelectricity revealed in the copolymer of vinylidenecyanide and vinylacetate. *Polym. J.* **1980**, *12*, 857–860.
- Furukawa, T.; Date, M.; Nakajima, K.; Kosaka, T.; Seo, I. Large dielectric relaxations in an alternate copolymer of vinylidene cyanide and vinyl acetate. *Jpn. J. Appl. Phys.* **1986**, *25*, 1178–1182.
- Seo, I. Piezoelectricity of vinylidene cyanide copolymers and their applications. *Ferroelectrics* **1995**, *171*, 45–55.
- Sakurai, M.; Ohta, Y.; Inoue, Y.; Chujo, R. An important factor generating piezoelectric activity of vinylidene cyanide copolymers. *Polym. Commun.* **1991**, *32*, 397–399.
- Tasaka, S.; Toyama, T.; Inagaki, N. Ferro- and pyroelectricity in amorphous polyphenylethnitrile. *Jpn. J. Appl. Phys.* **1994**, *33*, 5838–5844.
- Takahashi, T.; Kato, H.; Ma, S.P.; Sasaki, T.; Sakurai, K. Morphology of a wholly aromatic thermoplastic, poly(ether nitrile). *Polymer* **1995**, *36*, 3803–3808.
- Hall, H.K.; Chan, R.J.H.; Oku, J.; Hughes, O.R.; Scheinbeim, J.; Newman, B. Piezoelectric activity in films of poly(1-bicyclobutanecarbonitrile). *Polym. Bull.* **1987**, *17*, 135–136.
- Furukawa, T. Piezoelectricity in polymers. *IEEE Trans. Electr. Insul.* **1989**, *24*, 375–393.
- Bharti, V.; Nath, R. Quantitative analysis of piezoelectricity in simultaneously stretched and corona poled polyvinyl chloride films. *J. Appl. Phys.* **1997**, *82*, 3488–3492.
- Ounaies, Z.; Young, J.A.; Harrison, J.S. *An Overview of the Piezoelectric Phenomenon in Amorphous Polymers*; Amer. Chem. Society: Washington, DC, 1999.

35. Martin, A.J.P. Tribo-electricity in wool and hair. *Proc. Phys. Soc.* **1941**, *53*, 183.
36. Fukada, E. Piezoelectricity and Pyroelectricity of Biopolymers. In *Ferroelectric Polymers, Chemistry, Physics and Application*; Marcel Dekker, Inc.: New York, NY, 1995; 393.
37. Maeda, H.; Fukada, E. Effect of water on piezoelectric dielectric and elastic properties of bone. *Biopolymers* **1982**, *21*, 2055.
38. Date, M.; Takashita, S.; Fukada, E. Temperature variation of piezoelectric moduli in oriented poly(γ -methyl-L-glutamate). *J. Polym. Sci.* **1970**, *A-2* (8), 61.
39. Furukawa, T.; Fukada, E. Piezoelectric relaxation in poly(γ -methyl-L-glutamate). *J. Polym. Sci., Polym. Phys.* **1976**, *14*, 1979.
40. Ando, Y.; Fukada, E. Piezoelectric properties of oriented deoxyribonucleate films. *J. Polym. Sci., Polym. Phys.* **1976**, *14*, 63.
41. Fukada, E. Piezoelectric properties of poly-L-lactic acid. *Rep. Prog. Polym. Phys. Jpn.* **1991**, *34*, 269.
42. Ando, Y.; Fukada, E.; Glimcher, M.J. Piezoelectricity of chitin in lobster shell and apodeme. *Biorheology* **1977**, *14*, 175.
43. Fukada, E.; Yasuda, L. On the piezoelectric effect of bone. *J. Phys. Soc. Jpn.* **1957**, *12*, 1158.
44. Bassett, C.A.L.; Becker, R.O. Generation of electric potentials by bone in response to mechanical stress. *Science* **1962**, *137*, 1063.
45. Shamos, M.H.; Lavine, L.S.; Shamos, M.I. Piezoelectric effect in bone. *Nature* **1963**, 1978–1981.
46. Fukada, E. Piezoelectricity of Bone and Osteogenesis of Piezoelectric Films. In *Mechanisms of Growth Control*; Charles C. Thomas ed; Springfield, IL, 1981; 192.
47. Bar-Cohen, Y. *Electroactive Polymer (EAP) Actuators as Artificial Muscles: Reality, Potential and Challenges*; SPIE Press: Bellingham, WA, 2001; 615.
48. DeRossi, D.; Dario, P. *Medical Applications of Piezoelectric Polymers*; Gordon and Breach Science Publishers, 1988; 83.
49. Garner, G.M. *The Applications of Ferroelectric Polymers*; Blackie: London, UK, 1988; 190.
50. Meeker, T.R. *The Applications of Ferroelectric Polymers*; Blackie: London, UK, 1988; 305.
51. Yamaka, E. *The Applications of Ferroelectric Polymers*; Blackie: London, UK, 1988; 329.



Plasma Modification of Materials

Ricky K.Y. Fu

Paul Kim Ho Chu

Department of Physics & Materials Science, City University of Hong Kong, Kowloon, Hong Kong

INTRODUCTION

Biomaterials are typically defined as materials that constitute parts of biomedical implants and devices. Many synthetic materials currently used in biomedical applications include metals and alloys, polymers, and ceramics. Because of their different structures and properties, these materials have different applications. The ideal biomaterials should have good physicochemical properties, while the surface should possess good biocompatibility as the environment inside the body is chemically and mechanically active. Interactions between the implanted biomaterials and body tissues take place on the materials surface and, consequently, biochemical responses and reactions depend on the surface properties. Therefore, the implanted biomaterials must have good biocompatibility and in some instances need to integrate with the surrounding biological environment to restore adequate functions without releasing harmful chemical products or inducing deleterious effects, such as skin irritation, chronic inflammatory, and hypersensitivity responses. Surface modification of biomaterials has become an interesting topic in medicine as an effective means to improve the surface functions and lifetime of biomaterials.

Plasma surface modification (PSM) not only chemically and/or physically alters the surface composition and microstructures via processes such as etching, chemical reactions, sterilization, ion irradiation and bombardment, but also can synthesize a different structure, such as a thin film on the surface of the biomaterials via coating, grafting, and thin film deposition. The new film or surface can possess very different chemical and biological properties from the bulk materials. In these ways, the surface of the biomaterials can be made more biocompatible, while the favorable bulk attributes, such as strength, can be retained. In this entry, common plasma techniques and the properties of plasma-treated biomaterials are described. A number of the examples are related to metals and polymers that are the most widely accepted materials in medical devices, such as total joint replacements, orthopedic implants, dental prostheses, cardiovascular blood vessel stents, artificial heart valves, and breast implants. In all of these cases, the

aim of surface treatment is to enhance the surface properties, such as biocompatibility, bioactivity or bioinertness, corrosion and wear resistance, wettability, adhesion, and resistance against fatigue-related failures.

PLASMA CHARACTERISTICS, PLASMA DISCHARGE TECHNIQUES, AND PLASMA–SURFACE INTERACTIONS

Plasma

Plasma is often called the fourth state of matter. A plasma can be envisioned as a collection of electrons, singly and multiply charged positive and negative ions, neutral atoms, excited particles or radicals, electromagnetic radiation, molecules, and molecular fragments. An important feature of the plasma is that the positive and negative particles are in a state of charge equilibrium and the sum of the positive and negative charges in a sufficiently large volume is equal to zero. In fact, plasma technology is dry, environmentally friendly, and cost-efficient in a myriad of applications.

Plasma Discharge

To produce a plasma, electron separation from atoms or molecules in the gas state, that is, ionization, is required. When an atom or a molecule gains enough energy from an outside excitation source or via interactions (collisions) with one another, ionization occurs. In most practical situations, plasma is produced by an electrical discharge. There are a number of different discharge techniques including radiofrequency (RF) discharge, microwave discharge, glow discharge (GD), and arc discharge. Generally, these techniques can be classified into two categories: gaseous and metallic plasma discharges.

Radiofrequency GD

Radiofrequency discharges have been widely applied to the surface treatment of biomaterials and can be classified into two types according to the method that

couples the RF with the load: capacitive coupling and inductive coupling. The hardware typically consists of an RF generator, a matching network, and an antenna. Commonly, the generator operates at a frequency of 13.56 MHz and the pressure during the discharge is between 10^{-3} and 100 Torr. The plasma density, which is defined as the electron or ion density (equal in the quasi-neutral state), in a low-pressure (10^{-3} –1 Torr) radiofrequency GD varies from 10^9 to 10^{11} cm^{-3} , whereas that at a medium pressure (1–100 Torr) can reach 10^{12} cm^{-3} . The electron temperature in a low-pressure discharge is usually between 1 and 5 eV and the plasma potential of the RF discharge is typically in the range of 10 to 30 V above the wall potential and proportional to the electron temperature.

Vacuum arc

The vacuum arc technique is used to generate a metal ion plasma. It is sustained by materials originating from a cathode. In general, a cathodic arc plasma source is composed of two parts: plasma production unit and macroparticle filter. The arc discharge is characterized by a relatively high current (10 or 100 A) and a relatively low voltage (10–80 V) between the cathode and the anode. When a feedback mechanism is established in which a very small area of the cathode surface is heated by electron emission, more materials are ejected. The materials emitted from the cathodic spot are mainly ions and macroparticles. A curved magnetic filter is commonly utilized to eliminate these macroparticles from the cathodic arc plasma.

Corona and atmospheric discharge

Corona discharge appears as a luminous glow localized in space around a pointed tip in a highly nonuniform electric field. The voltage applied to the anode exceeds several kilovolts and the magnitude of the discharge current varies from 10^{-10} to 10^{-4} A. In the plasma near the tip, the density of the charge species decreases rapidly with distance from about 10^{13} to 10^9 cm^{-3} .

The most widely used atmospheric arc plasma source in biomaterials research is the plasma spray torch. A common configuration comprises a nozzle-shaped anode serving as a constrictor. To sustain a plasma torch, the discharge current and power density are very high. The plasma is formed between the cathode and the anode by an electrically conducting gas at $T^\circ > 8000$ K at pressure of 10^5 Pa. Owing to the high plasma flow velocity (nearly 1×10^3 m/sec) and high temperature in the plasma, almost all solid particles can be melted and a high velocity is acquired. Thus, atmospheric plasma torches are widely used in plasma spraying.

Plasma-Surface Interactions

The interactions between plasma particles and sample surface are critically dependent on the working pressure, plasma parameters, and mode of operations. In general, both chemical and physical interactions have a significant effect on the outcome of the plasma treatment. The plasma–surface interactions in ion implantation and thin film deposition process are briefly described here and readers are urged to peruse other references for more in-depth discussions.^[1–3]

Ion implantation

The important considerations in any description of ion–solid interactions are the deep distribution of the implanted ions, irradiation damage and sputtering, creation of new materials, and ion alloying. The implanted ion penetrates a solid and comes to rest via electron and nuclear energy loss resulting in collision cascades. Those collisions can cause sputtering, radiation-enhanced diffusion, ion backscattering, ion-induced chemical reactions, and defect generation. Sputtering sets the limit of the maximum concentration of implanted atoms that can be retained in the target. The yield of sputtered atoms, that is, the number of sputtered atoms per incident ion, typically varies from 0.5 to 20 depending on the primary ion species, ion energy, structure and composition of the target, and sample topography. As ion irradiation is quite efficient in forming vacancy–interstitial pairs, the effects of irradiation on the ordered alloy are described by the competing processes of chemical disordering, which is induced by atomic replacements resulting from displacement and cascade damage, and chemical ordering. The process is stimulated by radiation-enhanced diffusion. Besides, ion implantation offers the ability to alloy any elemental species into the near-surface region of any substrate, irrespective of thermodynamic criteria, such as solubility and diffusivity. These considerations coupled with the possibility of low-temperature processing have prompted explorations into applications, where the limitations of dimensional changes and possible delamination associated with conventional coatings are the practical concerns.

Thin film deposition

In this process, reactants diffuse from the plasma to the substrate surface giving rise to chemical reactions and thin film deposition. Deposition process parameters, such as the type of ions and their energy, flux, incident direction, and deposition temperature, play important roles in the film composition, structure, and many related properties, such as grain size, crystallographic



orientation, defects, stress concentration, film adhesion, and surface topography, all of which alter the mechanical, electrical, optical, as well as biomedical properties of the materials. Under certain ion energy bombardment, various surface features and microstructures, such as cones, pyramids, ridges, ledges, pits, and faceted planes, can form. It should also be noted that under ion bombardment, films frequently show preferred orientation. In the case of low atom mobility or low substrate temperature, films have a tilted, voided columnar structure, and ion bombardment generally causes a reduction in the film grain size. Energetic ions displacing atoms from equilibrium sites lead to a small grain as well as a defective and stressful film. The threshold ion energies and critical temperatures often delineate the limits of the competition between lattice damage and defect annealing effects. Therefore, a proper set of plasma parameter and deposition conditions is important to yield the desirable film structure and properties.

PLASMA TREATMENT OF BIOMATERIALS

As the surface of biomaterials plays an important role in their interactions with body fluids, tissues, and cells, the surface structures and properties are of paramount importance. They can be altered by plasma-based treatment processes, such as plasma sputtering and etching, implantation, deposition, coating, and spraying. Fig. 1 illustrates some of the surface properties that can be altered by PSM.

Plasma Sputtering and Etching

In plasma sputtering and etching processes, materials are removed from the surface by chemical reactions and/or physical sputtering. Commonly, inert gases, such as neon and argon, are used as the carrier gas and are responsible for physical ablation, while reactive plasma is employed for chemical etching of the surface. The plasma used in the process generally is produced by RF GD. The sputtering rate depends on the input power of the source, working pressure, and voltage applied to the substrate.

Plasma etching is a useful method to change the surface morphology of biomaterials. For instance, polymeric materials, such as polyurethane (PU), can be treated in an argon (Ar) and sulfur dioxide (SO₂) plasma to increase the surface wettability. The fibrinogen adsorption rate is reduced because of the etching effect that renders the surface rougher.^[4] This process can also be used to create the desirable microfeatures and macrofeatures on the biomaterials. These features are capable of regulating cell functions,

such as repellency, proliferation, differentiation, and apoptosis. The effects of the substratum topography on the cellular behavior, such as cell orientation, migration, and cytoskeletal arrangements, have attracted much attention. Investigations of the grooved surface reveal that cells aligned to the long axis of the grooves, often with organization of actin and other cytoskeletal elements in an orientation parallel to the grooves.^[5] As cell behaviors, such as size, shape, density, and geometry, can provide regular signals for cell survival and growth or apoptosis, the effectiveness of cell culturing and spreading rather than rounding on the specialized surface feature can modulate or switch cells from apoptosis to survival and growth.^[6] With regard to plasma-induced surface patterning and cell culturing, Schroder et al. have introduced a plasma etching process to induce micropatterning on polymeric materials surfaces as illustrated in Fig. 2.^[7] Their results show that ammonia plasma introduces amino groups to the surface to attain good cell adhesion, whereas hydrogen plasma can be used for surface passivation. The latter process can be used to remove surface functional groups and consequently, a higher degree of cell repellence is obtained. D. Braber et al. have produced microgradient textures on titanium and silicon rubber substrates for rat dermal fibroblasts (RDFs) culturing.^[8] The RDFs cellular behavior depends on the size of the grooves, and the proliferation rate of the RDFs is much higher on titanium than silicon rubber substrate. Other studies on surface activation of polyetheretherketone reveal that the surface topography can become irregular with high roughness, and inflammatory cell responses can be spatially controlled in a manner that can be ultimately exploited to improve the biocompatibility of medical devices.^[9] Hence, the plasma-based surface patterning technique enables visualization of the effects on cell functions and spatially controls the cellular microorganization.

Plasma Implantation

In plasma implantation, ions in the overlying plasma are accelerated by a high field toward the biomaterial surface for structures and properties modification. Generally, the ion penetration depth is a function of the implantation energy and also depends on the material properties, such as electrical conductivity and density. The effects of ion implantation are different for metals, ceramics, and polymers. In metal and ceramic biomaterials, the physical changes arise from atomic and nuclear collisions often leading to the formation of highly disordered and sometimes amorphous structures. Chemical changes arise from the formation of hard-phase precipitates or surface alloys. The physical and chemical changes typically combine

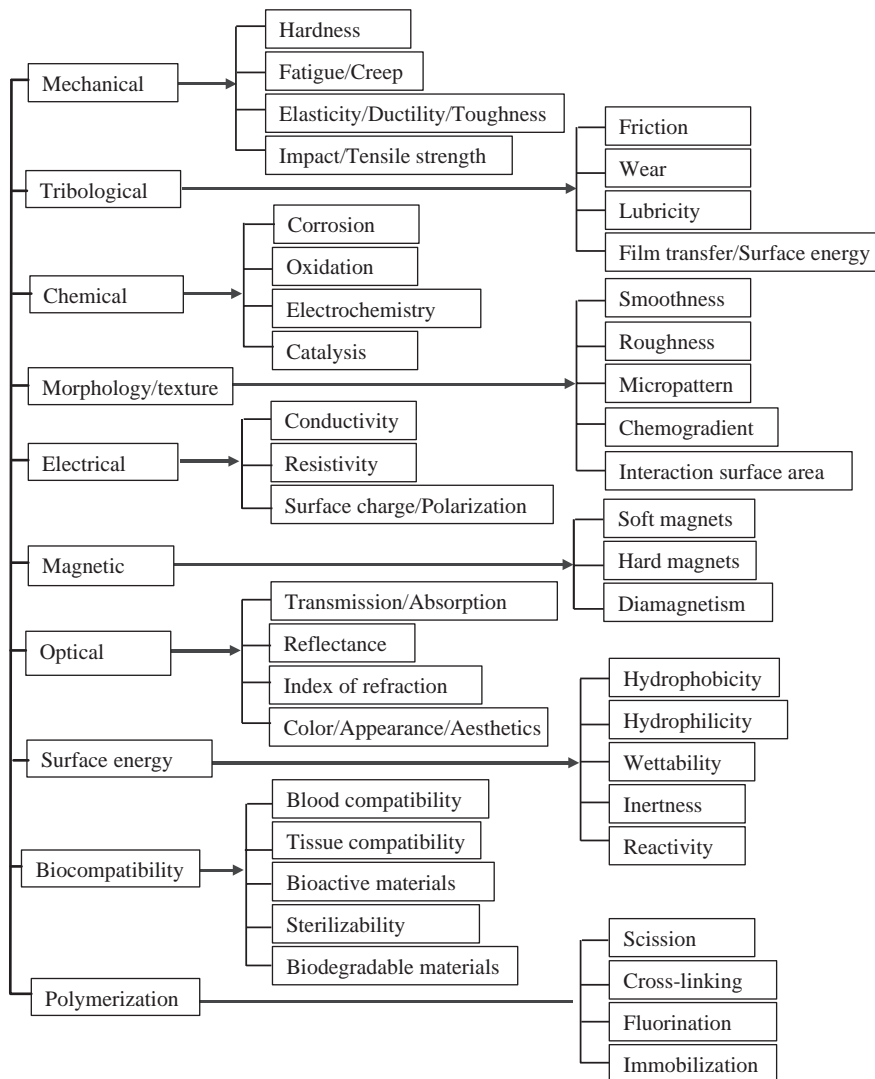


Fig. 1 Biomaterial surface properties altered by PSM.

to create surfaces that are harder and more resistant to wear and chemical attack without substantially altering the bulk properties. In polymeric biomaterials, the use of energetic ions and activated species in the ion implantation process introduces deeper ion penetration, cross-linking, and/or chain scission. It can also enhance or reduce the functional groups on the polymeric surfaces. This may thus help reduce the aging effects and degradation of the functional groups or chains. In addition, surface properties such as surface hardness, wear resistance, surface wettability, and surface tension can be improved.

Ion implantation of metallic biomaterials

Metallic biomaterials, such as stainless steels, cobalt-based and titanium-based materials, are commonly used in orthopedic devices, such as total knee, hip, shoulder, and ankle joint replacements, spinal implants, and bone fixators, dental implants to replace teeth/root

systems and bony tissues in oral cavities, as well as cardiovascular implants, such as artificial heart valves and blood vessel stents. Most orthopedic devices are designed to remove diseased areas of joints and replace them with implants designed specifically to restore stability and functions under adverse and stressed conditions. Materials chosen for the devices must meet the biocompatibility requirements of the immune system, the biological environment, corrosion issues, friction, and wear of the articulating surfaces, and implant fixation through either osseointegration or bone cement. Ion implantation using either conventional beam-line ion implantation (BLII) or plasma immersion ion implantation (PIII) can be used to modify the mechanical, physical, chemical, and biological properties of most metallic biomaterials.

Nitrogen Plasma Implantation. Metallic biomaterials after nitrogen implantation exhibit improved tribological characteristics and can better resist corrosion attacks due to changes in the surface composition and

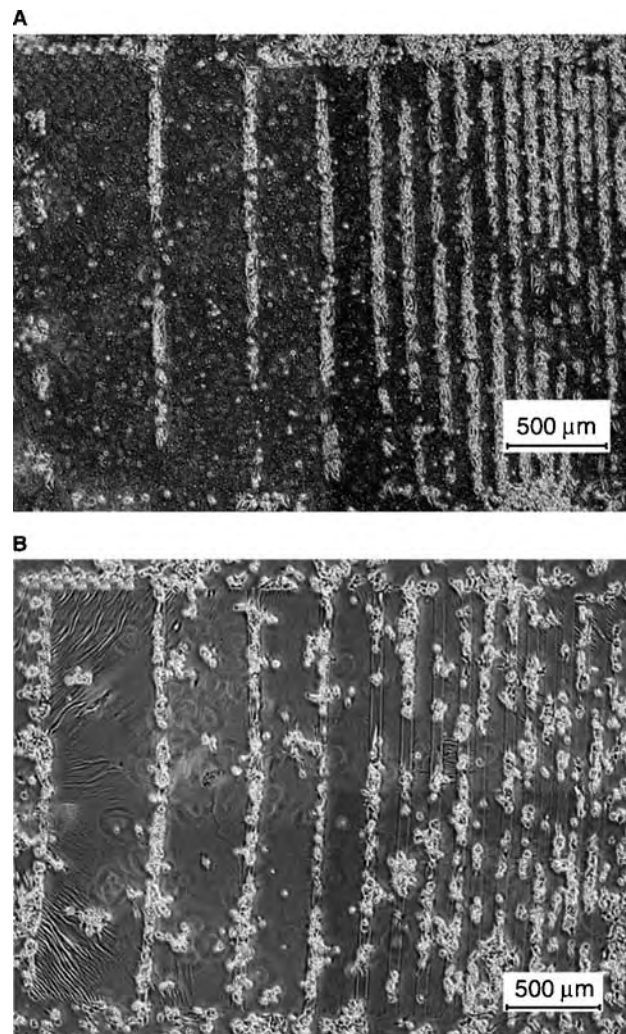


Fig. 2 Micropatterned growth of epithelial KB cells on (A) PET and (B) PHEMA. The areas of increased cell density were treated with ammonia plasma, while the areas with suppressed cell adhesion were additionally treated in hydrogen plasma.

formation of favorable microstructures. For instance, the wear enhancement of nitrogen-implanted materials stems from the formation of hard, secondary nitride-phase precipitates in the near-surface region. The high hardness of these nitride precipitates translates into significant strength and the matrix phase is further strengthened by the presence of strain fields around the precipitates that can impede dislocation movement. Improvement in the corrosion resistance arises from the high solubility of the implanted nitrogen leading to a more compositionally homogenous surface. In steels, it is generally accepted that the chromium-depleted zone adjacent to the grain boundaries is susceptible to intergranular corrosion and galvanic attack. Nitrogen implantation can significantly enhance the wear and corrosion properties of steels by the formation of hard CrN precipitates, expanded austenite transformation

into martensite during high-temperature treatment, production of a surface devoid of grain boundaries, and formation of a protective passivating film.

Nitrogen implantation has also been used to modify titanium and titanium alloys. Damdorenea et al. have implanted nitrogen into pure commercial titanium and compared the electrochemical behavior in a media of different pH.^[10] In a neutral pH saline solution, the implantation does not affect the corrosion resistance. In an acidic solution, the implanted titanium shows higher corrosion resistance and stability up to +1.5 V. The improved electrochemical behavior can be explained by the nitride phase acting as a barrier. They screen the Ti ions from the external environment, thereby inhibiting migration of the Ti ions and formation of TiO₂ at low potentials. Ueda et al. have performed nitrogen plasma implantation to improve the surface tribological properties of Ti6Al4V as artificial heart valve materials.^[11] The results show a well-behaved nitrogen profile (~40% peak atomic concentration) and nitrogen retention layer (~50 nm). After plasma implantation, the friction coefficient is reduced by about one-third compared to the untreated alloy and the hardness is increased by about 70%.

Oxygen Plasma Implantation. Oxygen implantation into titanium can induce the formation of titanium oxide and it is known to have varying stoichiometries. The common compounds are Ti₃O to Ti₂O, Ti₃O₂, TiO, Ti₂O₃, Ti₃O₅, and TiO₂. Titanium oxides are thermodynamically very stable and the Gibbs free energy of formation is highly negative in a variety of oxidation media, such as water or oxygen containing organic molecules. Poon et al. have conducted oxygen PIII on NiTi shape memory alloys used in orthopedic applications.^[12] It is observed that the oxide barrier layers can effectively resist the release of harmful Ni ions from the implants into the living tissues. They also compare the anticorrosion efficacy of oxide films produced by atmospheric pressure oxidation with that of oxygen PIII. The results show that the oxidized samples do not possess improved corrosion resistance and are even far worse than the untreated samples, while significant improvement in the corrosion resistance can be achieved by oxygen PIII, as shown in Fig. 3. In addition, in vitro biological tests reveal that the oxygen PIII surfaces are cytologically compatible, allowing the attachment and proliferation of osteoblasts. Oxygen implantation into NiTi forms titanium oxides so that the wear loss and the corrosion attack can be mitigated and osteoblasts can grow and spread on the oxides. Tan, Dodd, and Crone also employ oxygen PIII to modify the electrochemical properties of NiTi.^[13] Their results illustrate that both the corrosion and the wear-corrosion behaviors can be improved by heat treatment and an oxygen-modified surface.

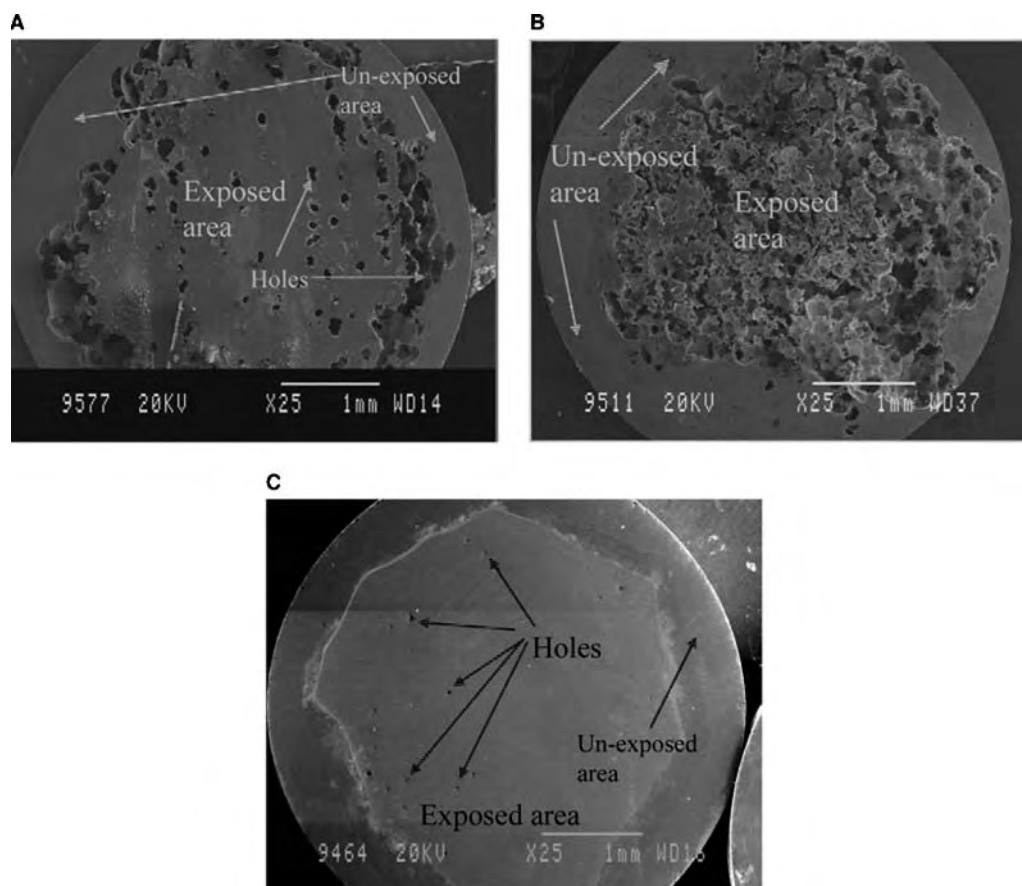


Fig. 3 Scanning electron micrograph images of NiTi surface after electrochemical tests based on the ASTM G5-94 and G61-86 protocols in a standard SBF at a pH of 7.42 and temperature of $37 \pm 0.5^\circ\text{C}$. A cyclic potential was scanned between -400 and $+1600$ mV at a rate of 600 mV/hr. (A) Untreated control, (B) atmospherically oxidized, and (C) oxygen-plasma implanted.

Metal Plasma Implantation. Biological friendly elements, such as calcium, phosphorus, and sodium, can be introduced into biomaterials to enhance the mechanical and chemical properties, surface bioactivity, and biocompatibility. Hanawa has studied calcium implantation into titanium to improve bone conductivity.^[14] A more positively charged Ca-implanted surface is achieved because of the dissociation of hydroxyl radicals and the release of Ca ions from the surface accelerates the formation of calcium phosphate, as illustrated in Fig. 4. Osteogenic cells are activated to form osteotissues and a large amount of new bone is observed to form early on calcium ion implanted titanium after implantation into rat tibia. A similar observation has been obtained from Ca PIII titanium surface.^[15] On exposure to air, the surface calcium is first oxidized and then reacts with water to form calcium hydroxide. After the Ca PIII titanium is soaked in simulated body fluids (SBFs), a bone-like apatite layer forms on the surface confirming the improvement of the bone conductivity.

Calcium ions are implanted into titanium to immobilize bisphosphonates on the surface.^[16] The

modified layer, which consists of CaTiO_3 and TiO_2 , is relatively stable so that immobilized bisphosphonate can remain on the modified surface. In the dental field, bisphosphonate prevents periodontal destruction and decreases the formation of dental calculus. Many of the bisphosphonates inhibit the crystal formation of calcium phosphate in vitro blocking the transformation of amorphous calcium phosphate into hydroxyapatite (HA) and the dissolution of these crystals. Therefore, the use of bisphosphonates is expected to promote osteogenesis at the bone tissue-implant interface by inhibiting the activity of osteoclasts, and immobilization of bisphosphonates on titanium implants is considered important.

The effects of phosphorus implantation on the corrosion resistance and biocompatibility of titanium have been investigated by Krupa et al.^[17] The process leads to amorphization of the surface layer and the formation of TiP as well as the increase of the corrosion resistance after short-term and long-term exposures. In the biological tests, osteoblasts show the ability to spread on the surfaces, but no difference is

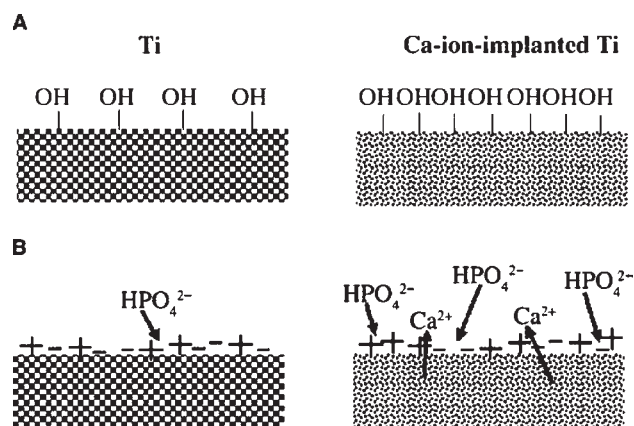


Fig. 4 Hydroxyl radicals (A) in air and (B) in electrical charges in body fluids on unimplanted titanium and calcium ion implanted titanium.

found between the modified and the control titanium. Calcium and phosphorus-coimplanted titanium can also promote the formation of HA on the surface after backfilling with oxygen and thermal treatment as the modified surface structure containing both Ca and P is close to that of apatite.

Sodium has been implanted into titanium to form sodium titanate in the near-surface region. The process produces morphologically rugged surfaces that induce apatite precipitation from SBF in immersion tests.^[18] The surface reactivity is associated with a titania hydrogel film formed by the hydrolysis of sodium titanate. The double layer on a negatively charged titania hydrogel surface tends to increase supersaturation that fosters apatite nucleation and precipitation near the surface.

Ion implantation of polymeric biomaterials

Polymers are probably the most widely used biomaterials. They are the materials of choice in cardiovascular devices, scaffolding implants for tissues, and ophthalmic applications. However, polymers at present have many deficiencies, such as low mechanical strength and poor wear resistance. Ion implantation is a useful solution to mitigate these deficiencies and satisfactory results have been achieved by ion implantation. Several common synthetic polymers are described here.

Polypropylene. Polypropylene has an exceptionally high flex life and excellent environmental stress-cracking resistance. It is commonly used in finger joint prostheses with an integrally molded hinge design. On the other hand, PP contains only carbon and hydrogen and it is quite hydrophobic, thereby limiting its potential applications. Li and Niu have performed

carboxy ion (COOH⁺) implantation at room temperature into PP and studied the in vitro effects of ion implantation on immune macrophages, 3T3 mouse fibroblasts, and human endothelial cells (ECs).^[19] COOH⁺ ion implantation is observed to break off some pendant groups of PP and induce unbound carbon atoms to combine with the implanted COOH⁺ ions and oxygen in air to form new polar O-containing groups, such as O–H, C=O, and C–O bonds. This gives rise to a more hydrophilic surface. The ensuing cell attachment tests show that fibroblasts and human ECs can be better cultured and proliferated on the implanted surface. Low macrophage attachment with normal cellular morphology has also been observed because of the formation of O-containing groups on the surface and the diminution of cell decomposition.

Polystyrene. Polystyrene (PS) has good transparency, high thermal stability, low specific gravity, and a relatively high modulus. It is often used in cell culturing and studies of mechanisms of the interactions between cells and artificial materials. Nakao et al. have studied neon ion (Ne⁺) bombardment of PS.^[20] Ne bombardment of PS destroys the benzene ring completely and induces the formation of an amorphous carbon phase with new functional groups, such as C–O and C=O. In cell adhesion measurements, the modified PS exhibits high adhesion and spreading of HeLa cells. In addition, Bacakova et al. have reported that fluorine ion (F⁺) implantation in PS can increase the polarity, hydrophilicity, electrical conductivity, and carbonization of polymeric surface.^[21] Endothelial cell culturing tests demonstrate a significant increase of cell attachment. In ELISA tests, the modified PS is found to enhance the expression of a cytoskeletal protein, vimentin, and protein of focal adhesion plaques, and talin.

Polymethylmethacrylate. Polymethylmethacrylate (PMMA) is an amorphous material with good resistance to diluted alkaline and inorganic solution. It also possesses high transparency, high refractive index, and good weathering properties. It has been used in contact lenses and implantable ocular lenses because of their excellent optical properties, dentures, and maxillofacial prostheses due to their good physical and coloring properties. Research interests have focused on extending the lifetime of PMMA implanted into the body with respect to ophthalmic intraocular lens (IOLs) applications. Li, Cui, and Gu have reported that the cleavage of pendants, surface oxidation, and formation of F-containing groups in F⁺ ion-implanted PMMA promote the reduction of platelets, neutral granulocytes, and macrophages adhesion on the surface of the IOLs.^[22] It is believed that fluorine ion bombardment increases the surface hydrophobicity, disfavoring cell

attachment. At the same time, no appreciable changes in the tensile strength and optical transmittance of the implanted PMMA have been observed.

Polytetrafluoroethylene. Polytetrafluoroethylene (PTFE) has a very low surface tension and friction coefficient. It also has the unusual property of being able to expand on a microscopic scale into a microporous material. Microporous PTFE or expanded PTFE (ePTFE) is used as subcutaneous augmentation material in various biomedical applications, such as vascular grafts and facial surgery. As ePTFE has a highly hydrophobic surface, it limits cell adhesion. Several studies have concentrated on improving the surface properties, such as wettability as well as blood and tissue compatibility. Colwell et al. have reported that N_2^+ , Ar^+ , and Ca^+ ion implantation can induce defluorination, resulting in the increase of surface roughness and the change in hydrophobicity depending on the implantation dose.^[23] Suzuki et al. have studied Ne^+ ion implantation into ePTFE for dura mater applications. It is found that ion implantation destroys the surface chemical structure of ePTFE, thereby producing C=O bonds.^[24] The in vivo study using rabbits shows that the internodal distance, density of the nodes in ePTFE, and C=O radicals induced by ion implantation are the major factors influencing cell invasion on the ion-implanted ePTFE.

Plasma Deposition and Coating

Plasma deposition is an important area in biomaterials engineering. In this process, a layer with properties distinctly different from those of the bulk materials can be synthesized. The common techniques include PIII-deposition, plasma-enhanced chemical vapor deposition, plasma grafting, and polymerization.

Carbon-based thin films for biomedical applications

Diamond-like carbon (DLC) films possess novel properties, such as chemical inertness, electrical resistance, optical transparency, high hardness, and low friction coefficient. Some representative uses of hard DLC coatings are: 1) hard carbon films on surgical instruments; 2) implanted fittings, such as infusion devices and connecting parts; and 3) semipermanent components, such as heart valves, stents, replacement joints, and ophthalmic devices. Hard, thin, and hermetic DLC films can act as nontoxic corrosion-resistant barriers against attack by microorganisms. They also resist biodeterioration. By suitably tailoring the surfaces, protein adhesion and cell growth can be promoted or retarded as required by the applications.

Diamond-like carbon consists of carbon and hydrogen and is a metastable form of amorphous carbon containing a hybrid structure of sp^2 and sp^3 bonds. The disordered structures and properties of DLC depend very much on the deposition techniques and fabrication parameters. Fourier transfer infrared (FTIR) and laser Raman spectroscopies are the common techniques used to characterize the structure and bonding of a-C:H. Hydrogenated DLC displays a wide FTIR absorption band. The C–H stretching modes are composed of three regions, the $\text{sp}^1\text{=C-H}$ modes centered at 3300 cm^{-1} , the $\text{sp}^2\text{=CH}_n$ modes from 2975 to 3085 cm^{-1} , and $\text{sp}^3\text{-C-H}_n$ modes from 2850 to 2955 cm^{-1} . Diamond-like carbon films with a hybrid sp^2/sp^3 structure show a D band (disorder) and a G band (graphite) signal at around 1350 and 1580 cm^{-1} , respectively, in the Raman spectrum. A common way to monitor the changes in the structures or chemical states of the DLC is to compare the shift of the G band and the ratio of the intensity of D band (I_D) to that of G band (I_G).

Diamond-like carbon coatings typically have good hemocompatibility, which is typically expressed by decreased thrombus formation. When exposed to blood, an increased ratio of albumin to fibrinogen adsorption, as well as decreased platelet activation, is observed on coated surfaces.^[25] The degree of blood compatibility is related to the electronic conduction, work function, and surface energy, which are dependent on the ratio of sp^3 to sp^2 in the DLC film. There have been attempts to modify the DLC structures by doping with biocompatible elements, such as Ti, Si, P, etc. Diamond-like carbon films are usually not cytotoxic and can be applied to chronic neuroprosthetic implants. The antibacterial adsorption behavior of DLC-coated polymers presented by Wang et al. shows that reduction in bacterial adhesion is due to energetically unfavorable adhesion on the DLC surface, as the free energy of adhesion (ΔF_{Adh}) predicts whether microbial adhesion is energetically favorable ($\Delta F_{\text{Adh}} < 0$) or not ($\Delta F_{\text{Adh}} > 0$).^[26] On the other hand, in some particular applications in which no cell or bacterial attachment to the DLC surface is desirable, certain cytotoxic elements, such as Ag, Cu, or V, can be embedded into the DLC films. This allows the DLC surface to have a tunable antibacterial effect. However, care should be exercised because in a biological media, the doping elements can be released to cause toxic reactions inhibiting cell proliferation and differentiation.

Titanium-based thin films in biomedical applications

Titanium-based films are normally used in orthopedic and dental prostheses as well as blood-contacting biomaterials. Rutile titanium oxide films generally

have blood compatibility better than that of clinical biomaterials, such as low-temperature isotropic carbon (LTIC). Depending on the deposition parameters, a wide range of phases, from amorphous to crystalline rutile, can be synthesized by plasma-deposition techniques. Although the crystal structure and crystallite size of Ti–O films appear to influence the activation of the plasmatic clotting system only to a minor degree, Ti–O films exhibit superior thromboresistant properties in long-term tests. In vitro blood compatibility investigation indicates that the TiO₂ film has a longer clotting time, lower hemolytic rate, less adherent platelets, as well as less aggregation and pseudopodium of the adherent platelet.^[27] In vivo implantation into dogs shows no thrombosis on the Ti–O films after 17–90 days, in contrast to that on the control of LTIC, as shown in Fig. 5. It is suggested that the significantly lower interfacial tension between the Ti–O films and blood and plasma proteins, and the semiconducting nature of the Ti–O films give rise to the improved hemocompatibility. Other studies also show that Ti–O films have the desired properties to promote growth of ECs on the surface without protein precoating.^[28] It is believed that seeding with ECs on vascular prostheses can overcome the generation of thrombi and emboli because of the nonthrombogenic properties of the cells in the natural state in contact with blood. Moreover, doping of Ti–O films with the appropriate elements can not only change the microstructures, but also enhance their mechanical properties and biological compatibility. For instance, tantalum-doped Ti–O films exhibit changes in the semiconductivity and physical properties that preclude fibrinogen from denaturation and, consequently, prevent blood coagulation.^[29] Phosphorus-doped Ti–O films also demonstrate improved clotting behavior.^[30] It is believed that P and its oxides in Ti–O can act as the biochemical functional groups that hinder electron exchange between proteins and the surface of materials.

Plasma grafting and polymerization

When polymeric materials are exposed to plasma, radicals are created in the surface chains of the polymer because of energetic irradiation. These surface radicals can initiate polymerization reactions when put in contact with monomers in the liquid or in the gas phase. As a result, a variety of desirable properties can be obtained via graft polymerization. A specific demand for a polymeric biomaterial surface is its compatibility with organs and cells. This leads to the approach of cell seeding before implantation as the natural state of seeded cells can inhibit the generation of thrombi and emboli on the surface of implants in contact with blood. Thus, a suitable interface between

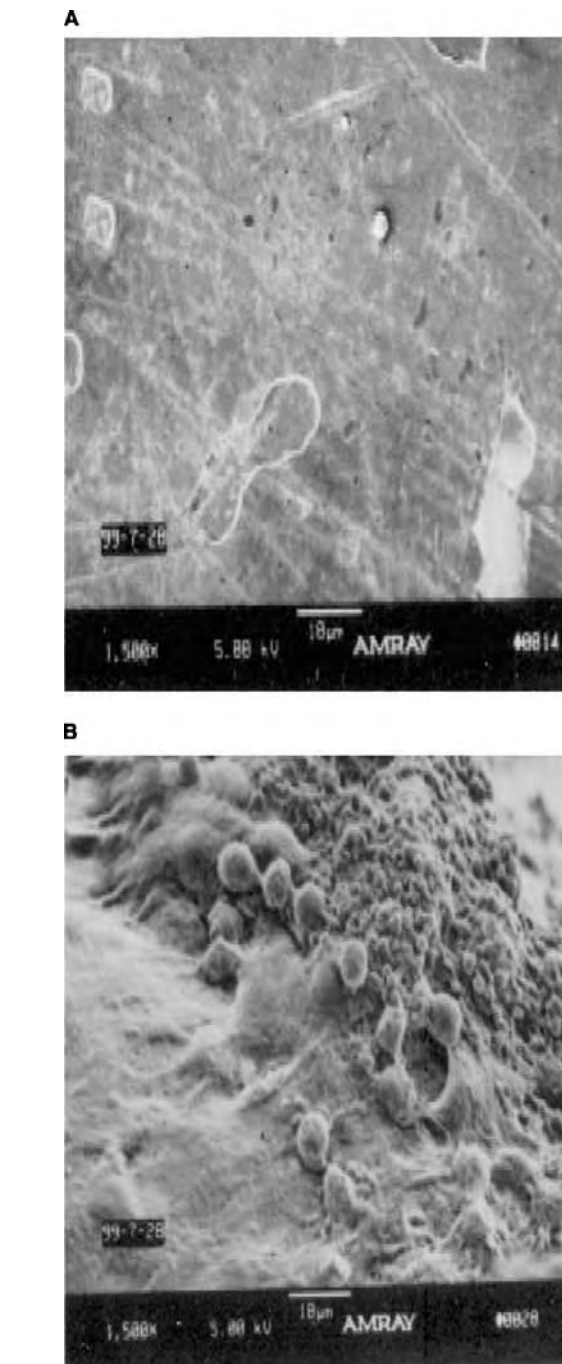


Fig. 5 Scanning electron micrographs of materials after 90 days of implantation into dogs. (A) Ti–O film and (B) LTIC.

the synthetic implants and the individual cells is required. Plasma graft polymerization is a useful technique to promote cell attachment. The cell seeding procedures are: 1) functionalization of the polymer surface with radicals by plasma irradiation, such as Ne, Ar, N₂, NH₃, or O₂; 2) immobilization of the suitable spacer molecule; 3) covalent coupling of the cell intergrin-mediator, such as extracellular matrix (ECM)

proteins of fibronectin, laminin, vitronectin, and von Willebrand factor; and (d) seeding with ECs. Fig. 6 illustrates the schematic of an interface for cell seeding.

Cell adhesion to polymeric materials is mediated by cell-surface receptors interacting with cell adhesion proteins bound to the material surface. To promote receptor-mediated cell adhesion, the polymer surface should mimic the extracellular matrix. Poly(vinylidene-fluoride) (PVDF) is a biomaterial designed for soft tissue applications, but it does not possess functional groups. Surface modification of PVDF can be performed by plasma-induced graft polymerization of hydrogels like polyacrylic acid spacer and immobilization of fibronectin.^[31] Klee et al. conclude that incorporation of fibronectin onto PVDF surface, either adsorptively or covalently, can induce cell attachment to PVDF. However, prolonged proliferation and survival of cells are seen only when the protein molecule is covalently bound to the surface. Therefore, long-term presence of the ECM protein on the polymer surface seems to be a prerequisite for successful application as a biomaterial.

Graft polymerization of functional monomers using plasma GD is very effective in increasing the immobilizing site on the segmented PU surface for heparin (Hep) immobilization. The blood compatibility of poly(ethylene oxide) (PEO)-grafted and Hep-immobilized PU has been investigated using in vitro blood interaction.^[32] Poly(ethylene oxide) is frequently used as a surface modifier or spacer, and Hep, which is well known as an anticoagulant, is often used as an immobilizing biomolecule. Kang et al. have observed that the activation of plasma proteins and platelets on PU is slightly suppressed by PEO grafting and significantly suppressed by Hep immobilization. A longer PEO biomolecule has better blood compatibility than a shorter PEO, whether or not it is directly grafted on the PU surface or used as a spacer between the PU surface and the Hep. Moreover, the release of cytokine interleukin-6 (IL-6) by the stimulation of peripheral blood mononuclear cell with Hep-immobilized PU can be highly suppressed.

Plasma Spraying

Plasma spraying includes atmospheric plasma spraying and vacuum plasma spraying. In plasma spraying, the powders of the sprayed materials are introduced into the plasma area of the plasma torch. Because of the high temperature and flux velocity of the plasma, the melted or the partially melted powders are accelerated toward the substrate at high speed forming a coating with a lamellae structure. Owing to the high temperature of the plasma core, plasma spraying offers the ability to deposit almost any metal and various combinations of materials. The high temperature enables the use of the coating materials with a high melting point, such as ceramics, cermets, and refractory materials.

Plasma spraying is widely used to produce bioactive ceramics, such as HA, $\text{Ca}_{10}(\text{PO}_4)_6(\text{OH})_2$, as interfacing osseoconductive layers on metallic surgical implants. The HA is one of the most attractive materials for human hard tissue implants because of its similar chemical composition (Ca/P ratio) to that of teeth and bones. An HA coating can accelerate the process of bone growth in the vicinity of the prosthesis. Therefore, HA has been clinically applied as a dense, sintered material and as a coating on bioinert metallic implants. However, the relatively poor bonding due to high residual stress between the plasma-sprayed HA coating and prostheses is one of the main disadvantages making it unsuitable for high-load applications. To achieve the mechanical characteristics needed for biomedical applications, blending with a tough phase is essential. Titanium metal, Ti-6Al-4V alloy, yttria-stabilized zirconia, Ni_3Al , and Al_2O_3 ceramics have been considered good candidates as the reinforcing phases. Another important concern is the resorption and degradation of HA coatings in a biological environment. This may lead to disintegration of the coating resulting in the loss of both the coating-substrate bonding strength and implant fixation. Generally, the amorphous and the metastable compounds are more soluble than the crystalline HA. Some post-deposition treatment processes have been proposed to

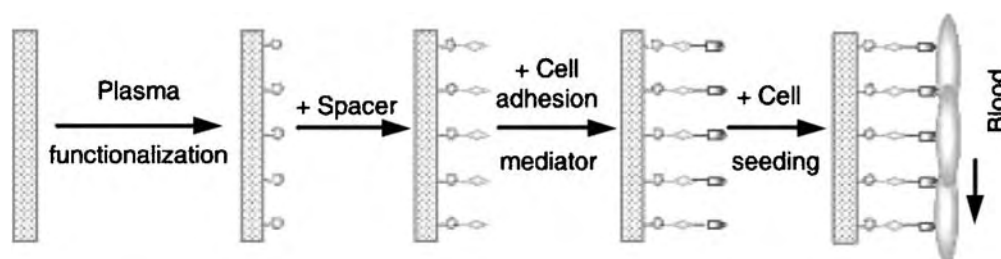


Fig. 6 Schematic illustration of plasma graft polymerization of an interface for cell seeding.

increase the crystallinity, including furnace annealing, laser irradiation, and vapor-flame treatment.^[33]

Other common plasma-sprayed coatings used on titanium alloys are Al₂O₃, ZrO₂, and TiO₂. These coatings are being used clinically because of their excellent wear resistance and corrosion resistance. However, Al₂O₃ and ZrO₂ coatings cannot bond directly to human bone tissues because of their bioinertness, thereby limiting their applications. Attempts have been made to plasma spray bioactive ceramics onto titanium implants.^[34] Some bioglass and glass-ceramic coatings, including CaO-SiO₂, have been found to bind to living bones and some ceramics including wollastonite (CaSiO₃) and dicalcium silicate (Ca₂SiO₄) have also been reported to have excellent bioactivity.^[35]

CONCLUSIONS

Plasma surface modification of biomaterials is an effective means to modify the surface of existing materials to enhance the physicochemical properties and optimize the biocompatibility and biofunctionality. There are many types of biomaterials and the design of each type of biomaterials or devices has its own challenge based on the intended function and biological uses. Much work has been conducted and more information pertaining to surface engineering of biomaterials can be found in Refs. [36–40]. More efforts are needed to further improve the functions and lifetime of the biomaterials. As the biomaterials field is interdisciplinary in nature, there must be a close collaboration among people in materials science, plasma physics, chemistry, biology, biomedical engineering, and medicine.

ARTICLES OF FURTHER INTEREST

Bioadhesion; Biomaterials: Protein-Surface Interactions; Surface Coating; Surface Modification; Surface Topography and Cell Behavior

REFERENCES

1. Grill, A. *Cold Plasma in Materials Fabrication: From Fundamentals to Applications*; IEEE Press: New York, 1994.
2. Lieberman, M.A.; Lichtenberg, A.J. *Principles of Plasma Discharges and Materials Processing*; John Wiley & Sons: New York, 1994.
3. Ohring, M. *The Materials Science of Thin Films*; Academic Press: Boston, 1992.
4. Klee, D.; Hocker, H. Polymers for biomedical applications: improvement of the interface compatibility. *Adv. Polym. Sci.* **2000**, *149*, 1–57.
5. Flemming, R.G.; Murphy, C.J.; Abrams, G.A.; Goodman, S.L.; Nealey, P.F. Effects of synthetic micro- and nano-structured surfaces on cell behavior. *Biomaterials* **1999**, *20* (6), 573–588.
6. Chen, C.S.; Mrksich, M.; Huang, S.; Whitesides, G.M.; Ingber, D.E. Micropatterned surfaces for control of cell shape, position, and function. *Biotechnol. Prog.* **1998**, *14* (3), 356–363.
7. Schroder, K.; Meyer-Plath, A.; Keller, D.; Ohl, A. On the applicability of plasma assisted chemical micropatterning to different polymeric biomaterials. *Plasma Polym.* **2002**, *7* (2), 103–125.
8. Den Braber, E.T.; Jansen, H.V.; De Boer, M.J.; Croes, H.J.E.; Elwenspoek, M.; Ginsel, L.A.; Jansen, J.A. Scanning electron microscopic, transmission electron microscopic, and confocal laser scanning microscopic observation of fibroblasts cultured on microgrooved surfaces of bulk titanium substrate. *J. Biomed. Mater. Res.* **1998**, *40* (3), 425–433.
9. Ha, S.W.; Hauert, R.; Ernst, K.H.; Wintermantel, E. Surface analysis of chemically-etched and plasma-treated polyetheretherketone (PEEK) for biomedical applications. *Surf. Coat. Technol.* **1997**, *96* (2–3), 293–299.
10. Damdorenea, J.D.; Conde, A.; Palacio, C.; Rodriguez, R. Modification of corrosion properties of titanium by N-implantation. *Surf. Coat. Technol.* **1997**, *91* (1–2), 1–6.
11. Ueda, M.; Silva, M.M.; Otani, C.; Reuther, H.; Yatsuzuka, M.; Lepienski, C.M.; Berni, L.A. Improvement of tribological properties of Ti6Al4V by nitrogen plasma immersion ion implantation. *Surf. Coat. Technol.* **2003**, *169*, 408–410.
12. Poon, R.W.Y.; Ho, J.P.Y.; Liu, X.Y.; Chung, C.Y.; Chu, P.K.; Yeung, K.W.K.; Lu, W.W.; Cheung, K.M.C. Anti-corrosion performance of oxidized and oxygen plasma-implanted NiTi alloys. *Mater. Sci. Eng. A* **2005**, *390* (1–2), 444–451.
13. Tan, L.; Dodd, R.A.; Crone, W.C. Corrosion and wear-corrosion behavior of NiTi modified by plasma source ion implantation. *Biomaterials* **2003**, *24* (22), 3931–3939.
14. Hanawa, T. In vivo metallic biomaterials and surface modification. *Mater. Sci. Eng. A* **1999**, *267* (2), 260–266.
15. Liu, X.; Poon, R.W.Y.; Kwok, S.C.H.; Chu, P.K.; Ding, C. Structure and properties of Ca-plasma-implanted titanium. *Surf. Coat. Technol.* **2005**, *191* (1), 43–48.
16. Yoshinari, M.; Oda, Y.; Ueki, H.; Yokose, S. Immobilization of bisphosphonates on surface modified titanium. *Biomaterials* **2001**, *22* (7), 709–715.
17. Krupa, D.; Baszkiewicz, J.; Kozubowski, J.A.; Barcz, A.; Sobczak, J.W.; Bilinski, A.; Lewandowska-Szumiel, M.; Rajchel, B. Effect of phosphorus-ion implantation on the corrosion resistance and biocompatibility of titanium. *Biomaterials* **2002**, *23* (16), 3329–3340.

18. Pham, M.T.; Maitz, M.F.; Matz, W.; Reuther, H.; Richter, E.; Steiner, G. Promoted hydroxyapatite nucleation on titanium ion-implanted with sodium. *Thin Solid Films* **2000**, *379* (1–2), 50–56.
19. Li, D.J.; Niu, L.F. Cell attachment of polypropylene surface-modified by COOH⁺ ion implantation. *Nucl. Instrum. Methods B* **2002**, *192* (4), 393–401.
20. Nakao, A.; Kaibara, M.; Iwaki, M.; Suzuki, Y.; Suzuki, Y.; Kusakabe, M. XPS and SERS studies of cell adhesion-controlled polymer modified by ion implantation. *Surf. Interface Anal.* **1996**, *24*, 252–256.
21. Bacakova, L.; Mares, V.; Lisa, V.; Svorcik, V. Molecular mechanisms of improved adhesion and growth of an endothelial cell line cultured on polystyrene implanted with fluorine ions. *Biomaterials* **2000**, *21*, 1173–1179.
22. Li, D.J.; Cui, F.Z.; Gu, H.Q. F⁺ ion implantation induced cell attachment on intraocular lens. *Biomaterials* **1999**, *20*, 1889–1896.
23. Colwell, J.M.; Wentrup-Byrne, E.; Bell, J.M.; Wielunski, L.S. A study of the chemical and physical effects of ion implantation of micro-porous and nonporous PTFE. *Surf. Coat. Technol.* **2003**, *168* (2–3), 216–222.
24. Suzuki, Y.; Iwaki, M.; Tani, S.; Oohashi, G.; Kamio, M. Ion implantation into ePTFE for application of a dural substitute. *Nucl. Instrum. Methods B* **2003**, *206*, 538–542.
25. Yang, P.; Kwok, S.C.H.; Fu, R.K.Y.; Leng, Y.X.; Wang, J.; Wan, G.J.; Huang, N.; Leng, Y.; Chu, P.K. Structure and properties of annealed amorphous hydrogenated carbon (a-C:H) films for biomedical applications. *Surf. Coat. Technol.* **2004**, *177–178*, 747–751.
26. Wang, J.; Huang, N.; Pan, C.J.; Kwok, S.C.H.; Yang, P.; Leng, Y.X.; Chen, J.Y.; Sun, H.; Wan, G.J.; Liu, Z.Y.; Chu, P.K. Bacterial repellence from polyethylene terephthalate surface modified by acetylene plasma immersion ion implantation–deposition. *Surf. Coat. Technol.* **2004**, *186* (1–2), 299–304.
27. Huang, N.; Yang, P.; Leng, Y.X.; Chen, J.Y.; Sun, H.; Wang, J.; Wang, G.J.; Ding, P.D.; Xi, T.F.; Leng, Y. Hemocompatibility of titanium oxide films. *Biomaterials* **2003**, *24* (13), 2177–2187.
28. Chen, J.Y.; Wan, G.J.; Leng, Y.X.; Yang, P.; Sun, H.; Wang, J.; Huang, N. Behavior of cultured human umbilical vein endothelial cells on titanium oxide films fabricated by plasma immersion ion implantation and deposition. *Surf. Coat. Technol.* **2004**, *186* (1–2), 270–276.
29. Chen, J.Y.; Leng, Y.X.; Tian, X.B.; Wang, L.P.; Huang, N.; Yang, P.; Chu, P.K. Antithrombogenic investigation of surface energy and optical bandgap and hemocompatibility mechanism of Ti(Ta⁺)₂O₂ thin films. *Biomaterials* **2002**, *23* (12), 2545–2552.
30. Tsyganov, I.; Maitz, M.F.; Wieser, E. Blood compatibility of titanium-based coatings prepared by metal plasma immersion ion implantation and deposition. *Appl. Surf. Sci.* **2004**, *235* (1–2), 156–163.
31. Klee, D.; Ademovic, Z.; Bosserhoff, A.; Hoecker, H.; Maziolis, G.; Erli, H.J. Surface modification of poly(vinylidene fluoride) to improve the osteoblast adhesion. *Biomaterials* **2003**, *24* (21), 3663–3670.
32. Kang, I.K.; Seo, E.J.; Huh, M.W.; Kim, K.H. Interaction of blood components with heparin-immobilized polyurethanes prepared by plasma glow discharge. *J. Biomater. Sci. Polym. Ed.* **2001**, *12* (10), 1091–1108.
33. Dong, Z.L.; Khor, K.A.; Quek, C.H.; White, T.J.; Cheang, P. TEM and STEM analysis on heat-treated and in vitro plasma-sprayed hydroxyapatite/Ti-6Al-4V composite coatings. *Biomaterials* **2003**, *24* (1), 97–105.
34. Liu, X.Y.; Zhao, X.B.; Fu, R.K.Y.; Ho, J.P.Y.; Ding, C.X.; Chu, P.K. Plasma-treated nanostructured TiO₂ surface supporting biomimetic growth of apatite. *Biomaterials* **2005**, *26* (31), 6143–6150.
35. Liu, X.; Tao, S.; Ding, C. Bioactivity of plasma sprayed dicalcium silicate coatings. *Biomaterials* **2002**, *23* (3), 963–968.
36. Sodhi, R.N.S. Application of surface analytical and modification techniques to biomaterials research. *J. Electron. Spectrosc. Phenom.* **1996**, *81* (3), 269–284.
37. Liu, X.; Chu, P.K.; Ding, C. Surface modification of titanium, titanium alloys, and related materials for biomedical application. *Mater. Sci. Eng. R.* **2004**, *47*, 49–121.
38. Chu, P.K.; Chen, J.Y.; Wang, L.P.; Huang, N. Plasma-surface modification of biomaterials. *Mater. Sci. Eng. R.* **2002**, *36*, 143–206.
39. Ratner, B.D.; Hoffman, A.S.; Schoen, F.J.; Lemons, J.E. *Biomaterials Science: An Introduction to Materials in Medicine*; Academic Press: New York, 1996.
40. Park, J.B.; Bronzino, J.D. *Biomaterials: Principles and Applications*; CRC Press: New York, 2003.

Polyamides (Synthetic and Natural)

Naveen C. Srivastav

Department of Laboratory Medicine and Pathology, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, Alberta, Canada

Babita Agrawal

Department of Surgery, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, Alberta, Canada

Rakesh Kumar

Department of Laboratory Medicine and Pathology, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, Alberta, Canada

INTRODUCTION

Polyamides are polymers consisting of R-CO-NH-R' monomeric units. Polyamides such as proteins, wool, and silk exist naturally or they can be made synthetically. Synthetic polyamides include nylon, Kevlar, Versamid, sodium polyaspartate, aromatic polyamides, pyrrole-imidazole polyamides, peptide nucleic acid (PNA), and peptidomimetics. A small group of polyamides are categorized in a separate group as polyamino acids to distinguish them from the proteins because of their different features of biosynthesis. Synthetic polyamides can be made by the condensation reaction of monomers and the polymer chain grows through the formation of amide bonds. The amide bond is formed by the reaction of a dicarboxylic acid or an acid chloride with a diamine. Synthetic polyamides can also be made from an amino acid or its lactam, which may be able to undergo a self-condensation reaction. Natural and synthetic polyamides are useful in modern industries and biomedical applications.

CLASSIFICATION OF POLYAMIDES

Both natural and synthetic polyamides are divided into the homopolyamides, consisting of one type of monomers, and the copolyamides, which are composed of more than one constituent. Chemically, polyamides can be divided into four categories: 1) polyamides made by the polymerization of lactams; 2) polyamides made by the condensation of a diamine with a dibasic acid; 3) polyamides made by self-condensation of amino acids; and 4) polyamides based on vegetable oil.

NATURAL POLYAMIDES

Proteins

Proteins are complex organic molecules built from 20 different amino acids, in which the amino acids are

joined through a peptide bond (an amide bond). Once linked in the protein chain, an individual amino acid is called a residue and the linked series of carbon, nitrogen, and oxygen atoms are known as the main chain or protein backbone. A typical protein consists of around 200–300 amino acids in a single chain. Titin is the largest protein in the cardiac and skeletal muscles. Titin contains 26,926 amino acid residues.^[1] The synthesis of proteins from amino acids occurs inside the cells using mRNA as a template and the process is called translation. The sequence of amino acids in a protein is defined by the genes encoded in the genetic code DNA. The structure of the protein can be determined using amino acid sequencing, X-ray crystallography, and nuclear magnetic resonance spectroscopy. Many proteins are enzymes and play a crucial role in cell metabolism. The proteins are also responsible for many physiochemical processes in living organisms, e.g., cell signaling, immune responses, cell adhesion, and the cell cycle. Proteins also play a major role in the structural support of the body as components of skin, hair, nails, tendons, muscles, and bones.^[2]

Many traditional proteins prepared by extraction of animal (e.g., collagen) and plant (e.g., soy or zein from corn) tissue are being chemically or physically modified for applications in the food and biotechnology industries. Fibrous proteins such as collagen and elastin have been used in biomaterials and tissue engineering.^[3] Using chemoselective ligation technology, chemoselective anchors can be incorporated into the protein backbone to change the pharmacokinetic properties, thus leading to the formation of new lead proteins.^[4]

In addition to the proteins, there is a small group of compounds in nature known as polyamino acids. Polyamino acids include some interesting compounds, such as polyglutamic acid (PGA), polylysine, and cyanophycin (Fig. 1). Polyglutamic acid was first detected in pathogenic Gram-positive bacterium *Bacillus anthracis* more than 60 years ago.^[5] Later it was

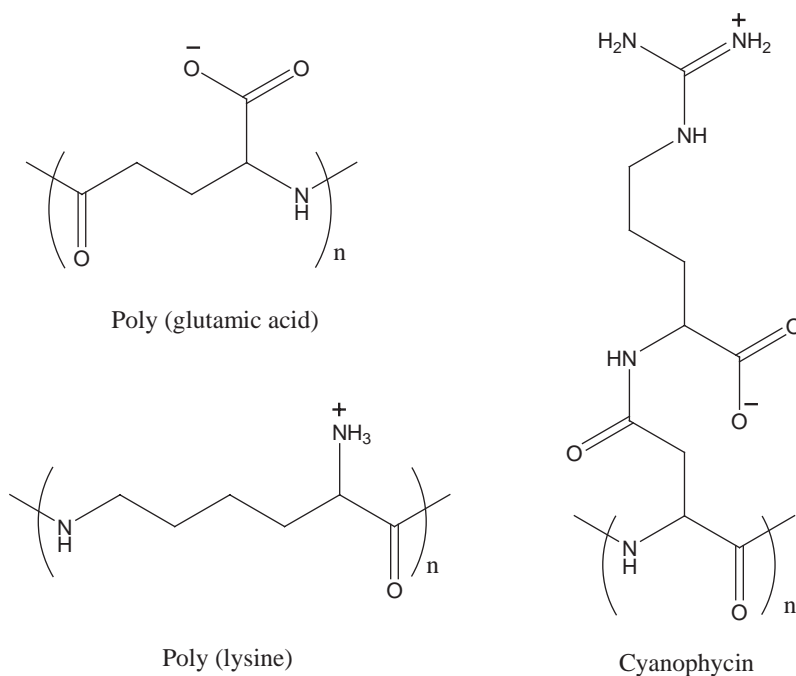


Fig. 1 Structures of polyglutamic acid, polylysine, and cyanophycin.

reported to be present in non-pathogenic Gram-positive strains of the genus *Bacillus*.^[6] In addition to the *Bacillus* spp., two halophilic eubacteria,^[7] a halophilic archaeobacterium,^[8] and few coelentrates (e.g., Hydra)^[9] have been found to produce PGA. PGA can act as a biodegradable and non-toxic substitute for the polyacrylate.^[10] In addition, it can be used as a biodegradable thickener, humectant, sustained release material, or drug carrier in food, cosmetics, or medicine. Among other applications, efforts have been made to convert the ammonium nitrogen present in liquid manure to a transient depot form of nitrogen using PGA-producing bacteria.^[11] Polylysine was isolated from *Streptomyces albulus* spp. around 40 years ago.^[12] Polylysine inhibited the growth of both Gram-positive and Gram-negative bacteria at lower concentrations.^[13] Furthermore, inactivation of bacteriophages was observed by this compound. Because of its antibiotic properties, polylysine can be used as a preservative in animal feed. Cross-linked polylysine is used as cationic adsorbent in pharmaceutical applications.^[14]

Cyanophycin, which is also referred to as cyanophycin granule polypeptide, a copolymer of aspartic acid and arginine, is a storage compound for carbon, nitrogen, and energy and exists as insoluble cytoplasmic inclusions in many cyanobacteria. In non-heterocyst-forming cyanobacteria, the cyanophycin granules are distributed in the protoplasts,^[15,16] whereas in heterocyst-forming cyanobacteria, cyanophycin granules are frequently present in the heterocysts, which are specialized for the fixation of molecular nitrogen. In those cells the cyanophycin

granules are primarily associated with the polar nodule and the neck, which connects the heterocyst to the adjoining vegetative cells.^[17] The cyanophycin granules were first isolated from *Anabaena cylindrica* by differential centrifugation.^[18] Purified cyanophycin can be chemically converted to a derivative with reduced arginine content, which then can be used in place of PGA in various technical processes, such as a biodegradable substitute for polyacrylate.^[19] Recently, cyanophycin synthesis has been established in several recombinant *Escherichia coli* strains harboring *cphA* from different sources, resulting in a maximum content of 26% (w/w) cell dry mass after growth on rich complex medium.^[20–22] Transgenic tobacco and potato plants accumulate high levels of cyanophycin. This could be a natural source of biodegradable, non-toxic plastic such as the polymer polyaspartate.^[23]

Silk

Silk is a natural protein fiber obtained from the larvae of the silk worm, *Bombyx mori*. Many other insects produce different types of silk, although none of them has been exploited for commercial purposes. Silk production is especially common in the Hymenoptera and is sometimes used in nest construction. Other types of arthropods produce silk, most notably various arachnids such as spiders (spider silk). Natural silk has long been used in making clothes, but it is also used for making parachutes, bicycle tires, comforter filling, and artillery gunpowder bags. Interestingly, early bullet-proof vests were also made from silk. Chinese medical

doctors have used natural silk to make prosthetic arteries. Recently, some bioengineers developed a genetically engineered chimeric material using a combination of nature's most remarkable materials, spider silk and silica. The resulting fusion protein has industrial and medical applications, such as growing bone tissues.^[24] Because of the strong and tough nature of spider silk, it has potential applications in the medical field for surgical sutures, artificial ligaments, and tendons. To test the possibility of using spider silk for surgical sutures, it was implanted subcutaneously in pigs. Within 2 weeks all the wound sites healed well, and few demonstrated any inflammation.^[25] Silk could also be used for soft body armor, oxygen-permeable contact lenses, biomaterial membranes, and even tethers for planes on aircraft carriers. Stem cell-based tissue engineering has been used to make biomaterials of silk fibroin scaffold. These biomaterials can be used as skeletal tissues like bone, ligament, and cartilage, as well as connective tissues like skin.^[26] Nexia Biotechnologies in Montreal successfully produced spider silk protein in transgenic goats. Milk produced by these goats contained the required protein.^[27]

Wool

Wool is a natural fibrous protein derived from the hair of sheep. It is also produced by other mammals, such as goats, rabbits, alpacas, and llamas. Natural wool is useful in making sweaters, gloves, coats, blankets, seat covers, boots, carpets, etc. It has been suggested that keratin extracts from hair and wool can be used as biomaterials. Further, it may be useful in biomedical applications for wound healing and bone regeneration, scaffolds for tissue engineering, and coating for medical devices.^[28]

Cyclotides

Cyclotides are unusual polypeptides found in a variety of tropical plants from the Rubiaceae and Violaceae families. They are head-to-tail cyclized through amide bonds. These cyclic polypeptides contain around 30 amino acid residues and possess strong chemical and biological stability toward enzymatic degradation because of the presence of a cyclic backbone. Cyclotides contain six conserved cysteine residues that form three disulfide bonds, which provide extra stability. These structural features make the cyclotides interesting for protein engineering. Use of stable cyclotides as templates for pharmacologically active sequences is being explored. In native medicine, cyclotides demonstrated a variety of biological properties, including antiviral (HIV), antimicrobial, and insecticidal activities, as well as neurotensin inhibition. An extract of the

plant *Oldenlandia affinis* has been used to accelerate childbirth, indicating that the cyclotide present in the plant affects contraction of the uterus. This is how the first cyclotide, Kalata B1, was discovered.^[29–31] Recently, researchers demonstrated that cyclotides possess cytotoxic activities.^[32,33]

SYNTHETIC POLYAMIDES

Nylons

The first chemical synthesis of a polyamide, namely nylon 6,6, was achieved by Carothers and coworkers at Du Pont in 1935. The polymer was produced commercially in 1939 and originally used for fibers. A year later, in 1940, Schlack discovered nylon 6 in Germany. Later, nylon 7 and nylon 9 were discovered as fiber-forming materials by Russian chemists. The production of nylon 11 and nylon 12 was reported in France in 1955 and 1960, respectively. The term nylon, originally a trade name, later became the generic name given to linear polyamides. Nylons can be made from single monomers such as amino acids or their cyclic lactams by self-polycondensation or by ring-opening polymerization, respectively. They can also be made from the polycondensation reaction of dicarboxylic acids and diamines. Nylons are designated by a numbering system. In the case of single monomer nylons, the digits that follow the nylon name represent the number of carbon atoms between two successive amide groups, such as nylon 6 and nylon 12. For nylons based on diamines and dicarboxylic acids, the numeral is composed successively of the number of carbons in the diamine and diacid molecules, respectively. Thus, nylon 6,12 is made of 1,6-hexamethylenediamine and 1,12-dodecanoic acid. The most important nylons produced commercially are nylon 4,6; nylon 6,6; nylon 6,9; nylon 6,10; nylon 6,12; nylon 6; nylon 11; and nylon 12.^[34]

Nylon 6 is prepared by hydrolytic or anionic ring-opening polymerization of caprolactam. The caprolactam is made from cyclohexane, phenol, or toluene. Nylon 6,6 is made by the polycondensation reaction of two monomers, each of which contain six carbon atoms. One of the monomers is a six-carbon acid with a $-\text{COOH}$ group at each end, hexanedioic acid (adipic acid). The other monomer is a six-carbon chain with an amino group, $-\text{NH}_2$, at each end, 1,6-diaminohexane (1,6-hexamethylenediamine).



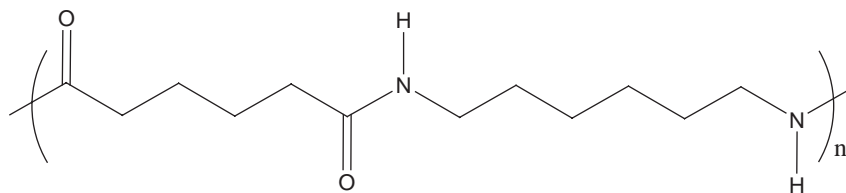


Fig. 2 Nylon 6,6.

When these two compounds polymerize, the amine and acid groups combine, each time losing a molecule of water. Because the acid is acidic and the amine is basic, they first react together to form a salt. The salt is then converted into nylon 6,6 by heating under pressure at 350°C (Fig. 2).

Similarly, nylon 4,6 is prepared from 1,4-diaminobutane and adipic acid. Other nylon salts are synthesized in the first step from hexamethylenediamine and the corresponding dicarboxylic acids, azelaic acid (obtained from oleic acid) and 9 carbon atoms for nylon 6,9; sebacic acid (obtained from castor oil) and 10 carbon atoms for nylon 6,10; and dodecandioic acid (obtained from butadiene) and 12 carbon atoms for nylon 6,12.

Nylon 11, also called Rilsan[®], is prepared by the self-condensation of ω -aminoundecanoic acid (11-aminoundecanoic acid), a crystalline solid. The raw material for the amino acid is castor oil, a vegetable product composed of the triglyceride of ricinoleic acid. The self-condensation of ω -aminoundecanoic acid is conducted in a continuous process at about 220°C, withdrawing water continuously. The latter stages of the process are conducted under reduced pressure. Nylon 11 is highly resistant and is attacked by very few solvents and chemicals.

Nylon 12 is synthesized by the hydrolytic polymerization of dodecanelactam prepared from butadiene by a multistage reaction. The process is conducted at about 300°C in the presence of aqueous phosphoric acid.

Nylons are linear thermoplastic polymers containing highly polar amide groups separated by a flexible aliphatic hydrocarbon chain. The carbonyl group of the amide linkage of a polyamide unit can become associated with the amino hydrogen of the amide linkage of an adjacent molecule through H-bonds. H-bonding affects several properties of these polymers. They have sharp melting points, high tensile strength, rigidity, hardness, high abrasion resistance, reduced solubility, and low thermal coefficient of linear expansion.^[34,35] When exposed to light, especially in the presence of air, there is a strong tendency for these polyamides (nylons) to lose tensile strength and to discolor. This is most likely a result of oxidation of amino groups and amide groups of the polymer molecule from effects of radiant energy on the polymer in contact with oxygen. Antioxidants of various kinds

are partially effective in stabilizing the nylons against the effects of oxygen and ultraviolet light. Some of these include mercaptobenzothiazole, syringic acid, and various phenolic antioxidants, chromium, and copper salts. The oxidation of polyamides is accelerated by heat or light or both. It is noteworthy that carbon black is an effective stabilizer against exposure to light and air, presumably because it absorbs radiation and thus slows the rate of deterioration.^[35]

Around 80% of nylon production goes into fiber applications such as clothing, tire cords, ropes, and carpets. Nylon 6 is considered the best material for shock-proof parts. Nylon 6,6 is used for high mechanical loads. Nylon 11 and 12 are preferred for cable sheathing, surgical sutures, and precision engineering parts. Nylons are extensively used in the automotive industry as parts of chassis, steering wheels, bumpers, housings, rear-view mirrors, parts for electrical equipment, transmission systems, and fuel supply system parts in the engine block and the lubricating system. In mechanical engineering nylons are used in gears, cams, bearings, etc. In telecommunications they are used for circuits and relays, contact makers, drive units, cable sheaths, and plug and socket devices. Nylons are also used in water heater tanks, heat exchangers, thermostat casings, fans, chairs, tables, combs, and sterilizable moldings for medical uses. Nylon films are used in packaging pharmaceutical and food products.^[34]

Polyamides have good biocompatibility with human tissue because of its similarity in chemical structure with the collagen protein. Recently, a new kind of tissue engineering scaffold material of needle-like nano-hydroxyapatite (n-HA) and polyamide (PA66) biocomposite was prepared using the cosolution coprecipitation method. Strong molecular interactions and chemical bondings exist between the n-HA and PA66 in the composite. This new scaffold has excellent mechanical properties close to that of natural bone.^[36]

Kevlar[®]

Kevlar[®] is a brand of synthetic polyamides made by DuPont. Kwolek and Blades developed Kevlar in 1965 at DuPont.^[37] Fibers of Kevlar consist of long molecular chains produced from polyparaphenylene terephthalamide. Kevlar is similar in structure to nylon 6,6, except that instead of the amide links joining

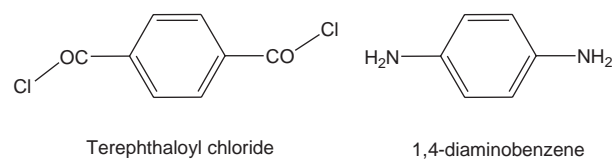


Fig. 3 Monomer units of Kevlar.

chains of carbon atoms together, they join benzene rings. The two monomers are terephthaloyl chloride and 1,4-diaminobenzene (Fig. 3).

When these monomers react and eliminate a molecule of hydrochloric acid as a byproduct, an amide bond is formed and the chain grows in length. The structure of Kevlar is illustrated in Fig. 4.

Many interchain bonds make Kevlar highly oriented and extremely strong. Kevlar derives a portion of its improved strength from intermolecular hydrogen bonds formed between the carbonyl groups and protons on neighboring polymer chains and the partial pi stacking of the benzenoid aromatic stacking interactions between stacked strands. These interactions result in a unique combination of properties. Kevlar has high tensile strength, low electrical conductivity, and excellent dimensional stability and is chemical and flame resistant. There are three common grades of Kevlar: pure Kevlar, Kevlar 29, and Kevlar 49.

Since its discovery, Kevlar has been used for many applications such as bulletproof vests, composites for boat construction, lightweight mountaineering ropes, lightweight skis, and helmets and racquets. Kevlar has played a significant role in many critical and diverse applications: ropes that secure airbags in the crucial landing apparatus of the Mars Pathfinder; shrapnel-resistant shielding in jet aircraft engines that can protect passengers in case an explosion occurs; and gloves that protect hands and fingers against cuts, slashes, and other injuries that often occur in glass and sheet metal factories.^[37]

Because biopolymers are now being used as replacements for ligaments, Wening et al. recently tested the cytotoxicity and mutagenicity of Kevlar 49 against six strains of *Salmonella typhimurium* and Chinese hamster cells, respectively, and reported that Kevlar 49 does not exhibit mutagenic or cytotoxic activity.^[38]

Polyamides are fairly readily attacked by strong acids, but are much more resistant to alkaline

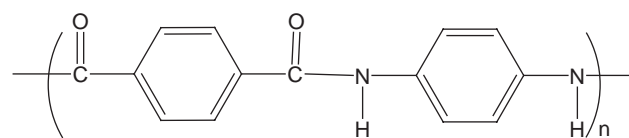


Fig. 4 Structure of Kevlar.

hydrolysis. Hydrolysis is faster at higher temperatures. Hydrolysis by water alone is so slow as to be completely unimportant. Kevlar is more resistant to hydrolysis than nylon.^[39]

Versamid

The synthetic polyamide resins made from condensation of polymerized vegetable oil acids and polyalkylene polyamines are marketed under the trade name Versamid[®].^[35] They are available in several grades, ranging from materials that are liquid at room temperature to hard, solid materials melting at temperatures as high as 185°C. The polyamide resins are considered internally plasticized by the long carbon chains and the large alkyl substituents on the chains of the dibasic acid portion. Versamids are classed as relatively low-melting, highly soluble polyamide resins of low to moderate molecular weight. Solid versamid resins are used as such or modified with plasticizers, waxes, and other resins in hot-melt cements, heat seal coatings, barrier coatings, inks, and specialty adhesives and coatings. The reactive Versamid polyamides may be used to modify the solid versamid resins or converted to thermoset resins by reaction with epoxy resins, heat-reactive phenolic resins, and other substances. The Versamid resins are frequently used as adhesives and include Versamid resins 930, 940, 950, and 100. The adhesives are used for a wide range of different purposes such as bookbinding adhesives, heat seal adhesives for bread wrappers, cigarette packages, packaged soups and packaged foods in paperboard boxes, and the manufacture of corrugated fiber boxes. Generally, they are used with flexible packaging materials including paper, glassine, cellophane, aluminum foil, and cellulose acetate films.^[35]

Versamid polyamide resins are useful in the field of printing inks, particularly for flexographic and roto-gravure printing and for clear overprint varnishes. Their use is closely tied to the needs of the flexible packaging industry because inks based on Versamid polyamides are largely used for printing on treated polyethylene and polypropylene, cellophane, coated cellophanes, saran, cellulose acetate, Mylar[®], and other plastic films as well as glassine, paper, and aluminium foil. Packages designed from printed stock are used for food products such as apples, carrots, and potatoes, for small articles of clothing, for household goods such as cellulose sponges, and for many other products.^[35]

Amino-containing polyamide resins of the fluid series, Versamid 115 and Versamid 125, find widespread use in casting and laminating applications. When blended with liquid epoxy resins, they harden into tough thermoset products, which are especially

well suited for use in plastic tools and dies, glass-fiber laminates, electrical potting and encapsulating, and general casting and molding. When the fluid Versamid resins are blended with liquid epoxy resins, a chemical reaction occurs between the amino groups of the polyamides and the epoxy groups of the epoxy resins. A cross-linked or 3-D polymer is formed. When the longer-chain reactive polyamides are cross-linked with the epoxy resins, a more flexible, resilient structure is obtained than that of the epoxy cross-linked with the shorter, more rigid molecules of the amine hardeners.

Versamids may be used alone or with plasticizers or other modifiers as surface coatings. In this form they are used most frequently as coatings for flexible substrates, including paper, cellophane, and aluminum foil. They are valued for their heat-seal properties and for their grease and water barrier properties. The soluble polyamides are applied over printed matter on packages as gloss overprint varnishes to protect the printing against abrasion and effects of oil or water, as well as to give an improvement in appearance, gloss, and depth of color.^[35]

Sodium Polyaspartate

Sodium polyaspartate (Fig. 5) is a condensation polymer based on aspartic acid. When aspartic acid is simply heated to 180°C, many condensation reactions occur and the equivalent of water is eliminated with each condensation. The product formed is water-insoluble polysuccinimide. In the subsequent step the polymer is reacted with sodium hydroxide in water, which results in partial cleavage of the amide bonds. Hydrolysis can occur from either of the two sites of the

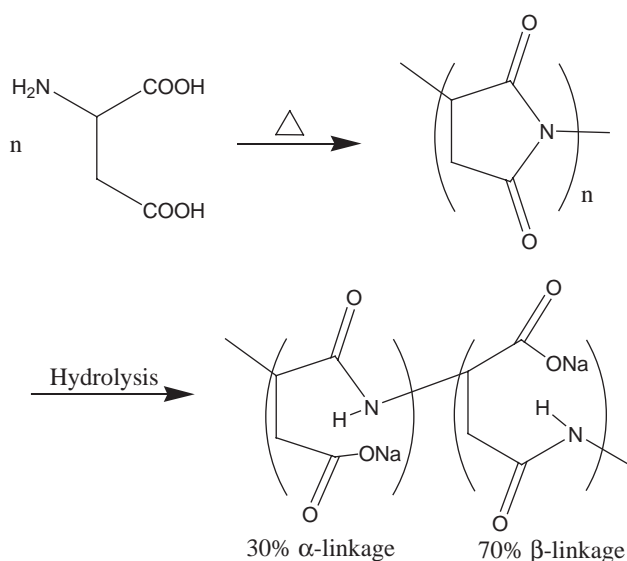


Fig. 5 Structure of sodium polyaspartate.

succinimide ring. Two different bonds (α and β) are hydrolyzed, resulting in a sodium polyaspartate copolymer with 30% α -linkages and 70% β -linkages.^[40] As a polymer of amino acid, it is an example of a polypeptide.

Sodium polyaspartate is used as a chelating agent. It can be applied to remove calcium phosphate deposits consisting of hydroxyapatite and brushite or dicalcium phosphate dihydrate from stainless-steel surfaces.^[41] Sodium polyaspartate is also present in mascara, shampoo, facial moisturizer, facial cleanser, and moisturizer.^[42] Sodium polyaspartate (PASP) acts as a stabilizer for indocyanine green (ICG), which is a widely used dye in cardiovascular, hepatic, and ophthalmological studies. The stock solution of ICG-PASP is stable and demonstrated fluorescence over a wide range of concentration for 2 weeks when stored in the dark.^[43] Recently, some bioinorganic nanocomposites have been synthesized in which the layered double hydroxides are intercalated with poly(α,β -aspartate). They have good interlayer spacing and modified textures. The bioinorganic hybrid material formed can be used as organic clays with bioactive, biocompatible, and biodegradable properties.^[44]

Pyrrole–Imidazole Polyamides

DNA acts as the storehouse of the biological information for all living organisms. Any chemical substance that can interfere with the four bases present in the DNA severely affects the cellular function of all organisms. Chemical substances that can recognize and bind DNA in a sequence-specific manner have enormous importance in modern biology and medicine. Dervan et al. developed the minor groove-binding hairpin polyamides that contain *N*-methylimidazole (Im)–*N*-methylpyrrole (Py)–hydroxypyrrrole (Hp), which uniquely recognize each of the four Watson-Crick base pairs.^[45,46] Hairpin polyamides are derived from distamycin A, which binds to A:T-rich sequences of DNA. Composed of Py and Im compounds, hairpin polyamides are synthetic ligands that have an affinity and specificity for DNA similar to that of many naturally occurring DNA binding proteins. An Im–Py recognizes G:C, whereas a Py–Im recognizes C:G. In contrast, a Py–Py pair degenerately recognizes A:T or T:A base pairs. An Hp and Py pair (Hp–Py) distinguishes T:A from A:T base pairs. The structure of pyrrole–imidazole polyamides is illustrated in Fig. 6.

The hairpin polyamides are cell permeable and bind to DNA at nanomolar concentrations, with binding coefficients similar to that of transcription factors. The solid phase synthesis of Py–Im polyamides was achieved in 1996 by *t*-Boc chemistry using a

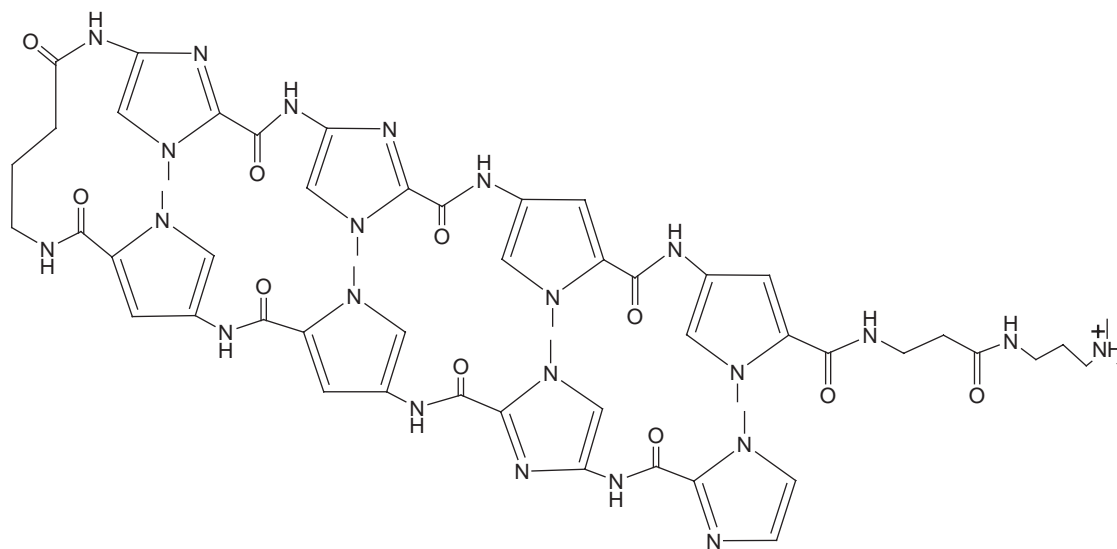


Fig. 6 Structure of pyrrole-imidazole polyamides.

combination of building blocks, Boc-Py-OBt ester and Boc-Im acid and Boc- β -alanine-Pam resin. Cleavage with *N,N*-dimethylaminopropylamine (Dp) affords a β -alanine-Dp tail at the C terminus of Py-Im polyamides. These polyamides can be used for regulation of gene expression and inhibition of gene expression and as gene activators. The cell permeability of polyamides is useful in cell culture experiments.^[47] Wang et al.^[48] recently reported the alkylation of DNA using a combination of the anticancer drug chlorambucil with a sequence-selective hairpin pyrrole-imidazole polyamide ImPy- β -ImPy- γ -ImPy- β -Dp. The alkylation occurs predominantly at the 5'-AGCTGCA-3' sequence, which represents the polyamide binding site. The alkylation inhibits the *in vitro* DNA replication, but not protein or RNA synthesis. Flow cytometric analysis demonstrated that the hybrid arrested cells in the G₂/M phase.^[48] Similarly, Wurtz et al. reported that hairpin polyamide-chlorambucil conjugate binds and selectively alkylates predetermined sites in the HIV-1 promoter at subnanomolar concentrations.^[49]

Aromatic Polyamides/Aramidics

According to the definition of the U.S. Federal Trade Commission, aromatic polyamides or aramid fibers are manufactured fibers that contain a long polyamide chain, with at least 85% of the amide linkages directly attached to the two aromatic rings.^[34] The aromatic rings impart rigidity to the polymer, and in the production of the fibers the stiff chains are generally aligned parallel to the fiber axis, thus leading to a modulus of elasticity, which can be as high as 130 GPa.

Although they are resistant to heat compared with other fibers, at temperatures higher than 300°C the fiber may lose most of its strength.^[50]

These aromatic polyamides cannot be synthesized by the direct condensation of amines and acids because the aromatic amines are less reactive and do not provide polymeric materials. The other hurdle is discoloration and decomposition in the molten state, making the bulk polycondensation technique impractical. Therefore, the aromatic polyamides are synthesized by a more costly solution and interfacial polycondensation technique. The diacid chlorides and diamines serve as starting materials.^[34]

The first aramidic fiber, poly-*m*-phenylene isophthalamide, was prepared in the 1960s by the polycondensation of isophthaloyl chloride with *m*-phenylenediamine (Fig. 7). It is commercialized as Nomex[®].

This fiber has been used in applications where high heat resistance is required, such as heat-protective clothing. At normal temperatures their chemical, mechanical, and electrical insulation properties are comparable with that of nylon 6 and nylon 6,6. The aramidic fibers are directly obtained by the polymerization process and not from fusion, because they degrade before fusion takes place.^[51]

Poly-*p*-phenylene terephthalamide is another aromatic polyamide fiber introduced to the market in 1972. It is prepared by the low-temperature polycondensation of terephthaloyl chloride with *p*-phenylenediamine in a solvent mixture of hexamethylphosphoramide and *N*-methylpyrrolidone in 2:1 ratio at -10°C. The fiber has the highest strength and density among commercially available fibers and is highly competitive with glass, carbon, and steel fibers. It is used as a filler in

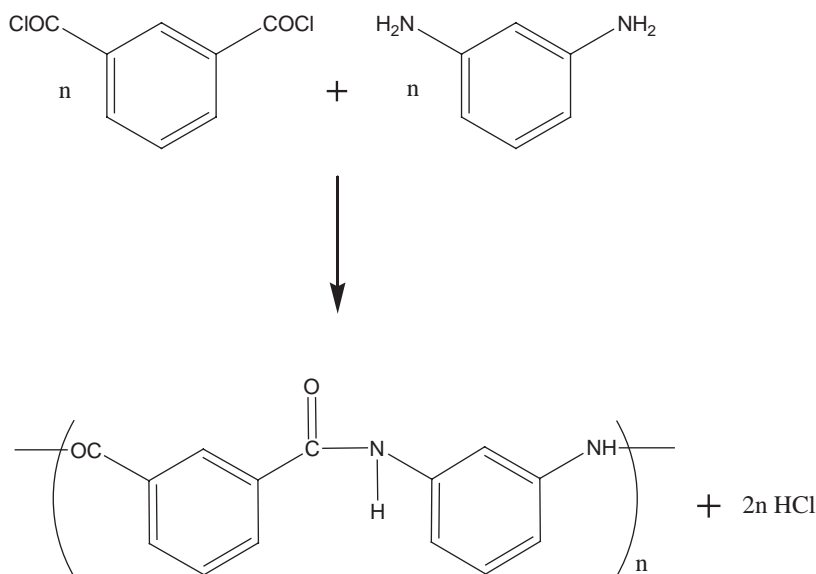


Fig. 7 Structure and synthesis of Nomex.

composite materials because of its lighter weight, greater stiffness, and high tensile strength. Other uses for poly-*p*-phenylene terephthalamide include reinforced composites for boat hulls, military protection helmets, lightweight bullet-proof jackets, and filament-wound rocket motors. The fiber is also used as reinforcement in radial tires.^[34]

Peptide Nucleic Acids

Peptide nucleic acids were originally designed as a ligand for the recognition of double-stranded DNA. The concept was to mimic an oligonucleotide binding to double-stranded DNA via Hoogsteen base pairing in the major groove. Peptide nucleic acid is a nucleic acid analog in which the sugar phosphate backbone of natural nucleic acid has been replaced by a synthetic peptide backbone, usually formed from *N*-(2-aminoethyl)-glycine units, resulting in an achiral and uncharged mimic. They do not exist in nature but are synthesized chemically (Fig. 8). The PNAs are chemically and enzymatically (proteases and nucleases) stable and thus are not prone to be degraded inside the living cell. Peptide nucleic acid is an excellent structural mimic of DNA (or RNA), and PNA oligomers are able to form very stable duplex structures with Watson–Crick complementary DNA, RNA (or PNA) oligomers and they can also bind to targets in duplex DNA by helix invasion. The hybrid complexes exhibit extraordinary thermal stability and unique ionic strength effects. Therefore, these molecules are of interest in many areas of chemistry, biology, and medicine including drug discovery, genetic diagnostics, molecular recognition, and the origin of life.

The first experiments conducted with homothymine PNAs clearly demonstrated that PNAs bound to double-stranded DNA in a sequence-specific manner. Two homothymine PNAs had formed a triplex with

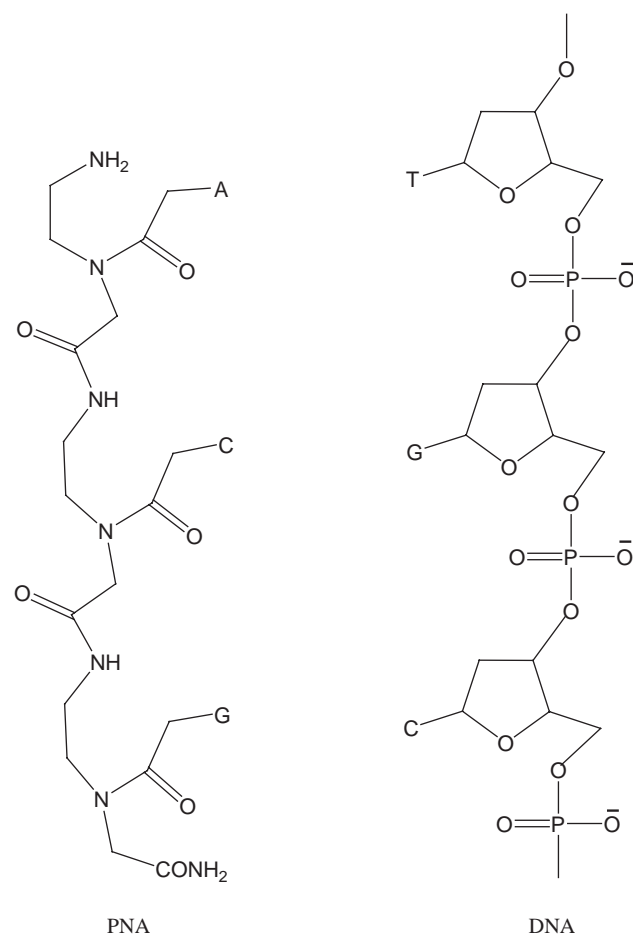


Fig. 8 Structure of PNA and DNA. A, adenine; T, thymine; G, guanine; C, cytosine.

the homoadenine target in the double-stranded DNA, while displacing the homothymine strand in the DNA target. Later it was revealed that PNAs with both purine and pyrimidine bases form stable duplexes with DNA and RNA, although not with the extremely high stability of the homopyrimidine 2PNA/DNA triplexes, but still more stable than the corresponding DNA/DNA and DNA/RNA duplexes.^[52]

In vitro studies indicate that PNA could inhibit both transcription and translation of genes to which it has been targeted, which holds promise for its use for antigene and antisense therapy. However, as with other high-molecular-weight drugs, the delivery of PNA, involving passage through the cell membrane, appears to be a general problem.^[53] The PNA oligomers can also be used to selectively suppress the growth of specific types of bacteria.^[54]

The unique physicochemical characteristics of PNAs led to the development of PNA probes for various molecular, genetics, and cytogenetics applications. Also, PNA-based hybridization technology is developing rapidly within the field of in situ fluorescence hybridization, indicating the great potential of PNA probes for chromosomal investigations.^[55]

Peptidomimetics

Peptides and proteins are important constituents of the human body and microorganisms such as bacteria and viruses, but their use in medicinal purposes suffers from many problems. They are poorly absorbed in the gastrointestinal tract and have a short plasma half-life when given in the form of injection. Therefore, attempts were made to design and synthesize molecular fragments that could replace parts of peptides. A peptidomimetic is a small protein-like chain designed to mimic a peptide. They typically arise from modification of an existing peptide to alter the molecule's properties. For example, they may arise from modifications to change the molecule's stability or biological activity. This can play a role in the development of drug-like compounds from existing peptides. These modifications involve changes to the peptide that will not occur naturally (such as altered backbones and the incorporation of non-natural amino acids). Alternatively, they could be large molecules that are completely unrelated to peptides but look like peptides. The peptidomimetics have some important features in their use as drugs or in biological application. They have conformationally restrained structures that can minimize binding to non-target receptors and enhance activity at the desired receptor. The addition of hydrophobic groups or replacement of amide bonds results in better transport properties through cellular membranes. Isosteres, cyclic peptides, and more

modification in the peptidomimetics reduce the rate of degradation by peptidases and other enzymes.^[56] In the future, these peptidomimetics would be designed to easily pass through cell membranes and to survive in the human body long enough to reach the biological target. Peptidomimetics will be useful for infectious diseases like malaria, tuberculosis, hepatitis C, and HIV and for non-infectious diseases like Alzheimer's and high blood pressure.^[57]

An example of peptidomimetics is an anticancer agent designed and synthesized to induce programmed cell death (apoptosis) of cancer cells. Walensky et al. synthesized a peptidomimetic called stabilized α -helix of BCL-2 domains, which was effective in inhibiting the growth of human leukemia cells xenografted in mice.^[58]

CONCLUSION

We summarized the polyamides into two broad categories—natural and synthetic. These polyamides play an important role in every aspect of life. They are crucial for the biological processes as enzymes and are important for the outer skeleton of humans and animals. Polyamides find strong applications in modern industries as nylons, Kevlar, sodium polyaspartate, Versamids, etc. They are also used in the biotechnological and pharmaceutical industries. The emerging fields of science such as tissue engineering, protein engineering, and biomaterial engineering are using these polyamides as scaffolds for various applications.

ARTICLES OF FURTHER INTEREST

Biofunctional Polymers; Biomimetic Materials; Collagen; Elastin; Protein Adsorption Modeling; Protein Delivery Systems

REFERENCES

1. <http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/P/Proteins.html> (accessed March 2007).
2. http://ull.chemistry.uakron.edu/genobc/Chapter_19/ (accessed March 2007).
3. Wang, X.; Kim, H.J.; Wong, C.; Vepari, C.; Matsumoto, A.; Kaplan, D.L. Fibrous proteins and tissue engineering. *Mater. Today* **2006**, *9* (12), 44–53.
4. Kochendoerfer, G.G. Site-specific polymer modification of therapeutic proteins. *Curr. Opin. Chem. Biol.* **2005**, *9* (6), 555–560.
5. Watson, D.W.; Cromartie, W.J.; Bloom, W.L.; Heckly, R.J.; McGhee, W.J.; Weissman, N. Studies on infection with *Bacillus anthracis*. V. The isolation of an inflammatory factory factor from crude extracts of lesions

- of *B. anthracis* infection and its biological and chemical relationship to glutamyl polypeptide. *J. Infect. Dis.* **1947**, *80*, 121–136.
6. Kambourova, M.; Tangney, M.; Priest, F.G. Regulation of polyglutamic acid synthesis by glutamate in *Bacillus licheniformis* and *Bacillus subtilis*. *Appl. Environ. Microbiol.* **2001**, *67*, 1004–1007.
 7. Kandler, O.; König, H.; Wiegel, J.; Claus, D. Occurrence of poly- γ -D-glutamic acid and poly- α -L-glutamine in the genus *Xanthobacter*, *Flexithrix*, *Sporosarcina* and *Planococcus*. *Syst. Appl. Microbiol.* **1983**, *4*, 34–41.
 8. Hezayen, F.F.; Rehm, B.H.A.; Tindall, B.J.; Steinbüchel, A. Transfer of *Natrialba asiatica* BIT to *Natrialba taiwanensis* sp. nov. and description of *Natrialba aegyptiaca* sp. Nov., a novel extremely halophilic, aerobic, non-pigmented member of archaea from Egypt which produces extracellular poly (glutamic acid). *Int. J. Syst. Evol. Microbiol.* **2001**, *51*, 1133–1142.
 9. Weber, J. Poly(γ -glutamic acid)s are the major constituents of nematocysts in Hydra (Hydrozoa, Cnidaria). *J. Biol. Chem.* **1990**, *265*, 9664–9669.
 10. Alford, D.D.; Wheeler, A.P.; Pettigrew, C.A. Biodegradation of thermally synthesized polyaspartate. *J. Environ. Polym. Degrad.* **1994**, *2*, 225–236.
 11. Pötter, M.; Oppermann-Sanio, F.B.; Steinbüchel, A. Cultivation of bacteria producing polyamino acids with liquid manure as carbon and nitrogen source. *Appl. Environ. Microbiol.* **2001**, *67*, 617–622.
 12. Shima, S.; Sakai, H. Polylysine produced by *Streptomyces*. *Agric. Biol. Chem.* **1977**, *41*, 1807–1809.
 13. Shima, S.; Matsuoka, H.; Iwamoto, T.; Sakai, H. Antimicrobial action of ϵ -poly-L-lysine. *J. Antibiot* **1984**, *37*, 1449–1455.
 14. Hirayama, C.; Sakata, M.; Nakamura, M.; Ihara, H.; Kunitake, M.; Todokora, M. Preparation of poly (ϵ -lysine) adsorbents and application to selective removal of lipopolysaccharides. *J. Chromatogr. B* **1999**, *721*, 187–195.
 15. Lang, N.J. The fine structure of blue-green algae. *Annu. Rev. Microbiol.* **1968**, *22*, 15–46.
 16. Tomaselli, F.L.; Margheri, M.C.; Pelosi, E. Ultrastructure of *Spirulina* in comparison with *Oscillatoria*. *Zentralbl. Bakteriell. Abt. II* **1976**, *131*, 592–601.
 17. Sherman, D.M.; Tucker, D.; Sherman, L.A. Heterocyst development and localization of cyanophycin in N_2 -fixing cultures of *Anabaena* sp. PCC 7120 (Cyanobacteria). *J. Phycol.* **2000**, *36*, 932–941.
 18. Simon, R.D. Cyanophycin granules from the blue-green alga *Anabaena cylindrical*: A reserve material consisting of copolymers of aspartic acid and arginine. *Proc. Natl. Acad. Sci. USA* **1971**, *68*, 265–267.
 19. Schwamborn, M. Chemical synthesis of polyaspartates: A biodegradable alternative to currently used polycarboxylate homo- and copolymers. *Polym. Degrad. Stab.* **1998**, *59*, 39–45.
 20. Aboulmagd, E.; Oppermann-Sanio, F.B.; Steinbüchel, A. Molecular characterization of the cyanophycin synthetase from *Synechocystis* sp. strain PCC6308. *Arch. Microbiol.* **2000**, *174*, 297–306.
 21. Oppermann-Sanio, F.B.; Hai, T.; Aboulmagd, E.; Hezayen, F.F.; Jossek, S.; Steinbüchel, A. Biochemistry of microbial polyamide metabolism. In *Biochemical Principles and Mechanisms of Biosynthesis and Biodegradation of Polymers*; Steinbüchel, A., Ed.; Wiley-VCH: Weinheim, 1991; 185–193.
 22. Ziegler, K.; Diener, A.; Herpin, C.; Richter, R.; Deutzmann, R.; Lockau, W. Molecular characterization of cyanophycin synthetase, the enzyme catalyzing the biosynthesis of the cyanobacterial reserve material multi-L-arginyl-poly-L-aspartate (cyanophycin). *Eur. J. Biochem.* **1998**, *254*, 154–159.
 23. Conrad, U. Polymers from plants to develop biodegradable plastics. *Trends Plant Sci.* **2005**, *10* (11), 511–512.
 24. <http://www.physorg.com/news70608222.html> (accessed February 2007).
 25. http://tc.engr.wisc.edu/Steuber/papers/2002/spider_p_1st.doc (accessed February 2007).
 26. Wang, Y.; Kim, H.-J.; Novakovic, G.V.; Kaplan, D.L. Stem cell-based tissue engineering with silk biomaterials. *Biomaterials* **2006**, *27* (36), 6064–6082.
 27. <http://www.mcgill.ca/files/ott/nexia.pdf> (accessed March 2007).
 28. http://www.aiche.org/uploadedFiles/SBE/DepartmentUploads/MayCEP_EngineeringBiomedicalMaterialsfromCornandWool.pdf (accessed February 2007).
 29. <http://www.cyclotide.com/what.html> (accessed March 2007).
 30. <http://www.medfarm.uu.se/forskning/program.php?vetenskapsid=2&forskomr=43&id=241&lang=en> (accessed February 2007).
 31. <http://www.cyclotide.com/uses.html> (accessed March 2007).
 32. Lindholm, P.; Goransson, U.; Johansson, S.; Claeson, P.; Gullbo, J.; Larsson, R.; Bohlin, L.; Backlund, A. Cyclotides: A novel type of cytotoxic agents. *Mol. Cancer Ther.* **2002**, *1* (6), 365–369.
 33. Svargard, E.; Goransson, U.; Hocaoglu, Z.; Gullbo, J.; Larsson, R.; Claeson, P.; Bohlin, L. Cytotoxic cyclotides from *Viola tricolor*. *J. Nat. Prod.* **2004**, *67* (2), 144–147.
 34. Feldman, D.; Barbalata, A., Eds. Polyamides & polyimides. In *Synthetic Polymers: Technology, Properties, Applications*; Chapman & Hall: London, 1996; 222–258.
 35. Floyd, D.E., Ed.; *Polyamide Resins*, 2nd Ed.; Reinhold: New York, 1966.
 36. Jie, W.; Yubao, L. Tissue engineering scaffold material of nano-apatite crystals and polyamide composite. *Eur. Poly. J.* **2004**, *40*, 509–515.
 37. <http://www.dupont.com/kevlar/whatiskevlar.html> (accessed February 2007).
 38. Wening, J.V.; Marquardt, H.; Katzer, A.; Jungbluth, K.H.; Marquardt, H. Cytotoxicity and mutagenicity of Kevlar[®]: An in vitro evaluation. *Biomaterials.* **1995**, *16*, 337–340.
 39. <http://www.chemguide.co.uk/organicprops/amides/polyamides.html> (accessed February 2007).
 40. Bennett, G.D. A green polymerization of aspartic acid for the undergraduate organic laboratory. *J. Chem. Educ.* **2005**, *82* (9), 1380–1381.
 41. Littlejohn, F.; Saez, A.E.; Grant, C.S. Use of sodium polyaspartate for the removal of hydroxyapatite/



- brushite deposits from stainless steel tubing. *Ind. Eng. Chem. Res.* **1998**, *37* (7), 2691–2700.
42. <http://www.ewg.org/reports/skindeep2/report.php?type=INGREDIENT&id=2065> (accessed February 2007).
 43. Maarek, J.M.; Holschneider, D.P.; Harimoto, J. Fluorescence of indocyanine green in blood: intensity dependence on concentration and stabilization with sodium polyaspartate. *J. Photochem. Photobiol. B* **2001**, *65* (2–3), 157–164.
 44. Whilton, N.T.; Vickers, P.J.; Mann, S. Bioinorganic clays: Synthesis and characterization of amino- and polyamino acid intercalated layered double hydroxides. *J. Mater. Chem.* **1997**, *7* (8), 1623–1629.
 45. Trauger, J.W.; Baird, E.E.; Dervan, P.B. Recognition of DNA by designed ligands at subnanomolar concentrations. *Nature* **1996**, *382*, 559–561.
 46. White, S.; Szewczyk, J.W.; Turner, J.M.; Baird, E.E.; Dervan, P.B. Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands. *Nature* **1998**, *391*, 468–471.
 47. Murty, M.S.R.C.; Sugiyama, H. Biology of *N*-methylpyrrole-*N*-methylimidazole hairpin polyamide. *Biol. Pharm. Bull.* **2004**, *27* (4), 468–474.
 48. Wang, Y.-D.; Dziegielewska, B.; Dervan, P.B.; Beerman, T.A. DNA crosslinking and biological activity of a hairpin polyamide–chlorambucil conjugate. *Nucleic Acids Res.* **2003**, *31* (4), 1208–1215.
 49. Wurtz, N.R.; Dervan, P.B. Sequence specific alkylation of DNA by hairpin pyrrole–imidazole polyamide conjugates. *Chem. Biol.* **2000**, *7* (3), 153–161.
 50. Zheng, Z.; Feldman, D. Synthetic fibre-reinforced concrete. *Prog. Polym. Sci.* **1995**, *20*, 185–210.
 51. <http://www.technica.net/NT/NT3/composites.htm> (accessed February 2007).
 52. <http://www.horizonpress.com/gateway/pna.html> (accessed February 2007).
 53. Ray, A.; Norden, B. Peptide nucleic acid (PNA): Its medical and biotechnical applications and promise for the future. *FASEB J.* **2000**, *14*, 1041–1060.
 54. Pooga, M.; Land, T.; Bartfai, T.; Langel, U. PNA oligomers as tools for specific modulation of gene expression. *Biomol. Eng.* **2001**, *17*, 183–192.
 55. Paulasova, P.; Pellestor, F. The peptide nucleic acids (PNAs): A new generation of probes for genetic and cytogenetic analyses. *Ann. Genet.* **2004**, *47* (4), 349–358.
 56. http://daecr1.harvard.edu/pdf/smnr_2000-2001_Peterson_Gretchen.pdf (accessed March 2007).
 57. http://www.medfarm.uu.se/forskning/program.php?ve_tenskapsid=2&forskomr=43&id=216&lang=en (accessed March 2007).
 58. Walensky, L.D.; Kung, A.L.; Escher, I.; Malia, T.J.; Barbuto, S.; Wright, R.D.; Wagner, G.; Verdine, G.L.; Korsmeyer, S.J. Activation of apoptosis in vivo by a hydrocarbon-stapled BH3 helix. *Science*. **2004**, *305* (5689), 1466–1470.

Polycaprolactone

J.E. Gough

School of Materials, University of Manchester, Manchester, Greater Manchester, U.K.

P. Christian

School of Chemistry, University of Manchester, Manchester, Greater Manchester, U.K.

INTRODUCTION

Polycaprolactone (PCL) is an aliphatic polyester and was initially investigated for use as a biodegradable packaging material after the realization of its degradability by microorganisms. PCL is also degraded by hydrolysis in physiological conditions and has therefore received a great deal of attention for use as an implantable material.

PCL is now one of several degradable polymers approved for use in the human body as drug delivery devices, adhesion barriers, sutures, and staples and is being extensively investigated as a biomaterial for tissue repair and regeneration. Degradable polymers have different degradation rates and are commonly investigated as blends as a method to tailor the degradation rate. The focus of this article is the recent use of PCL as a biomaterial.

SYNTHESIS AND STRUCTURE

Polymerization

PCL is a biodegradable polyester, which may be prepared by a variety of routes (cationic,^[1] anionic,^[2] and coordination polymerization^[3] as well as by free-radical initiation^[4] and is generally a tough white material, insoluble in water but soluble in a range of organic solvents. Commercially available material is generally a high-molecular-weight straight-chain polymer with a narrow molecular weight distribution. It has a low T_g of -60°C and a T_m of 60°C , the polymer being around 57% crystalline at an M_w of 120 kDa.^[1] Relatively, low-molecular-weight polymer was initially prepared by Carrothers who condensed capronic acid by simply heating it.^[5] The reaction results in the elimination of water. However, the reaction is an equilibrium and the reverse reaction is sufficiently dominant to make a significant impact on the molecular weight of the final material and therefore compromise its mechanical properties.

A more practical approach to its synthesis is via the ring opening polymerization (ROP) of a lactone,

(a cyclic ester). PCL may be prepared by anionic or cationic polymerization of ϵ -caprolactone (Fig. 1A). These reactions produce high-molecular-weight material with very little chain branching and therefore a high degree of crystallinity. High-molecular-weight material may also be prepared by free-radical polymerization of 2-methylene-1,3-dioxepane Fig. 1B.^[4] This method relies on a propagating primary carbon radical, which is very reactive and renders the polymerization reaction very sensitive to contaminants. The final product is a polymer with a significant amount of branching, which results owing to a radical back-biting mechanism. The branching necessarily results in a polymer with lower crystallinity. Generally the tin-catalyzed polymerization of the bulk monomer is used to prepare the materials, and this results in high-molecular-weight polymer (200 kDa) with a T_m of 60°C and 57% crystallinity. The T_g of the polymer is fairly low at -60°C . The combination of high crystallinity and low T_g results in a polymer with a Young's modulus of approximately 200–400 MPa, which may be easily softened upon gentle heating.

ϵ -Caprolactone can be copolymerized with other monomers such as ethylene oxide, chloroprene, tetrahydrofuran, D-valerolactone, 4-vinylanisole, styrene, methylmethacrylate, vinylacetate^[6] (and many others as shown in Table 1^[7–17]), but most relevant to use in the human body are copolymers of caprolactone and lactic acid.^[3,18–21] The mechanical properties of PCL compared to that of polylactic acid (PLA) are shown in Table 2. A recent article by Albertsson and Varma^[22] details ROP of lactones for a variety of biomedical applications.

Enzyme-Catalyzed Polymerization

PCL and copolymers have been successfully polymerized using enzymes.^[23] This is an advantage for biomaterials development, as enzymes are natural molecules and rule out the use of harsh often-toxic catalysts and monomers. He et al.^[10] demonstrated the polymerization of ϵ -caprolactone with monohydroxyl or dihydroxyl poly(ethylene glycol) (PEG) using Novozyme-435 as a catalyst (an immobilized lipase B isolated from *Candida*

P

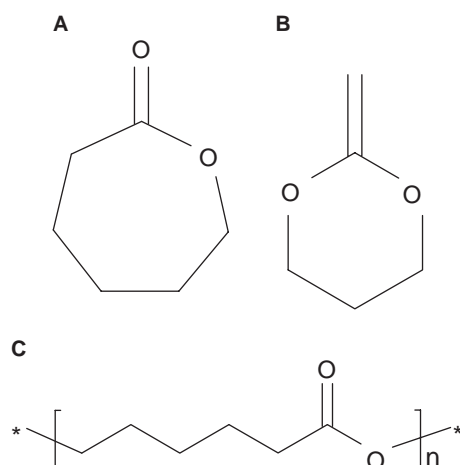


Fig. 1 (A) ϵ -caprolactone; (B) 2-methylene-1,3-dioxepane; and (C) PCL.

antartica). By altering the caprolactone:ethylene glycol ratio, control over molecular weight was obtained. These polymers were found to be semicrystalline and degradable by *Pseudomonas lipase*. ϵ -Caprolactone has also been copolymerized with D,L-lactide in a variety of ratios using the same enzyme.^[24]

DEGRADATION

Hydrolytic Degradation

Although PCL is biodegradable via hydrolysis of the ester linkages (Fig. 2), its degradation rate is very slow compared to materials such as PLA, as PCL has higher crystallinity and increased hydrophobicity, which reduces the diffusion rate of water through the polymer. PLA degrades rapidly and produces acidic degradation products that can lead to cell death and inflammation around the implant. Although the degradation of PCL is slower and produces less-acidic degradation products, the degradation may be too slow to allow efficient tissue regrowth. Ideally a biodegradable polymer will degrade at a rate similar to the rate of tissue formation. A great deal of research has therefore focused on producing PCL with a variable degradation rate and has aimed at the synthesis of copolymers^[25] or the manufacture of polymer blends.^[26]

The hydrolytic degradation may be catalyzed by either an increase or a decrease in the pH of the system and therefore may also be autocatalyzed by the resulting acidic by-products, eventually forming carbon dioxide and water. Degradation rate and mechanical strength can be altered by changes in molecular weight (lower molecular weight results in faster degradation) and crystallinity (amorphous regions degrade faster

than crystalline regions).^[6,7] Complete degradation of PCL is said to take more than two years *in vivo*^[27] and has a slower degradation rate than other biodegradable polymers.^[8]

When considering the biocompatibility of the polymer, it is very important to consider the catalyst. Many cationic or coordination polymerization reactions use compounds based on metals such as tin, titanium, or zinc. These metals remain in the polymer after processing and can have an adverse effect on cell proliferation. Whilst the use of anionic or free-radical polymerization might be used to circumvent the problem, anionic polymerizations are exceptionally sensitive to trace impurities and the free-radical polymerization relies on an unusual primary carbon radical, which is also very unstable. Furthermore the monomer for the free-radical reaction must be prepared and purified before the polymer may be produced.

There has been some success in using Lewis acids as catalysts for the polymerization. These can be based on catalysts such as BF_3 . The reaction proceeds at 80°C and tends to produce low-molecular-weight polymer ($< 30 \text{ kDa}$) unless a polyalcohol is added.^[28] However, the persistence of the catalyst is of particular use in the preparation of PCL with tuneable degradability. The BF_3 catalyses the hydrolysis reaction and can result in a rapid decrease in pH; treatment of the polymer by soaking in a potassium fluoride in ethanol solution allows the degradation to be retarded by converting the catalytic BF_3 species into an inactive KBF_4 salt. It also appears that the small amounts of fluoride and BF_4K , which remain, have little negative effect on cell proliferation.

While there has much interest in the biocompatibility of PCL in its pure form, it has also been used to prepare composites and, with the aid of rapid prototyping, patient-specific implants.^[29,30] The preparation of a truly biodegradable composite requires not only a biodegradable polymer but also a degradable reinforcing phase. The preparation of the final composite is not trivial. All the polymerization reactions are water sensitive and, owing to the adhesive properties of the polymer, moulds must be made from specialist materials such as polytetrafluoroethylene (PTFE). Furthermore PCL has a very high-melt viscosity ($\sim 50,000$ centipoise at 60°C). This makes simple injection molding of the composite almost impossible and results in poor fiber-polymer binding. To achieve reliable degradation and optimum mechanical properties, the composite must be prepared via a monomer transfer molding process (MTM). Particular success has been reported using the diethyl zinc-catalyzed polymerization of ϵ -caprolactone with vicryl mesh, and with BF_3 -catalyzed polymerization with phosphate glass fibers. During the MTM process, a reactive mixture of monomer and catalyst is prepared and then

Table 1 A review of several copolymer systems that include PCL segments

Polymer type	Notes	References
PCL-PEG multiblock	Prepared by butyltitanate-catalyzed reaction. Increased degradation rate in vitro and in vivo.	[7]
PCL-PEG multiblock	Prepared by tin alkoxide-catalyzed reaction. Increased photodegradation owing to PCL sensitization of the ether linkages.	[8]
PLA-PCL-PLA triblock	Prepared by zinc-catalyzed reaction. The polymers showed increased degradation rates compared to pure PCL but preferential degradation of the PLA segments.	[9]
PCL-PEG multiblock	Prepared by lipase-catalyzed reaction. There is little change in the enzymic degradation compared to PCL.	[10]
PCL-PLLA di- and multiblock	Prepared by macromonomer polymerization with a tin alkoxide catalyst. Tensile properties 1 kPa–20 MPa (strength); 600 kPa–221 MPa (modulus); 510–740 (% elongation at break). Triblock exhibits elastic properties and reasonable degradation over 150 days in phosphate-buffered saline	[11]
PAA-g-PCL graft copolymer	Prepared by radical-initiated macromonomer polymerization. Tensile properties 8–19 MPa (strength); 200–550 MPa (modulus). The polymers showed an increase in both modulus and strength with increasing graft length	[12]
PCL-ethyloxazoline multiblock	Prepared by dimethylaminopyridine-catalyzed macromonomer polymerization. Maximum strength 12 MPa, maximum elongation to break 1000%. The hydrogels exhibit excellent reversible swelling properties	[13]
PCL-hexamethyleneterephthalate multiblock	Prepared by a toluene diisocyanate macromonomer-coupling reaction. Produces a biphasic system with highly crystalline terephthalate regions	[14]
PCL-PLA multiblock	Prepared by high temperature transesterification of homopolymers. CAC dichloride is used to enable photocuring of the material. Modulus 50–270 MPa. Excellent shape memory properties are observed. The polymer shows 90% enzymic degradation over 27 days	[15]
PCL-PLA-PCL triblock	Prepared by anionic polymerization-catalyzed by calcium amide terminated with PEG. Exhibits a biphasic system, which may form micelles	[16]
PEG-PCL-PEG triblock	Prepared by macromonomer coupling and tin alkoxide polymerization and diisocyanate coupling. Maximum modulus 1.4 MPa, maximum viscosity 4.5 KPa. Degrades very slowly, showing only marginal molecular weight loss over 1 yr	[17]

injected into a mold containing the reinforcing fibers. The mould is then heated to the reaction temperature to promote the polymerization reaction. This approach resulted in good fiber mesh penetration and good composite properties.^[30]

Enzymatic Degradation

As discussed above, enzymes can be utilized to polymerize caprolactone and copolymers. Therefore there

exists the possibility of enzymatic degradation. Lipases have been used for such polymerization reactions and have therefore been used to study the degradation of PCLs and various polymer blends^[31] Two lipases have been studied, Novozyme 435 and Candida Rugosa, on PCL and poly (vinyl acetate) (PVAc) blends. Both the homopolymers were found to undergo chain scission, and the rate coefficient for specific chain scission of PCL was a function of the PVAc concentration. Enzyme degradation occurred over a period of 20 days at which time the enzymes became inactive.^[31]

Table 2 Mechanical properties of some common polymers compared to PCL

Polymer	Modulus (MPa)	Strength (MPa)
Polymethyl methacrylate	2500–3500	55–70
Poly(propene)	1100–1500	35
PLA	120–200	45–65
<i>Poly ε-caprolactone</i>	200–400	40–50

Macrophage Responses

Macrophage responses in vitro are vital to biomaterials research to enable prediction of a potential immune or inflammatory response in vivo. After implantation of a material, the macrophage and the neutrophil are the predominant cells in the resulting inflammatory response.



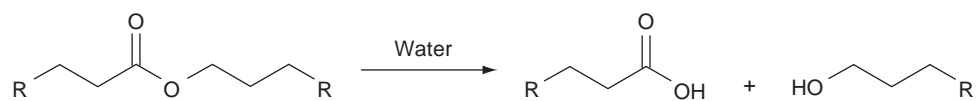


Fig. 2 Hydrolysis of PCL by water, resulting in an acid and an alcohol. (R indicates the rest of the polymer chain).

The murine BALB/c monocyte macrophage cell line J774A.1 has been investigated in response to PCL.^[32] No activation of the macrophages was observed as determined by interleukin 1- β and hydrogen peroxide release. Proliferation rate was slightly decreased, which may be beneficial as it is thought that macrophages arrive at their target site not only by chemotaxis but also by cell division.^[33] It has been demonstrated that macrophage proliferation as well as recruitment occurs in vivo around Dacron and expanded PTFE implants in Sprague Dawley rats,^[34] suggesting that proliferation could be involved in perpetuation of the chronic inflammatory response.

PCL AS A COMPOSITE

Biodegradable composites are a potential replacement of commonly used metallic implants. Such composites often consist of a polymer matrix with ceramic particles or fibers as reinforcing agents to strengthen the mechanical properties, and some of these are summarized in Table 3.^[30,35–40] There is, however, concern over poor adhesion between the fibers and the matrix of such composites.

Giang et al.^[41] have produced a poly(ϵ -caprolactone)/continuous bioglass fiber composite via the monomer transfer molding technique. The bioglass fibers were surface treated with an amine-ended silane to initiate polymerization of ϵ -caprolactone from the

fiber surface. This was found to improve the Young's modulus and flexural strength of the composite as well as increase water resistance.

The advantages of many composites for tissue engineering are that often a bioceramic is used within a degradable polymer matrix. The degradable polymers including PCL often have acidic degradation products, and so the bioceramic components often have a neutralizing effect owing to their dissolution products.

PCL composites have been prepared with phosphate glasses and have been shown to have neutral degradation products. The composition (Na_2O :0.55, CaO :variable, and P_2O_5 :0.45) of these composites was also shown to affect the degradation rate, therefore allowing tailored degradation rate for specific applications.^[42] PCL/hydroxyapatite composites have also been prepared,^[37] as have composites of poly(ϵ -caprolactone-co-DL-lactide) and bioactive glass (S53P4).^[43]

Composites of PCL with natural materials have been a focus of research over the last few years, for example, with collagen,^[44] silk fibroin,^[45] starch,^[46] and dextran.^[47] Inclusion of natural materials has a variety of advantages—the material becomes biomimetic, mimicking natural tissue if materials such as collagen are used and also will increase degradation potential by natural enzymes of the body such as collagenases. Composites with natural materials are therefore a useful addition to the range of tissue-engineering scaffolds.

Table 3 Properties of some composites containing PCL and reinforcing agents

Reinforcement	Notes	References
Silica: Nanofibers	Chemical grafting using a silane. 2.5% gives 330 MPa modulus	[35]
Chitin: 7 μm fibers	Prepared by melt blending. 45% gives 1.8 GPa modulus, 35 MPa flexural strength	[36]
Hydroxyl apatite: aggregated particles	Prepared by solution mixing and grafted by hydrolysis. An overall decrease in strength and modulus were observed	[37]
Hydroxyl apatite: particles	Prepared by solution mixing tensile modulus 500 MPa at 50%	[38]
Clay: nanoplatelets	Prepared by melt mixing. 10% reinforcement gave 400 MPa modulus, 17 MPa tensile strength	[39]
Polypropylene: in situ reinforcement	Prepared by melt flow mixing. Linear relationship between modulus and PP content 350–1000 MPa similarly for tensile strength 18–25 MPa	[40]
Silica: long fibers	Prepared by a material deposition method. Modulus 12 GPa, strength 168 MPa at 45%	[30]

Composites containing PCL have been investigated for many applications and will be discussed in relation to specific applications below.

APPLICATIONS

PCL is used and currently investigated for a number of applications in the body, mainly drug delivery and as a material for tissue repair. The applications are summarized in Table 4.

Drug Delivery

Controlled release devices work in several ways: via diffusion through a semipermeable polymer membrane, via the use of a porous polymer, and via biodegradation.

Some researchers are now working on self-regulated and responsive systems, which release drugs in response to biological need or stimulus, instead of a predetermined, unchangeable rate based on the polymer-release characteristics. These more recent systems differ from traditional devices as they can target specific tissues and regulate the supply of the active agent in proportion to the biological need, often triggered by a biological stimulus.

Owing to its slow degradation rate, PCL is ideally suited to long-term delivery over a period of more than one year. This has led to applications in various drug delivery systems as microspheres and nanospheres. Such micro-/nanospheres can be used as an injectable delivery system. A variety of drugs have been encapsulated in PCL for targeted drug delivery and for controlled drug release either in PCL alone or as copolymers.^[48]

An example: capronor

In 1979, the Research Triangle Institute patented the contraceptive implant, Capronor. It is an inch-long tube composed of PCL, which releases a steady dose of levonorgestrel (a synthetic progestogen) over 12–18 mo. Capronor has become favorable over its rival Norplant (based on poly(dimethylsiloxane), because it requires fewer implants, is biodegradable, is easy to insert, is less obvious under the skin, and can be formed via a variety of polymerization reactions.

Other examples of PCL as a drug delivery polymer

PCL nanoparticles^[49] and microspheres^[50] have been investigated for the delivery of insulin. Both appear to release insulin over prolonged periods of time, a characteristic necessary for the effective control of blood sugar levels.

PCL has also been developed as an anticancer drug delivery system in the form of methoxy PEG/poly(ϵ -caprolactone) amphiphilic block copolymeric nanospheres for tumor-specific folate-mediated targeting,^[51] nanoscopic micelles comprising block copolymers of poly(ϵ -caprolactone) and PEG for doxorubicin delivery,^[52] nanospheres for tamoxifen administration,^[53] and PCL nanospheres loaded with the pure antiestrogen RU 58668, a promising estrogen-dependent anticancer agent.^[54]

Other examples of PCL in drug delivery include the release of 5-fluorouracil-loaded PCL microspheres,^[55] ovalbumin (hydroxyl-functionalized glycerol poly(ϵ -caprolactone)-based microspheres and poly(*N*-isopropylacrylamide) (PNIPAAm) hydrogel),^[56] and



Table 4 Summary of the many applications of PCL and its composites/copolymers

Application	Composition of PCL, e.g., composite and copolymer	Form, e.g., nanospheres and fibrous	References
Drug delivery	PCL homopolymer, block copolymers of PCL and PEG hydroxyl-functionalized glycerol PCL, PNIPAAm hydrogel), PCL/PLGA	Microspheres, nanospheres, nanoscopic micelles, fibers	[48–60]
Skin	Collagen/PCL, PLGA/PCL, PCL homopolymer	Fibrous mat, knitted mesh, films	[61–64]
Nerve	PCL with LIF	Matrix	[65,66]
Cartilage	PCL/polyurethane, PCL/PLA	Porous, fibrous	[67–70]
Bone/bone-cartilage interface	PCL/bioactive glass, PCL/hydroxyapatite, PCL/tricalcium phosphates, PCL/chitosan	Matrix	[71–81]
Liver	Peptide-functionalized PCL/PEG, PCL/PGA	Nanofibers, fibrous mesh	[82–84]
Heart	PCL homopolymer	Nanofibrous mesh	[85]
Gene delivery	PCL/PEG, PCL homopolymer	Micelles, capsules	[87,88]

microspheres of PCL/poly(lactide-co-glycolide) (PLGA) loaded with contraceptive steroid drugs^[57] for the delivery of the lipophilic drugs levonorgestrel and ethinyl-estradiol. The hydrophilic macromolecule ovalbumin and the lipophilic drug progesterone have also been incorporated into PCL in the form of fibers produced by gravity spinning.^[58] The release behavior may be ideal for specific release of growth factors that are influential at specific time points in, for example, a wound-healing process.

PCL has also been studied as nanospheres coated with bioadhesive hyaluronic acid for ocular delivery,^[59] which again requires long-term release.

Sustained ex vivo skin antiseptic activity of chlorhexidine has been reported in PCL nanocapsules against *Staphylococcus epidermidis*.^[60] The capsules were 200–300 nm in diameter, and sustained release was demonstrated for up to 8 hr. The bactericidal activity was more effective compared to a solution of chlorhexidine, and a significant reduction in the number of colony forming units (CFU) (5 to $>1 \log_{10}$) was observed with chlorhexidine nanocapsules for 8 hr.

Tissue Engineering

Tissue engineering is a broad multidisciplinary subject aiming to regenerate damaged or diseased tissues via porous scaffold templates to direct new tissue formation. The production of porous PCL templates has therefore received a great deal of attention from researchers worldwide for a wide variety of tissue engineering applications owing to food and drug administration (FDA) approval and the need for cellular ingrowth into tissue-engineering scaffolds.

Porosity (in the range of $10 \mu\text{m}$) can be achieved by particulate leaching of salts.^[89] The process of electrospinning has gained renewed attention over the last five years and nano-to-microstructured biodegradable poly(L-lactide-co- ϵ -caprolactone) fabrics have been prepared^[90,91] as well as pure PCL electrospun fibers^[92,93] Supercritical CO_2 can also be utilized to produce porous scaffolds.^[94,95]

Freeform fabrication has also become a highly useful technique to produce patient-specific implants for tissue repair.^[96] This allows the production of specific geometries and architectures.

Skin

PCL has shown promise for skin tissue engineering. Skin was the first organ/tissue to be repaired using the tissue-engineering approach. However, many tissue-engineered skin equivalents have not worked well owing to the requirement for an active dermis and epidermis.

Composites of Collagen and PCL for Tissue Engineering of the Skin.^[61] Fibrous collagen mats were impregnated with PCL solutions, and after solvent evaporation were tested for both fibroblast and keratinocyte growth and compatibility. A 1:20 ratio of collagen:PCL was found to be optimal.

Matrices for dermal repair using fibroblasts grown on knitted poly(lactic-co-glycolic acid) (10:90)-poly(ϵ -caprolactone) (PLGA-PCL) mesh have been studied.^[62] PLGA-PCL mesh supported cell attachment, homogenous cell distribution, proliferation, and neo-tissue formation.

Khor et al.^[63] have investigated PCL as an epidermal equivalent. Human keratinocytes were cultured on solvent cast and biaxially stretched PCL films, which were found to promote good keratinocyte responses (proliferation to monolayers of cobblestone appearance) on 5–15 μm thick films, which is thought would make a better epidermal substrate owing to the better flexibility and strength to mass ratio compared to solvent cast films.

In another study,^[64] it was found that 5 μm thickness (via biaxial stretching) films have potential for skin tissue engineering when using dermal fibroblasts. Proliferation on the PCL films was superior to that on the polyurethane (PU) samples, therefore showing potential to be applied in tissue engineering a dermal equivalent or skin graft.

Nerve

Nerve repair via guided tissue regeneration based on fibrous or tubular tissue engineering materials is a relatively new area of research as an alternative to donor nerve graft harvesting.

PCL/ poly-DL-lactic acid (PDLLA) is one of several materials that shows promise for such applications,^[65] in the form of tubular structures with an internal diameter of 1.7 mm. These nerve guides were found to retain their mechanical properties and flexibility over the study period of 8–10 weeks, and mass loss was evident thereafter. Galla et al.^[66] developed a three-dimensional matrix, consisting of fibrin loaded with Schwann cells, with or without the neurotrophic factor leukemia inhibitory factor (LIF), and this matrix was used to fill tubular PCL conduits. This system was found to repair a 10 mm nerve gap in the buccal branch of a rat facial nerve.

Cartilage

Cartilage has a limited capacity for self-repair owing to its avascular nature. Therefore, a tissue-engineering approach may be ideal as a repair and regeneration strategy.

Longitudinal lesions in menisci are one of the most common surgical problems of the knee, in particular, the avascular part of the meniscus. PU with PCL, 1,4-butanediisocyanate, and 1,4-butanediol has been investigated in the form of a porous scaffold produced by liquid–liquid phase separation and salt leaching. Biocompatibility and meniscus-like tissue formation were shown in an animal model, and compression behavior was comparable to normal meniscus tissue.^[67] PCL in the form of fibrous scaffolds has also been investigated for cartilage applications,^[68] and with poly(L-lactide).^[69]

As with many tissue types, addition of growth factors to scaffolds can aid cell recruitment and differentiation. Huang, Hutmacher, and Lee^[70] demonstrated *in vivo* mesenchymal cell recruitment by a PCL scaffold loaded with transforming growth factor- β 1 (TGF- β 1). Chondrogenesis occurred after subperiosteal implantation.

Bone/cartilage interface

Owing to failure of integration of many materials developed for cartilage repair, the bone/cartilage interface may be a more successful way of regenerating cartilage using a tissue-engineering approach. This would also address the repair of osteochondral defects. An injectable bioactive glass (S53P4)/biodegradable polymer composite has been successful in an *in vivo* rabbit model for osteoconductivity and articular cartilage regeneration^[71] via colonization of chondroblasts and mature chondrocytes. The composite was injected as a viscous liquid or a moldable paste. These materials were injected at 47–50°C and then underwent solidification during cooling owing to the crystallization of the polymer matrix. Cao, Ho, and Teoh^[72] have also demonstrated success in engineering a bone/cartilage interface by using a three-dimensional PCL scaffold fabricated by fused deposition modeling. Fibers were oriented at 0, 60, and 120° in layers, resulting in a highly porous structure. Coculture of osteogenic cells in one-half of the scaffold and chondrogenic cells in the other half resulted in proliferation, migration, and integration at the interface.

Bone

Current surgical methods for the treatment of bone defects involve the use of autografts and allografts, which can cause donor-site morbidity and requires more than one operation. Recently tissue-engineered bone using a variety of materials has been generated.

A great deal of interest has focused on the use of PCL in bone defects owing to the search for a degradable bone repair material. To promote bone bonding, many composites of PCL with various bioceramics have been investigated.

PCL composite scaffolds for bone engineering have been developed with hydroxyapatite and tricalcium phosphates,^[73–75] novel hydroxyapatite/PCL composites with fluoridation of the hydroxyapatite,^[76] and tricalcium phosphates with recombinant human bone morphogenetic protein 2 (rhBMP-2) to stimulate bone formation^[77] with organosiloxanes, and^[78,79] PCL–chitosan matrices containing platelet derived growth factor-BB (PDGF-BB) were found to induce bone formation in a rat calvarial defect.^[80] In all studies, inclusion of the bioceramic component enhances bone cell responses and bone formation.

The potential of using PCL synthesized by a novel method, involving boron trifluoride dimethyl etherate polymerization and a potassium fluoride post-treatment, has been investigated for craniofacial bone repair.^[81] The method of production of PCL enabled control of the molecular weight and degradation time. Craniofacial osteoblast responses showed a preference for KF-treated PCL, where greater cell attachment and spreading were observed as well as synthesis of a collagen-I-rich-mineralized extracellular matrix.

Liver

Although liver has a capacity for regeneration, decreased hepatocyte adhesion to polymer materials for tissue engineering limits the function of tissue-engineered hepatic assist devices. To enhance hepatocyte adhesion, Carlisle et al.^[82] grafted RGD and YIGSR adhesion peptides to PCL in an attempt to mimic the extracellular matrix. Peptides were grafted using PEG–peptide links to allylamines coated on the PCL surface using pulsed plasma deposition. Adhesion on PCL surfaces was significantly enhanced by both YIGSR and RGD.

Primary rat hepatocytes were found to self-assemble into multicellular spheroids and maintain differentiated functions (ammonia metabolism, albumin secretion, and cytochrome P450 enzymatic activity) when cultured on PCL-co-ethyl ethylene phosphate nanofiber scaffold conjugated with galactose ligand.^[83]

A number of liver diseases target the biliary tree, and therefore bile duct tissue engineering has been investigated. Human biliary epithelial cells have been cultured on PGA/PCL fiber mesh scaffolds. Cells formed spheroidal aggregates, proliferated, and expressed long-term phenotypic stability for up to six months.^[84]

Heart

Heart tissue is not capable of regeneration after damage because of the terminally differentiated state of cardiomyocytes. Shin et al.^[85] have developed an *in vitro* system for engineered myocardium using electrospun nanofibrous meshes. Cardiac-specific proteins such

as α -myosin heavy chain, connexin43, and cardiac troponin were found to be successfully expressed. The results suggest that cardiac grafts can be matured in vitro to obtain sufficient function to enable implantation.

Gene Delivery/Therapy

The insertion of a functioning gene into the cells of a patient can correct single-gene defects such as errors of metabolism, and this process is known as gene therapy.^[86] Tissue-engineering scaffolds could be used as a source of gene transfer or as a source of genetically engineered cells. Ideally, autologous cells are used for all tissue-engineering applications; however, non-autologous can be used if immunoisolation techniques are performed. Novel biodegradable ternary copolymers hy-poly(ethylene imine) (PEI)-g-PCL-b-PEG have been developed for non-viral gene delivery.^[87]

Depending on their compositions, these polymers were completely soluble in water or form micelles. Complexation of plasmid DNA with various copolymers was investigated, and copolymer composition was found to affect the gene transfection efficiency. The ternary copolymers were found to have potential for gene delivery owing to their biocompatibility, biodegradability, and relatively high gene transfection efficiency, and their neutral surface charge allows potential intravenous delivery.

Chen et al.^[88] developed immunoisolation device based on PCL capsules to encapsulate human atrial natriuretic peptide-producing Chinese hamster ovary cells for implantation into hypertensive rats. These show promise in the study and treatment of renovascular hypertension and cardiovascular diseases.

CONCLUSIONS

PCL is a very versatile material and can also be prepared as a range of copolymers. However, its major role in the human body so far has been in controlled release devices. Over the last few years, extensive research has been performed to investigate the potential of PCL as a tissue-regeneration scaffold. The main tissue targets that have been investigated are bone, cartilage, skin, nerve, and liver.

ARTICLES OF FURTHER INTEREST

Biodegradable Polymers, An Overview; Biofunctional Polymers; Biorubber/Poly(Glycerol Sebacate) Polymers

REFERENCES

- Okamoto, Y. Cationic ring opening polymerisation of lactones in the presence of alcohol. *Makromol. Chem. Macromol. Symp.* **1991**, 42–3, 117–133.
- Morton, M.; Meiyuan, W. Anionic polymerization. *ACS. Symp. Ser.* **1985**, 285, 51–67.
- Feng, X.D.; Song, C.X.; Chen, W.Y.; Voong, S.T. Synthesis and evaluation of biodegradable block co-polymers of epsilon-caprolactone and DL-lactide. *J. Polym. Sci. C: Polym. Lett.* **1983**, 21, 593–600.
- Bailey, W.J.; Ni, Z.; Wu, S-R. Synthesis of poly- ϵ -caprolactone via a free radical mechanism-free radical opening polymerization of 2-methylene-1,3-dioxepane. *J. Polym. Sci. A: Polym. Chem.* **1982**, 20, 3021–3030.
- van Natta, F.J.; Hill, J.W.; Carothers, W.H. Studies of polymerization and ring formation. XXIII.1 ϵ -Caprolactone and its polymers. *J. Am. Chem. Soc.* **1934**, 56, 455.
- Pachence, J.M.; Kohn, J. Biodegradable polymers. In *Principles of Tissue Engineering*, 2nd Ed.; Academic Press: San Diego, USA, 2000; 263–277.
- Bei, J.-Z.; Li, J.-M.; Wang, Z.-F.; Le, J.-C.; Wang, S.-G. Polycaprolactone-Poly(ethylene-glycol) block copolymer. IV: biodegradation behavior in vitro and in vivo. *Polym. Adv. Technol.* **1997**, 8, 693–696.
- Bei, J.-Z.; He, W.; Hu, X.; Wang, S. Photodegradation behavior and mechanism of block copoly(caprolactone-ethylene glycol). *Polymer Degradation and Stability* **2000**, 67, 375–380.
- Huang, M-H.; Li, S.; Vert, M. Synthesis and degradation of PLA-PCL-PLA triblock copolymer prepared by successive polymerization of 3-caprolactone and DL-lactide. *Polymer* **2004**, 45, 8675–8681.
- He, F.; Li, S.M.; Vert, M.; Zhuo, R.X. Enzyme-catalyzed polymerization and degradation of copolymers prepared from epsilon-caprolactone and poly(ethylene glycol). *Polymer* **2003**, 44 (18), 5145–5151.
- Jeon, O.; Lee, S.-H.; Kim, S.H.; Lee, Y.M.; Kim, Y.H. Synthesis and characterization of Poly(L-lactide)-Poly(ϵ -caprolactone) multiblock copolymers. *Macromolecules* **2003**, 36, 5585–5592.
- Kim, C.-H.; Cho, K.Y.; Park, J.K. Effect of poly(acrylic acid)-g-PCL microstructure on the mechanical properties of starch/PCL blend compatibilized with poly(acrylic acid)-g-PCL. *Polym. Eng. Sci.* **2001**, 41, 542.
- Lee, S.C.; Kang, S.W.; Kim, C.; Kwon, I.C.; Jeong, S.Y. Synthesis and characterization of amphiphilic poly(2-ethyl-2-oxazoline)/poly(1-caprolactone) alternating multiblock copolymers. *Polymer* **2000**, 41, 7091–7097.
- Lefevre, C.; Villers, D.; Koch, M.H.J.; David, C. Synthesis and characterization of a crystallizable polycaprolactone/polyhexamethylene terephthalate block copolymer. *Polymer* **2001**, 42, 8769–8777.
- Nagata, M.; Sato, Y. Synthesis and properties of photocurable biodegradable multiblock copolymers based on poly(ϵ -caprolactone) and poly(L-lactide) segments. *J. Polym. Sci.* **2005**, 43, 2426–2439.
- Piao, L.; Dai, Z.; Deng, M.; Chen, X.; Jing, X. *Polymer*, **2003**, 44, 2025–2031.

17. Sosnik, A.; Chon, D. Poly(ethylene glycol)-poly(epsilon-caprolactone) block oligomers as injectable materials. *Polymer* **2003**, *44*, 7033–7042
18. Pitt, C.G.; Chasalow, F.I.; Hibionada, Y.M.; Klimas, D.M.; Schindler, A. Aliphatic polyesters 1. The degradation of poly(epsilon-caprolactone) in vivo. *J. Appl. Polym. Sci.* **1981**, *26*, 3779–3787.
19. Pitt, C.G.; Gratzl, M.M.; Kimmel, G.L.; Surles, J.; Schindler, A. Aliphatic polyesters 2. The degradation of poly(DL-lactide), poly(epsilon-caprolactone) and their copolymers in vivo. *Biomaterials* **1981**, *2* (4), 215–220.
20. Pitt, C.G.; Marks, T.A.; Schindler, A. *Controlled Release of Bioactive Materials*; Baker, R., Ed.; Academic Press: New York, 1980.
21. Pitt, C.G. Poly-epsilon-caprolactone and its copolymers. In *Biodegradable Polymers as Drug Delivery Systems*; Chasin, M., Langer, R., Eds.; Dekker: New York, 1990; 71–119.
22. Albertsson, A.C.; Varma, I.K. Recent developments in ring opening polymerization of lactones for biomedical applications. *Biomacromolecules* **2003**, *4* (6), 1466–1486.
23. Wahlberg, J.; Persson, P.V.; Olsson, T.; Hedenstrom, E.; Iversen, T. Structural characterization of a lipase-catalyzed copolymerization of epsilon-caprolactone and D,L-lactide. *Biomacromolecules* **2003**, *4* (4), 1068–1071.
24. Varma, I.K.; Albertsson, A.C.; Rajkhowa, R.; Srivastava, R.K. Enzyme catalyzed synthesis of polyesters. *Prog. Polym. Sci.* **2005**, *30* (10), 949–981.
25. Qian, H.T.; Bei, J.Z.; Wang, S.G. Synthesis, characterisation and degradation of ABA block copolymer of L-lactide and epsilon caprolactone. *Polym. Degrad. Stab.* **2000**, *68*, 423–429.
26. Domb, A.J. Degradable polymer blends 1. Screening of miscible polymers. *J. Polym. Sci. [A1]* **1993**, *31*, 1973–1981.
27. Smith, K.L.; Schimpf, M.E.; Thompson, K.E. Bioerodible polymers for delivery of macromolecules. *Adv. Drug. Deliv. Res.* **1990**, *4*, 343–357.
28. Christian, P.; Jones, I.A. Polymerisation and stabilisation of polycaprolactone using a Borontrifluoride/glycerol catalyst system. *Polymer* **2001**, *42*, 3989.
29. Corden, T.J.; Jones, I.A.; Rudd, C.D.; Christian, P.; Downes, S.; McDougall, K.E. Physical and biocompatibility properties of poly-epsilon-caprolactone produced using in situ polymerisation: a novel manufacturing technique for long-fibre composite materials. *Biomaterials* **2000**, *21*, 713.
30. Christian, P.; Jones, I.A.; Rudd, C.D.; Cambell, R.I.; Corden, T.J. Monomer transfer molding and rapid prototyping methods for fiber reinforced thermoplastics for biomedical applications. *Composites Part A* **2001**, *32*, 969.
31. Sivalingam, G.; Chattopadhyay, S.; Madras, G. Enzymatic degradation of poly(epsilon-caprolactone), poly(vinyl acetate) and their blends by lipases. *Chem. Eng. Sci.* **2003**, *58* (13), 2911–2919.
32. Gough, J.E.; Christian, P.; Unsworth, J.; Evans, M.P.; Scotchford, C.A.; Jones, I.A. Controlled degradation and macrophage responses of a fluoride-treated polycaprolactone. *J. Biomed. Mater. Res.* **2004**, *69A*, 17–25.
33. Ryan, G.B.; Spector, W.G. Macrophage turnover in inflamed connective tissue. *Proc. R. Soc. Lond. Ser. B.* **1970**, *175*, 269–292.
34. Hagerty, R.D.; Salzman, D.L.; Kleinert, L.B.; Williams, S.K. Cellular proliferation and macrophage populations associated with implanted expanded polytetrafluoroethylene and polyethyleneterephthalate. *J. Biomed. Mater. Res.* **2000**, *49*, 489–497.
35. Avella, M.; Bondioli, F.; Cannillo, V.; Errico, M.E.; Ferrari, A.M.; Focjer, B.; Malinconico, M.; Manfredini, T.; Montorsi, M. Preparation, characterisation and computational study of poly(epsilon-caprolactone) based nanocomposites. *Mater. Sci. Technol.* **2004**, *20*, 1340.
36. Chen, B.; Sun, K.; Ren, T. Mechanical and viscoelastic properties of chitin fiber reinforced poly(epsilon-caprolactone). *Eur. Polym. J.* **2005**, *41*, 453–457.
37. Choi, D.W.; Marra, K.G.; Kumta, P.N. Chemical synthesis of hydroxyapatite/poly(epsilon-caprolactone) composites. *Mater. Res. Bull.* **2004**, *39* (3), 417–432.
38. Hao, J.; Yuan, M.; Deng, X. Biodegradable and biocompatible nanocomposites of poly(caprolactone) with hydroxyapatite nanocrystals: thermal and mechanical properties. *J. Appl. Polym. Sci.* **2002**, *86*, 676–683.
39. Lepoittevin, B.; Devalckenaere, M.; Pantoustier, N.; Alexandre, M.; Kubies, D.; Calberg, C.; Jerome, R.; Dubois, P. Poly caprolactone/clay nanocomposite prepared by melt interchelation: mechanical, thermal and rheological properties. *Polymer* **2002**, *43*, 4017–4023.
40. Semba, T.; Kitagawa, K.; Endo, S.; Maeda, K.; Hamada, H. In situ fiber-reinforced composites from blends containing polypropylene and polycaprolactone. *J. Appl. Polym. Sci.* **2004**, *91*, 833–840.
41. Giang, G.; Evans, M.E.; Jones, I.A.; Rudd, C.D.; Scotchford, C.A.; Walker, G.S. Preparation of poly(epsilon-caprolactone)/continuous bioglass fibre composite using monomer transfer moulding for bone implant. *Biomaterials* **2005**, *26*, 2281–2288.
42. Prabhakar, R.L.; Brocchini, S.; Knowles, J.C. Effect of glass composition on the degradation properties and ion release characteristics of phosphate glass - polycaprolactone composites. *Biomaterials* **2005**, *26* (15), 2209–2218.
43. Jaakkola, T.; Rich, J.; Tirri, T.; Narhi, T.; Jokinen, M.; Seppala, J.; Yli-Urpo In vitro Ca-P precipitation on biodegradable thermoplastic composite of poly(epsilon-caprolactone-co-DL-lactide) and bioactive glass (S53P4). *Biomaterials* **2004**, *25* (4), 575–581.
44. Coombes, A.G.A.; Verderio, E.; Shaw, B.; Li, X.; Griffin, M.; Downes, S. Biocomposites of non-cross-linked natural and synthetic polymers. *Biomaterials* **2002**, *23* (10), 2113–2118.
45. Chen, G.; Zhou, P.; Mei, N.; Chen, X.; Shao, Z.Z.; Pan, L.F.; Wu, C.G. Silk fibroin modified porous poly(epsilon-caprolactone) scaffold for human fibroblast culture in vitro. *J. Mater. Sci.-Mater. Med.* **2004**, *15* (6), 671–677.
46. Yavuz, H.; Babac, C. Preparation and biodegradation of starch/polycaprolactone films. *J. Polym. Environ.* **2003**, *11* (3), 107–113.

47. Lemarchand, C.; Couvreur, P.; Besnard, M.; Costantini, D.; Gref, R. Novel polyester-polysaccharide nanoparticles. *Pharm. Res.* **2003**, *20* (8), 1284–1292.
48. Sinha, V.R.; Bansal, K.; Kaushik, R.; Kumria, R.; Trehan, A. Poly- ϵ -caprolactone microspheres and nanospheres: an overview. *Int. J. Pharm.* **2004**, *278* (1), 1–23.
49. Attivi, D.; Wehrle, P.; Ubrich, N.; Damge, C.; Hoffman, M.; Maincent, P. Formulation of insulin-loaded polymeric nanoparticles using response surface methodology. *Drug Dev. Ind. Pharm.* **2005**, *31* (2), 179–189.
50. Shenoy, D.B.; D'Souza, R.J.; Tiwari, S.B.; Udupa, N. Potential applications of polymeric microsphere suspension as subcutaneous depot for insulin. *Drug Dev. Ind. Pharm.* **2003**, *29* (5), 555–563.
51. Park, E.K.; Lee, S.B.; Lee, Y.M. Preparation and characterization of methoxy poly(ethylene glycol)/poly(ϵ -caprolactone) amphiphilic block copolymeric nanospheres for tumor-specific folate-mediated targeting of anticancer drugs. *Biomaterials* **2005**, *26* (9), 1053–1061.
52. Shuai, X.T.; Ai, H.; Nasongkla, N.; Kim, S.; Gao, J.M. Micellar carriers based on block copolymers of poly(ϵ -caprolactone) and poly(ethylene glycol) for doxorubicin delivery. *J. Control. Release* **2004**, *98* (3), 415–426.
53. Chawla, J.S.; Amiji, M.M. Cellular uptake and concentrations of tamoxifen upon administration in poly(ϵ -caprolactone) nanoparticles. *AAPS. Pharmsci.* **2003**, *5* (1), 1–7.
54. Ameller, T.; Marsaud, W.; Legrand, P.; Gref, R.; Renoir, J.M. Pure antiestrogen RU 58668-loaded nanospheres: morphology, cell activity and toxicity studies. *Eur. J. Pharm. Sci.* **2004**, *21* (2–3), 361–370.
55. Sastre, R.L.; Blanco, M.D.; Teijon, C.; Olmo, R.; Teijon, J.M. Preparation and characterization of 5-fluorouracil-loaded poly(ϵ -caprolactone) microspheres for drug administration. *Drug Dev. Res.* **2004**, *63* (2), 41–53.
56. Zhang, X.Z.; Lewis, P.J.; Chu, C.C. Fabrication and characterization of a smart drug delivery system: microsphere in hydrogel. *Biomaterials* **2005**, *26* (16), 3299–3309.
57. Dhanaraju, M.D.; Jayakumar, R.; Vamsadhara, C. Influence of manufacturing parameters on development of contraceptive steroid loaded injectable microspheres. *Chem. Pharm. Bull.* **2004**, *52* (8), 976–979.
58. Williamson, M.R.; Chang, H.I.; Coombes, A.G.A. Gravity spun polycaprolactone fibres: controlling release of a hydrophilic macromolecule (ovalbumin) and a lipophilic drug (progesterone). *Biomaterials* **2004**, *25* (20), 5053–5060.
59. Barbault-Foucher, S.; Gref, R.; Russo, P.; Guehot, J.; Bochot, A. Design of poly- ϵ caprolactone nanospheres coated with bioadhesive hyaluronic acid for ocular delivery. *J. Control. Release* **2002**, *83* (3), 365–375.
60. Lboutounne, H.; Chaulet, J.F.; Ploton, C.; Falson, F.; Pirot, F. Sustained ex vivo skin antiseptic activity of chlorhexidine in poly(ϵ -caprolactone) nanocapsule encapsulated form and as a digluconate. *J. Control. Release* **2002**, *82* (2–3), 319–334.
61. Dai, N.T.; Williamson, M.R.; Khammo, N.; Adams, E.F.; Coombes, A.G.A. Composite cell support membranes based on collagen and polycaprolactone for tissue engineering of skin. *Biomaterials* **2004**, *25* (18), 4263–4271.
62. Ng, K.W.; Khor, H.L.; Hutmacher, D.W. In vitro characterization of natural and synthetic dermal matrices cultured with human dermal fibroblasts. *Biomaterials* **2004**, *25* (14), 2807–2818.
63. Khor, H.L.; Ng, K.W.; Schantz, J.T.; Phan, T.T.; Lim, T.C.; Teoh, S.H.; Hutmacher, D.W. Poly(ϵ -caprolactone) films as a potential substrate for tissue engineering an epidermal equivalent. *Mater. Sci. Eng. C-Biomimetic Supramol. Syst.* **2002**, *20* (1–2), 71–75.
64. Woei, K.; Hutmacher, D.W.; Schantz, J.T.; Seng, C.; Too, H.P.; Chye, T.; Phan, T.T.; Teoh, S.H. Evaluation of ultra-thin poly(ϵ -caprolactone) films for tissue-engineered skin. *Tissue Eng.* **2001**, *7* (4), 441–455.
65. Meek, M.F.; Jansen, K.; Steendam, R.; van Oeveren; van Wachem, P.B.; van Luyn, M.J.A. In vitro degradation and biocompatibility of poly(DL-lactide- ϵ -caprolactone) nerve guides. *J. Biomed. Mater. Res. A* **2004**, *68A* (1), 43–51.
66. Galla, T.J.; Vedecnik, S.V.; Halbgewachs, J.; Steinmann, S.; Friedrich, C.; Stark, G.B. Fibrin/Schwann cell matrix in poly- ϵ -caprolactone conduits enhances guided nerve regeneration. *Int. J. Artif. Organs* **2004**, *27* (2), 127–136.
67. Heijkants, R.G.J.C.; van Calck, R.V.; De Groot, J.H.; Pennings, A.J.; Schouten, A.J.; van Tienen, T.G.; Ramrattan, N.; Buma, P.; Veth, R.P.H. Design, synthesis and properties of a degradable polyurethane scaffold for meniscus regeneration. *J. Mater. Sci. Mater. Med* **2004**, *15* (4), 423–427.
68. Li, W.J.; Danielson, K.G.; Alexander, P.G.; Tuan, R.S. Biological response of chondrocytes cultured in three-dimensional nanofibrous poly(ϵ -caprolactone) scaffolds. *J. Biomed. Mater. Res. A* **2003**, *67A* (4), 1105–1114.
69. Zhao, J.; Yuan, X.Y.; Cui, Y.L.; Ge, Q.B.; Yao, K.D. Preparation and characterization of poly(L-lactide)/poly(ϵ -caprolactone) fibrous scaffolds for cartilage tissue engineering. *J. Appl. Polym. Sci.* **2004**, *91* (3), 1676–1684.
70. Huang, Q.; Hutmacher, D.W.; Lee, E.H. In vivo mesenchymal cell recruitment by a scaffold loaded with transforming growth factor β 1 and the potential for in situ chondrogenesis. *Tissue Eng.* **2002**, *8* (3), 469–482.
71. Aho, A.J.; Tirri, T.; Kukkonen, J.; Strandberg, N.; Rich, J.; Seppala, J.; Yli-Urpo, A. Injectable bioactive glass/biodegradable polymer composite for bone and cartilage reconstruction: concept and experimental outcome with thermoplastic composites of poly(ϵ -caprolactone-co-D,L-lactide) and bioactive glass S53P4. *J. Mater. Sci.-Mater. Med.* **2004**, *15* (10), 1165–1173.
72. Cao, T.; Ho, K.H.; Teoh, S.H. Scaffold design and in vitro study of osteochondral coculture in a three-dimensional porous polycaprolactone scaffold fabricated by fused deposition modelling. *Tissue Eng.* **2003**, *9* (Suppl. 1), S103–S112.
73. Aydin, H.M.; Piskin, E.; Calimli, A. Microporous scaffolds from poly(lactide-co- ϵ -caprolactone) composites with hydroxyapatite and tricalcium phosphates

- using supercritical CO₂ for bone tissue engineering. *J. Bioact. Compat. Polym.* **2004**, *19* (5), 383–394.
74. Endres, M.; Hutmacher, D.W.; Salgado, A.J.; Kaps, C.; Ringe, J.; Reis, R.L.; Sittering, M.; Brandwood, A.; Schantz, J.T. Osteogenic induction of human bone marrow-derived mesenchymal progenitor cells in novel synthetic polymer-hydrogel matrices. *Tissue Eng.* **2003**, *9* (4), 689–702.
75. Looss, P.; Le Ray, A.M.; Grimandi, G.; Daculsi, G.; Merle, C. A new injectable bone substitute combining poly(ϵ -caprolactone) microparticles with biphasic calcium phosphate granules. *Biomaterials* **2001**, *22* (20), 2785–2794.
76. Kim, H.W.; Lee, E.J.; Kim, H.E.; Salih, V.; Knowles, J.C. Effect of fluoridation of hydroxyapatite in hydroxyapatite-polycaprolactone composites on osteoblast activity. *Biomaterials* **2005**, *26* (21), 4395–4404.
77. Rai, B.; Teoh, S.H.; Ho, K.H.; Hutmacher, D.W.; Cao, T.; Chen, F.; Yacob, K. The effect of rhBMP-2 on canine osteoblasts seeded onto 3D bioactive polycaprolactone scaffolds. *Biomaterials* **2004**, *25* (24), 5499–5506.
78. Rhee, S.H.; Lee, Y.K.; Lim, B.S. Evaluation of a novel poly(ϵ -caprolactone)-organosiloxane hybrid material for the potential application as a bioactive and degradable bone substitute. *Biomacromolecules* **2004**, *5* (4), 1575–1579.
79. Rhee, S.H. Bone-like apatite-forming ability and mechanical properties of poly(ϵ -aprolactone)/silica hybrid as a function of poly(ϵ -caprolactone) content. *Biomaterials* **2004**, *25* (7–8), 1167–1175.
80. Im, S.Y.; Cho, S.H.; Hwang, J.H.; Lee, S.J. Growth factor releasing porous poly(ϵ -caprolactone)-chitosan matrices for enhanced bone regenerative therapy. *Arch. Pharm. Res.* **2003**, *26* (1), 76–82.
81. Christian, P.; Jones, I.A. Polymerisation and stabilisation of polycaprolactone using a borontrifluoride–glycerol catalyst system. *Polymer* **2001**, *42* (9), 3989–3994.
82. Carlisle, E.S.; Mariappan, M.R.; Nelson, K.D.; Thomes, B.E.; Timmons, R.B.; Constantinescu, A.; Eberhart, R.C.; Bankey, P.E. Enhancing hepatocyte adhesion by pulsed plasma deposition and polyethylene glycol coupling. *Tissue Eng.* **2000**, *6* (1), 45–52.
83. Chua, K.N.; Lim, W.S.; Zhang, P.C.; Lu, H.F.; Wen, J.; Ramakrishna, S.; Leong, K.W.; Mao, H.Q. Stable immobilization of rat hepatocyte spheroids on galactosylated nanofiber scaffold. *Biomaterials* **2005**, *26* (15), 2537–2547.
84. Barralet, J.E.; Wallace, L.L.; Strain, A.J. Tissue engineering of human biliary epithelial cells on polyglycolic acid/polycaprolactone scaffolds maintains long-term phenotypic stability. *Tissue Eng.* **2003**, *9* (5), 1037–1045.
85. Shin, M.; Ishii, O.; Sueda, T.; Vacanti, J.P. Contractile cardiac grafts using a novel nanofibrous mesh. *Biomaterials* **2004**, *25* (17), 3717–3723.
86. Huard, J.; Fu, F.H. Eds. *Gene Therapy and Tissue Engineering in Orthopedic and Sports Medicine*; Birkhauser publishers: Boston, 2000.
87. Shuai, X.T.; Merdan, T.; Unger, F.; Wittmar, M.; Kissel, T. Novel biodegradable ternary copolymers hy-PEI-g-PCL-b-PEG: synthesis, characterization, and potential as efficient nonviral gene delivery vectors. *Macromolecules* **2003**, *36* (15), 5751–5759.
88. Chen, L.G.; Wang, Z.R.; Wan, C.M.; Chao, L.; Chao, J.; Xing, H.Y. Encapsulated transgene cells attenuate hypertension, cardiac hypertrophy and enhance renal function in Goldblatt hypertensive rats. *J. Gene. Med.* **2004**, *6* (7), 786–797.
89. Ciapetti, G.; Ambrosio, L.; Savarino, L.; Granchi, D.; Cenni, E.; Guizzardi, N.S.; Causa, F.; Giunti, A. Osteoblast growth and function in porous poly ϵ -caprolactone matrices for bone repair: a preliminary study. *Biomaterials* **2003**, *24*, 3815–3824.
90. Kwon, I.K.; Kidoaki, S.; Matsuda, T. Electrospun nano- to microfiber fabrics made of biodegradable copolyesters: structural characteristics, mechanical properties and cell adhesion potential. *Biomaterials* **2005**, *26* (18), 3929–3939.
91. Mo, X.M.; Xu, C.Y.; Kotaki, M.; Ramakrishna, S. Electrospun P(LLA-CL) nanofiber: a biomimetic extracellular matrix for smooth muscle cell and endothelial cell proliferation. *Biomaterials* **2004**, *25* (10), 1883–1890.
92. Hsu, C.M.; Shivkumar, S. Nano-sized beads and porous fiber constructs of poly(ϵ -caprolactone) produced by electrospinning. *J. Mater. Sci.* **2004**, *39* (9), 3003–3013.
93. Yoshimoto, H.; Shin, Y.M.; Terai, H.; Vacanti, J.P. A biodegradable nanofiber scaffold by electrospinning and its potential for bone tissue engineering. *Biomaterials* **2003**, *24* (12), 2077–2082.
94. Busby, A.J.; Zhang, J.X.; Roberts, C.J.; Lester, E.; Howdle, S.M. Novel nanostructured polymeric composites of polycaprolactone and ultra-high-molecular-weight polyethylene via a supercritical-fluid route. *Adv. Mater.* **2005**, *17* (3), 364.
95. Stassin, F.; Jerome, R. Contribution of supercritical CO₂ to the preparation of aliphatic polyesters and materials thereof. *Macromol Symposia* **2004**, *217*, 135–146.
96. Wang, F.; Shor, L.; Darling, A.; Khalil, S.; Sun, W.; Gucer, S.; Lau, A. Precision extruding deposition and characterization of cellular poly- ϵ -caprolactone tissue scaffolds. *Rapid Prototyping J.* **2004**, *10* (1), 42–49.



Poly(carbonate)urethanes

Michael Szycher

CT Biomaterials Division, CardioTech International, Inc., Woburn, Massachusetts, U.S.A.

INTRODUCTION

Polyurethanes have unique mechanical and biologic properties that make them ideal for many implantable devices. Polyurethanes display excellent mechanical properties such as high tensile strength, tear and abrasion resistance, and good blood compatibility. Accordingly, polyurethanes are widely used in medical implants such as cardiac pacemakers, catheters, cardiac assist devices, heart valves, and vascular grafts. The excellent mechanical properties of segmented polyurethanes are attributed to their two-phase morphology derived from microphase separation of soft and hard segments.

The current generation of commercially available polyurethanes is susceptible to certain in-vivo degradation mechanisms. For instance, poly(ester)urethanes are subject to hydrolytic degradation and are no longer used in long-term implanted devices. Poly(ether)urethanes, while hydrolytically stable, are subject to oxidative degradation in several forms, including environmental stress cracking (ESC) and metal ion oxidation (MIO). For these reasons, for long-term implantation, the poly(carbonate)urethanes are recommended.

POLY(URETHANE) POLYCARBONATES

The dictionary defines elastomer as a “material which at room temperature can be stretched repeatedly and upon immediate release will return to its approximate original length.” Since natural rubber was the original elastomer, in polymeric nomenclature, synthetic materials that approximate or exceed cured natural rubber in physical properties are called elastomeric. Because natural rubber exhibits such an excellent combination of physical properties, e.g., tensile strength of 4000 psi and 400 to 600% ultimate elongation, the term elastomeric grade is frequently used to characterize and describe those synthetic materials with the highest physical performance.

Among the highest performing biomedical-grade elastomers are the polyurethanes. Polyurethanes are block copolymers containing blocks of low molecular weight polyesters or polyethers linked together by a urethane group. These have the versatility of being either rigid, semirigid, or flexible. They have also

demonstrated excellent blood compatibility^[1] and outstanding hydrolytic stability,^[2] superior abrasion resistance, excellent physical strength, and high flexure endurance. No wonder the polyurethanes enjoy such great popularity among producers of medical prostheses. They bridge the gap among elastomers and elastoplastics, as shown in Fig. 1.

However, the use of the generic term polyurethane is deceiving in that all useful polyurethane polymers contain a minority of urethane functional groups. Thus, polyurethane is a term of convenience rather than accuracy since these polymers are not derived by polymerizing a monomeric urethane reactant, nor are they polymers containing primarily urethane linkages. In fact, in other groups such as ethers, carbonates are the majority linkages in the molecular chain. Although urethane linkages represent the minority of functional groups, as long as the polymers contain a significant number of urethane linkages, the name polyurethane is correctly ascribed to these polymers. Thus, the polyurethanes are a heterogeneous family of polymers, unlike PVC, polyethylene, and polystyrene. The polyurethanes comprise an array of different products, ranging from rigid foams to soft, millable grains.

Polyurethanes are based on three monomers—an isocyanate source, a macroglycol, and a chain extender or curative. As shown in Fig. 2, the isocyanates can be either aromatic or aliphatic and, as subsequent data will show, there is a profound difference between polymers made from either of these materials.

The macroglycols represent perhaps the single most important choice in the synthesis of medical-grade polyurethanes. The final physical and biological properties we associate with the polyurethanes depend for the most part on what type of macroglycol is used in the synthesis of medical-grade elastomers. Depending on the macroglycol used, the resultant polyurethanes can be either hydrophilic or hydrophobic. We will critically discuss these crucial monomers in greater detail subsequently.

Among chain extenders, there are two choices, either difunctional or multifunctional monomers. For the production of linear elastomers, only difunctional chain extenders are used. Of these, diols and diamines are by far the most important. Linear elastomers produce thermoplastic polyurethanes, whereas the diamines produce, thermosetting polyurethanes.

Bridges the Gap between Rubber and Plastics!

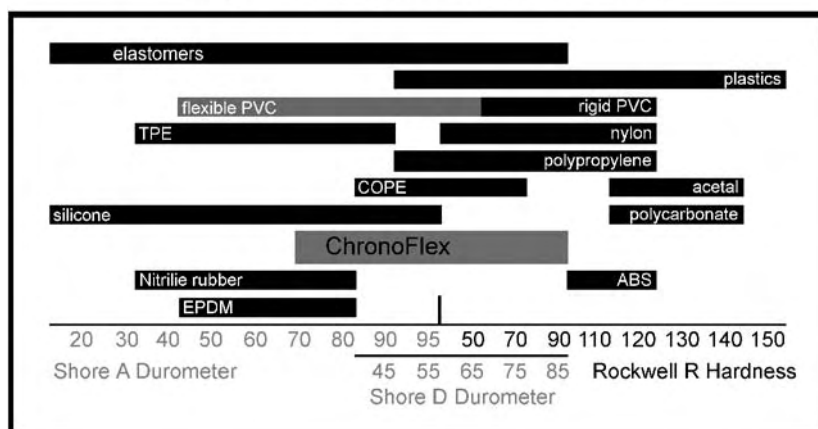


Fig. 1 ChronoFlex polyurethanes bridge the gap between rubber and elastoplastics because they can be synthesized in Durometer hardnesses ranging from a soft 75A to a hard 75D.

We will now discuss the three monomers that constitute the polyurethane elastomers, starting with the isocyanate sources. Table 1 presents the chemical structure of the prototype aromatic isocyanate example, methylene diphenyl diisocyanate, better known in the industry as MDI. MDI is a very quick-reacting isocyanate. In fact, it is so reactive that, in many cases, a catalyst is not required for synthesis if a diamine chain extender is used. On a practical basis, commercially available aromatic polyurethanes reacted with diols contain small amounts of a catalyst.

Thus, polyurethanes are a versatile family of polymers that comprise the widest array of properties, ranging from soft to hard, hydrophilic to hydrophobic, and thermoplastic to thermosets.

BIODURABILITY TESTING

Although poly(ether)urethanes are the materials of choice for a wide variety of medical implants, in

long-term applications, the poly(ether)urethanes degrade and cause malfunction or failure of the device. Degradation is usually apparent in terms of surface cracking, stiffening, erosion, or the deterioration of flexural strength. The mechanisms responsible for such degradation include environmental stress cracking, the generation of cracks and crazes produced by the biological environment acting on the polyurethane surface and metal ion oxidation, which is primarily restricted to pacemaker leads. It is generally accepted that ether linkages in the poly(ether)urethanes are the most vulnerable sites for initiation of degradation. Efforts have been made to overcome this problem by developing polyurethanes that are not exclusively based on polyether macroglycols, such as those disclosed in U.S. Patent 4,875,308 (Coury et al.), U.S. Patent 5,133,742 (Pinchuck), and U.S. Patent 5,109,077 (Wick). Nevertheless, the combination of degradation resistance, mechanical characteristics, processability, and clarity is suboptimal for certain applications.

Poly(carbonate) macrodiols have been used as reactive ingredients in the synthesis of block and segmented copolymer systems in high-performance polyurethane elastomers. Processes for preparing poly(carbonate) macrodiols on a range of bishydroxy alkylene compounds are disclosed in JP 62,241,920 (Toa Gosei Chemical Industry Co., Ltd.), JP 62,187,725

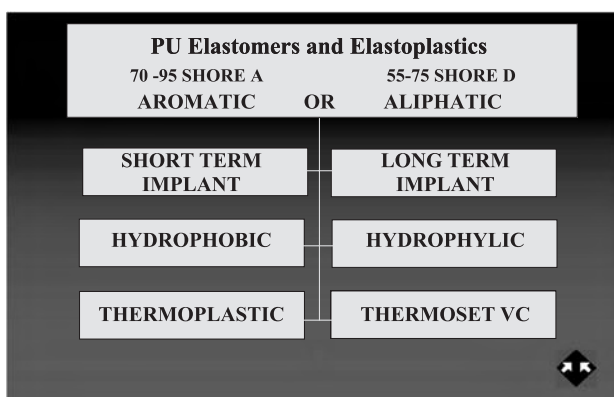


Fig. 2 The universe of polyurethanes, the most versatile polymer in the world.

Table 1 Urethane elastomer technology

Isocyanate	Macroglycol	Chain extender
Aliphatic	Poly(ester)	Diol
Aromatic	Poly(ether)	Diamines
	Poly(carbonate)	
	Silicone diols	
	Silicone amines	



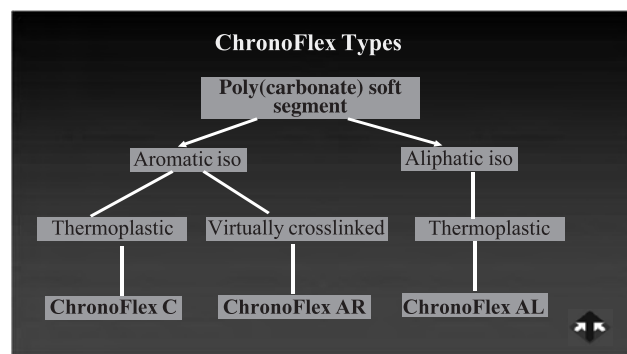


Fig. 3 CT Biomaterials produces three types of ChronoFlex biodurable polymers: 1) ChronoFlex C; 2) ChronoFlex AR; and 3) ChronoFlex AL. This figure shows the chemical constituents of these three types.

(Daicel Chemical Industries, Ltd.), DE 3,717,060 (Bayer AG), U.S. 4,131,731 (Beatrice Foods Company), and U.S. 5,171,830 (Arco Chemical Technology).

In U.S. Patents 5,254,662 and 5,863,627, Szycher disclosed biodurable, i.e., hydrolytically stable and enzymatically (biologically) stable polyurethane elastomers, in combination with a low modulus of elasticity and other desirable mechanical properties. The poly(carbonate)urethanes are prepared from the reaction of cycloaliphatic or aromatic diisocyanates with poly(carbonate) macrodiols extended with diols or diamines. The chemical composition of the ChronoFlex polymers is shown in Fig. 3.

ChronoFlex has been shown to be biodurable^[3] when tested according to the Stokes method.^[4] This method is an accelerated test designed to predict the long-term stability of polyurethanes in the corrosive environment of the human body.

Szycher was the first to discover the connection between polyether-based polyurethanes and biologically induced environmental stress cracking.^[5] This basic discovery led to the development of a poly(carbonate)urethane,^[6] which is not only biodurable but also displays a low modulus of elasticity. A low

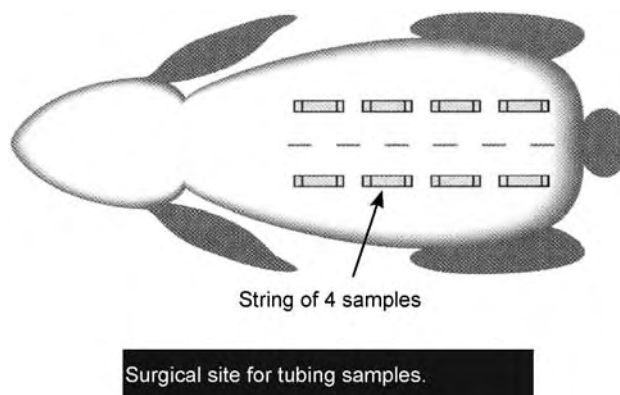


Fig. 4 Eight samples are implanted in rabbits for the accelerated test.

modulus of elasticity is an essential characteristic of polymers designed to interface with the cardiovascular system to enable the prostheses to have compliance with the host artery.

In this accelerated test, a polyurethane test tube is placed over a mandrel and stretched to different levels. The assembly of the stretched polymer and the mandrel is then implanted subcutaneously. Eight samples are implanted on either side of the vertebral column, as shown in Fig. 4.

Once an assembly of four test mandrels has been sterilized, they are implanted in experimental animals, e.g., rabbits. The mandrel dimensions are shown in Fig. 5. Each string of four samples is implanted subcutaneously between the natural plane that divides the hypodermis from the superficial muscular layer, as shown in Fig. 6.

After explantation, the mandrels are cleaned in Tergazyme™ solution for several days until all the biological tissue is gently removed. The specimens are then evaluated under scanning electron microscopy conditions to ascertain any surface cracking. Figures 7, 8, and 9 are microscans showing the results of implantation of three polyurethanes, Tecoflex, Pellethane, and ChronoFlex.

As shown, both Tecoflex and Pellethane have experienced catastrophic failure, secondary to surface

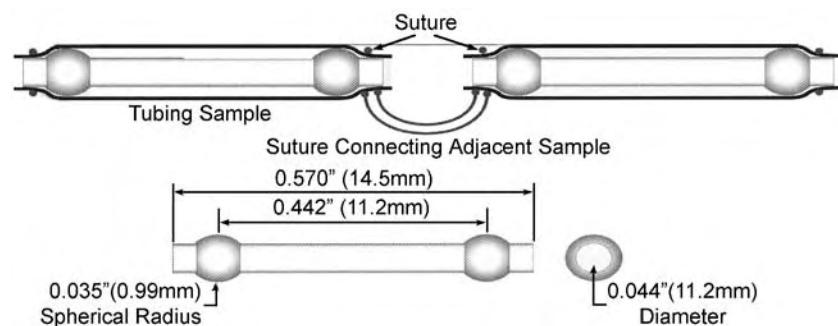


Fig. 5 Mandrel used to hold elongated tubing.

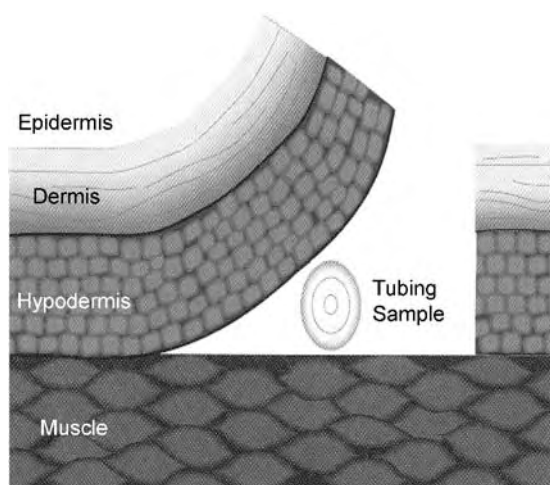


Fig. 6 Implantation of test specimen immediately below the hypodermis and above the muscle layer.

cracking. This is similar to the catastrophic failures experienced clinically when these poly(ether)urethanes are implanted for long periods of time. In contrast, the poly(carbonate)urethane ChronoFlex appears free of microcracks under the conditions of the test.

BIOCOMPATIBILITY TESTING

ChronoFlex polyurethanes have undergone rigorous in-vitro biocompatibility tests. While biocompatibility



Fig. 7 Catastrophic failure: TecoFlex EG-85A, an aliphatic poly(ether)urethane. After three months, the tubing is seen fractured. A small piece of tubing still hangs on the left side of the mandrel, maintained in place by the left suture.



Fig. 8 Pellethane 2363-80A. After three months, this aromatic poly(ether)urethane has undergone complete failure, secondary to microcracking. This is classical environmental-stress cracking (ESC), which is followed by fracture.

of polymers is highly application specific, ChronoFlex polymers have consistently performed well in contact with biological tissue and circulating blood.

Testing of ChronoFlex at independent laboratories has shown this family to be biocompatible when tested under USP Class VI standards, as shown in Table 2, which summarizes the results.

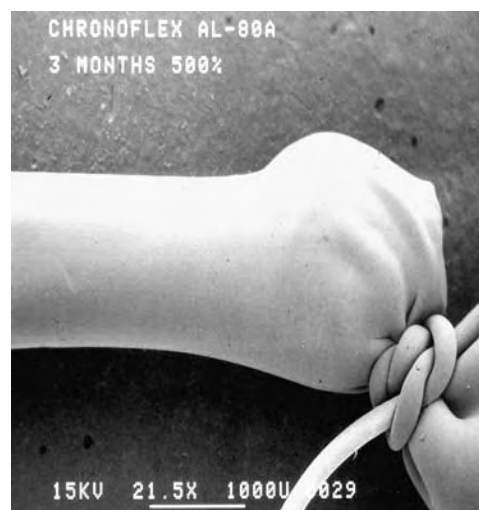


Fig. 9 ChronoFlex AL 80A poly(carbonate)urethane does not display any sign of ESC or failure after three months implantation at 500% elongation. It is our hypothesis that ether-free polyurethanes are biodegradable because no susceptible ether linkages are accessible to degrading enzymes.

Table 2 ChronoFlex biocompatibility testing

Test identification	Results
Agar overlay	Nontoxic
USP class VI	Nontoxic
MEM elution	Nontoxic
Hemolysis	Nonhemolytic
Ames mutagenicity	Nonmutagenic

AROMATIC VERSUS ALIPHATIC POLYURETHANES

As previously discussed, polyurethanes are divided into aromatic or aliphatic types on the basis of the presence or absence of benzene rings in the diisocyanate monomer used in the formulation. ChronoFlex C and AR use methylene bis(diphenyl) diisocyanate and are thus called aromatic. ChronoFlex AL is based on hydrogenated methylene bis(cyclohexyl) diisocyanate and is thus called aliphatic. This represents a major difference in formulation, which produces distinct differences between these two types of polyurethanes, as shown in Table 3.

Softening at Body Temperature

All polyurethanes soften within minutes of insertion into the body. Device manufacturers exploit this feature because it enhances patient comfort and reduces the risk to the vascular endothelium. Aliphatic polyurethanes display a significant lowering in the modulus of elasticity over corresponding aromatic polyurethanes. The actual softening at body temperature needs to be measured empirically since each formulation and hardness varies in its ability to soften.

Yellowing

Most medical-grade polyurethanes are clear to slightly yellow in their original state. However, aromatic-based polyurethanes turn dark yellow to amber as a result of processing, sterilization, or exposure to UV light.

Table 3 Physical differences between aromatic versus aliphatic polyurethanes

	Aromatic	Aliphatic
Softening at body temperature	Slight	Significant
Yellowing	Yes	No
Melt-processing temperature	High	Moderate
Solvent resistance	Good	Poor

While this discoloration, which is caused by the formation of chromophoric groups in the hard segment of the molecular chain, does not affect the physical properties of the product, it nevertheless affects the esthetics of the device. In contrast, aliphatic-based polyurethanes are more resistant to yellowing although they can also yellow slightly.

Melt-Processing Temperatures

Aliphatic polyurethanes melt at temperatures significantly lower than aromatics. For example, ChronoFlex AL elastomers are extruded at temperatures in the range of 350–365°F, whereas ChronoFlex C of equivalent durometer hardness extrudes, at nozzle temperatures in the range of 380–435°F. For medical applications, the lower the processing temperatures, the lower the rate of thermogradation since degradation follows Arrhenius' law, i.e., the rate of degradation doubles for each increment of 10°C. It follows that lower extrusion temperatures are advantageous in minimizing potential degradation that may be detrimental in the delicate biological environment.

Solvent Resistance

Aliphatic polyurethanes exhibit lower resistance to organic solvents and oils than aromatics. This is particularly true of the lower durometer aliphatics in which prolonged contact with solvents may lead to swelling and short-term contact may lead to surface tackiness.

MELT PROCESSING

ChronoFlex thermoplastic elastomers can be processed on conventional extrusion or injection-molding equipment. As with all polyurethanes, certain precautions need to be observed to ensure trouble-free operations.

Drying

ChronoFlex pellets should be dried to a moisture content equal to or less than 0.01% by weight prior to melt processing. Since ChronoFlex elastomers are highly hygroscopic, i.e., they attract and hold chemisorbed water, the pellets must be dried before extrusion. Unless pellets are properly dried, ChronoFlex tubing will exhibit haze and gels, and, under extreme moisture conditions, will appear foamy or hazy.

We recommend a machine-mounted, desiccant-type hopper drier that is capable of reaching and maintaining

a dew point of -40°F . Typically, the pellets must be dried for at least three hours with a drier inlet air temperature of 160°F . Temperatures above 170°F will likely cause pellet bridging due to surface blocking. If the pellets can be predried for extended periods, such as overnight, an inlet temperature of 150°F is generally sufficient.

Extruder

We suggest a single-stage, nonvented, 24/1 L/D extruder with a heavy-duty, DC-variable speed, high-torque drive. For example, a 2.5-inch extruder should have a 40–50 hp drive. The use of screens up to 325 mesh are common with ChronoFlex elastomers.

Screw

A general-purpose screw is recommended with an L/D ratio of 24:1 with a 12-6-6 screw for softer grades and a 6-9-9 screw for the harder grades.

Dies

A maximum drawdown ratio of 2:1 is recommended to reduce residual stresses and help reduce postextrusion shrinkage. All tooling should be streamlined to minimize material hang-up and possible polymer thermal degradation. Screws utilizing 303 or 304 stainless steel have proven acceptable in contact with ChronoFlex elastomers.

Extrusion Conditions

For extrusion, there should be four independently controlled heating sections—feed, transition, metering, and die. See Fig. 10 for recommended conditions for the softer ChronoFlex grades.

REGULATORY STATUS

ChronoFlex is used in catheters, ports, vascular grafts, and so forth. Their most common use is in long-term implants. They are appropriate for a variety of uses in which biodurability, performance, longevity, toughness, and high-stress/-strain properties are desired.

The U.S. Food and Drug Administration's (FDA) Center for Radiological Health regulates medical device applications. Raw materials, including polyurethanes, are not approved per se by the FDA. Raw materials are components of devices reviewed by FDA and are not approved individually. FDA evaluates the safety and effectiveness of a device for its intended use, and

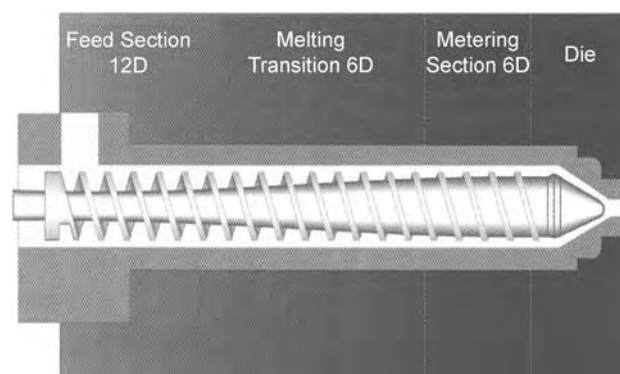


Fig. 10 A typical extruder that shows the four independent heating sections.

approvals are granted to the final product based on these considerations. It is the responsibility of the device manufacturer to determine the suitability of all components and raw materials to be used in the finished device.

The use of a medical-grade polyurethane to make an implant creates a need for toxicological data, which is generally provided by the manufacturer, as USP Class VI biocompatibility. Implant is defined as a device that is placed into a surgically or naturally formed cavity of the human body and is intended to remain implanted continuously for a period of 30 days or more.

CT Biomaterials has tested ChronoFlex in Class VI tests and are thus considered to have a low level of toxicity under the conditions of testing. Historically, this test has served as a standard to have a polyurethane considered viable as a raw material for use in implantable medical devices. Beyond USP Class VI testing, other standard test methods are used such as cytotoxicity, hemolysis, and physicochemical characteristics. The FDA has also adopted International Standards Organization (ISO) 10993: "Biological Evaluation of Medical Devices Part I—Evaluation and Testing," which, for many manufacturers, has replaced USP Class VI as the standard to have a polyurethane considered viable as a suitable raw material for use in medical devices.

CT Biomaterials has collected relevant toxicological data about ChronoFlex and has submitted a series of Device Master Files (MAFs). These files can be reviewed by the FDA in conjunction with a product review submitted by a device manufacturer that has been authorized by CT Biomaterials. CT Biomaterials maintains individual MAFs on ChronoFlex AL, AR, and C.

The MAFs submitted by CT Biomaterials on the ChronoFlex family of polyurethanes contain the following sections: 1) facilities; 2) manufacturing procedures; 3) synthesis details; and 4) QA/QC personnel



Table 4 ChronoFlex C

CardioTech International, Inc.		Biodurable poly(carbonate)urethane elastomers		
		Technical data sheet		
Properties	ASTM procedure	Typical values		
Hardness (Shore durometer)	ASTM D-2240	80A	55D	75D
Appearance	Visual	Slightly cloudy	Slightly cloudy	Slightly cloudy
Ultimate tensile strength (psi)	ASTM D-638 ^a	> 5000	> 6000	> 7000
Ultimate tensile strain (%)	ASTM D-638 ^a	400–490	365–440	255–320
100% Secant modulus (psi)	ASTM D-638 ^a	770–1250	1850–2200	5300–5700
300% Secant modulus (psi)	ASTM D-638 ^a	700–1400	1700–2000	2700–3200
Flexural strength (psi)	ASTM D-790	350	550	10,000
Flexural modulus (psi)	ASTM D-790	5500	9300	300,000
Melt index (g/10 min) 210°C; 2.17kg	ASTM D-1238	8	5	3
Vicat softening point (°F/°C)	ASTM D-1525	160/70	180/80	–
Water absorption	ASTM D-5170	1.2	1.0	0.8
Dielectric strength (volts/mil)	ASTM D-149	360	520	420
Specific gravity	ASTM D-792	1.2	1.2	1.2
Coefficient of friction (kinetic)	ASTM D-1894	1.5	0.8	0.64
Abrasion resistance (% loss at 1000 cycles)	ASTM D-1044	0.008	0.035	0.053
Melt-process. temp. (°F/°C)			375–430/190–220	
Recommended sterilization method			Gamma; E-beam; ethylene oxide	
Class VI biocompatibility test	U.S.P. XXII	Pass	Pass	Pass

^aSpecimens conditioned for 7 days at room temperature and 50% relative humidity.

involved in the manufacturing and testing of the polymers.

For example, ChronoFlex C displays tensile strengths in the range of 5000 to 7000 psi, depending on durometer hardness. This is one of the strongest polymers available on the market, and, when coupled with ultimate tensile strain (elongation) in the 300–490% range, it is considered among the toughest of the thermoplastic elastomers (Table 4).

ACKNOWLEDGMENTS

With minor revisions, this communication was published by the author as a “White Paper” on the web.

The author has collected information from many sources. The literature on medical grade polyurethanes is extensive. The reader is encouraged to refer to the original sources for additional details.

Pellethane is a registered trademark of Dow Chemical.

Tecoflex is a registered trademark of Viasys, Inc.

ChronoFlex is a tradename of CardioTech International, Inc.

ARTICLES OF FURTHER INTEREST

Biorubber/Poly(Glycerol Sebacate); Blood-Material Interactions; Elastomers, Biodegradable; Polyurethane Biomaterials; Silicones; Vascular Grafts

REFERENCES

1. Szycher, M.; Poirier, V.L. Synthetic polymers in artificial hearts: A progress report. *Ind. Chem. Eng.* **1983**, *22* (4), 558.
2. Szycher, M.; Dempsey, D.; Edwards, A. Hydrolytically and Proteolytically-Stable Polycarbonate Polyurethane Silicone Copolymers. US Patent 5,863,627.
3. Szycher, M.; Siciliano, A.A.; Reed, A.M. Polyurethanes in medical devices. *Med. Des. Mater.* **1991**, *1* (2), 19.
4. Stokes, K.B.; Urbanski, P.; Cobian, K. New Test Methods for the Evaluation of Stress Cracking and

Metal Catalyzed Oxidation in Implanted Polymers. In *Polyurethanes in Biomedical Engineering II*; Planck, H., et al., Eds.; Elsevier: Amsterdam, The Netherlands, 1987; 109–127.

5. Szycher, M. Biostability of polyurethanes: A critical review. *J. Biomater. Appl.* **1988**, *3* (2), 297–402.
6. Szycher, M. Biostable Polyurethane Products. US Patent 5,254,662.

FURTHER READING

- Bernacca, G.M.; O'Connor, B.; Williams, D.F.; Wheatley, D.J. Hydrodynamic function of polyurethane prosthetic heart valves: Influences of Young's modulus and leaflet thickness. *Biomaterials* **Jan 2002**, *23* (1), 45–50.
- Chen, C.T.; et al. Synthesis, characterization, and permeation properties of polyether-based polyurethanes. *J. Appl. Polym. Sci.* **1972**, *16*, 2105–2114.
- Farrar, D.J. The thoratec ventricular assist device: A paracorporeal pump for treating acute and chronic heart failure. *Semin. Thorac. Cardiovasc. Surg.* **Jul 2000**, *12* (3), 243–250.
- Farrar, D.J. Development of a prosthetic coronary artery bypass graft. *Heart Surg. Forum* **2000**, *3* (1), 36–40.
- Farrar, D.J.; et al. In vivo evaluations of a new thromboresistant polyurethane for artificial heart blood pumps. *J. Thorac. Cardiovasc. Surg.* **Feb 1988**, *95* (2), 191–200.
- Fields, C.; Cassano, A.; Allen, C.; Meyer, A.; Pawlowski, K.J.; Bowlin, G.L.; Rittgers, S.E.; Szycher, M. Endothelial cell seeding of a 4 mm ID polyurethane vascular graft. *J. Biomater. Appl.* **Jul 2002**, *17* (1), 45–70.
- Grasel, T.G.; et al. Surface properties and blood compatibility of polyurethaneureas. *Biomaterials* **Sep 1986**, *7*, 315–328.
- Hergenrother, R.W.; et al. Blood-contacting properties of 4,4'-polydimethylsiloxane polyurea-urethanes. *Biomaterials* **1994**, *15* (8), 635–640.
- Hutchinson, D.T.; Savory, K.M.; Bachus, K.N. Crack-growth properties of various elastomers with potential application in small joint prosthesis. *J. Biomed. Mater. Res.* **Oct 1997**, *37* (1), 94–99.
- Jahangir, A.R.; McClung, W.G.; Cornelius, R.M.; McCloskey, C.B.; Brash, J.L.; Santerre, J.P. Fluorinated surface-modifying macromolecules: Modulating adhesive protein and platelet interactions on a polyetherurethane. *J. Biomed. Mater. Res.* **Apr 2002**, *60* (1), 135–147.
- Lim, F.; et al. Synthesis, characterization and ex vivo evaluation of polydimethylsiloxane polyurea-urethanes. *Biomaterials* **1994**, *15* (6), 408–416.
- Nyilas, E.; et al. Development of Blood-Compatible Elastomers. V. Surface Structure and Blood Compatibility of Avcothane Elastomers. In *Biological Interactions at Polymer Surfaces*, J. Biomed. Mater. Res., Biomed. Mater. Symp., John Wiley & Sons, Inc.: New York, 1977; Vol. 8, 69–84.
- Okkema, A.Z.; et al. Bulk, surface, and blood-contacting properties of polyetherurethanes modified with polyethylene oxide. *J. Biomater. Sci., Polym. Ed.* **1989**, *1* (1), 43–62.
- Park, J.H.; Lee, K.B.; Kwon, I.C.; Bae, Y.H. PDMS-based polyurethanes with MPEG grafts: Mechanical properties, bacterial repellency, and release behavior of rifampicin. *J. Biomater. Sci. Polym. Ed.* **2001**, *12* (6), 629–645.
- Park, J.C.; Song, M.J.; Hwang, Y.S.; Suh Yonsei, H. Calcification comparison of polymers for vascular graft. *Med. J.* **Jun 2001**, *42* (3), 304–310.
- Paynter, R.W.; et al. The hydrolytic stability of mitrathane (a polyurethane urea)—An X-ray photoelectron spectroscopy study. *J. Biomed. Mater. Res.* **Jul 1988**, *22* (7), 687–698.
- Sakurai, S.; et al. Changes in structure and properties due to mechanical fatigue for polyurethanes containing poly(dimethylsiloxane). *Polymer* **1994**, *35* (3), 532–539.
- Speckhard, T.A.; et al. Ultimate tensile properties of segmented polyurethane elastomers: Factors leading to reduced properties for polyurethanes based on nonpolar soft segments. *Rubber Chem. Technol.* **1986**, *59*, 405–429.
- Perena, M.F.; Castillo, J.; Medrano, J.; De Gregorio, M.A.; Loras, E.; Cristobal, J.A. Nasolacrimal polyurethane stent placement: Preliminary results. *Eur. J. Ophthalmol.* **Jan–Mar 2001**, *11* (1), 25–30.
- Pulat, M.; Akdogan, A.; Ozkan, S. Microbiological interaction and diffusion properties of hydrophilic and hydrophobic PU membranes. *J. Biomater. Appl.* **Apr 2002**, *16* (4), 293–303.
- Salacinski, H.J.; Tai, N.R.; Punshon, G.; Giudiceandrea, A.; Hamilton, G.; Seifalian, A.M. Optimal endothelialisation of a new compliant poly(carbonate-urea)urethane vascular graft with effect of physiological shear stress. *Eur. J. Vasc. Endovasc. Surg.* **Oct 2000**, *20* (4), 342–352.
- Salacinski, H.J.; Punshon, G.; Krijgsman, B.; Hamilton, G.; Seifalian, A.M. A hybrid compliant vascular graft seeded with microvascular endothelial cells extracted from human omentum. *Artif. Organs* **Dec 2001**, *25* (12), 974–982.
- Salacinski, H.J.; Tai, N.R.; Carson, R.J.; Edwards, A.; Hamilton, G.; Seifalian, A.M. In vitro stability of a novel compliant poly(carbonate-urea)urethane to oxidative and hydrolytic stress. *J. Biomed. Mater. Res.* **Feb 2002**, *59* (2), 207–218.
- Seifalian, A.M.; Salacinski, H.J.; Punshon, G.; Krijgsman, B.; Hamilton, G. A new technique for measuring the cell growth and metabolism of endothelial cells seeded on vascular prostheses. *J. Biomed. Mater. Res.* **Jun 15, 2001**, *55* (4), 637–644.
- Soldani, G.; Varelli, G.; Minnocci, A.; Dario, P. Manufacturing and microscopical characterization of polyurethane nerve guidance channel featuring a highly smooth internal surface. *J. Biomater.* **Nov 1998**, *19* (21), 1919–1924.
- Sugimoto, K.; Hirota, S.; Imanaka, K.; Ohnishi, H.; Tomioka, H.; Tada, K.; Aoki, M.; Sugimura, K. Covered Gianturco stent for tracheal stricture: Application of polychlorovinylidene and polyurethane as covering materials. *Eur. Radiol.* **2001**, *11* (10), 1933–1938.
- Szycher, M.; Reed, A.M. In vivo testing of a biostable polyurethane. *J. Biomater. Appl.* **Oct 1991**, *6* (2), 110–130.



- Szycher, M.; et al. Blood compatible polyurethane elastomers. *J. Biomater. Appl.* **Oct 1987**, 2, 290–313.
- Tai, N.R.; Salacinski, H.J.; Edwards, A.; Hamilton, G.; Seifalian, A.M. Compliance properties of conduits used in vascular reconstruction. *Br. J. Surg.* **Nov 2000**, 87 (11), 1516–1524.
- Takahara, A.; et al. Surface molecular mobility and platelet reactivity of segmented poly(etherurethaneureas) with hydrophilic and hydrophobic soft segment components. *J. Biomater. Sci., Polym. Ed.* **1989**, 1 (1), 17–29.
- Takahara, A.; et al. Effect of soft segment chemistry on the biostability of segmented polyurethanes. II. In vitro hydrolytic degradation and lipid sorption. *J. Biomed. Mater. Res.* **1992**, 26, 801–818.
- Tanzi, M.C.; Fare, S.; Petrini, P. In vitro stability of polyether and polycarbonate urethanes. *J. Biomater. Appl.* **Apr 2000**, 14 (4), 325–348.
- Thomas, V.; Jayabalan, M.; Sandhya, S. Studies on polyurethane potting compound based on isocyanurate of aliphatic diisocyanate for fabrication of a haemodialyser. *J. Biomater. Appl.* **Oct 2000**, 15 (2), 86–105.
- Tiwari, A.; Salacinski, H.; Seifalian, A.M.; Hamilton, G. New prostheses for use in bypass grafts with special emphasis on polyurethanes. *Cardiovasc. Surg.* **Jun 2002**, 10 (3), 191–197.
- Van Den Hazel, S.J.; Mulder, C.J.; Den Hartog, G.; Thies, J.E.; Westhof, W. A randomized trial of polyurethane and silicone percutaneous endoscopic gastrostomy catheters. *Aliment. Pharmacol. Ther.* **Oct 2000**, 14 (10), 1273–1277.
- Visai, L.; Rindi, S.; Speziale, P.; Petrini, P.; Fare, S.; Tanzi, M.C. In vitro interactions of biomedical polyurethanes with macrophages and bacterial cells. *J. Biomater. Appl.* **Jan 2002**, 16 (3), 191–214.
- Ward, R.S. Surface modification prior to surface formation: Control of polymer surface properties via bulk composition, medical plastics and biomaterials. *Mater. Forum* **1995**, 34–41.
- Ward, R.S.; et al. Production of Biomedical I. Polymers, Silicone/Urethane Synergy in Avcothane Elastomers. In *Organometallic Polymers*; Carraher, E.C., Jr., Ed.; Academic Press: New York, 1978; 219–229.
- Ward, R.S.; et al. BPS-215M: A Polyurethaneurea for Biomedical Devices: Development and in vivo Testing in the Pierce–Donachy VAD, The 13th Annual Meeting of the Society for Biomaterials, New York, NY, USA, Jun. 2–6, 1987; 259.
- Ward, R.S.; et al. Development of a hybrid artificial pancreas with a dense polyurethane membrane. *ASAIO J.* **1993**, 39 (3), M261–M267.
- Ward, R.S.; et al. Improved Polymer Biostability via Oligomeric End Groups Incorporated During Synthesis. In *Polymeric Materials Science and Engineering*, Fall Meeting, Boston, MA, Aug. 23–27, 1998; 526–527.
- Wheatley, D.J.; Raco, L.; Bernacca, G.M.; Sim, I.; Belcher, P.R.; Boyd, J.S. Polyurethane: Material for the next generation of heart valve prostheses? *Eur. J. Cardiothorac. Surg.* **Apr 2000**, 17 (4), 440–448.
- White, K.A.; et al. Surface Modification of Segmented Polyurethaneureas via Oligomeric End Groups Incorporated During Synthesis. In *Surface Modification of Polymeric Biomaterials*; Rainer, B.D., et al., Eds.; Plenum Press: New York, NY, 1996; 27–33.
- Wiggins, M.J.; Wilkoff, B.; Anderson, J.M.; Hiltner, A. Biodegradation of polyether polyurethane inner insulation in bipolar pacemaker leads. *J. Biomed. Mater. Res.* **May 1, 2001**, 58 (3), 302–307.
- Yang, M.; Zhang, Z.; Hahn, C.; King, M.W.; Guidoin, R. Assessing the resistance to calcification of polyurethane membranes used in the manufacture of ventricles for a totally implantable artificial heart. *J. Biomed. Mater. Res.* **1999**, 48 (5), 648–659.

Polyelectrolyte Multilayers

Darwin R. Reyes

*Semiconductor Electronics Division, Electronics and Electrical Engineering Laboratory,
National Institute of Standards and Technology, Gaithersburg, Maryland, U.S.A.*

Laurie E. Locascio

*Molecular Spectrometry and Microfluidic Methods Group, Analytical Chemistry Division,
National Institute of Standards and Technology, Gaithersburg, Maryland, U.S.A.*

Michael Gaitan

*Semiconductor Electronics Division, Electronics and Electrical Engineering Laboratory,
National Institute of Standards and Technology, Gaithersburg, Maryland, U.S.A.*

INTRODUCTION

The process of self-assembly occurs naturally in living systems to form complex structures out of “random events” in an incredibly accurate way. Order out of disorder is achieved in a fashion that is analogous to a jigsaw puzzle where each piece fits together in a specific configuration by a process that is called molecular recognition. Self-assemblies and template-assisted assemblies have recently become a key approach to nanofabrication for the assembly of nanoarchitectures. In 1992, Decher and coworkers reported on a template-assisted assembly technique based on a process that relies on electrostatic interactions between polycations and polyanions to form ultrathin multilayers of polymers, known as polyelectrolyte multilayers (PEMs).^[1,2] This approach takes advantage of the spontaneous self-assembly of a polyelectrolyte on a substrate when immersed in a polyelectrolyte solution. The polyelectrolytes interact electrostatically with the surface adsorbing a polymer monolayer. The substrate is subsequently rinsed with water and immersed in an oppositely charge polyelectrolyte solution until a second layer is adsorbed. Then, the substrate with the two layers of adsorbed polymers is rinsed again with water. These steps can be repeated as many times as polyelectrolyte layers are required. This method is known as layer-by-layer deposition or electrostatic self-assembly. The approach overcomes the limitations encountered in covalent and coordination chemistries in terms of the required proximity of reactive and neighboring groups for the binding to occur.

A number of chemical interactions and materials have been used since the first report on the fabrication of PEMs appeared in the scientific literature.^[3] Initially, the multilayers were formed using electrostatic interactions, but later other types of chemical interactions (i.e., hydrogen bonding, covalent bonds, stereocomplex formation, etc.) have been used successfully. In a similar way, the number of different

chemical structures and materials used to form PEMs has been growing, and a wealth of new applications have emerged. Polyelectrolytes, macromolecules, biomacromolecules, and biological cells, among others, are some of the materials currently used when assembling PEMs. Also, a variety of new applications such as light emitting diodes, separation membranes, flow control, and cell adhesion and culture have been explored using this assembly technique. In this entry, we intend to cover the fundamental aspects of PEM structure and conditions for multilayer formation based on electrostatic interactions. In addition, the characterization techniques commonly used to analyze the topography and chemical interactions of the layers are described in a general way. Finally, we discuss some interesting applications and describe relevant aspects regarding the relation between the materials and the specific application.

POLYELECTROLYTE MULTILAYER STRUCTURE

The structure of PEMs, in the simplest film arrangement of same repeating pairs of polyanion and polycation, has been described as a three-characteristic zone model.^[4,5] The region closest to the substrate is known as Zone I and consists of one or just a few layers of polyelectrolytes. This zone is shaped by the characteristic properties of the substrate in terms of charge and roughness. The next stratum, Zone II, is considered to be neutral and exhibits bulk film properties. Last, surface Zone III (the outermost) can interact with the solution (or air). Though the structure of PEMs is subdivided into three zones, the transition between them is subtle. Fig. 1 shows an example of the buildup of the three zones and their relative thickness.

The presence of counterions in Zone I depends on the charge and roughness of the substrate, whereas they are believed to be everywhere in Zone III. The counterions



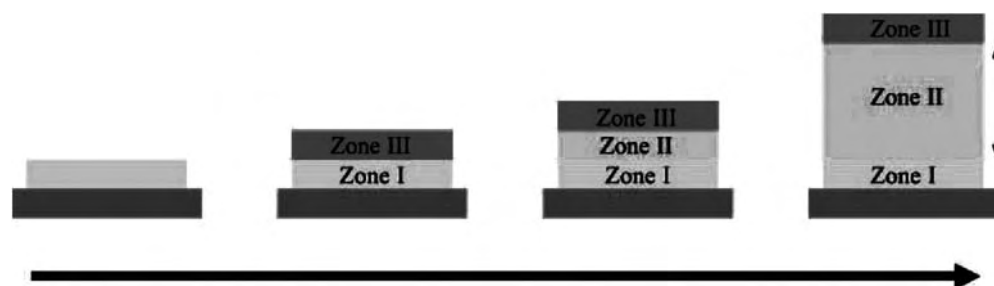


Fig. 1 Polyelectrolyte multilayers zone model.

in Zone III are associated with the excess charge created when charge overcompensation of the previous surface occurs owing to the adsorption of a new polymer layer. At the same time, it is deemed that layers distant from the surface will play a role in the charge excess observed at the surface by interspersing with the adjacent layers.^[6] These distinctive properties between the three zones play an important part in the swelling and deswelling of PEMs. Because Zone II is considered a neutral region, due to overall charge compensation, it could be expected to behave similarly to a polyzwitterion. Such a neutral zone should swell when immersed in an ionic solution and deswell in pure water. This behavior is better observed when the polyelectrolyte complex has been created in the presence of salt. When the PEMs are formed in pure water, the swelling is considerably less.^[7] On the other hand, Zone III is considered to behave as a polyion. Therefore, under pure water conditions Zone III would be expected to swell, and under ionic solution conditions it would be expected to deswell, as polyelectrolytes would do.

The formation of the three zones should be considered only when a minimum number of layers have been assembled. Zones I and III will be the first to form, and Zone II will not form until the other two zones (I and III) have attained their final thickness. After this point, Zones I and III are expected to maintain their thickness even when more layers are added to the multilayer film. Therefore, when a new layer is added to the surface of the film (Zone III), the material in the boundary between Zone II and III will move away from the coercing effect of Zone III and will become part of Zone II. In this way, the thickness of Zone III will remain the same, and that of Zone II will increase further.^[1,8]

CONDITIONS FOR MULTILAYER FORMATION AND MULTILAYER MATERIALS

Conditions for Multilayer Formation

In general, the formation of PEMs is based on the electrostatic attraction of oppositely charged polymers

that are sequentially stacked. Experimental conditions such as temperature, ionic strength, pH, flexibility of the polymers, and polyelectrolyte charge, among others, will affect the final structure of the polymers when shaping the multilayers. There have been several model representations for the formation of PEMs.^[9–12] The more prominent of them presents a mechanism of multilayer formation based on experimental data showing interpenetration of succeeding layers of weak and flexible polyelectrolytes at high ionic strength.^[11] To model the building up of multilayers in a simplified way, a couple of assumptions need to be made. The first of these is the assumption that the process is irreversible, so no desorption should occur due to equilibrium reactions, competition between the polyelectrolyte and salt ions, or detachment of the deposited polyion by interaction with the oppositely charged polyion in solution.^[6,13] Second, there are no strict boundaries between adjacent layers; instead, interpenetration of polymer chains takes place between neighboring films.^[14–16]

We will consider here flexible weak polyelectrolytes at high ionic strength to describe the multilayer formation.^[11] At high ionic strength, ions in solution exert a screening effect on the charged groups of the polyelectrolytes leading to short-range electrostatic interactions with the charged surface. The formation of the first layer starts with the adsorption of a first polymer containing multiple charges. This charge excess produces a reversal of the charge on the substrate surface. Nonelectrostatic interactions, such as van der Waals forces, hydrophobic interactions, etc., between the polyelectrolyte and the surface can contribute to the attraction forces taking part in the anchorage of the polyions to the surface. This can increase the amount of polyion adsorbed on the surface and in turn will produce a higher charge reversal. Because the first layer is the base for the successive adhesion of polyelectrolytes, it is practical to deposit a polyion with good adhesion characteristics first and change to the desired polyelectrolytes in the layers that follow. Once the first layer has been formed, an oppositely charged polyion solution is then placed in contact with the first layer to generate a second layer. Repulsions between the surface and the second polyion (both having the

same charge) will create a depletion layer in the proximity of the substrate surface.^[11] Moving outward from the end of the depletion layer to the outer side of the first polyion, complexation with the second polyelectrolyte occurs. The polyions in the first layer interact with the secondary polyions in a way that all are complexed and no further interaction with the secondary polyion can take place. On the other hand, the second polyelectrolyte interacts with the first polyion on one side and with the bulk solution on the other side. It can form large loops extending toward the solution, producing a formation similar to a double layer. Therefore, the amount of polyions in the first layer is limited by substrate surface constrictions, such as charge area and roughness, whereas the extent of the second layer is not limited because the side of the layer facing the bulk of the solution is in thermal equilibrium with it. Further layer depositions should follow the same interactions just described. Deposition of a larger number of layers will produce collective interactions between the layers in the bottom, middle, and top portions of the PEMs (Zones I, II, and III), creating different grouped characteristics. This has been discussed in the previous section.

This model is based on linear electrostatics and cannot be applied to excessively high local density charges. Therefore, it is valid only for weakly charged polyelectrolytes. An interesting feature of PEMs is that they can also be produced by replacing a polyelectrolyte layer for another charge species such as biological polymers (i.e., proteins, DNA) and non-biological chemicals (i.e., surfactants). The variety of charged species that can be used to create PEMs makes this type of assembly very flexible and broadens the number of applications where PEMs could be used. Examples of charged materials used for PEMs assembly will be discussed in the next section.

Multilayer Materials

Polyelectrolytes and other materials

The first work in what is known today as PEMs utilized negatively and positively charged polyions to produce an assembly of multiple nanosize layers in the Z dimension.^[17] Since then, a variety of charged materials have been used to create structures for specific applications. Even though PEMs buildup is not confined to the electrostatic interactions of oppositely charged species, most of the work carried out in this field relies on these types of interactions. Table 1 shows some of the most common polyion chemical structures reported in the literature. Nevertheless, there have been reports on PEM assemblies based on the interactions between polyelectrolytes with biomolecules and other

biological materials such as lipids, proteins, cells, and viruses.^[18–26]

MULTILAYER DEPOSITION TECHNIQUES

Even though the layer-by-layer deposition process commonly relies on electrostatic attractions, as already mentioned, several different deposition methods have been reported in the literature since the assembly of the first multilayer inorganic colloids by Iler.^[27] Some of these interactions include covalent bonds, hydrogen bonding, donor–acceptor interactions, and stereo-complex formations among others. In the following sections, a number of methods used for the fabrication of PEMs using some of the above-mentioned interactions will be described.

Dipping Technique (Total Surface Immersion)

The first studies on the sequential formation of monolayers of polyions (of up to 100 layers) and their multilayer structural organization were published by Decher and coworkers.^[1,17] This method has been widely used since its introduction and consists of the immersion of a charged substrate in a polyion solution A, of opposite charge. A rinsing step in water or salt solution follows the deposition of the first monolayer to remove unbound polyions. The substrate is again immersed in a polyion solution B, but this time with the polyion having the same charge as the substrate. A rinsing step follows the deposition of polyion B. These steps produce A/B bilayers, and further deposition steps using the same polyions will generate an (A/B)_n structure.

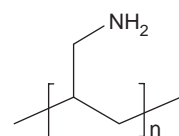
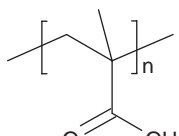
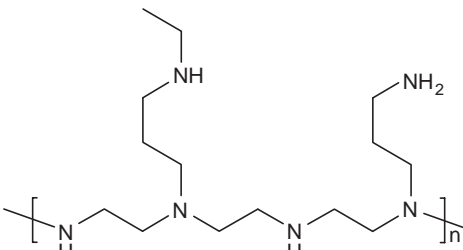
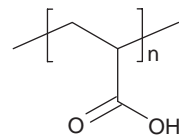
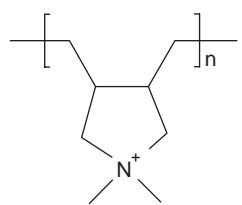
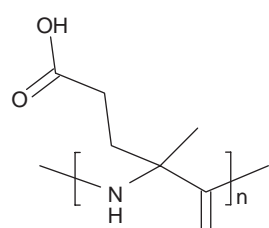
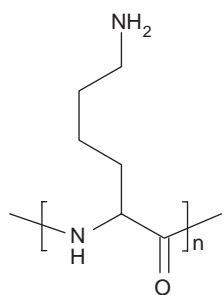
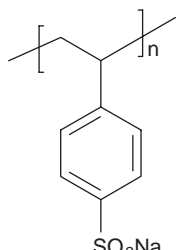
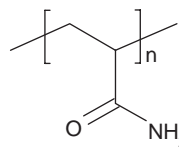
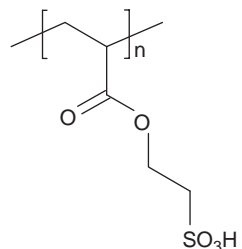
Other methods based on the same sequential deposition principle, but in restricted areas, have emerged. These methods have been developed mainly for the need of producing PEM areas of micrometer size.

Patterning Techniques

Stamping

This technique uses a polymer-on-polymer stamping procedure to pattern regions of weak PEM films on a substrate chemically.^[28–30] In this case, the polyions used were poly(allylamine hydrochloride) (PAH) and poly(acrylic acid) (PAA). A poly(dimethyl siloxane) (PDMS) stamp is immersed in the weak polyion for about 1 h. The stamp is then rinsed with water and dried. To finally obtain the pattern on the substrate, the stamp is brought into contact with the substrate for a short period of time, removed, and then rinsed with water. The thickness of the layers is adjusted by varying the pH of the polyion solution. Thicker layers

Table 1 Chemical structures of some common polyelectrolytes

Polycations	Polyanions
 <p data-bbox="268 464 627 498">Poly(allylamine hydrochloride) (PAH)</p>	 <p data-bbox="1137 487 1424 521">Poly(methacrylic acid) (PMA)</p>
 <p data-bbox="313 816 573 850">Poly(ethylene imine) (PEI)</p>	 <p data-bbox="1227 703 1442 737">Polyacrylic acid (PAA)</p>
 <p data-bbox="206 1099 609 1168">Poly(diallyl-dimethyl ammonium) chloride (PDADMAC)</p>	 <p data-bbox="1191 1145 1460 1179">Poly(glutamic acid) (PGA)</p>
 <p data-bbox="376 1530 537 1564">Poly(lysine) (PL)</p>	 <p data-bbox="1182 1485 1478 1519">Poly(styrene sulfonate) (PSS)</p>
 <p data-bbox="331 1769 573 1803">Polyacrylamide (PAAm)</p>	 <p data-bbox="1128 1871 1532 1905">Poly(3-sulfopropyl methacrylate) (PSPM)</p>

are obtained when the pH of the polyamine solution significantly increases, and, hence, the polyamine is less charged. More recently, a similar type of patterning approach was used to fabricate polymer–colloid templates and to control selective cell attachment.^[31]

Microfluidics

Microfluidic structures have been used as molds for flowing through polyelectrolytes to facilitate their deposition, in a patterned fashion, on different surfaces. Either stagnant or dynamic depositions have been achieved using this platform. The dynamic deposition has been carried out by spinning the polyelectrolytes through microchannels placed on a carboxylic terminated self-assembled monolayer (16-mercaptohexadecanoic acid) previously deposited on a gold surface.^[32] With this system, up to 40 layers were deposited in a very short time. In addition, sub-micrometer width lines can be patterned using this approach. Stagnant deposition of polyelectrolytes using microfluidic networks has been recently presented (D. R. Reyes, S. P. Forry, L. E. Locascio, M. Gaitan, in preparation).^[25] In this case, a microfluidic network is placed on an oxidized PDMS flat surface, and then the microchannels are filled with the polyelectrolyte solution. The polyelectrolytes are allowed to deposit on the PDMS surface and are later removed from the channels following a rinsing step with water. By repeating this cycle, PEM lines are patterned on this elastomer.

Substrate Materials

The motivation when choosing a certain material for PEM formation is primarily based on the feasibility of a strong interaction between the polyelectrolyte and the substrate, and on the purpose for the formation of the multilayers (i.e., detection in characterization studies, surface attachment of chemical or biological material, etc.). Thus, the physico-chemical characteristics play an important role in selecting a substrate. There are a variety of materials that have been used as substrates for multilayer formation, such as glass, silicon, gold, quartz, and polymers.

Electrostatic interactions are commonly used to build up PEMs. For this type of interaction to occur between a substrate and a polyelectrolyte, the substrate should either have an intrinsic charge under the deposition conditions or should be capable of acquiring surface charge by chemical or physical means (i.e., redox reactions with strong acids or bases, or treatment with a gas plasma). These treatments will increase the interaction between the substrate and the first layer

of polyelectrolyte, which is essential for a uniform PEMs assembly.

Glass and quartz are substrate materials suitable for detection methods such as UV–Vis spectroscopy and optical microscopy because they are transparent at these wavelengths. Silicon is compatible with detection techniques such as x-ray reflectivity, IR spectroscopy, and ellipsometry.^[2,6,17,33–39] Gold surfaces are commonly employed for electrochemical, quartz crystal microbalance (QCM), and surface plasmon resonance (SPR) detection techniques.^[40–51]

The functionalization of substrates is an alternative when optimizing the interaction between the substrate surface and a polyelectrolyte. For example, materials such as glass, quartz, and oxidized silicon wafers can be deprotonated by chemical treatment with a strong base to produce a negatively charged surface. On the other hand, the same surfaces could be positively charged by further reactions with silanes such as trialkoxy-3-aminopropylsilane.^[38,52,53] Positive and negative charges can be attained on gold surfaces as well when treated with thiols or disulfides compounds.^[54,55] Polydimethylsiloxane becomes amenable to polyelectrolyte deposition, by means of electrostatic forces, when treated with oxygen plasma and promptly immersed in polyelectrolyte solution. Polymeric surfaces such as poly(propylene), poly(styrene), poly(methyl methacrylate), and poly(ethylene terephthalate) can also be used as the substrate for PEM depositions with good results. Deposition on such uncharged surfaces is also an alternative if the appropriate hydrophobic polyelectrolytes combination is selected and no surface treatment is desired.^[56,57]

CHARACTERIZATION TECHNIQUES

Characterization of the multilayer's thickness, surface topography, and, in some cases, the internal structure interactions can be carried out either by stopping the deposition steps and analyzing multilayer formation as a dried sample or by direct measurement in solution. Analysis techniques such as UV–Vis spectroscopy, ellipsometry, and x-ray reflection have been commonly used to measure the thickness of layers deposited on solid substrates.^[2,6,17,33,34,38,39,51,58] Other techniques have been added to the repertoire of analytical methods to monitor these deposited layers, for example, QCM, SPR, scanning angle reflectometry, atomic force microscopy, Fourier transform infrared (FTIR) spectroscopy, and optical waveguide lightmode spectroscopy (D. R. Reyes, S. P. Forry, L. E. Locascio, M. Gaitan, in preparation).^[25,32,42–44,48,49,59–76] Refer to Table 2 for a brief description of the parameters obtained, workable film thickness range, and important references regarding these techniques.



Table 2 Characterization techniques

Techniques	Parameter measured	Film thickness range ^a	References
Ellipsometry	Change in intensity and state of polarization of a reflected light at a surface	From a few to 6000 nm	[6,33,38,39]
X-ray reflection	Electron density of the material, film thickness, and average roughness	From about 1 nm to 1 μ m	[2,17,34,58]
Quartz crystal microbalance	Changes in resonance frequency of a quartz crystal when a film of a material is adsorbed on the crystal surface	From < 1 nm to a few micrometers	[42,43,59–63]
Surface plasmon spectroscopy	Change in angle of incidence (needed to excite a surface plasmon) when the thickness of a material changes	From < 1 nm to ~200–300 nm ^[77,78]	[44,48,49,64,65]
Scanning angle reflectometry	Interface changes in the reflectivity coefficients of a layer deposited on a surface	From < 1 nm to 100–150 nm ^[79,80]	[66]
Atomic force microscopy	Attractive or repulsive forces occurring between a tip and a sample when the sample surface is being scanned; a topographic image of the surface is produced	From < 1 nm to about 5 or 6 μ m	(D.R. Reyes, S.P. Forry, L.E. Locascio, M. Gaitan, in preparation) ^[25,32,67,68]
Fourier transform infrared spectroscopy	Absorption of infrared wavelengths, due to functional groups, by the sample	N/A ^b	[69–73]
Optical waveguide lightmode spectroscopy	An evanescent field of a laser is coupled to a grating integrated in a planar optical waveguide, the measurement of the incoupling angles permits the monitoring of the adsorbed macromolecules	From < 3 nm to 100–200 nm ^[81] (for up to 400 nm, see Ref. [75].)	[66,74–76]

^aFilm thickness range reported came from commercial instruments unless a reference is cited.

^bFourier transform IR spectroscopy is used for chemical characterization of functional groups on the surface, and it will not provide thickness or morphology information unless it is combined with other techniques. The film thicknesses that could be analyzed with FTIR are in the range of a few to several hundred nanometers.

APPLICATIONS

Light Emitting Diodes (LED)

Organic polymeric materials have received interest for electronic and electrochromic applications because they offer the capability of forming complex circuits on low-cost plastic materials and flexible displays. Semi-conducting and electrochromic polymer devices have been developed, and complex circuits have been demonstrated by ink-jet processes.^[82–85] The ink-jet based processes can typically produce dots of material with a minimum diameter on the order of 10 μ m with 10 μ m spacing.

More recently, the polycation poly(phenylene vinylene) (PPV) has been deposited using layer-by-layer self-assembly processes in combination with the polyanions poly(methacrylic acid) (PMA) and poly(styrene

sulfonate) (PSS).^[86] It has been observed that the number of deposited bilayers contributes linearly to the amount of absorbance and photoluminescence of the device up to about 30 bilayers, although the thickness of each bilayer is reported to be constant up to 50 bilayers. It has also been reported that the type of anion chosen will affect the behavior and performance of the devices and that PMA is the polyanion of choice. Devices assembled with PPV/PMA were observed to perform with luminance levels of 20–60 cd/m^2 and with a rectification ratio of greater than 10^5 . These materials can be patterned on surfaces using selective deposition techniques that are based on using a mold to transfer (or stamp) the polymers on the surface.^[87]

An important advantage of using a layer-by-layer self-assembly process to construct polymer-based light emitting diodes is the molecular level control of film thickness. This degree of control has the potential to produce

devices of high quality and uniformity. However, methods based on PEMs have so far not yielded circuits of equal complexity to the ink-jet based processes.

Separations

The fabrication of membranes suitable for separations using layer-by-layer deposition of polyions has been used for a number of applications. Gas and aqueous samples have been subjected to selective permeability separations using these polyelectrolyte membranes, and evident improvements have been obtained.

With gases, asymmetric membranes combining PAH, PSS, and dimethyl silicone membranes have proven to have an increased permeability for CO₂ than for N₂ when compared with the same polyions on porous poly(propylene) membranes. Permeability values increased from 6.2 to 23.8 when the temperature was increased from 20°C to 50°C.^[88] Another alternative for the separation of the components of a gas sample is the use of a composite gas permeable membrane formed of PAH and PSS to modify poly(4-methyl-L-pentene). This asymmetric membrane was demonstrated to have 18,000 times better barrier capacity for nitrogen than for oxygen and hydrogen. When polyaniline and PSS were alternatively used, the permeability of the membrane did not change compared to the bare surface of poly(4-methyl-L-pentene).^[89]

Control of permeabilization properties for separation purposes in aqueous media can also be achieved with polyelectrolyte layers. For example, multilayer deposition of derivatized PAA/PAH films produces a layer with hydrophobic characteristics that passivates electrodes below this layer. Subsequent cross-linking of the polyions provides wider pH range stability. This allows for base-promoted hydrolysis, which generates a selective, hydrophobic layer. The new layer demonstrated better permeability toward Ru(NH₃)₆³⁺ when compared to Fe(CN)₆³⁻, as a result of newly created carboxylic groups.^[90]

The modification of capillary surfaces using layer-by-layer deposition was tested for the deposition of polybrene (polycation) and dextran sulfate (polyanion). In comparison with the conventional capillary modification, the layer-by-layer deposition technique showed remarkable improvements. Using this technique, the modified capillary surface tolerated 600 replicate analyses and proved to be stable under basic conditions (1 M NaOH) and acid concentrations of about 0.1 M HCl. Electroosmotic flows (EOF) independent of pH were observed at pHs in the range between 2 and 11. Also, the relative standard deviation was less than 1% between runs carried out on different days and different capillaries. Protein separations at

pHs near the proteins' isoelectric point were obtained as well as reproducible separations of samples with complex matrices such as plasma. Separations of cationic, neutral, and anionic amino acids, within the same electrophoretic run, were attained using capillary zone electrophoresis.^[91,92] In another study, commonly used polyelectrolytes such as PDADMAC and PSS showed the use of PEMs as modifying coating materials in conventional capillary separations. Extreme pH and ionic strength as well as dehydration and rehydration experiments were performed, proving these coatings to be stable under such conditions for the separation of a series of proteins.^[93]

Chiral separations have been demonstrated using optically active PEMs as membranes. L- and D-poly(lysine) (PL) and L- and D-poly(glutamic acid) (PGA) were used as the membrane components for the separation of chiral compounds. This approach showed a better chiral separation when a combination of the two optically active polyelectrolytes, L- and D-PL and L- and D-PGA, were used rather than only one. No selectivity for any of the chiral probes was observed when optically inactive polyelectrolytes were used. Control of the selectivity and flux of the optical species was achieved by regulating the salt concentration. Swelling of the PEMs generates reversible doping ion exchange/transport sites, which allows for a small or large flux of ions through the membranes depending on the size of the created sites. For a higher flux, a reduction of selectivity is observed, and the opposite effect is observed for a smaller flux.^[94]

Biomolecules, Cell Adhesion, and Cell Encapsulation

Cell adhesion on the extracellular matrix is mediated by cell surface proteins. These surface proteins include the cadherin, selectin, and integrin families.^[95] In addition to protein-specific interactions, charge interactions seem to have a role in the adherence of cells when growing on artificial surfaces.^[25,96-98] Cells were initially cultured on slightly negative glass surfaces, where they attached and spread out. Later, plastics such as polystyrene were used for cell culture after being treated with an electric ion discharge or high-energy ionizing radiation to promote cell adhesion.^[99] Most of the work done using PEMs as a biomaterial surface to culture cells was inspired and driven by the charge interactions between the surface of the PEMs and the charged surface of the cells. Even though charge interactions are thought to be important as a first mechanism when attaching and growing cells on PEMs, polyelectrolytes have also been modified with the RGD domain that is recognized by integrins.^[29] RGD is a known cell adhesion molecule, which is

composed of the amino acids arginine-glycine-aspartic acid and is found in the extracellular matrix protein fibronectin.^[100] Although not all polyelectrolytes are suitable as surface material for growing cells, there are a number of them that have shown good biocompatibility. In the next paragraphs we will discuss some examples of the polyelectrolytes, cells, and approaches used to culture cells on PEMs in macrocultures as well as in micropatterns.

A number of cells such as osteoblast-like cells, human periodontal ligament cells, chondrosarcoma cells, mammalian fibroblasts, and rat retinal cells have been cultured on a variety of polyelectrolytes. The intrinsic properties of PEMs, primarily their surface charge, the flexibility in using a wide variety of combinations of cations and anions, and the ease of deposition and patterning, have made them very convenient as a biomaterial scaffold for growing cells. PAH, PGA, PSS, and poly(L-lysine) (PLL) have shown good adhesive properties (similar to controls) for growing human osteoblast-like SaOS-2 cells and human periodontal ligament cells.^[101] Cell attachment of these two cell types was observed as fast as 20 min after seeding on PAH, PGA, and PLL. Results of cell viability, osteoblast phenotype expression, and inflammatory responses have indicated that PSS, PGA, and PLL can be considered as biocompatible coatings for bone tissue implants.

The organization of chondrosarcoma cell contacts in a three-dimensional matrix containing titanium beads modified with PLL or PAH and polyanions PGA or PSS was studied. Discrete focal spots of vinculin and actin cables were observed when chondrosarcoma cells adhered to uncoated or negatively charged surfaces. Also, cells plated onto these surfaces showed a twofold increase in phosphorylation of p44/42 MAPK/ERK. On the other hand, no cytoskeletal vinculin and actin organization was observed when cells were plated on a positively charged surface.^[102]

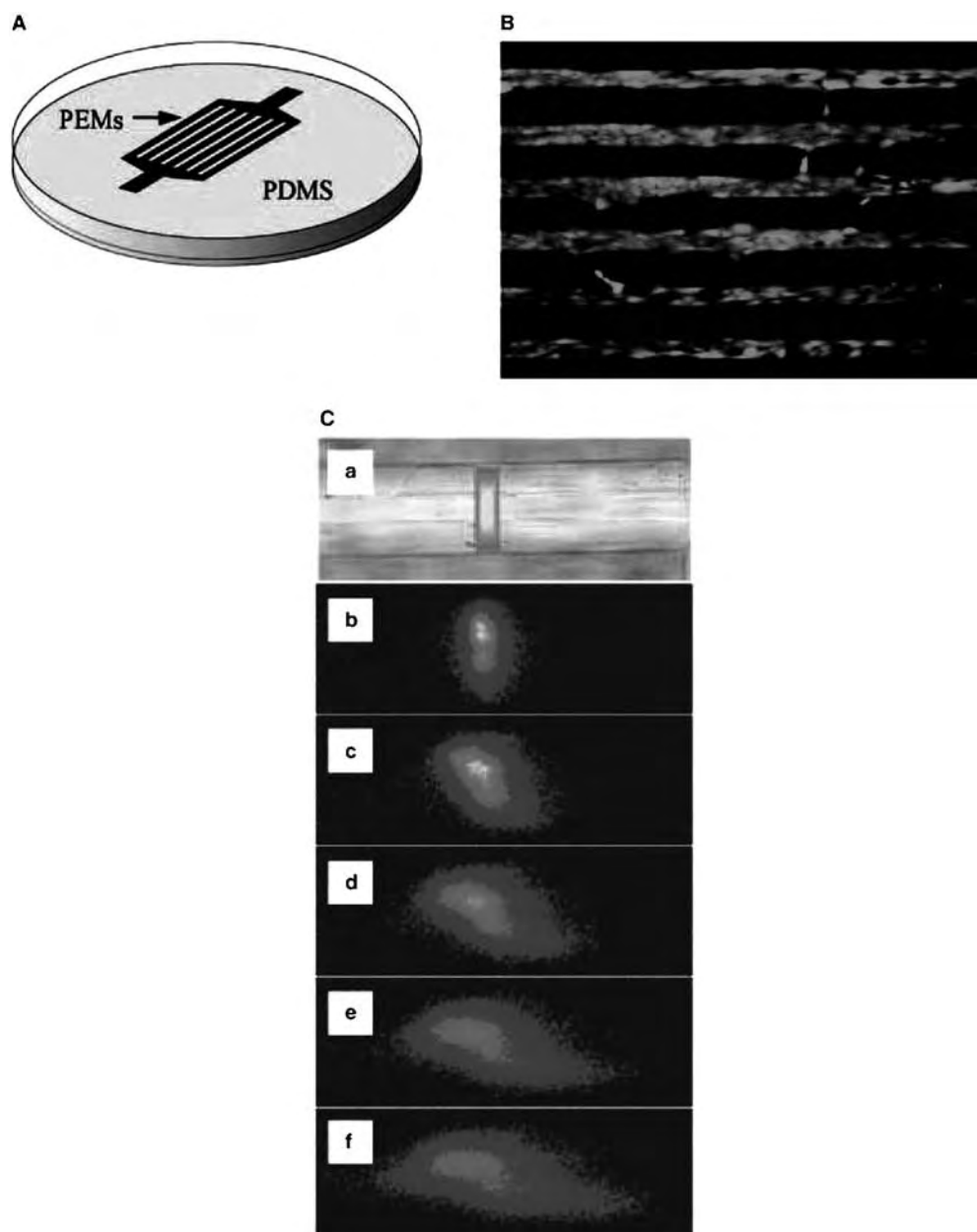
Another approach for using PEMs as adhesive material for cell attachment is the patterning of polyelectrolytes using stamping methods or microfluidic patterning. Polyelectrolytes have been patterned on various substrates to either attach cells or prevent them from attaching to specific areas on the substrate surface. While polyelectrolytes such as PAH, PSS, PLL, PEI, and many others are good materials for cell attachment, others such as PAA/polyacrylamide (PAAm) and PAA/PMA have been shown to be resistant coatings for mammalian fibroblasts.^[24]

Rat retinal (R28) cells have been successfully patterned on PAH and PEI, and the viability and morphology on such coatings have been studied.^[25] These cells attached and grew on patterned lines (approximately 55 μm wide) of these polyelectrolytes

deposited on PDMS surfaces. This approach took advantage of the reversible nature of the PDMS that, in oxidized form, has a highly charged negative surface, but reverts back to its natural uncharged hydrophobic state when stored in air. This characteristic of PDMS was used to create a PDMS substrate with two regions of different cell adhesion characteristics: 1) an oxidized surface onto which PEMs were patterned using microfluidic patterning to direct cellular deposition (the cell adhesion region) and 2) an uncharged region in the areas where PEMs were not deposited that reverted back to its hydrophobic state after exposure to air, (the cell resistant region) (see Fig. 2A). The cells attached and grew onto the areas where PEMs were deposited, whereas the cells were repelled from the uncharged PDMS areas (see Fig. 2B). Cells attached between 5 and 6 hr, the same time they normally attach in control culture dishes. A viability assay 24 h after seeding the cells proved they were viable, and their morphology remained the same on the areas larger than 55 μm when compared with the control experiments.

As previously mentioned, the adhesion of cells can also be directed by modifying the outermost polyelectrolyte on the PEMs with peptide sequences known to bind to protein receptors on the surface of the cells. An example of a well-known peptide sequence that binds surface proteins is the RGD motif. This motif was used to modify the outermost layer of PEMs, which contained PAH as the first layer and PAA (outermost polyelectrolyte) and PAAm.^[24] To control the ligand density (RGD motif), the pH of the PAH solution was changed during the stamping process. It was determined that a PAH solution buffered to a pH of 11 provided the highest functional group densities. The lowest functional group densities were obtained at a solution of pH 3.5. Cells tended to attach in higher numbers as the RGD densities increased. Interestingly, cells with nicely shaped focal adhesions and cell-spanning stress fibers were only found in line patterns of 25 μm or wider, indicating that geometry and surface functionality are both important in determining cell adhesion.

Recently, PEMs have been used to coat cells in suspension to encapsulate single cells, and the characteristics of these nanoencapsulated cells have been studied. Diaspro et al. coated *Saccharomyces cerevisiae* (yeast) cells by sequential incubation in suspension with alternating solutions of negatively charged (PSS) and positively charged (PAH) polyelectrolytes.^[21] In this method, the positively charged polyelectrolyte was incubated with the cells in the first step, taking advantage of the negative charges on the yeast cell surface to induce complete coating. As a result of this work, it was determined that single yeast cells could be effectively encapsulated by the PEMs, forming a



P

Fig. 2 Panel A shows the patterned PEMs (black) lines surrounded by the native hydrophobic PDMS surface (gray). Panel B shows the cell patterning occurring on the PEMs. The green color of the cells indicates that they are viable while attached on the PAH/PSS PEMs (11 layers). The dark areas are where the native PDMS is exposed (not covered by the PEMs). Panel C shows the velocity profile of fluorescent dye that was induced to become fluorescent upon exposure to a focused UV light source. In sequential images, a–f, the profile of the fluorescent dye is imaged as it moves through a microfluidic channel that was negatively charged on one side and positively charged (due to PEM deposition) on the other side. The images demonstrate flow in two directions in a single microfluidic channel. (Panels A and B, from Ref. [25]; Panel C, from Ref. [103].)

porous shell around the cell. Greater than 98% of the cells in suspension were successfully coated, and cells with four to six layers of polyelectrolytes were found to be intact and retained normal metabolic behavior. Encapsulated cells were able to divide, although at a slower rate than the uncoated yeast cells under the same conditions. It was postulated that the

polyelectrolyte shell hindered the first cell division, causing a reduction in cell proliferation.

Other bioapplications of PEMs include the deposition of biomolecules, such as enzymes and antigens, intercalated in multilayers. Preparation of active enzymes in thin films has been shown for horseradish peroxidase with a bipolar quaternary ammonium as

the polycation. In this example, up to 16 bilayers were deposited, and the activity of the enzyme was electrochemically measured using a gold electrode laying on the substrate surface.^[104] Another example of the combination of biomolecules with the multilayers' deposition is the fabrication of an immunosensor based on an antibody linked to a gold particle deposited alternately with cysteamine on a gold electrode. Using cyclic voltammetry and impedance spectroscopy measurements, this sensor showed a linear range from 5 to 400 $\mu\text{g/L}$ for the detection of human IgG and a detection limit of about 0.5 $\mu\text{g/L}$.^[105] The entrapment of proteins in hollow polyelectrolyte microcapsules is another interesting example of a biological application. A polyelectrolyte hollow structure can be obtained by deposition of multilayers on colloidal particles followed by the dissolution of the core particle.^[106] This provides an empty PEM cage or capsule with a thickness in the range of 5–50 nm. Using this approach, proteins like α -chymotrypsin have been encapsulated in microcapsules fabricated from polyions such as PAH and PSS.^[107,108] The encapsulated proteins demonstrated high enzymatic activity, as well as thermal and mechanical stability. This microcapsule serves as protection against high molecular weight inhibitors and also serves as a bioreactor.

Microfluidic Flow Control

One of the earliest uses of PEMs in microfluidics was for flow modulation in electroosmotically pumped systems. Electroosmotic pumping is the most common method for pumping solutions in a microfluidic format because of its unique ability to control very low-volume flow rates. Electroosmotic flow is flow that occurs when an electric field is applied across a microchannel or capillary that is filled with buffer solution. Electroosmotic flow is modulated by several parameters including buffer concentration, electric field strength, wall charge (positive or negative), and wall surface charge density. For instance, a higher field strength or a higher surface charge density results in faster flows, keeping all other conditions constant. Importantly, if the wall charge is positive, the ions that are attracted to the wall forming the double layer are negative, resulting in a flow toward the positive electrode. If the wall charge is negative, the ions attracted to the wall are positive, resulting in a bulk fluid flow toward the negative electrode.

As described previously, PEMs can be adsorbed to a variety of substrate materials that are used to fabricate microfluidic channels including glass, silicon, poly(methyl methacrylate), poly(carbonate), poly(styrene), and PDMS. In a microfluidic device, PEMs can be adsorbed to the surface of the microfluidic channel and

thus may be used to modulate and control EOF. For example, microfluidic devices fabricated in different types of polymer materials exhibit different EOF rates when all factors other than the material are held constant. However, when these polymer microchannels are coated with the same PEMs, the EOF rates in the different materials are altered and become the same.^[109] Laminar flow patterning can also be used to micro pattern one side of a microfluidic channel with PEMs while leaving the other side of the microchannel uncoated in its native state. If the native state of the polymer microchannel is negatively charged and the PEM coated side of the microchannel is positively charged, electroosmotic fluid flow can be propagated in the microchannel in two directions simultaneously (Fig. 2C).^[103] Polyelectrolyte multilayer coatings can be used in microfluidic devices to create unique microflow patterns, increase or decrease EOF, and eliminate differences in flow behavior in microchannels made from different materials.

CONCLUSIONS

Owing to the unique characteristics of PEM films as well as their simple self-assembly deposition processes, these nanoscale structures have found broad application in widely different fields ranging from electronics to analytical chemistry to biomaterials. Although this chapter is not comprehensive in covering all applications of PEMs, it should provide the reader with examples that demonstrate the enormous potential of these films. For other reviews on PEM materials, we refer the reader to Decher and coworkers and Bertrand et al.^[5,15,110]

ARTICLES OF FURTHER INTEREST

Self-Assembled Monolayers; Surface Coatings; Surface Modification; Surface Topography and Cell Behavior

REFERENCES

1. Decher, G.; Hong, J.D.; Schmitt, J. Buildup of ultrathin multilayer films by a self-assembly process. 3. Consecutively alternating adsorption of anionic and cationic polyelectrolytes on charged surfaces. *Thin Solid Films* **1992**, *210*, 831–835.
2. Decher, G.; Schmitt, J. Fine-tuning of the film thickness of ultrathin multilayer films composed of consecutively alternating layers of anionic and cationic polyelectrolytes. *Prog. Colloid Polym. Sci.* **1992**, *89*, 160–164.
3. Iler, R.K. Multilayers of colloidal particles. *J. Colloid Interface Sci.* **1966**, *21*, 569–594.

4. Ladam, G.; Schaad, P.; Voegel, J.C.; Schaaf, P.; Decher, G.; Cuisinier, F. In-situ determination of the structural-properties of initially deposited polyelectrolyte multilayers. *Langmuir* **2000**, *16*, 1249–1255.
5. Decher, G.; Schlenoff, J.B. In *Multilayer Thin Films*; Decher, G., Schlenoff, J.B., Eds.; Wiley-VCH Verlag: Weinheim, 2003.
6. Schlenoff, J.B.; Dubas, S.T. Mechanism of polyelectrolyte multilayer growth-charge overcompensation and distribution. *Macromolecules* **2001**, *34*, 592–598.
7. Sukhorukov, G.B.; Schmitt, J.; Decher, G. Reversible swelling of polyanion/polycation multilayer films in solutions of different ionic strength. *Bunsenges. Phys. Chem.* **1996**, *100*, 948–953.
8. Decher, G.; Eckle, M.; Schmitt, J.; Struth, B. Layer-by-layer assembled multicomposite films. *Curr. Opin. Colloid Interface Sci.* **1998**, *3*, 32–39.
9. Solis, F.J.; Olvera de la Cruz, M. Surface-induced layer formation in polyelectrolytes. *J. Chem. Phys.* **1999**, *110*, 11517–11522.
10. Netz, R.R.; Joanny, J.-F. Adsorption of semiflexible polyelectrolytes on charged planar surfaces: charge compensation, charge reversal, and multilayer formation. *Macromolecules* **1999**, *32*, 9013–9025.
11. Castelnovo, M.; Joanny, J.-F. Formation of polyelectrolyte multilayers. *Langmuir* **2000**, *16*, 7524–7532.
12. Messina, R. Polyelectrolyte multilayering on a charged planar surface. *Macromolecules* **2004**, *37*, 621–629.
13. Schlenoff, J.B.; Ly, H.; Li, M. Charge and mass balance in polyelectrolyte multilayers. *J. Am. Chem. Soc.* **1998**, *120*, 7626–7634.
14. Schmitt, J.; Grunewald, T.; Decher, G.; Pershan, P.S.; Kjaer, K.; Lösche, M. Internal structure of layer-by-layer adsorbed polyelectrolyte films—a neutron and x-ray reflectivity. *Macromolecules* **1993**, *26*, 7058–7063.
15. Decher, G. Fuzzy nanoassemblies: toward layered polymeric multicomposites. *Science* **1997**, *277*, 1232–1237.
16. Losche, M.; Schmitt, J.; Decher, G.; Bouwman, W.G.; Kjaer, K. Detailed structure of molecularly thin polyelectrolyte multilayer films on solid substrates as revealed by neutron reflectometry. *Macromolecules* **1998**, *31*, 8893–8906.
17. Decher, G.; Lvov, Y.; Schmitt, J. Proof of multilayer structural organization in self-assembled polycation-polyanion molecular films. *Thin Solid Films* **1994**, *244*, 772–777.
18. Michel, M.; Vautier, D.; Voegel, J.-C.; Schaaf, P.; Ball, V. Layer by layer self-assembled polyelectrolyte multilayers with embedded phospholipid vesicles. *Langmuir* **2004**, *20*, 4835–4839.
19. Ladam, G.; Gergely, C.; Senger, B.; Decher, G.; Voegel, J.-C.; Schaaf, P.; Cuisinier, J.G. Protein interactions with polyelectrolyte multilayers: interactions between human serum albumin and polystyrene sulfonate/polyallylamine multilayers. *Biomacromolecules* **2000**, *1*, 674–687.
20. Kumar, G.; Wang, Y.C.; Co, C.; Ho, C.-C. Spatially controlled cell engineering on biomaterials using polyelectrolytes. *Langmuir* **2003**, *19*, 10550–10556.
21. Diaspro, A.; Silvano, D.; Krol, S.; Cavalleri, O.; Gliozzi, A. Single living cell encapsulation in nano-organized polyelectrolyte shells. *Langmuir* **2002**, *18*, 5047–5050.
22. Richert, L.; Lavalle, Ph.; Vautier, D.; Senger, B.; Stoltz, J.-F.; Schaaf, P.; Voegel, J.-C.; Picart, C. Cell interactions with polyelectrolyte multilayer films. *Biomacromolecules* **2002**, *3*, 1170–1178.
23. Yang, S.Y.; Mendelsohn, J.D.; Rubner, M.F. New class of ultrathin, highly cell-adhesion-resistant polyelectrolyte multilayer with micropatterning capabilities. *Biomacromolecules* **2003**, *4*, 987–994.
24. Berg, M.C.; Yang, S.Y.; Hammond, P.T.; Rubner, M.F. Controlling mammalian cell interactions on patterned polyelectrolyte multilayer surfaces. *Langmuir* **2004**, *20*, 1362–1368.
25. Reyes, D.R.; Perruccio, E.M.; Becerra, S.P.; Locascio, L.E.; Gaitan, M. Micropatterning neuronal cells on polyelectrolyte multilayers. *Langmuir* **2004**, *20*, 8805–8811.
26. Yang, L.; Liang, H.; Angelilni, T.E.; Butler, J.; Coridan, R.; Tang, J.X.; Wong, G.C. Self-assembled virus-membrane complexes. *Nat. Mater.* **2004**, *3*, 615–619.
27. Iler, R.K. Multilayers of colloidal particles. *J. Colloid Interface Sci.* **1966**, *21*, 569–594.
28. Berg, M.C.; Choi, J.; Hammond, P.T.; Rubner, M.F. Tailored micropatterns through weal: polyelectrolyte stamping. *Langmuir* **2003**, *19*, 2231–2237.
29. Berg, M.C.; Yang, S.Y.; Hammond, P.T.; Rubner, M.F. Controlling mammalian cell interactions on patterned polyelectrolyte multilayer surfaces. *Langmuir* **2004**, *20*, 1362–1368.
30. Hammond, P.T.; Whitesides, G.M. Formation of polymer microstructures by selective deposition of polymer multilayers using patterned self-assembled monolayers as a template. *Macromolecules* **1995**, *28*, 7569–7571.
31. Zheng, H.P.; Berg, M.C.; Rubner, M.F.; Hammond, P.T. Controlling cell attachment selectively onto biological polymer-colloid templates using polymer-on-polymer stamping. *Langmuir* **2004**, *20*, 7215–7222.
32. Jang, H.; Kim, S.; Char, K. Multilayer line micropatterning using convective self-assembly in microfluidic channels. *Langmuir* **2003**, *19*, 3094–3097.
33. Kleinfeld, E.R.; Ferguson, G.S. Stepwise formation of multilayered nanostructural films from macromolecular precursors. *Science* **1994**, *265*, 370–373.
34. Losche, M.; Schmitt, J.; Decher, G.; Bouwman, W.G.; Kjaer, K. Detailed structure of molecularly thin polyelectrolyte multilayer films on solid substrates as revealed by neutron reflectometry. *Macromolecules* **1998**, *31*, 8893–8906.
35. Fahrhat, T.; Yassin, G.; Dubas, S.T.; Schlenoff, J.B. Water and ion pairing in polyelectrolyte multilayers. *Langmuir* **1999**, *15*, 6621–6623.
36. Muller, M.; Rieser, T.; Lunkwitz, K.; Berwald, S.; Meier-Haack, J.; Jehnichen, D. An in-situ ATR-FTIR study on polyelectrolyte multilayer assemblies on solid surfaces and their susceptibility to fouling. *Macromol. Rapid Commun.* **1998**, *19*, 333–336.



37. Muller, M.; Rieser, T.; Lunkwitz, K.; Meier-Haack, J. Polyelectrolyte complex layers: a promising concept for antifouling coatings verified by in-situ ATR-FTIR spectroscopy. *Macromol. Rapid Commun.* **1999**, *20*, 607–611.
38. Watanabe, S.; Regen, S.L. Dendrimers as building-blocks for multilayer construction. *J. Am. Chem. Soc.* **1994**, *116*, 8855–8856.
39. Keller, S.W.; Kim, H.-N.; Mallouk, T.E. Layer-by-layer assembly of intercalation compounds and heterostructures on surfaces—toward molecular beaker epitaxy. *J. Am. Chem. Soc.* **1994**, *116*, 8817–8818.
40. Hakkarainen, S.; Gilbert, S.L.; Kontturi, A.K.; Kontturi, K. Amperometric method for determining the degree of complexation of polyelectrolytes with cationic surfactants. *J. Colloid Interface Sci.* **2004**, *272*, 404–410.
41. Panambur, G.; Zhang, Y.B.; Yesayan, A.; Galstian, T.; Bazuin, C.G.; Ritcey, A.M. Preparation and characterization of polyion-complexed Langmuir-Blodgett films containing an NLO chromophore. *Langmuir* **2004**, *20*, 3606–3615.
42. Huang, H.X.; Qian, D.J.; Nakamura, N.; Nakamura, C.; Wakayama, T.; Miyake, J. Quartz crystal microbalance and electrochemical studies on the electrode modified by layer-by-layer multilayers of viologen polyelectrolytes. *Electrochim. Acta* **2004**, *49*, 1491–1498.
43. Hoffmannova, H.; Fermin, D.; Krtil, P. Growth and electrochemical activity of the poly-L-lysine|poly-L-glutamic acid thin layer films: an EQCM and electrochemical study. *J. Electroanal. Chem.* **2004**, *562*, 261–265.
44. Serpe, M.J.; Lyon, L.A. Optical and acoustic studies of pH-dependent swelling in microgel thin films. *Chem. Mater.* **2004**, *16*, 4373–4380.
45. Zhu, H.G.; Ji, J.; Shen, J.C. Biomacromolecules electrostatic self-assembly on 3-dimensional tissue engineering scaffold. *Biomacromolecules* **2004**, *5*, 1933–1939.
46. Haynie, D.T.; Balkundi, S.; Palath, N.; Chakravarthula, K.; Dave, K. Polypeptide multilayer films: role of molecular structure and charge. *Langmuir* **2004**, *20*, 4540–4547.
47. Welsh, E.R.; Schauer, C.L.; Santos, J.P.; Price, R.R. In situ cross-linking of alternating polyelectrolyte multilayer films. *Langmuir* **2004**, *20*, 1807–1811.
48. Caruso, F.; Niikura, K.; Furlong, D.N.; Okahata, Y. Ultrathin multilayer polyelectrolyte films on gold: construction and thickness determination. 1. *Langmuir* **1997**, *13*, 3422–3426.
49. Muller, M.; Meier-Haack, J.; Schwarz, S.; Buchhammer, H.M.; Eichhorn, E.J.; Janke, A.; Kessler, B.; Nagel, J.; Oelmann, M.; Reihs, T.; Lunkwitz, K. Polyelectrolyte multilayers and their interactions. *J. Adhesion* **2004**, *80*, 521–547.
50. Schoeler, B.; Poptoshev, E.; Caruso, F. Growth of multilayer films of fixed and variable charge density polyelectrolytes: effect of mutual charge and secondary interactions. *Macromolecules* **2003**, *36*, 5258–5264.
51. Schwarz, S.; Nagel, J.; Jaeger, W. Comparison of polyelectrolyte multilayers built up with polydiallyldimethylammonium chloride and poly(ethyleneimine) from salt-free solutions by in-situ surface plasmon resonance measurements. *Macromol. Symp.* **2004**, *211*, 201–216.
52. Dante, S.; Advincula, R.; Frank, C.W.; Stroeve, P. Photoisomerization of polyionic layer-by-layer films containing azobenzene. *Langmuir* **1999**, *15*, 193–201.
53. Keller, S.W.; Johnson, S.A.; Brigham, E.S.; Yonemoto, E.H.; Mallouk, T.E. Photoinduced charge separation in multilayer thin films grown by sequential adsorption of polyelectrolytes. *J. Am. Chem. Soc.* **1995**, *117*, 12879–12880.
54. Cheng, L.; Niu, L.; Gong, J.; Dong, S.J. Electrochemical growth and characterization of polyoxometalate-containing monolayers and multilayers on alkanethiol monolayers self-assembled on gold electrodes. *Chem. Mater.* **1999**, *11*, 1465–1475.
55. Dermody, D.L.; Peez, R.F.; Bergbreiter, D.E.; Crooks, R.M. Chemically grafted polymeric filters for chemical sensors: hyperbranched poly(acrylic acid) films incorporating beta-cyclodextrin receptors and amine-functionalized filter layers. *Langmuir* **1999**, *15*, 885–890.
56. Delcorte, A.; Bertrand, P.; Wischerhoff, E.; Laschewsky, A. Adsorption of polyelectrolyte multilayers on polymer surfaces. *Langmuir* **1997**, *13*, 5125–5136.
57. Laschewsky, A.; Wischerhoff, E.; Bertrand, P.; Delcorte, A. Polyelectrolyte multilayers containing photoreactive groups. *Macromol. Chem. Phys.* **1997**, *198*, 3239–3253.
58. Buscher, K.; Graf, K.; Ahrens, H.; Helm, C.A. Influence of adsorption conditions on the structure of polyelectrolyte multilayers. *Langmuir* **2002**, *18*, 3585–3591.
59. Serpe, M.J.; Lyon, L.A. Optical and acoustic studies of pH-dependent swelling in microgel thin films. *Chem. Mater.* **2004**, *16*, 4373–4380.
60. Zhu, H.G.; Ji, J.; Shen, J.C. Biomacromolecules electrostatic self-assembly on 3-dimensional tissue engineering scaffold. *Biomacromolecules* **2004**, *5*, 1933–1939.
61. Haynie, D.T.; Balkundi, S.; Palath, N.; Chakravarthula, K.; Dave, K. Polypeptide multilayer films: role of molecular structure and charge. *Langmuir* **2004**, *20*, 4540–4547.
62. Welsh, E.R.; Schauer, C.L.; Santos, J.P.; Price, R.R. In situ cross-linking of alternating polyelectrolyte multilayer films. *Langmuir* **2004**, *20*, 1807–1811.
63. Caruso, F.; Niikura, K.; Furlong, D.N.; Okahata, Y. Ultrathin multilayer polyelectrolyte films on gold: construction and thickness determination. 1. *Langmuir* **1997**, *13*, 3422–3426.
64. Schoeler, B.; Poptoshev, E.; Caruso, F. Growth of multilayer films of fixed and variable charge density polyelectrolytes: effect of mutual charge and secondary interactions. *Macromolecules* **2003**, *36*, 5258–5264.
65. Schwarz, S.; Nagel, J.; Jaeger, W. Comparison of polyelectrolyte multilayers built up with polydiallyldimethylammonium chloride and poly(ethyleneimine) from salt-free solutions by in-situ surface plasmon resonance measurements. *Macromol. Symp.* **2004**, *211*, 201–216.

66. Picart, C.; Ladam, G.; Senger, B.; Voegel, J.C.; Schaaf, P.; Cuisinier, F.J.G.; Gergely, C. Determination of structural parameters characterizing thin films by optical methods: a comparison between scanning angle reflectometry and optical waveguide lightmode spectroscopy. *J. Chem. Phys.* **2001**, *115*, 1086–1094.
67. McAloney, R.A.; Sinyor, M.; Dudnik, V.; Goh, M.C. Atomic force microscopy studies of salt effects on polyelectrolyte multilayer film morphology. *Langmuir* **2001**, *17*, 6655–6663.
68. McAloney, R.A.; Dudnik, V.; Goh, M.C. Kinetics of salts-induced annealing of a polyelectrolyte multilayer film morphology. *Langmuir* **2003**, *19*, 3947–3952.
69. Jaber, J.A.; Schlenoff, J.B. Polyelectrolyte multilayers with reversible thermal responsivity. *Macromolecules* **2005**, *38*, 1300–1306.
70. Sullivan, D.M.; Bruening, M.L. Ultrathin, cross-linked polyimide pervaporation membranes prepared from polyelectrolyte multilayers. *J. Membr. Sci.* **2005**, *248*, 161–170.
71. Sukhishvili, S.A.; Granick, S. Layered, erasable polymer multilayers formed by hydrogen-bonded sequential self-assembly. *Macromolecules* **2002**, *35*, 301–310.
72. Schwinte, P.; Voegel, J.C.; Picart, C.; Haikel, Y.; Schaaf, P.; Szalontai, B. Stabilizing effects of various polyelectrolyte multilayer films on the structure of adsorbed/embedded fibrinogen molecules: an ATR-FTIR study. *J. Phys. Chem. B.* **2001**, *105*, 11906–11916.
73. Muller, M.; Rieser, T.; Lunkwitz, K.; Berwald, S.; Meier-Haack, J.; Jehnichen, D. An in-situ ATR-FTIR study on polyelectrolyte multilayer assemblies on solid surfaces and their susceptibility to fouling. *Macromol. Rapid Commun.* **1998**, *19*, 333–336.
74. Heuberger, R.; Sukhorukov, G.; Voros, J.; Textor, M.; Mohwald, H. Biofunctional polyelectrolyte multilayers and microcapsules: control of non-specific and bio-specific protein adsorption. *Adv. Funct. Mater.* **2005**, *15*, 357–366.
75. Picart, C.; Gergely, C.; Arntz, Y.; Voegel, J.C.; Schaaf, P.; Cuisinier, F.J.G.; Senger, B. Measurement of film thickness up to several hundreds of nanometers using optical waveguide lightmode spectroscopy. *Biosensors Bioelectron.* **2004**, *20*, 553–561.
76. Vodouhe, C.; Schmittbuhl, M.; Boulmedais, F.; Bagnard, D.; Vautier, D.; Schaaf, P.; Egles, C.; Voegel, J.C.; Ogier, J. Effect of functionalization of multilayered polyelectrolyte films on motoneuron growth. *Biomaterials* **2005**, *26*, 545–554.
77. Green, R.J.; Frazier, R.A.; Shakesheff, K.M.; Davies, M.C.; Roberts, C.J.; Tendler, S.J.B. Surface plasmon resonance analysis of dynamic biological interactions with biomaterials. *Biomaterials* **2000**, *21*, 1823–1835.
78. Nelson, B.P.; Frutos, A.G.; Brockman, J.M.; Corn, R.M. Near-infrared surface plasmon resonance measurements of ultrathin films. 1. Angle shift and SPR imaging experiments. *Anal. Chem.* **1999**, *71*, 3928–3934.
79. Furst, E.M.; Pagac, E.S.; Tilton, R.D. Coadsorption of polylysine and the cationic surfactant cetyltrimethylammonium bromide on silica. *Ind. Eng. Chem. Res.* **1996**, *35*, 1566–1574.
80. Koper, G.J.M. Optical properties of colloidal films. *Colloids Surf. A Physicochem. Eng. Asp.* **2000**, *165*, 39–57.
81. Voros, J.; Ramsden, J.J.; Csucs, G.; Szendro, I.; De Paula, S.M.; Textora, M.; Spencer, N.D. Optical grating coupler biosensors. *Biomaterials* **2002**, *23*, 3699–3710.
82. Katz, H.E.; Lovinger, A.J.; Johnson, J.; Kloc, C.; Siegrist, T.; Li, W.; Lin, Y.-Y.; Dodabalapur, A. A soluble and air-stable organic semiconductor with high electron mobility. *Nature* **2000**, *404*, 478–480.
83. Burroughes, J.H.; Bradley, D.D.C.; Brown, A.R.; Marks, R.N.; Mackay, K.; Friend, R.H.; Burns, P.L.; Homes, A.B. Light-emitting diodes based on conjugated polymers. *Nature* **1990**, *347*, 539–541.
84. Sirringhaus, H.; Kawase, T.; Friend, R.H.; Shimoda, T.; Inbasekaran, M.; Wu, W.; Woo, E.P. High-resolution inkjet printing of all-polymer transistor circuits. *Science* **2000**, *290*, 2123–2126.
85. Bharathan, J.; Yang, Y. Polymer electroluminescent devices processed by inkjet printing: I. Polymer light-emitting logo. *Appl. Phys. Lett.* **1998**, *72*, 2660–2662.
86. Fou, A.C.; Onitsuka, O.; Ferreira, M.; Rubner, M.F. Fabrication and properties of light-emitting diodes based on self-assembled multilayers of poly(phenylene vinylene). *J. Appl. Phys.* **1996**, *79*, 7501–7509.
87. Jiang, X.; Hammond, P.T. Selective deposition in layer-by-layer assembly: functional graft copolymers as molecular templates. *Langmuir* **2000**, *16*, 8501–8509.
88. Stroeve, P.; Vasquez, V.; Coelho, M.A.N.; Rabolt, J.F. Gas transfer in supported films made by molecular self-assembly of ionic polymers. *Thin Solid Films* **1996**, *284–285*, 708–712.
89. Levasalmi, J.-M.; McCarthy, T.J. Poly(4-methyl-L-pentene)-supported polyelectrolyte multilayer films: preparation and gas permeability. *Macromolecules* **1997**, *30*, 1752–1757.
90. Dai, J.; Jensen, A.W.; Mohanty, D.K.; Erndt, J.; Bruening, M.L. Controlling the permeability of multilayered polyelectrolyte films through derivatization, cross-linking, and hydrolysis. *Langmuir* **2001**, *17*, 931–937.
91. Katayama, H.; Ishihama, Y.; Asakawa, N. Stable capillary coating with successive multiple ionic polymer layers. *Anal. Chem.* **1998**, *70*, 2254–2260.
92. Katayama, H.; Ishihama, Y.; Asakawa, N. Stable cationic capillary coating with successive multiple ionic polymer layers for capillary electrophoresis. *Anal. Chem.* **1998**, *70*, 5272–5277.
93. Graul, T.W.; Schlenoff, J.B. Capillaries modified by polyelectrolyte multilayers for electrophoretic separations. *Anal. Chem.* **1999**, *71*, 4007–4013.
94. Rmaile, H.H.; Schlenoff, J.B. Optically active polyelectrolyte multilayers as membranes for chiral separations. *J. Am. Chem. Soc.* **2003**, *125*, 6602–6603.
95. Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walters, P. Cell junctions, cell adhesion, and the extracellular matrix. In *Molecular Biology of the Cell*; 4th. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walters, P., Eds.; Garland Science: New York, 2002; 1065–1125.



96. Yavin, E.; Yavin, Z.J. Attachment and culture of dissociated cells from rat embryo cerebral hemispheres on polylysine-coated surface. *J. Cell Biol.* **1974**, *62*, 540–546.
97. Letourneau, P.C. Possible roles for cell-to-substratum adhesion in neuronal morphogenesis. *Dev. Biol.* **1975**, *44*, 77–91.
98. McKeehan, W.L.; Ham, R.G. Stimulation of clonal growth of normal fibroblasts with substrata coated with basic polymers. *J. Cell Biol.* **1976**, *71*, 727–734.
99. Freshney, R.I. *Culture of Animal Cells: A Manual of Basic Techniques*, 4th Ed.; Wiley-Liss: New York, 2000.
100. Brandley, B.K.; Schnaar, R.L. Covalent attachment of an Arg-Gly-Asp sequence peptide to derivatizable polyacrylamide surfaces—support of fibroblast adhesion and long-term growth. *Anal. Biochem.* **1988**, *172*, 270–278.
101. Tryoen-Toth, P.; Vautier, D.; Haikel, Y.; Voegel, J.C.; Schaaf, P.; Chluba, J.; Ogier, J. Viability, adhesion, and bone phenotype of osteoblast-like cells on polyelectrolyte multilayer films. *J. Biomed. Mater. Res.* **2002**, *60*, 657–667.
102. Vautier, D.; Hemmerle, J.; Vodouhe, C.; Koenig, G.; Richert, L.; Picart, C.; Voegel, J.C.; Debry, C.; Chluba, J.; Ogier, J. 3-D surface charges modulate protrusive and contractile contacts of chondrosarcoma cells. *Cell Motil. Cytoskeleton* **2003**, *56*, 147–158.
103. Barker, S.L.R.; Ross, D.; Tarlov, M.J.; Gaitan, M.; Locascio, L.E. Control of flow direction in microfluidic devices with polyelectrolyte multilayers. *Anal. Chem.* **2000**, *72*, 5925–5929.
104. Cui, L.; Tang, J.; Li, W.J.; Wang, Z.C.; Sun, C.Q.; Zhao, M.Y. Preparation of catalytically active enzyme thin film by alternate deposition of horseradish peroxidase and bipolar quaternary ammonium on solid surface. *Mater. Chem. Phys.* **2001**, *71*, 23–27.
105. Wang, M.; Wang, L.; Yuan, H.; Ji, X.; Sun, C.; Ma, L.; Bai, Y.; Li, T.; Li, J. Immunosensors based on layer-by-layer self-assembled Au colloidal electrode for the electrochemical detection of antigen. *Electroanalysis* **2004**, *16*, 757–764.
106. Donath, E.; Sukhorukov, G.B.; Caruso, F.; Davis, S.A.; Mohwald, H. Novel hollow polymer shells by colloid-templated assembly of polyelectrolytes. *Angew. Chem.* **1998**, *37*, 2201–2205.
107. Tiourina, O.P.; Antipov, A.A.; Sukhorukov, G.B.; Larionova, N.I.; Lvov, Y.; Mohwald, H. Entrapment of -chymotrypsin into hollow polyelectrolyte microcapsules. *Macromol. Biosci.* **2001**, *1*, 209–214.
108. Balabushevitch, N.G.; Sukhorukov, G.B.; Moroz, N.A.; Volodkin, D.V.; Larionova, N.I.; Donath, E.; Mohwald, H. Encapsulation of proteins by layer-by-layer adsorption of polyelectrolytes onto protein aggregates: factors regulating the protein release. *Biotechnol. Bioeng.* **2001**, *76*, 207–213.
109. Barker, S.L.R.; Tarlov, M.J.; Canavan, H.; Hickman, J.J.; Locascio, L.E. Plastic microfluidic devices modified with polyelectrolyte multilayers. *Anal. Chem.* **2000**, *72*, 4899–4903.
110. Bertrand, P.; Jonas, A.; Laschewsky, A.; Legras, R. Ultrathin polymer coatings by complexation of polyelectrolytes at interfaces: suitable materials, structure and properties. *Macromol. Rapid Commun.* **2000**, *21*, 319–348.

Polyepoxy Compound Fixation

Yasuharu Noishiki

Department of Surgery, Yokohama City University School of Medicine, Kanazawa-ku, Yokohama, Japan

Teruo Miyata

Koken Bioscience Institute, Koken Co. Ltd., Toshima-ku, Tokyo, Japan

INTRODUCTION

Characteristic advantages of polyepoxy compounds (PCs) as a cross-linker for biological materials are explained in this article. Glutaraldehyde (GA) has been used for fixation of biological materials in the fields of artificial organs, but GA has both advantages and disadvantages. Compared with GA, PCs have various advantages with little disadvantages. With these peculiarities, PCs are anticipated as a new cross-linker for biological materials. In this article, basic chemical reactions and mechanisms of PC cross-linking will be explained in detail as PCs are used in the fields of biological materials. Some examples of artificial organs cross-linked with PCs are also demonstrated.

CROSS-LINKING OF BIOLOGICAL MATERIALS FOR ARTIFICIAL ORGANS

Before the explanation of PCs,^[1,2] biological materials for artificial organs will be discussed. Numerous synthetic polymer materials have been used in the medical field and many artificial organs with sufficient mechanical strength and desirable shapes and sizes have been created. But, implantable materials with very fine mechanical and structural properties for host-cell migration and proliferation in order to create new hybrid artificial organs or tissue-engineered systems cannot be produced from synthetic materials. Biological materials have an extremely fine structure and unique properties that cannot be imitated with synthetic polymer materials. Therefore, some biological materials obtained from animal tissues have been used as raw materials for implantable artificial organs.^[3-6]

Unless they are autologous, biological materials for implantation need some modification such as chemical cross-linking in order to give toughness and durability in vivo, to prevent rapid biodegradability and to reduce antigenicity. Several chemical reagents, such as glutaraldehyde (GA), formaldehyde, hexamethylene diisocyanate, dialdehyde starch, etc. have been used for the fixation of biological materials. Among them, GA is the most frequently used.^[7-11] GA alone has been used

especially for the fixation of materials and products that are used for permanent implants such as biological heart valves.^[12-15] After fixation with GA, materials become mechanically stronger than the original natural tissue. Fixed materials are resistant to collagenase digestion.

Most biological tissues obtained from mammals contain collagen fibers. Because of collagenase digestion, the framework of the implanted tissue is destroyed, resulting in absorption of the tissue in vivo, which is a natural reaction to a foreign body. Collagenase digestion is the most common fate suffered by implanted biological materials. After GA fixation, not only collagen fibers but also most other tissue components become resistant to enzymatic digestion and unexpected reactions such as immunogenicity are reduced. However, GA has also severe side effects such as cytotoxicity, which cannot be ignored in implantable biomedical materials.^[16,17] For this reason, we needed to develop a new system of cross-linking. We introduced PC fixation technology into the medical field in 1986 in order to overcome this problem.^[1]

In this article, we would like to explain the principle of PC cross-linking and to introduce its applications to heart valves, blood vessel prostheses, antiadhesive membranes, and hemostatic materials.

PC AS A CROSS-LINKER FOR BIOLOGICAL MATERIALS

PCs are now widely used as cross-linking agents, adhesion promoters, stabilizers of textile and paper, and for many other industrial purposes.^[18,19] Among them, water soluble PCs are useful for the treatment of biological materials. High-quality water-soluble PCs are commercially available under the trade name Denacol, which is produced by Nagase ChemteX Corporation (Osaka, Japan).^[20] In this article, the water-soluble PCs, i.e., Denacol, will be introduced in detail.

Nagase ChemteX Corporation has developed and characterized PCs that include monofunctional and polyfunctional epoxy compounds, called by the generic name Denacol. Denacol, especially its polyfunctional



epoxy compounds, is not only water soluble but also highly reactive (the epoxy group of Denacol reacts easily with amino, carboxyl, and hydroxyl groups). With this water solubility and reactivity Denacol can be used as a cross-linking agent in water-based systems. Denacol has been applied in such industrial fields as textile finishing, paper finishing, paints, and adhesives. Denacol can be also used as a cross-linking agent and insolubilizing agent for various water dispersible or soluble high polymers, such as acrylic resin, polyvinyl-alcohol, latex, casein, starch, polyester, and polyurethane, which have functional groups containing active hydrogen, such as carboxyl, amino, and hydroxyl groups. Many properties such as water resistance, solvent resistance, adhesiveness, weathering property, and thermal resistance can be improved by cross-linking with Denacol.

The most representative Denacol compounds are sorbitol polyglycidyl ether, diglycerol polyglycidyl ether, glycerol polyglycidyl ether, trimethylolpropane polyglycidyl ether, neopentylglycol diglycidyl ether, 1,6-hexanediol diglycidyl ether, ethylene polyethylene glycol diglycidyl ether, and propylene polypropylene glycol diglycidyl ether. Their chemical structures and properties are shown in Table 1.^[20]

As shown in these structures, they have more than two epoxy groups in their molecular structure. These polyfunctional sites can work to cross-link biological materials in water-based systems, especially under physiological conditions.

These epoxy groups can react effectively under various conditions when certain catalysts are used. Table 2 shows examples of typical reactions with some reactants.^[20] Reaction speed and properties vary substantially according to the compound structure, solvents, catalysts, and reaction conditions such as temperature, pH, concentration, and duration. Details of epoxy group reaction and reactivity with various functional groups are given in a technical data sheet^[20] provided by Nagase ChemteX Corporation.

PC CROSS-LINKING OF COLLAGEN MOLECULES

PCs can cross-link not only collagen molecules, but also many other components of biological materials. We introduced the PC cross-linking technology in the medical field. Detailed information related to the application of PCs in the medical field is described elsewhere.^[21–26] In this article, we will explain the reaction of PCs with collagen fibers as an example, since most biological materials obtained from animals contain a high amount of collagen fibers.

Collagen is produced and released by cells such as fibroblasts and is a triple helical rodlike protein.

Collagen molecules contain three polypeptide chains (called alpha-chains) coiled in a triple helix. The amino acids of these polypeptide chains have characteristic features. Their most characteristic sites are the $-\text{COOH}$ and $\epsilon\text{-NH}_2$ groups. Figure 1 shows schematically a representative cross-linking reaction of PC with two collagen molecules, i.e., cross-linking. As shown in the Fig. 1, epoxy groups can react with both $-\text{COOH}$ and $\epsilon\text{-NH}_2$ groups. Before cross-linking, the total number of $\epsilon\text{-NH}_2$ groups is higher than the number of $-\text{COOH}$ groups. Therefore, collagen molecules have a positive charge under physiological conditions. Although it depends on the reaction conditions, $\epsilon\text{-NH}_2$ groups are more highly reactive than $-\text{COOH}$ groups. Therefore, most $\epsilon\text{-NH}_2$ groups are used up in the reaction. After the reaction, the total number of $\epsilon\text{-NH}_2$ groups is smaller than the number of $-\text{COOH}$ groups, and the collagen molecules have a negative charge under physiological conditions, and become less thrombogenic, because platelets have a negative charge, and repulsion of their negative charge prevents platelet aggregation on treated collagen molecules.

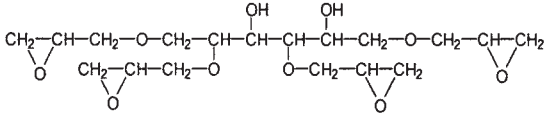
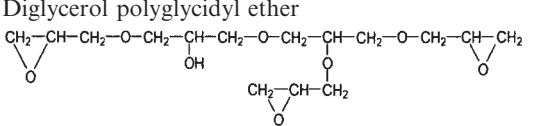
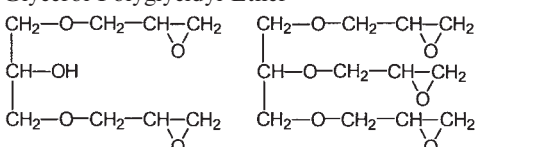
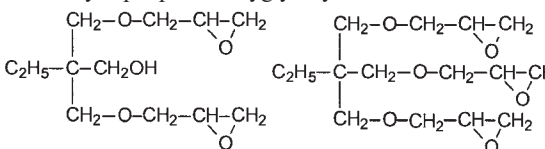
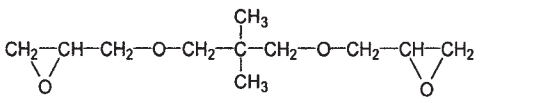
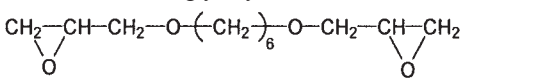
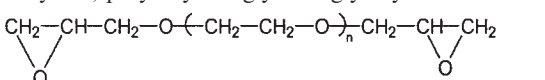
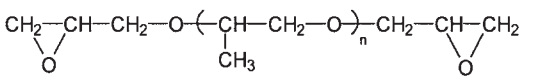
COMPARISON OF GA WITH PCs

Most biological materials for medical use have been cross-linked with GA. As described above, GA treatment has many advantages, but it also has disadvantages. The greatest advantage of the GA treatment is the stability of its reaction.^[7–9] The treatment makes materials stable *in vivo* after implantation. Materials cross-linked with GA show no biodegradability at all for a long period of time after implantation. The reaction can reduce the antigenicity of animal tissues.^[8,9]

As an effect, GA makes materials stiff. Their color becomes yellowish. Materials lose their original elasticity and softness. A major problem is the cytotoxicity of GA. GA is highly cytotoxic and induces various problems.^[16,17] After implantation, nonreactive GA molecules ooze out continuously for a long period of time from the cross-linked materials.^[17] GA molecules tend to polymerize themselves. Polymerized molecules are also the source of leaking nonreactive monomers. These are cytotoxic and damage surrounding cells *in vivo*.^[17]

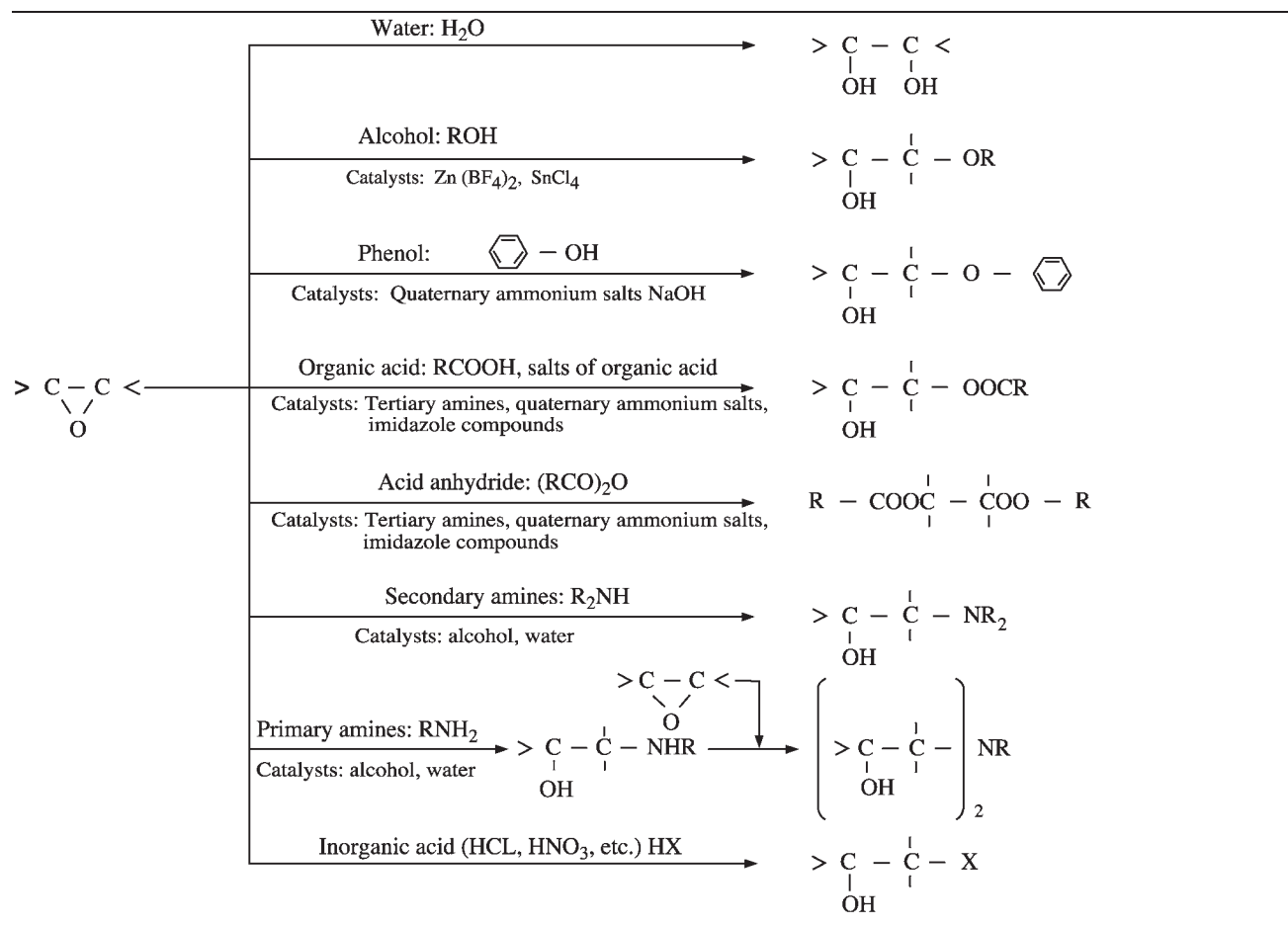
Host cells cannot migrate or proliferate on GA cross-linked materials. The materials cannot obtain cell compatibility. Sometimes, this side effect is desirable. In the case of cross-linked heart valves, host cells cannot adhere to the surface, and this can prevent pannus hyperplasia over the heart valve leaflets. But in the case of artificial blood vessels and pericardium, surrounding host cells cannot migrate into the cross-linked tissues, resulting in poor connection of the graft to the surrounding tissues.

Table 1 Chemical structures and properties of denacol compounds

Product name (TSCA no.)	WPE	Color	Viscosity	Water solubility (%)	Chemical name and structure
EX-611	170	5(G)	11,500	50	Sorbitol polyglycidyl ether
EX-612	170	2(G)	11,500	50	
EX-614	166	3(G)	17,000	78	
EX-614B	180	3(G)	4,000	90	
EX-622 (68412-01-1)	198	2(G)	10,000	Insoluble	
EX-421 (68134-62-3)	155	4(G)	500	81	Diglycerol polyglycidyl ether 
EX-313	141	20(A)	155	99	Glycerol Polyglycidyl Ether
EX-314 (25038-04-4)	145	20(A)	175	80	
EX-321 (30499-70-8)	145	20(A)	135	20	Trimethylolpropane Polyglycidyl Ether 
EX-211 (17557-23-2)	140	20(A)	14	26	Neopentylglycol diglycidyl ether 
EX-212 (16096-31-4)	150	30(A)	20	Insoluble	1,6-Hexanediol diglycidyl ether 
EX-810 (n = 1)	112	30(A)	15	100	Ethylene, polyethylene glycol diglycidyl ether
EX-811(n = 1) (2224-15-9)	135	10(A)	15	95	
EX-850(n = 2)	121	60(A)	14	100	
EX-851(n = 2)	154	10(A)	26	99	
EX-821(n = 4)	195	20(A)	66	100	
EX-830(n = 9)	262	100(A)	66	100	
EX-832(n = 9)	280	40(A)	100	100	
EX-841(n = 13)	382	40(A)	164	100	
EX-861(n = 22) (72207-80-8)	587	100(A)	mp.45°C	100	
EX-911(n = 1)	165	20(A)	20	75	Propylene, polypropylene glycol diglycidyl ether
EX-941(n = 2)	172	60(A)	22	80	
EX-920(n = 3)	180	100(A)	17	100	
EX-931(n = 11) (26142-30-3)	528	30(A)	120	20	

(Note) WPE: Weight per epoxy equivalent; Viscosity: mPa · S at 25°C; Color: (G) means Gardner Color, (A) means APHA color.

Table 2 Reaction of epoxy group



Usually, biological materials are very soft and elastic except for bone and other hard tissues. Materials cross-linked with GA become strong, but also stiff. This stiffness causes trouble when the materials are used at sites that require elasticity and softness. Some of these disadvantages stem from the molecular structure of GA. As shown in Fig. 2, GA is composed of a chain of five carbons.^[16] It is extremely stiff, and it is also hydrophobic. Usually, most biological materials are hydrophilic. The original properties are changed after cross-linking with GA. The chain of five carbons changes the original soft and hydrophilic properties into stiff and hydrophobic properties. This change is caused by the molecular structure of GA.

The molecular structure of PCs is different from that of GA. Some of the disadvantages mentioned previously can be overcome with the use of PCs. As shown in Fig. 2, PCs have ether bonds in their molecular structure.^[20] A nickname for the ether bond is free joint, and the molecules of PCs are flexible. Materials maintain their original softness and elasticity after cross-linking.^[1,2]

Another advantage of PCs is their hydrophilicity. As shown in Fig. 1, a hydroxyl group is created after reaction of PCs with $\epsilon\text{-NH}_2$ groups and carboxyl

groups.^[20] The epoxy circle is opened at the reaction and takes a -H from the $\epsilon\text{-NH}_2$ group of collagen molecules, resulting in the creation of a new -OH group at each reaction site. This gives the materials hydrophilicity. Usually, with this chemical reaction, PC cross-linking maintains the original hydrophilicity of the materials. Cross-linking with GA makes the materials hydrophobic. This is against nature.

Calcification of materials cross-linked with GA is another large disadvantage.^[27,28] The mechanism of calcification is not yet known, but GA-treated materials easily become calcified, especially when they are implanted in young patients. In the case of heart valves, a very high ratio of calcification has been reported in biological heart valves implanted in child patients.^[27,28] Experimentally, calcification is easily demonstrated in a subcutaneous implantation model in growing animals. Within three weeks, GA-cross-linked materials become calcified, but with PC cross-linking, the calcification ratio is extremely low. Imamura and his colleagues demonstrated a low calcification ratio in a PC cross-linked heart valve experiment, a subcutaneous implantation model with PC cross-linking compared with GA cross-linking.^[29] Okoshi and his colleagues also

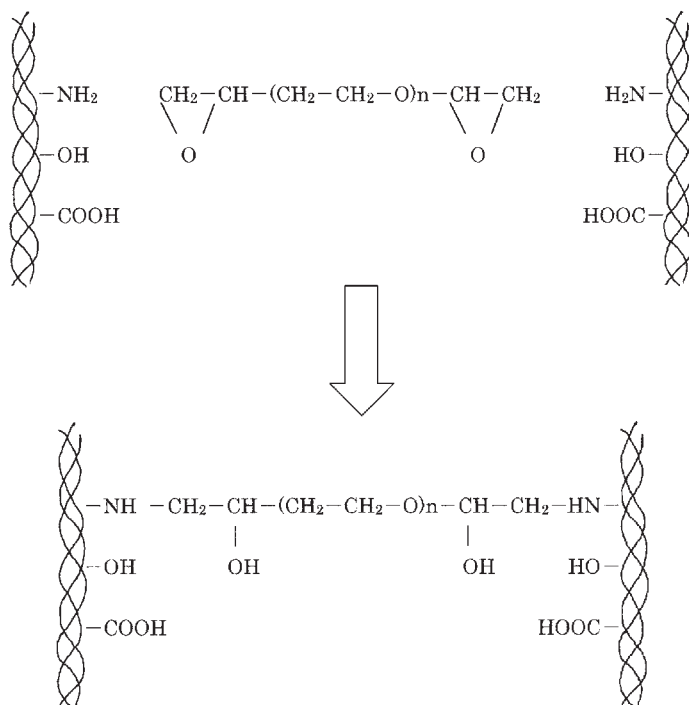


Fig. 1 Cross-linking of collagen molecules with a polyepoxy compound.

obtained similar results with the materials cross-linked with PC and GA. Less than 3% of the materials cross-linked with PC showed calcification compared with 100% for GA cross-linking.^[30]

Fine controlability of cross-linking ratio is also one of the advantages of PCs. Compared with GA, control of the reaction rate with PC cross-linking is easy.^[20] The reaction proceeds smoothly and relatively slowly compared with GA. The reaction with PC can be accelerated using catalysts and control of reaction conditions such as pH, temperature, reaction time, and concentration of the cross-linker. Therefore, the reaction ratio and speed of PC cross-linking is controllable.

Related to the cytotoxicity of cross-linking reagent, PC is less cytotoxic than GA. Besides, nonreactive epoxy groups can be easily reduced by heating in water, or adding some small-molecule amine solutions such as glycine. Therefore, the cytotoxicity of materials cross-linked with PCs is extremely low compared with that of materials treated with GA.

Murayama et al. studied the reduction of antigenicity and immunogenicity both quantitatively and

qualitatively in dogs and mice. They reported that treatments with either GA or PC reduced antigenicity and immunogenicity equally.^[31,32]

APPLICATION OF PC CROSS-LINKING TO THE FIELD OF HEART VALVES

Imamura created porcine heart valves cross-linked with PC.^[29,33] Their heart valves were elastic and soft, and there was no calcification during implantation.^[34-36] We also created a right ventricular-to-pulmonary artery conduit having a valve made from the jugular vein of a cow.^[37-39] This has a natural venous valve, but we adapted it for use as a pulmonary valve. The conduit was extremely soft and pliable after cross-linking with PC. It was possible to open and close the valve with less than a 5-mmHg pressure difference.^[39]

We have developed a new heparinization technique that uses protamine sulfate and sodium heparin to make biological materials antithrombogenic.^[40,41] Protamine is composed of positively charged amino acids such as lysine and arginine. Heparin contains many carboxyl groups and sulfonic groups, resulting in a negative charge under physiological conditions. Heparin and protamine can bind ionically. We used this ionic bond for heparinization of biological materials. The heart valve experiment performed by Ichikawa and his colleagues adopted our heparinization technology. The valve was implanted in the right ventricle in the pulmonary artery position. The technique prevented thrombus formation and calcification in animal experiments for more than one year. At the time of the heart valve

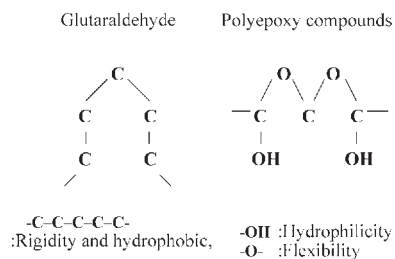


Fig. 2 Molecular structures of GA and PC.

Table 3 Mechanical properties of native and cross-linked arteries

Material	Property		
	Tensile strength (g/mm ²)	Elongation rate (%)	Compliance at 90 mmHg
Native artery	224	100	1.05
GA cross-linked A.	144	86	0.45
PC cross-linked A.	195	126	1.08

retrieval, the valve was extremely soft and elastic without calcification. Detailed data is available in their papers.^[38,39]

APPLICATION TO BLOOD VESSEL PROSTHESES

We developed a new vascular prosthesis made of an animal artery.^[42-46] The graft was heparinized by our own method described above and cross-linked with PC. The following is a summary of our animal experiments.

Canine carotid arteries were soaked in distilled water and then sonicated and squeezed in order to create an acellular matrix for biological vascular grafting. With this treatment, a tubular framework composed of elastic laminae and collagen fibers was obtained. A protamine sulfate solution was poured into the lumen of the tube. Then the tube was cross-linked with PC from the outer surface of the tube, which made the graft white, hydrophilic, soft, pliable, and biodurable. Table 3 shows the elasticity, tensile strength, and compliance of the PC cross-linked graft compared with the control GA cross-linked graft. The graft was dipped into a heparin solution, which allowed heparin to bind ionically. The cross-linking percentage of the ϵ -NH₂ group of collagen side chains is about 85%. Figure 3 shows a PC cross-linked graft (top) and a control GA cross-linked graft (bottom). The PC cross-linked graft is white, and its wall is thick due to water content. The GA cross-linked graft is yellow and its wall is thin.

Eighty grafts (3 mm internal diameter, 6 cm in length) were implanted in the carotid arteries of 40 dogs

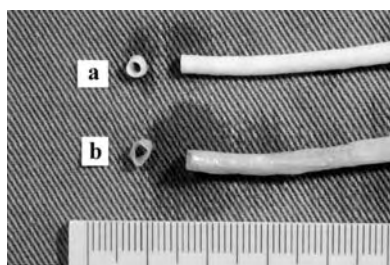


Fig. 3 Canine carotid arteries cross-linked by PC (top) and GA (bottom).

and retrieved from one to 389 days after implantation.^[1] Three grafts were occluded by graft infection and 77 were patent (96% patency). Sixteen control grafts cross-linked with GA were stiff. All of them were found to be occluded at the time of the graft retrieval. Measurement of the heparin content revealed that about 90% of the heparin was released from the graft wall within 1 month. After the release of heparin, the surface of the graft was completely covered with endothelial cells. The antithrombogenicity of the graft was due to the slow release of heparin, permitting the attachment of endothelial cells that have natural antithrombogenicity. The graft wall kept its soft and pliable properties even after long-term implantation.

In specimens examined six months or more after implantation, host cells such as fibroblasts and smooth muscle cells had migrated into the acellular matrix, i.e., the graft wall. Smooth muscle cells near the luminal surface were aligned in parallel to the circumference. Near the adventitia side, they were arranged in a longitudinal direction like a natural arterial wall. These results indicate that the antithrombogenicity produced by the combination of the slow release of heparin and endothelialization, together with the natural vessel compliance in the graft, were the major reason for its success as a small-caliber graft.

We also implanted the grafts in the aorta-coronary bypass position in eight dogs,^[47] and observed the grafts for up to four months. All the grafts were confirmed to be patent by autopsy or graft angiography.

These successful results were possible only with the combined use of the heparinization technique and PC cross-linking. With GA cross-linking, it would be impossible to obtain such good results.

APPLICATION TO A GROWABLE VASCULAR PROSTHESIS

Conventional artificial blood vessels retain their original dimensions, such as thickness, length, and internal diameter, after implantation. Therefore, they cannot be implanted in child or infant patients, because relative stenosis will occur as the child grows. In order to overcome this problem, we developed an expansible vascular

prosthesis that can grow *in vivo* after implantation.^[48,49] Our animal experiments were as follows.

Biological tubes obtained from humans or animals were used as the raw material for an inner graft. The graft was covered with an outer graft made of a polyester fiber mesh. The inner graft was characterized by measuring the rate of reaction of ϵ -NH₂ groups on the side chains in its collagen. The ratio was about 40%. If we used a fast cross-linking rate, the growth of the graft, i.e., inflation of the graft by blood pressure during implantation, could be made slow. If we used a slow cross-linking rate, the graft would grow fast after implantation. The graft growth was controllable with the cross-linking ratio of collagen. We used Denacol EX-861 (a trade name of Nagase ChemteX Corporation).^[20] Denacol EX-861 is a hydrophilic cross-linker with slightly high molecular weight. Therefore, it was easy to control the cross-linking rate. The collagen tube thus obtained (inner diameter 4 mm, length 6 cm) was covered with a polyester mesh (internal diameter 10 mm) and implanted in the descending aorta of four young dogs. The animals weighed from 4.5 to 7.5 kg. The grafts were retrieved from six months to three years after implantation. At the time of graft retrieval, the animals weighed from 8.5 to 10.3 kg. The average inner diameter of the graft was 9 mm. Growth of the graft was limited by the size of the polyester mesh that covered the inner graft. The expansible graft was made possible by the epoxy cross-linking technique, which can control the inflation rate, in conjunction with the outer polyester mesh tube, which can determine the final growth size.

APPLICATION TO A HEMOSTAT

Insoluble collagen of bovine dermis was solubilized by pepsin in acidic solution. Telopeptides, which are nonhelical parts at the terminals of the molecule, were digested by pepsin treatment. This solubilized collagen was named Atelocollagen.^[50] A 3% atelocollagen solution was extruded from a nozzle having 200 70- μ m pores into saturated aqueous sodium-sulfate solution. The resulting fibers were dipped in a PC solution, and washed with water. After freeze-drying the fibers, they were cut into small lengths and entangled with each other to create a cotton-like structure. This cotton-like material was used as a hemostat.^[51,52] Scanning electron microscope observation showed that each collagen fiber had very fine fissures on its surface, which were caused by cross-linking the collagen molecules of soluble collagen fibers. These very fine fissures increased the surface area, so that each fiber could absorb blood easily and could make contact with platelets and fibrinogen, resulting in acceleration of blood coagulation.

As stated above, the absorbing ratio *in vivo* of the collagen fibers can be controlled by varying the degree of PC cross-linking. The ratio of cross-linking, which is measured as the rate of reaction of ϵ -amino groups on the side chains in collagen, is about 30%. A higher amount of cross-linking will slow the absorbing ratio, and a lower amount of cross-linking will increase the absorbing ratio. Accordingly, the *in-vivo* absorbing ratio of the hemostatic material of our experiment can be conveniently adjusted, depending on the condition of the bleeding region where it is to be used. These fine controls of the *in-vivo* absorption are only possible with the use of the atelocollagen spinning technique and the PC cross-linking.

USE IN ANTIADHESIVE MEMBRANES

We also developed an antiadhesive collagen membrane. The membrane was heparinized with our own heparinization technique described previously, and cross-linked with GA. The membrane showed excellent antiadhesive properties on serous membrane peeled off surfaces of the large intestine of dogs. But even two years after operation, the membrane still remained on the colon surface. In order to make the membrane absorbible *in vivo*, we changed the cross-linking method from GA to PC.^[53,54] Thus, we could control the cross-linking rate in order to make the membrane absorbible *in vivo* after its antiadhesive function was finished. The following is a summary of the experiment.

A thin piece of human amnion was obtained aseptically and treated with 0.01% ficin in phosphate buffer, pH 7.4, at room temperature for 24 hours and washed with distilled water repeatedly. The membrane was heparinized with our own method using protamine sulfate impregnation. The membrane was cross-linked with 3% PC Denacol EX-810 (Nagase ChemteX Corporation, Osaka, Japan) for 12 hours. The cross-linking rate was about 30%. Then the membrane was dipped in a 1% heparin solution at pH 6.9 for 24 hours. The membrane was again washed in distilled water repeatedly, and preserved in a 70% solution of ethanol.

Ten mongrel dogs weighing 8–12 kg were used. Six of them were used for the heparinized membrane experiment, and the other four were used as controls. The peritoneal cavity of the animal was entered via a midline incision under general anesthesia. 5 cm² pieces of the colon serous membrane, which consists of a layer of mesothelial cells resting on loose connective tissue, were removed from each of the dogs and pieces of the membranes were sutured onto each site with intermittent sutures at 16 places per site. The animals were killed at from three to 687 days after surgery.

All the heparinized membranes prevented intestinal adhesion. The heparinized membranes were already

absorbed after a period of one month, but after periods of less than one month, the membranes remained on the colon surface. But, no adhesion was observed whether the membranes were absorbed or not. In the control experiment, all the areas covered with the GA-treated membranes had serious omentum adhesion.

The mechanism of the antiadhesion system was as follows. In the early stage of the intestinal adhesion, fibrin deposition on the operation sites always occurs. The fibrin deposit makes a fibrin network, which binds the wound surface to the nearby tissues. This fibrin network can be dissolved by fibrinolysis. But if fibroblasts infiltrate the fibrin network, they will synthesize collagen fibers around them. The collagen network cannot be dissolved by fibrinolysis. Next, capillary blood vessels follow the fibroblast migration. This accelerates fibroblast proliferation and collagen synthesis. In this process, adhesion tissue composed of fibroblasts and collagen fibers is created between the wound and the surrounding tissues.

If we can prevent the creation of a fibrin network, fibroblasts cannot migrate into the network, since they will have no anchoring sites and cannot synthesize collagen fibers between the wound surface and the surrounding tissues. Therefore, no adhesion tissue will be created. In this experiment, the slow release of heparin from the membrane effectively prevented fibrin deposition, thus preventing adhesion. Once adhesion was prevented, the membrane was no longer needed, and therefore it was designed to be dissolved one month after implantation.^[54] The area left after the membrane was dissolved was covered with mesothelial cells that have a natural antiadhesive property. The cross-linking ratio of collagen molecules was easily controlled with the PC cross-linking technique. This experiment showed one successful example of PC cross-linking of biological materials.

CONCLUSION

Compared with GA cross-linking, PC cross-linking proceeds relatively slowly, but it can be accelerated and the reaction rate is controllable. The low cytotoxicity of PC is also a great advantage over GA. This PC cross-linking method is not widely used yet, but in the near future, the method will become a major technique in the field of materials for medical use.

ACKNOWLEDGMENTS

We would like to express our grateful thanks to Dr. S. Shimizu, chief scientist of Nagase ChemteX Corporation, for his invaluable suggestions, and to the company

for supplying all information related to Denacol technology.

ARTICLES OF FURTHER INTEREST

Collagen; Collagen Fixation; Collagen Processing; Tissue Engineering of Blood Vessel; Vascular Grafts; Xenografts

REFERENCES

- Noishiki, Y.; Miyata, Y.; Kodaira, K. Development of a small caliber vascular graft by a new crosslinking method incorporating slow heparin release collagen and natural tissue compliance. *ASAIO Trans.* **1986**, *32*, 114–119.
- Kodaira, K.; Miyata, T.; Furuse, M.; Noishiki, Y. Characterization of collagenous materials cross-linked by polyepoxy compounds. *Jap. J. Artif. Organs* **1986**, *15*, 239–242.
- Dare, W.A.; Lewis, M.R. Further experiences with bovine arterial grafts. *Surgery* **1976**, *80*, 711–721.
- Dardik, I.; Dardik, H. Human umbilical cord vein as an aortic substitute in baboon. A preliminary report. *J. Med. Primatol.* **1973**, *2*, 269–301.
- Tice, D.A.; Zerbino, V. Clinical experience with preserved human allografts for vascular reconstruction. *Surg* **1972**, *72*, 260–267.
- Mazur, P. The freezing of biological systems. *Science* **1970**, *168*, 939–949.
- Bowes, J.H.; Carter, C.W. The reaction of glutaraldehyde with proteins and other biological materials. *J. R. Microsc. Soc.* **1966**, *85*, 193–200.
- Binet, J.P.; Carpentier, A.; Langlois, J.; Duran, C.; Colvez, P. Implantation of heterogenic valves in the treatment of aortic cardiopathies. *C. R. Acad. Sci. Hebd. Seances. Acad. Sci., D* **1965**, *261*, 5733–5734.
- Salgaller, M.L.; Bajpai, P.K. Immunogenicity of glutaraldehyde treated bovine pericardial tissue xenografts in rabbits. *J. Biomed. Mater. Res.* **1985**, *19*, 1–12.
- Hardy, P.M.; Nicholls, A.C.; Rydon, H.N. The nature of the cross-linking of proteins by glutaraldehyde. I. Interaction of glutaraldehyde with the amino-groups of 6-aminohexanoic acid and alpha-N-acetylic-lysine. *J. Chem. Soc., Perkin I* **1976**, *1976*, 958–962.
- Cheung, D.T.; Nimni, M.E. Mechanism of cross-linking of proteins by glutaraldehyde. II. Reaction with monomeric and polymeric collagen. *Connect. Tissue Res.* **1982**, *10*, 201–216.
- Binet, J.P.; Carpentier, A.; Langlois, J. Heterologous transplantation of the aortic valve. *Langenbecks Arch Chir.* **1966**, *316*, 800–802.
- Carpentier, A.; Lemaigre, G.; Robert, L.; Carpentier, S.; Dubost, C. Biological factors affecting long-term results of valvular heterografts. *J Thorac Cardiovasc Surg* **1969**, *58*, 467–483.
- Huges, H.; Tipton, L.S.; About-Enein, H.Y.; Duran, C.M.G. The reaction of glutaraldehyde with pericardial tissue. *J. Heart Valve Dis.* **1994**, *3*, 105–110.

15. Carpentier, A.; Deloche, A.; Relland, J.; Fabiani, J.N.; Forman, J.; Camilleri, J.P.; Soyer, R.; Dubost, C. Six-year follow up of glutaraldehyde preserved heterografts. *J. Thorac. Cardiovasc. Surg.* **1974**, *68*, 771–782.
16. Beauchamp, R.O., Jr.; St Clair, M.B.; Fennell, T.R.; Clarke, D.O.; Morgan, K.T. A critical review of the toxicology of glutaraldehyde. *Crit. Rev. Toxicol.* **1992**, *22*, 143–174.
17. Speer, D.P.; Chvapil, M.; Eskelson, C.D.; Ulreich, J. Biological effects of residual glutaraldehyde in glutaraldehyde-tanned collagen biomaterials. *J. Biomed. Mater. Res.* **1980**, *14*, 753–764.
18. Lee, H.; Nevill, K. *Handbook of Epoxy Resins*; McGraw-Hill Book Company: New York, NY, 1962; 1–8.
19. DiStasio, J.I. *Epoxy Resin Technology, Developments Since 1979*; Noyes Data Corporation: Park Ridge, NY, 1982; 1–25.
20. Nagase ChemeX Corporation. *Technical Data Sheet of Denacol. Water Soluble Cross-linking agent*; Tatsuno: Hyogo, Japan, 2002; 1–13.
21. Noishiki, Y.; Takahashi, K.; Yamamoto, K.; Mo, M.; Matsumoto, A.; Yamane, Y.; Miyata, T. A new cardiac wall substitute with high affinity for fibroblasts that can induce an endothelial cell lining. *ASAIO J.* **1994**, *40*, M751–M756.
22. Tomizawa, Y.; Noishiki, Y.; Okoshi, T.; Koyanagi, H. Development of a polyepoxy compound cross-linked heterologous connective tissue tube. *ASAIO J.* **1992**, *38*, M357–M361.
23. Peterson, R.C.; Tu, R.; Noishiki, Y.; McIntyre, J. Dynamic internal compliance measurements of fresh and fixed artery. *ASAIO Trans.* **1990**, *36*, 766–769.
24. Okoshi, T.; Noishiki, Y.; Tomizawa, Y.; Morishima, M.; Terada, R.; Koyanagi, H. Long-term results of a new antithrombogenic cardiac wall substitute. *ASAIO Trans.* **1989**, *35*, 391–395.
25. Tu, R.; Quijano, R.C.; Lu, C.L.; Shen, S.; Wang, E.; Hata, C.; Lin, D. A preliminary study of the fixation mechanism of collagen reaction with a polyepoxy fixative. *Int. J. Artif. Organs* **1993**, *16*, 537–544.
26. Tu, R.; Lu, C.L.; Thyagarajan, K.; Wang, E.; Nguyen, H.; Shen, S.; Hata, C.; Quijano, R.C. Kinetic study of collagen fixation with polyepoxy fixatives. *J. Biomed. Mater. Res.* **1993**, *27*, 3–9.
27. Hannah, H., III; Reis, R.L. Current status of porcine heterograft prostheses: A 5-year appraisal. *Circulation* **1976**, *54*, 27–31.
28. Perier, P.; Mihaileanu, S.; Fabiani, J.N.; Deloche, A.; Chauvaud, S.; Jindani, A.; Carpentier, A. Long-term evaluation of the Carpentier-Edwards pericardial valve in the aortic position. *J. Card. Surg.* **1991**, *6*, 589–594.
29. Imamura, E.; Sawatani, O.; Koyanagi, H.; Noishiki, Y.; Miyata, T. Epoxy compounds as a new cross-linking agent for porcine aortic leaflets: Subcutaneous implant studies in Rats. *J. Card. Surg.* **1989**, *4*, 50–57.
30. Okoshi, T.; Noishiki, Y.; Tomizawa, Y.; Morishima, M.; Taira, T.; Kawai, T.; Itoh, H.; Miyata, T.; Koyanagi, H. A new bioprosthetic cardiac valve with reduced calcification. *ASAIO Trans.* **1990**, *36*, 411–414.
31. Murayama, Y.; Sato, S.; Oka, T. Mechanical properties and immunogenicity of vascular xenografts. *Jpn. J. Surg.* **1991**, *21*, 533–541.
32. Murayama, Y.; Satoh, S.; Oka, T.; Imanishi, J.; Noishiki, Y. Reduction of the antigenicity and immunogenicity of xenografts by a new cross-linking reagent. *ASAIO Trans.* **1988**, *34*, 546–549.
33. Imamura, E.; Noishiki, Y.; Koyanagi, H.; Miyata, T.; Furuse, M. Bioprosthetic Valve. United States Patent, 5,080,697, Jan.14,1992.
34. Shen, S.H.; Sung, H.W.; Tu, R.; Hata, C.; Lin, D.; Noishiki, Y.; Quijano, R.C. Characterization of a polyepoxy compound fixed porcine heart valve bioprosthesis. *J. Appl. Biomater.* **1994**, *5*, 159–162.
35. Tu, R.; Shen, S.H.; Lin, D.; Hata, C.; Thyagarajan, K.; Noishiki, Y.; Quijano, R.C. Fixation of bioprosthetic tissues with monofunctional and multifunctional polyepoxy compounds. *J. Biomed. Mater. Res.* **1994**, *28*, 677–684.
36. Noishiki, Y.; Sung, H.; Hata, C.; McIntyre, J.; Shen, S.H.; Randall, P.; Lin, D.; Roland, M.; Quijano, R.C. Cross-Linking Characteristics of a Porcine Heart Valve Bioprosthesis Fixed with Epoxy Compound. In *New Horizons and the Future of Heart Valve Bioprostheses*, 1st Ed.; Gabbay, S., Frater, M., Eds.; Silent Partners, Inc.: Austin, TX, USA, 1994; 13–24.
37. Matsumoto, H.; Sugiyama, S.; Shibazaki, A.; Tanaka, R.; Takashima, K.; Noishiki, Y.; Yamane, Y. Experimental study of materials for patch graft on right ventricular outflow tract under extracorporeal circulation in dogs-comparison between Denacol EX-313-treated bovine jugular vein graft and expanded polytetrafluoroethylene (EPTFE) graft. *J. Vet. Med. Sci. Sep.* **2001**, *63*, 961–965.
38. Ichikawa, Y.; Noishiki, Y.; Kosuge, T.; Yamamoto, K.; Kondo, J.; Matsumoto, A. Use of a bovine jugular vein graft with natural valve for right ventricular outflow tract reconstruction: A one-year animal study. *J. Thorac. Cardiovasc. Surg.* **1997**, *114*, 224–233.
39. Ichikawa, Y.; Noishiki, Y.; Soma, T.; Ishii, M.; Yamamoto, K.; Takahashi, K.; Mo, M.; Kosuge, T.; Kondo, J.; Matsumoto, A. A new antithrombogenic RV-PA valved conduit. *ASAIO J.* **1994**, *40*, 714–718.
40. Noishiki, Y.; Miyata, T. A method to give antithrombogenicity to biological materials. *Jap. J. Artif. Organs* **1982**, *11*, 966–968.
41. Noishiki, Y.; Yamane, Y.; Miyata, T. Artificial Blood Vessel. European Patent 0,311,305B1, Sep. 29, 1988.
42. Hata, C.; Wang, E.; Noishiki, Y.; Tomizawa, Y.; Tu, R.; Sung, H.W.; Quijano, R.C. Evaluation of a polyepoxy compound fixed biological vascular prosthesis and an expanded polytetrafluoroethylene vascular graft. *Artif. Organs Jun.* **1992**, *6* (3), 263–266.
43. Tomizawa, Y.; Noishiki, Y.; Okoshi, T.; Koyanagi, H. Development of a small-caliber vascular graft with antithrombogenicity induced by extreme hydrophilicity. *ASAIO Trans.* **1988**, *34*, 644–650.
44. Wang, E.; Thyagarajan, K.; Tu, R.; Lin, D.; Hata, C.; Shen, S.H.; Quijano, R.C. Evaluation of collagen modification and surface properties of a bovine artery

- via polyepoxy compound fixation. *Int. J. Artif. Organs* **1993**, *16*, 530–536.
45. Wang, E.Y.; Giclas, P.C.; Tu, R.H.; Hata, C.; Quijano, R.C. A comparative study of complement activation by Denaflex, Bioflow, and BioPolyMerics vascular grafts. *ASAIO J.* **1993**, *39*, 691–694.
 46. Tu, R.; Wang, E.; Hata, C.; Shen, S.H.; Lin, D.; Quijano, R.C. A compliant biological vascular prosthesis. *Int. J. Artif. Organs* **1993**, *16*, 141–145.
 47. Nojiri, C.; Noishiki, Y.; Koyanagi, H. Aorta-coronary bypass grafting with heparinized vascular grafts in dogs. A preliminary study. *J. Thorac. Cardiovasc. Surg.* **1987**, *93*, 867–877.
 48. Noishiki, Y.; Yamane, Y.; Miyata, T. Long-term evaluation of a growable graft. *ASAIO Trans.* **1989**, *35*, 267–270.
 49. Noishiki, Y.; Yamane, Y.; Furuse, M.; Miyata, T. Development of a growable vascular graft. *ASAIO Trans.* **1988**, *34*, 308–313.
 50. Stenzel, K.H.; Miyata, T.; Rubin, A.L. Collagen as a biomaterial. *Annu. Rev. Biophys. Bioeng.* **1974**, *3*, 231–253.
 51. Kuru, S. Collagen Fiber Hemostatic Material and Method of Producing the Same. European Patent, 0,463,887 B1, July 1, 1991.
 52. Tomizawa, Y.; Noishiki, Y.; Endo, M.; Hashimoto, A.; Koyanagi, H. Evaluation of blood absorption, hemostatic ability and purity of a polyepoxy compound cross-linked cotton type collagen hemostat. *Kyobu Geka* **Feb. 1996**, *49*, 126–129.
 53. Miyata, T.; Fususe, M.; Yamane, Y.; Noishiki, Y. A biodegradable antiadhesion collagen membrane with slow release heparin. *ASAIO Trans.* **1988**, *34*, 687–691.
 54. Noishiki, Y.; Miyata, T. Antiadhesive collagen membrane with heparin slow release. *J. Bioact. Compat. Polym.* **1987**, *2*, 325–333.

Poly(Glycolic Acid)

Eugene D. Boland

Department of Biomedical Engineering, Virginia Commonwealth University, Richmond, Virginia, U.S.A.

Gary E. Wnek

Department of Chemical Engineering, Case Western Reserve, Cleveland, Ohio, U.S.A.

Gary L. Bowlin

Department of Biomedical Engineering, Virginia Commonwealth University, Richmond, Virginia, U.S.A.

INTRODUCTION

Poly(glycolic acid) (PGA) is a synthetic polymer that degrades to natural metabolic waste products. The chemical structure of glycolic acid can be written as HOCH₂COOH. Note that the molecular termination is a hydroxyl group on one end and a carboxyl group on the other. Glycolic acid, also known as hydroxyacetic acid, is found naturally in young plants and green fruits. Glycolate, which is the ester or salt form of the acid, is produced from ribulose 1,5 biphosphate, a Calvin cycle initiator, in chloroplasts and transferred to peroxisomes, where it is oxidized to glyoxylate, which can be transaminated to the amino acid glycine. It is further processed in the mitochondria, passed back to the peroxisome for transamination and reduction, and then finally passed back to the chloroplast as glycerate, where adenosine triphosphate (ATP) phosphorylates it. The 3-phospho-glycerate can then reenter the Calvin cycle for sugar production. This process, which is called photorespiration, is a seemingly wasteful side reaction of photosynthesis that consumes both oxygen and ATP and produces carbon dioxide and hydrogen peroxide. In human metabolism, the specific glycolic acid oxidase is not present, but it may be slowly oxidized and transaminated to glycine in a similar manner. This slow metabolism would lead to localized build up of glycolic acid. The ramifications of which will be discussed later as pertaining to the in-vivo tissue response.

SYNTHESIS OF PGA

After one has sufficient quantity of monomer from either an extraction and purification from plants^[1] or synthesis in the laboratory, the polymer can be synthesized. Two different methods can be employed to produce the polymer. The first is polycondensation. In condensation reactions, the monomer subunits react

and release a small molecule such as water.^[2] Because glycolic acid contains two functional groups that can participate in a condensation reaction, it can self-polymerize.^[2,3] The basic reactions involve creating dimers from two monomers, tetramers from two dimers, and so on to eventually form a polymer (Scheme 1).

The reaction is characterized by the early disappearance of monomers from the solution and the slow increase of high molecular weight species. Due to the random nature of the polymer formation, there is a broad molecular weight distribution throughout the reaction, and significant concentrations of high molecular weight polymers are difficult to achieve.^[2,3] Adding to the difficulty, common commercial esterification catalysts, such as antimony trioxide, also catalyze the reverse reaction, namely depolymerization.^[3] Moreover, glycolic acid has a great tendency to cyclodimerize to glycolide (Scheme 2), which thus ends its reactivity in condensation reactions.

Due to these drawbacks, a second method of preparing PGA is preferred, namely ring-opening polymerization (ROP) of glycolide (Scheme 3).^[3]

Depending on the catalyst used, these polymerization reactions can be anionic, cationic, or involve insertion mechanisms. Of these mechanisms, insertion using metal alkoxides or carboxylates as catalysts is the most commercially viable and therefore will be the only method discussed. The tin (Scheme 2) octoate catalyst is the most commonly used due to its acceptance by the U.S. Food and Drug Administration (FDA) as a food stabilizer.^[3] The polymerization of glycolide is typically carried out in bulk at 220°C for 4 hours. This results in a 96% conversion to polymer with molecular weights ranging from 10⁴ to 10⁶ Daltons.^[3] As is the case with polyesters in general, equilibration reactions such as ester interchange and chain unzipping may take place, the effects of which would be the formation of shorter chains and low molecular weight species such as cyclic dimers, trimers,



Poly(Glycolic Acid)



Scheme 1 Condensation polymerization of glycolic acid to PGA. Note that the number of repeating units (n) in the polymer is an average number.

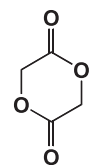
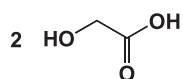
and, to a lesser extent, cyclic oligomers, and even monomers.^[2,3] These species can drastically affect the mechanical properties of the final products, and their production must be carefully monitored and avoided if possible. This ROP method can also be employed to form copolymers with other α -hydroxy acids of the general form $[-O-CH(R)-CO-]$. Other copolymerization methods include ABA block copolymer formations with a monoester group of poly(ethylene glycol) or fatty acids.^[4] These various copolymers of glycolide also have significant uses in medicine.

PGA CHARACTERIZATION

The regularity of the PGA backbone and its rather compact repeat unit suggests that the polymer will tend to at least partially crystallize. Indeed, early X-ray diffraction work reveals that the unit cell contains two antiparallel polymer chains and these chains have a sheetlike molecular arrangement of planar zig-zag conformations. Also, the C=O groups of adjacent polymers in the unit cell overlap (Fig. 1). The density of the crystal was determined to be rather high at 1.69 grams per cm^3 . It was suggested that the close approach of ester groups, which results in tight molecular packing, produces the abnormally high melting temperature and insolubility of the polymer.^[5]

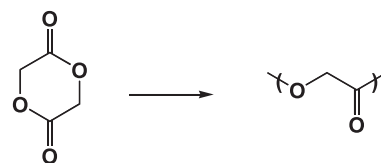
PGA DEGRADATION

Biodegradation is the destruction of one or more chemicals by biological processes. PGA is not removed



glycolide

Scheme 2 Dimerization of glycolic acid to glycolide.



Scheme 3 Ring-opening polymerization of glycolide to PGA.

by biodegradation alone but also by mechanisms of hydrolytic degradation, erosion, and bioabsorption. Hydrolytic degradation is the breakdown of a material by physical or chemical interactions with water, and bioabsorption is the uptake of material by a cell. Hydrolytic degradation is the predominant mechanism of PGA breakdown resulting in the loss of mechanical strength.^[6]

There have been many theories about the hydrolytic degradation of PGA. All the theories have had two central dogmas. First, the amorphous region of the polymer is quickly degraded, which results in a significant loss of mechanical properties in a matter of days to weeks. Second, the crystalline region of the remaining fragments degrade more slowly yet crystallinity remains measurable until the material is completely degraded. One such theory was formulated from a detailed multitime point analysis described by Hurrell and Cameron,^[6] a group that employed analysis techniques that included small and wide angle X-ray scattering (SAXS and WAXS, respectively), mass measurement, differential scanning calorimetry, pH measurement, and UV-spectrophotometry. One key finding was that, as the degradation proceeded, the intensity of the crystal peaks on the WAXS profiles increased with respect to the amorphous halo.^[6] This confirms earlier findings that the amorphous region degrades first and leaves the crystalline regions intact.^[3] This observation continued over the first 10 days, at which time, the crystallinity leveled off, and the amorphous halo was only faintly visible. Again, this observation is supported by clinical findings that PGA sutures lose most of their strength in the first two

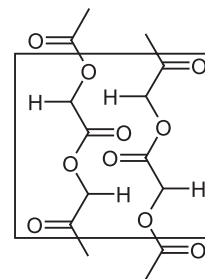


Fig. 1 Crystal structure of polyglycolide.

weeks.^[3] Another valuable observation is the rapid drop in the glass transition temperature over the first 6 hours and then it continues to drop at a much slower rate. This rapid drop may indicate that water had diffused to the center of the sample in the first 6 hours. This corresponds well with the previously reported diffusion constant of $5.0 \pm 1.0 \times 10^{-9} \text{ cm}^2 \text{ s}^{-1}$ of water into PGA.^[6] A final observation showed a dramatic drop in pH after about 10 days despite being in a buffered solution.^[6] This indicated that glycolic acid is released during the first 10 days, but the rate of release may increase dramatically around the 10-day point. In all the techniques applied to analyzing PGA degradation, a dramatic change in rate or morphology was noted at the 10-day point.

These observations and calculations helped devise the following four-stage degradation profile:

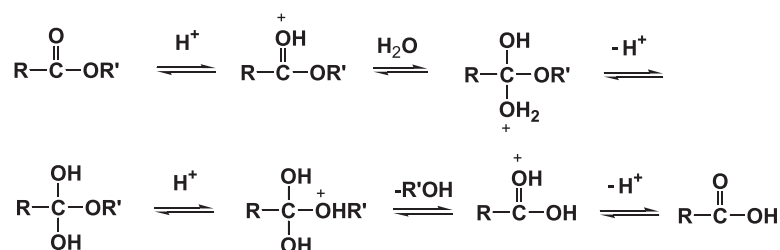
- The first stage of the degradation process is characterized by homogeneous diffusion of the hydrolytic solution into the sample. A sharp drop in the glass transition temperature characterizes this stage. For their model, this stage occurred over the first 6 hours.
- The second stage is that of homogeneous hydrolysis. During this stage, there is solution absorbed and relatively low mass loss. The chain hydrolysis reduces the polymer fractions to a critical molecular weight that is capable of diffusing from the solid. This stage took place over the first 10 days.
- The end of the second stage and beginning of the third stage is marked by a significant drop in the localized pH. This pH drop is a function of oligomer diffusion from the construct. Cameron theorizes that, as oligomers diffuse out and additional fluid diffuses in, a reaction-erosion front forms on the surface.^[6] This would expedite the degradation of the sample and address crystal breakdown. Due to the observation that cracks originate on the surface and proceed inward during hydrolysis, it is believed that these fronts proceed inward via hydrodynamic forces and eventually reach the opposing front at the center.

- The time at which these fronts meet is the onset of stage four. During stage four, only small crystalline fragments remain within the implant site. There is no structural integrity left in the implant, and the stage concludes at the removal of all traces of the PGA.^[6]

With hydrolytic degradation of polyesters, changing the pH of the hydrolytic media can significantly alter the rate of degradation. This would be an interesting aside if one were to assume all bodily tissues and fluids remained at a pH near 7.4. While this may be true for the majority of bulk tissues and fluids, it is certainly not the rule. Localized pH values of 1.0 are common in the stomach, but the pH in the small intestine and bladder often exceeds 10.0. In addition to regional pH differences, one must be aware of the local effects of inflammatory and foreign body responses. Depending on the cells recruited, local effectors can either raise or lower the pH, depending on the proteins involved. To compare pH effects, an experiment was developed to degrade PGA suture material in three different buffered solutions. The three different buffer solutions were maintained as acidic (pH = 5.25), slightly alkaline (pH = 7.44), and strongly alkaline (pH = 10.09). During the degradation process, samples were tested for strength retention, i.e., standard uniaxial fiber testing in a controlled environmental chamber.^[7] Other experimentation added scanning electron microscopy (SEM) to evaluate surface morphology during degradation.^[7,8] Initial results found that the suture material degraded significantly faster in the strong alkaline solution than in either of the less alkaline solutions. In fact, there was no significant difference in the degradation behavior of the acidic or slightly alkaline treated samples prior to 21 days of immersion. Two possible factors could be responsible for differences seen in this study, the so-called cage-effect in the crystalline regions and the pH effect on the ability to form hydrogen bonds.^[7]

The hydrolysis of simple organic esters in acidic media follows the general mechanism in Scheme 4.

It is believed that the reversal process of two fragmentary ends rejoining in the amorphous region is



Scheme 4 Acid-catalyzed ester hydrolysis.

insignificant due to chain mobility, but this cannot be said about the crystalline regions. Two chains held in a crystal lattice that are separated by hydrolytic cleavage have a greater chance to recombine. This is the so-called cage effect. If this does indeed exist, it could slow down the degradation in acidic or near-neutral pH buffer conditions. The cage effect may not even be restricted to the crystalline regions. If sufficiently immobile polymers, either due to high viscosity or high molecular weights, a cage could be created in amorphous regions as well.^[7,8]

The same cannot be said in the presence of a strong alkaline-buffered solution due to the irreversibility of hydrolysis.^[7-9] Alternatively, it has been reported that the capability of adjacent polymer chains to form hydrogen bonds is significantly reduced or even lost at higher pH levels. Since hydrogen bonding was previously identified as a contributing factor to PGA's crystallinity and stability,^[5] it would stand to reason that the elimination of those bonds could accelerate degradation. A relatively looser structure would be more open to attack.^[9]

It is likely that both of these mechanisms contribute to the hydrolytic degradation of PGA. The cage effect would be more pronounced in acidic or neutral attacks, and the reduction of pH-dependent hydrogen bonding would be more pronounced in strong alkaline attacks. The former mechanism preserves the crystalline regions while the latter rapidly degrades the amorphous regions. These mechanisms are closely tied to the experimental observations if we accept a chain scission model. This model states that chain scission during hydrolytic degradation begins at tie-chain segments at the point where the rigid crystals tie into the amorphous chains.^[7,8,17] These ends carry the load in the bulk polymer and are most susceptible to move during hydrolytic cleavage in acidic or neutral attacks. The tie-chains are also theorized to participate in extensive hydrogen bonding with the crystal lattices and are therefore prone to dissociate with strong alkaline attacks. Both outcomes can predict the loss of mechanical strength and elasticity observed experimentally.^[7,9]

Localized pH is not the only effectors that can change degradation rates in vivo. The role of enzymes has also been identified as contributing to the degradation of PGA. As stated earlier, humans do not possess the specific glycolic acid oxidase that can expedite breakdown and removal of PGA, but we do possess a number of nonspecific esterases and proteases that may be expressed by cells recruited in a foreign body response to implanted polymers.^[9] Experiments have demonstrated that certain families of enzymes increased the rate of degradation of PGA in vitro.^[10] Specifically, the presence of bromelain, esterase, or leucine aminopeptidase reduced the complete degradation time of

PGA sutures to less than one week from the previously reported values ranging from two to four weeks.^[3,4,10] This in-vitro work prompted an in-vivo exploration of implant sites. Large quantities of acid phosphatase and leucine aminopeptidase were found at the implant sites of orthopedic PGA fixatives. While the biodegradation properties of leucine aminopeptidase were seen in vitro, acid phosphatase did not enhance in-vitro degradation.^[10] The role of the acid phosphatase is unclear, but it has been speculated to play a role in absorption.^[10]

PGA APPLICATIONS

As previously mentioned, PGA is a semicrystalline material with a high melting temperature and low solubility. In fact, the only known means of solvent processing PGA is in hexafluoroisopropanol or hexafluoroacetone sesquihydrate in carbon tetrachloride.^[6,12] However, both of these solvents, have high toxicity levels appropriate for yarn or fabric production but not for drug delivery since the drug may react with the solvent or trace amounts of the solvent could be incorporated into the end product. Melt processing is the typical commercial method, specifically hot-drawn fibers or melt-pressed molding. With the former technique, fiber diameters as low as 10 μm can be produced, and the latter is capable of larger-sized constructs typically for orthopedic applications.^[3,11] These fibers can then be further processed through winding, weaving, and braiding techniques to produce currently marketed products such as suture and wound dressings.^[3,11,13] Several products have been commercially available since the early 1970s and have paved the way for new advances in PGA processing and products.

Current medical research involving PGA is in four key areas: drug delivery, orthopedic fixation, degradable sutures, and tissue engineering. As previously stated, PGA has many drawbacks as a drug delivery device except for its rapid and predictable degradation. To avoid toxicity of solvents, scientists sought a drug that was thermally stable at temperature compatible with melt-processing PGA.^[12] For this study, their drug of choice was theophylline, a bronchodilator that is used in the treatment of chronic asthma and pulmonary disease. The drug is water-soluble, weakly basic, and sublimates without degradation at temperature above PGA's processing temperature.^[12] Even at low concentrations, its inclusion prevented crystal formation in PGA, which may be evidence of even distribution throughout the polymer, and should give a more even dosage. Both of these theories proved correct and illustrate the strength of PGA as a drug-delivery polymer. A rather steady dispensing of the drug from day 6 to day 15 was observed,^[12] which

provided additional support to their previous theory on the four-stage degradation process of PGA.

The second area of interest is in orthopedic fixation devices, namely rods, plates, and screws. Two different uses are emerging. The first is as a binding agent with other polymers such as calcium phosphate fibers,^[14] and the second is as a self-reinforced polymer.^[15] In the former application, PGA is used to provide initial support that is gradually reduced until the calcium phosphate is carrying the full load. Nevertheless, a material with a two-week degradation time may not be adequate for bone repair. To slow the rate of hydrolytic degradation, researchers chose a novel surface plasma treatment, CH₄ and Parlene[®], to make the surface more hydrophobic.^[14] In doing so, they slowed the rate of media uptake and therefore slowed the first stage of degradation, which appears to be pivotal in the overall degradation rate. These plasma coatings were found to be nontoxic and thus provide a new method to decrease the degradation rate of PGA to counter the many factors that can increase the rate of degradation. Along a similar methodology, a method of coating self-reinforced PGA (sintered suture) with another slower degrading polymer is presented.^[15] The constructs were briefly submerged in chloroform containing polydioxanone, poly(hydroxybutyric acid), or poly(lactic acid). This was done to form a thin, slowly degrading coating that would prevent water uptake by the PGA construct. The idea may have been sound, but the application suffered cracking and delamination and was eventually ruled out since it did not statistically decrease the degradation rate of the constructs *in vivo*. It was concluded by this group that self-reinforced PGA materials degrade too rapidly to be effective internal fixatives for cortical bone fractures.^[15]

The third area of interest returns to sutures. As long as surgeons are cutting into the body, there will be a quest for the perfect suture. One that can be used anywhere, has strong knot retention, retains strength long enough for complete closure, and avoids adverse effects on healing. This has led researchers to reevaluate modifications to PGA that affect its strength and degradation rate. One such approach has been to anneal the suture material for different lengths of times and at different temperatures while subjecting the suture to a slight axial load.^[16] The idea is that the crystalline regions will align along the long axis of the suture, and polymer chains in the amorphous regions will straighten to increase strength and decrease the degradation rate. This was the response elicited from the testing. A very complex relationship emerged that linked annealing temperature and load with hydrolytic degradation and mechanical properties.^[16] Further evaluation will be necessary to test the limits of these relationships.

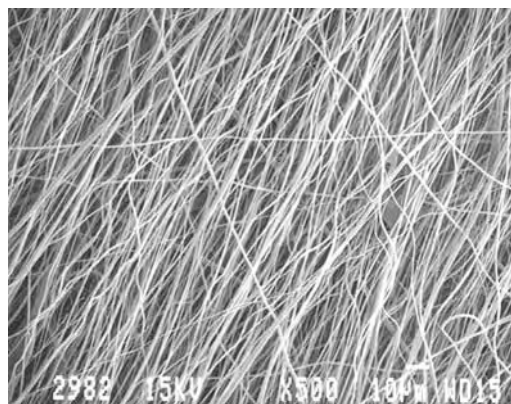


Fig. 2 Scanning electron micrograph of electrospun PGA fibers (scale bar = 10 microns).

The emerging discipline of tissue engineering is reaching into many aspects of anatomy and physiology—from wound closure to full organ replacement. Electrospinning is the deliberate charging of a polymer, either in solution or melt, by applying a high voltage potential between that solution and a grounded target. The phenomenon of combining electrospinning with tissue fabrication, which occurs when electrical forces at the surface of the polymer solution overcome the surface tension, creates a polymer solution jet. This jet produces fibers with diameters from microns to nanometers, specifically down to 50 nm.^[17] Many factors can affect the morphology of electrospun fibers although these variables have only recently been examined with PGA.

As described earlier, PGA can be tailored into many different scaffold configurations utilizing electrospinning. Electrospinning uses a static electric field to draw the polymer fibers out of solution rather than a mechanical drawing or ejection system previously described for fiber formation. PGA fibers can be produced from $1.2 \pm 0.4 \mu\text{m}$ diameters at a 140-mg/ml concentration in hexafluoroisopropanol to $110 \pm 40 \text{ nm}$ diameters at a 50-mg/ml concentration. Fiber diameter varied linearly with concentration. Fiber alignment, as depicted in the micrograph in Fig. 2, was achieved through rotation and translation of the grounded target. Fibrous mats produced in this manner demonstrated significant mechanical strength and anisotropic properties related to fiber orientation. It is speculated that these relationships can be exploited to develop a wide variety of biomimicking tissue-engineering scaffolds.^[18]

PGA REGULATORY ISSUES

Regulatory pathways have not always been as critical. The United States federal government did not tightly regulate initial PGA products, which were approved

for sale prior to the medical device amendment to the Food, Drug and Cosmetic Act in 1976; however, new devices or drugs will be much more thoroughly scrutinized for compliance to specific parts of the Good Manufacturing Practices as outlined in CFR title 21.^{[19]a} Even with this scrutiny, similar products can still be expedited through the regulatory process. And, as often the case, the first product approved becomes the standard to which all others are measured. Outside the United States, the European Union has established the European Medicines Evaluation Agency (EMA) to regulate the marketing of medical products as either a medical device or medicinal product. Similarly, the Japanese Pharmaceutical and Medical Safety Bureau (PMSB) requires demonstration of safety and efficacy of medical products regulated as either medical devices or pharmaceuticals. Regulation of PGA products are increasingly hard to characterize as devices, drugs, or biologics, and approval for use must be sought on a case-by-case, country-by-country basis. Further information about issues facing tissue-engineered products can be found in a World Technology Evaluation Center (WTEC) report on tissue engineering research.^[20]

One area that is always a concern of the regulatory bodies is the sterilization of products. Gamma irradiation is becoming a popular method to sterilize everything from latex gloves to lunch meats. Care should be taken before applying this technology to PGA constructs. It has been reported that gamma irradiation can significantly affect the degradation properties of PGA. Utilizing the techniques of electron spin resonance (ESR) analysis and scanning electron microscopy, chain scission, and degradation were monitored *in vitro*. Gamma irradiation of ester groups can enhance the susceptibility to form chain scissions even at a normal sterilization dose of 25 kilograys (kGy).^[21] At this dosage, total mechanical strength is lost during the first 10 days in buffered media. Significant decrease in the molecular weight after irradiation was measured, and ESR confirmed many chain scission radicals in the polymer chains. Although initial SEM analysis showed no difference before placing the samples in media, it became quickly apparent that irradiated samples degraded at a much faster rate and appeared to degrade by the same mechanism. This provides confirmation that merely inspecting for cracks cannot accurately predict damage caused by gamma irradiation. It would be advisable to select an alternate means of sterilization.^[21]

^aCurrent Good Manufacturing Practices for Medical Device Manufacturers, Code of Federal Regulations, Title 21 Part 820.

PGA IN-VIVO TISSUE RESPONSE

In addition to regulatory issues, such as sterilization, the greatest hurdle is often tissue compatibility. Work has been done to describe the immune response to PGA implants. In general, they are very acceptable. The principle cell type found in effusion samples around PGA implants was a small lymphocyte although a few monocytes were also found.^[22] This could be interpreted as a slight foreign body response to the implant, but the absence of neutrophils supports a noninfectious response, possibly a lymphocyte-mediated immunological reaction.^[22] These data can be taken to show PGA is relatively inert immunologically; however, inert may not be good enough for the next generation of tissue engineering materials. It would be desirable to create materials that are bioactive, *i.e.*, they cause a biological response. The desired response would be that of natural healing and incorporation—not encapsulation.

Many groups are racing to find the magic material that promotes natural healing, augments native tissues (if desired), or is capable of replacing native tissue with a new tissue construct. One such approach was to take highly porous PGA matrices of nonwoven 12-micron diameter fibers and implant them subcutaneously in a rat model to see if native tissue would grow into the matrix and form a new tissue mass. Initially, the matrices were encapsulated, but, with time, the capsule thinned and cells migrated into the matrix. Histological examination showed the cells migrated from the skeletal muscle rather than from the skin or adipose tissue.^[23] Much to researchers' dismay, the tissue receded as the matrix degraded. Upon complete degradation of the matrix, no trace of the cells that had previously ingrown could be found, and no internal scar was evident even though a fibrous capsule previously existed in that location. When comparing PGA to other matrix material, less neovascularization was noticed. It was postulated that the rapid release of glycolic acid into the local environment dropped the localized pH and down-regulated vascular formation.^[23]

Another approach taken with a PGA based product was pre-hydrolyzing the surface to promote cell adhesion.^[24] The premise of this research is that short-term exposure to an alkaline solution will begin hydrolyzing the surface of the matrix, improve its wetting behavior, and ultimately make it a more suitable cell culture surface. SEM analysis revealed surface erosion without crack formation from this processing. A 50% mass reduction in the fibers of the matrix was targeted (13- μ m diameter fibers reduced to 9 μ m). Wetting behavior was also effected by this treatment by changing the surface composition from 51% carbon and 49% oxygen to 49% carbon and 51% oxygen as measured by XPS.^[24] This increase in oxygen on the

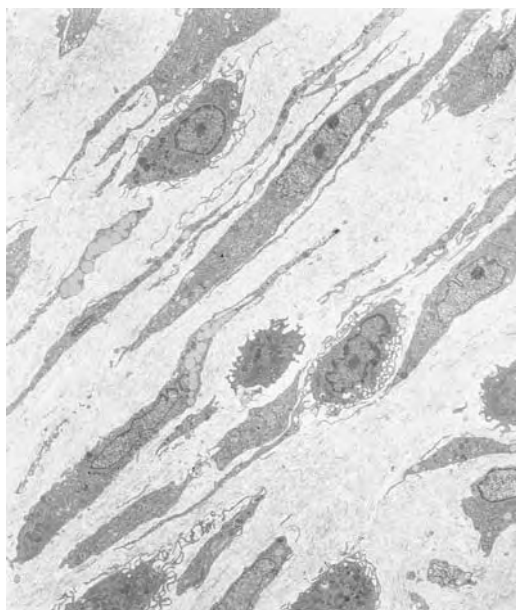


Fig. 3 Transmission electron micrograph of a PGA scaffold (1- μm untreated fibers) implanted for 2 weeks in a rat skeletal muscle model.

surface could account for the increased hydrophilic nature of the surface. Through this experimental technique, they were able to devise a way to improve static, in-vitro cell seeding densities.^[24] They did not expand to animal trials or do long-term seeding studies to see if the trend continued.

The final method to be presented is work being done to enhance the usefulness of PGA as a tissue-engineering polymer. The focus of this work can be surmised as encapsulation versus incorporation. Many researchers have shown that PGA is encapsulated by collagen during the degradation process. This may be acceptable for sutures but not a tissue scaffold. Building on work with electrospun PGA presented earlier, one would want to look at cellular response to the material/matrix construct. Scientists have described a size scale for tissue engineering claiming that cells cannot migrate through pores under 20 to 50 microns. They maintain that small pores promote encapsulation, and large pores promote incorporation.^[25] Nevertheless, unpublished data seems to contradict this dogma. Figures 3 and 4 are two transmission electron micrographs that show cells within an electrospun PGA matrix with pore sizes below 5 microns and below 0.8 μm , respectively.

Observationally, the micrograph in Fig. 3 shows a typical capsule formation by actively secreting fibroblasts while the micrograph in Fig. 4 shows no capsule and high cell densities laying down orderly matrices. The difference between the two is that the first is 1- μm fibers and the second is 0.2- μm fibers that were

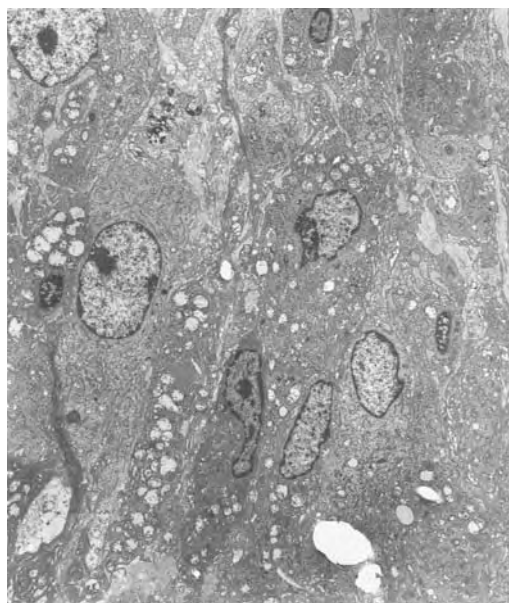


Fig. 4 Transmission electron micrograph of a PGA scaffold (0.3- μm fibers that were acid pretreated) implanted for 2 weeks in a rat skeletal muscle model.

pretreated in 11.7-molar hydrochloric acid. While other research presented claimed no difference in degradation with acidic versus neutral pH buffers, the material was not exposed to as strong an acid. These preliminary findings, if supported by additional analytical and animal experimentations, may bring PGA back into the forefront as a bioabsorbable polymer for medicine, specifically tissue engineering.

ARTICLES OF FURTHER INTEREST

Biodegradable Polymers, An Overview; Biofunctional Polymers; Bone Plates and Screws, Bioabsorbable; Elastomers, Biodegradable; Poly(lactic acid)s; Polymers; Sutures; Tissue Engineering Scaffolds

REFERENCES

1. Nelson, D.; Cox, M. Carbohydrate Biosynthesis. In *Lehninger: Principles of Biochemistry*, 3rd Ed.; Nelson, D., Cox, M., Eds.; Worth Publishers: New York, 2000; 722–764.
2. Alcock, H.; Lampe, F. *Contemporary Polymer Chemistry*; Prentice-Hall: New Jersey, 1981.
3. Wong, W.H.; Mooney, D.J. Synthesis and Properties of Biodegradable Polymers Used as Synthetic Matrices for Tissue Engineering. In *Synthetic Biodegradable Polymer Scaffolds*; Atala, A., Mooney, D., Vacanti, J., Langer, R., Eds.; Birhauser: Boston, 1997; 50–82.

4. Belenkaya, B.; Sakharova, V.; Sinevich, E.; Belousov, S.; Kuptsov, A. Synthesis and biodegradation of modified polyglycolide. *Polymer Prepr.* **1999**, *39* (2), 164.
5. Chatani, Y.; Suehiro, K.; Okita, Y.; Tadokoro, H.; Chujo, K. Structural studies of polyesters I. Crystal structure of polyglycolide. *Makromol. Chem.* **1968**, *113*, 215–229.
6. Hurrell, S.; Cameron, R. Polyglycolide: Degradation and drug release. Part I: Changes in morphology during degradation. *J. Mater. Sci. Mater. Med.* **2001**, *12*, 811–816.
7. Chu, C. The in vitro degradation of poly(glycolic acid) sutures—Effect of pH. *J. Biomed. Mater. Res.* **1981**, *15* (6), 795–804.
8. Chu, C.; Campbell, N. Scanning electron microscopic study of the hydrolytic degradation of poly(glycolic acid) suture. *J. Biomed. Mater. Res.* **1982**, *16* (4), 417–430.
9. Tomihata, K.; Suzuki, M.; Ikada, Y. The pH dependence of monofilament sutures on hydrolytic degradation. *J. Biomed. Mater. Res.* **2001**, *58* (5), 511–518.
10. Williams, D.; Mort, E. Enzyme-accelerated hydrolysis of poly(glycolic acid). *J. Bioeng.* **1977**, *1* (3), 231–238.
11. Barrows, T. Synthetic Bioabsorbable Polymers. In *High Performance Biomaterials: A Comprehensive Guide to Medical and Pharmaceutical Applications*; Szycher, M., Ed.; Technomic: Lancaster, PA, 1991; 243–257.
12. Hurrell, S.; Cameron, R. Polyglycolide: Degradation and drug release. Part II: Drug release. *J. Mater. Sci. Mater. Med.* **2001**, *12*, 817–820.
13. Eilbert, J.; McKinney, P.; Conn, J.; Binder, P.; Beal, J. Poly(glycolic acid) synthetic absorbable sutures. *Am. J. Surg.* **1971**, *121* (5), 561–565.
14. Ibnabddjalil, M.; Loh, M.; Chu, C.; Blumenthal, N. Effect of surface plasma treatment on the chemical, physical, morphological, and mechanical properties of totally absorbable bone internal fixation devices. *J. Biomed. Mater. Res.* **1994**, *28* (3), 289–301.
15. Vasenius, J.; Vainionpää, S.; Vihtonen, K.; Mäkelä, A.; Rokkanen, P. Comparison of in vitro hydrolysis, subcutaneous and intramedullary implantation to evaluate the strength retention of absorbable osteosynthesis implants. *Biomaterials* **1990**, *11* (7), 501–504.
16. Browning, A.; Chu, C.C. The effects of annealing treatments on the tensile properties and hydrolytic degradative properties of polyglycolic acid sutures. *J. Biomed. Mater. Res.* **1986**, *20* (5), 613–632.
17. Deitzel, J.; Kleinmeyer, M.; Harris, J.; Beck, D.; Tan, N. The effect of processing variables on the morphology of electrospun nanofibers and textiles. *Polymer* **2001**, *42*, 261–272.
18. Boland, E.D.; Wnek, G.E.; Simpson, D.G.; Pawlowski, K.J.; Bowlin, G.L. Tailoring tissue engineering scaffolds using electrostatic processing techniques: A study of poly(glycolic acid) electrospinning. *J. Macromol. Sci.* **2001**, *38* (12), 1231–1243.
19. Kitchens, W.H. Legal Issues Involved in Tissue Engineering: FDA Regulation of Tissue Engineering. In *Synthetic Biodegradable Polymer Scaffolds*; Atala, A., Mooney, D., Vacanti, J., Langer, R., Eds.; Birhauser: Boston, 1997; 33–49.
20. http://wtcc.org/loyola/te/final/te_final.pdf. (accessed Oct. 2002).
21. Babanalbandi, D.; Hill, D.; O'Donnell, J.; Pomery, P. An electron spin resonance analysis on γ -irradiated poly (glycolic acid) and its copolymers with lactic acid. *Polym. Degrad. Stab.* **1996**, *52*, 59–66.
22. Santavirta, S.; Kontinen, Y.; Saito, T.; Grönblad, M.; Partio, E.; Kemppinen, P.; Rokkanen, P. Immune response to poly(glycolic acid) implants. *J. Bone Joint. Surg. Br.* **1990**, *72* (4), 597–600.
23. Holder, W.; Gruber, H.; Moore, A.; Culberson, C.; Anderson, W.; Burg, K.; Mooney, D. Cellular ingrowth and thickness changes in poly-L-lactide and polyglycolide matrices implanted subcutaneously in the rat. *J. Biomed. Mater. Res.* **1998**, *41*, 412–421.
24. Goa, J.; Niklason, L.; Langer, R. Surface hydrolysis of poly(glycolic acid) meshes increases the seeding density of vascular smooth muscle cells. *J. Biomed. Mater. Res.* **1998**, *42* (3), 417–424.
25. Mooney, D.J.; Langer, R.S. Engineering Biomaterials for Tissue Engineering: The 10–100 Micron Size Scale. In *The Biomedical Engineering Handbook*; Bronzino, J.D., Ed.; CRC Press: Boca Raton, 1995; 1609–1618.

Poly(lactic acid)s

Michel Vert

CRBA-UMR CNRS 5473, Faculty of Pharmacy, Université Montpellier 1, Montpellier
Cedex 5, France

INTRODUCTION

Poly(α -hydroxy acid)s are aliphatic polyesters that derive from α -hydroxy acids. The main members of the family are poly(glycolic acid) (PGA) and poly(lactic acid) (PLA), also named polyglycolide and polylactide. In the biomedical field, poly(glycolic acid) and lactic acid-rich copolymers were first developed as bioresorbable sutures, thanks to the sensitivity of the corresponding polymer chains to water and in-vivo hydrolysis.^[1] Soon after, it was shown that polymers are more resistant to water than PGA and degrade more slowly in aqueous media.^[2,3] Furthermore, it was rapidly found that not only the degradation rate, but also most of the physical, thermal, and mechanical characteristics of poly(α -hydroxy acids) can be modified and somewhat adjusted via molecular chemical and configurational structures, a feature of great interest in the biomedical and the pharmaceutical fields.^[5,6] Indeed, lactic acid (2-hydroxy propanoic) is the simplest chiral hydroxy acid. It exists in two optically active optical isomers or enantiomers, namely L- (S- according to absolute configuration) and D- (R)-lactic acids. The stereocopolymerization of these two enantiomers can generate polymeric materials with similar chemical compositions, but very different physical properties, depending on the enantiomeric composition and the chiral lactyl unit distribution within the chains. Last but not least, lactic acid, like glycolic acid, is normally present in biochemical pathways of many cells and living organisms. Therefore, poly(lactic acid) (PLA) and lactic acid- and glycolic acid-containing copolymer chains (PLAGA) degrade to natural metabolic compounds. Lactic acid-based polymers were first studied from a scientific viewpoint^[2] until they became of industrial interest for bone surgery in the 1990s^[6] after their known properties had been considerably improved.^[3,7]

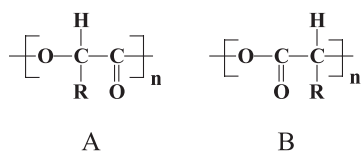
This contribution deals primarily with poly(α -hydroxy acid)s based on lactic acid enantiomers. However, many comments and characteristics will be of value for lactic acid-based copolymers involving glycolyl and/or other counits.

SYNTHESIS OF PLA POLYMERS

The prime precursor of lactic acid-based polymers is lactic acid, a compound of interest in the food industry. Lactic acid can be manufactured either by carbohydrate fermentation (L-enantiomer) or by chemical synthesis (racemic mixture of L- and D-enantiomers), although fermentation dominates.^[8] Lactic acid-based polymers were first synthesized at the time other polycondensates like Nylon[®]-type polyamide were investigated.^[9] The resulting poor polymers being rather sensitive to humidity and heat, they were considered useless. Fortunately, advances in polymer chemistry opened another route, namely the ring-opening polymerization of the cyclic dimer of lactic acid.^[10] PLA polymers and stereocopolymers correspond to the following general formula where R = methyl ($-\text{CH}_3$) (Scheme 1).

Formula A is commonly used mostly because it is more symmetrical and emphasizes the ester links. Formula B obeys the IUPAC rules. When it was realized in 1981 that the properties of a given member of the poly(α -hydroxy acid) family were batch-dependent, the use of the acronym PLAX was recommended to name PLA stereocopolymers and reflect the gross composition in repeating units, at least. In PLAX, X is the percentage of L-lactyl units present in the PLA chains.^[11] Since then, more than 20 factors have been identified that can affect the behavior of PLA polymers.^[12] Of course, it is not possible to reflect the contribution of all these factors within a simple acronym, but we strongly believe that everyone ought to use the PLAX acronym to make the comparison of literature data more consistent. The chirality of the lactic acid molecule makes PLA polymers rather special in the field of synthetic polymers. Being aware of the particularities is critical for applications in human therapy. Indeed LA-containing stereocopolymer and copolymer molecules are outstanding because of the presence of chiral centers along the chain. Indeed, chirality generates an extra source of differentiation between polymers having similar gross compositions in L-LA and D-LA units, namely the distribution of chiral centers. This distribution depends on many more or less interrelated factors including the gross composition, the mechanism





Scheme 1 Usual (A) and nomenclature official (B) general formula of PLA polymers.

chain formation, the nature of monomer, and the occurrence of configurational rearrangements caused by intrachain and interchain transesterification reactions.^[13,14]

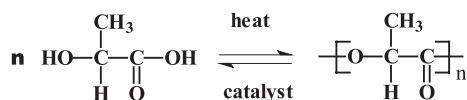
FORMATION OF LACTIC ACID-BASED POLYMERS

There are two main routes to make lactic acid-based polymers, namely the step-growth polymerization or copolycondensation of lactic acid enantiomers together and with other hydroxyacids, and the chain-growth polymerization or ring-opening polymerization of lactide cyclic dimers together or with other heterocyclic monomers.

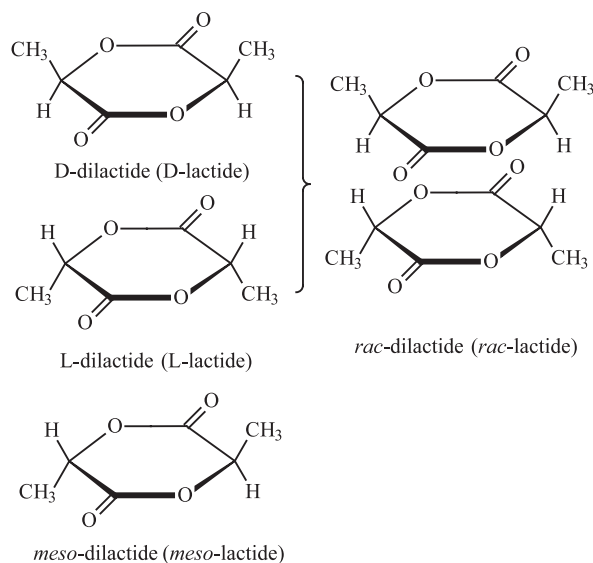
Step-Growth Polymerization of Lactic Acids

The step-growth polymerization of lactic acid corresponds to the following equilibrium (Scheme 2).

The resulting PLA homopolymers have low molecular weight (LMW) although high molecular weight polycondensates have been reported as feasible in literature, using particular catalytic and polycondensation conditions.^[15,16] When L- or the D-lactic acid is used, homo-oligomers are obtained that are semicrystalline, and oligo-L and oligo-D lactic acids have similar physical characteristics, provided molecular weights are comparable. When the two enantiomers are mixed in equal proportions (racemic lactic acid), oligomeric chains composed of L- and D-repeating units are obtained where these units are distributed at random. Poly(DL-lactic acid) polycondensates are thus amorphous. If L- and D-lactic acids are mixed in different proportions, stereocopolycondensates are obtained that can be amorphous or crystalline, depending on the proportion of the two enantiomeric repeating units. Under normal conditions, PLAX



Scheme 2 Poly(lactic acid)s obtained by polycondensation (step-growth polymerization) of lactic acids.



Scheme 3 The various lactide-type diastereoisomers.

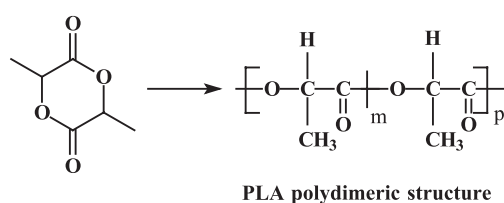
polycondensate chains are ended by OH and COOH groups in equivalent amounts.

Chain-Growth Polymerization of Lactides, the Cyclic Dimers Derived from Lactic Acids

To obtain high molecular weight chains, the most efficient route is the ring-opening polymerization (ROP) of lactides, the cyclic dimers of lactic acids. Because of their dimeric structure and because of the presence of two chiral centers, there are different lactide diastereoisomers as shown in Scheme 3.

The official IUPAC name of these cyclic dimers is 3,6-dimethyl-1,4-dioxane-2,5-diones, or dilactides. In literature, people generally use lactide for the sake of simplicity.

- D-lactide, L-lactide, and meso-lactide are composed of two D-lactyl, two L-lactyl, and one D- and one L-lactyl unit, respectively, whereas rac-lactide is a 50/50 mixture of L- and D-lactides.
- L-lactide and D-lactide are generated during the thermal degradation of lactic acid oligomers obtained by polycondensation of L-lactic acid and D-lactic acid, respectively. They melt at 97–98°C and are soluble in many organic solvents (acetone, ethyl acetate, and benzene, for instance).
- Rac-lactide (Mp = 126°C) is obtained by the same method starting from rac-lactic acid. This lactide can be readily purified by recrystallization from hot saturated solutions in organic solvents.
- Meso-lactide (Mp = 46°C) is usually extracted by crystallization from the hot solution remaining after recrystallization of rac-lactide and evaporation of



Scheme 4 Polydimeric poly(lactic acid)s obtained by ring opening polymerization.

most of the solvent. It is rather difficult to obtain pure meso-lactide.^[14]

The structure of the cyclic dimers composed of lactic acid enantiomers includes two ester bonds that are chemically equivalent. As soon as one of the ester bonds is broken due to the action of a reagent, the second ester bond is stabilized and can no longer be involved in the polymerization process, although it can be affected by transesterification reactions^[13,14,18] (Scheme 4).

The polymer chains that result from the ROP of lactides are actually polydimer chains according to chain growing by the addition of pairs of repeating units (Scheme 4).^[3,14] This particularity is the source of rather complex structures that are typical of lactic acid polymers obtained by ring-opening polymerization of lactides. The various structures and corresponding structure–properties relationships have been extensively described and discussed in reviews and encyclopedia.^[4,19–22]

The polymerization of lactides can be initiated by a great number of compounds that were identified very early.^[23] However, resulting polymers are not equivalent. Lactide ROP generally leaves an initiator or a catalyst residue at one terminal unit.^[24–26] This particularity is not observed when zinc metal or zinc lactate are used as initiators in bulk polymerization,^[27] all PLAX chain being terminated by OH and COOH end groups. Carboxylic chain ends being involved in the hydrolytic degradation of PLA chains, their presence is one of the major factors that can dramatically affect the degradation of PLA polymers, as discussed later in this article.

The initial list of compounds that are able to polymerize lactides was complemented by more exotic initiators such as alkaline earth derivatives^[28] and other complex catalysts,^[29] polymerizations being conducted either in solution or in bulk. Almost all the possible mechanisms, i.e., anionic, cationic, and coordination–insertion have been proposed, but not always conclusively demonstrated. Despite the number of papers appearing in literature, the mechanisms involved in many polymerization reactions are still a matter of debate, as is the case for stannous or tin octoate (tin(II)

2-ethyl hexanoate). Stannous octoate is largely used in academia and industry. However, fundamental investigations are usually conducted in solution whereas industry prefers bulk polymerizations. Zinc metal was selected in France many years ago.^[30] Nowadays, zinc lactate, which was shown to be the active species, is preferred. Zinc metal and zinc lactate provide polymers with rather different physico-chemical and degradation behaviors as compared with stannous octoate.^[31,32]

Both tin and zinc initiating systems are now approved by the FDA, the former having been the first approved for biomedical applications after being approved as a food additive, the second being a known oligo element. Industrially, the polymerization of lactides is generally carried out at temperatures ranging between 100 and 180°C. It is worth noting that at high temperatures, transesterification reactions can dramatically change the repeating unit distribution of the L and D repeating units of poly(lactic acid) derived from lactides and tend to randomize them although differences remain with respect to bernoullian statistics because transesterification reactions seem to be dependent on the configurations of the units vicinal to reacting ester bonds.^[14] Conversion is generally very high although the resulting polymers often contain residues of monomers and catalysts. For biomedical applications, it is recommended to purify once or twice the crude polymers by the dissolution–precipitation method. Careful purification is a critical factor that determines the behavior and the properties of PLA polymers, in particular, the degradation characteristics. Lactides can be copolymerized with many other heterocyclic monomers such as glycolide,^[33] benzyl malolactonate,^[34] and ϵ -caprolactone.^[35] A particular copolymer composed of largely alternating lactyl and glycolyl units was obtained by ROP of 3-methyl-1,4-dioxane-2,5-dione, a cyclic monomer composed of one lactic acid and one glycolic acid residue.^[36] Such copolymer does not seem to have been tested for biomedical applications yet.

High molar mass PLA polymers have been obtained by postcondensation of oligomeric polycondensates of lactic acids either by using a coupling reagent^[17–37] or a chain extender after appropriated functionalization of chain ends.^[38,39] In the latter case, the resulting polymer chains can contain nonmetabolite residues whose presence must be confronted to requirements of the biomedical applications.

Lactides can lead to diblock and triblock copolymers when polymerization is initiated by tin octanoate in the presence of monoalcohol or diol oligomers. The best example is that of methyl-OH poly(ethylene glycols) and diOH polyethylene glycols.^[39] Particular chain ends can be introduced if low molecular weight monoalcohols are used as cocatalysts in combination with stannous octoate^[23,25] or polymerization is initiated by alkoxides, for instance.^[40]

CHARACTERIZATION OF LACTIC ACID-BASED POLYMERS

PLA stereocopolymers can be characterized by using many techniques and literature is enriching with data issued from solution and solid-state NMR spectroscopies,^[41] ATD, DSC, and TGA,^[42] IRTF and Raman spectroscopies,^[43] MALDI-tof,^[44] and Capillary Zone Electrophoresis.^[45] Stereocopolymers composed of chains enriched in one of the enantiomeric lactyl units can be advantageously characterized by chiroptical properties, namely Optical Rotatory Dispersion and Circular Dichroism.

Because of the pair addition mechanism related to the dimeric structures of lactides that acts in favor of the formation of stereoregular chains, lactide-derived stereocopolymers exhibit higher stereoregularity than polycondensates issued from a bernouillian polycondensation of enantiomeric lactic acids. For instance, in the absence of transesterification reactions, DL-lactide leads to partially isotactic PLA₅₀, whereas meso-lactide leads to partially syndiotactic PLA₅₀.^[14] The use of stereoselective polymers can yield more or less semicrystalline racemic PLA₅₀, depending on the initiating system.^[29] PLA₀ and PLA₁₀₀ can form highly semicrystalline and high melting temperature stereocomplexes.^[17] Therefore, combining the effects of molar mass, molar mass distribution, composition in chiral units, distribution in chiral units, and morphology, PLA_x polymers can be made that cover a very large range of physical, thermal, and mechanical properties. The problem is controlling these properties because they can change dramatically during processing, regardless of the processing method, at the sterilization stage and even on storage.^[11]

Literature is sometimes misleading because authors do not identify the PLA polymer they worked with. The confusion starts at the level of the gross chain composition that can be defined either from the M/M or W/W composition of the monomer feed or from the analysis of chain composition by NMR. In the biomedical field, any scientific contribution dealing with PLA_x polymers should mention the identity of the investigated compounds as accurately as possible, a requirement underlined 20 years ago and still not respected.^[11]

DEGRADATION OF LACTIC ACID-BASED POLYMERS

Basically, the degradation of polymeric materials in an animal or human body can result either from cell-mediated or from chemistry-mediated cleavages of labile bonds or from both. In literature, there is a lot of confusion related to the absence of a standard terminology

and to the misuse of words such as degradable, bioabsorbable, biodegradable, bioresorbable, etc. The term degradation should be reserved to degradations in the absence of living cells or when the mechanism of degradation is unknown. Accordingly, biodegradation ought to be reserved to reflect the degradation due to the biological activity of cells. Bioresorbable should be reserved to qualify polymers that are proved eliminated from the human body after degradation and metabolization to CO₂ and H₂O, or because they can be excreted as small enough oligomers through kidney filtration.

For many years, the degradation in aqueous media, i.e., in vitro as well as in vivo, was regarded as homogeneous (bulk erosion),^[47] although surface erosion was claimed in a few cases.^[1] This was primarily due to the fact that molecular weight changes were monitored by viscometry. Since then, Size Exclusion Chromatography (SEC) has revealed that two populations of macromolecules are often observed when rather large-size devices are allowed to degrade in liquid media and in an animal body as well.^[48,49] Indeed, when a device made of a lactic acid-based polymer is placed in contact with an aqueous medium, water penetrates into the specimen and the hydrolytic cleavage of ester bonds starts and autocatalysis as well. Water absorption is thus a critical factor.^[50] For a time, the partially degraded macromolecules remain insoluble in the surrounding aqueous medium, and the degradation proceeds homogeneously. As soon as the molecular weight of some of the partially degraded macromolecules becomes low enough to allow dissolution of the formed oligomers in the surrounding aqueous medium, the diffusion of these oligomers starts within the whole bulk, with the soluble compounds moving slowly to and off the surface while they continue to degrade. This process, which combines diffusion, chemical reaction, and dissolution phenomena, results in a differentiation between the rates of degradation at surface and interior of the matrix.^[22] In vivo, where the medium is buffered, neutralization of terminal carboxyl groups might also contribute to discriminate the surface degradation rate.

Basically there are four main factors that condition the diffusion–reaction–dissolution phenomena: 1) the hydrolysis rate constant of the ester bond; 2) the diffusion coefficient of water within the matrix; 3) the diffusion coefficient of chain fragments within the polymeric matrix; and 4) the solubility of degradation products, generally oligomers, within the surrounding liquid medium from which penetrating water is issued.^[51]

During degradation, hollow structures can be observed depending on the initial configurational structures of polymer chains and on the existence of initially present or degradation-induced semicrystalline domains.^[48] The formation or not of crystalline residues was mapped and discussed in detail, at least for zinc-derived PLA_x and copolymers with glycolic acid.^[52]

The mechanism of heterogenous degradation contributed very much to the understanding of the effects of many other factors, namely matrix morphology, size and shape, presence of chemically reactive compounds within the matrix, and nature of the degradation medium.^[53] The heterogeneous degradation mechanism provided fruitful information on the behavior that can be expected from lactic acid-based polymers where additives or foreign molecules (oligomers, drug, initiator, and solvent residues) are present or have been introduced. For instance, one can mention the case of an entrapped drug in a PLA-based drug delivery system or when an implanted matrix absorbs compounds by absorption from surrounding body fluids. Indeed, any electrostatic interaction between a PLA matrix and acidic, basic, or amphoteric molecules can dramatically affect the degradation characteristics by changing the natural acid-base equilibrium due to the presence of chain end carboxylic groups. For acidic compounds, one can expect faster hydrolysis of ester bonds. In contrast, in the case of basic compounds, two effects can be observed: base catalysis of ester bond cleavage when the basic compound is in excess with respect to acid chain ends, and decrease of the degradation rate in the contrary.^[51] Examples were reported that deal with oligomers,^[54] residual monomers,^[55] coral,^[56] caffeine,^[57] and gentamycin.^[58] Reviews are available.^[12,59,60]

The heterogeneous degradation mechanism helped also very much in understanding the differences found between the degradation rates observed for films, powder, and microspheres issued from the same batch of PLA₅₀ polymer.^[61] The authors showed conclusively that the smaller the size of a polymer device, the slower the degradation rate. From the same logic, one can understand why porous systems degrade at a slower rate than plain ones, especially if dimensions of the considered device are submillimeter.

Recently, attention was paid to the effects of the initiator on the degradation characteristics of some PLA_x polymers. It has been shown that dramatic differences in repeating unit distribution can be observed in the case of stereocopolymers such as PLA₅₀ polymerized in the presence of stannous octoate or zinc metal.^[62] Later on, it was suggested that, in bulk polymerization of lactides, stannous octoate causes more or less chain end modifications by esterification of alcoholic end groups by octanoic acid and generates hydrophobic residues, which resist purification by precipitation from an organic solution with ethanol or water.^[63,64] Such modifications do not occur in the case of zinc metal or zinc lactate initiations. Whether similar effects can be observed in the case of polymers obtained using other initiating systems, or through solution polymerization, or in the presence of alcohols is still unknown. Indeed, scientists rarely investigated

the effect of resulting structures on degradation characteristics and comparison of literature data is seldom feasible significantly. Nevertheless, the influence of chain end hydrophobization on degradation characteristics of PLA_x polymers is well exemplified by the difference between the so-called H and non H polymers appearing in the catalog of one of the greatest suppliers of PLA_x polymers, namely Boehringer Ingelheim (Germany). At this point, people ought to keep in mind that degradation of lactic acid-based matrices and release of low molecular weight compounds from these matrices depend on experimental conditions and on device's history much more than for any other type of polymeric matrix. It is then strongly recommended again that careful identification of the investigated compounds be included in any paper in the field. Anyhow, this concept of heterogeneous degradation in aqueous media helped very much to understand the in-vivo behaviors of PLA polymers.

CHARACTERISTICS OF THE DEGRADATION OF PLA_x POLYMERS IN VIVO

Most of the basic information regarding biocompatibility, configuration-dependence of degradability and crystallinity, and routes of elimination after ¹⁴C labeling was identified very early.^[2] Until 1981, when it was shown that PLA₁₀₀ and LA-rich stereocopolymers prepared in the presence of zinc metal and carefully purified can last for a year in vivo with less than 10% molecular weight decrease,^[11] literature had reported much shorter half lifetime.^[65-67] Despite these discrepancies, it was conclusively shown that the in-vivo degradability of PLA polymers, copolymers, and stereocopolymers depends on the gross composition and on the content in chiral units.^[68,69] Since then, most of the contributions concluded to excellent biocompatibility of PLA polymers.^[70,71] Year after year, literature was enriched in information on devices made of PLA_x polymers, such as experimental implants for bone surgery^[3,72] or microspheres and nanospheres for drug delivery.^[60] From a general viewpoint, a majority of authors agree to say that poly(α -hydroxy acids) do not degrade enzymatically in an animal body. Discussions were generally based on the comparison of in-vivo and in-vitro degradation characteristics, differences being assigned to the action of enzymes. However the correlation cannot be straightforward. As the particularities of the chemistry-based hydrolytic degradation of PLA polymers showed, one can find many reasons to see different oligomer diffusion and solubilization between a device implanted in vivo and the same device allowed to age in an abiotic medium modeling body fluids. Phenomena like adsorption of proteins, absorption of lipids, and greater solubility of lactic

acid-based oligomers in blood are examples of sources of difference.^[73] Superoxide ions can also contribute if one refers to data reported in the case of sutures.^[74] Anyhow, it has been found that in-vitro and in-vivo degradation characteristics are comparable in the case of small cylinders made of lactic acid-based polymers implanted subcutaneously in rats.^[75] In agreement with these findings, parallel-sided plates implanted intramuscularly in rabbits and similar plates allowed to age in isoosmolar pH = 7.4 phosphate buffer showed similar degradation characteristics except shape modifications that were greater in vivo than in vitro according to the effect of the mechanical stresses generated by muscles.^[76] Nevertheless, the subject is the source of endless debates^[77,78] in the absence of careful characterization of the investigated compounds, including processing and sterilization protocols that are known to affect the initial properties of LA-based polymers. To model body fluids, it is recommended to take into account the physiological environmental pH, ionic strength, and temperature as much as possible. Using 0.13 M pH = 7.4 isosmolar phosphate buffer at 37°C is a current choice although the presence of phosphate ions can be perturbing. The use of any other pH = 7.4 isoosmolar buffered medium is possible.

The glass transition temperature of PLA_x stereocopolymers is in the range of 55–60°C, depending on the composition in chiral units.^[12] It is a critical factor in the case of implanted polymers because absorbed water acts as a plasticizer, thus resulting in a decrease of the glass transition temperature up to below body temperature (37°C). The diffusion of soluble oligomers or drug molecules outside a PLA_x matrix is also very much related to glass transition. Last but not least, raising the temperature to shorten degradation time can be the source of unexpected phenomena that are able to change the regime of degradation from typical of glassy state to typical of degradation in a rubbery state.

All these particularities are of great interest to control the lifetime and the degradation behavior of lactic acid-based polymers in vivo. Nowadays, everybody agrees to conclude that LA-based polymers are remarkably biocompatible. However, secondary and even tertiary inflammatory responses can be observed, respectively, when the release of soluble oligomers becomes critical as in the case of matrices remaining amorphous during the whole degradation process, and later on when crystalline microdomains are released under the form of tiny particles that are very inflammatory and degradation-resistant. Anyhow, once formed during degradation, any tiny lactic acid-based crystalline particulate will be phagocytosed sooner or later to undergo intracellular degradation.^[79] PLA stereocomplexes are normally more resistant to hydrolytic degradation than PLA₁₀₀ polymers, although contradictory data have been published recently.^[80] The best

method to monitor the fate of PLA polymers undergoing degradation within complex living media such as an animal body is radiolabeling. Recently, a method to tritiate lactide and glycolide by catalytic isotopic exchange in the presence of palladium on calcium carbonate was reported that is applicable to any lactic acid-based polymer.^[81,82]

LACTIC ACID-BASED POLYMERS' BIOMEDICAL APPLICATIONS

From a general viewpoint, PLA_x polymers have a high potential as degradable and bioresorbable polymers to make therapeutic devices aimed at helping the body to self repair, i.e., when the prosthetic system is required for a limited period of time.

It is the case of sutures. However, whereas the good fiber-forming, highly crystalline, and rapidly degradable poly(glycolic acid) is well adapted to make sutures, PLA_x semicrystalline polymers are not recommended for this kind of application. Indeed, PLA_x fibers are small-size devices and highly crystalline, two factors that work against the fast degradation requested for the healing of soft tissue. In contrast, rather long lasting PLA_x polymers are of interest to make degradable bone fracture fixation devices such as screws, plates, pins, staples, etc. Nowadays, several devices are on the market for the temporary internal fixation of bone fracture (screws, plates, pins) and anterior cruciate ligament autograft (interference screw) from several companies (Figs. 1 and 2).^[6,83] As early as 1978, a totally bioresorbable composite system composed of PGA fibers reinforcing PLA₁₀₀ matrices, the long-lasting matrix protecting the fastly degraded fibers, was patented.^[84] Later on, other composite systems were proposed such as self-reinforced PGA plates where PGA fibers were embedded in a PGA matrix.^[85] SR-PGA systems went up to clinical and industrial applications. The new generation

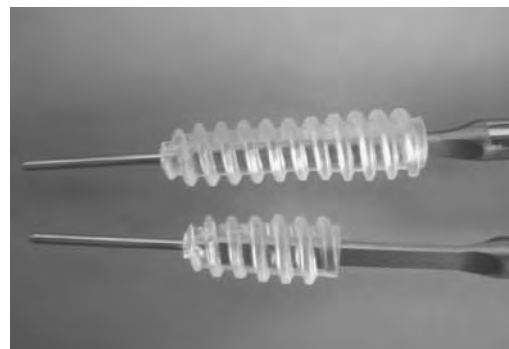


Fig. 1 Injection-molded PLA₉₈ bioresorbable screws for fixation of the bony graft in anterior cruciate ligament reconstruction. (Photo by PHUSIS France.)



Fig. 2 Intervertebral bioresorbable PLA₉₈ cage. (Photo by PHUSIS France.)

of self-reinforced devices is based on PLA polymers.^[86] PLA_x are also worthwhile compounds for bone tissue reconstruction,^[68,70] in particular, they have a potential in maxillar and mandibular reconstruction.^[87] Many potential applications have been proposed such as tiny tubes to help nerve regeneration.^[88]

In drug delivery, the literature is full of examples of PLA_x-based drug formulations and devices such as implants, microparticles, and solid nanoparticles or nanoaggregates made of self-assembling di- and tri-block copolymers.^[89,90] Several reviews have been issued that cover well a very rich field.^[5,60]

Nowadays, one of the most stimulating areas of potential applications is tissue engineering, a technique that consists in culturing cells on a degradable or biodegradable scaffold for the sake of implantation or for cell-growth stimulation and tissue reconstruction. Tissues like skin, bone, and cartilage are the favorite targets of scientists.^[91,92] PLA₅₀ plasticized with oligoPLA₅₀ has been successfully used in periodontal-guided tissue regeneration in animals.^[93] Among the various degradable and biodegradable scaffolds presently under prospect, those made of PLAGA 3D porous devices seem to be the most attractive.^[91,92,94,95] The potential applications have been reviewed.^[96] Another challenge that is considered is the development of bioresorbable stents.

Last but not least, high-scale production of low-cost PLA polymers is becoming a reality, thanks to U.S. and Japanese companies. However, the brands that are commercialized for environmental protection and applications^[97,98] are formulated with additives, and thus, they should be far from being comparable to biomedical grades and they should be considered with great care in case one wants to test their potential in relation to therapeutic applications in humans.

CONCLUSIONS

Lactic acid-based polymers have a great interest in the biomedical and pharmaceutical fields. For several

applications, they have reached the clinical and industrial stages. However, one must keep in mind that PLA is a generic name. Behind this name there are a lot of very different compounds for the reasons we tried to underline in this much-too-short presentation. The future is likely to be oriented toward medicated PLA matrices and devices, thanks to the ability of PLA polymers to accommodate antibiotic and/or antitumoral drugs as well as larger molecules like growth factors and other proteins. Another trend targets the development of functionalized lactic acid-based polymers that should pave the route to more hydrophilic systems, including hydrogels, to cross-linking and to chemical modification.

ARTICLES OF FURTHER INTEREST

Biodegradable Polymers, An Overview; Biofunctional Polymers; Bone Plates and Screws, Bioabsorbable; Elastomers, Biodegradable; Poly(Glycolic Acid); Polymers; Sutures; Tissue Engineering Scaffolds



REFERENCES

1. Kronenthal, R.L. Biodegradable Polymers in Medicine and Surgery. In *Polymers in Medicine and Surgery*; Kronenthal, R.L., User, Z., Martin, E., Eds.; Plenum Publishing Corp.: New York, 1974; 119–137.
2. Kulkarni, R.K.; Pani, K.; Neuman, C.; Leonard, F. Polylactic acid for surgical implants. *Arch. Surg.* **1966**, *93*, 839–843.
3. Vert, M.; Christel, P.; Chabot, F.; Leray, J. Bioresorbable Plastic Materials for Bone Surgery. In *Macromolecular Biomaterials*; Hastings, G.W., Ducheyne, P., Eds.; CRC Press: New York, 1984; 119–141.
4. Wise, D.L.; Fellman, T.D.; Sanderson, J.E.; Wentworth, R.L. Lactic/Glycolic Polymers. In *Drug Carriers in Biology*; Gregoriadis, G., Ed.; Academic Press: London, 1979; 237–270.
5. Holland, S.J.; Tighe, B.J.; Gould, P.L. Polymers for biodegradable medical devices. 1. The potential of polyesters as controlled macromolecular release systems. *J. Control. Release* **1986**, *4*, 155–180.
6. Barber, A.F. Resorbable fixation devices: A product guide. *Orthop. Spec. Ed.* **1998**, *4*, 11–17.
7. Kulkarni, R.K.; Moore, E.G.; Hegyeli, A.F.; Leonard, F. Biodegradable poly(lactic acid) polymers. *J. Biomed. Mater. Res.* **1971**, *5*, 169–181.
8. Holten, C.H. *Lactic Acid*; Verlag Chemie: Weinheim, 1971; 21–224.
9. Carothers, W.H.; Dorough, G.I.; Van Natta, F.J. Studies of polymerization and ring formation. X. The reversible polymerization of six-membered cyclic esters. *J. Am. Chem. Soc.* **193**, *54*, 761–772.
10. Lowe, C.H. Preparation of High Molecular Weight Polyhydroxyacetic Ester. U.S. patent 2,668,162, Appl. March 24, 1952.

11. Vert, M.; Chabot, F.; Leray, J.; Chrsitel, P. Biorresorbable polyesters for bone surgery. *Makromol. Chem., Suppl.* **1981**, *5*, 30–41.
12. Li, S.M.; Vert, M. Biodegradation of Aliphatic Polyesters. In *Degradable Polymers. Principles and Applications*; Scott, G., Ed.; Kluwer Academic Publ.: Dordrecht, The Netherlands, 2002; 71–132.
13. Lillie, E.; Schulz, R.C. ^1H and ^{13}C - $\{^1\text{H}\}$ -NMR spectra of stereocopolymers of lactide. *Makromol. Chem.* **1975**, *176*, 1901–1906.
14. Chabot, F.; Vert, M.; Chapelle, S.; Granger, P. Configurational structures of lactic acid stereocopolymers as determined by ^{13}C - ^1H -NMR. *Polymer* **1993**, *24*, 53–59.
15. Ajioka, M.; Enamoto, E.; Suzuki, K.; Yamaguchi, A. Basic properties of polylactic acid produced by the direct condensation polymerization of lactic acid. *Bull. Chem. Soc. Jpn.* **1995**, *68*, 2125–2131.
16. Moon, S.I.; Lee, C.W.; Miyamoto, M.; Kimura, Y. Melt polycondensation of L-lactic acid with Sn(II) catalysts activated by various proton acids: A direct manufacturing route to high molecular weight poly(L-lactic acid). *J. Polym. Sci., A, Polym. Chem.* **2000**, *38*, 1673–1679.
17. Ikada, Y.; Jamshidi, K.; Tsuji, H.; Hyon, S.H. Stereo-complex formation between enantiomeric poly(lactides). *Macromolecules* **1987**, *20*, 904–906.
18. Kricheldorf, H.R.; Kreiser, I. Poly lactones. 13. Transesterification of poly(L-lactide with poly(glycolide), poly(β -propiolactone) and poly(ϵ -caprolactone). *J. Macromol. Sci., Chem.* **1987**, *A24*, 1345–1358.
19. Garlotta, D. A literature review of poly(lactic acid). *J. Polym. Environ.* **2001**, *9*, 63–84.
20. Albertsson, A.-C.; Varma, I.K. Aliphatic Polyesters. In *Biopolymers: Biology, Chemistry, Biotechnology, Applications, Polyesters III*; Doi, Y., Steinbüchel, A., Eds.; Wiley-VCH, Verlag GmbH: Weinheim, Germany, 2002; 25–52.
21. Tsuji, H. Poly lactides. In *Biopolymers: Biology, Chemistry, Biotechnology, Applications, Polyesters III*; Doi, Y., Steinbüchel, A., Eds.; Wiley-VCH, Verlag GmbH: Weinheim, Germany, 2002; 129–178.
22. Vert, M. Polyglycolide and Copolyesters with Lactides. In *Biopolymers: Biology, Chemistry, Biotechnology, Applications, Polyesters III*; Doi, Y., Steinbüchel, A., Eds.; Wiley-VCH, Verlag GmbH: Weinheim, Germany, 2002; 179–202.
23. Kleine, J.; Kleine, H. Unber hochmolekulare, insbesondere optische aktive polyester des milchsäure, ein beitrage zur stereochemie makromolekularer verbindungen. *Makromol. Chem.* **1959**, *30*, 23–38.
24. Kricheldorf, H.R.; Kreiser-Saunders, I.; Stricker, A. Poly lactones 48. Sn-Oct2-initiated polymerizations of lactide: A mechanistic study. *Macromolecules* **2000**, *33*, 702–709.
25. Dubois, P.; Jacobs, C.; Jérôme, R.; Teyssié, P. Macromolecular engineering of polylactones and poly lactides. IV. Mechanism and kinetics of lactide homopolymerization by aluminum isopropoxide. *Macromolecules* **1991**, *24*, 2266–2270.
26. Kowalski, A.; Duda, A.; Penczek, S. Kinetics and mechanism of cyclic esters polymerization initiated with tin(II) octoate. 3. Polymerization of L-L-dilactide. *Macromolecules* **2000**, *33*, 7359–7370.
27. Schwach, G.; Coudane, J.; Engel, R.; Vert, M. Ring opening polymerization of DL-lactide in the presence of zinc-metal and zinc-lactate. *Polym. Int.* **1998**, *46*, 177–182.
28. Zhiyan-Zhong; Dijkstra, P.J.; Birg, C.; Westerhausen, M.; Feijen, J. A novel and versatile calcium-based initiator system for the ring-opening polymerization of cyclic esters. *Macromolecules* **2001**, *34*, 3863–3868.
29. Sarasua, J.R.; Prud'homme, R.E.; Wisniewski, M.; Le-Borgne, A.; Spassky, N. Crystallization and melting behavior of poly lactides. *Macromolecules* **1998**, *31*, 3895–3905.
30. Leray, J.; Vert, M.; Blanquaert, D. Nouveau matériau de prothèse osseuse et son application. FR7628629, September 20, 1976.
31. Vert, M.; Schwach, G.; Engel, R.; Coudane, J. Something new in the field of PLA/GA bioresorbable polymers. *J. Control. Release* **1998**, *53*, 85–92.
32. Schwach, G.; Coudane, J.; Engel, R.; Vert, M. Zn lactate as initiator of DL-lactide ring opening polymerization and comparison with Sn octoate. *Polym. Bull.* **1996**, *37*, 771–776.
33. Deasy, P.B.; Finan, M.P.; Meegan, M.J. Preparation and characterization of lactic/glycolic acid polymers and copolymers. *J. Microencapsul* **1989**, *6*, 369–378.
34. He, B.; Bei, J.; Wang, S. Synthesis and characterization of a functionalized biodegradable copolymer: Poly(L-lactide-co-RS- β -malic acid). *Polymer* **2003**, *44*, 989–994.
35. Barakat, I.; Dubois, Ph.; Grandsfils, Ch.; Jerome, R. Poly(ϵ -caprolactone-b-glycolide) and poly(D,L-lactide-b-glycolide) diblock copolyesters: Controlled synthesis, characterization, and colloidal dispersions. *J. Polym. Sci., A, Polym. Chem.* **2001**, *39*, 294–306.
36. Dong, C.M.; Qiu, K.Y.; Gu, Z.W.; Feng, X.D. Living polymerization of D,L-3-methylglycolide initiated with bimetallic (Al/Zn) μ -oxo alkoxide and copolymers thereof. *J. Polym. Sci., A, Polym. Chem.* **2001**, *39*, 357–367.
37. Inata, H.; Matsumura, S. Chain extenders for polyester. II: Reactivities of carboxyl-addition-type chain extenders; bis cyclic-imino-ethers. *J. Appl. Polym. Sci.* **1986**, *32*, 5193–5202.
38. Younes, H.; Cohn, D. Morphological study of biodegradable PEO/PLA block copolymers. *J. Biomed. Mater. Res.* **1987**, *21*, 1301–1316.
39. Kurcock, P.; Penczek, H.J.; Franek, J.; Jedlinski, Z. Anionic polymerization of lactones. XIV: Anionic block copolymerization of δ -valerolactone and L-lactide initiated with potassium methoxide. *Macromolecules* **1992**, *25*, 2285–2289.
40. Dubois, P.; Degée, P.; Ropson, N.; Jerome, R. Macromolecular Engineering of Polylactones and Poly lactides by Ring Opening Polymerization. In *Macromolecular Design of Polymeric Materials*; Hatada, K., Kitayama, T., Vogel, O., Eds.; Marcel Dekker Inc.: New York, 1997; 247–272.
41. Zell, M.T.; Padden, B.E.; Paterick, A.J.; Thakur, K.A.M.; Kean, R.T.; Hillmyer, M.A.; Munson, E.J.

- Unambiguous determination of the ^{13}C and ^1H NMR stereosequence assignments of polylactide using high resolution solution NMR spectroscopy. *Macromolecules* **2002**, *35*, 7700–7707.
42. Urayama, H.; Kanamori, T.; Kimura, Y. Microstructure and thermomechanical properties of glassy polylactides with different optical purity of the lactate units. *Macromol. Mater. Eng. (Print)* **2001**, *286*, 705–713.
 43. Kister, G.; Cassanas, G.; Vert, M. Effects of morphology, conformation and configuration on the IR and Raman spectra of various poly(lactic acid)s. *Polymer* **1998**, *39*, 267–273.
 44. Libiszowski, J.; Kowalski, A.; Duda, A.; Penczek, S. Kinetics and mechanism of cyclic esters polymerization initiated with covalent metal carboxylates. 5. End-group studies in the model ϵ -caprolactone and L,L-dilactide/tin(II) and zinc octoate/butyl alcohol systems. *Macromol. Chem. Phys. (Print)* **2002**, *203*, 1694–1701.
 45. Braud, C.; Devarieux, R.; Garreau, H.; Vert, M. Capillary electrophoresis to analyze water-soluble oligo (hydroxy acids) issued from degraded or biodegraded aliphatic polyesters. *J. Environ. Polym. Degrad.* **1996**, *4*, 135–148.
 46. Vert, M.; Sélégny, E. The optical properties of chiral centres as tools for studying irreversible or reversible reactions of synthetic polymers in solution. *J. Polym. Sci., C, Polym. Symp.* **1973**, *42*, 1239–1248.
 47. Pitt, C.G.; Gratzel, M.M.; Kimmel, G.L.; Surles, J.; Schindler, A. Aliphatic polyesters. 2. The degradation of poly(DL-lactide), poly (ϵ -caprolactone) and their complexes in vivo. *Biomaterials* **1981**, *2*, 215–220.
 48. Li, S.; Garreau, H.; Vert, M. Structure–property relationships in the case of degradation of solid aliphatic poly (α -hydroxy acids) in aqueous media: 3. Amorphous and semi-crystalline PLA100. *J. Mater. Sci., Mater. Med.* **1990**, *1*, 198–206.
 49. Li, S.; Garreau, H.; Vert, M. Structure–property relationships in the case of degradation of solid aliphatic poly (α -hydroxy acids) in aqueous media: 2. PLA37.5GA25 and PLA75GA25 copolymers. *J. Mater. Sci., Mater. Med.* **1990**, *1*, 131–139.
 50. Schmitt, E.A.; Flanagan, D.R.; Lindhart, R.J. Importance of distinct water environments in the hydrolysis of poly(DL-lactide-co-glycolide). *Macromolecules* **1994**, *27*, 743–748.
 51. Vert, M. Bioresorbable Synthetic Polymers and Their Operation Field. In *Biomaterials in Surgery*; Walenkamp, G., Ed.; Georg Thieme Verlag: Stuttgart, 1998; 97–101.
 52. Vert, M.; Li, S.; Garreau, H. Attempts to map structure and degradation characteristics of aliphatic polyesters derived from lactic and glycolic acids. *J. Biomater. Sci., Polym. Ed.* **1994**, *6*, 639–649.
 53. Vert, M. Degradation of Polymeric Biomaterials with Respect to Temporary Applications. In *Degradable Materials*; Baremberg, S.A., Brash, J.L., Narayan, R., Redpath, A.E., Eds.; CRC Press: Boca Raton, 1990; 11–37.
 54. Mauduit, J.; Perouse, E.; Vert, M. Hydrolytic degradation of films prepared from blends of high and low molecular weight poly(DL-lactic acid)s. *J. Biomed. Mater. Res.* **1996**, *30*, 201–207.
 55. Cordewiner, F.W.; Rozema, F.R.; Bos, R.R.; Boering, G. Material properties and tissue reaction during degradation of poly(96L/4D-lactide)—A study in vitro and in rats. *J. Mater. Sci., Mater. Med.* **1995**, *6*, 211–218.
 56. Li, S.M.; Vert, M. Hydrolytic degradation of coral/poly(DL-lactic acid) bioresorbable materials. *J. Biomater. Sci., Polym. Ed.* **1996**, *7*, 817–827.
 57. Li, S.M.; Girod-Holland, S.; Vert, M. Hydrolytic degradation of poly(DL-lactic acid) in the presence of caffeine base. *J. Control. Release* **1996**, *40*, 41–53.
 58. Mauduit, J.; Bukh, N.; Vert, M. Gentamycin/poly (lactic acid) blends aimed at sustained release local antibiotic therapy administered per-operatively. III. The case of gentamycin sulfate in films prepared from high and low molecular weight poly(DL-lactic acids). *J. Control. Release* **1993**, *25*, 43–49.
 59. Li, S.; Vert, M. Biodegradable Polymers: Polyesters. In *The Encyclopaedia of Controlled Drug Delivery*; Mathiowitz, Ed.; J. Wiley & Sons: New York, 1999; 71–93.
 60. Brannon-Peppas, L.; Vert, M. Poly(lactic and Polyglycolic Acids as Drug Delivery Carriers. In *Handbook of Pharmaceutical Controlled Release Technology*; Wise, L.R., Klibanov, D.L., Mikos, A., Brannon-Peppas, L., Peppas, N.A., Trantalo, D.J., Wnek, G.E., Michael, J., Yaszemski, M.J., Eds.; Marcel Dekker: New York, 2000; 90–130.
 61. Grizzi, I.; Garreau, H.; Li, S.; Vert, M. Biodegradation of devices based on poly(DL-lactic acid): Size dependence. *Biomaterials* **1995**, *16*, 305–311.
 62. Schwach, G.; Engel, R.; Coudane, J.; Vert, M. Stannous octoate versus zinc-initiated polymerization of racemic lactide: Effect of configurational structures. *Polym. Bull.* **1994**, *32*, 617–623.
 63. Zhang, X.; MacDonald, D.A.; Goosen, M.F.A.; McAuley, K.B. Mechanism of lactide polymerization in the presence of stannous octoate: The effect of hydroxy and carboxylic acid substances. *J. Polym. Sci., A, Polym. Chem.* **1994**, *32*, 2965–2970.
 64. Schwach, G.; Coudane, J.; Engel, R.; Vert, M. More about the initiation mechanism of lactide polymerization in the presence of stannous octoate. *J. Polym. Sci., A, Polym. Chem.* **1997**, *35*, 3431–3440.
 65. Cutright, D.E.; Perez, B.; Beasley, J.D.; Larson, W.J.; Rosey, W.R. Degradation rates of polymers and copolymer of polylactic and polyglycolic acids. *Oral Surg.* **1974**, *37*, 142–152.
 66. Miller, R.A.; Brady, J.M.; Cutright, D.E. Degradation rates of oral resorbable polyesters for orthopaedic surgery. *J. Biomed. Mater. Res.* **1977**, *11*, 711–719.
 67. Guilding, D.K.; Reed, A.M. Biodegradable polymers for use in surgery—Polyglycolic/poly(lactic acid) homo- and copolymers: 1. *Polymer* **1979**, *20*, 1459–1464.
 68. Sedel, L.; Chabot, F.; Christel, P.; de Charantenay, F.X.; Leray, J.; Vert, M. Les implants biodégradables en chirurgie orthopédique. *Rev. Chir. Orthop.* **1978**, *64*, 92–101.
 69. Reed, A.M.; Guilding, D.K. Biodegradable polymers for use in surgery—Poly(glycolic)/poly(lactic acid) homo and copolymers. *Polymer* **1981**, *22*, 494–498.
 70. Vert, M.; Li, S.M.; Spenlehauer, G.; Guerin, P. Bioresorbability and biocompatibility of aliphatic polyesters. *J. Mater. Sci., Mater. Med.* **1992**, *3*, 432–446.



71. Ignatius, A.A.; Claes, L.E. In vitro biocompatibility of bioresorbable polymers: Poly(L,DL-lactide) and poly(L-lactide-co-glycolide). *Biomaterials* **1996**, *17*, 831–839.
72. Hollinger, J.O. Preliminary report on the osteogenic potential of a biodegradable copolymer of polylactide (PLA) and polyglycolide (PGA). *J. Biomed. Mater. Res.* **1983**, *17*, 71–82.
73. Sharma, C.P.; Williams, D.F. The effects of lipids on the mechanical properties of polyglycolic acid structures. *Eng. Med.* **1981**, *10*, 8–10.
74. Lee, K.H.; Chu, C.C. The role of superoxide ions in the degradation of synthetic absorbable sutures. *J. Biomed. Mater. Res.* **2000**, *49*, 25–35.
75. Kenley, R.A.; Ott Lee, M.; Mahoney, T.R.; Sanders, L. Poly(lactide-co-glycolide) decomposition kinetics in vivo and in vitro. *Macromolecules* **1987**, *20*, 2398–2403.
76. Thérin, M.; Christel, P.; Li, S.; Garreau, H.; Vert, M. Degradation of massive poly(α -hydroxy acids) in aqueous living medium: In vivo validation of in vitro findings. *Biomaterials* **1992**, *13*, 594–600.
77. Gogolewski, S.; Jovanovic, M.; Perren, S.M.; Dillon, J.G.; Hughes, M.K. Tissue response and in vivo degradation of selected polyhydroxyacids: Polylactides (PLA), poly(3-hydroxybutyrate-co-3-hydroxyvalerate) (PHB/VA). *J. Biomed. Mater. Res.* **1993**, *27*, 1135–1148.
78. Mainil-Varlet, P.; Curtis, R.; Gogolewski, S. Effect of in vivo and in vitro degradation on molecular and mechanical properties of various low-molecular weight polylactides. *J. Biomed. Mater. Res.* **1997**, *36*, 360–380.
79. Tabata, Y.; Ikada, Y. Macrophage phagocytosis of biodegradable microspheres composed of L-lactic acid/glycolic acid homo and copolymers. *J. Biomed. Mater. Res.* **1988**, *22*, 837–858.
80. Serizawa, T.; Arikawa, Y.; Hamada, K.; Yamashita, H.; Fujikawa, T.; Kimura, Y.; Akashi, M. Alkaline hydrolysis of enantiomeric poly(lactide)s stereo-complexes deposited on solid substrates. *Macromolecules* **2003**, *36*, 1762–1765.
81. Dos Santos, I.; Morgat, J.L.; Vert, M. Hydrogen isotope exchange as a means of labeling lactides. *J. Labelled Compd Radiopharm.* **1998**, *41*, 1005–1012.
82. Dos Santos, I.; Morgat, J.L.; Vert, M. Glycolide deuteration by hydrogen isotope exchange using the HSCIE method. *J. Labelled Compd Radiopharm.* **1999**, *42*, 1093–1101.
83. Middleton, J.C.; Tipton, A.J. Synthetic biodegradable polymers as orthopaedic devices. *Biomaterials* **2000**, *21*, 2325–2346.
84. Vert, M.; Chabot, F.; Leray, J.; Christel, P. Nouvelles pièces d'ostéosynthèse, leur préparation et leur application. FR782987, October 20, 1978.
85. Törmälä, P.; Vasenius, J.; Vainionpää, S.; Laiho, J.; Pohjonen, T.; Rokkanen, P. Ultra high strength absorbable self-reinforced polyglycolide (SR-PGA) composite rods for internal fixation of bone fractures: In vitro and in vivo study. *J. Biomed. Mater. Res.* **1991**, *25*, 22.85.
86. Joukainen, A.; Pihlajamäki, H.; Makela, E.A.; Ashammakhi, N.; Viljanen, J.; Patiala, H.; Kellomäki, M.; Tormala, P.; Rokkanen, P. Strength retention of self-reinforced drawn poly-L/DL-lactide 70/30 (SR-PLA70) rods and fixation properties of distal femoral osteotomies with these rods. An experimental study on rats. *J. Biomater. Sci., Polym. Ed.* **2000**, *11*, 1411–1428.
87. Chanavaz, M.; Chabot, F.; Donazzan, M.; Vert, M. Further Clinical Applications of Bioresorbable PLA 37.5 GA 25 and PLA 50 Polymers for Limited Bone Augmentation and Bone Replacement in Human. In *Biological and Biomechanical Performance of Biomaterials*; Christel, P., Meunier, A., Lee, A.J.C., Eds.; Elsevier, 1986; 233–238.
88. Oudega, M.; Gautier, S.E.; Chapon, P.; Frago, M.; Bates, M.L.; Parel, J.-M.; Bunge, M.B. Axonal regeneration into Schwann cell grafts within resorbable poly(α -hydroxyacid) guidance channels in the adult rat spinal cord. *Biomaterials* **2001**, *22*, 1125–1136.
89. Anderson, J.M.; Shive, M.S. Biodegradation and biocompatibility of PLA and PLGA microspheres. *Adv. Drug Deliv. Rev.* **1997**, *28*, 5–24.
90. Brannon-Peppas, L. Recent advances on the use of biodegradable microparticles and nanoparticles in controlled drug delivery. *Int. J. Pharm.* **1995**, *116*, 1–9.
91. Agrawal, C.M.; Niedersuer, C.C.; Micallef, D.M.; Athanasiou, K.A. The Use of PLA-PGA Polymers in Orthopaedics. In *Encyclopedic and Book of Biomaterials and Bioengineering*; Marcel Dekker: New York, 1995; 2081–2115. Chap. 30.
92. Athanasiou, K.A.; Schmitz, J.P.; Agrawal, C.M. The effects of porosity on in vitro degradation of polylactic acid-polyglycolic acid implants used in repair of articular cartilage. *Tissue Eng.* **1998**, *4*, 53–63.
93. Robert, P.M.; Frank, R.M. Periodontal guided tissue regeneration with a new bioresorbable polylactic acid membrane. *J. Periodontol.* **1994**, *65*, 414–422.
94. Ishaug, S.L.; Yashemski, M.J.; Bizios, R.; Mikos, A.G. Osteoblast function on synthetic biodegradable polymers. *J. Biomed. Mater. Res.* **1994**, *28*, 1445–1453.
95. Holder, W.D., Jr.; Gruber, H.E.; Moore, A.L.; Culberson, C.R.; Anderson, W.; Burg, K.J.L.; Mooney, D.J. Cellular ingrowth and thickness changes in poly-L-lactide and polyglycolide matrices implanted subcutaneously in the rat. *J. Biomed. Mater. Res.* **1998**, *41*, 412–421.
96. Ikada, Y.; Tsuji, H. Biodegradable polyesters for medical and environmental applications. *Rapid Commun.* **2000**, *21*, 117–132.
97. Gruber, P.; O'Brien, M. Polylactides Nature Works™ PLA. In *Biopolymers: Biology, Chemistry, Biotechnology, Applications, Polyesters III*; Doi, Y., Steinbüchel, A., Eds.; Wiley-VCH, Verlag GmbH: Weinheim, Germany, 2002; 235–250.
98. Kawashima, N.; Ogawa, S.; Obuchi, S.; Matsuo, M.; Yagi, T. Polylactic Acid «Lacea». In *Biopolymers: Biology, Chemistry, Biotechnology, Applications, Polyesters III*; Doi, Y., Steinbüchel, A., Eds.; Wiley-VCH, Verlag GmbH: Weinheim, Germany, 2002; 251–274.

Poly(lactide-co-glycolic acid) (PLGA)

Konstantinos Avgoustakis

Laboratory of Pharmaceutical Technology, Department of Pharmacy, University of Patras, Patras, Rio, Greece

INTRODUCTION

Poly(lactide-co-glycolide) (PLGA) is a synthetic copolymer of lactic acid (α -hydroxy propanoic acid) and glycolic acid (hydroxy acetic acid). Lactic acid contains an asymmetric carbon atom, and therefore has two optical isomers: L(+) lactic acid and D(-) lactic acid. Lactic acid is present in nature as either an intermediate or an end product in carbohydrate metabolism. It is widely distributed in all living creatures (man, animals, plants, and microorganisms). Glycolic acid occurs in nature to a limited extent. Poly(lactide-co-glycolide) degrades *in vivo* to innocuous products. Its final degradation products are lactate (salt form of lactic acid) and glycolate (salt form of glycolic acid).

Poly(lactide-co-glycolide) can be synthesized by direct polycondensation of lactic acid and glycolic acid. However, the most efficient route to obtain high-molecular weight (MW) copolymers is the ring opening polymerization of lactide and glycolide. Lactide and glycolide are the cyclic diesters of lactic acid and glycolic acid, respectively, and they are prepared by thermal degradation of lactic and glycolic acid oligomers, respectively. Being a biocompatible and biodegradable polymer with adjustable properties and capable of being processed to form a variety of objects, PLGA finds extensive biomedical applications, such as sutures, orthopedic fixation devices, and drug delivery systems. Poly(lactide-co-glycolide) scaffolds are investigated for tissue engineering applications. The synthesis, properties, and biomedical applications of PLGA are discussed in this entry.

SYNTHESIS OF PLGA

Poly(lactide-co-glycolide) can be prepared by the direct polycondensation of lactic and glycolic acid (Scheme 1). The polycondensation reaction can be effected in solution or in the melt/solid state.

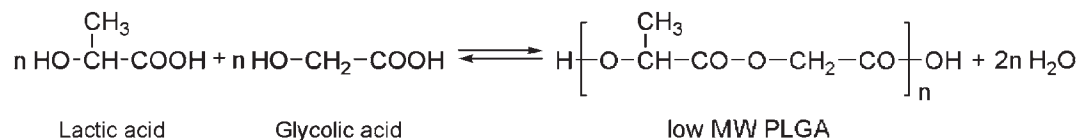
As a high degree of dehydration, required for the production of high-MW polymer, is difficult to achieve, polycondensation was considered to be an inadequate method to obtain high polymers.^[1] However, Ajioka et al. prepared high-MW (160 kDa) poly(L-lactic-co-glycolic acid) by azeotropic

dehydration of a solution of L-lactic acid and glycolic acid in diphenylether for 20–40 hr at 130°C using tin powder as catalyst.^[2] The solvents introduce complexity of both process control and purification of the end product. Therefore, polycondensation is preferred to be effected in the melt state. The melt polycondensation system involves two equilibria: the dehydration/hydration for ester formation and the ring/chain equilibrium for cyclic diester formation. In ordinary melt polycondensation, the cyclic diesters lactide and/or glycolide (Scheme 2) are formed by depolymerization because the reaction conditions of high temperature and high vacuum can induce not only dehydration but also formation of the cyclic diesters in equilibrium with the polymer. This prevents further growth of the polymer chain, and a high-MW polymer cannot be obtained.

To overcome this problem, a melt/solid polycondensation method was recently developed (Scheme 3).^[3] In the first step, oligomers are prepared by dehydration of monomers. In the second step, the oligomers are melt-polymerized under vacuum (10 Torr) using, for example, zinc acetate dihydrate or tin chloride dihydrate/*p*-toluenesulfonic acid as catalysts. A third step then follows in which condensation is continued at the solid state for 10 hr under high vacuum (0.5 Torr). It is believed that in the solid state, the polymerization reaction is favored over the depolymerization or other side reactions. With this melt/solid polycondensation method, polymers with MW comparable to the polymers produced by the ring opening polymerization of the cyclic diesters can be produced.

To prepare high-MW PLGA in a relatively short reaction time, it is necessary to proceed by ring-opening polymerization of the cyclic diesters, lactide, and glycolide (Scheme 4). Lactide and glycolide are produced by thermal degradation of lactic and glycolic acid oligomers, respectively, and are commercially available. As the lactide molecule bears two asymmetric carbon atoms, lactide exists in the form of three diastereoisomers (L-, D-, and *meso*-lactide) and a racemate (50/50 mixture of L- and D-lactides commonly referred to as DL-lactide) (Scheme 2). The ring opening polymerization reaction can be conducted at temperatures higher than the melting point of the polymer





Scheme 1 Polycondensation of lactic and glycolic acid.

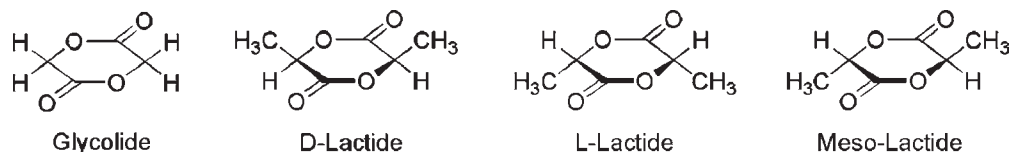
(melt polymerization), at temperatures lower than the melting point of the polymer (solid-state polymerization), or in solution.

Polymers having high MW can be synthesized using a variety of catalysts, such as powdered zinc, Lewis acids (e.g., zinc chloride and antimony trifluoride), or organometallic compounds (e.g., triethyl aluminum or stannous octoate).^[1,4,5] Stannous octoate is the most commonly used catalyst because of its acceptance by the U.S. Food and Drug Administration as a food stabilizer. When organometallic catalysts, such as diethyl zinc, are employed, the polymerization reaction is considered to follow a coordination mechanism.^[6] The initiation of the reaction is considered to involve a nucleophilic attack of the catalyst species on one of the lactide or glycolide carbonyls leading to the opening of the ring and formation of active species. Lactide/glycolide molecules are then coordinated to the active species to create the polymer chain (Scheme 5).

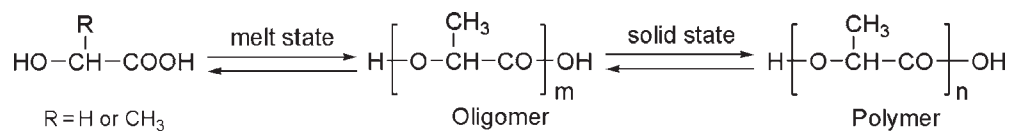
The situation with the stannous octoate-catalyzed polymerization is more complex. Based on kinetic and spectroscopic data, Kowalski, Duda, and Penczek concluded that stannous octoate is by itself neither an initiator nor a catalyst of lactide polymerization.^[7] Reaction of stannous octoate with hydroxyl containing compounds (e.g., alcohols and water), added in the polymerization feed or present as impurities of stannous octoate, provides the initiating species. Chain growth proceeds by a coordination–insertion mechanism in the $\cdots\text{Sn}-\text{OR}$ bond in a similar way to that shown for $\cdots\text{Zn}-\text{OR}$, as shown in Scheme 5.^[7,8] Transesterification reactions toward the conclusion of polymerization (i.e., at high conversion) may cause a broadening of the MW distribution.^[5,6] The polymerization was found to be first order with regard to the monomer (lactide/glycolide) concentration.^[6,7] The rate of polymerization was initially increased with the concentration of stannous octoate in the polymerization feed, but after reaching a maximum it eventually decreased with further increasing stannous octoate concentration.^[7]

Owing to chain growing by dimer (lactide or glycolide) addition (Scheme 5), the polymer chains that result from the ring opening polymerization of lactide and glycolide are actually polydimer chains. Thus, PLGA formed by ring-opening polymerization of lactide and glycolide is more accurately referred to as poly(lactide-*co*-glycolide). Because of the addition of pairs of repeating units during chain growth, and the existence of different stereoisomers of lactide, poly(lactide) (PLA) homopolymers and PLGA copolymers having different and rather complex configurational structures can be obtained through ring-opening polymerization. Using different mixtures of lactide stereoisomers, stereocopolymers (PLA and PLGA polymers containing L- and D-lactic acid units) having different physical and biological properties can be synthesized.^[4]

Lactide and glycolide of high purity are required to obtain high-MW copolymers. The exclusion of moisture is also essential because it may cause termination of the chain growth or hydrolysis of the monomers to acidic products. Thus, the polymerization is carried out either in high vacuum or in inert atmosphere (e.g., in nitrogen). In practice, the amount and type of the catalyst determine the temperature and the time required to produce high-MW polymer in high yield. As a rule, the higher the amount of catalyst the shorter the time required to produce high-MW polymer. Also, the amount of catalyst required to produce high-MW polymer decreases as the temperature of polymerization increases.^[9] When organometallic catalysts, such as triethyl aluminum and stannous octoate, are employed in lactide/glycolide polymerization, the MW of the polymer initially increases sharply with the catalyst level, and after reaching a maximum it eventually decreases with further increase of the catalyst level.^[1,9,10] The ascending part of the polymer MW vs. catalyst level curve, observed at low catalyst/monomer ratios, may be attributed to the increased rate of polymerization with increasing catalyst levels. The descending part of the polymer MW vs. catalyst-level



Scheme 2 Structures of glycolide and lactide.



Scheme 3 Melt/solid polycondensation of lactic and glycolic (R = H) acid.

curve, observed at high catalyst/monomer ratios, may be attributed to the formation of a higher number of initiating species at high catalyst levels. A higher number of initiating species would initiate a greater number of polymerization reactions, leading to the formation of relatively small chains, i.e., to a decreased MW of the synthesized polymer. In support of the latter consideration, Kricheldorf, Kreiser-Saunders, and Stricker, using relatively high stannous octoate proportions in L-lactide polymerizations at 180°C, found that the MWs of synthesized polymers paralleled the monomer/stannous octoate ratio.^[8] As glycolide is more reactive than lactide, PLGA copolymer chains may contain blocks of glycolide units rather than a random distribution of glycolide and lactide units.^[5] Indeed, PLGA copolymers produced by melt polymerization under vacuum were found to have heterogeneous microstructure, based on ¹³C-NMR spectra.^[9]

The ring-opening polymerization of lactide and glycolide in bulk (melt or solid polymerization) is usually carried out at temperatures ranging between 130°C and 220°C over a period of 2–6 hr. Stannous octoate is added at 0.03–0.1% w/w proportions. Under these conditions, copolymers with molecular weights in the range 10⁴–10⁵ are prepared at high conversions (higher than 90%).^[5] Long reaction periods at high temperatures may cause significant depolymerization (chain unzipping) and should therefore be avoided.^[9,10] It is pertinent to remember here that the cyclic diesters lactide and glycolide are prepared by thermal breakdown of low-MW PLA and PGA, respectively. For PLGA copolymers synthesized under identical conditions, increasing the glycolide content of the feed resulted in copolymer of lower MW, indicating that more severe reaction conditions are required to polymerize glycolide than lactide.^[9] A pilot-scale polymerization process for the preparation of PLGA has been described in detail.^[11]

As previously mentioned, minor reactions, such as ester interchange and chain unzipping, may take place during lactide/glycolide polymerization, which lead to

the formation of low-MW species (e.g., cyclic dimers, trimers, cyclic oligomers, and even monomers). These, together with unreacted monomers and catalyst residues, may adversely affect the properties of the final product, such as mechanical strength and degradation rate, and should therefore be removed by careful purification (once or twice) of the crude polymer. The purification can be effected by first dissolving the polymerization product in dichloromethane or chloroform and then precipitating the polymer in excess methanol or diethyl ether. The purified polymer is collected by filtration and dried. Alternatively, after the polymerization is complete, residual monomers are removed from the polymer melt by applying vacuum.^[11] The purified polymer is stored under vacuum or is packaged under nitrogen in moisture-proof packages.

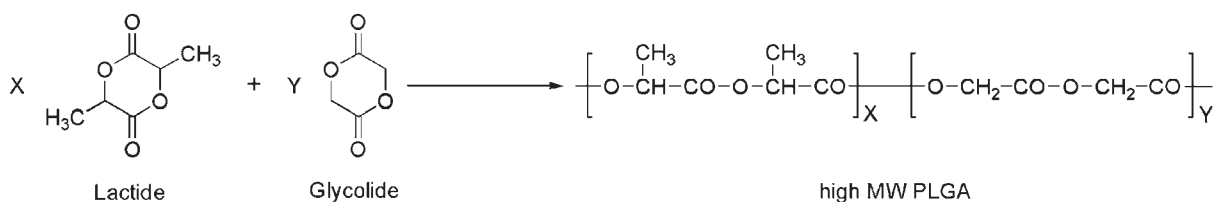
P

POLY(LACTIDE-CO-GLYCOLIDE) PROPERTIES

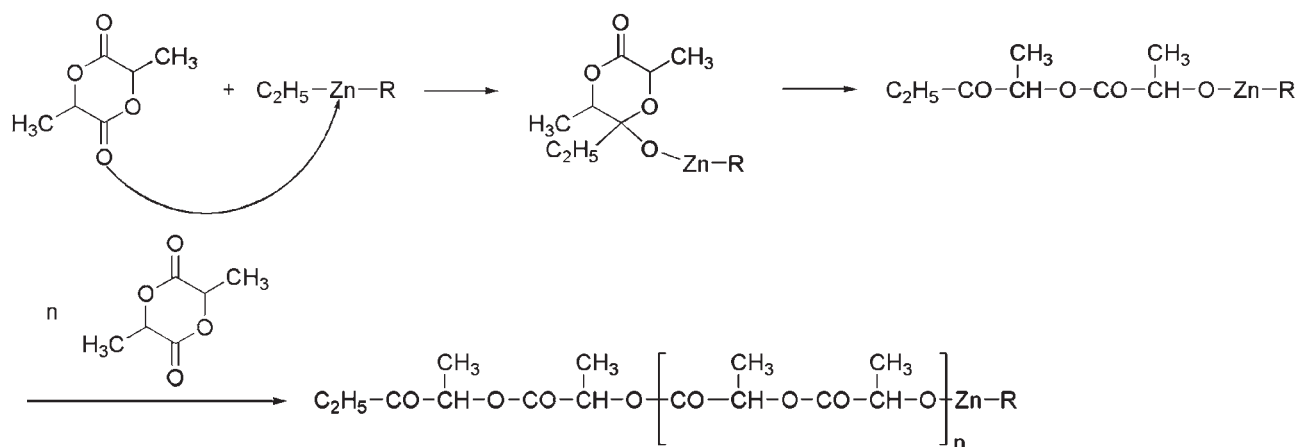
A distinct advantage of PLGA copolymers, and the one that characterizes this family of polymers, is that they can be designed so as to acquire specific, desired properties. A broad spectrum of performance characteristics can be obtained by careful manipulation of three key properties of the copolymer: composition (lactide/glycolide ratio), lactide stereoisomeric composition (L- or DL-lactide), and MW. The notation L-PLGA(*X*:*Y*) or DL-PLGA(*X*:*Y*) will be used in this entry to describe PLGA copolymers consisting of *X* mol% L- or DL-lactide, respectively, and *Y* mol% glycolide.

Solubility

Solubility in common organic solvents is important with regard to the ease of polymer characterization, processing, and application. Poly(lactide-co-glycolide) containing less than 50% glycolyl units (glycolic acid units) is soluble in most common organic solvents,



Scheme 4 Ring opening polymerization of lactide and glycolide.



Scheme 5 Coordination mechanism of polymerization of lactide.

such as halogenated hydrocarbons (chloroform and dichloromethane), ethyl acetate, acetone, dioxan, and tetrahydrofuran. Poly(lactide-*co*-glycolide) rich in glycolyl units (50% and higher) is insoluble in most organic solvents, and uncommon solvents, such as hexafluoroisopropanol, have been used in characterization and processing of PLGA with high content of glycolyl units.

Crystallinity

Crystallinity affects the rate of degradation and the mechanical properties of PLGA. Poly(lactide-*co*-glycolide) crystallinity depends on chemical composition (lactide/glycolide ratio) and stereoisomeric composition of lactide units in the copolymer.^[5,12] Poly(DL-lactide-*co*-glycolide) consisting of 0–75% glycolyl units is amorphous. Poly(L-lactide-*co*-glycolide) with 25–75% glycolyl units is also amorphous. The crystallinities, glass transition temperatures, and melting points of poly(L-lactide-*co*-glycolide) copolymers are presented in Fig. 1.

Thermal Stability

Poly(lactide-*co*-glycolide) copolymers are thermoplastic materials exhibiting adequate heat stability in the absence of moisture. Thus, they can be melt processed to produce sutures, orthopedic fixation devices, and drug delivery systems. Poly(lactide-*co*-glycolide) degrades to lactide and glycolide after prolonged heating above 200°C under nitrogen or vacuum.^[5] At lower temperatures, thermal degradation is a function of time and temperature and is considerably accelerated by impurities, residual monomers, and humidity.

Mechanical Properties

The mechanical properties (strength, toughness, and elasticity) of PLGA depend on composition (lactide/glycolide ratio), lactide stereochemistry, MW, and processing.^[12,13] Crystalline polymers, such as poly(L-lactide) and poly(glycolide), exhibit significantly higher mechanical strength than the amorphous poly(DL-lactide) and the amorphous or less crystalline PLGA. Certain mechanical properties of commonly used lactide/glycolide polymers are presented in Table 1.

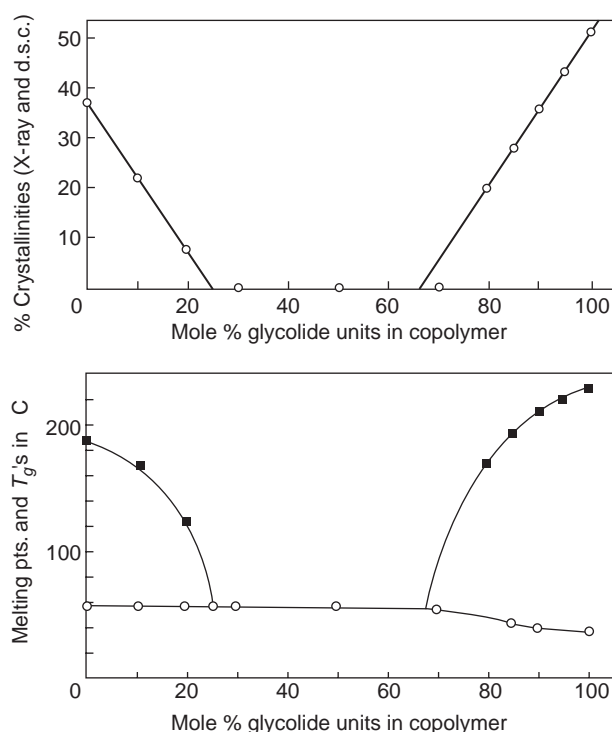


Fig. 1 Crystallinities (upper graph) and glass transition temperatures (○) and melting point (■) (lower graph) for lactide/glycolide polymers. (From Ref. [5].)

Table 1 Mechanical properties of commercially available lactide and glycolide homopolymers and copolymers

Polymer type	Tensile strength (psi $\times 10^{-3}$)	Elongation (%)	Modulus (psi $\times 10^{-5}$)
L-PLA	8–12	5–10	4–6
DL-PLA	4–6	3–10	2–4
DL-PLGA(85:15)	6–8	3–10	2–4
DL-PLGA(75:25)	6–8	3–10	2–4
DL-PLGA(65:35)	6–8	3–10	2–4
DL-PLGA(50:50)	6–8	3–10	2–4
PGA	>10	15–20	10

Information was provided by the suppliers.

Degradation

Poly(lactide-*co*-glycolide) degrades *in vitro* through hydrolysis of ester bonds. The final degradation products *in vitro* are lactic acid and glycolic acid. Poly(lactide-*co*-glycolide) degradation has been proposed to be a two-stage process.^[14] During the first stage, random hydrolytic cleavage of ester bonds occurs, leading to MW reduction. The chain scission process appears to follow first-order kinetics. The second stage is characterized by the onset of mass loss from the polymer matrix, accompanied by an increase in the rate of chain scission. Mass loss begins when the polymer has been degraded to a degree where chain fragments small enough to dissolve in the degradation medium and diffuse out from the polymer matrix have been generated. Because esterification of carboxylic acid chain ends (end-capping) to reduce the initial carboxylic end group concentration was shown to retard degradation, chain scission is believed to be autocatalyzed by the generated carboxylic acid end groups.^[14,15]

After the early work on PLGA, hydrolytic degradation was regarded as homogenous (bulk erosion).^[14] However, more recent research has shown that relatively large-sized PLGA devices degrade (both *in vitro* and *in vivo*) through a heterogenous process, with the degradation of the core being faster than that of the surface of the device.^[16] In Fig. 2, the core of poly(DL-lactide-*co*-glycolide) (50:50) matrices (cylindrical tablets) was completely degraded (dissolved away, so that the matrix appeared to be centrally hollow under the electron microscope) after 20 days' incubation in phosphate buffer, pH 7.4.^[17] Vert and coworkers introduced the concept of heterogenous PLGA degradation related to "diffusion–reaction–dissolution" phenomena: the soluble oligomers that eventually form during polymer degradation can escape faster from the surface than from the interior of the matrix, resulting in a smaller autocatalytic effect at the

carboxyl-depleted surface and, consequently, a lower degradation rate at the surface than in the interior of the matrix.^[16] According to the diffusion–reaction–dissolution model, relatively large PLA devices, which degrade via a heterogenous mechanism, were found to degrade faster than relatively small PLA devices that degrade via a homogenous mechanism.^[16]

In vivo degradation of PLGA proceeds through hydrolysis of ester bonds, as the *in vitro* degradation. No involvement of the living tissues (cells and enzymes) is considered to take place during the early stages of degradation, apart from that resulting from the foreign body response. During foreign body reaction, cells accumulating around the polymer implant may produce free radicals and acidic products that could accelerate polymer hydrolysis. Enzyme-catalyzed PLGA degradation *in vivo* has been claimed by some investigators based on differences between the *in vitro* and the *in vivo* degradation characteristics. Although degradation has been shown to be accelerated by enzymes *in vitro*, no convincing data have been presented for enzymatic hydrolysis of PLGA *in vivo*.^[18] Enzymes are not the only body constituent that could affect PLGA degradation. Phenomena like adsorption of proteins, absorption of lipids, and greater solubility of PLGA oligomers in blood may be the sources of differences between the *in vitro* and *in vivo* degradation.^[19] However, when degradation has advanced to the stage that implant coherence is lost and the matrix begins to fragment, living tissues participate actively in further polymer degradation; small enough fragments (less than 10 μm) are assimilated by phagocytes and are further hydrolyzed intracellularly to monomeric anions (lactate and glycolate).^[11,20] Lactate and glycolate are also the final products of PLGA extracellular hydrolysis. The final stage of *in vivo* degradation is elimination. During this stage, L-lactate is converted into carbon dioxide and pyruvate. Pyruvate then enters the Krebs cycle. D-Lactate is not metabolized and has only been detected in excreta.^[1,21] Glycolate is excreted



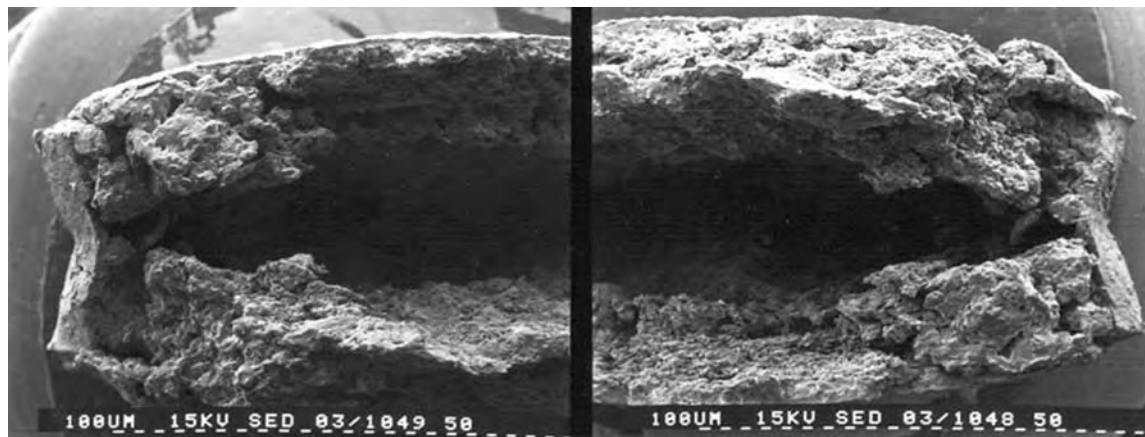


Fig. 2 Scanning electron microscope picture of a DL-PLGA (50:50) tablet after 20 days of incubation in phosphate buffer, pH 7.4 at 37°C. (From Ref. [17].)

directly in the urine or may be oxidized to glyoxylate, which is then converted to glycine, serine, and pyruvate. Pyruvate then can enter Krebs cycle, as before, with lactate.^[11] Several studies reported that PLGA degrades at similar rates in vitro (in media modeling body fluids) and in vivo. Other studies, however, found faster PLGA degradation in vivo than in vitro. This discrepancy is probably due to differences in the experimental conditions, degradation parameters measured, and types of polymers involved in these studies. From a PLGA device development perspective, an in vivo degradation rate higher to that measured in vitro in simulated body fluids should be taken into account.

The factors affecting the rate of PLGA degradation are well documented in the literature and are those related to the following.^[20,22,23]

1. Polymer characteristics, such as chemical composition (lactide/glycolide ratio) monomer distribution pattern, chain-ends chemical composition, MW, MW distribution, and polymer purity (e.g., residual monomers).^[24]
2. Device characteristics, such as size, shape, porosity, and presence of additives (e.g., acidic or basic compounds, plasticizers, or drugs).
3. Melt-processing, annealing, and sterilization.
4. Environment in which degradation takes place, such as site of in vivo implantation, and the pH and temperature of the degradation medium.

Chemical composition and lactide stereoisomeric composition are the most influencing factors for PLGA degradation as they determine, among other properties, the hydrophilicity and the crystallinity of the polymer. Among the first to evaluate the effect of composition on PLGA degradation rate in vivo were Miller, Brady, and Cutright.^[25] They found that PLGA (50:50) was the fastest degrading composition, with the

degradation rate being decreased when either lactide or glycolide content of the copolymer was increased (Fig. 3). Although the degradation rates reported may be a little higher than those obtained in recent years, probably due to better standardization of the experimental conditions and better characterization of polymers in more recent degradation studies, the trends shown in Fig 3 have been confirmed by studies conducted after the work of Miller, Brady, and Cutright.^[25]

The effect of composition on PLGA degradation (Fig. 3) can be explained based on the combined effects that the changes of composition have on copolymer hydrophilicity and crystallinity. An increase in hydrophilicity increases the rate of PLGA degradation, whereas an increase in crystallinity has the opposite effect. Hydrophilicity increases as the proportion of the more hydrophilic glycolide in the copolymer increases and crystallinity increases as the proportion of

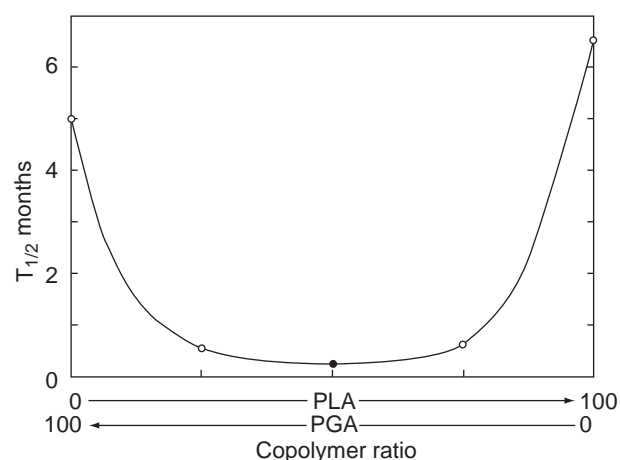


Fig. 3 Half-life in months of lactide/glycolide polymers implanted in rat tissue. (From Ref. [25].)

glycolide or L-lactide units exceeds 75% (Fig. 1). Increased hydrophilicity and decreased crystallinity account for the increased PLGA degradation rate as the lactide content decreases from 100 mol% to 50 mol% (Fig. 3). In addition, glycolic linkages are hydrolyzed faster than the lactic linkages and this may also contribute to the observed increase in degradation rate with increasing glycolide content of PLGA copolymers from 0 to 50 mol% glycolide (Fig 3).^[22] However, with a further increase of the glycolide content, crystallinity of the copolymer increases and outweighs the effect of increased copolymer hydrophilicity on degradation. The net result is a decrease in degradation rate as the glycolide content of PLGA increases from 50 to 100 mol% (Fig 3). By changing the lactide/glycolide ratio and lactide stereoisomeric composition, polymers having *in vivo* degradation times from a few weeks to more than 2yr can be produced (Table 2).

Not only the monomer proportion, but also the distribution pattern of the two monomers in the copolymer chains (copolymer “microstructure”) may affect the degradation of PLGA copolymers that are rich in glycolide. As glycolide is more reactive than lactide, these copolymers may contain blocks of glycolide rather than a random sequence of glycolide and lactide.^[5] At the same monomer ratio, the copolymers having a homogenous microstructure (random monomer sequence) are expected to degrade faster than the copolymers having a heterogenous microstructure.^[26]

Poly(lactide-*co*-glycolide) degradation time increases with the average MW of the copolymer.^[15,27] Apart from the average MW, the MW distribution may also affect PLGA degradation. A wide MW distribution would indicate the presence of relatively large numbers of carboxylic end groups, which can facilitate the autocatalytic degradation of the polymer chains.^[20]

The chemical composition of chain ends could significantly affect PLGA degradation. The chemical

composition of the chain terminus depends on the type of catalyst/initiator used in polymerization, but it can be modified after PLGA preparation. It has been reported, for example, that stannous octoate causes chain-end modification by esterification of alcoholic end groups by octanoic acid and generates hydrophobic residues.^[28] Such modifications do not occur with zinc metal or zinc lactate initiators, and, as a result, these initiators provide PLGA polymers with different physicochemical and degradation properties as compared with stannous octoate. The influence of the chain-end composition on PLGA degradation has been exploited by PLGA suppliers (e.g., Boehringer Ingelheim in Germany and Alkermes in the United States), which have commercialized copolymers with free carboxyl end groups exhibiting faster degradation than normal copolymers of the same composition but with ester end groups (Table 2).

Sterilization by ionizing γ - or β -rays has been shown to cause radiation dose-dependent degradation of PLGA.^[5,29] On the contrary, sterilization by ethylene oxide does not appear to adversely affect PLGA properties. Both γ irradiation and ethylene oxide have been applied to sterilize commercial PLGA devices. Recently, plasma sterilization was introduced as a sterilizing method for polyester devices. A low-temperature radiofrequency plasma treatment was applied to sterilize PLA pins. Treatment with oxygen or carbon dioxide plasmas at a power of 100 W for 15 min sterilized the pins efficiently. The method caused a slight increase of the MW of the polymer, whereas it did not adversely affect the crystallinity and the mechanical properties of the polymer.^[30]

Biocompatibility and Tissue Reactions

The biocompatibility and nontoxicity of PLGA polymers were first demonstrated with the application of these polymers in the production of biodegradable sutures. Dexon (PGA) and Vicryl L-PLGA (8:92) sutures have successfully been used in the clinic for more than 30 yr.

In general, host responses to a polymeric implant are affected by polymer physical and chemical properties and implant properties (volume, shape, and surface characteristics) and are considered to be tissue-dependent, organ-dependent, and species-dependent.^[20] Several studies have shown that PLGA implantation in bone or soft tissues of animals causes none or only a mild inflammatory response, which diminishes with time.^[31–33] No toxicity or allergic responses were observed.

Soft tissue responses to PLGA occur in three phases.^[20] The first phase includes the initiation, resolution, and organization of the acute and chronic

Table 2 Approximate resorption times of commercially available lactide and glycolide homopolymers and copolymers

Polymer	Biodegradation time (months)
L-PLA	>24
DL-PLA	12–16
DL-PLGA (85:15)	5–6
DL-PLGA (75:25)	4–5
DL-PLGA (50:50)	1–2
DL-PLGA (50:50) H ^a	<1
PGA	6–12

Information was provided by the suppliers.

^aCopolymer with free carboxyl end groups.



inflammatory response. A minimal inflammatory reaction with the presence of polymorphonuclear leukocytes, lymphocytes, plasma cells, and monocytes is observed. The second phase is characterized by the presence of foreign body reactions with the accumulation of monocytes, which differentiate to macrophages, which in turn may fuse or join together to form giant cells, and fibroblast infiltration, which leads to implant encapsulation (and progressive invasion) by connective tissue. Formation of blood capillaries (neovascularization), which surround and invade the implant, is also observed. The third phase is characterized by implant fragmentation and phagocytosis of small fragments by macrophages and foreign body giant cells. Also, fiber formation and neovascularization are enhanced and connective tissue replaces the resorbed polymer. The duration of the second and third phases depends on the rate of polymer degradation. The implants are finally totally resorbed and replaced by connective tissue. Similar responses are observed when PLGA is implanted intraosteally.^[31,33]

A mild foreign body reaction is elicited on PLGA implantation with accumulation of histiocytes and mast cells in close vicinity to the implant. A bony capsule builds around the implant with a fibrous interface separating the osteoid line from the polymer. Polymer degradation is accompanied by a fibrous and vascular invasion of the cracks that develop. The implant is progressively replaced by bone, and polymer fragments resulting from polymer degradation are surrounded by giant and mononucleated cells. Polymer debris is observed intracellularly. The implant is finally replaced by woven bone.

Clinical trials of PLGA application in man as sutures and fracture fixation devices (rods, pins, screws, and plates) have confirmed the biocompatibility of PLGA polymers.^[34-36] Late-stage severe foreign body reactions in about 5-8% of patients with PGA rod implants have been reported.^[11] The factors that may contribute to late-stage body reactions are numerous and complex. Polymer purity appears to be important. This is because a much lower incidence of late-stage

Table 3 Commercially available PLGA sutures, orthopedic fixation devices, and drug delivery systems

Trade name	Manufacturer	Application	Polymer type
<i>Sutures and suture anchors</i>			
Vicryl	Ethicon	Sutures	L-PLGA (8:92)
Polysorb	U.S. Surgical	Sutures	L-PLGA
Biologically Quiet Biosphere	Instrument Makar	Suture anchors	DL-PLGA (85:15)
Biologically Quiet Mini-screw	Instrument Makar	Suture anchors	DL-PLGA (85:15)
SD sorb 2 mm and 3 mm, SD sorb E-Z TAC	Surgical Dynamics	Suture anchors	L-PLGA (82:18)
<i>Orthopedic fixation devices</i>			
Biologically Quiet Interference Screw	Instrument Makar	Interference screw	DL-PLGA (85:15)
Lactosorb Screws and Plates	Biomet	Cranio-maxillofacial fixation	L-PLGA (82:18)
Biologically Quiet Staple	Instrument Makar	ACL reconstruction	DL-PLGA (85:15)
SD sorb Meniscal Staple	Surgical Dynamics	Meniscus repair	L-PLGA (82:18)
BiosorbPDX Screws and Tacks	Bionx	Cranio-maxillofacial fixation	Self-reinforced L-PLGA (80:20)
SmartPinPDX	Bionx	Fracture fixation	Self-reinforced L-PLGA (80:20)
<i>Drug delivery systems</i>			
Zoladex		Delivery of goserelin acetate in prostate cancer	DL-PLGA
Nutropin Depot		Delivery of human growth hormone in growth deficiencies	DL-PLGA
Trelstar Depot		Delivery of triptorelin pamoate in prostate cancer	DL-PLGA
Sandostatin LAR		Delivery of octreotide in acromegaly	DL-PLGA

ACL, Anterior cruciate ligament.

foreign body reactions has been observed with purified melt processed PLA than with bulk PLA, which may have contained a relatively high level of residual lactide.^[11] It has also been suggested that the occurrence of inflammatory reactions may depend on the anatomical region, the capacity of the tissue to clear the degradation products, and the volume of polymer implanted.^[37] Tiainen et al. noted that the complications in the use of resorbable polymeric devices reported in the literature mainly concerned homopolymers (L-PLA and PGA) and not PLGA copolymers.^[33]

BIOMEDICAL APPLICATIONS

Lactide/glycolide homopolymers and copolymers have been applied in the clinic as sutures, fixation devices in bone surgery (plates, screws, and pins), and drug delivery systems. Commercially available PLGA products are presented in Table 3. Poly(glycolide) and glycolide-rich poly(L-lactide-*co*-glycolide) have good fiber-forming properties and suitable (relatively fast) degradation rate to be applied as sutures.^[34]

The use of PLGA in bone surgery offers significant advantages. First, implant degradation obviates the need for a second surgical event for removal. Second, the implant can be engineered to degrade at a rate that will slowly transfer load to the healing bone, resulting in a healed bone, which is strong and has no tendency to refracture on implant removal. The latter is a serious problem with metal implants, where the load is exclusively carried by the rigid implant during bone healing. Crystalline PLA and L-lactide-rich poly(L-lactide-*co*-glycolide) have appropriate mechanical properties (high strength and tensile/flexural modulus), especially when reinforced, and the degradation rate is slow enough to be applied for orthopedic fixation.^[38,39] Self-reinforced, amorphous PLGA copolymers of adequate mechanical strength and appropriate rate of *in vivo* degradation have recently been developed as osteofixation materials in craniofacial surgery.^[33]

Amorphous PLGA polymers have received wide attention as excipients in controlled drug delivery systems and as antigen carriers in the development of novel vaccines.^[40-42] Apart from their established biocompatibility, the main reason behind the extensive investigation of PLGA polymers for drug delivery applications is their controlled degradation. By controlling the rate of *in vivo* degradation, the rate of drug release can be controlled and delivery systems with tailor-made drug release properties can be developed. For example, in the area of cancer treatment, a single once-a-month injection of a PLGA depot of goserelin provides a continuous supply of the required amount of drug for 1 mo and has replaced 30 daily injections

of the drug. Currently, a number of approved drug products in the market utilize PLGA polymers (Table 3).

Tissue engineering can be used to restore, maintain, or enhance tissues and organs. Poly(lactide-*co*-glycolide) porous scaffolds have been proposed as three-dimensional templates to guide tissue regeneration.^[43]

CONCLUSIONS

Poly(lactide-*co*-glycolide) copolymers represent an important family of polymers for medicine. They are biocompatible and biodegradable, degrading *in vivo* to innocuous products. In some cases, foreign body reactions on *in vivo* implantation can be minimized by careful control of polymer properties (e.g., purity or degradation rate). Poly(lactide-*co*-glycolide) properties can be tailored for the specific application by simply changing the lactide/glycolide ratio and/or lactide diastereoisomeric structure. Poly(lactide-*co*-glycolide) application in controlled drug delivery and orthopedics has grown in the last two decades and is expected to grow further in the years to come. Owing to the progress in molecular biology and genetics, an ever-increasing number of novel types of drugs (recombinant proteins and genes) will become available. Efficient *in vivo* administration of these drugs will require the development of novel delivery technologies involving, among others, PLGA carriers. Poly(lactide-*co*-glycolide)-based vaccines, reducing or even eliminating the need for booster immunizations or capable of inducing the appropriate type of immune responses (e.g., cytotoxic T cell responses against viral infections or cancer), will be clinically tested. Tissue engineering is the new field where PLGA copolymers are expected to make a major impact in the near future. Clinical application of PLGA scaffolds in tissue engineering is expected to be realized, when the progress in fabrication techniques and scaffolds properties will allow the PLGA scaffolds to meet important clinical design criteria, such as the ability to be fabricated during surgery and being tailored for specific applications. The scaffolds can be designed to release, at a controlled rate, growth factors that induce cellular differentiation and tissue growth *in vitro* or cell migration into the wound site *in vivo*.^[44] Copolymers of PLGA with other polymers may also find significant biomedical applications in the years to come. For example, block copolymers of PLGA with poly(ethylene glycol) have been applied for the preparation of surface engineered nanoparticulate drug carriers. These carriers exhibit long-circulation properties in blood following intravenous administration and are currently being investigated for drug-targeting applications.^[45]

ARTICLES OF FURTHER INTEREST

Biodegradable Polymers, An Overview; Biofunctional Polymers; Poly(Glycolic Acid); Poly(lactic acid)s

REFERENCES

- Wise, D.L.; Fellman, T.D.; Sanderson, J.E.; Wentworth, R.L. Lactic/glycolic acid polymers. In *Drug Carriers in Biology and Medicine*; Gregoriadis, G., Ed.; Academic Press: New York, 1979; 237–270.
- Ajioka, M.; Suizu, H.; Higuchi, C.; Kashima, T. Aliphatic polyesters and their copolymers synthesized through direct condensation polymerization. *Polym. Degrad. Stab.* **1998**, *39*, 137–143.
- Takahashi, K.; Taniguchi, I.; Miyamoto, M.; Kimura, Y. Melt/solid polycondensation to obtain high-molecular-weight poly(glycolic acid). *Polymer* **2000**, *41*, 137–143.
- Chabot, F.; Vert, M.; Chapelle, S.; Granger, P. Configurational structures of lactic acid stereocopolymers as determined by ^{13}C - ^1H n.m.r.. *Polymer* **1983**, *24* (2), 53–59.
- Gilding, D.K.; Reed, A.M. Biodegradable polymers for use in surgery-polyglycolic/poly(lactic acid) homo- and copolymers: 1. *Polymer* **1979**, *20* (12), 137–143.
- Dittrich, W.; Schulz, R.C. Kinetic und Mechanismus der ringöffnenden polymerization von L(-)-Lactid. *Angew. Macromol. Chem.* **1971**, *15*, 109–126.
- Kowalski, A.; Duda, A.; Penczek, S. Kinetics and mechanism of cyclic esters polymerization initiated with tin(II) octoate 3. Polymerization of L,L-dilactide. *Macromolecules* **2000**, *33*, 7359–7370.
- Kricheldorf, H.R.; Kreiser-Saunders, I.; Stricker, A. Poly(lactones) 48. SnOct₂-initiated polymerizations of lactide. *Macromolecules* **2000**, *33*, 702–709.
- Avgoustakis, K.; Nixon, J.R. Biodegradable controlled release tablets 1: preparative variables affecting the properties of poly(lactide-co-glycolide) copolymers as matrix forming material. *Int. J. Pharm.* **1991**, *70*, 77–85.
- Hyon, S.H.; Jamshidi, K.; Ikada, Y. Synthesis of poly(lactides) with different molecular weights. *Biomaterials* **1997**, *18* (22), 1503–1508.
- Perrin, D.E.; English, J.P. Polyglycolide and polylactide. In *Handbook of Biodegradable Polymers*; Domb, A.J., Kost, J., Wiseman, D.M., Eds.; In *Drug Targeting and Delivery*; Florence, A.T., Gregoriadis, G., Series Eds.; Harwood Academic Publishers: Amsterdam, 1997; 3–27.
- Vert, M.; Chabot, F.; Leray, J.; Christel, P. Stereoregular bioresorbable polyesters for orthopaedic surgery. *Macromol. Chem. Suppl.* **1981**, *5*, 30–41.
- Daniels, A.U.; Chang, M.K.; Andriano, K.P. Mechanical properties of biodegradable polymers and composites proposed for internal fixation of bone. *J. Appl. Biomater.* **1990**, *1* (1), 57–78.
- Pitt, C.G.; Gu, Z. Modification of the rates of chain cleavage of poly(ϵ -caprolactone) and related polyesters in the solid state. *J. Controlled Release* **1987**, *4*, 283–292.
- Tracy, M.A.; Ward, K.L.; Firouzabadian, L.; Wang, Y.; Dong, N.; Qian, R.; Zhang, Y. Factors affecting the degradation rate of poly(lactide-co-glycolide) microspheres in vivo and in vitro. *Biomaterials* **1999**, *20*, 1057–1062.
- Grizzi, I.; Garreau, H.; Li, S.; Vert, M. Hydrolytic degradation of devices based on poly(DL-lactic acid) size-dependence. *Biomaterials* **1995**, *16* (4), 305–311.
- Avgoustakis, K. In *Synthesis and Evaluation of Some Poly(Lactide-co-Glycolides) for Use in Sustained Release Tablets*. Ph.D. thesis, University of London, 1992.
- Williams, D.F. Biodegradation of surgical polymers. *J. Mater. Sci.* **1982**, *17*, 1233–1246.
- Vert, M. Polyglycolide and copolyesters with lactides. In *Biopolymers: Biology, Chemistry, Biotechnology, Applications, Vol. 4 Polyesters III*; Doi, Y., Steinbuechel, A., Eds.; Wiley-VCH Verlag GmbH: Weinheim, Germany, 2002; 179–202.
- Anderson, J.M.; Shive, S.M. Biodegradation and biocompatibility of PLA and PLGA microspheres. *Adv. Drug Deliv. Rev.* **1997**, *28*, 5–24.
- Drury, D.R.; Wick, A.N. Chemistry and metabolism of L(+) and D(-) lactic acids. *Am. J. Physiol.* **1956**, *184*, 1061–1069.
- Alexis, F. Factors affecting the degradation and drug-release mechanism of poly(lactic acid) and poly[(lactic acid)-co-(glycolic acid)]. *Polym. Int.* **2005**, *54*, 36–46.
- Vert, M. Degradation of polymeric systems aimed at temporary therapeutic applications: structure-related complications. *e-Polymers* **2005**, *8*, 1–10.
- Hyon, S.H.; Jamshidi, K.; Ikada, Y. Effects of residual monomer on the degradation of DL-lactide polymer. *Polym. Int.* **1998**, *46* (3), 196–202.
- Miller, R.A.; Brady, J.M.; Cutright, D.E. Degradation rates of oral resorbable implants (poly(lactates) and polyglycolates): rate modification with changes in PLA/PGA copolymer ratios. *J. Biomed. Mater. Res.* **1977**, *11*, 711–719.
- Dunn, R.L.; English, J.P.; Strobel, J.D.; Cowsar, D.R.; Tice, T.R. Preparation and evaluation of lactide/glycolide copolymers for drug delivery. In *Polymers in Medicine III*; Migliaresi, C., Nikolais, L., Giusti, P., Chiellini, E., Eds.; Elsevier: Amsterdam, 1988; 149–160.
- Kamei, S.; Inoue, Y.; Okada, H.; Yamada, M.; Ogawa, Y.; Toguchi, H. New method for analysis of biodegradable polyesters by high-performance liquid chromatography after alkali hydrolysis. *Biomaterials* **1992**, *13*, 953–958.
- Schwach, G.; Goudane, J.; Engel, R.; Vert, M. More about the initiation mechanism of lactide polymerization in the presence of stannous octoate. *J. Polym. Sci. A. Polym. Chem.* **1997**, *35*, 3431–3440.
- Loo, S.C.J.; Ooi, C.P.; Boey, Y.C.F. Radiation effects on poly(lactide-co-glycolide) (PLGA) and poly(lactide) (PLLA). *Polym. Degrad. Stab.* **2004**, *83*, 259–265.
- Gogolewski, S.; Mainil-Varlet, P.; Dillon, J.G. Sterility, mechanical properties, and molecular stability of poly(lactide) internal-fixation devices treated with low-temperature plasmas. *J. Biomed. Mater. Res.* **1996**, *32*, 227–235.

31. Vert, M.; Christel, P.; Chabot, F.; Leray, J. Bioresorbable plastic materials for bone surgery. In *Macromolecular Biomaterials*; Hastings, G.W., Ducheyne, P., Eds.; CRC Press: Boca Raton, FL, 1984; 119–142.
32. Visscher, G.E.; Robison, R.L.; Maulding, H.V.; Fong, J.W.; Pearson, J.E.; Argentieri, G.J. Biodegradation of and tissue reaction to 50:50 poly(DL-lactide-co-glycolide) microcapsules. *J. Biomed. Mater. Res.* **1985**, *19*, 349–365.
33. Tiainen, J.; Soini, Y.; Tormala, P.; Waris, T.; Ashammakhi, N. Self-reinforced poly(lactide)/poly(glycolide) 80/20 s crews take more than 11/2 years to resorb in rabbit cranial bone. *J. Biomed. Mater. Res. B Appl. Biomater.* **2004**, *70*, 49–55.
34. Racey, G.L.; Wallace, W.R.; Cavalaris, J.; Marguard, J.V. Comparison of a polyglycolic–poly(lactide) acid suture to black silk and plain catgut in human oral tissues. *J. Oral Surg.* **1978**, *36*, 766–770.
35. Rokkanen, P.; Vainionpaa, S.; Tormala, P.; Kilpikari, J.; Bostman, O.; Vihtonen, K.; Laiho, J.; Tamminmaki, M. Biodegradable implants in fracture fixation: early results of treatment of fracture of the ankle. *Lancet* **1985**, *1*, 1422–1424.
36. Ashammakhi, N.; Renier, D.; Arnaud, E.; Marshac, D.; Ninkovich, M.; Donaway, D.; Jones, B.; Serlo, W.; Laurikainen, K.; Tormala, P.; Waris, T. Successful use of Biosorb osteofixation devices in 165 cranial and maxillofacial cases: a multicenter report. *J. Craniofac. Surg.* **2004**, *15* (4), 692–701.
37. Athanasiou, K.A.; Niederauer, G.G.; Agrawal, C.M. Sterilization, toxicity, biocompatibility, and clinical applications of poly(lactide) acid/poly(glycolide) acid copolymers. *Biomaterials* **1996**, *17*, 93–102.
38. Rokkanen, P.; Bostman, O.; Hirvensalo, E.; Partio, E.K.; Makela, E.A.; Patiala, H.; Vihtonen, K. Bioabsorbable implants in orthopaedics. *Curr. Orthop.* **1999**, *13*, 223–228.
39. Middleton, J.C.; Tipton, A.J. Synthetic biodegradable polymers as orthopedic devices. *Biomaterials* **2000**, *21*, 2335–2346.
40. Jain, R.; Shah, N.H.; Malick, A.W.; Rhodes, C.T. Controlled drug delivery by biodegradable poly-(ester) devices: different preparative approaches. *Drug Dev. Ind. Pharm.* **1998**, *24* (8), 703–727.
41. Panyam, J.; Labhasetwar, V. Biodegradable nanoparticles for drug and gene delivery to cells and tissue. *Adv. Drug Deliv. Rev.* **2003**, *55* (3), 329–347.
42. Johansen, P.; Men, Y.; Merckle, H.P.; Gander, B. Revisiting PLA/PLGA microspheres: an analysis of their potential in parenteral vaccination. *Eur. J. Pharm. Biopharm.* **2000**, *50*, 129–146.
43. Mooney, D.J.; Baldwin, D.F.; Suh, N.P.; Vacanti, J.P.; Langer, R. Novel approach to fabricate porous sponges of poly(D,L-lactide-co-glycolic acid) without the use of organic solvents. *Biomaterials* **1996**, *17*, 1417–1422.
44. Griffith, L.G.; Naughton, G. Tissue engineering-current challenges and expanding opportunities. *Science* **2002**, *295*, 1009–1014.
45. Avgoustakis, K. Pegylated poly(lactide) and poly(lactide-co-glycolide) nanoparticles: preparation, properties, and possible applications in drug delivery. *Curr. Drug Deliv.* **2004**, *1* (4), 321–333.

FURTHER READING

- Tsuji, H. Poly(lactides). In *Biopolymers: Biology, Chemistry, Biotechnology, Applications, Vol. 4 Polyesters III*; Doi, Y., Steinbuchel, A., Eds.; Wiley-VCH, Verlag GmbH: Weinheim, Germany, 2002; 129–178.



Polymer Foams

V. Prasad Shastri

Vanderbilt University, Nashville, Tennessee, U.S.A.

Molly M. Stevens

Imperial College London, London, United Kingdom

INTRODUCTION

In this article, the fabrication of polymeric foams (cellular solids) will be discussed in general terms while providing examples of processes that are suitable for the processing of biodegradable polymers in the context of TE. Parameters that are used to describe foam properties will also be briefly discussed.

OVERVIEW

Polymer foams or polymeric cellular solids can be best described as a dispersion of gas in a solid. Polymer foams may also be described as an emulsion of air in a polymeric phase. The density of the foam dictates the major and minor phases of this air-polymer emulsion, which is a function of its porosity. The extent of porosity and the nature of the pores (gaseous voids) and their relationship to one another governs and dictates the changes in mechanical properties as a solid (dense) polymer evolves into a foam. This process of transitioning from a solid to a less dense gas-filled structure is generally referred to as foaming. Porosity can be introduced in a polymer matrix by a variety of means. The structural and mechanical characteristics of the foamed material are influenced to a significant extent by the foaming process. Because the uses of polymer foams extend from applications as mundane as wall insulations, chair cushions, vibrational dampeners, and shock absorbers to the more esoteric such as bioreactors for separations and scaffolds for tissue engineering (TE), the requirements with respect to the extent and nature of porosity vary over a wide range. Different polymers can be used to generate tailored foams for some of the aforementioned applications. The book on cellular solids by Gibson and Ashby is highly recommended for those seeking a more in-depth analysis and mathematical treatment of foam structure and properties.^[1]

CELLULAR SOLIDS

An important characteristic of a porous material is its relative density (ρ_{rel}). Relative density for a foam is defined as the ratio of the foam density to the density of the virgin material:

$$\rho_{rel} = \frac{\rho_{foam}}{\rho_{polymer}} \quad (1)$$

Foams can be characterized into distinct groups based on their ρ_{rel} . As the ρ_{rel} of a foam increases, the void (cell) walls thicken, the pores become less connected, and the porosity diminishes. Conversely, as the ρ_{rel} decreases, the void walls get thinner, voids (cells) become interconnected, and the porosity increases. As the cells become interconnected and cell walls coalesce, the architecture of the porous solid will evolve from one that is dominated by closed cells to one with an open-cell architecture, eventually resulting in a bicontinuous structure of polymer phase and gaseous phase. The evolution and morphology of the cells (pores) is dictated to a significant extent by the nature of the foaming process and the pore-forming agent, i.e., the porogen. Cellular solids by definition are porous materials with a ρ_{rel} less than 0.3. Above this density the morphology of a porous solid evolves from one of interconnected pores to isolated pores. The porosity of the porous polymer is defined as follows:

$$P = (1 - \rho_{rel}) \quad (2)$$

Cellular solids can possess both closed-cell and open-cell architecture. The transition from a closed-cell to an open-cell architecture occurs when adjacent cells fuse via a process called reticulation. This process is typically accompanied by an increase in porosity and can occur under certain foaming conditions as discussed later.

STRUCTURAL, MORPHOLOGICAL, AND MECHANICAL REQUIREMENTS OF POLYMER FOAMS FOR TE APPLICATIONS

The structural, morphological, and mechanical requirements for polymer foams when used as scaffolds for TE are quite different from those required for more traditional applications.

1. Polymer foam scaffolds must be derived from bicompatible degradable polymers to ensure cytocompatibility, correct evolution of tissue, and *in vivo* biocompatibility. In some instances, a nondegradable bicompatible polymer might be adequate. Applications for a nondegradable porous scaffold material may include cartilage, bone, and certain muscular regenerative applications. On the whole, a degradable polymer is preferred. As a consequence, the choice of synthetic polymers is significantly diminished; usable polymers include poly(esters) derived from α -hydroxyacids, α,ω -hydroxyacids such as caprolactones, orthoesters, and *p*-dioxanone; poly(anhydrides); poly(ethyleneglycol); and other polyols.
2. The pores in polymeric scaffolds for TE have to be highly interconnected to allow for easy diffusion of cells and nutrients during the tissue cultivation process.
3. The polymer content in a scaffold has to be minimal, *i.e.*, the porosity has to be relatively high to ensure rapid resorption of the scaffold so as to maximize the growth of the tissue.
4. The foam scaffold has to be mechanically robust enough to withstand both handling during and after association with mammalian cells and fluid-induced shear stress during cultivation in a bioreactor or after implantation in a tissue site of interest. The mechanical robustness of the scaffold becomes even more critical when mechanical stimulation is used as a means to stimulate tissue development.

FABRICATION OF POLYMER FOAMS

Gas Foaming

Polyurethane, polystyrene, and polyethylene foams are typically manufactured by a gas foaming process. In this process, gas bubbles are introduced within a polymer melt or monomer solution either by blowing an inert gas such as nitrogen into the melt (solution) or by the use of a chemical blowing agent.^[2] Chemical blowing agents are inorganic or organic compounds that liberate gaseous molecules such as carbon dioxide and nitrogen upon heating. In a variation of this

approach, gas bubbles can be generated *in situ* by reaction of one of the monomer components such as isocyanate with water to liberate carbon dioxide. Furthermore, volatile solvents such as halogenated hydrocarbons can be utilized to generate bubbles primarily composed of solvent vapors. The gas is then dispersed using mechanical means such as stirring, and the bubbles are allowed to nucleate and grow. Once the gas bubbles have stabilized, the polymer phase is either rapidly solidified or physically cross-linked, as in the case of polyurethane foams, to impart structural stability. The rate of nucleation and growth of gas bubbles and bubble size is a function of solution viscosity and surface tension. Due to the dynamics of bubble growth and fusion, gas foaming yields voids that have a polyhedral structure.

Porosity can also be introduced in a solid by first saturating the solid with an inert gas under high pressure and then relieving the pressure causing rapid expansion of the gaseous phase. Heating the polymer to its glass transition temperature, where it begins to exhibit viscoelastic flow, further enhances the segregation of the gaseous phase into stable bubbles. This process can yield microcellular plastics with isolated closed-pore and reticulated open-pore morphology. Mooney and coworkers have adapted this approach to produce poly(DL-lactic-co-glycolic acid) (PLGA) foams with partially open-cell architecture for TE applications.^[3] The advantage of this process with regard to the fabrication of scaffolds for TE is that it can be carried out in the absence of solvents. Residual solvents in polymer foams can pose biocompatibility problems *in vivo*.

Removal of a Fugitive Phase

Removal of a fugitive phase is one of the oldest techniques to create porosity in a matrix. It is the preferred method for the production of porous ceramics and metals. In this process, an organic component such as a polymer or wax serves as the fugitive phase and hence the porogen. The desired inorganic phase is dispersed with the organic fugitive phase and then heated above the pyrolysis temperature of the fugitive phase to generate the voids, with concurrent sintering of the inorganic phase. The advantage of this approach is that the porosity and pore morphology can be controlled *a priori*. Yet another advantage is that foams with both isolated closed-pore and open-pore morphology may be obtained. However, because synthetic polymers are not capable of surviving the pyrolysis step, other mechanisms to remove the fugitive phase have been explored.

Pore formation using a water-soluble fugitive phase

In one such approach, Mikos and coworkers used sodium chloride particles as a fugitive phase for the fabrication of porous poly(L-lactic acid) (PLLA) foams.^[4] In this particulate leaching approach, a composite of the polymer with the salt particles (porogen) is first prepared by dispersion of the salt particles in a polymer solution. Following drying of this composite, the salt porogen is leached out by dissolution in water. While this process yields foams with ρ_{rel} in the range of 0.08–0.15, the pore morphology is at best partially open, with cubic pores surrounded by wall leaflets. This pore morphology is a direct consequence of use of salt as porogen. Spherical pores with improved interconnectivity and open-pore architecture can be obtained by the substitution of gelatin spheres for the salt porogen. Significant limitations of this approach include: 1) the need for physical contact between the porogen particles to ensure efficient extraction of the porogen, which narrows the porosity range considerably; 2) high crystallinity in the polymer phase to ensure the structural integrity of the polymer-salt composite; and 3) small foam thickness typically 2–5 mm due to diffusional constraints and issues related to phase separation of the polymer and porogen. Thicker foams can be prepared by a variation of this approach where the polymer and salt particles are first sintered around the glass transition temperature of the polymer followed by annealing prior to leaching.

Nevertheless, the salt leaching technique remains an important development in the fabrication of biodegradable polymer foam scaffolds for TE and remains a popular method due to its simplicity. Recently, Chen et al. have utilized ice particles to achieve similar results.^[5]

Pore formation using a non-water-soluble fugitive phase

Recently, Shastri and coworkers reported the fabrication of polymeric cellular solids by using hydrocarbon particles (porogen) as a template.^[6] This approach differs from the fugitive phase removal process as applied to the fabrication of porous ceramics and metals in the following manner:

1. The fugitive phase is removed by extraction in a solvent that is a nonsolvent for the polymer at moderate temperatures ($<100^{\circ}\text{C}$). This ensures rapid solidification of the polymer around the hydrocarbon template particles due to precipitation of the polymer phase.
2. The fugitive phase is introduced into the polymer phase by simple dispersion of the porogen in a high

viscosity solution of the polymer, thus eliminating the need for sintering.

3. Because the polymer phase is a viscous solution, blending and dispersion of noncompatible inorganic and organic phases can be easily achieved without affecting overall macroporosity.

With this approach, polymer foams with porosity as high as 90% and compressive moduli of over 5 MPa have been fabricated using a variety of degradable and nondegradable polymers. Furthermore, because the polymer phase can be modified by an additive approach, and the processing conditions are anhydrous, polymer foams containing water-soluble and water-insoluble inorganic salts, polymers, and colloids can be fabricated with no significant loss in macroporosity. Since the hydrocarbon porogen is actively extracted by dissolution, by the appropriate choice of the hydrocarbon porogen, polymeric porous solids possessing either a continuous open-pore or an isolated closed-pore architecture may be prepared (Fig. 1). The former is ideally suited for TE applications and has been extensively explored in this context.^[7,8] Yet another interesting structural characteristic of foams produced by the hydrocarbon templating process is the presence of microporosity within the polymer phase, which is a direct consequence of both the nucleation and growth of solvent bubbles and the phase separation of the polymer during the extraction phase (Fig. 2). Because a viscous polymer solution is used to disperse the porogen particles, the success of the hydrocarbon templating approach is dependent on finding a solvent for the porogen that is a nonsolvent for the polymer and miscible with the theta (ideal) solvent for the polymer phase. For degradable polyesters and other common nondegradable polymers such as polystyrene, polycarbonate, and polymethylmethacrylate, low-boiling alkanes such as pentane and hexane satisfy the requirements for a nonsolvent as they are miscible with chlorinated hydrocarbons, which are typical solvents for these polymers. Miscibility with the polymer solvent is key in ensuring a rapid precipitation of the polymer phase around the porogen.

In a variation of the hydrocarbon templating approach, Ma and coworkers have used an assembly of spherical paraffin particles, annealed to form a continuous bonded phase as a template, to produce porous polymeric foams with spherical-interconnected pore architecture.^[9] Using this method they have produced foams of PLLA and PLGA with a pore distribution in the range of 100–500 microns. However, the spherical geometry of the porogen poses two constraints: 1) the porosity cannot be manipulated independent of the pore size; and 2) all the porogen particles have to be in intimate contact with each other to ensure integrity of the annealed paraffin template.

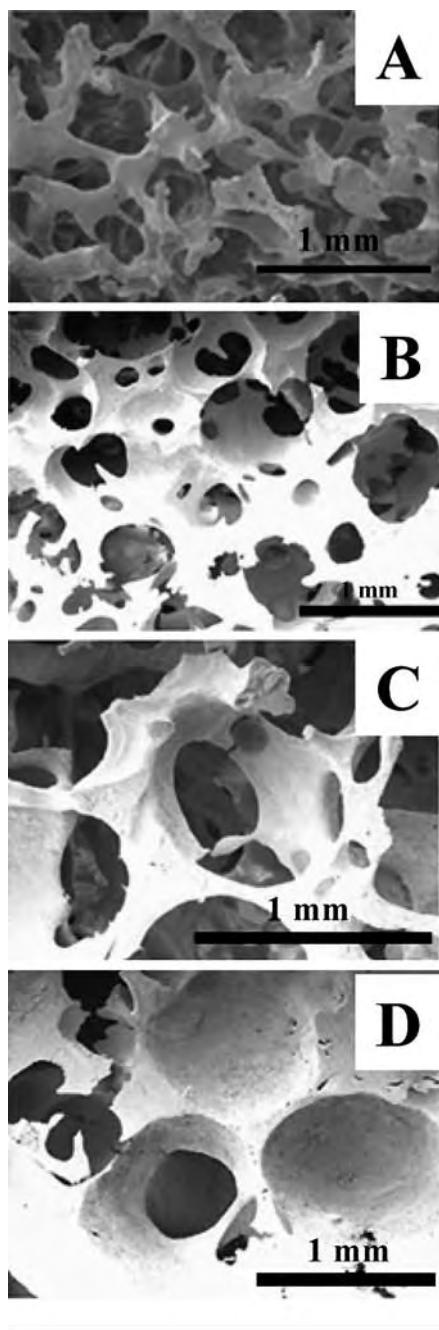


Fig. 1 Scanning electron micrographs of PLA and PLGA foam produced by the hydrocarbon templating process. Note the presence of a continuous polymer phase with open-pore architecture (A), spherical open-pore architecture (B and C), and isolate closed-pore architecture (D). The pore shape and interconnectivity are a function of hydrocarbon template morphology and concentration, respectively, in the polymer phase.

Phase Separation

Phase separation remains an easy and simple way of introducing porosity into a polymer phase. However,

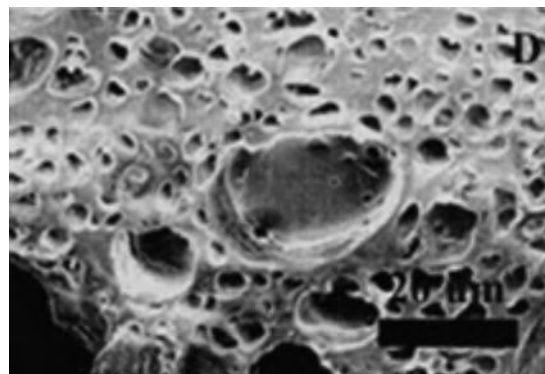


Fig. 2 Scanning electron micrograph of the cross-section of the polymer phase in PLGA foam produced by the hydrocarbon templating process, revealing the presence of a well-defined microporosity. The source of the micropores is most likely a combination of nucleation and growth of solvent bubbles and polymer phase separation.

controlling the segregation of the polymer and solvent phase is complex and typically results in narrow control over porosity and pore characteristics. The key to this process is finding a polymer-solvent combination that can yield a high-phase-segregated structure under conditions in which the solvent can be rapidly either evaporated or sublimed. Foams possessing microcellular architecture can be readily prepared by the phase separation approach.^[10] Some commonly used polymer solvent combinations are naphthalene:PLLA,^[11] supercritical carbon dioxide:PLGA,^[12] and dioxane:PLGA.^[13] Pore shapes can range from circular to elongated^[14] with high aspect ratios depending on the solvent:polymer combination. Due to the relatively mild processing conditions, proteins, peptides, and growth factors can be introduced directly into the polymer phase during the foaming process, thus enabling scaffold-based delivery of bioactives.^[11,12] Scaffold-based delivery of growth factors is an important aspect of TE.

Sintering

The process of sintering entails the thermally induced bonding of particles under pressure. Sintering is a very common approach for the production of porous ceramics. Because sintering requires intimate contact between particles, void volume is severely diminished, and therefore sintering yields articles that are quite dense, with porosity rarely exceeding 50%. Due to the low porosity, extensive pore interconnectivity is absent in sintered articles, making them less attractive for TE applications. In order to increase the porosity in a sintered article, one can combine a thermally stable, water-soluble fugitive phase (spacing agent) such as salt or gelatin with the polymer phase during the

sintering, thus creating an avenue for the introduction of additional porosity upon a removal of this phase.^[15] Porous hydroxyapatite cylinders are processed using this approach.

Textile Processing

Nonwoven fibrous meshes of poly(glycolic)acid (PGA) and PLLA having porosity >90% have been produced using standard textile processing techniques.^[16] Although polymer meshes and felts do not fit the definition of a polymer foam or cellular solid in that they lack a cellular structure, they do share some functional characteristics such as low thermal conductivity and acoustic dampening. Nonwoven PGA meshes have been extensively explored in TE applications with a high degree of success.^[17] Because meshes constitute a random arrangement of discrete polymer fibers, they are structurally weak, i.e., easily compressed. However, their structural stability can be vastly improved by bonding the fibers with a secondary polymer binder.^[18] Such reinforced structures are structurally closer to polymer foams as the bonding process results in the formation of struts (fibers) and vertices (bonded fiber clusters).

ARTICLES OF FURTHER INTEREST

Microporous Materials; Supercritical Fluid Processing; Tissue Engineering Scaffolds

REFERENCES

- Gibson, L.J.; Ashby, M.J. *Cellular Solids: Structure and Properties*, 2nd Ed.; Pergamon: Oxford, United Kingdom, 1997.
- Suh, K.W.; Skochdopole, R.E. *Encyclopedia of Chemical Technology*, 3rd Ed.; Kirk-Othmer, Ed.; John Wiley & Sons, 1980; Vol 2, 82.
- Mooney, D.J.; Baldwin, D.F.; Suh, N.P.; Vacanti, J.P.; Langer, R. Novel approach to fabricate porous sponges of poly(DL-lactic-co-glycolic acid) without the use of organic solvents. *Biomaterials* **1996**, *17* (14), 1417–1422.
- Mikos, A.G.; Thorsen, A.J.; Czerwonka, L.A.; Bao, Y.; Langer, R. Preparation and characterization of poly(L-lactic acid) foams. *Polymer* **1994**, *35*, 1068–1077.
- Chen, G.; Ushida, T.; Tateishi, T. Preparation of poly(L-lactic acid) and poly(DL-lactic-co-glycolic acid) foams by use of ice microparticles. *Biomaterials* **2001**, *22* (18), 2563–2567.
- Shastri, V.P.; Martin, I.; Langer, R. Macroporous polymeric foams by hydrocarbon templating. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97* (5), 1970–1975.
- Schaefer, D.; Martin, I.; Shastri, V.P.; Padera, R.F.; Langer, R.; Freed, L. E.; Vunjak-Novakovic, G. In vitro engineering of bone-cartilage composite tissues. *Biomaterials* **2000**, *21*, 2599–2606.
- Martin, I.; Shastri, V.P.; Padera, R.F.; Langer, R.; Yang, J.; MacKay, A.; Vunjak-Novakovic, G.; Freed, L.E. Selective in vitro differentiation of mammalian mesenchymal progenitor cells into three dimensional skeletal tissues. *J. Biomed. Mat. Res.* **2001**, *55* (2001), 229–235.
- Ma, P.X.; Choi, J.W. Biodegradable polymer scaffolds with well-defined interconnected spherical pore network. *Tissue Eng.* **2001**, *7* (1), 23–33.
- Zhang, R.; Ma, P.X. Poly(alpha-hydroxyl acids)/hydroxyapatite porous composites for bone-tissue engineering I. Preparation and morphology. *J. Biomed. Mat. Res.* **1999**, *44* (4), 446–455.
- Lo, H.; Kadiyala, S.; Guggino, S.F.; Leong, K.W. Poly(L-lactic acid) foams with cell-seeding and controlled release capacity. *J. Biomed. Mat. Res.* **1996**, *30* (4), 475–484.
- Hile, D.; Amirpour, M.L.; Akgerman, A.; Pishko, M.V. Active growth factor delivery from poly(DL-lactide-co-glycolide) foams prepared in super critical CO₂. *J. Control. Release* **2000**, *66* (2–3), 177–185.
- Schugens, C.; Maquet, V.; Grandfils, C.; Jerome, R.; Teysse, P. Polylactide macroporous biodegradable implants for cell transplantation. II. Preparation of polylactide foams by liquid-liquid phase separation. *J. Biomed. Mater. Res.* **1996**, *30* (4), 449–461.
- Teng, Y.D.; Lavik, E.B.; Qu, X.; Park, K.I.; Ourednik, J.; Zurakowski, D.; Langer, R.; Snyder, E.Y. Functional recovery following traumatic spinal cord injury mediated by a unique polymer scaffold seeded with neural stem cells. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99* (5), 3024–3029.
- Thomson, R.C.; Yaszemski, M.J.; Powers, J.M.; Mikos, A.G. Fabrication of biodegradable polymer scaffolds to engineer trabecular bone. *J. Biomater. Sci., Polym. Ed.* **1995**, *7* (1), 23–38.
- Freed, L.E.; Vunjak-Novakovic, G.; Biron, R.J.; Eagles, D.B.; Lesnoy, D.C.; Barlow, S.K.; Langer, R. Biodegradable polymer scaffolds for tissue engineering. *Biotechnology (NY)* **1994**, *2* (7), 689–693.
- Freed, L.E.; Marquis, J.C.; Nohria, A.; Emmanuel, J.; Mikos, A.G.; Langer, R. Neocartilage formation in vitro and in vivo using cells cultured on synthetic biodegradable polymers. *J. Biomed. Mat. Res.* **1993**, *27* (1), 11–23.
- Mikos, A.G.; Bao, Y.; Cima, L.G.; Ingber, D.E.; Vacanti, J.P.; Langer, R. Preparation of poly(glycolic acid) bonded fiber structures for cell attachment and transplantation. *J. Biomed. Mat. Res.* **1993**, *27* (2), 183–189.

Polymers

Gary E. Wnek

Department of Chemical Engineering, Case Western Reserve, Cleveland, Ohio, U.S.A.

INTRODUCTION

Eras of civilization have frequently been named for materials discovered and subsequently used extensively by humans (e.g., the Stone Age, the Bronze Age, etc.). Toward that end, the 20th century might appropriately be labeled as the Plastics Age or, somewhat more broadly, the Polymer Age. Whereas biological polymers (e.g., DNA, proteins) have been present since the origins of life, the development of synthetic polymers and the realization that long-chain molecules could exist are relatively recent milestones on the materials calendar. Indeed, it was only in the 1920s that the concept of a polymer or macromolecule came to be grudgingly accepted, slowly displacing the colloid theory that suggested that macromolecules were nothing more than aggregates of small molecules.^[1] The subsequent two decades witnessed the rapid development of many types of synthetic fibers, plastics, and elastomers, including Dacron, nylon, and Spandex. Also, guiding principles for the successful linking of monomers into high-molecular-weight polymers were established. The remainder of the 20th century saw explosive growth in new polymers and their applications. Today, polymers are ubiquitous, from commodity plastics such as plumbing pipe, food wrap, and trash bags, to highly specialized and tailored materials for use in aircraft components, personal computers, and health care.^[2]

A key reason for the broad acceptance and appeal of polymers is that they are the most versatile of all materials, spanning the range of mechanical properties from soft gels and rubbers to extremely strong fibers. This versatility is a direct consequence of the great latitude conferred by design of the repeat-unit structure and control of chain length (or molecular weight) and chain architecture. While most polymers are based on carbon backbones, inorganic backbones (e.g., silicones) add yet another dimension to the tailoring of material properties. A key aim of this article is therefore to introduce basic ideas of how polymer structure affects properties. A short discussion of the implications of these ideas in the design of polymeric biomaterials will follow. Several examples of such materials are discussed in some detail in additional

entries of the encyclopedia. This article concludes with a summary of opportunities and challenges in polymer science for the next two decades and beyond.

INTRODUCTION TO STRUCTURE/PROPERTY RELATIONSHIPS

Small Organic Molecules

The properties of small organic molecules (melting point, for example) depend intimately on the structure of the molecule and principally on four factors: 1) size; 2) symmetry; 3) internal flexibility (presence or lack of bond rotations); and 4) strength of intermolecular forces. A consideration of how these dictate properties of small molecules is instructive in understanding structure/property relations in polymers. Melting (or fusion) is a first-order thermodynamic transition, and the melting temperature, T_m , depends on the ratio of the enthalpy and entropy of fusion:

$$T_m = \Delta H_m / \Delta S_m \quad (1)$$

Melting points typically increase with molecular size, as evidenced by tracking the properties (gas \rightarrow liquid \rightarrow grease \rightarrow wax \rightarrow resin) of n-alkanes as a function of carbon number. Highly flexible molecules that are reasonably symmetric crystallize readily because the conformation(s) required in the crystal (for example, the extended zig-zag conformation for n-alkanes) are easily adopted. Such materials have rather low melting points even at high molecular weights, since considerable entropy (translational and rotational) is gained on melting and thus ΔS_m is large. Branching can lower melting points by reducing the ability of a molecule to pack tightly into a crystal, thus lowering the enthalpy of fusion. Moreover, branching introduces asymmetry that can inhibit crystallization when the material is rapidly cooled, the result being an amorphous solid or glass. In cases where branching is extensive and does not lead to lack of symmetry, melting points can actually increase vs. linear counterparts due to loss of conformational flexibility and hence lower entropies of fusion. Rigid molecules such as anthracene tend to



have high melting points due to a low entropy of fusion, as only translational entropy is primarily gained on melting. Similarly, molecules with strong intermolecular forces (e.g., hydrogen bonds) have high melting points, usually not due to the need to overcome these forces on melting but rather as the result of their persistence in the molten state affording low entropies of fusion. For this reason, benzoic acid has a relatively high melting point for its molecular weight. These same four factors applied to polymer repeat units, along with chain orientation and extent of crosslinking, are the primary determinants of polymer properties. These points will be illustrated with polyethylene and vinyl polymers derived from mono-substituted ethylenes as examples, followed by a brief discussion of polyesters.

Glass Transition Temperature

A glass is appropriately defined as material having the structural properties of a liquid, namely disorder, and the mechanical properties of a solid; it has a glass transition temperature, T_g , above which the material can flow. The idea of a melting point is easily appreciated but T_g is perhaps less familiar and thus deserves a brief discussion here. Cooling of a molten substance can lead to crystallization at T_m and is accompanied by a rapid drop in volume as shown in Fig. 1 (except for odd substances such as water). The crystalline substance will continue to drop in volume, albeit slowly, as

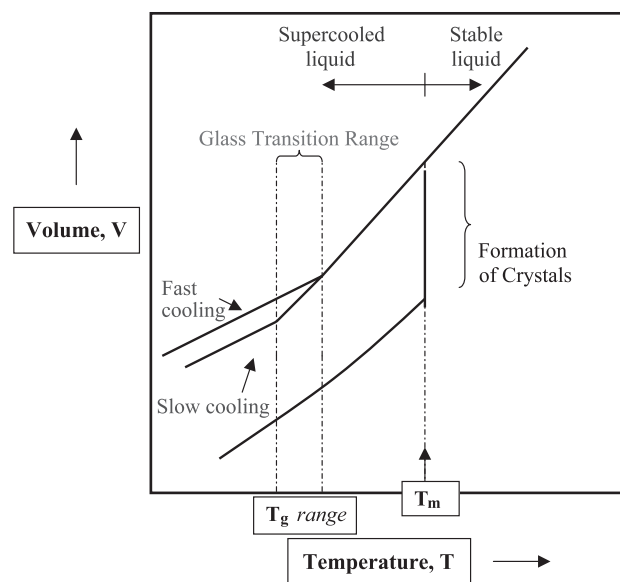


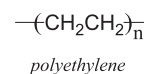
Fig. 1 Plot of specific volume vs. temperature, contrasting the formation of a glass vs. crystals from the melt.

determined by its coefficient of expansion. However, under certain circumstances it is possible to bypass T_m and form a supercooled liquid that eventually forms a glass. This process, known as vitrification, can occur if cooling is rapid, the viscosity of the melt is very high, and/or the molecular structure of the substance is either asymmetric or large and difficult to pack into a crystal. At T_g , molecular motion effectively ceases. However, unlike T_m , the value of T_g depends upon the cooling rate, with rapid cooling “freezing-in” higher energy molecular conformations and organizations. Therefore, it is common to see T_g reported as ranges, and it is important to specify cooling or heating rates in measurements of T_g .

Polymers^[1,2]

Polyethylene and polymers from mono-substituted ethylenes

Consider first molecular weight, using the n-alkane series as an example with high molecular weight and linear or high-density polyethylene (HDPE) as the upper limit.



Several important points derive from inspection of samples of polyethylene and some lower-molecular-weight solid n-alkanes such as paraffin wax (roughly 30–40 carbons). The wax, a highly crystalline substance, is easily broken into small pieces with a fingernail, whereas the polymer resists such damage and is seen to be a very tough flexible material. Molten wax has a low viscosity in contrast to polyethylene. More significantly, and unlike the wax, fibers can be drawn from the molten polyethylene very easily. The reason is that the molecules of polyethylene are long enough to physically entangle with one another, thus supporting the ability to form fibers from the melt. The entanglement molecular weight is roughly the length of a molecule at which the properties of polymers begin to manifest themselves. It is dependent upon the chemical structure of the polymer, but most entanglement molecular weights are in the range of 2000 to 20,000 g/mol.

As expected, the viscosity of a molten polymer will rise dramatically with molecular weight since chain entanglements will inhibit the ability of molecules to move past one another. Many other properties such as melting point and tensile strength do not continue to increase with molecular weight but reach an asymptote. (However, the tensile strength can be increased substantially by chain orientation.) This is because, at

long chain lengths, polymer chains behave as a collection of shorter independent segments. In the case of polyethylene, the asymptotic limit for T_m is about 140°C . This relatively low temperature is the result of high chain flexibility and thus a high ΔS_m .

Polymers such as linear polyethylene crystallize easily due to the highly symmetric repeat unit and high chain flexibility. However, polymers rarely crystallize completely. Crystallization can begin at multiple points along a long polymer chain, eventually trapping segments of disordered or amorphous chains between crystallites. Therefore, materials such as HDPE are semicrystalline (Fig. 2), with a T_m for the crystalline component and a T_g for the amorphous regions, with the degree of crystallinity being about 70–80%. The T_g for polyethylene is approximately -50 to -90°C , and hence this polymer at room temperature is a composite of crystals and liquid-like amorphous regions. The toughness of polyethylene at room temperature is due to the ability of the amorphous regions above their T_g to absorb mechanical impact. Semicrystalline polymers are typically translucent due to light-scattering by the crystallites. Their solubilities can effectively be nil until temperatures rise to close to T_m , and this is why HDPE containers can be used to store gasoline even though a mixture of alkanes might be expected to dissolve the polymer readily at room temperature because of similar molecular structures.

Mechanical extension of a semi-crystalline polymer typically leads to “necking-down” in the drawn region as chains are pulled out of crystallites and oriented in the stretch direction. Extensive orientation in gel-spun, ultrahigh-molecular-weight HDPE affords fibers the tensile strengths of which exceed that of steel on a weight basis.

Molecular symmetry can affect both the melting point and the degree of crystallinity. Low-density polyethylene (LDPE) has occasional short alkyl branches

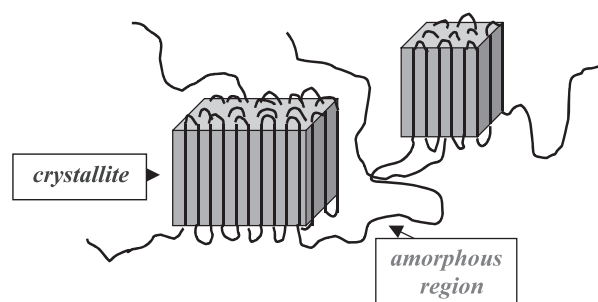
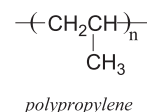


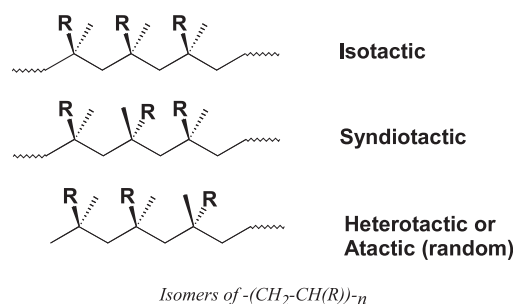
Fig. 2 Morphology of a typical semicrystalline polymer such as linear polyethylene. It is kinetically easier for the chains to fold on themselves during crystallization rather than form long extended structures.

that serve to disrupt chain packing and thus lower the degree of crystallinity to about 50% and the T_m to 105 – 110°C vs. about 140°C for the linear material. LDPE is softer (and tougher) than HDPE due to its lower crystallinity.

The effect of symmetry is particularly pronounced for polymers derived from mono-substituted ethylenes such as polypropylene.



This repeat unit as drawn is misleading in that there are two possible configurations at the methine carbon ($-\text{CH}-$) and the arrangement of these along a chain leads to three possible isomers (here $R = \text{methyl}$):



After polyethylene, polypropylene is the second most important polymer in terms of annual world production, and it is only the isotactic form of polypropylene that is produced. This material has a higher T_m (165 – 170°C) compared with HDPE, the principal reason being that the methyl groups reduce the conformational flexibility of the chain in the melt, affording a lower ΔS_m . Its T_g is about -10°C due again to restricted bond rotations, and hence, compared to polyethylene, it is not flexible at lower temperatures. Interestingly, the laboratory that discovered catalysts to produce HDPE many decades ago also produced isotactic polypropylene, but the material was ignored since it was not thought that a backbone with pendant methyl groups would have a higher melting point than polyethylene. As pointed out earlier in our discussion about small molecules, branching can lower T_m but that is principally if there is no regular placement of branches and no stereochemical control at the branch points. (This assumes that there are no strong intermolecular forces acting to depress the entropy of the melt; see the discussion on polylactide and polyglycolide.) Isotactic polypropylene has a very regular structure and its conformational flexibility dictates its thermal properties. The higher T_m makes this material useful in a variety of medical applications (e.g., plastic syringes and tubing) because it can be sterilized in an autoclave without coming too close to its T_m .

It is interesting to compare the structures of linear polyethylene and isotactic polypropylene in the crystalline regions. In polyethylene, the chains adopt a zig-zag conformation as is the case in crystals of small-molecule n-alkanes. In polypropylene, by contrast, the pendant methyl groups force the chains to be helical. An analogy exists with the secondary structure of polypeptides; for example, poly(glycine) prefers the β -sheet structure whereas polypeptides with bulky side chains adopt an α -helix.

Polystyrene is another important material derived from a mono-substituted ethylene where R is a benzene ring. Commercial polystyrene is exclusively the atactic form and, because this isomer is irregular, the polymer does not crystallize and therefore is transparent. This is an attractive property and accounts for the widespread use of polystyrene in petri dishes and related lab supplies. Its T_g is about 100°C due to restricted chain flexibility by virtue of the rather bulky phenyl substituents, and hence the material is rather brittle at room temperature. Like T_m , the T_g of a polymer reaches an asymptote with molecular weight, the 100°C value is the limiting T_g for atactic polystyrene.

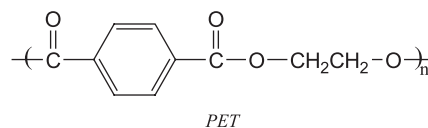
Poly(vinyl chloride) or PVC, where $R = \text{Cl}$, is another important atactic polymer. Besides its widespread use in plumbing pipe and connectors, it is a familiar lab staple in the form of flexible Tygon tubing. Tygon is PVC containing a plasticizer such as 2-diethylhexyl phthalate (DEHP). PVC has a T_g of about 85°C , significantly higher than that of polypropylene because of the slightly larger size of the pendant Cl but more significantly due to stronger intermolecular forces stemming from the polar C-Cl bond. At T_g , intermolecular forces need to be overcome to allow chain segments to move relative to one another. Addition of a plasticizer, which is typically a low-volatility liquid with good solubility in the polymer, serves to disrupt intermolecular forces and hence depress the T_g . In Tygon, the T_g is below room temperature. It is interesting that the material keeps its physical shape rather than forming a puddle of high-viscosity polymer plus plasticizer, and the explanation is that PVC contains a small amount of tiny crystallites (perhaps from short segments of syndiotactic repeat units) that act to hold the material together.

Another means of lowering T_g and thus increasing material flexibility is to copolymerize vinyl chloride with monomers such as vinyl acetate. At least some readers will recall long-playing phonograph records which were frequently made from such copolymers.

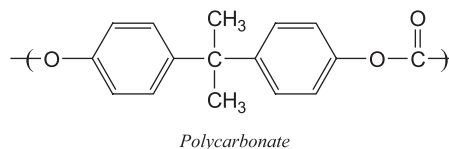
Polyesters

Polyesters represent an important class of materials with broad applications as biomaterials. Poly(ethylene

terephthalate), PET, is a ubiquitous material known as Mylar in film form and Dacron in fiber form. Woven Dacron mesh is used for large-diameter vascular grafts. PET, with its rather regular structure, is a semi-crystalline polymer with a T_m of about 265°C and a T_g near 85°C .

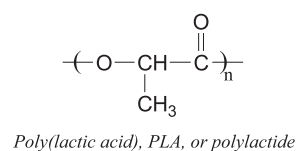
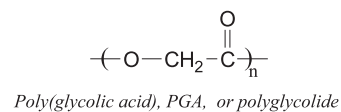


Polycarbonates are formally polyesters of carbonic acid, the most popular being the material with the trade names Lexan or Merlon. The repeat unit, which is symmetric, is more complex and less flexible than that of PET, and thus polycarbonate will typically form a glass (T_g ca. 150°C) on cooling; hence the polymer is transparent.



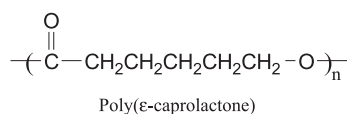
The polymer is rigid at room temperature, and it is tempting to think that all glassy polymers below their T_g might be brittle. Polycarbonate is remarkably resistant to impact vs., for example, glassy polystyrene, as evidenced by the properties of compact discs (polycarbonate) and their holders (polystyrene). Polycarbonate has sub- T_g short-range molecular motions associated with the flexible carbonate group that serve to absorb significant impact energy. Such energy-absorbing mechanisms are not available in polystyrene. Similar sub- T_g motions account for the flexibility of PET fibers and films at room temperature.

Both PET and polycarbonate hydrolyze very slowly and are not useful as biodegradable materials. However, simple aliphatic polyesters such as poly(glycolic acid) and poly(lactic acid) are well-known biodegradable polyesters. (They are sometimes referred to as polyglycolide and polylactide, respectively, because they are frequently prepared from dimers of the acids.)



PGA is semicrystalline (typically ca. 50% crystallinity) with a T_m of 225–230°C and a T_g of about 35–40°C. The small $-\text{CH}_2-$ group allows close packing in the crystal and strong dipole interactions among esters, and only rather unusual solvents such as fluorinated alcohols can dissolve the polymer. PLA, like polypropylene, has two possible configurations at the methine carbon. Since lactic acid is made naturally as the L-isomer, the pure (isotactic) polymer derived from it is called poly(L-lactic acid) or PLLA. This material is also semicrystalline (35–40%), but with a lower T_m (ca. 175°C) than PGA, presumably due to methyl groups acting to partially shield dipole-dipole interactions. The T_g of PLLA is around 60°C v 40°C for PGA, consistent with the idea that the methyl group increases rotational barriers along backbone bonds. Copolymers of the D and L isomers of lactic acid are readily made and these are amorphous as expected from their atactic structure. Also, copolymers of lactic and glycolic acids (PLGA) are of great interest because degradation rates can be controlled for a variety of applications in tissue engineering and drug delivery.

Dilution of the polar ester group by methylene units reduces intermolecular forces and thus lowers both T_g and T_m compared with PGA. An example is poly(ϵ -caprolactone), a semicrystalline polymer that has a T_m of 60°C and a T_g of about -60°C .



Cross-linking and Orientation

The molecular origins of structure/property relationships discussed above apply to both small molecules and polymers. However, two additional parameters affect polymers in unique ways. The first is cross-linking, which involves primary bonds (chemical cross-linking) or secondary bonds (physical cross-linking) and affords what are termed polymer networks. Light cross-linking of low- T_g polymers gives materials known as elastomers, and these have the typical properties of a rubber band. Cross-linking raises T_g , as chain motions are now restricted by network junction points. Heavy cross-linking can give very hard, brittle materials, as T_g will rise above room temperature and molecular motions will be too restricted to contribute significantly to impact energy absorption (e.g., toughness). A biopolymer example is keratin, the major constituent of hair and fingernails. Keratin in hair is lightly cross-linked via disulfides from coupling of neighboring cystine amino acids, and the ability to introduce a “permanent wave” takes advantage of the reversibility of disulfide

formation and cleavage via redox chemistry. Fingernails have a much higher disulfide cross-link density, and hence they are harder. (Tortoise shells are even more highly cross-linked.)

Polymer networks can be prepared by covalently linking chains into what is effectively a single molecule using high-energy radiation such as gamma rays, light, or a wide variety of chemical reactions (e.g., disulfide formation). It is also possible to exploit strong intermolecular interactions to afford materials with the properties of a polymer network, yet retaining the ability to reprocess the material upon heating. A classic example is the segmented polyurethanes, which contain a low- T_g block and are cross-linked via thermally labile hydrogen-bonded domains. These have the “snap” of a rubber band but are reprocessable, and are termed thermoplastic elastomers.

New Synthesis Strategies

Polymer chemists continue to explore and optimize polymerization reactions as well as discover new reaction schemes and catalysts that afford a tight molecular weight distribution (MWD), afford more precise control over backbone composition and repeat unit stereochemistry, and allow for the creation of novel chain architectures.

Particularly interesting advances have been made in the area of “living” polymerization, a term for a system where propagating chain ends remain active throughout the duration of the reaction. For example, it has long been thought that free radical polymerization would not be amenable to the characteristics of a living polymerization (e.g., a linear dependence of molecular weight with monomer conversion) due to the inherent high reactivity of radicals. However, it has been discovered over the last decade that it is indeed possible to achieve living polymerization character provided that the number of active radical species is very low at any given time. This can be done by reversibly capping active chain ends using stable nitroxide radicals, halides via redox chemistry with selected transition metals (called atom-transfer radical polymerization, or ATRP), or reversible addition-fragmentation polymerization (RAFT).^[3] On another front, it was believed about 15 years ago that polyethylene and polypropylene were mature technologies, although that has changed dramatically with the advent of new transition metal catalysts^[4] that, for example, afford fine control of the extent of branching (and hence the degree of crystallinity) of polymers of ethylene and small amounts of α -olefins. The new catalysts are also producing old polymers with a high control of stereospecificity. One recent example is syndiotactic polystyrene.



Finally, polymerization reactions are being developed that yield polymers having interesting chain architectures, such as block and graft structures. Chain architecture is important in that compositionally different regions of a polymer chain can contribute different properties. For example, styrene-butadiene-styrene triblock polymer will self-assemble upon cooling from the melt into various thermodynamically stable morphologies depending on the relative lengths of, and thus volume fractions of, the blocks. At about 20 vol% polystyrene, the material is composed of a low- T_g polybutadiene continuous phase and small (ca. 20-nm diameter) spherical domains of atactic polystyrene (T_g of 100°C). The domains are interconnected by virtue of the block polymer chain structure, and hence this material behaves as an elastomer with cross-linking via the glassy polystyrene domains. Such a material is a thermoplastic elastomer, because it can flow if heated above the T_g of polystyrene.

Whereas block and graft copolymers have been known for some time, new polymerization methods such as ATRP are expanding the number of examples of these systems.

More complex architectures, such as multiarm stars and dendrimers,^[5] are being targeted as polymerization chemistry becomes more versatile. Star polymers (multiarm structures emanating from a central core) are of interest principally for the large numbers of chain ends. For example, a 4-arm star polymer will typically have a lower T_g and a lower viscosity than a linear polymer of a similar total molecular weight. Dendrimers are a special class of star-like polymer that have regularly placed branched repeat units. Construction of such a polymer can begin with a core to which a monomer capable of branching is attached (divergent approach), or conversely by making highly branched arms and then connecting these to a core as a final step (convergent approach). Dendrimers typically have poor physical properties in the solid state due to the absence of chain entanglements. However, their interiors have open volume in which to bind drugs and an exterior with a very high concentration of functional groups due to the high degree of branching. These characteristics make dendrimers of great interest in medical applications, for example, as vehicles for delivery of genes to cells.

CHALLENGES AND OPPORTUNITIES: THE NEXT 20 YEARS

Polymer science and engineering will continue to be a key discipline for the design and development of new classes of materials with new or improved applications in medicine. One snapshot of where the field might be heading is derived from a recent report focusing on

similar challenges and opportunities for the Department of Defense,^[6] and a section of that report written by the author is amended here as it might apply to biomaterials development.

Modeling Will Be a Routine First Step

The creation of new polymers will begin with a broad evaluation of properties using high-level modeling and simulation to determine critical parameters (isomeric structure, molecular weight, degree of chain orientation) that influence a property of interest. In particular, modeling will be employed to predict complex organization of functional low-molar-mass molecules and polymers, as is beginning to be done for the difficult problem of predicting protein folding motifs. Modeling will also extend to synthesis routes to define the best approach, as well as to processing. Much guiding information will be in hand prior to doing “wet chemistry.”

Synthesis and Processing Will Become Seamless

Synthesis and processing are typically separate operations. This will change over the next two decades as opportunities to carry out simultaneous synthesis and processing emerge. This idea is not new (e.g., reaction injection molding or chemical vapor deposition), but it will become more widespread. Of particular interest will be polymers and small molecules that self-assemble into ordered molecular structures (e.g., liquid crystals) or morphological structures (e.g., block copolymers). Combinatorial synthesis of polymers will become more routine, and the emergence of combinatorial processing approaches to rapidly identify conditions for fabricating polymers to achieve maximum properties is anticipated.

Polymer Synthesis Will Boast Sequence Control

Nature has evolved machinery to precisely control the sequence of amino acid additions and hence the primary structure of polypeptides. However, conventional polymerization techniques have limited opportunities for the control the sequence of addition of two or more monomers. Block copolymers are possible due to successive addition of monomer charges to active chain ends, and alternating copolymers can be obtained under special circumstances. However, there currently is no viable means to prepare, for example, vinyl-type copolymers with sequence control (e.g., poly [(monomerA)₁-(monomerB)₂]_n). This is in stark contrast to peptide synthesis on ribosomes within

biological cells, which employ a template to code for specific amino acids that are enzymatically linked. A major opportunity and challenge presents itself for the sequence-controlled polymerization of a wide variety of monomers using systems that mimic the functions of ribosomes. Template-directed polymerizations are energetically uphill due to the decrease in entropy resulting from monomer ordering, and in living systems the requisite energy is supplied by ATP. Electrochemical polymer synthesis is an attractive option in that properly patterned electrodes may simultaneously serve as solid templates for synthesis and as energy sources. Other lessons from biology, such as self-assembly and development of hierarchical structures, will continue to be borrowed and built upon.

The Definition of Structure in Polymers Will Rival That of Low Molar Mass Materials

Polymers are by their very nature complex materials, typically having a distribution of chain lengths, isomer content, degrees of orientation, and fractional crystallinity (if any). Thus, structure in polymeric materials is frequently ill-defined compared with small molecules. This difference will gradually disappear over the next two decades as long-chain molecules are synthesized with greater compositional and structural precision. The implications will be significant in that very precise structure/property relationships will be possible, and maximization of properties through the integration of modeling, synthesis, and processing will be realized.

The Impact on Nanotechnology

Many common examples exist of nanoscale (one dimension of <100 nm) materials, including block copolymer films and collagen fibers that function as scaffolds for tissues and organs. However, there will be an increasing push to exploit the properties of individual molecules, or very small aggregates of molecules, for the next generation of multifunctional biomaterials. For example, carbon nanotubes and collections of only a few organic molecules are being studied as components of diodes, transistors, and memory elements. The ability to create well-defined organic polymer structures and manipulate these to form complex and functional arrangements will drive a revolution in information storage and processing, sensing, and communications. This effort is predicted to impact biomaterials and biomedical devices in many

ways, including the design of surfaces with specific and predictable protein adsorption characteristics, and stable and reliable interfaces between cells and electronic materials, both inorganic and organic.

Hybrid Materials Systems Will Be Ubiquitous

Opportunities will continue to arise for the use of polymers as components in hybrid materials systems along with metals, ceramics, and/or electronic materials. Familiar composite materials (e.g., graphite-reinforced epoxy) will see increased uses, but significant potential is seen for new materials combinations such as organic electroactive materials and silicon for hybrid electronic and optical devices in, for example, biosensors.

SUMMARY

Polymers represent the most versatile class of materials because of their adaptability to increasingly precise control of structure and properties. Thus, it is anticipated that polymers will continue to be exploited in many and, in some cases, highly specialized biomaterials applications.

ARTICLES OF FURTHER INTEREST

Biofunctional Polymers; Elastomers, Biodegradable; Hydrogels; Poly(Glycolic Acid); Poly(lactic acid)s; Polyurethane Biomaterials; Silicones; Ultra-High Molecular Weight Polyethylene (UHMWPE)

REFERENCES

1. Morawetz, H. *Polymers: The Origin and Growth of a Science*; Wiley-Interscience: New York, 1985.
2. Allcock, H.R.; Lampe, F.; Mark, J. *Contemporary Polymer Chemistry*, 3rd Ed.; Prentice Hall: Englewood Cliffs, NJ, 2003.
3. Qui, J.; Charleux, B.; Matyjaszewski, K. *Prog. Polym. Sci.* **2001**, *26*, 2083.
4. *Organometallic Catalysts and Olefin Polymerization*; Blom, R., Follestad, A., Rytter, E., Tilset, M., Ystenes, M., Eds.; Springer-Verlag: Berlin, 2001.
5. Roovers, J.; Comanita, B. *Adv. Polym. Sci.* **1999**, *142*, 180.
6. *Materials Research to Meet 21st-Century Defense Needs*; National Academies Press: Washington, DC, 2003.

Polymers Used in Tissue Engineering

Vasif Hasirci
Deniz Yucel

Department of Biological Sciences, Middle East Technical University, Ankara, Turkey

INTRODUCTION

Tissue engineering is an extension of the biomedical materials field and involves scaffolds (preferably biodegradable and polymeric) seeded with autogeneous cells. Polymers are large molecules, which can be found in the nature and also prepared in the laboratory (synthetic). Biocompatible and biodegradable polymers degrade to low-molecular-weight biocompatible products and are gradually replaced by the growing tissue. The goal is to have a properly and faster healing tissue. In this article the various polymer types used in tissue engineering are presented.

A polymer is a structurally ordered macromolecule made up of a large number of atoms.^[1] The molecules of a polymer have a relatively small unit, the monomer that repeats itself many times to form the polymer. The polymerization reaction takes place via addition or condensation of monomers to construct the main backbone of the polymer chain. Polymers are extensively used in medicine, in applications such as surgical devices, implants, drug delivery systems, biosensors, and tissue engineering.

A biomaterial is designed to substitute, complement, or completely take over the function of a diseased or incapacitated tissue or organ and polymers constitute a major portion of the biomaterials available for human healthcare alongside natural materials, metals, and the ceramics. Biomaterials are required even though the best solution for aiding or substituting tissue and tissue function is an autograft, because tissue transferred to the affected site from the patient's own body is either in limited quantity or not sufficiently "healthy" to transfer. Allografts (tissue from another member of the same species) and xenografts (tissue from another species) have a much lower success rate than with the autografts.

TISSUE ENGINEERING

Tissue engineering is by far the most exciting new field of biomaterials. It has gained so much popularity that it is being considered as a field of its own and the field is now frequently called "biomaterials and tissue

engineering," in contrast to "biomaterials" of a few years earlier. A variety of tissues are being engineered using this approach including artery,^[2] skin,^[3] cartilage,^[4] bone,^[5] and ligament.^[6] Several of these tissues are now in or are nearing clinical use.

A number of researchers have offered their definitions of what tissue engineering is. According to one, it is "the application of engineering principles to create devices for the study, restoration, modification, and assembly of functional tissues from native or synthetic sources".^[7] Another definition is that it is "the application of the principles and methods of engineering and the life science toward the fundamental understanding of structure–function relationships in normal and pathological mammalian tissue and the development of biological substitutes to restore, maintain, or improve functions".^[8] Two pioneers of the field, Langer and Vacanti (1993), have defined it as "an interdisciplinary field that applies the principles of engineering and the life sciences toward the development of biological substitutes that restore, maintain, or improve tissue function".^[9] It actually is the process of creating living, physiological, three-dimensional tissues and organs utilizing specific combinations of cells, cell scaffolds, and a variety of chemical and biological signals (Fig. 1).^[10] First, a source of cells derived from the patient or from a donor is required. These cells could be stem cells, the undifferentiated cells with capability to be converted into a variety of cell types, or differentiated cells of the desired tissue. Since most of the cells in our tissues have an extracellular matrix as a physical support, an anchorage site and source for signals, a number of biological signals inherent in these tissues need also to be provided. Thus, engineering of surfaces of 2-D and 3-D biodegradable physical supports is an essential task for the success of tissue engineering. The tissue growth in or on these scaffolds in the laboratory before being transferred to the recipient (the patient) is required for tissue engineering applications. The efficacy and safety of both the materials and the cells must be tested for their approval by regulatory authorities. Here, a crucial parameter is the interface between the two ingredients, and the reactions at the interface between the composite structure and the tissue.

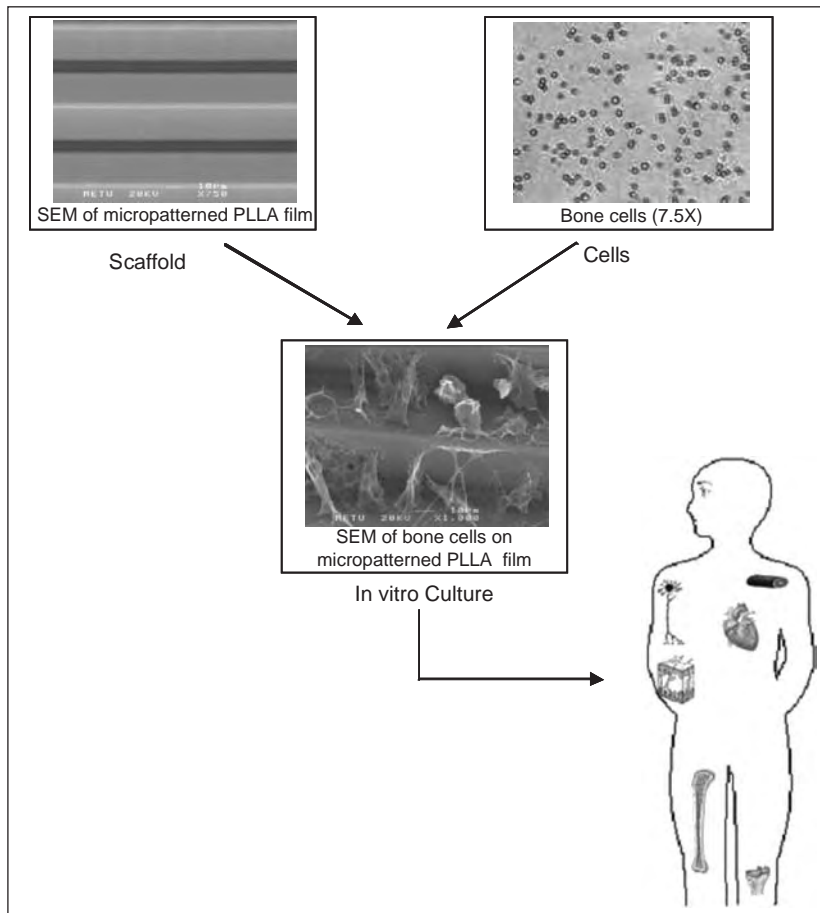


Fig. 1 Schematic presentation of tissue engineering concept using a bone tissue engineering model.

Cell Source

The cell sources used in tissue engineering are:

- Autologous cells (from the patient)
- Allogeneic cells (from a human donor)
- Xenogeneic cells (from another species)

The best cell source for tissue engineering is obviously the patient's own cells, in other words, autologous cells, especially because it eliminates the risk of rejection or adverse tissue reactions.^[11] The problem associated with autologous cells is the questionable healthiness of these cells and also the quantity of cells that could be harvested from a patient who probably is in need of extra healthy cells. Besides, these cells are at an advanced stage of development, and therefore, do not have so many cell divisions to go through. Until recently, it was thought that tissues of an adult would not have any stem cells and existence of such cells was only known in embryos and newborn animals. However, stem cell populations were found in many adult tissues and even amongst the nerve cells in the mammalian brain.^[12] Stem cells have a great potential for proliferation and differentiation. For example, the cells in the

bone marrow can differentiate into cells circulating in blood as well as into osteoblasts and chondrocytes.

Importance of ECM and Scaffold: Housing the Cells

Extracellular matrix (ECM) consists of proteins (collagen, laminin, fibronectin, etc.) and glycosaminoglycans (hyaluronic acid, chondroitin sulfate, heparan sulfate, etc.), is a shelter for cells and forms the skeleton of the tissues. These are large molecules, each with domains involved in cell recognition, adhesion, spreading, differentiation, and proliferation.^[13] The most important group for cell adhesion is the arginine–glycine–aspartic acid (RGD) sequence, which has been initially identified as a minimal sequence required to mediate cell adhesion to the ECM molecule fibronectin,^[14] and has since been found to be present in a wide range of molecules, including the blood circulating fibrinogen. After the discovery of RGD sequence, many such small adhesion-mediating peptide domains have been identified within ECM molecules.^[15] These peptides interact with a class of cell surface adhesion molecules called integrins.^[16] A member of the integrin family consists of a common

beta chain and variable alpha chain and is very diverse.^[17] Integrins are known to mediate many aspects of cell behavior through triggering a number of signaling pathways, some of which are related to cell adhesion, while others are related with differentiation. Therefore, an intense research activity is observed in the direction of influencing the integrin binding to material surfaces, which in return influence ECM ordering leading the cells to follow suit.

At this point ECM needs to be studied, for it is the natural material that carries the cells, and that is what a biodegradable carrier needs to mimic. The ECM is not only a surface on which cells reside but is an active source of physical and chemical signals influencing cell behaviour. Of a large number of scaffolds fabricated from various kinds of materials, a very large fraction are of polymeric origin because of their biodegradability and suitability for functionalization and attachment of biosignal molecules. Researchers, therefore, incorporate a variety of factors such as RGD sequences, growth factors and surface patterns on biodegradable polymeric films and foams.

Processed extracellular matrix (decellularized) scaffolds, such as small intestinal submucosa (SIS), urinary bladder submucosa (UBS),^[18] and porcine heart valves^[19] are also used as scaffolds; however, there are concerns over adverse immune response.^[19]

Properties of an Ideal Cell Carrier

The cell carrier, or as it is widely known, the scaffold, for tissue engineering should meet several design criteria:^[20]

1. Scaffold should preferably be biodegradable and eventually eliminated
2. Scaffolds should be biocompatible; neither the carrier, which is predominantly a polymer, nor its degradation products should provoke adverse responses (inflammation, toxicity, carcinogenicity, thrombosis, etc.) *in vivo*
3. Surface should have appropriate chemistry and physical properties that permit cell adhesion, cell growth, spreading, and regaining functions it had before harvesting
4. Scaffolds should be highly porous to provide space for cell adhesion and proliferation, ECM accumulation, and present minimal diffusional constraints during culture
5. Material should be conveniently processable into 2-D and 3-D forms, and mechanically strong enough for implantation into the recipient.

None of the currently available materials satisfies these requirements sufficiently, and therefore, surface

modification, attachment of biological cues, or formation of composites is employed. Thus, the tissue engineers started incorporating natural or modified cell-signalling molecules into the scaffolds to interact with specific receptors and elicit selective cellular responses, proliferation or differentiation. It has been recognized since the mid-1970s that three-dimensional structures are an important component of engineered tissue development.^[21] The properties, especially pore size, pore size distribution and connectivity, pore orientation, if the carrier is a mesh-like structure the fiber dimension, and form are important characteristics in addition to the chemical properties of the material.

MATERIALS USED IN BIOMATERIALS AND TISSUE ENGINEERING

Polymers, natural materials, metals, and ceramics have all contributed to the medical field.^[20] Metals and ceramics are not biodegradable, their processability is very limited, and the capacity of their surface or the bulk to be modified is very limited in comparison to polymers and natural materials. That is why polymeric materials have received more and increasing attention and are being widely used in tissue engineering.

Among the polymers suitable for tissue engineering are natural polymers, such as collagen, hyaluronic acid, chondroitin sulfate, agarose, and alginate.^[11] Synthetic biodegradable polymers, such as poly(lactide), poly(glycolide), polydioxanone and polycaprolactone, and nondegradable polymers like polyvinyl alcohol, polyethylene, polyethylene glycol, polypropylene oxide (Pluronic), and silicones are among the most frequently used. Rarely hydroxyapatite and bioglasses are used as tissue engineering carriers. Porous hydroxyapatite have been used to house osteoprogenitor cells derived from the periosteum or the bone marrow.^[22] Since the ceramic materials degrade over a long time *in vivo*, they are not widely used. In contrast, polymers could be designed to have a range of degradation time extending from very short (days) to long (several months). In addition, polymeric cell carriers could be processed in the form of fibers, fibrous meshes, porous sponges, or foams, which create an interconnective labyrinth of pores and channels just as needed for the cells.

Polymer Stability: Degradable and Nondegradable

Among the many properties desired in a tissue engineering carrier, degradability has an important place. Polymers degrade in the biological medium by simple and/or enzymatic hydrolysis. It is not a simple task to

identify the actual cause of degradation and the level of contribution of each degradation path, and therefore, in this manuscript biodegradation and degradation have been used interchangeably. Polymer degradation is influenced by the hydrophilicity in the side chains and the backbone, by the crystallinity and porosity.^[23] Ideally, a biodegradable polymer considered for medical applications should have adequate mechanical properties to match the application (strong enough for the duration of in vitro processing and for a certain time after implantation), and should be fully metabolized to biocompatible molecules once it degrades.

The rate and mode of degradation are important. Although most tissues, particularly bone, can tolerate degrading materials, even slowly degrading implants elicit a foreign body response. In bone, where bone tissue formation may be observed adjacent to a non-degradable implant, the mechanical properties of the final hybrid tissue may be poor compared to native tissue. Thus, materials degradable within an appropriate time frame are strongly preferred. Degradation is initiated when water penetrates the bulk of the polymeric system, attacking primarily the amorphous regions and converting long polymer chains into shorter, water soluble fragments by hydrolysis.^[24] This leads to an increase in overall crystallinity of the remaining polymer. Degradable polymers have two main routes of passive hydrolysis: random scission and unzipping. Both routes, however, involve chain scission (lysis of the polymer backbone by water) that leads to a variety of chain lengths depending on the mode of degradation. In the case of unzipping, hydrolysis starts from chain ends while in random scission bond breakage occurs randomly. In unzipping, the products are monomers and polymers with no chains of intermediate length, while in random scission the products are polymer chains cleaved statistically. Degradation may also be initiated by a biologically active process involving enzymes. Degradation also occurs on the side chains of the molecules but this category is not as important as the backbone cleavage.

Another category of materials are called “erodible” or “bioerodible” simply because they are removed by solubilization of the polymer rather than by bond breakage.^[24] This leads to removal of implants from the implantation site through loss of material from the implant surface.

POLYMER SOURCES

There are two sources of polymers used in tissue engineering: synthetic and natural (Table 1). The main biodegradable synthetic polymers include polyesters, polyanhydrides, polyorthoesters, polycaprolactone, polycarbonate, and polyfumarate, while the natural

origin polymers include collagen, alginate, agarose, hyaluronic acid derivatives, and chitosan. Among the man-made polymers, polyglycolide, polylactides, poly(caprolactone), and poly(dioxanone) constitute the major polymer groups that have been studied as matrix materials.

Important Synthetic Polymers Used in Tissue Engineering

Poly(α -hydroxyesters): polylactides (PLA) and polyglycolides (PGA)

Naturally occurring hydroxy acids, such as glycolic, lactic, and ϵ -caproic acids, have been polymerized to create an array of useful biodegradable polymers for a variety of medical applications.^[25] Polylactides, polyglycolides, and their random copolymers [Poly(lactide-*co*-glycolide) (PLGA)] belong to the family of aliphatic α -hydroxy polyesters (Fig. 2). They have been used extensively to fabricate scaffolds for engineering tissues.

PLA and PGA are often synthesized by ring-opening polymerization of the cyclic diesters, glycolide or lactide.^[26] Catalysts such as antimony, zinc, or lead are used in the polymerization process. The chemical structures of PLA and PGA are similar, except that the PLA has a methyl pendant group instead of a H in the backbone, and this difference in structure is the basis of all the significant differences observed in their properties, such as in degradation rate, T_g , T_m , and hydrophilicity. The higher hydrophobicity of PLA limits its water uptake and decreases its degradation rate.^[27,28] When a copolymer PLGA is formed, its properties depend on the ratio of LA and GA present.^[26] Passive hydrolytic scission, which involves random cleavage of polymer backbone by simple hydrolysis of the ester linkages, is the primary mode of degradation. Hydrolytic scission eventually forms the monomers, lactic acid (incorporated into the tri-carboxylic acid cycle and eliminated from the body as carbon dioxide via the lungs), and glycolic acid (excreted in the urine or incorporated into the tri-carboxylic acid cycle).

PLGA polymers have characteristic degradation ranging from days to years, depending on specific formulation, processing conditions, and initial molecular weight.^[10] Lactic acid is a chiral molecule, and therefore, exists in L and D forms. The L isomer is the biological metabolite. When this stereoisomerism is taken into consideration, the polymer poly(L-lactic acid) (PLLA) is a completely stereoregular polymer as poly(D-lactic acid) (PDLA). Poly(D,L-lactic acid) P(D,L-LA), on the other hand, is an amorphous polymer where D and L forms are mixed in the polymer

Table 1 Polymers and cell types used in some tissue engineering applications

Application area	Cell type	Polymer	Reference
Bone tissue engineering	Osteoblasts	PLA/calcium metaphosphate	[106]
		PLLA/hydroxyapatite	[107]
		Silk fibroin films	[64]
		Starch/cellulose acetate	[80]
		Chitosan/alginate	[81]
	Bone marrow stromal cells	RGD modified alginate hydrogel	[108]
		Poly(caprolactone)	[36]
		PHBV foams	[109]
		Calcium phosphate loaded collagen foams	[109]
		Starch/polycaprolactone blend	[110]
Cartilage tissue engineering	Chondrocytes	Collagen film crosslinked with carbodiimide and glutaraldehyde	[62]
		PLGA	[32]
		Poly(caprolactone)	[36]
	Mesenchymal stem cells	HYAFF 11	[77]
		PHBV	[88]
		Poly(NIPAAM-co-AAc)	[103]
		Poly(caprolactone)	[111]
Skin tissue engineering	Fibroblasts	Type I collagen matrix	[113]
		PLGA/collagen mesh	[112]
	Endothelial cells, keratinocytes and fibroblasts	Collagen/glycosaminoglycans	[90]
Blood vessel tissue engineering	Smooth muscle cells and endothelial cells	PGA mesh and PLGA film (respectively)	[31]
	Venous endothelial cells and vascular smooth muscle cells	PLGA nonwoven mesh	[114]
Liver tissue engineering	Hepatocytes	Alginate	[115]

chain even though the chemistry is that of pure polylactic acid. Since amorphous regions are more susceptible to penetration of water, the rate of hydrolysis is much higher when copolymers or nonstereoregular PLA are used. In the body, the initial degradation occurs until the molecular weight decreases to around 5000 Da, at which point biological degradation takes over.^[25] The final degradation and resorption of the poly(α -hydroxy acid) implants involves inflammatory cells (such as macrophages, lymphocytes, and neutrophils). These alpha-polyesters are hydrophobic, therefore, insoluble in water, and are degraded rapidly in acidic aqueous media. These thermoplastic polymers are generally formed into films, tubes, matrices, fibers, meshes, and open-cell foams (Fig. 3) using standard

processing techniques. These polymers form rigid films and foams which crumble upon degradation, sometimes creating crystalline particulates that remain in the implantation site for years, leading to mechanical damage or sterile sepsis.^[10] In addition, the acid nature of the products cause pH drop at the implantation site damaging cells and neighboring tissue.

PGA itself is highly crystalline (45%–55%), has a T_m of 220°C–225°C and a T_g of 35°C–40°C.^[24] Because of its high degree of crystallization, it is not soluble in most organic solvents; the exceptions are highly fluorinated organics such as hexafluoroisopropanol. PLLA is a semicrystalline polymer with 37% crystallinity, and has a T_m of 175°C–178°C and a T_g of 60°C–65°C, while P(d,l-LA) is an amorphous polymer.^[24]

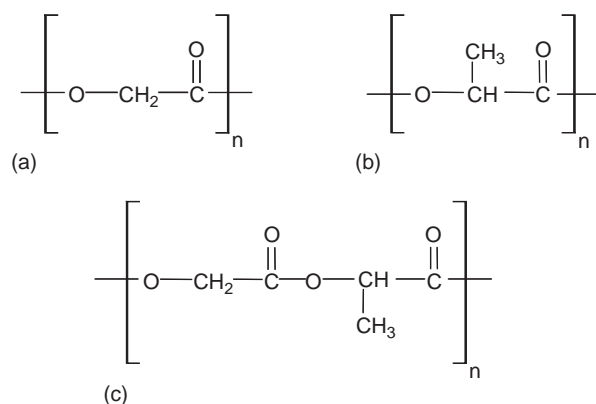


Fig. 2 Chemical structure of (A) PGA, (B) PLA, (C) PLGA.

The degradation rate of PLLA is much slower than that of P(d,l-LA), requiring more than 2 years to be completely absorbed. The 50:50 glycolide and lactide copolymer degrades faster than either homopolymer.^[24] The 10:90 copolymer (10% l-lactide and 90% glycolide) was developed by Ethicon as an absorbable suture material under the trade name Vicryl and degrades within 3–4 months. These polymers have gained the approval of the U.S. Food and Drug Administration for human clinical use in a variety of applications, including sutures.^[29] However, in spite of their wide use in a variety of biomedical applications, their acidic degradation products^[30] and lack of functional groups for further modification limit their usefulness.

The use of PLA and PLGA in tissue engineering scaffolds is very common. In order to give an idea, a few examples are cited. In constructing a tissue-engineered blood vessel, Niklason and Langer^[31] used PLGA- and PGA-based scaffolds and seeded endothelial cells onto PLGA films and bovine aortic smooth muscle cells onto PGA mesh. PGA meshes seeded with smooth muscle cells developed into a tissue-like appearance and dense, albeit disorganized, cell structure after 4 weeks in a static culture. Shi et al.^[32] constructed 90% porous PLGA (50:50) scaffolds through solvent casting/salt leaching to culture chondrocytes for cartilage restoration. Moreover, Hasirci et al.^[33] reported that disk-shaped PLGA (85:15) foams represented a promising scaffold for scaling up rat hepatocyte culture.

Poly(ϵ -caprolactone)

Poly(ϵ -caprolactone) (PCL) is an aliphatic polyester, that can be prepared by ring-opening polymerization of caprolactone like the PLA and the family (Fig. 4).^[24] It is semicrystalline, with a T_m of 59°C–64°C and a T_g of –60°C. The polymer can be degraded enzymatically as well as hydrolytically.^[25] Low molecular weight

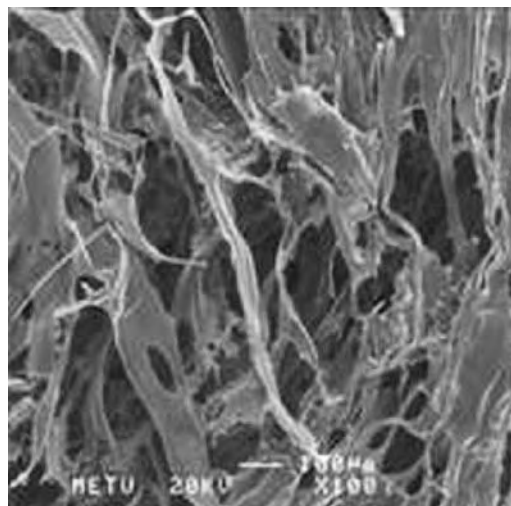


Fig. 3 Scanning electron micrograph of a PLLA foam.

PCL in particulate form is reported to be taken up by macrophages and degraded intracellularly.^[34] It has a degradation time of about 2 years, a period generally longer than desired for tissue engineering applications and slower than those of PGA or PLA.^[24] It, however, has the capability to form blends and copolymers with other biodegradable polymers and could therefore be rendered suitable for tissue engineering.

Williams et al.^[35] produced a computer-designed, porous PCL scaffold with selective laser sintering (SLS), a rapid prototyping technique. Bone morphogenetic protein-7 (BMP-7) transduced fibroblasts were seeded and the construct was implanted subcutaneously as potential bone and cartilage implants. The results indicated that this carrier was suitable for tissue engineering. Cao et al.^[36] designed an osteochondral construct by coculturing osteogenic and chondrogenic cells. The 3-D porous PCL polymer scaffold was prepared by fused deposition modeling. Both cell types proliferated, migrated, linked in their scaffold compartments, integrated at the interface, and produced different extracellular matrices. High osteocalcin and high alkaline phosphatase were found indicating the suitability of this system to osteochondral defect repair.

Polyanhydrides

The anhydride bond is more unstable against hydrolytic attack than esters, amides, and ethers.^[15] Their

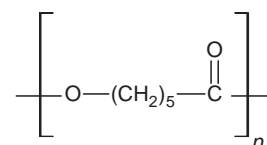


Fig. 4 Chemical structure of PCL.

P

relatively high degradation rate is both an advantage and a limitation. Degradation of polyanhydride devices is initiated by access of water into the device and can be controlled by the hydrophobicity of the monomers. They are known to be biocompatible,^[37] susceptible to hydrolysis, and degrade rapidly through surface erosion.^[26] Ibim et al.^[38,39] compared poly(anhydride-*co*-imide) polymers to PLGA and revealed poly(anhydride-*co*-imide) polymers to be equally biocompatible to PLGA and able to support endosteal and cortical bone regeneration. Thus, they identified poly(anhydride-*co*-imide) polymer scaffolds as a potential implant for orthopedic surgery, specifically in weight-bearing applications. Extensive engineering of copolymer anhydrides has yielded formulations with a range of mechanical properties, degradation times, and chemical reactivities.^[15]

Polyfumarates

Poly(propylene fumarate) (PPF) is a viable alternative to the frequently used polyesters, PLGA family, because of their biocompatibility and biodegradability in many biological applications. The main advantage of this polymer, however, is its unsaturated structure which makes it suitable for curing after implantation or processing.^[26] Thus, they are very suitable for injectable applications. It was reported to be evaluated as an injectable scaffold in the form of a composite with β -tricalcium phosphate (β -TCP).^[40]

Peter et al.^[41] using human BMP-2 releasing PLGA microspheres, reported that the mesenchymal cells seeded on PPF scaffolds showed an increased ALP activity, osteocalcin secretion, and mineralization. Payne et al.^[42] studied culturing of rat marrow stromal osteoblasts encapsulated in gelatin microparticles on crosslinked PPF and showed that these cells exhibited a high viability, proliferation, and phenotypic expression on PPF.

Synthetic poly(amino acids)

These constitute a class of polymers based on natural amino acids linked by amide and nonamide bonds, including ester, iminocarbonate, urethane, and carbonate to obtain polymers with desired properties, such as predetermined composition with desired sequences, chain length, T_g , crystallinity, and hydrophobicity.^[10] They are produced either through the use of molecular biology techniques or solid state synthesis, with the desired amino acid sequences to carry the pendant groups, sequences, or the properties needed for cell adhesion, spreading, and proliferation and differentiation.^[20] They are also designed to possess improved mechanical properties, easy processability,

appropriate stability, well defined composition.^[10] As a result of the production methods, they possess a fixed molecular weight or a narrow molecular weight distribution. Since the degradation products are naturally occurring amino acids, they are expected to cause a very low level of biological response.^[25]

Most poly(amino acids) are highly insoluble due to their high crystallinity and are nonprocessable materials.^[25] A major problem appears to be the difficulty in reproducing and controlling their degradation in vivo because of patient-to-patient variation. Another problem is the antigenicity of the poly(amino acid) polymers containing three or more amino acids.

Among the many polymers in this family, poly(tyrosine carbonates) show promise in orthopedic applications due to improved behavior of degradation products in bony sites.^[43] These polymers possess three potentially hydrolyzable bonds: amide, carbonate, and ester. The carbonate group hydrolyzes at a faster rate than the ester group, and the amide bond is not labile in vitro.^[43] By variation of the structure of the pendant R group, polymers with different mechanical properties, degradation rates as well as cellular response, could be prepared.

Betre et al.^[44] studied artificial elastin-like polypeptide (ELPs), composed of a pentapeptide repeat, Val-Pro-Gly-Xaa-Gly (Xaa is any amino acid except Pro), and cultured chondrocytes in ELP coacervates for cartilaginous repair purposes. The cells maintained a rounded morphology and chondrocytic phenotype (synthesis of ECM of sulfated glycosaminoglycans and collagen). Jeschke et al.^[45] showed that coating synthetic scaffolds with tailor-made, cyclic RGD-peptides helped maintain attachment of human and pig chondrocytes.

Holmes et al.^[46] synthesized self-assembling peptide scaffolds by solid-state-phase peptide synthesizer and revealed that these peptide-based scaffolds served as ideal substrates for neurite outgrowth and synapse formation.

Poly(ortho esters)

Poly(ortho esters) are a family of synthetic degradable polymers that have been under development for many years.^[25] Originally, poly(ortho esters) were prepared by the condensation of 2,2-diethoxytetrahydrofuran and a dialcohol.^[47] Upon hydrolysis, these polymers release acidic by-products that autocatalyze the degradation process, resulting in degradation rates that increase with time. A new type of poly(ortho ester) was synthesized based on the reaction of 3,9-bis(ethylidene 2,4,8,10-tetraoxaspiro [5,5] undecane) with various dialcohols.^[48] These poly(ortho esters) are not hydrolyzed to release acidic degradation products, and

therefore, do not degrade at an increasing rate due to autocatalysis.

Poly(ethylene oxide) (PEO) and poly(ethylene glycol) (PEG)

The ethylene oxide monomer is composed of an epoxide ring, as two corners of the molecule consist of $-\text{CH}_2-$ linkages, the third corner is an oxygen, $-\text{O}-$, and by addition of a catalyst, the monomers form a chain having the repeat unit $-\text{CH}_2-\text{CH}_2-\text{O}-$. Poly(ethylene glycol) (PEG, also known as PEO at high molecular weights) is hydrophilic, soluble in water and in organic solvents, including aromatic hydrocarbons (not aliphatics), lack of toxicity, and absence of antigenicity and immunogenicity.^[49] Considering these properties, PEG can be used for many biomedical and biotechnological applications, especially as a biomaterial.

Multiblock copolymers of PEO are being developed as prosthetic devices and artificial skin and as scaffolds for tissue engineering.^[50] These materials undergo both hydrolysis (via ester bonds) and oxidation (via ether bonds). Degradation rate is influenced by PEO molecular weight and content, and the copolymer with the highest water uptake degrades most rapidly.

PEO is currently FDA approved and is one of the most commonly used synthetic polymers for tissue engineering.^[51] PEO, and the chemically similar PEG, are hydrophilic polymers that can be photocrosslinked with UV and in the presence of photoinitiators after modifying each end of the polymer with either acrylates or methacrylates and hydrogels are then formed.

Thermally reversible hydrogels have also been formed from block copolymers of PEO and PLLA^[52] and of PEG and PLLA.^[53] Degradable PEO and PEG hydrogels have also been reported as block copolymers of PEO and PEG with PLA.^[54] A newer, synthetic hydrogel block copolymer, P(PF-*co*-EG) has been created for use as an injectable carrier for bone and blood vessel engineering.^[55]

Important Natural Polymers Used in Tissue Engineering

The naturally derived polymers consist of bioactive proteins (enzymes, growth factors), structural proteins (collagen, an ingredient of the ECM), and polysaccharides (glycosaminoglycans, alginic acid, chitosan). Among the first tissue regeneration materials are the collagen–glycosaminoglycan scaffolds of Yannas et al.^[21] who tried to create the dermis of skin. Since these materials are used after reconstitution, their natural properties are somewhat less satisfactory than before removal from their natural environment, and therefore need to be improved. Along this line, their degradation

rates are controlled and mechanical properties increased through chemical crosslinking by glutaraldehyde, formaldehyde, carbodiimide, epichlorohydrin, etc.

Animal origin/protein-based polymers

(a) *Collagen*. Collagen is the most abundant protein in animals (it makes up 30% of all proteins in the human body)^[25] and because of its high mechanical strength and good resistance to degradation, it has been utilized in a wide range of products in the medical industry. More than 20 genetically distinct collagens have been identified.^[56] They are mainly responsible for maintaining the structural integrity of the body and are mainly found in tendons, bone, cartilage, and skin. Collagen is arranged in fibrils with a 67 nm axial periodicity and the fibrils consist of triple helices approximately 300 nm in length and 1.5 nm in diameter. Only collagen types I, II, III, V, and XI self-assemble into fibrils. Collagen fibril formation is finalized by the removal of the telopeptides (the terminal procollagen peptides) by specific procollagen metalloproteinases. Some collagens form networks (types IV, VIII, and X), as is the basement membrane mostly made of type IV.^[56] Type I collagen occurs throughout the body, except in cartilage. Type II collagen occurs in cartilage, the developing cornea, and in the vitreous body of the eye. Type III collagen occurs mainly in the wall of blood vessels and hollow intestinal organs and copolymerizes with type I collagen. Collagens are mostly synthesized by the cells comprising the ECM: fibroblasts, myofibroblasts, osteoblasts, and chondrocytes.

Collagen can be obtained by extraction and purification of natural tissue or as specific recombinant polypeptides. Its mechanical properties can be controlled by crosslinking using chemical agents such as glutaraldehyde, formaldehyde, carbodiimide, or by physical treatments such as UV irradiation, freeze-drying, heating, and dehydration, and thus make it suitable for use in a variety of biomedical applications, including tissue engineering (Fig. 5).^[51] It has also been reported as blends with hyaluronic acid,^[57] chitosan,^[58] and PEO.^[59] Collagen is naturally degraded by metalloproteases, specifically collagenase, and serine proteases allowing for its degradation to be locally controlled by cells present in the engineered tissue.^[51]

Wakitani et al.^[60] seeded chondrocytes in collagen gels to repair articular cartilage defects. The chondrocytes maintained their phenotype and new cartilage was observed as early as 4 weeks. The grafted tissue remained as cartilage over 48 weeks. Orwin and Hubel^[61] investigated collagen sponges as carriers for epithelial and endothelial cells, and for keratocytes to construct corneal tissue. To overcome the loss of transparency caused by contraction of the collagen gel

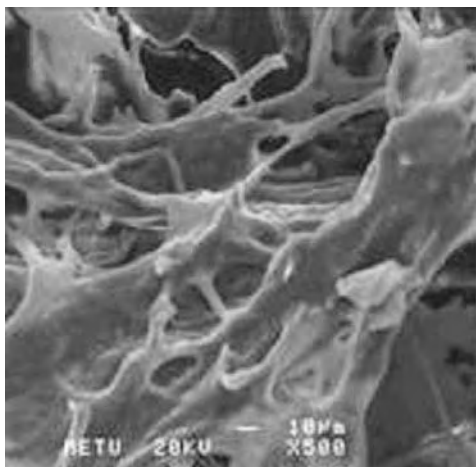


Fig. 5 Scanning electron micrograph of a collagen foam.

by keratocytes, they prepared a dispersion of collagen in dilute acid, pH 4, followed by freeze-drying. Thus, constructed acid-swollen sponges were expected to have less fibrillar structure, and therefore scatter less light than fibrils formed at neutral pH. Ber et al.^[62] investigated guiding osteoblast cells from rat bone marrow on micropatterned collagen films crosslinked with carbodiimide (EDC), glutaraldehyde or by dehydrothermal route (DHT). It was observed that the cells seeded on calcium phosphate deposited collagen films have a high osteoblastic phenotype expression (alkaline phosphatase activity), emphasizing the effect of mineral presence on osteoblast differentiation.

(b) *Silk*. Silks, are a group of fibrous proteins produced by a wide range of insects and spiders.^[63] They consist of helical, β -sheet, or cross- β -sheet secondary structures. Spider dragline silk from *Nephila clavipes*, and cocoon silk fibroin from the silkworm, *Bombyx mori*, have characteristic β -pleated sheet secondary structures with extended polypeptide chains, in which the carbonyl oxygens and amide hydrogens are at near right angles to the long axis of the chain and the sheet. The interchain interactions are through intra- and inter-hydrogen bonding and van der Waals interactions. Because of the impressive mechanical properties, silk proteins are important materials for use in biomaterials, especially in controlled release, as sutures and as scaffolds for tissue engineering.^[64] Silks are insoluble in most solvents, including water, dilute acid, and alkali. Silk has the advantage of being easily chemically modified due to the availability of amine and acid side chains on some of the amino acids. Its composition can be genetically tailored to provide it with moderate specific features, such as molecular weight, crystallinity, solubility, and slow rates of degradation in vitro and in vivo. There are, however, some concerns, such as inadequate removal of sericin from silkworm silk leads to biocompatibility problems, slow degradation of

crystalline (β -sheet) regions prolongs removal time, aborted proteolytic attack by macrophages and giant cells lead to encapsulation and the formation of a granuloma, and potential sensitization to silk fibroin resulting in an allergic response upon exposure to the biomaterial. On a brighter note, silk fibroin can be processed into foams, films, fibers, and meshes.

Minoura et al.^[65] compared silk fibroin and collagen films in terms of supporting attachment, spreading, and growth of the L-929 fibroblast and found them to be equivalent in their ability to support cell attachment and physiological morphology. Sofia et al.^[66] studied silk fibroin films in vitro by seeding with osteoblasts. The films were chemically bonded with RGD sequences to promote increased adhesion. The increase in alkaline phosphatase levels, upregulation of bone-specific transcripts, and calcification levels over 4 weeks showed the response of osteoblast cells to silk fibroin films. Unger et al.^[67] reported that various human cells from different tissues (e.g., epithelial cells, fibroblast, glial cells, keratinocyte, osteoblast) adhered and spread over a nonwoven silk fibroin net and maintained their growth, making it a potential cell carrier.

(c) *Fibrin*. The conversion of a soluble plasma protein fibrinogen to an insoluble fibrous protein fibrin is initiated by the enzyme thrombin.^[68] The polymerization step causes both end-to-end and side-to-side aggregation of fibrinogen to form fibrin fibers.

Ye et al.^[69] constructed a 3-D fibrin gel scaffold for blood vessel tissue engineering. Human myofibroblast cultures isolated from human aortic tissue were suspended in fibrinogen solution and seeded into 6-well culture plates for cell development for 4 weeks and supplemented with different concentrations of aprotinin. Results of tissue development studies on 3-D fibrin gel structures showed homogenous cell growth and confluent collagen production without any toxic degradation or inflammatory reactions. Pellegrini et al.^[70] showed that fibrin is a suitable substrate for keratinocyte cultivation and transplantation. Their results showed that when fibrin-cultured autografts bearing stem cells are applied on massive full-thickness burns, the “take” of keratinocytes is high, reproducible, and permanent.

(d) *Other protein-based polymers*. A series of recently introduced casein- and soy-based biodegradable thermoplastics have joined collagen as a source of natural protein-based biomaterials.^[23] They are reported to be less susceptible to thermal degradation, and therefore, can safely be processed at elevated temperatures as scaffolds for tissue engineering.

Animal origin/polysaccharide-based polymers

Polysaccharides have generally been underutilized as biomaterials.^[71] A number of studies are pointing to

the critical role of saccharide moieties in cell signaling and in immune recognition, thus increasing the awareness of this important category of polymers. Recently, the development of new automated synthesis techniques made production of biologically active oligosaccharides more plausible than before. One major property of polysaccharides is their capability to form hydrogels. They can form gels by hydrogen-bonded and by ionic interaction. Hydrogen-bonded gels are mainly formed by agarose via thermal gelation, and by chitosan via pH modification. Alginates and carrageenans, on the other hand, form ionically crosslinked hydrogels.

(a) *Chitosan*. Chitosan is obtained by partially deacetylating chitin that is found in arthropod exoskeletons.^[71] Commercially available preparations have degrees of deacetylation ranging from 50 to 90%. It is a linear copolymer of $\beta(1 \rightarrow 4)$ linked D-glucosamine and a variable number of N-acetyl-glucosamine. The average molecular weight of chitosan varies in the range 50–1000 kDa, depending on the source and preparation procedure. It is semicrystalline and the degree of crystallinity varies with the degree of deacetylation. Crystallinity is maximum for both chitin (0% deacetylated) and fully deacetylated (100%) chitosan, and minimum when there is an intermediate degree of deacetylation. Chitosan is normally insoluble in aqueous solutions above pH 7 but in dilute acids with pH lower than 5 it becomes soluble because of the protonation of the amine groups. Gel fibers can be drawn from viscous solutions to form high-strength fibers. Chitosan and its parent molecule, chitin, have been shown to exert chemoattractive effects on neutrophils in vitro and in vivo. In vivo, chitosan is degraded by enzymatic hydrolysis by lysozyme, which appears to target acetylated residues. The degradation products are chitosan oligosaccharides of variable length. The degradation rate appears to decrease with crystallinity and this, in return, is controlled by the degree of deacetylation. Highly deacetylated forms (>85%) exhibit the lowest degradation rates and may last several months in vivo. Derivatization modifies the molecular packing and increase the amorphous fraction, allowing more rapid degradation. Chitosan-based implants have been found to initiate minimal foreign body reaction with no major fibrous encapsulation.

Porous chitosan structures can be formed by freezing and lyophilizing chitosan-acetic acid solutions.^[72] Pore size can be controlled by varying the freezing rate and hence the ice crystal size.^[71] Pore orientation can be directed by controlling the geometry of thermal gradients during freezing. The mechanical properties of chitosan scaffolds are mainly dependent on the pore sizes and pore orientations.

Tuzlakoglu et al.^[73] constructed a 3-D chitosan fiber mesh for potential use in bone tissue engineering

applications. Chitosan fibers were produced by a wet spinning and they had enough tensile strength to be used to produce scaffolds with good mechanical performance. Osteoblasts directly cultured over chitosan fiber mesh scaffolds presented adequate morphology and proliferation. In the other bone tissue engineering application, Zhao et al.^[74] seeded osteoblasts on a 3-D hydroxyapatite/chitosan-gelatin network scaffold. Zhu et al.^[75] attached various amino acids (lysine, arginine, aspartic acid, phenylalanine) onto chitosan by using glutaraldehyde. These amino acid functionalized chitosan moieties were subsequently entrapped on PLA surfaces. The amino acid–chitosan–PLA membrane was shown to have a good cytocompatibility with chondrocytes.

(b) *Glycosaminoglycans*. Glycosaminoglycans (GAGs) are linear polymers of repeating disaccharide units of a uronic acid (e.g., glucuronic acid) and a hexosamine (e.g., N-acetyl-D-glucosamine).^[25] The major GAGs attached to the protein components of proteoglycans are chondroitin sulfate, hyaluronic acid, dermatan sulfate, keratan sulfate, and heparan sulfate. The most commonly used and the largest GAG is hyaluronic acid (hyaluronan), which is a linear, anionic polysaccharide with repeating N-acetylglucosamine and glucuronic acid units, and molecular weight in the range of 500 kDa to several thousand kDa. It can be isolated from natural sources (e.g., rooster combs) or produced by fermentation. Hyaluronic acid can be easily chemically modified, as by esterification of the carboxyl moieties or crosslinked to form hydrogels. The gels can be formed into sheaths, membranes, tubes, and particles of various shapes and sizes. No species variations have been found in the chemical and physical structure of hyaluronic acid. It is not antigenic, and elicits neither inflammatory nor foreign body reaction. Its main drawbacks in this respect are its residence time and the limited range of its mechanical properties.

Zavan et al.^[76] investigated whether an extracellular matrix obtained from fibroblasts cultured within scaffolds of HYAFF-11 (derived from the total esterification of hyaluronic acid with benzyl alcohol on the free carboxyl groups of glucuronic acid along the polymeric chain leading to an increase of the hydrophobic component of the polymer chain) could influence the proliferation rate and survival of hepatocytes. The hepatocytes exhibited a round morphology and re-established cell-to-cell contacts, and grew into aggregates of several cells along and/or among fibers in the fabric. Aigner et al.^[77] used benzyl ester of hyaluronic acid (HYAFF-11) as a scaffold for the culture of human nasoseptal chondrocytes for cartilage tissue engineering purposes. The scaffold allowed good viability and adhesivity of the cells. The cultured cells expressed cartilage-specific collagen type II, indicating

their ability to redifferentiate within the scaffold of HYAFF-11 and to retain a chondrocyte phenotype even after a long period under in vitro conditions.

Plant origin/polysaccharide-based polymers

(a) *Starch*. Starch is a plant-originated polysaccharide found in a variety of plants, including wheat, corn, rice, beans, and potatoes.^[78] It is a branched polymer formed by amylose as the linear portion (20% w/w) and the amylopectin branches. Granular starch is sometimes added to polyethylenes in order to make it biodegradable. Heating the starch in the presence of water during extrusion or injection molding causes the formation of a thermoplastic material that may be deformed during blending. This starch-based product is then blended with either natural or synthetic materials. Heating starch above its glass transition temperature breaks its molecular structure, allowing further bonding.^[79] Plastic materials from starch-based blends may be injection molded, extruded, blown, or compression molded.

Salgado et al.^[80] prepared 50:50 (wt%) blend of starch and cellulose acetate scaffolds to seed osteoblasts for bone tissue engineering. The scaffolds had adequate porosity and allowed the cells in the centre of the scaffolds to remain viable for more than 3 weeks. The maintenance of osteogenic phenotype and deposition of mineralized extracellular matrix showed the success of the study.

(b) *Alginate*. Alginate is a copolymer of β -D-mannuronic acid (M) and α -L-guluronic acid (G) isolated from seaweed. Gels are formed in the presence of divalent cations (eg. Ca^{+2}) and formed by covalently crosslinking alginate with adipic hydrazide and PEG using standard carbodiimide chemistry.^[51] It has advantages and disadvantages. It is widely available, relatively biocompatible and easily gels. However, its gels lose their integrity upon removal of the gelling cation by the biological medium. It is nondegradable in the human body and its composition and purity varies from batch to batch. The crosslinking density, and thus mechanical properties and pore size of the ionically crosslinked gels can be readily manipulated by varying the M to G ratio and molecular weight of the polymer chain.^[51]

Li et al.^[81] constructed a naturally derived chitosan and alginate polymers having significantly improved mechanical and biological properties as compared to its chitosan counterpart. They revealed that bone-forming osteoblasts readily attached to the chitosan-alginate scaffold, proliferated well, and deposited calcified matrix. The hybrid scaffold had a high degree of tissue compatibility in vivo and also promoted rapid vascularization and deposited connective tissue within the entire scaffold structure.

Polyesters of biological origin

Polyhydroxyalkanoates (PHAs) are thermoplastic, biodegradable, natural polyesters that are produced by numerous micro-organisms. They are intracellular storage polymers whose function is to provide a reserve of carbon and energy. The differences between various PHAs arise from the length of the pendant groups (Fig. 6). PHAs with short pendant groups are hard crystalline materials, whereas PHAs with longer pendant groups are elastomeric.^[82] Its high crystallinity (60%–90%) makes its hydrolysis by enzymes or physiological medium difficult.^[83] PHAs are combined or blended with other polymers and materials to improve properties, shape and/or performance.^[82] Research is mainly concentrated on two commercially available, more crystalline, more brittle poly(hydroxybutyrate) (PHB) (Fig. 6A) and the less crystalline, more flexible, more degradable copolymer of butyrate and valerate moieties, poly(3-hydroxybutyrate-co-3-hydroxyvalerate) (PHBV) (Fig. 6B). PHAs are fabricated in molded shapes, films, fibers (Fig. 7), woven and nonwoven materials, tubes, and composites.

PHB's degradation product, R-3-hydroxybutyric acid, is a normal constituent of blood at concentrations between 0.3 and 1.3 mM and is also found in the cell envelope of eukaryotes,^[83,84] and is therefore biocompatible. The polymer itself was not reported to be incompatible and a negative reaction cannot be expected due to the acidity of accumulated product because PHB has a rather low degradation rate in the body, extending into months and even years, depending on the crystallinity of the polymer. Thus, PHB can be accepted as a biocompatible material.

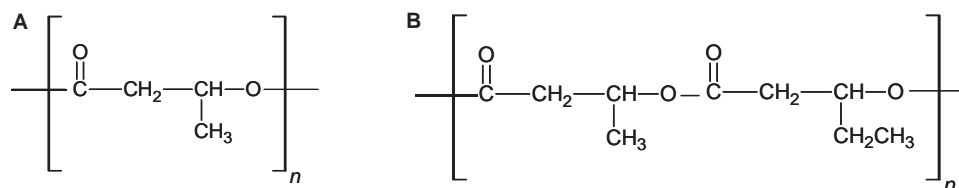


Fig. 6 Chemical structure of (A) PHB and (B) PHBV.

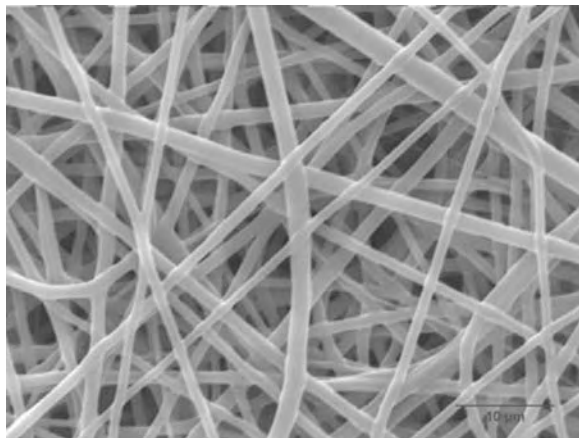


Fig. 7 Scanning electron micrograph of PHBV fibers.

Torun Kose et al.^[85,86] constructed macroporous PHBV matrices by solvent evaporation and solute leaching to cultivate osteoblasts for bone tissue engineering. It was revealed that the weight and density of the matrices were unchanged for a period of 120 days. Scanning electron microscopy results indicated that osteoblasts proliferated even in the inside of the matrices. In addition, growth of osteoblasts on matrices was observed after 60 days of incubation. Torun Kose et al.^[87] used PHBV matrices also for cartilage engineering and revealed that PHBV matrices show early cartilage formation resembling normal articular cartilage. In the late phase of the experiment, mature hyaline-like cartilage formation was observed in chondrocyte-seeded PHBV scaffolds. Sun et al.^[88] investigated a polymer scaffold system consisting of PHBV for possible application as a matrix for the three-dimensional growth of chondrocyte culture. PHBV scaffolds with *ca.* 80% porosity were constructed by compression molding, thermal processing, and salt leaching. They revealed that the proportion of chondrocytes in their natural round morphology increased and they could synthesize the type II collagen and cartilage-specific proteoglycans in the prolonged incubation time.

Modified Polymers of Natural and Synthetic Origin

When the polymer of choice has most of the properties needed, but is lacking in some others, one of the options is to modify the principal polymer by attaching functional groups or polymers via grafting through application of chemical compounds, with or without activation, with ionizing radiation, plasma, etc. The principal polymer defines the gross shape, size, and major mechanical properties of the final molecule, while the grafted component defines the chemistry and

mode of interaction with the biological interface. The synthetic macromolecules generally lack functional groups recognizable by the tissue, such as their lack of cell-recognition signals. In contrast, fibrinogen (Fb) or fibronectin (Fn) carry one of the many cell recognition sequences, such as RGD. Thus, the combination of the principal polymer, with the physical and mechanical properties required from a cell carrier, and the attachment of Fb or Fn, brings the advantage of cell recognition and thus creates a cell-adhering scaffold for use in tissue engineering.

Most approaches to tailored surface modification involves introduction of a specific adhesion peptide to the polymeric surface to enhance the degree of cell adhesion.^[15] A simple approach to this is attachment of a long hydrophobic peptide tail and spacer to the adhesion peptide. This, as well as some involving covalent linkage of peptides to the polymer backbone, introduces additional sites for adhesion allowing precise control over the surface properties. Cell migration and differentiation can be inhibited by strong adhesion of cells due to high ligand surface densities, and its control is of prime importance.

Besides protein- and peptide-attached synthetic polymers, sugar residue-bearing polymers offer a surface suitable for cell attachment and, thus, might be used to improve cell recognition in tissue engineering applications.^[89]

Modification of polymers to activate cell-material interface is very common in tissue engineering, and in order to give an idea a few examples are cited below. In the study of Grzesiak et al.^[90] collagen/glycosaminoglycan matrices modified with covalently-bonded RGD containing peptide were seeded with endothelial cells, keratinocytes, and fibroblasts. It was revealed that modification of the collagen/GAG matrices with RGD binding sites led to enhanced attachment of cells and was considered a potential structure for dermal regeneration. Santiago et al.^[91] modified poly (caprolactone) surfaces to promote the attachment and proliferation of adipose-derived stem cells (ASCs). The results showed that laminin-derived isoleucine-lysine-valine-alanine-valine (IKVAV) peptide sequences, covalently attached to an aminated polymer surface using carbodiimide chemistry, improved the attachment of ASCs to a tissue-engineered scaffold. For tissue-engineered vascular graft, He et al.^[92] designed a collagen-coated poly(l-lactic acid)-*co*-poly(ϵ -caprolactone) P(LLA-CL) (70:30) nanofiber mesh (NFM), which was fabricated using electrospinning, followed by plasma treatment and collagen coating. The collagen-coated P(LLA-CL) NFM enhanced the spreading, viability, and attachment of human coronary artery endothelial cells, and moreover, helped preserve their phenotype. Tezcaner et al.^[93] used PHBV film to serve as a temporary substrate for growing retinal pigment

epithelium (RPE) cells as an organized monolayer before their subretinal transplantation. The surface of the film was rendered hydrophilic by oxygen plasma treatment to increase the reattachment of D407 cells on the film surface. It was concluded that the oxygen plasma treated PHBV films were good substrates for use as temporary templates for subretinal transplantation of RPE cells.

Intelligent Polymers: Responsiveness Against T and pH

Intelligent materials are very useful in tissue engineering for a variety of reasons. They are known to help cell sheet removal from surfaces, control release of cell activating factors from microparticles, membranes, etc. or help load and unload cells and other support molecules onto the tissue engineering scaffold. Currently under development are anionic, cationic, and biomimetic hydrogels, controlled pore structures, and biodegradable polymers responding to specific biological conditions.^[23]

“Smart” or stimuli-sensitive materials can be either synthetic or natural and are generally polymer-based. Some smart polymers could undergo reversible gelation and some are covalently crosslinked to form irreversible hydrogels. Major examples of these kinds of polymers include poly(*N*-isopropylacrylamide) poly (NIPAAm), alginic acid and its univalent salts, chitosan, -carrageenan, and their combinations, such as collagen-acrylate and polyethylene glycol-peptides.^[94]

Smart polymeric materials are designed to respond to small changes in their environment with a considerable change in their properties.^[94] Smart systems change shape, surface characteristics, solubility, and form molecular self-assembly or sol-gel transition. Environmental stimuli include temperature, pH, chemicals, light, magnetic field and electric field.^[95]

The thermosensitive poly(NIPAAm) is one of the most commonly studied smart polymers, and various strategies for the synthesis of the hydrogel have been described.^[94] Efforts have been directed toward altering the swelling/shrinking behavior and preparing copolymers that also respond to other stimuli.

Poly(NIPAAm) and its copolymers are most often investigated for drug delivery^[96] and tissue engineering applications.^[97] An aqueous poly(NIPAAm) solution precipitates when the medium temperature is raised above 32°C, the phase transition temperature or the LCST below which favorable interactions via hydrogen bonding with water through the amide groups of the polymer, lead to dissolution of the polymer chains, and above the LCST, the hydrogen bonds are broken and water molecules are expelled from the polymer, resulting in the precipitation of the polymer.^[98] The

LCST of NIPAAm polymers can be controlled by copolymerizing it with other monomers with different hydrophobicities.^[99]

Hydrogels containing different LCST oligomer grafts show dual transitions with increasing temperature. Comb-type grafted hydrogels of crosslinked poly(NIPAAm), grafted with oligoNIPAAm^[100] or PEG,^[101] exhibited fast response to temperature changes. It was suggested that the grafted short chains of oligoNIPAAm contribute to fast dehydration, whereas the hydrophilic PEG chains provides a water channel for a fast deswelling mechanism.

Li et al.^[102] synthesized biocompatible, thermo-responsive ABA triblock copolymers where the A blocks were poly(NIPAAm) and the B blocks were poly(2-methacryloyloxyethyl phosphorylcholine) via atom transfer radical polymerization with a commercially available bifunctional initiator. These novel triblock copolymers are water-soluble in dilute aqueous solution at 20°C and pH 7.4, but form free-standing physical gels at 37°C. This gelation is reversible, and the gels are believed to contain nanosized micellar domains. It was suggested that the less viscous copolymer gels were worthy for tissue engineering applications such as cartilage regeneration. Au et al.^[103] studied the thermally reversible polymer gels poly(NIPAAm-*co*-AAc), as a scaffold for chondrocytic cells. The results indicated that chondrocytes propagated in the thermo-reversible polymers expressed enhanced or maintained expression of collagen type II and aggrecan. The results support that *N*-isopropylacrylamide and acrylic acid copolymer gel has potential use as a cell culture substrate and as a cell delivery vehicle. Cho et al.^[104] demonstrated that human mesenchymal stem cells (MSCs) could differentiate into chondrocytes in a thermosensitive chitosan-poly(NIPAAm) copolymer gel. They showed that the thermosensitive gel transition was optimal for in vivo applications because the LCST was 32°C. Thus, chondrogenic differentiated cells from MSCs with a thermosensitive polymer could be used as an injectable cell-polymer complex and the differentiation process was induced both in vitro and in vivo.

Besides temperature and pH responsive materials, bioresponsive polymers are developed in drug delivery and tissue engineering applications. Wang et al.^[105] constructed a bioresponsive, intelligent hydrogel based on a phosphoester-poly(ethylene glycol) polymer (PhosPEG) for application to marrow-derived mesenchymal stem cell (MSC) encapsulation and tissue engineering of bone. The degradation of the scaffold and the extend of calcification were controlled by alkaline phosphatase levels. It was revealed that phosphate-PEG-derived hydrogels, containing an intermediate concentration of phosphorus, promoted gene expression of bone-specific markers, secretion of

bone-related matrix, and mineralization and may have a potential impact on bone-engineering therapies.

Future Trends

Tissue engineering is in development along a number of avenues. One of these is nanotechnology, which is making a great impact. Through the use of nanotechnological approaches in the production of nanofibers or in the creation of nanomodified material surfaces, we are able to change the cell–material interface significantly. Already a number of studies involving the influence of nanofiber dimensions, and incorporated compounds such as hydroxyapatite and growth factors, have been published. Creation and examination of nanopatterned surfaces expand our knowledge of the nano-world but also bring up new questions to answer. Nanotechnology is also providing us nano-sized drug delivery systems, which are especially important in the delivery of growth factors to guide tissue regeneration.

Self-assembled monolayers, constructed using a variety of molecules, is another novel approach to surface modification as a “bottom-up approach.” This leads to both surfaces modified at a molecular level, as well as to intelligent, responsive surfaces.

Developments in stem cell biology and technology are also very important. Researchers use cell biology approaches to direct the stem cell differentiation routes and gain the ability to create a number of different tissue types from a few cell-scaffold constructs.

Finally, with the contribution of molecular biology and cell biology, we are able to identify and even to quantify the tissue regeneration level, thus decreasing the dependence on such qualitative techniques as histology. For example, identification and quantification of mRNA is one important step in this. When we learn more about the roles of individual growth factors, we measure their impact by determining their amounts in the tissue, determine their products (cell number, differentiation markers, etc.) and learn about the extent of regeneration. With all these developments in their infancy the future of tissue engineering appears very exciting.

CONCLUSION

Research on biomaterials is showing great advances especially in two directions: tissue engineering and nanobiomaterials. The polymers used in tissue engineering are playing their roles in this. Researchers are developing new polymers or copolymers using synthetics, biologicals, and synthetic analogs of biologicals, all designed to meet the new requirements of

tissue engineering such as responsiveness. Thus, this chapter could not be considered a complete review of the field; but rather just a glimpse to this rapidly expanding research area.

ACKNOWLEDGMENTS

We gratefully acknowledge the contributions of Pinar Zorlutuna (Fig. 3), Engin Vrana (Fig. 5), and Albana Ndreu (Fig. 7), who provided micrographs from their thesis studies.

ARTICLES OF FURTHER INTEREST

Biodegradable Polymers, An Overview; Biomaterials; Biomimetic Materials; Protein–Surface Interactions; Tissue Engineering



REFERENCES

1. Brady, J.E.; Holum, J.R. Organic compounds, polymers, and biochemicals, In *Chemistry: The Study of Matter and its Changes*, 2nd Ed.; Wiley: Canada, 1996; 914.
2. Niklason, L.E.; Gao, J.; Abbott, W.M.; Hirschi, K.K.; Houser, S.; Marini, R.; Langer, R. Functional arteries grown in vitro. *Science* **1999**, *284* (5413), 489–493.
3. Pomahac, B.; Svensjo, T.; Yao, F.; Brown, H.; Eriksson, E. Tissue engineering of skin. *Crit. Rev. Oral Biol. Med.* **1998**, *9* (3), 333–344.
4. Ma, P.X.; Langer, R. Morphology and mechanical function of long-term in vitro engineered cartilage. *J. Biomed. Mater. Res.* **1999**, *44*, 217–221.
5. Service, R.F. Tissue engineers build new bone. *Science* **2000**, *289* (5484), 1498–1500.
6. Lin, V.S.; Lee, M.C.; O’Neal, S.; McKean, J.; Sung, K.L.P. Ligament tissue engineering using synthetic biodegradable fiber scaffolds. *Tissue Eng.* **1999**, *5*, 443–452.
7. Williams, D.F. *The Williams Dictionary of Biomaterials*; University of Liverpool Press: Liverpool, UK, 1999; 318.
8. Skalak, R.; Fox, C.F. Tissue Engineering. Proceedings of a workshop held at Granlibakken, Lake Tahoe, CA, February 26–29, 1988; Alan Liss New York.
9. Langer, R.; Vacanti, J.P. Tissue engineering. *Science* **1993**, *260* (5110), 920–926.
10. Griffith, L.G. Emerging design principles in biomaterials and scaffolds for tissue engineering. *Ann. N.Y. Acad. Sci.* **2002**, *961*, 83–95.
11. Curtis, A.; Riehle, M. Tissue engineering: The biophysical background. *Phys. Med. Biol.* **2001**, *46*, 47–65.
12. Johansson, C.B.; Momma, S.; Clarke, D.L.; Risling, M.; Lendahl, U.; Frisen, J. Identification of a neural

- stem cell in the adult mammalian central nervous system. *Cell* **1999**, *96*, 25–34.
13. Mooney, D.; Hansen, L.; Vacanti, J.; Langer, R.; Farmer, S.; Ingber, D. Switching from differentiation to growth in hepatocytes: control by extracellular matrix. *J. Cell. Physiol.* **1992**, *151*, 497–505.
 14. Ruoslahti, E.; Pierschbacher, M.D. New perspectives in cell adhesion: RGD and integrins. *Science* **1987**, *238* (4826), 491–497.
 15. Griffith, L.G. Polymeric biomaterials. *Acta Mater.* **2000**, *48*, 263–277.
 16. Rouslahti, E. RGD and other recognition sequences for integrins. *Annu. Rev. Cell Dev. Biol.* **1996**, *12*, 697–715.
 17. Clark, E.A.; Brugge, J.S. Integrins and signal transduction pathways: the road taken. *Science* **1995**, *268* (5208), 233–239.
 18. Badylak, S.F. The extracellular matrix as a scaffold for tissue reconstruction. *Semin. Cell Dev. Biol.* **2002**, *13*, 377–383.
 19. Van Wachem, P.B.; Brouwer, L.A.; Zeeman, R.; Dijkstra, P.J.; Feijen, J.; Hendriks, M.; Cahalan, P.T.; van Luyn, M.J.A. Tissue reactions to epoxy-cross-linked porcine heart valves post-treated with detergents or a dicarboxylic acid. *J. Biomed. Mater. Res.* **2001**, *55* (3), 415–423.
 20. Chen, G.; Ushida, T.; Tateishi, T. Scaffold design for tissue engineering. *Macromol. Biosci.* **2002**, *2* (2), 67–77.
 21. Yannas, I.V.; Burke, J.F. Design of an artificial skin: basic design principles. *J. Biomed. Mater. Res.* **1980**, *14*, 65–81.
 22. Bonassar, L.J.; Vacanti, C.A. Tissue engineering: the first decade and beyond. *J. Cell. Biochem. Suppl.* **1998**, *30/31*, 297–303.
 23. Katz, J. Developments in Medical Polymers for Biomaterials Applications. *Medical Device & Diagnostic Industry Magazine*, January 2001.
 24. Middleton, J.C.; Tipton, A.J. Synthetic Biodegradable Polymers as Medical Devices. *Medical Plastics and Biomaterials Magazine (Materials)*, March 1998.
 25. Pachence, J.M.; Kohn, J. Biodegradable polymers for tissue engineering. In *Principles of Tissue Engineering*; Lanza, R.P., Langer, R., Chick, W.L., Eds.; Texas, USA, 1997; 273.
 26. Agrawal, C.M.; Ray, R.B. Biodegradable polymeric scaffolds for musculoskeletal tissue engineering. *J. Biomed. Mater. Res.* **2001**, *55*, 141–150.
 27. Gilding, D.K.; Reed, A.M. Biodegradable polymers for use in surgery-poly(glycolic)/poly(lactic acid) homo and copolymers. *Polymer* **1979**, *20*, 1459–1464.
 28. Reed, A.M.; Gilding, D.K. Biodegradable polymers for use in surgery- poly(glycolic)/poly(lactic acid) homo and copolymers: 2. In vitro degradation. *Polymer* **1981**, *22*, 494–498.
 29. Gilding, D.K. Biocompatibility of clinical implant materials. In *Williams, D.F., Ed.; Biodegradable Polymers*; CRC Press: Boca Raton, FL, 1981; Vol. II, 209–232.
 30. Bostman, O.M.; Pihlajamaki, H.K. Adverse tissue reactions to bioabsorbable fixation devices. *Clin. Orthop. Relat. Res.* **2000**, *371*, 216–217.
 31. Niklason, L.F.; Langer, R.S. Advances in tissue engineering of blood vessels and other tissues. *Transpl. Immunol.* **1997**, *5*, 303–306.
 32. Shi, G.; Cai, Q.; Wang, C.; Lu, N.; Wang, S.; Bei, J. Fabrication and biocompatibility of cell scaffolds of poly (L-lactic acid) and poly (L-lactic-co-glycolic acid). *Polym. Adv. Tech.* **2002**, *13*, 227–232.
 33. Hasirci, V.; Berthiaume, F.; Bondre, S.P.; Gresser, J.D.; Trantolo, D.J.; Toner, M.; Wise, D.L. Expression of liver-specific functions by rat hepatocytes seeded in treated poly(lactic-co-glycolic) acid biodegradable foams. *Tissue Eng.* **2001**, *7* (4), 385–394.
 34. Woodward, S.C.; Brewer, P.S.; Moatamed, F.; Schindler, A.; Pitt, C.G. The intracellular degradation of poly(epsilon-caprolactone). *J. Biomed. Mater. Res.* **1985**, *19*, 437–444.
 35. Williams, J.M.; Adewunmi, A.; Schek, R.M.; Flanagan, C.L.; Krebsbach, P.H.; Feinberg, S.E.; Hollister, S.J.; Das, S. Bone tissue engineering using polycaprolactone scaffolds fabricated via selective laser sintering. *Biomaterials* **2005**, *26*, 4817–4827.
 36. Cao, T.; Ho, K.-H.; Teoh, S.-H. Scaffold design and in vitro study of osteochondral coculture in a three-dimensional porous polycaprolactone scaffold fabricated by fused deposition modeling. *Tissue Eng.* **2003**, *9* (1), 103–112.
 37. Katti, D.S.; Lakshmi, S.; Langer, R.; Laurencin, C.T. Toxicity, biodegradation and elimination of poly-anhydrides. *Adv. Drug Deliv. Rev.* **2002**, *54*, 933–961.
 38. Ibim, S.E.; Uhrich, K.E.; Attawia, M.; Shastri, V.R.; El-Amin, S.F.; Bronson, R.; Langer, R.; Laurencin, C.T. Preliminary in vivo report on the osteocompatibility of poly(anhydride-co-imides) evaluated in a tibial model. *J. Biomed. Mater. Res.* **1998**, *43*, 374–379.
 39. Ibim, S.M.; Uhrich, K.E.; Bronson, R.; El-Amin, S.F.; Langer, R.S.; Laurencin, C.T. Poly(anhydride-co-imides): In vivo biocompatibility in a rat model. *Biomaterials* **1998**, *19*, 941–951.
 40. Peter, S.J.; Miller, S.T.; Zhu, G.; Yasko, G.; Mikos, A.G. In vivo degradation of a poly(propylene fumarate)/beta-tricalcium phosphate injectable scaffold. *J. Biomed. Mater. Res.* **1998**, *41*, 1–7.
 41. Peter, S.J.; Lu, L.; Kim, D.J.; Stamatias, G.N.; Miller, M.J.; Yaszemski, M.J.; Mikos, A.G. Effects of transforming growth factor b1 released from biodegradable polymer microparticles on marrow stromal osteoblasts cultured on poly(propylene fumarate) substrates. *J. Biomed. Mater. Res.* **2000**, *50*, 452–462.
 42. Payne, R.G.; McGonigle, J.S.; Yaszemski, M.J.; Yasko, A.W.; Mikos, A.G. Development of an injectable, in situ crosslinkable, degradable polymeric carrier for osteogenic cell populations. Part 3. Proliferation and differentiation of encapsulated marrow stromal osteoblasts cultured on crosslinking poly(propylene fumarate). *Biomaterials* **2002**, *23* (22), 4381–4387.

43. Muggli, D.S.; Burkoth, A.K.; Keyser, S.A.; Lee, H.R.; Anseth, K.S. Reaction behaviour of biodegradable, photocross-linkable polyanhydrides. *Macromolecules* **1998**, *31*, 4120–4125.
44. Betre, H.; Setton, L.A.; Meyer, D.E.; Chilkoti, A. Characterization of a genetically engineered elastin-like polypeptide for cartilaginous tissue repair. *Bio-macromolecules* **2002**, *3* (5), 910–916.
45. Jeschke, B.; Meyer, J.; Jonczyk, A.; Kessler, H.; Adamietz, P.; Meenen, N.M.; Kantlehner, M.; Goepfert, C.; Nies, B. RGD-peptides for tissue engineering of articular cartilage. *Biomaterials* **2002**, *23* (16), 3455–3463.
46. Holmes, T.C.; Lacalle, S.; Su, X.; Liu, G.; Rich, A.; Zhang, S. Extensive neurite outgrowth and active synapse formation on self-assembling peptide scaffolds. *PNAS* **2000**, *97* (12), 6728–6733.
47. Cho, N.; Heller, J. Drug delivery devices manufactured from polyorthoesters and polyorthocarbonates. US Patent 4093709, 1978.
48. Heller, J.; Penhale, D.; Helwing, R. Preparation of poly(ortho esters) by the reaction of kicetene acetals and polyols. *J. Polym. Sci.* **1980**, *18*, 619–624.
49. Guo, Q.; Slavov, S.; Halley, P.J. Phase behavior, crystallization, and morphology in thermosetting blends of a biodegradable poly(ethylene glycol)-type epoxy resin and poly(ϵ caprolactone). *J. Polym. Sci. Part B: Polym. Phys.* **2004**, *42*, 2833–2843.
50. Deschamps, A. Transactions of the Sixth World Biomaterials Congress, Society for Biomaterials: Minneapolis, 2000; 364
51. Drury, J.L.; Mooney, D.J. Hydrogels for tissue engineering: scaffold design variables and applications. *Biomaterials* **2003**, *24*, 4337–4351.
52. Jeong, B.; Bae, Y.H.; Lee, D.S.; Kim, S.W. Biodegradable block copolymers as injectable drug-delivery systems. *Nature* **1997**, *388*, 860–862.
53. Huh, K.M.; Bae, Y.H. Synthesis and characterization of poly(ethylene glycol)/poly(L-lactic acid) alternating multiblock copolymers. *Polymer* **1999**, *40*, 6147–6155.
54. Metters, A.T.; Anseth, K.S.; Bowman, C.N. Fundamental studies of a novel, biodegradable PEG-*b*-PLA hydrogel. *Polymer* **2000**, *41*, 3993–4004.
55. Suggs, L.J.; Mikos, A.G. Development of poly(propylene fumarate-co-ethylene glycol) as an injectable carrier for endothelial cells. *Cell Trans.* **1999**, *8*, 345–350.
56. Rosso, F.; Marino, G.; Giordano, A.; Barbarisi, M.; Parmeggiani, D.; Barbarisi, A. Smart materials as scaffolds for tissue engineering. *J. Cell. Physiol.* **2005**, *203*, 465–470.
57. Park, S-N.; Park, J-C.; Kim, H.O.; Song, M.J.; Suh, H. Characterization of porous collagen/hyaluronic acid scaffold modified by 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide cross-linking. *Biomaterials* **2002**, *23*, 1205–1212.
58. Tan, W.; Krishnaraj, R.; Desai, T.A. Evaluation of nanostructured composite collagen–chitosan matrices for tissue engineering. *Tissue Eng.* **2001**, *7*, 203–210.
59. Huang, L.; Nagapudi, K.; Apkarian, R.P.; Chaikof, E.L. Engineered collagen–PEO nanofibers and fabrics. *J. Biomater. Sci.: Polym. Ed.* **2001**, *12*, 979–993.
60. Wakitani, S.; Goto, T.; Young, R.; Mansour, J.; Goldberg, V.; Caplan, A. Repair of large full-thickness articular cartilage defects with allograft articular chondrocytes embedded in a collagen gel. *Tissue Eng.* **1998**, *4*, 429–444.
61. Orwin, E.; Hubel, A. In vitro culture characteristics of corneal epithelial, endothelial, and keratocyte cells in a native collagen matrix. *Tissue Eng.* **2000**, *6*, 307–319.
62. Ber, S.; Torun Kose, G.; Hasirci, V. Bone tissue engineering on patterned collagen films: an in vitro study. *Biomaterials* **2005**, *26*, 1977–1986.
63. Valluzzi, R.; Winkler, S.; Wilson, D.; Kaplan, D.L. Silk: Molecular organization and control of assembly. *Phil. Trans. R. Soc. Lond. B* **2002**, *357*, 165–167.
64. Altman, G.H.; Diaz, F.; Jakuba, C.; Calabro, T.; Horan, R.L.; Chen, J.; Lu, H.; Richmond, J.; Kaplan, D.L. Silk-based biomaterials. *Biomaterials* **2003**, *24*, 401–416.
65. Minoura, N.; Aiba, S.; Gotoh, Y.; Tsukada, M.; Imai, Y. Attachment and growth of cultured fibroblast cells on silk protein matrices. *J. Biomed. Mater. Res.* **1995**, *29*, 1215–1221.
66. Sofia, S.; McCarthy, M.B.; Gronowicz, G.; Kaplan, D.L. Functionalized silk-based biomaterials for bone formation. *J. Biomed. Mater. Res.* **2001**, *54*, 139–148.
67. Unger, R.E.; Wolf, M.; Peters, K.; Motta, A.; Migliaresi, C.; Kirkpatrick, C.J. Growth of human cells on a non-woven silk fibroin net: a potential for use in tissue engineering. *Biomaterials* **2004**, *25*, 1069–1075.
68. Walton, A.G.; Blackwell, J. Fibrous proteins and biopolymer models. In *Biopolymers*; Academic Press Inc.: New York, USA, 1973; p. 456.
69. Ye, Q.; Zund, G.; Benedikt, P.; Jockenhoevel, S.; Hoerstrup, S.P.; Sakyama, S.; Hubbell, J.A.; Turina, M. Fibrin gel as a three dimensional matrix in cardiovascular tissue engineering. *Eur. J. Cardiothorac. Surg.* **2000**, *17*, 587–591.
70. Pellegrini, G.; Ranno, R.; Stracuzzi, G.; Bondanza, S.; Guerra, L.; Zambruno, G.; Micali, G.; De Luca, M. The control of epidermal stem cells (holoclones) in the treatment of massive full-thickness burns with autologous keratinocytes cultured on fibrin. *Transplantation* **1999**, *68* (6), 868–879.
71. Suh, J.K.F.; Matthew, H.W.T. Application of chitosan-based polysaccharide biomaterials in cartilage tissue engineering: a review. *Biomaterials* **2000**, *21*, 2589–2598.
72. Madhally, S.V.; Matthew, H.W. Porous chitosan scaffolds for tissue engineering. *Biomaterials* **1999**, *20* (12), 1133–1142.
73. Tuzlakoglu, K.; Catarina, M.; Alves, C.M.; Mano, J.F.; Reis, R.L. Production and characterization of chitosan fibers and 3-D fiber mesh scaffolds for tissue engineering applications. *Macromol. Biosci.* **2004**, *4*, 811–819.
74. Zhao, F.; Yin, Y.; Lu, W.W.; Leong, J.C.; Zhang, W.; Zhang, J.; Zhang, M.; Yao, K. Preparation and histological evaluation of biomimetic three-



- dimensional hydroxyapatite/chitosan-gelatin network composite scaffolds. *Biomaterials* **2002**, *23*, 3227–3234.
75. Zhu, H.; Ji, J.; Lin, R.; Gao, C.; Feng, L.; Shen, J. Surface engineering of poly(D,L-lactic acid) by entrapment of chitosan-based derivatives for the promotion of chondrogenesis. *J. Biomed. Mater. Res.* **2002**, *62*, 532–539.
 76. Zavan, B.; Brun, P.; Vindigni, V.; Amadori, A.; Habeler, W.; Pontisso, P.; Montemurro, D.; Abatangelo, G.; Cortivo, R. Extracellular matrix-enriched polymeric scaffolds as a substrate for hepatocyte cultures: in vitro and in vivo studies. *Biomaterials* **2005**, *26* (34), 7038–7045.
 77. Aigner, J.; Tegeler, J.; Hutzler, P.; Campoccia, D.; Pavesio, A.; Hammer, C.; Kastenbauer, E.; Naumann, A. Cartilage tissue engineering with novel nonwoven structured biomaterial based on hyaluronic acid benzyl ester. *J. Biomed. Mater. Res.* **1998**, *42*, 172–181.
 78. Kolybaba, M.; Tabil, L.G.; Panigrahi, S.; Crerar, W.J.; Powell, T.; Wang, B. Biodegradable polymers: past, present, and future ASAE/CSAE North Central Intersectional Meeting, Fargo, North Dakota, October 3–4, 2003 ASAE Paper No. RRV03-0007.
 79. Jopski, T. Biodegradable plastics. *Starch* **1993**, *83* (10), 17–20.
 80. Salgado, A.J.; Gomes, M.E.; Chou, A.; Coutinho, O.P.; Reis, R.L.; Huttmacher, D.W. Preliminary study on the adhesion and proliferation of human osteoblasts on starch-based scaffolds. *Mater. Sci. Eng. C* **2002**, *20*, 27–33.
 81. Li, Z.; Ramay, H.R.; Hauch, K.D.; Xiao, D.; Zhang, M. Chitosan–alginate hybrid scaffolds for bone tissue engineering. *Biomaterials* **2005**, *26*, 3919–3928.
 82. Williams, S.F.; Martin, D.P.; Horowitz, D.M.; People, O.P. PHA applications: Addressing the price performance issue I. Tissue engineering. *Int. J. Biol. Macromol.* **1999**, *25*, 111–121.
 83. Zinn, M.; Witholt, B.; Egli, T. Occurrence, synthesis and medical application of bacterial polyhydroxyalkanoate. *Adv. Drug Deliv. Rev.* **2001**, *53*, 5–21.
 84. Adams, J.H.; Irving, G.; Koeslag, J.H.; Lochner, J.D.; Sandell, R.C.; Wilkinson, C. Beta-adrenergic blockade restores glucose's antiketogenic activity after exercise in carbohydrate-depleted athletes. *J. Physiol. Lond.* **1987**, *386*, 439–454.
 85. Torun Kose, G.; Kenar, H.; Hasirci, N.; Hasirci, V. Macroporous poly(3-hydroxybutyrate-co-3-hydroxyvalerate) matrices for bone tissue engineering. *Biomaterials* **2003**, *24* (11), 1949–1958.
 86. Torun Kose, G.; Korkusuz, F.; Korkusuz, P.; Purali, N.; Ozkul, A.; Hasirci, V. Bone generation on PHBV matrices: An in vitro study. *Biomaterials* **2003**, *24*, 4999–5007.
 87. Torun Kose, G.; Korkusuz, F.; Ozkul, A.; Soysal, Y.; Ozdemir, T.; Yildiz, C.; Hasirci, V. Tissue engineered cartilage on collagen and PHBV matrices. *Biomaterials* **2005**, *26*, 5187–5197.
 88. Sun, J.; Wu, J.; Li, H.; Chang, J. Macroporous poly(3-hydroxybutyrate-co-3-hydroxyvalerate) matrices for cartilage tissue engineering. *Eur. Poly. J.* **2005**, *41* (10), 2443–2449.
 89. Cho, C.S.; Seo, S.J.; Park, I.K.; Kim, S.H.; Kim, T.H.; Hoshiba, T.; Harada, I.; Akaike, T. Galactose-carrying polymers as extracellular matrices for liver tissue engineering. *Biomaterials* **2006**, *27* (4), 576–585.
 90. Grzesiak, J.J.; Pierschbacher, M.D.; Amodeo, M.F.; Malaney, T.I.; Glass, J.R. Enhancement of cell interactions with collagen/glycosaminoglycan matrices by RGD derivatization. *Biomaterials* **1997**, *18*, 1625–1632.
 91. Santiago, L.Y.; Nowak, R.W.; Peter Rubin, J.; Marra, K.G. Peptide-surface modification of poly(caprolactone) with laminin-derived sequences for adipose-derived stem cell applications. *Biomaterials* **2006**, *27* (15), 2962–2969.
 92. He, W.; Ma, Z.; Yong, T.; Teo, W.E.; Ramakrishna, S. Fabrication of collagen-coated biodegradable polymer nanofiber mesh and its potential for endothelial cells growth. *Biomaterials* **2005**, *26* (36), 7606–7615.
 93. Tezcaner, A.; Bugra, K.; Hasirci, V. Retinal pigment epithelium cell culture on surface modified poly(hydroxybutyrate-co-hydroxyvalerate) thin films. *Biomaterials* **2003**, *24*, 4573–4583.
 94. Roy, I.; Gupta, M.N. Smart polymeric materials: emerging biochemical applications. *Chem. Biol.* **2003**, *10*, 1161–1171.
 95. Gupta, M.N.; Roy, I. Applied biocatalysis: An overview. *Indian J. Biochem. Biophys.* **2002**, *39*, 220–228.
 96. Ganorkar, C.R.; Liu, F.; Baudy, M.; Kim, S.W. Modulating insulin-release profile from pH/thermosensitive polymeric beads through polymer molecular weight. *J. Control. Release* **1999**, *59*, 287–298.
 97. Shimizu, T.; Yamato, M.; Kikuchi, A.; Okano, T. Two-dimensional manipulation of cardiac myocyte sheets utilizing temperature-responsive culture dishes augments the pulsatile amplitude. *Tissue Eng.* **2001**, *7*, 141–151.
 98. Jeong, B.; Kim, S.W.; Bae, Y.H. Thermosensitive sol-gel reversible hydrogels. *Adv. Drug Deliv. Rev.* **2002**, *54*, 37–51.
 99. Yoshida, R.; Sakai, Y.; Okano, T.; Sakurai, Y. Modulating the phase transition temperature and thermosensitivity in N-isopropylacrylamide copolymer gels. *J. Biomater. Sci. Polym. Ed.* **1994**, *6*, 585–598.
 100. Yoshida, R.; Uchida, K.; Kaneko, Y.; Sakai, K.; Kiruchi, A.; Sakurai, Y.; Okano, T. Comb-type grafted hydrogels with rapid deswelling response to temperature changes. *Nature* **1995**, *374*, 240–242.
 101. Kaneko, Y. Rapid deswelling response of poly(N-isopropylacrylamide) hydrogels by the formation of water release channels using poly(ethylene oxide) graft chains. *Macromolecules* **1998**, *31*, 6099–6105.
 102. Li, C.; Tang, Y.; Armes, S.P. Synthesis and characterization of biocompatible thermo-responsive gelators based on ABA triblock copolymers. *Biomacromolecules* **2005**, *6*, 994–999.
 103. Au, A.; Ha, J.; Polotsky, A.; Krzyminski, K.; Gutowska, A.; Hungerford, D.S.; Frondoza, C.G. Thermally reversible polymer gel for chondrocyte culture. *J. Biomed. Mater. Res.* **2003**, *67A*, 1310–1319.

104. Cho, J.H.; Kim, S.-H.; Park, K.D.; Jung, M.C.; Yang, W.I.; Han, S.W.; Noh, J.Y.; Lee, J.W. Chondrogenic differentiation of human mesenchymal stem cells using a thermosensitive poly(*N*-isopropylacrylamide) and water-soluble chitosan copolymer. *Biomaterials* **2004**, *25*, 5743–5751.
105. Wang, D.A.; Williams, C.G.; Yang, F.; Cher, N.; Lee, H.; Elisseeff, J.H. Bioresponsive phosphoester hydrogels for bone tissue engineering. *Tissue Eng.* **2005**, *11* (1–2), 201–213.
106. Jung, Y.; Kim, S.-S.; Kim, Y.H.; Kim, S.H.; Kim, B.-S.; Kim, S.; Choi, C.Y.; Kim, S.H. A poly(lactic acid)/calcium metaphosphate composite for bone tissue engineering. *Biomaterials* **2005**, *26*, 6314–6322.
107. Ma, P.X.; Zhang, R.; Xiao, G.; Franceschi, R. Engineering new bone tissue in vitro on highly porous poly(-hydroxyl acids)/hydroxyapatite composite scaffolds. *J. Biomed. Mater. Res.* **2000**, *54* (2), 284–293.
108. Alsberg, E.; Anderson, K.W.; Albeiruti, A.; Franceschi, R.T.; Mooney, D.J. Cell-interactive alginate hydrogels for bone tissue engineering. *J. Dent. Res.* **2001**, *80*, 2025–2029.
109. Torun Kose, G.; Korkusuz, F.; Korkusuz, P.; Hasirci, V. In vivo tissue engineering of bone using poly(3-hydroxybutyric acid-*co*-3-hydroxyvaleric acid) and collagen scaffolds. *Tissue Eng.* **2004**, *10* (7), 1234–1250.
110. Gomes, M.E.; Sikavitsas, V.I.; Behraves, E.; Reis, R.L.; Mikos, A.G. Effect of flow perfusion on the osteogenic differentiation of bone marrow stromal cells cultured on starch-based three-dimensional scaffolds. *J. Biomed. Mater. Res.* **2003**, *67A*, 87–95.
111. Li, W.J.; Tuli, R.; Okafor, C.; Derfoul, A.; Danielson, K.G.; Hall, D.J.; Tuan, R.S. A three-dimensional nano fibrous scaffold for cartilage tissue engineering using human mesenchymal stem cells. *Biomaterials* **2005**, *26*, 599–609.
112. Chen, G.; Sato, T.; Ohgushi, H.; Ushida, T.; Tateishi, T.; Tanaka, J. Culturing of skin fibroblasts in a thin PLGA-collagen hybrid mesh. *Biomaterials* **2005**, *26*, 2559–2566.
113. Besseau, L.; Coulomb, B.; Lebreton-Decoster, C.; Giraud-Guille, M.M. Production of ordered collagen matrices for three-dimensional cell culture. *Biomaterials* **2002**, *23*, 27–36.
114. Wen, S.J.; Zhao, L.M.; Li, P.; Li, J.X.; Liu, Y.; Liu, J.L.; Chen, Y.C. Blood vessel tissue engineering: Seeding vascular smooth muscle cells and endothelial cells sequentially on biodegradable scaffold in vitro. *Zhonghua Yi Xue Za Zhi.* **2005**, *85* (12), 816–818.
115. Dvir-Ginzberg, M.; Gamlieli-Bonshtein, I.; Agbaria, R.; Cohen, S. Liver tissue engineering within alginate scaffolds: Effects of cell-seeding density on hepatocyte viability, morphology, and function. *Tissue Eng.* **2003**, *9* (4), 757–766.

Polytetrafluoroethylene

Robert R. Pavlis

Department of Chemistry, Pittsburg State University, Pittsburg, Kansas, U.S.A.

HISTORY

Tetrafluoroethylene is a gas having a molar mass of 100.016 g/mol, a boiling point of -76.3°C , and a melting point of -142.5°C . The critical temperature is -33.3°C , and the critical pressure is 39.52 atmospheres. Tetrafluoroethylene will readily undergo radical polymerization, forming an extremely important polymer that is the topic of this work.

Fig. 1 shows the chemical structure of polytetrafluoroethylene and represents its actual conformation in crystal structures. The dark spheres represent carbon atoms; the white spheres represent fluorine atoms.

Like many important discoveries, the discovery of the polymerization of tetrafluoroethylene was an accident. In 1936, one of DuPont's research chemists, Roy J. Plunkett, was engaged in research to develop new refrigerants. He and his colleagues produced many kilograms of tetrafluoroethylene, which they stored in cylinders which they kept packed in dry ice.

On April 6, Dr. Plunkett's assistant, Jack Rebock, attempted to get some of the material from one of the storage cylinders when he found there was no pressure in the cylinder. They weighed the cylinder and discovered that there was still material inside. When they removed the cylinder valve, they discovered that the cylinder was filled with a white powder that was insoluble in every common laboratory solvent that they tried. It was also remarkably unreactive, showing no reactivity toward strong acids and oxidants. The material was also remarkably thermally stable.

One of its peculiar characteristics that the group immediately seemed to recognize as important was that it was amazingly slippery.

The next year they applied for a patent, which was granted in 1941. During World War II, PTFE was used for several military applications. Its extraordinary chemical inertness resulted in its being used for lining containers for the uranium hexafluoride diffusion process used to separate uranium isotopes.

The "Teflon" trademark was registered by DuPont in 1945, and commercial products using it appeared the following year.

Dr. Plunkett worked on tetraethyl lead for most of the rest of his career. He retired from Dupont in 1975. He died in 1994.^[1]

CHEMICAL PROPERTIES

Polytetrafluoroethylene owes its remarkable chemical and thermal stability to several different factors: (1) The C–F bond strength for secondary carbon atoms is about 450 kJ/mol, far higher than the C–H bond strength. (2) The C–C bond strength in fluorocarbons is unusually high, while the C=C bond strength in fluoroalkenes is fairly low, creating an enthalpy of polymerization of about -163 kJ/mol and making this an extremely exothermic polymerization. (The enthalpies of polymerization of most other alkenes are about half this value; styrene, for example, is 70 kJ/mol. Ethylene is above average at 109 kJ/mol.) (3) Because both the C–F and the C–C bonds in PTFE are so strong, almost any reaction that would convert this material into something else requires a substantial input of energy to disrupt these strong bonds.

The extraordinarily high enthalpy of polymerization results in an extremely high ceiling temperature—well above 1100°C ! The thermodynamic values at 25°C are ($\Delta H = -163$ kJ/mol, $\Delta S = 112$ J/deg-mol, and $\Delta G = -121$ kJ/mol).^[2] (ΔH is enthalpy change, ΔS is entropy change, and ΔG is the change in Gibbs Free Energy; $\Delta G = \Delta H - T\Delta S$. At the ceiling temperature, $\Delta G = 0.0$.)

Quantity	Value
ΔH	-163 kJ
ΔS	112 J/deg-mol
$\Delta G (25^{\circ}\text{C})$	-121 kJ

Polytetrafluoroethylene chains consist of nothing but perfluoromethylene groups. Because these are highly symmetrical, the material can crystallize, forming a highly solid state polymer with a very high crystal lattice energy. Its melting point is 327°C . Amorphous polymer has a density of 2.00 g/cm³, but 2.30 g/cm³ is the solid state density.^[3] PTFE is a peculiar material with several first-order transitions (19 , 30 , and 90°) and several second-order transitions (-90 , -30 , and 130). (When materials undergo a first order transition, there is an immediate change of a physical property; but when they undergo a second order transition, there is a change in the rate of change of a property.)

Tables can be obtained from Web sources that list the important physical properties of various forms of this

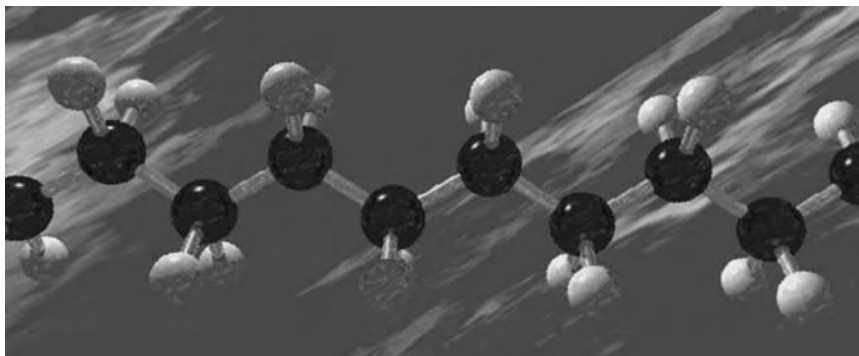


Fig. 1 Chemical structure of polytetrafluoroethylene

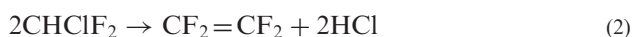
material. MatWeb Material Property Data has a particularly extensive set of tables of various commercial polytetrafluoroethylene products; in fact, they provide this data for over 100 different fluorinated polymers.^[4]

PRODUCTION OF THE MONOMER AND POLYMER

Tetrafluoroethylene is generally produced from chloroform. Two of its chlorine atoms are first exchanged by reacting it with a mixture of hydrogen fluoride and antimony pentafluoride, forming chlorodifluoromethane:



When chlorodifluoromethane is heated to about 950°C, it forms tetrafluoroethylene. The reaction mechanism for this process almost certainly involves first producing difluorocarbene, which immediately dimerizes to form tetrafluoroethylene:



The monomer is normally polymerized using emulsion polymerization under pressure in stainless steel vessels. Various water-soluble peroxides can be used as initiators for the radical polymerization process. It is, as mentioned above, highly exothermic. As the polymer forms, it separates. Because it is so extremely insoluble, it is easy to recover it completely:^[5]



PROBLEMS WITH TOXICITY

One of the most important commercial applications of this material is as a coating for cookware. Its thermal stability and its slippery nature make it possible to coat cooking surfaces with it, creating a “nonstick” coat.

Unfortunately, there have been some negative consequences to its use.

DuPont researchers and others have shown that high-temperature thermal degradation of PTFE in air produces a remarkably complex mixture of compounds that includes tetrafluoroethylene, hexafluoropropene, perfluorocyclobutane, perfluoroisobutene, carbonyl fluoride, perfluoromethane, trifluoroacetic acid, trifluoroacetyl fluoride, hydrogen fluoride, monofluoroacetic acid, and difluoroacetic acid, along with traces of other compounds. It also produces particulate matter. Several of these compounds are extraordinarily toxic—monofluoroacetic acid, for example, is far more toxic than hydrogen cyanide.

Polytetrafluoroethylene-coated cookware and household appliances are very popular. Occasionally, cooking accidents may result in these items being heated to temperatures at which at least some degradation begins to occur.

Birds have far more efficient respiratory systems than mammals.^[6] There have been many instances reported of birds being fatally injured by exposure to thermal degradation products from overheated PTFE. There have been similar problems in the poultry industry. When PTFE-coated heat lamps were used in poultry production,^[7] there were a massive number of deaths. These problems seem to occur when PTFE surfaces are heated to relatively low temperatures, perhaps only about 200°C. There is some evidence that the toxicity in birds is due to oligomers that are driven from the polymer mixture upon heating.^[8]

Because PTFE is produced by radical polymerization, typical samples contain a wide range of molecular weights. Fluorinated compounds tend to have unusually high vapor pressures compared to their molecular weights. The polymer molecules with lower molecular weights seem to distill out of the polymer mass when it is heated to relatively low temperatures. These molecules condense and produce a smoke with very fine particles. Strong heating may also cause rupture of a few of the polymer bonds, and thus produce even more polymer chains with low molecular weights.^[9]

Dr. Carl Johnson and coworkers^[10] reported that neither the smoke nor the other gases produced by heating PTFE are toxic enough to kill animals by themselves, but that the two together are extremely toxic. They suggested that toxic compounds become absorbed on the particles and are carried in this state to the lungs, where they cause severe tissue destruction. The particles need to be very small to be dangerous; those around 16 nm seem most toxic, and those larger than about 100 nm do not seem to be toxic to rats.

There have been many reports of human illness caused by breathing these fumes. One of the most famous incidents occurred when all the passengers and crew of a flight became ill. Not only did the passengers and crew become ill, but the four people who collected samples after the aircraft landed did as well. After ruling out other possible causes, the toxicity became linked to PTFE tape that was used to cover part of the exhaust manifold.^[11]

There have been many other incidences of what is often called "polymer fume fever." A paper by Zanen and Rietveld describes a case in which a woman became seriously ill after breathing fumes from a defective microwave oven.^[12]

Toxicity problems seem to become generally acute above the polymer melting point (327°C). In solid state PTFE, the high crystal lattice energy of the material reduces the vapor pressure and thus reduces the volatility of the lower-molecular-weight materials in the polymer.

It should be noted that pyrolysis of PTFE in vacuum results in its quantitative conversion to the monomer. It simply "unzips" to form it! The reactions that occur to form the complex and toxic mixture produced by overheating coated cookware only occur when the polymer is heated in air.^[13]

PERFLUOROCTANOIC ACID

In recent years, many news stories have appeared in various news media concerning possible environmental consequences that may result from the use of perfluorooctanoic acid, commonly called C-8 or PFOA.

Perfluorooctanoic acid is extremely stable, so its environmental life seems to be extremely long. It is used as a surfactant during the polymerization process. The U.S. Environmental Protection Agency Web site contains a risk assessment.^[14]

RELATED COMMERCIAL FLUORINATED POLYMERS

There are several other polymers similar to PTFE that have been developed and that are in current production. All are extremely chemically inert.

Polytetrafluoroethylene itself is a linear polymer with long chains of completely symmetrical difluoromethylene groups. This results in, as stated formerly, a high degree of crystallinity, a high crystal lattice energy, and an extremely high melting point. By introducing side chains or making copolymers, this symmetry can be reduced or destroyed.

Vinyl polymers with side chains can be isotactic, syndiotactic, or atactic. Atactic polymers are generally not solid state because they are not stereo regular. The polymers discussed below are produced in an atactic state, because the primary reason for using them rather than PTFE itself is to have a polymer without such a high degree of crystallinity.

Chlorotrifluoroethylene polymers have been produced commercially. The trade name of one such polymer is Kel-F 81. This material was formerly produced by 3M, but it was discontinued in the mid-1990s. Daikin currently produces this material under the trade name of Neoflon.

DuPont produces a copolymer of tetrafluoroethylene and ethylene. It is produced under the trade name Tefzel.

DuPont also produces a polymer called Teflon Fluoroethylene-propylene (FEP) that is a copolymer of hexafluoropropylene and tetrafluoroethylene. It can be processed using conventional thermoplastics processing equipment, because the trifluoromethyl side chains introduce asymmetric carbon atoms into the structure and thus prevent crystallization.

An essentially alternating 1:1 copolymer of ethylene and chlorotrifluoroethylene is produced by Ausimont under the trade name of Halar.

CONCLUSION

Polytetrafluoroethylene has become an extraordinarily important material since its discovery in 1936. Its chemical and thermal stability make it an ideal material for many applications.

Similar polymers with reduced stereo regularity are also used for many applications.

There has been concern in recent years about the toxic products that form when it is strongly overheated. This concern has been enhanced by the demonstrated toxicity toward birds. Many consumers try to avoid using products that use it. The author overheard a woman in a department store saying, "If these pans kill birds, they can't be good for my family either!"

Her concerns were probably an overreaction. Vast numbers of coated cookware has been produced, and the number of human illnesses reported is relatively small.

ARTICLES OF FURTHER INTEREST

Biomaterials: Protein–Surface Interactions; Polymers; Vascular Grafts; Vascular Grafts: Development Strategies; Vascular Grafts: Host Interactions

REFERENCES

1. <http://www.dupont.com/teflon/newsroom/history.html>.
2. Allcock, H.R.; Lampe, F.W. *Contemporary Polymer Chemistry*, 3rd Ed.; Prentice Hall: Englewood Cliffs, NJ, 1990; 284.
3. <http://www.polymerprocessing.com/polymers/PTFE.html>.
4. <http://www.matweb.com/>
5. Bahadur, P.; Sastry, N.V. *Principles of Polymer Science*; CRC Press: Boca Raton, FL, 2000.
6. Altman, R.; Chubb, S.; Dorrenstein, G.M.; Quesenberry, K. *Avian Medicine and Surgery*; W.B. Saunders, Inc.: Philadelphia, PA, 1997; 609.
7. Boucher, M.; Ehmler, T.J.; Bermudez, A.J. Polytetrafluoroethylene gas intoxication in broiler chickens. *Avian Dis.* **2000**, *44* (2), 449–453.
8. Seidel, W.C.; Scherer, K.V., Jr; Cline, D., Jr; et al. Chemical, physical, and toxicological characterization of fumes produced by heating tetrafluoroethene homopolymer and its copolymers with hexafluoropropene and perfluoro (propyl vinyl ether). *Chem. Res. Toxicol.* **1991**, *4* (2), 229–236.
9. Waritz, R.S. An industrial approach to evaluation of pyrolysis and combustion hazards. *Environ. Health Perspect.* **1975**, *11*, 197–202.
10. Johnston, C.J.; Finkelstein, J.N.; Mercer, P.; Corson, N.; Gelein, R.; Oberdorster, G. Pulmonary effects induced by ultrafine PTFE particles. *Toxicol. Appl. Pharmacol.* **2000**, *168* (3), 208–215.
11. Nuttall, J.B.; Kelly, R.J.; Smith, B.S.; Whiteside, C.K. *Aerospace Med.* **1964**, *July*.
12. Zanen, A.L.; Rietveld, A.P. Inhalation trauma due to overheating in a microwave oven. *Thorax* **1993**, *48* (3), 300–302.
13. Allcock, H.R.; Lampe, F.W. *Contemporary Polymer Chemistry*, 3rd Ed.; Prentice Hall: Englewood Cliffs, NJ, 1990; 173.
14. <http://www.epa.gov/opptintr/pfoa/>



Polyurethane Biomaterials

Joanna D. Fromstein
Kimberly A. Woodhouse

Department of Chemical Engineering and Applied Chemistry, University of Toronto,
Toronto, Ontario, Canada

INTRODUCTION

Polyurethanes (PUs), a family of polymers containing the urethane linkage, were first used in medicine in the late 1950s to construct a composite PU foam breast prosthesis. By the early 1960s, catheters and pacemaker lead insulators were also being made out of PU. The popularity of PUs for use in medical applications has risen steadily ever since. Polyurethanes are generally considered to have good biocompatibility, and PU chemistry is so varied that it enables fine-tuning of the chemical properties for any given application. Polyurethane copolymers or segmented PUs represent an important subclass of thermoplastic polymers that is frequently used in biomedical devices. These materials tend to be elastomeric, making them particularly well suited for soft tissue engineering applications, where a matrix that can withstand local mechanical forces is needed. The construction of a cardiac patch is one such example.

In this review, the structure and chemistry of PUs will be introduced, as well as a brief description of synthesis techniques. Next, the wide array of medical applications in which PUs are used is provided. Following this, the biocompatibility of these polymers is described, including the mechanisms of degradation and the general host response to PU-based materials. Finally, the bulk of this review pertains to the PUs currently in use and in development.

SEGMENTED PUs

Segmented PUs or PU copolymers are phase-segregated polymers composed of alternating polydispersed blocks of “soft” and “hard” segments. Almost all segmented PUs are synthesized using three components: a diisocyanate, a polyol, and a chain extender, usually a diamine or diol. The availability of a wide array of reactants enables fine-tuning of the physical and chemical properties of the PU for a given application.

Polyurethanes can be synthesized using either a one-step or a two-step process. Most biomedical PUs are

made using the two-stage condensation process (Fig. 1).^[1] This technique provides good control over polymer architecture.^[2] The first step of this method involves the reaction of an isocyanate with a polyol to form a prepolymer. This prepolymer is then reacted with a chain extender, usually a diamine or a diol, to create the final PU. The use of an amine chain extender will lead to the formation of urea bonds, whereas a diol chain extender produces urethane linkages. Table 1 details the chemical components of some common commercial biomedical PUs.

Unlike polyols and chain extenders, there are only a handful of isocyanates that are readily available for use in PU synthesis. Traditionally, methyl diisocyanate (MDI) and, to a lesser extent, toluene diisocyanate (TDI) have been used extensively in the biomedical PU industry.^[2,5] These aromatic isocyanates have been used because they result in materials with high Young's modulus, ultimate strength, and elasticity. Concern was raised, however, about the potential for these aromatic diisocyanates to degrade into suspected carcinogens methylene diamine (MDA) and toluene diamine (TDA).^[5] As a result, many medical PUs, especially those designed to be intentionally degradable, are now synthesized using aliphatic diisocyanates despite the poorer mechanical properties and relatively low risk presented by the TDA or MDA degradation products of implants.^[5] Some of the more common aliphatic isocyanates currently being used in medical grade PUs include: hexamethylene diisocyanate (HMDI), 4,4-methylene dicyclohexyl diisocyanate (H12MDI), lysine diisocyanate (LDI), and *trans* cyclohexane-1,4-diisocyanate (CHDI).^[2,3]

In the formation of biomedical PUs, soft segments are almost always diols. Commercially, polytetramethylene oxide (PTMO) and polycarbonate (PCN) diol are the most commonly used. These soft segments have been associated with improved long-term stability *in vivo*. Other popular soft segments used in research include: polyethylene oxide (PEO), poly(ϵ -caprolactone) (PCL), and polypropylene oxide (PPO).

The primary purpose of the chain extender is to extend the length of the hard segment and increase the hydrogen-bond density of the PU.^[3] Commercially

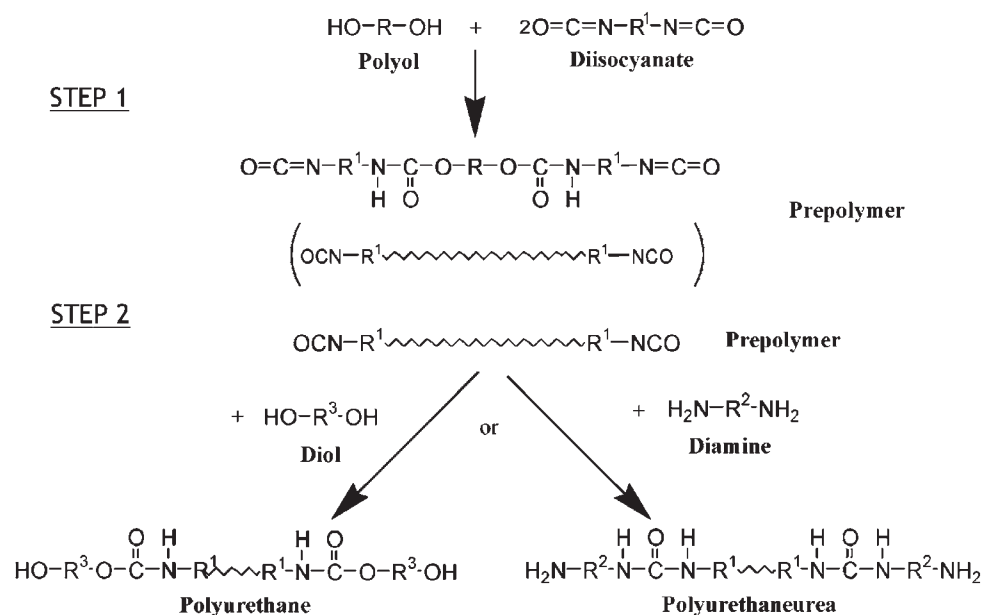


Fig. 1 The two-step PU synthesis process.

available medical PUs are generally made with 1,4-butanediol (1,4-BD) or ethylene diamine (ED) in this role. Ethylene glycol or oxide can be used as a chain extender as well; glycol-extended PUs tend to have greater flexibility but less strength than amine-extended analogs.^[3]

A segmented PU is made up of two distinct moieties within the polymer: a hard segment, comprised of the isocyanate and chain extender, and a soft segment containing the diol. These two segments are thermodynamically incompatible, and the extent of phase segregation can be varied by altering the molecular weights, chemistry, and relative percentages of the reactants.^[4] The segmented nature of elastomeric PUs are responsible for their mechanical properties^[2] with interchain hydrogen bonding playing an important role.^[1,2]

The phase-segregated morphology allows for re-arrangement of the polymer's surface depending on the

surrounding environment.^[1] In the case of a medical device or implant, this translates into an ability to change the protein and cellular response to the PU by controlling the surface morphology of the PU under physiological conditions.

MEDICAL APPLICATIONS OF PUs

Polyurethanes are the most extensively used synthetic polymers in biomedical applications.^[6] This popularity is a result of the versatility and recognized "biocompatibility" of PUs. Current and proposed uses of PUs in medical devices are summarized in Table 2. In addition to these applications, degradable PUs have tremendous potential for use in drug delivery and tissue-engineered constructs.

Table 1 Commercially available biomedical PUs

	Soft segment diol	Isocyanate	Chain extender	Manufacturer
Pellethane	PTMO	MDI	BD	DOW Chemical
Biomer	PTMO	MDI	ED	Ethicon
Corethane	HMEC	MDI	BD	Corvita Corp.
Tecoflex	PTMO	HMDI	BD	Thermedics Inc
Estane	Varied	Varied	Varied	BF Goodrich
Medadhere	Aliphatic	HMDI	Aliphatic	Medtronic

PTMO, polytetramethylene oxide; HMEC, hexamethylene ethylene carbonate; MDI, diphenylmethane diisocyanate; HMDI, hexamethylene diisocyanate; BD, butane diol; and ED, ethylene diamine.
(From Refs. [1,3,4].)

Table 2 Current biomedical uses of PUs

Catheters
Breast implants
Pacemaker lead insulators
Heart valves
Ventricular assist devices
Vascular grafts
Aortic grafts
Dialysis membranes
Intraortic balloons
Nerve guides
Meniscal reconstruction
Blood bags
Endotracheal tubes
Orthopaedic splints
Wound dressings
Sutures
Synthetic bile ducts
Blood pumps

(From Refs. [3,5,7–10].)

BIOCOMPATIBILITY AND TOXICITY

The adsorption of proteins from the blood or other biological fluids immediately following exposure is critical in determining the cellular interactions with the surface, including thrombus formation and bacterial infection.^[7] Bacterial infection represents one of the greatest risks in using biomaterials, with half of all hospital infections associated with implants and/or medical devices.^[9]

The term “biocompatibility” is used to assess the performance of a material in a given application. In the field of PUs, many investigators refer to PUs as simply being biocompatible. This is based on results that PUs perform relatively well *in vivo* under a wide variety of uses and that PUs generally have better blood and tissue compatibilities in comparison to numerous other synthetic polymers.

MECHANISMS OF PU BIODEGRADATION

It is widely acknowledged that traditional biomedical PU materials degrade in the biological environment.^[2] The ultimate biostability of PUs can be affected by the chemical structure, morphology, manufacturing, and processing conditions. The major mechanisms of degradation are hydrolysis and oxidation, both *in vitro* and *in vivo*.^[10] Environmental factors including biological fluids, proteins, cells, and mechanical forces, play an important role *in vivo* in the degradation of

PUs. Calcification can also play a role in the degradation of biomedical PUs.^[10]

Hydrolysis

The soft segment generally dominates the degradation characteristics of the PU and an increased soft segment content tends to be accompanied by higher degradation rates.^[5] Ninety percent of commercial PUs are polyester urethanes. These materials are commonly chosen for their low cost combined with high tensile strength, tear strength, and resilience (wear resistance). Unfortunately, polyester urethanes are highly susceptible to hydrolytic degradation.^[10] While polyether urethanes are less susceptible than their polyester counterparts, they too undergo hydrolysis when implanted long term.^[10]

Oxidation

While not as degradable by hydrolysis as polyester PUs, polyether-based PUs are particularly susceptible to oxidative degradation. Oxidation can be initiated by peroxides, free radicals, and enzymes, and can also be the result of metal-catalyzed oxidation (MO). Metal-catalyzed oxidation is most frequently associated with pacemaker leads, where the metallic components of the lead catalyze oxidative degradation of the PU. Metal ions are commonly released by reaction of the metal with radicals of hydrogen peroxide and superoxide produced by phagocytic cells during inflammation.^[3]

Environmental Stress Cracking

Environmental stress cracking (ESC) is an important problem with long-term PU implants.^[8] Fig. 2 is a scanning electron micrograph of a PU that has undergone ESC. The ESC manifests itself as crack lines that are usually found perpendicular to the direction of the stress being applied on the material.

Crack depths from ESC sometimes represent a significant portion of the material's thickness, and these cracks allow for tissue infiltration, often resulting in a well-integrated material.^[3] Some implanted PUs that are not under stress develop shallow cracks. In these cases, the cracking is considered to be the result of auto-oxidation, not ESC, and the cracking is generally limited to the surface. Although the mechanisms of ESC are still being elucidated, it has been determined that oxidation is the driving force behind the degradation. Of the PUs, PCN-based PUs are the least susceptible to ESC.^[8]

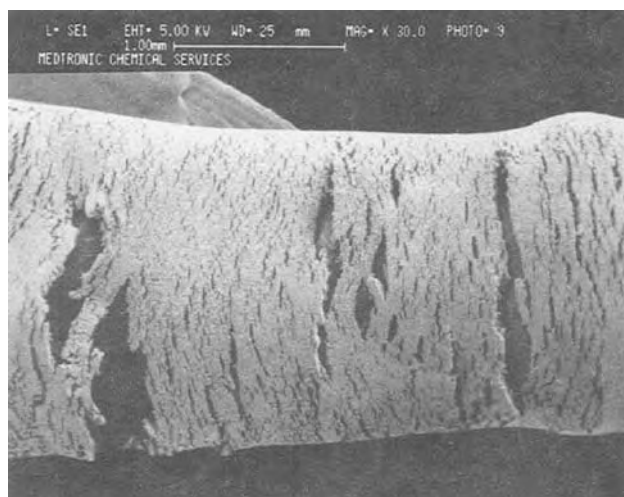


Fig. 2 A PU that has undergone ESC. (From Ref. [3].)

Enzyme and Cell-Mediated Degradation

Inflammation, one of the primary biological responses to implantation of a synthetic biomaterial, is characterized by the presence of leukocytes and macrophages.^[11] These cells have been shown to release a variety of hydrolytic enzymes. Within 48 hr of implantation, monocytes are the predominant cell type present at the PU interface.^[8] These cells change over time into monocyte-derived macrophages (MDMs), which remain present for the duration of the implant. Monocyte-derived macrophages are activated by adhesion to surfaces that are too large for them to engulf, leading to a process called frustrated phagocytosis, whereby oxidative species and hydrolytic enzymes are released in the vicinity of the PU.^[8] Substances released from phagocytes might include hydrogen ions, enzymes, and oxygen radicals.^[5] Activated MDMs also fuse to form multinucleated foreign body giant cells, the central mediators of the chronic inflammatory response to biomaterials. The adhesion and activation of MDMs can be altered by changing the chemistry of the biomaterial or by masking the PU with some other group, such as a protein layer.^[8] For more details on these methodologies, a discussion on surface modification will be presented later in this article.

Enzymatic cleavage is one of the main mechanisms by which both nondegradable and degradable PUs are degraded.^[5] Esterase, papain, bromelain, ficin, chymotrypsin, trypsin, and cathepsin C all have the ability to cleave some PUs, although the extent of degradation observed for stable PUs has been limited to date.^[5] The susceptibility to these enzymes appears to be related to soft segment chemistry and molecular weight. Trypsin has also been shown to degrade the ether linkages present in polyether soft segments.

Ratner, Gladhill, and Horbett^[12] subjected four PUs to enzymes and oxidizers to assess their biodegradation. The PUs investigated were Pellethane 2363, Estane 58409, and two materials synthesized using MDI, ED, and PPO (MW625 or 2000). The study involved two endopeptidases (papain and α -chymotrypsin), an exopeptidase (leucine aminopeptidase), and hydrogen peroxide. This study demonstrated that all three of the enzymes investigated degraded the test PUs to some extent. Of the PUs, the estane was found to degrade the least. In this study, the enzyme solutions contained 100 \times the amount necessary to cleave every bond in the PU backbone. Hence, although degradation was seen, the overall percentage of degradation was not very large, especially for such a high, nonphysiological presence of enzymes. The hydrogen peroxide degraded each of the PUs more significantly than any of the enzymes tested. Thus, the most important result of this study is the confirmation that PUs are susceptible to oxidation.

Tang, Labow, and Santerre^[13] investigated the effects of PU chemistry and morphology on degradation by cholesterol esterase. Results of this study indicated that, for PCN-based PUs, the hard segment content affects the hydrolytic stability of the material. Degradation of these urethanes is initiated at sites where there are low quantities of hydrogen-bonded carbonates and urethanes.

CURRENT TRENDS IN BIOMEDICAL PUs

Biostable PUs

Biodegradation is dependant in part on the degree of hydrogen bonding in the PU. Hence, the biostability of a PU can be improved by increasing the degree of hydrogen bonding. Gorna, Polowinski, and Gogolewski^[14] showed that the longer the diol chain, the lower the percentage of urethane linkages. The higher density of urethane linkages will enable increased hydrogen bonding. Consequently, using a lower molecular weight diol will result in a more biostable PU.

Polyurethanes based on aromatic isocyanates are more biostable than those synthesized using aliphatics.^[2,5] As a result, MDI has been the isocyanate of choice for biostable PU synthesis. Ester and ether groups have both been found to be susceptible to hydrolytic degradation. In addition, fewer ether linkages in the soft segment results in improved resistance to oxidation.^[2] Of the available diols, PTMO has been found to form PUs with the highest biostabilities.^[5]

Harder PUs have traditionally been found to be more stable *in vivo* than softer PUs.^[5] Tang, Labow, and Santerre^[13] have demonstrated that the higher the hard segment content, the more resistant the PU to

hydrolysis. It is suspected that the high degree of hydrogen bonding present within hard segments is responsible for this increased stability. Crystallinity of the soft segment may also improve the hydrolytic stability of PUs, for example, in polycaprolactone-containing PUs.^[13]

The PCN-based PUs appear to hold the most promise for use as biostable PUs. They have been shown to possess higher stability than traditional medical PUs, and there are already several PCN-PUs on the market, including: Corethane, Carbothane, Bionate, and Chronoflex.^[2,13] Carbonate linkages are more hydrolytically stable than esters because of the presence of oxygen molecules flanking both sides of the carbon.^[13] These oxygens form a quasi- π electron system, thereby stabilizing the central ketone group.^[13] In the short term, PCN-based PUs have demonstrated improved resistance to biodegradation and ESC as compared to PTMO-based PUs.^[8] The degradation of PCN urethanes appears to be the result of monocyte-derived macrophages and enzymes such as cholesterol esterase.^[2]

Other approaches explored include aliphatic PUs, polyether-based PUs, using higher macrodiols than PTMO, and siloxane-based PUs.^[2] Polyurethanes composed entirely of aliphatic components tend to be stiffer materials, and hence may not be appropriate for many biomedical applications. The polyether-based PUs developed by Gunatillake et al.^[2] were significantly more resistant to biodegradation than PTMO-based PUs. Siloxane-based PU elastomers are promising materials that incorporate many attractive properties of polydimethyl siloxane (PDMS) into the PU. These properties include: low toxicity, good hemocompatibility, good hydrolytic and oxidative stability, and low surface energy (nonadhesive).^[2] This means the siloxanes groups will migrate to the surface and help to prevent cell adhesion to the material. Both of these materials may become alternatives for PCN-based PUs.

Biodegradable PUs

Researchers have been working to develop PUs for use in drug delivery, tissue engineering, and regenerative medicine. For these applications, a biodegradable PU is desired. The ideal degradable PU scaffold would: be nontoxic (including the degradation products), have a degradation rate that correlates to tissue ingrowth, and have mechanical properties similar to that of the tissue it is replacing.

Guan et al.^[15,16] have developed a biodegradable poly(ester urethane) urea based on polycaprolactone or a copolymer of PCL and polyethylene glycol (PEO), and 1,4-diisocyanatobutane (BDI), and chain extended

with putrescine. These reactants were purposely chosen so that the PU would degrade into nontoxic degradation products, including putrescine, which is already present in vivo, and is a suspected mediator of cell growth and differentiation. The investigators have successfully synthesized materials with tensile strengths on the order of 1–2 MPa.^[15]

Skarja and Woodhouse designed a family of degradable PUs based on LDI, and L-phenylalanine-based diester chain extender and either PCL or PEO as the soft segment. The specially synthesized chain extender was added to incorporate enzymatic susceptibility into the hard segment. Chymotrypsin-like enzymes were found to successfully degrade the PU in a surface-limited fashion (Fig. 3). The PCL and PEO-based versions of these PUs were blended together, by Fromstein and Woodhouse, to create a range of PUs that are more susceptible to degradation, while maintaining the mechanical properties imparted by the PCL soft segment.^[17] By varying the soft segment molecular weight and chemistry, the degradation rates in vitro

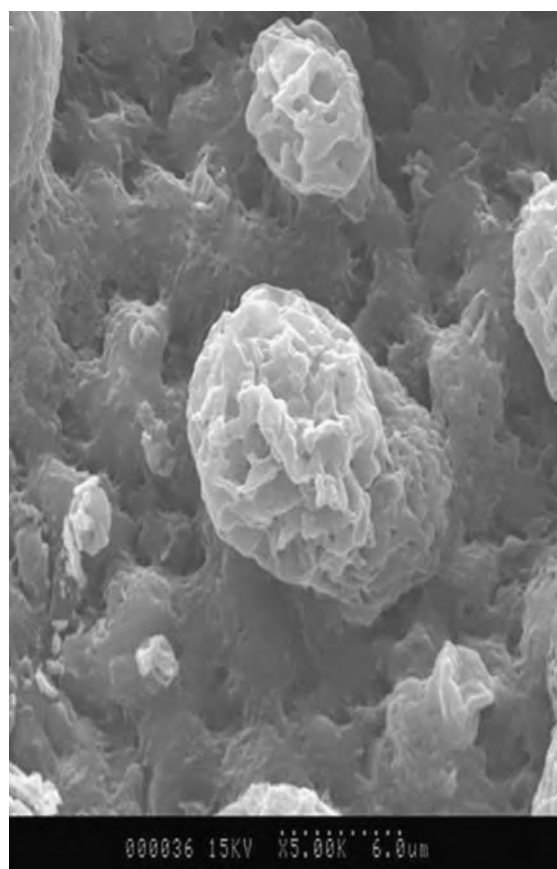


Fig. 3 A scanning electron micrograph of one of Skarja and Woodhouse's degradable PUs following 28 days of exposure to chymotrypsin. Degradation was surface limited, and the amorphous regions were worn down first, leaving nubs of crystalline spherulites sticking out. (Courtesy of G. Skarja.)

ranged from a matter of days to several years. Investigation into the degradation products of this family of PUs indicated that the presence of phenylalanine in the hard segment led to the PU backbone being cleaved not only at the ester bonds, but also at the urea bonds neighboring the phenylalanine residues.^[18]

Zhang et al.^[19] have also designed a PU that will degrade into nontoxic products that the body will recognize. Their material is synthesized using LDI, glycerol, and ascorbic acid. Degradation products from this PU do not result in a change in pH, and initial cell culture results using osteoblastic precursor cells looked promising. One reason for the selection of ascorbic acid was its ability to promote osteoblast differentiation. Although this group is not the only group using PUs for orthopedic applications, the vast majority of PU research is focused on more flexible tissues.

Another attempt at using natural components to produce a degradable PU was made by Ganta et al.^[6] who reacted 2-methyl pentane-diisocyanate and sucrose. The polymerization reaction was stopped via the addition of water to create a porous three-dimensional PU foam. These foams were tested both *in vivo* and *in vitro*, and were found to be biocompatible, and did not appear to elicit a significant inflammatory response.

Covolan et al. fabricated a PU using HMDI, PEO (MW400), PCL (MW530), and chain extended with 1,4-BD.^[20] These materials demonstrated good cell compatibility as assessed by adhesion and spreading of mouse embryo fibroblasts, but the degradative properties have not yet been reported.

Polyurethanes can also be formed into hydrogel materials that may be useful for cell encapsulation, soft tissue engineering, and drug delivery applications.^[21] Petrini et al. synthesized a hydrogel using PEO, HMDI, and diazobicyclooctane. Although the resulting material was water insoluble, the hydrophilic nature of the PEO soft segment encouraged hydrolytic degradation of the PU hydrogel, making this material a candidate for shorter term implantations.

Processing Techniques

The processing methods used to create a medical product can have an effect on the biological response to the PU. This is a result of the ability of the segments within the PU to arrange themselves differently depending on their treatment during processing. As a significant amount of the research currently involving PUs has to do with tissue regeneration, many researchers have been investigating various methods of fabricating porous three-dimensional scaffolds that can serve as temporary matrices for seeded or growing cells. Common methods of making porous materials include: foaming, thermally induced phase

separation (TIPS), particulate leaching, and electrospinning.^[15,16,22–24] The TIPS (Fig. 4) is a popular processing method that requires very little equipment. Essentially, the PU is dissolved in an organic solvent such as dioxane and then the PU solution is frozen and then subsequently freeze dried. The resulting materials have an open pore structure, and depending on the freezing temperature, PU concentration, and size of container, the physical characteristics of the scaffold can be altered.

Electrospinning PUs for use in tissue engineering has become a popular method of scaffold formation. Unlike the other techniques used, electrospinning forms scaffolds with fiber thicknesses on the micro- and nanoscale. It is hypothesized that these electrospun meshes therefore appear a lot closer to natural tissue or extracellular matrix architecture (Fig. 5).

In a recent study, Kidoaki, Kwon, and Matsuda^[22] made scaffolds for use in tissue engineering by cospinning a PU and PEO into microfibers. The incorporation of PEO was intended as a leachable that would be hydrolyzed from the construct prior to implant. The microvoids formed by the elimination of PEO are hypothesized to provide pores via which cells can migrate into the construct. The merits of this approach have not yet been confirmed.

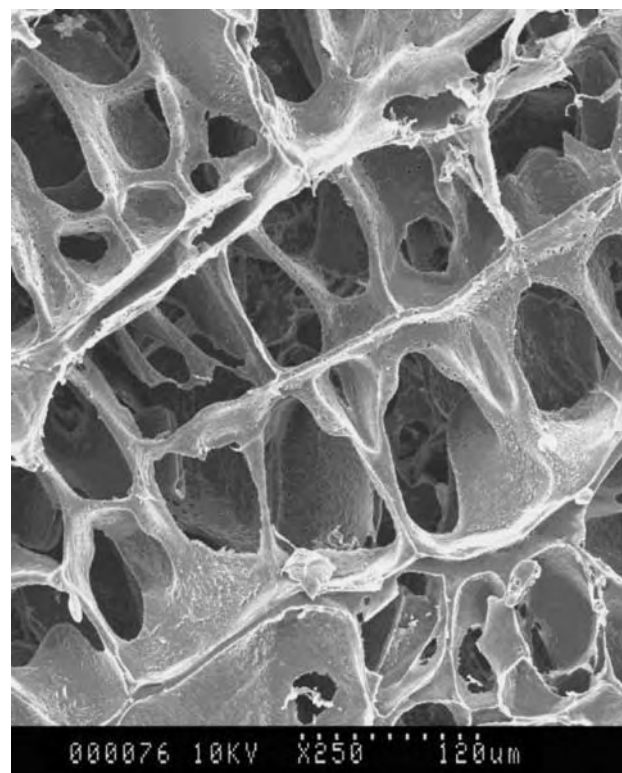


Fig. 4 Scanning electron microscopy of a biodegradable PU scaffold for tissue engineering fabricated using TIPS. (Courtesy of C. Alperin.)

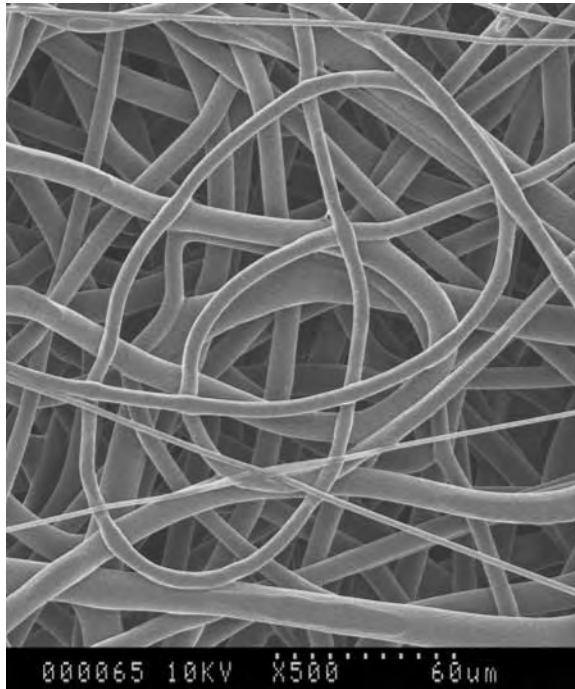


Fig. 5 An electrospun mat of a biodegradable PU developed by Skarja and Woodhouse. Matrices like these are being used to seed cardiomyocytes for use as a cardiac patch. (Photo courtesy of J. Fromstein and material electrospun by D. Rockwood and J. Rabolt, University of Delaware.)

The phase-segregated nature of segmented PUs can have a significant impact on the biological response to the material.^[25] Takahashi, Kita, and Kaibara were able to improve the blood compatibility of Pellethane, simply by thermally annealing the material. It is believed that the thermal annealing encourages the surface of the PU to become more hydrophobic and therefore prevents the activation of adhered platelets.^[25] This study also determined that there was a different biological reaction to a PU depending on whether the PU was in contact with air or with a solid surface during fabrication/molding. This reinforces the point that the processing method plays an important role in the compatibility of the PU.

Biomedical PU Modifications

Antibacterial PUs

One area that has received some attention over the past decade is the development of a PU that has antimicrobial agents incorporated into the material in an attempt to minimize the risk of implant-associated infection. Grapski and Cooper,^[9] for example, synthesized a series of nonleaching biocidal PUs containing quaternized pyridine residues in the chain

extender. Results from this study indicated that the PUs exhibited bactericidal activity against *Staphylococcus aureus* and bacteriostatic activity against *Escherichia coli*.

Woo, Mittleman, and Santerre^[26] used a similar strategy, incorporating the antimicrobial agent into the PU backbone. In their work, however, they designed a biodegradable polyester urethane that would intentionally release antimicrobial agents into the implant site. Their strategy relied on the inflammatory process to trigger degradation of the polymer and to control the magnitude of the response.

Improving biological response to PUs

Biocompatibility is a property that will change for a given application depending on what biological areas the material will be exposed to. The interfaces to which medical devices are exposed can be grouped into three categories: 1) direct contact with blood (e.g., a vascular graft or heart valve); 2) direct contact with solid tissue (e.g., fatty, connective, and muscle); and 3) contact with other bodily fluids (e.g., urine). When modifications of the PUs are performed, these biomaterial interfaces must be taken into account.

Blood–Material Interfaces. Blood-compatible surfaces are defined as materials that do not induce a large number of platelets and fibrinogen molecules to adhere, and do not activate platelets during blood contact.^[4] The segmented PUs have traditionally been considered to have relatively good blood compatibility. This property may be attributable in part to the phase-segregated nature of the PU, although there does not appear to be a direct effect of hard-to-soft segment ratio on blood compatibility.^[4] To improve blood–material interactions, a number of techniques have been investigated. Some of the more important findings are summarized below.

Grasel and Cooper^[4] showed that lower degrees of phase segregation in a PU system based on PTMO and PPO led to the formation of relatively large thrombi during blood contact. This study also showed that changing the degree of phase segregation by changing the ratio of PTMO to PPO was a primary determinant for blood compatibility. These results indicate that the segmented morphology of PUs may play a central role in hemocompatibility.

Many researchers have utilized PEO either as the soft segment incorporated in the polymer backbone or grafted onto the PU surface.^[4,10,27] These strategies result in surfaces that are hydrophilic in nature and have low interfacial surface energy, high chain mobility, and molecular chain conformation.^[27] Polyethylene oxide had been a popular choice for biomaterials because it is

resistant to protein adsorption, platelet adhesion, and bacterial adhesion.^[4,27] For example, Corneille et al.^[7] incorporated PEO into a standard PTMO-based PU, and found that the addition of PEO yielded a polymer that was nonadhesive for cells, but was relatively cyto-compatible. This PU holds promise for use in tissue engineering applications.

There remains a debate over the benefits of hydrophobicity. Some researchers have found that hydrophobic surfaces are hemocompatible.^[10] Others, however, feel that hydrophobic biomaterials initiate a foreign body reaction.^[23] Although a foreign body response will be initiated by any synthetic material, whether the extent of this response is greater than for hydrophilic surfaces is not known with any certainty.

A method of surface modification that has been popular for use with PUs is the use of surface-modifying additives (SMAs) also called surface-modifying macromolecules.^[28–31] These synthetic surface modifiers are designed to have chemistries similar to that of the PU they are being incorporated into. This similar chemistry enables hydrogen bonding between the PU and SMA to occur, thereby firmly anchoring the modifiers to the material. Surface-modifying additives make use of groups such as PEO or fluorine tails to ensure that the SMA is predominantly present at the surface of the biomaterial rather than in the bulk. The presence of these additives at the PU surface has been shown to improve blood compatibility.^[29,30]

Another method of ensuring good blood compatibility is to achieve a covering of endothelial cells (ECs) on the PU.^[32] Unmodified PUs do not tend to support endothelial cell adhesion and proliferation.^[25] An intact endothelial layer would mask the underlying PU from the blood, mimicking the natural lining of a blood vessel. The ECs are negatively charged, and release antiplatelet and antithromogenic factors into the vicinity.^[33] Zhu et al. published a study recently, in which porous PU (Estane) tubes were covalently coated with collagen, gelatin, or chitosan. The resulting materials were then seeded with human umbilical vein endothelial cells (HUVECs), and cultured for six days in vitro. The authors were able to achieve a viable intima on the luminal side of the PU scaffold, indicating that these modified PUs hold promise for use in vascular tissue engineering.

Tissue–PU Interactions. When it comes to material–tissue interactions, we are usually trying to encourage cells to adhere to and grow on the PU. In the case of a medical device such as a wound dressing a minimal foreign body response is desired. The response could be prevented most successfully by the formation of an endothelial cell layer along the surface of the implant. Similarly, for a tissue engineering construct, it is also important to encourage the growth and adhesion of

the cell type(s) native to the tissue being replaced. The second goal in tissue engineering can sometimes be overlooked, but is equally important to the long-term survival of an implanted construct, angiogenesis. To ensure that a newly populated construct will remain viable, scientists should consider adding angiogenic factors to their construct in an attempt to address this issue.

By controlling hydrophobicity of the polymer surface, the adherence of cells can also be controlled.^[6] A hydrophobic surface will encourage the absorption of a passive layer of proteins like albumin upon implantation.^[4] Sanders et al.^[23] have been charging the surface of PU scaffolds using plasma coatings in an attempt to affect cell adhesion, proliferation, and migration. To date they have not found conclusive evidence that they can reproducibly control the biological response using surface charge alone. Others have found that positively charged surfaces do enhance endothelial cell migration and spreading.^[33]

One of the most popular strategies for modulating cell and tissue response to foreign materials involves the use of natural materials to “mask” the surface or to mimic extracellular matrix. Researchers have used adsorption, grafting, and SMAs as methods of distributing the biological materials across the PU surface. They have tried using proteins, such as fibronectin, and peptides, such as RGD and YIGSR, that are known to affect cell adhesion.

In one study, Wang et al.^[33] used SMAs to coat their PU with RGD peptide and with the individual amino acids lysine, arginine, glycine, and aspartic acid. Results from this work showed that similar endothelial cell growth and adherence was achieved with both the peptides and the basic amino acids (lysine and arginine). Interestingly, they tried dispersing a 1:1:1 mixture of the individual amino acids making up the RGD across the surface of the PU. The response of the ECs to this mixed surface was significantly lower than to the RGD-covered PU, indicating that while basic coatings comprised a single amino acid may perform equally well to an RGD-covered surface, the RGD sequence is only effective as a sequence, and not as three separate amino acids.^[33]

Guan et al.^[16] found that cell adhesion to their degradable putrescine-based poly(ether ester urethane) ureas was low compared to tissue culture polystyrene. In an attempt to improve the biological response, they modified by immobilizing an adhesion peptide (RGDS) onto the surface using a spacer molecule and radio-frequency glow discharge. A marked improvement in cell adhesion was achieved.

Kidoaki, Kwon, and Matsuda.^[22] have recently published results of layering electrospun PU (micro-scale) with sheets of electrospun collagen and gelatin (nanoscale). In addition to the spinning of flat sheets,

tubular structures with an inner (luminal) coating of collagen and an outer layer of PU were produced.

Jun and West^[27] recently published results of a new PU synthesized using MDI, PEO as the soft segment, and a peptide chain extender containing the sequence YIGSR sequence. The incorporation of a bioactive peptide and PEO was expected to result in improved endothelial cell adhesion. Their results indicated that in addition to improved cell adhesion and spreading, the material also suppressed platelet adhesion. This improved blood response is attributed to the presence of PEO.

Van Wachem et al.^[34] investigated tissue integration into Pellethane that had collagen I immobilized on the surface using acrylic acid spacer arms. Collagen-immobilized PU was implanted subcutaneously in rats, using the acrylic acid-grafted Pellethane and original Pellethane as controls. Results indicated that the collagen-covered PU incited a far more intense foreign body reaction than the controls, as evidenced by increased fibrin deposition and accumulation of granulocytes and macrophages. At the end of six weeks, however, the collagen-immobilized samples had the thinnest capsule formed around them, and demonstrated the most adherence of connective tissue.^[34] It is important to note that the collagen used in this study was from a bovine source, and so there is an additional cross-species response that may have played a role in the elevated initial foreign body response.

Increased metabolite transport may occur in scaffolds with increased hydrophilicity.^[15] This could result not only in increased degradation, but also in a more viable environment for cells deeper within the construct. This could be a way to improve angiogenesis and cell infiltration within the center of the scaffold. The encouragement of angiogenesis and cellular penetration are primary goals in the design of any tissue-engineered construct.^[6]

Exposure to Biological Fluids. In some cases, an implant will be interfacing with tissue on one side and bodily fluid on the other. A good example of this would be a bladder patch. As the implant will be exposed to urine, it is important to select a material that will not degrade if a permanent implant is sought. In addition, the material should not induce calcification. A PU/PEO hybrid hydrogel, Aquavene, has been tested for degradation in artificial urine, and results indicated that degradation was not evident.^[35]

CONCLUSIONS

Although PUs have played an important part in biomedical applications since the 1960s, until recently

their use was restricted to nondegradable device applications including catheters, wound dressings, coatings on pacemaker leads, and ventricular assist devices. Researchers are now designing PUs specifically for a given application and are expanding the use of these polymers into the arena of tissue engineering. Features being incorporated through the flexible chemistry of these segmented PUs include: tunable degradation rates, bioactive degradation products, resistance to hydrolysis for permanent implants, and modification of surface charge and morphology to support cellular interactions. With their diverse mechanical, chemical, and physical properties, and relative ease of processing, PUs are likely to play a major role in medical implants into the future.

ARTICLES OF FURTHER INTEREST

Biomimetic Materials; Elastomers, Biodegradable; Poly(carbonate)urethanes; Polymers; Tissue Engineering of Heart Valves; Vascular Grafts

REFERENCES

1. Guignot, C.; Betz, N.; Legendre, B.; Le Moel, A.; Yagoubi, N. Influence of filming process on macromolecular structure and organization of a medical segmented polyurethane. *J. Appl. Polym. Sci.* **2002**, *85* (9), 1970–1979.
2. Gunatillake, P.A.; Martin, D.J.; Meijs, G.F.; McCarthy, S.J.; Adhikari, R. Designing biostable polyurethane elastomers for biomedical implants. *Aust. J. Chem.* **2003**, *56* (6), 545–557.
3. Lamba, N.M.K.; Woodhouse, K.A.; Cooper, S.L. *Polyurethanes in Biomedical Applications*; CRC Press LLC: Boca Raton, 1998.
4. Grasel, T.G.; Cooper, S.L. Surface-properties and blood compatibility of polyurethaneureas. *Biomaterials* **1986**, *7* (5), 315–328.
5. Pinchuk, L. A review of the biostability and carcinogenicity of polyurethanes in medicine and the new generation of 'biostable' polyurethanes. *J. Biomat. Sci. Polym. Ed.* **1994**, *6* (3), 225–267.
6. Ganta, S.R.; Piesco, N.P.; Gassner, R.; Motta, L.F.; Papworth, G.D.; Stolz, D.B.; Watkins, S.C.; Agarwal, S. Vascularization and tissue infiltration of a biodegradable polyurethane matrix. *J. Biomed. Mat. Res. A* **2003**, *64A* (2), 242–248.
7. Corneillie, S.; Lan, P.N.; Schacht, E.; Davies, M.; Shard, A.; Green, R.; Denyer, S.; Wassall, M.; Whitfield, H.; Choong, S. Polyethylene glycol-containing polyurethanes for biomedical applications. *Polym. Int.* **1998**, *46*, 251–259.
8. Matheson, L.A.; Santerre, J.P.; Labow, R.S. Changes in macrophage function and morphology due to

- biomedical polyurethane surfaces undergoing biodegradation. *J. Cell. Physiol.* **2004**, *199* (1), 8–19.
9. Grapski, J.; Cooper, S.L. Synthesis and characterization of non-leaching biocidal polyurethanes. *Biomaterials* **2001**, *22*, 2239–2246.
 10. Jayabalan, M.; Lizymol, P.P.; Thomas, V. Synthesis of hydrolytically stable low elastic modulus polyurethane-urea for biomedical applications. *Polym. Int.* **2000**, *49* (1), 88–92.
 11. Labow, R.S.; Erfle, D.; Santerre, J.P. Neutrophil-mediated degradation of segmented polyurethanes. *Biomaterials* **1995**, *16*, 51–59.
 12. Ratner, B.D.; Gladhill, K.W.; Horbett, T.A. Analysis of in vitro enzymatic and oxidative degradation of polyurethanes. *J. Biomed. Mat. Res.* **1988**, *22*, 509–527.
 13. Tang, Y.; Labow, R.S.; Santerre, J.P. Enzyme-induced biodegradation of polycarbonate polyurethanes: dependence of hard-segment concentration. *J. Biomed. Mat. Res.* **2001**, *56*, 516–528.
 14. Gorna, K.; Polowinski, S.; Gogolewski, S. Synthesis and characterization of biodegradable poly(caprolactone urethane)s. I. Effect of the polyol molecular weight, catalyst and chain extender on the molecular and physical characteristics. *J. Polym. Sci. A Polym. Chem.* **2001**, *40*, 156–170.
 15. Guan, J.J.; Fujimoto, K.L.; Sacks, M.S.; Wagner, W.R. Preparation and characterization of highly porous, biodegradable polyurethane scaffolds for soft tissue applications. *Biomaterials* **2005**, *26* (18), 3961–3971.
 16. Guan, J.J.; Sacks, M.S.; Beckman, E.J.; Wagner, W.R. Biodegradable poly(ether ester urethane)urea elastomers based on poly(ether ester) triblock copolymers and putrescine: synthesis, characterization and cytocompatibility. *Biomaterials* **2004**, *25* (1), 85–96.
 17. Fromstein, J.D.; Woodhouse, K.A. Elastomeric biodegradable polyurethane blends for soft tissue applications. *J. Biomat. Sci. Polym. Ed.* **2002**, *13* (4), 391–406.
 18. Elliott, S.L.; Fromstein, J.D.; Santerre, J.P.; Woodhouse, K.A. Identification of biodegradation products formed by L-phenylalanine based segmented polyurethaneureas. *J. Biomat. Sci. Polym. Ed.* **2002**, *13* (6), 691–711.
 19. Zhang, J.Y.; Doll, B.A.; Beckman, E.J.; Hollinger, J.O. A biodegradable polyurethane-ascorbic acid scaffold for bone tissue engineering. *J. Biomed. Mat. Res. A* **2003**, *67A* (2), 389–400.
 20. Covolan, V.L.; Di Ponzio, R.; Chiellini, F.; Fernandes, E.G.; Solaro, R.; Chiellini, E. Polyurethane based materials for the production of biomedical materials. *Macromol. Symp.* **2004**, *218*, 273–282.
 21. Petrini, P.; Fare, S.; Piva, A.; Tanzi, M.C. Design, synthesis and properties of polyurethane hydrogels for tissue engineering. *J. Mat. Sci. Mat. Med.* **2003**, *14* (8), 683–686.
 22. Kidoaki, S.; Kwon, I.K.; Matsuda, T. Mesoscopic spatial designs of nano- and microfiber meshes for tissue-engineering matrix and scaffold based on newly devised multilayering and mixing electrospinning techniques. *Biomaterials* **2005**, *26*, 37–46.
 23. Sanders, J.E.; Sanders, J.E.; Lamont, S.E.; Karchin, A.; Golledge, S.L.; Ratner, B.D. Fibro-porous meshes made from polyurethane micro-fibers: effects of surface charge on tissue response. *Biomaterials* **2005**, *26* (7), 813–818.
 24. Stankus, J.J.; Guan, J.J.; Wagner, W.R. Fabrication of biodegradable elastomeric scaffolds with sub-micron morphologies. *J. Biomed. Mat. Res. A* **2004**, *70A* (4), 603–614.
 25. Takahashi, A.; Kita, R.; Kaibara, M. Effects of thermal annealing of segmented polyurethanes on surface properties, structure and antithrombogenicity. *J. Mat. Sci. Mat. Med.* **2002**, *13*, 259–264.
 26. Woo, G.; Mittleman, M.; Santerre, J.P. Synthesis and characterization of a novel biodegradable antimicrobial polymer. *Biomaterials* **2000**, *21*, 1235–1246.
 27. Jun, H.-K.; West, J.L. Modification of polyurethaneurea with PEG and YIGSR peptide to enhance endothelialization without platelet adhesion. *J. Biomed. Mat. Res. B Appl. Biomater.* **2005**, *72B*, 131–139.
 28. Ho, J.Y.; Matsuura, T.; Santerre, J.P. The effect of fluorinated surface modifying macromolecules on the surface morphology of polyethersulfone membranes. *J. Biomat. Sci. Polym. Ed.* **2000**, *11* (10), 1085–1104.
 29. Jahangir, A.R.; McClung, W.G.; Cornelius, R.M.; McCloskey, C.B.; Brash, J.L.; Santerre, J.P. Fluorinated surface-modifying macromolecules: modulating adhesive protein and platelet interactions on a polyether-urethane. *J. Biomed. Mat. Res.* **2002**, *60* (1), 135–147.
 30. Jahangir, R.; McCloskey, C.B.; McClung, W.G.; Labow, R.S.; Brash, J.L.; Santerre, J.P. The influence of protein adsorption and surface modifying macromolecules on the hydrolytic degradation of a poly(ether-urethane) by cholesterol esterase. *Biomaterials* **2003**, *24* (1), 121–130.
 31. Ward, R.S.; Tian, Y.; White, K.A. Improved polymer biostability via oligomeric end groups incorporated during synthesis. *Abstr. Pap. Am. Chem. Soc.* **1998**, *216*, U907–U907.
 32. Zhu, Y.B.; Gao, C.Y.; He, T.; Shen, J.C. Endothelium regeneration on luminal surface of polyurethane vascular scaffold modified with diamine and covalently grafted with gelatin. *Biomaterials* **2004**, *25* (3), 423–430.
 33. Wang, D.A.; Feng, L.X.; Ji, J.; Sun, Y.H.; Zheng, X.X.; Elisseff, J.H. Novel human endothelial cell-engineered polyurethane biomaterials for cardiovascular biomedical applications. *J. Biomed. Mat. Res. A* **2003**, *65A* (4), 498–510.
 34. van Wachem, P.B.; Hendricks, M.; Blaauw, E.H.; Dijk, F.; Verhoeven, M.; Cahalan, P.T.; van Luyn, M.J.A. Electron microscopic observations on tissue integration of collagen-immobilized polyurethane. *Biomaterials* **2002**, *23* (6), 1401–1409.
 35. Gorman, S.P.; Tunney, M.M.; Keane, P.F.; VanBladel, K.; Bley, B. Characterization and assessment of a novel poly(ethylene oxide)/polyurethane composite hydrogel (Aquavene[®]) as a ureteral stent biomaterial. *J. Biomed. Mat. Res.* **1998**, *39*, 642–649.



Positron Emission Tomography (PET)

Anna-Liisa Brownell

Department of Radiology, Massachusetts General Hospital, Boston, Massachusetts, U.S.A.

INTRODUCTION

The concept of positron emission-based imaging was developed in 1951^[1] and the first human studies were published in 1953.^[2] Since then, fundamental technical development has taken place in all scientific fields including electronics, computer science, biology, and biochemistry. These developments have also significantly advanced the development of positron emission-based imaging techniques. Even though positron imaging technology is more than 50 years old, its use in the clinical environment has been recognized only in the past 10 years. Interestingly, the earliest applications of positron imaging were tumor studies in the brain.^[2] The main clinical use, as well as a major research interest today, is also related to cancer. In the early applications long-lived positron emitters such as copper-64 and arsenium-72 were used for tumor detection, whereas today the most important positron ligand is a glucose analogue, fluorine-18 labeled 2-fluoro-2-deoxy-D-glucose. The limited availability of suitable positron ligands has been the biggest delay to clinical applications.

Technical development of positron emission imaging has gone through many steps after starting as two-dimensional imaging using one pair of single detectors.^[2] The major developments have been in detector design including crystal material, crystal size, light detection, photomultiplier development, and overall detection geometry. However, the most important aspect has been the development of mathematical algorithms. This development turned positron imaging into positron emission tomographic (PET) imaging.^[3–8] Further development of tomographic image reconstruction including filtering techniques has followed the advances in computer science.

Recently, PET imaging systems have also been developed for specific applications. In addition to self-standing, whole-body PET scanners, combined PET and X-ray computed tomographic (CT) systems are commercially available. These systems can combine high resolution anatomical information provided by CT to the functional information of PET. Also, organ-specific systems are developed along with many different designs for the imaging of small animal models.^[9,10]

Studies in the small animal models and advances in gene research have introduced a new level of biological molecular imaging and biomarker has become a magic

term in noninvasive imaging techniques. PET imaging techniques have high detection sensitivity; thus, only tracer amounts of radiolabeled ligand or drug is needed for imaging and the homeostasis of the biological system will not be changed. PET imaging techniques can be used to investigate physiological functions related to blood flow, metabolism, or receptor binding; to track cells or reporter genes; and to test drug candidates. Radiolabeled biomarkers have significantly enhanced the role of PET imaging, especially in drug development, as well as in developing new sensitive diagnostic methods and/or techniques to follow the response of therapeutic regimen.

TECHNICAL ASPECTS OF PET

Positron Emission and Annihilation

The nuclear origin of the positron is a proton. Under the influence of other nucleons the proton is converted into a neutron, a positron, and a neutrino. The immediate result of the breakdown is that the positron and the neutrino are ejected from the nucleus while the neutron remains (Fig. 1). The released positron carries kinetic energy depending on the binding energy in the nucleus and loses it while traveling in the surrounding media. The annihilation reaction is a result of a collision between the positron and a negatively charged electron. In annihilation the masses of the two particles completely turn into energy. The energy is divided equally between the two photons (each of 511 keV), traveling to the opposite directions ($180 \pm 0.25^\circ$) of each other. In PET imaging these two emitted photons are detected as a line between a detector pair (Fig. 1).

Light Detection

In the earliest applications (Fig. 2), sodium iodide (NaI) was used as a crystal material to absorb the emitted photon (released in annihilation) and turn it into light (Fig. 2). With thallium, enhanced NaI is still used as crystal material in instrument design. The drawbacks of NaI(Tl) are its hygroscopicity and low melting point. There are a number of other crystals that are better in these aspects and have better light production.

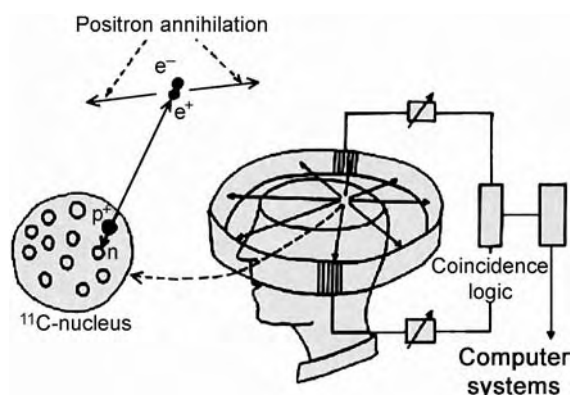


Fig. 1 Schematic diagram of the process where a positron is released from the nucleus of the radioactive marker. After that, the released positron travels in the surrounding media and loses its energy, finally colliding with a negatively charged electron. In this process, annihilation, the masses of the two particles completely turn into energy, which is divided between the two photons (each 511 KeV) traveling in opposite directions ($180 \pm 0.25^\circ$) of each other. This released photon pair can hit the scintillation crystals connected with coincidence logic and can be detected as a line between the crystals. This resultant information is used to generate images of activity distribution.

Characteristics of some crystals used in the PET imaging devices are presented in Table 1. Barium fluoride (BrF) is the fastest of the crystals listed in Table 1. It is used in PET designs based on time-of-flight detection. Bismuth-germanate (BGO) and lutetium oxyorthosilicate (LSO)^[11] are presently the most commonly used crystal materials in commercial PET devices. Of these two crystals, LSO is superior, having 5 times more light output than BGO and 7.5 times faster decay of light than BGO, resulting in a combined speed and light output improvement of 37.5 times over BGO.^[11]

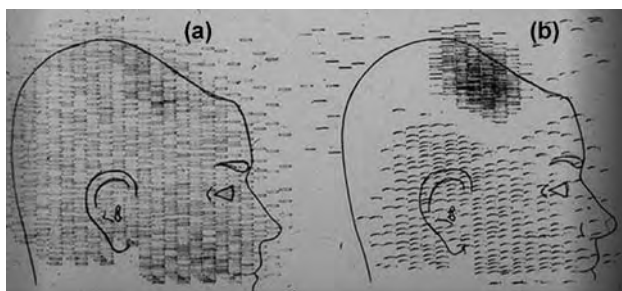


Fig. 2 Coincidence scan (a) of a patient showing recurrence of tumor under the previous operation site, and an unbalance scan (b) showing asymmetry to the left. The site of the previous tumor is indicated by the indentation in the skull outline, which resulted from the absence of the bone flap. (Reproduction of a picture published in Ref. [2].)

Table 1 Characteristics of the most important scintillators used for detector material

Quantity	Unit	Na(Tl)	BGO	LSO	BaF
Density	g/cm ³	3.67	7.13	7.4	4.89
Radiation length	cm	2.6	1.11	1.2	2.06
Decay constant	ns	230	300	40	0.6
Light yield	relative	100	15		5
Peak emission	nsec	415	480	420	210
Radiation hardness	log (rad)	3	6	7	7
Hygroscopicity	yes	no	no	no	no
Melting point	°C	65	1,000	1,000	135

(From Ref. [11].)

Detector Design

To obtain optimal resolution and sensitivity, crystals are cut into different sizes and shapes and packed in different ways to enhance the special characteristics of PET imaging devices. In the earlier designs, the crystals were large. In PCI (Positron Camera I) and PCII (Positron Camera II) PET scanners, constructed at Massachusetts General Hospital, the diameter of the crystals was $\frac{3}{4}$ " and the thickness $2\frac{1}{4}$ ". In PCI the photomultiplier tubes (PMTs) were arranged in a square pattern and in PCII in a hexagonal pattern over the crystals. Since the system resolution correlates with the crystal size and is generally almost equal to it, the crystal size has been decreased significantly. In the present commercial scanners, the crystal pattern is typically rectangular or square. In the commercial human scanners, the size of the crystal surface is between 4 and 7 mm, while the thickness is between 20 and 30 mm (Table 2). In the animal scanners, the crystal size is smaller to obtain a higher resolution. In the commercial animal scanners, the crystal size is typically $2 \times 2 \times 10$ mm (Concord Microsystems; P4 and R4), and in some in-house built scanners, even as small as $1 \times 1 \times 5$ mm.^[10]

Because it was not possible to reduce the size of PMTs in the same scale as crystal size, the concept of analogue coding^[13] and block detector^[14] was developed. In the first PET device using analogue coding, crystals were placed on a circular light guide and one PM tube was looking for four crystals. Crystals were 20 mm long and 30 mm thick and the width varied from 4 mm at the front to 4.5 mm at the light guide, permitting the detectors to be assembled in a continuous ring.^[15] The position of the crystal, from where light was emitted, was determined by the ratio of the adjacent photomultiplier signals. The first block detector had 32 crystals for four PMTs.^[16] Photomultipliers are also of different shapes and sizes. In addition to the traditional

Table 2 Characteristics of some commercially available PET systems

Model	Advance NXi	ECAT		ALLEGRO	C-PET plus
		EXACT HR+	ECAT ACCEL		
<i>Detectors</i>					
Number of detector rings	18	32	24	N/A	N/A
Number of crystals	12,096	18,432	9,216	17,864	6
Crystal material	BGO	BGO	LSO	GSO PIXELAR	Curved NaI(Tl) CCT
Crystal size (mm)	3.9×8.2×30	4.05×4.39×30	6.75×6.75×25	4×6×20	500×300×25
Number of PMTs	672 Dual	1,152	576	420	288
<i>Physical parameters</i>					
Axial spatial resolution (mm)	4	4.2	4.3	4.2	5.5
Transaxial spatial resolution (mm)	4.8	4.6	6.2	4.8	5.0
System sensitivity (2D) kcps/μCi/cc	200 (trues)	200 (trues + singles)	180 (trues + singles)	N/A	N/A
System sensitivity (3D) kcps/μCi/cc	1,060 (trues)	900	780	800 trues	450 trues
<i>Image reconstruction</i>					
Filtered backprojection	Yes	Yes	Yes	Yes	Yes
Iterative reconstruction	WA-OSEM FORE + WLS-OSEM	WA-OSEM	WA-OSEM	FORE/ OSEM	FORE/OSEM

(From Ref. [12].)

tube model there are flat area-sensitive PMTs. Table 2 summarizes detector designs of some commercial PET devices.^[12]

Gantry Assembly

The earliest PET cameras had a two-detector bank design and the camera heads rotated at 180° (Fig. 3).^[17–19] This first design was followed by several single-ring designs^[20–22] and hexagonal designs having six detector banks.^[23] After that, several multiring systems were developed, and finally separate rings have been replaced

with detector blocks. Also, a cylindrical design has been built.^[24] To obtain a large covering for the whole-body cameras, the two-detector bank design has been redesigned with large rectangular detector heads. Because of the large variation in the gantry design of the present PET devices, the standards to evaluate performance of the imaging systems have been changed accordingly.^[25]

Resolution

Resolution in positron imaging is limited by positron range, small angle deviation, sensitivity, count-rate

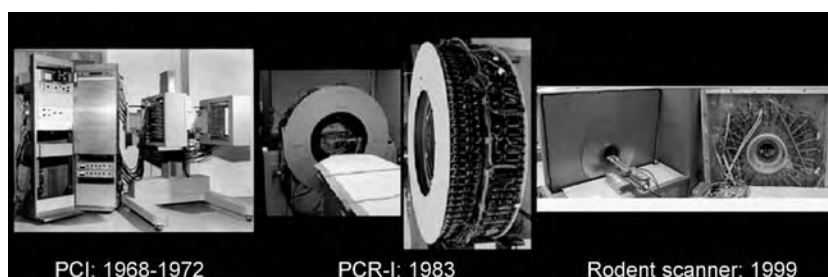


Fig. 3 Development of the PET designs at the Massachusetts General Hospital. On the left is PCI, the first positron emission tomograph with a two-detector bank design, built in 1968–1972. The spatial resolution of this device was 10 mm. In the middle is the first PET device that used analogue coding, built in 1983. The spatial resolution of this device was 4.5 mm. On the right, is a high resolution tomograph for rodent studies, built in 1999. The spatial resolution at the center is 1.16 mm.

capability, and sampling. Each of these factors must be considered in the optimization of an instrument design. Spatial resolution of the tomograph determines how small of a structure can be observed. Structures that are smaller than the spatial resolution cannot be perfectly distinguished from adjacent structures and are blurred by the activity from near-by structures. Spatial resolution is an important factor, especially in the quantification of radioactivity concentration. The size of the object should be at least two times the spatial resolution to avoid partial volume effects.^[26] If the object is smaller, there will be an underestimation in activity concentration. However, on the visualization point, partial volume effect may not be a problem if there is high activity or high contrast between the lesion and the background activity.

Spatial resolution is considered separately for in-plane (transverse) and out-of-plane (axial) directions. In-plane resolution is determined by the spread of the point or line source in the direction of a line from the center of the gantry to the position of the source (radial) and perpendicular to it (tangential). When the source is at the center of the gantry, radial and tangential resolution should be the same and would give the best value for the spatial resolution. Normally the spatial resolution is given at that point as a width of the half maximum.

Out-of-plane (axial) resolution refers to the slice thickness and ability to separate structures in the axial direction. According to the latest National Electrical Manufacturers Association (NEMA) recommendations,^[25] the spatial resolution should be measured with a point source in all three directions, rather than by a line source, to allow more comparable results between two-dimensional and three-dimensional imaging modes of PET. The spatial resolution of the positron imaging devices has improved from 30 mm in the early scanners^[2] to 4.5 mm in commercial whole-body scanners (Table 2). In commercial small-animal scanners, the resolution is 1.8 mm (primate microPET, P4, Concord Microsystems) and as good as 1.16 mm in the special in-house built rodent PET scanner.^[10]

Sensitivity

Sensitivity of the tomograph is determined as a ratio of the coincidence count rate and the radioactivity concentration. Annihilation photons are emitted in all directions, and only a small percent of photons can be detected in the tomograph. Depending on the width of the detector bank, fewer than 10% of the emitted photons are detected. If the efficiency of the scintillation crystal is 80%, the coincidence detection efficiency in two crystals is 64% and the PET device is able to detect about 6.4% of the annihilation events. If the radius of

the gantry is 50 cm, it can be calculated that only 1% of the emitted photons can be detected in a 1-cm-wide slice.

The sensitivity of the PET scanner is measured with a phantom, normally using ^{18}F as a positron emitter. Only a small amount of radioactivity is used in these measurements to minimize random coincidences and dead time losses. According to the latest NEMA recommendation,^[25] the sensitivity should be measured using a 70-cm long tube filled with radioactivity and surrounded by several sleeves to create stepwise differences in attenuation. The attenuation-corrected counting rate is mathematically calculated from the observed data and used to determine sensitivity. However, if the length of the field of view of the PET scanner is less than 19 cm, the sensitivity can be measured with a phantom of 20-cm diameter and 19-cm length.

Data Acquisition

Data acquisition is a compromise of the characteristics of the tomograph and the requirement of the biological application. Injected activity (radiation dose) and acquisition time are often limited by the biological system. The acquisition parameters in the system are adjusted with these guidelines. The old classical data acquisition method is two-dimensional, which means that the data are acquired in a plane. The recent technical development in detector design has enabled three-dimensional data acquisition, which means that data can be acquired both in the crossplane and in the axial direction. Three-dimensional data acquisition increases sensitivity, but internal scatter and random coincidences will also be higher than in two-dimensional acquisitions. Additional technical factors affecting data acquisition and final image quality are sampling, statistical noise, scatter, and attenuation.

Sampling

Sampling describes the density of coincidence lines in the field-of-view. The linear sampling distance (d) sets a resolution limit for any imaging system. This limit is known as Nyquist frequency ($d \leq (1/2v_{\text{max}})$) and means that the highest spatial frequency component to be recovered from the data has to be sampled with at least two points per cycle. Axial sampling can be improved by the motion of the detector head in the axial direction by half of the slice thickness.^[27] In-plane sampling can be enhanced by wobbling, the circular motion of the detector head by 1–2 cm off-center.^[28]

Statistical noise

During data acquisition, the collected counts are subject to the statistical noise, since radioactive disintegrations



occur at random and it is not possible to determine when a given atom will disintegrate. The detected counts (N) have a standard deviation \sqrt{N} and the statistical noise is determined as $1/\sqrt{N}$. Statistical uncertainty increases with low counts. However, obtaining high counts might be limited by radiation dose or imaging time. Long imaging times increase the possibility for motion-induced artifacts.

Scatter

The scattering of annihilation photons leads to falsely positioned coincidence events. They contribute to the noise and loss of the contrast in PET images. Scattered coincidences cannot be distinguished from true coincidences electronically, and there are no exact methods for correction of scatter in biological studies. However, there are several methods for partial correction.^[29,30] Internal scatter in biological systems is often a burden if some low activity areas are investigated near an area with a high accumulation of activity. An example of this problem is studies of glucose metabolism in the heart. After administration of the radiolabeled ligand ^{18}F -2-fluoro-2-deoxy-D-glucose (^{18}F -FDG), the high activity will be first in the blood pool inside the ventricles, but after 30 minutes the activity will be metabolized into the cardiac muscle and cleared off from the blood pool. In the image, the high scatter is first at the cardiac muscle because of the high activity inside the walls. After 30 minutes the high scatter will be in the lumen because of the high activity in the cardiac walls. In this special case, the scatter-created malfunction is called spillover.

Attenuation

Photons created through annihilation lose energy when passing through the tissue depending on the thickness and density of the tissue. Attenuation correction is necessary to obtain activity distribution, especially if surrounding tissues have asymmetrical bone structures. Measurement for attenuation correction can be done in two ways. The most accurate method is to perform a separate imaging for transmission and correct the sinograms with the transmission data. This is done in most commercial scanners by using rotating line or point sources. The other method is to use a mathematical approach based on determining the length of the photon range inside the surface contours. In combined PET-CT imaging systems, transmission image is conducted by CT.

Image Reconstruction

Data can be acquired both in two- and/ three-dimensional mode. However, independently of the data

acquisition mode, reconstruction is often done in a two-dimensional mode to obtain transverse, coronal, and/or sagittal slices with a variety of slice thicknesses depending on the statistical count distribution and biological questions. To reconstruct a data set in two-dimensional mode after three-dimensional acquisition requires the rebinning of the sinograms. The sinograms are corrected for uniformity, scatter, and attenuation before reconstruction of images. Also, corrections for acquisition time, decay, and injected radioactivity can be included with the corrections before reconstruction.

Reconstruction algorithms have been developed for both two- and three-dimensional approaches. The most commonly used algorithm is filtered backprojection. Also, convolution backprojection can be used to reconstruct two-dimensional data sets.^[4] Resolution in filtered backprojection, as well as in convolution backprojection, is affected mostly by the selection of the filter function and its cut-off value, which depends on the noise level of the data. Also, several arithmetic iterative reconstruction techniques are available in commercial PET scanners for two- and three-dimensional data analyses including the basic ordered-subsets expectation-maximization (OSEM) algorithm, attenuation-weighted OSEM (WA-OSEM) and segmented attenuation/weighted OSEM (WLS-OSEM) as well as Fourier binning OSEM (FORE). Even though the OSEM algorithm is used to suppress image noise and to make non-negative value images, it is not superior to the filtered backprojection in parametric images.^[31] Selection of the reconstruction algorithm depends mainly on the biological application.

Quantification

For accurate quantification, a good contrast is required to outline the borders of the different tissues. With the present multimodality imaging techniques it is possible to fuse PET images with anatomical magnetic resonance images and obtain the needed anatomical borderlines. A number of different biological models have been developed for quantification of PET imaging studies so that the final results can express biological function in metric units. There are generally approved models to quantify hemodynamics^[32] and glucose metabolism^[33] and several models for receptor binding, depending on binding mechanism.^[34-36] With the present PET imagers, the resolution is high enough to minimize partial volume effects.

BIOLOGICAL APPLICATION OF PET

The variety of biological applications of PET imaging is based on the superior sensitivity of detection of radioactivity. The obvious applications are studies of

Table 3 A summary of the most commonly used radioligands in PET imaging and their biological target

Physiological function	Radioligand	Disease or target	Reference
Hemodynamics			
Blood flow	C ¹⁵ O ₂ , H ₂ ¹⁵ O	Stroke, infarct, activation,	[32,37–41]
Blood volume	C ¹⁵ O, ¹¹ CO	cancer, myocardium	
Perfusion	¹³ NH ₃		
Metabolism			
Glucose metabolism	¹⁸ F-FDG	Cancer, stroke, infarct,	[33,42–44]
Oxygen metabolism	¹⁵ O ₂	Alzheimer's disease, Huntington's disease	
Amino acid metabolism	¹¹ C-methionine	Cancer	[45]
Fatty acid metabolism	¹¹ C-palmitic acid, ¹⁸ F-fluoro-6-thia-heptadecanoic acid	Myocardium	[46,47]
Receptor function			
Dopamine transporter	¹¹ C-CFT, ¹⁸ F-L-dopa	Parkinson's disease,	[34,48–50]
Dopamine D1 receptor	¹¹ C-SCH 23390	schizophrenia, drug addiction	
Dopamine D2 receptor	¹¹ C-raclopride		
5-HT (1A) receptor	¹¹ C-WAY 100635	Schizophrenia	[51]
Benzodiazepine receptor	¹¹ C-flumazenil	Epilepsy	[52]
Opioid receptor	¹¹ C-diprenorphine, ¹¹ C-carfentanil	Pain	[53,54]
Estrogen receptor	16 alpha-[¹⁸ F]fluoro-17beta-estradiol	Cancer	[55]
Molecular mechanism			
Proliferation potential	3-deoxy-3- ¹⁸ F-fluoro-thymidine	Cancer	[56]
DNA synthesis	¹¹ C-thymidine	Cancer	[57]
Herpes simplex virus type 1 thymidine kinase (HSV 1-tk)	¹⁸ F-5-fluorouracil, 8- ¹⁸ F-fluoroganciclovir, ¹⁸ F-acycloguanosine, ¹²⁴ I-FIAU	Cancer	[58–60]

functional mechanisms, where a tracer amount of radioactivity is introduced into the system without changing its physiological environment. Biological applications have developed together with radiopharmaceutical development, especially with the development of chemical synthesis and labeling techniques. Carbon-11 ($t_{1/2} = 20$ min) and fluorine-18 ($t_{1/2} = 110$ min) are the most commonly used positron emitters in these labeling procedures. Recently, there is also a trend to develop labeling techniques for small peptides using copper-64 ($t_{1/2} = 12.8$ hours) and iodine-124 ($t_{1/2} = 4.2$ days). The commonly used radiopharmaceuticals and their targeted applications are summarized in Table 3.

Imaging of Hemodynamics

Stroke

Detection of hemodynamic changes in brain blood flow has been one of the earliest clinical application of PET techniques.^[32,37] These techniques were used intensively in stroke studies to localize a site and extend, as well as

follow up, the recovery.^[38,39] Recently, magnetic resonance imaging (MRI) techniques have overcome many of these applications.^[61] When separate information is needed for blood flow or blood volume changes and the possible connection to oxygen consumption, only PET studies can provide information on all these hemodynamic aspects.^[62]

Activation studies

PET studies have been extensively used to evaluate hemodynamic changes induced by stimulations to evaluate the activation and connectivity in different brain circuitries. These stimulations can be introduced through all the senses, but visual stimulation is the most commonly used technique, combined with different tasks.^[40,63] Also, pure electrophysiological techniques can be used to activate specific brain areas.^[64]

Infarct

PET studies of hemodynamic changes in the heart are often combined with studies of glucose utilization to



obtain more clinically relevant information, especially differentiation of ischemic and infarcted tissue areas (Fig. 4).^[65] If only blood flow studies are done, it is necessary to repeat the studies during rest and exercise or combine blood flow studies with evaluation of wall motion in order to differentiate viable versus nonviable tissue areas.^[66]

Cancer detection

Most tumor tissues have an enhanced growth rate and they require a lot of energy, creating a need for an increased blood supply. This can be introduced by increased blood flow and/or blood volume. Blood flow is a sensitive indicator of changed hemodynamic requirements. PET studies of tumor blood flow and volume have been used as an indicator of efficacy of the treatment with antiangiogenic agents. A study of combretastatin, an antivascular agent, showed a 30–60% reduction in tumor blood flow in patients treated with a high dose of the drug.^[67] Similar studies were done to evaluate the antiangiogenic activity of endostatin.^[68] This type of assessment can be a new method to determine the efficacy and optimal dosage of antiangiogenic agents.

Imaging of Metabolism

The availability of fluorine-18-labeled 2-fluoro-2-deoxy-D-glucose (¹⁸F-FDG) has significantly enhanced PET

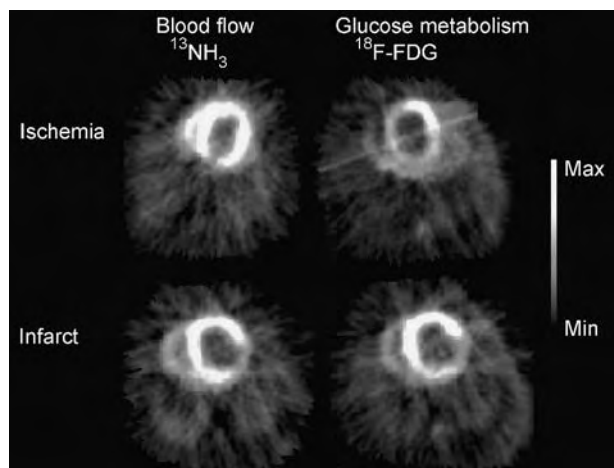


Fig. 4 Experimental studies of blood flow and glucose metabolism in the heart demonstrate the concept of matching information (infarct) and mismatching information (ischemia). In the ischemic process, the blood supply (flow) was decreased in the apical area (upper row, left), which shows an enhanced accumulation of ¹⁸F-2-fluoro-2-deoxy-D-glucose (upper row, right). In the infarct (lower row), both blood flow and glucose metabolism are similarly decreased.

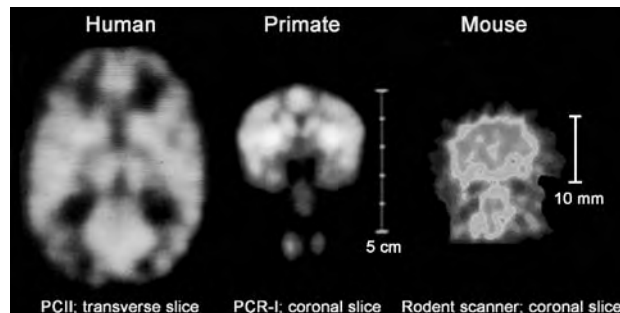


Fig. 5 Studies conducted of glucose metabolism with ¹⁸F-2-fluoro-2-deoxy-D-glucose in normal human, primate, and mouse brain. The human study shows a transverse slice imaged with PCII (spatial resolution of 10 mm). The primate study shows a coronal slice imaged with PCR-I (spatial resolution of 4.5 mm) and the mouse study shows a coronal slice imaged by the rodent scanner (Fig. 3). These images demonstrate the effect of the instrumentation development on the biological outcome.

studies based on glucose metabolism, as well as instrumentation development to the level of the concept of FDG-PET. The most important application of FDG-PET is in cancer studies, especially in whole-body imaging in order to find metastases.

Glucose is the main source of energy in the brain (Fig. 5). Therefore, all the changes in functionality create some changes in regional glucose metabolism. Even though glucose metabolism is a sensitive indicator of changed metabolism, it is not a selective method. Thus, the clinical relevance has to be obtained with other additional bases.

The tracer kinetic model of glucose consists of intravascular tracer, a tissue precursor pool in exchange with blood, and a metabolic pool representing a tracer that has been chemically transformed.^[69] The compartments are interrelated by the transport constants so that k_1 and k_2 represent exchange between blood and tissue (in the blood-brain barrier) and k_3 and k_4 represent the rate of tracer phosphorylation by hexokinase and dephosphorylation by glucose-6-phosphatase, respectively. There is an additional parameter, a lumped constant, which correlates the steady-state rate of tracer (¹⁸F-FDG) phosphorylation to the rate of glucose phosphorylation.^[33] The lumped constant has different values in different organs and it can also be affected by disease-induced stress in the tissue.^[69–71]

Glucose metabolic rate can be used as an index of malignancy of brain tumors,^[42] as well as to predict the response of the therapy.^[72] A study of high-grade gliomas showed that a decrease in ¹⁸F-FDG accumulation after 7 days of treatment could predict radiological response seen at 2 months.^[73]

Using ¹⁸F-FDG, it is possible to obtain invaluable information on all other biological studies except those

that are based on specific receptor recognition mechanisms.

Imaging of Receptor Function

Imaging of the receptor function has significantly enhanced the understanding of underlying pathophysiological mechanisms, especially in many neurological disorders. There are a number of specific ligands to investigate dopaminergic and serotonergic receptor functions (Table 3). PET studies of dopaminergic function have enabled an early diagnosis of Parkinson's disease,^[48] as well as the development and testing of therapeutic regimens for repairing the function via either pharmacological or surgical methods, including deep brain stimulation^[74] or cell transplantation.^[75] Dopaminergic function is affected in neurodegenerative disorders, schizophrenia,^[76] drug addiction,^[77] and in many psychological conditions such as anger,^[78] anxiety, and hyperactivity.^[79]

PET imaging is also used to assess the kinetics of drug–receptor interaction to determine the presence and level of in-vivo expression of receptors in tumors without the necessity of biopsy and to aid in selection of the appropriate therapy. PET studies of breast tumors using ¹⁸F-estrogen (FES), a radioligand for the estrogen receptor, correlate well with estrogen receptor status obtained from pathological samples.^[80,81] This correlation allows longitudinal in-vivo follow up of hormone receptor status and aid in the choice of therapy.

Imaging of Molecular Mechanism

In molecular imaging, ¹⁸F-labeled compounds with high specific activity have a central role, since biological processes and incorporation of ligands into the cells, genes, or DNA may take hours. For example, ¹⁸F gas is used in the synthesis of ¹⁸F-labeled acycloguanosine analogues and other compounds used to image Herpes simplex virus 1 thymidine kinase to follow gene therapy,^[58,59] and in the production of ¹⁸F-labeled 3'-deoxy-3'-fluorothymidine (FLT) to image proliferation potential in the cell.^[56] DNA synthesis can be imaged by ¹¹C-labeled thymidine.^[57] PET studies with ¹¹C-thymidine showed in patients receiving a variety of chemotherapy regimens a 100% decrease in ¹¹C-thymidine flux constant among responders, compared to a 35% decrease in partial responders and an even smaller decrease in non-responders after 7 days of treatment.^[82] Thus PET studies can be used to evaluate cell-based responses in chemotherapy.

Drug Development

PET has the potential to be a powerful tool in drug discovery and development. PET imaging is able to noninvasively assess drug distribution and action at the molecular level. Positron emitters can be introduced to the drug candidates and used as radiolabeled ligands. It has been well established that substituting a stable atom with a radioisotope of the same element does not affect the physicochemical, pharmacokinetic, or biological properties of a drug.^[83] Preliminary studies indicate that dynamic PET imaging, using repeated images over time, can be a valuable technique for defining the time course of uptake and retention of radiolabeled anticancer drugs in tumors and in the surrounding normal tissue in patients.^[84] These drugs are designed to inhibit key processes in cancer initiation and progression: angiogenesis, proliferation, avoidance of cell death or apoptosis, invasion and metastasis, and transduction of signals that modulate these processes. Many of these drugs are inhibitors of cell surface receptors or inhibitors of intracellular enzymes that transmit signals from receptors to the nucleus.

The current generation of PET scanners provide whole-body imaging and good spatial resolution to evaluate drug effects in small targets, and only trace amounts (nano or picograms) of the labeled ligand (drug) will be administered into the patient.

CONCLUSION

Development of positron imaging instrumentation has been a challenging effort in the past 50 years. Now is the time to establish its biological use. The future technical development relies mainly on developing faster and more efficient light detection, as well as even faster electronics. These aspects would increase sensitivity and improve image quality. In the clinical environment, PET has established its power in cancer studies. However, these studies utilize only a fraction of the possibilities that can be obtained by PET imaging. In the future, the main challenge of PET imaging is in drug development based on different biological and functional mechanisms. The new ligands and drugs can be used in humans to enhance early diagnosis and predict and follow the progression of the therapy. The biggest challenge is to develop new radioactive ligands to investigate in vivo different receptor and enzyme functions. For example, the amino acid glutamate mediates most of the excitatory synaptic transmissions in the brain and the glutamatergic pathways are involved in diverse processes like epilepsy, ischemic brain disease, pain, learning, etc. However, there is no radioactive ligand available for human studies to investigate in vivo glutamate receptor functions. The potential applications of

PET imaging in drug development have just begun to be studied and the preliminary successes of these approaches are evidence of the future power of PET imaging.

ACKNOWLEDGMENT

I would like to acknowledge Dr. Gordon L. Brownell and Mr. Charles Burnham for the brief discussions about PET instrumentation development.

ARTICLES OF FURTHER INTEREST

Magnetic Resonance Imaging; Medical Imaging, 3-D; Microcomputed Tomography and Its Applications; Real Time Tomographic Reflection

REFERENCES

- Sweet, W.H. The use of nuclear disintegration in diagnosis and treatment of brain tumors. *N. Engl. J. Med.* **1951**, *245*, 875–878.
- Brownell, G.L.; Sweet, W.H. Localization of brain tumors with positron emitters. *Nucleonics* **1953**, *11*, 40–45.
- Chesler, D.D. Three-dimensional activity distribution from multiple positron scintigraphs. *J. Nucl. Med.* **1971**, *12*, 347–348.
- Chesler, D.A. Positron Emission Tomography and Three-Dimensional Reconstruction Technique. In *Tomographic Imaging in Nuclear Medicine*; Freedman, G.S., Ed.; The Society of Nuclear Medicine: New York, 1973; 176–183.
- Ambrose, J. Computerized transverse axial scanning (tomography): Part II. Clinical application. *Br. J. Radiol.* **1973**, *46*, 1023–1047.
- Hounsfield, G.N. Computerized transverse axial scanning (tomography). Part 1: Description of system. Part II: Clinical applications. *Br. J. Radiol.* **1973**, *46*, 1016–1022.
- Cormack, A.M. Representation of a function by its line integrals, with some radiological applications. *J. Appl. Phys.* **1973**, *34*, 2722–2727.
- Cormack, A.M. Reconstruction densities from their projections, with applications in radiological physics. *Phys. Med. Biol.* **1973**, *18*, 195–207.
- Cherry, S.R.; Shao, Y.; Silverman, R.W.; Chatziioannou, A.; Meadows, K.; Siegel, S.; Farguhar, T.; Young, J.; Jones, W.F.; Newport, D.; Moyers, C.; Andreaco, M.; Paulus, M.; Binkley, D.; Nutt, R.; Phelps, M.E. MicroPET: A high resolution PET scanner for imaging small animals. *IEEE Trans. Nucl. Sci.* **1997**, *44*, 1161–1166.
- Correia, J.; Burnham, C.A.; Kaufman, D.; Fischman, A. Development of a small animal PET imaging device with resolution approaching 1 mm. *IEEE Trans. Nucl. Sci.* **1999**, *46* (3), 631–635.
- Melcher, C.L. Lutetium Orthosilicate Single Crystal Scintillator Detector. US Patent 4,958,080, 1990.
- www.itnonline.net. “Product Comparison Chart” or www.reillycomm.com/cgi-bin/proj/compare2.cgi.
- Burnham, C.A.; Bradshaw, J.; Kaufman, D.; Chesler, D.A.; Brownell, G.L. Positron source position sensing detector and electronics. US Patent 4,531,058, July 23, 1985.
- Casey, M.; Nutt, R. A multislice two-dimensional BGO detector system for PET. *IEEE Trans. Nucl. Sci.* **1986**, *33*, 760–763.
- Brownell, G.L.; Burnham, C.A.; Chesler, D.A. High Resolution Tomograph Using Analog Coding. In *The Metabolism of the Human Brain Studies with Positron Emission Tomography*; Greitz, T., Ingvar, D.H., Widen, L., Eds.; Raven Press: New York, 1985; 13–19.
- Schmand, M.; Dahlbom, M.; Eriksson, L.; Casey, M.E.; Andreaco, M.S.; Vagneur, K.; Phelps, M.E.; Nutt, R. Performance of a LSO/NaI(Tl) phoswich detector for a combined PET/SPECT imaging system. *J. Nucl. Med.* **1998**, *39* (5), 9.
- Brownell, G.L. US Atomic Energy Commission 1971. (Inspection document).
- Brownell, G.L.; Burnham, C.A. MGH Positron Camera. In *Tomographic Imaging in Nuclear Medicine*; Freedman, G.S., Ed.; The Society of Nuclear Medicine: New York, 1973; 154–164.
- Chesler, D.A.; Hoop, B., Jr.; Brownell, G.L. Transverse section imaging of myocardium with $^{13}\text{NH}_4$. *J. Nucl. Med.* **1973**, *14*, 623.
- Robertson, J.S.; Marr, R.B.; Rosenblum, M.; Radeka, V.; Yamamoto, Y.L. 32-Crystal Positron Transverse Section Detector. In *Tomographic Imaging in Nuclear Medicine*; Freedman, G.S., Ed.; The Society of Nuclear Medicine: New York, 1973.
- Cho, Z.H.; Chan, J.K.; Eriksson, L. Circular ring transverse axial positron camera for 3-dimensional reconstruction of radionuclide distribution. *IEEE Trans. Nucl. Sci.* **1976**, *23*, 613–623.
- Derenzo, S.E.; Budinger, T.F.; Cahoon, J.L.; Greenberg, W.L.; Huesman, R.; Vuletich, T. The Donner 280-crystal high resolution positron tomograph. *IEEE Trans. Nucl. Sci.* **1979**, *26*, 2790–2793.
- Ter-Pogossian, M.M.; Phelps, M.E.; Hoffman, E.J. A positron emission transaxial tomograph for nuclear medicine imaging (PETT). *Radiology* **1975**, *114*, 89–98.
- Brownell, G.; Burnham, C.A.; Stearns, C.W.; Chesler, D.A.; Brownell, A.-L.; Palmer, M. Development in high-resolution positron emission tomography at MGH. *Int. J. Imaging Syst. Technol.* **1989**, *1*, 207–217.
- NEMA. *Performance Measurements of Positron Emission Tomographs*; National Electrical Manufacturers Association: Rosslyn, VA, 2001.
- Hoffman, E.; Huang, S.C.; Phelps, M.E. Quantitation in positron emission tomography. I—Effect of object size. *J. Comput. Assist. Tomogr.* **1979**, *3*, 299–308.
- Ficke, D.C.; Beecher, D.E.; Bergmann, S.R.; Hoffman, G.R.; Hood, J.T.; Ter-Pogossian, M.M. In *Performance Characterization of a Whole Body PET System Designed for Dynamic Cardiac Imaging*, IEEE Nuclear Science Symposium and Medical Imaging Conference, Santa Fe, NM, 1991.

28. Bohm, C.; Eriksson, L.; Bergstrom, M.; Litton, J.; Sundman, R.; Singh, M. A computer assisted ring-detector positron camera system for reconstruction tomography of the brain. *IEEE Trans. Nucl. Sci.* **1978**, *25*, 624–637.
29. Bendriem, B.; Soussaline, F.; Campagnolo, R.; Verrey, B.; Wajnberg, P.; Syrota, A. A technique for correction of scattered radiation in a PET system using time-of-flight information. *J. Comput. Assist. Tomogr.* **1986**, *10*, 287–295.
30. Bergstrom, M.; Eriksson, L.; Bohm, C.; Blomqvist, G.; Litton, J. Correction for scattered radiation in a ring detector positron camera by integral transformation of the projections. *J. Comput. Assist. Tomogr.* **1983**, *7*, 42–50.
31. Oda, K.; Toyama, H.; Uemura, K.; Ikoma, Y.; Kimura, Y.; Senda, M. Comparison of parametric FBP and OS-EM reconstruction algorithm images for PET dynamic study. *Ann. Nucl. Med.* **2001**, *15*, 417–423.
32. Subramanyam, R.; Alpert, N.M.; Hoop, B., Jr.; Brownell, G.L.; Taveras, J.M. A model for regional cerebral oxygen distribution during continuous inhalation of $^{15}\text{O}_2$, C^{15}O , and C^{15}O_2 . *J. Nucl. Med.* **1978**, *19*, 43–53.
33. Phelps, M.E.; Huang, S.C.; Hoffman, E.J.; Selin, C.; Sokoloff, L.; Kuhl, D. Tomographic measurement of local cerebral glucose metabolic rate in humans with (F-18-2-fluoro-2-deoxy-D-glucose: Validation of method. *Ann. Neurol.* **1979**, *6*, 371–388.
34. Farde, L.; Hall, H.; Ehrin, E.; Sedvall, G. Quantitative analysis of D2 dopamine receptor binding in the living human brain by PET. *Science* **1986**, *231*, 258–260.
35. Delforges, J.; Loc'h, C.; Hantraye, P.; Stulzaf, O.; Khalili-Varasteh, M.; Maziere, M.; Syrota, A.; Maziere, B. Kinetic analysis of central (76Br)bromolisuride binding to dopamine D2-receptor studies by PET. *J. Cereb. Blood Flow Metab.* **1991**, *11*, 914–925.
36. Gjedde, A.; Wong, D.F. Modeling Neuroreceptor Binding of Radioligands in vivo. In *Quantitative Imaging: Neuroreceptors Neurotransmitters and Enzymes*; Frost, J., Wagner, H.W., Jr., Eds.; Raven Press: New York, 1990; 51–79.
37. Jones, T.; Chesler, D.A.; Ter-Pogossian, M.M. The continuous inhalation of oxygen-15 for assessing regional oxygen extraction fraction in the brain of man. *Br. J. Radiol.* **1976**, *49*, 339–343.
38. Lenzi, G.I.; Frackowiak, R.S.J.; Jones, T. Cerebral oxygen metabolism and blood flow in human cerebral ischemic infarction. *J. Cereb. Blood Flow Metab.* **1982**, *2*, 321–335.
39. Rubin, G.; Firlik, A.D.; Levy, E.I.; Pindzola, R.R.; Yonas, H. Relationship between cerebral blood flow and clinical outcome in acute stroke. *Cerebrovasc. Disord.* **2000**, *10* (4), 298–306.
40. Brooks, D.J. Cerebral blood flow activation studies in Parkinson's disease. *Parkinson's Dis. Adv. Neurol.* **2001**, *86*, 225–235.
41. Akutsu, Y.; Harumi, K.; Michihata, T.; Watanabe, T.; Yamanaka, H.; Okazaki, O.; Kashida, M.; Hasegawa, M.; Katagiri, T. Correlations between resting regional wall motion and regional myocardial blood flow (at rest and during exercise) in infarct-related myocardium—A study with ^{13}N ammonia positron emission tomography. *Jpn. Circ. J.* **1997**, *61*, 665–672.
42. Di Chiro, G.; DeLaPaz, R.L.; Brooks, R.A.; Sokoloff, L.; Kornblith, P.L.; Smith, B.H.; Patronas, N.J.; Kufta, C.V.; Kessler, R.M.; Johnston, G.S.; Manning, R.G.; Wolf, A.P. Glucose utilization of cerebral gliomas measured by ^{18}F fluorodeoxyglucose and positron emission tomography. *Neurology* **1982**, *32* (12), 1323–1329.
43. Di Chiro, G. Which PET radiopharmaceutical for brain tumors? *J. Nucl. Med.* **1991**, *32* (7), 1346–1348.
44. Frackowiak, R.; Lenzi, G.L.; Jones, T.; Heather, J.D. Quantitative measurement of regional cerebral blood flow and oxygen metabolism in man using ^{15}O and positron emission tomography: Theory, procedure and normal values. *J. Comput. Assist. Tomogr.* **1980**, *4*, 727–736.
45. Ericson, K.; Lilja, A.; Bergstrom, M.; Collins, V.P.; Eriksson, L.; Ehrin, E.; von Holst, H.; Lundqvist, H.; Langstrom, B.; Mosskin, M. Positron emission tomography with ^{11}C -methyl-L-methionine, ^{11}C -D-glucose and ^{68}Ga -EDTA in supratentorial tumors. *J. Comput. Assist. Tomogr.* **1985**, *9*, 683–689.
46. Geltman, E.M. Assessment of myocardial fatty acid metabolism with 1- ^{11}C -palmitate. *J. Nucl. Cardiol.* **1994**, *1* (Suppl), 15–22.
47. Taylor, M.; Wallhaus, T.R.; Degrado, T.R.; Russell, D.C.; Stanko, P.; Nickles, R.J.; Stone, C.K. An evaluation of myocardial fatty acid and glucose uptake using PET with ^{18}F -fluoro-6-thia-heptadecanoic acid. *J. Nucl. Med.* **2001**, *42*, 55–62.
48. Frost, J.J.; Rosier, A.J.; Reich, S.G.; Smith, J.S.; Ehlers, M.D.; Snyder, S.H.; Ravert, H.T.; Dannals, R.F. Positron emission tomographic imaging of the dopamine transporter with ^{11}C -WIN 35,428 reveals marked decline in mild Parkinson's disease. *Ann. Neurol.* **1993**, *34*, 423–431.
49. Hantraye, P.; Brownell, A.-L.; Elmaleh, D.R.; Spealman, R.D.; Wullner, U.; Brownell, G.L.; Isacson, O. Dopamine fiber detection by ^{11}C -CFT and PET in a primate model of parkinsonism. *NeuroReport* **1992**, *3*, 265–268.
50. Weeks, R.A.; Piccini, P.; Harding, A.E.; Brooks, D.J. Striatal D1 and D2 dopamine receptor loss in asymptomatic mutation carriers of Huntington's disease. *Ann. Neurol.* **1996**, *40* (1), 49–54.
51. Juhasz, C.; Nagy, F.; Watson, C.; da Silva, E.A.; Muzik, O.; Chugani, D.C.; Shah, J.; Chugani, H.T. Glucose and ^{11}C flumazenil positron emission tomography abnormalities of thalamic nuclei in temporal lobe epilepsy. *Neurology* **1999**, *53*, 2037–2045.
52. Martinez, D.; Hwang, D.; Mawlawi, O.; Slifstein, M.; Kent, J.; Simpson, N.; Parsey, R.V.; Hashimoto, T.; Huang, Y.; Shinn, A.; Van Heertum, R.; Abi-Dargham, A.; Caltabiano, S.; Malizia, A.; Cowley, H.; Mann, J.J.; Laruelle, M. Differential occupancy of somatodendritic and postsynaptic 5HT(1A) receptors by pindolol: A dose-occupancy study with ^{11}C WAY 100635 and positron emission tomography in humans. *Neuropsychopharmacology* **2001**, *24* (3), 209–229.
53. Jones, A.; Kitchen, N.D.; Watabe, H.; Cunningham, V.J.; Jones, T.; Luthra, S.K.; Thomas, D.G. Measurement of



- changes in opioid receptor binding in vivo during trigeminal neuralgic pain using [11C] diprenorphine and positron emission tomography. *J. Cereb. Blood Flow Metab.* **1999**, *19*, 803–808.
54. Saji, H.; Tsutsumi, D.; Magata, Y.; Iida, Y.; Konishi, J.; Yokoyama, A. Preparation and biodistribution in mice of [11C]carfentanil: A radiopharmaceutical for studying brain mu-opioid receptors by positron emission tomography. *Ann. Nucl. Med.* **1992**, *6*, 63–67.
 55. Mortimer, J.E.; Dehdashti, F.; Siegel, B.A.; Katzenellenbogen, J.A.; Fracasso, P.; Welch, M.J. Positron emission tomography with 2-[18F]fluoro-2-deoxy-D-glucose and 16alpha-[18F]fluoro-17beta-estradiol in breast cancer: Correlation with estrogen receptor status and response to systemic therapy. *Clin. Cancer Res.* **1996**, *2*, 933–939.
 56. Shields, A.F.; Grierson, J.R.; Dohmen, B.M.; Machulla, H.J.; Stayanoff, J.C.; Lawhorn-Crews, J.M.; Obradovich, J.E.; Muzik, O.; Mangner, T.J. Imaging proliferation in vivo with [F-18]FLT and positron emission tomography. *Nat. Med.* **1998**, *4* (11), 1334–1336.
 57. Wells, P.; Gunn, R.N.; Alison, M.; Steel, C.; Golding, M.; Ranicar, A.S.; Brady, F.; Osman, S.; Jones, T.; Price, P. Assessment of proliferation in vivo using 2-[(11C)]thymidine positron emission tomography in advanced intra-abdominal malignancies. *Cancer Res.* **2002**, *62* (20), 5698–5702.
 58. Tjuvajev, J.; Chen, S.H.; Joshi, A.; Joshi, R.; Guo, Z.S.; Balatoni, J.; Ballon, D.; Koutcher, J.; Finn, R.; Woo, S.L.; Blasberg, R.G. Imaging adenoviral-mediated herpes virus thymidine kinase gene transfer and expression in vivo. *Cancer Res.* **1999**, *59*, 5186–5193.
 59. Gambhir, S.; Barrio, J.R.; Herschman, H.R.; Phelps, M.E. Assays for noninvasive imaging of reporter gene expression. *J. Nucl. Med.* **1999**, *26*, 481–490.
 60. Gambhir, S.; Barrio, J.R.; Wu, L.; Iyer, M.; Namavari, M.; Satyamurthy, N.; Bauer, E.; Parrish, C.; MacLaren, D.C.; Borghei, A.R.; Green, L.A.; Sharfstein, S.; Berk, A.J.; Cherry, S.R.; Phelps, M.E.; Herschman, H. Imaging of adenoviral-directed herpes simplex virus type 1 thymidine kinase reporter gene expression in mice with radiolabeled ganciclovir. *J. Nucl. Med.* **1998**, *39*, 2003–2011.
 61. Schellinger, P.D.; Jansen, O.; Fiebich, J.B.; Pohlers, O.; Rysse, H.; Heiland, S.; Steiner, T.; Hacke, W.; Sartor, K. Feasibility and practicality of MR imaging of stroke in the management of hyperacute cerebral ischemia. *Am. J. Neuroradiol.* **2000**, *21* (7), 1184–1189.
 62. Mintun, M.A.; Lundstrom, B.N.; Snyder, A.Z.; Vlassenko, A.G.; Shulman, G.I.; Raichle, M.E. Blood flow and oxygen delivery to human brain during functional activity: Theoretical modeling and experimental data. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 6859–6864.
 63. Ito, H.; Takahashi, K.; Hatazawa, J.; Kim, S.G.; Kanno, I. Changes in human regional cerebral blood flow and cerebral blood volume during visual stimulation measured by positron emission tomography. *J. Cereb. Blood Flow Metab.* **2001**, *21*, 608–612.
 64. Fukuda, M.; Mentis, M.; Ghilardi, M.F.; Dhawan, V.; Antonini, A.; Hammerstad, J.; Lozano, A.M.; Lang, A.; Lyons, K.; Koller, W.; Ghez, C.; Eidelberg, D. Functional correlates of pallidal stimulation for Parkinson's disease. *Ann. Neurol.* **2001**, *49*, 155–164.
 65. Marshall, R.C.; Tillisch, J.H.; Phelps, M.E.; Huang, S.C.; Carson, R.; Henze, E.; Schelbert, H.R. Identification and differentiation of resting myocardial ischemia and infarction in man with positron computed tomography, 18F-labeled fluorodeoxyglucose and N-13 ammonia. *Circulation* **1983**, *67* (4), 766–778.
 66. Kitsiou, A.N.; Bacharach, S.L.; Bartlett, M.L.; Srinivasan, G.; Summers, R.M.; Quyyumi, A.A.; Dilsizian, V. 13N-ammonia myocardial blood flow and uptake: Relation to functional outcome of asynergic regions after revascularization. *J. Am. Coll. Cardiol.* **1999**, *33*, 678–686.
 67. Anderson, H.; Yap, J.; Price, P.M. Measurement of tumour and normal tissue (NT) perfusion by positron emission tomography (PET) in the evaluation of anti-vascular therapy: Results in the phase I study of Combretastatin A4 Phosphate (CA4P). *Proc. Ann. Meet. Am. Soc. Clin. Oncol.* **2000**, *9*, 179a.
 68. Herbst, R.S.; Mullani, N.A.; Davis, D.W.; Hess, K.R.; McConkey, D.J.; Charnsangavej, C.; O'Reilly, M.S.; Kim, H.W.; Baker, C.; Roach, J.; Ellis, L.M.; Rashid, A.; Pluda, J.; Bucana, C.; Madden, T.L.; Tran, H.T.; Abbruzzese, J.L. Development of biologic markers of response and assessment of antiangiogenic activity in a clinical trial of human recombinant endostatin. *J. Clin. Oncol.* **2002**, *20*, 3804–3814.
 69. Sokoloff, L.; Reivich, M.; Kennedy, C.; Des Rosiers, M.H.; Patlak, C.S.; Pettigrew, K.D.; Sakurada, O. The (C-14) deoxy glucose method for the measurement of local cerebral glucose utilization: Theory, procedure, the normal values in the conscious and anesthetized albino rat. *Neurochemistry* **1977**, *28*, 897–916.
 70. Ratib, O.; Phelps, M.E.; Huang, S.C.; Henze, E.; Selin, C.E.; Schelbert, H.R. Positron tomography with deoxyglucose for estimating local myocardial glucose metabolism. *J. Nucl. Med.* **1982**, *23*, 577–586.
 71. Kapoor, R.; Spence, A.M.; Muzi, M.; Graham, M.M.; Abbott, G.L.; Krohn, K. Determination of the deoxyglucose and glucose phosphorylation and the lumped constant in rat brain and a transplantable rat glioma. *J. Neurochem.* **1989**, *53*, 37–44.
 72. Bokenmeyer, C.; Kollmannsberger, C.; Oechsle, K.; Dohmen, B.M.; Pfannenberger, A.; Claussen, C.D.; Bares, R.; Kanz, L. Early prediction of treatment response to high-dose salvage chemotherapy in patients with relapsed germ cell cancer using [18F]FDG-PET. *Br. J. Cancer* **2002**, *86*, 506–511.
 73. Brock, C.S.; Young, H.; O'Reilly, S.M.; Matthews, J.; Osman, S.; Evans, H.; Newlands, E.S.; Price, P.M. Early evaluation of tumour metabolic response using [18F]fluorodeoxyglucose and positron emission tomography: A pilot study following the phase II chemotherapy schedule for temozolomide in recurrent high-grade gliomas. *Br. J. Cancer* **2000**, *82* (3), 608–615.
 74. Benabid, A.L.; Pollak, P.; Gao, D.; Hoffman, D.; Limousin, P.; Gay, E. Chronic electrical stimulation of the ventralis intermedialis nucleus of the thalamus as a

- treatment of movement disorders. *J. Neurosurg.* **1996**, *84*, 203–214.
75. Lindvall, O.; Sawle, G.; Widner, H.; Rothwell, J.C.; Bjorklund, A.; Brooks, D.; Brundin, P.; Frackowiak, R.S.; Marsden, C.D.; Odin, P.; Rehncrona, S. Evidence of long-term survival and function of dopaminergic grafts in progressive Parkinson's disease. *Ann. Neurol.* **1994**, *32*, 172–180.
76. Frankle, W.G.; Laruelle, M. Neuroreceptor imaging in psychiatric disorders. *Ann. Nucl. Med.* **2002**, *16*, 437–446.
77. Volkow, N.D.; Fowler, J.S.; Wang, G.J. Role of dopamine in drug reinforcement and addiction in humans: Results from imaging studies. *Behav. Pharmacol.* **2002**, *13*, 355–366.
78. Dougherty, D.D.; Shin, L.M.; Alpert, N.M.; Pitman, R.K.; Orr, S.P.; Lasko, M.; Macklin, M.L.; Fischman, A.J.; Rauch, S.L. Anger in healthy men: A PET study using script-driven imagery. *Biol. Psychiatry* **1999**, *46*, 466–472.
79. Volkow, N.D.; Fowler, J.S.; Wang, G.; Ding, Y.; Gatley, S.J. Mechanism of action of methylphenidate: Insights from PET imaging studies. *J. Atten. Disord.* **2002**, *6* (Suppl. 1), 31–43.
80. Dehdashti, F.; Mortimer, J.E.; Siegel, B.A.; Griffeth, L.K.; Bonasera, T.J.; Fusselman, M.J.; Detert, D.D.; Cutler, P.D.; Katzenellenbogen, J.A.; Welch, M.J. Positron tomographic assessment of estrogen receptors in breast cancer: Comparison with FDG-PET and in vitro receptor assays. *J. Nucl. Med.* **1995**, *36*, 1766–1774.
81. Dehdashti, F.; Flanagan, F.L.; Mortimer, J.E.; Katzenellenbogen, J.A.; Welch, M.J.; Siegel, B.A. Positron emission tomographic assessment of “metabolic flare” to predict response of metastatic breast cancer to antiestrogen therapy. *Eur. J. Nucl. Med.* **1999**, *26*, 51–56.
82. Young, H.; Baum, R.; Cremerius, U.; Herholz, K.; Hoekstra, O.; Lammertsma, A.A.; Pruim, J.; Price, P. Measurement of clinical and subclinical tumour response using [18F]-fluorodeoxyglucose and positron emission tomography: Review and 1999 EORTC recommendations. European Organization for Research and Treatment of Cancer (EORTC) PET Study Group. *Eur. J. Cancer* **1999**, *35*, 1773–1782.
83. Mintun, M.A.; Welch, M.J.; Siegel, B.A.; Mathias, C.J.; Brodack, J.W.; McGuire, A.H.; Katzenellenbogen, J.A. Breast cancer: PET imaging of estrogen receptors. *Radiology* **1988**, *169*, 45–48.
84. Gupta, N.; Price, P.M.; Aboagye, E.O. PET for in vivo pharmacokinetic and pharmacodynamic measurements. *Eur. J. Cancer* **2002**, *38*, 2094–2107.

Protein Adsorption Modeling

Alan H. Goldstein
Alastair N. Cormack

New York State College of Ceramics at Alfred University, Alfred, New York, U.S.A.

INTRODUCTION

This article deals with theoretical aspects of the initial events underlying protein adsorption to the surface of materials used in the body (biomaterials). A wide range of approaches to simulate protein adsorption are currently under investigation. However, there is no *prima facie* reason to treat the forces controlling these interactions differently from *in-vivo* processes whereby proteins functionally adsorb to a wide range of naturally occurring surfaces, including the surfaces of other proteins. These processes would include: formation of protein quaternary structures, assembly of supramolecular complexes, and both peripheral and integral association with biological membranes, to name a few. In this article, we will consider the physicochemical bases of the initial adsorption process, briefly survey current strategies for computer simulation of adsorption, and then provide a specific example of the use of molecular mechanics and dynamics for atomistic simulation of protein adsorption to a materials surface in an aqueous environment.

OVERVIEW: THE PROPERTIES OF BIOMATERIALS ARE OFTEN THE PROPERTIES OF THEIR SURFACES

A Protein Is a Material

At this point in the biotechnology revolution, the time has come for biomaterials scientists and structural biochemists to work closely together to use existing atomistic/molecular computer modeling systems to develop a detailed understanding of the interface formed between a protein initially folded into its native tertiary structure and a biomaterials surface. Our database of known protein structures is growing dramatically, as is our ability to predict protein structures from genetic sequence data and atomic composition. The future of this branch of science is clearly targeted toward a comprehensive predictive capability based on primary protein sequence data. While reaching this goal may lie somewhat in the future, it is appropriate for biomaterials scientists to make a parallel effort to bring current materials computer modeling systems

into the mix so that we can begin to develop simulation capabilities with respect to the materials–protein interface. Such efforts must, of course, be paralleled by laboratory studies designed to test the output of such models.

Biomolecular–Materials Composites

It is interesting to ask this question: What engineering challenge do the following emerging technologies have in common?

- Fabrication of new biomimetic/bioreactive materials.
- Optimization of DNA and protein microarray density and sensitivity.
- Engineering of optimized microfluidics systems.
- Development of tissue engineering scaffolds, and hybrid devices such as stents or controlled drug release implants that include both biological and nonbiological materials.

One answer is that the optimized design and fabrication (engineering) of these systems require an understanding of the biomolecular–materials interface and, more specifically, an atomistic/molecular picture of the nanostructure of this interface. While this concept applies to any biomaterial that uses or interacts with biological macromolecules, this article will focus on proteins and their adsorption to surfaces of inorganic materials.

Both the fields of protein engineering and materials engineering have virtually unlimited potential to provide significant advances in science, medicine, and industry. It is logical to propose that, as our understanding of protein structure–function relationships continues to grow, the status of protein engineering will begin to approach that of the more mature fields of engineering where our knowledge of fundamental physicochemical properties permits precise structures to be created *de novo* with prescribed functions.^[1,2] A mature protein engineering capability will allow us to understand: 1) how proteins interact with nonbiological materials (classic biomaterials); and 2) how to optimize construction of hybrid materials comprised of proteins and nonbiological materials. A mature protein engineering

capability will also allow us to fabricate hybrid structures comprised of proteins and other materials such as polymers and glasses. Such composites will often be developed as biomaterials. However, one definition of a biomaterial endorsed by a consensus of experts is “a nonviable material used in a medical device intended to interact with biological systems.”^[3] This is really too broad a definition within the context of advanced materials science and biomedical engineering. We propose the name protein–material composite (or more generally biomolecular–material composite), to define a new class of materials based on emerging technologies that will allow researchers to fabricate entirely new compounds with unique or dramatically enhanced properties of great utility in science, medicine, and industry.

Examples of these biomolecular–materials composites already exist. The DNA microarray, usually a DNA–glass composite material, represents a dramatic example. Likewise, the line between microfluidic and nanofluidic systems will mainly be drawn based on interfacial phenomena that control the movement of biological macromolecules when the dimensions of the transport system are on the same scale as the molecules to be transported. From a bioengineering standpoint, an important first step in optimizing these types of systems is gaining the ability to understand and visualize what is occurring at the biomolecular–materials interface. This, in turn, requires the integration of existing atomistic computer simulation systems for materials and biopolymers (e.g., proteins, nucleic acids) to optimize our modeling and predictive capacity. Fabrication followed by micro- and nanostructural characterization will be required for validation of these models which, in turn, will finally result in production and use of these new biopolymer–material composites.

Molecular Modeling of Proteins

The numbers of proteins whose structures are known at the atomic–coordinate level is increasing at a dramatic rate (cf., Refs. [1,2], and <http://www.rcsb.org>). Of equal importance, structure–function analyses have revealed that the virtually unlimited range of protein functions results from permutations and combinations of a relatively limited set of structural motifs.^[1] Structural stability for a given protein is almost always imparted by some type of folding involving alpha-helices and beta-sheets packed together into a stable three-dimensional conformation based on energy minimization within the protein’s native environment. These folding motifs are well characterized for a large number of protein families and, in fact, are so common that we can now recognize superfamilies with over 200 members that use almost identical folding motifs to maintain structural stability (e.g., the globin superfamily). Atomistic and molecular

modeling of proteins has now reached the point where protein engineering is possible. Proteins with enhanced thermal stability or modified substrate specificity are now produced routinely. Catalytic antibody technology and other strategies make it possible to design protein surfaces to mimic nonbiological catalysts (cf., Ref. [2] and references therein).

Ab initio prediction of the native conformation of a protein remains one of the great challenges to modern biochemistry and molecular biology.^[2] Computer simulation is essential in our attempts to understand processes such as protein folding. There is reason to be optimistic with respect to the ultimate development of an ability to predict the native conformation of a protein from its amino acid sequence. Recently, Duan and Kollman^[4] have folded a 36 residue protein fragment by molecular dynamics simulation into a structure that resembles the native state. But even while significant barriers remain to the prediction of protein structure from calculations based on the primary amino acid sequence, X-ray crystal structure data (as well as NMR and other experimental techniques) currently provide the capability to produce accurate three-dimensional pictures of a wide variety of protein structures. It is reasonable to suggest that available high-resolution structures include all the spatial and electrochemical information necessary to predict how a protein will bind to a surface and/or the type of material surface required to retain, or even enhance, protein function. With the growing availability of these databases, it is proposed that the time is right for a concentrated effort aimed at modeling protein–materials interactions.

Is Protein Adsorption to a Material Surface Unique as a Physicochemical Phenomenon?

Whereas complex chemical reactions may occur between a biomolecule and other materials, there is no reason to assume that there are any unique forces governing the initial protein adsorption to a material’s surface. A protein is, after all, a polymeric material whose structure and properties are determined mainly by surface-mediated (interfacial) phenomena. From protein folding itself, to the binding of a substrate in the active site of an enzyme, to the polymerization of structural proteins such as actin to form microfilaments, standard noncovalent interactions play the roles of paramount importance. These interactions include all the usual suspects: hydrogen bonding, electrostatic bonding, and van der Waals interactions. A fourth category, often referred to as hydrophobic interactions, is really a manifestation of a balance between the previous three types of forces, an observation that is not meant to trivialize the importance of the hydrophobic effect in the aqueous environment of living systems.



When a protein approaches a materials surface the result is almost invariably adsorption.^[3] Adsorption literally means to gather on a surface in a condensed layer (be the condensate a gas, liquid, or dissolved substance). While undoubtedly practiced since antiquity, the term adsorption was first coined in the late 19th century, although the process itself was not widely appreciated until the 1940s and 1950s when activated carbon was first used for municipal water treatment. It is of interest to note that this term is not used with any degree of rigor and this lack of specificity has evolved, quite probably, from that early use of activated carbon as a common but nonspecific adsorbent. Like the famous quote by a Supreme Court justice about pornography, we know adsorption when we see it. However, what is actually occurring at the surface is usually a mystery. The related term chemisorption, often defined as adsorption involving a chemical linkage between adsorbent and adsorbate, is also notably lacking in mechanistic information. Like most of the biological processes involving proteins (catalysis itself being the major exception), the initial adsorption of a protein to a surface is assumed to be based on multiple weak noncovalent interactions. The number and types of such interactions will be a function of the protein, the materials surface, and the solvent (in addition to any effects of temperature and pressure).

SIMULATION MODELING OF PROTEIN ADSORPTION

General Considerations

In this section, we will review some characteristic examples of computer simulation strategies employed to model protein adsorption. Most molecular modeling studies involve three stages. In the first stage, methodology is selected within which to describe the intra- and intermolecular interactions in the system. The two methodologies used in molecular modeling are quantum mechanics and classical mechanics. These methodologies enable the energy of any arrangement of the atoms and molecules in the system to be calculated, and allow the modeler to determine how the energy of the system varies as the positions of the atoms and molecules change. The second stage of a molecular modeling study is the calculation itself, involving techniques such as energy minimization, molecular dynamics or Monte Carlo simulation, or a conformational search. The results of the calculation must be analyzed, not only to obtain properties, but also to check that it has been performed properly. Finally, experimental validation of the simulated output is required. This may, in fact, be the limiting factor in evaluating the utility of most theoretical studies.

As we will discuss, there are basically two approaches to the simulation of protein adsorption: the continuum and the atomistic. The continuum approach is based on solving Poisson's equation for a model of the electrostatic potential in and around a protein near a charged surface, using dielectric constants for both the protein and the electrolyte. The atomistic approach, as its name implies, treats the system atomistically. This strategy includes explicit atom-atom interactions, which are basically summed to provide the energy of the system, and include both Molecular Mechanics/Molecular Dynamics (MM/MD) and quantum mechanics. Usually, MM and MD are coupled iteratively in order to simulate molecular motion and energy minimization during the adsorption process. In general, the large size of the system (number of atoms) requires that quantum approaches are semiempirical and are reserved for small organics (e.g., amino acids or polypeptides) although larger systems are being explored, as will be discussed subsequently.

The Continuum Approach

In a classic example of the continuum approach, Yoon and Lenhoff^[5] attempted the computation of the electrostatic interaction energy between a protein and a charged surface. The interaction potential energy between the molecule and the surface was used to determine equilibrium states or to follow the dynamics by integration of the equation of motion for the molecule. The interactions were considered to be entirely electrostatic in nature, but were analyzed using continuum models in which the overall interaction was decomposed into parts, e.g., van der Waals (VDW) interaction, electrostatic double-layer interaction, hydrophobic interaction. The authors noted that these are usually treated as being independent and additive despite the fact that they might not be mutually exclusive, and alluded to the first successful example of this approach as being DLVO (Derjaguin-Landau-Verwey-Overbeek) theory, which is central to much of colloid science. In DLVO theory,^[5] the VDW and electrostatic double-layer interactions are combined. These authors made use of full three-dimensional protein structure to account for the anisotropy of the shape and function of the protein molecule in determining its electrostatic double layer interaction with a charged surface. Numerical results were presented for ribonuclease A, with and without the presence of a charged surface, and the potential field information was used to determine the electrostatic interaction energy. These authors used a continuum formulation that assumed the potential Φ^i in the interior B^i of the biomolecule to be governed by the Poisson equation with the dielectric constant taken to have a uniform value ϵ^i .

A set of n point charges q_k were placed at positions x_k inside the molecule and the linearized Poisson–Boltzmann equation was assumed to govern the potential Φ^e in the exterior region B^e .

Protein interactions with synthetic polymers have been extensively investigated by numerous techniques. The general consensus appears to be that, in most cases, binding results in monolayer formation and strong hydrophobic interactions are crucial for observed irreversibility. The most important question to be solved is that of the orientation and tertiary structure on the polymer surface. Waldman-Meyer and Knippel^[6] generated a surface charge density model for the structure and orientation of polymer-bound proteins. These workers used negatively charged polystyrene (PS) beads. Their model was based on a layer of maximal packing density able to cover the total polymer surface. They assigned 15 \AA^2 per sulfate group on PS, so that a typical $5 \mu\text{m}$ microsphere would have a $\sigma' = -38 \times 10^{-3} \text{ C/m}^2$ (where σ' = surface charge density). This, in turn, meant that the ionogenic part of the surface would be only 1.8%. The authors concluded that, to account for observed full coverage, interactions must be mainly entropy-driven. Measured adsorption values of approximately 1.8 mg/m^2 for human serum albumin (HSA, a classic aqueous protein), and approximately 12 mg/m^2 for immunoglobulin G (IgG, a much more hydrophobic protein), indicated close-packed monolayers of adjacent HSA or protruding IgG, respectively. Likewise, Lu and Park^[7] examined protein adsorption on polymer surfaces and calculated adsorption energies. These workers looked at the adsorption of hen eggwhite lysozyme (HEWL), trypsin, an immunoglobulin, Fab, and hemoglobin on five polymer surfaces: polystyrene, polyethylene, polypropylene, poly(hydroxyethyl methacrylate), and poly(vinyl alcohol). They used an approach (later modified by Noinville et al., which will be discussed subsequently) that calculated binding energies from -850 kJ mol^{-1} to $+600 \text{ kJ mol}^{-1}$ with an average interaction energy of -155 kJ mol^{-1} ($-36.89 \text{ kcal mol}^{-1}$). These workers also cited microcalorimetric experimental data that had shown a mean net heat of adsorption of approximately 7000 kJ mol^{-1} ($1666 \text{ kcal mol}^{-1}$) for human γ -(7s)-globulin adsorbed on glass. Noinville et al.^[8] extended this work by modeling protein adsorption on polymer surfaces through a large number of initial interaction states. Protein-poly(vinylimidazole) polymer interaction energies were calculated for 2592 configurations scanning (according to the authors) the whole space of possible orientations from the bulk down to the contact with the adsorbent. Their approach assumed that the total interaction potential energy was equal to the sum of the individual pair potentials between the atoms of the protein and those of the solid surface. The protein was treated as a rigid molecule and the effect of the electrolyte solution

was accounted for by a simple distance-dependent function. For hen eggwhite lysozyme (HEWL), minimum interaction energies ranged from $-7.39 \text{ kJ mol}^{-1}$ to $+69.97 \text{ kJ mol}^{-1}$ ($-1.76 \text{ kcal mol}^{-1}$ to $+16.65 \text{ kcal mol}^{-1}$). For calcium-depleted α -lactalbumin (ALC), minimum interaction energy ranged from $-133.8 \text{ kJ mol}^{-1}$ to $+12.06 \text{ kJ mol}^{-1}$ ($-31.84 \text{ kcal mol}^{-1}$ to $+2.9 \text{ kcal mol}^{-1}$). For comparison, the experimentally measured ΔG of adsorption for Ca-depleted ALC was $-0.85 \text{ kJ mol}^{-1}$ ($-0.2 \text{ kcal mol}^{-1}$).

Song and Forciniti^[9] used Monte Carlo simulations of peptide adsorption on solid surfaces, based on the assumption that hydrophobic interactions, electrostatic interactions, and van der Waals forces are the main driving forces for protein adsorption on solid surfaces. They recognized that one of major problems encountered in the simulation of biological systems is the large number of atomic pair interactions involved. To compensate for limited simulation capability, water was treated as a dielectric continuum, and proteins were represented by reduced atomic coordinates or lattice protein models. Proteins were further assumed to have a rigid structure in solution. Computer simulation studies of protein adsorption became even more complex with the inclusion of solid surfaces. The authors concluded that protein adsorption onto hydrophobic surfaces may be entropically driven with water playing the dominant role. The argument is that release of water of hydration from the hydrophobic surface upon adsorption increases the entropy of the system. This implies that the entropy gain by the solvent is greater than the entropy loss caused by protein localization at the surface. Using Monte Carlo methods, paired correlation function (PCF) analyses showed that the solvent must be considered when adsorption phenomena are simulated. Water molecules remained trapped, even when considering adsorption of a fully ionized aspartate residue to a surface of uniform positive charge. Our own results are in agreement with much of this argument (see the subsequent discussion).

The Importance of the Solvent

The problem of including water (or other electrolyte) in the simulation is not trivial. Daura et al.^[10] addressed the question of a solvent requirement for peptide folding per se. Their results clearly showed that the solvent cannot be ignored in folding simulations and that there are significant difficulties in implicitly modeling solvent effects in an atomic interaction function for the solute. Likewise, Andricioaei and Karplus^[11] conducted a calculation of entropy from covariance matrices of the atomic fluctuations during protein folding. These authors noted that the small free energy of protein

folding (on the order of 10 kcal mol^{-1} at 300 K for many proteins) involved a near cancellation between a large decrease in enthalpy and a large decrease in entropy. A corresponding balance occurs in ligand binding so that we may assume, at least *prima facie*, that a similar balance occurs in the initial adsorption of proteins to biomaterials surfaces.

Bujnowski and Pitt^[12] studied the water structure around enkephalin near a polyethylene (PE) surface using molecular dynamics. A molecular model was constructed simulating a pentapeptide protein, leucine enkephalin, near a crystalline PE surface surrounded by explicit water molecules. A molecular dynamics simulation of the model was performed to examine the forces exerted on the protein by the surface and the water molecules. These workers used the MD software packages of Insight II[®] and Discover[®], with a simulation temperature of 298 K. The system was relaxed during an energy minimization for 300 time steps, equilibrated for another 1000 time steps, and then allowed to run for 12,000 time steps, which simulated a real time of 18 ps (at 1.5 fs time step). However, the authors locked the α carbons in place to prevent the protein from making any gross translations or rotations. The results of the simulations showed that there is a relatively strong spatial and orientational structuring of the water directly adjacent to the PE surface, and a lesser amount of structuring near the surface. This triple layer effect is, of course, familiar to students of classical electrochemical theory,^[13] but does not appear explicitly in most computer simulations. As with Daura et al.,^[10] these authors concluded that solvent interactions in the system played a key role in the proteins' behavior near a surface. This conclusion is further strengthened by the data presented by Ben-Tal et al.^[14] These workers examined the theoretical and experimental aspects of a related problem involving binding of small basic peptides to membranes containing acidic lipids. They directly measured the binding of the polypeptides lys3, lys5, and lys7 to vesicles containing acidic phospholipids in a 100 mM monovalent aqueous electrolyte. The standard Gibbs free energy for the binding of these peptides was found to be 3, 5, and 7 kcal mol^{-1} , respectively. The electrostatic free energy of interaction was considered to arise from both long-range coulombic attraction between the positively charged peptide and the negatively charged lipid bilayer, and a short-range Born repulsion. These workers considered that, in this system, each basic residue contributes approximately 1 kcal mol^{-1} to the membrane binding energy under physiological conditions. An energy minimum was achieved when approximately 2.5 Å (i.e., one layer of water) existed between the van der Waals surfaces of the peptide and the lipid bilayer, reemphasizing the importance of a proper treatment of the solvent.

The Quantum Mechanics Approach to Modeling Protein Adsorption

An example of the semiempirical quantum mechanics approach may be seen in the work of West et al.^[15] The hypothesis for the development of the study was that, to understand and predict protein adsorption behavior, it was necessary first to understand the adsorptive behavior of each individual midchain amino acid segment making up the protein. By incrementally increasing the number of amino acid residues experimentally studied and numerically modeled, the authors conceived that protein adsorption behavior itself would be eventually understood. Their model predicted direct hydrogen bonding of lysine hydrogen to oxygen of the silica substrate (modeled as a ring structure) with an average adsorption value of $-3.6 \text{ kcal mol}^{-1}$. This is approximately $4 \times$ the accepted value for an H-bond in an aqueous environment. The authors then used molecular mechanics in order to show that the helical structure of the polylysine would allow only two active amine sites per polylysine 16mer. When the % silica was reduced to reflect actual composition (from 100% to 71–74%), a calculated theoretical value of approximately $-0.34 \text{ kcal mol}^{-1}$ was developed. This value was close to the measured adsorption isotherm values of $-0.23 \text{ kcal mol}^{-1}$ for a midchain lysine residue bound to silica spheres.

Because the experimental values were obtained with polymers of lysine and molecular mechanics predicted that only one in eight residues would be able to interact with the silica surface, the interaction of a single midchain L-Lysine with the silica surface was assumed by West et al.^[15] to be approximately $8 \times$ the average value of the experimentally measured value ($8 \times -0.23 \text{ kcal mol}^{-1}$) = $-1.84 \text{ kcal mol}^{-1}$. The experimental component of this study utilized 6-, 13-, 42-, and 120-mers of poly-L-lysine in PBS at pH 7.4.

A potential difficulty with the output of the QM simulation is that this particular work predicted direct contact between the amino acid residues (in this case poly-L-lysine) and the materials surface. Based on the work discussed in the preceding section, ignoring the contributions of the solvent may be a serious error, given the dominant role of water in both the native protein conformation and the biomaterials surface structure. The interested reader is referred to Latour, Jr.^[16] for further discussions of this strategy.

Forcefield-Based Molecular Mechanics and Molecular Dynamics to Simulate Protein Adsorption to Biomaterials Surfaces in Solvent

Heuristically, it is useful to consider protein adsorption as a specialized case of protein structure–function relationships. As noted earlier, surface-mediated interactions

hold the key to most protein functions, from binding of substrate prior to catalysis, to association with specific regions of the cell membrane *in vivo*. If we assume that protein adsorption to a biomaterials surface is a subset of this larger group of phenomena, then it is quite reasonable to apply many of the existing rules of protein structure and function to protein adsorption to a biomaterial. Of course, this does not mean that the problem is simple, but rather that no new (or hidden) variables require explicit treatment *a priori* in order to accurately simulate initial adsorption events. These are still treacherous waters, especially when one begins to validate the adsorption model, as shown by the study of Daggett and Levitt.^[17] These workers developed a model of the molten globule state of a protein from molecular dynamics simulations. Simulations proceeded with 2-fs time steps up to 550 ps with *T* up to 423 K. With *T* at 423 K, the average rms displacement of α carbon from the crystal structure coordinates increased to 5.1 Å. While disulfide bonds strongly constrained the protein at high *T* ($C\alpha$ rms deviation $\sim 50\%$ higher in the reduced protein at 423 K than in the oxidized protein), some simulated displacements still were as high as 18 Å. In spite of these large simulated displacements, experimental measurement and MD modeling were in agreement with respect to the amount of secondary structure retained, and even added, at these higher temperatures. Experimental measurements showed that secondary structure content ranged from 60%–200% when the native protein was thermally denatured to the molten globule state. This was seen in the MD simulations as well. As the authors pointed out, if the amount of secondary structure were used as a measure of unfolding, both simulations and experimental techniques such as circular dichroism (CD) could conceivably show almost no unfolding at 423 K (equivalent to 150°C).

These types of data should be taken as a warning to modelers with respect to validation of simulated protein adsorption to materials. Protein biochemists generally consider that folding to the molten globule state *in vivo* is very fast (milliseconds at most), whereas attainment of proper tertiary interactions is slow. One could assume that the reverse is true as well. This would indicate that it is highly unlikely that proteins will denature past the molten globule state on biomaterials surfaces, at least during the initial adsorption phase. In all probability, the protein will assume some slightly modified version of the native tertiary structure.

Protein Adsorption: A Standardized Simulation Strategy for Molecular Mechanics/Molecular Dynamics

One goal of our current work on protein adsorption to biomaterials has been to develop a generally applicable,

standardized approach to the simulation of protein adsorption. A prerequisite for this is the use of widely available software systems so that a relatively large number of workers may compare results within a unified theoretical framework. Toward this end, a step-by-step tutorial describing our entire algorithm has been made available at www.bmes.alfred.edu.

The essential problem associated with modeling the interaction between a protein and inorganic surface is the determination of suitable interatomic potentials. For the simulation of protein structure, a number of satisfactory and proven potential functions exist. We currently use the intermolecular and nonbonded potentials known as the consistent-valence forcefield, or CVFF.^[18] The magnesium oxide surface shown on our tutorial Web site is accurately modeled in a particularly simple manner wherein the atoms are considered completely nonpolarizable with integer atomic charges of +2.0 and –2.0 for the magnesium and oxygen ions, respectively.^[19]

The challenge is that well-established potentials are not available for describing the interactions between proteins and inorganic materials. One starting point is to consider all interactions between the protein and magnesium atoms to be purely coulombic, with the interactions between the atoms of the protein and the oxygen atoms of the inorganic surface represented by Lennard-Jones potentials^[18,20] with a single oxygen parameter that serves the various interactions through a combinatorial procedure. This tactic has been applied with some success in the case of organic molecules interacting with silica-related inorganic materials, such as zeolites.^[21–23] Unfortunately, with this approach, unreasonably close binding was found to occur between the magnesium atoms and the nitrogen, carboxylate oxygen, and carbonyl oxygen within the protein molecule since the coulombic interactions are, in these cases, attractive so that short-range potentials acting between the magnesium atoms of the surface, and nitrogen and oxygen atoms within the protein, cannot be neglected.

As discussed in preceding sections, we have concluded that proper treatment of the aqueous solvent is absolutely essential for a valid simulation. Another reason to choose MgO for our adsorption study was because a great deal of effort has been invested in the simulation of the interaction between water and magnesium oxide surfaces.^[24–26] For example, the work of McCarthy et al.^[25] provides highly tractable potential functions that are fully compatible with the CVFF forcefield as well as the surface model in use. However, these researchers represented the water molecule by a rigid SPC model,^[27] which we felt was too restrictive, so we chose to use the more flexible TIP3P model of water in our work. Details of the short-range potential parameters for the magnesium oxide surface as well as

for the adjusted surface–water interactions utilized in our present study are available by using the <Bio-materials Modeling> link at www.bmes.alfred.edu or by going directly to the modeling tutorial at: http://biot.alfred.edu/%7Elewis/BPTI_WEB_1/front_page.html.

The parameters apply to a simple Buckingham potential model,^[19] which means that it is only necessary to establish the interaction potentials between the surface and the remaining protein molecule atom types. Our first approximation was to use the potential parameters for the surface–water interaction for all other forms of hydrogen and oxygen within the protein. Also, interactions with nitrogen atoms were assumed to be equivalent to those with oxygen interactions, as has been done before.^[28] Finally, all other nonbonded interactions between magnesium and the protein molecule (i.e., carbon, sulfur, and phosphorus) were treated as purely coulombic while the single, CVFF oxygen potential type was used to represent nonbonded interactions with the surface oxygen atoms.

The goal is to simulate the interaction between a single small protein molecule and an (infinite) inorganic surface, within an aqueous environment. To this end, an eight-layer, magnesium oxide system, $67.3380 \text{ \AA} \times 67.3380 \text{ \AA}$ in surface dimensions, with periodic boundary conditions, was chosen to serve as the surface. The bottom three layers of the system were fixed at their bulk positions to more fully represent the effects of an infinite bulk away from the surface. Infinite space above the magnesium oxide was modeled by adjusting the *z*-direction boundary condition so as to leave the eventual surface–protein–water system effectively uncoupled from its periodic images. A *z*-dimension of 86.8 \AA was used for this purpose. The space above the surface is where the protein and its aqueous environment were placed.

The energy of interaction between protein molecule and the surface is usually evaluated as the difference between the energies of an interacting protein–surface system and a noninteracting system (in which the protein and surface are at, essentially, infinite separation) with both systems in aqueous environments. We and others (as per the previous discussion) have found that the interaction between protein and the inorganic surface is directly mediated by the intervening water molecules. Therefore, water is explicitly included as part of the system. In our work, all the energy calculations employed a group-based direct sum method with a 14.0 \AA cutoff. Therefore, rather than completely fill the space above the surface in the simulation box with water molecules, a 15.0 \AA layer of water was used to approximate an aqueous environment. Naturally, there remained a vacuum–water boundary at which some water molecules are not within an aqueous environment at all and we would expect the energy of these to deviate slightly from bulk values. However, considerable computational expense is saved in this way.

The usual method of evaluating the energy of the noninteracting system would involve calculating the energy of a single protein molecule encased within a 15.0 \AA layer of water and adding to that the energy, calculated separately, for the surface system covered with a 15.0 \AA layer of water. The interacting system energy would then be established by evaluating the energy of the combined protein–inorganic surface system collectively coated with a 15.0 \AA layer of water. The disadvantage of this approach may be seen when the energy of the water is evaluated. The numbers of water molecules within the noninteracting and interacting systems are not automatically equivalent. Even if they were deliberately made equal, the water–vacuum interface would not be the same in the two systems. Therefore, our noninteracting system was constructed within a single periodic system as will be described subsequently.

The BPTI (Bovine Pancreatic Trypsin Inhibitor) crystal structure 6PTI was downloaded in PDF format from the Protein Data Bank: <http://www.rcsb.org/pdb/>.

The BPTI file was modified as necessary (hydrogens added, capping, pH set to 7.0, etc.) and energy minimization and molecular dynamics on the hydrated molecule were conducted using the Accelrys Discover[®] software package with the CVFF forcefield function. The minimum energy structure was then inserted into the simulation box containing the magnesium oxide surface. To create the noninteracting system, the protein molecule was placed a minimum of 20 \AA away from the surface; the remainder of the simulation box was filled with water molecules to within a minimum of 10 \AA above the uppermost segments of the protein molecule. Figure 1 shows a postminimization visualization of the simulation box. The complete pre- and postminimization simulation boxes may be viewed at: http://biot.alfred.edu/%7Elewis/BPTI_WEB_1/system_set_up.html.

This resulted in a total of 7541 water molecules being included in the system. This system was then subject to energy minimization (coupled with MD as described subsequently) to find the minimum energy of the noninteracting system. The interacting system was constructed in an equivalent way, with the periodic system containing an equal number of water molecules, but in which the protein is much closer to the surface. This system was also subject to energy minimization (coupled with MD as described subsequently) to find the minimum energy of the noninteracting system. The resulting systems both contained a single, well-defined water–vacuum boundary that is nearly identical in each case and this fact allowed for a straightforward comparison of the two. The interaction energy is then quite simply the difference between the total energies of the interacting and noninteracting systems.

The amorphous nature of the protein and water molecules requires special handling for the energy calculations. Local minimization techniques alone are not sufficient to acquire a reasonable estimate of the minimum energy of such a system. For this purpose, an improved estimate was obtained by alternating 2.0 ps molecular dynamics simulations at 320.0 K with local minimization until consistent values were achieved. Consistency was typically obtained within five iterations and the lowest energy thus obtained over five additional iterations was taken as the minimum value. Detailed results of these investigations will be published elsewhere.^[29] However, Fig. 1 shows a characteristic output from our standardized protein adsorption modeling system. The protein bovine pancreatic trypsin inhibitor (BPTI) is adsorbed to the magnesium oxide surface in water with a binding energy of $-100.5 \text{ kcal mol}^{-1}$ for this starting orientation. This minimized structure is equivalent to System_2 at: http://biot.alfred.edu/%7EEllewis/BPTI_WEB_1/results.html.

There are several instructive aspects to the output of this simulation. First, it is significant to note that the simulation accurately produces double layer waters at the highly polar MgO surface. This aqueous double layer, in turn, mediates the initial adsorption of the BPTI. Tyrosine21 in the amino acid sequence of BPTI has been identified, to facilitate analysis of Fig. 2 versus Fig. 1. The orientation of Tyrosine21 also shows

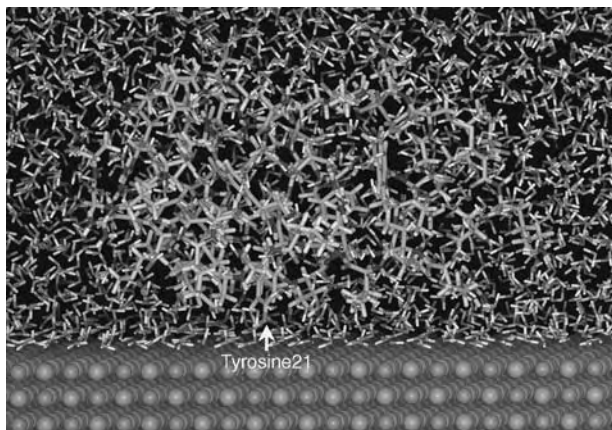


Fig. 1 Visualization of initial adsorption of the protein BPTI to magnesium oxide in pure water. This output shows one of many possible orientations of the protein relative to the surface. We have chosen to designate this as orientation number 2, which is identical to System_2 at: http://biot.alfred.edu/%7EEllewis/BPTI_WEB_1/results.html. Full materials and methods used for the simulation may be found at the linked Web pages. The protein is adsorbed to the hydrated magnesium oxide surface with a free energy of binding of approximately $-100.5 \text{ kcal mol}^{-1}$.

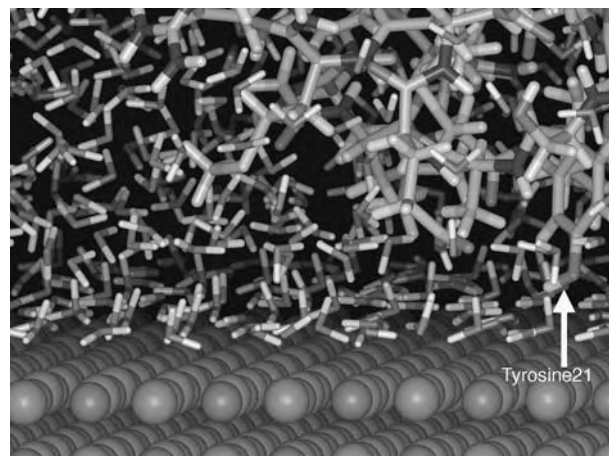


Fig. 2 Close-up of System_2 showing the hydrogen bond formed by Tyrosine21 of BPTI with double-layer waters bound tightly to the magnesium oxide surface. As discussed in the text, it is useful to note the clustering of protein hydrogens around the aromatic ring of this R group (hydrophobic shielding) as well as the difference in orientation of secondary hydration waters in regions of the magnesium oxide covered by the adsorbed protein versus the part of the magnesium oxide surface in contact with bulk solvent.

that even the most tightly bound residues of the protein (those within 5 Å of the material surface) initially interact through the double layer. Fig. 2 is a close-up of the region from Tyrosine21 along the MgO:BPTI interface into bulk solution. Several structure–function components of the simulation are revealed. Note the clustering of BPTI hydrogens around the aromatic ring of the Tyrosine R group. This hydrophobic shielding, in turn, allows the OH group of this residue to directly interact, presumably by H-bonding with double layer water at the surface. Therefore, in our simulations, the double layer mediates the initial adsorption event. Following the interface from adsorbed protein into bulk solution (right to left), it may be seen that, as expected, the secondary waters begin to take up positions intermediate in orientation between the protein surface and the bulk solvent while the double-layer waters remain intact. We believe that this is the first forcefield-based biomaterials modeling system to consider adequately the solvent in generating an atomistic picture of initial adsorption events. Note the emphasis on the term initial. Dramatic postadsorption rearrangement of proteins are routinely reported in the literature. Such rearrangement may certainly include direct contact and/or covalent bond formation between the biomolecule and the material surface and even changes in the chemical nature of the entire interphase region, as is known to occur with Bioglass[®].^[14]

CONCLUSION

Characterization of the structure and functionality of biomolecular–materials composites will play a key role in the development of both next-generation materials for biomedical engineering and biotechnology systems including genomics- and proteomics-enabling instrumentation. Given the importance of the problem, it is essential that standardized computer simulation software be developed that, in turn, may be used as CAD tools for the engineering of these advanced biomaterials and other devices. Based on the theoretical considerations and calculations described here, it is concluded that, conceptually, protein adsorption will be governed by the same structure–function rules that are known to apply in other areas of protein physical biochemistry. What is lacking is a seamless integration of current highly sophisticated protein and materials modeling software systems that explicitly takes the solvent into account. A brief summary of several current approaches to simulating adsorption have been presented, including efforts in the authors' laboratory to use customized commercially available software suites to simulate initial protein adsorption events.

ACKNOWLEDGMENTS

This work was supported in part by National Institutes of Health AREA Grant 1r15GM55947-0121 and the Norman and Evelyn Fierer endowment at Alfred University. The authors wish to acknowledge the essential contributions of Mr. Raymond Lewis to the simulation work presented in the last section.

ARTICLES OF FURTHER INTEREST

Biomaterials: Protein–Surface Interactions; Surface Coatings; Surface Modification; Surface Topography and Cell Behavior

REFERENCES

1. Branden, C.; Tooze, J. *Introduction To Protein Structure*, 1st Ed.; Garland: New York, 1991; 1–410.
2. Cleland, J.L. Introduction to Protein Engineering. In *Protein Engineering: Principles and Practice*, 1st Ed.; Cleland, J.L., Craik, C.S., Eds.; Wiley-Liss: New York, 1996; 1–32.
3. Horbett, T.A.; Ratner, B.D.; Schakenraad, J.M.; Schoen, F.J. *Biomaterials Science: An Introduction to Materials in Medicine*, 1st Ed.; Ratner, B.D., Hoffman, A.S., Schoen, F.J., Lemons, J.E., Eds.; Academic Press: New York, 1996; 133–141.
4. Duan, Y.; Kollman, P.A. Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. *Science* **1998**, *282*, 740–744.
5. Yoon, B.J.; Lenhoff, A.M. Computation of the electrostatic interaction energy between a protein and a charged surface. *J. Phys. Chem.* **1992**, *96*, 3130–3134.
6. Waldman-Meyer, H.; Knippel, E. A surface charge density model for structure and orientation of polymer-bound proteins. *J. Colloid Interface Sci.* **1992**, *148*, 508–516.
7. Lu, D.R.; Park, K. Protein adsorption on polymer surfaces: Calculation of adsorption energies. *J. Biomater. Sci., Polym. Ed.* **1990**, *1*, 243–260.
8. Noinville, V.; Vidal-Madjar, C.; Sebille, B. Modeling of protein adsorption on polymer surfaces: Computation potential. *J. Phys. Chem.* **1995**, *99*, 1516–1522.
9. Song, D.; Forciniti, D. Monte Carlo simulations of peptide adsorption on solid surfaces. *J. Chem. Phys.* **2001**, *115*, 8089–8900.
10. Daura, X.; Mark, A.E.; van Gunsteren, W.F. Peptide folding simulations: No solvent required? *Comput. Phys. Commun.* **1999**, *123*, 97–102.
11. Andricioaei, I.; Karplus, M. On the calculation of entropy from covariance matrices of the atomic fluctuations. *J. Chem. Phys.* **2001**, *115*, 6289–6292.
12. Bujnowski, A.M.; Pitt, W.G. Water structure around enkephalin near a PE surface: A molecular dynamics study. *J. Colloid Interface Sci.* **1998**, *203*, 47–58.
13. Bockris, J.O'M.; Reddy, A.K.N.; Gamboa-Aldeco, M. *Modern Electrochemistry: Ionics*, 2nd Ed.; Plenum Pub. Corp.: New York, 2001; 1–2053.
14. Ben-Tal, N.; Honig, B.; Peitzsch, R.M.; Denisov, G.; McLaughlin, S. Binding of small basic peptides to membranes containing acidic lipids: Theoretical models and experimental results. *Biophys. J.* **1996**, *71*, 561–575.
15. West, J.K.; Latour, R., Jr.; Hench, L.L. Molecular modeling study of adsorption of poly-L-lysine onto silica glass. *J. Biomed. Mater. Res.* **1997**, *37*, 585–591.
16. Latour, R.A., Jr. Molecular modeling of biomaterial surfaces. *Curr. Opin. Solid State Mater. Sci.* **1999**, *4*, 413–417.
17. Daggett, V.; Levitt, M.A. Model of the molten globule state from molecular dynamics simulations. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 5142–5146.
18. Dauber-Osguthorpe, P.; Roberts, V.A.; Osguthorpe, D.J.; Wolff, J.; Genest, M.; Hagler, A.T. Structure and energetics of ligand binding to proteins: *E. coli* dihydrofolate reductase-trimethoprim, a drug-receptor system. *Proteins* **1988**, *4*, 31–37.
19. Lewis, G.V.; Catlow, C.R.A. Potential models for ionic oxides. *J. Phys. C. Solid State Phys.* **1985**, *18*, 1149.
20. Kittel, C. *Introduction to Solid State Physics*, 7th Ed.; John Wiley and Sons, Inc.: New York, 1996; 56–65.
21. Kiselev, A.V.; Lopatkin, A.A.; Shulga, A.A. Molecular statistical calculation of gas adsorption by silicalite. *Zeolites* **1985**, *5*, 261.
22. Catlow, C.R.A.; Freeman, C.M.; Vessal, B.; Tomlinson, S.M.; Leslie, M. Molecular dynamics studies of hydrocarbon diffusion in zeolites. *J. Chem. Soc., Faraday Trans.* **1991**, *87*, 1947.
23. Sastre, B.G.; Catlow, C.R.A.; Chica, A.; Corma, A. Molecular dynamics of C7 hydrocarbon diffusion in

- ITQ-2. The benefits of zeolite structures containing accessible pockets. *J. Phys. Chem., B* **2000**, *104*, 416.
24. Langel, W.; Parrinello, M.J. Ab Initio molecular dynamics of H₂O adsorbed on solid MgO. *Chem. Phys.* **1995**, *103*, 3240.
25. McCarthy, M.I.; Schenter, G.K.; Scamehorn, C.A.; Nicholas, J.B. Structure and dynamics of the water/MgO interface. *J. Phys. Chem.* **1996**, *100*, 16989.
26. De Leeuw, N.H.; Watson, G.W.; Parker, S.C. Atomistic simulation of the effect of dissociative adsorption of water on the surface structure and stability of calcium and magnesium oxide. *J. Phys. Chem.* **1995**, *99*, 17219.
27. Berendsen, H.J.C.; Postma, J.P.M.; von Gunsteren, W.F.; Hermans, J. *Intermolecular Forces*; Pullman, B., Ed.; Reidel: Dordrecht, Holland, 1981; 331.
28. Westwood, A.D.; Youngman, R.A.; Mccartney, M.R.; Cormack, A.N.; Notis, M.R. Oxygen incorporation in AlN via extended defects: I refinement of a structural model for the planar inversion domain boundary (IDB). *J. Mater. Res.* **1995**, *10*, 1270.
29. Cormack, A.N.; Lewis, R.J.; Goldstein, A.H. In preparation.

Protein and Cell Signaling with Biomaterials: Influence of Surfactants

Solomon W. Leung

College of Engineering and Biomedical Research Institute, Idaho State University, Pocatello, Idaho, U.S.A.

James C. K. Lai

Pharmaceutical Sciences Department, College of Pharmacy and Biomedical Research Institute, Idaho State University, Pocatello, Idaho, U.S.A.

INTRODUCTION

This article traces the milestones of how the two fields—1) proteins and cell signaling and 2) surfactants and biomaterials—converge into the new field of roles of surfactants in protein and cell signaling with biomaterials. It highlights biomedical and bioengineering importance of naturally occurring surfactants and the more established industrial, biomedical, bioengineering, and pharmaceutical applications of surfactants. It discusses the key advances in roles of surfactants in protein and cell signaling and points out the need for simpler model systems to elucidate the roles of surfactants in protein and cell signaling to facilitate new biomaterials and tissue engineering research and development. It summarizes some of the exciting new data obtained using enzymes as model systems to elucidate the effects of surfactants on protein functions and looks to the future of advances in biomaterials, biomedical and tissue engineering, and pharmaceutical research and applications once the roles of surfactants in protein and cell signaling are better defined.

HISTORICAL PERSPECTIVE AND DEVELOPMENT

We briefly discuss the biomedical importance of naturally occurring surfactants in different, better characterized physiological and cellular systems and indicate how biomedical scientists and biomedical engineers have borrowed from nature in designing biomedical, industrial, and related applications. This historical perspective allows us to pinpoint some of the developmental milestones leading to the new field of surfactant roles in protein and cell signaling with biomaterials.

Naturally Occurring Surfactants (Endogenous Biological Surfactants)

Extracellular surfactants

The better known and physiologically important extracellular surfactants include 1) bile acids; 2) lung surfactants; 3) joint/tendon lubricants; 4) certain types and fractions of extracellular matrix; and 5) microbial surfactants.

Bile Acids. These complex organic acids are synthesized in the liver from cholesterol and can be conjugated with taurine or glycine to form secondary bile acids. Bile acids play important functions in lipid digestion and absorption. These functions are attributable to the ability of bile acids to form micelles.^[1]

Surfactants of Lungs and Other Tissues/organs. A lung surfactant is a complex mixture of phospholipids (~80%), cholesterol (~10%), and proteins (~10%). This surfactant is mainly synthesized and secreted by Type II alveolar cells of the lung. Because surfactants act to lower the surface tension, the lung surfactant helps the lungs to adjust dynamically to their rates of inflation and deflation during the breathing process.^[2] Additionally, other tissues/organs are recently known to secrete surfactants: The gastrointestinal surfactant secreted by the apical epithelial cells of the gastrointestinal tract may be involved in lubricating the gastrointestinal tract; eustachian tube contains a surface tension-lowering substance; and surfactants are also detected on the surface of peritoneal, pericardial, and pleural mesothelial cells.^[2]

Joint/tendon Lubricants. Surfactants play important lubricant roles in sites in the body where tissues slide over each other with ease (e.g., tendon/tendon sheath and joints).^[3,4]

Certain Types and Fractions of Extracellular Matrix. The complex extracellular matrix material secreted by many mammalian cell types forms an intricate scaffold onto which such cells congregate and form a three-dimensional matrix.^[5] Some of the “ground material” of this scaffold show surfactant properties.^[5,6]

Microbial Surfactants. Other than mammalian cells and tissues, diverse genera of bacteria also produce a host of biosurfactants, which can act as models for the design of biodegradable and environmentally friendly synthetic surfactants and related chemicals.^[7]

Intracellular surfactants

Cytomatrix. The intracellular syntheses, assembly, and ultimately the targeting of lung surfactants for release into the extracellular space constitute an excellent model system for understanding how secretory cells make and secrete surfactants. “Surfactant phospholipids are synthesized in the type II epithelial cells of the lung. The lipids and surfactant proteins are assembled in intracellular storage organelles, called lamellar bodies, and are subsequently secreted into the alveolar space. Within this extracellular space surfactants undergo several transformations. First, the lamellar bodies unravel to form a highly organized lattice-like lipid: protein structure tubular myelin. Second, the organized structures, in particular tubular myelin, adsorb to form a lipid at the air-liquid interface of the alveoli.”^[8]

Nuclear Matrix. Unlike the metabolism and trafficking of lung surfactants in the cytomatrix prior to its secretion into the extracellular space, localization and/or function of surfactants in the nuclear matrix have not been defined. However, surfactant protein A (SP-A) was found in the nuclei of type II alveolar cells^[9] even though the role of SP-A in alveolar nuclei was unknown.

Industrial Formulations and Applications of Surfactants

Industrial applications of surfactants are widespread and economically significant; many applications involve synthetic polymers. In addition to biomedical applications, surfactants are used for paint mixing, water and wastewater treatment, hazardous waste remediation (mostly to mobilize water insoluble substances), and food processing.

Biomedical and Bioengineering Use and Applications of Surfactants

Historically, the empirical applications of surfactants in the pharmaceutical and biotechnology industries are

among the best known industrial applications. However, the molecular mechanisms underlying these applications are either the proprietary information of the companies that hold the patents for such surfactant applications or not well understood.

Biomedical uses and applications

Drug Formulation and Drug Delivery. Surfactants have been extensively employed in drug formulation. Included as pharmaceutical ingredients, surfactants have been employed as emulsifying agents, surface active agents, tablet coating agents, and tablet lubricants.^[10] Surfactants are also employed as soft contact lens cleaners.^[10]

In addition to the formation of liposomes as drug delivery device, surfactants have also been used to deliver viral vaccines,^[11] to formulate lipid microspheres to target drugs to the liver,^[12] and to target immunoliposomes to brain glial tumors.^[13] These are just a few of the numerous current examples of applications of surfactants in drug delivery.

Bioengineering uses and applications of surfactants

Protein Engineering. The traditional and well-established applications of surfactants in protein engineering are the large-scale purification of cellular proteins and solubilization of membrane-bound proteins. There are several new and exciting applications of surfactants in protein engineering. One such application involves enzyme-mediated site-specific poly(ethylene-glycol) (PEG)ylation of proteins to improve the applicability of PEG-protein conjugates for clinical use.^[14] Another new development is the reverse micellar extraction for downstream processing of proteins/enzymes in purification of proteins/enzymes.^[15] The third exciting area is to exploit the catalytic capabilities of enzymes when enzymes are entrapped in reverse micelles so that water insoluble substrates can be metabolized by the elevated enzymatic activities.^[16]

CURRENT ADVANCES IN ROLES OF SURFACTANTS IN PROTEIN AND CELL SIGNALING

As discussed in detail in a related article, recent advances in cell and molecular biology as well as biotechnology have pointed to proteins having two important roles in cell signaling, namely intercellular and intracellular roles. Furthermore, recent developments in cell and tissue engineering research also exemplify the functional and biomedical engineering importance of these two roles of proteins in cell signaling.

Roles of Surfactants in Protein and Cell Signaling: The Intercellular or Extracellular Perspective

“Protein signaling is crucial to regulate cell phenotype and thus engineered tissue structure and function.”^[17] This assertion highlights the role of protein cell signaling in phenotypic expression of cells in culture or in engineered tissue. For example, proteins adsorbed onto substrates on which osteoblasts are cultured provide the integrins on cell membranes of osteoblasts with ligands for interactions, thereby facilitating the integrins’ control of cell behavior through bidirectional signaling pathways.^[18] Another example is the dependence of the phenotypic expression of human chondrocytes on the structure of the biomaterial substrate on which the chondrocytes are cultured.^[19] Thus, depending on the chemical nature of the biomaterial substrate and substrate interaction with focal adhesions on the chondrocytes, the chondrocytes can develop either differentiated or dedifferentiated phenotypic characteristics.^[19] Although protein–protein and protein–substrate interactions are instrumental in cell signaling, the roles of surfactants in mediating these signaling protein interactions have not been fully elucidated.

Even though the roles of surfactants in protein and cell signaling are at present poorly defined, one may use the example of the better-characterized lung surfactant to deduce such putative roles of surfactants. A lung surfactant is a macromolecular complex containing mainly phospholipids and proteins (~10% of lung surfactant weight). There are four main surfactant-associated proteins: SP-A and SP-D are hydrophilic proteins, whereas SP-B and SP-C are hydrophobic proteins.^[2,20] Lung surfactants are synthesized by type II epithelial cells lining the alveoli and stored in lamellar bodies (which are typical intracellular inclusion bodies). The secreted, alveolar form of the surfactant is a complex mixture of lamellar bodies, tubular myelin (which are very organized structures), and monolayered and multilayered, phospholipid-rich sheets and vesicles.^[20] “The multiplicity of these structural forms is influenced by the stoichiometry of lipids and specific proteins, by mechanical forces exerted on the material during the respiratory cycle, and by the uptake, recycling, and degradation of subgroups of particles by respiratory epithelial cells and alveolar macrophages. Alveolar macrophages regulate catabolism of both lipids and proteins—activities that are under strict control of the signaling of granulocyte-macrophage colony-stimulating factor (GM-CSF).”^[20]

As the preceding summary by Whitsett and Weaver^[20] clearly indicates that although the underlying cellular and molecular mechanisms are far from being fully characterized, the secreted lung surfactant phospholipid and protein components play vital signaling

roles in their interactions with each other in the respiratory cycle. Furthermore, the signaling of GM-CSF also regulates the metabolism of and interactions between the components of the lung surfactant (i.e., phospholipids and proteins).^[20] Thus, it is conceivable that other surfactants secreted into the extracellular space (e.g., the gastrointestinal surfactant and joint/tendon lubricants) may share protein and signaling mechanisms similar to those of the lung surfactant.

Roles of Surfactants in Protein and Cell Signaling: The Intracellular Perspective

Even though the evidence for the extracellular roles of surfactants in protein and cell signaling is scant, from the better characterized example of the lung surfactant one can at least derive some generalizable principles as to how surfactants may play roles in protein and cell signaling. In contrast with the extracellular roles of surfactants in protein and cell signaling, the evidence for the intracellular roles of surfactants in protein and cell signaling is even more tenuous. Nevertheless, one can again look to the better-characterized lung surfactant as an example for some mechanistic clues. Both SP-B and SP-C are synthesized in alveolar type II cells as much larger precursor proteins, transported from the endoplasmic reticulum to the Golgi apparatus, then incorporated into multivesicular bodies, and ultimately packaged in lamellar bodies where the active SP-B peptide is stored with SP-C and surfactant phospholipids.^[20] Subsequently, the contents of lamellar bodies are secreted into the extracellular air space, and they interact with SP-A to form tubular myelin.^[20] During the transit from multivesicular bodies to lamellar bodies, both precursor proteins are proteolytically cleaved to form the corresponding SP-B and SP-C.

It is important to note that SP-B is crucial to the proper proteolytic cleavage of the SP-C precursor to form SP-C, the organization of surfactant phospholipids in lamellar bodies, the formation of tubular myelin in the alveolus, and the production of surfactant films that are able to reduce surface tension.^[20] From our knowledge of the intracellular protein and phospholipid trafficking in secretory cells, it is evident that multiple signaling pathways/mechanisms are involved in the transfer of proteins and phospholipids from the endoplasmic reticulum to the Golgi apparatus prior to the exocytotic release of these macromolecular complexes. Consequently, although the signaling mechanisms for the SP-B and the SP-C trafficking from the endoplasmic reticulum to the Golgi apparatus per se have not been fully elucidated, one could reasonably infer the possibility that some signaling pathways/

mechanisms are likely to regulate such intracellular processes.

Need for Simpler Model Systems to Elucidate the Roles of Surfactants in Protein and Cell Signaling

As discussed in the previous sections, evidence is accumulating for the fact that surfactants likely play some roles in protein and cell signaling. However, it is important to emphasize that it is imperative to understand surfactant–protein interactions in molecular and biophysical terms prior to assigning which surfactant–protein interaction could lead to a particular form of cell signaling. Consequently, it is noteworthy that to elucidate the molecular details of surfactant–protein interactions, one needs to design more simplified model systems whereby one could systematically investigate surfactant–protein interactions under vigorously controlled experimental conditions. To address this need, we have employed enzymes as model proteins to systematically investigate surfactant–protein interactions (see the following section).

PROTEIN AND CELL SIGNALING WITH BIOMATERIALS: INFLUENCE OF SURFACTANTS

Surfactants can drastically affect the functions of many proteins, thereby directly or indirectly influence functions of membranes, cells, and organs. Numerous studies have purportedly elucidated the effects of surfactants on cells or membrane-bound proteins. Those studies employed heterogeneous or complex systems that contained many macromolecular structures and/or organelles. Consequently, it is difficult to extrapolate the results of such studies to delineate the effects of surfactants on single proteins. Thus, predicting how surfactants affect cellular protein functions (e.g., enzymatic activity or transport) with existing knowledge is nearly impossible.

Surfactant Effects on Protein Functions

Limitations of previous studies in this area

Historically, there have been many studies on surfactant effects on protein functions using enzymes as model proteins in the biochemical literature. However, most of such earlier studies focused on employing surfactants to release enzymes from their membrane-bound or membrane-associated sites or using surfactants to aid enzyme purification process. Thus, few, if any, of such studies were primarily aimed at a

systematic investigation of the effects of a variety of surfactants on protein functions.

Among model proteins, lactate dehydrogenase (LDH) has been extensively characterized. Many of its catalytic properties and kinetic constants are established. Thus, LDH is an excellent model protein for elucidating the effects of surfactants on protein functions.^[21] Although there have been several studies on the effects of cationic and anionic surfactants on the catalytic activity of purified LDH, most of these studies employed either comparatively narrow ranges of surfactant concentrations,^[22–24] or a very few surfactant concentrations but with a diverse group of cationic and anionic surfactants.^[25] Clearly, even for a well-characterized protein such as LDH, there is a need for systematic and comprehensive studies on the effects of surfactants on its function.

Need for comprehensive and systematic studies to elucidate surfactant effects on protein functions

A systematic understanding of how surfactants affect enzymes is important because many of our body functions or cell signaling are directly or indirectly modulated by naturally occurring surfactants. Our simplified approach of employing enzymes as model proteins allows us to systematically elucidate surfactant effects on protein functions.^[21] Through this approach we can monitor enzymatic activities when enzyme proteins are exposed to surfactants under different but more vigorously controlled conditions thereby permitting us to gain more insights into the dynamics of surfactants' interactions with functionally important enzymes and other proteins. Consequently, the results from studies employing this experimental approach will facilitate better designs of systems of enzyme reactors^[15,26,27] and have important practical implications and applications in protein and tissue engineering^[28,29] as well as drug targeting and delivery in pharmaceutical research and development.^[30,31]

We have selected three well-characterized and metabolically important enzymes—glutamate dehydrogenase (GDH), LDH, and malate dehydrogenase (MDH)—as model proteins to examine the effects of surfactants on protein functions (especially catalytic activity) and on protein transport (see section “Surfactant Effects on Protein Interfacial Transport Across a Semipermeable Membrane”). We have studied anionic surfactants with decreasing hydrophilicity (IB-45, MA-80-I, OT-75, and TR-70) and a small-molecular-weight cationic surfactant (c-573). As a “standard” for comparison with results obtained using these surfactants, we have conducted similar studies with the well-characterized nonionic surfactant, Triton X-100.

Effect of various surfactants and pH on LDH activity

When LDH is exposed to the anionic surfactants IB-45, MA-80-I, OT-75, and TR-70 (in increasing order of hydrophobicity) at pH 6.95, its activity exhibits bands of surfactant concentration dependence, with the most hydrophobic TR-70 exerting the most inhibitory effect (Fig. 1).

At concentrations below or near 1 ppm, Triton X-100 (a nonionic surfactant) exerts an inhibitory effect on LDH activity similar to those induced by the anionic surfactants investigated (Fig. 1). However, between 1 and 10 ppm, Triton X-100 slightly depresses LDH activity (from 15% to 4% depression). Between 50 and 500 ppm, Triton X-100 actually induces increases in LDH activity, with the maximum increase to 120% at 150 ppm (Fig. 1). Thus, the concentration-related effect of Triton X-100 on LDH activity at pH 6.95 is different from those induced by the anionic surfactants on LDH.

These results (Fig. 1) indicate that prediction of changes in enzymatic activities with respect to surfactant concentration is impossible without detailed experimental observations. Moreover, the direction of changes in enzymatic activity induced by a surfactant is just hard to predict even though most enzymatic activity fluctuations are noted with surfactant concentrations less than 10 ppm (Fig. 1).

In general, a systematic pH dependent pattern in LDH activity with various surfactant concentrations is not clearly evident (data not shown). Nevertheless, surfactant concentration-related LDH activity at pH 6.5 is generally lower than the corresponding values at the two higher pHs (6.95 and 7.4), especially at surfactant concentrations less than 10 ppm (data not shown).

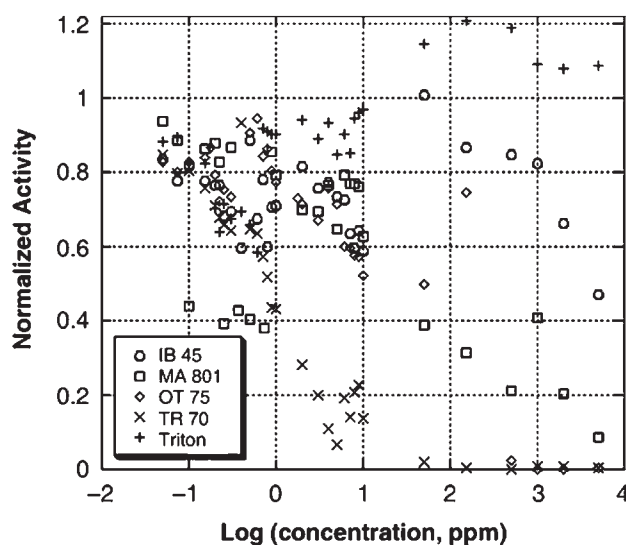


Fig. 1 Effects of surfactants on LDH activities at pH 6.95. (From Ref. [21].)

Moreover, LDH activity is usually higher in the presence of the hydrophilic surfactant IB-45 than those in the presence of the hydrophobic surfactant TR-70. For the hydrophobic surfactant TR-70, the pattern of pH-related changes in LDH activity at various surfactant concentrations is quite different from those noted for IB-45 and Triton X-100 (data not shown).

Effects of hydrophilic and hydrophobic surfactants on enzymes of various molecular masses

For GDH, the enzyme with the largest molecular mass among the three enzymes examined, its activity is the most stable in the presence of the hydrophilic surfactant IB-45 (Fig. 2). Glutamate dehydrogenase activity varies between 84% and 104% of control values at IB-45 levels below 10 ppm (Fig. 2). When IB-45 levels increase, GDH activity exhibits surfactant-related decreases and is about 47% of the control activity at 5000 ppm of IB-45, at which IB-45 concentration, the LDH activity also shows about the same activity. Interestingly, as GDH activity commences to show surfactant concentration-related decreases, LDH and MDH (MDH, the enzyme with the lowest molecular mass among the three enzymes examined) activities show a “step up” when the IB-45 concentrations are higher than 10 ppm; LDH and MDH activities are 101% and 76%, respectively, of the corresponding control values at 50 ppm of IB-45. It appears that, MDH, the enzyme with the smallest molecular mass, exhibits more distinctive bands of activity changes with increasing concentration of IB-45 (Fig. 2). As the molecular mass of the enzyme increases, the enzymatic activity varies less unpredictably with increasing IB-45 concentrations.

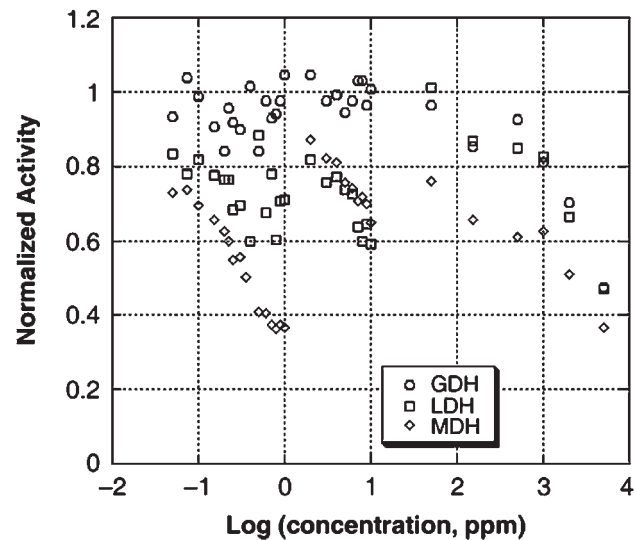


Fig. 2 Effects of surfactants on different sizes of enzymes at pH 6.95 in IB-45 (hydrophilic surfactant). (From Ref. [21].)

Both smaller mass enzymes (i.e., LDH and MDH) behave similarly in the presence of the hydrophobic TR-70. At TR-70 levels lower than 10 ppm, LDH and MDH activities show almost linear decreases with logarithmic increases in the TR-70 concentration. Similar to the effect of IB-45 on GDH, GDH activity is not significantly affected by the hydrophobic TR-70 until the surfactant concentration is higher than 10 ppm, at which point GDH activity shows TR-70 concentration-dependent decreases and is drastically decreased at TR-70 levels higher than 1000 ppm.^[21]

The effects of the nonionic Triton X-100 on the three enzymes of different molecular masses are somewhat the reverse of those noted with the hydrophobic surfactant TR-70, especially at concentration above 10 ppm. Generally, the presence of Triton X-100 tends to enhance enzymatic activities; however, its effect on GDH activity is minimal. Although there are some decreases in activities of the enzymes at Triton X-100 levels less than 10 ppm, MDH and LDH activities are actually more than 100% at Triton X-100 concentrations between 50 and 1000 ppm.^[21]

As one might predict, the activity behavior of LDH (whose molecular mass is midway between those of MDH and GDH) is usually between that of the enzyme with the smallest molecular mass (i.e., MDH) and that of the enzyme with the largest molecular mass (i.e., GDH).^[21]

Time course of LDH activity in the presence of a hydrophilic, a hydrophobic, or a nonionic surfactant at pH 6.95

In the presence of the hydrophilic surfactant IB-45, LDH activity does not markedly change with time. In general, LDH activity tends to decrease in a surfactant concentration-dependent and a time-dependent manner although the decreases in activity beyond the first 60 min of exposure to IB-45 are modest.^[21]

The time course of LDH activity at pH 6.95 varies depending on the concentration of the hydrophobic surfactant TR-70 (Fig. 3). At TR-70 levels lower than 10 ppm, LDH activity does not markedly decrease after the first two hours. However, at TR-70 levels higher than 10 ppm, LDH activity shows a wider range of time-related decreases (Fig. 3), a range much bigger than that exhibited by LDH exposed to the hydrophilic IB-45.^[21] At TR-70 levels higher than 500 ppm, LDH becomes totally inactive.

In the presence of Triton X-100 (a nonionic surfactant), LDH activities show surfactant concentration-dependent and time-dependent variations that are similar to those observed with LDH exposed to the hydrophobic TR-70, with the largest range of 42%. However, unlike the effect of TR-70 on LDH activity, Triton X-100 tends to induce increases in LDH activity

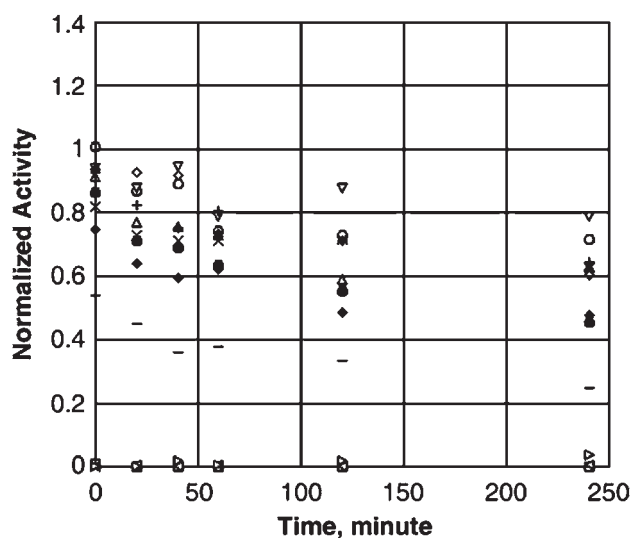


Fig. 3 Lactate dehydrogenase activities with various surfactants over time at pH 6.95 in TR-70 (hydrophobic surfactant). (○) 0.05 ppm, (▽) 0.1 ppm, (◇) 0.5 ppm, (×) 1 ppm, (+) 5 ppm, (Δ) 7 ppm, (●) 10 ppm, (◆) 50 ppm, (—) 100 ppm, (□) 500 ppm, (◁) 1000 ppm, (▷) 5000 ppm. Average relative error of each datum was $<\pm 10\%$. (From Ref. [21].)

with exposure time.^[21] Thus, the findings summarized above suggest that hydrophilic and hydrophobic anionic as well as nonionic surfactants exert dissimilar concentration-related and time-dependent effects on LDH activity.

Effect of a low-molecular-weight cationic surfactant (c-573) on activities of enzymes of different molecular masses at pH 6.95

As discussed above, the generalization that surfactants tend to exert a stabilizing effect on enzymes with larger molecular mass (e.g., GDH) is also valid for depicting the effect of a low-molecular-weight cationic surfactant (c-573) on enzymes of different molecular masses (data not shown). At concentrations below 100 ppm, c-573 tends to induce an enhancement of GDH activity relative to this surfactant's effect on LDH and MDH whose molecular masses are lower than that of GDH. As c-573 concentration exceeds 100 ppm, GDH activity gradually becomes inhibited, similar to the inhibitory effect of this surfactant on LDH and MDH. In general, the smaller the enzyme protein molecule, the more the fluctuation of its activity for the range of surfactant concentration tested.

Comparison of effect of surfactants with different charges on LDH activity at pH 6.95

When LDH activity monitored in the presence of a hydrophilic anionic surfactant (i.e., IB-45) is compared

to that monitored in the presence of a low-molecular-weight cationic surfactant (c-573) at the same concentration range, it is evident that both types of surfactant show similar inhibitory effects on LDH activity except that at levels below 1 ppm the c-573-concentration-dependent variations in LDH activity are less marked compared to the corresponding variations in LDH activity in the presence of IB-45 (data not shown). At levels between 1 and 10 ppm, the inhibitory effect of c-573 on LDH is less pronounced compared to that of IB-45 (data not shown).

Surfactant Effects on Protein Interfacial Transport Across a Semipermeable Membrane

Surfactants can induce an increase or inhibit the permeation rate of proteins across semipermeable membranes, thereby directly or indirectly exert some control of cell signaling. We have employed a U-Tube separation cell to study the effects of surfactants with various hydrophilicities on LDH permeability across a 10- μm thick polycarbonate membrane with 1- μm pores.^[32] We summarize and discuss some of our findings as follows.

Effects of anionic and nonionic surfactants on permeation LDH through a semipermeable membrane at pH 6.5, 6.95, and 7.4

To simulate the intracellular conditions in which the functions of cells can occur with the two extremes of intracellular pH, namely between pH 6.5 and 7.4, we have carried out experiments at pH 6.5, 6.95, and 7.4. We monitored the appearance of LDH in the left half-cell at various times after the enzyme had been administered into the right half-cell.

At pH 6.5, the solution without any added surfactant allows the least amount of LDH to permeate the left half-cell, whereas the solution containing the most hydrophilic surfactant (namely, IB-45) allows most LDH to permeate but the solution containing the

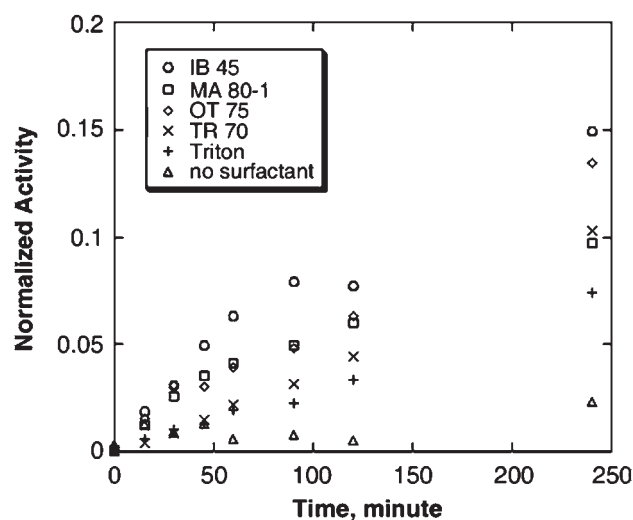


Fig. 4 Permeation of LDH in the presence of various surfactants with 10 mM of phosphate buffer. Membrane pore size 1 μm , surfactant concentration at the right side is 0.1 ppm at pH 6.5.

nonionic surfactant (Triton X-100) is the least effective in promoting LDH permeation (Fig. 4). Interestingly, the results obtained at pH 7.4 are exactly in the reversed order: The solution without surfactant allows the highest amount of LDH permeation through the membrane; the solution containing IB-45 allows the least amount of LDH permeation; and among the solutions containing different surfactants, the solution containing Triton X-100 is the one that exerts the least “inhibition” of LDH permeation.

At pH 6.95, which is the normal intracellular pH for most mammalian cells, solutions containing surfactants allow higher LDH mass permeation than the solutions without surfactants do. At this pH, solutions containing hydrophilic or nonionic surfactants are more effective than other surfactant-containing solutions in enhancing LDH permeability. Table 1 lists the permeabilities of LDH determined in some of our experiments.

Table 1 Permeabilities (cm/sec) of LDH with various surfactants and pHs through the 10- μm thick polycarbonate membrane

	pH 6.5	pH 6.95	pH 7.4
IB-45	9.15E-05	1.89E-04	3.25E-05
MA-80-I	5.94E-05	1.66E-04	3.37E-05
OT-75	5.91E-05	7.02E-05	3.94E-05
TR-70	4.13E-05	7.63E-05	3.50E-05
Triton X-100	3.21E-05	1.01E-04	5.01E-05
No surfactant	1.03E-05	7.86E-05	5.44E-05

Surfactant concentration was 0.1 ppm.

CONCLUSIONS

This article traces the milestones of how the two fields—1) proteins and cell signaling and 2) surfactants and biomaterials—converge into the new field of roles of surfactants in protein and cell signaling with biomaterials. It highlights biomedical and bioengineering importance of naturally occurring surfactants and the more established industrial, biomedical, bioengineering, and pharmaceutical applications of surfactants. It discusses the key advances in roles of surfactants in protein and cell signaling and points out the need for simpler model systems to elucidate the roles of surfactants in protein and cell signaling to facilitate new biomaterials and tissue engineering research and development.

We have selected three well-characterized and metabolically important enzymes—GDH, LDH, and MDH—as model proteins to examine the effects of surfactants on protein functions (especially catalytic activity) and on protein transport.

There is no generalizing trend about the pH effect on the surfactant concentration dependence of the activity of enzymes of different molecular masses. Functional changes of enzymes (e.g., enzymatic activity) are dependent on time of exposure to the surfactant, surfactant concentration, and hydrophilicity of the surfactant. As a factor, pH modulates enzymatic activity only when enzymes are exposed to selected surfactants.

Surfactants appear to exert a “stabilizing effect” on enzymes of larger molecular mass (e.g., GDH). The activity of such an enzyme is less affected by pH, surfactant hydrophilicity, or surfactant concentration. Activity of enzymes of larger molecular mass is more predictable with variation of the concentration of surfactant employed. There are no sudden swifts of enzymatic activity with small changes of surfactant concentration. For an enzyme of smaller molecular mass (i.e., MDH), there are sudden swifts of enzymatic activity at low surfactant concentrations (less than 10 ppm). In general, enzymes are more active in hydrophilic than in hydrophobic surfactants. This difference in enzymatic activity is more prominent at high surfactant concentrations.

Triton X-100, a nonionic surfactant, generally induces increases in enzymatic activities, usually at a concentration higher than 10 ppm. The enhancement in enzymatic activity is less than 50%.

Surfactant effects in modulating activities of enzymes critically depend on contact time with surfactants. Thus, any comparisons of activity of enzymes exposed to a surfactant must take into account the time factor.

Caution must be exercised when using enzymatic activities to calculate mass balance for enzymes, especially in the presence of surfactants and when time is a variable during enzymatic assays. Thus, because of these other factors, changes in the activities of enzymes may become nonlinear with mass calculation.

Anionic and nonionic surfactants have definite effects on permeability of enzymes through semipermeable membranes. Their effects are dependent on surfactant hydrophilicity and are not linearly related to pH.

Until the structures of protein enzymes are known, it would be difficult to investigate the mechanisms of how these proteins interact with surfactants. However, we observed that increase of ionic surfactant concentrations exerted bands of activity suppressions to enzymes, and increase of nonionic surfactants at high concentrations slightly enhanced enzyme activities. We speculate that the enzyme active sites are charged; therefore, charged surfactants would tie up (bind with) the active sites and inhibit the catalytic reactions of enzymes. As the concentration of surfactants increases to certain level, surfactants reconfigure the orientation of enzymes thus create bands of activities of enzymes. For nonionic surfactants, we speculate that the surfactant actually “relaxes” the enzymes to expose more active sites, or it captures the reactive species (pyruvate, for example) and creates a high reactive site with enzymes, similar to the surfactant effect in photocatalytic reactions.^[33]

Results from ongoing studies on surfactant effects on protein functions employing enzymes as model proteins indicate that such simpler experimental systems allow us to investigate systematically the interactions between surfactants and proteins at the molecular level. Such studies will further our understanding of the surfactant–protein interactions, ultimately leading to the elucidation of the roles of surfactants in protein and cell signaling with biomaterials. Moreover, this type of fundamental studies will make significant contributions toward cell and tissue engineering, drug formulation and targeting, and water/waste water treatment.

ARTICLES OF FURTHER INTEREST

Biofilms; Biomaterials: Protein–Surface Interactions; Bioadhesion; Phosphorylcholine (PC) Technology; Protein Adsorption Modeling; Self-Assembled Monolayers; Surface Modification

ACKNOWLEDGMENTS

Our studies were supported, in part, by NIH Grant #P20 RR16454 from the Idaho BRIN Program of the National Center for Research Resources at the National Institutes of Health, and an FRC grant and a URC grant from Idaho State University.

REFERENCES

1. Hardikar, W.; Suchy, F.J. *Hepatobiliary function. Medical Physiology: A Cellular and Molecular Approach*,

- 1st Ed.; Boron, W.F., Boulpaep, E.L., Eds.; Saunders: Philadelphia, 2003, Chapter 45, 975–1002.
2. Khubchandani, K.R.; Snyder, J.M. Surfactant protein A (SP-A): the alveolus and beyond. *FASEB J.* **2001**, *15*, 59–69.
 3. Hills, B.A. Boundary lubrication in vivo. *Proc. Inst. Mech. Eng. H* **2000**, *214* (1), 83–94.
 4. Mills, P.C.; Hills, Y.; Hills, B.A. Surface-active phospholipid (surfactant) in equine tendon and tendon sheath fluid. *N. Z. Vet. J.* **2005**, *53* (2), 154–156.
 5. Anderson, E.H.; Ruegsegger, M.A.; Murugesan, G.; Kottke-Marchant, K.; Marchant, R.E. Extracellular matrix-like surfactant polymers containing arginine-glycine-aspartic acid (RGD) peptides. *Macromol. Biosci.* **2004**, *4* (8), 766–775.
 6. Wang, Y.C.; Ho, C.C. Micropatterning of proteins and mammalian cells on biomaterials. *FASEB J.* **2004**, *18* (3), 525–527.
 7. Maier, R.M. Biosurfactants: evolution and diversity in bacteria. *Adv. Appl. Microbiol.* **2003**, *52*, 101–121.
 8. Veldhuizen, R.; Possmayer, F. Phospholipid metabolism in lung surfactant. *Subcell. Biochem.* **2004**, *37*, 359–388.
 9. Bakewell, W.E.; Smith, G.J.; Miller, B.E.; Viviano, C.J.; Hook, G.E. Confocal laser scanning immunofluorescence microscopy of the pulmonary surfactant system. Association of surfactant protein A with the nucleus of the alveolar type II cell. *Lab. Invest.* **1993**, *68* (5), 566–576.
 10. Allen, L.V., Jr.; Popovich, N.G.; Ansel, H.C. *Ansel's Pharmaceutical Dosage Forms and Drug Delivery Systems*, 8th Ed.; Lippincott Williams & Wilkins: Philadelphia, 2005.
 11. Chattaraj, S.C.; Das, S.K. Physicochemical characterization of influenza viral vaccine loaded surfactant vesicles. *Drug Deliv.* **2003**, *10* (2), 73–77.
 12. Abrol, S.; Trehan, A.; Katare, O.P. Formulation, characterization, and in vitro evaluation of silymarin-loaded lipid microspheres. *Drug Deliv.* **2004**, *11* (3), 185–191.
 13. Chekhonin, V.P.; Zhirkov, Y.A.; Gurina, O.L.; Ryabukhin, I.A.; Lebedev, S.V.; Kashparov, I.A.; Dmitriyeva, T.B. PEGylated immunoliposomes directed against brain astrocytes. *Drug Deliv.* **2005**, *12* (1), 1–6.
 14. Sato, H. Enzymatic procedure for site-specific PEGylation of proteins. *Adv. Drug Deliv. Rev.* **2002**, *54* (4), 487–504.
 15. Krishna, S.H.; Srinivas, N.D.; Raghavarao, K.S.; Karanth, N.G. Reverse micellar extraction for downstream processing of proteins/enzymes. *Adv. Biochem. Eng. Biotechnol.* **2002**, *75*, 119–183.
 16. Orlich, B.; Schomacker, R. Enzyme catalysis in reverse micelles. *Adv. Biochem. Eng. Biotechnol.* **2002**, *75*, 185–208.
 17. Boonthekul, T.; Mooney, D.J. Protein-based signaling systems in tissue engineering. *Curr. Opin. Biotechnol.* **2003**, *14* (5), 559–565.
 18. Siebers, M.C.; ter Brugge, P.J.; Walboomers, X.F.; Janeen, J.A. Integrins as linker proteins between osteoblasts and bone replacing materials. A critical review. *Biomaterials* **2005**, *26* (2), 137–146.
 19. Mahmood, T.A.; de Jong, R.; Riesle, J.; Langer, R.; van Blitterswijk, C.A. Adhesion-mediated signal transduction in human articular chondrocytes: the influence of biomaterial chemistry and tenascin-C. *Exp. Cell Res.* **2004**, *301*, 179–188.
 20. Whitsett, J.A.; Weaver, T.E. Hydrophobic surfactant proteins in lung function and disease. *N. Engl. J. Med.* **2002**, *347* (26), 2141–2148.
 21. Leung, S.W.; Lai, J.C.K. Differential effects of anionic surfactants on activities of GDH, LDH, and MDH. *Biochem. Eng. J.* **2005**, *25* (1), 79–88.
 22. Sanford, K.; Meyer, D.; Mathison, M.; Figueras, J. Selective inactivation of lactate dehydrogenase isoenzymes with ionic surfactants. *Biochemistry* **1981**, *20*, 3207–3214.
 23. Lehnert, T.; Berlet, H.H. Selective inactivation of lactate dehydrogenase of rat tissues by sodium deoxycholate. *Biochem. J.* **1979**, *117*, 813–818.
 24. Vincenzini, M.; Favilli, F.; Treves, C.; Vanni, P. Specific interaction among some enzymes and sodium dodecyl sulfate. *Life Sci.* **1982**, *31*, 463–470.
 25. Chander, A.; Sen, N.; Spitzer, A. Synexin and GTP increase surfactant secretion in permeabilized alveolar Type II cells. *Am. J. Physiol. (Lung Cellular and Molecular Physiology)* **2001**, *280*, 991–998.
 26. Castro, G.R.; Knubovets, T. Homogeneous biocatalysis in organic solvents and water-organic mixtures. *Crit. Rev. Biotechnol.* **2003**, *23*, 195–231.
 27. Kirk, O.; Borchert, T.V.; Fuglsang, C.C. Industrial enzyme applications. *Curr. Opin. Biotechnol.* **2002**, *13*, 345–351.
 28. Cartmell, J.S.; Dunn, M.G. Effect of chemical treatments on tendon cellularity and mechanical properties. *J. Biomed. Mat. Res.* **2000**, *49*, 134–140.
 29. Kellam, B.; De Bank, P.A.; Shakesheff, K.M. Chemical modification of mammalian cell surfaces. *Chem. Soc. Rev.* **2003**, *32*, 327–337.
 30. Levis, S.R.; Deasy, P.B. Pharmaceutical applications of size reduced grades of surfactant co-processed microcrystalline cellulose. *Int. J. Pharm.* **2001**, *230*, 25–33.
 31. Varshosaz, J.; Pardakhty, A.; Hajhashemi, V.I.; Najafabadi, A.R. Development and physical characterization of sorbitan monoester niosomes for insulin oral delivery. *Drug Deliv.* **2003**, *10*, 251–262.
 32. Leung, S.W.; Rivera, G.; Lai, J.C.K. *Effects of surfactants on membrane separation of proteins at trace levels for drinking water*, Proceedings of the 3rd Leading-Edge Conference on Water and Wastewater Treatment Technologies, June 6–8, 2005; Sapporo, Japan.
 33. Chu, W.; Jafvert, C.T. Photodechlorination of polychlorobenzene congeners in surfactant micelle solution. *Environ. Sci. Technol.* **1994**, *28*, 2415–2422.

Protein and Cell Signaling with Biomaterials: Interfacial Transport

James C. K. Lai

Pharmaceutical Sciences Department, College of Pharmacy and Biomedical Research Institute, Idaho State University, Pocatello, Idaho, U.S.A.

Solomon W. Leung

College of Engineering and Biomedical Research Institute, Idaho State University, Pocatello, Idaho, U.S.A.

INTRODUCTION

This article traces how the field of protein and cell signaling with biomaterials originated from the biomedical research on protein and cell signaling to the current advances in biomaterial design and its applications in cell and tissue engineering. Despite the paucity of relevant data in the literature on this subject, this entry identifies the importance of interfacial transport in protein and cell signaling and reports on some fundamental studies concerning interfacial protein transport. It discusses some of the key thermodynamic, kinetic, and molecular properties of a novel membrane system for protein transport. It highlights many application areas of interfacial protein transport, including, but not limited to, development of new techniques for drug separation; improved, more efficient apparatus design for dialysis; engineering applications in water and wastewater treatment; and applications in artificial organs and cell and tissue engineering.

HISTORICAL PERSPECTIVE AND DEVELOPMENT

This introductory section traces the origin of the field of protein and cell signaling with biomaterials from the biomedical research on protein and cell signaling to the current concepts in biomaterial design and their applications in tissue engineering. Ultimately, this historical perspective highlights how interfacial transport plays an important role in protein and cell signaling with biomaterials.

Protein and Cell Signaling

The development of cell and molecular biology has led to the elucidation of two roles of proteins in cell signaling—intercellular and intracellular. In intercellular signaling, the coupling of specialized proteins in membrane dynamics is involved in cell-to-cell recognition,

adhesion, and signaling; this intercellular protein signaling also plays an important role in apparently unrelated but physiologically significant cell fusion functions, such as phagocytosis, cell migration, axonal growth, and synaptogenesis.^[1] Among the key factors of the extracellular environment influencing these cell-to-cell protein signaling functions are pH, ionic strength, and extracellular matrix. On the other hand, intracellular signaling implicates plasma membrane-bound proteins (e.g., receptors, transporters, membrane-bound or membrane-associated enzymes), which transmit a series of signaling cascades within cells resulting in, for instance, intracellular protein trafficking and/or alterations in regulation of gene expression.^[2] This membrane mechanism highlights the importance of transmembrane events including, but not limited to, ligand–receptor interaction and interfacial transport (or transport through membrane pores and ion channels). This type of intracellular protein signaling can be influenced by changes in intracellular environment, such as the changes in intracellular pH and ionic milieu, resulting in alterations in binding capacities and other molecular properties of intracellular protein signals. Such changes in intracellular protein signals are ultimately responsible for controlling intracellular protein trafficking and gene expression.

Biomaterials scientists and biomedical engineers have exploited the understanding of the extracellular and intracellular roles of proteins in cell signaling in biomedical engineering research and development (see below). However, there has been a paucity of systematic studies to elucidate how extracellular and intracellular milieu exerts effects on protein signals at the molecular level (see below).

Protein and Cell Signaling with Biomaterials

Biomaterials and tissue engineering research and development

Advances in biomaterials have exploited different aspects of protein and cell signaling. An important area



where this knowledge is applied is in drug formulation and drug delivery. For example, a protein drug (e.g., trastuzumab) that targets to cancer cells recognizes surface receptors on cancer cells. As a result of this recognition of surface receptors on the cancer cells, the protein drug can interact with the cancer cells and leads to killing of the cancer cells (see Ref. [3] and references therein). Another example is artificial bone/joint design, in which the biologically competent and compatible integrins (i.e., cell adhesion proteins) are used as linker proteins between osteoblasts and bone replacing materials.^[4] This tissue engineering example is based on new knowledge of how the integrins control cell behavior through bidirectional signaling mechanisms/pathways.^[4] Similarly, other applications of protein-based signaling include, but are not limited to, the manipulation of cell signaling employing growth factors and peptides in tissue engineering, whereby the damaged tissues are replaced using either transplanted or host cells.^[5]

Despite significant advances in biomaterials and tissue engineering, there are methodological issues that preclude the design of new and fully biocompatible biomaterials employed in cell and tissue engineering. These methodological issues revolve around the central question of “biocompatibility.” To more adequately address the more global aspects of biocompatibility, researchers have focused their attention on enhancing their understanding of: 1) protein and cell interactions with biomaterials; 2) protein and cell signaling with biomaterials; and 3) more biocompatible biomimetic.^[6–8] (A detailed discussion of these topics is, however, beyond the scope of this entry.)

Protein and cell interactions with biomaterials

Earlier, biomaterial scientists, biochemists, and biomedical engineers had focused their attention on broad biocompatibility issues by addressing protein conformational considerations, as proteins interact with biomaterial surfaces. Even then, researchers conjectured that to really understand protein and cell signaling with biomaterials, one must get down to the more fundamental level of being able to understand what are the molecular interactions between a protein and the surface it adsorbs to, and the conformational changes of the protein as it adsorbs to the same surface.

As extensively discussed by Lu, Kadambi, and Lai, proteins in solution assume one of two different secondary structures, i.e., the α -helix or the β -sheet (or some combination of the two structures), depending on the number and the type of amino acids that constitute the proteins.^[9] For a protein in bulk solution, its conformation is determined by the internal

stabilization forces, including hydrogen bonding, hydrophobic interactions, electrostatic interactions, dispersion forces, and conformation entropy (see Ref. [9] for a detailed discussion). Molecular modeling of the interaction potential, hydrophobic interactions, and conformational changes has been employed to determine the conformational changes of a protein adsorbed onto a surface.^[9] Thus, molecular modeling and various experimental approaches have been used to elucidate how the conformational changes of a protein, as it adsorbs onto a surface, play a role in the protein's interaction with the surface, thereby imparting the signaling function of the protein (see Ref. [9] for a detailed discussion). The nature of the biomaterial surface (e.g., its hydrophobicity or hydrophilicity) determines how the protein, on adsorption onto the surface, interacts with the surface.^[9] From the results of these earlier studies emerges the notion that this protein–surface interaction determines the way the protein functions as a signaling molecule. Furthermore, these earlier studies pointed to the fact that there is a need for new mechanistic studies to further elucidate the binding characteristics and interactions of the adsorbed protein with cells at the solution–material interfaces.^[9]

Importance of interfacial transport in protein and cell signaling with biomaterials

As alluded to above, recent advances in protein and cell signaling with biomaterials have emphasized on functions of cell surface proteins but very few studies have addressed interfacial transport of proteins in this context.^[3–5] Nevertheless, earlier studies on interactions of adsorbed proteins with cells clearly suggested an important function for interfacial transport in protein and cell signaling with biomaterials.^[9] Consequently, any studies to elucidate mechanisms underlying interfacial transport of proteins will shed new light on protein and cell signaling with biomaterials.

INTERFACIAL TRANSPORT

Introduction

Biological interfacial transport must involve some kind of membrane barrier. Incidentally, membrane research and applications are technologies identified by the National Science Foundation (NSF) and the Department of Energy as the emerging technologies of the future.^[10] Recent developments in nanotechnologies broaden the applications of membrane science even faster. Armed with a better understanding of surface phenomena and science, membrane researchers are

exploring many applications that were overlooked previously, especially in biomedical fields. Thus, the newly formed biomedical research and engineering division at the National Institutes of Health (NIH) has targeted nanotechnology as one of its main programmatic emphases; evidently, this strategy was directed, in part, to address the gaps in membrane research applications in the biomedical engineering field. Because of the complex nature of membrane phenomena, both NSF and NIH have advocated that any research and applications of membrane sciences have to adopt an interdisciplinary approach and strategy.

As most of our bodily functions are dependent on the interactions of various membrane systems, a better understanding of how the selectivities and permeabilities of biomolecules and drugs across these membranes are controlled allows biomedical researchers and bioengineers to come up with improved designs of membrane systems, such as those in hemodialysis and detoxification of body fluids. Similarly, the pharmaceutical industry can greatly exploit this technology in varieties of chemical separations they routinely carry out.

Both the NSF and the NIH have recently emphasized the importance of membrane-associated phenomena and the underlying biophysical and molecular mechanisms. While most previous and current research directions in this field are concerned with the transport of chemicals across the membranes, very few studies address the transfer across membranes of macromolecules, such as enzymes and other proteins, because most studies have focused on the transfer of small molecules (e.g., sugars, amino acids, and small-molecular weight drugs).

Membrane research has been conducted by diverse investigators for some time, especially in the biomedical sciences.^[11-14] A substantial number of reports can be found in the literature about liquid membrane systems, mostly in the biological sciences.^[13-16] However, most of these observations were based on limited data employing biological systems, which are heterogeneous and contain multiple, complex experimental variables. Consequently, these reports are mainly statements of facts or qualitative descriptions of phenomena; few fundamental physical and thermodynamic properties and related principles can be deduced from such observations because of the complexity of their membrane systems, which invariably contained very heterogeneous biological membrane components.

On the other hand, most of the membrane research in the engineering domain in the past emphasized on industrial applications, such as purification of gaseous mixtures, desalination of brackish water, and food processing.^[17]

In the past, many research groups (e.g., those located at the University of Texas at Austin, McMaster University, University of Ottawa, and University of

Colorado at Boulder) concentrated their research efforts on membrane formulation and industrial applications. More recent development of nanotechnology expands the traditional membrane research to include more explorations in biomedical applications.

Up to now, the literature on surface phenomena and behavior and on thermodynamic properties of membranes is scant. The exceptions are the contributions of the group led by Martin and Mitchell, and Nishizawa, Menon, and Martin, who have pioneered the method of creating conductive membrane mesh by electrochemical deposition and have begun to quantify transport properties of the unique membrane(s) they designed.^[18-21] On the other hand, information on protein transport across semipermeable membranes is nearly nonexistent.^[22]

To understand how proteins are transported intercellularly and intracellularly, and how these transport mechanisms are functionally coupled to cell signaling, we must gain more fundamental insights into the thermodynamic properties that are the driving forces of these transport mechanisms. In general, because biological membrane systems usually contain a heterogeneous, multifunctional mixture of membrane components (e.g., membranes of different sizes, shapes, density, and chemical composition) and subsystems, it is impractical and methodologically difficult at this time to try to harness such multicomponent mechanisms all at once in a multicellular system design. Thus, it would be more realistic to investigate the thermodynamic properties of interfacial transport employing a simple, more homogeneous membrane system in which the experimental conditions can be much more vigorously controlled, and the environmental factors and other test parameters can be evaluated one at a time. Through the latter approach one can construct a comprehensive picture of the transport properties of biomaterials of interest.

Current Applications of Interfacial Transport

A comprehensive review of applications of interfacial transport in biomedical engineering, environmental engineering, and biotechnology is beyond the scope of this entry. Nevertheless, it is important to briefly point out some of the rapidly growing areas where applications of interfacial transport are making significant inroads and contributions.

Critical issues in applications of interfacial transport in artificial organs and tissue engineering

Although the idea of creating artificial organs dated back several decades, development of functionally competent artificial organs is still in its infancy except



for the development of the artificial kidney. As indicated above (in the section “Biomaterials and Tissue Engineering Research and Development”), the successes of a designed artificial organ and an engineered tissue critically depend on the ability to engineer biomaterial surfaces that specially target certain functionally important and biocompatible protein–ligand interactions and signaling cascades to achieve a strategy for optimizing cellular responses in biomaterials and tissue engineering applications.^[3–5] Furthermore, implicit in the latter assertion is the implication that success in artificial organ and tissue engineering design is equally contingent on a strategy to optimize interfacial transport of biomolecules. For example, from the functional point of view of designing and manufacturing artificial organs, it is imperative to have a fundamental understanding of the transport properties between the organ surface/membrane and the blood vessels that supply the organ with oxygen and nutrients.^[23] In other words, we must have a good understanding of the transport properties between the cell membranes of the organ and the blood vessels. Therefore, the development of artificial membranes (which allow the systematic investigation of such transport phenomena) becomes the forefront of research and development of artificial organs.

There are two important aspects of the development of artificial membranes we must resolve—the mechanistic and the functional aspect. In the last decade or so, several research groups have been working on the first aspect: the most advanced artificial membranes are currently made from polymer(s) and metal composite.^[24–26] However, the functional aspect of the development of artificial membranes is yet to be addressed.

What appears to hold up the advances in the functional aspect of the development of artificial membranes can be traced to two critical issues. First, the membrane properties of organs vary depending on their functions and there is no single artificial membrane—for that matter even a series of artificial membranes—that can represent all organ surfaces/membranes. Second, our body fluids are very complex heterogenous systems and contain multiple, chemically dissimilar components. Consequently, the transport of components in our body fluids across organ membrane systems is very complex and is very difficult to simulate employing artificial membranes. Hence, we must resolve these two critical issues before significant strides can be made in research and development of artificial organs and tissue engineering.

Applications of interfacial transport in water and wastewater treatment

There are discharges of industrial hormones from the dairy and poultry industries, excretions of vitamins and

enhancement hormones from humans in sewer systems, and discharges from the pharmaceutical industry. This new category of contaminants (which are quite unlike the traditional humic materials that perpetually exist in most of the surface waters and wastewaters) mainly comprises small molecular weight protein and other organic molecules. These organic contaminants eventually find their way into our water supply systems. As recycling of spent water for potable use is becoming a reality, the understanding of how these organic molecules affect our water treatment processes is becoming important, especially for the emerging contaminants, such as antibiotics and steroid hormones; treatment standards for these organic compounds are soon to be realized.^[27] As most of these molecules exist in rather low concentrations ($<10^{-6}$ M) and in soluble forms, treatment and removal of these molecules are more difficult than that of the suspended solids. Artificial membranes can be used to remove these soluble biomolecules. Thus, a better understanding of interfacial transport of biomaterials can help develop more efficient water and wastewater treatment systems.

Interfacial Transport Across a Membrane

Before examining how one can study interfacial transport across a membrane, one needs to be cognizant of some basic definitions and theoretical considerations in this field.

Some basic definitions and theoretical considerations

Flux of i component across a membrane can be defined by:

$$F_i^0 = L_{ii}^0 \phi(\Delta G_i) + \sum_j L_{ij}^0 \phi(\Delta G_j) \quad (1)$$

The first term represents the driving force of the i th component, which is a function of free-energy difference across the membrane and the phenomenological flux coefficient (L_{ii}^0) and includes the diffusion coefficient and others. The second term represents the coupled flow of component i with other components and their phenomenological flux coefficients.

From thermodynamics utilizing the Gibbs equation, the entropy production rate can be written as:^[28]

$$\phi = T \frac{d\Delta S}{dt} = J\Delta P + I\Delta E$$

where T is temperature, S is entropy, J is generalized flux of molecules, P is pressure, I is electric current, E is electrical potential, and t is time.

The linear equations corresponding to currents and fluxes can be written as:

$$I = L_{11}\Delta E + L_{12}\Delta P \quad (2)$$

$$J = L_{21}\Delta E + L_{22}\Delta P \quad (3)$$

where $L_{ij} = L_{ji}$, and they are phenomenological coefficients. Eq. (2) states that electrical potential and pressure gradient can cause electric current, and Eq. (3) states that both electric potential and pressure gradient are responsible for the flux or bulk volume flow.

Using Onsager relationship, Eqs (2) and (3) can be combined and simplified to:

$$\left(\frac{\Delta E}{\Delta P}\right)_{I=0} = -\left(\frac{J}{I}\right)_{\Delta P=0} \quad (4)$$

$$\left(\frac{\Delta P}{\Delta E}\right)_{J=0} = -\left(\frac{I}{J}\right)_{\Delta E=0} \quad (5)$$

Therefore, pressure, electrical current, and electrical potential can be manipulated to control fluxes of molecules going across both sides of a membrane.

Design of a separation cell to investigate interfacial transport

As indicated in the previous section, pressure, electrical current, and electrical potential are some of the critical variables that can determine the fluxes of molecules across a membrane. To evaluate how biomaterials behave thermodynamically during interfacial transport based on Eqs. (1), (4), and (5), one needs to design a comparatively simple artificial membrane system in which the experimental conditions can be vigorously controlled. (The importance and utility of such a membrane system have already been discussed earlier.) Thus, to address this experimental design need, we have developed a simple separation cell (Fig. 1) that allows us to systematically investigate interfacial transport of biomaterials. In the series of studies to be discussed below, we address several key parameters that may exert effects on the interfacial transport of proteins. We chose three well-characterized enzyme proteins as the surrogate biomaterials: L-glutamate dehydrogenase (GDH) [EC 1.4.1.3; G-2626; Sigma Chemical Co., St. Louis, Missouri, U.S.A.; from bovine liver, molecular mass (MM) = 2,200,000], L-lactate dehydrogenase (LDH) (EC 1.1.1.27; L-2518; Sigma Chemical Co.; from rabbit muscle, MM = 135,000), and L-malate dehydrogenase (MDH) (EC 1.1.1.37; M-2634; Sigma Chemical Co.; from porcine heart,

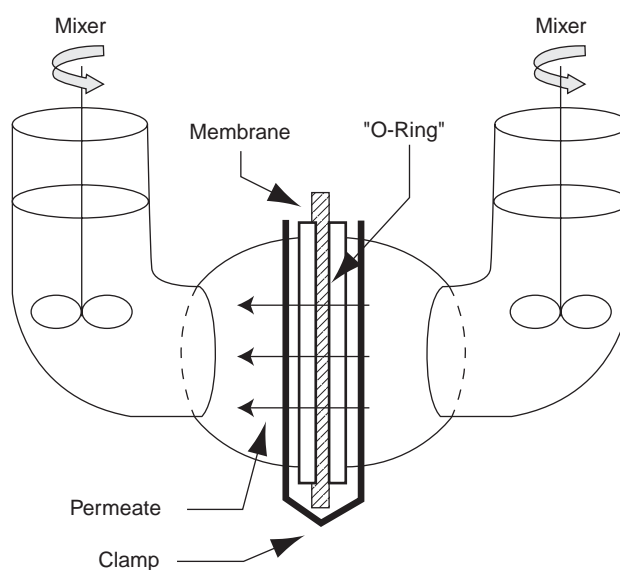


Fig. 1 The separation cell with a 1 μm semipermeable polycarbonate membrane.

MM = 70,000). These enzyme molecules were chosen as a representative of proteins with small (MDH), intermediate (LDH), or larger (GDH) MMs. Furthermore, these enzymes were selected because they are important in several key metabolic pathways in mammalian cells and, as already discussed above, systematic elucidation of interfacial transport of these enzyme proteins employing an interdisciplinary approach will shed new light on interfacial transport in protein and cell signaling with biomaterials.

Effect of protein concentrations on protein permeability through an artificial membrane

For this study, the surrogate protein is LDH. Fig. 2 is the atomic force microscopy (AFM) image of a polycarbonate membrane employed for this study. As shown in Fig. 3, when protein concentration is the sole driving force in the system, the amount of LDH protein that passes through the membrane increases with concentration, but not necessarily proportionally. Observation of the medium concentration curve ($1.00 \times 10^{-8} \text{ M}$) at 2 hr shows a decrease in concentration; we believe it was the effect of artifact(s). A parallel run at pH 6.9 of the same concentration shows that LDH concentration continues to increase at the 2 hr period.

Permeability Through an Artificial Membrane of Proteins Having Different MMs. For this study, the surrogate proteins employed are MDH (having the lowest MM of the three proteins), LDH (having the intermediate MM), and GDH (having the highest

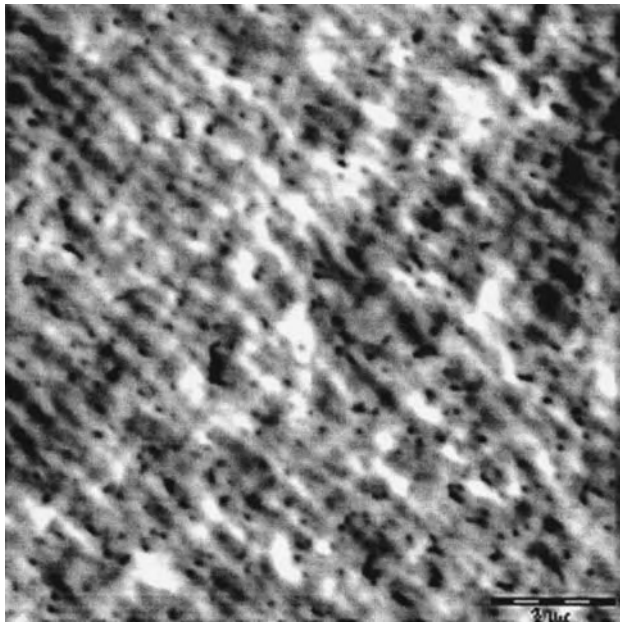


Fig. 2 Atomic force microscopy of a 0.1 μm polycarbonate membrane in LDH solution with 10 mM phosphate buffer.

MM). As shown in Fig. 4, other than the concentration difference, the size of a protein molecule would also influence its effective permeation. If as the hydrodynamical theory predicts, the diffusivity is reversibly proportional to the molecular radius and directly proportional to permeability, then our results are the same as that predicted.^[29] L-Malate dehydrogenase, the smallest molecule among the three proteins studied, has the highest permeability. In addition, even when

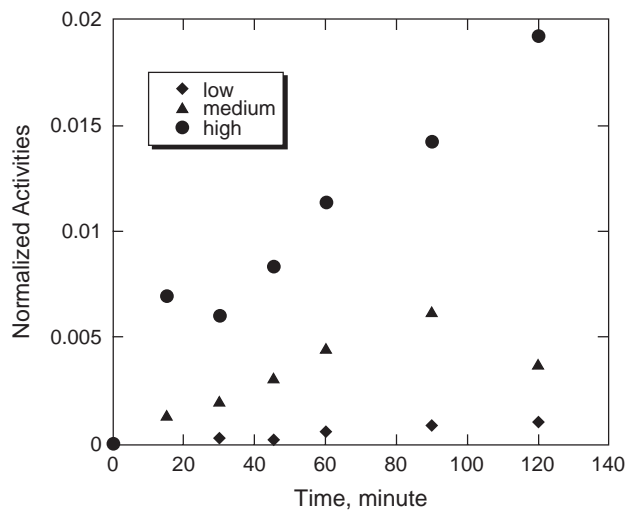


Fig. 3 L-Lactate dehydrogenase permeation at various initial concentrations (low = 2.50×10^{-9} M, medium = 1.00×10^{-8} M, and high = 2.00×10^{-8} M) at pH 7.4 in 2 hr.

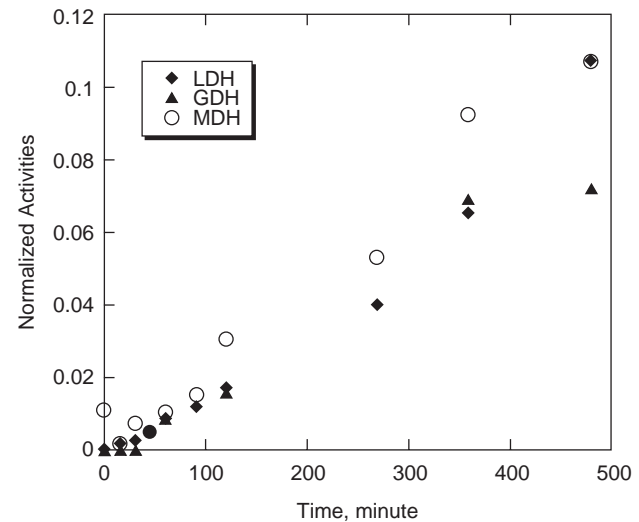


Fig. 4 Permeability of GDH (MM = 2,200,000), LDH (MM = 135,000), and MDH (MM = 70,000) at pH 6.95 in solution with 10 mM of phosphate buffer and 100 mM of KCl through a membrane with pore size of 0.1 μm . The initial concentration of GDH was 1.42×10^{-7} M, that of LDH was 2.86×10^{-8} M, and that of MDH was 1.82×10^{-7} M.

their concentrations are low, larger protein molecules tend to form multimeric molecular aggregates, which certainly hinder their microfluidic passage through the membrane. Furthermore, for biological protein molecules in general, their activities behave differently even at the same pH, thereby further complicating the separation process.

Effect of pH on Permeability Through an Artificial Membrane of Proteins Having Different MMs. For this study the surrogate proteins employed are MDH (having the lowest MM of the three proteins), LDH (having the intermediate MM of the three proteins), and GDH (having the highest MM of the three proteins). As mentioned previously, proteins, such as enzymes, behave differently at various pHs.^[30] Although the proteins used in this study are among the better-characterized ones, the exact molecular structure of these proteins in solution are not well defined and many molecular properties have not been elucidated. Therefore, we investigated the effect of pH on protein permeation through an artificial membrane that could eliminate many complications associated with biological membranes. As shown in Fig. 5A, permeation of MDH (the smallest) is more pronounced at a higher pH (7.4), and its permeation decreases at a lower pH (6.51). However, for LDH (the intermediate) (Fig. 5B), its permeation is highest at pH 6.95 and lowest at pH 7.4. Fig. 5C shows that GDH (the largest) has the highest permeation at pH 6.5 and the lowest at pH 6.95, although the membrane pore size for the GDH

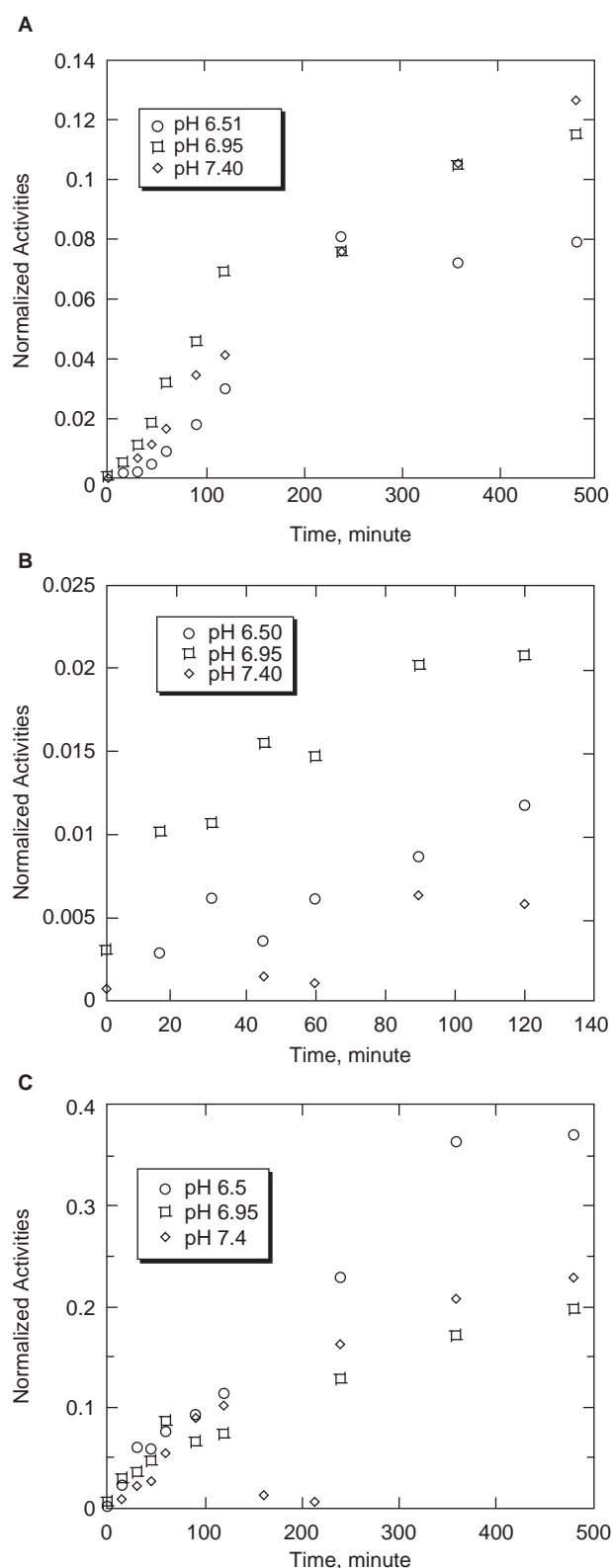


Fig. 5 (A) pH effect on permeability of MDH. All solutions started with same initial concentration (1.82×10^{-7} M). (B) pH effect on permeability of LDH. All solutions started with same initial concentration (2.86×10^{-7} M). (C) pH effect on permeability of GDH. All solutions started with same initial concentration (1.42×10^{-7} M).

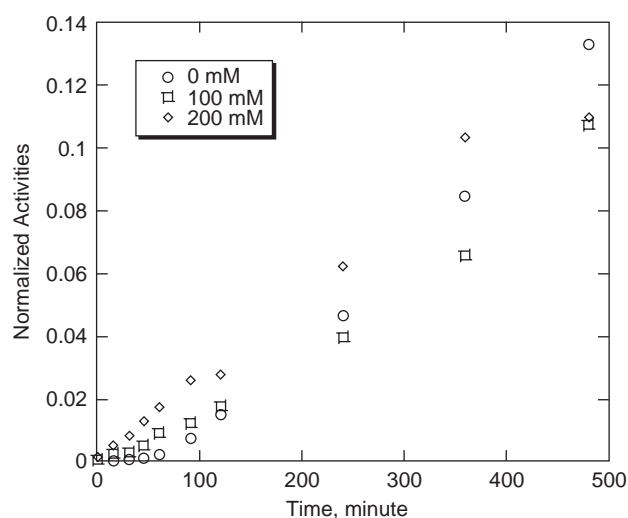


Fig. 6 Effect of electrolyte on LDH permeability. Initial concentration of LDH for all solutions was 2.86×10^{-7} M. Concentrations of KCl were 0, 100, and 200 mM.

experiments was 10 times that employed in the experiments with LDH and MDH. Therefore, if a membrane is used to separate biological protein molecules, one must take into consideration how the pH will affect their permeabilities to obtain the optimal result. It should be noted that the pH effect on protein molecules is different from the pH effect caused by hydrolysis on membranes made from organic ester(s).^[31]

Some factors we are neglecting here in our observations are the effect of pH on the surface charge of the membrane and the surface charge (zeta potential) of the protein molecules.^[32] Since the size (MMs) of protein molecules can vary more than orders of magnitude, it is reasonable to expect that their pI's of zero potential are different. When viewed in conjunction with the potential of significant sorption of protein on the surface of the membrane, the effect of pH on interfacial transport of protein is complex. A complete understanding of the phenomenon requires a detailed study of multiple-surface phenomena (see above) and is thus beyond the scope of our discussion.

Effect of Electrolyte Concentration on Protein Permeability Through an Artificial Membrane.

For this study the surrogate protein is LDH. From the results of our limited number of observations, there is no obvious effect of KCl as an electrolyte on the permeability of LDH (Fig. 6). As the concentration of electrolytes in the solution increases, the electrostatic interactions between the electrolytes and LDH molecules also increase.^[33] Thus, mobility of LDH molecules will decrease as a consequence. Hence, one would expect that the solution with no KCl would allow for the highest permeation rate, but this was not

obvious from the results of our experimental conditions. On the contrary, the results of our experiments appear to indicate that the solution with no KCl allows for the highest increase of permeation rate with time.

Effect of Membrane Pore Size on Protein Permeability Through an Artificial Membrane.

For this study the surrogate proteins employed are MDH (having the lowest MM of the three proteins), LDH (having the intermediate MM), and GDH (having the highest MM). Fig. 7A shows the permeation rate of LDH through membranes with pore sizes of 0.1, 1, and 5 μm . Our results show that permeation rate increases with pore size, although the rate of increase is not linearly

proportional to the pore size. When the membrane pore size approaches the size of the molecule to be separated, other factors, such as surface potential and charge of electrolytes in the solution will become critical.^[34] As shown in Fig. 7A, the results suggest that when the pore size reaches a certain threshold, the increase in permeation rate with pore size is reduced significantly. This could be the consequence of two possible effects. First, accumulation of LDH molecules in the pore tunnels of the membrane blocks a certain amount of LDH going through in the case of the larger pore membranes, but LDH molecules would not accumulate in the pore tunnels of the smaller pore membrane because of higher mass flux going through

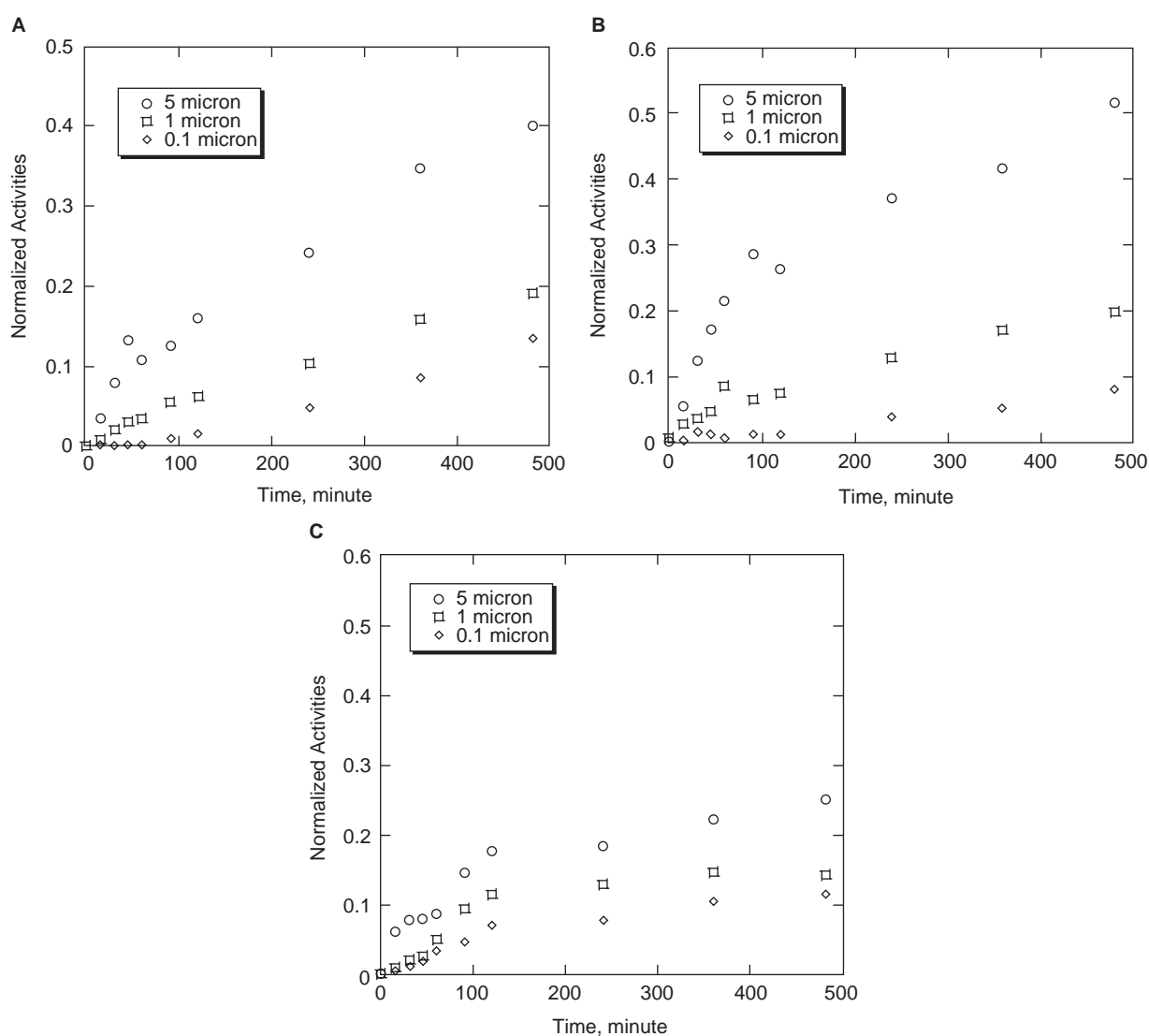


Fig. 7 (A) Permeability of LDH through different membrane pore sizes of 0.1, 1, and 5 μm . Initial concentration of LDH was 2.86×10^{-7} M in 10 mM of P-buffer with no KCl at pH 6.95. (B) Permeability of GDH through different membrane pore sizes of 0.1, 1, and 5 μm . Initial concentration of GDH was 1.42×10^{-7} M in 10 mM of P-buffer with no KCl at pH 6.95. (C) Permeability of MDH through different membrane pore sizes of 0.1, 1, and 5 μm . Initial concentration of MDH was 1.82×10^{-7} M in 10 mM of P-buffer with no KCl at pH 6.95.

the tunnels. Second, after enough LDH mass has been passed through the membrane from the concentrated side of the half-cell, the driving force of mass differential for the larger pore membrane is significantly less than that of the smaller pore membrane. Thus, slower permeation rate is the result. However, the first speculative mechanism cannot be verified until we can physically observe the LDH at the molecular level on the surface and pore tunnels of the membrane.

Observations on GDH (Fig. 7B) and MDH (Fig. 7C) show the same consistent results as those for LDH (Fig. 7A). It should be noted that the mass flux of MDH going through the membrane is much less than that of LDH or GDH, although MDH was the smallest of the three protein molecules. Evidently, membrane pore size is not the only controlling mechanism for mass permeation through the membrane. As technologies such as AFM, scanning electrochemical potential microscope, and other analytical instruments become more developed, verification of the permeation mechanism will soon be possible.

Effect of Gold-Palladium Alloy Coating of Membrane on Protein Permeability. The conditions of this set of experiments were essentially identical to those shown in Fig. 7A with LDH, except for the 0.1 nm of gold-palladium coating on the membranes. This set of experiments was included in this study to examine if a metal-coated membrane would behave differently from that of a noncoated one. Coating a membrane would alter the surface properties of the membrane drastically; thus, studies with a coated membrane can improve our understanding of ion transport in cells and organ surface. As shown in

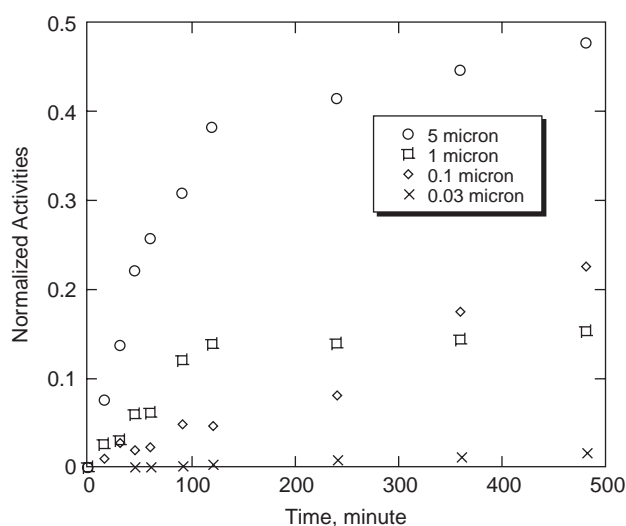


Fig. 8 Effect of gold-palladium coating on LDH permeability. Conditions of solutions in this figure were the same as in Fig. 7 except the gold-palladium coating.

Fig. 8, the permeation rate through the membrane with 0.1 μm pore size supersedes that through the membrane with 1 μm pore size at the later stage of the experiment. It appears that the metallic coating has allowed an increase in the permeation rate through the 0.1 μm membrane. The coating has also allowed increases in the permeation rate through the membranes of pore sizes 1 and 5 μm for the first 120 min, and then the protein permeation slows down immensely, especially that through the 1 μm membrane. This latter phenomenon could be caused by several factors, such as sorption and surface charge, as indicated by the elegant studies of Nishizawa, Menon, and Martin.^[19] This metal-coated membrane is conductive; thus, the surface charge of the membrane can be changed according to the potential being applied. Therefore, one can exploit the electrokinetic effect on the transport properties of proteins with this “smart membrane.”

Permeabilities of Proteins Having Different MMs. For this study the surrogate proteins employed are MDH (having the lowest MM of the three proteins), LDH (having the intermediate MM), and GDH (having the highest MM). For solute flux through a membrane with concentration gradient being the sole driving force, the solute flux can be written as:^[35]

$$F_i^0 = D_i^0 K_d \frac{\Delta C_i}{\delta_m} = K_p \Delta C_i$$

where D_i^0 is effective diffusion coefficient of species i , K_d is distribution coefficient, δ_m is thickness of membrane, and ΔC_i is the concentration difference across the membrane for species i . K_p is the coefficient of permeability and usually a constant in the range of interest for the membrane. For the conditions given in Fig. 4, K_p of GDH, LDH, and MDH were calculated to be 1.81×10^{-4} cm/sec, 2.64×10^{-4} cm/sec, and 3.97×10^{-4} cm/sec, respectively. As expected, the permeability of these protein molecules decreases with the increase of their MMs.

CONCLUSIONS

The field, as depicted by the title of this entry, “Protein and Cell Signaling with Biomaterials: Interfacial Transport,” is so new that there are many aspects where the lack of fundamental data does not allow us to make generalizations. One major point we would like to make at this juncture is that significant advances in this field can only be made if current and future studies adopt an interdisciplinary approach. The multidisciplinary nature of this field demands that biomedical engineers, biomedical researchers, and



chemists, and other physical scientists work together synergistically to address the fundamental issues. After all, in membrane separation technology, many of the interfacial and surface phenomena have been overlooked and are not well understood. Moreover, of the two major aspects of artificial membrane research (namely, the mechanistic and functional aspects), only the mechanistic aspect has been investigated in the last decade or so. By contrast, the functional aspect of artificial membrane research has not been addressed until recently.

In this entry, we trace the origin of the field of protein and cell signaling with biomaterials from the biomedical research on protein and cell signaling to the current concepts in biomaterial design and their applications in tissue engineering. We emphasize how and why interfacial transport is important in protein and cell signaling with biomaterials. We further advocate that interfacial transport in biological systems is too complex and not easily understood. However, we can begin to unravel such complexity by designing a more simple artificial membrane system to simulate discrete paradigms of the biological membrane systems. The availability of such a simpler artificial membrane system allows us to systematically investigate, in an interdisciplinary manner, the fundamentals of interfacial transport under vigorously controlled experimental conditions.

We report herein some of the results of our ongoing studies employing a simple separation cell to investigate qualitatively and quantitatively how large proteins are separated by artificial membranes at low concentrations under well-defined controlled conditions. To date, our results indicate the following: permeation of proteins increases with protein concentration; increase of ionic strength under the experimental conditions (≤ 200 mM) does not seem to induce a decrease in protein permeability; and protein permeation in general increases with increase in membrane pore size and decreases with increase in molecular mass of the protein. For biological macromolecules, which possess multiple functional groups, surface interactions with membranes are complex; prediction of the trend of permeability with variation of pH is difficult because of variations of zeta potential and surface charge that are chemical species dependent, even under a restrictive controlled environment. Change of surface potential can alter the order of size effect when molecules are passing through the microchannels in membranes (permeabilities).

The study of permeabilities of macromolecules and other chemical species in solution are valuable in the development of a treatment system using membrane technology to reclaim spent water for potable use, as well as for the development of tissue engineering. While artificial organs constitute an imminent

development in the biotechnology forefront, behavior of protein transport across artificial membranes must be well understood for this development to be successful. Thus, the results of our ongoing studies employing simple artificial membranes are highly relevant to the development of bioengineering, tissue engineering, and water/wastewater treatment applications. Clearly, for the rapid advances of such applications, more systematic studies using artificial membranes are warranted.

ACKNOWLEDGMENTS

We thank Garnet Rivera, John Cleary, and Kurino Ashizawa for their assistance in some of the studies reported in this entry. Our studies were supported, in part, by NIH Grant #P20 RR16454 from the Idaho BRIN Program of the National Center for Research Resources at the National Institutes of Health, and an FRC grant and an URC grant from Idaho State University.

ARTICLES OF FURTHER INTEREST

Biomaterials: Protein–Surface Interactions; Integrins; Mass Transfer in Tissue Function: Roles

REFERENCES

1. Chen, E.H.; Olson, E.N. Unveiling the mechanisms of cell-cell fusion. *Science* **2005**, *308* (5720), 369–373.
2. Wang, L.; Arcasoy, M.O.; Watowich, S.S.; Forget, B.G. Cytokine signals through STAT3 promote expression of granulocyte secondary granule proteins in 32D cells. *Exp. Hematol.* **2005**, *33* (3), 308–317.
3. Yamaguchi, Y.; Hironaka, K.; Okawaki, M.; Okita, R.; Matsuura, K.; Ohshita, A.; Toge, T. HER2-specific cytotoxic activity of lymphokine-activated killer cells in the presence of trastuzumab. *Anticancer Res.* **2005**, *25* (2A), 827–832.
4. Siebers, M.C.; ter Brugge, P.J.; Walboomers, X.F.; Jansen, J.A. Integrins as linker proteins between osteoblasts and bone replacing materials. A critical review. *Biomaterials* **2005**, *26* (2), 137–146.
5. Boonthekul, T.; Mooney, D.J. Protein-based signaling systems in tissue engineering. *Curr. Opin. Biotechnol.* **2003**, *14* (5), 559–565.
6. Orban, J.M.; Marra, K.G.; Hollinger, J.O. Composition options for tissue-engineered bone. *Tissue Eng.* **2002**, *8* (4), 529–539.
7. Vogel, V.; Baneyx, G. The tissue engineering puzzle: a molecular perspective. *Annu. Rev. Biomed. Eng.* **2003**, *5*, 441–463.

8. Ratner, B.D.; Bryant, S.J. Biomaterials: where we have been and where we are going. *Annu. Rev. Biomed. Eng.* **2004**, *6*, 41–75.
9. Lu, D.R.; Kadambi, B.N.; Lai, J.C.K. Protein and cell interactions with biomaterials—protein conformational considerations. In *Encyclopedic Handbook of Biomaterials and Bioengineering, Part A: Materials*, 1st Ed.; Wise, D.L., Trantolo, D.J., Altobelli, D.E., Yaszemski, M.J., Gresser, J.D., Schwartz, E.R., Eds.; Marcel Dekker, Inc.: New York, 1995; Vol. 1, 305–328.
10. C & EN. In *Chemical and Engineering News*, **2000**, Mar 6, 60–61.
11. Hwang, S.-T.; Kammermeyer, K. *Membranes in Separations*; Wiley-Interscience Publications: New York, 1975.
12. Tamagawa, H.; Sakurai, M.; Inoue, Y. Theoretical study of intermolecular interactions at the lipid water interface. 2. Analysis based on the Poisson–Boltzmann equation. *J. Phys. Chem. B.* **1997**, *101* (24), 4817–4825.
13. Tohda, K.; Umezawa, Y.; Yoshiyagawa, S. Cation permselectivity at the phase boundary of ionophore-incorporated solvent polymer membranes as studied by optical second harmonic generation. *Anal. Chem.* **1995**, *67*, 570–577.
14. Bailey, A.L.; Cullis, P.R. Membrane fusion with cationic liposomes: effects of target membrane lipid composition. *Biochemistry* **1997**, *36*, 1628–1634.
15. Morf, W.E. *The Principles of Ion-Selective Electrodes and of Membrane Transport*; Elsevier: New York, 1981.
16. Coombs, H.; Miller, A.J.; Sanders, D. Disruptive effects of protein on performance of liquid membrane-based microelectrodes. *Am. J. Physiol.* **1994**, *267* (pt 1), C1027–C1035.
17. Ibanez, J.G.; Batina, N.; Gutierrez-Granados, S.; Fernandez, L. Conducting polymers for environmental remediation. *Proc. ACS Environ. Chem.* **1998**, *38* (suppl. 1), 42–43.
18. Martin, C.R.; Mitchell, D.T. Nanomaterials in analytical chemistry. *Anal. Chem.* **1998**, *70* (9), 322A–327A.
19. Nishizawa, M.; Menon, V.P.; Martin, C.R. Metal nanotubule membranes with electrochemically switchable ion-transport selectivity. *Science* **1995**, *268*, 700–702.
20. Jirage, K.B.; Hulteen, J.C.; Martin, C.R. Nanotubule-based molecular-filtration membranes. *Science* **1997**, *278* (24), 655–658.
21. Yamada, K.; Gasparac, R.; Martin, C.R. Electrochemical and transport properties of templated gold/polypyrrole-composite microtubule membranes. *J. Electrochem. Soc.* **2004**, *151* (1), 14–19.
22. Leung, S.W.; Lai, J.C.K. Macromolecular diffusion through artificial membranes. Proceedings of WTA International Conference on Nano- and Microparticles in Water and Waste Water Treatment, Zurich, Switzerland, Sep 21–24, 2003, 363–370.
23. Bar-Cohen, Y. *Electroactive Polymer (EAP) Actuators as Artificial Muscles, Reality Potential, and Challenges*; SPIE Press, 2001.
24. Shahinpoor, M.; Kim, K.J. The effect of surface-electrode resistance on the performance of ionic polymer-metal composite (IPMC) artificial muscles. *Smart Mater. Struct. J.* **2000**, *9* (4), 543–551.
25. Nemat-Nasser, S. Micromechanics of actuation of ionic polymer-metal composites. *J. Appl. Phys.* **2002**, *92* (5), 2899–2915.
26. Baughman, R.H. Muscles made from metal. *Science* **2003**, *300* (Apr 11), 268–269.
27. Metcalf and Eddy Inc. In *Wastewater Engineering, Treatment, and Reuse*; McGraw Hill: New York, 2003.
28. DeGroot, S.R. *Thermodynamics of Irreversible Processes*; North Holland: Amsterdam, 1961, 186.
29. Skilland, A.H.P. *Diffusional Mass Transfer*; Wiley and Sons: New York, 1974.
30. White, J.S.; White, D.C. *Source Book of Enzymes*; CRC Press: New York, 1997.
31. Vos, K.D.; Burris, F.O.; Riley, R.L. Kinetic study of the hydrolysis of cellulose acetate in the pH range of 2–10. *J. Appl. Polym. Sci.* **1966**, *10*, 825–832.
32. Weber, W. *Physicochemical Processes for Water Quality Control*; Wiley and Sons: New York, 1972.
33. Snoeyink, V.L.; Jenkins, D. *Water Chemistry*; Wiley and Sons: New York, 1980.
34. Kesting, R.E. *Synthetic Polymeric Membranes, a Structure Perspective*; Wiley and Sons: New York, 1985.
35. Lonsdale, H.; Meerten, U.; Riley, R.L. Transport properties of cellulose acetate osmotic membranes. *J. Appl. Polym. Sci.* **1965**, *9*, 1341–1362.

FURTHER READING

Leung, S.W.; Lai, J.C.K. Protein and cell signaling with biomaterials: influence of surfactants. In *Encyclopedia of Biomaterials and Biomedical Engineering*; Wnek, G.E., Bowlin, G.L., Eds.; Marcel Dekker, Inc.: New York, 2005.

Protein Delivery Systems

Nicholas A. Peppas
Cristina Donini

*Departments of Chemical and Biomedical Engineering and Division of Pharmaceutics,
University of Texas at Austin, Austin, Texas, U.S.A.*

INTRODUCTION

More effective and tailored drug delivery systems (DDS) are continuously investigated, tested, and, eventually, marketed. In the last 30 yr, the pharmaceutical industry has made substantial steps in this direction. Drug products containing proteins and peptides, as therapeutic agents, are a perfect example of these two intimately correlated aspects: route of administration and quality of life. In the past 60 yr, hormones, serum proteins, and enzymes have been extensively used in therapy as active ingredients.^[1,2] But recently, molecular biology and, especially, recombinant deoxyribonucleic acid (DNA) technology have made tremendous progress resulting in the availability of a wide variety of peptide- and protein-based therapeutics for poorly controlled diseases. For this class of drugs, however, the parenteral route represents today the predominant route of administration, as the bioavailability of alternative routes is extremely low and variable.

The development of a successful non-parenteral delivery system for proteins implies challenges including the protection of the therapeutic agent from hostile physiological conditions, otherwise detrimental for the biological activity, and gentle manufacturing conditions to maintain the original molecular conformation of the protein. One such agent widely studied as model protein is insulin, a 51 amino acid polypeptide, used in therapy for the treatment of diabetes mellitus. Diabetes mellitus is a group of metabolic disorders characterized by hyperglycemia that affects a sizeable portion of the population worldwide, estimated to increase in the next 20 yr of more than 30% reaching the size of an epidemic.^[3] Furthermore, the complications coupled with it are debilitating and at times fatal. For these reasons, a chronic therapy is required. To maintain normal glucose levels in the blood and mimic the physiological profile of the secretion of the protein, frequent daily injections (up to four) of insulin are needed.

One of the most promising new methods of non-parenteral protein delivery is oral delivery. For this process, various types of specialized polymeric carriers are used. For example, hydrogels have been studied as suitable candidates for oral drug delivery of peptides.

Particularly the complexing hydrogels that respond to pH changes in the surrounding environment have been extensively studied. For example, we developed new cross-linked copolymers^[4] composed of poly(ethylene glycol) (PEG) grafted on poly(methacrylic acid) (PMAA) backbone, henceforth designated as P(MAA-g-EG), by UV-initiated bulk-free radical polymerization. They possess the unique characteristic to form interpolymer complexes in response to changes in pH of the environmental fluids. The interpolymer complexes are formed owing to hydrogen bonding between the protonated carboxylic acid group of PMAA and the ether group of the PEG chain. In particular, they remain in the collapsed state at low pH values, similar to those found in the stomach. Therefore, the release of the protein is impeded, and the protein entrapped in the network can be protected from the undesirable conditions found in the stomach. When the pH rises up to approximately 6, the network swells. This unusual behavior seems very promising for an oral formulation that would be able to release the labile therapeutic agent only in a more favorable environment such as that encountered in the upper small intestine. The expansion of the system would allow the polymeric chains to interact with the mucosal lining, causing the tight junctions to slightly separate, thus enhancing the transport of the protein into the systemic circulation.^[5]

PROTEIN STRUCTURE AND FORMULATION DESIGN

Proteins and Peptides: The Internal Architecture

Proteins are high molecular weight molecules, characterized by high-ordered, hierarchically organized structure. It is possible to describe their structure at four separate but interdependent levels: primary, secondary, tertiary, and quaternary structures, mainly stabilized by weak physical forces.^[6] The primary structure is essentially the amino acid sequence connected one by one via peptide bonds to form a polypeptide chain. This “linear” sequence is responsible for determining the higher levels of organization of the molecule.

A single change in the primary structure can have a profound outcome on the overall arrangement and function of the protein. For instance, hemoglobin, the oxygen carrier in the blood, is a tetramer consisting of two α chains and two β chains. A change in the sixth residue of a β chain from glutamic acid (glu) to valine (val) causes deformation of the red blood cells, an extremely serious condition called *sickle cell anemia*.

The second level in the protein organization hierarchy is known as secondary structure, where chains of amino acid residues are further organized in space forming regularly repeating patterns. The most common secondary structures in proteins are the α -helix, a rod-shaped structure wrapped around an imaginary cylinder and stabilized via hydrogen bonds mainly, and the β -pleats, where the backbone amino acids atoms form a zigzag pattern (the β sheet). The structure of many proteins consists mainly of α -helices, some others mainly of β sheets, and still others a combination of both, the random coils.

The global structure of a single polypeptide chain is known as tertiary structure. Groups of α helices or β sheets (domains) fold with one another to form a three-dimensional architecture. Hydrogen bonding involving groups from both the peptide backbone and the side chains is important in stabilizing the tertiary structure. In addition, an important role is also played by covalent bonding (disulfide bridges) and hydrophobic interactions between side chains. The polypeptide chain folds such that the side chains of non-polar residues are hidden within the structure and the ones carrying the polar moieties are exposed on the outer surface.

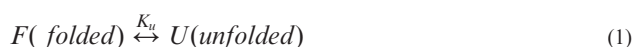
Finally, the quaternary structure refers to the arrangement of two or more polypeptide chains (tertiary structures) into a multisubunit structure. However, not all proteins exhibit quaternary structure. Usually, each polypeptide within a multisubunit protein folds quite independently into a stable tertiary structure, and the folded subunits associate with each other to form the final structure. Quaternary structures are stabilized mainly by non-covalent interactions such as hydrogen bonding, van der Waals interactions, and ionic bonding. The hierarchy described above generates a complex architecture that dictates the biological activity of the protein. Here lies the great difference between proteins and low molecular weight drugs. The large amount of functional groups and the complex and labile three-dimensional structure render them prone to a variety of degradation and inactivation pathways. In general, the biological activity can be lost because of chemical or physical alterations. Examples of chemical alterations that can be detrimental for the protein are reactions of deamidation. Glutamine and asparagine under certain condition can undergo the deamidation reaction that can strongly affect the activity of the protein. Oxidation and reduction can usually affect free cysteine.^[6,7]

Also, owing to the primary reactions of certain groups, reactions of aggregation and internal cross-linking can occur thus leading to significant loss of biological activity. A common example of cross-linking within the protein subunits is the formation of disulfide bridges or the rearrangement of the existing S-S bonds.^[8]

Another important chemical reaction in terms of loss of activity is proteolysis. This reaction is owing to the breakdown of the peptide bonds because of harsh conditions such as long exposure to extreme pHs or high temperature or, again, because of proteolytic enzymes. They may also undergo adsorption processes that can irreversibly denature them. A polypeptide chain generally folds into a sole three-dimensional conformation, unique in ensuring biological activity. Examples of physical alterations are conformational changes and non-covalent aggregation (self-association). Therefore, a wrong handling or manufacturing procedure may lead to irreversible changes in the arrangement of polypeptide chains, which become inactive.

For this reason, the formulation and drug delivery of such pharmaceuticals represent a tremendous challenge for the pharmaceutical scientist. The stringent interplay between the biological activity and the final conformation of such a complex structure is the crucial issue to be addressed in protein drug delivery. It is well known that environmental conditions such as temperature, moisture content, salinity, or pressure can have a profound, and, at times, detrimental reflection on the internal architecture of these macromolecules. The protein conformation, indeed, exhibits a marginal stability.

The protein stability has several definitions.^[8-10] Generally it can be described in terms of *thermodynamic stability* and *kinetic stability*. A protein unfolds and refolds following a two-state process, the equilibrium of which can be described as



where K_u is the equilibrium constant for unfolding. The rate of unfolding and the reversibility of the process reported in Eq. (1) are the key issues. In other words, a protein may be fully unfolded and inactive in certain conditions; yet, it will refold and fully recover its activity in the "original" or "ideal" condition. Hence, from a functional point of view, this may be all that is required for the molecule to be classified as thermostable. However, from a thermodynamic point of view it would be classified as thermally unstable.

Irreversible loss of protein folded structure is represented by



where k_i is the rate constant for some irreversible inactivation processes such as aggregation, disulphide exchange, proteolysis, irreversible subunits dissociation, chemical degradation, and so forth.

In light of the aforementioned features, protein stability must be investigated at molecular level, taking into consideration that all the inactivation processes described above can occur during preparation, storage, and release from the dosage form. Thus, the formulation has to be designed in such a way that these delicate structures are stable during preparation, purification, formulation, storage, and administration to the patient. The structure of the whole molecule should be preserved during all stages ensuring proper pharmacokinetic and pharmacodynamic profiles. Furthermore, therapeutic proteins may need special formulation procedures to protect them against the physiological conditions they have to face before reaching the target.

Protein Absorption and Formulation Issues

To understand the limiting step of administration of a protein as drug and, on the whole, of a therapeutic agent, one must understand the concept of drug kinetics. In short, pharmacokinetics refers to what human body does toward a medicine. It involves the processes of drug absorption into the body, distribution to various tissues, metabolism, and excretion.

Absorption can be defined as the process by which a drug passes from its site of administration into the bloodstream. When the therapeutic reaches the systemic circulation, the blood carries the drug to all the organs, including those on which the molecule acts. The rate, ease, and degree of absorption are all related to the route of administration, to the physicochemical properties of the drug itself, and to the formulation intended as the actual dosage form.

The routes of administration can be broadly classified as enteral and parenteral. The enteral route refers to those dosage forms wherein the drug is ultimately absorbed within the intestine such as oral and rectal. However, the term “enteral” is used more broadly to refer to any route of administration that involves absorption anywhere from the mouth to the rectum. Thus, sublingual administration is often categorized as an enteral route of administration.^[11]

Parenteral, on the other side, means introduced by way of other routes outside of the intestines and not by way of the alimentary tract. There are several sites of administration via this route. The three primary routes of parenteral administration are subcutaneous, intramuscular, and intravenous, although there are others such as intracardiac and intraspinal. In the intravenous administration, the drug is injected directly into the blood, most commonly into a peripheral vein. The

advantages of such a mean are the rapid achievement of concentration in the blood and precision in delivering the effective dosage. Among the disadvantages, however, are the high initial concentration that can fall into toxic levels, especially for drugs with a low therapeutic index, the fact that once the drug is injected there is no retreat, and, not least, the need of skilled personal.

Generally, the delivery of proteins is performed by parenteral administration (subcutaneous or intramuscular injections). The problem with this kind of administration is related to the short half-time of the protein, which results in the necessity of frequent administrations to obtain therapeutic effectiveness. Obviously, multiple injection regimens do not meet patient acceptance. To avoid lower compliance, one approach involves producing smart parenteral delivery systems that reduce the frequency of injection by providing a sustained release of the drug over time. Examples of this type of devices could be found in insulin pumps,^[12] which provide a continuous subcutaneous infusion of the protein trying to mimic the physiological circadian rhythm of the protein. A needle is inserted only every three or four days giving the patient a wide flexible lifetime. The use of genetically modified protein analogs can lead to an increase of the low availability of the injected protein. Again, implantable infusion devices could represent a reasonable intelligent system for delivering proteins without the need of multiple injections.

However, the development of non-parenteral protein delivery systems remains an intriguing challenge for many investigators in the pharmaceutical field as a reasonable alternative to the parenteral route.

A wide range of dosage forms have been proposed for the delivery of proteins via enteral routes such as nasal, pulmonary, buccal, rectal, vaginal, transdermal, and oral routes. Among these, the oral administration has always represented the “holy grail” for proteins and peptides delivery, especially in chronic treatment. Compared with others, enteral routes, oral administration has always been considered the most natural, convenient, and safe mean of administration. The advantage in terms of patient compliance is amplified by the potential cost saving to the health-care industry, as no skilled personnel or sterile processes are required.

So far, in despite of all efforts spent in reaching the target of safe and effective oral formulations for peptides and proteins, relatively small progress has been made.^[13] The main reason for the absence, to date, of a successful oral protein formulation can be ascribed to the characteristic of the drug molecule itself. To have adequate bioavailabilities, the protein should possess robustness toward the highly acidic environment of the stomach, should be able to pass across intestinal cell membranes, and it should have endurance to destruction by hepatic metabolism (first-pass effect). In most

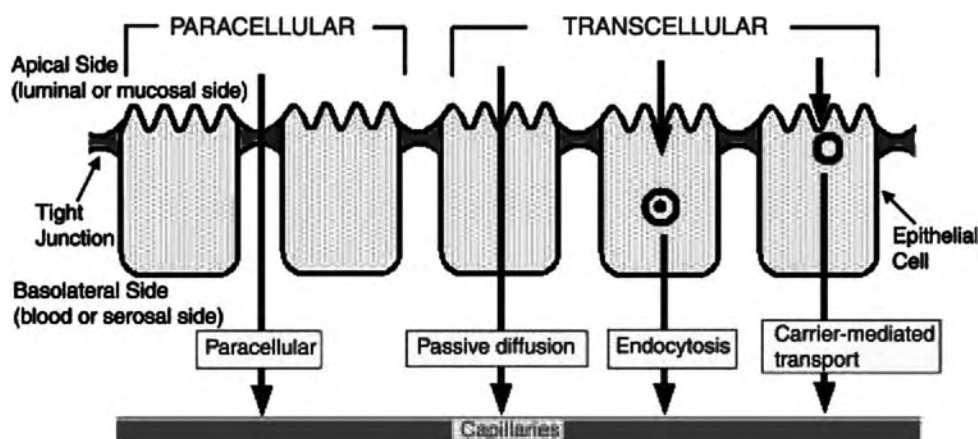


Fig. 1 Types of transport across intestinal membrane.

cases, such drugs do not have intrinsic characteristics to withstand these conditions.

Proteins must bypass physical barrier, enzymatic digestion, chemical degradation, and intrinsic low stability.^[14] The former is represented by the single layer of intestinal epithelial cells that form the barrier between the lumen of the gastrointestinal (GI) tract and the body fluids. Generally, transport of molecules across the intestinal epithelium can be achieved by three transepithelial pathways: transcellular (across the cell membranes), carrier-mediated active or facilitated transport, and paracellular transport (i.e., through the intercellular spaces between adjacent cells) (Fig. 1).

For many drugs, the predominant route of transport across the intestinal epithelium is the transcellular route. The molecule diffuses through the cell from one side to the other. This usually happens for molecules such as oxygen and carbon dioxide. In the carrier-mediated transport, the molecule interacts with binders exposed on the surface of the cell membrane. A complex is formed and this crosses the lipid layer to the inside of the cell. The bound molecule complex then dissociates inside the cells, and the free molecule moves to the other side of the cell where the same mechanism takes place and allows the molecule to reach the outside environment.

Paracellular transport, on the other hand, occurs when the molecules move through the layer by passing between adjacent cells. Small hydrophilic compounds and charged molecules are mainly transported via paracellular route. Epithelial tissue is tightly bound together into sheets, called epithelia, and specialized junctions between the cells enable this tissue to form barriers to the movement of water, solutes, and cells from one lumen side to the systemic circulation. The presence of these “occluding” junctions, known as tight junctions, often hampers the paracellular path. Under various physiological conditions, material transport across tight junctions occurs in a regulated manner.

Absorption of glucose from the intestine, for instance, occurs via paracellular way, across tight junctions, particularly when luminal glucose concentrations are elevated, saturating the Na^+ -glucose cotransporter in the apical plasma membrane. However, under normal conditions, this way is restricted to molecules with molecular radii less than 11 \AA and therefore is not accessible to large compounds such as peptides and proteins.^[15]

The Role of Enzymes and Environmental pH in Protein Absorption

The intestinal absorption of drugs can be reduced by the presence of efflux proteins like P-glycoproteins and/or metabolism in the gut.

The proteolytic degradation in the GI tract is another important impediment that limits the effectiveness of peptide- and protein-based delivery systems. Enzymes located throughout this part of the body are capable of denaturing or destroying these fragile molecules. Proteases are the enzymes deputed for the digestion of proteins in the stomach and they represent the first enzymatic barrier that the protein has to face along the GI tract. Passing the pylorus and arriving in the small intestine, pancreatic enzymes such as trypsin, chymotrypsin, and carboxypeptidases are other enzymes responsible for degradation of ingested proteins. Finally, degradation also can take place on the brush-border membrane of the intestinal cells because of various peptidases. It is very unlikely that such a labile molecule can withstand all these possible degradation. A protection against proteolytic degradation, therefore, is necessary.

The sharp pH gradients that characterize the GI tract are another problem to bear in mind. The gastric pH has been measured to range from 1.4 up to 2.1 on an empty stomach. It can reach pH values around 4 or 5

when food is present.^[16] In such conditions, reactions of aggregation, deamidation, can easily occur and can be detrimental for the biological activity of the drug.

Another hurdle in developing an effective oral delivery formulation may be the harsh conditions of the manufacturing process. As described above, proteins have very large globular structure, a complex and fragile internal architecture that defines their unique biological activity, chemically weak bonds and numerous reactive groups on their side chains. All these aspects render these compounds difficult to handle, sensitive to the interactions with excipients. Any disruption in one of the levels of the molecule architecture (primary, secondary, or tertiary structure) can lead to inactivation, reversible or irreversible, of the protein. These disruptions may occur upon slight changes in the microenvironment. A rise in the temperature, a change of the pH, the presence of a particular solvent, and a change in the crystalline form of the protein represent a common scenario during the manufacturing process of a formulation. Therefore, it is clear that an accurate evaluation of the physicochemical characteristics of the protein, a careful analysis of the interaction between the drug molecule and the excipients, and the impact of the manufacturing processes on the protein structure are important issues in developing a new formulation.

Several strategies to improve oral absorption of peptides and proteins have been investigated to target both the low physiological permeability and the stability of these compounds.

The main approaches include the addition of penetration enhancers to increase the permeability across the epithelium, the use of proteolytic inhibitors within the final dosage form to surmount the enzymatic degradation, and gastroenteric coating to protect the drug against the acidic environment of the stomach.

A common approach to overcome the low epithelial permeability has been the use of permeation enhancers such as bile salts, fatty acids, calcium chelating agents, surfactants, and the like. These compounds apparently have the ability to loosen the intestinal tight junctions and therefore improve the paracellular transport.

Calcium chelators administered together with proteins improved significantly the absorption through the epithelium. The improvement observed was owing to subtraction of calcium ions from the tight junctions, resulting in the opening of water channels.^[17,18]

Bile salts and their derivatives have been also widely used to enhance the penetration through the cells membrane. They are amphiphilic compounds that above the critical micellar concentration tend to self-associate forming vesicles referred to as micelles, where the hydrophobic groups orientate themselves toward the inside creating a lipophilic core, whereas the hydrophilic groups are exposed to the medium rendering the whole structure water soluble.

The mechanism by which these compounds increase the absorption of proteins and peptides through the intestinal mucosa has not been fully understood. However, it appears to be a combination of mechanisms. The formation of aqueous channels in the cell membrane can be a reason for the increase in cell membrane permeability. Ion pairing between bile salts and proteins and peptides that become more lipophilic could be another effect.^[19] Also micellar solutions have shown the capability to increase the fluidity of the mucosal membrane and consequently enhance protein absorption.^[20] A variety of bile salts like taurocholate, cholate, and glycocholate have been tested through the years, and they have been found to be equally effective in enhancing protein absorption.

Surfactants are also used for improving penetration through physiological mucosa. These agents have the peculiarity to lower the surface tension of water and promote micelles formation. They modify the properties of biological membranes and alter their permeability and transport characteristics.^[21] Sodium lauryl sulphate, saponin, and Laureth-9 are some examples of surfactants used for improving absorption across biological membranes. They have all shown to have membrane-solubilizing effects toward membranes.

In any case, the ability of all these agents in enhancing epithelial penetration is related to serious and at times irreversible modifications or disruptions of the integrity of biological membranes. The potential toxicity of these compounds, therefore, has worried the scientific community. One big threat of such excipients is their ability in altering the membrane-barrier functions. Moreover, rendering the paracellular way pervious would mean exposing the systemic circulation indiscriminately to the intestinal tract content, including toxins and undesired compounds.^[22] Considering the gravity of the damages that this can cause (solubilize and therefore extract important components from the cell membrane and irreversibly lose pivotal barrier functions of the mucosa), the acceptance to these compounds as excipients in drug formulation is very much unlikely.

Other ways have been investigated to improve the epithelial penetration of proteins and peptides. A target widely tackled has been the opening of the tight junctions. Recently, zonula occludens toxin, designated as ZOT, a protein elaborated by *Vibrio cholerae*, has been discovered.^[23] The protein has shown the particular ability to interact with specific surface receptors present in the mature cells of the small intestine, but not in the colon. *In vitro* and *in vivo* studies have demonstrated that ZOT effects are reversible and that the compound is not cytotoxic. Therefore, it seems to be a promising alternative to other enhancers, but the studies on this direction are still in progress. Up to date, however, bioavailabilities reached using these expedients are still too low.

Enteric coating is another approach widely used in pharmaceuticals for providing protection of drug against the harsh environment of the GI tract. It has been successfully used for small drug molecules and it has been extensively investigated for proteins and peptides as well. The concept behind it is to encapsulate the protein in pH-sensitive materials and release it only in a more “friendly” environment. Polyacrylic polymers have been used for this purpose. Eudragit[®] polymers are a well-known example of enteric coatings used in pharmaceutical industry and it has been proposed also for a “site-specific” insulin release.^[24] However, the bioavailability obtained with this strategy was only around 0.1%.

Methods of Elimination of Proteolytic Activity

Numerous strategies have been attempted for inhibiting proteolytic enzymes responsible of the inactivation of proteins. For example, it has been shown that the use of protease inhibitors such as bacitracin, soybean trypsin inhibitor, and camostat mesilate could improve the intestinal absorption of these compounds.^[25] Increasing attention has been focused on methods of administering proteins and peptides continually for prolonged time periods and in a controlled manner. The primary method of accomplishing this controlled release has been through incorporating these bioactive agents into polymeric particles.

As a matter of fact, polymers are gaining increasing importance in pharmaceutical technology owing to their great versatility in applications. First introduced in the medical field such as suture means and the like, they have become quickly excipients widely used as drug encapsulants for protecting an active agent during its passage through the body until its release, as vehicles of drug delivery and as means for controlling the drug release. A conventional tablet usually releases the drug in regions of the body where it may not be at its most effective, typically reaching a maximum concentration, which then decays exponentially until the next administration. In some cases, this is not the best pharmacokinetic profile. This can be overcome to some extent by sustained/retarded release, and/or selective delivery of the drug to the targeted organs. The use of polymers in pharmaceutical formulation has paved the way for controlled release. One of the most remarkable features of these compounds is the possibility to tailor the characteristics of the final polymeric system upon need, combining different monomers or polymers.

There are several formulations containing microspheres loaded with proteins on the market.^[26] Lupron Depot was one of the first controlled-release delivery devices for the administration of leuproreline acetate,

an LH–RH analogue, and contained microspheres. De-Capeptyl SR is another example of a formulation containing microspheres of a lactide–glycolide copolymer loaded with triptorelin, belonging to the same pharmaceutical category of leuproline and so on. All these systems are nevertheless designed for injections or implantable devices.

Nowadays, polymeric matrices are receiving significant attention for the oral delivery of proteins. For example, Lowe and Temple^[27] studied nanoparticles of polyacrylamide (PAAm) and poly(isobutyl cyanoacrylate) for the oral delivery of human calcitonin and insulin. These two different polymers have shown the ability to reduce enzymatic degradation in the GI tract.

Yet, the most exciting systems in the controlled release area lie in the responsive delivery systems.^[28] The advantages over other controlled release systems include the release of the therapeutic agent only in response to an external stimulus such as change in pH, temperature, ionic strength, and electromagnetic radiation in the surrounding environment. This behavior can be of particular interest when a constant concentration of the drug in the blood is not desirable, while the release of the drug would be ideal if occurring as a response to a physiological change. The most used example is the delivery of insulin. In patients affected by diabetes mellitus, insulin has to be released in response to an increase in the blood glucose level, to mimic the physiological release from the pancreas.

Many other examples of such ideal disease treatment could be found in hormone administration. Hormones have usually a circadian release into the body. Hence, the ideal delivery of such compounds should be a stimuli-sensitive release, instead of a “conventional” controlled release. This kind of devices is also known as “smart” DDS. Hydrogels belongs to this category. Both cationic and anionic hydrogels are becoming very popular in this field.^[29] Structure and properties of these polymer networks can be manipulated to achieve the desired features. Permeability and swelling behavior are indeed strongly related to the chemical nature of the polymers composing the final network.

An example of such network is pH sensitive hydrogels. These polymeric materials contain ionizable side groups such as carboxylic acids or amine groups that provide them the peculiar characteristic of responding to pH changes.

The most studied include PAAm, poly(acrylic acid) (PAA), and PMAA.^[30] These copolymers exhibit pH-dependent swelling behavior and are able to release drugs depending on the pH of the environmental fluid. These systems seem to be promising for the formulation of protein- and peptide-based drugs.

pH-Responsive Complexing Hydrogels as Oral Delivery Systems for Peptides and Proteins Administration

The characteristic pH profile of the GI tract makes pH-sensitive polymeric systems promising candidates for the oral delivery of peptides and proteins, owing to their ability to respond to the surrounding environment. Hydrogels, in particular, have captured many consensuses among researchers because of their benign characteristics. They can be defined as three-dimensional networks physically or chemically cross-linked. In any case, their cross-linked nature renders them insoluble. Owing to their hydrophilicity, they can absorb large amounts of water.^[31] Hydrogel's structure can be designed according to the desired application. Reacting, for example, monomers with ionic functional groups will provide the final network with environmental sensitivity, whereas other functional groups could allow the carrier to target a specific area of the body and protect the therapeutic agent from otherwise detrimental physiological conditions.

Hydrogels contain hydrophilic functional groups such as $-\text{OH}$, $-\text{COOH}$, $-\text{CONH}_2$, $-\text{CONH}$, and $-\text{SO}_3\text{H}$,^[32] which allow large amounts of water to be incorporated in the polymer. However, the presence of cross-links, as already mentioned, renders the polymer networks water insoluble. According to the ionic nature of the pendant groups, hydrogels can be classified as anionic gels or cationic gels. Hydrogels containing pending groups such as $-\text{OH}$, $-\text{COOH}$, and $-\text{SO}_3\text{H}$ are classified as anionic, whereas cationic gels may contain amine groups.^[33]

Changes in the environment such as pH and ionic strength will turn the pendant functional groups in their unionized or ionized forms according to the ionic nature and this will affect their swelling behavior. For anionic hydrogels, the functional groups are weak acids and therefore are unionized below their $\text{p}K_a$. When the surrounding pH is increased above the $\text{p}K_a$, the functional moiety ionizes and the electrostatic repulsion causes the network to swell. The opposite occurs with cationic hydrogels. The characteristic equilibrium swelling behavior of the two types of hydrogels is shown in Fig. 2. This mechanism of swelling is often reversible. Many hydrogels can fully recover from deformation when the external stimulus terminates.

A promising class of stimuli-sensitive hydrogels is complexing hydrogels. Interpolymer complexes occur within the network owing to interactions between repeating units in the side chains. The resulting tie points act as additional physical cross-links.

These systems can be used to protect therapeutic agents such as proteins from the harmful physiological conditions that could degrade such agents before they produce their therapeutic effects.^[31,34]

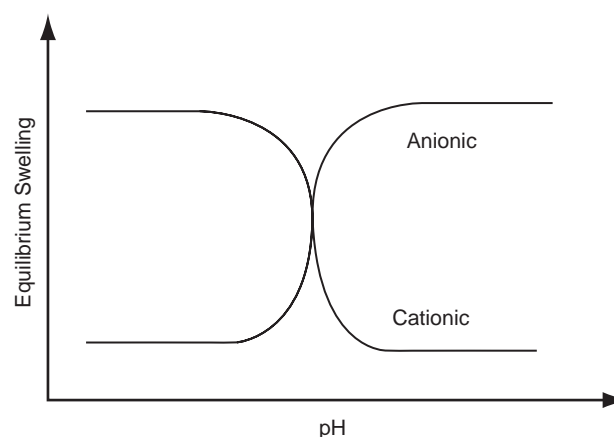


Fig. 2 Equilibrium swelling of anionic and cationic hydrogels as a function of pH.

PAA hydrogels that show pH-sensitivity have been utilized as stimuli-sensitive mucoadhesive carriers. We have studied these systems in detail^[35] and we have examined the effect of incorporation of PEG grafts into a PAA network. The polymeric network was prepared by free-radical solution polymerization. Different methacrylates with pendent PEG chains of various lengths were used as monomers in addition to the acrylic acid. Infrared spectroscopy was used to investigate the presence of hydrogen bonding in the networks. It is the sensitivity of this bond to the external environment that provides the pH-dependent swelling behavior to it and, in general, to all the hydrogen-bonding interpolymer complexes. The copolymer networks of poly(acrylic acid-g-polyethylene glycol) showed a pH-sensitive swelling behavior related to the length of the PEG graft. Networks with the largest PEG grafts showed the highest sensitivity to external signals.

Other pH-sensitive complexation hydrogels studied in our group, which have been used as DDS are the hydrogels of poly(methacrylic acid-g-ethylene glycol), henceforth designated as P(MAA-g-EG).^[3,36–38] Interpolymer complexes in this system are formed owing to hydrogen bonding between the etheric oxygen of the graft PEG chain and the acidic groups of MAA.

Furthermore, we confirmed^[39] that at low pH, the network is collapsed because of the formation of temporary physical cross-links. In Fig. 3 the effect of pH is reported on complexation–decomplexation mechanism of the polymer network. P(MAA-g-EG) hydrogels showed reversible, pH-dependent swelling behavior. This is owing to hydrogen bonding between the hydrogen in the carboxylic acid groups of the PMAA and the oxygen in the grafted PEG chains. At low pH values, which are lower than $\text{p}K_a$ of the carboxylic acid, the carboxylic acid groups of the PMAA protonate and hydrogen bonding can occur resulting in collapsed hydrogels. At higher pH values, these acid groups

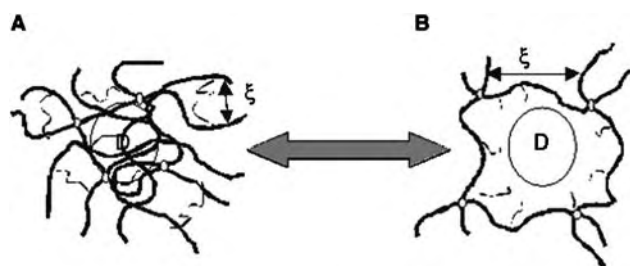


Fig. 3 Network structural changes because of variations on environmental pH. At low pH (A) interactions between the tethered grafts with the protonized ionic moieties increase. The formation of interpolymer complexes occurs and the system is in the collapsed state. Higher pH (B) disrupts the interpolymer complexes and ionic moieties deprotonize leading to extensive swelling.

become ionized and then the hydrogen bonding breaks down. As a result, the hydrogels swell to a high degree because of electrostatic repulsion within the copolymers (Fig. 3) and water is allowed to enter.

Our group studied in detail the properties and stability of the polymer complexes formed in the P(MAA-g-EG) system. We investigated the equilibrium swelling behavior of the P(MAA-g-EG) hydrogels as a function of pH.^[40] The equilibrium swelling ratio, q , of the samples was calculated as a ratio of the equilibrium weight to the weight of the dry sample. At low pH values, when complexation occurs, the polymer networks were collapsed and the swelling ratio was low, while at high pH values, *circa* 8, the swelling ratio increased and reached its maximum value.

We were able to produce microspheres^[41] of poly(methacrylic acid-g-ethylene oxide) [P(MAA-g-EO)] network by using a suspension polymerization reaction. The reaction was carried out in silicone oil at high temperature and led to the formation of copolymer microspheres. The system showed the same pH-dependent swelling behavior as before. The aforementioned systems have been studied as possible intelligent carriers for the administration of peptides and proteins. These pH-sensitive hydrogels can provide protection to the drug loaded into the system and they can target the release of it into a more suitable environment such as the upper intestine.

In the work of Drummond et al.^[42] the uptake and release kinetics of budesonide from P(MAA-g-EG) systems in vitro as well as the pharmacokinetics following nasal administration were investigated. The loading efficiencies of budesonide into the polymers were examined as a function of concentration of ethanol in water. Ethanol was necessary to help the solubility of budesonide but hindered hydrogel swelling at pH 7.2. It was found out that the maximum loading was obtained using a 25% ethanol solution. A constant

plasma concentration was attainable for 45 min after nasal administration of the system.

We also investigated^[43] the use of P(MAA-g-EG) hydrogels as oral delivery vehicles for insulin. P(MAA-g-EG) microparticles were obtained by crushing films of copolymers and, subsequently, loading them with insulin by imbibition. The polymeric microspheres were then administered orally to healthy and diabetic Wistar rats. In the acidic environment of the stomach, the gels were unswollen. The insulin remained in the gel and was protected from proteolytic degradation. In the basic and neutral environments of the intestine, the complexed network dissociated and this resulted in rapid gel swelling and insulin release. Within 2 hr of administration of the insulin-containing polymers, strong dose-dependent hypoglycemic effects were observed in both healthy and diabetic. These effects lasted for up to 8 hr following administration.

In another work from Lowman,^[44] it was concluded that the addition of PEG grafted to the gels was crucial for the in vivo delivery of the protein. PEG chains participate actively in forming the interpolymer complexes, but also act as peptide stabilizers and enhance the mucoadhesive characteristics of the polymer network when in the swollen state. This is owing to the fact that when the network is in the relaxed state, PEG chains could reach out and act as anchors, thus promoting adhesion. On the other hand, in the complexed state, the grafts cannot fully extend, impeding to penetrate the mucosa.

We examined the feasibility^[45] of using the P(MAA-g-EG) as a carrier for the oral administration of salmon calcitonin. They found out that the loading efficiency of the system was highly affected by the amount of solvent used during hydrogel preparation. They explained the transport mechanism as relaxation controlled. The diffusion coefficient was estimated using a heuristic Fickian/relaxational model.

In addition, we conducted studies^[46] to investigate the P(MAA-g-PEG) hydrogel ability to inhibit proteolytic enzymes. It was found that the calcium-binding properties of the copolymer inhibited the enzymatic activities of digestive enzymes. This study was crucial in showing that the material was also capable of protecting the insulin inside by deactivating some of the digestive enzymes in the stomach.

Insulin: Oral Delivery Challenges

Insulin is a suitable example of all the efforts spent in finding a successful way for the development of an oral delivery system.^[24,44,47–51] An oral formulation of this protein, indeed, has been a challenging and hard-to-pin-down goal since the initial discovery in 1922. Since then, the protein has been administered by injections; yet,

there has always been a strong tendency to move toward more patient-friendly administrations. The emerging insulin delivery technologies are focusing on achieving similar efficacy to injectable insulin formulations, yet offering a higher degree of convenience and quality lifestyle to the patient. Although many alternative routes of insulin delivery are being developed or undergoing clinical trials, none of them have been commercialized yet owing to efficacy, safety, and cost considerations.

The human insulin gene is expressed in the β cells of the pancreas.^[52] The secretion and excretion of this hormone is triggered by glucose concentration to maintain normal glucose levels in the body. Insulin is synthesized from the β cells of the pancreas from a single chain precursor known as preproinsulin that is cleavage of the prepeptide sequence to generate proinsulin and eventually in the Golgi apparatus cut off of C-peptide and transformed into insulin. This protein comprises two peptide chains (A and B) connected to each other via disulfide bridges. In addition, the A chain contains another disulfide bridge within the chain.

Both the final form of the protein and the C-peptide are stored in secretory granules of the β cells and released after physiological stimulus.

As mentioned before, the trigger stimulus for insulin synthesis and secretion is represented by glucose. In detail, release of the hormone occurs in a biphasic fashion. When the glucose concentration outside the cell increases, the sugar molecules are transported inside the cell via an insulin-dependent glucose transporter, GLUT-2. This produces an immediate release of insulin stored in the secretory granules. If the physiological stimulus persists, a delayed response occurs and insulin is hence synthesized.

Insulin is an anabolic hormone indispensable for glucose transport across the membrane, formation of glycogen in the liver and in the skeletal muscles, and conversion of glucose into triglycerides.

Insulin of different animal species has similar biologic activity and differs only in the sequence of three amino acids in the A chain-source. Insulin is obtained from beef or pork pancreas. Human insulin is a product of recombinant DNA technology or chemical modification of pork insulin and is structurally identical to human insulin.

Insulin is the mainstay of treatment for diabetes mellitus, a group of metabolic disorders characterized by prejudiced glucose utilization, ensuing in hyperglycemia.^[53,54] Diabetes affects more than 150 million of people worldwide. It is in fact one of the most widespread and serious chronic diseases. Each year, approximately 800,000 people are diagnosed with diabetes, and its occurrence has been increasing steadily over the last half of this century.^[3] Incidence is expected to continue rising with the aging US

population and the increasing prevalence of obesity among Americans.

The disease is characterized by abnormalities in the metabolism of carbohydrates, proteins, and lipids as a consequence of deficiency in the synthesis, secretion, or function of insulin. All these anomalies lead to increased glucose levels, coupled with both macrovascular and microvascular complications during chronic treatments including coronary artery disease, peripheral vascular disease, nephropathy, neuropathy, and retinopathy. There are multiple causes for this disease. Each can be categorized under different forms of diabetes. A first gross classification distinguishes between primary or idiopathic diabetes mellitus, the most common, and secondary diabetes. The latter one includes forms of hyperglycemia associated with identifiable primary causes such as pancreatitis, pancreatectomy, and hormonal tumors that result in destruction of pancreatic islets.

Primary diabetes, on the other hand, has been further divided by the National Diabetes Data Group of the National Institute of Health into two variants: Type 1 and Type 2. Type 1 diabetes, formerly called insulin-dependent diabetes mellitus (IDDM) or juvenile-onset diabetes, is an autoimmune disease that results when the body's immune system attacks and destroys its own insulin-producing β cells in the pancreas. People with Type 1 diabetes need daily injections of insulin to survive. Type 1 diabetes develops most often in children and young adults and accounts for 5–10% of diagnosed diabetes in U.S.A. Type 2 diabetes, also called non-IDDM (NIDDM) or adult-onset diabetes, is characterized by non-responsiveness to insulin. It occurs when the β islet cells of the pancreas secrete too little insulin and/or the body cannot utilize insulin effectively. About 90% of people with diabetes have Type 2 disease. It is more common among people who are over 40 yr old, obese, or have a family history of diabetes.

Gestational diabetes develops or is discovered during pregnancy and usually disappears following pregnancy. Women who have had gestational diabetes are at greater risk for developing Type 2 diabetes later in their lives.^[3] Insulin is indispensable for Type 1 patients that do need to take exogenous insulin on a regular basis to control their blood glucose levels. Patients affected by Type 2 can delay the intake of the protein in the first stage of the disease, when they can control glucose levels with hypoglycemic agents, exercise, and diet. However, owing to progressive nature of the disease, they eventually require insulin administration to manage hyperglycemia owing to exhaustion of the pancreas. Insulin, therefore, is the indispensable treatment for diabetes mellitus.

Insulin was discovered in 1921 by Best and Banting at the University of Toronto, Canada. On January 23rd 1922, Leonard Thompson, only 13 yr old, was the

first patient saved by insulin injection. For more than a decade, after its discovery, unmodified ordinary insulin obtained from pancreatic extracts was the only preparation available.^[55] An attempt of administering the protein orally was performed. Nevertheless, no improvements in the diabetes symptoms were noticed. Since then, insulin has been administered by injections. At the present time, the available options to diabetic patients range between insulin syringes, insulin pens, insulin pumps, and jet injectors. Each of the available technologies appeals to a section of the insulin-dependent diabetics; however, the associated discomfort and inconvenience has raised the dissatisfaction among the users. The reduced compliance leads to poor glucose control and increase in diabetes-associated complications. This is a cause for concern because of the high incidence of diabetes worldwide. Recent research findings suggest that in the year 2000, over 20 million people were diabetics in U.S.A. As stated before, insulin syringes were introduced 75 yr ago and are still the most popular insulin delivery option.

Alternate Routes of Insulin Delivery

There are three main routes for insulin delivery being developed by researchers. The first is insulin via injection. It assures that insulin reaches the blood stream in an active form and in quick dosage patterns.

A few variations of the injection technique have been developed, most notably the insulin pump and insulin pens.^[56–58] Insulin pumps are mechanical pumps attached to a catheter. The flow of insulin into the body can be directly controlled by the pump speed and concentration of the insulin solution. They have been shown to accurately deliver insulin in its active form. The problem with this system is the size of the pump. Also, the catheter needs to be continually cleaned and sterilized when the pump is disconnected to avoid infection. The insulin pen is a small pen-shaped device with a small needle that can administer exact amounts of insulin into the body. While it is easier to use than standard needles, it still relies on injection for insulin delivery into the body.

The second route of insulin delivery system being developed is through transdermal systems.^[59] The use of ultrasound enhancement has led to the development of this system. Transdermal transport is usually limited to small weight compounds, but low-frequency ultrasound shows a several orders of magnitude increase in the insulin transport across the skin. Researchers continue to improve this process, but it is currently not efficient enough for clinical use.

The third process is transmucosal delivery of insulin into the blood stream. There are a number of mucosal

surfaces possible for transport: nasal, buccal, pulmonary, oral, vaginal, and rectal.^[58] Nasal, pulmonary, and oral are the leading candidates for transmucosal delivery.

Numerous researchers have developed delivery devices for insulin via the nasal route.^[19,60–62] A significant amount of insulin reaches systemic circulation through the nasal cavity. Researchers were able to monitor a response in the glucose levels of the body after dosage. Nonetheless the insulin bioavailability was too low. Further research into enhancement of insulin transport via the nasal passage has shown an increase in blood serum insulin when chemical enhancers were used over more traditional nasal administration.^[63]

The first method of oral insulin delivery was with the addition of proteolytic inhibitors such as diisopropyl-fluorophosphate^[64] or naturally occurring aprotinin.^[65] While some decrease in blood glucose levels was obtained, these reductions were far short of the decrease required.

Research on encapsulating insulin into liposomes^[66] showed large decreases in the blood glucose level, but the results were inconsistent.

Touitou and Rubinstein^[24] developed a polyacrylate coating designed to dissolve at the pH of the colon. The particles were able to decrease blood glucose concentrations for 3 to 4 hr in rats. This study was of historical importance as it was the first use of polyacrylate for oral protein delivery. However, as already mentioned, the bioavailability achievable with this system was still too low to be considered for a marketed product.

A wide range of biodegradable and conventional polymers have been investigated as possible oral insulin delivery devices.^[14,51] These systems must protect the insulin as it passes through the stomach, and were designed to release their insulin and successfully transport it across the GI lining. Most systems developed to this point either protected the insulin while in transit or aided in the transport of insulin across the cell layer of the upper small intestine, none have accomplished both. Not designing the system to deal with both challenges caused failure and resulted in little active insulin getting from the device into the blood stream.

Nagai and coworkers^[50] added a protease inhibitor to aid in protecting the insulin while in transit. They found an increase of nearly four times the insulin reaching the blood stream compared to the system without the protease inhibitor. The bioavailability of the insulin their system delivered was near zero. This study proved that adding proteolytic inhibitors alone would not protect the insulin. Thus, protecting the insulin from proteolytic degradation was not the only important concern of insulin degradation en route to the blood stream.

CONCLUSIONS

In conclusion from these studies it is clear that the oral (transmucosal) delivery of proteins has been a leading problem in drug delivery, yet one that has not had any significant solutions up to now. We believe that development of new oral protein delivery devices based on pH-sensitive or hydrogen bonding-complexing and pH-responsive gel carriers interspersed with bioactive protein is a promising solution. Insulin, calcitonin or other proteins can be loaded into polymeric microspheres or microparticles. In these systems, in the acidic environment of the stomach, the gels are unswollen due to the formation of polymer complexes.^[67-74] The protein remains in the gel and is protected from proteolytic degradation. In the basic and neutral environments of the intestine, the complexes dissociate resulting in rapid gel swelling and protein release^[75-80]. These system exhibit significant mucoadhesion in the upper small intestine and show inhibitory properties on the proteolytic activity of enzymes in the GI tract.

ARTICLES OF FURTHER INTEREST

Biomaterials: Protein-Surface Interactions; Controlled Release; Drug Delivery, Controlled

REFERENCES

1. Lee, V.H.L. Changing needs in drug delivery in the era of peptide and protein drugs. In *Peptide and Protein Drug Delivery*; Lee, V.H.L., Ed.; Marcel Dekker Inc.: New York, NY, 1991; 1-56.
2. Lee, V.H.L.; Dodd-Kashi, S.; Grass, G.M.; Rubas, W. Oral route of peptide and protein drug delivery. In *Protein and Peptide Drug Delivery*; Lee, V.H.L., Ed.; Marcel Dekker Inc.: New York, NY, 1991; 691-740.
3. Anonymous. Report of the expert committee on the diagnosis and classification of diabetes mellitus. *Diabetes Care* **1997**, *20*, 1183-1197.
4. Klier, J.; Peppas, N.A. Controlled release by using poly (methacrylic acid-g-ethylene glycol) hydrogels. *J. Control Release* **1991**, *16*, 203-214.
5. Peppas, N.A.; Keys, K.B.; Torres-Lugo, M.; Lowman, A.M. Poly(ethylene glycol)-containing hydrogels in drug delivery. *J. Control Release* **1999**, *62*, 81-87.
6. Schwendeman, S.P.; Cardamone, M.; Klibanov, A.; Langer, R.; Brandon, M.R. Stability of proteins and their delivery from biodegradable polymer microspheres. In *Microparticulate Systems for the Delivery of Proteins and Vaccines. Drugs and the Pharmaceutical Sciences*; Cohen, S., Bernstein, H., Eds.; , 1996; 77, 1-87.
7. Stryer, L. *Biochemistry*; W.H. Freeman and Company: New York, 1995.
8. Oswein, J.Q.; Shire, S.J. *Protein and Peptide Drug Delivery*; Lee, V.H.L., Ed.; Marcel Dekker Inc.: New York, NY.
9. Shirley, B.A. *Protein Stability and Folding: Theory and Practice*; Humana Press, 1995.
10. Volkin, D.B.; Middaugh, C.R. *Stability of Protein Pharmaceuticals. Part A: Chemical and Physical Pathways of Protein Degradation*; Ahern, T.J., Manning, M.C., Eds.; Plenum: New York, 1992.
11. Ansel, H.C. *Introduction to Pharmaceutical Dosage Forms, Chap. 3*; Lea & Febiger, 1991.
12. Jeandidier, N.; Boivin, S. Current status and prospects of parenteral insulin regimens, strategies and delivery systems for diabetes treatment. *Adv. Drug Del. Rev.* **1999**, *35*, 179-198.
13. Carino, G.P.; Mathiowitz, E. Oral insulin delivery. *Adv. Drug Del. Rev.* **1999**, *35*, 249-257.
14. Fix, J.A. Oral controlled release technology for peptides: status and future prospects. *Pharm. Res.* **1996**, *13*, 1760-1764.
15. Fasano, A.; Uzzau, U. Modulation of intestinal tight junctions by zonula occludens toxin permits enteral administration of insulin and other macromolecules in an animal model. *J. Clin. Invest.* **1997**, *99*, 1158-1164.
16. Dressman, J.B.; Berardi, R.R.; Dermentzoglou, L.C.; Russel, T.L.; Schmaltz, S.P.; Barnett, J.L.; Jarvenpaa, K.M. Upper gastrointestinal pH in young, health men and women. *Pharm. Res.* **1990**, *7*, 756-761.
17. Yamashida, S.; Saitoh, K.; Nakanishi, M.; Nadai, T.; Kimura, T. Characterization of enhanced intestinal permeability; electrophysiological study on the effects of diclofenac and ethylenediaminetetraacetic acid. *J. Pharm. Pharmacol.* **1985**, *37*, 512.
18. Geary, R.S.; Schlameus, H.W. Vancomycin and insulin used as models for oral delivery of peptides. *J. Control. Rel.* **1993**, *23*, 65-74.
19. Hinchcliffe, M.; Illum, L. Intranasal insulin delivery and therapy. *Adv. Drug. Del. Rev.* **1999**, *35*, 199-234.
20. Schilling, R.; Mitra, A.K. Intestinal mucosal transport of insulin. *Int. J. Pharm.* **1990**, *62*, 53-64.
21. Gibaldi, M.; Feldman, S. Mechanism of surfactant effects on drug absorption. *J. Pharm. Sci.* **1970**, *59*, 579-588.
22. Hochman, J.; Artusson, P. Mechanism of absorption enhancement and tight junction regulation. *J. Contr. Rel.* **1994**, *29*, 253-267.
23. Fasano, A. Novel approaches for oral delivery of macromolecules. *J. Pharm. Sci.* **1998**, *87*, 1351-1356.
24. Touitou, E.; Rubinstein, A. Targeted enteral delivery of insulin to rats. *Int. J. Pharm.* **1986**, *30*, 95-99.
25. Yamamoto, A.; Taniguchi, T.; Rikyu, K.; Tsuji, T.; Fujita, T.; Murakami, M. Effects of various protease inhibitors on the intestinal absorption and degradation of insulin in rats. *Pharm. Res.* **1994**, *11*, 1496-1500.
26. Johnson, O.L.; Tracy, M.A. Peptide and protein drug delivery. In *Encyclopedia of Controlled Release*; Mathiowitz, E., Ed.; Wiley: New York, NY, 1999; Vol. II, 816-833.
27. Lowe, P.J.; Temple, C.S. Calcitonin and insulin in isobutyl cyanoacrylate nanocapsules: protection against



- proteases and effect on intestinal absorption in rats. *J. Pharm. Pharmacol.* **1994**, *46*, 547–552.
28. Brannon-Peppas, L. Polymers in controlled drug delivery. *Medical Plastics and Biomaterials* **1997**, *4*, 34–44.
 29. Sahlin, J.J.; Peppas, N.A. Hydrogels as muco- and bioadhesive materials: a review. *Biomaterials* **1996**, *17*, 1553–1561.
 30. Peppas, N.A. Hydrogels and drug delivery. *Curr. Opin. Coll. Inter. Sci.* **1997**, *2*, 531–537.
 31. Kim, S.W.; Bae, Y.H.; Okano, T. Hydrogels: swelling, drug loading, and release. *Pharm. Res.* **1992**, *9*, 283–290.
 32. Lowman, A.M.; Peppas, N.A. Hydrogels. In *Encyclopedia of Controlled Release*; Mathiowitz, E., Ed.; Wiley: New York, NY; Vol. 1, 397–418.
 33. Bell, C.L.; Peppas, N.A. Biomedical membranes from hydrogels and interpolymer networks. *Adv. Polym. Sci.* **1995**, *22*, 126–175.
 34. Khare, A.R.; Peppas, N.A. Swelling/deswelling of anionic copolymer gels. *Biomaterials* **1995**, *16*, 559–567.
 35. Peppas, N.A. *Hydrogels in Medicine and Pharmacy, Vol. 1 Fundamentals*; Peppas, N.A., Ed.; CRC Press: Boca Raton, FL, 1986.
 36. Peppas, N.A.; Bures, P. Molecular dynamics of pH-sensitive hydrogels based on poly(acrylic acid). *Polym. Prepr.* **1999**, *40* (2), 506–507.
 37. Peppas, N.A.; Klier, J. Controlled release by using poly(methacrylic acid-g-ethylene glycol) hydrogels. *J. Control Release* **1991**, *16*, 203–214.
 38. Bell, C.L.; Peppas, N.A. Water, solute and protein diffusion in physiologically-responsive hydrogels of poly(methacrylic acid-g-ethylene glycol). *Biomaterials* **1996**, *17*, 1203–1218.
 39. Lowman, A.M.; Peppas, N.A. Solute transport analysis in pH-responsive, complexing hydrogels of poly(methacrylic acid-g-ethylene glycol). *J. Biomat. Sci. Polym. Edn.* **1999**, *10*, 999–1009.
 40. Peppas, N.A.; Lowman, A.M. Pulsatile drug delivery based on a complexation-decomplexation mechanism. In *Intelligent Materials for Controlled Release*; Dinh, S.M., DeNuzzio, J.D., Comfort, A.R., Eds.; ACS Symposium Series: Washington, DC, 1999; Vol. 728, 30–42.
 41. Bell, C.L.; Peppas, N.A. Equilibrium and dynamic swelling of polyacrylates. *Polym. Eng. Sci.* **1996**, *36*, 1856–1861.
 42. Drummond, R.K.; Klier, J.; Alameda, J.A.; Peppas, N.A. Preparation of poly(methacrylic acid-g-ethylene oxide) microspheres. *Macromolecules* **1989**, *22*, 3816–3818.
 43. Nakamura, K.; Maitani, Y.; Lowman, A.M.; Takayama, K.; Peppas, N.A.; Nagai, T. Uptake and release of budesonide from mucoadhesive, pH-sensitive copolymers and their application to nasal delivery. *J. Control Release* **1999**, *61*, 329–335.
 44. Lowman, A.M.; Morishita, M.; Kajita, M.; Nagai, T.; Peppas, N.A. Oral delivery of insulin using pH-responsive complexation gels. *J. Pharm. Sci.* **1999**, *88*, 933–937.
 45. Lowman, A.M.; Peppas, N.A.; Morishita, M.; Nagai, T. Novel bioadhesive complexation network for oral protein drug delivery. In *Materials for Controlled Release Applications*; McCullough, I., Shalby, S., Eds.; ACS Symposium Series: Washington, DC, 1999.
 46. Torres-Lugo, M.; Peppas, N.A. Molecular design and in vitro studies of novel pH sensitive hydrogels for oral delivery of calcitonin. *Macromolecules* **1999**, *32*, 6646–6651.
 47. Constantino, H.R.; Liauw, S.; Mitragotri, S.; Langer, R.; Klivanov, A.M.; Sluzky, V. The pharmaceutical development of insulin: historical perspectives and future directions. In *Therapeutic Protein and Peptide Formulation and Delivery*; Shahrokh, Z., Sluzky, V., Cleland, J.L., Shire, S.J., Randolph, T.W., Eds.; American Chemical Society: Washington, DC, 1997; 29–66.
 48. Saffran, M.; Sudesh Kumar, G.; Savariar, C.; Burnhan, J.C.; Willims, F.; Neckers, D.C. A new approach to the oral administration of insulin and other peptide drugs. *Science* **1986**, *233*, 1081–1084.
 49. Saffran, M. Oral colon-specific drug delivery with emphasis on insulin. In *Oral Colon-Specific Drug Delivery*; Friend, D.R., Ed.; CRC Press: Boca Raton, FL, 1992; 115–142.
 50. Morishita, I.; Morishita, M.; Takayama, K.; Machida, Y.; Nagai, T. Hypoglycemic effect of novel oral microspheres of insulin with protease inhibitor in normal and diabetic rats. *Int. J. Pharm.* **1992**, *78*, 9–16.
 51. Carino, G.P.; Jacob, J.S.; Mathiowitz, E. Nanospheres based oral insulin delivery. *J. Control Release* **2000**, *65*, 261–269.
 52. Ashcroft, F.; Ashcroft, S. *Insulin: Molecular Biology to Pathology*; Oxford University Press: New York, 1992.
 53. Goodman Gilman, A. Insulin, oral hypoglycemic agents, and the pharmacology of the endocrine pancreas. In *The Pharmacological Basis of Therapeutics*; McGraw-Hill, Ed.; 2001.
 54. *Harrison's Principles of Internal Medicine, Genetics of Diabetes Mellitus*; McGraw-Hill Ed., 2001, <http://books.mcgraw-hill.com/medical/harrisons/index.html>.
 55. Trehan, A.; Ali, A. Recent approaches in insulin delivery. *Drug Dev. Ind. Pharm.* **1998**, *26*, 589–597.
 56. Rohde, T.D.; Buchwald, H. Implantable pumps: recent progress and anticipated future advances. *ASAIO Journal*, **1992**, *38*, 772–778.
 57. Charles, M.A.; Selam, J.L. Devices for insulin administration. *Diabetes Care* **1990**, *13*, 955–979.
 58. Ziv, E.; Hoffman, A. Pharmacokinetic consideration of new insulin formulations and routes of administrations. *Clin. Pharmacokinet.* **1997**, *33*, 285–301.
 59. Mitragotri, S.; Blankshtein, D.; Langer, R. Ultrasound-mediated transdermal protein delivery. *Science* **1995**, *269*, 850–853.
 60. Nagai, T.; Nishimoto, Y.; Nambu, N.; Suzuki, Y.; Sekine, K. Powder dosage form of insulin for nasal administration. *J. Control Release* **1984**, *1*, 15–22.
 61. Longenecker, J.P. Nazlin[®]-transnasal systemic delivery of insulin. In *Delivery Systems for Peptide Drugs*; Davis, S.S., Illum, L., Tomlinson, E., Eds.; Plenum: New York, 1986; 211–220.
 62. Moses, A.C.; Gordon, G.S.; Carey, M.C.; Flier, J.S. Insulin administered intranasally as an insulin-bile

- salt aerosol-effectiveness and reproducibility in normal and diabetic subjects. *Diabetes* **1983**, *32*, 1040–1047.
63. Fernandez_Urrusuno, R.; Calvo, P.; Reuman-Lopez, C.; Vila-Jato, J.L.; Alonso, M.J. Enhancement of nasal absorption of insulin using chitosan nanoparticles. *Pharm. Res.* **1993**, *16*, 1576–1581.
64. Danforth, E.; Ur, R.; Moore, O. Intestinal absorption of insulin in the rat. *Endocrinology* **1959**, *65*, 118.
65. Laskowski, M.; Haessier, H.A.; Mieon, R.R.; Peanasky, R.J.; Laskowski, M. Effect of trypsin inhibitor on passage of insulin across the intestinal barrier. *Science* **1958**, *127*, 1115.
66. Patel, H.M.; Ryman, B.E. Oral administration of insulin by encapsulation within liposomes. *FEBS Lett.* **1976**, *62*, 60–74.
67. Mehnert, W.; Mader, K. Solid lipid nanospheres. Production, characterization and applications. *Adv. Drug Del. Rev.* **2001**, *47*, 165–196.
68. De Jaeghere, F.; Doelker, E.; Gurny, R. Nanospheres. In *Encyclopedia of Controlled Drug Delivery*; Mathiowitz, E., Ed.; John Wiley & Sons, Inc.: New York, 1999; Vol. 2, 641–664.
69. Mittal, K.L.; Kumar, T.M. *Handbook of Microemulsion Science and Technology*; Marcel Dekker Eds.: New York, NY, 1999.
70. Wisendanger, R. *Scanning Probe Microscopy and Spectroscopy. Methods and Applications*; Cambridge University Press: Cambridge, 1994.
71. Sudol, E.D. Dispersion polymerization. In *Polymeric Dispersions: Principles and Applications*; Asua, J.M., Ed.; Kluwer Academic Publishers: Boston, 1997; 141–154.
72. Finsy, R. Particle sizing by quasi-elastic light scattering. *Adv. Coll. Inter. Sci.* **1999**, *52*, 79–143.
73. Flory, P.J. *Principles of Polymer Chemistry*; Cornell Univ. Press: Ithaca, New York, 1953.
74. Suzuki, M. *Gels Handbook, The Fundamentals*; Osada, Y., Kajiwara, K., Eds.; Academic Press: New York, 2001; 228–234.
75. Odian, G. *Principle of Polymerization*; John Wiley & Sons, Inc.: New York, 1990.
76. Arshady, R. Suspension, emulsion and dispersion polymerization: a methodological survey. *Colloid Polym. Sci.* **1990**.
77. Goldstein, J.I.; Newbury, D.E.; Echlin, P.; Joy, D.C.; Romig, A.D.; Lyman, C.E.; Fiori, C.; Lifshin, E. *Scanning Electron Microscopy and X-Ray Microanalysis*; Plenum Press, 1992.
78. Kreuter, J. Nanospheres. In *Colloidal Drug Delivery Systems*; Kreuter, J., Ed.; Marcel Dekker: New York, NY, 1994; 219–342.
79. Weiner, B.B. *Modern Methods of Particle Size Analysis*; Barth, H.G., Ed.; John Wiley & Sons: New York, 1984; 93–116.
80. Madsen, F.; Peppas, N.A. Complexation graft copolymer networks: swelling properties, calcium binding, and proteolytic enzyme inhibition. *Biomaterials* **1999**, *20*, 1701–1708.

Pyrolytic Carbon

Robert B. More

RBMore Associates, Austin, Texas, U.S.A.

George Sines

Dept. of Materials Science and Engineering, School of Engineering, University of California, Los Angeles, California, U.S.A.

Ling Ma

International Rectifier Corporation, El Segundo, California, U.S.A.

Jack C. Bokros

Medical Carbon Research Institute, Austin, Texas, U.S.A.

INTRODUCTION

Pyrolytic carbon (PyC), a material for long-term blood-contacting implants, has enjoyed great success. Over the past three decades, the use of PyC heart valve implants has become commonplace. The total number of components used in heart valve replacements alone is estimated to exceed 4 million with a total clinical experience of about 20 million patient years.

More recently, the successful use of PyC in small joint replacements over the past 20 years has led to the approval of such devices by the U.S. Food and Drug Administration. This event, in addition to recent improvements in PyC technology, has placed a critical focus on PyC as a candidate material for other implant applications. Although materials-related issues involving strength and durability have limited possibilities for many devices such as spinal prostheses and other joint replacements, the increased strength and toughness of recently developed PyC materials expands the horizon for new applications. The superior wear and fatigue properties of PyC and its stiffness match with bone further entice its increased use in load-bearing orthopedic implants.

PyC is formed by the high-temperature pyrolysis, or thermal decomposition, of source hydrocarbons and the subsequent recrystallization of elemental carbon. PyC is characterized by high temperature stability, high mechanical strength, and high chemical resistance. Furthermore, PyC exhibits excellent stability, wear resistance, fatigue resistance, and compatibility in the highly aggressive biological environment. Alloying elements such as silicon or boron may be added to the source hydrocarbon to produce carbide inclusions within the carbon matrix. Unlike other pyrolytic materials that tend to be highly anisotropic, the biomedical-grade PyC materials are deliberately prepared as isotropic materials.

Critical information needed for implant design using PyC is given in this text including physicochemical and

mechanical properties, manufacturing methods, and microstructural details. The intent is to provide the background, tools, and references for prospective designers and to spark their imaginations for new applications.

HISTORY

In 1964, abstract number 84—"The Anticlot Properties of Graphite Coatings on Artificial Heart Valves"—appeared in the journal *Carbon*, Volume 1, No. 3, April 1964. This abstract by Vincent L. Gott, James D. Whifler, Robert C. Dutton, Donald E. Koepke, Ronald L. Daggett, and William P. Young, all of the Department of Surgery at the University of Wisconsin, had been presented at the Sixth American Carbon Conference in Pittsburgh, PA, in June 1963. At the time, 95 percent of all heart valve implants in animals failed within two weeks following implant because of blood clotting (thrombosis). Their approach to the problem was to evaluate various plastics and coatings as candidates for use in artificial hearts and artificial valves in the hopes of finding a satisfactory clot-repelling material. Of all the materials tested, a colloidal graphite coating gave overwhelmingly better results than any of the others tested.

Another researcher, Jack C. Bokros at General Atomic Company in San Diego, happened to read this abstract. Bokros, a metallurgist by training, was working with PyC as coatings for nuclear fuel particles used in gas-cooled fission reactors. He was intrigued by Gott's work and contacted him to suggest trying PyC in the anticlot testing. This was the beginning of a collaboration that was to revolutionize heart valve technology and ultimately make the prospect of clinical heart valve implantation the commonplace event that it is today.

Development and Utilization

In the 1960s the study of biomaterials was in its infancy, and the only approach to solving materials problems for implants was by trial and error. There were designs for prosthetic heart valves and methods for implantation, but the effective lifetime of the implants was severely limited. Blood clots would form on the implants that interfered with function, or the materials would fail due to fatigue, by distortion due to sterilization, or by degradation due to exposure to the body fluids. Heart valve replacement lifetime was much shorter than the patient's expected lifetime.

Gott tested the PyC materials provided by Bokros and found excellent thromboresistance,^[1] prompting the collaborators to set out to optimize the material for heart valve applications. Their initial technical problem was to increase the PyC coating thickness from the micrometer thicknesses that Bokros had been using at General Atomic Company to levels on the order of 0.5 mm needed for the structural requirements for heart valves. To attain the thicker coatings, the coating reaction needed to be prolonged.

Pyrolytic carbons are prepared by heating a hydrocarbon such as propane to approximately 1200 to 1400°C in the absence of oxygen. At high temperature, the hydrocarbon breaks down to free radicals, which then "polymerize" into a solid. In order to produce isotropic PyC, a vertical fluidized bed reactor is needed.^[2] The reactor consists of a vertical tube containing a fine dispersion of refractory material particles, which are levitated using an inert carrier gas, such as nitrogen or helium, fed into the bottom of the tube. Induction coils on the outside of the tube heat the contents. At temperature, a hydrocarbon gas is added to the mixture and the pyrolysis reaction takes place. See Fig. 1.

This process can be described as a chemical vapor deposition because all of the particles within the bed become coated with PyC. Specific components are manufactured by placing preformed substrates within the fluidized bed that subsequently become coated with PyC during the reaction. The resulting components are composites consisting of an outer shell of PyC encasing and bonded to an inner refractory substrate such as fine-grained graphite. See Fig. 2.

The reactor depicted in Fig. 1 is a simple fluidized bed reactor. With prolonged coating reaction time, the bed surface area grows as additional carbon is deposited until fluidization ceases and the bed collapses. Bed collapse limited the coating thickness that could be applied in the simple bed reactor to micrometer levels. It was found that the coating reaction could be prolonged to give the needed thickness by simultaneously adding new bed particles and by withdrawing coated bed particles to maintain the reaction in a quasi steady

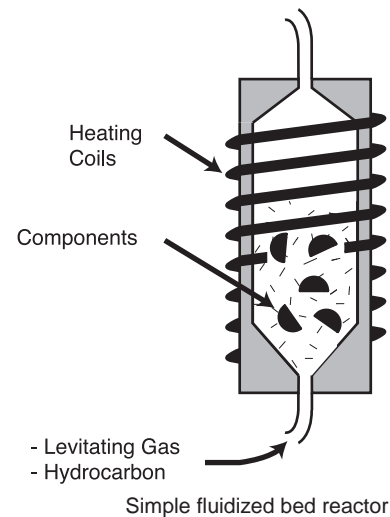


Fig. 1 A schematic diagram of a simple fluidized bed reactor. Components are levitated in the fluidized bed and coated with PyC as the source hydrocarbon is pyrolyzed. The reactor is heated using induction coils.

state.^[3] Because it was not possible to dynamically measure the bed surface area, the addition of new and removal coated bed particles was based on best guesses; consequently, there could be considerable variation in the properties of the resulting coatings. A quasi-steady-state reactor is depicted in Fig. 3.

The second technical problem was to ensure consistent production of pyrolytic carbon with the desired hardness and strength. Coating properties are dictated by several variables: bed temperature, gas composition, total gas flow rate, and bed surface area. See Fig. 3. With the reaction controls available in the 1960s, it was possible to control all of the variables except for the surface area of the fluidized bed of particles. Bed area

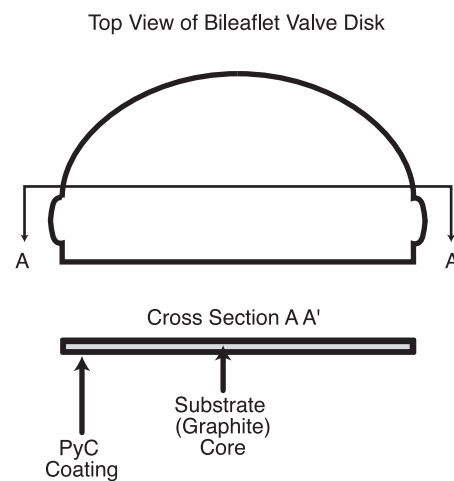


Fig. 2 Composite structure of PyC components. The PyC coating encases and is bonded to an inner substrate such as fine-grained graphite.

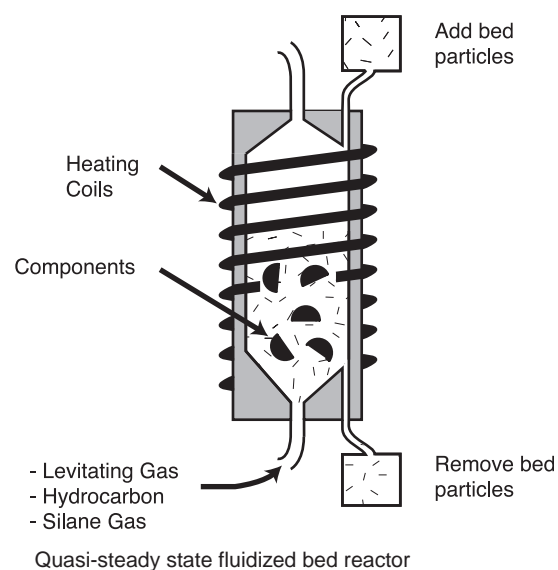


Fig. 3 A quasi-steady-state fluidized bed reactor. Bed particles are added and withdrawn in a semiempirical fashion to prolong the coating reaction. Silane gas is added to produce silicon carbide inclusions in the PyC coating.

was critical because it determined the resulting coating microstructure, strength, and hardness. Lack of bed area control was compensated for by the addition of methyl trichlorosilane gas. The silane gas decomposed during the pyrolysis reaction and formed fine silicon carbide inclusions dispersed within the pyrolytic carbon matrix. Silicon carbide inclusions increased coating hardness as needed for the consistent manufacture of heart valve structural components.^[3]

From Gott's tests, pure carbon displayed the best biocompatibility, whereas pure silicon carbide was thrombogenic, but silicon carbide in small amounts codeposited with PyC was found to be tolerated.^[1] A practical compromise between thrombogenicity and consistent mechanical properties was reached with a 4- to 12-weight-percent silicon-pyrolytic carbon alloy. This material was patented and given the trade name Pyrolite carbon.

Late-1960s state-of-the-art valve designs consisted of three major variations of a metallic cage with a single polymeric occluder, e.g., either a ball or disk poppet or a tilting disk^[4] as shown in Fig. 4. Typical heart valve occluder geometries consisted of spherical balls, lenticular poppets, or circular disks. Occluders were mounted in ringlike metallic cages or valve orifices, with struts that captured and guided the occluder during the opening and closing motions. Metals used for orifices were principally titanium or cobalt chrome alloys.

The use of Pyrolite carbon increased rapidly in existing mechanical heart valve designs as a replacement for the polymeric occluders beginning with the ball-in-cage DeBakey Surgitool design first implanted

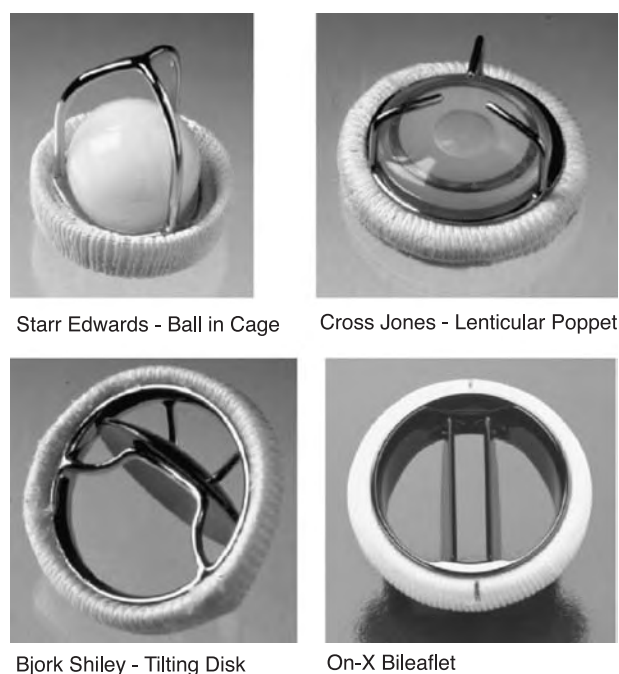


Fig. 4 Different heart valve design types: (top left) Starr-Edwards silicone rubber ball in metallic cage (1964); (top right) Cross Jones polymeric lenticular poppet in metallic cage (1967); (bottom left) Bjork Shiley tilting disk, originally a polyacetyl disk in a metallic cage (1969), shown here with a PyC disk (1975); (bottom right) On-X prosthetic heart valve, all-carbon bileaflet valve (1996).

clinically in 1968. The superior durability, stability, and biocompatibility of Pyrolite carbon eliminated the causes of early valve failure, thus allowing valves to endure for the patient's lifetime.^[5] Utilization of PyC as a material for mechanical heart valves was declared an "exceptional event."^[6]

In order to meet the demand for heart valve components, General Atomic Company formed the Medical Products Division, which in 1979 was sold to Intermedics, Inc. The new entity, renamed Carbomedics, was the sole source of Pyrolite carbon in the United States and a number of other countries worldwide until the original patent expired in the early 1990s.

The General Atomic/Carbomedics Pyrolite component manufacturing process consisted of several steps:

1. Fine-grained graphite substrate preforms were machined and included in the fluidized bed reactor to be coated with carbon. The substrates had the desired component shape but were under-sized to accommodate the added coating thickness.
2. The substrates were coated with silicon-alloyed Pyrolite carbon in the reactor to an excess coating thickness to compensate for the poor control of coating thickness.

3. Excess coating was removed by abrasive grinding.
4. The components were polished to a fine surface finish.

Because of the problems of wear and durability experienced with polymeric valve components, valves were designed with freely rotating circular disks to distribute the wear circumferentially rather than being concentrated at fixed contact points. Evaluations of PyC for use in the replacement of polymeric disks revealed that although the wear rate of PyC caused by metal was much less than that observed for polymers, it was still necessary to use the rotating circular disk concepts for the existing Lillehie–Kaster and Bjork–Shiley designs.

It was recognized in the 1970s that ball-in-cage and poppet designs produced lateral flow and tilting disk designs produced skewed asymmetric flow.^[7] In order to improve the flow, the obstructive designs that used poppets or tilting disk were rejected, and conceptualization was redirected toward bileaflet designs that centralized the flow. But bileaflet designs required fixed pivots and hence resulted in localized wear. Extensive testing using a variety of materials in contact with PyC indicated that PyC wearing against itself produced low wear rates,^[5] whereas metals wearing on PyC caused the PyC to wear much more rapidly. Consequently, efforts were directed to adapt the early Lillehie–Kalke bileaflet design, previously made using titanium, to an all-PyC design. The culmination of this effort resulted in the Saint Jude Medical (SJM) valve. The bileaflet SJM valve was implanted in 1976 and became the de facto standard mechanical valve during the 1980s.

During the 1980s several other mechanical valves that used PyC were developed. Sorin Biomedica in Italy produced a monoleaflet valve^[8] and a bileaflet valve^[9] using their own PyC (the Pyrolite carbon patents held by Carbomedics were not enforceable in Italy). The Sorin bileaflet valve was not all carbon; it had PyC leaflets and a titanium orifice. To improve blood compatibility, the Sorin titanium orifice was coated with a very thin layer of vapor-deposited carbon.^[10]

In the United States, two other bileaflet designs, the Duromedics and the Carbomedics valves, were introduced. These valves were designed using cylindrical orifices without upstream protrusions (like the SJM pivot guards) that included the pivot recesses. The protrusions were required in the SJM design because there was, at the time, no way to machine the pivot recesses on the inside of a straight cylinder and then grind them to final size after coating. Accordingly, a new manufacturing process was devised in which the configuration required to be on the inside surface of an orifice was machined as a negative on the outside of a graphite mandrel. After the PyC was deposited on the

mandrel, the mandrel was removed^[11] leaving a free-standing, all-PyC orifice. The shortcoming of the mandrel process was that during the initial moments of coating when the PyC layer is very thin (and in tension), the impacts of the mandrels with one another during the levitation causes impact defects that are exposed on the inside surface of the orifice once the mandrel is removed. These defects potentially reduce the strength and surface quality of orifices produced on a mandrel.

In the early 1990s the original General Atomic Company patents for Pyrolite carbon expired. Heart valve manufacturers that had relied upon Carbomedics as the sole source for Pyrolite valve components developed their own facilities to produce Pyrolite-like carbon. The FDA responded with guidance for demonstrating equivalence given in Appendix B of the Replacement Heart Valve Guidance—Draft document, Oct. 14, 1994. This document may be downloaded from the FDA Web site: <http://www.fda.gov/cdrh/ode/3751.html>.

During the early 1990s, Bokros and members of his original team from General Atomic left Carbomedics and formed a new company, Carbon Implants Incorporated (CII). At CII they reviewed the PyC reaction process, applying 1990s-level technological advances in transducers and computerized process control and were able to devise a method for dynamically measuring bed surface area. The bed surface area was the poorly controlled variable in the original General Atomic process that required the addition of silicon to maintain hardness and prevented the possibility of coating a substrate precisely to size, thus eliminating the final grinding operations.

With the new process controls, they attained a true steady-state reactor (Fig. 5), which made it possible to produce pure pyrolytic carbon. The resulting pure carbon PyC was stronger, tougher, and more deformable than the silicon-alloyed PyC. The improved PyC was patented and named On-X carbon.^[12]

In 1994, a new company was formed, Medical Carbon Research Institute, to design and produce a new bileaflet valve that took full advantage of the breakthrough in process technology. With the exception of Medical Carbon Research Institute's On-X heart valve, the original 1960s technology developed at General Atomic continues to be used in the manufacture of prosthetic heart valves, using the silicon-alloyed Pyrolite or Pyrolite-like carbon.

The new technology that allowed precise control of the deposition process provided a number of technological advances in the design of heart valve replacement.

First, the precise control allowed the consistent and reproducible production of a pure carbon that was

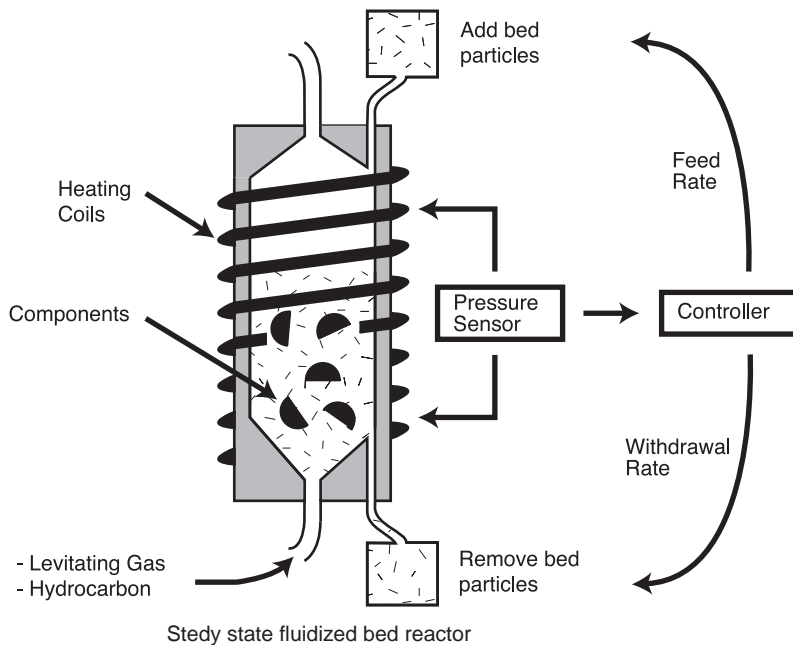


Fig. 5 True steady-state fluidized bed reactor. The coating reaction rate and bed surface area are analytically controlled by addition and withdrawal of bed particles. Such controls improved the quality of the PyC, which eliminated the need for the silicon carbide additive.

stronger and tougher than the silicon-containing material and with the hardness required to attain the wear resistance needed for use in heart valve replacements.

Second, the precision of the process allowed for the coating of pivot configurations inside of an orifice to produce finished dimensions that required no final grinding or otherwise shaping. This coat-to-size ability resulted in a surface quality not attainable with the older technology.

Third, the elimination of the silicon carbide addition to the PyC improved the toughness so that the design of heart valve replacement orifices was not restricted to straight cylinders (which are hemodynamically deficient) and allowed for the realization of orifices that are shaped more like the natural ones with a flared inlet.

Orthopedic Applications

The potential for PyC in orthopedic applications was recognized early on because of PyC's advantages over metallic alloys and polymers.^[13,14]

1. A modulus of elasticity similar to bone
2. Excellent wear characteristics
3. Excellent fatigue endurance
4. Excellent biocompatibility with bone and hard tissue
5. Fixation by direct bone apposition

Studies by Cook et al.^[15] of hemijoint implants of a PyC femoral head in the canine hip observed a greater

potential for acetabular cartilage survival in PyC than for cobalt-chromium-molybdenum alloy and titanium alloy femoral heads. PyC demonstrated significantly lower levels of gross acetabular wear, fibrillation, eburnation, glycosaminoglycan loss, and subchondral bone change than the metallic alloys.

In-vitro and clinical in-vivo PyC orthopedic implant studies during the 1970s through the early 1990s were summarized by Tian et al.,^[14] with the conclusion that PyC demonstrates both good biocompatibility and good function in clinical applications.

Between 1979 and 1987, 151 PyC metacarpophalangeal (MCP) finger joint replacements were implanted in 53 patients at the Mayo Clinic, Rochester, Minnesota.^[16] Follow-up after 10 years demonstrated excellent performance. These results were used in part to justify the FDA premarket approval (PMA) for the Ascension Orthopedics semiconstrained, uncemented MCP finger joint replacement, PMA P000057, Nov. 2001. The Ascension PMA summary of safety and effectiveness may be downloaded from the FDA Web site: <http://www.fda.gov/cdrh/pdf/p000057b.pdf>.

Although the potential for PyC in orthopedic applications is excellent, it can be more expensive than metals or polymers, so an economic prejudice exists against its use when less expensive materials were available and performed adequately at least in the near term. However, the recent demonstrations of long-term success with PyC MCP joint replacements, in combination with persistent materials-related lifetime limitations in metallic and polymeric devices, counter this economic prejudice. The cost of redo surgery to replace failed devices well exceeds the initial cost of

materials. New devices are currently being considered in which metals or polymers may be inappropriate.

PYC DESIGN CONSIDERATIONS

Materials selection for an implant is a complex issue that involves biocompatibility together with mechanical and physical compatibility. Data that a designer needs in order to consider PyC for an implant are summarized as follows.

Biocompatibility

Preclinical biocompatibility tests currently required by the FDA and European Community (International Standards Organization ISO 10993) for rigid material implants are given in Table 1 along with results for PyC. No significant reactions were observed.^[12] These tests do not detect a difference between silicon-alloyed and pure carbon PyC.

For long-term critical implants, a preclinical animal implant evaluation of a device is required. The animal tests model the implants and materials in the intended application as closely as possible. Examples may be found in FDA PMA Summaries of Safety and Effectiveness (SSE). See the Ascension MCP SSE, <http://www.fda.gov/cdrh/pdf/p000057b.pdf>, and the On-X valve SSE, <http://www.fda.gov/cdrh/pdf/p000037.html>, for examples. The FDA Web site has a searchable database that provides guidance documents and SSEs for a variety of devices: <http://www.fda.gov/cdrh/devadvice/pma/index.html>.

For any device, it is highly recommended to obtain regulatory agency guidance documents that may exist as a means of determining design testing and performance requirements. The International Standards Organization (ISO) also publishes guidance for devices. In many cases, a guidance document may not yet exist, but a review of one for a related device provides a good idea of the information needed. An excellent

review of theory and observations of PyC biocompatibility is given in *Cells and Materials* 1996;6(4):271–356, which is a special issue on “Pyrolytic Carbon: Biomedical Uses and Biocompatibility.”

Physical and Mechanical Properties

Carbon has an allotropic crystallographic structure and occurs as a spectrum of ordered or disordered (turbostratic) structures varying from tetrahedral covalent bonded diamond to layered hexagonal covalent in-plane and Van der Waals interplane bonded graphite. Each carbon allotrope has unique structure-dependent physical and mechanical properties. Carbon, rather than being a single material, is actually an entire family of materials: coal, charcoal, fullerenes, diamond, glassy carbon, etc. Thus, PyC biocompatible and mechanical properties that support implant applications are unique and cannot be assumed to apply for another carbon form just because it happens to be of the carbon family.

Pyrolytic carbon is a hard, rigid linear-elastic material and exhibits no appreciable plastic deformation. As most brittle materials, PyC has scatter in the value of its tensile strength. Strength scales with effective stressed volume or effective stressed area according to the Weibull relationship:

$$\sigma_2/\sigma_1 = (K_1 V_1/K_2 V_2)^{1/m}$$

where σ and KV are the strength and effective stressed volume of specimens 1 and 2 and m is the Weibull modulus of approximately 10 for PyC.^[17] Strain to failure also scales according to the Weibull relation. Elastic properties given in Table 1 are from four-point-bend, third-point-loading substrated specimens of dimensions 25 mm by 6 mm by 0.4 mm and effective stressed volume of 1.93 mm³. Thus, strength is determined by the specific device component design. For a new device, static and dynamic strength testing is required by regulatory agencies.

Table 1 Biological testing of pure PyC

Test description	Protocol	Results
Klingman maximization	ISO/CD 10993-10	Grade 1; not significant
Rabbit pyrogen	ISO/ DIS 10993-11	Nonpyrogenic
Intracutaneous injection	ISO 10993-10	Negligible irritant
Systemic injection	ANSI/AAMI/ISO 10993-11	Negative—same as controls
Salmonella typhimurium reverse mutation assay	ISO 10993-3	Nonmutagenic
Physicochemical	USP XXIII, 1995	Exceeds standards
Hemolysis—rabbit blood	ISO 10993-4/NIH 77-1294	Nonhemolytic
Elution test (L929 mammalian cell culture)	ISO 10993-5, USP XXIII, 1995	Noncytotoxic

The value of Young's modulus approximates that of cortical bone (~ 17 GPa). The similarity in modulus allows a much greater degree of stiffness matching between implants and bone. Interfacial stress concentrations that can contribute to implant loosening and bone reabsorption are minimized.^[13,14] Strain-to-failure values of 1.58% for the pure PyC are large relative to the strain-to-failure values of approximately 0.2% typical of ceramics.

Hardness is determined using a diamond penetrant indenter with a 500 g load. An interesting property of PyC is that it fully recovers the indentation, so hardness readings are made using a "record" of the indentation in a formvar coating over the test part surface. Density is low; therefore, valve moving components approach neutral buoyancy so component mass does not interfere with actuation by blood flow.

Fracture mechanics values cited were determined using miniature compact tension specimens and will be discussed subsequently in more detail. The graphite substrates were removed for the worst-case fracture mechanics value estimates using monolithic PyC. The friction coefficient for PyC bearing against PyC in blood plasma is low; the value of 0.15 given in Table 2

is a conservative overestimate. Wear resistance is given as the wear volume produced per linear distance using a rotating disk (PyC) pressed under load into a stationary PyC flat. However, it is important to note that wear and endurance are device design-specific tests. For example, heart valve designs are tested to 600 million cycles—full opening and closing under worst-case physiological flow conditions. Other devices must be tested in a manner simulating conservative end-use conditions. The Ascension MCP, for example, was cycled between 8 and 80 pounds load for 10 million cycles in a joint simulator.

Contemporary Long-Term Implant Design Criteria

The metallic strut failures in the convexo-concave Björk–Shiley heart valve during the late 1980s (see <http://www.me.utexas.edu/~uer/heartvalves/summary.html>) led to more stringent design evaluation criteria being required for long-term implants. In particular fracture mechanics, damage-tolerant design methods were strongly recommended in addition to the traditional stress-life (S–N) techniques.^[18] The first example of a damage-tolerant analysis for a heart valve was given in 1986 by Ritchie and Lubbock,^[19] but for metallic, not PyC, components. The late-1980s failure of PyC components in the Baxter–Edwards Duro-medics valve^[20] drove the development of methodology for measuring fracture mechanics properties of PyC.^[21] More recently, circa 2000, the FDA and ISO 5840 standards are requiring risk analyses also, but these are of a broader scope than damage-tolerant analyses.

Damage-tolerant life analyses for PyC long-term implant components have some unique aspects that merit discussion. Because PyC heart valve components and PyC layer thicknesses tend to be small, typical nondestructive evaluation techniques lack sufficient resolution and thus are not particularly helpful. Rather, a proof test approach has been taken. In this approach, every component is proof tested. Components survive the proof test provided that all flaws in the material are beneath the critical value (as defined by the fracture toughness K_{Ic}). Subsequently, flaws of the critical value are assumed to exist in the proof-tested components and to propagate under the worst-case service loads using the fatigue crack velocity, and the time-to-failure or lifetime is calculated. The existence of a fatigue threshold is disallowed, which requires that very small cracks behave in the same way as large cracks. This distinction will be discussed shortly, and it will be demonstrated that small cracks indeed behave as large cracks. Finally, an example of a damage-tolerant analysis will be given.

Table 2 Mechanical properties of biomedical carbons

Property	Pure PyC	Typical Si-Alloyed PyC
Flexural strength (MPa)	493.7 ± 12	407.7 ± 14.1
Young's modulus (GPa)	29.4 ± 0.4	30.5 ± 0.65
Strain-to-failure (%)	1.58 ± 0.03	1.28 ± 0.03
Hardness (DPH, 500 g load)	235.9 ± 3.3	287 ± 10
Density (g/cm ³)	1.93 ± 0.01	2.12 ± 0.01
Poisson's ratio	0.28 ± 0.04	0.22 ± 0.01
Fracture toughness (MPa m ^{1/2})	1.68 ± 0.05	1.17 ± 0.17
Fatigue threshold (MPa m ^{1/2})	1.11	0.7
Fatigue crack velocity (m/cycle)	$4.0 \times 10^{-15} \Delta K^{70.3}$	$4.15 \times 10^3 \Delta K^{88.9}$
Coefficient of friction	0.15	0.15
Wear resistance	$< 1.23 \times 10^{-6}$ mm ³ /km	
Residual stress (MPa)	18.2	28.6

(From Ref. [9].)



Fracture Mechanics

The primary role of fracture mechanics is to calculate the strength of a component containing a flaw with the basic assumption that all engineering structures are damaged and contain flaws or cracks. A “stress intensity factor” K describes the interaction of applied stress, σ_a , with a crack of a given size, a , and geometry, Q , and is defined in combined stress and flaw size units as $\text{MPa m}^{0.5}$ or $\text{ksi in.}^{0.5}$:

$$K = Q\sigma_a(\pi a)^{0.5}$$

The stress intensity factor K_I for mode-I tensile crack has a critical value K_{Ic} at which the component fractures in a single loading cycle, somewhat analogous to a fracture strength.

Stress intensities are defined for a given crack geometry and size relative to the component geometry and stress field. Many geometry factors have been determined and may be found in handbook tables or calculated analytically.^[22] For a through thickness transverse crack of length $2a$ oriented perpendicular to the stress direction in a tensile stress field, σ , where w is the specimen width, the critical stress intensity is given as

$$K_{Ic} = \sigma(\pi a \sec(\pi a/w))^{0.5}$$

A somewhat different formula is used for the edge crack of the compact disk tensile specimen. The K_{Ic} obtained using the appropriate formula for a compact disk, which had values ranging from 10 to 11 mm,^[12] gave an average value of $K_{Ic} = 1.67 \text{ MPa m}^{1/2}$. The question arises whether this behavior would extend down to the size of a flaw that might exist in a leaflet of a heart valve. Crack lengths of 10 to 11 mm are large relative to PyC coating thicknesses and heart valve components.

One way to make a crack in carbon is to apply a load to it by a Vickers diamond indenter. This gives a four-sided pyramidal indentation, and if the material is brittle, cracks extend from the four corners of the indentation. It was found that the smallest crack length that could be obtained by the Vickers indenter in PyC was 0.5 mm.

To make a shorter crack, it was necessary to use an indenter with a sharper edge and a more acute angle. The Berkovich diamond indenter proved satisfactory. It induced projected crack lengths from 138 to 169 micrometers, which were reasonable relative to coating thicknesses and component sizes. The Berkovich cracks were about 1/3 the length of the smallest Vickers crack and about 1/150 the size used in the compact tensile specimen. The pure carbon PyC specimens were loaded to failure to determine K_{Ic} . Averaged over five

tests, the K_{Ic} value was $1.76 \text{ MPa}\sqrt{\text{m}}$ with a standard deviation of 0.152. This is only slightly higher than the value of 1.67 obtained for the much larger cracks of the compact tension specimen.

Fatigue Tests

Three early studies—Schoen, Shim, and Kepner et al.^[23–25]—showed that the S–N curves for silicon-alloyed PyC were almost horizontal and within the scatter band of the static strength for up to 10^7 cycles and for Kepner et al. up to 10^8 cycles. The S–N curves appear to be horizontal, but it is our opinion that the cyclic stresses did not initiate the fractures. Rather, inadvertent stress excursions due to machine wear, machine vibration, and accidental shock initiated the fractures.

In the study of silicon-alloyed PyC, a specimen was tested at a strain of $10,000 \mu\epsilon$, $R = 0.05$, for 1.66×10^9 cycles without any indication of failure. (R is the ratio of the minimum stress of a cycle to the maximum.) The $10,000 \mu\epsilon$ strain is 94% of its average static fracture strain. For one lot of the silicon-alloyed carbon, the average static fracture strain was $10,610 \mu\epsilon$ with a std. dev. of $603 \mu\epsilon$. An unalloyed PyC specimen was cycled for 6.73×10^8 cycles without any sign of failure at $13,700 \mu\epsilon$, $R = 0.5$, which was 90% of the static fracture strength. It had survived an initial stress excursion of $14,200 \mu\epsilon$. (For this lot, the average strength was $15,200 \mu\epsilon$ with a std. dev. of $2400 \mu\epsilon$.)

Stepped Fatigue Test

From the three early tests cited previously, it appears that the fatigue strength, even at high cycles, is equal to the static strength. In order to demonstrate this more clearly, a novel testing procedure was developed. Tests on silicon-alloyed and unalloyed PyC are presented in detail in Ref. [26]. Static tests on 10 unalloyed specimens gave an average of $14,600 \mu\epsilon$ with a std. dev. of $2000 \mu\epsilon$; lowest value = $11,200$, highest value = $18,100$. It was found that cyclic stressing for 10^6 cycles in the scatter band of the static strength did not affect the subsequent distribution of the static strength.

Propagation and Non-Propagation from Microcracks

The average critical stress intensity factor for brittle fracture from microcracks in unalloyed PyC was $1.76 \text{ MPa}\sqrt{\text{m}}$. A reference level, σ_{ref} , for the threshold value for cyclic crack growth was taken to be 90% of this, i.e., $1.58 \text{ MPa}\sqrt{\text{m}}$. This value corresponds to an elastic strain of $4300 \mu\epsilon$ (133 MPa) and a crack length of $138 \mu\text{m}$. Considering the statistical variation of the

Table 3 Threshold study for cyclic crack growth

Interval	Cycles	% of σ_{ref}	Crack length (μm)		Growth?
			Initial	Final	
1	10^6	90	169	223	Yes
2	10^6	80	223	223	No
3	10^6	90	223	223	No
4	10^6	103	223	223	No
5	10^6	117	223	240	Yes (slightly)
6	10^6	107	240	240	No

properties, it is likely that this value might be in the range of the threshold value for crack growth under cyclic stress. Five specimens had cracks with lengths from 136 to 170 μm induced by the Berkovich indenter. This study is discussed in detail in Ref. [26]. The results on one of the specimens, given in Table 3, illustrate the behavior.

The test procedure was to subject the specimen to intervals of 10^6 cycles at different amplitudes of stress intensity factor near the reference value of $1.58 \text{ MPa}\sqrt{\text{m}}$ and see whether the crack grew or did not grow. For this particular sample, the following tests were performed and results obtained.

Notice that for the crack to grow in interval 5, the stress had to be raised to 117% of the reference value, thus making for this specimen the value for crack growth 105% above the value for brittle, fast fracture. Of course, it should have failed in fast fracture if there had not been a statistical variation of both the threshold value for growth and the critical value for fracture. In interval 6, the crack did not grow. The threshold value for crack growth is above this value. The value of the stress intensity amplitude for interval 6 was $1.69 \text{ MPa}\sqrt{\text{m}}$. The threshold for this specimen is above this value, but below $1.76 \text{ MPa}\sqrt{\text{m}}$.

The question that must be answered is, why did the crack grow in interval 1? It is known that there is a residual tangential stress field around an indentation and that it rapidly diminishes the farther the region is from the indentation. In the first interval, the crack grew under both the applied stress and this residual stress until it left the residual stress field.

This portion of the study answered two questions.

1. Is the crack tip made by the Berkovich indenter as sharp as the crack tip grown under cyclic stress? In this study, one might conclude that the indenter-induced crack tip is even sharper, but we believe this is the effect caused by the local residual tensile stress around the indentation.
2. Is linear elastic fracture mechanics valid over the range of crack sizes investigated in this series of studies from 11 mm down to 100 μm ? Based on the

values obtained, the answer is yes. In Ref. [27], a study with a transmission electron microscope at ultrahigh magnification showed that the structure of the carbon was in the nanometer range, thus indicating that, compared to the structure, 100 μm is also a large crack.

Structure

The structure was studied by transmission electron microscopy in the phase contrast mode at magnification ranging from 800,000 X to 4.5 million X. At the lower magnification, microbeads with an average diameter of 0.5 μm could be discerned. The pattern within the beads resembles that of a sliced onion.^[24] The beads might also be the tops of cones observed by Kaae.^[28] At a magnification of 1 million X, a spot pattern appears with all the spots having an average diameter of about 40 Å.^[29] See Fig. 6. This is about the size of the microdomains measured by X-ray diffractometry. The fact that the fatigue strength at high cycles could not be distinguished from the static strength leads us to believe that the isotropic PyC is

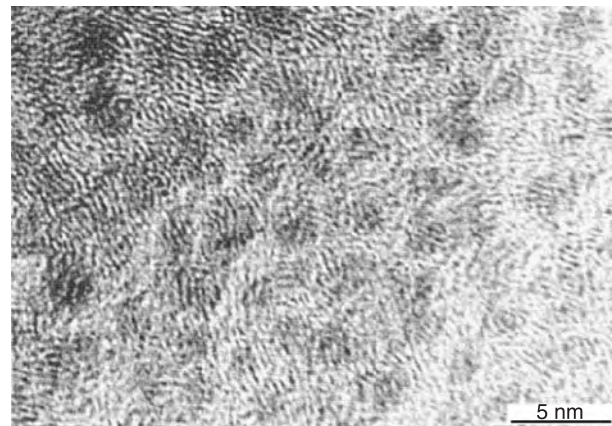


Fig. 6 Transmission electron microscope micrograph of unalloyed PyC (1,119,300 \times magnification). The spots are turbostratic carbon crystallites with an average size of 40 angstroms.

not subject to fatigue. Most theories of cyclic fatigue depend on the action of mobile dislocations. No dislocations were observed, and if they did exist, the extremely small size of the microdomains (4 nanometers) would make them immobile.

SUMMARY

The elastic strain at fracture for the unalloyed, isotropic PyC fatigue specimens was 1.46% at a stress of 448 MPa. The Young's modulus was 30.57 GPa and Poisson's ratio 0.28.

It was found that cyclic stressing within the scatter band of the static strength did not initiate cracks; induced cracks did not grow under cyclic stressing if below a threshold stress intensity factor, and this threshold stress intensity factor was valid over a wide range of crack lengths. Cyclic stressing did not reduce the static strength. No failures occurred when batches of 29 specimens were tested above the service stress at 6×10^8 cycles. The service stress is a small fraction of the fracture strength, thus a very high probability of survival can be ensured by a proof test of the assembly.

DESIGN ANALYSES EXAMPLES

Long-term implant devices in the United States require FDA premarket approval (PMA); a summary of safety and effectiveness (SSE) for each new device is prepared and made available to the public through the Freedom of Information Act. Many are directly available for download on the FDA Web site, <http://www.fda.gov/cdrh/>. While these documents typically give the results of analytic design evaluations, the exact details of the analyses may not be available because they are often considered proprietary information. However, for heart valves, several informative studies have been published in a special issue of the *Journal of Heart Valve Disease* 1996;5 Supplement I, which focused on fatigue and wear in mechanical heart valves. The best example from this issue, because it is the closest approximation of actual analyses used for a PMA, is given by Ryder and Cao.^[30]

Critical input information for Ryder and Cao's analysis includes estimates of the worst-case in-service loads (300 mm Hg pressure), which are used in finite element analyses to determine worst-case in-service stresses, σ_a , on the components. A proof test is then designed to apply a proof stress, σ_p , that defines an initial flaw size, a_i , in the component.

$$a_i \leq 1/\pi(K_{1c}/Q\sigma_p)^{0.5}$$

The initial flaw is then assumed to propagate with each loading cycle N by Paris law fatigue crack propagation,

$$da/dN = C\Delta K^m$$

where C is the fatigue constant and m the fatigue exponent. The crack grows until it reaches the final size, a_f , that satisfies the K_{1c} fracture criteria at the worst-case in-service stress:

$$a_f \geq 1/\pi(K_{1c}/Q\sigma_a)^{0.5}$$

The number of cycles to fracture, N_f , is the component lifetime and is found by integrating the Paris law crack growth rate:

$$N_f = 2/(C(m-2)a_i^{(m-2)/2})(K_{1c}/Q\sigma_a\pi^{0.5})^m$$

Cao's form of the Paris law fatigue crack velocity equation as given above is a simplification that recognizes that the initial flaw is much smaller than the final flaw size, and that stresses in heart valves range from zero to a maximum tensile value, without compression. For these reasons, Cao's fracture mechanics parameters as cited in Ref. [30] have slightly different values than those given in Table 2.

A heart is assumed to beat 40 million times a year, so the number of cycles divided by 40 million gives the damage-tolerant life in years. A summary of Cao and Ryders' results are given below in Table 4.

The calculated lifetimes are striking by their magnitude compared with an expected human lifetime. However, the small crack and stepped fatigue crack propagation results presented above demonstrate that crack propagation occurs only within the scatter band for the single cycle strength. Furthermore, the existence of a fatigue threshold has been ignored. Stresses may seem high relative to the values given in Table 2, but the effective stressed volumes are very different from those of the standard flexural stress test specimen and failure stresses scale by the effective stressed volume Weibull relationship. The Carbomedics valve design also survived the 600 million cycle accelerated test in severe loading conditions. Stress corrosion cracking has not been mentioned because the stress corrosion cracking rates for PyC are well below those for fatigue crack propagation.

Safe margins from fatigue failure demand good component design and require that no external damage is induced while handling the components during manufacture or implant. Fatigue crack propagation has been observed only when a crack has been deliberately induced in a component, as is done in the preparation of fracture mechanics specimens, or because of damage induced by an external mechanism,

Table 4 Carbomedics valve damage-tolerant lifetimes

Component and location	Proof stress (MPa)	Initial flaw size (μm)	Design stress (MPa)	Final flaw size (μm)	Lifetime (years)
Leaflet B-datum	167	21.5	47	250	1.8×10^{30}
Leaflet ear shoulder	64	143	33	431	4.6×10^{16}
Leaflet tip	157	24.3	40	325	5.2×10^{32}
Leaflet ear	530	2.67	310	7.8	7.8×10^{11}
Orifice pivot	530	2.67	310	7.8	7.8×10^{11}
Orifice groove	67	163	30	565	1.1×10^{20}
Orifice rim	78	98	34	414	3.0×10^{20}

such as cavitation, as occurred in the Baxter-Edwards failures. PyC does not have a known intrinsic damage generation or accumulation mechanism, such as dislocation pileup as occurs in metallic materials.

Other devices such as the Ascension MCP that are not considered life-sustaining do not necessarily require the rigorous damage-tolerant design analysis. For orthopedic devices, a conservative strength-and endurance-based evaluation can be sufficient. However, the individual designer must rationalize and defend the worst-case in-service loads chosen to define strength and endurance tests.

CONCLUSION

PyC has enjoyed considerable success as a material for long-term critical implants over the past three decades. Clearly, substantial progress has been made recently in PyC manufacturing technology and in the understanding of its behavior in complex, long-term implant biomedical applications. Factors such as PyC biocompatibility, successful clinical experience, improved mechanical strength, stiffness approximating that of bone, durability, and fatigue resistance support the use of PyC in new devices for long-term implants. The variability of carbon structures together with the dependence of properties on structure suggests that the pursuit of new, yet-to-be-discovered carbon structures also offers a viable means to provide further advances in many diverse areas, including implant design.

ARTICLES OF FURTHER INTEREST

Heart Valves, Mechanical; Thrombosis

REFERENCES

1. LaGrange, L.D.; Gott, V.L.; Bokros, J.C.; Ramos, M.D. Compatibility of Carbon and Blood, Chapter 5. Artificial Heart Program Conference, National Heart

- Institute Artificial Heart Program, Washington, DC, June 9–13, 1969; Johnson Hegyeli, R., Ed.; 47–58.
2. Bokros, J.C. Deposition, Structure and Properties of Pyrolytic Carbon. In *Chemistry and Physics of Carbon*; Walker, P.L., Ed.; Marcel Dekker, Inc.: New York, 1969; Vol. 4, 1–118.
3. Akins, R.J.; Bokros, J.C. The deposition of pure and alloyed isotropic carbons and steady state fluidized beds. *Carbon* **1974**, *12*, 439–452.
4. Brewer, L.A. *Prosthetic Heart Valves*; Charles C Thomas Pub: Springfield, IL, 1969.
5. Bokros, J.C.; Haubold, A.D.; Akins, R.J.; Campbell, A.C.; Griffin, C.D.; Lane, E. The Durability of Mechanical Heart Valve Replacements: Past Experience and Current Trends. In *Replacement Cardiac Valves*; Bodnar, E., Frater, R.W.M., Eds.; Pergamon Press: New York, 1991; 21–48.
6. Sadeghi, H. Dysfonctions des protheses valvulaires cardiaques et leur traitement chirurgical. *Schwiez. Med. Wschr.* **1987**, *117*, 1665–1670.
7. Black, M.M.; Ccochrane, T.; Lawford, P.V.; Reul, H.; Yoganathan, A. Design and Flow Characteristics. In *Replacement Cardiac Valves*; Bodnar, E., Frater, R.W.M., Eds.; Pergamon Press: New York, 1991; 1–20.
8. Milano, A.; Bortolotti, U.; Mazzucco, A.; Mossuto, E.; Testolin, L.; Thiene, G.; Gallucci, U. Heart valve replacement with the Sorin tilting-disc prosthesis. *J. Thorac. Cardiovasc. Surg.* **1992**, *103*, 267–275.
9. Vallana, F.; Rinaldi, S.; Galletti, P.M.; Nguyen, A.; Piwnica, A. Pivot design in bileaflet valves. *ASAIO Trans.* **1992**, *38*, M600–M606.
10. Arru, P.; Santi, M.; Vallana, F.; Majini, G.; Ottaviani, G.; Paccagnella, A. A New PyC Film for Biomedical Applications. In *High Tech Ceramics*; Vincenzini, P., Ed.; Elsevier Science Pub. B.V.: Amsterdam, 1987; 117–126.
11. Bokros, J.C.; Atkins, R.J. Process for Depositing Pyrolytic Carbon Coatings. US patent 3,977,896, August 31, 1976.
12. Ely, J.L.; Emken, M.R.; Accuntius, J.A.; Wilde, D.S.; Haubold, A.D.; More, R.B.; Bokros, J.C. Pure pyrolytic carbon: Preparation and properties of a new material, On-X carbon, for mechanical heart valve prostheses. *J. Heart Valve Dis.* **1998**, *7* (6), 626–632.

13. Bokros, J. Carbon biomedical devices. *Carbon* **1977**, *15*, 355–371.
14. Tian, C.L.; Hetherington, V.J.; Reed, S. A review of pyrolytic carbon: Application in bone and joint surgery. *J. Foot Ankle Surg.* **1993**, *32* (5), 490–498.
15. Cook, S.D.; Thomas, K.A.; Kester, M.A. Wear characteristics of the canine acetabulum against different femoral prostheses. *J. Bone Jt. Surg.* **1989**, *71B*, 189–197.
16. Cook, S.D.; Beckenbaugh, R.D.; Redondo, J.; Popich, L.S.; Klawitter, J.J.; Linscheid, R.L. Long term follow-up of pyrolytic carbon metacarpophalangeal implants. *J. Bone Jt. Surg.* **1999**, *81A* (5), 635–648.
17. More, R.B.; Kepner, J.L.; Strzepa, P. Hertzian Fracture in Pyrolite Carbon. In *Bioceramics Volume 6*; Ducheyne, P., Christiansen, D., Eds.; Butterworth-Heinemann: Oxford, 1993; 225–228.
18. Chwirut, D.J.; Regnault, W.F. Fracture mechanics principals applied to implant medical devices—A review. *Med. Prog. Technol.* **1988**, *14*, 193–203.
19. Ritchie, R.O.; Lubock, P. Fatigue life estimation procedures for the endurance of a cardiac valve prosthesis: Stress/life and damage tolerant analyses. *J. Biomech. Eng.* **1986**, *108*, 153–160.
20. Klepetko, W.; Moritz, A.; Mlczech, J.; Schurawitzki, H.; Domanig, E.; Wolner, E. Leaflet fracture in Edwards–Duromedics bileaflet valves. *J. Thorac. Cardiovasc. Surg.* **1989**, *97*, 90–94.
21. Ritchie, R.O.; Dauskardt, R.H.; Weikang, Yu; Brendzel, A.M. Cyclic fatigue-crack propagation and fracture toughness behavior in pyrolytic carbon-coated graphite for heart valve applications. *J. Biomed. Mater. Res.* **1990**, *24*, 189–206.
22. Tada, H.; Paris, P.C.; Irwin, G.R. *Stress Analyses of Cracks Handbook*, 2nd Ed.; Paris Publications/Del research Corp.: St. Louis, MO, 1985.
23. Schoen, F.J. On the fatigue behavior of pyrolytic carbons under cyclic loading. *Carbon* **1973**, *11*, 413–414.
24. Shim, H.S. The behavior of isotropic pyrolytic carbons under cyclic loading. *Biomater. Med. Dev. Artif. Organs* **1974**, *2* (1), 55–65.
25. Kepner, J.; Haubold, A.D.; Beavan, L.A. Cyclic Fatigue Testing of Pyrolytic Carbon, 41st Pacific Coast Regional Meeting of the American Chemical Society, San Francisco, October, 1990.
26. Ma, L.; Sines, G. Unalloyed pyrolytic carbon for implanted mechanical heart valves. *J. Heart Valve Dis.* **1999**, *8*, 578–585.
27. Ma, L. George Sines. Fatigue behavior of a pyrolytic carbon. *J. Biomed. Mater. Res.* **2000**, *51* (1), 61–68.
28. Kaae, J.L. The mechanism of the deposition of a pyrolytic carbon. *Carbon* **1985**, *23* (6), 665–673.
29. Ma, L. *Studies on Pyrolytic Carbons for Biomedical Applications*; Dissertation. University of California: Los Angeles 1997.
30. Ryder, J.K.; Cao, H. Structural integrity assessment of heart valve prostheses: A damage tolerance analysis of the carbomedics prosthetic heart valve. *J. Heart Valve Dis.* **1996**, *5* (Suppl I), S86–S96.

Rapid Prototyping

Shaochen Chen
Carlos A. Aguilar
Yi Lu

Department of Mechanical Engineering, University of Texas at Austin, Austin, Texas, U.S.A.

INTRODUCTION

The term rapid prototyping (RP) refers to a broad category of techniques that can automatically construct physical prototypes from computer-aided design (CAD) models. Rapid prototyping provides the unique opportunity to quickly create functional prototypes of highly complex designs in an additive fashion. This layered manufacturing is also known as solid freeform fabrication (SFF), desktop manufacturing, and computer-aided manufacturing (CAM). There are several methods of RP, but universal to all of them is the basic approach they use, which can be described in three phases:^[1,2]

1. A geometric model is constructed on a CAD/CAM system. Because the manufacturing technique relies on a layer-by-layer additive process, the model must be represented as closed surfaces, which define an enclosed volume.
2. Next, the CAD model is converted into an STL (STereoLithography) file format, which approximates the surfaces of the model by polygons. The more complex the geometry of the model, the larger the number of polygons needed to characterize the surface. Curved surfaces usually require more polygons to fit and in turn produce larger files.
3. A computer program reads the generated STL file and “slices” the model into a finite set of layered cross sections. Each layer is created individually from liquid, powder, or solid material and stacked onto the previous layer with each layer joined to its neighboring surfaces. The created three-dimensional (3-D) model resembles the geometry of the CAD structure.

There are a variety of methods that involve SFF, but several approaches have found frequent use in the biomedical arena. In this entry, the RP fabrication techniques are categorized by virtue of their state of material. The methods can be classified into three categories: 1) liquid-based systems; 2) powder-based systems; and 3) solid-based systems (see Fig. 1). The

liquid-based systems cure liquid material into a solid state through interaction with a laser beam. The most popular liquid-based technique used for biomedical applications is stereolithography (SL). The powder-based systems rely on a laser beam to melt and fuse powder grains or selectively bind powder particles using liquid binders. Frequently used powder-based techniques in biomedical applications are selective laser sintering (SLS) and 3-D printing. The solid-based systems extrude melted material in filament form and build 3-D shapes by layering filament in predesigned patterns. A popular solid-based method, fused deposition modeling (FDM), has gained interest recently for several biomedical engineering uses. Each of the aforementioned techniques will be discussed in further detail in the subsequent sections.

There are several advantages of RP systems, and the most prevalent is the ability to rapidly produce functional prototypes in small quantities.^[3] This benefit has greatly improved the efficacy of surgical planning and education by providing a convenient platform for rapidly creating anatomical models.^[4] An excellent example of this is demonstrated in a recent study, where a patient was involved in a car accident and suffered from a serious bone fracture of the upper and lateral orbital rim in the skull.^[5] As the first reconstructive surgery had proven unsuccessful, it was necessary to perform subsequent surgery to transplant an artificial bone into the hole. This would require the surgeons to carve an artificial bone during the operation and try and fit it into the hole. Given that the operation would be very time-consuming because of the fitting and refitting of the artificial implant, RP was applied to create a prototype of the patient’s skull. The model was then used to prepare an artificial bone that would fit the hole properly. The use of RP model not only greatly reduced the surgery time and risk of infection but also improved its accuracy. In another example of successful surgical planning using RP, a patient was diagnosed with a cancerous bone tumor in his temple area. Initially, the surgery plan was to access the tumor through the frontal, highly functional area of the brain. As a consequence of the surgery, the patient would lose vision in one eye and several motor

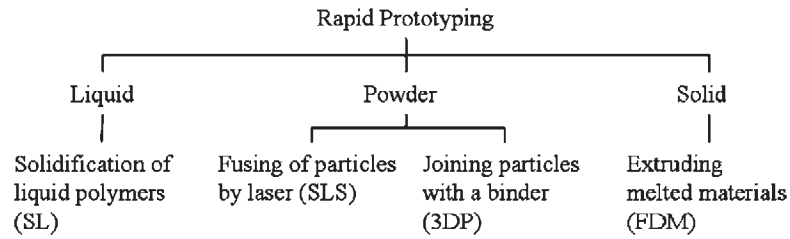


Fig. 1 Rapid prototyping systems.

functions would be impaired.^[6] Using RP technology, a replica of the patient's skull was manufactured from a series of 2-D computed tomography (CT) scans of the patient's skull. With the model, the surgeon was able to recognize that the tumor could be treated by entering through the patient's jawbone and that an alternative approach could be used for treatment. Using the alternative method, the risk of harming the eye and motor functions could be avoided. Other case studies relating to bone tumors have been reported with similar successful results.^[7,8]

The uses of RP for biomedical applications are extended beyond surgical planning and education. The single-operative RP processes also offer an excellent tool to aid in the development of new medical devices such as targeted drug delivery devices. An excellent example of this is illustrated in a recent study, where RP was used to manufacture morphologically detailed human tracheobronchial airway models in various disease states based on high-resolution anatomical imaging data.^[9] These anatomical detailed casts of respiratory tract structures that corresponded to diseased states from exposure to toxic or pathogenic aerosols were then used to explore new methods of targeted drug delivery via inhaled aerosols. These case studies are exceptional examples of the potential impact of RP for biomedical applications. Rapid prototyping processes also provide avenues to develop custom implants and prostheses, new scaffolds for the guided development of highly structured tissues, organs, and bones, next-generation in vivo controlled-release vehicles, and multilayer microfluidic networks. In the following sections, we will further explain several RP techniques and several examples of how RP can play an effective role in the biomedical industry.

STEREOLITHOGRAPHY

Stereolithography is perhaps the most widely used RP technique. It was commercialized by 3D Systems of Valencia, California, U.S.A.

Stereolithography is a liquid-based method that allows 3-D microfabrication in a room-temperature environment.^[1,10,11] This method relies on a photosensitive monomer resin or a photocrosslinkable polymer resin, which solidifies when exposed to ultraviolet

(UV) or visible light. The photosensitive monomers cure via photopolymerization, while the photocrosslinkable polymers become solids or hydrogels by forming cross-linking networks.

A typical SL system consists of a container filled with liquid photopolymer resin, a computer-controlled stage immersed in the resin, and a laser scanner, as shown in Fig. 2. The process begins with a 3-D solid model designed in CAD. The model is numerically sliced into a series of 2-D layers with an equal thickness in an STL file format.^[12] The code generated from each sliced 2-D file is then executed to scan a UV laser beam in the X-Y plane on the surface of the resin. The interaction with the laser beam causes the liquid to solidify in areas where the laser strikes. After the layer is completely traced and cured by the laser beam, the stage is lowered into the vat a distance equal to the thickness of a layer. However, as the resin is generally quite viscous, tools were developed to increase the speed of the process. Early SL systems drew a knife-edge over the surface to smoothen it. More recently, pump-driven recoating systems have been utilized. The scanning and recoating steps are repeated until the object is completely fabricated.

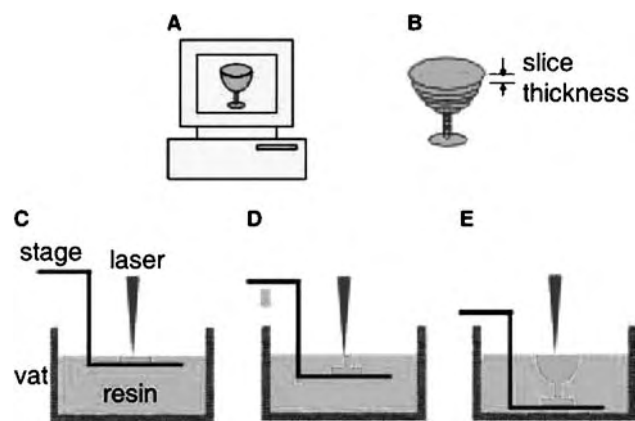


Fig. 2 Principle of stereolithography. (A) CAD model design. (B) STL file generation. (C) Laser scanning to photopolymerize the first layer of the 3-D structure. (D) The stage was lowered to produce the second layer of the 3-D structure. (E) Layer-by-layer selective photopolymerization to produce the entire 3-D structure.

The resolution of the conventional SL is limited by the spot size of the laser beam to approximately 25 μm . To fabricate smaller and more complicated geometries, such as tissue engineering scaffolds, minor modifications are made. The laser beam must be focused into a 1–2 μm spot by an objective lens. As it is relatively difficult to scan the objective lens, a 3-D stage is employed to move in the X – Y – Z directions instead. A line width of roughly smaller than 5 μm has been achieved using this approach.^[13]

This layer-by-layer liquid micromanufacturing system enables the fabrication of complex internal features, such as intricate passageways for microfluidic devices and curved surfaces, to be accurately produced.^[14] Furthermore, the approach can easily incorporate different proteins and microparticles containing polymer solutions for each layer (or even for partial layers). This allows SL the unique ability to rapidly create a precise spatial distribution of biochemical microenvironments within a single scaffold or system.

There are several factors that influence the quality of the SL generated products, and the photocurable polymer resin is the most sensitive to processing. Additionally, as the products are used for biomedical scaffold purposes, the demands of the resin are as follows: 1) biocompatibility and, in some cases, biodegradability; 2) low viscosity; 3) stability under visible light; 4) fast cross-linking kinetics; 5) compatibility with 3-D polymerization; and 6) reasonably good accuracy and mechanical strength in the cured state.

Several systems that meet the above demands have been investigated recently and include, poly(ethylene glycol) (PEG)-modified systems cross-linked with a photoinitiator, typically acetophenone (UV light) or eosin Y, and triethanolamine (visible blue light).^[15] Another system, comprising poly(ethylene glycol) dimethacrylate (PEGDMA) hydrogel polymer, has been used in a variety of *in vitro* and *in vivo* applications and has been proven to have no adverse effect on cells for a sufficient period of time. The polymer PEG, which is hydrophilic, highly wettable, and cell-adhesion resistant, can be modified with proteins [i.e., arginine-glycine-aspartic acid (RGD)] to enhance cell adhesion. Ten percent (w/v) resin of PEG has a viscosity close to that of water. Increasing the concentration leads to higher viscosity. The resin is stable under the normal experimental conditions, though sunlight should be minimized because it carries UV light. The curing speed depends on the UV sensitivity of the resin. High sensitivity indicates that a fast processing speed could be achieved. The resin can be cured only at the laser focal point and the uncured resin is reasonably stable under scattered or reflected UV light. Typical lateral resolution ranges from tens to hundreds of micrometers for PEGDMA systems. Though the resulting scaffold is relatively soft and fragile in aqueous environment, it has good tensile strength when water is extracted.

The successful fabrication of scaffolds using PEGDMA (MW 1000) along with cyto-compatible photoinitiator, 1-[4-(2-hydroxyethoxy)-phenyl]-2-hydroxy-2-methyl-1-propane-1-one, Darocure 2959, initiated by a frequency tripled Nd:YAG laser (355 nm in wavelength) was demonstrated recently.^[16] Poly(ethylene glycol) acrylates, covalently modified with the cell adhesive peptide RGD or the ECM component heparin sulfate, was incorporated within the scaffolds to facilitate cell attachment and to allow spatial sequestration of heparin-binding growth factors. Fluorescently labeled polymer microparticles and basic fibroblast growth factor (FGF-2) were chosen to illustrate the capability of SL to spatio-temporally pattern scaffolds. The results demonstrated a precise, predesigned distribution of single or multiple factors within the single 3-D structure with specific internal architectures. Functionalization of these scaffolds with RGD and heparan sulfate allowed efficient cell attachment and spatial localization of growth factors. Such patterned scaffolds might provide effective systems to study cell behavior in complex microenvironments and could eventually lead to engineering of complex, hybrid tissue structures through predesigned, multilineage differentiation of a single stem cell population.

The unique advantages of SL as a suitable technique for scaffolding lie in the following bases. First of all, the inherent photochemical process is attractive for the production of tissue engineering scaffolds because it has a controlled reaction initiation and termination, a relative short reaction time, and spatial control. Second, SL can be performed at mild conditions, such as room temperature and under physiologic pH, which allow incorporation of biologically active molecules and *in situ* cell seeding. Third, photopolymers can be modified with various peptides and copolymers, for instance, to promote or inhibit cell adhesion.

While SL does offer advantages for scaffolding, several issues must be addressed for subsequent use. First, as *in situ* cell seeding or *in vivo* scaffolding requires nontoxic precursors, initiators, and residues, a low concentration of initiators is essential. The inclusion of initiators frequently results in prolonged fabrication times and reduced control over resolution and composition of the scaffold. In addition, prolonged exposure to UV lasers may cause degradation in some bioactive molecules and reduce the capacity of the scaffolds to promote cell proliferation.

Selective Laser Sintering

Selective laser sintering is a popular RP powder-based technique. It was developed and patented at the University of Texas at Austin and commercialized by the DTM Corporation (now 3D Systems) in 1987.

The SLS process is similar to SL in that it creates 3-D objects layer by layer. However, SLS utilizes materials in powder form instead of liquid and uses heat generated from a CO₂ laser. The method begins by depositing a thin layer of material in powder form into a container. The layer of powder is then selectively irradiated by a CO₂ laser to the point of melting, fusing the powder particles to form a solid mass, as shown in Fig. 3 below. The intensity of the laser beam is modulated to melt the powder only in areas defined by the part's geometry, while the rest of the material remains in its original powder form. When the cross-sectional layer is solidified, sequential layers of powder are deposited via a roller mechanism on top of the previously scanned layer. The steps are repeated with each layer fusing to the layer below it until the 3-D model is completely built. As the material used in the SLS process is in powder form, the powder not melted or fused during the processing serves as a customized, built-in support structure. Therefore, there is no need to create support structures within the design phase or during processing and no support structure to remove when the part is complete. There is a broad range of materials that can be processed using SLS and no postprocessing or postcuring is required.

Much like SL, SLS provides an economical and efficient method to construct scaffolds for guided development of human tissues and organs. As SLS is also a layer-by-layer additive process, the design and fabrication of anatomically shaped scaffolds with varying internal and external architectures is possible.^[17] Additionally, as SLS uses materials in powder form, the precise control over pore interconnectivity, permeability, and stiffness of the scaffold can be achieved. The control of porosity arises from space between the individual granules of powder. The manipulation of these properties may optimize cell infiltration and adhesion, as well as mass transport of nutrients and metabolic waste throughout the scaffold. The SLS technique was recently

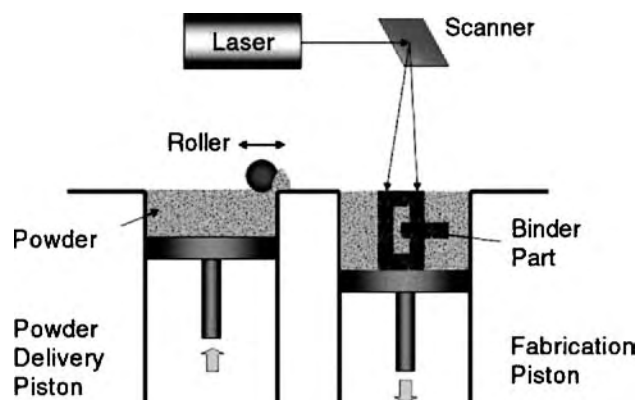


Fig. 3 Principle of selective laser sintering.

exhibited to engineer bone tissue scaffolds using the common biodegradable polymer, poly(ϵ -caprolactone) (PCL).^[18] Scaffolds were manufactured in a variety of designs and the biological properties of each were evaluated by seeding each with bone morphogenetic protein-7 (BMP-7) transduced human fibroblasts. The generated tissues were then studied using micro-computed tomography (μ CT) scans and histology. The results showed that PCL scaffolds fabricated via SLS possessed similar mechanical properties to trabecular bone and enhanced tissue in-growth in vivo.

Recently, it was demonstrated that a derivative of the SLS approach could be used to fabricate bioactive scaffolds using a layer of carbon, a strong absorber of laser radiation, deposited on a biodegradable polymer, poly(D,L-lactic) acid (DLPLA).^[19] In this experiment, near-infrared ($\lambda = 970$ nm) laser radiation, which DLPLA polymer particles do not absorb, was used to initiate the sintering process between a small quantity of carbon microparticles that were homogeneously distributed on the surface of the DLPLA particles. This limited the melting process to the surfaces of each particle and possibly opens up the SLS process to a range of polymers that could not previously be processed by conventional SLS. Furthermore, as the laser melts only the surface of the particles, delicate bioactive species such as enzymes or growth factors trapped within each particle can retain their activity throughout the processing.

The SLS technique has also been realized to fabricate controlled-release drug-delivery devices.^[20] A device composed of nylon was fabricated by selectively controlling the position and composition of the material. The drug model used for infiltrating the sample matrix was Fluka methylene blue dye. The results showed a gradient release profile and indicated that the release mechanism can be tailored to fit the specific application by controlling the device wall composition, anisotropy, and microstructure.

The SLS process also provides the ability to design and optimize subject-specific prototypes of orthotic devices that serve to compensate for a variety of injuries, disorders, and diseases. This was exhibited recently, where a patient suffered from a large frontal cranium defect after complications from a previous meningioma tumor surgery.^[2] The surgery left the patient with a missing cranial section, which caused the patient's head to look slightly deformed. A corrective surgery would have required a titanium-mesh plate to be hand-formed by the surgeon during the operation using trial and error. Using a CT scan of the affected area, a 3-D CAD generated model could be made, and fabricated using SLS. The SLS mold could then be used to mechanically press the titanium-mesh plate to the anatomical profile of the missing cranium section. The SLS technique improved the fit of the implant and significantly reduced the operation time.

Some disadvantages of the SLS technique are surface roughness, which can influence the cellular behavior in a scaffold, and trapped powder, which can affect the scaffold stiffness. Another disadvantage of the SLS process is the smallest attainable feature size, which is primarily governed by powder particle size, focused laser beam diameter, and heat transfer in the powder bed.

3-D Printing

Three-dimensional printing (3-DP) was invented and patented at the Massachusetts Institute of Technology. It was subsequently licensed and further developed by Z Corporation (Massachusetts, U.S.A.).

Three-dimensional printing has demonstrated the capability of fabricating microstructures and controlling local composition with a high resolution in the interior of the component.^[21] Similar to SL and SLS, a computer model (CAD) of the desired part is generated and a slicing algorithm draws detailed information for every layer. Just as in SLS, the process begins with a thin distribution of powder spread over the surface of a powder bed. Using a technology similar to ink-jet printing, a binder material selectively joins particles where the object is to be formed. A piston that supports the powder bed and the part-in-progress lowers so that the next powder layer can be spread and selectively joined. This layer-by-layer process is repeated until the part is completed (Fig. 4). The local microstructure within the component can be controlled either by changing the binder, which is printed, or by changing the printing parameters during the component construction. Although the system was initially designed for metallic or ceramic particles, powders of various polymeric biomaterials can be used directly. To print tissue engineering scaffolds, liquid solvent binder can be printed onto a powder bed of porogens (dissolvable particles that creates pores in a polymer matrix) and polymer particles. The solvent will dissolve the polymer and evaporate, and the polymer will reprecipitate to

form solid structures. The final porosity is achieved after particulate leaching and solvent removal.

A large combination of natural and synthetic biomaterials can be selectively bound using organic or water-based binders. Synthetic polymers include poly(lactic-co-glycolic acid) (PLGA), PCL, PLA, polyethylene oxide (PEO), with common organic binders such as chloroform, dichloromethane. Cornstarch, dextran, and gelatin have also been used in conjunction with water-based binders. Fine biomaterial powders are usually made by cryogenic milling with liquid nitrogen. The powders are then vacuum dried and sieved to obtain a uniform size.

Devices consisting of PCL and PEO were fabricated to demonstrate control of drug delivery profiles by controlling the position, composition, and microstructure.^[21] The top and bottom layers of the device were constructed by binding PCL powder into thin solid layers. A cellular-type pattern was printed with PEO, which has a faster degradation rate, in the intermediate layers. Dyes, which represent drugs, were selectively placed within the cells manually.

The fabrication sequence is described as follows: first, a bead of PCL powder (45–75 μm) is spread in a thin layer on the piston. Droplets of 20% acid-modified PCL LPS-60HP (PCL-LPS)/chloroform binder solution are deposited on the PCL powder layer through a 45 μm orifice plate to form a solid PCL sheet. Poly(ϵ -caprolactone) particles bind to each other according to the pattern of the printhead motion as the binder contacts the PCL layer. Lines of PCL are formed by moving the printhead in a linear motion. A dense PCL sheet is formed by repeatedly printing lines directly adjacent to each other. The piston is then lowered and ready for the next layer of powder. The second layer begins by spreading PEO powders (75–150 μm) on the top of the solid PCL sheet. A square grid pattern of 5 cells by 5 cells can be made by printing six parallel lines at 3 mm interline spacing, and then repeating another set of six lines at 90° to the first set by rotating the piston. Twenty-five cells, which serve as drug reservoirs, are

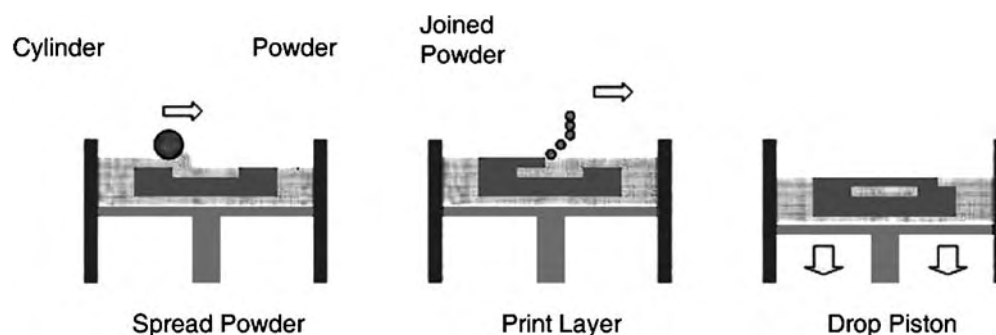


Fig. 4 Principle of 3-D printing.

formed by 20 layers of PEO grid pattern. Different dyes are selectively dropped into the unbound PEO powder before a top layer of dense PCL sheet is printed to seal the cells. Different release mechanisms can be achieved by controlling device wall composition, anisotropy, and microstructure with this technique.

Recently, a unique, heterogenous, osteochondral scaffold, which could be an effective tool for treating articular defects, was developed using the 3-DP process. The scaffold consisted of two distinctive regions that differed in material composition, porosity, macro-architecture, and mechanical properties. The cartilage region was 90% porous and composed of D,L-PLGA/L-PLA, with macroscopic staggered channels to facilitate homogenous cell seeding. The cloverleaf-shaped bone portion was 55% porous and consisted of a L-PLGA/calcium phosphate tribasic composite, designed to maximize bone ingrowth while maintaining critical mechanical properties. The transition region between these two sections contained a gradient of materials and porosity to prevent delamination. It was shown that chondrocytes preferentially attached to the cartilage portion of the device and cartilage formed during a 6-week in vitro culture period. The tensile strength of the bone region was similar in magnitude to fresh cancellous human bone, suggesting that these scaffolds have desirable mechanical properties for in vivo applications, including full joint replacement.

The fabrication of complex-shaped scaffolds has also been demonstrated using an indirect approach of the 3-DP process, where molds are printed and the final materials are cast into the mold cavity.^[22] To demonstrate the ability of this approach to build complex-shaped scaffolds from common biodegradable polymers with large pore size directly from medical imaging data, a zygomatic arch defect was fabricated. To evaluate the resolution that is possible with the indirect approach, scaffolds with small villi architecture were constructed. As the indirect approach involves the use of molding and mold release materials, the study addressed the removal of these additional materials by evaluating cell proliferation in the demolded scaffolds. To demonstrate the benefits of printing small features, cell attachment and proliferation in various regions of the scaffolds were also studied and showed cell growth in culture. This technique may be a useful adjunct for the fabrication of complex scaffolds for tissue engineering.

Postprocessing is necessary to remove the residual solvents and porogens, enhance the strength of the structures, or absorb functional molecules. Chlorinated solvents are undesirable because they may bring potential health risks. Those solvents are difficult to extract completely. An extraction method based on liquid CO₂ has been developed.^[23] Using water as a solvent eliminates the extraction procedure.

The as-solidified structures are vacuum dried at an elevated temperature to enhance the structural rigidity. Chemical modification may be performed, for instance, by infiltrating polymer structures with functional molecules to adjust the hydrophilicity.

In general, 3-DP is a fast and accurate process. It has precise control over composition, anisotropy, and microstructure. A wide range of materials can be used, provided they can be processed into powder form. It is possible to incorporate biomolecules, which may be carried by the binders, in the structures. However, the surface of the structures is usually powdery and the trapped powder is hard to completely remove.

Fused Deposition Molding

Fused deposition modeling is a widely used RP process that was developed by Scott Crump and commercialized by Stratasys (Minnesota, U.S.A.).

In this method (see Fig. 5) a spool of material is heated into a semiliquid state and dispensed into an extrusion head. The semiliquid material is extruded through the head and deposited into the desired shape. The platform, onto which the material is deposited, is maintained at a lower temperature so that the material cures immediately on contact after exiting the heated chamber. After the completion of each layer, the platform lowers, and the extrusion head deposits the next layer on top of the previous one until the 3-D model is completed. In newer FDM machines, two materials are fed through the extrusion head. One is used to produce the geometric model and the other is used to construct support structures. The second material forms a bond with the primary model material and is washed away when the 3-D model is processed. The advantages of using FDM include the speed and safety of the machine.

Fused deposition modeling has been demonstrated as a viable method for manufacturing functional

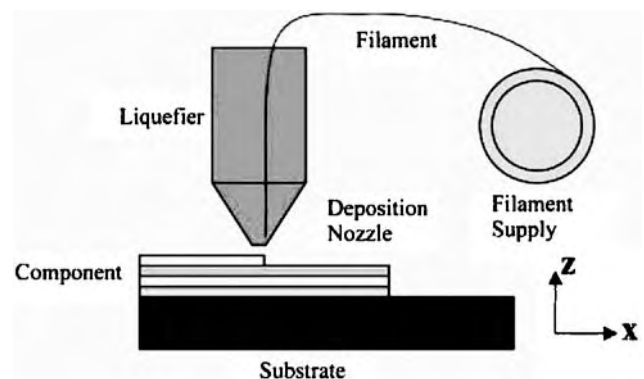


Fig. 5 Principle of fused deposition molding.

Table 1 Comparison of different rapid prototyping technologies in biomedical applications

RP system	Resolution (μm)	Material	Strength	Weakness	Reference
Stereolithography	100	PEG-acrylate	Biofriendly liquid environment; control over reaction; high resolution; can incorporate biomolecules and cells in situ	Relatively slow processing; UV exposure may cause degradation	[10–17]
Selective laser sintering	400	PCL, PEEK–HA	Fast processing; control over porosity; no postprocessing	High temperature; material must be in powder form; rough surface finish; trapped powder	[17–21]
Three-dimensional printing	100	PLGA, PEO, starch-based polymer	Fast processing; nonorganic binder is possible; no supporting structure needed; control over porosity	Material must be in powder form; rough surface finish; trapped powder	[17, 22]
Fused deposition molding	250	PCL, PCL–TCP	Can be operated in office-type environment; good mechanical strength	High temperature; simple geometry can be obtained	[23–26]

PEEK, polyetheretherketone.

scaffolds for tissue engineering. Tissue-engineering scaffolds with fully interconnected pore networks based on PCL were fabricated using FDM.^[24] The scaffolds were then seeded with human fibroblasts from the anterior cruciate ligament of a 34-year-old male. The 3-D structures showed excellent cell proliferation, differentiation, and extracellular matrix production. An adapted FDM method was recently applied in a similar fashion to create scaffolds. This method used a microsyringe to expel a dissolved polymer under constant pressure to form a desired pattern. Using this microextrusion approach, resolution down to 20 μm can be achieved.^[25] A modified FDM method produced PLGA scaffolds with a lateral resolution of up to 5 μm . The layer-by-layer fabrication method allows the design of a pore morphology that varies across the structure.^[26] Researchers have also used FDM to manufacture composite scaffolds of PCL, PCL–Tricalcium phosphate (TCP) and PCL–hydroxyapatite (HA). The scaffolds were seeded with human mesenchymal progenitor cells and proliferation toward and onto the surfaces were detected.

Another version of FDM uses a microsyringe to construct 3-D heterogenous multilayer tissue-like structures inside microchannels.^[27] The process involves a

three-step approach: 1) immobilization of a cell–matrix assembly; 2) cell–matrix contraction; and 3) pressure-driven microfluidic delivery using a syringe. Using this approach, patterning of biological microstructures can be used to form “neotissue” from different types of cells and biopolymer components in one single system. Thus, the approach provides a novel solution to fabricate hybrid biopolymers and hierarchical tissue structures for tissue engineering and basic cell biology.

Fused deposition modeling provides the unique ability to use almost any material but is limited by its narrow processing window, shrinkage, and high operating temperature, which could prevent the inclusion of biomolecules into the material. The approach also requires the input material in a particular size and geometry. (For a comparison of different rapid prototyping technologies in biomedical applications see Table 1.)

CONCLUSIONS

In addition to the aforementioned techniques, other plausible approaches for developing medical devices are available. One such method was recently developed as a means to create biodegradable microfluidic devices, using

a rational scaffold design.^[28] The novel approach used a thermal fusion bonding technique to rapidly fabricate 3-D monolithic microdevices with various internal geometries. The researchers stacked micropatterned thin films of the common biodegradable polymer PLGA, to create multilayer, microfluidic networks with extruded inlets and outlets without leaks and occlusions. The method also presents a viable platform for generating drug delivery vehicles with microscale cavities, capsules and microchannels for controlled-release locations, species and kinetics in complex biomolecular environments.

Another excellent study developed a technique derived from conventional integrated circuit technologies to photopattern hydrogels.^[29] The method began with a polytetrafluoroethylene (Teflon) base filled with polymer solution loaded with cells. Ultraviolet light was then shone through a patterned template atop the thin film, curing the exposed polymer that was seeded with cells. Using this layer-by-layer method, the construction of complex networks of cells and materials can be built by using various templates and depositing different layers atop one another.

Ultimately, RP platforms serve as an invaluable tool to achieve exciting new functionality in the fields of microfluidics, cell-support scaffold development, and controlled drug-release devices. The techniques are rapid, versatile, and low-cost approaches to precision manufacturing, thus making them ideal for large-scale engineering of complex, high-performance medical devices.

ACKNOWLEDGMENT

The authors appreciated financial support from the U.S. National Science Foundation (DMI 0093364).

ARTICLES OF FURTHER INTEREST

Photopolymerization; Tissue Engineering, Microscale; Tissue Engineering of Microvascular Networks

REFERENCES

- Chua, C.K.; Leong, K.F. *Rapid Prototyping: Principles and Applications in Manufacturing*; John Wiley: Singapore, 1997.
- Chua, C.K.; Leong, K.F.; Lim, C.S. *Rapid Prototyping: Principles and Applications*; World Scientific: Singapore, 2003.
- Beaman, J.J.; Barlow, J.W.; Bourell, D.L.; Crawford, R.H.; Marcus, H.L.; McAlea, K.P. *Solid Freeform Fabrication: A New Direction in Manufacturing*; Kluwer: Dordrecht, 1997.
- Liew, C.L.; Leong, K.F.; Chua, C.K.; Du, Z. Dual material rapid prototyping techniques for the development of biomedical devices. Part 1: space creation. *Int. J. Adv. Manuf. Technol.* **2001**, *18* (10), 717–723.
- Adachi, J.; Hara, T.; Kusu, N.; Chiyokura, H. Surgical simulation using rapid prototyping. *Proceedings of the 4th International Conference on Rapid Prototyping*, 1993; 135–142.
- Mahoney, D.P. Rapid prototyping in medicine. *Computer Graphics World* **1995**, *18* (2), 42–48.
- Jacobs, A.; Hammer, B.; Niegel, G.; Lambrecht, T.; Schiel, H.; Hunziker, M.; Steinbrich, W. First experience in the use of stereolithography in medicine. *Proceedings of the 4th International Conference on Rapid Prototyping*, 1993; 121–134.
- Swaelens, B.; Kruth, J.P. Medical applications in rapid prototyping. *Proceedings of the 4th International Conference on Rapid Prototyping*, 1993; 107–120.
- Clinkenbeard, R.E.; Johnson, D.L.; Parthasarathy, R.; Altan, M.C.; Tan, K.H.; Park, S.M.; Crawford, R.H. Replication of human tracheobronchial hollow airway models using a selective laser sintering rapid prototyping technique. *J. Am. Ind. Hygiene Assoc.* **2002**, *63* (2), 141–150.
- Cho, D.W.; Lee, I.H. An investigation on photopolymer solidification considering laser irradiation energy in micro-stereolithography. *Microsystem Technol.* **2004**, *10*, 592–698.
- Cooke, M.N.; Fisher, J.P.; Dean, D.; Rimnac, C.; Mikos, A.G. Use of stereolithography to manufacture critical-sized 3D biodegradable scaffolds for bone ingrowth. *J. Biomed. Mater. Res. Part B: Appl. Biomater.* **2002**, *64* (B), 65–69.
- Zhang, X.; Jiang, X.N.; Sun, C. Micro-stereolithography of polymeric and ceramic microstructures. *Sensors Actuators* **1999**, *77*, 149–156.
- Cabral, J.T.; Hudson, S.D.; Harrison, C.; Douglas, J.F. Frontal photopolymerization for microfluidic applications. *Langmuir* **2004**, *20*, 10020–10029.
- Fisher, J.P.; Dean, D.; Engel, P.S.; Mikos, A.G. Photoinitiated polymerization of biomaterials. *Annu. Rev. Mater. Res.* **2001**, *31*, 171–181.
- Mapili, G.; Lu, Y.; Chen, S.C.; Roy, K. Laser-layered micro-fabrication of spatially patterned, functionalized tissue engineering scaffolds. *J. Biomed. Mater. Res.* *in press*.
- Bryant, S.J.; Nuttelman, C.R.; Anseth, K.S. Cyto-compatibility of UV and visible light photoinitiating systems on cultured NIH/3T3 fibroblasts in vitro. *J. Biomater. Sci. Polym. Edn.* **2000**, *11* (5), 439–457.
- Yeong, W.Y.; Chua, C.K.; Leong, K.F.; Chandrasekaran, M. Rapid prototyping in tissue engineering: challenges and potential. *Trends Biotechnol.* **2004**, *22* (12), 643–652.
- Williams, J.M.; Adewunmi, A.; Schek, R.M.; Flanagan, C.L.; Krebsbach, P.H.; Feinberg, S.E.; Hollister, S.J.; Das, S. Bone tissue engineering using polycaprolactone scaffolds fabricated via selective laser sintering. *Biomaterials* **2005**, *26* (23), 4817–4827.
- Antonov, E.H.; Bagratashvili, V.N.; Whitaker, M.J.; Barry, J.J.A.; Shakesheff, K.M.; Kononov, A.N.

- Popov, V.K.; Howdle, S.M. Three-dimensional bioactive and biodegradable scaffolds fabricated by surface-selective laser sintering. *Adv. Mater.* **2005**, *17* (3), 327–330.
20. Cheah, C.M.; Leong, K.F.; Chua, C.K.; Low, K.H.; Quek, H.S. Characterization of microfeatures in selective laser sintered drug delivery devices. *Proc. Inst. Mech. Engrs. H: J. Eng. Med.* **2002**, *216* (H6), 369–383.
 21. Wu, B.M.; Borland, S.W.; Giordano, R.A.; Cima, L.G.; Sachs, E.M.; Cima, M.J. Solid free-form fabrication of drug delivery devices. *J. Controlled Release* **1996**, *40*, 77–87.
 22. Lee, M.; Dunn, J.C.D.; Wu, B.M. Scaffold fabrication by indirect three-dimensional printing. *Biomaterials* **2005**, *26*, 4281–4289.
 23. Sherwood, J.K.; Riley, S.L.; Palazzolo, R.; Brown, S.C.; Monkhouse, D.C.; Coates, M.; Griffith, L.G.; Landeen, L.K.; Ratcliffe, A. A three-dimensional osteochondral composite scaffold for articular cartilage repair. *Biomaterials* **2002**, *23*, 4739–4751.
 24. Hutmacher, D.W.; Schantz, T.; Zein, I.; Ng, K.W.; Teoh, S.H.; Tan, C.K. Mechanical properties and cell cultural response of polycaprolactone scaffolds designed and fabricated via fused deposition modeling. *J. Biomed. Mater. Res.* **2001**, *55* (2), 203–216.
 25. Vozzi, G.; Flaim, C.; Ahluwalia, A.; Bhatia, S. Fabrication of PLGA scaffolds using soft lithography and microsyringe deposition. *Biomaterials* **2003**, *24* (14), 2533–2540.
 26. Zein, I.; Hutmacher, D.W.; Tan, K.C.; Teoh, S.H. Fused deposition modeling of novel scaffold architectures for tissue engineering applications. *Biomaterials* **2002**, *23* (4), 1169–1185.
 27. Tan, W.; Desai, T.A. Layer-by-layer microfluidics for biomimetic three-dimensional structures. *Biomaterials* **2004**, *25* (7–8), 1355–1364.
 28. King, K.R.; Wang, C.C.J.; Kaazempur-Mofrad, M.R.; Vacanti, J.P.; Borenstein, J.P. Biodegradable microfluidics. *Adv. Mater.* **2004**, *16* (22), 2007–2012.
 29. Valerie, A.L.; Sangeeta, N.B. Three-dimensional photopatterning of hydrogels containing living cells. *Biomed. Microdev.* **2002**, *4*, 257–266.

FURTHER READING

- Curtis, A.; Wilkinson, C. Nanotechniques and approaches in biotechnology. *Trends Biotechnol.* **2001**, *19*, 97–101.
- LaVan, D.A.; McGuire, T.; Langer, R. Small-scale systems for in vivo drug delivery. *Nat. Biotechnol.* **2003**, *21* (10), 1184–1191.
- Lu, Y.; Chen, S.C. Micro and nano-fabrication of biodegradable polymers for drug delivery. *Adv. Drug Deliv. Rev.* **2004**, *56*, 1621–1633.
- Yeong, W.Y.; Chua, C.K.; Leong, K.F.; Chandrasekaran, M. Rapid prototyping in tissue engineering: challenges and potential. *Trends Biotechnol.* **2004**, *22* (12), 643–652.

Real Time Tomographic Reflection

Vikram S. Chib

Northwestern University, Chicago, Illinois, U.S.A.

George D. Stetten

Carnegie Mellon Robotics Institute, University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.

R

INTRODUCTION

Medical imaging modalities are important clinical and diagnostic tools in medicine. Ultrasound, computerized tomography (CT), and magnetic resonance imaging (MRI) provide images that can be used to localize structure and function in the human body; thus, they provide physicians the ability to determine patient abnormalities. However, conventional methods of displaying these imaging modalities are limited since they only allow for viewing images independent from the patient.

OVERVIEW

While this may not be an issue during non-invasive diagnostic imaging, it becomes one when these images are being used for guidance of surgical procedures. During these situations, the doctor is not able to use direct hand-eye coordination. Instead, the doctor must look at a display without obvious relation to the position and orientation of his surgical instruments in the operating field. This leads to difficulty in the surgeon's ability to utilize medical image guidance.

In order to allow surgeons a more natural feel during image-guided surgical procedures, there has been a recent push to develop augmented reality systems that seamlessly merge medical images with the direct vision of the surgeon. It is hoped that such systems will enable doctors to employ a more direct form of hand-eye coordination. In essence, medical augmented reality strives to give the surgeon a sort of X-ray vision in order to see directly inside the patient.

While many forms of augmented reality require a head-mounted or tracking apparatus, this article describes a particularly simple method called Real Time Tomographic Reflection (RTTR). A specific implementation of RTTR is called the sonic flashlight™. The sonic flashlight can project a virtual image of an ultrasound scan into its proper visual location within the patient and does not require any tracking or head-mounted apparatus. This system permits the operator to guide a needle through the skin by aiming directly at the ultrasound image and using natural

hand-eye coordination rather than looking away from the patient at a screen. It provides an intuitive merger of the visual exterior of the patient with the ultrasound image in situ, which can be shared by others standing next to the operator. The sonic flashlight may increase accuracy, safety, and speed for a wide variety of diagnostic and invasive procedures and enable them to be performed by an operator with less training.

REAL TIME TOMOGRAPHIC REFLECTION

Augmented Reality and Head-Mounted Displays

The main approach to fusing images with a direct view of the patient derives from research in augmented reality, so named to differentiate it from the more widely known virtual reality. After all, there is nothing virtual about real image data from real patients. Fuchs, Sauer, and others have developed augmented reality systems based on a head-mounted display (HMD), a type of device that generally replaces direct human vision with miniature video cameras and display screens mounted in front of each eye.^[1-5] The HMD permits the merger of a video view of the patient with image data, such as an ultrasound scan, that is rendered from an appropriate perspective. That perspective is determined by tracking both the HMD and the ultrasound transducer, which is generally held by the wearer of the HMD. Primarily because of their complexity, these systems have not been widely accepted. Significant problems remain unresolved, such as limited peripheral vision, resolution lower than human vision, and latency in the image registration. The weight, isolation, and need for a tether have also discouraged clinical use. Some of these problems will undoubtedly be addressed in time, but the HMD system remains inherently a complex and challenging approach.

Image Overlay

To reduce the apparatus that the operator must wear, DiGioia and his group at the Carnegie Mellon

Robotics Institute have developed a system called image overlay.^[6,7] They place the patient beneath a large, half-silvered mirror. Above the mirror, a flat panel monitor is mounted displaying 3-D renderings of computerized tomography (CT) data. The operator looks down through the mirror at the patient and sees the reflected CT rendering superimposed on the patient. The operator wears only a small, head-tracking optical transmitter, which is required to determine the proper perspective from which the CT data must be rendered for his/her particular viewpoint. A second tracking device must be attached to the patient in order to achieve proper registration between the patient and the CT data that was previously acquired. Special liquid crystal display (LCD) shutter glasses are needed if stereoscopic visualization is desired.

Real Time Tomographic Reflection (RTTR)

Real Time Tomographic Reflection (RTTR) represents a further simplification to DiGioia's image overlay. With restriction to a single tomographic slice (e.g., ultrasound) and by strategically positioning the transducer, half-silvered mirror, and display, RTTR eliminates the need for tracking the operator. This is possible because RTTR produces a virtual image. The word virtual is used here in its classical sense: the reflected image is optically indistinguishable from light-emitting voxels suspended in space. A virtual image is stable and independent of viewer location. Furthermore, it produces natural stereoscopic cues without requiring LCD shutter glasses. The operator does not wear any special apparatus, and several viewers can experience the same illusion simultaneously.

This approach is shown in Fig. 1. An ultrasound transducer acquires a tomographic slice representing a set of 3-D voxels in a plane. The image of that slice, which is displayed at its correct size on a flat-panel display, is reflected to occupy the same physical space within the patient. The patient is viewed through the half-silvered mirror. The reflected image is correctly located and independent of viewer location. Since the ultrasound transducer is an integral part of the apparatus, the merger is independent of patient location as well. The image will change during an invasive procedure in order to show the results of the procedure.^[8-14]

Masamune et al. have used a similar approach, which they call slice display, on CT data.^[15] In their first apparatus, the CT scanner was not an integral part; thus, independent registration of patient location was still required. They have since produced a real-time version in which the mirror and monitor are located within the CT scanner.^[16] Similar systems have been proposed previously, though, not for image-guided

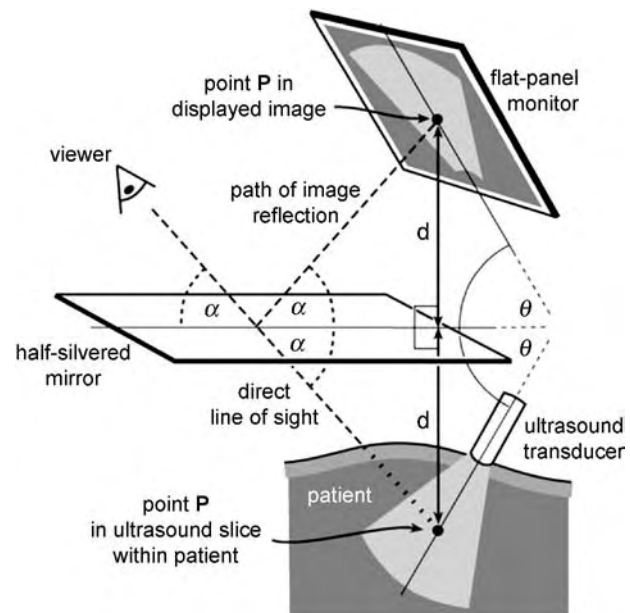


Fig. 1 Geometric relationships for Real Time Tomographic Reflection (RTTR). The mirror must bisect the angle between the slice and the monitor. On the monitor, the image must be correctly scaled, translated, and rotated so that each point in the image is paired with its corresponding point in the slice to define a line segment perpendicular to and bisected by the mirror. By fundamental laws of optics, the ultrasound image will appear at its physical location— independent of viewer position.

intervention, including an ultrasound-based system by Hofstein.^[17]

Figure 2 shows an implementation of RTTR in which a human hand is seen with an ultrasound transducer pressed against the soft tissue between the thumb and index finger. While not a common target for clinical ultrasound, the hand was chosen for this example because it clearly demonstrates successful alignment. The external surfaces of the hand are consistently located with structures in the ultrasound image. The photograph cannot convey the strong sense, which is derived from stereoscopic vision, that the reflected image is located within the hand. This sense is intensified by head motion because the image remains properly aligned from different viewpoints. To one experiencing the technique in person, ultrasound targets within the hand are clearly accessible to direct percutaneous injection, biopsy, or excision. The device was named the sonic flashlight because the transducer appears to illuminate the interior of the hand with the ultrasound data. A U.S. Patent (6599247) on RTTR was issued in 2004.

Geometric Transform for Calibration

Clearly, an accurate method for calibration is required. Without it, the sonic flashlight would be unable



Fig. 2 View of a hand through the half-silvered mirror of the sonic flashlight.

to guide interventional procedures.^[14] Calibration requires careful consideration of the degrees of freedom in the registration process. The challenge is to make each pixel in the virtual image occupy and, therefore, appear to emanate from its actual 3-D location in the slice. We first only consider the geometric transform of a rigid body, i.e., we assume that the ultrasound slice is displayed without distortion at its correct scale on a perfectly flat monitor. The geometric transform that would be required to superimpose the virtual image onto the slice can be represented as two sets of translations and rotations. Each of which has three degrees of freedom (Fig. 3). The first (two rotations and one translation) allows the

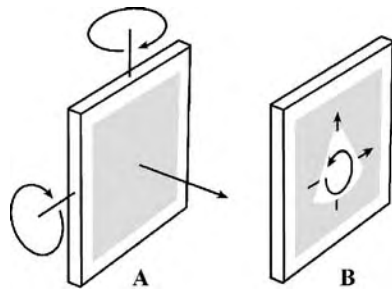


Fig. 3 Geometric transforms for (A) physically moving the display and (B) moving the image on the screen.

flat-panel display to be moved into its correct plane, which makes the virtual image coplanar with the actual ultrasound slice. We can achieve this by physically moving the display, the ultrasound transducer, or the mirror. The second (two translations and one rotation) can be achieved simply by adjusting the image of the ultrasound slice on the flat-panel monitor.

Affine Transform for Calibration

The assumption that both the ultrasound slice and the flat-panel monitor are truly flat is a safe one and makes a rigid-body (geometric) transform sufficient to move the monitor into the correct plane. However, correctly adjusting the image on the screen actually requires more degrees of freedom than those provided by the geometric transform—if only because the image must be scaled to its proper size. Extending the geometric transform to an affine transform allows us to scale the ultrasound slice to its correct size, adjust its aspect ratio, and correct for skewing. The calibration process becomes a matter of finding the optimal parameters for that affine transform. Mapping location (x,y) to (x',y') with an affine transform is accomplished by multiplying the homogeneous form of (x,y) by a 3×3 matrix \mathbf{A} .

$$\begin{bmatrix} x'_1 & x'_2 & x'_3 \\ y'_1 & y'_2 & y'_3 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ 1 & 1 & 1 \end{bmatrix}$$

An affine transform is capable of mapping any triangle onto any other triangle. If the apices of both triangles are known (and not colinear), we can find an explicit solution for the six unknown elements of the matrix \mathbf{A} . Calibration can thus be achieved by visually aligning three point targets with their appearances in the virtual ultrasound image.

CONCLUSION

Superimposing ultrasound images on human vision using Real Time Tomographic Reflection (RTTR) may improve an operator's ability to find targets while avoiding damage to neighboring structures and while generally facilitating interpretation of tomographic images by relating them spatially to external anatomy. As such, it holds promise for increasing accuracy, ease, and safety during percutaneous biopsy of suspected tumors, amniocentesis, fetal surgery, brain surgery, insertion of catheters, and many other interventional procedures.

ARTICLES OF FURTHER INTEREST

Magnetic Resonance Imaging; Medical Imaging, 3-D; Microcomputed Tomography and Its Applications; Positron Emission Tomography (PET)

REFERENCES

1. State, A.; Livingston, M.; Garret, W.; Hirota, G.; Whitton, M.; Pisano, E.; Fuchs, H. Technologies for Augmented Reality Systems: Realizing Ultrasound-Guided Needle Biopsies. In *Computer Graphics Proceedings, Annual Conference Series*, Proceedings of SIGGRAPH, New Orleans, LA, 1996; 439–446.
2. Fuchs, H.; State, A.; Pisano, E.; Garret, W.; Hirota, G.; Livingston, M.; Whitton, M.; Pizer, S. Towards Performing Ultrasound-Guided Needle Biopsies from Within a Head-Mounted Display. In *Visualization in Biomedical Computing*; Springer Verlag: Hamburg, Germany, 1996; 591–600.
3. Fuchs, H.; Livingston, M.; Raskar, R.; Colucci, D.; Keller, K.; State, A. Augmented Reality Visualization for Laparoscopic Surgery. In *Lecture Notes in Computer Science*, Proceedings of MICCAI, Cambridge, MA, 1998; 934–943.
4. Sauer, F.; Khamene, A.; Bascle, B.; Schimmang, L.; Wenzel, F.; Vogt, S. Augmented Reality Visualization of Ultrasound Images: System Description, Calibration, and Features. In *IEEE ACM International Symposium on Augmented Reality*; IEEE Computer Society: New York, 2001; 30–39.
5. Azuma, R.T. A survey of augmented reality. *Teleoperators and virtual environments*. **1997**, *6* (4), 355–385.
6. DiGioia, A.; Colgan, B.; Koerbel, N. Computer-Aided Surgery. In *Cybersurgery*; Satava, R., Ed.; Wiley: New York, 1998; 121–139.
7. Blackwell, M.; Morgan, F.; DiGioia, A. Augmented reality and its future in orthopaedics. *Clin. Orthop. Relat. Res.* **1998**, *345*, 111–122.
8. Stetten, G.; Chib, V.; Tamburo, R. System for Location-Merging Ultrasound Images with Human Vision. In *Applied Imagery Pattern Recognition (AIPR) Workshop*; IEEE Computer Society: Washington, DC, 2000; 200–205.
9. Stetten, G.; Chib, V. *Real Time Tomographic Reflection with Ultrasound: Stationary and Hand-Held Implementations*; Carnegie Mellon University, Nov. 2000. Robotics Institute Technical Report # CMU-RI-TR-000-28. http://www.ri.cmu.edu/pubs/pub_3412.html#text_ref.
10. Stetten, G.; Chib, V. Overlaying ultrasound images on direct vision. *J. Ultrasound Med.* **2001**, *20* (1), 235–240.
11. Stetten, G.; Chib, V. Magnified Real-Time Tomographic Reflection. In *Lecture Notes in Computer Science*, Proceedings of MICCAI, Springer Verlag, 2001; 683–690.
12. Stetten, G.; Shelton, D.; Chang, W.; Chib, V.; Tamburo, R.; Hildebrand, D.; Lobes, L.; Sumkin, J. Towards a Clinically Useful Sonic Flashlight. In *Proceedings of the IEEE International Symposium on Biomedical Imaging*; IEEE Computer Society: Washington, DC, 2002.
13. Chang, W.; Stetten, G.; Lobes, L.; Shelton, D.; Tamburo, R. Guidance of retrobulbar injection with real time tomographic reflection. *J. Ultrasound Med.* **2002**, *21*, 1131–1135.
14. Stetten, G.; Chib, V.; Hildebrand, D.; Burse, J. Real Time Tomographic Reflection: Phantoms for Calibration and Biopsy. In *IEEE ACM International Symposium on Augmented Reality*; IEEE Computer Society: New York, 2001; 11–18.
15. Masamune, K.; Masutani, Y.; Nakajima, S.; Sakuma, I.; Dohi, T.; Iseki, H.; Takakura, K. Three-Dimensional Slice Image Overlay System with Accurate Depth Perception for Surgery. In *Lecture Notes in Computer Science*, Proceedings of MICCAI, Pittsburgh, PA, 2000; 395–402.
16. Masamune, K.; Fichtinger, G.; Deguet, A.; Matsuka, D.; Taylor, R. An Image Overlay System with Enhanced Reality for Percutaneous Therapy Performed Inside CT Scanner. In *Lecture Notes in Computer Science*, Proceedings of MICCAI, Springer Verlag, 2002; 77–84.
17. Hofstein, S.R. Ultrasonic Scope. US Patent 956899, April 29, 1980.

Reconstructive Biomaterials

Mutaz B. Habal

Tampa Bay Craniofacial Center, Tampa, Florida, U.S.A.

R

INTRODUCTION

Today, the application and utility of biomaterials are essential for standard-of-care treatment of patients with extensive craniofacial trauma and reconstructive surgery. It is, therefore, imperative that surgeons using biomaterials have a comprehensive knowledge and understanding of the implications, applications, and effects of these materials when used in the living biologic systems. The biochemistry and composition vary from system to system, so that the details of the variables become as important as the mechanics of application. By definition, most biomaterials in use today are inert. They do not initiate any major inflammatory response or reaction and are easy for the practicing surgeon treating facial trauma.

In the past, controversy arose regarding the nature of the application of biomaterials in patients with acute trauma because of the high degree of contamination found in these traumatic wounds and limited the use to elective and clean wounds. The difference in the use of biomaterials in elective craniofacial surgery and its application in traumatic injuries are related to considerations of clean vs. clean/contaminated or grossly contaminated wounds of the facial skeleton. It was thought that the use of biomaterials in clean/contaminated (type II) or grossly contaminated (type III) wounds should be delayed, and immediate application limited to the clean conditions such as type I wounds. Other issues involving the use of biomaterials in facial surgery were related to open vs. closed wounds and the situation of internal contamination from the para nasal sinuses. These concerns compelled surgeons to defer the use of all biomaterials in craniofacial surgery related to primary trauma, and it had become a major principle that such biomaterials should only be used in the former type wounds rather than the latter two types. These concerns were alleviated as early as the 1970s when more relaxed views were applied owing to the understanding of the nature of the increased blood flow in the craniofacial region in contrast to that of other parts of the body. The presence of a good blood supply to the head and neck area may be the factor that differentiates trauma surgery in the craniofacial region from that in other regions of the body.

Biomaterials used today in facial trauma are divided into two categories: biomaterials used in immediate

trauma repair and those used for delayed reconstruction. Biomaterials are also used operatively for immediate skeletal fixation after reduction of the traumatized bone and placing them in their anatomic locations, as well as for delayed recreation and reconstruction of facial skeletal structures. In secondary reconstruction of traumatic defects, biomaterials are used as contour-fitting materials, and these procedures are most often deferred until after immediate surgical repair is done. Fixation of the craniofacial skeleton has undergone a major evolution in the past 30 years. However, our search continues for the ideal fixation device. The future of all skeletal fixation devices lies in the use of new biomaterials and new combinations of these materials to maximize the utility. In addition, the use of contour-fitting biomaterial will eventually follow the same process as in the past—that of the time element involved in the body's ability to incorporate these materials as a component that is not easily differentiated from adjacent tissue.

The last important concept that we need to go over in this entry is that of regeneration, as it affects the use of biomaterials in the craniofacial region. Multifunctional biomaterial implants have the ability to allow the surrounding bone to regenerate new bone, while using the applied material as a scaffold. A knowledge of the longevity of all the applied biomaterials is essential, as well as an understanding of the processes that are involved in total elimination by hydrolysis, incorporation, or regeneration.^[1,2]

SAFETY CONSIDERATIONS

The applications and the use of the biomaterials in any biologic system are always associated with a major question: Are those biomaterials safe to use or are they capable of causing harm? It is also of importance to know if there are any by-products that will be, under any circumstances, harmful. Safety issues are the fundamental focus of the Food and Drug Administration (FDA) and its European equivalents. These agencies study all biomaterials that are to be used for medical purposes, their efficacy, and the animal studies required before the premarket studies are designed to collect data from clinical applications. Data are then analyzed, and if the biomaterial is found to be safe and

efficacious, it is then released for marketing to doctors' offices and the hospitals for use in human patients. If there is any controversy regarding the applications clinically or any public concerns, the FDA will name a panel of experts to conduct a hearing before the application is allowed. All the biomaterials referred for discussion in this article have passed through this process and their final applications are dependent on the ability of the agencies to allow such a system to be used.

FIXATION OF THE FACIAL SKELETON

A review of the history of skeletal fixation is helpful to the understanding of the situations in which we work today. Skeletal fixation in the craniofacial region has gone through many advances in the past few decades. Most of these advances have followed major international conflicts involving the complex treatment of large numbers of casualties.

Initially, fractures of the craniofacial skeleton were treated without fixation by allowing the bones to heal in open soft tissue, then performing the repair at a later time. Fractures were also treated with closed reduction after manipulation. The next development was the use of an external apparatus for fixation. This method was useful until the external fixation was removed and the repaired structures collapsed again; these are referred to as relapsed segments.

Internal fixation then came into practice and required the use of rigid fixation. Use of the plating system began at the turn of the century, with the use of stainless steel plates. The use of vitalium and titanium plating followed. In the latter part of the last century, the use of resorbable plating systems evolved and has advanced to their present status today. Resorbable plating systems remain the state-of-the-art for skeletal fixation. The evolution of biocompatible resorbable polymers offers surgeons of today a new array of options for craniofacial skeletal fixation. Some of the potential benefits of resorbable polymers include greater ease and accuracy of contour adaptation, clear radiographic presentation owing to the absence of x-ray scatter, elimination of the need for secondary surgeries for device removal, and reduced risk of stress shielding of the underlying bone. Known as polyesters, these copolymers have chemical, physical, material, mechanical, and biologic properties different from those of metal fixation devices. Knowledge of these differences will facilitate the utilization of resorbable implants in fixation for craniofacial trauma.

Among the bioresorbable polyester facial fixation devices approved for clinical use by the FDA, copolymers of lactides and glycolides are available. The first copolymer of L-lactide and glycolide

(LactoSorb™, W. Lorenz, Jacksonville, Florida, U.S.A.) was approved by the FDA in 1996. The lactide in LactoSorb is a homopolymer of the levo form. The ratio of the L-lactide monomer to the glycolide monomer is 82:18 in poly(L-lactide-co-glycolide), to take advantage of glycolide's rapid degradation time. Strength declines to approximately 70% by six to nine weeks and resorption is complete by 12 months. The final resorption is dependant on the purity of the active form supplied by the manufacturer and is still slightly variable between the batches used.

Approved more recently in 1998 is a copolymer produced from a mixture of 70% L-lactide monomer and 30% D,L-lactide monomer (MacroPore®, MacroPore Biosurgery, Inc., San Diego, California, U.S.A.). This 70:30 ratio in poly(L-lactide-co-D,L-lactide) (PDLA) retains approximately 70% of its initial strength after nine months, and approximately 50% after 12 months, with resorption completed by 24–36 months. Additional resorbable polyesters from Bionx, Leibinger (Delta System and the New Delta System); Synthes (Resorbable System); KLS Martin (Resorb-X); and Inion (two systems) are all FDA approved and available for surgeons' to use. The differences among these systems are the ratios of the copolymers used in the compositions, which affect their longevity, a consideration of importance to surgeons. The deciding factor in which system to use is the individual surgeon's preference and ease in clinical application as well as the results obtained.

In view of these considerations, the primary focus in this article will be on the use of PDLA for skeletal repair and fixation, owing to the extensive clinical experience of the authors and to the commonalities this system shares with other systems that use similar materials in different compositions. The acronym PDLA, L-La/D,L-La copolymer, will also be utilized to designate a copolymer of the two monomers, L-lactide and D,L-lactide. In the interest of brevity, the copolymers are referred to as PDLA with the ratio given as well for accuracy.

PHYSICAL PROPERTIES

Poly(L-lactide), which has a high degree of crystallinity, is characterized by its strength and long degradation time. Conversely, a polymer created from D,L-lactide has little strength and degrades rapidly. Combining L-lactide and D-lactide polymers results in a copolymer with the intermediate characteristics of strength for 6–9 months and resorption in 24–36 months. In addition, the copolymer is optically clear and non-crystalline, resulting in minimal foreign body reactions by tissue. It is well noted that, even within a given copolymer, strength and degradation characteristics

can vary according to the degree of polymerization and purity within the copolymers. Therefore, the manufacturer must maintain this within the desired range. A common measure of the degree of polymerization is called intrinsic viscosity (IV) and, for any given polymer, the IV correlates with molecular weight. To measure IV, the polymer is dissolved in a standardized, known amount of chloroform and then passed through a viscometer. The length of time that it takes for passage is used to calculate the IV.

At sufficiently high temperatures, all materials change from hard to soft and finally to liquid. The temperature at which a material changes from hard to soft is known as the glass transition temperature (T_g). For 70:30 PDLLA, the T_g is 55°C (131°F), thus allowing heat to be utilized for contouring these implants.

Contouring an orbital floor liner illustrates this property. After making a template of the orbital floor, the template is held against the orbital floor liner (Fig. 1), then placed in a water bath and heated above T_g . The floor liner becomes soft in a few seconds and simply drapes over the template when lifted from the water bath. In a few more seconds, the floor liner cools below T_g and can be removed from the template. The liner is then ready to be placed in the patient. It is useful to note that 70:30 poly (L-lactide-co-D,L-lactide) has shape memory, and if placed back in the water bath it will return to its original contour thus enabling additional opportunities to recontour it. If only a portion of an implant needs to be recontoured, only that portion needs to be placed back in the water bath. Cyclic heating of 70:30 PDLLA to 70°C can be performed multiple times with no change in material strength or IV (Fig. 2).

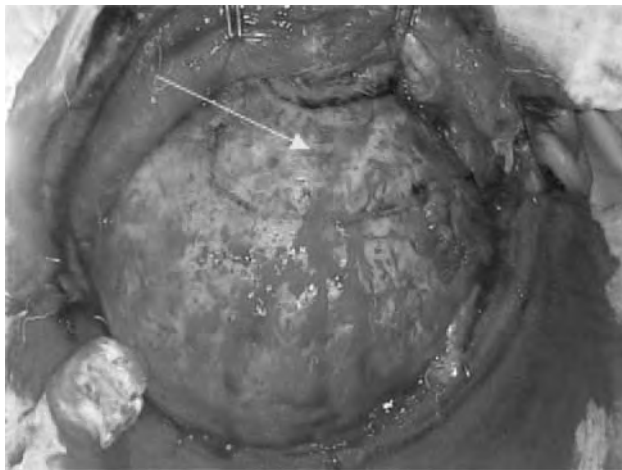


Fig. 1 Skull defect.

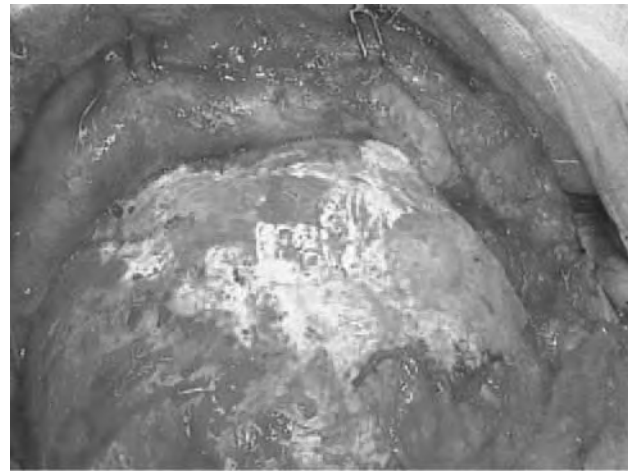


Fig. 2 Defect repaired with the new biomaterials tricalcium phosphate.

R

CHEMICAL PROPERTIES

When lactic acid undergoes polymerization, ester bonds are formed and H_2O is released. Therefore, lactide copolymers are also known as polyesters. Resorption of lactide copolymers takes place as a reversal of this process, with sorption of H_2O and scission of the ester linkages. This bulk hydrolysis of lactide copolymer implants continues until single lactic acid molecules are released, which are then metabolized into glucose or into CO_2 and H_2O via the Krebs tricarboxylic acid cycle.

A variety of factors are known to affect the rate of lactide copolymer resorption, such as temperature and the internal chemical environment. A higher IV or molecular weight means that there are more ester linkages that undergo scission, and this process results in a longer resorption time. A larger implant size or volume will also require more scission before implant resorption can be completed. If the polymer is packed more tightly in an orderly crystalline pattern, there is less space for H_2O access, and resorption will take longer than for non-crystalline implants. As hydrolysis occurs both on the implant surface and within its interior, implant-increased porosity will increase surface area, facilitate H_2O access, and decrease resorption time. The molecular configuration of copolymers may alter resorption time. Greater vascularity of the implant host site as well as flexural bending from functional loading appears to be associated with an increased rate of hydrolysis.

TOXICOLOGY

The toxicology of lactides has been of minimal concern, owing to the relatively small volumes of implant

material, slow degradation rates, and short serum half-lives. The serum half-life of the levo form is 15 min; for the D,L form it is 22 min. The normal resting blood lactate level is 1.1–2.8 mM/L. After muscular activity, it will rise 10-fold to 10–23 mM/L. If the assumption is made that degradation occurs over two months, with first order kinetics and a half-life of 74 hr, a 100 g implant would release 0.18 mM/L of lactide acid in the first minute, far less than the changes resulting from muscular activity. Two of the largest sheets of 70:30 L-La/D,L-La copolymer weigh only 18 g, and degradation actually takes place over a much longer 18- to 36-month time interval. Even with first order kinetics starting instantly, the 18 g of lactide copolymer would result in an increase in blood lactate levels of only 1.1%.

HISTOLOGY

The histological responses to 70:30 L-La/D,L-La copolymer have been well studied. There is an initial acute inflammatory response following implantation. By 72 hr, there is a narrow zone of fibrinous exudate, edematous granulation tissue, and a modest degree of fibroblast proliferation. By 7–14 days, the granulation tissue has matured into a thin, cellular, fibrovascular capsule.^[3,4] Measurements of *in vivo* tissue pH adjacent to 70:30 L-La/D,L-La copolymer implants have detected no change during degradation.^[5,6]

MECHANICAL PROPERTIES

Comparing some of the mechanical properties of 70:30 L-La/D,L-La copolymer, bone, and steel, the tensile strength of lactide is approximately 30% of the strength of bone. With a tensile strength of 70 MPa, lactide materials can readily be designed to accommodate the failure loads for non-weight-bearing bones. When designed as 1.0 mm thick plates, the tensile strength is approximately 190 N or 45 lbs. When metal screws are over torque, the threads strip the bone. When lactide screws are overtorqued, the heads shear off. The shear strength of 2.0-mm screws is approximately 85 N or 20 lbs. As the 70:30 L-La/D,L-La copolymer undergoes hydrolysis, its mechanical strength will decrease. At three months, strength remains near 100%, decreasing to 90% at six months, 70% at nine months, 50% at 12 months, and 0% by 18 months.

PLATE SELECTION FOR FRACTURE FIXATION

As with all fixation systems, clinical experience eventually determines the choice of implant design that

is most likely to produce a successful outcome. As a starting point, it is recommended that the surgeon select a copolymer design one size larger than the titanium system currently employed. An example from the authors' experience is the substitution of titanium with polylactide (PLA) for genioplasty, when the anterior part of the mandible needs configuration change. A traditional metal plate with two holes on each side and 2.0 mm screws would normally have been used; however, this patient insisted on resorbable over metal fixation. To ensure adequate stability, the 70:30 L-La/D,L-La copolymer plate was contoured from a piece of 1.0 mm thick mesh with four holes on each side and attached with 2.4 mm screws. While this design may be excessive in strength, it is appropriate to be conservative until clinical experience is acquired (Fig. 3).

CLINICAL EXPERIENCE

The following comprises a review of the authors' clinical experience with biomaterial use in craniofacial fixation.

Resorbable Fixation

Zygomatic and midface fractures

The natural extension of surgical experience with metal plates is to use resorbable plates of similar designs, such as in malar fracture fixation. The principles of exposure and reduction of fractures remain unchanged when resorbable plates are employed. The holes for lactide screws require tapping and maximal screw insertion torque. The plates can be precontoured by dipping in a hot water bath and, after inserting one or two screws, be further contoured by using a flat heat applier (Fig. 3A), a gauze sponge dipped in the warm water bath, or by dripping warm water onto the plate. Resorbable plates are easily trimmed *in situ* using a cautery tip attachment (Fig. 3B). Resorbable mesh may be used in place of metal mesh, such as in an orbital floor fracture (Fig. 3C). Resorbable mesh is easily trimmed with scissors. It is helpful to leave the leading edge of the mesh uncut, so that it can be inserted with less likelihood of engaging adjacent soft tissues. Fragments of bone can be incorporated into either resorbable plates or meshes.

With increased experience, it became apparent to the authors that resorbable mesh offered more sites for fixation and provided greater strength when contoured in three dimensions. By allowing custom shapes to be cut with multiple mesh spaces for screw or tack fixation, the surgeon is freed from the geometric restrictions of limited selection for metal plate designs.

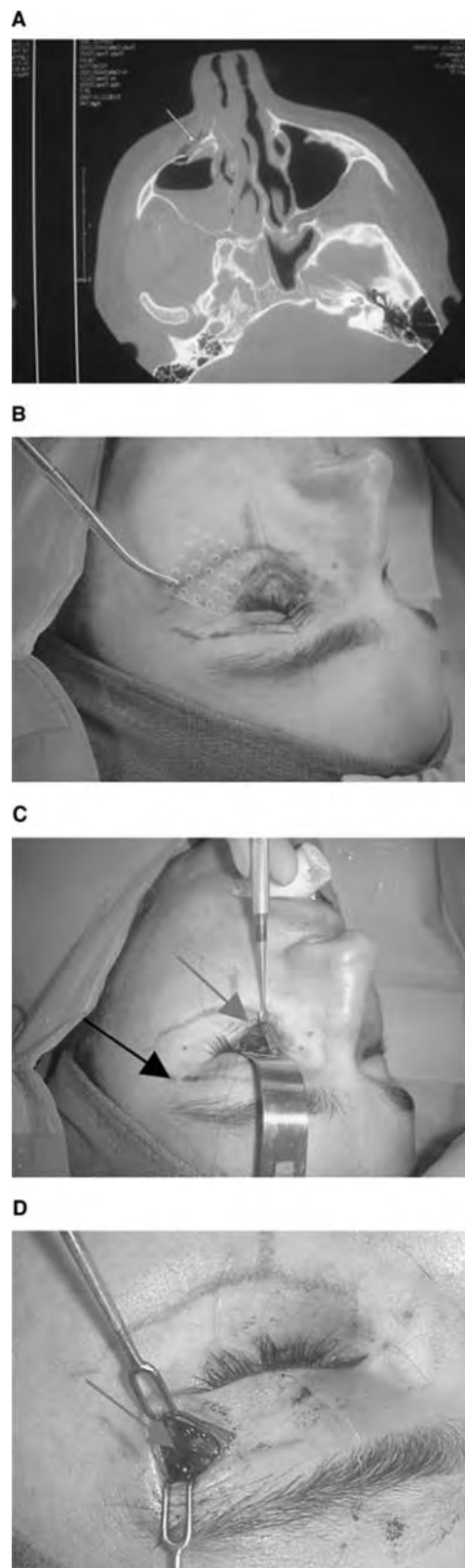


Fig. 3 Orbital fracture treated with bioresorbable plating system, a new biomaterial PLA. (A) Patient with a fractured orbit; (B) PLA panel shaped to inserted in the orbit via transconjunctival; (C) PLA inside the orbit; and (D) PLA fixation transpalpebral.

Combined with the speed and ease of adapting mesh to three-dimensional contours using warm water, it is not surprising that the use of resorbable mesh with screw or tack fixation has become even more prevalent than resorbable plates in our craniofacial and mid-facial reconstructions.

R

Fronto-orbital fractures

The properties of resorption and the ease of contouring have enabled complex fixation requirements of fronto-orbital fractures to be readily accomplished. By using the resorbable mesh as a template, accuracy of bone fragment fixation can be facilitated. In a two-year-old girl kicked in the forehead by a horse, the neurosurgeon outlined the planned craniotomy. Prior to making the bone cuts, a sheet of resorbable mesh was contoured over the fracture site to match the normal right forehead. The fragments were then reduced into this template on the back table, fixated with tacks, and then returned to the patient after completion of the neurosurgical repair. Additional tacks placed around the mesh periphery secured the reduction, restoring this girl's forehead symmetry.

When trauma results in boney defects, easily contoured resorbable mesh can be shaped into a template for the bone graft reconstruction. In a 11-year-old girl, anterior cranial base fractures were sustained in a motor vehicle accident. Following acute debridement she was transferred for definitive care of her cerebrospinal fluid leaks and fronto-orbital bone defects. To provide neurosurgical access to the anterior cranial base, the remaining frontal bone flap and orbital rim were removed. Contoured resorbable meshes were then used to rebuild the fronto-orbital construct from split cranial bone grafts from the flap, along with the portions of the orbital rim that had been removed. A nice fronto-orbital contour was achieved after the construct was fixated around its periphery.

Fixation of secondary craniofacial trauma during reconstruction

Perhaps no area of secondary craniofacial trauma reconstruction has been as challenging as the cranio-orbital fractures of the complex curvatures and multiple planes involved in the frontal bone, the orbital rims, and orbital roof. The availability of easily contoured, large resorbable sheets makes it possible to form templates using in situ anatomy for ex vivo back table assembly of the bone graft reconstruction. In a 59-year-old man, kicked in the right fronto-orbital area by a horse, osseous debridement was carried out at the time of initial neurosurgical repair. Three months later, he underwent reconstruction of his defect, which

included absence of the frontal bone; supraorbital, lateral, and medial orbital rims; and orbital roof. A full thickness bone flap from the right parietal skull was split, and the inner table was replaced and fixated with a resorbable mesh that was contoured before the craniotomy. A template of the desired cranio-orbital defect was formed using a sheet of 0.75 mm thick resorbable mesh, and the cranial bone graft was then cut to fit into the mesh template and fixated with tacks. The construct was fitted into the defect and secured around the periphery with additional tacks. Three months later, the cranio-orbital contour was excellent, the patient's preoperative diplopia completely resolved, and a ptosis correction was scheduled to complete the reconstruction.

EVOLUTION

Our experience is primarily with a pediatric craniofacial population in a practice that stresses treatment of the traumatic injuries of the craniofacial skeleton. However, our application of the techniques of fixation is similar to controlled osteotomies, non-controlled osteotomies, as well as multiple craniofacial fractures. Over time, our initial utilization of preformed plate design has been replaced by the use of sheets of mesh, which are easily cut to fit the requirements of repair for a given fracture. This choice is a matter of preference. With complex reconstructions, we have found that easily contoured resorbable mesh permits quick creation of templates that guide bone graft assembly to a precise and accurate reconstruction. We anticipate that techniques will continue to evolve, as more experience is gained with the properties of resorbable copolymers used for craniofacial fracture fixation and reconstruction.

OTHER BIOMATERIALS UTILITY

The above descriptions of biomaterials reflect those that are used most often in clinical practice. Today there are many surgeons who prefer the use of metallic implants, particularly titanium, for the correction of the facial fractures. Two points, however, must be stressed. First, in children, it is the standard-of-care to use bioresorbable plates and screws to mechanically stabilize the craniofacial skeleton, and it is impermissible to do otherwise. These patients have a higher possibility of screw migration, growth disturbances, and a need for unnecessary secondary operations. Second, for the adult population, repeated surgeries may also be avoided with the use of bioresorbable plating systems.

Other biomaterials are used to a lesser extent in the patients with craniofacial trauma for immediate repair and for reconstruction. However, their use is mostly limited to delayed reconstruction and, very rarely, such biomaterials are applied in the primary care of patients with multiple facial injuries. Fixation of the skeleton in the correction of facial fractures is the primary use for bioresorbable copolymers. When a defect in the craniofacial areas must to be reconstructed, the use of bone grafting is the treatment of choice for those patients. However, if that treatment is not an option or will precipitate a major disability for the patient, the surgeon should be knowledgeable in the applications of those biomaterials in use and familiar with the indications and contraindications of all such biomaterials.

Poly-Methyl Methacrylate

The use of methyl methacrylate was popular after World War II. However, analyzing that experience, we have learned that such biomaterial should not be used in type II wounds or in any contaminated areas of the craniofacial skeleton. Infection and fistulization often result, even when the biomaterial utilized for the repair has been present for over 30 years. The utility of the application of methylmethacrylate is in covering large defects in the skull. It is also used to protect the brain when skull defects are present. The reported incidence of infection varies owing to the longevity of use and because of the transient nature of the population in this country. The lack of standardized records, continuity of care, and information reporting contribute to the inaccuracy of existing data. Methyl-methacrylate is also used to correct contour abnormalities in the patients with trauma. The same principles apply for the avoidance of use in contaminated areas and type II and III wounds. This biomaterial application should always follow a list of strict indications and contraindications to achieve the best possible outcome for trauma patients. The last aspect that should be noted is that reconstruction of large defects in the skull may achieve the immediate purpose for the desired result, but the patient will end up with a shatter effect if the area is subjected to other blunt trauma. Reinforcement with metallic mesh did not alleviate that issue.

Silicones

Silicone rubber in the form of room-temperature vulcanizing silicone or prefabricated silicone components was used in skull repair, but was found to have no added value. Use of silicone was very popular in the early 1970s, and served in reconstruction to fill defects in areas of traumatic injury or for contour

abnormalities. The applications for silicone today are limited. This material is capable of causing major complications in dynamic sites owing to its ability to erode the underlying bone.

Polyether-Urethan Terephthalate and Med Pore

Both materials are used in a fashion similar to silicone but have less utility in trauma patients. We have extensive experience in the use of the polyurethane mesh as a contour and enhancer for bone regeneration for large defects in the skull. However, large sheets of bioresorbable biomaterial have replaced its use by obviating the need for a secondary surgery to remove the implant.^[7]

Bone Substitutes

To complete this discussion, we must also note the various bone substitutes that are in use today and available in most operating rooms. Bone substitutes react with the surrounding tissues to make the application more appropriate. Their main use is for filling defects that need a scaffold. These biomaterials do not stand sheer force, even after they consolidate, and are mechanically unstable. A second, more useful, application is for contour abnormalities.

We are not going to put a recommendation on any one of these biomaterials, but we will list them in a systematic way that will help those interested to know the difference between each component.

Biomaterials as bone substitutes

1. Osteo Inductive: Demineralized bone implants vary from 17% to 99%. Use should be species specific, so that there is no crossing of species boundaries (Dyna-graft 17%-).
2. Osteo Conductive: Calcium sulfate, calcium phosphate, and calcium carbonate combinations (Norian, Mimex, Bonepaste).
3. Bioactive and Interactive: Composition is changed when in contact with the biological system (Bioglas).
4. Tissue Engineering: Active cells and active scaffold will be needed in this unit of tissue repair. Scaffolds used vary from hydroxy appetite to collagen

to most of the degradable carriers. The stem cells used must be autologous in nature. Such uses in the trauma patients are under study today and have limited applications.

R

ARTICLES OF FURTHER INTEREST

Collagen; Collagen Fixation; Tissue Engineering

REFERENCES

1. Habal, M.B. Absorbable plating system-discussion of, resorbable plate fixation in pediatric craniofacial surgery. *Plast. Reconstr. Surg.* **1997**, *100*, 8–11.
2. Habal, M.B.; Pieterzak, W.S. Key points in the fixation of the craniofacial skeleton with absorbable biomaterial. *J. Craniofac. Surg.* **1999**, *10*, 491–499.
3. Kulkarni, R.K.; Pani, K.C.; Neuman, C.; Leonard, F. Polylactic acid for surgical implants. *Arch. Surg.* **1966**, *93*, 839–843.
4. Gogolewski, S.; Jovanovic, M.; Perren, S.M.; Dillon, J.G.; Hughes, M.K. Tissue response and in vivo degradation of selected polyhydroxyacids: polylactides, poly(3-hydroxybutyrate), and poly(3-hydroxybutyrate-co-3-hydroxyvalerate). *J. Biomed. Mater. Res.* **1993**, *27*, 1135–1148.
5. Van der Elst, M.; Dijkema, A.R.; Klein, C.P.; Patka, P.; Haarman, H.J. Tissue reaction on PLLA versus stainless steel interlocking nails for fracture fixation: an animal study. *Biomaterials* **1995**, *16*, 103–106.
6. Mainil-Varlet, P.; Rahn, B.; Gogolewski, S. Long-term in vivo degradation and bone reaction to various polylactides. *Biomaterials* **1997**, *18*, 257–266.
7. Habal, M.B. Absorbable, invisible and flexible plating system for the craniofacial skeleton. *J. Craniofac. Surg.* **1997**, *8*, 121–126.

BIBLIOGRAPHY

- Habal, M.B.; Reddi, A.H. *Bone Grafts and Bone Substitutes*; WB Saunders: New York, 1992.
- Tharanon, W.; Sinn, D.P.; Hobar, P.C.; Sklar, F.H.; Salomon, J. Surgical outcomes using bioabsorbable plating systems in pediatric craniofacial surgery. *J. Craniofac. Surg.* **1998**, *9*, 441–444.

Rotary Cell Culture Systems

Stephen S. Navran

Synthecon, Inc., Houston, Texas, U.S.A.

INTRODUCTION

The original concept of the rotary cell culture system (RCCS) was developed at NASA's Johnson Space Center in 1996. It was conceived as a means of carrying cultured cells into space on Space Shuttle flights to study the effects of microgravity at the cellular level. At the time, it was becoming apparent that microgravity had a variety of negative physiological consequences for the astronauts as the duration of space flights increased. By studying the effects of microgravity on cultured cells aboard the Space Shuttle, NASA hoped to gain insight into the molecular mechanisms of the effects of microgravity on the astronauts, and to devise countermeasures.

The designers of the original RCCS (often referred to as the Rotating Wall Vessel, RWV, U.S. Patent 5,026,650)^[1] envisioned the RWV as a device to protect delicate cell cultures from the mechanical forces generated during launch and landing of the Space Shuttle. To accomplish this task, the designers attempted to create an artificial microgravity-like environment. They placed cells in a cylindrical, zero-headspace culture vessel and rotated the vessel at speed sufficient to maintain the cells in a suspended state, in the absence of large shear forces such as those found in spinner flasks. By randomizing the gravitational vector, cells can be cultured in a near free fall environment, simulating, in some respects, the microgravity environment of space. Oxygenation was accomplished by adding a coaxial tubular core covered with a silicon membrane, through which an oxygen/carbon dioxide mixture was pumped to allow bubble-free diffusion of gas into the culture media.

During initial ground-based testing of the bioreactor at NASA, it was noted that when dispersed anchorage-dependent cells were introduced into the RCCS, cells were observed to aggregate and form larger, three-dimensional structures resembling tissues.^[2] It was at this stage that NASA investigators recognized that the RCCS could be an entirely new approach to cell culture. Since 1992, NASA has devoted significant resources to promoting ground-based applications of the RCCS. NASA currently funds approximately 40 investigator-initiated projects per year based on the RCCS. This article will briefly summarize the theoretical background of the RCCS, which has

recently been reviewed in depth;^[3] describe the operation of the bioreactor; and provide specific examples of cell culture applications performed in the RCCS.

THEORETICAL BACKGROUND

In the first analysis of particle motion in the RCCS done at NASA, the simplifying assumption was made that cells cultured in the device would behave as small, spherical particles.^[4] As we will see below, cells often aggregate to form larger, tissuelike structures that may deviate from the theoretical ideal. Nevertheless, this analysis is a reasonable starting point for understanding the principles of cell culture in the rotating bioreactor vessel.

In designing the bioreactor, the overall goal was to minimize the shear forces that were known to cause cellular damage in conventional stirred or sparged suspension bioreactors.^[5,6] The RCCS was designed with several features intended to minimize fluid shear.^[3]

1. The fluid flow within the vessel is near-laminar, or solid-body, under most conditions, preventing the shear produced by turbulent flow.
2. The culture medium is mixed by the rotation of the vessel, avoiding the need for mechanical stirring.
3. The vessel is completely filled with media to prevent bubble formation.
4. A coaxial silicon membrane core, rotating at the same speed as the vessel, is used to oxygenate the media.

The primary determinants of fluid flows producing shear across the surface of cultured cells are gravitational, centrifugal, and Coriolis forces.^[4] Figure 1 shows a particle, P, rotating in a fluid medium at an angular velocity, ω . The gravity-induced sedimentation vector, V_s , is resolved into radial and tangential components, V_{sr} and V_{st} . The Coriolis force, V_{ct} , is tangential to the particle, whereas the centrifugal force, V_{cr} , is radial and outwardly directed. As the particle is rotated in the RCCS, it accelerates until it reaches its terminal velocity, V_s , where the force of gravity is opposed by centrifugal, Coriolis, and shear forces. The Stokes equation for a particle falling in a fluid

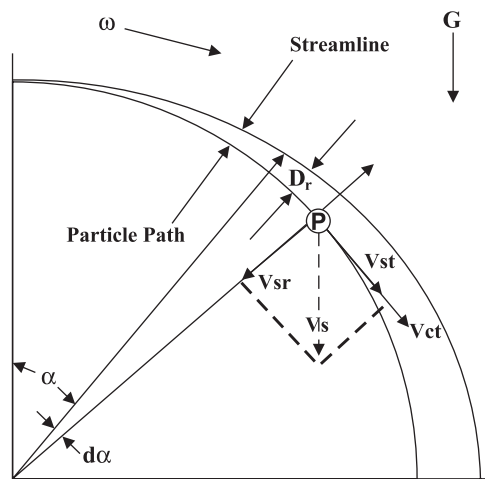


Fig. 1 Vector velocity diagram. The forces acting on a particle rotating in a fluid are shown. Gravity-induced sedimentation V_s is resolved into radial V_{sr} and tangential V_{st} components. Centrifugal force results in an outwardly directed particle velocity, V_{cr} . The Coriolis motion V_{ct} is tangential to the particle.

medium is

$$V_s = \frac{2\Delta dgr^2}{9\nu}$$

where V_s is the terminal velocity of the particle, Δd is the difference in density between the particle and the culture medium, g is the acceleration of gravity, r is the radius of the particle and ν is the viscosity of the medium. Therefore, the major determinants of the sedimentation rate, which is directly proportional to shear force, are gravity and the radius of the particle.

In practice, the rotational speed of the bioreactor vessel is set so that the sedimentation rate of the particle in the cell culture medium is minimized. The viscosity and density of the cell culture medium are not usually varied. As the particles increase in size, the rotational speed must be increased to keep the particles in suspension. Centrifugal force is then increased, causing particle impacts on the vessel wall. The increased rotation also causes an increase in shear. These two factors produce cell damage and effectively limit the size of cell aggregates to less than 1 cm in diameter.^[7] These limitations notwithstanding, the RCCS is a unique tool for constructing macroscopic, three-dimensional tissuelike aggregates that have proven highly useful in a variety of research applications, as detailed below.

Another important factor to be considered is mass transport of nutrients. In other types of suspension culture systems, such as spinner flasks or bubble sparged vessels, the extreme turbulence created to keep the cells in suspension ensures that the cells obtain

adequate nutrients and oxygen mostly by convective bulk flow, but at the expense of cellular injury from the fluid shear forces and bubbles. In true microgravity, where the terminal velocity of the cells is near zero, mass transport is accomplished primarily by diffusion. In ground-based applications of the RCCS, mass transport is accomplished by a combination of bulk flow and diffusion, without the damaging influences of fluid shear stress.

THE RCCS IN CELL BIOLOGY AND TISSUE ENGINEERING

The original purpose of the RCCS was to maintain cell cultures on the Space Shuttle before and during lift off. Once in orbit, the culture chamber served only as a containment vessel. Later, as the utility of the technology became apparent in ground-based experiments, the RCCS began to be thought of as a “microgravity simulator” that could be used to study the effects of microgravity on human physiology, as it impacted the astronauts. The immunosuppression experienced by astronauts on the Space Shuttle is an example. Studies on the ground using the RCCS have shown deficits in T-lymphocyte function similar to those found in the astronauts when subjected to simulated microgravity.^[8,9] More recent studies have used the bioreactor to devise a potential countermeasure for this immune dysfunction.^[10,11] However, a study comparing gene expression in the RCCS at normal gravity with that in the true microgravity of space demonstrated that there are considerable differences,^[12] suggesting that the RCCS may not be a good model of microgravity.

Most of the applications for the RCCS involve the ability of the bioreactor to assemble and maintain three-dimensional cell aggregates that closely resemble native tissue. When considering the utility of the RCCS as a tool to generate in-vitro models, an important question is whether the 3-D constructs produced are sufficiently differentiated. While the answer to this question is not complete, there are a number of studies which suggest that 3-D tissue aggregates grown in the RCCS are more differentiated than 2-D cultures.^[13-17] In general, these studies have dealt with a limited number of markers of differentiation. Ideally, the 3-D tissue constructs made in the RCCS should be compared to the native tissue. Such a study has recently been performed by Nakamura et al.,^[17] comparing a 3-D carcinoma model grown in the RCCS with actual tumor tissue, as well as with 2-D cultured tumor cells. The results showed that the 3-D construct was very similar to tumor tissue in those markers that were examined. The limitation of all the aforementioned studies is that they looked at only a few markers at

most. With the emergence of genomic and proteomic tools to study global gene and protein expression, it is now possible to compare large numbers of markers simultaneously. These technologies will undoubtedly be applied to the question of the degree of differentiation of 3-D culture in the RCCS.

The fact that the RCCS is capable of producing differentiated 3-D tissuelike constructs suggests that this technology will have widespread use in tissue engineering, including replacements for injured organs, drug development, and toxicity testing. A number of investigators have undertaken studies aimed at developing such applications (see examples in the following sections).

Methods of Cell Culture in the RCCS

The RCCS bioreactors were designed in two basic formats. One is the Slow Turning Lateral Vessel (STLV), illustrated in Fig. 2. This vessel has a cylindrical shape with a central core covered by a silicon membrane. The unit is placed in a humidified incubator with the appropriate gas mixture. The incubator gas is pumped through the core and diffuses through the silicon membrane without producing cell-damaging bubbles. Cells can be introduced into the vessel through a large fill port on one end cap. There are also two luer lock syringe ports on the end cap, through which bubbles removed and samples withdrawn. When feeding is required, the vessel is placed on its end under a laminar flow hood, and the cells or cell aggregates are allowed to settle to the bottom. The media can then be removed through the large fill port without disturbing the cells, and fresh media can be added.

The other basic RCCS bioreactor format is the High Aspect Ratio Vessel (HARV), as shown in Fig. 3. As its name implies, the HARV is a disk-shaped vessel.

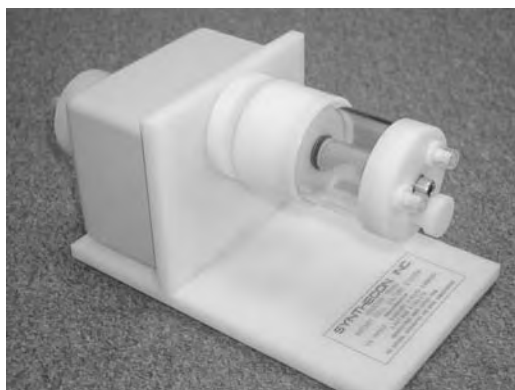


Fig. 2 The STLV. The vessel is cylindrical and has a central core covered with a silicon membrane. The gas mixture is pumped through the core and diffuses into the vessel.



Fig. 3 The HARV. The vessel is disc-shaped and has a silicon membrane on the back. The gas mixture is pumped into the back of the vessel.

The gas-permeable silicon membrane is on the back of the vessel, and the fill port and syringe ports are located on the front face. The choice of which vessel to use is often not obvious, depending on unknown characteristics of the cell type being cultured. One general principal that seems to hold true is that non-adherent cells seem to grow better in the HARV.

In addition to the two basic configurations, the RCCS was also designed as a perfusion bioreactor (Fig. 4). With a rotating coupling placed on each end cap, the vessel can be continuously perfused with media. This type of reactor requires a filter to be placed on the effluent side of the vessel to prevent the cells from passing into the flow loop.

The RCCS can culture both suspension-dependent and anchorage-dependent cells, as well as tissue explants. Suspension cells that do not aggregate can be grown in the RCCS at very low rotational speeds

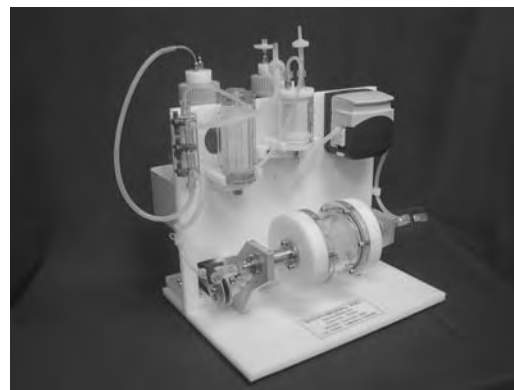


Fig. 4 The perfused RCCS. The medium is circulated through the vessel from a feed bottle. The media can be perfused into a collection bottle or recirculated. Gas exchange occurs in a section of coiled silicon tubing in the flow loop.

(<10 rpm), because their sedimentation velocity is minimal compared to three-dimensional cell aggregates. This type of culture allows for the lowest shear force possible in ground-based rotary bioreactors. Anchorage-dependent cells may be cultured in two ways. One way is to introduce them into the bioreactor vessel as individual, dispersed cells. During the initial phase of culture, some cell types may readily adhere to each other, forming three-dimensional aggregates. In other cases, substrates such as microcarriers or biocompatible polymers are required for initial cell attachment. Once attached, the cells proliferate and often adhere to cells on other microcarriers, forming large, three-dimensional aggregates of microcarriers and cells. As the aggregates grow and become visible to the naked eye, the rotational speed must be adjusted to maintain the aggregates in suspension. There are no absolute guidelines for determining whether anchorage-dependent cells will self-assemble or will require a substrate. The behavior of each cell type must be determined empirically. However, increasing the initial seeding density tends to favor self-assembly, due to the increased probability of cell-to-cell contacts.

More recently, the RCCS has found a new use: the culture of tissue explants. When explants are cultured in two-dimensional systems such as flasks or dishes, the integrity of the tissue is not maintained. Some cells migrate out of the explant onto the surface and proliferate, while most remain within the solid tissue mass and readily die off. The result is not representative of the original cell population. In the RCCS, however, small explants up to 2–3 mm in thickness can survive intact for extended periods of time, up to months in some cases (see Refs. [23] and [27]). Furthermore, new growth can often be observed. Explant culture has been shown to provide an excellent model for the study of complex cell interactions while maintaining the normal extracellular environment.

Coculture of multiple cell types is another area where the RCCS offers some unique advantages. Dispersed cells of two or more types can be introduced into the RCCS and assembled into a three-dimensional aggregate that closely approximates the native tissue cytoarchitecture. An example is shown in Fig. 5. In this study, done in collaboration with Dr. Arun Rajan at Baylor College of Medicine, two cell lines derived from pancreatic islets of Langerhans were assembled into artificial islets in the RCCS. The cell lines, alpha-TC6 and HIT (alpha and beta cells, respectively), make up the majority of the normal islet. When seeded into the RCCS, the cells assemble into spherical aggregates 200–300 microns in diameter and then remodel, so that the central core of the aggregate contains the beta cells surrounded by a layer of alpha cells, as in the normal islet. Functional studies have shown that these isletlike structures were capable of secreting insulin and

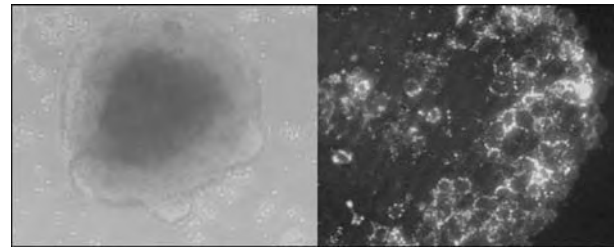


Fig. 5 Artificial islet assembled in the RCCS. A beta cell line, HIT, and an alpha cell line, alpha TC 6, were introduced simultaneously into the RCCS. After 2–3 days aggregates of 200–300 microns were assembled (left). Immunostaining (left) of these aggregates showed that the beta cells were in the center of the aggregates (bright staining) while the alpha cells were on the periphery (fainter staining).

glucagon when stimulated with the appropriate secretagogues.^[18]

Selected Applications

For a cell culture technology with such a short history, the RCCS has proven remarkably versatile. The following examples, in addition to those already mentioned, are chosen to illustrate the diverse areas in which the RCCS has been used. A complete bibliography is available at www.synthecon.com.

Cartilage

Chondrocytes, when cultured in two dimensional monolayers, dedifferentiate into a fibroblastlike phenotype and lose the ability to express cartilage-specific extracellular matrix proteins. However, when these cells are transferred to the RCCS, they spontaneously aggregate and re-express cartilage-specific proteins.^[19] Other investigators^[20,21] have used artificial polymer scaffolding such as polyglycolic acid to grow chondrocytes. Compared to static culture or spinner flasks, the RCCS produced constructs with more rapid tissue growth and higher levels of glycosaminoglycans and total collagen. The mechanical properties of the RCCS-cultured cartilage constructs more closely resembled the mechanical properties of native cartilage than did those of the other culture systems.

Cancer

Three-dimensional cultures of carcinomas in the RCCS have been shown to be superior models for the study of prostate cancer and have been used in the development

of therapies.^[22] Explants of human melanoma cultured in the RCCS maintain their structure and function for weeks, providing a unique model of human melanoma that mimics the *in vivo* environment more closely than conventional culture methods.^[23] Cocultures of cervical cancer and endothelial cells have been grown in the RCCS to study tumor angiogenesis.^[24]

Insect cells

Insect cells grown in the RCCS were maintained in the stationary phase two to three times longer than cells grown in shaker flasks.^[25] Cell death was reduced 20- to 90-fold in the RCCS. The differences were attributed to the reduction of hydrodynamic shear forces in the RCCS compared to the shaker flask.

Infectious disease (bacterial and viral)

A 3-D model of intestinal epithelium was developed in the RCCS to study the pathogenesis of *Salmonella*.^[26] Compared to static, 2-D culture, the RCCS produced a more differentiated tissue model. *Salmonella* was observed to infect the 3-D culture in a different way than it infected monolayer cells. There were marked differences in cytokine expression in the two models. The results showed that the RCCS provided a better means of studying microbial infectivity from the point of view of host–pathogen interaction.

Human lymphoid tissue explants (tonsil) maintained in the RCCS have been used to study the pathogenesis of HIV infection.^[27] Blocks of excised lymphoid tissue maintained viability and cytoarchitecture for more than three weeks in the RCCS. Both B and T lymphocytes were able to move freely between the lymphoid tissue and the growth media. The lymphoid tissue in the RCCS could be infected by several strains of HIV-1. It was suggested that such a model system could be used to study HIV pathogenesis as well as potential therapies.

Liver

Primary human liver cells have been cultured in the RCCS, forming 3-D spheroids up to 1.0 cm in diameter.^[28] The constructs were composed of hepatocytes and biliary epithelial cells arranged as bile duct-like structures. Scanning electron microscopy showed clusters of hepatocytes surrounded by extracellular matrix. When hepatocytes were cultured with endothelial cells, albumin mRNA expression was induced. Metabolic studies on porcine hepatocytes cultured in the RCCS demonstrated that the cells maintained normal metabolism for 21 days.^[29]

Conventional culture systems usually cannot maintain hepatocytes for that length of time. Development is currently underway to use RCCS technology for extracorporeal liver assist devices to treat liver failure.

Neural tissue

Murine neural precursor cells cultured in the RCCS formed macroscopic three-dimensional structures that were many times larger than those formed in static culture.^[30] The structures consisted of a surface layer of proliferating cells (nestin- and pcna-positive) and an inner core of more differentiated cells (beta tubulin III- and glial fibrillary acidic protein-positive). The results suggested that the RCCS may be a useful tool for production of transplantable neural tissue.

Pancreatic islets

The failure of islet transplants is the result of both poor initial condition of islets after isolation from the pancreas and rejection by the recipient's immune system. A period of *in vitro* culture markedly improves the success rate of islet transplantation. Islets cultured in the RCCS were shown to be superior—in terms of survival after transplantation—to those maintained in static culture.^[31] Furthermore, these studies showed that diabetic animals were immunotolerant to islets cultured in the RCCS, but not to static cultured islets. Other investigators have utilized the RCCS for coculturing islets and sertoli cells,^[32] which can confer immunotolerance to the sertoli-islet construct. Currently, studies are underway to test the feasibility of using the RCCS to enhance the success of human islet transplantation. Because of the shortage of donor islets, new sources of islets are being explored, including xenotransplants and islets assembled from stem cells. The RCCS can clearly be useful in both of these applications.

SUMMARY

The rotary cell culture system is a relatively new cell culture technology utilizing a zero-headspace, horizontally rotating vessel to grow cells in a low shear, free-fall environment. Under these conditions, anchorage-dependent cells will form three-dimensional aggregates either with themselves or on microcarriers or scaffolding. The aggregates can produce extracellular matrix and remodel into tissue-like structures that have differentiated functions similar to the native tissue. Until recently, the RCCS has been used primarily for basic science applications. However, in the emerging field of tissue engineering, several

applications utilizing tissues generated in the RCCS are approaching the stage of human testing.

The future of the RCCS will almost certainly co-evolve with tissue engineering technology for therapeutic applications. The exciting work now going on in stem cell biology will undoubtedly have a great impact on this area, providing virtually unlimited numbers of cells for production of almost any tissue. The RCCS is ideally suited to play an important role in the development of this technology because it can grow, assemble, and differentiate tissues originated from stem cells.

ARTICLES OF FURTHER INTEREST

Bioreactors; Cardiac Patch Engineering; Tissue Engineering

REFERENCES

- Schwarz, R.P.; Wolf, D.A.; Trinh, T.T. Horizontally Rotated Cell Culture System with a Coaxial Tubular Oxygenator. US Patent Number 5,026,650, 1991.
- Schwarz, R.P.; Goodwin, T.J.; Wolf, D.A. Cell culture for three-dimensional modeling in rotating wall vessels: An application of simulated microgravity. *J. Tissue Cult. Methods* **1992**, *14*, 51–58.
- Hammond, T.G.; Hammond, J.M. Optimized suspension culture: The rotating wall vessel. *Am. J. Physiol. Renal Physiol.* **2001**, *281*, F12–F25.
- Wolf, D.A.; Schwarz, R.P. Analysis of gravity-induced particle motion and fluid perfusion flow in the NASA-designed rotating zero head space tissue culture vessel. NASA Tech. Pap. **1991**, *3143*.
- Cherry, R.S.; Papoutsakis, E.T. Physical mechanisms of cell damage in microcarrier cell culture bioreactors. *Biotechnol. Bioeng.* **1988**, *32*, 1001–1004.
- Cherry, R.S.; Hulle, C.T. Cell death in the thin films of bursting bubbles. *Biotechnol. Prog.* **1992**, *8*, 11–18.
- Wolf, D.A.; Schwarz, R.P. Experimental measurement of the orbital paths of particles sedimenting within a rotating viscous fluid as influenced by gravity. NASA Tech. Pap. **1992**, *3200*.
- Cooper, D.; Pellis, N. Suppressed PHA activation of T lymphocytes in simulated microgravity is restored by direct activation of protein kinase C. *J. Leucoc. Biol.* **1998**, *63*, 550–562.
- Pellis, N.; Goodwin, T.; Risin, D.; McIntyre, B.; Pizzini, R.; Cooper, D.; Baker, T.; Spaulding, G. Changes in gravity inhibit lymphocyte locomotion through type I collagen. *In Vitro Cell. Dev. Biol., Anim.* **1997**, *33*, 398–405.
- Yamauchi, K.; Hales, N.W.; Robinson, S.M.; Niehoff, M.L.; Ramesh, G.T.; Pellis, N.R.; Kulkarni, A.D. Nutrition requirement of nucleotide for cellular immunity in simulated microgravity. *J. Appl. Physiol.* **2002**, *93*, 161–166.
- Kulkarni, A.D.; Yamauchi, K.; Hales, N.W.; Ramesh, V.; Ramesh, G.T.; Sundaresan, A.; Andrassy, R.J.; Pellis, N.R. Nutrition beyond nutrition: Plausibility of immunotropic nutrition for space travel. *Clin. Nutr.* **2002**, *21*, 231–238.
- Hammond, T.G.; Benes, E.; O'Reilly, K.C.; Wolf, D.A.; Linnehan, R.M.; Taher, A.; Kaysen, J.H.; Allen, P.L.; Goodwin, T.J. Mechanical culture conditions affect gene expression: Gravity-induced changes on the space shuttle. *Physiol. Genomics* **2000**, *3*, 163–173.
- Goodwin, T.J.; Schroeder, W.F.; Wolf, D.A.; Moyer, M.P. Rotating-wall vessel coculture of small intestine as a prelude to tissue modeling: Aspects of simulated microgravity. *Proc. Soc. Exp. Biol. Med.* **1993**, *202*, 181–192.
- Becker, J.L.; Prewitt, T.L.; Spaulding, G.F.; Goodwin, T.J. Three dimensional growth and differentiation of ovarian tumor cell line in high aspect rotating wall vessel: Morphologic and embryologic considerations. *J. Cell. Biochem.* **1993**, *51*, 283–289.
- Freed, L.E.; Hollander, A.P.; Barry, J.R.; Langer, R.; Vunjak-Novakovic, G. Chondrogenesis in a cell-polymer bioreactor system. *Exp. Cell Res.* **1998**, *240*, 58–65.
- Jessup, J.M.; Brown, D.; Fitzgerald, W.; Ford, R.D.; Nachman, A.; Goodman, T.J.; Spaulding, G. Induction of carcinoembryonic antigen expression in a three dimensional culture system. *In Vitro Cell. Dev. Biol., Anim.* **1997**, *33*, 352–357.
- Nakamura, K.; Kuga, H.; Morisaki, T.; Baba, E.; Sato, N.; Mizumota, K.; Sueishi, K.; Tanaka, M.; Katano, M. Simulated microgravity culture system for a 3-D carcinoma model. *BioTechniques* **2002**, *33*, 1068–1076.
- Rajan, A.; Navran, S. Microgravity tissue engineering of pancreatic islets. *Diabetes* **1999**, *48* (Suppl. 1), A340.
- Marlovits, S.; Tichy, B.; Truppe, M.; Gruber, D.; Schlegel, W. Collagen expression in tissue engineered cartilage of aged human articular chondrocytes in a rotating bioreactor. *Int. J. Artif. Organs* **2003**, *26*, 319–330.
- Vunjak-Novakovic, G.; Martin, I.; Obradovic, B.; Treppo, S.; Grodzinsky, A.J.; Langer, R.; Freed, L.E. Bioreactor cultivation conditions modulate the composition and mechanical properties of tissue-engineered cartilage. *J. Orthop. Res.* **1999**, *17*, 130–138.
- Vunjak-Novakovic, G.; Obradovic, B.; Martin, I.; Freed, L.E. Bioreactor studies of native and tissue engineered cartilage. *Biorheology* **2002**, *39*, 259–268.
- O'Connor, K.C. Three-dimensional cultures of prostatic cells: Tissue models for the development of novel anti-cancer therapies. *Pharm. Res.* **1999**, *16*, 486–493.
- Licato, L.L.; Prieto, V.G.; Grimm, E.A. A novel pre-clinical model of human malignant melanoma utilizing bioreactor rotating-wall vessels. *In Vitro Cell. Dev. Biol., Anim.* **2001**, *37*, 121–126.
- Chopra, V.; Dinh, T.V.; Hannigan, E.V. Three dimensional endothelial-tumor epithelial cell interactions in human cervical cancers. *In Vitro Cell. Dev. Biol., Anim.* **1997**, *33*, 432–442.
- Cowger, N.L.; O'Connor, K.C.; Bivins, J.E. Influence of simulated microgravity on the longevity of insect-cell culture. *Enzyme Microb. Technol.* **1997**, *20*, 326–332.

26. Nickerson, C.A.; Goodwin, T.J.; Terlonge, J.; Ott, C.M.; Buchanan, K.L.; Uicker, W.C.; Emami, K.; LeBlanc, C.L.; Ramamurthy, R.; Clarke, M.S.; Vanderburg, C.R.; Hammond, T.; Pierson, D.L. Three-dimensional tissue assemblies: Novel models for the study of *Salmonella enterica* serovar Typhimurium pathogenesis. *Infect. Immun.* **2001**, *69*, 7106–7120.
27. Margolis, L.B.; Fitzgerald, W.; Glushakova, S.; Hatfill, S.; Amichay, N.; Baibakov, B.; Zimmerberg, J. Lymphocyte trafficking and HIV infection of human lymphoid tissue in a rotating wall vessel. *AID Res. Hum. Retroviruses* **1997**, *13*, 1411–1420.
28. Yoffe, B.; Darlington, G.J.; Soriano, H.E.; Krishnan, B.; Risin, D.; Pellis, N.R.; Khaoustov, V.L. Cultures of human liver cells in simulated microgravity environment. *Adv. Space Res.* **1999**, *24*, 829–836.
29. Dabos, K.J.; Nelson, L.J.; Bradnock, T.J.; Parkinson, J.A.; Sadler, I.H.; Hayes, P.C.; Plevris, J.N. The simulated microgravity environment maintains key metabolic functions and promotes aggregation of primary porcine hepatocytes. *Biochim. Biophys. Acta* **2001**, *1562*, 119–130.
30. Low, H.P.; Savarese, T.M.; Schwartz, W.J. Neural precursor cells form rudimentary tissue-like structures in a rotating wall vessel bioreactor. *In Vitro Cell. Dev. Biol., Anim.* **2001**, *37*, 141–147.
31. Rutsky, L.P.; Bilinski, S.; Kloc, M.; Phan, T.; Zhang, H.; Katz, S.M.; Stepkowski, S.M. Microgravity culture conditions reduces immunogenicity and improves function of pancreatic islets. *Transplantation* **2002**, *74*, 13–21.
32. Cameron, D.F.; Hushen, J.J.; Nazian, S.J.; Willing, A.; Saporta, S.; Sanberg, P.R. Formation of sertoli cell-enriched tissue constructs utilizing simulated microgravity technology. *Ann. N.Y. Acad. Sci.* **2001**, *944*, 420–428.

Self-Assembled Monolayers

Koichi Kato

Institute for Frontier Medical Sciences, Kyoto University, Sakyo-ku, Kyoto, Japan



INTRODUCTION

A self-assembled monolayer (SAM) is a thin, organic film that is formed spontaneously by adsorption of pre-designed organic precursors from the solution or vapor phase to the surface of an appropriate solid substrate. The process starts with chemisorption of headgroup possessed by precursor molecules and eventually produces a two-dimensional, highly ordered, thin layer. Many different combinations of a headgroup and substrate have been employed, such as organosilicons on hydroxylated surfaces (e.g., silica on silicon, alumina on aluminum, glass, etc.); alkanethiols on gold, silver, and copper; dialkyl sulfides and dialkyl disulfides on gold; amines on platinum; and carboxylic acids on aluminum oxide and silver. Among these combinations, alkanethiols on gold and organosilicons on glass have been studied most extensively in the field of biomaterials research.

The processes to obtain SAMs have many characteristic features and advantages over other methods of surface preparation for basic and applied studies on biomaterial surfaces. First of all, one can easily obtain a well-defined planar surface. Such a surface can be used as a simple, rigid model of organic biomaterials. As described subsequently, it is easy to prepare SAMs with various functionalities at the topmost surface in a well-controlled density and distribution by selecting a type of self-assembling molecules and compositions in the case of mixed SAMs. Other important advantage is the possibility of applying a wide variety of techniques for the characterization of SAM surfaces, such as infrared spectroscopy (IR), second harmonic generation (SHG), sum frequency generation (SFG), surface enhanced-Raman scattering (SERS), high-resolution electron energy loss spectroscopy (HREELS), near-edge X-ray absorption fine structure (NEXAFS), X-ray photoelectron spectroscopy (XPS), surface plasmon resonance spectroscopy (SPR), and scanning probe microscopy (SPM). As can be seen in recent publications, SAM surfaces have found many applications in the two-dimensional micropatterning of chemicals, biomolecules, and live cells.

The following deals with fundamental aspects of SAM formation and its supramolecular structure. It focuses on two representative cases of SAM systems: alkanethiol and organosilicon monolayers. The overview

of the biotechnological applications of SAMs will be given in the last section. Langmuir–Blodgett films of amphiphiles and monolayers consisting of fatty acids are also classical issues in surface engineering, but they are not dealt with in this article.

SAM FORMATION

As mentioned in the preceding section, various combinations of self-assembling molecules and substrates have been employed for the preparation of SAMs.^[1,2] Table 1 shows the example of various SAM systems appearing in literature. The critical steps of SAM formation involve: 1) chemisorption of headgroup to the substrate surface and 2) two-dimensional organization through the lateral interchain interactions. Both steps are an exothermic process, and the former serves to initiate the short-range, dispersive, London-type, van der Waals interactions between self-assembling molecules.^[2] Fig. 1 simply depicts a self-assembling building block that consists sequentially of a surface-active headgroup chemisorbed at the substrate surface, an alkyl or derivatized alkyl chain that is capable of lateral interactions with the neighboring counterparts, and a terminal group exposed to the outer phase. Another important backbone for SAM molecules can be a chain that contains aromatic rings. The physicochemical properties of the three components as well as a type of substrate and medium used, e.g., solvents in the case of chemisorption from solution, have synergic effects on the entire process of SAM formation and the properties of resulting monolayers.

The following two sections are devoted to the formation of alkanethiol monolayers on noble metal surfaces and organosilicon monolayers on hydroxylated surfaces.

Alkanethiol on Metal

SAMs made of organosulfur compounds, including alkanethiol, dialkyl disulfide, dialkyl sulfide, aromatic thiol, alkyl xanthate, and dialkylthiocarbamate, have been studied extensively.^[37] The key step of monolayer formation is the coordination of sulfur to the surface of metals such as gold, silver, copper, platinum, mercury, iron, γ -Fe₂O₃, GaAs, and InP. An amine-terminated

Table 1 Example of various SAM systems appearing in literature

Anchoring group	Assembling molecule	Substrate	Reference
Carboxylate	<i>n</i> -Alkanoic acid	Al, Al ₂ O ₃	[3–6]
	<i>n</i> -Alkanoic acid	Ag	[7]
	<i>n</i> -Perfluorocarboxylic acid	Ag	[8]
Amine	ω -substituted alkyl amine	Pt	[9]
	ω -substituted alkyl amine	ITO (tin doped indium oxide)	[10]
Siloxane	<i>n</i> -Alkyltrichlorosilane and its derivative	Si, oxidized Si, glass	[11–15]
	<i>n</i> -Alkyltrichlorosilane	Au	[16]
	<i>n</i> -Alkyltrichlorosilane derivative	TiO ₂	[17]
Thiolate	<i>n</i> -Alkanethiol	Au	[18–25]
	Dialkyl sulfide and disulfide	Au	[19,24,26,27]
	Aromatic thiol and disulfide	Au	[28,29]
	Thiol-terminated peptide	Au	[30]
	<i>n</i> -Alkanethiol	Ag	[31,32]
	<i>n</i> -Alkanethiol	Cu	[33]
	<i>n</i> -Alkanethiol	Pt	[34–36]

alkyl chain, even though this does not contain sulfur, forms a SAM on a platinum surface, which is similar to the case with alkanethiols. The surface of gold is the most frequently studied as a substrate for SAMs constructed with alkanethiols due to ease of preparation, well-defined structure, and absence of an oxidized layer.

The surface of gold can be prepared by vapor-phase evaporation of polycrystalline gold onto a glass plate properly cleaned by, e.g., treatment with piranha solution (7:3 concentrated H₂SO₄/30% H₂O₂). Since adhesion at the interface between glass and gold is poor, a thin layer of titanium or chromium is coated prior to evaporation of the gold to avoid delamination of the gold. Annealing or a simple flame treatment of a thin film of gold formed by evaporation is performed to obtain the Au(111) surface with a well-defined lattice structure.

So far, there are many types of alkanethiol used in the study of basic and applied research on SAMs on

gold. Some of them are commercially available, and others must be synthesized in-house. One can use both the oxidized (disulfide) and reduced (thiol) forms with similar results with respect to the structure of SAMs formed. An intermitting chain is an *n*-alkyl chain with a carbon number of, in practice, 1–20 or derivatized alkyl chains containing an aromatic ring, unsaturated alkene, etc. The length of *n*-alkyl chain has great effect on the kinetics of SAM formation and the order and stability of the formed monolayers. An alkanethiol with a longer alkyl chain forms a SAM more rapidly than that with a shorter chain. It is known that the presence of a polar group in the chain has a great effect on the second step of SAM formation, i.e., ordering of chemisorbed alkanethiol, due to attractive interchain forces. Depending on the research interests, many different functional groups can be incorporated to the terminal of *n*-alkanethiol. Proper functional groups, such as carboxylic acids, provide the possibility to derivatize the SAM surface for further applications. The example is modification with poly(ethylene glycol) chains; cell-binding peptides such as RGD (Arg-Gly-Asp); biotin, which strongly binds to avidin and streptavidin; histidine-tag chelating groups; porphyrin; etc.

The procedure employed for SAM preparation is very simple. In a typical case, a plate coated with a thin layer of gold is just immersed in a dilute solution of alkanethiol to initiate SAM formation. The plate is then washed with solvent to remove weakly adsorbed alkanethiol. Usually, the reaction is carried out under ambient conditions. The concentration of alkanethiol solution has an effect on the kinetics of SAM formation. As shown in Fig. 2,^[22] depending on the

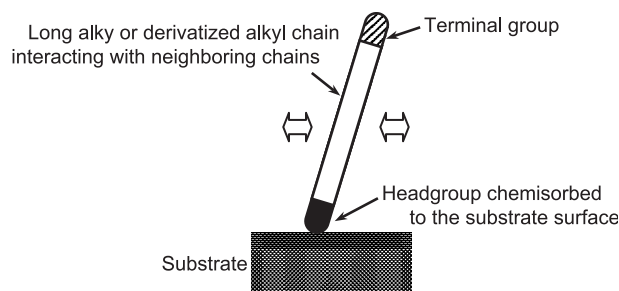


Fig. 1 A self-assembling precursor molecule consisting of a headgroup, a long alkyl chain, and a terminal group.

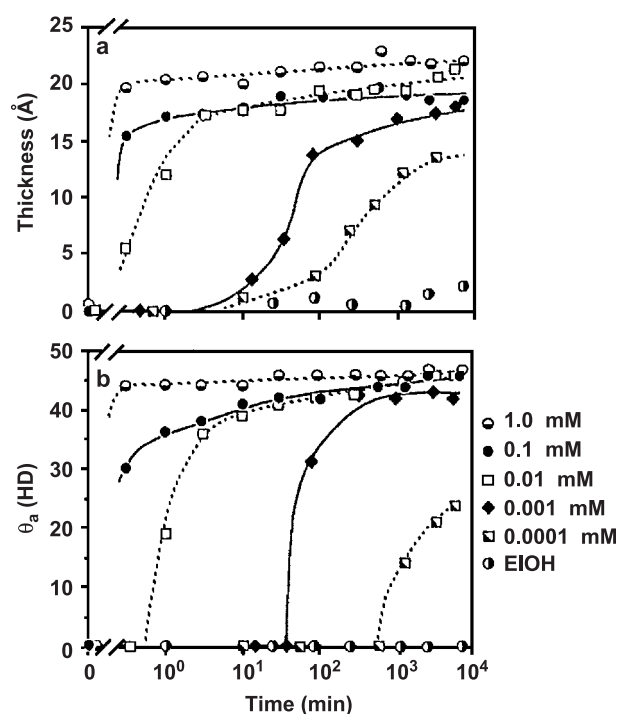
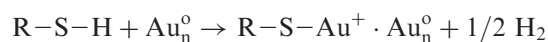


Fig. 2 Kinetics of adsorption of octadecanethiol from ethanol as a function of concentration: a) ellipsometric thickness and b) advancing contact angles. (Reprinted with permission from *J. Am. Chem. Soc.* **1989**, *111*, 321–335. Copyright (1989) American Chemical Society.^[22])

concentration of alkanethiol solution, time dependence of the ellipsometric thickness and the advancing contact angle is largely altered. These measurements represent the surface coverage and also the order of alkanethiol molecules. It can be seen that higher concentration gives rise to the rapid formation of SAM. The most popular solvent used in alkanethiol monolayer preparation is ethanol. Other organic solvents can also be used if they give an isotropic solution and have no specific interactions to the alkanethiol and the substrate. In case of formation of mixed SAM, by which different alkanethiols simultaneously form SAM from the mixed solution, small difference in the solubility power may produce a mixed SAM with an enriched single component compared to the other.

When the bare surface of gold is brought into contact with an alkanethiol solution, thiol is considered to react with gold through the oxidative addition of the S–H bond followed by reductive elimination of the hydrogen, which results in the formation of thiolate.^[2]



The energy associated with the sulfur–gold bond formation is estimated to be 40 kcal/mol^[2] while that with

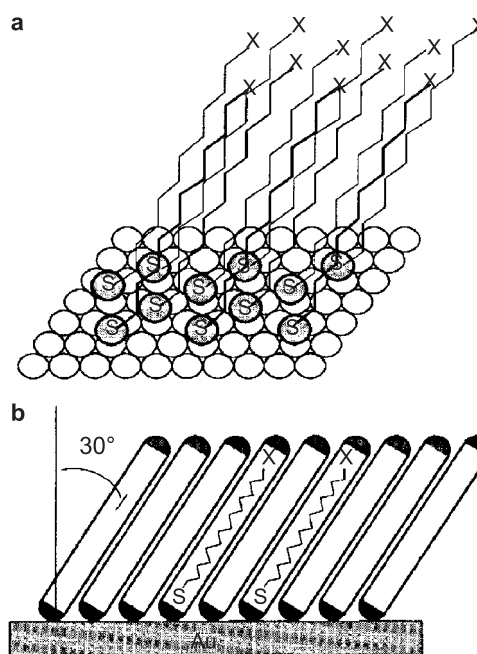


Fig. 3 Representation of a self-assembled monolayer (SAM) of alkanethiolates on the surface of Au(111). (Reprinted with permission from *Trends in Biotechnology* **1995**, *13*, 228–235. Copyright (1995) Elsevier.^[38])

interchain van der Waals interactions of alkyl chains is at the order of few (< 10) kcal/mol.^[2] The sulfur–gold bond is robust enough to fix alkanethiol molecules to the surface of gold, but the bonding energy is not necessarily high compared to other covalent bonds, which, in turn, gives the opportunity to be cleaved by site-restricted oxidation via, e.g., UV radiation through a photomask with a negative pattern that leaves a patterned SAM.

The structure of SAM that is organized on the surface of Au(111) has been extensively studied by means of electron diffraction, helium diffraction, atomic force microscopy, and ultrahigh vacuum scanning tunneling microscopy. As shown in Fig. 3,^[38] the sulfur atoms of alkanethiol coordinate to the hollow three-fold sites of the Au(111) surface on which the gold atoms are arranged in a hexagonal packing. The alkyl chains of alkanethiol are closely packed by interchain interactions and tilted approximately 30° from the normal to the surface. Readers should see Refs. [1,2,37] for more comprehensive consideration on the equilibrium and nonequilibrium structure of alkanethiol monolayers.

SAM of Organosilicon Derivatives on Hydroxylated Surface

Another SAM system important in the biotechnological application is a monolayer built up by the class of

organosilicon derivatives, R_nSiX_{3-n} ($n=1-3$). X is chloride or alkoxy, and R is an aliphatic or an aromatic hydrocarbon chain terminated with a functional group. The formation of SAM with organosilicon derivatives is primarily based on the chemisorption of a chloride or alkoxy terminal, which forms a covalent linkage of Si–O at the hydroxylated surface that has –OH groups, such as SiO_2 , SnO_2 , and TiO_2 . A glass is one of the most practical substrates. The Si–O linkage is also formed between neighboring organosilicons to form a two-dimensional layered, cross-linked network. Compared to other systems, e.g., a monolayer formed by fatty acid molecules by self-assembly, such a network is considered to provide a durable surface modification.

In the case of SAMs of organosilicon derivatives, the formation of an ordered, closely packed monolayer relies upon two steps: 1) chemisorption of self-assembling molecules associated with chemical bonding and cross-linking and 2) reorganization of pre-adsorbed molecules through lateral van der Waals interactions between hydrocarbon chains. Typically, a substrate having a hydroxylated surface, such as a properly cleaned glass, is immersed in a solution of organosilicon, such as alkyltrichlorosilane, in an organic solvent to obtain organosilicon SAMs.

It is often assumed that the chlorine or alkoxy are split off and the molecules are bound to the surface and each other by a Si–O–Si linkage. However, reality seems to be much more complicated. The process is reported to take place through two pathways: 1) direct chemisorption of a monomeric organosilicon and 2) adsorption of pre-formed oligomeric derivatives formed by Si–O–Si condensation in bulk solution. Such two different processes have implications in the structure of SAM formed, and it is often argued that organosilicon SAM is inherently disordered than, e.g., alkanethiol monolayers, in which molecules have more freedom to establish long-range order. Assuming that a monolayer completely covers the substrate surface and forms in situ Si–O bonds, there would not be enough space for all-trans hydrocarbon chains to organize a closely packed, ordered structure. According to Schreiber,^[1] octadecyltrichlorosilane (OTS; $CH_3-(CH_2)_{17}-SiCl_3$) forms a SAM on oxidized silicon with a thickness of circa 2.5 nm with a very limited correlation length. The alkyl chains are packed mostly in the all-trans conformation with a tilt angle of at most 20° normal to the surface, which exposes the terminal methyl group.

BIOTECHNOLOGICAL APPLICATIONS

As described in the preceding discussion, SAMs can be prepared by relatively simple processes and provide well-characterized, durable, organic thin layers on the surface of a variety of substrates. These advantages make SAMs

versatile modification techniques that have a potential to be utilized in various applications, such as coatings for the control of corrosion, wetting, friction, lubrication, and adhesion. Functionalized SAMs can also be used for graft-polymerization of organic monomers from the surfaces. The reactions and reactivity in SAMs were reviewed in Checik et al.,^[39] which discusses the growth of surface-attached polymers, intrafilm reactions, chemistry, photochemistry, and reactivity issues. Primarily due to the following two reasons, SAM is important in biotechnological applications. The surface of SAMs provides the model system in which very complex biological events taking place under in vivo and other quasiphysiological conditions can be simply simulated in vitro. The other advantage of using SAMs is concerned with much more practical applications as a component of biosensors and biochips.

It is often in biomaterials research that a detailed study on biomolecular interactions at the surface of artificial devices is affected by several intrinsic factors, such as surface roughness, dynamic property of surface molecules, and less controllable surface chemistry. Using SAM as a model surface can potentially circumvent these problems. In addition, various analytical methodologies are available for the precise study on surface events at a high resolution and sensitivity. They include quartz crystal microbalance, ellipsometry, and internal reflectance fluorescence. Besides them, scanning probe microscopy (SPM) has received considerable attention due to its potential to analyze in situ a broad range of biological objects. This is mostly owing to the specific feature that the SPM can be operated in aqueous media in contrast to other conventional techniques such as scanning electron microscopy (SEM). Currently, the SPM operated in aqueous media provides molecular-level images with a high signal-to-noise ratio that is comparable to those obtained by electron microscopy. Another powerful technique is surface plasmon resonance (SPR) spectroscopy, which takes advantage of a physical phenomenon of SPR to detect tiny changes in the refractive index of a medium near a thin metal film.^[40] The alkanethiol monolayer formed on gold is frequently employed because it is well suited for the configuration of SPR spectroscopy. Protein adsorption, protein–protein interactions, and DNA–DNA hybridization have been studied by many research groups using this technique.^[41]

Figure 4^[42] shows the example of SPR measurements in which interactions of serum complement fragment C3b were analyzed on mixed alkanethiol SAMs to investigate the effect of the surface density of hydroxyl group on the binding of complement fragment C3b to the surface as a result of complement activation. Information on the complement activation by biomaterial surfaces has important implications in various clinical settings, such as open heart surgery, blood transfusion

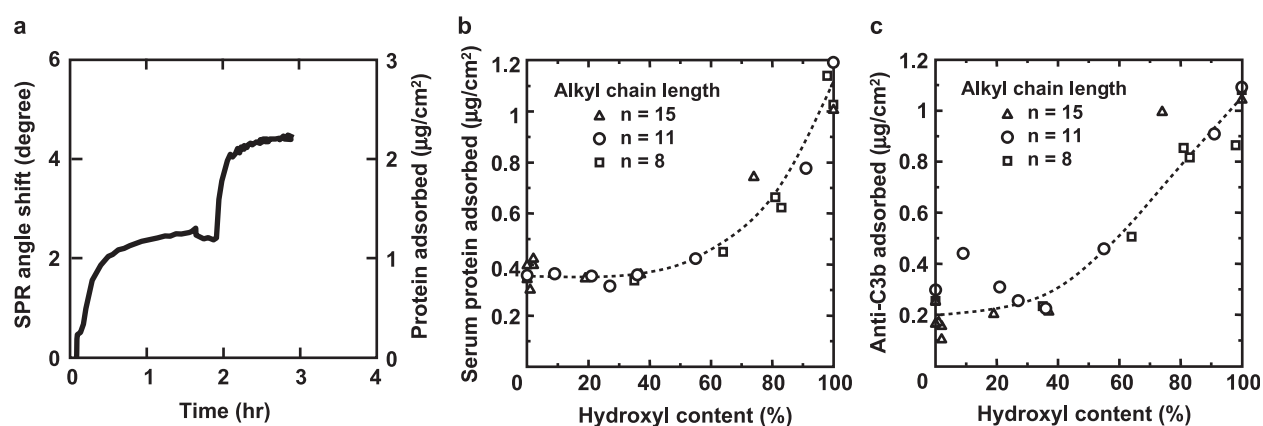


Fig. 4 a) Time curve of serum protein adsorption onto a SAM terminated with hydroxyl groups and anti-human C3b antibody binding to the protein adlayer formed on the SAM. The surface plasmon resonance angle shift was determined during these processes; b) The amount of serum proteins adsorbed on the mixed SAMs as a function of the hydroxyl content; c) The amount of anti-C3b antibody bound to the protein adlayer. (Reprinted with permission from *J. Biomed. Mater. Res.* **2003**, *66A*, 669–676. Copyright (2003) John Wiley & Sons, Inc.^[42].)

medicine, and extracorporeal immunotherapy. The alkanethiols used for preparation of mixed SAMs were 11-mercapto-1-undecanol (HO-(CH₂)₁₁-SH), *n*-nonanethiol (CH₃-(CH₂)₈-SH), *n*-dodecanethiol (CH₃-(CH₂)₁₁-SH), and *n*-hexadecanethiol (CH₃-(CH₂)₁₅-SH). A series of mixed SAMs with different molar ratios was prepared using the hydroxyl-terminated alkanethiol and one of the alkanethiols with a methyl end group. For SPR measurement, a p-polarized He-Ne laser beam ($\lambda = 632.8$ nm) was illuminated to the backside of the glass plate through a glass prism (the Kretschmann configuration) at an incident angle near the SPR angle while a protein solution was allowed to pass through a flow cell. The intensity of a reflected light was determined by a PMT detector. The immobilization of C3b on the SAMs was assessed from specific binding of an anti-C3b antibody. As can be seen in Fig. 4a, a small amount of protein interacting with the surface could be detected in real time under the physiological condition. The results show that an increase in the hydroxyl content gave rise to an enhanced deposition of serum proteins (Fig. 4b). As is evident from Fig. 4c, the enhancement is likely due to binding of C3b upon complement activation at the surface through the alternative pathway.

Currently, the SPR technique is widely used in biomolecular interaction analysis at SAM surfaces. The next generation of the SPR measurement involves a biochip analysis by which multiple biomolecular interactions are analyzed in parallel using a microarray format or microfluidics prepared by utilizing SAM techniques. To realize such a high-throughput analysis, an imaging apparatus was developed^[43] by taking advantage of the SPR induced by a collimated, expanded, and p-polarized light source and a CCD camera as a detector. Surface-plasmon field-enhanced fluorescence

microscopy^[44] was also reported for the parallel multipot detection of biomolecules bound to SAM-based microarrays.

A key step to analyze biomolecular interactions is the immobilization of biomolecules (ligand) onto the surface of SAMs.^[45] A variety of chemistries has been applied to protein immobilization onto SAM surfaces. A frequently used method for efficient protein immobilization is to activate carboxylic acid with the use of carbodiimide. First, a carboxylic acid-terminated alkanethiol monolayer is prepared and then activated by carbodiimide. Water-soluble and -insoluble carbodiimides, *N*-3-dimethyl-aminopropyl-*N'*-ethylcarbodiimide (EDC) and dicyclohexylcarbodiimide (DCC), can be used in an aqueous media and organic solvent, respectively. It is known that carbodiimide-mediated coupling reactions are much more efficient with the concomitant use of *N*-hydroxysuccinimide (NHS) or *N*-hydroxysulfosuccinimide (sulfo-NHS). Figure 5 represents the reaction scheme for the activation of carboxylate by using EDC plus NHS and the subsequent coupling of an amine-containing ligand. The advantage of this modification lies in the long half life of the intermediate active NHS-ester formed during activation with EDC, which, in turn, gives rise to a substantially increased coupling yield. Since the remaining activity of proteins should be kept as high as possible after their immobilization on the SAM surface, several new methods have been proposed. A biotin-terminated alkanethiol can be used to prepare the molecular assembly of a streptavidin layer. Biotinylated proteins and nucleic acids can then be coupled to the immobilized streptavidin, which makes use of a highly specific biotin-streptavidin interaction.^[46] However, chemical modification may lead to inactivation of proteins, and the orientation of protein molecules cannot be controlled

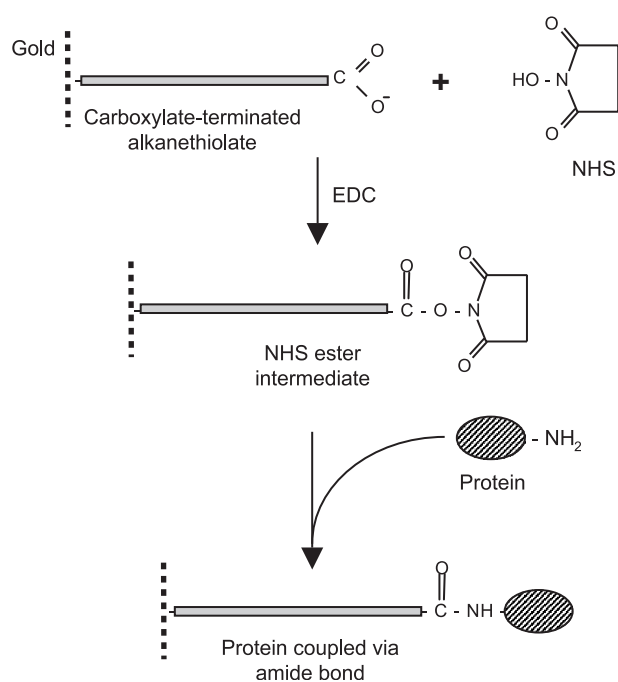


Fig. 5 Reaction scheme for the immobilization of protein to the surface of a carboxylate-terminated SAM through an active *N*-hydroxysuccinimide ester intermediate.

during immobilization reaction. The sophisticated way to overcome these problems may be to use an alkanethiol that is terminated with a Ni(II)-chelating nitrilotriacetic acid, which selectively binds proteins with a histidine tag.^[47] This method takes advantage of the fact that recombinant fusion proteins tagged with a peptide of six histidines interact with two vacant sites on Ni(II) ions.

Recently, there are many publications on the micro-contact printing (μCP) method.^[38] This method is an attractive way to create a micropatterned monolayer of alkanethiols on gold and has found a variety of biotechnological applications. An alkanethiol solution is inked and dried on an elastomeric polydimethylsiloxane stamp with a micropattern that is prepared using a rigid, master template. Then, the stamp is pressed onto the gold-coated surface to transfer the alkanethiol to the gold surface. A SAM of the second alkanethiol with chemical properties different from the stamped one is then formed from solution to the surrounding region, which produces a micropatterned SAM in which two distinct regions are located, e.g., lines, dots, and any other shapes. With the μCP , one can easily obtain a micropatterned SAM with a feature of sizes ranging down to $1\ \mu\text{m}$. It is claimed that the method can be performed at a lower cost and with higher resolution than other processes such as UV lithography in which a restricted area of a pre-formed homogeneous SAM is oxidized with a UV light that is radiated through a patterned mask.^[38] Recently, a patterned protein nanostructure was fabricated on a SAM using dip-pen nanolithography (DPN).^[48] The method involves the fabrication of chemically reactive nanoscale feature by patterning a SAM of a COOH-terminated alkanethiol on a gold substrate by DPN, which is followed by covalent immobilization of a high-affinity, small-molecule ligand on the nanopatterned SAM and subsequent molecular recognition of its protein-binding partner from solution.

The patterned SAMs containing reactive regions can be further utilized to create the pattern of live cells. Figure 6 shows photographs of acute lymphoblastic

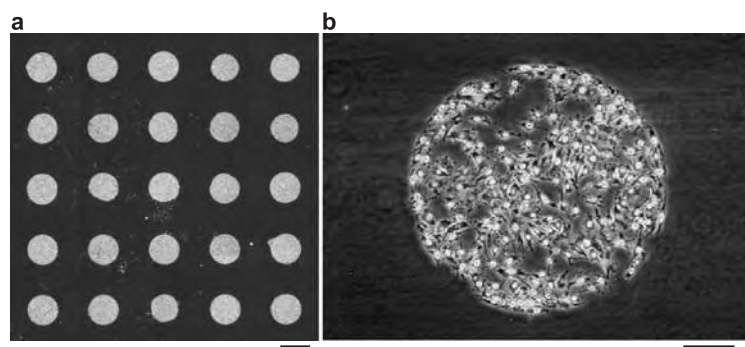


Fig. 6 a) Acute lymphoblastic leukemia cells (CCRF-CEM) attached to the patterned SAM. The homogeneous SAM of *n*-hexadecanethiol ($\text{CH}_3(\text{CH}_2)_{15}\text{SH}$) formed on a gold-coated glass plate was firstly patterned by UV lithography and then a SAM of carboxylic acid-terminated alkanethiol ($\text{HOOC}(\text{CH}_2)_{10}\text{SH}$) was formed in the irradiated regions. The carboxylic acid present in discrete regions was activated with DCC and NHS, and a monoclonal antibody to T-cell antigen, CD5, was covalently immobilized to the activated regions. After blocking the surrounding methyl-presenting SAM with bovine serum albumin, cells were plated to the SAM surface with patterned antibody. Cells are attached to 5×5 circular spots (1 mm diameter) with immobilized antibody. b) High-magnification image of a single spot (1 mm diameter) with mouse stromal cells, PA6. A patterned SAM was prepared by the same fashion as for leukemia cells, but fibronectin was immobilized to activated spots instead of antibody. Bar: (a) 1 mm, (b) $200\ \mu\text{m}$.

leukemia cells and mouse stromal cells that are attached to the patterned SAM surface. Specific interactions between cells and immobilized biomolecules are used to achieve regionally restricted cell attachment while cell adhesion is prevented on the surrounding regions by blocking with nonspecific molecules such as albumin and polymeric or oligomeric ethylene glycol. For the specific interactions, extracellular matrix proteins, e.g., collagen, fibronectin, and laminin; cell adhesion peptides, e.g., RGD peptide; and antibodies are immobilized to the reactive regions in order to capture cells expressing integrins and surface antigens. Using the techniques for preparing patterns of live cells, many attempts have been made to manipulate cell growth and axonal extension. The coculture system of two different cell types can be also established by making use of a patterned SAM, which provides a simple model for mechanistic studies of heterotypic cell-cell signaling in two-dimensional cellular organization.^[49] The highly integrated cell arrays prepared by patterning SAM will also provide promising platforms for the cell-based, high-throughput, functional analysis in which live cells are used to massively screen thousands of drugs, proteins, and genes in parallel.

CONCLUSION

Self-assembled monolayers are a thin organic film that are spontaneously formed by chemisorption and ordering of precursor molecules at the solid surface. Two important SAM systems, including the monolayers of alkanethiol on gold and organosilicon monolayers on hydroxylated surfaces, have been extensively studied. The kinetics of film formation and structural aspects are now well understood for these two systems. SAM is one of the best systems to establish a model mimicking biomolecular interactions at surfaces. Further modification of SAM terminals as well as patterning provides a broad range of platforms for biotechnological applications such as biomaterials, biosensors, and biochips.

ARTICLES OF FURTHER INTEREST

Polyelectrolyte Multilayers; Surface Coatings; Surface Modification; Surface Topography and Cell Behavior

REFERENCES

- Schreiber, F. Structure and growth of self-assembling monolayers. *Prog. Surf. Sci.* **2000**, *65*, 151–256.
- Ulman, A. Self-Assembled Monolayers. In *An Introduction to Ultrathin Organic Films from Langmuir–Blodgett to Self-Assembly*; Academic Press: San Diego, CA, 1991; 237–304.
- Allara, D.L.; Nuzzo, R.G. Spontaneously organized molecular assemblies. 1. Formation, dynamics, and physical properties of *n*-alkanoic acids adsorbed from solution on an oxidized aluminum surface. *Langmuir* **1985**, *1*, 45–52.
- Allara, D.L.; Nuzzo, R.G. Spontaneously organized molecular assemblies. 2. Quantitative infrared spectroscopic determination of equilibrium structures of solution-adsorbed *n*-alkanoic acids on an oxidized aluminum surface. *Langmuir* **1985**, *1*, 52–66.
- Ogawa, H.; Chihara, T.; Taya, K. Selective monomethyl esterification of dicarboxylic acids by use of monocarboxylate chemisorption on alumina. *J. Am. Chem. Soc.* **1985**, *107*, 1365–1369.
- Cohen, S.R.; Naaman, R.; Sagiv, J. Thermally induced disorder in organized organic monolayers on solid substrates. *J. Phys. Chem.* **1986**, *90*, 3054–3056.
- Schlotter, N.E.; Porter, M.D.; Bright, T.B.; Allara, D.L. Formation and structure of a spontaneously adsorbed monolayer of arachidic on silver. *Chem. Phys. Lett.* **1986**, *132*, 93–98.
- Chau, L.-K.; Porter, M.D. Composition and structure of spontaneously adsorbed monolayers of *n*-perfluorocarboxylic acids on silver. *Chem. Phys. Lett.* **1990**, *176*, 198–204.
- Brito, R.; Rodríguez, V.A.; Figueroa, J.; Cabrera, C.R. Adsorption of 3-mercaptopropyltrimethoxysilane and 3-aminopropyltrimethoxysilane at platinum electrodes. *J. Electroanal. Chem.* **2002**, *520*, 47–52.
- Oh, S.-Y.; Yun, Y.-J.; Kim, D.-Y.; Han, S.-H. Formation of a self-assembled monolayer of diaminododecane and a heteropolyacid monolayer on the ITO Surface. *Langmuir* **1999**, *15*, 4690–4692.
- Sagiv, J. Organized monolayers by adsorption. 1. Formation and structure of oleophobic mixed monolayers on solid surfaces. *J. Am. Chem. Soc.* **1980**, *102*, 92–98.
- Netzer, L.; Sagiv, J. A new approach to construction of artificial monolayer assemblies. *J. Am. Chem. Soc.* **1983**, *105*, 674–676.
- Maoz, R.; Sagiv, J. Penetration-controlled reactions in organized monolayer assemblies. 2. Aqueous permanganate interaction with self-assembling monolayers of long-chain surfactants. *Langmuir* **1987**, *3*, 1045–1051.
- Tillman, N.; Ulman, A.; Schildkraut, J.S.; Penner, T.L. Incorporation of phenoxy groups in self-assembled monolayers of trichlorosilane derivatives. Effects on film thickness, wettability, and molecular orientation. *J. Am. Chem. Soc.* **1988**, *110*, 6136–6144.
- Tillman, N.; Ulman, A.; Penner, T.L. Formation of multilayers by self-assembly. *Langmuir* **1989**, *5*, 101–111.
- Finklea, H.O.; Robinson, L.R.; Blackburn, A.; Richter, B.; Allara, D.; Bright, T. Formation of an organized monolayer by solution adsorption of octadecyltrichlorosilane on gold: Electrochemical properties and structural characterization. *Langmuir* **1986**, *2*, 239–244.
- Cossement, D.; Pierard, C.; Delhalle, J.; Pireaux, J.-J.; Hevesi, L.; Mekhalif, Z. Synthesis of 6-(1'-pyrrolyl)-*n*-hexyltrichlorosilane and 6-(1'-pyrrolyl)-*n*-hexyltrimethoxysilane: XPS characterization of their monolayers self-assembled on polycrystalline titanium surfaces. *Surf. Interface Anal.* **2001**, *31*, 18–22.

18. Porter, M.D.; Bright, T.B.; Allara, D.L.; Chidsey, C.E.D. Spontaneously organized molecular assemblies. 4. Structural characterization of *n*-alkyl thiol monolayers on gold by optical ellipsometry, infrared spectroscopy, and electrochemistry. *J. Am. Chem. Soc.* **1987**, *109*, 3559–3568.
19. Strong, L.; Whitesides, G.M. Structures of self-assembled monolayer films of organosulfur compounds adsorbed on gold single crystals: Electron diffraction studies. *Langmuir* **1988**, *4*, 546–558.
20. Rubinstein, I.; Steinberg, S.; Tor, Y.; Shanzer, A.; Sagiv, J. Ionic recognition and selective response in self-assembling monolayer membranes on electrodes. *Nature* **1988**, *332*, 426–429.
21. Chidsey, C.E.D.; Liu, G.-Y.; Rowntree, P.; Scoles, G. Molecular order at the surface of an organic monolayer studied by low energy helium diffraction. *J. Chem. Phys.* **1989**, *91*, 4421–4423.
22. Bain, C.D.; Troughton, B.; Tao, Y.T.; Evall, J.; Whitesides, G.M.; Nuzzo, R.G. Formation of monolayer films by the spontaneous assembly of organic thiols from solution onto gold. *J. Am. Chem. Soc.* **1989**, *111*, 321–335.
23. Bain, C.D.; Whitesides, G.M. Formation of monolayers by the coadsorption of thiols on gold: variation in the length of the alkyl chain. *J. Am. Chem. Soc.* **1989**, *111*, 7164–7175.
24. Whitesides, G.M.; Laibinis, P.E. Wet chemical approaches to the characterization of organic surfaces: Self-assembled monolayers, wetting, and the physical-organic chemistry of the solid–liquid interface. *Langmuir* **1990**, *6*, 87–96.
25. Nuzzo, R.G.; Dubois, L.H.; Allara, D.L. Fundamental studies of microscopic wetting on organic surfaces. 1. Formation and structural characterization of a self-consistent series of polyfunctional organic monolayers. *J. Am. Chem. Soc.* **1990**, *112*, 558–569.
26. Nuzzo, R.G.; Fusco, F.A.; Allara, D.L. Spontaneously organized molecular assemblies. 3. Preparation and properties of solution adsorbed monolayers of organic disulfides on gold surfaces. *J. Am. Chem. Soc.* **1987**, *109*, 2358–2368.
27. Troughton, E.B.; Bain, C.D.; Whitesides, G.M.; Nuzzo, R.G.; Allara, D.L.; Porter, M.D. Monolayer films prepared by the spontaneous self-assembly of symmetrical and unsymmetrical dialkyl sulfides from solution onto gold substrates: structure, properties, and reactivity of constituent functional groups. *Langmuir* **1988**, *4*, 365–385.
28. Ulman, A. Self-assembled monolayer of 4-mercaptobiphenyls. *Acc. Chem. Res.* **2001**, *34*, 855–863.
29. Aslam, M.; Bandyopadhyay, K.; Vijayamohan, K.; Lakshminarayanan, V. Comparative behavior of aromatic disulfide and diselenide monolayers on polycrystalline gold films using cyclic voltammetry, STM, and quartz crystal microbalance. *J. Colloid Interface Sci.* **2001**, *234*, 410–417.
30. Clegg, R.S.; Reed, S.M.; Hutchison, J.E. Self-assembled monolayers stabilized by three-dimensional networks of hydrogen bonds. *J. Am. Chem. Soc.* **1998**, *120*, 2486–2487.
31. Fenter, P.; Eisenberger, P.; Li, J.; Camillone, N., III; Bernasek, S.; Scoles, G.; Ramanarayanan, T.A.; Liang, K.S. Structure of octadecyl thiol self-assembled on the silver(111) surface: An incommensurate monolayer. *Langmuir* **1991**, *7*, 2013–2016.
32. Ehler, T.T.; Malmberg, N.; Noe, L.J. Characterization of self-assembled alkanethiol monolayers on silver and gold using surface plasmon spectroscopy. *J. Phys. Chem., B* **1997**, *101*, 1268–1272.
33. Laibinis, P.E.; Whitesides, G.M. Self-assembled monolayers of *n*-alkanethiolates on copper are barrier films that protect the metal against oxidation by air. *J. Am. Chem. Soc.* **1992**, *114*, 9022–9028.
34. Lang, P.; Mekhalif, Z.; Rat, B.; Garnier, F. Self-assembled alkylthiols monolayers onto platinum; influence of the adsorbed oxygen. *J. Electroanal. Chem.* **1998**, *441*, 83–93.
35. Li, Z.; Chang, S.-C.; Williams, R.S. Self-assembly of alkanethiol molecules onto platinum and platinum oxide surfaces. *Langmuir* **2003**, *19*, 6744–6749.
36. Brito, R.; Tremont, R.; Feliciano, O.; Cabrera, C.R. Chemical derivatization of self-assembled 3-mercaptopropionic and 16-mercaptopentadecanoic acids at platinum surfaces with 3-aminopropyltrimethoxysilane: A spectroscopic and electrochemical study. *J. Electroanal. Chem.* **2003**, *19*, 53–59.
37. Ulman, A. Formation and structure of self-assembled monolayers. *Chem. Rev.* **1996**, *96*, 1533–1554.
38. Mrksich, M.; Whitesides, G.M. Patterning self-assembled monolayers using microcontact printing: A new technology for biosensors. *Trends Biotechnol.* **1995**, *13*, 228–235.
39. Chechik, V.; Crooks, R.M.; Stirling, C.J.M. Reactions and reactivity in self-assembled monolayers. *Adv. Mater.* **2000**, *12*, 1161–1171.
40. Davies, J.; Faulkner, I. Surface Plasmon Resonance—Theory and Experimental Considerations. In *Surface Analytical Techniques for Probing Biomaterial Process*; Davies, J., Ed.; CRC Press: Boca Raton, FL, 1996; 67–87.
41. Ivarsson, B.; Malmqvist, M. Surface Plasmon Resonance. Development and Use of Biacore Instruments for Biomolecular Interaction Analysis. In *Biomolecular Sensors*; Gizeli, E., Lowe, C.R., Eds.; Taylor & Francis: London, 2002; 241–268.
42. Hirata, I.; Hioki, Y.; Toda, M.; Kitazawa, T.; Murakami, Y.; Kitano, E.; Kitamura, H.; Ikada, Y.; Iwata, H. Deposition of complement protein, C3b, on mixed self-assembled monolayers carrying surface hydroxyl and methyl groups studied by surface plasmon resonance. *J. Biomed. Mater. Res.* **2003**, *66A*, 669–676.
43. Nelson, B.P.; Frutos, A.G.; Brockman, J.M.; Corn, R.M. Near-infrared surface plasmon resonance measurements of ultrathin films. 1. Angle shift and SPR imaging experiments. *Anal. Chem.* **1999**, *71*, 3928–3934.
44. Liebermann, T.; Knoll, W. Parallel multispot detection of target hybridization to surface-bound probe oligonucleotides of different base mismatch by surface-plasmon field-enhanced fluorescence microscopy. *Langmuir* **2003**, *19*, 1567–1572.

45. Ferretti, S.; Payner, S.; Russell, D.A.; Sapsford, K.E. Self-assembled monolayers: A versatile tool for the formation of bio-surfaces. *Trends Anal. Chem.* **2000**, *19*, 530–540.
46. Schaeferling, M.; Schiller, S.; Paul, H.; Kruschina, M.; Pavlickova, P.; Meerkamp, M.; Giammasi, C.; Kambhampati, D. Application of self-assembly techniques in the design of biocompatible protein microarray surfaces. *Electrophoresis* **2002**, *23*, 3097–3105.
47. Sigal, G.B.; Bamdad, C.; Barberis, A.; Strominger, J.; Whitesides, G.M. A self-assembled monolayer for the binding and study of histidine-tagged proteins by surface plasmon resonance. *Anal. Chem.* **1996**, *68*, 490–497.
48. Hyun, J.; Ahn, S.J.; Lee, W.K.; Chilkoti, A.; Zauscher, S. Molecular recognition-mediated fabrication of protein nanostructures by dip-pen lithography. *Nano Lett.* **2002**, *2*, 1203–1207.
49. Yousaf, M.N.; Houseman, B.T.; Mrksich, M. Using electroactive substrates to pattern the attachment of two different cell populations. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 5992–5996.

Shape Memory Metals

Marie Haidopoulos

Fatiha El Feninat

Diego Mantovani

Biomaterials and Bioengineering Laboratory, Laval University, Quebec City, Quebec, Canada

INTRODUCTION

Shape Memory Metals (SMM) have certain original properties, particularly an ability to return to a memorized shape in response to a simple change in temperature. The shape-memory effect was first observed in copper–zinc and copper–tin alloys by Greninger and Mooradian in 1938, but it was only in the early 1960s that Buehler and his colleagues discovered and patented Nitinol (N*ickel* T*itanium* N*aval* O*rdnance* L*aboratory*), a nickel–titanium alloy created in the Naval Ordnance Laboratory. Buehler suggested using SMMs in dentistry to make implants that would return to a pre-set shape. Much work has since been done to exploit the high potential of these materials for biomedical applications. They currently seem to hold the most promise in radiology, cardiovascular applications, urology, and other medical applications for use as prostheses, tissue connectors, and endovascular stents. In fact, SMMs have come into wide use because of their exceptional superelasticity, their shape memory, their good resistance to fatigue and wear, and their relatively good biocompatibility.

For biomedical applications, the most attractive SMM is Nitinol, an almost equiatomic nickel–titanium alloy. This SMM, however, is 55% nickel by weight and may thus have allergic, toxic, or carcinogenic effects. For short-term implantation, in-vitro and clinical data strongly support Nitinol as a safe biomaterial, at least as good as stainless steel or titanium alloys, such as Ti₆Al₄V. For long-term implantation, Nitinol's performance has yet to be proven and most of its interactions with living tissues and their effects when implanted have yet to be elucidated. Thus, the biocompatibility of SMMs is still doubtful. Biocompatibility is mainly surface-related, the surface being the part of the material to come into contact with cells, tissues, and body fluids. Many different processes have been developed to modify the surface in order to control the host's reaction to the material and to keep it from degrading in the body's inner environment. The main modifications involve insulating the material from the body by altering the nature of the surface while inducing or not inducing a specific response from

surrounding cells in the body. The metallic surface is treated to modify its topography, porosity, and chemical composition or it is coated with thin layers of polymer, ceramic, or other biocompatible and bioactive materials such as hydroxyapatite.

In the following sections, we will briefly explain the main actions and specific properties of SMMs, we will introduce some medical applications for SMMs, and, finally, we will review the importance of modifying the interfacial properties of SMMs for long-term implantation in humans.

PROPERTIES

The principal properties of SMMs are governed by phase transitions between the austenite (high-temperature phase) and the martensite (low-temperature phase). These properties are explained by a solid-state transition called the martensitic thermoelastic transformation (from the austenitic phase to the martensitic phase and vice versa). Four temperatures characterize this transition:

- Martensite start (M_s): the temperature at which the martensite layer spontaneously appears.
- Martensite finish (M_f): the temperature at which the entire specimen is transformed into martensite.
- Austenite start (A_s): the temperature at which the austenite layer spontaneously appears.
- Austenite finish (A_f): the temperature at which the entire specimen is transformed into austenite.

The behavior of the material is dictated by three parameters: stress (σ), strain (ϵ , or length (L)), and temperature (T) (Table 1). Changes in these parameters reflect changes in a class of thermomechanical cycles that are responsible for the specific characteristics of SMMs. The three parameters interrelate in a complex manner: It is thus difficult to predict the behavior of SMMs for any specific application. Most of the current applications use alloys that allow two of the three parameters to be varied, with the third one held

Table 1 Parameters that control the behavior of SMMs

Parameters	Definitions	Abbreviations
Stress (σ) (on the material)	Ratio between the force F and the surface S on which the force is applied: $\sigma = F/S$	F = Force
Strain (ε)	Observed relative displacement, for example, lengthening: $\Delta L = L - L_0$; $\varepsilon = \Delta L/L_0 = (L - L_0)/L_0$	L = Length
Temperature (T)	Temperature of the material's environment	T = Temperature

S

constant by a specific mix of alloy elements and thermomechanical treatment.

The relationship between phase transitions and temperature must be understood if we wish to control the properties of SMMs. Figure 1 presents an overview of SMM behavior when stress, strain, and temperature are varied:

- A one-way shape-memory effect (depicted in (1) and (2)), where changes in shape ($L \downarrow$) are regulated by the martensite-to-austenite transition.
- A two-way shape-memory effect, with a learning process by mechanical cycles in (1) and a memory effect apparent in (2), where changes in shape ($L \downarrow$ then \uparrow) are regulated by phase transitions (first, martensite-austenite; then, austenite-martensite).
- A superelastic effect (3), where the deformations ($L \downarrow$ then \uparrow) are regulated by phase transitions

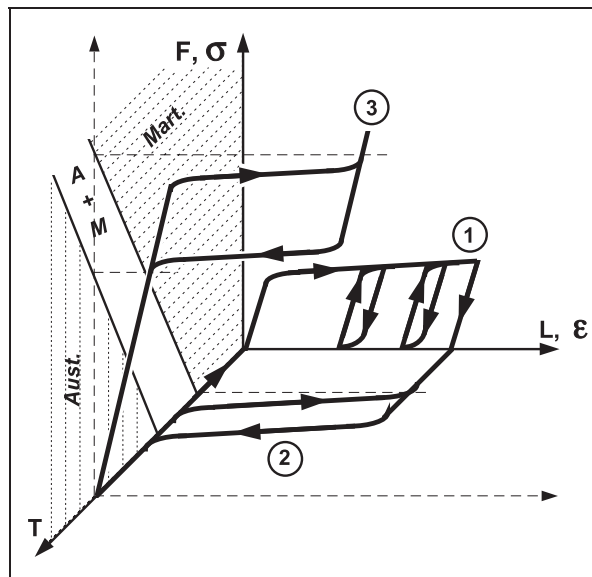


Fig. 1 Review of principal properties of SMMs: One-way memory (outer line of (1) and (2)); two-way memory with learning (1) and the memory effect itself (2); pseudoelastic effect (3); superthermic effect, shifted (2) toward non-zero force values; rubber-like effect (learning cycle loops (1)).

(first, austenite-martensite; then, martensite-austenite).

- A superthermic effect, which is seen when the curve (2) is displaced from the zero plane ($L, T, F = 0$) to non-zero values of force. In this case, external constraints replace the internal constraints accumulated during the learning process.
- A rubber-like effect is seen in the repeated mechanical learning cycles (1).

It has to be noted that under certain conditions (aging temperature, time and cold working) an R-phase can appear upon cooling and heating and prior to the martensitic and austenitic phase transformation, respectively. The presence of this phase is determined by a sharp increase in the material electrical resistance, essentially due to a distortion of the austenitic cubic lattice.^[1] The R-phase transition behaves much like the martensitic transition, i.e., it exhibits superelasticity and shape-memory effects, but much lower recoverable strain.

The observed effects are quite noticeable when the SMM is in a single phase (pure austenite or pure martensite). These effects require the presence of a specific phase, so that when the state is biphasal (austenite + martensite), the same effects occur but are less significant. A more detailed discussion of how SMMs work is provided in another review article.^[2]

BIOMEDICAL APPLICATIONS

Because SMMs can exhibit one-way or two-way shape memory and display a superelastic effect, they are particularly well suited for the design of surgical devices.^[3] In fact, devices with shape memory very often make surgery easier, shorter, and safer for both the patient and the surgeon. Extensive reviews of SMM biomedical applications have already been published.^[2,3] We will review only the latest devices and applications.

Atherosclerosis involves localized narrowing of blood vessels and, in extreme cases, blockage of blood flow to the artery. Atherosclerosis may affect any

artery, including the carotid, coronary, renal, abdominal aorta, iliac, femoral, and lower arteries. Traditionally, treatment would require by-passes and endarterectomies, but nowadays these medical interventions are no longer the only ones available. Dilations with or without arterectomy are frequently performed when the lesions are localized. Inserting a stent maintains the artery diameter and prevents blockage from recurring in the same area and/or prevents fibrous hyperplasia at the site of an anastomosis. Such stents are deployed up to their functional diameter by expanding a balloon made of stainless steel, tantalum, titanium, or SMMs. If SMMs are used, the heat of the bloodstream can expand the stent to its functional diameter, thus preventing any initial mechanical damage. Such stents are called self-expandable. However, it has to be underlined that the elastic recovery in blood contact does not necessarily require 37 degrees for final shape, and that most stent applications use SMMs' superelastic properties. Use of self-expandable stents in vascular endosurgery has increased exponentially over the last 10 years, despite a lack of scientific consensus over their long-term safety.^[4,5] Although their clinical effectiveness is well established, further study is needed to explain the superiority of shape-memory alloys over other materials.^[6] In particular, the effect of disturbances in blood flow, the compatibility of these materials with blood and their long-term biostability, their bio-functionality, the possible need for anti-coagulant treatments, and the patient's biological and immunological responses must be studied. SMMs hold considerable potential, especially in the manufacture of stents, for treating cardiovascular diseases. Moreover, self-expanding metallic stents have also proven to be effective in noncardiovascular procedures, such as palliation of nonoperable malignant dysphagia and esophagus obstruction,^[7] although tumor overgrowth and incomplete initial expansion have been observed in some cases.^[8]

A recently developed SMM device displays interesting properties for scoliosis correction.^[9] It is a rod that was initially engineered to bring about a gradual three-dimensional correction of a scoliotic deformity. In the martensitic phase, it can easily be twisted into the shape of the scoliotic spine and attached to the spine. After heating (austenitic phase), it will regain its original shape and rigidity. The desired shape can be pre-set by bending the rod beyond its reversible limit after heating. This design induces a bending moment and a torque that correct the lateral curvature and the axial rotation, respectively. A comparable study on monkeys found that a Nitinol rod could completely correct scoliosis with a Cobb angle of 43 degrees after four weeks.^[10] There were, however, significantly higher nickel concentrations in the urine (ten-fold increase).

In surgical correction of maxillofacial fractures, shape-memorizing devices are simple to implant and effective in stabilizing fracture surfaces. Memory materials not only simplify the surgical operation but also shorten the time of rehabilitation and promote rapid bone healing, as shown by Itró et al.^[11] The main advantage lies in their complete functional independence. They need no energy source for operation, unlike conventional medical materials. This property makes them much easier to monitor during implantation. Moreover, the superelastic properties are much more common than thermally induced shape recovery for *in vivo* devices such as maxillofacial fracture reduction or stabilization.

SMMs have found widespread use in various dental and orthodontic applications.^[12] Nitinol is ideal for a new form of intermittent orthodontic treatment because of its ability to exert a mild recovery force within a normal range of oral temperatures. This recovery force is reversible, durable, and reproducible.^[13]

SMMs are increasingly being used in traumatology, dentistry, and pharmacology as surgical tools, orthodontic arch wires, implantable delivery systems and other applications, according to a review by Moorleghem et al.^[14] In the area of managing ureteral obstruction due to malignancy or a recurrent benign disease, Kulkarni et al. have tested a Nitinol device for four years and have shown it to be significantly better than devices made from conventional materials.^[15]

Despite their interesting mechanical properties, SMMs may not be completely safe over a long period of time. The long-term implications of implanting them should be methodically assessed and tested with a view to getting unequivocal answers. Since certain elements and their compounds, such as nickel and chromium, increase tissue reactivity and sensitivity and may induce cytotoxic, carcinogenic, or mutagenic effects, we should develop new techniques to modify the SMM surface and make implantation safer by enhancing compatibility and by minimizing the possibility of nickel leaching into the human body. In the next section, we will review the major techniques for modifying the surface of SMMs. We will focus on Nitinol, the most widely used SMM.

SURFACE MODIFICATIONS

The state of a material's surface depends not so much on the phase state of the material as a whole, but rather on the way its surface has been treated. It is well known that the natural surface of Nitinol consists mainly of carbon contaminants, stable titanium oxides (TiO₂), and smaller amounts of nickel oxides (NiO and Ni₂O₃).^[16] Nitinol's resistance to corrosion is mainly due to a TiO₂-based surface layer and its specific

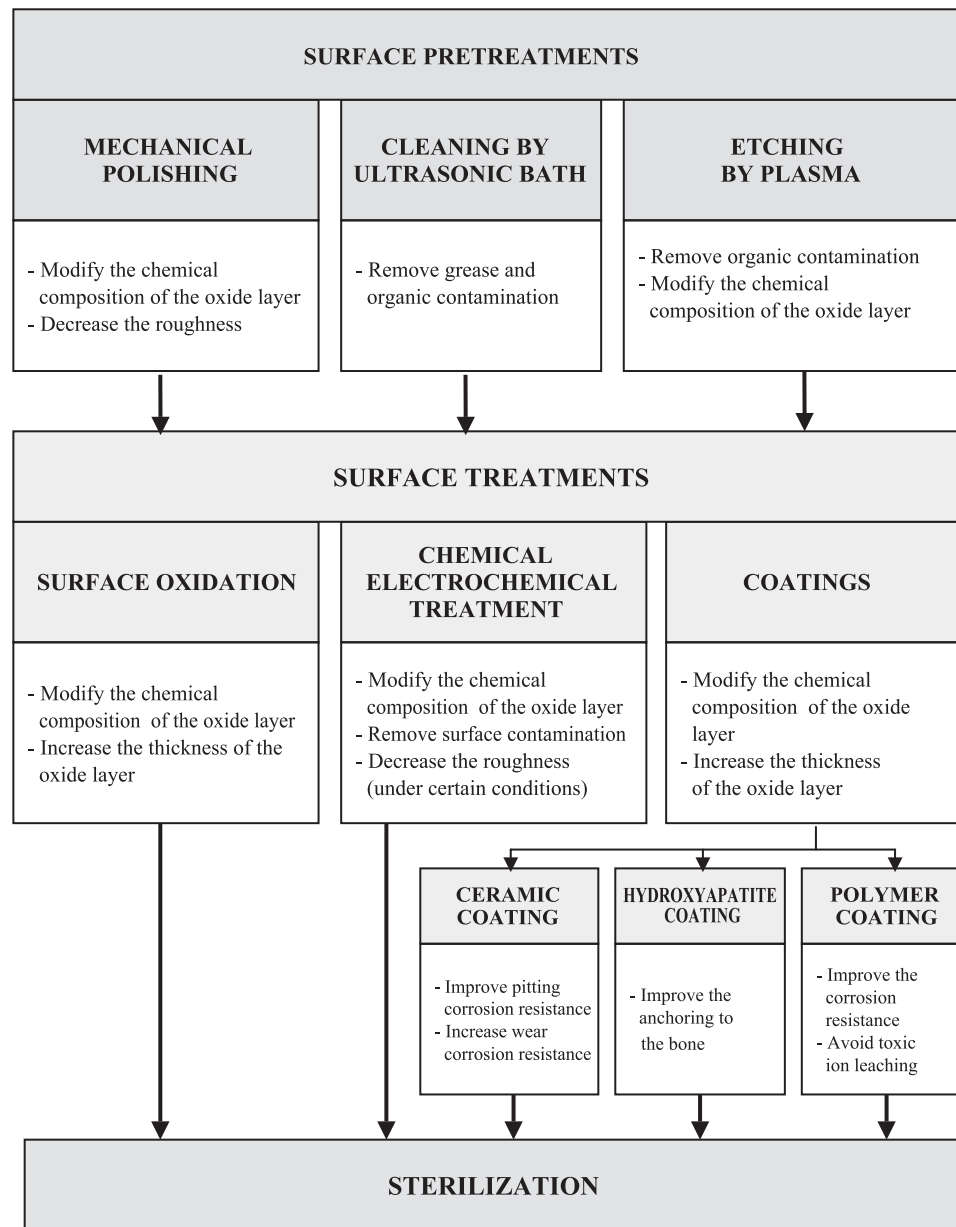


Fig. 2 Procedures for SMM surface modification before implantation: from pretreatment to sterilization.

properties, including its ability to form a calcium-phosphate layer when the device is implanted in living tissues.^[17] The surface chemistry and the amount of Ni may vary over a wide range depending on the preparation method.^[18]

For biomedical applications, a wide variety of Nitinol surface treatments are usually performed to improve corrosion resistance (and thereby prevent release of Ni ions), and/or to improve wear resistance or host-cell response. These treatments include (Fig. 2):

1. Preliminary treatments, including polishing, ultrasonic bath, and physical etching.

2. Surface oxidation.
3. Chemical and electrochemical treatments.
4. Surface coating (with ceramics, hydroxyapatite, or polymers).
5. Sterilization.

Preliminary Treatments: Polishing, Ultrasonic Bath, and Physical Etching

Surface chemistry and surface topography depend strongly on the way the material has been treated and this treatment will determine all of the interactions

between the implanted material and the body. In particular, corrosion resistance is reduced by a poor surface finish. The passive layer is affected by surface imperfections and impurities, which increase the rate of local corrosion.^[19] Consequently, for the most common applications, Nitinol surfaces are pretreated before any further surface treatment. Three preliminary treatments are commonly used: (mechanical or electrochemical) polishing, cleaning in an ultrasonic bath,^[20,21] and physical etching by argon plasma.^[21] For simple (i.e., flat) geometries, the surfaces are usually mechanically polished through a standard metallographic procedure by using a series of increasingly less abrasive SiC papers, followed by a micron alumina polishing powder to obtain a mirror finish.^[22] Roughness is reduced to an average vertical variation of around 10 nm after mechanical polishing, with typical Ti and Ni surface concentrations being around 10% and 25%, respectively, (or less).^[23] However, most of the commercially available implants are electrochemically polished, the geometries being more complexes. By this treatment, surface roughness results are decreased and the surface chemical composition modified. This point will be further discussed in a later section. After polishing, the surfaces are usually cleaned ultrasonically in cleaning reagents such as acetone, ethanol, and isopropanol to remove grease, organic contamination, and debris. Usually, this procedure is followed by rinsing with deionized water and finally drying with compressed air. The last pretreatment is argon plasma etching, to clean the surface before any further vacuum treatment, such as plasma polymerization.^[21]

Surface Oxidation

The formation of a passive TiO₂ film is responsible for the good corrosion resistance of Nitinol. Therefore, to improve Nitinol's corrosion resistance, an initial strategy is to optimize the quality of the oxide layer, e.g., by increasing or altering its chemical composition or by making it more uniform. The surface can be oxidized in various media. Some techniques promote the formation of highly adhesive passive films that should not crack as the device expands at body temperature. These techniques have been admirably reviewed by Shabalovskaya.^[24]

Oxidation processes may be dry or wet. In dry processes, the oxide layer is formed in an oxidative or steam atmosphere. Such Nitinol oxide films are essentially made of the dominant oxide TiO₂, although other Ti and Ni oxides, and elemental Ni, are present as well. At room temperature, the oxide layers are incomplete. At high temperatures, the surface is heterogeneous and composed of various oxides that make it more

vulnerable to corrosion. When oxide film coatings are produced by heat treatment of Nitinol in a low vacuum, they display better corrosion resistance and a higher potential for biocompatibility, and cracks do not appear after shape-memory deformation.^[25] At high temperatures, however, oxidation leads to an increase in elemental Ni at the surface, which could become a source of potential toxicity.^[23] Steam-treated surfaces have a higher Ni concentration than surfaces oxidized in boiling water, because Ni accumulates in the surface sublayers during the autoclaving procedure. The quality of treated surfaces is directly related to the conditions of steam treatment because high Ni concentrations and humidity will lead to the formation of Ni hydroxides during the steam process. This technique does not seem to improve surface biocompatibility (because of the presence of Ni compounds), and the high temperatures may also impair shape memory and superelasticity. Oxidation in boiling water, with preliminary chemical etching in an HF/HNO₃ solution, results in surfaces with a very low Ni concentration.^[24] The Ni concentration and the Ti/Ni ratio will strongly depend on how long the treatment lasts. The longer the boiling, the higher the Ni concentration at the surface. The boiling procedure includes selective Ni etching in a chemical solution, and the leaching of Ni into the water with simultaneous oxidation of the titanium results in a completely disordered surface. This amorphous surface state contributes to excellent corrosion resistance and biological response.

Chemical and Electrochemical Treatments

Chemical etching is generally used to remove surface contamination. One of the most popular solutions is a mixture of various ratios of HF, HNO₃, and H₂O. Shabalovskaya et al. have shown that this treatment produces a surface especially rich in elemental Ni, an inappropriate result if one wishes to increase biocompatibility.^[22] Another popular way to passivate Nitinol surfaces is to use nitric acid. Surfaces of polished Nitinol wires have been studied after heat treatment and passivation in a nitric acid solution.^[26] Improved corrosion resistance is obtained after heat treatment alone and heat treatment followed by passivation. The oxide composition seems to be more important than its thickness alone in improving corrosion resistance. The layer on both heat-treated and passivated wires consists primarily of TiO₂ with some NiO, but the passivated wires are enriched with Ni and NiO at their oxide-metal interface and with Ni at their outer layer. When the wires are immersed in a solution for a long period, the quantity of nickel released is lower for passivated samples and decreases over time.

Electropolishing represents a valid alternative to passivate Nitinol surfaces because it is already used

and well controlled for other materials, such as stainless steel. Electropolishing removes the native oxide layer and induces the formation of a new, more uniform, and smoother, almost nickel-free surface that greatly improves the implant's biocompatibility.^[18,27,28] Electropolished surfaces are extremely smooth and have better mechanical properties than surfaces chemically treated by an HF/HNO₃ solution or mechanically polished.^[29] Although electropolishing and nitric passivation have many advantages over other methods and hold much promise for commercial applications, the improvement they bring about is rather limited. At most, electropolishing lowers the passive current density by 1 order of magnitude.^[20]

Ceramic Coatings

Coating surfaces with a thin ceramic layer not only enhances resistance to pitting and crevice corrosion but also increases wear resistance as well. One of the most popular ways to deposit thin ceramic films, such as TiN and TiCN, on Nitinol surfaces is chemical vapor deposition.^[25] Although thin ceramic films exhibit better corrosion and wear resistance, they are less uniform and cohesive as a result of cracks that form during the extensive deformations induced by the shape-memory effect. Nitinol's shape-memory effect may produce a recoverable strain as high as 8%, which is much higher than the elastic strain possible with ceramic coatings. Several other techniques that form a thin hard layer on the alloy surface, such as ion implantation or laser treatment, are under investigation.

Ion nitriding at high temperatures can improve the surface properties of Nitinol. The treatment involves amorphization of the surface. Two layers are formed, an outer layer of TiN and an inner one of Ti₂Ni. Together, the two adjacent layers markedly increase surface hardness and surface wear-induced corrosion resistance.^[30] The coating is hard and strongly adherent to the substrate, but no tests involving substantial deformations have yet been performed. The shot-peening process can also improve the mechanical properties at the surface of SMMs.^[31] This technique and N⁺ ion implantation have been compared to surface microhardness and the coefficient of friction, because both techniques aim to produce an amorphous surface. The two treatments produce similar improvements in surface properties, with no direct effects on bulk properties. The presence of TiN precipitates within the surface oxide layer tends to make the ion-implanted layer harder, and less than 3 atomic percent of Ni is detectable at the surface after these treatments. The downside of ion implantation is that geometrically complex implants soon become a very complicated challenge.

A more recent technique is the Laser Melting Surface (LMS). It has the advantage of being very simple, allowing surface modification without affecting bulk properties. It forms a highly adhesive corrosion-resistant layer by immediately transferring the laser energy to the lattice and thereby quickly melting the near-surface regions. With this technique, the magnitude of the corrosion and the passivation current densities is reduced by an order of 2. LMS also induces thickening of the TiO₂-based oxide layer and, concurrently, nitrogen incorporation and surface homogenization. Nevertheless, it is very expensive and poorly suited for geometrically complex devices.^[32]

S

Hydroxyapatite Coatings

A smart Nitinol-based composite biomaterial has been developed with a hydroxyapatite coating.^[33] It allows the surrounding bone to colonize and encapsulate the implant. The main problem is the metal-hydroxyapatite interface rather than the bone-hydroxyapatite one. Although several treatments have been investigated, such as dip coating, electrophoretic deposition, and hot isostatic pressure, no significant solution has been found to date. Plasma spray coating greatly increases adhesion, probably by forming chemical bonds at the hydroxyapatite-Nitinol interface and by dissipating strain energy through stress-induced martensite formation and/or martensite reorientation.^[32] Despite having the appropriate mechanical properties, hydroxyapatite tends to dissolve when in contact with body fluids, probably due to its relative miscibility with them.^[23]

Polymer Coatings

To provide a protective barrier that will contain potential release of Ni and so protect the implant from corrosion and the patient from toxic contamination, the SMM can be evenly coated with a thin polymer film. Polymers seem to make a good coating for shape-memory material for three main reasons. First, they are more biocompatible than metals and therefore better suited for being in contact with living tissues. Second, they tolerate a wide range of elastic deformation and can bend with the shape-memory deformation without cracking. Finally, it is possible to create a plasma-deposited polymer layer that is uniform and adheres well to the metal substrate. Thus, an evenly deposited polymer film may provide implants with outstanding corrosion resistance.

Villiermaux et al. have studied the improvement in the corrosion resistance of Nitinol staples by applying an RF plasma polymerized-polytetrafluoroethylene

(PPTFE) coating.^[21] This coating had very good adhesion but lacked cohesiveness. It did, however, efficiently protect the substrate from pitting and decreased the corrosion rate. A PPTFE film can lower the current corrosion density by 1 order of magnitude although it cannot reduce the passive current density. The deposited film must be uniform. When damaged, its corrosion resistance can easily decrease in comparison to that of untreated samples. Moreover, the coating can bend with the bending of a substantial shape-memory deformation without cracking. Finally, plasma treatment is appropriate for geometrically complex implants. In contrast to fluorine-containing monomers, organosilicon is more environmentally suitable and less corrosive for vacuum equipment. Yang et al. have studied DC plasma polymerized hexamethyldisilazane (PHMDSN) coatings with a view to improving the corrosion resistance of Nitinol.^[20] At DC voltages higher than 1000 volts, PHMDSN coatings protect against corrosion in Ringer's solution and can be deformed by up to 2%. The corrosion current density of Nitinol can be lowered by 4 orders of magnitude—four times more than with an RF PPTFE coating.^[21] It is worth noting that before each plasma deposition, for both PPTFE and PHMDSN coatings, the samples were pretreated with successive mechanical polishing, degreasing in an ultrasonic bath, rinsing with distilled water, and drying with a stream of compressed air, and finally cleaning with argon plasma under vacuum. Plasma coatings for shape-memory alloys seem to be a promising and efficient way to protect implants from the severe environment of living human tissues.

Sterilization

Sterilization is the stage that all materials must undergo before being implanted into the human body. It affects the surface finish and corrosion resistance of metallic devices and, as with any surface treatment, we need to investigate effects and assess potential consequences. Tabrizian et al.^[27] and Thierry et al.^[28] have studied the effects of sterilization on Nitinol surfaces that had been mechanically polished prior to electropolishing. After sterilization by ethylene oxide, a common technique for sterilizing metallic implants, no changes were observed in the corrosion behavior of electropolished samples. The improvement in corrosion resistance due to electropolishing still remained even after sterilization by ethylene oxide. But some surface modifications were reported. After sterilization, the oxide layer was thicker and its nickel content higher and not uniform. Other sterilization processes, such as steam autoclave and peracetic acid sterilization, have been tested and shown to modify the corrosion resistance of Nitinol.

Other Treatments

Some other treatments can provide a Nitinol surface with very specific and biomedically useful properties. Phosphorus-ion implantation can be used to make radioactive stents.^[34] The purpose is to prevent cells from dividing, forming tissue hyperplasia, and thereby inducing restenosis. Phosphorus-ion implantation can be used to improve the hemocompatibility of Nitinol.

Gold coating can be used to improve the radiopacity of Nitinol medical devices, such as filters and stents. A thick gold layer adheres well to the substrate, and should be no thicker than around 4 microns to allow elastic deformation of the material. An electropolishing pretreatment before gold coating is recommended. The downside is that hydrogen uptake must be controlled to avoid Nitinol embrittlement.^[35]

CONCLUSIONS

Because conventional materials do not exhibit memory-related and superelastic properties, SMMs have attracted much interest. They offer new possibilities in the designing of new devices for unconventional purposes. Many applications have already been developed and we strongly believe that many more are on the way. With their unconventional properties, these alloys have made possible some of the boldest applications since the 1970s while clearing the way for further technological advances. All potential applications, however, must meet the social requirements of our age. First, they must be inexpensive, especially for medical/surgical uses. Second, they must be safe for human implantation. All implants must be biocompatible, biofunctional, and biodurable over the short, medium, and long term, depending on the application. Until now, no study has unequivocally proven that Nitinol devices will be completely safe for the human body and still functional after long-term implantation.^[36] If we wish to go on using SMMs because of their special mechanical properties, we will have to be sure of their safeness for the patient's health. We will probably have to focus on understanding their surface nature, tailoring their properties, and developing specific surface treatments or coatings that will protect the human body from any potential toxic contamination.

ACKNOWLEDGMENTS

This work has been funded by NSERC-Canada and FQRNT-Quebec.

ARTICLES OF FURTHER INTEREST

Artificial Muscles; Distal Protection Devices; Nickel and Nickel Alloys; Orthodontic Wires; Stents

REFERENCES

- Uchil, J.; Mohanchandra, K.P.; Mahesh, K.K.; Kumara, K.G. Thermal and electrical characterization of R-phase dependence on heat-treat temperature in Nitinol. *Phys. Rev., B, Condens. Matter* **1998**, *253*, 83–89.
- Mantovani, D. Shape memory alloys for biomedical applications. *J. Min. Met. Mater. Soc.* **2000**, *52* (10), 36–44.
- El Feninat, F.; Laroche, G.; Fiset, M.; Mantovani, D. Shape memory materials for biomedical applications. *Adv. Eng. Mater.* **2002**, *4* (3), 91–104.
- Blum, U. Endoluminal stent-grafts for infrarenal abdominal aortic aneurysms. *N. Engl. J.* **1997**, *336*, 13–20.
- Drugacz, J.; Lekston, Z.; Morawiec, H.; Januszewski, K. Use of TiNiCo shape-memory clamps in the surgical treatment of mandibular fractures. *J. Oral Maxillofac.* **1995**, *53*, 665–671.
- Thurnher, S.A.; Grabenwöger, M. Endovascular treatment of thoracic aortic aneurysms. *Eur. Radiol.* **2002**, *12*, 1370–1387.
- Christie, N.A.; Buenaventura, P.O.; Fernando, H.C.; Nguyen, N.T.; Weigel, T.L.; Ferson, P.F.; Luketich, J.D. Results of expandable metal stents of malignant oesophageal obstruction in 100 patients: Short-term and long-term follow up. *Ann. Thorac. Surg.* **2001**, *71* (6), 1801–1802.
- May, A.; Selmaier, M.; Hochberger, J.; Gossner, L.; Muhldorfer, S.; Hahn, E.G.; Ell, C. Memory metal stents for palliation of malignant obstruction of the oesophagus and cardia. *Gut* **1995**, *37* (3), 309–313.
- Veldhuizen, A.G.; Sanders, M.M.; Cool, C.J. A scoliosis correction device based on SMM. *Med. Eng. Phys.* **1997**, *19* (2), 171–179.
- Mastumoto, K.; Tajima, N.; Kuwahara, S. Correction of scoliosis with shape-memory alloys. *J. Jpn. Orthop. Assoc.* **1993**, *67* (4), 267–274.
- Itro, A.; Garau, V.; Tararo, G. Experience with a rigid fixation device in maxillofacial surgery using shape-memory clips. *Minerva Stomatol.* **1997**, *46* (7–8), 381–389.
- Iijima, M.; Endo, K.; Ohno, H.; Yonekura, Y.; Mizoguchi, I. Corrosion behaviour and surface of orthodontic Ni–Ti alloy wires. *Dent. Mater.* **2001**, *20* (1), 103–113.
- Fukuizumi, M.; Kakigawa, H.; Kozono, Y. Utility of Ni–Ti shape memory orthodontic wire. *Dent. Mater.* **1999**, *18* (4), 413–424.
- Van Moorleghem, W.; Chandrasekaran, M.; Reynaerts, D.; Peirs, J.; Van Brussel, H. Shape memory and superelastic alloys: The new medical materials with growing demand. *Biomed. Mater. Eng.* **1998**, *8* (2), 55–60.
- Kulkarni, R.; Bellamy, E. Nickel–titanium shape memory alloy Memokath 051 ureteral stent for managing long-term ureteral obstruction: 4-year experience. *J. Urol.* **2001**, *166* (5), 1750–1754.
- Ryhänen, J. Biocompatibility of nitinol. *Min. Invas. Ther. Allied Technol.* **2000**, *9* (2), 99–106.
- Wever, D.J.; Veldhuizen, A.G.; De Vries, J.; Busscher, H.J.; Uges, D.R.A.; Van Horn, J.R. Electrochemical and surface characterization of a nickel–titanium alloy. *Biomaterials* **1998**, *19* (7–9), 761–769.
- Trigwell, S.; Hayden, R.D.; Nelson, K.F.; Selvaduray, G. Effects of surface treatment on the surface chemistry of Nitinol alloy for biomedical applications. *Surf. Interface Anal.* **1998**, *26*, 483–489.
- Es-Souni, M.; Es-Souni, M.; Fisher-Brandies, H. On the properties of two binary NiTi shape memory alloys. Effects of surface on the corrosion behaviour and in vitro biocompatibility. *Biomaterials* **2002**, *23*, 2887–2894.
- Yang, M.-R. DC-plasma-polymerized hexamethyldisilane coatings of an equiatomic TiNi shape memory alloy. *Surf. Coat. Technol.* **2000**, *127*, 274–281.
- Villiermaux, F.; Tabrizian, M.; Yahia, L.H.; Czeremuszkin, G.; Piron, D.L. Corrosion resistance improvement of NiTi osteosynthesis staples by plasma polymerized tetrafluoroethylene coating. *Biomed. Mater. Eng.* **1996**, *6* (4), 241–254.
- Shabalovskaya, S.A.; Anderegg, J.W. Surface spectroscopic characterization of TiNi nearly equiatomic shape memory alloys for implants. *J. Vac. Sci. Technol., A* **1995**, *13* (5), 2624–2632.
- Shabalovskaya, S.A. Surface, corrosion and biocompatibility aspects of Nitinol as an implant material. *Biomed. Mater. Eng.* **2002**, *12*, 69–109.
- Shabalovskaya, S.A. Biological aspects of TiNi alloy surfaces. *J. Phys., III* **1995**, *18*, 1199–1204.
- Kimura, H.; Sohmura, T. Surface coating on TiNi shape memory implant alloys. *J. Osaka Univ. Dent. Sch.* **1987**, *27*, 211–223.
- O'Brien, B.; Carroll, W.M.; Kelly, M.J. Passivation of nitinol wire for vascular implants—A demonstration of the benefits. *Biomaterials* **2002**, *23* (8), 1739–1748.
- Tabrizian, M.; Thierry, B.; Savadogo, O.; Yahia, L.H. Surface characteristics of sterilized electropolished NiTi shape memory alloy as biomaterials. *Proc. SPIE Int. Soc. Opt. Eng.* **1999**, *3670*, 106–114.
- Thierry, B.; Tabrizian, M.; Trepanier, C.; Savadogo, O.; Yahia, L.H. Effect of surface treatment and sterilization processes on the corrosion behavior of NiTi shape memory alloy. *J. Biomed. Mater. Res.* **2000**, *51* (4), 685–693.
- Miao, W.; Mi, X.; Zhu, M.; Guo, J.; Kou, Y. Effect of surface preparation on mechanical properties of a NiTi alloy. *Mater. Sci. Forum* **2002**, *394–395*, 173–176. (Shape Memory Materials and Its Applications).
- Starosvetsky, D.; Gotman, I. TiN coating improves the corrosion behavior of superelastic NiTi surgical alloy. *Surf. Coat. Technol.* **2001**, *148* (2–3), 268–276.
- Green, S.M.; Grant, D.M.; Wood, J.V. XPS characterisation of surface modified Ni–Ti shape memory alloy. *Mater. Sci. Eng., A* **1997**, *A224* (1–2), 21–26. (Structural Materials: Properties, Microstructure and Processing).
- Villiermaux, F.; Tabrizian, M.; Yahia, L.H.; Meunier, M.; Piron, D.L. Excimer laser treatment of NiTi



- shape memory alloy biomaterials. *Appl. Surf. Sci.* **1997**, *109–110*, 62–66.
33. Wen, X.; Wang, X.; Zhang, N. Microrough surface of metallic biomaterials: A literature review. *Biomed. Mater. Eng.* **1996**, *6*, 173–189.
 34. Zhao, X.K.; Cai, W.; Tian, Y.; Zhao, L. Microstructure and hemocompatibility of a phosphorus ion-implanted TiNi shape-memory alloy. *Mater. Sci. Forum* **2002**, *394–395*, 153–156. (Shape Memory Materials and Its Applications).
 35. Steegmueller, R.; Wagner, C.; Fleckenstein, T.; Schuessler, A. Gold coating of nitinol devices for medical applications. *Mater. Sci. Forum* **2002**, *394–395*, 161–164. (Shape Memory Materials and Its Applications).
 36. Riepe, G.; Heintz, C.; Kaiser, E.; Chakfe, N.; Morlock, M.; Delling, M.; Imig, H. What can we learn from explanted endovascular devices? *Eur. J. Vasc. Endovasc. Surg.* **2002**, *24* (2), 117–122.

Silicones

Kenneth J. Wynne

Department of Chemical Engineering, Virginia Commonwealth University, Richmond, Virginia, U.S.A.

James M. Lambert

SiTech, LLC, Richmond, Virginia, U.S.A.



INTRODUCTION

Silicones, or properly poly(diorganosiloxane)s, describe a class of polymeric materials with high chemical stability that are widely used in biomedical applications. The chemistry, properties, surface science, and processing of silicone materials used in these applications are summarized herein. References within provide the reader access to the vast literature of poly(diorganosiloxane) materials.

BACKGROUND

It is important initially to distinguish the terms silicone and silicon^[1] and to note that silicones (polymers) contain the element silicon. The silicone designation comes from Kipping, who in the early part of the 20th century knew of the -O-Si- chain structure but anticipated facile conversion to structures analogous to ketones.^[2] This supposition proved incorrect, but the term silicones remains. In the context of similar terminology, silica (silicon dioxide, SiO_2) is used as a reinforcing filler for silicones. Amorphous, high-surface-area fumed silica is often the filler of choice.

The chemistry and materials science of silicones is a relatively mature area. For the basic chemistry of silicones, the chapter by Barry in the 1962 Stone and Graham text *Inorganic Polymers* is still worthwhile reading.^[2] Another *Inorganic Polymers* contains a chapter on silicones referencing many review articles,^[3] and a recent symposium volume on silicones and silicone-modified materials is available.^[4] The LaVier review^[1] is a useful introduction, as is the article by Arkles.^[5] Chapter 2 (Silicone Chemistry) in the NAS report *Safety of Silicone Breast Implants* gives good insight into the evolution of chemistry and processing for silicone materials.^[6] For implant medical devices, biocompatibility and biodurability are determined by chemistry, processing, and fabrication. Chapter 3 (Implant Catalog) in the NAS report just cited presents detailed information on the evolution of implant construction and the rationale for many changes.

PDMS STRUCTURE

The most commonly encountered poly(diorganosiloxane) structure is poly(dimethylsiloxane) (PDMS), which contains the $\text{-Me}_2\text{SiO-}$ repeat unit (Fig. 1).

The space-filling model of a PDMS chain with 12 repeat units (Fig. 2) was generated using Chem3DTM and partly energy-minimized with an MM2 calculation. This representation allows visualization of the effect of lack of substituents on oxygen and of the larger size of Si in the main chain compared to carbon chain polymers. In addition, the Si–O–Si bond angle is larger (143°) than the tetrahedral (109°) C–O–C or C–C–C bond angle.^[7] This combination of larger Si–O bond distance and Si–O–Si bond angle accounts for the lack of restriction of rotation about the Si–O–Si bond and the extremely low glass transition temperature (T_g) of PDMS elastomers (-120°C). The low T_g , nonpolar character (due to the masking of polar Si–O bonds by the methyl groups), and high thermal stability are hallmarks of the PDMS structure.

NOMENCLATURE

The nomenclature and shorthand notation used to describe silicones have been historically useful, if not highly rigorous. For the purposes of this article, the frequently employed M, D, T, Q notation will be used to denote Si bound to 1, 2, 3, and 4 oxygens, respectively. It can be assumed that methyl groups are present on the silicon atom unless indicated otherwise by superscript notation. Thus, a D silicon that has two Si–O bonds and methyl and vinyl groups is denoted D^{Vi} . Further examples are octamethyltrisiloxane and octamethylcyclotetrasiloxane, written as MDM and D_4 , respectively. Table 1 provides shorthand for some common silicones encountered as biomaterials.

MACROMONOMERS AND LOW-MOLECULAR-WEIGHT FLUIDS: SYNTHESIS

Equilibration reactions are used industrially to make liquid polymer precursors or macromonomers, or for

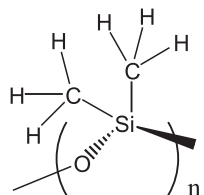


Fig. 1 A representation of the $-\text{Me}_2\text{SiO}-$ repeat unit in poly(dimethylsiloxane).

thermosetting elastomeric materials.^[2] Equilibration reactions are also used to make low-molecular-weight (MW) fluids such as trimethylsilyl-terminated PDMS (TMS-PDMS), known as silicone oil. Macromonomers and low-MW fluids are both useful in medical applications, but in completely different ways.

Macromonomers

A macromonomer is an oligomer or higher-MW polymer that has at least two functional groups. The term telechelic is also used for large (tele) difunctional (chelice) oligomers or low-molecular-weight polymers.^[8,9] Common telechelic functional groups include silanol ($\text{Si}-\text{OH}$), acetoxy ($\text{Si}-\text{O}-\text{C}(\text{O})\text{CH}_3$), vinyl ($\text{Si}-\text{CH}=\text{CH}_2$), and aminopropyl ($\text{Si}-\text{CH}_2\text{CH}_2\text{CH}_2\text{NH}_2$). An example of equilibration for the synthesis of terminally functionalized vinyl polymer, $\text{M}^{\text{Vi}}\text{D}_n\text{M}^{\text{Vi}}$, is shown in Scheme 1.

In this example, 1,3-divinyltetramethyldisiloxane is used as a chain stopper, so that vinyl groups occur as end groups in the macromonomer. The ratio of D_4 to functional disiloxane controls the molecular weight of the macromonomer. Reports on the synthesis and characterization of hydrosilation- and condensation-cured elastomers sometimes utilize macromonomers of low (2K–20K) MW.^[10] Macromonomers used in the formation of elastomers for biomedical applications generally have average molecular weights in the range of 25,000 to 400,000 or higher. Fluid viscosities are often cited by composition formulators as a measure of molecular weight.^[11]

The use of D_4^{Vi} , D_4 , and hexamethyldisiloxane gives D^{Vi} , or side-chain vinyl groups substituted randomly

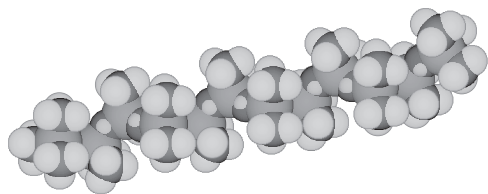


Fig. 2 Space-filling model of a PDMS chain with 12 silicon groups. The chain ends with a trimethylsilyl moiety.

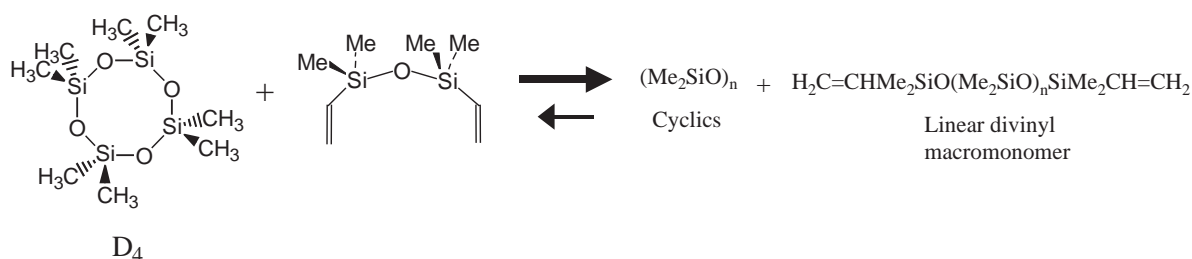
Table 1 Functional groups commonly encountered in silicones used for biomaterials

Moiety	R	Symbol
$\text{R}(\text{CH}_3)_2\text{SiO}-$	H, Vi, OH, OAc	M^{R}
$\begin{array}{c} \text{CH}_3 \\ \\ -\text{O}-\text{Si}-\text{O}- \\ \\ \text{R} \end{array}$	H, Vi	D^{R}
$\begin{array}{c} \text{CH}_3 \\ \\ -\text{O}-\text{Si}-\text{O}- \\ \\ \text{R} \end{array}$	$-\text{CH}_2\text{CH}_2\text{CF}_3$	D^{f}
$\begin{array}{c} \text{C}_6\text{H}_5 \\ \\ -\text{O}-\text{Si}-\text{O}- \\ \\ \text{R} \end{array}$	CH_3	D^{MePh}
$\begin{array}{c} \text{C}_6\text{H}_5 \\ \\ -\text{O}-\text{Si}-\text{O}- \\ \\ \text{R} \end{array}$	C_6H_5	D^{Ph}

along the main chain (Scheme 2). The average degree of functionalization is controlled by the mole ratio of D_4^{Vi} to D_4 . Side-chain vinyl may be used to change the crosslinking architecture or to obtain higher network crosslink density.

The small backward arrow in Schemes 1 and 2 indicates that there is always a significant concentration of lower-molecular-mass linear and cyclic species present.^[12] Numerous studies of these cyclic and linear species have been conducted. One recent review provides a summary of this work and a number of references for further reading.^[13] Low-molar-mass linears and cyclics are usually removed by distillation. The weight percent of low-molar-mass species varies considerably depending on the particular commercial resin involved. Polydimethylsiloxane species are very soluble in liquid and supercritical CO_2 .^[14] An innovative way to fractionate and purify PDMS macromonomers is by supercritical CO_2 extraction.^[15] For biomedical applications, the special purification methods used to reduce volatile constituents will be discussed later. Depending on the application, the presence of cyclics (nonfunctional oils) can be troublesome.

Macromonomers made by equilibration reactions have broad molecular weight distributions. Several methods may be used to make narrow-MW distribution macromonomers. Disilanol-terminated PDMS macromonomers can be prepared from hexamethylcyclotrisiloxane using a modification of Lee's method.^[16] This method utilizes a benzyltrimethylammonium bis(o-phenylenedioxy)phenylsiliconate catalyst, with dimethyl sulfoxide (DMSO) and water as cocatalysts. Disilanol-terminated PDMS macromonomers are converted to divinyl analogs by reaction with vinyltrimethylchlorosilane in toluene, using pyridine as an HCl acceptor. PDMS macromonomer polydispersities (PDs) of about 1.1 are obtained, compared to PDs of 2.0 or greater for commercially available macromonomers.^[17] Combining low- and high-MW macromonomers improves the mechanical properties of silicone elastomers.^[10]



Scheme 1

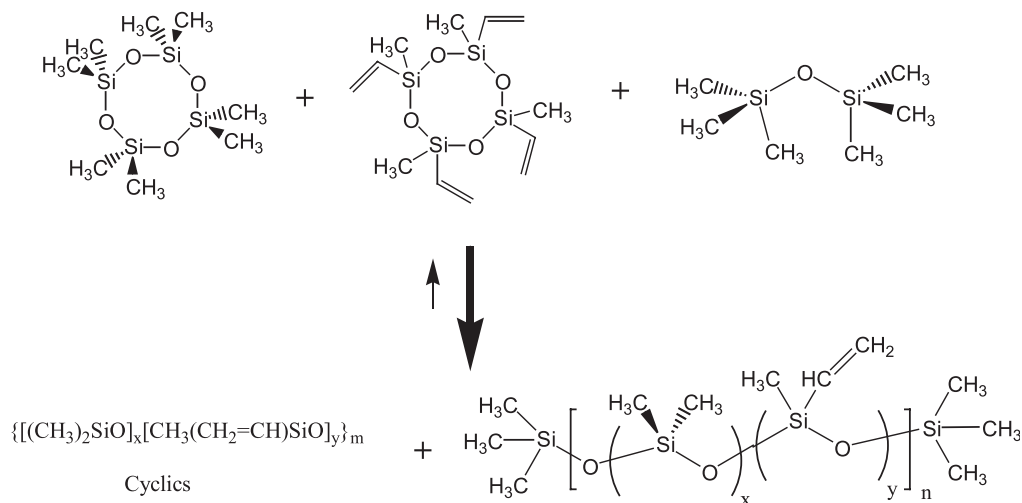
Another class of macromonomers made via equilibration reactions incorporates (trifluoropropyl) methylsiloxane $[-(\text{CF}_3\text{CH}_2\text{CH}_2)(\text{CH}_3)\text{SiO}-]$ groups. With increased fluoro-organo content, elastomers derived from these macromonomers become increasingly oleophobic. The starting material for $-(\text{CF}_3\text{CH}_2\text{CH}_2)(\text{CH}_3)\text{SiO}-$ containing macromonomers is the cyclic trimer D_3^f , $[(\text{CF}_3\text{CH}_2\text{CH}_2)(\text{CH}_3)\text{SiO}]_3$. A variety of polymerization catalysts may be used,^[18] some of which differ from those employed in dimethylsiloxane polymerizations. The choice of catalyst and reaction conditions is important to obtaining the desired polymer.^[19] Brown and Carmichael reported that in the case of reactions carried out between 333K and 423K using 100% methyltrifluoropropyl siloxane cyclic trimer and potassium hydroxide catalyst, the concentration of linear polymer was only approximately 14% in the bulk equilibrate, with cyclic oligomers making up the remaining 86%.

Low-Molecular-Weight Fluids

Equilibration reactions are also used to make low-MW fluids such as silicone oil, that is, trimethylsilyl-terminated

poly(dimethylsiloxane), $(\text{Me}_3\text{SiO}(\text{Me}_2\text{SiO})_n\text{SiMe}_3)$. Because of high thermal and chemical stability, silicone oil finds applications in numerous mechanical, electrical and chemical processing applications. TMS-PDMSs are chemically nonfunctional macromolecules that are used as swelling agents in silicone gels. In the same way that water is absorbed by hydrophilic polymers to generate soft hydrogels, lightly crosslinked PDMS (addition cure, *vide infra*) absorbs TMS-PDMSs to form a soft PDMS-gel.

Another class of low-MW fluids used in biomedical applications incorporates (trifluoropropyl) methylsiloxane $[-(\text{CF}_3\text{CH}_2\text{CH}_2)(\text{CH}_3)\text{SiO}-]$ groups. These fluids are similar in nature to the trifluoropropyl-containing macromonomers described above, except that they typically incorporate trimethylsilyl chain ends. The mole fraction of (trifluoropropyl)methylsiloxane groups can range from as low as a few percent to 100%, depending upon application requirements. For certain applications, many properties of fluoro-silicone fluids are superior to PDMS fluids in particular lubricity, solvent resistance, and performance across a wide range of temperatures.^[20]



Scheme 2 Synthesis of the co-dimethyl-methylvinyl macromonomer with trimethylsilyl terminal groups.

THERMOSETTING ELASTOMERS

The term thermoset is used to describe polymer chain network structures formed by covalent bonds through chemical reactions.^[21] Thermosets may have one or two components depending on the crosslinking chemistry. For thermosetting elastomers, at least one starting component must have a low glass-transition temperature, typically below -50°C . For PDMS elastomers, a filler is incorporated to improve mechanical properties. For convenient mixing, starting components must be flowable liquids, despite the presence of added filler. These considerations set the MW range of PDMS macromonomers produced by the equilibration reactions discussed earlier. For thermosetting PDMS elastomers, macromonomer molecular weights are usually in the range of 25,000 to 400,000 amu or higher.

Because of the extraordinary flexibility of the PDMS chain, even 10,000,000-centipoise macromonomers with degrees of polymerization of 5,000 or higher are flowable liquids (but slow!). Such high-molecular-weight macromonomers are employed in elastomers prepared by hydrosilation and peroxide cures discussed in the next sections. A wide range of macromonomer molecular weights is commercially available. Different macromonomer molecular weights may be employed in thermosets (and the thermoplastics that will be discussed later) to study surface properties or control mechanical properties.^[10,22]

Hydrosilation Cure

Vinyl functionalized telechelics are important in hydrosilation-cured (addition-cured) thermosetting elastomers for biomedical applications. There are several ways that network formation can be accomplished utilizing hydrosilation, which is the addition of a hydrosilane, i.e., a Si-H-containing species, to the vinyl group. One approach is shown in Scheme 3, where the Si-H functionalities are pendant to the main chain. Permutations on this approach, using functionality that is present as both end and pendant groups, are readily envisaged. Because the crosslinking reaction involves the formation of Si-C and C-H bonds, networks generated by hydrosilation cure are extremely chemically stable, particularly to hydrolysis, and well-suited to many biomedical applications. It has been reported that hydrosilation-cured coatings show no drop in receding contact angle or measurable mass loss after two months of immersion in water.^[23,24]

In carrying out addition cure, close attention must be paid to the Si-H-to-vinyl ratio. For a given catalyst concentration, the optimum Si-H-to-vinyl ratio is determined empirically, due to broad molecular weight distributions and side reactions. One approach to

determining the optimum ratios is minimization of soluble sol.^[17] Such a procedure minimizes non-network species that could leach.

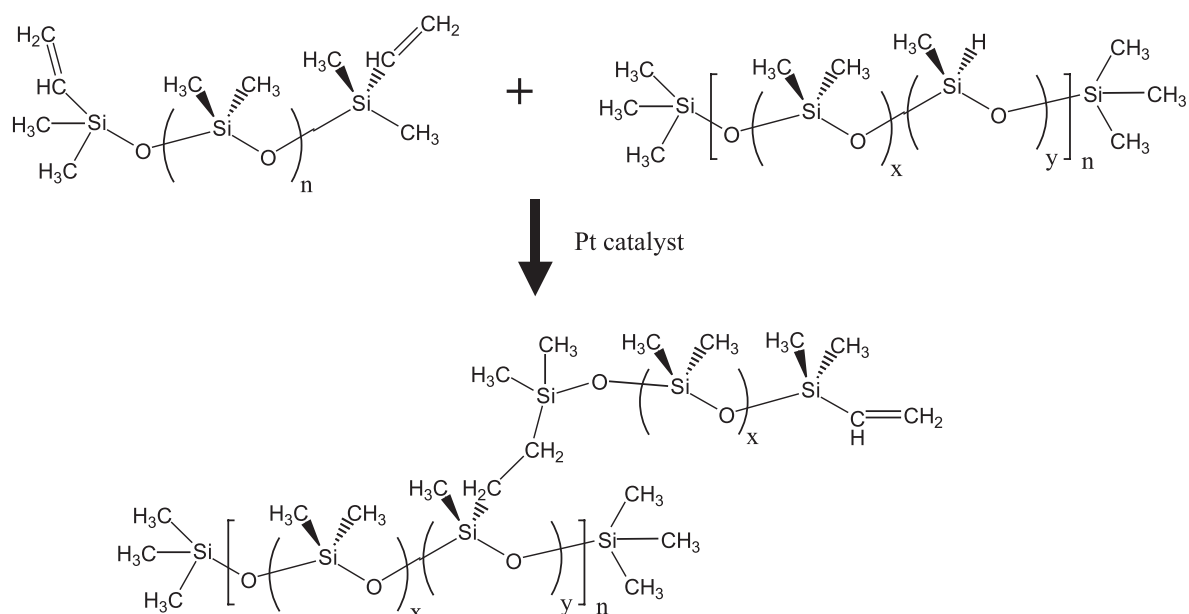
Often, a divinyl(tetramethyldisiloxane) complex of platinum is used for hydrosilation.^[25] At ambient temperature or above, solutions of this catalyst may be unstable with respect to Pt(0) colloid formation. Cold storage or periodic replacement are two ways to keep the catalyst active. The catalyst remains in the network structure, though at a very low level. A recent study indicated that there is no evidence that residual Pt causes enhanced allergic reactions in women with breast implants.^[26]

A new platinum(0)-carbene complex reported by Markó et al. reduces the tendency to form platinum colloids and gives improved isomer selectivity.^[27] However, the authors report that the reactivity of this catalyst is lower relative to the vinyl-platinum complex. There have as yet been no specific reports of the use of this platinum(0)-carbene catalyst in biomedical silicone applications.

Peroxide Cure

Silicone elastomer networks can also be formed via free radical crosslinking processes catalyzed by organic peroxides. This was the first method invented for silicone elastomer synthesis, and it is still in use today for certain biomedical applications. As shown in Scheme 4, the process involves cleaving the oxygen-oxygen bond of an aryl peroxide catalyst at elevated temperatures to produce a free radical (Ar-O^{\bullet}). This free radical reacts with an organic group in the silicone polymer to produce a carbon-centered free radical (Si-R^{\bullet}), which then reacts further to yield a crosslink with another silicone polymer chain via carbon-carbon bond formation. A detailed description of the mechanism(s) of this process can be found elsewhere.^[28] The vast majority of peroxide-cured elastomers for medical applications utilize high-molecular-weight PDMS with a small amount of vinyl functionality (usually at the chain termini) and the peroxide catalyst bis-2,4-dichlorobenzoyl peroxide. The cure temperatures for this system are somewhat lower than for those employing other peroxide catalysts or with PDMS containing no vinyl functional groups. However, the product of bis-2,4-dichlorobenzoyl peroxide catalyst decomposition is 2,4-dichlorobenzoic acid, which must be removed in a high temperature postbaking procedure to avoid "blooming" of the material on the surface of the elastomer over time. This is responsible, in part, for the trend toward use of silicones that employ other cure systems.

Although the physical properties of silica-filled peroxide-cured elastomers differ in some respects from hydrosilation-or condensation-cured elastomers,^[29] they

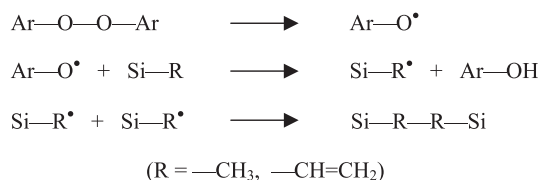


Scheme 3 Network formation through the reaction of a terminal vinyl group with a side-chain Si-H group. The depiction shows a chain-end/side-chain reaction, which leads to network formation.

offer unique processing benefits.^[12] Uncured, silica-filled, catalyzed pre-elastomer stock, known in industry as high-consistency elastomer, can be stored as a stable mixture for months prior to use. The combination of filler loading, filler treatment (hydrophobization), and extremely high-molecular-weight polymer yields high green strength that, coupled with rapid cure rates, allows complex shapes and profiles formed from the material to retain their physical dimensions throughout the molding and curing processes. This dimensional stability during cure is especially beneficial for continuous extrusion and curing processes employed in the manufacture of high-volume products such as silicone tubing.

Condensation Cure

Condensation cure means that a volatile product (condensate) is formed. Water is always a product in the presence of moisture, whereas the organic product depends on the condensation precursor. Scheme 5 shows acetic acid as the cocondensation product stemming from the reaction of an alkyl triacetoxysilane, $\text{RSi}[(\text{OC}(\text{O})\text{CH}_3)_3]$

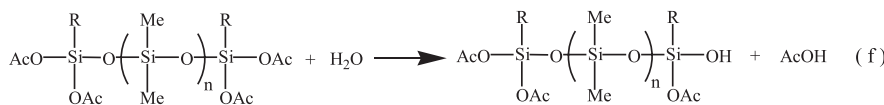
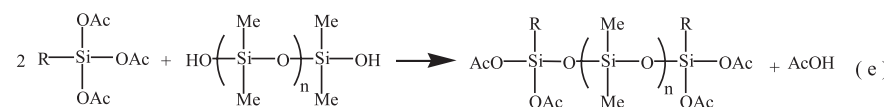
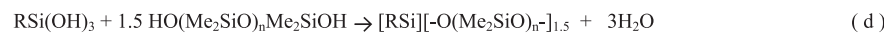
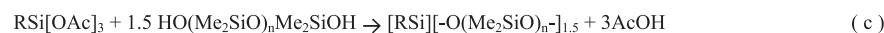
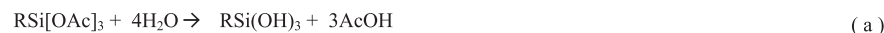


Scheme 4 General mechanism of the peroxide-catalyzed silicone elastomer cure.

or $\text{RSi}[\text{OAc}]_3$, with hydroxy-terminated PDMS. Ethanol is another common condensation product when silicon ethoxides are employed.

In contrast to the stoichiometric hydrosilation cure (one Si-H reacts with one Si-CH=CH₂), condensation cure is nonstoichiometric. Thus, it is important to note that the examples in Scheme 5 are ideal, or illustrative, because the composition of the hybrid product is determined by the experimental ratio of $\text{RSi}[(\text{OC}(\text{O})\text{CH}_3)_3]$ to Si-OH groups in the macromonomer. The term hybrid product indicates that the elastomer consists of a PDMS network and a siliceous domain (SD). In Scheme 5, the alkyl triacetoxysilane is both crosslinker and SD precursor. That is, $\text{RSi}[(\text{OC}(\text{O})\text{CH}_3)_3]$ is trifunctional and crosslinks the PDMS network by reaction with Si-OH, as well as undergoing self-condensation to form an SD.

The amount of crosslinker/siliceous domain precursor (C-SD-P) used is arbitrary but limited to certain bounds, as will be discussed further. C-SD-P in excess of that required by the macromonomer forms increasing volume fractions of SD. Care must be used in selection of the C-SD-P precursor. For example, tetraethoxysilane (TEOS) is volatile under these non-solvent reaction conditions. Under open conditions, product stoichiometry reflects up to 90% TEOS loss by volatilization.^[30] The absence of quantitative evaluation of composition confounds the significance of morphological studies on TEOS-cured materials.^[31] Condensation cure of PDMS at ambient temperature by the relatively involatile C-SD-P poly(diethoxysiloxane) produces materials of reproducible and predictable composition whose stability to mass loss in



A

B

Scheme 5 Examples of typical condensation-cure reactions.

water is highly dependent on the C–SD–P concentration.^[32] Information on stability to water is not available for other C–SD–Ps, but is an important issue for practical biomedical applications.

In Scheme 5, the composition of the siliceous domain (SD) is shown as $\text{RSiO}_{1.5}$. Equations for condensation cure are often represented in a way that indicates complete condensation, for example TEOS to SiO_2 .^[4] However, with the polydiethoxysiloxane C–SD–P precursor and ambient temperature cure conditions,²⁹Si solid-state NMR spectroscopy showed that the SD composition is approximately $\text{SiO}_{1.5}(\text{OH})$.^[32] Although the organic condensate ethanol is completely evolved within 24 hours, the crosslinked network formed by condensation cure at ambient temperature creates a kinetic barrier to the formation of water. The presence of the SD adds an element of compositional complexity to condensation cure that is not present in hydrosilation cure.

Condensation-cure systems for biomedical applications often use trifunctional crosslinkers, such as the alkyltriacetoxysilanes ($\text{RSi}(\text{OC}(\text{O})\text{CH}_3)_3$) illustrated in Scheme 5. Methyltri(acetoxy)silane (MTA) has a molecular weight of 220 g/mol, a melting point of 40°C, and boils at 95°C at 9 mm pressure. Given its relative low volatility, loss of MTA by volatilization during cure certainly does not occur to the extent observed for TEOS.^[30] Whether some small amount of volatilization occurs is not known. One notable example of the use of alkyltriacetoxysilane cure in medical devices is the silicone elastomer envelopes used in some saline-filled breast implants.^[33] Dramatic improvements have been reported in abrasion resistance and bi durability of these acetoxy condensation-cured shells, compared to those produced by hydrosilation cure. The improvement is specific to the saline implant envelopes and stems from envelope-to-envelope self-abrasion (such as at folds).

This wear mechanism is not operative for envelopes in silicone gel-filled implants, because the silicone gel acts as a lubricant.^[6,34,35]

In the presence of excess alkyltriacetoxysilane (ATAS), Si–OH macromonomer endgroups react to form alkyl(diacetoxy) siloxy-terminated chains (Scheme 5e). Upon exposure to moisture, Si-acetoxy groups in structures such as **A** are replaced by Si–OH (Scheme 5f). In the course of crosslinking and SD formation, Si–OH either condenses with another Si–OH group, **B**, or reacts with another Si-acetoxy (ATAS, **A**). A catalyst may be used to accelerate cure.

Each acetoxy group is liberated as acetic acid, which volatilizes under the ambient cure conditions driving condensation. This system is sufficiently reactive to cure sections with a thickness of approximately 1 cm or less. Above this thickness, the inability of moisture and acetic acid to diffuse in and out of the material on a practical time scale limits cure effectiveness. Excellent thick-section cure is a significant advantage of hydrosilation cure over acetoxy cure.

The ratio of ATAS to Si–OH is determined empirically for specific end-use applications, with the aim of optimizing physical properties and cure rate. Too low a ratio results in a crosslinked gel even in a sealed storage container, but too high a ratio results in slow cure rates under normal “moisture cure” conditions. Experience has led to a ratio of at least 10:1, as discussed further in the following text. Hydrolysis and condensation of excess ATAS lead to crosslinking and the formation of a siliceous phase. In contrast to the TEOS and poly(diethoxysiloxane) cure, the siliceous phase from ATAS contains T silicons.

Mallam, et al., have elegantly described the nature of the siliceous component for a hydroxy-terminated PDMS (M_n , 40K)–ethyltriacetoxysilane (4 wt%)

system.^[36] The mole ratio of acetoxy groups to Si–OH for this composition was approximately 10:1. Using light and neutron scattering, condensed ETAS aggregates with a mean radius of approximately 80 nm were found. Only about one-third of the ETAS was accounted for by these aggregates, with the remainder composed of oligomers and monomers. One can readily imagine the presence of a dispersed ethylsiliceous phase, composed of all manner of $\text{RSi}(\text{O}-\text{X})_3$ species, with $\text{X} = \text{H}, -\text{O}-\text{Si}(\text{C}_2\text{H}_5)(\text{O}-)_2, -\text{O}-[\text{Si}(\text{CH}_3)_2\text{O}]_n, \text{Si}(\text{CH}_3)_2\text{O}-$. In industrial formulations, the ratio of acetoxy groups to Si–OH may be considerably higher than 10:1. However, with PDMS macromonomers having extremely high MW, the weight fraction of the siliceous phase in the final material is low (ca. 3–5%).

Fillers

Hydrosilation-cured or condensation-cured (3–4 wt% siliceous phase) polysiloxane networks are inherently weak materials. To achieve the desired properties for most applications, silicone elastomers must be reinforced with fillers.^[37] The customary filler for biomedical silicone elastomers is noncrystalline (amorphous) fumed silica,^[6] but, reinforcing silica fillers may also be formed in situ.^[38,39] Fumed silica (SiO_2) consists of nanoscale-sized particles and particle clusters,^[40] the surfaces of which are usually treated with a hydrophobizing agent to moderate reinforcing properties. Commonly used silica treating agents include hexamethyldisilazane (HMDZ), octamethylcyclotetrasiloxane (D_4), dimethyldichlorosilane (DDS), and various polydimethylsiloxane oligomers, either with or without functional groups such as Si–OH. During these treatments, surface Si–OH groups are converted to siloxane groups, rendering the surface nominally hydrophobic. There is a large amount of information about these treated silicas available in refereed publications,^[37,40–44] the patent literature^[45–47] and manufacturers' product literature.^[48–58]

Treatment of fumed silica with HMDZ leads to a hydrophobic surface on which the majority of the surface Si–OH groups have been replaced by Si–O–SiMe₃ groups.^[41] This process can be conducted either in a separate pretreatment step prior to the compounding of the silica with siloxane polymer or in situ during the compounding process. Pretreatment is

employed when functional groups in the siloxane polymer may be reactive toward the HMDZ, as in the case of –Si–OH–terminated PDMS.^[12]

Treatment of fumed silica with difunctional agents also produces a hydrophobic surface, but in these cases the majority of the surface Si–OH groups are replaced by dimethylsiloxane oligomers. Again, the silica treatment can be carried out either in a pretreatment step prior to mixing with polymer or during the compounding step. Treating agents include dimethyldichlorosilane,^[53,54] dimethylsiloxane cyclics^[49–51] (e.g., D_4 or mixed cyclics), and PDMS.^[55–58] In situ treatment processes generally employ low-molecular-weight functionalized linear siloxane oligomers such as $(\text{RO})\text{Me}_2\text{Si}-(\text{Me}_2\text{SiO}-)_x-\text{SiMe}_2(\text{OR})$, where $\text{R} = \text{H}, \text{CH}_3$, etc., and $x =$ approximately 3–10.^[12]

An interesting alternative to fumed silica is the recent employment of nanoclays to provide nanocomposite hybrids. An example is the work of Giannelis, who reported that the addition of only a few percent nanoclays dramatically improves mechanical properties. A few percent nanoclay filler gives mechanical properties equivalent to the usual 25–30% addition of fumed silica.^[59] This is a promising result that may find medical materials applications.

POLYURETHANES AND UREAS

Copolymerization utilizing siloxane monomers and macromonomers (telechelics) and other modifications has led to a wide variety of polymer structures for biomedical applications that incorporate a silicone segment.^[60,61] Silicone polyurethanes have been increasingly investigated because of their durability, improved chemical stability, and good biocompatibility.^[62,63] Silicone polyurethanes (PUs) are readily prepared from precursors to hard and soft segments^[64] and are often referred to as segmented copolymers. Siloxane polyurethane ureas (PUUr) are prepared similarly, for example by using an α,ω -aminopropyl-dimethylsilyl substituted macromonomer (Fig. 3). Silicone polyurea ureas (PUrUrs) utilize an amine chain extender (instead of a diol), as well as the alkylamine-terminated silicone macromonomer. With the incorporation of urea functionality, better phase separation is observed due to

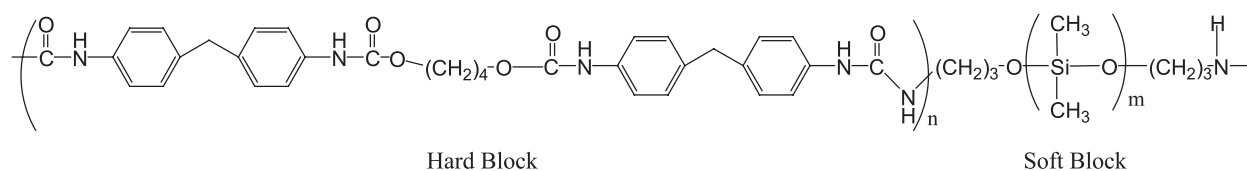


Fig. 3 General structure for a polyurethane urea derived from α,ω -bis-(3-aminopropyl) poly(dimethylsiloxane) (soft block) and methylene bis-p-phenyl isocyanate (MDI) and butane diol (BD).



Fig. 4 PDMS-PU heart valve on a poly(etheretherketone) frame, sewing ring attached, explanted after six months, in the mitral position. (From Ref. [68].)

enhanced hard-block hydrogen bonding.^[65,66] Therefore, for a given hard-block concentration, mechanical properties are better in the order $UrUr > UUr > UU$.

Because of their toughness and biostability, silicone polyurethane ureas have been used to fabricate heart valve leaflets.^[67] A mechanical and morphological study of biostable polyurethane urea heart valve leaflets explanted from sheep was reported by Bernacca.^[68] The valve was explanted after six months in the mitral position and there was “no sign of significant lipid uptake over this time.” SEM images on an unimplanted control were identical to the explanted leaflet. Figure 4 shows a picture of the explanted leaflet valve. These results are important, because some studies of silicone-polyurethane ureas, with regard to platelet deposition, and silicone-polyurethanes, with respect to lipophilicity, were not encouraging.^[69,70]

Instead of using the pure PDMS PU, various PDMS-containing segmented copolymers may be added in small amounts to conventional polymers as “surface modifying additives.”^[71,72] The SMA approach may be useful in large-volume applications such as tubing, where cost is an issue.

SURFACE SCIENCE

The surface science of poly(dimethylsiloxane)-containing materials is fundamental to the understanding of interfacial interactions relevant to biocompatibility. The poly(dimethylsiloxane) moiety, even when present in small amounts as an additive,^[72,73] dominates the surface and interfacial science. New experimental methods such as sum frequency generation (SFG) infrared spectroscopy^[71,74] provide in situ measurements that

promise better information on the chemical interface between polymers and biological systems. Better understanding of the chemical and morphological nature of this interface will lead to rational chemical and processing approaches to improve desired properties.

Contact angles have been used extensively to characterize PDMS implants or implant analogues.^[75–77] Water contact angles were used by Kennan et al. to estimate the degradation of medical-grade silicones under oxidizing conditions.^[78,79]

Recently, a method for measuring intrinsic advancing (θ_{adv}) and receding (θ_{rec}) water contact angles was reported by Uilk for hydrosilation-cured PDMS coatings that are analogues of biomedical materials.^[17] Static and dynamic contact angle (DCA) methods were used to evaluate coatings prepared from commercial divinyl-terminated PDMS (C-DVPDMS) and synthesized low-polydispersity divinyl-terminated PDMS (N-DVPDMS). DCA measurements showed that coatings prepared with both C-DVPDMS and N-DVPDMS contaminated the water surface during analysis. For ambient temperature cure, the rate of contamination was C-DVPDMS > N-DVPDMS. Methods for acquiring intrinsic contact angles on the PDMS coatings include using a large-surface-area reservoir or changing the water reservoir between DCA cycles. Intrinsic contact angles for hydrosilation-cured PDMS coatings are θ_{adv} , 118° and θ_{rec} , 83° . After contamination of the water surface, the contact angles change to θ_{adv} , 108° and θ_{rec} , 87° .

Changes in wetting behavior due to water contamination stand in contrast to changes in wetting behavior that occur due to surface structure reorganization. Surface reorganization in water is usually characterized by decreasing θ_{adv} and θ_{rec} with immersion time. Such changes are observed for polyurethanes, acrylates, and other polymers that can hydrogen-bond with water.^[80–82] Favorable enthalpy drives reorganization, with the rate depending on surface mobility. However, if the coating does not contaminate water and if surface mobility is sufficient, intrinsic θ_{adv} and θ_{rec} are restored, usually by drying in air. Until contamination effects occur, C-PDMS coatings display stable intrinsic values for θ_{adv} and θ_{rec} . Contact angle hysteresis ($35^\circ \pm 4^\circ$) is therefore also stable, because PDMS has a very low T_g , making the timescale of surface reorganization many orders of magnitude faster than the switch from advancing to receding force-distance determinations.

Although these were not especially purified macromonomers as discussed in the following section on purity requirements, the water/saline contamination effect observed for PDMS coatings has implications for biomaterials evaluation. The importance of the

protocol for detecting contamination rests in the likelihood that if a polymer coating contaminates water or saline in a DCA experiment, it is likely to contaminate a host in an implant application.

The dibutyltin diacetate-catalyzed reaction of (tridecafluoro-1,1,2,2-tetrahydrooctyl) triethoxysilane (FTEOS) with silanol-terminated poly(dimethylsiloxane), $\text{HO}(\text{Me}_2\text{SiO})_n\text{H}$, has been used to prepare surface-modified siloxane networks (ambient temperature cure). Surface phase separation occurs as a function of FTEOS concentration. FTEOS-6x compositions (where 6x is the initial SiOEt/SiOH ratio) exhibited micron-scale islands of fluorinated siliceous phase (FSP) surrounded by a siloxane-rich sea (Fig. 5), whereas the surface of FTEOS-12x was featureless due to the formation of a continuous surface FSP. Dynamic contact angle (DCA) analysis on FTEOS-12x confirmed the presence of an FSP monodomain characterized by high θ_{adv} ($135\text{--}136^\circ$) and low θ_{rec} ($55\text{--}56^\circ$), similar to previously reported values for fluoroalkyl side chain polymers.^[83] Long-term testing showed that the coatings undergo slow chemical degradation by water, with the rate depending on composition. FTEOS-12x is chemically stable in water due to the continuous FSP overlayer.

As noted earlier, poly(diethoxysiloxane) is a useful, nonvolatile condensation-cure SD precursor.^[32] With Si-OEt from poly(diethoxysiloxane) and Si-OH from $\text{HO}(\text{Me}_2\text{SiO})_n\text{H}$, compositions with SiOEt/SiOH ratios from 5 to 35 (ES40-5x-ES40-35x) were prepared. Both light microscopy and TM-AFM show that PDMS-SD compositions 5x and higher slowly develop islandlike surface features while stored at ambient conditions (Fig. 6). After a one month cure, PDMS-SD surface features are $\leq 1\ \mu\text{m}$, whereas after 6 months feature size is broadly distributed up to approximately $10\ \mu\text{m}$. In contrast, surface features developed quickly (within 24 hr) with the PDMS-FSD hybrids noted earlier. TM-AFM demonstrated that PDMS-SD surface features are often poorly adherent and easily moved compared to the robust phase-separated surface structure of PDMS-FSD materials. The stability in water of

PDMS-SD materials is strongly compositionally dependent. Over a period of 70 days, the ES40-14x composition showed the greatest resistance to mass loss, with a mass loss rate of 0.08 wt% per month.

In summary, fundamental surface science studies show that addition^[17] and condensation cured^[32] PDMS surfaces can be “alive” in the sense of diffusion of water-contaminating species. Condensation-cured PDMS shows the development of surface features due to the reactive phase-separation processes and a compositionally dependent mass loss in water.^[32] Further research is clearly warranted on specially purified resins and filled compositions that closely parallel those used in biomedical applications, so as to better understand the efficacy of purification protocols and effects on biocompatibility.

PURITY REQUIREMENTS

Medical devices intended for use in the human body are regulated by government agencies around the world. The Food and Drug Administration (FDA) is the regulatory body for devices marketed in the United States. The FDA provides guidance information for manufacturers of these medical devices.^[84] The FDA also lists selected medical device agencies outside the United States and their Websites that contain information on the requirements for marketing medical devices in those countries.^[85] Specific requirements may vary somewhat from country to country, but in general they are based upon performance of the device in the application for which it is designed. One important factor associated with the performance of an implantable device and, more specifically, its biocompatibility and toxicology, is the purity of the materials from which the device is manufactured. Silicone is no exception.

Because of its long history in the marketplace, silicone has been one of the most intensively studied biomaterials ever developed. Various characteristics of the specific silicone materials employed in devices for

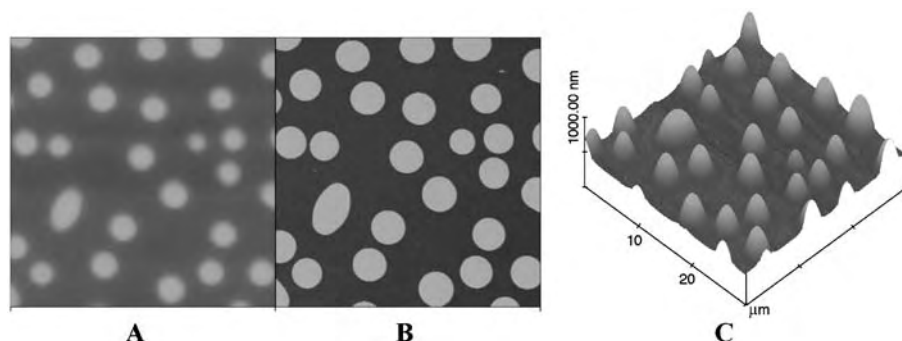


Fig. 5 Tapping mode atomic force microscopy images of FTEOS 6x. **A**, 2D height; **B**, phase contrast, and **C**, 3D height. Scans are $30 \times 30\ \mu\text{m}$.

S

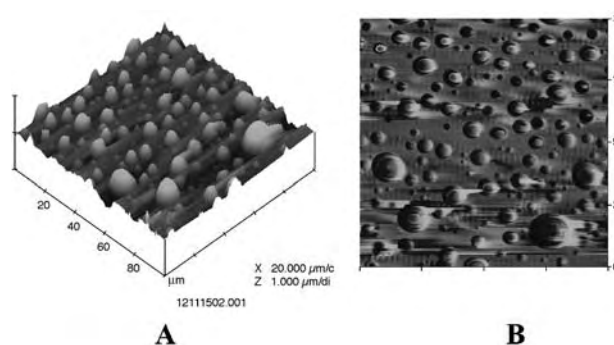


Fig. 6 TM-AFM of poly(diethoxysiloxane)-cured hydroxy-terminated PDMS (27K) with Si–OEt/Si–OH = 10. Scans are $100 \times 100 \mu\text{m}$. **A**, 3D height; **B**, phase contrast.

human use must be assessed by a device manufacturer as part of the process of applying for approval to market a new product. Some of the characteristics that are assessed for certain devices may be specified in guidance documents, such as the FDA guidance on breast implants.^[86] Materials composing finished sterilized devices are thoroughly characterized by the device manufacturer, and the data from these characterization activities are provided to the regulatory agencies as part of the approval process for each new device. In general, attributes that are often assessed fall into the following categories: trace/low-level nonsilicone components present by design (e.g., catalysts); trace/low-level contaminants, impurities, or by-products (e.g., heavy metals); low-molecular-weight and/or volatile species; and extractable materials. Because these attributes are controlled by the silicone manufacturer, the device manufacturer develops specifications for the material. This ensures that the material attributes are suitable for the ultimate use intended for the device.

Silicone material manufacturers are not subject to regulation by the agencies that govern device manufacturing. However, silicone manufacturers often voluntarily develop quality systems that are audited by quality standards organizations such as the International Organization for Standardization (ISO).^[87] An ISO9000 certificate of conformity provides a high level of assurance that a silicone manufacturer will deliver consistently high-quality materials to its customers. Strategies employed by silicone manufacturers for controlling material purity are guided by the principles of their quality systems. These strategies are implemented in at least four different areas of the overall manufacturing process: 1) raw material purity control; 2) control of the silicone material processing environment; 3) in-process purification steps; and 4) selection of packaging materials. Each of these is discussed in greater detail in the following sections.

Raw Material Purity Control

The list of raw materials used by silicone manufacturers includes silane or silicone monomers, fillers, filler treating agents, solvents, catalysts, and cross-linkers. A control scheme for raw materials may consist of several key elements, as shown in Fig. 7.

Carefully documented specifications including the purity requirements for each of these materials are the first control point in the overall purity control scheme. For example, solvents used to disperse silicone intended for dip-coating operations must be of high purity, because the coating process typically involves removal of the solvent after film formation. Any nonvolatile contaminants, such as heavy metals in the solvent would therefore likely concentrate in and remain with the silicone in the device. Thorough analysis of several lots of a raw material from each prospective vendor, along with audits of their facilities and quality systems, is typically carried out prior to approving those vendors as raw material suppliers. Testing at this stage would likely include analyses not routinely performed on all incoming raw material lots, but important to overall characterization and assurance of final product purity nonetheless. After vendor approval, inspection of each lot of purchased raw material may include routine purity and identity checks prior to acceptance for use. In addition to incoming inspection data generated by the silicone manufacturer, a vendor's certified material data may also be accepted as evidence of acceptability after initial vendor approval is complete. Periodic checks of vendor-certified data may be conducted to provide ongoing assurance of vendor compliance.

Control of the Silicone Processing Environment

Several elements compose an effective control strategy for preventing contamination of silicone materials

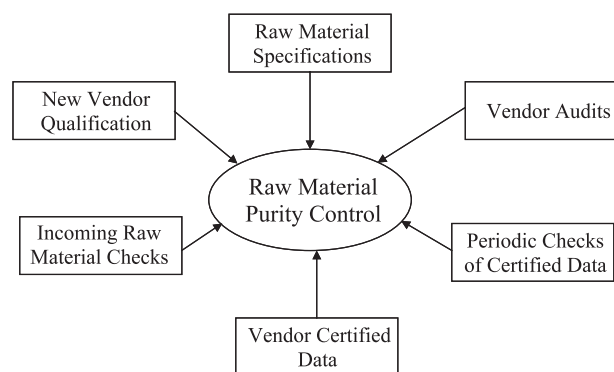


Fig. 7 Important elements in a raw material control scheme.

while they are being processed. These include atmospheric controls, equipment controls, and process controls. With respect to atmospheric controls, the nature of certain silicone materials and the manufacturing processes in which they are employed may require that they be processed within an environment that has particulate controls in place. For example, HEPA filtration, so-called white rooms, or clean rooms may be utilized to prevent particulate contamination of the silicones from ambient air. Equipment design, maintenance, and cleanliness are important factors associated with avoiding contamination by foreign materials or incompatible silicone materials that were processed in the same equipment. Mixers are used in many different places in silicone processing, and a poorly designed or maintained agitator seal, for example, might allow foreign materials to contaminate the silicone. Equipment cleaning procedures are carefully designed and conducted to prevent cross-contamination with other materials that may be processed in the same equipment. Process controls are important in preventing conditions that might generate unwanted byproducts or degradation products. For example, high-temperature processing requires the use of an inert atmosphere to prevent undesired oxidation of the silicone. Oxidative degradation is a well-known phenomenon^[88] that can be prevented by controlling process conditions carefully.

In-Process Purification Steps

After raw material and environment controls, in-process purification steps are important for ensuring silicone material acceptability. Two of the most commonly employed in-process purification methods are filtration and distillation/stripping.

Filtration

Two similar types of filtration are useful: simple physical filtration and adsorptive filtration. In either case, the choice of filtration material is important to ensure compatibility with silicone materials and the absence of leachables that might compromise purity. Physical filters remove particulates above a specified size that may be present in the various silicone materials. Examples include cartridge filters for lower-viscosity materials, such as solvent dispersions or silicone gel precursors, and screen packs placed at the outlets of extruders in which high-viscosity materials, such as high-consistency elastomer stocks, are processed. These filters are commonly used to remove incidental particulate contamination or agglomerates of fumed silica that might remain after mixing the silica into silicone polymer. Adsorptive

filtration may be used to remove color- or odor-causing species from the silicone materials. Activated carbon is widely used in the chemical manufacturing industry^[89] and is commonly used as an adsorbent in silicone manufacturing. Species removed include, for example, polymerization catalyst residues. Tetramethylammonium hydroxide (TMAH)^[90] is a commonly used polymerization catalyst in silicone manufacturing because of its high activity and its transient nature. After polymerizations using TMAH are complete, the process temperature is increased to promote internal rearrangement of the TMAH to trimethylamine and methanol, which are subsequently stripped from the polymer under vacuum (vide infra). Trimethylamine has a typical, strong amine odor that can be detected at extremely low levels. A carbon-impregnated cartridge filter may be used to remove the last traces of amine odor from the silicone polymers.

Distillation/stripping

As described earlier, one of the consequences of industrially practiced equilibrium-controlled silicone polymerization is the presence of a substantial fraction (> 10% by weight) of silicone cyclics at equilibrium and—to a degree that depends on the average molecular weight of the polymer produced—low-molecular-weight linear siloxane oligomers. The most abundant cyclics are often dimethylsiloxane cyclics (D_x ; $x=4, 5, 6\dots$). Lacking functionality, they are potential extractables if not removed from biomedical elastomers. A commonly employed method for removal of these cyclics and low-molecular-weight oligomers is distillation or stripping. These steps are often carried out at elevated temperature and/or low pressure. For the most demanding applications, the levels of D_x can be reduced to < 100 ppm by weight relative to the polymer by thin-film or wiped-film evaporation at high temperatures (> 150°C) and low pressures (< 50 millitorr).^[12]

Packaging Materials

Proper selection and control of packaging materials for silicones is important to ensure purity. Packaging at the silicone materials stages can include small vials and jars holding a few grams, plastic sheeting for high-consistency elastomers, and bulk containers holding as much as 1000 liters or more. The possible presence of particulate matter in the packaging must be addressed to prevent contamination. Careful consideration of packaging materials is also necessary to ensure that foreign substances do not leach into the silicone. Steel drums (55-gallon capacity) are often used to package dispersions of silicones in solvents, and the drum liners must be checked for degradation by aggressive

solvents. In addition, certain silicones, particularly acetoxy-endblocked silicones, contain fairly aggressive components that react with or corrode unprotected steel. A suitable drum lining material, such as a phenolic epoxy, is required for acetoxy-cured silicone in xylene to ensure corrosion and leaching resistance.

APPLICATIONS, BIOCOMPATIBILITY, AND BIODURABILITY

Table 2 outlines some of the hundreds of biomedical applications for silicones. Most of these applications have been adequately described in readily available sources.^[5,91,92]

Determining the parameters that are important in biocompatibility and biodurability remains an ongoing

challenge. With regard to long-term biocompatibility, no evidence has been found suggesting an association between silicone breast implants and connective tissue diseases.^[6]

Brandon et al. have conducted systematic studies of shells from explanted silicone gel breast implants.^[93] Their study addressed the failure mechanisms associated with silicone gel breast implants. The results indicated that silicone gel implants have remained intact for as long as 32 years in vivo and that “degradation of the shell mechanical and chemical properties is not a primary mechanism for silicone gel breast implant failure.”

The foregoing studies attest to the intrinsically safe and robust nature of silicone biomaterials.^[6,68,93] This information must be tempered with the realization that there are still important gaps in our knowledge

Table 2 Selected examples of applications of biomedical silicones according to material type and molecular architecture

Material type	Molecular architecture	Applications examples
Fluids	Uncross-linked, unfilled linear polysiloxanes	<ul style="list-style-type: none"> • Implants: retinal tamponade, gel swelling agents • Lubricants: silicone valves, syringes, needles, condoms • Antifoams: gastric “antigas” agents
Responsive gels	Lightly cross-linked, unfilled macromonomers	<ul style="list-style-type: none"> • Soft tissue implant filler: breast implants, testicular implants
Cohesive gels & films	Moderately cross-linked, unfilled macromonomers	<ul style="list-style-type: none"> • Wound care: gel sheeting
Low-durometer (soft) elastomers	Cross-linked macromonomers with low filler loading	<ul style="list-style-type: none"> • Membranes: blood oxygenation, dialysis • Encapsulants: cochlear implants, pacemakers/defibrillators • Skin contact adhesives: transdermal drug delivery, electromedical devices, ostomy appliances, surgical drapes, tapes
High-strength elastomers	Cross-linked macromonomers with moderate to high-level filler loading	<ul style="list-style-type: none"> • Plastic surgery: soft tissue implant envelopes, tissue expanders, maxillofacial implants • Ophthalmology: punctum plugs, intraocular lenses • Otolaryngology: tympanostomy tubes, tracheostomy tubes • Urology: penile implants, catheters, incontinence devices • Neurology: hydrocephalic shunts • Orthopedics: joint/cartilage/bone replacements • Cardiology: artificial heart valves/coating, pacemaker lead coatings • Others: integral septa for implanted devices (drug delivery devices, tissue expanders, vascular infusion ports, etc.), catheters, balloons, shunts, infusion sleeves, wound drains, laparoscopic devices, splints

concerning the hemocompatibility and histocompatibility of silicone biomaterials, as well as interfacial phenomena such as the release and diffusion of siloxane-containing species. Basic knowledge concerning these phenomena, the promise of molecular-level characterization of the polymer-biological interface, and ways to improve the mechanical properties of silicone biomaterials without compromising biocompatibility will no doubt lead to more widespread applications, with attendant improvements in the quality of life.

ACKNOWLEDGMENTS

Kenneth J. Wynne thanks The National Science Foundation (DMR 0207560) for supporting of aspects of surface science reported herein. The authors thank Dr. Iskender Yilgor for reading the manuscript and for helpful suggestions.

ARTICLES OF FURTHER INTEREST

Burn Dressing; Contact Lenses: Silicone Hydrogels; Polymers

REFERENCES

- Levier, R.R.; Harrison, M.C.; Cook, R.R.; Lane, T.H. What is silicone. *Plast. Reconstr. Surg.* **1993**, *92*, 163–167.
- Barry, A.J.; Beck, H.N. *Inorganic Polymers*; Stone, F.G.A. Graham, W.A.G., Eds.; Academic Press: New York, 1962; 189–320.
- Mark, J.E.; Allcock, H.R.; West, R. *Inorganic Polymers*; Prentice-Hall, Inc.: Englewood Cliffs, NJ, 1992.
- Clarson, S.J.; Smith, S.D.; Dyke, M.E. *Synthesis and Properties of Silicones and Silicone-Modified Materials*; American Chemical Society: Washington, D.C., 2003.
- Arkles, B. Look what can make out of silicones. *Chemtech* **1983**, *13*, 542–555.
- Bondurant, S.; Ernster, V.; Herdman, R. *Safety of Silicone Breast Implants*; National Academy Press: Washington, DC, 1999.
- Flory, P.J. *Statistical Mechanics of Chain Molecules*; Wiley-Interscience: New York, 1969.
- Allcock, H.R.; Prange, R. Properties of poly(phosphazene-siloxane) block copolymers synthesized via telechelic polyphosphazenes and polysiloxane phosphoranimines. *Macromolecules* **2001**, *34*, 6858–6865.
- Morita, T.; Maughon, B.R.; Bielawski, C.W.; Grubbs, R.H. A ring-opening metathesis polymerization (ROMP) approach to carboxyl- and amino-terminated telechelic poly(butadiene)s. *Macromolecules* **2000**, *33*, 6621–6623.
- Mark, J.E. Elastomeric networks with bimodal chain-length distributions. *Accounts Chem. Res.* **1994**, *27*, 271–278.
- Bondurant, S.; Ernster, V.; Herdman, R. Silicone Chemistry. In *Safety of Silicone Breast Implants*; Bondurant, S., Ernster, V., Herdman, R., Eds.; National Academy Press: Washington, DC, 1999; 39–53.
- Compton, R.A. Silicone manufacturing for long-term implants. *J. Long-Term Eff. Med. Implants* **1997**, *7*, 29–54.
- Clarson, S.J. *Cyclic Polymers*, 2nd Ed.; Semlyen, J.A., Ed.; Kluwer, Dordrecht: The Netherlands, 2000; 161–183.
- Shiho, H.; DeSimone, J.M. Preparation of silicone-graft copolymers by homogeneous radical copolymerization in supercritical carbon dioxide. *J. Polym. Sci., Part A, Polym. Chem.* **2000**, *38*, 1139–1145.
- Yilgör, I.; McGrath, J.E.; Krukonis, V.J. Novel supercritical fluid techniques for polymer fractionation and purification. II. fractionation and characterization of functional siloxane oligomers. *Polym. Bull.* **1984**, *12*, 499–506.
- Lee, C.L.; Johansson, O.K. Polymerization of hexamethylcyclotrisiloxane with a biscatecholsiliconate. II. Kinetics of polymerization. *J. Polym. Sci., Polym. Chem.* **1976**, *14*, 729.
- Uilk, J.M.; Mera, A.E.; Fox, R.B.; Wynne, K.J. Hydrosilation-cured poly(dimethylsiloxane) networks: Intrinsic contact angles via dynamic contact angle analysis. *Macromolecules* **2003**, *36*, 3689–3694.
- Razzano, J.S.; Gosh, N.E. Process for Producing Fluorosilicon Polymers. U.S. 6,492,479, 2002.
- Brown, E.D.; Carmichael, J.B. Distribution in 3,3,3-trifluoropropylmethylsiloxane polymers. *J. Polym. Sci., Polym. Lett.* **1965**, *3*, 473–482.
- Boutevin, B.; Pietrasanta, Y. *Progress in Organic Coatings*; Elsevier: New York, 1985; Vol. 13, 297–331.
- Allcock, H.R.; Lampe, F.W.; Mark, J.E. *Contemporary Polymer Chemistry*, 3rd Ed.; Pearson-Prentice Hall: Upper Saddle River, NJ, USA, 2003.
- Chen, X.; Gardella, J.A.; Ho, T.; Wynne, K.J. Surface-composition of a series of dimethylsiloxane urea urethane segmented copolymers studied by electron-spectroscopy for chemical-analysis. *Macromolecules* **1995**, *28*, 1635–1642.
- Mera, A.E.; Fox, R.; Bullock, S.; Wynne, K.J. Surface properties and fouling release behavior of poly(dimethylsiloxane) networks. *Proc. 21st Ann. Mtg. Adh. Soc.* **1998**, 138–140.
- Bullock, S.; Johnston, E.E.; Willson, T.; Gatenholm, P.; Wynne, K.J. Surface science of a filled polydimethylsiloxane-based alkoxy silane-cured elastomer: RTV11. *J. Colloid Interfacial Sci.* **1999**, *210*, 18–36.
- Stein, J.; Lewis, L.N.; Gao, Y.; Scott, R.A. In situ determination of the active catalyst in hydrosilylation reactions using highly reactive Pt(0) catalyst precursors. *J. Am. Chem. Soc.* **1999**, *121*, 3693–3703.
- Arepalli, S.R.; Bezabeh, S. Allergic reaction to platinum in silicone breast implants. *J. Long-Term Eff. Med. Implants* **2002**, *12*, 299–306.
- Marko, I.E.; Sterin, S.; Buisine, O.; Mignani, R.; Branlard, P.; Tinant, B.; Declercq, J.P. Selective and efficient platinum(0)-carbene complexes as hydrosilylation catalysts. *Science* **2002**, *298*, 204–206.

28. Thomas, D.R. *Siloxane Polymers*; Clarkson, S.J., Semlyen, J.A., Eds.; Prentice Hall: Englewood, NJ, 1993; 567–615.
29. Rich, J.; Cella, J.; Lewis, L.; Stein, J.; Singh, N.; Rubinsztajn, S.; Wengrovius, J. *Kirk-Othmer Encyclopedia of Chemical Technology*, 5th Ed.; Kroschwitz, J.I., Howe-Grant, M., Eds.; John Wiley & Sons, Inc.: New York, 1997; Vol. 22, 82–142.
30. Wynne, K.J.; Ho, T.; Johnston, E.E.; Myers, S.A. Surface science and stability of networks prepared from hydroxy-terminated polydimethylsiloxane and methyltriethoxysilane. *Appl. Organomet. Chem.* **1998**, *12*, 763–770.
31. Viers, B.D.; Sukumaran, S.; Beaucage, G.; Mark, J.E. On the observation of novel morphologies in "model" PDMS elastomers—Evidence of trapped spinodal structures. *Polym. Preprints* **1997**, *38*, 333–334.
32. Ulk, J.; Bullock, S.; Johnston, E.; Myers, S.A.; Merwin, L.H.; Wynne, K.J. Surface science of elastomeric coatings prepared from alpha, omega-dihydroxypoly (dimethylsiloxane) and the ethoxysiloxane mixture "ES40". *Macromolecules* **2000**, *33*, 8791–8801.
33. Bondurant, S.; Ernster, V.; Herdman, R. Implant Catalogue. In *Safety of Silicone Breast Implants*; Bondurant, S., Ernster, V., Herdman, R., Eds.; National Academy Press: Washington, DC, 1999; 54–79.
34. Gutowski, K.A.; Mesna, G.T.; Cunningham, B.L. Saline-filled breast implants: A plastic surgery educational foundation multicenter outcomes study. *Plast. Reconstr. Surg.* **1997**, *100*, 1019–1027.
35. Mladick, R.A. No-touch submuscular saline breast augmentation technique. *Aesthet. Plast. Surg.* **1993**, *17*, 183–192.
36. Mallam, S.; Hecht, A.-M.; Geissler, E.; Pruvost, P. Structure of swollen polydimethylsiloxane gels. *J. Chem. Phys.* **1989**, *91*, 6447–6454.
37. Cochrane, H.; Lin, C.S. The influence of fumed silica properties on the processing, curing, and reinforcement properties of silicone-rubber. *Rubber Chem. Technol.* **1993**, *66*, 48–60.
38. Kwan, K.S.; Harrington, D.A.; Moore, P.A.; Hahn, J.R.; Degroot, J.V.; Burns, G.T. Synthesis and use of colloidal silica for reinforcement in silicone elastomers. *Rubber Chem. Technol.* **2001**, *74*, 630–644.
39. Rajan, G.S.; Sur, G.S.; Mark, J.E.; Schaefer, D.W.; Beaucage, G. Preparation and characterization of some unusually transparent poly(dimethylsiloxane) nanocomposites. *J. Polym. Sci., Part B, Polym. Phys.* **2003**, *41*, 1897–1901.
40. Barthel, H.; Rosch, L.; Weis, J. Fumed Silica—Production, Properties, and Applications. In *Organosilicon Chemistry II: From Molecules to Materials*; Auner, N., Weis, J., Eds.; VCH Publishers: Weinheim, 1996; 761–778.
41. Mathias, J.; Wannemacher, G. Basic characteristics and applications of aerosil 30. The chemistry and physics of the aerosil surface. *J. Colloid Interface Sci.* **1988**, *125*, 61–68.
42. Hertl, W.; Hair, M.L. Reaction of hexamethyldisilazane with silica. *J. Phys. Chem.* **1971**, *75*, 2181–2185.
43. Gun'ko, V.M.; Vedamuthu, M.S.; Henderson, G.L.; Blitz, J.P. Mechanism and kinetics of hexamethyldisilazane reaction with a fumed silica surface. *J. Colloid Interface Sci.* **2000**, *228*, 157–170.
44. Boonstra, B.B.; Cochrane, H.; Dannenberg, E.M. Reinforcement of silicone rubber by particulate silica. *Rubber Chem. Technol.* **1975**, *48*, 558–576.
45. Lucas, G. Process of Treating Silica with a Siloxane and Product Thereof. U.S. 2,938,009, 1960.
46. Kobayashi, H.; Ohnishi, M. Method for Modifying the Surface of Finely Divided Silica. US 4,849,022, 1989.
47. Brown, E.D.; Hyde, J.F. Process for Rendering Inorganic Powders Hydrophobic. U.S. 3,334,062, 1967.
48. <http://w1.cabot-corp.com/controller.jsp?N=23+1001+4294967082+4294967073&entry=product> (accessed August 2003).
49. http://www2.sivento.de/sivento/silica/html/e/solutions_products/datasheet/silica_datasheet.asp?id=44&definition=2 (accessed August 2003).
50. http://www2.sivento.de/sivento/silica/html/e/solutions_products/datasheet/silica_datasheet.asp?id=518&definition=2 (accessed August 2003).
51. http://www2.sivento.de/sivento/silica/html/e/solutions_products/datasheet/silica_datasheet.asp?id=538&definition=2 (accessed August 2003).
52. http://www2.sivento.de/sivento/silica/html/e/solutions_products/datasheet/silica_datasheet.asp?id=57&definition=2 (accessed August 2003).
53. http://www2.sivento.de/sivento/silica/html/e/solutions_products/datasheet/silica_datasheet.asp?id=46&definition=2 (accessed August 2003).
54. http://www2.sivento.de/sivento/silica/html/e/solutions_products/datasheet/silica_datasheet.asp?id=522&definition=2 (accessed August 2003).
55. http://www.wacker.com/internet/webcache/en_US/PTM/TM/Wacker_HDK/Hydrophob/WACKER_HDK_H15.pdf (accessed August 2003).
56. http://www.wacker.com/internet/webcache/en_US/PTM/TM/Wacker_HDK/Hydrophob/WACKER_HDK_H18.pdf (accessed August 2003).
57. http://www.wacker.com/internet/webcache/en_US/PTM/TM/Wacker_HDK/Hydrophob/WACKER_HDK_H20.pdf (accessed August 2003).
58. http://www.wacker.com/internet/webcache/en_US/PTM/TM/Wacker_HDK/Hydrophob/WACKER_HDK_H30.pdf (accessed August 2003).
59. Burnside, S.D.; Giannelis, E.P. Synthesis and properties of new poly(dimethylsiloxane) nanocomposites. *Chem. Mater.* **1995**, *7*, 1597–1600.
60. Hron, P. Hydrophilisation of silicone rubber for medical applications. *Polym. Int.* **2003**, *52*, 1531–1539.
61. Abbasi, F.; Mirzadeh, H.; Katbab, A.A. Modification of polysiloxane polymers for biomedical applications: A review. *Polym. Int.* **2001**, *50*, 1279–1287.
62. Gunatillake, P.A.; Martin, D.J.; Meijs, G.F.; McCarthy, S.J.; Adhikari, R. Designing biostable polyurethane elastomers for biomedical implants. *Aust. J. Chem.* **2003**, *56*, 545–557.
63. Lamba, N.M.K.; Woodhouse, K.A.; Cooper, S.L. *Polyurethanes in Biomedical Applications*; CRC Press: Boca Raton, FL, 1998; 15.
64. Yilgor, E.; McGrath, J.E. Polysiloxane containing copolymers: A survey of recent developments. *Adv. Polym. Sci.* **1988**, *86*, 1–86.

65. Garrett, J.T.; Runt, J.; Lin, J.S. Microphase separation of segmented poly(urethane urea) block copolymers. *Macromolecules* **2000**, *33*, 6353–6359.
66. Yilgör, E.; Yilgör, I. Hydrogen bonding: A critical parameter in designing silicone copolymers. *Polymer* **2001**, *42*, 7953–7959.
67. Pathiraja, A.; Gunatillake, S.J.; McCarthy, G.F.; Adhikari, M.; Adhikari, R. *Siloxane-containing polyurethane-urea compositions*, AU PCT/AU00/00345; 2000.
68. Bernacca, G.M.; Straub, I.; Wheatley, D.J. Mechanical and morphological study of biostable polyurethane heart valve leaflets explanted from sheep. *J. Biomed. Mater. Res.* **2002**, *61*, 138–145.
69. Hergenrother, R.W.; Yu, X.-H.; Cooper, S.L. Blood-contacting properties of polydimethylsiloxane polyurea-urethanes. *Biomaterials* **1994**, *15*, 635–640.
70. Takahara, A.; Hergenrother, R.W.; Coury, A.J.; Cooper, S.L. In vitro evaluation of the biostability of segmented polyurethanes. *Artif. Organs* **1990**, *14*, 87–90.
71. Chen, Z.; Ward, R.; Tian, Y.; Malizia, F.; Gracias, D.H.; Shen, Y.R.; Somorjai, G.A. Interaction of fibrinogen with surfaces of end-group-modified polyurethanes: A surface-specific sum-frequency-generation vibrational spectroscopy study. *J. Biomed. Mater. Res.* **2002**, *62*, 254–264.
72. Ho, T.; Wynne, K.J. Mechanism and kinetics of hexamethyldisilazane reaction with a fumed silica surface blends of a siloxane/urea/urethane-segmented copolymer with an IPDI-based polyetherurethane. *Polym. Adv. Technol.* **1994**, *6*, 25–31.
73. White, K.A.; Ward, R.S.; Gill, R.S.; Lim, F.; Coviello, S.K. *Surface Modification of Polymeric Biomaterials*; Ratner, B.D., Castner, D.G., Eds.; Plenum: NY, 1997; 27–33.
74. Amitay-Sadovsky, E.; Komvopoulos, K.; Ward, R.; Somorjai, G.A. Surface reordering of stretched polyurethane block copolymer films studied by sum frequency generation vibrational spectroscopy. *J. Phys. Chem., B* **2003**, *107*, 6377–6383.
75. Waters, M.; Jagger, R. Improved wettability of an experimental silicone rubber denture soft lining material. *J. Biomed. Mater. Res.* **1999**, *48*, 765–771.
76. Waters, M.G.J.; Jagger, R.G.; Polyzois, G.L. Wettability of silicone rubber maxillofacial prosthetic materials. *J. Prosthet. Dent.* **1999**, *81*, 439–443.
77. Everaert, E.P.; Vandermei, H.C.; Devries, J.; Busscher, H.J. Hydrophobic recovery of repeatedly plasma-treated silicone-rubber. 1. Storage in air. *J. Adhes. Sci. Technol.* **1995**, *9*, 1263–1278.
78. Kennan, J.J.; Peters, Y.A.; Swarthout, D.E.; Owen, M.J.; Namkanisorn, A.; Chaudhury, M.K. Effect of saline exposure on the surface and bulk properties of medical grade silicone elastomers. *J. Biomed. Mater. Res.* **1997**, *36*, 487–497.
79. Batich, C.; DePalma, D.; Marotta, J.; Latorre, G.; Hardt, N.S. Silicone degradation reactions. *Immunol. Silicones* **1996**, *210*, 13–23.
80. Tingey, K.G.; Andrade, J.D. Probing surface microheterogeneity of poly(ether urethanes) in an aqueous environment. *Langmuir* **1991**, *7*, 2471–2478.
81. Tretinnikov, O.N.; Ikada, Y. Dynamic wetting and contact-angle hysteresis of polymer surfaces studies with the modified Wilhelmy balance method. *Langmuir* **1994**, *10*, 1606–1614.
82. Makal, U.; Wynne, K.J. Unpublished Results.
83. Katano, Y.; Tomono, H.; Nakajima, T. Surface property of polymer films with fluoroalkyl side chains. *Macromolecules* **1994**, *27*, 2342–2344.
84. <http://www.fda.gov/opacom/morechoices/industry/guidedc.htm> (accessed August 2003).
85. <http://www.fda.gov/cdrh/international/nonusregulators.html> (accessed August 2003).
86. <http://www.fda.gov/cdrh/ode/guidance/1354.html> (accessed August 2003).
87. <http://www.iso.ch/iso/en/ISOOnline.frontpage>.
88. Atkins, D.C.; Murphy, C.M.; Saunders, C.E. Polymethylsiloxanes: Thermal and oxidation stabilities. *Ind. Eng. Chem.* **1947**, *39* (11), 1395–1401.
89. Baker, F.S.; Miller, C.E.; Repik, A.J.; Tolles, E.D. Activated Carbon. In *Kirk-Othmer Encyclopedia of Chemical Technology*, 4th Ed.; Kroschwitz, J.I., Howe-Grant, M., Eds.; John Wiley & Sons, Inc.: New York, 1997; Vol. 4, 1015–1037.
90. Gilbert, A.K.K.S.W. Transient catalysts for the polymerization of organosiloxanes. *J. Polym. Sci., Polym. Lett.* **1959**, *40*, 35–58.
91. McMillin, C.R. Elastomers for biomedical applications. *Abstr. Papers Am. Chem. Soc.* **1994**, *67*, 417–446.
92. Yoda, R. Elastomers for biomedical applications. *J. Biomater. Sci., Polym. Ed.* **1998**, *9*, 561–626.
93. Brandon, H.J.; Jerina, K.L.; Wolf, C.J.; Young, V.L. Biodurability of retrieved silicone gel breast implants. *Plast. Reconstr. Surg.* **2003**, *111*, 2295–2306.

Silks

Juming Yao
Tetsuo Asakura

Department of Biotechnology, Tokyo University of Agriculture and Technology, Koganei,
Tokyo, Japan

INTRODUCTION

Silks represent a unique and important class of structural proteins in nature, providing a wide range of native functions as protective membranes to withstand environmental insults during development, high strength netting to entrap insects, and lifelines to support spiders. Silk proteins are usually produced within specialized glands after biosynthesis in epithelial cells, followed by secretion into the lumen of these glands where the proteins are stored prior to spinning into fibers. As fibrous proteins, silks differ widely in amino acid composition, structure, and properties depending on the specific source. Generally, silks are characterized by a combination of highly repetitive primary sequences that lead to significance in secondary structure, which provides unique mechanical properties. Because of these impressive properties, combined with their biocompatibility and relative environmental stability, silks have been used as an important set of material options in the fields of controlled release, biomaterials, and scaffolds for tissue engineering for decades, in addition to their uses as textile materials. There are thousands of silk-spinning insects and spiders, yet only a few have been investigated in detail.

Bombyx mori SILK

The most extensively characterized silk is from the domesticated silkworm *Bombyx mori*, which has a number of desirable properties, i.e., luster, comfort, high strength, and elasticity for textile.^[1] *B. mori* silk has been found to have potential applications in the development and construction of various biotechnological and biomedical devices.^[2–4] World silk production has declined since the 1930s because of competition from less expensive synthetic fibers. Silkworms produce silk during one stage in their lifecycle. The lifecycle of *B. mori* lasts 55–60 days and passes through a series of developmental stages or molts. Cocoon formation occurs at around 26 days in the cycle, during the later stages of the fifth instar. This stage is characterized by the extrusion of silk from a spinneret located in the head and drawing of the fiber by a characteristic figure-of-eight head movement. *B. mori* silk consists of

two fibroin protein threads adhered together with a sericin protein gum resulting in a single thread about 10 to 25 μm in diameter.^[5] The morphology of silk fibroin threads is triangular in cross-section. Here, we would like to give an overview of the key component in *B. mori* silk, fibroin, which has been studied intensively.

Bombyx mori Silk Structure

Primary structure

The unique properties of *B. mori* silk fibroin fiber should be attributed to the distribution of crystalline and amorphous domains formed in the process of spinning by protein–protein interactions. The silk fibroin from *B. mori* comprises high-molecular-weight polypeptides containing a predominance of the amino acids glycine, alanine, and serine. The overall composition of silk fibroin in mol% consists of glycine (42.9%), alanine (30.0%), serine (12.2%), tyrosine (4.8%), and valine (2.5%).^[1,5] Acidic side-chain groups predominate by 2.3-fold over basic side-chain groups. The silk fibroin consists of two primary peptides, one approximately 325,000 Da (heavy chain) and the other 25,000 Da (light chain). Using the cDNA sequencing method, Zhou et al., employing shotgun sequencing strategy combined with traditional physical map-directed sequencing of the fibroin gene of the heavy chain, predicted the presence of unusual repeat sequences in the silk fibroin,^[6] shown in Fig. 1. The silk fibroin protein can be described as a block copolymer containing crystalline domains (regions of the polypeptide containing the short side-chain amino acids), interrupted with amorphous domains consisting of the bulkier side-chain amino acids. The crystalline domain of silk fibroin is predominantly constituted with the highly repetitive hexameric AGSGAG motif. A more extended sequence has been reported for the crystalline domain precipitated after chymotrypsin hydrolysis, as GAGAGSGAA (SGAGAG)_nY, where *n* is usually 6–10. This sequence accounts for 55% of fibroin. Other characteristics of the primary structure in *B. mori* silk fibroin are repeating motifs such as AGYGAG and AGVGYGAG, as well as the irregular unit GAAS.^[6,7]

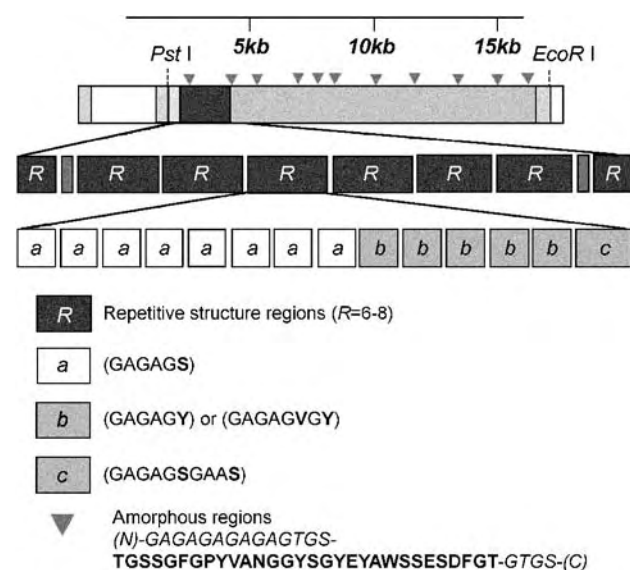


Fig. 1 Schematic representation of the organization of the *B. mori* silk fibroin heavy chain gene. (From Refs. [6,7].)

Secondary structure

Two key crystalline structures have been identified for *B. mori* silk fibroin, silk I before spinning and silk II after spinning the silk fiber. The corresponding structures have been investigated by X-ray fiber diffraction, electron diffraction, conformational energy calculations, infrared spectroscopy, and nuclear magnetic resonance (NMR). Despite a long history of studying silk I, its structure determination was difficult because any attempt to induce a macroscopic orientation of the sample for X-ray diffraction, electron diffraction, or solid state NMR readily causes a conversion of the silk I form to the silk II form. Recently, Asakura et al. have resolved the molecular conformation of silk I as a “repeated β -turn type II” (Fig. 2), using solid-state NMR methods such as 2-D spin diffusion NMR under off magic angle spinning, Rotational Echo Double Resonance (REDOR), and ^{13}C chemical shift data.^[8] Instead of *B. mori* silk fibroin, the sequential model peptide (Ala-Gly)₁₅ was used. The torsional angles of (Ala-Gly)₁₅ with silk I structure were determined as $(-60^\circ \pm 5^\circ, 130^\circ \pm 5^\circ)$ and $(70^\circ \pm 5^\circ, 30^\circ \pm 5^\circ)$ for Ala and Gly residues, respectively. The formation of the intramolecular hydrogen bonding along the chain was confirmed from REDOR NMR by determination of the interatomic distance between the nitrogen and carbon atoms composing the intramolecular hydrogen bonding.

Concerning the structure of silk II, Marsh et al. were the first to propose an antiparallel β -sheet model based on a fiber diffraction study of native *B. mori* silk fibroin fiber.^[9] Although there were some reports suggesting the possession of some intrinsic structural disorder,

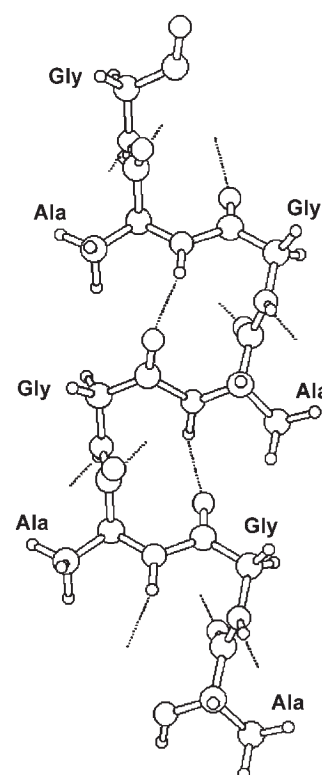


Fig. 2 The conformation of a repeated β -turn type II-like molecule as a model for silk I proposed by Asakura et al. (From Ref. [8].)

they are basically in accordance with Marsh’s model. This is because X-ray diffraction cannot yield any results on the amorphous domains of silk fibroin in silk II form, in which the repeated Ala-Gly sequences are known to contain interspersed Tyr and Val residues as shown in Fig. 1. However, by using ^{13}C solid-state NMR, Asakura et al. have clarified not only the heterogeneous structure of crystalline domains, as well as a heterogeneous structure of the natural fiber from *B. mori* silk fibroin with silk II form, but also the detailed molecular conformation and packing of the protein chains in the crystalline domains.^[10] The relative proportions of the various heterogeneous components in the native *B. mori* silk fiber were determined from their relative peak intensities, as shown in Fig. 3. Although there are some contents of distorted structure, nearly 70% of the repeated sequences in the crystalline domains are present as two types of antipolar antiparallel β -sheet structures with different packing model.

Bombyx mori Silk Fiber Formation

The unique properties of silk fibers result from the primary and higher-order structures of the protein as well as the processing conditions used by the organism

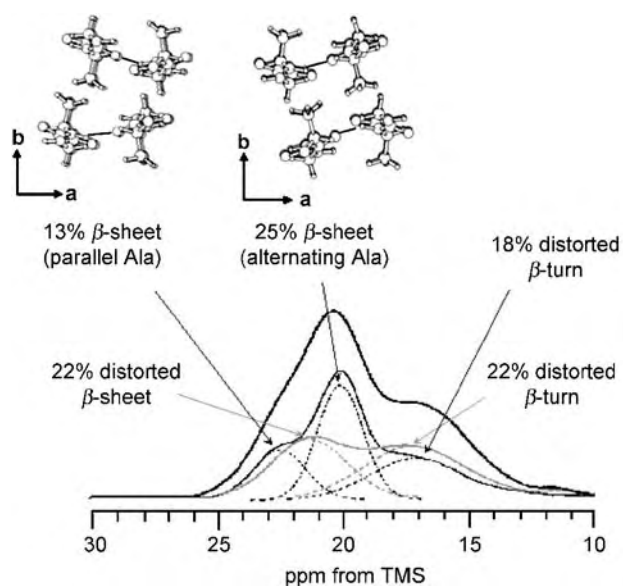


Fig. 3 Expanded Ala C β peaks of native *B. mori* silk fibroin fiber, the Cp fraction, and their difference spectrum. Shown as dotted lines underneath are the spectral deconvolutions with Gaussian peaks. The detailed peak assignments are described in Ref. [10].

to spin fibers. The processing conditions function to produce molecular orientation in the protein chains. This occurs at ambient temperatures in aqueous solution, unlike the conditions employed in synthetic fiber processing.

B. mori silk fibroin is produced in a single-cell type lining a pair of long tubelike glands located on each side of the silkworm. The gland consists of three relatively distinct regions, the posterior section where the fibroin is synthesized by epithelial cells, the middle gland where the fibroin product is stored and the sericin is synthesized but separated from the fibroin, and the anterior section where mixing of fibroin and sericin occurs and the two glands join to produce a single duct leading to the spinneret. The concentration of fibroin is approximately 30% when it is stored in middle gland. The anterior division of the silkworm gland is very narrow, 0.05 to 0.3 mm in diameter, which results in a high shear rate of 2 to 400/s at 1.0 cm/s spinning rate.^[11] The soluble silk in the silk gland undergoes a conformational change owing to mechanical shear, drawing at the spinneret, and loss of water.

The fiber formation mechanism has been studied in vivo and in vitro through the use of a variety of techniques including X-ray crystallography, infrared spectroscopy, and NMR spectroscopy. Magoshi et al. first studied liquid crystalline behavior of silkworm silk and reported that fibroin from the anterior portion of *B. mori* silk gland exhibited liquid crystalline properties.^[12] A nematic liquid crystalline phase was presumed

by the authors, which can be drawn by hand at 7 mm/s, similar to the approximate rate of silkworm fiber spinning to form an orientated β -sheet conformation. They also studied the viscosity of liquid silk at different locations within the silk gland and found a decrease in viscosity as the silk progressed towards the anterior division of the gland, suggesting an ordering of structure. The pH of fibroin solution decreased to 4.0 as it approached the anterior division of the gland, which correlates with the formation of β -form crystals at pH 4.5.^[12]

Iizuka reported on shear rates required to convert the soluble silk to the β -form in the anterior portion of silk gland leading to the spinneret, and the acceleration of this effect by calcium ions.^[11] On the basis of the solution properties and evidence of ordered structure within the silk gland, it was speculated that the silkworm simulates a liquid-crystal spinning process during spinning silk fiber. At an assumed spinning rate of 1.0 cm/s, the author chose a critical shear rate of aqueous silk to be 10^2 to 10^3 /s for polymer concentrations between 15 and 30% (wt). Microfibrils formed owing to structural transition, with increasing shear rate as the solution processed toward the spinneret.

Asakura et al. have determined the molecular conformations of *B. mori* silk fibroin before (silk I structure) and after spinning (silk II structure) using solid-state NMR technique as mentioned above. To reproduce and validate the structural transition of *B. mori* silk fibroin from silk I to silk II, a molecular dynamics simulation was performed, starting from 16 fully hydrated (Ala-Gly)₆ molecules in the silk I conformation.^[13] The parameters were chosen to mimic the effects of dehydration, shear force, and stretching during spinning in the silkworm, which were proposed to induce the structural transition from silk I to silk II. After releasing the shear and stretch forces applied during the MD simulation, water was removed and the system allowed to equilibrate. The resulting ϕ and ψ angles of the Ala residues are plotted, before and after the equilibration period, respectively. A considerable amount of antiparallel β -sheet structure emerges around $\phi = -150^\circ$ and $\psi = 150^\circ$. Additionally, a significant proportion of the Ala residues assumes a β -turn conformation, as in silk I. The relative areas of these two predominant conformations are found to be 65% and 35%, when counting the number of data points within the characteristic ranges of $\phi = -150^\circ \pm 20^\circ$ and $\psi = 150^\circ \pm 20^\circ$ for β -sheets, and $\phi = -60^\circ \pm 30^\circ$ and $\psi = 130^\circ \pm 30^\circ$ for β -turns. This result agrees well with the respective ratios of 68% and 32% determined for the quasicrystalline Cp fraction in the silk II form.^[10] Thus, the structural transition from silk I in aqueous solution into solid-state silk II, i.e., the fiber formation of *B. mori* silk fibroin, could be reproduced by MD simulation.

***Bombyx mori* Silk Properties**

General characteristics

B. mori silk fibers are insoluble in water, dilute acids and alkali, and most organic solvents. In addition, silks are resistant to most proteolytic enzymes, and are hydrolyzed by concentrated sulfuric acid. Lithium bromide (9.0 M) solubilizes silk fibroin fibers, while specific mixtures of acids solubilize silk fibers without decreasing the molecular mass of the polymer. Silk assembles into a compact conformation as shown by intrinsic viscosity measurements. For example, a solution containing fibroin of molecular size 300 kDa exhibits a viscosity significantly below values obtained for comparably sized and smaller sized synthetic polymers. Silk fibers are hygroscopic, with a moisture region of 10 to 15%. The specific gravity for *B. mori* silk is 1.25.

Mechanical properties

B. mori silks exhibit an unusual combination of strength and toughness which sets them apart from other natural and synthetic fibers. The tensile strength of *B. mori* silk fibroin fiber is approximately 650 MPa, the modulus is approximately 15 Gpa, and the elongation to break is 15–20%.^[5] Silk fibers represent an exception to most materials where strength and modulus increase with increased loading rates, while elongation decreases. With silks, increasing rates of loading result in increased elongation, so that work to rupture increases at higher rates of loading. At low loading rates, *B. mori* silk has a higher modulus than nylon but lower strength, while at high loading rates silk out-performs nylon.

Utility of *Bombyx mori* Silk in Biomedical Applications

Sutures

Silk has been used most extensively as sutures for wound ligation and has become the most common natural suture used in the biomedical industry, surpassing even collagen.^[14] During decades of use, silk fibers have proven to be effective in many clinical applications. In general, sutures should be strong, handle easily, and form secure knots. Virgin silk (silk fibroin containing sericin gum), like most proteins, is a potential allergen causing a Type I allergic response in some cases and inducing asthma and specific upregulation in IgE levels. Wen et al. showed the extracted sericin was responsible for sensitization by skin testing of 64 children with asthma.^[15] Further biochemical analysis by Zaoming et al. concluded that upregulated IgEs were

produced in response to the sericin, clarifying the role of these contaminants as the main allergenic agent in silk and not the core fibroin fibers.^[16] As sutures, the silk fibroin fibers are usually coated with waxes or silicone to enhance material properties and reduce fraying; these sutures are commonly referred to as black braided silk (e.g., Perma-Hand™). The transition from virgin silk to black braided sutures occurred in the late 1970s and early 1980s; however, virgin silk is still commercially available today. Although the use of virgin silk negatively impacted the general acceptance of this biomaterial from the surgical practitioner perspective in the earlier stage, to date, there is no report found that implicates black braided silk in inducing hypersensitivity and allergic reactions. All biomaterials derived from a nonautologous source will elicit some level of foreign body response following implantation in vivo. In a comprehensive blind study, Setzen et al. statistically analyzed tissue responses to 11 types of sutures subcutaneously implanted in the rabbit.^[17] They concluded that the inflammatory response induced by silk was not greater than that in response to common absorbable sutures 30 days post implantation. Santin et al. also found that silk was no more thrombogenic than common biomedical materials such as polystyrene and poly(2-hydroxyethyl methacrylate).^[18]

Silk is degradable but over longer time periods due to proteolytic degradation usually mediated by a foreign body response. Silk fibers lose the majority of their tensile strength within one year in vivo, and fail to be recognized at the site within two years. In general, silk is slowly absorbed in vivo. The rate of absorption is dependent upon the implantation site, mechanical environment, and variables related to the health and physiological status of the patient, and the type (e.g., virgin silk and black braided fibroin) and the silk fiber size. Furthermore, alterations in silk processing may cause conformational changes in the protein structure potentially increasing or decreasing susceptibility to degradation. However, none of these variables has been studied in detail, so that it is difficult to gain a clear understanding of the relationships between structure, processing, and degradability.

Biosensors

In addition, it has been demonstrated that silk can be used as a substrate for enzyme immobilization in biosensors. In the preparation of an enzyme-immobilized film with fibroin, simultaneous insolubilization of the water-soluble film and immobilization of the enzyme in the film is possible without any chemical reagents for cross-linking.^[3] The sensor was prepared with a glucose oxidase (GOD)-immobilized silk fibroin film attached on an oxygen electrode surface and assembled in the

apparatus. Moreover, a fourfold increase in sensitivity was observed when the glucose sensor was made with the GOD immobilized on nonwoven silk fabrics compared with a GOD-immobilized silk fibroin film.^[19] The membrane potential of the GOD-immobilized silk fibroin film is induced by an enzymatic reaction after the addition of the substrate glucose. An application for the process has been shown by the preparation of a new type of glucose sensor that functions without an oxygen electrode. Biophotosensors have also been made using fibroin films with immobilized peroxidase.

The application of the immobilized peroxidase (POD) *Bombyx mori* silk fibroin membrane to biophotosensors was attempted for the purpose of determining the concentration of hydrogen peroxide generated by the luminol reaction.^[20] POD was entrapped in the silk fibroin membrane without any chemical cross-linking agents. The response time and intensity of the photocurrent from the photodiode determined after the luminol reaction depends on the character of the immobilized POD-membrane. Namely, when POD is distributed asymmetrically in the membrane, the response is very fast and strong compared with POD distributed uniformly in the membrane. The photocurrent after the luminol reaction increases linearly with the concentration of hydrogen peroxide in the aqueous solution, indicating the applicability, of such a silk fibroin membrane to biophotosensors.

Qian et al. described a procedure to prepare an amperometric H₂O₂ sensor, based on the immobilized peroxidase in regenerated silk fibroin membrane, in the use of new methylene blue N as an electron transfer mediator.^[21] The characteristics of the sensor with respect to linearity, response time, effect of pH and temperature, stability, and reproducibility were investigated. The sensor was highly sensitive to H₂O₂ with a detection limit of 1.0×10^{-7} M and a response time of less than 40 s.

Scaffolds for tissue engineering

The emergence of tissue engineering has increased the demand for a diverse set of biomaterials to support the development of tissues in vitro prior to implantation in vivo. The biomaterial or matrix plays a key role in communicating or transducing environmental cues to cells seeded within or on the matrix. Silk fibroin offers versatility in matrix scaffold designs for a number of tissue engineering needs in which mechanical performance and biological interactions are major factors for success, including bone, ligaments and tendons, blood vessels, and cartilage.^[4,14,22,23] Silk fibroin can be processed into foams, films, fibers, and meshes.

Inouye et al. demonstrated the utility of a silk fibroin film in culturing animal cells in comparison to a

collagen matrix.^[22] Films of both fibroin and collagen supported equivalent cell growth after 5 days in cell culture conditions. Comparative studies on the ability of fibroin and sericin films in supporting attachment, spreading, and growth of the L-929 fibroblast cell line support that extracted fibroin free of sericin was a suitable matrix for cell and tissue culture. Moreover, Sofia et al. chemically decorated the silk fibroin films with integrin recognition peptide (RGD) to promote integrin interactions for adhesion; the induction of bone formation in vitro was significantly enhanced.^[4] Furuzono et al. developed apatite-silk composites by using an alternate soaking process, that is, a calcium solution and a phosphate solution were applied alternatively to a silk fabric.^[23] Thus, this method may provide a bioactive material with Ca²⁺ ions that is similar to the natural bone apatite.

Recent research with silk has focused on the development of a wire rope matrix for the development of autologous tissue-engineered anterior cruciate ligaments using a patient's own adult stem cells. Human bone marrow stromal cells cultured in collagen gels were grown in a mechanically dynamic environment relevant to that present in vivo leading to the formation of ligament fibroblasts and mature tissue development in vitro. However, the poor integrity of the collagen gels and the demanding and dynamic intraarticular mechanical and biochemical environment of the knee have prompted renewed interest in silk as a long term absorbable material with good mechanical integrity and biocompatibility.

Molecular Design and Synthesis of Silk-Like Materials

The design and synthesis of artificial proteins is an emerging area of research with important implications for structural biology, materials science, and biomedical engineering. Significant progress has been reported in the design of fibrous proteins that adopt predictable secondary structures and have higher order protein folding. With virtually absolute control of sequence, chain length, and stereochemical purity, the artificial proteins can be designed to represent a new class of macromolecular materials, with properties potentially quite different from those of the synthetic polymers currently available and in widespread use.

Tirrell's group was interested in artificial fibrous proteins that are able to form crystals of defined thickness and bearing functional groups at their surface.^[24] For instance, the Glu residues were inserted in (Ala-Gly)_n at intervals varying between 8 and 14. The big size of Glu compared to Ala and Gly should prevent its insertion in the lamellar crystal. Therefore, when crystals form, the Glu residues must be found on the

surface. This is facilitated by the fact that, among the 20 amino acids, Glu is the residue that has the lowest tendency to form β -sheet. Cappello et al. focused on the production of synthetic protein analogs of *B. mori* silk, and other well known structural protein materials and tried to reproduce defined properties of their natural counterparts.^[25] Also, by combining structural blocks from different natural proteins, they hoped to obtain new properties found in neither the synthetic homoblock polymers nor the natural proteins.

Asakura's group has designed and expressed in *Escherichia coli* a number of artificial fibrous proteins derived from *B. mori* silk, as well as other proteins, such as *Samia cynthia ricini* and spider silks, and elastin.^[26,27] The examples of amino acid sequence of designed proteins are as follows, where single amino acid codes are used:

[GGAGSGYGGGYGHGYGSDGG(GAGAGS)₃AS]_n
 [(GVPGV)₂GG(GAGAGS)₃AS]_n
 [TGRGDSPA[(GVPGV)₂GG(GAGAGS)₃AS]_m]_n
 [TGRGDSPAGG(GAGAGS)₃AS]_n
 [(AGSGAG)₃YGGLGSQGAGRAS]_n
 [(A)₁₈GVGAGYGAGAGYGVGAGYGVGAGVGYGA-
 GAGYAS]_n
 [DGG(A)₆GGAASGAGYGA(GGA)₄GAGYGAAS]_n
 [DGG(A)₁₂GGAASGAGYGA(GGA)₄GAGYGAAS]_n
 [GG(A)₆GGAASGAGYGA(GGA)₄GAGYGAAS]_n
 [DGG(A)₁₂GGAASGAGYGVGAGAGSGAGAGSGA-
 GAGYGAAS]_n
 [DGG(A)₆GGAASGAGY(GAGAGS)₂GAGAGY-
 GAAS]_n
 [DGG(A)₁₂GGAASGAGY(GAGAGS)₂GAGAGY-
 GAAS]_n
 [DGG(A)₁₂GGAASGAGY(GAGAGS)₂GAGAGY-
 GAAS]_n
 [DGG(A)₆GGAASGAGY(GAGAGS)₂GAGAGY-
 GAAS]_n

The target genes were cloned in pET30a plasmids (Novagen, Inc.) under control of strong bacteriophage T7 and translation signals, which contain kanamycine-resistant gene (*kan^r*), *lacI* coding gene, an origin of replication (*ori*), and an f1 origin of replication that allows the production of single stranded plasmid DNA for mutagenesis and sequencing applications (Fig. 4). For protein production, the authors transferred recombinant plasmids to *E. coli* strain BL21(DE3)pLysS. In this strain, a gene encoding T7 RNA polymerase is incorporated into the bacterial chromosome under *lacUV5* control, and protein production is induced by the addition of IPTG. The pLysS plasmid provides low levels of T7 lysozyme, which inhibits T7 RNA polymerase and suppresses the basal level of protein expression. The objective of this research is to determine the interactions between different amino acid sequences, as well as the effect of single amino acid residues. These could tell us which amino acid sequence can be introduced into the fibrous proteins to control macromolecular structure and provide novel chemical or biological activities while

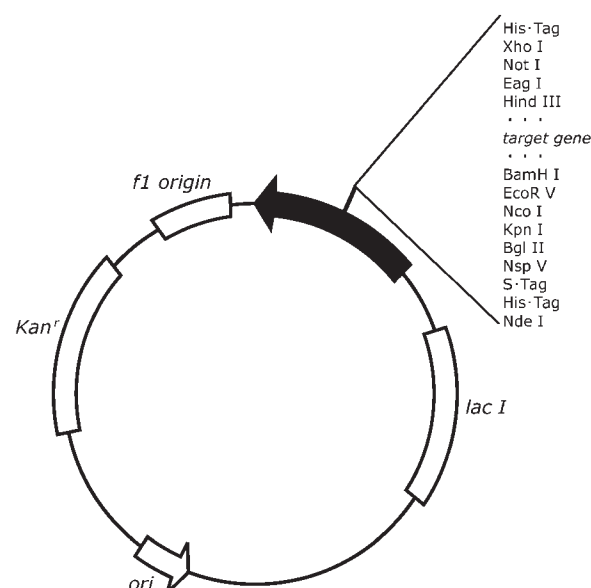


Fig. 4 Structure of an expression vector used for the production of recombinant protein in *E. coli*.

maintaining the physical properties of the structural blocks.

WILD SILKWORM SILKS

Many kinds of wild silkworm produce silks, including *Antheraea yamamai*, *Antheraea pernyi*, *Antheraea mylitta*, *Samia cynthia ricini*, etc. Wild silkworm silks also consist mainly of silk fibroin and sericin, but the content of sericin is relatively low, as is expected from the fact that the size of the middle region of the silk gland is more slender than that of the posterior region of the gland. The major amino acid residues of these wild silkworm silk fibroins are alanine and glycine; as with *B. mori*, the sum of these two amino acid residues composes 74–82% of the silk fibroin. However, the relative content of alanine is larger than that of glycine, in contrast to the case of *B. mori* silk fibroin, where the reverse is the case.

The primary structures of *A. pernyi* and *S.c. ricini* silkworm silk fibroins have recently been shown to be very similar to each other. Namely, the silk fibroins mainly consist of the repeated similar sequences by about 80–100 times where there are alternative appearances of polyalanine region of (Ala)_n, where $n = 10\text{--}14$, and glycine-rich region. Structural analyses of the aforementioned two kinds of silk fibroin prepared from liquid silk in silkgland and silk fibroin fibers suggest that there is a structural transition from α -helix to β -sheet occurring in the polyalanine region.^[28] Recent results suggest that the structural transition also occurs in the glycine-rich region of *A. pernyi* and *S.c. ricini* silk fibroins.

The characteristics of wild silkworm fibroin fibers are considerably different from those of *B. mori* silk fibroin fibers. For example, the stress–strain curves are quite different. The curve is a monotonic rapid increase in the case of *B. mori* silk fiber, while the curve shows a slow increase with a flat region in the case of wild silkworm silk fibers.^[1,6] Compared with *B. mori* silk fibers, the wild silkworm silk fibers have much lower solubility, so that the aqueous solution of silk fibroin is usually prepared from the liquid silk stored in the posterior region of the silk gland directly. The cast film from this aqueous solution mainly takes α -helix structure with high thermal stability, which can be changed to β -sheet with stretching or immersing in methanol. Such kinds of films can also be used for the immobilization of GOD and applied to a glucose sensor.

SPIDER SILKS

In the past decade, spider silks have been the most attractive fibrous proteins because of their remarkable mechanical properties. All spiders can make seven kinds of silk, one type to wrap its prey, one to wrap its eggs, and five to construct the web. Each type is biologically engineered to perform its task to perfection. The different silk proteins contain repetitive amino acid sequences that vary slightly, depending on their use. These small changes in sequence can confer different mechanical properties.

Most molecular and structural investigations of spider silks have focused on dragline silk, which often refers to a major ampullate silk because it is produced in the major ampullate glands, and flagelliform silk, which is the core fiber of the capture spiral. Dragline silk, used to anchor the web and form its spokes, has a higher tensile strength than high-tensile steel of the same weight. Capture silk, which is used to snare insects, is very sticky and can stretch to three times its length before breaking. Generally, dragline silk contains multiple repeats of motifs that include an 8–10 residue–long polyalanine block and a 24–35 residue–long glycine-rich block. When fibroins are spun into fibers, the polyalanine blocks form β -sheet crystals that cross-link the fibroins into a polymer network with great stiffness, strength, and toughness.^[29] On the other hand, like all silks, the flagelliform silk is composed largely of repeated motifs. The dominant repeat of this protein is Gly-Pro-Gly-Gly-X (X is mainly Ala, Ser, Tyr, or Val), which can appear up to 63 times in tandem arrays. This motif likely forms Pro²-Gly³ type II β -turns and the resulting series of the concatenated β -turns are thought to form a β -spiral. This spring-like helix is proposed to be responsible for the elasticity of this protein.^[30]

Spiders cannot be farmed like silkworms because of their aggressive and territorial nature; therefore, materials scientists have tried to biomimic the process of spider silk production in the laboratory. Lazaris et al. describe the production of a number of dragline silk proteins in mammalian cells and show, for the first time, that harvested recombinant proteins can be spun into strong, lightweight fibers.^[26] Although the fibers are not as good as natural dragline silk, some improvements might be gained from making larger proteins by spinning the constituent proteins of natural dragline silk.

CONCLUSION

Although the substitution of natural materials has been underway for several decades, natural silks remain commercially important because of their unique properties, as well as their relative environmental stability and biocompatibility, and (in some cases) consumer preferences. During the past decade a great deal of progress has been made in understanding silk genetic and protein structures. On the other hand, materials scientists have long been fascinated by structure–function relationships in silk proteins. The illustration of the sequences of cDNAs and genomic clones encoding silks can therefore be readily exploited to construct genetically engineered silklike proteins using synthetic oligonucleotides. Cloning and expression of native and synthetic silks have been achieved in a variety of host systems. The studies on natural silks summarized here may be able to provide important insights into the nature of macromolecular materials, both natural and synthetic.

ACKNOWLEDGMENTS

TA acknowledges the support from the Asahi Glass Foundation and the Insect Technology Project, Japan.

ARTICLES OF FURTHER INTEREST

Bifunctional Polymers; Polyamides (Synthetic and Natural); Sutures

REFERENCES

1. Asakura, T.; Kaplan, D.L. Silk Production and Processing. In *Encyclopedia of Agricultural Science*; Arutzen, C.J., Ed.; Academic Press: New York, 1994; Vol. 4, 1–11.
2. Kaplan, D.L.; Mello, C.M.; Arcidiacono, S.; Fossey, S.; Senecal, K.; Muller, W. Silk. In *Protein Based Materials*.

- McGrath, K., Kaplan, D.L., Eds.; Birkhauser: Boston, 1998; 103–131.
3. Demura, M.; Asakura, T. Immobilization of glucose oxidase with *Bombyx mori* silk fibroin by only stretching treatment and its application to glucose sensor. *Biotechnol. Bioeng.* **1989**, *33*, 598–603.
 4. Sofia, S.; McCarthy, M.B.; Gronowicz, G.; Kaplan, D.L. Functionalized silk-based biomaterials for bone formation. *J. Biomed. Mater. Res.* **2001**, *54* (1), 139–148.
 5. Hojo, N., Ed. *Zoku Kenshi no Kozo (Structure of Silk Fibers)*; Shinshu University: Ueda, Japan, 1980.
 6. Zhou, C.-Z.; Confalonieri, F.; Medina, N.; Zivanovic, Y.; Esnault, C.; Yang, T.; Jacquet, M.; Janin, J.; Duguet, M.; Perasso, R.; Li, Z.-G. Fine organization of *Bombyx mori* fibroin heavy chain gene. *Nucleic Acids Res.* **2000**, *28* (12), 2413–2419.
 7. Asakura, T.; Sugino, R.; Yao, J.; Takashima, H.; Kishore, R. Structural analysis of semi-crystalline *Bombyx mori* silk fibroin chain with Silk I and Silk II forms by ^{13}C , ^{15}N and ^2H stable isotope labeling, conformation-dependent chemical shifts and solid state NMR spectroscopy. *Biochemistry* **2002**, *41* (13), 4415–4424.
 8. Asakura, T.; Ashida, J.; Yamane, T.; Kameda, T.; Nakazawa, Y.; Ohgo, K.; Komatsu, K. A repeated β -turn structure in poly(Ala-Gly) as a model for silk I of *Bombyx mori* silk fibroin studied with two-dimensional spin-diffusion NMR under off magic angle spinning and rotational echo double resonance. *J. Mol. Biol.* **2001**, *306* (2), 291–305.
 9. Marsh, R.E.; Corey, R.B.; Pauling, L. An investigation of the structure of silk fibroin. *Biochim. Biophys. Acta* **1955**, *16*, 1–34.
 10. Asakura, T.; Yao, J.; Yamane, T.; Umemura, K.; Ulrich, A.S. Heterogeneous structure of silk fibers from *Bombyx mori* resolved by ^{13}C solid-state NMR spectroscopy. *J. Am. Chem. Soc.* **2002**, *124*, 8794–8795.
 11. Iizuka, E. Silk thread: Mechanism of spinning and its mechanical properties. *J. Appl. Polym. Sci.* **1985**, *41*, 173–185.
 12. Magoshi, J.; Magoshi, Y.; Nakamura, S. Crystallization, liquid crystal, and fiber formation of silk fibroin. *J. Appl. Polym. Sci.* **1985**, *41*, 187–204.
 13. Yamane, T.; Umemura, K.; Nakazawa, Y.; Asakura, T. Molecular dynamics simulation of conformational change of poly(Ala-Gly) from repeated β -turn type II to β -sheet in relation to fiber formation mechanism of *Bombyx mori* silk fibroin. *Macromolecules* **2003**, *36*, 6766–6772.
 14. Altman, G.H.; Diaz, F.; Jakuba, C.; Calabro, T.; Horan, R.L.; Chen, J.; Lu, H.; Richmond, J.; Kaplan, D.L. Silk-based biomaterials. *Biomaterials* **2003**, *24* (3), 401–416.
 15. Wen, C.M.; Ye, S.T.; Zhou, L.X.; Yu, Y. Silk-induced asthma in children: A report of 64 cases. *Ann. Allergy* **1990**, *65* (5), 375–378.
 16. Zaoming, W.; Codina, R.; Fernandez-Caldas, E.; Lockey, R.F. Partial characterization of the silk allergens in mulberry silk extract. *J. Investig. Allergol. Clin. Immunol.* **1996**, *6* (4), 237–241.
 17. Setzen, G.; Williams, E.F. Tissue response to suture materials implanted subcutaneously in a rabbit model. *Plast. Reconstr. Surg.* **1997**, *100* (7), 1788–1795.
 18. Santin, M.; Motta, A.; Freddi, G.; Cannas, M. In vitro evaluation of the inflammatory potential of the silk fibroin. *J. Biomed. Mater. Res.* **1999**, *46* (3), 382–389.
 19. Asakura, T.; Kitaguchi, M.; Demura, M.; Sakai, H.; Komatsu, K. Immobilization of glucose oxidase on nonwoven fabrics with *Bombyx mori* silk fibroin gel. *J. Appl. Polym. Sci.* **1992**, *46*, 49–53.
 20. Demura, M.; Asakura, T.; Nakamura, E.; Tamura, H. Immobilization of peroxidase with a *Bombyx mori* silk fibroin membrane and its application to biophoto-sensors. *J. Biotech.* **1989**, *10*, 113–119.
 21. Qian, J.; Y.; Liu, H.; Yu, T.; Deng, J. An amperometric new methylene blue N-mediating sensor for hydrogen peroxide based on regenerated silk fibroin as an immobilization matrix for peroxidase. *Anal. Biochem.* **1996**, *236* (2), 208–214.
 22. Inouye, K.; Kurokawa, M.; Nishikawa, S.; Tsukada, M. Use of *Bombyx mori* silk fibroin as a substratum for cultivation of animal cells. *J. Biochem. Biophys. Methods* **1998**, *37* (3), 159–164.
 23. Furuzono, T.; Taguchi, T.; Kishida, A.; Akashi, M.; Tamada, Y. Preparation and characterization of apatite deposited on silk fabric using an alternate soaking process. *J. Biomed. Mater. Res.* **2000**, *50* (3), 344–352.
 24. Krejchi, M.T.; Atkins, E.D.T.; Waddon, A.J.; Fournier, M.J.; Mason, T.L.; Tirrell, D.A. Chemical sequence control of β -sheet assembly in macromolecular crystals of periodic polypeptides. *Science* **1994**, *265* (5177), 1427–1432.
 25. Cappello, J.; Ferrari, F. Microbial Production of Structural Protein Polymers. In *Plastics from Microbes*; Mobley, D.P., Ed.; Hanser: New York, 1994; 35–92.
 26. Asakura, T.; Tanaka, C.; Yang, M.; Yao, J.; Kurokawa, M. Production and characterization of a silk-like hybrid protein, based on the polyalanine region of *Samia cynthia ricini* silk fibroin and a cell adhesive region derived from fibronectin. *Biomaterials* **2004**, *25*, 617–624.
 27. Yao, J.; Asakura, T. Synthesis and structural characterization of silk-like materials incorporated with elastic motif. *J. Biochem.* **2003**, *133*, 147–154.
 28. van Beek, J.D.; Beaulieu, L.; Schafer, H.; Demura, M.; Asakura, T.; Meier, B.H. Solid-state NMR determination of the secondary structure of *Samia cynthia ricini* silk. *Nature* **2000**, *405* (6790), 1077–1079.
 29. Simmons, A.; Michal, C.; Jelinski, L. Molecular orientation and two-component nature of the crystalline fraction of spider dragline silk. *Science* **1996**, *271* (5245), 84–87.
 30. Hayashi, C.Y.; Lewis, R.V. Evidence from flagelliform silk cDNA for the structural basis of elasticity and modular nature of spider silks. *J. Mol. Biol.* **1998**, *275* (5), 773–784.



Small-Angle Light Scattering Methods for Soft Connective Tissue Structural Analysis

Michael S. Sacks

Engineered Tissue Mechanics Laboratory, McGowan Institute for Regenerative Medicine, University of Pittsburgh, and Department of Bioengineering, University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.

INTRODUCTION

Quantification of the architecture of the soft tissue fibrous extracellular matrix is a critical step in furthering our understanding of changes in tissue function in health and disease. Moreover, many soft tissues, either in their native or processed forms, are increasingly utilized in medical therapies and devices. In the technique of small-angle light scattering (SALS), monochromatic light is passed through a thin specimen. The spatial intensity distribution of the resulting scattered light represents the sum of all structural information within the light beam envelope. For soft tissue membranes or where thin tissue sections can be prepared, SALS provides a straightforward and rapid means to map the fiber architecture. SALS has been shown to measure fiber orientation accurately for up to a tissue thickness of at least 500 μm to an angular resolution of 1° and a spatial resolution of 250 μm . The resulting information can be used to understand native tissue function, as well as to aid the development of engineered tissues. In this article, the SALS technique is reviewed as a general method to quantify connective tissue fibrous structure. In addition to providing important specific application information, it is also hoped that this chapter will serve as a general guideline for the application of SALS in connective tissue analysis.

BACKGROUND

It is well known that the collagen and elastin fibers that compose soft connective tissues can efficiently withstand tensile forces, but have low torsional and flexural stiffness.^[1] Thus, directions in which fibers are oriented can be identified with the directions in which the tissue is able to withstand the greatest tensile stresses. Gross fiber orientation thus leads to an understanding and predictability of the mechanical properties of soft tissues. Although available microscopic techniques can provide quantitative information on fiber structure,^[2] this information is by definition highly localized, making it difficult to quantify larger scale fiber

structural features. Thus we need a rapid method to map the complete gross fiber structure of planar tissues.

Several methods have been previously used for the analysis of the fiber architecture in soft tissues.^[2] Image processing techniques, which use either optical^[3,4] or scanning electron microscopic images,^[5-7] have been used to provide highly localized information on the size and orientation of fibers. However, these are destructive approaches requiring time-consuming tissue preparations, and are thus not suitable to large-scale mapping studies. Small angle X-ray scattering techniques have been used to analyze the collagen molecular architecture.^[2,8-10] In this technique, scattering patterns are recorded on film that requires long exposure time, and the technique suffers from similar mapping limitations of the electron microscopy methods. Polarized light microscopy has often been used to assess the gross collagen fiber architecture of soft tissues by taking advantage of their inherent birefringence.^[2,11] For example, Hilbert et al. applied polarized light microscopy to virgin and accelerated tested bioprosthetic heart valve tissues to visualize fiber directions and crimp.^[12] They were able to make accurate crimp measurements and could obtain general information on fiber orientation. However, this technique is difficult to apply as a detailed mapping method since it requires extensive manual operation, and the inability to distinguish multiple layers of fibers deep within the tissue accurately limits measurement to the more superficial layers.

Recently, Tower and Tranquillo developed a novel technique for extending polarized light fiber alignment methodologies by generating polarized light retardation and alignment direction directly from single images.^[13,14] The primary advantage of this approach is the ability to quantify preferred fiber directions of a complete specimen image without scanning. However, this approach is limited in that while it provides preferred fiber direction and an *index* of the strength of orientation, it does not provide the actual fiber angular distribution. Further, as in all polarized light techniques, it is only applicable to tissues composed exclusively of collagen fibers because of the need for fiber optical birefringence inherent in the technique.

In the SALS technique, monochromatic light is passed through a thin specimen. The spatial intensity distribution of the resulting scattered light represents the sum of all structural information within the light beam envelope. For soft tissues, Helium–Neon laser light is generally preferred since its wavelength ($\lambda = 632.8$ nm) is within an order of magnitude of the diameter of the collagen and elastin fibers and is unpolarized so that there is no directional dependence to any particular angular orientation. SALS has been used to quantify the fibrous structure of synthetic polymers^[15–19] and biopolymers such as collagen films.^[20,21] Moreover, Kronick and Buechler^[22] found close agreement between the fiber orientation measurements made by SALS and X-ray diffraction, indicating that SALS was an inexpensive alternative to X-ray diffraction for quantitative measurements of fibrous materials.

SALS has also been used extensively to quantify the structure of native connective tissues. Kronick and Sacks^[23] used SALS to quantify structural variations in calf skin suffering from a genetic collagen fiber pathology. Ferdman and Yannas^[24] have used SALS for skin scar analysis and were able to detect structural variations successfully in regions of scar from normal skin. Several investigators have also made extensive use of SALS to investigate the interrelationship between the layered structure of the collagen fibers and resultant light scattering behavior in the cornea.^[25–30] Over the past decade, Sacks and colleagues have applied SALS to several tissues, including the diaphragmatic central tendon,^[31] bovine pericardium,^[32–34] cranial dura mater,^[35] tissue-engineered skin,^[36] native and processed small intestinal submucosa,^[37,38] the native aortic^[39] and bioprosthetic valve,^[40,41] the native mitral valve,^[42] and the effects of TGF β on skin wound healing.^[43]

In addition to quantifying tissue structure, the fiber architectural information obtained from SALS can be related to the direction and degree of tissue mechanical anisotropy. For example, the maps of the collagen fiber structure of the diaphragmatic central tendon,^[31] and in particular the volume fraction of the collagen fiber populations, were consistent with the direction and degree of mechanical anisotropy.^[44] More mechanistic approaches were taken by Billiar and Sacks,^[45] who developed a method for measuring fiber kinematics in two-dimensional soft collagenous tissues. The technique combined SALS and biaxial stretch controlled by simultaneous optical strain measurement. Findings for the aortic valve and bovine pericardium indicate that local fiber kinematics are highly tissue specific, and that local fiber deformations may not always follow global tissue strains (i.e., are generally nonaffine). The mobility of the fibers in each tissue seems to be specialized to perform a distinct physiological function

and underscores the importance of measuring fiber kinematics for each specific tissue type that is to be modeled.

The availability of quantitative fiber architectural information obtainable by SALS motivated a recent study by Sacks^[46] wherein a structural constitutive model formulation incorporating the discrete quantitative fiber architectural data derived from SALS was developed. Demonstration of the model was performed using existing biaxial mechanical and fiber orientation data for native bovine pericardium.^[34,47] The structural constitutive model accurately predicted the complete measured biaxial mechanical response. An important aspect of this approach is that only a single equibiaxial test to determine the effective fiber stress–strain response and the SALS-derived fiber orientation distribution were required to determine the complete planar biaxial mechanical response. This study demonstrated that the fiber architectural data obtained from SALS can be directly utilized in mechanical modeling of soft tissue multiaxial behavior.

These cited studies underscore the utility of SALS as a robust technique for obtaining fiber architectural information in soft connective tissues. Moreover, direct utilization of the architectural information into structural constitutive models of soft tissue behavior can avoid ambiguities in material characterization. This is an important step since full realization of structural approaches is limited without direct quantitative structural information for direct implementation or to validate model predictions. However, as with any experimental approach, the results can depend on the exact methodologies utilized. In the following discussion, a detailed description of the instrumentation and analysis techniques utilized in our laboratory is presented. In particular, detailed validation and resolution determination of the SALS technique are demonstrated in an effort to provide investigators important information on how to apply SALS effectively, as well as to unify its application overall.

EXPERIMENTAL SYSTEM AND ANALYSIS METHODS

Introduction

Early SALS devices utilized simple photographic imaging systems or rotating photodiode arrays to digitize the scattered light intensity, as well as requiring manual positioning of the specimen.^[24] Further, analysis of the scattered light intensity data was performed manually.^[24] These methods of data acquisition and analyses are clearly unsuitable for practical applications of mapping large areas of tissue. In our laboratory, we developed a high-speed system that is



able to rapidly raster-scan tissues and automated data analyses. The following is largely drawn from our previous work^[48] and updated where appropriate.

Hardware Description

The SALS device consists of an unpolarized 4-mW HeNe laser ($\lambda = 632.8$ nm, Uniphase, Manteca, CA), spatial filter-beam expander, specimen positioner, projection screen, and a CCD NTSC-compatible auto-iris video camera (Sony Inc., Japan) equipped with a telecentric lens to minimize image distortion (Edmund Scientific, Barrington, NJ), all mounted on a rigid optical platform (Fig. 1). Since the Helium-Neon laser beam has a Gaussian intensity profile, the beam diameter is defined as the diameter of the laser beam that contains all but $1/e^2$ (i.e., 13.5%) of the total radiant laser beam power.^[49] Thus, all reported beam diameters contain 86.5% of the total beam intensity. As the laser light passes through the tissue, it is scattered by the fibrous structures, and the resultant scattering pattern is cast onto the projection screen (Fig. 1). The video image of the scattered light is digitized to a resolution of 640×480 8-bit pixels using

a PCI-bus image grabber board mounted in the personal computer (PC) (Fig. 1). Note that higher resolutions can be used (e.g., $1024 \times 1024 \times 12$ bits), but were found to be unnecessary. Use of the PCI-bus image grabber allowed analysis of the most recent acquired image while the device moved to the next test position and acquired the next image.

The sample positioner allows for two axes of movement in a plane perpendicular to the optic axis (Fig. 1). Both axes utilize a precision linear travel system using 16 threads/inch lead screws with anti-backlash nuts. Positioning of each axis is accomplished by rotating the lead screws using two PC-controlled stepper motors (Arrick Robotics, Hurst, TX) with 400 steps/rev resolution, resulting in a net movement resolution of $4 \mu\text{m}/\text{step}$ and a total travel range of 25 cm along each axis. All motion control, video processing, and data analysis is performed by custom written C++ programs using the GLUT graphics interface software (www.opengl.org). This system allows for video acquisition, initial data analysis (see forthcoming discussion), storage of the analyzed data to the hard disk, and movement to the next test location, all in 0.25 to 0.5 seconds, depending on the distance between test locations.

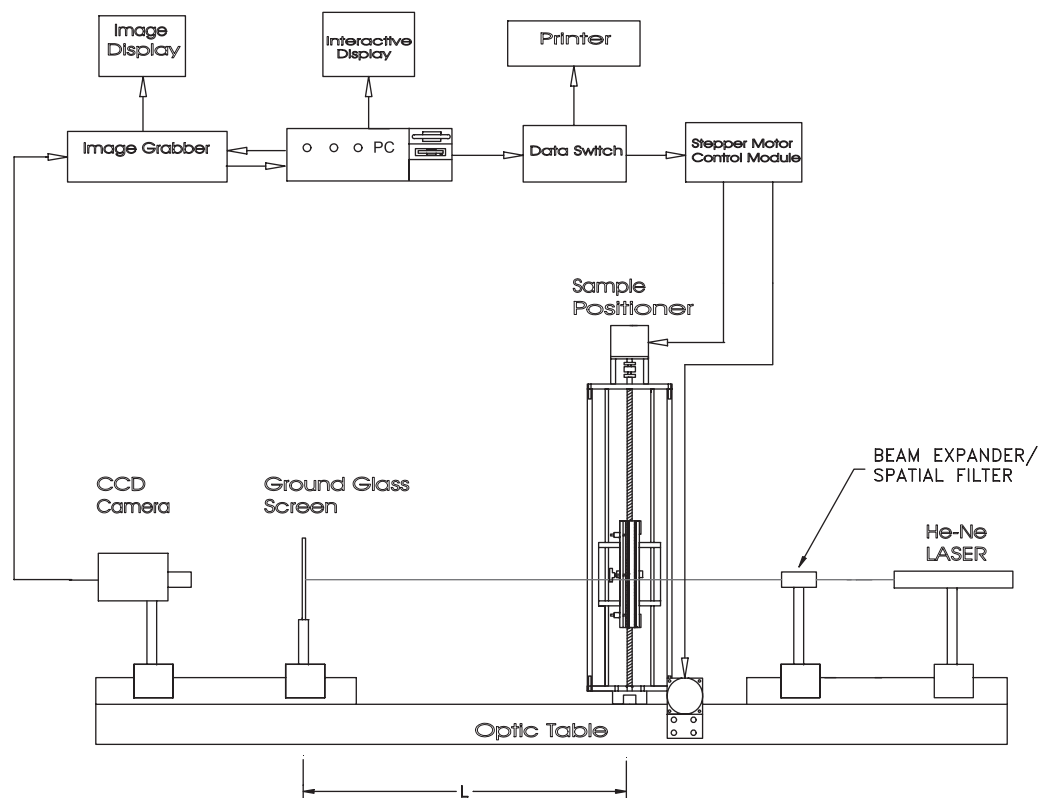


Fig. 1 A schematic of the SALS device, which consists of an unpolarized 4-mW HeNe laser, spatial filter-beam expander, sample positioner, projection screen, and a CCD video camera. The sample positioner allows for two axes of movement in a plane perpendicular to the optic axis.

Analysis of the Light Scattering Pattern

Assuming no optical interactions, a fibrous network optically behaves like a two-dimensional assembly of single slits. This is in contrast to polarized light techniques that require inherent birefringence of the constituent fibers. The Fraunhofer diffraction for a single slit of width w is given by

$$I(\theta) = I_0 \left(\frac{\sin \alpha}{\alpha} \right)^2 \quad (1)$$

$$\alpha = \frac{\pi w}{\lambda} \sin(\theta)$$

where I is the scattered light intensity measured along the scattered angle θ , and I_0 is the incident light intensity.^[50] For two slits separated by a distance d , $I(\theta)$ is now given by

$$I(\theta) = I_0 \left(\frac{\sin \alpha}{\alpha} \right)^2 \cos^2(\beta) \quad (2)$$

$$\beta = \frac{\pi d}{\lambda} \sin(\theta)$$

If the width of the separation of the slits is comparable to the slit width ($w/d \cong 1$), the intensity variations due to slit and slit separation are approximately equal, where w now represents the average of the slit size and interslit spacing. In our experience, this is a very good approximation for dense connective tissues. Although possibly limiting accurate determination fiber diameters from the rate of decrease of the scattered light from the optic axis (see Guinier for related work on X-rays^[51]), all orientation information is preserved since the fibers and the spaces between the fibers have the identical orientations.

Thus, in the case of dense fibrous tissues, the effective slit width is an average of the fiber diameters and the spaces between the fibers. According to single-slit diffraction theory, light is scattered in a direction perpendicular to the fiber axis,^[50] although there will be no observable minima due to the distribution of slit widths (i.e., fiber diameter and interfiber spaces) in the tissue. Each fiber, having an angular orientation Φ , will thus contribute its scattered light intensity at an angular orientation of $\Phi + 90^\circ$. As long as the fiber diameter and interfiber spacing distribution are independent of Φ (as has been our experience in dense connective tissues), the angular distribution of scattered light at a constant R , $I(\Phi)$ represents the angular distribution of fibers about the optic axis within the light beam envelope (Fig. 2a).

To extract the angular fiber orientation information, scattered light intensity values are taken from the digitized SALS image at 1° increments through the

angle Φ about a circle corresponding to a constant radius R centered on the optical axis (Fig. 2a). The extracted intensity distribution, $I(\Phi)$, is taken as the average detected intensity using a 3×3 -pixel grid centered on the current pixel location. Because light is scattered perpendicular to the fiber axes, all $I(\Phi)$ values are actually acquired at $\Phi + 90^\circ$, so that $I(\Phi)$ represents the true fiber orientation (Fig. 2a).

Since the light scattering pattern is symmetric about the preferred fiber direction, the $I(\Phi)$ distribution is repeated twice over the full 360° range (Fig. 2b). The two intensity minima demarcate where the axis of symmetry occurs in the scattering pattern, with the data from either symmetric segment usable for analysis (Fig. 2b). From each symmetric 180° segment, the preferred fiber directions, degree of symmetry of the fiber distribution about the preferred direction, and the degree of fiber orientation can be determined^[22,31] (Fig. 2c). First, we define the preferred fiber direction as the distribution centroid (Φ_c):

$$\Phi_c = \frac{\sum_{\Phi=\Phi_{\min}+1}^{\Phi_{\min}+180} \left[(\Phi - \Phi_{\min}) \frac{I(\Phi) + I(\Phi - 1)}{2} \right]}{\sum_{\Phi=\Phi_{\min}+1}^{\Phi_{\min}+180} \left[\frac{I(\Phi) + I(\Phi - 1)}{2} \right]} \quad (3)$$

where the summations are computed directly from a 180° segment (i.e., Φ_{\min} to $\Phi_{\min} + 180^\circ$) of the SALS data using increments of 1° (the contiguous segment closest to 0° is chosen arbitrarily). Φ_c thus represents the overall preferred direction of the fiber network and does not require the need of any distribution model such as a Gaussian as used in Ref. [31].

If the distribution is symmetrical about the Φ_c , the intensity peak Φ_p will coincide with Φ_c (Fig. 2c). Deviation from the condition of symmetry is defined as the distribution skew, $\Phi_s = \Phi_p - \Phi_c$ (Fig. 2c). More traditional methods of defining a distribution skew using distribution moments^[52] were considered, but the above definition provides a simple and physically interpretable approach. Note that large values of Φ_s will generally indicate that the fiber angular distribution cannot be considered a single population, suggesting the existence of multiple layers of fibers with substantially different orientations.

The width of the $I(\Phi)$ distribution is proportional to the degree of fiber orientation: Highly oriented fiber networks result in a very narrow peak, while less well oriented fibers yield a broader peak. We have previously used the $(\cos^2 \Phi)$ to compute the degree of orientation normalized to a percentile scale.^[31] However, the $(\cos^2 \Phi)$ has no direct physical interpretation and is a nonlinear function of the distribution width. We have defined a physically intuitive

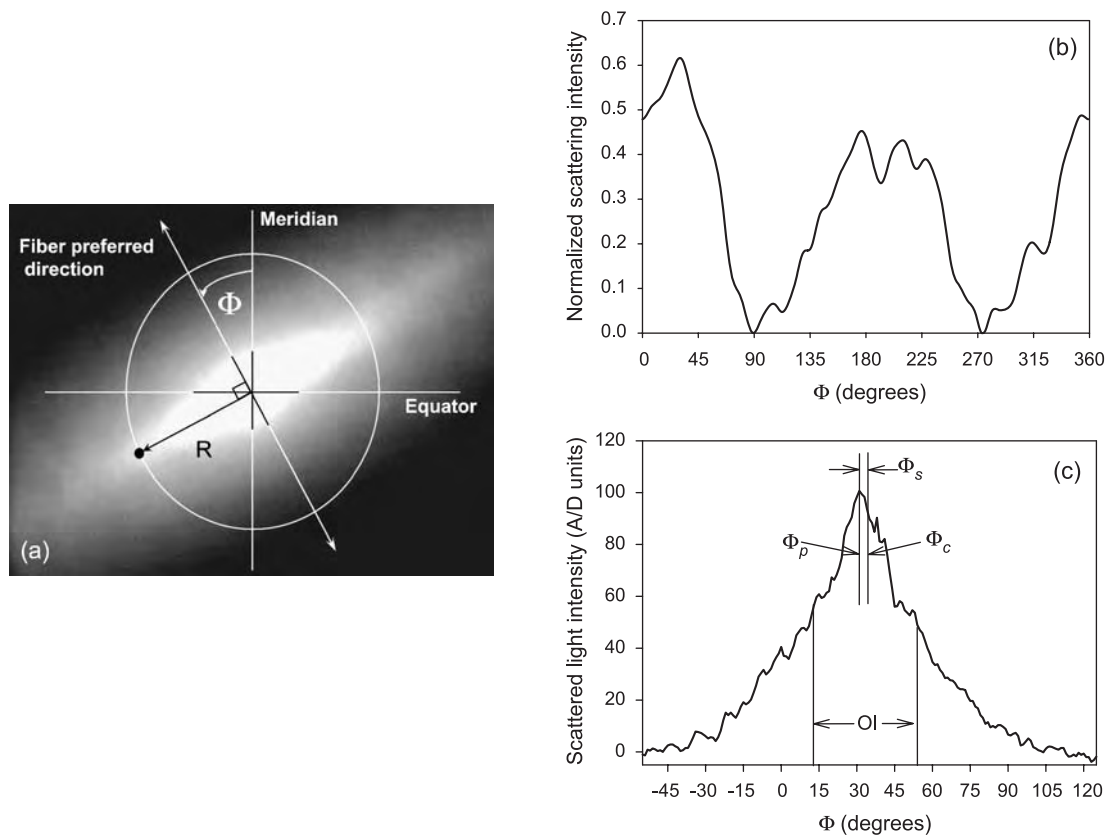


Fig. 2 (a) A digitized image of the scattered light intensity from a porcine aortic valve. The $I(\Phi)$ data is extracted from the SALS image along a circle of radius R centered on the optic axis (dot on circle), corresponding to a scattering angle $\theta = \tan^{-1}(R/L)$, where L is the distance from the specimen to the projection screen (Fig. 1). (b) The complete 360° extracted $I(\Phi)$ data showing the two intensity minima demarcating a 180° contiguous data segment. (c) locations of Φ_p , Φ_c , and the definition of the distribution skew, $\Phi_s = \Phi_p - \Phi_c$. Because light is scattered perpendicular to the fiber axes, all angular measurements are automatically shifted by 90° so that $I(\Phi)$ represents the true fiber orientation.

orientation index, OI, as the angle that contains one half of the total area under the $I(\Phi)$ distribution, representing 50% of the total number of fibers (Fig. 2c). Thus, highly oriented fiber networks will have low OI values, while more randomly oriented networks will have larger values. Note that the OI is computed directly from $I(\Phi)$ without the need of any distribution model.

CALIBRATION AND SENSITIVITY

Introduction

Evaluation of the SALS technique capabilities should include: 1) assessment of the angular and spatial resolutions, 2) sensitivity to the presence of multiple fiber populations, 3) effects of beam diameter, and 4) effect of tissue thickness. In the following section, methods and results to determine these parameters are presented.

Angular Position Accuracy and Sensitivity to Multiple Populations

To assess angular positional accuracy, two 10-mm square sections of bovine tendon (chosen for its uniformity in fiber structure, see the appendix in Ref. [31]), one at 100- μm and the other at 500- μm thickness, were mounted on a rotatable positioning device (0.1° accuracy) and tested at angular positions ranging from 0°–90°. SALS data was acquired over the entire specimen using a 1-mm spaced rectilinear grid and analyzed as discussed previously. Results indicated that the relationship between true angular position and mean Φ_c was consistently linear and demonstrated an accuracy of 1° (Fig. 3). This accuracy was observed for both the 100- μm and 500- μm thicknesses, indicating that SALS can maintain a 1° accuracy for a tissue thickness of 500 μm .

When two fiber populations exist, the resultant SALS data will contain two overlapping populations.^[28] To determine the sensitivity to multiple

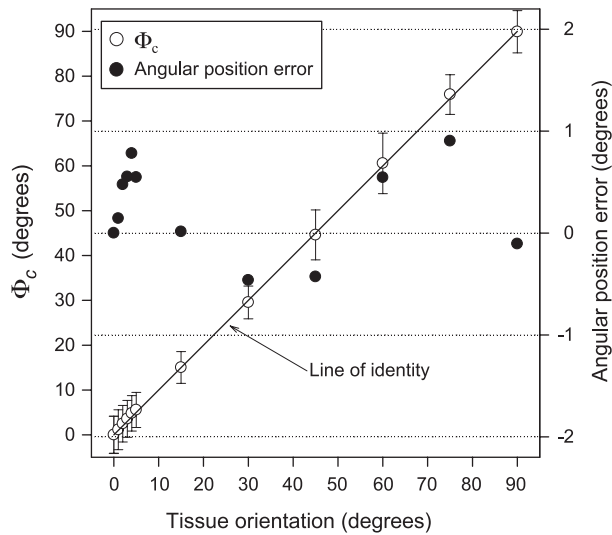


Fig. 3 Mean Φ_c and angular position error vs. specimen orientation Φ for a 500- μm bovine tendon section for $0^\circ \leq \Phi \leq 90^\circ$. For all orientations, there was close agreement between measured and actual (line of identity) Φ values, along with angular position errors all below 1° . These results indicate a 1° angular resolution in Φ_c .

populations, two 200- μm bovine tendon sections prepared as above were placed partially overlapping each other in the SALS device. One specimen was rotated with respect to the other for values of the difference in preferred directions, $\Delta\Phi_c$, ranging from 0° – 90° . $I(\Phi)$ was acquired for each $\Delta\Phi_c$ value from the overlapping area of the sections using a 254- μm spaced testing rectilinear grid. Since we assumed no optical interaction between fibers, the resultant scattering pattern of the overlapping sections should be the sum of the two individual layers. To evaluate this assumption, we summed the SALS data from the individual bovine tendon sections at the same $\Delta\Phi_c$ values used in the overlapped sections to generate synthetic multiple population SALS data. Both the Φ_s and OI were computed from the synthetic data sets and compared to the actual overlapping tissue SALS measurements.

When multiple populations occurred, the gradual separation of the fiber populations becomes clearly visible in the intensity distribution curves as $\Delta\Phi_c$ increases (Fig. 4). There was also close agreement between the experimental and synthetic curves, indicating that the $I(\Phi)$ response is the algebraic sum of

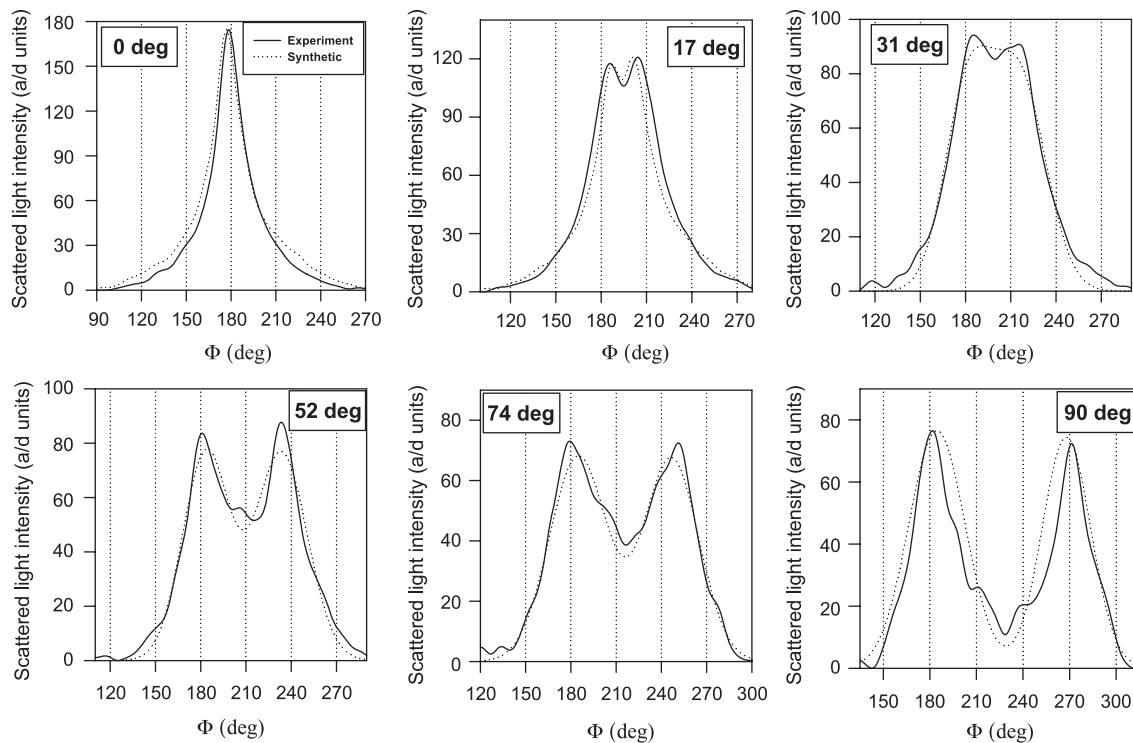


Fig. 4 $I(\Phi)$ vs. Φ data for two overlapping 200- μm bovine tendon sections rotated 0° , 17° , 31° , 52° , 74° , and 90° with respect to their preferred directions, showing both the experimentally measured and synthetically generated data. The second specimen's fiber orientation distribution can be seen to become increasingly distinct as the angle between the two specimens increases. Good agreement can be seen between the experimentally measured and synthetically generated data for all values of Φ .

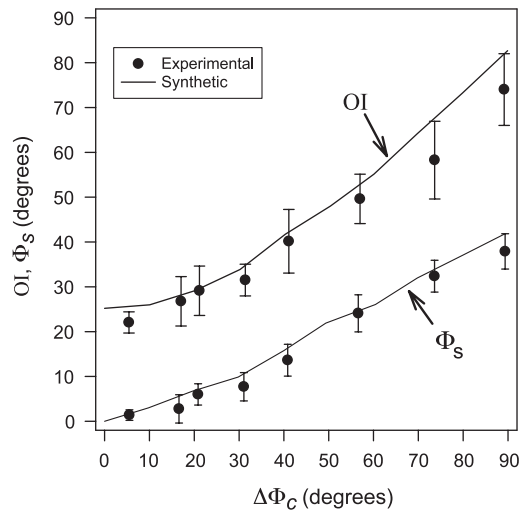


Fig. 5 Comparison of the measured and synthetic SALS analysis for OI and Φ_s using the SALS data of Fig. 4 showing good agreement between experimental and synthetic data. These results indicate that the SALS device can detect multiple populations to $\sim 5^\circ$ separation in their preferred directions.

the scattered light contributions from the individual bovine tendon sections (Fig. 4). In order to determine the sensitivity of the SALS parameters (Φ_c , Φ_s , and OI) to multiple populations, the parameters were plotted against the angle between the populations $\Delta\Phi_c$ (Fig. 5). Both Φ_s and OI predictably increased at larger $\Delta\Phi_c$ values and showed close agreement between the synthetic and experimental data over the complete range of $\Delta\Phi_c$ (Fig. 5). The values for Φ_c computed from the synthetic data were within $\pm 3^\circ$ of the actual value also over the complete range of $\Delta\Phi_c$. These results indicate that the SALS device can accurately detect the presence of multiple populations when there is at least a $\sim 5^\circ$ separation in their preferred directions.

Effects of Tissue Thickness

As tissue thickness increases, multiple scattering effects caused by light rescattered by succeeding collagen fiber layers also increase, causing potential loss of structural information.^[51,53] It was previously determined using bovine tendon, a dense collagenous tissue, that thicknesses up to 400 μm do not detectably effect SALS measurements.^[31] In order to assess the effects of tissue thickness more fully, SALS data was acquired from a 19-mm \times 19-mm \times 0.7-mm thick specimen of bovine pericardium using a 2.5-mm increment grid. Next, the specimen was frozen and sectioned at 25- μm increments using a freezing microtome, with sections mounted on a glass slide. SALS data was acquired for each individual section using the same scanning grid, with care taken to ensure consistent registration of the SALS test locations for all sections. The SALS data from intact

tissue and the individual sections were then compared to determine differences in the SALS data between the individual sections and the intact, full-thickness tissue.

Results indicated that the overall fiber orientation of the bovine pericardial section changed gradually from the outer (epipericardial) to the inner (visceral) surfaces (Figs. 6a,b,c). This can be seen to occur at single test locations where the $I(\Phi)$ intensity distribution can be seen to change with thickness (Figs. 6a,b,c). Since no optical interactions between the light scattered from each fiber are assumed, the total light scattering pattern should be the algebraic sum of the contributions from each individual fiber layer. To verify this assumption, the SALS data from two locations near the center of each 25- μm section was averaged and compared with the SALS data for sections obtained before sectioning at the same location (Fig. 6d). The close agreement between the averaged and intact SALS data indicates that the SALS pattern is the algebraic sum of the contribution of each layer in the tissue and that multiple scattering effects due to tissue thickness were negligible.

A final comment on tissue thickness effects is in order. As previously mentioned, as tissue thickness increases, multiple scattering effects will tend to predominate. One of the effects of multiple scattering is the generation of a diffuse light intensity component, which produces the equivalent of a DC component in the $I(\theta)$ distribution.^[31] For highly aligned tissues (e.g., tendon), this problem can be easily circumvented by subtracting the minimum scattered light intensity from the $I(\theta)$ distribution.^[31] However, for tissues with more broad or complex fiber orientation distributions, this method could potentially remove both the diffuse light contribution and some fiber orientation information. In our extensive experience utilizing SALS for soft connective tissues, we have observed that these effects can be checked for by performing multiple-section studies as described above. Moreover, specific adjustments to the optical hardware (especially the camera light sensitivity and shutter speed) can be made using tissues with highly aligned fibers (as done using tendon in Refs. [31] and [48]) to be sure that diffuse light component is removed for the scattered light contribution without affecting fiber orientation measurements. As with any experimental measurement, it is suggested that verification studies as described above be performed for each tissue type to ensure accuracy and correct interpretation of the SALS data.

Effect of Beam Diameter on SALS Spatial Resolution

The effect of beam diameter on SALS spatial resolution was also determined using the same two bovine

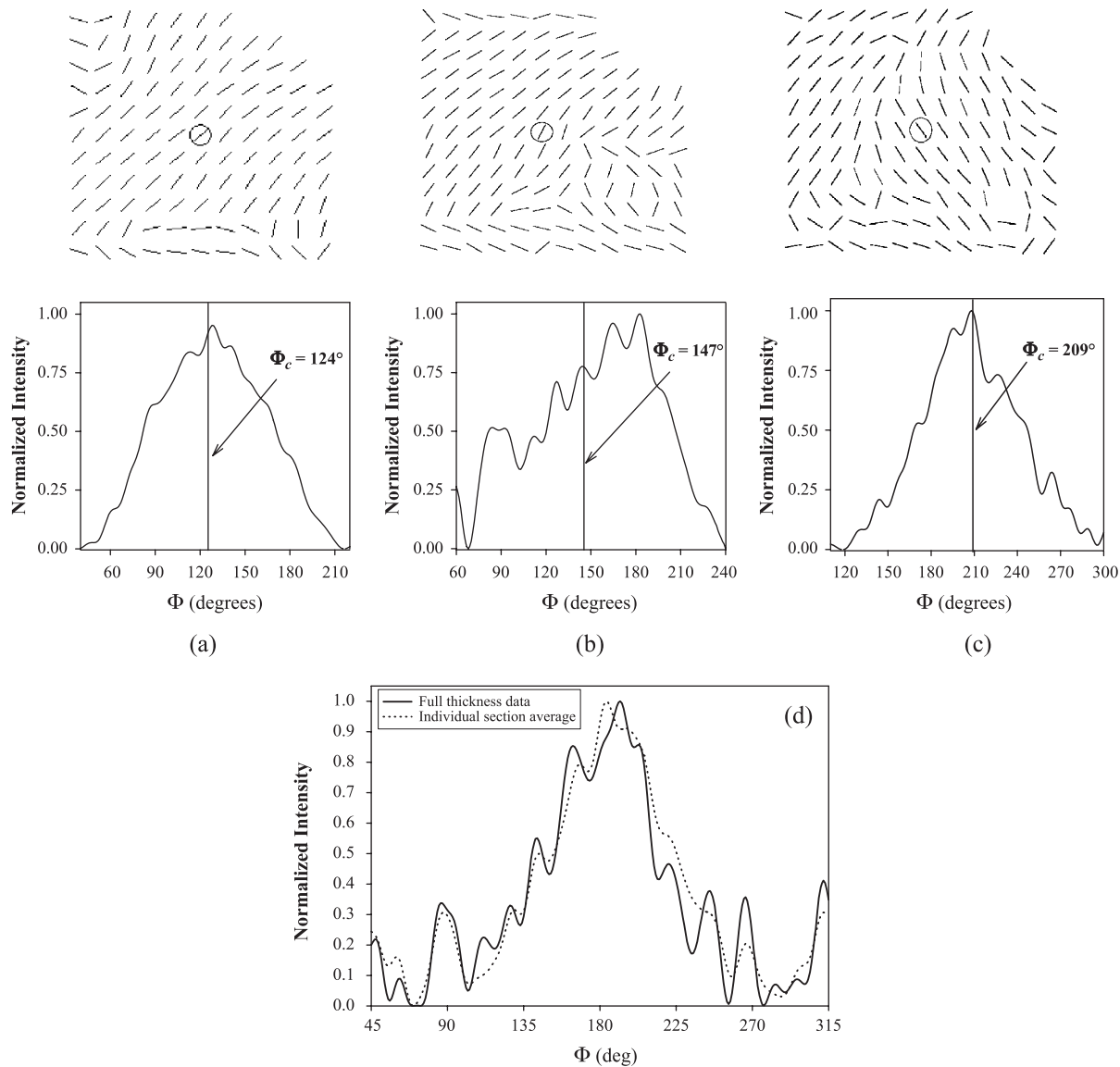


Fig. 6 Fiber preferred direction vector plots of sequential 25- μm sections taken from a single bovine pericardium specimen at depths of (a) epipericardial side, (b) 225 μm , and (c) 375 μm (visceral side). $I(\Phi)$ data from each location indicated by the circle is shown in the plots below each vector plot. (d) Averaged normalized $I(\Phi)$ data from each 25- μm bovine pericardium section shown along with the data from the intact specimen for the location indicated by the circle. The close agreement between the averaged and intact SALS data indicates that the SALS pattern is the algebraic sum of the contribution of each layer in the tissue.

pericardial sections, which were scanned using a 254- μm rectilinear grid and 100- μm , 350- μm , and 500- μm beam diameters. As beam diameter is decreased, speckle (i.e., self-interference) becomes more prominent, which appears in the SALS data as random noise that will reduce accuracy (Fig. 7). As the beam diameter increases, speckle diminishes and the scattering pattern becomes more regular. At a $1/e^2$ beam diameter of 600 μm , the least speckle was observed. In our experience with dense collagenous tissues, the best data is obtained with a beam diameter of at least 250 μm .

However, increasing beam diameter will reduce spatial resolution by averaging SALS results over larger tissue areas. In our 1997 paper, it was determined that the SALS device can distinguish tissue structures to an effective spatial resolution of $\pm 254 \mu\text{m}$ with a 600- μm $1/e^2$ beam diameter.^[48] Subsequent improvements to the spatial filter/beam expander have indicated that when using a beam diameter of at least 250 μm , an equivalent spatial resolution of 250 μm is obtained. Overall, a general rule of thumb is that distance between scans should not be less than the $1/e^2$ beam diameter.

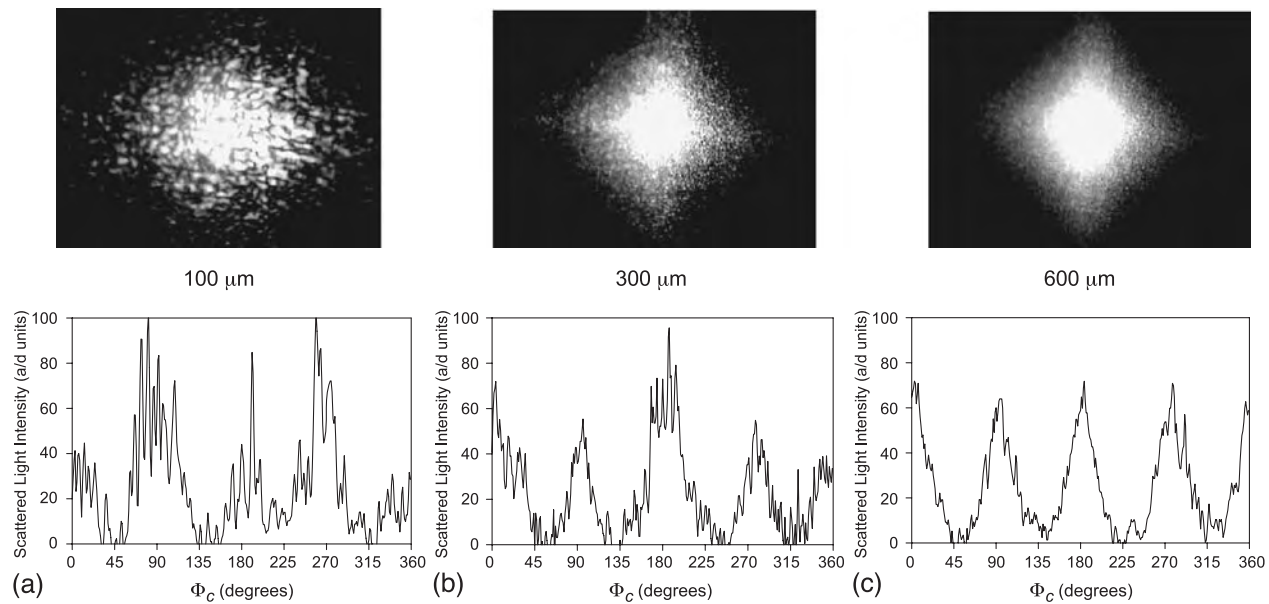


Fig. 7 Effects of laser beam diameter on the $I(\Phi)$ vs. Φ data for (a) 100 μm , (b) 350 μm , and (c) 600 μm $1/e^2$ beam diameters using two overlapping 200- μm bovine tendon sections rotated 90° with respect their preferred directions. For the 100- μm beam diameter, speckle interference noticeably degraded the $I(\Phi)$ data.

Effects of Fiber Diameter and Interfiber Spacings

As stated above, it is assumed that the fiber diameter distribution is *independent* of fiber orientation Φ . As long as this assumption is true, the angular distribution of scattered light intensity accurately represents the angular distribution of fiber orientations. Based on Fraunhofer diffraction theory (Eq. 1), larger fibers tend to scatter light closer to the beam axis and smaller fibers farther out. Thus if the fiber diameter distribution was a function of Φ , then the $I(\Phi)$ data would tend to change with θ . We examined the $I(\Phi)$ data, normalized to account for the decrease in absolute scattered light intensity for $2^\circ \leq \theta \leq 7^\circ$, to validate this assumption for bovine pericardium and porcine aortic valve leaflets. Both tissue types demonstrated no detectable change in the $I(\Phi)$ intensity distribution for all values of θ , indicating that $I(\Phi)$ is independent of fiber diameter and interfiber spacing for these tissues, *so that the fiber diameter distribution does not affect their fiber orientation measurements*. This result is consistent with the fiber morphology of these tissues, which does not indicate that the collagen bundle diameters are a function of their orientation.^[31] We have observed this same behavior in all tissues we have encountered.

Effects of Collagen Crimp

Application of SALS in determining the fiber architecture in collagenous tissues is complicated by the

presence of fiber crimp. To help clarify exactly how crimp can affect the SALS data, the following graphical explanation is presented. Consider a perfectly straight fiber with no crimping. The resulting $R(\theta)$ will be a nearly perfect intensity spike (Fig. 8a). For a crimped fiber, each segment of the fiber will scatter light according to its local orientation with respect to the overall fiber axis (Fig. 8b). Similar to a

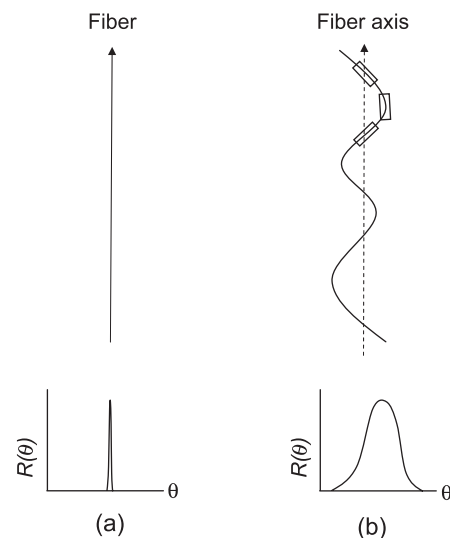


Fig. 8 (a) Representation of a crimped collagen fiber along with its fiber axis and corresponding $R(\theta)$. (b) Same fiber shown in its stretched state with its corresponding $R(\theta)$. The small boxes outline the conceptual linear segments that make up the fiber and add to the total $R(\theta)$.

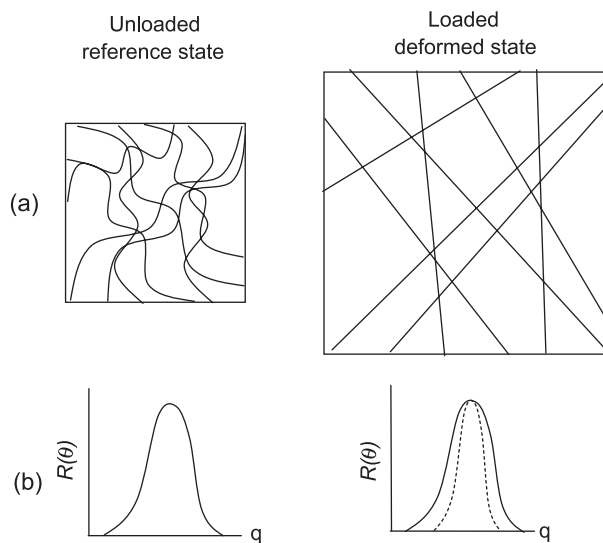


Fig. 9 (a) A schematic showing the fiber network in the reference (left) and equibiaxially deformed (right) states, where there is no fiber rotation. The corresponding $R(\theta)$ for both states is shown in (b). In the situation shown here, the uncrimping of the fibers causes $R(\theta)$ to narrow [dotted distribution shown in rightmost $R(\theta)$ plot in (b)]. When uncrimping does not affect the SALS data, then no changes in $R(\theta)$ will occur.

single fiber, the effect of crimp on a fiber network is to broaden $R(\theta)$ (Figs. 9 and 10). However, when a tissue is subjected to equibiaxial strain (i.e., same strain in both directions), the constituent fibers will stretch but not rotate. Thus, equibiaxial strain offers a straightforward method to assess the effects of fiber crimp on the accuracy of $R(\theta)$. This is shown graphically in Fig. 10a, where fibers in the unloaded reference state are crimped, and then subjected to equibiaxial strain to the point where the crimp has been removed in the deformed state. Here, if crimp affected the SALS data, its effect would be to broaden $R(\theta)$.

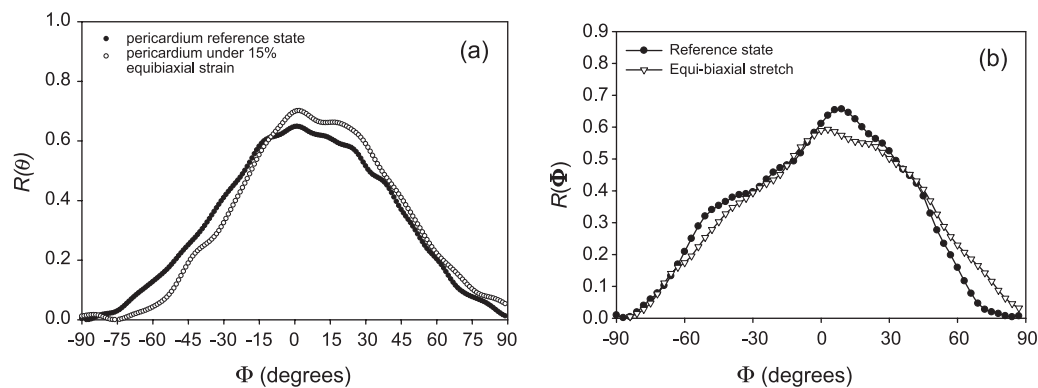


Fig. 10 Changes in $R(\theta)$ under equibiaxial strain for (a) bovine pericardium and (b) small intestinal submucosa. For both tissues, no detectable changes in $R(\theta)$ were observed, indicating that collagen fiber crimp did not affect the accuracy of the SALS measurements.

Interestingly, we have observed no change in $R(\theta)$ in either porcine intestinal submucosa^[38] or bovine pericardium^[46] (Fig. 10). This result is most likely due to the fact that the collagen crimp amplitude is small compared to the overall fiber architecture and tissue dimensions. That is, given the very large number of fibers through a typical thickness of $\sim 400\ \mu\text{m}$ (resulting in ~ 800 – 1000 fibers through thickness) and that the variations in fiber orientations are large with respect to dispersion due to crimp, crimp effects tend to be averaged out.

Maps of Collagen Fiber Architecture

One of the main advantages of SALS is the ability to produce fiber architectural maps rapidly to exact anatomical scale. An example of this is presented for the aortic valve in Fig. 11. Here, fiber-preferred directions can be seen to course smoothly along the circumferential direction in both valves, although the preferred directions are more ordered in the valve leaflet fixed at 4-mmHg transvalvular pressure (Fig. 11c,d). However, there was a dramatic change in OI with only a 4-mmHg pressure change, demonstrating substantial mobility of the fiber network in the PAV at low transvalvular pressures. The nodulii of Aranti can be seen as a more randomly oriented region, and the free edge is clearly delineated in both leaflets as a band of highly oriented fibers along the upper leaflet boundary.

APPLICATION TO ENGINEERED TISSUES

In addition to the analysis of native tissue, SALS can also be useful in the development of engineered tissues. An example of this is a study performed in our lab on DermagraftTM (Advanced Tissue Sciences, La Jolla, CA),

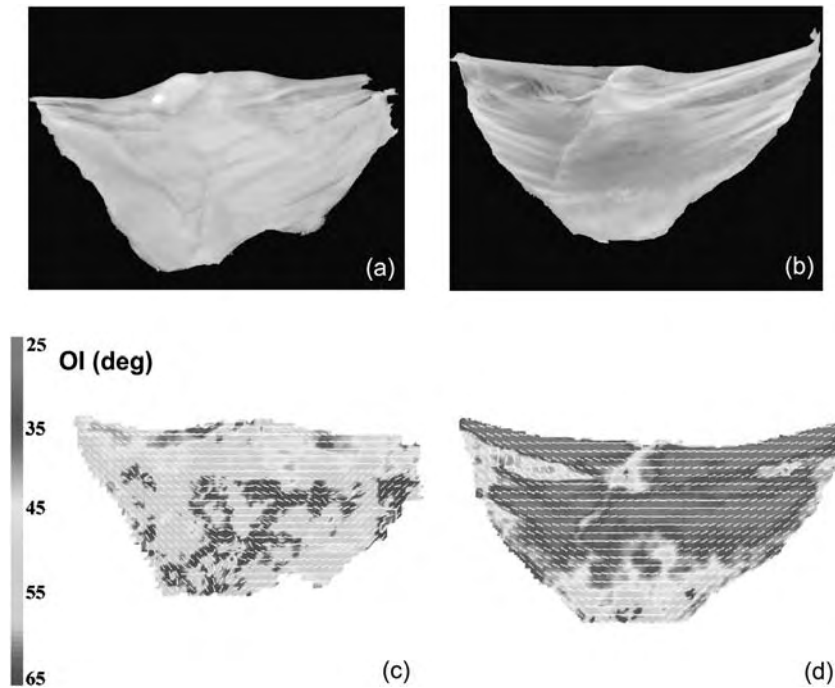


Fig. 11 Two PAV leaflets fixed at (a) 0 mmHg and (b) 4 mmHg. Each leaflet was SALS tested using a 254- μm spaced rectilinear grid. The corresponding Φ_c data are shown for each valve as vector plots superposed over color fringe of the OI for (c) 0 mmHg and (d) 4 mmHg, with vector data shown for every third test location for clarity. Note the dramatic change in OI with only a 4-mmHg pressure change, demonstrating substantial mobility of the fiber network at low transvalvular pressures.

which consists of human neonatal dermal fibroblasts grown on a woven mesh of biodegradable polyglycolic acid/polylactic acid copolymer. As the fibroblasts proliferate throughout the mesh scaffolding, they secrete collagen and other human dermal matrix proteins to produce a living dermal tissue.^[36]

Experimental details have been previously presented.^[36] Briefly, SALS studies were performed on five 102-mm \times 152-mm Dermagraft specimens taken from the same cell line and production run, stored at -80°C , and then thawed. Light scattering by the collagen fiber specimens was enhanced using Picro-sirius red collagen stain. Unlike native tissues, Dermagraft SALS signal contains contributions from both the collagen and scaffold fibers. The scaffold used to fabricate Dermagraft is a structurally regular woven fabric (Fig. 12, left), which produces a unique SALS pattern with four peaks corresponding to each of the four sides (Fig. 12, right).

Assuming the fiber orientation contributions from each of the four mesh sides and collagen fibers to be Gaussian distributed, and that the principle of optical superposition applies, the total scattered light intensity distribution $I(\Phi)$ can be represented as

$$I(\Phi) = c \sum_{i=1}^4 \left\{ I_p(i) \exp \left(- \left(\frac{\Phi - \Phi_p(i)}{\sigma(i)} \right)^2 \right) \right\} + I_p(5) \exp \left(- \left(\frac{\Phi - \Phi_p(5)}{\sigma(5)} \right)^2 \right) \quad (4)$$

where for the i th Gaussian population $I_p(i)$ is the peak intensity, $\Phi_p(i)$ is the direction of preferred orientation, $\sigma(i)$ the Gaussian population variance, and c a constant required to compensate for the lower light scattering power (i.e., the ratio of incident light over the scattered light) of the mesh fibers. The second term on the right-hand side of Eq. 4 is the fifth Gaussian population representing the collagen component of the SALS signal, from which the local collagen fiber orientations can be obtained.

To obtain the Gaussian fit parameters for the untreated mesh, SALS data from 50 randomly chosen locations from 10 untreated mesh specimens were averaged together. At each angular position Φ , the resulting mean $I(\Phi)$ data was fit to Eq. 4, using the Levenberg–Marquardt algorithm,^[54] with the collagen parameters temporarily removed from the fit equation and c forced to unity. This allowed us to model accurately ($r^2=0.98$) the untreated mesh SALS signal (Fig. 11b), and the fitted parameters were then used in the complete form of Eq. 4.

Equation 4 fitted the Dermagraft SALS signal well ($r^2>0.95$), revealing the collagen population to be the main contributor to the SALS signal (Fig. 11b). The fitted mesh opening side angles (Φ_p) had similar mean values with small standard errors between specimens, as well as comparable mesh vertex angles $\Delta\Phi$ computed by subtracting the two side angles at each vertex. Thus, mesh distortions during specimen processing were small. Results from Eq. 4 showed that the collagen $\Phi_p \approx 90^\circ$ (Fig. 12, left), indicating that the Dermagraft collagen fibers have a preferred direction

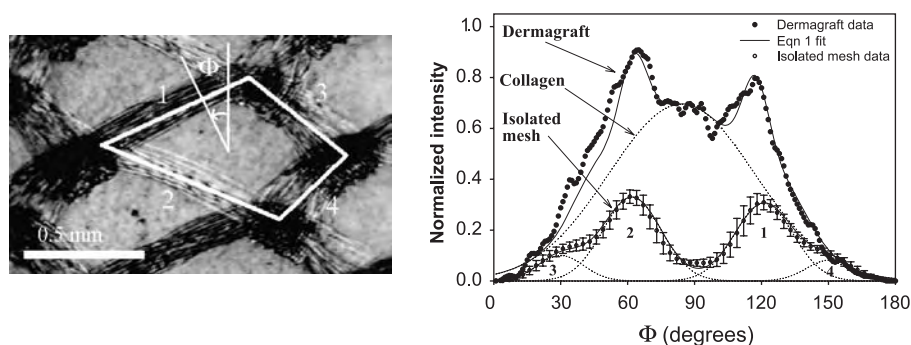


Fig. 12 (left) Photomicrograph of Dermagraft showing a kite-shaped idealization of the mesh openings, (right) an example of Eq. 4 to fit the SALS data for Dermagraft. Intensity values have been normalized to range from 0 to 1 to demonstrate more clearly that collagen was the largest contributor to the Dermagraft SALS signal. The numbers within each untreated mesh Gaussian population indicate the corresponding mesh opening side of the left image.

nearly parallel to the long dimension of the kite-shaped scaffold opening. This study demonstrated the utility of SALS as a rapid and inexpensive technique for the evaluation of gross collagen fiber architecture in engineered tissues.

SUMMARY

This chapter demonstrated that the SALS technique lends itself to rapid analysis of large areas of tissue without the need for expensive equipment and time-consuming tissue preparations. The present device provides high resolution information on local fiber orientation to an angular resolution of $\sim 1^\circ$ and a minimum recommend spatial resolution equal to the minimum $1/e^2$ beam diameter of $250\ \mu\text{m}$. Another important aspect of SALS demonstrated in this study is that the distribution of scattered intensity represents the algebraic sum of the contributions from each tissue layer. Thus, SALS measurements represent the total structure of the tissue throughout its thickness, with each layer contributing in proportion of its thickness to the total thickness. Overall, the SALS technique lends itself to the study of a wide range of soft tissue structures in health and disease.

ACKNOWLEDGMENTS

The generous support of the A.W. Ford Foundation for funding the construction of the SALS instrumentation, the American Heart Association (National, Florida, and Pennsylvania-Delaware affiliates), and the National Institutes of Health/NHLBI is gratefully acknowledged. The author would like to thank Dr. Paul L. Kronick of the USDA-ERRC (Philadelphia, PA) for introducing him to the use of SALS as a method for connective tissue structural analysis. The

author is also indebted to Drs. David B. Smith and Kristen L. Billiar, as well as Mr. Jonathan Grashow for their invaluable assistance in the SALS device design and fabrication and the development of sophisticated control and analysis software. MSS is an Established Investigator of the American Heart Association.

ARTICLES OF FURTHER INTEREST

Collagen; Elastin; Extracellular Matrix Scaffolds; Tissue Engineering Scaffolds

REFERENCES

1. Fung, Y.C. *Biomechanics: Mechanical Properties of Living Tissues*; Springer Verlag: New York, 1993.
2. Hukins, D.W.L. Collagen Orientation. In *Connective Tissue Matrix*; Hukins, D.W.L., Ed.; Verlag Chemie, 1984; 211–240.
3. Whittaker, P.; Canham, P.B. Demonstration of quantitative fabric analysis of tendon collagen using two-dimensional polarized light microscopy. *Matrix* **1991**, *11*, 56–62.
4. Whittaker, P.; Kloner, R.A.; Boughner, D.R.; Pickering, J.G. Quantitative assessment of myocardial collagen with picrosirius red staining and circularly polarized light. *Basic Res. Cardiol.* **1994**, *89*, 397–410.
5. Chaudhuri, S.; Nguyen, H.; Rangayyan, R.M.; Walsh, S.; Frank, C.B. A fourier domain directional filtering method for analysis of collagen alignment in ligaments. *IEEE Trans. Biomed. Eng. BME* **1987**, (34), 509–518.
6. Frank, C.; MacFarlane, B.; Edwards, P.; Rangayyan, R.; Liu, Z.Q.; Walsh, S.; Bray, R. A quantitative analysis of matrix alignment in ligament scars: A comparison of movement versus immobilization in an immature rabbit model. *J. Orthop. Res.* **1991**, *9*, 219–227.

7. Liu, Z.Q.; Rangayyan, R.M.; Frank, C.B. Statistical analysis of collagen alignment in ligaments by scale-space analysis. *IEEE Trans. Biomed. Eng.* **1991**, *38*, 580–587.
8. Kastelic, J.; Baer, E. Deformation of Tendon Collagen. In *The Mechanical Properties of Biological Materials*; Vincent, J.F., Currey, J.D., Eds.; Society for Experimental Biology: Great Britain, 1980; 397–433.
9. Purslow, P.P.; Bigi, A.; Ripamonti, A.; Roveri, N. Collagen fibre reorientation around a crack in biaxially stretched materials. *Int. J. Macromol.* **1984**, *6*, 21–25.
10. Sasaki, N.; Odajima, S. Stress-strain curve and young's modulus of a collagen molecule as determined by the x-ray diffraction technique. *J. Biomech.* **1996**, *29*, 655–658.
11. Baer, E.; Cassidy, J.J.; Hiltner, A. Hierarchical Structure of Collagen and Its Relationship to the Physical Properties of Tendon. In *Collagen*; Nimni, M.E., Ed.; CRC Press: Boca Raton, 1988; 177–199.
12. Hilbert, S.L.; Ferrans, V.J.; Swanson, W.M. Optical methods for the nondestructive evaluation of collagen morphology in bioprosthetic heart valves. *J. Biomed. Mater. Res.* **1986**, *20*, 1411–1421.
13. Tower, T.T.; Tranquillo, R.T. Alignment maps of tissues: I. Microscopic elliptical polarimetry. *Biophys. J.* **2001**, *81* (5), 2954–2963.
14. Tower, T.T.; Tranquillo, R.T. Alignment maps of tissues: II. Fast harmonic analysis for imaging. *Biophys. J.* **2001**, *81* (5), 2916–2971.
15. Borch, J.; Sundarajan, P.R.; Marchessault, R.H. Light scattering by cellulose. Iii-Morphology of wood. *J. Polym. Sci.* **1971**, *9*, 313–329.
16. Muggli, R.; Marton, R. Light scattering by cellulose. V.-Anisotropy scattering by wood fibers. *J. Polym. Sci.* **1971**, *36*, 121–139.
17. Raman, C.V.; Bhat, M.R. The structure and optical behavior of some natural and synthetic fibers. *Proc. Acad. Sci.* **1954**, *440*, 109–116.
18. Stein, R.S.; Erhardt, P.; van Aartsen, J.J.; Clough, S. Theory of light scattering from oriented and fiber structures. *J. Polym. Sci.* **1966**, *13*, 1–35.
19. Stein, R.S.; Wilson, P.R. Scattering of light by polymer films possessing correlated orientation fluctuation. *J. Appl. Physiol.* **1962**, *33*, 1914–1922.
20. Chien, J.C.W.; Chang, E.P. Small-angle light scattering of reconstituted collagen. *Macromolecules* **1972**, *5*, 610–617.
21. Moritani, M.; Hayashi, N.; Utsuo, A.; Kawai, H. Light-scattering patterns from collagen films in relation to the texture of a random assembly of anisotropic rods in three dimensions. *Polym. J.* **1971**, *2*, 74–87.
22. Kronick, P.L.; Buechler, P.R. Fiber orientation in calfskin by laser light scattering or x-ray diffraction and quantitative relation to mechanical properties. *J. Am. Leather Chem. Assoc.* **1986**, *81*, 221–229.
23. Kronick, P.L.; Sacks, M.S.; Dahms, M. Vertical fiber defect quantified by small angle light scattering. *Connect. Tissue Res.* **1991**, *27*, 1–13.
24. Ferdman, A.G.; Yannas, I.V. Scattering of light from histologic sections: A new method for the analysis of connective tissue. *J. Invest. Dermatol.* **1993**, *100*, 710–716.
25. McCally, R.L.; Farrell, R.A. Small-angle light scattering and birefringence properties of chick cornea. *J. Refract. Surg.* **1999**, *15* (6), 706–710.
26. McCally, R.L.; Farrell, R.A. Structural implications of small-angle light scattering from cornea. *Exp. Eye Res.* **1982**, *34* (1), 99–113.
27. Andreo, R.H.; Farrell, R.A. Corneal small-angle light-scattering theory: Wavy fibril models. *J. Opt. Soc. Am.* **1982**, *72* (11), 1479–1492.
28. Bettelheim, F.A.; Kumbar, M. An interpretation of small-angle light-scattering patterns of human cornea. *Investig. Ophthalmol. Vis. Sci.* **1977**, *16* (3), 233–236.
29. Bettelheim, F.A.; Magrill, R. Small-angle light-scattering patterns of corneas of different species. *Investig. Ophthalmol. Vis. Sci.* **1977**, *16* (3), 236–240.
30. Bettelheim, F.A.; Kaplan, D. Small angle light scattering of bovine cornea as affected by birefringence. *Biochim. Biophys. Acta* **1973**, *313* (2), 268–276.
31. Sacks, M.S.; Chuong, C.J. Characterization of collagen fiber architecture in the canine central tendon. *J. Biomech. Eng.* **1992**, *114*, 183–190.
32. Sacks, M.S.; Chuong, C.J.; More, R. Collagen fiber architecture of bovine pericardium. *ASAIO* **1994**, *40*, M632–M637.
33. Hiester, E.D.; Sacks, M.S. Optimal bovine pericardial tissue selection sites. Ii. Cartographic analysis. *J. Biomed. Mater. Res.* **1998**, *39* (2), 215–221.
34. Hiester, E.D.; Sacks, M.S. Optimal bovine pericardial tissue selection sites. I. Fiber architecture and tissue thickness measurements. *J. Biomed. Mater. Res.* **1998**, *39* (2), 207–214.
35. Hamann, M.C.; Sacks, M.S.; Malinin, T.I. Quantification of the collagen fibre architecture of human cranial dura mater. *J. Anat.* **1998**, *192* (Pt. 1), 99–106.
36. Sacks, M.S.; Chuong, C.J.; Petroll, W.M.; Kwan, M.; Halberstadt, C. Collagen fiber architecture of a cultured dermal tissue. *J. Biomech. Eng.* **1997**, *119* (1), 124–127.
37. Sacks, M.S.; Gloeckner, D.C. Quantification of the fiber architecture and biaxial mechanical behavior of porcine intestinal submucosa. *J. Biomed. Mater. Res.* **1999**, *46*, 1–10.
38. Gloeckner, D.C.; Sacks, M.S.; Billiar, K.L.; Bachrach, N. Mechanical evaluation and design of a multilayered collagenous repair biomaterial. *J. Biomed. Mater. Res.* **2000**, *52* (2), 365–373.
39. Sacks, M.S.; Smith, D.B.; Hiester, E.D. The aortic valve microstructure: Effects of transvalvular pressure. *J. Biomed. Mater. Res.* **1998**, *41* (1), 131–141.
40. Sacks, M.S.; Smith, D.B. Effects of accelerated testing on porcine bioprosthetic heart valve fiber architecture. *Biomaterials* **1998**, *19* (11–12), 1027–1036.
41. Sacks, M.S.; Schoen, F.J. Collagen fiber disruption occurs independent of calcification in clinically explanted bioprosthetic heart valves. *J. Biomed. Mater. Res.* **2002**, *62* (3), 359–371.
42. Kunzelman, K.S.; Sacks, M.S.; Cochran, R.P.; Eberhart, R.C. Mitral valve leaflet collagen distribution by laser analysis. *Seventh South. Biomed. Eng. Conf.* **1988**, 82–85.

43. Bowes, L.E.; Jimenez, M.C.; Hiester, E.D.; Sacks, M.S.; Brahmawari, J.; Mertz, P.; Eaglstein, W.H. Collagen fiber orientation as quantified by small angle light scattering in wounds treated with transforming growth factor-beta2 and its neutralizing antibody. *Wound Repair Regen.* **1999**, *7* (3), 179–186.
44. Chuong, C.J.; Sacks, M.S.; Johnson, R.L.; Reynolds, R.C. On the anisotropy of the diaphragmatic central tendon. *J. Biomech.* **1991**, *24*, 563–576.
45. Billiar, K.L.; Sacks, M.S. A method to quantify the fiber kinematics of planar tissues under biaxial stretch. *J. Biomech.* **1997**, *30* (7), 753–756.
46. Sacks, M.S. Incorporation of experimentally-derived fiber orientation into a structural constitutive model for planar collagenous tissues. *J. Biomech. Eng.* **2003**, *125* (2), 280–287.
47. Sacks, M.S.; Chuong, C.J. Orthotropic mechanical properties of chemically treated bovine pericardium. *Ann. Biomed. Eng.* **1998**, *26* (5), 892–902.
48. Sacks, M.S.; Smith, D.B.; Hiester, E.D. A small angle light scattering device for planar connective tissue microstructural analysis. *Ann. Biomed. Eng.* **1997**, *25* (4), 678–689.
49. Marshall, G.E. Gaussian Laser Beam Diameters and Divergence. In *Optical Scanning*; Marshall, G.E., Ed.; Marcel Dekker: New York, 1991; 1–11.
50. Halliday, D.; Resnick, R. *Physics*; John Wiley and Sons: New York, 1960.
51. Guinier, A.; Fournet, G. *Small Angle Scattering of X-Rays*; Wiley: New York, 1955.
52. Mendenhall, W.; Sincich, T. *Statistics for the Engineering and Computer Sciences*; Dellen: San Francisco, 1988.
53. Cowley, J.M. Principles of Image Formation. In *Introduction to Analytical Electron Microscopy*; Hren, J.J., Goldstein, J.I., Joy, D.C., Eds.; Plenum Press: New York, 1979; 1–42.
54. Press, W.H.; Flannery, B.P.; Teukolsky, S.A.; Vetterling, W.T. *Numerical Recipes in C*; Cambridge University Press: Cambridge, 1988.



Spine Biomechanics

Clayton Adam

Pediatric Spine Research Group, School of Engineering Systems, Queensland University of Technology, Brisbane, Queensland, Australia

INTRODUCTION

Spine biomechanics is the study of stability, motion, and deformation in the human spine, generated by both internal and external forces acting on and through spinal connective tissues. The primary aim in studying spine biomechanics is to understand the underlying mechanical aspects of various spinal disorders and their clinical treatments, so that adverse outcomes can be avoided and optimal prevention or treatment strategies can be developed.

Theoretical tools used in the study of spine biomechanics include the equations of Newtonian mechanics and mathematical principles from continuum mechanics, porous media mechanics, fracture mechanics, and numerical analysis. Computational mechanics techniques, in particular the finite element method, have been widely applied to the analysis of spinal problems owing to their ability to solve previously intractable governing differential equations for complex three-dimensional shapes, time-varying forces, and nonlinear materials. The validation and application of theories have been accompanied by *in vitro* tests of both animal and human spines to determine stiffness, strength, strain, and failure loads of spinal segments and surgical implants. To a lesser extent, *in vivo* instrumentation has been used to measure strains and pressures in the spine during daily activities. Finally, medical imaging technologies such as biplanar radiography, computed tomography, and magnetic resonance imaging have allowed three-dimensional imaging of the spine for accurate assessment of both anatomy and kinematics.

For the purposes of this article, three categories of spinal disorder have been chosen: 1) intervertebral disc (IVD) degeneration; 2) spinal fractures; and 3) spinal deformity. To better understand the mechanics of the pathological spine, it is also important to study the behavior of the healthy spine; therefore the kinematics and kinetics of the healthy spine are discussed first, followed by a summary of tissue mechanical properties (stiffness and strength). The subsequent sections describe biomechanical aspects of pathology and treatment for each category mentioned above. Finally, a brief overview of likely future developments in the field is given. Owing to the breadth of the field, it is not

possible to mention every avenue of biomechanical investigation being undertaken in the spine, or even all the major ones. Notable areas not discussed in this article include biomechanics of the ribcage and sternum (which have a significant stiffening effect on the thoracic spine), biofluid mechanics of cerebrospinal fluid, spinal cord rheology, and tissue mechanics at the cellular level. The article assumes a basic knowledge of spinal anatomy and some understanding of the mechanical concepts of force, moment, stiffness, strength, pressure, stress, and strain.

THE HEALTHY SPINE

Kinematics

Spinal kinematics is the study of spinal movement, without reference to the forces producing the motion. The healthy human spine provides a central, flexible support structure to the rest of the body, and the complex spinal motions occurring during everyday activity can be considered in terms of primary motions in each of three orthogonal planes: midsagittal, coronal, and axial. In the midsagittal plane, spinal flexibility is asymmetric about the neutral standing position, with substantial forward flexion possible, but limited backward extension owing to the semirigid constraint of the zygapophyseal (facet) joints. Ranges of motion for both lateral bending and axial rotation are approximately the same on left and right sides.

When considering spinal motions, the vertebrae are rigid to a first approximation so that overall movement can be described in terms of the response of individual “motion segments” (also referred to as a functional spinal unit), each comprising two vertebrae connected by an IVD and the posterior complex of zygapophyseal joints, ligaments, and thoracolumbar fascia. For a single motion segment, movements in a particular plane can be described in terms of rigid-body translations and rotations of one vertebra relative to the other. The rotation component occurs about an instantaneous axis of rotation (IAR). The IAR locations have been extensively studied because shifts in the IAR can indicate abnormal kinematics.^[1] For example, variations in IAR position during spinal movement

have been used as indicators of spondylolysis and have been shown to correlate with spinal pain.^[2]

Fig. 1 shows IAR locations for the cervical, thoracic, and lumbar spines.^[1] As suggested by its name, the IAR is not stationary, but translates during motion with the locus of IAR points traced by a particular overall movement referred to as a centrode. The polycentric behavior of spinal motion segments has important implications for the design of total disc

replacement (TDR) devices. Approximate ranges of motion per segment in combined flexion/extension are 10–20° in the cervical spine, 4–6° in the upper thoracic spine, 9–12° in the lower thoracic spine, and 12–17° in the lumbar spine. For lateral bending to one side only, ranges of motion are 4–11° in the cervical spine, 5–9° in the thoracic spine, and 3–8° in the lumbar spine. For axial rotation to one side only, ranges of motion are 2–7° in the cervical spine, 2–9° in the thoracic spine, and 1–2° in the lumbar spine.

Spinal movements during everyday activities are complex combinations of primary motions in the sagittal, coronal, and axial planes. However, spinal kinematics cannot simply be considered in terms of superposition of independent primary motions, as the anatomy of the spinal joint leads to coupling between different degrees of freedom of a motion segment. For example, lateral bend and axial rotation, flexion and lateral bend, and flexion and axial rotation are all strongly coupled in the lumbar spine.^[3]

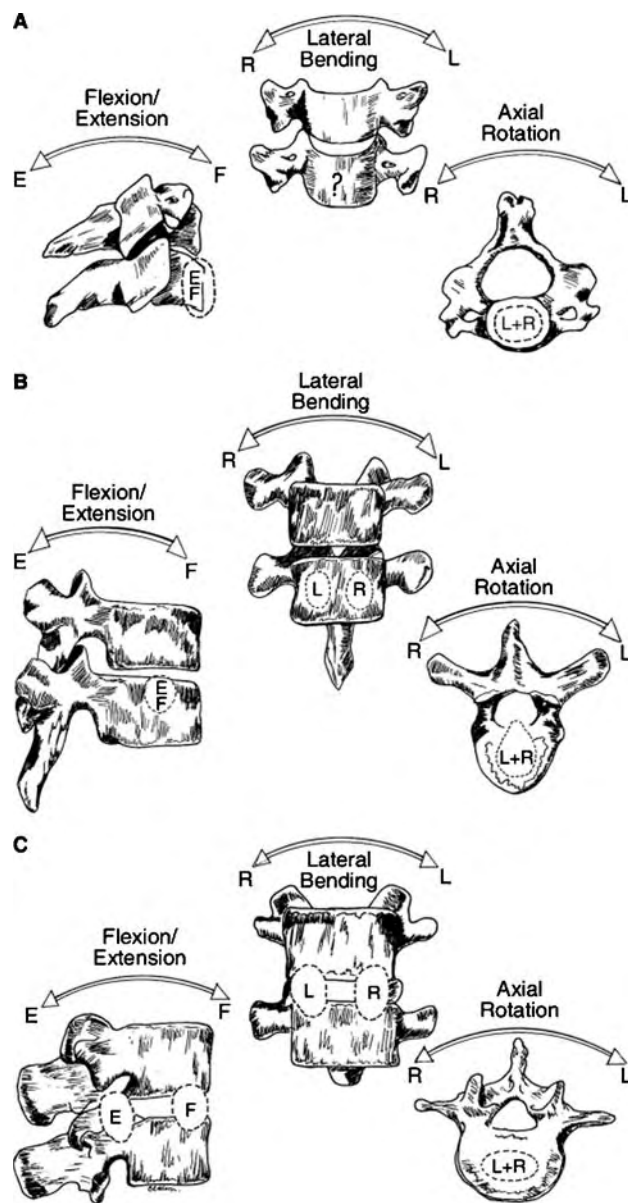


Fig. 1 IAR locations for the (A) cervical, (B) thoracic, and (C) lumbar spines for combined flexion (F)/extension (E), left (L) and right (R) lateral bending, and left and right axial rotation. (Adapted from White, A.A., III; Panjabi, M.M. Spinal kinematics. The research status of spinal manipulative therapy. NINCDS Monograph [No.15], US Department of Health, Education and Welfare: Washington, DC, 1975; 93.)

Kinetics

Spinal kinetics is the study of the forces and moments acting on the spine to either stabilize or generate motion. The earliest quantitative studies of compressive forces acting in the spine were disc pressure measurements performed by Nachemson,^[4] in which a pressure-sensing needle was inserted into the IVDs of volunteers, and the pressures generated in the nucleus pulposus during everyday tasks were measured. During relaxed standing, the compressive force at the base of the spine is equal to approximately 70% of bodyweight. Reclining reduces this pressure by 50–80%, unsupported sitting increases it by 40%, and combined forward flexion/rotation while holding a 10 kg weight increases disc pressures by 400%.

Muscle force measurements are important in static and dynamic analysis of the spine, allowing estimation of the internal forces and moments acting on passive osseoligamentous structures and the resulting tissue level stresses and strains. Muscle forces acting on the spine have been estimated via electromyography (EMG) and using mathematical models. The EMG measurements provide information on muscle activation levels during various tasks and, when combined with knowledge of muscle cross-sectional areas, can be used to infer muscle forces. The EMG measurements of the back muscles are complicated by the complexity of muscular anatomy near the spine, uncertainties in muscle cross-sectional area, and nonstraight lines of action for muscles that contact other structures between their attachment points. In recent years, magnetic resonance imaging has allowed more accurate determination of muscle cross-sectional areas, and

mathematical trunk models have incorporated more realistic muscular lines of action.

The first three-dimensional mathematical model for the whole trunk was developed by Schultz et al.,^[5] and included the rectus abdominus, oblique internus and externus, erector spinae, and latissimus dorsi muscles. Subsequent models have included dynamic (inertial) effects, lumped parameter representations of discs, ligaments, and facets, and more detailed muscle fascicle representations.^[6] Trunk models including muscles are usually statically indeterminate for a given posture, and therefore require the assumption of some muscle activation strategy (e.g., minimizing the sum of cubed muscle stresses) for solution.^[7] Dynamic activities result in higher spinal loads than static lifting,^[8] and intra-abdominal pressures do not appear to contribute substantially during dynamic lifting tasks.

Tissue Stiffness and Strength

Spinal soft tissues undergo large deformations. Pearcy and Tibrewal^[9] measured disc deformations of 35% for anterior compression and 60% for posterior extension under flexion/extension movements. They also reported a 370% increase in interspinous distance for the same motion.

The ability of spinal structures to withstand internally and externally applied forces is expressed in terms of *stiffness* (force per unit displacement) and *strength* (force required to produce failure). The stiffness and strength of ligaments, discs, and vertebrae depend, in turn, on the mechanical (stress vs. strain) properties and architecture of their constituent soft and bony tissues. In analytical and numerical studies of spine biomechanics, tissue stress-strain response is

often assumed to be linear. Although this may be a reasonable assumption for small incremental deformations, tissue material properties are generally nonlinear, exhibiting viscoelasticity, creep, hysteresis, and plasticity (permanent deformation on overloading). The nonlinear behavior of ligaments and IVDs is especially important owing to the large strains encountered in these tissues, and hyperelastic constitutive models have been developed for the annulus fibrosus.^[10] Tissue material properties are also generally anisotropic and inhomogeneous, although the common assumptions of isotropy and homogeneity in finite element studies of the spine can be somewhat justified by the need to obtain solutions for stress and strain in tissues within computing power constraints, and the fact that spinal loading for a particular structure is often predominantly unidirectional.

The stiffness and strength of spinal tissues has been assessed in numerous studies, both by direct (force vs. displacement) and by indirect experimental techniques such as indentation and ultrasound. Combined microcomputed tomography and voxel-based finite element techniques have also been used to accurately determine the stiffness and strength of vertebral trabecular bone for particular specimen architectures.^[11] Table 1 presents some of the material properties that have been used in mathematical models of spinal tissue behavior. Interspecimen variability as well as differences in testing methodology can lead to substantial ranges in reported mechanical properties between studies. For example, Adams and Dolan^[16] report a compressive strength range for the human lumbar vertebral body between 2 and 14 kN.

In addition to individual tissue properties, measurements of the stiffness of whole motion segments have been made.^[17] Some in vitro experiments on

Table 1 Summary of material properties used in mathematical models of spinal tissue behavior

Structure	Material Properties	References
Anterior longitudinal ligament	$E = 20 \text{ MPa}$, $\nu = 0.3$	[12]
Posterior longitudinal ligament	$E = 70 \text{ MPa}$, $\nu = 0.3$	[12]
Ligamentum flavum	$E = 50 \text{ MPa}$, $\nu = 0.3$	[12]
Capsular ligament	$E = 20 \text{ MPa}$, $\nu = 0.3$	[12]
Intertransverse ligament	$E = 50 \text{ MPa}$, $\nu = 0.3$	[12]
Interspinous ligament	$E = 28 \text{ MPa}$, $\nu = 0.3$	[12]
Superspinous ligament	$E = 28 \text{ MPa}$, $\nu = 0.3$	[12]
Zygapophyseal joint cartilage	$E = 10 \text{ MPa}$, $\nu = 0.4$	[13]
Vertebral trabecular bone	$E = 35\text{--}154 \text{ MPa}$	[14]
Vertebral cortical bone	$E = 12 \text{ GPa}$, $G = 4.6 \text{ GPa}$, $\nu = 0.30$	[15]
IVD annulus ground substance	$E = 4.2 \text{ MPa}$, $G = 1.6 \text{ MPa}$, $\nu = 0.45$	[15]
IVD collagenous fibers	$\sigma = 23,000\epsilon^{1.9}$ (for outer fiber layers)	[15]

E = modulus of elasticity, G = shear modulus, ν = Poisson's ratio, σ = stress, ϵ = strain.

whole motion segments have suggested the existence of a “neutral zone,” defined as a region of little or no resistance to motion in the middle of an intervertebral joints’ range of movement. However, the neutral zone does not appear to exist in all planes.^[18] Spinal joint stiffness increases rapidly near the anatomically imposed limits of movement, and the high forces required to flex motion segments beyond these limits lead to damage of the connective tissues (facet joints, discs, and ligaments).

BIOMECHANICS OF DISC DEGENERATION

The IVDs are the largest avascular structures in the body, withstanding high static and dynamic compressive and shear forces, particularly in the lumbar spine. The biomechanics of IVD degeneration in the lumbar spine has received much attention because of the great economic and social importance of lower back pain.

A number of studies have shown strong correlations between disc degeneration and back pain, and several mechanisms have been postulated for this connection; loss of hydrostatic pressure owing to a reduction in proteoglycan content of the disc may encourage nerve ingrowth,^[16] altered stress distributions owing to loss of disc height may exert additional force on facet joints and sensitive surrounding tissues causing pain,^[16,19] and disc prolapse may exert direct pressure on the spinal cord and nerve roots. Fig. 2 shows the clear structural differences between healthy and degenerate discs.

The structural characteristics of disc degeneration include loss of disc height, the presence of radial, concentric, or rim lesions in the annulus fibrosus, and prolapse of the nucleus pulposus. Disc prolapse is associated with disc height reduction, and loss of disc height following prolapse has been shown to continue even after surgical intervention.^[20] Loss of mechanical integrity in lumbar IVDs subjected to anterior shear forces (primarily L4-L5 and L5-S1) can lead to instability and spondylolisthesis (anterior slip of one vertebra relative to the inferior adjacent vertebra).

The causes of disc degeneration are not well understood, although biochemical, genetic, and mechanical factors are all likely to contribute. Owing to the lack of blood supply, adequate nutrient flow through the vertebral endplates into the disc is important for maintenance of healthy IVDs. Although some authors have suggested that convective nutrient transport owing to cyclic mechanical loading is important, nutrient flow into the disc appears to be predominantly diffusion driven.^[21] If current levels of nutrient demand within the disc cannot be met because of occlusion of pores in the osseous vertebral endplate or inadequate diffusion rates, cellular matrix degradation

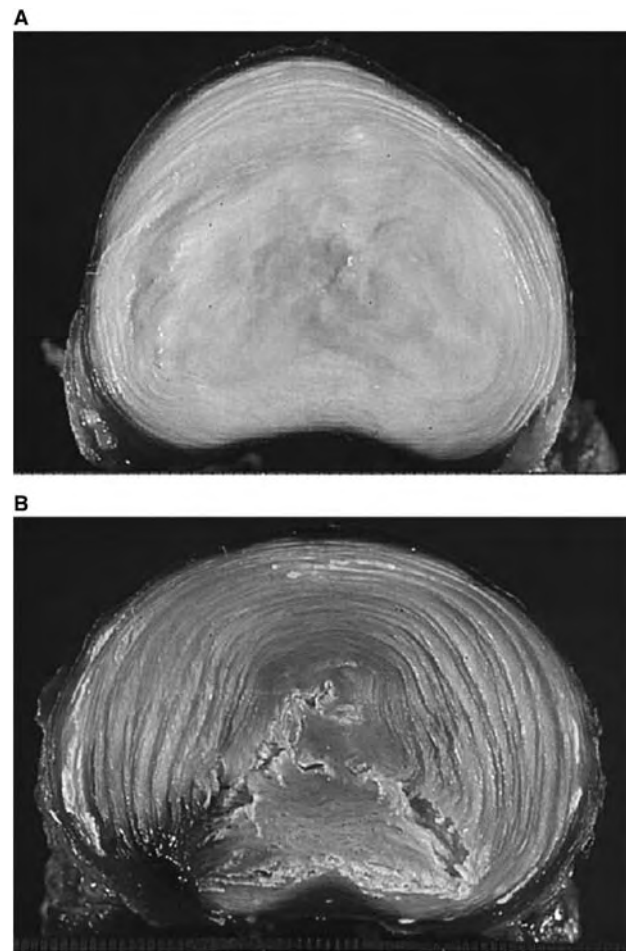


Fig. 2 Axial sections showing (A) healthy and (B) degenerate lumbar IVDs. (Photo courtesy of Professor N. Fazzalari.)

leading to localized tissue damage and annular lesions will result. Studies using progressive dissection of annular fiber layers or artificially induced lesions have reported the loss in overall disc stiffness and strength associated with a particular level of mechanical damage.^[22] Owing to the impracticability of direct measurement, finite element studies have been used to infer stress and strain levels in the disc. Under direct compression, the inner annulus appears most highly stressed;^[23,24] however maximum stresses occurring at the outer annulus have been reported for torsional loading.^[25] Existing studies show no clear consensus on the link between predicted annular stresses and the location of clinically observed disc lesions (which often occur in the posterolateral annulus).

Spinal Fusion

Spinal fusion is currently the primary surgical intervention for pain and instability associated with

end-stage disc degeneration. Fusion constructs are required to impart stability and limit relative motion between the vertebrae to be fused. A range of surgical techniques, as well as numerous different implant designs are used; however some general biomechanical principles apply in the successful creation of a surgical construct that can withstand spinal loads in the three to six months required for bony fusion to occur.

Interbody fusion constructs are susceptible to subsidence failure, in which the graft or implant penetrates into the vertebral body, causing a loss of disc space height and kyphosis at the fused level. Pearcy, Evans, and O'Brien^[26] have shown that to avoid endplate subsidence with rectangular autograft spacers, the area of the graft spacers needs to cover at least 40% of the total endplate area. More recent finite element studies of interbody fusion mechanics^[27] have predicted stress levels owing to contact between interbody implants and the vertebral body (Fig. 3). Although the stiffness and strength of the implant material is clearly an important consideration in interbody implant design, the notion that matching implant elastic modulus to that of cortical bone will result in an optimal load transfer is incorrect. Rather, the contact stress distribution between an implant and its adjacent vertebral endplate is dependent on the initial shapes and the respective bending stiffnesses (EI , where E is the modulus of elasticity and I is the second moment of area of the member cross section) of the two structures.

For posterior rod fusion constructs, stiffness and strength can be analyzed using structural mechanics techniques.^[28] Rod diameter is usually the primary factor affecting construct stiffness and strength, because maximum bending and torsional stress in a rod of circular cross section vary with the inverse cube of rod diameter.

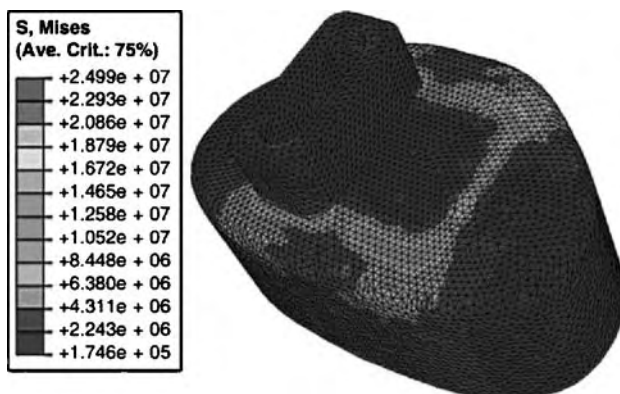


Fig. 3 Predicted von Mises stress (in Pascals) for finite element simulation of an interbody fusion implant contacting a vertebral body with applied compressive force of 610 N to simulate relaxed standing. (From Ref. [27].)

Performing a spinal fusion in either the cervical or lumbar spine not only effectively eliminates motion at the segment in question, but also affects the biomechanics of the surrounding spinal structures. In particular, spinal fusion may place increased mechanical demands on the adjacent IVDs, leading to accelerated degeneration at these levels, which eventually requires further intervention. Although reported incidence levels vary, so-called adjacent segment disease^[29] is widely regarded to be a largely unavoidable consequence of eliminating relative motion at one segment of a multijointed system.

Disc Arthroplasty

A current field of intense interest in spine biomechanics is the area of disc arthroplasty, or replacement of an entire IVD with an artificial joint that maintains mobility at the operated level. As previously discussed, the spinal motion segment is a complex joint, and current implant designs can only reproduce some aspects of the original joint's kinematics. Very little systematic investigation of the biomechanics of disc arthroplasty has been undertaken to date; however one recent study suggests that less-constrained implant designs allow a greater range of motion, are less sensitive to positioning inaccuracies during surgery, and apply less abnormal loading to facet joints than designs that constrain rotation about a single fixed point.^[30] Several longer term studies have shown that the motion permitted by TDR implants decreases over time, eventually becoming "kinematically equivalent to a fusion" in a significant proportion of TDRs.^[31] Whether or not disc arthroplasty avoids the adjacent segment degeneration associated with disc arthrodesis is unknown at this stage.^[32]

BIOMECHANICS OF SPINAL FRACTURES

Spinal fractures can occur either when the spine is subjected to a single traumatic loading event, or when repeated loading causes an accumulation of damage at a faster rate than at which the bone tissue can heal itself. The latter case refers to "stress fractures," which are often associated with certain repetitive occupational or sporting activities.

Damage of vertebral trabecular bone in the form of trabecular microfractures is most likely a normal physiological occurrence, and healing microcracks appear to be part of the bone modeling and remodeling process. However, when the spine is subjected to loads significantly higher than those encountered during normal activity, gross fractures can occur. Under large compressive loads, the vertebral endplates are

particularly susceptible to fractures, leading to cracking, fragmentation, and disc protrusion into the vertebral body. Endplate fractures are often undetectable on radiographs, and are a common feature of moderate energy impacts (e.g., ejection seat injuries).

Vertebral burst fractures are associated with high energy impacts such as falls and motor vehicle accidents, in which high internal fluid pressures “burst” the vertebral body, comminuting vertebral bone and causing retropulsion of bony fragments into the neural canal. Experimental attempts to simulate burst fracture in the laboratory are characterized by significant variability, although for younger spines with healthy IVDs the sequence of events seems to begin with fracture of the vertebral endplate under high intradiscal pressures, followed by herniation of the nucleus pulposus into the vertebral body, vertebral fracture, and fragment expulsion. Biomechanical investigations using strain-gauged cadaveric motion segments have indicated that facet joint loading also plays an important role in burst fracture mechanics,^[33] and that sagittal orientation (flexed, neutral or extended) affects resulting fracture patterns. Specimens tested in extension exhibit more severe fracture characteristics. In vitro experiments have also shown that burst fractures tend not to occur in the presence of preexisting disc degeneration and osteoporosis, suggesting that the elderly are prone to compression rather than burst fractures. The isolated vertebral body collapse in osteoporotic compression fractures carries less risk of bone fragment retropulsion and spinal cord damage.^[34] Metastatically involved vertebrae are at particular risk of burst fracture, and a recent experimental and finite element investigation concluded that tumor size was the principal risk factor for burst fracture initiation, followed by loading magnitude and bone density.^[35]

Other spinal fractures can occur owing to repetitive overuse, particularly in the case of nonconditioned individuals where bone microarchitecture has not had the opportunity to adapt to newly imposed repetitive loads.^[36] Stress fractures in butterfly swimmers and cricket fast bowlers are two examples of fatigue damage occurring because of microcrack coalescence over a period of extended cyclic overloading. Principles of linear elastic fracture mechanics have been applied to study fatigue crack growth in both trabecular and cortical bone, with Nalla et al.^[37] finding that the Paris law applied well to bone

$$\frac{da}{dN} = C(\Delta K)^m$$

where da/dN is the crack growth rate per cycle, C is a constant, ΔK is the alternating stress intensity factor, and m is the Paris law exponent, ranging between 4.4 and 9.5 for human cortical bone.

One of the most common age-related degenerative skeletal pathologies is osteoporosis, in which vertebral trabecular bone loss can eventually lead to vertebral compression fractures in both the thoracic and lumbar spine. Studies of vertebral load distribution have shown that in a healthy vertebral body, between 35% and 55% of vertical compressive load is carried by the trabecular bone. The loss of trabecular bone will therefore have major effects on vertebral load-bearing ability. Osteoporosis is characterized by both reduction in trabecular strut diameter and loss of struts. Structurally, loss of struts has a much more significant effect on the reduction of overall bone stiffness and strength, with Silva and Gibson^[38] reporting a 70% reduction in strength owing to a 10% decrease in bone volume by strut removal, vs. only a 20% decrease in strength for the same decrease in bone volume owing to strut thinning.

Vertebral compression fractures are commonly treated by injection of polymethylmethacrylate cements (vertebroplasty). Cement injection may be preceded by inflation of a compliant balloon in the collapsed vertebra to restore height (kyphoplasty). In vitro and computational studies of the biomechanics of vertebroplasty have concluded that while vertebroplasty can restore vertebral height and reduce the likelihood of subsequent fracture at the treated level, the high stiffness of the augmented vertebra exerts increased stresses on neighboring structures, with greater risk of damage to adjacent vertebral bodies.^[39]

BIOMECHANICS OF SPINAL DEFORMITY

Spinal deformities such as scoliosis and hyperkyphosis can be caused by a range of factors including congenital malformation, disease, neoplasm, and neuromuscular disorders, or as a secondary biomechanical manifestation of some other cause such as osteoporotic compression fractures or leg length inequality. However, the majority (70%) of spinal deformities are classified as idiopathic scoliosis.

There is still much debate regarding the etiology of “idiopathic” scoliosis, and although a number of promising hypotheses are currently being investigated as to the causes of the initial abnormal growth, there is general consensus that progression (worsening of the abnormal curve) in scoliosis is primarily a biomechanical phenomenon. The laterally curved spine leads to abnormal loading under body weight and muscle forces, and this abnormal loading causes further asymmetrical bone growth, which then worsens the deformity—an unstable “vicious cycle.”^[40]

One biomechanical theory for the initiation of scoliosis is the anterior overgrowth hypothesis,^[41] in which the anterior (vertebral) column of the spine

grows faster than the posterior column, generating internal compression in the vertebral column owing to the restraint of the posterior joint complex, and eventually leading to buckling of the entire spinal column in its least stable plane—the coronal plane. Numerous authors have made reference to the Euler equation of column buckling to describe the onset of scoliotic deformity, and Ogilvie^[42] has pointed out that girls with idiopathic scoliosis tend to have longer, more slender spinal columns, which would be conducive to buckling according to Euler's equation

$$P_{cr} = K \frac{EI}{L^2}$$

Where P_{cr} is the critical buckling load for the column, K is a constant dependent on the constraint conditions at the ends of the column, E is the modulus of elasticity of the column material, I is the second moment of area of the column cross section (with the term EI representing the bending stiffness of the column), and L is the effective length of the column. Although it may be conceptually useful, the Euler buckling equation is only valid for idealized geometry, loading conditions, and isotropic linear elastic material properties, and thus cannot be used in a quantitative sense in the study of spinal deformity.

Recent animal experiments by Stokes using in vivo rat and calf tail models have shown that the rate of vertebral growth is strongly dependent on applied load, with growth being accelerated by up to 30% above controls with a 0.1 MPa tensile stress, and growth reduced by 40–50% relative to controls with 0.2 MPa compressive stress. These results provide significant insight into the mechanical modulation of vertebral growth by endochondral ossification, because the wedge-shaped growth of vertebrae in progressive scoliosis can be predicted by models that account for the compressive pressure dependence of bone growth.^[40] Modulation of growth by mechanical forces is often referred to as the Heuter-Volkman law. Several studies have used finite element techniques to implement mechanically modulated growth in three-dimensional spine models.^[43,44] Although the models are capable of predicting scoliosis progression during bone growth and remodeling, uncertainties in model parameters can overshadow model sensitivity to small initial perturbations in spinal geometry.

Treatment of spinal deformities, whether conservative or surgical, currently involves application of corrective forces to the spine in an attempt to restore normal geometry and prevent further asymmetric growth. Biomechanically, the problem of straightening a bent column can be approached by distraction (stretching), bending in the opposite direction to the deformity, or compression on the convex side of the deformity. The use of orthoses for nonsurgical treatment of scoliosis is

an example of the second method, where the brace applies a lateral corrective force at the apex of the deformity, resisted contralaterally above and below the apical level, effectively forming a three-point bending construct. The biomechanics of bracing has been investigated both by direct measurement of pressures between the brace and skin, and using numerical simulations with personalized geometry obtained from three-dimensional imaging techniques.^[45]

Surgical treatment of spinal deformities involves the use of rods, hooks, screws, and wires to straighten and stabilize the spine. Harrington distraction rods are effective at correcting large scoliotic deformities, but as the curve severity decreases, larger distraction forces are required to generate the same corrective bending moment, with associated risk of neurological deficit owing to stretching of the spinal cord and loss of sagittal plane spinal profile. The forces acting on spinal implants in vivo have been measured directly using strain gauges, and the stiffness of various implant constructs has been assessed with in vitro experiments.^[46] The stiffness of implant constructs is particularly important for correction with fusion, because relative motion between vertebrae to be fused must be limited to allow bone growth to occur.

Anterior implant systems for scoliosis correction rely primarily on tension in the implant applying compressive forces to the convex side of the vertebral column. Both cables and rods have been used in this manner, although rods are loaded in both tension and bending, therefore allowing potentially greater curve correction, albeit with greater risk of rod fracture. Fracture of anterior rods appears to be caused by pseudarthrosis (fibrous rather than bony fusion) at an attempted fusion level, such that the rod experiences high cyclic stresses and eventually fails owing to bending fatigue several years after surgery. Anterior constructs are held in place by cancellous bone screws inserted laterally through the vertebral body. The stability of these screws is related to both cancellous bone shear strength and area of screw thread in contact with the bone; thus bone density and screw diameter are the two primary determinants of bone screw pullout strength for anterior scoliosis constructs.

Posterior correction of spinal deformity involves the positioning of implants adjacent to the posterior spinal elements, attached to the laminae or pedicles with screws, hooks, and wires. Earlier posterior constructs using two longitudinal rods were improved by the addition of transverse cross-members, allowing better correction of vertebral rotation and stability in torsion. Finite element models have also been used to simulate scoliosis correction using posterior implant constructs, and the development of patient-specific computer simulations of spinal deformity correction is an area of current research activity (Fig. 4).

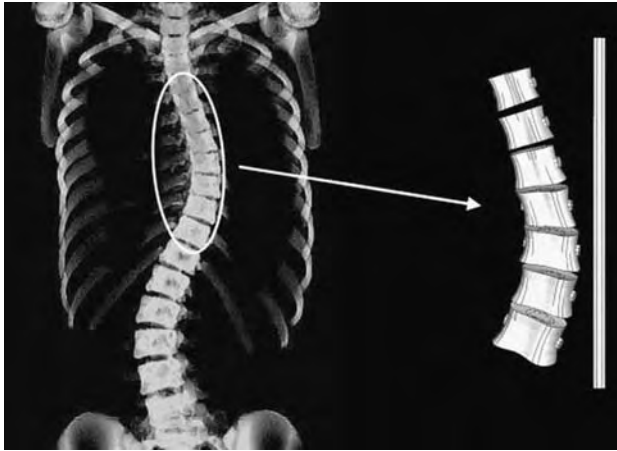


Fig. 4 Coronal computed tomography image of idiopathic scoliosis patient and patient-specific solid model of vertebral column geometry from T5-T11 (with anterior correction rod) for finite element simulation to predict biomechanical outcomes of corrective surgery.

CONCLUSIONS

This article has briefly summarized three of the major areas of interest in spine biomechanics; disc degeneration, spinal fractures, and spinal deformity. The kinematics, internal forces, and tissue material properties of the healthy spine were also discussed. Important biomechanical principles for spinal implant design include the need to consider joint kinematics and the alterations in load distribution and tissue level stresses that fusion constructs or enforced changes in joint kinematics will produce. The mechanics of vertebral burst and compression fractures were discussed, in particular the importance of trabecular bone in supporting spinal compressive loads, and the potential damage to adjacent vertebrae caused by compression fracture augmentation with vertebroplasty or kyphoplasty. Biomechanical hypotheses for the initiation and progression of spinal deformity were summarized, and principles of deformity correction with orthoses or corrective implants were identified. Future advances in spine biomechanics are likely to lie at the interfaces of engineering mechanics and biology, particularly in the areas of tissue engineering and mechanobiology. Further understanding of the role of mechanical forces in modulating tissue growth and remodeling will lead to improved interventions: minimally invasive and preserving normal kinematics as closely as possible. The wide use of numerical simulation in spine biomechanics will continue, with more robust algorithms and greater computing power leading to patient-specific applications of spinal finite element analysis. Finally, further advances in medical imaging technology will enhance our ability to accurately measure spinal anatomy and movement for both healthy and pathological states.

ACKNOWLEDGMENTS

The author would like to thank Professor Mark Percy, Professor John Evans, Dr Geoffrey Askin, and Mrs Maree Izatt for their contributions to this article.

S

ARTICLES OF FURTHER INTEREST

Bone Remodeling; Finite Element Analysis; Hard Tissue: Biomaterials Interaction; Hard Tissue Elastic Properties; Healing of Bone and Connective Tissues

REFERENCES

1. White, A.A.; Panjabi, M.M. *Clinical Biomechanics of the Spine*, 2nd Ed.; J.B Lippincott Company: Philadelphia, 1990.
2. Bogduk, N.; Amevo, B.; Percy, M. A biological basis for instantaneous centres of rotation of the vertebral column. *Proc. Inst. Mech. Eng. H* **1995**, *209* (3), 177–183.
3. Russell, P.; Percy, M.J.; Unsworth, A. Measurement of the range and coupled movements observed in the lumbar spine. *Br. J. Rheumatol.* **1993**, *32* (6), 490–497.
4. Nachemson, A.L. Disc pressure measurements. *Spine* **1981**, *6* (1), 93–97.
5. Schultz, A.B.; Andersson, G.B.J.; Ortengren, R.; Haderspeck, K.; Nachemson, A. Loads on the lumbar spine. Validation of a biomechanical analysis by measurements of intradiscal pressures and myoelectric signals. *J. Bone Joint Surg. Am.* **1982**, *64A* (5), 713–720.
6. Cholewicki, J.; McGill, S.M. Mechanical stability of the in vivo lumbar spine: implications for injury and chronic low back pain. *Clin. Biomech.* **1996**, *11*, 1–15.
7. Stokes, I.A.F.; Gardner-Morse, M. Lumbar spinal muscle activation strategies predicted by multi-criteria cost function. *J. Biomech.* **2001**, *34*, 733–740.
8. Han, J.S.; Goel, V.K.; Ahn, J.Y.; Winterbottom, J.; McGowan, D.; Weinstein, J.; Cook, T. Loads in the spinal structures during lifting: development of a three-dimensional comprehensive biomechanical model. *Eur. Spine J.* **1995**, *4* (3), 153–168.
9. Percy, M.J.; Tibrewal, S.B. Lumbar intervertebral disc and ligament deformations measured in vivo. *Clin. Orthop. Relat. Res.* **1984**, *191*, 281–286.
10. Natali, A.N. A hyperelastic and almost incompressible material model as an approach to intervertebral disc analysis. *J. Biomed. Eng.* **1991**, *13* (2), 163–168.
11. Keaveny, T.M.; Morgan, E.F.; Niebur, G.L.; Yeh, O.C. Biomechanics of trabecular bone. *Ann. Rev. Biomed. Eng.* **2001**, *3*, 307–333.
12. Lu, M.; Hutton, W.C. Do bending, twisting, and diurnal fluid changes in the disc affect the propensity to prolapse? A viscoelastic finite element model. *Spine* **1996**, *21*, 2570–79.

13. Kumaresan, S.; Yoganandan, N.; Pintar, F.A. Finite element modeling approaches of human cervical spine facet joint capsule. *J. Biomech.* **1998**, *31*, 371–376.
14. Linde, F. Elastic and viscoelastic properties of trabecular bone by a compression testing approach. *Dan. Med. Bull.* **1994**, *41* (2), 119–138.
15. Shirazi-Adl, S.A.; Suresh, C.S.; Ahmed, A.M. Stress analysis of the lumbar disc-body unit in compression. *Spine* **1984**, *9* (2), 120–132.
16. Adams, M.A.; Dolan, P. Spine biomechanics. *J. Biomech.* **2005**, *38* (10), 1972–1983.
17. Stokes, I.A.; Gardner-Morse, M.; Churchill, D.; Laible, J.P. Measurement of a spinal motion segment stiffness matrix. *J. Biomech.* **2002**, *35* (4), 517–521.
18. Thompson, R.E.; Barker, T.M.; Pearcy, M.J. Defining the neutral zone of sheep intervertebral joints during dynamics motions: an in vitro study. *Clin. Biomech.* **2003**, *18* (2), 89–98.
19. Evans, J.H. Clinical biomechanics of the spine. In *Biomechanics in Orthopedics*; Niwa, S., Perren, S.M., Hattori, T., Eds.; Springer-Verlag: Tokyo, 1992; 31–45.
20. Tibrewal, S.B.; Pearcy, M.J. Lumbar intervertebral disc heights in normal subjects and patients with disc herniation. *Spine* **1985**, *10* (5), 452–454.
21. Urban, J.P.G.; Smith, S.; Fairbank, J.C.T. Nutrition of the intervertebral disc. *Spine* **2004**, *29* (23), 2700–2709.
22. Thompson, R.E.; Pearcy, M.J.; Barker, T.M. The mechanical effects of intervertebral disc lesions. *Clin. Biomech.* **2004**, *19* (5), 448–455.
23. Shirazi-Adl, S.A.; Shrivastava, S.C.; Ahmed, A.M. Stress analysis of the lumbar disc-body unit in compression: a three-dimensional nonlinear finite element study. *Spine* **1984**, *9* (2), 120–134.
24. Fagan, M.J.; Julian, S.; Siddall, D.J.; Mohsen, A.M. Patient specific spine models. Part 1: Finite element analysis of the lumbar intervertebral disc—a material sensitivity study. *Proc. Inst. Mech. Eng.* **2002**, *216H*, 299–314.
25. Ueno, K.; Liu, Y.K. A three-dimensional nonlinear finite element model of lumbar intervertebral joint in torsion. *J. Biomech. Eng.* **1987**, *109*, 200–209.
26. Pearcy, M.J.; Evans, J.H.; O'Brien, J.P. The load bearing capacity of vertebral cancellous bone in interbody fusion of the lumbar spine. *Eng. Med.* **1983**, *12* (4), 183–184.
27. Adam, C.J.; Pearcy, M.J.; McCombe, P. Stress analysis of interbody fusion—finite element modeling of intervertebral implant and vertebral body. *Clin. Biomech.* **2003**, *18*, 265–272.
28. Carson, W.L.; Asher, M.A. Parameters influencing spine implant construct axial-flexion stiffness. Proceedings of the Harrington Spine Symposium, Kansas City, Kansas, July 28–30, 2005; Asher, M.A., Ed.; University of Kansas Medical Center: Kansas, 2005; 83 pp.
29. Hilibrand, A.S.; Robbins, M. Adjacent segment degeneration and adjacent segment disease: the consequences of spinal fusion? *Spine* **2004**, *4* (6 suppl), 190S–194S.
30. Huang, R.C.; Girardi, F.P.; Cammisa, F.P.; Wright, T.M. The implications of constraint in lumbar total disc replacement. *J. Spinal Disord. Tech.* **2003**, *16* (4), 412–417.
31. Huang, R.C.; Wright, T.M.; Panjabi, M.M.; Lipman, J.D. Biomechanics of nonfusion implants. *Orthop. Clin. N. Am.* **2005**, *36*, 271–280.
32. German, J.W.; Foley, K.T. Disc arthroplasty in the management of the painful lumbar motion segment. *Spine* **2005**, *30*, S60–S67.
33. Langrana, N.A.; Harten, R.D.; Lin, D.C.; Reiter, M.F.; Lee, C.K. Acute thoracolumbar burst fractures: a new view of loading mechanisms. *Spine* **2002**, *27* (5), 498–508.
34. Shirado, O.; Kaneda, K.; Tadano, S.; Ishikawa, H.; McAfee, P.C.; Warden, K.E. Influence of disc degeneration on mechanism of thoracolumbar burst fractures. *Spine* **1992**, *17* (3), 286–292.
35. Whyne, C.M.; Hu, S.S.; Lotz, J.C. Burst fracture in the metastatically involved spine: development, validation, and parametric analysis of a three-dimensional poroelastic finite-element model. *Spine* **2003**, *28* (7), 652–660.
36. Huiskes, R.; Ruimerman, R.; van Lenthe, G.H.; Janssen, J.D. Effects of mechanical forces on maintenance and adaptation of form in trabecular bone. *Nature* **2000**, *405* (6787), 704–706.
37. Nalla, R.K.; Kruzic, J.J.; Kinney, J.H.; Ritchie, R.O. Aspects of in vitro fatigue in human cortical bone: time and cycle dependent crack growth. *Biomaterials* **2005**, *26* (14), 2183–2195.
38. Silva, M.J.; Gibson, L.J. Modeling the mechanical behavior of vertebral trabecular bone: effects of age-related changes in microstructure. *Bone* **1997**, *21* (2), 191–199.
39. Wilcox, R.K. The biomechanics of vertebroplasty: a review. *Proc. Inst. Mech. Eng. H* **2004**, *218* (1), 1–10.
40. Stokes, I.A.F. How vicious is the ‘vicious cycle’? Proceedings of the International Research Society of Spinal Deformities Symposium, Vancouver, Canada, June 10–12, 2004; Sawatsky, B.J., Ed.; International Research Society of Spinal Deformities: Vancouver, 2004; 331 pp.
41. Millner, P.A.; Dickson, R.A. Idiopathic scoliosis: biomechanics and biology. *Eur. Spine J.* **1996**, *5* (6), 362–373.
42. Ogilvie, J.W. Biomechanics. In *Moe's Textbook of Scoliosis and Other Spinal Deformities*; Bradford, D.S., Lonstein, J.E., Moe, J.H., Ogilvie, J.W., Winter, R.B., Eds.; W.B. Saunders Company: Philadelphia, 1987; 7–24.
43. Goto, M.; Kawakami, N.; Azegami, H.; Matsuyama, Y.; Takeuchi, K.; Sasaoka, R. Buckling and bone modeling as factors in the development of idiopathic scoliosis. *Spine* **2003**, *28* (4), 364–370.
44. Villemure, I.; Aubin, C.E.; Dansereau, J. Biomechanical simulations of the spine deformation process in adolescent idiopathic scoliosis from different pathogenesis hypotheses. *Eur. Spine J.* **2004**, *13* (1), 83–90.
45. Perie, D.; Aubin, C.E.; Petit, Y. Personalized biomechanical simulations of orthotic treatment in idiopathic scoliosis. *Clin. Biomech.* **2004**, *19* (2), 190–195.
46. Panjabi, M.M.; Abumi, K.; Duranceau, J.; Crisco, J.J. Biomechanical evaluation of spinal fixation devices: II. Stability provided by eight internal fixation devices. *Spine* **1988**, *13* (10), 1135–1140.

Stem Cells

Andrea Liebmann-Vinson

Sharon C. Presnell

BD Technologies, Research Triangle Park, North Carolina, U.S.A.

S

INTRODUCTION

The words “stem cell” were once reserved for special cells that are capable of giving rise to all cells of an organism, such as embryonic stem cells, or all cells of a tissue, such as the hematopoietic stem cells. However, the classical definition of stem cell has evolved as we have learned more about cellular plasticity and discovered the presence of multipotential stem-like cells in many adult tissues. In today’s world, words such as transdifferentiation and dedifferentiation seem less like science fiction and more like real phenomena. In the past decade, we have read reports of remarkable cells from adult bone marrow and muscle able to repair damage and regenerate function in critical organs such as liver, brain, and heart.^[1–3] One study offered evidence that a single bone marrow–derived cell when transplanted into a recipient could home to a variety of tissues and differentiate into several functional cell types including liver, lung, GI tract, and skin.^[4] Likewise, researchers have demonstrated that embryonic stem cells can be expanded in vitro and differentiated into a wide variety of cell types in vitro and in vivo.^[5] Stem cells have the potential to profoundly impact the development of new therapeutics for a wide variety of human diseases. In this article, we will discuss the key properties of stem cells, as well as some potential pitfalls regarding their use in human therapeutics. Furthermore, we will provide examples and discussion of the importance of interactions between stem cells and biomaterials.

THE STEM CELL PHENOTYPE

In the broadest sense, a stem cell is a cell that is capable of self-renewal and differentiation into at least one functional cell type. If we think of stem cells existing along a continuum of plasticity, embryonic stem (ES) cells would be on one end of the spectrum, while other cells with limited differentiation potential (e.g., committed progenitors of the hematopoietic system) would be on the other end. There are some markers that have been suggested for stem cells, and these markers are routinely used as tools to selectively isolate ES cells, mesenchymal stem cells (MSCs), and hematopoietic

stem cells (HSCs). The NIH has provided an informative web publication that outlines the markers, research progress, and potential issues surrounding these specific stem cell types, as well as others (please refer to www.nih.gov/news/stemcell/scireport.htm). For a partial list of markers associated with some specific stem cells, see Table 1. Although the presence of specific markers on a particular cell may give an indication of the stem-like phenotype of a cell, it would be perhaps naïve to assume that these markers are all-encompassing or are absolutely definitive. A cell’s ability to self-renew and differentiate in vivo is a better indication of its stem-like nature than expression of specific markers in vitro.

ANALYSIS OF STEM CELLS

Classical analysis of well-characterized stem cells such as ES cells, MSCs, and HSCs usually involves detection of specific markers on the cells’ surface using a panel of antibodies and flow cytometric analysis (see Table 1 for listing of classical markers). There are many other ways in which cells that are believed to be stem-like in nature can be characterized in vitro. When normal tissues such as liver, kidney, prostate, and intestine are dissociated mechanically or enzymatically, it is possible to obtain short-term cultures of the differentiated cells of that tissue. However, establishment of cell strains or lines that can be propagated in vitro through multiple passages and for longer periods of time is difficult without genetic manipulation. In primary cultures of some tissues, such as adult rat liver, colonies of small epithelial cells will emerge as the majority of the fully differentiated cells die; these small cells can be propagated long-term in vitro but they express few markers and do not secrete albumin or have characteristics consistent with mature liver cells (hepatocytes).^[7] However, upon in vitro treatment with a potent differentiating agent (sodium butyrate) or after transplantation into the liver of syngeneic rats, the cells differentiate into functional hepatocytes.^[8] Thus, the mere fact that these cells could emerge and be propagated from a primary culture of cells from a normal adult tissue is a sign that they are self-renewing, a characteristic of

Table 1 Marker profile of embryonic stem cells (ES), hematopoietic stem cells (HSC), and mesenchymal stem cells (MSC)

Stem cell type	Putative marker profile	
	Negative	Positive
ES	SSEA-1	SSEA-3, SEA-4, TRA-1-60, TRA-1-81, Alkaline phosphatase, Oct 4, Telomerase
HSC	CD38, Thy-1 (-/low), Lin	CD34, c-kit (CD117), CD133
MSC	Lin	CD34, Sca-1, Thy-1, c-kit (CD117)

The information in this table was compiled from Ref. [6].

stem cells. The additional fact that the cells can be stimulated to differentiate into specialized cell types by changing their microenvironment in vitro and in vivo provides additional evidence of their stem-like nature. Functional analyses such as these are as critical as marker expression in evaluating the stem-like potential of a particular cell. Experiments aimed at in-vivo analysis of stem cells deserve a few words of caution. It has come to light recently that transplanted cells (usually tagged with a genetic marker such as beta galactosidase or a fluorescent protein) may fuse with host cells, creating cells that contain genetic material from both cells.^[9,10] The phenomenon of cell fusion may lead to the misinterpretation of results, as host stem cells that may have picked up the genetic tag during the fusion process can be mistaken for the transplanted cells. Ideally, the transplanted cells should be recovered from the in-vivo environment and compared genetically to a sample of the cells taken prior to transplantation; evidence of fusion, such as an increase in ploidy, could be detected. Further protection against misinterpretation of results related to transplantation of stem cells can be achieved by transplanting the cells into a disease model and looking for functional correction of the disease state rather than simple appearance of cells.

SOURCES OF STEM CELLS

Stem cells have been isolated from embryonic tissues and from a variety of adult tissues, including blood, bone marrow, muscle, liver, pancreas, and intestine.^[11] An interesting debate has arisen over the possibility that stem cells that reside in adult tissues are recruited to those tissues as needed from the bone marrow. In 1999, researchers at the University of Pittsburgh demonstrated, in a rodent model of liver damage, that

progenitor cells in the bone marrow were able to home to the site of damage in the liver and participate in the repair process, ultimately giving rise to cells with a phenotype consistent with normal functional hepatocytes in the liver.^[1] Subsequent studies from a group at Albert Einstein University suggested rescue of lethally irradiated mice after transplantation of a single progenitor cell from the bone marrow; cells derived from the transplanted cell were subsequently identified in many tissues throughout the animal, including liver, lung, intestine, and skin.^[4] Since these seminal studies, there have been a large number of similar studies whereby transplantation of stem cells has resulted in the incorporation of those cells into host tissues and subsequent rescue of disease phenotypes. These studies have included rescue of animals from liver failure, cardiac failure, diabetes, and neurological deficits.^[12] Recent research from the University of North Carolina has now demonstrated that stem-like cells isolated from unmanipulated adult rat liver can differentiate into functional hematopoietic cells when transplanted into the bone marrow.^[13] Taken together, these studies provide strong evidence that stem cells within adult tissues are somewhat interchangeable (Fig. 1). Regardless of their origin, adult tissue stem cells represent a viable source of cells for research and for therapeutic applications.

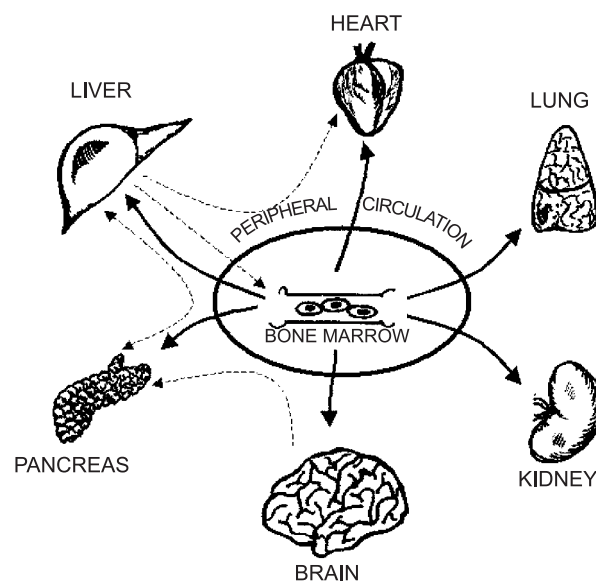


Fig. 1 The interchangeable stem cell hypothesis. Primitive stem cells may exist within the bone marrow that are capable of homing to tissues that need repair or regeneration (solid lines). Likewise, progenitor cells from an adult organ, such as liver, may act as a source of re-generation-competent cells for another tissue, including the bone marrow (dotted lines).

HOW DO STEM CELLS DIFFERENTIATE?

If stem cells are indeed interchangeable, how do they know what cell types to become? If cells from the bone marrow can differentiate into hepatocytes, why don't we find hepatocytes in the bone marrow? The answer to this question lies within each individual cell and its ability to interpret the signals from its surroundings. Each cell in the body exists within a microenvironment consisting of the unique set of neighboring cells, structural molecules, and soluble factors, all of which are able to initiate signaling cascades that ultimately affect the gene expression profile and phenotype of that cell. No two microenvironments in the body are exactly the same, so it is reasonable to assume that the same stem cell transplanted to two distinct microenvironments (such as the liver and the bone marrow) will receive tissue-specific signals that govern the growth and differentiation of that cell. There are many examples of cells exhibiting distinct growth and/or differentiation characteristics upon transplantation into different tissues. Therefore, it is reasonable to assume that the general properties of stem cells that reside within tissues are distinct from stem cells that reside in the bone marrow, even if they are interchangeable or have a precursor/progeny relationship. Returning to the continuum of plasticity described previously, bone marrow-derived stem cells may have a wide differentiation potential, but they must go through many steps in order to be differentiated into highly specialized cells of a tissue. In contrast, stem cells that reside in an adult tissue may have a more narrow range of plasticity, but they are only a step or two away from differentiating into the specialized cells of that tissue, thus making them an interesting option when tissue-specific therapies are developed. For example, in a normal mammal, the fully mature hepatocytes of the liver are able to act as the stem cells of the liver; in the event that 2/3 of the liver is removed surgically, the remaining hepatocytes proliferate to restore the functional mass of the liver within days.^[14] Grompe et al. have demonstrated that, in a damage model where the host hepatocytes are unable to proliferate, transplantation of fully differentiated hepatocytes results in a full restoration of liver mass and alleviation of the disease state.^[15] Serial transplantation experiments demonstrated that fully differentiated hepatocytes could undergo 7.3×10^{20} -fold expansion to restore liver mass and provide correction of liver disease.^[15] In contrast, transplantation of less differentiated stem-like cells from liver or from extrahepatic sources leads to incomplete restoration of liver mass and a slower recovery.^[16] This is likely due to the necessity for the transplanted stem-like cells to receive the appropriate signals from the hepatic microenvironment that allow the cells to differentiate into

hepatocyte-like cells; until these signals are received, interpreted, and implemented, the stem-like cells are incapable of initiating repair. These microenvironmental signals that govern the differentiation state of transplanted cells are important considerations in tissue engineering as new tools are developed for the expansion, differentiation, and in-vivo delivery of cells for therapy (see the following section).

S

BIOMATERIALS AND STEM CELLS

Cell–Biomaterial Interactions

Biological tissues and organs consist of cells surrounded by a complex molecular framework that is known as the extracellular matrix (ECM). ECM is a complex mixture of different macromolecules, such as collagens, proteoglycans, glycoproteins, and elastin and one of its important roles is to provide tissues with the appropriate structural integrity.^[17] In addition, cell–ECM interactions, mediated by cell surface receptors, such as integrins and syndecans, have been shown to promote signaling pathways that influence key cell function, such as proliferation, migration, and differentiation.^[17,18]

Due to the important role of ECM in the development and maintenance of living tissues and organs, the development of an artificial substitute for ECM has been a focus of biomaterial research with the goal to model and understand cell function and eventually to enable reconstruction or regeneration of damaged tissues and organs. The approaches range from polymer surfaces coated with different ECM molecules to surfaces modified with peptide sequences from ECM molecules that have been identified as cell-adhesion motifs. These approaches have led to the development of a number of three-dimensional matrices, so called scaffolds, which are commonly used today in biomaterial research and tissue engineering. In the following, we review a few examples of stem cell research that utilize such biomaterials.

Applications of Biomaterials in Stem Cell Research

Hematopoietic stem cells

Hematopoietic stem cells (HSCs) reside in the human bone marrow and they give rise to all blood cells, including lymphocytes, red blood cells, platelets, and leukocytes. Cell transplants containing HSCs, e.g., bone marrow transplants, are currently used to restore the hematopoietic system of cancer patients whose bone marrow is compromised as a result of radiation- or

chemotherapy. HSCs are rare cells and transplants are hindered by the availability of a sufficient amount of these cells. Current research efforts are focused on in-vitro expansion of HSCs while maintaining the stem cell character, as measured by the ability to restore hematopoietic function.

So far, in-vitro HSC expansion has only been achieved with mouse bone marrow in long-term cultures on stromal feeder layers.^[19] For clinical applications, however, HSC expansion cultures need to be well defined; thus, an undefined stromal layer of potentially animal origin is unacceptable. Attempts to substitute the stromal layer with growth factor cocktails have shown only partial success. In another approach, Rosenzweig et al. investigated the influence of topographical cues provided by an inert, three-dimensional, tantalum-coated biomaterial scaffold in the absence of exogenous growth factors (but presence of serum) on hematopoietic progenitor cells.^[20] More hematopoietic progenitor cells were maintained in culture on the three-dimensional scaffold compared to a two-dimensional culture on stroma or two-dimensional culture on the scaffold material, indicating that HSC phenotype may be maintained in short-term cultures on three-dimensional scaffolds. In a later study, hematopoietic progenitor cells were cultured on the same (fibronectin) scaffold but in the presence (or absence) of various growth factors.^[21] The results from that study suggested that a synergy exists between use of a three-dimensional scaffold and low-dose cytokine cocktails, resulting in the preservation and expansion of both more mature and immature hematopoietic progenitor cells.

Mesenchymal stem cells

Mesenchymal stem cells (MSCs) also originate from the bone marrow and can give rise to chondrocytes, osteoblasts, adipocytes, myoblasts, and even neural cells.^[22] MSCs present a potential cell source for cartilage and bone repair and have thus been the subject of a number of tissue engineering studies. Martin et al. described the in-vitro formation of cartilaginous and bone-like tissue when culturing chick-embryo bone-marrow stromal cells on three-dimensional polyglycolic acid scaffolds in the presence of fibroblast growth factor-2.^[23] Similarly, bovine MSCs were found to form either cartilaginous or bone-like tissue when cultured on poly(lactic-co-glycolic acid) and poly(ethylene glycol) scaffolds in the presence of the right stimulants.^[24] MSCs in a collagen matrix, seeded onto a pretensioned polygluconate suture, were found to aid tendon repair in vivo when the MSCs-collagen-suture construct was implanted into a tendon gap model.^[25,26]

Neural stem cells

Neural stem cells (NSCs) are of interest as a potential cell source for the repair of spinal cord or brain injury, following stroke, for example, or as cells used for therapies to alleviate neurodegenerative diseases such as Parkinson's disease or Alzheimer's. Some promising tissue engineering approaches have been reported. For example, a polymer scaffold seeded with NSCs and implanted into a spinal-cord defect of a rat was found to promote long-term function as exhibited in weight-bearing hindleg stepping.^[27] When seeded onto a polyglycolic acid mesh, NSCs were found to profoundly adhere, grow, and migrate in vitro.^[28] When this cell-scaffold construct was implanted into an extensively injured mouse brain, NSCs were found to differentiate, and interactions with the surrounding host tissue occurred, leading to the formation of a neuronal fiber web composed of both host and implanted cells. In this example, the scaffold helped to bridge a large lesion in the mouse brain and thus supported the formation of new tissue that structurally integrated with the host tissue.

Preadipocytes

Adipose tissues (soft or fat tissue) are composed of fragile nonproliferating mature adipocytes and proliferating preadipocytes that differentiate into adipocytes. Adipose tissue augmentation for congenital deformities, posttraumatic repair, cancer rehabilitation, and soft tissue defects in general is challenging because of the fragility of the soft tissue and often muscle or other materials, including silicon and collagen have been used as substitutes.^[29] Tissue engineering approaches to adipose tissues have focused on using preadipocytes in combination with different scaffolds. Attachment, proliferation, and differentiation of preadipocytes on fibronectin-coated polytetrafluoroethylene mesh was reported by Kral et al.^[30] Similarly, a poly(lactic-co-glycolic) acid scaffold was found to be a suitable substrate for preadipocyte differentiation.^[31]

CONSIDERATIONS OF STEM CELLS FOR IN-VIVO THERAPY

Stem cells are a promising tool in the field of regenerative medicine. The therapeutic value of stem cells is clear, given their critical use in reconstitution of bone marrow after chemotherapy for hematopoietic cancers.^[32] There is much excitement about the use of stem cells in the therapy of other diseases such as diabetes, neurodegenerative disorders, and cardiac

dysfunction. However, the clinical reality of cell therapy requires further consideration. Most studies today are conducted using syngeneic animal models, whereby cells from the donors are a near-perfect match to cells of the diseased host, thus enabling transplantation of cells without immune suppression. Immune-deficient models, such as athymic mice, SCID mice, and Rag1 mice, enable xenotransplantation without immune suppression, so that assessment of human cells in an in-vivo environment is possible.

Clinical therapies for human diseases will require that the cells that are delivered to the diseased individual are: 1) autologous, i.e., isolated from that same individual; 2) delivered along with immune suppressants to prevent rejection; 3) enclosed within a barrier that prevents immune-mediated destruction while retaining full function; or 4) delivered along with a regimen aimed at providing immune tolerance to those cells. The down side of immune suppressive regimens as they are given today is the ultimate complications related to those drugs, which include unpleasant daily side effects and the greatly increased risk of developing cancers such as lymphomas. The development of situation-specific immune suppressants, as opposed to the broad-spectrum steroidal approach to immune suppression, has led to fewer side effects and risks.^[33] The ideal solution is to deliver cells that were isolated from the diseased individual or a living related donor, so that little or no immune suppression is required. However, this strategy may not be effective in diseases with an autoimmune component, such as diabetes, where autoantibodies involved in the disease mechanism may continue to destroy any insulin-producing cells that are transplanted. Coating or encapsulation of transplanted cells in semipermeable membranes is an option that is being explored for applications in diabetes; however, many encapsulation strategies have been abandoned due to poor performance and reproducibility.^[34]

Another question arises as we consider the implications of transplanting stem cells for disease therapy—how will these cells function in vivo after transplantation? There are many reports of stem cells that are able to participate in regeneration after tissue damage, as well as reports of stem cells that can be pushed to differentiate into a specific phenotype through modulation of culture conditions in vitro. However, the phenotypic characteristics of stem cells in vivo are modulated by their microenvironment, as discussed above. The site of delivery as well as accessibility of the cells to the vascular system will influence their ability to perform the desired function in vivo. In many cases, it is not feasible for the therapeutic cells to be delivered directly to the site of injury. For example, islet cells transplanted for the purpose of type I diabetes therapy are delivered to the liver, not the

pancreas.^[35] Although many animal models involve delivery of stem cells to the bone marrow, homing of the transplanted cells from the bone marrow or peripheral blood to the site of tissue injury may not be sufficient to provide a therapeutic benefit.^[10] The ability of therapeutic cells to achieve and maintain a functional state of differentiation will depend on the consistent presence of the environmental signals that modulate cellular phenotype. In the growing fields of stem cells and biomaterials, the challenge may be to develop engineered biomaterials that support tissue-specific differentiation of stem cells and provide the structure and signals necessary to maintain the desired phenotype in vitro and in vivo.

S

CONCLUSION

The science of stem cells has captured the interest of many in the past decade. New stem cells and markers are being discovered throughout the world across a variety of scientific disciplines. The potential value of stem cells in therapeutics has brought them to the forefront of research in regenerative medicine, with active research in a vast array of disease applications, including diabetes, neurodegenerative disorders, cardiovascular disorders, and genetic diseases. As demonstrated by the discussion presented in this article, many of the dogmatic assumptions regarding stem cells have fallen away in recent years, giving rise to an emerging theory that stem cells can exist in many adult tissues and are characterized by varying degrees of phenotypic plasticity. The field of regenerative medicine stands to benefit considerably as we gain knowledge related to the identification, isolation, propagation, differentiation, and in-vivo delivery of stem cells.

ARTICLES OF FURTHER INTEREST

Biomedical Ethics; Cryopreservation of Living Cells; Hematopoietic Stem Cells and Assays; Histogenesis; Osteogenic Progenitor Cells of Bone

REFERENCES

1. Petersen, B.E.; Bowen, W.C.; Patrene, K.D.; Mars, W.M.; Sullivan, A.K.; Murase, N.; Boggs, S.S.; Greenberger, J.S.; Goff, J.P. Bone marrow as a potential source of hepatic oval cells. *Science* **1999**, *284* (5417), 1168–1170.
2. Orlic, D.; Hill, J.M.; Arai, A.E. Stem cells for myocardial regeneration. *Circ. Res.* **2002**, *91* (12), 1092–1102.

3. LaBarge, M.A.; Blau, H.M. Biological progression from adult bone marrow to mononucleate muscle stem cell to multinucleate muscle fiber in response to injury. *Cell* **2002**, *111* (4), 589–601.
4. Theise, N.D.; Badve, S.; Saxena, R.; Henegariu, O.; Sell, S.; Crawford, J.M.; Krause, D.S. Derivation of hepatocytes from bone marrow cells in mice after radiation-induced myeloablation. *Hepatology* **2000**, *31* (1), 235–240.
5. Trounson, A. Human embryonic stem cells: Mother of all cell and tissue types. *Reprod. Biomed. Online* **2002**, *4* Suppl. 1, 58–63.
6. Department of Health and Human Services. *Stem Cells: Scientific Progress and Future Research Directions*; June 2001. www.nih.gov/news/stemcell/scireport.htm.
7. Grisham, J.W.; Coleman, W.B.; Smith, G.J. Isolation, culture, and transplantation of rat hepatocytic precursor (stem-like) cells. *Proc. Soc. Exp. Biol. Med.* **1993**, *204* (3), 270–279.
8. Coleman, W.B.; Smith, G.J.; Grisham, J.W. Development of dexamethasone-inducible tyrosine aminotransferase activity in WB-F344 rat liver epithelial stemlike cells cultured in the presence of sodium butyrate. *J. Cell. Physiol.* **1994**, *161* (3), 463–469.
9. Spees, J.L.; Olson, S.D.; Ylostalo, J.; Lynch, P.J.; Smith, J.; Perry, A.; Peister, A.; Wang, M.Y.; Prockop, D.J. Differentiation, cell fusion, and nuclear fusion during ex vivo repair of epithelium by human adult stem cells from bone marrow stroma. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (5), 2397–2402.
10. Wang, X.; Montini, E.; Al-Dhalimy, M.; Lagasse, E.; Finegold, M.; Grompe, M. Kinetics of liver repopulation after bone marrow transplantation. *Am. J. Pathol.* **2002**, *161* (2), 565–574.
11. Presnell, S.C.; Petersen, B.; Heidaran, M. Stem cells in adult tissues. *Semin. Cell Dev. Biol.* **2002**, *13* (5), 369–376.
12. Henningson, C.T., Jr.; Stanislaus, M.A.; Gewirtz, A.M. Embryonic and adult stem cell therapy. *J. Allergy Clin. Immunol.* **2003**, *111* (2 Suppl.), S745–S753.
13. Kirby, S.L.; Bentley, S.; Fry, J.; Walton, W.; Desai, N.; Latour, A.; Coleman, W.; Anderson, P.; Grisham, J.; Malouf, N. Hematopoietic cell transdifferentiation of cloned adult live-derived stem cells. *Proc. Am. Soc. Cell Biol.* **2002**, *A680*.
14. Boura, P.; Kountouras, J.; Lygidakis, N.J. Tumor immunity and immunotherapy. *Hepatogastroenterology* **2001**, *48* (40), 1040–1044.
15. Overturf, K.; al-Dhalimy, M.; Ou, C.N.; Finegold, M.; Grompe, M. Serial transplantation reveals the stem-cell-like regenerative potential of adult mouse hepatocytes. *Am. J. Pathol.* **1997**, *151* (5), 1273–1280.
16. Overturf, K.; Al-Dhalimy, M.; Finegold, M.; Grompe, M. The repopulation potential of hepatocyte populations differing in size and prior mitotic expansion. *Am. J. Pathol.* **1999**, *155* (6), 2135–2143.
17. Haralson, M.A.; Hassell, J.R. The Extracellular Matrix—An Overview. In *Extracellular Matrix—A Practical Approach*; Haralson, M.A., Hassell, J.R., Eds.; Oxford University Press: Oxford, 1995; 1–30.
18. Akiyama, S.K.; LaFlamme, S.E. Bioadhesion and cell behavior. *Colloids Surf., B: Biointerfaces* **1994**, *2*, 241–250.
19. Dexter, T.M. Haemopoiesis in long-term bone marrow cultures. A review. *Acta Haematol.* **1979**, *62* (5–6), 299–305.
20. Rosenzweig, M.; Pykett, M.; Marks, D.F.; Johnson, R.P. Enhanced maintenance and retroviral transduction of primitive hematopoietic progenitor cells using a novel three-dimensional culture system. *Gene Ther.* **1997**, *4*, 928–936.
21. Banu, N.; Rosenzweig, M.; Kim, H.; Bagley, J.; Pykett, M. Cytokine-augmented culture of haematopoietic progenitor cells in a novel three-dimensional cell growth matrix. *Cytokine* **2001**, *13* (6), 349–358.
22. Dominici, M.; Hofmann, T.J.; Horwitz, E.M. Bone marrow mesenchymal cells: Biological properties and clinical applications. *J. Biol. Regul. Homeost. Agents* **2001**, *15* (1), 28–37.
23. Martin, I.; Padera, R.; Vunjak-Novakovic, G.; Freed, L.E. In vitro differentiation of chick embryo bone marrow stromal cells into cartilaginous and bone-like tissues. *J. Orthop. Res.* **1998**, *16*, 181–189.
24. Martin, I.; Shastri, V.P.; Padera, R.F.; Yang, J.; Mackay, A.J.; Langer, R.; Vunjak-Novakovic, G.; Freed, L.E. Selective differentiation of mammalian bone marrow stromal cells cultured on three-dimensional polymer foams. *J. Biomed. Mater. Res.* **2001**, *55* (2), 229–235.
25. Young, R.G.; Butler, D.L.; Weber, W.; Caplan, A.I.; Gordon, S.L.; Fink, D.J. Use of mesenchymal stem cells in a collagen matrix for Achilles tendon repair. *J. Orthop. Res.* **1998**, *16*, 406–413.
26. Awad, H.A.; Butler, D.L.; Harris, M.T.; Ibrahim, R.E.; Wu, Y.; Young, R.G.; Kadiyala, S.; Boivin, G.P. In vitro characterization of mesenchymal stem cell-seeded collagen scaffolds for tendon repair: Effects of initial seeding density on contraction kinetics. *J. Biomed. Mater. Res.* **2000**, *51* (2), 233–240.
27. Teng, Y.D.; Lavik, E.B.; Qu, X.L.; Park, K.I.; Ourednik, J.; Zurakowski, D.; Langer, R.; Snyder, E.Y. Functional recovery following traumatic spinal cord injury mediated by a unique polymer scaffold seeded with neural stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (5), 3024–3029.
28. Park, K.I.; Teng, Y.D.; Snyder, E.Y. The injured brain interacts reciprocally with neural stem cells supported by scaffolds to reconstitute lost tissue. *Nat. Biotechnol.* **2002**, *20* (11), 1111–1117.
29. Katz, A.J.; Lull, R.; Hendrick, M.H.; Futrell, J.W. Emerging approaches to the tissue engineering of fat. *Clin. Plast. Surg.* **1999**, *26* (4), 587–603.
30. Kral, J.G.; Crandall, D.L. Development of a human adipocyte synthetic polymer scaffold. *Plast. Reconstr. Surg.* **1999**, *104*, 1732–1738.
31. Patrick, C.W.; Chauvin, P.P.; Hobbey, J.; Reece, G.P. Preadipocytes seeded PLGA scaffolds for adipose tissue engineering. *Tissue Eng.* **1999**, *5* (2), 139–151.
32. Hurd, D.D. Bone marrow transplantation for cancer: An overview. *Recent Results Cancer Res.* **1993**, *132*, 1–14.

33. Trofe, J.; Buell, J.F.; First, M.R.; Hanaway, M.J.; Beebe, T.M.; Woodle, E.S. The role of immunosuppression in lymphoma. *Recent Results Cancer Res.* **2002**, *159*, 55–66.
34. Orive, G.; Hernandez, R.M.; Gascon, A.R.; Calafiore, R.; Chang, T.M.; De Vos, P.; Hortelano, G.; Hunkeler, D.; Lacik, I.; Shapiro, A.M.; Pedraz, J.L. Cell encapsulation: Promise and progress. *Nat. Med.* **2003**, *9* (1), 104–107.
35. Ryan, E.A.; Lakey, J.R.; Paty, B.W.; Imes, S.; Korbitt, G.S.; Kneteman, N.M.; Bigam, D.; Rajotte, R.V.; Shapiro, A.M. Successful islet transplantation: Continued insulin reserve provides long-term glycemic control. *Diabetes* **2002**, *51* (7), 2148–2157.

S

Stent Grafts, Endovascular

Timothy A.M. Chuter

Division of Vascular Surgery, University of California San Francisco, San Francisco, California, U.S.A.

INTRODUCTION

As the name suggests, this type of endovascular prosthesis combines two elements, a stent and a graft. The graft conveys blood from one end of the prosthesis to the other, isolating the surrounding artery from the circulation. The stent framework supports the position, shape, and dimensions of the graft, at the same time generating a seal between the ends of the graft and the surrounding artery.

STENT-GRAFT APPLICATIONS

Endovascular repair abdominal aortic aneurysm (AAA) was the first,^[1] and remains the commonest, application of stent-graft technology. Once deployed, the stent-graft performs much like a surgical graft, bridging the dilated segment of distal aorta, or aneurysm, and excluding it from direct arterial flow. In theory, the excluded aneurysm is protected from arterial pressure, dilatation, and rupture. The surgical and endovascular methods of repair differ in the means of graft insertion and attachment. Whereas the surgical graft is sutured into the open aneurysm, the stent-graft is introduced and secured without opening the abdomen or interrupting aortic flow. As a result, endovascular aneurysm repair causes less physiological stress, leading to lower morbidity and quicker recovery.

Because the primary goal is to exclude the aneurysm from the circulation, continuing perfusion of the space outside the stent-graft (endoleak) is regarded as a failure (Fig. 1A–D). However, not all forms of endoleak have the same effect on aneurysm pressure,^[2] dilatation,^[3] and risk of rupture.^[4] High pressure, aneurysm pulsatility, aneurysm dilatation, and continuing risk of rupture are associated with direct endoleak, either at the implantation site (type I) or through a gap in the stent-graft (type III). Intermediate values are associated with indirect endoleak (type II) through collateral routes to lumbar or inferior mesenteric branches of the aneurysm.^[4] In addition, porous stent-grafts are prone to type IV endoleak, with diffuse leakage through the stent-graft, and endotension, with pressurization in the absence of endoleak.

Short-term success depends on atraumatic transfemoral insertion, accurate deployment, and secure

hemostatic attachment proximal and distal to the aneurysm. Unfortunately, many patients with aneurysms have narrow tortuous iliac arteries and short angulated implantation sites. Some stent-grafts, and their delivery systems, cope well with these forms of anatomic distortion; others do not. During the decade since the first clinical application of stent-grafts for aneurysm repair, successful features have been copied and unsuccessful features have been eliminated. With a few notable exceptions, such as the Trivascular device, stent-grafts have undergone a process of convergent evolution so that current systems have many features in common. Nevertheless, every element of stent-graft design offers a wide variety of options, each with important consequences for stent-graft performance. Examples include: stent-graft configuration, tubular or branched; stent-graft construction, one piece (unibody) or multi-component (modular); stent type, balloon-expanded or self-expanding; stent materials, stainless steel, Elgiloy, or Nitinol; stent-graft attachment, barbs or suprarenal stents; stent distribution, fully-stented or partially stented; and graft materials, PTFE or polyester.

STENT-GRAFT BRANCHING

The shape of the stent-graft is dictated by the location of the aneurysm or, more precisely, by the location of the nondilated implantation sites proximal and distal to the aneurysm. The goal in endovascular aneurysm repair is to exclude the aneurysm, at the same time maintaining downstream flow. When the aneurysm has important branches, the stent-graft must also have branches. The simplest example of this principle is bifurcated stent-graft repair of an AAA involving the aortic bifurcation. A more extensive aneurysm, which encroaches on the origins of the renal or internal iliac arteries, requires a more complicated multibranched form.

Unbranched Stent-Grafts

Abdominal aorta

The original AAA stent-grafts were unbranched,^[1] with a single lumen connecting the one proximal orifice to one distal orifice. The proximal orifice was implanted between the renal arteries and the aneurysm

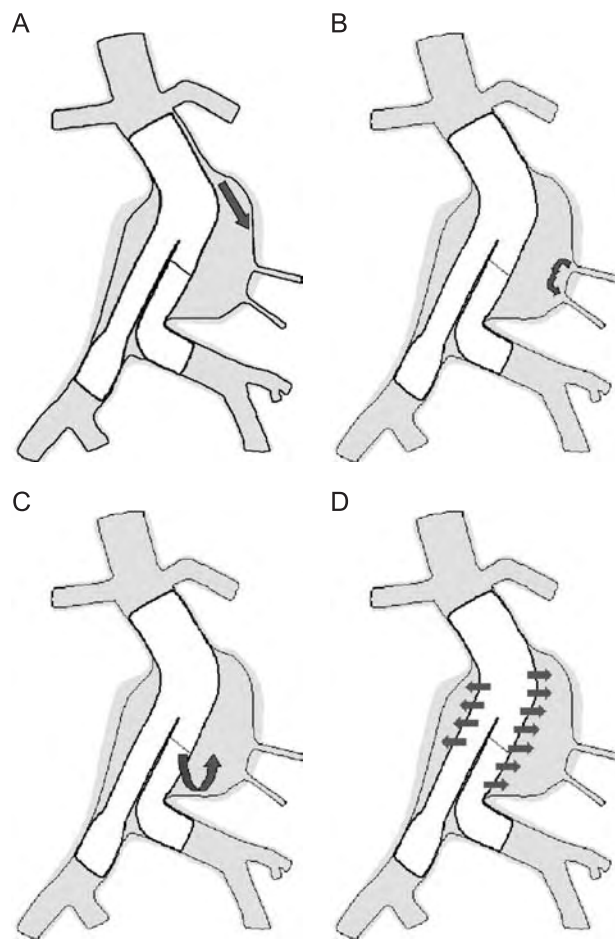


Fig. 1 A. Type I endoleak between the graft and an implantation site. B. Type II endoleak between two branches of the aneurysm. C. Type III endoleak between two components of the stent-graft. D. Type IV endoleak through many small holes in the stent-graft.

in a nondilated segment referred to as the neck, and the distal orifice was implanted between the aneurysm and the aortic bifurcation. The aorto-aortic approach proved to be of limited use on infrarenal aorta because the aneurysm typically involves the entire distal aorta.^[5] Most cases have absolutely no nondilated distal aortic implantation site. Attempts to use short, wide, conical, thrombus-lined implantation sites resulted in a high rate of detachment and leakage. Stent-grafts that were implanted in a cuff of thrombus often showed no signs of leakage into the aneurysm, and yet the aneurysms continued to enlarge. The probable explanation for this phenomenon of “endotension” was transmission of pressure through the rim of thrombus.

Thoracic aorta

The descending thoracic aorta is bounded proximally by the subclavian artery and distally by the celiac

artery. The lack of major branches makes this segment of the aorta a better location for unbranched stent-grafts.^[6] The endovascular repair of thoracic aortic aneurysms is a particularly appealing alternative to open surgery because the patients are often old, sick, and female. They tolerate thoracotomy and supraceliac interruption of aortic flow poorly. Consequently, surgical repair has high mortality and morbidity rates.

In cases of intramural hematoma, penetrating ulcer, and isolated saccular aneurysm, stent-graft implantation is quick, easy, and effective. However, typical thoracic aortic aneurysms are far more difficult to treat. The implantation sites are often short, wide, angulated, and widely separated. Large-diameter implantation sites require large-diameter stent-grafts, which in turn require large stiff delivery systems that may be difficult to insert. Small iliac arteries often preclude transfemoral insertion of such large sheaths, especially in women. The alternative is insertion into the common iliac artery or aorta, either directly or through a surgically-created conduit. Larger delivery systems also tend to be stiffer, leading to difficulty in cases of arterial tortuosity. Three problem areas are the iliac arteries, the aorta at the diaphragm, and the aortic arch. Stiff guidewires and trackable delivery systems are helpful in the iliac arteries, which can straighten, but of less use in the less mobile segments of the aorta. Angulated implantation sites make specific demands of stent-graft performance that are discussed in more detail later in the context of stent-graft attachment and flexibility.

One way to lengthen the proximal implantation site is to cover the subclavian artery. In most cases, the lack of direct prograde flow into the subclavian artery has no serious effects. In the presence of a left internal mammary graft to the coronary circulation, or a dominant left vertebral artery, bypass from the left carotid artery to the subclavian artery should precede endovascular stent-graft implantation. Otherwise, subclavian artery bypass will need to be performed later, if the endovascular repair causes left upper extremity ischemia.

It is not yet clear whether stent-grafts will have any role in the treatment of aortic dissection. Currently, the acute management of aortic dissection depends on the location of the dissection and the presence of ischemic complications. Type A dissections, proximal to the subclavian artery, require urgent surgical repair. Type B dissections, distal to the subclavian artery, can be managed medically with antihypertensive agents unless there are signs of end-organ ischemia, in which case stent-grafts have been used to cover an intimal tear in the proximal descending thoracic aorta and to direct blood into the true lumen.^[7] Reports of false lumen thrombosis and healed dissection suggest that this approach may have a role, but there are problems. For

one, there are often multiple intimal tears producing multiple communications between the true and false lumens. A stent-graft that covers them all is more likely to exclude flow through intercostal arteries to the anterior spinal cord, causing paraplegia.

Bifurcated Stent-Grafts

Although as many as 80% of open surgical AAA repairs can be performed using aorto-aortic tube grafts, the aorto-aortic stent-graft is rarely successful in the long term. The anatomic requirements for surgical anastomosis and endovascular stent-graft implantation are very different. Endovascular stent-grafts need more than a sewing ring for secure hemostatic implantation. These days, few manufacturers even produce unbifurcated stent-grafts for AAA repair.

The first bifurcated stent-grafts were inserted whole (unibody) and pulled into position using cross-femoral catheters.^[8] However, most current systems employ a different (modular) approach whereby the bifurcated stent-graft is assembled in situ from two or more components. The relative advantages and disadvantages of the modular and unibody approaches are discussed in detail below in the section on stent-graft construction.

Multibranch Stent-Grafts

Indispensable aortic branches in the arch, proximal abdomen, and pelvis have in the past defined the limits of endovascular aneurysm repair. Conventional endovascular repair of aneurysms that involve these areas requires prior surgical bypass into the affected branches.^[9] Alternatively, the stent-graft must be provided with branches to feed these arteries.^[10-13]

Like bifurcated stent-grafts, the more complicated multibranch versions divide into two broad groups, unibody and modular. Inoue et al.^[13] have used multibranch unibody stent-grafts to treat aneurysms of the aortic arch, thoracoabdominal aorta, and common iliac arteries. Their success is a reflection of enormous ingenuity and technical skill. In a unibody stent-graft, multiple branches require multiple control wires or catheters. There is an exponential increase in the degree of complexity with every additional branch. Cases with three or more branches have been few and the complication rates high. Others have limited the unibody approach to the provision of an internal iliac branch, but failure to implant the internal iliac branch was common in the presence of usual iliac tortuosity or internal iliac stenosis.

We have experimented with modular multibranch stent-grafts in the aortic arch, thoracoabdominal aorta, and iliac bifurcation. Our approach is to implant

an aortic component with short branches, which we then extend into the branch arteries using simple tubular bridging stent-grafts.

Thoracoabdominal aorta

In the thoracoabdominal aorta, the aortic component has one orifice proximally, four small side branches (one for each visceral artery), and one orifice distally for a standard infrarenal type reconstruction. All five branches point distally. The four side branches are catheterized in turn from above through the brachial artery. Once the catheter has been passed out through the branch of the stent-graft, it is directed into the corresponding branch of the aorta. We then exchange the catheter over a stiff guidewire for a guiding sheath and the delivery system for the side-branch extension. We have experimented with a variety of stent-grafts for use as the extension. The basic requirements are a low profile, a trackable delivery system, and a flexible stent-graft of predictable length. In our experience, the two best are the self-expanding PTFE-covered Smartstent (Cordis), no longer available, and the balloon-expanded PTFE-covered JoStent (JoMed). Both work well in the short-to-medium term, but the balloon-expanded JoStent is vulnerable to permanent kinking or compression where the celiac artery passes through the diaphragm.

Aortic arch

Our aortic arch system is not really multibranch, but bifurcated. We insert the stent-graft through the right carotid artery, leaving its trunk in the ascending aorta, its long narrow limb in the innominate (brachiocephalic) artery, and its short wide limb in the aorta. Flow through the innominate artery is distributed to the rest of the brachiocephalic circulation through a previously created carotid-carotid bypass and left subclavian-carotid transposition. A large-diameter tubular stent-graft is inserted transfemorally and used to extend the short wide limb into the descending thoracic aorta.

Iliac bifurcated stent-grafts

In cases of bilateral common iliac aneurysm, we preserve flow to an internal iliac artery by performing an endovascular reconstruction of the common iliac artery using a bifurcated stent-graft. We place the long limb of a bifurcated component into the ipsilateral external iliac artery, the short limb just above the internal iliac orifice, and the trunk in the proximal common iliac artery. A cross-femoral sheath is inserted

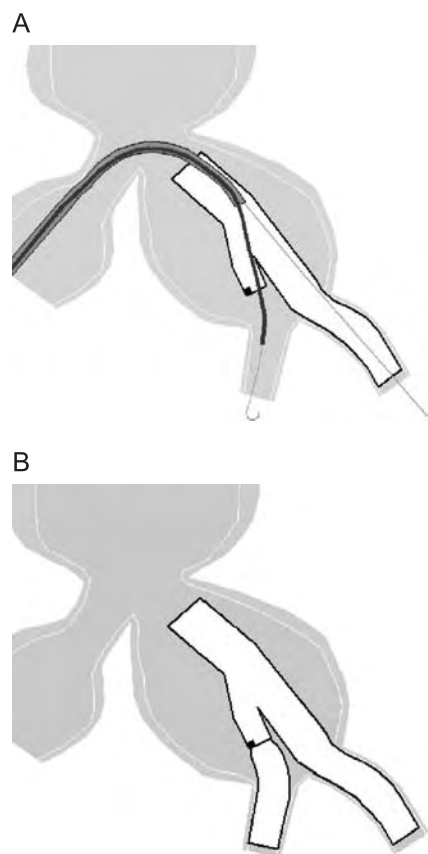


Fig. 2 A. Catheters, sheaths, and guidewires within the iliac bifurcation component. B. The iliac bifurcation component with its internal iliac extension in place.

from the contralateral side over a cross-femoral guidewire. A second puncture of the sheath valve provides access for insertion of a catheter and guidewire, which are directed through the short limb of the bifurcated component into the internal iliac artery. This bridging catheter is replaced over-the-wire for a bridging sheath, followed by a bridging stent-graft of the same type used in thoracoabdominal reconstruction, only a little wider (Fig. 2). The trunk of the bifurcated component then serves as the distal implantation for a conventional bifurcated reconstruction of the infrarenal aorta. The eventual result is a stent-graft that bifurcates once in the aorta and again in the common iliac artery. It has one proximal implantation site in the infrarenal neck and as many as four distal implantation sites in the internal and external iliac arteries.

Fenestrated Stent-Grafts

One solution to the problem of implantation site encroachment is to provide each vital branch of the

aneurysm with a branch of the stent-graft. Another is to perfuse each arterial branch through a hole, or fenestration, in the stent-graft. The danger is that blood will also flow through the fenestration into the aneurysm. One solution is to pin the orifice of the stent-graft fenestration to the orifice of the branch artery using a flared bridging stent.^[14] This has the added benefit of fixing the fenestration in position so that neither malposition nor migration ever interrupts branch artery perfusion. Although fenestrated stent-grafts require precise sizing, manufacture, and deployment, the approach is now practiced at 35 centers in Australia and Europe. The modified Zenith device is manufactured by Cook, Inc., Australia, according to the measurements entered on a standard order form.

S

MODULARITY

Each branch point in a stent-graft represents the intersection of two separate lines of insertion. In a unibody stent-graft, the device is deposited along one line of insertion and its branches pulled into the other line of insertion by a system of catheters or wires.^[8] This translocation cannot occur with the limb in its expanded state, so there also has to be a mechanism to maintain the limb in a compressed state until it reaches the target artery. In a modular stent-graft, separate components take different lines of insertion. The point of intersection is the point at which these components join one another to form the final composite stent-graft.

Because the unibody stent-graft has no connections, there is no possibility of component separation or intercomponent leakage, as there is with the modular stent-graft. In all other regards, the modular stent-graft is superior to its unibody equivalent. Not needing to position catheters makes the delivery system narrower and simpler with less to pack and less to fail. Modular systems are more versatile and more flexible. Each component has a range of lengths and widths, so that every additional component adds two dimensions to the matrix of possible combinations. For example, a stent-graft with three components, each having five possible lengths and five possible diameters has 1565 possible combinations. A unibody stent-graft could achieve such a wide range of configurations only if it were custom-made. Besides, variability in the length of overlap between components provides another opportunity to adjust the dimensions of a modular stent-graft, which is particularly advantageous because it allows in situ adjustment.

The modular approach facilitates the development of new stent-graft configurations and applications. The new stent-graft can be constructed by modifying or adding a single component. This keeps to a minimum

the amount of new untested technology and reduces the degree of uncertainty regarding its performance. In this way, the design of device can advance through a series of small relatively predictable steps. The corresponding advances in implantation technique proceed in a similar way.

These advantages account for the overall ascendancy of the modular technique. The first unibody bifurcated stent-grafts were developed in 1991 and used clinically in 1993.^[8] Although this device was abandoned in 1996, a similar unibody stent-graft, the Ancure (Guidant), was manufactured until 2003.^[15] The Ancure stent-graft was stable and effective. However, a narrow range of sizes and a potential for device malfunction during implantation limited its application, especially outside the United States where more versatile alternatives were readily available. In early 1994, Lawrence-Brown and Hartley began a series of in vitro experiments with various bifurcated stent-grafts of their own design. Problems with the unibody bifurcated approach prompted them to adopt a modular system. The first clinical implantation of this bifurcated modular stent-graft followed in August 1994. Their device, the Zenith (Cook, Inc.), has continued to evolve ever since to become the most widely used stent-graft in the world, with multiple versions for increased flexibility, additional branches, and fenestrations.

New Designs

Despite its many disadvantages, the unibody approach still has many adherents who point to high rates of component separation, secondary type III endoleak, and aneurysm rupture with the use of modular devices, the Vanguard (Boston Scientific) in particular.^[16] This group, most of whom are surgeons, regard the potential for component separation as a fundamental flaw of the modular approach, but that is not the case. Inter-component connections are not necessarily unstable. Even a simple friction joint can be made more stable simply by increasing the length of the overlap zone. This design feature has the added benefit of increasing the range of in situ length adjustment through a variable degree of tromboning one component inside the other. Furthermore, there are many ways of making the junction even more secure using barbs, abrupt changes in stent-graft profile, or combinations of interlocking stents, plastic rings, and cuffs.

GRAFT FABRIC

The evolutionary history of the stent-graft is apparent in the choice of materials for the conduit or graft. The

two current alternatives, PTFE and polyester, are the same materials used in conventional surgical grafts, and their characteristics reflect the desired goals of surgical graft implantation. Both are porous enough to allow connective tissue ingrowth and impervious enough to prevent bleeding. Endovascular applications probably have different requirements for optimal performance. For example, the lining of an aneurysm has no tissue that would ever propagate into the interstices of the graft, nor does it require immediate hemostatic sealing. The long-term goal is to exclude the aneurysm from pressure, but there is no risk of trans-graft bleeding because the space around the graft is already bounded by the intact walls of an artery, however dilated and weakened by the lytic processes of aneurysm formation.

Polyester

The most widely used stent-graft (Zenith) is constructed from conventional woven polyester surgical grafts, the primary design objective being durability. Most other stent-grafts use thinner fabrics in an attempt to minimize delivery system size. The reduction in wall thickness reflects a smaller fiber diameter (Dernier size) and a tight weave to limit porosity.

PTFE

The PTFE fabrics used in stent-graft construction all have the familiar node/fibril structure of conventional grafts. However, the relative dimensions of these features vary widely, as does the overall thickness of the membrane. Thin-walled PTFE differs from woven fabric in several important regards.

PTFE tends to have less suture retention. Sutured attachment points between the stent framework and graft need to be at the margins of the fabric, or other places where a hole will not impair the barrier function of the graft. Alternatively, the stent-graft attachment needs to be suture-free. One approach is to sandwich the stent between adherent layers of PTFE. The Excluder (WL Gore), for example, traps a sinusoid Nitinol stent between the primary graft and an outer wrap of PTFE ribbon, whereas the Fluency (Bard, Inc.) has inner and outer tubes of PTFE. The opposite approach is employed in the JoMed stent-graft with inner and outer stents separated by a single tube of PTFE.

PTFE has a low friction surface. A PTFE-covered stent-graft cannot rely on frictional forces to resist migration. This point was illustrated well by the original Endologix (Bard) stent-graft, which migrated immediately, and only stopped moving when its bifurcation straddled the aortic bifurcation; hence, the

need for an extension cuff in a large proportion of cases.

Porosity

The optimal porosity of an endovascular fabric is unknown. In vitro and animal models differ from clinical circumstances in important ways, rendering experimental findings of uncertain predictive value. Experimental work has shown, however, that short wide thrombus channels between the aortic lumen and the aneurysm transmit pressure, whereas long thin ones do not.

The most porous stent-graft ever put to clinical use was a Strecker stent that was co-knit with polyester. This proved to be an effective barrier to flow in experimental and clinical studies, yet it had no effect on aneurysm pressure or dilatation. The thrombus seal eliminated endoleak, but permitted endotension (see Stent-Graft Technology). Among current stent-grafts, the highest rates of endotension are seen with AneuRx and Excluder.^[17] Both grafts are porous, although in different ways.

The AneuRx stent-graft has many thousands of suture holes where the stents are attached to the outer surface of the graft, disrupting the weave pattern. A new category of endoleak (type IV) was created to describe the diffuse leakage of contrast through the walls of this graft at the time of implantation. Although type IV endoleaks tend to resolve as these holes fill with thrombus, the holes do not disappear. Indeed, they tend to get larger due to suture-induced fiber separation, especially with the loose weave pattern of the predicate design.^[18] The occasional result was a recurrence of diffuse endoleak that looked like a water sprinkler at the time of conversion. More often there was no leakage, yet aneurysms treated with the AneuRx device shrink at a lower rate than with other stent-grafts, and some even dilate.^[19]

The Excluder stent-graft has no suture holes and the pores in its fabric are very small. Nevertheless, aneurysm shrinkage rates are low and dilatation rates high. In my opinion, this graft sometimes behaves as a kind of semipermeable membrane, allowing the aneurysm to distend with transudate. A similar phenomenon is known to occur following surgical implantation of PTFE grafts, whereby clear fluid collects around the graft.

Erosion

The barrier and attachment functions of the stent-graft require different mechanical properties, hence, the combination of a metal stent with a fabric graft. However, the combination of two such different

components is a potential source of instability in the long term. Every point of contact between the two has to be secured to prevent damage to the graft. Most current designs do so using multiple sutures, yet the sutures themselves can disrupt the structure of the graft. The Vanguard experience provided the most extreme example of stent-induced fabric erosion.^[16] The polypropylene sutures holding the stent framework in proper alignment broke, allowing stent apices into contact with the flimsy polyester fabric. This is a failure mode not seen with any of the current systems, but fabric erosion is far from rare. For example, explanted AneuRx stent-grafts have more than two fabric holes per device.^[18] These holes are rarely large enough to cause endoleak, but they may be large enough (1 mm) to pressurize the aneurysm sac.

New Designs

The graft component of a typical stent-graft serves two distinct functions: it isolates the aneurysm from the circulation and it connects the stents. These functions demand different properties. The barrier function requires that the graft be impervious to blood flow and pressure, whereas the structural function requires the strength and durability to withstand large hemodynamic forces, continual movement, and a compliance mismatch. Current stent-grafts combine these functions in one fabric tube, which must be both strong and impervious. An alternative is to divide the functions of the graft between two layers of graft material, each with properties tailored to optimize its role. The outer (structural) layer would serve as a foundation garment for the stents, just as a soldier's jacket holds the plates of body armor in place. This layer need not be impervious to blood and it need not be constructed from the polyester or PTFE of traditional graft material. The inner layer, freed from the need to support the stents, could be very thin-walled. The two layers would come together at the implantation sites to provide optimal attachment and sealing with or without additional stents.

A quite different approach involves integrating the stent and graft into a more homogeneous structure. The Wallgraft and Corvita stent-grafts (both from Boston Scientific) exemplify this approach. They suffer from the undesirable characteristics of the braided stent, such as insecure graft/artery attachment, weak expansion, and unpredictable shortening, but they do show the advantageous effects of integrated structure on stent-graft durability.

STENTS

The two basic types of stent are self-expanding, which opens through elastic deformation toward an

expanded “resting” diameter, and balloon-expanded, which undergoes plastic deformation to an expanded state under the action of an intraluminal balloon. Both types have been incorporated into stent-grafts, although the self-expanding type predominates. Differences between balloon-expanded and self-expanding stents have significant effects on the way they function in endovascular aneurysm repair.

Balloon-Expanded Stents

Balloon-driven stent expansion stretches both the stent and the surrounding artery. When the balloon deflates, the artery contracts onto the relatively unyielding stent, generating high frictional forces and encouraging stent incorporation into the arterial wall. Once expanded, the balloon-expanded stent has stable dimensions, so it reinforces the structure of the artery and limits dilatation. Since the mechanisms of attachment (friction and incorporation) both depend on arterial stretching, precise sizing of the aorta and balloon are important. Too little oversizing causes insecure attachment and leakage; too much causes the aorta to tear.

Self-Expanding Stents

At the time of implantation, self-expanding stents open to fill all the available space, and may continue to expand with the surrounding artery after implantation. Assuming the graft is oversized, the stent pushes continually outwards against the artery. The forces involved are low compared to the forces generated by arterial pressure. It is unlikely that stent expansion would cause implantation site dilatation, but it is equally unlikely that the self-expanding stent will resist the natural tendency toward arterial dilatation the way a balloon-expanded stent does. Since high degrees of oversizing rarely cause leakage, the exact dimensions of the stent-graft are less critical. An association between oversizing and neck dilatation migration would suggest otherwise. However, it is far from clear that oversizing causes migration or neck dilatation. Experienced users apply high degrees of oversizing to accommodate irregular or thrombus-lined necks, both of which are likely causes of implantation site instability. Less than 20% oversizing is associated with an increased risk of implantation site (type I) endoleak. For this reason, most patients are treated with stent-grafts 28 mm or more in diameter.

Stent Breakage

Stent materials include Nitinol, 300 series stainless steel, and Elgiloy. Late-occurring stent fracture has

been reported with all current stent-grafts. However, some are more prone to breakage than others.

The early Nitinol stents were a particular problem.^[20] Nitinol is an attractive option in the manufacture of self-expanding stents. Its shape memory properties allow the creation of complex forms from Nitinol wire or laser-etched tubes of Nitinol. Much has been made of the temperature-dependent transition of Nitinol from plastic to superelastic, but this phenomenon probably plays a more important role in intellectual property lawsuits than in stent-graft performance. While the low-temperature plasticity certainly helps in loading, subsequent stent expansion has little to do with temperature change. The stent-graft warms to blood temperature well before it is released from the delivery system unless iced saline is infused constantly through a side port, which it never is.

The surface of Nitinol passivates by forming a layer of titanium oxide, which gives raw Nitinol its characteristic black, or dark gray, appearance. On a microscopic level, raw unpolished Nitinol has many surface irregularities, each of which becomes the nidus for a stress fracture or erosive pit when the stent is exposed to the pulsatile movements and corrosive fluids of the arterial circulation. All of the early Nitinol-based stent-grafts suffered from this problem. Examples include Vanguard,^[16] AneuRx,^[18] and Talent.^[21] The Vanguard device had many other problems and was abandoned. The other two are still available, but their stents have undergone a number of changes, including better electropolishing. Time will tell whether these changes produce the desired improvement in durability.

Steel and Elgiloy stents are on the whole more durable than those of Nitinol, with a few notable exceptions. The original United States multicenter study of the Lifepath (Baxter) was stopped because each balloon-expanded wireform (Baxter’s term for a stent) was susceptible to multiple fractures, causing complete loss of proximal attachment. This study has now resumed following changes in stent manufacture. Rumor has it that scratches in the stent surface were causing stress fractures. In the early Zenith (Cook, Inc., Australia) experience, a high rate of fractures was seen at the distal end of the main body, just above the bifurcation.^[22] This finding agrees well with the observed movement of this stent on cine-fluoroscopy, and with an analysis of the hemodynamic forces acting at the bifurcation. The substitution of two short stents for the single long stent of the original design seems to have solved the problem.

Corrosion of welded and soldered joints represents another potential cause of late failure. Examination of explanted Zenith stent-grafts, for example, showed corrosion of all solder points. Accelerated durability

testing produced similar findings. However, the process is considered too slow to pose a threat to the structural integrity of the device.^[22] Engineers involved in stent-graft design have long avoided placing different alloys in close proximity for fear of Galvanic corrosion. On the other hand, most physicians have not been so scrupulous. Nitinol-based stent-grafts have been reinforced with stents of Elgiloy and steel, and stainless steel-based stent-grafts have been reinforced with stents of Elgiloy and Nitinol, all without a single reported case of stent fracture in a region of stent material admixture.

Future Designs

The ideal stent for use in a stent-graft has the following features: 1) high enough hoop strength to press the graft firmly against the walls of the implantation site, creating a hemostatic seal in inducing arterial ingrowth, but not so high that it accelerates implantation site dilatation; 2) high ratio of diameter to length, so that the stent-graft conforms well to noncylindrical implantation sites; 3) high expansion ratio for larger arteries and small delivery systems; 4) multiple anchor points for easy attachment of the fabric component and reduced potential for fabric erosion; 5) no sharp points that might impinge on the fabric, causing erosion; and 6) durability in the face of wide-ranging pulsatile movement.

None of the current designs meets all these criteria. A full discussion of potential solutions to all of the above is clearly beyond the scope of this article, but one fundamental (universal) problem deserves a special mention. Stents expand and contract by deformation of the angle between adjacent struts. The smaller this area is, the less durable the stent; the larger this area is, the more space the stent occupies when compressed. No such compromise limits deformation of the struts. Stents that expand by strut deformation are both durable and highly expansile. A strut of this type functions like a long bow or a carriage spring, with a wide range of movement at the ends, but minimal focal deformation at any point throughout its length.

FIXATION

Blood flow and pressure subject stent-grafts to caudally directed forces of up to 10N.^[23] The forces increase with increasing flow rate, proximal diameter, angulation, diameter change, and branching of the stent-graft. All of these features are common in aortic applications. Thoracic stent-grafts have the highest diameters, flow rate, and angulation, whereas aortic stent-grafts have a branch point with an abrupt

diameter reduction. The tendency toward caudal migration is resisted by a combination of barb-mediated anchoring, friction, arterial ingrowth, and column strength. Of these, only barb-mediated anchoring appears to provide secure long-term fixation.

Barbs

The importance of barb-mediated fixation is illustrated by comparing the two stent-grafts first approved for use in the United States, the Ancure^[15] and the AneuRx.^[18,24] The Ancure had no column strength to transmit support from the iliac implantation sites, no external stents to become incorporated, little radial expansion to promote friction, and little oversizing to accommodate implantation site diameter. Its only apparent means of attachment was a series of balloon-driven barbs projecting almost at right angles from its orifices, and yet large prospective multicenter studies showed the migration rate to be very low. The AneuRx, on the other hand, was (and still is) sheathed entirely in strong self-expanding external stents, which provided column strength by resisting both bending and longitudinal compression. It lacked only barbs, and yet many studies have shown migration rates that rise steadily from the time of implantation. At three years the migration rate is 20–30%, depending on study design.^[18,19,24]

Column Strength

Longitudinal connections between stents resist compression and bending, and allow the distal end of the stent-graft to support the position of the proximal end, a feature known as column strength. In practice, column strength contributes little to the stability of proximal stent position because high column strength deprives the stent-graft of the flexibility it needs to accommodate the tortuosity of the typical aneurysmal infrarenal aorta. In cases of neck angulation, the rigid stent-graft often fails to adopt the axis of the neck, causing leakage, or fails to remain within the neck, causing migration. The resulting type I endoleak causes aneurysm pressurization and rupture, as exemplified by high rates of rupture following repair with early versions of the AneuRx.

Length is another factor in the relationship between stent-graft flexibility and stability. This is particularly important in the graft limbs, which can vary in length according to stent-graft design. Short stiff limbs are more stable than long flexible limbs (Fig. 3). Most stent-grafts have long limbs and a short body,^[16] whereas others, such as the Zenith,^[22] have short limbs and a long body, a combination that has proven to be far more stable.

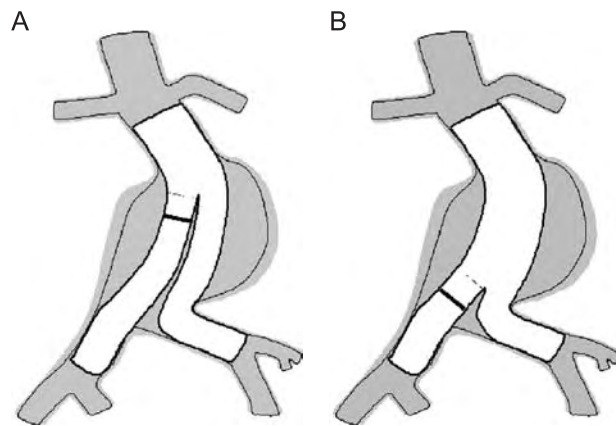


Fig. 3 A. A bifurcated stent-graft with a short trunk and long (unstable) limbs. B. A bifurcated stent-graft with a long trunk and short (stable) limbs.

Incorporation

Myointimal hyperplasia usually produces some incorporation of the uncovered, or external, proximal stent, especially when the stent is pressed forcibly into the surrounding artery. However, the effect on migration rate is uncertain. The best-incorporated of the self-expanding stents, in the AneuRx device, has high migration rates, whereas the least-incorporated, the Ancure, has low migration rates.

Suprarenal Stent

The adventitial layer of the aorta between the renal and the celiac arteries is richly endowed with elastin, which resists the dilating tendency that elsewhere leads to aneurysm formation. Consequently, this segment of the aorta is the most stable. Several stent-grafts take advantage of this by seating an uncovered stent in this area. The main concern of suprarenal stent attachment is the possible effect on renal function. However, mathematical modeling, animal experimentation, and clinical observation have all shown no significant effect.

Migration Rates

Although the Ancure is no longer manufactured, barb-mediated fixation is a prominent feature of several other devices, most notably the Zenith. While the Ancure barbs were driven into the aortic wall by a high-pressure balloon, the Zenith barbs deploy through caudal migration of the proximal stent at the time of, or shortly after, implantation. The inflation of a low-pressure compliant balloon may assist this process by transmitting hemodynamic forces through

the occlusive balloon to the stent-graft, but the caudally-directed barbs are not forced into the wall by the outward expansion of the balloon alone. In addition to the presence of 10–12 barbs, the Zenith is secured by a large uncovered suprarenal stent, robust stent expansion, and routine oversizing. Three-year data from the ASRNP (Australian Government) registry, four-year data from the Eurostar registry, and five-year data from UCSF all show very low rates of significant migration.

Most other stent-graft designs have migration rates somewhere between Zenith and AneuRx. Examples include: Talent, which has a suprarenal stent and significant column strength, but no barbs;^[21] and the Excluder,^[25] which has barbs, but no suprarenal stent. Other devices, such as the Cordis and Trivascular models, are too new for long-term data, although their features give some indication of expected behavior.

Endovascular Rescue of Stent-Graft Migration

The large numbers of AneuRx stent-grafts already implanted in the United States, the high rates of migration at long follow-up,^[18,25] and the serious consequences of migration are all causes for concern. The only remedy is early detection and re-intervention. On plain abdominal radiographs, migration is more apparent as a change in the shape and orientation of the stent-graft than a change in its position. Apparent movement of the stent-graft is often nothing more than an artifact caused by small changes in angle or magnification.^[26] On the other hand, bowing of the stent-graft usually means that one, or the other, end of the stent-graft is migrating into the aneurysm. Changes in aneurysm shape and length are far less likely causes.

Early detection of migration allows the opportunity for endovascular correction using an additional stent-graft to bridge the gap between the original stent-graft and its implantation site. The new stent-graft needs a better means of attachment than the original, or migration is likely to recur. Ideally, the new stent-graft should be implanted above the bifurcation of the original to maintain prograde flow to both lower extremities. This may not be possible when the original stent-graft was of a short-body/long-leg configuration, as is the case with AneuRx. Under these circumstances the safest approach may be aorto-uniiliac reconstruction with femorofemoral bypass and contralateral common iliac occlusion.

SEALING

Orificial stents are the sutures of the nonvascular device. They push the graft against the inner wall of the

implantation site, creating a seal. In this role, a stent on the inside performs better than a stent on the outside, and a short stent conforms better to a non-cylindrical implantation site better than a long one. Perhaps an even better seal is obtained through the hydrostatic action of an intramural balloon rather than a stent (Trivascular).

Many stent-grafts, such as Zenith and Talent, separate the sealing and attachment functions between two stents, one above the renal arteries uncovered by graft fabric, and the other within the graft fabric below the renal arteries. The Trivascular device takes this approach to an extreme, dispensing with the sealing stent altogether. The main advantage is staged delivery, whereby the graft itself contains no stents and occupies little space; hence, the low profile of the Trivascular delivery system.

COMPLETE STENT SUPPORT

The original concept of stent-grafts was a direct descendant of the traditional surgical approach with the stents taking the place of sutures. Indeed, the first Lazarus patent had no real stents, just a ring of staples. In the evolution of the EVT (later Ancure) device, these were connected together to form a rudimentary stent at each orifice of the graft, but there were no stents in between,^[27] nor were there any stents in the body of the original Parodi aorto-aortic device.^[1] More extensive stent support first appeared with the advent of modular bifurcated stent-grafts, which had to remain stable during re-instrumentation. This remains the primary role of the stent skeleton in fully-stented devices. Once the modular device is assembled, the stents provide radial support, helping to prevent kinks. They also provide longitudinal dimensional stability, especially in devices that rely on column strength, but the necessary rigidity can be a disadvantage in the presence of aortic tortuosity.

In general, flexibility and stable length are mutually exclusive. When a stent-graft bends, the outer surface lengthens relative to the inner surface in direct proportion to the stent-graft's diameter. The wider the stent-graft, the greater the length differential. If the stent-graft lacks the capacity for changing length, one of two things will happen. Either the stent-graft will resist bending, or its walls will come together in a kink. A kink resolves the disparity by eliminating, or greatly reducing stent-graft diameter in the plane of the bend.

Most stent-grafts are inserted in an extended state, and shorten as they are deployed inside tortuous arteries. It is important to remember this when selecting the stent-graft, especially in thoracic aortic applications. The required length corresponds to the dimensions of the outer curvature, not the center line.

The wider the artery, the greater the effect. In most cases, the necessary shortening is accomplished by crumpling the fabric between stents. The greatest flexibility is seen in stent-grafts, such as the Zenith Flex, with wide gaps and short stents. Alternatively, the stents themselves can overlap, as in the case of the Excluder stent-graft.

Although this capacity for shortening enhances flexibility, it makes the stent-graft unable to resist compressive loading. Such stent-grafts cannot simply be pushed from a sheath without collapsing the stent-graft, which increases the forces of deployment and imparts excessive redundancy. The original Stentor (later Vanguard) stent-grafts suffered from these problems.^[28] The deployed length was as much as 3 cm shorter than the nominal length. Subsequent lengthening in situ caused bowing and kinking of the graft limbs.

The Talent device incorporates a single longitudinal strut, which stabilizes the overall length of the device, while retaining flexibility. When the strut is on the outer aspect of the bend, the fabric on the inner curve crumples, the stents come together, and the stent-graft flexes. However, when the strut ends up on the inner aspect curve, there is no capacity for shortening. Furthermore, there is no capacity for lengthening on the outer aspect of the curve because the graft fabric is inelastic and there is no redundancy at the time of deployment. The frequent result is buckling of the longitudinal strut, which eventually breaks.^[21] In theory, the solution is always to implant the stent-graft with the strut on the outer aspect. In practice, this requirement just adds another potential source of failure.

Stents limit the flexibility of the stent-graft by fixing the longitudinal dimensions of opposing walls, but they also add radial support, which helps to prevent kinking. The unsupported stent-graft limbs of the Ancure had the highest rates of occlusion.^[29,30] The rates for fully-stented devices are generally less than 5%.

DELIVERY SYSTEM DIAMETER

Endovascular aneurysm repair is commonly referred to as percutaneous, which is not strictly true. Current stent-graft delivery systems make a large hole in the femoral artery, which has to be repaired when the delivery system is removed. The simplest, most widely-practiced method of repair involves surgical exposure and suture of the arterial access site. Alternatively, sutures can be inserted percutaneously at the access site before delivery system insertion using a special system of catheters and needles. At the end of the procedure, the sutures are tied with a sliding knot, closing the

defect in the femoral artery or the overlying femoral sheath.^[31] The expense and potential for failure have limited use of the percutaneous suturing system. Consequently, most stent-grafts are inserted through surgically exposed femoral arteries. That may change if smaller-diameter systems, such as the Trivascular device, prove to be effective.

CONCLUSION

Much of what we know about stent-graft performance is the result of a vast human experiment. Stent-graft design has benefited greatly from the painful lessons of the early experience. We hope that future advances will rely more on improved preclinical testing, based on a better understanding of materials science and the long-term demands of the endovascular environment. Yet fundamental assumptions in the modeling and testing of device durability remain poorly founded. For example, we still know very little about the pressure changes within the aneurysm sac after endovascular repair, or their effects on the trans-graft pressure, hemodynamic force, and pulsatile movement of a stent-graft.

ARTICLES OF FURTHER INTEREST

Controlled Release; Polymers; Stents

REFERENCES

1. Parodi, J.C.; Palmaz, J.C.; Barone, H.D. Transfemoral intraluminal graft implantation for abdominal aortic aneurysms. *Ann. Vasc. Surg.* **1991**, *5*, 491–499.
2. Sonesson, B.; Dias, N.; Malina, M.; Olofsson, P.; Griffin, D.; Lindblad, B.; Ivancev, K. Intra-aneurysm pressure measurements in successfully excluded abdominal aortic aneurysm after endovascular repair. *J. Vasc. Surg.* **Apr 2003**, *37* (4), 733–738.
3. Resch, T.; Ivancev, K.; Lindh, M.; Nyman, U.; Brunkwall, J.; Malina, M.; Lindblad, B. Persistent collateral perfusion of abdominal aortic aneurysm after endovascular repair does not lead to progressive change in aneurysm diameter. *J. Vasc. Surg.* **Aug 1998**, *28* (2), 242–249.
4. Harris, P.L.; Vallabhaneni, S.R.; Desgranges, P.; Becquemin, J.P.; van Marrewijk, C.; Laheij, R.J. Incidence and risk factors of late rupture, conversion, and death after endovascular repair of infrarenal aortic aneurysms: The EUROSTAR experience. European collaborators on stent/graft techniques for aortic aneurysm repair. *J. Vasc. Surg.* **2000**, *32* (4), 739–749.
5. Chuter, T.A.; Green, R.M.; Ouriel, K.; DeWeese, J.A. Infrarenal aortic aneurysm structure: Implications for transfemoral repair. *J. Vasc. Surg.* **Jul 1994**, *20* (1), 44–49. discussion 49–50.
6. Dake, M.D.; Miller, D.C.; Semba, C.P.; Mitchell, R.S.; Walker, P.J.; Liddell, R.P. Transluminal placement of endovascular stent-grafts for the treatment of descending thoracic aortic aneurysms. *N. Engl. J. Med.* **2001**, *331*, 1729–1734.
7. Dake, M.D.; Kato, N.; Mitchell, R.S.; Semba, C.P.; Razavi, M.K.; Shimono, T.; Hirano, T.; Takeda, K.; Yada, I.; Miller, D.C. Endovascular stent-graft placement for the treatment of acute aortic dissection. *N. Engl. J. Med.* **1999**, *340*, 1546–1552.
8. Chuter, T.A.; Risberg, B.; Hopkinson, B.R.; Wendt, G.; Scott, R.A.; Walker, P.J.; Viscomi, S.; White, G. Clinical experience with a bifurcated endovascular graft for abdominal aortic aneurysm repair. *J. Vasc. Surg.* **Oct 1996**, *24* (4), 655–666.
9. Criado, F.J.; Barnatan, M.F.; Rizk, Y.; Clark, N.S.; Wang, C.F. Technical strategies to expand stent-graft applicability in the aortic arch and proximal descending thoracic aorta. *J. Endovascular Ther.* **2002**, *9*, II32–II38.
10. Chuter, T.A.M.; Gordon, R.L.; Reilly, L.M.; Goodman, J.D.; Messina, L.M. An endovascular system for thoracoabdominal aortic aneurysm repair. *J. Endovasc. Ther.* **2001**, *8*, 25–33.
11. Schneider, D.B.; Curry, T.K.; Reilly, L.M.; Kang, J.W.; Messina, L.M.; Chuter, T.A.M. Branched endovascular repair of aortic arch aneurysm. *J. Vasc. Surg.* **2003**, *37*, 855.
12. Abraham, C.Z.; Reilly, L.M.; Schneider, D.B.; Dwyer, S.; Sawhney, R.; Messina, L.M.; Chuter, T.A. A modular multi-branched system for endovascular repair of bilateral common iliac artery aneurysms. *J. Endovascular Ther.* **2003**, *10*, 203–207.
13. Hosokawa, H.; Iwase, T.; Sato, M.; Yoshida, Y.; Ueno, K.; Tamaki, S.; Inoue, K. Successful endovascular repair of juxtarenal and suprarenal aortic aneurysms with a branched stent graft. *J. Vasc. Surg.* **May 2001**, *33* (5), 1087–1092.
14. Browne, T.F.; Hartley, D.; Purchas, S.; Rosenberg, M.; van Schie, G.; Lawrence-Brown, M. A fenestrated covered suprarenal aortic stent. *Eur. J. Vasc. Surg.* **1999**, *18*, 445–449.
15. Makaroun, M.S. The ancure endografting system: An update. *J. Vasc. Surg.* **2001**, *33* (Suppl. 2), S129–S134.
16. Becquemin, J.P.; Lapie, V.; Favre, J.P.; Rousseau, H.; for the French Vanguard Study Group. Midterm results of a second generation bifurcated endovascular graft for AAA repair. The French vanguard trial. *J. Vasc. Surg.* **1999**, *30*, 209–217.
17. van der Laan, M.; Prinssen, M.; Bertges, D.; Makaroun, M.S.; Blankensteijn, J.D. Does the type of endograft affect AAA volume change after endovascular aneurysm repair? *J. Endovasc. Ther.* **2003**, *10*, 406–410.
18. Zarins, C.K.; for the AneuRx Clinical Investigators. The US AneuRx clinical trial: 6-year clinical update 2002. *J. Vasc. Surg.* **2003**, *37*, 904–908.
19. Sternbergh, W.C., 3rd; Connors, M.S., 3rd; Tonnessen, B.H.; Carter, G.; Money, S.R. Aortic aneurysm sac shrinkage after endovascular repair is device-dependent:

- A comparison of Zenith and AneuRx endografts. *Ann. Vasc. Surg.* **2003 Jan**, *17* (1), 49–53.
20. Heintz, C.; Riepe, G.; Birken, L.; Kaiser, E.; Chakfe, N.; Morlock, M.; Delling, G.; Imig, H. Corroded nitinol wires in explanted aortic endografts: an important mechanism of failure? *J. Endovasc. Ther.* **2001**, *8*, 248–253.
 21. Criado, F.J.; Fairman, R.M.; Becker, G.J.; for Investigators of Talent LPS Pivotal Clinical Trial Talent LPS AAA stent-graft: Results of a pivotal clinical trial. *J. Vasc. Surg.* **2003**, *37*, 709–715.
 22. Cook Zenith P.M.A. *Panel Presentation*. April 10 2003. Gaithersberg, MD.
 23. Liffman, K.; Lawrence-Brown, M.M.; Semmens, J.B.; Bui, A.; Rudman, M.; Hartley, D.E. Analytical modeling and numerical simulation of forces in an endoluminal graft. *J. Endovasc. Ther.* **2001**, *8*, 358–371.
 24. Zarins, C.K.; Bloch, D.A.; Crabtree, T.; Matsumoto, A.H.; White, R.A.; Fogarty, T.J. Stent graft migration after endovascular aneurysm repair: Importance of proximal fixation. *J. Vasc. Surg.* **2003**, *38*, 1264–1272.
 25. Matsumura, J.S.; Katzen, B.T.; Hollier, L.H.; Dake, M.D. Update on the bifurcated EXCLUDER endoprosthesis: Phase I results. *J. Vasc. Surg.* **2001**, *33* (Suppl. 2), S135–S145.
 26. Hodgson, R.; McWilliams, R.G.; Simpson, A.; Gould, D.A.; Brennan, J.A.; Gilling-Smith, G.L.; Harris, P.L. Migration versus apparent migration: Importance of errors due to positioning variation in plain radiographic follow-up of aortic stent-grafts. *J. Endovasc. Ther.* **Oct 2003**, *10* (5), 902–910.
 27. Moore, W.S. The EVT tube and bifurcated endograft systems: Technical considerations and clinical summary. *J. Endovasc. Surg.* **1997**, *4*, 182–194.
 28. White, G.H.; May, J.; Waugh, R.; Harris, J.P.; Chaufour, X.; Yu, W.; Stephen, M.S. Shortening of endografts during deployment in endovascular AAA repair. *J. Endovasc. Surg.* **Feb 1999**, *6* (1), 4–10.
 29. Carpenter, J.P.; Neschis, D.G.; Fairman, R.M.; Barker, C.F.; Golden, M.A.; Velazquez, O.C.; Mitchell, M.E.; Baum, R.A. Failure of endovascular abdominal aortic aneurysm graft limbs. *J. Vasc. Surg.* **Feb 2001**, *33* (2), 296–302. discussion 302–3.
 30. Franssen, G.A.J.; Desgranges, P.; Laheij, R.J.F.; Haris, P.L.; Becquemin, J. Frequency, predictive factors, and consequences of stent-graft kink following endovascular AAA repair. *J. Endovasc. Ther.* **2003**, *10*, 913–918.
 31. Lee, do Y.; Kang, S.G.; Choi, D.; Lee, G.H.; Maeda, M.; Roh, B.S.; Won, J.W.; Kim, C.W.; Kim, E.S.; Song, H.Y. Percutaneous modular stent-grafts in the treatment of abdominal aortic aneurysms. *J. Endovasc. Ther.* **Aug 2003**, *10* (4), 752–759.

Stents

Benjamin Thierry

Department of Biomedical Engineering, McGill University, Montreal, Quebec, Canada

Luc Bilodeau

Montreal Heart Institute, Montreal, Quebec, Canada

Maryam Tabrizian

Department of Biomedical Engineering, McGill University, Montreal, Quebec, Canada

INTRODUCTION

Stents are metallic implantable tubular devices (Fig. 1) used as mechanical scaffolds to the vascular wall during revascularization procedures of coronary or peripheral arteries or lumen enlargement of hollow structure such as the tracheobronchial tree biliary duct and urinary tract. 316L stainless steel, nitinol or NiTi alloy, tantalum, and cobalt-chromium-based alloys are the materials most commonly used in the design of endovascular devices. The beneficial effect of coronary stenting over conventional balloon angioplasty has been demonstrated in the BELgian–Netherlands STENT (BENESTENT-I) and STent REStenosis Study (STRESS) studies in the early 1990s. In these landmark studies, restenosis, defined as the renarrowing of a vessel (definition angiographic: >50% diameter stenosis at 6 months follow-up), was significantly reduced. Since then, intensive research in the field of adjunctive pharmacological therapies, implantation techniques, and stent design have further reduced the clinical outcomes associated with stent implantation, namely subacute thrombosis and the need for target vessel revascularization (restenosis). Stenting is now a well established technique in endovascular interventions, representing as much as 70–85% of all percutaneous coronary interventions and about 50% of peripheral interventions, with about 1,000,000 stents being placed each year worldwide. Currently, more than 50 types of stents, and, more recently, stent-grafts with various designs (mesh structure, coil, slotted tube, ring, etc.) and compositions, have received FDA approval or CE marking. The worldwide market is estimated to be U.S. \$5 billion in 2002 for both coronary and peripheral interventions, with an annual growth rate of 5%. The recent availability of drug-eluting stents with perspectives of even greater restenosis reduction potential is likely to change the features of interventional cardiology.

HISTORICAL PERSPECTIVES

Charles Dotter introduced the idea of endovascular stents (Fig. 1) in 1964 and demonstrated in 1983 its

feasibility in the revascularization of diseased arteries. Stents were initially investigated to prevent recoil and cover intimal tears in order to prevent acute closure of balloon-revascularized arteries. In 1986, Jacques Puel performed the first human coronary stent implantation in Toulouse (France) using an ancestor of the self-expandable Wallstent. Since then, the BENESTENT and STRESS studies have shown the beneficial effects of elective coronary stenting following balloon angioplasty, versus angioplasty alone, in coronary revascularization (Table 1).^[1,2] The beneficial effect of stent over balloon angioplasty has been attributed to larger acute lumen dimensions and to the prevention of constrictive remodeling. The Palmaz–Schatz stent was the first FDA-approved device in 1994. Two major breakthroughs have clearly expanded the use of stents in percutaneous revascularization procedures and significantly improved the initial outcomes of the procedures: intravascular ultrasound (IVUS) and antiplatelet therapy. IVUS studies have demonstrated the importance of *optimal* stent deployment (symmetry, struts apposition, minimal area) and the need for high-pressure balloon postinflation to achieve proper stent expansion within the vessel.^[3] The optimal stent deployment concept suppressed the need for anti-vitamin K therapy (Warfarin) and in turn reduced the risk of bleeding. The benefit of antiplatelet therapy over traditional therapy, based on heparin and coumadin, along with aspirin and dipyridamole, namely a significant reduction in subacute thrombosis and bleeding, was demonstrated in the randomized ISAR and STARS trials (Table 1). Current revascularization procedures with stents are characterized by thrombotic occlusion rates lower than 2% and very high procedural success rates (>95%).

CLINICAL INDICATIONS

Stents reduce the risk of acute vessel closure and its consequence, the need for emergency coronary artery bypass graft surgery, and have been demonstrated to improve clinical outcomes in specific conditions. From

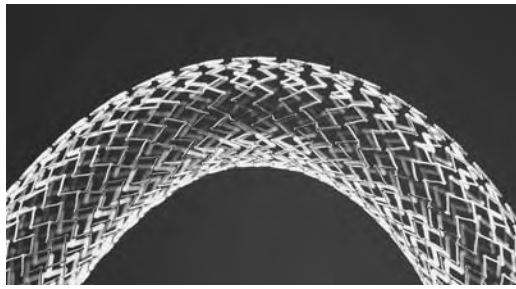


Fig. 1 Current stent designs and materials allow increased flexibility. Nitinol SMART stent from Nitinol Devices & Components.

randomized trials, those conditions are the treatment of abrupt or threatened vessel closure during angioplasty, primary reduction in restenosis in de novo focal lesions in vessels > 3 mm in diameter, focal lesions in saphenous vein grafts, and chronic total occlusions.^[4] Stents are also successfully used in many other conditions such as restenotic lesions, acute coronary syndromes (unstable angina and acute myocardial infarction), multivessel diseases, and the treatment of iliac and renal artery stenosis. However, their benefit in bifurcating lesions, in small vessel diameters (<3 mm), in diabetic patients, and in diseases of the femoral, popliteal, and carotid arteries is still being debated.

The use of stents is likely to increase due to the recent availability of devices designed to treat these specific conditions, such as small artery or vein graft disease, as well as the broadening of the available stent properties.

COMPLICATIONS

In spite of major technical advances in the past two decades, endovascular revascularization procedures remain limited by the cascade of events subsequent to device implantation.

S

Subacute Thrombosis

Early stent implantations were plagued by a high rate of thrombotic occlusions (8–10%) of the revascularized vessel, mainly due to suboptimal deployment of the devices and its inherent thrombogenicity. As previously mentioned, high-pressure deployment techniques combined with enhanced antiplatelet therapy have drastically reduced the risk of vessel thrombotic occlusion (1.7%). The thrombogenicity of the metallic alloys may still contribute to platelet activation, and in turn initiate neointimal hyperplasia.

Table 1 Selected randomized clinical trials

Trial		Composite events	Restenosis
BENESTENT	Balloon	AC = 2.7%	32%
	Stent	AC = 3.5%	22%
STRESS	Balloon	AC = 1.5%	42%
	Stent	AC = 3.4%	31%
ISAR	ASA + Triclo	D, MI, Rev = 1.6%	—
	ASA only	—	—
	ASA + Coumarin	D, MI, Rev = 6.2%	—
STARS	ASA + Triclo	D, MI, Rev = 0.6%	—
	ASA only	D, MI, Rev = 3.6%	—
	ASA + Coumarin	D, MI, Rev = 2.4%	—
SISA	Balloon	AC = 7.1%	33%
	Stent	AC = 3%	28%
SIRIUS	Stent	—	35%
	Sirolimus-coated stent	—	3%
TAXUS	Stent	AC = 0%	11%
	Taxol-coated stent	AC = 0%	0%

Abbreviations—BENESTENT: BELgian NETHERLANDS STENT; STRESS: STent REStenosis Study; ISAR: Intra-coronary Stenting and Antithrombotic Regimens; STARS: STent Anticoagulation Regimen Study; SISA: Stenting In Small Arteries; SIRIUS: SIRoIImUS; ASA: Aspirin; Triclo: Triclopidine; AC: Acute Closure; D: Death; Rev: Revascularization.

In-Stent Restenosis

Despite reduced restenosis rates in short lesions such as demonstrated in randomized trials, stenting is still associated with late in-stent restenosis. Based on clinical studies, restenosis rate can range from 8 to 50% depending on lesion complexity, reference vessel diameter plaque burden, and systemic conditions such as diabetes.^[5,6] In-stent restenosis is a multifactorial phenomenon that involves thrombus formation, inflammation, cellular proliferation, and extracellular matrix synthesis within the treated vessel.^[7] The cascade of events is initiated by arterial injury following angioplasty and is usually limited to the first 6 months following the procedure. Thrombus growth resulting from exposure of the extracellular matrix in the endothelium-denudated vessel wall and from the presence of the metallic device leads to the release of biological factors, including PDGF- β , TGF- β , bFGF, thrombin, etc., which in turn can initiate medial smooth muscle cell migration (SMC) and proliferation. Medial injury with disruption of the internal elastic lamina also directly stimulates SMC migration and proliferation. This cell growth is associated with extracellular matrix secretion that further participates in the neointimal thickening. The presence of inflammatory cells has been reported in addition to SMCs in the surrounding tissues.^[8] A strong correlation between the extent of medial damage, inflammation, and restenosis has been reported from human pathological studies.^[7] There is growing evidence that neointimal hyperplasia following stent implantation is strongly related to a prolonged inflammatory state of the stented artery.

Issues such as stent embolization, side branch occlusion, coronary perforation, or infection of the stented vessels are infrequent complications but could have potentially catastrophic effects.

METALLIC ALLOYS USED IN ENDOVASCULAR PROCEDURES

While many efforts have focused on the development of devices with improved design, biofunctionality, and surface properties, only a few alloys are routinely employed for stent manufacturing. They are generally made from one of three alloy classes: 316L stainless steel, nitinol (or NiTi alloy), and, to a lesser extent, cobalt–chromium alloys such as Elgiloy[®].

The differences in the mechanical properties of these materials strongly influence the properties of the devices and thus possibilities in terms of design. Stainless steel–based stents are balloon-expandable devices that require a plastic deformation to be deployed in the treated artery. Nitinol is gaining popularity in the

design of multiple permanent vascular implants, such as coronary and peripheral stents, atrial septal defect occlusion systems, and vena cava filters, due to its unique mechanical properties. These devices can be designed as self-expandable due to the superelasticity of nitinol. Tantalum was also used in the manufacture of coronary stents such as the Wiktor-GX but was nearly abandoned due to its exaggerated radiopacity. Cobalt–chromium alloy stents are currently under investigation (Multi-link Vision, Guidant and S8, Medtronic AVE). These alloys enable thinner strut thickness compared with stainless steel–based devices without reducing radial strength and radiopacity. They may also be less prone to biocorrosion than stainless steel devices.

New materials are being investigated to improve mechanical properties, radiopacity, MRI compatibility, and hemocompatibility.^[9,10] Differences in RX radiopacity and magnetic resonance compatibility inherent in the materials themselves translate clinically into differences in the visibility of the devices. Radiopaque markers, mainly gold and platinum, may be used to improve the visibility of these devices (e.g., beStent with proximal and distal gold markers).

BIOMECHANICAL CONSIDERATIONS

“The greater the lumen the better” concept that has prevailed based on minimal lumen diameter as the strongest predictive factor for restenosis, has led to the idea that all stents were equals in terms of clinical outcomes. As a consequence, stents were evaluated on their potential to achieve maximal lumen gain. There is now evidence that stent design also strongly influences the outcome of the revascularization procedure. Recent studies have indeed pointed out the dependency of thrombus formation and neointimal proliferation on the type and design of the stent.^[11–13] Follow-up angiography in a cohort of 3370 patients has shown that stent design was the second strongest risk factor, after artery size, with restenosis ranging from 20.0% to 50.3%, depending on the stent type.^[13] It seems that reduced strut thickness represents the crucial factor for restenosis prevention. Much work remains, however, to determine precisely which features of the stent are detrimental to implantation.

The ideal stent should have a low profile, good trackability, and flexibility with a high and well distributed radial strength for optimal scaffolding (Table 2). The delivery system should be highly reliable and allow precise placement of the device. The improvements in both delivery system and device properties have drastically increased the conditions wherein stents can be used, e.g., the availability of a stent designed specifically for small arteries. The beStent (4-crown

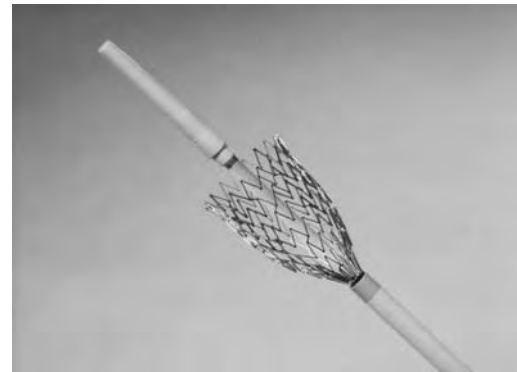
Table 2 Characteristic of the ideal stent

Low profile
Reliable delivery system
Trackability
Elasticity
Flexibility
Conformity
Uniform radial strength
Availability for various lesions
Thromboresistance
Biocompatibility
Corrosion resistance
Radiopacity

from Medtronic AVE) was compared in randomized studies to balloon angioplasty and proved to be superior in the treatment of <3 mm vessels. In addressing a wide variety of vessels and indications, the concept of an ideal stent has been evolving toward the idea that a different stent would best suit each indication. Nonetheless, it is likely that the choice of the best device will continue to rely on the clinician's expertise on a case-to-case basis.

The properties of balloon-expandable stents are mainly related to their design. Three designs are widely used, i.e., slotted tubes, corrugated rings and, to a lesser extent, coils. Slotted tube stents are the most widely used devices (beStent 2, Multilink Tetra and Zeta, BiodivYsio, etc.) and are characterized by good radial strength but less conformability compared with ring stents (AVE S670-S7, NIR, etc.), which are made of repeating cells and usually display good flexibility and are more suitable for the treatment of lesions requiring good conformability. New slotted tube design devices with thinner stent struts, such as the PURA AS, the Sorin Carbostent, or the 0.08-mm Tenax, could also perform well in such applications.

Nitinol-based devices rely on the mechanical properties of the alloy to be self-expanded within the vessel (Fig. 2). Comprehension of the material's characteristics is necessary in order to understand the device's properties. Self-expansion has been reported to reduce endothelium damage^[14] by minimizing the need for balloon postdilatation and allowing more uniform radial expansion along with less longitudinal shortening, in comparison with balloon-expandable devices.^[15] The effects of the continual expansion of oversized nitinol stents (10–15%) have been demonstrated to accommodate neointimal formation, but this remains controversial.^[16] Conversely, the over-expansion required to compensate for the acute recoil of balloon stents can be expected to increase injury to the vessel. Therefore, selection of a properly sized stent will influence its performance. As defined by D. Stoeckel, the

**Fig. 2** Deployment of a self-expandable stent.

crush-recoverability and the biased stiffness (or stress hysteresis) of NiTi alloys are particularly well adapted to the design of peripheral endovascular devices.^[17] Despite higher radial strength, stainless steel devices can be permanently deformed if outside forces exceed their buckling resistance. Nitinol devices, however, will return to their deployed shape even if completely crushed, which may be extremely important in superficial vessels such as the carotid artery.

SURFACE TREATMENTS

The surface properties of an endovascular device may strongly influence the outcome of a procedure. As stated in the FDA's cardiology guidance documents, particular attention should be paid to the surface treatment of endovascular devices.^[18] Surface properties such as roughness, surface energy, and chemistry are known to mediate the compatibility of blood-contacting surfaces. We have previously shown that, independent of other surface characteristics, the material itself affects the hemocompatibility of the device.^[19]

In addition, metallic materials used in endovascular applications are passive alloys or metals that impart their corrosion resistance and surface properties from surface oxide layers. In the presence of oxygen, oxides spontaneously grow on the alloy surface, e.g., Cr₂O₃ for stainless steel and Elgiloy[®], TiO₂ for nitinol or Ta₂O₅, and TaOH for tantalum. Therefore, the biostability of metal implants often relies on appropriate processing and a surface finish that reinforces the native oxide layer. Treatments such as acid passivation or electropolishing, used extensively in the final manufacturing process for endoluminal stents, as prescribed in the ASTM F86 standard,^[20] have been reported to improve the stability of materials such as nitinol and stainless steel significantly. Electropolishing removes the native oxide to form a more homogeneous oxide layer, resulting in a smoother surface, with roughness

in the range of nanometers, suitable for blood contacting devices.

INORGANIC COATINGS

Inorganic coatings have been investigated on endovascular stents, such as diamondlike carbon, hydrogen-rich amorphous silicon carbide (a-SiC:H) and amorphous titanium oxide.^[21–23] Such inorganic coatings improve the overall biocompatibility of the stent with the advantage of low thrombogenicity in comparison with uncoated devices. In addition, titanium–nitric-oxide–coated stainless steel stents significantly reduced neointimal hyperplasia in an experimental model.^[24] Along with increased biocompatibility, these inorganic coatings may also improve significantly the corrosion properties and in turn reduce the release of ions in the surrounding vessel wall. Five inorganic-coated stents are available: Tenax (Biotronik, Germany), Sirius Carbostent (Sorin Biomedica Cardio, Italy), Diamond Flex (Phystis, Germany), BioDiamond (Plasma Chem, Germany), and the MAC carbon stent (AMG, Germany). To date, the long-term beneficial effects of inorganic coatings have not been demonstrated by randomized trials. Gold coatings have also been used with the expectation that the inertness of gold would increase the overall biocompatibility while increasing the radiopacity of the device. Clinical studies have, however, reported an increased likelihood for restenosis that may be related to a coating defect rather than the gold itself.^[25] This strongly suggests that, along with biomechanical considerations, the processing itself could influence the outcome of the procedure.

THROMBORESISTANT COATINGS

Many approaches have been investigated to reduce thrombogenicity of metallic endovascular stents, using heparin, phosphorylcholine, or r-hirudin coatings, for instance. While these coatings have been proven to reduce thrombus formation on the device and as a consequence lower the subacute thrombosis rate in an animal model, the beneficial long-term clinical impact of these devices remains uncertain. Heparin-coated stents are commercially available and have been investigated in many experimental and clinical studies. For example, the heparin-coated BX Velocity stent, wherein heparin is immobilized on the surface and remains free to inhibit thrombi formation, has recently been accepted by the FDA. These devices are well tolerated but failed to prevent neointimal hyperplasia.^[5,6] Along with its anticoagulant properties, heparin has been shown, *in vitro*, to have an anti-proliferative action on SMCs. However, local delivery

of heparin from either delivery catheters or stents did not prove to be beneficial. The lack of activity of heparin-coated devices is likely to be related to sub-optimal tissue concentration of the drug. Phosphorylcholine-coated *Biodiv Ysio* stents also hold promise in the revascularization of small arteries, which remains a very challenging issue. Phosphorylcholine coating has been demonstrated *in vitro* to prevent platelet adhesion to the metallic surface. The phosphorylcholine-based coating is also being investigated as a drug delivery platform through hydrophobic interactions.^[26,27]

MEMBRANE-COVERED STENTS

As initially described by Parodi, stent grafts or membrane-covered stents are subclasses of stents developed for the endovascular treatment of abdominal aortic aneurysms (AAA).^[28] Stent-grafts are made of a conventional segment of bypass graft supported at the ends by expandable metallic stents. Membrane-covered stents are endovascular stents covered with or covering a synthetic membrane (Fig. 3). The membrane is usually made of a conventional surgical material such as Dacron, ePTFE, or polyurethane. These devices are currently being used or are in clinical trials for aneurysmal or AV-fistula repair, arterial rupture sealing,

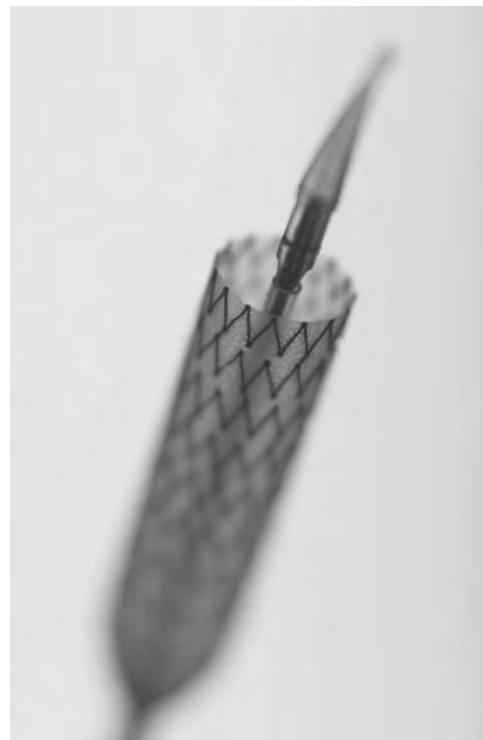


Fig. 3 Membrane-covered stent during its deployment.

coronary and peripheral occlusive diseases, and for the treatment of aortocoronary vein graft disease.^[29–32] The underlying concept resides in vessel wall exclusion by complete semi-porous materials. In the particular setting of saphenous vein graft disease, membrane-covered stents, either PTFE sandwiched between two stainless steel layers (JOGRAFT, JOMED) or stainless steel sandwiched between two PTFE layers (SYM-BIOT, Boston Scientific), are being investigated with the goal of reducing the risk of distal embolization of the friable plaque and reducing restenosis.

Chronic inflammatory reactions and delayed endothelialization of the membrane have been reported following implantation. The impaired endothelialization of the synthetic membranes could translate into higher risk of subacute thrombosis rates. One-year implantation studies, however, by Virmani et al. have demonstrated good results with ePTFE-nitinol grafts, with complete endothelialization observed within 6 months.^[33] Issues such as graft migration, graft fretting against the vascular wall, extravasation of blood into the aneurysm sac after stent-graft deployment, so-called endoleak, and structural failure of the graft, such as those reported by Guidoin et al., however, give rise to some concerns about the long-term efficacy of these devices.^[34]

RADIOACTIVE STENTS

Brachytherapy, i.e., the use of radiotherapy to inhibit neointimal proliferation, represents the treatment of choice if in-stent restenosis occurs. Radiation treatments inhibit the ability of cells to clonogenically divide, thus preventing acute cellular proliferation and migration, usually triggered by vascular wall injury. It is delivered in situ via a catheter-based, removable radioactive source (Ribbon or Beads Beta source or Gamma source—deeper penetration with radiosafety concerns). Stents have also been investigated as a radiation delivery platform with the idea that prolonged exposure of the arterial wall to low radiation doses may be beneficial.^[35] Most of the experiences have been performed using 32-phosphorus ion implanted stents. Dose-dependant inhibition has been reported with the β -emitting ³²P stents.^[36] Edge restenosis caused by intense radiation decay at stent outer limits and high rate of subacute thrombosis have limited their development.^[37] The hot-ends approach has been proposed to circumvent the edge-effect by reducing the dose's fall-off at the stent margins. Conversely, the investigation of radionuclides with more appropriate physical properties (γ radiation, shorter half-life, etc.) could renew interest in the field. The emergence of efficient antirestenosis drug-eluting stents will likely

restrict the use of radiation therapy as a mode of treatment of restenosis.

BIODEGRADABLE STENTS

While the scaffolding effect of the stent is beneficial in the short term, the presence of metallic devices may mediate chronic inflammatory reactions and in turn stimulate restenosis. To circumvent the latter, the concept of a biodegradable stent has been proposed. This stent should display the same mechanical properties, e.g., radial strength, of a metallic implant but then slowly disappear (biodegrade) once the artery is stabilized. Tamai and coworkers have recently reported the first clinical application of biodegradable stents made of poly-L-lactic acid (PLLA) monofilament. Follow-up at 6 months indicated an acceptable rate of restenosis (10.6%).^[38] The potential of local delivery of biologically active components from such biodegradable stents may favor their development.

It is noteworthy that the intriguing concept of a metallic biodegradable stent based on corrosion biodegradation of magnesium or iron (Fe > 99.8%) has been recently introduced, respectively, by Heublein et al. and Peuster et al.^[39,40]

DRUG-ELUTING STENTS

The development of drug-eluting stents that could efficiently deliver biologically active drugs to the vascular wall has been the object of much investigation in recent years. Early works using various drugs and coatings were often disappointing due to poor drug efficacy, uncertain drug dosage or required duration of delivery, and most commonly unstable and proinflammatory polymeric coating. The excellent results recently reported using antiproliferative drugs, such as sirolimus and paclitaxel, have, however, generated hope that restenosis will no longer compromise endovascular revascularization.^[41–44] These cytotoxic hydrophobic drugs derived from cancer chemotherapy or kidney allograft rejection treatment can be retained in the vascular wall when eluted from the stent and can selectively target proliferating cells within the artery. The mechanical scaffold offered by the metallic stent conjugated with the antiproliferative properties of these drugs translate, at least in initial clinical studies, into very low restenosis rates. The CE Mark and FDA approval of the sirolimus-coated Cypher™ stent (Cordis) and the enthusiastic results of the US SIRIUS trial with this device have further motivated ever increasing activity in the field of drug-eluting stents. Sirolimus is a natural macrocyclic lactone with potent antiproliferative and antiinflammatory properties. Sirolimus eluted from the



stent platform binds to SMCs' cytosolic receptor proteins and inhibits cyclin-dependent kinase complexes, thus mediating cell arrest in the late G1 phase. Paclitaxel prevents microtubule organization and as a consequence inhibits several steps in SMC migration and proliferation. Along with sirolimus and paclitaxel, many drugs are currently under investigation for stent-based delivery. These include everolimus, tacrolimus, estradiol, dexamethasone, angiopeptin, batimastat, and c-myc antisense oligonucleotide.

Regardless of the type of drug, proper coating and drug delivery platform remain the key issues in optimizing the antirestenotic effect of these devices. Two different approaches are currently being used to achieve delivery of the drugs: polymer-coated devices (Cordis/sirolimus Cypher stent and Boston Scientific/paclitaxel stent) and direct elution from dip-coated stents (Cook/paclitaxel stent). The potential long-term degradation and inflammatory effect of the polymer coatings versus control of the release rate are the key elements in the debate between these two approaches. The disappointing results obtained in the Guidant's DELIVER trial, using Cook Inc.'s ACHIEVE™ stent, which did not employ a polymer coating to elute the drug, suggest that sustained release is highly desirable to optimize the beneficial action of the drug. However, much work remains to be done to assess the long-term safety of these new devices and determine their clinical potential. Along with possible damaging effects to the artery of the antiproliferative components, the biocompatibility of the coated surfaces requires further investigation and potentially optimization. In particular, delayed reendothelialization and late thrombosis seem to be a potential outcome of these antiproliferative strategies.^[45] The very low restenosis rates observed so far (<5%) in clinical trials using drug-eluting stents may just represent a delayed healing, and long-term studies in humans are required to show whether the benefit is maintained.

CONCLUSIONS

Over the past decade, stenting has become the procedure of choice in many intraarterial revascularization procedures. Advances in stent technology are likely to increase their use further in an always growing number of conditions. In particular, new devices with excellent mechanical properties based on improved design or new materials offer exciting possibilities to treat challenging pathologies. The availability of devices designed to specific situations such as small vessel diseases or bifurcation lesions is also noteworthy. Drug-eluting stents that locally deliver therapeutic amounts of antiproliferative drugs may solve the issue of restenosis that has continuously plagued stent

implantation. Despite the increase in cost associated with the use of these devices, the elimination of restenosis will translate in more percutaneous transluminal treatment of occlusive vascular and nonvascular disorders, whereas surgical treatment modalities will be reserved for a minority of extremely complex diseases. The success of the pharmaceutical strategies combined with stents has, however, increased the hope that pharmaceutical strategies alone will be able to cure endovascular diseases. Systemic therapies with new drugs, nanoparticles-based local drug delivery, and gene therapy, for example, are promising research areas that could in the future reduce the need for the mechanical scaffold to the vascular wall offered by the stents.

ACKNOWLEDGMENTS

The authors are grateful to Dr. J. Silver from Nitinol Devices & Components for photography of the SMART stent. We also wish to thank Mrs. P. Cap for her help in reviewing the manuscript.

ARTICLES OF FURTHER INTEREST

Controlled Release; Polymers; Shape Memory Metals; Stent Grafts, Endovascular

REFERENCES

1. Serruys, P.W.; de Jaegere, P.; Kiemeneij, F.; Macaya, C.; Rutsch, W.; Heyndrickx, G.; Emanuelsson, H.; Marco, J.; Legrand, V.; Materne, P. A comparison of balloon-expandable-stent implantation with balloon angioplasty in patients with coronary artery disease. Benestent Study Group. *N. Engl. J. Med.* **1994**, *331*, 489–495.
2. Fischman, D.L.; Leon, M.B.; Baim, D.S.; Schatz, R.A.; Savage, M.P.; Penn, I.; Detre, K.; Veltri, L.; Ricci, D.; Nobuyoshi, M. A randomized comparison of coronary-stent placement and balloon angioplasty in the treatment of coronary artery disease. Stent Restenosis Study Investigators. *N. Engl. J. Med.* **1994**, *331*, 496–501.
3. Moussa, I.; Di Mario, C.; Di Francesco, L.; Reimers, B.; Blengino, S.; Colombo, A. Subacute stent thrombosis and the anticoagulation controversy: Changes in drug therapy, operator technique, and the impact of intravascular ultrasound. *Am. J. Cardiol.* **1996**, *78*, 13–17.
4. Gershlick, A.H. Role of stenting in coronary revascularisation. *Heart* **2001**, *86*, 104–112.
5. Babapulle, M.N.; Eisenberg, M.J. Coated stents for the prevention of restenosis: Part II. *Circulation* **2002**, *106*, 2859–2866.
6. Babapulle, M.N.; Eisenberg, M.J. Coated stents for the prevention of restenosis: Part I. *Circulation* **2002**, *106*, 2734–2740.

7. Welt, F.G.; Rogers, C. Inflammation and restenosis in the stent era. *Arterioscler. Thromb. Vasc. Biol.* **2002**, *22*, 1769–1776.
8. Grewe, P.H.; Deneke, T.; Machraoui, A.; Barmeyer, J.; Muller, K.M. Acute and chronic tissue response to coronary stent implantation: Pathologic findings in human specimen. *J. Am. Coll. Cardiol.* **2000**, *35*, 157–163.
9. Bhargava, B.; De Scheerder, I.; Ping, Q.B.; Yanming, H.; Chan, R.; Soo Kim, H.; Kollum, M.; Cottin, Y.; Leon, M.B. A novel platinum–iridium, potentially gamma radioactive stent: Evaluation in a porcine model. *Catheter. Cardiovasc. Interv.* **2000**, *51*, 364–368.
10. van Dijk, L.C.; van Holten, J.; van Dijk, B.P.; Matheijssen, N.A.; Pattynama, P.M. A precious metal alloy for construction of MR imaging-compatible balloon-expandable vascular stents. *Radiology* **2001**, *219*, 284–287.
11. Colombo, A.; Stankovic, G.; Moses, J.W. Selection of coronary stents. *J. Am. Coll. Cardiol.* **2002**, *40*, 1021–1033.
12. Kandzari, D.E.; Tchong, J.E.; Zidar, J.P. Coronary artery stents: Evaluating new designs for contemporary percutaneous intervention. *Catheter. Cardiovasc. Interv.* **2002**, *56*, 562–576.
13. Kastrati, A.; Mehilli, J.; Dirschinger, J.; Pache, J.; Ulm, K.; Schuhlen, H.; Seyfarth, M.; Schmitt, C.; Blasini, R.; Neumann, F.J.; Schomig, A. Restenosis after coronary placement of various stent types. *Am. J. Cardiol.* **2001**, *87*, 34–39.
14. Harnek, J.; Zoucas, E.; Carlemalm, E.; Cwikiel, W. Differences in endothelial injury after balloon angioplasty, insertion of balloon-expanded stents or release of self-expanding stents: An electron microscopic experimental study. *Cardiovasc. Interv. Radiol.* **1999**, *22*, 56–61.
15. Carter, A.J.; Scott, D.; Laird, J.R.; Bailey, L.; Kovach, J.A.; Hoopes, T.G.; Pierce, K.; Heath, K.; Hess, K.; Farb, A.; Virmani, R. Progressive vascular remodeling and reduced neointimal formation after placement of a thermoelastic self-expanding nitinol stent in an experimental model. *Catheter. Cardiovasc. Diagn.* **1998**, *44*, 193–201.
16. McLean, D.R.; Eiger, N.L. Stent design: Implications for restenosis. *Rev. Cardiovasc. Med.* **2002**, *3 Suppl. 5*, S16–S22.
17. Stoeckel, D. Nitinol medical devices and implants. *Minim. Invasive Ther. Allied Technol.* **2000**, *9*, 81–88.
18. Chwirut, D.J.; Oktay, H.S.; Ryan, T. Nitinol Interventional Cardiology Devices—FDA’s Perspective. In *Shape Memory and Superelastic Technologies*; Pelton, A., Hodgson, D., Russell, S., Duerig, T., Eds.; SMST: Santa Clara, 1997; 369–377.
19. Thierry, B.; Merhi, Y.; Bilodeau, L.; Trépanier, C.; Tabrizian, M. Nitinol versus stainless steel stents: Acute thrombogenicity study in an ex vivo porcine model. *Biomaterials* **2002**, *23*, 2997–3005.
20. ASTM F86, Standard Practice for Surface Preparation and Marking of Metallic Surgical Implants. In *Annual Book of ASTM standards: Medical Devices and Services*; American Society for Testing and Materials, 1995; Vol. 13.01, 6–8.
21. Nan, H.; Ping, Y.; Xuan, C.; Yongxang, L.; Xiaolan, Z.; Guangjun, C.; Zihong, Z.; Feng, Z.; Yuanru, C.; Xianghuai, L.; Tingfei, X. Blood compatibility of amorphous titanium oxide films synthesized by ion beam enhanced deposition. *Biomaterials* **1998**, *19*, 771–776.
22. Gutensohn, K.; Beythien, C.; Bau, J.; Fenner, T.; Grewe, P.; Koester, R.; Padmanaban, K.; Kuehnl, P. In vitro analyses of diamond-like carbon coated stents. Reduction of metal ion release, platelet activation, and thrombogenicity. *Thromb. Res.* **2000**, *99*, 577–585.
23. Heublein, B.; Ozbek, C.; Pethig, K. Silicon carbide-coated stents: Clinical experience in coronary lesions with increased thrombotic risk. *J. Endovascular Surg.* **1998**, *5*, 32–36.
24. Windecker, S.; Mayer, I.; De Pasquale, G.; Maier, W.; Dirsch, O.; De Groot, P.; Wu, Y.P.; Noll, G.; Leskosek, B.; Meier, B.; Hess, O.M. Stent coating with titanium–nitride-oxide for reduction of neointimal hyperplasia. *Circulation* **2001**, *104*, 928–933.
25. Kastrati, A.; Schomig, A.; Dirschinger, J.; Mehilli, J.; von Welser, N.; Pache, J.; Schuhlen, H.; Schilling, T.; Schmitt, C.; Neumann, F.J. Increased risk of restenosis after placement of gold-coated stents: Results of a randomized trial comparing gold-coated with uncoated steel stents in patients with coronary artery disease. *Circulation* **2000**, *101*, 2478–2483.
26. New, G.; Moses, J.W.; Roubin, G.S.; Leon, M.B.; Colombo, A.; Iyer, S.S.; Tio, F.O.; Mehran, R.; Kipshidze, N. Estrogen-eluting, phosphorylcholine-coated stent implantation is associated with reduced neointimal formation but no delay in vascular repair in a porcine coronary model. *Catheter. Cardiovasc. Interv.* **2002**, *57*, 266–271.
27. Lewis, A.L.; Tolhurst, L.A.; Stratford, P.W. Analysis of a phosphorylcholine-based polymer coating on a coronary stent pre- and post-implantation. *Biomaterials* **2002**, *23*, 1697–1706.
28. Parodi, J.C.; Ferreira, M. Historical prologue: Why endovascular abdominal aortic aneurysm repair? *Semin. Interv. Cardiol.* **2000**, *5*, 3–6.
29. Kribs, S. Endovascular stent grafting: A review. *Can. Assoc. Radiol. J.* **2001**, *52*, 145–152.
30. Baldus, S.; Koster, R.; Reimers, J.; Kahler, J.; Meinertz, T.; Hamm, C.W. Membrane-covered stents: A new treatment strategy for saphenous vein graft lesions. *Catheter. Cardiovasc. Interv.* **2001**, *53*, 1–4.
31. Stockx, L. Stent-grafts in the superficial femoral artery. *Eur. J. Radiol.* **1998**, *28*, 182–188.
32. Kessel, D.O.; Wijesinghe, L.D.; Robertson, I.; Scott, D.J.; Raat, H.; Stockx, L.; Nevelsteen, A. Endovascular stent-grafts for superficial femoral artery disease: Results of 1-year follow-up. *J. Vasc. Interv. Radiol.* **1999**, *10*, 289–296.
33. Virmani, R.; Kolodgie, F.D.; Dake, M.D.; Silver, J.H.; Jones, R.M.; Jenkins, M.; Gillespie, D.L. Histopathologic evaluation of an expanded polytetrafluoroethylene–nitinol stent endoprosthesis in canine iliofemoral arteries. *J. Vasc. Interv. Radiol.* **1999**, *10*, 445–456.
34. Guidoin, R.; Marois, Y.; Douville, Y.; King, M.W.; Castonguay, M.; Traore, A.; Formichi, M.; Staxrud, L.E.; Norgren, L.; Bergeron, P.; Becquemin, J.P.;

- Egana, J.M.; Harris, P.L. First-generation aortic endografts: Analysis of explanted Stentor devices from the EUROSTAR Registry. *J. Endovascular Ther.* **2000**, *7*, 105–122.
35. Chorny, M.; Fishbein, I.; Golomb, G. Drug delivery systems for the treatment of restenosis. *Crit. Rev. Ther. Drug Carr. Syst.* **2000**, *17*, 249–284.
36. Drachman, D.E.; Simon, D.I. Restenosis: Intracoronary brachytherapy. *Curr. Treat. Options Cardiovasc. Med.* **2002**, *4*, 109–118.
37. Serruys, P.W.; Kay, I.P. I like the candy, I hate the wrapper: The (32)P radioactive stent. *Circulation* **2000**, *101*, 3–7.
38. Tamai, H.; Igaki, K.; Kyo, E.; Kosuga, K.; Kawashima, A.; Matsui, S.; Komori, H.; Tsuji, T.; Motohara, S.; Uehata, H. Initial and 6-month results of biodegradable poly-L-lactic acid coronary stents in humans. *Circulation* **2000**, *102*, 399–404.
39. Heublein, B.; Rhode, R.; Huasdorf, G.; Hartung, W.; Haverich, R. Biocorrosion, a new principle for temporary cardiovascular implants? *Eur. Heart J.* **2000**, *21*, 286.
40. Peuster, M.; Wohlsein, P.; Brugmann, M.; Ehlerding, M.; Seidler, K.; Fink, C.; Brauer, H.; Fischer, A.; Hausdorf, G. A novel approach to temporary stenting: Degradable cardiovascular stents produced from corrodible metal—results 6–18 months after implantation into New Zealand white rabbits. *Heart* **2001**, *86*, 563–569.
41. Rensing, B.J.; Vos, J.; Smits, P.C.; Foley, D.P.; van den Brand, M.J.; van der Giessen, W.J.; de Feijter, P.J.; Serruys, P.W. Coronary restenosis elimination with a sirolimus eluting stent: First European human experience with 6-month angiographic and intravascular ultrasonic follow-up. *Eur. Heart J.* **2001**, *22*, 2125–2130.
42. Heldman, A.W.; Cheng, L.; Jenkins, G.M.; Heller, P.F.; Kim, D.W.; Ware, M., Jr.; Nater, C.; Hruban, R.H.; Rezai, B.; Abella, B.S.; Bunge, K.E.; Kinsella, J.L.; Sollott, S.J.; Lakatta, E.G.; Brinker, J.A.; Hunter, W.L.; Froehlich, J.P. Paclitaxel stent coating inhibits neointimal hyperplasia at 4 weeks in a porcine model of coronary restenosis. *Circulation* **2001**, *103*, 2289–2295.
43. Fox, R. American Heart Association 2001 scientific sessions: Late-breaking science—drug-eluting stents. *Circulation* **2001**, *104*, E9052.
44. Sousa, J.E.; Costa, M.A.; Abizaid, A.C.; Rensing, B.J.; Abizaid, A.S.; Tanajura, L.F.; Kozuma, K.; Van Langenhove, G.; Sousa, A.G.; Falotico, R.; Jaeger, J.; Popma, J.J.; Serruys, P.W. Sustained suppression of neointimal proliferation by sirolimus-eluting stents: One-year angiographic and intravascular ultrasound follow-up. *Circulation* **2001**, *104*, 2007–2011.
45. Liistro, F.; Colombo, A. Late acute thrombosis after paclitaxel eluting stent implantation. *Heart* **2001**, *86*, 262–264.

Sterilization of Biomedical Materials

Francesco Tassarolo

Department of Physics, University of Trento, Trento, Italy

Giandomenico Nollo

Department of Physics, University of Trento, and FBK-Fondazione Bruno Kessler, Trento, Italy

S

INTRODUCTION

The development of more complex medical devices and specific targeting nanodrugs requires materials able to guarantee special mechanical performance, physical–chemical characteristics, and properties of interaction with the biological environment and tissue interfaces.

Beside the specific design, shape, dimension, and functionality of the medical device, as well as surface and bulk properties of biomaterials, are fundamental parameters to obtain the desired diagnostic or therapeutic effect. In addition, many applications require devices and preparation to guarantee an absence of contaminant species in the finished product. Therefore, there is a need to process materials and release medical devices that satisfy the condition of sterility.

Sterilization is indeed the best practice to limit or prevent risk of infection related to the use of medical devices. For instance, the contamination of perimplant tissue is the foremost reason for complications post-intervention and often triggers the failure of implanted devices in the short and medium term. Sterility is, therefore, mandatory for every critical medical device that enters into contact with anatomic sterile districts (internal organs, blood tissue, etc.), but is also frequently desirable for semicritical devices used on integral mucosas of the digestive tract and respiratory apparatus.

Processes and technologies developed to guarantee the destruction or inactivation of microorganisms are based on physical or chemical agents and, often, on the synergic combination of two or more treatments. Sterilization aims to deactivate cell processes involved in metabolism or genetic transcription, inducing cell death or stopping the replication ability of contaminant species. Because many microorganisms, especially bacterial spores, have excellent resistance to environmental stresses, the achievement of sterility requires exposure to aggressive treatments and conditioning in artificial environmental conditions that are entirely different from what the material is designed for. Currently all available techniques induce, to different extents, some modifications in the sterilized material.

The development of sterile medical devices requires conciliating complex characteristics of biocompatibility, bioreabsorption, biomimesis, etc., to sterilization treatment. This paper aims to address the selection of the most appropriate sterilization technique by fixing some important procedural steps. An overview of the currently available sterilization methods, with discussion of some of the effects of traditional and experimental sterilization processes on medical devices, is presented.

BIOMATERIALS AND MEDICAL DEVICES

A material is of biomedical interest in accordance with its specific clinical use and the purpose for which it has been conceived, projected, and realized. Material design and destination of use are identified in the finished medical device. The medical device is then usually sterilized before use. The sole alternative to sterilization is to process sterile raw material in an aseptic environment from the inception to the packaging of the final product. The medical device includes all chemical, physical, and functional properties, in addition to hygienic requirements. The current legislation does not regulate the materials per se, but regulates their final shape with potentials and risks as medical devices.

The European Union and the United States have similar classifications for medical device risk that account for the global risk associated with the device including the infection risk, malfunctioning risk, electrical risk, and personnel risk. For sterilization purposes, the Spaulding classification, based on infective risk, is pertinent.^[1] According to this classification, medical devices are grouped into critical, semicritical, and non-critical and sterilization treatment is required, advisable, or not necessary, as detailed in Table 1.

STERILIZATION PROCESS

The final result of the sterilization process requires the absolute absence of living forms from the surface and

Table 1 Infective risk classification of medical devices according to Spaulding^[1]

Application	Level of risk	Spaulding classification	Recommendation
In close contact with a break in the skin or mucous membrane	High	Critical	Sterilization
For introduction into sterile body areas			
In contact with intact mucous membranes	Intermediate	Semicritical	High-level disinfection or sterilization
Prior to use in immunocompromised patients			
In contact with healthy skin	Low	Non-critical	Low-level disinfection

bulk of the item. The World Health Organization (WHO) states that sterilization is any chemical or physical process that implies the destruction of all living forms, including bacteria, fungi, virus, and spores. The achievement of the sterility status comes from a probabilistic process, which takes into account presterilization procedures of cleaning, disinfection, and packaging, the effectiveness and correct use of sterilizing media and apparatus, and the suitability of storage conditions and times.

WHO's definition simplifies the concept of sterility that can be operatively defined only on a statistical basis. To this end, the sterility assurance level (SAL) has been defined as the probability of finding a living microorganism in the sterilizer load after treatment. The SAL for a validated sterilization process is 10^{-6} . This requirement demands a minimal reduction in bioburden equal to $6+x$ decades, where x is the expected microorganism load before sterilization, expressed in decades. This complies with the need to reduce the microbial load and assures a SAL of 10^{-6} (Fig. 1).

As a result, any microbiologic test developed to assess sterilization effectiveness should utilize a minimal bioburden greater than 10^6 microorganisms. Procedures for testing effectiveness of sterilization treatments usually aim to reveal a reduction greater than 6 log in microorganisms selected among the most resistant species to the specific sterilization process.

PRESTERILIZATION PROCEDURES

From the operative definition of sterilization follows the importance of reducing the initial bioburden as much as possible before sterilization. Presterilization procedures account for any cleaning and disinfection process able to minimize inorganic and organic soil and microbiologic burden before sterilization.^[2] In the case of device reuse, the possibility of iatrogenic transmission must also be considered and the first treatment immediately after clinical use (usually

referred as decontamination) is mandatory to minimize the risk of disease transmission to health care professionals.

Cleaning and disinfection should be pursued using detergents (solvents, surfactants, enzymatic solutions, etc.) and biocides (chlorine, compounds, polyphenolic,

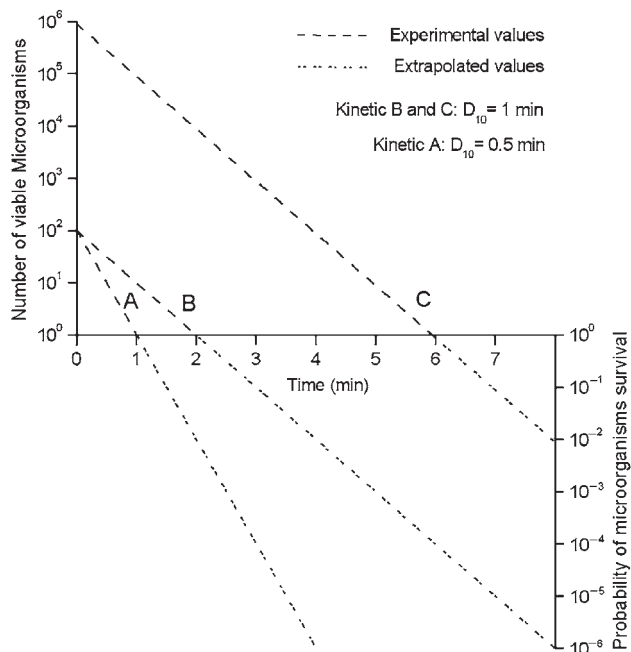


Fig. 1 Schematic representation of three possible sterilization kinetics and their efficiency with respect to SAL. Differences in the D_{10} value (time required to reduce the bioburden by 10 times) between kinetic A and B guarantee a SAL of 10^{-6} in 4 and 8 min, respectively, when considering an initial bioburden of 10^2 microorganisms. For higher initial bioburden (kinetic C), the same D_{10} value of kinetic B allows a SAL of 10^{-2} in the same time frame (8 min), which is considered insufficient for sterilization purposes. For a standardized process, inactivation kinetics usually follow simple exponential decay, which results in linearity in a semilog scale, but significant deviation from this ideal time course must be considered for non-conventional sterilization techniques.

aldehydes, etc.), respectively. Structured protocols and proper sequences of treatments should be considered to optimize presterilization protocols (Fig. 2).^[3]

Chemical properties of solutions could be highly enhanced by mechanical cleaning (spray washing, brushing, turbulation, ultrasonic cleaning, etc.). Special attention must be given in flushing and rinsing narrow and long lumens, cavities, and holes. Suitable rinsing and drying should guarantee the removal of biologic and chemical residuals, any eventual particulate, and water from exposed and inner surfaces. Effective drying is strictly required for low-temperature sterilization using ethylene oxide or hydrogen peroxide gas plasma to avoid the production of toxic byproducts (ethylene chlorohydrin, ethylene glycol) or inactivation of the sterilization efficiency, respectively. The residual biologic burden before sterilization should be lower than 10^2 microorganisms per device. Any other non-viable soil, especially organic material, albumin, nuclei acids, sugar, and starch, should be reduced to a minimum to avoid local inhibition of

chemicals, shielding of microorganisms, and a general inefficiency of the process.

MAIN STERILIZATION TECHNIQUES

A number of sterilization techniques are currently available, including traditional methods by saturated steam, ethylene oxide, ionizing radiation, and more recent processes based on low-temperature plasmas in the vapor or gas phase.

Despite this large panel of techniques, an elective process able to guarantee the sterilization of all materials while preserving them from modification does not exist. Each process presents pros and cons, and many undesirable effects are related to the incompatibility between the material properties and the process parameters of the sterilization technique.

In researching and selecting the most suitable sterilization technique for a specific device, it is mandatory

S

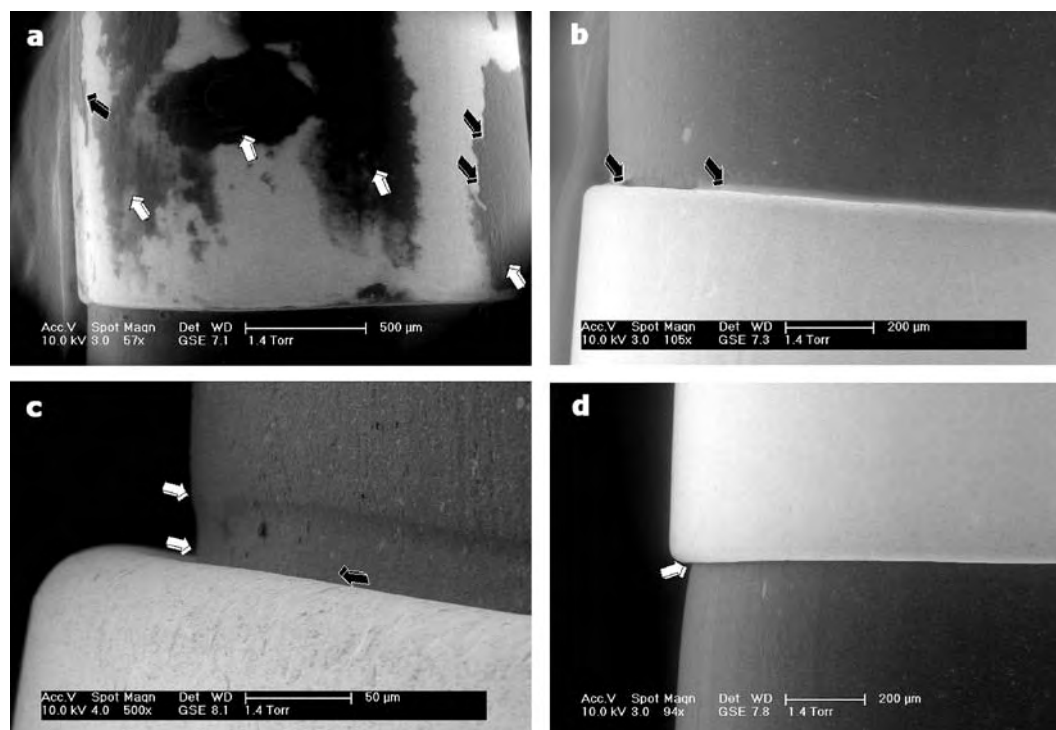


Fig. 2 Low-vacuum scanning electron microscopy after different decontamination-cleaning protocols on cardiac catheters soiled by bacterial spiked blood. Significant differences in cleaning efficiency are visible at the interface between the two device components: the polyurethanic shaft (dark gray) and platinum electrode (light gray). (a) Chlorine solution followed by enzymatic solution: massive blood residuals (white arrows) on the electrode surface. Black arrows indicate thick deposits. (b) Enzymatic solution followed by chlorine solution: meniscus-like residual at the electrode–polymer interface. Black arrows indicate a clean segment between the remaining soiled interfaces. (c) Polyphenolic emulsion: optimal cleaning with no residuals (black arrow). Electrode imprint and dislodging on the catheter shaft (white arrows) promote interstitial soiling. (d) Polyphenolic emulsion followed by enzymatic solution: optimal cleaning of the critical interface area (white arrow). Images are presented at different magnifications to highlight different levels of residual soil. From Tessarolo et al.^[3]

to perform in-depth compatibility analyses of the material with the operative parameter of the process.

Saturated Steam

Water-saturated steam at high temperature (121–134°C) and pressure (1.1–2.1 bar) is the most common sterilization technique used in hospitals and health care facilities. Different from the less effective treatment using unsaturated steam, a prevacuum is required in autoclave before the injection of water vapor. Steam is non-toxic, inexpensive, and has a high sporicidal effect with short application times (4–15 min). Steam is highly recommended for sterilizing heat-resistant material and metals. Conversely, steam is inapplicable to heat-sensitive materials. The elevated humidity and pressure can induce softening, degradation, and hydrolysis in many biomedical polymers.

Changes in the mechanical properties, surface degradation, and byproduct formation has been reported on metallic, polymeric, and biologic biomaterials.^[4,5] Because of its numerous advantages, steam has also been applied to biomedical polymers. Vascular polytetrafluoroethylene prosthesis, polyvinyl chloride hematic circuit, and polyurethanes used for pacemaker cables, ventricular-assisted devices and catheters have been sterilized by steam autoclaving. Although the thermal stress could be sustained without compromising polymer characteristics, low-temperature sterilization techniques are generally preferable for polymeric materials, especially in the case of multiple sterilizations.

Dry Heat

Dry heat sterilization is obtained in an oven (Pasteur stove) by hot air that destroys pathogens by oxidation, volatilization of light components, and deep dehydration of microorganisms. The slow penetration rate requires the achievement of a high-temperature (150–170°C) for a prolonged time (60–150 min). Dry heat has several constraints and limitations because of the high temperature and time required for microorganism inactivation. Only thermoresistant materials can be processed, and usually steel and glass are sterilized using this technique. Different from saturated steam, dry heat does not induce superficial corrosion and deposits and allows the effective sterilization of complex designs with closed cavities.

Ethylene Oxide

Ethylene oxide (EtO) sterilization is a commonly used sterilization technique for processing thermosensitive

products in the medical device industry. The sterilization process includes a vacuum phase in autoclave and a subsequent injection of EtO gas in concentrations ranging from 600 to 1200 mg/L. The chamber is kept between 30 and 50°C with a relative humidity of 40%–50%. The cycle lasts for 2–8 h according to the specific sterilization protocol and sterilizer load.

EtO is a non-colored, flammable, explosive gas, as well as an extreme irritant to skin and mucosa. The addition of carbon dioxide or chlorofluorocarbons lowers the flammability and explosive risk, but complicates environmental compatibility and treatment of exhausted gas.

Bactericidal, sporicidal, and virucidal effect result from the alchilation of sulphhydrylic, carboxylic, and hydroxylic groups in nucleic acids, which induces cell death.^[6]

The main advantage of EtO sterilization is related to the low-temperature process and the wide range of material compatibility. Conversely, the sterilization cycle should include a postprocess aeration for detoxification of gas residual that can last 12 h in forced systems or even 14 days if spontaneous. The Food and Drug Administration fixes the maximal level of EtO in a ready-to-use device at 25 mg/L, and the high permeability of polymeric materials to EtO requires monitoring of the residual levels. Handlos and coworkers^[7] demonstrated that the degassing process is mainly a diffusion-controlled phenomenon. Time courses of degassing are known for a wide range of materials, and the forced degassing procedure assures a time-effective process. However, EtO systems require costly installation, and concerns remain about potential exposure of personnel to the gas, which is highly toxic, mutagenic, and a suspected carcinogenic.^[7]

Although the exposure of polymers to EtO sterilization usually induces slight effects on molecular weight, degradation rate, mechanical properties, and surface chemistry,^[8] byproduct formation and alchilation reactions were observed with urethanic and urea groups.^[9] In specific biomedical applications, morphological and compositional degradation of surfaces were also reported.^[5,9–11]

Plasmas

Cold plasma techniques are currently the most effective alternative to EtO for low-temperature sterilization of thermosensitive materials. Cold plasma is a partly ionized gas including ions, electrons, ultraviolet photons, and neutral reactive species such as radicals and excited molecules generated by the application of an electric or magnetic field in specific gas or vapor phases. In the early 1990s two commercial plasma sterilization systems were marketed under the

trademark Sterrad™ and Plazlyte™. These technologies are based on hydrogen peroxide and a hydrogen peroxide-peracetic acid mixture, respectively.^[10] Their biocidal effectiveness is mostly related to the chemical sterilization phase obtained by the vaporization of the chemical solutions in an evacuated autoclave. The plasma phase is then used in both treatments to neutralize the residual active species and to provide a sterile device in a short processing time with no need for aeration or detoxification. Sterrad™ and Plazlyte™ should therefore be considered a low-temperature (> 55°C) chemical sterilization, different from the pure plasma-based techniques that are in development. Although various authors reported surface oxidation and roughening of biopolymers by these commercial plasmas (Fig. 3),^[10,12,13] no marked modifications in thermal or mechanical properties existed on the polymers. The release of 4,4-methylene dianiline was not revealed in repeatedly sterilized polyurethanes.^[10]

The high compatibility with a large number of polymers and the short sterilization cycle makes these techniques available for difficult tasks such as reprocessing single-use medical devices.^[13]

Ionizing Radiations

Sterilization by exposure to ionizing radiation is widely used in the medical device industries as a room temperature treatment and end-point process of chain production. Both γ -photons from ^{60}Co or ^{137}Cs sources and high-energy electrons from accelerators are ascribed to this technique. γ -Rays are able to deeply penetrate low-density materials, such as polymers, whereas high-energy electrons usually interact with the superficial layer of materials. To assure the homogeneous exposure of products to the required dose (a dose of 25 kGy is generally considered effective),

S

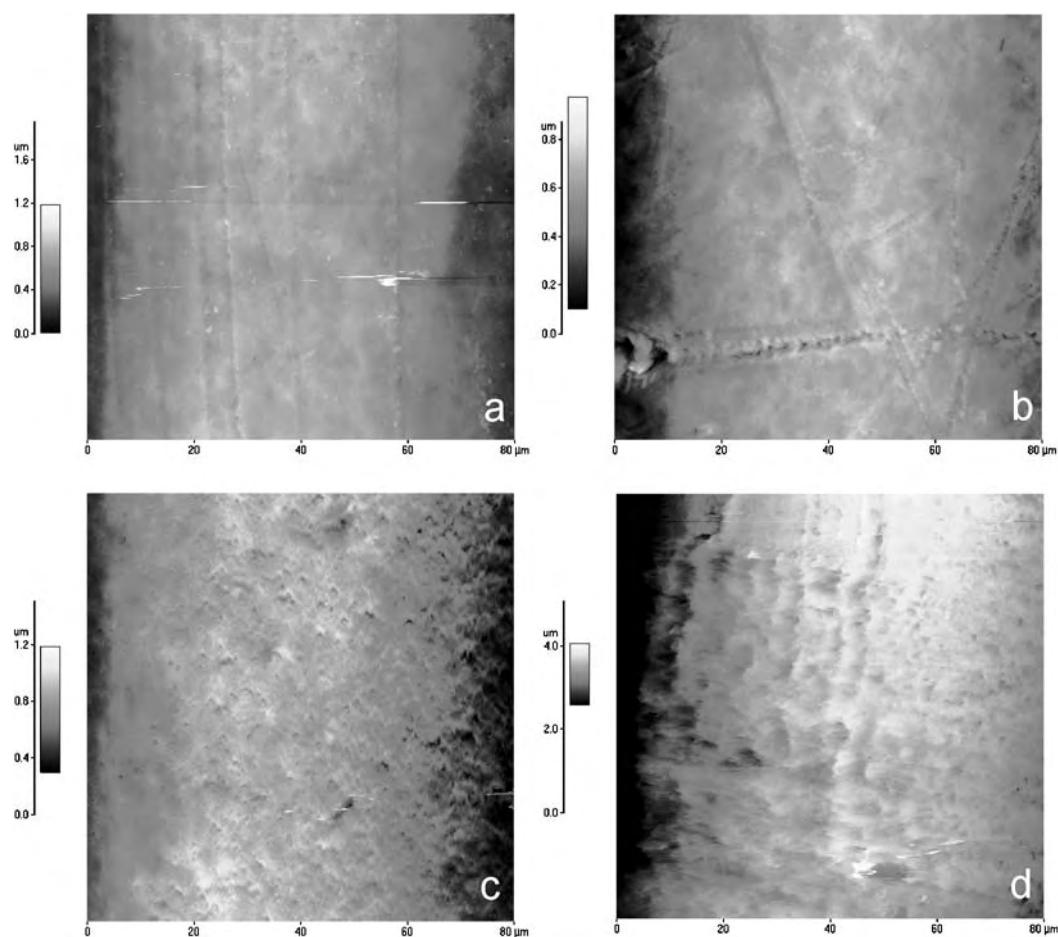


Fig. 3 Atomic force microscopy $80 \times 80 \mu\text{m}$ scans on the external surface of a polyurethane shaft for endovascular devices. (a) New and sterilized (b) one, (c) four, and (d) eight times by hydrogen peroxide gas plasma (Sterrad™ technology). Changes in nanoroughness of the shaft surface are evident after four cycles of reprocessing, documenting the etching effect of the plasma sterilization technique. From Tessarolo et al.^[13]

a relative displacement of materials with respect to the source is usually provided.

Although this technique allows treatment of large amounts of products in a time-effective manner, the application is limited by the huge investment required to realize the treatment facility and manage the radiation hazards.

Several papers have reported posttreatment modifications of sterilized polymers, such as increased brittleness, discoloration,^[14] hardening,^[8] softening, and variations in melting temperature and molecular weight.^[8,15] These modifications are the consequence of two opposing phenomena involved in the polymer radiation process: chain scission and cross-linking. Chain scission produces low-molecular-weight fragments with unsaturated chemical bonds and possible gas formation. Significant decreases in strain resistance and changes in crystallinity were reported in γ -radiated biopolymers after chain scission.^[8] Cross-linking usually results in an initial increase of tensile stress and decreased impact resistance. Polymers containing aromatic groups are typically less alterable than aliphatic chains, while terminal methyl groups are often cross-linked.

Superficial defects are usually enlarged after irradiation,^[9] and oxidation of superficial layers in biomedical polymer sterilized in aerobic environments was reported.^[8]

γ -Radiation can induce the production of toxic byproducts such as 4,4-methylene dianiline after decomposition of high-molecular-weight polyurethanes^[11] or other low-molecular-weight degradation compounds, which can induce cellular lysis. Resterilization by EtO of polyvinyl chloride devices previously treated by γ -radiation should be avoided because of the production of toxic epichloridrine.

Steam Formaldehyde

An effective method of sterilization by steam-formaldehyde was originally reported by Vink.^[16] The treatment is performed in an autoclave at about 0.1 bar with alternated injections of saturated steam and vaporized formalin. Several injections and evacuations of the chamber are required because of the low penetration ability of formalin with respect to EtO. Microorganism inactivation by formaldehydes, similar to other aldehydes, proceeds by the high chemical reactivity of carbonylic terminals able to alchilate aminic, hydroxylic, and sulphhydrylic groups of cell constituents.

Steam-formaldehyde sterilization is non-explosive and non-flammable and presents a markedly lower detection level by smelling than EtO, which allows earlier detection in the case of leakage and autoclave malfunctioning.

Formaldehyde residues have been detected on medical devices after steam-formaldehyde sterilization and the amount of residual is strongly related to the polymeric material and process parameters with a crucial role of the detoxifying and cleaning phases obtained by steam injections.^[16] Aldehyde residuals can be observed both as monomeric formaldehyde adsorbed into the materials and paraformaldehyde on the surface. Paraformaldehyde can be created by dissolution of formaldehyde on a moistened surface with the formation of polymeric hydrates. After polymer dehydration, paraformaldehyde can precipitate. As a result, hydrophilic and permeable polymers (polyamides, polyurethanes) with the possibility of absorbing larger percentages of water retain a greater amount of aldehydes than hydrophobic and non-permeable polymers.^[17]

Glutaraldehyde and Other Liquid Chemical Sterilizations

A class of chemicals such as glutaraldehyde, orthophthalaldehyde, performic acid, and peracetic acid, although defined as disinfectants because they are unable to realize a rapid inactivation of bacterial spores, can obtain sterilization through prolonged exposure time.^[18]

Glutaraldehyde is characterized by a marked reactivity similar to that described for formaldehyde, but with higher polymerization properties as a result of the double-ended molecule. Glutaraldehyde is usually available as an aqueous solution to be activated by regulating the pH. The strongest biocidal activity is reached at alkaline pH and increases with temperature to a maximum at 70°C.

Devices should be immersed completely in the sterilizing solution and then a careful rinsing with sterile water must be applied to remove chemical residual from the processed device. When compared to widespread disinfectants such as chlorine, glutaraldehyde does not affect metal components. Mixtures of glutaraldehyde and other disinfectants agents such as phenolic compounds are often used to obtain synergic biocidal effects.

Liquid-chemical sterilization techniques are widely used in hospitals for rapid low-temperature sterilization of thermosensitive materials, but do not offer the same hygienic assurance as heat-based or radiation-based techniques.^[2]

Other Available Sterilization Techniques

Several other ad hoc methodologies for sterilization are available. Microwaves are frequently used for rapid sanitization of infected waste, but significant

application for sterilization of catheters, air ventilators, and polymeric tubing are possible. X-rays obtained from Röntgen tubes or converted from e-beam accelerators are also available for low-temperature sterilization of thermosensitive materials such as polymers.

Mechanical filtration is a possible alternative for sterilization of soluble substances that are thermally and chemically sensitive and would be damaged by heat, radiation, or chemical sterilization. This method is widely used for protein solutions and pharmaceutical compounds. A filter with pore size 0.2 μm will effectively remove bacteria. If viruses must also be removed, a much smaller pore size, around 20 nm, is required.

NEW AND FORTHCOMING STERILIZATION TECHNIQUES

The increasing development of new biomedical materials with structural and superficial characteristics for highly specialized applications necessitates the parallel research and development of new sterilization techniques based on different chemical and physical principles that can assure strong biocidal effects as well as limited modification on sterilized materials.

Experimental Plasmas

Numerous studies are being conducted on plasmas for surface modification that demonstrate good biocidal properties with respect to bacteria and viruses. These techniques, properly called plasma treatments, are based on the exposure of materials to ions and activated molecules obtained by exciting selected mixtures of gases with radiofrequency or microwave electromagnetic fields. Low-temperature sterilization under direct exposure to the plasma could be realized by direct current, radiofrequency (1–100 MHz), or microwave (MW; > 300 MHz) discharges. The plasma condition can be obtained by MW excitation in the pulsed mode or by a very-low-frequency oscillating field (1–10 Hz) as needed to operate the corona discharge mode or dielectric barrier mode.^[19]

The choice of gas in the efficiency of spore inactivation, which is the most resistant contaminating agent to plasma, has special relevance. A high efficiency in spore inactivation was demonstrated independent of the type of discharge for O₂, N₂, SO₂, SF₆, NO₂, H₂O, H₂O₂, CO₂, halogens, air, and aldehydes.

Beside the chemical and physical etching, a non-negligible biocide role is attributed to ultraviolet radiation. Highly energetic UV photons (3.3–6.2 eV) provide a marked biocide effect because they correspond

with the frequency of maximum absorbance in DNA and nucleic acids.

An in-depth revision of the operative parameters involved in plasma sterilization was reported by Lerouge and collaborators,^[19] who demonstrated how gas composition, field power, excitation frequency, inlet gas flux, chamber pressure, and geometric factors could be modified to optimize the biocidal effect and minimize material alteration.

S

Electron Cyclotron Resonance Plasma

Chao and coworkers proposed a sterilization plasma reactor based on the electron cyclotron resonance (ECR) effect.^[20] When plasma is generated and maintained by MW, the application of an external magnetic field induces rotational trajectories in the free electrons. Variation of the magnetic field creates a variation in the angular frequency of electrons and a resonance with MW excitation can be obtained. In ECR mode the energy for sustaining plasma is minimal and a strong etching effect could be obtained. ECR plasmas offer a high ionization rate, low average temperature, and a large operative range in pressure and temperature.

The main disadvantage of plasma sterilizations is the low penetration of the active species, which are deactivated in the first layer of material. Similarly, shadowing effects must be considered and the presence of packaging markedly generally reduces the sterilization efficiency. Some studies proposed the use of an additional irradiation by UV in the range of 254 nm, which is the most efficient against microorganisms,^[19] but the large application of pure plasmas for batch sterilization requires a deeper understanding of the process.

Supercritical CO₂

Treatment by supercritical CO₂ was successfully applied to the inactivation of many bacterial species and yeasts in food,^[21] but was considered also for the sterilization of biomaterials.^[22] The treatment of materials in operative conditions near the critical region requires temperatures and pressures able to induce supercritical fluidity. The inactivation of *Escherichia coli* was reported under moderate pressure conditions (73.8 bar) and temperatures (31.1°C). Higher temperature and pressure were required for spore inactivation.^[21] Spilimbergo and coworkers demonstrated that the sterilization effect is mainly based on the diffusion of CO₂ into the cell and the extraction of cytoplasmic material through the membrane.^[22] The amphiphilic properties of supercritical CO₂ allow the rapid diffusion through and the fluidification of the cell

membrane. Cytoplasmic pH is therefore altered and the structure is compromised, inducing cell death.

Treatment of sensitive biopolymers (polylactic and polyglycolic acid polymers) did not demonstrate any macroscopic effect of degradation or variation in mass or molecular weight.

Additional studies are required to validate the operative protocol against many pathogenic species, but supercritical CO₂ sterilization could be of interest for treating sensitive material and drugs because of its intrinsic low reactivity to non-severe chemical-physical conditions and its high penetrating capability into cavities of the medium.

Pulsed Electric Field

Sterilization treatments based on high-intensity pulsed electric field (PEF) are realized by the application of high-voltage pulses (20–80 kV/cm) to substrates interposed between electrodes for less than 1 sec at room temperature.

These techniques are promising methods for sterilizing foods and the ability to inactivate large number of microbial species^[23] suggests the possibility of sterilization of biomedical materials as well.

The mechanism of microbial inactivation is mainly the result of membrane disruption following the application of the electric field. Breakage of the bacterial membrane is reached for a transmembrane potential of about 1 V, which, for *E. coli*, corresponds to about 10 KV/cm.^[23] This process has been called electroporation and induces cell swelling and membrane disruption.

The increase in pulse number usually increases the biocidal effect, but the first pulse mainly affects bacterial viability. Microorganism inactivation can also be improved by combining PEF with other physical or chemical stresses such as variation of pH, ionic strength, and pressure.

Spilimbergo et al. proposed the combination of PEF and supercritical CO₂ to inactivate bacterial spores at temperature lower than 40°C^[22] by favoring the penetration of the supercritical fluid with the electroporation process.

Little information is available regarding PEF treatment of biomedical materials, but its good compatibility and low alteration of food properties indicate a possible compatibility with thermosensitive materials and biomedical polymers.

STERILIZATION OF NON-CONVENTIONAL AGENTS

Non-conventional pathogenic agents, such as abnormal prions and bacterial endotoxins, are not

considered in the former definition of sterilization because they are non-viable contaminants. However, the increasing incidence of prion-related disease and the possibility of pyrogenic adverse events require treatments and techniques able to significantly reduce these contaminants. In this context the term sterilization is usually extended and refers to a process able to guarantee a reduction in the contaminating agent for more than 6 decades. Prions and endotoxins are notoriously highly resistant to inactivation by standard sterilization processes. The most common and safe approach to sterilization of these agents uses thermal techniques (dry heat or saturated steam) for a treatment time ranging from 3 to 6 times the normal sterilization cycle for microorganisms.^[24] Specific chemical treatments are also available but require harsh working conditions. Plasma sterilization, because of its peculiar etching properties able to inactivate and remove contaminants, could have some potential in this field.^[25]

A POSSIBLE APPROACH TO THE DEVELOPMENT OF A STERILE DEVICE

The usual approach to the individuation of an elective sterilization technique is based on a wide range of multidisciplinary experimentation. Material differences pre- and posttreatment and modifications induced by different sterilization techniques should be elicited to determine the technique that minimizes material alterations with respect to the final device use.

In-depth studies are required for some classes of devices to establish the maximum number of sterilization or reprocessing sustainable under safe and effective conditions. In this process of analytical characterization, experimental parameters such as biocompatibility, bioactivity, cell and bacteria adhesion, mechanical and rheological properties, and byproducts release are usually considered.

Close to the traditional process *ex post*, in which the objective is to identify the correct form of sterilization for minimizing modifications on the finished material, is the *ex ante* approach. The in-depth knowledge of sterilization techniques and their actions on materials could be used to define *ex ante* the sterilization treatment by considering it an integral part of device production. In a similar process, sterilization is not an additional and external step, but a functional phase to the implementation of desired parameters. Unavoidable modifications induced by sterilization could be turned into expected ameliorations of the finished device.

The possibility of obtaining material improvement and the availability of a ready-to-use sterile device are the immediate advantages of implementing a sterilization protocol as an integral part of the production

process. An example is represented by sterilization of the ultrahigh-molecular-weight polyethylene for joint prostheses. γ -Radiation in a non-oxidizing environment allowed proper modulation of cross-linking without increased polymer oxidation, therefore improving the mechanical properties of the sterile finished device.

In this field, the potential properties and parameters of plasma-based sterilization techniques might represent a source of new and interesting solutions for both sterilization and material processing. Plasmas are already widely used for surface functionalization and modification of biomaterials, important phases in the production process. The effort currently in progress to verify the efficacy of these techniques on microorganism inactivation indicates great potential in production-sterilization combined systems.

CONCLUSIONS

No available sterilization technique is applicable to all medical devices, and each sterilization method induces modifications to some extent in the chemical, physical, and thermomechanical properties of biomedical materials. Sterilization treatments can affect device properties and in vivo functionality, making the choice of the correct sterilization treatment a crucial step in medical device development and application.

ACKNOWLEDGMENT

The authors thank Dr. Iole Caola of the Microbiology and Virology Laboratory, Department of Medicine Laboratory, APSS, Trento, for her valuable support in reviewing the manuscript.

ARTICLE OF FURTHER INTEREST

Infection of Medical Devices

REFERENCES

- Spaulding, E.H. Chemical disinfection and antisepsis in the hospital. *J. Hosp. Res.* **1972**, *9*, 5–31.
- Russel, A.D.; Hugo, W.B.; Ayliffe, G.A.J. *Principles and Practice of Disinfection, Preservation and Sterilization*, 3rd; Blackwell Science: London, 1999.
- Tessarolo, F.; Caola, I.; Fedel, M.; Stacchiotti, A.; Caciagli, P.; Guarrera, G.M.; Motta, A.; Nollo, G. Different experimental protocols for decontamination affect the cleaning of medical devices. A preliminary electron microscopy analysis. *J. Hosp. Infect.* **2007**, *65*, 326–333.
- Baier, R.E.; Meyer, A.E.; Akers, C.K.; Natiella, J.R.; Meenaghan, M.; Carter, J.M. Degradative effects of conventional steam sterilization on biomaterial surfaces. *Biomaterials* **1982**, *3*, 241–245.
- Zhang, Y.Z.; Bjursten, L.M.; Freij-Larsson, C.; Kober, M.; Wesslen, B. Tissue response to commercial silicone and polyurethane elastomers after different sterilization procedures. *Biomaterials* **1996**, *17*, 2265–2272.
- Dempsey, D.J.; Thirucote, R.R. Sterilization of medical devices: A review. *J. Biomater. Appl.* **1989**, *3*, 454–523.
- Handlos, V. Kinetics of the aeration of ethylene oxide sterilized plastics. *Biomaterials* **1980**, *1*, 149–157.
- Goldman, M.; Pruitt, L. Comparison of the effects of gamma radiation and low temperature hydrogen peroxide gas plasma sterilization on the molecular structure, fatigue resistance, and wear behavior of UHMWPE. *J. Biomed. Mater. Res.* **1998**, *40*, 378–384.
- Abraham, G.A.; Frontini, P.M.; Cuadrado, T.R. Physical and mechanical behaviour of sterilized biomedical segmented polyurethanes. *J. Appl. Polym. Sci.* **1997**, *65*, 193–191.
- Lerouge, S.; Tabrizian, M.; Wertheimer, M.R.; Marchand, R.; Yahia, L. Safety of plasma-based sterilization: Surface modifications of polymeric medical devices induced by Sterrad and Plazlyte processes. *Biomed. Mater. Eng.* **2002**, *12*, 3–13.
- Hirata, N.; Matsumoto, K.; Inishita, T.; Takenaka, Y.; Suma, Y.; Shintani, H. Gamma ray radiation, autoclave and ethylene oxide sterilization to thermosetting polyurethane: Sterilization to polyurethane. *Radiat. Phys. Chem.* **1995**, *46*, 377–381.
- Meijs, G.F.; McCarthy, S.J.; Rizzardo, E.; Chen, Y.C.; Chatelier, R.C.; Brandwood, A.; Schindhelm, K. Degradation of medical-grade polyurethane elastomers: The effect of hydrogen peroxide in vitro. *J. Biomed. Mater. Res.* **1993**, *27*, 345–356.
- Tessarolo, F.; Ferrari, P.; Nollo, G.; Motta, A.; Migliaresi, C.; Zennaro, L.; Rigo, A.; Guarrera, G.M. Evaluation and quantification of reprocessing modification in single use devices in interventional cardiology. *Appl. Surf. Sci.* **2004**, *238*, 341–346.
- Sturdevant, M. Plastics in medicine: How sterilisation changes long-term resin properties. *Plast. Eng.* **1991**, *47*, 27–32.
- Hooper, K.A.; Cox, J.D.; Kohn, J. Comparison of the effect of ethylene oxide and gamma-irradiation on selected tyrosine derived polycarbonates and poly(L-lactic acid). *J. Appl. Polym. Sci.* **1997**, *63*, 499–510.
- Vink, P. Residual formaldehyde in steam-formaldehyde sterilized materials. *Biomaterials* **1986**, *7*, 221–224.
- Kanemitsu, K.; Kunishima, H.; Saga, T.; Harigae, H.; Imasaka, T.; Hirayama, Y.; Kaku, M. Residual formaldehyde on plastic materials and medical equipment following low-temperature steam and formaldehyde sterilization. *J. Hosp. Infect.* **2005**, *59*, 361–364.
- Rutala, W.A.; Weber, D.J. New disinfection and sterilization methods. *Emerg. Infect. Dis.* **2001**, *7*, 348–353.

19. Lerouge, S.; Wertheimer, M.R.; Yahia, L.H. Plasma sterilization: A review of parameters, mechanisms and limitations. *Plasmas Polym.* **2001**, *6*, 175–188.
20. Chau, T.T.; Kao, K.C.; Blank, G.; Madrid, F. Microwave plasmas for low temperature dry sterilization. *Biomaterials* **1996**, *17*, 1273–1277.
21. Dillow, A.K.; Dehghani, F.; Hrkach, J.S.; Foster, N.R.; Langer, R. Bacterial inactivation by using near- and supercritical carbon dioxide. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10344–10348.
22. Spilimbergo, S.; Dehghani, F.; Bertucce, A.; Foster, N.R. Inactivation of bacteria and spores by pulse electric field and high pressure CO₂ at low temperature. *Biotechnol. Bioeng.* **2002**, *82*, 118–125.
23. Castro, A.J.; Barbosa-Canovas, G.V.; Swanson, B.G. Microbial inactivation of foods by pulsed electric fields. *J. Food Process. Pres.* **1993**, *17*, 47–73.
24. Vadrot, C.; Darbord, J.C. Quantitative evaluation of prion inactivation comparing steam sterilization and chemical sterilants: Proposed method for test standardization. *J. Hosp. Infect.* **2006**, *64*, 143–148.
25. Tessarolo, F.; Caola, I.; Nollo, G.; Antolini, R.; Guarrera, G.M.; Caciagli, P. Efficiency in endotoxin removal by a reprocessing protocol for electrophysiology catheters based on hydrogen peroxide plasma sterilization. *Int. J. Hyg. Environ. Health* **2006**, *209*, 557–565.

Stress Fractures

Marko Pećina
Ivan Bojanić
Alan Ivković

Department of Orthopaedic Surgery, School of Medicine, University of Zagreb,
Zagreb, Croatia

S

INTRODUCTION

Stress fracture represents overuse injury of bone. It can be defined as a partial or a complete bone fracture that results from repeated application of a stress lower than the stress required to fracture the bone in a single loading. The first clinical description of stress fracture is credited to Breithaupt, a Prussian military physician. In 1855, he described the clinical picture and symptoms of metatarsal bone stress fractures in soldiers and termed the condition march fracture. The present understanding of stress fractures and the bone's response to loading has been advanced by numerous studies investigating the incidence of stress fractures in military personnel and athletes, and stress fractures have been described in nearly every bone of the human body. Stress fractures usually occur when excessive, repetitive loads on the bone cause imbalance between bone resorption and bone formation. Most common stress fractures occur in the lower-extremity weight-bearing bones, followed by pars interarticularis stress fractures of the lumbar spine, and then on some bones of the upper extremity. The origins of stress fractures are most likely site specific and depend on the bone density and geometry, the direction of the load, the vascular supply to the bone, the surrounding muscular attachments, the skeletal alignment, and the type of athletic activity. In addition to mechanical influences, systemic factors, including nutritional deficiencies, hormonal imbalances, collagen abnormalities, and metabolic bone disorders, can contribute to the development of stress fractures. Epidemiology, predisposing factors, pathophysiology, diagnostics, and treatment options of the stress fractures will be discussed in this entry.

EPIDEMIOLOGY

Stress fractures are common problems in sports medicine, comprising between 1.1% and 10% of all athletic injuries, while in the population of runners, they account for 15% of all injuries (although the data may vary from one medical institution to another).^[1–5] The

global incidence of bone stress injuries in sports has been estimated to be 2–4% in men and 7% in women.^[6]

Age

The average age of patients ranges from 19 to 30 yr. Considering the present-day tendency of including children and adolescents in top athletic activities, the number of overuse injuries in pediatric athletes is increasing. Nevertheless, stress fractures in pediatric athletes are less common than in adults.^[7] Hulkko and Orava surveyed 368 patients with stress fractures and found 9% among children younger than 15 yr of age, and 32% among those aged 16–19 yr.^[8] In addition, the distribution of stress fractures is somewhat different in pediatric athletes. The tibia is the predominant site of stress fractures in pediatric athletes with an incidence of 51%.^[8] Other sites, in descending order, are the fibula 20%, pars interarticularis 15%, femur 3%, metatarsal 2%, and the tarsal navicular 2%.^[8] The so-called epiphyseal stress fractures have been reported on the humerus, olecranon, distal radius, distal femur, and proximal tibia, although when compared to apophyseal overuse injuries (Osgood-Schlatter's disease, Sever's disease, etc.) they occur much more rarely.^[9]

Sex

Various studies have found that women have a higher incidence of stress fractures compared to men.^[10,11] Identical military conditions and physical training have shown that women are 3–12 times more often affected by stress fractures than men.^[12,13] This increased risk ranging from 1.5 to 3.5 times has also been observed in athletic populations.^[14] This difference may be due to biomechanical differences between the sexes and the female athletic triad (disordered eating, amenorrhea, and osteoporosis).^[15,16] Multiple studies have demonstrated that stress fractures occur more commonly in women with amenorrhea or oligomenorrhea than in eumenorrheic women.^[17] Significant sex differences have been noted in the site distribution of stress fractures. The tibia is the most common site

for stress fractures in men and women who participate in jumping and running sports. However, studies suggest that women develop more metatarsal and pelvic stress fractures and fewer fibular fractures than men.

Race

In a questionnaire-based study, the prevalence of stress fractures was much lower among black athletes than among white athletes.^[12] Military studies also have suggested this.^[13] These studies surmised that lower stress fracture risk in blacks may be related to their higher bone density or to their different biomechanical features that may protect against stress fracture development.

Type of Activity

The site and frequency of stress fractures in athletes depend to a large extent on the type of activity. Evidence shows that certain stress fracture sites are reported more often in some athletes, than in others, e.g., ribs stress fracture in golfers and rowers, tarsal navicular stress fracture in sprinters and high jumpers, etc. (see Table 1).

PREDISPOSING FACTORS

A cause must be sought for every overuse injury and predisposing factors should be carefully analyzed for every individual athlete. The cause may be quite obvious, such as a sudden increase of training quantity, poor footwear, or an obvious biomechanical abnormality. On the other hand, it can be subtler, such as running on a cambered surface, muscle imbalance, or leg-length discrepancy. The causes of overuse injuries are traditionally divided into intrinsic factors, which pertain to the individual itself and cannot be easily changed, and into extrinsic factors, which can be more easily influenced and modified.

Intrinsic Factors

Biomechanical factors

Research evidence shows that certain anatomical deviations on lower extremities, e.g., genu varum, tibia vara, forefoot varus, flatfeet, and cavus feet, by altering the distribution of load, contribute to the development of stress fracture.^[19] Varus alignments are common in athletes with stress fractures, especially with

Table 1 Sports and activities commonly associated with various stress fractures sites

Site of stress fracture	Sport or activity
Coracoid process of scapula	Trap shooting
Humerus	Throwing and racquet sports
Ulna—olecranon	Javelin throwing and pitching
Ulna—shaft	Racquet sports
Radius	Gymnastics
Ribs (I)	Basketball and ballet
Ribs (IV–IX)	Rowing and golf
Pars interarticularis	Gymnastics, dancing, weightlifting, and pole vaulting
Sacrum	Long-distance and marathon running
Pubic ramus	Long-distance and marathon running
Femur—neck	Long-distance running and gymnastics
Femur—shaft	Long- and middle-distance running and hurdling
Tibia—shaft	Running and ballet
Tibia—midshaft	Basketball, volleyball, pole vaulting, and high jumping
Tibia—medial malleolus	Basketball, high and long jumping, and sprinting
Fibula	Running, gymnastics, and figure skating
Tarsal navicular	Sprinting, hurdling, basketball, and high and long jumping
Metatarsal—general	Running, dancing, and gymnastics
Metatarsal—second	Ballet
Metatarsal—fifth	Basketball, soccer, football, and throwing (javelin and hammer)
Sesamoid bones of great toe	Running, dancing, gymnastics, and basketball

(From Ref. [18].)

femoral and tarsal stress fractures. Cavus feet is found more often in athletes with stress fracture of the metatarsal bones and the femur, while flatfeet with excessive pronation significantly contributes to the development of stress fractures of the tibia, fibula, and tarsal bones. Leg-length discrepancy may also predispose to the development of stress fracture. For instance, stress fractures of the femur, tibia, and metatarsal bone occur more often in longer legs, while stress fractures of the fibula are more common in shorter legs.

Muscles

The condition and tone of the skeletal muscles, i.e., their strength, are directly related to stress fracture. During a rapid increase of physical activity, the muscular tone also increases resulting in stresses stronger than those that the bone can withstand and a stress fracture may occur. In contrast to this, muscular weakness may predispose to the development of stress fracture by causing an increase or redistribution of stress to bone. Fatigued muscles have decreased contractility and their ability to absorb shock is diminished causing altered stress distribution and high-compression loads and shear stresses. Increased tightness and poor flexibility of muscles and tendons have also been regarded as major causes of chronic stress conditions. This is precisely why stretching and strengthening exercises of muscles are essential.

Extrinsic Factors

Training errors

Training errors are the most frequently encountered causes of stress fractures. There is vast research evidence showing that training errors cause stress fractures in as many as 22–75% of cases.^[5,10,14] They are characterized by “mileage mania” (excessive mileage), extremely intensive training, or too rapid changes in both the qualitative and the quantitative aspects of training. Studies in military recruits and athletes have revealed that modification of the training regimen to include rest periods intermittently with periods of strenuous activity significantly decreased stress fracture risk by providing time for bone microdamage to be repaired and by decreasing the load applied to the bone.

Surface

The type of ground for running (tartan, asphalt, grass, etc.), i.e., the surface quality of a playground (parquet, concrete, asphalt, etc.), is extremely important in the event of stress fracture.^[14] Running on hard surfaces

may be associated with an increased risk of stress fractures because of the mechanical shock introduced to the bone. During the assessment of two groups of runners having stress fractures, it has been found that 50% of them were running on hard surface.^[20] Running on an uneven surface may increase the risk of stress fracture by causing increased muscle fatigue and redistributing load to the bone. Running along the margin of the road, i.e., on a sloped surface, leads to the so-called short-long leg syndrome, which may cause tibial stress fracture.



Footwear

Athletic footwear is designed to reduce impact on ground contact and provide stability by controlling foot and ankle motion. Shoe age has been shown to be a better indicator of shock absorbing quality than shoe cost. It has been shown that training in shoes older than 6 mo increases the risk of stress fracture. There has been no association, however, between shoe cost and stress fracture risk. But new shoes can also lead to the development of stress fractures. Increased cushioning, which reduces the body's sensory-feedback mechanism (originating in the plantar surface of the foot), can lead to injury during running. This sensory disturbance has been described as a pseudoneuropathic condition; it also applies to other changes (running on a different surface, using a different tennis racquet, trying a different pitch, etc.).

PATHOPHYSIOLOGY

Bone is an adaptable tissue capable of repair, regeneration, and remodeling of itself in response to environmental, particularly mechanical, signals. Bone is exposed to both stress (i.e., load) and strain (i.e., deformation) with weight-bearing exercise, and it remodels in direct response to the forces applied. If the bone is subjected to repetitive cyclical loading, at a certain point (physiological threshold) reparative mechanisms are overwhelmed, microdamage begins to accumulate, and the process culminates in a stress fracture. Two competing, but not mutually exclusive, theories may explain the development of stress fractures.^[21] One theory holds that during the initial increase in exercise activity, the osteoblastic activity lags behind osteoclastic activity by a few weeks resulting in a period during which the bone is more susceptible to injury. Torsional and bending stress forces from repetitive use result in microfractures that consolidate into stress fractures. The other theory emphasizes strong and repetitive stress on the bone at the insertion point of muscles resulting

in focal bending stresses beyond the ability of the bone to tolerate.

DIAGNOSTICS

Similar to other overuse injuries, when attempting to diagnose a stress fracture, it is essential to have the possibility of such a condition constantly in mind. Thus, clinical diagnosis represents the basic procedure followed later on by other diagnostic methods.

Clinical Diagnosis

The main symptom of stress fracture is pain.^[1,3,5,7,21–23] Classically, the patient describes a history of insidious onset of pain in the area of the fracture that has its onset during activity, but disappears with rest. If the activity is continued, the pain may begin earlier, at an increasing intensity, in the course of that activity and persists after the activity. Later on, the pain increases so much that activity has to be stopped. Finally, the pain may be present despite cessation of the activity and at rest. It should be noted that night pain is a frequent complaint. A comprehensive history should include a review of the exercise or the training program, in particular any recent changes in the type or level of activity, as well as changes in surface, equipment (especially shoes), and technique. A history of previous similar injury or any other musculoskeletal injury should be obtained. In addition, the patient's general health, medications, diet, occupation, and personal habits should be assessed. A careful menstrual history should be obtained from all female patients. The hallmark of stress fracture is localized tenderness to palpation at the fracture site. Point tenderness may be best provoked over bones that can be easily palpated, such as the metatarsal bones or the fibula. For bones that are deep, such as the pelvis or femoral neck and shaft, pain may be elicited through gentle range of motion or specific diagnostic tests. Deformity is usually absent at the fracture site. Physical examination may also show swelling and erythema. Percussion of the affected bone distal to the fracture site causes pain. Muscle atrophy, weakness, and limited range of motion are usually not present. If the diagnosis is still in doubt after palpation, applying stress to the area involved will help to reproduce a patient's symptoms. Activities, such as walking, walking on toes, or running in place, can reproduce a patient's symptoms in the physician's office. Passive stretching can also elicit pain in the affected area. In this case, muscle pull rather than weight bearing induces stress on the bone. As part of a thorough physical examination, the practitioner

should assess the athlete's flexibility, lower-limb alignment (including leg lengths), foot structure, and motor function.

Radiography

Standard x-ray imaging may well show the bone lesions in only 30–70% of stress fracture cases.^[3,5,24] It has also been found that from the onset of pain to the moment when alterations become visible on a standard x-ray film, it takes from 2 weeks to even 3 mo, depending on the fracture site. For certain locations, e.g., tarsal bones, femur, or spine, it is often not possible to identify any changes whatsoever, even if followed up for a longer period of time. The first noticeable radiographic signs of stress fracture in cortical bone are periosteal reaction (in diaphyseal fractures) and medullar sclerosis and endosteal thickenings (in metaphyseal fractures). These changes are not so easily recognized, i.e., they can be identified with greater certainty only on tangential radiographs or tomograms. The initial changes of cancellous bone stress fractures, known as trabecular fractures, are characterized by hardly noticeable flake-like patches of new bone that may be seen 10–21 days following the onset of symptoms (see Fig. 1).^[5] Later, the patches transform into wider cloud-like areas of mineralized bone, and then an area of focal linear sclerosis oriented perpendicular to trabeculae appears within the bone, being among the most characteristic signs of cancellous bone stress fracture. Plain radiographs are necessary in the initial assessment to rule out other causes of localized bone pain, such as complete fracture, infection, or malignancy. Special views must be used in some cases, e.g., orthograde projection for tarsal navicular bone, oblique views for lumbar spine, etc. It is important to remember that stress fractures are dynamic in nature, and the radiographic appearance is also a continuum. However, despite a negative radiograph, the diagnosis of a stress fracture should be considered if the prodromal signs and symptoms and clinical index of suspicion are still high.^[3,5,24]

Scintigraphy

A triple-phase bone scintigraphy is highly sensitive, but it lacks specificity and anatomic resolution. On account of higher specificity of x-ray findings, bone scan is recommended only when the radiograph seems negative in spite of reasonable doubt that stress fracture might be the issue. According to this, the triple-phase technetium 99m diphosphate bone scan is the current gold standard for the diagnosis of stress fracture.^[25] Bone scan may demonstrate bony changes within 6–72 hr following the appearance of initial stress



Fig. 1 Radiogram (AP view) demonstrating stress fracture of the fibula (arrow).

fractures symptoms. A typical scintigraphic finding of a stress fracture is characterized by an oval or a spindle, sharply bordered focus of increased uptake that involves one cortex or occasionally extension of the width of the bone (see Fig. 2). In stress fractures, all three phases (the angiogram phase, the blood-pool phase, and the delayed image phase) of the triple-phase bone scan are positive.^[25] Some other overuse soft-tissue injuries would be positive only in the angiogram and blood-pool phase, thereby enabling bony pathology and pathology involving soft tissue to be differentiated. The bone scan is diagnostic of the stress fracture if focal isotope uptake is seen in the area of interest on the third phase of the scan. The asymptomatic accumulation of radionuclide, the so-called stress reaction, only indicates remodeling of the bone caused by a physical stress. Because scintigraphic finding may remain positive for several months, clinical and radiographic findings are used as discriminative ones for the follow-up of stress fracture treatment.

Computed tomography

Computed tomography (CT) may be used in conjunction with scintigraphy to fully delineate the

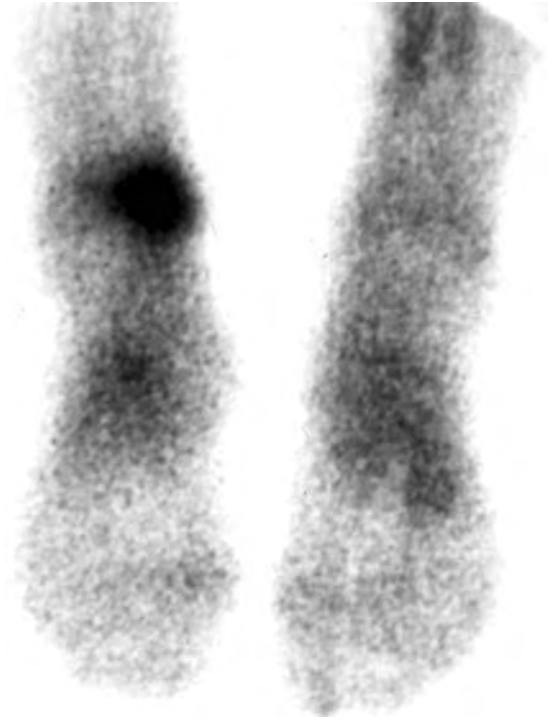


Fig. 2 Bone scan (AP view) demonstrates focal uptake in the right medial malleolus.

pathologic process. Computed tomography is very useful for differentiating conditions that may mimic stress fracture: various malignancies, osteomyelitis with Brodie's abscess, etc. Computed tomography scans can show stress fractures that are not apparent on plain radiographs and may also be used to evaluate healing when plain radiographs do not adequately demonstrate the fracture. They are particularly helpful in identifying stress fractures of the sacrum, pelvis, and spine, and in proposing treatment of tarsal navicular stress fracture (see Fig. 3).^[26] A disadvantage of CT is that, in many circumstances, scanning is limited to the axial plane, which does not allow views of transversely oriented fractures. An alternative, helical, or spiral CT can be used to minimize this limitation.

Magnetic resonance imaging

Magnetic resonance imaging (MRI) is proving useful in identifying stress fractures.^[24,27] It is highly sensitive for the diagnosis of stress injury to bone, and usually allows depiction of abnormalities several weeks before the development of radiographic alterations. It has comparative sensitivity to bone scintigraphy for detection of bone abnormalities. Besides having higher specificity than bone scintigraphy in distinguishing bone involvement from soft-tissue injuries, MRI is helpful in grading the stage of certain stress fractures and, therefore, predicting the time to recovery. In

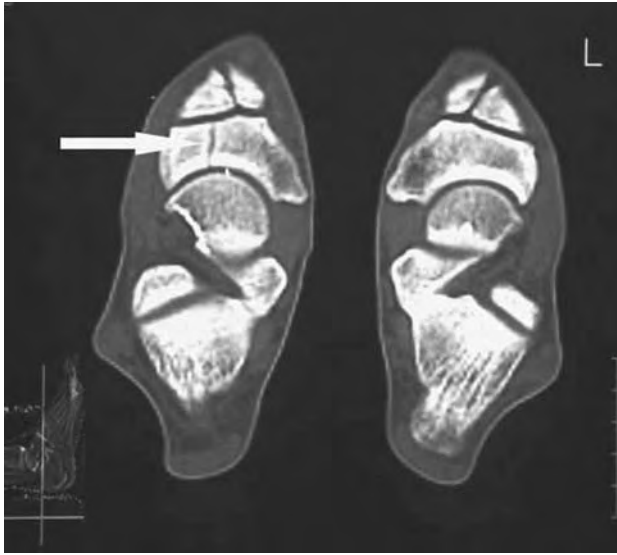


Fig. 3 Computed tomography scan shows navicular tarsal stress fracture (arrow).

addition, MRI avoids exposure to ionizing radiation and requires less time to complete the test than triple-phase bone scan. Recent advances in MR availability, including the increasing availability of relatively low-cost dedicated extremity scanners, likely will expand further the application of MR for the detection of bone abnormalities, and could challenge conventional radiography as the primary diagnostic tool used to detect stress injuries to the bone.

DIFFERENTIAL DIAGNOSIS

The two major reasons for failure to diagnose a stress fracture are neglect of it as a diagnostic possibility and acceptance of negative radiography, which has been taken prematurely or examines the region inadequately. The differential diagnosis for stress fractures is extensive and includes stress reaction, periostitis, muscle strain, infection, neoplasm, bursitis, exertional compartment syndrome, and nerve entrapment. In children and adolescents, it is of utmost importance to differentiate stress fracture from osteogenic sarcoma and Ewing's tumor. A follow-up roentgenogram, repeated after 2–3 weeks, usually solves the problem of diagnosis. Looser's zones of transformation should be differentiated from stress fractures. They are incomplete stress fractures, which heal with callus, lacking in calcium. This is a pseudo-fracture seen in osteomalacia. On x-ray, it appears as a thin, translucent band, about 2 mm in width, which runs perpendicular to the surface of the bone extending from the cortex inward. Another condition that should be differentiated from stress fractures is Milkman's

syndrome—a generalized bone disease marked by multiple transparent stripes of absorption in the long and flat bones. Differential diagnosis of stress fracture and overuse injuries of soft tissue cannot be based solely on the clinical findings and radiological diagnosis; it also requires scintigraphy, and sometimes even MRI. A negative scintigraphy rules out bone stress injuries with great certainty. For example, in medial-tibial stress syndrome, an increase in the uptake is seen in a large area on the surface of the shinbone. The signal is seen only in the collecting phase, but not with angiography or bone-pole phases in connection with bone stress injuries. Magnetic resonance imaging shows swelling only in the periosteal area.^[5] Osteoid osteoma is commonly mistaken for a stress fracture, because it presents with pain and discrete focal area of increased uptake on bone scintigraphy. Two distinguishing features of osteoid osteoma are presence of night pain and relief of pain with the use of aspirin.

TREATMENT

The foundation of treatment for stress fractures is activity modification. Interruption of the cycle of overload is essential to permit the bone to heal. The main treatment of stress fractures is rest from the offending athletic activity; a concept known as “relative rest.”^[5,7,22,23] There is a group of stress fractures, called high-risk stress fractures, that requires additional treatment to relative rest.^[28] High-risk stress fractures include those in the femoral neck (tension side), patella, anterior cortex of the tibia, medial malleolus, talus, tarsal navicular, fifth metatarsal, and great toe sesamoids.^[26] The duration of treatment may vary according to the individual patient, the bone involved, and the particular sport. For most patients with stress fractures, the period of relative rest may last 4–12 weeks. Compliance is critical to the success of treatment. In some cases where the activities of daily living are painful, it may be necessary for the patient to be non-weight bearing or partially weight bearing on crutches for a period of time. Exercise is very important to prevent detraining in affected athletes. General conditioning can be maintained by exercising other areas of the body, and by partaking in alternate training, such as water running, swimming, cycling, etc. The alternate activity must not produce pain or tenderness. When the athlete has been free from pain for 2–3 weeks and full-weight bearing is normal, they may gradually return to their sport. The athlete must have no bony tenderness at the time of resumption of the sport. Furthermore, the athlete must be cautioned to resume the sport at a frequency, intensity, and duration well below the level that previously produced symptoms. In general, increasing the volume or

intensity of training, no more than 10% per week, will avoid recurrence of the stress fracture. The patient must be sequentially monitored during recovery. If they develop any pain during athletic participation or tenderness afterward, they should decrease or modify the activity for an additional 2-week period.^[26] In rare instances, surgical treatment may be indicated if there is a high-risk fracture, the stress fracture has progressed to a complete fracture, or the fracture has become an established nonunion.

OVERVIEW OF THE MOST COMMON STRESS FRACTURES

Various bones have very different mechanical and biological individual properties that affect their healing and treatment. Only the most common stress fractures will be discussed in this section, but the physician should be aware that the general treatment principles mentioned below apply to all stress fractures.

Tibia

The tibial shaft is the most common stress fracture site.^[3,5,10,22] Unfortunately, shin pain is common among athletes and can result from a variety of causes, including tibial periostitis (i.e., shin splints) and exertional compartment syndromes (a potentially serious condition). A careful history is necessary in distinguishing between these entities. Pain that occurs early in the exercise program and then improves with ongoing activity suggests periostitis. Pain precipitated by the exercise that worsens progressively with continued activity may herald a stress fracture. Stress fractures of the tibial shaft on the posteromedial aspect (compression side of the bone) are most common in sports that include running, such as distance running, soccer, basketball, and ballet dancing.^[3,5,10,22] The main symptom is pain appearing during activity and disappearing during rest. With time, it may be present with activities of daily living, or even at night. Physical examination reveals painful tenderness to palpation and percussion at the fracture site, while at a later stage callus may be palpated. Occasionally, swelling is present. Radiological diagnosis is possible only 2–5 weeks following the appearance of pain, when callus as well as the fracture line may be seen (usually oblique radiographs are also required) (see Fig. 4).

Besides, for early diagnosis, the bone scan helps to differentiate stress fracture from the posterior tibialis muscle syndrome, tibial periostitis, and exertional chronic compartment syndrome. Treatment is rest, in which a period of non-weight bearing on crutches for pain relief is sometimes required, until the patient is



Fig. 4 Radiogram (AP view) of the tibia demonstrating stress fracture at the proximal metaphyseal-diaphyseal junction (arrow).

pain free and has no bony tenderness. Most tibial stress fractures heal within 6–8 weeks, with an average time to return to full sports, of 8–12 weeks. The atypical stress fracture of the anterior cortex of the midshaft of the tibia (tension side of the bone) occurs almost exclusively in athletes performing repetitive jumping and leaping activities. This type of stress fractures has been reported in basketball players, runners, pole vaulters, high-jumpers, volleyball players, handball players, figure skaters, and ballet dancers.^[17,29] These stress fractures require prolonged healing time. Initial treatment consists of rest with or without immobilization for a minimum of 3–6 mo. If after 3–6 mo there is no evidence of healing either clinically or radiologically, surgical management is indicated. Another variety of atypical tibial stress fracture occurs in the medial malleolus. Stress fractures of the medial malleolus have been reported in athletes participating in running and jumping activities, i.e., in basketball players, hurdlers, high- and long-jumpers, sprinters, and gymnasts. Nondisplaced stress fractures of the medial malleolus should be treated with immobilization for a period of 6 weeks, while displaced fractures require open reduction and internal fixation.

Metatarsal Bones

The metatarsals are the second most frequent stress fracture site, common especially among military

recruits, distance runners, and ballet dancers. The second metatarsal is injured most frequently of all the metatarsals, followed in order by the third, first, fourth, and fifth metatarsals. In more than 80% of metatarsal bone stress fractures, the shaft of the second and the third metatarsal bone is affected.^[30] Less common locations include the base of the second metatarsal and the base of the fifth metatarsal bone. Metatarsal shaft stress fractures are typical in runners, dancers, basketball players, soccer players, figure skaters, and gymnasts. Diagnosis is very simple once the stress fracture is suspected. Patients complain of forefoot pain that is aggravated by running, jumping, or dancing activity. Although the pain is not severe at first, it gradually worsens when activity is undertaken. Physical examination reveals tenderness over the involved bone and sometimes a mass can be palpated along the shaft of the involved bone, which represents fracture callus. Radiographic changes may not be evident for at least 2–3 weeks following the onset of patients' symptoms, but thereafter a subtle fracture line with periosteal reaction may be present. If radiographs do not confirm that a stress fracture is present, a bone scan can be conducted. Rest from athletic activities for 4–6 weeks and immobilization in a boot cast with a heel for walking (2–4 weeks) yield good results, i.e., complete bone healing. Patients are observed clinically and allowed to commence a graduated activity program once the symptoms have resolved and there is no local tenderness at the fracture site. Stress fractures of the proximal fifth metatarsal distal to the tuberosity, also known as Jones' fracture, should be pointed out as a separate entity, especially as it tends toward delayed union, nonunion, and refracture. The Jones' fracture is most common in basketball players, football players, soccer players, runners, and throwers (javelin and hammer). The main symptom is pain in the lateral foot region. Physical examination reveals pain on palpation in the proximal end of the fifth metatarsal (about 1.5 cm

distally from the tuberosity). Pain is exacerbated by inversion of the foot. Based on radiography staging, immobilization in a short-leg cast with non-weight bearing for 6 to 8 weeks is suggested for early lesions, and surgery is suggested for advanced lesions as well as for nonunions.^[31] Choices for surgical intervention include intramedullary screw fixation or corticocancellous bone grafting (see Fig. 5). Currently, the most popular choice in athletes is intramedullary screw fixation.

Femur

A stress fracture of the femur may be located in the neck, the shaft, and supracondylar and condylar regions. The most common stress fracture of the femur is that of its neck, but not for its frequency as much as for the resulting complications: delayed union, nonunion, dislocation of fragments, and avascular femoral head necrosis.^[32] Owing to this, the primary goal of management should be to prevent complications through early diagnosis and careful treatment. Stress fracture of the femoral neck most often occurs in military recruits during basic training and in long-distance runners. However, it has also been reported in basketball players and gymnasts. The initial and the most frequent symptom is anterior groin (inguinal) pain. The pain is often exacerbated by weight bearing or activity, resulting in an antalgic gait. Night pain is occasionally present. Physical examination reveals pain at the extremes of gentle passive hip motion and slight limitation of flexion and internal rotation. Bone tenderness is often minimal owing to the depth of soft tissue overlying the femoral neck. Axial compression or percussion over the greater trochanter may also elicit pain. An additional test that has been found to be helpful is the "hop test" in which the patient attempts to hop on the injured leg, inevitably reproducing the pain that these patients experience if an undisplaced



Fig. 5 Postoperative roentgenogram of the stress fracture of the fifth metatarsal bone (Jones fracture) (white arrow) managed with intramedullary screw fixation.

stress fracture is present. If physical findings suggest a femoral neck stress fracture, it should be considered present until proved otherwise. If the diagnosis is suspected, radiographs may be obtained, keeping in mind that changes (when present) normally lag behind the onset of symptoms by 2–4 weeks (see Fig. 6).

Bone scan or MRI should be performed to make an early, definitive diagnosis (see Fig. 7). Compared with bone scan, MRI has a similar sensitivity, but an improved specificity, and is becoming the diagnosis procedure of choice. In the case of a normal radiograph and a positive bone scan, or of sclerosis without overt fracture line, nonoperative treatment will suffice. This comprises absolute bed rest until the pain resolves, followed by advancement from partial- to full-weight bearing on crutches as symptoms permit. Once the patient is free of pain, he is allowed to progress to a cane and then to unprotected weight bearing. A progressive walking and then running program is prescribed and the patient is returned over several months to full activity. If the patient presents with or develops an undisplaced or displaced cortical crack more aggressive treatment is required, i.e., internal fixation. Displaced femoral neck stress fracture is an orthopedic emergency. Early, accurate reduction and internal fixation are essential.

Lumbar Spine

Stress fractures of the lumbar spine occur most often in the pars interarticularis, although they may also be



Fig. 6 AP radiograph of the left hip demonstrates stress fracture of the femoral neck (arrow).

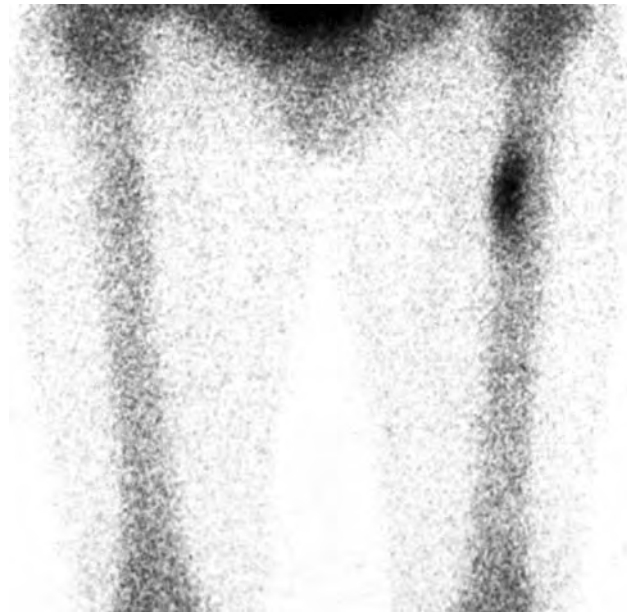


Fig. 7 Bone scan demonstrating intense radiotracer uptake on the proximal third of the femoral shaft.

located in the pedicle or lamina (see Fig. 8). The most common vertebral level affected is the L5 segment followed by the L4 and L3 levels. Stress fracture of the pars interarticularis is especially common in sports that involve repeated hyperextension of the lower back, such as gymnastics, dancing, weightlifting, hockey, pole vaulting, football, basketball, and running.^[33]

The basic signs of pars interarticularis stress fracture comprise absence of periosteal reaction and the great possibility of refracturing, i.e., development of spondylosis and, sometimes even, spondylolisthesis. The patients typically present with localized, unilateral low-back pain. The pain worsens with sports activities, especially extension and hyperextension, and is relieved by rest. The one-leg lumbar hyperextension test (commonly called the stork test), in which the patient is asked to hyperextend the back while standing on the ipsilateral leg of the affected side, usually reproduces the pain. They also often have diminished straight-leg raising secondary to hamstrings spasm. Radiographic assessment should include anteroposterior, lateral, and oblique views of the lumbar spine. Many stress fractures may be too subtle to detect radiographically, and require nuclear imaging. Bone scanning with single-photon emission tomography has been particularly helpful in delineating stress injury to the pars interarticularis.^[34] Advanced imaging with CT or MRI may be needed to ascertain the acuity of the lesion, assist in identifying a particular pars lesion as potentially symptomatic, and exclude other spinal pathology that may be present. Early recognition of the fracture and rest from athletic activities together with thoracolumbar orthosis are essential for good treatment

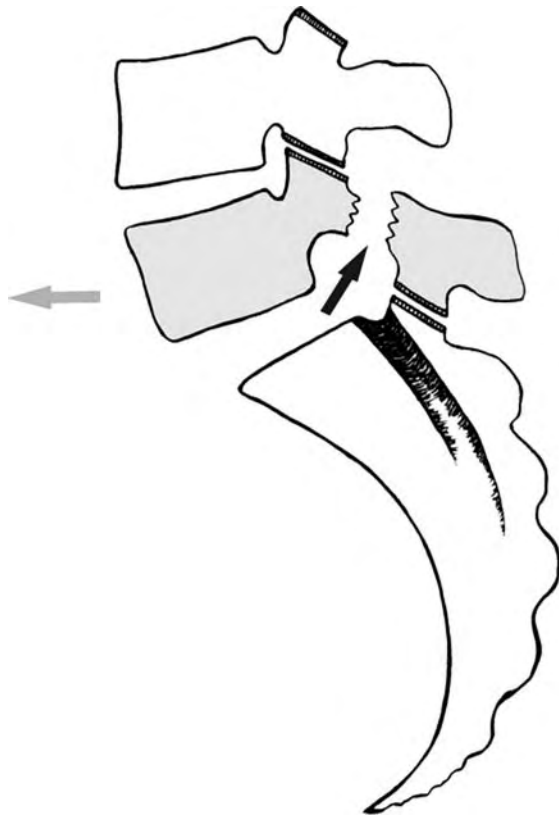


Fig. 8 Spondylolytic defect in the vertebral arch (black arrow) causes spondylolisthesis (gray arrow) of the fifth lumbar vertebra.

outcome.^[35] When the athlete is pain free, he or she may begin a graduated rehabilitation program emphasizing a spinal stabilization program consisting of abdominal and paraspinal muscle strengthening.

CONCLUSIONS

Stress fractures can cause long-lasting damage, especially in athletes and military recruits, if not diagnosed and treated properly. Diagnosis has been traditionally based on clinical, radiographic, and scintigraphic examinations, but MRI has become increasingly important. High-resolution MRI is particularly valuable for the grading of bone stress injuries. The clinician should be aware of the wide range of bone stress injuries and available diagnostic methods. Early diagnosis is essential for avoiding long-lasting complications. Most bone stress injuries heal with closed treatment, but surgery is necessary in some cases. They heal well if the diagnosis is not delayed and the treatment is adequate. As in most of the injuries belonging to the overuse group, prevention is the keyword. To prevent stress fractures and other overuse injuries, it is of utmost importance that the athlete, coach, and

physician cooperate. It is essential to identify the group at higher risk of stress fracture, most commonly, athletes with biomechanical irregularities in the lower extremities. These athletes should be kept under constant surveillance and provided with various orthotic devices. In extreme cases, surgery may be required to correct biomechanical disorders. Female athletes and their coaches and parents need to be alerted to the adverse effects of eating disorders and hormonal abnormalities. Identification of female athletes with “female athlete triad” is especially important so that early interventions may be initiated. Furthermore, athletic and training activities should be adapted to the age and abilities of individual athletes, i.e., maximum load should be achieved gradually. Also, rest periods should be incorporated with periods of strenuous activity. Athletic requisites are also very important for the prevention of stress fractures, especially the condition of the running shoes and quality of playground surface. An important preventive measure also entails stretching and strengthening exercises, which also diminish stressful forces on the bone.

ARTICLES OF FURTHER INTEREST

Bone Fracture Fixation; Bone Remodeling; Zirconia Ceramics

REFERENCES

1. Lassus, J.; Tulikoura, I.; Konttinen, Y.T.; Salo, J.; Santavirta, S. Bone stress injuries of the lower extremity: a review. *Acta Orthop. Scand.* **2002**, *73*, 359–368.
2. Brukner, P.; Bradshaw, C.; Khan, K.; White, S.; Crossley, K. Stress fractures: a review of 180 cases. *Clin. J. Sport Med.* **1996**, *6*, 85–89.
3. Jones, B.H.; Harris, J.M.; Vinh, T.; Rubin, C. Exercise-induced stress fractures and stress reactions of bone: epidemiology, etiology, and classification. *Exerc. Sport Sci. Rev.* **1989**, *17*, 379–422.
4. Matheson, G.O.; Clement, D.B.; McKenzie, D.C.; Taunton, J.E.; Lloyd-Smith, D.R.; MacIntyre, J.G. Stress fractures in athletes. A study of 320 cases. *Am. J. Sports Med.* **1987**, *15*, 46–58.
5. Bennell, K.L.; Brukner, P.D. Epidemiology and site specificity of stress fractures. *Clin. Sports Med.* **1997**, *16*, 179–196.
6. Johnson, A.W.; Weiss, C.B.; Wheeler, D.L. Stress fractures of the femoral shaft in athletes—more common than expected. *Am. J. Sports Med.* **1994**, *22*, 248–256.
7. Coady, C.M.; Michelli, L.J. Stress fractures in the pediatric athlete. *Clin. Sports Med.* **1997**, *16*, 225–238.
8. Hulkko, A.; Orava, S. Stress fractures in athletes. *Int. J. Sports Med.* **1987**, *8*, 221–226.

9. Ivković, A. Overuse injuries in young athletes. In *Overuse Injuries of the Musculoskeletal System*, 2nd Ed.; Pećina, M.M., Bojanić, I., Eds.; CRC Press: Boca Raton, 2004; 373–389.
10. Dusek, T.; Pecina, M.; Loncar-Dusek, M.; Bojanić, I. Multiple stress fractures in a young female runner. *Acta Chir. Orthop. Traumatol. Cech.* **2004**, *71*, 308–310.
11. Dusek, T. Influence of high intensity training on menstrual cycle disorders in athletes. *Croat. Med. J.* **2001**, *42*, 79–92.
12. Brudvig, T.J.S.; Gudger, T.D.; Obermeyer, L. Stress fractures in 295 trainees. A one-year study of incidence as related to age, sex, and race. *Mil. Med.* **1983**, *148*, 666–667.
13. Jones, B.H.; Bovee, M.W.; Harris, J.M.; Cowan, D.N. Intrinsic risk factors for exercise-related injuries among male and female army trainees. *Am. J. Sports Med.* **1993**, *21*, 705–710.
14. Brunet, M.E.; Cook, S.D.; Brinker, M.R.; Dickinson, J.A. A survey of running injuries in 1505 competitive and recreational runners. *J. Sports Med. Phys. Fitness* **1990**, *30*, 307–315.
15. Ivkovic, A.; Bojanic, I.; Ivkovic, M. Trijas sportašica. *Liječ. Vjesn.* **2001**, *123*, 200–206.
16. Arendt, E.A. Stress fractures and the female athlete. *Clin. Orthop.* **2000**, *372*, 131–138.
17. Dubravčić-Šimunjak, S.; Pećina, M.; Kuipers, H.; Moran, J.; Hašpl, M. The incidence of injuries in elite junior figure skaters. *Am. J. Sports Med.* **2003**, *31*, 511–517.
18. Bojanić, I.; Pećina, H.I.; Pećina, M. Prijelomi zamora. *Arh. Hig. Rada Toksikol.* **2001**, *52*, 471–482.
19. Krivickas, L.S. Anatomical factors associated with overuse sports injuries. *Sports Med.* **1997**, *24*, 132–146.
20. Walter, S.D.; Hart, L.E.; McIntosh, J.M.; Sutton, JR. The Ontario cohort study of running-related injuries. *Arch. Intern. Med.* **1989**, *149*, 2561–2564.
21. Pećina, M.M.; Bojanić, I. Stress fractures. In *Overuse Injuries of the Musculoskeletal System*, 2nd Ed.; Pećina, M.M., Bojanić, I., Eds.; CRC Press: Boca Raton, 2003; 315–349.
22. Knapp, T.P.; Garrett, W.E., Jr. Stress fractures: general concepts. *Clin. Sports Med.* **1997**, *16*, 339–356.
23. Reeder, M.T.; Dick, B.H.; Atkins, J.K.; Pribis, A.B.; Martinez, J.M. Stress fractures. Current concepts of diagnosis and treatment. *Sports Med.* **1996**, *22*, 198–212.
24. Spitz, D.J.; Newberg, A.H. Imaging of stress fractures in the athlete. *Radiol. Clin. North Am.* **2002**, *40*, 313–331.
25. Deutsch, A.L.; Coel, M.N.; Mink, J.H. Imaging of stress injuries to bone: radiography, scintigraphy, and MR imaging. *Clin. Sports Med.* **1997**, *16*, 275–290.
26. Bojanić, I.; Pećina, M.M. Traitement conservateur des fractures de contrainte du scaphoïde tarsien chez le sportif. *Rev. Chir. Orthop.* **1997**, *83*, 133–138.
27. Ishibashi, Y.; Okamura, Y.; Otsuka, H.; Nishizawa, K.; Sasaki, T.; Toh, S. Comparison of scintigraphy and magnetic resonance imaging for stress injuries of bone. *Clin. J. Sport Med.* **2002**, *12*, 79–84.
28. Boden, B.P.; Osbahr, D.C. High-risk stress fractures: evaluation and treatment. *J. Am. Acad. Orthop. Surg.* **2000**, *8*, 344–353.
29. Pećina, M.; Bojanić, I.; Dubravčić, S. Stress fractures in figure skaters. *Am. J. Sports Med.* **1990**, *18*, 227–279.
30. Weinfeld, S.B.; Haddad, S.L.; Myerson, M.S. Metatarsal stress fractures. *Clin. Sports Med.* **1997**, *16*, 319–338.
31. Nunley, J.A. Fracture of the base of the fifth metatarsal. The Jones fracture. *Orthop. Clin. North Am.* **2001**, *32*, 171–180.
32. Egol, K.A.; Koval, K.J.; Kummer, F.; Frankel, V.H. Stress fractures of the femoral neck. *Clin. Orthop.* **1998**, *348*, 72–78.
33. Moeller, J.L.; Rifat, S.E. Spondylolysis in active adolescent. *Phys. Sports Med.* **2001**, *29*, 27–32.
34. Sys, J.; Michielsen, J.; Bracke, P.; Martens, M.; Verstreken, J. Nonoperative treatment of active spondylolysis in elite athletes with normal x-ray findings: literature review and results of conservative treatment. *Eur. Spine J.* **2001**, *10*, 498–504.
35. Bellah, R.D.; Summerville, D.A.; Treves, S.T.; Micheli, L.J. Low-back pain in adolescent athletes: detection of stress injury to the pars interarticularis with SPECT. *Radiology* **1991**, *180*, 509–512.

Supercritical Fluid Processing

Ihtesham U. Rehman

Department of Materials, Interdisciplinary Research Centre in Biomedical Materials, Queen Mary University of London, London, U.K.

Jawwad A. Darr

Department of Chemistry, University College London, London, U.K.

Alireza Moshaverinia

Section of Oral Biology, College of Dentistry, Ohio State University, Columbus, Ohio, U.S.A.

INTRODUCTION

Biomaterials are synthetic and natural materials that are in contact with biological systems or used in medical devices. The field of biomaterials has continually evolved over almost more than half a century and has found many uses in medicine and biology.^[1] The cohesion of these applications is the interaction between synthetic or modified natural materials and biological systems. Biomaterials are not often used on their own but are more commonly combined into medical devices such as dental fillings, bone grafts, wires plates and pins for bone repair, artificial skin, artificial blood vessels, total artificial hearts, pacemakers, total artificial joint replacements, and implants.^[1]

The high toxicity of some organic solvents in food processing, drug formulation, and biomedical materials processing created interest in developing new and clean technologies for processing.^[2] Supercritical fluids such as carbon dioxide and water have been particularly of interest^[7] for many years. The commercial applications of supercritical fluid (SCF) technologies remained restricted, however, because of high costs of equipment and unfamiliarity with equipment operation. More recently, SCF processing contributed to both the production of biocommodities, such as extraction of fermentation products, and the production of high-value-added biomolecules such as proteins via purification and processing.^[3] In the following chapter the fundamentals of supercritical CO₂ and water-based technologies will be explained, and different applications will be illustrated.

SUPERCritical FLUIDS

An SCF is any substance for which the temperature and pressure are higher than critical values. Below the critical point, the fluid can coexist in both gas and liquid phases, but above the critical point there can be

only one phase.^[3] SCFs are also of interest because their properties are intermediate between gases and liquids and are readily adjustable. The pressure–temperature phase diagram for a pure substance (CO₂) is presented in Fig. 1, where the four phases of a substance (solid, liquid, gas, and SCF) are illustrated. In the SCF state the physical properties of the fluid, such as density and viscosity, become transient between that of the gas and liquid phases of the substance.^[3]

Unlike that of liquid solvents, the density of SCFs can be changed by altering the pressure and temperature of the fluid and therefore the dissolution power of a SCF can be enhanced. SCFs can have gas-like diffusivities, which have important implications for reaction kinetics, as well as liquid-like densities, which allow the deliverance for many compounds.^[4]

The most widely used SCF is carbon dioxide because of its relatively low critical pressure and temperature (T_c of 31.1°C and P_c of 7.38×10^4 MPa, respectively) and because it is relatively non-toxic and inflammable.^[7] Table 1 presents commonly used SCFs and their related critical values. Consideration of the properties of supercritical CO₂ suggests that it is suitable for extraction of different materials and it can be used as a reaction medium for polymerization. Supercritical carbon dioxide is an excellent non-polar solvent for many organic compounds; it has solvent power similar to a light hydrocarbon for most solutes. However, fluorinated compounds are often more soluble in supercritical CO₂ than in hydrocarbons. This increased solubility is important for polymerization reactions (such as biopolymers or scaffolds for tissue engineering). Organic compounds such as alkenes, alkanes, aromatics, ketones, and alcohols up to a relative molecular mass of around 400 dissolve in SCF CO₂. Very polar molecules, such as amino acids, sugars, and most inorganic salts, are insoluble in this fluid.^[4] However, there are some disadvantages in the usage of SCF technology as a reaction medium, such as expensive equipment and a complicated system thermodynamics.^[5]

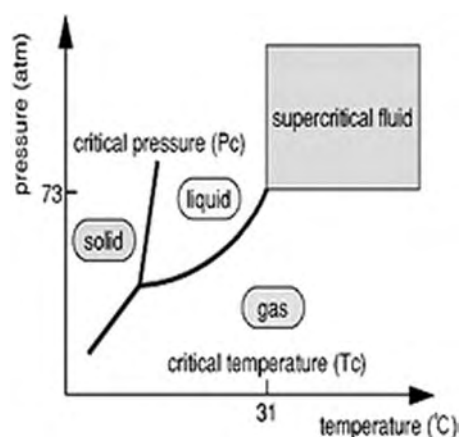


Fig. 1 Phase diagram for CO₂. T_c , critical temperature; P_c , critical pressure.

Another commonly used fluid in supercritical technology is water. Its critical temperature is 374°C and critical pressure is 220 bar. Water has been growing in importance as a medium for chemical reactions recently because of the potential for synthesis of different materials. The most important advantage of supercritical water technology is the possibility of varying the properties of the reaction medium over a wide range, basically by altering the pressure and temperature and optimizing the reaction in this way without changing solvent. Moreover, the reaction kinetics can be influenced in the supercritical region by varying the pressure. Additionally, many non-polar organic substances such as cyclohexane and different kinds of gases are highly soluble in supercritical water (critical pressure and temperature = 374 and $P_c = 2.2 \times 10^4$ MPa, respectively) so that mass transfer restrictions according to phase boundaries do not apply. However, there are disadvantages to working at high pressures, such as high investment and possibly energy costs. The other problem is related to corrosion, and expensive alloys frequently are used.

One of the main applications of SCFs regarding biomaterials is the precipitation of solid products from SCF solutions. Different techniques provide a high

degree of control over the precipitation process, such as rapid expansion, antisolvent precipitation, and atomization of fluid or liquid-like phase.^[6] Considerable efforts have been directed toward the processing of biomaterials and pharmaceuticals using SCF by using different methods. The fundamentals of these methods are mentioned below.

Rapid Expansion of Supercritical Solutions (RESS)

RESS has a considerable role in precipitation of different materials. By application of RESS, a large number of fine-powdered materials can be precipitated. The apparatus used for this procedure consists of a reservoir of supercritical solution and an expansion nozzle. Supercritical containing solute is subjected to rapid expansion over the nozzle by reducing the pressure. This process results in supersaturation and particle formation. The advantage of the RESS technique is that it has the ability to produce narrow particle size distribution, which is important in drug delivery systems. However, the main disadvantage of RESS is that it can only apply to certain kinds of materials that are soluble in supercritical CO₂. Therefore, a large amount of fluid is needed in order to produce a small amount of particles. As a result, the technique is relatively expensive and can be used for laboratory-based processes involving high-value-added products such as pharmaceuticals.^[6] The RESS process can produce pure and high-quality drug particles. However, as mentioned previously, because of the extremely low solubility of polar drugs in supercritical CO₂ (SC-CO₂), RESS has limited commercial applicability. To overcome this major limitation, a modified process for rapid expansion of supercritical solution with solid cosolvent (RESS-SC) is recommended, which uses a solid cosolvent. The recommended RESS-SC process solid cosolvent delays the particle growth and, as a result, hinders the formation of small nanoparticles.^[6,7] During the rapid expansion of solutions in SCF CO₂ (RESS), the sudden loss of solvent powder results in the precipitation of fine particles.

Gas Antisolvent (GAS)

The applications of SCF CO₂ as an antisolvent have become a main area of industrial interest. The use of supercritical antisolvents is similar to layering poor solvents above a solution to induce precipitation and crystallization in normal synthetic chemistry. There are different types of this method, all of which involve injecting the organic solution into a vessel containing either dense CO₂ or co-injection of CO₂ and organic

Table 1 Supercritical parameters for commonly used supercritical fluids

Substance	Critical temperature, T_c (°C)	Critical pressure, P_c (bar)
CO ₂	31.1	73.8
Fluoroform	25.9	48.2
Water	374.0	220.6
Ammonia	132.4	113.2
Ethane	32.2	48.7

solution in a typical vessel. The advantage of this method is that the flow rate of the fluid and the temperature and density of the supercritical CO₂ can be optimized in order to control the morphology and particle size of the precipitated materials.^[10] The principle of this method is similar to that of conventional liquid antisolvent crystallization. Regarding high miscibility of organic solvent in SCF, volume expansion of the solvent leads to a simultaneous reduction in the density and solvent capacity. As a result, the levels of particle formation, supersaturation, and solution recrystallization increase. This method is suitable for crystallization of materials that are slightly soluble in SCF.^[7]

Supercritical Antisolvent (SAS)

In the SAS process, a solution of a drug that is dissolved in organic solvent is sprayed through a capillary tube into a continual SCF stream. Controlled-release drug delivery and protein particles preparation are examples of applications of the SAS method.^[7]

APPLICATIONS OF SCFS IN BIOMATERIALS

Processing of Biological Materials

Over the past 3 decades, supercritical CO₂ has been used for the production, extraction, and isolation of valuable natural or synthetic biomaterial compounds. Supercritical CO₂ was selective in the separation, extraction, and preparation of desired biomedical materials for different applications. In this section some of the applications of supercritical technology in biomedical materials are illustrated briefly.

Processing of biological molecules

Recently, different studies mentioned the application of supercritical CO₂ technology in processing biomolecules such as proteins, phospholipids, and DNA. Some examples of these procedures are mentioned here. Webb et al.^[8] reported new methodologies for dissolving ionic and polar species in supercritical CO₂, which, as a result, led to a broadening of the range of applications for SCF technology. Webb et al. used fluorinated surfactants in order to prepare water in carbon dioxide microemulsions to permit solubilization of ionic and biological species such as phospholipids. They also reported the preparation of supercritical CO₂ soluble metal precursors such as metal complexes that can be impregnated efficiently into polymeric substrates. The clean synthesis of such

composites can be applied to the development of improved medical implants.^[8]

The major components of cellular membranes are phospholipids, biological surfactants that are able to form unilamellar and multilamellar vesicles. Liposomes, which have applications in controlled drug delivery, have been formed using purified phospholipids. Badens et al.^[9] applied the RESS method in order to produce microparticles of lecithin. In their experiments ethanol was used as a cosolvent to dissolved lecithins in pressurized CO₂ in the RESS process. Spherical, fairly agglomerated particles in the 1- to 50-nm range were obtained from this procedure under proper conditions. Whereas increasing the liquid flow rate decreased the particle size and tightened the particle size division, there was no effect of pressure on particle size. This was recognized to better atomization and, therefore, enhanced mass transfer and supersaturation at higher liquid flow rates. Ethanol was not detected in the processed lecithins when analyzed for remaining solvent content via infrared spectroscopy.^[9]

Plasmid DNA, in the form of vaccines for mucosal administration and intranasal inhalation, has been considered a promising therapy for treatment of diseases such as diabetes and Parkinson's. There is increased interest in producing plasmid DNA powders. As an example of these recent works, Tsevistas et al.^[10] studied the effect of SCF CO₂ on extraction of DNA. They concluded that exposure to SCF CO₂ in the absence of free water did not degrade plasmid DNA. Isopropanol-modified SCF CO₂ was used to produce plasmid DNA micropowders. Tsevistas et al.^[10] precipitated pSV β , a 6.9-kb plasmid, by supercritical CO₂ technology. They reported that SCF technology is a promising approach in the processing of biomaterials such as DNA.^[10]

Proteins are currently one of the most important families of biomolecules in the pharmaceutical industry. Many studies demonstrated the ability of supercritical technologies to prepare and precipitate different types of proteins that can be used in several therapeutic methods. Sievers et al.^[11] demonstrated that supercritical CO₂ technology can be applied as an aerosolization agent in order to prepare protein microparticles such as bovin serum albumin. They used a bubble drying technique in their experiments, which has the advantage of using pure aqueous systems without the addition of organic solvents.^[11]

SCF technology has the ability to swell and plasticize a broad range of amorphous polymers. A plasticized polymer has a reduced viscosity and allows for incorporation of soluble material into the swollen polymer. Subsequent depressurization entraps the insoluble material and creates a porous structure. This process has been exploited to incorporate a variety of inorganics, pharmaceuticals, and biomolecules in

biodegradable polymeric matrices that can be used in both tissue engineering and drug delivery. Howdle et al.^[12] used this technique in encapsulation of ribonuclease A, catalase, and β -D-galactosidase in poly(lactic acid) (PLA). Release studies of preserved protein demonstrated an initial burst of less than 10% in the first 2 days, followed by zero order of release for 70 days.^[12]

Bacterial inactivation (sterilization) and enzyme denaturation

The effect of SCF on enzyme activity and enzymatic reaction for food and pharmaceutical processing is typically explored. In presence of a CO₂ supercritical situation, 10 MPa pressure is required to denature monomeric proteins; on the other hand, hydrostatic pressures of more than 600 MPa are required to have the same effect on the proteins. Enzymatic denaturation by SCF-CO₂ in the food industry caused increased digestibility, foaming, and emulsifying in the treated food.

Current methods of sterilization, which include ethylene oxide, γ -radiation, electron-beam, steam, and hydrogen peroxide plasma, have limitations regarding their biomedical applications. These methods altered the structure and characteristics of materials, especially when applied to thermally and hydrolytically sensitive polymers. Pressures slightly above 10⁻² MPa inhibited a range of fungi, bacteria, and yeasts. Supercritical CO₂ treatment provides an effective method of sterilization for thermally and hydrolytically labile biomedical materials by inactivation of a variety of bacterial microbes. Carbon dioxide has unique properties that make it an appealing medium for sterilization. At relatively low pressures and temperatures carbon dioxide transitions to a supercritical state, often referred to as a dense-phase gas. The use of supercritical CO₂ as a sterilizing medium is all the more appealing because it is non-toxic and easily removed by simple depressurization and outgassing. It has the ability to eliminate pathologic and spoilage microorganisms from food products, polymers, and other biomaterials.^[12]

Extracting of biologically active substances

Extraction is one of the oldest commercial applications involving supercritical carbon dioxide.^[13] The increasing demand for natural food antioxidants encouraged worldwide researchers to extract biologically active substances from a variety of raw plant materials. Supercritical CO₂ allows the extraction of active ingredients from herbs and plants with a better imitation of flavor or fragrance in comparison with conventional

operations. As a result of the operation at decreased temperature, thermal degradation and decomposition of labile compounds are avoided, while the lack of oxidative factors such as light and oxygen prevents oxidation reactions. This latter point is of particular interest for the extraction of antioxidants, warranting the preservation of their biological properties.^[13]

Pharmaceutical Formulations

Conventional methods for particle formation, micronization, coprecipitation, impregnation, and encapsulation of pharmaceuticals can be challenging because the heat and mechanical stresses involved can cause thermal and chemical degradation of the drug and change the properties of the drug. In addition, application of large amounts of solvent and surfactants-emulsifiers can lead to undesirable levels of residual impurities, which require further purification procedures. With regard to its advantages, such as being environmentally benign and having adjustable solvent properties and produced particles that are free from solvent residues, supercritical CO₂ technology appeared to be a promising substitute for the solvents that are currently used in the pharmaceutical industry. SCF technology provides a clean way to produce active particles for controlled delivery in the body process of thermally labile biological materials. Freeze drying, spray drying, precipitation, and other conventional methods for powder production have significant disadvantages for this application. SCF technology facilitates controlled particle size formation and solvent removal from the products.^[14]

Drug delivery and controlled release of drugs

In advanced drug delivery technologies, new approaches for the delivery of pharmaceuticals will be required. For example, beside traditional methods, inhalation techniques will be applied more for drug delivery into the bloodstream. This new methodology requires injection devices in order to shoot the small particles of the drug through the skin of the patient instead of being applied with a needle. Particles that can be used in this method should have a specific size range and well-defined particle size distribution. Different processes use supercritical CO₂ and have been used for particle formation, such as RESS and GAS recrystallization.^[14]

The application of SCF such as supercritical CO₂ provided an effective and clean solvent substitute for drug and polymer processing. Gong et al.^[15] impregnated indomethacin (IM, a non-steroidal antiinflammatory drug) into chitosan thermosets for the preparation of controlled release formulations. The

results indicated that the SCF impregnation process resulted in IM being amorphyously dispersed within the chitosan matrix. Fourier transform infrared spectroscopy data implied that the aliphatic carbonyl group of IM interacted with the amine group of the chitosan backbone. In vitro dissolution studies (via UV-vis spectroscopy) demonstrated that the dissolution rate of IM significantly increased after processing in supercritical CO₂.^[15] In another study by Gong et al.^[16] supercritical carbon dioxide (SC-CO₂) was used to prepare coprecipitates of IM and poly(vinylpyrrolidone) (PVP) in order to enhance the dissolution rate of the drug. The coprecipitates of IM and PVP at various proportions were prepared using a stirred batch reactor containing SC-CO₂ as a gas-saturated solution. The results indicated that as the PVP weight fraction decreased, IM displayed an increasing amount of crystalline material. The dissolution rate of IM was increased with the incorporation of PVP. IM and PVP at various weight fractions exhibited comparatively higher dissolution rates than crystalline IM alone. The supercritical CO₂ technology produced a solvent-free, completely amorphous porous IM solid dispersion with a fast dissolution pace.^[15,16]

Particle size reduction of drugs; Preparation of drug nanoparticles

Particle size is an important factor that influences the rate of dissolution of a drug in biological fluids and therefore has a major effect on bioactivity and bioavailability. Numerous researchers utilized SCF technologies as substitute approaches to reduce the particle size of drugs. Compared with conventional methods, the SCF process is advantageous because it has better control over particle size distribution and morphology. In different reviews, the temperature and pressure mainly determine the particle morphology. Pressure and temperature have a large effect on SCF CO₂ density; these parameters are expected to have a strong effect on the resulting properties of the particles. The temperature is also expected to affect the solubility of the protein, the antisolvent in the solvent, and the solubility of the solvent in the antisolvent, thus modifying the precipitation kinetics and crystal growth.

One of the main applications of SCF CO₂ processing is the production of nanosize particles. SCF has been used in order to synthesize nanoparticles such as titanium dioxide. This technology is especially interesting for the development of dry powder inhalation systems and dry powder injection systems in the pharmaceutical field, as well as in useful biomaterials such as nanohydroxyapatite. Thakur and Gupta^[17] reported a new process for phenytoin drug synthesis using menthol solid cosolvent in a modified RESS

method. The results of this study demonstrated that this technique has the capability of nanosize particle formation. The average particle size of synthesized phenytoin in the conventional RESS process was 200 nm, whereas upon the addition of cosolvent in the RESS modified process, the average particle size decreased to 120 nm at 96 bar and 45°C. By increasing the pressure and temperature, the particle size decreased.^[17]

Recently, supercritical CO₂ was used to prepare silver nanoparticles within commercially available silicone and polyurethane polymers in a controlled and reproducible manner. Silver is a non-toxic, naturally occurring metal that can kill many harmful microorganisms and is a very good antibacterial agent. This alternative approach seems to be a promising way to manufacture materials with controlled loadings and a high surface area of silver. Moreover, it provides an alternative to the use of conventional antibiotics and the inherent problems of antibiotic resistance as a route to prevent infection in medical devices.^[14]

Stabilization of proteins in dry powder formulations

Therapeutic proteins have become vital in the treatment of many diseases. Their formulation in dry form is often required to improve their stability. Traditional procedures for protein drying, such as freeze-drying or spray-drying processes, are often harmful to labile proteins. Alternative drying techniques have been explored for drying pharmaceutical protein formulations, each of which has its specific advantages and disadvantages. Among the available drying methods, the SCF drying procedure is popular because it is cost effective, has mild conditions for the procedure, and has the capability of producing microparticulate protein powders. Supercritical CO₂ mixing technology has also been applied in the production of release systems for controlled delivery of proteins or other bioactive components.^[15]

Tissue Engineering, Medical Implants, and Bone Grafts

Tissue engineering

Certain tissues and organs cannot heal effectively by themselves after injuries and, in order to restore their function, they require special treatments. A standard therapy in this case is transplantation but because of a shortage of donor tissue, not enough patients can receive the organ or tissue transplantations they need to survive. Therefore, researchers recently concentrated on developing methods for tissue regeneration. Three-dimensional polymer scaffolds with porous structures

can be used as cell supports to offer mechanical stability and structural guidance and to permit seeding of a high density of cells before or after transplantation into the body. Drugs, nutrients, hormones, and growth factors can be inserted into these porous appliances to promote cell differentiation and then vascularization of the growing tissue.

Conventional techniques for preparation of polymer scaffolds or particles typically use organic solvents and high temperatures that may be harmful to adherent cells, nearby tissues, or biologically active factors. Therefore, the use of SCF, in particular supercritical CO₂ technology, is an appealing alternative approach for a wide range of biomedical applications in tissue engineering.^[18]

The advantages of supercritical technology in comparison with other conventional methods in preparation of scaffolds in tissue engineering include the absence of organic solvents, the ability to incorporate delicate biological matter without loss of activity, and control over the morphology of an internal porous architecture. Tissue engineering scaffolds must possess an interconnected pore structure and high degree of porosity, which ensures adequate cell attachment throughout the 3-D matrix. Porosity is created during the SCF process by the thermodynamic instability that results following depressurization around the liquidated polymer. The porosity and pore structure of the resultant scaffold can be controlled by altering the amount of gas added and the release rate from the polymer.^[18]

Synthesis of biopolymers. A number of manufacturing techniques have been applied to produce scaffolds from synthetic and naturally derived biodegradable polymers such as PLA, poly(glycolic acid), collagen, and chitosan, in addition to inorganic materials such as hydroxyapatite.

Gas foaming using supercritical technology is a substitute method applied to create porous structures for scaffolds in tissue engineering. Pores are created from a gas-polymer solution by sudden depressurization, which induces bubble nucleation. Foaming with CO₂ has the ability to eliminate the need for organic solvents and high temperatures for plasticization, and is a process that is best suited for biomolecules.

Application of this technique can result in a wide range of morphologies and porosities. Bioactive elements such as proteins or growth factors can also be added in this method. Gas foaming is a solvent-free technique used to produce porous structures with a wide range of characteristics. These characteristics seem adequate for bone tissue engineering applications. Polyurethanes have been successfully used as a biomaterial with uses ranging from heart catheters to bone scaffolds. Porous polyurethane has been synthesised at 40°C using gas foaming method, which

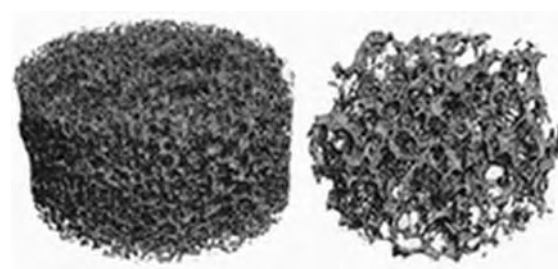


Fig. 2 Microcomputerized tomography image of a PEMA/THFMA scaffold processed in SC-CO₂.^[18]

compared relatively well to conventional solvent polymerization at 130°C. Barry et al. utilized supercritical CO₂ to fabricate porous poly(ethylmethacrylate/tetrahydrofurfuryl methacrylate) (PEMA/THFMA, a non-degradable polymer proposed as a support material for cartilage repair) foams (Fig. 2) with controlled porosity, pore geometry, and interconnectivity.^[21] The cross-linked polymer system was plasticized with supercritical CO₂ at 100 bar for a period of 8 h, followed by controlled depressurization. Total porosity was greater than 80%, with at least 60% being open porosity. The results were significantly greater than the 10%–30% porosity obtained in other conventional polymer processing systems. Pore sizes ranged between 35 and 400 mm depending on the pressure, exposure time, and depressurization time.^[18]

Different techniques have been used to produce porous polymers and hydrogels such as freeze drying, particulate leaching, gas-foaming methods, and the use of gas-forming agents. Hydrogels with broad macropore size distributions and a mixture of open and closed pores are often produced as a result of these techniques, whereas emulsion templating techniques can sometimes produce polymers with a more defined and controlled pore structure. Butler et al. used carbon dioxide in water emulsion templating to successfully produce porous polyacrylamide polymers. However, because of the slow polymerization rate of hydroxyethyl acrylate (HEA) and the destabilization of the CO₂-in-water emulsion, they were not successful in synthesising hydrogels based on HEA and hydroxyethyl methacrylate using the same method.^[22]

Partap et al.^[19] reported an emulsion templating “oil-in-water” technique to prepare highly porous alginate hydrogels (Fig. 3) using application of supercritical water technology by coupling a templating oil-in-water emulsion reaction with an internal gelation reaction to lock in the structure from the internal oil. They mentioned that both an organic solvent (isooctane) and a dense-phase CO₂ (above the critical point of CO₂) have been used as the internal oil phase for the templating step. The results demonstrated that they developed a new generic method of producing

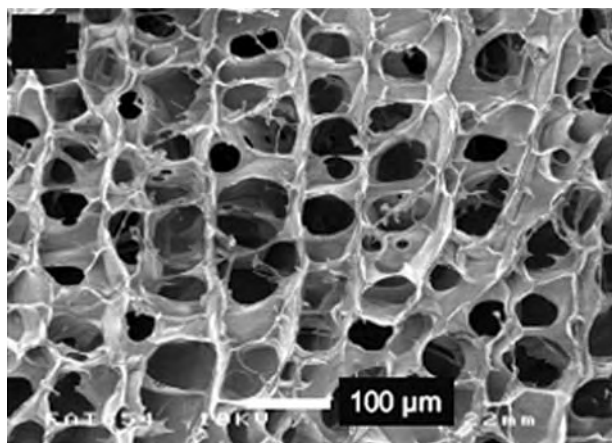


Fig. 3 Scanning electron microscopy image of a sample of synthesized porous calcium alginate hydrogels.^[19]

porous alginate hydrogels by combining emulsion templating with an internal gelation reaction. Highly porous alginate hydrogels with a narrow range of macropore sizes (without the use of an organic solvent) were produced in their experiments, as presented in Fig. 3.^[19] A narrow range of macropore sizes (without the use of an organic solvent) was produced using the new RET (reactive emulsion templating) reaction, which utilizes carbon dioxide in water emulsion to template the pores. The RET hydrogels exhibit an open, well-interconnected pore network with a narrow pore-size distribution that is potentially suitable for tissue-engineering applications [85%–90% total porosity and less than 250 μm (23–43 μm) mean pore diameter].

Medical implants

Different researchers demonstrated that supercritical technology can be used for surface modification of biomedical devices. One recent work involved the application of SCF for surface modification of coronary stents with a drug-eluting polymer matrix. Drug-eluting stents reduced early-state restenosis in clinical trials. Intrinsic process control limitations of current solvent-based methods, as well as challenges associated with maintaining structural and morphological integrity of the therapeutic agents deposited during the coating process, have been significant challenges to date. Improvements in surface modification technologies such as uniformity, conformity, and control of drug morphology, along with the use of other potential classes of active therapeutic agents, may help to address these issues. SCF-based technology has the ability to deliver the next generation of drug-eluting stents that can afford control over drug morphology and may allow manufacturing processes to apply

separate and independent therapies within a single, multitherapy coating. The morphology of the drug is maintained by avoiding the need to dissolve the drug in a liquid solvent, which is compulsory in traditional processes. One advantage of the application of this method is that it provides the possibility of increasing the constancy of the material.

The other application of SCF technology is the coating of orthopedic implant replacements such as hip implants. Failure of these implants is common and is caused by gradual wearing of the polymer surface. As the ball segment of the implant moves around in the socket of the implant, the polymer degrades at the load-bearing surface through the process of adhesive wear. This leads to production of submicrometer-size particles from the coated layer of the implant, which the body simply cannot remove. Macrophage cells in the body unsuccessfully try to digest and remove the particles, which leads to a severe toxic response and finally to rejection of the implant. Webb et al.^[8] attempted to develop SCF methods to improve the adhesive wear properties of the implant surface by application of certain metals, metal oxides, and metal sulfides as fillers in polymeric substrates that can considerably enhance the wear lifetime. They applied silver-based precursors to the preparation of silver composites by impregnation from supercritical CO_2 solution. Results of preliminary mechanical testing on the composites indicated that there was an improvement in adhesive wear properties. In addition, the presence of silver particles led to visibility of the desired part of the implant in X-ray imaging.^[8]

Bone graft materials

Hydroxyapatite (HAp) is a promising material used in orthopaedics as a bone graft material. There are many

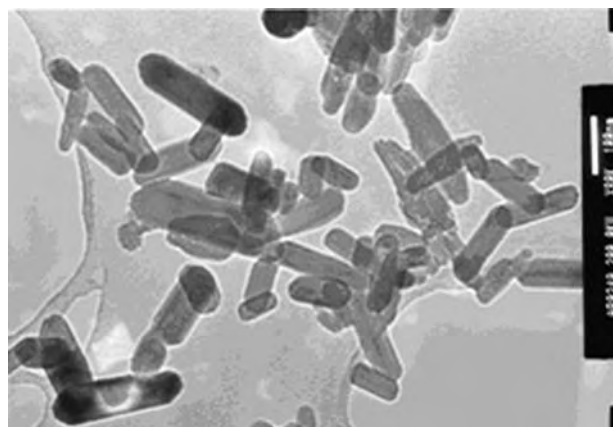


Fig. 4 Transmission electron micrograph of hydroxyapatite nanorod made using supercritical water at 400°C and 24 MPa in the hydrothermal flow system.^[20]

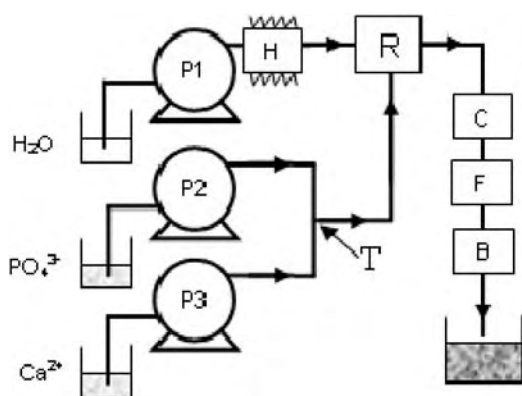


Fig. 5 Scheme of the three-pump (P1, P2, and P3) hydrothermal flow system used for the synthesis of nanohydroxyapatite. P, pump; C, cooling; F, filter; B, backpressure regulator; R, reactor; T, Tee piece mixer.^[20]

techniques for synthesis of this bioceramic, such as wet chemical synthesis and the sol-gel technique. Chaudhry et al.^[20] synthesised nanoparticles of HAp, which can be used as a synthetic bone replacement material or for coating of implants, using a three-pump hydrothermal flow process (up to 400°C and at 24 MPa) with the application of supercritical water technology. The process used by Chaudhry et al. decreased the time required for maturation of the reagents from over 18 h to less than a few seconds and avoided the need for further crystallization heat treatment steps such as sintering. The results demonstrated that the product was obtained as a highly crystalline and phase pure material, without the need for an aging step or subsequent heat treatment.

Fig. 4 illustrates the transmission electron microscopy micrograph of the nanohydroxyapatite synthesized using this method, whereas Fig. 5 presents a schematic figure of the hydrothermal flow system used for synthesis of hydroxyapatite nanoparticles by application of the supercritical water technology.^[20]

Haque et al. synthesized high-pressure carbonated apatite in supercritical carbon dioxide with high carbonate content and good mechanical and biological properties. They also used reduced calcination temperature and tried to change the reagent ratios to ensure phase purity. In another study, Haque et al. synthesized calcium phosphates in SC-CO₂ at low pH without the use of ammonia solution to allow the pH to drop below 9.

CONCLUSIONS

SCFs (particularly CO₂ and water) have covered a wide range of chemical reactions and processing in the pursuit of advanced biomaterials and natural product

extraction. In particular, supercritical carbon dioxide is a good non-toxic alternative to organic solvents in the manufacture of biomaterials with well-defined properties. By controlling the characteristics of biomaterials (size, size distribution, porosity, surface area), SCF processes can allow a better match to clinical requirements.

The next decade may see an expansion in the number of applications of SCFs in biomaterials, and much promise has already been evidenced. For example, many opportunities exist for the manufacture of more robust and well-defined tissue engineering scaffolds in which degradation and other properties must be acutely controlled. The authors of this review do not doubt that the application of suitable SCF technologies could hold the answer to this and other challenges in the biomaterials field.

S

ARTICLES OF FURTHER INTEREST

Microporous Materials; Polymer Foams; Polymers; Tissue Engineering Scaffolds

REFERENCES

1. Park, J.B. *Biomaterials: An Introduction*, 2nd Ed.; Plenum Press: New York, 1992.
2. Mohamed, R.S.; Mansoori, G.A. *The Use of Supercritical Fluid Extraction Technology in Food Processing, Featured Article—Food Technology Magazine*; The World Markets Research Centre: London, UK, June 2002.
3. Sihvonen, M.; Javenpaa, E.; Hietaniemi, V.; Huopalahti, R. Advances in supercritical carbon dioxide technologies. *Trends Food Sci. Technol.* **1999**, *10*, 217–222.
4. McHugh, M.A. *Supercritical Fluid Extraction: Principles and Practice*, 2nd Ed.; Butterworths Publishers: Boston, MA, USA, 1986.
5. Woods, H.M.; Silva, M.M.; Nouvel, C.; Shakesheff, K.M.; Howdle, S.M. Materials processing in supercritical carbon dioxide: surfactants, polymers and biomaterials. *J. Mater Chem.* **2004**, *14*, 1663–1678.
6. Phelps, C.L.; Smart, N.G.; Wai, C.M. Past, present, and possible future application of supercritical fluid extraction technology. *J. Chem. Educ.* **1996**, *73* (12), 1163–1168.
7. Darr, J.A.; Poliakoff, M. New directions in inorganic and metal–organic coordination chemistry in supercritical fluids. *Chem. Rev.* **1999**, *99*, 495–541.
8. Webb, B.B.; Marr, P.C.; Parsons, A.J.; Gidda, H.S.; Howdle, S.M. Dissolving biomolecules and modifying biomedical implants with supercritical carbon dioxide. *Pure Appl. Chem.* **2000**, *72* (7), 1347–1355.
9. Badens, E.; Magnan, C.; Charbit, G. Microparticles of soy lecithin formed by supercritical processes. *Biotechnol. Bioeng.* **2001**, *72*, 194–204.

10. Tsevistas, M.; Levy, M.S.; Lo-Yim, M.Y.A.; O'Kennedy, P.; York, P.; Humphery, G.O.; Hoare, M. The formation of plasmid DNA loaded pharmaceutical powders using supercritical fluid technology. *Bio-technol. Bioeng.* **2001**, *72* (1), 12–18.
11. Sievers, R.E.; Karst, U.; Milewski, P.D.; Sellers, S.P.; Miles, B.A.; Schaefer, J.D.; Stoldt, C.R.; Xu, C.Y. Formation of small droplet aerosols assisted by supercritical carbon dioxide. *Aerosol. Sci. Technol.* **1999**, *30*, 3–15.
12. Dixon, N.M.; Kell, D.B. The inhibition by CO₂ of the growth and metabolism of microorganisms. *J. Appl. Bacteriol.* **1989**, *67*, 109–136.
13. Díaz-Reinoso, B.; Moure, A.; Domínguez, H.; Parajó, J.C. Supercritical CO₂ extraction and purification of compounds with antioxidant activity. *J. Agric. Food Chem.* **2006**, *54* (7), 2441–2469.
14. Howdle, S.M.; Watson, M.S.; Whitaker, M.J.; Popov, W.K.; Davies, M.C.; Mandel, F.S.; Shakesheff, K.M. Supercritical fluid mixing: Preparation of thermally sensitive polymercomposites containing bioactive materials. *Chem. Commun.* **2001**, 109–110.
15. Gong, K.; Darr, J.A.; Rehman, I.U. Supercritical fluid assisted impregnation of indomethacin into chitosan thermosets for controlled release applications. *Int. J. Pharmaceut.* **2005**, *315*, 93–98.
16. Gong, K.; Viboonkiat, R.; Rehman, I.U.; Buckton, G.; Darr, J.A. Formation and characterization of porous indomethacin–PVP coprecipitates prepared using solvent-free supercritical fluid processing. *J. Pharm. Sci.* **2005**, *94* (12), 2583–2590.
17. Thakur, R.; Gupta, R.B. Formation of phenytoin nanoparticles using rapid expansion of supercritical solution with solid cosolvent (RESS–SC) process. *Pharmaceut. Nanotechnol.* **2006**, *38* (2), 190–199.
18. Quirk, R.A.; France, R.M.; Shakeshe, K.M.; Howdle, S.M. Supercritical fluid technologies and tissue engineering scaffolds. *Curr. Opin. Solid State Mater. Sci.* **2004**, *8*, 313–321.
19. Partap, S.; Rehman, I.; Jones, J.R.; Darr, J.A. Supercritical carbon dioxide in water emulsion-templated synthesis of porous calcium alginate hydrogels. *Adv. Mater.* **2006**, *18*, 501–504.
20. Chaudhry, A.A.; Haque, S.; Kellici, S.; Boldrin, P.; Rehman, I.; Khalid, F.A.; Darr, J.A. Instant nanohydroxyapatite: a continuous and rapid hydrothermal synthesis. *Chem. Commun.* **2006**, *4* (21), 2286–2288.
21. Barry, J.J.A.; Gidda, H.S.; Scotchford, C.A.; Howdle, S.M. Porous methacrylate scaffolds: supercritical fluid fabrication and in vitro chondrocyte responses. *Biomater.* **2004**, *25* (17), 3559–3568.
22. Butler, R.; Davies, C.M.; Cooper, A.I. Emulsion Templating Using High Internal Phase Supercritical Fluid Emulsions. *Adv. Mater.* **2001**, *13* (19), 1459–1463.

Surface Coatings

Shih-Horng Su
Suzanne Conroy
Tung-Liang Lin
Min-Shyan Sheu
Ih-Houng Loh

AST Products, Inc., Billerica, Massachusetts, U.S.A.

S

INTRODUCTION

Surface coatings have a long history in medical products. Earlier, colorful coatings served chiefly in decorative or identification roles on pills. Later, surface coatings became a way to modify the properties of medical devices and implants. By using different bioactive components, coatings can enhance the surface wettability, electrical conductivity, X-ray radiopacity, and echogram visibility. Coatings can impact physical properties such as coefficient of friction and thus enhance maneuverability without loss in device functionality. Anticrustation coatings reduce the surface mineralization on ureteral stents to keep the lumens open for urine flow. Thromboresistant coatings prevent blood clotting on the device surfaces to maintain oxygen-exchange efficiency. Antimicrobial coatings reduce infectious complications to prolong the useful life of devices and implants. Finally, coatings that promoted cellular adhesion enable bone implants to have stronger fixation. These applications and others illustrated that surface coatings were essential to the continual improvement of medical products.

A recent development in coating technology is the use of drug delivery mechanisms to promote selected biological functions. Drug functionality elevates the importance of surface coatings to the point that the coating is an equal contributor to the therapeutic goal, as is the naked device. This statement is dramatically demonstrated by the success of drug-eluting stents. Sirolimus-eluting coatings boost the clinical outcomes of stenting to levels that never have been achieved with bare stents. In the near future, surface coatings will be the essential players in fulfillment of physicians' therapeutic goals.

SURFACE COATINGS AND MEDICAL DEVICES

Medical device and implant manufacturing frequently includes surface coatings steps. Surface coatings provide add-on benefits to enhance the surface properties without significantly altering the bulk properties of

medical products. From a methodology viewpoint, a coating passively or actively modifies the surface properties of medical devices and implants. The so-called passive method alters selected properties of the substrate surfaces without delivery of drugs or other agents. In contrast, the active approach delivers bioactive compounds that influence biological responses in order to achieve desirable outcomes. From a strategic aspect, surface coatings generally offer benefits such as improved maneuverability, functionality, and additional therapeutic effects for medical devices. Regarding maneuverability, tailor-made surface coatings alter surface wettability, electric insulation, and coefficient of friction of medical devices. Moreover, certain surface coatings can improve the image quality of medical devices under X-ray fluoroscope and ultrasound set-ups and thus greatly further the surgical outcomes. Regarding functionality, surface coatings can prevent mineralization, thrombosis, and microbial colonization on devices. Coatings can promote cell adhesion to ensure that the devices' performance is not compromised due to foreign body reactions.

Drug-eluting coatings are the latest development in the medical device coating industry. Surface coatings act as reservoirs for local drug release. Properly formulated surface coatings enable delivery of medications to site-specific lesions and thus achieve desirable therapeutic long-term results. In this section, applications of surface coatings and methods of making such coatings are broadly reviewed.

Surface Coating Improves the Maneuverability of Medical Devices

For smooth procedures, unique surface properties including wettability, electric isolation, radiopacity, echo visibility, and lubricity are desirable for many medical devices.

Wettability is an important physical factor for designing membrane-type biosensors or filtration devices. Enhanced surface wettability shortens the analysis time,

improves biosensor accuracy, and promotes cell viability for filtration devices.^[1] Several methods including corona discharge, gas plasma exposure, flame treatment, and acid etching are available to improve the wettability of polymeric porous materials. However, the wettability introduced by these processes dissipates after treatment, particularly when exposed to aqueous washing, elevated temperatures, or even simple storage in air. A more stable surface coating-based polyethylenimine (PEI) grafting has been developed that creates permanent wettability. The process of PEI coating includes plasma pretreatment and wet solution immersion. Substrates are first primed with O₂/CO₂ plasma under mild conditions. PEI is then grafted to the activated surface via ionic binding. The wettability of membranes is maintained for more than a year.

During conventional manufacturing of cardiac rhythm managing devices (implanted pulse generators (IPGs), for example), superior electric insulation coatings are needed to prevent local muscle stimulation and myopotential sensing in unipolar pacing. Parylene (polyparaxylylene) coating has been selected to meet the performance requirements for decades due to its inert/biocompatible properties (noncytotoxic, non-systemic toxic, and hemocompatible). However, conventional parylene coating processed via vapor deposition polymerization (VDP) has always been plagued by poor substrate adhesion. A modified coating process has been developed to overcome adhesion limitations.^[2,3] The new process integrates traditional parylene VDP process with plasma molecular activation. Substrate surfaces are first primed with hydrogen plasma. Then, without breaking containment, parylene is deposited on the primed surfaces. Consequently, molecular-level binding is created between the substrate and parylene. Coatings generated via the process uniformly and completely cover substrates and can range from 2 to 25 μm in thickness. The shelf life is more than two years.

Radiopacity is another important physical characteristic for devices including guidewires, catheters, and metal stents used in interventional cardiology. Radiopacity improves devices' image quality and allows physicians to have a reliable method to trace the exact location of a guidewire or catheter through vasculature. This assists in the successful evaluation of lesions as well as the accurate deployment of stents. Radiopacity was conventionally achieved by applying metal bands or implanting noble metals such as gold, tantalum, iridium-platinum and platinum, or heavy metal salts such as barium sulfate, onto guidewires and into catheters. Ion beam-assisted deposition (IBAD) is a traditional way to apply those metals onto devices. However, clinical practitioners were concerned about cellular debris entrapment with the metal band and had difficulties with the transparent unfilled portions of

catheters. Coatings produced by IBAD are also plagued by unreliable adhesion. A new alternative to IBAD process, microfusion, was developed to answer these questions.^[4] Briefly, in microfusion, devices are first placed in the unbalanced magnetron filed and set up as the negative bias. Ions are injected through unbalanced magnetron sputtering and bombard the device surfaces. The negative-bias voltage on the substrate controls the traveling speed of ions. The processing time determines the coating thickness. Consequently, coating adhesion is stronger and uniformity is better. Microfusion has the advantage of providing multiple ions deposition from multiple orientations simultaneously.

Another image-enhancing surface coating is known as echo coating. Echo coatings were developed to enhance the ultrasonic image quality of needles and interventional devices. The echogenicity is a result of microbubbles generated by the coating matrix. Through a dip-coating and curing processes, a polyurethane-based microporous coating is created. Microcavities in the coating are on the scale of a few microns and have minimal impact on coating lubricity. Under agitation, microbubbles exit the polymer matrix, become excellent ultrasound reflectors, and help to dramatically enhance the image in Doppler sonogram. The improved ultrasonic images increase the technical success rate of the examination by reducing examination time and patient discomfort.

Less friction and more lubricity are common characteristics required for guidewires, catheters, probes, needles, and implantable ports. Two types of lubricious coatings are common. The first type is dry-film lubricious coating. Parylene and expanded polytetrafluoroethylene (ePTFE, Teflon[®]) are frequently used for dry-film lubricious coatings. Both coatings are about equivalent in the measurement of static and dynamic coefficients. Dry-film lubricious coatings are ideal for tiny lead wires requiring electric insulation, chemical resistance, and abrasion resistance.

When in contact with sensitive, fluid-rich tissue/organ systems, surfaces with better lubricity are needed. Slippery-when-wet coatings were specifically developed to meet the challenges. The principle of slippery-when-wet coatings is to employ water molecules as lubricants. Typically, lubricious coatings are polyurethane or poly(ethylene-co-acrylic acid) based. During the preparation of coating, the resins are cross-linked either through chemical or photochemical reaction, and polyvinylpyrrolidone (PVP) is incorporated via a network interpenetration mechanism (Fig. 1). The polyurethane or poly(ethylene-co-acrylic acid) network provides good adhesion and allows the PVP to protrude in order to adsorb water molecules. Since PVP is highly hydrophilic, it produces a microlayer of water that reduces the wet coefficient of friction to nearly zero. This minimizes the damage/invasion that

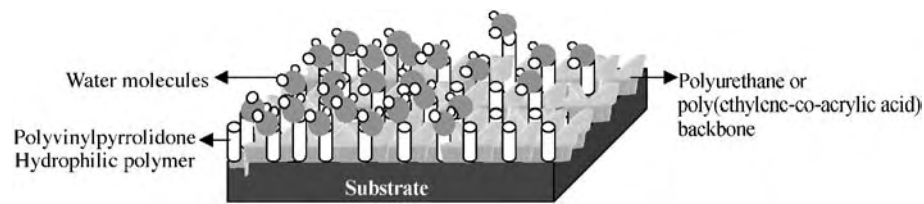


Fig. 1 Slippery-when-wet coating. Polyurethane or poly(ethylene-co-acrylic acid) matrix provides strong adhesion for coating and substrate. Embedded hydrophilic polyvinylpyrrolidone protrude outward to grab water molecules at contact with body fluid. This watery microlayer increases the surface lubricity of the substrate.

imposed onto the epithelium and mucus layers during the device positioning and greatly assists physicians in the navigation of devices through tortuous vasculature.

Surface Coatings Ensure the Functionality of Medical Devices

Medical devices and implants are developed to improve life quality in many ways. Good maneuverability enables physicians to have smooth operations in positioning medical devices or implants and thus minimize patient discomfort or injury. Equally important is to ensure that devices and implants function as designed after procedures. Devices and implants inevitably have contact with body fluids and tissues. In some cases, devices fail due to natural biochemical reactions. Encrustation on urinary stents and catheters is an example. Encrustation is the mineralization resulted from pH change in urine. Moderate to severe encrustation blocks the urine flow and predisposes stents and catheters to replacement. In other situations, cell or bacterial adhesion endangers the functionality of implants. For example, platelet adhesion on cardiopulmonary bypass (CPB) leads to thrombosis and further blocks the blood flow, which jeopardizes the successful surgery outcomes. Postsurgical lens epithelial cell deposition on intraocular lens (IOL) results in failed implantation. Bacterial adhesion on catheters leads to nosocomial catheter-associated infections, which compromise the functionality of catheters. In these cases, minimal bacterial/cell adhesion on devices and implants are most desired. In other cases, implants are expected to coordinate with cell/tissue responses to achieve better performance. For instance, orthopedic metal implants need better bone ingrowth (osteogenicity) for stronger fixation.

Biomaterials that devices and implants are made from frequently do not have the surface properties to respond properly to physiological environments. For example, polyurethane has a higher tendency to cause encrustation than does silicon, and polymethylmethacrylate (PMMA) IOLs induce more epithelialization than do

polyacrylate ones. Surface coatings bridge the device functionality and cell/tissue responses and allow for a greater range of choices in base material selection. Both active and passive surface coatings are applied to prevent compromised functionality.

Antienrustation Coating

Mineral encrustation is a common complication associated with the use of urethral stents and catheters. In encrustation, components in the urine deposit, grow into crystals, and eventually encrust the surface of the device. The mechanism has many variables. It is understood, however, that in many *in vivo* systems, bacterial colonization generates biofilms that become mineralized. In sterile urine, the development of encrustation has been shown to be dependent on urinary properties such as pH and ionic strength as well as the biomaterials properties. Urine is generally free from bacteria in ureteral stents and thus in the stent environment, it is the chemistry of the urine whose variables dominate. In infected urine, urease produced by the adhered bacteria hydrolyzes the urea to produce ammonia. This elevates urine pH, a condition that thermodynamically favors the precipitation of magnesium and calcium in the form of struvite ($\text{NH}_4\text{MgPO}_4 \cdot 6\text{H}_2\text{O}$) and hydroxyapatite ($\text{Ca}_5(\text{PO}_4)_3\text{OH}$). These minerals are two of the major components of encrustation found in urinary prosthetics. Encrustation frequently leads to the blockage and fracture of the devices. Conventional strategies have focused on minimizing the effect of one variable without addressing the others. For example, use of different biomaterials had limited effect on minimizing encrustation. Hydrophilic surfaces (poly(ethylene oxide)/polyurethane composite hydrogel) discourage bacterial adherence and hence the formation of biofilm. However, they have shown no significant effect on encrustation reduction. Other approaches currently being evaluated include coating the surface with glycosaminoglycan heparin and with urease inhibitors such as acetohydroxamic acid and fluorofamide. While promising *in vitro* data has been published, *in vivo* results are pending.^[5,6]

Thromboresistant Coatings

Medical devices involving blood contact often fail due to the formation of blood clots on the surface. For example, cardiopulmonary bypass, an extracorporeal blood circulating system, maintains the circulation throughout the body and provides nutrients and oxygen necessary for life, while the cardiothoracic surgeons are fixing cardiac defects. The efficiency of CPB can be impaired by the adhesion of blood cells on biomaterial surfaces. Catastrophic surgery outcomes may result from consequential severe thrombosis and inflammation. During surgery, heparin is routinely administered to prevent blood clotting. Heparin binds to antithrombin III (ATIII) to become an acting inhibitor to thrombin, factor X, and other enzymes (IX, XI, & XII). Even so, clotting on the surface of devices still occurs while the heparin activity decayed. To better prevent the incidence, permanent and drug-release types of antithrombotic coatings were developed to render the device surfaces thromboresistant. Heparin covalent bonded surfaces are one of the most popular permanent antithrombotic coatings. The use of heparin-coated biomaterial surfaces in CPB has been shown to decrease platelet adhesion and inflammatory responses resulting from the extensive contact between blood and artificial surfaces.

The two most commonly used heparin-coated systems for CPB are the Carmeda BioActive Surface (CBAS) and the Duraflo II coating. The Carmeda thromboresistant coating process involves covalent binding of biomaterial surfaces and degraded heparin attached to spacer arms (100 Å in length) at endpoint. As a result, the active sequence of heparin is available for binding with ATIII. This prevents surface thrombosis formation. The competing Duraflo II heparin coating ionically binds alkylbenzyltrimethyl ammonium-heparin complex onto the biomaterial surfaces. Slowly released heparin lifts the thromboresistant capability. Optimized coating formula and processing retard heparin leaching from Duraflo II coating. The heparin release capability enables Duraflo II to be distinguished from BioActive surface.

Duraflo II heparin coating is one example of a drug-release-type antithrombotic coating. The application of drug-release-type antithrombotic coatings is also popular for indwelling catheters. For indwelling catheters used in the vascular system, antithrombotic and lubricous properties are equally desired. Lubricious hydrogel coatings, as described earlier, with releasable heparin salt ingredients meet the dual needs of lubricity and thromboresistance. When the coating is hydrated, the microwater layer on the catheter surface promotes the lubricity. Meanwhile, a diffusion mechanism drives the release of heparin from the polymer matrix into the surface water layer.

Released heparin then further prevents blood clotting on the device surfaces.

Anti Cell Adhesion Coating

Although heparin coatings are well known for their thromboresistant functionality, attempts to apply heparin coatings to reduce postoperational cell adhesion on intraocular lenses (IOL) have also been tried.^[7,8] IOL implantation is the most effective surgical procedure in treating cataracts. Unfortunately, postsurgery epithelial cell adhesion to the lens is a major cause of IOL implantation failure. In some clinical reports, heparin coatings on polymethylmethacrylate (PMMA) IOLs have been reported to reduce cell adhesion and to maintain the vision-rehabilitation function of IOLs. However, the overall efficacy of heparin surface-modified IOLs remains controversial. Alternatives to heparin coatings on IOLs are being researched, with albumin-affinity coatings emerging as one with a promising future.^[9] To generate such a coating, PMMA IOLs are passivated with cibacron blue dye, an affinitive compound to albumin. Lenses are first aminated via plasma surface treatment. Amine functional groups on PMMA IOLs are covalently bound with aldehyde groups on the modified cibacron blue dextran. The uniqueness of this coating is its high affinity to albumin in tissue fluid, which enables coated IOLs to be covered by albumin in seconds. When albumin covers the surface, other proteins that promote cell adhesion cannot attach. Thus, the IOL surface becomes less favorable to cell adhesion. The efficacy of the coating is theoretically long-lasting since the attachment of albumin on the coating is reversible and thus the albumin does not become denatured. Larger scale trials of this albumin-affinity cibacron blue dextran coating on live subjects are worthy of future follow-up.

Antimicrobial Coating

Infectious complication frequently predisposes medical device and implant to failure. Various antimicrobial coatings have been developed to address the problem. In general, there are two types of antimicrobial coatings: passive and active ones. Passive antimicrobial coating permanently binds antimicrobial agents onto the device/implant surface and does not allow their release. Microorganisms are killed on contact. Active antimicrobial coatings use the principles of drug-controlled release. They incorporate chemical compounds into the coating matrix for release into the vicinity of devices/implants to perform bactericidal functions.

Silver coatings were typical examples of passive antimicrobial coatings. Silver has been used as an antiseptic compound since the 14th century. Silver coatings kill microorganisms by denaturing membrane protein/polysaccharide upon contact with the silver surface. Elemental silver (Silzone coating) was permanently coated onto the sewing cuff of mechanical valves via ion beam-assisted deposition in order to reduce the risk of prosthetic valve endocarditis (PVE). Passive antimicrobial silver-coated mechanic valves failed to achieve clinical satisfaction because of complications irrelevant to infection. Other applications like permanent silver coatings on surgical tools and supplies show dramatically reduced risk of infection.

Drug release-type antimicrobial coatings are constructed by infusing a polymer matrix with antiseptic components. The polymer matrix can be hydrogel or nonhydrogel-based. During solution preparation, polymer solutions (solvent-based or water-based) are homogenized with various antimicrobial actives (antibiotics and antiseptics). After curing, the polymer matrix tightly attaches to coating substrates. Simple diffusion or ion exchange at the interface of body fluid and coating drive the release of antimicrobial agents. One good example of a nonhydrogel type, antibiotics containing coating that relies on the diffusion mechanism, is minocycline and rifampin containing PVB-coVA-coVA (polyvinyl butyryl-co-vinyl alcohol-co-vinylacetate) coating. Central venous catheters coated with minocycline and rifampin have proven efficacy in clinical environments. Minocycline is primarily bacteriostatic and exerts its antimicrobial effect by inhibiting protein synthesis. Minocycline is an active against a wide range of

gram-negative and gram-positive organisms. Rifampin inhibits bacterial DNA-dependent RNA polymerase activity and is bactericidal in nature. Rifampin possesses a broad spectrum of activity against a wide range of gram-positive and gram-negative bacteria. The combination of minocycline and rifampin has a strategic advantage since microorganisms developing simultaneous resistance to the two antibiotics have not been reported. An unfortunate characteristic of diffusion-dependent release mechanisms is that a high initial loading of active agents is required to prolong the effectiveness. As a result, an initial spike followed by exponential decay is the characteristic release profile.

Another example of antimicrobial coatings is hydrogel-based coatings that release antiseptic compounds in a controlled manner. Ion-exchange mechanisms have been frequently adopted to achieve controlled-release outcomes. In general, antiseptic agents (silver salts, for example) are first ionized in proper solvent systems and homogenized with negatively charged polymers (carboxylated polyurethane, for example). Under mild conditions, ionic bonds are formed between silver ions and the carboxyl groups on the polyurethane. Afterward, the polyurethane containing silver ions is cross-linked and polyvinylpyrrolidone (PVP) is incorporated into the matrix via the previously described interpenetrating network mechanism. At the moment of tissue contact, water molecules adsorbed by PVP hydrate the coating. Ion exchange then occurs between Ag^+ in the polyurethane matrix and Na^+/K^+ in the fluid, as illustrated in Fig. 2. Released Ag^+ in the surrounding devices performs the antimicrobial function. Na^+/K^+ concentration in

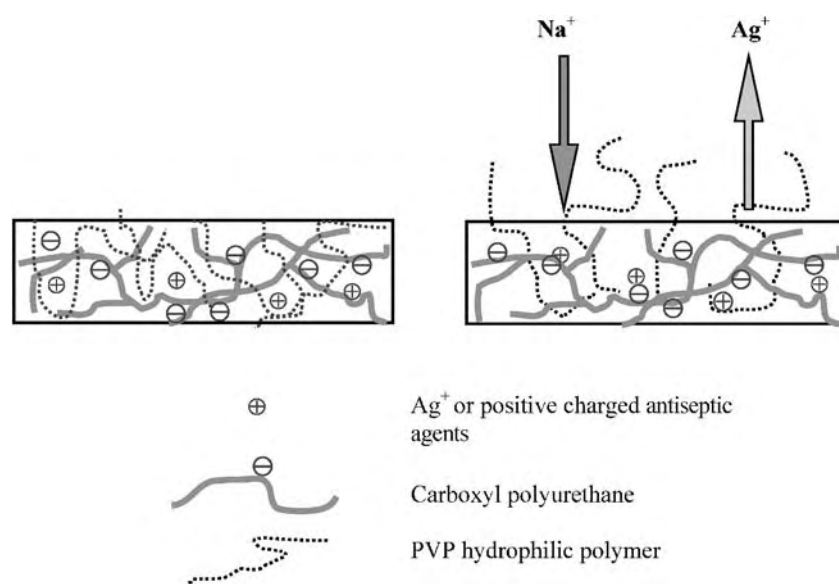


Fig. 2 Controlled-release antimicrobial coating. The mechanism of controlled, long-term release of silver ions from antimicrobial coating. Silver ions or antiseptic agents first couple with the negatively-charged polyurethane matrix. PVP results in a swollen, gellike watery layer. The ion exchange between silver and sodium ions occurs in the watery layer.

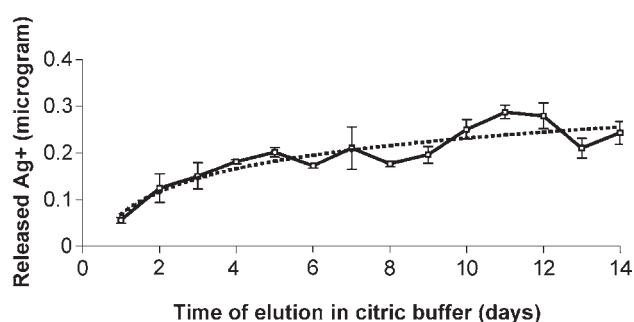


Fig. 3 Silver ion release profile from the controlled-release antimicrobial coating. The silver release profile was conducted by incubating antimicrobial coating in citric buffer at 37°C for two weeks. Buffer solution was replaced daily and the amount of released silver ions was determined by measuring the light absorbance of Bromopyrogallol red-1,10 Phenanthroline-Ag⁺ complex in extracting solutions. The solid line represents the actual measurements and the dotted line indicates the numerical regression result.

the tissue fluid determines the release rate of Ag⁺. Na⁺/K⁺ concentration is normally at a constant level in tissue fluid, which results in constant release profile of Ag⁺ Fig. 3.

Osteoconductivity-Promoting Coating

In some cases, cells' adhesion onto the surface of implants is desired to ensure the functionality of implants. Orthopaedic metal implants for fracture repair (bone nails, plates, and meshes) and joint replacement (knee, hip, and elbow replacement) have been successful because they exhibit two characteristics: superior mechanical functionality and the absence of adverse effects on surrounding soft tissues. However, reported proximal bone resorption, osteolysis, and fibrous ingrowth have frequently compromised the functionality of implants. Hydroxyapatites (HA) coating have been introduced to avoid those complications by enhancing fixation. Hydroxyapatites coatings are applied through a well-established process on both metal and polymer substrates. Among the many techniques used to fabricate HA coatings, plasma-spraying has become the most popular method due to the availability of technology and equipment, cost-effectiveness, efficacy, and reproducibility. Hydroxyapatites had been selected for coating materials due to its excellent osteoconductivity. Bone implants with higher osteoconductivity enabled the growth of bone tissue into gaps with a depth of up to 1 mm without significant formation of intermediate fibrous tissue.^[10] Better ingrowth or ongrowth of bone tissue on implants resulted in stronger fixation.

Surface Coatings Provide Therapeutic Benefits

Drug-release antimicrobial coatings prevent bio-material-related infections and ensure the functionality of medical devices/implants. The success of drug-release antimicrobial coatings reveals a possibility that surface coatings could play an active role in the attainment of dramatic therapeutic results. The development of drug-eluting stents is one vivid example.

Abundant evidence shows that restenosis following stent placement entails a series of interrelated host responses: thrombosis, inflammation, and smooth muscle cell (SMC) proliferation. Stent treatments have addressed each of these three responses. The earliest view held that metal stent strut-induced thrombogenesis ultimately resulted in restenosis. Thus, major efforts were made to reduce the thrombogenic property of metal stents. As thrombosis was controlled, it became clear that other (inflammatory) processes governed the restenosis rate. Broader strategies, employing both passive and active surface modification, were called for to reduce the host response. The so-called passivation methods altered selected properties of the stent surface without delivery of drugs or other agents. These techniques were expected to remove surface initiators of the host response, although less intense cellular responses to the surface may continue. In contrast, active surface modification provided a drug-release matrix to block or otherwise alter cellular responses. Significant reduction of platelet aggregation, inflammation, and SMC proliferation was achieved by some of these means.

Surface Passivation

The passivation approach is frequently employed to generate a thromboresistant coating. Both organic and inorganic compounds are used. Among many non-polymeric organic chemicals, phosphorylcholine (PC) has had the most attention as a stent passivation method. PC mimics phosphatidylcholine, a major component of the outer membrane layer of cells that inhibits the activation of formed and unformed elements of blood. There is good evidence to show that PC-coated stents reduce thrombotic complications. A recent clinical trial of the PC-coated Bio-divYsio stent yielded encouraging results. In six-month implantations, angiographic evidence of stenosis decreased from 89 to 5.6%.^[11,12] However, the PC coating is subject to erosion. Optimization of the PC-coating and longer-term studies are needed to prove its superiority.

Organic polymeric materials are also used for stent surface modification. Both biostable and biodegradable polymer coatings have been studied. Most polymer-coated stents are in the developmental stage and

frequently have been tested only in experimental animal models. Biostable polymer coatings include expanded polytetrafluoroethylene (ePTFE), polyester (Dacron), a blend of methylmethacrylate (MMA) and 2-hydroxyethyl methacrylate (HEMA), fluorine-acrylstyrene-urethane-silicon (FASUS) copolymer, parylene, and amphiphilic polyurethane. Variable success rates in reducing thrombus formation or neointimal growth have been reported. Biodegradable polymer poly(organo)phosphazene and poly(L lactic acid) (PLLA)-coated stents have also been reported to be less thrombogenic. Despite the short-term successes, a major concern for all these polymer coatings is the longer-term tissue/materials response. As noted earlier, mild to severe foreign-body responses associated with biodegradable and biostable polymer coatings have been observed. In some cases, an overwhelming inflammatory response has been closely linked to later restenosis, an unacceptable outcome.

Inorganic compounds—including gold, silicon carbide, hydrogenated carbon, and titanium nitride/oxide—have been tested as thromboresistant surfaces for metal stents. Gold is believed to be a biocompatible material. However, gold-coated stents revealed poor thromboresistance in an *ex vivo* human stasis model. The 12-month outcome from a randomized trial of a gold-coated stent yielded less favorable neointimal proliferation.^[13] Silicon carbide, postulated as a less thrombogenic semiconducting material, was shown to inhibit IgG, fibrinogen, and fibronectin adsorption. Fibrinogen is easily activated when it contacts a metallic surface, and activated fibrinogen, in the presence of thrombin, can be converted into fibrin to start the clotting process. Patients implanted with silicon carbide-coated stents showed less acute and subacute stent thrombosis, even when only antiplatelet medication was received postprocedure. This finding was confirmed in a larger scale randomized trial. However, the long-term benefit of silicon carbide coating was not confirmed in the same study.^[14] Hydrogenated diamondlike carbon and titanium nitride oxide are other inert materials that have been applied as stent coatings. By means of plasma or electronic beam techniques, diamondlike nanocomposites and titanium nitride/oxide coatings were deposited onto stent surfaces. Significant reductions of adherent thrombus on both coated stents were demonstrated in a short-term, porcine model study. Reduction of neointimal proliferation was observed only for the titanium nitride/oxide coating.

In general, metal stent surface modifications by these passive approaches achieve only a certain degree of thromboresistance, without significant reduction of neointimal hyperplasia. Thus, approaches that are more aggressive must be taken in order to reach a satisfactory outcome.

Active Surface Modification

The active approach aims for tissue responses beyond thrombosis resistance, since thromboresistant surfaces are achievable with some surface passivation methods. A number of active surface modification approaches are based on the successful development of stent-coating procedures. These stent coatings function as reservoirs for specific therapeutic agents. The rationale is that drugs are gradually and locally released from the coating matrix and taken up by surrounding tissues. A reduction in neointimal hyperplasia is expected, at least for the period that the beneficial drug actions remain in effect and perhaps beyond, if healing processes are sufficiently advanced. An optimal combination of drugs and coating matrix maximizes the local therapeutic benefit, thereby maintaining the patency of the stented vessel.

A variety of pharmacologic agents have been used to prevent thrombosis and restenosis. The agents are loaded in bulk by physical entrapment or by chemically bonding to the polymer chain.^[15] Four types of drugs have been studied in experimental animal models or human clinical trials: anticoagulants, antiplatelet, antiproliferative, and immunosuppressive agents.

The rationale for immobilized heparin, the most intensively investigated stent-coating agent, is to inhibit thrombin-dependent positive feedback reactions at the stent surface. Heparin-coated stents have been used in a primary antirestenosis strategy (BENESTENT II, MENTOR) and in the context of primary angioplasty for acute myocardial infarction (PAMI stent pilot trial) and total coronary occlusion (TOSCA).^[16–19] Although thrombogenicity was reduced in the BENESTENT II study, investigators doubted the importance of the heparin coating. In the longer term (six-month follow-up), the heparin coating did show improved angiographic evidence of restenosis. Stents coated with the anticoagulant hirudin have been shown to reduce neointimal formation. Despite these findings, it is now generally believed that anticoagulant coatings are unnecessary when proper stent deployment techniques are coupled with effective antiplatelet medication, e.g., aspirin and ticlopidine.

Antiplatelet agents employed in drug-eluting stents include nitric oxide donors, glycoprotein IIb/IIIa receptor antagonist/antibodies, and thrombin inhibitors. Significant reductions in platelet deposition and thrombosis have been demonstrated for several of these agents in animal studies. Unfortunately, no indication was revealed that the palliation of thrombosis by these means resulted in the inhibition of neointimal growth.

Antiproliferative drug-eluting stents combat restenosis by directly inhibiting vessel smooth muscle cell (SMC) proliferation. Drugs designed to inhibit SMC

migration, or to arrest the SMC cell cycle at the G2/M transition have been studied.^[20,21] Stents coated with such drugs proved to reduce neointimal growth in animal studies. Paclitaxel-eluting stents attracted particular attention. A paclitaxel-eluting stent was the first device reported to have six-month inhibitory efficacy on neointimal growth.^[22] This six-month efficacy was reproduced in a human clinical trial, although the antiproliferative efficacy was not maintained in the 12-month follow-up, as measured by angiographic restenosis.^[23] Postponed thrombosis and uncompleted wound healing were later reported in more thorough animal examinations.^[24] The paclitaxel-eluting stent experience nevertheless implies that the optimal combination of drug and coating matrix may yield a long-term, restenosis-free therapy.

A fourth, exciting area of progress concerns stents coated with immunosuppressive drugs. These coatings attack neointimal growth in a different way, possessing both potent antiinflammatory and immunosuppressive effects. Such drugs have been shown to decrease the migration and functional capabilities of inflammatory cells. Some of these drugs prevent cell cycle progression in the G1 stage, and consequently inhibit cell proliferation. In a pig coronary artery model, methylprednisolone-eluting stents showed promise in reducing the foreign-body reaction and restenosis at six weeks. In one month, porcine and rabbit models, a rapamycin (sirolimus)-eluting stent, inhibited restenosis. Moreover, the long-term behavior of the sirolimus-eluting stent is as encouraging as its short-term property. Multiple experiences in human clinical trials show that the antirestenotic efficacy of the sirolimus-eluting stents is sustained for 12 months.^[25–27] In contrast to their opinion of paclitaxel-eluting stents, interventional cardiologists are optimistic about the 12-month follow-up angiographic outcome with this device.

The benefit of drug-release therapeutic coatings is evident. Metal stents were invented as supportive mechanical structures to rescue artery occlusions, but never obtain satisfactory clinical outcomes. Thanks to the active, drug-eluting strategy of surface coatings, coated stents ignite a promising future in overcoming restenosis.

CONCLUSION

As recently as just 20 years ago, surface coatings were considered mere supplements to medical devices and implants. As knowledge accumulated about host responses to biomaterials, surface coatings became more indispensable to the manufacture of safe medical devices and implants. Coatings provide benefits of lubricity, electric isolation, wettability, enhanced

imaging, antiencrustation, thromboresistance, anti or promoting cell adhesions, antiinfection, and anti-restenosis. The recent clinical success of drug-eluting stents moves the application of surface technology forward. Surface coatings are beginning to be regarded as equally important to the success of the device as the physical device itself. In the near future, surface coatings are expected to be an important partner in the achievement of desired therapeutic results in many diseases.

ARTICLES OF FURTHER INTEREST

Antimicrobial/Antibiotic (Infection Resistance) Materials; Bioadhesion; Biofilms; Biomaterials: Protein–Surface Interactions; Cardiac Catheters; Diamond and Diamond-Like Carbons; Phosphorylcholine (PC) Technology; Plasma Modification of Materials; Protein Adsorption Modeling; Self-Assembled Monolayers; Surface Modification

REFERENCES

1. Catapano, G.; Di Lorenzo, M.C.; Della Volpe, C.; De Bartolo, L.; Migliaresi, C. Polymeric membranes for hybrid liver support devices: The effect of membrane surface wettability on hepatocyte viability and functions. *J. Biomater. Sci., Polym. Ed.* **1996**, *7* (11), 1017–1027.
2. Loh, I.H.; Hudson, D.M. US patent 5,355,832, December 15, 1992.
3. Loh, I.H.; Hudson, D.M. US patent 5,447,799, October 18, 1994.
4. Sahagian, R. Critical insight: Marking devices with radiopaque coatings. *Med. Device and Diag. Ind.* **May 1999**.
5. Hildebrandt, P.; Sayyad, M.; Rzany, A.; Schaldach, M.; Seiter, H. Prevention of surface encrustation of urological implants by coating with inhibitors. *Biomaterials* **2001**, *22* (5), 503–507.
6. Morris, N.S.; Stickler, D.J. The effect of urease inhibitors on the encrustation of urethral catheters. *Urol. Res.* **1998**, *26* (4), 275–299.
7. Petrus, M.; Czajkowski, G. Cellular response one year after extracapsular cataract extraction—comparison of regular and heparin surface modified PMMA intraocular lenses. *Klin. Oczna* **2000**, *102* (3), 165–168.
8. Dada, T.; Dada, V.K.; Sharma, N.; Vajpayee, R.B. Primary posterior capsulorhexis with optic capture and intracameral heparin in paediatric cataract surgery. *Clin. Experiment Ophthalmol.* **2000**, *28*, 361–363.
9. Tang, L.; Su, S.H.; Xiao, T.L. Adsorbed Fibrin(ogen) Triggers Posterior Capsule Opacification in Young Rabbits. *Proceedings of the Association for Research in Vision and Ophthalmology*, Fort Lauderdale, Florida, May 5–10, 2002; LE128.

10. Strnad, Z.; Strnad, J.; Urban, K. Effect of plasma-sprayed hydroxyapatite coating on the osteoconductivity of commercially pure titanium implants. *Int. J. Oral. Maxillofac Implants.* **2000**, *15* (4), 483–490.
11. Atalar, E.; Haznedaroglu, I.; Aytemir, K.; Aksoyek, S.; Ovunc, K.; Oto, A.; Ozmen, F. Effects of stent coating on platelets and endothelial cells after intracoronary stent implantation. *Clin. Cardiol.* **2001**, *24*, 159–164.
12. Grenadier, E.; Roguin, A.; Hertz, I.; Peled, B.; Boulos, M.; Nikolsky, E.; Amikam, S.; Kerner, A.; Cohen, S.; Beyar, R. Stenting very small coronary narrowings (<2 mm) using the biocompatible phosphorylcholine-coated coronary stent. *Cath. Cardiovasc. Interv.* **2002**, *55*, 303–308.
13. Kastrati, A.; Schomig, A.; Dirschinger, J.; Mehilli, J.; von Welser, N.; Pache, J.; Schuhlen, H.; Schilling, T.; Schmitt, C.; Neumann, F.J. Increased risk of restenosis after placement of gold-coated stents: results of a randomized trial comparing gold-coated with uncoated steel stents in patients with coronary artery disease. *Circulation* **2000**, *101*, 2478–2483.
14. Kalnins, U.; Erglis, A.; Dinne, I.; Kumsars, I.; Jegere, S. Clinical outcomes of silicon carbide coated stents in patients with coronary artery disease. *Med. Sci. Monit.* **2002**, *8*, PI16–20.
15. Peng, T.; Gibula, P.; Yao, K.; Goosen, M.F.A. Role of polymers in improving the results of stenting in coronary arteries. *Biomaterials* **1996**, *17*, 685–694.
16. Serruys, P.W.; Emanuelsson, H.; van der Giessen, W. Heparin-coated Palmaz-Schatz stents in human coronary arteries: Early outcome of the Benestent-II pilot study. *Circulation* **1996**, *93*, 412–422.
17. Vrolix, M.C.; Legrand, V.M.; Reiber, J.H.; Grollier, G.; Schali, M.J.; Brunel, P.; Martinez-Elbal, L.; Gomez-Rocio, M.; Bar, F.W.; Bertrand, M.E.; Colombo, A.; Brachman, J. Heparin-coated Wiktor stents in human coronary arteries (MENTOR trial). MENTOR trial investigators. *Am. J. Cardiol.* **2000**, *86*, 385–389.
18. Stone, G.W.; Brodie, B.R.; Griffin, J.J.; Costantini, C.; Morice, M.C.; St. Goar, F.G.; Overlie, P.A.; Popma, J.J.; McDonnell, J.; Jones, D.; O'Neill, W.W.; Grines, C.L. Clinical and angiographic follow-up after primary stenting in acute myocardial infarction: the Primary Angioplasty in Myocardial Infarction (PAMI) stent pilot trial. *Circulation* **1999**, *99*, 1548–1554.
19. Buller, C.E.; Dzavik, V.; Carere, R.G.; Mancini, G.B.; Barbeau, G.; Lazzam, C.; Anderson, T.J.; Knudtson, M.L.; Marquis, J.F.; Suzuki, T.; Cohen, E.A.; Fox, R.S.; Teo, K.K. Primary stenting versus balloon angioplasty in occluded coronary arteries: the Total Occlusion Study of Canada (TOSCA). *Circulation* **1999**, *100*, 236–242.
20. Mooradian, D.L.; Fernandes, B.; Diglio, C.A.; Lester, B.R. Angiopeptin (BIM23014C) inhibits vascular smooth muscle cell migration in vitro through a G-protein-mediated pathway and is associated with inhibition of adenylyl cyclase and cyclic AMP accumulation. *J. Cardiovasc. Pharmacol.* **1995**, *25*, 611–618.
21. Sindermann, J.R.; Skaletz-Rorowski, A.; Bartels, A.; Hohage, H.; Plenz, G.; Schmidt, A.; Breithardt, G. Paclitaxel and cyclosporine a show supra-additive antiproliferative effects on smooth muscle cells by activation of protein kinase C. *Basic Res. Cardiol.* **2002**, *97*, 125–131.
22. Drachman, D.E.; Edelman, E.R.; Seifert, P.; Groothuis, A.R.; Bornstein, D.A.; Kamath, K.R.; Palasis, M.; Yang, D.; Nott, S.H.; Rogers, C. Neointimal thickening after stent delivery of paclitaxel: change in composition and arrest of growth over six months. *J. Am. Coll. Cardiol.* **2000**, *36*, 2325–2332.
23. Liistro, F.; Stankovic, G.; Di Mario, C.; Takagi, T.; Chieffo, A.; Moshiri, S.; Montorfano, M.; Carlino, M.; Briguori, C.; Pagnotta, P.; Albiero, R.; Corvaja, N.; Colombo, A. First clinical experience with a paclitaxel derivate-eluting polymer stent system implantation for in-stent restenosis: Immediate and long-term clinical and angiographic outcome. *Circulation* **2002**, *105*, 1883–1886.
24. Farb, A.; Heller, P.F.; Shroff, S.; Cheng, L.; Kolodgie, F.D.; Carter, A.J.; Scott, D.S.; Froehlich, J.; Virmani, R. Pathological analysis of local delivery of paclitaxel via a polymer-coated stent. *Circulation* **2001**, *104*, 473–479.
25. Morice, M.C.; Serruys, P.W.; Sousa, J.E.; Fajadet, J.; BanHayashi, E.; Perin, M.; Colombo, A.; Schuler, G.; Barragan, P.; Guagliumi, G.; Molnar, F.; Falotico, R. A randomized comparison of a sirolimus-eluting stent with a standard stent for coronary revascularization. *N. Engl. J. Med.* **2002**, *346*, 1773–1780.
26. Sousa, J.E.; Costa, M.A.; Abizaid, A.C.; Rensing, B.J.; Abizaid, A.S.; Tanajura, L.F.; Kozuma, K.; Van Langenhove, G.; Sousa, A.G.; Falotico, R.; Jaeger, J.; Popma, J.J.; Serruys, P.W. Sustained suppression of neointimal proliferation by sirolimus-eluting stents: one-year angiographic and intravascular ultrasound follow-up. *Circulation* **2001**, *104*, 2007–2011.
27. Serruys, P.W.; Degertekin, M.; Tanabe, K.; Abizaid, A.; Sousa, J.E.; Colombo, A.; Guagliumi, G.; Wijns, W.; Lindeboom, W.K.; Ligthart, J.; de Feyter, P.J.; Morice, M.C. Intravascular ultrasound findings in the multicenter, randomized, double-blind RAVEL (Randomized study with the sirolimus-eluting VELOCITY balloon-expandable stent in the treatment of patients with de novo native coronary artery Lesions) trial. *Circulation* **2002**, *106*, 798–803.



Surface Modification

Sally L. McArthur

Department of Engineering Materials, University of Sheffield, Sheffield, United Kingdom

Keith M. McLean

Ian Wark Laboratories, Commonwealth Scientific & Industrial Research Organisation (CSIRO), Molecular Science, Clayton, Victoria, Australia

INTRODUCTION

While materials selection for most biomedical devices is generally based upon bulk properties, it is the surface's properties that may ultimately determine the success or failure of the device. Today the need for efficient and effective methods of surface modification for biomedical applications is becoming increasingly evident. Advances in biological array, diagnostics, and tissue engineering are currently hampered by problems that can, in large part, be controlled by tailoring the surface properties. Surface modification aims to tailor the surface characteristics of a material for a specific application without detrimentally affecting the bulk properties. At present a range of biological, physical, and chemical methods are used to effect surface modifications on biomedical devices and biomaterials. This article introduces and discusses the application of a number surface modification techniques used in both industrial production and fundamental research of materials for biomedical applications.

SURFACE MODIFICATION METHODS

Surface modification generally falls into one of two categories: 1) modification of the existing surface, e.g., etching or chemical modification or 2) coating the surface with a different material, e.g., plasma deposition or polymer grafting. As illustrated in Fig. 1, these approaches can be used to affect a range of properties, including wettability, permeability, biostability and/or chemical inertness, adhesion, biocompatibility, topography, electrical characteristics, and optical and frictional properties.

While complicated and expensive surface modification techniques may be applicable in fundamental research, these may be impractical when considering a commercial application. Table 1 lists some of the common surface modification techniques used at present in biomedical engineering and the materials properties they influence. A successful commercial product therefore requires any surface modification to be one that is readily transferable to a larger scale and

preferably to a variety of substrates. In addition, any surface modification should be physically, chemically, and biologically characterized. A range of highly sensitive surface analytical methods exist and their applications in biomedical and biomaterials applications have been recently reviewed by Castner and Ratner.^[36]

RADIATION-BASED SURFACE MODIFICATION METHODS

Radiation-based surface modification techniques, including gamma, ultraviolet (UV), laser, electron, and ion beams, can be used to surface modify biomedical materials.^[18] Specific applications range from ion implantation for improved wear, corrosion and biocompatibility of metal alloys in orthopedics and dentistry to surface chemical patterning via photolithography and radiation-initiated polymer grafting.

Ion Beam Techniques

Ion beam techniques can be considered in one of two categories: 1) thin film formers, e.g., ion beam-assisted deposition (IBAD); and 2) surface chemical modifiers, e.g., ion implantation. In both applications, an ion source is used to produce the appropriate ionized species that are then separated and subsequently accelerated in an electric field. The ions reach the surface with energies between 20 and 200 keV,^[37] which enables the delivered ion beam to induce modifications of the macromolecular structure through gas evolution, formation of double bonds, chain scissions, and cross-linking over a thickness corresponding to the penetration depth of the ions.^[21] Major parameters affecting ion implantation are the ion species, the beam energy and dose, and the beam current density.^[20]

In the case of metals and ceramics, the atomic and nuclear collisions induced by ion implantation lead to the formation of highly disordered and sometimes amorphous structures. Chemical processes result in the formation of new surface phases with significantly different mechanical properties from the bulk.^[20] The formation

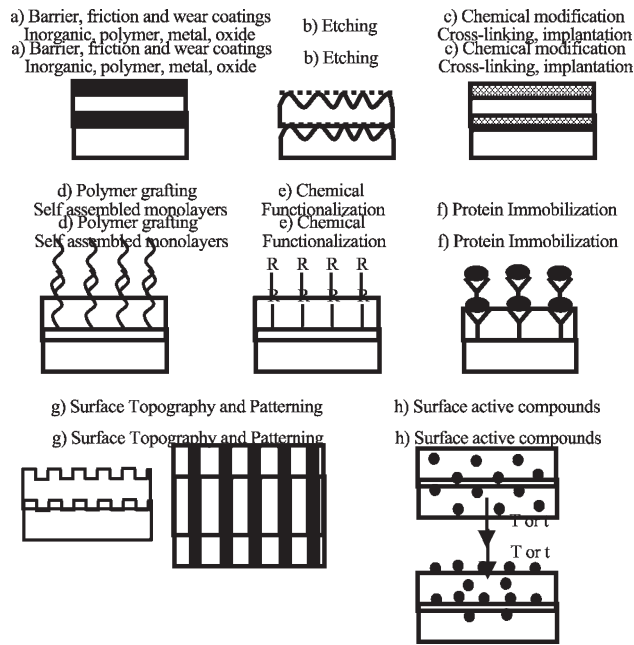


Fig. 1 Surface modification methods used in biomaterials and biomedical engineering.

of hard phase nitride, carbide, and oxide precipitates hardens the surface, which improves the wear resistance of titanium and titanium alloy orthopedic implants.^[38] Changes induced in the crystalline structures of titanium alloys 316L stainless steels and Co–Cr alloys have also been shown to improve friction properties.^[20,39]

In the case of polymers, chain cross-linking and scission compete during the ion implantation process.^[21] Cross-linking results in surfaces with higher wear resistance, but polymer degradation can occur if processing parameters are not well-controlled.^[20] Ion implantation has been used to increase surface hydrophilicity of silicone catheters to reduce the incidence of thrombus formation^[40] and improve cell adhesion, proliferation, and spreading on polyurethane and polyethersulphone.^[41] Nitrogen ion implantation of polycarbonate results in complex three-dimensional, cross-linked surface, which results in increased solvent resistance and significant improvements in the wear and friction characteristics.^[42]

Ion implantation can also be used to introduce specific chemical species at the surfaces of polymers, ceramics, and metals.^[19,20,38] The wear, corrosion resistance, and bone conductivity of titanium alloys have all been shown to be improved with nitrogen, calcium, and/or phosphorus ion implantation, respectively.^[38,43] In polymeric materials, the incorporation of silver (Ag) ions have been used to impart both the antimicrobial properties that are typically associated with silver and reduce the occurrence of thrombus formation on indwelling catheters.^[22,44]

The second category of ion beam process is the thin film former. The bone conductivity, corrosion, and

wear resistance of metal substrates can be improved with the formation of apatite, TiO₂, and TiN coatings.^[38] Traditional plasma spraying of these materials results in a film with controlled chemical composition but poor adhesion properties.^[38] Ion beam-assisted deposition (IBAD) combines physical vapor deposition with ion beam implantation. An electron beam evaporator or plasma vapor deposition (PVD) is used to generate the metallic or inorganic coating atoms as gaseous ions are simultaneously accelerated towards the surface. Ions interact with the coating atoms, which drives them into the surface, and produces a low-stress, uniform, and adherent coating.^[20]

IBAD has been used to produce hydroxyapatite coatings on titanium alloys, alumina, and UHMWPE with a greater coating/substrate adhesive strength than the traditional plasma spraying methods.^[45] Thin, clear sealant coatings of alumina have also been developed to prevent oxygen and water vapor permeation into polymers.^[20,37] The process has also been used to produce titanium alloy coatings on Co–Cr substrates for improved cellular integration in orthopedic applications^[6] and C–N coatings that are both biocompatible and wear-resistant.^[20] Commercially, IBAD is used to produce antimicrobial silver coatings that will adhere to a range of polymeric substrates^[21,22,37] and diamondlike carbon (DLC) coatings that are chemically inert, optically transparent, extremely hard, and have a low friction coefficient.^[20] These DLC coatings are used to treat the bearing surfaces of orthopedic implants to improve wear and friction properties and reduce the incidence of wear debris.^[37]

Laser, Ultraviolet, and Gamma Radiation Techniques

Controlled exposure of a surface to excimer lasers and UV lamps can produce a range of surface chemical modifications, including etching, deposition, and polymerization.

UV and laser treatments have been demonstrated to introduce oxygen containing species on the surface of expanded-polytetrafluoroethylene (e-PTFE)^[18] and poly(ethylene terephthalate) (PET).^[17] This improves surface wettability and changes surface morphology plus improves cell proliferation and growth on surfaces that would normally display poor cell growth characteristics.^[18] Gamma irradiation and vacuum ultraviolet (VUV) radiation have both been used to improve the wear resistance on the load-bearing surfaces of orthopedic implants and cross-link UHMWPE to reduce the incidence of wear-related failure.^[46]

Radiation techniques have long been recognized as a method of initiating polymer and biomolecule grafting on a variety of polymeric substrates.^[18] In the

Table 1 Methods and effects of surface modification techniques used for biomaterials

Method	Modification	Properties	References
Plasma Techniques			
Polymerization	Organic and inorganic coatings with tailored physicochemical properties.	Barrier coatings (thermal and chemical). Improved electrical, optical, and abrasion resistance properties. Control of chemical functionality, cell, and protein adhesion.	[1–3]
Plasma immersion or source Ion implantation (PIII)	Chemical modifications. Deposition of metals layers.	Wear resistance and improved friction properties for metals ceramics and polymers. Improved biocompatibility.	[4–6]
Sputtering and etching	Chemical and physical deposition and etching.	Surface cleaning, introduction of topographical features and chemical functionality.	[1,3,7,8]
Spraying	Deposition of metals and inorganic coatings.	Enhanced corrosion and ablation resistance. Improved biocompatibility.	[9]
Corona discharge	Surface oxidation.	Surface chemical functionalization and biocompatibility.	[8,10]
Grafting copolymerization	Creation of radicals to initiate polymerisation.	Polymer grafting for enhanced biocompatibility and nonfouling.	[10–12]
UV and Gamma and Laser irradiation	Chemical functionalization and etching.	Polymer grafting, enhanced biocompatibility, introduction of topographical features.	[13–18]
Ion implantation	Implantation of new chemical species. Changes in surface structure.	Improved wear and friction properties. Improved passivation and biocompatibility of polymers and metals.	[19–21]
Ion beam-assisted deposition (IBAD)	Metallic and inorganic coatings with good adhesive properties.	Enhanced biocompatibility, antimicrobial properties, friction, wear, and chemical stability.	[20–22]
Wet Chemistry			
Polymer grafting	Physical and chemical.	Nonfouling and biomimetic surfaces. Enhanced hydrophilicity, functionality, and biocompatibility.	[23–26]
Biomolecule attachment	Chemical and biological.	Biomimetic surfaces, introduction of specific biological function, and activity.	[27–30]
Molecular imprinting	Chemical and topographical.	Cell and protein selectivity and improved biological functionality.	[31,32]
Self assembly	Chemical and biological.	Introduce functionality, nonfouling, and biomimetic properties.	[33–35]

majority of the studies, radiation grafting has been implemented as a means of improving the biocompatibility of the device surface.^[13,15,47] Gamma, laser, and UV radiation have been used to immobilize poly(ethylene oxide) (PEO) containing Pluronic[®] surfactants,^[48,49] hydroxethyl methacrylate (HEMA),^[15] n-isopropylacrylamide (NIPPAAm),^[50] and acrylic acid^[14,47] to a variety of biomaterials.

There is increasing interest in the spatial control of protein and cell adhesion for biosensors and tissue engineering. The introduction of masked irradiation and the ability to focus the energy source has enabled laser and UV radiation techniques to revolutionize the spatial manipulation of surface topography and chemistry.^[14] By incorporating photosensitive groups into the polymer backbone, Ward et al. have shown that it is possible to use a mask and a UV source to micropattern a range of biomedical polymers, including poly(ethylene glycol) and create surfaces that are capable of selective protein adsorption and cell adhesion.^[51] In other studies, laser ablation has been used to micropattern surfaces to control cell location and shape^[52] while a UV laser coupled to a scanning near-field optical microscope has been used to induce photo-oxidation and nanometer scale patterns in self-assembled monolayers.^[53] Laser excimer beam technology has also been used to introduce high-resolution topography on polymeric surfaces to investigate the effects of three-dimensional structures on cell/surface interactions and as a possible route to controlling cell proliferation and growth.^[54]

PLASMA SURFACE MODIFICATION TECHNIQUES

Plasmas consist of highly excited atomic, molecular, ionic, and radical species. The excitation can be achieved using gaseous (e.g., radio frequency, corona, or microwave), metallic, and laser-based sources. The resulting gaseous environment is highly reactive and capable of eliciting a range of plasma–surface interactions, even on inert materials such as ceramics. Plasma-based treatments enable adhesive metallic, ceramic, and polymeric coatings to be applied to a range of materials. In specific formats, they can be used to etch, cross-link and oxidize materials, and introduce specific chemical functionality at the material interface. These factors make plasma techniques particularly useful for the surface modification of biomedical materials, e.g., in the oxidation of polystyrene for tissue culture applications. In this section, two techniques, plasma source and plasma immersion ion implantation (PSII and PIII) and plasma modification/polymerisation, are discussed. For further detail of the

application of plasmas to biomaterials, the reader is referred to recent reviews by Chu,^[1] Chan,^[8] and Favia.^[3]

Plasma Ion Implantation Methods

Plasma ion implantation techniques, such as plasma source ion implantation (PSII) and plasma immersion ion implantation (PIII), are able to introduce elements into the surface without thermodynamic constraints.^[1] They have one major advantage over the standard ion implantation methods discussed in the previous section. They are not line of sight techniques, a factor that enables the modification of complex shapes commonly found in biomedical applications.^[5]

In PIII and PSII, samples are surrounded by high-density plasma and pulse-biased to a high negative potential relative to the chamber wall. Ions generated in the plasma are accelerated across the sheath formed around the samples and implanted into the surface. The gaseous plasma can be induced using a variety of sources, including radiofrequency and microwave. Typically, oxygen is used in these systems to produce Ti–O coatings on the surface of titanium alloy (Ti₆Al₄V and NiTi shape memory alloy) implants. The resulting coating improves the mechanical properties, wear resistance, corrosion, and biocompatibility of the device.^[4,5,55] Treatments using argon, xenon, oxygen, and nitrogen have also been used to improve the electrical and adhesion properties of a number of polymers.^[56,57]

Traditionally it is very difficult to make a good adherent metal layer on a polymer due to the lack of chemical bonds. Combinations of gas plasma systems with metallic plasma from vacuum or cathodic arc plasma sources enable a number of elements to be introduced to the plasma simultaneously. These plasma-assisted processes allow interface mixing, which results in a coating with low intrinsic stress, a factor that reduces the probability of delamination on both polymeric and metallic substrates.^[58] This approach has been used to surface modify skeletal prosthetic implants with titanium alloy coatings for cell recruitment while maintaining the mechanical properties of the Co–Cr substrate^[55] and deposit carbon and Ti–N coatings on both metals and polymers from improved wear and scratch resistance.^[58]

Plasma Modification and Polymerization Methods

Radio frequency glow discharge (RFGD) or plasma polymers have been applied to biomaterials and biomedical devices since the 1960s^[59] (Table 2). An electric field at radio frequency is applied across a vapor or

Table 2 Applications of RFGD plasma polymers and surface modification in in-vitro and in-vivo biomedical applications

Ophthalmology—IOLs, Contact lenses and Artificial cornea ^[7,60–62]
Vascular prostheses ^[63]
Catheters ^[64] and electrodes ^[65]
Blood compatible coatings ^[1,66]
Cell culture surfaces ^[67,68]
Biosensor chip coatings ^[69,70]
Model surfaces for biological interaction studies ^[71–76]
Functionalized substrates for:
Biomolecule immobilization ^[28,77–81]
Polymer grafting and copolymerization ^[82]
Surface chemical patterning ^[83,84]

gas-filled chamber, ionizing a fraction of the molecules, and generating electrons, ions, free radicals, photons, and molecules in both ground and excited states within the gas plasma.

Plasmas can be used to modify surfaces by etching or coating surfaces by deposition. In the case of surface modification, the interaction between plasma and substrate leads to two competitive reactions, modification and degradation. When degradation is prominent, etching will occur on the polymer surface. The rate of materials loss depends on both the plasma energy and the nature of the polymer substrate.^[1] Plasma etching or cleaning is commonly used to remove contaminants from the surface and improve the adhesive properties of both metals and polymers. With all plasma treatments, there is a possibility that surface reactions will result in unexpected surface bound

species. The selection of the treatment gas and plasma conditions is critical in ensuring that the reaction products do not readsorb to the cleaned surface.^[85] In some instances, a specific phase of the material may be more susceptible to the plasma, which results in an etched surface with the chemistry of the plasma-resistant material.

In the modification regime, chemical species are introduced to the surface of the material. These non-depositing plasmas typically use inert gases such as argon, ammonia, oxygen, nitrogen, air, and water to influence surface chemistry and topography. A common application for these types of chemical modifications are as pre-activation step for polymer copolymerization in which the plasma creates radicals in the polymer chains that can then be used to initiate polymerization when placed in contact with monomers in the liquid or gas phase and form grafted copolymers at the surface.^[8] Specific gases such as oxygen or nitrogen can also be used to produce TiO₂ and TiN modifications on the surfaces of titanium alloy orthopedic implants.^[85] Plasma treatments are also used in a range of biomedical applications to introduce specific functional groups at the interface that enable, e.g., biomolecule and drug immobilization.^[81,86] The presence of specific chemical species on plasma-modified surfaces has also been shown to influence cell attachment and growth in cell culture.^[87]

In the case of surface-depositing plasmas, the deposition process involves the volatilization of a liquid monomer into an evacuated process chamber. Excitation of the monomer results in reactive species impinging on a surface within the plasma zone, which creates reactive sites that may be available for the covalent attachment of other species. The process results in the deposition of a polymeric film from the vapor phase.^[73] As detailed in Table 3, plasma deposition has

Table 3 Advantages of RFGD plasma deposited films in biomedical applications

Conformal Films: Due to their gaseous nature, plasma can penetrate complex structures and conform to their surface.
Substrate Adhesion: The polymer films adhere well to a range of substrates because free electrons impinge on the surface and create sites for covalent attachment.
Low Leachables: The polymer is highly cross-linked and so contains negligible amounts of low molecular weight material that may cause adverse biological responses. The presence of the film may in some cases prevent leaching of additives from the bulk material.
Barrier Function: The densely cross-linked and pinhole-free nature of these coatings make them adaptable to barrier and protective applications.
Sterility: When removed from the reactor system, the coatings are sterile.
Range of Chemical Functionalities: A wide range of monomers are available that are able to introduce a variety of chemical functionality.
Surface Specificity: The film thickness can be closely controlled, which results in a negligible effect on bulk mechanical properties.
Rapid processing: Films are easily prepared following process optimization.

(From Refs. [7,73].)

a number of specific advantages over alternate surface modification techniques with a key strength lying in its ability to form an adherent coating on a range of materials.

A wide range of monomers have been used to produce plasma polymer coatings with tailored surface characteristics. Organo-silicon monomers such as silanes and disiloxanes are used to produce coatings with thermal and chemical resistance as well as specific electrical optical properties. Fluorine- and hydrocarbon-containing monomers are used to produce hydrophobic surfaces while acid-containing monomers produce hydrophilic coatings.^[1] Amine and acid-rich plasma polymers are used to provide a stable functionalized surface for the attachment of biomolecules^[28,77-80] while a range of monomers can be used to functionalize surfaces prior to wet chemical grafting of polymers.^[94,95] In general, plasma polymers do not reproduce the chemistry of the monomer. Instead, the process tends to form complex, cross-linked polymers with random structures that incorporate various structural fragments of the monomer.^[73]

Plasma-polymer surfaces that retain the functionality and specific characteristics of their monomer are of increasing interest and include poly(ethylene oxide)like surfaces for protein resistance and nonfouling,^[72,88] thermally responsive N-isopropylacrylamide (NIP-PAAm) surfaces,^[89] and hydrogel hydroxyethyl methacrylate (pHEMA).^[90,91] Monomer characteristics can be retained by modifying the deposition condition by using lower powers or pulsed systems^[7,88-90,92] or copolymerising the functionalized monomer with a diluent hydrocarbon monomer.^[93] The resulting materials have been shown to retain higher monomer functionality and, in some cases, specific physicochemical properties normally associated with multi-step polymer grafted surfaces. The development of single-step plasma-polymer films present distinct advantages for the production of biomedical devices.

WET CHEMISTRY

Surface Modification by Polymer Grafting

Polymer grafting provides a further, versatile approach to the modification of surfaces. Grafted surfaces can be produced by direct coupling of existing polymer molecules to surfaces containing functional groups or graft polymerization of monomers. Graft polymerization methods include chemical graft polymerization and grafting using high-energy radiation—both ionizing and non-ionizing—or oxidizing agents such as ozone or acid treatments. Detailed reviews on polymer grafting and the physicochemical characteristics of graft layers have been written by Uyama,^[94] Kato,^[12]

Kang,^[95] Claesson,^[96] and Bamford.^[97] Grafting techniques for polymeric biomaterials have been used to generate surfaces that exhibit nonfouling properties, selective protein adsorption, and improved biocompatibility. In the case of nonfouling surfaces, a number of water soluble polymers have been used including polyacrylamide, poly(ethylene glycol), polysaccharides, and phosphorylcholine. Grafting has been achieved by covalent coupling, surface graft polymerization, surface segregation, and interpenetration of the surface. A number of reviews specifically focused on surface modification for nonfouling have been written recently by Kingshott,^[25] Malmsten,^[98] and Hoffman.^[24]

Factors known to affect the properties of a grafted polymer layer include polymer molecular weight, the extent of chain branching, the density of functional and charged groups on both the substrate and polymer, and the solubility of the polymer.

Self-Assembled Monolayers

Self-assembled monolayers (SAMs) also provide a flexible approach to the creation of surface chemical and biological functionality. SAMs are highly ordered, surface coatings that form spontaneously on certain substrates, e.g., alkanethiols ($\text{HS}(\text{CH}_2)_n\text{X}$) on gold or silver.^[34] SAMs can also be formed from the interactions of chlorosilanes with OH-terminated oxide surfaces^[99] and a number of acids on metal oxide surfaces.^[33] SAMs provide well-defined surface structures, can be patterned using microcontact printing and photolithography,^[100] and have chemistries amenable to a range of modifications.^[23,101,102] SAMs have been widely used as model surfaces in the study of protein adsorption and cell adhesion^[23,103] and the investigation of cell adhesion peptide sequences,^[101,104] but they have also found commercial applicability in biosensors and nucleotide arrays.^[33] There has been a considerable research effort in recent years investigating the formation and structure of SAMs that are capable of resisting protein adsorption by the incorporation of derivatives of poly(ethylene glycol) and other specific end groups. The results of which have been reviewed by Ostuni.^[23]

Biomolecule Attachment and Biomimetic Surfaces

Coatings incorporating biologically active and/or inactive molecules to generate specific, predictable, and controlled responses in the biological environment have been of considerable interest in recent biomaterials research. These so-called intelligent materials are designed to provide specific signals that result in

biologically functional surfaces.^[36,102,105,106] A wide range of biomolecules, including proteins, peptides, polysaccharides, lipids, and oligonucleotides, have been immobilized on surfaces with the aim of eliciting specific, predictable, and controlled biological responses. Of particular interest in the biomaterials field has been the incorporation of peptide sequences (e.g., RGD) from the receptor-binding domains of adhesion proteins in order to promote cell adhesion.^[105,107] The success of strategies designed to present biological ligands is maximized if the immobilized molecule is coupled to a surface capable of preventing nonspecific adsorption. The subsequent biological response is influenced by the presentation, average density, and the spatial distribution of the immobilized molecule.^[105]

Biomolecule immobilization methodologies encompass both physical and chemical techniques and range in their degree of specificity and resulting functionality. In their simplest incarnation, nonspecific adsorption results in a surface with randomly oriented molecules deposited in a noncontrolled manner. While these coatings have displayed some short-term functionality, their efficacy is diminished by denaturation and biomolecule desorption. Covalent immobilization has also been used to couple biomolecules to suitable surface functionalities using a range of chemistries.^[108] Chemical modification of the surface to introduce functional groups can be achieved by using a variety of the techniques outlined in previous sections, including plasma polymerization, and photochemical grafting. But, nonspecific covalent immobilization gives little control over orientation or activity of the biomolecule. The introduction of polymer or amino acid spacer chains with reactive functional groups enables the biomolecule to be attached in a manner that can provide greater access to specific sites on the biomolecule, which reduces the incidence of denaturation and thus maintains biomolecule activity. Specific examples of this type of approach are the site-specific modification of proteins with cysteine or biotin tags that enables immobilization of the proteins in specific orientations on gold- and streptavidin-modified surfaces.^[109]

CONCLUSION

An extensive range of surface modification techniques can be used to affect changes in surface properties of biomaterials while maintaining the bulk properties of the material. Several of these modification techniques have found application in commercial devices and are likely to continue to be used for some considerable time. However, the focus of surface modification is increasingly directed toward the design of bioengineered surfaces that are able to elicit and control biological responses.

Recent years have seen an increased focus on the study and design of surfaces incorporating specific recognition molecules that are precisely oriented and at an appropriate density for optimal biological function. In addition, there is increasing interest in surfaces that borrow from structures observed in nature, so-called biomimetic surfaces. Examples of these include protein template imprinted structures and tethered lipid bilayers. At the same time, well-defined model systems continue to give insight into the parameters that control biological responses to implanted materials. Surfaces based on the principles of self-assembly provide valuable insight into interfacial interactions and are likely to find increasing application for in-vitro and diagnostics applications. In the same vein, nonfouling surfaces continue to be critical to the development of a range of in-vitro biomedical applications, including diagnostics and microfluidic devices.

ARTICLES OF FURTHER INTEREST

Biomaterials: Protein–Surface Interactions; Gradient Surfaces: Preparation, Characterization, and Interactions with Biological Species; Phosphorylcholine (PC) Technology; Photopolymerization; Plasma Modification of Materials; Polyelectrolyte Multilayers; Protein Adsorption Modeling; Self-Assembled Monolayers; Surface Coatings; Thrombosis

REFERENCES

1. Chu, P.K.; Chen, J.Y.; Wang, L.P.; Huang, N. Plasma-surface modification of biomaterials. *Mater. Sci. Eng., R* **2002**, *36* (5–6), 143–206.
2. Griesser, H.J.; Chatelier, R.C.; Gengenbach, T.R.; Vasic, Z.R.; Johnson, G.; Steele, J.G. Plasma surface modifications for improved biocompatibility of commercial polymers. *Polym. Int.* **1992**, *27*, 109–117.
3. Favia, P.; d'Agostino, R. Plasma treatments and plasma depositions of polymers for biomedical applications. *Surf. Coat. Technol.* **1998**, *98*, 1102–1106.
4. Mandl, S.; Sader, R.; Thorwarth, G.; Krause, D.; Zeilhofer, H.F.; Horch, H.H.; Rauschenbach, B. Investigation on plasma immersion ion implantation treated medical implants. *Biomol. Eng.* **2002**, *19* (2–6), 129–132.
5. Tan, L.; Crone, W.C. Surface characterization of NiTi modified by plasma source ion implantation. *Acta Mater.* **2002**, *50*, 4449–4460.
6. Howlett, C.R.; Zreiqat, H.; Wu, Y.; McFall, D.W.; McKenzie, D.R. Effect of ion modification of commonly used orthopedic materials on the attachment of human bone-derived cells. *J. Biomed. Mater. Res.* **1999**, *45* (4), 345–354.

7. Ratner, B.D. Plasma deposition for biomedical applications: A brief review. *J. Biomater. Sci., Polym. Ed.* **1992**, *4* (1), 3–11.
8. Chan, C.M.; Ko, T.M.; Hiraoka, H. Polymer surface modification by plasmas and photons. *Surf. Sci. Rep.* **1996**, *24* (1–2), 3–54.
9. Polzonetti, G.; Iucci, G.; Frontini, A.; Infante, G.; Furlani, C.; Avigliano, L.; Del Principe, D.; Palumbo, G.; Rosato, N. Surface reactions of a plasma-sprayed CaO–P₂O₅–SiO₂-based glass with albumin, fibroblasts and granulocytes studied by XPS, fluorescence and chemiluminescence. *Biomaterials* **2000**, *21* (15), 1531–1539.
10. Seto, F.; Muraoka, Y.; Sakamoto, N.; Kishidu, A.; Akashi, M. Surface modification of synthetic fiber nonwoven fabrics with poly(acrylic acid) chains prepared by corona discharge induced grafting. *Angew. Makromol. Chem.* **1999**, *266*, 56–62.
11. Zhang, F.; Kang, E.T.; Neoh, K.G.; Wang, P.; Tan, K.L. Surface modification of stainless steel by grafting of poly(ethylene glycol) for reduction in protein adsorption. *Biomaterials* **2001**, *22* (12), 1541–1548.
12. Kato, K.; Uchida, E.; Kang, E.T.; Uyama, Y.; Ikada, Y. Polymer surface with graft chains. *Prog. Polym. Sci.* **2003**, *28* (2), 209–259.
13. Porte-Durrieu, M.C.; Aymes-Chodur, C.; Vergne, C.; Betz, N.; Baquey, C. Surface treatment of biomaterials by gamma and swift heavy ions grafting. *Nucl. Instrum. Methods, B* **1999**, *151* (1–4), 404–415.
14. Welle, A.; Gottwald, E. UV-based patterning of polymeric substrates for cell culture applications. *Biomed. Microdevices* **2002**, *4* (1), 33–41.
15. Khorasani, M.T.; Mirzadeh, H.; Sammes, P.G. Laser surface modification of polymers to improve biocompatibility: HEMA grafted PDMS, in vitro assay—III. *Radiat. Phys. Chem.* **1999**, *55* (5–6), 685–689.
16. Hunt, J.A.; Williams, R.L.; Tavakoli, S.M.; Riches, S.T. Laser surface modification of polymers to improve biocompatibility. *J. Mater. Sci., Mater. Med.* **1995**, *6* (12), 813–817.
17. Dadsetan, M.; Mirzadeh, H.; Sharifi-Sanjani, N.; Daliri, M. Cell behavior on laser surface-modified polyethylene terephthalate in vitro. *J. Biomed. Mater. Res.* **2001**, *57* (2), 183–189.
18. Benson, R.S. Use of radiation in biomaterials science. *Nucl. Instrum. Methods, B* **2002**, *191*, 752–757.
19. Zhao, Q.; Zhai, G.J.; Ng, D.H.L.; Zhang, X.Z.; Chen, Z.Q. Surface modification of Al₂O₃ bioceramic by NH₂⁺ ion implantation. *Biomaterials* **1999**, *20* (6), 595–599.
20. Cui, F.Z.; Luo, Z.S. Biomaterials modification by ion-beam processing. *Surf. Coat. Technol.* **1999**, *112* (1–3), 278–285.
21. Davenas, J.; Thevenard, P.; Philippe, F.; Arnaud, M.N. Surface implantation treatments to prevent infection complications in short term devices. *Biomol. Eng.* **2002**, *19* (2–6), 263–268.
22. Bambauer, R.; Mestres, P.; Schiel, R.; Schneidewind, J.M.; Latza, R.; Bambauer, S.; Sioshansi, P. Surface treated catheters with ion beam based process for blood access. *Ther. Apher.* **2000**, *4* (5), 342–347.
23. Ostuni, E.; Chapman, R.G.; Holmlin, R.E.; Takayama, S.; Whitesides, G.M. A survey of structure–property relationships of surfaces that resist the adsorption of protein. *Langmuir* **2001**, *17*, 5605–5620.
24. Hoffman, A.S. Non-fouling surface technologies. *J. Biomater. Sci., Polym. Ed.* **1999**, *10* (10), 1011–1014.
25. Kingshott, P.; Griesser, H.J. Surfaces that resist bioadhesion. *Curr. Opin. Solid State Mater. Sci.* **1999**, *4*, 403–412.
26. Holland, N.B.; Qiu, Y.X.; Ruegsegger, M.; Marchant, R.E. Biomimetic engineering of non-adhesive glycocalyx-like surfaces using oligosaccharide surfactant polymers. *Nature* **1998**, *392* (6678), 799–801.
27. Massia, S.P.; Stark, J. Immobilized RGD peptides on surface-grafted dextran promote biospecific cell attachment. *J. Biomed. Mater. Res.* **2001**, *56* (3), 390–399.
28. Puleo, D.A.; Kissling, R.A.; Sheu, M.S. A technique to immobilize bioactive proteins, including bone morphogenetic protein-4 (BMP-4), on titanium alloy. *Biomaterials* **2002**, *23* (9), 2079–2087.
29. Yoshinari, M.; Matsuzaka, K.; Inoue, T.; Oda, Y.; Shimon, M. Bio-functionalization of titanium surfaces for dental implants. *Mater. Trans.* **2002**, *43* (10), 2494–2501.
30. Groves, J.T.; Mahal, L.K.; Bertozzi, C.R. Control of cell adhesion and growth with micropatterned supported lipid membranes. *Langmuir* **2001**, *17* (17), 5129–5133.
31. Bures, P.; Huang, Y.B.; Oral, E.; Peppas, N.A. Surface modifications and molecular imprinting of polymers in medical and pharmaceutical applications. *J. Control. Release* **2001**, *72* (1–3), 25–33.
32. Shi, H.; Tsai, W.B.; Garrison, M.D.; Ferrari, S.; Ratner, B.D. Template-imprinted nanostructured surfaces for protein recognition. *Nature* **1999**, *398* (6728), 593–597.
33. Textor, M.; Ruiz, L.; Hofer, R.; Rossi, A.; Feldman, K.; Hahner, G.; Spencer, N.D. Structural chemistry of self-assembled monolayers of octadecylphosphoric acid on tantalum oxide surfaces. *Langmuir* **2000**, *16* (7), 3257–3271.
34. Harder, P.; Grunze, M.; Dahint, R.; Whitesides, G.M.; Laibinis, P.E. Molecular conformation in oligo(ethylene glycol)-terminated self-assembled monolayers on gold and silver surfaces determines their ability to resist protein adsorption. *J. Phys. Chem., B* **1998**, *102* (426–436).
35. Hubbard, J.B.; Silin, V.; Plant, A.L. Self assembly driven by hydrophobic interactions at alkane-thiol monolayers: Mechanism of formation of hybrid bilayer membranes. *Biophys. Chem.* **1998**, *75* (3), 163–176.
36. Castner, D.G.; Ratner, B.D. Biomedical surface science: Foundations to frontiers. *Surf. Sci.* **2002**, *500* (1–3), 28–60.
37. Sioshansi, P.; Tobin, E.J. Surface treatment of biomaterials by ion beam processes. *Surf. Coat. Technol.* **1996**, *83* (1–3), 175–182.
38. Hanawa, T. In vivo metallic biomaterials and surface modification. *Mater. Sci. Eng., A: Struct.* **1999**, *267* (2), 260–266.

39. Maruyama, M.; Capello, W.N.; D'Antonio, J.A.; Jaffe, W.L.; Bierbaum, B.E. Effect of low-friction ion-treated femoral heads on polyethylene wear rates. *Clin. Orthop. Relat. Res.* **2000**, *1* (370), 183–191.
40. Frank, J.L.; Garb, J.L.; Halla, B.; Reed, W.P., Jr. Ionic implantation of silicone chronic venous access devices does not alter thrombotic complications: A double-blinded, randomized clinical trial. *Surgery* **2001**, *129* (5), 547–551.
41. Pignataro, B.; Conte, E.; Scandurra, A.; Marletta, G. Improved cell adhesion to ion beam-irradiated polymer surfaces. *Biomaterials* **1997**, *18* (22), 1461–1470.
42. Guzman, L.; Celva, R.; Miotello, A.; Voltolini, E.; Ferrari, F.; Adami, M. Polymer surface modification by ion implantation and reactive deposition of transparent films. *Surf. Coat. Technol.* **1998**, *104*, 375–379.
43. Krupa, D.; Baszkiewicz, J.; Kozubowski, J.; Barcz, A.; Sobczak, J.; Bilinski, A.; Rajchel, B. The influence of calcium and/or phosphorus ion implantation on the structure and corrosion resistance of titanium. *Vacuum* **2001**, *63* (4), 715–719.
44. Sato, H.; Tsuji, H.; Ikeda, S.; Ikemoto, N.; Ishikawa, J.; Nishimoto, S. Enhanced growth of human vascular endothelial cells on negative ion (Ag^-)-implanted hydrophobic surfaces. *J. Biomed. Mater. Res.* **1999**, *44* (1), 22–30.
45. Cui, F.Z.; Luo, Q.L.; Feng, J. Highly adhesive hydroxyapatite coatings on titanium alloy formed by ion beam assisted deposition. *J. Mater. Sci., Mater. Med.* **1997**, *8*, 403–405.
46. Kurtz, S.M.; Muratoglu, O.K.; Evans, M.; Edidin, A.A. Advances in the processing, sterilization, and cross-linking of ultra-high molecular weight polyethylene for total joint arthroplasty. *Biomaterials* **1999**, *20* (18), 1659–1688.
47. Wetzels, G.M.; Koole, L.H. Photoimmobilisation of poly(n-vinylpyrrolidinone) as a means to improve haemocompatibility of polyurethane biomaterials. *Biomaterials* **1999**, *20* (20), 1879–1887.
48. McPherson, T.B.; Shim, H.S.; Park, K. Grafting of PEO to glass, nitinol, and pyrolytic carbon surfaces by gamma irradiation. *J. Biomed. Mater. Res.* **1997**, *38* (4), 289–302.
49. Kidane, A.; McPherson, T.; Shim, H.S.; Park, K. Surface modification of polyethylene terephthalate using PEO–polybutadiene–PEO triblock copolymers. *Colloid Surf., B* **2000**, *18* (3–4), 347–353.
50. Curti, P.S.; De Moura, M.R.; Radovanovic, E.; Rubira, A.F.; Muniz, E.C. Surface modification of polystyrene and poly(ethylene terephthalate) by grafting poly(n-isopropylacrylamide). *J. Mater. Sci., Mater. Med.* **2002**, *13* (12), 1175–1180.
51. Ward, J.H.; Bashir, R.; Peppas, N.A. Micropatterning of biomedical polymer surfaces by novel UV polymerization techniques. *J. Biomed. Mater. Res.* **2001**, *56* (3), 351–360.
52. Thissen, H.; Hayes, J.P.; Kingshott, P.; Johnson, G.; Harvey, E.C.; Griesser, H.J. Nanometer thickness laser ablation for spatial control of cell attachment. *Smart Mater. Struct.* **2002**, *11* (5), 792–799.
53. Sun, S.; Chong, K.S.; Leggett, G.J. Nanoscale molecular patterns fabricated by using scanning near-field optical lithography. *J. Am. Chem. Soc.* **2002**, *124* (11), 2414–2415.
54. Duncan, A.C.; Weisbuch, F.; Rouais, F.; Lazare, S.; Baquey, C. Laser microfabricated model surfaces for controlled cell growth. *Biosens. Bioelectron.* **2002**, *17* (5), 413–426.
55. Leng, Y.X.; Chen, J.Y.; Zeng, Z.M.; Tian, X.B.; Yang, P.; Huang, N.; Zhou, Z.R.; Chu, P.K. Properties of titanium oxide biomaterials synthesized by titanium plasma immersion ion implantation and reactive ion oxidation. *Thin Solid Films* **2000**, *377*, 573–577.
56. Lee, Y.; Han, S.; Lim, H.; Kim, Y.; Kim, H. Surface analysis of polymers electrically improved by plasma-source ion-implantation. *Anal. Bioanal. Chem.* **2002**, *373* (7), 595–600.
57. Shin, G.H.; Lee, Y.H.; Lee, J.S.; Kim, Y.S.; Choi, W.S.; Park, H.J. Preparation of plastic and biopolymer multilayer films by plasma source ion implantation. *J. Agric. Food. Chem.* **2002**, *50* (16), 4608–4614.
58. Bilek, M.M.M.; McKenzie, D.R.; Tarrant, R.N.; Lim, S.H.M.; McCulloch, D.G. Plasma-based ion implantation utilizing a cathodic arc plasma. *Surf. Coat. Technol.* **2002**, *156* (1–3), 136–142.
59. Ratner, B.D.; Chilkoti, A.; Lopez, G.P. Plasma Deposition and Treatment for Biomedical Applications. In *Plasma Deposition, Treatment, and Etching of Polymers*; d'Agostino, R., Ed.; Academic Press Inc.: San Diego, 1990; Vol. 1, 463–516.
60. Latkany, R.; Tsuk, A.; Sheu, M.-S.; Loh, I.-H.; Trinkaus-Randall, V. Plasma surface modification of artificial corneas for optimal epithelialisation. *J. Biomed. Mater. Res.* **1996**, *36*, 29–37.
61. Dai, L.; Zientek, P.; St. John, H.A.W.; Pasic, P.; Chatelier, R.C.; Griesser, H.J. Surface Modification of Contact Lenses: Making Contact Lenses Wettable. In *Surface Modification of Polymeric Biomaterials*; Ratner, B., Castner, D., Eds.; Plenum Press: New York, 1996; 147.
62. Weikart, C.M.; Matsuzawa, Y.; Winterton, L.; Yasuda, H.K. Evaluation of plasma polymer-coated contact lenses by electrochemical impedance spectroscopy. *J. Biomed. Mater. Res.* **2001**, *54* (4), 597–607.
63. Lumsden, A.B.; Chen, C.; Coyle, K.A.; Ofenloch, J.C.; Wang, J.H.; Yasuda, H.K.; Hanson, S.R. Nonporous silicone polymer coating of expanded polytetrafluoroethylene grafts reduces graft neointimal hyperplasia in dog and baboon models. *J. Vasc. Surg.* **1996**, *24* (5), 825–833.
64. Yasuda, H.; Gazicki, M. Biomedical applications of plasma polymerization and plasma treatment of polymer surfaces. *Biomaterials* **1982**, *3* (2), 68–77.
65. Hahn, A.W.; Yasuda, H.K.; James, W.J.; Nichols, M.F.; Sathir, R.K.; Sharma, A.K.; Pringle, O.A.; York, D.H.; Charlson, E.J. Glow discharge polymers as coatings for implanted devices. *Biomed. Sci. Instrum.* **1981**, *17*, 109–113.
66. Nilsson-Ekdahl, K.; Nilsson, B.; Golander, C.G.; Elwing, H.; Lassen, B.; Nilsson, U.R. Complement activation on radio frequency plasma modified polystyrene surfaces. *J. Colloid Interface Sci.* **1993**, *158*, 121–128.

67. Griesser, H.J.; Chatelier, R.C.; Gengenbach, T.R.; Johnson, G.; Steele, J.G. Growth of human cells on plasma polymers: Putative role of amine and amide groups. *J. Biomater. Sci., Polym. Ed.* **1994**, *5*, 531–554.
68. Haddow, D.B.; Steele, D.A.; Short, R.D.; Dawson, R.A.; Macneil, S. Plasma-polymerized surfaces for culture of human keratinocytes and transfer of cells to an in vitro wound-bed model. *J. Biomed. Mater. Res.* **2003**, *64* (1), 80–87.
69. Mar, M.N.; Ratner, B.D.; Yee, S.S. An intrinsically protein-resistant surface plasmon resonance biosensor based upon a RF-plasma-deposited thin film. *Sens. Actuators, B: Chem.* **1999**, *54* (1–2), 125–131.
70. Nakamura, R.; Muguruma, H.; Ikebukuro, K.; Sasaki, S.; Nagata, R.; Karube, I.; Pedersen, H. A plasma-polymerized film for surface plasmon resonance immunosensing. *Anal. Chem.* **1997**, *69*, 4649–4652.
71. Kiaei, D.; Hoffmann, A.; Horbett, T. Radio-frequency gas discharge (RFGD) fluorination of polymers: Protein and cell interactions at RFGD-fluorinated interfaces. *Radiat. Phys. Chem.* **1995**, *46* (2), 191–197.
72. Lopez, G.P.; Ratner, B.D.; Tidwell, C.D.; Haycox, C.L.; Rapoza, R.J.; Horbett, T.A. Glow discharge plasma deposition of tetraethylene glycol dimethyl ether for fouling-resistant biomaterial surfaces. *J. Biomed. Mater. Res.* **1992**, *26*, 415–439.
73. Griesser, H.J. Gas plasma techniques for the preparation of hydrophilic polymer surfaces. *Mater. Forum* **1990**, *14*, 192–202.
74. Lassen, B.; Malmsten, M. Competitive protein adsorption at radio frequency plasma polymer surfaces. *J. Mater. Sci., Mater. Med.* **1994**, *5*, 662–665.
75. Tang, L.; Wu, Y.; Timmons, R.B. Fibrinogen adsorption and host tissue responses to plasma functionalized surfaces. *J. Biomed. Mater. Res.* **1998**, *42*, 156–163.
76. Shen, M.C.; Martinson, L.; Wagner, M.S.; Castner, D.G.; Ratner, B.D.; Horbett, T.A. PEO-like plasma polymerized tetraglyme surface interactions with leukocytes and proteins: In vitro and in vivo studies. *J. Biomater. Sci., Polym. Ed.* **2002**, *13* (4), 367–390.
77. Bae, J.; Seo, E.; Kang, I. Synthesis and characterization of heparinized polyurethanes using plasma glow discharge. *Biomaterials* **1999**, *20*, 529–537.
78. Griesser, H.J.; McLean, K.M.; Beumer, G.J.; Gong, X.; Kingshott, P.; Johnson, G.; Steele, J.G. Surface immobilization of synthetic proteins via plasma polymer interlayers. *Mater. Res. Symp. Proc.* **1999**, *544*, 9.
79. Hendricks, S.K.; Kwok, C.; Shen, M.; Horbett, T.A.; Ratner, B.D.; Bryers, J.D. Plasma-deposited membranes for controlled release of antibiotic to prevent bacterial adhesion and biofilm formation. *J. Biomed. Mater. Res.* **2000**, *50* (2), 160–170.
80. Narayanan, P. Surface functionalization by RF plasma treatment of polymers for immobilization of bioactive-molecules. *J. Biomater. Sci., Polym. Ed.* **1994**, *6* (2), 181–193.
81. Sodhi, R.N.S.; Sahi, V.P.; Mittelman, M.W. Application of electron spectroscopy and surface modification techniques in the development of anti-microbial coatings for medical devices. *J. Electron Spectrosc.* **2001**, *121* (1–3), 249–264.
82. Lens, J.P.; Terlingen, J.G.A.; Engbers, G.H.M.; Feijen, J. Preparation of heparin-like surfaces by introducing sulfate and carboxylate groups on poly(ethylene) using an argon plasma treatment. *J. Biomater. Sci., Polym. Ed.* **1998**, *9* (4), 357–372.
83. Goessl, A.; Garrison, M.D.; Lhoest, J.B.; Hoffman, A.S. Plasma lithography-thin film patterning of polymeric biomaterials by RF plasma polymerization I: Surface preparation and analysis. *J. Biomater. Sci., Polym. Ed.* **2001**, *12* (7), 721–738.
84. Ohl, A.; Schroder, K. Plasma-induced chemical micro-patterning for cell culturing applications: A brief review. *Surf. Coat. Technol.* **1999**, *119*, 820–830.
85. Aronsson, B.O.; Lausmaa, J.; Kasemo, B. Glow discharge plasma treatment for surface cleaning and modification of metallic biomaterials. *J. Biomed. Mater. Res.* **1997**, *35* (1), 49–73.
86. Terlingen, J.G.; Feijen, J.; Hoffman, A.S. Immobilization of surface active compounds on polymer supports using a gas discharge process. *J. Biomater. Sci., Polym. Ed.* **1992**, *4* (1), 31–33.
87. Pu, F.R.; Williams, R.L.; Markkula, T.K.; Hunt, J.A. Effects of plasma treated PET and PTFE on expression of adhesion molecules by human endothelial cells in vitro. *Biomaterials* **2002**, *23* (11), 2411–2428.
88. Beyer, D.; Knoll, W.; Ringsdorf, H.; Wang, J.-H.; Timmons, R.; Sluka, P. Reduced protein adsorption on plastics via direct plasma deposition of triethylene glycol monoallyl ether. *J. Biomed. Mater. Res.* **1996**, *36*, 181–189.
89. Pan, Y.V.; Wesley, R.A.; Luginbuhl, R.; Denton, D.D.; Ratner, B.D. Plasma polymerized n-isopropylacrylamide: Synthesis and characterization of a smart thermally responsive coating. *Biomacromolecules* **2001**, *2* (1), 32–36.
90. Tarducci, C.; Schofield, W.C.E.; Badyal, J.P.S. Monomolecular functionalization of pulsed plasma deposited poly(2-hydroxyethyl methacrylate) surfaces. *Chem. Mater.* **2002**, *14* (6), 2541–2545.
91. Kim, D.D.; Takeno, M.M.; Ratner, B.D.; Horbett, T.A. Glow discharge plasma deposition (GDPD) technique for the local controlled delivery of hirudin from biomaterials. *Pharm. Res.* **1998**, *15* (5), 783–786.
92. Calderon, J.G.; Harsch, A.; Gross, G.W.; Timmons, R.B. Stability of plasma-polymerized allylamine films with sterilization by autoclaving. *J. Biomed. Mater. Res.* **1998**, *42*, 597–603.
93. Beck, A.J.; Jones, F.R.; Short, R.D. Plasma copolymerization as a specific route to the fabrication of new surfaces with controlled amounts of specific chemical functionality. *Polymer* **1996**, *37*, 5537–5539.
94. Uyama, Y.; Kato, K.; Ikada, Y. Surface Modification of Polymers by Grafting. In *Advances in Polymer Science*; Springer-Verlag: Berlin, 1998; Vol. 137, 3–39.
95. Kang, E.T.; Zhang, Y. Surface modification of fluoropolymers via molecular design. *Adv. Mater.* **2000**, *12*, 1481–1494.
96. Claesson, P.M. Poly(ethylene oxide) surface coatings: Relations between intermolecular forces, layer structure



- and protein repellency. *Colloid Surf., A* **1993**, *77*, 109–118.
97. Bamford, C.H.; Al-Lamee, K.G. Studies in polymer surface modification and grafting for biomedical uses. 2. Application to arterial blood filters and oxygenators. *Polymer* **1996**, *37*, 4885–4889.
98. Malmsten, M.; Emoto, K.; Van Alstine, J. Effect of chain density on inhibition of protein adsorption by poly(ethylene glycol) based coatings. *J. Colloid Interface Sci.* **1998**, *202*, 507–517.
99. Filippini, P.; Rainaldi, C.; Ferrante, A.; Mecheri, B.; Gabrielli, G.; Bombace, M.; Indovina, P.L.; Santini, M.T. Modulation of osteosarcoma cell growth and differentiation by silane-modified surfaces. *J. Biomed. Mater. Res.* **2001**, *55* (3), 338–349.
100. Whitesides, G.M.; Ostuni, E.; Takayama, S.; Jiang, X.; Ingber, D.E. Soft lithography in biology and biochemistry. *Annu. Rev. Biomed. Eng.* **2001**, *3*, 335–373.
101. Houseman, B.T.; Mrksich, M. The role of ligand density in the enzymatic glycosylation of carbohydrates presented in self-assembled monolayers of alkanethiolates on gold. *Angew. Chem. Int. Ed.* **1999**, *38* (6), 782–785.
102. Mrksich, M. What can surface chemistry do for biology? *Curr. Opin. Biotechnol.* **2002**, *6*, 794–797.
103. Jung, D.R.; Kapur, R.; Adams, T.; Giuliano, K.A.; Mrksich, M.; Craighead, H.G.; Taylor, D.L. Topographical and physicochemical modification of material surface to enable patterning of living cells. *Crit. Rev. Biotechnol.* **2001**, *21* (2), 111–154.
104. Roberts, C.; Chen, C.S.; Mrksich, M.; Martichonok, V.; Ingber, D.E.; Whitesides, G.M. Using mixed self-assembled monolayers presenting RGD and (EG)₃OH groups to characterize long-term attachment of bovine capillary endothelial cells to surfaces. *J. Am. Chem. Soc.* **1998**, *120*, 6548–6555.
105. Hubbell, J.A. Bioactive biomaterials. *Curr. Opin. Biotechnol.* **1999**, *10*, 123–129.
106. Tirrell, M.; Kokkoli, E.; Biesalski, M. The role of surface science in bioengineered materials. *Surf. Sci.* **2002**, *500*, 61–83.
107. Sakiyama-Elbert, S.E.; Hubbell, J.A. Functional biomaterials: Design of novel biomaterials. *Annu. Rev. Mater. Res.* **2001**, *31*, 183–201.
108. Hermanson, G.T. *Bioconjugate Techniques*; Academic Press: London, 1996; Vol. 1.
109. Peluso, P.; Wilson, D.S.; Do, D.; Tran, H.; Venkatasubbaiah, M.; Quincy, D.; Heidecker, B.; Poindexter, K.; Tolani, N.; Phelan, M.; Witte, K.; Jung, L.S.; Wagner, P.; Nock, S. Optimizing antibody immobilization strategies for the construction of protein microarrays. *Anal. Biochem.* **2003**, *312*, 113–124.

Surface Topography and Cell Behavior

Douglas W. Hamilton
Salem Ghrebi
Hugh Kim
Babak Chehroudi
Donald M. Brunette

*Department of Oral, Biological, and Medical Sciences, University of British Columbia,
Vancouver, British Columbia, Canada*

S

INTRODUCTION

It has been long recognized that mammalian cells are sensitive to topographical features of their underlying substratum, whether they be other cells, extracellular matrix proteins, or implant surfaces. With the increased use of artificial materials in medicine, understanding which stimuli promote desirable reactions from cells in surrounding tissues has become critical to improve implant performance. When considering suitable materials for implantation, two key aspects have been identified that influence cellular response, namely, the chemistry and the topography of the material. Initial research models used to investigate topography were criticized owing to the difficulty in the separation of the effects of material chemistry from topography. However, by the harnessing of techniques used in the microelectronic industry, the fabrication of substrata with precisely defined topographies and uniform chemistry has been developed. Through the use of such substrata, the effects of topography on cell behavior have been identified on well-characterized materials. The aim of this entry is to review the types of surface topographies currently investigated and the types of reactions such substrata evoke from cells and/or tissues.

SIGNIFICANCE OF TOPOGRAPHICAL REACTION

The importance of topographical reactions and gaining an understanding of the mechanisms lying behind it has been obvious in development and wound repair. Directed cell migration is essential during development, where the movement of cells from one part of the embryo to another must be under precise control.^[1] A number of cues including cell–cell interactions and topography have been identified as playing key roles in this patterned migration. In addition, topographical control of cell migration is also thought to play a part

in pathological processes such as cancer metastasis, where local invasion of tumor cells often occurs along aligned sheets of collagen.^[2] In such in vivo situations, however, it is uncertain whether topography alone is responsible for the migration of the cells. The possibility that extracellular matrix chemistry or chemotactic gradients could be influential cannot be excluded.

FIRST OBSERVATIONS OF TOPOGRAPHICAL REACTIONS

Ross Harrison first observed the direct reaction of cells to topography.^[3] He cultured cells on spider webs, and observed that the cells became oriented and their movement guided by the fine silk threads. The control of cell behavior by fibers was confirmed by Weiss, who cultured cells on networks of fibrin in plasma clots.^[4] He observed that when two explants of chick embryonic tissue were placed a small distance apart in a plasma clot culture, growth from each explant was directed along the fibers. Weiss termed this phenomenon of cell orientation by the underlying topography “contact guidance.”

TYPES OF SURFACE TOPOGRAPHY

Much of the early work carried out on topography employed natural substrata, structures that were derived from biological materials. The reasoning behind this was that such substrata could be of biological importance in understanding guidance cues during development.^[1] Unfortunately, there are disadvantages in using biological materials. First, the chemistry of the material could affect the behavior of the cells, and consequently, any results observed may not be due to topography alone.^[1,5] Second, the chemistry of such substrata could be variable. Despite the disadvantages, certain biological substrata have been shown to align

cells in vitro, proteins such as collagen (when aligned), and fibronectin fibrils. Neuronal cell processes, neural crest cells, and leukocytes have all been shown to react to these substrata, cell shape being oriented and locomotion guided. However, to fully investigate the effects of topography, substrata with well-defined features and, more importantly, uniform chemistry needed to be designed.

Microfabrication and Defined Surface Features

For many years, little work was carried out using artificial topographies in vitro to investigate cellular reaction to surface morphology, mainly because of lack of suitable materials that had well-defined and uniform surface morphology. Methods of fabrication employed lacked flexibility in dimensions and the surfaces were complicated owing to heterogeneities introduced during fabrication. With the advent of the use of techniques harnessed from the microelectronic industry, it has become possible to produce structures with precisely defined surfaces in cell biological studies, thus allowing the reaction of cells to defined topographic features to be observed.^[6]

Synthetic materials such as silicon and glass, cellulose acetate, epoxy, araldite, polystyrene and metals such as titanium and titanium alloys have all been employed to produce precisely characterized topographies.^[5] Manufactured substrata have the advantage that the material can be designed to precise dimensions and can be fully characterized in terms of

surface chemistry. Using such methods of fabrication, fibers of glass can be drawn out to precise dimensions, and structures such as ridges and grooves can be made using photolithography and etching techniques in silicon. The process of photolithography is shown diagrammatically in Fig. 1.

Topography Vs. Chemistry

One argument that needs to be addressed at the outset is whether the methods of fabrication of topography result in the production of altered chemistry across the surfaces, raising the possibility that the chemistry influences cell reaction. This area has been the subject of much research and has been addressed in great detail elsewhere. One approach to alleviate this problem is to coat the material with titanium, making it unlikely that differences in chemistry would be present.^[6] Another approach would be to blanket etch the substratum after topographic patterning by etching. For a more in-depth review of alternate approaches, see Curtis and Wilkinson.^[5]

Single Steps (Cliffs)

Single steps or cliffs are one of the simplest structures to employ and have both an acute and obtuse edge. Clark et al. found that a variety of cell types reacted to this type of structure.^[7] They noted that cell response was influenced by the direction from which the cells

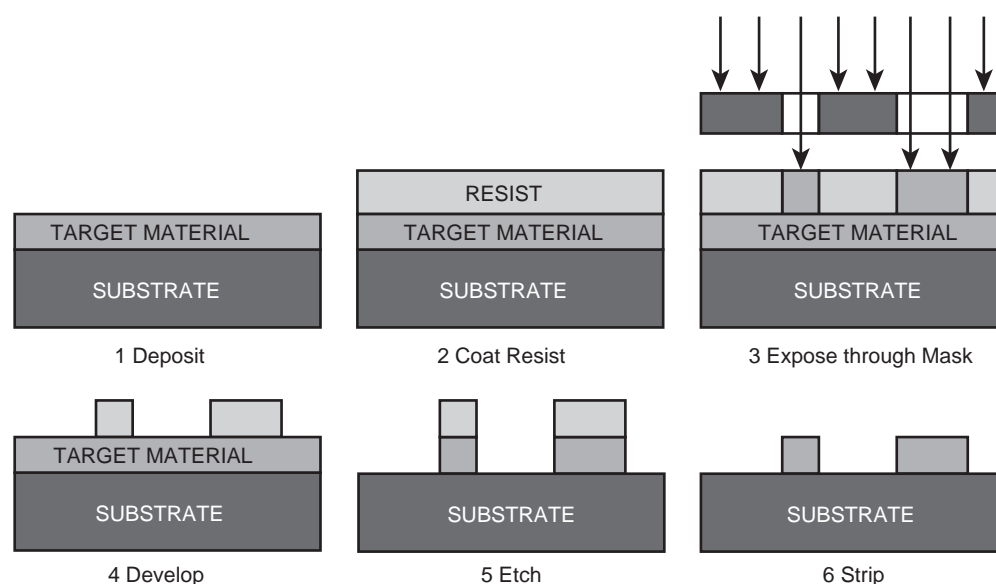


Fig. 1 Schematic representation of the photolithography process. The target surface (1) is cleaned and resist applied to the target material (2). A photomask is then created and developed onto the substrate (3), through the mask (4). The substrate is then etched (5), and the resist is removed using acetone (6), leaving the desired topography etched onto the substrate.

approached the cue. For certain heights of cliffs, the cells could climb up the step, but could not descend from the top. Ridges have also been employed that are essentially double cliffs. Depending on spacing, cells responded by accumulating on the ridges and extending along them. In some instances, the cells were guided in their migration by the ridges, a phenomenon termed ridge walking.^[1]

Fibers

The use of fibers in the field of regenerative medicine first appeared in a study by Jenkins and colleagues in 1977, with a view to applications in ligament repair and regeneration.^[8] Since this time, fibers have been produced using materials such as titanium, poly-glycolic acid, poly-lactic acid, polypropylene, glass, chitosan, ceramics, as well as natural materials such as fibronectin and collagen (Fig. 2A). The advantages of fiber-based systems include being able to extrude the fibers to any length required as well as being able to knit the fibers into meshes.^[9] Another major advantage of fiber-based systems is the ability to create a porous structure without any loss of mechanical strength of the material, thus producing an implant that is able to withstand physiological biomechanical stimuli as well as promote cellular ingrowth. The tendency of cells to align with the long axis of the fibers, regardless of the material chemistry, mimics the *in vivo* architecture of cells in tissues such as tendons, ligaments, and striated muscle.

Repeating Grooved Substrata

Initial attempts to fabricate grooves relied on techniques such as commercially available ruling engines or methods used to produce sound recording techniques.^[10,11] However, such methods often left features that were highly variable, not only from a topographical perspective, but also in surface chemistry. With the use of photolithography, followed by reactive ion etching, it became possible to fabricate highly ordered topographies with precise dimensions, such as repeating groove and ridge structures.^[6] Processes such as etching allow exact control over the depth, repeat spacing, and shape of the substratum. The techniques rely on the etching of materials such as silicon along precisely defined planes. By selecting silicon with different crystal orientations, it is possible to fabricate grooves of rectangular or V-shapes. Grooves are defined by the depth, width, and pitch (repeat spacing).

The application of V-shaped grooves produced from silicon (1,0,0) to study cell response on substrata was first reported by Brunette, Kenner, and Gould.^[6] Rectangular grooves could be produced using silicon

110. Since then, many studies of the effect of grooved substrates of varying dimensions on cell behavior have been reported. Increased adhesion, alignment, migration, and changes in gene expression have all been reported in various cell types cultured on grooves (for details see section “Cell Behavior Influenced by Substratum Topography”). In general, cells align to the groove long axis (Fig. 2B) and migrate in the same direction; depth seems to be the most important determinant in alignment. However, recent investigations have shown that certain cell types, such as chondrocytes, do not show the typical reaction to grooves as exemplified by fibroblasts or osteoblast.^[12,13] However, repeating grooved structures have proven to be an invaluable tool in understanding cell reaction to topography, and further, have translated well to the *in vivo* environment (see section “*In Vivo* Applications of Cell–Topography Interactions”).

Pits/Holes

An attractive possibility of using micromachined topographies is the ability to change the microenvironment around the implant surface, possibly providing conditions to promote implant integration. The use of pitted surface and holes has shown great potential in this regard, especially related to macrophage biology and osseointegration. Dow et al. reported on the “trapping” of macrophages in pit-type structures, opening the intriguing possibility of being able to control the positioning of macrophages around implant surfaces (Fig. 2C).^[14] As will be highlighted later in this entry, the benefits of macrophages around newly implanted constructs is of great importance, especially with regard to implant integration into bone.

Microenvironment has been shown to play a pivotal role in matrix and mineral deposition. Qu, Chehroudi, and Brunette, demonstrated that osteoblasts grown in truncated pits would form mineral within 2 weeks of culture, with the mineral deposited toward the bottom of the pit (Fig. 2D).^[15] In this study, the pits had a depth of 120 μm , and the investigators suggested that the cells effectively conditioned the immediate environment with growth factors and matrix, which promoted osteogenesis. It is likely that any surface that promotes mineralization on the implant in this way will stand to promote long-term integration and successful retention of the implant. In addition, recent work in our laboratory using pits on the order of 10 μm in depth has promoted very similar results, suggesting that it could simply be the microenvironment directly adjacent to the cells that is of importance. Although the optimal dimensions are not known, pits and hole-type topographies show great promise for promoting implant integration into bone.

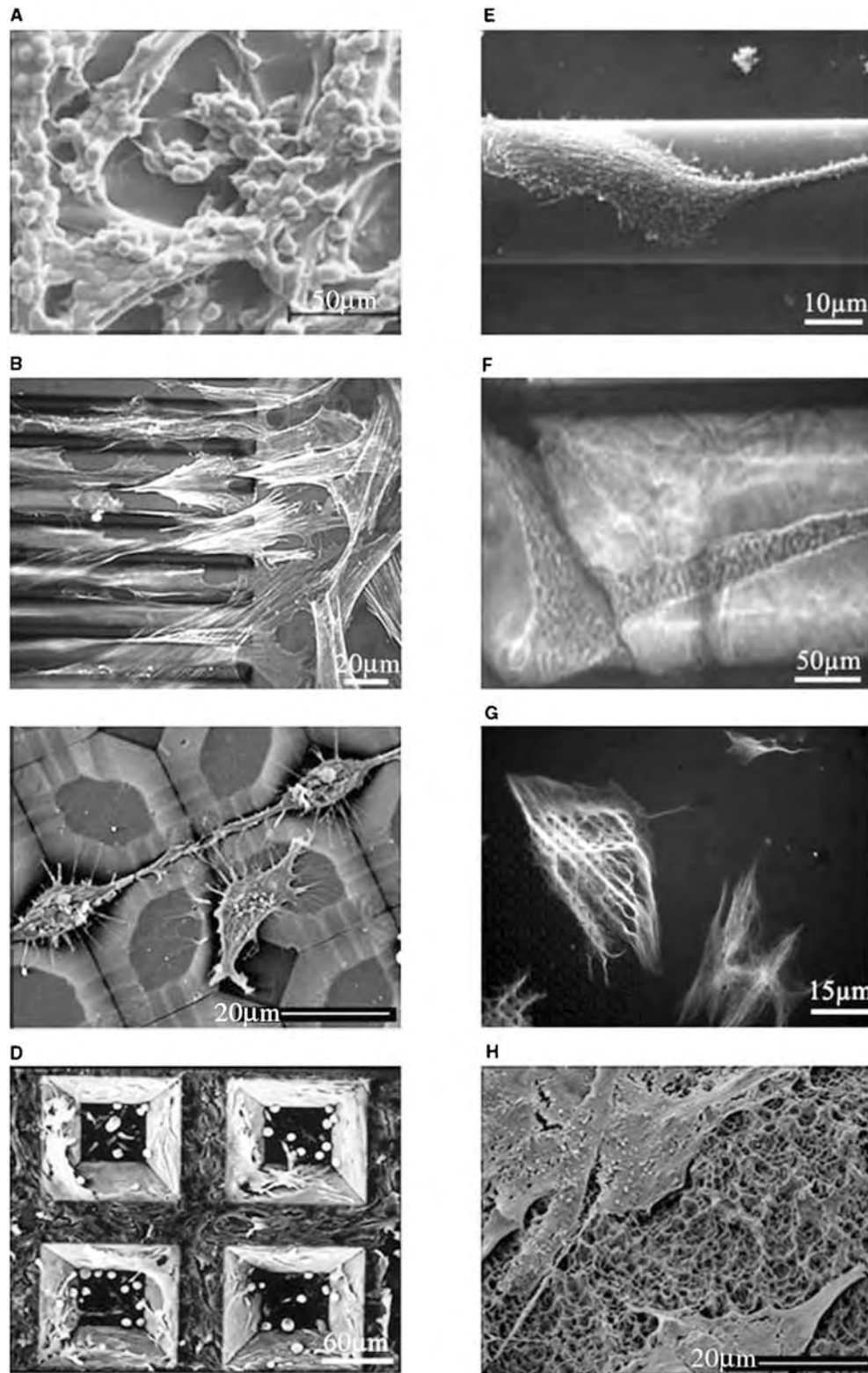


Fig. 2 Representative images of surface topographies currently investigated in research laboratories. (A) Epithelial cells growing on fibers, (B) actin staining of mesenchymal stem cells growing on repeating grooves, (C), P388D1 macrophages growing in cell traps, (D), rat calvarial osteoblasts growing in “pit”-type topographies, (E) fibroblasts growing on a convex topography, (F) ovine chondrocytes cultured in 240 μm -diameter quartz tube, (G) microtubule staining of human gingival fibroblasts growing on a discontinuous edged topography, and (H) human gingival fibroblasts cultured on SLA roughened titanium. The figure also highlights different methods for visualizing the cells. A, C, D, E, G, and H are scanning electron microscope images, B and G are epifluorescence, and F is phase contrast microscopy.

Tunnels/Tubes

Several cell types within the body, such as endothelial cells in blood vessels, live within tube-like structures. There has, however, been only scant attention paid to the response of cells to concave surfaces, in contrast to numerous reports of cell response to convex surfaces (Fig. 2E). In perhaps the most comprehensive examination to date, Riehle et al. investigated chondrocytes, osteoblasts, endothelia, and fibroblasts, and noted a differential response to concave features based on cell type.^[16] Initially, all cell types attached to the inner surface of the quartz tubes, including those with a diameter as small as 25 μm , although no specific orientation or alignment was noted. However, macrophages and fibroblasts were observed to become apoptotic within a relatively short space of time. In contrast, chondrocytes and osteoblasts thrived in such environments for up to 21 days, irrespective of the tube diameter. Both cell types formed large, motile cords, (Fig. 2F) and in the case of osteoblasts, large mineralized deposits were observed.

Discontinuous Substratum

Recently, we have begun investigation into the behavior of cells on microfabricated substratum that are discontinuous at the micrometer level.^[17] The rationale for designing such topographies was to understand further the mechanisms that underlie the process of contact guidance. Traditionally, the theory of contact guidance centers on the presence of a continuous edge, guiding a cell in a specified direction. The design of the discontinuous edged topographies (DES) is such that essentially the topographies comprised repeated open square boxes, with gaps at each corner. Examples of the DES topographies are shown in Fig. 2G. Our recent investigations using fibroblasts and epithelial cells have shown that DES topographies have great potential to be used not only *in vitro* to investigate cell guidance (Fig. 3A), but also *in vivo* to organize tissue around percutaneous implants (Fig. 3B).

Roughened Surfaces

Surface roughness was defined by Curtis and Clark as a randomly produced topography.^[1] The characterization of rough surfaces is a complex affair, as it is dependent on the scale length and the resolution of the measuring system, which may be based on mechanical, optical, or electron optical principles. The most widely accepted characterization is the average vertical amplitude or R_A value, which is calculated from an average of the feature sizes. In general, the higher the R_A value, the rougher is the surface. For example, a surface with

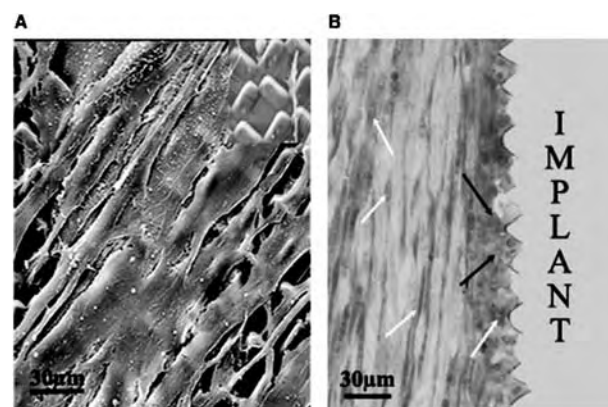


Fig. 3 Discontinuous edged surface topographies produced by microfabrication. (A) Human gingival fibroblasts cultured for 4 days on the DES topography. Note the alignment of the cells at diagonals to the ridges. (B) Toluidine blue staining of a DES topography implanted subcutaneously in a rat for 1 week. Black arrows indicate the presence of macrophages close to the implant surface. The white arrows highlight the fibrotic tissue composed of matrix and fibroblasts.

an R_A value of 0.6 μm would be considered to be rougher than one with an R_A value of 0.3 μm .

The simplest way of producing a rough surface is to rub the material with abrasive particles. A harsher method often used on metallic surfaces such as titanium and its oxides is to blast the material with abrasive grits such as aluminum or titanium oxide, giving rise to materials that have pits and peaks distributed more or less randomly.^[5] This process causes local deformations in the surface, and the R_A value depends on the size of the particle used. Other techniques to produce rough surfaces include titanium plasma spraying, which involves the coating of melted droplets of titanium onto the material surface, which then cool to form protruding structures (R_A value of $\approx 5.85 \mu\text{m}$). Acid etching, commonly accomplished using an aqueous mixture of nitric acid and hydrofluoric acid at a ratio of 10:1, results in a surface that has an R_A value of $\approx 0.59 \mu\text{m}$. Roughening procedures can also be used in combination. For example, blasting of titanium, followed by etching with acid, produces a surface (SLA) with an R_A value of $\approx 4.39 \mu\text{m}$.

The rationale for using rough surfaces is that the contact area between the tissue and the implant surface is increased, thus facilitating interaction between the tissues and the device. As will be discussed in the section "Cell Behavior Influenced by Substratum Topography," certain cell types appear to be more attracted to rough surfaces, a phenomenon termed rugophilia, but roughness also affects cell adhesion and spreading (Fig. 2H). Although most studies indicate that osteoblasts prefer rough surfaces, one study found that an

R_A value of $0.15\ \mu\text{m}$ promotes better cell attachment than R_A values of up to $1.2\ \mu\text{m}$. It appears that the geometry of the features as well as the overall roughness affect cell behavior. It has also been observed that with increasing time, enhanced mineralization and matrix deposition occur on rougher surfaces.

Nanotopography Vs. Microtopography

As techniques used to produce microfabricated topography evolve, the range of dimensions that can be fabricated increases. One area on which many research laboratories are currently focused is the use of nanometric sized surface features, such as pillars and holes. The logic of using such small features is that they are in the same size range as molecules such as collagen and other extracellular matrix molecules that cells will encounter within the body. While providing an excellent model for understanding cell sensing of small features, their applicability in implanted devices may be limited as matrix absorption at the tissue–implant interface would most likely mask such small topographical features, which are in the region of $10\text{--}100\ \text{nm}$ in size. One desired reaction at the implant surface is good mechanical interlocking, especially in bone contacting regions, because a major cause of implant failure is micromotion and friction at the tissue–implant interface. The implants that incorporate micrometer sized

pores, grooves, or even random rough surfaces have generally been found to lead to better implant integration. However, further investigations will be required to delineate the optimal size and geometry for optimal interaction with diverse tissues.

CELL BEHAVIOR INFLUENCED BY SUBSTRATUM TOPOGRAPHY

Cell Types Sensitive to Topography

Epithelial cells, fibroblasts, osteoblasts, chondrocytes, nerve axons, and macrophages are some of the cell types that have been demonstrated to be sensitive to topography.^[5] Table 1 lists cell types that have been shown to respond to substratum topography.

Adhesion

It has been long known that rougher surfaces generally promote adhesion of some cell types such as osteoblasts, whereas cells such as fibroblasts prefer smooth surfaces. When cells attach to a biomaterial, they demonstrate structures such as close contacts or focal adhesions. Cells that form a large number of focal adhesions are generally accepted to be better attached, to the material, although focal adhesion formation

Table 1 Cell types shown to be sensitive to substratum topography

Cell type	Reference
Astrocytes	Baac et al., 2004 (Mat. Sci. Eng.)
Chondrocytes	Riehle et al., 1998 (Exp. Biol. online); Hamilton et al., 2005a,b (Cell Biol. Int.)
Osteoblasts	Qu et al., 1996 (Oral Dis.); Wieland et al., 2005 (Biomaterials)
Neurons	Rajnicek et al., 1997a,b (J. Cell Sci.)
Fibroblasts	Brunette 1986 (Exp. Cell Res.); Walboomers et al., 1998 (Biomaterials)
Epithelium	Brunette 1986 (Exp. Cell Res.); Hamilton and Brunette 2005 (Exp. Cell Res.)
Neutrophils	Meyle et al., 1995 (J. Biomed. Mater. Res.); Chang et al., 2003 (J. Vasc. Surg.)
Smooth muscle cells	Thakar et al., 2003 (Biochem. Biophys. Res. Commun.)
Platelets	Kikuchi et al., 2005 (Biomaterials)
Endothelium	Riehle et al., 1998 (Exp. Biol. online); Barbucci et al., 2003 (J. Mater. Sci. Mater. Med.)
Macrophages	Meyle et al., 1995 (J. Biomed. Mater. Res.); Wojciak-Stothard et al., 1996 (Exp. Cell Res.)
Oligodendrocytes	Webb et al., 1995 (J. Cell Sci.)
Schwann cells	Miller et al., 2001 (Biomaterials)
Myoblasts	Williamson and Coombes, 2004 (Biomaterials)
Keratinocytes	Lauer et al., 2001 (Biomaterials)
Keratocytes	Teixiera et al., 2004 (J. Biomed. Mater. Res.)
Mesenchymal stem cells	Walboomers et al., 1999 (Biomaterials)
Mesenchymal progenitors	Wood et al., 1988 (J. Cell Sci.)

impacts on other cellular processes, as will be further highlighted in this section. The general rule to follow is that surface topography influences cell adhesion, although the nature and extent of the influences depends on the cell type involved, as well as the dimensions of the substratum topography.

The differential response of cell types to topography opens the possibility of using surfaces to select certain populations of cells based on their adhesive properties. For example, when considering surfaces to promote osseointegration, it would be anticipated that the desired populations of osteoblast would preferentially adhere to rough surfaces, in preference to cell types such as epithelium or fibroblasts. Macrophages are also known to prefer rougher surfaces, and indeed have their secretory profile altered on such surfaces.^[18] Indeed, it has recently been suggested that the presence of macrophages around bone contacting implants may be advantageous, and this possibility will be discussed later in the entry.

Alignment and Orientation

Within many tissues, the architecture or spatial arrangement is an important determinant of function. When considering cartilage, for instance, the cells and matrix are positioned in a way to allow the dispersion of compressive forces without causing damage to the tissue.^[12] However, when cartilage is damaged, the ability to perform this function is lost owing to the alteration in tissue architecture. Many tissues such as ligaments, tendons, and muscle rely on the specific position of the cells with regard to the surrounding matrix, and in many cases the cells have a determined alignment and orientation within the tissues.

Fibers, and repeating grooved substrates, in particular, have been shown to have a pronounced effect on cell and tissue alignment. For most cell types, it has been found that groove depth is a more important determinant of cell alignment than groove width. However, cells such as chondrocytes and macrophages have been shown to align to gratings only 80 nm in depth. Different cell types also vary in their reaction time to grooved substrates, with macrophages responding within 5 min, fibroblasts within 2 hr, and chondrocytes up to 6 hr. In general, cells align parallel to the groove long axis, although in the presence of chemical patterns superimposed perpendicular to the grooves, alignment of cells becomes dependent on the dimensions of the grating.

Contact Guidance

Contact, or topographic guidance refers to the tendency of cells to be guided in their direction of

locomotion by a continuous, physical edge, whether this is a groove, fiber, or other type of topographical structure. Many different cell types have been observed to show contact guidance, although cells such as chondrocytes have been observed to show guided migration without aligning parallel to the groove long axis.^[12,13] Controlled guidance of cells *in vivo* is of great importance during development, and has also been demonstrated to play a role at the tissue/implant interface. However, the exact mechanisms underlying contact guidance still remain unclear.

Differences in the adhesiveness of grooves, inflexibility of the cytoskeleton, and confinement of focal adhesions, have been put forward as drivers of contact guidance.^[19–21] Dunn and Heath suggested that substratum morphology and shape imposed restrictions on the formation of linear bundles of microfilaments located between focal adhesions of the leading lamella and the nuclear region.^[20] These bundles were thought to be relatively inflexible, so any attempt by the cell to accommodate curvature or angles, would result in the shortening of the leading lamella, which is important in cell spreading and locomotion.

Although the theory of Dunn and Heath seemed to explain cell reaction to topography on single-feature surfaces, O'Hara and Buck stated that it did not explain cell reaction to substratum with multiple features like grooved surfaces.^[21] Using multiple grooved substrata with features smaller than cell size, they showed that cells were sufficiently rigid to bridge over grooves, such that cell–substratum contact was confined to the surfaces of the ridges between the grooves. Brunette, Kenner, and Gould, however, found that focal adhesions could bend round the edges of grooves.^[22] Curtis and Clark suggested that cells simply aligned actin filaments and focal adhesions at discontinuities in the substrata.^[1] However, Oakley et al. subsequently demonstrated microtubules (MT) to be more sensitive to topography than F-actin, and that MT played key roles in elongation, alignment, and directed migration of cells in response to grooved substrata.^[23] Interestingly, a recent paper by Andersson et al. suggested that elongation and guidance of cells only occurs in the presence of a continuous edge.^[24]

Gap Guidance

Our laboratory is currently investigating the response of cells to discontinuous topographies, with an emphasis on understanding the conditions for the oriented guidance of cells.^[17] Results with gingival fibroblasts and periodontal ligament epithelial cells have shown that a continuous edge may not be a prerequisite to guiding a cell in a specific direction. We have termed this phenomenon “gap guidance” to



distinguish it from contact guidance.^[17] Analysis of cytoskeletal components suggests that a primary driver of gap guidance is focal adhesion stability. Reaction of cells to topographies such as grooves has been hypothesized to be a probabilistic affair, based on the likelihood that a cell will form a stable adhesion in any specified direction.^[25] Further results using DES have added further weight to this hypothesis. If focal adhesion stability is favored in any given orientation upon attachment and spreading, it is likely that the direction of locomotion will similarly follow that direction. Our results using DES have shown that cytoskeletal reaction is secondary to focal adhesion formation in determining alignment, but it is likely the cytoskeleton plays a larger role in migration and spreading.

Contact Guidance: Gene and Protein Expression

Although the general behavior of cells on topographical cues has been well described, the alterations in cell physiology at the molecular level remain elusive. Initial attempts to decipher altered gene expression on topographical cues required gene selection and analysis. Chou et al. showed that grooved substrata evoked an increase in fibronectin mRNA and its stability, as well as increased deposits of fibronectin in the secreted matrix.^[26] Recent developments in microarray technology have permitted more comprehensive investigations into genetic regulation by topography. Dalby and colleagues used microarray technology to compare the gene expression of human fibroblasts on grooved and flat surfaces, and demonstrated a large alteration in the pattern of expression.^[27] Interestingly, they noted that 24 hr postplating, many genes showed an upregulation, but at 5 days, many more genes were downregulated. For example, at 24 hr, genes such as the extracellular matrix protein, fibronectin, the ribosomal proteins 60S, 40S, the cytoskeletal component tubulin, and signal transduction molecules such as ras and protein kinase C were all significantly upregulated, but by 5 days were downregulated below the level of the controls. Such observations suggest that certain genes are required for cellular reaction to topography initially, but with increasing time are not required. However, it has been shown that cells will still respond to grooved substrates, even in the presence of cycloheximide, which blocks protein synthesis. Therefore, new protein synthesis is not required for cells to adhere and align on these types of substrates. At present, the role of such gene changes in cellular response to surface topography requires more investigation.

IN VIVO APPLICATIONS OF CELL—TOPOGRAPHY INTERACTIONS

Implants can be classified based on their chemical composition, topographical features, and additionally, the type of tissue with which they will interface. Implants can be fully (artificial hips, breast implants) or partially embedded (dental implants, catheters). For the purpose of this discussion we will limit our attention to partially embedded implants, specifically dental or percutaneous devices that penetrate the epithelium.

Dental implants face complex problems because they interface with three different tissue types, bone, epithelium, and connective tissue. In vitro experiments have shown that each cell type demonstrates preferential attachment to different topographical cues: bone (rough surfaces), epithelium (smooth), and fibroblasts (smooth, grooved). Therefore, the design of dental implants might benefit by incorporating multiple topographies to enhance the attachment of the different cell types at desired locations.

From a clinical standpoint, a vulnerable area of a dental implant, or any percutaneous device, is the region where the implant penetrates the epithelium. One reason for the failure of such devices is related to the downgrowth of epithelium (Fig. 4), which can lead to marsupialization, permigration, and avulsion at the implant tissue interface. Therefore, one could conceive of altering the topography on the implant surface that is intended for connective tissue to a surface that epithelial cells dislike, so as to impede this downgrowth. Chehroudi and colleagues in 1989 demonstrated a striking example of in vivo contact guidance by implanting V-shaped grooves of 10 μm depth with the grooves oriented either parallel or perpendicular to the long axis of the implant.^[28] As predicted from in vitro studies, the horizontally oriented grooves inhibited, whereas the vertically oriented grooves promoted, epithelial downgrowth. Additionally, it was demonstrated that not only are the fibroblasts aligned by the grooves, but so is the collagen and other matrix laid down by the cells.

Titanium and its alloys are commonly used in dental applications as they have good mechanical properties and biocompatibility. Such materials can be roughened through processes such as blasting, and acid etching, which is commonly done to ensure good integration in bone. The use of topographies produced by these processes has been successful clinically. The most common reason given for this excellent integration of dental implants with rough surfaces with bone was preferential adhesion of osteoblasts onto the rough material. However, osteoblasts are unlikely to be the first cell type to encounter the implant surface. One of the first cell types to be shown to attach, as well as to persist at the implant surface, is the macrophage.

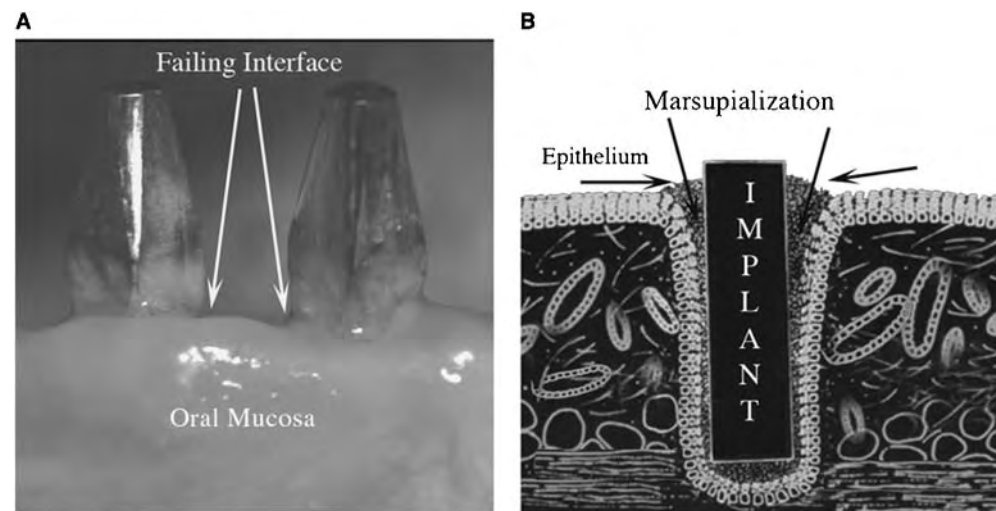


Fig. 4 Representative images of the failing interface in dental implants. (A) The rapid failure (4 mo) of an osteointegrated implant. Bacterial plaque (white material at base of the implant) can be seen, and the surrounding gingival tissue is inflamed. (Image courtesy of Dr. B. Chehroudi, Scident Family Dental Clinic). (B) Schematic representation of marsupialization of an implant. (Modified from Grosse-Siestrup et al. *J. Biomed. Mater. Res.* **1984**). Epithelium from the surface of the gingival grows down the side of the implant preventing optimal connective tissue attachment and osseointegration. Bacteria will then accumulate within the pocket, and the implant is eventually extruded from the tissue.

Although the presence of macrophages has always been associated with wound healing, they have always been treated with suspicion at the implant surface, as cells from the monocytic family are known to form osteoclasts. There has also been coassociation of macrophages with deposition of fibrous tissue and the formation of a capsule. Preliminary results from our laboratory suggest that the presence of rough surfaces at bone contacting surfaces could advantageously influence macrophage activities (unpublished data). In vitro, it was demonstrated that rough surfaces such as SLA stimulate ED1 macrophages (recruited

macrophages) to secrete interleukins 1β , and 6, as well as tumor necrosis factor α and monocyte chemoattractant protein 1.^[18] Recent in vivo work has shown that these same ED1 positive macrophages are present at the implant surface for up to 1 week postimplantation (Fig. 5A). These findings open the possibility that macrophages on some surfaces, could contribute to successful osseointegration.

Although, roughened surfaces produce osseointegration, little information exists as to how connective tissue responds in vivo to the presence of such surfaces. Recent investigations in our laboratory have studied

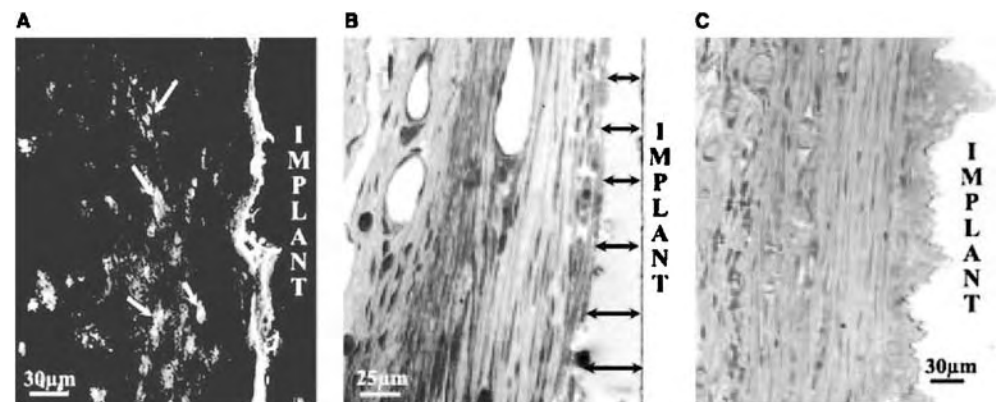


Fig. 5 In vivo performance of various surface roughened materials. (A) The presence of ED1 macrophages in connective tissue 1 week postimplantation of a titanium coated SLA implant in a subcutaneous model. Macrophages are visualized using epifluorescent techniques. (B) Polished and (C) SLA titanium implants placed in connective tissue at 6 weeks postimplantation. Note the separation from the implant on the polished surface. Tissue is stained with toluidine blue.

the influence of different roughened surfaces on connective tissue formation at the interface of tissue and implant. The implants were placed subcutaneously in rats for between 1 and 11 weeks, and analysis comprised attachment, capsule thickness, and degree of separation of the tissue from the implant. Overall, surface roughness promoted stable connective tissue attachment, and a significantly lower magnitude and frequency of separation at the implant/tissue interface (see Figs. 5 and 5D). In addition, rougher surfaces had a thinner capsule present. Interestingly, acid etched surfaces, which have a significantly lower R_A value than SLA, performed comparably with the rougher surfaces, suggesting that not only roughness but also feature geometry affects cell reactions.

A possible explanation for the successful rough surface connective tissue integration is that surface topography can promote fibronectin expression in gingival fibroblasts, and correlates closely with cell shape, and in particular thickness. In many cellular systems, changes in cell shape are associated with altered cell behavior. Wieland et al. demonstrated that fibroblasts had a significantly greater thickness on rough surfaces compared with smooth.^[29] It is possible that the cell attachment on rough surfaces results from the absorption of the fibronectin produced by active cells to the implant surface, followed by deposition of collagen and other connective tissue matrix components, leading to good mechanical interlocking.

CONCLUSIONS

Although cell behavior has been well documented on surface topography in vitro (cell adhesion, shape, migration, cytoskeletal function), two areas will require further investigation:

1. The long-term function of different surface topographies in vivo.
2. Further analysis of the alterations in cellular function by topography at the molecular level.

Given that different cell types show desirable reactions to different surface topographies, the potential exists to produce implants with topographies at different locations that interact appropriately with the tissue type with which they interface. Further utilization of micromachining techniques will enhance the possibility of producing tissue-specific "designer" implants to further the field of regenerative medicine. As more precise methods of fabrication have evolved, so has our understanding of the role of topography in cellular response, particularly at the implant interface. Topography not only allows development of more suitable implant designs, but further provides

information on the mechanisms cells utilize to sense their surrounding environment.

ACKNOWLEDGMENTS

The authors would like to thank Mr Andre Wong for his invaluable assistance on technical issues. The authors wish to acknowledge the financial support of the Canadian Institute of Health Research and the ITI Foundation.

ARTICLES OF FURTHER INTEREST

Gradient Surfaces: Preparation, Characterization, and Interactions with Biological Species; Polyelectrolyte Multilayers; Protein Adsorption Modeling; Surface Coatings; Surface Modification

REFERENCES

1. Curtis, A.S.G.; Clark, P. The effects of topographic and mechanical properties of materials on cell behavior. *Crit. Rev. Biocomp.* **1990**, *5*, 343–362.
2. Walter, J.B.; Israel, M.S. *General Pathology*, 6th Ed.; Churchill Livingstone: Edinburgh, 1987, 387.
3. Harrison, R.G. On the stereotropism of embryonic cells. *Science* **1911**, *34*, 279–381.
4. Weiss, P. Experiments on cell and axon orientation in vitro: the role of colloidal exudates in tissue organization. *J. Exp. Zool.* **1945**, *100*, 353–386.
5. Curtis, A.S.G.; Wilkinson, C.D.W. Reactions of cells to topography. *J. Biomater. Sci. Polym. Ed.* **1998**, *9*, 1313–1329.
6. Brunette, D.M.; Kenner, G.S.; Gould, T.R.L. Grooved titanium surfaces orient growth and migration of cells from human gingival explants. *J. Dent. Res.* **1983**, *62*, 1045–1048.
7. Clark, P.; Connolly, P.; Curtis, A.S.G.; Dow, J.A.; Wilkinson, C.D. Topographical control of cell behavior. I. Simple step cues. *Development* **1987**, *99*, 439–448.
8. Jenkins, D.H.; Forster, I.W.; McKibbin, B.; Ralis, Z.A. Induction of tendon and ligament formation by carbon implants. *J. Bone Joint Surg. Br.* **1977**, *59*, 53–57.
9. Paquay, Y.C.; de Ruyter, J.E.; van der Waerden, J.P.; Jansen, J.A. Wound healing phenomena in titanium fibre mesh: the influence of the length of implantation. *Biomaterials* **1997**, *18* (2), 161–166.
10. Curtis, A.S.G.; Varde, M. Control of cell behavior: topological factors. *J. Nat. Can. Inst.* **1964**, *33*, 15–26.
11. Rovinsky, Y.A.; Slavnya, I.L. Spreading of fibroblast-like cells on grooved substrates. *Exp. Cell. Res.* **1974**, *84*, 199–206.
12. Hamilton, D.W.; Riehle, M.O.; Monaghan, W.; Curtis, A.S.G. Articular chondrocyte passage number: influence on adhesion, migration, cytoskeletal organization

- and phenotype in response to micro- and nanometric topography. *Cell Biol. Int.* **2005**, *29* (6), 408–421.
13. Hamilton, D.W.; Riehle, M.O.; Rappuoli, R.; Monaghan, W.; Barbucci, R.; Curtis, A.S.G. The response of primary articular chondrocytes to micrometric topography and sulfated hyaluronic acid based matrices. *Cell Biol. Int.* **2005**, *29* (8), 605–615.
 14. Dow, J.A.T.; Clark, P.; Connolly, P.; Curtis, A.S.; Wilkinson, C.D. Novel methods for the guidance and monitoring of single cells and simple networks in culture. *J. Cell Sci.* **1986**, *8* (suppl.), 55–79.
 15. Qu, J.; Chehroudi, B.; Brunette, D.M. The use of micromachined surfaces to investigate the cell behavioural factors essential to osseointegration. *Oral Dis.* **1996**, *2*, 102–115.
 16. Riehle, M.; Ferris, D.; Hamilton, D.; Curtis, A.S.G. Cell behavior in tubes. *Exp. Biol. Online* **1998**, *3* (2); <http://link.springer.de/link/service/journals/00898/tocs/t8003100.htm> accessed Apr 2005.
 17. Hamilton, D.W.; Brunette, D.M. Gap guidance of fibroblasts and epithelial cells by discontinuous surface topographies. *Exp. Cell Res.* **2005**, *309* (2), 429–437.
 18. Refai, A.K.; Textor, M.; Brunette, D.M.; Waterfield, J.D. Effect of titanium surface topography on macrophage activation and secretion of proinflammatory cytokines and chemokine. *J. Biomed. Mat. Res.* **2004**, *70* (2), 194–205.
 19. Weiss, P.; Taylor, A.C. Fish Scales as substratum for uniform orientation of cells in vitro. *Anat. Rec.* **1956**, *124*, 381.
 20. Dunn, G.A.; Heath, J.P. A new hypothesis of contact guidance in tissue cells. *Exp. Cell Res.* **1976**, *101*, 1–14.
 21. O'Hara, P.T.; Buck, R.C. Contact guidance in vitro: a light, transmission and scanning electron microscope study. *Exp. Cell Res.* **1979**, *121*, 235–249.
 22. Brunette, D.M. Spreading and orientation of epithelial cells on grooved substrata. *Exp. Cell Res.* **1986**, *167*, 203–217.
 23. Oakley, C.; Brunette, D.M. The sequence of alignment of microtubules, focal contacts and actin filaments in fibroblasts spreading on smooth and grooved titanium substrata. *J. Cell Sci.* **1993**, *106*, 343–354.
 24. Andersson, A.M.; Olsson, P.; Lidberg, U.; Sutherland, D. The effects of continuous and discontinuous groove edges on cell shape and alignment. *Exp. Cell Res.* **2003**, *288*, 177–188.
 25. Brunette, D.M. Fibroblasts on micromachined substrata orient hierarchically to grooves of different dimensions. *Exp. Cell Res.* **1986**, *164*, 11–26.
 26. Chou, L.; Firth, J.D.; Uitto, V.-J.; Brunette, D.M. Substratum surface topography alters cell shape and regulates fibronectin mRNA level, mRNA stability, secretion and assembly in human fibroblasts. *J. Cell Sci.* **1995**, *108* (4), 1563–1573.
 27. Dalby, M.J.; Riehle, M.O.; Yarwood, S.J.; Wilkinson, C.D.; Curtis, A.S. Nucleus alignment and cell signaling in fibroblasts: response to a micro-grooved topography. *Exp. Cell Res.* **2003**, *284* (2), 274–282.
 28. Chehroudi, B.; Gould, T.R.; Brunette, D.M. Effects of a grooved titanium-coated implant surface on epithelial cell behavior in vitro and in vivo. *J. Biomed. Mater. Res.* **1989**, *23*, 1067–1085.
 29. Wieland, M.; Chehroudi, B.; Textor, M.; Brunette, D.M. Use of Ti-coated replicas to investigate the effects on fibroblast shape of surfaces with varying roughness and constant chemical composition. *J. Biomed. Mat. Res.* **2002**, *60* (3), 434–444.

Sutures

C.C. Chu

Department of Textiles and Apparel & Biomedical Engineering Program, Cornell University, Ithaca, New York, U.S.A.

INTRODUCTION

Every wound requires the use of biomaterials to close for subsequent successful healing. The complexities involved in wound healing—such as the various degrees of wound strength during the process of healing, the involvement of different type of tissues in a wound, the presence of biomaterials, or the variety of surgical wounds each with its own healing problems—call for different types of wound closure materials. The choice of these biomaterials is based largely on the type of wound and surgeons' preferences.

OVERVIEW

Wound closure biomaterials are generally divided into three major categories: sutures, staples/ligating clips, and tissue adhesives. The sutures have the longest use history, have received the most attention, and are the most widely used in wound closure. Ligating clips and staples facilitate anastomosis with minimal trauma, necrosis, or interruption of tissue function and their use has steadily increased in specific clinical conditions, particularly with the availability of synthetic absorbable ligating clips and staples. Tissue adhesives are the least frequently used for wound closure at the present time, even though they received considerable attention in the 1960s; however, some new tissue adhesives have recently received increasing attention. There are several reviews about wound closure biomaterials.^[1–9] In this article, the focus is on suture-based wound closure biomaterials because they are the most frequently used and studied. All aspects of suture-based wound closure biomaterials, their classification, and their chemical, physical, mechanical, biological, and biodegradative properties will be concisely covered.

A suture is a strand of material, either natural or synthetic, used to ligate blood vessels and to fasten tissue together. Textile materials were the earliest and most frequently used materials for surgical wound closure. Linen was used as a suture material as early as 4000 years ago. Since then, numerous materials have been used as ligatures and sutures: iron wire, gold, silver, dried gut, horse hair, strips of hide, bark fibers, silk, linen, and tendon. Among them, catgut and silk

dominated the suture market until 1930. The introduction of steel wire and synthetic nonabsorbable fibers such as nylon, polyester, and polypropylene during and after World War II greatly expanded the chemical composition of suture materials. During the early 1970s, the introduction of two synthetic absorbable suture materials, Dexon[®] (Davis & Geck, Danby, CT) and Vicryl[®] (Ethicon, Inc., Somerville, NJ), opened a new milestone for suture materials. Owing to their precisely controlled manufacturing processes and uniform and reproducible properties, these synthetic absorbable sutures have received a great deal of attention from both surgeons and researchers. Since then, several new synthetic absorbable suture materials such as PDSII[®] (Ethicon), Maxon[®] (US Surgical/Davis & Geck), Monocryl[®] (Ethicon), and Biosyn[®] (US Surgical, Norwalk, CT) have become commercially available. The most important advantage of synthetic absorbable sutures is their reproducible and predictable degradability inside a biological environment. This property enables the sutures to minimize chronic undesirable tissue reactions after the sutures have lost most of their mechanical properties. The latest introduction of sutures is the one having multifunction, and a typical example is Triclosan-coated Vicryl, Vicryl Plus[®], which has antimicrobial capability to reduce the chance of wound infection.^[10,11] Today, surgeons can choose between a large number of suture materials with various chemical, physical, mechanical, and biological properties.

NEEDLES

The surgical needle, to which a suture is attached, has the primary function of introducing the suture through the tissues to be brought into apposition. Ideally, the needle has no role in wound healing but inappropriate needle selection can prolong the operating time and/or damage tissue integrity, leading to such complications as tissue necrosis, wound dehiscence, bleeding, leakage of anastomoses, and poor tissue apposition. Clearly, needle selection is an important factor in suturing and the progress of wound healing. For maximum effectiveness, the surgical needle must be able to carry the suture material through tissue with minimal trauma

Table 1 Needle selection criteria

Minimal tissue trauma
High sharpness (acuity)
Corrosion resistance
High strength
Stable shape
Proper balance
Abrasion resistance
Smooth profile

and, to achieve this goal, needles are required to satisfy several criteria shown in Table 1.

Needles are commonly fabricated from stainless steel, a material that has high strength, is readily available, presents few manufacturing problems, can be polished to a smooth finish, and is relatively inexpensive. Because the gripping action of needle holders and the repeated grip-release action during surgical procedures can damage the needle surface, wear and abrasion resistance is a necessary prerequisite for surgical needles.

A wide variety of surgical needles exists but all types have basically three components: the point, the body, and the eye, swage, or attachment end. Differences between needles arise from variations and modifications within these basic parameters.

The most important biomechanical performance parameters for surgical needles are acuity or sharpness, bending resistance, and ductility. Needle acuity is the most important biomechanical parameter because the needle is required to penetrate tough or fibrous tissue without undue trauma, creating a pathway through which the body of the needle and the attached suture will then pass with minimum drag. Further, the relative ease of tissue penetration with a sharp needle facilitates surgical dexterity, at the same time reducing the risk of needle bending or deformation during use because less force has to be applied to a sharp needle.

COMMERCIAL SUTURES, THEIR SIZES, AND THEIR PHYSICAL CONFIGURATIONS

Suture materials are generally classified into two broad categories: absorbable and nonabsorbable. Absorbable suture materials generally lose their entire or most of their tensile strength within two to three months; those that retain most of their initial strength longer than two to three months are considered nonabsorbable. The absorbable suture materials are catgut (collagen sutures derived from sheep intestinal submucosa), reconstituted collagen, polyglycolide (Dexon[®],

Dexon II[®], Dexon S[®]), poly(glycolide-lactide) random copolymer (Vicryl[®]), antimicrobial-coated Vicryl[®] (Vicryl Plus[®]), poly-p-dioxanone (PDS[®], PDSII[®]), poly(glycolide-trimethylene carbonate) block copolymer (Maxon[®]), poly(glycolide-ε-caprolactone) (Monocryl[®]), and Glycolide-dioxanone-trimethylene carbonate block copolymer (Biosyn[®]). The nonabsorbable sutures are divided into the natural fibers (i.e., silk, cotton, linen) and man-made fibers (i.e., polyethylene, polypropylene, polyamide, polyester, polytetrafluoroethylene (Gore-Tex[®]), and stainless steel). Table 2 summarizes all of the commercial suture materials that are available mainly in the United States, Europe and the Pacific area, their generic and trade names, their physical configurations, and their manufacturers. Fig. 1 shows the scanning electron images of some of the most popular commercial sutures.

Suture materials are also classified according to their size. Currently, two standards are used to describe the size of suture materials: USP (United States Pharmacopoeia) and EP (European Pharmacopoeia).^[9] Table 3 summarizes both EP and USP standards. The USP standard is more commonly used. In the USP standard, the size is represented by a series of combinations of two Arabic numbers: a zero and any number other than zero, such as 2-0 (or 2/0). The larger the first number, the smaller the suture material diameter. Sizes greater than 0 are denoted by 1, 2, 3, etc. This standard size also varies with the type of suture material.

In the EP standard, the code number ranges from 0.1 to 10. The corresponding minimum diameter (mm) can be easily calculated by taking the code number and dividing by 10. The EP standard does not separate natural from synthetic absorbable sutures as the USP standard does.

In terms of the physical configuration of suture threads, they can be classified into monofilament, multifilament, twisted, and braided. Suture materials made of nylon, polyester, and stainless steel are available in both multifilament and monofilament forms. Catgut, reconstituted collagen, and cotton are available in twisted multifilament form; Dexon[®], Vicryl[®], Monosyn[®], Polysorb[®], PolySyn FA[®], Safil[®], BioSorb[®], silk, and polyester-based and polyamide-based suture materials are available in the braided multifilament configuration. PDS[®], Maxon[®], Monocryl[®], Biosyn[®], Caprosyn[®], MonoPlus[®], polypropylene, and Gore-Tex[®] (polytetrafluoroethylene) suture materials exist in monofilament form only. Stainless steel metallic suture materials can be obtained in either monofilament or twisted multifilament configurations. Another unique physical configuration of suture material is available in polyamide (nylon 6) and has the trade name Supramid[®]; it has a twisted core covered by a jacket of the same material.

Table 2 List of commercial sutures, their trade names, and their manufacturers

Generic name	Trade name	Physical configuration	Surface treatment	Manufacturer
<i>Natural absorbable sutures</i>				
Catgut			Plain and chromic	Surgical Specialties Corporation
Catgut			Plain and chromic	Dyneke Sutures
Catgut		Multifilament	Monofilament finish, plain, and chromic	SURU
Collagen—bovine	Surgical gut		Plain, chromic, and mild chromic	USS/DG
Collagen—bovine, ovine	Surgical gut		Plain and chromic	Ethicon
<i>Synthetic absorbable sutures</i>				
Glycolic acid and trimethylene carbonate—copolymer	MAXON	Monofilament	Clear or dyed green	USS/DG
Glycolic acid and trimethylene carbonate—copolymer	MAXON CV	Monofilament	Clear or dyed green	USS/DG
Glycolic acid—homopolymer	DEXON II	Braided	Dyed green or bicolored or undyed; coated with polycaprolate	USS/DG
Glycolic acid—homopolymer	DEXON S	Braided	Dyed green or undyed	USS/DG
Glyconate	Monosyn	Mid-term braided	Dyed violet or undyed	B. Braun Melsungen AG
Lactomer	POLYSORB	Braided	Dyed violet or undyed	United States Surgical
Poliglecaprone 25	MONOCRYL	Monofilament	Dyed or undyed	Ethicon
Polydioxanone	PDS II	Monofilament	Dyed or undyed	Ethicon
Polyester—glycolide (60%), dioxanone (14%), trimethylene (26%)	BIOSYN	Monofilament	Dyed violet or undyed	USS/DG
Polyglactin 910	Coated VICRYL	Braided	Dyed or undyed	Ethicon
Polyglactin 910	VICRYL RAPIDE	Braided	Undyed	Ethicon
Polyglactin 910	Coated VICRYL Plus Antibacterial	Braided	Dyed violet or undyed; coated with glycolide and lactide (polyglactin 370) and calcium stearate	Ethicon
Polyglycolic acid	PolySyn FA	Braided	Undyed	Surgical Specialties Corporation
Polyglycolic acid (PGA)	SURUCRYL	Braided	Coated with polycaprolactone calcium stearate	SURU
Poly(L-lactide/glycolide) copolymer	Panacryl	Braided	95% lactide/ 5% glycolide and coated with 90% caprolactone/ 10% glycolide copolymer	Ethicon (discontinued in late 2003)
Polyglycolic acid—low molecular weight	Safil—Quick	Mid-term braided	Coated	B. Braun Melsungen AG
Polyglycolic acid—pure	Safil—green	Mid-term braided	Coated; dyed green	B. Braun Melsungen AG
Polyglycolic acid—pure	Safil—violet	Short-term braided	Coated; dyed violet or undyed	B. Braun Melsungen AG

(Continued)

Table 2 List of commercial sutures, their trade names, and their manufacturers (*Continued*)

Generic name	Trade name	Physical configuration	Surface treatment	Manufacturer
Polyglytone 6211 synthetic polyester	CAPROSYN	Monofilament	Undyed	United States Surgical
Poly-p-dioxanone	MonoPlus	Long-term monofilament	Dyed violet	B. Braun Melsungen AG
	BioSorb	Braided	Dyed green	Alcon Laboratories
	BioSorb Coated	Braided	Dyed green; coated with polycaprolate	Alcon Laboratories
<i>Nonabsorbable sutures</i>				
316L stainless steel	Surgical Stainless Steel	Monofilament and multifilament		Ethicon
Corrosion-resistant steel	Steelex	Twisted or monofilament		B. Braun Melsungen AG
<i>Nonabsorbable sutures</i>				
Fibroin-natural	SOFSILK	Braided	Dyed black with logwood extract; coated with special wax	USS/DG
Linen	Linatrix	Twisted	Natural white	B. Braun Melsungen AG
Nylon	NYLENE	Monofilament	Dyed blue	Dyne Sutures
Nylon		Monofilament	Dyed black	Surgical Specialties Corporation
Nylon 6 and 6.6	ETHILON		Dyed green or black and undyed clear	Ethicon
Nylon 6 and 6.6	MONOSOF	Monofilament	Dyed black with logwood extract or undyed; silicone coated	USS/DG
Nylon 6 and 6.6	DERMALON	Monofilament	Dyed blue; silicone coated	USS/DG
Nylon 6 and 6.6	SURGILON	Braided	Dyed blue; silicone coated	USS/DG
Nylon		Monofilament	Dyed black	Alcon Laboratories
Polyethylene terephthalate	TI·RON	Braided	Dyed blue or undyed; silicone coated or uncoated	USS/DG
Poly(vinylidene fluoride) and polyvinylidene fluoride-co-hexafluoropropylene)	PRONOVA		Dyed blue	Ethicon
Polyamide (nylon)	SURULON	Monofilament	Dyed blue/black	SURU
Polyamide 6	Dafilon	Monofilament	Dyed blue	B. Braun Melsungen AG
Polyamide 6.6	Dafilon	Monofilament	Dyed black	B. Braun Melsungen AG
Polyamide 6/6.6	Supramid	Pseudomonofilament-core polyamide 6.6	Cover polyamide 6	B. Braun Melsungen AG
Polybutester	NOVAFIL	Monofilament	Dyed and undyed	USS/DG
Polybutester	VASCUFIL	Monofilament	Dyed blue; coated	USS/DG
Polybutylene terephthalate polyester	Miralene	Monofilament	Dyed blue	B. Braun Melsungen AG
Polyester	SURGIDAC	Braided	Dyed green; coated	USS/DG

(Continued)

Table 2 List of commercial sutures, their trade names, and their manufacturers (*Continued*)

Generic name	Trade name	Physical configuration	Surface treatment	Manufacturer
Polyester	SURUPOL	Braided	Dyed green/white, silicone coated	SURU
Polyester		Braided or monofilament	Dyed white or green	Alcon Laboratories
Polyester	Polyviolene	Braided	Dyed green or white	Surgical Specialties Corporation
Polyether	DYLOC	Monofilament	Dyed blue	Dyneke Sutures
Polyethylene terephthalate polyester	Dagrofil HRT	Braided	Uncoated; dyed green	B. Braun Melsungen AG
Polyethylene terephthalate polyester	Synthofil	Braided	Coated; dyed green or undyed	B. Braun Melsungen AG
Polyethylene terephthalate polyester	PremiCron	Braided	Silicone coated; dyed green or white	B. Braun Melsungen AG
Polyethylene terephthalate polyester	ETHIBOND	Braided	Dyed green; coated with polybutylate	Ethicon
Polypropylene	Premilene	Monofilament	Dyed blue w/ copper phthalocyanine	B. Braun Melsungen AG
<i>Nonabsorbable sutures</i>				
Polypropylene	PROLENE		Dyed blue	Ethicon
Polypropylene	SURGIPRO	Monofilament	Dyed blue or undyed	United States Surgical
Polypropylene	SURGIPRO II	Monofilament	Dyed blue or undyed	United States Surgical
Polypropylene	SURULENE	Monofilament	Dyed blue	SURU
Polypropylene		Monofilament	Dyed blue	Alcon Laboratories
Polypropylene		Monofilament	Dyed blue	Surgical Specialties Corporations
Polyvinylidene fluoride	RADENE	Monofilament	Dyed blue	Dyneke Sutures
Polyvinylidene fluoride	VILENE	Monofilament	Dyed blue	Dyneke Sutures
Siliconized polyester	POLYFLEX	Braided	Dyed black	Dyneke Sutures
Siliconized polyester	DYFLEX	Braided	Dyed green	Dyneke Sutures
Siliconized polyester	TEFLEX	Braided	White	Dyneke Sutures
Silk	SURUSIL	Braided	Dyed black, treated with wax	SURU
Silk		Twisted or braided	Dyed black or white	Alcon Laboratories
Silk		Braided	Dyed black	Surgical Specialties Corporation
Silk-fibroin	PERMA-HAND	Natural waxes and gums removed	Dyed black and undyed; special wax	Ethicon
Silk-fibroin	PERMA-HAND-virgin	Sericin gum not removed		Ethicon
Silk-natural	Silram	Braided	Coated; dyed black	B. Braun Melsungen AG
Silk-natural	Virgin Silk	Twisted	Dyed methylene blue	B. Braun Melsungen AG
Stainless steel	FLEXON	Multifilament	FEB polymer coating	USS/DG
Stainless steel	Steel	Monofilament		USS/DG
Treated silk	DYSILK	Braided	Dyed black	Dyneke Sutures
USP8/0 to 10/0 in various materials	MICROFLEX			Dyneke Sutures

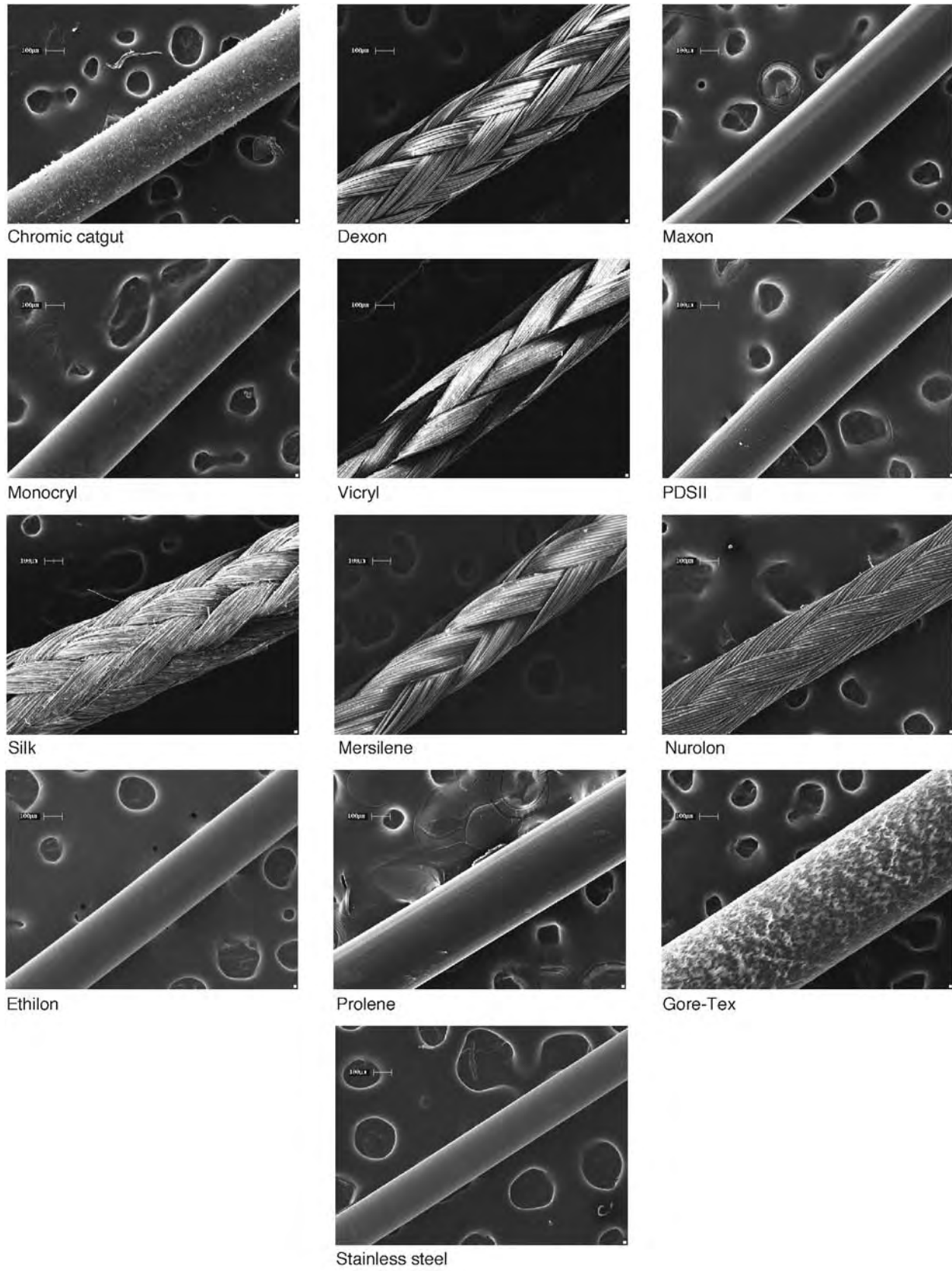


Fig. 1 Scanning electron images of some commercial sutures.

Table 3 USP suture size classification

USP size codes		EP size codes	Suture diameter (mm)
Nonsynthetic absorbable materials	Nonabsorbable and synthetic absorbable materials	Absorbable and nonabsorbable materials	Min. Max.
	11/0	0.1	0.01–0.019
	10/0	0.1	0.02–0.029
	9/0	0.3	0.03–0.039
	8/0	0.4	0.04–0.049
8/0	7/0	0.5	0.05–0.069
7/0	6/0	0.7	0.07–0.099
6/0	5/0	1	0.10–0.14
5/0	4/0	1.5	0.15–0.19
4/0	3/0	2	0.20–0.24
3/0	2/0	2.5	0.25–0.29
2/0	0	3	0.30–0.39
0	1	4	0.40–0.49
1	2	5	0.50–0.59
2	3	6	0.60–0.69
3	4	7	0.70–0.79
4	5	8	0.80–0.89
5	6	9	0.90–0.99
6	7	10	1.00–1.09

Suture materials are frequently coated to enhance their handling properties, particularly to reduce tissue drag when passing through the needle tract and to ease the sliding of knots down during knotting (i.e., knot tie-down). Although nonabsorbable bee's wax, paraffin wax, silicone, and polytetrafluoroethylene (Teflon) are the traditional coating materials, new coating materials have been reported, particularly those that are absorbable. This is because the coating materials used for absorbable sutures must be absorbable; traditional nonabsorbable coating materials such as wax are not appropriate for absorbable sutures.^[12–14] Furthermore, absorbable coating materials have better tissue biocompatibility because of the lack of chronic tissue reaction. There are basically two types of absorbable coating materials: water-soluble and water-insoluble. Water-insoluble coating materials have chemical constituents similar to those in the suture and they are broken down by hydrolysis. They remain on the suture surface longer than water-soluble coatings. A typical example is polyglactin 370 used for Vicryl suture. Dexon II sutures have a polycaprolate coating which is water-insoluble. Water-soluble coating materials dissolve promptly to reveal the underneath uncoated suture after wound closure. A typical example is poloxamer 188 found on Dexon Plus. There is, however, one technical concern about using water-soluble coating

material in actual wound closure. Suture materials are frequently soaked in saline after their removal from packages before use. Some or the bulk of water-soluble coating materials might be removed by this routine soaking practice. Thus, it is important to minimize the time of soaking when dealing with water-soluble coated suture materials. Multifilament sutures are more commonly coated than monofilament sutures. For example, multifilament Vicryl and Dexon Plus or II have coating materials applied, whereas monofilament PDS and Maxon sutures have no coatings.

Although coating of suture materials facilitates easy passage through tissue, it frequently results in poor knot security. For example, Dexon Plus and coated Vicryl require four or five square throws to form secure square knots, whereas the uncoated Dexon and Vicryl sutures form secure knots with only two throws (1 = 1).^[15,16] Water-soluble coating materials such as poloxamer 188, found on Dexon Plus, do not suffer from the adverse effects of water-insoluble coating materials on knot security.

There are several other patented procedures and materials reported to improve knot tie-down performance (the ease of sliding a knot down the suture into place during knotting) or/and knot security (the ability of a knot to hold after knotting).^[17–20] In general, a coating designed to improve knot tie-down can reduce knot security. It is difficult to achieve both ease of knot tie-down and enhanced knot security of sutures. There are very few reported treatments that would achieve these two contradictory and mutually exclusive properties. One of them is the use of a combination of both coating and textured yarns.^[17] The coating materials used in that study included sucrose fatty acid ester, bees wax, paraffin wax, polytetrafluoroethylene, silicone, poly(oxyethylene-co-oxypropylene), polyglactin-370, and gelatin. Other recently reported absorbable but water-insoluble coating materials that can improve knot tie-down and knot security are high-molecular-weight poly- ϵ -caprolactone, a copolymer of at least 90% by weight of caprolactone and 10% at most of other biodegradable monomers such as glycolide, lactide, and their derivatives,^[18,19] or a random copolymer of 25–75% by weight of glycolide and the rest of trimethylene carbonate.^[20] The observed improved knot tie-down and knot security was attributed to deep penetration and even distribution of the coating materials into the interstices of suture filaments.

ESSENTIAL PROPERTIES OF SUTURE MATERIALS

There are four essential properties of suture materials: physical and mechanical, handling, biological, and biodegradative. Table 4 summarizes the characteristics

Table 4 Four major categories of the characteristics of suture materials

Physical/mechanical	Handling	Biocompatibility	Biodegradation
USP vs. EP size (diameter)	Pliability	Inflammatory reaction	Tensile breaking strength and mass loss profiles
Mono vs. multifilament	Memory	Propensity toward wound infection, calculi formation, thrombi formation, carcinogenicity, allergy	Biocompatibility of degradation products
Tensile breaking strength and elongation	Knot tie-down		
Modulus of elasticity	Knot slippage		
Stiffness	Tissue drag		
Stress relaxation and creep			
Capillarity			
Swelling			
Coefficient of friction			

S

of each of the four essential properties. It is important to recognize that these characteristics are interrelated. For example, capillarity of a suture material, a physical/mechanical property, is closely related to the ability of the suture to transport bacteria which is a biological property. The modulus of elasticity, a physical/mechanical property, is frequently used to relate to pliability of sutures, a handling property. A brief description of each of those essential properties will be discussed subsequently; they are listed in Tables 5–9. Readers should be aware of the fact that the data in the tables will vary, depending on specific clinical and/or physical environments that suture materials are subjected to and the constant refining of manufacturing processes by suture manufacturers.

Physical and Mechanical Properties

Physical and mechanical properties are probably the most important ones in terms of suture function, i.e., to close wounds and carry physiologic load during healing. They include those related to strength, stiffness, viscoelasticity, coefficient of friction, compliance, size, form (monofilament or multifilament), fluid absorption, transport, etc. Strength includes knotted and unknotted (straight pull) tensile strength, modulus of elasticity (relating to stiffness), elongation at break, and toughness.

The tensile strength character is the most frequently reported and studied physical/mechanical property of suture materials. Because tensile strength is expressed in terms of the cross-sectional area of the material, it is normalized based on the dimension of the material and hence can be used to compare sutures having different

chemical structure or/and sizes. Tensile breaking force, however, does not take into account suture size (i.e., diameter). Thus, a larger size suture will have a higher tensile breaking force than the same suture of a smaller size, even though the two sutures may have the same tensile strength. Therefore, a meaningful comparison of tensile breaking force of several sutures should be made using the same suture size (diameter) and form. In addition, knotted tensile strength or breaking force is frequently lower than for the unknotted suture. Strength values are obtained in either dry or wet conditions. The ASTM testing conditions that stipulate the cross-head speed, gage length, temperature ($21 \pm 1^\circ\text{C}$), and humidity ($65 \pm 2\%$) of testing room should be followed, if possible, for obtaining and reporting strength data.

Bending stiffness is a complex mechanical phenomenon and it closely relates to the handling characteristic of suture materials, particularly knot security. There are very few reported data describing bending stiffness of sutures. Most reported stiffness of sutures was derived from the modulus of elasticity obtained from a tensile strength test. Because a knot involves bending of suture strands, stiffness based on modulus of elasticity may not adequately represent the performance of knot strength, security, and tie-down. There are two reported studies of bending stiffness of sutures.^[23,24]

Chu et al.'s study^[23] of bending stiffness was based on the force required to bend a suture to a predetermined angle. The measured bending force was converted to flexural stiffness in pounds/in² according to an ASTM formula. Braided sutures were generally more flexible than monofilament sutures of equivalent size, irrespective of their chemical constituents. Coated

Table 5 Mechanical Properties of Absorbable Sutures^a

Class (chemical name)	Commercial name	Break strength straight pull (MPa)	Break strength knot pull (MPa)	Elongation to break (%)	Young's modulus (GPa) (psi) ^b
Catgut		310–380	110–210	15–35	2.4 (358,000)
Regenerated collagen					
Poly (<i>p</i> -dioxanone)	PDS [®] , PDSII [®]	450–560	240–340	30–38	1.2–1.7 (211,000)
Poly(glycolide-co-trimethylene carbonate)	Maxon [®]	540–610 570–910	280–480 300–400	26–38 18–25	3.0–3.4 (380,000) 7–14
Poly(glycolide-co-lactide) or Polyglactin 910	Vicryl [®]				
Polyglycolide-co- ϵ -caprolactone or Polyglecaprone 25	Monocryl [®] Dexon S [®]	654–882 760–920		67–96 18–25	(113,000) 7–14
Poly(glycolic acid) or Polyglycolide	Dexon Plus [®]				

^aMechanical properties presented are typical for sizes 0 through 3–0 but may differ for finer or larger sizes. (*Partial source*: Ref. [14].)

^bData in () are in psi units. (*Source*: Ref. [21].)

sutures had a significantly higher bending stiffness than the corresponding uncoated ones. This increase in bending stiffness is attributable to the loss of mobility of constituent fibers under bending force. An increase in suture size significantly increased their stiffness, and the magnitude of increase depended on the chemical constituent of the suture. The large porous volume inherent in the Gore-Tex monofilament suture was the reason for its lowest flexural stiffness. The second bending stiffness study of a few sutures^[24] was based on attaching a constant weight to each of the two ends of a bending suture and measuring the distance between these two ends after one minute of loading. Bending stiffness data from these two studies agree with each other that braided sutures are generally more flexible than monofilamentous sutures, and that the Gore-Tex suture has the lowest bending stiffness.

Suture compliance (also referred to as elasticity) is a mechanical property that closely relates to the ease of a suture to elongate under a tensile force. It is believed that the level of suture compliance should contribute to the compliance of tissues at the anastomotic site. Suture compliance is particularly important in surgery where there is a tubular or vascular anastomosis. Compliance mismatch between a vascular graft and host tissue has long been suggested as one of several factors contributing to graft failure.^[25] Compliance mismatch at the anastomotic site constitutes a major component of overall compliance mismatch associated with vascular grafts. Because sutures are the only foreign materials in the anastomotic site, it is expected that a wide range of suture compliance might result in

different levels of anastomotic compliance. There is only one reported study that examined the effect of suture compliance on the compliance of arterial anastomotic tissues closed with two sutures vastly different in compliance: 6/0 Novafil and Prolene.^[26] Novafil is an elastomeric suture made from polybutester and is characterized by a high elongation at low tensile force, low modulus of elasticity, and high hysteresis, whereas the Prolene suture has relatively higher modulus of elasticity, low elongation at low tensile force, and low hysteresis. In a clinical condition of minimal tubular compliance and diameter mismatch such as artery-artery anastomoses, a far more compliant anastomosis was achieved with Novafil ($5.9 \pm 2.0\%$) than with Prolene ($3.3 \pm 0.6\%$) suture. Thus, arterial anastomoses closed with a more compliant suture such as Novafil produced on average over 75% more arterial anastomotic compliance than those closed with the less compliant Prolene suture.

Handling Properties

Handling properties describe those that relate to the feel of suture materials by surgeons during wound closure. It is the only category of suture property that is the most difficult to evaluate objectively. Handling properties include pliability (or stiffness), ease of knot tie-down, knot security, packaging memory, surface friction, viscoelasticity, tissue drag, etc. They are directly and indirectly related to physical/mechanical properties of a suture. For example, the term pliability

Table 6 Mechanical properties of nonabsorbable sutures^a

Class (chemical name)	Commercial name	Break strength straight pull (MPa)	Break strength knot pull (MPa)	Elongation to break (%)	Young's modulus (GPa)
Cotton & linen	Cotton	280–390	160–320	3–6	5.6–10.9
Silk	Silk	370–570	240–290	9–31	8.4–12.9
	Surgical Silk				
	Dermal [®]				
	Virgin Silk				
Polypropylene	Surgilene [®]	410–460	280–320	24–62	2.2–6.9
	Prolene [®]				
Nylon 66 & Nylon 6	Surgilon [®]	460–710	300–330	17–65	1.8–4.5
	Dermalon [®]				
	Nurolon [®]				
	Ethilon [®]				
	Supramid [®]				
Poly[(tetramethylene ether)terephthalate-co-tetramethylene terephthalate]	Novafil [®]	480–550	290–370	29–38	1.9–2.1
Poly(butylene terephthalate)	Miralene [®]	490–550	280–400	19–22	3.6–3.7
Poly(ethylene terephthalate)	Dacron [®] Ethiflex [®]	510–1,060	300–390	8–42	1.2–6.5
	Ti.Cron [®]				
	Polydek [®]				
	Ethibond [®]				
	Tevdek [®]				
	Mersilene [®]				
	Mirafil [®]				
Stainless steel	Flexon [®]	540–780	420–710	29–65	200
	Stainless Steel				
	Surgical S.S.				

^aMechanical properties presented are typical for sizes 0 through 3–0 but may differ for finer or larger sizes. *Source:* Ref. [14].

of a suture is a subjective description of how easily a person can bend it and hence relates to the surgeon's feel of a suture during knot tying. It is directly related to the bending modulus of a suture and indirectly to the coefficient of friction. Packaging memory, another

handling property that indirectly relates to pliability, is the ability to retain the kink form of sutures after unpacking them. The ability to retain such kink form after unpacking would make surgeons' handling of sutures more difficult during wound closure, particularly tying a knot. This is because sutures with high memory such as nylon, polypropylene, PDS, and Maxon tend to loosen their knots as they try to return to their kink form from packaging. Thus, packaging memory should be as low as possible. In general, monofilament sutures have more packaging memory than braided ones. The three exceptions are the newly available Monocryl, Biosyn, and Gore-Tex monofilament sutures which were reported to have exceptionally low packaging memory. The easiest means to evaluate packaging memory of sutures is to hang them in air and measure the time required to straighten out the kink form from packaging.

Table 7 Relative tissue reactivity to sutures

Tissue reactivity level	Nonabsorbable	Absorbable
Most →		Catgut
	Silk, cotton	
	Polyester coated	Dexon and Vicryl
	Polyester uncoated	Maxon, PDS, Monocryl
Least →	Nylon	
	Polypropylene	

Source: Ref. [1].

Table 8 General comparison of absorbable sutures

Trade name	Configuration	Tensile		Tissue reactivity	Handling	Knot security	Memory	Absorption	Degradation mode	Comments
		strength	strength							
Collagen (plain)	Twisted	Poor (0% at 2-3 wks)		Moderate	Fair	Poor	Low	Unpredictable (12 weeks)	Proteolytic	Less impure than surgical gut
Collagen (chromic)	Twisted	Poor (0% at 2-3 wks)		Moderate	Fair	Poor	Low	Unpredictable (12 weeks)	Proteolytic	Less impure than surgical gut
Surgical gut (plain)	Twisted	Poor (0% at 2-3 wks)		High	Fair	Poor	Low	Unpredictable (12 weeks)	Proteolytic	May be ordered as "fast-absorbing gut" (Ethicon) for percutaneous sutures
Surgical gut (chromic)	Twisted	Poor (0% at 2-3 wks)		Moderately high	Fair	Fair	Low	Unpredictable (14-80 days)	Proteolytic	Darker, more visible (Davis & Geck); mild or extra chromatization (Davis & Geck)
Coated vicryl	Braided	Good (50% at 2-3 wks)		Low	Better	Fair	Low	Predictable (80 days)	Hydrolytic	Clear, violet, coated
Dexon "S"	Braided	Good (50% at 2-3 wks)		Low	Fair	Good	Low	Predictable (90 days)	Hydrolytic	Uncoated
Dexon plus	Braided	Good (50% at 2-3 wks)		Low	Better	Fair	Low	Predictable (90 days)	Hydrolytic	Clear, green, coated
PDS	Monofilament	Better (50% at 5-6 wks)		Low	Fair	Poor	High	Predictable (180 days)	Hydrolytic	Violet, clear
Maxon	Monofilament	Better (50% at 4-5 wks)		Low	Fair	Fair	High	Predictable (210 days)	Hydrolytic	Green, clear
Biosyn	Monofilament	Good (5% at 2-3 weeks)		Low	Good		Low	Predictable	Hydrolytic	Second lowest in modulus of elasticity
Monocryl	Monofilament	Good (50% at 1-2 wks)		Low	Good	Fair	Low	Predictable (119 days)	Hydrolytic	Gold

(Partial source: Refs. [1] and [22].)

Table 9 General comparison of nonabsorbable sutures

Generic or trade name	Configuration	Tensile strength	Tissue reactivity	Handling	Knot security	Memory	Comments
Cotton	Twisted	Good	High	Good	Good	Poor	Obsolete
Silk	Braid	Good	High	Good	Good	Poor	Predisposes to infection; does not tear tissue; D & G suture is silicone-treated; Ethicon is waxed
Ethilon	Monofilament	High	Low	Poor	Poor	High	Cuts tissue; nylon 6.6; black, clear, or green
Dermalon	Monofilament	High	Low	Poor	Poor	High	Nylon 6.6
Surgamid	Monofilament or braid	High	Low	Poor	Poor	High	Nylon 6.6
Nurolon	Braid	High	Moderate	Good	Fair	Fair	May predispose to infection; black or white; waxed; nylon 6.6
Surgilon	Braid	High	Moderate	Fair	Fair	Fair	Nylon 6.6
Prolene	Monofilament	Fair	Low	Poor	Poor	High	Very low coefficient of friction; cuts tissue; blue or clear
Surgilene	Monofilament	Fair	Low	Poor	Poor	High	–
Dermalene	Monofilament	Good	Low	Poor	Poor	High	–
Novafil	Monofilament	High	Low	Fair	Poor	Low	Blue or clear
Mersilene	Braid	High	Moderate	Good	Good	Fair	Green or white
Dacron	Braid	High	Moderate	Good	Good	Fair	–
Polyviolene	Braid	High	Moderate	Good	Good	Fair	Green or white
Ethibond	Braid	High	Moderate	Good	Good	Fair	Green or white
Tri-Cron	Braid	High	Moderate	Poor	Poor	Fair	–
Polydek	Braid	High	Moderate	Good	Good	Fair	–
Tevdek	Braid	High	Moderate	Poor	Poor	Fair	–
Fore-Tex	Monofilament	Fair–Poor	Low	Very good	Fair	Poor	No packing memory and knot construction has virtually no effect on breaking strength
Stainless steel	Monofilament, twist, or braid	High	Low	Poor	Good	Poor	May kink

(Data partially from Ref. [1].)

Knot tie-down and security describe how easily a surgeon can slide a knot down to the wound edge and how well the knot will stay in position without untying or slippage. This handling property relates to surface and mechanical properties of sutures. The relatively smooth surface of a monofilament or coated braided

suture will have a better knot tie-down than a suture with a rough surface such as an uncoated braided suture, if everything else is equal. The coefficient of friction of sutures also relates to knot tie-down and security. A linear relationship between knot security and coefficient of friction was reported by Herman.^[27]

A high coefficient of friction makes knot tie-down difficult but leads to a more secure knot. This is because a high friction suture provides additional frictional force to hold the knot together. This high friction suture surface also makes the passage of suture strands difficult during knot tie-down. It thus appears that knot tie-down and knot security are two contradictory requirements. There is no reported standard test for evaluating knot tie-down capacity. However, Tomita et al. recently reported a method to objectively quantify the knot tie-down capacity of 2/0 silk, polyester sutures, Gore-Tex, and an experimental ultrahigh-molecular-weight polyethylene suture (Nesplon®).^[24]

Biological Properties

Biocompatibility of suture materials describes how sutures, which are foreign materials to the body, can affect surrounding tissues and how the surrounding tissues can affect the properties of sutures. Thus, biocompatibility is a two-way relationship. The extent of tissue reactions to sutures depends largely on the chemical nature of the sutures and their degradation products if they are absorbable. Sutures from natural sources such as catgut and silk usually provoke more tissue reactions than synthetic ones because of the availability of enzymes to react with natural biopolymers. Besides the more important chemical factors and physical form, the amount and stiffness of suture materials have been reported to elicit different levels of tissue reactions. For example, a stiff suture will result in stiff projecting ends in a knot where cut. These stiff ends can irritate surrounding tissues through mechanical means, a problem associated with some monofilament sutures but generally not found in braided multifilament sutures.

Because the quantity of a buried suture relates to the extent of tissue reaction, it is a well-known practice in surgery to use as little suture material as possible, such as a smaller knot or a smaller size, to close wounds. The use of a smaller size of suture for wound closure, without detriment to the provision of adequate support to wounds and without cutting through wound tissue, is due to the square relationship between diameter and volume which suggests that a slight increase in suture size or diameter would increase its volume considerably.

There are two basic means to study biocompatibility of suture materials: cellular response and enzyme histochemistry. The former is the most frequently used and provides information about the type and density of inflammatory cells at a suture site. In the cellular response approach, sutures without tension are implanted in the gluteal muscle in small animals such as the rat. This implantation site has given a very consistent reproducible cellular response for valid comparisons, even though it is not a common site for suture in surgery.

However, Walton recently questioned the use of this common test procedure, particularly in orthopedic surgery,^[28] due to the observed inflamed nature of the postoperative synovial tissue and the mechanically stressed nature of the suture. Histological stains with a variety of dyes such as the most frequently used H & E are the standard methods of evaluation of cellular activity at the suture sites. Figure 2 is a typical example of histological photomicrographs of PDS and Maxon sutures at 35 days postimplantation in a variety of tissues.^[29] In addition to a qualitative description of cellular activities, tissue response can also be graded by the most frequently used and accepted Sewell et. al. method or its modification.^[30]

The enzyme histochemical approach is a more objective, quantitative, consistent, and reproducible method than cellular response. Enzyme histochemistry is based on the fact that any cellular response to a foreign material is always associated with the presence of a variety

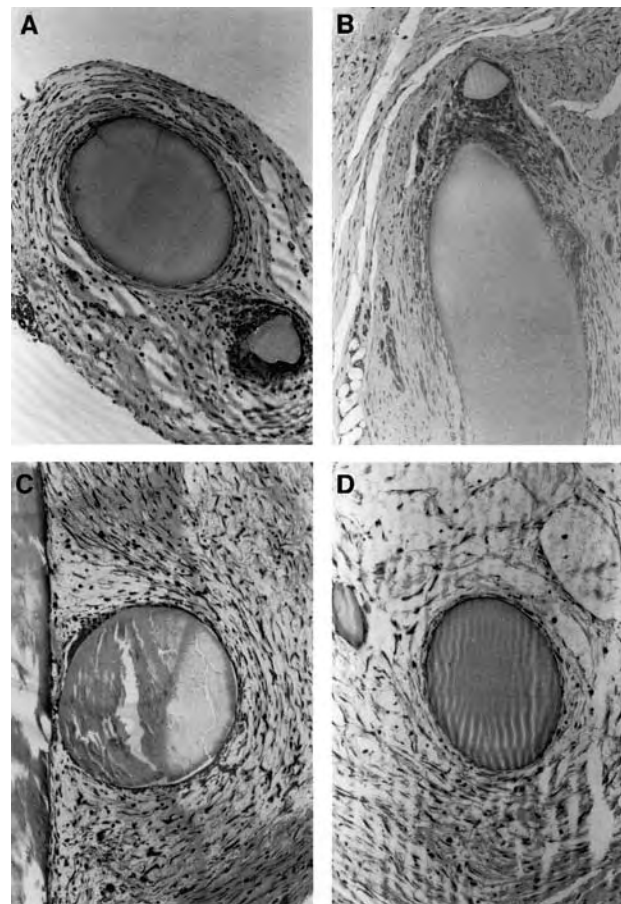


Fig. 2 Light histologic photomicrographs of tissue adjacent to PDS and Maxon sutures at 3 and 5 days postimplantation in a variety of tissues of New Zealand White rabbit ($\times 130$). (A) PDS in peritoneum; (B) PDS in fascia; (C) Maxon in peritoneum; (D) Maxon in fascia. (From Ref. [29].)

of enzymes; however, this approach is more tedious and requires more sophisticated facilities and better experience. The data obtained provide additional insight into the functions of those cells appearing during various stages of wound healing. The enzymatic activity of a suture implant site is quantified by microscopic photometry of a cryostat section of the tissue. For example, van Winkle et al. reported that the high level of cellular response to silk suture observed from histological study is confirmed in their enzyme histochemical study.^[31–33] Enzyme histochemistry is also useful for studying the biodegradation mechanism of absorbable sutures.

The normal tissue reaction to sutures can be grouped into three stages, according to the time for the appearance of a variety of inflammatory cells.^[1,31,33,34] They are: initial infiltration of polymorphonuclear leukocytes, lymphocytes and monocytes during the first three to four days (i.e., acute response); appearance of macrophages and fibroblasts from day 4 to day 7; and beginning maturation of fibrous connective tissue formation with chronic inflammation after the seventh to the tenth day. During the first seven days post-implantation, there is virtually no difference in normal tissue reaction between synthetic absorbable and non-absorbable sutures. However, a slightly higher inflammatory reaction to synthetic absorbable sutures can persist for an extended period until they are completely absorbed and metabolized, whereas synthetic non-absorbable sutures, in general, are characterized by a minimal chronic inflammatory reaction with a thin fibrous connective tissue capsule surrounding the sutures usually by 28 days postimplantation.

In addition to the normal tissue reactions to sutures, there are several adverse tissue reactions that are suture- and site-specific. Some examples include urinary stone or calculi formation, granuloma formation, thrombogenicity, propensity toward wound infection, and recurrence of tumor after radical surgery and allergy.

Monofilament sutures are considered to be a better choice than multifilament ones in closing contaminated wounds. This is because not only do multifilament sutures elicit more tissue reactions which may lessen

tissue ability to deal with wound infections but also multifilament sutures have a capillary effect that could transport microorganisms from one region of the wound to another. The reason that multifilament sutures generally elicit more tissue reactions than their monofilament counterparts is because inflammatory cells are able to penetrate into the interstitial space within a multifilament suture and invade each filament. Such an invasion by inflammatory cells, well evident in histological pictures, cannot occur in monofilament sutures. Thus, the available surface area of a suture to tissue should bear a close relationship to the level of tissue reaction that a suture could elicit.

S

Biodegradation and Absorption Properties

Biodegradation and absorption properties are the most important issue of absorbable sutures, but are far less relevant for most nonabsorbable sutures, particularly those of synthetic nature. This biodegradation property is also responsible for the fact that absorbable sutures do not elicit permanent chronic inflammatory reactions found with nonabsorbable sutures. The most important characteristics in biodegradation and absorption of sutures are the strength and mass loss profiles and biocompatibility of degradation products. Although there is a wide range of strength and mass loss profiles among the available absorbable sutures, they have one common characteristic: Strength loss always occurs much earlier than mass loss, as shown in Table 10. This suggests that absorbable sutures retain a large portion of their mass in tissue while they have already lost all their mechanical properties required to provide support for tissues during wound healing. Thus, an ideal absorbable suture should have matched mass loss and strength loss profiles, and none of the current commercial absorbable sutures can achieve this ideal biodegradation property.

The observed wide range of strength and mass loss profiles among the available absorbable sutures is attributable not only to the chemical differences among

Table 10 Absorption delay of commercial synthetic absorbable sutures

Suture materials	Time to complete loss of tensile strength (days)	Time to complete mass absorption (days)	Absorption delay (days)	Useful lifetime (%) ^a
Dexon [®]	28	50–140	22–112	20–56
Vicryl [®]	28	90	62	31
PDS [®]	63	180–240	117–170	26–35
Maxon [®]	56	210	155	27
Monocryl [®]	21	90–119	69–98	18–23

The higher percentage denotes the better absorbable suture.

^aThe ratio of [the time to complete loss of tensile strength] to [the time to complete mass absorption].

the absorbable sutures but also to a variety of intrinsic and extrinsic factors, such as pH, electrolytes, stress applied, temperature, γ irradiation, superoxide, microorganisms, and tissue type, to name a few. Among these intrinsic and extrinsic factors, the role of superoxide on the biodegradation of absorbable sutures appears to be one of the most interesting factors because of the unusually fast loss of mechanical integrity and unique surface morphology.^[35] For example, at a 0.005 M superoxide ion concentration and room temperature, the five synthetic absorbable sutures retained 20–70% of their original tensile breaking force at the end of 24 hrs as shown in Fig. 3. The bulk of the loss of tensile breaking force of these sutures occurred during the initial two-hour period. The order of tensile breaking force of these five absorbable suture materials at this relatively high superoxide ion concentration was the same as at the lower superoxide ion concentration: Monocryl[®] > Maxon[®] > Vicryl[®] > Dexon[®] > PDS II[®]. It is important to know that there would be no change in tensile breaking force of these absorbable sutures in regular buffered saline medium at 25°C for as long as days.^[9]

Upon biodegradation, absorbable sutures have shown quite interesting surface morphology and some examples are shown in Fig. 4. For example, multifilament Dexon sutures that were subjected to γ -irradiation treatment and hydrolytic degradation in buffer solution showed very regular circumferential surface cracks along the longitudinal fiber axis and had

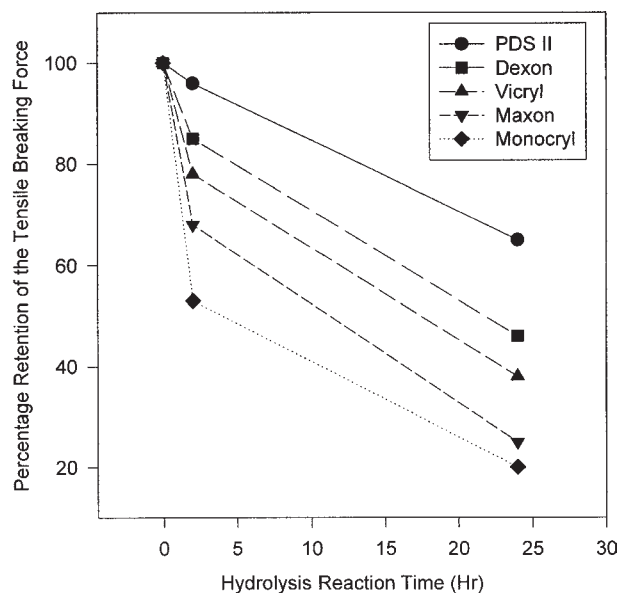


Fig. 3 The percentage of retention of tensile breaking force of five 2/0 synthetic absorbable sutures upon 0.005 M superoxide ion-induced hydrolytic degradation at 25°C. ● - PDS; ■ - Dexon; ◆ - Monocryl; ▲ - Vicryl; ▼ - Maxon. (From Ref. [35].)

the appearance of corn-like structure (Figs. 4A and B). Upon γ -irradiation treatment and hydrolytic degradation in buffer solution, monofilament Maxon sutures, however, showed both circumferential and longitudinal surface cracks (Fig. 4C) and subsequent peeling off of these surface cracks (Fig. 4D). The appearance of moon-crater-shaped impressions of various sizes (about 10–100 μ m diameter) on a Monocryl[®] suture at a superoxide ion concentration of > 0.005 M (Fig. 4E) is unique because such circular impressions were never observed in the hydrolytic degradation of all existing absorbable sutures in a conventional buffered saline medium or in vivo. The formation of moon crater-shaped impressions on Monocryl and Maxon sutures deviates from the conventional understanding of the anisotropic character of fibers. In the reported morphological studies of all existing absorbable sutures in conventional buffer media,^[9] the most common surface morphological characteristic upon hydrolytic degradation of suture fibers is the formation of circumferential or/and longitudinal surface cracks that are consistent with the anisotropic character of fibers. It is not fully understood at this stage how superoxide ion-induced degradation could lead to such unusual surface morphology on Monocryl and Maxon sutures.

The biocompatibility of degradation products is usually not a problem because all existing absorbable sutures are made from well-known biocompatible glycolide, lactide, and their derivatives. However, biocompatibility of degradation products also depends on the rate of their accumulation in the surrounding tissues. This implies that the ability of the surrounding tissues to actively remove and metabolize degradation products is essential. Such a metabolism depends on the extent of blood circulation in the tissue. A well-vascularized tissue can remove degradation products as fast as they are released from an absorbable suture and subsequently metabolized, which can minimize tissue reactions to degradation products.

Due to their ability to release degradation products, absorbable sutures have recently been studied as a vehicle to deliver a variety of biochemicals such as growth factors to facilitate wound healing or antibiotics to combat wound infection. This new approach will increase the value of absorbable sutures and extend their function beyond the traditional role of wound closure. A typical example is Vicryl Plus which has an antimicrobial agent coating.^[10,11]

Biodegradation properties are usually examined in vitro or/and in vivo. In the in vitro environment, the most commonly used medium is phosphate-buffered physiological saline, pH 7.44 at 37°C. However, other buffers such as Tris or body fluids such as urine, bile, and synovial fluids have been used. In the in vivo environment, unstressed absorbable sutures are normally

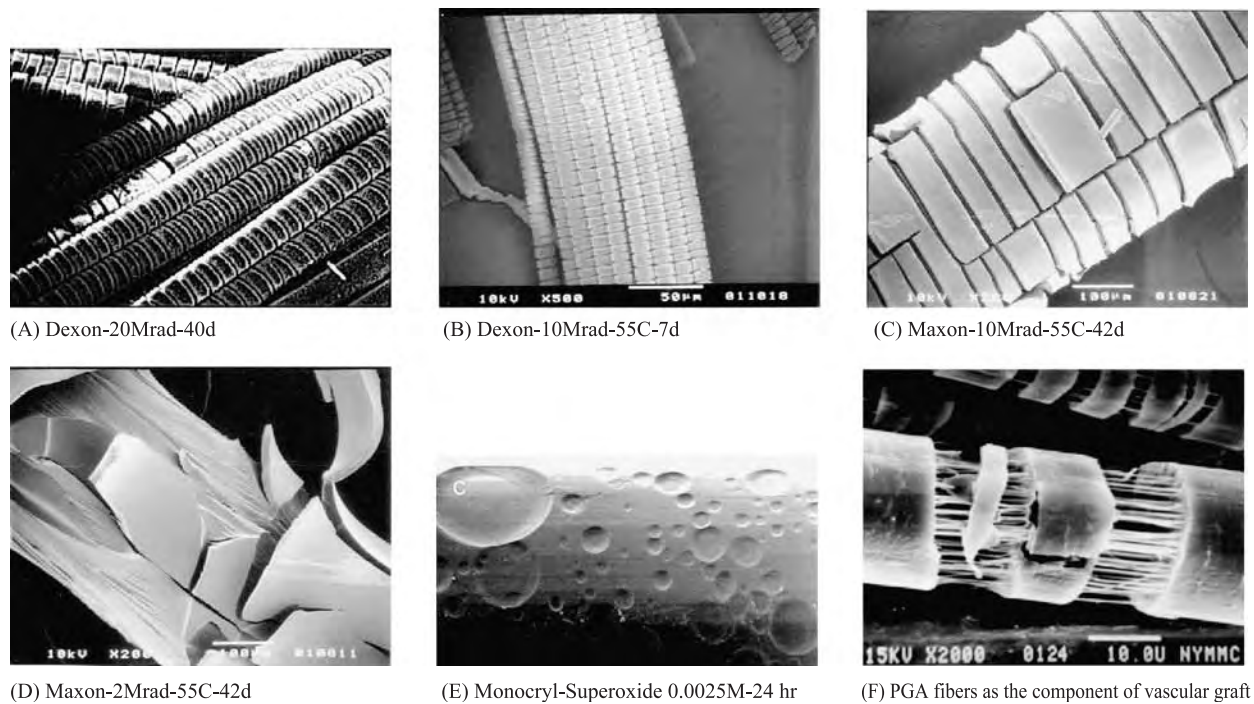


Fig. 4 Scanning electron images of degradation of some commercial absorbable sutures. (A) 2/0 Dexon after 20 Mrad γ irradiation at room temperature and 40 days in vitro in buffer at 37°C; (B) 2/0 Dexon after 10 Mrad γ irradiation at 55°C and 7 days in vitro in buffer at 37°C; (C) 2/0 Maxon after 10 Mrad γ irradiation at 55°C and 42 days in vitro in buffer at 37°C; (D) 2/0 Maxon after 2 Mrad γ irradiation at 55°C and 42 days in vitro in buffer at 37°C; (E) 2/0 Monocryl upon 0.005 M superoxide ion-induced hydrolytic degradation at 25°C; (F) Polyglycolide fibers as the component of woven vascular grafts upon in vitro hydrolytic degradation in buffer of pH 7.4 at 37°C.

implanted in rat gluteal muscle for predetermined periods of implantation. The sutures retrieved at various periods of immersion or implantation are then subjected to evaluation of their mechanical and physical properties to assess their changes with time. The degree of absorption in vivo is evaluated by the change in suture cross-sectional area, and the level of tissue reaction is assessed by a histological method and/or enzyme histochemistry.

SUMMARY

Suture materials are the most frequently used biomaterials because every wound must be closed. Suture materials are made from either natural or synthetic fibers and there is a wide range of properties of suture materials available for surgeons to choose from. Among suture properties, four properties are essential: physical and mechanical, handling, biological, and biodegradative. These properties are interrelated. The trend of current research development of suture materials is toward value-added products, such as the

incorporation of antimicrobial agents into sutures to ward off wound infection.

ARTICLES OF FURTHER INTEREST

Collagen; Glues; Orthopedic Fixation Devices; Poly(Glycolic Acid); Poly(lactic acid)s; Silks

REFERENCES

1. Bennett, R.G. Selection of wound closure materials. *J. Am. Acad. Dermatol.* **1988**, *18* (4), 619–637.
2. Guttman, B.; Guttman, H. Sutures: Properties, Uses, and Clinical Investigation. In *Polymeric Biomaterials*; Dumitriu, S., Ed.; Marcel Dekker: New York, 1994; Chapter 10 page range.
3. Stone, I.K. Suture materials. *Clin. Obstet. Gynecol.* **1988**, *31* (3), 712–717.
4. Chu, C.C. Suture Materials. In *Encyclopedia of Materials Science and Engineering*; Beaver, M.B., Ed.; Pergamon Press: New York, 1986; Vol. 6, 4826–4832.
5. Chu, C.C. The Degradation and Biocompatibility of Suture Materials. In *CRC Critical Reviews in*

- Biocompatibility*; Williams, D.F., Ed.; CRC Press: Boca Raton, FL, 1985; Vol. 1, 261–322.
6. Chu, C.C. Survey of Clinically Important Wound Closure Biomaterials. In *Biocompatible Polymers, Metals, and Composites*; Szycher, M., Ed.; Society for Plastics, Engineers, Technomic: Westport, CT, 1983; 477–523. Chapter 22.
 7. Chu, C.C. Biodegradable Suture Materials: Intrinsic and Extrinsic Factors Affecting Biodegradation Phenomena. In *Encyclopedic Handbook of Biomaterials and Bioengineering*; Wise, D.L., Altobelli, D.E., Schwartz, E.R., Yszemski, M., Gresser, J.D., Trantolo, D.J., Eds.; Marcel Dekker: New York, 1995; Vol. 1, 543–688.
 8. Chu, C.C. Degradation and Biocompatibility of Synthetic Absorbable Suturematerial: General Biodegradation Phenomena and Some Factors Affecting Biodegradation. In *Biomedical Applications of Synthetic Biodegradable Polymers*; Hollinger, J., Ed.; CRC Press: Boca Raton, FL, 1995; 103–128. Chapter 5
 9. Chu, C.C.; von Fraunhofer, J.A.; Greisler, H.P. *Wound Closure Biomaterials and Devices*; CRC Press: Boca Raton, FL, 1997.
 10. Storch, M.; Perry, L.C.; Davison, J.M.; Ward, J.J. A 28-day study of the effect of coated Vicryl Plus antibacterial suture (coated Polyglactin 910 suture with triclosan) on wound healing in Guinea pig linear incisional skin wounds. *Surg. Infect.* **2002**, *3* (Supplement), 89–98.
 11. Rothenburger, S.; Spangler, D.; Bhende, S.; Burkley, D. In vitro antimicrobial evaluation of coated Vicryl Plus antibacterial suture (coated Polyglactin 910 with triclosan) using zone of inhibition assays. *Surg. Infect.* **2002**, *3* (Supplement), 79–87.
 12. Conn, J., Jr.; Beal, J.M. Coated Vicryl synthetic absorbable sutures. *Surg. Gynecol. Obstet.* **1980**, *150* (6), 843–844.
 13. Mattei, F.V. Absorbable Coating Composition for Sutures. US Patent 4,201,216, 1980.
 14. Casey, D.J.; Lewis, O.G. Absorbable and Nonabsorbable Sutures. In *Handbook of Biomaterials Evaluation: Scientific, Technical, and Clinical Testing of Implant Materials*; von Recum, A.F., Ed.; Macmillan Publishing: New York, 1986; 86–94. Chapter 7
 15. Rodeheaver, G.T.; Thacker, J.G.; Owen, J.; et al. Knotting and handling characteristics of coated synthetic absorbable sutures. *J. Surg. Res.* **1983**, *35*, 525–530.
 16. Rodeheaver, G.T.; Thacker, J.G.; Delich, R.F. Mechanical performance of polyglycolic acid and polyglactin 910 synthetic absorbable sutures. *Surg. Gynecol. Obstet.* **1981**, *153*, 835–841.
 17. Kawai, T.; Matsuda, T.; Yoshimoto, M. Coated Sutures Exhibiting Improved Knot Security. US Patent 4,983,180, 1991.
 18. Messier, K.A.; Rhum, J.D. Caprolactone Polymers for Suture Coating. US Patent 4,624,256, 1986.
 19. Bezwada, R.S.; Hunter, A.W.; Shalaby, S.W. Copolymers of ϵ -caprolactone, glycolide and glycolic acid for suture coatings. US Patent 4,994,074, 1991.
 20. Wang, D.W.; Casey, D.J.; Lehmann, L.T. Surgical Suture Coating. US Patent 4,705,820, 1987.
 21. Bezwada, R.S.; et al. Monocryl, a new ultra-pliable absorbable monofilament suture derived from ϵ -caprolactone and glycolide. **1994**.
 22. Roby, M.S.; et al. US Patent, 5,403,347, April 4, 1995.
 23. Chu, C.C.; Kizil, Z. Quantitative evaluation of stiffness of commercial suture materials. *Surg. Gynecol. Obstet.* **1989**, *168* (3), 233–238.
 24. Tomita, N.; Tamai, S.; Morihara, T.; Ikeuchi, K.; Ikada, Y. Handling characteristics of braided suture materials for tight tying. *J. Appl. Biomater.* **1993**, *4* (1), 61–65.
 25. Abbott, W.M.; Megerman, J.; Hasson, J.E.; L'Italien, G.; Warnock, D. Effect of compliance mismatch upon vascular graft patency. *J. Vasc. Surg.* **1987**, *5* (2), 376–382.
 26. Megerman, J.; Hamilton, G.; Schmitz-Rixen, T.; Abbott, W.M. Compliance of vascular anastomoses with polybutester and polypropylene sutures. *J. Vasc. Surg.* **1993**, *18* (5), 827–834.
 27. Herman, J.B. Tensile strength and knot security of surgical suture materials. *Am. Surg.* **1971**, *37*, 209–217.
 28. Walton, M. Strength retention of chromic gut and monofilament synthetic absorbable suture materials in joint tissues. *Clin. Orthop. Relat. Res.* **1989** (242), 303–310.
 29. Metz, S.A.; Chegini, N.; Masterson, B.J. In vivo tissue reactivity and degradation of suture materials: A comparison of Maxon and PDS. *J. Gynecol. Surg.* **1989**, *5*, 37–46.
 30. Sewell, W.R.; Wiland, J.; Craver, B.N. A new method of comparing sutures of bovine catgut with sutures of bovine catgut in three species. *Surg. Gynecol. Obstet.* **1955**, *100*, 483–494.
 31. Van Winkle, W.; Salthouse, T.N. *Biological Response to Sutures and Principles of Suture Selection*; Ethicon: Somerville, NJ, 1976; 1–20.
 32. Salthouse, T.N.; Matlaga, B.F. Significance of cellular enzyme activity at nonabsorbable suture implant sites: Silk, polyester and polypropylene. *J. Surg. Res.* **1975**, *19*, 127–132.
 33. Salthouse, T.N. Biocompatibility of Sutures. In *Biocompatibility in Clinical Practice*; Williams, D.F., Ed.; CRC Press: Boca Raton, FL, 1982; Vol. 1, 12–32.
 34. Madsen, E.T. An experimental and clinical evaluation of surgical suture materials, I and II. *Surg. Gynecol. Obstet.* **1953**, *97*, 73–80.
 35. Lee, K.-H.; Chu, C.C. The effect of superoxide ions in the degradation of five synthetic absorbable suture materials. *J. Biomed. Mater. Res.* **2000**, *49* (1), 25–35.

Telemedicine

Brett M. Harnett

Virginia Commonwealth University, Richmond, Virginia, and University of Cincinnati, Cincinnati, Ohio, U.S.A.

Charles R. Doarn

Virginia Commonwealth University, Richmond, Virginia, and University of Cincinnati, Cincinnati, Ohio, U.S.A.

Ronald C. Merrell

Virginia Commonwealth University, Richmond, Virginia, U.S.A.

T

INTRODUCTION

Telemedicine or telehealth is the integration of telecommunications and information systems to support the delivery of healthcare, regardless of the patient's location. The use of sensors to collect physiologic signs and transmission of these data were hallmarks of the space program from the time of Sputnik 2 in 1957. This telemetry in space science led to critical care telemetry in standard medicine and the possibility of telemedicine. Although satellite telecommunications prompted telemedicine's genesis in the 1970s (STAR-PAHC),^[1] videoconference consultation by microwave transmission made its debut in 1955 with a psychiatric program in Nebraska.^[2] Telemedicine was also used via radio from Antarctica to Australia in the 1920s.

It is safe to state that the medical industry is changing. The advent of affordable computer technology and ubiquitous telecommunications topologies has created a new paradigm in healthcare capabilities, cost models, and challenges. The prospect of telemedicine holds the opportunity of increasing the quality and quantity of healthcare for millions of people by eliminating the barriers of distance and time.

Telemedicine need not, nor should not be, a radical change from traditional medicine. Familiar environments, existing protocols, and standard tools need to be integrated into clinical practice so clinicians and patients are encouraged with little inconvenience. The implementation of telemedicine is finding its way into the mainstream of the medical community.^[3]

OVERVIEW

To date, hundreds of funded telemedicine programs have proven to be successful in terms of increasing the quality of care. However the cost structure to support long-term and self-sustainment is still in debate.^[4-6] Studies in the area of prison telemedicine have largely

proved to be financially feasible while other areas have not.^[7,8] Ironically, the slow adaptation of telehealth in the mainstream industry is not due to the technology but rather to barriers of policies, reimbursement, and legal issues.^[9,10] Until these issues are resolved, the progression and acceptance is expected to be slow.

As clinical expertise is ferried across state lines and international borders, the need for electronic medical records becomes more apparent. With the data collection and distribution capabilities now available, the aptitude of informatics becomes plausible and a potentially powerful tool for telemedicine providing comprehensive electronic information about patients. There are many acronyms associated with names being debated, such as the Electronic Medical Record (EMR), Electronic Health Record (EHR), Computerized Patient Record (CPR), and others. The names all mean basically the same thing; although, some may imply a broader capability to include practice management capabilities such as billing. The terms stem from a variety of sources, but the Institute of Medicine is likely to have the most influence on the eventual official name—if there ever is one. The challenge remains on how the EMR will look and feel.

CLINICAL APPLICATION

Early telemedicine matured as a video-teleconferencing (VTC) through a Massachusetts General Hospital program to Logan International Airport Medical Station in 1967 and the remarkable work on Space Lab in 1983.^[11-13] The application of telemedicine has broadened from space medicine to maritime program, rural program, and support after international disasters.^[14] Telemedicine can be applied in a wide variety of medical disciplines. The technology is an enabling factor for improved clinical care.^[15]

The elements of telemedicine remain the same. For example, there must be a competent capture of data by

sensors or devices at the site of the patient. Sensors or devices may include video, audio, physiologic data, laboratory data, text, or radiographic files. Although analog data may be the source, digital readings are now almost universal with some form of compression or perhaps data interpretation at the source. Transmission can be via electromagnetic spectrum or fixed mediums such as copper or fiber optics. Various transmission technologies may interface for long distance communication—from the very slow to the very fast. Each digital phase or byte is represented as eight bits. Therefore, the speed to transfer data is a function of the number of bits measured per second. For example, to transfer a text file of 10,000 characters, which is equivalent to approximately 1500 words, with a cellular connection of 9600 bps will take just over eight seconds. A high-speed network will move it in a split second.

The data must convey an accurate depiction of the situation. The consultant must interpret, manipulate, and formulate the data with confidence and competence in the context of medical professionalism. Thereby, the consultant may ethically render a professional opinion or therapeutic plan for a patient never seen or touched. This leap of professional faith calls for consummate technical success and transparency for the consultant. The consultant response must be an action or series of actions, which can actually be carried out at the distant site. It makes little sense to recommend a drug that is not available or an operation that is beyond the training level or expertise of the practitioner. The recommendation may be in text or require effectors such as robotics. The lag time becomes critical when the consultant must await arrival of images, transmission of robotic instructions, and return of the image as altered by the robotic action. Network jitter and delays over 500 milliseconds for this segment exceeds the hand-eye coordination of surgeons.^[16] Performance of a robotic transatlantic operation by Dr. Jacques Marescaux in 2001 had a lag of only 155 milliseconds and was completely successful. Robotics solutions for Mars travelers with a 22-minute transmission—at the furthest point with our planet—is completely impractical.^[17]

Telemedicine has several components, including sensors, transmitters, responders, and effectors. In order to practice medicine in a system of telemedicine, even seamless communications cannot overcome the challenge. Medical practitioners have been implicated to human touch and uniquely personal patient interaction. In order to practice medicine in the partnership with engineers, technology, and unseen colleagues, there must be process simulators to prepare the team for competent professional performance. This may be simply familiarity with software or extensive training in a common medical language, robotics, and virtual reality. Through training, telesurgery, teleradiology, and telepathology have all seen success and look to greater ones.

TYPES OF TELEMEDICINE

Real Time

Real-time telemedicine can include a number of tools, including VTC, the plain old telephone system (POTS), Web-based chat environments, and wireless options such as Short Message Services (SMS).

Before computers were affordable and the Internet was popularized, the primary mechanism for telemedicine was VTC. While VTC matured in the boardroom of large corporations, it was being identified as a requisite component for telemedicine and for good reason. Here was an ideal solution for a physician and patient separated by great distance to interact. A call from a referring doctor to a consulting doctor using the same type of equipment corporate officers was using was a tremendous boon to the remote consult.^[18] In this modality, consults are performed by a referring physician, i.e., the location of the patient, and a consulting physician usually represents a specialty in support of the patient.

While VTC can be a powerful tool for remote consultation, the costs can be prohibitive.^[19] Dedicated hardware includes a coder-decoder (CODEC), camera, microphone, speaker, and a projection device or large-screen monitor. In addition, acquisition devices are required to provide the consulting site with additional data and images to make a diagnosis and recommendations. CODECS can be both hardware and software.

VTC facilities are not always readily available, and the equipment and technical support can be expensive to operate. VTC uses various communications capabilities such as Integrated Services Digital Network (ISDN). ISDN is a switched circuit topology that is managed by the local telecommunication company (TELCO.) These circuits behave like the traditional POTS, but, instead of routing analog signals, its signals are digital. This separate network requires installation of dedicated lines to a particular location where the CODEC is connected. The standard for videoconferencing over ISDN is H.320.

ISDN lines are available as either Basic Rate Interface (BRI) or Primary Rate Interface (PRI). Each BRI consists of a pair of 64 Kbps circuits providing 128 Kbps plus a D channel that negotiates the connection. BRIs can be configured in 128 Kbps increments. A PRI is a single 1.544 mega/million bits per second (Mbps), which is bundled as 23 BRIs plus the D channel. This channel can be fractionalized to support communications with other sites, depending on what bandwidth is available. For example, if Site A has three BRIs (384 Kbps) and it connects to Site B that has a PRI, the connection is negotiated through the D channel, and the systems agree that the highest speed they can connect is 384 Kbps. When a multipoint

conference is required, i.e., more than two participants, a specialized piece of equipment called a bridge or multipoint conferencing unit (MCU) is required. An important note, ISDN services are not available everywhere; therefore, its use depends on the capabilities at corroborating sites.

This is why significant bandwidth is required for VTC. VTC uses various compression schemes such as H.261 to minimize bandwidth requirements, even on broadband links. For example, images are sent as either 352×288 pixels, which is known as CIF, or 176×144 , which is known as QCIF. Using QCIF in 256-color mode, which requires eight bits per pixel, will require $176 \times 144 \times 8 = 202,752$ bits per frame. At 15 frames per second, this equates to $15 \times 202,752 = 3,041,280$ bits per second, which requires approximately 3 Mbps of bandwidth. Fortunately most CODECs employ other compression techniques, such as only sending the portion of the frame that changes, which is known as Deltas.

How much bandwidth is needed for clinical consultations? This is debatable and depends on the type of service, but 384 Kbps, in most cases, provides the quality for clinical application in real-time sessions.^[20] This amount of bandwidth provides full-motion video and acceptable audio. Because ISDN is billed like a normal telephone call, long distance charges accrue from the originator of the call. Also, there are initial setup fees and monthly costs for ISDN service. ISDN is often referred to as bursty bandwidth because it provides bandwidth only when needed.

By comparison, the Internet can also be a network conduit for videoconferencing. Connection to the Internet is virtually mandatory these days. ISDN is a luxury and must be justified by clinical application and participating partner sites. Since a facility is more likely to have copious bandwidth to the Internet than ISDN, videoconferencing over IP networks, which is referred to as H.323, may be a sensible option but not without challenges. The Internet uses protocols such as Transmission Control Protocol/Internet Protocol (TCP/IP) to manage data packets from one point to another. But, unlike transmission of an E-mail message, for example, which is composed and sent to a mail server to await the recipient's attention, VTC is synchronous. The packets traverse the network and are either delivered or simply expire.

In client/server environments such as the Internet, TCP is primarily responsible for transferring data over the IP network. Native to TCP are various components such as error correction and flow control—packet management. These protocols ensure that data packets arrive at their intended destination. E-mail is a good example. Data are sent in packets from one host to another. The receiving host replies to the sending host that it received the data properly or not. In the event

packets were lost or damaged in transit, the sending host retransmits the subject packets. This process is asynchronous, which means that there is no critical timing or order that the packets must be received. For most asynchronous communications, this is perfectly acceptable and preferred. However, in synchronous communications in which timing is critical such as voice over IP or videoconferencing, packet loss is extremely problematic.

IP-based VTC doesn't use TCP. It uses User Datagram Protocol (UDP) instead, which does not employ error correction. If packets are lost or damaged, they are not resent. That part of a given conversation is over. This results in dropped packets that cause dropped audio and jerky video. Another challenge presented by network topologies and real-time communications is network jitter. Jitter is the term used to describe the inconsistent flow rate of data packets over a defined path. Latency is created due to two reasons, distance and routing mechanisms. Routing mechanisms refer to the various devices that a packet must traverse to reach its destination. As packets pass through routers and other Level 3 devices, which are called hops, delays occur much like traffic at a busy intersection. This latency creates problems for synchronous communications such as image pixelization, in which portions of the image are missing or out of place, as well as audio interruptions—especially with low-bandwidth connections. More bandwidth helps because of lessened congestion, data can be compressed less, and losses that affect critical components, such as audio, are not as noticeable.^[21,22]

Nevertheless, H.323 is highly scalable, affordable, and easy to implement. As bandwidth becomes more available and affordable, H.323 will likely become more embraced.

Store-and-Forward

In many telemedicine consultation scenarios, it may be more effective—in cost, ease, and performance to use store-and-forward techniques. Examples include E-mail, File Transfer Protocol (FTP), Web services, and even the facsimile machine and documents generated by software that converts to the Group 3 fax standard. The Internet has enabled a new dimension for exchange of knowledge. TCP/IP is natively an ideal protocol for the exchange of medical data. Store-and-forward employs the asynchronous advantages of TCP/IP.

Store-and-forward is extremely useful when non-emergent conditions are addressed. Some of the most common store-and-forward applications include radiology, dermatology, pathology, and ophthalmology. The migration to EMRs will significantly increase the



viability of store-and-forward by sharing entire patient histories, current medications, allergies, laboratory values, and digital images.^[23] One area that can be used in both a real-time and store-and-forward mode is a practice called whiteboarding. Whiteboarding is accomplished over the Internet by the T.120 standard. Videoconferencing applications such as NetMeeting[®], ProShare[®], and CU-SeeMe[®] provide this function as an augment to the live video. T.120 allows you to place a selected image or part of image onto the whiteboard where each participant can annotate by using different color highlighting tools, text, or pointers. This is used in real-time modes to show a remote physician where to place a bandage, take a biopsy, or, in the case of surgical telementoring, where to—or not to—make a cut. Images can be saved as various types of formats and saved or imported into a patient's EMR for later review.

Digital images are created by a group of pixels (picture elements). Each pixel represents a single spot of color and value, which is much like the grain on photographic film. When enlarged too much, an image can lose focus because the individual pixels are noticeable. Since an image is made up of a number of these pixels, digital images can be large. This factor is one of the primary constraints in the ability to easily capture, manipulate, and transfer digital images.

However, not all pixels contain the same amount of information. With still images, each pixel is represented by a certain number of bits, and this is referred to as bit-depth. The greater the bit-depth, the more colors and values a pixel can represent. A true-color image is 24 bits, which means that, for each pixel, there are 24 bits of data or 8 bits each for red, green, and blue. A grayscale image, such as an X-ray, has only eight bits per pixel because there is no color data.

It is important to understand the concept of resolution as it applies to digital imaging if acceptable results are desired. Different devices such as scanners, printers, and cameras have different resolutions and different ways of specifying resolution. Higher resolution means more data, but it might also be a waste of space and time. Using too low of a resolution can cause the image to be unusable. There are different types of formats that have been used on the Web for years, such as Joint Photographic Experts Group (JPEG), Graphics Interchange Format (GIF), Portable Network Graphic (PNG), Tag Image File Format (TIFF), Windows Bitmap (BMP), and Digital Imaging and Communications in Medicine (DICOM), which is the standard format endorsed by the American College of Radiology.

NETWORKS AND TELECOMMUNICATIONS

There is a primary difference between traditional circuit-switched networks such as POTS and packet-switched

networks such as the Internet. Circuit-switched means that, for a given network session, a dedicated or virtual circuit is established for the parties involved. This allows for a relatively secure session, and its link is not shared. Conversely, packet-switched networks, which are akin to interstate highways, share the link and are not secure in native format. The advantage of packet-switched networks is the scalability, ease of use, and low cost.

It is important to realize that the shared network resources of the Internet make it easy to communicate with anyone in the world and there is no guaranteed level of service. Unlike a switched circuit link that is facilitated by one or more TELCOs that guarantee a certain level of service, the Internet is not managed by anyone or even few entities. It is a loose consortium of providers, carriers, and users. Therefore, it is considered a "best effort" network with no definable Quality of Service (QoS) level.^[24] The best strategy to increasing your own QoS for telemedicine applications is to increase the bandwidth of the Internet connection and a redundancy through different carriers and perhaps even different topologies.

Wide Area Networks (WANs) were once the standard for long distance data networks. Before the Internet, WANs were established by technologies such as frame relay, leased lines, or ISDN. WANs were very expensive due to the dedication of resources. Today, the Internet and telecommunications backbones and protocols such as Asynchronous Transfer Mode (ATM) have helped the Internet to become the de facto global WAN through whatever topology is appropriate for a particular facility.^[25] As the world becomes increasingly connected, access to partners around the globe becomes more prevalent. More Internet Service Providers (ISPs) and Points of Presence (POPs) provide on-ramps to the Internet by allowing more people to potentially have access to telemedicine. An example of how a telemedicine network hub and spoke system is utilized is represented in (Fig. 1).

Example Hub and Spoke Topology

Wireless cellular networks have provided telecommunications to areas that previously had none. In developed countries, copper has been the mainstay for voice and data communications. Capacity has been increased geometrically through broadband technologies, including fiber, microwave, satellites, Digital Subscriber Lines (DSL), cable, and more intelligent switching. In developing countries, terrestrial copper has been virtually skipped over in favor of cellular. The cost of implementation can be significantly lower and highly scalable, depending on radio frequency allocation. This brings us to a critical discussion known as the last mile.

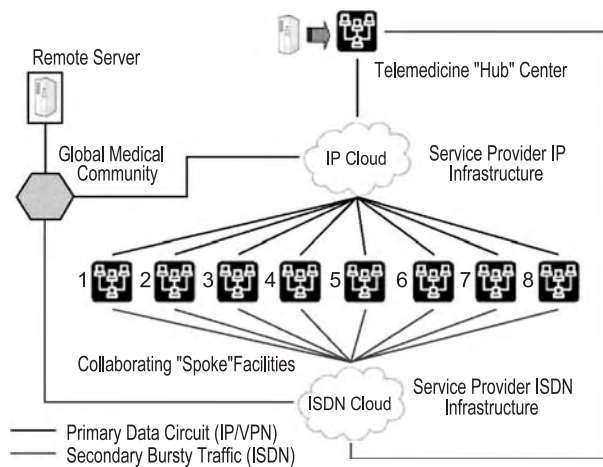


Fig. 1 Utilization of a telemedicine network hub and spoke system.

Despite these technological advances and low-cost phones, bandwidth to remote areas remains a challenge. For effective telemedicine, enough bandwidth must be available at the location of the patient, or capabilities are very limited. Although the regional POP may have a tail circuit to the Internet with capacity in the megabit range, what eventually reaches the point of care or the last mile may only be that provided by the cellular service. This also assumes the provider has a digital signal either on top of an analog wave or digital as the base. This in many cases is only 9600–19,200 bps, which is enough for relatively small amounts of data transfer but not video.

Wireless systems are generalized in terms of capacity and capabilities. They are categorized as generations (G). For example, early analog systems usually referred to as Advanced Mobile Phone Services (AMPS) were considered first generation or 1G. As digital services began to arise, such as Global System for Mobile communications (GSM) and Time Division Multiple Access (TDMA), they were referred to as 2G. Services and bandwidth continue to improve through technologies such as Code Division Multiple Access (CDMA). We are currently in a transition between 2.5G and 3G, and development of 4G is already underway.

3G services promise a quantum leap in digital capacity—on the order of 128–384 Kbps. As these capacities of bandwidth are reached, the last mile will be enabled with enough bandwidth to execute many of the telemedicine tools that will be needed to provide the quality of care that so many people desperately need.

HARDWARE

Before the Internet was popularized and desktop PCs lacked the power to handle the processing and WAN

tasks, dedicated machines running UNIX[®] workstations were needed.^[26] Of course, there are now sufficient resources on typically configured computers. Today, it is the Microsoft Windows[®] platform that is ubiquitous and, to a smaller extent, the Macintosh[®] OS. However, standard protocols and open source options such as Linux[®] may change this.

Minimum Hardware Requirements

When discussing hardware requirements, there are three areas to consider, referring sites, consulting sites, and the servers also called back ends.

The client/server modality of the Internet is ideal for the requirements of medical professionals. It is flexible, scalable, and affordable. Dedicated computers, which are used for telemedicine applications, are no longer the norm. From a commercial standpoint, telemedicine systems must be designed to utilize existing systems.

It goes without saying that the dominance of Microsoft in the personal computing market weighs heavily on the development of telemedicine. First of all, there is the issue of familiarity. People are used to the Windows environment and economies of scale have made the affordability factor low. Using the lowest common denominator for platforms and tools should help to accelerate growth in this industry. No longer will dedicated UNIX machines be required for telemedicine consultations. Beyond the specialty equipment noted subsequently, the computing and network architectures need to be what is readily available and preferably already installed.^[27,28] Purchasing and maintaining separate systems is economically unfeasible.^[29]

The needs of a site such as a remote clinic that requests consultations differs from a location that is strictly a consulting site. In some cases, the site is both. For example, a remote clinic that has a primary care physician may need assistance when diagnosing a neurological problem. A patient such as this would be better served by a virtual visit to a consulting site where a neurologist is available. A requesting or referring site usually needs more computing power and peripherals.

For consulting sites, computers need to be of average configuration, such as those that may be on your desk right now. These machines can readily access patient data presented by the referring sites. In some cases, special software may need to be installed but the push in the industry is standards compliance and browser-based applications.

Peripherals

Data, audio, and video are key components that travel over the networks, but the sources of acquisition are

T

relegated to the peripherals. During a consult, other acquisition devices besides the primary camera on the CODEC may be required. Examples include a multi-purpose portable camera that can be modified as a dermascope, otoscope, ophthalmoscope, etc. Other tools may include a digital stethoscope, ultrasound, pulse oximeter, blood pressure cuff, etc. Often, a document camera will be used to capture radiographs unless the image is already digital and can be transferred electronically. A flatbed scanner is also an affordable tool that can be useful.

Composite patient monitors such as a Propaq[®] or Criticare[®] are common in medical facilities and have data/serial ports, but, unfortunately, these devices do not yet fully support standardized output strings for sharing of data with remote sites. The IEEE 1073 Medical Information Bus standard, which has been ratified since 1994, is not yet embraced by the industry despite the fact that this standard will allow all types of medical device to fully interoperate and share vital data with remote sites.^[30]

Portable Devices

The definition of portable device has changed in recent years. For many years, the term portable device meant a cart on wheels that could be moved within a facility. These carts still exist today and are used in many cases for specialty data collection, video, or patient education.

Another more recent definition is a Personal Digital Assistant (PDA). PDAs are handheld devices such as Palm[®] computing devices or Microsoft's PocketPC[®]. These PDAs are becoming very popular, especially in medical schools where students carry and use the devices for the most basic tasks such as contacts and calendaring. But, they are increasingly embracing the applications designed for medicine.

SYSTEM ARCHITECTURE

Front-end software is the equivalent to the personal interaction of the patient and physician. For over a decade, computer users have become accustomed to the familiar Microsoft Windows OS. At least for now, this has become, in most cases, the front end for the telehealth community. Platform-independent applications such as Java that run either as applications or applets within a Web page are slowly gaining favor over applications written for a specific platform.

Back ends refer to the nomenclature associated with the data management that resides on a server. Any Web server can be considered a back end in the broadest sense but, in healthcare, a back end usually

refers to patient data from the clinical information system, which is often older and established legacy systems. Where existing clinical information systems are not available, stand-alone databases may be used for specific telemedicine practice. These back ends provide the net-enabled tools that make integrated telemedicine systems operational. This data links the participants and related services associated with the healthcare continuum.

Between the Web server and the data repository are processes called middleware. Middleware is the glue that binds the dynamic data to the telemedicine participants. The clients are permitted access to the databases, which may or may not be physically on that same server, through programming interfaces such as Common Gateway Interface (CGI)[®], Practical Extraction & Reporting Language (Perl)[®], Hypertext Preprocessor (PHP)[®], Cold Fusion[®], Java[®], and others. This business logic, as it is often referred, permits dynamic access to add, change, and delete information in a database.

Medical Informatics

There are new opportunities provided through the use of database tools that are referred to as medical informatics. Medical informatics is an emerging discipline that can be defined as the study, compilation, and implementation of structures and algorithms to improve communication, understanding, and management of medical information. As data is collected through a variety of forms and locations onto central or distributed servers, the ability to aggregate and evaluate data is of great value. There are various types of database systems available, such as flat file, Relational Database Management Systems (RDMS), object oriented (OO), hierarchical, and native eXtensible Markup Language (XML). These databases provide access to archives of medical knowledge and patient records.

Lessons learned from corporate America are beginning to take hold in the medical community. Warehoused data can be de-identified, i.e., remove patient-identifying information such as name, address, social security number, etc., aggregated, and evaluated for clinical trends or patterns. This can be manipulated in many different ways, such as to support evidence-based medicine and studies in epidemiology.

MIGRATION AND STANDARDS ADHERENCE

One of the major problems facing this industry is the interoperability of legacy systems with newly deployed, Web services.^[31] There is much debate about the

migration from paper-based record to electronic versions. To do this requires agreement by the industry on both form and function. With over 170 medical specialties and sectors such as nursing, laboratory, administrators, technicians, and physicians weighing in, the standardized EMR has an arduous path ahead. To date, at least the messaging format is largely agreed upon—XML in concert with Health Level Seven (HL7) and associated schemas.^[32]

In fact, the U.S. government has recently issued governmentwide standards for coding and sharing medical records. These standards are now mandatory for all federal agencies and stem the work of the Defense, Health and Human Services, and Veterans Affairs departments. The initiative will require agencies initially to adopt five standards:

- For messaging, the Health Level 7 standard, which includes XML, to share patient information, such as order entry, scheduling, tests and coordination of admittance, discharge, and transfer records during inpatient care.
- For ordering drugs from pharmacies, standards created by the National Council on Prescription Drug Programs.
- For systems integration, IEEE 1073 (The Medical Information Bus) to standardize how medical devices interface with computer systems.
- For diagnostic imaging, DICOM to capture, store, and share information from imaging devices.
- For laboratory records, the Logical Observation Identifier Name Codes (LOINC) to support the exchange of clinical lab results.

The legacy systems in use today use a variety of data formats such as Electronic Data Interchange (EDI), but many are proprietary in nature and not originally designed to interoperate with other systems. But, for EMRs to take hold, other issues need to be addressed. For example, who will be responsible for storing, maintaining, and hosting EMRs? Logic dictates that the common denominator is the patient. A medical encounter involves a physician(s) and the patient. It stands to reason that the patient should be empowered with control of his or her own EMR, but patients, as a whole, do not have the resources or expertise for the task. Because an EMR and its evolutionary future will require constant updates, insurance/payee information, and security components, a centralized facility such as an application service provider (ASP) or perhaps the government will need to assume the responsibility. But, is this what patients want?

Patients should be able to provide a doctor with his or her most recent and comprehensive medical history—no matter where they will be. Because all medical facilities do not necessarily have access to the

Internet or choose not to—for medical data—because of the requirements of the Health Insurance Portability and Accountability Act (HIPAA), the EMR may best be a smart card or chip the patient carries that can be easily read by any computer with a Universal Serial Bus (USB) port, card reader, or PDA. But again, if this record needs to be updated, should the patient be responsible for this? It is unlikely, and a solution needs to be found.

TRAINING AND DISTANCE LEARNING

Training and education are very important aspects of telemedicine. Telemedicine systems must be designed to be transparent to the user. The tools and interface must be highly intuitive and easy to master. Software developers straddle a difficult line in trying to provide feature-rich applications while maintaining a user-friendly environment.

Despite all efforts, the data requirements and multitude of tools required will invariably require some degree of training. In addition to the interface itself, it is recommended that those involved in telemedicine become familiar with the technology. Clinicians and even administrators can find that training in the very basics of telecommunications, data storage, imaging, and digital protocol can be a valuable augment to the people using the system. This type of training need not be exhaustive rather than a cursory overview of how it works. Training can be either live or online.

In concert with telemedicine as a clinical tool, distance learning or distance education is a complementary adjunct. The same networks that enable the host of clinical activities also provide the conduit for sharing of educational components. A good example is Continuing Medical Education (CME.) Physicians must continually keep their skill levels current, and CME is a primary vehicle for that activity. Other examples include demonstration of new technologies and even the delivery of customized modules to remote field clinicians. For example, in a remote clinic that is far from a host facility, a primary care physician cannot possibly be an expert in all areas of medicine. Yet, he or she is faced with these challenges regularly. By utilizing telemedicine consultative services, customized educational modules can be sent to address issues at hand, such as how to best handle issues like an outbreak of an infectious disease.

SECURITY

Security is increasingly becoming an integral component of telemedicine systems. Entire books are written on security and will not be addressed in detail



here. However, a few words are warranted. In the United States, HIPAA has created a stir in the medical community. Increased vigilance towards patient privacy is included in the regulations and that means increased security. This turns out to be just as much a perception issue as an actual one. As firewalls are erected and encryption is added using virtual private networks (VPNs), public/private keys and high-strength encryption algorithms, the perception is that it can still be intercepted and cracked.^[33] But, the benefits of electronic patient data and medical resources will outweigh these perceived and actual risks.^[34]

CONCLUSION

Telemedicine is a digital event that renders patient data into a simulacrum or hologram of the actual patient.^[35] The very name telemedicine feels increasingly anachronistic as we try to apply all the fabulous possibilities of informatics to the primitive videoconferences and telemetry, which initially defined telemedicine. Perhaps informatics will subsume telemedicine as a subheading. Certainly, the technology engineering, which makes telemedicine possible, is not the outcome of medical research. Telemedicine applications have called for great energy and creativity in medical science. However, the basic research in that of engineering, and it is rare that medicine is even the second or third application of basic discoveries. Medicine will continue to benefit from engineering technology that is informally applied vigorously to health issues by multidisciplinary research.

ARTICLES OF FURTHER INTEREST

Computer-Assisted Surgery; Eye Tracking: Characteristics and Methods; Eye Tracking: Research Areas and Applications; Medical Robotics

REFERENCES

1. Bashshur, R. Technology Serves the People: The Story of a Cooperative Telemedicine Project by NASA, the Indian Health Service and the Papago People, Superintendent of Documents. US Government Printing Office: Washington, D.C., 1980; 110 p.
2. Benschoter, R.A. *Educational Broadcasting; CCTV-Pioneering Nebraska Medical Center*, Oct. 1971; 1-3.
3. Berek, B.; Canna, M. Telemedicine on the move: Health care heads down the information superhighway. *Hosp. Technol. Ser.* **1994**, *13* (6), 1-65.
4. Lobley, D. The economics of telemedicine. *J. Telemed. Telecare* **1997**, *3* (3), 117-125.
5. Norris, T.E.; Hart, G.L.; Larson, E.H.; Tarczy-Hornoch, P.; Masuda, D.L.; Fuller, S.S.; House, P.J.; Dyck, S.M. Low-bandwidth, low-cost telemedicine consultations in rural family practice. *J. Am. Board Fam. Pract.* **Mar.-Apr. 2002**, *15* (2), 123-127.
6. Stensland, J.; Speedie, S.M.; Ideker, M.; House, J.; Thompson, T. The relative cost of outpatient telemedicine services. *Telemed. J.* **1999**, *5* (3), 245-256. Fall.
7. Zincone, L.H., Jr.; Doty, E.; Balch, D.C. Financial analysis of telemedicine in a prison system. *Telemed. J.* **1997**, *3* (4), 247-255. Winter.
8. McCue, M.J.; Mazmanian, P.E.; Hampton, C.; Marks, T.K.; Fisher, E.; Parpart, F.; Krick, R.S. The case of Powhatan Correctional Center/Virginia Department of Corrections and Virginia Commonwealth University/Medical College of Virginia. *Telemed. J.* **1997**, *3* (1), 11-17. Spring.
9. Wilson, M. The future of telemedicine. *Stud. Health Technol. Inform.* **2002**, *80*, 129-136.
10. Picot, J. Telemedicine and telehealth in Canada: Forty years of change in the use of information and communications technologies in a publicly administered health care system. *Telemed. J.* **1998**, *4* (3), 199-205. Fall.
11. Murphy, R.L.; Bird, K.T. Telediagnosis: A new community health resource; observations on the feasibility of telediagnosis based on 1000 patient transactions. *Am. J. Public Health* **Feb. 1974**, *64* (2), 113-119.
12. Murphy, R.L.; Fitzpatrick, T.B.; Haynes, H.A.; Bird, K.T.; Sheridan, T.B. Accuracy of dermatologic diagnosis by television. *Arch. Dermatol.* **Jun. 1972**, *105* (6), 833-835.
13. Ortner, J. Spacelab concept and its utilization for science. *Acta Astronaut.* **Jan.-Feb. 1975**, *2* (1-2), 1-13.
14. Doarn, C.R.; et al. Telemedicine in the US Space Program. *Telemed. J.* **Spring 1998**, *4* (1), 19-30.
15. Rossi Mori, A.; Consorti, F.; Galeazzi, E. Standards to support development of terminological systems for healthcare telematics. *Methods Inf. Med.* **Nov. 1998**, *37* (4-5), 551-563.
16. Rovetta, A.; Bejczy, A.K.; Sala, R. Telerobotic surgery: Applications on human patients and training with virtual reality. *Stud. Health Technol. Inform.* **1997**, *39*, 508-517.
17. Harnett, B.M.; Doarn, C.R.; Russell, K.M.; Kapoor, V.; Merriam, N.R.; Merrell, R.C. Wireless telemetry and internet technologies for medical management: A Martian analogy. *Aviat. Space Environ. Med.* **2001**, *72* (12), 1125-1131.
18. Darkins, A.; Fisk, N.; Garner, P.; Wootton, R. Point-to-point telemedicine using the ISDN. *J. Telemed. Telecare* **1996**, *1* (2 Suppl), 82-83.
19. Wu, T.K.; Liu, J.L.; Tschai, H.J.; Lee, Y.H.; Leu, H.T. An ISDN-based telemedicine system. *Digit. Imaging* **Aug. 1998**, *11* (3 Suppl 1), 93-95.
20. Chan, F.Y.; Whitehall, J.; Hayes, L.; Taylor, A.; Soong, B.; Lessing, K.; Cincotta, R.; Cooper, D.; Stone, M.; Lee-Tannock, A.; Baker, S.; Smith, M.; Green, E.; Whiting, R. Minimum requirements for remote realtime

- fetal tele-ultrasound consultation. *J. Telemed. Telecare* **1999**, *5* (3), 171–176.
21. Turner, J.; Brick, J.; Brick, J.E. MDTV Telemedicine Project: Technical considerations in videoconferencing for medical applications. *Telemed. J.* **1995**, *1* (1), 67–71. Spring.
 22. Broderick, T.J.; Harnett, B.M.; Doarn, C.R.; Rodas, E.B.; Merrell, R.C. Real-time Internet connections: Implications for surgical decision making in laparoscopy. *Ann. Surg.* **Aug. 2001**, *234* (2), 165–167.
 23. Feduykin, I.V.; Reviakin, Y.G.; Orlov, O.I.; Doarn, C.R.; Harnett, B.M.; Merrell, R.C. Experience in the application of Java technologies in telemedicine. *J. Appl. Health Technol. eHealth Int.* **2002**, *1*, 3.
 24. Rosser, J.C.; Bell, R.L.; Harnett, B.M.; Rodas, E.; Murayama, M.; Merrell, R.C. Use of mobile low-bandwidth telemedical techniques for extreme telemedicine application. *Am. Coll. Surg. J.* **1999**, *189* (4), 397–404.
 25. Harnett, B.; Angood, P.; Merriam, N.; Satava, R.; Doarn, C.R.; Merrell, R.C. The benefits of integrating internet technology with standard communications for telemedicine in extreme environments. *Aviat. Space Environ. Med.* **2001**, *72* (12), 1132–1137.
 26. Ruffin, M. Telemedicine: Where is technology taking us? *Physician Exec.* **Dec. 1995**, *21* (12), 43–44.
 27. Worth, E.R.; Patrick, T.B.; Klimczak, J.C.; Reid, J.C. Cost-effective clinical uses of wide-area networks: Electronic mail as telemedicine. *Proc. Annu. Symp. Comput. Appl. Med. Care* **1995**, 814–818.
 28. Della Mea, V. Internet electronic mail: A tool for low-cost telemedicine. *Telemed. Telecare* **1999**, *5* (2), 84–89.
 29. High, W.A.; Houston, M.S.; Calobrisi, S.D.; Drage, L.A.; McEvoy, M.T. Assessment of the accuracy of low-cost store-and-forward teledermatology consultation. *J. Am. Acad. Dermatol.* **May 2000**, *42* (5 Pt 1), 776–783.
 30. Kennelly, R.J. Improving acute care through use of medical device data. *Int. J. Med. Inf.* **Feb. 1998**, *48* (1–3), 145–149.
 31. Rossi Mori, A.; Consorti, F.; Galeazzi, E. Standards to support development of terminological systems for healthcare telematics. *Methods Inf. Med.* **Nov. 1998**, *37* (4–5), 551–563.
 32. Chapman, S. XML and HIPAA: The data wave of the future? Web-enable your data to strategically position your organization for future e-health initiatives. *Health Manag. Technol.* **Feb. 2003**, *24* (2), 14, 16.
 33. Wiederhold, G. Future of security and privacy in medical information. *Stud. Health Technol. Inform.* **2002**, *80*, 213–229.
 34. Moody, M. HIPAA strengthens business case for electronic report distribution systems. *J. Healthc. Inf. Manag.* **2002**, *16* (3), 47–51. Summer.
 35. Satava, R.M. Virtual reality surgical simulator. *Surg. Endosc.* **1993**, *7*, 203–205.

Tendons and Ligaments, Mechanical Testing Of

Richard E. Debski

Susan M. Moore

Eric J. Rainis

Musculoskeletal Research Center, Department of Bioengineering, University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.

INTRODUCTION

The injury and subsequent repair of biological soft tissues of the human body continue to be significant problems in the fields of orthopedics and bioengineering. With the increased interest in athletic activities, severe soft-tissue injuries are becoming more prevalent in the general population.^[1] Such injuries can cause acute or even chronic instability in the diarthrodial joints of the musculoskeletal system. Given the increasing life span of humans, and the increase of incidence of injuries to soft tissues, prevention of their chronic instability is becoming more significant.

Tendons and ligaments are important structures that guide motion as well as provide stability in diarthrodial joints. Injury to these structures can lead to significant changes in joint motion and morbidity for patients. This article will focus on the biomechanics of tendons and ligaments: Section “Structural Framework” will review the structure and function of the tendons and ligaments, while section “Uniaxial Tensile Testing” will focus on uniaxial testing regimes to assess their properties. Finally, sections “Multiaxial Mechanical Testing” and “Viscoelasticity” will describe recent developments in testing techniques and their impact on computational modeling.

STRUCTURAL FRAMEWORK

Tendons

Tendons are bands of tough, inelastic fibrous tissue with closely packed, parallel collagen fiber bundles that connect bone to muscle and are composed mainly of water, type-I collagen, and cells called tenocytes. Tenocytes are specialized fibroblasts responsible for the maintenance of the collagen structure. Minor fibrillar collagens, fibril-associated collagens, and proteoglycans are present in small quantities and are critical for tendon structure. Most of the strength of tendon is because of the parallel, hierarchical arrangement of densely packed collagen fibrils.

The origin of a tendon is the region where it joins to a muscle and it inserts into bone at an enthesis, where the collagen fibers are mineralized and integrated into bone tissue. Tenocytes produce collagen molecules that aggregate end-to-end and side-to-side to produce collagen fibrils. Bundles of these fibrils are then organized by tenocytes to form collagen fibers—the primary load-bearing elements in the tissue. These fibers coalesce into macroaggregates, which are bounded by connective tissue endotendon and are termed “fascicles.” Finally, groups of fascicles are bounded by the epitendon and peritendon to form the tendon.

Ruptures and/or tears of tendons upset the balance between mobility and stability. These injuries often result in abnormal kinematics and damage to other tissues around the joint that may lead to morbidity, pain, and osteoarthritis. The incidence of tendon injury has increased over the last few decades, and it is estimated that they account for 30–50% of all injuries related to sports.^[2] Tendon injuries are typically a result of repetitive loading to the tendon, rather than traumatic loading, which is more characteristic of ligament injuries.

Ligaments

Ligaments are a sheet or band of tough, inelastic fibrous tissue with parallel collagen fiber bundles that connect bone to bone. Ligaments are composed of water and densely packed collagen fibers that run in the longitudinal direction, parallel to the axis of loading. Non-collagenous proteins such as elastin, a fibrillar protein (less than 1%), and proteoglycans are also present. The collagen fibrils in ligaments are arranged in such a way that an increase in tensile force results in the recruitment of more fibrils to help resist the increased load. This arrangement is known as a crimp pattern, which can be described as a series of springs with differing lengths in parallel (Fig. 1). The crimped nature of the fibrils serves as a guide to joint motion and provides restraint at extremes of joint motion. During normal activity, ligaments can be easily elongated to maintain normal kinematics and allow the joint to move easily and smoothly.

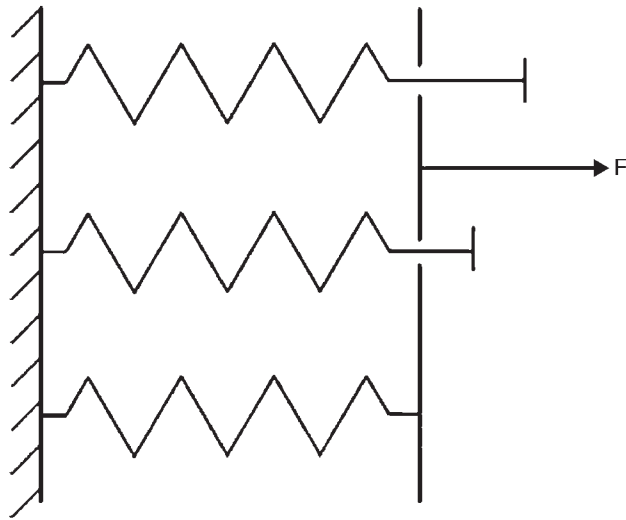


Fig. 1 Schematic of the crimped pattern of collagen fibers in a ligament.

With the application of higher loads, such as those found during exercise, the stiffness of the ligament increases to prevent any excessive motion in the joint. However, if the applied load exceeds the maximum limits of the ligament, as may occur during athletic events, ligament damage may occur.

Rupture or tears of ligaments can cause a joint to become unstable. These injuries often affect the kinematics of the joint and subsequently cause injury to surrounding tissues, leading to morbidity, pain, and osteoarthritis. Similar to tendon injuries, the incidence of ligament injury has also increased and is estimated to be 2 out of 1000 people in the general population per year, for the knee alone. These injuries primarily involve the medial collateral ligament (MCL) or the anterior cruciate ligament (ACL), accounting for as much as 90% of all sports-related ligament injuries.^[3]

UNIAXIAL TENSILE TESTING

What are Tensile Properties

Many researchers have examined the tensile properties of tendons and ligaments.^[4–20] The structural properties of a bone–ligament–bone complex are represented by parameters such as stiffness, ultimate load, and energy absorbed to failure that are obtained from the load–elongation curve (Fig. 2). Stiffness is simply the slope of the linear portion of the curve, ultimate load is the point at which the tissue fails, largest load achieved, and energy absorbed to failure is the area underneath the load–elongation curve from the onset of loading to the ultimate load.

Mechanical properties of a tendon and ligament substance are a measure of tissue quality and are

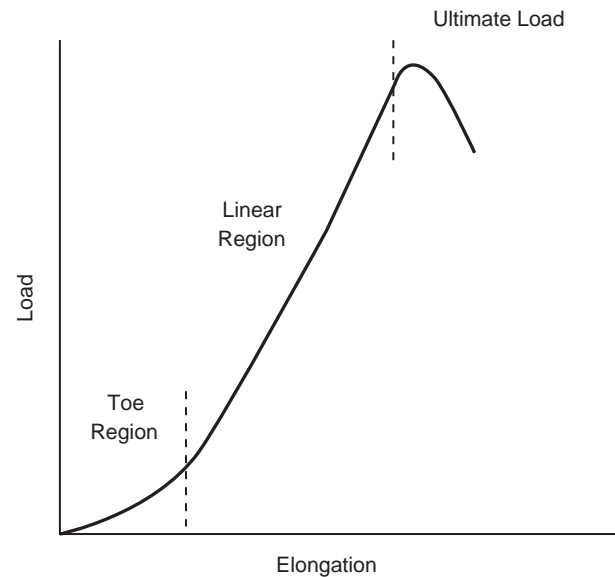


Fig. 2 Load–elongation curve depicting the three regions observed for tendons and ligaments: 1) toe region; 2) linear region; and 3) ultimate load.

represented by the tangent modulus, tensile strength, ultimate strain as well as the strain energy density, which are obtained from the stress–strain curve. Like stiffness, modulus is the slope of the linear portion of the curve. Ultimate stress is the maximum amount of stress achieved before failure, while ultimate strain is the strain in the tissue when ultimate stress is achieved. Finally, the strain energy density is the area under the stress–strain curve until failure.

A bone–ligament–bone complex allows for easy gripping of the bones thus minimizing any risk of slippage during testing. For this reason, it is convenient to test the mechanical properties of the tissue substance using a bone–ligament–bone complex. However, there are some instances when it is necessary to excise a tissue sample and clamp directly to the tissue substance making slippage a concern. As it is assumed that no slippage occurs during testing, one must take precautions to ensure a rigid fixation. In addition, a uniform loading distribution is assumed across the cross-section of the tissue. One method used to ensure this uniform load is to create a tissue sample with a large aspect ratio or length to width ratio. An aspect ratio of at least five is desired when performing such experiments. For an example of a typical tensile testing setup, see Fig. 3.

Testing Procedure for Structural Properties of a Bone–Ligament–Bone Complex

Structural properties can be obtained by performing a uniaxial tensile test, in which the load is measured

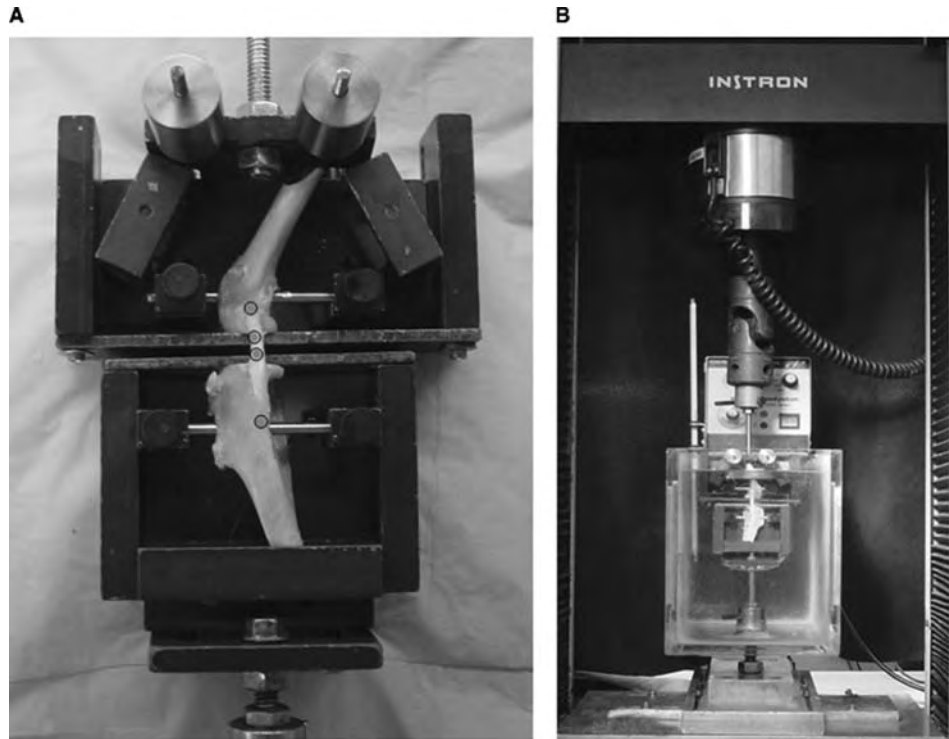


Fig. 3 (A) Close-up of bone–ligament–bone complex where strain markers have been placed on the tissue midsubstance and insertion sites and (B) view of experimental set-up for tensile testing.

simultaneously with the elongation. These properties depend on the geometry of the ligament in addition to the properties of the bony insertion sites. The load–elongation curve can be divided into three separate regions (Fig. 2). The first region, or toe region, has a low stiffness and is nonlinear. This behavior has been attributed to the crimp pattern of the collagen fibrils and to the nonuniform recruitment of the individual fibers. The next region is linear and has a higher stiffness, resulting from the fibers actually being stretched. The third region, or failure region, is initiated by the load at which the fibers can no longer withstand the force that is being applied to them. This point is known as the ultimate load. On additional loading, the bone–ligament–bone complex fails.

As stated previously, the stiffness is calculated using the slope of the linear portion of the load–elongation curve. However, owing to variability of the load–elongation curves between specimens, it is necessary to utilize a reproducible technique to calculate the stiffness for each specimen. One such technique is to define the stiffness as the slope of the load–elongation curve between two elongations that are within the linear region for all specimens. Another technique fits a line to the linear region of the curve and removes data points on either end until the R^2 value reaches 0.8.^[21,22] The ultimate load is defined as the maximum load achieved and the energy absorbed to failure is simply obtained by calculating the area under the curve, where the upper bound is the ultimate load. The ultimate

elongation is then represented by the elongation at failure.

Testing Procedure for Obtaining Mechanical Properties of the Tissue Midsubstance

Mechanical properties of a tissue midsubstance can also be obtained by performing a simple uniaxial tensile test. As stated above, the mechanical properties (or quality) of the midsubstance itself can be obtained by normalizing the tensile load by the cross-sectional area (stress) and by normalizing the change in elongation in a defined region of the ligament midsubstance by the initial length (strain).

A common approach to mechanical testing is to use dog-bone tissue samples instead of bone–ligament–bone complexes.^[21–23] A dog-bone sample provides a good aspect ratio and increases the chance of a midsubstance failure site, as well as ensuring a uniform load across the tissue (Fig. 5). As there is no bone present when testing a dog-bone sample, the ends of the tissue must be held rigidly by clamps. Therefore, precautions must be made to eliminate slipping of the tissue from the clamps, which is not a problem when using bone–ligament–bone complexes.

To determine the stress in the midsubstance, it is crucial that the cross-sectional area is obtained for the midsubstance. There are two types of techniques used to obtain the cross-sectional area of the tissue: contact

and noncontact.^[10,24,25] Contact methods include the use of molding techniques, digital vernier calipers, and area micrometers.^[17,26] These techniques introduce error, as they either assume a rectangular cross-section or force the tissue into a rectangular cross-section.^[14,27–29] To minimize these errors, there has been a recent push toward noncontact methods for obtaining the cross-sectional area. These techniques include the shadow amplitude method,^[30] the profile method,^[10] and the use of a light source.^[31] Another technique that is utilized is a laser micrometer system.^[32,33] This method has been shown to be highly accurate and reproducible, and the reconstructed shapes match histologic sections quite well. A new laser-reflectance system has also been developed to measure cross-sectional shape and area of soft tissues with a concave surface.^[34]

Several methods also exist to measure the strain in the midsubstance of the tissue. Again there are contact^[35,36] and noncontact^[5,37,38] methods that can be used. An example of a contact method is a differential voltage reluctance transducer. This device can be used to calculate the strain by suturing two metal probes on the surface of the tissue, and based on the distance those probes move, strain can be calculated.^[35,36] Because contact methods may alter the properties of the tissue owing to deformation, there has been a large push toward noncontact methods. More specifically, optical tracking is being used, because it requires no direct contact with the specimen. When using optical tracking, two or more markers are drawn or fixed to the surface of the tissue by means of Verhoeff's stain,^[5] elastin stain,^[37] or reflective tape^[38] to serve as gauge lengths. Camera systems are used to record marker motion, and the distance between the markers can be expressed as a function of the gauge length, corresponding to the strain in the tissue.^[4,39–41] As stated above, the tangent modulus of the tissue is calculated using the slope of the stress–strain curve. This can be performed using the same methods described in section “Testing Procedure for Obtaining Mechanical Properties of the Tissue Midsubstance” to calculate stiffness. To calculate ultimate strain and tensile strength, the strain and stress at the point just before failure is used, respectively.

Factors Affecting Tendon and Ligament Properties

There are several factors that affect the tensile properties of tendons and ligaments including hydration, temperature, strain rate, age, and orientation. The *in vitro* environment has profound effects on ligament behavior. Maintenance of specimen hydration is of primary importance to the study of soft tissues.

Temperature has also been shown to be important,^[13] although to a lesser degree. Increases in temperature from 2°C to 37°C resulted in a decrease in stiffness and a decrease in areas of hysteresis. These factors have led many investigators to immerse their specimens in a physiologic saline bath in which the pH and temperature are closely regulated.^[6,12,22,42,43]

The tensile properties of ligament are only affected minimally by strain rate. Two groups (3.5 mo old and 8.5 mo old) were used in a study of the rabbit femur–MCL–tibia complex (FMTC), with elongation rates of 0.01%/sec to over 200%/sec, corresponding to 0.008 and 113 mm/sec. For the young group, the structural properties of the FMTC and the mechanical properties of the MCL substance in the prefailure range were found to increase by approximately 300%, which was minimal considering that the strain rate had been increased by several decades. The older group followed similar trends, but the increases were even smaller. Regardless of the strain rate, however, the entire younger group failed by tibial avulsion, while the older group failed by ligamentous disruption either at midsubstance or near the tibial insertion.^[44] Although clinical injuries occur at much faster strain rates, much experimental work is performed at slow rates, based on the limitations of the testing equipment. These findings^[44] imply that age of the animals, rather than strain rate, is a much more important factor when performing such tests.

Age has been shown to have varying effects on ligaments and tendons. The significance of aging has been investigated in the rabbit MCL at 3, 5, 6, 12, 36, and 48 months of age. The tensile properties of the MCL remained relatively constant up to 12 months, with a slight decline at 48 months. The structural properties of the ACL from young donors (ages 22–35 yr) were also found to be significantly higher than those of older donors. The stiffness and ultimate load are also higher in the younger donors, by more than threefold.^[14] Age has similarly been shown to have an effect on the properties of the glenohumeral capsule.^[20] On the other hand, no significant differences were found between two different age groups (29–50 yr and 64–93 yr) of the human patellar tendon when comparing ultimate strength. The ultimate strengths for the younger and older group were 64.7 ± 15 and 53.6 ± 10 MPa, respectively.^[45]

Because bone–ligament–bone complexes are non-uniform in geometry and shape, the direction of applied force and the initial position of the ligament are important for its load–elongation behavior. Applying the force in the direction of anatomical loading allows a greater proportion of the fiber bundles to be loaded. In a study on young human femur–ACL–tibia complexes (FATC), paired specimens were studied with one knee being tested along the anatomic axis of the

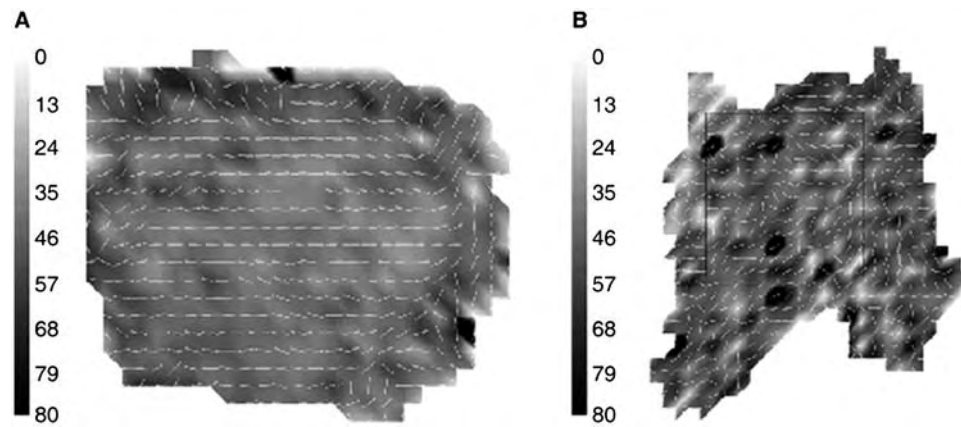


Fig. 4 Collagen fiber of the long head of the biceps tendon (A) aligned along the longitudinal axis of the tendon and axillary pouch of the glenohumeral capsule (B). White dashes indicate average local alignment of the fibers while the gray scale indicates the variability of the fiber alignment in that location.

ACL and the other along the tibial axis.^[33] In the anatomic orientation, the natural insertion angles of the ACL were maintained, allowing for a uniform load distribution across the ligament, whereas for the tibial orientation, the insertion angles were not maintained, resulting in a nonuniform load and thus a higher number of insertion site failures. The stiffness and strength for the FATC's tested in the anatomic orientation were significantly greater than those in the tibial orientation. Similar results have been observed for the FATC of the porcine, canine, and rabbit models.^[13,46]

A more recent study^[47] has analyzed the effects of multiple freeze–thaw cycles on the tensile properties of the rabbit MCL, an issue originally examined by Woo et al.^[5] The results showed that careful postmortem freezing for up to three months did not significantly alter the structural properties of the bone–ligament–bone complex and the mechanical properties of the ligament midsubstance. This finding is very important as many biomechanical tests of soft tissues are generally done after a period of postmortem storage by freezing.

MULTIAXIAL MECHANICAL TESTING

Ligament and Tendon Anisotropy

The mechanical response of an isotropic material is independent of the direction of loading. However, ligaments and tendons have demonstrated anisotropic behavior. Thus, their mechanical response is dependent upon the direction of loading. A special case of anisotropy is transverse isotropy, indicating an axis of material symmetry exists, which causes the mechanical response to be independent of the loading direction in

the plane transverse to this axis (i.e., isotropic in the plane transverse to the axis of material symmetry).

Tendons and ligaments are considered transversely isotropic materials. This is owing to the fact that traditional ligaments (i.e., ligaments that primarily experience uniaxial loading) and tendons such as the MCL of the knee and the long head of the biceps tendon have a preferred collagen fiber alignment^[38,43,48] resulting in an axis of material symmetry. The preferred alignment for such ligaments and tendons is along their longitudinal axis such that loads are transmitted between bones or between bone and muscle along their length. More recently, it has been shown that the capsular tissues such as the glenohumeral joint capsule may not necessarily have a distinct preferred collagen fiber alignment.^[49] Using the small angle light scattering technique,^[50] Debski et al.^[49] determined the local average collagen fiber direction of the long head of the biceps tendon and



Fig. 5 View of the posterior capsule illustrating the orientation of the dog-bone-shaped (midsubstance 12.5 mm × 2.5 mm) transverse and longitudinal tissue samples with respect to the posterior band (PB-IGHL).

Table 1 The mechanical properties for transverse axillary pouch and longitudinal axillary pouch tissue samples (mean \pm SD)

	Modulus (MPa)	Ultimate stress (MPa)	Ultimate strain (%)	Strain energy density (MPa)
Transverse ($n = 10$)	5.4 ± 2.9^a	0.8 ± 0.4^a	23.5 ± 11.5	10.8 ± 8.5
Longitudinal ($n = 10$)	14.8 ± 13.1	2.0 ± 1.0	33.3 ± 23.6	21.1 ± 15.4

^aSignificant differences ($p < 0.05$) in modulus and ultimate stress were detected. Transverse and longitudinal refers to those tissue samples taken perpendicular and parallel to the longitudinal axis of the tissue, respectively.

that of the axillary pouch of the glenohumeral capsule (Fig. 4). For the long head of the biceps tendon and the axillary pouch, the percentage of the tissue where 50% of the fibers were oriented within 25° – 45° of their local average direction was $61.6 \pm 15.2\%$ and $23.2 \pm 8.5\%$, respectively. Although the axillary pouch did not demonstrate a clear axis of material symmetry, it is important to note that the behavior of this tissue is still anisotropic as demonstrated in a recent study of its bidirectional mechanical properties.^[22] Differences in the ultimate stress and tangent modulus were detected between tissue samples tested in two perpendicular directions (Table 1).

As ligaments and tendons are anisotropic materials, constitutive models that mathematically describe their behavior under loading conditions are much more complex than the simple constitutive model of an isotropic elastic material (i.e., $\sigma = E\epsilon$). Many researchers have successfully used transverse isotropy to describe the behavior of tendons and ligaments.^[11,51–60] The physical interpretation for transversely isotropic symmetry is that of a matrix reinforced with a single fiber family whereby the matrix properties, fiber–matrix interaction, and fiber–fiber interactions give rise to the material behavior.^[58] Therefore, it has recently been advocated that to sufficiently characterize the three-dimensional material behavior of tendons and ligaments, it is necessary to collect experimental data on the multiaxial quasistatic and viscoelastic properties of these tissues.^[58] It should be noted that multiaxial testing has also been commonly performed on other soft tissues such as heart and vascular tissues.^[61–64]

Multiaxial Testing

Bidirectional mechanical testing

The simplest form of multiaxial testing is determining the bidirectional properties of soft tissues whereby the properties are determined in two perpendicular directions via simple uniaxial tensile tests. The two directions of interest are the axis of symmetry (i.e., axis collagen fibers are aligned along) and the direction perpendicular to this axis. These two perpendicular directions are of interest, as results from these

experiments, along with knowledge of the collagen fiber arrangement, can then be used to justify describing the tissue as transversely isotropic.

Testing Procedure. As tendons and ligaments do not insert into bone or muscle in the direction perpendicular to their longitudinal axes, it is necessary to harvest tissue samples from the ligament such that clamps can be used to grip the tissue itself. To maintain a good aspect ratio, dog-bone-shaped tissue samples are therefore harvested from the midsubstance of the tissue (Fig. 5).^[21–23] Thus, only the mechanical properties of the ligament substance, and not the structural properties of the bone–ligament–bone complex, can be determined with this testing configuration. As was discussed previously (see section “Testing Procedure for Structural Properties of a Bone–Ligament–Bone Complex”), one assumption for a uniaxial test is that the tissue sample is uniformly loaded. Thus, it is necessary to maintain an aspect ratio of at least five. The mechanical properties of the tissue can be determined in the two perpendicular directions once the cross-sectional area and strain are determined (see section “Testing Procedure for Obtaining Mechanical Properties of the Tissue Midsubstance”). Statistical differences between the two perpendicular directions demonstrate that the tissue is anisotropic.

One limitation to this testing methodology is that the interfiber and transfiber bonds between the collagen fibers may have been broken when the tissue samples were excised from the midsubstance. Moreover, while the dog-bone-shaped tissue sample makes it possible to uniformly load ligaments or tendons with complex geometry, it is important to understand whether or not the tissue is hetero- or homogeneous, as the location that the tissue sample is harvested from may affect the resulting properties. An additional limitation is that this methodology does not evaluate the tissue’s response to shear loading. Understanding the response of tendons and ligaments to shear and incorporating this response into constitutive models may be necessary depending on the functional role of the tissue and the loading conditions to be modeled.

Recent Studies. Bidirectional mechanical tests have recently been performed on the MCL of the knee^[23]

and on the axillary pouch^[21] and posterior regions^[22] of the glenohumeral joint capsule. For the MCL, the ability of three different constitutive models to describe its behavior was evaluated. Thus, dog-bone-shaped tissue samples were harvested in the directions parallel and perpendicular to the longitudinal axis of the MCL. A load to failure test was performed on each of the tissue samples, and a nonlinear regression was performed to determine the material coefficients for three different constitutive models.^[16,65] All three models were able to describe the behavior of the MCL along its longitudinal axis; however, only one of the three models provided a good description of the MCL's behavior in the direction perpendicular to the longitudinal axis. These comparisons are extremely valuable, as finite element models are being increasingly used to describe the stress and strain distributions of soft tissues under clinically relevant loading conditions.^[59,66–68]

Within our research center, bidirectional mechanical tests have also been used for the glenohumeral joint capsule.^[21,22] In these studies, dog-bone-shaped tissue samples were harvested from the axillary pouch and posterior region in the directions parallel and perpendicular to the longitudinal axes of the thickenings of the glenohumeral joint capsule. A load to failure test was performed; the ultimate stress and strain, tangent modulus, and strain energy density were determined in the two perpendicular directions. Unlike traditional ligaments such as the MCL, the differences in the mechanical properties of the two perpendicular directions were not as apparent. Results for the axillary pouch, MCL of the knee, and other ligaments are shown in Table 2.^[22] While a significant difference in the ultimate stress and tangent modulus was found, no differences were detected in the ultimate strain and

strain energy density. The ratio of the tangent moduli (axillary pouch, 3.3 ± 2.8 ; posterior capsule, 4.8 ± 4.2) was an order of magnitude less than the ratio for the MCL^[30] of the knee,^[23] a ligament that has been successfully modeled as transversely isotropic.^[57–59,71] While it is clear that the glenohumeral joint capsule is anisotropic, before utilizing a transversely isotropic constitutive model to describe the behavior of the glenohumeral joint capsule, the use of this model should be validated using a more complex multiaxial testing scheme that includes shear loading.

Shear loading

A more complex multiaxial testing scheme that includes shear loading may be necessary to appropriately evaluate the ability of various constitutive models to describe the behavior of some tendons and ligaments such as the glenohumeral joint capsule. Additionally, evaluating the response of tendons and ligaments to shear loading will allow further elucidation into the functional role these soft tissues play in joint mobility and stability, respectively.

Testing Procedure. Recently, Weiss, Gardiner, and Bonifasi-Lista^[58] have described a testing procedure for ligaments that utilizes shear loading to determine the coefficients of a constitutive model. Rectangular-shaped tissue samples are harvested such that the preferred direction of collagen fiber organization is parallel to one axis of the rectangle. The initial width, height, and thickness of the tissue sample are then recorded. Using a materials testing device, the tissue sample is first preconditioned via a sawtooth displacement profile applied for 10 cycles. The shear

Table 2 Summary of data reported for the axillary pouch, MCL of the knee, interosseous ligament of the forearm, hip capsule, and the cruciate ligaments of the knee

Tissue type	Tangent modulus (MPa)	Ratio of tangent modulus	Ultimate stress (MPa)	Ultimate Strain (%)
Transverse axillary pouch ^[22]	5.4 ± 2.9	3.3 ± 2.8	0.8 ± 0.4	23.5 ± 11.5
Longitudinal axillary pouch ^[22]	14.8 ± 13.1		2.0 ± 1.0	33.3 ± 23.6
Transverse medial collateral ligament	11.0 ± 3.6	30	1.7 ± 0.5	11.7 ± 0.9
Longitudinal medial collateral ligament (knee) ^[23]	332.2 ± 58.3		38.6 ± 4.8	17.1 ± 1.5
Transverse interosseous ligament	1.1 ± 1.6	385	0.1 ± 0.1	N/A
Longitudinal interosseous ligament (forearm) ^[69]	431.3 ± 321.3		52.7 ± 33.5	
Anterior longitudinal axillary pouch ^[9]	30.3 ± 10.6	N/A	5.5 ± 2.0	15.1 ± 5.7
Posterior longitudinal axillary pouch ^[9]	41.9 ± 12.5		5.6 ± 1.9	9.9 ± 5.3
Longitudinal hip capsule ^[70]	Range 76–286	N/A	Range 2.0–6.0	Range 6.2–25.3
Average of anterior and posterior cruciate ligament and lateral collateral ligament ^[8]	345 ± 22	N/A	36 ± 3	15 ± 1

strain rate used for the MCL was 1.3%/sec with a strain amplitude of $\tan(\theta) = 0.4$ based on the measured width of each tissue sample (Fig. 6). Data from the 10th cycle can then be used when determining the parameters of constitutive models.^[57] The clamp reaction force and the crosshead displacement of the materials testing device are continuously recorded during testing. The applied shear strain can then be computed using the tissue samples initial dimensions and crosshead displacement measurements.

Owing to the inhomogeneous deformation during shear loading, a parameter estimation technique^[58] can then be used in conjunction with the finite element method to estimate the parameters for the specified constitutive model. Using this combined approach, the experimental shear loading is simulated using the finite element method, and the parameters of the specified constitutive model are iteratively adjusted until the force–displacement behavior of the finite element predictions match that of the experimentally collected data.

The methodology described above requires that the tissue samples harvested are planar with a relatively uniform thickness and fiber orientation. For many tendons and ligaments, this sort of geometry does not exist. Moreover, this method is dependent upon the optimization algorithm implemented for the parameter estimation. Therefore, it is important to validate the estimated parameters by using them to predict the results of an additional loading condition.

Recent Studies. A combined experimental and computational approach has been used to determine the

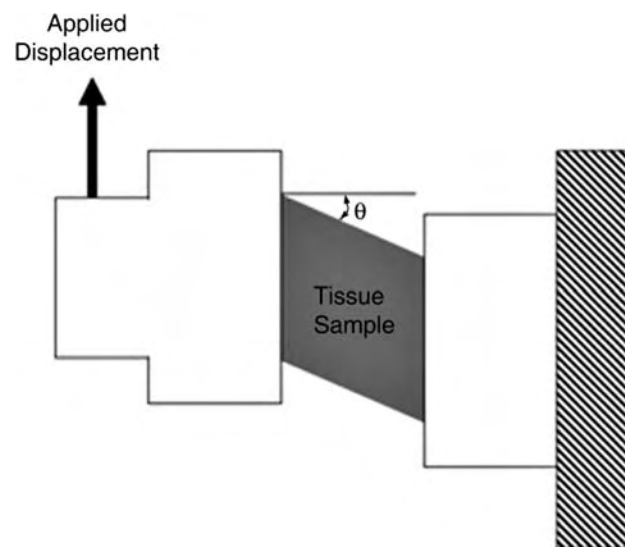


Fig. 6 Schematic illustrating θ , which is determined from the initial clamp-to-clamp length of the tissue sample and the applied displacement.

parameters for a transversely isotropic hyperelastic constitutive model^[58] of the MCL of the knee [Eq. (1)].^[57]

$$W = F_1(\tilde{I}_1, \tilde{I}_2) + F_2(\tilde{\lambda}) + U(J) \quad (1)$$

The functions F_1 and F_2 represent the matrix and fiber family strain energies, respectively and compose the entire deviatoric response of the material. An exponential matrix strain energy was utilized based on preliminary data. \tilde{I}_1 and \tilde{I}_2 are the deviatoric invariants of the right deformation tensor, and $\tilde{\lambda}$ is the deviatoric local fiber stretch. $U(J)$ governs the dilational response of the material.

Using the shear testing protocol described above, experimental force–displacement curves were obtained. The finite element method was then utilized to simulate the experiments. Subsequently, a nonlinear parameter estimation technique coupled with the finite element simulations optimized the parameters of the constitutive model based on the sum of squares difference of the force–displacement curves from the experimental data and finite element model. The nonlinear optimization program was successfully able to minimize the difference between the experimental and finite element force–displacement data for all of the tissue samples tested with R^2 values that were greater than 0.995 in all cases. Therefore, this combined experimental and computational approach was a good method by which the parameters of constitutive models can be determined. Additionally, no sensitivity to changes in the starting values for the constitutive model parameters was detected indicating that an exponential matrix strain energy was appropriate for the human MCL of the knee.

VISCOELASTICITY

Experimental Models

Viscoelastic materials such as tendons and ligaments^[72] are those that have viscous as well as elastic properties. Thus, both time- and history-dependent properties such as hysteresis, creep, and stress relaxation are displayed. This behavior is because of the interactions between their constituents (e.g., proteins, ground substance, and water). A material displays hysteresis when the loading and unloading curves do not follow the same path owing to internal energy dissipation. Creep is also displayed by tendons and ligaments as they elongate over time under a constant load. Finally, stress relaxation is displayed, with a decline in stress observed over time under a constant elongation. Physically, viscoelastic properties play an important role in tendon and ligament function by allowing

relaxation to occur during walking or jogging or by helping to prevent fatigue failure. Therefore, understanding the viscoelastic response of tendons and ligaments has important clinical significance. For example, to select the best graft material to replace a ruptured ligament, such as the ACL of the knee, it is important to select a material that has the same functional range as the ligament being replaced.

Testing procedure

Experimentally, these viscoelastic properties are evaluated using several standard methodologies that apply either a tensile^[15,19,42,45,73,74] or shear load.^[58] Tissue samples are prepared, preconditioned, and the force, stress, strain, and cross-sectional area are determined as was described previously in sections “Testing Procedure for Obtaining Mechanical Properties of the Tissue Midsubstance” and “Factors Affecting Tendon and Ligament Properties,” respectively. However, the loading conditions used to determine the viscoelastic properties differ from those used to determine the failure properties of the tissue samples (i.e., ultimate stress, ultimate strain, and strain energy density).

Hysteresis is evaluated by loading and then unloading the tissue sample. The area of hysteresis is obtained by subtracting the area under the curve obtained during the unloading of the tissue sample from the area under the curve obtained during the loading of the tissue sample. As an example of an experimental protocol used to test the rabbit MCL, one study subjected the tissue samples to 10 cycles of loading and unloading between 0 and 1.0 mm of extension at 1.0 cm/min.^[15] Using this experimental design, the decrease in the amount of hysteresis between the first and tenth cycle could then be compared.

Creep can be evaluated with either a static load or by cyclically loading the tissue sample. For example, a study investigating creep of the rabbit MCL^[75] applied a 14 MPa stress to the tissue samples and held the resulting load constant for 20 min to evaluate static creep. For cyclic creep, the tissue samples were cycled 30 times at 1 Hz between the tare stress and 14 MPa of stress. The amount of creep is obtained by subtracting the elongation achieved when the load is initially applied from the elongation that results after holding the tissue sample at that constant load for the specified time. Using this experimental protocol, the ligament length increased by approximately 10%.

As with creep, the amount of stress relaxation can be determined during either a static elongation or while cycling between two levels of elongation. For example, the static stress relaxation of the human patellar tendon in the toe and linear regions has been

investigated.^[45] The tendon was stretched at 100 mm/min to a stress level of 1 MPa and 4 MPa, respectively, and held at the resulting elongation for 15 min. A 30-min recovery period was allowed between tests. For cyclic stress relaxation, the tissue samples were cyclically extended between the two elongations for 15 cycles at a rate of three cycles per minute. Using this experimental protocol, a decrease in stress of approximately 50% was observed.

The elongations or loads used during tests for hysteresis, creep, and stress relaxation are all selected such that the tissue samples do not approach failure. Thus, the magnitude of the elongations and loads used to determine viscoelastic properties must first be determined from preliminary load-to-failure data.

Computational Modeling

Quasilinear viscoelastic theory

Quasilinear Viscoelastic (QLV) theory was developed by Dr. Y. C. Fung.^[76] The QLV model has been widely used to describe the viscoelastic properties of tendons^[77,78] and ligaments^[79–82] as well as many other soft tissues. The theory makes the assumption that the elastic and time-dependent relaxation functions can be combined in a convolution integral. This results in a one-dimensional general viscoelastic model as shown in Eq. (2).

$$\sigma(t) = \int_{-\infty}^t G(t - \tau) \frac{\partial \sigma^e(\varepsilon)}{\partial \varepsilon} \frac{\partial \varepsilon}{\partial \tau} d\tau \quad (2)$$

Using experimentally collected data, the coefficients of the QLV constitutive model can be derived. With the tissue-specific coefficients, the tendon or ligament behavior under other loading conditions can be predicted.

Recent advances

Recently, three issues regarding the QLV theory have been of interest:^[74,75,81,83,84] 1) application of parameters obtained during stress relaxation experiments to the prediction of creep behavior; 2) the assumption that the relaxation function is only dependent upon time and not strain level; and 3) the assumption that a step change in strain is applied, which is not possible experimentally.

To address the first issue, Thornton et al.^[75] obtained the parameters for the QLV theory using stress relaxation and creep experiments at low stresses for the rabbit MCL. These parameters obtained for both types of experiments predicted different behaviors. Dr. Y. C. Fung also noted that this may occur when using the

QLV theory as it is applied to certain tissues and has stated that he believes that creep is fundamentally more nonlinear, and may in fact not obey the quasi-linear hypothesis.^[76]

For the rat MCL, the rate of creep has been shown to be dependent upon stress level and that the rate of relaxation is dependent upon strain level.^[74] However, this is not admitted in QLV theory. Therefore, at least for the rat MCL model, QLV may not be appropriate when predicting the viscoelastic response over several strain levels. Instead, it has been suggested that the modified superposition method^[85] be utilized. Moreover, to account for this limitation, an analytical expression capable of determining the QLV parameters under an incremental strain history has been developed.^[84] Thus, an incremental strain history can be applied to tissue samples experimentally and the amount of stress relaxation across several strain levels can be utilized to determine a representative amount of relaxation.

Finally, the time-dependent stress relaxation of a tissue is normalized by the stress at the time of a step input of strain in the QLV theory. However, experimentally this is impossible and results in error when determining the coefficients. An improved method to analyze the stress relaxation of tissues following a finite ramp time based on the QLV theory has recently been published.^[81] With this method, the convoluted QLV constitutive equation was simultaneously curve-fit to the ramping and relaxation portions of the data. As both the ramping and the relaxation portions are considered, an instantaneous ramp time is no longer necessary. While this has been evaluated previously,^[86,87] the advantage of this new study is that extremely long ramping times are now possible. Therefore, this adaptation to the traditional QLV theory allows for a great deal of experimental flexibility.

CONCLUSIONS

This article has presented a broad review of the history of mechanical testing of tendons and ligaments, as well as more recent research. Expertise in evaluating the biomechanical properties in tendons and ligaments has been continually evolving with advances in technology. These advances include noncontact strain and cross-sectional area measurements of tissue samples. In addition to assessing the biomechanical properties of tendons and ligaments using traditional uniaxial experiments, there has been a recent interest in performing multiaxial experiments to gain a greater understanding of the function of tendons and ligaments. These experiments can provide a wealth of knowledge regarding the anisotropic nature of tendons

and ligaments and are extremely useful when establishing the parameters of constitutive models. The QLV theory has also been applied to the study of tendons and ligaments. While this theory is an extremely powerful analytical tool, it is important for researchers to understand the limitations and assumptions of such a model to ensure that their application is appropriate.

Subject-specific modeling of diarthrodial joints and their surrounding tissues has been gaining popularity. Thus, evaluating the biomechanical properties of tissue samples under uniaxial or multiaxial loading conditions on a subject-specific basis may be necessary for these models to adequately predict the tensile, shear, and/or viscoelastic response of these tissues. Furthermore, there exists a need to assess the ability of various constitutive models to predict the response of tendons and ligaments to applied loads or displacements. This demonstrates the importance of using multiaxial mechanical tests in conjunction with computational methods as a means for establishing the parameters of constitutive models.

ACKNOWLEDGMENTS

The authors acknowledge funding from the National Institutes of Health AR- 050218.

ARTICLES OF FURTHER INTEREST

Collagen; Ligament Repair: Animal Models; Tissue Engineering of Ligament; Tissue Engineering of Tendons

REFERENCES

1. Adirim, T.A.; Cheng, T.L. Overview of injuries in the young athlete. *Sports Med.* **2003**, *33* (1), 75–81.
2. Kannus, P. Tendons—a source of major concern in competitive and recreational athletes. *Scand. J. Med. Sci. Sports* **1997**, *7* (2), 53–54.
3. Griffin, L.Y.; Agel, J.; Albohm, M.J.; Arendt, E.A.; Dick, R.W.; Garrett, W.E.; Garrick, J.G.; Hewett, T.E.; Huston, L.; Ireland, M.L.; Johnson, R.J.; Kibler, W.B.; Lephart, S.; Lewis, J.L.; Lindenfeld, T.N.; Mandelbaum, B.R.; Marchak, P.; Teitz, C.C.; Wojtys, E.M. Noncontact anterior cruciate ligament injuries: risk factors and prevention strategies. *J. Am. Acad. Orthop. Surg.* **2000**, *8* (3), 141–150.
4. Woo, S.L. Mechanical properties of tendons and ligaments. I. Quasi-static and nonlinear viscoelastic properties. *Biorheology* **1982**, *19* (3), 385–396.
5. Woo, S.L.; Orlando, C.A.; Camp, J.F.; Akeson, W.H. Effects of postmortem storage by freezing on ligament tensile behavior. *J. Biomech.* **1986**, *19* (5), 399–404.

6. Woo, S.L.; Debski, R.E.; Withrow, J.D.; Janauschet, M.A. Biomechanics of knee ligaments. *Am. J. Sports Med.* **1999**, *27* (4), 533–543.
7. Bonifasi-Lista, C.; Lake, S.P.; Small, M.S.; Weiss, J.A. Viscoelastic properties of the human medial collateral ligament under longitudinal, transverse and shear loading. *J. Orthop. Res.* **2005**, *23* (1), 67–76.
8. Butler, D.; Kay, M.; Stouffer, D. Comparison of material properties in the fascicle-bone units from human patellar tendon and knee ligaments. *J. Biomech.* **1986**, *19*, 425–432.
9. Bigliani, L.U.; Pollock, R.G.; Soslowky, L.J.; Flatow, E.L.; Pawluk, R.J.; Mow, V.C. Tensile properties of the inferior glenohumeral ligament. *J. Orthop. Res.* **1992**, *10* (2), 187–197.
10. Gupta, B.N.; Subramanian, K.N.; Brinker, W.O.; Gupta, A.N. Tensile strength of canine cranial cruciate ligaments. *Am. J. Vet. Res.* **1971**, *32* (1), 183–190.
11. Hannafin, J.A.; Arnoczky, S.P. Effect of cyclic and static tensile loading on water content and solute diffusion in canine flexor tendons: an in vitro study. *J. Orthop. Res.* **1994**, *12* (3), 350–356.
12. Woo, S.L.; Hollis, J.M.; Adams, D.J.; Lyon, R.M.; Takai, S. Tensile properties of the human femur-anterior cruciate ligament-tibia complex. The effects of specimen age and orientation. *Am. J. Sports Med.* **1991**, *19* (3), 217–225.
13. Woo, S.L.; Lee, T.Q.; Gomez, M.A.; Sato, S.; Field, F.P. Temperature dependent behavior of the canine medial collateral ligament. *J. Biomech. Eng.* **1987**, *109* (1), 68–71.
14. Woo, S.L.; Ohland, K.J.; Weiss, J.A. Aging and sex-related changes in the biomechanical properties of the rabbit medial collateral ligament. *Mech. Ageing Dev.* **1990**, *56* (2), 129–142.
15. Woo, S.L.; Orlando, C.A.; Gomez, M.A.; Frank, C.B.; Akeson, W.H. Tensile properties of the medial collateral ligament as a function of age. *J. Orthop. Res.* **1986**, *4* (2), 133–141.
16. Weiss, J.A. *A Constitutive Model and Finite Element Representation for Transversely Isotropic Soft Tissues*; Bioengineering, University of Utah: Salt Lake City, 1994.
17. Race, A.; Amis, A.A. The mechanical properties of the two bundles of the human posterior cruciate ligament. *J. Biomech.* **1994**, *27* (1), 13–24.
18. Pollock, R.G.; Wang, V.M.; Bucchieri, J.S.; Cohen, N.P.; Huang, C.Y.; Pawluk, R.J.; Flatow, E.L.; Bigliani, L.U.; Mow, V.C. Effects of repetitive sub-failure strains on the mechanical behavior of the inferior glenohumeral ligament. *J. Shoulder Elbow Surg.* **2000**, *9* (5), 427–435.
19. Lynch, H.A.; Johannessen, W.; Wu, J.P.; Jawa, A.; Elliott, D.M. Effect of fiber orientation and strain rate on the nonlinear uniaxial tensile material properties of tendon. *J. Biomech. Eng.* **2003**, *125* (5), 726–731.
20. Lee, T.Q.; Dettling, J.; Sandusky, M.D.; McMahon, P.J. Age related biomechanical properties of the glenoid-anterior band of the inferior glenohumeral ligament-humerus complex. *Clin. Biomech.* **1999**, *14* (7), 471–476.
21. Moore, S.M.; McMahon, P.J.; Azemi, E.; Debski, R.E. Bi-directional mechanical properties of the posterior region of the glenohumeral capsule. *J. Biomech.* **2005**, *38* (6), 1365–1369.
22. Moore, S.M.; McMahon, P.J.; Debski, R.E. Bi-directional mechanical properties of the axillary pouch of the glenohumeral capsule: implications for surgical repair. *J. Biomech. Eng.* **2004**, *126* (2), 284–288.
23. Quapp, K.M.; Weiss, J.A. Material characterization of human medial collateral ligament. *J. Biomech. Eng.* **1998**, *120* (6), 757–763.
24. Woo, S.L.; Akeson, W.H.; Jemmott, G.F. Measurements of nonhomogeneous, directional mechanical properties of articular cartilage in tension. *J. Biomech.* **1976**, *9* (12), 785–791.
25. Njus, G.; NM, N. A non-contact method for determining cross sectional area of soft tissues. *Orthop. Res. Soc.* **1986**, *11*, 126.
26. Race, A.; Amis, A.A. Cross-sectional area measurement of soft tissue. A new casting method. *J. Biomech* **1996**, *29* (9), 1207–1212.
27. Allard, P.; Thiry, P.S.; Bourgault, A.; Drouin, G. Pressure dependence of “the area micrometer” method in evaluation of cruciate ligament cross-section. *J. Biomed. Eng.* **1979**, *1* (4), 265–267.
28. Ellis, D. Cross sectional area measurements for tendon specimens—a comparison of several methods. *J. Biomech.* **1969**, *2*, 175–186.
29. Walker, L.; Harris, E.; Benedict, J. Stress strain relationship in human cadaveric plantaris tendon—a preliminary study. *Med. Elect. Biol. Eng.* **1964**, *2*, 31–38.
30. Ellis, D.G. Cross-sectional area measurements for tendon specimens: A comparison of several methods. *J. Biomech* **1969**, *2*, 175–186.
31. Iaconis, F.; Steindler, R.; Marinozzi, G. Measurements of cross-sectional area of collagen structures (knee ligaments) by means of an optical method. *J. Biomech.* **1987**, *20* (10), 1003–1010.
32. Lee, T.Q.; Woo, S.L. A new method for determining cross-sectional shape and area of soft tissues. *J. Biomech. Eng.* **1988**, *110* (2), 110–114.
33. Woo, S.L.; Danto, M.I.; Ohland, K.J.; Lee, T.Q.; Newton, P.O. The use of a laser micrometer system to determine the cross-sectional shape and area of ligaments: a comparative study with two existing methods. *J. Biomech. Eng.* **1990**, *112* (4), 426–431.
34. Chan, S.; Livesay, G.; Morrow, D.; Woo, S.L.-Y. The development of a low cost laser reflectance system to determine the cross sectional shape and area of soft tissues. *ASME Adv. Bioeng. San Francisco* **1995**, *31*, 123–124.
35. Beynon, B.D.; Fleming, B.C.; Johnson, R.J.; Nichols, C.E.; Renstrom, P.A.; Pope, M.H. Anterior cruciate ligament strain behavior during rehabilitation exercises in vivo. *Am. J. Sports Med.* **1995**, *23* (1), 24–34.
36. Beynon, B.; Howe, J.G.; Pope, M.H.; Johnson, R.J.; Fleming, B.C. The measurement of anterior cruciate ligament strain in vivo. *Int. Orthop.* **1992**, *16* (1), 1–12.
37. Harner, C.D.; Xerogeanes, J.W.; Livesay, G.A.; Carlin, G.J.; Smith, B.A.; Kusayama, T.; Kashiwaguchi, S.;

- Woo, S.L. The human posterior cruciate ligament complex: an interdisciplinary study. Ligament morphology and biomechanical evaluation. *Am. J. Sports Med.* **1995**, *23* (6), 736–745.
38. Scheffler, S.U.; Clineff, T.D.; Papageorgiou, C.D.; Debski, R.E.; Benjamin, C.; Woo, S.L. Structure and function of the healing medial collateral ligament in a goat model. *Ann. Biomed. Eng.* **2001**, *29* (2), 173–180.
 39. Lam, T.C.; Frank, C.B.; Shrive, N.G. Calibration characteristics of a video dimension analyser (VDA) system. *J. Biomech.* **1992**, *25* (10), 1227–1231.
 40. Smutz, W.P.; Drexler, M.; Berglund, L.J.; Growney, E.; An, K.N. Accuracy of a video strain measurement system. *J. Biomech.* **1996**, *29* (6), 813–817.
 41. Yin, F.C.; Tompkins, W.R.; Peterson, K.L.; Intaglietta, M. A video-dimension analyser. *IEEE Trans. Biomed. Eng.* **1972**, *19* (5), 376–381.
 42. Abramowitch, S.D.; Yagi, M.; Tsuda, E.; Woo, S.L. The healing medial collateral ligament following a combined anterior cruciate and medial collateral ligament injury—a biomechanical study in a goat model. *J. Orthop. Res.* **2003**, *21* (6), 1124–1130.
 43. Whittaker, P.; Canham, P.B. Demonstration of quantitative fabric analysis of tendon collagen using two-dimensional polarized light microscopy. *Matrix* **1991**, *11* (1), 56–62.
 44. Woo, S.L.; Peterson, R.H.; Ohland, K.J.; Sites, T.J.; Danto, M.I. The effects of strain rate on the properties of the medial collateral ligament in skeletally immature and mature rabbits: a biomechanical and histological study. *J. Orthop. Res.* **1990**, *8* (5), 712–721.
 45. Johnson, G.A.; Tramaglino, D.M.; Levine, R.E.; Ohno, K.; Choi, N.Y.; Woo, S.L. Tensile and viscoelastic properties of human patellar tendon. *J. Orthop. Res.* **1994**, *12* (6), 796–803.
 46. Figgie, H.E., 3rd; Bahniuk, E.H.; Heiple, K.G.; Davy, D.T. The effects of tibial-femoral angle on the failure mechanics of the canine anterior cruciate ligament. *J. Biomech.* **1986**, *19* (2), 89–91.
 47. Moon, D.; Woo, S.L.; Gabriel, M.; Abramowitch, S.; Takakura, Y. The effects of refreezing on the viscoelastic and tensile properties of ligaments. *Journal of Biomechanics* **2005**.
 48. Parry, D.A. The molecular and fibrillar structure of collagen and its relationship to the mechanical properties of connective tissue. *Biophys. Chem.* **1988**, *29* (1–2), 195–209.
 49. Debski, R.E.; Moore, S.M.; Mercer, J.L.; Sacks, M.S.; McMahon, P.J. The collagen fibers of the anteroinferior capsulolabrum have multi-axial orientation to resist shoulder dislocation. *J. Shoulder Elbow Surg.* **2003**, *12* (3), 247–252.
 50. Sacks, M.S.; Smith, D.B.; Hiester, E.D. A small angle light scattering device for planar connective tissue microstructural analysis. *Ann. Biomed. Eng.* **1997**, *25* (4), 678–689.
 51. Hirokawa, S.; Tsuruno, R. Three-dimensional deformation and stress distribution in an analytical/computational model of the anterior cruciate ligament. *J. Biomech.* **2000**, *33* (9), 1069–1077.
 52. Hurschler, C.; Loitz-Ramage, B.; Vanderby, R., Jr. A structurally based stress-stretch relationship for tendon and ligament. *J. Biomech. Eng.* **1997**, *119* (4), 392–399.
 53. Kohles, S.S.; Thielke, R.J.; Vanderby, R., Jr. Finite elasticity formulations for evaluation of ligamentous tissue. *Biomed. Mater. Eng.* **1997**, *7* (6), 387–390.
 54. Lanir, Y. A microstructure model for the rheology of mammalian tendon. *J. Biomech. Eng.* **1980**, *102* (4), 332–339.
 55. Puso, M.A.; Weiss, J.A. Finite element implementation of anisotropic quasi-linear viscoelasticity using a discrete spectrum approximation. *J. Biomech. Eng.* **1998**, *120* (1), 62–70.
 56. Simbeya, K.W.; Shrive, N.G.; Frank, C.B.; Matyas, J.R. *Recent Advances in Computer Methods in Biomechanics and Biomedical Engineering*; Gordon and Breach: Swansea, U.K., 1992; 240–249.
 57. Weiss, J.A.; Maker, B.N.; Govindjee, S. Finite element implementation of incompressible, transversely isotropic hyperelasticity. *Computer Methods Appl. Mech. Eng.* **1996**, *135*, 107–128.
 58. Weiss, J.A.; Gardiner, J.C.; Bonifasi-Lista, C. Ligament material behavior is nonlinear, viscoelastic and rate-independent under shear loading. *J. Biomech.* **2002**, *35* (7), 943–950.
 59. Gardiner, J.C.; Weiss, J.A. Subject-specific finite element analysis of the human medial collateral ligament during valgus knee loading. *J. Orthop. Res.* **2003**, *21* (6), 1098–1106.
 60. Yin, L.; Elliott, D.M. A biphasic and transversely isotropic mechanical model for tendon: application to mouse tail fascicles in uniaxial tension. *J. Biomech.* **2004**, *37* (6), 907–916.
 61. Vande Geest, J.P.; Sacks, M.S.; Vorp, D.A. The effects of aneurysm on the biaxial mechanical behavior of human abdominal aorta. *J. Biomech.* **2005**.
 62. Vande Geest, J.P.; Sacks, M.S.; Vorp, D.A. Age dependency of the biaxial biomechanical behavior of human abdominal aorta. *J. Biomech. Eng.* **2004**, *126* (6), 815–822.
 63. Sacks, M.S.; Sun, W. Multiaxial mechanical behavior of biological materials. *Ann. Rev. Biomed. Eng.* **2003**, *5*, 251–284.
 64. Sacks, M.S.; Chuong, C.J. Biaxial mechanical properties of passive right ventricular free wall myocardium. *J. Biomech. Eng.* **1993**, *115* (2), 202–205.
 65. Lanir, Y. Constitutive equations for the lung tissue. *J. Biomech. Eng.* **1983**, *105* (4), 374–380.
 66. Debski, R.E.; Weiss, J.A.; Newman, W.J.; Moore, S.M.; McMahon, P.J. Stress and strain in the anterior band of the inferior glenohumeral ligament during a simulated clinical examination. *J. Shoulder Elbow Surg.* **2005**, *14* (1 suppl S), 24S–31S.
 67. Wang, D.H.; Makaroun, M.S.; Webster, M.W.; Vorp, D.A. Effect of intraluminal thrombus on wall stress in patient-specific models of abdominal aortic aneurysm. *J. Vasc. Surg.* **2002**, *36* (3), 598–604.
 68. Song, Y.; Debski, R.E.; Musahl, V.; Thomas, M.; Woo, S.L. A three-dimensional finite element model of the human anterior cruciate ligament: a computational

- analysis with experimental validation. *J. Biomech.* **2004**, *37* (3), 383–390.
69. Stabile, K.J.; Pfaeffle, H.J.; Weiss, J.A.; Gabriel, M.T.; Tomanio, M.M.; Fischer, K.J. Longitudinal and transverse mechanical properties of the interosseous ligament of the forearm. *Bioengineering Conference ASME*, 2001; 363–364.
 70. Hewitt, J.; Guilak, F.; Glisson, R.; Vail, T.P. Regional material properties of the human hip joint capsule ligaments. *J. Orthop. Res.* **2001**, *19* (3), 359–364.
 71. Gardiner, J.C.; Weiss, J.A. Simple shear testing of parallel-fibered planar soft tissues. *J. Biomech. Eng.* **2001**, *123* (2), 170–175.
 72. Cohen, R.E.; Hooley, C.J.; McCrum, N.G. Viscoelastic creep of collagenous tissue. *J. Biomech.* **1976**, *9* (4), 175–184.
 73. Abramowitch, S.D.; Woo, S.L.; Clineff, T.D.; Debski, R.E. An evaluation of the quasi-linear viscoelastic properties of the healing medial collateral ligament in a goat model. *Ann. Biomed. Eng.* **2004**, *32* (3), 329–335.
 74. Provenzano, P.; Lakes, R.; Keenan, T.; Vanderby, R., Jr. Nonlinear ligament viscoelasticity. *Ann. Biomed. Eng.* **2001**, *29* (10), 908–914.
 75. Thornton, G.M.; Oliynyk, A.; Frank, C.B.; Shrive, N.G. Ligament creep cannot be predicted from stress relaxation at low stress: a biomechanical study of the rabbit medial collateral ligament. *J. Orthop. Res.* **1997**, *15* (5), 652–656.
 76. Fung, Y.C. *Biomechanics: Mechanical Properties of Living Tissues*, 2nd Ed.; Springer: New York, NY, 1993.
 77. Thomopoulos, S.; Williams, G.R.; Gimbel, J.A.; Favata, M.; Soslowsky, L.J. Variation of biomechanical, structural, and compositional properties along the tendon to bone insertion site. *J. Orthop. Res.* **2003**, *21* (3), 413–419.
 78. Elliott, D.M.; Robinson, P.S.; Gimbel, J.A.; Sarver, J.J.; Abboud, J.A.; Lozzo, R.V.; Soslowsky, L.J. Effect of altered matrix proteins on quasilinear viscoelastic properties in transgenic mouse tail tendons. *Ann. Biomed. Eng.* **2003**, *31* (5), 599–605.
 79. Kwan, M.K.; Lin, T.H.; Woo, S.L. On the viscoelastic properties of the anteromedial bundle of the anterior cruciate ligament. *J. Biomech.* **1993**, *26* (4–5), 447–452.
 80. Funk, J.R.; Hall, G.W.; Crandall, J.R.; Pikley, W.D. Linear and quasi-linear viscoelastic characterization of ankle ligaments. *J. Biomech. Eng.* **2000**, *122* (1), 15–22.
 81. Abramowitch, S.D.; Woo, S.L. An improved method to analyze the stress relaxation of ligaments following a finite ramp time based on the quasi-linear viscoelastic theory. *J. Biomech. Eng.* **2004**, *126* (1), 92–97.
 82. Woo, S.L.; Gomez, M.A.; Akeson, W.H. The time and history-dependent viscoelastic properties of the canine medial collateral ligament. *J. Biomech. Eng.* **1981**, *103* (4), 293–298.
 83. Provenzano, P.P.; Lakes, R.S.; Corr, D.T.; R, R., Jr. Application of nonlinear viscoelastic models to describe ligament behavior. *Biomech. Model Mechanobiol.* **2002**, *1* (1), 45–57.
 84. Sarver, J.J.; Robinson, P.S.; Elliott, D.M. Methods for quasi-linear viscoelastic modeling of soft tissue: application to incremental stress-relaxation experiments. *J. Biomech. Eng.* **2003**, *125* (5), 754–758.
 85. Findley, W.N.; Lai, J.S.; Onaran, K. *Creep and Relaxation of Nonlinear Viscoelastic Materials*; Dover: New York, 1973.
 86. Myers, B.S.; McElhaney, J.H.; Doherty, B.J. The viscoelastic responses of the human cervical spine in torsion: experimental limitations of quasi-linear theory, and a method for reducing these effects. *J. Biomech.* **1991**, *24* (9), 811–817.
 87. Nigul, I.; Nigul, U. On algorithms of evaluation of Fung's relaxation function parameters. *J. Biomech.* **1987**, *20* (4), 343–352.

Thrombosis

Mukul S. Goel
Scott L. Diamond

*Institute for Medicine and Engineering, Department of Chemical Engineering,
University of Pennsylvania, Philadelphia, Pennsylvania, U.S.A.*

T

INTRODUCTION

Blood coagulation is initiated by the disturbance of the hemostatic balance following injury, inflammation, endothelial dysfunction, or contact with an artificial surface and involves significant interplay between activation of the enzymatic and cellular processes, adhesion of blood cells, their aggregation, and hemodynamics. Clotting on the inner wall of blood vessels is the major cause of cardiovascular diseases such as myocardial infarction, stroke, and deep vein thrombosis. The formation of homotypic and heterotypic aggregates of activated platelets and neutrophils and their adhesion through selectins, GPIb, and integrin-mediated pathways enhances thrombin generation and subsequent fibrinogen polymerization under coagulating flow conditions. At the same time, cell adhesion and aggregation, typically requiring cell activation, are also influenced by proteins of the blood coagulation cascade.

In addition to platelets, leukocytes and erythrocytes have roles to perform in thrombosis; how they interact with platelets under coagulating flow conditions to stimulate and enhance thrombotic processes has been less well understood. This article seeks to highlight how individual platelets and neutrophils, neutrophil-platelet interactions, and interactions of erythrocytes with these cells activate and promote thrombosis following contact of blood with a defined surface. We intend to provide a fundamental review relevant to biomaterial thrombosis and inflammation-coagulation crosstalk.

COAGULATION BIOLOGY

The Reaction Cascade

Distinct from biomaterial thrombosis, the tissue factor pathway of the coagulation cascade is activated as blood comes in contact with tissue factor following vessel-wall injury. Upon contact with blood, tissue factor forms a complex with endogenous factor VIIa that catalyzes the activation of factor X to Xa. After Xa is formed, it binds factor Va released from platelet

stores. The prothrombinase complex (XaVa) on platelets activates prothrombin to thrombin, which cleaves fibrinogen to fibrin monomer. Fibrin monomer polymerizes to form fibrin polymer, which is cross-linked by thrombin-activated factor XIIIa.

The coagulation cascade can also be triggered on the contact of blood with the subendothelium tissues or an artificial surface. This pathway, known as the contact (or intrinsic) pathway of coagulation, involves formation of factor Xa from a series of reactions initiated by factor XII activation on artificial surfaces. Once formed, factor XIIa can activate prekallikrein to kallikrein to amplify factor XIIa formation. Factor XIIa then activates factor XI to XIa, which, in turn, activates factor IX to IXa. Activated platelets participate in the assembly of the intrinsic tenase (IXa/VIIIa) and prothrombinase (Xa/Va) complexes by providing anionic phospholipid binding sites for coagulation proteins. Following an initial production of minuscule amounts of thrombin through the tissue factor pathway, factor XI may be activated to XIa on activated platelets by thrombin to trigger the intrinsic pathway in the absence of the upstream contact pathway proteins, factor XII, prekallikrein, and high molecular weight kininogen.^[1] To amplify coagulation, thrombin can activate factor V and factor VIII (to Va and VIIIa), and factor Xa can activate factor VII and factor V (to VIIa and Va). By releasing factor V, which may even be partially active,^[2] activated platelets contribute to prothrombinase complex formation for an amplified production of thrombin.

While considerable interplay exists between extrinsic and intrinsic pathways (e.g. TF: VIIa activation of factor IX) and the two pathways are not independent, the concept of TF-dependent and non-TF-dependent initiation of factor Xa production remains an important distinction. Formation of factor Xa is generally viewed as the rate-controlling step in thrombin production.

Platelet Function

Platelet activation, adhesion, and aggregation play a major role in coagulation events. Platelets can be

activated by pathological high shear flow ($>3000\text{ s}^{-1}$) or by agonists (e.g., adenosine diphosphate (ADP), thrombin, thromboxane, collagen) released or produced as a result of coagulation reactions or exposed to the blood. The full activation of the platelets involves shape change and release of ADP, as well as activation of glycoprotein IIb/IIIa receptors. The aggregation of platelets is mediated through fibrinogen crossbridging of GPIIb/IIIa and von Willebrand Factor (vWF) interactions with GPIb α -IX-V (i.e., GPIb) and GPIIb/IIIa. Once adherent, platelets can recruit other platelets from flowing blood through fibrinogen crossbridging of GPIIb/IIIa and vWF interactions with GP IbIX and GP IIb/IIIa. Fibrinogen binds to GPIIb/IIIa in a specific, saturable manner via the dodecapeptide sequence $\gamma_{400-411}$ on the D domain of fibrinogen. Upon platelet activation, GPIIb/IIIa receptors undergo conformational changes to become competent receptors for soluble fibrinogen and plasma proteins such as vWF. Fibrinogen binds to GP IIb/IIIa receptors (about 50,000 binding sites/platelet) with a dissociation constant (K_d) $\sim 0.1\ \mu\text{M}$.

The platelet GP IIb/IIIa-fibrinogen interaction is not fully understood and important questions remain concerning the interaction of resting platelets with surface-immobilized fibrinogen. Resting platelets do not appear to bind to adsorbed fibrinogen during short-lived ($<50\text{ m sec}$) collisions with fibrinogen-coated beads in a shear flow.^[3] However, in whole-blood flow over fibrinogen-coated surfaces, the fibrinogen may be sufficient to capture free-flowing, resting platelets at low shear rates via the $\gamma_{400-411}$ but not Arg-Gly-Asp (RGD) sequences in the carboxy terminus of the α -chain of fibrinogen. These differences may be due to: 1) more favorable cell adhesion mechanics due to wall-directed red blood cell (RBC) particle forces; 2) low-level activation near the wall from RBC-released ADP; 3) differences in contact-area development and force mechanics of doublet formation in suspension flow versus cell-wall interactions; 4) surface-mediated accumulation over several minutes of ~ 0.1 to 1% of activated platelets in drawn anticoagulated blood; and/or 5) surface-dependent conformational changes of adsorbed fibrinogen. Does GPIIb/IIIa-dependent platelet capture to a growing thrombus require platelet activation? This is a fundamental issue in the dynamics of thrombus growth. The measurement of probabilities of capture, firm arrest, and detachment as a function of interaction time often requires high temporal resolution exceeding that of standard video microscopy. Given the potential for multivalent interaction of polymeric vWF to bind GPIb on resting platelets by accumulating monovalent interactions of micromolar affinity, it also remains unexplained why vWF does not aggregate resting platelets under normal flow. Enhancement of vWF

function at various shear rates might be a result of alteration in soluble vWF structure or aggregation state; the GPIb avidity state; and/or the collisional mechanics of doublet formation, a hydrodynamic threshold effect, as in Finger et al.^[4] and Lawrence et al.^[5] Doggett et al. observed that a critical amount of flow above 70 s^{-1} is needed for platelet GPIb binding to recombinant vWF-A1 domain.^[6]

Neutrophil Function

The neutrophils may participate in coagulation by various potential mechanisms. While neutrophils may promote blood coagulation by the deencryption of cell-surface tissue factor^[7] or the transfer of tissue factor to platelets through microparticles,^[8] they may also release proteases that activate platelets^[9] or bind factor X via Mac-1 (CD11b/CD18).^[10] Neutrophil elastase may take an active role in the deencryption of tissue factor in the plasma by cleaving tissue-factor pathway inhibitor (TFPI),^[11] and this deencrypted tissue factor may become associated on the surface of platelets^[12] to promote fibrin formation.

In the baboon femoral arteriovenous shunt Dacron graft model,^[13] monoclonal antibody GA6 against P-selectin reduced indium-labeled leukocyte accumulation by 60% over a 2 hr time frame. GA6 Fab2' fragment reduced leukocyte accumulation by 80% over 2 hr. Anti-P-selectin had no effect on platelet accumulation in the Dacron graft, but dramatically reduced fibrin accumulation by 70%. A marked reduction of fibrin deposition by anti-P-selectin was seen as early as 5 minutes after the initiation of flow. The authors indicated that monocyte and leukocyte expression of tissue factor might be the cause of this increased fibrin deposition. This explanation does not explain the dramatic reduction of fibrin accumulation by anti-P-selectin compared to control at 5 minutes, because expression of tissue factor requires gene activation, transcription, synthesis, and display—a sequence of events that results in elevated tissue factor on monocytes only after 20 minutes. Neutrophils may elevate the production of thrombin at the site of the clot via cross-sectional effects for enhanced platelet capture, tissue factor expression, Mac-1 binding of factor X, the release of proteases such as elastase and cathepsin G, or the display of anionic phospholipid for coagulation pathway assembly. The relative roles of these potential pathways are not fully understood at various flow conditions.

Neutrophil homotypic interactions

Neutrophils present L-selectin on the tips of the microvilli. ICAM-3 and L-selectin are present on

resting neutrophils in a competent state for binding. Both L-selectin and ICAM-3 are shed rapidly following neutrophil activation. Upon activation, only about 10% of displayed Mac-1 and LFA-1 is active. LFA-1 (CD11a/CD18) binding to ICAM-3, as well as Mac-1 (CD11b/CD18) binding to an unknown receptor, can facilitate integrin-mediated homoaggregation at low shear rates between 100 and 400 s⁻¹. Above 100 s⁻¹, L-selectin plays an important role in fMLP-stimulated neutrophil homoaggregation. The collision efficiency increases with shear rate to a maximum of $\epsilon=0.4$ at 400 s⁻¹ and then declines to $\epsilon=0.15$ at 3000 s⁻¹.^[14] This is consistent with observations of CD62P,E,L-mediated rolling displaying a maximum with shear stress, likely due to shear or stress-mediated delivery of microvilli to the rolling contact area^[5] or cellular deformation at high flow. Direct imaging of the neutrophil rolling contact area has been achieved^[15] to indicate the role of neutrophil membrane deformation and membrane tether formation in regulating adhesion to P-selectin. Similarly, increases in neutrophil membrane rigidity can affect P-selectin-PSGL-1 bond cluster lifetime due to reduction in microvilli extension.^[16]

Platelet–neutrophil heterotypic interactions

Recently, the laboratory of Yale Nemerson demonstrated that an inhibitor of tissue factor can greatly reduce blood clotting on purified collagen surfaces lacking tissue factor.^[7] This data strongly supports the role of blood-borne tissue factor that is potentially in a cryptic state on neutrophils but is exposed upon activation. Other non-tissue factor mechanisms of neutrophil activation of blood clotting have been described but not prioritized quantitatively. Apoptotic neutrophils or activated neutrophils can display anionic phospholipids, detectable by annexin V binding, which may serve as a cofactor in coagulation factor assembly. Due to the massive size of the neutrophil surface, a few activated neutrophils may be potent catalytic participants in platelet accumulation.^[17]

Elastase is known to cause exposure of GPIIb/IIIa on platelets by surface proteolysis.^[18] Similarly, cathepsin G can activate platelets. Yamamoto also noted the prothrombotic tendency of leukocytes in coagulation assays that was antagonized by the neutrophil protease inhibitor, eglin C.^[19] They noted that the procoagulant tendency of activated leukocytes outweighed antiplatelet effects due to NO production or arachidonic acid metabolite exchange with platelets. Neutrophil cathepsin G can also activate factor X bound to Mac-1.^[10] The competition and kinetic efficacy of factor X, fibrinogen, and fibrin-binding Mac-1, as well as the prevailing levels of cathepsin

G or elastase under flow conditions, remain largely unquantified from a kinetic point of view.

Coagulation and fibrin deposition under flow

De Groot et al.^[20] used a calibrated staining assay to measure deposition of peroxidase-labeled fibrin on endothelial cell matrix (containing tissue factor) from heparinized whole blood. They found that fibrinopeptide A generation was not a function of shear rate at 300 or 1300 s⁻¹. Fibrin deposition to the surface was greater at 300 s⁻¹ than at 1300 s⁻¹. Fibrin monomer (fibrinogen des-A) concentration was greater in the perfusion fluid at 1300 s⁻¹. In this study, the perfused blood was anticoagulated with heparin, which may emphasize the role of fibrin-bound thrombin that is resistant to heparin-mediated inhibition. The impaired fibrin polymerization at high shear was suggested to be due to increased removal of the monomer from the surface, but the physical hydrodynamics of this are poorly understood. This study is in contrast to a study with nonanticoagulated whole blood perfused over partially denuded rabbit aortas where platelet and fibrin deposition both increased with shear rates up to 1500 s⁻¹.^[21] Fibrin deposition and its shear dependency are highly sensitive to the presence or absence of cellular contributions and prevailing anticoagulation conditions. Neutrophil adhesion to preformed endothelial matrix and fibrin under shear flow was studied recently by Kuijper et al.^[22] Although neutrophils can adhere to fibrin via Mac-1 (CD11b/CD18) at shear stresses up to 20 dynes/cm², soluble fibrinogen inhibited the interaction. Fibrin-mediated steric hindrance of neutrophil attachment to TNF- α -treated endothelium was noted by Kirchofer et al.^[23] As a bulky structure, it is possible that fibrin can prevent selectin interactions with endothelium that would normally facilitate rolling and subsequent firm adhesion.

A potential role for red-blood-cell adhesion in thrombosis

Palabrica et al.^[13] noted that anti-P-selectin appeared to reduce, through an unknown mechanism, the content of red blood cells on a Dacron-graft thrombus as indicated by scanning electron microscopy of the clot surfaces. This observation raises the issue of RBCs binding to clots or activated leukocytes. The RBCs help push platelets to the reactive wall, as well as leak ADP under high-shear conditions. While numerous studies have looked at thrombospondin and CD36-dependent red-blood-cell adhesive dynamics during sickle cell anemia, malaria, and diabetes, little is known about the adhesive interaction of normal RBC

during thrombosis. Treatment of healthy RBC with $2\ \mu\text{M}$ a23817 and $0.5\ \text{mM}$ CaCl_2 causes the loss of lipid asymmetry in up to 40% of the RBC as indicated by phosphatidylserine (PS) exposure detectable by FITC-annexin V binding.^[24] In their work, CD36 was not detected on ionophore-treated normal RBC. PS exposure by RBC is sufficient for expression of procoagulant activity, as well as alternative complement pathway activation as indicated by C3b deposition.^[25] It has been shown that PS on RBC was sufficient for the binding of RBC to endothelial matrix or purified thrombospondin. Interestingly, even normal RBC could bind immobilized thrombospondin at a level \sim ten-fold greater than RBC binding to albumin. This suggests that RBC-binding thrombospondin may be important in the adhesion of normal RBC to activated spread platelets that release thrombospondin and display CD36. Platelet-released thrombospondin is an important mediator of platelet aggregation via a CD36-dependent mechanism. A role for platelet-RBC interaction in sickle cell disease has been reported to be partially mediated by thrombospondin.^[26] Platelet-RBCs are routinely observed (and ignored) in flow cytometry of anticoagulated whole blood. RBC captured in the early thrombus may display PS due to alternative complement pathway activation, and thus become procoagulant and adhesive for platelet thrombospondin. These mechanisms may be especially operative at low-shear rates if binding is extremely weak and transient.

Complement System/Activation

The complement pathway also plays a role in biomaterial function, thrombosis, and biomaterial infections. Bacterial killing in coagulating blood involves numerous interactions among coagulation factors, platelets, neutrophils, bacteria, and the complement system. Infections in coagulated blood present an important class of clinical complications in the context of artificial biomaterials, infective endocarditis, mechanical and bioprosthetic heart valve endocarditis, and wound healing. Yet, little is known about the role and dynamics of various biochemical or biological participants in regulating bacterial survival during blood coagulation. Through the classical or alternative pathway, complement-based lysis of bacteria depends on the generation of C3b and C5b on bacterial surfaces, thus triggering the formation of the membrane attack complex (C5b-9). The classical pathway is calcium and magnesium-dependent, while the alternative pathway is magnesium-dependent. Gram-negative bacteria such as *Escherichia coli* are potent activators of the alternative pathway. Also, C3b and iC3b act as opsonins; C3b binds neutrophil complement receptor

CR1, while iC3b (but not C3b) binds neutrophil CR3 (Mac-1) and CR4.

Early work on platelet activation and aggregation in the presence of complement has demonstrated that thrombin can act on the platelet surface as a low-activity C3 convertase and high-activity C5 convertase. In solution, thrombin cleaves C3 (presumably at the Arg₆₉-Ala₇₀ bond) to produce fragments roughly the size of C3a and C3b that have no anaphylatoxin or hemolytic activity, respectively. Thrombin can very rapidly cleave C5 to a C5b-like fragment and a C5a-like fragment that lacks chemotactic activity unless further digested by thrombin for 16 hr. In fact, levels of anaphylatoxin antigens C3a, C4a, and C5a are ten-, four-, and 60-fold higher, respectively, in freshly prepared serum (from whole blood maintained for 30 min at room temperature in glass tubes) than in EDTA-treated plasma, demonstrating that coagulation can activate complement. However, this level of generation of C3a, C4a, and C5a in serum represents only a \sim 3 to 10% conversion of C3, C4, and C5. Also, thrombin can cleave C9 to C9a and C9b, which remain associated without loss of hemolytic function.

BIOMATERIAL THROMBOSIS

Fibrin Formation by Adherent Neutrophils

Implantation of biomaterials is generally followed by adsorption of plasma proteins such as albumin, fibrinogen, vWF, and vitronectin onto the artificial surface. Neutrophils, via Mac-1 (CD11b/CD18), can adhere to immobilized fibrinogen. The regional neutrophil density can increase by an inflammatory response that may even be triggered by biomaterial-adsorbed fibrinogen.^[27] Once adherent, neutrophils can mediate the capture of flowing neutrophils near the surface via a L-selectin-PSGL-1 (P-selectin glycoprotein ligand 1) mechanism.

Our studies have shown that under venous flow conditions, individual surface adherent neutrophils can promote fibrin deposition in the presence of the intrinsic coagulation protein factor XIIa.^[28] On the perfusion of platelet-free plasma (PFP) over fibrinogen-adherent neutrophils at a wall shear rate (γ_w) of $62.5\ \text{s}^{-1}$, dense fibrin deposition in close proximity to the neutrophils, especially downstream of each cell, was observed (Fig. 1A). The time required for this fibrin deposition on neutrophils was prolonged by 40 to 50 min by inhibition of factor XIIa with corn trypsin inhibitor (CTI), demonstrating the role of contact pathway in this fibrin accumulation around neutrophils. Fig. 1B illustrates no fibrin formation at $t = 35\ \text{min}$ in the presence of CTI.

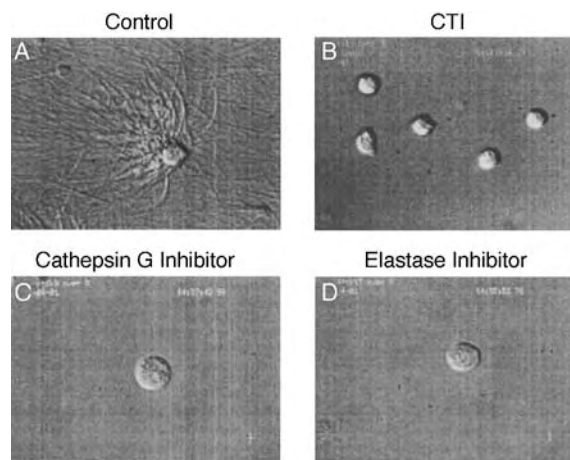


Fig. 1 XIIa-dependent fibrin formation around neutrophils. As recalcified citrated PFP was perfused over surface-adherent neutrophils at a shear rate of 62.5 s^{-1} for 20 min, dense fibrin deposition was observed around each neutrophil, especially downstream of each cell (A). No fibrin was formed even at $t = 35$ min in the presence of CTI (B). Fibrin accumulation was attenuated when either cathepsin G inhibitor (C) or elastase inhibitor (D) was present in the PFP being perfused over neutrophils (also treated with the matching inhibitor). Flow is from right to left.

One of the mechanisms through which individual fibrinogen-adherent neutrophils promote fibrin deposition is capture of short fibrin protofibrils flowing in recalcified plasma in contact with in-vitro surfaces. This mechanism may involve Mac-1 (CD11b/CD18)-mediated capture and, to a much lesser extent, capture by nonspecific cross-sectional capture effects that is analogous to particle capture by a sphere under flow. However, fibrin capture is not the sole process responsible for fibrin deposition. In addition to blockade of Mac-1 and inhibition of XIIa, the inhibition of either one of the two neutrophil proteases, cathepsin G and elastase (Fig. 1C and 1D), also attenuated fibrin formation around adherent neutrophils.^[28] This demonstrated that fibrinogen-adherent neutrophils can promote FXIIa-dependent thrombin generation and subsequent fibrin formation in a platelet-independent manner through pathways mediated by the released proteases elastase and cathepsin G.

The reduction in fibrin formation by either the blockade of Mac-1, inhibition of XIIa, or inhibition of either one of two proteases, cathepsin G and elastase, may be attributed to the suppression of prothrombinase formation (factor Xa and factor Va) on the neutrophil surface (Fig. 2). Because cathepsin G has been found to activate Mac-1-bound factor X,^[10] decline in factor Xa levels can be expected by inhibition of cathepsin G or the blockade of Mac-1.

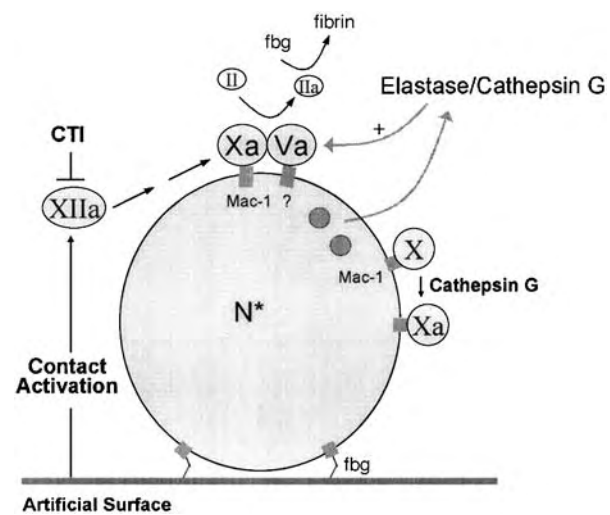


Fig. 2 Procoagulant activity on the neutrophil surface. Individual neutrophils may enhance coagulation in a platelet-independent manner through prothrombinase formation on the neutrophil surface. Abbreviations: fbg, fibrinogen; CTI, corn trypsin inhibitor; XIIa, factor XIIa; X, factor X; Xa, factor Xa; Va, factor Va; II, prothrombin; IIa, thrombin; N*, activated neutrophil.

Moreover, inhibition of elastase (or cathepsin G) may suppress factor V activation by these proteases,^[29] resulting in reduced levels of factor Va. Inhibition of factor XIIa would reduce the formation of the prothrombinase complex (XaVa) by reducing XIa and consequently IXa and then Xa production.

While blockage of Mac-1 attenuated fibrin formation on neutrophils, it is difficult to distinguish fully the role of Mac-1 on fibrin capture versus its role in thrombin formation, because Mac-1 binds fibrin(ogen) and factor X to the neutrophil.^[10] Since neutrophils can bind high molecular-weight kinogen (which can bind prekallikrein and factor XI), as well as factor XII,^[30] antibodies against CD11b/CD18 will clearly interfere with fibrin binding but may also attenuate pathways leading to Xa and thrombin formation on the neutrophil surface. Also, factor XIIa is a known activator of neutrophils, and CTI would be expected to attenuate kallikrein-mediated release of elastase by neutrophils exposed to recalcified plasma.^[31]

In the absence of platelet deposition and red-cell motion, fibrin accumulation over neutrophils in the presence of contact activation was inversely correlated with wall shear rate.^[28] The density of fibrin deposited around each neutrophil at a wall shear rate of 250 s^{-1} was significantly lower than that at a wall shear rate of 62.5 s^{-1} . Under the conditions present in these experiments, an increase in flow rate may decrease the time available for interaction between the Mac-1 receptors and the complementary fibrin domain or enhance the off-rate of binding. Alternatively, dilution

T

effects of key factors (e.g., Va, Xa, or thrombin) may be enhanced at high flow. Under physiological conditions, the presence of red cells and their motion would considerably influence coagulation. Their presence would increase the collision frequency of cells with the wall to enhance platelet aggregation and platelet deposition, respectively. Additionally, rouleaux formation would support vessel-wall adhesion of neutrophils through their outward displacement under flow.^[32]

Neutrophil Enhancement of Platelet-Dependent Coagulation

Platelets can adhere to plasma proteins immobilized on biomaterials through the same mechanisms that are involved in their adhesion to the vessel wall. However, unlike collagen, fibrinogen provides a surface onto which platelets can spread, but do not transform into fully procoagulant structures on their own. Our results have shown that adherent neutrophils can interact with fibrinogen-adherent platelets to turn them into activated procoagulant structures.^[28] As platelet-free plasma was perfused over a neutrophil-coated surface, a platelet-coated surface, and a surface coated with both neutrophils and platelets, dense fibrin deposition was observed only on the latter surface suggesting that neutrophil-platelet aggregation and adhesion can have profound roles in promoting coagulation under venous flow conditions even in the presence of factor XIIa-inhibition by CTI. This fibrin formation was reduced by specific peptide inhibitors against elastase and cathepsin G, indicating that the interaction of neutrophils with platelets to facilitate fibrin formation is mediated through these neutrophil proteases.

As we compared fibrin formation on fibrinogen-adherent neutrophil-platelet cell mixtures to that on collagen-adherent platelets (no neutrophils), we found the amounts of fibrin deposited to be equivalent (Fig. 3); neutrophils on a fibrinogen-coated surface had enhanced platelet-dependent fibrin formation to the levels supported by platelets alone on a collagen surface. Because fibrinogen-adherent platelets, unlike collagen-adherent platelets, are not fully activated procoagulant structures, these observations demonstrated that neutrophils, even in the absence of collagen, can activate platelets on their own and subsequently turn them into fully coagulating structures.

Rather than being localized on and near neutrophils, the formed fibrin deposited uniformly over the entire platelet surface of the flow chamber. This observation suggests that neutrophil-promoted fibrin formation on platelets is a consequence of some intercellular signaling event and/or catalysis of coagulation biochemistry between neutrophils and

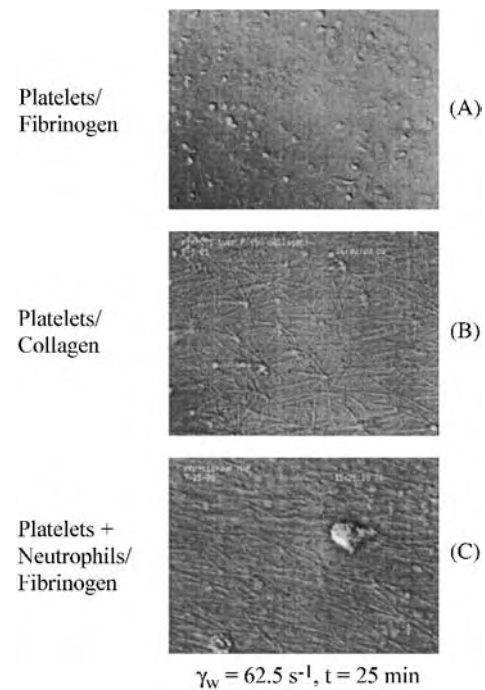


Fig. 3 Neutrophil promotion of fibrin formation in a platelet-dependent manner. On the perfusion of PFP at a shear rate of 62.5 s^{-1} for 25 min, while collagen-adherent platelets (B) supported more fibrin formation than fibrinogen-adherent platelets (A), neutrophils enhanced fibrin formation on fibrinogen-coated platelets (C) to levels supported by collagen-adherent platelets.

platelets that is mediated by elastase/cathepsin G. Several pathways exist by which neutrophils may trigger thrombin production and subsequent fibrin formation in a platelet-dependent manner, including: 1) elastase^[33] or cathepsin G^[9] enhancement of activation of spread platelets; or 2) elastase/cathepsin G cleavage of plasma zymogens such as factor V^[29] and/or X.^[10] While mechanisms of P-selectin-mediated adhesion under flow are eliminated in these studies, P-selectin-dependent platelet-neutrophil signaling did not appear to play a role in the process. We have also found that among the two neutrophil proteases, cathepsin G has a more potent role in enhancing platelet dependent coagulation.^[34] A coupled activity by both cathepsin G and elastase is not a prerequisite to achieve full activation of fibrinogen-adherent platelets; neutrophil cathepsin G alone can fully activate platelet-dependent coagulation. Moreover, platelet activation is the dominant mechanism through which cathepsin G turns fibrinogen-adherent platelets procoagulant. Using fluorimeter assay, the burst time (t_{50} , the time to reach 50% conversion of a fluorogenic thrombin substrate) for thrombin generation was measured to quantify the effect of elastase and cathepsin G on coagulation.^[34] While the addition of cathepsin G to PRP reduced the peak thrombin

Table 1 Effects of cathepsin G and elastase on thrombin generation

Environment		Burst time (t_{50}), min ^a
PRP, CTI	No protease	53 ± 1
	Cathepsin G (100 μM)	18 ± 1
	Elastase (100 μM)	53 ± 19
	Cathepsin G + Elastase	17 ± 1
PFP, CTI	Cathepsin G + Elastase	> 120

PRP=platelet-rich plasma, PFP=platelet-free plasma; CTI (corn trypsin inhibitor) to inhibit factor XIIa from the contact system.

^aData shown represent the means ± S.D. of the observations.

generation from 53 min to 18 min, the addition of elastase did not make a difference (Table 1).

The addition of phosphatidylserine/phosphatidylcholine vesicles (PSPC) alone to CTI-treated platelet-free plasma produced little thrombin, indicating that mere phosphatidylserine exposure during platelet activation is not sufficient for coagulation of the plasma.^[34] When neutrophil elastase was added to PFP containing PSPC, thrombin generation was also minimal. However, a detectable level of slow thrombin production was observed when cathepsin G or elastase/cathepsin G was added to PFP/CTI containing PSPC. These observations demonstrated that the presence of a negatively charged phospholipid surface is not sufficient to initiate coagulation. Cathepsin G is cationic and the slight increase in thrombin levels on addition of PSPC supports the view that cathepsin G-mediated cleavage of coagulation factors on a negatively charged surface may help support, in a secondary role, cathepsin G-mediated platelet activation. In comparison to cathepsin G-enhanced thrombin generation in PFP containing PSPC, a five times higher thrombin production in PRP (presence of platelets) indicated that a platelet has additional means, beyond merely presenting negative phospholipid, through which it responds to cathepsin G stimulation. These results were in agreement with studies by Sumner et al.,^[35] which show that PS exposure does not correlate with factor Xa or thrombin production on platelets and suggest that surface participants other than PS are involved in coagulation.^[36] Whereas display or release of active tissue factor by neutrophils^[7,8] to induce coagulation is possible, tissue factor did not appear to have a role in this process.^[28,34]

Adhesion of Normal Erythrocytes

While it is known that sickle and infected erythrocytes can directly adhere to the endothelium, a common

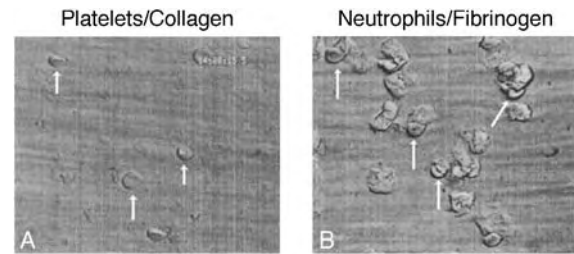


Fig. 4 Erythrocyte capture by activated platelets and neutrophils. As washed red cells were perfused over collagen-adherent platelets (A) or fibrinogen-adherent fMLP-treated neutrophils (B) at a wall shear rate of 50 s^{-1} , firm adhesion of red cells (arrows) was observed.

assumption is that normal red blood cells are passive during coagulation events and become entrapped in fibrin. We have examined capture and adhesion of normal erythrocytes to surface-adherent neutrophils and platelets under low-flow conditions as a mechanism of RBC accumulation distinct from passive entrapment within fibrin fibers.^[37] As washed erythrocytes were perfused over collagen-adherent platelets at a wall shear rate (γ_w) of 50 s^{-1} for 5 min to investigate the interactions between erythrocytes and platelets under flow, many events of red cell capture, transient tethering, and/or firm adhesion (Fig. 4A) to platelets were observed. To examine the dependence of erythrocyte adhesion on platelet activation, erythrocyte adhesion on collagen-adherent platelets was compared to that over fibrinogen-adherent platelets. The density of red cells captured was four-fold higher on collagen-adherent platelets, demonstrating that red-cell adhesion to platelets at depressed venous flow conditions is regulated by the platelet activation state. Moreover, red-cell adhesion to fibrinogen-adherent platelets that had been directly stimulated with the GPVI agonist convulxin was 2.7-fold higher than to unstimulated fibrinogen-adherent platelets.^[37] This verified the role of platelet activation in mediating these red-cell adhesion events.

Experiments in which potential mediators of these interactions, such as P-selectin, CD36, GPIb, GP IIb/IIIa, TSP, or vWF, were blocked by antibodies demonstrated partial roles for CD36 and GPIb in adhesion of red cells to activated platelets.^[37] When fibrinogen (3 mg/ml) was added to the red cells prior to their perfusion over platelets, cell adhesion was significantly blocked. This demonstrated that at physiological concentrations, fibrinogen plays a protective role in preventing red-cell adhesion to collagen-activated platelets. Red-cell adhesion to platelets was Ca^{2+} - or Mg^{2+} -dependent.

When washed red cells were perfused over fibrinogen-adherent neutrophils at a wall shear rate of 50 s^{-1} for 5 min, a few transient adhesion events were

observed between red cells and fibrinogen-adherent neutrophils. However, when neutrophils were activated with N-formyl-Met-Leu-Phe (fMLP), the number of pausing events and firmly adherent red cells (Fig. 4B) to these activated neutrophils dramatically increased. fMLP caused a sixfold increase in RBC adhesion to neutrophils. Erythrocyte adhesion to fMLP-activated neutrophils at depressed venous flows, unlike RBC adhesion to collagen-activated platelets, was not blocked by soluble fibrinogen. Adhesive interactions between erythrocytes and activated neutrophils were found to be mediated by Mac-1 (CD11b/CD18). One of the receptors on the red cell was detected to be ICAM-4 (LW blood group). Blocking antibodies against CD36 or TSP did not show any effect. Consistent with β_2 -integrin function, the adhesion of erythrocytes to activated neutrophils was markedly inhibited by EDTA.

From adhesion assays conducted at $\gamma_w = 25, 50, 75$ and 100 s^{-1} , we observed that the number of adherent red cells to collagen-adherent platelets decreased as the shear rate increased.^[37] These observations indicate that the RBC adhesion to platelets is significant only at depressed venous flow conditions (below 100 s^{-1}). Similarly, RBC adhesion to fMLP-treated neutrophils was inversely correlated to shear rate, and it was significantly reduced at γ_w above 100 s^{-1} .

While it is known that vascular compression devices help prevent deep vein thrombosis (DVT), the molecular mechanisms by which they do so have been unclear. Our observation that firm adhesion of red cells to platelets and neutrophils is efficient at shear rates below 100 s^{-1} may explain this phenomenon. Reduced shear rates can allow red-cell adhesion to the cells adherent to inflamed venular endothelium causing occlusion and pain. Vascular compressions may lead to a slight increase in blood movement and wall shear rate and shear stress, which, as explained previously, may impede erythrocyte capture and/or reverse adhesion and subsequent vasoocclusion, thus providing DVT prophylaxis to the patient.

CONCLUSION

Adherent neutrophils, being more concentrated at sites of inflammation, have a considerable impact on biomaterial thrombosis. Fibrinogen-adherent neutrophils are sufficient to initiate and amplify coagulation on a biomaterial surface in the presence of activated XIIa. Additionally, the effects of neutrophil-platelet interactions, which primarily occur during adhesion and heterotypic aggregation, on subsequent fibrin formation have been reviewed in a system that decouples flow-regulated adhesion events from flow-regulated coagulation biochemistry. Inflammatory processes,

causing neutrophil or platelet arrest and activation on an activated endothelium or a biomaterial surface, can be followed by receptor-mediated capture of red blood cells under low flow conditions. In addition to biomaterial thrombosis, these events may be clinically relevant in situations devoid of collagen or tissue factor such as those involving endothelial dysfunction and neutrophil activation during deep vein thrombosis.

ARTICLES OF FURTHER INTEREST

Angiogenesis Inhibitors; Blood-Material Interactions; Blood Purification; Distal Protection Devices; Fibrin; Fibrin Sealants

REFERENCES

1. Baglia, F.A.; Walsh, P.N. Prothrombin is a cofactor for the binding of factor XI to the platelet surface and for platelet-mediated factor XI activation by thrombin. *Biochemistry* **1998**, *37* (8), 2271–2281.
2. Monkovic, D.D.; Tracy, P.B. Functional characterization of human platelet-released factor V and its activation by factor Xa and thrombin. *J. Biol. Chem.* **1990**, *265* (28), 17132–17140.
3. Bonnefoy, A.; Hantgan, R.; Legrand, C.; Frojmovic, M.M. A model of platelet aggregation involving multiple interactions of thrombospondin-1, fibrinogen, and GPIIb/IIIa receptor. *J. Biol. Chem.* **2001**, *276* (8), 5605–5612.
4. Finger, E.B.; Puri, K.D.; Alon, R.; Lawrence, M.B.; von Andrian, U.H.; Springer, T.A. Adhesion through L-selectin requires a threshold hydrodynamic shear. *Nature* **1996**, *379* (6562), 266–269.
5. Lawrence, M.B.; Kansas, G.S.; Kunkel, E.J.; Ley, K. Threshold levels of fluid shear promote leukocyte adhesion through selectins (CD62L,P,E). *J. Cell Biol.* **1997**, *136* (3), 717–727.
6. Doggett, T.A.; Girdhar, G.; Lawshe, A.; Schmidtke, D.W.; Laurenzi, I.J.; Diamond, S.L.; Diacovo, T.G. Selectin-like kinetics and biomechanics promote rapid platelet adhesion in flow: The GPIIb(alpha)-vWF tether bond. *Biophys. J.* **2002**, *83* (1), 194–205.
7. Giesen, P.L.; Rauch, U.; Bohrmann, B.; Kling, D.; Roque, M.; Fallon, J.T.; Badimon, J.J.; Himer, J.; Riederer, M.A.; Nemerson, Y. Blood-borne tissue factor: Another view of thrombosis. *Proc. Natl. Acad. Sci., U. S. A.* **1999**, *96* (5), 2311–2315.
8. Rauch, U.; Bonderman, D.; Bohrmann, B.; Badimon, J.J.; Himer, J.; Riederer, M.A.; Nemerson, Y. Transfer of tissue factor from leukocytes to platelets is mediated by CD15 and tissue factor. *Blood* **2000**, *96* (1), 170–175.
9. LaRosa, C.A.; Rohrer, M.J.; Benoit, S.E.; Rodino, L.J.; Barnard, M.R.; Michelson, A.D. Human neutrophil cathepsin G is a potent platelet activator. *J. Vasc. Surg.* **1994**, *19* (2), 306–318.
10. Plescia, J.; Altieri, D.C. Activation of Mac-1 (CD11b/CD18)-bound factor X by released cathepsin G defines

- an alternative pathway of leukocyte initiation of coagulation. *Biochem. J.* **1996**, *319* (3), 873–879.
11. Higuchi, D.A.; Wun, T.C.; Likert, K.M.; Broze, G.J., Jr. The effect of leukocyte elastase on tissue factor pathway inhibitor. *Blood* **1992**, *79* (7), 1712–1719.
 12. Zillmann, A.; Luther, T.; Muller, I.; Kotzsch, M.; Spannagl, M.; Kauke, T.; Oelschlagel, U.; Zahler, S.; Engelmann, B. Platelet-associated tissue factor contributes to the collagen-triggered activation of blood coagulation. *Biochem. Biophys. Res. Commun.* **2001**, *281* (2), 603–609.
 13. Palabrica, T.; Lobb, R.; Furie, B.C.; Aronovitz, M.; Benjamin, C.; Hsu, Y.M.; Sajer, S.A.; Furie, B. Leukocyte accumulation promoting fibrin deposition is mediated in vivo by P-selectin on adherent platelets. *Nature* **1992**, *359* (6398), 848–851.
 14. Taylor, A.D.; Neelamegham, S.; Hellums, J.D.; Smith, C.W.; Simon, S.I. Molecular dynamics of the transition from L-Selectin to β_2 -Integrin dependent neutrophil adhesion under defined hydrodynamic shear. *Biophys. J.* **1996**, *71* (6), 3488–3500.
 15. Schmidtke, D.W.; Diamond, S.L. Direct observation of membrane tethers formed during neutrophil attachment to platelets or P-selectin under physiological flow. *J. Cell Biol.* **2000**, *149* (3), 719–730.
 16. Park, E.Y.; Smith, M.J.; Stropp, E.S.; Snapp, K.R.; DiVietro, J.A.; Walker, W.F.; Schmidtke, D.W.; Diamond, S.L.; Lawrence, M.B. Comparison of PSGL-1 microbead and neutrophil rolling: Microvillus elongation stabilizes P-selectin bond clusters. *Biophys. J.* **2002**, *82* (4), 1835–1847.
 17. Laurenzi, I.J.; Diamond, S.L. Monte Carlo simulation of the heterotypic aggregation kinetics of platelets and neutrophils. *Biophys. J.* **1999**, *77* (3), 1733–1746.
 18. Si-Tahar, M.; Pidard, D.; Balloy, V.; Moniatte, M.; Kieffer, N.; Van Dorsselaer, A.; Chignard, M. Human neutrophil elastase proteolytically activates the platelet integrin α IIb β_3 through cleavage of the carboxyl terminus of the α IIb subunit heavy chain. Involvement in the potentiation of platelet aggregation. *J. Biol. Chem.* **1997**, *272* (17), 11636–11647.
 19. Yamamoto, J.; Ishii, I.; Okada, Y.; Yamashita, T.; Ridler, C.D.; Gorog, P.; Kovacs, I.B. Effect of leukocyte products on platelet thrombus formation, coagulation and spontaneous thrombolysis, as measured from native human blood in vitro. *Thromb. Res.* **1993**, *71* (4), 281–287.
 20. Tijburg, P.N.; Ijsseldijk, M.J.; Sixma, J.J.; de Groot, P.G. Quantification of fibrin deposition in flowing blood with peroxidase-labeled fibrinogen. High shear rates induce decreased fibrin deposition and appearance of fibrin monomers. *Arterioscler. Thromb.* **1991**, *11* (2), 211–220.
 21. Ouriel, K.; Donayre, C.; Shortell, C.K.; Cimino, C.; Donnelly, J.; Oxley, D.; Green, R.M. The hemodynamics of thrombus formation in arteries. *J. Vasc. Surg.* **1991**, *14* (6), 757–762.
 22. Kuijper, P.H.; Gallardo Torres, H.I.; Lammers, J.W.; Sixma, J.J.; Koenderman, L.; Zwaginga, J.J. Platelet and fibrin deposition at the damaged vessel wall: Cooperative substrates for neutrophil adhesion under flow conditions. *Blood* **1997**, *89* (1), 166–175.
 23. Kirchhofer, D.; Sakariassen, K.S.; Clozel, M.; Tschopp, T.B.; Hadvary, P.; Nemerson, Y.; Baumgartner, H.R. Relationship between tissue factor expression and deposition of fibrin, platelets, and leukocytes on cultured endothelial cells under venous blood flow conditions. *Blood* **1993**, *81* (8), 2050–2058.
 24. Manodori, A.B.; Barabino, G.A.; Lubin, B.H.; Kuypers, F.A. Adherence of phosphatidylserine-exposing erythrocytes to endothelial matrix thrombospondin. *Blood* **2000**, *95* (4), 1293–1300.
 25. Test, S.T.; Mitsuyoshi, J. Activation of the alternative pathway of complement by calcium-loaded erythrocytes resulting from loss of membrane phospholipid asymmetry. *J. Lab. Clin. Med.* **1997**, *130* (2), 169–182.
 26. Wun, T.; Paglieroni, T.; Field, C.L.; Welborn, J.; Cheung, A.; Walker, N.J.; Tablin, F. Platelet-erythrocyte adhesion in sickle cell disease. *J. Investig. Med.* **1999**, *47* (3), 121–127.
 27. Hu, W.J.; Eaton, J.W.; Ugarova, T.P.; Tang, L. Molecular basis of biomaterial-mediated foreign body reactions. *Blood* **2001**, *98* (4), 1231–1238.
 28. Goel, M.S.; Diamond, S.L. Neutrophil enhancement of fibrin deposition under flow through platelet-dependent and -independent mechanisms. *Arterioscler. Thromb. Vasc. Biol.* **2001**, *21* (12), 2093–2098.
 29. Allen, D.H.; Tracy, P.B. Human coagulation factor V is activated to the functional cofactor by elastase and cathepsin G expressed at the monocyte surface. *J. Biol. Chem.* **1995**, *270* (3), 1408–1415.
 30. Henderson, L.M.; Figueroa, C.D.; Muller-Esterl, W.; Bhoola, K.D. Assembly of contact-phase factors on the surface of the human neutrophil membrane. *Blood* **1994**, *84* (2), 474–482.
 31. Wachtfogel, Y.T.; Kucich, U.; James, H.L.; Scott, C.F.; Schapira, M.; Zimmerman, M.; Cohen, A.B.; Colman, R.W. Human plasma kallikrein releases neutrophil elastase during blood coagulation. *J. Clin. Invest.* **1983**, *72* (5), 1672–1677.
 32. Goldsmith, H.L.; Turitto, V.T. Rheological aspects of thrombosis and haemostasis: Basic principles and applications. ICTH-report—Subcommittee on rheology of the International Committee on Thrombosis and Haemostasis. *Thromb. Haemost.* **1986**, *55* (3), 415–435.
 33. Selak, M.A. Neutrophil elastase potentiates cathepsin G-induced platelet activation. *Thromb. Haemost.* **1992**, *68* (5), 570–576.
 34. Goel, M.S.; Diamond, S.L. Neutrophil cathepsin G promotes prothrombinase and fibrin formation under flow by activating fibrinogen-adherent platelets. *J. Biol. Chem.* **2003**, *278* (11), 9458–9463.
 35. Sumner, W.T.; Monroe, D.M.; Hoffman, M. Variability in platelet procoagulant activity in healthy volunteers. *Thromb. Res.* **1996**, *81* (5), 533–543.
 36. Monroe, D.M.; Hoffman, M.; Roberts, H.R. Platelets and thrombin generation. *Arterioscler. Thromb. Vasc. Biol.* **2002**, *22* (9), 1381–1389.
 37. Goel, M.S.; Diamond, S.L. Adhesion of normal erythrocytes at depressed venous shear rates to activated neutrophils, activated platelets, and fibrin polymerized from plasma. *Blood* **2002**, *100* (10), 3797–3803.

Tinnitus Devices

Martin L. Lenhardt

Biomedical Engineering, Otolaryngology and Emergency Medicine, Virginia Commonwealth University, Richmond, Virginia, U.S.A.

INTRODUCTION

Tinnitus is the perception of a sound in the absence of external stimulation. The tinnitus is as real to the individual as that induced by external sound. Exposure to noise, ototoxic drugs, or even aging is commonly associated with the development of tinnitus. Clinically, tinnitus is characterized as either subjective or objective. Objective tinnitus, detectable by stethoscopic examination, is generally due to peripheral vascular abnormalities and will not be the topic of this entry. Most tinnitus is subjective, that is, nonpulsatile (varying with the heart beating) and not audible by others. Approximately 75% of those affected by subjective tinnitus adapt, but for the remaining 25%, tinnitus can be severe and debilitating to the point of suicide.

Masking is the auditory phenomenon in which a sound becomes imperceptible in the presence of another sound. The closer the masker is in frequency to the tone to be masked, the less masker energy is required. Low-frequency tones can mask high frequencies if there is sufficient masker energy. This effect is termed upward spread of masking and is important in hearing aid processing algorithms. Conceptually, wearable devices could be developed that would eliminate the perception of tinnitus.

OVERVIEW

Tinnitus is characterized by subjective pitch and loudness; both parameters can be assessed by matching tinnitus to an external tone. Although the process is tedious, reliable data can be obtained. The majority of tinnitus patients pitch-match their tinnitus to frequencies above 6 kHz. Most, but not all, have mild to moderate high-frequency hearing loss.^[1] In the acute stage, tinnitus is usually associated with an insult to the ear in the form of noise trauma or ototoxic drug exposure. There is generally an association of the hearing loss frequencies and the frequencies matched to tinnitus pitch,^[2] which is usually tonal in quality but can also be matched to narrowband noise. Tinnitus is typically matched in loudness to an intensity of an external tone of about 10 dB; however, tinnitus

annoyance may be very high. In designing a device to treat tinnitus sound, its associated annoyance must be addressed. Because there is no tinnitus cure, one standard form of treatment is the use of tinnitus maskers that produce a broad band of continuous external noise that masks or partially covers the tinnitus, but which ear to fit?

Tinnitus can be perceived as being lateralized to one ear or more often perceived binaurally. What is astonishing about the perception of tinnitus in one or both ears is that the ear is rarely the tinnitus source.^[3] Tinnitus arises not in the ear but in the brain. Imaging studies have conclusively demonstrated an auditory cortical activation.^[4] In the case of lateralized tinnitus, the ear opposite the involved cortex appears to be the site of tinnitus, giving the illusion of the ear as the site of tinnitus. Thus the lateralized tinnitus perception suggests peripheral involvement as if it sounds like the tinnitus is in one ear. If the ear is the site of tinnitus, a masker should provide relief by masking according to conventional auditory theory. Because multiple neural sites are involved in generating and processing tinnitus,^[4,5] the masking picture is far more complicated.

The vast majority of potential tinnitus device users have long-term or chronic tinnitus that can also be associated with clinical depression. There may be no direct cause-effect relationship between tinnitus and depression, but rather some common biofactors may predispose an individual to both conditions. Medical treatment strategies have been multifaceted, addressing the treatment of depression and tinnitus with pharmaceuticals, maskers, and behavioral therapy. The question arises as to what treatment outcomes should be selected to measure effectiveness of tinnitus devices. Two obvious measures are device purchases and device returns.

Approximately 1 of 6 patients evaluating maskers actually purchase, and the return rate is about 50%—hardly a ringing endorsement for the technology.^[6] Conventional tinnitus maskers address the initial symptom, but appear not to address the core complaint of tinnitus patients, the annoyance. The irritation, aggravation, and frustration associated with the tinnitus must be addressed in neural tinnitus prosthesis design. More than masking is needed.

HISTORY OF MASKING

The application of acoustic masking to tinnitus is attributed to the pioneering work of Vernon^[7] and his colleagues at Oregon Health Sciences University in the mid 1970s. Vernon readily acknowledges Jean-Marie Gaspar Itard as the first to apply the concept of acoustic masking to tinnitus around 1825. Itard suggested sitting by a green-wood fire; its hissing and popping would mask high-frequency tinnitus, whereas seasoned wood would produce a roar for masking lower-pitched tinnitus. Furthermore, Itard recognized that the sound of falling water is a very effective masker. Thus the principle of frequency-specific masking for tinnitus was known for a century and a half before the availability of wearable technology.

Vernon^[7] was also aware of the use of hearing aids to mask or suppress tinnitus in some patients, and hearing aids could potentially be fitted with masking circuits, thus functioning as both hearing aids and maskers. The first wearable tinnitus devices resembled behind-the-ear hearing aids and were manufactured by Zenith. The hearing aid microphone was removed and replaced with a broadband noise circuit. Early models produced a steady broadband sound, whereas the output of later models was matched to the tinnitus frequency and some even employed phase cancellation techniques. This entry will detail the evolution of such devices from simple maskers to instruments essential in comprehensive tinnitus management.

Conventional masking refers to the phenomenon of the perceptual elimination of one sound in the presence of a second tone. Typically, the masker and the sound to be masked must be similar in frequency content to keep the power requirements low. This was very important during the initial phase of device development in the 1970s because battery life was far more limited than it is today. Frequency matching of masker and maskee is based on efficiency in inner stimulation. The inner ear is effectively divided into critical bands of filters. If the masker and maskee fall within the same band, masking is accomplished with very little energy. The filtering capability of the basilar membrane of the inner ear is maximized when an externally applied tone falls within the same critical band associated with the tinnitus sound. If masking frequencies are systematically varied from low to high, the resulting values of masking are what is termed a psychophysical tuning curve, which is similar to what would be expected from an eighth (auditory) nerve afferent fiber associated with that critical band.^[8-10]

Unfortunately, masking of tinnitus appears to be quite different from conventional psychophysical masking. Investigators^[11,12] using maskers of varying bandwidths have shown that tinnitus is not masked as a tone in many subjects. Instead, a wide range of

frequencies typically provides roughly equivalent masking. These wide tuning curves suggest that tinnitus is not processed as if it were a pure tone. Somehow the tinnitus and the masker interact in the central nervous system. This was the first evidence that suggested tinnitus masking may have a central origin and that tinnitus itself is possibly generated centrally. This conclusion is reinforced by the phenomenon of contralateral masking; in some patients, masking noise presented to the contralateral ear is as effective as noise presented to the ipsilateral ear, so the interaction of tinnitus and masking noise must be at some point in the auditory pathway where there is binaural interaction.^[9] Evidence was mounting that tinnitus masking involves the nervous system and therefore a device that provides central masking was needed.

The observation that the frequency spectra of the masker are not always important in providing masking in some cases is compounded further by the observation that tinnitus can be masked by noise not containing any of the tinnitus frequencies. Kitajima et al.^[13] found that when a narrow band of noise around the tinnitus frequency had been removed (notched noise), masking was as effective as a very narrow band of noise covering the tinnitus frequency. This suggests that tinnitus can be masked by sounds higher or lower than the tinnitus frequency.

Determining which ear to fit with the masker, given the ambiguity of the effects of central masking, i.e., contralateral masking being as effective as ipsilateral masking, can be solved using bone conduction. Bone conduction involves placing the transducer on the skull, thereby efficiently oscillating the inner ear fluids. Bone conduction is a means to deliver broadband sound as part of a masking and habituation tinnitus therapy treatment.^[14] Bone conduction stimulates both ears because there is little attenuation across the head in the audiometric frequencies.^[15] Johnson and Hughes^[16] showed that, because the central neurological processes underlying tinnitus, correlated noise presented simultaneously to the two ears may be more effective at tinnitus suppression than two separate maskers in the ears.

Masker fitting is complex in regard to tinnitus; nonetheless, masking is recognized as an effective therapy.^[17] Masker effectiveness is often based on continuation of masking after the masker is removed, a perception of quiet termed residual inhibition (RI). RI is frequently evoked but is variable in duration ranging from seconds to weeks.^[18,19] A report that only 15 minutes of masking resulted in RI for the entire day encourages masker use. Clearly the RI is a neural phenomenon.^[20] The dependence of RI on masker characteristics such as center frequency, bandwidth, intensity, duration, and stimulating parameters are still not well understood, nor are the mechanisms for the

wide variability in RI for individuals. Thus there is considerable experimentation in both the selection of the type of masking sound and the delivery method employed in an attempt to maximize RI. Four types of maskers that have been shown to have importance in modern tinnitus treatment are electrical stimulation, acoustic masking, vibration masking, and high-frequency/ultrasonic masking.

TRANSDERMAL ELECTRICAL STIMULATION

Tinnitus is present in many severely hearing-impaired and totally deaf patients^[21] with the incidence as high as 87%.^[22] Because many severely hearing-impaired people seek hearing improvement by means of a cochlear implant that stimulates the nerve fibers in the absence of sensory cells, a logical measurable outcome of implantation procedure is the assessment of electrical stimulation on tinnitus. In some cases pre-operative electrical stimulation is applied to the cochlear wall to determine if the nerve fiber population can be activated by the implant. The electrical stimulation of the cochlear wall is in the region of high-frequency coding since it is the most accessible, i.e., nearest to the middle ear. Electrical stimulation of the cochlear wall can suppress tinnitus in 69% of the deaf patients. Cochlear implantation is not as effective, completely suppressing tinnitus in only 35% of the deaf patients and partially suppressing it in 42%. Residual inhibition for several hours after deactivation of the implant is possible.^[22] These findings suggest that activation of the nerve, including the high-frequency fibers, suppresses tinnitus and provides RI.

Years before the availability of cochlear implants, a form of transdermal electrical stimulation was explored to treat tinnitus. Transdermal electrical stimulation is applied by electrodes placed on the preauricular and postauricular regions and on the two mastoids (see Fig. 1). Electrical stimulation consisted of slowly varying tones between 0.2 and 20 kHz multiplied by a 60-kHz carrier (full amplitude modulation). This approach literally passed the audio frequencies under the skin via external electrodes placed on both mastoids. Electrical stimulation in this mode provided 60% positive success in suppressing tinnitus.^[23] An additional study^[24] revealed that in 50 patients tested, 14 (28%) obtained relief that met the criterion of a reduction in the tinnitus of 40% or more, with RI usually extended for several hours. Further, there was only one positive response in the placebo trial; nonetheless, the authors concluded electrical transcutaneous stimulation was not practical. A more powerful randomized, double-blind crossover study resulted in a reduction in tinnitus severity in only two of 20 patients with the active device and four of



Fig. 1 Headset with electrodes for transdermal electrical stimulation for tinnitus treatment. High frequencies were modulated on an ultrasonic carrier with the device (Theraband) capacitive coupled to the head during stimulation.

20 patients with the placebo device.^[25] In a large series of 500 patients treated with electrical stimulation twice weekly for a total of six to ten visits,^[26] 53% reported a significant benefit, defined as an improvement of at least two points on a 10-point scale of tinnitus intensity. Despite some positive and some ambiguous results, the transdermal electrical stimulation device was withdrawn from the market. This was unfortunately premature because electrical stimulation in the form of FDA-approved cochlear implants would provide the physiological rationale for transdermal tinnitus treatment. The transdermal stimulation provided additional design value in regard to demonstrating that high-frequency stimulation is possible without distortion associated with air conduction hearing. Although not specifically recognized at the time, high-frequency stimulation is an essential element in modern masking and treatment of high-frequency tinnitus.

ACOUSTIC MASKER DESIGNS

Masking a low-level high-pitch tinnitus percept at first does not seem to be a formidable task. If, for example, tinnitus is matched to an 8-kHz tone, delivering a tone or narrow band of noise to the cochlea at 10 dB above threshold should accomplish the predictable masking. Assuming that the tinnitus patient prefers the masking to the presence of tinnitus, providing a tinnitus masker should be the obvious simple solution. Surprisingly, about one-third of the tinnitus patients report no masking even if the masker is at very high intensities. One-third report masking with frequencies much different in frequency from their pitch match and only one-third respond as predicted from classical auditory masking theory. It appears that tinnitus masking

involves other processes different from classical masking that likely involve different pathways in the brain.

The first approach in developing an acoustic masker was to select a form of stimulation that would be effective in the majority of cases, and the logical candidate was broadband noise. The concept of masking an internally generated sound with an external sound didn't occur to researchers in the field until the 1970s. Dr. Jack Vernon, the pioneer in tinnitus masking,^[17] relates his experience with a physician suffering tinnitus who, when standing in one location near a water fountain, declared his tinnitus had disappeared. The sound of falling water was replicated in the laboratory and the synthetic waterfall was found to be effective as a masker of tinnitus. The challenge was to incorporate that masking sound into a wearable device.

The faucet test, as it were (I can't hear you while the water is running), is an excellent means of getting a preliminary assessment on the maskability of tinnitus. Water-generated noise is an outstanding source of almost flat and full-spectrum noise (Fig. 2). Synthetic water noise, essentially white, was the first masking stimulus incorporated into wearable and desk-version maskers and it is still used today.

The logical candidate for a personal wearable masker was a modified hearing aid. A prototype wearable noise generator was evaluated by Vernon's group in 1974. Although the circuit was capable of wide-frequency response, the high-frequency output was limited by the driver characteristics (<3 kHz) of the hearing aid. In the digital realm, a white noise chip will provide full high-frequency content, but the miniature speaker limits high-frequency fidelity. Most maskers are limited to about 6 kHz of effective output,

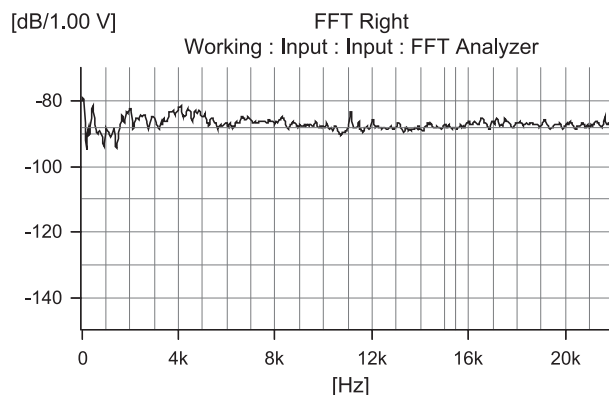


Fig. 2 Spectrum of running water from a faucet. The value of -70 dBV corresponds to 80 dB SPL (C) at one meter. The flat spectrum extending into the ultrasonic frequencies is the reason this sound is effective in initial tinnitus masking.

which is not surprising in that the middle ear resonance is about 3 kHz and its high frequency rolls off at about 26 dB/octave. Additionally, most tinnitus sufferers have high-frequency hearing loss placing increased intensity demands on high-frequency maskers, which can result in distortion and possible device rejection. The goal of wearable synthetic waterfalls proved difficult to obtain with conventional hearing aid technology. (See *Hearing Mechanisms*, p. 702)

FM MASKING

There are alternative choices to wearable maskers; desktop maskers and pillow maskers are available that generally provide a broad range of acoustic masking but are also limited in the high frequencies by their transducer response. A mistuned FM radio can provide white noise amplified through a high-fidelity speaker system. Although the mistuned radio advice has often been offered to tinnitus sufferers, its general utility and long-term acceptance is low. There persists the occasional report of a tinnitus sufferer able to fall asleep only while sitting in front of a full-volume television producing white noise after station sign-off.

PERSONAL STEREOS

Vernon readily appreciated that wearable maskers of mid-1970 vintage had intrinsically limited high-frequency response, acknowledging that patients report their tinnitus matches frequencies above 6 kHz. Vernon recorded higher frequencies on cassette tape, creating a type of desk-based masker for home use. The high-frequency response of most cassette players was usually about 16 kHz, but again the limiting element was the response of personal earphones, even some rated to 20 kHz. The high-frequency response of the stereo earphones varies by construction. An example of a relatively inexpensive set versus high-fidelity earphones is depicted in Fig. 3. The high-fidelity earphones (right panel) can extend the frequency response to a degree. Even with exact reproduction, wavelengths above 14 kHz and canal dimensions interact, depending on individual geometry, resulting in cancellation of some high frequencies.

With the advent of relatively inexpensive compact disks (CD) and players, high-frequency masking could be achieved by a "maskman." Digital sampling rates of 44.2 kHz are commonly used in CD recording that extend the frequency range to about 20 kHz (anti-aliasing filters limit the upper frequencies). Nonetheless, the concept of a wearable maskman never achieved popularity in spite of being a much cheaper alternative to tinnitus maskers.^[27] Perhaps wearable

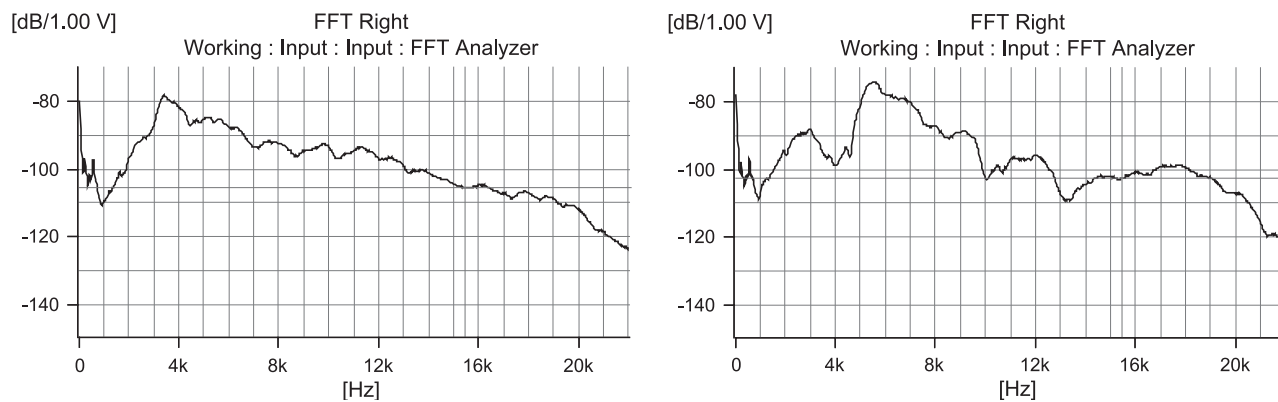


Fig. 3 Frequency response of an inexpensive headset (left panel) and high-fidelity earphones (right panel). Overall response is similar in the low and mid frequencies and not that different at 16 kHz. Ear canal acoustics, middle ear mechanics, and sensory hearing loss all contribute to inefficient high-frequency audibility by air conduction.

maskers and stereos were providing more high-frequency output, but full high-frequency fidelity is not obtainable. The need for full treble in tinnitus treatment has been demonstrated only in the past few years. It is the lack of high frequencies delivered to the ear that may determine the success of tinnitus masking; however, some other central factor must be involved in tinnitus persistence.^[28] For these reasons, alternative vibration-based maskers were developed.

VIBRATION MASKERS

There is one report^[29] in the literature of 60% success in tinnitus relief using low-frequency vibration as a form of acoustic masking in a small sample of patients. The approach was exploited by De Mino,^[56] who developed a vibrating rod that is applied to the skin behind the ear (mastoid area). The stimuli consisted of a 60-Hz fundamental with adjustable harmonics up to 1000 Hz. The device (Aurex3) provides air-conducted, bone-conducted, and tactile stimulation and received FDA approval [510(K)] for tinnitus masking. The device was not well accepted, possibly due to cosmetic reasons. It was assumed that the acoustic energy delivered by bone conduction was the principle feature related to successful masking; however, the key effect may not have been auditory at all.

The somatosensory system has been shown to have inputs into the auditory system as low as the first synapse in the dorsal cochlear nucleus. Stimulation of the muscle body behind the ear (postauricular muscle) results in inhibition in the dorsal cochlear nucleus.^[30] The tinnitus inhibition effect is not due to air-conducted hearing, as removing the transducer from the muscle eliminated the effect. Light touch is also ineffective, suggesting that touch or refocusing

attention is not the responsible mechanism. Thus it appears that the Aurex3 may be effective when the vibrating rod stimulates the postauricle muscle and not the ear directly. The salient point is that if tinnitus can be modified by a sensory system different from audition, it is clear that complete masking may require multisensory stimulation such as by sound and muscle vibration.^[31] Tinnitus inhibition via the somatosensory system provides fast immediate relief, but there is no RI. In contrast, very-high-frequency vibration has no tactile component and can provide long-term tinnitus relief. High success in masking tinnitus and producing RI was reported with low-frequency ultrasound (20–40 kHz) delivered through bone conduction.^[32] Lenhardt et al.^[33] demonstrated that hearing through bone conduction can extend up to about 100 kHz.

HIGH AUDIO AND ULTRASONIC BONE VIBRATION MASKERS

A new trend in tinnitus treatment is the use of very-high-frequency maskers, including ultrasound, delivered by bone conduction.^[18,32] Masking and long-term inhibition may involve inducing neuroplastic changes in the brain. The bone conduction application of either high audio frequencies (10–20 kHz) or ultrasound (20–100 kHz) provides binaural high-frequency stimulation for high-pitched tinnitus masking, a requirement recognized by Vernon decades ago.

The principle obstacle in delivering high frequencies to the head was the development of a suitable transducer. An aluminum ceramic (PZT) bimorph with a resonance frequency of about 40 kHz in air was designed to stimulate the head from 5–45 kHz. There are multiple transducer resonances, which are compounded with head resonances,^[22] making calibration

a challenge. Nonetheless, the transducer is light and comfortable and has a relatively flat frequency response when mass-loaded to the head.^[18]

Bone conduction calibration standards are not available for high audio and ultrasonic frequencies. Reference can be made to calibration procedures based in part on bone-anchored measurements for some frequencies,^[34] but ultrasonic bone conduction standards do not exist.^[35] The transducer can be placed on a water surface in a small tank, assuming brain and water impedances are similar.^[36] This approach provides reliable results using a high-frequency hydrophone, but provides no exact measurement of the energy acting in the ear in that the boundary condition of skin and bone are ignored in this approach. Alternatively, acceleration can be measured and this proves to be the best method of calibration.

High-frequency accelerometers have impedances closer to bone than to brain. Using acceleration (m/sec^2), the hearing threshold for bone conduction can be referenced as -30 dB relative to $1 \text{ m}/\text{sec}^2$ from 0.25 to 6 kHz . This reference is applicable to higher frequencies, including ultrasound. A standard point of measurement is 1 gravity unit (g) rms, or $9.81 \text{ m}/\text{sec}^2$.^[37] One g is also the U.S. Occupational Safety and Health Administration (OSHA) hearing protection reference for body-coupled ultrasound.^[38]

Tones or narrow bands of noise can be applied to the head as ultrasonic bone conduction tinnitus maskers. This would simply be a variation on acoustic maskers, but the delivery method would be bone conduction. However, if the masker is similar in pitch and intensity to the patient's tinnitus, most patients would reject the substitution of just one annoying sound for a masker. If, however, the sound stimulus has an attention-focusing temporal quality, e.g., music, rejection is rare.

Thus if music or similar sound is multiplied by a high-frequency or ultrasonic carrier and the carrier and

lower sideband are canceled, a high-frequency patterned stimulation will result. Patterned auditory stimulation in the high audio range (10 – 20 kHz), using bone conduction, has been shown to be an efficient means of high-frequency tinnitus masking.^[18] It differs from the conventional masker in that little of the sound overlaps the range of reported tinnitus (Fig. 4, right panel). In this technique, amplitude-modulated and processed music is presented through a bone conduction transducer at a low-level 12 -dB sensation level that is 12 dB above threshold, for periods of 30 – 60 minutes (Fig. 4, right panel). The goal is to induce changes in the central nervous system mechanisms of tinnitus, resulting in long-term inhibition.

In the case of high-frequency bone-conducted stimulation, not all patients with a similar profile experienced immediate masking; in fact, immediate masking is generally a positive predictor of long-term relief. The overall percentage of tinnitus relief was 60 – 70% .^[18] If the masker is shifted into the ultrasonic range, tinnitus relief is also reported.^[32] Ultrasound, as will be discussed later, maps onto the high audio area of the inner ear, so both high audio and ultrasonic stimulation have similar sites of peripheral and central action.

HOW IS ULTRASONIC HEARING POSSIBLE?

Reports that humans can hear ultrasound have appeared more than a dozen times over the last half-century.^[33] Dr. Roger Maass detailed the phenomenology of ultrasonic hearing, perceptible only as bodily (bone conduction) vibration, with pitch characterized as similar to audiofrequencies in the 10 – 15 kHz range, and with crude frequency discrimination. Maass further reported some deaf subjects could also hear ultrasound at high levels. These observations were

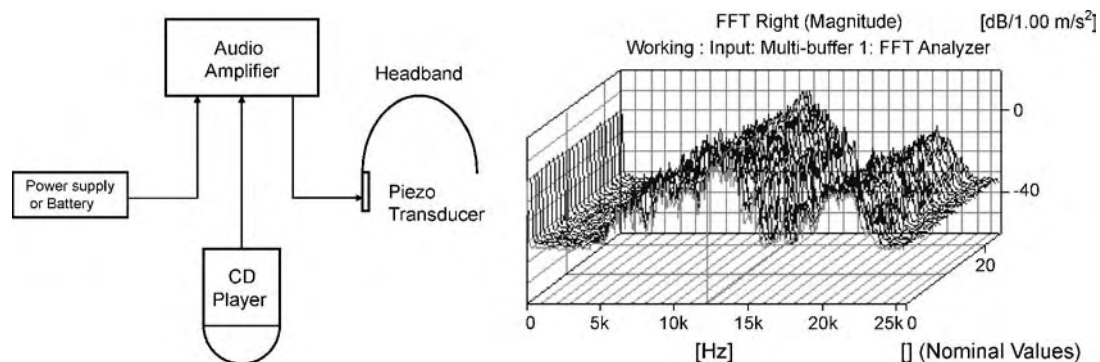


Fig. 4 The high-audiofrequency tinnitus device (UltraQuiet™) consists of a piezoceramic aluminum transducer, a compact high-frequency amplifier, and a playback system (left panel). The vibratory output of the device is depicted in the right panel. The suppressed carrier is at 16 kHz . Filtering limits the upper sideband energy to 20 kHz . Peak energy is in the region of brain resonance to utilize natural amplification.

T

confirmed, and speech discrimination in deaf subjects was verified using modulated ultrasound.^[33] By means of magnetoencephalography, it has been determined that the primary auditory cortex is active in ultrasonic hearing in normal hearing subjects, activating sites associated with very-high-frequency hearing.^[39] A lower-frequency cortical site for ultrasound for the deaf was observed, the significance of which will be discussed later in regard to neural plasticity.^[40]

When ultrasonic energy is applied to the head, some form of demodulation occurs. The site of the demodulator could be any place, in that only some form of half-wave rectification induced by a non-linearity is required. However, the pitch of ultrasound suggests the site must have a resonant frequency above 10 kHz. A potential candidate for the ultrasound demodulator is the brain itself. The calculated brain resonant frequency, assuming a homogenous fluid-filled spherical model,^[41,42] using the formula:

$$F = c/2\pi r$$

where F is the fundamental frequency of the sound generated by the sphere, c is the velocity of sound in brain tissue (1460 m/sec),^[41,42] and r is the sphere radius, and assuming a 7-cm head radius, is 13.4 kHz. Neglected in the model are the skull and skin, together acting as a boundary condition. A higher resonant frequency^[41] of $F=15.6$ kHz would be predicted with the boundary condition. Because the brain is not truly spherical, the exact brain resonant frequency is probably a value between the free and boundary states. Individual brain resonant frequencies are likely to fall between 11 and 16 kHz, with the exact value determined by skull geometry.^[15]

If the brain is set into vibration by ultrasound, the sound produced by brain resonance will propagate to the ear via fluid channels. The most likely fluid channel is the cochlea aqueduct that allows direct fluid communication with the brain via the cerebral spinal fluid.^[43-45] If the brain—and not the skull—is vibrated, auditory nerve discharges verify the brain-ear mechanism. Transcranial ultrasonic Doppler imaging has given additional support to brain demodulation. Patients report hearing a high audio sound, much like tinnitus, when the beam is focused in the center of the brain, but the sound disappears when the beam is focused on the ear.^[46]

The position (P) of maximal stimulation of each frequency on the basilar membrane can be calculated using a formula devised by Fay:^[47]

$$P = \frac{\log_{10}[\text{fHz}/.008 \text{ fHz max}] + 1^*}{2.1}$$

where P is the proportion of base to apex on the basilar membrane, fHz is the frequency of interest, and fHz max is the maximal audible frequency by air conduction in young adults.

Further, the position (P) on the basilar membrane is determined by:

$$P = 1 - P \times \text{cochlear length (31.5 mm)}$$

The positions on the basilar membrane for brain resonance would fall between 1.5 mm from the base of the cochlea for 16 kHz to 3.9 mm for 11 kHz. These data clearly support the hypothesis of physical demodulation of ultrasound by brain resonance and the detection of this resonance in the base of the cochlea. The spherical model data also support the theory that ultrasound, regardless of its frequency, stimulates an area on the basilar membrane that codes the fundamental resonant frequency on the brain. Further, ultrasound is a method to deliver high-frequency tinnitus masking.

APPLICATIONS OF ULTRASONIC TINNITUS MASKERS TO THE DEAF

With progressive high-frequency hearing loss, pitch perceptual range collapses. Ultrasonic thresholds also increase with the degree of high-frequency hearing loss.^[33] Presumably, the increased ultrasonic energy increases the displacement spread on the basilar membrane toward the apex, thus accounting for the observation that the ultrasonic pitch is related to the highest frequency detectable by air conduction. In the case of severe deafness, with ultrasonic thresholds approximately 30 dB above normal, insufficient surviving hair cells exist in the apex to detect the basilar membrane motion. Lenhardt et al.^[33] argued that the saccule may be stimulated in the case of severe deafness. Assuming that maximal displacement of the basilar membrane is at the place corresponding to the brain's fundamental frequency, bulk inner-ear fluid displacement, permitted by compliant oval and round windows at very high intensities (100 SL+), could conceivably create fluid flow in the saccule, not unlike the Tullio effect.^[48] The very short cilia of saccular hair cells, not mass-loaded by gel or otoconia^[49] in the striola region, are likely fluid velocity-sensitive. Stimulating the saccule—an organ having input into the auditory pathways in mammals^[50]—may explain ultrasonic detection in the deaf. Treating tinnitus in severely deafened individuals with ultrasound has not been systematically studied, which is surprising due to the high incidence of tinnitus in this population^[21] and the lack of alternatives to implant electrical stimulation. An ultrasonic tinnitus masker is commercially available.

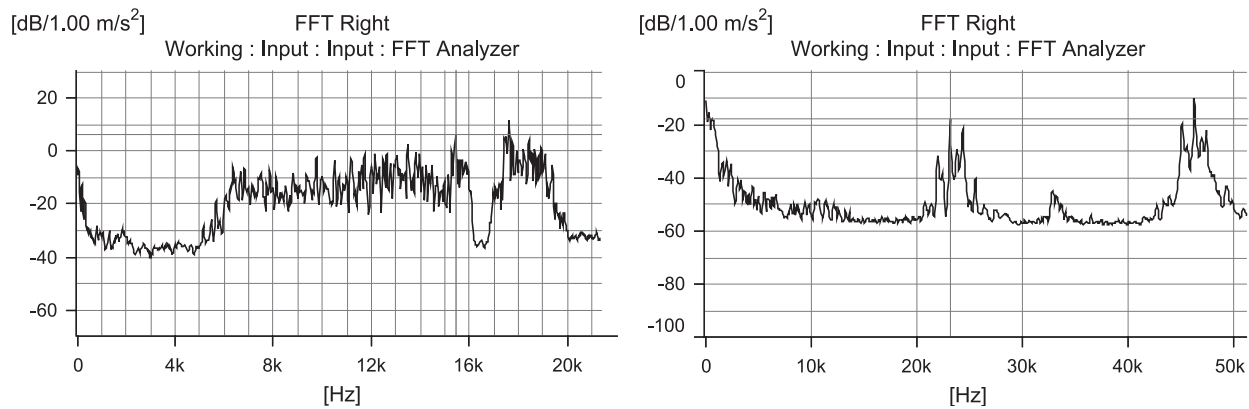


Fig. 5 Pulsed pattern bone-conduction stimulation with a high-audiofrequency device (left panel) and an ultrasonic device (right panel). The high-audio stimulation is double sideband, whereas the ultrasound is full amplitude modulation. Calibration reference is vibration in acceleration re: 1 m/s².

NEURAL PLASTICITY AND CORTICAL REMAPPING; THE MISSING ELEMENT IN TINNITUS DEVICE DESIGN

Traditional acoustic maskers have been available for more than a quarter century, yet masking only, when it occurs, seems insufficient for long-term tinnitus treatment. The compelling evidence for central processing in tinnitus suggests maskers must be designed to also address how the brain responds to tinnitus.

Conventional wisdom until very recently was that the adult human brain is hard-wired. Imaging studies have confirmed considerable cortical plasticity after sensory loss. That is to say, the brain is dynamic and it adjusts for loss in the peripheral sense organs. In monkeys,^[51] the tonotopic map (frequencies laid out spatially in the neuroaxis preserving the spatial coding in the ear) reorganized months after cochlear deafening. The deprived area of the cortex reorganized and became responsive only to intact areas of the cochlea. What this means is that the brain hears only those frequencies transduced by the ear. If stimulation ceases in a particular frequency region, the neurons in the brain sensitive to the damaged peripheral region change their frequency response.

Tinnitus is usually associated with hearing loss and hence cortical reorganization. Muhnickel et al.^[52] explored the reorganization of the auditory cortex in tinnitus using magnetoencephalography, and found a marked shift of the cortical representation of the tinnitus frequency into an area adjacent to the expected tonotopic location. There was also a strong positive correlation between the strength of the tinnitus and the amount of cortical reorganization. Plastic changes in the auditory neural axis in severe tinnitus, particularly in the auditory cortex, may play a role in the continued perception of tinnitus by adding salience to the experience.^[53,54] In tinnitus patients the tinnitus

frequency area expands to more than twice the size. Aberrant tinnitus neural reprogramming can possibly be reversed by increasing high-frequency stimulation (i.e., with frequencies above the tinnitus frequency), for auditory learning in primates has indeed been shown to expand the frequency map,^[52,55] suggesting external stimulation can change the brain. High-frequency stimulation (high audio and ultrasound) has been shown to not only mask tinnitus but also produce varying degrees of residual inhibition.^[18,32] The spectra of these masking stimuli are depicted in Fig. 5.

Tinnitus maskers have evolved from just broadband noise generators to brain stimulators essential in comprehensive tinnitus management. The next generation of devices will be smart systems with capacities for multisensory and multimodal stimulation, likely providing greater relief (RI) for a condition that will only increase in prevalence with an aging population.

ARTICLES OF FURTHER INTEREST

Hearing Mechanisms; Inner Ear Implants

REFERENCES

1. Vernon, J.A. Common Errors in the Use of Masking for Relief of Tinnitus. In *Tinnitus, Diagnosis/Treatment*; Shulman, A., Aran, J.M., Tonndorf, J., et al. Eds.; Lea & Febiger Inc.: Philadelphia, 1991; 50–66.
2. Lockwood, A.H.; Salvi, R.J.; Burkard, R.F. Tinnitus. *N. Engl. J. Med.* **2002**, *347* (12), 904–910.
3. Shulman, A.; Strashun, A.M.; Afriyie, M. SPECT imaging of brain and tinnitus—Neurotoxic/neurologic implications. *Int. Tinnitus J.* **1995**, *1* (1), 13–29.
4. Lockwood, A.H.; Salvi, R.J.; Coad, M.L. The functional anatomy of tinnitus: Evidence for limbic system

- links and neural plasticity. *Neurology* **1998**, *1* (50), 114–120.
5. Shulman, A. A final common pathway for tinnitus—The medial temporal lobe system. *Int. Tinnitus J.* **1995**, *2* (1), 115–126.
 6. Vernon, J.A.; Meikle, M.B. Tinnitus Masking. In *Tinnitus Handbook*; Tyler, R.S., Ed.; Singular: San Diego, 2000; 313–355.
 7. Vernon, J. The history of masking as applied to tinnitus. *J. Laryngol. Otol. Suppl.* **1981**, *1* (4), 76–79.
 8. Warren, R.M. *Auditory Perception: A New Analysis and Synthesis*; Cambridge University Press: New York, 1999.
 9. Feldmann, H. Homolateral and contralateral masking of tinnitus by noise bands and pure tones. *Audiology* **1971**, *4* (10), 138–144.
 10. Tyler, R.S.; Conrad-Arnes, D.; Smith, P. Postmasking effects of sensorineural tinnitus: A preliminary investigation. *J. Speech Hear. Res.* **1984**, *1* (27), 466–474.
 11. Penner, M.J. Masking of tinnitus and central masking. *J. Speech Hear. Res.* **1987**, *1* (30), 147–152.
 12. Smith, P.A.; Parr, V.M.; Lutman, M.E.; Coles, R.R.A. Comparative study of four noise spectra as potential tinnitus maskers. *Br. J. Audiol.* **1991**, *1* (25), 25–34.
 13. Kitajima, K.; Kitahara, M.; Kodama, A. Can tinnitus be masked by band erased filtered masker? Masking tinnitus with sounds not covering the tinnitus frequency. *Am. J. Otol.* **1997**, *1*, 203–206.
 14. Holgers, K.M.; Hakansson, B.E. Sound stimulation via bone conduction for tinnitus relief: A pilot study. *Int. J. Audiol.* **2002**, *41* (5), 293–300.
 15. Cai, Z.; Richards, D.G.; Lenhardt, M.L.; Madsen, A.G. Response of human skull to bone conducted sound in the audiometric to ultrasonic range. *Int. Tinnitus J.* **2002**, *8* (1), 1–8.
 16. Johnson, R.M.; Hughes, F.M. Diotic Versus Dichotic Masking of Tinnitus: Tinnitus 91, Proceedings of the Fourth International Tinnitus Seminar; Aran, J.-M., Dauman, R., Eds.; Kugler Publications: New York, 1991; 387–390.
 17. Vernon, J.A. *Tinnitus Treatment and Relief*; Vernon, J.J., Ed.; Allyn and Bacon: Boston, 1998.
 18. Goldstein, B.A.; Shulman, A.; Lenhardt, M.L.; Richards, D.G.; Madsen, A.G.; Guinta, R. Long-term inhibition of tinnitus by UltraQuiet therapy: Preliminary report. *Int. Tinnitus J.* **2001**, *2* (7), 122–127.
 19. Hazell, J.W.P.; Wood, S. Tinnitus masking—A significant contribution to tinnitus management. *Br. J. Audiol.* **1981**, *4* (15), 223–230.
 20. Terry, A.M.P.; Jones, D.M.; David, B.R.; Slater, R. Parametric studies of tinnitus masking and residual inhibition. *Br. J. Audiol.* **1983**, *1* (17), 245–256.
 21. Tyler, R.S. Tinnitus in the profoundly hearing impaired and the effects of cochlear implants. *Ann. Otol. Rhinol. Laryngol.* **1995**, *165* (104), 25–30.
 22. Ito, J.I.; Sakakihara, L. Suppression of tinnitus by cochlear implantation. *Am. J. Otolaryngol.* **1994**, *1* (15), 145–148.
 23. Shulman, A.; Tonndorf, J.; Goldstein, B. Electric tinnitus control. *Acta Oto-Laryngol.* **1985**, *3–4* (99), 3–4, 318–325.
 24. Vernon, J.A.; Fenwick, J.A. Attempts to suppress tinnitus with transcutaneous electrical stimulation. *Otolaryngol. Head Neck Surg.* **1985**, *3* (93), 9, 385–389.
 25. Dobie, R.A.; Hoberg, K.E.; Rees, T.S. Electrical tinnitus suppression: A double blind crossover study. *Otolaryngol. Head Neck Surg.* **1986**, *3* (95), 319–323.
 26. Steenerson, R.L.; Cronin, G.W. Treatment of tinnitus with electrical stimulation. *Otolaryngol. Head Neck Surg.* **1999**, *5* (121), 511–513.
 27. Al-Jassim, A.H. The use of Walkman mini-stereo system as a tinnitus masker. *J. Laryngol. Otol.* **1988**, *102* (1), 27–28.
 28. Shulman, A.; Strashun, A. Descending auditory system/cerebellum/tinnitus. *Int. Tinnitus J.* **1999**, *5* (1), 92–106.
 29. Thomas, M.; Laurell, G.; Lundeberg, T. Vibratory stimulation as a treatment alternative in patients with tinnitus. *Ear Nose Throat J.* **1989**, *68* (11), 810–814, 817–818.
 30. Kanold, P.; Young, E.D. Proprioceptive information from the pinna provides somatosensory input to cat dorsal cochlear nucleus. *J. Neurosci.* **2001**, *19* (21), 7848–7858.
 31. Lenhardt, M.L.; Goldstein, B.A.; Shulman, A.; Guinta, R. Use of high-frequency and muscle vibration in the treatment of tinnitus. *Int. Tinnitus J.* **2003**, *1* (9), 1, 32–36.
 32. Meikle, M.B.; Edlefsen, L.L.; Lay, J.W. Suppression of Tinnitus by Bone Conduction of Ultrasound. In *Abstracts of the 21st Annual Meeting of the Association for Research in Otolaryngology*; , 1999; 223.
 33. Lenhardt, M.L.; Skellett, R.; Wang, P.; Clarke, A.M. Human ultrasonic speech perception. *Science* **1991**, *5015* (253), 82–85.
 34. Hakansson, B.; Tjellstrom, A.; Rosenhall, U. Acceleration levels at hearing threshold with direct bone conduction vs. conventional bone conduction. *Acta Oto-Laryngol.* **1985**, *3–4* (100), 240–252.
 35. *Reference Equivalent Threshold Force Levels for Audiometric Bone Vibrators (ANSI S3.26)*; American National Standards Institute: New York, 1981.
 36. Corso, J. Bone-conduction thresholds for sonic and ultrasonic frequencies. *J. Acoust. Soc. Am.* **1963**, *4* (35), 1738–1743.
 37. Lenhardt, M.L.; Richards, D.G.; Madsen, A.G.; Goldstein, B.A.; Shulman, A.; Guinta, R. Measurement of bone conduction levels for high frequency. *Int. Tinnitus J.* **2002**, *1* (8), 1, 9–12.
 38. *Occupational Safety and Health Administration Technical Manual*; Ultrasonics, U.S. Department of Labor: Washington, DC, 2002. Section III, Chapter 5, Subchapter V.
 39. Hosoi, H.; Imaizumi, S.; Sakaguchi, T. Activation of the auditory cortex by ultrasound. *Lancet* **1998**, *9101* (351), 496–497.
 40. Imaizumi, S.; Hosoi, H.; Sakaguchi, T. Ultrasound activated the auditory cortex of profoundly deaf subjects. *NeuroReport* **2001**, *12* (3), 583–586.
 41. Borth, D.E.; Cain, C.A. Theoretical analysis of acoustic signal generation in materials irradiated with microwave energy. *IEEE Trans. Microwave Theor. Tech.* **1977**, *525* (11), 944–954.

42. Olsen, R.G.; Lin, J.C. Microwave pulse induced acoustic resonances in spherical models. *IEEE Trans. Microwave Theor. Tech.* **1981**, *29* (10), 114–117.
43. Ranke, O.F. Physiologie des Gehors. In *Gehor Stimme-Sprache*; Ranke, O.F., Lullies, H., Eds.; Springer: Berlin, 1953; 3–110.
44. Tonndorf, J. A new concept of bone conduction. *Arch. Otol.* **1968**, *2* (87), 49–54.
45. Tonndorf, J. Bone conduction. Studies in experimental animals. *Acta Oto-Laryngol., Suppl.* **1966**, *213*.
46. Magee, T.R.; Davies, A.H. Auditory phenomena during transcranial Doppler insonation of the basilar artery. *J. Ultrasound Med.* **1993**, *12* (12), 747–750.
47. Fay, R.R. Structure and Function in Sound Discrimination Among Vertebrates. In *The Evolutionary Biology of Hearing*; Webster, D.E., Fay, R.R., Popper, A.N., Eds.; Springer-Verlag: New York, 1992; 245–250.
48. Watson, S.R.; Halmagyi, G.M.; Colebatch, J.G. Vestibular hypersensitivity to sound (Tullio phenomenon): Structural and functional assessment. *Neurology* **2000**, *3* (54), 722–728.
49. Lim, D. The Development and Structure of the Otoconia. In *Ultrastructural Atlas of the Inner Ear*; Friedmann, I., Ballantyne, J., Eds.; Butterworth: London, 1984; 245–269.
50. Sheykhleslami, K.; Kaga, K. The otolithic organ as a receptor of vestibular hearing revealed by vestibular-evoked myogenic potentials in patients with inner ear anomalies. *Hear. Res.* **2002**, *165* (1–2), 62–67.
51. Schwaber, M.K.; Garraghty, P.E.; Kaas, J.H. Neuroplasticity of the adult primary auditory cortex following cochlear hearing loss. *Am. J. Otol.* **1993**, *3* (14), 252–258.
52. Muhlneckel, W.; Elbert, T.; Taub, E.; Flor, H. Reorganization of auditory cortex in tinnitus. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *17* (95), 10340–10343.
53. Mirz, F.G.; Jedde, A.; Ishizu, K.; Pedersen, C.B. Cortical networks subserving the perception of tinnitus. *Acta Oto-Laryngol., Suppl.* **2000**, (543), 241–243.
54. Stegiade, M. Coherent continuations and short term plasticity in corticothalamic networks. *Tr. Neurosci.* **1999**, *1* (22), 337–345.
55. Menning, H.; Roberts, L.E.; Pantev, C. Plastic changes in the auditory cortex induced by intensive frequency discrimination training. *NeuroReport* **2002**, *4* (11), 817–822.
56. Di Mino, A. Electronic Stimulation System for Treating Tinnitus Disorders. United States Patent 6,210,321, April 3, 2001.

Tissue Engineering

François A. Auger

Lucie Germain

LOEX, Hôpital du Saint-Sacrement du CHA, and Department of Surgery, Laval University, Quebec City, Quebec, Canada

INTRODUCTION

There is little doubt that tissue engineering is a revolutionary addition to the therapeutic armamentarium of medicine. The dilemma of adequately repairing either failing or traumatized organs has been looming larger as patients either become older or are in dire need of grafts. Compounding some of the intrinsic problems of transplantation is the chronic shortage of tissues and organs. Tissue engineering allows the hope of a regular creation of spare parts for the human body. This is a most significant approach to reconstruct, replace, or repair organs in a way that could not be foreseen 25 years ago.

Reconstructive medicine is, in a way, not a very recent concept. If one stays away from punctilious definitions, one of its forms, reconstructive surgery, has been practiced for quite some time, with a surge of development after the Second World War. In 1970s, the development of microsurgery allowed distant tissue transfer and reimplantation.^[1-5] Since then, the introduction of various biomaterials has allowed vast and diversified types of reconstruction of the human body. Vascular grafts and prosthetic articulation are two prominent examples.^[6]

However, tissue engineering does open a radically new chapter in reconstructive medicine, for it is now deemed possible to reconstruct in the laboratory human living tissues and organs for either *in vivo*, *ex vivo*, and even *in vitro* applications.^[7-13] This new domain of biotechnology is remarkably multidisciplinary, bringing together cell and molecular biologists, biochemists, engineers, pharmacologists, physicians, among others.

When the aim of tissue engineers is to obtain grafts for *in vivo* applications, then the biological and mechanical functions are of utmost importance. In some subdivisions of the field, one can essentially choose between the biological function, as in cell therapy, or mainly the mechanical function, such as happens with the use of tissue templates^[14] (Fig. 1).

Tissue-engineered substitutes are three-dimensional reconstructions that can be implanted into the human body, leading to rapid host integration and acceptance. These substitutes must have at least minimal biological and mechanical functions for such a reparative role.

HISTORICAL PERSPECTIVE

Even though the field of tissue engineering is relatively young, it has already enjoyed a fascinating evolution over the last quarter of a century. Tissue engineering has also been considered one of the most influential new technologies for the future of biomedicine.^[15-17]

The development of tissue engineering can be seen as having two phases. The phase of exponential development and potential application is still continuing to evolve. But it seems reasonable to also identify a second phase, brought about by a flurry of discoveries about stem cells. Even though these cells had been known of for many years and had certainly been involved in many tissue-engineered efforts, some very important aspects of embryonic stem cells were revealed. Indeed, the isolation and stable culturing of either the totipotent or germinal stem cell was a pivotal breakthrough.^[18-20] In parallel, adult stem cells were found to be much more ubiquitous and to have more lineage plasticity than previously thought.^[21]

Since 1998, these discoveries have brought a tremendous amount of energy and expectation to the field of tissue engineering and reconstruction. Indeed, their therapeutic potential since then has been estimated to be greatly enhanced.

APPROACHES TO TISSUE-ENGINEERED SUBSTITUTES

The different approaches to tissue-engineered substitutes involve deceptively few elements: cells and various natural or artificial matrices combined in such manner to obtain a tissue substitute. These tissue substitutes can then in turn be sequentially integrated to create more complex organs (Fig. 2).

Three main approaches are being utilized for tissue engineering. The first is the seeding of cells into various gels. The seminal work by Dr. Eugene Bell and collaborators has set the stage for various applications of such an approach with collagen gels.^[11] Previous work had already indicated that cells could be cultivated on different collagen gels. Furthermore, it was then shown that the incorporation within the depth of these matrices with various types of cells was possible.^[22,23]

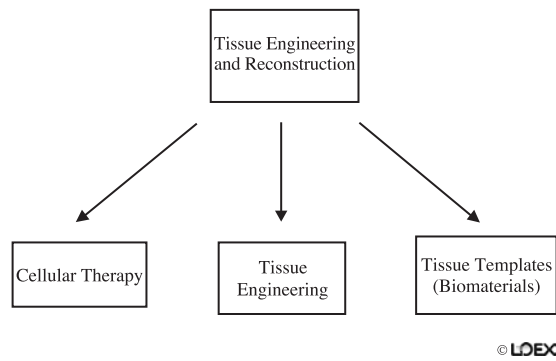


Fig. 1 Overview of the different approaches leading to tissue engineering and reconstruction.

However, the true integration of cells into gels allowed them to reorganize the surrounding matrix.^[12,24–30] A few investigators are pursuing this technology with either collagen gels or other matrix components such as fibrin.

Even though this approach allows the creation of many substitutes and can recreate excellent *in vitro* models, it has a few drawbacks. The main one is the weak mechanical resistance of the obtained substitutes. The structural integrity may be sufficient for such tissues as skin, but not so for substitutes in the vascular or orthopedic system. This problem has been addressed with some original solutions, such as glycation and magnetic alignment of the collagen fibers.^[31–33] It remains to be seen if these modifications are truly applicable to a clinical setting.

A second approach entails the seeding of cells into scaffolds. These scaffolds can be of a natural or bio-synthetic origin. The cells then thrive in the porous material and secrete various amounts of extracellular matrix depending on their nature. Much of the original concept was developed by Dr. Robert Langer's group at MIT.^[34] It also builds on the work of Dr.

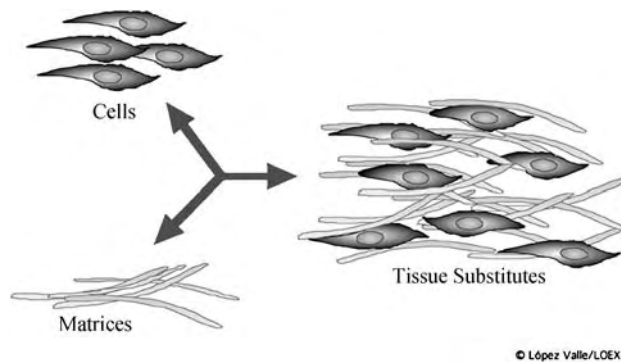


Fig. 2 Simplified illustration of the components necessary for tissue engineering.

Burke,^[35,36] Dr. Yannas, and Dr. Odile Damour,^[37,38] with the creation of sponge-like structures from mainly collagenous material. Most of the MIT work has centered around synthetic scaffolds such as PGA [Poly(glycolic acid)],^[34] There have been countless modifications and additions to the different types of synthetic materials used over the last decade in this approach of tissue engineering. The obvious advantage of this approach is the immediate creation of a three-dimensional structure that already has significant structural properties. However, the intrinsic nature of most of these polymers, which are suture materials, entails slow degradation with an ensuing lowering of the pH of surrounding tissues. This leads to a slow but rather protracted low-level inflammatory process. Many groups are thus attempting to alter the chemical nature of these biomaterials for not only better tissue acceptance, but also better control over the resorption time frame and acidic conditions. These efforts would inhibit the inflammatory process and ensure that the disappearance of the material is synchronous with the integration and constructive phase of the graft. How successful these attempts are will be apparent within the next few years.

More recently, the LOEX group in Canada has developed a different approach to tissue engineering. It has some roots in the very initial phases of *in-vitro* tissue reconstruction as exemplified by Dr. Howard Green with cultured epithelia for burn patient therapy,^[39] combined with many observations about the reorganization of the extracellular matrix in various substitutes. In this approach, various types of cells, mostly of mesenchymal origin, are grown in such a fashion within a culture flask that they literally embed themselves in their very own extracellular matrix. Among many factors, the addition of sodium ascorbate allows the significant appearance of the various components of the extracellular matrix. These sheets are then either stacked or rolled to obtain various tissue substitutes.^[17] The most convincing demonstration was the re-creation of a totally biological vascular substitute. The final substitute was not only three-dimensional and three-cellular, but also had very valuable mechanical and biological properties. The main advantage in this approach is the absence of extraneous collagens and any synthetic material.^[40] This method is truly close to some of the *in vivo* processes found during organogenesis *in utero*.

This team has also championed for many years the use of self-generated mechanical forces with additional mechanical stresses also applied to these substitutes. This combination not only leads to a significant *in vitro* cellular and extracellular matrix reorganization, but also achieves a most physiological final result.^[29,30] It is felt that such substitutes could be grafted with minimal inflammatory response, but also with a rather

unique state of readiness for *in vivo* integration and remodeling.

However, one must also note that this approach is time-consuming. Thus, it will demand some form of automation and acceleration of maturation time for the culture of these substitutes. On the other hand, they can be considered, at the present time, as the purest form of *in vitro* models for various physiological, pathophysiological, pharmaceutical, and toxicological studies.

CRITICAL INTEGRATIVE ASPECTS OF TISSUE-ENGINEERED SUBSTITUTES

As tissue engineering enters the clinical area in a more and more significant fashion each passing year, some very critical aspects of tissue integration must be addressed.

Two of these critical aspects, namely vascularization and innervation, have not received all the attention they deserve. This may be due to the fact that most of the efforts leading to clinically applicable substitutes centered at first on skin grafts. Then, because of their very nature, the animal experiments did not reveal the need to take into account these two important aspects of tissue integration.

In a way, this may be seen as quite surprising since reconstructive surgery devotes a lot of attention to the revascularization and reinnervation of autologous or allogeneic grafts. This is evidently less true if a graft has a nearly exclusively mechanical function such as an aortic synthetic graft or hip prosthesis. But even then some surgeons and scientists have pointed out that some biological responses to the graft should be enhanced for a better integration. As an example, the team of Dr. Zilla has extensively studied and optimized the seeding of endothelial cells on the internal side of vascular grafts.^[41] The field of tissue engineering is now much more attuned to the necessity of responding to such a challenge.

In regard to vascularization, even with a paucity of results at first, the approach of stimulating the ingrowth of blood vessels into solid organs has not been successful. Such organs rapidly, within hours or even minutes, demand blood irrigation for survival and proper function. This necessity was the basis for a new endothelialized skin substitute developed at LOEX.^[42] Drawing on clinical lessons showing that cadaver skin could rapidly take and demonstrate a capillary blood flow, the team of Dr. Auger has strived to recreate a microvascular system in the dermal component of skin substitutes. Such a design was even more plausible with understanding of the phenomenon that allows the rapid take of cadaver skin in burn patients. Inosculation, the physiological anastomosis between blood

vessels of the graft and the wound bed, was only recently clearly described in an animal model.^[43] Thus, at first using a scaffold technique, this group has combined in the dermal layer fibroblasts and endothelial cells in such a fashion that a capillary-like system was reconstituted. A series of analyses, including histology, immunohistology and electron microscopy, has shown the microvascular nature of this endothelialized skin equivalent. Furthermore, preliminary grafting experiments in animals have shown that blood flow within the graft is reestablished at a faster rate than in substitutes without a capillary-like system.^[44]

There is now a flurry of interest in such microvascularization in various soft tissues.^[45] But the next level of complexity will undoubtedly obtain for solid organs the full spectrum, from small arteries to capillaries and finally outflow small-diameter veins. This certainly is a lofty goal and will necessitate a tremendous multidisciplinary effort.

Of a less immediate nature is the matter of tissue substitute reinnervation. Too little attention has been focused on this element if the final goal is full tissue integration. Most tissues have many types of neurological receptors that play an important role in not only the biological function but also ultimately their very own homeostasis. Thus, these receptors allow the body to receive the appropriate messages, from nociception to positional information, which then translate into local and frequently systemic reactions. This loop is an integral part of the true physiological function, we are hoping for, in order for tissue-engineered organs to obtain their full potential. Not only will the appropriate neurological signals be sent by the grafted tissues, but many physiological responses will ensue or be set in motion, such as vasoreactivity, hormonal secretion, and enzymatic delivery.

Until now efforts have been mostly centered on evaluating the reappearance of the neural network in various tissues. Once again, the most studied tissue has been the skin. Early experiments have shown the level and time schedule of pain perception reappearance in the first generation of epidermal tissue grafts for burn patients.^[46,47]

Our LOEX group has now focused its observation on some reconstructed tissues such as full-thickness reconstructed skin substitutes for reinnervation after animal implantation. The results are quite encouraging.^[48] But once again the challenges will be much more complex in large solid organs with various cell types. This is even more true when there is a specialized nervous tissue within the organ with a precise physiological function, such as the cardiac pacing system.

Another layer of complexity is added when one considers secretory or filtrating organs such as the pancreas,

liver, or spleen. The successful three-dimensional arrangement of such organs is at the present time a daunting task, but the future may hold answers that are not yet foreseeable. Some investigators have targeted embryonic stem cells as an answer to such complex issues. It remains to be seen if the science of tissue engineering will ever find the right condition for such a demanding differentiation program to occur *ex vivo*.

ENABLING TECHNOLOGIES

Some technologies can be seen as being of incontrovertible value on the path to better tissue engineering. We shall review a few of them.

Tissue Morphometrics

Some scientists in the field of tissue engineering having noted the lack of quantitative and qualitative tissue values and analysis and have proposed measures to remedy this perceived shortcoming.

With an appropriate database for various morphometric tissue parameters, the science of reconstructive medicine should be more accurate and lead to better results. This ongoing effort is certainly quite interesting, and its impact should be felt in the near future according to its proponents.

New Biomaterials

The search for innovative biomaterials seems to be developing along two pathways that will probably intertwine. The first avenue is a worldwide effort, already alluded to, in searching for biomaterials that have better integrative properties. This entails a minimal inflammatory reaction to such materials combined with optimal resorption time depending on the targeted tissues or organs. One must note that inflammation will always be present since any surgical operation entails such a reaction. However, it is quite probable that the least reaction, the better in this context.

The second avenue is to attach signaling molecules to the utilized biomaterial. This can be as direct as a slow, sequential release of various cytokines or growth factors. However, the more technically sophisticated approach of adding DNA sequencing for appropriate messaging may be of high value for all those seeking to induce favorable integration of these biomaterials.^[49]

Mechanical Stimulation

In the last decade it has become quite clear that nearly all body tissues can attain a higher level of complexity

by being submitted to mechanical stimulation.^[12] Although such a concept was rather evident in the orthopedic field, where load bearing is known to be of great importance in tissue differentiation, it also turned out to be valuable for other tissues, such as blood vessels, heart valves, and skin.^[26,32,34]

Bioreactors

If one wants to combine the previous technology with any significant scaling-up process, then bioreactors are a pivotal step in tissue engineering. A great part of the answer lies not only in obtaining large amount of cells, but also exposing them to the appropriate mechanical conditions.

Furthermore, this is a critical pathway for automating at least some of the processes involved in re-creating tissue substitutes.

Gene Therapy

If the combining of DNA into biomaterial has already been proposed, one must also look at gene therapy as an additional and more permanent technology for obtaining better tissue-engineered organs.

For example, in re-creating the pancreas, at the present time a high number of Langerhans cells are necessary for an appropriate level of insulin secretion. If these cells were amenable to higher levels of secretion, this would be of great advantage in the treatment of diabetic patients.

Another example is reconstructed blood vessels with anti-atherosclerotic molecules secreted by either smooth muscle cells or endothelial cells. Here, the patient most prone to restenosis could benefit from a superior type of graft. However, this technology brings these substitutes into an entirely different regulatory environment. It may be a step that will be valuable, but the time frame of such a step is difficult to predict today.

This does not deter many groups working on various types of “universal” donor cells. Such cells, devoid of specific immunological markers, usually targeted in rejection, would allow the creation of well-tolerated substitutes. But once again, this task may entail more complexity than meets the eye.^[50]

MULTIDISCIPLINARY TEAMWORK IN TISSUE ENGINEERING

It seems to us that the best way to meet all these challenges is to combine various types of expertise that have not always been tied together in the past. Many



regional and national programs in tissue engineering are addressing this issue.

The LOEX group has always brought together a triad of specialties into its tissue engineering efforts:

- Cell and molecular biologists
- Bio-engineers
- Biomedical clinical investigators

When we set about the project of re-creating ligaments, the orthopedic surgeons declared the anterior cruciate ligament to be of the utmost clinical importance. They also gave the team a sense of alternative therapies and what additional advantages a tissue-engineered substitute should bring. This enabled the team to reach a set of biological and mechanical criteria for each substitute. This combination has proven to be quite fruitful in forging ahead with tissue engineering projects that are clinically focused. Thus, as a first step the biomedical clinical investigators help the whole team set the stage in pinning down the most significant goals to be achieved for a particular substitute. This interaction helps in determining the most acute need in tissue reconstruction for any body system. Then the cell biologists strive to reconstruct the organ or tissue with what is deemed to be the best tissue engineering approach. The pluses and minuses of each approach have already been presented and will, in a sense, dictate the best fit. One of the most crucial roles of the cell biologists at this point is to obtain the various cells necessary for the reconstruction. Their challenge is not only to find appropriate sources of cells, but also to establish efficacious extraction methods. The purity of the cells is of paramount importance if the tissue engineering effort is to be successful.

Afterward, the expansion of cells must be efficacious enough to obtain the amount of cells and the appropriate quality necessary for the reconstruction. The cell and molecular biologists will have also settled on some functional biological requirements for the tissue-engineered organs. This may be a cellular function such as the antithrombotic properties of endothelial cells or an exocrine secretory function such as insulin for pancreatic Langerhans cells.

The role of bio-engineers is also of great value. They frequently can create or refine bioreactors to facilitate and accelerate the expansion of the necessary cells. They maintain a close relationship with the biologists so that the processes are not deleterious to the cellular phenotypes and function. But usually at the inception of a project, the biomechanical engineers are in close interaction to create different types of apparatus that will induce the appropriate mechanical strains in the reconstructed tissues. These strains may appear passively or be induced by machines that, for example, bring pulsatile flow to a reconstructed blood vessel.

There is now clear evidence that such mechanical preparation is advantageous not only for the structural integrity, but also for some functions of tissue-engineered substitutes. Many cells have been shown to acquire their *in vivo* phenotype when exposed *in vitro* to strains similar to those encountered in the human body.^[12,30,34,51]

All this multidisciplinary work forms a working loop of interaction between teams. This is quite understandable, as advances in each sector have some bearing on the other members of the team.

There is also a two-step paradigm to our efforts in a given tissue engineering project. The first step, while taking into account what has been previously described, aims at recreating a three-dimensional structure that will have a satisfactory histological aspect. However, the biological and mechanical properties are not the main purpose. The team strives to demonstrate the possibility of assembling the various cells and tissues to form an organ that will have acceptable tissular organization and enough stability to maintain it in culture. The second step is optimization of the desired mechanical and biological functions. This leads to an intensive collaboration between the biologists and the engineers.

This two-step paradigm can be well observed in our own effort at recreating a tissue-engineered blood vessel (TEBV). Our first publication, using the self-assembly approach to tissue engineering, had allowed us to recreate all three layers of a blood vessel.^[40] The histological analysis revealed well-defined adventitia, media, and intima with no crossover contamination from the different cell types between layers. Furthermore, a few very important functions were shown to be present, such as incorporation of acetylated LDL by endothelial cells and antithrombogenic activity.^[51] But no vasoreactivity was noted, and the mechanical resistance was very weak. This TEBV was thus not regarded as acceptable for grafting. The second generation, utilizing the self-assembly approach, answered those pitfalls in an interesting manner. The vasoreactivity of these vessels was clearly established as their supraphysiological burst pressure settled around 2500 mm Hg.^[40,52] This is a fitting example to a stepwise road to the completion of a tissue-engineered organ *in vitro*. Thereafter the role of the clinician is of paramount importance in creating protocols for grafting these substitutes first in animals and then into the human host.

CONCLUSION

Tissue engineering has opened fascinating perspectives in the biomedical armamentarium. The astounding clinical demand for such biotechnological solutions ensures that this research domain will continue to forge ahead.

On the other hand, the road to success has also a few meanders that should be carefully addressed if this field is to keep its pace. There is a tremendous level of trepidation related to the expanding knowledge of both embryonic and adult stem cells. The question surrounding the ethical acceptability of using embryonic stem cells is already partly out of the scientists' hands. Various nations will have to define the boundaries of acceptable utilization of embryonic stem cells.

Next we will have to solve the important aspect of host acceptance of the tissue-engineered graft. There is no problem for an autologous construct; however, allogeneic grafting brings about the eventual problem of rejection. If tissue engineers want to stay clear of the immunosuppressive drug used in classical organ transplantation, new solutions will be necessary. The two main efforts have, up to now, aimed at either rendering the graft less of a target or inducing tolerance in the receiver. Here, tissue engineers are accompanied by all the transplantation community, thus heightening the possibilities of successes.

Finally, the regulatory aspect will eventually become clearer. The examining agencies in many countries are responding to these new therapies. However, the channel of communication must be well established between regulators and tissue engineers to ensure a safe but not too constraining introduction of this form of therapy into the clinical arena.

Whatever lays ahead for tissue engineering, it is already clear that this field will alleviate much suffering and improve life for many patients around the world.

ACKNOWLEDGMENTS

The authors are grateful to all LOEX investigators and their external collaborators in the development of reconstructed living substitutes by tissue engineering, and to all members of the LOEX laboratory for their kind help, advice, and technical assistance in relation to the work presented in this review. Further information can be found at <http://www.fmed.ulaval.ca/loex>.

ARTICLES OF FURTHER INTEREST

Biologic and Synthetic Apatites; Biomedical Ethics; Bioreactors; Extracellular Matrix Scaffolds; Tissue Engineering Scaffolds

REFERENCES

- Saies, A.D.; Urbaniak, J.R.; Nunley, J.A.; Taras, J.S.; Goldner, R.D.; Fitch, R.D. Results after replantation and revascularization in the upper extremity in children. *J. Bone Jt. Surg., Am.* **1994**, *76* (12), 1766–1776.
- Qing-tai, L.; Chang-qing, J.Z.; Ke-fei, Y.; Sha-ling, C.; Jan, L.; Zun-ying, L. Successful replantation in 10-digit complete amputations. *Plast. Reconstr. Surg.* **1996**, *98* (2), 348–353.
- Chen, Z.; Zhang, J. Replantation of severed limbs: Current status and prospects. *Chin. Med. J. (Engl.)* **1999**, *112* (10), 914–917.
- Caroli, A.; Adani, R.; Castagnetti, C.; Squarzina, P.B.; Pancaldi, G.; Cristiani, G. Replantation and revascularization of large segments of the hand and forearm. *Ital. J. Orthop. Traumatol.* **1991**, *17* (4), 433–447.
- Buncke, H.J. Microsurgery—retrospective. *Clin. Plast. Surg.* **1986**, *13* (2), 315–318.
- Clowes, A.W.; Kohler, T.R. Anatomy, Physiology, and Pharmacology of the Vascular Wall. In *Vascular Surgery: A Comprehensive Review*, 5th Ed.; Moore, W.S., Ed.; W.B. Saunders Company: Los Angeles, 1998; 35–43.
- Whalen, E.; Donnelly, T.A.; Naughton, G.; Rheins, L.A. The development of three-dimensional in vitro human tissue models. *Hum. Exp. Toxicol.* **1994**, *13*, 853–859.
- Germain, L.; Remy-Zolghadri, M.; Auger, F. Tissue engineering of the vascular system: From capillaries to larger blood vessels. *Med. Biol. Eng. Comput.* **2000**, *38* (2), 232–240.
- Langer, R.; Vacanti, J.P. Tissue engineering. *Science* **1993**, *260*, 920–926.
- Naughton, G.K.; Mansbridge, J.N. Human-based tissue-engineered implants for plastic and reconstructive surgery. *Clin. Plast. Surg.* **1999**, *26* (4), 579–586.
- Bell, E.; Ehrlich, H.P.; Buttle, D.J.; Nakatsuji, T. Living tissue formed in vitro and accepted as skin-equivalent tissue of full thickness. *Science* **1981**, *211*, 10452–10454.
- Goulet, F.; Rancourt, D.; Cloutier, R.; Germain, L.; Poole, A.; Auger, F.A. Tendons and Ligaments. In *Principles of Tissue Engineering*; Lanza, R., Langer, R., Vacanti, J., Eds.; Academic Press Ltd.: San Diego, 2000; 711–722.
- Paquette, J.-S.; Goulet, F.; Boulet, L.-P.; Laviolette, M.; Tremblay, N.; Chakir, J.; Germain, L.; Auger, F.A. Three-dimensional production of bronchi in vitro. *Can. Respir. J.* **1998**, *5*, 43.
- Badylak, S.F.; Kropp, B.; McPherson, T.; Liang, H.; Snyder, P.W. Small intestinal submucosa: A rapidly resorbed bioscaffold for augmentation cystoplasty in a dog model. *Tissue Eng.* **1998**, *4* (4), 379–387.
- Rawe, J. What Will be the 10 Hottest Jobs?. In *Time*; May 22, 2000; 54–55.
- Arnst, C.; Carey, J. Biotech Bodies. In *Business Week*; July 27, 1998; 56–63.
- Auger, F.A.; Rémy-Zolghadri, M.; Grenier, G.; Germain, L. The self-assembly approach for organ reconstruction by tissue engineering. *E-biomed* **2000**, *1*, 75–85.
- Shamblott, M.J.; Axelman, J.; Wang, S.; Bugg, E.M.; Littlefield, J.W.; Donovan, P.J.; Blumenthal, P.D.; Huggins, G.R.; Gearhart, J.D. Derivation of pluripotent stem

- cells from cultured human primordial germ cells. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95* (23), 13726–13731.
19. Thomson, J.A.; Itskovitz-Eldor, J.; Shapiro, S.S.; Waknitz, M.A.; Swiergiel, J.J.; Marshall, V.S.; Jones, J.M. Embryonic stem cell lines derived from human blastocysts. *Science* **1998**, *282*, 115–117.
 20. Bishop, A.E.; Buttery, L.D.; Polak, J.M. Embryonic stem cells. *J. Pathol.* **2002**, *197* (4), 424–429.
 21. Poulosom, R.; Alison, M.R.; Forbes, S.J.; Wright, N.A. Adult stem cell plasticity. *J. Pathol.* **2002**, *197* (4), 441–456.
 22. Auger, F.A.; Lopez-Valle, C.A.; Guignard, R.; Tremblay, N.; Noël, B.; Goulet, F.; Germain, L. Skin equivalent produced with human collagens. *In Vitro Cell. Dev. Biol., Anim.* **1995**, *31*, 432–439.
 23. Karasek, M.A.; Charlton, M.E. Growth of post-embryonic skin epithelial cells on collagen gels. *J. Invest. Dermatol.* **1971**, *56* (3), 205–210.
 24. Auger, F.; Rouabhia, M.; Goulet, F.; Berthod, F.; Moulin, V.; Germain, L. Tissue-engineered human skin substitutes developed from collagen-populated hydrated gels: Clinical and fundamental applications. *Med. Biol. Eng. Comput.* **1998**, *36*, 801–812.
 25. Auger, F.A.; Pouliot, R.; Tremblay, N.; Guignard, R.; Noel, P.; Juhasz, J.; Germain, L.; Goulet, F. Multistep production of bioengineered skin substitutes: Sequential modulation of culture conditions. *In Vitro Cell. Dev. Biol., Anim.* **2000**, *36* (2), 96–103.
 26. Lopez Valle, C.A.; Auger, F.A.; Rompré, P.; Bouvard, V.; Germain, L. Peripheral anchorage of dermal equivalents. *Br. J. Dermatol.* **1992**, *127*, 365–371.
 27. Bouvard, V.; Germain, L.; Rompre, P.; Roy, B.; Auger, F.A. Influence of dermal equivalent maturation on the development of a cultured skin equivalent. *Biochem. Cell. Biol.* **1992**, *70* (1), 34–42.
 28. Tamariz, E.; Grinnell, F. Modulation of fibroblast morphology and adhesion during collagen matrix remodeling. *Mol. Biol. Cell* **2002**, *13* (11), 3915–3929.
 29. Goulet, F.; Germain, L.; Caron, C.; Rancourt, D.; Normand, A.; Auger, F. Tissue-Engineered Ligament. In *Ligaments and Ligamentoplasties*; Yahia, L., Ed.; Springer-Verlag: Berlin, 1997; 367–377.
 30. Goulet, F.; Germain, L.; Rancourt, D.; Caron, C.; Normand, A.; Auger, F.A. Tendons and Ligaments. In *Textbook of Tissue Engineering*; Lanza, R., Langer, R., Chick, W., Eds.; Academic Press Ltd.: San Diego, 1997; 633–644.
 31. Girton, T.S.; Oegema, T.R.; Grassl, E.D.; Isenberg, B.C.; Tranquillo, R.T. Mechanisms of stiffening and strengthening in media-equivalents fabricated using glycation. *J. Biomech. Eng.* **2000**, *122* (3), 216–223.
 32. Tranquillo, R.T.; Girton, T.S.; Bromberek, B.A.; Triebes, T.G.; Mooradian, D.L. Magnetically orientated tissue-equivalent tubes: Application to a circumferentially orientated media-equivalent. *Biomaterials* **1996**, *17* (3), 349–357.
 33. Seliktar, D.; Black, R.A.; Vito, R.P.; Nerem, R.M. Dynamic mechanical conditioning of collagen-gel blood vessel constructs induces remodeling in vitro. *Ann. Biomed. Eng.* **2000**, *28* (4), 351–362.
 34. Niklason, L.E.; Gao, J.; Abbott, W.M.; Hirschi, K.K.; Houser, S.; Marini, R.; Langer, R. Functional arteries grown in vitro. *Science* **1999**, *284*, 489–493.
 35. Yannas, I.V.; Burke, J.F. Design of an artificial skin: Basic design principles. *J. Biomed. Mater. Res.* **1980**, *14*, 65–81.
 36. Yannas, I.V.; Lee, E.; Orgill, D.P.; Skrabut, E.M.; Murphy, G.F. Synthesis and characterization of a model extracellular matrix that induces partial regeneration of adult mammalian skin. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86* (3), 933–937.
 37. Berthod, F.; Hayek, D.; Damour, O.; Collombel, C. Collagen synthesis by fibroblasts cultured within a collagen sponge. *Biomaterials* **1993**, *14*, 749–754.
 38. Berthod, F.; Sahuc, F.; Hayek, D.; Damour, O.; Collombel, C. Deposition of collagen fibril bundles by long-term culture of fibroblasts in a collagen sponge. *J. Biomed. Mater. Res.* **1996**, *32*, 87–94.
 39. Green, H.; Kehinde, O.; Thomas, J. Growth of cultured human epidermal cells into multiple epithelia suitable for grafting. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76* (11), 5665–5668.
 40. L'Heureux, N.; Pâquet, S.; Labbé, R.; Germain, L.; Auger, F.A. A completely biological tissue-engineered human blood vessel. *FASEB J.* **1998**, *12*, 47–56.
 41. Meinhart, J.G.; Deutsch, M.; Fischlein, T.; Howanietz, N.; Froschl, A.; Zilla, P. Clinical autologous in vitro endothelialization of 153 infrainguinal ePTFE grafts. *Ann. Thorac. Surg.* **2001**, *71* (5 Suppl.), S327–S331.
 42. Black, A.F.; Berthod, F.; L'Heureux, N.; Germain, L.; Auger, F.A. In vitro reconstruction of a human capillary-like network in a tissue-engineered skin equivalent. *FASEB J.* **1998**, *12*, 1331–1340.
 43. Young, D.M.; Greulich, K.M.; Weier, H.G. Species-specific in situ hybridization with fluorochrome-labeled DNA probes to study vascularization of human skin grafts on athymic mice. *J. Burn Care Rehabil.* **1996**, *17*, 305–310.
 44. Auger, F.A.; Black, A.; Berthod, F.; L'Heureux, N.; Germain, L. In Vitro Reconstruction of a Human Capillary-Like Network in a Tissue-Engineered Skin Equivalent. *World Congress on Medical Physics and Biomedical Engineering, Nice, 1997*.
 45. Supp, D.M.; Wilson-Landy, K.; Boyce, S.T. Human dermal microvascular endothelial cells form vascular analogs in cultured skin substitutes after grafting to athymic mice. *FASEB J.* **2002**, *16* (8), 797–804.
 46. Compton, C.C.; Gill, J.M.; Bradford, D.A.; Regauer, S.; Gallico, G.G.; O'Connor, N.E. Skin regenerated from cultured epithelial autografts on full-thickness burn wounds from 6 days to 5 years after grafting. A light, electron microscopic and immunohistochemical study. *Lab. Invest.* **1989**, *60* (5), 600–612.
 47. English, K.B.; Stayner, N.; Krueger, G.G.; Tuckett, R.P. Functional innervation of cultured skin grafts. *J. Invest. Dermatol.* **1992**, *99* (2), 120–128.
 48. Gingras, M.; Paradis, I.; Berthod, F. Nerve regeneration in a collagen-chitosan tissue-engineered skin transplanted on nude mice. *Biomaterials* **2003**, *24*, 1653–1661.
 49. Richardson, T.P.; Murphy, W.L.; Mooney, D.J. Polymeric delivery of proteins and plasmid DNA for tissue

- engineering and gene therapy. *Crit. Rev. Eukaryot. Gene Expr.* **2001**, *11* (1–3), 47–58.
50. Rouabhia, M.; Germain, L.; Bergeron, J.; Auger, F.A. Allogeneic-syngeneic cultured epithelia. *Transplantation* **1995**, *59* (9), 1229–1235.
51. L'Heureux, N.; Germain, L.; Labbe, R.; Auger, F.A. In vitro construction of a human blood vessel from cultured vascular cells: A morphologic study. *J. Vasc. Surg.* **1993**, *17* (3), 499–509.
52. L'Heureux, N.; Stoclet, J.C.; Auger, F.A.; Lagaud, G.; Germain, L.; Andriantsitohaina, R. A human tissue-engineered vascular media: A new model for pharmacological studies of contractile responses. *FASEB J.* **2001**, *15*, 515–524.

Tissue Engineering: AC Electrokinetics

Gerard H. Markx

Microstructures and Microenvironment Research Group, School of Chemical Engineering and Analytical Science, The University of Manchester, Manchester, U.K.

Anne-Marie Buckle

School of Life Sciences, The University of Manchester, Manchester, U.K.

INTRODUCTION

The characterization, separation, and positioning of cells are all subjects of importance in tissue engineering. Alternating current (AC) electrokinetics, the movement of particles (including cells) in AC electric fields, can be used to do all of this—and more! To observe AC electrokinetic effects, very high electric field strengths are needed. Such electric fields are readily produced with quite small voltages using microelectrodes. Such microelectrodes are nowadays readily produced using microfabrication techniques borrowed from the electronics industries. Alternating current electrokinetic techniques are highly flexible because electrical signals are easy to control, individual electrodes can be addressed, and the direction of induced particle movement can depend on the AC frequency. The small scales at which electrokinetic devices can be produced makes AC electrokinetics particularly suitable for the manipulation and characterization of single cells and the separation of small numbers of cells. In addition, because arrays of electroactive regions are easily produced, the method is suitable for manipulating many cells simultaneously. Especially interesting for tissue engineering is the emerging use of AC electrokinetic microdevices for two- and three-dimensional (3D) patterning of cells. This could be used to create microniches and microenvironments *in vitro*.

WHAT IS AC ELECTROKINETICS?

The term AC electrokinetics describes the movement of particles, including cells, in AC electric fields.^[1–4] With its direct current (DC) counterpart electrophoresis, ultrasound, and laser trapping AC electrokinetics forms a set of physical techniques that can be used for the manipulation of cells.^[5–7] Each physical manipulation technique has its own advantages and disadvantages (Table 1). In an alternating electric field the electric field continuously changes direction, from a few times per second or less to many million times per second. Unlike electrophoresis, which occurs in DC electric fields, particle movement is not dominated by

the surface charge of the particles, but by dipole effects. A dipole is a particle in which equal but opposite electrical charges are present at two opposite sides of the particle. Dipoles can be permanent, or induced by external electric fields. How easy it is to induce a dipole in a particle depends on its polarizability. The polarizability of materials is in turn determined by their conductivity and dielectric constant (permittivity). For many materials, and cells in particular, the values of the conductivity and permittivity change with frequency. As a result, the AC electrokinetic behavior of the cells is highly dependent on the frequency used.^[2,3] At low frequencies (typically <10 kHz), polarization mechanisms dominate that are caused by the surface charge on the cell surface and the ion double layers formed on the surface. At midrange frequencies (10 kHz–100 MHz), the cell electrical properties are dominated by interfacial polarization mechanisms occurring at the interfaces between the cell membrane, the cytoplasm, and the medium. At higher frequencies, polarization of the molecules and atoms of which a cell consists dominates. The midfrequency range, at which (Maxwell–Wagner-type) polarization occurs at the interfaces of the structures that constitute a cell, is the one that is most used and studied in AC electrokinetics. Because an intact and viable cell is surrounded by a membrane barrier that has a low dielectric constant (~ 6) and low conductivity ($\sim 10^{-7}$ S/m) compared to the cytoplasm (dielectric constant/permittivity ~ 50 – 200 , conductivity ~ 0.1 S/m) and the suspending medium (dielectric constant of pure water ~ 79 , conductivity depends on the amounts of dissolved salts), a cell has very interesting dielectric properties that are highly frequency dependent (Fig. 1).

Although it is possible to observe AC electrokinetic effects using relatively large electrodes (centimeters) and large voltages (kilovolts), it is often more convenient to demonstrate AC electrokinetic effects using microelectrodes (typically 10–100 μm) and relatively modest voltages (typically 0–20 $V_{\text{pk-pk}}$).^[8,9] The added advantage of using microelectrodes is that heating effects are reduced because heat is more quickly dissipated. Microelectrodes are most conveniently made

Table 1 Comparison of different physical particle manipulation techniques*Electrophoresis*

Directional; electrodes needed; manipulation of many cells simultaneously possible; medium conductivity important; heating effects important

AC electrokinetics

Large frequency range (Hz–GHz); directional (positive and negative forces); frequency-dependent movement; separation of cells according to different electrical properties possible; conductivity affects movement; low conductivity often necessary; manipulation directly in growth medium difficult; high spatial resolution; microelectrodes needed in close vicinity of cells; high control of cell movement; possible to manipulate many cells simultaneously

Ultrasound

Standing waves; mainly nondirectional, so limited range of movements; manipulation can be done directly in growth medium; limited frequency range (typically 1–10 MHz); separation of cells according to differences in acoustic properties difficult; low spatial resolution; ultrasound sources can be placed at long distances; manipulation of many cells simultaneously possible

Optical traps (optical tweezers)

Limited range of movements; high spatial resolution; high control over cell movement; manipulation can be done directly in growth medium; laser light source can be at a long distance away from cells; possibility to manipulate many cells simultaneously limited

All techniques suffer to some extent from heating effects; these are particularly important for electrical techniques when used directly in growth medium.

using photolithography, although other techniques such as laser ablation are also sometimes used. Electrodes have been made of a variety of materials, including gold, titanium, chromium, aluminum, platinum, indium tin oxide (ITO). Electrodes of titanium and platinum tend to be more durable, while electrodes from ITO are both durable and transparent. More often than not the substrate is glass, although other materials such as plastic and silicon can and have also been used. Electrode shape determines the electric field pattern. Electrode size is typically a small number of multiples of the diameter of the particle (cell) being investigated. Structures can be made in which each electrode is individually addressed, but often electrodes are of a design in which the electrodes are connected to common bars to simplify the electrode construction process. Some designs make use of insulating layers deposited on or near electrodes to create local high-field regions. Examples of electrode designs are shown in Fig. 2. Typically, a signal generator producing sinusoidal signals is used to generate the electric fields. For some effects, for example, electrorotation (ROT) and traveling wave dielectrophoresis (DEP), sources that have four outputs, which are 90° phase shifted relative to each others are normally used.

DESCRIPTION OF DIFFERENT AC ELECTROKINETIC TECHNIQUES

A number of different AC electrokinetic techniques exist.^[1,2] Fig. 3 shows some of the techniques.

Electrostriction

Electrostriction refers to the mechanical distortion of particles in electric fields. Electrostriction of animal cells can be readily observed at high electric field strengths, and shows up as a distortion of the cell. If the electric field becomes too strong the cells can be destroyed.^[10]

Electro-orientation

Electro-orientation refers to the orientation of nonspherical particles in electric fields. At different frequencies, in strong electric fields, nonspherical cells would orient with different axes along electric field lines.^[3,11] For example, ellipsoidal erythrocytes with three different axes of rotation have been shown to orient along three different axes at different frequencies.^[11] The technique has been used in cell characterization, and could potentially be used in the creation of tissues containing oriented cells.^[12]

Normal DEP

Dielectrophoresis, also called normal dielectrophoresis to distinguish it from traveling wave dielectrophoresis, is the most studied AC electrokinetic effect. (Normal) DEP is the movement of particles in nonuniform electric fields. It has been used for cell characterization, cell separation, particle trapping, and the production of structured biomaterials, as well as bringing cells together for (electro)fusion.^[1,8,13–32] Positive and

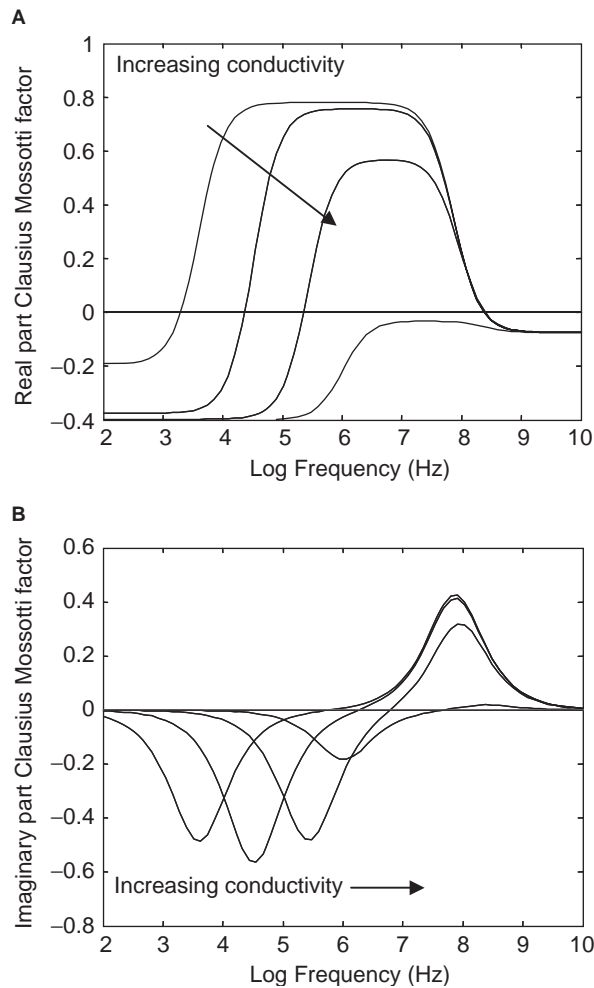


Fig. 1 Real (A) and imaginary (B) parts of the Clausius–Mossotti factor $(\epsilon_p^* - \epsilon_m^*) / (\epsilon_p^* + 2\epsilon_m^*)$ calculated for a mammalian cell using the multishell model. (From Ref. [13].) ϵ_p^* is the complex permittivity of the particle (cell), ϵ_m^* the complex permittivity of the suspending medium. $\epsilon^* = \epsilon - j\sigma/\omega$, where ϵ is the permittivity, σ the conductivity, and ω the frequency (in radians per second). $j = \sqrt{-1}$. The values for the properties of the different structures in the cell (cytoplasm, membrane) were taken as in Wang et al., except for $\epsilon_{\text{cytoplasm}} = 60$. (From Ref. [34].) Conductivity values used were 10^{-3} , 10^{-2} , 10^{-1} , and 1 S/m. The real and imaginary parts of the Clausius–Mossotti factor can in turn be used to calculate the different electrokinetic effects such as the dielectrophoretic force, electrorotation rate, etc. (From Ref. [2].) The dielectrophoretic force is determined by the real part of the Clausius–Mossotti factor. As a result, (A) clearly shows the effects of frequency and conductivity on the dielectrophoretic force.

negative DEP is possible, depending on whether particles move toward high-field regions between the electrodes (positive DEP) or away from them (negative DEP) (Fig. 4). Whether positive or negative DEP occurs depends on the polarizability of the particle relative to its suspending medium.^[9] If it is more

polarizable, positive DEP occurs, if less, negative. For animal cells to show positive DEP it is necessary to suspend the cells in low-conductivity iso-osmotic buffers. Negative DEP can be induced in low-conductivity buffers using low frequencies (typically below 50 kHz) or high frequencies (typically over 50 MHz), or at any frequency at very high medium conductivities (see Fig. 1). Especially at low conductivities and low frequencies, strong forces are exerted on the cell membrane, which, if strong enough, can disrupt cells.^[32] The low concentration of essential ions and other media components can also affect the cells adversely, and the exposure of the cells to these conditions should be kept to a minimum. At high medium conductivities the temperature of the suspending medium can be significantly increased by the applied electrical currents.^[17] It is, therefore, often necessary at high medium conductivities to limit the exposed electrode surface area, making it more difficult to manipulate many cells simultaneously.^[17] When applying electric fields over longer periods of time, hydrogen peroxide can be produced, which can affect cells adversely. Catalase can be added to the medium to overcome this problem.^[33]

Electrorotation

Electrorotation is the rotation of particles in electric fields.^[13,34] The ROT rate is dependent on the imaginary part of the Clausius–Mossotti factor (Fig. 1).^[2] Clockwise and anticlockwise rotation is possible, depending on the polarizability of the particle relative to that of the suspending medium. Although ROT can also sometimes be observed in nonphase shifted electric fields, it is best observed by applying four 90° phase-shifted sinusoidal signals to four electrodes surrounding a particle under investigation (see Fig. 3D). The technique is particularly useful for the characterization of the electrical properties of single cells. By measuring the ROT rate at different frequencies ROT spectra can be obtained, from which the electrical properties of the cell can be determined.

Traveling Wave Dielectrophoresis, Traveling Wave Devices

Traveling wave dielectrophoresis (TWD) refers to the lateral movement of particles in phase-shifted electric fields.^[35] Such phase-shifted electric fields can be produced by applying 90° phase-shifted electric fields to electrodes which are placed next to each other. By extending the electrode arrays, repeating the application of the 90° phase-shifted electric fields to every four consecutive electrodes (see Fig. 3E), railway track-like arrays can be produced, which can be used to

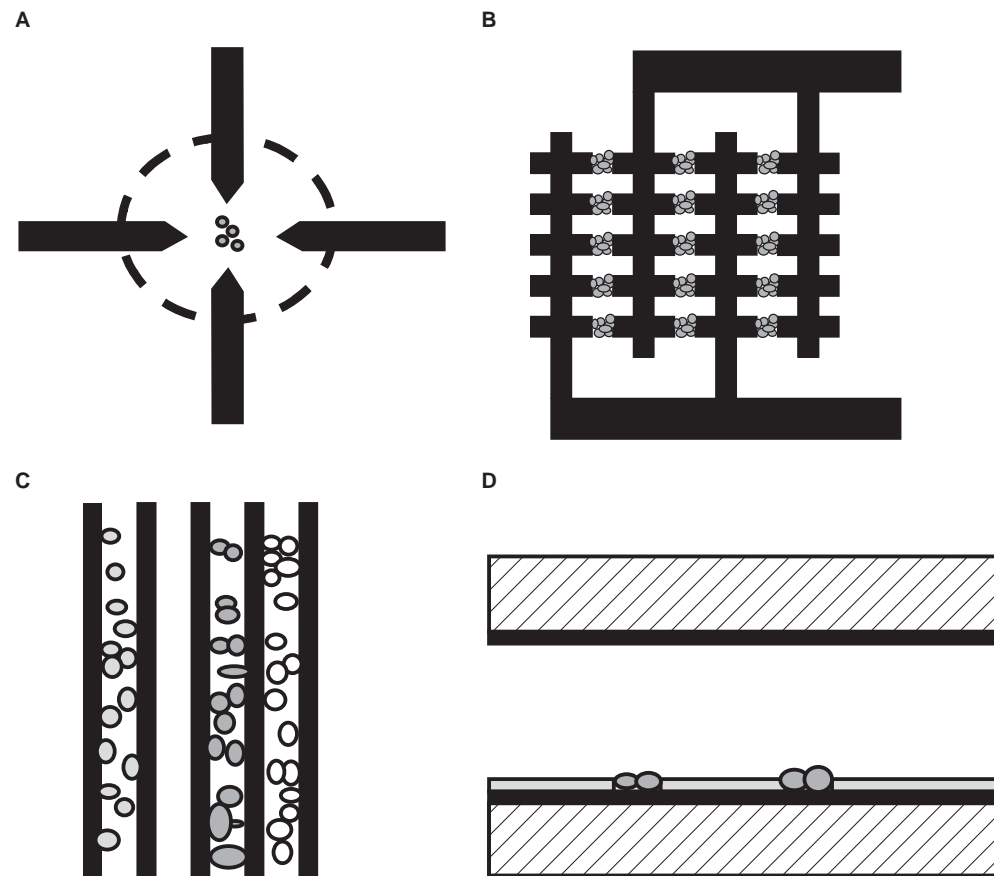


Fig. 2 Some examples of electrodes used for the patterning of cells. (A) Top view of a trap based on negative dielectrophoresis as used by Heida and coworkers. (From Refs. [20,21].) The system is able to trap cells by negative dielectrophoresis directly in growth medium. To prevent excessive heating, the electrodes outside the center are covered by an insulating layer. (B) Top view of an interdigitated oppositely castellated electrode system used to form cell aggregates by positive dielectrophoresis. (From Refs. [26–30].) Cells are suspended in low-conductivity buffer. (C) Top view of a parallel electrode system. (From Refs. [26–30].) Each electrode can be addressed individually, attracting cells suspended in low-conductivity buffers by positive dielectrophoresis. (D) Side view of a section through a parallel plate chamber as used by Albrecht et al. (From Ref. [25].) The top chamber is formed by an ITO covered slide, the bottom layer by an ITO covered slide, on which an electrically insulating layer of photoresist is to form local high-field regions to which cells suspended in low-conductivity buffers can be attracted by positive dielectrophoresis. Other electrode arrangements can be found in the literature. (From Refs. [19,22,23].)

characterize and separate particles, or to bring them together. Like ROT, it is dependent on the imaginary part of the Clausius–Mossotti factor.^[36]

APPLICATIONS OF AC ELECTROKINETICS IN TISSUE ENGINEERING

Alternating current electrokinetics has found the following main applications:

Characterization of Cell Properties

Typically, ROT and DEP are used to determine a cell's electrical properties. The dielectrophoretic force

and/or ROT rate is determined as a function of frequency, and a model (such as the multishell model) used to describe the relation between the properties of a cell's substructures and its overall particle electrical properties.^[13] The resulting model is fitted to the data in the dielectrophoretic or ROT spectrum to obtain the values of the different parameters. For mammalian cells, which have no cell wall, the most important substructures are the cell membrane and the cytoplasm; the nucleus and other membranous structures inside the cells only make a contribution to the cell dielectric properties if they occupy a large part of the cytoplasm. The integrity of the cell membrane is very important. The cell membrane normally has a very low conductivity and dielectric constant (permittivity) compared to the cytoplasm and suspending medium.

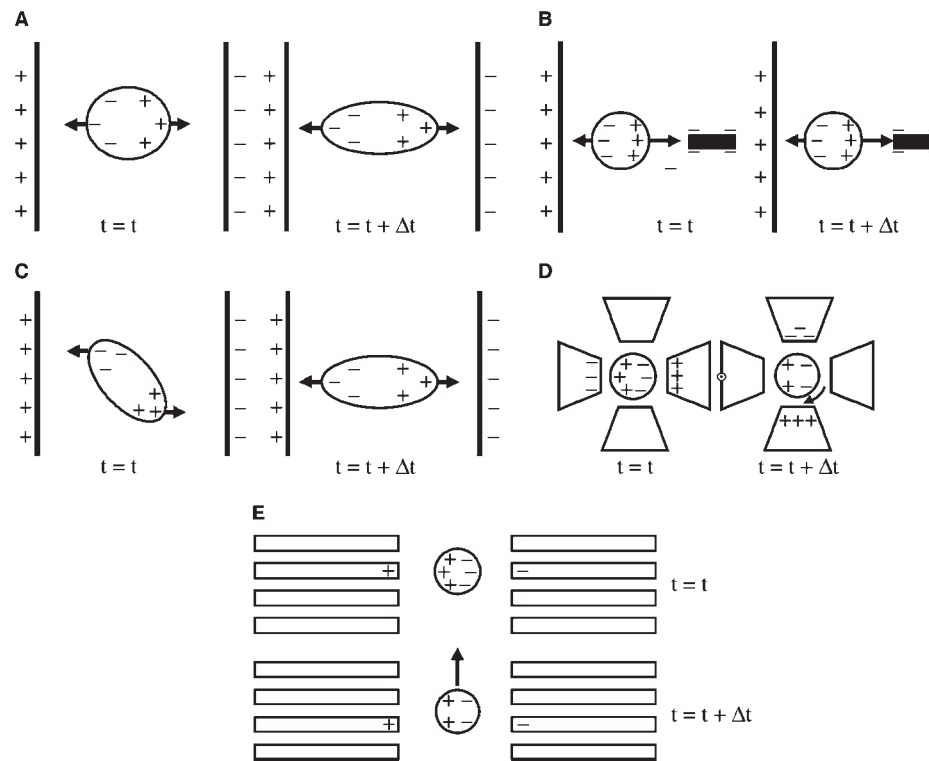


Fig. 3 Outline of the main AC electrokinetic effects. (A) Electrostriction, the distortion of particles by applied electric fields. (B) Dielectrophoresis, the movement of particle in nonuniform electric fields. (C) Electro-orientation, the orientation of elongated particles in electric fields. (D) Electrorotation, the rotation of particles in electric fields. At $t = t$ a dipole is induced in the particle. Switching the electric field at $t = t + \Delta t$ is instantaneous. The dipole induced in the particle does not disappear as quickly as the electric field is switched, and the interaction between the applied field and the remaining charges on the particle cause the particle to rotate. (E) Traveling wave dielectrophoresis. As in electrorotation, the electric field is switched from one set of electrodes to another. The interaction between the applied electric field and the dipole remaining in the particle causes it to move.

Any changes in its properties, for example, permeability, can be readily detected.^[13] The membrane capacitance is also strongly dependent on the presence of blebs and other forms of folding of the membrane.^[34] Apoptosis and necrosis are both associated with large changes in the internal conductivity and membrane properties, which lead to large differences in their AC electrokinetic behavior.^[13] Changes detected in the dielectric properties of cell during the transition from the noncancerous to the cancerous state have been associated with changes in membrane properties.^[14]

Separation of Cells

Differences in the electrical properties between cells, caused by differences in their structure and composition, lead to differences in dielectrophoretic behavior. These are sometimes large enough for their separation. Typically, in dielectrophoretic separations large arrays of electrodes of the interdigitated or interdigitated castellated type are used, covering the

bottom of a long, narrow chamber. Cells are either attracted to the electrodes by positive DEP, or pushed away to low-field regions between or above the electrodes by negative DEP. Selected cell populations are then moved out of the chamber. In dielectrophoretic field-flow fractionation the parabolic velocity profile of a fluid flowing through the chamber is used to advantage to improve the separation process and selectively move cells with different dielectrophoretic behaviors through the chamber at different speeds.^[14] In TWD no flow is needed, but the cells can be made to move to different parts of the separation chamber using the electric fields only. Typical examples of successful separations are those of dead and live cells, stem cells from nonstem cells, and cancer cells from noncancer cells.^[14]

Sorting of Cells

Dielectrophoretic forces produced by microelectrodes in a flow of suspended particles allow particles to be diverted into different streamlines, and thereby led

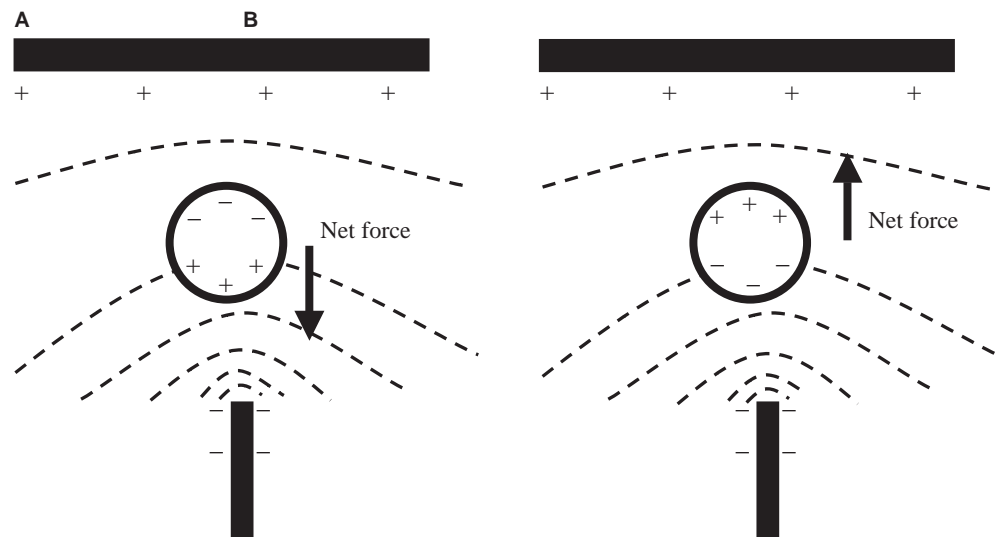


Fig. 4 Outline of positive and negative dielectrophoresis. The indented lines indicate the equipotential lines. Where the lines are closest, the electric field is highest. Positive dielectrophoresis (A) occurs when the particle is more polarizable than the surrounding medium. The dipole formed as a result of the polarization of the particle and the medium is aligned with the electric field, and the particle moves to high-field regions. Negative dielectrophoresis (B) occurs when the particle is less polarizable than the medium. The dipole is aligned opposite to the electric field, and the particle moves to the low-field regions. Note that for both positive and negative dielectrophoresis, when the electric field direction reverses, the induced dipole also reverses, and the direction of the net force stays the same. Thus, dielectrophoresis can occur in DC as well as AC electric fields.

into different channels. This approach could lead to miniaturized cell sorters.^[18] Microfluidic cell sorters, especially when combined with high-resolution analysis systems, could be particularly useful when handling relatively rare or valuable cells such as stem cells.^[18]

Trapping of Cells/Production of Structured Biomaterials

Alternating current electrokinetics can be used to trap (single) cells, to pattern cells on a surface, and to form aggregates of cells. The potential to use AC electrokinetics to produce tissue-like materials was first demonstrated by Pohl in 1972, who was able to trap and orient bacterial cells at an electrode surface, and then immobilize them using a concentrated aqueous solution of Pluronic.^[12] Pluronic is a block-copolymer of polyethylene and polypropylene oxide, which is fluid at 0–5°C, but sets at 35–40°C. Later, the introduction of microelectrodes led to an interest in trapping of single cells.^[17] Progress in the use of AC electrokinetics for patterning has been relatively slow, but in recent years the use of AC electrokinetics for the patterning of cells and the creation of aggregates and tissue-like materials has been receiving increasing attention.^[9,18–31]

Entrapment of the cells can be achieved with either positive or negative DEP; both methods have their own advantages and disadvantages. When positive

DEP is used, the cells are suspended in low salt buffers and attracted to high-field regions at electrode edges, typically at a frequency of 1 MHz. With negative DEP, cells are pushed into low-field regions between the electrodes. Trapping with positive DEP tends to be more rapid than trapping with negative DEP.^[24] In a typical experiment one starts with suspended cells, and uses DEP to guide cells to and trap them in specific areas on a surface. Although it is possible to continuously trap cells using an electric field only, it is often more convenient (and when using positive DEP, essential) to immobilize the cells once they have reached their destination.^[17,20,21] Sometimes cells can be made to adhere to a surface.^[19] In other cases, cells have to be immobilized in other ways. Embedding in hydrogels is used, but cross-linking is also possible.^[24,25,30] The use of hydrogels and cross-linking has the advantage that 3D physical microenvironments are created, which mimic the cell's natural microenvironment more closely than 2D systems.

The potential to use AC electrokinetics to create artificial neural networks has received a lot of attention. Heida and coworkers and other researchers have used negative DEP to create artificial neural networks by directing neural cells toward the center of a quadrupolar electrode system using negative DEP.^[20–22] It was possible to do the experiments directly in growth medium by insulating the electrode areas that were not used directly for cell manipulation, thus reducing heat production by the electric field.^[20,21] Other groups have

sought to use AC electrokinetics to pattern cells and recreate the microenvironment found in actual tissues. Matsue et al.^[31] used negative DEP to pattern cells, and subsequently immobilized them in a gel. Markx et al. used positive DEP to create multilayered aggregates of cells.^[26,27,30] By addressing individual interdigitated electrodes in an electrode array the group was able to put different cell types next to each other.^[26] Initially, the techniques were only used for microorganisms, but lately they have been adapted for use with mammalian cells.^[26,27,29,30] Cell immobilization was achieved with hydrogels or flocculating agents. Albrecht et al. have described a similar approach, using photopolymerizable PEG to immobilize cells.^[25] Instead of addressing individual electrodes, the approach taken was to pattern the cell with DEP, and then to immobilize cells in specific regions in a gel by exposing a photopolymerizable prepolymer solution through a mask, and subsequently wash away the cells and precursor in the unexposed regions. Finally, trapping of cells by DEP allows the production of large arrays of cell aggregates or even single cells, with an aggregate or cell in a specified location.^[19,25–27,29] This has many potential applications in biosensors and high-throughput analytical devices.

CONCLUSIONS

The application of AC electrokinetics in tissue engineering is still in its infancy and has to date been limited, but is likely to expand significantly in the near future. Tissues are made of specific combinations of specialized cells. Alternating current electrokinetics is unique in that it can be used for the separation of the specialized cells of which tissues are made, their characterization, as well as the actual formation of tissue-like materials. In our opinion, the layering of multiple cell types in direct contact with each other is a particularly important emerging area in AC electrokinetics. Although much early work was done with more robust microorganisms, the use of AC electrokinetics in building tissue equivalents from suspended mammalian cells is increasingly becoming a reality. The application of AC electrokinetics to building 3-D mammalian systems is particularly important because it is clear that cells placed in 3-D culture systems behave significantly differently from cells in 2-D cultures.^[37] Cells patterned using AC electrokinetics are potentially, particularly, useful for defining molecular interactions between different cell types. Aggregates of bacterial cells with defined internal architectures, made using DEP, have already been shown to be capable of responding to molecular signals and the exchange of nutrients.^[29,38] This approach may also be useful for studying the role of soluble mediators in mammalian

cells systems. Layering cells in adjacent positions could also enable the study of direct cell–cell interactions such as those mediated by adhesion molecules or ligand-receptor pairs.

Finally, while AC electrokinetics in itself is an excellent technique for manipulating cells, its usefulness could be even further enhanced by combining electrokinetic approaches with other physical manipulation techniques such as laser tweezers. For example, AC electrokinetic techniques are very suitable for the creation of cell arrays, while laser tweezers are excellent at maneuvering individual cells into specific positions. Combining the strengths of AC electrokinetics and laser tweezers in this way has many potential applications, in particular in the area of stem cell biology and the study of the stem cell microenvironment.

ACKNOWLEDGMENTS

We wish to thank the Wellcome Trust, Leukaemia Research Fund, Searle, British Heart Foundation, BBRSC and EPSRC for funding.

ARTICLE OF FURTHER INTEREST

Tissue Engineering

REFERENCES

1. Hughes, M.P. *Nanoelectromechanics in Engineering and Biology*; CRC Press: Boca Raton, 2003.
2. Markx, G.H.; Davey, C.L. The dielectric properties of biological cells at radio-frequencies: applications in biotechnology. *Enzyme Microbiol. Technol.* **1999**, *25* (3–5), 161–171.
3. Jones, T.B. *Electromechanics of Particles*; Cambridge University Press: Cambridge, New York, 1995.
4. Pethig, R.; Kell, D.B. The passive electrical properties of biological systems: their significance in physiology, biophysics and biotechnology. *Phys. Med. Biol.* **1987**, *32* (8), 933–970.
5. Ozkan, M.; Pisanic, T.; Scheel, J.; Barlow, C.; Esener, S.; Bhatia, S.N. Electro-optical platform for the manipulation of life cells. *Langmuir* **2003**, *19* (5), 1532–1538.
6. Bazou, D.; Kuznetsova, L.A.; Coakley, W.T. Physical environment of 2-D animal cell aggregates formed in a short pathlength ultrasound standing wave trap. *Ultrasound Med. Biol.* **2005**, *31* (3), 423–430.
7. Grier, D.G. A revolution in optical manipulation. *Nature* **2003**, *424* (6950), 810–816.
8. Pohl, H.A. *Dielectrophoresis*; Cambridge University Press: Cambridge, New York, 1978.
9. Pethig, R.; Huang, Y.; Wang, X.B.; Burt, J.P.H. Positive and negative dielectrophoretic collection of

- colloidal particles using interdigitated castellated electrodes. *J. Phys. D: Appl. Phys.* **1992**, *25* (5), 881–888.
10. Sukhorukov, V.L.; Mussauer, H.; Zimmermann, U. The effect of electrical deformation forces on the electroporabilization of erythrocyte membranes in low- and high-conductivity media. *J. Membrane Biol.* **1998**, *163* (3), 235–245.
 11. Miller, R.D.; Jones, T.B. Electro-orientation of ellipsoidal erythrocytes. *Biophys. J.* **1993**, *64* (5), 1588–1595.
 12. Pohl, H.A. Electrical forming of masses of living cells. *J. Colloid Interface Sci.* **1972**, *39* (2), 437–438.
 13. Huang, Y.; Hölzel, R.; Pethig, R.; Wang, X.B. Differences in the AC electrokinetics of viable and non-viable yeast cells determined through combined dielectrophoresis and electrorotation studies. *Phys. Med. Biol.* **1992**, *37* (7), 1499–1517.
 14. Gascoyne, P.R.C.; Vykoukal, J. Particle separation by dielectrophoresis. *Electrophoresis* **2002**, *23* (13), 1973–1983.
 15. Docoslis, A.; Kalogerakis, N.; Behie, L.A. Dielectrophoretic forces can be safely used to retain viable cells in perfusion cultures of animal cells. *Cytotechnology* **1999**, *30* (1–3), 133–142.
 16. Wang, X.B.; Huang, Y.; Burt, J.P.H.; Markx, G.H.; Pethig, R. Selective dielectrophoretic confinement of bioparticles in potential energy wells. *J. Phys. D: Appl. Phys.* **1993**, *26* (8), 1278–1285.
 17. Fuhr, G.; Glasser, H.; Müller, T.; Schnell, T. Cell manipulation and cultivation under a.c. electric field influence in highly conductive culture media. *Biochim. Biophys. Acta* **1994**, *1201* (3), 353–360.
 18. Müller, T.; Pfennig, A.; Klein, P.; Gradl G. Jäger, M.; Schnelle, T. The potential of dielectrophoresis for single-cell experiments. *IEEE Eng. Med. Biol. Mag.* **2003**, *22* (6), 51–61.
 19. Gray, D.S.; Tan, J.L.; Voldman, J.; Chen, C.S. Dielectrophoretic registration of living cells to a micro-electrode array. *Biosens. Bioelectron.* **2004**, *19* (12), 1765–1774.
 20. Heida, T.; Rutten, W.L.C.; Marani, E. Dielectrophoretic trapping of dissociated fetal cortical rat neurons. *IEEE Trans. Biomed. Eng.* **2001**, *48* (8), 921–930.
 21. Heida, T.; Vulto, P.; Rutten, W.L.; Marani, E. Viability of dielectrophoretically trapped neural cortical cells in culture. *J. Neurosci. Methods* **2001**, *110* (1–2), 37–44.
 22. Prasad, S.; Yang, M.; Zhang, X.; Ozkan, C.S.; Ozkan, M. Electric field assisted patterning of neuronal networks for the study of brain functions. *Biomed. Microdevices* **2003**, *5* (2), 125–137.
 23. Yu, Z.; Xiang, G.; Pan, L.; Huang, L.; Yu, Z.; Xing, W.; Cheng, J. Negative dielectrophoretic force assisted construction of ordered neuronal networks on cell positioning bioelectronic chips. *Biomed. Microdevices* **2004**, *6* (4), 311–324.
 24. Albrecht, D.R.; Sah, R.L.; Bhatia, S.N. Geometric and material determinants of patterning efficiency by dielectrophoresis. *Biophys. J.* **2004**, *87* (4), 2131–2147.
 25. Albrecht, D.R.; Tsang, V.L.; Sah, R.L.; Bhatia, S.N. Photo- and electropatterning of hydrogel-encapsulated living cell arrays. *Lab Chip* **2005**, *5* (1), 111–118.
 26. Alp, B.; Stephens, G.M.; Markx, G.H. Formation of artificial, structured microbial consortia (ASMC) by dielectrophoresis. *Enzyme Microbiol. Technol.* **2002**, *31* (1–2), 35–43.
 27. Alp, B.; Andrews, J.S.; Mason, V.P.; Wolowacz, R.; Markx, G.H. Building structured biomaterials using AC electrokinetics. *IEEE Eng. Med. Biol. Mag.* **2003**, *22* (6), 91–97.
 28. Markx, G.H.; Andrews, J.S.; Mason, V.P. Towards microbial tissue engineering? *Trends Biotechnol.* **2004**, *22* (8), 417–422.
 29. Mason, V.P.; Markx, G.H.; Thompson, I.P.; Andrews, J.S.; Manefield, M. Colonial architecture in mixed species assemblages affects AHL mediated gene expression. *FEMS Microbiol. Lett.* **2005**, *244* (1), 121–127.
 30. Verduzco-Luque, C.E.; Alp, B.; Stephens, G.M.; Markx, G.H. Construction of biofilms with defined internal architecture using dielectrophoresis and flocculation. *Biotechnol. Bioeng.* **2003**, *83* (1), 39–44.
 31. Matsue, T.; Matsumoto, N.; Uchida, I. Rapid micro-patterning of living cells by repulsive dielectrophoretic force. *Electrochim. Acta* **1997**, *42* (20–22), 3251–3256.
 32. Neil, G.A.; Zimmermann, U. Electrofusion. *Methods Enzymol.* **1993**, *221*, 171–196.
 33. Wang, X.; Yang, J.; Gascoyne, P.R.C. Role of peroxide in AC electrical field exposure effects on Friend murine erythroleukemia cells during dielectrophoretic manipulations. *Biochim. Biophys. Acta* **1999**, *1426* (1), 53–68.
 34. Wang, X.B.; Huang, Y.; Gascoyne, P.R.C.; Becker, F.F.; Hölzel, R.; Pethig, R. Changes in Friend murine erythroleukaemia cell membranes during induced differentiation determined by electrorotation. *Biochim. Biophys. Acta* **1994**, *1193* (2), 330–344.
 35. Pethig, R.; Talary, M.S.; Lees, R.S. Enhanced travelling-wave dielectrophoresis with signal superposition. *IEEE Eng. Med. Biol. Mag.* **2003**, *22* (6), 43–50.
 36. Huang, Y.; Tame, J.A.; Pethig, R. Electrokinetic behaviour of colloidal particles in travelling electric fields: studies using yeast cells. *J. Phys. D: Appl. Phys.* **1993**, *26* (9), 1528–1535.
 37. Schmeichel, K.L.; Bissell, M.J. Modeling tissue-specific signaling and organ function in three dimensions. *J. Cell Sci.* **2003**, *116* (12), 2377–2388.
 38. Andrews, J.A.; Mason, V.P.; Thompson, I.P.; Stephens, G.M.; Markx, G.H. Construction of artificially structured microbial consortia (ASMC) using dielectrophoresis: examining bacterial interactions via metabolic intermediates within environmental biofilms *J. Microbiol. Methods* **2005**, *in press*.

FURTHER READING

- Chethana, S.; Patil, G.; Raghavarao, K.S.M.S. Magnetic and electrophoretic cell separation. 10.1081/E-EBBE-120020410.
- Madou, M. Microelectromechanical system (MEMS) manufacturing. 10.1081/E-EBBE-120013935.
- Tan, W.; Desai, T. Tissue Engineering, Microscale. 10.1081/E-EBBE-120012936.



Tissue Engineering of Bladder

Anthony Atala

Wake Forest Institute for Regenerative Medicine, Winston-Salem, North Carolina, U.S.A.

INTRODUCTION

The bladder is exposed to a variety of possible injuries from the time the fetus develops, and may require replacement or repair. Currently, gastrointestinal segments are commonly used as tissues for bladder replacement or repair. However, gastrointestinal tissues are designed to absorb specific solutes, whereas bladder tissue is designed for the excretion of solutes. When gastrointestinal tissue is in contact with the urinary tract, multiple complications may ensue, such as infection, metabolic disturbances, urolithiasis, perforation, increased mucous production, and malignancy.^[1-4] Due to the problems encountered with the use of gastrointestinal segments, numerous investigators have attempted alternative materials for tissue regeneration and engineering for bladder replacement or repair.

MATRICES FOR BLADDER REGENERATION

Over the last few decades, several bladder wall substitutes have been attempted with both synthetic and organic materials. The first application of a free tissue graft for bladder replacement was reported by Neuhoff in 1917, when fascia was used to augment bladders in dogs.^[5] Since that first report, multiple other free graft materials have been used experimentally and clinically, including bladder allografts, pericardium, dura, and placenta.^[6-9] Synthetic materials, which have been tried previously in experimental and clinical settings, include polyvinyl sponge, tetrafluoroethylene (Teflon), collagen matrices, vicryl matrices, and silicone.^[10-13] Most of these attempts have usually failed due to either mechanical, structural, functional, or biocompatibility problems. Usually, permanent synthetic materials succumb to mechanical failure, and urinary stone formation and degradable materials lead to fibroblast deposition, scarring, graft contracture, and a reduced reservoir volume over time.

There has been a resurgence in the use of various collagen-based matrices for tissue regeneration. Fresh placental membranes have been used experimentally in dogs.^[9] Biodegradable pericardial implants were used for bladder augmentation in dogs.^[9] Functionally, the implanted bladders showed adequate capacity for up to 36 months, but grossly, they were noted to undergo

graft shrinkage. Histologically, the epithelial layer was present, but the muscular layer was absent. Bladder grafts, initially used experimentally in 1961, have been used recently by various investigators.^[6,14-16] In multiple studies using different materials as the acellular bladder graft for cystoplasty, the urothelial layer was able to regenerate normally, but the muscle layer, although present, was not fully developed.^[14-17]

Small-intestinal submucosa (SIS), a biodegradable, acellular, xenogeneic collagen-based tissue-matrix graft, has been shown to promote regeneration of a variety of host tissues, including blood vessels and ligaments. The matrix is derived from pig small intestine in which the mucosa is mechanically removed from the inner surface and the serosa and muscular layer are removed from the outer surface.

Animal studies have shown that the SIS matrix used for bladder augmentation is able to regenerate *in vivo*. Augmentation cystoplasty with SIS was performed in eight dogs, which were followed for 15 months.^[17] Preoperative mean bladder capacity was 51 ml compared to a postoperative mean capacity of 55 ml. The mean maximal voiding pressure was 52 cm H₂O preoperatively and 45 cm postoperatively. Histologically, the transitional layer was the same as that of the native bladder tissue, but as with other collagen matrices used experimentally, the muscle layer was not fully developed. A large amount of collagen was interspersed between a smaller number of muscle bundles. A computer-assisted image analysis demonstrated a decreased muscle-to-collagen ratio with a loss of the normal architecture in the SIS regenerated bladders.

In-vitro contractility studies performed on the SIS regenerated dog bladders showed a decrease in maximal contractile response by 50% from those of normal bladder tissues.^[17]

TISSUE ENGINEERING OF BLADDERS USING CELL TRANSPLANTATION

Engineering tissue using selective cell transplantation may provide a means to create functional new bladder segments.^[18] In tissue engineering, donor tissue is dissociated into individual cells, which are either implanted directly into the host or expanded in culture, attached to a support matrix, and reimplanted after

expansion. The implanted tissue can be either heterologous, allogeneic, or autologous. Ideally, this approach might allow lost tissue function to be restored or replaced in toto and with limited complications.^[18] The approach, which we have followed to bioengineer urologic tissues, involves the eventual use of autologous cells, thus avoiding rejection, wherein a biopsy of tissue is obtained from the host and the cells are dissociated and expanded in vitro, reattached to a matrix, and implanted into the same host.^[15,18–37]

The success of using cell transplantation strategies for bladder reconstruction depends on the ability to use donor tissue efficiently and to provide the right conditions for long-term survival, differentiation, and growth. We have achieved an approach to bladder tissue regeneration by patching isolated cells to a support structure, which would have suitable surface chemistry for guiding the reorganization and growth of the cells. The supporting matrix is composed of biodegradable artificial or natural polymers which can allow cell survival by diffusion of nutrients across short distances once the cell-support matrix is implanted. The cell-support matrix becomes vascularized in concert with expansion of the cell mass following implantation.

Cell Expansion

One of the initial limitations of applying tissue engineering techniques to the bladder was the previously encountered inherent difficulty of growing genitourinary-associated cells in large quantities. In the past, it was believed that urothelial cells had a natural senescence, which was hard to overcome. Normal urothelial cells could be grown in the laboratory setting, but with limited expansion. A system of urothelial cell harvest was created in our laboratory that does not use any enzymes or serum and has a large expansion potential.^[26] Urothelial cells are expanded until the desired cell numbers are obtained. Using these methods of cell culture, it is possible to expand a urothelial strain from a single specimen that initially covers a surface area of 1 cm² to one covering a surface area of 4202 m² (the equivalent area of one football field) within 8 weeks.^[26] The muscle cells are processed by the tissue explant technique and are expanded in a similar manner. We have shown that normal human bladder epithelial and muscle cells can be efficiently harvested from surgical material, extensively expanded in culture, and their differentiation characteristics, growth requirements, and other biological properties studied.^[15,26,27,29–40]

Cell Delivery Vehicles

As described initially by Judah Folkman in 1973, cells or tissue cannot be implanted in volumes greater than

3 mm³.^[41] Nutrition and gas exchange are limited by this maximum diffusion distance. If cells were implanted in volumes greater than 3 mm³, only the cells on the surface would survive, and the central cell core would necrose due to a lack of vascularity.^[41] The cell delivery vehicles we have used are designed for cell attachment and capillary infiltration to the interstitial spaces after implantation in vivo.^[15,24,25,27,29–32,34–37] In this manner, large numbers of cells can be implanted with maximal survival.

Both synthetic and natural biodegradable materials have been used as cell delivery vehicles.^[15,24,25,27,29–32,34–37] We have used resorbable biodegradable polymers, which are preferable because permanent polymers, such as silicone and Teflon, carry the risks of infection, calcification, and unfavorable connective tissue responses.^[18] The cell delivery scaffolds used experimentally in our laboratory include poly(glycolic acid) (PGA) polymer scaffolds alone and with copolymers of poly(L-lactic acid) and poly(DL-lactide-co-glycolide).^[24,25,36] These polymers have many desirable features: They are biocompatible and processible. Degradation occurs by hydrolysis, and the time sequence can be varied from weeks to over a year by manipulating the ratio of monomers and by varying the processing conditions. We have also used processed collagen derived from allogeneic donor bladder submucosa.^[15] These biodegradable matrices, both natural and artificial, can be readily modified, depending on their intended application, into a variety of shapes and structures, including small-diameter fibers and porous films, to various levels of rigidity or elasticity, and to different time sequences for degradation.

The porosity, pore size distribution, and continuity dictate the interaction of the biomaterials and transplanted cells with the host tissue. The appropriate porosity and branching pattern of the matrix results in the formation of a capillary network in the developing tissue.^[15,24,25,27,29–32,34–37] Vascularization of the engineered tissue is required to meet the metabolic requirements of the tissue and to integrate it with the surrounding host.

Formation of Bladder Tissue Ex Situ

Urothelial and muscle cells can be expanded in vitro, seeded onto the polymer scaffold, and allowed to attach and form sheets of cells. The cell-polymer scaffold can then be implanted in vivo. We have performed a series of in-vivo urologic-associated cell-polymer experiments. Histologic analysis of human urothelial, bladder muscle, and composite urothelial and bladder muscle-polymer scaffolds, implanted in athymic mice and retrieved at different time points, indicated that viable cells were evident in all three

experimental groups.^[25] Implanted cells oriented themselves spatially along the polymer surfaces. The cell populations appeared to expand from one layer to several layers of thickness with progressive cell organization with extended implantation times. Cell-polymer composite implants of urothelial and muscle cells, retrieved at extended times (50 days), showed extensive formation of multilayered sheet-like structures and well-defined muscle layers. Polymers seeded with cells and manipulated into a tubular configuration showed layers of muscle cells lining the multilayered epithelial sheets. Cellular debris appeared reproducibly in the luminal spaces, suggesting that epithelial cells lining the lumina are sloughed into the luminal space. Cell polymers implanted with human bladder muscle cells alone showed almost complete replacement of the polymer with sheets of smooth muscle at 50 days. This experiment demonstrated, for the first time, that composite tissue engineered structures could be created *de novo*. Prior to this study, only single cell-type tissue-engineered structures had been created.

Formation of Bladder Tissue In Situ

In order to determine the effects of implanting engineered tissues in continuity with the urinary tract, an animal model of bladder augmentation was utilized.^[15] Partial cystectomies, which involved removing approximately 50% of the native bladders, were performed in 10 beagles. In 5, the retrieved bladder tissue was microdissected and the mucosal and muscular layers separated. The bladder urothelial and muscle cells were cultured using the techniques described above. Urothelial and smooth muscle cells were harvested and expanded separately. A collagen-based matrix, derived from allogenic bladder submucosa, was used for cell delivery. This material was chosen for these experiments due to its native elasticity. Within six weeks, the expanded urothelial cells were collected as a pellet. The cells were seeded on the luminal surface of the allogenic bladder submucosa and incubated in serum-free keratinocyte growth medium for five days. Muscle cells were seeded on the opposite side of the bladder submucosa and subsequently placed in DMEM supplemented with 10% fetal calf serum for an additional five days. The seeding density on the allogenic bladder submucosa was approximately 1×10^7 cells/cm².

Preoperative fluoroscopic cystography and urodynamic studies were performed in all animals. Augmentation cystoplasty was performed with the matrix with cells in one group, and with the matrix without cells in the second group. The augmented bladders were covered with omentum in order to facilitate angiogenesis to the implant. Cystostomy catheters were used for urinary diversion for 10 to 14 days. Urodynamic

studies and fluoroscopic cystography were performed at one, two, and three months postoperatively. Augmented bladders were retrieved two ($n=6$) and three ($n=4$) months after surgery and examined grossly, histologically, and immunocytochemically.

Bladders augmented with the matrix seeded with cells showed a 99% increase in capacity compared to bladders augmented with the cell-free matrix, which showed only a 30% increase in capacity (Fig. 1). Functionally, all animals showed normal bladder compliance as evidenced by urodynamic studies; however, the remaining native bladder tissue may have accounted for these results. Histologically, the retrieved engineered bladders contained a cellular organization consisting of a urothelial-lined lumen surrounded by submucosal tissue and smooth muscle. However, the muscular layer was markedly more prominent in the cell-reconstituted scaffold.^[15]

Most of the free grafts (without cells) utilized for bladder replacement in the past have been able to show adequate histology in terms of a well-developed urothelial layer; however, they have been associated with an abnormal muscular layer which varies in terms of its full development.^[19,20] It has been well established for decades that the bladder is able to regenerate generously over free grafts. Urothelium is associated with a high reparative capacity.^[42] Bladder muscle tissue is less likely to regenerate in a normal fashion. Both urothelial and muscle in growth are believed to be initiated from the edges of the normal bladder toward the region of the free graft.^[43,44] Usually, however, contracture or resorption of the graft has been evident. The inflammatory response toward the matrix may contribute to the resorption of the free graft.

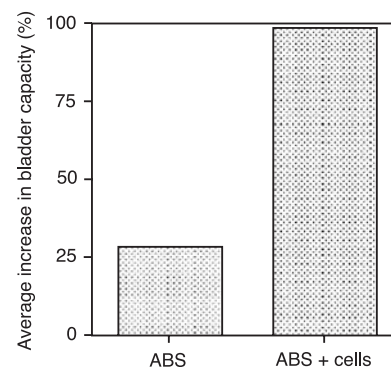


Fig. 1 Bladders augmented with a collagen matrix derived from bladder submucosa seeded with urothelial and smooth muscle cells (ABS + cells) showed a 100% increase in capacity compared to bladders augmented with the cell-free allogenic bladder submucosa (ABS), which showed only a 30% increase in capacity within three months after implantation. (Reproduced with permission, from Ref. [15].)

We hypothesized that building the three-dimensional structure construct *in vitro*, prior to implantation, would facilitate the eventual terminal differentiation of the cells after implantation *in vivo*, and would minimize the inflammatory response toward the matrix, thus avoiding graft contracture and shrinkage. This study demonstrated that there was a major difference evident between matrices used with autologous cells (tissue engineered) and matrices used without cells.^[15] Matrices implanted with cells for bladder augmentation retained most of their implanted diameter, as opposed to matrices implanted without cells for bladder augmentation, wherein graft contraction and shrinkage occurred. The histomorphology demonstrated a marked paucity of muscle cells and a more aggressive inflammatory reaction in the matrices implanted without cells. Of interest is that the urothelial cell layers appeared normal, even though the underlying matrix was significantly inflamed. We further hypothesized, therefore, that having an adequate urothelial layer from the outset would limit the amount of urine contact with the matrix, and would therefore decrease the inflammatory response, and that the muscle cells were also necessary for bioengineering, since native muscle cells are less likely to regenerate over the free grafts. Further studies performed in our laboratory confirmed this hypothesis.^[36] Thus, it appears that the presence of both urothelial and muscle cells on the matrices we used for bladder replacement are important for successful tissue bioengineering.

Bladder Replacement Using Tissue Engineering

The results of our initial studies showed that the creation of artificial bladders may be achieved *in vivo*; however, it could not be determined whether the functional parameters noted were due to the augmented segment or the intact native bladder tissue. To better address the functional parameters of tissue-engineered bladders, an animal model was designed which required a subtotal cystectomy with subsequent replacement with a tissue-engineered organ.^[36]

A total of 14 beagle dogs underwent a trigone-sparing cystectomy. The animals were randomly assigned to one of three groups. Group A ($n=2$) underwent closure of the trigone without a reconstructive procedure. Group B ($n=6$) underwent reconstruction with a cell-free bladder-shaped biodegradable polymer. Group C ($n=6$) underwent reconstruction using a bladder-shaped biodegradable polymer that delivered autologous urothelial cells and smooth muscle cells. The cell populations had been separately expanded from a previously harvested autologous bladder biopsy. Preoperative and

postoperative urodynamic and radiographic studies were performed serially. Animals were sacrificed at 1, 2, 3, 4, 6, and 11 months postoperatively. Gross, histological, and immunocytochemical analyses were performed.^[36]

The cystectomy-only controls and polymer-only grafts maintained average capacities of 22% and 46% of preoperative values, respectively. An average bladder capacity of 95% of the original pre-cystectomy volume was achieved in the tissue-engineered bladder replacements. These findings were confirmed radiographically (Fig. 2). The subtotal cystectomy reservoirs, which were not reconstructed, and polymer-only reconstructed bladders showed a marked decrease in bladder compliance (10% and 42%). The compliance of the tissue-engineered bladders showed almost no difference from preoperative values that were measured when the native bladder was present (106%). Histologically, the polymer-only bladders presented a pattern of normal urothelial cells with a thickened fibrotic submucosa and a thin layer of muscle fibers. The retrieved tissue-engineered bladders showed a normal cellular organization, consisting of a trilayer of urothelium, submucosa, and muscle (Fig. 3).

Immunocytochemical analyses for desmin, alpha-actin, cytokeratin 7, pancytokeratins AE1/AE3, and uroplakin III confirmed the muscle and urothelial phenotype. S-100 staining indicated the presence of neural structures. The results from this study showed that it is possible to tissue-engineer bladders that are anatomically and functionally normal.^[36]

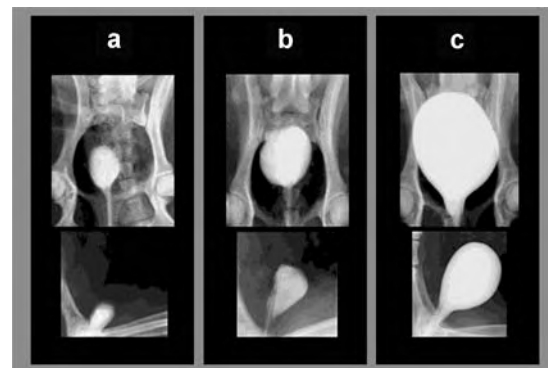


Fig. 2 Radiographic cystograms in beagles 11 months after subtotal cystectomy without reconstruction (a), with reconstruction using a polymer without cells (b), and with reconstruction using a polymer and cell-seeded tissue-engineered organ (c). Trigone-sparing cystectomy organs retained a small-sized reservoir. Tissue-engineered neobladders showed a normal configuration and a larger capacity than the trigones grafted with polymer only. (Reproduced with permission, from Ref. [36].)

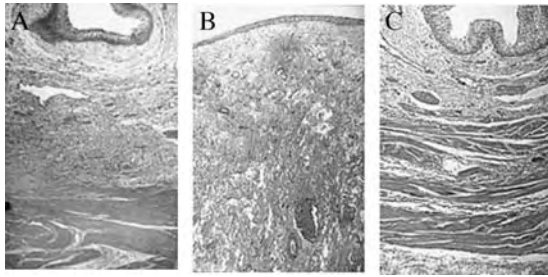


Fig. 3 Hematoxylin and eosin histological results six months after surgery (original magnification: $\times 250$). A: Normal canine bladder. B: The bladder dome of the cell-free polymer-reconstructed bladder consists of normal urothelium over a thickened layer of collagen and fibrotic tissue. Only scarce muscle fibers are apparent. C: The tissue-engineered neo-organ shows a histomorphologically normal appearance. A trilayered architecture consisting of urothelium, submucosa, and smooth muscle is evident. (Reproduced with permission, from Ref. [36].)

Fetal Bladder Tissue Engineering

The prenatal diagnosis of patients with bladder disease is now more prevalent. Prenatal ultrasonography allows for a thorough survey of fetal anatomy.

The evolution of prenatal diagnosis led to the use of intervention before birth to reverse potentially life-threatening processes. However, the concept of prenatal intervention is not limited to this narrow group of indications. A prenatal rather than postnatal diagnosis of exstrophy may be beneficial under certain circumstances. There is now a renewed interest in performing a single-stage reconstruction in some patients with bladder exstrophy. Limiting factors for following a single- or multiple-stage approach include findings of a small, fibrotic bladder patch without either elasticity or contractility, or a hypoplastic bladder.

Today's technological and scientific advances may facilitate the future prenatal management of patients with bladder disease. Having a ready supply of urologic-associated tissue for surgical reconstruction at birth may be advantageous. Theoretically, once the diagnosis of bladder exstrophy is confirmed prenatally, a small bladder and skin biopsy could be obtained via ultrasound guidance. These biopsy materials could then be processed and the different cell types expanded *in vitro*. Using tissue engineering techniques developed at our center and described previously, reconstituted bladder and skin structures *in vitro* could then be readily available at the time of birth for a one-stage reconstruction, allowing for adequate anatomic and functional closure.

Toward this end, we conducted a series of experiments using fetal lambs.^[29,30,32] Bladder exstrophy was

created surgically in ten 90–95-day gestation fetal lambs. The lambs were randomly divided into two groups of five. In group 1, a small fetal bladder specimen was harvested via fetoscopy. The bladder specimen was separated, and muscle and urothelial cells were harvested and expanded separately under sterile conditions in a humidified 5% CO₂ chamber, as previously described. Seven to 10 days prior to delivery, the expanded bladder muscle cells were seeded on one side and the urothelial cells on the opposite side of a 20 cm² biodegradable poly(glycolic acid) polymer scaffold. After delivery, all lambs in group 1 had surgical closure of their bladders using the tissue-engineered bladder tissue. No fetal bladder harvest was performed in the group 2 lambs, and bladder exstrophy closure was performed using only the native bladder. Cystograms were performed three and eight weeks after surgery. The engineered bladders were more compliant ($p=0.01$) and had a higher capacity ($p=0.02$) than the native bladder closure group. Histologic analysis of the engineered tissue showed a normal pattern, indistinguishable from native bladder at two months.^[30] Similar prenatal studies were performed in lambs with engineering of skin for reconstruction at birth.^[32]

The studies described above show that the potential for replicating this technology in humans is possible. Certainly, other tissues such as cartilage, corpora cavernosa, and bowel can be harvested and expanded in the same manner. Similar studies addressing these tissues are now in progress in our laboratory.

In addition to being able to manage the bladder exstrophy complex *in utero* with tissue engineering techniques, one could manage patients after birth in a similar manner, when a prenatal diagnosis is not assured. In these instances, bladder tissue biopsies could be obtained at the time of the initial surgery. Different tissues could be harvested and stored for future reconstruction, if necessary. A tissue bank for exstrophy complex patients could preserve different cell types indefinitely. Gene therapy could also be achieved by transfecting the cells prior to seeding and implantation.^[27]

CONCLUSION

Even though the use of bowel for bladder tissue replacement was first proposed over 100 years ago, it remains the gold standard today, despite its known problems. It is evident that urothelial to urothelial anastomoses are preferable functionally. Multiple new methods and techniques are becoming available for bladder augmentation. Most of the new techniques rely on the use of native urologic tissues, either partially or fully, in an effort to avoid the problems associated with

the use of bowel in the urinary tract. Hopefully, as experience is gained with the recent clinical and experimental approaches, a larger armamentarium of technologies and methods for bladder augmentation will become available.

Recently, functional bladder tissue has been engineered using selective cell transplantation. The approach that has been followed to bioengineer bladder tissue involves the use of autologous cells, thus avoiding rejection, wherein a biopsy of tissue is obtained from the host, the cells dissociated and expanded in vitro, reattached to a matrix, and implanted into the same host. Clinical trials using bioengineered tissues for bladder augmentation are currently being arranged.

ARTICLES OF FURTHER INTEREST

Extracellular Matrix Scaffolds; Stem Cells; Tissue Engineering Scaffolds

REFERENCES

- McDougal, W.S. Metabolic complications of urinary intestinal diversion. *J. Urol.* **1992**, *147*, 1199.
- Atala, A.; Hendren, H.; Bauer, S.; Retik, A. The effect of gastrocystoplasty on bladder function. *J. Urol.* **1993**, *149*, 1099.
- Kaefer, M.; Tobin, M.; Hendren, H.; Bauer, S.B.; Peters, C.A.; Atala, A.; Colodny, A.H.; Mandell, J.; Retik, A.B. Continent urinary diversion: The children's hospital experience. *J. Urol.* **1997**, *157*, 1394.
- Kaefer, M.; Hendren, H.; Bauer, S.; Goldenblatt, P.; Peters, C.; Atala, A.; Retik, A. Reservoir calculi: A comparison of reservoirs constructed from stomach and other enteric segments. *J. Urol.* **1998**, *160*, 2187.
- Neuhof, H. Fascial transplantation into visceral defects: An experimental and clinical study. *Surg. Gynecol. Obstet.* **1917**, *25*, 383.
- Tsuji, I.; Ishida, H.; Fujieda, J. Experimental cystoplasty using preserved bladder graft. *J. Urol.* **1961**, *85*, 42.
- Kambic, H.; Kay, R.; Chen, J.F.; Matsushita, M.; Harasaki, H.; Zilber, S. Biodegradable pericardial implants for bladder augmentation: A 2.5-year study in dogs. *J. Urol.* **1992**, *148*, 539–543.
- Kelami, A.; Ludtke-Handjery, A.; Korb, G.; Roll, J.; Schnell, J.; Danigel, K.H. Alloplastic replacement of the urinary bladder wall with lyophilized human dura. *Eur. Surg. Res.* **1970**, *2*, 195.
- Fishman, I.J.; Flores, F.N.; Scott, B.; Spjut, H.J.; Morrow, B. Use of fresh placental membranes for bladder reconstruction. *J. Urol.* **1987**, *138*, 1291.
- Gleeson, M.J.; Griffith, D.P. The use of aloplastic biomaterials in bladder substitution. *J. Urol.* **1992**, *148*, 1377.
- Bona, A.V.; De Gresti, A. Partial substitution of urinary bladder with Teflon prosthesis. *Minerva Urol.* **1966**, *18*, 43.
- Monsour, M.J.; Mohammed, R.; Gorham, S.D.; French, D.A.; Scott, R. An assessment of a collagen/vicryl composite membrane to repair defects of the urinary bladder in rabbits. *Urol. Res.* **1987**, *15*, 235.
- Rohrmann, D.; Albrecht, D.; Hannappel, J.; Gerlach, R.; Schwarzkopp, G.; Lutzeyer, W. Alloplastic replacement of the urinary bladder. *J. Urol.* **1996**, *156*, 2094.
- Probst, M.; Dahiya, R.; Carrier, S.; Tanagho, E.A. Reproduction of functional smooth muscle tissue and partial bladder replacement. *Br. J. Urol.* **1997**, *79*, 505.
- Yoo, J.J.; Meng, J.; Oberpenning, F.; Atala, A. Bladder augmentation using allogenic bladder submucosa seeded with cells. *Urology* **1998**, *51*, 221.
- Sutherland, R.S.; Baskin, L.S.; Hayward, S.W.; Cunha, G.R. Regeneration of bladder urothelium, smooth muscle, blood vessels, and nerves into an acellular tissue matrix. *J. Urol.* **1996**, *156*, 571–577.
- Kropp, B.P.; Rippy, M.K.; Badylak, S.F.; Adams, M.C.; Keating, M.A.; Rink, R.C.; Thor, K.B. Small intestinal submucosa: Urodynamic and histopathologic evaluation in long term canine bladder augmentations. *J. Urol.* **1996**, *155*, 2098–2104.
- Atala, A.; Lanza, R. Preface. In *Methods of Tissue Engineering*; Atala, A., Lanza, D., Eds.; Academic Press: San Diego, CA, 2001; xli.
- Atala, A. Commentary on the replacement of urologic associated mucosa. *J. Urol.* **1995**, *156*, 338.
- Atala, A. Autologous cell transplantation for urologic reconstruction. *J. Urol.* **1998**, *159*, 2.
- Atala, A.; Cima, L.G.; Kim, W.S.; Page, K.T.; Vacanti, J.P.; Retik, A.B.; Vacanti, C.A. Injectable polymers seeded with chondrocytes as a therapeutic approach. *J. Urol.* **1993**, *150*, 745.
- Atala, A.; Kim, W.S.; Paige, K.T.; Vacanti, C.A.; Retik, A.B. Endoscopic treatment of vesicoureteral reflux with a chondrocyte-alginate suspension. *J. Urol.* **1994**, *152*, 641.
- Kershen, R.T.; Atala, A. Advances in injectable therapies for the treatment of incontinence and vesicoureteral reflux. *Urol. Clin. North Am.* **1999**, *26*, 81–94.
- Atala, A.; Vacanti, J.P.; Peters, C.A.; Mandell, J.; Retik, A.B.; Freeman, M.R. Formation of urothelial structures in vivo from dissociated cells attached to biodegradable polymer scaffolds in vitro. *J. Urol.* **1992**, *48*, 658.
- Atala, A.; Freeman, M.R.; Vacanti, J.P.; Shepard, J.; Retik, A.B. Implantation in vivo and retrieval of artificial structures consisting of rabbit and human urothelium and human bladder muscle. *J. Urol.* **1993**, *150*, 608.
- Cilento, B.G.; Freeman, M.R.; Schneck, F.X.; Retik, A.B.; Atala, A. Phenotypic and cytogenetic characterization of human bladder urothelia expanded in vitro. *J. Urol.* **1994**, *152*, 655.
- Yoo, J.J.; Atala, A. A novel gene delivery system using urothelial tissue engineered neo-organs. *J. Urol.* **1997**, *158*, 1066.

28. Machlouf, M.; Atala, A. Emerging concepts for tissue and organ transplantation. *Graft* **1998**, *1*, 31.
29. Atala, A. Tissue Engineering Techniques for Closure of Bladder Exstrophy: An Experimental Animal Model. In *The Exstrophy-Epispadias Complex*; Gearhart, J., Ed.; Plenum Press: New York, 1998; 63–64.
30. Fauza, D.O.; Fishman, S.; Mehegan, K.; Atala, A. Videofoscopically assisted fetal tissue engineering: Bladder augmentation. *J. Pediatr. Surg.* **1998**, *33*, 7.
31. Yoo, J.J.; Lee, I.; Atala, A. Cartilage rods as a potential material for penile reconstruction. *J. Urol.* **1998**, *160*, 1164.
32. Fauza, D.O.; Fishman, S.; Mehegan, K.; Atala, A. Videofoscopically assisted fetal tissue engineering: Skin replacement. *J. Pediatr. Surg.* **1998**, *33*, 377.
33. Amiel, G.E.; Atala, A. Current and future modalities for functional renal replacement. *Urol. Clin. North Am.* **1999**, *26* (1), 235–246.
34. Yoo, J.J.; Park, H.J.; Lee, I.; Atala, A. Autologous engineered cartilage rods for penile reconstruction. *J. Urol.* **1999**, *160*, 1164.
35. Park, H.J.; Kershen, R.; Yoo, J.J.; Atala, A. Reconstitution of human corporal smooth muscle and endothelial cells in vivo. *J. Urol.* **1999**, *162*, 1106.
36. Oberpenning, F.O.; Meng, J.; Yoo, J.; Atala, A. De novo reconstitution of a functional urinary bladder by tissue engineering. *Nat. Biotechnol.* **1999**, *17*, 2.
37. Atala, A. Future perspectives in reconstructive surgery using tissue engineering. *Urol. Clin. North Am.* **1999**, *26* (1), 157–165.
38. Tobin, M.S.; Freeman, M.R.; Atala, A. Maturational response of normal human urothelial cells in culture is dependent on extracellular matrix and serum additives. *Surg. Forum* **1994**, *45*, 786.
39. Freeman, M.R.; Yoo, J.J.; Raab, G.; Soker, S.; Adam, R.M.; Schneck, F.X.; Renshaw, A.A.; Klagsbrun, M.; Atala, A. Heparin-binding EGF-like growth factor is an autocrine factor for human urothelial cells and is synthesized by epithelial and smooth muscle cells in the human bladder. *J. Clin. Invest.* **1997**, *99* (5), 1028.
40. Nguyen, H.T.; Park, J.M.; Peters, C.A.; Adam, R.A.; Orsola, A.; Atala, A.; Freeman, M.R. Cell-specific activation of the HB-EGF and ErbB1 genes by stretch in primary human bladder cells. *In Vitro Cell. Dev. Biol.* **1999**, *35*, 371–375.
41. Folkman, J.; Hochberg, M.M. Self regulation of growth in three dimensions. *J. Exp. Med.* **1973**, *138*, 745.
42. de Boer, W.I.; Schuller, A.G.; Vermay, M.; van der Kwast, T.H. Expression of growth factors and receptors during specific phases in regenerating urothelium after acute injury in vivo. *Am. J. Path.* **1994**, *145*, 1199.
43. Baker, R.; Kelly, T.; Tehan, T.; Putman, C.; Beaugard, E. Subtotal cystectomy and total bladder regeneration in treatment of bladder cancer. *J. Am. Med. Ass.* **1955**, *168*, 1178.
44. Gorham, S.D.; French, D.A.; Shivas, A.A.; Scott, R. Some observations on the regeneration of smooth muscle in the repaired urinary bladder of the rabbit. *Eur. Urol.* **1989**, *16*, 440.

Tissue Engineering of Blood Vessel

Douglas Hamilton
David Vorp

Departments of Surgery and Bioengineering, Vascular Biomechanics and Vascular Tissue Engineering Laboratories, McGowan Institute for Regenerative Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.

T

INTRODUCTION

A suitable conduit for use in bypassing vessels less than 6 mm in diameter remains elusive; in recent years, attention has turned to the field of tissue engineering for potential new treatments. The focus of this article is to review some of the design criteria that have been followed in attempting to produce the first clinically acceptable tissue-engineered blood vessel (TEBV), and to highlight rate-limiting barriers to achieving this goal.

OVERVIEW

Anatomy of the Vasculature

Within the body, nutrients and gases contained in the blood are transported to the body through the vascular system, which is further subdivided into the pulmonary system (to and from the lungs) and the systemic system (to and from the rest of the body). Each of these subsystems is composed of arteries, which deliver blood from the heart to the various tissues and organs; and veins, which return blood back to the heart. Arteries are characterized as pulsatile, high-pressure, and high-flow, whereas veins are characterized as pulseless, low-pressure, and low-flow. Within the various tissues and organs, arteries and veins are separated by a microvascular network of arterioles, capillaries, and venules. Due to the varying pressures and flows in the different parts of the vascular system, the precise structure and function of the vessels differ, although the overall structure of veins and arteries is similar in many ways.

Although the wall thickness varies between arteries and veins, they are both composed of three distinct layers. The intima forms the blood contacting layer and consists of a continuous single layer of endothelial cells that is attached to a layer of connective tissue, composed mainly of elastin and other matrix proteins to form the internal elastic lamina. This layer of connective tissue separates the intima from the cell-rich media. This layer is composed of multiple layers of concentrically arranged smooth muscle cells and collagen and elastin networks. The outermost layer is the adventitia, which

is composed entirely of connective tissue, myofibroblasts, capillaries, and nerves. These three layers are most obvious in the larger diameter veins and arteries.

Clinical Need for Vascular Replacement

According to the American Heart Association, cardiovascular disease claimed nearly 1 million lives in the United States in 1996 alone.^[1] One-half of these are due to coronary artery disease, which is currently the number-one cause of death in the western world. The main cause of coronary artery disease is arteriosclerosis (hardening of the arteries), a disorder characterized by progressive thickening and hardening of the walls of the arteries, causing a decrease in or loss of blood circulation (Fig. 1). In addition, erosion or disruption of the calcified plaques within the arterial wall may trigger the coagulation cascade, resulting in thrombotic occlusion and, often, sudden cardiac death. Since the days of Alexis Carrell, the standard surgical solution to cardiovascular disease has been vascular replacement, or bypass.

The Need for a Tissue-Engineered Blood Vessel

The saphenous vein is the most widely used conduit for bypass grafting,^[2] despite its short- and long-term failure remaining a significant clinical problem.^[3] Diffuse intimal thickening occurs in all autologous saphenous vein grafts within the first year, and late failure of grafts, usually 5–10 years after implantation, is associated with intimal hyperplasia (IH) and atherosclerotic lesions.^[4] Development of intimal thickening is regarded as an unavoidable adaptation of the saphenous vein to the changed hemodynamic conditions.^[5] The process is associated with a migration of adventitial myofibroblasts and smooth muscle cells from the adventitia and media to the intima, where the cells proliferate and change from a contractile to a synthetic phenotype. This arterIALIZATION of the vessel may initially increase patency, but ultimately, intimal hyperplasia is often involved in occlusion of the vessels.

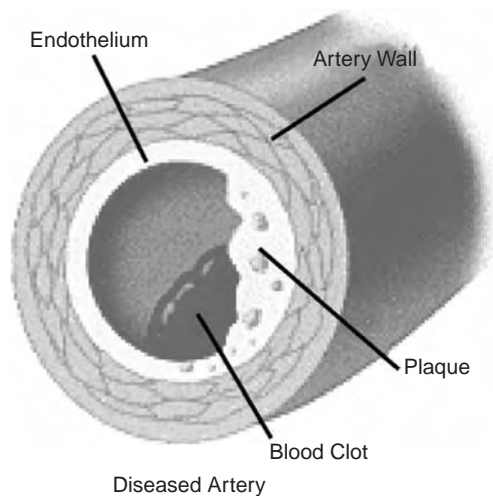


Fig. 1 Stenosed coronary artery. (From Texas Heart Institute at <http://www.tmc.edu/thi/cad.html>.)

There are limitations on alternative conduits for coronary bypass. Arterial grafts are the preferred conduits due to improved long-term patency,^[6] but they are limited due to their short size, small diameter, and availability, because many patients require multiple grafts and synthetic grafts have had only limited success, mainly in larger-diameter vessels.^[7] These clinical shortcomings in the treatment of vascular disease guided many attempts in the development of alternative bypass grafts, including synthetic grafts^[7] and grafts composed of decellularized porcine and human vessels.^[8]

The first report on the use of synthetic grafts was in 1954 when Voorhees showed the potential of using Vinyon-N cloth tubes to repair arterial defects. Subsequent investigators established the potential use of polymeric materials, like Dacron (polyethylene terephthalate), polyesters, and ePTFE (poly-tetrafluoroethylene), to form either permanent or resorbable replacements for the damaged vessels (Fig. 2). In cases where the graft can be of large diameter (greater than 6 mm), the synthetic material has been extremely effective. The main clinical problem that faces surgeons using synthetic graft materials is that they often cause severe inflammatory and thrombotic reactions. In small-diameter synthetic vascular grafts, fibrous deposition and thrombus formation will eventually lead to stenosis of the graft. To counteract such reactions, attempts were made to modify the graft materials.^[9] For example, early attempts at promoting endothelialization of synthetic vascular grafts and stents came in the form of EC seeding and sodding, or modifying the surface of polymer graft materials to promote the attachment of EC. However, successful reendothelialization of synthetic vascular grafts is still unrealized, making their use for small-diameter vascular bypass suboptimal.

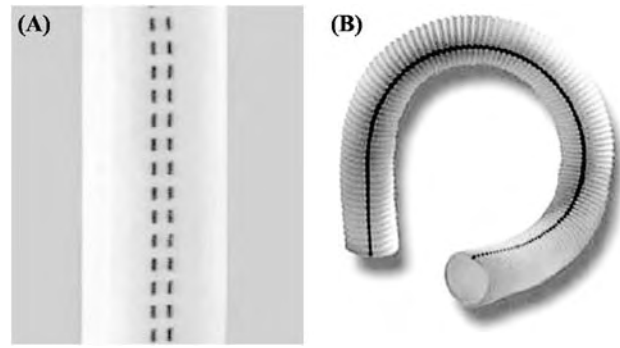


Fig. 2 (A) Dacron (polyethylene terephthalate) and (B) unsealed woven polyester vascular graft. (From www.vascutek.com.)

Due to these inherent limitations on synthetic grafts, alternative strategies led to the development and use of vessels derived from either humans (allogeneic) or animals such as pigs (xenogeneic).^[10] In order to be clinically viable, such grafts must be decellularized to prevent undesirable immune responses that would lead to failure of the grafts upon implantation. Either detergent (Triton X-100 or sodium dodecyl sulfate) or nondetergent techniques (most commonly enzymatic treatments) are employed, that do not alter the structure of the tissue but could affect the performance of the graft. Although such grafts possess a structure that is initially mechanically stable, they lack a cellular component necessary for vasoactivity and antithrombogenicity. In a recent study, Conklin et al. evaluated the potential of using decellularized porcine carotid arteries cross-linked with heparin to reduce thrombogenicity.^[11] They demonstrated that when implanted in a dog model, they became populated with smooth muscle, and at 2 months, an endothelial layer was evident in the lumen. However, although successful in animal models, the use of decellularized grafts remains problematic in humans. Dardik et al. evaluated the feasibility of glutaraldehyde-treated human umbilical cord veins as grafts in humans, and most failed within 2 years.^[10] Failure of these grafts usually occurs due to late aneurysmal degradation and dilation. Similarly, studies using decellularized homologous saphenous veins for coronary artery bypass reported patency rates at 1 year of only 35%.^[8] These suboptimal patency rates highlight the limitations of using decellularized matrices for vascular applications.

DESIGN CONSIDERATIONS FOR A TISSUE-ENGINEERED BLOOD VESSEL

Many scientists working in the field of regenerative medicine feel that a tissue-engineered vascular graft

will be one of the first successful clinical applications of tissue engineering.^[12] This section highlights the criteria governing current TEBV requirements and the techniques used to fabricate them, and describes recent attempts at using such techniques.

Mimicry of Native Vascular Tissue

There are certain basic properties of the native vasculature that a successful tissue-engineered blood vessel must achieve. These can be subdivided into mechanical and biological properties.^[13] First, the vessel must be of sufficient mechanical strength to not rupture when subjected to changing blood pressures and flows. This means that any technique used to fabricate the blood vessel must result in a conduit that will function immediately upon implantation. It is also essential that the TEBV match the mechanical properties of the vessel to which it is being grafted, because compliance mismatch at the anastomosis is thought to be involved in the development of intimal hyperplasia and subsequent graft failure.^[14] From a biological standpoint, it is essential that any TEBV be antithrombogenic. Finally, a major requirement (and a major limitation) in current vascular grafts is that the vessel must be vasoresponsive in order to function under changing blood pressures and flows. To date, no vascular graft has been developed that matches the specified criteria; this is where the field of tissue engineering holds great promise.

Considerations for TEBV Fabrication

Cell source

In a recent review, Nerem highlighted cell source as one of the keys to a successful vascular tissue-engineering approach.^[13] In order to be clinically viable, any vascular tissue-engineering technique should utilize a method by which autologous cells may be easily harvested and used. Most previous studies on TEBV fabrication have used either terminally differentiated or nonautologous cells.^[15–17] The use of nonautologous cells is limited by difficulties in harvest or isolation, inability of the cells to expand, and immunorejection issues. Niklason et al. recently reported that after in-vitro expansion (as required for tissue-engineering applications), terminally differentiated cells derived from a 2-year-old human aorta were incapable of expanding into a functional tissue,^[18] emphasizing the potential limitations on autologous cells. Attempts to transfect the cells with telomerase to stimulate cell proliferation has had limited success, and would not be practical in the clinical setting within a foreseeable time-frame.

There is clearly a need for an alternative cell source for TEBV development. One such approach utilized by Julie Campbell is to use the cells present in the peritoneal cavity to form blood vessels.^[19] This new procedure involves growing tissue around an artificial tube placed inside a patient's abdominal cavity and then using the tissue to replace diseased or damaged blood vessels. Peritoneal myofibroblasts and macrophages attach to the tube and form granulation tissue composed of collagen and elastin, which is subsequently covered in a layer of mesothelial cells that express von Willebrand factor. Vessels formed in this manner have shown patency in rabbit and rat models for up to 4 months, with the constituent cells assuming a smooth musclelike phenotype with time.^[19] However, the use of such vessels in clinical trials has yet to be achieved; other cell sources must also be considered.

A second potential source of cells is stem, or progenitor, cells, which have shown great potential for use in tissue-engineering applications, and may circumvent many of the problems associated with cell sourcing. The notion that primitive progenitor cells exist in adult humans was first suggested in 1867 by Cohnheim.^[20] Bone marrow-derived progenitor cells (BMPCs) are the best-characterized progenitor cells due to the focus on treating bone marrow-associated diseases such as leukemia. Bone marrow is known to contain primitive and multipotent progenitor cells, identifiable by cell surface markers that are capable of differentiating into several different hematopoietic and mesenchymal lineages. Specifically, in-vitro conditions have been identified that stimulate the differentiation of these progenitor cells, not only along hematopoietic lineages, but also to cardiovascular cell types such as cardiomyocytes.^[21] In addition, new data are emerging that suggest that BMPCs possess the ability to differentiate to smooth muscle cells.^[22]

Noishiki and colleagues reported that ePTFE grafts treated with bone marrow had increased neovascularization^[23] and improved endothelialization after 3 weeks following implantation in a canine model, resulting in improved patency at 3 and 6 months compared to untreated grafts.^[24] More recently, several in vivo studies have shown that BMPCs can contribute to blood vessel formation within the heart after myocardial infarction.^[25] Highly-enriched hematopoietic stem cells (CD34(-)/low, c-Kit(+), Sca-1(+)) were injected into a mouse model and subsequently migrated into ischemic cardiac muscle and blood vessels, contributing to the repair of the infarction and revascularization. Furthermore, a recent study has shown that circulating bone marrow-derived progenitor cells can contribute to neointima formation after injury.^[26] In this study, a mouse model of iliac artery injury was transfused with nucleated bone marrow cells from donor mice, and after 4 weeks, the arterial lumen was obliterated by a cell-rich

neointima, with cells containing smooth-muscle α actin, approximately half of which were of donor origin. These results suggest that bone marrow-derived cells are recruited in vascular healing as a complementary source of smooth musclelike cells, and are capable of differentiation in a biomechanically dynamic environment. Although these studies suggest that *in vivo* delivery of BMPCs could be useful in cardiovascular regeneration, the intrinsic factors that govern the recruitment and differentiation of these cells are yet to be elucidated *in vitro*. Factors that must be considered relate not only to cytokine and growth factor-mediated differentiation, but also to the influence of mechanical stimuli, as the cardiovascular system is one of the most dynamic environments in the body.^[27] The primary limitation on *in vivo* differentiation is that it is impossible to control the exact biochemical and biomechanical interactions (both magnitude and history) that stimulate differentiation.

The advantages to using BMPCs are numerous. Given the potential they show for cardiovascular tissue engineering, they offer a real alternative to terminally differentiated smooth muscle. First, they may be easily isolated from a patient for whom a vascular procedure requiring cells is intended. Furthermore, BMPCs are also known to produce growth factors that promote SMC proliferation and migration. This includes vascular endothelial growth factor (VEGF), basic fibroblast growth factor (bFGF), and platelet-derived growth factor-BB (PDGF-BB). In addition, we have completed experiments that suggest that BMPCs may differentiate to a smooth muscle lineage when grown under the appropriate biomechanical stimuli.^[28] When grown under 10% cyclic strain at 1 Hz for 7 days, BMPCs upregulated smooth muscle-specific genes (namely, smooth muscle α actin and h1-calponin) (Fig. 3). Furthermore, strain caused a cessation of proliferation and caused the cells to align perpendicular to the direction of strain. Given these important new findings in combination with information already known about BMPCs, they clearly show great promise as a cell source for vascular tissue-engineering applications, and as an alternative cell source for TEBV fabrication.

Construct design

One of the main areas of focus in vascular tissue engineering is developing the optimal method of forming a tubular construct. Factors that must be considered include the ability to control the inner diameter, wall thickness, and length.^[13] Following is a summary of various approaches to forming a tubular construct.

Collagen-Based Scaffolds. Initial construct design focused on collagen as the basis for TEBV fabrication

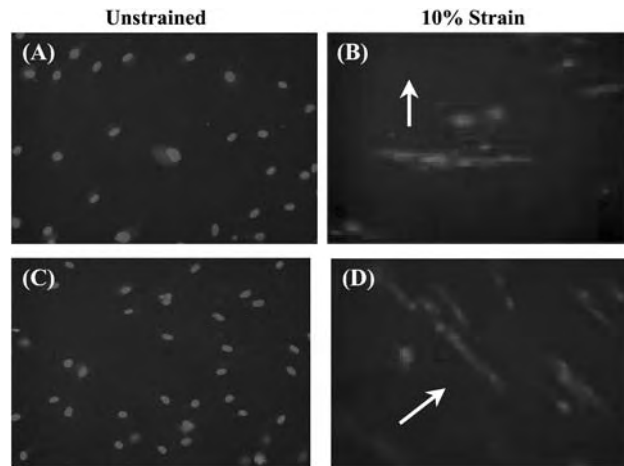


Fig. 3 Detection of smooth muscle α actin in (A) unstrained and (B) strained cultures, and h1-calponin in (C) unstrained and (D) strained cultures. Nuclei are labeled blue and stain is red. Note that only strained cultures are positive for each marker. Arrows indicate direction of strain. Mag \times 20.

because it is the predominant extracellular matrix (ECM) protein found in native arteries. Collagen has many potential advantages. For example, it is a natural material that can be readily isolated and formed into a tubular structure.^[12] In a 1986 paper that many credit as the impetus for the vascular tissue-engineering field, Weinberg and Bell reported the production of a two-layer TEBV. This was achieved with an annular mold, and featured a medial layer containing type I collagen and SMCs, and an adventitial layer containing collagen and fibroblasts. Seeding the luminal surface of the resulting tube with EC formed the intimal layer. However, Weinberg and Bell made an important observation about the histology of their TEBV: The SMC and collagen were oriented longitudinally rather than in the predominant circumferential orientation found in native blood vessels.^[29] It is believed that this longitudinal orientation of the tissue constituents compromised the strength of the TEBV. Furthermore, circumferential alignment ensures mechanical loading of SMC in a pressurized construct, and this is thought to be imperative to maintaining SMC phenotype.^[30]

Subsequent investigators have used a molding technique similar to Weinberg and Bell's, but have attempted to increase the strength of the collagen constructs using several different approaches. First, most cell types will contract collagen gels and this has been used to the advantage of being able to compact the collagen against a mandrel^[16] that, when removed, forms the lumen of the vessel. Other techniques include cross-linking of the collagen using glycation^[31] or enzymatic methods^[32] (Fig. 4). Other studies have

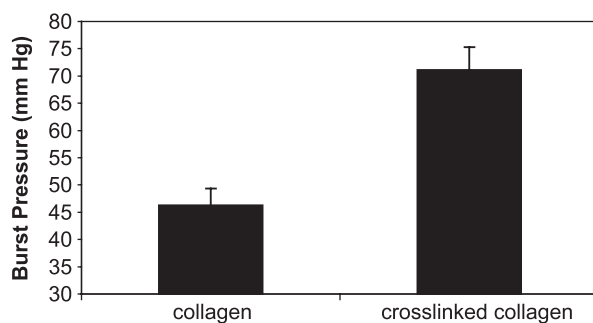


Fig. 4 Effect of enzymatic cross-linking with transglutaminase on collagen construct burst strength. (From Ref. [32].)

investigated the influence of collagen concentration on the mechanical strength of TEBVs. For example, Hirai et al. demonstrated that the burst pressures of their collagen-based tubes increased with increasing collagen concentration, although they only succeeded in improving the pressure to 110 mm Hg.^[33] Subsequent attempts to organize the collagen used magnetic pre-alignment, but this worked only in constructs without a central mandrel,^[34] limiting the formation of a central lumen in the TEBV.

Although collagen gels have taken precedence in the development of a TEBV, other studies are showing the potential of native, collagen-based tissues such as small intestine submucosa used as the basis for a graft. In one approach, Huynh et al. used the submucosa derived from porcine intestine to produce an acellular graft, which showed patency in a rabbit model for 13 weeks.^[35] Other studies have shown the potential of submucosa used in combination with cells as a potential graft, with patency rates in a rat model of up to 8 weeks.^[36]

As many advantages as collagen has as a scaffold, it also confers certain limitations on the fabrication of a fully-functional TEBV. It is known that cells grown on collagen gels show a tendency to switch off collagen synthesis,^[37] and show reduced proliferation, which may be major limiting factors when vascular tissue-engineering applications require remodeling and growth within the construct. However, it is clear that collagen—either in gel or native form—shows good potential as the basis for fabrication of a TEBV.

Cellular-Based Approaches. By definition, the field of tissue engineering is based on the utilization of cells and constructs to produce functional tissues that can be reimplanted into the body.^[38] In 1998, L'Heureux and colleagues produced a TEBV that utilizes human vascular SMCs cultured with ascorbic acid to produce a cohesive cellular sheet.^[16] This sheet, as well as a similar sheet of human fibroblasts, was wrapped

around a cylindrical mandrel to produce a tubular construct with an inner SMC-based media and an outer fibroblast-based adventitia. After maturation, the tubular support was removed and endothelial cells were seeded on the lumen. The TEBV resembled a native human artery, and its histology showed well-defined intima, media, and adventitia (Fig. 5). However, although positive for smooth muscle α actin and desmin, SMC density was lower than is evident in a normal artery. A confluent endothelial layer was observed on the luminal surface, although upon implantation in a dog model, the grafts remained patent for only 7 days with significant thrombus formation. Burst strength measurements showed that the TEBVs had good mechanical properties, bursting at higher pressures than the human saphenous vein. Although limited by poor thrombogenicity, this study clearly shows the potential of forming TEBVs with good mechanical properties using a cellular-based approach.

Scaffold-Based Approaches. Although the study by L'Heureux and colleagues demonstrated the ability to generate TEBVs using a purely cellular-based approach, most vascular tissue-engineering approaches have relied on some form of scaffold—natural or synthetic—to form the basis of the graft. Many of the techniques employed in the fabrication of a tissue-engineered blood vessel revolve around mechanical integrity of the graft, (incorporation of polymer scaffolds) such that it will not rupture upon implantation when faced with arterial levels of pressures and flows or faced with cyclic deformations. Furthermore, in many vascular tissue-engineering applications, mechanical conditioning is implemented in the fabrication process, and the presence of synthetic materials allows such regimens to be applied without compromising the TEBV. The synthetic materials used as constructs have varied, ranging from biodurable materials such as Dacron to biodegradable polymers such as polyglycolic acid (PGA). Niklason et al. created a TEBV from bovine aortic SMCs and a PGA mesh.^[17] In this approach, the SMCs were seeded onto PGA tubes and then placed in a bioreactor for 8 weeks under pulsatile radial forces. The resulting TEBVs demonstrated contractile responses to pharmacological agents and contained SMCs that maintained markers of differentiation such as h1-calponin and myosin heavy chain (Fig. 6). In addition, the TEBVs showed good mechanical properties, such that they could be implanted in a porcine model where they showed patency for 4 weeks. However, long-term patency of this TEBV remains to be proved, and complications such as the production of undesirable breakdown products from the polymer scaffold must be addressed. Although the field of polymer development has evolved significantly, there remains a need to develop materials

T

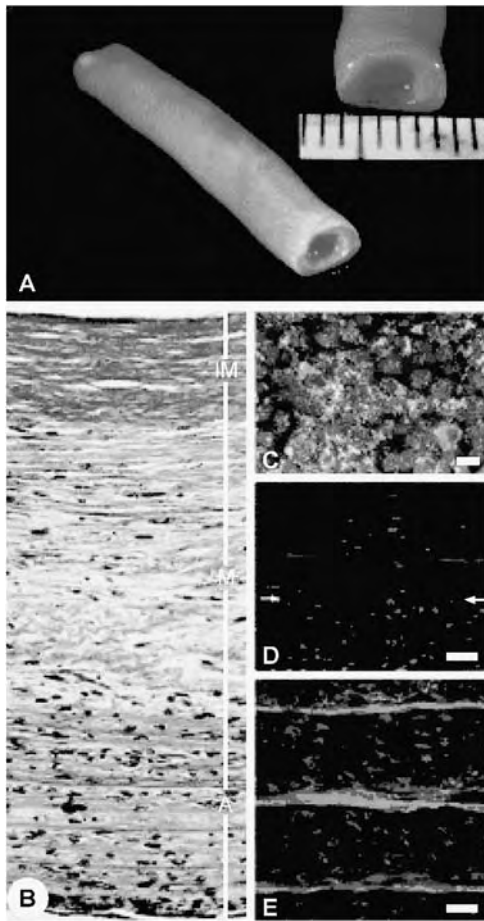


Fig. 5 Organization of the TEBV developed by L'Heureux et al.^[16] (A) Macroscopic view of a mature TEBV. The vessel is self-supporting when removed from culture medium (open lumen = 3 mm). (B) Paraffin cross section of the vascular wall stained with Masson's trichrome shows collagen in blue-green and cells in dark purple. Aside from an oversized internal elastic lamina (IM = 125 μ m), the histology is similar to that of a muscular artery with a large media (M = 320 μ m) and a surrounding adventitia (A = 235 μ m). (C) Frontal view of the endothelium seeded on the IM. Cytoplasmic green fluorescence reveals cell viability, metabolic activity, and degree of confluence. Red fluorescence (DII-ac-LDL uptake) confirms cell viability and the endothelial nature of the cells. Blue fluorescence shows the characteristic von Willebrand factor expression in ECs (orange = red + green; pink = red + green + blue). Scale bar = 25 μ m. (D) The media-adventitia junction (arrows) stained for desmin (nuclei are stained blue). Scale bar = 50 μ m. (E) The adventitia double-stained for elastin (green) and vimentin (red). Scale bar = 50 μ m. (From Ref. [16].)

that exhibit both acceptable biomechanical properties and biocompatibility.^[39] Factors that must be considered include matching polymer degradation rates to ECM production and enhancing adhesion and differentiation.

Mechanical conditioning

The vascular system is an extremely dynamic environment, wherein cells in the blood vessel wall are exposed to pulsatile pressure and flow, causing cyclic deformations and fluid shear stress. In addition, in the coronary system, blood vessels are exposed to cyclic bending, twisting, and stretching due to the beating of the heart.^[40] Smooth muscle and endothelial cells respond to these stresses and strains to maintain tissue in a quiescent state.^[41]

Biomechanical stimulation also has a profound effect on cellular phenotype. It is well-documented in the literature that a loss of cyclic strain causes an alteration in SMC from contractile to synthetic phenotype, thereby causing a loss of function within the vessel.^[42] Therefore, if a TEBV is fabricated using differentiated SMCs but is then placed in static culture, it may lose its phenotypic profile. Several laboratories have focused on the design of bioreactors to apply physiological biomechanics to TEBVs. Niklason et al. employed such a system to apply 5% radial distension through inflation of a silicone mandrel (Fig. 7). When static and conditioned vessels were compared, the latter had a higher collagen content, a histological appearance closely resembling native tissue, and a burst strength of 2000 mm Hg, compared with 300 mm Hg for static controls. In addition, a study by Seliktar and colleagues investigated the effect of cyclic strain on SMC/collagen constructs over an 8-day culture period, and observed increased contraction and mechanical strength that correlates with circumferential cellular and collagen orientation.^[43] This is an interesting contrast to the observations of Weinberg and Bell, whose constructs had longitudinal cellular and collagen orientations and subsequently exhibited low burst strength. It thus appears that the application of strain not only influences cellular phenotype, but also tissue structure and matrix composition.

Although studies to date have focused on the application of circumferential strains, future studies should consider the effects of physiologically relevant mechanical stimuli. For example, if a TEBV is required for coronary bypass, then it may be appropriate to condition it with coronary-consistent cyclic stretching, bending, and twisting in addition to circumferential straining. To this end, our laboratory has developed vascular perfusion systems that allow the application of precisely controlled pressures and flows,^[40,44,45] along with cyclic stretching, twisting, and bending.

CURRENT LIMITATIONS OF TEBVs

None of the TEBV approaches to date has become clinically viable. This section discusses the major

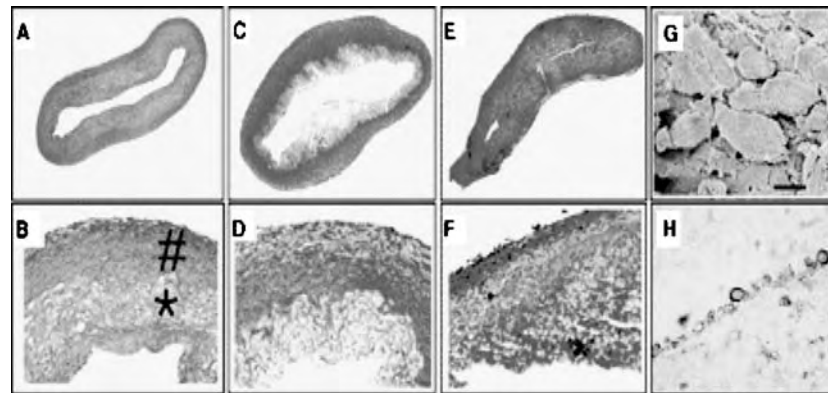


Fig. 6 Histology of the engineered blood vessels developed by Niklason et al.^[18] (A and B) Pulsed vessel cultured for 8 weeks; (A) Verhoeff's elastin stain [original magnification (orig. mag.) $\times 20$]; (B) Masson's trichrome stain; collagen stains blue (orig. mag. $\times 100$); (C and D) Nonpulsed vessel cultured for 8 weeks; (C) Verhoeff's stain (orig. mag. $\times 20$); (D) Masson's stain (orig. mag. $100 \times$); (E and F) Pulsed vessel without medium supplementation; (E) Verhoeff's stain (orig. mag. $\times 20$); (F) Masson's stain (orig. mag. $\times 100$); (G) Scanning electron microscopy of the endothelial cell layer in an engineered vessel. Scale bar, $10 \mu\text{m}$; (H) Immunoperoxidase staining for PECAM antigen reveals an EC monolayer on the vessel lumen (orig. mag. $\times 1000$). (From Ref. [17].)

limitations of these approaches and potential strategies to circumvent them.

Clearly, the most prominent shortcomings of previously described TEBVs—whether with or without synthetic materials—is the requirement of antithrombogenicity. Within a blood vessel, the endothelial layer functions to prevent the adhesion of platelets and other blood components such as macrophages. The endothelial cells prevent such adhesions by not only acting as a physical barrier, but also by continued release of factors such as nitric oxide and von

Willebrand factor. In many of the tissue-engineered blood vessels fabricated to date, patency has been maintained for only a few days after implantation, and thrombus deposition has been evident in the endothelial layer upon explantation. Although this thrombogenicity is likely multifactorial, the fact that nonautologous cells were used in most cases highlights this as a potential limitation in vascular tissue-engineering applications. Even though nonautologous cells may be shown to be partially functional (e.g., the production of von Willebrand factor by the endothelial layer of a TEBV), they may not be fully functional. Furthermore, autologous endothelial cells isolated from, e.g., adipose tissue, also often prove to be problematic for reasons such as an inability to expand in culture, undesirable phenotypic changes, and an inability to withstand blood vessel stresses.

Another major problem in developing a TEBV is the possible mismatch of mechanical properties between graft and host artery.^[46] Such a mismatch can lead to elevated suture line stresses as well as elevated shear stress levels and turbulent flow, which can in turn lead to development of anastomotic intimal hyperplasia, the most common failure mechanism of bypass grafts. The challenge in engineering a TEBV, therefore, includes ensuring that it has biomechanical properties similar to a native artery's.

Another rate-limiting barrier to the fabrication of TEBVs is the length of time required to expand the cells, seed the constructs, and mechanically condition them. Current time frames to fabricate a TEBV are in the order of 10 to 15 weeks—clearly too long for many patients in need of bypass surgery. For any TEBV

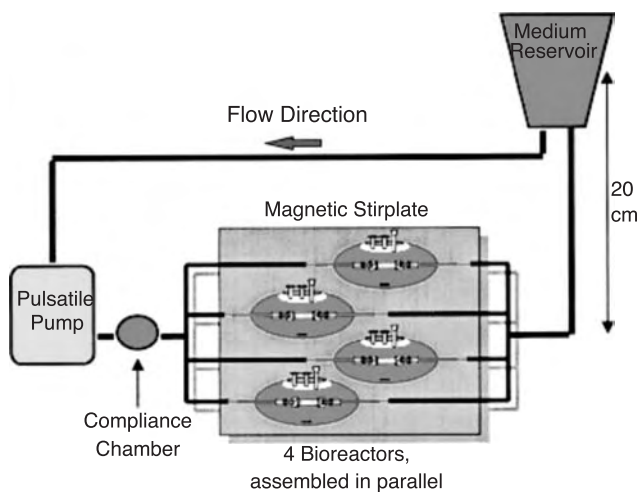


Fig. 7 Biomimetic system for blood vessel culture used by Niklason et al.^[18] The bioreactors and fluid reservoir were fitted to provide gas exchange, and buffer flowed through highly distensible silicone tubing inserted through the vessel lumen. (From Ref. [17].)

application to be clinically successful, such fabrication times must be minimized. One potential solution may be to optimize the use of mechanical conditioning,^[43] which accelerates remodeling.

Finally, TEBVs' greatest limitation may be the lack of performance data in a clinical setting, or even in an animal model. It is important to demonstrate that the constructs have the ability to remodel and repair themselves, while resisting the development of pathologies such as thrombosis, arteriosclerosis, and intimal hyperplasia.

CONCLUSION

Treatment modalities developed to date fall short of what is required, and vascular disease remains the greatest killer in the Western world. In recent years, the design and fabrication of the TEBV has come a long way, but the challenges remain large (although not insurmountable). The significant hurdles that remain to be overcome include cell sourcing (differentiated cells vs. progenitor cells) and construct design (natural materials vs. synthetic). An area that deserves further investigation is the use of biomechanical conditioning as a means to reduce fabrication time. With improvements rapidly occurring in these areas, and with increased understanding of cell biology and biomechanics, it is likely that the first clinically acceptable TEBV will be developed in the near future.

ARTICLES OF FURTHER INTEREST

Bioreactors; Blood Vessel Mechanics; Collagen; Tissue Engineering Scaffolds; Vascular Grafts; Vascular Grafts: Development Strategies; Vascular Grafts: Host Interactions

REFERENCES

1. American Heart Association *Heart and Stroke Statistical Update*; 2001. Dallas.
2. von Segesser, L.K. *Arterial Grafting for Myocardial Revascularization: Indications, Surgical Techniques and Results*; Springer Verlag: New York, 1991.
3. Pelletier, L.C. The Saphenous Vein Graft: What Have We Learned from the Past 25 Years. In *Conduits for Myocardial Revascularization*; Carrier, M., Pelletier, L.C., Eds.; RG Landes Co.: Austin, 1993; 3–34.
4. Yamada, T.; Shiraiishi, R.; Taki, K.; Nakano, S.; Tokunaga, O.; Itoh, T. Immunohistochemical and ultrastructural examination of smooth muscle cells in aortocoronary saphenous vein grafts. *Angiology* **1997**, *48* (5), 381–390.
5. Mavromatis, K.; Fukai, T.; Tate, M.; Chesler, N.; Ku, D.N.; Galis, Z.S. Early effects of arterial hemodynamic conditions on human saphenous veins perfused ex vivo. *Arterioscler. Thromb. Vasc. Biol.* **2000**, *20* (8), 1889–1895.
6. Predel, H.G.; Yang, Z.; von Segesser, L.; Turina, M.; Buhler, F.R.; Luscher, T.F. Implications of pulsatile stretch on growth of saphenous vein and mammary artery smooth muscle. *Lancet* **1992**, *340* (8824), 878–879.
7. Williams, S.K.; Rose, D.G.; Jarrell, B.E. Microvascular endothelial cell seeding of ePTFE vascular grafts: Improved patency and stability of the cellular lining. *J. Biomed. Mater. Res.* **1994**, *28* (2), 203–212.
8. Gelbfish, J.; Jacobowitz, I.J.; Rose, D.M.; Connolly, M.W.; Acinapura, A.J.; Zisbrod, Z.; Lim, K.H.; Cappabianca, P.; Cunningham, J.N. Cryopreserved homologous saphenous vein: Early and late patency in coronary artery bypass surgical procedures. *Ann. Thorac. Surg.* **1986**, *42* (1), 70–73.
9. Greisler, H.P. *New Biologic and Synthetic Vascular Prosthesis*; RG Landes: Austin, TX, 1991.
10. Dardik, H.; Ibrahim, I.M.; Dardik, I. Modified and unmodified umbilical vein allografts and xenografts as arterial substitutes: Morphologic assessment. *Surg. Forum* **1975**, *26*, 286–287.
11. Conklin, B.S.; Richter, E.R.; Kreutziger, K.L.; Zhong, D.S.; Chen, C. Development and evaluation of a novel decellularized vascular xenograft. *Med. Eng. Phys.* **2002**, *24* (3), 173–183.
12. Ratcliffe, A. Tissue engineering of vascular grafts. *Matrix Biol.: J. Int. Soc. Matrix Biology* **2000**, *19* (4), 353–357.
13. Nerem, R.M.; Seliktar, D. Vascular tissue engineering. *Annu. Rev. Bioeng.* **2001**, *3*, 225–243.
14. Salacinski, H.J.; Goldner, S.; Giudiceandrea, A.; Hamilton, G.; Seifalian, A.M.; Edwards, A.; Carson, R.J. The mechanical behavior of vascular grafts: A review. *J. Biomater. Appl.* **2001**, *15* (3), 241–278.
15. Weinberg, C.B.; Bell, E. A blood vessel model constructed from collagen and cultured vascular cells. *Science* **1986**, *231* (4736), 397–400.
16. L'Heureux, N.; Paquet, S.; Labbe, R.; Germain, L.; Auger, F.A. A completely biological tissue-engineered human blood vessel [see comments]. *FASEB J.* **1998**, *12* (1), 47–56.
17. Niklason, L.E.; Gao, J.; Abbott, W.M.; Hirschi, K.K.; Houser, S.; Marini, R.; Langer, R. Functional arteries grown in vitro [see comments]. *Science* **1999**, *284* (5413), 489–493.
18. Niklason, L. Progress in the Engineering of Small Caliber Arterial Prostheses. In *Engineering Tissue Growth International Conference & Exposition*; 2002Pittsburgh
19. Campbell, J.H.; Efendy, J.L.; Campbell, G.R. Novel vascular graft grown within recipient's own peritoneal cavity. *Circ. Res. (Online)* **1999**, *85* (12), 1173–1178.
20. Cohnheim, J. *Arch. Pathol. Anat. Physiol. Klin. Med.* **1867**, *40*, 1.

21. Fukuda, K. Development of regenerative cardiomyocytes from mesenchymal stem cells for cardiovascular tissue engineering. *Artif. Organs* **2001**, *25* (3), 187–193.
22. Arakawa, E.; Hasegawa, K.; Yanai, N.; Obinata, M.; Matsuda, Y. A mouse bone marrow stromal cell line, TBR-B, shows inducible expression of smooth muscle-specific genes. *FEBS Lett.* **2000**, *481* (2), 193–196.
23. Noishiki, Y.; Tomizawa, Y.; Yamane, Y.; Matsumoto, A. Autocrine angiogenic vascular prosthesis with bone marrow transplantation [see comments]. *Nat. Med.* **1996**, *2* (1), 90–93.
24. Noishiki, Y.; Yamane, Y.; Tomizawa, Y.; Matsumoto, A. Transplantation of autologous tissue fragments into an e-PTFE graft with long fibrils. *Artif. Organs* **1995**, *19* (1), 17–26.
25. Hirschi, K.; Goodell, M. Common origins of blood and blood vessels in adults? *Differ. Res. Biol. Divers.* **2001**, *68* (4–5), 186–192.
26. Han, C.I.; Campbell, G.R.; Campbell, J.H. Circulating bone marrow cells can contribute to neointimal formation. *J. Vasc. Res.* **2001**, *38* (2), 113–119.
27. Nerem, R.M.; Girard, P.R. Hemodynamic influences on vascular endothelial biology. *Toxicol. Pathol.* **1990**, *18* (4 Pt 1), 572–582.
28. Hamilton, D.W.; Maul, T.M.; Vorp, D.A. Characterization of the response of bone marrow derived progenitor cells to cyclic strain: Implications for vascular tissue engineering applications. *Tissue Eng. in press*.
29. Sumpio, B.E. *Hemodynamic Forces and Vascular Cell Biology*; R.G. Landes Co.: Austin, TX, 1993.
30. Kanda, K.; Matsuda, T.; Oka, T. Mechanical stress induced cellular orientation and phenotypic modulation of 3-D cultured smooth muscle cells. *ASAIO J.* **1993**, *39* (3), M686–M690.
31. Girton, T.S.; Oegema, T.R.; Tranquillo, R.T. Exploiting glycation to stiffen and strengthen tissue equivalents for tissue engineering. *J. Biomed. Mater. Res.* **1999**, *46* (1), 87–92.
32. Orban, J.M.; Wilson, L.B.; Kofroth, J.A.; El-Kurdi, M.S.; Maul, T.M.; Vorp, D.A. Cross linking of collagen gels by transglutaminase. *J. Biomed. Mater. Res.* revision under review.
33. Hirai, J.; Kanda, K.; Oka, T.; Matsuda, T. Highly oriented, tubular hybrid vascular tissue for a low pressure circulatory system. *ASAIO J.* **1994**, *40* (3), M383–M388.
34. Tranquillo, R.T.; Girton, T.S.; Bromberek, B.A.; Tribes, T.G.; Mooradian, D.L. Magnetically orientated tissue-equivalent tubes: Application to a circumferentially orientated media-equivalent. *Biomaterials* **1996**, *17* (3), 349–357.
35. Huynh, T.; Abraham, G.; Murray, J.; Brockbank, K.; Hagen, P.O.; Sullivan, S. Remodeling of an acellular collagen graft into a physiologically responsive neovessel. *Nat. Biotechnol.* **1999**, *17* (11), 1083–1086.
36. Badylak, S. The extracellular matrix as a scaffold for tissue reconstruction. *Semin. Cell Dev. Biol.* **2002**, *13* (5), 377.
37. Grassl, E.D.; Oegema, T.R.; Tranquillo, R.T. Fibrin as an alternative biopolymer to type-I collagen for the fabrication of a media equivalent. *J. Biomed. Mater. Res.* **2002**, *60* (4), 607–612.
38. Nerem, R.M. Tissue engineering in the USA. *Med. Biol. Eng. Comput.* **1992**, *30* (4), CE8–CE12.
39. Greisler, H.P.; Gosselin, C.; Ren, D.; Kang, S.S.; Kim, D.U. Biointeractive polymers and tissue engineered blood vessels. *Biomaterials* **1996**, *17* (3), 329–336.
40. Vorp, D.A.; Severyn, D.A.; Steed, D.L.; Webster, M.W. A device for the application of cyclic twist and extension on perfused vascular segments. *Am. J. Physiol.* **1996**, *270* (2 Pt 2), H787–H795.
41. Shyy, J.Y. Mechanotransduction in endothelial responses to shear stress: Review of work in Dr. Chien's laboratory. *Biorheology* **2001**, *38* (2–3), 109–117.
42. Williams, B. Mechanical influences on vascular smooth muscle cell function. *J. Hypertens.* **1998**, *16* (12 Pt 2), 1921–1929.
43. Seliktar, D.; Black, R.A.; Vito, R.P.; Nerem, R.M. Dynamic mechanical conditioning of collagen-gel blood vessel constructs induces remodeling in vitro. *Ann. Biomed. Eng.* **2000**, *28* (4), 351–362.
44. Vorp, D.A.; Severyn, D.A.; Mears, J. An experimental system to expose perfused vascular segments to cyclic bending ex vivo. *Adv. Bioeng.* **1997**, *36*, 115–116.
45. Vorp, D.A.; Peters, D.G.; Webster, M.W. Gene expression is altered in perfused arterial segments exposed to cyclic flexure ex vivo. *Ann. Biomed. Eng.* **1999**, *27* (3), 366–371.
46. Tai, N.R.; Salacinski, H.J.; Edwards, A.; Hamilton, G.; Seifalian, A.M. Compliance properties of conduits used in vascular reconstruction. *Br. J. Surg.* **2000**, *87* (11), 1516–1524.

Tissue Engineering of Bone

L. Di-Silvio

Institute of Orthopaedics & Musculo-Skeletal Science, University College London, Royal National Orthopaedic Hospital Trust, Middlesex, and Biomaterials Science, The Guy's, King's, and St. Thomas' Hospital Medical and Dental Institute, King's College London, London, United Kingdom

N. Gurav

Institute of Orthopaedics & Musculo-Skeletal Science, University College London, Royal National Orthopaedic Hospital Trust, Middlesex, United Kingdom

Eleftherios Tsiridis

The Whittington Hospital, University College London, London, and Institute of Orthopaedics, Royal National Orthopaedic Hospital, Stanmore, United Kingdom

INTRODUCTION

Tissue engineering is a new development in biomedicine, involving a series of strategies using biologically based mechanisms to repair and heal damaged and diseased tissue. The key elements include a specific living cell type (or several cell types), a material scaffold that forms a supporting structure for culturing the cells in vitro and surgical delivery in vivo to the patient, and, for the majority of mammalian cell types, a growth stimulus.

Tissue loss or end organ failure resulting from trauma or disease is a major health problem. Examples of bone-related illness include osteoarthritis, cancer, trauma, fetal defects, and metabolic disorders. Currently available methods include mechanical devices or artificial prostheses, which do not repair the tissue or organ function and do not integrate with the host tissue. In the aging population worldwide, the incidence of bone-related illness is more prevalent than ever before, and bone transplantation is a commonly required surgical procedure. In the UK there are approximately 150,000 fractures (wrist, vertebral, and hip) due to osteoporosis, with high rates of associated morbidity and mortality. In the United States bone tissue is used in over 450,000 graft procedures, and total national health care costs for these patients exceed \$120 billion per year.^[1] Hence, there is an increasing demand for developing bone augmentation strategies. Bone tissue engineering is an emerging interdisciplinary field seeking to address needs for the development of biological substitutes, by combining the principles of orthopedic surgery with basic science (such as chemistry and cellular and molecular biology) and engineering (chemical and structural), to develop strategies for the restoration, maintenance, or improvement of tissue form and function. Bone is a living structure and as such, has the ability to regenerate through a dynamic biological integration of biochemical, cellular,

and hormonal processes. However, if the lesion is above critical size, the site will not regenerate but will become scarred, leading to nonunion at the site. To restore normal function, tissue regeneration has to occur; hence new bone tissue has to be surgically introduced to bridge the defect.

OVERVIEW

There are approximately 500,000 surgical procedures performed each year in the United States that require bone substitutes. These substitutes include autografts, allografts, and synthetic materials. The autograft, which involves harvesting viable tissue from one site in the patient and transplanting it into another part of the same patient, is considered the current gold standard against which other bone transplantation choices are weighed. Commonly used autograft procedures in orthopedics include spinal fusion and the treatment of nonunion fracture, where bone is usually taken from the patient's iliac crest and used to stabilize the defect site. It has proved highly successful due to its immunological compatibility and direct transfer of osteogenic cells and osteoinductive cues. However, it has limited therapeutic potential, as donor tissue is usually short in supply, collection is painful, and there are risks of donor site morbidity and infection with the patient having two operations. Furthermore, it is highly dependent on the quality and quantity of patient bone that can be harvested. The allograft, tissue transplanted from a different donor host, provides an abundant supply of both cortical and trabecular bone; however, it can cause an immune response and potentially transmit viral diseases. In addition, its processing removes all cells and bioactive molecules, compromising its incorporation into existing host tissue. The natural skeletal repair process of fracture

healing utilizes bone-forming cells, a cartilaginous scaffold upon which the new woven bone forms, and bioactive molecules to direct the repair sequence. An ideal augmentation would be “de-novo” tissue which has the same function, structure, and mechanical integrity as the original tissue that was lost. The cartilaginous scaffold recognized and ossified by the body is porous and allows vascularization and cellular ingrowth. Hence, the vast majority of engineered approaches are based on porous scaffolds. Bone analogue materials such as porous hydroxyapatite, composites, and degradable polymers such as poly(glycolic acid) and poly(lactic acid) are also being developed. In the main, these materials may be osteoconductive in nature; few, however, provide osteoinductivity.

Bone morphogenesis is a sequential cascade with three key phases: chemotaxis and mitosis of mesenchymal cells, differentiation of the mesenchymal cells initially into cartilage, and replacement of the cartilage by bone. The natural repair mechanism of bone utilizes bone-forming cells, a cartilaginous scaffold upon which the new woven bone is formed, and bioactive molecules to direct the repair process. Tissue engineering of bone endeavors to provide some or all of these components.

CELLS

Osteogenic function is necessary for biomaterial integration into the host environment. In bone engineering there are basically four types of cells that can be used: unfractionated fresh bone marrow; purified, culture-expanded mesenchymal stem cells (MSCs); embryonic stem cells (ESCs); and differentiated osteoblasts.

Unfractionated bone marrow autotransplantation was one of the very first techniques to be used in bone tissue engineering. This was an empirical procedure and seemed ideal, as bone marrow is rich in osteoprogenitor cells and osteogenic precursors such as bone morphogenetic proteins (BMPs). The procedure involves taking bone marrow from a donor site (the iliac crest) and transplanting it into the defect site. The procedure was relatively simple, inexpensive, and did not require FDA regulation. This type of approach in a rat model has shown favourable results.^[2-5] However, limited sources and the scarcity of osteogenic cells after aging, disease, and irradiation prevented its continued widespread use.^[6,7]

MSCs are immature and undifferentiated cells that are obtained from bone marrow and the periosteum. They have the capacity for extensive replication without differentiation and they possess multilineage developmental potential, making them a powerful research tool.^[8] An important issue with respect to these cells is understanding how to induce them and

controlling their differentiation into the desired cell type. Stem cells are either pluripotent or multipotent depending on their lineage. These stem cells are able to differentiate into fibroblastic, adipogenic, reticular, and most osteogenic cells.^[9,10] They are capable of multiple passages without loss of stem cell characteristics and tend not to de-differentiate unless exposed to specific biochemical and mechanical cues, in which case they can be directionally differentiated.^[11] Furthermore, the cells can be grown in a bioreactor, which provides an ideal environment, where a permeable membrane allows gaseous exchange and the microgravity environment created by the rotation reduces flow stress on the construct and encourages the cells to remain within the construct. After a given time period, tissue regeneration is observed and the engineered tissue can be transplanted into the body and integrated in a functional manner with existing tissue.

Stem cells have the advantage that they can be harvested in small quantities to be expanded in culture and, being autologous, they present the ideal cell type. The use of stem cells, however, relies on being able to identify adult stem cells and isolate them in sufficient quantities. This can be a lengthy procedure, and in older patients, where cell numbers and proliferative capacities may be lower, this can be a limiting factor. However, development of selective monoclonal antibodies that bind to surface antigens will allow the selection of specific subsets of osteoprogenitor cells that can be used in combination with osteoconductive materials for bone formation at specific target sites. Jaiswal et al. reported passing MSCs through 30 population doublings in vitro without loss of osteogenic potential.^[12] The long-term biological effects of the stem cells at the implant site, as well as issues of cell plasticity, remain largely unknown.^[13] Hence, it can be said that stem cell technology is still in its infancy, and the long-term biological interaction of these cells with the scaffold material itself has yet to be fully elucidated.

Stem cells can be subdivided into adult stem cells and embryonic stem cells. Embryonic stem cells (ESCs) are usually isolated from the inner wall of the pre-implantation blastocysts. These embryonic cells in the early developmental stages are more proliferative and are pluripotent due to their indefinite amplification, without the risk of de-differentiation.^[14] ESCs are nonautologous, as they contain a haploid set of chromosomes from a different non-self genetic parent; hence they can be immunoreactive. However, recent advances in gene transfer, MHC manipulation, and nuclear cloning allow the formation of autologous-like ESCs.^[10] The ethics behind their use, however, remains an area of consternation—in particular the use of embryos solely for research purposes.^[15] Stem cells may have the potential to be used to generate entire skeletal tissues without the need for scaffolds.^[16]

Thomson et al. showed that embryonic stem cells (from blastocysts) can proliferate indefinitely and have the ability to become any desired cell type. These cells may induce an immune response, but again, their use presents a number of ethical problems.^[17] Much work remains to be done on the use of stem cells.

Genetically modified cells are the newest cell-based technique for bone regeneration, first described by Lieberman et al.^[18] This technique involves the genetic modification of isolated and expanded cells to produce populations of progenitor cells that over-express certain signaling molecules.^[19] For example, Breitbart et al. showed that via retroviral transduction, cells were able to express BMP-2.^[20] Musgrave and coworkers used direct adenoviral-mediated gene therapy to deliver active BMP-2 and produce bone in skeletal muscles, thus eliminating the problems associated with delivery of BMP-2 to the required site as well as the need for cell expansion.^[21] However, limitations concerning this approach involve several safety issues. The possibility of immunogenicity in vivo is a concern as is the fate of the adenoviral cells. Moreover, the oncogenic biosafety fears must be answered before potential therapeutic models can be established clinically. The major challenge will be to demonstrate bone repair in clinical models without an immunological reaction.^[22] Differentiated osteoblasts are essentially committed mesenchymal cells that have been directed down the osteogenic lineage. This has the advantage of rapid and effective repair of bone defects, as the cells are already differentiated. The main disadvantage is that they have a limited capacity for proliferation, as they are able to perform only a certain number of replication cycles before the problem of de-differentiation arises.

SCAFFOLDS

Central to the formation of new bone tissue is the role of the scaffold; this should either induce the formation

of bone from the surrounding tissue, or act as a carrier or template for implanted bone cells or other osteoinductive agents. An ideal biomaterial should be osteoconductive, osteoinductive, and porous for cellular infiltration. Furthermore, it should be biocompatible, mechanically stable with respect to the native bone, and biodegradable to prevent the presence of foreign material in the body for prolonged periods of time. Other considerations include ease of sterilization without loss of properties.

Biomaterials can be divided in cellular and acellular materials. Acellular materials are usually solid, absorbable fillers that are biodegradable, or porous scaffolds that allow bone ingrowth into the construct. Cellular materials are designed to be seeded with cells. The first acellular material was demineralized bone matrix (DBM); this had both osteoconductive and osteoinductive properties. In early studies inflammatory cells and MSCs were observed to infiltrate the DBM, and within three weeks endochondral ossification was observed.^[23] The role of the scaffold is to provide a template to define the structure of the growing tissue and ensure effective delivery of the osteogenic cells, either singly or in combination with growth factors (Fig. 1). Following implantation the scaffold provides an initial biomechanical support until the cells produce adequate extracellular matrix. Alone, they overcome the need for donor tissue in autologous bone grafts; however, their success in large bone defects is limited as they lack the osteogenic and osteoinductive properties of bone autografts and growth factors. Scaffolds in bioengineered tissues are important not only in terms of structural support for cells, but also for providing signaling cues in influencing cell proliferation and differentiation.

The success of cell-based therapies is very much dependent on the ability of the transplanted cells to be delivered and retained at the site where they are required. Based on the functional requirements of the cells, the method of delivery must be able to encourage

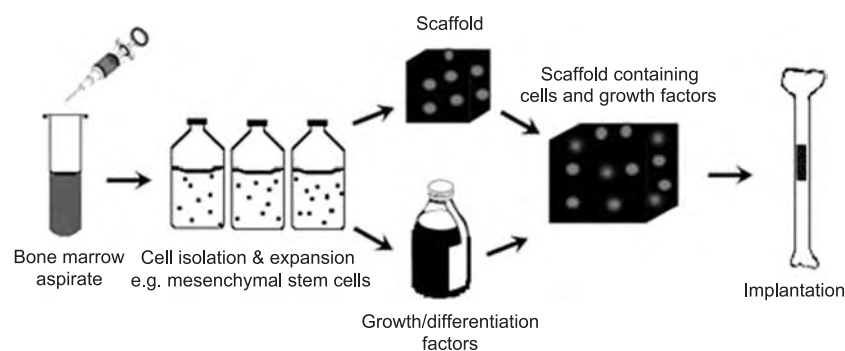


Fig. 1 The optimal tissue-engineered construct requires three essential elements: a suitable cell source, growth and differentiation factors, and a scaffold matrix to support the attachment, migration, and proliferation of these cells.

angiogenesis, be osteoinductive, and be biodegradable so that it can be replaced by new bone formation. Materials commonly used include ceramics, polymers, and composites. More specifically, bone tissue-engineered systems have included demineralized bone tissue, collagen composites, fibrin, and calcium phosphate. In the case of degradable polymers, it is important that scaffold degradation occur at the same rate as tissue regeneration, to ensure that only natural tissue will remain in the body. An ideal scaffold should provide cell anchorage sites, mechanical stability, and structural guidance. Numerous types of scaffolds are being investigated, with biodegradability a common characteristic. Pore size, macro- and microporosity, interconnectivity, and tensile strength are all important factors affecting physical properties as well as cell number and function.

Current approaches to bone tissue engineering involve the use of MSCs to form the three-dimensional (3-D) shape and internal architecture of bone. In order for these exogenously cultured cells to function, they need to be combined with a carrier material before implantation. This will give rise to the 3-D structure where a vascular bed can be established and also act as a supporting scaffold for the transplanted cells to differentiate. This construct with the appropriate biological growth factor can be used to effectively repair defects of critical sizes that would otherwise not heal. Gundle et al. showed that osteoprogenitor cells from humans seeded on the scaffold and given the appropriate osteogenic signaling (e.g., dexamethasone and beta-glycerophosphate) could produce bone tissue *in vivo*.^[24] Bone is made up of an extracellular matrix (ECM) consisting of a composite fibrous network of collagen within a hard matrix of calcium/phosphorus. Bone cells exist within this ECM and are responsible for its generation and remodeling. The physical communication between cells and the ECM directly and indirectly impacts cell shape and function, and these signals are all necessary cues for normal cellular activity. These actions are regulated by protein-based molecules: growth factors. Vacanti et al. reported that for cell colonies to grow into 3-D functional tissues, external cues in the form of mechanical, electrical, structural, and chemical signals mimicking the local ECM are required.^[25]

The similarity of calcium-based ceramics such as hydroxyapatite (HA) or tricalcium phosphate to bone composition and their osteoconductive properties have allowed them to be used extensively as coatings on implants. HA has been successful in increasing implant fixation integrity and has been tried with culture-expanded MSCs.^[26] Although HA is highly biocompatible, it is very brittle and is mechanically too weak to take full compressive loads. It has been shown that the lack of completely interconnected pores, lack of

osteoinductivity, and slow resorption leads to sub-optimal results where only bony ingrowth is observed on the porous surfaces instead of true bone regeneration.^[27]

Pore size is important in determining cellular ingrowth, growth factor release, and vascularization. Opinions on pore size vary; Burg et al. have suggested an optimal pore size between 200 and 400 μm .^[28] Pineda et al. however, have claimed that pore sizes up to 200 μm in polyester membranes achieve the best bone ingrowth.^[29] The porosity of the scaffold has a significant effect on its mechanical integrity, and a compromise is necessary between this and its biological activity. Scaffold macroporosity is a critical factor in cell migration and bone matrix elaboration *in vitro*. In addition, angiogenesis has to occur and could be achieved by incorporating angiogenic factors into the scaffold. Ceramic-based cellular techniques have shown the greatest promise, in particular HA-based constructs, in the fabrication of sintered porous hydroxyapatite (SPHA).^[30] This material consists of repetitive sequences of concavities and is intrinsically osteoinductive without the need for growth factors. SPHA is a so-called smart material. Such materials are gene and cell activating, and it is thought that they attract BMPs and other osteogenic factors intrinsically into the material cavities by adsorption.

Three-dimensional porous structures also offer a desirable scaffold. These can be tailored to achieve a specific pore size (300–400 μm) and high interconnectivity, thus allowing a high degree of cellular ingrowth. These types of systems—for example, microsphere—have the advantage that they can be injected directly into the site and then solidify *in situ*; however, this limits their capacity to be shaped for a specific site.^[31] Other materials, such as polymers, are being tested as potential scaffolds; examples include blends of starch with cellulose acetate poly(lactic acid) and ethylene vinyl alcohol and poly(caprolactone), poly(glycolic acid) (PGA), and poly(lactic-co-glycolic acid) (PLGA). Variable processing methods have been used depending on the desired porosity, degradability, and other properties. This allows them to be tailored to specific applications depending on the mechanical properties and the degradation kinetics required. The one disadvantage of these polymers is their lack of mechanical competence in load-bearing bone applications. Surface-eroding polymers such as poly(ortho-esters) may be advantageous in load-bearing applications as only the surface of these materials degrades, leaving the bulk of the material to provide mechanical strength and reducing the risk of implant failure.^[32] Degradation kinetics applies to all biodegradable biomaterials and leads to a number of changes in both the polymer surface and the local environment.^[33] A further consideration of degradable materials is the constant change in surface topography and energy, both of

which have an effect on cell attachment and proliferation. For example, nanotopography has been shown to improve cell proliferation rates on polymer-based materials, whereas surface energy has been shown to effect cell adhesion to surfaces and even their orientation.^[34]

A more adaptive scaffold could be achieved if it could be made responsive to the mechanical environment in which it is placed. Work in the literature has indicated that mechanoreceptive stimulation of cells can be achieved by incorporating slow-release agonists in the polymer scaffolds. Biomimetic materials aim to provide a biological scaffold to promote integration of host bone with implanted cells and growth factors. For example, it has been shown *in vitro* that mechanically active scaffolds, where a voltage-gated calcium channel inhibitor is integrated into PLLA with a cyclically loaded environment, can increase collagen type I expression and increase matrix production.^[35] This work is in an early stage and has not been performed *in vivo*, where the mechanical stimuli will be more randomly distributed across the material.

ECM would seem to be the most natural scaffold, since its composition adapts depending on the site and physiological requirements of the host tissue. Furthermore, it contains all the necessary proteins for cell adhesion (fibronectin, laminin), proliferation, and differentiation, as well as those required for structural support such as collagen and GAGs. *In vivo*, ECM has been shown to be rapidly biodegradable and transient in terms of functionality.^[36] The major disadvantage of ECM, however, is its limited supply.

A greater understanding of biomineralization has initiated developments in biomimetic materials fabricated according to biological principles and processes of self-assembly and self-organization. Biomimetic materials are being developed whereby the desirable properties of natural, biodegradable, and ceramic-based techniques are combined. For example, arginine-glycine-aspartic acid (RGD) sequence peptides have been incorporated onto scaffold surfaces to increase cell adhesion, proliferation, and biocompatibility.^[37] RGD is the minimal amino acid sequence needed to promote cell adhesion and is found in many ECM proteins. Hrkach et al. have prepared comb-like graft copolymers with lysine or aspartic acid side chains, which have demonstrated increased cell attachment compared to unmodified surfaces.^[38] Another promising approach has been the use of supercritical fluid technology in the development of porous biodegradable scaffolds.^[39] This technology does not involve solvents or thermal processing, thus allowing growth factors to be incorporated into the scaffold at construction. Other bioactive particles such as HA could also be incorporated; this would overcome the problem of mechanical competence.

The synthesis of complex inorganic forms that mimic natural structures offers exciting avenues for the chemical construction of macrostructures and a new generation of biologically and structurally related bone analogues for tissue engineering.^[40] Although these appear to offer attractive possibilities, biomimetic scaffolds require clinical evaluation, with little known about their immunoreactivity, biocompatibility, and effects on stem cell fate and host osseointegration of such constructs.

SIGNALING MOLECULES

Growth factors (GFs) are proteins secreted by cells that act as signaling molecules on the appropriate target cell to carry out a specific function. They act as a cellular communications network and influence critical functions such as cell proliferation, matrix production, and differentiation of tissues. They have been shown to play a significant role in bone and cartilage formation, fracture healing, and the repair of other musculoskeletal tissues. The advent of recombinant proteins has resulted in considerable interest in their use as therapeutic agents in the treatment of skeletal injuries and disease. A GF may have effects on multiple cell types and may induce several different cellular responses in a variety of tissues. Once a GF binds to a target cell receptor, it induces a ligand-receptor interaction. This induces an intracellular signal transduction system to the nucleus and results in a biological response. These interactions can be very specific, with a specific GF binding to a single cellular receptor, or complex, with one or more GFs binding to one or more receptors in order to produce an effect.

A number of GFs have been shown to be expressed during fracture healing; these include transforming growth factor- β (TGF- β), bone morphogenetic proteins (BMPs), fibroblast growth factor (FGF), platelet-derived growth factor (PDGF), and insulin-like growth factor (IGF). Osteoinduction is therefore mediated by numerous growth factors, and on this basis, it is thought that these GFs may have potential use as therapeutic agents in bone healing and tissue engineering. Hauschka et al. showed that osteogenesis is in part due to the combination of actions of several growth factors acting at specific stages in the different cells.^[41] Bone is rich in many growth-stimulating factors, but the most important group is the TGF- β superfamily, of which the BMPs are the most significant. There are 15 BMPs (BMP-2–BMP-16) which are known to regulate bone formation via stem cell recruitment, proliferation, and differentiation. The biological action of BMPs are mediated via specific BMP receptors. There are two types of BMP receptors, type I and type II, which are serine/threonine protein kinases.

These kinases are enzymes that phosphorylate proteins called Smads and activate them. The active Smads are then translocated to the nucleus, where they participate in the transcriptional regulation of the expression of genes involved in cartilage and bone formation.^[42]

BMPs have been extensively studied and used in clinical trials with variable success depending on the mode of delivery.^[43,44] Human BMPs are difficult to obtain and purify from bone donors, and BMP-7 (also known as OP-1) and BMP-2 are currently synthesized via recombinant technology. BMPs have already been successful in trials in vivo. It has been shown using native human BMP that they have potential in the treatment of established nonunions and spinal fusions.^[45] The importance of BMPs in fracture healing was demonstrated by Bostrom et al. when they were localized to fracture callus.^[46] Einhorn et al. showed that injection of rhBMP-2 into a standard fresh-fracture model can accelerate healing, which was achieved two weeks earlier than in the controls.^[47] Evidence for its proposed use in tissue-engineered bone was provided by the study of Wang et al., who showed that BMPs caused commitment and differentiation of multipotential stem cells into osteoprogenitor cells.^[48] Wozney et al. was also able to show new bone formation using BMPs.^[49] Extensive work by Ripamonti using BMP-7 has shown that it elicits heterotopic bone formation and complete healing of 25 mm-diameter critical-sized defects in primates in 90 days.^[50] In another study, Johnson and Urist used BMPs to lengthen femoral nonunion fracture.^[51]

Recombinant BMP-2 has been extensively used for in vitro and in vivo studies and has been shown to induce cell differentiation and form endochondral bone in ectopic and heterotopic locations.^[52] Although BMPs are at the forefront of reconstructive orthopedic surgery,^[53] there are difficulties in integrating them into a drug delivery system. Purified BMPs are hydrophilic and, when incorporated in the body, tend to disperse rather than act locally. Hence to slow the release down, they need to be complexed to a carrier. This adds a further problem in that proteins are immunogenic.

The osteoinductive capacity of BMPs has been extensively demonstrated in preclinical studies, and the efficacy of BMP-2 and BMP-7 for the treatment of orthopedic patients is now being evaluated clinically.^[54]

Much remains to be known about the cascade of growth factors necessary for tissue regeneration. Growth factors are species and dose dependent and require careful targeting within the host environment. Large bone defects can be repaired only if scaffolds are able to cause ingrowth of host tissue into the structure and become fully integrated with the adjoining tissue. A major problem, however, is that these GFs require a carrier that will allow their release in a controlled manner. Kaplan et al. showed that too low a dose

resulted in poor bone formation with reduced mechanical strength, while high doses may inhibit osteogenesis or lead to bone growth at sites outside the graft boundaries.^[55] These factors are easily diffusible and have a short half-life.

Scaffold-based drug delivery is important, with numerous promising studies well advanced. However, it is limited by conflicting requirements of scaffold and drug delivery needs. For example, microgeometry is important in BMP-induced osteogenesis.^[56] A large pore size promotes vascularization and osteogenesis but also causes greater drug release; the hydrophobicity of a scaffold ensures gradual biodegradation and mechanical competence but binds the drugs less readily. Many novel drug delivery systems which can control drug release kinetic externally are being developed. Smart polymers utilizing hydrogels, alginate, and gelatin microspheres, which are responsive to thermal, mechanical (ultrasound), and pH stimuli, are being investigated.

CONCLUSION

Orthopedic surgery has evolved over the last 25 years from the age of resection, reconstruction, and replacement to the age of regeneration. Bone tissue engineering provides a means of evolving therapies using the body's own cells. The requirements include a scaffold that is immunologically inert and allows the incorporation and release of osteoinductive proteins. Furthermore, it should allow cellular ingrowth with angiogenesis occurring in a timely fashion within the scaffold construct in order to restore proper function. The complex problem of the scaffold is one of the main limiting factors, and the interaction of stem cells with biomaterials still requires a greater understanding. Although some progress has been made in this field, bone tissue engineering is still in its infancy. A basic understanding of the spatial and temporal distribution of cells and growth factors necessary for osteogenesis remains to be developed. Modern technologies like molecular cloning and gene therapy may further enhance the ability of BMPs to play an important role in the clinical practice of the 21st century. The clinical application of bone-engineered constructs has the potential to greatly improve the treatment of conditions requiring bone repair. Enhancing our understanding of these factors will assist in the development and use of engineered systems in orthopedic applications.

ARTICLES OF FURTHER INTEREST

Biologic and Synthetic Apatites; Biphasic Calcium Phosphate (BCP) Bioceramics: Chemical, Physical,

and Biological Properties; Bone Remodeling; Orthodontic Wires; Osteogenic Progenitor Cells of Bone; Poly(Glycolic Acid); Poly(lactic acid)s; Stem Cells; Tissue Engineering Scaffolds

REFERENCES

- Vaccaro, A.R. The role of the osteoconductive scaffold in synthetic bone graft. *Orthopedics* **2002**, *25* (5), s571–s578.
- Tomin, E.; Lane, J.; Nakamichi, K.; Hsu, J.; Schneider, K.; Weiland, A. Performance of molded vascularized bone grafts induced by BMP-2 and marrow in a rat segmental defect model. *Trans. Orthop. Res. Soc.* **1999**, *24*, 621.
- Beresford, J.N. Osteogenic stem cells and the stromal system of bone and marrow. *Clin. Orthop.* **1989**, *240*, 270–280.
- Connolly, J.F. Injectable bone marrow preparations to stimulate osteogenic repair. *Clin. Orthop.* **1995**, *313*, 8–18.
- Paley, D.; Young, M.C.; Wiley, A.M.; Fornasier, V.L.; Jackson, R.W. Percutaneous bone marrow grafting of fractures and bony defects. *Clin. Orthop.* **1986**, *208*, 300–312.
- Liang, C.T.; Barnes, J.; Sedor, J.G.; Quartuccio, H.A.; Bolander, M.; Jeffrey, J.J.; Rodan, G.A. Impaired bone activity in aged rats: Alterations at the cellular and molecular levels. *Bone* **1992**, *13* (6), 435–441.
- Quarto, R.; Thomas, D.; Liang, C.T. Bone progenitor cell deficits and the age-associated decline in bone repair capacity. *Calcif. Tissue Int.* **1995**, *56* (2), 123–129.
- Kuehnle, I.; Goodell, M.A. The therapeutic potential of stem cells from adults. *BMJ* **2002**, *325* (7360), 372–376.
- Bianco, P.; Riminucci, M.; Kuznetsov, S.; Robey, P.G. Multipotential cells in the bone marrow stroma: Regulation in the context of organ physiology. *Crit. Rev. Eukaryot. Gene Expr.* **1999**, *9* (2), 159–173.
- Vats, A.; Tolley, N.S.; Polak, J.M.; Buttery, L.D. Stem cells: Sources and applications. *Clin. Otolaryngol.* **2002**, *27* (4), 227–232.
- Haynesworth, S.E.; Carrino, D.A.; Caplan, A.I. Characterization of the core protein of the large chondroitin sulfate proteoglycan synthesized by chondrocytes in chick limb bud cell cultures. *J. Biol. Chem.* **1987**, *262* (22), 10574–10581.
- Jaiswal, N.; Haynesworth, S.E.; Caplan, A.I.; Bruder, S.P. Osteogenic differentiation of purified, culture-expanded human mesenchymal stem cells in vitro. *J. Cell. Biochem.* **1997**, *64* (2), 295–312.
- Rose, F.R.; Oreffo, R.O. Bone Tissue Engineering: Hope vs. hype. *Biochem. Biophys. Res. Commun.* **2002**, *292* (1), 1–7.
- Keller, G.M. In vitro differentiation of embryonic stem cells. *Curr. Opin. Cell. Biol.* **1995**, *7* (6), 862–869.
- McLaren, A. Ethical and social considerations of stem cell research. *Nature* **2001**, *414* (6859), 129–131.
- Oreffo, R.O.; Triffitt, J.T. Future potentials for using osteogenic stem cells and biomaterials in orthopedics. *Bone* **1999**, *25* (2), 5S–9S.
- Thomson, J.A.; Itskovitz-Eldor, J.; Shapiro, S.S.; Waknitz, M.A.; Swiergiel, J.J.; Marshall, V.S.; Jones, J.M. Embryonic stem cell lines derived from human blastocysts. *Science* **1998**, *282* (5391), 1145–1147.
- Lieberman, J.R.; Daluiski, A.; Stevenson, S.; Wu, L.; McAllister, P.; Lee, Y.P.; Kabo, J.M.; Finerman, G.A.; Berk, A.J.; Witte, O.N. The effect of regional Gene Therapy with bone morphogenetic protein-2-producing bone-marrow cells on the repair of segmental femoral defects in rats. *J. Bone Jt. Surg., Am.* **1999**, *81* (7), 905–917.
- Scaduto, A.A.; Lieberman, J.R. Gene Therapy for osteoinduction. *Orthop. Clin. North Am.* **1999**, *30* (4), 625–633.
- Breitbart, A.S.; Grande, D.A.; Mason, J.M.; Barcia, M.; James, T.; Grant, R.T. Gene-enhanced tissue engineering: Applications for bone healing using cultured periosteal cells transduced retrovirally with the BMP-7 gene. *Ann. Plast. Surg.* **1999**, *42* (5), 488–495.
- Musgrave, D.S.; Bosch, P.; Ghivizzani, S.; Robbins, P.D.; Evans, C.H.; Huard, J. Adenovirus-mediated direct Gene Therapy with bone morphogenetic protein-2 produces bone. *Bone* **1999**, *24* (6), 541–547.
- Asahara, T.; Kalka, C.; Isner, J.M. Stem cell therapy and gene transfer for regeneration. *Gene Ther.* **2000**, *7* (6), 451–457.
- Urist, M.R.; Silverman, B.F.; Buring, K.; Dubuc, F.L.; Rosenberg, J.M. The bone induction principle. *Clin. Orthop.* **1967**, *53*, 243–283.
- Gundle, R.; Joyner, C.J.; Triffitt, J.T. Human bone tissue formation in diffusion chamber culture in vivo by bone-derived cells and marrow stromal fibroblastic cells. *Bone* **1995**, *16* (6), 597–601.
- Vacanti, C.A.; Bonassar, L.J. An overview of Tissue Engineered bone. *Clin. Orthop.* **1999**, *367*, S375–S381.
- Bruder, S.P.; Fink, D.J.; Caplan, A.I. Mesenchymal stem cells in bone development, bone repair, and skeletal regeneration therapy. *J. Cell. Biochem.* **1994**, *56* (3), 283–294.
- Petite, H.; Viateau, V.; Bensaid, W.; Meunier, A.; de Pollak, C.; Bourguignon, M.; Oudina, K.; Sedel, L.; Guillemain, G. Tissue-engineered bone regeneration. *Nat. Biotechnol.* **2000**, *18* (9), 959–963.
- Burg, K.J.; Porter, S.; Kellam, J.F. Biomaterial developments for bone Tissue Engineering. *Biomaterials* **2000**, *21* (23), 2347–2359.
- Pineda, L.M.; Busing, M.; Meinig, R.P.; Gogolewski, S. Bone regeneration with resorbable polymeric membranes. III. Effect of poly(L-lactide) membrane pore size on the bone healing process in large defects. *J. Biomed. Mater. Res.* **1996**, *31* (3), 385–394.
- Ripamonti, U.; Crooks, J.; Rueger, D.C. Induction of bone formation by recombinant human osteogenic protein-1 and sintered porous hydroxyapatite in adult primates. *Plast. Reconstr. Surg.* **2001**, *107* (4), 977–988.
- Borden, M.; Attawia, M.; Khan, Y.; Laurencin, C.T. Tissue Engineered microsphere-based matrices for bone repair: Design and evaluation. *Biomaterials* **2002**, *23* (2), 551–559.

32. Andriano, K.P.; Tabata, Y.; Ikada, Y.; Heller, J. In vitro and in vivo comparison of bulk and surface hydrolysis in absorbable polymer scaffolds for tissue engineering. *J. Biomed. Mater. Res.* **1999**, *48* (5), 602–612.
33. Hollinger, J.O.; Battistone, G.C. Biodegradable bone repair materials. Synthetic polymers and ceramics. *Clin. Orthop.* **1986**, *207*, 290–305.
34. Dalby, M.J.; Riehle, M.O.; Johnstone, H.; Affrossman, S.; Curtis, A.S. In vitro reaction of endothelial cells to polymer demixed nanotopography. *Biomaterials* **2002**, *23* (14), 2945–2954.
35. Yang, Y.; Magnay, J.L.; Cooling, L.; El, H.A. Development of a 'mechano-active' scaffold for tissue engineering. *Biomaterials* **2002**, *23* (10), 2119–2126.
36. Badylak, S.F.; Kropp, B.; McPherson, T.; Liang, H.; Snyder, P.W. Small intestinal submucosa: A rapidly resorbed bioscaffold for augmentation cystoplasty in a dog model. *Tissue Eng.* **1998**, *4* (4), 379–387.
37. Quirk, R.A.; Chan, W.C.; Davies, M.C.; Tendler, S.J.; Shakesheff, K.M. Poly(L-lysine)-GRGDS as a biomimetic surface modifier for poly(lactic acid). *Biomaterials* **2001**, *22* (8), 865–872.
38. Hrkach, J.S.; Ou, J.; Lotan, N.; Langer, R. Synthesis of poly(L-lactic acid-co-L-lysine) graft copolymers. *Macromolecules* **1995**, *28* (13), 4736–4739.
39. Cooper, A.I. Polymer synthesis and processing using supercritical carbon dioxide. *J. Mater. Chem.* **2000**, *10* (2), 207–234.
40. Green, D.; Walsh, D.; Mann, S.; Oreffo, R.O. The potential of biomimesis in bone Tissue Engineering: Lessons from the design and synthesis of invertebrate skeletons. *Bone* **2002**, *30* (6), 810–815.
41. Hauschka, P.V.; Chen, T.L.; Mavrakos, A.E. Polypeptide growth factors in bone matrix. *Ciba Found. Symp.* **1988**, *136*, 207–225.
42. Reddi, A.H. Bone morphogenetic proteins: From basic science to clinical applications. *J. Bone Jt. Surg., Am.* **2001**, *83A* (Suppl. 1 (Pt. 1)), S1–S6.
43. Centrella, M.; Horowitz, M.C.; Wozney, J.M.; McCarthy, T.L. Transforming growth factor-beta gene family members and bone. *Endocr. Rev.* **1994**, *15* (1), 27–39.
44. Wozney, J.M. Bone morphogenetic proteins. *Prog. Growth Factor Res.* **1989**, *1* (4), 267–280.
45. Johnson, E.E.; Urist, M.R.; Finerman, G.A. Bone morphogenetic protein augmentation grafting of resistant femoral nonunions. A preliminary report. *Clin. Orthop.* **1988**, *230*, 257–265.
46. Bostrom, M.P.; Lane, J.M.; Berberian, W.S.; Missri, A.A.; Tomin, E.; Weiland, A.; Doty, S.B.; Glaser, D.; Rosen, V.M. Immunolocalization and expression of bone morphogenetic proteins 2 and 4 in fracture healing. *J. Orthop. Res.* **1995**, *13* (3), 357–367.
47. Einhorn, T.A.; Trippel, S.B. Growth factor treatment of fractures. *Instr. Course Lect.* **1997**, *46*, 483–486.
48. Wang, E.A.; Israel, D.I.; Kelly, S.; Luxenberg, D.P. Bone morphogenetic protein-2 causes commitment and differentiation in C3H10T1/2 and 3T3 cells. *Growth Factors* **1993**, *9* (1), 57–71.
49. Wozney, J.M. Overview of bone morphogenetic proteins. *Spine* **2002**, *27* (16 Suppl. 1), S2–S8.
50. Ripamonti, U.; Ramoshebi, L.N.; Matsaba, T.; Tasker, J.; Crooks, J.; Teare, J. Bone induction by BMPs/OPs and related family members in primates. *J. Bone Jt. Surg., Am.* **2001**, *83-A* (Suppl. 1 (Pt. 2)), S116–S127.
51. Johnson, E.E.; Urist, M.R. One-stage lengthening of femoral nonunion augmented with human bone morphogenetic protein. *Clin. Orthop.* **1998**, *347*, 105–116.
52. Aspenberg, P.; Turek, T. BMP-2 for intramuscular bone induction: Effect in squirrel monkeys is dependent on implantation site. *Acta Orthop. Scand.* **1996**, *67* (1), 3–6.
53. Einhorn, T.A. Problems with delayed and impaired fracture healing remain a challenge to the orthopedic trauma surgeon. *Orthop. Traumatol.* **1997**, *11* (4), 243.
54. Einhorn, T.A.; Lee, C.A. Bone regeneration: New findings and potential clinical applications. *J. Am. Acad. Orthop. Surg.* **2001**, *9* (3), 157–165.
55. Kaplan, F.S.; Tabas, J.A.; Zasloff, M.A. Fibrodysplasia ossificans progressiva: A clue from the fly? *Calcif. Tissue Int.* **1990**, *47* (2), 117–125.
56. Tsuruga, E.; Takita, H.; Itoh, H.; Wakisaka, Y.; Kuboki, Y. Pore size of porous hydroxyapatite as the cell-substratum controls BMP-induced osteogenesis. *J. Biochem. (Tokyo)* **1997**, *121* (2), 317–324.

Tissue Engineering of Bone Marrow

Athanassios Mantalaris

Department of Chemical Engineering & Chemical Technology, Imperial College London, London, United Kingdom

Nicki Panoskaltsis

Department of Haematology, Imperial College London, London, Harrow, United Kingdom

J. H. David Wu

Departments of Chemical Engineering, Biomedical Engineering, and Microbiology & Immunology, University of Rochester, Rochester, New York, U.S.A.

INTRODUCTION

Hematopoietic stem cell cultures have generated intense clinical, commercial, and academic interest due to their potential applications in cell-based therapies, such as genetic, cellular, tissue engineering, transfusion support, and immunotherapy. Furthermore, the increasing therapeutic applications of mesenchymal stem cells (MSCs) in gene therapy, bone marrow (BM) transplantation, and tissue engineering exploit current knowledge in molecular biology and biomaterial science^[1] and have greatly expanded the potential of therapeutic use of BM-based stem cells. However, numerous problems remain to be addressed and new methodologies are needed for the well-controlled and reproducible culture of clinically relevant numbers of hematopoietic stem cells (HSCs), MSCs, and their derivatives. This article will focus on the BM micro-environment, its complex three-dimensional (3-D) structure, and its role in hematopoiesis. In addition, we will review the culture parameters critical for the *ex vivo* hematopoietic cell cultures, aiming at developing clinically desirable “designer transplants.”^[2]

HEMATOPOIESIS AND BONE MARROW

Hematopoiesis is a three-stage continual process consisting of the proliferation, commitment, and differentiation of the hematopoietic cells. It takes place in the BM, which is located in the medullary cavity of bone. In normal adults, the BM produces approximately 2.5 billion red blood cells, 2.5 billion platelets, and 1 billion granulocytes per kilogram of body weight per day, the rate of production adjusting to actual needs and varying from nearly zero to many times normal.^[3] All blood cells are derived from a common precursor, the HSC (Fig. 1). Self-renewal of the HSC ensures maintenance of the stem cell pool and hence continuous blood formation throughout life. Commitment of the HSC to differentiation is the first

step in a series of cell divisions through which specialization and differentiation along a single lineage occur resulting in the production of mature blood cells. Within the three stages of hematopoiesis, extensive expansion and maturation occur; it has been proposed that a single HSC is capable of more than 50 cell divisions to generate up to 10^{15} cells.^[4]

Bone marrow is a complex tissue with an elaborate architecture that contains the hematopoietic cells and the supporting stroma consisting of reticular cells, osteocytes, adipocytes, the vascular endothelium, and the extracellular matrix (ECM).^[5] Hematopoiesis is regulated by the hematopoietic inductive micro-environment (HIM), a concept that was introduced in the 1960s to explain the restriction of adult hematopoiesis to the BM.^[6] The HIM (Fig. 2) consists of the stromal cells, which through their intimate physical contact with the hematopoietic cells, the ECM, and the growth factors (GF) they secrete, create this unique 3-D structural and biochemical micro-environment that regulates, through positive and negative factors, the survival, proliferation, and differentiation of the HSCs. In vivo the HIM supports and regulates hematopoiesis, which takes place in spaces surrounded by BM sinusoids that form a 3-D spatially designed microvascular network.^[7] The stroma, composed of endothelial cells that line the marrow sinuses, adventitial reticular cells, macrophages, and adipocytes,^[8] liberates GFs in the immediate environment.^[9] Marrow stromal cells, originally examined because of their critical role in the formation of the HIM, have come to center stage with the recognition that they are the source of the stem/progenitor cells of skeletal tissues. Recent data have also suggested an unexpected plasticity of marrow stromal cells in that they may be able to differentiate into neural tissue or muscle.^[10]

The ECM is a rich heterogeneous carpet of a broad class of macromolecules such as proteoglycans or glycosaminoglycans, adhesive proteins such as fibronectin and laminin, and collagen (types I, III, and IV).^[11]

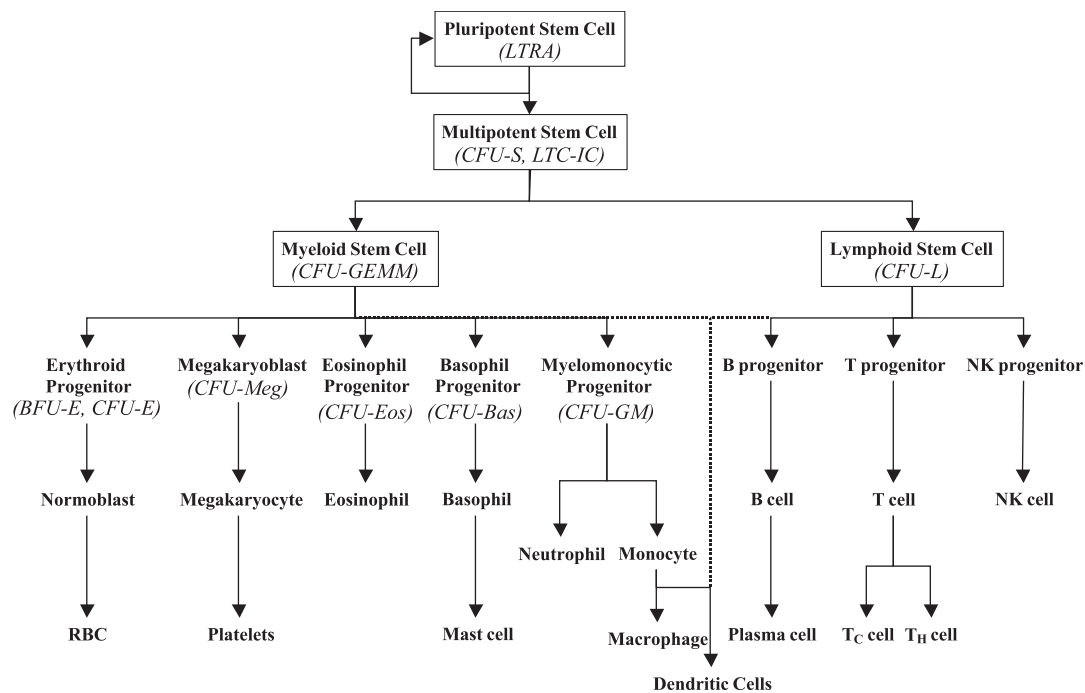


Fig. 1 The hematopoietic system hierarchy, including stem and progenitor cell assays. Hematopoiesis is a three-stage process, which includes proliferation, commitment, and differentiation. LTRA: long-term repopulation assay; CFU-S: colony-forming unit-spleen; LTC-IC: long-term culture-initiation cell; CFU-GEMM: CFU-granulocyte/erythrocyte/macrophage/megakaryocyte; CFU-L: CFU-lymphocyte; CFU-Meg: CFU-megakaryocyte; CFU-Eos: CFU-eosinophil; CFU-Bas: CFU-basophil; CFU-GM: CFU-granulocyte/macrophage.

Collagen fibers of types I and III are produced by fibroblasts and provide the structural framework for the HIM.^[3] Laminin and fibronectin contain the universal attachment peptide, RGD, and bind hematopoietic cells to collagen and GFs. Proteoglycans may also bind GFs and modulate hematopoiesis.^[3] The ECM plays an important role in hematopoietic regulation by localizing cytokines and hematopoietic cells to specific 3-D microenvironmental niches,^[3] thus establishing different HIMs through binding of

particular cell lineages to distinct ECM components^[12-14] and providing scaffolding that lends support and cohesiveness for the marrow structure. For example, the ECM modulates cell maturation by selectively binding to erythroid^[14] and granulocyte precursors.^[13] The ECM may also determine cell adhesion and homing.^[8] For example, CD34⁺ cells have multiple adhesion receptors^[3] enabling them to attach to cellular and ECM components within the marrow.^[15] These receptor-ligand interactions permit close cell-cell or

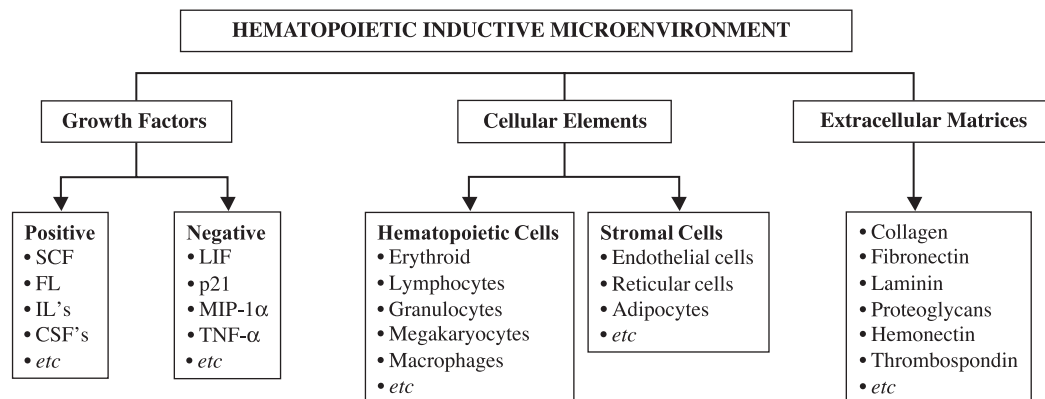


Fig. 2 The components of the hematopoietic inductive microenvironment. SCF: stem cell factor; FL: flt-3 ligand; ILs: interleukins; CSFs: colony-stimulating factors; LIF: leukemia inhibitory factor; p21: protein 21; MIP-1 α : macrophage inflammatory protein-1 α ; TNF- α : tumor necrosis factor- α .



cell-matrix interactions and are essential both for the retention of stem cells and for their maturation and development.^[9,16]

A model for hematopoietic regulation has been proposed according to which rare niches within the HIM specifically promote the self-renewal of HSCs.^[17] Space within each niche acts as the limiting factor for stem cell renewal. When sufficient numbers of self-renewing divisions have occurred and the niche is filled with stem cells, the excess cells are pushed into a neighboring niche, which fosters the production of differentiated progeny by these cells. Eventually, mature blood cells exit the marrow space and enter the blood circulation. These selective microenvironments may be due to the presence and/or absence of positive and negative GFs, as well as to concentration gradients created by membrane-bound cytokines that may instruct HSCs to differentiate or remain quiescent.^[18,19] The cellular and molecular natures of each niche remain to be determined. Development of culture systems recapitulating these specialized microenvironments in vitro may help verify the niche concept and reveal the niche requirements.^[20]

CULTURE PARAMETERS

The BM culture system supports in vitro hematopoiesis, which can be influenced by culture parameters including pH, local oxygen concentration, the composition and concentration of GFs, the supply of nutrients, and the cellular and molecular surrounding of the cells.^[21] Due to the heterogeneous cell population with a wide range of maturation stages, defining optimal culture parameters for ex vivo expansion of hematopoietic stem and progenitor cells is challenging. The effort can be further complicated by the complex kinetics, transient nature, and intricate interactions among culture parameters.^[22]

Culture Media

Traditional media for hematopoietic cultures contain xenogenic supplements such as fetal calf and horse sera (at a total of 25% v/v) or bovine serum albumin, usually supplemented with a cocktail of cytokines.^[23–25] Serum has a number of potential advantages for culture performance: serving as a source of essential nutrients, hormones, and GFs, and protecting the cell against damage.^[26] The presence of serum in hematopoietic culture media has also been shown to influence the induction or enhancement of various cytokines^[27,28] and to contribute to the growth and survival of hematopoietic progenitor cells.^[25] However, serum contains a wide array of proteins that may interact

with cell signaling components and alter the physiological/physiochemical properties of the culture. Serum supplementation in the cultures of CD34⁺ BM cells facilitates greater expansion of granulocytic and monocytic lineages due to the presence of transforming growth factor- β (TGF- β),^[29] which inhibits the expansion of erythroid and megakaryocytic lineages.^[30–32] Furthermore, serum additives are problematic for clinical applications due to the risk of contamination by infectious agents and foreign proteins and the lack of well-defined and standardized media composition.^[33] For these reasons, low concentrations (1–2%) of autologous serum or plasma have been used to substitute for animal sera for supporting the expansion of hematopoietic cells in culture.^[34,35]

A number of serum-free media have been developed, which, when supplemented with the appropriate cytokines, support the ex vivo expansion of HSCs and hematopoietic progenitor cells (HPCs).^[33,36–41] In a direct comparison, serum-free media enhanced the clonogenic potential of human BM CD34⁺ cells whereas serum-containing media enhanced total cell expansion.^[41] Serum-free media are generally considered suitable for stroma-free cultures because serum aides the growth and maintenance of the stromal layer,^[42] and serum-free media do not support the attachment of an adherent cell layer to the surface of the tissue culture dish, even after pretreatment of the culture surface with poly-D-lysine, collagen, or fibronectin.^[36]

Growth Factors

Hematopoiesis is regulated by a large number of GFs (cytokines), primarily the interleukins (ILs) and colony-stimulating factors (CSF), with multiple and overlapping functions (Fig. 3).^[43] HSCs and HPCs are thought to express most cytokine receptors, yet they are selective in the cytokines to which they respond.^[44] Growth factors have been used to stimulate or potentiate the proliferation of stem or progenitor cells in the various lineages (Table 1). A key feature of a typical hematopoietic culture medium is the combination of cytokines it delivers to the cells and how the concentrations of these cytokines are maintained over time, both influencing the cumulative response obtained.^[45]

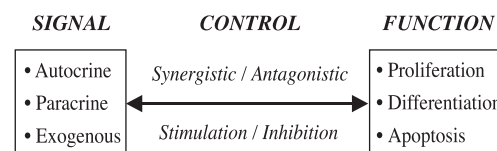


Fig. 3 Regulation of hematopoiesis by cytokine networks.

Table 1 Examples of cytokines with stimulating action on hematopoietic cell proliferation

Cell type	Cytokines
B-cell	IL-7, SCF, IL-2, IL-5, FL
T-cell	IL-2, SCF, FL
NK-cell	IL-2, IL-12
Erythroid	Epo, SCF, IL-3
Megakaryocyte	TPO, CSF, LIF, Epo, IL3, IL-6, IL-11
Neutrophil	G-CSF, SCF, GM-CSF, M-CSF, IL-3, IL-6
Eosinophil	IL-5, GM-CSF
Basophil	IL-3, SCF
Macrophage	M-CSF, GM-CSF, IL-3

Abbreviations: NK-cell: natural killer cell; IL: interleukin; SCF, stem cell factor; FL: flt-3 ligand; Epo: erythropoietin; TPO: thrombopoietin; CSF: colony-stimulating factor; LIF: leukemia inhibitory factor; G-CSF: granulocyte-colony stimulating factor; GM-CSF: granulocyte/macrophage-colony stimulating factor; M-CSF: macrophage-colony stimulating factor.

Cytokine receptors belong to three main families: the tyrosine kinase receptors, the hematopoietic GF (or cytokine) receptors, and the gp130 receptor.^[45] c-kit, the receptor for stem cell factor (SCF),^[46] and flk-2/flt-3, the receptor for flt-3 ligand (FL),^[47] are examples of the tyrosine kinase receptors. Both are expressed on very primitive hematopoietic cells and are implicated in the regulation of early stages of hematopoiesis.^[48–50] Ligands for the cytokine receptor family include IL-3 and thrombopoietin (TPO), both of which are active on primitive progenitors. IL-3 is required, along with SCF and flt-3 ligand, to maximize the expansion of human long-term culture-initiating cells (LTC-ICs) from CD34⁺/CD38[−] cells isolated from BM.^[51] TPO supports the proliferation and long-term maintenance of primitive hematopoietic progenitors and repopulating HSCs when used in combination with other stimulatory cytokines.^[52,53] The gp130 receptor family includes receptors for IL-6, IL-11, oncostatin-M, and leukemia inhibitory factor (LIF). IL-6 has direct proliferative effects on hematopoietic cells and interacts synergistically with other GFs to stimulate myeloid proliferation.^[54] IL-11 is produced by stromal cells and has been found, in synergy with SCF and flt-3 ligand, to stimulate the ex vivo expansion of adult murine stem cells.^[55] The addition of LIF to purified CD34⁺/Thy-1⁺ cells on a murine BM-derived stromal cell line caused the expansion of these cells, which in synergy with other GFs such as IL-3, IL-6, SCF, and granulocyte/macrophage-colony-stimulating factor (GM-CSF) resulted in a 150-fold expansion of cells retaining the CD34⁺/Thy-1⁺ phenotype.^[56]

Although most studies have focused on hematopoietic-related cytokines, novel factors, not originally identified as a result of their effects on hematopoietic cells, have been investigated. Bone morphogenetic proteins (BMPs) can modulate the proliferation and differentiation of primitive human hematopoietic cells (CD34⁺/CD38[−]/Lin[−]).^[57] Recently, the addition of the Notch ligand, Jagged-1, to cultures of primitive progenitor cells induced the survival and expansion of repopulating HSCs.^[58] Aside from growth-promoting cytokines, a number of growth-inhibitor cytokines, such as TGF- β , macrophage inflammatory protein-1 α (MIP-1 α), and tumor necrosis factor- α (TNF- α), have been shown to affect hematopoiesis.^[44] Although the role of the growth-inhibitory cytokines in vivo is not known, in vitro production of inhibitory proteins may be one reason for the transient expansion observed in many cultures.^[2] In long-term cultures, TGF- β is responsible for the quiescent state of the progenitor five to seven days after the last medium exchange.^[59] Similarly, addition of MIP-1 α to cultures prevents cells from cycling.^[60] Hence, addition of MIP-1 α and platelet factor-4 to IL-3-containing cultures of human hematopoietic cells resulted in the preservation of LTC-ICs for at least eight weeks.^[61]

Inoculum Density

Long-term hematopoietic cultures have demonstrated the influence of seeding density, GF supplements, and the rate of medium exchange on progenitor output.^[62] Consideration needs to be given not only to cell density, but also to the composition of stem, progenitor, and accessory cells, as well as the combination and concentration of GFs in the medium.^[63] A number of reports have examined the effect of inoculum density on the ex vivo expansion of human hematopoietic cells using static cultures. Koller et al.^[64] demonstrated that high seeding density cultures of BM mononuclear cells (MNCs) (e.g., 5×10^5 cells/well) resulted in higher total cell and colony-forming cell (CFC) numbers, whereas lower seeding densities (e.g., 5×10^4 cells/well) led to higher cell expansion and greater depletion of CFCs. Similarly, Poloni et al.^[65] demonstrated that decreasing the seeding density from 3×10^4 to 1.5×10^3 cells/ml enhanced total cell and progenitor expansion by up to 20-fold and 3-fold, respectively, while maintaining the LTC-ICs. The effect of preformed stroma on culture performance and its relation to seeding density was also examined.^[64] It was demonstrated that when BM MNCs were cultured at sufficiently high seeding densities (e.g., 5×10^5 cells/well), the culture became stroma-independent in terms of LTC-IC maintenance and expansion. On the other hand, when the seeding density of MNCs was lowered

T

(e.g., 5×10^4 cells/well), their behavior greatly resembled cultures of CD34-enriched and stroma-free cells, with greater variability in culture output and dependence on accessory stromal cells.

Studies on the rate and net expansion of LTC-ICs and CFCs in stirred suspension bioreactors have also exemplified the dependency on seeding density, with the highest cell density ($0.1\text{--}1 \times 10^6$ cells/ml) of BM MNCs resulting in a significant net expansion of primitive HPCs.^[62] The typical inoculum range reported has been in the order of $1\text{--}3 \times 10^6$ MNCs/ml and $0.5\text{--}1 \times 10^4$ CD34⁺ cells/ml.^[62,64,66,67] However, the range can vary depending on the culture system, the use of preformed stromal “feeder” layer, or the use of scaffolds. Therefore, it is important that individual laboratories determine the initial inoculum concentrations that generate optimal cell-outputs of the desired cells from a given starting population under specific experimental conditions.

Feeding Schedule and Culture Duration

The *in vivo* BM microenvironment has a perfusion rate of $0.1 \text{ ml/cm}^3/\text{min}$,^[68] wherein hematopoietic cells are subjected to continuous feeding of nutrients and the removal of metabolic wastes. To mimic such an environment, various feeding strategies have been employed in the cultivation of hematopoietic cells.^[66,69–73] The major role of medium exchange is to overcome the limitation of growth^[69] due to the depletion of substrates and the accumulation of the metabolic by-products such as lactate that result in a drop in culture pH,^[74] and growth inhibitory substances. Mass transfer limitations of oxygen can also restrict cell expansion.^[62,64,75] In static cultures, especially at high cell densities of BM MNCs in the presence of preformed stroma, frequent medium exchange enhanced the cell and colony-forming unit-granulocyte/macrophage (CFU-GM) outputs.^[64] Perfusion systems offer several advantages over static cultures in terms of automated, continuous, and controlled medium exchange. Perfusion cultures of peripheral blood (PB) MNCs or CD34⁺ cells have demonstrated an increase in LTC-IC numbers in comparison to static cultures^[66] and have been shown to enhance the expansion of total and CFCs.^[73] Finally, perfusion cultures have the added advantage over frequently-fed static cultures in that during feeding they do not result in the depopulation of nonadherent cell fractions and they require less handling.^[66]

The duration of hematopoietic cultures is especially relevant for clinical applications. In CD34⁺ *ex vivo* expansion from mobilized peripheral blood (MPB), a two-phase growth was observed with a pre-CFU phase during the first 7–10 days and a phase past day 10 where CFU-GM decreased and mature myeloid cell

numbers increased.^[76] Many clinically relevant studies have reported a decrease in progenitor production during days 12–15 of culture, probably due to inadequate feeding protocols and high cell concentrations, indicating that expansion in static liquid cultures should be limited to 7–14 days.^[77] However, several preclinical studies have reported continuously increasing levels of expansion of all hematopoietic lineages for several months.^[78]

pH

Hematopoietic cultures, unless controlled, experience varying pH with time. Hence, the *ex vivo* expansion of hematopoietic cells is affected by gradients in pH^[79] that may alter cell physiology, receptor expression,^[80,81] and cell differentiation.^[80–87] The optimum extracellular pH is dependent on the hematopoietic cell lineage,^[80–87] with enhanced proliferation of phytohemagglutinin-activated T cells at pH 7.0–7.2, accelerated erythroid differentiation at pH 7.6,^[82–84] optimal formation of CFU-GM at pH 7.2–7.4,^[82] and enhanced differentiation, polyploidization, and apoptosis of megakaryocytes at pH 7.4–7.6.^[86] PB-derived macrophage progenitors (CFU-M) appear to be more sensitive to acidic pH than granulocytic progenitors (CFU-G), yet cord blood (CB)-derived progenitor cells are more resistant to pH variation in methylcellulose cultures.^[82] Expression of cell surface receptors, which mediate cell adhesion and signal transduction, is also pH-dependent. For example, the receptor for granulocyte-colony-stimulating factor (G-CSF) is expressed optimally at pH 7.22,^[80] whereas expression of the IL-2 receptor is optimal at pH 7.0.^[81]

Oxygen Tension

Oxygen tension affects the growth of hematopoietic progenitor cells,^[88–90] and increases the responsiveness of cells to cytokines^[90,91] and the ability of these cells to produce cytokines.^[92] The concentration of oxygen within the microenvironment is more important than the gas phase oxygen level.^[8] *In vivo*, BM oxygen tension is between 10 and 50 mm Hg.^[8,88] *In vitro*, low oxygen concentration (2–7%) enhances the size and number of hematopoietic colonies in semisolid cultures,^[88,89,91,93–95] and the optimal expansion of human BM CFU-neutrophil, macrophage, and eosinophil progenitors (CFU-C), burst-forming unit-erythroids (BFU-E), and multipotential progenitors (CFU-mix) occurs at 7% O₂.^[88] The production of progenitor and mature cells is likewise dependent upon oxygen tension,^[88–91,95,96] with the formation of CB BFU-E favored in hypoxic conditions (1%), whereas cultures

exposed to 20% oxygen produced mature erythrocytes.^[89] Similarly, the size and number of CFU-megakaryocytes (CFU-MKs) was enhanced at low oxygen tension (5%), whereas high oxygen tension (20%) enhanced CD41⁺ megakaryocyte generation, high ploidy, proplatelet formation, and apoptosis.^[96] This *in vitro* situation may resemble the *in vivo* situation, where megakaryocyte maturation and platelet production occur adjacent to the BM sinus, which is exposed to venous pO₂ (8%). In contrast, early progenitors and stem cells reside in the core of the bone marrow hematopoietic compartment (BMHC), where the average concentration of O₂ is 1–5%.^[97]

Stimulation under low oxygen tension can, in part, be attributed to the increased responsiveness of the cells to cytokines^[90,95,96,98] and to the production of hematopoietic GFs.^[92] Thus, oxygen tension may regulate the production of erythropoietin (Epo) by macrophages^[92] as well as the formation of CFU-erythroid (CFU-E) and BFU-E by altering the sensitivity of these colonies to Epo.^[91] Therefore, there may exist an optimal oxygen tension based on the cytokines used and on the developmental state/potential of the cells.

In vitro, the availability of oxygen depends on culture configuration, cell density, and gas-phase oxygen tension.^[42] At early stages of culture, low oxygen tension is sufficient for the production and maintenance of progenitor cells. As cell numbers increase, oxygen becomes a limiting factor. Depending on the culture system used, oxygen may be maintained within the system by increasing gas-phase oxygen tension, agitation rates, or perfusion rates. However, increasing the gas-phase oxygen tension to overcome the increased demand due to increased cell density can result in oxygen toxicity. Increased feeding and the use of perfusion are preferred alternatives to oxygenation because they do not result in cell damage.^[95] The generation of reactive oxygen intermediates, as a consequence of cellular metabolism and the accumulation of neutrophils, macrophages, and mature myeloid cells, results in the progressive decline of progenitor cell production.^[91,99] To reduce oxygen toxicity in response to high oxygen tension, antioxidants, such as reduced glutathione, vitamin E, and 2-mercaptoethanol,^[88,91] as well as scavengers of hydroxyl radicals and hydrogen peroxide, such as mannitol and catalase, may be used.^[99]

CONCLUSIONS

The success of clinical stem cell transplantation depends on the supply of sufficient numbers of high-quality repopulating HSCs. In addition, a number of clinical strategies that combine HSC with

accessory and mesenchymal stem cells are being examined. Although significant progress has been made in characterizing the factors that govern hematopoiesis, *ex vivo* expansion of hematopoietic cells appears to be a technique still in its infancy. Nevertheless, the *ex vivo* expansion of hematopoietic cells for clinical use may hold potential.

The operational control of hematopoiesis *in vitro* presents a challenge in terms of the cell source and type, culture system, culture medium, and operating conditions.^[22] Even slight deviations in the culture parameters can affect the final cell output. Factorial and composite designs may be utilized to increase the efficiency of investigations on the effects of stimulatory (or inhibitory) factors and culture parameters on hematopoietic culture systems, and to reveal unexpected interactions that would otherwise be missed by conventional analyses.^[100,101] The ultimate challenge in developing successful *ex vivo* HSC expansion processes, however, lies in devising convenient assays that faithfully reflect the homing, function, and repopulating activities of stem cells in human subjects.

T

ABBREVIATIONS

BFU-E	burst-forming unit-erythroid
BM	bone marrow
BMP	bone morphogenetic protein
CB	cord blood
CFC	colony-forming cell
CFU	colony-forming unit
CFU-Bas	CFU-basophil
CFU-C	CFU-neutrophil, macrophage, and eosinophil progenitor
CFU-E	CFU-erythroid
CFU-Eos	CFU-eosinophil
CFU-GEMM	CFU-granulocyte/erythrocyte/macrophage/megakaryocyte
CFU-GM	CFU-granulocyte/macrophage
CFU-L	CFU-lymphocyte
CFU-M	CFU-macrophage
CFU-Meg or CFU-MK	CFU-megakaryocyte
CFU-mix	CFU- multipotential progenitor
CFU-S	CFU-spleen
CSF	colony-stimulating factor
ECM	extracellular matrix
Epo	erythropoietin
FL	flt-3 ligand

G-CSF	granulocyte-colony-stimulating factor
GF	growth factor
GM-CSF	granulocyte/macrophage-colony-stimulating factor
HIM	hematopoietic inductive microenvironment
HPC	hematopoietic progenitor cell
HSC	hematopoietic stem cell
IGF-1	insulin-like growth factor-1
IL	interleukin
LIF	leukemia inhibitory factor
LTC-IC	long-term culture-initiating cell
M-CSF	macrophage-colony stimulating factor
MIP-1 α	macrophage inflammatory protein-1 α
MNC	mononuclear cell
MPB	mobilized peripheral blood
MSC	mesenchymal stem cell
SCF	stem cell factor
TGF- β	transforming growth factor- β
TNF- α	tumor necrosis factor- α
TPO	thrombopoietin
2-ME	2-mercaptoethanol
3-D	three-dimensional

ACKNOWLEDGMENT

This work was supported by grants to JHDW from NSF BES-9631670 and NASA NAG9-1360. AM would like to thank the "Alexander S. Onassis" Foundation for providing a scholarship.

ARTICLES OF FURTHER INTEREST

Extracellular Matrix Scaffolds; Hematopoietic Stem Cells and Assays; Stem Cells; Tissue Engineering of Bone Marrow, Culture Systems; Tissue Engineering Scaffolds

REFERENCES

- Kaplan, A.; Bruder, S.P. Mesenchymal stem cells: Building blocks for molecular medicine in the 21st century. *Trends Mol. Med.* **2001**, *7* (6), 259–264.
- Madlambayan, G.J.; Rogers, I.; Casper, R.F.; Zandstra, P.W. Controlling culture dynamics for the expansion of hematopoietic stem cells. *J. Hematother. Stem Cell Res.* **2001**, *10*, 481–492.
- Abboud, C.N.; Lichtman, M.A. Structure of the Marrow and the Hemopoietic Microenvironment. In *Hematology*; Beutler, E., Ed.; McGraw-Hill: New York, 2001; 29–58.
- McNiece, I.; Briddell, R. Ex vivo expansion of hematopoietic progenitor cells and mature cells. *Exp. Hematol.* **2001**, *29*, 3–11.
- Charbord, P. Microenvironmental Cell Populations Essential for the Support of Hematopoietic Stem Cells. In *Hematopoiesis: A Developmental Approach*; Zon, L.I., Ed.; Oxford University Press: New York, 2001; 691–701.
- Trentin, J.J. Influence of Hematopoietic Organ Stroma (Hematopoietic Inductive Microenvironments) on Stem Cell Differentiation. In *Regulation of Hemopoiesis*; Gordon, A.S., Ed.; Appleton-Century-Crofts: New York, 1970; 161–186.
- Naito, K.; Tamahashi, N.; Chiba, T.; Kaneda, K.; Okuda, M.; Endo, K.; Yoshinaga, K.; Takahashi, T. The microvasculature of the human bone marrow correlated with the distribution of hematopoietic cells. A computer-assisted three-dimensional reconstruction study. *Tohoku J. Exp. Med.* **1992**, *166*, 439–450.
- Koller, M.R.; Palsson, B.O. Tissue engineering: Reconstitution of human hematopoiesis ex vivo. *Bio-technol. Bioeng.* **1993**, *42*, 909–930.
- Dorshkind, K. Regulation of hemopoiesis by bone marrow stromal cells and their products. *Annu. Rev. Immunol.* **1990**, *8*, 111–137.
- Bianco, P.; Riminucci, M.; Gronthos, S.; Robey, P.G. Bone marrow stromal cells: Nature, biology, and potential applications. *Stem Cells* **2001**, *19*, 180–192.
- Campbell, A.; Wicha, M.S. Extracellular matrix and the hematopoietic microenvironment. *J. Lab. Clin. Med.* **1988**, *112* (2), 140–146.
- Dao, M.A.; Hashimo, K.; Kato, I.; Nolte, J.A. Adhesion to fibronectin maintains regenerative capacity during ex vivo culture and transduction of human hematopoietic stem and progenitor cells. *Blood* **1998**, *92*, 4612–4621.
- Campbell, A.; Long, M.W.; Wicha, M.S. Haemonectin, a bone marrow adhesion protein specific for cells of granulocyte lineage. *Nature* **1987**, *329*, 744–746.
- Weistein, R.; Riordan, M.A.; Wenc, K.; Kreczko, S.; Zhou, M.; Dainiak, N. Dual role of fibronectin in hematopoietic differentiation. *Blood* **1989**, *73*, 111–116.
- Liesveld, J.; Winslow, J.M.; Kempski, M.C.; Ryan, D.H.; Brennan, J.K.; Abboud, C.N. Adhesive interaction of normal and leukemic human CD34⁺ myeloid progenitors-role of marrow stromal, fibroblast, and cytomatrix components. *Exp. Hematol.* **1991**, *9*, 63–70.
- Verfaillie, C. Soluble factor(s) produced by human bone marrow stroma increase cytokine-induced proliferation and maturation of primitive hematopoietic progenitors while preventing their terminal differentiation. *Blood* **1993**, *82*, 2045–2053.
- Weissman, I. Developmental switches in the immune system. *Cell* **1994**, *76*, 207–218.
- Aiuti, A.; Friedrich, C.; Sieff, C.; Gutierrez-Ramos, J. Identification of distinct elements of the stromal microenvironment that control human hematopoietic

- stem/progenitor cell growth and differentiation. *Exp. Hematol.* **1998**, *26*, 143–157.
19. Quesenberry, P.; Crittenden, R.; Lowry, P.; Kittler, E.; Rao, S.; Peters, S.; Ramshaw, H.; Stewart, F. In vitro and in vivo studies of stromal niches. *Blood Cells* **1994**, *20*, 97–106.
 20. Mantalaris, A.; Keng, P.; Bourne, P.; Chang, A.Y. C.; Wu, J.H.D. Engineering a human bone marrow model: A case study on ex vivo erythropoiesis. *Biotechnol. Prog.* **1998**, *14*, 126–133.
 21. Noll, T.; Jelinek, N.; Schmidt, S.; Biselli, M.; Wandrey, C. Cultivation of hematopoietic stem and progenitor cells: Biochemical engineering aspects. *Adv. Biochem. Eng.* **2002**, *74*, 111–128.
 22. Nielsen, L.K. Bioreactors for hematopoietic cell culture. *Annu. Rev. Biomed. Eng.* **1999**, *1*, 129–152.
 23. Douay, L. Experimental culture conditions are critical for ex vivo expansion of hematopoietic cells. *J. Hematother. Stem Cell Res.* **2001**, *10*, 341–346.
 24. Alley, C.; MacDermott, R.P.; Stewart, C.C. The effect of serum on human marrow mononuclear cell proliferation and maturation. *J. Reticuloendothel. Soc.* **1983**, *34*, 271–278.
 25. Bohmer, R. Interaction of serum and colony-stimulating factor for survival of a factor-dependent hemopoietic progenitor cell line. *J. Cell. Physiol.* **1989**, *139*, 531–537.
 26. Collins, P.; Miller, W.M.; Papoutsakis, E.T. Ex vivo culture systems for hematopoietic cells. *Curr. Opin. Biotechnol.* **1996**, *7* (2), 223–230.
 27. Guba, S.; Gottschalk, L.R.; Jing, Y.H.; Mulligan, T.; Emerson, S.G. Bone marrow stromal fibroblasts secrete interleukin-6 and granulocyte-macrophage colony-stimulating factor in the absence of inflammatory stimulation: Demonstration by serum-free bioassay, enzyme-linked immunosorbent assay, and reverse transcriptase polymerase chain reaction. *Blood* **1992**, *80* (5), 1190–1198.
 28. Tsukada, J.; Misago, M.; Kikuchi, M.; Sato, T.; Ogawa, R.; Ota, T.; Oda, S.; Morimoto, I.; Chiba, S.; Eto, S. Interaction between recombinant human erythropoietin and serum factor(s) on murine megakaryocyte colony formation. *Blood* **1992**, *80* (1), 37–45.
 29. Lill, M.C.; Lynch, M.; Fraser, J.K.; Chung, G.Y.; Schiller, G.; Glaspy, J.A.; Souza, L.; Baldwin, G.C.; Gasson, J.C. Production of functional myeloid cells from CD34-selected hematopoietic progenitor cells using a clinically relevant ex vivo expansion system. *Stem Cells* **1994**, *12*, 626–637.
 30. Fortunel, N.; Hatzfeld, A.; Hatzfeld, J.A. Transforming growth factor-B: Pleiotropic role in the regulation of hematopoiesis. *Blood* **2000**, *96* (6), 2022–2036.
 31. Krystal, G.; Lanm, V.; Dragowska, W.; Takahashi, C.; Appel, J. Transforming growth factor beta 1 is an inducer of erythroid differentiation. *J. Exp. Med.* **1994**, *180*, 851–860.
 32. Dybedal, I.; Jacobsen, S.E. Transforming growth factor beta (TGF-beta), a potent inhibitor of erythropoiesis: Neutralizing TGF-beta antibodies show erythropoietin as a potent stimulator of murine burst-forming unit erythroid colony formation in the absence of a burst-promoting activity. *Blood* **1995**, *86*, 949–957.
 33. Sandstrom, C.; Miller, W.M.; Papoutsakis, E.T. Review: Serum-free media for cultures of primitive and mature hematopoietic cells. *Biotechnol. Bioeng.* **1994**, *43*, 706–733.
 34. Rosenzweig, M.; Canque, B.; Gluckman, J.C. Human dendritic cell differentiation pathway from CD34⁺ hematopoietic precursor cells. *Blood* **1996**, *87*, 535–544.
 35. Brugger, W.; Mocklin, W.; Heimfeld, S.; Berenson, R.J.; Mertelsmann, R.; Kanz, L. Ex vivo expansion of enriched peripheral blood CD34⁺ progenitor cells by stem cell factor, interleukin-1 (IL-1), IL-6, IL-3, interferon γ and erythropoietin. *Blood* **1993**, *81*, 2579–2584.
 36. Brown, R.; Xu, F.S.; Dusing, S.K.; Li, Q.; Fischer, R.; Patchen, M. Serum-free conditions for cells capable of producing long-term survival in lethally irradiated mice. *Stem Cells* **1997**, *15*, 237–245.
 37. Lansdorp, P.M.; Lansdorp, W.; Dragowska, W. Long-term erythropoiesis from constant numbers of CD34⁺ cells in serum-free cultures initiated with highly purified progenitor cells from human bone marrow. *J. Exp. Med.* **1992**, *175*, 1501–1509.
 38. Mobest, D.; Mertelsmann, R.; Henschler, R. Serum-free ex vivo expansion of CD34⁺ hematopoietic progenitor cells. *Biotechnol. Bioeng.* **1998**, *60* (3), 341–347.
 39. Mobest, D.; Goan, S.; Junghahn, I.; Winkler, J.; Fichtner, I.; Becker, M.; De Lima-Hahn, E.; Mertelsmann, R.; Henschler, R. Differential kinetics of primitive hematopoietic cells assayed in vitro and in vivo during serum-free suspension culture of CD34⁺ blood progenitor cells. *Stem Cells* **1999**, *17*, 152–161.
 40. Yamaguchi, M.; Hiramaya, F.; Kanai, M.; Sato, N.; Fukazawa, K.; Yamashita, K.; Sawada, K.; Koike, T.; Kuwabara, M.; Ikeda, H.; Ikebuchi, K. Serum-free coculture system for ex vivo expansion of human cord blood primitive progenitors and SCID mouse-reconstituting cells using human bone marrow primary stromal cells. *Exp. Hematol.* **2001**, *29*, 174–182.
 41. Almeida-Porada, G.; Brown, R.L.; MacKintosh, F.R.; Zanjani, E.D. Evaluation of serum-free culture conditions able to support the ex vivo expansion and engraftment of human hematopoietic stem cells in human-to sheep xenograft model. *J. Hematother. Stem Cell Res.* **2000**, *9*, 683–693.
 42. McAdams, T.A.; Winter, J.N.; Miller, W.M.; Papoutsakis, E.T. Hematopoietic cell culture therapies (Part II): Clinical aspects and applications. *Trends in Biotechnol.* **1996**, *14* (10), 388–396.
 43. McAdams, T.; Sandstrom, C.E.; Miller, W.M.; Bender, J.G.; Papoutsakis, E.T. Ex vivo expansion of primitive hematopoietic cells for cellular therapies: An overview. *Cytotechnology* **1995**, *18*, 133–146.
 44. Verfaillie, C.M. Ex Vivo Expansion of Hematopoietic Stem Cells. In *Hematopoiesis: A Developmental Approach*; Zon, L.I., Ed.; Oxford University Press: New York, 2001; 119–129.

45. Audet, J.; Zanstra, P.W.; Eaves, C.J.; Piret, J.M. Advances in hematopoietic stem cell culture. *Curr. Opin. Biotechnol.* **1998**, *9*, 146–151.
46. Broudy, V.C. Stem cell factor and hematopoiesis. *Blood* **1997**, *90*, 1354–1364.
47. Lyman, S.D.; Williams, D.E. Biology and potential clinical applications for flt3 ligand. *Curr. Opin. Hematol.* **1995**, *2*, 177–181.
48. Brandt, J.; Bhalla, K.; Hoffman, R. Effects of interleukin-3 and c-kit ligand on the survival of various classes of human hematopoietic progenitor cells. *Blood* **1994**, *83* (6), 1507–1514.
49. Lyman, S.D.; Jacobsen, S.E. c-kit ligand and Flt3 ligand: Stem/progenitor cell factors with overlapping yet distinct activities. *Blood* **1998**, *91*, 1101–1134.
50. Shah, A.J.; Smogorzewska, E.M.; Hannum, C.; Crooks, G.M. Flt3 ligand induces proliferation of quiescent human bone marrow CD34⁺CD38⁻ cells and maintains progenitor cells in vitro. *Blood* **1996**, *87*, 3563–3570.
51. Petzer, A.L.; Zandstra, P.W.; Piret, J.M.; Eaves, C.J. Differential cytokine effects on primitive (CD34⁺CD38⁻) human hematopoietic cells: Novel responses to flt3-ligand and thrombopoietin. *J. Exp. Med.* **1996**, *183*, 2551–2558.
52. Won, J.H.; Cho, S.D.; Park, S.K.; Lee, G.T.; Baick, S.H.; Suh, W.S.; Hong, D.S.; Park, H.S. Thrombopoietin in synergy with other cytokines for expansion of cord blood progenitor cells. *J. Hematother. Stem Cell Res.* **2000**, *9*, 465–473.
53. Kobayashi, M.; Laver, J.H.; Kato, T.; Miyazaki, H.; Ogawa, M. Thrombopoietin supports proliferation of human primitive hematopoietic cells in synergy with steel factor and/or interleukin-3. *Blood* **1996**, *88*, 429–436.
54. Ogawa, M.; Clark, S.C. Synergistic interaction between interleukin-6 and interleukin-3 in support of stem cell proliferation in culture. *Blood Cells* **1988**, *14*, 329–335.
55. Holyoake, T.L.; Freshney, M.G.; McNair, L.; Parker, A.N.; McKay, P.J.; Steward, W.P.; Fitzsimons, E.; Graham, G.J.; Pragnell, I.B. Ex vivo expansion with stem cell factor and interleukin-11 augments both short-term recovery posttransplant and the ability to serially transplant marrow. *Blood* **1996**, *87*, 4589–4595.
56. Shih, C.-C.; et al. A secreted and LIF-mediated stromal cell-derived activity that promotes ex vivo expansion of human hematopoietic stem cells. *Blood* **2000**, *95*, 1957–1966.
57. Bhatia, M.; et al. Bone morphogenetic proteins regulate the developmental program of human hematopoietic stem cells. *J. Exp. Med.* **1999**, *189*, 1139–1148.
58. Karanu, F.N.; et al. The notch ligand jagged-1 represents a novel factor of human hematopoietic stem cells. *J. Exp. Med.* **2000**, *192*, 1365–1372.
59. Eaves, C.J.; et al. Mechanisms that regulate the cell cycle status of very primitive hematopoietic cells in long-term human marrow cultures. II. Analysis of positive and negative regulators produced by stromal cells within the adherent layer. *Blood* **1991**, *78*, 110–117.
60. Broxmeyer, H.E.; et al. Comparative analysis of the human macrophage inflammatory protein family of cytokines (chemokines) on proliferation of human myeloid progenitor cells. Interacting effects involving suppression, synergistic suppression, and blocking of suppression. *J. Immunol.* **1993**, *150*, 3448–3458.
61. Verfaillie, C.M.; et al. MIP-1 α combined with IL-3 conserves primitive human LTBMIC-IC for at least 8 weeks in ex vivo “stroma-non-contact” cultures. *J. Exp. Med.* **1994**, *179*, 643–649.
62. Zandstra, P.W.; Eaves, C.J.; Piret, J.M. Expansion of hematopoietic progenitor cell populations in stirred suspension bioreactors of normal human bone marrow cells. *Bio/Technology* **1994**, *12*, 909–914.
63. Zandstra, P.W.; Petzer, A.L.; Eaves, C.J.; Piret, J.M. Cellular determinants affecting the rate of cytokine depletion in cultures of human hematopoietic cells. *Biotechnol. Bioeng.* **1997**, *54*, 58–68.
64. Koller, M.; Manchel, I.; Palsson, M.A.; Maher, R.J. Different measures of human hematopoietic cell culture performance are optimized under vastly different conditions. *Biotechnol. Bioeng.* **1996**, *50*, 505–513.
65. Poloni, A.; Giarratana, M.C.; Kobari, L.; Firat, H.; Bouchet, S.; Gorin, N.C.; Douay, L. The ex vivo expansion capacity of normal human bone marrow cells is dependent on experimental conditions: Role of cell concentration, serum and CD34⁺ cell selection in stroma-free cultures. *Hematol. Cell. Ther.* **1997**, *39*, 49–58.
66. Sandstrom, C.E.; et al. Effects of CD34⁺ cell selection and perfusion on ex vivo expansion of peripheral blood mononuclear cells. *Blood* **1995**, *86*, 958–970.
67. Koller, M.R.; Manchel, I.; Newsom, B.S.; Palsson, M.A.; Palsson, B.O. Bioreactor expansion of human bone marrow: Comparison of unprocessed, density-separated and CD34-enriched cells. *J. Hematother.* **1995**, *4*, 159–169.
68. Martiat, P.; Ferrant, A.; Cogneau, M.; Bol, A.; Rodhain, J.; Michaux, J.L.; Sokal, G. Assessment of bone marrow blood flow using positron emission tomography: No relationship with bone cellularity. *Br. J. Haematol.* **1987**, *66*, 307–310.
69. Caldwell, J.; Locey, B.; Clarke, M.F.; Emerson, S.G.; Palsson, B.O. Influence of medium exchange schedules on metabolic, growth, and GM-CSF secretion rates of genetically engineered NIH-3T3 cells. *Biotechnol. Prog.* **1991**, *7*, 1–8.
70. Caldwell, J.; Palsson, B.O.; Locey, B.; Emerson, S.G. Culture perfusion schedules influence the metabolic activity and granulocyte-macrophage colony-stimulating factor production rates of human bone marrow stromal cells. *J. Cell. Physiol.* **1991**, *147*, 344–353.
71. Koller, M.R.; Bender, J.G.; Miller, M.E.; Papoutsakis, E.T. Expansion of primitive human hematopoietic progenitors in a perfusion bioreactor system with IL-3, IL-6, and stem cell factor. *Bio/Technology* **1993**, *11*, 358–363.
72. Martinson, J.; Unverzagt, K.; Schaeffer, A.; Smith, S.L.; Loudovaris, M.; Schneidkraut, M.J.; Bender, J.G.; Van Epps, D.E. Neutrophil precursor

- generation: Effect of culture conditions. *J. Hematother.* **1998**, *7*, 463–471.
73. Schwartz, R.; Palsson, B.O.; Emerson, S.G. Rapid medium perfusion rate significantly increases the productivity and longevity of human bone marrow cultures. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 6760–6764.
 74. Patel, S.D.; Papoutsakis, E.T.; Winter, J.N.; Miller, W.M. The lactate issue revisited: Novel feeding protocols to examine inhibition of cell proliferation and glucose metabolism in hematopoietic cell cultures. *Biotechnol. Prog.* **2000**, *16*, 885–892.
 75. Oh, D.; Koller, M.R.; Palsson, B.O. Frequent harvesting from perfused bone marrow cultures results in increased overall cell and progenitor expansion. *Biotechnol. Bioeng.* **1994**, *44*, 609–616.
 76. Haylock, D.N.; To, L.B.; Dowse, T.L.; Juttner, C.A.; Simmons, P.J. Ex vivo expansion and maturation of peripheral blood CD34⁺ cells into the myeloid lineage. *Blood* **1992**, *80*, 1405–1412.
 77. Bregni, M.; Magni, M.; Siena, S.; Di Nicola, M.; Bonadonna, G.; Gianni, A.M. Human peripheral blood hematopoietic progenitors are optimal targets of retroviral-mediated gene transfer. *Blood* **1992**, *80*, 1418–1422.
 78. Piacibello, W.; Sanavio, F.; Garetto, L.; Severino, A.; Dane, A.; Gammaitoni, I.; Aggietta, M. Differential growth factor requirement of primitive cord blood hematopoietic stem cells for self-renewal and amplification versus proliferation and differentiation. *Leukemia* **1998**, *12*, 718–727.
 79. Akatov, V.; Lezhnev, E.I.; Vexler, A.M.; Kublik, L.N. Low pH value of pericellular medium as a factor limiting cell proliferation in dense cultures. *Exp. Cell Res.* **1985**, *160*, 412–418.
 80. Hevehan, D.; Papoutsakis, E.T.; Miller, W.M. Physiologically significant effect of pH and oxygen tension on granulopoiesis. *Exp. Hematol.* **2000**, *28*, 267–275.
 81. Carswell, K.; Papoutsakis, E.T. Extracellular pH affects the proliferation of cultured human T-cells and their expression of the interleukin-2 receptor. *J. Immunol.* **2000**, *23* (6), 669–674.
 82. McAdams, T.A.; Miller, W.M.; Papoutsakis, E.T. Variations in culture pH affect the cloning efficiency and differentiation of progenitor cells in ex vivo haematopoiesis. *Br. J. Haematol.* **1997**, *97*, 889–895.
 83. Endo, T.; Ishibashi, Y.; Okana, H.; Fukumaki, Y. Significance of pH on differentiation of human erythroid cell lines. *Leuk. Res.* **1994**, *18*, 49–54.
 84. McAdams, T.A.; Miller, W.M.; Papoutsakis, E.T. pH is a potent modulator of erythroid differentiation. *Br. J. Haematol.* **1998**, *103*, 317–325.
 85. Fischkoff, S.; Pollak, A.; Gleich, G.J.; Testa, J.R.; Misawa, S.; Reber, T.J. Eosinophilic differentiation of the human promyelocytic leukemia cell line. *J. Exp. Med.* **1984**, *160*, 179–196.
 86. Yang, H.; Miller, W.M.; Papoutsakis, E.T. High pH promotes megakaryocytic maturation and apoptosis. *Stem Cells* **2002**, *20*, 320–328.
 87. Tennant, G. Control of pH in human long-term bone marrow cultures with low-glucose medium containing zwitterion buffer lengthens the period of haematopoiesis activity. *Br. J. Haematol.* **2000**, *109*, 785–787.
 88. Ishikawa, Y.; Ito, T. Kinetics of hematopoietic stem cells in a hypoxic culture. *Eur. J. Haematol.* **1988**, *40*, 126–129.
 89. Cipolleschi, M.; D'Ippolito, G.; Bernabei, P.A.; et al. Severe hypoxia enhances the formation of erythroid bursts from human cord blood cells and the maintenance of BFU-E in vitro. *Exp. Hematol.* **1997**, *25*, 1187–1194.
 90. Ivanovic, Z.; Belloc, F.; Faucher, J.; et al. Hypoxia maintains and interleukin-3 reduces the pre-colony-forming cell potential of dividing CD34⁺ murine bone marrow. *Exp. Hematol.* **2002**, *30*, 67–73.
 91. Rich, I.; Kubanek, B. The effect of reduced oxygen tension on colony formation of erythropoietic cells in vitro. *Br. J. Haematol.* **1982**, *52*, 579–588.
 92. Rich, I. A role for the macrophage in normal hemopoiesis. II. Effect of varying physiology oxygen tensions on the release of hemopoietic growth factors from bone-marrow-derived macrophage in vitro. *Exp. Hematol.* **1986**, *14*, 746–751.
 93. Cipolleschi, M.G.; Dello Sbarba, P.; Olivotto, M. The role of hypoxia in the maintenance of hematopoietic stem cells. *Blood* **1993**, *82*, 2031–2037.
 94. Ivanovic, Z.; Bartolozzi, B.; Bernabei, P.A.; et al. Incubation of murine bone marrow cells in hypoxia ensures the maintenance of marrow-repopulating ability together with the expansion of committed progenitors. *Br. J. Haematol.* **2000**, *108*, 424–429.
 95. Koller, M.; et al. Reduced oxygen tension increases hematopoiesis in long-term culture of human stem and progenitor cells from cord blood and bone marrow. *Exp. Hematol.* **1992**, *20*, 264–270.
 96. Mostafa, S.; Miller, W.M.; Papoutsakis, E.T. Oxygen tension influences the differentiation, maturation and apoptosis of human megakaryocytes. *Br. J. Haematol.* **2000**, *111*, 879–889.
 97. Pennathur-Das, R.L.; Levitt, L. Augmentation of in vitro human marrow erythropoiesis under physiological oxygen tensions is mediated by monocytes and T lymphocytes. *Blood* **1987**, *69* (3), 899–907.
 98. Naldini, A.; Carraro, F.; Silvestri, S.; et al. Hypoxia affects cytokine production and proliferation responses by human peripheral mononuclear cells. *J. Cell. Physiol.* **1997**, *173*, 335–342.
 99. Meagher, R.; et al. An analysis of the multilineage production of human hematopoietic progenitors in long-term bone marrow culture: Evidence that reactive oxygen intermediates derived from mature phagocytic cells have a role in limiting progenitor cell self-renewal. *Blood* **1988**, *72*, 273–281.
 100. Zandstra, P.W.; et al. Cytokine manipulation of primitive human hematopoietic cell self-renewal. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 4698–4703.
 101. Zandstra, P.W.; et al. Ontogeny-associated changes in the cytokine responses of primitive human hematopoietic cells. *Br. J. Haematol.* **1998**, *101*, 770–778.

Tissue Engineering of Bone Marrow, Culture Systems

Athanassios Mantalaris

Department of Chemical Engineering & Chemical Technology, Imperial College London, London, United Kingdom

Nicki Panoskaltsis

Department of Haematology, Imperial College London, London, United Kingdom

J. H. David Wu

Departments of Chemical Engineering, Biomedical Engineering, and Microbiology & Immunology, University of Rochester, Rochester, New York, U.S.A.

INTRODUCTION

Following the development of a stroma-dependent long-term murine bone marrow culture system by Dexter and colleagues,^[1] conventional bone marrow cultures have been on flat two-dimensional surfaces of tissue culture dishes. This methodology has been adapted for human long-term bone marrow culture.^[2] The major challenge for the *ex vivo* expansion of hematopoietic stem and progenitor cells is to provide a suitable culture environment that would support the expansion of clinically relevant cell numbers of the desired cell type(s) from relatively small donor samples.^[3]

STROMAL SUPPORT

There are two main approaches for *in vitro* hematopoietic cultures, namely, stroma-free and stroma-dependent. For clinical use, stroma-free cultures are more easily standardized and maintained.^[4] However, stroma-based cultures can facilitate a 600-fold expansion of human cord blood (CB) CD34⁺ cells and the enhanced expansion of colony-forming unit cells (CFU-Cs) and cobblestone area-forming cells (CAFCs).^[5,6] They can also provide a more physiologically relevant microenvironment to the hematopoietic stem and progenitor cells. Stromal feeders have also been used to support the retroviral transduction of autologous hematopoietic stem cells (HSCs)^[7] and to purge contaminating leukemic cells from chronic myelogenous leukemia (CML)^[8] and acute myelogenous leukemia (AML)^[9] autografts *ex vivo*. An advantage of using autologous feeders is that the graft intended for clinical applications will not be contaminated with allogeneic or xenogeneic feeder cells.^[4] However, *ex vivo* hematopoietic support may be defective due to exposure of the stroma donor to chemotherapy and radiation prior to harvest of the

autologous stroma feeders^[10] or due to inherent defects in the stromal cells as a result of the disease.^[11] To circumvent these potential problems, murine or human stromal feeder cells have been cloned from other sources.^[4]

STATIC CULTURES

Different devices have been developed to facilitate *ex vivo* hematopoietic culture, including vessels for static cultures and bioreactors (Table 1). Static culture systems, such as well-plates, T-flasks, or gas-permeable blood bags, are the most widely used devices.^[12,13]

Gas-permeable polypropylene bags provide a closed and easily manipulated system for the expansion of cells for clinical applications. In these systems, an optimal seeding density of 5×10^4 bone marrow (BM) CD34⁺ cells/ml or 1×10^4 CB CD34⁺ cells/ml in serum-free medium and a 14-day culture period resulted in the expansion of total cells, CD34⁺ cells, clonal progenitors (CFU-granulocyte/macrophage [GM] and burst-forming unit-erythroid [BFU-E]), and long-term culture-initiation cells (LTC-ICs).^[14] An increased seeding density induced the production of granulocytic cells with terminal differentiation, but the overall extent of cell expansion decreased.^[15] In contrast to flask cultures, gas-permeable blood bags do not support the development of stromal cells. Despite their widespread applications, static cultures are limited to relatively low-density cultures with low total cell output.^[12,13] The drawbacks of the blood bags include the difficulties in monitoring and control, the need for repeated handling during feeding, and the limited surface area available for growth and expansion of cells.^[13]

Bioreactor culture systems were developed to overcome the limitations of static cultures. They enable better control of culture parameters and automated feeding strategies, allowing a hands off operation and

Table 1 Summary of the culture systems used for cultivation of hematopoietic stem and progenitor cells

Culture system	Comments	Ref.
Static	<ul style="list-style-type: none"> Well-plates, T-flasks, or gas-permeable blood bags Lack of on-line monitoring and control for dissolved oxygen and pH Lack of support for automated feeding/exchange of culture medium 	[12,13]
Suspension (Stirred bioreactor)	<ul style="list-style-type: none"> Provision of a well-controlled, closed, and reproducible environment Cell-cell and cell-matrix interactions are not reproduced Generation of shear stress through agitation Easy scale-up 	[25,31]
Perfusion		
Hollow-fiber	<ul style="list-style-type: none"> Continually perfused culture medium Shear-free environment Spatial inhomogeneities for oxygen, pH, nutrients, and metabolites Difficulty in monitoring and controlling the culture environment Scale-up difficulties Hematopoietic cell expansion questionable 	[25]
Perfusion chambers	<ul style="list-style-type: none"> Continually perfused culture medium Relatively homogeneous bulk culture medium Support the formation of a stromal layer if desired Support higher expansion of LTC-IC numbers for both MNC and CD34⁺ cultures Problem with over diluting endogenously-produced factors Medium-intensive culture system Scale-up not straightforward due to the system geometry Problems with monitoring and control of the environment near cell periphery Clinical application (Aastrom Biosciences) 	[16–19,22,23,28]
Airlift	<ul style="list-style-type: none"> High cell density growth Supports the formation of stromal cells on fiberglass beads Adverse effects such as bubble shear and breakage caused by gas sparging Lower expansion of hematopoietic cell culture compared to static control 	[25,58]

Abbreviations: LTC-IC: long-term culture-initiating cell; MNC: mononuclear cell.

lower risk of contamination,^[3,16] and are able to support high cell densities and accessory cells. One such bioreactor system (Aastrom Biosciences, Ann Arbor, Michigan), which is an automated perfusion system, has obtained market approval from the Food and Drug Administration in the United States and a CE (Conformite Europeenne) mark in Europe.^[17,18]

PERFUSION CULTURES

Perfusion bioreactors have been developed to continually replace culture medium, thus removing any harmful metabolites and maintaining a constant supply of growth nutrients. However, care must be taken not to over-dilute endogenously produced soluble stimulatory compounds^[19] or to mechanically damage

the cells. Perfusion bioreactors have also been used to analyze the consumption and production of growth factors and the relationship of these parameters to specific cell production.^[20] Changes in the schedule of the medium perfusion may influence growth factor secretion by stromal cells.^[21] This suggests that it may be possible to provide culture conditions in which the stroma is maintained in a state that supports hematopoiesis but negates the need of growth factor supplementation. Perfusion bioreactor systems have been shown to support the expansion and maintenance of LTC-ICs and colony-forming cells (CFCs).^[16,17,22,23] Currently, at least three types of perfusion bioreactor systems have been reported, namely, the flatbed perfusion system,^[16,17,22] the grooved-bed perfusion system,^[23,24] and the hollow-fiber system.^[25]

Flatbed perfusion systems contain culture chambers that are inclined with the outlet above to reduce the removal of nonadherent cells during perfusion. The chambers are designed to maintain low laminar flow.^[22] They have been used to expand mononuclear cells (MNCs) in which a stromal layer developed^[16] or CD34-enriched cultures on preformed stroma.^[26] To facilitate stromal cell attachment, medium flow is started after a 24-hour initial period without perfusion.^[17] Nonadherent cells are harvested by aspirating the medium, whereas adherent cells are removed after trypsinization.^[17] The flatbed perfusion systems have been shown to facilitate the expansion of all types of human hematopoietic progenitor cells (HPCs) and HSCs from BM, CB, and mobilized peripheral blood (MPB).^[22,27] A 20–25-fold expansion of unselected human BM MNCs and a 10–30-fold expansion of CFU-GM over a two-week culture period^[16] have been observed. A BM aspirate of 10–15 ml can produce 3 billion cells in total and a 7.5-fold expansion of LTC-ICs.^[17] On preformed stromal cells, a 7-day culture of CB MNCs can support an 11-fold expansion of CFU-GM, a 2.5-fold expansion of BFU-E, and a 2.4-fold expansion of CFU-mix.^[22] Harvesting half the cell population every 3 to 4 days may further enhance the expansion of adult human bone marrow MNCs in perfusion bioreactors.^[28] Hence, when supplemented with cytokines and under optimized conditions, flatbed perfused bioreactors support expansion of LTC-ICs and CFU-GM, suggesting that adult-sized grafts from CB and MPB can be obtained. These culture systems are being evaluated in clinical trials, and initial results advocate that human repopulating stem cells can be preserved.^[4,29]

Grooved-bed perfusion bioreactors facilitate stroma-free cultures and hold more potential for clinical applications. The chambers are designed with multiple microgrooves, 200 μm across and 200 μm deep, separated every 200 μm . Because the microgrooves are perpendicular to the direction of flow, a higher perfusion rate (2.5 ml/min) can be accommodated, compared with the flatbed reactor, without disruption of the cells. However, the cells are isolated from the main convective flow of the medium, which creates inhomogeneities in concentration and presents difficulties in the monitoring and control of the cellular microenvironment.^[24] Grooved-bed perfusion bioreactors have supported the expansion of CFU-GM (19.4-fold) and can maintain LTC-IC (64% of input after a 15-day culture)^[23,24] from MPB MNC as well as from CB.

Hollow-fiber perfusion bioreactors offer a shear-free environment for cell proliferation and have been investigated for their potential applications in the expansion of HSCs.^[25] The bioreactor employed by Sardonini et al. had a 1.0-ft² fiber surface area with an

extracapillary space of 12 ml, which was seeded with 4.4×10^7 MNCs. Medium was perfused into the hollow-fiber cartridge at a rate of 50 ml/min. In contrast to the continuous flatbed and grooved-bed perfusion culture systems, the hollow-fiber bioreactor showed no observable expansion of hematopoietic cells when compared to the control (static) cultures.^[25] In addition, hollow-fiber bioreactor systems present inherent difficulties in scale-up because the spaces between the fiber modules are not kept constant, resulting in fluctuations in oxygen transfer and difficulties in monitoring of the culture.

STIRRED SUSPENSION CULTURES

Stirred culture systems are well characterized and widely used for the culture of both microbial and mammalian cells. They provide a well-controlled, closed, and reproducible culture environment and are superior to most static and perfusion systems for sampling ease, data collection, and control of culture parameters. Furthermore, stirred bioreactors could be advantageous for clinical applications in terms of the scale anticipated for clinical cultures.^[30] BM MNCs have been successfully propagated in serum-containing stirred culture systems;^[25,31] MPB and CB MNCs and CD34⁺ cells have been propagated in both serum and serum-free stirred cultures.^[13] Natural killer cells isolated from peripheral blood mononuclear cells have also been cultivated in stirred culture bioreactors.^[32] However, since hematopoietic cells are relatively sensitive to shear, low shear rates are required to avoid physical damage,^[13,25,31] altered expression of surface markers and cytokine receptors, and changes in cellular metabolism.^[33] Although macrophages tolerate suspension, fibroblasts and endothelial cells are unlikely to perform properly in stirred cultures.^[34] Therefore, agitation could have a profound effect on the type of cells that will expand and the extent of this expansion. Since suspension culture systems do not mimic the complex three-dimensional (3-D) microenvironment found in BM, it is not clear if extensive stem cell self-renewal can be sustained in the stirred culture. Expansion of cell numbers in such stirred systems can be achieved by the addition of microspheres that provide a scaffold for HSCs.^[25]

THREE-DIMENSIONAL CULTURES

Whereas *in vivo* hematopoiesis is regulated by cell-cell, cell-matrix, and growth factor interactions, most *ex vivo* culture systems have focused on the addition of cytokines alone and, hence, are insufficient to promote optimal blood cell development and retain long-term

generative capacity.^[35,36] Traditional stroma-dependent culture systems are not designed to enhance cell-cell and cell-matrix interactions or to create a spatial distribution of stromal and hematopoietic cells that mimic the *in vivo* 3-D configuration. Novel 3-D culture systems (Table 2) have been developed based on the hypothesis that growth in 3-D culture may more faithfully replicate the *in vivo* marrow hematopoietic inductive microenvironment (HIM) required for self-renewal and differentiation of stem and progenitor cells.^[37-48] We developed a long-term bone marrow culture system using porous microspheres as scaffolding to support 3-D growth configuration.^[40,41] Highly porous scaffolds used in these systems allow high-density cell growth across a higher specific surface area than the conventional static (monolayer) culture, thus creating a spatial distribution of stromal and hematopoietic cells.^[49] This spatial distribution facilitates the localization and delivery of cells and growth factors to the desired sites on the scaffolds and mimics the biological and mechanical function of the extracellular matrix (ECM).^[50] Furthermore, the interconnecting pores enable the dispersed cells to rearrange into functioning tissue units.^[48]

The scaffolds used for bone marrow tissue engineering applications are either naturally derived materials, such as macroporous collagen carriers,^[41,51] or synthetic polymers, such as porous polyvinyl formal (PVF),^[48] nonwoven polyethylene terephthalate (PET),^[49] porous biomatrix (Cellfoam™),^[45] tantalum-coated porous biomaterial (TCPB),^[44] porous gelatin microspheres (CultiSpher G),^[43] nylon filtration screen,^[37-39] polyester nonwoven fiber disks,^[52,53] and cellulose porous microspheres.^[40]

Macroporous collagen carriers have been used for the cultivation of CB CD34⁺ cells in coculture with the murine cell line S1/S1 in a fixed-bed reactor with continuous perfusion.^[51] The collagen microspheres supported the development of the murine stromal cell line, which produced ECM proteins. This system resulted in 100-, 114-, and 6.4-fold increases in MNCs, CFU-GMs, and CAFs, respectively. Research on synthetic polymers (in static culture systems) includes the use of collagen-coated nylon mesh as a template to grow BM cells, maintaining multilineage differentiation but with declining CFU numbers.^[37-39] A PVF resin has been utilized as a 3-D scaffold for the proliferation and differentiation of murine BM cells,^[48] which resulted in enhanced cellular proliferation and increased number of CFUs in the presence of growth factors and in the absence of preformed stroma. Nonwoven polyethylene terephthalate fabric^[49] has been used to simulate the marrow microenvironment for the expansion of CB progenitor cells. Such cultures enhanced cell-cell and cell-matrix interactions and allowed the 3-D distribution of

stromal and hematopoietic cells. In the presence of exogenous cytokines, 62.1-, 7.0-, and 6.0-fold increases in total cells, CD34⁺, and CFUs were obtained, respectively. Human BM CD34⁺ cells have also been cultured on inert Cellfoam™,^[45] which resulted in enhanced colony formation and LTC-IC in both cytokine-supplemented and cytokine-free cultures when compared with plastic cultures (using either preformed stroma or fibronectin-coated plastic well-dishes). A comparison between polyester nonwoven disks (Fibra cel) and porous cellulose carriers using mouse bone marrow showed that the weekly output of total hematopoietic cells increased at week 3 in the culture using polyester nonwoven disks along with the proportion of progenitor cells (BFU-E and CFU-GM) after week 1 in contrast to the output of the porous cellulose carriers.^[53] A tantalum-coated porous scaffold has been shown to improve retroviral transduction and maintain hematopoietic progenitor cells.^[44,54] This scaffold, in the absence of exogenous cytokines, supported up to a 1.5-fold expansion of BM CD34⁺ cells as well as a 6.7-fold increase in CFUs.^[44]

We have developed packed-bed bioreactors, both continuous and batch-fed, that use porous collagen or cellulose microspheres to grow both murine and human BM cells.^[40,41] These cultures supported multilineage differentiation in the absence of exogenous growth factors and a 3-D growth configuration that mimicked that of BM *in vivo*. In the presence of cytokines, including erythropoietin (Epo), this system supported extensive erythropoiesis when compared with a traditional flask culture under lower, more physiological concentrations of Epo,^[40] indicating that the culture system better mimics the HIM *in vivo* and is a more physiological marrow model. Furthermore, CultiSpher G microspheres, coated with preformed BM stroma have been used in a stirred tank culture system for the expansion of CB MNCs in the presence of exogenous cytokines.^[43] This system supported 7.7-, 23.3-, and 9.6-fold increases in total cells, CFU-C, and CD34⁺ cells, respectively, as compared to the static control after a 12-day culture period. Distinct hematopoietic islands were observed on the scaffolds, indicating the localization and compartmentalization of the progenitor cells.

Biodegradable porous scaffolds can be prepared in different ways.^[55-57] Wang et al. manufactured and evaluated bioactive and biodegradable materials and scaffolds, both porous or nonporous,^[42] and found that incorporation of bioceramic particles in the biodegradable polymers would render the composites bioactive and enhance their ability to induce the formation of bonelike apatite on their surfaces. These biodegradable porous materials may be employed for hematopoietic cell culture.

Table 2 Summary of the 3-D scaffolds used for the growth and expansion of HSCs and HPCs

Scaffold	Experimental condition	Comments	Ref.
Porous polyvinyl formal (PVF)	Culture system: static Surface treatment: type I collagen Porosity: 90% Pore size: 130 μm Cytokines: IL-3, IL-6, Epo, SCF	<ul style="list-style-type: none"> • Cell expansion without a preformed stroma • Increased cell proliferation efficiency • Increased CFUs and HPCs 	[48]
Nonwoven polyethylene terephthalate (PET) fabric	Culture system: static Surface treatment: hydrolysis step to improve adhesion Porosity: 85% Pore size: 10–60 μm Cytokines: TPO, Flk-2/flt-3 ligand	<ul style="list-style-type: none"> • Supports CB stroma development, which does not contain adipocytes (compared to BM stroma) • Increase in total cell number and CFUs in the absence of GFs. • Enhanced long-term expansion in the presence of GFs • 62.1-, 7-, and 6-fold increase in total cells, CD34⁺, and CFUs in the presence of GFs 	[49]
Porous biomatrix (Cellfoam™)	Culture system: static Surface treatment: tantalum, fibronectin, and collagen Porosity: 90% Pore size: not reported Cytokines: none and IL-3, IL-6, SCF, Flk-2/Flt-3 ligand	<ul style="list-style-type: none"> • Increased CFUs and LTC-ICs in both cytokine-free and cytokine-containing cultures compared with 2-D cultures 	[45]
Tantalum-coated porous biomaterial (TCPB)	Culture system: static Surface treatment: tantalum and fibronectin or collagen Porosity: 90% Pore size: not reported Cytokines: none	<ul style="list-style-type: none"> • No preformed stroma requirement • 1.5- and 6.7-fold increase in HPCs and CFCs, respectively • Supports progenitor viability and multipotency 	[44]
Nylon filtration screen	Culture system: static Surface treatment: collagen Porosity: not reported Pore size: 210 μm Cytokines: none	<ul style="list-style-type: none"> • Supports stroma development • Supports multilineal differentiation and proliferation of immature cells. 	[37–39]
Polyester nonwoven disks	Culture system: static or stirred Surface treatment: not reported Porosity: not reported Pore size: not reported Cytokines: none	<ul style="list-style-type: none"> • Supports stroma development • Increase in weekly output of total hematopoietic cells in wk 3 	[52,53]
Porous gelatin microspheres (CultiSpher G)	Culture system: stirred flask Surface treatment: not reported Porosity: not reported Pore size: not reported Cytokine: IL-3, IL-6, SCF	<ul style="list-style-type: none"> • Preformed stroma • 7.7-, 23.3-, and 9.6-fold increase in total cells, CFU-C, and CD34⁺ cells compared to static control • Distinct hematopoietic islands were observed 	[43]
Porous cellulose microspheres	Culture system: packed-bed bioreactor Surface treatment: not reported Porosity: 95% Pore size: 300 μm Cytokine: none or IL-3, Epo, SCF, GM-CSF, IGF-I	<ul style="list-style-type: none"> • Supports 3-D growth configuration • Supports multilineal differentiation • Supports erythropoiesis in the presence of GFs 	[40,41]

(Continued)

Table 2 Summary of the 3-D scaffolds used for the growth and expansion of HSCs and HPCs (*Continued*)

Scaffold	Experimental condition	Comments	Ref.
Macroporous collagen carriers	Culture system: packed-bed bioreactor Treatment: not reported Porosity: 85% Pore size: 50 μm Cytokine: IL-3, SCF, TPO, Flt-3 ligand	<ul style="list-style-type: none"> • Supports preformed stroma development and ECM production • 100-, 114-, and 6.4-fold increase in MNCs, CFU-GMs, and CAFCS 	[51]

Abbreviations: IL: interleukin; Epo: erythropoietin; SCF: stem cell factor; GFs: growth factors; CFUs: colony-forming units; HPC: hematopoietic progenitor cell; CB MNCs: cord blood mononuclear cells; TPO: thrombopoietin; BM: bone marrow; LTC-IC: long-term culture-initiating cell; CFCs: colony-forming cells; CFU-C: CFU-neutrophil, macrophage, and eosinophil progenitor; GM-CSF: granulocyte/macrophage-colony-stimulating factor; IGF-1: insulin-like growth factor-1; CFU-GM: CFU-granulocyte/macrophage.

Although the use of a scaffold provides greater surface area for cell attachment than the two-dimensional monolayer culture systems and enhances spatial cell-cell and cell-matrix interactions, several issues, including those of scale-up, monitoring, and sampling, remain to be addressed.

One other bioreactor configuration used for the expansion of hematopoietic stem and progenitor cells was the airlift packed-bed bioreactor system, which was designed with a central draft tube.^[25,58] The annular region of the bioreactor was packed with fiberglass disks with a mean diameter of 13.5 μm and a surface area of 109 cm^2 , which were used to establish preformed stroma designed to mimic the *in vivo* environment. The medium was circulated upward through the draft tube by gas sparging and then downward through the packed bed of stacked fiber disks. A 0.5-L system resulted in the total production of 3.6×10^8 BM cells over an 11-week culture period. However, the fiberglass disks did not simulate the natural marrow microenvironment, and adverse effects of bubble shear and breakage caused by gas sparging further hindered cell expansion.^[58]

CONCLUSIONS

The engineering of optimal hematopoietic cell culture systems will require the design of culture systems that mimic the *in vivo* self-regulating bone marrow environment, operate under reliable and reproducible conditions, and offer a broad spectrum of possibilities for different culture strategies as well as the cultivation of various cell types—from stem cells to differentiated cells for gene, cellular, and tissue therapies. Furthermore, the requirement for good manufacturing practices (GMP) compliance of *ex vivo* expanded HSCs for clinical application involves the development of suitable, closed culture systems that can be easily controlled and monitored.

ACKNOWLEDGMENTS

This work was supported by grants to JHDW from NSF BES-9631670 and NASA NAG9-1360. AM would like to thank the Alexander S. Onassis Foundation for providing a scholarship.

ARTICLES OF FURTHER INTEREST

Bioreactors; Extracellular Matrix Scaffolds; Hematopoietic Stem Cells and Assays; Rotary Cell Culture Systems; Tissue Engineering of Bone Marrow; Tissue Engineering Scaffolds

REFERENCES

1. Dexter, T.M.; Allen, T.D.; Lajtha, L.G. Conditions controlling the proliferation of haematopoietic stem cells *in vivo*. *J. Cell. Physiol.* **1977**, *91*, 335–344.
2. Gartner, S.; Kaplan, H.S. Long-term culture of human bone marrow cells. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 4756–4759.
3. Cabral, J.M.S. *Ex vivo* expansion of hematopoietic stem cells in bioreactors. *Biotechnol. Lett.* **2001**, *23*, 741–751.
4. Verfaillie, C.M. *Ex Vivo Expansion of Hematopoietic Stem Cells*. In *Hematopoiesis: A Developmental Approach*; Zon, L.I., Ed.; Oxford University Press: New York, 2001; 119–129.
5. Koller, M.; Bender, J.G.; Papoutsakis, E.T.; Miller, W.M. Effects of synergistic cytokine combinations, low oxygen, and irradiated stroma on the expansion of human cord blood progenitors. *Blood* **1992**, *80* (2), 403–411.
6. Kanai, M.; Hirayama, F.; Yamaguchi, M.; Ohkawara, J.; Sato, N.; Fukazawa, K.; Yamashita, K.; Kuwabara, M.; Ikeda, H.; Ikebuchi, K. Stromal cell-dependent *ex vivo* expansion of human cord blood and augmentation of transplantable stem cell activity. *Bone Marrow Transplant.* **2000**, *26*, 837–844.

7. Emmons, R.V.; Doren, S.; Zujewski, J.; Cottler-Fox, M.; Carter, C.S.; Hines, K.; O'Shaughnessy, J.A.; Leitman, S.F.; Greenblatt, J.J.; Cowan, K.; Dunbar, C.E. Retroviral gene transduction of adult peripheral blood or marrow-derived CD34⁺ cells for six hours without growth factors or on autologous stroma does not improve marking efficiency assessed in vivo. *Blood* **1997**, *89*, 4040–4046.
8. Barnett, M.J.; Eaves, C.J.; Phillips, G.L.; Gascoyne, R.D.; Hogge, D.E.; Horsman, D.E.; Humphries, R.K.; Klingemann, H.G.; Lansdorp, P.M.; Nantel, S.H.; Eaves, A.C. Autografting with cultured marrow in chronic myeloid leukemia: Results of a pilot study. *Blood* **1994**, *84*, 724–732.
9. Coutinho, L.H.; Testa, N.G.; Chang, J.; Morgenstern, G.; Harrison, C.; Dexter, T.M. The use of cultured bone marrow cells in autologous transplantation. *Prog. Clin. Biol. Res.* **1990**, *333*, 415–432.
10. Mayani, H. Composition and function of the hematopoietic microenvironment in human myeloid leukemia. *Leukemia* **1996**, *10*, 1041–1047.
11. Bhatia, M.; McGlave, P.B.; Dewald, G.W.; Blazar, B.R.; Verfaillie, C.M. Abnormal function of the bone marrow microenvironment in chronic myelogenous leukemia: Role of malignant stromal macrophages. *Blood* **1995**, *85*, 3636–3645.
12. Chabannon, C.; Olivero, S.; Blaise, D.; Maraninchi, D.; Viens, P. Ex vivo expansion of human hematopoietic progenitors and cells to support high-dose chemoradiation therapy: Five years of clinical experience. *Cytokines Cell. Mol. Ther.* **2000**, *6*, 97–108.
13. Collins, P.C.; Miller, M.E.; Papoutsakis, E.T. Stirred culture of peripheral and cord blood hematopoietic cells offers advantages over traditional static systems for clinically relevant applications. *Biotechnol. Bioeng.* **1998**, *59*, 534–543.
14. Douay, L. Culture Conditions for Ex Vivo Expansion of Hematopoietic Primitive Cells. In *Autologous Blood and Marrow Transplantation X: Proceedings of the Tenth International Symposium*; Dicke, K., Keating, A., Eds.; Carden Jennings Publishing Co., Ltd.: Charlottesville, Virginia, 2001; 604–614.
15. Giarratana, M.-C.; Kobari, L.; Neildez Nguyen, T.M.A.; Firat, H.; Bouchet, S.; Lopez, M.; Gorin, N.-C.; Thierry, D.; Douay, L. Cell culture bags allow a large extent of ex vivo expansion of LTC-IC and functional mature cells which can subsequently be frozen: Interest for large-scale clinical application. *Bone Marrow Transplant.* **1998**, *22* (7), 707–715.
16. Palsson, B.O.; Paek, S.-H.; Schwartz, R.M.; Palsson, M.A.; Lee, G.-M.; Silver, S.; Emerson, S.G. Expansion of human bone marrow progenitor cells in a high cell density continuous perfusion system. *Bio/Technology* **1993**, *11*, 368–372.
17. Koller, M.; Emerson, S.G.; Palsson, B.O. Large-scale expansion of human stem and progenitor cells from bone marrow mononuclear cells in continuous perfusion cultures. *Blood* **1993**, *82* (2), 378–384.
18. Koller, M.R.; Manchel, I.; Maher, R.J.; Goltry, K.L.; Armstrong, D.; Smith, A.K. Clinical-scale human umbilical cord blood cell expansion in a novel automated perfusion culture system. *Bone Marrow Transplant.* **1998**, *21*, 653–663.
19. Collins, P.; Miller, W.M.; Papoutsakis, E.T. Ex vivo culture systems for hematopoietic cells. *Curr. Opin. Biotechnol.* **1996**, *7* (2), 223–230.
20. Koller, M.R.; Bradley, T.R.; Palsson, B.O. Growth factor consumption and production in perfusion cultures of human bone marrow correlate with specific cell production. *Exp. Hematol.* **1995**, *23*, 1275–1283.
21. Caldwell, J.; Palsson, B.O.; Locey, B.; Emerson, S.G. Culture perfusion schedules influence the metabolic activity and granulocyte-macrophage colony-stimulating factor production rates of human bone marrow stromal cells. *J. Cell. Physiol.* **1991**, *147*, 344–353.
22. Koller, M.R.; Bender, J.G.; Miller, M.E.; Papoutsakis, E.T. Expansion of primitive human hematopoietic progenitors in a perfusion bioreactor system with IL-3, IL-6, and stem cell factor. *Bio/Technology* **1993**, *11*, 358–363.
23. Sandstrom, C.E.; Bender, J.G.; Miller, M.E.; Papoutsakis, E.T. Development of a novel perfusion chamber to retain nonadherent cells and its use for comparison of human “mobilized” peripheral blood mononuclear cell cultures with and without irradiated bone marrow stroma. *Biotechnol. Bioeng.* **1996**, *50*, 493–504.
24. Horner, M.; Miller, M.E.; Ottino, J.M.; Papoutsakis, E.T. Transport in a grooved perfusion flat-bed bioreactor for cell therapy applications. *Biotechnol. Prog.* **1998**, *14*, 689–698.
25. Sardonini, C.A.; Wu, Y.J. Expansion and differentiation of human hematopoietic cells from static cultures through small-scale bioreactors. *Biotechnol. Prog.* **1993**, *9*, 131–137.
26. Koller, M.R.; Manchel, I.; Newsom, B.S.; Palsson, M.A.; Palsson, B.O. Bioreactor expansion of human bone marrow: Comparison of unprocessed, density-separated and CD34-enriched cells. *J. Hematother.* **1995**, *4*, 159–169.
27. Van Zant, G.; Rummel, S.A.; Koller, M.R.; Larson, D.B.; Drubachevsky, I.; Palsson, M.A.; Emerson, S.G. Expansion in bioreactors of human progenitor populations from cord blood and mobilized peripheral blood. *Blood Cells* **1994**, *20*, 482–491.
28. Oh, D.J.; Koller, M.R.; Palsson, B.O. Frequent harvesting from perfused bone marrow cultures results in increased overall cell and progenitor expansion. *Biotechnol. Bioeng.* **1994**, *44*, 609–616.
29. Jaroscak, J.; Goltry, K.; Smith, A.; Waters-Pick, B.; Martin, P.L.; Driscoll, T.A.; Howrey, R.; Chao, N.; Douville, J.; Burhop, S.; Fu, P.; Kurtzberg, J. Augmentation of umbilical cord (UCB) transplantation with ex vivo-expanded UCB cells: Results of a phase I trial using the AastromReplicell System. *Blood* **2003**, *101* (12), 5061–5067.
30. Zandstra, P.W.; Nagy, A. Stem cell bioengineering. *Annu. Rev. Biomed. Eng.* **2001**, *3*, 275–305.
31. Zandstra, P.W.; Eaves, C.J.; Piret, J.M. Expansion of hematopoietic progenitor cell populations in stirred suspension bioreactors of normal human bone marrow cells. *Bio/Technology* **1994**, *12*, 909–914.

32. Pierson, B.A.; Europa, A.F.; Hu, W.S.; Miller, J.S. Production of human natural killer cells for adoptive immunotherapy using a computer-controlled stirred-tank bioreactor. *J. Hematother.* **1996**, *5*, 474–483.
33. McDowell, C.L.; Papoutsakis, E.T. Increased agitation intensity increases CD13 receptor surface content and mRNA levels, and alters the metabolism of HL60 cells cultures in stirred tank bioreactors. *Biotechnol. Bioeng.* **1998**, *60*, 239–250.
34. Nielsen, L.K. Bioreactors for hematopoietic cell culture. *Annu. Rev. Biomed. Eng.* **1999**, *1*, 129–152.
35. Dao, M.A.; Hashimo, K.; Kato, I.; Nolte, J.A. Adhesion to fibronectin maintains regenerative capacity during ex vivo culture and transduction of human hematopoietic stem and progenitor cells. *Blood* **1998**, *92*, 4612–4621.
36. Chute, J.P.; Saini, A.A.; Kampen, R.L.; Wells, M.R.; Davis, T.A. A comparative study of the cell cycle status and primitive cell adhesion molecule profile of human CD34⁺ cells cultured in stroma-free versus porcine microvascular endothelial cell cultures. *Exp. Hematol.* **1999**, *27*, 370–379.
37. Naughton, B.A.; Naughton, G.K. Hematopoiesis on nylon mesh templates—Comparative long-term bone-marrow culture and the influence of stromal support cells. *Ann. N.Y. Acad. Sci.* **1989**, *554*, 125–140.
38. Naughton, B.A.; Jacob, L.; Naughton, G.K. A three-dimensional culture system for the growth of hematopoietic cells. *Prog. Clin. Biol. Res.* **1990**, *333*, 435–445.
39. Naughton, B.A.; Tjota, A.; Sibanda, B.; Naughton, G.K. Hematopoiesis on suspended nylon screen-stromal cell microenvironments. *J. Biomech. Eng., Trans. ASME* **1991**, *113*, 171–177.
40. Mantalaris, A.; Keng, P.; Bourne, P.; Chang, A.Y.C.; Wu, J.H.D. Engineering a human bone marrow model: A case study on ex vivo erythropoiesis. *Biotechnol. Prog.* **1998**, *14*, 126–133.
41. Wang, T.-Y.; Brennan, J.K.; Wu, J.H.D. Multilineal hematopoiesis in a three-dimensional murine long-term bone marrow culture. *Exp. Hematol.* **1995**, *23*, 26–32.
42. Wang, M.; Chen, L.J.; Ni, J.; Weng, C.; Yve, Y. Manufacturing and evaluation of bioactive and biodegradable materials and scaffolds for tissue engineering. *J. Mater. Sci., Mater. Med.* **2001**, *12*, 855–860.
43. Xiong, F.; Chen, Z.; Liu, H.; Xu, Z.; Li, X. Ex vivo expansion of human umbilical cord blood hematopoietic progenitor cells in a novel three-dimensional culture system. *Biotechnol. Lett.* **2002**, *24*, 1421–1426.
44. Bagley, J.; Rosenzweig, M.; Mark, D.F.; Pykett, M.J. Extended culture of multipotent hematopoietic progenitors without cytokine augmentation in a novel three-dimensional device. *Exp. Hematol.* **1999**, *27*, 496–504.
45. Banu, N.; Rosenzweig, M.; Kim, H.; Bagley, J.; Pykett, M. Cytokine-augmented culture of haematopoietic progenitor cells in a novel three-dimensional cell growth matrix. *Cytokine* **2001**, *13* (6), 349–358.
46. Knospe, W.; Hussein, S.G.; Fried, W. Hematopoiesis on cellulose ester membranes. XI. Induction of new bone and a hematopoietic microenvironment by matrix factors secreted by marrow stromal cells. *Blood* **1989**, *74* (1), 66–70.
47. Levee, M.; Lee, G.-M.; Paek, S.-H.; Palsson, B.O. Microencapsulated human bone marrow cultures: A potential culture system for the clonal outgrowth of hematopoietic progenitor cells. *Biotechnol. Bioeng.* **1994**, *43*, 734–739.
48. Tun, T.; Miyoshi, H.; Aung, T.; Takahashi, S.; Shimizu, R.; Kuroha, T.; Yamamoto, M.; Ohshima, N. Effect of growth factors on ex vivo bone marrow cell expansion using three-dimensional matrix support. *Artif. Organs* **2002**, *26* (4), 333–339.
49. Li, Y.; Kniss, D.A.; Yang, S.-T.; Lasky, L.C. Human cord cell hematopoiesis in three-dimensional nonwoven fibrous matrices: In vitro simulation of the marrow microenvironment. *J. Hematother. Stem Cell Res.* **2001**, *10*, 355–368.
50. Godbey, W.; Atala, A. In vitro systems for tissue engineering. *Ann. N.Y. Acad. Sci.* **2002**, *961*, 10–26.
51. Jelinek, N.; Schmidt, S.; Hilbert, U.; Thoma, S.; Biselli, M.; Wandrey, C. Novel bioreactors for the ex vivo cultivation of hematopoietic cells. *Eng. Life Sci.* **2002**, *2* (1), 15–18.
52. Sasaki, T.; Takagi, M.; Yoshida, T. 3D culture of murine hematopoietic cells with spatial development of stromal cells in nonwoven fabrics. *Cytotherapy* **2002**, *4* (3), 285–291.
53. Tomimori, Y.; Takagi, M.; Yoshida, T. The construction of an in vitro three-dimensional hematopoietic microenvironment for mouse bone marrow cells employing porous carriers. *Cytotechnology* **2000**, *34*, 121–130.
54. Rosenzweig, M.; Pykett, M.J.; Marks, D.F.; Johnson, R.P. Enhanced maintenance and retroviral transduction of primitive hematopoietic progenitor cells using a novel three-dimensional culture system. *Gene Ther.* **1997**, *4*, 928–936.
55. Washburn, N.; Simon, C.G.; Tona, A.; Elgendy, H.M.; Karim, A.; Amis, E.J. Co-extrusion of biocompatible polymers for scaffolds with co-continuous morphology. *J. Biomed. Mater. Res.* **2002**, *60*, 20–29.
56. Liao, C.-J.; Chen, C.-F.; Chen, J.-H.; Chiang, S.-F.; Lin, Y.-J.; Chang, K.-Y. Fabrication of porous biodegradable polymer scaffolds using a solvent merging/particulate leaching method. *J. Biomed. Mater. Res.* **2002**, *59*, 678–681.
57. Lin, H.-R.; Kuo, C.-J.; Yang, C.Y.; Shaw, S.-Y.; Wu, Y.-J. Preparation of macroporous biodegradable PLGA scaffolds for cell attachment with the use of mixed salts as porogen additives. *J. Biomed. Mater. Res.* **2002**, *63*, 271–279.
58. Highfill, J.; Haley, S.D.; Kompala, D.S. Large-scale production of murine bone marrow cells in an airlift packed bed bioreactor. *Biotechnol. Bioeng.* **1996**, *50*, 514–520.

Tissue Engineering of Breast

Cheryl T. Gomillion

Cheryl A. Parzel

Department of Bioengineering, Clemson University, Clemson, South Carolina, U.S.A.

Richard L. White, Jr.

Division of Surgical Oncology, Carolinas Medical Center, Charlotte, North Carolina, U.S.A.

Karen J.L. Burg

Department of Bioengineering, Clemson University, Clemson, South Carolina, U.S.A.

INTRODUCTION

An estimated 269,000 new cases of breast cancer were diagnosed in women in 2005.^[1] Breast cancer is the most frequently diagnosed cancer in American women, accounting for approximately 31% of new cancer cases yearly.^[1] Breast cancer poses an enormous threat because of the high concentration of local blood and lymphatic vessels through which cancer can metastasize. The actively differentiating epithelial cells that line the breast ductal structures are responsive to hormonal changes and thus are susceptible to mutation during cell division cycles.^[2] Upon formation of epithelial-derived breast carcinomas, women have several treatment options, including radiation therapy, chemotherapy, hormonal therapy, and surgery, the latter involving the removal of a tumor and a surrounding margin of normal breast tissue. Management of breast cancer is often a multistep process, generally consisting of primary, neoadjuvant (presurgical), and adjuvant (postsurgical) therapies. Neoadjuvant therapies, which may be local or systemic, have been used successfully to shrink tumors prior to removal, whereas adjuvant treatments provide a mechanism of targeting and eradicating malignant cells residing at surgical margins or cells that have metastasized. Although the use of neoadjuvant therapies does not affect patient survival outcome, increased use of such therapies is correlated with an increase in the lumpectomy rate. These treatment methods have been proved effective as preventive therapies in high-risk patients to avert tumor recurrence or disease of the contralateral breast.^[3] The removal or partial removal of a breast by way of mastectomy or lumpectomy can be taxing, physically and psychologically; nevertheless, because of the effectiveness of surgical tumor removal, lumpectomies and mastectomies are almost always part of a treatment regimen.

Following surgical tumor removal, women may select among a variety of reconstruction procedures. Because of the complications associated with current reconstruction methods, researchers are investigating methods to restore tissue function and aesthetics permanently with biological substitutes.

The discussion of the role of tissue engineering and biomaterials to breast cancer applications in this article is threefold. First, the development of a breast reconstruction procedure using biocompatible and biodegradable materials could result in the development of an implant that can stimulate the regeneration of host adipose cells and fill a soft-tissue void in the breast. With successful implantation and cellular control, permanent biological restoration of adipose tissue can be achieved using autologous patient cells. Second, the use of these materials as drug-delivery vehicles could lead to prevention of tumor recurrence in the breast. Last, the construction of *in vitro* engineered breast tissue has the potential to aid in cancer research and drug testing by providing a model in which concurrent cellular processes that lead to tumor growth and metastasis may be monitored.

CURRENT TREATMENT AND RECONSTRUCTION OPTIONS

Breast cancer treatments are multifaceted and vary widely from patient to patient. Most treatment plans, however, consist of surgical tumor removal coupled with presurgical or postsurgical therapies. The two major types of surgeries are lumpectomies and mastectomies. A lumpectomy is often preferred by patients because unlike a mastectomy, which results in the loss of form and function due to a high degree of soft-tissue removal, a lumpectomy is a breast-conserving surgery. To help patients cope with the psychological, emotional, and physical aspects of a mastectomy, breast reconstruction procedures are widely performed. Two primary categories of breast reconstruction methods exist: prosthetic and autologous.

Prosthetic Methods

Prosthetic or acellular methods of reconstruction include breast prostheses and breast implants, and tissue expansion devices followed by the insertion of implants. Breast prostheses are custom-molded to be

worn outside the body and thus are not permanent corrections of breast defects. Furthermore, prostheses can limit the physical activities of patients and may give breasts a distorted appearance.

Permanent alloplastic implants are constructed from a silicone shell filled with either silicone or saline. The first step in implant reconstruction (if it is not carried out immediately following tumor removal) is the expansion of the skin and muscle in the breast region. An expander is placed under the chest muscle, and saline is added gradually over the course of weeks or months until the skin and muscle have stretched sufficiently for the implant to be placed. Implants, generally placed within 6–8 months of initial tissue expansion, often require replacement after several years, and with them come a series of possible complications, including bleeding, fluid collection, and capsular contraction due to scarring and fibrosis. Additionally, there is a low level of concern that implants can obstruct mammographic images and potentially prevent detection and diagnosis of recurring tumors.^[4]

Autologous Methods

Traditional autologous methods often consist of the use of flaps of tissue removed from the abdomen, back, or buttocks that are transferred to the defect site to reconstruct the breast mound. Autologous donor tissue can remain coupled to its original blood supply, or it can be completely excised and subsequently reattached to the vascular network in the breast region. Complications of autologous flap procedures include scarring, infection, and donor-site morbidity.^[5] Soft-tissue transplants similarly can be used to fill tissue voids following tumor removal. Adipose tissue can be removed from a fatty region of the body and transplanted by way of injection to a defect site. The concept of autologous fat transplantation is logical in that fatty tissue is replaced with fatty tissue, but the use of this technique has not been consistently successful in practice. Unfortunately, adipocytes are frequently damaged as a result of harvesting methods such as liposuction. These damaging effects, in addition to poor vascularization and integration of the transplanted tissue, often lead to rapid resorption of more than 50% of the original graft volume and eventual deformation of the transplant site.^[6]

To minimize the effects of cellular and tissue resorption, researchers have more recently focused on developing cellular scaffolds, where cells are seeded onto an absorbable template or scaffold. The scaffold is designed to induce cell growth and to absorb in the body gradually once implanted, paralleling tissue growth. The scaffold material provides a surface to

which the cells attach and therefore diminishes the likelihood of cellular resorption. The selection of cell type and cell source and the design of the optimal scaffold material are the subjects of many research programs.

BREAST RECONSTRUCTION AND TISSUE ENGINEERING

Significant improvements have been made over the past decade in current prosthetic and autologous reconstruction procedures with regard to aesthetics and ensuing complications. Tissue engineering practices, however, have the potential to provide patients with a minimally invasive reconstruction alternative that will render current complications obsolete. A permanent, biologically based restoration of breast tissue will also preclude the need for multiple surgeries and recovery time. Current methods for breast reconstruction typically involve invasive medical procedures that can be costly and, at times, ineffective in achieving the desired shape, size, texture, and volume of the natural breast tissue. So that the patient avoids pain, scarring, and recovery time associated with multiple procedures, tissue engineering strategies are being investigated.

Anatomy of the Breast

The framework of the breast, shown in Fig. 1, is comprised of three major tissue types: glandular, fibrous, and adipose. Among these highly innervated tissues is a local network of blood and lymphatic vessels. Also contained within the breast are 15–20 lobes that branch into smaller lobules and that have individual excretory ducts, also known as lactiferous tubules. Lobes are connected by the fibrous tissue of the breast and are separated by intervals of adipose tissue. The volume of the breast is composed primarily of adipose tissue. The fundamental unit of adipose tissue is a fat lobule, which consists of approximately 10^2 – 10^3 mature adipocytes. Comprised of mostly ductal and lobed structures, the mammary glands are the major functional units of the breast and are embedded in a layer of contractile epithelial cells.^[7]

Cellular Components of Adipose Tissue

Adipose tissue is a highly specialized connective tissue that serves to insulate and cushion the body. The principal cellular component of adipose tissue is a collection of lipid-filled cells known as adipocytes, which are adhered to a collagen fiber matrix. Other cellular components of adipose tissue include



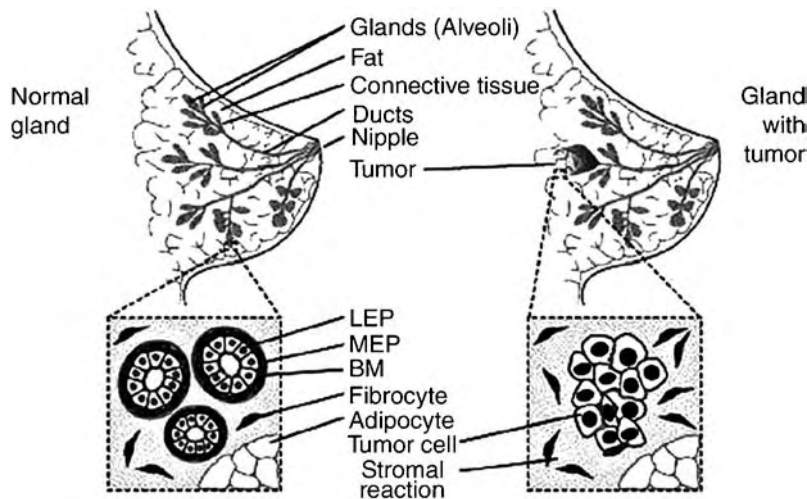


Fig. 1 Structure and 3-D model of mammary gland. Luminal epithelial cells (LEP) forming the acinus are regulated by homotypic interactions with neighboring luminal cells, by heterotypic interactions with myoepithelial cells (MEP), and by interactions with the basement membrane. The bilayered structure is surrounded by a complex basement membrane and a stroma comprised of stromal extracellular matrix, fibroblasts, and adipocytes, as well as nerves, blood, and lymphatic vessels (not depicted). This structure is disrupted in tumors.

fibroblastic connective tissue cells, leukocytes, macrophages, and preadipocytes (committed precursor cells that, while accumulating lipid, will differentiate into mature adipocytes; the classic, mature adipocyte found in the white adipose tissue of the breast has one large fat vacuole and a sliver-shape nucleus located against the perimeter of the cell). Although white adipose tissue is not as highly vascularized as its brown counterpart, each fat cell in the white adipose tissue construct is in contact with at least one capillary, which provides a vascular network for continued growth of the tissue. As a result, the vascular cells that make up this network are of great significance to adipose tissue engineering.

Cell Sources

As noted earlier, several problems are associated with the use of autologous adipose tissue transplantations, including migration of the fat graft tissue, destruction of adipocytes during harvesting methods, and tissue resorption due to insufficient revascularization. One possible solution to these problems may be the use of preadipocytes rather than mature adipocytes or autologous adipose tissue. Preadipocytes are mesenchymal adipocyte precursor cells that are committed to becoming mature adipose cells. The fusiform shape of preadipocytes is maintained throughout the proliferation process until growth-arrest initiates the onset of differentiation or maturation. Cells in the differentiation stage begin to accumulate triglycerides, transform into a more spherical morphology, and begin to express genes typical of mature adipocytes. A differentiation-inducing medium containing insulin, dexamethasone, and 3-isobutyl-1-methylxanthine (IBMX) can accelerate *in vitro* phenotypic changes.^[8]

It has been shown that preadipocytes more readily survive transplantation conditions than do mature adipocytes.^[9] For this reason, a supply of preadipocytes is now needed for use in studies and research in the areas of tissue regeneration. Research assessing the use of stem cells in producing adipose tissue has become increasingly common. Because fat tissue located in different areas in the body behaves differently, it is likely that a ready supply of cells that displays characteristics of multipotency (i.e., cells that can give rise to numerous mature cell types) will be needed. Stem cells not only provide, in theory, an unlimited supply of cells, but also increase the ability of researchers to define and control cellular processes.

One of the basic requirements for the development of suitable tissue replacements is an adequate source of viable cells to be used for tissue production. Cell types that have the potential to be used for repair and regeneration include mature cells obtained directly from a patient and stem cells (adult or embryonic). Adult stem cells are located in niches of the body—specifically in the brain, bone marrow, mammary tissue, gums, and intestines, among others. Marrow stromal or mesenchymal stem cells, which arise from haematopoietic (blood) precursor cells, are capable of differentiating into osteoblasts, chondrocytes, and adipocytes. The use of mature cells obtained from the patient minimizes the need for immunosuppressive therapy postimplantation; these cells are differentiated and committed to a specific cell type, however, and thus provide little potential for further proliferation. Embryonic stem cells, on the other hand, are by definition a population of cells that can be harvested from the inner cell mass of a blastocyst (16-cell mass) and are capable of providing replacement cells for a specific differentiated cell type. Embryonic stem cells are different from other cell types due to two primary distinguishing characteristics. First, embryonic stem

cells are capable of self-renewal in long-term cultures. Second, embryonic stem cells have the potential to differentiate into numerous specialized cell types when exposed to a variety of epigenetic factors.

Scaffold Constructs

Scaffolds used in tissue engineering may be implanted either by surgical incision or by injection. The materials used for constructing the scaffolds, whether they are implanted or injected, must possess several essential properties. These materials should be biocompatible, biodegradable, easily manufactured and sterilized, easily managed during surgical procedures, and capable of guiding the differentiation, migration, and proliferation of cells. Additionally, they should possess mechanical properties similar to those of natural tissue and have nontoxic byproducts. To allow adequate cellular growth, the construct design should be porous to allow the permeation of nutrients and waste products in and out of the construct. Thus, the size of the scaffold should also be taken into consideration, not only with respect to the method of implantation, but also with respect to cellular nourishment and mass transport through the scaffold.

In Vitro Studies

All cells in the human body exhibit an identical genotype—that is, they carry the same genetic information within their nuclei. Epigenetic regulators such as cell–extracellular matrix interactions, cell–cell interactions, growth factors, electrical stimuli, and physical force are

all responsible for cellular expression of a specific phenotype. Therefore, it is crucial to understand the environmental factors responsible for cellular processes and protein production so that *in vitro* culture systems can adequately mimic the *in vivo* environment. Though static culture conditions remain important stepping stones, dynamic bioreactors are becoming more prominent in the laboratory because they better simulate *in vivo* flow, pressure, and diffusion characteristics.

Prior to implantation of scaffolds, it is important to conduct *in vitro* studies to understand how material properties and surface chemistry affect a specific cell type. The response of preadipocytes when seeded on a 2-D polystyrene surface, for example, drastically differs from the response when seeded on 3-D spherical surfaces.^[8] Through scanning electron microscopy (SEM) images, shown in Fig. 2, significant differences in the morphology of various 3-D surfaces are observed. Lipid production, cell number, cell viability, and gene expression can all be altered by the capability of adipocytes to form adhesions to surfaces and the orientation with which they do so. Consequently, material characterization techniques are important so that gene expression can be predicted and better controlled. Among the most common material properties to consider are denaturation and melting temperatures, molecular weight, chemistry, polydispersity index, stiffness, surface roughness, viscosity, and microscopic structure.

Cellular response of adipose precursor cells can also be influenced by stimuli existing in co-culture systems. It has been shown, for example, that the presence of MAC-T cells (normal bovine mammary epithelial cells)

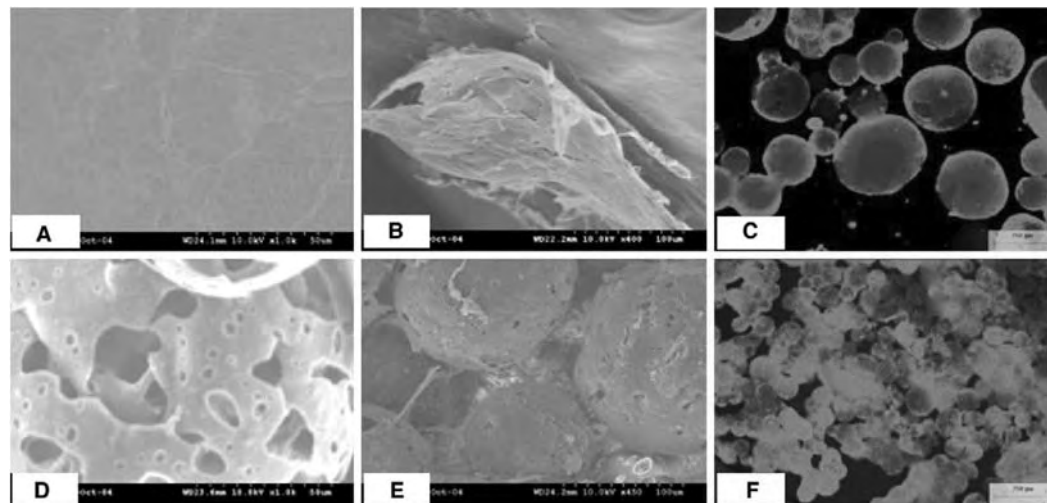


Fig. 2 Preadipocyte culture on poly(lactide) (PL) and gelatin microcarriers. (A) Scanning electron microscopy (SEM) image of PL bead surface. (B) SEM image of 3T3-L1 cells attached to PL beads. (D) SEM image of gelatin microcarrier surface. (E) SEM image of 3T3-L1 cells attached to gelatin microcarriers. (C) and (F) Results of LIVE/DEAD viability assay—viable cells attached to PL and gelatin surfaces, respectively, fluoresce green.

in vitro inhibits the differentiation of D1 (murine marrow stromal stem cells) into adipocytes, even in the presence of an adipogenic differentiation inducer.^[10] It has been hypothesized that pH changes in culture medium resulting from proliferation of MAC-T cells are responsible for reduced lipid formation.

CURRENT ADIPOSE TISSUE ENGINEERING STRATEGIES

One proposed strategy for engineering adipose tissue is through the isolation of preadipocytes and endothelial cells, procured through enzymatic digestion of fat obtained from patients during liposuction or biopsy.^[9] A microvascular network is formed by the capillary endothelial cells, and this network and the preadipocytes are cultured with the appropriate growth factors on a polymeric scaffold.^[6] The scaffold would be implanted into the patient to fill the defect site; in the case of breast tissue reconstruction, the cellular scaffold would be placed in the breast envelope. Ideally, as the scaffold gradually absorbs, the preadipocytes would grow, proliferate, and eventually mature into adipose tissue. In a study conducted by Patrick and associates, differentiation of preadipocytes into mature adipocytes was demonstrated using *in vitro* cultures of preadipocytes harvested from the epididymal fat pads of Sprague-Dawley and Lewis rats.^[11] Adipocyte production was identified through the use of histological staining and fluorescence labeling, and the cells, seeded on polylactide-based scaffolds, were viable for extended periods postimplantation. Modifications to this scaffold-guided form of breast tissue regeneration have been proposed as well. In a patent issued to Vacanti and associates, a hydrogel polymer solution used as an injectable polymer is suggested rather than the use of the typical scaffold construct.^[12] Burg and coworkers have suggested the concurrent use of absorbable tissue expanders to create more space for the implantation of cellular scaffolds and to allow the serial development of breast tissue in large defects.^[13]

In another study, Cui and coworkers assessed steroid-induced adipogenesis of a multipotent mesenchymal stem-cell line obtained from mouse bone marrow.^[14] The bone-marrow cell line, cloned from mouse bone-marrow stroma, was treated with the steroid dexamethasone, which caused the cells to produce vesicles that contained triglyceride and to express adipocyte lipid binding protein (aP2), a gene that is expressed in mature fat cells. The results of the study indicated that the dexamethasone, which was applied at varying concentrations and for varying lengths of time, stimulated the differentiation of cells in bone-marrow stroma into adipocytes and the accumulation of fat in the marrow itself, suppressing differentiation into

osteoblasts. Further investigation into the properties of this steroid and its effect on stem cells and their production of adipose tissue, along with investigation of the other methods for adipose tissue engineering, could lead to a more efficient method of regenerating tissue for breast reconstructions.

To date, *de novo* lipogenesis in animal models has been minimal, and it has not yet been determined whether the presence of fat tissue in an implant region is due to migration of preadipocytes via induction by growth factors [i.e., basic fibroblast growth factor (bFGF)] or a result of proliferation of implanted precursor cells.^[15] Cell labeling prior to implantation may be a necessary next step to better understand the proliferative capacity of implanted cells and their capability to form new fat tissue.

DEVELOPMENT OF INJECTABLE COMPOSITE SYSTEMS

Early efforts to engineer breast tissue were limited due to insufficient vascularization of scaffold materials and the inability of larger scaffolds to fill defect sites properly. As an alternative to the common synthetic materials used for scaffold production, natural injectable materials—including collagens, alginates, and gelatins—have been suggested for use in breast tissue reconstruction. Injectable scaffolds have adaptable geometries; when injected collectively, they are able to conform to an infinite amount of defect shapes and sizes, and they require minimally invasive surgical procedures for implantation. Additionally, injection of collective groups of scaffolds allows the incorporation and timed release of bioactive molecules in specific scaffolds, provides support for anchorage-dependent cells, allows mass and fluid transfer between scaffolds, and allows the inclusion and tailoring of individual scaffolds to accommodate multiple cell types. Natural scaffolds degrade into natural metabolites and provide biologically familiar surfaces to which cells will readily attach; they are quickly remodeled in the body, however, usually in an unpredictable manner. Synthetic materials, by contrast, while not providing the familiar surface, can easily be processed to provide the appropriate mechanical properties and degradation times. Thus, a combination of natural and synthetic materials of injectable size can provide an optimized device for breast reconstruction.

Both preadipocytes and adipocytes are anchorage-dependent cells that must be seeded on an appropriate scaffold surface that provides traction for cell differentiation and proliferation to occur. Because preadipocytes and adipocytes are generally function limited when embedded directly in a gel, an injectable composite system has been proposed and tested.^[16] In this system, cells are

seeded on biodegradable beads of an injectable size to create cell carriers; then these cellular constructs are mixed with a hydrogel delivery medium, resulting in a composite that may be injected into a patient through a syringe at the defect site. Hydrogel-based microspheres can be constructed with an electrostatic bead generator to facilitate ejection through a small-gauge needle. Synthetic beads, such as those comprised of polylactide, can be fabricated using an oil-in-water emulsification process. Cell-seeding protocols can be carried out following the harvest, separation, and expansion of preadipocytes. After fat tissue is removed from the host animal in a biopsy procedure, cells can be isolated following a collagenase digestion process during which extracellular matrix proteins are removed. Cells are expanded in 2-D, polystyrene culture vessels, with the intention of preventing terminal differentiation of preadipocytes; at confluency, cells can be seeded on injectable scaffolds. The scaffolds and delivery medium will then degrade over time as the implanted cells proliferate and form new tissue. The most advantageous aspect of this technology is that these injectable devices require a minimally invasive surgical procedure, effectively eliminating any additional trauma a patient will have to undergo following cancer treatment. Further studies are required to define parameters fully regarding cell seeding densities that will provide optimal conditions for inducing *de novo* adipogenesis and sustaining long-term viability of fat tissue.

LARGE-ANIMAL MODEL SELECTION

Before any of these tissue engineering technologies may be applied to human medicine, they must be evaluated within appropriate animal models. Several researchers have demonstrated the ability to regenerate adipose tissue to repair soft-tissue defects in small animals, such as rats and mice. Halberstadt and coworkers have also demonstrated the feasibility of adipose tissue regeneration using an ovine model.^[17] Further investigation of this technology within a large animal that has biological characteristics comparable with those of humans is required, however. An animal model's mammary-tissue anatomy and physiology must be similar to those of humans, and the animal must be large enough to display similar mass transport and diffusion rates.^[18,19] The size of the animal is also important because a selected area for implantation of cellular constructs must be large enough to ensure space for sufficient replication of *in vivo* experimentation. Last, the life span of the selected model should be long enough to allow extensive studies during the relevant phase.

Bovine animal models have been used in the evaluation of tissue engineered devices for breast tissue

reconstruction.^[20,21] The bovine mammary gland and the normal human breast have a similar anatomical structure and tissue type. Histologically, bovine mammary tissue is more similar to that of humans than is the mammary tissue of traditional animal models, such as mice and rats. Ductal structures in the human and cow, for example, are surrounded by relatively dense stromal tissue, unlike the ducts in the mouse, which are almost completely enclosed by adipocytes. Regarding injectable tissue engineered constructs, the dairy cow could serve as a more clinically relevant animal. A larger defect site that would more closely resemble those that follow lumpectomies could be mimicked in the bovine model. Likely, large defects will require a slow progression of *de novo* adipogenesis to allow for neovascularization and integration with host tissue. This can be achieved using a system of serial injections that a smaller, traditional animal model could not likely support. If fat tissue can be engineered during a series of injections, problems related to mass transport, maturation and integration of newly formed tissue, and inflammation can be minimized.

FUTURE AREAS OF DEVELOPMENT

Drug Delivery

Further development of adipose tissue engineering strategies for breast tissue reconstruction may lead to more complex systems that incorporate therapeutic components. Studies have shown that common adjuvant therapies, both systemic and localized, greatly reduce the risk of tumor recurrence, for example.^[22] Since 1985, mortality rates in the United States have declined approximately 8%, largely due to better use of adjuvant therapies.^[23] Tumors can recur, stemming either from cells that were not removed during surgical treatments or from those that have metastasized. Currently, the majority of women opt for lumpectomy (breast-conserving surgery) because this procedure maintains the appearance of the breast. This surgical method may not result in a cancer-free surgical margin, requiring the sacrifice of additional breast tissue.^[24] Despite resection to cancer-free margins, local recurrence remains an issue. Thus, the incorporation of antitumor agents with implantable breast tissue engineering devices that can complement traditional adjuvant treatment regimens may further aid in minimizing local recurrence.

Current adjuvant treatments include chemotherapy, radiation therapy, and hormonal therapy. Polychemotherapeutic treatments consist of a series of systemically administered drugs. Currently, 100 different types of drugs are used for chemotherapy.^[25] The target of a chemotherapeutic drug is a rapidly dividing cell.

Unfortunately, drugs cannot always differentiate between normal and tumorigenic cells; thus, a balance must be met to destroy cancerous cells while preventing extraneous side effects. Most drugs target cells in the M (mitosis) or S (DNA synthesis) phases of the cell cycle when cells are actively dividing, as opposed to the resting phases (G_0 , G_1 , G_2) of the cell cycle.

Like chemotherapy, radiation therapy acts on cells that are actively dividing in the M and S stages of the cell cycle. Unlike chemotherapy, however, radiation therapy is not administered systemically but is used to treat localized regions of the body. Post-mastectomy radiation therapy (PMRT) practices have changed over the years. Generally, PMRT is administered only to women at high risk for tumor recurrence. This risk includes tumors that are originally 5 cm or larger, tumors that are affixed to the skin or chest wall, and tumors whose surgical margins test positive for residual cancer. Furthermore, PMRT is often used in conjunction with polychemotherapeutic treatments. There are several categories of radiation, all with specific energies and applications.^[26] Some common types of radiation include high-energy photons from radioactive sources, electron beams, protons, and neutrons. Radiation is administered as either external beam or internal radiation, external beam being the most widely used. Toxicity associated with radiation therapy includes fibrosis of the treated area, color changes, and (rarely) persistent discomfort. These and other toxicities are usually confined to the area of treatment.

Hormonal treatments are not cell cycle dependent. Rather, the goal of hormonal therapy is to compete with estrogen for estrogen receptor binding sites or to reduce the amount of circulating estrogen.^[27] Therefore, hormonal therapies are given only to patients whose tumor cells express estrogen receptors. The primary advantages of hormonal therapies are that they have relatively low toxicity and result in an approximately 50% decrease in risk of recurrence. Tamoxifen is the most widely used hormonal therapy. Tamoxifen blocks estrogen receptor sites and thus prevents a pathway leading to estrogen-regulated genes that contribute to cell proliferation. The duration of tamoxifen treatments is optimal at about 5 years postsurgery. Common side effects of this treatment include estrogen withdrawal-related hot flashes and vaginal bleeding, as well as depression, headache, insomnia, and dizziness.^[26] Recent evidence has demonstrated that aromatase inhibitors such as arimidex have a better toxicity profile and modest improvement in efficacy.^[28]

The ability to deliver drug-loaded devices with minimally invasive techniques to the breast locale could lead to prevention of systemic drug circulation and toxicity to healthy tissues. It is believed that the physiological changes that occur in most cancerous

tissues can actually facilitate the positive effects of drugs delivered to the area. The “leaky vasculature” that results from increased angiogenesis works synergistically with poor lymphatic transport to allow for better drug retention and potency in the tumor region.^[29] In fact, gold-coated polymeric nanoshells have been tailored to respond to specific wavelengths of light such that they may be injected into the vasculature, allowed to “leak” into the tumor area, and then heated via the external application of light. Such “homing” devices and timed-release polymeric drug delivery systems are frequent subjects of research investigations; few studies have assessed such systems as components in tissue engineered devices, however.

3-D *In Vitro* Testing Systems

Tissue engineering strategies may also be employed in the development of *in vitro* breast tissue models that have potential for use in testing of drug therapies and vaccines.^[10] With these models, researchers can also study the physical and chemical interactions that occur among cells and matrix components to better understand the progression of transformed epithelial cells into tumors and the ultimate metastasis of tumor cells. Thus, methods of breast tissue engineering may be useful in the treatment and prevention of breast cancer. Using 3-D breast tissue systems, breast cancer cells and normal human mammary cells can be observed in a variety of relevant conditions to deduce why variations in subsequent cellular processes and matrix mechanics lead to cancer progression and, thus, what therapies will be useful in prevention as well as a cure.

CONCLUSION

Several critical challenges remain before breast tissue regeneration can be put to practical use. First, more basic tissue engineering methods and procedures must be developed. Without suitable biomaterials, cell supplies, and methodologies for producing the needed tissue, the breast tissue engineering process cannot be completed. Along with obtaining a sufficient supply of cells, it must be determined whether stem cells are in fact the best source for cells, and if so, the most clinically and scientifically sensible stem cell retrieval site must be ascertained. Additionally, the controversies that currently surround the use of stem cells, particularly embryonic stem cells, must be resolved before the research can proceed in an efficient manner. Other challenges include perfecting the engineered adipogenesis process. The use of preadipocytes has yielded positive results, but there are still unanswered questions—for example, how one can identify whether the

preadipocyte cells do in fact form mature adipocytes, and how one can determine what characteristics truly affect the performance of the preadipocytes during the adipogenesis process (i.e., do cells from different locations in the body respond the same way to culturing techniques?). Other challenges include the formation of functional glandular and tubular structures, the use of coculture environments, and the integration of the implanted tissue with the host vasculature. Ideally, the investigation of these ideas will one day result in a precise, clear-cut method for breast tissue regeneration that will eliminate the need for painful and invasive reconstruction, and restore the aesthetics and function of the breast.

ARTICLES OF FURTHER INTEREST

Tissue Engineering; Tissue Engineering, Fat

REFERENCES

- American Cancer Society. Breast Cancer Facts and Figures. Available at: www.cancer.org. Accessed on September 2006.
- Tozeren, A.; Coward, C.W.; Petushi, S.P. Origins and evolution of cell phenotypes in breast tumors. *J. Theor. Biol.* **2005**, *233* (1), 43–54.
- Available at: www.cancer.gov/cancertopics/factsheet/Therapy/adjuvant-breast. Accessed on May 2006.
- Taylor, C.W.; Horgan, K.; Dodwell, D. Oncological aspects of breast reconstruction. *Breast* **2005**, *14* (2), 118–130.
- Noda, S.; Eberlein, T.J.; Eriksson, E. Breast reconstruction. *Cancer* **1994**, *74*, 376–380.
- Holder, W.D., Jr; Gruber, H.E.; Moore, A.L.; et al. Cellular ingrowth and thickness changes in poly-lactide and polyglycolide matrices implanted subcutaneously in the rat. *J. Biomed. Mater. Res.* **1998**, *41* (3), 412–421.
- Gray, H. The Urogenital System, 30th Ed. In *Anatomy of the Human Body*; Lea & Febiger: Philadelphia, PA, 1985; 1581–1586.
- Cavin, A.N. *Adipocyte Response to Injectable Breast Tissue Engineering Scaffolds*; Clemson University Press: Clemson, SC, 2005.
- Patrick, C.W., Jr; Mikos, A.G.; McIntire, L.V. Tissue engineered adipose tissue, 1st Ed. In *Frontiers in Tissue Engineering*; Pergamon: Oxford, UK, 1998; 369–382.
- Burg, K.J.L.; Cavin, A.N.; Gomillion, C.T.; et al. 3-D Test Systems: New Tools for Unlocking the Mysteries of Breast Cancer. Presented at: Mammary Gland Biology Gordon Research Conference, Providence, Rhode Island, 2005.
- Patrick, C.W., Jr; Chauvin, P.B.; Hopley, J.; et al. Preadipocyte seeded PLGA scaffolds for adipose tissue engineering. *Tissue Eng.* **1999**, *5* (2), 139–151.
- Vacanti, J.P.; Atala, A.; Mooney, D.J.; et al. Breast Tissue Engineering. U.S. Patent 5,716,404, February 10, 1998.
- Burg, K.J.L.; Halberstadt, C.R.; Holder, W.D., Jr. Absorbable Tissue Expander. U.S. Patent 6,206,930, March 27, 2001.
- Cui, Q.; Wang, G.J.; Balian, G. Steroid-induced adipogenesis in a pluripotential cell line from bone marrow. *J. Bone Joint Surg. Am.* **1997**, *79* (7), 1054–1063.
- Tabata, Y.; Miyao, M.; Inamoto, T.; et al. De novo formation of adipose tissue by controlled release of basic fibroblast growth factor. *Tissue Eng.* **2000**, *6* (3), 279–289.
- Burg, K.J.L.; Austin, C.E.; Culberson, C.R.; et al. A novel approach to tissue engineering: injectable composites. *Transactions of the 2000 World Biomaterials Congress*; Kona: Hawaii, 2000.
- Halberstadt, C.; Austin, C.; Rowley, J.; et al. A hydrogel material for plastic and reconstructive applications injected into the subcutaneous space of a sheep. *Tissue Eng.* **2002**, *8* (2), 309–319.
- Patrick, C.W. Breast tissue engineering. *Annu. Rev. Biomed. Eng.* **2004**, *6*, 109–130.
- Patrick, C.W., Jr. Adipose tissue engineering: the future of breast and soft tissue reconstruction following tumor resection. *Semin. Surg. Oncol.* **2000**, *19* (3), 302–311.
- Gomillion, C.T.; Burg, K.J.L. Evaluation of tissue engineered injectable devices for breast tissue engineering. In *Transactions of the 30th Annual Meeting of the Society for Biomaterials*; Memphis: Tennessee, 2005.
- Parzel, C.A.; Cavin, A.N.; Ellis, S.E.; et al. *Analysis of bovine intradermal tissue response to injectable tissue engineering composites*. In *Transactions of the 30th Annual Meeting of the Society for Biomaterials*; Memphis: Tennessee, 2005.
- Jatoi, I.; Miller, A.B. Why is breast-cancer mortality declining? *Lancet Oncol.* **2003**, *4* (4), 251–254.
- Muss, H.B. Adjuvant therapy for older women with breast cancer. *Breast* **2003**, *12* (6), 550–557.
- Singletary, S.E. Surgical margins in patients with early-stage breast cancer treated with breast conservation therapy. *Am. J. Surg.* **2002**, *184* (5), 383–393.
- Available at: www.cancer.org/docroot/ETO/content/ETO_1_2X_Chemotherapy_What_It_Is_How_It_Helps.asp. Accessed on May 2006.
- Dow, K.H. *Contemporary Issues in Breast Cancer: A Nursing Perspective*, 2nd Ed.; Jones and Bartlett Publishers International: London, UK, 2004; 349.
- O'Regan, R.; Jordan, V.C.; Gradishar, W.J. Tamoxifen and contralateral breast cancer. *J. Am. Coll. Surg.* **1999**, *188* (6), 678–683.
- Chung, C.T.; Carlson, R.W. The role of aromatase inhibitors in early breast cancer. *Curr. Treat. Options Oncol.* **2003**, *4* (2), 133–140.
- Maillard, S.; Ameller, T.; Gauduchon, J.; et al. Innovative drug delivery nanosystems improve the anti-tumor activity in vitro and in vivo of anti-estrogens in human breast cancer and multiple myeloma. *J. Steroid. Biochem. Mol. Biol.* **2005**, *94* (1–3), 111–121.

Tissue Engineering of Cardiac Muscle

Wolfram-Hubertus Zimmermann
Thomas Eschenhagen

*Institute of Experimental and Clinical Pharmacology, University-Hospital
Hamburg-Eppendorf, Hamburg, Germany*

INTRODUCTION

Cardiac tissue engineering might offer new solutions for the treatment of heart disease. Ideally, tissue engineered heart muscle could be utilized to replace diseased myocardium and thereby refurbish the heart with contractile elements. To achieve this goal, engineered heart muscle should exhibit structural and functional properties of native myocardium, particularly synchronous contractions and active force development should be present. Additionally, *in vitro* designed heart muscle constructs must have the propensity to integrate into the tissue architecture *in vivo*, remain vital, and contract synchronously with the host myocardium. From a clinical perspective, engineered heart muscle must be autologous or at least allogeneic to avoid immunorejection and should not induce a significant inflammatory response. Further, issues of arrhythmia induction and potential thrombogenicity of graft materials have to be considered. In this article, we will give an overview of current methods in cardiac muscle engineering. Issues that we feel are critical for cardiac muscle engineering in general and specifically for clinical applications will be discussed.

METHODS IN CARDIAC TISSUE ENGINEERING

Pioneering experiments in cardiac regeneration were performed by Bader and Oberpriller in the late 1970s.^[1] In their study, partial regeneration of injured newt ventricles was observed after autologous implantation of minced myocardium. However, grafted tissue fragments remained morphologically and functionally separate from the native myocardium. Recently, the regenerative capacity of ventricular myocardium was demonstrated in zebrafish.^[2] In contrast to amphibians and fish, mammals do not exhibit a similar regenerative capacity. However, survival of engrafted tissue fragments from fetal myocardium of rats and humans could be demonstrated for up to 65 days *in situ*.^[3]

Lately, studies to regenerate myocardium in mammals have been expanded by true tissue engineering approaches (Table 1).^[4-13] Alginate, polyglycolic acid (PGA), and collagen have been utilized as scaffold materials to construct heart muscle. Different groups

including our own demonstrated that cardiac myocytes from neonatal rats and embryonic chicken can be reconstituted to three-dimensional (3-D) cardiac tissue-like constructs. Animal studies indicate that engineered heart muscle can survive and remain contractile *in vivo*. Importantly, the door to potential clinical applications has recently been opened by the demonstration that autologous bovine cardiac myocytes can be derived from fetal tissue generated by somatic nuclear transfer and can be used for cardiac tissue engineering.^[14]

Three principally different strategies have been employed to engineer heart muscle constructs: 1) seeding of solid matrices with cardiac cells; 2) mixing of cardiac cells with initially liquid extracellular matrix components; and 3) stacking of cardiac myocyte monolayers (Fig. 1).

Cell Seeding on Solid Matrices

Langer and Vacanti pioneered the field of matrix-based tissue engineering in the early 1990s.^[15] Seeding of preformed matrices offers intriguing perspectives for tissue regeneration *in vivo*. Ideally, matrix constructs can be formed in every shape and size and potentially even as whole organs. Cell seeding would then vitalize these nonvital matrices to yield biomaterial with organotypic properties. Successful application of this concept has been reported for the reconstitution of cartilage, bone, liver, intestine, and urologic tissues.^[16] In cardiovascular tissue engineering, the construction of artificial valves and vessels has been quite successful, and ongoing preclinical trials in large animal models will decide whether or not clinical studies are warranted in the near future.^[17,18] Unfortunately, utilization of preformed matrices to engineer myocardium has been rather disappointing so far. Mainly, lack of contractile function, poor tissue morphology, and size limitations have been setbacks in the field. Several reasons might account for the failure of the classic tissue engineering approach in cardiac muscle engineering: 1) Engineered matrices resemble diffusion barriers and limit nutrition and oxygen supply in thick constructs (> 100–200 μm); 2) currently employed scaffolds appear not to promote the organization of

Table 1 Cardiac tissue engineering approaches

Strategies in cardiac tissue engineering	Matrix material
<i>Seeding of solid/preformed matrices</i>	
Akins et al., 1999	Polystyrene beads/ collagen threads
Carrier et al., 1999	Polyglycolic acid
Li et al., 1999	Gelatin mesh
Leor et al., 2000	Alginate
Kofidis et al., 2002	Collagen fleece
Van Luyn et al., 2002	Modified collagen
<i>Liquid cell/matrix reconstitution mixture</i>	
Eschenhagen et al., 1997	Collagen type 1
Zimmermann et al., 2000	Collagen type 1 + matrigel
<i>No matrix/stacking of monolayers</i>	
Shimizu et al., 2002	No matrix

Strategies in cardiac tissue engineering. Various scaffold materials have been employed as solid and sometimes preformed matrices, or as initially liquid matrices that contain solubilized extracellular matrix components. Solid matrices serve as an attachment substratum and guidance for the growth of seeded cells, whereas cell entrapment in liquid matrix-cell mixtures is assumed to promote the intrinsic capacity of neonatal rat cardiac myocytes to organize spontaneously into a three-dimensional contractile tissue.

cardiac cells in 3-D cardiac muscle constructs in vitro; 3) differentiation of cardiac myocytes in preformed matrices does not proceed to an adult phenotype; and 4) contractile function and actively developed forces of tissue constructs are modest or not detectable.^[4,7–10] The hope is that chemical or physical modifications of solid matrices will result in optimized scaffolds for cardiac muscle engineering.

Liquid Cell Matrix Mixtures

Solubilized scaffold material has been used as an alternative approach in cardiac muscle engineering by our group.^[5,6,12,19] The principle of this method is that neutralized collagen type I and extracellular matrix proteins (matrigel) are mixed with isolated heart cells from neonatal rats. After being cast in circular molds with a central Teflon rod, this liquid reconstitution mixture entraps the cardiac myocytes in a 3-D geometry and coalesces to strongly contracting and well-differentiated ring-shaped engineered heart tissue (EHT). Large 3-D constructs can be generated by merging several single EHTs. The following factors have been found to be crucial for EHT construction: 1) batch-controlled collagen type 1 from rat tail; 2) matrigel (only in rat EHT);^[3] 3) horse serum (10%) and chick embryo extract (2%);^[20] 4) culture under

mechanical load;^[19] 5) a circular geometry rather than lattices; and 6) utilization of mixed heart cell populations. Under these conditions, strongly contracting (up to 3 mN) and morphologically highly-differentiated muscle constructs can be engineered. Myocytes within EHTs form a network of interconnected myocyte strands throughout the EHT (diameter of ~1 mm).^[12] Single myocyte strands reach a diameter of 20–200 μm . This size limitation is likely to be caused by an oxygen diffusion deficit in thick reconstituted muscle strands. Similar limits in tissue oxygen diffusion have been observed in isolated heart muscle strips.^[21] Despite the possible oxygen diffusion limitation in single reconstituted muscle strands, we could not observe a spatial preference of myocyte strands within EHTs. This indicates that the loose matrix in contrast to the dense myocyte strands does not constitute a significant diffusion barrier in EHTs. Another aspect that indicates normoxic conditions is the high degree of cardiac myocyte differentiation and an organotypic organization in EHTs. This is supported by ultrastructural analysis of cell volume fractions and sarcomere organization of cardiac myocytes in EHT.^[12] Interestingly, myocyte differentiation surpassed the degree of differentiation that was found in neonatal rat myocardium, indicating the importance of 3-D growth possibilities and matrix factors for cardiac myocyte development. Beside the development of a differentiated cardiac myocyte strand network, EHTs exhibit cardiac organoid structures formed by nonmyocytes (e.g., endo- and epicardial surface lining and primitive capillaries).

Stacking of Monolayers

An alternative nonmatrix-based method in cardiac tissue engineering has been developed by Shimizu and colleagues.^[11] This method utilizes single detached cardiac myocyte monolayers in a stacked culture format. Resulting cardiac myocyte sandwiches developed diffuse connexin 43 junctions within the construct. Further, synchronized contractions were observed in vitro and after subcutaneous implantation in vivo. A potential advantage of this technique might be the possibility of stacking monolayers of different cell types and thus reconstituting an organoid tissue construct composed of all cell species that make up the myocardium.^[22] In line with this reasoning, addition of endothelial cells might facilitate the induction of vascularization, and fibroblasts may increase the stability of engineered constructs by producing extracellular matrix. Yet, it would be anticipated that size limitations do also apply to the sandwich culture format. Maximal thickness of four layer constructs has been reported at 70 μm (four monolayers). Whether

T

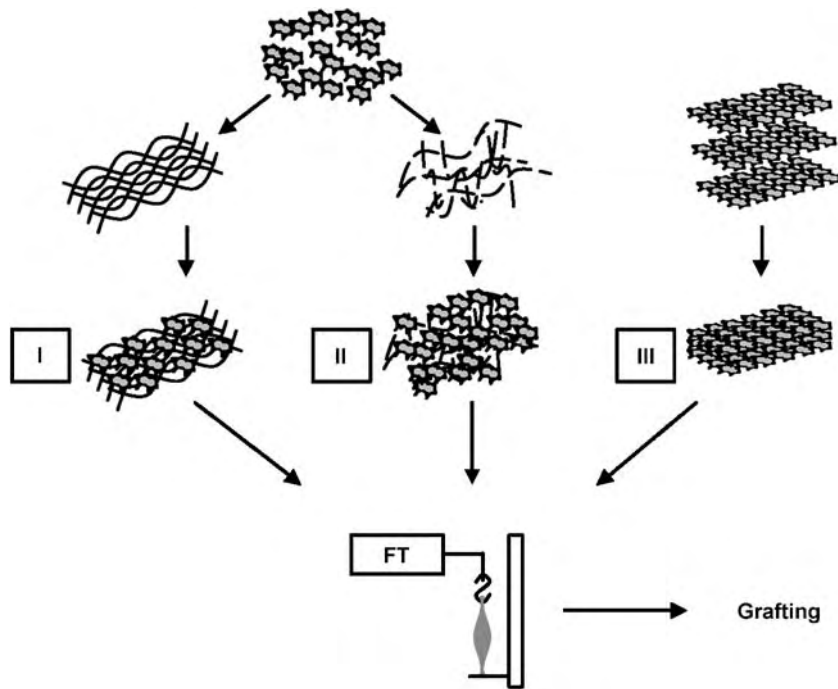


Fig. 1 Different strategies in cardiac tissue engineering. (I) Seeding of solid/preformed matrices with cardiac myocytes;^[1,5-7,13] (II) mixing of liquid, not preformed extracellular matrix components, with cardiac myocytes;^[2,3,12] and (III) stacking of detached cardiac myocyte monolayers^[11] have been employed as main approaches in cardiac tissue engineering. Contractility of muscle constructs can be characterized by isometric force measurements *in vitro*. Subsequently, contractile heart muscle constructs with defined mechanical properties may be used for cardiac regeneration *in vivo*. FT: force transducer.

this technique will be suitable to construct thick viable cardiac grafts (> 200–300 μm and eventually up to several millimeters) remains unclear.

SCAFFOLD MATERIALS IN CARDIAC TISSUE ENGINEERING

A broad range of synthetic polymers such as polyglycolic acid and biomaterials such as alginate, collagen, or gelatine have been employed in the aforementioned tissue engineering approaches (Table 1) and have been extensively reviewed elsewhere.^[16] In general, scaffold materials should be nontoxic, biodegradable, and biocompatible.

Synthetic Polymers

Synthetic polymers have following advantages: 1) Chemical properties are defined; 2) batch-to-batch variations can be reduced to a minimum; 3) geometry can be chosen as required; 4) negligible immune responses are expected; and 5) chemical modifications allow bonding of bioactive compounds that might facilitate survival, differentiation, or angiogenesis *in vitro* and *in vivo*. On the one hand, synthetic polymers are by nature unphysiological. Conceptually, synthetic materials serve initially as an attachment substratum and should guide cell growth but are to be replaced by a physiological matrix over time. On the other hand, biodegradation of synthetic polymers can induce

inflammatory responses to the original material or byproducts of its degradation. Changes in local pH, foreign body reactions, and secretion of inflammatory cytokines may cause unwanted side effects.

Biomaterials

Biomaterials are naturally occurring extracellular matrix components that can be chosen according to the desired application. Yet, the physiological complexity of extracellular matrix composition makes it difficult to mimic an organ-specific environment. Therefore, it is not clear whether biomaterials as currently employed serve a more specific purpose than synthetic scaffolds. Biomaterials may have an advantage of inducing only mild inflammatory responses *in vivo*. On the other hand, the composition of biomaterials strongly depends on the isolation procedure and exhibits batch-to-batch variations. The varying content of extracellular matrix components and growth factors might exert desired and undesired effects on tissue formation that are difficult to define. Additionally, utilization of biomaterials may be associated with the risk of infectious disease transmission.

Choosing the Right Scaffold for Cardiac Tissue Engineering

In cardiac tissue engineering, the choice of scaffold material, with regard to its physical properties, might represent a more crucial issue than in other approaches

in tissue engineering, including artificial cartilage, bone or skin, and organs with a mostly humoral or detoxifying function such as pancreas or liver. On the one hand, scaffold materials in cardiac tissue engineering would have to withstand tremendous mechanical stress *in vivo*. On the other hand, engineered cardiac constructs must remain flexible and extensible at the same time. Importantly, physical matrix properties should not compromise cardiac myocyte contractility and allow transmission of contractile force via the matrix.

It is apparent that none of the currently employed materials or approaches offers an ideal solution. Whereas chemical and physical modification of scaffold materials might yield further optimized matrices, alternative approaches exist. One is to use scaffolds that are derived from intact organs of donor animals by chemical decellularization and then repopulated with host-derived cells. This approach is currently employed in tissue engineering of heart valves. Additional approaches are to utilize liquid matrix protein mixtures that may not serve as attachment substratum but maintain a favorable growth milieu for cardiac myocytes and thus promote spontaneous cardiac organogenesis, or to avoid scaffold material completely.

As outlined earlier, our own approach is based on the concept of self-organization of isolated cardiac myocytes into cardiac tissue in a primarily liquid growth milieu. The propensity of cardiac myocytes from neonatal rats to reorganize into cardiac tissue structures *in vitro* has been demonstrated.^[7] A main requirement for successful EHT construction is to maintain a high cardiac cell density in a defined volume of concentrated collagen/matrigel and thus allow for an intimate cell–cell contact. During 10–14 days in culture, EHTs undergo extensive remodeling. This includes changes in matrix structure and composition as well as in cell differentiation and organization (Fig. 2). As observed in monolayer cultures, embryonic chick or neonatal rat cardiac myocytes start to spread, redifferentiate, and contract spontaneously in EHTs. In contrast to monolayers, cell growth occurs in three dimensions. Over time cardiac myocytes organize to highly differentiated interconnected muscle strands and develop cell–cell contacts via fascia adherens, gap junctions, and desmosomes. Additionally, nonmyocytes within EHTs remodel and partially replace the extracellular matrix (fibroblasts, macrophages), forming a complete epi- and endocardial surface lining (endothelial cells, smooth muscle cells, cells of leukocytotic origin) and primitive capillaries (endothelial cells).^[12] Thus, our data support the notion that cardiac cells from embryonic chick and neonatal rats have the intrinsic capacity to form new cardiac tissue, and that this process is merely stimulated and directed by the collagen/matrigel-containing environment. A disadvantage of this approach is that

geometric shape and size cannot be controlled as easily as when employing preformed fabricated scaffolds. A compromise is the use of different casting molds and the merging of several ring-shaped EHTs to larger patches. An advantage of EHTs is that stretch devices can be utilized to allow culture under defined mechanical load. The importance of stretch/load has been substantiated by our recent findings that contractile force, morphology, and tissue formation of EHTs can be markedly improved by phasic stretch.^[19] In contrast, when EHTs are left unloaded during culture, they quickly retract, exhibit thickening of the matrix, and stop beating.

CELLS IN CARDIAC TISSUE ENGINEERING

Reconstruction of viable myocardium after a large myocardial infarction in humans will require approximately 1 billion cardiac myocytes.^[23] Hence, success in cardiac tissue engineering will critically depend on the availability of large quantities of, ideally, autologous cardiac myocytes. Unfortunately, cardiac myocytes withdraw from the cell cycle before or shortly after birth. This precludes the use of biopsy material for cardiac myocyte propagation. Whether strategies to induce cell cycle reentry can resolve this limitation remains to be elucidated.^[24] Embryonic and adult stem cells might represent an alternative to primary cells.^[25,26] Despite the proof of embryonic stem cell differentiation into cardiac myocytes and some evidence for adult stem cell transdifferentiation, enthusiasm is limited because cardiac lineage (trans) differentiation appears to be a rare event. The key questions involve: 1) means to upscale the yield of cardiac myocyte differentiation from embryonic stem cells; 2) factors that induce transdifferentiation of adult stem cells; and 3) factors that induce terminal differentiation of stem cell-derived cardiac myocyte. Other problems associated with (embryonic) stem cells are tumor formation and immunorejection. Tumor formation might be a question of cell dose and the proportion of remaining pluripotent stem cells in transplanted cardiac cell populations. Whether depletion of remaining multipotent stem cells can resolve this issue remains to be demonstrated. Immunorejection might be controlled by utilization of a primarily autologous cell source (adult stem cells), tolerance induction,^[27] or nuclear somatic transfer to generate cells of the recipient's genotype.^[14] Intense basic research is necessary to find answers to these questions before advancing into the clinic.

Another issue that might be important for cardiac tissue engineering is the cellular composition of cardiac grafts. Physiologically, endothelial cells, fibroblasts, smooth muscle cells, neural cells, and leukocytes

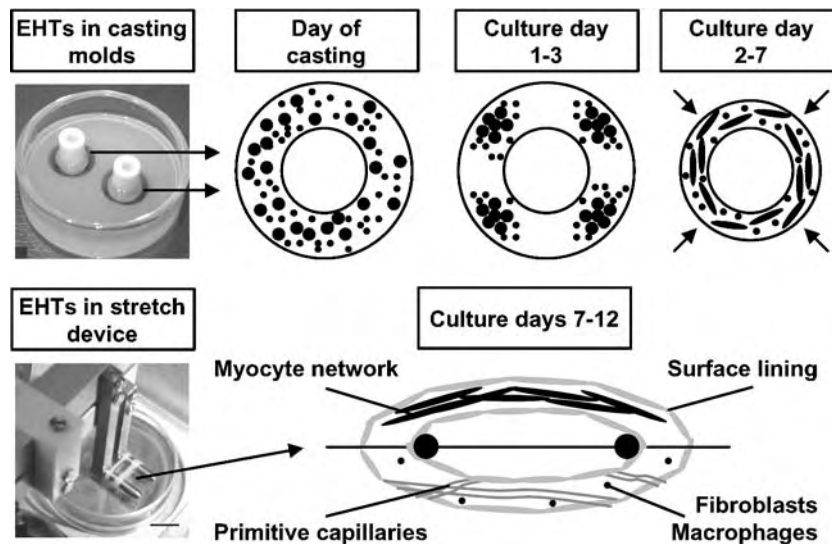


Fig. 2 Engineered heart tissue (EHT) development. Original photographs depict EHTs in casting molds (culture days 0–7) and in a stretch device (culture days 7–12). Initially, heart cells from neonatal rats are entrapped in a liquid collagen/matrigel mixture. Cell clustering and spontaneous contractions of clusters are observed on culture days 1–3. At the same time, EHTs start to coalesce around central Teflon spacers; cardiomyocytes elongate, and begin to form a synchronously contracting muscle network (culture days 2–7). Subsequently, EHTs are subjected to phasic strain (culture days 7–12) to further improve tissue formation and contractility. Formation of organoid structures is summarized in the bottom sketch: Cardiac myocytes develop a dense myocyte network throughout the entire EHT. Nonmyocytes reconstitute primitive capillaries and a surface lining composed of endothelial cells, smooth muscle cells, and cells of leukocytotic origin. Fibroblasts and macrophages are scattered throughout the EHT and contribute to matrix remodeling by synthesizing extracellular matrix proteins and sequestering cell debris.

constitute about 70% of the total cell number in the working myocardium.^[22] Thus, cardiac muscle engineering is more likely to be successful if various cardiac cell species are being considered and brought together in a physiological context. *In vitro* studies have demonstrated the role of nonmyocytes in cardiac development and function.^[28] Whether nonmyocytes will be beneficial in cardiac tissue engineering remains to be clarified. This notion is supported by preliminary studies from our laboratory. When EHTs were constructed from a native heart cell rather than from a purified cardiac myocyte population, contractile force was markedly improved. Presently, the exact contribution of each single cell type to tissue formation is under investigation.

FUNCTIONAL REQUIREMENTS

Conceptually, most cardiac muscle engineering approaches aim at improving systolic function of diseased ventricles. We believe that to achieve this goal, engineered heart muscle needs to be fairly differentiated, to beat in synchrony, to feature systolic and diastolic properties of native myocardium, to develop sufficient contractile force, to have a critical size, and

to have the propensity to integrate into the host tissue architecture. In fact, noncontractile cardiac grafts have been implanted and have improved diastolic properties.^[9] This was explained by a beneficial effect on scar remodeling and prevention of aneurysm formation in infarcted hearts. Yet, systolic properties were not changed in this study.

So far, strong synchronous contractions have, to our knowledge, been shown only for stacked monolayers and EHTs (Table 2). Another study that employed cell seeding of preformed matrices has also described synchronously contractile muscle constructs, but systolic and diastolic contractile properties were markedly inferior to the aforementioned methods.^[8] Besides absolute contractile forces, an important parameter of contractile function is the ratio of systolic (twitch tension; TT) and diastolic force (resting tension; RT). Physiologically, the TT/RT ratio of healthy cardiac muscle is > 1 , even at modest inotropic stimulation. This ratio is lower in diseased hearts as a consequence of fibrotic remodeling of the myocardium. EHTs develop a TT of > 2 mN with a TT/RT of > 3 under maximal inotropic stimulation with calcium. Monolayer stacks develop comparable TT but RT has not been reported.^[11] In cell-seeded collagen fleeces, the TT/RT ratio was apparently low, indicating a high

Table 2 Reported contractile properties of tissue engineered heart muscle

	TT (mN)	RT (mN)	TT/RT
Eschenhagen et al., 1997	0.34±0.05	2.81±0.11	0.12
Fink et al., 2000	1.91±0.16	nd	nd
Zimmermann et al., 2000	0.51±0.13	0.63	0.81
Kofidis et al., 2002	0.019±0.004	0.23	0.08
Shimizu et al., 2002	1.18±0.26	nd	nd
Zimmermann et al., 2002	0.75±0.11	0.15±0.02	5

Systolic force (twitch tension; TT) and diastolic force (resting tension; RT) were measured under isometric conditions^[2,3,5,12,19] or in vivo.^[11] If possible, TT/RT ratios were calculated from available data. Nd: not determined.

matrix rigidity.^[8] High passive diastolic forces, on the one hand, could have a beneficial effect on scar remodeling and possibly prevent ventricle enlargement and aneurysm formation. On the other hand, they may contribute to increased myocardial stiffness in vivo and thus would have to be considered as rather detrimental. Systematic implantation studies will have to examine this issue.

PRECLINICAL IN VIVO APPLICATIONS

Studies to test the feasibility of cardiac muscle constructs as graft material have been performed almost exclusively in rats.^[9–11,20,29] So far, most studies have addressed issues of in vivo survival, maintenance of electrical activity, and contractile function in healthy animals. Grafting was performed subcutaneously,^[10,11] into the peritoneal cavity,^[20] or on beating hearts of healthy rats.^[29] The question of in vivo survival is not trivial because implanted muscle constructs are not immediately connected to the host vasculature and must be nurtured by diffusion until vascularization of implants has occurred. Despite this apparent caveat, different groups have shown that engineered cardiac muscle constructs do survive in vivo and retain their electromechanical activity.^[9–11,20,29] Presently, it is unclear if and to what extent cell death occurs immediately after implantation of cardiac grafts as a consequence of hypoxia or reduced nutrient supply. Vascularization of implants has been reported to occur rather fast and thus might rescue surviving cells. In EHT implants, survival of well-organized cardiac muscle has been observed predominantly in close vicinity of new ingrown vessels. Here, grafted cardiac myocytes showed ultrastructural evidence of terminal differentiation and were indistinguishable from native cardiac myocytes.^[29] Additionally, ingrowth of well-organized nerve bundles

composed of myelinated and nonmyelinated nerve fibers was observed in EHTs. So far, we have not analyzed the origin of the nerve bundles and whether innervation has any functional consequences.^[29] Contractile function of grafted EHTs has been observed for at least eight weeks in vivo, and ongoing grafting studies will shed light on the consequences of EHTs for the function of infarcted hearts.

Important issues for cardiac muscle grafting are inflammation and immune responses. Although a transient inflammation caused by the surgical intervention cannot be completely avoided, additional inflammatory responses to scaffold materials or its degradation products may be minimized by optimizing the physical and chemical properties of scaffold materials, by using autologous biomaterials, or by avoiding scaffold material altogether. Immune responses in recent grafting studies have been controlled by utilization of nude rats^[11] or immunosuppression.^[29] A new and exciting, but technically challenging and ethically questionable, approach is to utilize autologous donor animals that have been developed by nuclear cloning.^[14]

CONSIDERATIONS FOR CLINICAL APPLICATIONS

Predictions on the future of cardiac muscle engineering in practical medicine remain necessarily speculative at present. We believe that the principal techniques of graft construction are likely to be resolved. However, identification of a suitable cell source for human applications will remain a key issue. Whether embryonic stem cells from presently available lines, adult stem cells, or ethically controversial nuclear cloning approaches will offer a perspective remains to be shown. As mentioned earlier, all of these approaches are hampered by a low cardiac myocyte yield, the potential of tumor formation, and, in the case of adult stem cells, the remaining uncertainty about true transdifferentiation versus cell fusion or simple phagocytosis.^[30] Other areas of intense research will be the risk of arrhythmia induction by spontaneously depolarizing cells, inflammation, immune responses, vascularization, functional integration, thrombogenicity of graft material, and tumor formation. Some of these caveats might be addressed by electrical pacing of implanted constructs, utilization of purely autologous scaffolds and cells; integration of vessels in cardiac constructs in vitro; bonding of factors that improve cell survival, differentiation, and function; utilization of nonthrombogenic scaffolds; and stringent cell purification when stem cell-derived cardiac myocytes are employed.

Eventually, all approaches in cardiac regeneration have to be compared to standard treatment. Potential advantages of cardiac tissue engineering for cardiac regeneration might be: 1) Grafting of cardiac constructs would resemble a causal therapy; 2) graft design can be performed according to patients' needs; 3) functional and structural properties of cardiac grafts can be controlled prior to grafting; 4) surgical grafting of engineered heart muscle is likely to prevent washout of cells as it occurs in simple cell implantation; and 5) it would ideally allow for true replacement of a ventricle wall, reconstruction of congenital malformations, and possibly construction of a complete heart, although, the latter aspect is, at least in our view, rather unlikely to be achieved with currently employed strategies. A first step to advance presently employed approaches would be to provide proofs for the concept of cardiac regeneration in small animal models and subsequently in large animals such as pigs, sheep, or dogs.

ARTICLES OF FURTHER INTEREST

Cardiac Bioelectricity; Cardiac Patch Engineering; Electrocardiography; Poly(Glycolic Acid); Poly(lactic acid)s; Tissue Engineering, Microscale; Tissue Engineering Scaffolds

REFERENCES

- Bader, D.; Oberpriller, J.O. Repair and reorganization of minced cardiac muscle in the adult newt (*Notophthalmus viridescens*). *J. Morphol.* **1978**, *155* (3), 349–357.
- Poss, K.D.; Wilson, L.G.; Keating, M.T. Heart regeneration in zebrafish. *Science* **2002**, *298* (5601), 2188–2190.
- Leor, J.; Patterson, M.; Quinones, M.J.; Kedes, L.H.; Kloner, R.A. Transplantation of fetal myocardial tissue into the infarcted myocardium of rat. A potential method for repair of infarcted myocardium? *Circulation* **1996**, *94* (9 Suppl), II332–II336.
- Carrier, R.L.; Papadaki, M.; Rupnick, M.; Schoen, F.J.; Bursac, N.; Langer, R.; Freed, L.E.; Vunjak-Novakovic, G. Cardiac tissue engineering: cell seeding, cultivation parameters, and tissue construct characterization. *Biotechnol. Bioeng.* **1999**, *64* (5), 580–589.
- Eschenhagen, T.; Fink, C.; Remmers, U.; Scholz, H.; Wattochow, J.; Weil, J.; Zimmermann, W.; Dohmen, H.H.; Schafer, H.; Bishopric, N.; Wakatsuki, T.; Elson, E.L. Three-dimensional reconstitution of embryonic cardiomyocytes in a collagen matrix: a new heart muscle model system. *Faseb J.* **1997**, *11* (8), 683–694.
- Zimmermann, W.H.; Fink, C.; Kralisch, D.; Remmers, U.; Weil, J.; Eschenhagen, T. Three-dimensional engineered heart tissue from neonatal rat cardiac myocytes. *Biotechnol. Bioeng.* **2000**, *68* (1), 106–114.
- Akins, R.E.; Boyce, R.A.; Madonna, M.L.; Schroedl, N.A.; Gonda, S.R.; McLaughlin, T.A.; Hartzell, C.R. Cardiac organogenesis in vitro: reestablishment of three-dimensional tissue architecture by dissociated neonatal rat ventricular cells. *Tissue Eng.* **1999**, *5* (2), 103–118.
- Kofidis, T.; Akhyari, P.; Boublik, J.; Theodorou, P.; Martin, U.; Ruhparwar, A.; Fischer, S.; Eschenhagen, T.; Kubis, H.P.; Kraft, T.; Leyh, R.; Haverich, A. In vitro engineering of heart muscle: artificial myocardial tissue. *J. Thorac. Cardiovasc. Surg.* **2002**, *124* (1), 63–69.
- Leor, J.; Aboulafia-Etzion, S.; Dar, A.; Shapiro, L.; Barbash, I.M.; Battler, A.; Granot, Y.; Cohen, S. Bioengineered cardiac grafts: A new approach to repair the infarcted myocardium? *Circulation* **2000**, *102* (19 Suppl 3), III56–III61.
- Li, R.K.; Jia, Z.Q.; Weisel, R.D.; Mickle, D.A.; Choi, A.; Yau, T.M. Survival and function of bioengineered cardiac grafts. *Circulation* **1999**, *100* (19 Suppl), II63–II69.
- Shimizu, T.; Yamato, M.; Isoi, Y.; Akutsu, T.; Setomaru, T.; Abe, K.; Kikuchi, A.; Umezu, M.; Okano, T. Fabrication of pulsatile cardiac tissue grafts using a novel 3-dimensional cell sheet manipulation technique and temperature-responsive cell culture surfaces. *Circ. Res.* **2002**, *90* (3), e40.
- Zimmermann, W.H.; Schneiderbanger, K.; Schubert, P.; Didie, M.; Munzel, F.; Heubach, J.F.; Kostin, S.; Neuhuber, W.L.; Eschenhagen, T. Tissue engineering of a differentiated cardiac muscle construct. *Circ. Res.* **2002**, *90* (2), 223–230.
- van Luyn, M.J.; Tio, R.A.; Gallego y van Seijen, X.J.; Plantinga, J.A.; de Leij, L.F.; DeJongste, M.J.; van Wachem, P.B. Cardiac tissue engineering: characteristics of in unison contracting two- and three-dimensional neonatal rat ventricle cell (co)-cultures. *Biomaterials* **2002**, *23* (24), 4793–4801.
- Lanza, R.P.; Chung, H.Y.; Yoo, J.J.; Wettstein, P.J.; Blackwell, C.; Borson, N.; Hofmeister, E.; Schuch, G.; Soker, S.; Moraes, C.T.; West, M.D.; Atala, A. Generation of histocompatible tissues using nuclear transplantation. *Nat. Biotechnol.* **2002**, *20* (7), 689–696.
- Langer, R.; Vacanti, J.P. Tissue engineering. *Science* **1993**, *260* (5110), 920–926.
- Vacanti, J.P.; Langer, R.; Upton, J.; Marler, J.J. Transplantation of cells in matrices for tissue regeneration. *Adv. Drug. Deliv. Rev.* **1998**, *33* (1–2), 165–182.
- Sodian, R.; Hoerstrup, S.P.; Sperling, J.S.; Daebritz, S.; Martin, D.P.; Moran, A.M.; Kim, B.S.; Schoen, F.J.; Vacanti, J.P.; Mayer, J. E., Jr. Early in vivo experience with tissue-engineered trileaflet heart valves. *Circulation* **2000**, *102* (19 Suppl 3), III22–III29.
- Steinhoff, G.; Stock, U.; Karim, N.; Mertsching, H.; Timke, A.; Meliss, R.R.; Pethig, K.; Haverich, A.; Bader, A. Tissue engineering of pulmonary heart valves on allogenic acellular matrix conduits: In vivo restoration of valve tissue. *Circulation* **2000**, *102* (19 Suppl 3), III50–III55.

19. Fink, C.; Ergun, S.; Kralisch, D.; Remmers, U.; Weil, J.; Eschenhagen, T. Chronic stretch of engineered heart tissue induces hypertrophy and functional improvement. *Faseb J.* **2000**, *14* (5), 669–679.
20. Eschenhagen, T.; Didie, M.; Munzel, F.; Schubert, P.; Schneiderbanger, K.; Zimmermann, W.H. 3D engineered heart tissue for replacement therapy. *Basic Res. Cardiol.* **2002**, *97* (Suppl 1), I146–I152.
21. Gulch, R.W.; Ebrecht, G. Mechanics of rat myocardium revisited: investigations of ultra-thin cardiac muscles under high energy demand. *Basic Res. Cardiol.* **1987**, *82* (Suppl 2), 263–274.
22. Nag, A.C.; Zak, R. Dissociation of adult mammalian heart into single cell suspension: An ultrastructural study. *J. Anat.* **1979**, *129* (3), 541–559.
23. Gepstein, L. Derivation and potential applications of human embryonic stem cells. *Circ. Res.* **2002**, *91* (10), 866–876.
24. Pasumarthi, K.B.; Field, L.J. Cardiomyocyte cell cycle regulation. *Circ. Res.* **2002**, *90* (10), 1044–1054.
25. Kehat, I.; Kenyagin-Karsenti, D.; Snir, M.; Segev, H.; Amit, M.; Gepstein, A.; Livne, E.; Binah, O.; Itskovitz-Eldor, J.; Gepstein, L. Human embryonic stem cells can differentiate into myocytes with structural and functional properties of cardiomyocytes. *J. Clin. Invest.* **2001**, *108* (3), 407–414.
26. Jiang, Y.; Jahagirdar, B.N.; Reinhardt, R.L.; Schwartz, R.E.; Keene, C.D.; Ortiz-Gonzalez, X.R.; Reyes, M.; Lenvik, T.; Lund, T.; Blackstad, M.; Du, J.; Aldrich, S.; Lisberg, A.; Low, W.C.; Largaespada, D.A.; Verfaillie, C.M. Pluripotency of mesenchymal stem cells derived from adult marrow. *Nature* **2002**, *418* (6893), 41–49.
27. Fandrich, F.; Lin, X.; Chai, G.X.; Schulze, M.; Ganten, D.; Bader, M.; Holle, J.; Huang, D.S.; Parwaresch, R.; Zavazava, N.; Binas, B. Preimplantation-stage stem cells induce long-term allogeneic graft acceptance without supplementary host conditioning. *Nat. Med.* **2002**, *8* (2), 171–178.
28. Gray, M.O.; Long, C.S.; Kalinyak, J.E.; Li, H.T.; Karliner, J.S. Angiotensin II stimulates cardiac myocyte hypertrophy via paracrine release of TGF-beta 1 and endothelin-1 from fibroblasts. *Cardiovasc Res* **1998**, *40* (2), 352–363.
29. Zimmermann, W.H.; Didie, M.; Wasmeier, G.H.; Nixdorff, U.; Hess, A.; Melnychenko, I.; Boy, O.; Neuhuber, W.L.; Weyand, M.; Eschenhagen, T. Cardiac grafting of engineered heart tissue in syngenic rats. *Circulation* **2002**, *106* (12 Suppl 1), I151–I1157.
30. Wurmser, A.E.; Gage, F.H. Stem cells: Cell fusion causes confusion. *Nature* **2002**, *416* (6880), 485–487.

Tissue Engineering of Cartilage

J.N. Amritha De Croos

Rita A. Kandel

Bioengineering of Skeletal Tissues Team, Mount Sinai Hospital, Toronto, Ontario, Canada

INTRODUCTION

Articular cartilage is a unique tissue found at the articulating ends of bones whose functions are the distribution of applied load to the underlying subchondral bone and to provide, along with the synovial fluid, a low-friction interface between the contacting surfaces in the joint.^[1] As cartilage does not have a blood supply, it relies on the surrounding synovial fluid for nutrition. As well as being avascular, it has neither a lymphatic supply nor neural innervation. Adult articular cartilage has a limited ability to spontaneously heal when damaged by trauma or disease. A variety of reasons for this lack of reparative response have been postulated, including the inability of chondrocytes to migrate into the site of injury, the avascular nature of cartilage, and the absence of a fibrin clot scaffold into which the cells can migrate.^[2,3] However, other factors may also influence the repair process as chondrocytes have been shown to be able to reconstitute at least part of their extracellular matrix (ECM) in vitro as well as in vivo with the appropriate treatment.^[4,5] The morbidity, pain, and limitation of movement that arises as a result of joint disease are significant problems for both the individual and the health care system. Arthritis affects some 40 million Americans, or 15% of the US population.^[6] And, with an aging population, this number is only expected to increase. There is a great need to develop new approaches to treat damaged cartilage, especially focal defects or during the early stages of arthritis.

Currently, replacement of an articulating joint with a synthetic prosthesis represents the optimal treatment for end-stage joint disease. Primary hip replacements placed since the late 1980s have shown reasonable success rates. However, even apparently successful implants have their limitations as failure rates of up to 20% have been reported after 10–20 years.^[7–9] To circumvent the problems associated with the use of prostheses, several biological treatments have been developed to repair focal cartilage defects.^[10] One of these involves the transplantation of osteochondral plugs that have been obtained from donor sites within the same joint into a defect site (osteochondral transfer/mosaicplasty). Several plugs are usually required to repair a single defect that may be several centimeters in

diameter. The major disadvantages of this approach are the resulting donor site morbidity (i.e., a hole remains in the donor site) and difficulty in placing the plugs congruent with the joint surface. The long-term success of this technique is also questionable.^[1,10] Alternatively, large articular shell grafts (articular cartilage with subjacent thin layer of bone at least 1 cm in thickness) have been utilized with success, but as these are allografts they are limited in supply and can be associated with disease transmission.^[11] Perforation (microfracture) of the subchondral bone has also been used to induce cartilage repair. However, as this response is a result of ingrowth of cells from the bone marrow this treatment results in fibrocartilage repair, rather than hyaline cartilage, and as it does not have the necessary load-bearing capability it will degenerate over time.^[12]

As all of the above treatments have limitations, current investigations have focused on developing strategies to replace damaged tissue with tissue-engineered cartilage as an alternative, novel therapeutic approach. Tissue engineering emerged in the late 1980s and has as its goal the regeneration of tissue using biomaterials and/or cells.^[13] This approach has been investigated intensively, and there is now proof of concept to support cell-based regeneration of tissue.^[14–16] However, there are still some limitations, and the goal of current research related to tissue engineering of cartilage is to identify the conditions that will result in the formation of tissue that approximates in vivo cartilage and will be suitable to use in repairing joint defects. This tissue could also be used as a model to study articular cartilage development and disease mechanisms, and for in vitro testing of new therapeutic agents. This review will provide a summary of the characteristics and composition of native articular cartilage, an overview of the methods used currently to bioengineer cartilage, an approach to evaluate the bioengineered cartilage, and a discussion of the problems associated with generating tissue in vitro.

ARTICULAR CARTILAGE: COMPOSITION

Chondrocytes are the cellular component of articular cartilage, which synthesize, secrete, and organize ECM

molecules. Chondrocytes make up about 10% (or less) by volume of cartilage.^[17] An important feature of chondrocytes is that they are isolated from each other and surrounded by abundant ECM. This matrix prevents cell migration likely contributing to the limited repair potential of cartilage. For a review of chondrocytes, their function, and origin, see Ref. [18].

In humans, articular cartilage has a high (60–80% of wet weight) water/fluid content and varies in thickness from 2 to 4 mm depending on the location in the joint. The ECM makes up about 20–40% of the wet weight of articular cartilage and consists predominantly of proteoglycans (15–30% of dry weight) and collagens (50–70% of dry weight).^[17,19] Glycoproteins and other non-collagenous proteins make up approximately 15–20% of the dry weight.^[17,19] Aggrecan is the main proteoglycan found in articular cartilage. Aggrecan monomers bind to hyaluronan, an interaction stabilized by link protein, and form large aggregates that are distributed throughout the collagen network. Smaller proteoglycans such as biglycan, decorin, and fibromodulin can bind to other components of the ECM, contributing to matrix stability. The predominant collagen is type II, but collagen types V, VI, IX, X, and XI are also found, but to a lesser extent. Type I and III collagens originally not thought to be a component of articular cartilage have also been detected in small amounts. Proteoglycans are hydrophilic and entrap water, allowing articular cartilage to withstand compressive forces. The collagen fibers constrain the proteoglycans and provide the tensile strength required by cartilage.

As well, there are numerous other non-collagenous proteins present in the ECM (see review^[20]). One such protein is cartilage oligomeric matrix protein (COMP), which has been shown to be important in sensing mechanical loading.^[21] COMP along with the other non-collagenous proteins aids in the proper functioning of articular cartilage, but is present in lower amounts than proteoglycans and collagens. The importance of these non-collagenous matrix molecules has been demonstrated by studies of knockout mice or naturally occurring mutations in matrix molecules that occur in humans. For example, individuals with a mutation in COMP develop pseudoachondroplasia, which is characterized by a lack of type II collagen fibril organization, although the role that COMP plays in collagen organization and pseudoachondroplasia is not known.^[22]

ARTICULAR CARTILAGE: ORGANIZATION

Articular cartilage can be divided into four distinct zones: A) superficial; B) middle; C) deep; and D) calcified (Fig. 1). Each of these areas has a composition

and organization that contributes to normal articular cartilage function. There is evidence suggesting that each zone plays a role in response to loading. It is possible that for bioengineered cartilage to function properly it may be necessary to reconstruct all these zones.

Superficial Zone

This zone is the upper-most layer and forms the gliding surface of the cartilage. It has a lower-water content than the deeper zones of cartilage.^[17] The collagen fibers are arranged parallel to the articular surface,^[17,23] and morphologically, the chondrocytes in this zone are flatter. The superficial zone has the highest content of type II collagen of the four zones and the lowest content of proteoglycans. Type I and III collagens have also been detected in this region. Decorin and biglycan are the predominant proteoglycans in this zone. Superficial zone protein (SZP), a proteoglycan involved in lubrication, is unique to the superficial zone at the joint–cartilage interface.^[24] The organization and composition of the superficial zone likely aids in resisting the shear stresses produced during articulation.^[17]

Middle Zone

This zone (also called the transitional, intermediate, or middle zone) is the largest of the zones in terms of thickness.^[17,23] The chondrocytes of the middle zone are spherical. The collagen fibers in the middle zone tend to be randomly oriented and surround the cell. Type II collagen content is higher than in the deep zone. Aggrecan is the major proteoglycan, and the proteoglycan content is higher in the middle zone relative to the superficial zone. The middle zone also has a characteristic protein, cartilage intermediate layer protein, that increases in amount with age,^[25] although its function is not fully characterized.

Deep Zone

In the deep zone (also called the radial zone), the collagen fibers are oriented perpendicular to the subchondral plate and cross into the underlying calcified cartilage. The radial orientation of collagen fibers contributes to the compressive resistance characteristic of this zone and allows a stronger adherence of the articular cartilage to the subchondral bone.^[26] The chondrocytes can be aligned in columns perpendicular to the subchondral plate. These cells retain some of the properties of hypertrophic-like chondrocytes, which



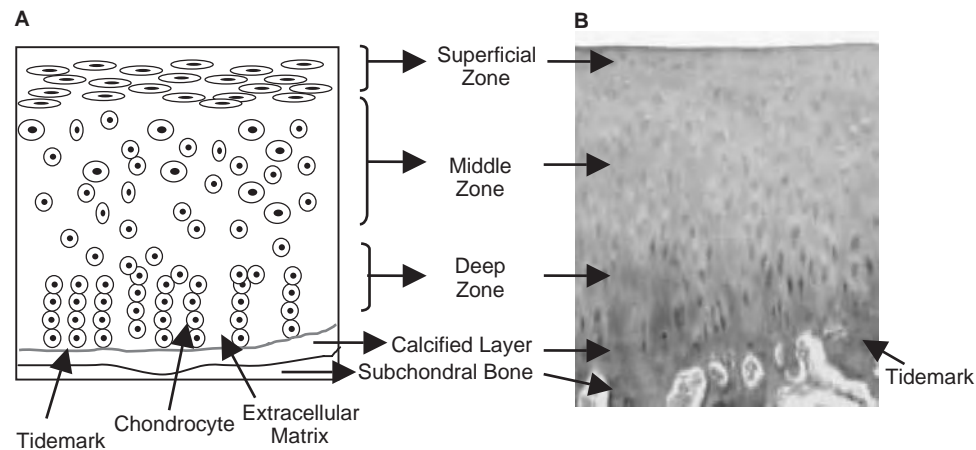


Fig. 1 Schematic of articular cartilage. (A) Diagram of articular cartilage outlining the various zones and key features of articular cartilage. (B) Photomicrograph of human articular cartilage and subchondral bone following formalin fixation, decalcification, and paraffin embedding (hematoxylin and eosin, magnification $\times 25$).

include expression of alkaline phosphatase and type X collagen.^[27]

Just below the deep zone is the tidemark, which is the interphase between the deep zone and calcified layer. The tidemark is known to contain crystals of calcium salts and hyaluronan but otherwise is poorly characterized.

Calcified Layer

Just below the deep zone is the calcified layer that consists of hyaline cartilage that has undergone mineralization. During development the cells of the calcified layer are active but become quiescent in adult cartilage. This zone anchors the hyaline cartilage to the underlying subchondral bone and is involved in the transmission of forces across the joint. The calcified layer stiffness is less than bone most likely because of the arrangement of collagen fibers.^[28] The elastic modulus, which is a measure of the stiffness of the tissue, is approximately 0.32 GPa, which is an order of magnitude greater than the non-mineralized hyaline cartilage but is an order of magnitude less than bone.^[28] It has been hypothesized that this stepwise transition in stiffness is very important for the proper functioning of cartilage and distribution of forces across the joint and is often not considered for tissue bioengineering articular cartilage.

In summary, the goal of tissue engineering cartilage is to generate articular cartilage that will function similarly to the native tissue and to do so may require regeneration of cartilage with all of these zones. It is well known that repair of cartilage defects with fibrocartilage, which does not have the organization and

composition of hyaline articular, will degrade over time with weight bearing as it does not have the necessary load-bearing capability.^[10]

ARTICULAR CARTILAGE: FUNCTION

Articular cartilage is important for smooth, painless articulation. The tissue is constructed to allow for this. The superficial zone is organized, as described above, to support this function. As well, chondrocytes in this region (superficial zone) produce a factor (lubricin or SZP) that is critical to proper cartilage function.^[24] Cartilage tribology is not well understood and likely is important for tissue-engineered cartilage.

Cartilage is exposed to a variety of static and cyclic loads and can sustain loads that compress the cartilage as much as 45% (with contact pressures up to 14 MPa).^[29,30] The mechanical properties of the cartilage can vary depending on its composition and location in the joint. When cartilage is loaded, for instance, during walking there is an initial increase in pressure which squeezes the fluid out of the solid matrix,^[23] however, the exudation is affected by the relatively low fluid permeability of the ECM. The subchondral bone and the surrounding adjacent cartilage confine the cartilage serving to restrict mechanical deformation.^[31] As the fluid is expressed, the ECM of the superficial layer and to some extent the upper middle (transitional) zone consolidate, the chondrocytes in this zone will undergo compressive deformation, and the cartilage experiences an increase in the negative charge density because of the fixed proteoglycans as well as an increase in osmotic pressure. Given the high (60–80%) water content of articular

cartilage, the compressive strains can be substantial.^[17] The compressive forces are resisted by the solid phase and the fluid pressures. In addition, the rolling action of joint surfaces produce cyclic tensile strains that are tangential and outside the contact area of articular cartilage.^[32] Tissue-engineered cartilage must be able to withstand these forces. However, what is not known is the minimal mechanical properties the tissue should have to ensure that the tissue-engineered cartilage will function properly over the long term.

Loading of articular cartilage is also important for nutrition and cartilage health. As mentioned above, cartilage is avascular and alymphatic and therefore relies on diffusion of nutrients from the synovial fluid. Mechanical loading facilitates water/fluid movement and thus distribution of nutrients to the chondrocytes embedded in the ECM as well as removal of waste products from the tissue. It has been shown that *in vitro* cartilage tissue formation is improved when there is unrestricted exchange of nutrients and/or solutes that may explain in part the positive effect of forming cartilage in bioreactors observed under many experimental conditions.^[33,34]

ARTICULAR CARTILAGE: CELL SIGNALING

As chondrocytes are surrounded by ECM (and are not in contact with other cells) they perceive changes through effects on molecules in the matrix and translate those signals into gene and protein changes (Fig. 2). Thus, tissue-engineered cartilage that does not have the appropriate composition when subjected to weight bearing could activate a chondrocyte response that would potentially result in either tissue degradation or production of abnormal matrix molecules. The end result would be degradation of the bioengineered cartilage. Much research has focused on the cell surface molecules involved in the signaling process activated by matrix molecules.^[35,36] Integrins, one such molecule, are transmembrane proteins made up of an alpha and beta subunit, which function as receptors for extracellular matrix molecules such as collagen. Integrins also contain a cytoplasmic tail that interacts with the actin cytoskeleton, and this allows for outside-in signaling by the extracellular matrix.^[36,37] Other membrane surface receptors found in chondrocytes such as CD44 and annexin V (anchoring CII) are also involved in signaling and act through a pathway different than that used by integrins.^[35]

After activation of cell surface receptors, the signal is conducted inwards through secondary messengers. Regulation of cellular function is then exerted by transcription factors that are induced by these environmental cues and can enhance or inhibit transcription of

a gene. Generally, transcription factors are phosphorylated at certain protein residues by kinases such as mitogen activated protein kinases (MAPK). Once phosphorylated, the transcription factor undergoes a conformational change that allows it to bind to a DNA consensus sequence and regulate transcription.^[38] Recent studies have highlighted the mechanosensitivity of the MAPK p38, extracellular-signal-regulated kinase (ERK)1/2, and stress activated protein kinase/cjun amino terminal kinase (SAPK/JNK) in cartilage explants.^[39] Apart from the MAPK, there are other pathways that may play a role in chondrocyte signaling such as the RhoA/RhoA kinase (ROCK) pathway.^[40,41] This cell-signaling pathway involves RhoA, a small G-protein in the Rho family that regulates cell morphology via actin cytoskeleton reorganization in response to extracellular signals. The RhoA/ROCK pathway has not been well studied in articular cartilage, but it has been shown that chondrocyte cytoskeleton changes involved both integrins and association with RhoA.^[37] The sensitivity of the chondrocyte to its environment emphasizes the importance of selection of the appropriate material on which to grow the chondrocytes as well as the need to generate cartilage with the appropriate composition.

T

APPROACHES TO TISSUE ENGINEERING CARTILAGE

Tissue engineering of articular cartilage can be performed using a variety of strategies that can be divided into three categories as shown in Fig. 3. Tissue formation may occur *in vitro* or *in vivo* depending on the approach selected.

Cells

In this approach, cartilage repair is encouraged by placing a large number of cells in a defect. To accomplish this, the patient's cartilage is harvested, and cells are isolated and grown in monolayer culture (2D) to expand the number of chondrocytes [Fig. 3(A)]. The cells are then transplanted into a cartilage defect under a flap, which can be either periosteum or a collagen membrane and which serves to confine cells in the defect (autologous chondrocyte implantation).^[10,14–16] Alternatively, the cells can be seeded onto the membrane prior to implantation.^[42] The contribution of these cells in repairing is somewhat controversial. Although a recent study showed that the cells remain in the defect,^[43] it is possible that the mesenchymal cells from the periosteum and/or the underlying bone marrow (as the subchondral bone can be breached during the procedure) may also

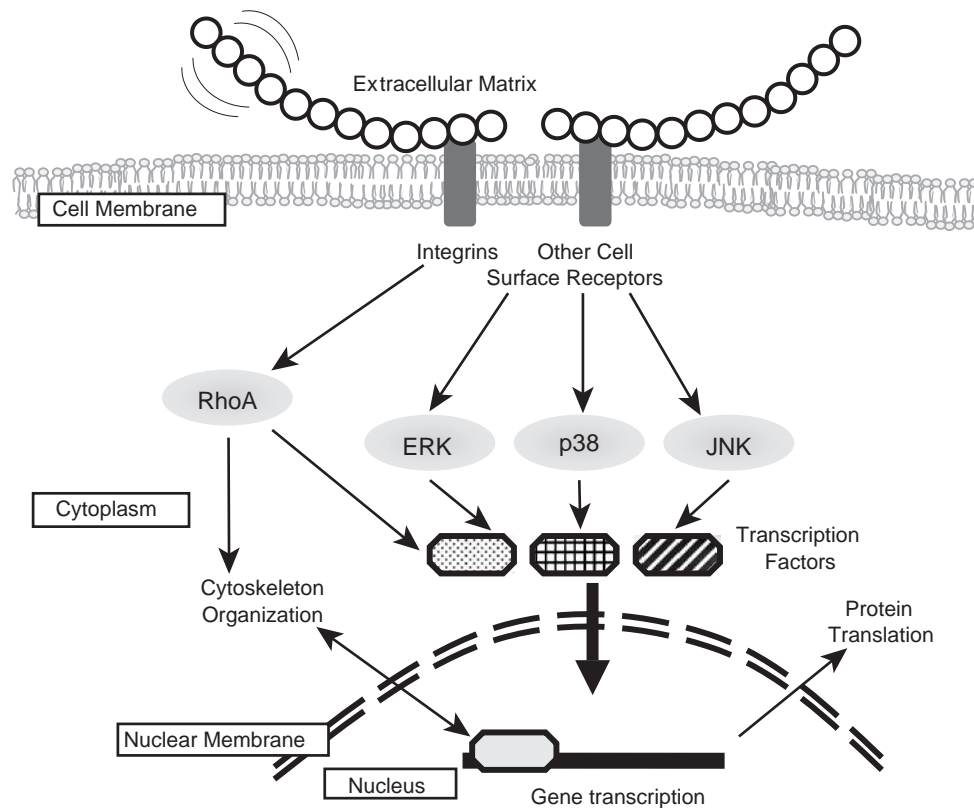


Fig. 2 Integration of cell signaling in chondrocytes. Changes in the ECM are perceived by receptors, such as integrins, on the cell surface. These signals are relayed inwards through secondary messengers ultimately resulting in gene and protein changes.

contribute to the repair as histological studies have documented the presence of fibrocartilage repair in many of these defects. This treatment is useful in alleviating pain but is not without its disadvantages.^[15] As the resulting tissue is at least partially fibrocartilage it may not remain intact over the long term. More recently, it was shown in a goat model that there was a loss of bone underlying the implant, which could negatively impact on implant survival and have long-term effects on joint function because of the altered biomechanics.^[44]

The benefits of growing cells in a 2D monolayer are that it is easy to manipulate, to expand cell number, and to treat cells with exogenous growth factors or other forms of stimulation when compared to cells grown in 3D culture. However, a major limitation of this approach is maintenance of the chondrocyte phenotype. Chondrocytes grown in monolayer at low density and/or after several passages will dedifferentiate. If this approach is to be used for tissue engineering, then monolayer conditions need to be identified that allow for expansion of the number of chondrocytes while retaining phenotype as well as maintenance of phenotype after transplantation of these cells.

Scaffolds

In this approach, tissue engineering cartilage entails the use of scaffolds. Scaffolds [Fig. 3(B)] are typically porous materials that may be untreated or coupled to chemicals, growth factors, oligonucleotides, silencing RNA oligonucleotides, plasmid DNA, or peptides with the goal of stimulating cell attachment and/or ingrowth as well as matrix accumulation.^[45] They may or may not (cell-free scaffolds) be seeded with cartilage-forming cells and either are placed directly into defects and the cartilage forms *in vivo* or the cell-seeded scaffold is grown in culture to form cartilage prior to implantation (see reviews^[45-48]). The porosity of the scaffold is important to facilitate cell permeation (*in vitro*) or ingrowth (*in vivo*, if necessary) and matrix accumulation as well as mass transport of nutrients into and waste products out of the tissue. Scaffolds can be made from a variety of materials ranging from naturally occurring materials such as collagen or silk to synthetic polymers such as polylactic acid.^[45,49] These scaffolds can be made in a variety of ways including standard solvent casting particulate leaching foam methods, melt molding, electrospinning, and solid freeform fabrication.^[45] They can be formed with

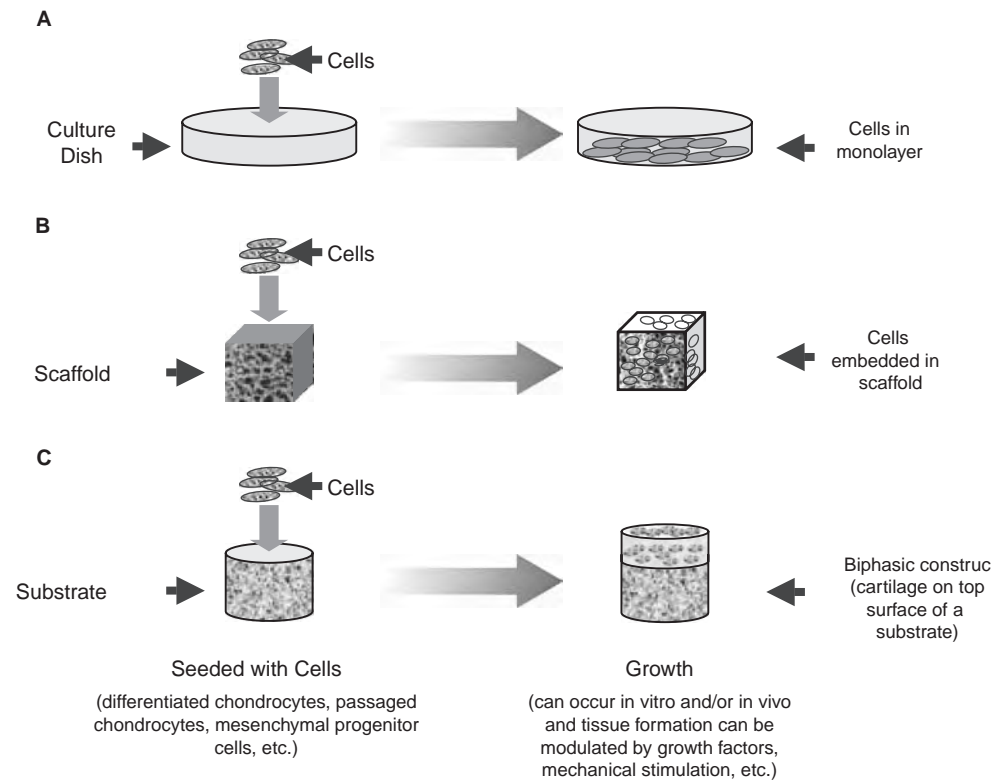


Fig. 3 Schematic of the three main methods that are used to tissue engineer articular cartilage. The duration of the culture period can vary considerably between all three methods; (A) Cells grown in monolayer are expanded and then harvested. The cells can be placed into articular cartilage defect; (B) Tissue-forming cells are seeded into a 3D scaffold. Cartilage may be generated in vitro or in vivo; and (C) Cartilage is formed on the top (the intended articulation) surface of a substrate. One example of this approach, and demonstrated in this figure, is to place chondrocytes on the substrate and cartilage tissue is allowed to form in vitro. The biphasic construct has a cartilage layer integrated with the top subsurface of the bone interfacing component. This composite construct can then be transplanted into a full thickness cartilage defect.

variable percent porosity and pore sizes ranging up to the micron size and in various physical forms such as a fleece or hydrogel. Much attention is now being directed toward the suitability of nanomaterials for cartilage tissue engineering.^[50] As well, design of “smart” scaffolds is also of great interest. For example, polymeric carriers can be created that have pH-sensitive membrane destabilizing activity to enhance delivery of biomolecules but are able to switch conformation with changes in pH so that they are not toxic to the cell^[51] or alternatively cross-linking oligopeptide metalloprotease substrates into 3D networks together with integrin-binding domains to create a condition that favors cell ingrowth.^[52] It would appear that some of the major benefits derived from growing chondrocytes in scaffolds are the 3D structure, which provides a setting conducive to the maintenance of cell phenotype, initial mechanical strength, and enhancement of cell retention in a defect.

Studies have shown that material geometry, pore size, and composition can affect the matrix molecules

produced by the chondrocytes, which could result in part from the interactions of chondrocytes with the material. To properly form articular cartilage it is necessary to ensure that the appropriate scaffold is selected as it is necessary to establish the conditions that will stimulate cells to produce the appropriate matrix molecules to ensure proper cell signaling and thus proper cartilage function. This includes selecting a material whose surface chemistry and roughness favors cell attachment and proper cell morphology (spread but not flattened with cortical actin distribution), which facilitates maintenance of chondrocyte phenotype, has an architecture that allows for formation of cartilage tissue with the organization described above, and has a degradation rate that matches the rate of tissue formation and does not incite an inflammatory reaction when implanted. Unfortunately, it appears that identification of optimal scaffold properties must be determined experimentally for each material developed.

Although most materials used are biodegradable, there are reports of non-biodegradable scaffolds, such

T

as agarose gels, that have been used with some success. Although cells seeded into agarose are distributed uniformly, their phenotype being maintained and can be injected into defects, the agarose can induce a giant cell foreign body response, suggesting that there are potential problems with its use.^[53]

There are several limitations in using scaffolds for cartilage repair. At present when cell-seeded scaffolds are allowed to form tissue *in vitro* prior to implantation the resultant tissue is of inferior quality with low collagen content and a fraction of the mechanical properties when compared to *in vivo* cartilage.^[54-57] This may be because of the absence of weight bearing during tissue formation, which is known to effect gene expression and matrix synthesis and can enhance nutrient diffusion. However, other factors such as scaffold breakdown can affect tissue formation. For example, as poly-L-lactic acid or its copolymer poly D,L-lactic glycolic acid degrade they release lactic acid, resulting in an acidic environment that can inhibit collagen synthesis by the cells and/or disrupt proteoglycan aggregation.^[58,59] As well, identifying the optimal scaffold for cartilage tissue engineering is still the subject of numerous investigations.

Substrates for Forming Biphasic Constructs

A third approach for cartilage tissue engineering is to generate biphasic constructs composed of a cartilage tissue phase overlying a substrate phase that serves as the bone interfacing component ("osteochondral-type" or biphasic constructs) [Fig. 3(C)]. Similar to scaffolds, substrates are porous and usually biodegradable, and can also be coupled to chemicals or biomolecules. However, it differs from scaffolds in that the cells or tissue is placed on the top surface and not seeded

throughout the substrate, and the regenerated cartilage is anchored to the intended articulation surface of the substrate. The substrate both supports cartilage formation and facilitates fixation after implantation by ingrowth of bone into the pores. An important characteristic of this substrate is that it is fully porous or the chondrocytes will not form cartilage.^[60] As well, the size and organization of the pores is important as they must allow for fluid flow but not allow (or severely restrict) cell infiltration throughout the full thickness of the substrate. The substrates have been made from many of the natural and synthetic materials utilized for scaffolds. However, as these are bone interfacing they must have sufficient strength for weight bearing, and so ceramics, processed bone, mineral reinforced polymers or collagen, and coral are used more commonly as substrates.^[10,61-63]

The cartilage component can be formed in several ways. Cartilage-forming cells can be seeded into a scaffold that is then affixed to the substrate either by sutures^[62] or by an adhesive such as fibrin glue,^[63] or if the substrate is a polymer by a concentration gradient of material properties.^[64] Alternatively, the cells can be placed directly on the top surface of the substrate. The biphasic construct can then be implanted, and cartilage tissue forms *in vivo*, or they can be grown in culture and cartilage allowed to form prior to implantation. An example of the latter approach entails the use of the ceramic porous calcium polyphosphate (CPP), which has mechanical properties approximating cancellous bone.^[65] Articular chondrocytes are placed on the intended articulation surface of the substrate, and as the cartilage forms *in vitro* the tissue integrates with the top subsurface of the scaffold (Fig. 4). The cartilage tissue that forms contains type II collagen and large proteoglycans similar to native cartilage.^[57] There are several potential benefits in using this type of biphasic

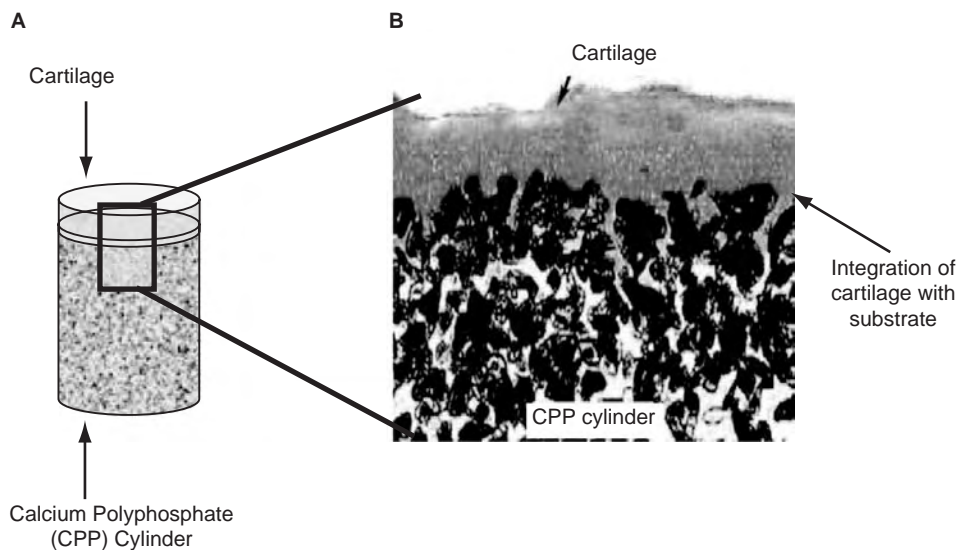


Fig. 4 An example of articular cartilage formed *in vitro* on a biodegradable substrate. (A) Diagram of the biphasic construct. (B) Photomicrograph of the histological appearance of the eight-week old cartilage grown on the surface of a porous CPP cylinder after formalin fixation and processing in Spurr resinTM.

construct to repair focal cartilage defects. Cartilage grown on substrates has the potential to form a stratified structure that more closely resembles native tissue. As the cartilage is already formed, lateral integration with the adjacent host cartilage is possible immediately upon implant insertion. As the substrate (CPP) is porous, bone can grow into pores not filled by cartilage after implantation, resulting in secure implant fixation. The CPP is biodegradable and will ultimately be replaced by bone, and as the biomaterial is made of calcium and phosphate the breakdown products do not incite an inflammatory reaction.^[65,66] However, similar to cartilage formed in scaffolds, the cartilage tissue and the cartilage–substrate interface formed after four weeks in culture have only a fraction of the mechanical properties of native tissue.^[57,67]

METHODS TO ASSESS BIOENGINEERED CARTILAGE

As there are many ways to tissue engineer cartilage, it is necessary to evaluate whether the approach under development generates cartilage with the appropriate qualities. Assessment of bioengineered cartilage can be done either following *in vitro* growth if appropriate or after implantation into the defect. Some of the methods that can be utilized are described below.

Pre-transplantation Assessment

Initial tests of the suitability of a biomaterial involve evaluation to determine if the material is toxic to chondrocytes. The 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrasodium bromide (MTT) assay, which measures mitochondrial activity, can be used to quantify cell viability. Live–dead assay (Molecular Probes) and confocal microscopy allow for visualization of the cells as well as quantification of the number of dead cells. Cellularity can also be determined using Hoechst dye binding assay to quantitate DNA content.^[68] Scanning electron microscopy is suitable to evaluate cell shape. Light microscopical examination is necessary to determine if cartilage tissue has formed and whether it approximates the appearance of *in vivo* tissue. Fig. 4 shows the histological appearance of the cartilage formed by chondrocytes seeded on top of a substrate. The type of molecules and their distribution in the tissue can be determined using immunohistochemical staining with antibodies reactive to the different macromolecules. The total amount of proteoglycan and collagen accumulated can be measured using biochemical assays (dimethylmethylene blue assay for glycosaminoglycans to estimate proteoglycan content^[69] and chloramine-T/Ehrlich's assay for

OH-proline to estimate collagen content).^[70] As treatment of cells with growth factors or other methods of stimulation can elicit a change in extracellular matrix synthesis these effects can be evaluated by semi-quantifying newly synthesized macromolecules and changes in gene expression.

Post-transplantation Assessment

After transplantation, histological evaluation should be used to assess the morphology and integration of the regenerated cartilage. As well, this will allow determination of the host response to the implant. If the tissue has formed *in vivo*, then it is necessary to characterize the tissue and confirm that it has the composition of hyaline cartilage and not fibrocartilage. The mechanical properties of the implant should be determined and compared to the native cartilage. This can also be evaluated *in situ* over time using arthroscopy and indentation probes.^[71,72]

MAJOR PROBLEMS ASSOCIATED WITH TISSUE ENGINEERING OF CARTILAGE

Appropriate Cell Source for Tissue Engineering Cartilage

One of the major limitations of cartilage bioengineering is obtaining sufficient numbers of chondrocytes to form enough cartilage to resurface a defect several centimeters in diameter. Although, cartilage and neonatal chondrocytes are considered immune privileged,^[73] recent studies have shown that implantation of cartilage between different species can incite an immune response that can compromise the implant.^[74] This suggests that xenogeneic sources may not be an option to obtain tissue. Development of transgenic pigs with “humanized” tissues has the potential to be an alternative source of cells, but this is still in the experimental phase because of the risk of viral disease transmission from pigs to humans. Current investigations into ways of preventing infection, such as RNA silencing, may prove successful, and in the future “humanized” animals may be able to provide chondrocytes or even cartilage for implantation. Thus, the major focus has been on developing a source of autologous cells. There are several ways to approach this.

Differentiated chondrocytes

To acquire chondrocytes, cells are isolated from cartilage obtained from around the defect or from a non-weight-bearing region. However, as only small amounts of tissue can be harvested a limited numbers



of cells are obtained, and as the individual gets older (+40 years) this problem is exacerbated as cell yield appears to decrease with age.^[75] In addition, cells from older individuals appear to be less robust in their ability to synthesize and accumulate extracellular matrix as determined by agarose gel culture studies.^[76] It has also been shown that chondrocytes obtained from arthritic cartilage, as may surround a joint defect, produce less extracellular matrix.^[75] There have been numerous investigations into culture methods that will allow for reliable expansion of chondrocyte number by cell proliferation while still preserving chondrogenic potential such that the cells can redifferentiate to chondrocytes under the appropriate conditions. Another approach to circumvent this problem has been to induce redifferentiation of the cells after the cell number has been expanded. This has been done by harvesting the cells from monolayer culture and then growing them under 3D conditions, at high cell density or on a material such as alginate.^[77] An alternative approach has been to harvest cells after growth on alginate and then culture them scaffold free on filter inserts to form tissue (the alginate-recovered-chondrocyte or ARC method).^[78]

Stem cells

Stem cells and progenitor cells are present in most if not all adult tissues and could serve as a source of cells. Stem cells give rise to progenitor cells and retain the capability of self-renewal. Progenitor cells can proliferate, have a limited capacity for self-renewal, and have the potential to differentiate. For example, mesenchymal stem cells (MSCs) can differentiate toward osteoblasts, chondrocytes, and adipocytes. MSCs, although initially identified in bone marrow, have since been isolated from a variety of tissues including periosteum, cancellous bone, cartilage, and even adipose tissue.^[79] The number of progenitor cells in a tissue is small (0.001–0.01% of nucleated cells in marrow)^[80] and varies from individual to individual and is influenced negatively by increasing age and presence of disease.^[81] Most of the work is focused on MSCs, and their role in tissue engineering is preliminary in part because of our limited understanding of the biology of these cells and because the markers used to separate out these cells from a mixed cell population have not been well defined for each subset. However, MSCs after expansion and seeding into a scaffold are able to differentiate toward chondrocytes under the appropriate culture conditions (see review^[81]). MSCs have also been used to repair cartilage defects with promising results,^[81,82] however, the cells can differentiate toward bone and cartilage. The use of MSCs avoids potential donor site morbidity that can

result from the harvesting of autologous cartilage, and as they escape immune surveillance, allogeneic cells could potentially be used.

An alternative to placing MSCs directly into a defect is to implant a scaffold that has been modified to stimulate MSC in situ to proliferate and migrate into the scaffold.^[83] Progenitor cells present in the superficial zone of cartilage may be particularly susceptible to this type of stimulus.^[10,84] As well, injection of stem cells into the knee joint has resulted in homing of cells to damaged cartilage and may be another mode of delivery of these cells to a defect.^[85]

Improving the Quality of Bioengineered Cartilage

Despite the development of potentially suitable biomaterials for tissue engineering, there still remains the problem that the articular cartilage generated in vitro does not mimic native cartilage in terms of composition, organization, and mechanical properties. There are several tactics that show promise in improving the quality of the tissue formed.

Mechanical stimulation

Under physiological conditions articular cartilage in vivo is exposed to complex loads and it has been shown that cartilage development and maintenance is regulated by these forces. Several studies have demonstrated that loading chondrocytes in a controlled manner in vitro leads to improved tissue formation with better physical properties. For example, Waldman et al. (2003) demonstrated that intermittent application of either cyclic compressive or shearing forces (400 cycles, 1 Hz) increased the amount of cartilage that formed and improved at least twofold both the load-bearing capacity and stiffness compared to unstimulated tissues formed on the CPP ceramic substrate.^[86] This would suggest that formation of tissue in the appropriate mechanical environment could be used to influence the quality of the cartilage that forms.

Growth factors

Addition of media supplements such as insulin-like growth factor I, osteogenic protein 1, and/or transforming growth factor- β (TGF- β) can improve the quality of in vitro-formed cartilage by stimulating production of specific matrix molecules.^[87,88] It may be that the addition of TGF- β coupled with mechanical stimulation may be an appropriate way to improve extracellular matrix accumulation.^[88]

Subpopulations of chondrocytes

Studies in our laboratory have shown that the quality of the cartilage formed *in vitro* can be influenced by the zone of the native cartilage from which the chondrocytes were obtained to establish the cultures.^[89] Chondrocytes obtained from the deep zone of cartilage accumulated more extracellular matrix and formed tissue that was thicker than that formed by cells harvested from full thickness or the mid/deep zones of articular cartilage. However it was the tissue formed by the mid/deep zone cells that had the best mechanical properties.^[89] This suggests that the use of chondrocyte subpopulations may be an interesting approach for tissue engineering. A study by Klein et al. demonstrated that it is possible to generate cartilage tissue *in vitro* by layering the cells from different zones in an appropriate sequential manner from deep to superficial to mimic the architectural features of the *in vivo* tissue.^[90] The stratified tissue although deficient in physical properties had a protein distribution similar to native tissue.^[90] Furthermore, using subpopulations will also allow for formation of tissue-engineered cartilage with a deep mineralized zone to recreate the soft tissue (cartilage) –hard tissue (bone) interface critical to proper cartilage function.^[89] So it may be advantageous to use chondrocyte subpopulations to modulate the composition and mechanical properties of the regenerated tissue.^[89,91]

CONCLUSIONS

There has been much research in tissue engineering of articular cartilage for joint repair. Current studies are aimed at improving the quality of tissue generated. However even when the optimal conditions to bio-engineer cartilage are developed there will be other problems to overcome such as integration of the repair tissue with the surrounding recipient tissues (cartilage and/or bone) after implantation and ensuring its long-term functionality *in vivo* under weight-bearing conditions in a “hostile” environment likely containing cytokines and metalloproteases because of the underlying cartilage damage.

ACKNOWLEDGMENTS

Owing to page constraints, we have been limited in the number of references that could be included, but we wish to acknowledge the many contributors to this field. As well, we thank all the students, fellows, and technologists in the Kandel, Pilliar, and Grynepas laboratories for the stimulating discussions and their

hard work, as well as CIHR, The Arthritis Society, and NSERC for their support.

ARTICLES OF FURTHER INTEREST

Tissue Engineering; Tissue Engineering of Bone; Tissue Engineering of Tendons

REFERENCES

- Hunziker, E.B. Articular cartilage repair: basic science and clinical progress. A review of the current status and prospects. *Osteoarthr. Cartil.* **2002**, *10* (6), 432–463.
- Hunziker, E.B. Articular cartilage repair: problems and perspectives. *Biorheology* **2000**, *37* (1–2), 163–164.
- Kinner, B.; Capito, R.M.; Spector, M. Regeneration of articular cartilage. *Adv. Biochem. Eng. Biotechnol.* **2005**, *94*, 91–123.
- Boyle, J.; Luan, B.; Cruz, T.F.; Kandel, R.A. Characterization of proteoglycan accumulation during formation of cartilagenous tissue *in vitro*. *Osteoarthr. Cartil.* **1995**, *3* (2), 117–125.
- Chuma, H.; Mizuta, H.; Kudo, S.; Takagi, K.; Hiraki, Y. One day exposure to FGF-2 was sufficient for the regenerative repair of full-thickness defects of articular cartilage in rabbits. *Osteoarthr. Cartil.* **2004**, *12* (10), 834–842.
- Lawrence, R.C.; Helmick, C.G.; Arnett, F.C.; Deyo, R.A.; Felson, D.T.; Giannini, E.H.; Heyse, S.P.; Hirsch, R.; Hochberg, M.C.; Hunder, G.G.; Liang, M.H.; Pillemer, S.R.; Steen, V.D.; Wolfe, F. Estimates of the prevalence of arthritis and selected musculoskeletal disorders in the United States. *Arthritis Rheum.* **1998**, *41* (5), 778–799.
- Barrack, R.L. Early failure of modern cemented stems. *J. Arthroplasty* **2000**, *15* (8), 1036–1050.
- Berry, D.J.; Harmsen, W.S.; Ilstrup, D.; Lewallen, D.G.; Cabanela, M.E. Survivorship of uncemented proximally porous-coated femoral components. *Clin. Orthop. Relat. Res.* **1995**, (319), 168–177.
- Jacobsson, S.A.; Djerf, K.; Wahlstrom, O. Twenty-year results of McKee-Farrar versus Charnley prosthesis. *Clin. Orthop. Relat. Res.* **1996**, (329 Suppl.), S60–S68.
- Redman, S.N.; Oldfield, S.F.; Archer, C.W. Current strategies for articular cartilage repair. *Eur. Cell Mater.* **2005**, *9*, 23–32.
- Shasha, N.; Aubin, P.P.; Cheah, H.K.; Davis, A.M.; Agnidis, Z.; Gross, A.E. Long-term clinical experience with fresh osteochondral allografts for articular knee defects in high demand patients. *Cell Tissue Bank* **2002**, *3* (3), 175–182.
- Smith, G.D.; Knutsen, G.; Richardson, J.B. A clinical review of cartilage repair techniques. *J. Bone Joint Surg. Br.* **2005**, *87* (4), 445–449.
- Langer, R.S.; Vacanti, J.P. Tissue engineering: the challenges ahead. *Sci. Am.* **1999**, *280* (4), 86–89.
- Brittberg, M.; Lindahl, A.; Nilsson, A.; Ohlsson, C.; Isaksson, O.; Peterson, L. Treatment of deep cartilage

- defects in the knee with autologous chondrocyte transplantation. *N. Engl. J. Med.* **1994**, *331* (14), 889–895.
15. Browne, J.E.; Anderson, A.F.; Arciero, R.; Mandelbaum, B.; Moseley J.B., Jr.; Micheli, L.J.; Fu, F.; Erggelet, C. Clinical outcome of autologous chondrocyte implantation at 5 years in US subjects. *Clin. Orthop. Relat. Res.* **2005**, (436), 237–245.
 16. Peterson, L.; Minas, T.; Brittberg, M.; Nilsson, A.; Sjogren-Jansson, E.; Lindahl, A. Two- to 9-year outcome after autologous chondrocyte transplantation of the knee. *Clin. Orthop. Relat. Res.* **2000**, (374), 212–234.
 17. Hasler, E.M.; Herzog, W.; Wu, J.Z.; Muller, W.; Wyss, U. Articular cartilage biomechanics: theoretical models, material properties, and biosynthetic response. *Crit. Rev. Biomed. Eng.* **1999**, *27* (6), 415–488.
 18. Archer, C.W.; Francis-West, P. The chondrocyte. *Int. J. Biochem. Cell Biol.* **2003**, *35* (4), 401–404.
 19. Heinegard, D.; Oldberg, A. Structure and biology of cartilage and bone matrix noncollagenous macromolecules. *FASEB J.* **1989**, *3* (9), 2042–2051.
 20. Roughley, P.J. Articular cartilage and changes in arthritis: noncollagenous proteins and proteoglycans in the extracellular matrix of cartilage. *Arthritis Res.* **2001**, *3* (6), 342–347.
 21. Giannoni, P.; Siegrist, M.; Hunziker, E.B.; Wong, M. The mechanosensitivity of cartilage oligomeric matrix protein (COMP). *Biorheology* **2003**, *40* (1–3), 101–109.
 22. Hecht, J.T.; Nelson, L.D.; Crowder, E.; Wang, Y.; Elder, F.F.; Harrison, W.R.; Francomano, C.A.; Prange, C.K.; Lennon, G.G.; Deere, M. Mutations in exon 17B of cartilage oligomeric matrix protein (COMP) cause pseudoachondroplasia. *Nat. Genet.* **1995**, *10* (3), 325–329.
 23. Wong, M.; Carter, D.R. Articular cartilage functional histomorphology and mechanobiology: a research perspective. *Bone* **2003**, *33* (1), 1–13.
 24. Schumacher, B.L.; Block, J.A.; Schmid, T.M.; Aydelotte, M.B.; Kuettner, K.E. A novel proteoglycan synthesized and secreted by chondrocytes of the superficial zone of articular cartilage. *Arch. Biochem. Biophys.* **1994**, *311* (1), 144–152.
 25. Lorenzo, P.; Aman, P.; Sommarin, Y.; Heinegard, D. The human CILP gene: exon/intron organization and chromosomal mapping. *Matrix Biol.* **1999**, *18* (5), 445–454.
 26. Li, L.; Buschmann, M.D.; Shirazi-Adl, A. The role of fibril reinforcement in the mechanical behavior of cartilage. *Biorheology* **2002**, *39* (1–2), 89–96.
 27. Oegema T.R., Jr.; Thompson R.C., Jr. The zone of calcified cartilage: its role in osteoarthritis. In *Articular Cartilage and Osteoarthritis*; Kuettner, K., Ed.; Raven Press: New York, U.S.A., 1992; 319–331.
 28. Mente, P.L.; Lewis, J.L. Elastic modulus of calcified cartilage is an order of magnitude less than that of subchondral bone. *J. Orthop. Res.* **1994**, *12* (5), 637–647.
 29. Park, S.; Hung, C.T.; Ateshian, G.A. Mechanical response of bovine articular cartilage under dynamic unconfined compression loading at physiological stress levels. *Osteoarthr. Cartil.* **2004**, *12* (1), 65–73.
 30. Thambyah, A.; Goh, J.C.; De, S.D. Contact stresses in the knee joint in deep flexion. *Med. Eng. Phys.* **2005**, *27* (4), 329–335.
 31. Mow, V.C.; Holmes, M.H.; Lai, W.M. Fluid transport and mechanical properties of articular cartilage: a review. *J. Biomech.* **1984**, *17* (5), 377–394.
 32. Carter, D.R.; Wong, M. Modelling cartilage mechanobiology. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2003**, *358* (1437), 1461–1471.
 33. Darling, E.M.; Athanasiou, K.A. Articular cartilage bioreactors and bioprocesses. *Tissue Eng.* **2003**, *9* (1), 9–26.
 34. Vunjak-Novakovic, G.; Martin, I.; Obradovic, B.; Treppo, S.; Grodzinsky, A.J.; Langer, R.; Freed, L.E. Bioreactor cultivation conditions modulate the composition and mechanical properties of tissue-engineered cartilage. *J. Orthop. Res.* **1999**, *17* (1), 130–138.
 35. Knudson, W.; Loeser, R.F. CD44 and integrin matrix receptors participate in cartilage homeostasis. *Cell Mol. Life Sci.* **2002**, *59* (1), 36–44.
 36. Millward-Sadler, S.J.; Salter, D.M. Integrin-dependent signal cascades in chondrocyte mechanotransduction. *Ann. Biomed. Eng.* **2004**, *32* (3), 435–446.
 37. Clancy, R.M.; Rediske, J.; Tang, X.; Nijher, N.; Frenkel, S.; Philips, M.; Abramson, S.B. Outside-in signaling in the chondrocyte. Nitric oxide disrupts fibronectin-induced assembly of a subplasmalemmal actin/rho A/focal adhesion kinase signaling complex. *J. Clin. Invest.* **1997**, *100* (7), 1789–1796.
 38. Whitmarsh, A.J.; Davis, R.J. Regulation of transcription factor function by phosphorylation. *Cell Mol. Life Sci.* **2000**, *57* (8–9), 1172–1183.
 39. Fanning, P.J.; Emkey, G.; Smith, R.J.; Grodzinsky, A.J.; Szasz, N.; Trippel, S.B. Mechanical regulation of mitogen-activated protein kinase signaling in articular cartilage. *J. Biol. Chem.* **2003**, *278* (51), 50,940–50,948.
 40. Wang, G.; Woods, A.; Sabari, S.; Pagnotta, L.; Stanton, L.A.; Beier, F. RhoA/ROCK signaling suppresses hypertrophic chondrocyte differentiation. *J. Biol. Chem.* **2004**, *279* (13), 13,205–13,214.
 41. Woods, A.; Wang, G.; Beier, F. RhoA/ROCK signaling regulates Sox9 expression and actin organization during chondrogenesis. *J. Biol. Chem.* **2005**, *280* (12), 11,626–11,634.
 42. Willers, C.; Chen, J.; Wood, D.; Xu, J.; Zheng, M.H. Autologous chondrocyte implantation with collagen bioscaffold for the treatment of osteochondral defects in rabbits. *Tissue Eng.* **2005**, *11* (7–8), 1065–1076.
 43. Dell'Accio, F.; Vanlauwe, J.; Bellemans, J.; Neys, J.; De Bari, C.; Luyten, F.P. Expanded phenotypically stable chondrocytes persist in the repair tissue and contribute to cartilage matrix formation and structural integration in a goat model of autologous chondrocyte implantation. *J. Orthop. Res.* **2003**, *21* (1), 123–131.
 44. Vasara, A.I.; Hyttinen, M.M.; Lammi, M.J.; Lammi, P.E.; Langsjo, T.K.; Lindahl, A.; Peterson, L.; Kellomaki, M.; Kontinen, Y.T.; Helminen, H.J.; Kiviranta, I. Subchondral bone reaction associated with chondral defect and attempted cartilage repair in goats. *Calcif. Tissue Int.* **2004**, *74* (1), 107–114.

45. Frenkel, S.R.; Di Cesare, P.E. Scaffolds for articular cartilage repair. *Ann. Biomed. Eng.* **2004**, *32* (1), 26–34.
46. Hollister, S.J. Porous scaffold design for tissue engineering. *Nat. Mater.* **2005**, *4* (7), 518–524.
47. Shin, H.; Jo, S.; Mikos, A.G. Biomimetic materials for tissue engineering. *Biomaterials* **2003**, *24* (24), 4353–4364.
48. Yarlagadda, P.K.; Chandrasekharan, M.; Shyan, J.Y. Recent advances and current developments in tissue scaffolding. *Biomed. Mater. Eng.* **2005**, *15* (3), 159–177.
49. Altman, G.H.; Diaz, F.; Jakuba, C.; Calabro, T.; Horan, R.L.; Chen, J.; Lu, H.; Richmond, J.; Kaplan, D.L. Silk-based biomaterials. *Biomaterials* **2003**, *24* (3), 401–416.
50. Ma, Z.; Kotaki, M.; Inai, R.; Ramakrishna, S. Potential of nanofiber matrix as tissue-engineering scaffolds. *Tissue Eng.* **2005**, *11* (1–2), 101–109.
51. Stayton, P.S.; El Sayed, M.E.; Murthy, N.; Bulmus, V.; Lackey, C.; Cheung, C.; Hoffman, A.S. ‘Smart’ delivery systems for biomolecular therapeutics. *Orthod. Craniofac. Res.* **2005**, *8* (3), 219–225.
52. Rosso, F.; Marino, G.; Giordano, A.; Barbarisi, M.; Parmeggiani, D.; Barbarisi, A. Smart materials as scaffolds for tissue engineering. *J. Cell Physiol.* **2005**, *203* (3), 465–470.
53. Rahfoth, B.; Weisser, J.; Sternkopf, F.; Aigner, T.; von der Mark, K.; Brauer, R. Transplantation of allograft chondrocytes embedded in agarose gel into cartilage defects of rabbits. *Osteoarthr. Cartil.* **1998**, *6* (1), 50–65.
54. Beekman, B.; Verzijl, N.; Bank, R.A.; von der Mark, K.; TeKoppele, J.M. Synthesis of collagen by bovine chondrocytes cultured in alginate; posttranslational modifications and cell-matrix interaction. *Exp. Cell Res.* **1997**, *237* (1), 135–141.
55. Martin, I.; Obradovic, B.; Treppo, S.; Grodzinsky, A.J.; Langer, R.; Freed, L.E.; Vunjak-Novakovic, G. Modulation of the mechanical properties of tissue engineered cartilage. *Biorheology* **2000**, *37* (1–2), 141–147.
56. Riesle, J.; Hollander, A.P.; Langer, R.; Freed, L.E.; Vunjak-Novakovic, G. Collagen in tissue-engineered cartilage: types, structure, and crosslinks. *J. Cell Biochem.* **1998**, *71* (3), 313–327.
57. Waldman, S.D.; Gryn timer, M.D.; Pilliar, R.M.; Kandel, R.A. Characterization of cartilagenous tissue formed on calcium polyphosphate substrates in vitro. *J. Biomed. Mater. Res.* **2002**, *62* (3), 323–330.
58. Kohn, D.H.; Sarmadi, M.; Helman, J.I.; Krebsbach, P.H. Effects of pH on human bone marrow stromal cells in vitro: implications for tissue engineering of bone. *J. Biomed. Mater. Res.* **2002**, *60* (2), 292–299.
59. Sittering, M.; Reitzel, D.; Dauner, M.; Hierlemann, H.; Hammer, C.; Kastenbauer, E.; Planck, H.; Burmester, G.R.; Bujia, J. Resorbable polyesters in cartilage engineering: affinity and biocompatibility of polymer fiber structures to chondrocytes. *J. Biomed. Mater. Res.* **1996**, *33* (2), 57–63.
60. Spiteri, C.G.; Pilliar, R.M.; Kandel, R.A. Substrate porosity enhances chondrocyte attachment, spreading, and cartilage tissue formation in vitro. *J. Biomed. Mater. Res. in press.*
61. Alhadlaq, A.; Mao, J.J. Tissue-engineered osteochondral constructs in the shape of an articular condyle. *J. Bone Joint Surg. Am.* **2005**, *87* (5), 936–944.
62. Schaefer, D.; Martin, I.; Jundt, G.; Seidel, J.; Heberer, M.; Grodzinsky, A.; Bergin, I.; Vunjak-Novakovic, G.; Freed, L.E. Tissue-engineered composites for the repair of large osteochondral defects. *Arthritis Rheum.* **2002**, *46* (9), 2524–2534.
63. van Susante, J.L.; Buma, P.; Homminga, G.N.; van den Berg, W.B.; Veth, R.P. Chondrocyte-seeded hydroxyapatite for repair of large articular cartilage defects. A pilot study in the goat. *Biomaterials* **1998**, *19* (24), 2367–2374.
64. Sherwood, J.K.; Riley, S.L.; Palazzolo, R.; Brown, S.C.; Monkhouse, D.C.; Coates, M.; Griffith, L.G.; Landeen, L.K.; Ratcliffe, A. A three-dimensional osteochondral composite scaffold for articular cartilage repair. *Biomaterials* **2002**, *23* (24), 4739–4751.
65. Pilliar, R.M.; Filiaggi, M.J.; Wells, J.D.; Gryn timer, M.D.; Kandel, R.A. Porous calcium polyphosphate scaffolds for bone substitute applications—in vitro characterization. *Biomaterials* **2001**, *22* (9), 963–972.
66. Gryn timer, M.D.; Pilliar, R.M.; Kandel, R.A.; Renlund, R.; Filiaggi, M.; Dumitriu, M. Porous calcium polyphosphate scaffolds for bone substitute applications in vivo studies. *Biomaterials* **2002**, *23* (9), 2063–2070.
67. Hamilton, D.J.; Seguin, C.A.; Wang, J.; Pilliar, R.M.; Kandel, R.A. Formation of a nucleus pulposus-cartilage endplate construct in vitro. *Biomaterials* **2006**, *27* (3), 397–405.
68. Kim, Y.J.; Sah, R.L.; Doong, J.Y.; Grodzinsky, A.J. Fluorometric assay of DNA in cartilage explants using Hoechst 33258. *Anal. Biochem.* **1988**, *174* (1), 168–176.
69. Goldberg, R.L.; Kolibas, L.M. An improved method for determining proteoglycans synthesized by chondrocytes in culture. *Connect. Tissue Res.* **1990**, *24* (3–4), 265–275.
70. Woessner, J.F. Determination of hydroxyproline content in connective tissues. In *The Methodology of Connective Tissue Research*; Hall, D.A., Ed.; Joynson-Bruvvers Limited: Oxford, 1976; 227–234.
71. Duda, G.N.; Kleemann, R.U.; Bluecher, U.; Weiler, A. A new device to detect early cartilage degeneration. *Am. J. Sports Med.* **2004**, *32* (3), 693–698.
72. Lyyra, T.; Jurvelin, J.; Pitkanen, P.; Vaatainen, U.; Kiviranta, I. Indentation instrument for the measurement of cartilage stiffness under arthroscopic control. *Med. Eng Phys.* **1995**, *17* (5), 395–399.
73. Newman, A.P. Articular cartilage repair. *Am. J. Sports Med.* **1998**, *26* (2), 309–324.
74. Stone, K.R.; Ayala, G.; Goldstein, J.; Hurst, R.; Walgenbach, A.; Galili, U. Porcine cartilage transplants in the cynomolgus monkey. III. Transplantation of alpha-galactosidase-treated porcine cartilage. *Transplantation* **1998**, *65* (12), 1577–1583.
75. Martin, J.A.; Buckwalter, J.A. Aging, articular cartilage chondrocyte senescence and osteoarthritis. *Biogerontology* **2002**, *3* (5), 257–264.



76. Tran-Khanh, N.; Hoemann, C.D.; McKee, M.D.; Henderson, J.E.; Buschmann, M.D. Aged bovine chondrocytes display a diminished capacity to produce a collagen-rich, mechanically functional cartilage extracellular matrix. *J. Orthop. Res.* **2005**, *23* (6), 1354–1362.
77. Homicz, M.R.; Chia, S.H.; Schumacher, B.L.; Masuda, K.; Thonar, E.J.; Sah, R.L.; Watson, D. Human septal chondrocyte redifferentiation in alginate, polyglycolic acid scaffold, and monolayer culture. *Laryngoscope* **2003**, *113* (1), 25–32.
78. Masuda, K.; Sah, R.L.; Hejna, M.J.; Thonar, E.J. A novel two-step method for the formation of tissue-engineered cartilage by mature bovine chondrocytes: the alginate-recovered-chondrocyte (ARC) method. *J. Orthop. Res.* **2003**, *21* (1), 139–148.
79. Minguell, J.J.; Erices, A.; Conget, P. Mesenchymal stem cells. *Exp. Biol. Med. (Maywood)* **2001**, *226* (6), 507–520.
80. Pittenger, M.F.; Mackay, A.M.; Beck, S.C.; Jaiswal, R.K.; Douglas, R.; Mosca, J.D.; Moorman, M.A.; Simonetti, D.W.; Craig, S.; Marshak, D.R. Multilineage potential of adult human mesenchymal stem cells. *Science* **1999**, *284* (5411), 143–147.
81. Barry, F.P.; Murphy, J.M. Mesenchymal stem cells: clinical applications and biological characterization. *Int. J. Biochem. Cell Biol.* **2004**, *36* (4), 568–584.
82. Tatebe, M.; Nakamura, R.; Kagami, H.; Okada, K.; Ueda, M. Differentiation of transplanted mesenchymal stem cells in a large osteochondral defect in rabbit. *Cytotherapy* **2005**, *7* (6), 520–530.
83. Bensaid, W.; Triffitt, J.T.; Blanchat, C.; Oudina, K.; Sedel, L.; Petite, H. A biodegradable fibrin scaffold for mesenchymal stem cell transplantation. *Biomaterials* **2003**, *24* (14), 2497–2502.
84. Dowthwaite, G.P.; Bishop, J.C.; Redman, S.N.; Khan, I.M.; Rooney, P.; Evans, D.J.; Haughton, L.; Bayram, Z.; Boyer, S.; Thomson, B.; Wolfe, M.S.; Archer, C.W. The surface of articular cartilage contains a progenitor cell population. *J. Cell Sci.* **2004**, *117* (Pt 6), 889–897.
85. Murphy, J.M.; Fink, D.J.; Hunziker, E.B.; Barry, F.P. Stem cell therapy in a caprine model of osteoarthritis. *Arthritis Rheum.* **2003**, *48* (12), 3464–3474.
86. Waldman, S.D.; Spiteri, C.G.; Grynepas, M.D.; Pilliar, R.M.; Hong, J.; Kandel, R.A. Effect of biomechanical conditioning on cartilaginous tissue formation in vitro. *J. Bone Joint Surg. Am.* **2003**, *85-A* (Suppl. 2), 101–105.
87. Darling, E.M.; Athanasiou, K.A. Biomechanical strategies for articular cartilage regeneration. *Ann. Biomed. Eng.* **2003**, *31* (9), 1114–1124.
88. Mauck, R.L.; Nicoll, S.B.; Seyhan, S.L.; Ateshian, G.A.; Hung, C.T. Synergistic action of growth factors and dynamic loading for articular cartilage tissue engineering. *Tissue Eng.* **2003**, *9* (4), 597–611.
89. Waldman, S.D.; Grynepas, M.D.; Pilliar, R.M.; Kandel, R.A. The use of specific chondrocyte populations to modulate the properties of tissue-engineered cartilage. *J. Orthop. Res.* **2003**, *21* (1), 132–138.
90. Klein, T.J.; Schumacher, B.L.; Schmidt, T.A.; Li, K.W.; Voegtline, M.S.; Masuda, K.; Thonar, E.J.; Sah, R.L. Tissue engineering of stratified articular cartilage from chondrocyte subpopulations. *Osteoarthr. Cartil.* **2003**, *11* (8), 595–602.
91. Kim, T.K.; Sharma, B.; Williams, C.G.; Ruffner, M.A.; Malik, A.; McFarland, E.G.; Elisseeff, J.H. Experimental model for cartilage tissue engineering to regenerate the zonal organization of articular cartilage. *Osteoarthr. Cartil.* **2003**, *11* (9), 653–664.

Tissue Engineering of Cornea

Lucie Germain

Departments of Surgery and Ophthalmology, and Laboratoire d'Organogenese Experimentale (LOEX), Hopital du Saint-Sacrement du CHA, LOEX, Hôpital du Saint-Sacrement du CHA, and Laval University, Quebec City, Quebec, Canada

Claude J. Giasson

Department of Ophthalmology, and Organogenese Experimentale (LOEX), Hopital du Saint-Sacrement du CHA, LOEX, Hôpital du Saint-Sacrement du CHA, and Laval University, Quebec City, Canada, and Ecole d' Optometrie, University of Montreal, Montreal, Canada

Patrick Carrier

Departments of Surgery and Ophthalmology, and Laboratoire d'Organogenese Experimentale (LOEX), Hopital du Saint-Sacrement du CHA, LOEX, Hôpital du Saint-Sacrement du CHA, and Laval University, Quebec City, Quebec, Canada

Sylvain L. Guérin

Department of Ophthalmology and Laboratoire d'Endocrinologie Moleculaire, LOEX, Hôpital du Saint-Sacrement du CHA, and Laval University, Quebec City, Quebec, Canada

Christian Salette

Department of Ophthalmology, Unité de recherche en Ophtalmologie, Centre de Recherche du CHUL, Hôpital du Saint-Sacrement du CHA, and Laval University, Quebec City, Quebec, Canada

François A. Auger

Departments of Surgery and Ophthalmology, and Laboratoire d'Organogenese Experimentale (LOEX), Hopital du Saint-Sacrement du CHA, LOEX, Hôpital du Saint-Sacrement du CHA, and Laval University, Quebec City, Quebec, Canada



INTRODUCTION

The cornea is the transparent barrier between the eye and the environment. Tissue-engineered corneas are currently developed to replace wounded or diseased corneas. Various experimental applications are also foreseen for these tissues reconstructed in vitro by tissue engineering. This article covers the first human corneas reconstructed by tissue engineering from normal human cells and the different models used for the production of human and animal corneas in vitro. Corneal injury and the activation of the complex wound-healing mechanisms are also addressed. Finally, we will attempt to provide the reader with a brief look toward the future of corneal tissue engineering, including the challenges that lie ahead as well as the potential experimental and clinical applications of this field.

CORNEAL TISSUE RECONSTRUCTION

Living tissue reconstruction rests mainly on a good choice of cells and biomaterials because these components will affect the quality of the final product. The ideal cell source should provide cells with extensive proliferation potential and appropriate differentiation

abilities. Stem cells possess such characteristics. After isolation and culture, cells can be combined with biomaterials to produce three-dimensional constructs.

Cell Source

The corneal epithelium is a transparent mucous tissue composed of a nonkeratinizing-stratified squamous epithelium (Fig. 1) comprising five cell layers centrally and up to 10 peripherally. It is surrounded by a small transitional zone, the limbus, and by the conjunctiva. The latter will not be discussed further since this nontransparent tissue represents a different type of epithelium containing goblet cells that cannot restore a normal central cornea. Corneal cells from the basal layer are polygonal and become more flattened in the suprabasal layers. All cells are nucleated. The stromal tissue underlying the epithelium as well as the basement membrane varies according to the three regions of the cornea. The limbus and conjunctiva are highly vascularized. In contrast, blood vessels are absent in the central cornea. However, endothelial cells form a monolayer underlying the stroma in this region. This cell layer is very important as it is involved in the maintenance of corneal transparency by regulating its

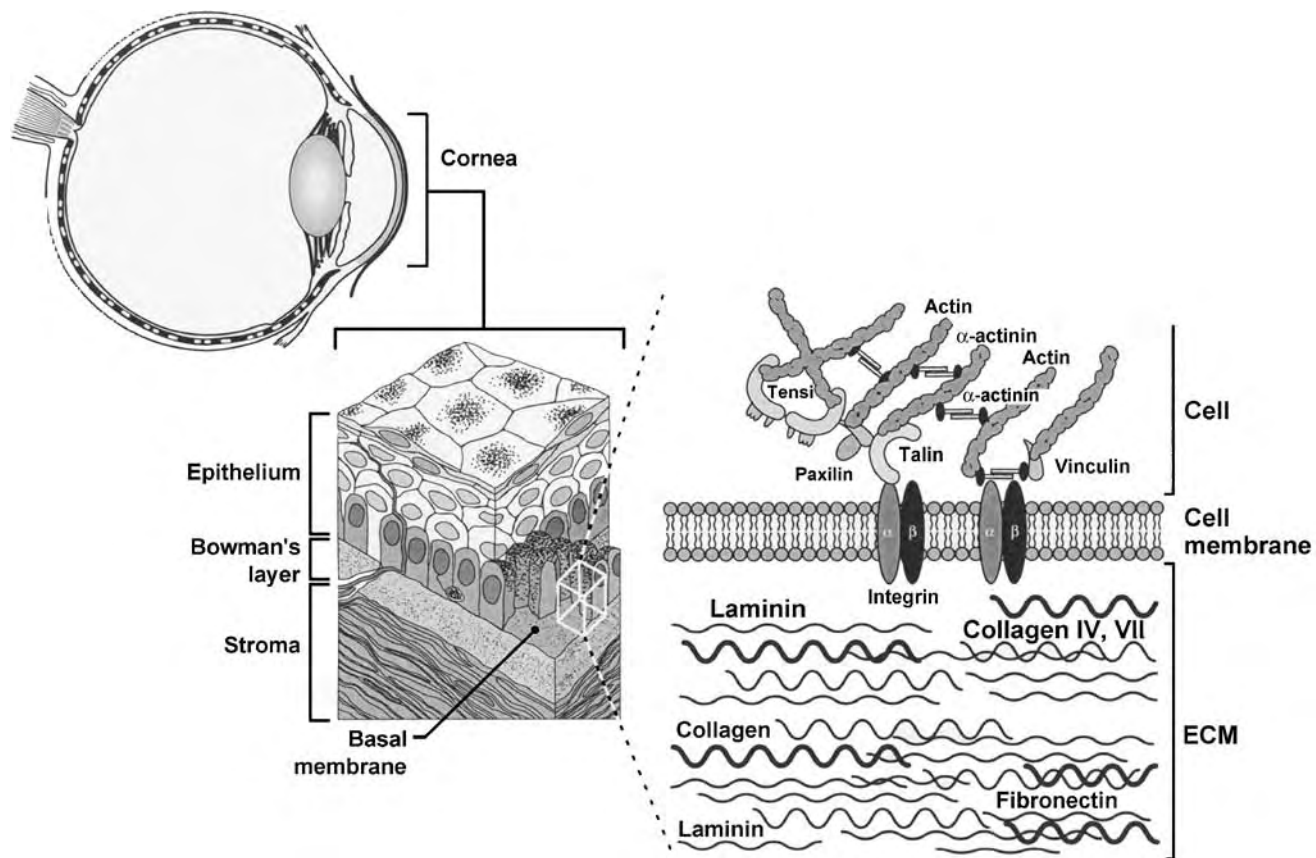


Fig. 1 Organization of the corneal epithelium and its interaction with basement membrane components. The stratified corneal epithelium is depicted along with the Bowman's layer and the most outer part of the corneal stroma. The enlarged section of the basal corneal epithelial cells illustrates how membrane-bound integrins bridge the components of the ECM (for instance, laminin, fibronectin, collagen IV, and collagen VII) to the actin filaments through their interaction with cytoskeletal elements (talin and α -actinin). Formation of functional focal adhesion structures also requires the participation of other cytoskeletal components such as tensin, vinculin, and paxillin. The end result of integrins clustering at the cell membrane and the stabilization of their intracellular connection with actin filaments through the selective use of these cytoskeletal elements is the formation of an adhesive junction that can withstand stable attachment of the cell to the ECM. (Taken from Ref. [6] with permission.)

degree of hydration. Therefore, the central and limbal cornea comprise three cell types: epithelial cells, fibroblasts (also called keratocytes), and endothelial cells. The first step of living tissue engineering is determining the best cell source for tissue reconstruction, the ideal source being able to regenerate the tissue for a lifetime as well as producing a differentiated progeny. Then, tissue engineering requires the isolation, culture, and expansion of each cell type on plastic substrates before three-dimensional tissue reconstruction.

Epithelial cells

The corneal epithelium is a self-renewing tissue in which the balance between proliferation and differentiation

maintains the homeostasis. However, the stem cells, ultimately responsible for repopulating the cornea, are not uniformly distributed on its surface. They are restricted to the limbus, the transitional zone surrounding the central cornea.^[1] They migrate in a swirling circular movement to the basal layer of the central cornea and then move through the suprabasal compartment toward the corneal surface.^[2] All cells from the central cornea express the keratin pair, K3/K12, which is typical of terminal differentiation of the corneal epithelium.^[3] Despite this keratin expression, cells from the basal layer of the central cornea can proliferate, but only for a limited number of divisions; they are transitional cells and not stem cells. In contrast, basal cells do not contain K3/K12, but express K19 in the region containing stem cells, the limbus.^[4-6] Interestingly, this keratin is expressed in stem cells of the skin.^[7] Keratin

19 was shown to possess distinct assembly properties and to form narrower filaments *in vitro*. Thus, keratin 19 may impart unique properties to the basal cells expressing it.^[8]

Human epithelial cells can be isolated from the central cornea or from the limbus by dispase digestion. They are cultured with the method originally described by Green for the clonal growth of skin keratinocytes.^[9] The coculture with an irradiated fibroblasts feeder layer in a medium containing growth factors allows expansion of corneal epithelial cells *in vitro*.^[10,11] In addition to the promotion of epithelial cell growth, the presence of irradiated fibroblasts limits the outgrowth of human fibroblasts that may have been isolated with the epithelial cells. The high concentration of calcium in the culture medium sustains cell differentiation and allows the formation of desmosomes essential to the production of a multilayered epithelium.

In culture, human epithelial cells form colonies that are several layers thick. The proliferative cells are located at the edge of the colonies, in the basal layer where they attach to the plastic substrate. The differentiated cells can detach from the plastic substrate and leave the basal layer to form the suprabasal layers where cells are bound to their neighbors by desmosomes. Therefore, cultures performed under these conditions contain a mixture of epithelial cells with various differentiation levels. The behavior of K19 positive limbal cells in culture remains to be

determined. Interestingly, cells isolated from the limbus grow much better in culture than cells from the central cornea, suggesting that stem cells are preserved in culture. Cells from the central cornea have a short lifetime in culture as they can barely be subcultured. In contrast, limbal cells can be subcultured for several passages (up to seven passages, depending on the donor). Moreover, cells isolated from the central cornea present morphological characteristics consistent with their low proliferative potential. They form small colonies containing large cells with a low nucleus-to-cytoplasm ratio, indicating that they are differentiated. In contrast, limbal cells produce large colonies consisting of small cells that possess a high nucleus-to-cytoplasm ratio (Fig. 2). These characteristics are typical of less differentiated cells.

The lifespan of primary cultured human corneal epithelial cells varies with the donors. This variation has been correlated with expression of the transcription factors of the Sp1 family.^[12] Therefore, the expression of Sp1/Sp3 might represent a good predictor for selecting human corneal epithelial cell cultures that are most likely to proliferate, stratify, and differentiate properly when used for the production of reconstructed corneal substitutes.

Within the limbus, a higher colony-forming efficiency is attributed to cells that have been isolated from the superior and temporal quadrants compared to cells from nasal or inferior limbus,^[13] indicating that

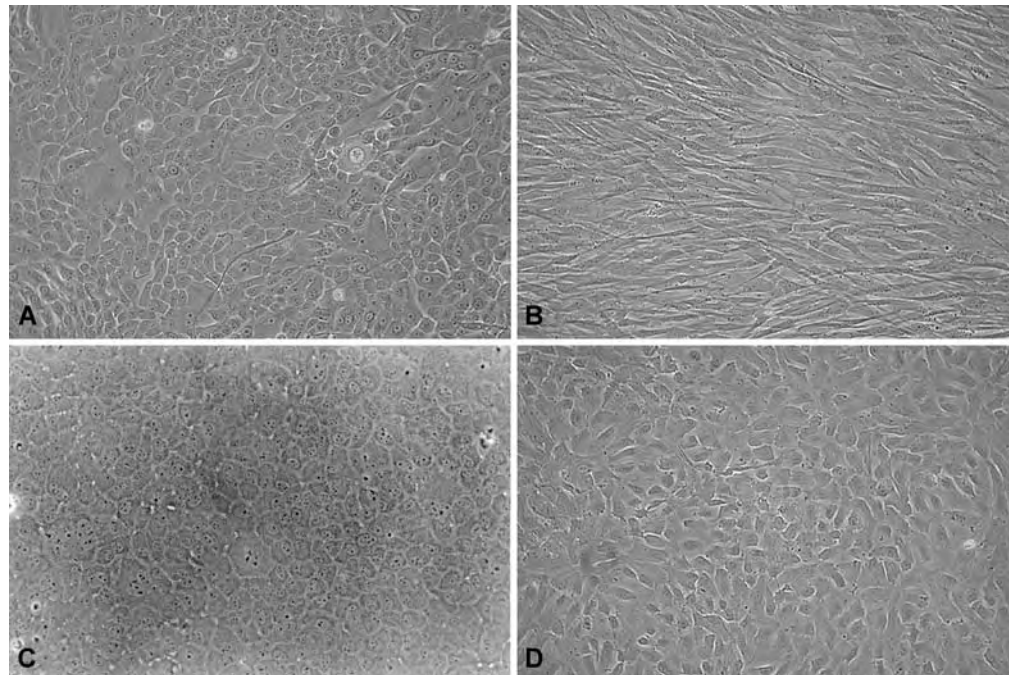


Fig. 2 Phase contrast micrographs of epithelial cells isolated from limbus cultured on plastic substrate with irradiated mouse 3T3 fibroblasts (A), corneal stromal cells (B), and bovine (C) or human (D) corneal endothelial cells. Note that limbal epithelial cells form large colonies of small cells surrounded by irradiated fibroblasts. Magnification = $\times 70$.

biopsies should preferably be taken from these two former sites. The preferential localization of stem cells in the superior quadrant is consistent with a better protection against environmental and mechanical injuries provided by the eyelids in this region.

Keratocytes

Human corneal keratocyte cultures can be established from the corneal stroma that is left after dispase digestion and removal of the epithelium and endothelium. Stromal explants are seeded in DMEM (Dulbecco–Vogt modification of Eagle’s culture medium) supplemented with 10% fetal calf serum and antibiotics. Stromal cells are subcultured in the same medium and undergo more than nine passages. In this culture medium, stromal cells adopt a fibroblastic morphology after subculturing (Fig. 2).

Alternatively, human corneal keratocytes can be obtained either after digestion of bare corneal stroma with collagenase or from stromal cell contamination in epithelial cell cultures by changing the culture medium for keratocyte culture medium. Epithelial cells are lost in subcultures because this medium is not adequate for their proliferation.

Endothelial cells

As opposed to the other corneal cell types, mature human corneal endothelium has a very limited capacity of division under normal conditions. However, these mesothelial cells, arrested in the G1 phase of the cell cycle, have retained a proliferative potential, as indicated by positive reaction against cyclins D, E, and A.^[14] As with epithelial cells, incubation in a dispase solution allows the endothelial sheet to separate easily from the underlying stroma. The endothelium is grown in a 1:1 combination of medium M199: Ham’s F12 supplemented with fetal calf serum, ascorbic acid, insulin, transferrin, penicillin, and gentamicin.^[15] Under these conditions, we were able to grow corneal cells from very young donors (2 or 3 years old). On the other hand, endothelial cells from adult corneas are more difficult to grow in culture. Alternatively, DMEM, supplemented with fetal calf serum, endothelial cell growth supplement, penicillin, and gentamicin, may be used to culture bovine endothelial cells (Fig. 2).

Clinical Application of Corneal Epithelial Cells Cultured on Fibrin Gels

The expansion of corneal epithelial cells in culture was applied for the treatment of patients suffering from conjunctivalization following chemical burn of the

cornea.^[16] Corneal cells harvested from an intact portion of the limbus from the contralateral eye were expanded in culture as described earlier. After the primary culture, cells were seeded on an acellular fibrin gel and cultured until confluence. The grafting of cultured epithelium resulted in a permanent corneal epithelium and the improvement of the visual quality for the patients.^[16] These results indicate that stem cells are conserved in cultures and remain functional after grafting.

Corneal Tissue Reconstruction Using Collagen Gels

The collagen gel model allows the production of living corneal tissue comprising both the reconstructed stroma and epithelium. This model was first optimized for skin.^[17,18] The reconstructed stroma is produced by mixing stromal cells with collagen. The epithelium is obtained by culturing corneal epithelial cells on the mature reconstructed stroma (Fig. 3).

Several models have been produced *in vitro* using animal corneal cells. Animal models with two or three reconstructed layers (endothelium, stroma, and epithelium) were produced with rabbit or bovine cells. Zieske et al. (1994) used rabbit corneal epithelial cells over rabbit stromal keratocytes included in a collagen gel with or without an underlying layer of immortalized mouse corneal endothelial cells.^[19] Minami et al. (1993) reported another model with the three cell types of bovine origin.^[20] The following paragraphs present the first model reconstructed from normal human corneal stromal and epithelial cells.

The first step is producing the reconstructed stroma. Corneal stromal cells are mixed with bovine type I collagen, adjusting the pH to 7.4. The mixture is then poured into a petri dish containing an anchorage ring and incubated 1 h at 37°C for gelation. Alternatively, bovine collagen can be replaced by human type I and type III collagen to produce an entirely human substitute.^[11] The addition of a filter paper ring allows the anchorage of the gel to prevent the inevitable contraction of floating collagen gels that is proportional to the cell number and inversely proportional to collagen concentration.^[11] The anchorage was designed to keep a constant surface area and to reduce the thickness without interference of the peripheral filter paper with the cells. The collagen gel produced by this process contains dispersed living stromal cells. The reconstructed stroma, (Fig. 4) resembles the corneal stroma, but the Descemet’s and Bowman’s membranes are absent in this first model of normal corneal cells.

The production of the epithelium was initiated by cultured epithelial cells from the limbus. They were chosen because of their greater proliferative potential. They are seeded on mature reconstructed stroma

HUMAN CORNEA RECONSTRUCTION BY THE COLLAGEN GEL METHOD

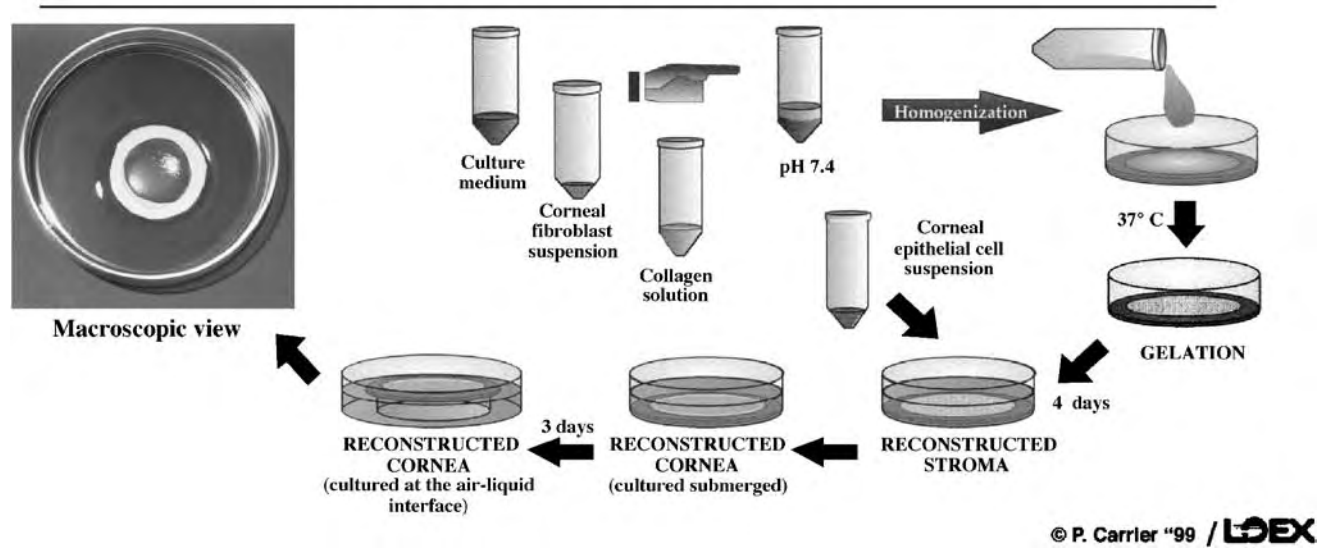


Fig. 3 Schematic representation of the production of reconstructed cornea by the collagen gel model, showing also a macroscopic view of the reconstructed cornea. The original color figure has been published in *Science* (1999) **284**, 423.

(4-day culture in stromal cell culture media). With the addition of epithelial cells, the culture medium is changed for epithelial cell culture media, which contains EGF, insulin, hydrocortisone, cholera toxin, and serum.^[11] Under these conditions, epithelial cells attach to the reconstructed stroma and uniformly cover it. After 3 days of culture submerged in the medium, it forms a multilayered epithelium with a basal layer containing cuboidal cells. Cells become elongated as they progress into the suprabasal layers (Fig. 4). The cultures can be raised at the air-liquid interface to promote epithelial cell differentiation, as shown for reconstructed skin,^[21,22] and later used for reconstructed cornea.^[6,19,20] Obviously, human or bovine endothelial cells may be seeded over this stromal equivalent covered with epithelial cells. The medium is changed afterward to a 1:1 mixture of epithelial and endothelial medium.^[23] In an animal model, the addition of endothelial cells under a collagen matrix improves differentiation of corneal epithelium as well as the distribution of keratocytes within the gel.^[19]

The resulting reconstructed cornea presents a translucent macroscopic aspect (Fig. 3). The thickness of the reconstructed epithelium was found to be close to that of normal corneas upon histological examination (Fig. 4).

Corneal Tissue Reconstruction Using Collagen Sponges

Another method to produce artificial corneas was described by Griffith et al.^[24] It takes advantage of

natural or synthetic polymers as a scaffold to seed cell lines of transformed human cells for the production of a tissue.^[24,25] Collagen-based polymers are cross-linked with glutaraldehyde or other cross-linkers to produce a sponge into which virally transformed keratocytes are seeded. The addition of transformed corneal endothelial and epithelial cells leads to artificial corneas that presented functional properties as shown by a decrease in transparency following their response to chemical wounding.

Corneal Tissue Reconstruction by the Self-Assembly Approach

The self-assembly approach constitutes a truly new concept for tissue engineering. We have originally designed it to reconstruct tissues in a manner akin to the embryonic development phenomenon, which would result in the creation of living substitutes as similar as possible to their physiological *in vivo* counterparts. Indeed, in this approach, cells produce their own extracellular matrix and organize it into a structured three-dimensional network. Addition of the adequate cell types results in the production of complex tissue with the expected histological, mechanical, and functional properties, as was previously demonstrated by our group for blood vessels and skin.^[21,26] There is no need for providing exogenous collagen or synthetic material.^[21] The vascularization system can also be reconstructed, as demonstrated by the formation of a

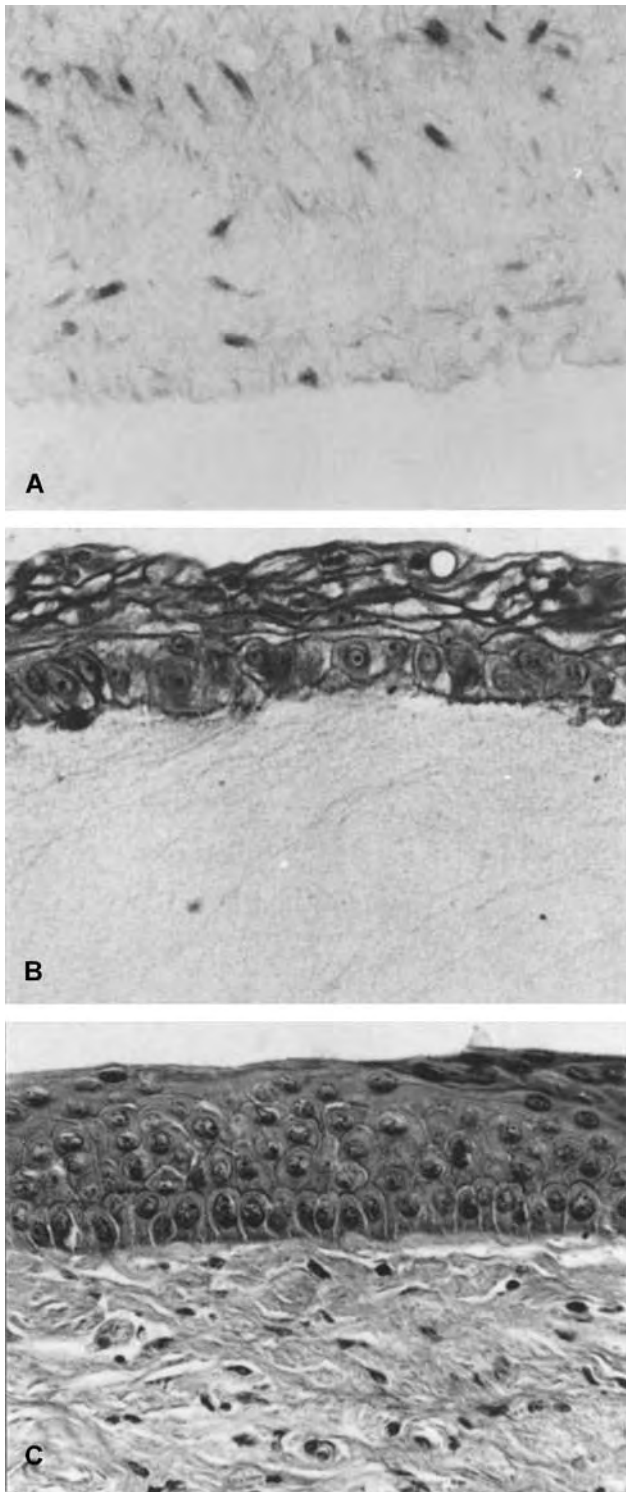


Fig. 4 (A) Masson's trichrome staining of reconstructed human stroma produced with bovine collagen. Note the presence of keratocytes and extracellular matrix. Masson's trichrome staining of reconstructed human corneas produced with human collagen (B) and normal human cornea in situ (C). Histological sections of reconstructed corneas after 3 days of epithelium culture under submerged conditions. Magnification = $\times 243$. (Taken from Ref. [11] with permission.)

capillary-like network in our skin construct.^[27] This promising new technology is now applied for the reconstruction of human cornea^[28,29] and results in a transparent tissue (Fig. 5). In this model, the stromal cells produce a dense matrix (Fig. 6B). Moreover, the epithelial-mesenchymal interactions lead to the formation of a complete basement membrane with all expected ultrastructural components (lamina lucida, lamina densa, hemidesmosomes) (Fig. 6A).

WOUND HEALING

Corneal Injury and Activation of the Wound-Healing Mechanisms

One major aim in developing a tissue-engineered human reconstructed cornea is its use as a corneal substitute for those individuals afflicted with severe corneal wounds. Corneal wounds, either superficial with damage limited to the epithelium or extensive with defects of basement membrane and stromal lamellae, account for approximately 37% of all visual disabilities and 23% of medical consultations for ocular problems in North America. Severe recurrent and persistent corneal wounds are most commonly secondary to recurrent erosion, mild chemical burns, superficial herpetic infections, neuroparalytic cornea, autoimmune diseases, and stromal ulcerations due to viral or bacterial infections or to severe burns (reviewed in Reim et al., 1997).^[30] Many of these corneal wounds and scarring diseases persist for weeks and months or else recur frequently and can progress to corneal perforation. Those eyes that survive perforation will eventually develop corneal scars, which will compromise vision and may also alter the mobility of the globe and conjunctiva.

The use of techniques for refractive surgery such as excimer laser photorefractive keratectomy (PRK), automated lamellar keratoplasty (ALK), and laser in situ keratomileusis (LASIK) has become popular worldwide. Individual patients have shown an aggressive wound-healing response with excessive myopic regression and severe corneal haze after PRK as well as a severe form of lamellar keratitis after LASIK.^[31] Stromal keratocytes, together with epithelial cells, are also likely to play a critical function in defining the visual outcome of refractive surgery. Corneal injury also induces a series of phenotypic changes, ultimately resulting in the transformation of stromal keratocytes into a fibroblastic cell type that shares characteristics of myofibroblasts. This transformation process and the subsequent physiological activity of these cells ultimately determine the final outcome of refractive surgical procedures.

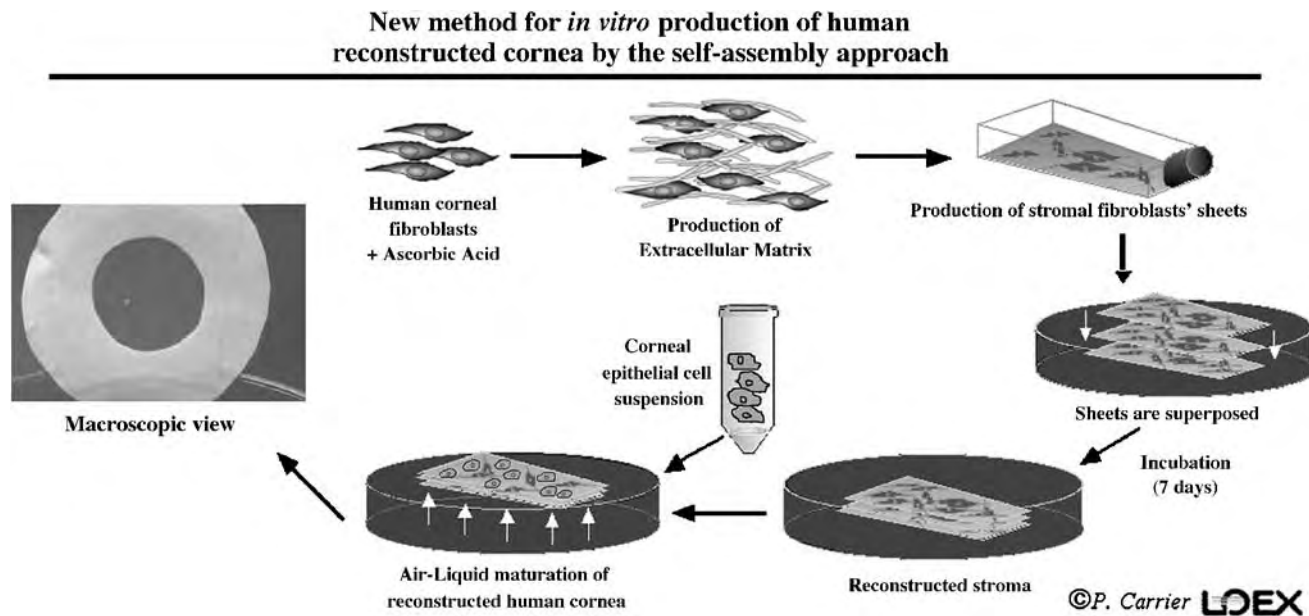


Fig. 5 Schematic representation of the production of reconstructed cornea by the self-assembly approach, showing also a macroscopic view of the reconstructed cornea.

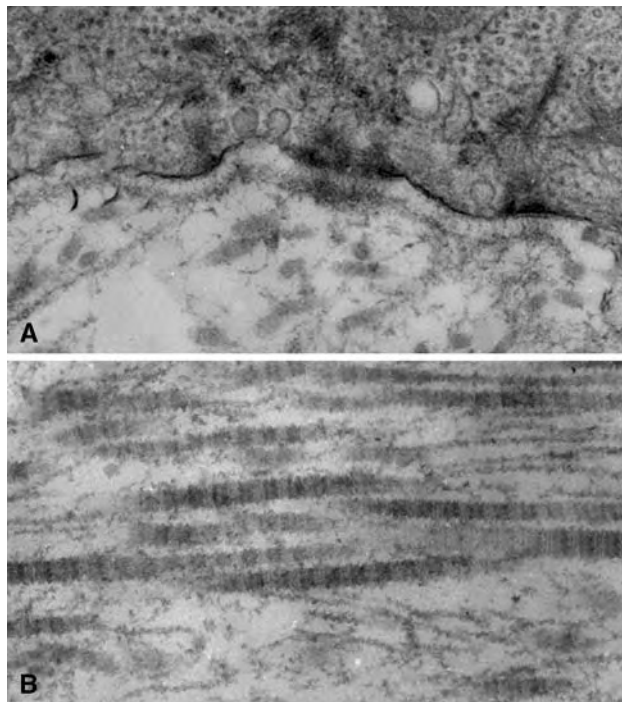


Fig. 6 Transmission electron microscopy of reconstructed human skin tissue by the self-assembly approach, showing the presence of the basement membrane between the stroma and the epithelium (A), and the extracellular matrix organization including the typical pattern of collagen bundles (B).

Wound healing in the cornea is a complex process that requires the formation of scar tissue as well as the restoration and maintenance of the transparency of the cornea. Healing of the ocular surface is closely related to the regeneration of the corneal and conjunctival epithelia. Tissue repair obviously requires cell migration, proliferation, and adhesion. Cell adhesion and migration in turn depend on the synthesis and assembly of the extracellular matrix (ECM). The ECM is a complex, cross-linked structure made up of various proteins such as fibronectin (FN), laminin (LM), collagen, and vitronectin. It organizes the geometry of normal tissues through cell-cell and cell-ECM adhesion. Both FN and LM are adhesion proteins identified as potential wound-healing agents because of their cell-attachment, migration, differentiation, and orientation properties. Cells bind the different adhesion proteins with membrane proteins better known as the integrin superfamily of receptors, which recognizes the RGD tripeptide that is shared by many adhesion proteins. These integrins bridge the cytoskeletal elements of cells to the ECM and thus promote cell-ECM and cell-cell interactions (Fig. 1).

Integrin Expression in Normal Cornea and in Wound Healing

Integrins are now recognized as the main class of cell adhesion receptors for the various components of the

ECM. Therefore, the patterns of integrin expression of human corneal substitutes must be very similar to normal human cornea in order to save its many adhesive and migratory characteristics. The integrin superfamily contains at least eight types of receptor subfamilies, which differ from one another only by the nature of their β subunit. To date, 18 α and 8 β integrin subunits that can heterodimerize into more than 25 known integrin receptors have been described.^[32] Both the ECM and integrins collaborate to regulate gene expression associated with cell growth, differentiation, and survival. Therefore, the interactions between ECM and integrins exert profound influences on these major cellular programs by altering, through a number of signal transduction pathways, the transcription of genes. The binding of ECM components, such as FN, with their corresponding integrin receptors is known to trigger the activation of different intracellular signaling mediators. For example, the activation of the MAPK signal transduction pathway is of particular interest because it links integrin-mediated signaling to transcriptional regulation of genes, which are crucial for cell growth and differentiation through the activation of distinct transcription factors such as ELK, c-Jun, c-Myc, ETS, and Sp1. These factors play pivotal roles in controlling the transcription of genes required for the cellular programs mentioned earlier.

Many integrins have been identified in the corneal epithelium including alpha subunits $\alpha 2$, $\alpha 3$, $\alpha 4$, $\alpha 5$, $\alpha 6$, αv , and $\alpha 9$, and beta subunits $\beta 1$, $\beta 4$, and $\beta 5$ (for a review, see Elnor and Elnor, 1996).^[33,34] The increase of FN and LM expression in corneal wound healing was postulated to be coordinated with the expression of their corresponding integrin receptors, $\alpha 4\beta 1$ and $\alpha 5\beta 1$ (recognized as the major FN receptors), as well as $\alpha 6\beta 4$ and $\alpha 6\beta 1$ (LM receptors) (for a review, see Refs. [35,36]). For instance, the integrin $\alpha 5\beta 1$ was shown to be present during corneal wound healing after radial keratectomy and $\alpha 6$ was detectable around the entire cell up to the leading edge of the migrating epithelium. The use of a debridement wound model in the mouse cornea clearly provided evidence that the expression of both $\alpha 6$ and $\alpha 9$, as well as that of $\beta 4$, is closely associated with wound healing, resulting in a significant increase in the level of expression of these integrins at both the mRNA and protein levels (Ref. [37] and the cited references therein). The appearance and disappearance of $\alpha 5$ and $\alpha 6$ integrin subunits are thus well coordinated with FN and LM in the cornea.

Extracellular Matrix Components: Expression in Normal and Healing Cornea

In the unwounded rat eye, FN was observed by immunohistological staining at the level of the corneal

epithelium basement membrane whereas LM is found in the subepithelial region, the basal aspect of the basal epithelial cells and the Descemet's membrane.^[35,36] Shortly after corneal injury, the basal cells that border the injured area and stromal keratocytes start producing massive amounts of FN and its subepithelial immunohistological staining progressively decreases once the wound is reepithelialized. In corneal keratectomy wounds, the presence of LM was demonstrated along the wound bed. In addition, it was shown that scarred regions of keratoconus corneas had greater than normal staining of the epithelial basement membrane for FN and LM.^[38,39] Moreover, FN and LM are expressed at levels greatly above normal in the healing corneal stroma of alkali-burned corneas.^[40]

The expression of FN and LM was found to increase in tear fluid as well as in corneas after PRK or LASIK treatments.^[41-43] With increased postoperative time, FN levels diminished and the change in expression of $\beta 4$ became linear. It is therefore clear that FN and LM, as well as their respective receptors, are very important components of a proper regeneration and adhesion of corneal epithelium.

Integrin and Extracellular Matrix Expression in Reconstructed Human Cornea

Integrins and basement membrane components were visualized by immunohistofluorescence labeling of human cornea reconstructed by the collagen gel model because these constituents are necessary for a good tissue cohesion. The expression of the $\beta 1$, $\alpha 3$, $\alpha 5$, and $\alpha 6$ integrin subunits are detected in the epithelium, but the expression of the $\alpha 4$ integrin subunit is not.^[6,11]

The basement membrane components, laminin, type VII collagen, and fibronectin are already detectable in the reconstructed cornea after 3 days of culture. A discontinuous pattern of laminin is visualized at the basement membrane. Laminin, as well as type VII collagen labeling, is present in the cytoplasm of the epithelial basal cells, indicating an active production of ECM proteins. Fibronectin is also present at the epithelium-stromal junction in addition to its detection within the stroma.^[11]

FUTURE CHALLENGES AND APPLICATIONS

Future Challenges

The biggest challenge in the tissue engineering of reconstructed human cornea is to combine the improvement of the models while conserving transparency. All components of the cornea—the epithelium, the stroma, and Bowman's and Descemet's membranes, as well as the endothelium—are organized to allow light to pass

through. The corneal epithelium with its relative thinness, the absence of cornified cells, and the smooth and linear junction that contributes to its transparency has been reproduced in vitro by tissue engineering with human^[11,24] and animal^[19,20,44] cells. The absence of blood vessels in the cornea facilitates its reconstruction in vitro.

The stroma represents a greater challenge because its transparency is due to the particular organization of its collagen fibers. The keratocytes oriented parallel to the collagen fibers form lamellae, which are superposed with others in a regular order. The collagen in each lamella is perpendicular to one of the adjacent lamellae. The uniform diameter and regular spacing of the fibers, in which proteoglycans play a role, also contribute to the transparency of the stroma. Further characterization of the ultrastructural features and the extracellular matrix composition of the reconstructed stroma will be necessary. The reconstructed cornea could be used as a model to evaluate the extent to which mechanical stimulation and matrix constituents (proteoglycans) are necessary for a particular stromal organization.

The Bowman's membrane is composed of randomly oriented delicate collagen fibers and is therefore less complex. The Descemet's membrane is a basal lamina elaborated by the underlying corneal endothelial cells that should reappear in vitro after the addition of endothelial cells to the reconstructed tissue.

The state of hydration of the cornea is another factor of utmost importance for its transparency. The semipermeable endothelium plays an active role in keeping the cornea underhydrated. Therefore, the addition of corneal endothelial cells to the reconstructed cornea is an important optimization step that could be accomplished because the culture of human endothelial corneal cells is possible^[45-47] and the reconstruction of corneal tissue with animal endothelial cells^[19,20,44] has been reported.

The level of optimization necessary may depend on the application foreseen for the human reconstructed cornea. The complete model with all three layers might not be required for in vitro applications such as the evaluation of the effect of mechanical forces on the extracellular matrix organization. Since the maturation and differentiation of reconstructed tissues continue after transplantation, immature tissues could be grafted for some clinical applications.

Experimental Applications

An understanding of the cellular and molecular mechanisms underlying the adhesion, migration and differentiation properties of the various cell types encountered in the cornea will lead to predictable, safe refractive surgical and visual results. The availability of

a human model will be much closer to the real in vivo tissue and would clearly represent a more appropriate alternative to the use of whole human tissues. No such human model could be exploited for this type of study until we recently succeeded in tissue engineering human corneal equivalents, which show appropriate histology and expression of basement membrane components and integrins.^[11] Corneal endothelial cells of either human or bovine origin have recently been added to this corneal equivalent.^[23] Not only are such tissue-engineered corneal substitutes valuable tools for physiological, toxicological, and pharmacological studies, but they also represent outstanding models for the study of stem cells or gene expression. This approach may be helpful to study the factors that modulate integrin expression at the promoter level such as growth factors and cytokines, as well as ECM components.

The development of a reconstructed cornea is also of great interest in industrial applications where alternatives for the Draize test are sought. This procedure evaluates the toxicity of drugs and household products by exposing the cornea of rabbits to various concentrations of the test chemical or product to determine the concentration resulting in irritation, toxicity, or alteration in transparency. The use of corneal equivalents for these tests would contribute to reducing the large number of animals required. In addition, this approach is not biased by potential intrinsic differences in the reaction of the animal compared to human. Therefore, there is a need to find alternative new methods that will result in a reduction in the use of laboratory animals. In June 2000, a new law (*directive européenne* 97/18/CEE) banished in France the use of animals in the testing of cosmetics. Promising results have been obtained with the collagen sponge model including transformed human epithelial cells, keratocytes, and endothelial cells. The chemical-induced change in transparency of reconstructed corneas is indeed a criterion of the original Draize test.^[24]

Clinical Applications

A reconstructed human cornea could be very valuable in two different clinical fields. First, a tissue-engineered reconstructed anterior cornea could be grafted for patients suffering from corneal epithelial stem cell aplasia. The loss of stem cells in the limbal epithelium may lead to functional blindness as it impedes with cell renewal leading to neovascularization, persistent epithelial defects, scarring, ulceration, and corneal perforation. The only possible cure is transplanting limbal stem cells from autologous or allogeneic tissues.^[48,49] In the short term, tissue engineering will be most advantageous when performed for lamellar keratoplasty autologous transplantation, as only a small limbal

tissue biopsy is needed because the epithelial culture can be expanded from a 1 mm² biopsy sample and transplanted on the patient.^[50] As a long-term goal for reconstructed whole cornea production, the supplementary autologous biopsy that may be required to isolate keratocytes and endothelial cells remains a challenge both technically and in terms of controlled cell differentiation.

Second, the tissue-engineered cornea might be used as a specialized living wound dressing in applications related to wound healing. The cultured tissue would not be grafted, but instead temporarily placed over the wounds to improve healing. Under these circumstances, healing is initiated by the patient's own stem cells remaining in the limbal cornea whereas the tissue-engineered cornea protects the wound and secretes factors that promote healing. Such a clinical application has proven useful for cutaneous wounds. Indeed, tissue-engineered epidermis, or skin, has been reported to accelerate the healing of second-degree burns or cutaneous ulcers.

Thus, the development of reconstructed cornea by tissue engineering could bring an additional and valuable therapeutic approach to the present clinical arsenal for the treatment of various corneal diseases.

ACKNOWLEDGMENTS

The authors wish to acknowledge Drs. Richard Bazin and Alain Rousseau (Département d'Ophtalmologie, Université Laval), and Dr. Hélène Boisjoly (Département d'Ophtalmologie, Université de Montréal) for critically reviewing this manuscript. We also thank Marcelle Giasson, Rina Guignard, and Éric Grandbois for technical assistance and Claude Marin for photographic assistance. This study was supported by a grant from the Fonds de la Recherche en Santé du Québec (F.R.S.Q.) and the Réseau de Recherche en Vision du F.R.S.Q. L.G. is holder of the Canadian Research Chair in Stem Cells and Tissue Engineering from the Canadian Institutes for Health Research. C.S. and S.L.G. received scholarships from the F.R.S.Q. P.C. is the recipient of a studentship from the F.R.S.Q.

ARTICLES OF FURTHER INTEREST

Collagen; Corneal Implants; Fibrin; Fibrin Sealants; Ocular Implants; Stem Cells

REFERENCES

1. Kruse, F.E. Stem cells and corneal epithelial regeneration. *Eye* **1994**, *8*, 170–183.
2. Thoft, R.A.; Friend, J. The X, Y, Z hypothesis of corneal epithelial maintenance. *Invest. Ophthalmol. Vis. Sci.* **1983**, *24* (10), 1442–1443.
3. Schermer, A.; Galvin, S.; Sun, T.-T. Differentiation-related expression of a major 64 K corneal keratin in vivo and in culture suggests limbal location of corneal epithelial stem cells. *J. Cell Biol.* **1986**, *103* (1), 49–62.
4. Lauweryns, B.; Van Den Oord, J.J.; Missotten, L. The transitional zone between limbus and peripheral cornea. An immunohistochemical study. *Invest. Ophthalmol. Vis. Sci.* **1993**, *34* (6), 1991–1999.
5. Kasper, M.; Stosiek, P.; Lane, B. Cytokeratin and vimentin heterogeneity in human cornea. *Acta Histochem.* **1992**, *93* (2), 371–381.
6. Germain, L.; Carrier, P.; Auger, F.A.; Salesse, C.; Guérin, S.L. Can we produce a human corneal equivalent by tissue engineering? *Prog. Retin. Eye Res.* **2000**, *19* (5), 497–527.
7. Michel, M.; Torok, N.; Godbout, M.-J.; Lussier, M.; Gaudreau, P.; Royal, A.; Germain, L. Keratin 19 as a biochemical marker of skin stem cells in vivo and in vitro: Keratin 19 expressing cells are differentially localized in function of anatomic sites, and their number varies with donor age and culture stage. *J. Cell. Sci.* **1996**, *109* (Pt. 5), 1017–1028.
8. Fradette, J.; Germain, L.; Seshaiiah, P.; Coulombe, P.A. The type I keratin 19 possesses distinct and context-dependent assembly properties. *J. Biol. Chem.* **1998**, *273* (52), 35176–35184.
9. Rheinwald, J.G.; Green, H. Serial cultivation of strains of human epidermal keratinocytes: The formation of keratinizing colonies from single cells. *Cell* **1975**, *6*, 331–343.
10. Lindberg, K.; Brown, M.E.; Chaves, H.V.; Kenyon, K.R.; Rheinwald, J.G. In vitro propagation of human ocular surface epithelial cells for transplantation. *Invest. Ophthalmol. Vis. Sci.* **1993**, *34* (9), 2672–2679.
11. Germain, L.; Auger, F.A.; Grandbois, E.; Guignard, R.; Giasson, M.; Boisjoly, H.M.; Guérin, S.L. Reconstructed human cornea produced in vitro by tissue engineering. *Pathobiology* **1999**, *67* (3), 140–147.
12. Gaudreault, M.; Carrier, P.; Larouche, K.; Leclerc, S.; Giasson, M.; Germain, L.; Guerin, S.L. Influence of sp1/sp3 expression on corneal epithelial cells proliferation and differentiation properties in reconstructed tissues. *Invest. Ophthalmol. Vis. Sci.* **2003**, *44* (4), 1447–1457.
13. Pellegrini, G.; Golisano, O.; Paterna, P.; Lambiase, A.; Bonini, S.; Rama, P.; De Luca, M. Location and clonal analysis of stem cells and their differentiated progeny in the human ocular surface. *J. Cell Biol.* **1999**, *145* (4), 769–782.
14. Joyce, N.C.; Navon, S.E.; Roy, S.; Zieske, J.D. Expression of cell cycle-associated proteins in human and rabbit corneal endothelium in situ. *Invest. Ophthalmol. Vis. Sci.* **1996**, *37* (8), 1566–1575.
15. Engelmann, K.; Bednarz, J.; Bohnke, M. Endothelial cell transplantation and growth behavior of the human corneal endothelium. *Ophthalmologie* **1999**, *96* (9), 555–562.

16. Rama, P.; Bonini, S.; Lambiase, A.; Golisano, O.; Paterna, P.; De Luca, M.; Pellegrini, G. Autologous fibrin-cultured limbal stem cells permanently restore the corneal surface of patients with total limbal stem cell deficiency. *Transplantation* **2001**, *72* (9), 1478–1485.
17. Auger, F.A.; Lopez-Valle, C.A.; Guignard, R.; Tremblay, N.; Noël, B.; Goulet, F.; Germain, L. Skin equivalent produced with human collagens. *In vitro Cell. Dev. Biol., Anim.* **1995**, *31*, 432–439.
18. Bell, E.; Ehrlich, H.P.; Buttle, D.J.; Nakatsuji, T. Living tissue formed in vitro and accepted as skin-equivalent tissue of full thickness. *Science* **1981**, *211* (4486), 10452–11054.
19. Zieske, J.D.; Mason, V.S.; Wasson, M.E.; Meunier, S.F.; Nolte, C.J.M.; Fukai, N.; Olsen, B.R.; Parenteau, N.L. Basement membrane assembly and differentiation of cultured corneal cells: Importance of culture environment and endothelial cell interaction. *Exp. Cell Res.* **1994**, *214*, 621–633.
20. Minami, Y.; Sugihara, H.; Oono, S. Reconstruction of cornea in three-dimensional collagen gel matrix culture. *Invest. Ophthalmol. Vis. Sci.* **1993**, *34* (7), 2316–2324.
21. Michel, M.; L'Heureux, N.; Pouliot, R.; Xu, W.; Auger, F.A.; Germain, L. Characterization of a new tissue-engineered human skin equivalent with hair. *In vitro Cell. Dev. Biol., Anim.* **1999**, *35* (6), 318–326.
22. Pruniéras, M.; Régnier, M.; Woodley, D. Methods for cultivation of keratinocytes with an air–liquid interface. *J. Invest. Dermatol.* **1983**, *81* (1), 28s–33s.
23. Giasson, C.J.; Carrier, P.; Germain, L. Collagen-based corneal equivalents made with human cells as a model to study cell–cell contact. *Invest. Ophthalmol. Vis. Sci.* **2003**, *44*. E-Abstract 927.
24. Griffith, M.; Osborne, R.; Munger, R.; Xiong, X.; Doillon, C.J.; Laycock, N.L.; Hakim, M.; Song, Y.; Watsky, M.A. Functional human corneal equivalents constructed from cell lines. *Science* **1999**, *286* (5447), 2169–2172.
25. Griffith, M.; Hakim, M.; Shimmura, S.; Watsky, M.A.; Li, F.; Carlsson, D.; Doillon, C.J.; Nakamura, M.; Suuronen, E.; Shinozaki, N.; Nakata, K.; Sheardown, H. Artificial human corneas: Scaffolds for transplantation and host regeneration. *Cornea* **2002**, *21* (Suppl. 7), S54–S61.
26. L'Heureux, N.; Pâquet, S.; Labbé, R.; Germain, L.; Auger, F.A. A completely biological tissue-engineered human blood vessel. *FASEB J.* **1998**, *12*, 47–56.
27. Rochon, M.-H.; Roberge, C.J.; Auger, F.A.; Germain, L. Normal Epithelial Cells Modulate Capillary Size by Secreting VEGF in a Human In Vitro Angiogenesis Model Produced by the Self-Assembly Method. *J. Invest. Dermatol.* *121* (1): p. 1, abst 0002. In *International Investigative Dermatology*; April 30–May 4 2003; Miami Beach, Florida.
28. Germain, L.; Carrier, P.; Giasson, M.; Grandbois, E.; Auger, F.A.; Guerin, S.L. A new method for in vitro production of human reconstructed cornea by tissue engineering. *Invest. Ophthalmol. Vis. Sci.* **1999**, *40*, S329.
29. Carrier, P.; Deschambeault, A.; Talbot, M.; Guerin, S.L.; Auger, F.A.; Germain, L. Expression of extracellular matrix components in reconstructed human cornea produced by tissue engineering. *Invest. Ophthalmol. Vis. Sci.* **2002**, *43*. E-Abstract 4166.
30. Reim, M.; Kottek, A.; Schrage, N. The cornea surface and wound healing. *Prog. Retin. Eye Res.* **1997**, *16* (2), 183–225.
31. Smith, R.J.; Maloney, R.K. Diffuse lamellar keratitis. A new syndrome in lamellar refractive surgery. *Ophthalmology* **1998**, *105* (9), 1721–1726.
32. Plow, E.F.; Haas, T.A.; Zhang, L.; Loftus, J.; Smith, J.W. Ligand binding to integrins. *J. Biol. Chem.* **2000**, *275* (29), 21785–21788.
33. Giancotti, F.G. Complexity and specificity of integrin signalling. *Nat. Cell Biol.* **2000**, *2* (1), E13–E14.
34. Elner, S.G.; Elner, V.M. The integrins superfamily and the eye. *Invest. Ophthalmol. Vis. Sci.* **1996**, *37* (5), 696–701.
35. Zieske, J.D. Extracellular matrix and wound healing. *Curr. Opin. Ophthalmol.* **2001**, *12* (4), 237–241.
36. Ahmadi, A.J.; Jakobiec, F.A. Corneal wound healing: Cytokines and extracellular matrix proteins. *Int. Ophthalmol. Clin.* **2002**, *42* (3), 13–22.
37. Stepp, M.A.; Zhu, L. Upregulation of $\alpha 9$ integrin and tenascin during epithelial regeneration after debridement in the cornea. *J. Histochem. Cytochem.* **1997**, *45* (2), 189–201.
38. Kenney, M.C.; Nesburn, A.B.; Burgeson, R.E.; Butkowski, R.J.; Ljubimov, A.V. Abnormalities of the extracellular matrix in keratoconus corneas. *Cornea* **1997**, *16* (3), 345–351.
39. Tuori, A.; Virtanen, I.; Aine, E.; Uusitalo, H. The expression of tenascin and fibronectin in keratoconus, scarred and normal human cornea. *Graefes Arch. Clin. Exp. Ophthalmol.* **1997**, *235* (4), 222–229.
40. Saika, S.; Kobata, S.; Hashizume, N.; Okada, Y.; Yamanaka, O. Epithelial basement membrane in alkali-burned corneas in rats. *Immunohistochemical study. Cornea* **1993**, *12* (5), 383–390.
41. Latvala, T.; Barraquer-Coll, C.; Tervo, K.; Tervo, T. Corneal wound healing and nerve morphology after excimer laser in situ keratomileusis in human eyes. *J. Refract. Surg.* **1996**, *12* (6), 677–683.
42. Anderson, J.A.; Binder, P.S.; Rock, M.E.; Vrabcic, M.P. Human excimer laser keratectomy. Immunohistochemical analysis of healing. *Arch. Ophthalmol.* **1996**, *114* (1), 54–60.
43. Winter, M.; Behrendt, S.; Binder, P.S.; Duncker, G.I. Ultrastructural and immunohistochemical findings after linear excimer laser keratectomy. *J. Refract. Surg.* **1997**, *13* (1), 60–68.
44. Schneider, A.I.; Maier-Reif, K.; Graeve, T. Constructing an in vitro cornea from cultures of the three specific corneal cell types. *In vitro Cell. Dev. Biol., Anim.* **1999**, *35* (9), 515–526.
45. Chen, K.H.; Azar, D.; Joyce, N.C. Transplantation of adult human corneal endothelium ex vivo: A morphologic study. *Cornea* **2001**, *20* (7), 731–737.
46. Blake, D.A.; Yu, H.; Young, D.L.; Caldwell, D.R. Matrix stimulates the proliferation of human corneal

- endothelial cells in culture. *Invest. Ophthalmol. Vis. Sci.* **1997**, *38*, 1119–1129.
47. Engelmann, K.; Firedl, P. Optimization of culture conditions for human corneal endothelial cells. *In vitro Cell. Dev. Biol., Anim.* **1989**, *25* (11), 1065–1072.
48. Pfister, R.R. Corneal stem cell disease: Concepts, categorization, and treatment by auto- and homo-transplantation of limbal stem cells. *CLAO J.* **1994**, *20* (1), 64–72.
49. Holland, E.J.; Schwartz, G.S. The evolution of epithelial transplantation for severe ocular surface disease and a proposed classification system. *Cornea* **1996**, *15* (6), 549–556.
50. Pellegrini, G.; Traverso, C.E.; Franzi, A.T.; Zingirian, M.; Cancedda, R.; De Luca, M. Long-term restoration of damaged corneal surfaces with autologous cultivated corneal epithelium. *Lancet* **1997**, *349* (9057), 990–993.

Tissue Engineering of the Esophagus

Stephen F. Badylak
Thomas W. Gilbert
Alejandro Nieponice

McGowan Institute for Regenerative Medicine, University of Pittsburgh, Pittsburgh,
Pennsylvania, U.S.A.

T

INTRODUCTION

The default mechanism of wound repair in the esophagus of adult mammals is scar tissue formation with stricture. Tissue engineering and regenerative medicine strategies for reconstitution of functional esophageal tissue have been largely unexplored or not reported. The clinical problems related to the esophagus, animal models of esophageal pathology, and an overview of regenerative medicine and tissue engineering principles as they relate to replacement and reinforcement of injured or missing esophageal tissue will be discussed, with an emphasis upon scaffold and cell-based approaches.

CLINICAL RELEVANCE

Congenital pathology such as esophageal atresia or acquired pathology such as esophageal cancer are currently treated by surgical resection,^[1,2] with morbidity rates exceeding 50%.^[3] Translocation of the stomach, large bowel, small bowel, and skin have been utilized in attempts to replace the resected esophageal tissue.^[4,5] The results of past and current surgical therapy have been less than ideal with high morbidity and mortality.^[6,7]

The current standard of care for esophageal cancer or Barrett's esophagus with high-grade dysplasia is esophagectomy with a "gastric pull-up," a procedure in which the stomach is shaped into a tube and mobilized through the mediastinum and anastomosed to the remaining cervical esophagus (Fig. 1).^[4] During this procedure, the blood supply to the stomach is compromised, and the resulting poorly vascularized anastomotic site has a reduced healing capacity that contributes to complications such as mediastinal leaks or strictures that hinder the normal postoperative recovery.^[7,8] Following esophagectomy with gastric pull-up, patients suffer from delayed swallowing or chest discomfort because of gastric distention within the mediastinal cavity.^[9,10] Anastomotic leakage because of dehiscence of the tissues is one of the most common complications associated with esophageal surgery that contributes to postoperative morbidity.^[11,12] Technical-related complications

such as anastomotic site leakage are considered an independent risk factor for the prognostic outcome.^[13] Another common complication related to suboptimal healing is stricture at the anastomotic site. Almost 50% of patients require one or more postoperative endoscopic dilations to treat stricture.^[12]

Non-malignant pathologies of the esophagus such as traumatic injury, fistulae, or small congenital defects would benefit greatly from effective tissue engineering approaches that could close defects without stricture and/or prevent resection of esophageal tissues with the need for an anastomosis.^[14] The mechanisms of esophageal scarring and stricture formation are largely unknown, although they are likely related to the presence of an acidic environment, proteolytic enzymes, and molecular mediators of inflammation that lead to fibroblast proliferation and contracture.^[15-19]

Esophageal stricture can be triggered by partial mucosal resection, which is a therapeutic intervention for the treatment of Barrett's disease, a preneoplastic condition that can increase the risk of developing esophageal adenocarcinoma by 50-fold.^[20,21] In the past decade, two minimally invasive approaches have yielded promising results for the treatment of Barrett's esophagus. The first involves endomucosal resection (EMR) of the affected area to remove the affected epithelium. The main limitation of this procedure is stricture formation, which, in turn, is dependent on the extent of the resection. The risk of stricture increases significantly when the segment is longer than 30 mm or when the resection involves more than three-quarters of the inner circumference of the esophagus.^[22-24] The second minimally invasive approach for treatment of Barrett's is ablation of the pathological tissue by photodynamic therapy (PDT). PDT is not as strongly associated with stricture as EMR, but the risk of perforation increases with the depth of the ablation. Complete removal of the esophageal mucosa by PDT requires several endoscopic sessions to be effective without complications,^[25] while superficial ablation of the esophageal mucosa increases the risk of leaving metaplastic glands beneath the neoe epithelium, thereby increasing the risk of recurrent disease or progression to neoplasia.^[26,27]

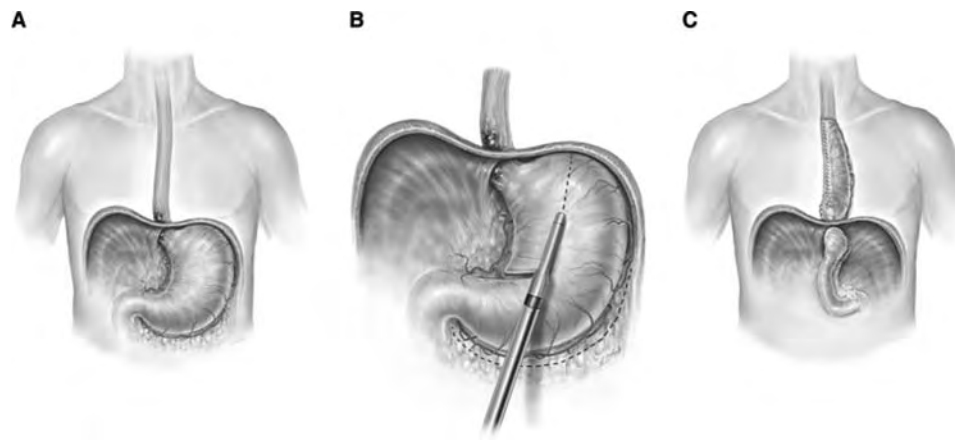


Fig. 1 Schematic of the gastric “pull-up” procedure. (A) Normal anatomy of the upper gastrointestinal tract with high-grade dysplasia in the distal esophagus. (B) The esophagus and part of the gastric fundus are surgically excised to remove the tumor. The remaining stomach is shaped into a tubular form. (C) The tubularized stomach is passed through the mediastinum and anastomosed to the remaining proximal esophagus.

These negative outcomes could potentially be mitigated by employing a tissue engineering/regenerative medicine approach for reinforcement or replacement of functional esophageal tissue.

ANIMAL MODELS

Although interactions between various cell populations and scaffold materials can be studied *in vitro*, preclinical *in vivo* studies are essential to study esophageal regeneration. Complex interactions such as cell–cell and cell–matrix interactions in the presence of inflammation, fibroblastic proliferation, and scarring that characterize the esophageal response to injury in mammals cannot be replicated in *in vitro* studies. Animal models have been used extensively in the field of regenerative medicine of the digestive tract.^[28] One of the earliest reports involved the repair of a duodenal wound with jejunal serosa using mongrel dogs.^[29] Subsequently, reconstructive strategies have been investigated for almost the entire length of the digestive tract, including the esophagus, stomach, small bowel, and large bowel.^[28,30–34] Two of the most commonly used animal models for esophageal studies are the mongrel dog and the opossum. The canine model of the esophagus is appropriate because the canine esophagus is similar to that of humans, dogs are a convenient size for surgical models, they are readily available for research purposes, and their slow growth rate and typical life span are suitable for long-term follow-up studies.^[35] Limitations of the canine model include the relatively high cost compared to rodent studies.

The other model employed in esophageal studies is the opossum. Several studies have been published, showing the similarity of the opossum esophageal

histology to that of the human esophagus, especially with regard to blood supply.^[8,36,37] Although, there are no studies using the opossum in the tissue engineering/regenerative medicine literature, it has been used as a valid model to understand the complications of gastrointestinal tract reconstruction after radical esophagectomy. Limitations with this species include size of the animal and its esophagus, availability of the animal model, and behavior of the species.

The choice of an appropriate animal model in any experimental design is very important. Factors to be considered include resemblance of the animal model to human anatomy and function, and the response to injury both in normal and pathological conditions. Misinterpretation of excellent healing capacity by certain species can lead to erroneous human clinical translation.

ESOPHAGEAL TISSUE ENGINEERING STRATEGIES

Synthetic Scaffolds

A variety of scaffold materials have been considered for tissue engineering of esophageal tissue, including synthetic and naturally occurring scaffolds, with and without the addition of cells. Recently, there have been a number of studies evaluating viability and differentiation of esophageal cells on various synthetic scaffold materials. Poly-L-lactic acid (PLLA), poly-lactic-glycolic acid (PLGA), and polycaprolactone/PLLA (PCL/PLLA), alone or surface modified with collagen and fibronectin, have been shown to support growth of esophageal cells including epithelial cells, smooth muscle cells, and fibroblasts.^[38–40] Rat esophageal organoid units isolated by digestion of normal esophageal tissue

by dispase and collagenase have been seeded onto a PLGA scaffold for implantation within the omentum of an allogeneic recipient.^[31] After four weeks, the organoid graft was removed from the omentum and used to replace a segment of the esophagus in the same animal. The result was a functional repair of the esophagus that was histologically similar to the normal esophagus, although it was significantly dilated. Staining of the cells in the organoid units with green fluorescent protein prior to implantation allowed for verification that these cells were responsible for reconstituting the esophageal tissue.

Collagen Type I Scaffolds

A large body of literature describes the replacement of the esophagus with a composite collagen matrix and silicone stent.^[41–46] This composite approach has shown moderate success when the stent is left in place for a minimum of three weeks after surgery. Prior to three weeks, the mortality rate was nearly 100% and was associated with stricture and only partial epithelialization.^[43] With longer stenting of four to six weeks, survival has been observed for as long as 26 weeks after surgery, although some stenosis and contracture were still present.^[44] The remodeled tissue consisted of longitudinally oriented collagenous tissue without smooth muscle or glandular structures. The degree of success when stenting for three weeks compared to four weeks or greater was associated with the extent of reepithelialization.^[47]

Type I collagen scaffolds have also been studied with cell seeding prior to implantation. A sheet-like construct of esophageal epithelial cells, aortic smooth muscle cells, and dermal fibroblasts seeded on Type I collagen was implanted onto the latissimus dorsi of an athymic rat.^[48] The resulting remodeled graft had a structure similar to that of normal esophageal tissue. However, the tissue was not transplanted to the esophageal location so functionality was not determined. In summary, Type I collagen scaffolds show some promise for replacement of esophageal tissue. The work with collagen-based scaffolds also emphasizes the importance of complete reepithelialization to prevent stricture formation.

Cell Sheets

A cell-based approach for tissue reconstruction has been investigated for a number of tissues, including the corneal epithelium, urothelium, and periodontal ligament.^[49–52] Cell sheets have been formed by growing cells on specially designed culture dishes, which allow an intact, confluent monolayer to be detached from the surface of the dish. As no enzymatic treatments are used, the resulting cell sheet retains the extracellular

matrix (ECM) components produced by the cells and is robust enough to be handled and delivered to the site of implantation. Although such cell sheets have shown promise for reepithelialization of the esophagus after EMR, controlled studies are forthcoming. (Yamato, M. Personal communication, 2005.)

Naturally Occurring Scaffolds

Biologic scaffolds, such as ECM, that are derived from various tissues have shown a great deal of promise for esophageal reconstruction. An *in vitro* study showed that PLLA, PLGA, and PCL/PLLA (all synthetic scaffold materials) were inferior to AlloDerm (LifeCell Corporation, Branchburg, NJ), an ECM scaffold derived from human skin, when evaluated for epithelial organization and stratification.^[38] Naturally derived ECM scaffolds have also been investigated for *in vivo* esophageal tissue engineering. Surgically created full thickness defects in the cervical esophagus measuring 5 cm in length and encompassing 40–50% of the circumference or 100% of the circumference were repaired with small intestinal submucosa or urinary bladder submucosa. Well-organized, site-specific remodeling with esophageal tissue that was fully epithelialized and contiguous with normal esophageal tissue was present by 50 days after surgery in the partial circumferential defects. However, the full circumferential defects healed with stricture within 35 days after surgery.^[53]

In a subsequent study that used an ECM composed of the basement membrane and tunica propria of the porcine urinary bladder [urinary bladder matrix (UBM)], it was shown that critically sized, full circumferential defects could be repaired with minimal stricture formation if adjacent autologous smooth skeletal muscle tissue was placed in direct apposition to the ECM scaffold at the time of surgery.^[30] The maintenance of the basement membrane on UBM provides an ideal substrate for epithelial layer development, and this, combined with the presence of skeletal muscle cells, may have facilitated the epithelialization necessary to prevent or minimize tissue contracture.^[54] Spontaneous motility of the muscle within the remodeled section was shown by esophograms and endoscopy, although the motility did not appear to be synchronous with the native esophagus. The mechanical behavior of the graft was also evaluated, and it was shown that while the biologic scaffold material was much stiffer and stronger than the native tissue at the time of surgery, the remodeled tissue approached the mechanical behavior of the normal esophagus by three months after implantation. In summary, it appears that a regenerative medicine approach that includes both ECM scaffolds and autologous muscle cells holds promise for functional esophageal replacement.

T

CONCLUSIONS

Complete replacement of the esophagus after esophagectomy and reinforcement of the esophagus after anastomosis or EMR are important clinical applications for which tissue engineering/regenerative medicine approaches hold promise. Rapid epithelialization of a graft has been shown to be essential for the development of functional tissue without stricture. It appears that the presence of parenchymal cells associated with a biologic scaffold is important for the repair of the esophagus. However, it is not clear which cell types are most important. It has yet to be determined whether it is better to introduce the cells prior to in vivo transplantation or to utilize an autologous cell source in situ at the time of surgery. In summary, tissue engineering/regenerative medicine approaches for esophageal reconstruction appear promising, but significant work remains to be done before human clinical application becomes reality.

ARTICLES OF FURTHER INTEREST

Tissue Engineering; Tissue Engineering of the Intestine

REFERENCES

1. Aprodu, S.G.; Gavrilesco, S.; Savu, B.; Munteanu, V.; Vlad, A. Indications for esophageal replacement in children. *Rev. Med. Chir. Soc. Med. Nat. Iasi* **2004**, *108* (4), 805–808.
2. Orringer, M.B.; Marshall, B.; Iannettoni, M.D. Transhiatal esophagectomy: clinical experience and refinements. *Ann. Surg.* **1999**, *230* (3), 392–400; discussion 400–403.
3. Sauvanet, A.; Mariette, C.; Thomas, P.; Lozac'h, P.; Segol, P.; Tiret, E.; Delperio, J.R.; Collet, D.; Leborgne, J.; Pradere, B.; Bourgeon, A.; Triboulet, J.P. Mortality and morbidity after resection for adenocarcinoma of the gastroesophageal junction: predictive factors. *J. Am. Coll. Surg.* **2005**, *201* (2), 253–262.
4. Orringer, M.B.; Stirling, M.C. Cervical esophagogastric anastomosis for benign disease. Functional results. *J. Thorac. Cardiovasc. Surg.* **1988**, *96* (6), 887–893.
5. Yildirim, S.; Koksall, H.; Celayir, F.; Erdem, L.; Oner, M.; Baykan, A. Colonic interposition vs. gastric pull-up after total esophagectomy. *J. Gastrointest. Surg.* **2004**, *8* (6), 675–678.
6. Dimick, J.B.; Wainess, R.M.; Upchurch, G.R., Jr.; Iannettoni, M.D.; Orringer, M.B. National trends in outcomes for esophageal resection. *Ann. Thorac. Surg.* **2005**, *79* (1), 212–216; discussion 217–218.
7. Iannettoni, M.D.; Whyte, R.I.; Orringer, M.B. Catastrophic complications of the cervical esophagogastric anastomosis. *J. Thorac. Cardiovasc. Surg.* **1995**, *110* (5), 1493–1500; discussion 1500–1501.
8. Reavis, K.M.; Chang, E.Y.; Hunter, J.G.; Jobe, B.A. Utilization of the delay phenomenon improves blood flow and reduces collagen deposition in esophagogastric anastomoses. *Ann. Surg.* **2005**, *241* (5), 736–745; discussion 745–747.
9. Chattopadhyay, T.K.; Shad, S.K.; Kumar, A. Intra-gastric bile acid and symptoms in patients with an intrathoracic stomach after oesophagectomy. *Br. J. Surg.* **1993**, *80* (3), 371–373.
10. Koh, P.; Turnbull, G.; Attia, E.; LeBrun, P.; Casson, A.G. Functional assessment of the cervical esophagus after gastric transposition and cervical esophagogastric anastomosis. *Eur. J. Cardiothorac. Surg.* **2004**, *25* (4), 480–485.
11. Bhasin, D.K.; Sharma, B.C.; Gupta, N.M.; Sinha, S.K.; Singh, K. Endoscopic dilation for treatment of anastomotic leaks following transhiatal esophagectomy. *Endoscopy* **2000**, *32* (6), 469–471.
12. Orringer, M.B.; Marshall, B.; Iannettoni, M.D. Eliminating the cervical esophagogastric anastomotic leak with a side-to-side stapled anastomosis. *J. Thorac. Cardiovasc. Surg.* **2000**, *119* (2), 277–288.
13. Rizk, N.P.; Bach, P.B.; Schrag, D.; Bains, M.S.; Turnbull, A.D.; Karpeh, M.; Brennan, M.F.; Rusch, V.W. The impact of complications on outcomes after resection for esophageal and gastroesophageal junction carcinoma. *J. Am. Coll. Surg.* **2004**, *198* (1), 42–50.
14. Inci, I.; Ozcelik, C.; Nizam, O.; Balci, A.E.; Eren, N.; Ozgen, G. Traumatic oesophageal perforation. *Scand. Cardiovasc. J.* **1997**, *31* (2), 97–100.
15. Dunne, D.P.; Paterson, W.G. Acid-induced esophageal shortening in humans: a cause of hiatus hernia? *Can. J. Gastroenterol.* **2000**, *14* (10), 847–850.
16. Huang, J.Q.; Hunt, R.H. pH, healing rate, and symptom relief in patients with GERD. *Yale J. Biol. Med.* **1999**, *72* (2–3), 181–194.
17. Paterson, W.G. Role of mast cell-derived mediators in acid-induced shortening of the esophagus. *Am. J. Physiol.* **1998**, *274* (2 Pt 1), G385–G388.
18. Sampliner, R.E. Reduction of acid exposure and regression of Barrett's esophagus. *Dig. Dis.* **2000**, *18* (4), 203–207.
19. Zhang, X.; Geboes, K.; Depoortere, I.; Tack, J.; Janssens, J.; Sifrim, D. Effect of repeated cycles of acute esophagitis and healing on esophageal peristalsis, tone, and length. *Am. J. Physiol. Gastrointest. Liver Physiol.* **2005**, *288* (6), G1339–G1346.
20. Sampliner, R.E. Prevention of adenocarcinoma by reversing Barrett's esophagus with mucosal ablation. *World J. Surg.* **2003**, *27* (9), 1026–1029.
21. Stein, H.J.; Feith, M.; Siewert, J.R. Malignant degeneration of Barrett's esophagus: clinical point of view. *Recent Results Cancer Res.* **2000**, *155*, 42–53.
22. Giovannini, M.; Bernardini, D.; Moutardier, V.; Monges, G.; Houvenaeghel, G.; Seitz, J.F.; Delperio, J.R. Endoscopic mucosal resection (EMR): results and prognostic factors in 21 patients. *Endoscopy* **1999**, *31* (9), 698–701.
23. Giovannini, M.; Bories, E.; Pesenti, C.; Moutardier, V.; Monges, G.; Danisi, C.; Lelong, B.; Delperio, J.R. Circumferential endoscopic mucosal resection in

- Barrett's esophagus with high-grade intraepithelial neoplasia or mucosal cancer. Preliminary results in 21 patients. *Endoscopy* **2004**, *36* (9), 782–787.
24. Katada, C.; Muto, M.; Manabe, T.; Boku, N.; Ohtsu, A.; Yoshida, S. Esophageal stenosis after endoscopic mucosal resection of superficial esophageal lesions. *Gastrointest. Endosc.* **2003**, *57* (2), 165–169.
 25. Fitzgerald, R.C. Ablative mucosectomy is the procedure of choice to prevent Barrett's cancer. *Gut* **2003**, *52* (1), 16–17.
 26. Luman, W.; Lessels, A.M.; Palmer, K.R. Failure of Nd-YAG photocoagulation therapy as treatment for Barrett's oesophagus—a pilot study. *Eur. J. Gastroenterol. Hepatol.* **1996**, *8* (7), 627–630.
 27. Sharma, P.; Bhattacharyya, A.; Garewal, H.S.; Sampliner, R.E. Durability of new squamous epithelium after endoscopic reversal of Barrett's esophagus. *Gastrointest. Endosc.* **1999**, *50* (2), 159–164.
 28. Chen, M.K.; Beierle, E.A. Animal models for intestinal tissue engineering. *Biomaterials* **2004**, *25* (9), 1675–1681.
 29. Kobold, E.E.; Thal, A.P. A simple method for the management of experimental wounds of the duodenum. *Surg. Gynecol. Obstet.* **1963**, *116*, 340–344.
 30. Badylak, S.F.; Vorp, D.A.; Spievack, A.R.; Simmons-Byrd, A.; Hanke, J.; Freytes, D.O.; Thapa, A.; Gilbert, T.W.; Nieponice, A. Esophageal reconstruction with ECM and muscle tissue in a dog model. *J. Surg. Res.* **2005**, *128*, 87–97.
 31. Grikscheit, T.; Ochoa, E.R.; Srinivasan, A.; Gaijsert, H.; Vacanti, J.P. Tissue-engineered esophagus: experimental substitution by onlay patch or interposition. *J. Thorac. Cardiovasc. Surg.* **2003**, *126* (2), 537–544.
 32. Grikscheit, T.; Srinivasan, A.; Vacanti, J.P. Tissue-engineered stomach: a preliminary report of a versatile in vivo model with therapeutic potential. *J. Pediatr. Surg.* **2003**, *38* (9), 1305–1309.
 33. Grikscheit, T.C.; Ochoa, E.R.; Ramsanahie, A.; Alsberg, E.; Mooney, D.; Whang, E.E.; Vacanti, J.P. Tissue-engineered large intestine resembles native colon with appropriate in vitro physiology and architecture. *Ann. Surg.* **2003**, *238* (1), 35–41.
 34. Kaihara, S.; Kim, S.S.; Benvenuto, M.; Choi, R.; Kim, B.S.; Mooney, D.; Tanaka, K.; Vacanti, J.P. Successful anastomosis between tissue-engineered intestine and native small bowel. *Transplantation* **1999**, *67* (2), 241–245.
 35. Rashid, S.T.; Salacinski, H.J.; Hamilton, G.; Seifalian, A.M. The use of animal models in developing the discipline of cardiovascular tissue engineering: a review. *Biomaterials* **2004**, *25* (9), 1627–1637.
 36. Forse, R.A.; MacDonald, P.H.; Mercer, C.D. Anastomotic and regional blood flow following esophagogastrectomy in an opossum model. *J. Invest. Surg.* **1999**, *12* (1), 45–52.
 37. Zijlstra, F.G.; Hynna-Liepert, T.T.; Dinda, P.K.; Beck, I.T.; Paterson, W.G. Gastrointestinal blood flow in the opossum with special reference to the esophagus. *Can. J. Physiol. Pharmacol.* **1990**, *68* (9), 1221–1225.
 38. Beckstead, B.L.; Pan, S.; Bhrany, A.D.; Bratt-Leal, A.M.; Ratner, B.D.; Giachelli, C.M. Esophageal epithelial cell interaction with synthetic and natural scaffolds for tissue engineering. *Biomaterials* **2005**, *26* (31), 6217–6228.
 39. Zhu, Y.; Chan-Park, M.B.; Chian, K.S. The growth improvement of porcine esophageal smooth muscle cells on collagen grafted poly (dl-lactide-co-glycolide) membrane. *J. Biomed. Mater. Res. B Appl. Biomater.* **2005**, *75B*, 193–199.
 40. Zhu, Y.; Chian, K.S.; Chan-Park, M.B.; Mhaisalkar, P.S.; Ratner, B.D. Protein bonding on biodegradable poly(L-lactide-co-caprolactone) membrane for esophageal tissue engineering. *Biomaterials* **2006**, *27* (1), 68–78.
 41. Takimoto, Y.; Okumura, N.; Nakamura, T.; Natsume, T.; Shimizu, Y. Long-term follow-up of the experimental replacement of the esophagus with a collagen-silicone composite tube. *Asaio J.* **1993**, *39* (3), M736–M739.
 42. Takimoto, Y.; Nakamura, T.; Teramachi, M.; Kiyotani, T.; Shimizu, Y. Replacement of long segments of the esophagus with a collagen-silicone composite tube. *Asaio J.* **1995**, *41* (3), M605–M608.
 43. Takimoto, Y.; Nakamura, T.; Yamamoto, Y.; Kiyotani, T.; Teramachi, M.; Shimizu, Y. The experimental replacement of a cervical esophageal segment with an artificial prosthesis with the use of collagen matrix and a silicone stent. *J. Thorac. Cardiovasc. Surg.* **1998**, *116* (1), 98–106.
 44. Yamamoto, Y.; Nakamura, T.; Shimizu, Y.; Matsumoto, K.; Takimoto, Y.; Kiyotani, T.; Sekine, T.; Ueda, H.; Liu, Y.; Tamura, N. Intrathoracic esophageal replacement in the dog with the use of an artificial esophagus composed of a collagen sponge with a double-layered silicone tube. *J. Thorac. Cardiovasc. Surg.* **1999**, *118* (2), 276–286.
 45. Yamamoto, Y.; Nakamura, T.; Shimizu, Y.; Takimoto, Y.; Matsumoto, K.; Kiyotani, T.; Yu, L.; Ueda, H.; Sekine, T.; Tamura, N. Experimental replacement of the thoracic esophagus with a bioabsorbable collagen sponge scaffold supported by a silicone stent in dogs. *Asaio J.* **1999**, *45* (4), 311–316.
 46. Yamamoto, Y.; Nakamura, T.; Shimizu, Y.; Matsumoto, K.; Takimoto, Y.; Liu, Y.; Ueda, H.; Sekine, T.; Tamura, N. Intrathoracic esophageal replacement with a collagen sponge—silicone double layer tube: evaluation of omental-pedicle wrapping and prolonged placement of an inner stent. *Asaio J.* **2000**, *46* (6), 734–739.
 47. Takimoto, Y.; Teramachi, M.; Okumura, N.; Nakamura, T.; Shimizu, Y. Relationship between stenting time and regeneration of neoesophageal submucosal tissue. *Asaio J.* **1994**, *40* (3), M793–M797.
 48. Hayashi, K.; Ando, N.; Ozawa, S.; Kitagawa, Y.; Miki, H.; Sato, M.; Kitajima, M. A neo-esophagus reconstructed by cultured human esophageal epithelial cells, smooth muscle cells, fibroblasts, and collagen. *Asaio J.* **2004**, *50* (3), 261–266.
 49. Akizuki, T.; Oda, S.; Komaki, M.; Tsuchioka, H.; Kawakatsu, N.; Kikuchi, A.; Yamato, M.; Okano, T.; Ishikawa, I. Application of periodontal ligament cell sheet for periodontal regeneration: a pilot study in beagle dogs. *J. Periodontal Res.* **2005**, *40* (3), 245–251.
 50. Nishida, K.; Yamato, M.; Hayashida, Y.; Watanabe, K.; Yamamoto, K.; Adachi, E.; Nagai, S.; Kikuchi, A.; Maeda, N.; Watanabe, H.; Okano, T.; Tano, Y. Corneal reconstruction with tissue-engineered cell sheets

- composed of autologous oral mucosal epithelium. *N. Engl. J. Med.* **2004**, *351* (12), 1187–1196.
51. Shiroyanagi, Y.; Yamato, M.; Yamazaki, Y.; Toma, H.; Okano, T. Urothelium regeneration using viable cultured urothelial cell sheets grafted on demucosalized gastric flaps. *BJU Int.* **2004**, *93* (7), 1069–1075.
52. Yang, J.; Yamato, M.; Kohno, C.; Nishimoto, A.; Sekine, H.; Fukai, F.; Okano, T. Cell sheet engineering: recreating tissues without biodegradable scaffolds. *Biomaterials* **2005**, *26* (33), 6415–6422.
53. Badylak, S.; Meurling, S.; Chen, M.; Spievack, A.; Simmons-Byrd, A. Resorbable bioscaffold for esophageal repair in a dog model. *J. Pediatr. Surg.* **2000**, *35* (7), 1097–1103.
54. Brown, B.; Lindberg, K.; Reing, J.; Stolz, D.B.; Badylak, S.F. The basement membrane component of biologic scaffolds derived from extracellular matrix. *Tissue Eng.* **2006**, *12* (3), 519–526.

Tissue Engineering, Fat

Barbara Weiser
Markus Neubauer
Achim Göpferich
Torsten Blunk

*Department of Pharmaceutical Technology, University of Regensburg,
Regensburg, Germany*

T

INTRODUCTION

There is a tremendous demand in reconstructive and plastic surgery for suitable grafts to repair soft tissue defects. Conventional inanimate implants represent only suboptimal solutions. Fat is the natural filler tissue for soft tissue defects. Surgeons already use autologous adipose tissue grafts, but current procedures are all afflicted with different disadvantages. Tissue engineering, that is, the *in vitro* and *in vivo* generation of biologically functional tissue substitutes, provides new strategies to overcome the existing limitations. Currently, two major approaches are pursued to create new adipose tissue: the implantation of cell-containing constructs and the induction of *de novo* genesis of adipose tissue *in vivo*. Therefore, the three classical components of tissue engineering, that is, cells, scaffolds, and tissue-inducing substances, are combined in various ways. Ongoing research is furthermore focused on the crucial process of vascularization of the generated fat. Additionally, engineered adipose tissue constructs have recently been established as a model system for basic research. The model especially considers three-dimensional cell–cell and cell–extracellular matrix (ECM) interactions and is proposed to be utilized in investigations on the endocrine functions of adipose tissue and fat-related diseases, such as obesity.

BACKGROUND

Clinical Need

A transplantable and biologically functional adipose tissue graft is especially required in reconstructive and plastic surgery. It is needed to compensate various soft tissue defects, for example, after tumor resections. In 2004 alone, over 60,000 clinically indicated breast reconstructions were reported by the American Society of Plastic Surgery and in the same year over 16,000 so far commonly used breast implants had to be removed again because of incompatibilities and other

problems.^[1] Furthermore, fat equivalents are needed to reconstruct congenital abnormalities and traumatic soft tissue defects, especially large burns or chronically open wounds. Cosmetic and aesthetic plastic surgeons would also prefer a natural implant to compensate contour deformities.^[2]

Current Surgical Techniques and Their Limitations

One approach to soft tissue replacement is to use alloplastic surrogates, such as Teflon and silicone implants, or allogeneic or xenogeneic materials, such as calcium hydroxyl apatite, hyaluronic acid (HA), collagen, and decellularized tissue matrix implants.^[1,3] The greatest drawbacks of these grafts are fibrous capsule formation, allergic or foreign body reactions, suboptimal mechanical properties, distortion and migration of the implant, and long-term resorption.^[2]

Another strategy to augment soft tissue defects is to use autologous tissue, which is regarded to be the ideal implant in terms of biocompatibility, immune response, and avoidance of tissue rejection. Smaller deformities are currently treated with isolated injections of mature adipocytes. However, as lipid-loaded adipocytes are very fragile, the majority of the fat cells are mechanically damaged during aspiration or preparation. The remaining cells tend to form cysts or localized necrosis postinjection. Free fat grafts also tend to get necrotic, as vascularization of the grafts is too slow leading to insufficient oxygen supply. Poor vascularization also includes limited transport of nutrients as well as deficient removal of metabolic waste products that may also contribute to the poor survival of the transplants. Therefore, these methods have an unpredictable outcome with regard to transplant shrinkage, and repeated surgery might become necessary. To overcome the limitations of insufficient neovascularization, pedicled flaps can be used. These grafts consist of subcutaneous fat, fasciae, and, if greater stability and integrity is necessary, muscle. The

vessels of the pedicled grafts are microsurgically integrated into the preexisting vasculature. These grafts show better survival rates; however, the surgical procedure is complicated and, therefore, cost-intensive. Furthermore, the acquisition of the flap might cause severe donor site morbidity, such as hypertrophic scar formation and constricted motility of the affected area.^[2]

Origin and Characteristics of Adipose Tissue

When seeking the optimal method to engineer adipose tissue as an alternative therapeutic approach, it is necessary to be aware of the typical characteristics of fat, its natural development, and its biological functions.

There are two types of adipose tissue: brown adipose tissue and white adipose tissue (WAT). Brown adipose tissue is an organ that dissipates energy in the form of heat and represents only a small portion of the total body fat.^[4] This work focuses only on WAT, which is the sole object of the current adipose tissue engineering research. White adipose tissue is located throughout the body, but most of it is situated subcutaneously and viscerally.

On the cellular level, WAT mainly consists of lipid-loaded mature adipocytes, adipocyte precursor cells and, as it is highly vascularized, various blood cells, epithelial cells, pericytes, and fibroblasts (Fig. 1).^[4] The origin of adipocytes is still not fully understood. Pluripotent embryonic stem cells are thought to give rise to multipotent mesenchymal stem cells (MSCs).

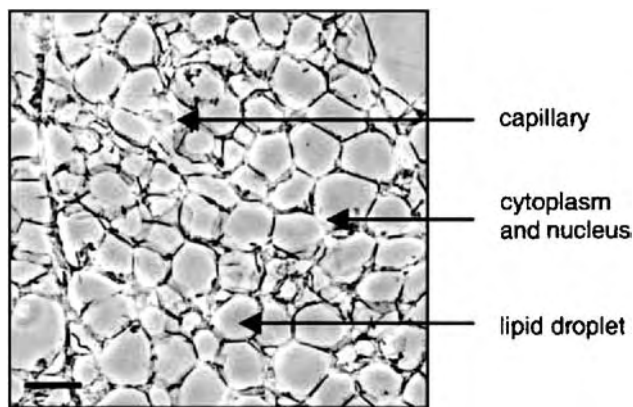


Fig. 1 Native fat (H&E staining). Accumulation of fat droplets that coalesce to one large droplet during adipocyte cell maturation leads to the displacement of the cytoplasm and the nucleus to the edge of the cell. This morphology is described as the signet ring form. Hexagonal-shaped adipocytes are surrounded by basement membrane. Capillaries are dispersed throughout the tissue. Bar = 50 μ m. (From Ref. [5].)

Further downstream the adipocytic lineage, it has been established that the commitment of MSCs leads to unipotent adipocyte progenitor cells.^[6] On the right environmental stimulation, these so-called preadipocytes can differentiate and mature into adipocytes.^[7] Many inducing factors, signal transduction pathways, and molecular events of adipogenesis are well characterized and summarized in various reviews.^[8,9] However, there are also mechanisms in adipose differentiation that have not been fully elucidated yet, such as the contribution of cell–cell or cell–ECM interactions.^[10,11] Adipose tissue development begins during fetal development; postnatal fat growth occurs either by enlargement of differentiated adipocytes (hypertrophy) or by recruitment of new adipocytes (hyperplasia).^[4,8]

The functions of fat are to serve as filling and isolation material and as a storage organ for energy homeostasis. In recent years, fat has additionally been established to be an active endocrine and immunological organ. In particular, adipocytes secrete various cytokines, hormones, and other biochemically active molecules involved in the regulation of insulin sensitivity, glucose homeostasis, inflammation, energy balance, lipid metabolism, reproduction, and vascular homeostasis.^[12]

STRATEGIES OF ADIPOSE TISSUE ENGINEERING

The aim of adipose tissue engineering approaches is to generate fat that is equivalent to native adipose tissue with regard to its cellular and extracellular structure and composition as well as its biological functions. Furthermore, it has to be nonimmunogenic to avoid allergic reactions and tissue rejection. In addition, it is important for clinical purposes to produce tissue volumes sufficient for the intended therapeutic use and to be able to create different implant shapes. The developed fat needs to be well vascularized and it must not be resorbed over time.

To comply with these requirements, presently two major strategies are pursued that differ primarily in whether cells are transplanted or not (Fig. 2). The first approach to generate fat equivalents is to implant cell-containing constructs. Cells are isolated from the desired donor tissue and, if necessary, propagated *in vitro* until a sufficient number of cells are obtained. After seeding onto or into a scaffold that functions as cell carrier, the construct is implanted. Alternatively, the cell-carrier constructs can be further cultivated and differentiated *in vitro* and implanted at later time points. Additionally, it is possible to administer

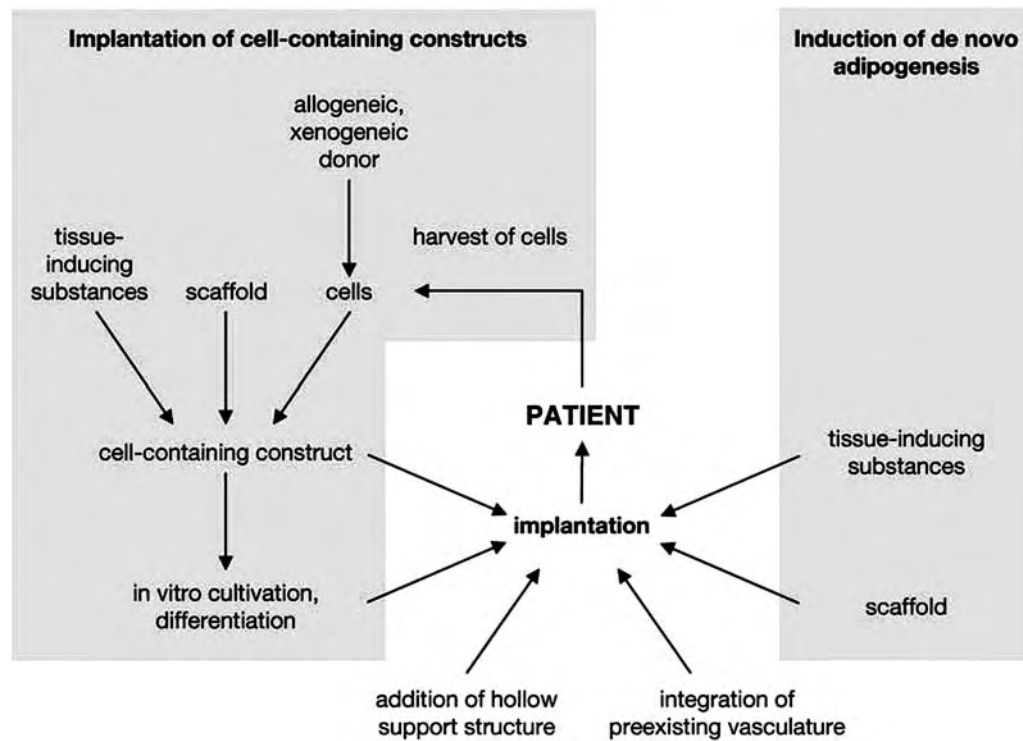


Fig. 2 Different strategies for the tissue engineering of fat. Left: implantation of cell-containing constructs; right: induction of de novo adipogenesis. The integration of preexisting vasculature and the addition of hollow support structures are options applicable in both strategies.

growth factors or tissue-inducing substances to enhance adipogenesis or also angiogenesis, either by simple injection into the recipient site or by application of a controlled release device. Furthermore, vascularization (i.e., oxygen supply, nutrient supply, and removal of waste products) can be enhanced by microsurgical integration of the preexisting vasculature.

The second strategy of fat engineering is to induce de novo adipogenesis *in vivo*. Here, only adipose tissue-inducing substances are administered. These substances then facilitate the *in vivo* development of preferably vascularized adipose tissue. Also in this case, the application of supporting scaffolds and the integration of preexisting vasculature are additional options. In comparison with the first strategy, here no cells are transplanted. This can be advantageous in future clinical applications because problems concerning the harvest and immunological compatibility of the cells would be avoided and no cost- and time-intensive *in vitro* cell culture would be necessary.

In the following chapters the main components of both strategies, i.e., cells, scaffolds, tissue-inducing substances, and vascularization, are discussed in detail; here, the corresponding adipose tissue engineering studies are only mentioned. Subsequently, an extensive

discussion of the approaches and achievements of these publications is provided.

Cell Sources

The ideal cells for the engineering of fat should fulfill the following criteria: The cells need to be autologous or, at least, nonimmunogenic. The cells have to be available in sufficient quantities either by direct harvest of enough material or by their propagation *in vitro* without the loss of their potential to differentiate into adipocytes. The harvest should be possible with minimally invasive procedures to ensure minimal donor site morbidity. Potential cell types generally can be found along the adipocytic lineage.

Mature adipocytes

As already mentioned, mature adipocytes are very fragile and easily traumatized during cell preparation. It is principally possible to maintain and propagate differentiated adipocytes *ex vivo*; however, these cultures are difficult to handle *in vitro* because of the buoyancy of lipid-loaded cells.^[13,14] The use of mature adipocytes is, therefore, not a prevalent method.^[7]

Preadipocyte cell lines

Other possible cell sources are preadipocyte cell lines, for example, 3T3-L1 and 3T3-F442A.^[5,15-17] These cells are commercially available and can easily be propagated and differentiated into adipocytes in vitro. Major drawbacks of cell lines are their aneuploid status and xenogeneic origin. A clinical use of preadipocyte cell lines is therefore unlikely. However, in contrast to primary cell sources, cell lines are reproducibly uniform in their process and status of differentiation. Therefore, they are an appropriate cell type to establish the principles of fat engineering and they are well suited for model systems in basic research.^[5,7,16,17]

Adipose-derived progenitor cells

The cell source that is currently utilized in most cell-based approaches and that is estimated most promising for clinical purposes is adipose-derived progenitor cells.^[18-27] Adipose tissue contains a cellular fraction that can be isolated by enzymatic digestion of fat tissue, optional differential centrifugation, and subsequent adherence to plasticware surfaces. This fraction of cells has been variously termed adipose-derived (adherent) stromal cells, processed lipoaspirate cells, adipose-derived (adult) stem cells, and preadipocytes.^[28] Today this cell pool is considered to be a mixture of, among others, noncommitted multipotent stem cells that can differentiate at least along all the mesenchymal lineages, and committed unipotent preadipocytes that can only differentiate into adipocytes. At the moment, no distinct markers are established, which could facilitate the strict separation between stem cells and preadipocytes. Therefore, this cell pool will be subsumed as adipose-derived progenitor cells in this work.

Adipose-derived progenitor cells are easily available from lipoaspirates or liposections. A device that combines direct harvest of lipoaspirates and isolation of cells has already been developed.^[29] A shortage of autologous donor tissue is unlikely for most patients. Furthermore, several in vivo studies support the general feasibility of allogeneic adult stem cell transplantation.^[6] Allogeneic transplantations carry the risk of immune rejection by the host and transfer of infectious agents. Nevertheless, the possible use of allogeneic cells would enable the fabrication of off-the-shelf transplants eliminating the need to perform elective liposuction to obtain autologous cells and facilitating faster treatment. Therefore, it is currently discussed whether also an allogeneic transplantation of adipose-derived progenitor cells is possible.^[28]

Adipose-derived progenitor cells have the capacity to proliferate in vitro over several passages with low

levels of senescence and with conservation of their potential to differentiate into adipocytes.^[30,31] However, it has to be considered that the self-renewing and differentiation capacity of this heterogenous cell pool may vary depending on the donor itself (sex, age, and health status) and the donor site (epididymal, mammary, subcutaneous, omentum, etc.).^[32] Fat depots at different sites of the body can vary substantially with regard to fat cell morphology and function. These differences are discussed to be not solely a result of site-specific environmental factors (nutrient availability, innervation, anatomic constraints, etc.) modulating the development of mature adipocytes but also a result of the site-specific characteristics of preadipocytes the mature fat cells originate from.^[33]

Bone marrow-derived MSCs

Whereas they have been applied in the engineering of several other tissues for years, only recently, bone marrow-derived MSCs have also been used for adipose engineering purposes in different studies.^[34,35] Mesenchymal stem cells can be isolated from bone marrow aspirates.^[6] The easiest way to separate MSCs from other cellular fractions, that is, primarily the hematopoietic fraction, is by their tendency to adhere to tissue culture plastic. Other possibilities to further purify MSCs are differential centrifugation and fluorescent activated cell sorting or magnetic activated cell sorting. However, no selective surface markers to identify the "pure" MSC have been established yet.

Bone marrow-derived MSCs were reported to have a high self-renewal potential in vitro.^[36,37] They differentiate into adipocytes in vitro and in vivo on adipogenic stimuli. Mesenchymal stem cells are discussed to be nonimmunogenic in allogeneic transplantations.^[6] In contrast to adipose-derived progenitor cells, harvest of bone marrow-derived MSCs is more inconvenient for the donor and only smaller amounts are accessible. The properties of the isolated MSC pool are also donor-dependent.

Embryonic stem cells

Embryonic stem cells have not been used for the tissue engineering of fat yet, but have to be mentioned as a possible cell source in the future. The major advantage of embryonic stem cells is their almost infinite proliferation potential and their capacity to differentiate into a wide variety of cell types, such as adipocytes.^[38,39] Ethical concerns and legal constraints limit the application and availability of human embryonic stem cells at present. Nevertheless, cells from other species, such as mouse embryonic stem cell lines, are readily available for basic research purposes.

Scaffold Materials

A scaffold can serve as a template for the desired three-dimensional shape of the engineered tissue. Current opinion favors biodegradable scaffold materials that provide initial, transient stability. The newly developed tissue, that is, the cells and secreted ECM, is responsible for the long-term maintenance. Therefore, it is necessary to adjust the degradation kinetics to the tempo of tissue development. The bulk material and its biodegradation products need to be biocompatible. Further requirements are adequate mechanical properties; very rigid scaffolds might not be appropriate to augment soft tissue defects. The surface structure and chemistry should allow cell attachment and facilitate differentiation and maintenance of the biological functions of the cells. Another very critical feature is the microarchitecture of the scaffolds. High porosity and interconnectivity are necessary to guarantee an even distribution of the cells, capillary ingrowth, and sufficient nutrient supply. A specific characteristic of adipocytes is that they enlarge to over 100 μm in diameter during maturation and potentially change their size because of their lipid storage function in energy homeostasis. Therefore, adequately large pore sizes within the scaffold have to be provided to enable full differentiation of the cells.^[22]

Synthetic materials

Possible scaffold materials include synthetic biodegradable and absorbable polymers, such as polylactic acid (PLA), polyglycolic acid (PGA), and copolymers thereof, which are Food and Drug Administration (FDA) approved and clinically widely used as sutures. Disadvantages of these materials are their acidic degradation products, which may affect cell functions and the lack of chemically reactive groups that are easily accessible for surface modifications. For adipose tissue engineering, these polymers have been processed by salt-leaching techniques or solid-lipid templating, which leads to relatively stiff scaffolds.^[18,19,34,40] Softer constructs are achieved by the use of meshes fabricated with textile technologies.^[5,16] Polytetrafluoroethylene (PTFE) is a non-biodegradable polymer that has already been used as filling material in aesthetic surgery. Polytetrafluoroethylene scaffolds proved useful to seed preadipocytes if coated with ECM proteins.^[24]

Semisynthetic and natural materials

One approach is to mimic the native cellular environment of fat cells by the use of ECM compounds naturally occurring in adipose tissue, such as collagen

or further compounds of the basement membrane. Further options are materials based on extracellular components of other tissues, such as fibrin glue or HA esters. These materials are characterized by low toxicity and, if autologous material is used, by a low inflammatory response. Furthermore, most ECM compounds can be processed into gels, which allow minimally invasive application. Drawbacks, however, include batch-to-batch variations, easy denaturation, and difficult chemical modification due to their complex structure. The relatively low mechanical stability of ECM-based scaffolds may not be a major problem in soft tissue engineering.

Collagen has been used as gels or as porous sponges produced by freeze-drying of diluted solutions.^[20,22,23] It has been reported that collagens enhance the adipose conversion of precursor cells.^[41] Major disadvantages of collagens are their potential immunogenicity and relatively high price. Matrigel is a basement membrane extract derived from Engelbreth-Holm-Swarm mouse sarcoma, which is rich in ECM. The main components of Matrigel are laminin, collagen type IV, proteoglycans, and entactin, which are all components of the basement membrane that surround adipocytes in native tissue. Furthermore, it contains different growth factors, such as platelet-derived growth factor (PDGF), transforming growth factor (TGF- β), and basic fibroblast growth factor (bFGF); however, the exact composition is not defined. Matrigel promotes cell attachment of adipocyte precursor cells, adipogenesis, and angiogenesis in vivo.^[15,42,43] However, it is still unclear to what extent this effect can be attributed to the presence and architecture of the ECM compounds or to the existence of the adipogenic and the angiogenic growth factors. It has to be noted that as Matrigel is a mouse tumor product, it does not appear appropriate for clinical purposes. Fibrin glue is already FDA approved for a variety of clinical applications and has also been used for adipose tissue engineering purposes.^[26,27,44] It can be produced from autologous blood, thereby avoiding tissue reactions. Fibrin is dissolved rapidly in vivo by proteases, but residence time can be prolonged by adjusting the process parameters during gelation and addition of protease inhibitors.^[45,46] Hyaluronic acid is highly soluble and, therefore, rapidly resorbed in vivo. But esterification of HA leads to insoluble polymers that can be easily processed into sponges or fiber meshes suitable for fat engineering.^[20,21] An advantage of HA-based scaffolds might be that degradation products of HA are known to enhance angiogenesis in vivo.^[47]

Besides ECM compounds, other biological materials, such as alginates, can be used as scaffold materials. Alginates are cheap, biocompatible, and nontoxic. But as they are strongly hydrophilic, protein adsorption is reduced. Therefore, chemical modifications, such

as the linkage of lectins or RGD-containing adhesion peptides, are necessary to enable cell–material interactions.^[25]

Growth Factors and Adipogenic Inducers

On the one hand, the scaffold itself as well as its degradation products can enhance adipose tissue development, as discussed in the previous entry. On the other hand, there are a great number of publications that identify growth factors and other biological or pharmaceutical agents that influence proliferation and differentiation of adipocyte precursor cells and adipose tissue development. In some cases, the studies come to contradictory results, most probably because of the fact that different cell lines and primary cell types have been studied under diverse culture conditions. A summary of growth factors and adipogenic inducers that can be used for adipose tissue engineering purposes is given in the following.

One important step during the generation of cell-based constructs is the efficient expansion of precursor cells *in vitro*. Factors that enhance *in vitro* proliferation of adipose-derived progenitor cells and bone marrow-derived MSCs include insulin-like growth factors (IGF-1, IGF-2), acidic FGF, bFGF, epidermal growth factor (EGF), and PDGF; TGF- β 1 is reported to stimulate the proliferation of adipose-derived progenitor cells, but inconsistent results exist concerning the growth of MSCs.^[29]

In vitro adipogenesis of adipose-derived progenitor cells and bone marrow-derived MSCs can be induced by administration of a hormonal cocktail typically composed of insulin or IGF-1, a glucocorticoid (corticosterone, dexamethasone, hydrocortisone, etc.), and cAMP or an agent increasing intracellular cAMP levels, such as phosphodiesterase inhibitors (isobutylmethylxanthine and forskolin).^[8] Furthermore, agonists of peroxisome proliferator-activated receptor gamma (PPAR γ), a key transcription factor during adipogenesis, also can be used either alone or in combination with the mentioned hormonal cocktail to induce adipogenesis. The PPAR γ agonists include thiazolidinediones (rosiglitazon, troglitazon), indomethacin, and prostaglandins of the D and J series that have recently been found to be natural ligands of PPAR γ .^[48] Growth hormone and thyroid hormone have been reported to enhance adipogenesis of preadipocyte cell lines, but they have no or even a negative effect on the differentiation of primary precursors.^[8] The influence of bFGF on adipogenesis is controversially discussed. Basic FGF has been considered to inhibit adipogenesis of preadipocyte cell lines and primary adipose-derived progenitor cells.^[8] However, recent studies have shown a distinct increase of

adipogenesis of MSCs when bFGF was administered.^[34,49] Most of the knowledge about adipogenic inducers has been gained *in vitro*. The physiological relevance of the inducing hormones is still not fully clear. Nevertheless, administration of IGF-1 and insulin can induce *de novo* adipogenesis *in vivo* and prolong survival of newly generated or transplanted mature adipose tissue.^[50,51] *In vivo* application of bFGF also induces *de novo* adipogenesis.^[23,26,42,50,52] However, it is unclear whether bFGF has a direct adipogenic effect or adipogenesis is triggered through neoangiogenesis. (For a more detailed discussion of the relation between adipogenesis and angiogenesis, see the next entry.) Besides the search for inducing agents, the way of *in vivo* application has to be considered. Several of the aforementioned studies revealed that simple injection of the substrates is less effective than application of devices that enable the controlled release over a defined and adjustable period of time.^[23,42]

Vascularization

As already mentioned above, vascularization is a crucial requirement for adequate development and long-term survival of engineered adipose tissue. Fat is a highly vascularized organ with every adipocyte attached to at least one capillary. Moreover, the metabolism of adipose tissue is even higher than that of muscle, if normalized to the amount of cytoplasm. Thus, it is reasonable that adipose tissue easily gets necrotic in avascular areas.^[2] Furthermore, fetal and adult adipogenesis is spatially and temporally closely related to angiogenesis.^[53,54] Both processes are coordinated via an intensive cross-talk between endothelial cells and adipocytes. Endothelial cells secrete various soluble and matrix-bound factors which promote proliferation as well as differentiation of adipose-derived progenitor cells; one of these factors has been identified as acidic FGF.^[55–57] Platelet-derived growth factor is also found to be involved in the cross-talk, as it is secreted by endothelial cells and enhances adipose precursor cell proliferation.^[29,58] In turn, established adult adipose tissue modulates angiogenesis via several secreted factors, such as vascular endothelial growth factor (VEGF), angiopoietin-1, and others.^[53,59]

Different strategies are pursued to enhance vascularization of the engineered adipose tissue. One approach that has been adopted from surgical procedures is the integration of preexisting capillary structures into an implanted scaffold.^[52,60] The vascular pedicle provides initial oxygen and nutrient supply and is the origin for sprouting capillaries. Another strategy is the administration of angiogenic growth factors, such as VEGF and bFGF, which recruit

endothelial cells and their precursors and enhance sprouting of new vessels.^[26,42,50,52] Effects of growth factors might be improved via a prolonged application by the use of controlled release devices or gene transfer. Vascularization can also be achieved by transplantation of endothelial cells, which then form vascular networks within the newly formed tissue.^[27]

CURRENT APPROACHES AND ACHIEVEMENTS

In Vitro Studies

In vitro cultures have been used for different purposes in adipose tissue engineering. Several researchers utilized in vitro studies to explore the suitability of new scaffold materials with regard to their interaction with cells. von Heimburg et al. seeded adipose-derived progenitor cells onto different sponges consisting either of collagen or of HA benzyl esters (HYAFF[®] 11) of various degrees of esterification and different pore sizes.^[21,22] In analog studies, it was shown that adipose-derived progenitor cells adhere also to the salt-leached PLGA scaffolds or PTFE scaffolds.^[18,24] The latter have to be coated with ECM components to enable cell attachment; fibronectin turned out to be more effective than albumin and collagen. Furthermore, all of these studies proved that the precursor cells differentiate into adipocytes and develop a mature phenotype within the scaffold. Hong et al. established that also human bone marrow-derived MSCs are capable of differentiating into mature adipocytes in a three-dimensional context.^[35] However, none of the aforementioned studies could show in vitro formation of coherent tissue.

Other studies aimed at engineering mature adipose tissue in vitro. Neubauer et al. seeded bone

marrow-derived MSCs onto PLGA scaffolds.^[34] On repeated hormonal induction, MSCs matured into adipocytes; administration of bFGF significantly enhanced adipogenesis. After 4 weeks of cultivation tissue-like structures with mature unilocular adipocytes could be detected at least in the outer areas of the scaffolds, whereas isolated immature, multivacuolar adipocytes were observed in the inner parts. So far only one approach achieved the development of coherent adipose tissue in vitro.^[5] Constructs consisting of 3T3-L1 preadipocytes and PGA nonwoven fiber meshes were cultivated over 5 weeks in vitro. Cells differentiated into mature adipocytes exhibiting the typical signet ring form and the constructs developed into coherent fat pads histologically comparable to native adipose tissue (Fig. 3A).

In Vivo Studies

Implantation of cell-containing constructs

The first study conducting the implantation of cell-containing constructs to engineer adipose tissue was presented by Patrick et al.^[18] Adipose-derived progenitor cells seeded onto salt-leached PLGA scaffolds were subcutaneously implanted into rats. After 5 weeks, clusters of vascularized adipose tissue had developed within the implant with remnants of scaffold material dispersed throughout the construct. Further in vivo investigation revealed that the implants as well as the newly formed adipose tissue were completely resorbed after 4 mo.^[19] As this time point correlated with the degradation kinetics of the scaffold material, the engineered tissue was probably absorbed when the supporting structure broke down. In comparable studies, von Heimburg et al. used HYAFF

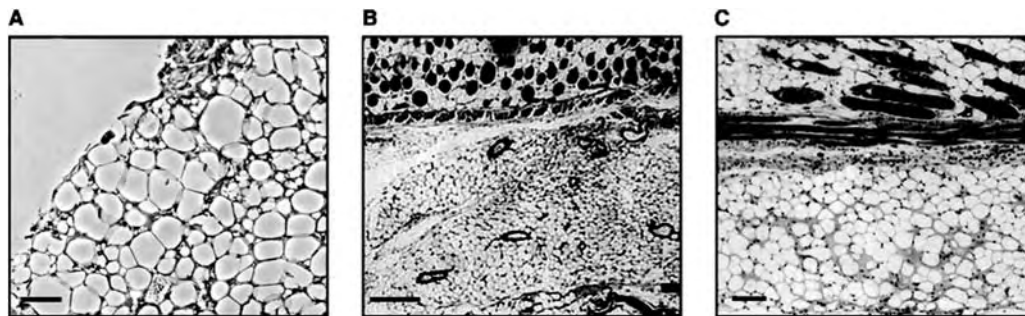


Fig. 3 Adipose tissues engineered with different strategies (H&E staining). (A) Coherent fat pads generated in vitro. Preadipocytes (3T3-L1) were seeded onto PGA nonwoven fiber meshes, hormonally induced, and further cultivated for 5 weeks. Bar = 50 μ m. (B) Formation of adipose tissue (lower part of the image) in the subcutis of BALB/c mice 6 weeks after implantation of collagen sponges incorporating human adipose-derived precursor cells and gelatin microspheres containing 1 μ g of bFGF. Bar = 300 μ m. (C) De novo formation of adipose tissue (lower part of the image) in the subcutis of BALB/c mice 15 weeks after implantation of Matrigel mixed with gelatin microspheres incorporating 0.1 μ g of bFGF. Bar = 150 μ m. (From Refs. [5,23,42].)

sponges, HYAFF nonwoven fiber meshes, and collagen sponges as cell carriers.^[20,22] Scaffolds were seeded with human adipose-derived progenitor cells and implanted subcutaneously for up to 12 weeks. Only within HYAFF sponges could areas of vascularized adipose tissue be detected. In contrast, implant sizes of HYAFF nonwoven fiber meshes and collagen sponges drastically decreased over *in vivo* time. The fiber meshes had collapsed during implantation; thus, the space available for preadipocyte growth and differentiation was significantly reduced. After explantation of collagen sponges, mature adipocytes were only located in the surface areas of the scaffolds. This was probably due to the fact that cells could not fully penetrate the collagen sponges because of their small pore size (40–100 μm). Another material used for the delivery of adipose-derived progenitor cells is alginate modified with a covalently bound cell-adhesion peptide (arginine-glycine-aspartic acid, RGD).^[25] Progenitor cells suspended in alginate–RGD also developed into adipose structures after subcutaneous implantation into sheep.

The application of growth factors in conjunction with cell-containing constructs can enhance *in vivo* adipose tissue formation. Human adipose-derived progenitor cells seeded on collagen sponges and implanted into the subcutis of mice for 6 weeks gave rise to adipose tissue only if bFGF was administered.^[23] Best results were achieved when bFGF was released over an extended period of time from gelatin microparticles (Fig. 3B). Yuksel et al. used PLGA/PEG microparticles to locally release insulin, IGF-1, and bFGF.^[50] All three growth factors were reported to increase the survival of transplanted diced mature adipose tissue fragments for up to 12 weeks, but bFGF led to infiltration with interadipocyte stromal tissue. Generation of adipose tissue with the help of growth factors resulted in the development of more coherent fat pads; however, also in these cases the long-term maintenance has not been established yet.

Whereas, in the studies described above, constructs were implanted shortly after cell seeding, a different approach is to implant cell-containing constructs that already have been induced and cultivated *in vitro*.^[5] Implantation of coherent fat pads, generated from 3T3-L1 preadipocytes and PGA nonwoven meshes (see previous entry) resulted in vascularization and further development of the constructs. The size of the engineered fat was maintained *in vivo* over 3 weeks, but a long-term investigation was not conducted.

The mentioned approaches show that adipose tissue engineering based on the application of cell-containing constructs is, in principle, possible with various types of scaffolds. However, volumes of the constructs were smaller than 1 ml in all cases. Furthermore, long-term survival (>6 mo) of engineered fat was not achieved or

not investigated in these studies. One important reason why newly generated fat is absorbed is the insufficient neovascularization of the constructs. Schoeller et al. showed that adipose-derived progenitor cells can develop into mature adipose tissue that persists over 12 mo if implanted into well-vascularized capsules.^[44] These capsules were formed by inducing a foreign-body reaction through implantation of silicon tubes. This study clearly points out that a sufficient blood supply is necessary for long-term survival, but it is questionable if this technique of capsule formation is clinically applicable.

Induction of de novo adipogenesis

Application of bFGF in combination with Matrigel but without the addition of any cells proved to induce de novo development of adipose tissue in different studies.^[42,43] The newly formed fat was a well-vascularized coherent tissue consisting of mature adipocytes (Fig. 3C). Long-term release of bFGF from gelatin microspheres enhanced the effect of bFGF. Matrigel and bFGF are both known to enhance angiogenesis. Further proadipogenic effects of bFGF are currently being discussed. It is not fully understood yet whether the combination of Matrigel and bFGF induces adipogenesis directly or via the development of capillary structures within the transplant. Interestingly, neither injection of bFGF nor Matrigel alone could induce de novo adipogenesis; however, there are no obvious explanations for this synergy. Despite the good quality of the de novo engineered fat, it has to be mentioned once more that Matrigel is a tumor mouse product and, therefore, not suitable for clinical applications. Besides the effects of bFGF, Yuksel et al. reported that long-term delivery of insulin or IGF-1 also induced de novo adipogenesis.^[51] Especially, coimplantation of insulin- and IGF-1-loaded PLGA/PEG microspheres led to the development of mature vascularized adipose tissue. However, similar to the studies in which cell-containing constructs were implanted, volumes generated via de novo adipogenesis also did not exceed 1 ml and no long-term study has proved the survival yet.

Implantation of synthetic hollow support structures

A recent trend in adipose tissue engineering is to implant synthetic hollow support structures consisting of silicone, PGA/PLA, or polycarbonate (Fig. 4A).^[26,52,60] These approaches take advantage of the natural characteristic of fat to fill void spaces. Furthermore, such support structures could protect the developing fat from potentially deforming forces. On the one hand, the structures were placed around

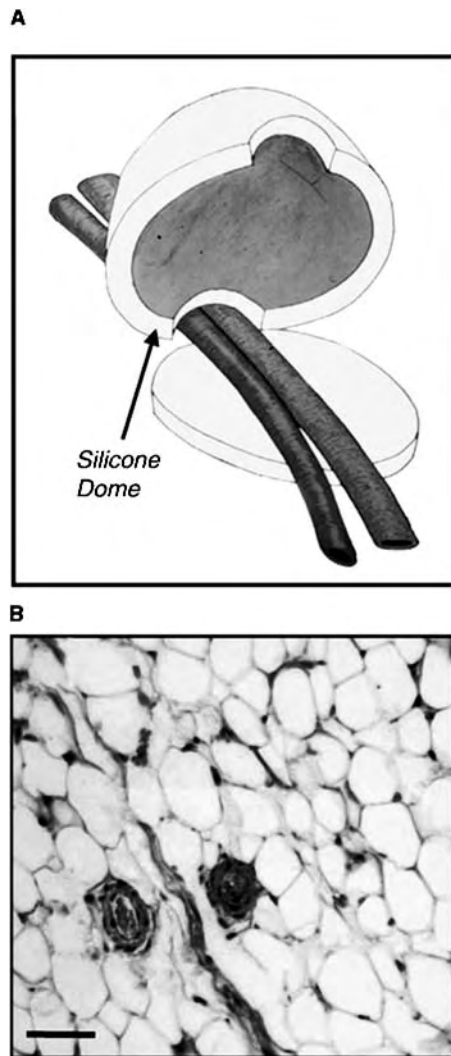


Fig. 4 Application of hollow support structures. (A) Diagram of dome-shaped construct placed around preexisting vasculature. (B) Adipose tissue engineered within a hollow silicon support structure (H&E staining). Injection of Matrigel and bFGF into the mold induced de novo formation of mature fat. Bar = 200 μm . (From Ref. [52].)

existing vasculature and filled either with Matrigel and bFGF or with fibrin, Matrigel, or PLGA.^[52,60] Invasion and differentiation of preadipocytes led to de novo adipose tissue formation within the capsule (Fig. 4B). On the other hand, injection of preadipocytes suspended in fibrin glue in dome-shaped structures also gave rise to mature adipose tissue.^[26] It was reported in a conference contribution that by this principle it was possible to generate larger adipose tissue volumes in the order of 50 ml in a pig model.^[60] However, specific procedures and results have not been further published yet. Though current results are quite promising, it has to be considered that these approaches require complex surgical procedures

during implantation; furthermore, biodegradable structures have to be developed to avoid the otherwise inevitable explantation of the support structures.

TISSUE-ENGINEERED FAT FOR BASIC RESEARCH

Besides the interest of plastic and reconstructive surgeons, fat is also the topic of research concerning obesity and related disorders, such as accelerated atherosclerosis, increased risk of thrombosis, hypertension, insulin resistance, and type 2 diabetes. Fat has been established to be an active endocrine organ involved in many important physiological and pathological processes.^[12] (See also entry “Origin and Characteristics of Adipose Tissue.”) Therefore, rapidly increasing efforts have been made to investigate the basic understanding of adipogenesis, fetal and adult adipose tissue development, and regulation of adipocyte functions. Most of the knowledge about molecular regulation of adipocytes has been acquired using two-dimensional *in vitro* cell cultures with preadipocyte cell lines, such as 3T3-L1, 3T3-F442A, or Ob-17. Two-dimensional cultures are appropriate model systems for many questions; however, they only marginally reflect the physiological environment of adipose tissue. Recent research has suggested that adipogenesis as well as regulatory or secretory functions of adipocytes are influenced by tissue inherent factors, such as cell–cell or cell–matrix interaction.^[8,53] Tissue-engineered model systems could overcome existing experimental limitations by providing three-dimensional fat constructs that partially mimic the *in vivo* environment. Recently, Fischbach et al. presented a system to develop coherent adipose tissue *in vitro* (see entry “*In vitro* studies”).^[5] This study reported several differences in adipocyte morphology and function depending on whether cells were cultivated in two-dimensional culture plates or within three-dimensional constructs (Fig. 5). Although no coherent tissue was presented, differences in adipocyte marker gene and protein expression between two-dimensional and three-dimensional cultures were also found by Kang, Xie, and Kniss, who recently established a similar model consisting of 3T3-L1 cells cultivated on fibrous PET scaffolds.^[17] A drawback of the presented models might be that immortalized cell lines have different properties from primary cells with regard to proliferation potential and differentiation capacity. However, on the other hand, the use of a cell line provides reproducible and standardized conditions. Therefore, these and other tissue-engineered systems may be used to address the relevance of interactions between adipocytes and their surroundings and may contribute to the further elucidation of the important secretory

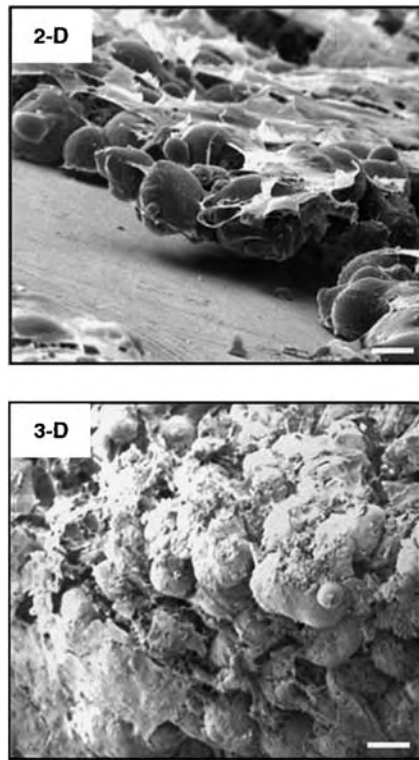


Fig. 5 Comparison of cell and tissue development in two- and three-dimensional cell culture systems. Scanning electron micrograph images of 3T3-L1 cells grown in vitro in two-dimensional cell culture or three-dimensional tissue-engineered constructs. Cells in two-dimensional monolayers are loosely connected to each other and ECM sheets are preferentially deposited on the cell surface. In contrast, three-dimensional differentiated adipocytes are situated within a coherent environment featuring three-dimensional cell-cell and cell-ECM interactions. Bar = 20 μm . (From Ref. [5].)

functions of adipose tissue. In turn, the knowledge gained on microenvironmental characteristics can contribute to improving the engineering of biologically functional adipose tissue.

CONCLUSIONS

To date, several approaches have been investigated to engineer adipose tissue in vivo and in vitro. Different strategies have been presented including the implantation of cell-containing constructs and the de novo generation of fat. In principle, both strategies enabled the generation of fat constructs. However, there are still major questions to be solved until clinical application becomes a reality. First, methods have to be established that facilitate the long-term survival of the engineered constructs. The problem of fat implant resorption that has bothered plastic surgeons for over

almost 100 yr has not been solved satisfactorily by tissue engineers yet. This problem is closely related to the necessity to enhance early and sufficient vascularization of the constructs. Furthermore, it will become inevitable to investigate not only the histological characteristics but also the functionality of the newly generated tissue. To date, only small volumes of engineered fat have been generated. These constructs might be available for clinical purposes in the near future, such as aesthetic facial augmentations. In contrast, it will take more time until it is possible to generate fat transplants sufficient for the treatment of larger defects, for example, after breast tumor resections. Finally, similar to the situation in other tissues, it may be considered if the use of autologous material is strictly necessary or if allogeneic or xenogeneic cells and matrices may also be applicable, enabling, in the long run, the development of off-the-shelf products. Until engineered adipose tissue becomes a widely applicable alternative in plastic and reconstructive surgery, much more intensive research by clinicians, bioengineers, and life scientists is required.

ARTICLE OF FURTHER INTEREST

Tissue Engineering

REFERENCES

1. Statistics of the American Society of Plastic Surgeons (ASPS), 2004; http://www.plasticsurgery.org/public_education/Statistical-Trends.cfm (accessed Apr 2005).
2. Patrick, C.W., Jr; Chauvin, P.B.; Robb, G.L. Tissue engineered adipose tissue. In *Frontiers in Tissue Engineering*, 2nd Ed.; Patrick, C.W., Jr. Mikos, A.G., McIntire, L.V., Eds.; Elsevier Science Ltd.: Oxford, 1998; 369–382.
3. Ashinoff, R. Overview: soft tissue augmentation. *Clin. Plast. Surg.* **2000**, *27* (4), 479–487.
4. Ailhaud, G.; Grimaldi, P.; Negrel, R. Cellular and molecular aspects of adipose tissue development. *Annu. Rev. Nutr.* **1992**, *12* (1), 207–233.
5. Fischbach, C.; Spruss, T.; Weiser, B.; Neubauer, M.; Becker, C.; Hacker, M.; Gopferich, A.; Blunk, T. Generation of mature fat pads in vitro and in vivo utilizing 3-D long-term culture of 3T3-L1 preadipocytes. *Exp. Cell Res.* **2004**, *300* (1), 54–64.
6. Baksh, D.; Song, L.; Tuan, R.S. Adult mesenchymal stem cells: characterization, differentiation, and application in cell and gene therapy. *J. Cell Mol. Med.* **2004**, *8* (3), 301–316.
7. Beahm, E.K.; Walton, R.L.; Patrick, C.W., Jr. Progress in adipose tissue construct development. *Clin. Plast. Surg.* **2003**, *30* (4), 547–558.

8. Gregoire, F.M.; Smas, C.M.; Sul, H.S. Understanding adipocyte differentiation. *Physiol. Rev.* **1998**, *78* (3), 783–809.
9. Rosen, E.D.; Spiegelman, B.M. Molecular regulation of adipogenesis. *Annu. Rev. Cell Dev. Biol.* **2000**, *16* (1), 145–171.
10. Smas, C.M.; Sul, H.S. Control of adipocyte differentiation. *Biochem. J.* **1995**, *309* (3), 697–710.
11. Boudreau, N.J.; Jones, P.L. Extracellular matrix and integrin signalling: the shape of things to come. *Biochem. J.* **1999**, *339* (3), 481–488.
12. Guerre-Millo, M. Adipose tissue hormones. *J. Endocrinol. Invest.* **2002**, *25* (10), 855–861.
13. Sugihara, H.; Yonemitsu, N.; Miyabara, S.; Toda, S. Proliferation of unilocular fat cells in the primary culture. *J. Lipid Res.* **1987**, *28* (9), 1038–1045.
14. Tholpady, S.S.; Aojanpong, C.; Llull, R.; Jeong, J.H.; Mason, A.C.; Futrell, J.W.; Ogle, R.C.; Katz, A.J. The cellular plasticity of human adipocytes. *Ann. Plast. Surg.* **2005**, *54* (6), 651–656.
15. Kawaguchi, N.; Toriyama, K.; Nicodemou-Lena, E.; Inou, K.; Torii, S.; Kitagawa, Y. Reconstituted basement membrane potentiates in vivo adipogenesis of 3T3-F442A cells. *Cytotechnology* **1999**, *31* (3), 215–220.
16. Fischbach, C.; Seufert, J.; Staiger, H.; Hacker, M.; Neubauer, M.; Gopferich, A.; Blunk, T. Three-dimensional in vitro model of adipogenesis: comparison of culture conditions. *Tissue Eng.* **2004**, *10* (1–2), 215–229.
17. Kang, X.; Xie, Y.; Kniss, D.A. Adipose tissue model using three-dimensional cultivation of preadipocytes seeded onto fibrous polymer scaffolds. *Tissue Eng.* **2005**, *11* (3–4), 458–468.
18. Patrick, C.W., Jr.; Chauvin, P.B.; Hobbey, J.; Reece, G.P. Preadipocyte seeded PLGA scaffolds for adipose tissue engineering. *Tissue Eng.* **1999**, *5* (2), 139–151.
19. Patrick, C.W., Jr.; Zheng, B.; Johnston, C.; Reece, G.P. Long-term implantation of preadipocyte-seeded PLGA scaffolds. *Tissue Eng.* **2002**, *8* (2), 283–293.
20. von Heimbarg, D.; Zachariah, S.; Low, A.; Pallua, N. Influence of different biodegradable carriers on the in vivo behavior of human adipose precursor cells. *Plast. Reconstr. Surg.* **2001**, *108* (2), 411–420.
21. Halbleib, M.; Skurk, T.; de Luca, C.; von Heimbarg, D.; Hauner, H. Tissue engineering of white adipose tissue using hyaluronic acid-based scaffolds. I: in vitro differentiation of human adipocyte precursor cells on scaffolds. *Biomaterials* **2003**, *24* (18), 3125–3132.
22. von Heimbarg, D.; Kuberka, M.; Rendchen, R.; Hemmrich, K.; Rau, G.; Pallua, N. Preadipocyte-loaded collagen scaffolds with enlarged pore size for improved soft tissue engineering. *Int. J. Artif. Organs* **2003**, *26* (12), 1064–1076.
23. Kimura, Y.; Ozeki, M.; Inamoto, T.; Tabata, Y. Adipose tissue engineering based on human preadipocytes combined with gelatin microspheres containing basic fibroblast growth factor. *Biomaterials* **2003**, *24* (14), 2513–2521.
24. Kral, J.G.; Crandall, D.L. Development of a human adipocyte synthetic polymer scaffold. *Plast. Reconstr. Surg.* **1999**, *104* (6), 1732–1738.
25. Halberstadt, C.; Austin, C.; Rowley, J.; Culberson, C.; Loebbeck, A.; Wyatt, S.; Coleman, S.; Blacksten, L.; Burg, K.; Mooney, D.; Holder, W., Jr. A hydrogel material for plastic and reconstructive applications injected into the subcutaneous space of a sheep. *Tissue Eng.* **2002**, *8* (2), 309–319.
26. Cho, S.W.; Kim, S.S.; Won Rhie, J.; Mi Cho, H.; Yong Choi, C.; Kim, B.S. Engineering of volume-stable adipose tissues. *Biomaterials* **2005**, *26* (17), 3577–3585.
27. Borges, J.; Mueller, M.C.; Padron, N.T.; Tegtmeier, F.; Lang, E.M.; Stark, G.B. Engineered adipose tissue supplied by functional microvessels. *Tissue Eng.* **2003**, *9* (6), 1263–1270.
28. Gimble, J.M. Adipose tissue-derived therapeutics. *Expert. Opin. Biol. Ther.* **2003**, *3* (5), 705–713.
29. Katz, A.J.; Llull, R.; Hedrick, M.H.; Futrell, J.W. Emerging approaches to the tissue engineering of fat. *Clin. Plast. Surg.* **1999**, *26* (4), 587–603.
30. Rodriguez, A.M.; Elabd, C.; Amri, E.Z.; Ailhaud, G.; Dani, C. The human adipose tissue is a source of multipotent stem cells. *Biochim.* **2005**, *87* (1), 125–128.
31. Zuk, P.A.; Zhu, M.; Mizuno, H.; Huang, J.; Futrell, J.W.; Katz, A.J.; Benhaim, P.; Lorenz, H.P.; Hedrick, M.H. Multilineage cells from human adipose tissue: implications for cell-based therapies. *Tissue Eng.* **2001**, *7* (2), 211–228.
32. Kirkland, J.L.; Hollenberg, C.H.; Gillon, W.S. Age, anatomic site, and the replication and differentiation of adipocyte precursors. *Am. J. Physiol. Cell Physiol.* **1990**, *258* (2), C206–C210.
33. Tchkonina, T.; Tchoukalova, Y.D.; Giorgadze, N.; Pirtskhalava, T.; Karagiannides, I.; Forse, R.A.; Koo, A.; Stevenson, M.; Chinnappan, D.; Cartwright, A.; Jensen, M.D.; Kirkland, J.L. Abundance of two human preadipocyte subtypes with distinct capacities for replication, adipogenesis, and apoptosis varies among fat depots. *Am. J. Physiol. Endocrinol. Metab.* **2005**, *288* (1), E267–E277.
34. Neubauer, M.; Hacker, M.; Bauer-Kreisel, P.; Weiser, B.; Fischbach, C.; Schulz, M.B.; Gopferich, A.; Blunk, T. Adipose tissue engineering based on mesenchymal stem cells and basic fibroblast growth factor in vitro. *Tissue Eng. in press*.
35. Hong, L.; Peptan, I.; Clark, P.; Mao, J.J. Ex vivo adipose tissue engineering by human marrow stromal cell seeded gelatin sponge. *Ann. Biomed. Eng.* **2005**, *33* (4), 511–517.
36. Bruder, S.P.; Jaiswal, N.; Haynesworth, S.E. Growth kinetics, self-renewal, and the osteogenic potential of purified human mesenchymal stem cells during extensive subcultivation and following cryopreservation. *J. Cell Biochem.* **1997**, *64* (2), 278–294.
37. Gregory, C.A.; Prockop, D.J.; Spees, J.L. Non-hematopoietic bone marrow stem cells: molecular control of expansion and differentiation. *Exp. Cell Res.* **2005**, *306* (2), 330–335.
38. Dani, C.; Smith, A.G.; Dessolin, S.; Leroy, P.; Staccini, L.; Villageois, P.; Darimont, C.; Ailhaud, G. Differentiation of embryonic stem cells into adipocytes in vitro. *J. Cell Sci.* **1997**, *110* (11), 1279–1285.

39. Dani, C. Embryonic stem cell-derived adipogenesis. *Cells Tissues Organs* **1999**, *165* (3–4), 173–180.
40. Hacker, M.; Tessmar, J.; Neubauer, M.; Blaimer, A.; Blunk, T.; Gopferich, A.; Schulz, M.B. Towards biomimetic scaffolds: anhydrous scaffold fabrication from biodegradable amine-reactive diblock copolymers. *Biomaterials* **2003**, *24* (24), 4459–4473.
41. Ibrahimi, A.; Bonino, F.; Bardou, S.; Ailhaud, G.; Dani, C. Essential role of collagens for terminal differentiation of preadipocytes. *Biochem. Biophys. Res. Commun.* **1992**, *187* (3), 1314–1322.
42. Kimura, Y.; Ozeki, M.; Inamoto, T.; Tabata, Y. Time course of de novo adipogenesis in matrigel by gelatin microspheres incorporating basic fibroblast growth factor. *Tissue Eng.* **2002**, *8* (4), 603–613.
43. Kawaguchi, N.; Toriyama, K.; Nicodemou-Lena, E.; Inou, K.; Torii, S.; Kitagawa, Y. De novo adipogenesis in mice at the site of injection of basement membrane and basic fibroblast growth factor. *Proc. Natl. Acad. Sci. USA* **1998**, *95* (3), 1062–1066.
44. Schoeller, T.; Lille, S.; Wechselberger, G.; Otto, A.; Mowlavi, A.; Piza-Katzer, H. Histomorphologic and volumetric analysis of implanted autologous preadipocyte cultures suspended in fibrin glue: a potential new source for tissue augmentation. *Aesthetic Plast. Surg.* **2001**, *25* (1), 57–63.
45. Helgerson, S.L.; Seelich, T.; di Orio, J.P.; Bittner, K.; Spaethe, R. Fibrin. In *Encyclopedia of Biomaterials and Biomedical Engineering*; Wnek, G., Bowlin, G., Eds.; Marcel Dekker, Inc.: New York, 2004; 1, 603–610.
46. Meinhart, J.; Fussenegger, M.; Hobling, W. Stabilization of fibrin-chondrocyte constructs for cartilage reconstruction. *Ann. Plast. Surg.* **1999**, *42* (6), 673–678.
47. West, D.C.; Kumar, S. Hyaluronan and angiogenesis. *Ciba Found. Symp.* **1989**, *143*, 187–201.
48. Desvergne, B.; Wahli, W. Peroxisome proliferator-activated receptors: nuclear control of metabolism. *Endocr. Rev.* **1999**, *20* (5), 649–688.
49. Neubauer, M.; Fischbach, C.; Bauer-Kreisel, P.; Lieb, E.; Hacker, M.; Tessmar, J.; Schulz, M.B.; Gopferich, A.; Blunk, T. Basic fibroblast growth factor enhances PPARgamma ligand-induced adipogenesis of mesenchymal stem cells. *FEBS Lett.* **2004**, *577* (1–2), 277–283.
50. Yuksel, E.; Weinfeld, A.B.; Cleek, R.; Wamsley, S.; Jensen, J.; Boutros, S.; Waugh, J.M.; Shenaq, S.M.; Spira, M. Increased free fat-graft survival with the long-term, local delivery of insulin, insulin-like growth factor-I, and basic fibroblast growth factor by PLGA/PEG microspheres. *Plast. Reconstr. Surg.* **2000**, *105* (5), 1712–1720.
51. Yuksel, E.; Weinfeld, A.B.; Cleek, R.; Waugh, J.M.; Jensen, J.; Boutros, S.; Shenaq, S.M.; Spira, M. De novo adipose tissue generation through long-term, local delivery of insulin and insulin-like growth factor-1 by PLGA/PEG microspheres in an in vivo rat model: a novel concept and capability. *Plast. Reconstr. Surg.* **2000**, *105* (5), 1721–1729.
52. Walton, R.L.; Beahm, E.K.; Wu, L. De novo adipose formation in a vascularized engineered construct. *Microsurgery* **2004**, *24* (5), 378–384.
53. Hausman, G.J.; Richardson, R.L. Adipose tissue angiogenesis. *J. Anim. Sci.* **2004**, *82* (3), 925–934.
54. Rupnick, M.A.; Panigrahy, D.; Zhang, C.Y.; Dallabrida, S.M.; Lowell, B.B.; Langer, R.; Folkman, M.J. Adipose tissue mass can be regulated through the vasculature. *Proc. Natl. Acad. Sci. USA* **2002**, *99* (16), 10730–10735.
55. Varzaneh, F.E.; Shillabeer, G.; Wong, K.L.; Lau, D.C.W. Extracellular matrix components secreted by microvascular endothelial cells stimulate preadipocyte differentiation in vitro. *Metabolism* **1994**, *43* (7), 906–912.
56. Hutley, L.J.; Herington, A.C.; Shurety, W.; Cheung, C.; Vesey, D.A.; Cameron, D.P.; Prins, J.B. Human adipose tissue endothelial cells promote preadipocyte proliferation. *Am. J. Physiol. Endocrinol. Metab.* **2001**, *281* (5), E1037–E1044.
57. Hutley, L.; Shurety, W.; Newell, F.; McGeary, R.; Pelton, N.; Grant, J.; Herington, A.; Cameron, D.; Whitehead, J.; Prins, J. Fibroblast growth factor 1: a key regulator of human adipogenesis. *Diabetes* **2004**, *53* (12), 3097–3106.
58. Barleon, B.; Borges, J.; Lauchart, M.; Herzog, C.; Tegtmeier, F.; Walgenbach, K.J.; Andree, C.; Bittner, K.; Stark, G.B.; Marme, D. Angiogenesis in fat tissue: characterization of the molecular and cellular interactions between preadipocytes, adipocytes, and endothelial cells in vitro. In *Cells Tissues Organs* 2000, 2nd BioValley Tissue Engineering Symposium, Freiburg, Germany, Nov 25–27, 1999; Vol. 166, suppl. 1, Abstract B403, 103. .
59. Dallabrida, S.M.; Zurakowski, D.; Shih, S.C.; Smith, L.E.; Folkman, J.; Moulton, K.S.; Rupnick, M.A. Adipose tissue growth and regression are regulated by angiopoietin-1. *Biochem. Biophys. Res. Commun.* **2003**, *311* (3), 563–571.
60. Dolderer, J.H.; Findlay, M.W.; Cooper-White, J.; Thompson, E.W.; Trost, N.; Hennessy, O.; Penington, A.; Morrison, W.A. Making tissue engineering for breast reconstruction a reality. In International Conference “Strategies of Tissue Engineering,” Wuerzburg, Germany, Jun 2004. *Cytotechnology* **2004**, *6* (3) Abstract O-36, 263–263.

Tissue Engineering of Heart Valves

Ivan Vesely

Biomedical Engineering, The Cleveland Clinic Foundation, Cleveland, Ohio, U.S.A.

INTRODUCTION

Valvular heart disease is a slow, debilitating process during which the cardiac valves become stenosed or incompetent. Unless the diseased valve is repaired or replaced with either a bioprosthesis or mechanical artificial valve, this disease eventually results in cardiopulmonary failure and death. Mechanical valves are typically constructed from pyrolytic carbon, and bioprostheses are fabricated from glutaraldehyde-tanned bovine pericardium or whole porcine aortic valves (Fig. 1). These devices, however, are imperfect. Mechanical valves are rigid, require chronic anticoagulant therapy, and can fail suddenly and catastrophically.^[1,2] The rigid construction can lead to perivalvular leaks, and the chronic anticoagulation is associated with cumulative morbidity and mortality.^[3] The rates of anticoagulant-related morbidity can be as high as 5% per year,^[4] essentially guaranteeing an untoward event within 20 years. Animal tissue valves or bioprostheses, on the other hand, are flexible and do not require anticoagulation. Unfortunately, they have poor long-term durability. The bioprostheses fail through calcification and rupture of the valve cusps,^[5] and eventually need to be replaced. Reoperation to replace a failed bioprosthetic valve, however, carries significant operative risk. When the chest is re-opened to access the heart, the adhesions between heart and thoracic cavity can cause rupture of the heart. Furthermore, removing a failed bioprosthesis is very time-consuming due to considerable tissue ingrowth into the sewing ring. This requires additional time on cardiac bypass and increases the chance of complications. Because the operative mortality to replace a failed bioprosthetic valve can be as high as 20% for the elderly,^[6,7] bioprostheses are intended to be implanted only once and should last the life of the patient. Because they last only 10 to 12 years,^[3,8-10] they are used only in the elderly who are not expected to outlive their bioprosthetic valve. However, due to their excellent short-term function and lack of long-term morbidity, bioprostheses are considered by many to be the ideal replacement heart valve. Surgeons would prefer to implant a tissue valve, but must resort to a mechanical prosthesis if patients are too young and at risk of outliving their bioprostheses. If it wasn't for their poor long-term durability, bioprosthetic valves could be used in practically all patients. At present, the

most durable of the tissue valves appears to be the allograft, a transplanted human aortic valve.^[11] Because donor human valves are in short supply, however, the majority of the tissue valves implanted are made from glutaraldehyde-tanned porcine aortic valves or bovine pericardium. Indeed, the bioprosthesis is making a strong comeback in its popularity relative to the mechanical valve. Surgeons and patients alike are becoming less tolerant of accepting the complications of anticoagulant therapy and in the United States today, more tissue valves than mechanical valves are implanted every year. Tissue valves have been recognized as the valve of choice in the United States for the majority of cardiac valve patients, in spite of their limited durability.

T

THE CURRENT PROBLEM

A more critical need for tissue valves exists for children and young adults. A large fraction of valve replacement occurs for congenitally bicuspid aortic valves that need to be replaced at an early age.^[12] Unfortunately, the implanted bioprostheses calcify in a few years in young adults, and in as little as a few months in children.^[13] A conventional bioprosthesis is therefore not an option for these patients; they must receive a mechanical valve and be placed on permanent anticoagulant therapy. Warfarin therapy requires a strict protocol of daily oral doses and periodic testing of blood coagulation at the hospital. Side effects include bruising and uncontrolled bleeding and even strokes and infarcts. This seriously decreases the quality of life of active people. Warfarin is also teratogenic and unsuitable for women in child-bearing years. These patients must therefore be on daily injections of heparin and face major complications during labor. For this smaller but very significant subpopulation of patients, current prosthetic valve solutions are simply not acceptable. There is therefore a critical need to develop better, more long-lived tissue valves that do not succumb to early degenerative failure.

Mechanisms of Bioprosthetic Valve Failure

Bioprosthetic heart valves degenerate through calcification and mechanical damage. Large calcific growths

can impede leaflet motion causing stenosis, but often result in leaflet tearing and hence massive regurgitation. For many years, calcification has been promoted as the main mechanism of valvular degeneration, with tearing thought to be secondary to calcification.^[1] Opinion has recently shifted toward recognizing that mechanical damage and calcification can be two separate processes, each with different initiators.^[5] In the current generation of porcine bioprosthetic valves, calcific degeneration appears to be a relatively minor issue, with the majority of the damage occurring strictly for mechanical reasons.^[14,15]

Role of Mechanical Damage in Calcification

Pathologic examinations of a large number of prosthetic valves with torn leaflets have shown that free-edge tears occur frequently in the absence of calcification^[16–18] in the areas that experience highly localized mechanical forces. Therefore, a purely mechanical mode of failure must also be responsible for a large percentage of bioprosthetic valve failures. The prosthetic valve cusps experience high transient stresses during valve closure, and cyclic flexural and compressive stresses during valve opening. Although it was initially believed that tensile stresses induced during valve closure were responsible for cuspal tearing, studies have shown that tissue rupture is related to flexural fatigue^[14,16,19] and the presence of compressive stresses on the inner surface of the bent tissue.^[20] Others have theorized that the cyclic compressive stressing of the valve cusps disrupts the collagen architecture^[21] and exposes calcium-binding sites believed to exist on the collagen molecule.^[22,23] This stimulates calcification, disrupts the material, and may promote further calcification.^[24,25]

Mechanical Basis for Valvular Degeneration

The relevance of mechanical degeneration to tissue engineering is the experience with aortic valve allografts. These valves are harvested from cadavers and

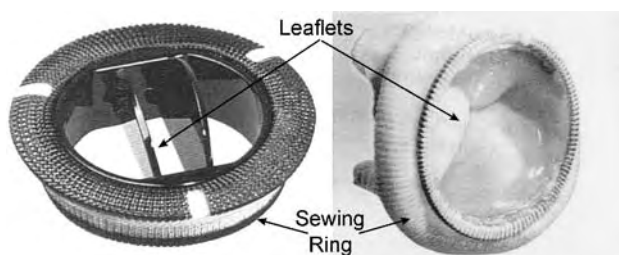


Fig. 1 Images of a typical mechanical valve (left) and a porcine bioprosthetic valve (right), also called a xenograft.

implanted into patients without any chemical preservation. These valves do not contain viable cells in the long term^[26] and therefore do not remodel, yet remarkably they do not disintegrate and fail rapidly. Aortic valve allografts, in fact, last longer in most patients than do glutaraldehyde-preserved xenografts, and are considered the optimal biologic valve.^[11] Not surprisingly, the mechanical properties of allograft valves are nearly identical to those of the native aortic valve.^[27,28] Their success as functional, load-bearing, biomechanical devices results not from chemical stabilization like the bioprostheses nor from an ability to repair and remodel like other tissues in the body, but rather from their inherent structural strength, durability, and fatigue resistance. Allograft valves are therefore durable only because they have the appropriate microstructure and mechanical properties.

AORTIC VALVE CUSP MICROSTRUCTURE AND MICROMECHANICS

Aortic valve cusps consist of three layers of morphologically distinct tissue: the fibrosa, spongiosa, and ventricularis^[29] (Fig. 2). Toward the aortic surface is the fibrosa, consisting mainly of collagen fiber bundles arranged primarily in a circumferential direction; that is, running from commissure to commissure. Starting at one commissure, these fiber bundles spread out into a somewhat isotropic mesh near the belly of the cusp, then combine again into clearly visible bundles toward the opposite commissure (Fig. 3). The ventricularis consists mainly of sheet elastin with some collagen arranged in a highly folded manner. It is much thinner than the fibrosa and does not have the rough surface topography. Separating the fibrosa and ventricularis is the very loosely organized spongiosa, consisting of collagen, elastin, and glycosaminoglycans (GAGs). The GAGs bind water readily and give the spongiosa a gelatinous, watery consistency. The specific function of the spongiosa is not yet well understood but is believed to be a buffer zone that enables the localized movement and shearing between the fibrosa and the ventricularis during loading and unloading. The spongiosa may also act as a reservoir for GAGs, which impart viscoelasticity to the valve material. The ridges on the fibrosal surface of the leaflet are not simply the relief formed by large collagen fiber bundles, but rather the radial folding of the fibrosa.^[30] When separated from the ventricularis, the fibrosa spontaneously flattens out and lengthens by 30%. The corrugations in the compressed fibrosa are maintained by virtue of its attachment to the ventricularis, which is under tension.^[31] When this tension is relieved, the ventricularis shrinks by 10%.^[32] The complex morphologic differentiation

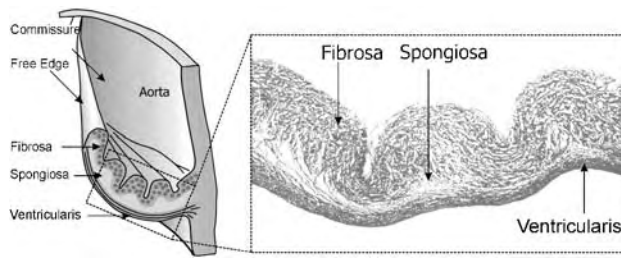


Fig. 2 Diagram of a cutaway through the central portion of a valve cusp (left), and a histologic image (right) showing the three major layers.

between the principal layers of the valve cusps suggests that the fibrosa, spongiosa, and the ventricularis perform specific functions.

Indeed, the mechanical properties of the fibrosa and the ventricularis are quite different. When the aortic valve cusps were carefully dissected to separate the layers and the fibrosa and ventricularis tested separately, it was found that the fibrosa has a much lower radial extensibility than the ventricularis.^[30] This was indeed surprising because the two layers extend simultaneously during valve closure. We subsequently determined that the fibrosa is so extensible in situ (up to 60% strain) because of its naturally folded configuration. Because the fibrosa elongates following separation from the ventricularis, its loading curve is missing the long, unfolding region that it would normally exhibit. Similarly, the ventricularis appears more extensible when tested alone, because it is no longer preloaded in tension. Subsequent studies have reconstructed the load transfer that occurs from the ventricularis to the fibrosa during valve closure and confirm that the unique mechanical properties of the aortic valve cusps result in part from the multilayered nature of the cusp and the folded configuration of the fibrosa.^[33]

Aortic Valve Collagen

As a whole, the aortic valve cusp is made up of 55% collagen and 5% to 13% elastin,^[34,35] with the remainder consisting of interstitial macromolecules, including proteoglycans. Aortic valve collagen has been studied extensively over the years.^[36-41] The aortic valve contains mainly type I collagen (about 3/4 by content) with some type III collagen (about 1/4). Type V collagen has also been found in the cardiac valves; no doubt, these tissues contain many of the other newly discovered lesser collagens that are found in many different types of connective tissues.

Collagen fiber bundles are the main load-bearing components of the aortic valve. The collagen fibers

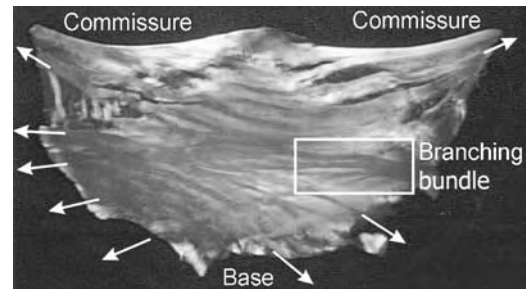


Fig. 3 Aortic valve cusp transilluminated with polarized light, demonstrating the pattern of collagen fiber bundles. The arrows denote the lines of force along these bundles.

bear the membrane stresses in the belly of the cusp, collect it, and direct it to the aortic wall by way of branching bundles, as seen in Fig. 3. The tension on these large collagen chords is substantial and actually deforms the aortic root during valve closure.^[42] In the belly of the cusp there exists an almost isotropic mesh of collagen that is held corrugated by its attachment to the ventricularis. The collagenous structures within the belly of the cusp are therefore under axial compression by way of their connection to the elastin structures, most of them in the ventricularis.^[33] These connections between the fibrosa and the ventricularis are substantial and pass through the spongiosa (Fig. 4).

In order to collect and channel the stress from the belly of the cusp, the collagen fibers are highly branched, having a fairly consistent pattern (Fig. 5). In Table 1, we see that most of the fiber bundles are either primary or secondary, with some tertiary. In the second table, however, we see that almost half of the bundles contain no bifurcations and that a roughly equal number of primary bundles that contain both secondary and tertiary branches. Roughly half the cusp, therefore, has bundles that run from attachment

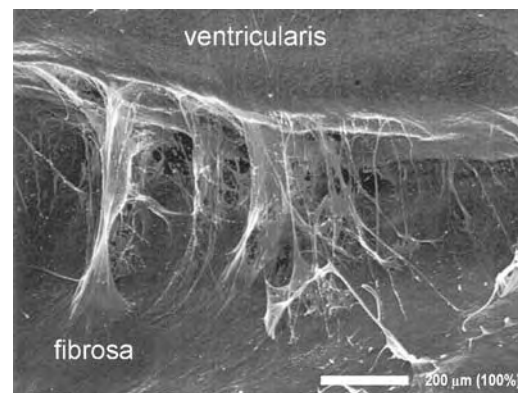


Fig. 4 Detail of elastin fibers joining the fibrosa to the ventricularis.

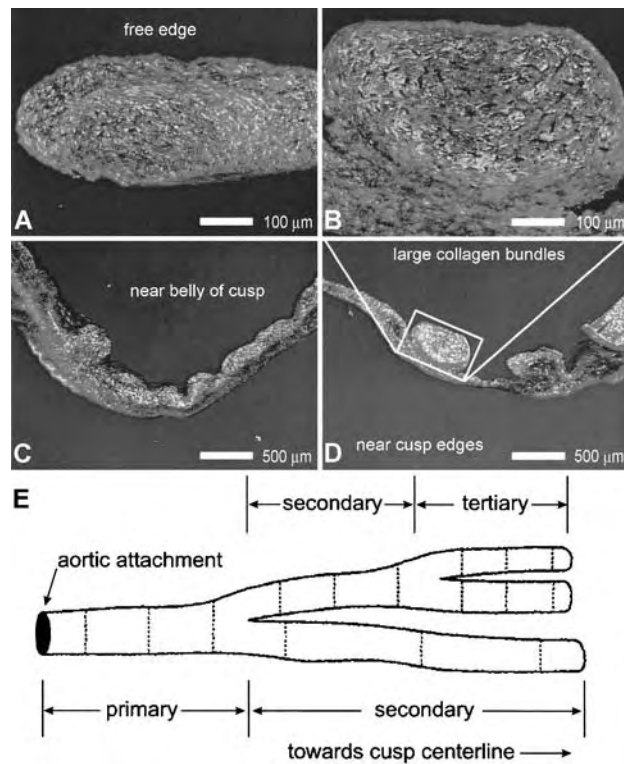


Fig. 5 Details of collagen fiber bundle morphology. In the belly of the cusp, the fibers are not pronounced (C). As the fibers progress toward the edges, they become larger and raised from the surface (B,D). The free edge of the cusp is very dense (A). Morphologic measurements (E) demonstrate a consistent pattern of equally bifurcating collagen fiber bundles.

to attachment to the aorta without any major branching. These data therefore provide some key design features for a composite, tissue-engineered aortic valve cusp.

Aortic Valve Elastin

The elastin component of the aortic valve is only about 1/2000th as stiff as the collagen component, and

Table 1 Collagen fiber bundle measurements

Type of bundle	Bundles per cusp	S.D.
Primary	11.6	2.9
Secondary	13.1	3.0
Tertiary	8.1	4.5
Primary—no bifurcations	5.1	2.7
Primary with secondary	3.6	1.4
Primary with sec. and tertiary	3.0	1.4

therefore does not participate in load bearing during valve closure.^[43] Elastin structures, however, do affect aortic valve mechanics, as evidenced by the preload of the fibrosa and ventricularis.^[31,33] Aortic valve elastin can be visualized through NaOH digestion and scanning electron microscopy. In the aorta, elastin occurs typically as concentric sheets, often with pronounced fenestrations.^[44] In the aortic valve, elastin structures are considerably more complex. In general, aortic valve cusp elastin occurs as continuous amorphous masses or as networks of discrete fibers. Amorphous elastin is found as a continuous sheet less than 1 μm thick, whose surface detail varies from smooth to slightly textured under high magnification, sometimes with integrated fibers visible on the surface (Fig. 6A). The other type, fibrillar elastin (Fig. 7A), is not formed into large continuous sheets, but is visible as individual fibers organized into structures ranging from loosely connected individual elements to complex, intricate meshes. Individual fibers can be as large as 5 μm in diameter, but most are less than 1 μm across. The sheet elastin often transitions to fibrillar elastin, with

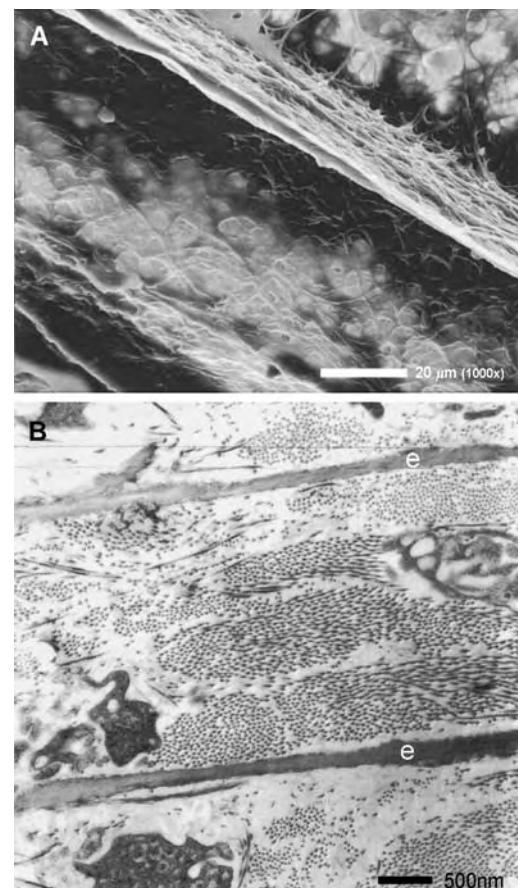


Fig. 6 Morphology of sheet elastin viewed using (A) SEM and (B) TEM. Sheets of elastin clearly run between bundles of collagen.

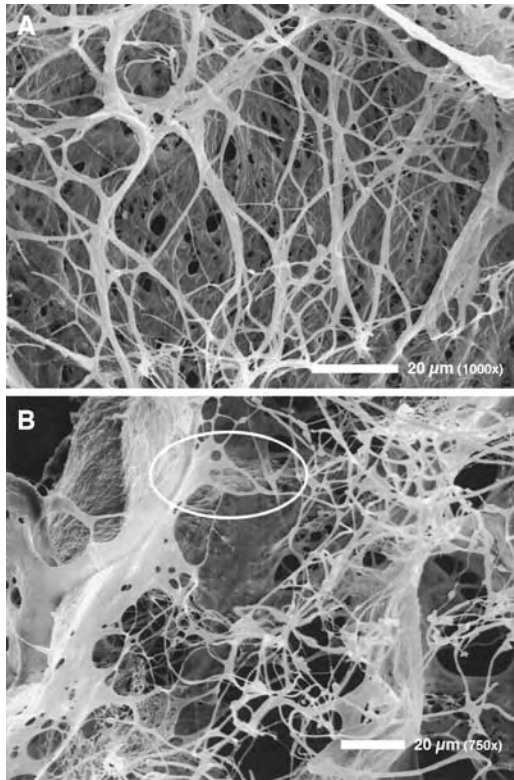


Fig. 7 (A) Typical morphology of fibrillar elastin; and (B) transitions (shown circled) from sheet to elastin fibers. The configuration of the fibers suggests that elastin bears load.

branching structures that are clearly organized to bear, channel, and distribute load between the fibrillar and sheet elastin (Fig. 7B).

Each of the three morphologically different structures—the fibrosa, spongiosa, and ventricularis—have all of these elastin types in varying amounts and conformations. The ventricularis contains a continuous sheet of dense elastin that stretches over the entire layer, becoming a fibrous mesh toward the edges. On the inside of the leaflet, elastin fibers protrude from the ventricularis, branch and anchor into the sheet like roots, and progress toward the fibrosa through the spongiosa (Fig. 7).

The elastin of the fibrosa is more delicate and complex than in the ventricularis. The fibrosa has a highly organized network of tubes that radiate out from the central region of the cusp toward the line of attachment of the cusp to the aortic wall. These tubes are typically 100 to 500 μm in diameter, and their walls are made of dense or amorphous sheet elastin (Fig. 8A). Smaller tubes up to 80 μm in diameter stretch circumferentially across the middle of the cusp, are packed densely together parallel to each other, and are oriented from commissure to commissure.^[35] Figure 8B is a schematic diagram of the elastin tube structures in the fibrosa.

We now know that elastin is not primarily confined to the ventricularis, as suggested previously,^[45,46] but is widely distributed throughout the aortic valve in what appear to be highly functional elements. In the ventricularis, the substantial amount of sheet elastin generates the preload that holds the fibrosa in its wrinkled configuration (see Fig. 2). Within the fibrosa, the most striking feature is the presence of directional tubes, the largest collection located around the line of attachment of the cusp to the aorta, with the tubes directed toward the center of the cusp. TEM images have shown that these tubes indeed surround the large collagen fiber bundles that emerge from the valve cusp and anchor into the aortic wall. These fiber bundles follow the lines of principal stress and move and respond to stress as the cusp is loaded. As the valve is loaded in systole, the cusps extend radially by over 30% on average,^[47] and considerably more in certain areas. Such radial extension is made possible by an unraveling of the corrugations in the central part of the fibrosa and by the pulling apart of the collagen fiber bundles. The elastin network of the cusp—particularly the struts between the elastin tubes (Fig. 8A)—therefore stores energy during diastolic loading, and returns it to the collagen network in systole when the cusp is

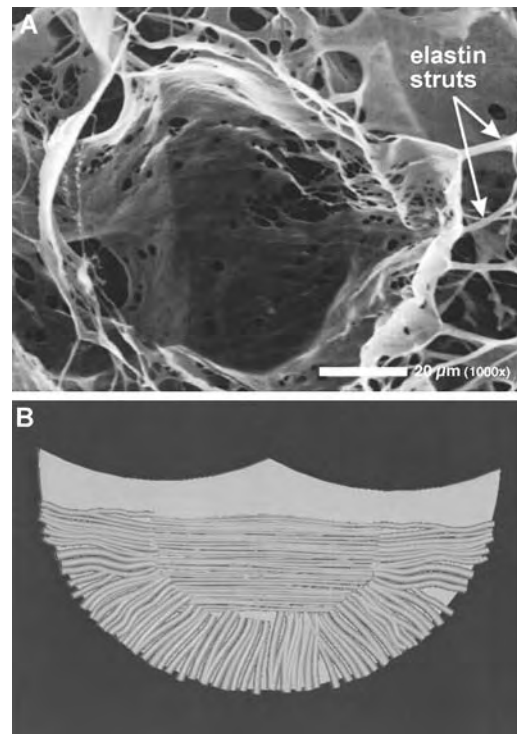


Fig. 8 Typical morphology of elastin tubes from the central area of the fibrosa. (A) These tubes are connected by elastin struts and arranged so that they radiate outward from the leaflet. (B) Each tube envelops a collagen fiber bundle and links the bundles together.

unloaded. The energy stored in the stretched elastin is therefore used to return the collagen structure to its resting configuration.

Cellular Population

Both allografts and xenografts lack a viable cell population. Over time, both allografts and xenografts appear to lose collagen fiber crimp, have a collapsed central spongiosa, and stain negatively for glycosaminoglycans. Thus, the ground substance associated with valve leaflet viscoelasticity and the overall collagen fiber morphology eventually disappear from these acellular tissues. Indeed, we have shown^[48,49] that glycosaminoglycans completely disappear from implanted bioprosthetic heart valves. Because aortic valve interstitial cells have been shown to have considerable biosynthetic activity in some animal models,^[50] it can be argued that the restoration and maintenance of a viable cell population in bioprosthetic heart valves is one way of prolonging their functional lifespan from the mechanical viewpoint. We recognize, however, that very little is known about the biology of aortic valve interstitial cells. Indeed, any new findings are very important to the field of tissue engineering. Currently, it is understood that there may be at least three subtypes of valvular interstitial cells, characterized primarily by the differential expression of marker proteins such as contractile filaments.^[51–53] How these different cells contribute to the synthesis of interstitial GAGs is presently unknown. The rapid loss of GAGs from transplanted xenograft and allograft valves, however, suggests that GAG synthesis is likely one of their prime functions.

AN ARGUMENT FOR PROPER MATRIX MECHANICS

Collagen, elastin, and GAGs are the structural components of essentially all of the body's connective tissues. What is important to the proper function of each of these diverse connective tissues is not the mere presence of their structural components—not even their correct relative amounts—but rather their functional organization and their mutual interactions. The aortic valve has nearly identical amounts of collagen and elastin as does the mitral valve, but they clearly do not function the same. A mitral valve leaflet cannot be transplanted into the aortic position and be expected to function properly in the long term. Indeed, the inability of transplanted connective tissues to function long-term in a foreign environment has been clear for many years, given the experience of the autologous pericardial valves pioneered by Dr. Jack Love.^[54]

Bovine pericardium has been used to make very successful bioprosthetic heart valves for a number of years.^[55] Their primary disadvantage, however, is that they are antigenic and need to be cross-linked with glutaraldehyde. As described above, glutaraldehyde treatment leads to calcification and to eventual valve failure. It was thought that autologous pericardium would be a better alternative to bovine pericardium, because it is not antigenic and hence does not need to be cross-linked with aldehydes. When Dr. Love began experimenting with such valves, he found that autologous pericardium rapidly contracted into dense scar tissue and the valve failed to function. Indeed, the only way that contraction could be prevented was to kill all the cells in the pericardium by a brief immersion in glutaraldehyde. The importance of killing the cells was later confirmed by Duran^[56] with alcohol-treated pericardial patches. The mechanism of this cell-mediated contraction is only now being elucidated.^[57]

The presence of cells in engineered tissues can be both beneficial and detrimental. We know that cells respond to the strains of their underlying matrix.^[58–60] Clearly, when cells populate a matrix that experiences mechanical strains different from what the cells expect, they respond in unpredictable ways. Abnormal strains lead to abnormal cell activity.^[61] When used to fabricate a prosthetic valve, living pericardium is subjected to loads very different from those imposed on it during normal function and it therefore experiences abnormal loads. Accordingly, the pericardial cells in the autologous valve do not respond in the appropriate way and are likely responsible for the rapid contraction and scarlike remodeling of the pericardial valve leaflets. By way of analogy, cells isolated from heart valves, seeded on the tissue-engineered substrates that do not induce the proper strains in the cell cytoskeleton during valve loading, will receive the wrong signals. These cells are therefore unlikely to behave properly in the tissue-engineered valves and should not be expected to remodel the degrading matrix into a functioning heart valve. Although the creation of aortic valvelike tissue has been reported in the literature,^[62–64] closer examination of the histological images by an experienced observer reveals a scarlike matrix that bears little resemblance to the finely structured matrix of the native aortic valve.

Surprisingly, the experience with acellular matrix xenograft valves has also been disappointing. A fully-formed aortic valve is perhaps the best matrix for a tissue-engineered valve. If such a matrix could be obtained from an animal, for example, and its antigenicity reduced by means other than cross-linking, it would function much like a human allograft. Indeed, if the patient's own cells were to repopulate this graft, then a hybrid-autograft/xenograft bioprosthesis could be created.^[65] This approach was first tried in vascular grafts with some success^[66] and later applied to pericardium in

an attempt to fabricate prosthetic leaflets.^[67] The approach involves first extracting the porcine cells and their membrane-associated antigens with various detergents to clean out the valve. Such a valve could then be cultured with cells from the recipient to create a hybrid living human/porcine valve. Although we pursued this approach in the early 1990s,^[65] we had little success. We were not able to fully remove all the cellular debris and still maintain proper matrix mechanics. However, Goldstein, working for CryoLife Inc., persevered and eventually succeeded in developing such a valve, which was eventually implanted into patients.^[68] Indeed, this device received CE Mark approval for use in the pulmonary position of pediatric patients in October 17, 2000. Recently, personal communications from Cryolife have confirmed that 5 patients in Australia have received this valve, and apparently all 5 have had to have the valves removed for complications resulting from uncontrolled scarring or inappropriate remodeling. Because of the proprietary nature of this technology, no published information is available regarding its rumored failure, and these reports are difficult to confirm. However, these anecdotes are consistent with the experience of Love^[54] (inappropriate matrix leads to scarring), and of others who have attempted the hybrid valve approach and were unsuccessful. Besides CryoLife, the other prominent company that attempted this was LifeCell Corp.^[69] Their approach apparently failed in sheep experiments because of highly variable fibrosis and never progressed to clinical trials. Clearly, for some unknown reason, porcine aortic valve scaffolds experience variable fibrosis and do not support a stable population of recipient cells once implanted.

Further discouraging the use of the decellularized scaffolds are findings that human aortic valve allografts do not naturally repopulate. In the aortic valve allograft, there exists perhaps the most perfect matrix for recolonization by host cells. Yet in a matter of months, these valves become totally acellular.^[26,70] The viable host cells quickly die and recipient cells do not repopulate this seemingly ideal matrix. Over time, the allograft valves fail from progressive mechanical damage, possibly mediated by a low-grade foreign-body response.^[71,72]

There is no question that both biodegradable matrices and acellular matrix xenografts are exciting technologies that may eventually be made to work. A third alternative to fabricating a tissue-engineered aortic valve has been pursued in our laboratory.

THE BIOMIMETIC CONSTRUCT

Our concept for a tissue-engineered heart valve is to build up a permanent, nonbiodegradable, composite

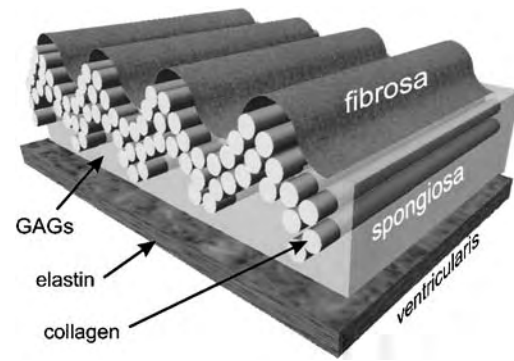


Fig. 9 Schematic diagram showing the arrangement of collagen, elastin, and GAGs within the valve cusp.

structure from biological molecules, one that closely mimics the structure and function of the native aortic valve. We believe that for the aortic valve, remodeling of the matrix is not essential. Indeed, the highly cross-linked nature of native aortic valve collagen indicates that it is very mature, very old, and does not remodel. Our new understanding of the role of valvular interstitial cells, as producers of glycosaminoglycans, suggests that the main function of these cells is to replenish the GAG stores that are rapidly depleted during cyclic loading.^[48,49,73] A biomimetic construct modeled along the lines of the schematic shown in Fig. 9 would be fabricated in the laboratory using conventional engineering approaches, would be ready to bear cardiac loads immediately upon implantation, would not need to remodel, would respond mechanically to loads in a manner identical to that of native valves, and would thus stimulate the valvular interstitial cells in a manner that is much more natural than the biodegradable matrices. Such a construct would therefore have a lower risk of succumbing to the nonspecific scarring and fibrosis that have affected current versions of valvular scaffolds. We acknowledge, however, that cellular response to strains in the underlying matrix is very much a function of the cells selected for the construct. At this stage, we are focusing on the structure and function of the matrix only, choosing cells that will be most suitable for the fabrication of the matrix. The choice of which cells to use for constructs that are to be implanted into animal models will be made later, when the engineering parameters of our approach are finalized. Part of the ongoing problem with tissue engineering of heart valves is that cardiovascular tissues are far more complex than initially envisioned and have a capacity for self-repair that is far less effective than tissue-engineering principles demand.

Our approach to developing a tissue-engineered aortic valve is not to rely on remodeling of a degrading scaffold, but rather to fabricate the cusp from its basic

building blocks: 1) branching collagen fiber bundles; 2) elastin sheets and tubes; and 3) a highly hydrated GAG matrix. Following are descriptions in greater detail of these specific technologies.

Collagen Fiber Bundles

Our approach to fabricating collagen fiber bundles is based on the principle of directed collagen gel shrinkage.^[74–77] The principle involves first mixing soluble, fibrillar collagen with cells; neutralizing the mixture to induce gelation; and maintaining the system in culture. When the system gels, cells become entrapped within the collagen gel and begin to interact with the collagen fibrils.^[78–83] These cells reorganize the surrounding collagen matrix, contracting it and excluding water. In many ways, this in-vitro contraction mimics wound healing in vivo.^[84] When the gel is mechanically constrained from shrinking in one direction, the collagen fibrils align in the direction of constraint.^[83,85] Highly-aligned, compacted, collagenous constructs can thus be fabricated.

First, neonatal rat aortic smooth muscle cells (SMC) were harvested by a method outlined by Oakes et al.^[86] to establish primary culture. Once these cells are amplified, they are detached from the culture dishes using trypsinization, counted, centrifuged, and added to the collagen suspension at a cell-seeding concentration of 1.0 million cells/ml. The collagen suspension consists of sterile acid-soluble type I collagen (BD Biosciences, Rat tail; 3.94 mg/ml, 0.02 N acetic acid) at an initial concentration of 2.0 mg/ml. The cell/collagen suspension is then pipetted into the wells and incubated at 37°C. Within several hours, the collagen gel forms and attaches to the microporous holders at the ends of the wells. These anchors prevent longitudinal contraction and allow shrinkage to occur only transverse to the long axis of the wells.^[87]

After 8 weeks of culture, the original transparent gel becomes a dense, cylindrical construct with well-aligned collagen fibrils (Fig. 10). The final constructs have the typical nonlinear stress/strain curve of tendonous materials, an extensibility of 14%, a stiffness of 5 MPa and failure strength of 1.1 MPa. Although the stiffness and strength are still lower than what is required, our constructs are already 10 to 100 times stronger than similar collagen-based materials fabricated previously.^[88,89] Ultrastructural analyses have shown that the main reason for the good strength of our constructs is the very high collagen fibril density. Because the constructs are relatively simple, one-dimensional collagen bundles, they compact from two directions, producing an area shrinkage ratio that is greater than 99% (from an area of 324 mm² to less than 1 mm²). When fully compacted, the collagen-fiber

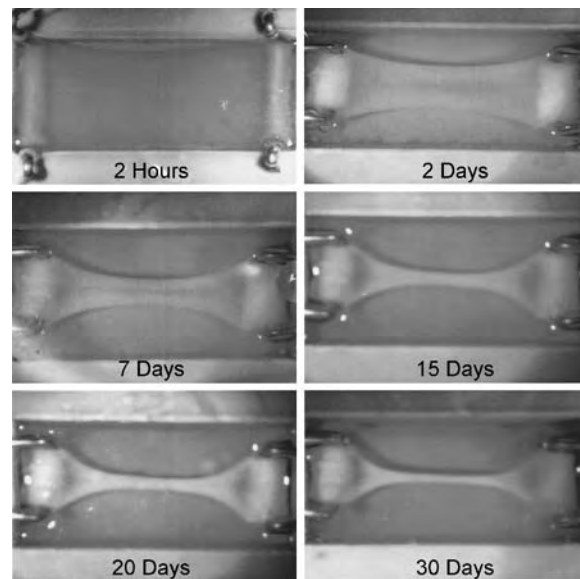


Fig. 10 Images showing how collagen gels attached to holders and shrank over time to form thin, dense collagen fiber bundles.

density visually approaches that of mitral valve chordae (Fig. 11). The obvious differences are less collagen fiber crimp than is present in most collagenous structures and lower failure strength, possibly because of less cross-linking of collagen fibrils by proteoglycans, or other matrix molecules.

Because collagen fiber bundles in heart valves are branched, we have also attempted to fabricate branched structures, with good results (Fig. 12). The key to these geometries is to ensure that tension is properly controlled during the shrinkage process so that the constructs do not tear away from the holders during the early stages. As the images demonstrate, different branching angles and parent/daughter ratios can be engineered with the use of appropriate molds.

Cross-Linked Hyaluronan

Hyaluronan is a glycosaminoglycan polymer with a repeating disaccharide structure (glucuronic acid- β 1, 3-N-acetylgalactosamine- β 1,4-)_n, where *n* can reach 25,000 or more. In solution, hyaluronan forms large, random coil structures, which occupy large solvent volumes. When constrained within a matrix—for example, within a collagen network—hyaluronan exerts a swelling pressure on the matrix that traps water, gives tissues compressive resistance, and imparts viscoelastic properties. Hyaluronan also plays a role during embryonic cardiac development, as invading cells transform cardiac jelly into specialized cardiac structures, like the myocardium and the cardiac

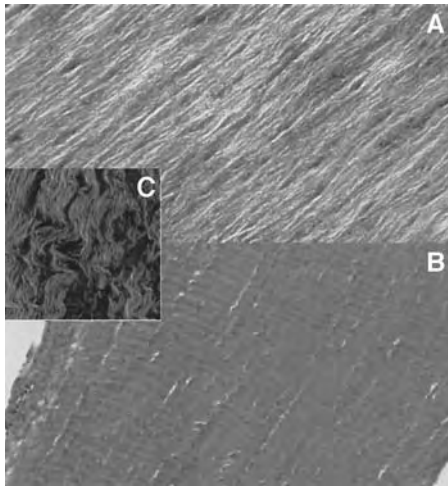


Fig. 11 Typical morphology of collagen fibers (A) in the constructs and (B) in mitral valve chordae. Note the high density of fibers in the constructs. The inset (C) shows that constructs with crimped collagen can be fabricated.

valves.^[90] Hyaluronan is exceptionally biocompatible.^[91–93] Unlike collagen, which has both tissue- and species-specific markers, hyaluronan exhibits structural homology across species. For example, hyaluronan made by bacteria is the same as synthetic hyaluronan.^[94] Because of its interesting viscoelastic properties and its broad biocompatibility, hyaluronan has been used in a number of clinical applications. Perhaps the greatest contribution to bringing hyaluronan to commercial use has been made by Dr. Endre Balazs, the founder of Biomatrix Inc. (Ridgefield, N.J.), now owned by Genzyme. Dr. Balazs pioneered the cross-linking of hyaluronan with divinyl sulfone^[95] and began using it for viscosurgery of the eye.^[96] Other approaches to cross-linking hyaluronan have been attempted, including reactions involving hydrazides,^[97] carbodiimides,^[98] and esters.^[99] However, some problems encountered with the use of these cross-linkers include the extreme conditions needed for inducing cross-linking, the potential toxicity of the gelled product or of the cross-linker, and the inability of the cross-linking chemistry to accommodate desired characteristics of the gels, such as mechanical characteristics similar to native tissues and the ability to support cell attachment. We found none of these problems with the use of divinyl sulfone, and thus have based our hylan work on it. One important advantage of divinyl sulfone is that it utilizes the OH groups on the glycosaminoglycan chains, thereby leaving the more biologically active groups (carboxylate and N-acetyl groups) unaltered. The new molecules therefore have high biocompatibility and unaltered biological activity. It should also be noted that even though divinyl

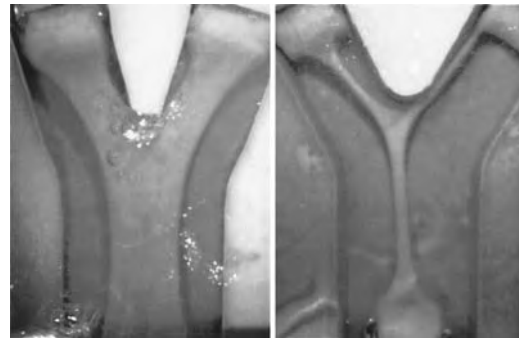


Fig. 12 Images showing branched constructs at 1 day and at 8 weeks.

sulfone is extremely toxic in its unreacted state, the reaction products (bis-ethyl-sulfone crosslinks) are nontoxic and nonleachable when incorporated as a cross-linker within hylan gels.^[100] Hylan gels incorporating such cross-linked chemistry are approved in a number of product formulations for long-term clinical use, including injections into the knee for the treatment of osteoarthritis or under the skin to correct facial defects.^[96] In the long term, divinyl sulfone-crosslinked hyaluronan persists in the body unaltered. It does not elicit a foreign-body reaction, and it does not calcify. These are features very desirable for use in cardiac valve prostheses.

We fabricated hylan gels using the method of Balazs,^[95] cast them onto the bottom of 24-well tissue culture plates, and coated them with cell-binding matrix factors such as collagen (type I, from rat tail), extracellular matrix (ECM) gel (from mouse sarcoma), laminin, and fibronectin (from bovine plasma). Relative to uncoated controls, these matrix-coated gels induced much better cell attachment and proliferation.^[101] Because particulate gels induced greater cell attachment, regardless of the amount of adhesion protein bound, we sought to manipulate the texture of the hylan gels to maximize cell attachment and cell penetration.

Previously, several researchers used low-intensity ultraviolet radiation as a cross-linking agent in the preparation of natural and synthetic biomedical polymers. Bryant et al.^[102] showed that PVA scaffolds exposed to low-intensity ultraviolet radiation are cytocompatible. Ultraviolet irradiation was also shown to enhance the attachment and chemotaxis of melanocytes on extracellular matrix proteins.^[103] We therefore subjected DVS-cross-linked hydrogels to UV light and carried out similar cell adhesion studies.

As in the previous study, neonatal rat aortic smooth muscle cells did not attach to unmodified hylan gels. They attached sparsely to oven-dehydrated gels and were rounded and nonproliferative. However, cells attached readily to irradiated gels, appeared healthy

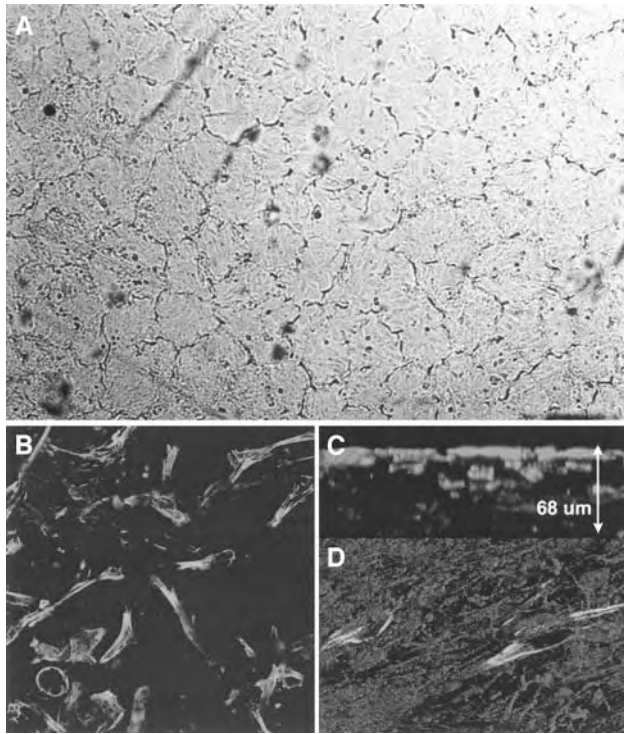


Fig. 13 (A) Light microscopy image of the surface of a UV-irradiated gel showing the cracks; (B) confocal images showing a mid-section; (C) transverse z -section of a reconstructed stack of images; and (D) section showing the penetration of cells into the interior. In this image, hyaluronan fluoresces red.

and highly proliferative, and formed multiple confluent layers on the surface within 10 days of seeding.^[101] Cells also migrated into the interior of the UV-exposed gels. Confocal microscopy indicated that in many cases, the cells populated the entire thickness of the gels (Fig. 13C). Cells within the interior appeared highly extended, often as long as 75 micrometers. Cells appeared to migrate into the interior of gels via cracks in the gel that were visible with light microscopy (Fig. 13A). Nonirradiated gels did not demonstrate this topography. It is also noteworthy that these high levels of adhesion and proliferation were achieved without any precoating with matrix factors. Dehydration and UV irradiation were all that was necessary to change the hylans from an inert, watery gel to an elastic substrate that readily supports cellular adhesion and infiltration.

Elastin Structures

As previously mentioned, above, elastin structures are extremely important in the mechanical function of the aortic valve cusp. They exist as fibers, sheets, and tubes

that link together the major load-bearing collagenous structures. Accordingly, we have explored the means by which such elastin structures could be fabricated in the laboratory. We have been quite successful in growing elastin sheets atop hylan gels and around collagen fiber bundles, and have thus created large elastin sheets and elastin tubes (Figs. 14 and 15). These sheets of elastin were isolated through NaOH digestion and imaged with SEM. In most areas, they were solid and continuous (Fig. 15B), and in others, the sheets had fenestrations, typical of those seen in aorta and in the ventricularis of the aortic valve (Fig. 15C). Although we currently do not know why an elastin sheath would form around our collagen constructs in vitro, we speculate that the greater cell density or

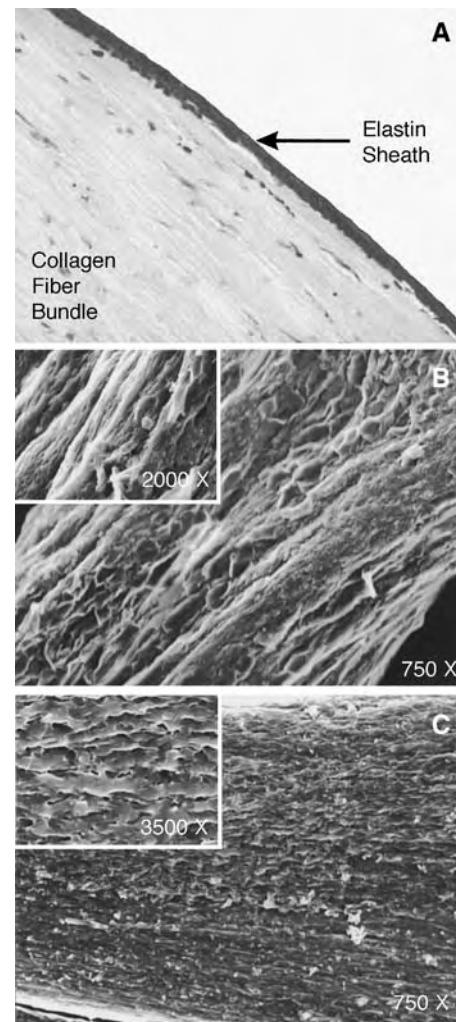


Fig. 14 (A) Histologic section of a collagen fiber bundle construct showing the elastin sheath that has formed on its surface, (B) SEM images of NaOH-digested samples of constructs; and (C) native chordae showing the surface topology of the elastin sheaths.

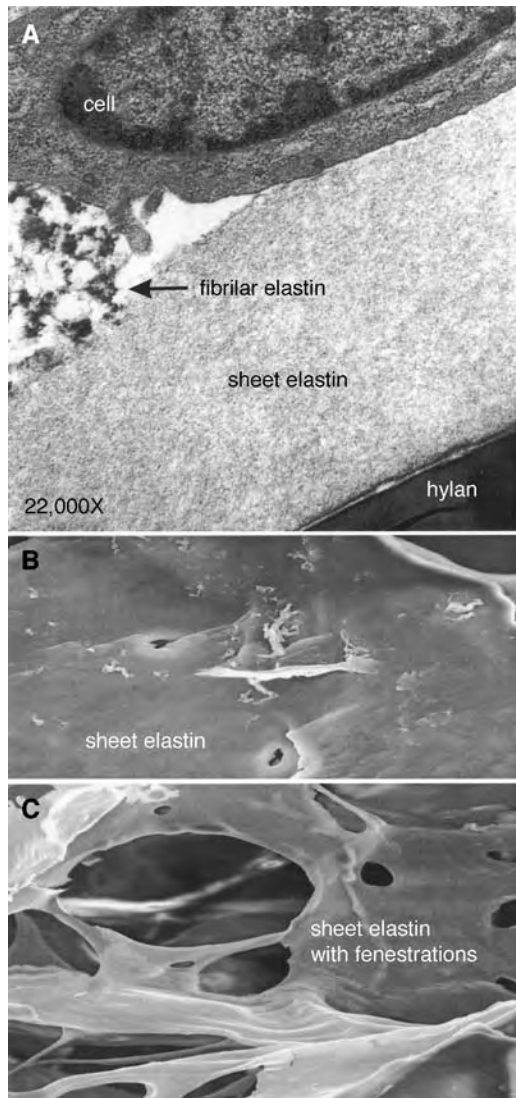


Fig. 15 (A) TEM image of a cell secreting fibrillar elastin adjacent to sheet elastin that has formed on UV-irradiated hylan gels, (B) SEM images of smooth and (C) fenestrated sheet elastin isolated through digestion in NaOH.

greater tension induced at the surface of the construct may have some role in the formation of the outer elastin sheath.

SUMMARY

Although we have not yet succeeded at assembling all of these structural components, we have established that such a structure can at least be developed in vitro. Our approach is very different from the traditional tissue-engineering approach, in which cells are seeded onto a biodegradable scaffold and a functional organ

is expected to emerge as the scaffold degrades. The advantage of our proposed technique is that it will rely less on the remodeling capabilities of living tissues and more on conventional engineering methodologies. As such, this approach is likely to be more predictable in fabricating the desired types of tissues.

It is unclear which of the three approaches to tissue engineering of heart valves will have success first. Certainly, the road thus far has been longer than initially anticipated. Maybe the true benefit of the so-called tissue-engineered revolution is that researchers will finally look at native heart valve tissue and native cells and begin to study the system with heart valve disease in mind. Maybe the therapy of the future will be not the replacement of diseased valves with tissue-engineered prostheses, but rather the treatment of valvular disease with cells and molecules to slow down and potentially even reverse the course of the disease itself, thus avoiding a prosthesis altogether.

ARTICLES OF FURTHER INTEREST

Collagen; Elastin; Heart Valve, Bioprosthetic; Heart Valve Failure, Bioprosthetic; Heart Valves, Mechanical; Hyaluronan; Tissue Engineering of Cardiac Muscle; Tissue Engineering Scaffolds

REFERENCES

1. Schoen, F.J. Cardiac valve prostheses: Review of clinical status and contemporary biomaterials issues. *J. Biomed. Mater. Res.* **1987**, *21* (A1 Suppl), 91–117.
2. Gonzalez-Lavin, L.; et al. Strut fracture and other events after valve replacement with the 60 degree convexoconcave Bjork–Shiley prosthesis. *Circulation* **1987**, *76* (3 Pt 2), III137–III140.
3. Hammermeister, K.E.; et al. A comparison of outcomes in men 11 years after heart-valve replacement with a mechanical valve or bioprosthesis. Veterans affairs cooperative study on valvular heart disease [see comments]. *N. Engl. J. Med.* **1993**, *328* (18), 1289–1296.
4. Jamieson, W.R.; et al. Multiple mechanical valve replacement surgery comparison of St. Jude Medical and CarboMedics prostheses. *Eur. J. Cardio-thorac. Surg.* **1998**, *13* (2), 151–159.
5. Schoen, F.J.; Levy, R.J. Tissue heart valves: Current challenges and future research perspectives. *J. Biomed. Mater. Res.* **1999**, *47* (4), 439–465.
6. Biglioli, P.; et al. Reoperative cardiac valve surgery: A multivariable analysis of risk factors. *Cardiovasc. Surg.* **1994**, *2* (2), 216–222.

7. McGrath, L.B.; et al. Perioperative events in patients with failed mechanical and bioprosthetic valves. *Ann. Thorac. Surg.* **1995**, *60* (2 Suppl), S475–S478.
8. Bernal, J.M.; et al. Valve-related complications with the Hancock I porcine bioprosthesis. A twelve- to fourteen-year follow-up study. *J. Thorac. Cardiovasc. Surg.* **1991**, *101* (5), 871–880.
9. Burdon, T.A.; et al. Durability of porcine valves at fifteen years in a representative North American patient population. *J. Thorac. Cardiovasc. Surg.* **1992**, *103* (2), 238–251. Discussion 251–252.
10. Grunkemeier, G.L.; Bodnar, E. Comparative assessment of bioprosthesis durability in the aortic position. *J. Heart Valve Dis.* **1995**, *4* (1), 49–55.
11. O'Brien, M.F.; et al. The homograft aortic valve: A 29-year, 99.3% follow up of 1,022 valve replacements. *J. Heart Valve Dis.* **2001**, *10* (3), 334–344. Discussion 335.
12. Sabet, H.Y.; et al. Congenitally bicuspid aortic valves: A surgical pathology study of 542 cases (1991 through 1996) and a literature review of 2,715 additional cases. *Mayo Clin. Proc.* **1999**, *74* (1), 14–26.
13. Dunn, J.M. Porcine valve durability in children. *Ann. Thorac. Surg.* **1981**, *32*, 357–368.
14. Vesely, I.; Barber, J.E.; Ratliff, N.B. Tissue damage and calcification may be independent mechanisms of bioprosthetic heart valve failure. *J. Heart Valve Dis.* **2001**, *10* (4), 471–477.
15. Sacks, M.S.; Schoen, F.J. Collagen fiber disruption occurs independent of calcification in clinically explanted bioprosthetic heart valves. *J. Biomed. Mater. Res.* **2002**, *62* (3), 359–371.
16. Ishihara, T.; et al. Structure and classification of cuspal tears and perforations in porcine bioprosthetic cardiac valves implanted in patients. *Am. J. Cardiol.* **1981**, *48*, 665–678.
17. Pomar, J.L.; et al. Late tears in leaflets of porcine bioprostheses in adults. *Ann. Thorac. Surg.* **1984**, *37* (1), 78–83.
18. Stein, P.D.; et al. Relation of calcification to torn leaflets of spontaneously degenerated porcine bioprosthetic valves. *Ann. Thorac. Surg.* **1985**, *40* (2), 175–180.
19. Vesely, I.; Boughner, D.; Song, T. Tissue buckling as a mechanism of bioprosthetic valve failure. *Ann. Thorac. Surg.* **1988**, *46* (3), 302–308.
20. Krucinski, S.; et al. Numerical simulation of leaflet flexure in bioprosthetic valves mounted on rigid and expansile stents. *J. Biomech.* **1993**, *26* (8), 929–943.
21. Thubrikar, M.J.; et al. Role of mechanical stress in calcification of aortic bioprosthetic valves. *J. Thorac. Cardiovasc. Surg.* **1983**, *86*, 115–125.
22. Schoen, F.J.; Tsao, J.W.; Levy, R.J. Calcification of bovine pericardium used in cardiac valve bioprostheses. Implications for the mechanisms of bioprosthetic tissue mineralization. *Am. J. Pathol.* **1986**, *123* (1), 134–145.
23. Ferrans, V.J.; et al. Calcific deposits in porcine bioprostheses. Structure and pathogenesis. *Am. J. Cardiol.* **1980**, *46*, 721–734.
24. Gabbay, S.; et al. Do heart valve bioprostheses degenerate for metabolic or mechanical reasons? *J. Thorac. Cardiovasc. Surg.* **1988**, *95* (2), 208–215.
25. Sabbah, H.N.; Hamid, M.S.; Stein, P.D. Mechanical stresses on closed cusps of porcine bioprosthetic valves: Correlation with sites of calcification. *Ann. Thorac. Surg.* **1986**, *42* (1), 93–96.
26. Schoen, F.J.; Mitchell, R.N.; Jonas, R.A. Pathological considerations in cryopreserved allograft heart valves. *J. Heart Valve Dis.* **1995**, *4* (Suppl 1), S72–S75. Discussion S75–S76.
27. Vesely, I.; et al. Mechanical testing of cryopreserved aortic allografts. Comparison with xenografts and fresh tissue. *J. Thorac. Cardiovasc. Surg.* **1990**, *99* (1), 119–123.
28. Vesely, I.; Casarotto, D.C.; Gerosa, G. Mechanics of cryopreserved aortic and pulmonary homografts. *J. Heart Valve Dis.* **2000**, *9* (1), 27–37.
29. Gross, L.; Kugel, M.A. Topographic anatomy and histology of the valves in the human heart. *Am. J. Pathol.* **1931**, *7*, 445–456.
30. Vesely, I.; Noseworthy, R. Micromechanics of the fibrosa and the ventricularis in aortic valve leaflets. *J. Biomech.* **1992**, *25* (1), 101–113.
31. Vesely, I.; Lozon, A. Natural preload of aortic valve leaflet components during glutaraldehyde fixation: Effects on tissue mechanics. *J. Biomech.* **1993**, *26* (2), 121–131.
32. Vesely, I.; Lozon, A.; Talman, E. Is zero-pressure fixation of bioprosthetic valves truly stress free? *J. Thorac. Cardiovasc. Surg.* **1993**, *106* (2), 288–298.
33. Vesely, I. Reconstruction of loads in the fibrosa and ventricularis of porcine aortic valves. *ASAIO J.* **1996**, *42* (5), M739–M746.
34. Bashey, R.I.; Torii, S.; Angrist, A. Age-related collagen and elastin content of human heart valves. *J. Gerontol.* **1967**, *22* (2), 203–208.
35. Scott, M.J.; Vesely, I. Morphology of porcine aortic valve cusp elastin. *J. Heart Valve Dis.* **1996**, *5* (5), 464–471.
36. Kanke, Y.; et al. Biochemical study of cardiac valvular tissue. Biosynthesis in vitro of hexosamine-containing substances in bovine heart valve. *Biochem. J.* **1971**, *124* (1), 207–214.
37. Jimenez, S.A.; Bashey, R.I. Solubilization of bovine heart-valve collagen. *Biochem. J.* **1978**, *173* (1), 337–340.
38. Bashey, R.I.; Bashey, H.M.; Jimenez, S.A. Characterization of pepsin-solubilized bovine heart-valve collagen. *Biochem. J.* **1978**, *173* (3), 885–894.
39. Nimni, M.E. The molecular organization of collagen and its role in determining the biophysical properties of the connective tissues. *Biorheology* **1980**, *17* (1–2), 51–82.
40. Viidik, A. Interdependence Between Structure and Function in Collagenous Tissues. In *Biology of Collagen*; Academic Press, 1980; 257–280.
41. Hilbert, S.L.; Ferrans, V.J.; Swanson, W.M. Optical methods for the nondestructive evaluation of collagen morphology in bioprosthetic heart valves. *J. Biomed. Mater. Res.* **1986**, *20* (9), 1411–1421.

42. Vesely, I. Aortic root dilation prior to valve opening explained by passive hemodynamics. *J. Heart Valve Dis.* **2000**, *9* (1), 16–20.
43. Vesely, I. The role of elastin in aortic valve mechanics. *J. Biomech.* **1998**, *31* (2), 115–123.
44. Roach, M.R.; Song, S.H. Arterial elastin as seen with scanning electron microscopy: A review. *Scanning Microsc.* **1988**, *2* (2), 994–1004.
45. Broom, N.; Christie, G. The Structure/Function Relationship of Fresh and Glutaraldehyde-Fixed Aortic Valve Leaflets. In *Cardiac Bioprostheses*; Cohn, L., Galluci, V., Eds.; Yorke Medical Books: New York, 1982.
46. Clark, R.E.; Finke, E.H. Scanning and light microscopy of human aortic leaflets in stressed and relaxed states. *J. Thorac. Cardiovasc. Surg.* **1974**, *67* (5), 792–804.
47. Adamczyk, M.M.; Lee, T.C.; Vesely, I. Biaxial strain properties of elastase-digested porcine aortic valves. *J. Heart Valve Dis.* **2000**, *9* (3), 445–453.
48. Mako, W.J.; et al. Loss of Glycosaminoglycans (GAGs) From Implanted Bioprosthetic Heart Valves. In *70th Scientific Sessions of the American Heart Association*; 1997.
49. Grande-Allen, K.J.; et al. Loss of chondroitin 6-sulfate and hyaluronan accompany structural deterioration of porcine bioprosthetic valves. *J. Biomed. Mater. Res.* **2002**. accepted.
50. Mulholland, D.L.; Gotlieb, A.I. Cell biology of valvular interstitial cells. *Can. J. Cardiol.* **1996**, *12* (3), 231–236.
51. Taylor, P.M.; Allen, S.P.; Yacoub, M.H. Phenotypic and functional characterization of interstitial cells from human heart valves, pericardium and skin. *J. Heart Valve Dis.* **2000**, *9* (1), 150–158.
52. Roy, A.; Brand, N.J.; Yacoub, M.H. Molecular characterization of interstitial cells isolated from human heart valves. *J. Heart Valve Dis.* **2000**, *9* (3), 459–464. Discussion 464–465.
53. Della Rocca, F.; et al. Cell composition of the human pulmonary valve: A comparative study with the aortic valve—the VESALIO Project. *Vitalitate Exornatum Succedaneum Aorticum labore Ingenoso Obtinebitur.* *Ann. Thorac. Surg.* **2000**, *70* (5), 1594–1600.
54. Love, J.W. *Autologous Tissue Valves*; R.G. Landes Company: Austin, TX, 1993.
55. Marchand, M.; et al. Twelve-year experience with Carpentier–Edwards PERIMOUNT pericardial valve in the mitral position: A multicenter study. *J. Heart Valve Dis.* **1998**, *7* (3), 292–298.
56. Cheung, D.T.; et al. Behavior of vital and killed autologous pericardium in the descending aorta of sheep. *J. Thorac. Cardiovasc. Surg.* **1999**, *118* (6), 998–1005.
57. Marchion, D.C.; et al. Serum components stimulate pericardial tissue contraction. *J. Heart Valve Dis.* **2002**, *11* (2), 283–287.
58. Eastwood, M.; McGrouther, D.A.; Brown, R.A. Fibroblast responses to mechanical forces. *Proc. Inst. Mech. Eng., H* **1998**, *212* (2), 85–92.
59. Chiquet, M. Regulation of extracellular matrix gene expression by mechanical stress. *Matrix Biol.* **1999**, *18* (5), 417–426.
60. Elder, S.H.; et al. Chondrocyte differentiation is modulated by frequency and duration of cyclic compressive loading. *Ann. Biomed. Eng.* **2001**, *29* (6), 476–482.
61. Mudera, V.C.; et al. Molecular responses of human dermal fibroblasts to dual cues: Contact guidance and mechanical load. *Cell Motil. Cytoskelet.* **2000**, *45* (1), 1–9.
62. Zund, G.; et al. The in vitro construction of a tissue engineered bioprosthetic heart valve. *Eur. J. Cardiothorac. Surg.* **1997**, *11* (3), 493–497.
63. Shinoka, T.; et al. Tissue engineering heart valves: Valve leaflet replacement study in a lamb model. *Ann. Thorac. Surg.* **1995**, *60* (6 Suppl), S513–S516.
64. Hoerstrup, S.P.; et al. Functional living trileaflet heart valves grown in vitro. *Circulation* **2000**, *102* (19 Suppl 3), III44–III49.
65. Vesely, I.; Noseworthy, R.; Wilson, G. Development of a hybrid xenograft/autograft aortic valve bioprosthesis. *Can. J. Cardiol.* **1991**.
66. Wilson, G.J.; et al. Acellular matrix allograft small caliber vascular prostheses. *ASAIO Trans.* **1990**, *36* (3), M340–M343.
67. Courtman, D.W.; et al. Development of a pericardial acellular matrix biomaterial: Biochemical and mechanical effects of cell extraction. *J. Biomed. Mater. Res.* **1994**, *28* (6), 655–666.
68. Goldstein, S.; et al. Transpecies heart valve transplant: Advanced studies of a bioengineered xeno-autograft. *Ann. Thorac. Surg.* **2000**, *70* (6), 1962–1969.
69. Livesey, S.; Boerboom, L.; Coleman, C. Decellularized Porcine Heart Valves as Templates for Host Cell Repopulation. In *Workshop on Prosthetic Heart Valves: Future Directions*; Hilton Head: South Carolina, 1998.
70. Mitchell, R.N.; Jonas, R.A.; Schoen, F.J. Pathology of explanted cryopreserved allograft heart valves: Comparison with aortic valves from orthotopic heart transplants. *J. Thorac. Cardiovasc. Surg.* **1998**, *115* (1), 118–127.
71. Human, P.; Zilla, P. Characterization of the immune response to valve bioprostheses and its role in primary tissue failure. *Ann. Thorac. Surg.* **2001**, *71* (5 Suppl), S385–S388.
72. Rajani, B.; Mee, R.B.; Ratliff, N.B. Evidence for rejection of homograft cardiac valves in infants. *J. Thorac. Cardiovasc. Surg.* **1998**, *115* (1), 111–117.
73. Vyavahare, N.; et al. Mechanisms of bioprosthetic heart valve failure: Fatigue causes collagen denaturation and glycosaminoglycan loss. *J. Biomed. Mater. Res.* **1999**, *46* (1), 44–50.
74. Bell, E.; et al. Living tissue formed in vitro and accepted as skin-equivalent tissue of full thickness. *Science* **1981**, *211* (4486), 1052–1054.
75. Huang, D.; et al. Mechanisms and dynamics of mechanical strengthening in ligament-equivalent fibroblast-populated collagen matrices. *Ann. Biomed. Eng.* **1993**, *21* (3), 289–305.

76. Butler, D.L.; Awad, H.A. Perspectives on cell and collagen composites for tendon repair. *Clin. Orthop.* **1999**, (367 Suppl), S324–S332.
77. Barocas, V.H.; Tranquillo, R.T. An anisotropic biphasic theory of tissue-equivalent mechanics: The interplay among cell traction, fibrillar network deformation, fibril alignment, and cell contact guidance. *J. Biomech. Eng.* **1997**, *119* (2), 137–145.
78. Bell, E.; Ivarsson, B.; Merrill, C. Production of a tissue-like structure by contraction of collagen lattices by human fibroblasts of different proliferative potential in vitro. *Proc. Natl. Acad. Sci. U. S. A.* **1979**, *76* (3), 1274–1278.
79. Harris, A.K.; Stopak, D.; Wild, P. Fibroblast traction as a mechanism for collagen morphogenesis. *Nature* **1981**, *290* (5803), 249–251.
80. Stopak, D.; Harris, A.K. Connective tissue morphogenesis by fibroblast traction. I. Tissue culture observations. *Dev. Biol.* **1982**, *90* (2), 383–398.
81. Grinnell, F.; Lamke, C.R. Reorganization of hydrated collagen lattices by human skin fibroblasts. *J. Cell. Sci.* **1984**, *66*, 51–63.
82. Guidry, C.; Grinnell, F. Studies on the mechanism of hydrated collagen gel reorganization by human skin fibroblasts. *J. Cell. Sci.* **1985**, *79*, 67–81.
83. Tranquillo, R.T.; Durrani, M.A.; Moon, A.G. Tissue engineering science: Consequences of cell traction force. *Cytotechnology* **1992**, *10* (3), 225–250.
84. Grinnell, F. Fibroblasts, myofibroblasts, and wound contraction. *J. Cell Biol.* **1994**, *124* (4), 401–404.
85. Grinnell, F. Fibroblast-collagen-matrix contraction: Growth-factor signalling and mechanical loading. *Trends Cell Biol.* **2000**, *10* (9), 362–365.
86. Oakes, B.W.; et al. The synthesis of elastin, collagen, and glycosaminoglycans by high density primary cultures of neonatal rat aortic smooth muscle. An ultrastructural and biochemical study. *Eur. J. Cell Biol.* **1982**, *27*, 34–46.
87. Shi, Y.; Vesely, I. The Use of Directed Collagen Gel Shrinkage in the Development of Tissue Engineered Mitral Valve Chordae. In *1st Biennial Meeting of the Society for Heart Valve Disease*; 2001. London, England.
88. Girton, T.S.; Oegema, T.R.; Tranquillo, R.T. Exploiting glycation to stiffen and strengthen tissue equivalents for tissue engineering. *J. Biomed. Mater. Res.* **1999**, *46* (1), 87–92.
89. Seliktar, D.; et al. Dynamic mechanical conditioning of collagen-gel blood vessel constructs induces remodeling in vitro. *Ann. Biomed. Eng.* **2000**, *28* (4), 351–362.
90. Eisenberg, L.M.; Markwald, R.R. Molecular regulation of atrioventricular valvuloseptal morphogenesis. *Circ. Res.* **1995**, *77* (1), 1–6.
91. Larsen, N.E.; et al. Hylan gel biomaterial: Dermal and immunologic compatibility. *J. Biomed. Mater. Res.* **1993**, *27*, 1129–1134.
92. Larsen, N.E.; et al. Biocompatibility of Hylan Polymers in Various Tissue Compartments. In *Polymers in Medicine and Pharmacy*; Mikos, A.G., Ed.; Materials Research Society: Pittsburgh, 1995; 149–153.
93. Denlinger, J.L.; El-Mofty, A.A.; Balazs, E.A. Replacement of the liquid vitreous with sodium hyaluronate in monkeys. II. Long-term evaluation. *Exp. Eye Res.* **1980**, *31* (1), 101–117.
94. Laurent, T.C. Biochemistry of hyaluronan. *Acta Otolaryngol., Suppl.* **1987**, *442*, 7–24.
95. Balazs, E.A.; Leshchiner, A. *Crosslinked Gels of Hyaluronic Acid and Products*; Biomatrix Inc., 1984.
96. Balazs, E.A.; Denlinger, J.L. Clinical uses of hyaluronan. *Ciba Found. Symp.* **1989**, *143*, 265–275. Discussion 275–280.
97. Vercruyse, K.P.; et al. Synthesis and in vitro degradation of new polyvalent hydrazide cross-linked hydrogels of hyaluronic acid. *Bioconj. Chem.* **1997**, *8* (5), 686–694.
98. Tomihata, K.; Ikada, Y. Crosslinking of hyaluronic acid with water-soluble carbodiimide. *J. Biomed. Mater. Res.* **1997**, *37* (2), 243–251.
99. Mälson, T.; Lindqvist, B. Crosslinked hyaluronate gels their use and method for producing them. **1998**.
100. Balazs, E.A.; et al. Matrix engineering. *Blood Coagul. Fibrinolysis* **1991**, *2* (1), 173–178.
101. Ramamurthi, A.; Vesely, I. Smooth muscle cell adhesion on crosslinked hyaluronan gels. *J. Biomed. Mater. Res.* **2002**, *60* (1), 195–205.
102. Bryant, S.J.; Nuttelman, C.R.; Anseth, K.S. The effects of crosslinking density on cartilage formation in photocrosslinkable hydrogels. *Biomed. Sci. Instrum.* **1999**, *35*, 309–314.
103. Neitmann, M.; et al. Attachment and chemotaxis of melanocytes after ultraviolet irradiation in vitro. *Br. J. Dermatol.* **1999**, *141* (5), 794–801.

Tissue Engineering of the Intestine

James C.Y. Dunn

Department of Bioengineering, and Department of Surgery, University of California, Los Angeles, California, U.S.A.

Benjamin M. Wu

Department of Biomedical Engineering, and Department of Materials Science & Engineering, University of California, Los Angeles, California, U.S.A.

T

INTRODUCTION

The intestine is a complex organ that integrates the functions of many types of cells and tissues. Current attempts to regenerate the intestine have yielded promising results. Nevertheless, significant challenges remain before intestinal tissue engineering can become clinically useful. A review of the state of the art and a discussion on the challenges ahead will be presented in this article.

BACKGROUND

A variety of human diseases require the extensive removal of the intestine. For the small intestine, this can occur in the setting of necrotizing enterocolitis, intestinal volvulus, mesenteric ischemia, and Crohn's disease.^[1-3] For the large intestine, total proctocolectomy is frequently performed for patients with familial polyposis and ulcerative colitis. Although supportive measures and reconstructive procedures are available for these patients, the current therapies suffer from significant shortcomings.

Short gut syndrome occurs in patients with inadequate lengths of functional small intestine to maintain normal digestion and absorption. Although the prognosis of patients with short gut syndrome depends on several variables, including the presence of the ileocecal valve and colon and the underlying etiology of the disease, the length of the remaining small intestine is arguably the most significant factor. In many series of patients with short gut syndrome, it has been shown that their survival depends on the length of the remaining small intestine.^[3] While the exact length of the remaining intestine may be difficult to determine owing to its elastic and dynamic nature, it is generally accepted that short gut syndrome occurs when more than two-thirds of the small intestine is removed. In adults, as a general rule, 200 cm of the small intestine is needed for adequate bowel function. When the remaining length of the small intestine is less than 100 cm, the mortality of the patients begins to rise.

The therapy for short gut syndrome consists of providing parenteral nutrition, optimizing the health of the remnant intestine, and enhancing bowel adaptation.^[4] Besides the estimated cost of \$200,000 per patient each year, the use of long-term parenteral nutrition is associated with significant complications such as catheter sepsis and liver cirrhosis. Surgical procedures to optimize the intestinal function include creating intestinal valves and recirculating intestinal loops, interposing a colonic segment in an isoperistaltic or antiperistaltic fashion, and tapering and lengthening dilated intestinal segments.^[5-9] Most of these procedures have met with only limited success.^[10,11] The outcome following intestinal transplantation is improving; but it is still limited by complications of immunosuppression and donor availability.^[12]

For patients who need the removal of their entire large intestine, reconstructive surgery utilizing the distal small intestine to create an ileoanal pouch is now the standard of care.^[13] Nevertheless, the ileoanal pouch procedure is associated with dilation, inflammation, stool frequency, and pelvic sepsis. The frequency of complications associated with the ileoanal pouches is over 50% in 10 years, and the removal of the pouch is required so that a permanent ileostomy is necessary in some cases.^[14,15] The colonic metaplasia and the chronic inflammation of the pouches have been postulated to increase long-term neoplasia.^[16]

Tissue engineering offers the potential to regenerate an autologous intestine. For patients with short gut syndrome, this would represent a new, life-saving therapy that will allow them to be weaned from parenteral nutrition and to eat normally. For patients with colonic failure, this would avoid much of the morbidity associated with the current restorative procedures. Although a complex organ, portions of the intestine had been regenerated using cells, scaffolds, and signals.

INTESTINAL ARCHITECTURE

The intestine is derived from the lateral folding of the trilayered embryonic disc to form a tube during

development. The innermost aspect of the intestine is lined with epithelial cells facing the lumen of the gut. These cells originated from the embryonic endodermal cells. Beneath the intestinal epithelium lie the lamina propria, the submucosa, and the smooth muscle layers, all derived from the embryonic mesoderm. The serosa is the outermost layer of the intestine and is penetrated by the mesentery blood vessels and lymphatics. Dispersed throughout the intestine are intrinsic and extrinsic autonomic nerves, both of ectodermal origin. The intestine lengthens during fetal development and postnatal growth. The human fetus in the second trimester has 150 cm of intestines which double to 300 cm in full-term newborns.^[17] Thereafter the intestine continues to elongate, reaching an average length of 600 cm in adulthood, although there is a wide range from 350 to 700 cm, depending on the age, height, weight, and race.^[18]

A review of the normal intestinal histology reveals the layered architecture of the intestinal wall. The epithelium is the first one-cell thick layer followed by several other layers including the lamina propria, the submucosal plexus, the circular smooth muscle, the myenteric plexus, the longitudinal smooth muscle, and the serosa. Each layer consists of its own specialized cells with their unique environment. The intestinal epithelium covers the mucosa with a layer of columnar absorptive cells, goblet cells, endocrine cells, epithelial stem cells, and Paneth cells. The intestinal epithelial stem cells reside near the bottom of crypts and give rise to all of the other cell types in the epithelium. Epithelial cells generated in the crypts migrate toward the villi where they undergo maturation, followed by apoptosis over the course of a few days. The absorptive cells are highly polarized. They possess apical microvilli that form the brush border and basally located nuclei. The brush border contains digestive enzymes and nutrient transporters, and the basolateral surface contains adenosine triphosphate-dependent pumps that maintain the ionic gradient inside the cells to facilitate the directional transport. The goblet cells are variably dispersed among the absorptive cells. They secrete a mucopolysaccharide complex believed to act as a protective layer for the intestinal epithelium. The endocrine cells are located in both the villi and crypts. Their basally located granules contain gastrointestinal hormones such as somatostatin vasoactive intestinal peptide and motilin. The Paneth cells are located in the lowest part of the crypts. They also secrete proteins such as defensins that have a wide range of antibiotic effects.^[19]

The lamina propria lies beneath the intestinal epithelium and is the core of the villi. A rich network of lymphatics and blood vessels is located in the lamina propria to transport the absorbed nutrients into the systemic circulation. There are also numerous intraepithelial lymphocytes and lymphoid nodules in this region, and they provide immunity against the intestinal

flora and secrete immunoglobulin A. The submucosa is rich in elastic fibers. In the duodenum, the submucosa is filled with secretory glands that produce an alkalotic viscous fluid to neutralize the gastric acid.

The peristaltic action of the intestine results from the rhythmic contraction of two smooth muscular layers coordinated by the enteric nervous system and paced by the interstitial cells of Cajal. The smooth muscle cells are oriented along two axes with one arranged circumferentially around the intestine and the other longitudinally along the intestine. They can be identified by immunostaining for α smooth muscle actin and smooth muscle myosin heavy chain. The enteric nervous system is organized as a network with ganglia at the nodal points. It has been estimated that the enteric nervous system contains as many neurons as the central nervous system. The myenteric plexus lies in between the two smooth muscle layers and stimulates the contraction of the smooth muscle cells. The interstitial cells of Cajal are considered to be the pacemaker cells of gastrointestinal tract. They form contacts with the smooth muscle cells and enteric neurons and are responsible for the generation of rhythmic slow waves of the intestine. The serosa is the outermost layer and consists of a squamous epithelium that covers the entire gastrointestinal tract and is continuous with the peritoneum.

FEASIBILITY STUDIES

The intestine is a complex organ composed of many cell types, including epithelial cells, immune cells, endothelial cells, smooth muscle cells, pacemaker cells, and nerve cells. These cells are highly organized into tissues that are intricately integrated with each other to bring about the coordinated function of the intestine. Although the regeneration of such complex organ may seem unlikely at first, there is some evidence in the literature that documented the feasibility of intestinal tissue engineering.

Much of the early work on intestinal regeneration is focused on the epithelium, which is known to possess stem cells with the tremendous ability for growth. The growth of intestinal mucosa in ectopic locations is observed in many clinical settings, such as their aberrant presence in fistulous tracts in Crohn's disease. When an opening created in the small intestine is patched closed using the serosa from another intestinal segment, the epithelial cells migrate onto the new surface and form structures resembling the normal intestinal mucosa.^[20] Such "neomucosa" can be grown on a variety of surfaces, including biological and prosthetic materials.^[21,22] Although the clinical utility of the neomucosa is limited, as the newly generated epithelium contracted over time,^[23] the innate ability of the intestinal epithelial cells

to grow and to organize sets the stage for the initial efforts in intestinal tissue engineering.

By using the neonatal small intestine as the starting source, one can procure rodent intestinal organoids that are made up of heterogeneous cell types.^[24] These organoids were attached to a biodegradable polymeric scaffold following isolation and were implanted into the omentum of syngeneic rats.^[25] After six weeks, a cystic structure formed in the peritoneal cavity. The inner lining of such cysts resembled epithelial cells of the small intestine, and the outer layer was composed of fibrovascular tissues. When the opened cysts were anastomosed to the side of an intestinal segment, the lining cells blossomed into structures resembling intestinal crypts and villi.^[26] When the function of these cysts was tested in an animal model of short gut syndrome, the weight loss observed in rats with surgically created short gut syndrome was lessened in the presence of the cyst.^[27] These results established the feasibility of small intestinal tissue engineering.

Using an analogous technique for the large intestine, colonic tissues can also be regenerated from organoid units procured from both neonatal and adult animals.^[28] After four weeks of implantation into the omentum, a 4-cm cyst formed in the peritoneal cavity. Such cysts could be anastomosed to native intestine and exhibited similar histological architecture as the native colon. Furthermore, such tissue-engineered colons could decrease the output from an ileostomy in rats, suggesting absorptive function *in vivo*. Although these demonstrations are impressive, several major challenges still need to be addressed before intestinal tissue engineering can be applied to patients.

CELLS

Before intestinal tissue engineering can be brought into the clinical arena, it is necessary to find an alternate cell source for tissue regeneration. For the small intestine, the established technique employed organoid units pooled from several neonatal rat donors to procure sufficient cell mass to generate a single cyst. When the adult small intestine was used, the procured organoids were not able to generate similar intestinal cysts after implantation into the omentum. This reflects the tremendous growth potential of cells obtained from neonatal animals. There is, however, no readily available neonatal human small intestine that can be used in a similar fashion. In contrast, for the large intestine, the established technique was successful in employing either neonatal or adult rats as the donor for intestinal organoids. However, patients who might benefit from tissue-engineered colons will have diseases that affect the entire colon, such as familial polyposis. The diffuse nature of the disease will not permit the use of the

autologous colon as the source for cell procurement. Therefore, the search for a substitute cell source is a critical next step in intestinal tissue engineering.

For the intestine to function properly, it must possess adequate peristaltic function to transport the enteric content down the gastrointestinal tract. This is evident from a variety of human diseases where a histologically normal intestine fails to function because of poor intestinal motility. The current tissue-engineered intestine has an adequately regenerated mucosa, but the presence of a functional neuromuscular layer has not been well documented. Although some studies to date have shown the presence of smooth muscle cells and ganglia, no documentation of peristalsis has been demonstrated. Analogous to the intestine in Hirschsprung's disease, such tissue-engineered structures would not be clinically useful without motility.

Embryonic stem cells have the potential to generate every cell type in the body. Using murine embryonic stem cells, Yamada et al. were able to generate an *in vitro* gut-like organ.^[29] These dome-like structures possessed an inner epithelial layer and an outer smooth muscle layer. They also exhibited coordinated peristalsis and showed immunoreactivity for interstitial cells of Cajal and neurons. Despite this tremendous potential, the acceptability of therapeutic cloning remains highly controversial. Neural stem cells derived from the central nervous system were transplanted into the gastrointestinal tract.^[30] Such neural stem cells expressed the receptors for the glial cell derived nerve growth factor and could produce nitric oxide *in vitro*. After transplantation, these cells continued to express the neural nitric oxide synthase and survived for at least eight weeks.

Besides embryonic stem cells, postnatal stem cells have also been utilized for intestinal tissue engineering. Gastrointestinal smooth muscle cells have been isolated and grown in culture.^[31] They are capable of expansion *in vitro*, although the cells tend to lose differentiation during the growth phase. Cells from the enteric nervous system are derived from the neural crest cells during development. It has been shown that neural crest stem cells persist in the intestine of postnatal animals.^[32] These cells can be isolated using an antibody to the surface receptor p75. The p75-positive cells can be expanded *in vitro*, and they can be differentiated into neuronal and glial cells. The enteric neuronal cells do not innervate the smooth muscle cells by the classic neuromuscular junction. Instead, neurotransmitters diffuse from the nerve terminal toward the smooth muscle cells. The interstitial cells of Cajal can be identified by the surface marker c-kit, a receptor tyrosine kinase for stem cell factor. These cells can also be isolated and cultured *in vitro*,^[33] and the cultured cells exhibit rhythmic spontaneous action potentials. For a tissue-engineered intestine to have adequate peristaltic action, it is likely that all of these cellular elements

need to be present and integrated in the regenerated structure.

SCAFFOLDS

The scaffold allows the design of the overall shape, provides a surface for cellular attachment, and can simulate an environment that is conducive for tissue regeneration. Both natural and polymeric materials have been used for intestinal tissue engineering.^[25,34] Microscopic features, such as the intestinal villi, can be replicated using three-dimensional printing techniques.^[35] It is likely that both physical and biological cues can be incorporated into the design of the scaffold to facilitate tissue engineering.^[36]

Although started as a cylindrical scaffold seeded with intestinal organoid units, the cysts generated in the omentum did not retain the tubular shape.^[25] The spherical shape of the cysts does not allow for ready scale-up to clinically useful dimensions. By increasing the diameter of the cyst to increase the surface area, the dilated cysts would cause stasis and bacterial overgrowth within the lumen and would be considered as diverticula of the intestine, which is clinically a pathological condition.

Analogous to its normal state, the ideal geometry for intestinal tissue engineering is a cylindrical structure that scales with length of the scaffold. When a tubularized scaffold is interposed between intestinal segments, intestinal obstruction occurs because of the lack of adequate motility in the scaffold.^[22,37] When an acellular, tubularized scaffold was interposed between intestinal segments isolated from intestinal continuity, a new intestinal epithelium regenerated over the tubular construct over 4–16 weeks.^[22,34] The spontaneous epithelialization is not surprising, given the ability of the normal intestinal epithelium to turnover. Underneath the epithelium, however, there is little smooth muscle regeneration. When the scaffolds were seeded with mesenchymal stem cells, smooth muscle was still not regenerated after epithelialization was complete.^[38] The lack of multiple intestinal layers implies that other environmental cues are needed for the regeneration of the smooth muscle and the enteric nervous system.

SIGNALS

Besides the cells and the scaffolds, the proper signals are needed to guide tissue regeneration. A myriad of factors have been shown to stimulate the growth of the intestinal villi and crypts in animal models,^[39] including epidermal growth factor,^[40,41] hepatocyte growth factor,^[42,43] glucagon-like peptide 2,^[44,45] interleukin 11,^[46] and insulin-like growth factor 1.^[47–50] Although most of

these factors clearly stimulate the growth of intestinal epithelial cells, the optimal dosage and the exact timing for the delivery of these factors remain to be determined. Furthermore, it may be unnecessary to provide any trophic factors for the intestinal epithelial cells, as evidenced by their propensity to grow spontaneously in conditions without added growth factors as stated previously. On the other hand, intestinal smooth muscle cells and enteric neurons have much less potential to grow in the current models of regeneration. The delivery of factors that will direct the regeneration of these cellular elements is likely to have a significant impact on the progress of intestinal tissue engineering.^[51–55]

Because of its central role in providing the source of energy, the intestine is intricately integrated with the other organ systems. In addition to the regeneration of the intestine itself, other crucial elements are needed to reconnect the intestine with the body. Most importantly, the engineered construct needs vascularization to provide the pathway for nutrient exchange. Angiogenesis does occur in the intestinal cysts generated by the current technique; however, there are significant quantitative differences between the normal growing intestine and the tissue-engineered construct.^[56] In particular, the levels of vascular endothelial growth factor and basic fibroblast growth factor are low in the engineered intestine, and these may serve as the targets for growth factor delivery. Another important pathway to establish is the lymphatic system, which transports chylomicrons back into the circulation. Lymphangiogenesis also occurs in the intestinal cysts, but their function has not been assessed.^[57]

The small intestine also grows in response to applied mechanical forces.^[58] The capacity of the bypassed ileum in pigs can be expanded to 12 times the original volume in eight weeks by using a round, intraluminal balloon expander.^[59] Similarly, the expansion of the jejunal capacity in rabbits using an oval, intraluminal tissue expander was described.^[60] When an intussusception valve was created to obstruct the jejunum partially in rats, the proximal jejunum dilated in response to the increased intraluminal pressure.^[61] All of these studies showed that the intestine was not simply stretched thin but rather intestinal hypertrophy and hyperplasia had occurred in response to the applied force. Although successful in increasing the total surface area by increasing the intestinal diameter, these techniques resulted in dilated segments of the small intestine that would not function well because of stasis and bacterial overgrowth.

There were also attempts directed at lengthening the small intestine by mechanical stretching without causing simultaneous dilatation. The elongation of the ileum in rabbits was achieved by the placement of a round, extraluminal tissue expander adjacent to a fixed intestinal segment.^[62] Because of geometric considerations, the maximal gain in length was limited using this

technique. A more recent study used an external distractor sutured to the serosa of the jejunum in rabbits and reported a near doubling of the initial length after three weeks of gradual distraction.^[63] When the distraction device was removed, the intestinal segment remained longer than its initial length. Another study employed intraluminal expanders to lengthen bypassed ileal segments in rabbits and reported 50% lengthening after two weeks.^[64] They also observed persistent lengthening after the intraluminal expanders were removed. Although promising, none of these studies examined the in vivo functions of the mechanically lengthened bowel.

CONCLUSIONS

There is yet much that we do not understand about intestinal regeneration. The signals critical to the process are unknown, the role of scaffolding remains to be determined, and the necessary cell types are not well defined. New knowledge gained from studying the development of the intestine will yield further insights into nature's intestinal tissue engineering so that better therapies can be developed for patients with intestinal failure.

ACKNOWLEDGMENT

This work was supported by a grant from the Fubon Foundation.

ARTICLE OF FURTHER INTEREST

Tissue Engineering of the Small Intestine

REFERENCES

- Georgeson, K.E.; Breaux, C.W., Jr. Outcome and intestinal adaptation in neonatal short-bowel syndrome. *J. Pediatr. Surg.* **1992**, *27*, 344–350.
- Koffeman, G.I.; Van Gemert, W.G.; George, E.K.; Veenendaal, R.A. Classification, epidemiology and aetiology. *Best Pract. Res. Clin. Gastroenterol.* **2003**, *17* (6), 879–893.
- Messing, B.; Crenn, P.; Beau, P.; Boutron-Ruault, M.C.; Rambaud, J.C.; Matuchansky, C. Long-term survival and parenteral nutrition dependence in adult patients with the short gut syndrome. *Gastroenterology* **1999**, *117* (5), 1043–1050.
- Wilmore, D.W. Indication for specific therapy in the rehabilitation of patients with the short gut syndrome. *Best Pract. Res. Clin. Gastroenterol.* **2003**, *17* (6), 895–906.
- Benedetti, E.; Panaro, F.; Holterman, M.; Abcarian, H. Surgical approaches and intestinal transplantation. *Best Pract. Res. Clin. Gastroenterol.* **2003**, *17* (6), 1017–1040.
- Bianchi, A. Intestinal loop lengthening-A technique for increasing small intestinal length. *J. Pediatr. Surg.* **1980**, *15*, 145–151.
- Kim, H.B.; Fauza, D.; Garza, J.; Oh, J.T.; Nurko, S.; Jaksic, T. Serial transverse enteroplasty (STEP): a novel bowel lengthening procedure. *J. Pediatr. Surg.* **2003**, *38* (3), 425–429.
- Kimura, K.; Soper, R.T. A new bowel elongation technique for the short gut syndrome using the isolated bowel segment Iowa models. *J. Pediatr. Surg.* **1993**, *28*, 792–794.
- Panis, Y.; Messing, B.; Rivet, P.; Coffin, B.; Hautefeuille, P.; Matuchansky, C.; Rambaud, J.C.; Valleur, P. Segmental reversal of the small bowel as an alternative to intestinal transplantation in patients with short gut syndrome. *Ann. Surg.* **1997**, *225*, 401–407.
- Bianchi, A. Experience with longitudinal intestinal lengthening and tailoring. *Eur. J. Pediatr. Surg.* **1999**, *9*, 256–259.
- Thompson, J.S. Surgical approach to the short gut syndrome: procedures to slow intestinal transit. *Eur. J. Pediatr. Surg.* **1999**, *9*, 263–266.
- Mittal, N.K.; Tzakis, A.G.; Kato, T.; Thompson, J.F. Current status of small bowel transplantation in children: update 2003. *Pediatr. Clin. North Am.* **2003**, *50* (6), 1419–1433.
- Fonkalsrud, E.W.; Bustorff-Silva, J. Reconstruction for chronic dysfunction of ileoanal pouches. *Ann. Surg.* **1999**, *229*, 197–204.
- Meagher, A.P.; Farouk, R.; Dozois, R.R.; Kelley, K.A.; Pemberton, J.H. J ileal pouch-anal anastomosis for chronic ulcerative colitis: complications and long-term outcome in 1310 patients. *Br. J. Surg.* **1998**, *85*, 800–803.
- Stahlberg, D.; Gullberg, K.; Liljeqvist, L.; Hellers, G.; Lofberg, R. Pouchitis following pelvic pouch operation for ulcerative colitis. Incidence, cumulative risk, and risk factors. *Dis. Colon Rectum* **1996**, *39*, 1012–1018.
- Shepherd, N.A. The pelvic ileal reservoir: apocalypse later? *BMJ*, **1990**, *301*, 886–887.
- Touloukian, R.J.; Smith, G.J. Normal intestinal length in preterm infants. *J. Pediatr. Surg.* **1983**, *18*, 720–723.
- Weaver, L.T.; Austin, S.; Cole, T.J. Small intestinal length: a factor essential for gut adaptation. *Gut* **1991**, *32* (11), 1321–1323.
- Bevins, C.L. Events at the host-microbial interface of the gastrointestinal tract V. Paneth cell alpha-defensins in intestinal host defense. *Am. J. Physiol. Gastrointest. Liver Physiol.* **2005**, *289* (2), G173–G176.
- Binnington, H.B.; Sumner, H.; Lesker, P.; Alpers, D.A.; Ternberg, J.L. Functional characteristics of surgically induced jejunal neomucosa. *Surgery* **1974**, *75* (6), 805–810.
- De Ugarte, D.A.; Choi, E.; Weitzbuch, H.; Wulur, I.; Caulkins, C.; Wu, B.; Fonkalsrud, E.W.; Atkinson, J.B.; Dunn, J.C. Mucosal regeneration of a duodenal defect using small intestine submucosa. *Am. Surg.* **2004**, *70* (1), 49–51.

22. Thompson, J.S.; Kampfe, P.W.; Newland, J.R.; Vanderhoof, J.A. Growth of intestinal neomucosa on prosthetic materials. *J. Surg. Res.* **1986**, *41* (5), 484–492.
23. Thompson, J.S.; Hollingsed, T.C.; Saxena, S.K. Prevention of contraction of patched intestinal defects. *Arch. Surg.* **1988**, *123* (4), 428–430.
24. Evans, G.S.; Flint, N.; Somers, A.S.; Eyden, B.; Potten, C.S. The development of a method for the preparation of rat intestinal epithelial cell primary cultures. *J. Cell Sci.* **1992**, *101*, 219–231.
25. Choi, R.S.; Riegler, M.; Pothoulakis, C.; Kim, B.S.; Mooney, D.; Vacanti, M.; Vacanti, J.P. Studies of brush border enzymes, basement membrane components, and electrophysiology of tissue-engineered neointestine. *J. Pediatr. Surg.* **1998**, *33* (7), 991–997.
26. Kim, S.S.; Kaihara, S.; Benvenuto, M.S.; Choi, R.S.; Kim, B.S.; Mooney, D.J.; Vacanti, J.P. Effects of anastomosis of tissue-engineered neointestine to native small bowel. *J. Surg. Res.* **1999**, *87* (1), 6–13.
27. Grikscheit, T.C.; Siddique, A.; Ochoa, E.R.; Srinivasan, A.; Alsberg, E.; Hodin, R.A.; Vacanti, J.P. Tissue-engineered small intestine improves recovery after massive small bowel resection. *Ann. Surg.* **2004**, *240* (5), 748–754.
28. Grikscheit, T.C.; Ochoa, E.R.; Ramsanahie, A.; Alsberg, E.; Mooney, D.; Whang, E.E.; Vacanti, J.P. Tissue-Engineered large intestine resembles native colon with appropriate in vitro physiology and architecture. *Ann. Surg.* **2003**, *238* (1), 35–41.
29. Yamada, T.; Yoshikawa, M.; Takaki, M.; Torihashi, S.; Kato, Y.; Nakajima, Y.; Ishizaka, S.; Tsunoda, Y. In vitro functional gut-like organ formation from mouse embryonic stem cells. *Stem Cells* **2002**, *20* (1), 41–49.
30. Micci, M.A.; Learish, R.D.; Li, H.; Abraham, B.P.; Pasricha, P.J. Neural stem cells express RET, produce nitric oxide, and survive transplantation in the gastrointestinal tract. *Gastroenterology* **2001**, *121* (4), 757–766.
31. Brittingham, J.; Phiel, C.; Trzyna, W.C.; Gabbeta, V.; McHugh, K.M. Identification of distinct molecular phenotypes in cultured gastrointestinal smooth muscle cells. *Gastroenterology* **1998**, *115* (3), 605–617.
32. Kruger, G.M.; Mosher, J.T.; Bixby, S.; Joseph, N.; Iwashita, T.; Morrison, S.J. Neural crest stem cells persist in the adult gut but undergo changes in self-renewal, neuronal subtype potential, and factor responsiveness. *Neuron* **2002**, *35* (4), 657–669.
33. Li, C.X.; Liu, B.H.; Tong, W.D.; Zhang, L.Y.; Jiang, Y.P. Dissociation, culture and morphologic changes of interstitial cells of Cajal in vitro. *World J. Gastroenterol.* **2005**, *11* (18), 2838–2840.
34. Hori, Y.; Nakamura, T.; Matsumoto, K.; Kurokawa, Y.; Satomi, S.; Shimizu, Y. Tissue engineering of the small intestine by acellular collagen sponge scaffold grafting. *Int. J. Artif. Organs.* **2001**, *24* (1), 50–54.
35. Lee, M.; Dunn, J.C.; Wu, B.M. Scaffold fabrication by indirect three-dimensional printing. *Biomaterials.* **2005**, *26* (20), 4281–4289.
36. Nishimura, I.; Garrell, R.L.; Hedrick, M.; Iida, K.; Osher, S.; Wu, B. Precursor tissue analogs as a tissue-engineering strategy. *Tissue Eng.* **2003**, *9* (suppl 1), S77–S89.
37. Chen, M.K.; Badylak, S.F. Small bowel tissue engineering using small intestinal submucosa as a scaffold. *J. Surg. Res.* **2001**, *99* (2), 352–358.
38. Hori, Y.; Nakamura, T.; Kimura, D.; Kaino, K.; Kurokawa, Y.; Satomi, S.; Shimizu, Y. Experimental study on tissue engineering of the small intestine by mesenchymal stem cell seeding. *J. Surg. Res.* **2002**, *102* (2), 156–160.
39. Thiesen, A., Jr.; Drozdowski, L.; Iordache, C.; Neo, C.C.; Woudstra, T.D.; Xenodemetropoulos, T.; Keelan, M.; Clandinin, M.T.; Thompson, A.B.R. Adaptation following intestinal resection: mechanisms and signals. *Best Pract. Res. Clin. Gastroenterol.* **2003**, *17* (6), 981–995.
40. Dunn, J.C.; Parungo, C.P.; Fonkalsrud, E.W.; McFadden, D.W.; Ashley, S.W. Epidermal growth factor selectively enhances functional enterocyte adaptation after massive small bowel resection. *J. Surg. Res.* **1997**, *67* (1), 90–93.
41. Huang, F.S.; Kemp, C.J.; Williams, J.L.; Erwin, C.R.; Warner, B.W. Role of epidermal growth factor and its receptor in chemotherapy-induced intestinal injury. *Am. J. Physiol. Gastrointest. Liver Physiol.* **2002**, *282* (3), G432–G442.
42. Kuenzler, K.A.; Pearson, P.Y.; Schwartz, M.Z. Hepatocyte growth factor pretreatment reduces apoptosis and mucosal damage after intestinal ischemia-reperfusion. *J. Pediatr. Surg.* **2002**, *37*, 1093–1097.
43. Ohda, Y.; Hori, K.; Tomita, T.; Hida, N.; Kosaka, T.; Fukuda, Y.; Miwa, H.; Matsumoto, T. Effects of hepatocyte growth factor on rat inflammatory bowel disease models. *Dig. Dis. Sci.* **2005**, *50* (5), 914–921.
44. Kitchen, P.A.; Goodlad, R.A.; FitzGerald, A.J.; Mandir, N.; Ghatei, M.A.; Bloom, S.R.; Berlanga-Acosta, J.; Playford, R.J.; Forbes, A.; Walters, J.R. Intestinal growth in parenterally-fed rats induced by the combined effects of glucagon-like peptide 2 and epidermal growth factor. *JPEN* **2005**, *29* (4), 248–254.
45. Shin, E.D.; Estall, J.L.; Izzo, A.; Drucker, D.J.; Brubaker, P.L. Mucosal adaptation to enteral nutrients is dependent on the physiologic actions of glucagon-like peptide-2 in mice. *Gastroenterology* **2005**, *128* (5), 1340–1353.
46. Fiore, N.F.; Ledniczky, G.; Liu, Q.; Orazi, A.; Du, X.; Williams, D.A.; Grosfeld, J.L. Comparison of interleukin-11 and epidermal growth factor on residual small intestine after massive small bowel resection. *J. Pediatr. Surg.* **1998**, *33* (1), 24–29.
47. Steeb, C.B.; Trahair, J.F.; Tomas, F.M.; Read, L.C. Prolonged administration of IGF peptides enhances growth of gastrointestinal tissues in normal rats. *Am. J. Physiol.* **1994**, *266* (6 Pt 1), G1090–G1098.
48. Lund, P.K. Molecular basis of intestinal adaptation: the role of the insulin-like growth factor system. *Ann. N Y Acad. Sci.* **1998**, *859*, 18–36.
49. Peterson, C.A.; Gillingham, M.B.; Mohapatra, N.K.; Dahly, E.M.; Adamo, M.L.; Carey, H.V.; Lund, P.K.; Ney, D.M. Enterotrophic effect of insulin-like growth factor-I but not growth hormone and localized expression of insulin-like growth factor-I, insulin-like growth factor binding protein-3 and -5 mRNAs in jejunum of parenterally fed rats. *JPEN* **2000**, *24* (5), 288–295.

50. Theiss, A.L.; Fruchtman, S.; Lund, P.K. Growth factors in inflammatory bowel disease: the actions and interactions of growth hormone and insulin-like growth factor-I. *Inflamm. Bowel Dis.* **2004**, *10* (6), 871–880.
51. Wang, J.; Niu, W.; Nikiforov, Y.; Naito, S.; Chernausek, S.; Witte, D.; LeRoith, D.; Strauch, A.; Fagin, J.A. Targeted overexpression of IGF-I evokes distinct patterns of organ remodeling in smooth muscle cell tissue beds of transgenic mice. *J. Clin. Invest.* **1997**, *100*, 1425–1439.
52. Knott, A.W.; Juno, R.J.; Jarboe, M.D.; Profitt, S.A.; Erwin, C.R.; Smith, E.P.; Fagin, J.A.; Warner, B.W. Smooth muscle overexpression of IGF-I induces a novel adaptive response to small bowel resection. *Am. J. Physiol. Gastrointest. Liver Physiol.* **2004**, *287* (3), G562–G570.
53. Kuemmerle, J.F.; Zhou, H.; Bowers, J.G. IGF-I stimulates human intestinal smooth muscle cell growth by regulation of G1 phase cell cycle proteins. *Am. J. Physiol. Gastrointest. Liver Physiol.* **2004**, *286* (3), G412–G419.
54. Walters, J.R. Cell and molecular biology of the small intestine: new insights into differentiation, growth and repair. *Curr. Opin. Gastroenterol.* **2004**, *20* (2), 70–76.
55. Kleber, M.; Lee, H.Y.; Wurdak, H.; Buchstaller, J.; Riccomagno, M.M.; Ittner, L.M.; Suter, U.; Epstein, D.J.; Sommer, L. Neural crest stem cell maintenance by combinatorial Wnt and BMP signaling. *J. Cell Biol.* **2005**, *169* (2), 309–320.
56. Gardner-Thorpe, J.; Grikscheit, T.C.; Ito, H.; Perez, A.; Ashley, S.W.; Vacanti, J.P.; Whang, E.E. Angiogenesis in tissue-engineered small intestine. *Tissue Eng.* **2003**, *9* (6), 1255–1261.
57. Duxbury, M.S.; Grikscheit, T.C.; Gardner-Thorpe, J.; Rocha, F.G.; Ito, H.; Perez, A.; Ashley, S.W.; Vacanti, J.P.; Whang, E.E. Lymphangiogenesis in tissue-engineered small intestine. *Transplantation* **2004**, *77* (8), 1162–1166.
58. Park, J.; Puapong, D.P.; Wu, B.M.; Atkinson, J.B.; Dunn, J.C. Enterogenesis by mechanical lengthening: morphology and function of the lengthened small intestine. *J. Pediatr. Surg.* **2004**, *39* (12), 1823–1827.
59. Stark, G.B.; Dorer, A.; Walgenbach, K.J.; Grunwald, F.; Jaeger, K. The creation of a small bowel pouch by tissue expansion: an experimental study in pigs. *Langenbecks Arch. Chir.* **1990**, *375* (3), 145–150.
60. Narayan, D.; Castro, A.; Jackson, I.T.; Herschman, B. Tissue expanders in the gut: a histological and angiographic study. *J. R. Coll. Surg. Edinb.* **1992**, *37*, 402–404.
61. Kawaguchi, A.L.; Dunn, J.C.; Lam, M.; O'Connor, T.P.; Diamond, J.; Fonkalsrud, E.W. Glucose uptake in dilated small intestine. *J. Pediatr. Surg.* **1998**, *33* (11), 1670–1673.
62. Manders, E.K.; Saggars, G.C.; Diaz-Alonso, P.; Finn, L.; Sipio, J.C.; Glumac, T.; Au, V.K.; Wong, R.K.M.; Mottaleb, M. Elongation of peripheral nerve and viscera containing smooth muscle. *Clin. Plast. Surg.* **1987**, *14* (3), 551–562.
63. Printz, H.; Schlenzka, R.; Requadt, P.; Tscherny, M.; Wagner, A.C.; Eissele, R.; Rothmund, M.; Arnold, R.; Goke, B. Small bowel lengthening by mechanical distraction. *Digestion* **1997**, *58* (3), 240–248.
64. Chen, Y.; Zhang, J.; Qu, R.; Wang, J.; Xie, Y. An animal experiment on short gut lengthening. *Chin. Med. J.* **1997**, *110* (5), 354–357.

Tissue Engineering of Ligament

E. Gentleman

G. A. Livesay

K. C. Dee

E. A. Nauman

Department of Biomedical Engineering, Lindy Boggs Center, Suite 500, Tulane University, New Orleans, Louisiana, U.S.A.

INTRODUCTION

Ligaments perform essential stabilization roles in synovial joints by maintaining bone alignment and guiding motion. Although many ligaments in the body are regularly injured, those found in the knee joint have received particular attention because they are subjected to large tensile loads and surgical repair often results in poor outcomes. Efforts to create tissue-engineered ligaments can be guided by a thorough understanding of the physiological utility and needs of the natural tissue, and, regardless of the physiological location of the specific ligament to be replaced, must meet a number of functional challenges including biocompatibility, appropriate fixation to/within bones, and adequate mechanical strength and nutrient transport. Recent advances strongly suggest that tissue engineering will soon offer functional tissue replacements without many of the adverse side effects associated with current treatment options.

ANTERIOR CRUCIATE LIGAMENT INJURIES AND RECONSTRUCTIONS

Because knee injuries are common and often debilitating, much of the ligament tissue engineering research has focused on the knee. In the knee joint, the anterior cruciate ligament (ACL) prevents excessive anterior translation of the tibia and is crucial for normal joint function. Acute injuries to the ACL often result in instability, decreased joint function, and, over time, degenerative diseases such as osteoarthritis. This is of considerable concern because it has been estimated that as many as 150,000 Americans suffer an ACL injury each year,^[1] and the number of injuries appear to be on the rise.^[2] Furthermore, when fully or even partially disrupted, the ACL demonstrates a very poor capacity to heal, which often necessitates surgical reconstruction to restore function to the knee joint.^[2]

Long-term solutions to nonhealing ligament injury include surgical reconstructions with autografts, allografts, xenografts and synthetic grafts, or physical therapy. Surgical reconstruction is usually recommended for

active patients who suffer an ACL injury because the instability resulting from an ACL-deficient joint often precludes participation in rigorous sports or recreation activities. However, such surgical reconstructions are not an ideal long-term solution, as some studies have shown that patients who undergo ACL reconstructions have no better long-term results than those who opt not to have surgery.^[3] Autografts used in ACL reconstruction, most commonly the central third of the patellar tendon (PT), require prolonged recovery times, result in weakening of other knee structures and patellar pain, and still often lead to long-term degenerative problems such as osteoarthritis.^[4-6] Allografts and xenografts have had some success, but they are difficult and expensive to procure and still carry a risk of disease transmission.^[2,4] Some studies have also cited marked inflammatory and rejection responses to these grafts.^[1,6,7] Furthermore, required sterilization treatments may adversely affect the mechanical properties of allograft and xenograft tissues, and therefore lead to failure of the grafts.^[4] Synthetic materials such as Dacron (polyester) and Gore-Tex (PTFE) have also been used for ACL reconstructions, but long-term success rates have been poor due to inflammation and foreign-body reaction, material degradation, and eventual prosthesis failure.^[2,6]

Other Ligament Injuries

Injuries to ligaments other than the ACL often result in poor clinical outcomes, although not generally as severe as the long-term results associated with ACL injuries. Lateral ankle injuries, which are most frequently to the anterior talofibular ligament, account for approximately 25% of all musculoskeletal injuries with some 23,000 injuries daily in the United States.^[8] Treatments for lateral ankle sprains range from minimal to no treatment to surgery, with better outcomes for some remedies than others. However, regardless of treatment, many patients still complain of the ankle giving way and pain.^[9] Other studies of lateral ankle sprains find similar results. Ankle sprains often lead to persistent disability, including loss of function and pain.^[10]

The shoulder joint is also a common site for ligament injury. Within the joint, four major glenohumeral ligaments guide joint motion, and, most importantly, check excessive translations and rotations of the joint at extreme positions. Not surprisingly, glenohumeral ligaments tears have been implicated in acute shoulder joint dislocations.^[11] In fact, the glenohumeral joint is the most commonly dislocated joint in the body.^[12] Injuries to the glenohumeral ligaments often accompany damage to other soft tissues in the shoulder joint as the result of a dislocation and frequently result in long-term problems, most commonly increased joint laxity. Pathological joint laxity may result in chronic dislocations, which occur in young patients 66–95% of the time after an acute injury,^[13,14] and the necessity for surgery to prevent further dislocations. Although surgery often prevents further dislocations, almost all patients suffer a reduced range of shoulder motion. Other possible complications include osteoarthritis and recurrent pain, which may be severe enough to necessitate a shoulder arthroplasty.^[15]

Tissue Engineering for Ligament Replacement

Tissue engineering offers the possibility of creating functional soft tissue replacements to treat nonhealing ligament injuries without many of the adverse side effects associated with current reconstructive options. The fundamental basis of tissue engineering for ligament applications is the development of a biocompatible, resorbable scaffold that can be seeded with functioning cells of a relevant phenotype. Upon implantation, these viable cell/scaffold constructs should support joint function and, over time, gradually transfer mechanical loads born by the scaffold biomaterial to the newly developing neoligamentous tissue. Ultimately, the cell/scaffold construct should be replaced by biologically and mechanically functioning ligament tissue.

STRUCTURE AND FUNCTION OF NATURAL LIGAMENTS

Composition and Structure

Normal ligament is composed of approximately 64% water and combined with types I and III collagen, elastin, proteoglycans, glycosaminoglycans, fibroblasts, and small amounts of fibronectin. However, when dry weight is considered, collagen comprises some 70 to 80% of the tissue.^[16] Biochemical analyses have also shown that, of the collagen in ligament, 90%

is type I and approximately 10% is type III.^[17] Elastin composes about 5% of most ligaments and proteoglycans make up less than 1% of the tissue.^[17]

The hierarchical organization of collagen is perhaps the most important feature of ligament tissue with regard to creating functional tissue replacements since this multiscale architecture provides the structural support that defines the primary role of the ligament. At the lowest level of the collagen hierarchy is the α chain, a long polypeptide chain wound in a left-handed helix (Fig. 1). Three α chains are arranged in a larger right-handed triple helix to form a tropocollagen molecule, a rodlike structure approximately 1.5 nm in diameter and 300 nm in length. Tropocollagen molecules are packed together in a quarter stagger arrangement to form microfibrils, which are aligned roughly in parallel along the long axis of the ligament. Microfibrils aggregate to form fibers, and fibers form subfascicles. Subfascicles are surrounded by a connective tissue sheath called the endotenon.^[6,18] Fascicles, which are composed of subfascicles, are likewise surrounded by a connective tissue sheath called the epitenon.^[6] At the highest levels of the hierarchy, fascicles form the ligaments, which are, in turn, surrounded by a thicker structure of loose connective tissue called the paratenon that is enclosed by the synovial sheath.^[19]

T

Physiological Source of Mechanical Properties

Much of the characteristic strength of ligaments is attributed to intramolecular and intermolecular cross-linking within and among, respectively, the tropocollagen molecules that compose the ligament. Aldol cross-links, which constitute the intramolecular cross links, occur between aldehydes on the α chains of tropocollagen molecules.^[20] Conversely, intermolecular cross-links between adjacent tropocollagen molecules form as Schiff bases in the reaction of allysine with lysine or hydroxylysine.^[20]

A feature called crimp is generally considered to be responsible for the mechanical behavior observed for ligaments at low loads and extensions. Crimp refers to the roughly sinusoidal undulations or waviness present in many collagenous connective tissues, including ligament, when viewed under polarized light. Researchers have hypothesized that crimp is responsible for the characteristic toe-region or initial nonlinear portion of a ligament's load-elongation curve (Fig. 2) as crimp appears to straighten out when small loads are applied to the tissue. In this way, crimp enables the ligament to stretch a small amount before it develops substantial force. Presumably, this phenomenon allows the joint surfaces and ligaments to passively guide the motion of the joint through common activities, and for ligament stiffness to increase markedly under extreme

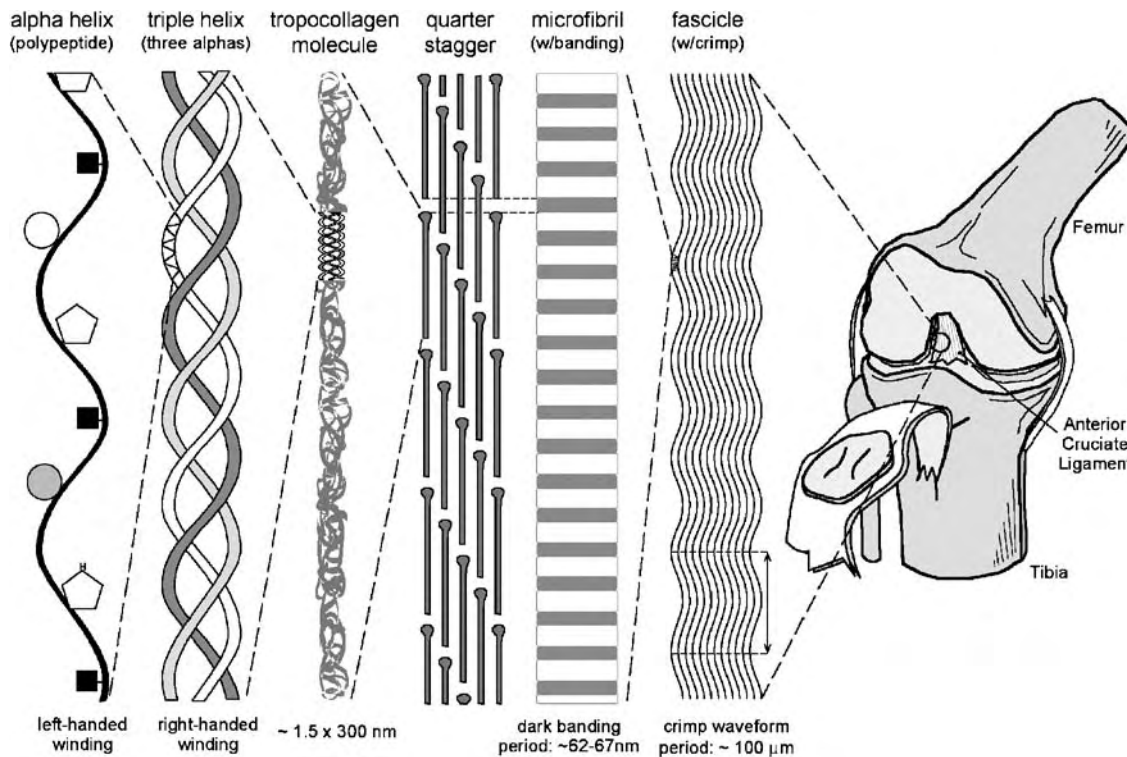


Fig. 1 Schematic diagram of the hierarchical organization present in collagenous materials, such as ligament, from polypeptide α chain to a functional structure in the knee joint. Human right knee shown with patella and patellar tendon reflected to allow visualization of the anterior cruciate ligament.

motions to protect the joint. It should be noted that crimp period and amplitude vary among tissues. Measurements of crimp within the rabbit ACL reveal a crimp period of 45–60 μm and an amplitude of 5 μm . Within the medial collateral ligament (MCL) however, the crimp period is 45 μm and the amplitude is 10 μm .^[20] Its importance to tissue engineering is not yet

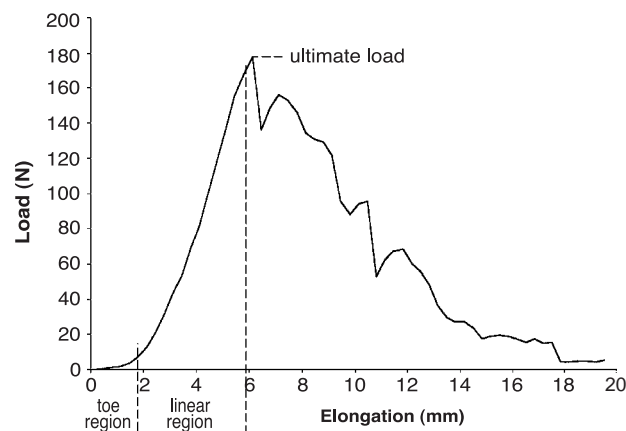


Fig. 2 Typical load-elongation curve for porcine femur-ACL-tibia complex demonstrating toe-region and ultimate failure. (Courtesy of James B. Raasch III.)

known. That is, it may be possible to use a synthetic scaffold with a nonlinear constitutive response to mimic the toe-region and subsequent stiffening of the ligament. But, it is also possible that crimp serves a physiological purpose related to mechanotransduction, cell-to-cell signaling, or nutrient transport that must be considered in the final design of a successful tissue-engineered ligament replacement.

Although the hierarchical collagen structure of ligaments has been a primary focus in the creation of functional ligament reconstructions,^[21] the ground substance in this tissue also appears to play a vital role in joint mechanics. In collagenous tissues, such as ligament, ground substance refers to the portion of the tissue containing the large polysaccharides, glycosaminoglycans and proteoglycans, and tissue fluid. Although the ground substance comprises just 1% of the dry weight of ligament, the majority of the water contained within the wet tissue is associated with the ground substance. Glycosaminoglycans are highly negatively charged and contain numerous hydroxyl groups that attract water through hydrogen bonding. It has long been hypothesized that the viscoelastic or time-dependent properties, i.e., creep and stress relaxation, of ligaments are due to complex water-ground substance interactions.

CELLS

Ligament fibroblasts are interspersed within fiber bundles along the tissue and possess numerous cellular organelles, which indicates high levels of metabolic activity.^[16] These fibroblasts make numerous contacts with nearby cells through fine processes that allow the cells to communicate via gap junctions.^[22] However, fibroblast morphology and size vary among ligaments, and these differences are often attributed to variations in the ligament's function and in-vivo environment. For example, fibroblasts within the MCL of the knee joint are rod- or spindle-shaped and have lengths of approximately 15 μm .^[20] Conversely, ACL fibroblasts, which resemble fibrocartilage cells, are round to ovoid in shape with diameters of 5–8 μm .^[20] Within ligaments, fibroblasts play a direct role in synthesizing collagen. α -chain helices are manufactured inside the cells and then passed through the cellular membrane and into the extracellular space where they are modified into tropocollagen molecules and then assembled into fibrils, as described previously. Thus, fibroblasts are responsible for creating the collagen that comprises the majority of the ligament tissue and in remodeling it both on a regular basis and after injury.

LIGAMENT-BONE INSERTIONS

Ligaments form junctions with bones at entheses or insertions. At these junctions, the ligament transitions from soft tissue to hard, bone tissue over a distance of less than one millimeter.^[23] Histological analyses have shown four distinct regions of the ligament insertion (Fig. 3). In a typical direct insertion, Zone 1 is ligament tissue and contains fibroblasts. Zone 2 is composed of fibrocartilage and contains fibrocartilage cells. Zone 2 is separated from Zone 3, which also contains fibrocartilage cells, by a distinctive tidemark, which indicates the presence of mineralized fibrocartilage. Zone 4 is composed of bone and contains characteristic osteocytes. The fibrocartilage cells of Zones 2 and 3 are oval in shape^[24] and do not communicate with other cells via gap junctions. This means that there is no communication between the osteocytes of the bone and the fibroblasts of the ligament. In addition to the communication barrier across the ligament insertion, there is also a vascular barrier. The uncalcified fibrocartilage zone of the ligament insertion does not contain blood vessels and therefore completely separates the vasculature of the ligament from that of the bone. It has been hypothesized that this bone-soft tissue interface, combined with the compression-tolerance properties of the fibrocartilage, allows for dissipation of stress in the joint and therefore protects the ligament from injury.^[25]

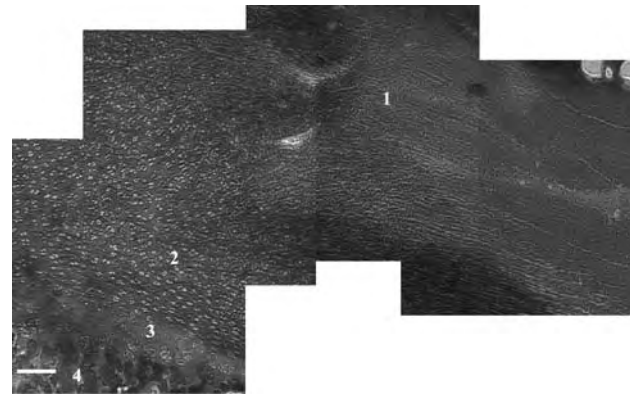


Fig. 3 Fibular insertion of a porcine lateral collateral ligament (LCL). Numbers denote the four zones of the insertion, as discussed in the text. Bone (at the lower left of the picture) stains red and blue, fibrocartilage stains light and dark blue, and ligament stains red (at the upper right). Scale bar = 150 μm . Phase contrast photograph. Modified Mallory's Stain. (Figure courtesy of Alycia Wanat.)

There are two distinct types of insertions found in the body, direct and indirect. The more common of the two, direct insertions, which are sometimes referred to as fibrocartilage entheses or chondral insertions, are most frequently observed in joints where the ligament attaches to the epiphyses of long bones.^[26] Direct insertions include both the tibial and femoral insertions of the ACL and the femoral insertion of the MCL. Indirect insertions, which are also known as fibrous entheses or periosteal insertions, tend to attach to the shafts of long bones.^[26] The tibial insertion of the MCL is an indirect insertion as it fans and attaches along the shaft of the tibia.

ROLES IN JOINT GUIDANCE

Ligaments traverse joints and connect bone to bone. The principal role of ligaments in the body is to guide joint motion; however, ligaments also play a critical role in stabilizing joints by preventing excessive movement that would result in injury. Some ligaments also provide structural support for internal organs. At low loads, the characteristic crimp structure of ligaments permits them to guide motion without placing undue stress on the soft tissue. Conversely, their toughness and strength allow them to restrain excessive movement and thus prevent joint damage at high loads. In the knee, two collateral and two cruciate ligaments guide joint motion and limit rotation and translation of the femur on the tibia in all possible directions of motion. In the glenohumeral joint,

ligaments play similar, crucial roles for normal function. Given the lack of bony constraint near the glenohumeral joint, the capsuloligamentous structures, i.e., ligaments that are intrinsically attached to the capsule, in concert with other soft tissue structures, play a vital role by guiding the joint through its wide range of motion without allowing excessive motion. Similar roles are attributed to ligaments in joints such as the elbow, ankle, wrist, and spine.

Mechanical Properties and Loading

Although significant efforts have been put forth to definitively establish the mechanical properties of ligaments of interest, particularly the ACL, controversy still remains. Some concerns arise from differences in testing protocols between research groups since specimen orientation, strain rate, hydration, and other factors affect the apparent mechanical properties of ligaments. However, even when testing conditions are consistent, many researchers have observed significant biological variability between specimens. Some of this variability has been conclusively attributed to factors such as donor age,^[27,28] physical immobilization, and exercise regimen.^[29,30] Noyes^[31] reported that the ACL supports about 445 N under normal in-vivo conditions and fails at approximately 1730 N. Woo^[28] found the linear stiffness of the human ACL from younger donors to be 242 ± 28 N/mm with an ultimate load of approximately 2160 N when tested along the anatomical axis of the specimen. However, these values were found to be significantly lower for both older specimens (Table 1) and specimens tested along the tibial axis, which was consistent with the results of other studies.^[34] In one of the few studies reporting mechanical properties, Noyes and Grood^[27] documented the elastic moduli for ACL specimens from younger and older donors to be 111 ± 26 MPa and 65.3 ± 24 MPa, respectively (Table 1).

Proprioception

In addition to their primary roles in joint stability and guidance of joint motion, ligaments also appear to play a proprioceptive role. That is, nerves within ligaments are involved in afferent functions that convey information to the brain regarding position, motion, and acceleration of the joint.^[32,35–38] Researchers who study the glenohumeral joint have indicated that proprioception may help prevent injuries by signaling appropriate muscular responses that actually protect capsuloligamentous structures in the joint.^[33,39] Numerous studies of the knee joint have also shown that patients with both ACL-deficient joints and ACL reconstructions have a poorer sense of joint position than patients with intact ligaments.^[40–42] The lack of a functional proprioceptive response may contribute to the inferior kinematics and function of the joint that are often implicated in degenerative diseases and poor long-term outcomes for patients with both reconstructed and ACL-deficient joints.

DESIGN CONSIDERATIONS FOR TISSUE-ENGINEERED LIGAMENTS

The European Society for Biomaterials defines biocompatibility as “the ability of a material to perform with an appropriate host response in a specific situation.”^[43] Normally, the concepts of biocompatibility and appropriate host response are framed only in terms of the chemistries of biological systems and materials. That is, a biomaterial that does not initiate a damaging inflammatory or coagulation response is generally considered to be biocompatible. However, for many tissues and tissue replacements, such as ligaments and ligament analogues, mechanical function and tissue integration are also essential components of acceptable performance. For example, an ACL replacement graft with incorrect mechanical

Table 1 Structural and mechanical properties of human ligaments

Ligament	Ultimate load (N)	Ultimate stress (MPa)	Linear stiffness (N/mm)	Elastic modulus (MPa)
Human ACL (16–26 years) ^[27]	1725 ± 269	37.8 ± 9.3	182 ± 33	111 ± 26
Human ACL (48–86 years) ^[27]	734 ± 266	13.3 ± 5.0	129 ± 39	65.3 ± 24.0
Human ACL (22–35 years) ^[28]	2160 ± 157	–	242 ± 28	–
Human ACL (60–97 years) ^[28]	658 ± 129	–	180 ± 25	–
Human LCL ^[61]	309 ± 91		58.1 ± 22.8	
Human Patellar Ligament (age 24.9 ± 4.4 years) ^[62]	–	53.4 ± 7.2	–	363.3 ± 93.7
Human MCL ^[63]	–	38.6 ± 4.8	–	332.2 ± 58.3

properties will cause joint dysfunction, which possibly leads to meniscal damage, further knee joint instability, and the initiation of osteoarthritis,^[44] all of which are certainly inappropriate host responses. Likewise, a tissue replacement that is chemically biocompatible but fails to integrate with the natural tissue cannot be functional in the long term. Put simply, a ligament replacement that does not create a strong fixation with the surrounding bone will not be able to support load. Furthermore, any engineered tissue that doesn't permit nutrient transport via revascularization will eventually develop necrotic regions.

Therefore, a more inclusive consideration of biocompatibility for tissue-engineered ligament replacements should include multiple aspects of appropriate biological function *in vivo*—appropriate chemical and ultrastructural composition, clinically-desirable interactions with surrounding cells and tissues yielding appropriate wound healing and tissue integration, and mechanical/structural properties that allow normal joint function.

Integration with Bone Tissue

The ability of the ligament analogue to quickly affix to bone *in vivo* is one of the most important aspects to consider when creating an effective tissue-engineered construct. Bone heals readily when damaged or grafted, which is a potentially advantageous trait when attempting to integrate an engineered tissue with the body's natural tissues. Without strong bone fixation to the implant, even a mechanically competent engineered ligament will fail because the function of this tissue is ultimately to transmit forces across joints. Ligament analogue-bone integration is therefore a fundamental aspect of appropriate biological function. Since bone heals well, a truly biocompatible construct will take advantage of this response and direct the initiation of bone growth in a way that quickly attaches the engineered tissue to the constituent bone. This can be achieved either by specifically incorporating a segment of bone tissue at the end of the implant *in vitro* or allowing the natural bone tissue to simply assimilate the engineered ligament tissue upon implantation.

Goulet et al.^[45] have attempted to create a bone-ligament interface in their tissue-engineered ligament constructs by casting a ligament fibroblast-seeded gel between partially demineralized human bone anchors. While this approach did yield constructs histologically similar to natural ligament, the functional biocompatibility associated with the ligament-bone integration approach was not assessed. That is, true functional biocompatibility would have been demonstrated by a strong and quick incorporation of the cell-seeded collagen gel with the bone anchors. While

mechanical testing results were not reported for this study, the authors reported a “lack of strength” as the main drawback of their work.^[46] Using a similar approach, Altman et al.^[47] placed coral and cancellous bone anchors at the ends of cell-seeded collagen gel constructs, but they also did not report resulting mechanical properties of the construct. As it turns out, physical placement of bone on the ends of tissue-engineered ligament constructs may not be necessary for appropriate bone fixation. In a series of studies in a rabbit model, Dunn et al.^[48] reported some bone growth into drilled femoral tunnels eight weeks after implantation of collagen fiber constructs despite the lack of any bone tissue-engineering on the fiber constructs prior to implantation.

Despite the critical nature of the bone-ligament interface, surprisingly little research has focused on designing insertion replacements. The type of insertion to be replaced, *i.e.*, direct or indirect, should guide the structure of any such replacements. For example, a relatively hard, porous material that is formed of bone, hydroxyapatite, metals or alloys, etc. would likely be the best choice for anchoring direct insertion replacements and mimicking the natural anchoring structure. A shallow bone tunnel filled with collagen fibers may suffice for indirect insertions. Regardless of the type of insertion to be replaced, functional integration with the surrounding bone and maintenance of mechanical functionality over time are critical aspects that must be achieved for ultimate clinical use.

Maintenance of Mechanical Strength Over Time

Much research on ligament tissue engineering to date has focused on creating a functional construct with regard to the initial mechanical properties. That is, research has focused on creating a ligament-like scaffold that has mechanical properties upon implantation that are sufficiently similar to the natural ligament that there need be no reliance on tissue ingrowth or *in-vivo* strengthening prior to the construct being able to bear loads. Without sufficient mechanical properties upon implantation, a tissue-engineered ligament could easily fail before the body has a chance to integrate or remodel the new tissue. However, simply immobilizing the joint after the implantation of a tissue-engineered ligament to allow for strengthening is not advisable since studies of autograft ACL reconstructions have suggested that patients can achieve better outcomes if the joint is mobilized as soon as possible after surgery.^[49] Concomitant with the requirement of sufficient initial mechanical strength is the need for a mechanically sound implant to maintain strength in the *in-vivo* environment while being remodeled and incorporated into the surrounding tissue. As this is

associated with the natural healing process, it is not only of concern for tissue-engineered constructs. Orthopedic surgeons report that patellar tendon autografts weaken at about six weeks postimplantation^[46] as the surrounding tissues remodel.

In a 1993 study, Dunn and colleagues^[50] created collagen fibers for ligament tissue engineering with the goal of finding the optimal fiber diameter and cross-linking technique to maximize fiber mechanical properties and degradation rates. They found that low-diameter (20–50 μm), dehydrothermal-cyanamide cross-linked fibers yielded high tensile strengths (75–110 MPa) but quick degradation rates (typically within four weeks) and a minimal inflammatory response upon implantation in a rabbit model. On the other hand, glutaraldehyde cross-linked scaffolds did not degrade. Instead, they elicited a strong inflammatory response. Altman et al.^[21] took a different approach to creating a ligament analogue with tensile properties similar to natural ligament by forming silk-fiber scaffolds. They reported ultimate tensile loads of these constructs to be 2337 N, which is similar to human ACL,^[27,28] however, the biocompatibility and degradation rates of these scaffolds were not assessed.

A number of studies have examined the use of polymer scaffolds in ligament tissue engineering. However, many researchers have avoided this approach because the degradation of these constructs can result in acidic byproducts, which can potentially injure cells and surrounding tissue. Lin et al.^[5] seeded ligament fibroblasts on treated polyglycolic acid (PGA) fibers (Dexon II) and reported increased cell proliferation after applying a mechanical stimulus, i.e., fluid mixing, and growth factors. Tissue formation was observed within the constructs due to fibroblast extracellular matrix secretion. No data were reported on the mechanical properties of these constructs, their chemical biocompatibility, or degradation rates. In a similar study, Cao et al.^[51] attempted to construct tissue-engineered tendons with fibroblast-seeded PGA scaffolds. When implanted into the subcutaneous space of nude mice, the constructs appeared histologically similar to natural tendon after six weeks and had mechanical properties on the same order as natural tendon.

Researchers have also attempted to create tissue-engineered ligaments from collagen gels. However, the major drawback of this approach is that, because constructs lack the fibrous structure present in natural ligament, it is very difficult to achieve sufficient tensile strength for ligament application. Huang et al.^[52] applied mechanical tensile loads to fibroblast-seeded collagen gels and did achieve success in cell-mediated strengthening of the constructs; however, the mechanical properties of the constructs remained orders of magnitude lower than those of a natural ligament.

Altman et al.^[47] and Goulet et al.^[45,46] have also attempted to create tissue-engineered ligaments by casting a cell-seeded collagen gel between bone anchors and subsequently applying a mechanical tensile stimulus to the construct. The resulting mechanical properties of these tissues were not reported.

Our research group has attempted a combined approach to ligament tissue engineering by merging reconstituted collagen fiber scaffolds with cell-seeded collagen gels to create a composite material that mimics the composition of natural tissue. In this model, the reconstituted collagen fibers act as the collagen fibers in the natural tissue by supporting tensile loads while the cell-seeded collagen gel acts as the ground substance of the ligament by retaining water and the constituent cells. We produced reconstituted collagen fibers in a similar manner to Dunn and colleagues^[44,48,50,53–57] and cross-linked them in a solution of 1% (w/v) 1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride in water. Fibers were then assembled into bundles and placed in a custom-built Teflon mold where they were surrounded with a cell-seeded collagen gel. After 25 days in culture, compared to nonseeded controls, the elastic modulus of the constructs had nearly doubled from 49.6 ± 3.3 MPa to 83.4 ± 10.8 MPa. These constructs have not yet been tested over long periods of time or in vivo (Fig. 4).

It should be noted that characterizing the modulus and strength of the ligament is not sufficient for ultimate clinical use. Because the toe-region (Fig. 2) of

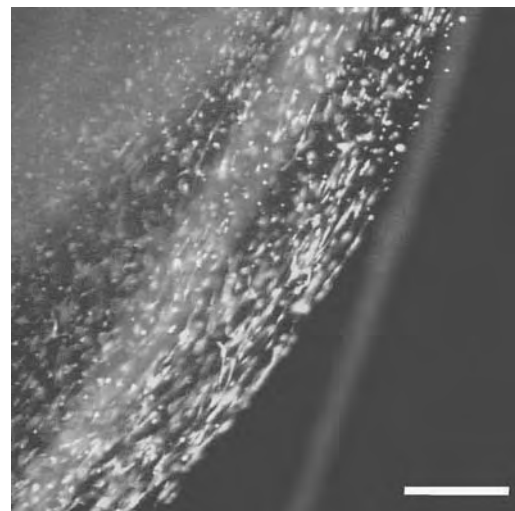


Fig. 4 Fluorescent micrograph of cell-seeded collagen gel/collagen fiber composite after four days in culture. Original seeding density was 5×10^5 cells/mL. Collagen fibers fluoresce a bright red color. Living cells are stained green. Live/Dead Viability/Cytotoxicity Kit, Molecular Probes, Eugene, OR. Scale bar = 200 μm .

ligament load-elongation behavior is thought to be critical to guiding joint motion, a tissue-engineered ligament analogue cannot be considered successful unless it responds appropriately to relatively small stretches. In this context, reproducing the toe-region seems to be a more important consideration than precisely matching the modulus of the linear portion of the stress-strain curve to an individual patient.

Allowing Vascularization and Nutrient Transport

As stated previously, bone is not the main source of nutrients for ligament since the insertion creates a communication and vascular barrier between the bone and ligament tissues. Therefore, an effective engineered ligament tissue would presumably have to rely on vasculature or nutrient transport from the synovium to support constituent cells, just as natural ligaments do. Such nutrient transport can potentially take place via two methods. The engineered tissue either must have a vascular structure that can quickly integrate with the native blood vessels, or must be permeable enough upon implantation that constituent cells can survive via diffusion of nutrients from the synovial fluid while angiogenesis of the engineered tissue takes place. Both of these vascularization methods are biocompatibility issues. An engineered tissue that is well accepted by the body will be incorporated and vascularized whereas a nonbiocompatible tissue will experience degradation and fibrous encapsulation that will inhibit angiogenesis. Very few ligament tissue engineering studies have examined the issues of cell viability in analogues implanted in an animal model or the ingrowth of vasculature upon implantation. Bellincampi et al.^[58] implanted cell-seeded collagen fibers scaffolds in the knee joints of rabbits and observed seeded cells on the surface of collagen fiber scaffolds as late as six weeks after implantation, but very few cells were present between adjacent fibers. This seems to indicate that cells could survive on the surface of the implant because they were in close contact with nutrients from the synovial fluid. Cells within the scaffolds may not have been able to survive to the same extent because the scaffold was neither particularly permeable nor did it contain a microstructure to act as vasculature. Weadock et al.^[59] examined tissue ingrowth on collagen films treated with different cross-linking techniques in a rat model and found that cyanamide and severely dehydrated cross-linking techniques allowed for capillary growth in collagen implants within 14 days.

Avoiding Undue Inflammation

A major challenge of ligament tissue engineering is achieving a balance between creating a construct that

has sufficient mechanical properties to function effectively in a joint while not eliciting an inflammatory response. Much research has focused on using collagen fiber constructs to achieve this goal. Collagen fibers are an obvious choice because they mimic the structure of the natural tissue, they're biodegradable,^[56,59] and they have a low antigenicity. However, collagen fiber constructs require chemical cross-linking to both mimic cross-linking within the collagen of the natural tissue and provide strength.

In a series of papers focused on ligament tissue engineering, Dunn and coworkers^[48,53,54,60] attempted to strengthen their constructs by using glutaraldehyde cross-linking. However, they report that, upon implantation, the collagen fiber constructs "appeared to have caused an acute inflammatory response." Other observations included fibrous encapsulation of the implant^[53,60] and the presence of mononuclear and multinuclear inflammatory cells.^[48,53,54] Wasserman et al.^[57] found similar results for dehydrothermal and cyanamide cross-linked collagen fiber scaffolds implanted in tendon defects of rabbits. Cross-linking collagen fiber scaffolds with ultraviolet light led to scaffolds that did not initiate an inflammatory response, i.e., seeded cells survived on them for at least four weeks when implanted in the knee joints of rabbits, but they were completely resorbed after eight weeks.^[58] Carbodiimide cross-linking of type I collagen has yielded conflicting results with regard to initiation of an inflammatory response. Dunn et al.^[53] report infiltration of their carbodiimide cross-linked fiber prosthesis used to reconstruct a rabbit ACL by inflammatory cells. On the other hand, Kato et al.^[60] performed a similar study, but they implanted the prostheses in defects created in rabbit Achilles tendons without observing an inflammatory response. Weadock et al.^[59] concluded that carbodiimide should be chemically biocompatible because the only byproduct of its degradation is water-soluble urea.

Appearing Immunologically Acceptable

Tissue engineering efforts for ligament applications should develop not only a scaffolding material to provide functional support for the joint but also a means to remodel and incorporate that tissue with constituent cells once in vivo. Therefore, an ideal replacement would neither elicit an immune response because of the scaffolding material nor because of the constituent cells. The body's immune system recognizes any implanted material—whether it be cells, natural proteins, or a synthetic material—as self or nonself. Even among members of the same species, the immune system can distinguish self-cells from

nonselself-cells. Therefore, without some means to repress or deceive the immune system, the only way to incorporate constituent cells into a tissue-engineered ligament is to use the patient's own cells. However, the means to carry out this procedure has received very little attention because so many questions remain to be answered. For example, is it necessary for donor cells to come from the same tissue that they intend to replace? That is, are ACL fibroblasts necessary to tissue engineer an ACL? Bellincampi et al.^[58] used both ACL fibroblasts and skin fibroblasts to seed collagen fiber scaffolds for ligament tissue engineering. These constructs were later implanted in the knee joints of rabbits, and cell viability was assessed at up to eight weeks. In this study, the researchers removed the rabbits' ACLs to obtain the fibroblasts that were later seeded on the collagen fiber constructs. This would not be ideal with a human patient; however, as two surgical procedures—one to collect cells and the other to implant the cell-seeded construct—would be required. Furthermore, a patient with an older ACL injury may have no tissue because of resorption, and thus no ACL fibroblasts remaining in his or her joint. Although Bellincampi's data suggest that fibroblast source has little influence on cell viability once implanted in the joint, since fibroblasts from the rabbit ACL fared no better than skin fibroblasts upon implantation in the rabbit knee joint, more studies are needed to conclusively answer questions about appropriate cell sources for human tissue replacements.

Since harvesting cells from a patient's ligament tissue or even via a skin biopsy for example are not ideal solutions for obtaining cells to make a tissue-engineered construct, one logical step to avoid an immune response is to use cell lines from another species that have been transformed to be nonimmunogenic. Stem cells are another potential source for non-immunogenic cells for ligament tissue engineering. Both embryonic and adult stem cells offer the possibility of creating a tissue-engineered ligament with nonimmunogenic cells that are also morphologically and functionally ligament fibroblasts. Adult stem cells are fairly easily obtained from either the peripheral blood or the bone marrow of an adult's posterior iliac crest without significant discomfort, harvest site morbidity, or surgical invasion. These adult stem cells have the potential to differentiate into fibroblasts that could then be seeded on a construct to be subsequently implanted in a joint to repair a ligament. On the other hand, embryonic stem cells would have to be harvested from another member of the same species. While these cells have greater potential for differentiation, without some sort of intervention, they would still pose the same immunogenic difficulties that are encountered whenever a tissue transplant takes place. Furthermore, ethical issues involved with these cell lines.

CONCLUSION

As many as 150,000 Americans injure their ACL every year.^[1] Unfortunately, due to poor vascularization and the substantial mechanical demands to which this tissue is subjected, the ACL does not heal of its own accord. Moreover, surgical reconstruction with tissue replacements in the form of patellar tendon or hamstring grafts require long recuperation times. Similar problems are associated with the posterior cruciate ligament in the knee as well as ligaments in the ankle, shoulder, and elbow. Tissue engineering offers the possibility of creating functional tissue replacements to treat these types of injuries without many of the adverse side effects associated with current reconstructive options. The fundamental basis of ligament tissue engineering is the development of a biocompatible, resorbable, cell-seeded scaffold that, once implanted, can provide functional support while gradually transferring the mechanical loads born by the scaffold to neoligament tissue. Recent work indicates that cell-seeded collagen or silk fibers organized into a three-dimensional matrix can provide the appropriate initial stiffness. Skin biopsies or adult stem cell-derived fibroblasts yield host-specific cells, which thereby obviates the concern over the body's immune response. However, more research is required to optimize fixation of the ligament analogue to the native bone and find a means of introducing a nutrient supply. Likewise, the physiological role of crimp should be investigated and elucidated. These issues are all critical aspects of biocompatibility and must be solved in order to engineer functional ligament analogues.

ARTICLES OF FURTHER INTEREST

Bioreactors; Host Reactions; Tendons and Ligaments, Mechanical Testing Of; Tissue Engineering of Bone; Tissue Engineering of Microvascular Networks; Tissue Engineering Scaffolds; Tissue Engineering of Tendons

REFERENCES

1. Silver, F.H. *Biomaterials, Medical Devices and Tissue Engineering: An Integrated Approach*, 1st Ed.; Chapman & Hall: London, 1994.
2. Koski, J.A.; Ibarra, C.; Rodeo, S.A. Tissue-engineered ligament: Cells, matrix, and growth factors. *Orthop. Clin. North Am.* **2000**, *31*, 437–452.
3. Daniel, D.M.; Stone, M.L.; Dobson, B.E.; Fithian, D.C.; Rossman, D.J.; Kaufman, K.R. Fate of the ACL-injured

- patient. A prospective outcome study. *Am. J. Sports Med.* **1994**, *22*, 632–644.
4. Johnson, R.J.; Beynon, B.D.; Nichols, C.E.; Renstrom, P.A. The treatment of injuries of the anterior cruciate ligament. *J. Bone Jt. Surg., Am.* **1992**, *74*, 140–151.
 5. Lin, V.S.; Lee, M.C.; O'Neal, S.; McKean, J.; Sung, K.L. Ligament tissue engineering using synthetic biodegradable fiber scaffolds. *Tissue Eng.* **1999**, *5*, 443–452.
 6. Silver, F.H.; Tria, A.J.; Zawadsky, J.P.; Dunn, M.G. Anterior cruciate ligament replacement: A review. *J. Long-Term Eff. Med. Implants* **1991**, *1*, 135–154.
 7. Arnoczky, S.P.; Warren, R.F.; Ashlock, M.A. Replacement of the anterior cruciate ligament using a patellar tendon allograft. An experimental study. *J. Bone Jt. Surg., Am.* **1986**, *68*, 376–385.
 8. Robins, C.; Seaton, T.L. What is the optimal treatment for lateral ankle ligament ruptures? *J. Fam. Pract.* **2000**, *49*, 885.
 9. Pijnenburg, A.C.; Van Dijk, C.N.; Bossuyt, P.M.; Marti, R.K. Treatment of ruptures of the lateral ankle ligaments: A meta-analysis. *J. Bone Jt. Surg., Am.* **2000**, *82*, 761–773.
 10. Gerber, J.P.; Williams, G.N.; Scoville, C.R.; Arciero, R.A.; Taylor, D.C. Persistent disability associated with ankle sprains: A prospective examination of an athletic population. *Foot Ankle Int.* **1998**, *19*, 653–660.
 11. Symeonides, P.P. The significance of the subscapularis muscle in the pathogenesis of recurrent anterior dislocation of the shoulder. *J. Bone Jt. Surg., Br.* **1972**, *54*, 476–483.
 12. Wirth, M.A.; Rockwood, C.A. Traumatic Glenohumeral Instability: Pathology and Pathogenesis. In *The Shoulder: A Balance of Mobility and Stability*; Matsen, F.A., Fu, F.H., Hawkins, R.J., Eds.; American Academy of Orthopaedic Surgeons: Rosemont, IL, 1993.
 13. McLaughlin, H.L.; McLellan, D.I. Recurrent anterior dislocation of the shoulder: II. A comparative study. *J. Trauma* **1967**, *7*, 191–201.
 14. Simonet, W.T.; Cofield, R.H. Prognosis in anterior shoulder dislocation. *Am. J. Sports Med.* **1984**, *12*, 19–24.
 15. Lusardi, D.A.; Wirth, M.A.; Wurtz, D.; Rockwood, C.A., Jr. Loss of external rotation following anterior capsulorrhaphy of the shoulder. *J. Bone Jt. Surg., Am.* **1993**, *75*, 1185–1192.
 16. Amiel, D.; Frank, C.; Harwood, F.; Fronck, J.; Akeson, W. Tendons and ligaments: A morphological and biochemical comparison. *J. Orthop. Res.* **1984**, *1*, 257–265.
 17. Frank, C.; Woo, S.L.; Andriacchi, T.; Oakes, B.; Dahners, L.; DeHaven, K.; Lewis, J.; Sabiston, P. Normal Ligament: Structure, Function and Composition. In *Injury and Repair of the Musculooskeletal Soft Tissues*; Woo, S.L.-Y., Buckwalter, J.A., Eds.; American Academy of Orthopaedic Surgeons: Park Ridge, IL, 1988.
 18. Danylchuk, K.D.; Finlay, J.B.; Krcek, J.P. Microstructural organization of human and bovine cruciate ligaments. *Clin. Orthop.* **1978**, *131*, 294–298.
 19. Smith, B.A.; Livesay, G.A.; Fu, F.H. Biomechanics of the Knee. In *Ligaments of the Knee*; Tria, A.J., Ed.; Churchill Livingstone: New York, 1995.
 20. Amiel, D.A.; Billings, E., Jr; Akeson, W.H. Ligament Structure, Chemistry and Physiology. In *Knee Ligaments: Structure, Function, Injury and Repair*; Akeson, W.H., O'Connor, J., Daniel, D.M., Eds.; Raven Press, 1990.
 21. Altman, G.H.; Horan, R.L.; Lu, H.H.; Moreau, J.; Martin, I.; Richmond, J.C.; Kaplan, D.L. Silk matrix for tissue engineered anterior cruciate ligaments. *Bio-materials* **2002**, *23*, 4131–4141.
 22. Ralphs, J.R.; Benjamin, M. Cell and Matrix Organization in Tendons and Ligaments. In *Biology of the Synovial Joint*; Archer, C.W., Caterson, B., Benjamin, M., Ralphs, J.R., Eds.; Harwood Academic Publishers, 1999.
 23. Smith, B.A.; Livesay, G.A.; Woo, S.L. Biology and biomechanics of the anterior cruciate ligament. *Clin. Sports Med.* **1993**, *12*, 637–670.
 24. Cooper, R.R.; Misol, S. Tendon and ligament insertion. A light and electron microscopic study. *J. Bone Jt. Surg., Am.* **1970**, *52*, 1–20.
 25. Woo, S.L.; Maynard, J.; Butler, D.; Lyon, R.; Torzilli, P.; Akeson, W.; Cooper, R.; Oakes, B. Ligament, Tendon and Joint Capsule Insertions to Bone. In *Injury and Repair of the Musculooskeletal Soft Tissues*; Woo, S.L., Buckwalter, J.A., Eds.; Am. Acad. Orthop. Surg.: Park Ridge, IL, 1988.
 26. Benjamin, M.; Ralphs, J.R. The cell and developmental biology of tendons and ligaments. *Int. Rev. Cyt.* **2000**, *196*, 85–130.
 27. Noyes, F.R.; Grood, E.S. The strength of the anterior cruciate ligament in humans and Rhesus monkeys. *J. Bone Jt. Surg., Am.* **1976**, *58*, 1074–1082.
 28. Woo, S.L.; Hollis, J.M.; Adams, D.J.; Lyon, R.M.; Takai, S. Tensile properties of the human femur-anterior cruciate ligament-tibia complex. The effects of specimen age and orientation. *Am. J. Sports Med.* **1991**, *19*, 217–225.
 29. Tipton, C.M.; James, S.L.; Mergner, W.; Tchong, T.K. Influence of exercise on strength of medial collateral knee ligaments of dogs. *Am. J. Physiol.* **1970**, *218*, 894–902.
 30. Woo, S.L.; Gomez, M.A.; Sites, T.J.; Newton, P.O.; Orlando, C.A.; Akeson, W.H. The biomechanical and morphological changes in the medial collateral ligament of the rabbit after immobilization and remobilization. *J. Bone Jt. Surg., Am.* **1987**, *69*, 1200–1211.
 31. Noyes, F.R.; Butler, D.L.; Grood, E.S.; Zernicke, R.F.; Hefzy, M.S. Biomechanical analysis of human ligament grafts used in knee-ligament repairs and reconstructions. *J. Bone Jt. Surg., Am.* **1984**, *66*, 344–352.
 32. Griffin, L.Y.; Agel, J.; Albohm, M.J.; et al. Noncontact anterior cruciate ligament injuries: Risk factors and prevention strategies. *J. Am. Acad. Orthop. Surg.* **2000**, *8*, 141–150.
 33. Grigg, P. The Role Of Capsular Feedback And Pattern Generators In Shoulder Kinematics. In *The Shoulder: A Balance of Mobility and Stability*; Matsen, F.A., Fu, F.H., Hawkins, R.J., Eds.; American Academy of Orthopaedic Surgeons: Rosemont, IL, 1993.
 34. Butler, D.L.; Grood, E.S.; Noyes, F.R.; Sodd, A.N. On the interpretation of our anterior cruciate ligament data. *Clin. Orthop.* **1985**, *196*, 26–34.

35. Doukas, W.C.; Speer, K.P. Anatomy, pathophysiology, and biomechanics of shoulder instability. *Orthop. Clin. North Am.* **2001**, *32*, pp. vii, 381–391.
36. Lephart, S.M.; Kocher, M.S. The Role of Exercise in the Prevention of Shoulder Disorders. In *The Shoulder: A Balance of Mobility and Stability*; Matsen, F.A., Fu, F.H., Hawkins, R.J., Eds.; American Academy of Orthopaedic Surgeons: Rosemont, IL, 1993.
37. Schultz, R.A.; Miller, D.C.; Kerr, C.S.; Micheli, L. Mechanoreceptors in human cruciate ligaments. A histological study. *J. Bone Jt. Surg., Am.* **1984**, *66*, 1072–1076.
38. Schutte, M.J.; Dabezies, E.J.; Zimny, M.L.; Happel, L.T. Neural anatomy of the human anterior cruciate ligament. *J. Bone Jt. Surg., Am.* **1987**, *69*, 243–247.
39. Myers, J.B.; Lephart, S.M. Sensorimotor deficits contributing to glenohumeral instability. *Clin. Orthop.* **2002**, *400*, 98–104.
40. Fischer-Rasmussen, T.; Jensen, P.E. Proprioceptive sensitivity and performance in anterior cruciate ligament-deficient knee joints. *Scand. J. Med. Sci. Sports* **2000**, *10*, 85–89.
41. Fremerey, R.W.; Lobenhoffer, P.; Zeichen, J.; Skutek, M.; Bosch, U.; Tscherne, H. Proprioception after rehabilitation and reconstruction in knees with deficiency of the anterior cruciate ligament: A prospective, longitudinal study. *J. Bone Jt. Surg., Br.* **2000**, *82*, 801–806.
42. Friden, T.; Roberts, D.; Ageberg, E.; Walden, M.; Zatterstrom, R. Review of knee proprioception and the relation to extremity function after an anterior cruciate ligament rupture. *J. Orthop. Sports Phys. Ther.* **2001**, *31*, 567–576.
43. Black, J. In *Biological Performance of Materials: Fundamentals of Biocompatibility*, 2nd Ed.; Marcel Dekker: New York, NY, 1992.
44. Dunn, M.G.; Liesch, J.B.; Tiku, M.L.; Zawadsky, J.P. Development of fibroblast-seeded ligament analogs for ACL reconstruction. *J. Biomed. Mater. Res.* **1995**, *29*, 1363–1371.
45. Goulet, F.; Rancourt, D.; Cloutier, R.; Germain, L.; Poole, A.R.; Auger, F.A. Tendons and Ligaments. In *Principles of Tissue Engineering*, 2nd Ed.; Lanza, R.P., Langer, R., Vacanti, J., Eds.; Academic Press: San Diego, 2000.
46. Goulet, F.; Germain, L.; Rancourt, D.; Caron, C.; Normand, A.; Auger, F.A. Tendons and Ligaments. In *Principles of Tissue Engineering*; Lanza, R.P., Langer, R., Chick, W.L., Eds.; Academic Press: San Diego, 1997.
47. Altman, G.H.; Horan, R.L.; Martin, I.; Farhadi, J.; Stark, P.R.; Volloch, V.; Richmond, J.C.; Vunjak-Novakovic, G.; Kaplan, D.L. Cell differentiation by mechanical stress. *FASEB J.* **2002**, *16*, 270–272.
48. Dunn, M.G.; Maxian, S.H.; Zawadsky, J.P. Intraosseous incorporation of composite collagen prostheses designed for ligament reconstruction. *J. Orthop. Res.* **1994**, *12*, 128–137.
49. McCarthy, M.R.; Yates, C.K.; Anderson, M.A.; Yates-McCarthy, J.L. The effects of immediate continuous passive motion on pain during the inflammatory phase of soft tissue healing following anterior cruciate ligament reconstruction. *J. Orthop. Sports Phys. Ther.* **1993**, *17*, 96–101.
50. Dunn, M.G.; Avasarala, P.N.; Zawadsky, J.P. Optimization of extruded collagen fibers for ACL reconstruction. *J. Biomed. Mater. Res.* **1993**, *27*, 1545–1552.
51. Cao, Y.; Vacanti, J.P.; Ma, X.; Paige, K.T.; Upton, J.; Chowanski, Z.; Schloo, B.; Langer, R.; Vacanti, C.A. Generation of neo-tendon using synthetic polymers seeded with tenocytes. *Transplant. Proc.* **1994**, *26*, 3390–3392.
52. Huang, D.; Chang, T.R.; Aggarwal, A.; Lee, R.C.; Ehrlich, H.P. Mechanisms and dynamics of mechanical strengthening in ligament-equivalent fibroblast-populated collagen matrices. *Ann. Biomed. Eng.* **1993**, *21*, 289–305.
53. Dunn, M.G.; Tria, A.J.; Kato, Y.P.; Bechler, J.R.; Ochner, R.S.; Zawadsky, J.P.; Silver, F.H. Anterior cruciate ligament reconstruction using a composite collagenous prosthesis. A biomechanical and histologic study in rabbits. *Am. J. Sports Med.* **1992**, *20*, 507–515.
54. Goldstein, J.D.; Tria, A.J.; Zawadsky, J.P.; Kato, Y.P.; Christiansen, D.; Silver, F.H. Development of a reconstituted collagen tendon prosthesis. *J. Bone Jt. Surg.* **1989**, *71*, *AM.*, 1183–1191.
55. Kato, Y.P.; Christiansen, D.L.; Hahn, R.A.; Shieh, S.J.; Goldstein, J.D.; Silver, F.H. Mechanical properties of collagen fibres: A comparison of reconstituted and rat tail tendon fibres. *Biomaterials* **1989**, *10*, 38–42.
56. Law, J.K.; Parsons, J.R.; Silver, F.H.; Weiss, A.B. An evaluation of purified reconstituted Type I collagen fibers. *J. Biomed. Mater. Res.* **1989**, *23*, 961–977.
57. Wasserman, A.J.; Kato, Y.P.; Christiansen, D.; Dunn, M.G.; Silver, F.H. Achilles tendon replacement by a collagen fiber prosthesis: Morphological evaluation of neotendon formation. *Scanning Microsc.* **1989**, *3*, 1183–1200.
58. Bellincampi, L.D.; Closkey, R.F.; Prasad, R.; Zawadsky, J.P.; Dunn, M.G. Viability of fibroblast-seeded ligament analogs after autogenous implantation. *J. Orthop. Res.* **1998**, *16*, 414–420.
59. Weadock, K.; Olson, R.M.; Silver, F.H. Evaluation of collagen crosslinking techniques. *Biomater. Med. Dev. Artif. Organs* **1983**, *11*, 293–318.
60. Kato, Y.P.; Dunn, M.G.; Zawadsky, J.P.; Tria, A.J.; Silver, F.H. Regeneration of Achilles tendon with a collagen tendon prosthesis. Results of a one-year implantation study. *J. Bone Jt. Surg., Am.* **1991**, *73*, 561–574.
61. Sugita, T.; Amis, A.A. Anatomic and biomechanical study of the lateral collateral and popliteofibular ligaments. *Am. J. Sports Med.* **2001**, *29*, 466–472.
62. Staubli, H.U.; Schatzmann, L.; Brunner, P.; Rincon, L.; Nolte, L.P. Mechanical tensile properties of the quadriceps tendon and patellar ligament in young adults. *Am. J. Sports Med.* **1999**, *27*, 27–34.
63. Quapp, K.M.; Weiss, J.A. Material characterization of the human medial collateral ligament. *J. Biomech. Eng.* **1998**, *120*, 757–763.

Tissue Engineering of Liver

Alexander B. G. Sevy

Katherine M. Kulig

Massachusetts General Hospital, Boston, Massachusetts, U.S.A.

Joseph P. Vacanti

Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts, U.S.A.

T

INTRODUCTION

Liver-tissue engineering has the potential to prevent and treat liver diseases in a large number of patients worldwide. Hepatic failure can be caused by viral infection, drug reactions, toxins, and vascular complications. Currently, the only treatment for liver disease is orthotopic liver transplantation. However, a shortage of donors has led to a growing waiting list for those in need of hepatic transplantation. In the United States alone, from July 2001 to June 2002, only 5261 of 17,379 patients on the waiting list received a liver transplant, while 1861 died while waiting.^[1] Liver failure also generates a cost of \$ 9 billion each year in the United States.^[2]

The clinical options for hepatic transplantation include cadaveric whole-organ transplantation, reduced-size liver transplantation (RSLT), and split-liver transplantation. Recently, living donor transplantation in children and adults has been introduced. Because only 10–20% of the liver is required for life-sustaining function, and because the liver has regenerative properties *in vivo*, one donor liver can be divided to treat two patients. Because livers are not genetically identical, immunosuppressive drugs are required throughout the transplant recipient's lifetime. Despite vast improvements in therapy and antirejection medications, the immunosuppressant drugs often cause toxic damage to the liver and kidneys, requiring another transplant in later years.

OVERVIEW

The low ratio of available organs to needed organs has sparked the investigation of alternate treatments that would tackle not only the problems of entire liver replacement, but also the onset of liver disease and genetic disorders. Scientists have developed bioreactors to study the growth and function of hepatocytes in three-dimensional configurations. The bioreactor is a continuous flow system that allows for nutrients and oxygen to reach the hepatocytes seeded on a polymer scaffold while carbon dioxide and wastes are removed.

Bioreactors increase the survival and function of hepatocytes by acting as a circulation support system similar to that of the liver. Liver-assist devices (LADs) employ bioreactor advances by using hollow-fiber cartridges in which hepatocytes are seeded to treat liver failure patients in the short term.

The ultimate long-term alternative to transplantation may be an engineered whole-liver replacement. Our laboratory has used the bioreactor paradigm to design a microfabricated liver system that functions as an implantable whole liver replacement. This approach utilizes MicroElectroMechanical Systems (MEMS) technology to etch the vascular network necessary to support viability and function of hepatocytes *in vitro*.^[3] Our group has endothelialized this vascular bed, which can support hepatocytes seeded within an adjacent parenchymal chamber.^[4]

The multidisciplinary field of hepatic-tissue engineering involves cellular developmental biology, molecular biochemistry, polymer chemistry, physics, and other engineering sciences like computational biology and MEMS. Synergy has improved within tissue engineering as researchers make and employ advances in these varied areas. Due to the extensive scope of the field, of hepatic tissue engineering, this article is unable to incorporate every investigator's work and subfield, but presents a brief view of the background, accomplishments, and future prospects of the field.

THE HEPATIC ENVIRONMENT

The liver is a complex organ composed of endothelial cells and parenchyma cells such as hepatocytes, bile duct cells, stellate cells, oval cells (Ito cells), fibroblasts, and Kupffer cells.^[5,6] Most tissue-engineering research focuses on the hepatocytes that constitute 60% of the liver.^[3] This functional hepatic mass carries out toxin removal; metabolic function, including P-450 activity, glycogen storage, and gluconeogenesis for blood glucose regulation; synthesis of urea; and release of proteins, carbohydrates, and lipids.^[7,8] Many functional assays have determined that hepatocytes—both mature

and progenitor cells—exhibit certain functions such as albumin production when perfused from the liver or from a cell line, and cultured in both static^[9,10] and continuous flow systems.^[11,12]

The hepatocyte environment has many important factors that influence the cells' growth and functional, divisional, and apoptotic properties. However, a combination of components is required to alter the pathway of the hepatocytes.^[13,14] Hepatocytes will exhibit normal functions in a homogenous culture with added mitogens but will exhibit increased functional activity when cocultured with hepatic associated cells, which mimic their natural environment. This recreated hepatic cellular environment fosters heterotypic and homotypic cell–cell junctions and allows the cells to communicate via intercellular matrix interactions.^[15] Communication also occurs through insoluble molecules found in the extracellular matrix (ECM).^[13] The surrounding cells and matrix play an intricate role in determining through which pathway the cells will proceed.^[16] Hepatocytes are able to turn on growth and differentiation genes depending on with which ECM molecules the cells adhere or are in contact. These include laminin and type IV collagen (basement membrane), fibronectin, and type I collagen (interstitial matrix).^[13] In addition, binding of cell adhesion receptors (e.g., integrins) can lead to mechanical stresses that affect signaling pathways controlling cells' growth, differentiation, motility, and apoptosis.^[17] Pathways of replication and function are also dependent on the size and shape of the colony. At low densities of cells within a colony, the cells have a tendency to proliferate, expanding the colony to higher densities where they will differentiate and carry out normal metabolic functions.^[14]

Hepatic Interactions

Hepatocytes require external factors including environmental markers, growth factors, and cell signaling from the ECM. Researchers integrate these various environmental factors in culture systems and investigate their effects on growth and differentiation and function of cultured hepatocytes. Many of these growth factors and mediators work in a series of or in combination with other factors in initiating cellular pathways. Factors that are known to control the growth and proliferation of the cell include Epidermal Growth Factor (EGF), Fibroblastic Growth Factor (FGF),^[18] Transforming Growth Factor alpha (TGF- α), Interleukin-6 (IL-6),^[19] and Hepatocyte Growth Factor (HGF).^[10,20,21] Differentiation and function are stimulated by dimethyl sulfoxide (DMSO), nicotinic acid,^[22] corticosteroids, and amino acids.^[19] HGF and Vascular Endothelial Growth Factor (VEGF) mediate

angiogenesis. In addition to proliferation and angiogenesis, cell motility is also mediated by HGF.^[23] These signals and growth factors are among the many that contribute to cell signaling.

Hepatic Cell Source

Despite all the known factors that promote cell growth and differentiation, it is important to choose a source of cells with predictable phenotype and robust activity of desired traits. The forerunners for this role in the field of hepatic-tissue engineering are mature hepatocytes, oval/progenitor cells (primary hepatocytes, small hepatocytes)^[10,24] hematopoietic cells,^[24,25] and stem cells.^[25]

Mature hepatocytes currently in use are immortalized carcinoma cell lines such as SV40 Tag^[5,25] and HepG2.^[24,25] There are advantages in using an immortalized cell line: They are of a human cell line, exhibit a high degree of differentiated hepatocyte function, and are easy to culture in large quantities.^[24,25] It is imperative to maintain safety when using cell lines, and to avoid exposing patients to tumor-derived cells.^[5,25] Studies have shown that rat primary hepatocytes outperform human mature hepatocytes (HepG2) in regard to function and differentiation.^[24] However, obtaining a pure culture of primary rat hepatocytes is problematic, and for long-term survival it is difficult to extend the life of the rat primary hepatocyte through several divisions.^[10] Human primary hepatocytes have proven to be even more challenging to maintain in culture.^[25,26] Stem cells are being considered for their ability to differentiate into a normal liver cell exhibiting hepatic functions while maintaining the proliferating and differentiating abilities of a parenchymal cell.^[5,25] Evidence has shown that hepatic stem cells can be derived from other sources, such as the bone marrow,^[26,27] in addition to those found within the liver (oval cells). The sequences and combinations of signals needed to cue these cells into growth and differentiation are currently being studied.

Also debated as a source of cells are those cells used for xenotransplantation. Porcine hepatocytes have been shown to exhibit qualities and functions similar to human hepatocytes, such as urea synthesis, albumin production, and P-450 cytochrome activity, with the same therapeutic effects.^[8,28] Porcine hepatocytes have the clear advantage of being readily available (unlike human hepatocytes, which are scarce due to the lack of donor organs).^[8] There are many risks in using animal cells, however including the difficulty of overcoming immune rejection and the chance that zoonosis from viruses may transfer to the patient.^[25]

Coculture of Hepatocytes

The interactions between hepatocytes and cocultured cells can alter the hepatocytes' phenotype and stimulate or inhibit specific functional pathways. Rat primary hepatocytes cultured with fibroblasts^[19] or endothelial cells^[15] were able to maintain their viability and function. Cocultures with bone marrow stromal cells^[29] or nonparenchymal cells^[20] facilitated proliferation and differentiation. Epithelial cell and mesenchymal cell coculture have promoted the differentiation of hepatocytes.^[19] Studies have consistently shown beneficial results in the presence of specific cocultures that may help to recreate the natural environment of developing or functioning hepatocytes more effectively.

SCAFFOLDS FOR LIVER-TISSUE ENGINEERING

The model for tissue engineering combines a bioactive cellular component with a structural biodegradable scaffold. Once the desired cells have been isolated, the scaffolds play numerous essential roles in the task of creating functioning tissues and organs. Most cells—including hepatocytes—require attachment to a surface to function properly. In addition to providing an attachment surface, scaffolds help control three-dimensional configuration and growth, affect differentiation and function with surface characteristics and growth factors, and provide a carrier for implantation.^[30] To meet requirements for nontoxicity and nonimmunogenicity, early scaffolds borrowed technology from Food and Drug Administration (FDA)-approved sutures, wound dressings, and hemostatic agents.^[31] However, scientists are developing new biomaterials and are able to more precisely create and control the environment to which cells are exposed in tissue-engineered constructs.

Tissue-engineering biomaterials research focuses on the development and testing of new materials with different mechanical and chemical properties, degradation times, and cell-material interactions, and studies processing techniques that control microstructures and macrostructures to best promote the growth and function of specific tissues. There are many classifications of scaffolds—including gel or solid, biologic or synthetic, and biodegradable or nondegradable but biocompatible (see Table 1). Hybrid constructs have also been developed that combine more than one class of scaffold to vary characteristics of different regions of the constructs.

Poly(glycolic acid) (PGA) and poly(lactic acid) (PLA) are derived from resorbable sutures and have been used since the beginnings of the modern field of tissue engineering. They have crystalline and amorphous forms and can be combined in different ratios to control degradation characteristics. Depending on the desired

Table 1 Examples of different types of tissue-engineering scaffolds

Gels	Hydrogel, collagen gel, and alginate gel
Solids	All other non-gels listed below
Biologic	Small intestinal submucosa (SIS), and other decellularized matrices
Synthetic	All polymers
Biodegradable	Polyglycolic acid (PGA), poly-L-lactic acid (PLLA), polylactide-co-glycolic acid (PLGA), polycaprolactone (PCL), and polyhydroxybutyrates (PHB), Biorubber
Nondegradable and biocompatible	Polydimethylsiloxane (silicone or PDMS), polycarbonate (PC), polysulfone (PS), polyethylenes (PE), polyurethanes (PU), and polyesters

(From Refs. [31–33].)

conditions for development and the rate of ECM deposition, scientists can vary the degradation rate to allow for complete degradation of the polymer either entirely in vitro prior to implantation or after 2 weeks to 2 years in vivo. PLA, PGA, and PLGA have been effective and are in use in clinical applications of tissue engineering;^[31] however, these polymers have limitations due to their brittle nature and acidic byproducts. Numerous other polymers have recently been in the forefront of tissue engineering, including PCL, PHB, and pseudopoly amino acids that have different and more desirable mechanical, chemical, and degradation properties for different tissue applications.^[31] Novel polymers like biorubber^[32] have been developed with varying elasticity and strength for applications like heart- and lung-tissue engineering, and may also play a role in the future of hepatic-tissue engineering.

COATINGS AND FACTORS TO MODIFY SCAFFOLD AND HEPATOCYTE ENVIRONMENT

Scientists often turn to nature for guidance; tissue engineers attempt to recreate physiological conditions for cells by mimicking the environment—including the ECM—to which the cells are exposed in vivo. Attempts to recreate the ECM found in the liver have focused largely on surface characteristics to enhance interactions with cells for increased adhesion, proliferation, spreading, and function while reducing levels of dedifferentiation. Several prominent scaffold modifications include pore size, micropatterning, integration of adhesion factors, and surface coating with factors like collagen, fibronectin, and laminin.

Tissue repair normally takes place in the presence of glycosaminoglycans (GAGs)—especially hyaluronan (HA)—and ECM proteins like collagen, fibronectin, and elastin. In addition to utilizing genetic engineering to employ bacteria to produce ECM proteins or oligopeptides to make the scaffold more suitable for cell attachment, scientists can alter the genes of the cells to attach to a specific scaffold surface. Hepatocyte membrane integrins bind to adhesion domains in ECM proteins composed of short peptide sequences. The tripeptide sequence Arg-Gly-Asp (RGD)—an adhesion domain found in collagen, fibronectin, and vitronectin—and the pentapeptide sequence Tyr-Ile-Gly-Ser-Arg (YIGSR)—present in laminin—are two peptide sequences commonly being attached or incorporated on the surfaces of polymer scaffolds. Researchers have used a variety of intermediate molecules to integrate adhesion peptides to polymer scaffolds. RGD and YISGR have been linked to polyethyleneglycol (PEG), which is bound to ally-amine coating on polymers (PCL, PLLA) via pulsed plasma deposition.^[33] Comb polymers can be used to covalently bind adhesion molecules or other ligands to the surface of scaffolds.^[34,35]

Growth factors such as VEGF^[21] can be added as soluble factors or incorporated on the surface for interactions following cell attachment, in the bulk of degradable polymer used as scaffolding, or in microspheres that are slowly released over varying timescales. One area of concern is the acidic microenvironment created by the breakdown products of scaffolds like PLGA and its effect on cells and the bioactivity of growth factors, drugs, and other candidates for slow release. Researchers are currently developing and testing scaffolds that are more compatible with growth factors, drugs, and cells.

Another expanding area in material science research is the development of so-called smart polymers that are able to respond to changes in their environment such as to the pH, temperature, or concentration of a ligand. By employing a temperature-sensitive coating that has varying temperature-dependent cell adherence properties, researchers have been able to release cells grown to confluence from the culture surface in an intact sheet, avoiding damage to the matrix between cells that results from the use of trypsin in traditional cell passaging.^[36]

IMMUNE RESPONSE

For long-term replacement tissue-engineered implants, it is important to minimize immune responses to tissue-engineered constructs that can lead to inflammation, rejection, and cytokine production. Unlike allogenic transplantation or xenotransplantation, tissue engineering can employ autologous cells that do not require

the administration of immunosuppressive therapies that are toxic to both the liver and kidneys. In constructs engineered from the patient's own cells, the cellular component will be recognized as "self" and will not induce rejection. However, current applications with bioartificial liver-assist devices employ either allogenic human or xenogenic porcine hepatocytes. Efforts have been made to immunologically isolate these cells using blood separation and a series of membranes to prevent the release of any foreign hepatocytes into the blood, and to filter out those that may contaminate the blood or plasma within the device.

Whereas the cell source can be derived from the patient, most forms of tissue-engineered scaffolds are artificial. Because any artificial material implanted over an extended period will induce a foreign-body reaction that leads to scar tissue formation and rejection, it is imperative to use biodegradable polymers in tissue engineering for long-term therapeutic treatments. However, even with biodegradable polymers, there are still potential immunogenic problems. Residual monomer is thought to be a major cause of the rejection of biodegradable polymers, so care must be taken to purify polymers implanted either during or following synthesis. When solvent-casting polymers, extra care must be taken to remove all residual solvents, many of which are toxic to cells. Small intestinal submucosa (SIS) is a natural scaffold devoid of cells derived from the intestinal wall ECM that induces lower immunologic reactions. In SIS, immunoglobulin A (IgA) may suppress immune response despite also having the galactose α -1, 3 epitope (α -gal), which is normally attacked by human immune systems by the naturally expressed anti- α -gal antibody. SIS implants in mice induce levels of interleukins IL-4 and IL-10, but not the interferon IFN- γ , and the antibodies mounted are only of the immunoglobulin IgG1 class.^[37]

HEPATIC BIOREACTORS

The purpose behind creating the bioreactor was to develop a larger, functioning, viable, and proliferating three-dimensional aggregate of hepatocytes required for support and eventual replacement of hepatic functions in an implantable liver construct. A basic bioreactor functions as a continuous flow system that circulates media in lieu of blood through tubing, an oxygenator (lung), and enclosed capsule that contains a scaffold with engrafted hepatocytes in a three-dimensional configuration (see Fig. 1). Hepatocytes in bioreactors have been shown to exhibit normal hepatocyte function for longer periods than seen in a static two-dimensional culture. Researchers of hepatic bioreactors are experimenting to develop more efficient and effective configurations to allow for longer survival and specific function for larger masses of hepatocytes. As the aggregate of cells increases

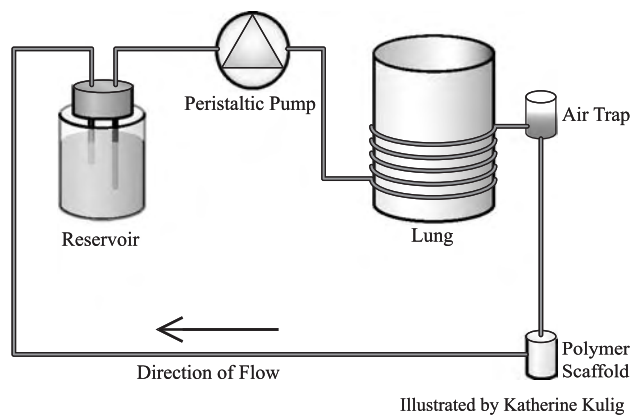


Fig. 1 Bioreactor for in-vitro culture of hepatocyte polymer constructs. (From Refs. [6,11,12].)

in size, fewer cells survive due to limited nutrients, waste removal, and oxygen exchange. Whereas the cells in the peripheral zones thrive with readily available and accessible nutrients and gases, the cells within the core of the construct often have an increased mortality rate due to the impedance to mass transfer. These necessary functions occur in vivo with the assistance of a vascular system. In order to improve the life span and normal hepatic functions of hepatocytes in an aggregate, all cells must be within 100–300 μM of a blood supply.^[38] The bioreactor system supplies the needed surface for attachment via the polymer construct (PGA, PLLA, PLGA) of optimal porosity and through dynamic seeding techniques and even distribution of cells while angiogenesis occurs.^[11] The flow-through design has enabled many investigators to demonstrate long-term viability and increased production of hepatic products such as albumin, cytochrome P-450 activity, and other metabolic processes. The benefits of long-term survival and growth provide insight into the need for tissue replacement, and also act as an in-vitro liver replicate for future drug testing. (See Fig. 1.)

LIVER-ASSIST DEVICES

An expanding area of research that overlaps and has synergy with hepatic bioreactors for tissue engineering is the development of liver-assist devices (LADs). LADs are being tested in both laboratories and clinical trial devices as a gateway treatment for patients with FHF to support them on a timescale of weeks until a liver becomes available for transplant. These systems may also be beneficial in the short term for patients excluded as candidates for transplant.

The primary functional goal of the devices is to detoxify toxins accumulating in the blood stream because of hepatic failure. Laboratories are developing a variety

of LADs that are not yet in clinical trials, using different cell sources and membrane configurations such as a spirally-wound polyester fabric with direct hepatocyte-blood contact^[28] or a multicoaxial set of four porous tubes.^[39] Devices in clinical trials draw from Wolf's idea and combine the hollow fibers used in kidney dialysis cartridges technology with added living hepatocytes to restore some of the biologic function of the liver.^[40] (See Fig. 2.) Many of the configurations incorporate activated charcoal to help remove toxins, in addition to oxygenators and heaters to support the hepatocytes. The hepatocytes are typically seeded on either the fiber surface or in microcarrier beads in the extrafiber space.^[41] The compartment seeded with hepatocytes can also be separately perfused with media to both support hepatocyte function and help remove toxins from the device. High levels of toxins in the blood are thought to be the main reason for the development of encephalopathy from liver failure. In one configuration, hepatocytes are seeded within the hollow-fiber lumen in a collagen gel that contracts 60% to allow flow through the lumen that is independent of the extrafiber flow;^[42] this serves as the basis for Algenix's LiverX200 that will soon begin phase I trials. LADs currently in clinical trials are the HepatAssist[®],^[43–45] ELAD[™] (Extracorporeal Liver-Assist Device),^[46–48] MELS (Modular Extracorporeal Liver Support),^[49–51] and BLSS (Bioartificial Liver Support System).^[52–55] The cells utilized in these clinical trials are human hepatic cell line HepG2/C3A cells (ELAD), human primary hepatocytes harvested from untransplantable livers (MELS), or porcine, hepatocytes (HepatAssist, BLSS). With one exception, the BALs all separate the patient's blood and perfuse only the plasma. The exception is the BLSS, which perfuses with the patient's whole blood through the porous polysulfone (HepatAssist, MELS) or cellulose acetate (ELAD, BLSS) hollow fibers of dialysis cassettes that are seeded with hepatocytes in the extrafiber space. In perfusing only the plasma, precautions are taken to isolate these cells and possible pathogens (including viral particles from the patient's blood) by employing filters with pores 0.2 to 1.5 microns in diameter. Although these devices have been successful for short periods, fully implantable replacement systems are required for long-term therapy.

TISSUE-ENGINEERED LIVER REPLACEMENT

A functional replacement liver has long been the goal of the field of tissue engineering. Greater success has been achieved with other tissues, such as cartilage and skin, due in large part to their relatively low metabolic rates and thin dimensions, respectively. With the major limitation being sufficient circulation for three-dimensional hepatic tissue, approaches such as direct

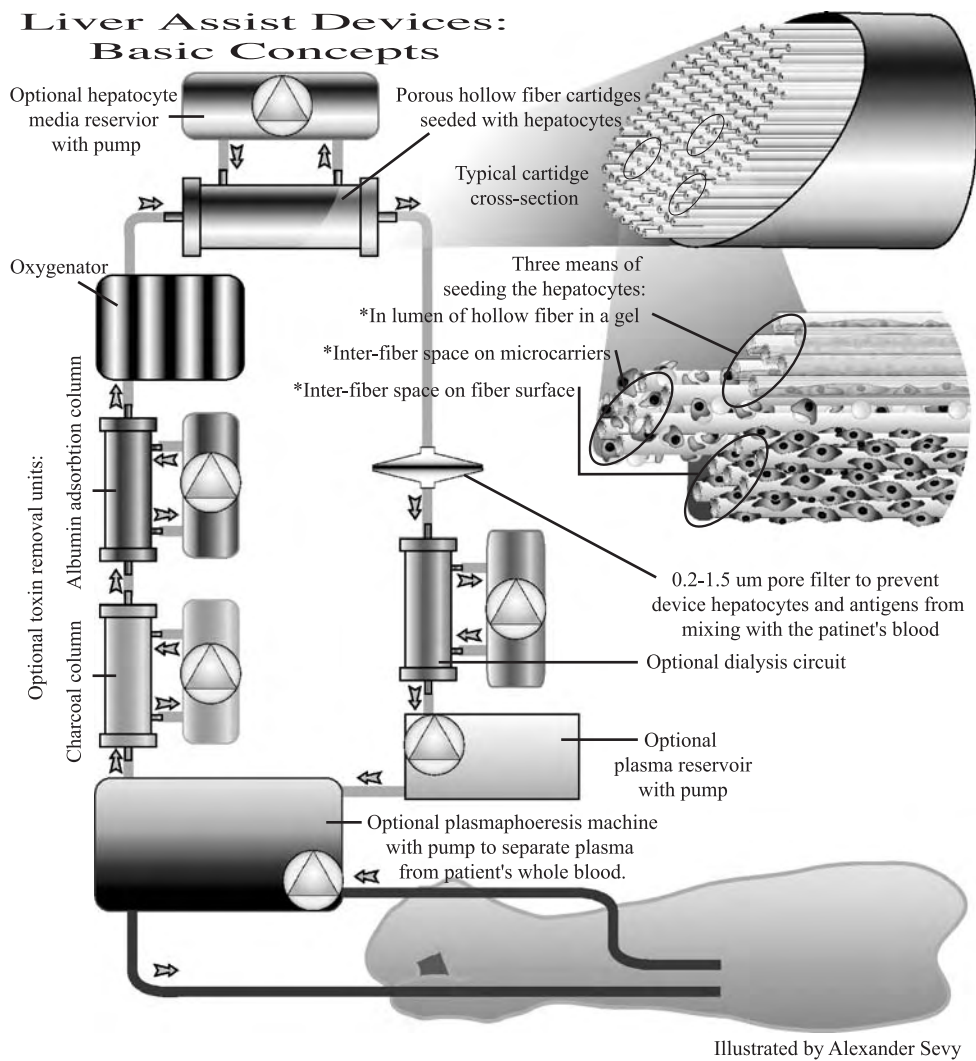


Fig. 2 A schematic drawn from multiple LADs designed for clinical trial. (From Refs. [42–55].)

hepatocyte injection into existing vascular beds, micro-carrier attachment, and scaffold support and attachment have been investigated as possible means of hepatic functional support and eventual replacement.

One method of vascularizing tissue-engineered liver constructs is to promote angiogenesis into a non-vascularized implant. Angiogenesis research advances have induced an increase up to 260% in matrix neovascularization after one week with the addition to the construct of VEGF and endothelial cells.^[56,57] Adding or integrating angiogenic factors such as VEGF to the environment of hepatocyte constructs has resulted in improved function and viability over constructs without induced angiogenesis. However, successful results in vivo have been limited to thicknesses that would not be feasible for a clinical replacement mass of hepatocytes.

Another approach for gaining sufficient circulation is to inject cells (with or without a gel) into an area of the

body that is already sufficiently vascularized to support the additional hepatocytes. Direct hepatocyte injection has focused on the degree of liver function replaced and the cells' ability to sustain the metabolic functions of the liver. The cells have been injected directly into vascular beds, organs, and body cavities. The spleen and its related peritoneal cavities have been successful in maintaining transplanted hepatocytes.^[5] Cell transplantation has been effective at curing some forms of genetic diseases with specific hepatic function defects; however, the mass of cells that can be injected and supported by the existing vasculature is not sufficient for liver replacement. To improve the attachment and survival of human hepatocytes transplanted into athymic rats, our laboratory provided hepatotrophic stimulation with a portocaval shunt and prevascularized polymer sponges; placing the constructs in subcutaneous tissues and between mesenteric leaves allowed for the concomitant

ingrowth of fibroblasts and vascularization of the construct.^[58]

A prototypical implantable tissue-engineered liver would have three external connections, a blood inlet and outlet, and an outflow for bile, with a preexisting internal network of blood vessels and ducts branching down or almost down to the capillary and ductile level. Whereas complex bioreactors have been designed that can support complex three-dimensional configurations of cells, nonprevascularized hepatocyte constructs do not survive *in vivo* because they do not have access to a sufficient supply of oxygen, nutrients, and waste removal. The mass of hepatocytes required to support a large animal or human would require a rapid blood supply throughout the construct for the cells to survive and function. To solve the problem of mass transport in a thick, metabolically active whole-organ replacement, our laboratory has developed a tissue-engineered vascular bed designed to support metabolic tissues.^[3] Three-dimensional printing is capable of producing complex three-dimensional branching networks, but currently lacks the resolution to fabricate capillaries. Using photolithographic techniques borrowed from the computer chip industry, our group etched patterns of vascular beds with capillary-scale branches on silicon wafers. These patterns can be transferred to biocompatible materials such as polydimethylsiloxane (PDMS) or biodegradable polymers like PLGA. The channels can be sealed and seeded with endothelial cells that have been grown to confluence. Stacks of these vascular beds can then be interposed with layers of parenchymal cells such as hepatocytes and communicate through membranes that allow the transfer of oxygen, nutrients, and waste. (See Fig. 3.) We have recently grown hepatocytes that are supported by a microfabricated vascular bed that communicates

through a porous membrane with the hepatocyte compartment.^[4]

CONCLUSION

Although the modern field of tissue engineering has made great advances in the past thirty years, there is no functional tissue-engineered liver replacement construct to date. However, a number of promising developments and research in diverse fields show clinical improvements in trials and are advancing the field toward a therapeutic alternative to liver transplantation. Progress in understanding cell developmental and molecular biology, polymer scaffold development and testing, development of human liver bioreactors for drug testing, application of gene therapy, and means to store three-dimensional tissue constructs will provide improved methods of treating patients with liver diseases.

As the need for the optimal cell source is considered for research and tissue engineering, ultimate control over the destiny of the cell—be it differentiation or proliferation—is the factor that determines which type of cell line is used. Globally, embryonic, fetal, and adult stem cells are being studied to further characterize their potential and ability to functionally repopulate or replace diseased tissues. As more is understood about stem cells, regulatory and ethical principles must be developed and instituted to prevent therapeutic misuse or premature clinical application. In addition to cell biology research, novel scaffold materials development and testing are essential for the advancing field of tissue engineering. Scaffolds with more ideal physical properties and surface characteristics are only part of the picture; other factors in the hepatocytes' environment must continue to be

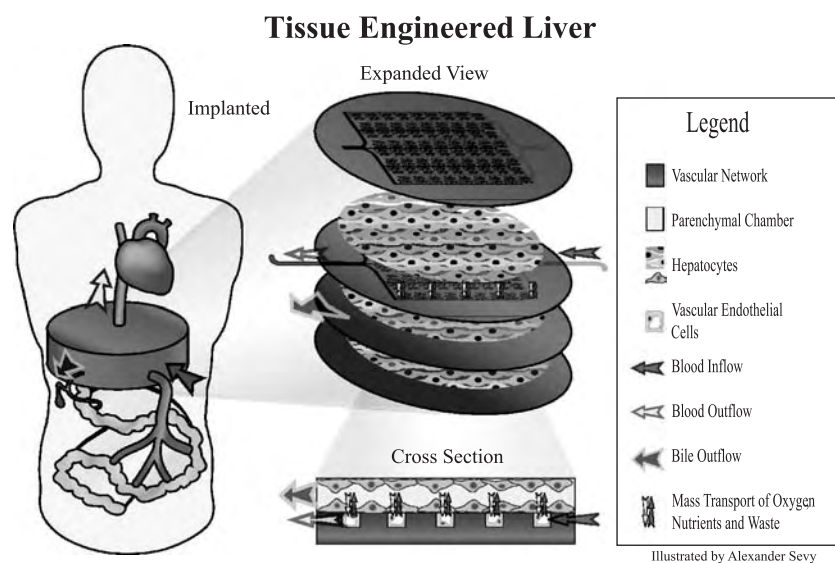


Fig. 3 Tissue-engineered liver replacement. (From Refs. [3,4].)

developed to better mimic conditions in vivo. Used in concert with cell transplantation or in cell-polymer constructs, gene therapy may offer powerful means to treat patients with genetic defects that impair liver function. Gene therapy may help prevent immunosuppression if tissue-engineered liver constructs were to house a patient's conditioned cells. Storage of tissue-engineered constructs may become possible as vitrification and desiccation protocols advance with fewer harmful cryoprotectants (such as disaccharides), and as damage to cells and their membranes is minimized.^[59] In addition to future storage facilities, a novel approach to drug testing will also become available.

The Hepatitis C virus (HCV) is the leading cause of liver failure in the United States; in one year those infected with HCV accounted for 8000 to 10,000 deaths.^[25] Unfortunately, there are difficulties in developing drugs to treat the disease; because HCV infects liver tissue only in humans, animal models are ineffective for testing HCV drugs. Researchers can use bioreactors with HCV-infected human hepatocytes to develop effective treatment. The liver also plays a key role in the breakdown of other drugs such that in vitro and implant models with human hepatocytes will help screen drugs that may be either harmful to or inactivated by the liver.

The expanding field of hepatic-tissue engineering has already shown clinical benefit and promises to provide even more robust long-term support for patients with liver disease. In this exciting and multidisciplinary field, it is imperative to collaborate and to integrate a wide array of information in the quest for the holy grail of a tissue-engineered replacement liver. Advances leading to the realization of a replacement liver will serve not only to benefit patients with liver failure, but also to advance the development of therapies for other diseases.

ACKNOWLEDGMENTS

We would like to thank Wing S. Cheung, Tracy Clark, and Joseph W. Carlson.

ARTICLES OF FURTHER INTEREST

Bioreactors; Extracellular Matrix Scaffolds; Liver, Bio-Artificial; Microelectromechanical Systems (MEMS) Manufacturing; Poly(Glycolic Acid); Poly(lactic acid)s; Tissue Engineering Scaffolds

REFERENCES

1. Scientific Registry of Transplant Recipients Oct. 31, 2002 <http://www.ustransplant.org/facts.html>.
2. Asonuma, K.; Gilbert, J.C.; Stein, J.E.; Takeda, T.; Vacanti, J.P. Quantification of transplanted hepatic mass necessary to cure the Gunn rat model of hyperbilirubinemia. *J. Pediatr. Surg.* **1992**, *27* (3), 298–301.
3. Kaihara, S.; Borenstein, J.; Koka, R.; Lalan, S.; Ochoa, E.R.; Ravens, M.; Pien, H.; Cunningham, B.; Vacanti, J.P. Silicon micromachining to tissue engineer branched vascular channels for liver fabrication. *Tissue Eng.* **2000**, *6* (2), 105–117.
4. Cheung, W.; Borenstein, J.; Kaazempur-Mofrad, M.; Shin, M.; Sevy, A.; Kulig, K.; Vacanti, J.P. Development of an Implantable Tissue-Engineered Liver Device with a Vascular Network of Channels Co-Cultured with Human Hepatocytes and Human Microvascular Endothelial Cells. Presented at the American Transplant Congress, Washington, D.C., May 30–June 4, 2003.
5. Gupta, S.; Malhi, H.; Gorla, G.R. Re-engineering the liver with natural biomaterials. *Yonsei Med. J.* **2000**, *41* (6), 814–824.
6. Powers, M.J.; Domansky, K.; Kaazempur-Mofrad, M.R.; Kalezi, A.; Capitano, A.; Upadhyaya, A.; Kurzawski, P.; Wack, K.E.; Beer Stolz, D.; Kamm, R.; Griffith, L.G. A microfabricated array bioreactor for perfused 3D liver culture. *Biotechnol. Bioeng.* **2002**, *78* (3), 257–269.
7. Davis, M.W.; Vacanti, J.P. Toward development of an implantable tissue-engineered liver. *Biomaterials* **1996**, *17* (3), 365–372.
8. te Velde, A.A.; Ladiges, N.C.J.J.; Flendrig, L.M.; Chamuleau, R.A.F.M. Functional activity of isolated pig hepatocytes attached to different extracellular matrix substances. Implication for application of pig hepatocytes in a bioartificial liver. *J. Hepatol.* **1995**, *23* (2), 184–192.
9. Mitaka, T.; Kojima, T.; Mizuguchi, T.; Mochizuki, Y. Growth and maturation of small hepatocytes isolated from adult rat liver. *Biochem. Biophys. Res. Commun.* **1995**, *214* (2), 310–317.
10. Mitaka, T.; Mizuguchi, T.; Sato, F.; Mochizuki, C.; Mochizuki, Y. Growth and maturation of small hepatocytes. *J. Gastroenterol. Hepatol.* **1998**, *13* (Suppl.), S70–S77.
11. Kim, S.S.; Sundback, C.A.; Kaihara, S.; Benvenuto, M.S.; Kim, B.; Mooney, D.J.; Vacanti, J.P. Dynamic seeding and in vitro culture of hepatocytes in a flow perfusion system. *Tissue Eng.* **2000**, *6* (1), 39–44.
12. Powers, M.J.; Janigian, D.M.; Wack, K.E.; Baker, C.S.; Beer Stolz, D.; Griffith, L.G. Functional behavior of primary rat liver cells in a three-dimensional perfused microrarray bioreactor. *Tissue Eng.* **2002**, *8* (3), 499–512.
13. Mooney, D.; Hansen, L.; Vacanti, J.; Langer, R.; Farmer, S.; Ingber, D. Switching from differentiation to growth in hepatocytes: Control by extracellular matrix. *J. Cell. Physiol.* **1992**, *151* (3), 497–505.
14. Sugimoto, S.; Mitaka, T.; Ikeda, S.; Harada, K.; Ikai, I.; Yamaoka, Y.; Mochizuki, Y. Morphological changes induced by extracellular matrix are correlated with maturation of rat small hepatocytes. *J. Cell. Biochem.* **2002**, *87* (1), 16–28.

15. Goulet, F.; Normand, C.; Morin, O. Cellular interactions promote tissue-specific function, biomatrix deposition and junctional communication of primary cultured hepatocytes. *Hepatology* **1988**, *8* (5), 1010–1018.
16. Hansen, L.K.; Mooney, D.J.; Vacanti, J.P.; Ingber, D.E. Integrin binding and cell spreading on extracellular matrix act at different points in the cell cycle to promote hepatocyte growth. *Mol. Biol. Cell* **1994**, *5* (9), 967–975.
17. Ingber, D. Mechanical signaling. *Ann. N.Y. Acad. Sci.* **2002**, *961*, 162–163.
18. Kim, W.J.H. Cellular signaling in tissue regeneration. *Yonsei Med. J.* **2000**, *41* (6), 692–703.
19. Bhandari, R.N.B.; Riccalton, L.A.; Lewis, A.L.; Fry, J.R.; Hammond, A.H.; Tendler, S.J.B.; Shakesheff, K.M. Liver tissue engineering: A role for co-culture systems in modifying hepatocyte function and viability. *Tissue Eng.* **2001**, *7* (3), 345–357.
20. Michalopoulos, G.K.; Bowen, W.C.; Zajac, V.F.; Beer-Stolz, D.; Watkins, S.; Kostrubsky, V.; Strom, S.C. Morphogenetic events in mixed cultures of rat hepatocytes and nonparenchymal cells maintained in biological matrices in the presence of hepatocyte growth factor and epidermal growth factor. *Hepatology* **1999**, *29* (1), 90–100.
21. Mooney, D.J.; Kaufmann, P.M.; Sano, K.; Schwendeman, S.P.; Majahod, K.; Schloo, B.; Vacanti, J.P.; Langer, R. Localized delivery of epidermal growth factor improves the survival of transplanted hepatocytes. *Biotechnol. Bioeng.* **1996**, *50* (4), 422–429.
22. Mitaka, T.; Mikami, M.; Sattler, G.L.; Pitot, H.C.; Mochizuki, Y. Small cell colonies appear in the primary culture of adult rat hepatocytes in the presence of nicotinamide and epidermal growth factor. *Hepatology* **1992**, *16* (2), 440–447.
23. Xin, X.; Yang, S.; Ingle, G.; Zlot, C.; Rangell, L.; Kowalski, J.; Schwall, R.; Ferrara, N.; Gerritsen, M.E. Hepatocyte growth factor enhances vascular endothelial growth factor-induced angiogenesis in vitro and in vivo. *Am. J. Pathol.* **2001**, *158* (3), 1111–1120.
24. Nyberg, S.L.; Rimmel, R.P.; Mann, H.J.; Peshwa, M.V.; Hu, W.; Cerra, F.B. Primary hepatocytes outperform HepG2 cells as the source of biotransformation functions in a bioartificial liver. *Ann. Surg.* **1994**, *220* (1), 59–67.
25. Allen, J.W.; Bhatia, S.N. Engineering liver therapies for the future. *Tissue Eng.* **2002**, *8* (5), 725–737.
26. Faris, R.A.; Konkin, T.; Halpert, G. Liver stem cells: A potential source of hepatocytes for the treatment of human liver disease. *Artif. Organs* **2001**, *25* (7), 513–521.
27. Noishiki, Y. Dreams for the future in the field of in vivo tissue engineering. *Artif. Organs* **2001**, *25* (3), 159–163.
28. Flendrig, L.M.; la Soe, J.W.; Jörning, G.G.A.; Steenbeek, A.; Karlsen, O.T.; Bovée, W.M.M.J.; Ladiges, N.C.J.J.; te Velde, A.A.; Chamuleau, R.A.F.M. In vitro evaluation of a novel bioreactor based on an integral oxygenator and a spirally wound nonwoven polyester matrix for hepatocyte culture as small aggregates. *J. Hepatol.* **1997**, *26* (6), 1379–1392.
29. Mizuguchi, T.; Hui, T.; Palm, K.; Sugiyama, N.; Mitaka, T.; Demetriou, A.A.; Rozga. Enhanced proliferation and differentiation of rat hepatocytes cultured with bone marrow stromal cells. *J. Cell. Physiol.* **2001**, *189* (1), 106–119.
30. Terada, S.; Sato, M.; Sevy, A.; Vacanti, J.P. Tissue engineering in the twenty-first century. *Yonsei Med. J.* **2000**, *41* (6), 685–691.
31. Griffith, L.G. Emerging design principles in biomaterials and scaffolds for tissue engineering. *Ann. N.Y. Acad. Sci.* **2002**, *961*, 83–95.
32. Wang, Y.; Guillermo, A.A.; Sheppard, B.J.; Langer, R. A tough biodegradable elastomer. *Nat. Biotechnol.* **2002**, *20* (6), 602–606.
33. Carlisle, E.S.; Mariappan, M.R.; Nelson, K.D.; Thomes, B.E.; Timmons, R.B.; Constantinescu, A.; Eberhart, R.C.; Bankey, P.E. Enhancing hepatocyte adhesion by pulsed plasma deposition and polyethylene glycol coupling. *Tissue Eng.* **2000**, *6* (1), 45–52.
34. Irvine, D.J.; Mayes, A.M.; Griffith, L.G. Nanoscale clustering of RGD peptides at surfaces using comb polymers. 1. Synthesis and characterization of comb thin films. *Biomacromolecules* **2001**, *2* (1), 85–94.
35. Irvine, D.J.; Ruzette, A.V.; Mayes, A.M.; Griffith, L.G. Nanoscale clustering of RGD peptides at surfaces using comb polymers. 2. Surface segregation of comb polymers in polylactide. *Biomacromolecules* **2001**, *2* (2), 545–556.
36. Prestwich, G.D.; Matthew, H. Hybrid, composite, and complex biomaterials. *Ann. N.Y. Acad. Sci.* **2002**, *961*, 106–108.
37. Metzger, D.W. Immune responses to tissue-engineered extracellular matrix used as a bioscaffold. *Ann. N.Y. Acad. Sci.* **2002**, *961*, 335–336.
38. Vacanti, J.P.; Morse, M.A.; Saltzman, W.M.; Domb, A.J.; Perez-Atayde, A.; Langer, R. Selective cell transplantation using bioabsorbable artificial polymers as matrices. *J. Pediatr. Surg.* **1988**, *23* (1 pt 2), 3–9.
39. Wolfe, S.P.; Hsu, E.; Reid, L.M.; Macdonald, J.M. A novel multi-coaxial hollow fiber bioreactor for adherent cell types. Part 1: Hydrodynamic studies. *Biotechnol. Bioeng.* **2002**, *77* (1), 83–90.
40. Wolf, C.F.; Munkelt, B.E. Bilirubin conjugation by an artificial liver composed of cultured cells and synthetic capillaries. *Trans. Am. Soc. Artif. Intern. Organs* **1975**, *21*, 16–27.
41. Arnaut, W.S.; Mosconi, A.D.; Barbout, R.L.; Demetriou, A.A. Development of bioartificial liver: Bilirubin conjugation in Gunn rats. *J. Surg. Res.* **1990**, *48* (4), 379–382.
42. Nyberg, S.L.; Shirabe, K.; Peshwa, M.V.; Sielaff, T.D.; Crotty, P.L.; Mann, H.J.; Rimmel, R.P.; Payne, W.D.; Hu, W.S.; Cerra, F.B. Extracorporeal application of a gel-entrapment, bioartificial liver: Demonstration of drug metabolism and other biochemical functions. *Cell Transplant* **1993**, *2* (6), 441–452.
43. Chen, S.C.; Hewitt, W.R.; Watanabe, F.D.; Eguchi, S.; Kahaku, E.; Middleton, Y.; Rozga, J.; Demetriou, A.A. Clinical experience with a porcine hepatocyte-based liver support system. *Int. J. Artif. Organs* **1996**, *19* (11), 664–669.

44. Detry, O.; Arkadopoulos, N.; Ting, P.; Kahaku, E.; Watanabe, F.D.; Rozga, J.; Demetriou, A.A. Clinical use of a bioartificial liver in the treatment of acetaminophen-induced fulminant hepatic failure. *Am. Surgeon* **1999**, *65* (10), 934–938.
45. Watanabe, F.D.; Mullon, C.J.-P.; Hewitt, W.R.; Arkadopoulos, N.; Kahaku, E.; Eguchi, S.; Khalili, T.; Arnaout, W.; Shackleton, C.R.; Rozga, J.; Solomon, B.; Demetriou, A.A. Clinical experience with a bioartificial liver in the treatment of severe liver failure—A phase I clinical trial. *Ann. Surg.* **1997**, *225*, 484–491.
46. Sussman, N.L.; Gislason, G.T.; Kelly, J.H. Extracorporeal liver support. Application to fulminant hepatic failure. *J. Clin. Gastroenterol.* **1994**, *18*, 320–324.
47. Sussman, N.L.; Kelly, J.H. Extracorporeal liver support: Cell-based therapy for the failing liver. *Am. J. Kidney Dis.* **1997**, *30* (5), S66–S71.
48. Ellis, A.J.; Hughes, R.D.; Wendon, J.A.; Dunne, J.; Langley, P.G.; Kelly, J.H.; Gislason, G.T.; Sussman, N.L.; Williams, R. Pilot-controlled trial of the extracorporeal liver assist device in acute liver failure. *Hepatology* **1996**, *24*, 1446–1451.
49. Gerlach, J.C. Long-term liver cell cultures in bioreactors and possible application for liver support. *Cell Biol. Toxicol.* **1997**, *13* (4–5), 349–355.
50. Gerlach, J.C.; Lemmens, P.; Schon, M.; Janke, J.; Rossaint, R.; Busse, B.; Puhl, G.; Neuhaus, P. Experimental evaluation of a hybrid liver support system. *Transplant. Proc.* **1997**, *29* (1–2), 852.
51. Sauer, I.M.; Gerlach, J.C. Modular extracorporeal liver support. *Artif. Organs* **2002**, *26* (8), 703–706.
52. Mazariegos, G.V.; Kramer, D.J.; Lopez, R.C.; Shakil, A.O.; Rosenbloom, A.J.; DeVera, M.; Giraldo, M.; Grogan, T.A.; Zhu, Y.; Fulmer, M.L.; Amiot, B.P.; Patzer, J.F., II. Safety observations in phase I clinical evaluation of the excorp medical bioartificial liver support system after the first four patients. *ASAIO* **2001**, *47*, 471–475.
53. Mazariegos, G.V.; Patzer, J.F., II; Lopez, R.C.; Giraldo, M.; Devera, M.E.; Grogan, T.A.; Zhu, Y.; Fulmer, M.L.; Amiot, B.P.; Kramer, D.L. First clinical use of a novel bioartificial liver support system (BLSS). *Am. J. Transplant* **2002**, *2* (3), 260–266.
54. Patzer, J.F., II; Mazariegos, G.V.; Lopez, R. Preclinical evaluation of the Excorp Medical Inc, bioartificial liver support system. *J. Am. Coll. Surg.* **2002**, *195* (3), 299–310.
55. Kuddus, R.; Patzer, J.F., II; Lopez, R.; Mazariegos, G.V.; Meighen, B.; Kramer, D.J.; Rao, A.S. Clinical and laboratory evaluation of the safety of a bioartificial liver assist device for potential transmission of porcine endogenous retrovirus. *Transplantation* **2002**, *73* (3), 740–749.
56. Richardson, T.P.; Peters, M.C.; Ennett, A.B.; Mooney, D.J. Polymeric system for dual growth factor delivery. *Nat. Biotechnol.* **2001**, *19* (11), 1029–1034.
57. Peters, M.C.; Poverini, P.J.; Mooney, D.J. Engineering vascular networks in porous polymer matrices. *J. Biomed. Mater. Res.* **2002**, *60* (4), 668–678.
58. Fontaine, M.; Schloo, B.; Jenkins, R.; Uyama, S.; Hansen, L.; Vacanti, J.P. Human hepatocyte isolation and transplantation into an athymic rat, using prevascularized cell polymer constructs. *J. Pediatr. Surg.* **1995**, *30* (1), 56–60.
59. Toner, M.; Kocsis, J. Storage and translational issues in reparative medicine. *Ann. N.Y. Acad. Sci.* **2002**, *961*, 258–262.

Tissue Engineering, Microscale

Wei Tan

University of Illinois at Urbana-Champaign, Urbana, Illinois, U.S.A.

Tejal Desai

Department of Biomedical Engineering, Boston University, Boston, Massachusetts, U.S.A.

T

INTRODUCTION

The last decade has seen the dramatic development of biomedical technologies for organ replacement. The ability to perform tissue and organ replacement has led to an improvement in quality of life and healthcare. Some of this success can be attributed to the interdisciplinary approaches of tissue engineering. The goal of tissue engineering is to augment, replace, or restore complex human tissue function by combining synthetic and living components in appropriate configurations and environmental conditions.^[1] Scientists with very diverse backgrounds collaborate with engineers to develop tissue analogs. There are generally three key components to consider in any tissue-engineered construct: the cells, the scaffolding or biomaterial construct, and the signaling molecules. Cells are an active living component in the tissue-engineered product. Cellular processes such as adhesion, migration, growth, secretion, and gene expression are triggered or influenced by the three-dimensional organization of the surrounding cells, scaffolds, and signaling molecules. Cells play a central role in building and maintaining specific tissue function, but when they are separated from their natural living environments and cultured *in vitro*, the physiological cellular activities are either diminished or reduced. Thus, it is critically important to develop a fundamental understanding on how the surrounding microenvironments induce and maintain the stability of the cellular functions, and then apply the understanding in designing biomimetic environments to encourage isolated cells to recover their *in-vivo* functions.^[2,3] Scaffolding or biomaterials for tissue engineering are not just inert supports in tissue-engineered constructs as they were used in the early stage of tissue engineering.^[4] The impact of the microstructure, surface chemistry, surface topography, and tissue-integrating ability has been increasingly realized. The precise control of the interactions between cells and scaffolds has received more and more attention. Signaling molecules mediate interactions and may induce cells to generate signals in mechanical, biological, chemical, and electrical forms for communications. The three-dimensional distribution of these molecules should be well controlled. From an engineering standpoint, it is important to choose an appropriate method

for synthesizing or assembling all the components into a biological and medically meaningful structure that will confer desired functions.

HISTORICAL BACKGROUND

Although tissue engineering is a relatively new field, a multitude of strategies has been developed. The design of many tissue-engineered systems, such as the vascular system, bone, and skin, evolved from bioinert prostheses to bioinert surface-coated prostheses, and then to biointeractive and biomimetic prostheses. The approach of bioinert prostheses uses only scaffolding materials, for example, ePTFE and dacron grafts used to replace arteries in cardiovascular surgery.^[5,6] The majority of these grafts not only lack typical components of healthy native tissues, but also lack the ability to restore tissue integrity through quiescent scar formation. A further step in tissue engineering is to reconstruct cell-seeded biomaterials via surface coatings on the biomaterials.^[7,8] In spite of many successes, inflammation and adverse healing responses in these grafts are found to be a serious problem. Therefore, biointeractive or biomimetic prostheses, which have the hope of subduing adverse responses and facilitating the healing process, have been extensively developed in recent years.^[9,10]

A fundamental issue in any design of biointeractive or biomimetic prostheses is the dimensional scale of the interactions. It is known that tissue function is highly related to the tissue structure. The way in which cells interact with other cell types, the surrounding scaffold, and the signaling factors, is dependent on the complex micro- to nano-scale tissue structure. In addition, the effects of physical forces or electrical signals are also dependent on the tissue microstructure. Researchers have attempted to control all these conditions, collectively called microenvironments, precisely *in vitro* using various forms of nano- or microtechnology. When studying cell function and designing tissue replacements, microtechnology can help to replicate the cells' *in-vivo* dimensional, configurational, and compositional environment, thus creating more physiologically relevant models of cell and tissues.^[11-13] Although the promise of tissue engineering is rapidly becoming a

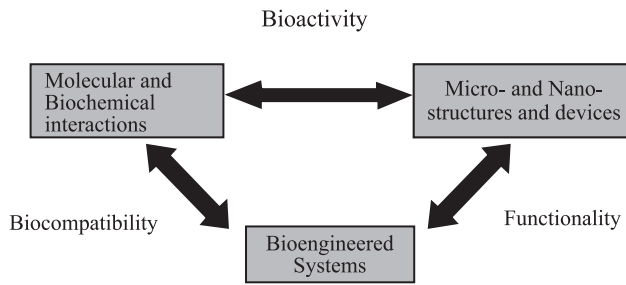


Fig. 1 The influence of the microscale on cellular and subcellular function.

reality, there are several challenges to overcome in the near future. This article will concentrate on one aspect of those challenges, namely the development of tissue-engineering constructs that more closely replicate complex tissue architecture and arrangement. It is through the incorporation of micro- and nanoscale features that we can significantly affect cellular and subcellular function (Fig. 1).

OVERVIEW OF MICROTECHNOLOGY

Traditional microfabrication techniques, such as micromachining using silicon wafer or glass, have been used widely in biological applications, but recently researchers have developed nano- and microtechnology using natural and synthetic polymers specifically for biology and medicine. Here the microtechnology, a collective term for micromachining, surface engineering, and soft lithography, will be briefly reviewed.

Micromachining

The processes of microfabrication and micromachining use photolithography, along with the deposition and selective etching of thin film layers, to create microscale (and even nanoscale) features on a given substrate. Structural materials, as opposed to sacrificial layers, complex features such as microchannels and micropores, as well as other two- and three-dimensional shapes, can be chemically etched in substrates. Using a series of combinations of additive and subtractive microfabrication processing, complex structure with multiple components can be fabricated. Micromachining has conferred many enhancing and enabling qualities to the devices. However, for tissue engineering purposes, sometimes silicon and glass are not ideal choices, so they should be either modified or replaced with biocompatible materials.

Surface Engineering

Surface engineering is the process of modifying the substrate surface with surface modification techniques.

Surface chemical modification of a substrate is one key to long-term interaction with biological systems. Control over biomaterial surface chemistry and atomic microstructure can result in a regulated cell response.^[14] In general, surface modifications fall into two categories: 1) chemically or physically altering the atoms, compounds, or molecules in the existing surface (treatment, etching, chemical modification), or 2) overcoating the existing surface with a material having a different composition (coating, grafting, thin film deposition) (see *Surface Modification*). After surface engineering, the surface can be either biointeractive or bioinert, providing more compatibility for tissue-engineering prosthesis.

Soft Lithography

Soft lithography employs polydimethyl siloxane (PDMS) elastomer to fabricate microstamps, microfluidic channels, and micromolds (Fig. 2).^[13,15] Although the original master is created via conventional photolithography, once the master is created, it can be reused many times outside of a clean room. Thus, the technique is cheaper and more amenable to different substrates^[16] and can be used on nonplanar surfaces as well.^[17]

Novel techniques using soft lithography have been developed and used extensively in biology and medicine. For example, microcontact printing technique employs a microstamp to transfer microscale patterns of chemicals or biomolecules onto a substrate, thus creating patterned surfaces.^[18–20] Microfluidic technique employs one or more paralleled or networked microchannels to pattern proteins or cells onto a substrate.^[21,22] Fluid flow at the microscale exhibits unique phenomena that can be leveraged to create multiple chemical environments.^[23] Laminar flow has been used with PDMS microchannels to adjacently pattern various proteins and cells on the culture substrates.^[24,25] Formation of microscale chemical gradients was created by microfluidics, and the neurons put into the environments with the gradients of chemicals oriented and differentiated according to the gradients.^[26,27] Three-dimensional microfluidics can pattern multiple cell types on the substrate.^[28] Detailed reviews on soft lithography have been given by Whitesides.^[13]

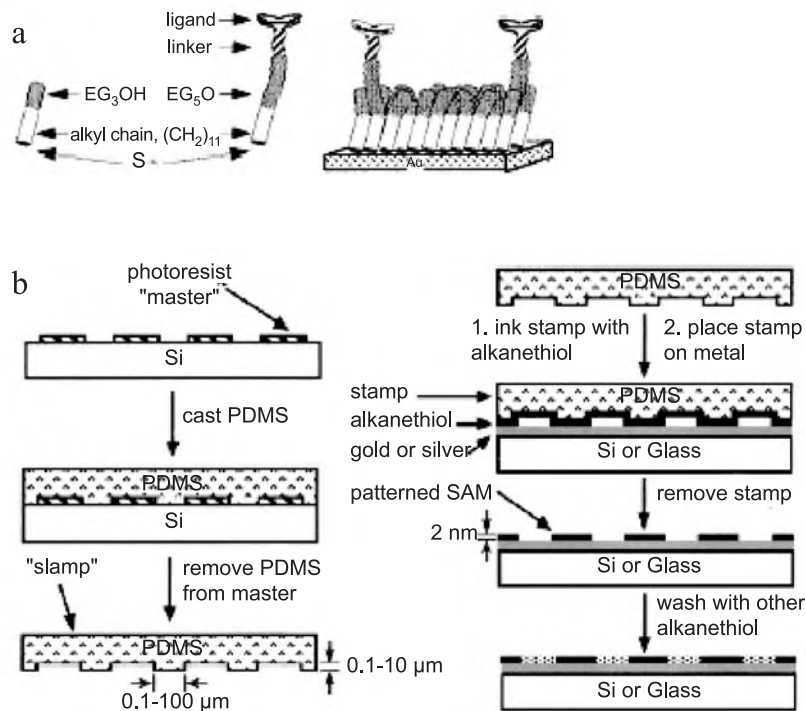


Fig. 2 Examples of surface engineering and soft lithography (a) self-assembled monolayers; (b) microcontact printing. (From Ref. [16].)

ACHIEVING CONTROL BY MICROT TECHNOLOGY

Equally as important as fabrication techniques, the next issue is the application of microtechnology tools in tissue engineering to achieve control over the microstructure and microenvironment, while at the same time developing an understanding of the control mechanism that will impact the tissue-engineering designs. Specifically, techniques that can be used to achieve topographical, spatial, chemical, and immunological controls over tissue-engineered constructs will be addressed. Though not highlighted here, the physical force is also a potential aspect for microcontrolling the environments of cells, for example, micro- and nanoscale mechanical stresses (mechanophysical patterning) generated by cell–biomaterial interactions, since they may have significant effects on phenotypic behavior.^[29]

Topographical Control

Topographic modulation of the tissue response can be one of the most important considerations during the design and manufacture of biomaterials.^[14,30] A tissue-engineered construct that has a well controlled microstructure will better maintain cell morphology, differentiation, and functionality over long periods of time. A key feature of these constructs will be the replication of in-vivo geometry and dimensional size scale that will aid in the maintenance of an in-vivo–like cell phenotype. Silicon micromachining techniques make it

relatively easy to generate the complex geometries on the nano- and microscale that are required to model physiological environments in living tissue. However, recently novel approaches have been developed to create microstructured polymers,^[31–35] which has opened up unique opportunities in substrate development for tissue-engineering applications. The development of microfabricated constructs using natural and polymeric biomaterials may allow us to engineer highly controlled interfaces that are: 1) more physiologically relevant and hierarchically arranged, and 2) biologically, mechanically, and optically compatible.

By combining living cells and two- and three-dimensional substrates with microfabricated topographical features, investigators have begun to address fundamental issues such as cell adhesion, orientation, and migration, as well as tissue integration and vascularization. For example, it has been shown that microtexture has significant effects on cell proliferation and adhesion.^[36] Surfaces with microtopography can augment fibroblast attachment and induce contact guidance of these cells.^[37] Furthermore, DNA profiles were influenced by microtexture and the fact that cells on 10- μm –wide grooves proliferated less than cells on 2- and 5- μm grooves.^[35,38] The arrangement of cells in controlled two- and three-dimensional arrangements has been shown to have beneficial effects on cell differentiation, maintenance, and functional longevity.^[39] In addition, many researchers have demonstrated that topography can influence and direct cell migration. The width and spacing of the lines can be used to control the

speed of cell movement. For example, etched parallel grooves can control cell spreading and migration in the axis of the lines, even when the depth and spacing of the grooved features are as small as 500 nm, about ten times smaller than the typical length of a mammalian cell.^[40] Therefore, the design of microtextured tissue-engineering matrices provides more physiological conditions under which to maintain cells.

Spatial Control

Creating a tissue-engineered product requires specific biological considerations regarding cell type, scaffold materials, bioreactor design, and environmental conditions (signaling proteins, neighboring cells, and physical stimuli). Thus, another important challenge in tissue engineering is to control the spatial organization of cells in their synthetic microenvironments. This is especially vital in the bioengineering of tissues that have multiple cell types and require precisely defined cell–cell and cell–substrate interactions. For example, the functioning of tissue such as retinal, cardiac, and vascular is dependent on the controlled orientation of multiple cell types. A key determination in the engineering of these tissues is how, and to what extent, this microenvironment, be it a two-dimensional surface or a three-dimensional matrix, can be controlled *in vitro* to recreate *in-vivo*-like architecture.

Microtechnology offers the potential to control cell–substrate, cell–cell, and cell–medium interactions on a micrometer scale. Microengineering the interactions between cells and their microenvironments is reviewed by Folch and Toner.^[2] Compared to the traditional cell culture where no or little control is able to be exerted on cells, defined architectures yield more predictable patterns of growth and differentiation. A control over the spatial assignment of cells and their environments effectively mimic the natural ordering of tissues.

Chemical Control

Microtechnology has also been used in the identification and application of defined chemistries in the extracellular microenvironment. The chemical patterning of adhesive and nonadhesive regions onto a substrate at the resolution of single cells has relied on a variety of techniques such as photolithographic patterning of surface chemistries with either covalent linking or chemical adsorption of the protein,^[41] microcontact printing of self-assembled monolayers of different chemistries or proteins,^[18–20] and solution coating of microfabricated surfaces.^[42] Using these approaches, substrates (typically planar) can be designed to promote cell attachment or cell repulsion, provided that the appropriate cell-specific ligands are used. Ultimately, these

approaches may provide clues about chemical cues that direct tissue assembly and regeneration.

Recent advances in cell biology have highlighted the importance of the ligand–receptor interaction for controlling cellular adhesion. Therefore, approaches have been developed to coat surfaces with extracellular matrix protein to create adhesive regions^[43] or to link integrin ligands covalently to the surface. The well known RGD cell adhesion ligand provides a simple mechanism of modifying surfaces on the microscale to mimic the extracellular matrix. Peptides that contain the Arg–Gly–Asp (RGD) attachment site, together with the integrins that serve as receptors for them, constitute a major recognition system for cell adhesion and cell–substrate interaction.^[44] It has been shown that substratum adhesiveness can be micropatterned using peptides to create a model for morphogenetic cues controlling cell behavior.^[45]

Another way to microengineer the interface is by covalently or noncovalently immobilizing bioinert hydrophilic polymer chains or heparin (anticoagulant) molecules onto the material surface. For example, polyethylene glycol (PEG), a water soluble, nontoxic, and nonimmunogenic polymer, has been shown to significantly decrease both protein and cell adsorption on several biomaterials.^[46,47] PEG can be micropatterned with RDG, or other molecular peptides, to create multifunctional surfaces that both repel unwanted proteins and promote cell interaction. By using functional groups such as PEG, surfaces can be tailor made to elicit specific cell and protein response as well as decrease adsorption of nonspecific biologicals. Furthermore, these linkage groups can be further modified by attaching biomolecules, immunomolecules, or recognition factors in order to create more biocompatible and biofunctional interfaces.^[48,49]

Immunological Control

Another key factor influencing the success of tissue-engineered constructs is the interaction with the immune response. The immune response (cell-mediated and humoral) accounts for both short- and long-term biological effects such as hypersensitivity, foreign body reactions/fibrous encapsulation, and ultimately rejection. With regard to acute rejection that occurs due to antibody recognition and complement fixation, investigators have attempted to hinder passage of these effectors by encapsulating cells with size-selective polymer membranes. Such immunoisolation membranes are able to shield transplanted cells via an immunoprotective semipermeable shell. Several diseases are being investigated for treatment by cell immunoisolation, including diabetes, growth hormone deficiency, and Parkinson's disease.^[50–53] Micro- and nanotechnology

may offer a better tool to fabricate membranes of specific pore size, allowing one to optimize the biocapsules specifically for the encapsulation of pancreatic cells.^[54] Such membranes are capable of completely blocking immunomolecules (with sizes of ≥ 15 nm), while allowing the passage of insulin and nutrients (with sizes of ≤ 6 nm).

DEVELOPMENT OF MICROSCALE TISSUE ENGINEERING IN DIFFERENT TISSUES

Most microtechnologies have been directed toward basic cell studies or cellular array for drug screening, as opposed to real tissue-engineered products. Tissue-engineered products are far more complicated than just cell patterns or controlled cell–cell interactions on a flat surface. Tissues and organs are very heterogeneous. They are composed of layers, membranes, cells, and channels. Each function requires the appropriate components: cells, and materials of appropriate sizes (typically micro- or nanoscale). Through these heterogeneous components, sophisticated functions can be achieved. Tissue function is modulated by an intricate architecture of cells and biomolecules on a micrometer scale. With the incorporation of microtechnology into biology, it is now possible to design surfaces that reproduce some of the aspects of that architecture.

The following paragraphs will review the efforts made toward microscale tissue-engineered products in some biological systems: pancreatic, hepatic, cardiovascular, skin, and neural. Although design principles and control mechanisms are shared in all the approaches, due to the difference in structure and functionality between organs, some microtechnology might be more advantageous for particular tissues.

Pancreatic Tissue Engineering

Desai et al.^[55,56] have created microfabricated nanoporous biocapsules for the transplantation of insulin-secreting cells. To turn the microfabricated biocapsules into a tissue-engineered product, several tissue-engineering related issues have been taken into consideration: immunoprotection of encapsulated cells, biocompatibility of the implant, and the effect of microenvironment on cell viability and functions.^[57]

A schematic diagram of the microfabricated biocapsule is shown in Fig. 3. The basic technology that was developed for the pores themselves was the use of a sacrificial oxide sandwiched between silicon layers, thus defining a space that could be opened by a subsequent etching of the oxide in hydrogen-fluoride. To create the final immunoisolation capsule, the silicon substrate is etched up to the membrane, leaving a cavity in the wafer with an encapsulating immunoisolation membrane. Pancreatic islets are placed within the membrane-bound cavities and the two separate microfabricated membranes are bonded together.

The functionality of insulinoma cells, mouse beta cell tumor (betaTC-3) cells, encapsulated in microfabricated biocapsules and unencapsulated cells, has been evaluated in terms of insulin secretion. The amounts were similar in concentration and time release suggesting that glucose was able to pass through the membrane pores sufficiently to stimulate islets for insulin production. Figure 4 shows the typical insulin release profile in response to stimulatory (16.7 mM) glucose. At the same time, the biocapsule excludes the passage of IgG.

The ability to prolong pancreatic islet cells' viability and functionality within the immuno-isolation device can be achieved by providing cells within a three-dimensional matrix that would ideally mimic the biological extracellular matrix^[58] and therefore the in-vivo

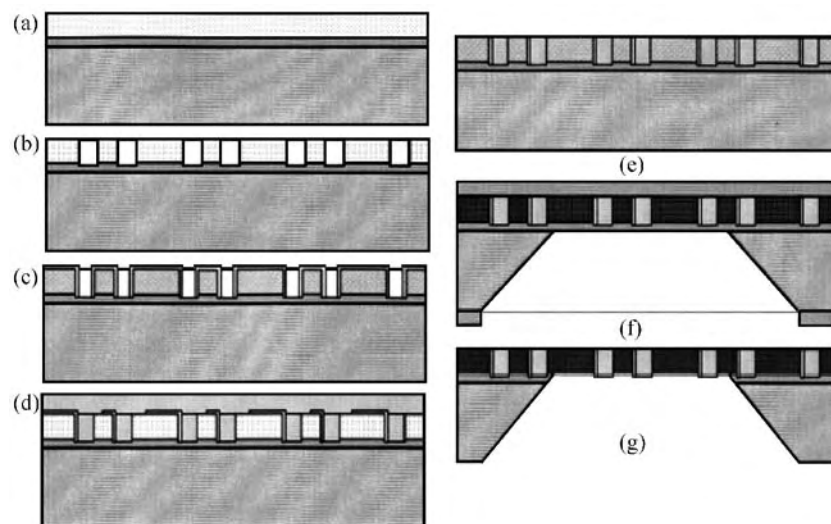


Fig. 3 Fabrication protocol for nanopore membranes: (a) growth of buried nitride layer and base layer deposition; (b) hole definition in base layer; (c) growth of thin sacrificial oxide to determine channel size; (d) pattern of anchor points and deposition of polysilicon plug layer; (e) planarization through plug layer to base; (f) deposition of protective nitride layer, backside patterning, and through-wafer etch; and (g) final release of structure in HF.

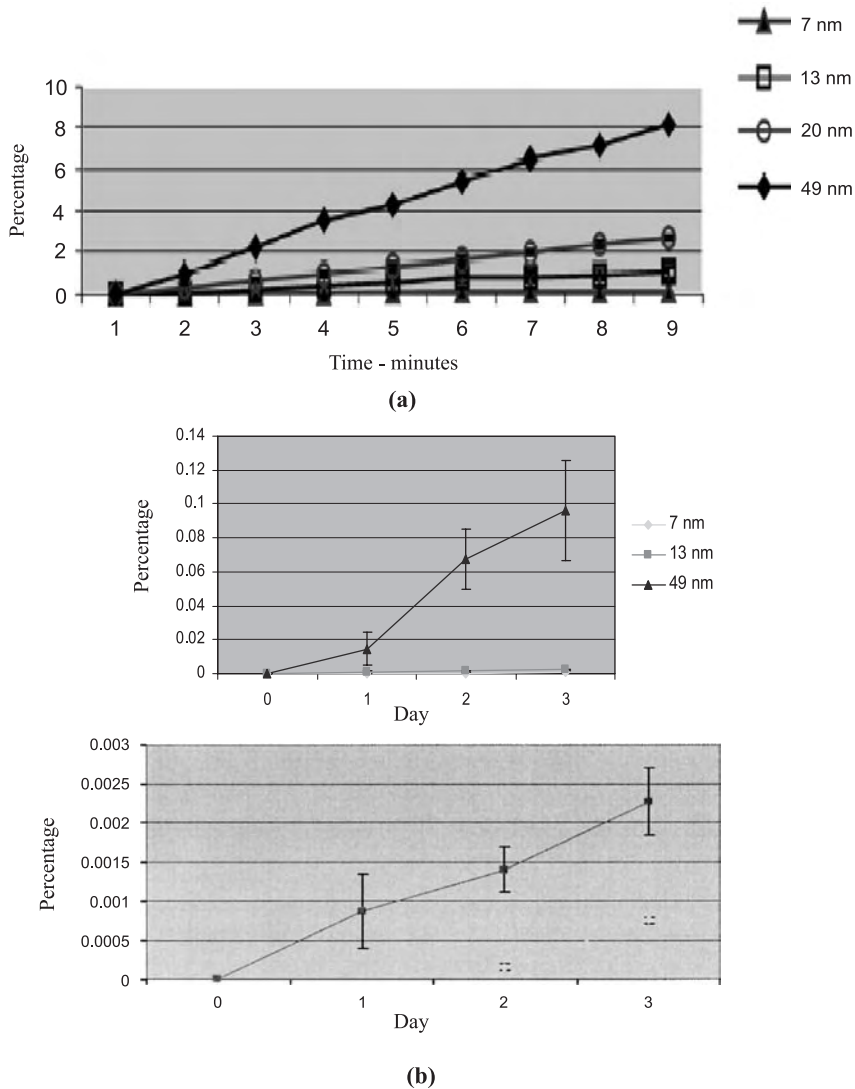


Fig. 4 Microfabricated nanoporous membranes provide immunoprotection to the capsuled cells. Diffusion profiles of glucose (a) and immunoglobulin G (b).

architecture of the endocrine pancreas. BetaTC-3 cells were cultured within three-dimensional matrices and compared with cells grown in standard conditions. At day 6, a rearrangement of the cell spatial distribution is observed in the collagen–chitosan matrix and small cell clusters are visible (Fig. 5). Collagen seemed to perform better as a matrix material in the biocapsule, preventing aggregation and maintaining a homogeneous cell distribution. Higher diffusion was associated with collagen and collagen–chitosan, while little diffusion was detected from biocapsules with a simple cell suspension (Fig. 5). The use of matrices, and collagen in particular, proved to augment the functionality of betaTC-3 cells in microfabricated biocapsules.

In order to retain the same performances in vivo as in vitro, the biocapsule must be fully biocompatible and able to integrate into the host body, eliciting a minimal foreign body response. The degree of the foreign body response is known to depend upon several parameters, such as size of the device, the outer morphology of the

capsule, the surgical technique, and the site of implantation. For example, the responses of the biocapsule sutured to the abdominal wall or within the omentum were quite different. And PEG-modification of the membranes significantly improved biocompatibility (Fig. 6).

An immunoprotective capsule has the advantage of being able to use different cell sources other than the cells originated from the host, such as xenogeneic cells and tumor cell lines. Therefore, the cell source for a tissue-engineered product is significantly increased.

Liver Tissue Engineering

The liver is a multicellular complex organ. The adult liver provides a scaffold for many complex cell–cell interactions that allow for effective, coordinated organ function. Coculture is emerging as a necessity in successful designs of the bioartificial liver (see *Liver, Bioartificial; Tissue Engineering of Liver*). Bhatia et al.

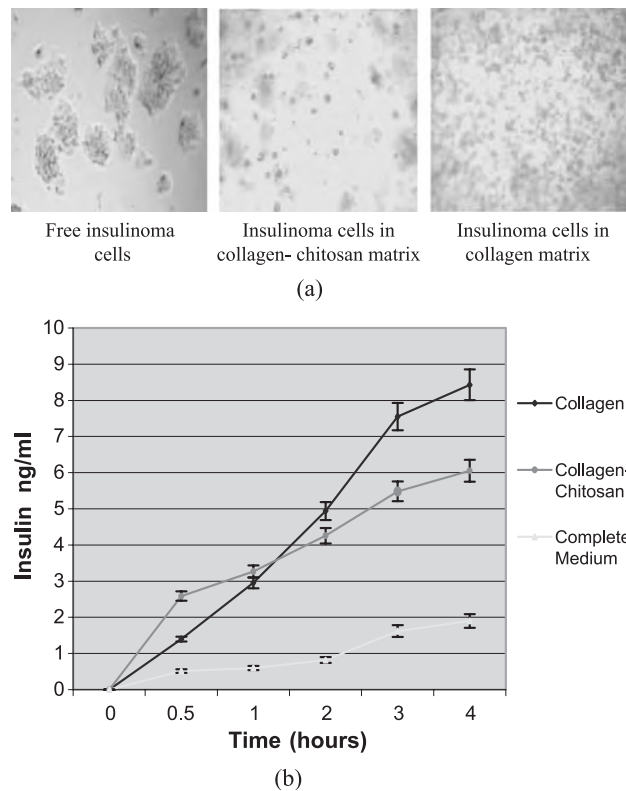


Fig. 5 (a) Cell distribution in different living conditions at day 6. (b) Insulin secretory profile through microfabricated biocapsules.

investigated microscale cell-cell interactions in cocultures for enhancing the liver-specific functions of cells. They developed a technique for micropatterning two different cell types based on strategies for surface modification with aminosilanes linked to biomolecules and the manipulation of serum content of cell culture media. Through this, the spatial organization of hepatocytes and fibroblasts could be precisely controlled. This approach could eliminate variations present in random mixed cocultures. The amount of heterotypic interface and homotypic interface, as well as ratios between two cell types, can be modulated precisely

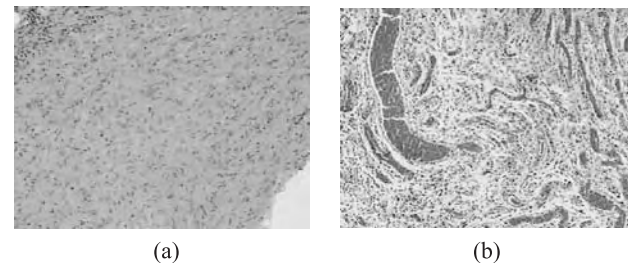


Fig. 6 Microscopic appearance of tissue surrounding the subcutaneous implants retrieved from Lewis rat after 17 days. Two types of implants were used: unmodified silicon membrane (a) and PEG-modified membrane (b).

(Fig. 7). As a result of the micropatterned coculture, liver-specific function, measured by albumin and urea synthesis, is dependent on the amount of heterotypic interface in the coculture, and albumin production is localized to hepatocytes at the interface. Increase in hepatocyte/fibroblast interaction achieved through micropatterning produced higher long-term function of cultures with similar cell populations but different spatial arrangements.

Another approach, named a vascularized artificial liver, used micromachining technologies on silicon and Pyrex surfaces to generate the branched architecture of vascular network. Hepatocytes and endothelial cells were cultured and subsequently lifted as single-cell monolayers from the two-dimensional molds. Hepatocytes maintained albumin production on the monolayer. The lifted monolayers were then folded into compact three-dimensional tissues with the intent to produce an artificial liver.^[54]

Cardiovascular System Tissue Engineering

The cardiovascular system is a multicomponent, hierarchical system providing for blood circulation. The heart works like a pump to generate forces by beating, and hierarchical vascular systems, including arteries,

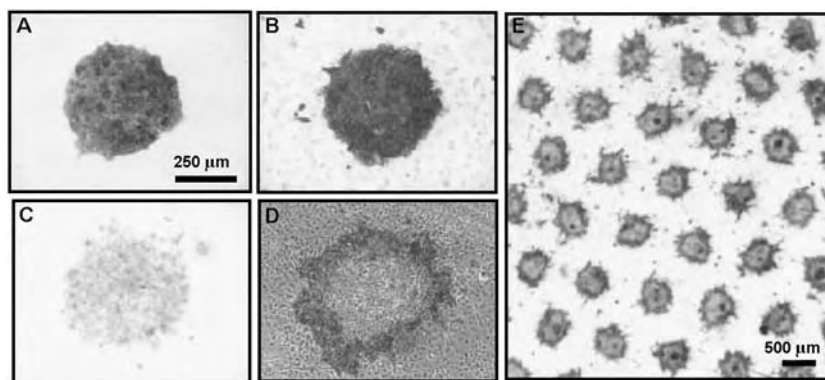


Fig. 7 An example of cocultured cells for microscale liver engineering. Compared to micropatterned hepatocytes (alone) on days 1 and 6 (A, C), micropatterned cocultures on days 1 and 6 (B, D) did not display a uniform decline in albumin content.



veins, and capillaries, serve as pipes and tubing for the passage of the blood and exchange of substances between blood and tissues. Tissue engineering of a heart is clearly an ambitious undertaking. But the problem can be broken down into more manageable component tissues and interim milestones.

Forces for the beating of the heart are generated by the contraction of myocytes, the major cell type forming the heart. Therefore, the morphology, orientation, polarization, and mechanobiology of the myocytes in culture are very important. Recapitulating physiological gene expression in cell culture by development of a microfabricated three-dimensional system has been investigated. Both myocytes and nonmyocytes (mainly fibroblasts) from neonatal rat hearts were studied on micropegged and microgrooved substrata. The procedure for generating the microtextured substrate is shown in Fig. 8. Some unique phenotype characteristics were found when cells were cultured on the microtextured surfaces, but not in random culture. It has been observed that three-dimensional surface topography of culture substrates significantly affects in-vitro cardiac myocyte orientation (Fig. 9)^[31] and attachment. Cardiac myocytes prefer to terminate on vertical structures with a subsequent increase in cell height, which is more representative of in-vivo behavior. By creating culture surfaces with topographic features corresponding to cellular dimensions (10–50 μm size range), cells seem to exhibit a more in-vivo-like cellular morphology. The microtopography provides anisotropic or directional growth for cells and thus can recreate tissue architecture at the cellular and subcellular level in a reproducible fashion, facilitating the culture and maintenance of differentiated cell states. The expressions of N-cadherin and connexin-43 for myocytes cultured on a specific microgroove feature were found to be close to those in neonate tissues.^[59]

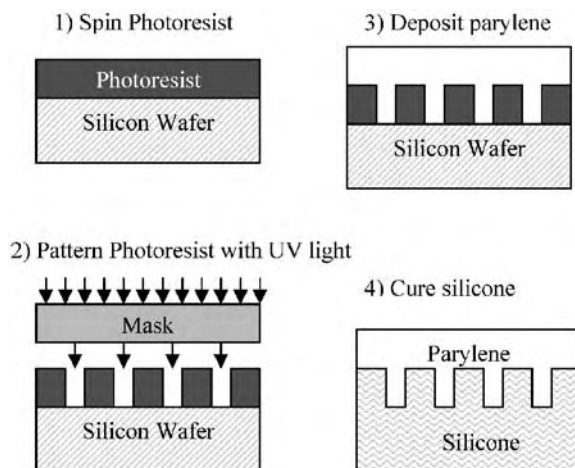


Fig. 8 Example of a process for creating microtextured polymers.

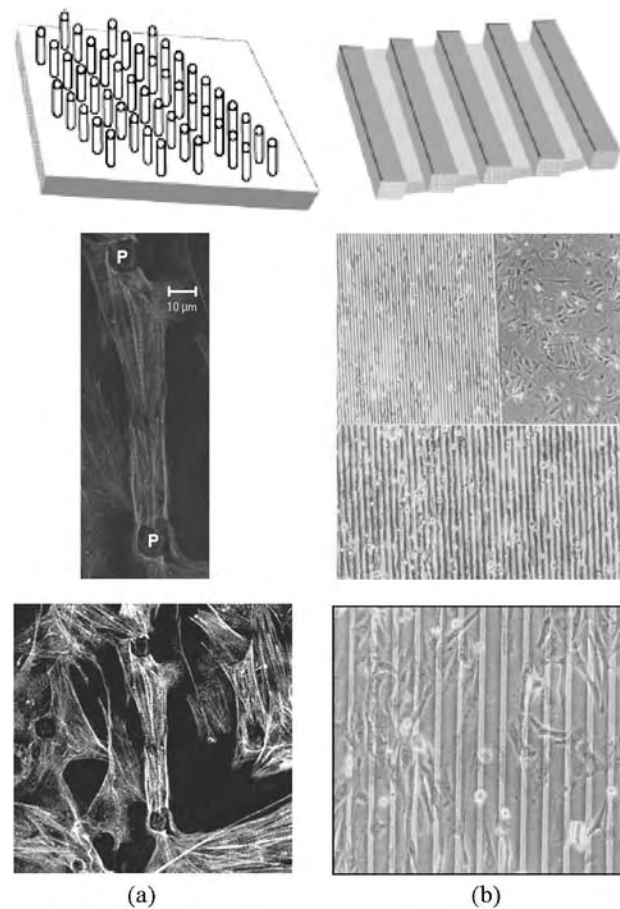


Fig. 9 Micropegs (a) and microgrooves (b) can be used to create microarchitectures that direct cardiac myocyte attachment and morphology.

The normal arteries comprise three lamellar layers: the adventitia, the medial, and the endothelium or intima. The adventitia layer, mainly composed of fibroblasts and collagen, harbors blood and nerve supply; the medial layer, composed of smooth muscle cell and laminin, contributes the bulk of the mechanical strength and controls the vessel size in response to external stimuli; and the endothelium, composed of endothelial cells and basement membrane, prevents spontaneous blood clotting. Mimicry of some or all of the properties of the three layers of a healthy artery has been the strategy of many arterial tissue engineering approaches (see *Tissue Engineering of Blood Vessel*). Current biomimetic approaches of vascular tissue engineering to reproduce the three-tunic structure and mimic some properties are primarily focused on the macroscale; however, microvascular tissue engineering has had less progress. Because of their extremely small size and fragility, it is difficult to assemble very small vessels, such as capillaries, individually in vitro and then later implant them where they are needed. Instead, the approaches considered to date rely on bulk culturing of large

quantities of capillary endothelial cells. Dike et al.^[60] described a microcontact printing technique using SAM for patterning substrates for bovine capillary endothelial cell culture that allowed the control of positions and sizes of attached cells. But this study focused primarily on the demonstration of the underlying technology instead of cell and tissue engineering. Recently, Tan and Desai^[61] developed a layer-by-layer microfluidics technique that provides a tool for bioengineering microscale vascular system with three-layer structure. This method involves a surface modification process and a layer-by-layer microfluidics technique. By taking advantages of matrix contraction by cells, cell-matrix assemblies were delivered and positioned inside the microchannels, thus forming a multilayer configuration inside the channels (Fig. 10). The multilayer having mimetic structure of an artery can be considered a neotissue *in vitro*. The matrix materials that were used for building up the structure were found to influence the 3-D cell migration and the neotissue architecture. It was also shown that the neotissue configurations affect the nature and degree of homotypic and heterotypic cell-cell interactions, which may impact the biological function of the tissue (e.g., ICAM-1 expression in endothelial cells).

Neural Tissue Engineering

Due to its bioelectronic properties, the nervous system is the first biological system in which researchers

extensively employed microtechnology for the construction of artificial implants,^[62] for example, brain stem implants, cochlear implants, retinal prostheses, and cortical prostheses. In the nervous system, specific neurons form complicated and sensitive structures to sense the environments and transmit and process information. To maintain proper signal processing, the intricate network of connections among the neurons is extremely essential. Microtechnology can fabricate structures of a size comparable with the size of the neurons, thus controlling the connections on a one-to-one basis. Furthermore, electrodes with even smaller feature sizes that can interact with nerve fibers can also be fabricated. Microtechnology makes the neuron culture meaningful for signal processing. Microelectrodes for sensing and recording the electrical signals and thus stimulating nerve cells are popular approaches in neural engineering. But recently, the culture of micropatterned neurons using soft lithography has gained more attention.^[63,64] Thus, combining microelectrodes with micropatterned neurons is a new approach in the tissue engineering of the nervous system. An example of this approach is recent work from Lauer et al.,^[65] in which they developed a setup to perform aligned microcontact printing of extracellular matrix proteins on microelectronic devices in order to guide the growth of electrogenic cells specifically to these sensitive spots (Fig. 11). Aligned neuronal growth on patterned devices was demonstrated with dissociated hippocampal neurons.

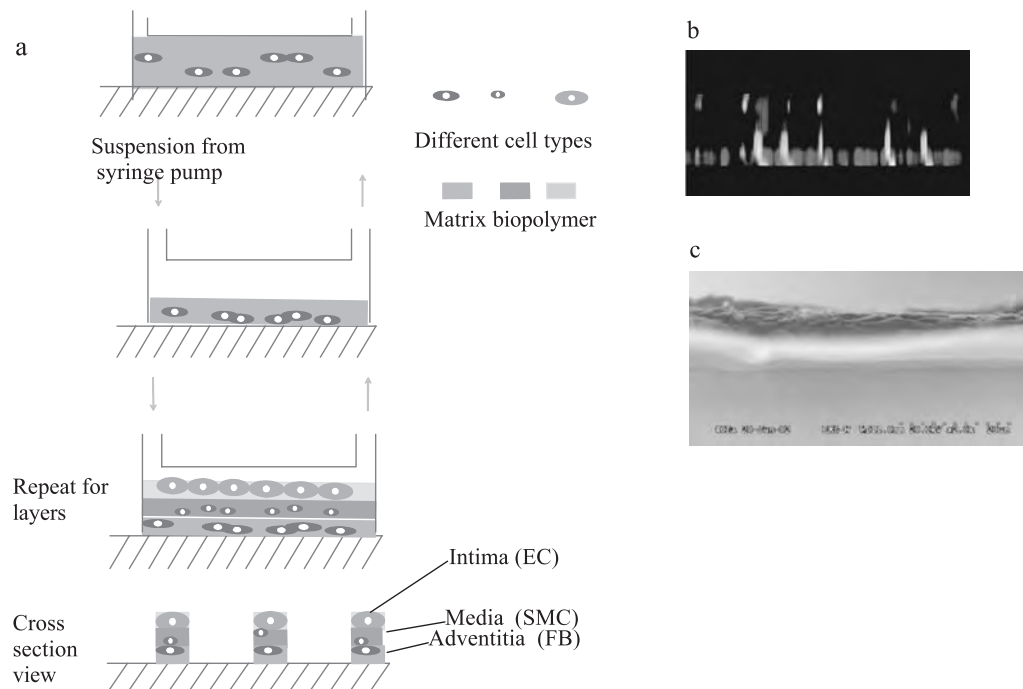


Fig. 10 Multilayer microfluidic approach for the 3-tunic vascular system: (a) setup protocol; (b) reconstituted fluorescence image shows a 3-layer structure; (c) SEM picture shows a 2-layer structure.

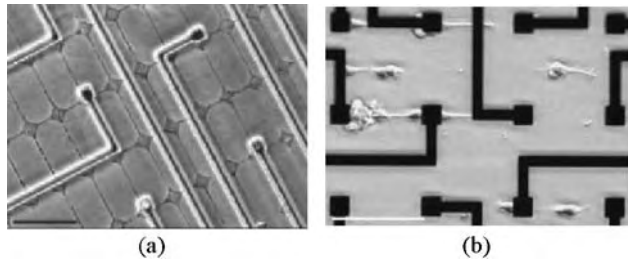


Fig. 11 An example of microscale tissue engineering of neural system: (a) a laminin grid pattern aligned to the microelectrodes of a microelectronic recording device; (b) Hippocampal neurons cultured on the pattern for 2 days.

Skin Tissue Engineering

As an effort to microscale tissue engineering of skin, Toner's group produced analogs of the basal lamina with complex topographic features using a microfabrication approach.^[66] Microfabricated pattern, produced by laser machining, was used to mold PDMS elastomer replicates, which were then used as templates for the production of thin membranes of collagen or gelatin. The pattern has features of ridges and channels with a depth scale at 40 to 210 μm and length scales (40 to 210 μm) similar to invaginations found in a native basal lamina. Both the surface of the PDMS and the highly porous collagen-GAG sponge were shown to replicate the features on the master faithfully. Keratinocytes were seeded and cultured onto the collagen or gelatin membrane for composite dermal analogs/equivalents. The equivalents were cultured at the air-liquid interface to induce the differentiation and cornification of the epidermal layer. The keratinocytes formed a translucent, hydrophobic epidermis with gross topographic features that correlated with the microfabricated surface of the membrane. The epidermal layer was found to be composed of differentiated and stratified keratinocytes that conformed to the surface of the skin equivalent, forming ridgelike structures comparable in dimension to those observed in native skin (Fig. 12). And they also found that stratification was enhanced in the deeper

channels, which may be due to the effects of the topographical microenvironment on keratinocyte differentiation/proliferation.

CONCLUSIONS AND FUTURE DIRECTIONS

Although different technologies have been employed in microscale tissue engineering, such as membranes with nanopores, microelectrodes, microfluidic channels, microtextured surfaces, and functional thin-film substrates, they all share the same goal—to fabricate functional biomimetic implantable replacements for tissue. These microscale formats, reminiscent of the precise in-vivo assembly in the tissues, have provided creative tissue equivalents having some enabling functions that macroscale products cannot achieve. Overall, microscale tissue engineering has many advantages over the traditional tissue engineering strategies: 1) It recapitulates complex tissue functions that cannot be easily mimicked with presently employed macromaterials and devices; 2) it provides an implantable biointerface; 3) it provides controlled and predictable interactions and phenomena between cells and materials. Although we are getting closer to the clinical applications of tissue-engineered constructs, many challenges and opportunities remain for the future.

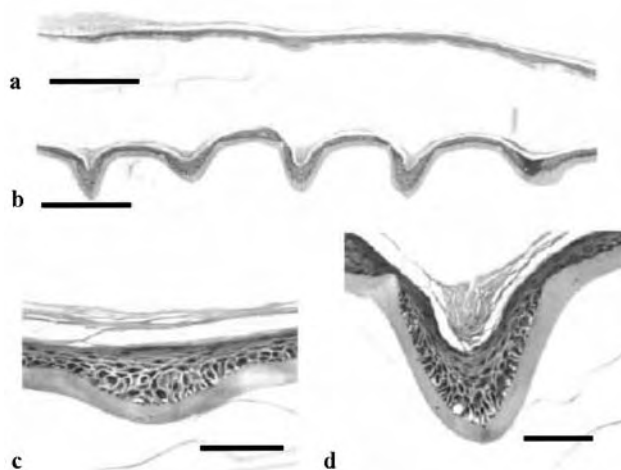


Fig. 12 Skin equivalent with a microfabricated analog of the basal lamina. Keratinocytes were seeded on the surface of a composite dermal analog and grown at the air/liquid interface. Low magnification of a skin equivalent with the shallowest (a) or deepest set of channels (b). High magnification micrographs show differentiated and stratified keratinocytes in individual channels of a microfabricated membrane (c, d).

Multidimensional: 2-D, 3-D, and 4-D

It is now conceivable to achieve biological control by using micro/nanotechnologies that are able to attach, position, and manipulate individual cells and proteins in vitro in two dimensions. However, tissue-engineered products are far more complicated than one or two types of cells attaching and growing on top of a solid surface; they are rather well designed, precisely controlled, heterogeneous and multicomponent 3-D entities. Therefore, although promising, microscale tissue engineering still has a long way to go to achieve clinical and commercial success. Future developments in the design of tissue substitutes will be aided by more understanding of the effects of 3-D scaffolds and tissue architecture on cell function.

The engineering of many tissue structures, such as the branching architecture that is found in many tissues or the intricate network architecture of the nervous system, will probably require methods for introducing and changing molecular signals during the process of neotissue development.^[67] For example, changes in agent concentration with time can influence the differentiated state of cells within tissues. Smart ways to control the microscale features or micropatterned cells or molecules with time may be essential for the next generation of tissue-engineered products. Time adds another dimensional challenge to future research. Therefore, spatio-temporal challenges to microscale tissue engineering are in front of us. A three-dimensional tissue replacement with long lifetime will be a better tissue-engineered product.

Convergence with Other Technique Advancements

Microtechnology alone can not recapitulate adult tissue structure and function. Convergence with other advancements in tissue engineering may open up more opportunities for microscale tissue engineering. For example, the interests of stem cells for cell sourcing are increasing dramatically recently. With the recent advances in stem cell biology, stem cells may be a popular cell source for tissue-engineered products in the future. The primary challenge for stem cell biology is to determine biological and physical cues leading to differentiation of stem cells into functional tissue units (see *Stem Cells*). As reviewed above, microtechnologies offer tools to well controlled microenvironments so that they could be not only useful in developing a detailed understanding of the cues for differentiation of stem cells, but also helpful in guiding the development of cells into the desired tissue product.

Furthermore, many new types of biomaterials are under development. Nanostructured biomaterials

fabricated by self-assembly methods can be used in the tissue engineering fields as well.^[68] With the advance of combinatorial approaches, production of novel materials increases dramatically, allowing the production of more biocompatible polymers suitable for microscale tissue engineering.

In-Vitro Research into In-Vivo Implants

Finally, tissue-engineered products, different from other in-vitro applications such as drug screening, should be aimed at integrating with in-vivo tissues and replacing damaged tissue. The future challenges lies in assembling and applying our collective knowledge, in engineering, biology, chemistry, and physics, to develop functional and clinically relevant tissue substitutes. It is also hoped that we can explore the use of these new tools not only to create more complex substrates for tissue engineering applications, but also to begin to understand fundamental processes in cell biology better.



ARTICLES OF FURTHER INTEREST

Cardiac Patch Engineering; Liver, Bio-Artificial; Microelectromechanical Systems (MEMS) Manufacturing; Nerve Guides; Neuroprostheses; Tissue Engineering; Tissue Engineering of Cardiac Muscle; Tissue Engineering of Skin

REFERENCES

1. Langer, R.; Vacanti, J.P. Tissue Engineering. *Science* **1993**, *260*, 920–925.
2. Folch, A.; Toner, M. Microengineering of cellular interactions. *Annu. Rev. Biomed. Eng.* **2000**, *2*, 227–256.
3. Bhatia, B.N.; Chen, C.S. Tissue engineering at the micro-scale. *Biomed. Microdevices* **1999**, *2*, 131–144.
4. Griffith, L.G. Emerging design principle in biomaterials and scaffolds for tissue engineering. *Ann. N.Y. Acad. Sci.* **2002**, *969*, 83–95.
5. Wilson, G.J.; MacGregor, D.C.; Klement, P.; Weber, B.A.; Binnington, A.G.; Pinchuk, L. A compliant Corethane/Dacron composite vascular prosthesis. Comparison with 4-mm ePTFE grafts in a canine model. *ASAIO J.* **1993**, *39*, M526–M531.
6. Cantelmo, N.L.; Quist, W.C.; Lo Gerfo, F.W. Quantitative analysis of anastomotic intimal hyperplasia in paired Dacron and PTFE grafts. *J. Cardiovasc. Surg.* **1989**, *30*, 910–915.
7. Godbey, W.T.; Atala, A. In vitro systems for tissue engineering. *Ann. N.Y. Acad. Sci.* **2002**, *961*, 10–26.
8. Rutkowski, G.E.; Heath, C.A. Development of a bioartificial nerve graft. II. Nerve regeneration in vitro. *Biotechnol. Prog.* **2002**, *18*, 373–379.

9. Zilla, P.; Greisler, H.P. *Tissue Engineering of Vascular Prosthetic Grafts*; R G Landes Company: Austin, TX, 1999.
10. Nerem, R.M.; Seliktar, D. Vascular tissue engineering. *Annu. Rev. Biomed. Eng.* **2001**, *3*, 225–243.
11. Desai, T.A. Micro- and nanoscale structures for tissue engineering constructs. *Med. Eng. Phys.* **2000**, *22*, 595–606.
12. Prokop, A. Bioartificial organs in the twenty-first century: Nanobiological devices. *Ann. N.Y. Acad. Sci.* **2001**, *944*, 472–490.
13. Whitesides, G.M.; Ostuni, E.; Takayama, S.; Jiang, X.; Ingber, D.E. Soft lithography in biology and biochemistry. *Annu. Rev. Biomed. Eng.* **2001**, *3*, 335–373.
14. Ito, Y. Surface micropatterning to regulate cell functions. *Biomaterials* **1999**, *20*, 2333–2342.
15. Xia, Y.; Whitesides, G.M. Soft lithography. *Angew. Chem. Int. Ed.* **1998**, *37*, 550–575.
16. Kane, R.S.; Takayama, S.; Ostuni, E.; Ingber, D.E.; Whitesides, G.M. Patterning proteins and cells using soft lithography. *Biomaterials* **1999**, *20*, 2363–2376.
17. Patel, N.; Padera, B.; Sanders, G.H.W.; Cannizzaro, S.M.; Davies, M.C.; Langer, R.; Roberts, C.J.; Tendler, S.J.B.; Williams, P.M.; Shakesheff, K.M. Spatially controlled cell engineering on biodegradable polymer surfaces. *FASEB J.* **1998**, *12*, 1447–1454.
18. Mrksich, M.; Whitesides, G.M. Using self-assembled monolayers to understand the interactions of man-made surfaces with proteins and cells. *Annu. Rev. Biophys. Biomol. Struct.* **1996**, 155–178.
19. Chen, C.S.; Mrksich, M.; Huang, S.; Whitesides, G.M.; Ingber, D.E. Micropatterned surfaces for control of cell shape, position, and function. *Biotechnol. Prog.* **1998**, *14*, 356–363.
20. Kumar, A.; Whitesides, G. Features of gold having micrometer to centimeter dimensions can be formed through a combination of stamping with an elastomeric stamp and an alkanethiol ink followed by chemical etching. *Appl. Phys. Lett.* **1993**, *63*, 2002.
21. Folch, A.; Toner, M. Cellular micropatterns on biocompatible materials. *Biotechnol. Prog.* **1998**, *14*, 388–392.
22. Raty, S.; Davis, J.A.; Beebe, D.J.; Rodriguez-Zas, S.L.; Wheeler, M.B. Culture in microchannels enhances in vitro embryonic development of preimplantation mouse embryos. *Theriogenology* **2001**, *55*, 241.
23. Beebe, D.J.; Mensing, G.A.; Walker, G.M. Physics and applications of microfluidics in biology. *Annu. Rev. Biomed. Eng.* **2002**, *4*, 261–286.
24. Takayama, S.; Ostuni, E.; LeDue, P.; Naruse, K.; Ingber, D.; Whitesides, G. Laminar flows: Subcellular positioning of small molecules. *Nature* **2001**, *411*, 1016.
25. Takayama, S.; McDonald, J.C.; Ostuni, E.; Liang, M.N.; Kenis, P.J.A.; Ismagilov, G.V.; Whitesides, G. Patterning cells and their environments using multiple laminar fluid flows in capillary networks. *Proc. Natl. Acad. Sci.* **1999**, *96*, 5545–5548.
26. Dertinger, S.K.; Jiang, X.; Li, Z.; Murthy, V.N.; Whitesides, G.M. Gradients of substrate-bound laminin orient axonal specification of neurons. *Proc. Natl. Acad. Sci.* **2002**, *99*, 12542–12547.
27. Li Jeon, N.; Baskaran, H.; Dertinger, S.K.; Whitesides, G.M.; Van de Water, L.; Toner, M. Neutrophil chemotaxis in linear and complex gradients of interleukin-8 formed in a microfabricated device. *Natl. Biotechnol.* **2002**, *20*, 826–830.
28. Chiu, D.T.; Jeon, N.L.; Huang, S.; Kane, R.S.; Wargo, C.J.; Choi, I.S.; Ingber, D.E.; Whitesides, G.M. Patterned deposition of cells and proteins onto surfaces by using three-dimensional microfluidic systems. *Proc. Natl. Acad. Sci.* **2000**, *97*, 2408–2413.
29. Wang, N.; Ostuni, E.; Whitesides, G.M.; Ingber, D.E. Micropatterning tractional forces in living cells. *Cell Motil. Cytoskelet.* **2002**, *52*, 97–106.
30. Curtis, A.; Wilkinson, C. Topographical control of cells. *Biomaterials* **1997**, *18*, 1573–1583.
31. Deutsch, J.; Motlagh, D.; Russell, B.; Desai, T.A. Fabrication of microtextured membranes for cardiac myocyte attachment and orientation. *J. Biomed. Mater. Res. (Appl Biomater)* **2000**, *53*, 267–275.
32. Brunette, D.M. Fibroblasts on micromachined substrata orient hierarchically to grooves of different dimensions. *Exp. Cell Res.* **1986**, *164*, 11–26.
33. Schmidt, J.A.; von Recum, A.F. Texturing of polymer surfaces at the cellular level. *Biomaterials* **1991**, *12*, 385–389.
34. Van Kooten, T.G.; Whitesides, J.F.; von Recum, A.F. Influence of silicone surface texture on human skin fibroblast proliferation as determined by cell cycle analysis. *J. Biomed. Mater. Res.* **1998**, *43*, 1–14.
35. Kapur, R.; Spargo, B.J.; Chen, M.S.; Calvert, J.M.; Rudolph, A.S. Fabrication and selective surface modification of 3-dimensionally textured biomedical polymers from etched silicon substrates. *J. Biomed. Mater. Res.* **1996**, *33* (4), 205–216.
36. Den Braber, E.T.; de Ruijter, J.E.; Ginsel, L.A.; von Recum, A.F.; Jansen, J.A. Orientation of ECM protein deposition, fibroblasts cytoskeleton, and attachment complex components on silicone microgrooved surfaces. *J. Biomed. Mater. Res.* **1998**, *40* (2), 291–300.
37. Meyle, J.; Gultig, K.; Nisch, W. Variation in contact guidance by human cells on a microstructured surface. *J. Biomed. Mater. Res.* **1995**, *29* (1), 81–88.
38. Chou, L.; Firth, J.D.; Uitto, V.J.; Brunette, D.M. Substratum surface topography alters cell shape and regulates fibronectin mRNA level, mRNA stability, secretion and assembly in human fibroblasts. *J. Cell. Sci.* **1995**, *108* (4), 1563–1573.
39. Kapur, R.; Calvert, J.M.; Rudolph, A.S. Electrical, chemical, and topological addressing of mammalian cells with microfabricated systems. *J. Biomech. Eng.* **1999**, *121* (1), 65–72.
40. Flemming, R.G.; Murphy, C.J.; Abrams, G.A.; Goodman, S.L.; Nealey, P.F. Effects of synthetic micro- and nano-structured surfaces on cell behavior. *Biomaterials* **1999**, *20* (6), 573–588.
41. Drumheller, P.; Hubbell, J. Surface Immobilization of Adhesion Ligands for Investigations of Cell-Substrate Interactions. In *The Biomedical Engineering Handbook*; Bronzino, J., Ed.; CRC Press: Boca Raton, FL, 1995.

42. Zhang, M.; Ferrari, M. Hemocompatible polyethylene glycol films on silicon. *Biomed. Microdevices* **1998**, *1*, 81–89.
43. Hubbell, J.A. Biomaterials in tissue engineering. *Biotechnology* **1995**, *13*, 565–576.
44. Ruoslahti, E. RGD and other recognition sequences for integrins. *Annu. Rev. Cell Dev. Biol.* **1996**, *12*, 697–715.
45. Britland, S.; Clark, P.; Connolly, P.; Moores, G. Micropatterned substratum adhesiveness: A model for morphogenetic cues controlling cell behaviour. *Exp. Cell. Res.* **1992**, *198*, 124–129.
46. Aldenhoff, Y.B. Studies on a new strategy for surface modification of polymeric biomaterials. *J. Biomed. Mater. Res.* **1992**, *26*, 917–928.
47. Tseng, Y.C.; Park, K. Synthesis of photoreactive poly(ethyleneglycol) and its application. *J. Biomed. Mater. Res.* **1992**, *26*, 373–391.
48. Zhang, M.; Desai, T.A.; Ferrari, M. Proteins and cells on PEG immobilized silicon surfaces. *Biomaterials* **1998**, *19*.
49. Sharma, S.; Popat, K.C.; Desai, T.A. Controlling nonspecific protein interactions in silicon biomicrosystems with nanostructured poly(ethylene glycol) films. *Langmuir* **2002**, *18*, 8728–8731.
50. Lanza, R.P.; Chick, W.L. Encapsulated cell therapy. *Sci. Am. Sci. Med.* **1995**, 16–25.
51. Colton, C.K. Implantable biohybrid artificial organs. *Cell Transplant* **1995**, *4*, 415–436.
52. Fu, X.W.; Sun, A.M. Microencapsulated parathyroid cells as a bioartificial parathyroid. In vivo studies. *Transplantation* **1989**, *47*, 432–435.
53. Aebischer, P.; Winn, S.R.; Galletti, P.M. Transplantation of neural tissue in polymer capsules. *Brain Res.* **1988**, *448*, 364–368.
54. Kaihara, S.; Borenstein, J.; Koka, R.; Lalan, S.; Ochoa, E.R.; Ravens, M.; Pien, H.; Cunningham, B.; Vacanti, J.P. Silicon micromachining to tissue engineer branched vascular channels for liver fabrication. *Tissue Eng.* **2000**, *6*, 105–117.
55. Desai, T.A. Microfabricated interfaces: New approaches in tissue engineering and biomolecular separation. *Biomolecular Engineering* **2000**, *17*, 23–26.
56. Desai, T.A. Microfabrication technology for pancreatic cell encapsulation. *Exp. Opin.* **2002**, *2*, 633–646.
57. Leoni, L.; Desai, T.A. Nanoporous biocapsules for the encapsulation of insulinoma cells: Biotransport and biocompatibility considerations. *IEEE Trans. Biomed. Eng.* **2001**, *48* (11), 1335–1341.
58. Tan, W.; Krishnaraj, R.; Desai, T.A. Evaluation of nanostructured composite collagen–chitosan matrices for tissue engineering. *Tissue Eng.* **2001**, *7*, 203–211.
59. Motlagh, D.; Hartman, T.J.; Desai, T.A.; Russell, B. Microfabricated grooves recapitulate neonatal myocyte connexin43 and N-cadherin expression and localization. *Biomaterials* **2002**, *in press*.
60. Dike, L.E.; Chen, C.S.; Mrksich, M.; Tien, J.; Whitesides, G.M.; Ingber, D.E. Geometric control of switching between growth, apoptosis, and differentiation during angiogenesis using micropatterned substrates. *In vitro Cell. Dev. Biol. Anim.* **1999**, *35*, 441–448.
61. Tan, W.; Desai, T.A. Microfluidic patterning of cellular biopolymer matrices for biomimetic three-dimensional structures. *Biom. Microdev.* **2003**, *5*, 235–244.
62. Heiduschka, P.; Thanos, S. Implantable bioelectric interfaces for lost nerve functions. *Prog. Neurobiol.* **1998**, *55*, 433–461.
63. Branch, D.W.; Wheeler, B.C.; Brewer, G.J.; Leckband, D.E. Long-term maintenance of patterns of hippocampal pyramidal cells on substrates of polyethylene glycol and microstamped polylysine. *IEEE Trans. Biom. Eng.* **2000**, *47*, 290–300.
64. Rajnicek, A.; Britland, S.; McCaig, C. Contact guidance of CNS neurites on grooved quartz: Influence of groove dimensions, neuronal age and cell type. *J. Cell Sci.* **1997**, *110*, 2905–2913.
65. Lauer, L.; Ingebrandt, S.; Scholl, M.; Offenhausser, A. Aligned microcontact printing of biomolecules on microelectronic device surfaces. *IEEE Trans. Biomed. Eng.* **2001**, *48*, 838–842.
66. Pins, G.D.; Toner, M.; Morgan, J.R. Microfabrication of an analog of the basal lamina: Biocompatible membranes with complex topographies. *FASEB J.* **2000**, *14*, 593–602.
67. Saltzman, W.M.; Olbricht, W.L. Building drug delivery into tissue engineering. *Nat. Rev.* **2002**, *1*, 177–186.
68. Hartgerink, J.D.; Beniash, E.; Stupp, S.I. Self-assembly and mineralization of peptide–amphiphile nanofibers. *Science* **2001**, *294*, 1684–1687.

Tissue Engineering of Microvascular Networks

J. T. Borenstein

Charles Stark Draper Laboratory, Cambridge, Massachusetts, U.S.A.

E. J. Weinberg

Charles Stark Draper Laboratory, and Massachusetts Institute of Technology, Cambridge, Massachusetts, U.S.A.

M. R. Kaazempur-Mofrad

Massachusetts Institute of Technology, Cambridge, Massachusetts, and Massachusetts General Hospital, and Harvard Medical School, Boston, Massachusetts, U.S.A.

Joseph P. Vacanti

Massachusetts General Hospital, and Harvard Medical School, Boston, Massachusetts, U.S.A.

INTRODUCTION

Insufficient donor tissue and organs available for transplantation represents a crisis of immense proportions, and the problem is steadily worsening. Several potential solutions to this problem are being explored, including mechanical devices, xenotransplants, and tissue engineering. Tissue engineering promises to provide a fully natural solution to the organ shortage, and progress with tissues such as engineered skin, bone, and cartilage has been rapid. However, the field of vital organ tissue engineering presents additional challenges, including the requirement for complex arrangements of multiple cell types in a specific microarchitectural arrangement as well as the requirement for vascularization. Numerous approaches toward this latter goal are being pursued, including the introduction of suitable angiogenic growth factors as well as engineered methodologies that integrate computational fluid dynamics and microfabrication. This article addresses the application of microfabrication technology toward the engineering of vascular networks and vital organ assist and replacement devices. Fluid dynamic models for the vasculature are described, followed by a discussion of applicable microfabrication techniques, and finally examples of cell culture experiments in microfabricated devices are given.

THE REQUIREMENT FOR VASCULARIZATION

The most significant challenge facing vital organ transplant medicine is an ever-increasing shortage of donor organs.^[1] Alternative approaches involving mechanical devices^[2] and xenotransplants^[3] are progressing, as is the emerging field of tissue engineering.^[4,5] While vital

organ tissue engineering faces significant challenges related to cell biology^[6] and vascularization,^[7] engineered skin and cartilage are now in clinical use, and development of engineered blood vessels, bone, tubular structures, and corneas is rapidly advancing.^[8–10] For complex tissues such as vital organs, however, progress has been slowed because of challenging oxygen and nutrient transport requirements which severely limit the thickness of viable engineered tissues. Nature solves this transport problem by developing a vast circulatory network predominated by approximately 10-micron-diameter capillaries in a process known as angiogenesis. One approach for reproducing this microcirculation in engineered tissues and organs utilizes the introduction of angiogenic growth factors, most recently in a stepwise fashion in order to produce blood vessels with long-term viability and appropriate interconnectedness.^[11] A second approach relies on microfabrication technology to produce an engineered organ scaffold which is then populated with endothelial cells.^[12] This second approach—the generation of a microvascular network through the use of fluid dynamic models and microfabrication technology—is the focus of discussion in this article.

ENGINEERING VASCULARIZED TISSUES AND ORGANS

An engineered approach for the generation of vascularized tissues and organs consists of several well-defined steps. The process begins with the generation of a fluid dynamic model which approximates the behavior of blood flow in the target organ, and which accounts for blood's complex nonlinear blood rheology and two-phase composition.^[13] This model forms the basis of the

lithographic masks used to pattern the blood vessel network in the micromachining process. Master molds are then produced using high-aspect-ratio micromachining methods, typically either by silicon etching or by a thick photoresist process. Polymer films are cast in the master molds, and are integrated to form thick three-dimensional (3-D) scaffolds for cell seeding and expansion. Initial work has focused on the fabrication of scaffolds produced using either nondegradable biocompatible templates or biodegradable templates based on copolymers of polylactic and polyglycolic acid or other resorbable polymer systems. Microfluidic networks have been designed, built, and tested, and have produced highly uniform flow patterns that mimic large-scale physiologic properties such as total flow rate and small-scale phenomena such as fluid velocity and hematocrit (blood's fraction of red blood cells) in the capillaries.^[14] Early cell seeding results demonstrate successful attachment and proliferation of endothelial cells in microfluidic channels as small as 30 microns in diameter.^[15,16] The ultimate goal is to scale up the technology to a full 3-D system with the blood volume and cell density required to support a complete organ.

FLUID DYNAMICS OF THE VASCULATURE

Creation of an engineered vasculature begins with the fabrication of a microfluidic network that mimics physiological vasculature. Design of an appropriate network requires understanding of the geometry and flow characteristics of physiological vasculatures and the ability to create a network having the desired properties.

Information on the basic flow behavior in physiological vasculature, such as the total flow rate through an organ is widely available.^[17] Researchers have also used a variety of techniques to determine more specific flow details, such as the flow and pressure distribution in an individual organ and the specific vessel and branching geometries throughout an organ.^[18] Generally an engineered vasculature must match the geometric and flow dynamic distributions of physiological vasculature, including the Murray's hypothesis that the wall shear remain constant throughout the network.^[19]

Modeling Blood Flow Through a Single Vessel

Before modeling a network of blood vessels, we must understand the flow in a single vessel. Blood flow in a vessel is viscous and laminar, so the flow is governed by the viscosity of blood and the geometry of the vessel.

Blood is highly non-Newtonian, meaning its viscosity varies. This is due to the presence of solids, mostly red blood cells. The viscosity of blood varies with the concentration of these cells. The concentration of red

blood cells is known as the hematocrit, and this will vary depending on the size of the blood vessel and the pattern of vessel intersections that the blood flows through. These effects have been examined by analyzing blood flow under well-defined laboratory conditions to determine empirical relations for an effective viscosity for blood in various situations. For blood flowing through vasculature, Casson's equation is commonly used:

$$\mu = \left(k_1 + \frac{k_2}{\sqrt{U/d}} \right)^2 \quad (1)$$

which gives the viscosity μ in terms of the bulk flow velocity U , the vessel diameter d , and two empirical constants k_1 and k_2 .

Once the viscosity is known, the vessel can be represented by a fluid resistance. For a cylindrical vessel, the resistance is

$$R = \frac{8\mu L}{\pi r^4} \quad (2)$$

where L is the length of the vessel and r is the radius of the vessel. Microfabricated vessels are often rectangular, and a similar resistance equation is available for rectangular vessels. Resistance of vessels having other cross-sections can be found by numerical integration of the Poisson equation over the cross section.

The vessel walls may be distensible, in which case the vessel shape will change as the fluid pressure inside varies. Analytical solutions for the geometry shape are available for cylindrical vessels and can be found for other shapes by numerical solution of the plane strain equations.

Modeling Flow Through a Network of Blood Vessels

Flow through a network is modeled by representing each individual vessel with a resistance and solving the entire network, analogous to the modeling of a circuit of electrical resistors. Once the resistance value of a vessel is known, the pressure drop across the vessel is related to the flow rate through the vessel by:

$$P_1 - P_2 = QR \quad (3)$$

where P_1 and P_2 are the pressures at each end of the vessel, R is the resistance of the vessel, and Q is the flow rate of the vessel. Nodal formulation is used to combine all of the individual resistance equations into a matrix equation describing all of the flows in the network. This matrix equation can be solved by Sparse Gaussian Elimination or the Conjugate Residual Method to

determine all of the pressures, which can be substituted back in to find all of the flow rates. All other flow properties—namely, the flow velocity and wall shear—can be found from the flow rates. Thus this solution allows the determination of all flow properties in a network where the geometry is known. If distensibility effects are included, the network must be solved repeatedly as the geometry changes,^[20] and the matrix must be expanded and solved repeatedly to include non-Newtonian effects.^[13]

For designing a network, however, the desired flow properties are generally known and the geometry is unknown. Network design can be achieved two ways. One way is to rearrange the matrix equation to solve for unknown resistances in terms of known pressures and flowrates, then determine the geometry from the resistances. The other approach is to guess an initial geometry and solve for flow behavior, then modify the geometry and solve again, iterating until the desired flow properties are achieved. A sample vascular network design generated using the models described above is shown in Fig. 1.

MICROFABRICATION TECHNOLOGY FOR SCAFFOLD FABRICATION

Micromachining technology—the core fabrication process used to build microelectrical mechanical system (MEMS) structures and devices—has a resolution limit in the range of 0.1 μm , two orders of magnitude smaller than the dimension of the capillaries. Many MEMS devices have structural dimensions as large as one centimeter, yet still maintain critical dimensions and tolerances to ± 0.1 micron.^[21,22] This span of five orders of magnitude demonstrates the applicability of MEMS technology to the formation of vascularized

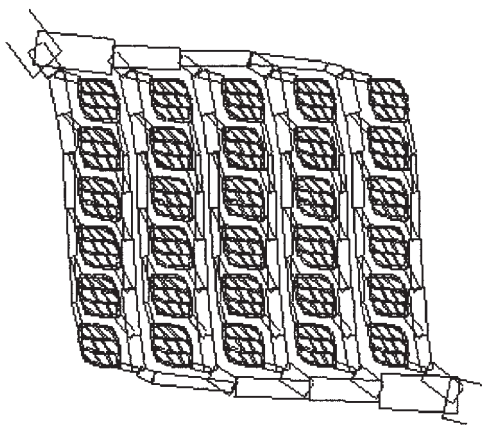


Fig. 1 Schematic of sample microvascular network based on computational fluid dynamic models designed to replicate organ vascular physiology.

tissues with organized size scales ranging from overall dimensions of tens of centimeters down to cellular dimensions of ~ 5 microns. Photolithographic patterns may be transferred into vascular molds through etching processes, a set of highly developed techniques for producing trenches with a wide range of depths, cross-sections, and geometries in semiconductors, glasses, and polymeric materials. Cellular interactions with vascular networks may be controlled by varying the geometries involved;^[23,24] this behavior has been demonstrated and exploited as part of an important phenomenon known as contact guidance.^[25–27] Micromachining is ideally suited to explore cell-surface interactions as a function of material properties, surface texture, and trench geometry.

Early efforts to produce microvascular networks in vitro utilized silicon micromachined wafers as a platform to demonstrate the basic principle.^[7] These approaches involved the application of photolithographic patterns to silicon wafer surfaces, or the transfer of lithographically patterned features into etched channels in the wafers, to generate a template for the growth of endothelial cells. Results obtained for culturing of endothelial cells in circular channels showed that virtually confluent coatings could be readily achieved in a matter of hours, using a variety of surface coatings such as collagen to promote adhesion and to provide a supporting matrix. Whereas these early studies provided useful preliminary data demonstrating the fundamental utility of microfabrication techniques for producing simulated microvascular networks, the application of polymer replica molding technology represented the most significant advance toward the construction of 3-D vascularized tissue engineered constructs. In this case, silicon micromachined wafers are used as master molds for the production of multiple polymeric films in a layering process designed to produce a large-scale vascularized network for engineering complex tissues and organs. Micromachining process technology is described in numerous references,^[28–30] here the basics of the tools necessary to generate vascularized scaffold structures are reviewed. A schematic of the process sequence is shown in Fig. 2.

Lithographic Techniques

Photolithography technology is the standard method for patterning structures and devices in applications ranging from standard microelectronics to MEMS sensors and actuators and, more recently, bioanalytical devices and systems. For most applications in bioanalysis and tissue engineering, standard contact lithographic processes, in which photomasks are brought into contact with substrates to be patterned, provide sufficient dimensional tolerances. Fundamentally, the

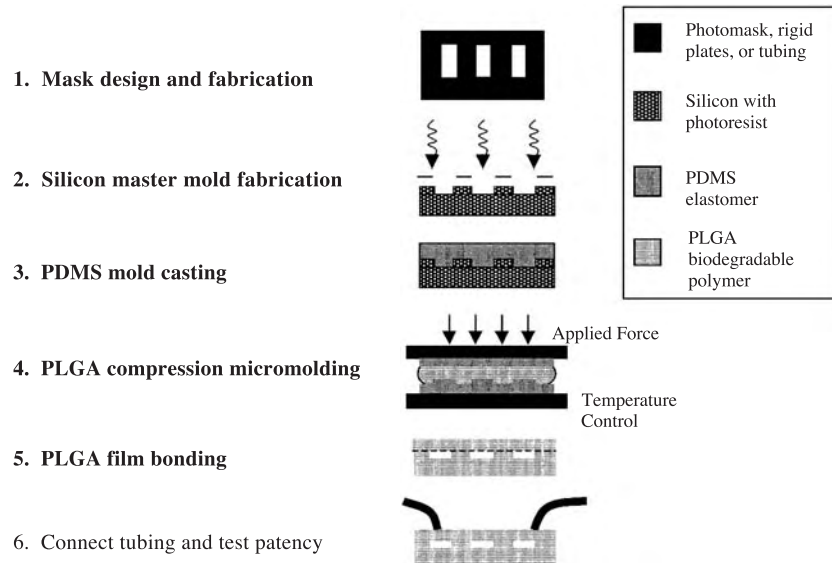


Fig. 2 Process schematic for the microfabrication of polymeric films for vascularized scaffolds. (Courtesy Ref. [15].)

lithographic process proceeds by the generation of a photomask containing a pattern such as the design discussed in the earlier section on fluid dynamics, and the repeated application of this photomask pattern to substrates for device fabrication.

Photomask fabrication techniques vary depending on the requirements of the application, but two general approaches are most often used for biological applications: 1) use of a chrome-metallized glass plate that has been written with an electron beam, typically with resolution between 0.1 and 0.5 microns; and 2) overhead transparency-style masks, written with a high-definition laser printer, typically with resolution no better than 10–20 microns. Although the former technique provides higher resolution, mask turnaround times are days rather than hours, and mask costs are $10\text{--}50 \times$ higher. In either case, masks are generated and then used to produce patterned silicon wafers either for direct cell culture or for the subsequent generation of polymer molded structures for cell culture.

Photolithographic processes involve the use of either positive or negative photoresists which, when applied in thin controlled layers to wafer surfaces by spin-coating, are sensitive to exposure to specified light sources to which they are exposed. Photomasks in contact with wafer surfaces provide light-exposed and light-blocked regions that alternately expose and protect features on the wafers. Once developed and hard-baked, these photoresists serve as protective layers enabling patterns to be transferred into the wafer through subsequent etching processes. Etching processes take one of two forms, either wet chemical etching or dry plasma etching. The former approach is not as useful for tissue engineering applications because of problems with undercutting (dimensional changes during etching) and with etch profiles. In the next section plasma etching techniques will be described.

Plasma Etching

Two types of plasma etching processes are available for generating silicon master molds. The first of these, isotropic reactive ion etching, was developed for applications in the semiconductor industry and is capable of producing trenches with curved sidewalls and depths of up to 100 microns using conventional photoresist technology. Typically, the etch gases used are sulfur hexafluoride or carbon tetrafluoride, and etch rates are usually in the range of 1 micron per minute. The principal advantage of isotropic plasma etching lies in the curved geometry obtained, which produces microchannels in the molded polymer films that are more amenable to continuous lining with endothelium. A significant disadvantage of this approach is the roughly one-to-one increase in feature width with etched depth, a result of the isotropic nature of the process, which causes features to expand beyond their intended lateral dimension during trench formation.

A second class of plasma etching techniques is typically referred to as Deep Reactive Ion Etching (DRIE), a process capability developed in only the last several years for applications in the field of MEMS. Although several variants of this technology exist, the primary embodiment is known as the Bosch process, first developed by Robert Bosch, GmbH, with process tools commercialized by Surface Technology Systems and other companies.^[31,32] This technique produces etched sidewalls that are exactly vertical, and retains the feature width on the photomask very tightly, even as the etched depth increases. Aspect ratios of 25:1 (depth: width) can be obtained using this process, and depths of several hundred microns can be readily achieved. The process consists of repeated cycling between an etching step and a so-called passivation step in which a fluoropolymer is deposited relatively uniformly over the

surface. During the subsequent etching step, the passivating fluoropolymer is removed from the bottom of the trench but persists long enough to protect the sidewalls from being attacked, thereby producing the vertical etch geometry and nearly perfect anisotropy. One important feature to note regarding DRIE etch profiles is the presence of scalloped sidewalls resulting from the cyclic etch-passivation process.

Polymer Replica Molding

Once silicon master molds have been generated using one of the lithographic or etching processes described earlier, polymer films patterned with the reverse image of the mold may be repeatedly generated from the master mold.^[33–35] The only treatment required for replication may be the application of a release agent to the surface of the master mold, depending upon the technique used to generate the master. Typically, lithographically patterned masters do not require release agent application, but etched molds with microscale surface roughness often do require it.

Several techniques are available for polymer film replication, including hot embossing, solvent casting, spin casting, and compression molding, among others.^[36] These are described in detail in the literature; for purposes of brevity, solvent casting of the classic biopolymer poly(dimethylsiloxane) (PDMS) will be presented, and for an example of biodegradable polymer microfabrication, compression molding of poly(lactic) co-glycolic acid (PLGA) will be discussed.

Biocompatible Polymer Replica Molding

The biocompatible polymer PDMS has become a standard for many microfluidic applications. A standard technique for the production of PDMS films using replica molding involves the use of solvent casting, a simple technique capable of producing films as thin as 5 microns with features less than a micron in lateral dimension. The material is flexible, strong, and transparent; is readily amenable to a variety of sterilization methods; and, if cured appropriately, will retain feature dimensions from the master mold to a high degree of accuracy. Typical PDMS molding begins with the mixture of the resin and the curing agent, followed by thorough mixing and vacuum degassing. Once the prepolymer has been prepared, it is poured over the silicon micromachined mold wafer, allowed to settle, and baked at elevated temperature. At this stage the PDMS film is ready to peel directly from the master mold wafer; the thickness of the polymer film is controlled by the volume of the prepolymer solution applied to the wafer. For microvascular network fabrication, PDMS may be used directly as a structural material, or may be

used as a transfer mold for subsequent generation of biodegradable polymer layers, as described in the next section.

Biodegradable Polymer Microfabrication

Ultimately, the utility of microvascular networks lies in their ability to support organ-specific cells and to perform functions as an assist or replacement device. Therefore the preferred platform for a microvascular network is a biodegradable polymer construct, which will provide a scaffolding for cell placement and a mechanical support structure until the growing tissue achieves sufficient strength and viability to sustain itself without the scaffold. Several polymers have been demonstrated as biodegradable scaffolds,^[37] including Poly(Glycolic Acid) (PGA) and Poly(Lactic Acid) (PLA), the copolymer Poly(Lactic-co-Glycolic Acid) (PLGA), polycaprolactone, Poly(Vinyl Alcohol) (PVA), and biorubber. The standard material for biodegradable scaffold fabrication has been PLGA, due to U.S. FDA acceptance, low cost, and availability, and positive results in many early applications. For replica molding, PLGA films must be produced using an intermediate transfer mold, because PLGA adheres strongly to silicon surfaces and because it is not sufficiently flexible to be peeled away from a rigid mold. Solvent casting and spin casting approaches have been utilized with some success; however, compression molding of thin films has resulted in the highest degree of fidelity of feature transfer.^[38]

Compression molding begins with the production of a PDMS transfer mold film, as described earlier. Solid polymer, usually in the form of small beads, is manually deposited on the PDMS transfer layer, which itself is placed on a metal chuck for mechanical support. A second unpatterned PDMS layer is attached to the underside of a second metal chuck and is brought down into contact with the first film through the layer of PLGA beads. Care must be taken to ensure that the two PDMS layers are coplanar in order to achieve a flat PLGA film. The assembly is then heated to a temperature well above the glass transition temperature of the PLGA, sufficient to enable viscous flow of the material so that pattern replication of fine features can take place. Once this viscous flow has occurred, controlled levels of high pressure are applied to the stack, resulting in a thin flat PLGA film with pattern transfer. The assembly is then cooled and the PDMS master is peeled away from the PLGA film. Figure 3 shows an image of a PLGA microvascular network.

Assembly of Microvascular Networks

The foregoing discussion has described some of the necessary components of the process of constructing

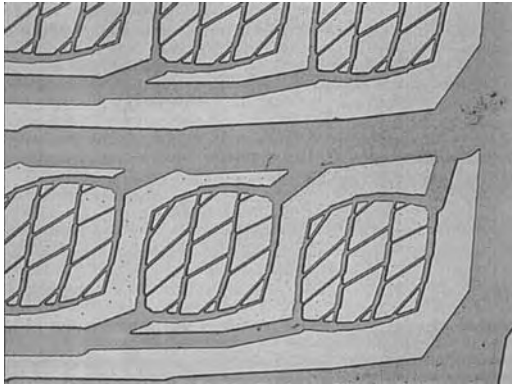


Fig. 3 PLGA vascular network layer produced by compression molding from a replica-molded PDMS transfer mold.

engineered microvascular networks; what remains is 3-D assembly into single level or multilayer vascular beds, and seeding of these beds with endothelium to form a functional microengineered vasculature. In this section a simple assembly process is described; more complex approaches resulting in constructs with improved physiological properties have been investigated, but are beyond the scope of this discussion.

A single-layer microvascular network containing a two-dimensional (2-D) array of interconnected channels may be constructed by simply bonding a polymer film patterned with trenches to a flat film and sealing the two films together. For the nondegradable polymer PDMS, this is a very straightforward exercise, as bonding of such layers may be accomplished simply by exposing the two adjacent faces to an oxygen plasma and joining them to form a permanent bond. Inlets and outlets to these 2-D networks are made using direct syringe interfaces, bonding of molded tubing sections, or other means. Such 2-D constructs can then be seeded with endothelium to form the simplest possible microvascular networks.

A 3-D microvascular network can be constructed by stacking multiple layers of vessel networks with fluidic connections for vertical integration. In the simplest embodiment, polymer films containing 2-D channel networks can be stacked with a single vertical connecting tube for the inlet and one corresponding outlet tube.^[16] This configuration is illustrated in Fig. 4; the vertical inlet and outlet tubes are arranged diagonally across a roughly square vessel network array. Although this type of integration scheme greatly simplifies the fabrication process, fluid dynamic models indicate that the resulting 3-D network differs in substantive ways from physiological networks. Principal differences include the density of capillaries relative to larger vessels, a key determinant of oxygen transport properties as most of the oxygen exchange occurs at the capillaries, and the distribution of shear stress in the network, a

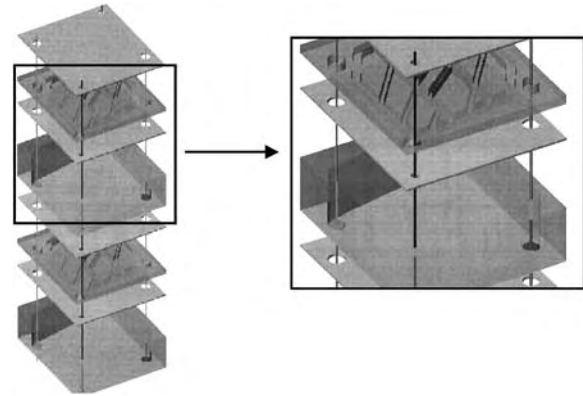


Fig. 4 Schematic of integration process for three-dimensional scaffold for organ tissue engineering.

major factor in endothelial cell viability as the main mechanical signal to initiate the biochemical cascade of events through the process known as mechanotransduction.^[39] Therefore, physiological 3-D models require the use of more complex integration techniques in which vertical interconnections are more numerous and occur over a much wider range of vessel diameter scales. One such 3-D integration scheme consists of vertical interconnects arranged in a dense array across each horizontal channel network. Such integration schemes are more challenging from a fabrication standpoint, due to the requirement for high alignment accuracy. However, double-sided alignment tools capable of submicron placement are applicable to construction of these more complex 3-D microfluidic devices.

Porous Microfabricated Devices

The foregoing discussion pertains to the design and construction of microvascular networks for tissue engineering, in which a fluid model is applied to a tissue or organ and a polymeric scaffold is fabricated in accordance with the model. Such structures can be seeded with various cell types in order to produce engineered blood vessel networks, but functional tissues and organs require the integration of organ-specific cell types into the microvascular system. Compartments machined into the engineered scaffolds can be used to house these organ-specific cells; the primary challenges are to arrange these cells so that they retain their function during *in vitro* culturing and following implantation, and to provide for appropriate oxygen and nutrient transport from the engineered microcirculation to the organ-specific cells. Substantial progress in reproducing the microenvironment necessary to maintain hepatocyte function has been reported,^[6,40,41] and recent reports highlight the importance of the 3-D aspect of the local environment.^[42,43] Oxygen transport can be provided by

utilizing polymeric materials with high oxygen diffusion coefficients,^[44] but fully functional cocultured systems will likely require the use of nanoporous structures through which large molecules can pass. With degradable scaffolds, once the polymer has been resorbed, cell-cell junctions in the endothelium will govern the transport of oxygen, nutrients, and waste products.

Porous polymer scaffolds have served as the foundation for an entire class of tissue engineering techniques, in which cell-seeded scaffolds are cultured and implanted for tissue replacement. Methods for providing porous films from PLGA, for instance, have recently been developed.^[45] Microporous and nanoporous films can be processed using the microfabrication techniques described earlier, with some modifications to the fabrication process. For instance, thermal fusion bonding may disturb the nanopore geometries unless additional steps are taken to control the thermal excursions experienced by the device during processing. Ultimately, porous structures will play an extremely important role in engineered scaffolds for vascularized tissue engineering.

ENDOTHELIAL CELL SEEDING OF MICROFABRICATED NETWORKS

Once polymer scaffolds designed to replicate organ microvasculature have been constructed, the next step is to seed them with endothelial cells to generate artificial blood vessel networks. Early efforts involving microfabrication for tissue engineering focused on simple static seeding, in which patterned networks were seeded with endothelial cells with no flow. In more recent work, the use of dynamic cell seeding has been shown to produce confluent cell coverage; this process introduces cells through an inlet port of the scaffold; they are given the opportunity to attach to the walls of the bifurcating channel networks as they proceed towards the outlet end.

Prior to cell culturing, devices are sterilized in an autoclave and perfused with a solution to aid cell adhesion, such as poly-L-lysine, collagen, gelatin, or fibronectin.^[15] The network is then seeded with a cell suspension of approximately 1×10^7 cells/mL using a positive displacement or peristaltic pump at flow rates of roughly $100 \mu\text{L}/\text{min}$. The bioreactor, consisting of a pump, oxygenator, bubble trap, and a reservoir of sterile culture medium, is shown in Fig. 5. The system is then run closed-loop in an incubator at standard culture settings for periods of several weeks. A micrograph shown in Fig. 6 illustrates confluent endothelial cell attachment after two weeks in culture. In the initial stages, single cells and aggregates begin to attach. After several days, channels approach near-confluence, and finally, after four weeks, surfaces in most regions of

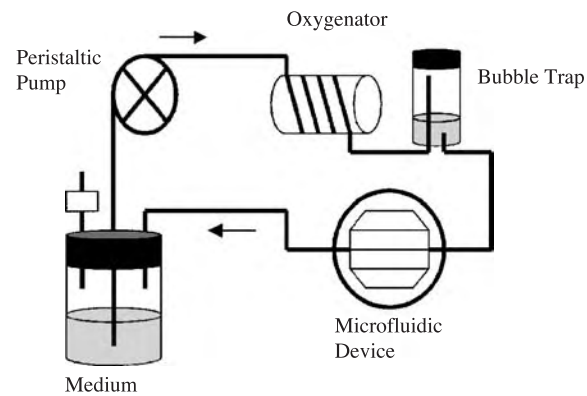


Fig. 5 Bioreactor for endothelial cell culturing in microvascular scaffolds. (Courtesy Ref. [12].)

small model networks are fully covered with a single layer of endothelial cells. One salient observation regarding the culture experiments is the fact that autoclave sterilization generates no measurable pattern distortion. Additionally, channels coated with cell adhesion molecules exhibit enhanced cell attachment when compared to phosphate-buffered saline coated control channels. Experiments have successfully concluded after four weeks without the occurrence of occlusion or contamination.

APPROACHES FOR WHOLE ORGAN ENGINEERING

This description represents the first step in the development of a new microfabrication technology for vascularized tissue engineering. Silicon micromachining technology is well-suited toward the generation of templates for endothelial cell growth due to its high precision, design flexibility, and the ability to form specific geometric shapes by varying process recipes. Replica molding of biocompatible polymers on these

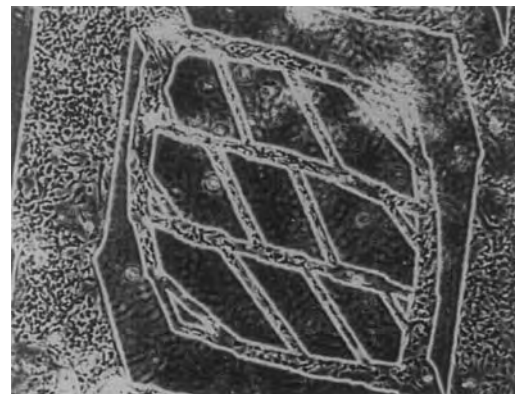


Fig. 6 Endothelial cells lining microvascular channels after two weeks of culturing.

high-definition silicon micromachined substrates enable the production of sheets of patterned materials suitable for cell seeding, attachment, and growth. Recent work has resulted in the extension of this replica molding technology to biodegradable materials.

Cell culturing has been shown to produce scaffolds capable of being seeded with endothelial cells to near confluence. Although cells have been introduced into microchannels for sorting or analysis purposes, they are not typically cultured long-term in microfluidic devices. We have successfully seeded and cultured endothelial cells in microfabricated channels for periods as long as four weeks, demonstrating that cells successfully attach, proliferate, and migrate in closed channels with small geometries. This is a critical milestone toward the ultimate goal of fabricating microvasculature and enabling the development of bulk organ tissue engineering. The two-phase composition (red blood cells and plasma) of blood and its nonlinear rheology present several complications at the microcirculation level that must be considered in order to build a full model of the fluid dynamics of the vasculature.^[13] Several phenomena must be considered, including interactions between red blood cells and the walls, leading to axial migration, dynamic reduction of hematocrit (the Fahraeus effect), correlations between red blood cells and velocity, reduction of red blood cell transit time, reduction of hematocrit contained in vessels relative to discharge hematocrit, and reduction of flow resistance below the level expected based on bulk flow (Fahraeus–Lindquist effect). These phenomena are all incorporated into our computational model and will be fully utilized in our design of the model systems in order to build channel networks that replicate physiological behavior more accurately.

Microfluidic devices such as those described in this work may be used to maintain and support large numbers of cells due to the favorable surface-to-volume ratio and short diffusion distances. As a longer-range goal, such devices may be used to engineer blood vessels; however, many challenges remain. Multiple cell types will be required, and the roles of extracellular matrix and cell–cell junctions will need to be determined. Ultimately, the physiology of the vasculature must be reproduced at both the macroscale and the microscale.

Full realization of tissue-engineered organs grown in the lab will require the use of microfabrication techniques that produce complex 3-D templates for cell adhesion and growth. Techniques for integrating multiple layers of microfluidic scaffolds will be required, as will methods for generating interpenetrating arrays of channels for multiple cell types. Efforts are now focused on developing 3-D fabrication techniques based on replica molding to produce scaffolds for vascular networks integrated with parenchymal cells to produce complex organs such as the kidney and liver.

CONCLUSIONS

Progress in the field of tissue engineering has so far been limited by the size and scale of tissues that contain cells with high nutrient and oxygen requirements. Advances in microfabrication technology represent an important opportunity for the field, because these fabrication techniques are capable of producing structures at size scales commensurate with the requirements of physiological systems in a range of biocompatible and biodegradable materials. These fabrication techniques are scalable and cost-effective, and will ultimately play a role in the development of tissue-engineered devices for therapeutic applications in the field of transplant medicine.

ACKNOWLEDGMENTS

The authors would like to thank the members of the MEMS Technology Group at Draper Laboratory—in particular, Edward Barnard, Connie Cardoso, Isaac Costa, Brian Orrick, and Mert Prince—for their efforts in fabricating micromachined silicon devices and polymer replica molds. In addition, we would like to thank Dr. Wing Cheung and Dr. Michael Shin of the Massachusetts General Hospital and their colleagues in the Organ Fabrication Group, Kimberly Morgan of Boston University, and Kevin King of MIT. This work was partially sponsored by the Department of the Army, Cooperative Agreement DAMD-17-02-2-0006. The content of this article does not necessarily reflect the position or the policy of the government, and no official endorsement should be inferred.

ARTICLES OF FURTHER INTEREST

Angiogenesis Inhibitors; Biorubber/Poly(Glycerol Sebacate); Microelectromechanical Systems (MEMS) Manufacturing; Poly(Glycolic Acid); Poly(lactic acid)s; Tissue Engineering, Microscale; Vascularization of Engineered Constructs

REFERENCES

1. Blood Weekly. In *U.S. Waiting List Tops 75,000*; NewsRx.net: Atlanta, GA, May 10, 2001.
2. Dowling, R.D.; Gray, L.A., Jr; Etoch, S.W.; Laks, H.; Marelli, D.; Samuels, L.; Entwistle, J.; Couper, G.; Vlahakes, G.J.; Frazier, O.H. The AbioCor implantable replacement heart. *Ann. Thorac. Surg.* **Jun 2003**, 75 (6 Suppl), S93–S99.
3. Seebach, J.D.; Schneider, M.K.; Comrack, C.A.; LeGuern, A.; Kolb, S.A.; Knolle, P.A.; Germana, S.;

- DerSimonian, H.; LeGuern, C.; Sachs, D.H. Immortalized bone-marrow derived pig endothelial cells. *Xenotransplantation* **2001**, *8*, 48.
4. Langer, R.; Vacanti, J.P. Tissue engineering. *Science* **1993**, *260*, 920.
 5. Vacanti, J.P.; Langer, R. Tissue engineering: The design and fabrication of living replacement devices for surgical reconstruction and transplantation. *Lancet* **1999**, *354* (Suppl I), 32SI–34SI.
 6. Bhatia, S.N.; Yarmush, M.L.; Toner, M. Controlling cell interactions by micropatterning in co-cultures: Hepatocytes and 3T3 fibroblasts. *J. Biomed. Mater. Res.* **1997**, *34*, 189.
 7. Kaihara, S.; Borenstein, J.T.; Koka, R.; Lalan, S.; Ochoa, E.R.; Ravens, M.; Pien, H.; Cunningham, B.; Vacanti, J.P. Silicon micromachining to tissue engineer branched vascular channels for liver fabrication. *Tissue Eng.* **2000**, *6*, 105.
 8. Bell, E.; Ehrlich, P.; Buttle, D.J.; Nakatsuji, T. Living tissue formed in vitro and accepted as skin-equivalent of full thickness. *Science* **1981**, *221*, 1052–1054.
 9. Parenteau, N.; Sabolinski, N.M.P.; Nolte, S.; Oleson, C.; Kriwet, M. Biological and physical factors influencing the successful engraftment of a cultured human skin substitute. *Biotechnol. Bioeng.* **1996**, *52*, 3–14.
 10. Niklason, L.E.; Gao, J.; Abbott, W.M. Functional arteries grown in vivo. *Science* **1999**, *284*, 489–493.
 11. Richardson, T.P.; Peters, M.C.; Ennett, A.B.; Mooney, D.J. Polymeric system for dual growth factor delivery. *Nat. Biotechnol.* **2001**, *19*, 1029.
 12. Borenstein, J.T.; Terai, H.; King, K.R.; Weinberg, E.J.; Kaazempur-Mofrad, M.R.; Vacanti, J.P. Microfabrication technology for vascularized tissue engineering. *Biomed. Microdevices* **2002**, *4* (3), 167–175.
 13. Kaazempur-Mofrad, M.R.; Vacanti, J.P.; Kamm, R.D. Computational Modeling of Blood Flow and Rheology in Fractal Microvascular Networks. In *Computational Fluid and Solid Mechanics*; Bathe, K.J., Ed.; Elsevier Science, Ltd.: Oxford, 2001; 864–867.
 14. Kaazempur-Mofrad, M.R.; et al. . (to be published).
 15. King, K.R.; Wang, C.C.; Shin, M.; Vacanti, J.P.; Borenstein, J.T. *Biodegradable Polymer Microfluidics for Tissue Engineering Microvasculature*, MRS Symp. Proc., MRS Press, 2002; Vol. 729, U1.3.
 16. Borenstein, J.T.; Cheung, W.; Hartman, L.; Kaazempur-Mofrad, M.R.; King, K.R.; Sevy, A.; Shin, M.; Weinberg, E.J. *Living Three-Dimensional Microfabricated Constructs For The Replacement Of Vital Organ Function*; IEEE MEMS: Boston, MA, 2003; 1754.
 17. Guyton, A.C.; Hall, J.E. *Textbook of Medical Physiology*; W.B. Saunders: Philadelphia, PA, 2000.
 18. Kassab, G.S.; Lin, D.H.; Fung, Y.C. Morphometry of pig coronary venous system. *Am. J. Physiol.* **1994**, *267*, H2100.
 19. Murray, C.D. The physiological principle of minimum work, I: The vascular system and the cost of blood volume. *Proc. Natl. Acad. Sci.* **1926**, *12*, 207–214.
 20. Weinberg, E.J.; Kaazempur-Mofrad, M.R.; Borenstein, J.T. Numerical Model of Flow in Distensible Microfluidic Network. In *Computational Fluid and Solid Mechanics*; Bathe, K.J., Ed.; Elsevier Science, Ltd.: Oxford, 2003; Vol. 2, 1569–1572.
 21. Gabriel, K.J. Microelectromechanical systems. *Proc. IEEE* **1998**, *86*, 1534.
 22. Borenstein, J.T.; Gerrish, N.D.; Currie, M.T.; Fitzgerald, E.A. A New Ultra-Hard Etch—Stop Layer for High Precision Micromachining. In *Technical Digest. IEEE International MEMS 99 Conference. Twelfth IEEE International Conference on Micro Electro Mechanical Systems*; IEEE: Piscataway, NJ, 1999; 205.
 23. Damji, A.; Weston, L.; Brunette, D.M. Directed confrontations between fibroblasts and epithelial cells on micromachined grooved substrata. *Exp. Cell. Res.* **1996**, *228*, 114.
 24. Den Braber, E.T.; Ruijter, J.E.D.; Ginsel, L.A.; Recum, A.F.V.; Jansen, J.A. Orientation of ECM protein deposition, fibroblast cytoskeleton, and attachment complex components on silicon microgrooved surfaces. *J. Biomed. Mater. Res.* **1998**, *40*, 291.
 25. Chen, C.S.; Mrksich, M.; Huang, S.; Whitesides, G.M.; Ingber, D.E. Geometric control of cell life and death. *Science* **1997**, *276*, 1425.
 26. Curtis, A.; Wilkinson, C. Topographical control of cells. *Biomaterials* **1997**, *18*, 1573.
 27. Flemming, R.G.; Murphy, C.J.; Abrams, G.A.; Goodman, S.L.; Nealey, P.F. Effects of synthetic micro- and nano-structured surfaces on cell behavior. *Biomaterials* **1999**, *20*, 571.
 28. McWhorter, P.J.; Frazier, A.B.; Rai-Choudhury, P. Micromachining and Trends for the Twenty-First Century. In *Handbook of Microlithography, Micromachining and Microfabrication*; Rai-Choudhury, P., Ed.; SPIE Press: Bellingham, WA, 1997; Vol. 3.
 29. Kendall, D.L.; Malloy, K.J.; Fleddermann, C.B. Critical Technologies for the Micromachining of Silicon. In *Semiconductors and Metals*; Faber, K.T., Malloy, K.J., Eds.; Academic Press: New York, 1992; Vol. 37, 193.
 30. Madou, M. *Fundamentals of Microfabrication*; CRC Press: Boca Raton, FL, 2002.
 31. Ayon, A.A.; Chen, K.-S.; Lohner, K.A.; Spearing, S.M.; Sawin, H.H.; Schmidt, M.A. Deep reactive ion etching of silicon. *Mat. Res. Soc. Symp. Proc.* **1999**, *546*, 51.
 32. Hynes, A.M.; Ashraf, H.; Bhardwaj, J.K.; Hopkins, J.; Johnston, I.; Sheperd, J.N. Recent advances in silicon etching for MEMS using the ASETM process. *Sens. Actuators, A* **1999**, *74*, 13.
 33. Jo, B.-H.; Beebe, D.J. Fabrication of three-dimensional microfluidic systems by stacking molded PDMS layers. *SPIE* **1999**, *3877*, 222.
 34. Quake, S.R.; Scherer, A. From micro- to nanofabrication with soft materials. *Science* **2000**, *290*, 1536.
 35. Whitesides, G.M.; Stroock, A.D. Flexible methods for microfluidics. *Phys. Today* **2001**, *54*, 42.
 36. Thomson, R.C.; Yaszemski, M.J.; Mikos, A.G. Polymer Scaffold Processing. In *Principles of Tissue Engineering*; Lanza, R., Langer, R., Chick, W., Eds.; Landes & Co., 1997.
 37. Wang, Y.; Guillermo, A.; Ameer, A.; Sheppard, J.; Langer, R. A tough biodegradable elastomer. *Nat. Biotechnol.* **2002**, *20*, 602–606.

38. King, K.R.; Wang, C.; Vacanti, J.P.; Borenstein, J.T. Submitted for publication.
39. Davies, P.F. Flow-mediated endothelial mechanotransduction. *Physiol. Rev.* **1995**, *75*, 519–560.
40. Bhatia, S.N.; Chen, C.S. Tissue engineering at the micro-scale. *Biomed. Microdevices* **1999**, *2*, 131–144.
41. Griffith, L.G.; Naughton, G. *Science* **2002**, *295*, 1009–1014.
42. Poznansky, M.C.; Evans, R.H.; Foxall, R.B.; Olszak, I.T.; Plascik, A.H.; Hartman, K.E.; Brander, C.; Meyer, T.H.; Pykett, M.J.; Chabner, K.T.; Kalams, S.A.; Rosenzweig, M.; Scadden, D.T. Efficient generation of human T cells from a tissue-engineered thymic organoid. *Nat. Biotechnol.* **2000**, *18*, 729.
43. Webb, D.J.; Horwitz, A.F. New dimensions in cell migration. *Nat. Cell Biol.* **2003**, *5*, 690.
44. LeClerc, E.; Sakai, Y.; Fujii, T. “Cell culture in 3-dimensional microfluidic structure of PDMS.” *Biomed. Microdevices.* **2003**, *5*, 109–114.
45. *Biomaterials for Tissue Engineering*; Wong, J.Y., Plant, A.L., Schmidt, C.E., Shea, L., Coury, A.J., Chen, C.S., Eds.; MRS Press: Warrendale PA, 2004; Vol. EXS-1, F3.5.

Tissue Engineering of Pancreas

Hiroo Iwata

Institute for Frontier Medical Sciences, Kyoto University, Kyoto, Sakyo-Ku, Japan

INTRODUCTION

Type I diabetes mellitus, which is caused by autoimmune destruction of pancreatic β -cells, is a devastating condition that affects millions of peoples worldwide. The use of conventional insulin therapy for control of glucose levels cannot inhibit development of serious chronic complications such as nephropathy, retinopathy, and cardiovascular disease in diabetic patients. Tight glycemic control through intensive insulin therapy can deter the chronic complications of diabetes. However, it places an enormous burden on patients, and intensive insulin therapy is associated with a high risk of severe hypoglycemia, which can also be life-threatening. Alternative treatments for type I diabetes are transplantation of a whole pancreas or islets of Langerhans (islets), which is made of endocrine tissue and secretes insulin. Although whole pancreas transplantation offers the possibility of excellent glycemic control in a physiologic manner, it also subjects patients to the risks of major surgery and complications related with pancreatic juice produced by acinar tissue. Islet transplantation has been expected to be able to offer the possibility of internal glycemic control without subjection of the patient to the risks of major surgery. However, low yield of islet isolation and deterioration of islet functions caused by conventional immunosuppressive regimens limited clinical success of islet transplantation.

A major breakthrough in clinical islet transplantation was made in recent years. Islet transplantation led to insulin independence and normal glucose metabolism in patients with type 1 diabetes mellitus, and these patients remained independent of insulin within a follow-up period up to 20 months after islet transplantation. Key innovations are comprised in the immunosuppressive regimen combining anti-IL-2 receptor monoclonal antibody, sirolimus, and tacrolimus without corticosteroid therapy. In addition, cold ischemia time was kept short, islets were grafted freshly without culturing *in vitro*, and mass of transplanted islets exceeded 9000 islet equivalents per kilogram body weight of the recipient. However, it still subjects patients to the peril of immunosuppression. The ability to safely engraft cells into patients is of particular importance to diseases like type I diabetes that can be effectively treated by traditional means.

Therefore, much effort has been focused on the development of the methods to transplant islets without

employment of immunosuppression. A tissue-engineering method, i.e., bioartificial pancreas (BAP), is the most promising approach to realize this requirement.

BIOARTIFICIAL PANCREAS (BAP)

In BAP, islets are enclosed into a semipermeable (SP) membrane prior to transplantation (Fig. 1). Immunoisolation is based on the premise that islets, which were once sequestered within a SP membrane, are protected from the host immune system. It can inhibit occurrence of adverse host immune reactions against the transplanted islets. Islets enclosed can survive under sufficient supply of oxygen and nutrients through the membrane and must be capable of detecting blood glucose levels and producing a sufficient and properly timed level of mature insulin to maintain glycemic control. Various types of BAP devices have been examined. These include spherical microcapsules, diffusion chambers, and devices anastomosed to the vascular system as arteriovenous shunts or capillary units (Fig. 2).

Membranes

A SP membrane is a common important structural unit in BAP. Immunogenicity of islets, i.e., allogeneic (allo-) or xenogeneic (xeno-), is a key factor for governing the selection of a membrane with the appropriate semipermeability. Allo-transplantation is organ and tissue transplantation between genetically different individuals within one species in which the predominant cause of graft rejection is activation of cellular immunity by interactions of host T cells. On the other hand, xeno-transplantation is transplantation between individuals from different species while humoral immunity including antibodies and complement proteins is thought to play a major role in the graft rejection. In the allo-transplantation, for a recipient without preformed antibodies, a membrane that can mechanically inhibit contact of the host immune cells with the graft is expected to effectively protect the graft from rejection. On the other hand, in xeno- and allo-transplantation, for a recipient carrying natural or preformed antibodies against the donor tissue, a SP membrane must be designed not only to inhibit cellular

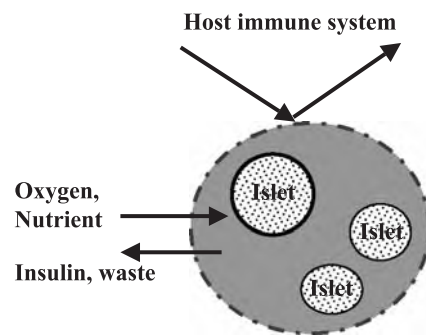


Fig. 1 Basic idea of a bioartificial pancreas (BAP).

contacts, but it also can inhibit permeation of antibodies and complement proteins.

Distinct different types of SP membranes, hydrogel (HG) membranes and ultrafiltration (UF) membranes have been employed for immunoisolation since the early 1970s. In the HG membrane, the increase in the solute molecular weight results in reduction in the diffusion coefficient due to interaction of the solute molecule with the HG polymer network. UF membranes are microporous in character and mechanically sieve molecules according to their molecular weight. For both types of membranes, contact inhibition between the host immune cells and the graft can be easily attained. However, there is still some ambiguity in immunoisolation efficacy of both types of the membranes against humoral immunity.

BAP Devices

Microcapsule

The microcapsule fabricated from alginate and poly-L-lysine (PLL) was originally described by Lim and Sun in 1980.^[1] Microcapsules made of alginate-PLL polyion complex membranes have extensively been reviewed.^[2] Cut-off molecular weights of the membranes can be easily controlled by changing the molecular weights of the polyelectrolytes and the time for polyion complex formation. They do not induce inflammatory reactions. The viability of enclosed islet cells is high. Many groups have examined immunoisolation efficacy of the microcapsules made of alginate-PLL polyion complex membranes. Results in diabetic models using rodents indicated that these BAP can function for periods of several months to more than one year. However, its immunoisolation efficacy has not been consistency in other diabetic models. Long-term survival of microencapsulated porcine islets transplanted into diabetic Cynomolgus monkeys was reported.^[3] Unfortunately, other groups have not yet reproduced

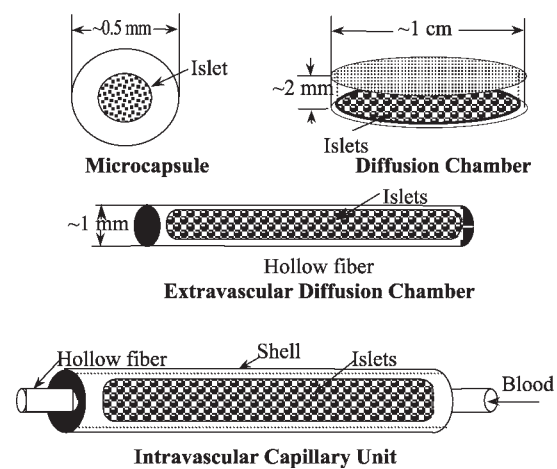


Fig. 2 Schematic representations of various types of bioartificial pancreas (BAP).

the result. Its clinical experience will be discussed subsequently.

One major drawback of this type of polyion complex microcapsules is low mechanical strength. When implanted into recipients, they are seriously deformed and sometimes are even ruptured in the peritoneal cavity. To overcome this problem, microbeads made of HGs such as agarose^[4] and barium-alginate were introduced.^[5] Functions of microencapsulated islets in agarose HG as BAP were extensively examined in vitro and in vivo in Japan.^[6] The ability of the agarose microcapsule to achieve marked prolongation of alloislet graft is shown in Fig. 3. NOD mice that received a transplantation of 1500 nonencapsulated, i.e., naked, islets from C3H mice into the peritoneal cavity demonstrated normoglycemia, but its blood glucose levels reverted to the preoperative high level after 10 days of the transplantation. On the other hand, C3H islets enclosed agarose microcapsules can reverse hyperglycemia for 190 days. Spherical microcapsules containing allo-islets appear to have high therapeutic potential.

Diffusion chamber

Diffusion chambers include planar chambers, flat sheet devices, and cylindrical hollow fiber macrocapsules. Chambers can be implanted in various sites such as the peritoneal cavity, the subcutaneous site, or the renal capsule. They can be also readily retrieved in a relatively simple way. Microporous millipore membrane^[7] and microporous polycarbonate membrane (nucleopore membranes)^[8] have been used to fabricate planar chambers. Usually, in chambers, islets are immobilized in HGs such as collagen, calcium-alginate, or agarose to prohibit aggregation of encapsulated tissue into large

T

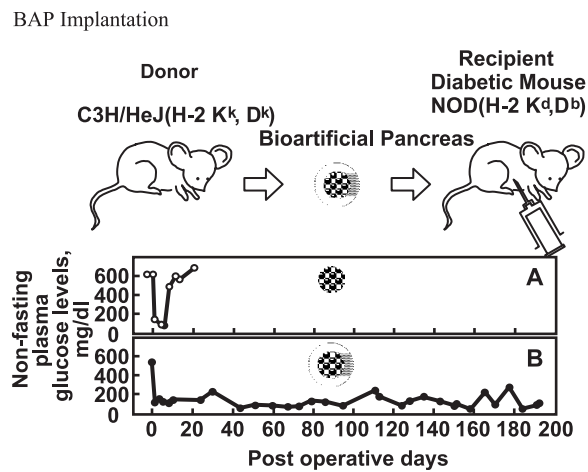


Fig. 3 Plasma glucose levels of diabetic NOD mice receiving 1500 naked islets (A) or microencapsulated islets (B) into the peritoneal cavity. Islets from C3H mice were enclosed in 5% agarose hydrogel microcapsules.

clumps. In the recent 10 years, the hollow fiber geometry has been preferred over the planar chamber due to its smaller foreign body response. In studies, hollow fibers made of poly(acrylonitrile-co-vinylchloride) (Poly(AN-co-VC))^[9] and sodium-methallyl sulphonate, AN69^[10] have been employed. Hollow fiber devices made of Poly(AN-co-VC) demonstrated the feasibility of long-term immunoisolation of islets and the long-term biocompatibility in rodent and canine models. The device with canine islets can provide correction of hyperglycemia in dogs with surgically induced diabetes for a period of more than one year without the use of immunosuppressive drugs. The outer surface showed only scattered foci of macrophages and lymphocytes.^[9] The major drawbacks of this device are that hollow fibers, as a consequence of their shape, tend to break when forced to bend under physiological stress, and that enormous lengths of fibers would be required to enclose a sufficient number of islets to normalize blood glucose levels of human patients.

Intravascular capillary unit

The original device was composed of a number of small-diameter, hollow fibers enclosed in one, large-diameter housing tube.^[11] The hollow fibers were made of Poly(AN-co-VC), that were similar to those used in diffusion chamber devices. Islets were loaded in a space between the hollow fibers and the housing tube. Although the design allows close contact between the islets and the blood flow, the duration of normoglycaemia is usually restricted to several hours due to blood clotting in the lumen of hollow fibers. The intravascular device was modified to be composed of a single, coiled and

tubular membrane with an internal diameter of five to six millimeters.^[12] The tubular membrane has been modified, but it is composed of Poly(AN-co-VC) with a nominal molecular weight cutoff of 50 kDa. The device with allo-islets demonstrated a long-term normoglycemia in a canine model without clotting of blood in the lumen of the tubular membrane. The intravascular devices have shown some degree of success. However, the complications associated with thrombus formation and tissue hyperplasia at the blood vessel anastomosis sites limit its clinical application because the ability to safely engraft cells into patients is of particular importance to diseases like type I diabetes that can be effectively treated by traditional means.

Immunoisolation Efficacy of Membranes Against Humoral Immunity

Prevention of immunocompetent cells entering into the islet compartment is easily attained; thus, immunoisolation of allo-islets is successfully achieved using various artificial membranes. Restoration of normal glucose metabolism by BAP containing allo-islets has already been reported in various animal diabetic models as mentioned in the preceding discussion. However, the supply of donor tissue remains quite inadequate if BAP enclosing allo-islets were to be developed as effective therapy for clinical treatment of human diabetic patients. Xeno-islets from pigs and cows have been considered as potential sources of islets for BAP preparation. One of the major aims for development of BAP devices is to overcome the vigorous humoral and cellular immune responses associated with transplantation of xeno-tissue. Use of porcine and bovine islets is expected to overcome severe shortage of human pancreatic tissue. Various BAP devices, which allow xeno-transplantation in rodent models, have been reported.^[3,4] However, reports on successful immunoisolation of xeno-islets in higher animals, such as dog and monkey, have been limited. Immunoisolation efficacy of membranes, which have been used for preparation of various BAP devices, will be discussed in some detail.

Different types of SP membranes, such as HG membranes and UF membranes, have been employed for BAP preparation. UF membranes with a nominal cut-off molecular weight of approximately 50 kD have been preferentially used to prepare BAP devices with xeno-islets because it has been believed that the membrane allows passage of oxygen and nutrients, but it restricts passage of antibodies of which molecular weights are much higher than 50 kD. The representative membrane employed is Amicon's XM-50 membrane made of (Poly(AN-co-VC)). Researchers working on BAP believe that the solute permeability through a

membrane that is determined by pressure-driven sieving experiments can predict its immunoisolation efficacy. However, it should be carefully borne in mind that the fractionation efficiency of a membrane is known to highly depend on separation processes employed.

In-vitro permeation studies of solutes through the XM-50 membrane were performed under both pressure-driven filtration and concentration-driven permeation processes.^[13] The XM-50 membrane can reject 99.4% of bovine serum albumin (MW = 69 kD) in the filtration solutions under the pressure-driven filtration. It effectively fractionated the solutes according to their molecular weights near 50 kD by UF, as claimed in the supplier's catalog. On the other hand, the XM-50 membrane could not fractionate the solutes on the basis of molecular size under the diffusion-controlled process. When a device made of XM-50 membrane is used in vivo, molecules permeate through the membrane, which is driven by the concentration difference. The UF membrane, such as XM-50 membrane, loses the molecular size cut-off property in the diffusion-driven process and, thus, cannot protect xenografts by fractionating microsolute and immune biomacromolecules according to their molecular weight. The immunoisolation efficiency of the XM-50 membrane has been overemphasized by many researchers. The standard pressure-driven measurements are inadequate to predict the rate or selectivity in diffusion-based devices, and its diffusive transport properties must be determined for a membrane to be used for immunoisolation.

The other type of membrane that has been used for immunoisolation is a HG membrane. It is made of cross-linked, water-soluble polymers. It does not have distinct molecular weight cut-off, and its selectivity is characterized by the diffusion coefficients of solutes. The increase in the molecular weight results in reduction in the diffusion coefficient due to more prominent interaction of the solutes with the polymer network. The time during which a solute diffuses through a membrane with thickness of L and then reaches to the membrane surfaces is roughly expressed by

$$t = \langle L^2 \rangle / aD$$

in which D is the diffusion coefficient of the solute through the HG membrane and a is a number determined by the shape of device, e.g., $a=2$ for a planar chamber or $a=6$ for a sphere. The thickness of the membrane of a planar chamber is assumed to be one millimeter. The diffusion coefficient of antibody through the membrane should be less than 10^{-13} cm²/sec to prolong the time for more than one year. During this time, antibody reaches the surface cells. However, it is quite difficult to render the diffusion coefficient of antibody through the HG membrane that is less than

the 10^{-13} cm²/sec, which simultaneously maintains the rapid insulin release through the same membrane in response to glucose stimulation.

Microencapsulated xeno-islets sustained normoglycemia for long periods in diabetic rodents without immunosuppression.^[6,14,15] Hollow fibers fabricated from an acrylic copolymer, XM-50, were used to encapsulate rat islets for transplantation in diabetic mice.^[16] The fibers effectively inhibit rejection, and xeno-islets in the fibers survived and functioned to maintain normoglycemia. Hyperglycemia returned when the fibers were removed at 60 days. Encapsulation in SP membranes effectively protects xeno-islets from immune destruction in the rodent models. On the other hand, reports on successful immunoisolation of xeno-islets in higher animals have been limited.^[2] Discrepancy observed between rodents and higher animals as recipients might be explained by taking stability of complement proteins into consideration. Antibody binding to antigens on the xeno-cell surface alone cannot give lethal damage to the cells. For destruction of the xeno-cells, antigen-antibody immune complexes should activate the complement system. Some proteins of the rodent complement system, such as C1, C2, and C5, are less stable than those of higher animals. Complement proteins take a certain amount of time to permeate through the membrane and reach the cell surface. It might be expected that labile complement proteins are denatured and thus lose the cytolytic activity during permeation.

At this point, it can be summarized that neither UF membranes or HG membranes can effectively fractionate the solutes on the basis of molecular size and thus cannot immunoisolate xeno-cells from the host immune system for a long period.

Porcine Endogenous Retrovirus Infection

The possible use of porcine islets in preparation of BAP has been actively examined to overcome severe shortage of human islets. However, in 1997, it was reported that porcine endogenous retrovirus (PERV) genomes can be infectious transmitted to human cells in culture.^[17] After this report, many groups have been examining whether PERV genomes can be infectious transmitted to human cells. There are some ambiguities in the possibility of PERV infection from porcine islets to human patients. However, microchimerism and PERV transmission were frequently observed in both mouse and human tissues during grafting of porcine cells into nude mice bearing human tumours and sometimes during porcine islet xenograft in this model.^[18] This strengthens the notion that there is a risk of transmitting PERV during porcine islet xenograft. Even though porcine islet cells, even with PERV

inserts and transcripts, have very little probability of transmitting PERV to human cells during coincubation, it is difficult to rationalize the use of porcine islets for treatment of human diabetic patients. The ability to safely engraft cells into human tissues is of particular importance to insulin-dependent diabetes.

Xeno- β -cells are potential source of cells for BAP. However, tissue from animals carries the risk of transmitting novel infectious agents into humans. As the preceding discussion demonstrated, islets from xeno-animals are more immunogenic than allo-animals; thus, it is difficult to find membranes that can effectively protect xeno-islets reproducibly in higher animals for a long period. In conjunction of these two facts, new sources of insulin-producing human cells must be exploited for treatment of human diabetic patients.

NEW CELL SOURCES

Stem Cell

Two functions are essential to any cells for use in BAP preparation. The cells must be capable of detecting blood glucose levels and releasing a sufficient and properly timed level of mature insulin to maintain glycemic control. Other desirable properties include plentiful and reproducible production of these cells. Stem cells offer a potential solution to these requirements. In theory, stem cells can proliferate in their undifferentiated state and then be converted to take on required functions by the use of genetic and epigenetic manipulations.

Embryonic stem cell

Embryonic stem cells (ES cells) are self-renewing elements that can generate the most cell types in the body. ES cells are an attractive potential source of insulin-releasing tissue for BAP preparation and other cell transplantation therapies. Recent studies, using mouse ES cells as an experimental system, demonstrated the promise of ES cells as a potential source of cells for the treatment of diabetes. Cell selection strategies using insulin promoter to derive the expression of antibiotic resistance genes was employed in derivation of insulin-producing cells from ES cells.^[19] These ES-derived insulin-secreting cells regulated insulin secretion in vitro in the presence of various secretagogues and contained roughly 90% of the insulin content of a normal mouse islet. Clusters obtained from this clone were implanted in the spleen of streptozotocin-induced diabetic animals. Transplanted animals correct hyperglycemia within one week and restore body weight in four weeks.

Another group generates cells with insulin and other pancreatic endocrine markers from mouse ES cells by a multistep cell culture strategy.^[20] The multistep strategy begins with the production of a highly enriched population of nestin-positive cells normally found as neural precursor cell from embryoid bodies. A critical step in this strategy is to plate the embryoid bodies into a serum-free medium (ITSFn) in which many other cell types die, which thus increases the proportion of nestin-positive cell. These cells were then expanded in the presence of a mitogen, a basic fibroblast growth factor (bFGF), in another serum-free medium and followed by mitogen withdrawal to promote cessation of cell division and then differentiation. The culture media, including a B26 supplement and nicotinamide, improved the yield of insulin-producing cell clusters. Insulin secretion occurs in response to physiologically appropriate glucose concentrations. However, insulin-positive cells in these cultures contains 50 times less insulin per cell than normal islet cells.

To increase population of insulin-positive cells in the culture and insulin content in the cells, various modifications have been introduced in the culture condition. In-vitro treatment of human fetal pancreas with nicotinamide and an inhibitor of phosphoinositide 3-kinase (PI3K), LY294002, increased total endocrine cell number and insulin content while reducing DNA synthesis. By referring to this fact, a PI3K inhibitor, in spite of a B26 supplement, was added to the culture medium during development of insulin positive cells from ES cells. The production of insulin at increased levels was greatly improved by this treatment.^[21]

Ectopic transcription factor expression constitutes a powerful genetic approach to stem cell differentiation. Pdx1 and Pax4 are essential for proper β -cell development. Constitutive expression of Pax4 in combination with the multistep cell culture strategy facilitates ES cell differentiation into the pancreatic lineage, which leads to the formation of isletlike spheroid structures that produce increased levels of insulin.^[22]

Insulin-producing cell clusters derived from human ES cells are a more promising alternative for human islets for BAP preparation. Recent data demonstrates that insulin-producing cells can be also derived from human ES cells. Using human ES cells in both adherent and suspension culture conditions, spontaneous in-vitro differentiation of ES cells were observed, and cells with characteristics of insulin-producing β -cells were generated.^[23] Immunohistochemical staining for insulin was observed in a surprisingly high percentage of cells. Secretion of insulin into the medium was observed in a differentiation-dependent manner and was associated with the appearance of other β -cell markers using reverse transcriptase-polymerase chain reaction (RT-PCR). These findings validate the human ES cell

as a possible future source of insulin-releasing tissue for preparation of BAP.

Most of the methods employed for derivation of insulin-releasing cells from ES cells have been developed on the basis of the multistep cell culture method originally reported by McKay et al.^[20] As mentioned above, efficient differentiation of ES cells to insulin-releasing cells have been reported. However, Rajagopal and Melton et al.^[24] supposed that the cells derived from ES cell concentrate the insulin from the culture medium rather than produce hormone from their careful examination of the cells derived from five ES cell lines, both murine and human. They mentioned that it is difficult to exclude the possibility that paucity of β -cell differentiation is due to cell line variability or suboptimal culture condition, but they did conclude that insulin staining alone can overestimate genuine β -cell differentiation when exogenous insulin is present in the culture medium. In spite of the results, Melton remains optimistic and mentioned that no clear data persuades him that it will be impossible to turn ES cells into insulin-producing cells. Much more careful studies should be done to derive insulin-releasing cells from ES cells by employing C-peptide staining, electron microscope, Northern analysis, metabolic labeling, and so forth.

Adult stem cells

Pancreatic stem cells have previously been well reviewed.^[25,26] The pancreas is composed of the exocrine (duct and acinar) and endocrine compartments, but, embryonically, all the cells are differentiated from the duct epithelium. In addition, throughout much of life, there is continued addition of new small islets neogenesis or differentiation from ductal progenitors. Adult, somatic, or nonembryonic stem cells may possess some advantages over ES cells as a cell source for BAP preparation, such as exclusion of the possibility of teratoma formation.

It was demonstrated that cells from pancreatic ductal epithelia of mice proliferate extensively in culture and have the ability to form large numbers of islet like cell cluster in vitro.^[27] The proliferation and differentiation capacity of these cells may warrant their classification as stem cells. These islets produce insulin I and II in addition to glucagon and somatostatin. With the addition of nicotinamide and a glucose load, these structures significantly up-regulate insulin production and can decrease blood glucose levels following transplantation into diabetic NOD mice.

Isletlike clusters of pancreatic endocrine cells were derived from human pancreatic tissue containing ductal fragment and acinar tissue.^[28] Pancreatic tissue from human islet isolations were cultured under conditions that allowed expansion of the ductal cells as a

monolayer in which the cells were overlaid with a thin layer of Matrigel. With this manipulation, 50- to 150- μ m diameter isletlike clusters of pancreatic endocrine cells budded from the monolayer of epithelial cells. Over three to four weeks in culture, the insulin content per flask increased 10- to 15-fold as the DNA content increased up to sevenfold. The human islet buds exposed to stimulatory 20 mM glucose had a 2.3-fold increase in secreted insulin. Thus, duct tissue from human pancreas can be expanded in culture and then be directed to differentiate into glucose-responsive islet tissue in vitro.

The need for transplantable human islets has stimulated efforts to expand existing pancreatic islets and/or grow and differentiate stem cells. These approaches may provide a potential new source of pancreatic islet cells for BAP preparation.

Genetically Engineered Insulin Releasing Cells

For preparation of BAP, plentiful and reproducible production of cells, which can produce a sufficient and properly timed level of mature insulin in response to changes of glucose levels, is required. Genetic engineering of non- β -cells for glucose-responsive insulin secretion offers some promise in developing BAL. But, two major obstacles should be overcome. Most somatic cells simply engineered using proinsulin cDNA cannot process proinsulin into mature insulin and thus release proinsulin. The timing of insulin delivery is crucial for optimal regulation of glucose homeostasis. However, most insulin-releasing cells derived from non- β -cells hardly timely regulate insulin secretion in response to glucose stimulation. Preparation of insulin-releasing cells that is capable of processing and releasing insulin in such a way that normal glucose homeostasis is maintained is crucial for preparation of BAP.

Various attempts have been done to derive non- β -cells that release mature insulin in response to glucose stimulation. In some attempts, insulin gene was driven to express with various glucose-sensitive promoter elements.^[29,30] However, insulin secretion dynamics is not adequate for glycemic regulation. The timing of insulin delivery is crucial for optimal regulation of glucose homeostasis. The slow time course of transcriptional control by glucose makes synchronizing insulin production with the periodic fluctuation in blood glucose levels an extremely difficult task. To overcome this difficulty, some endocrine cell lines were examined. The secretion kinetics of some endocrine cells closely parallels that of insulin, and those endocrine cells can process proinsulin. K cells located primarily in stomach, duodenum, and jejunum secrete the glucose-dependent insulinotropic polypeptide. A tumor-derived K-cell line

was examined to produce human insulin by providing the cells with the human insulin gene linked to the five-foot regulatory region of the gene-encoding GIP.^[31] NCI-H716 intestinal cell line driven from a poorly differentiated caecal adenocarcinoma was also employed.^[32]

All of these works have been done using immortalized cell lines. Those gave good models of gene therapy for diabetes. However, the obtained insulin-releasing cells could not directly apply to the preparation of BAP because cell proliferation in BAP could not be easily controlled.

BAP Made of Insulin-Releasing Cells from Stem Cell

In clinical islet transplantation, approximately 20 milliliters of a packed tissue of islets have been infused to the liver through the main portal vein.^[33] As demonstrated in the preceding discussion, the insulin release ability of the cells derived from ES and somatic stem cells is much less than that of islets. For example, cell clusters derived from ES cells produced insulin at 13% or less of levels found in pancreatic islets.^[21] From simple calculation, we find out that more than 150 milliliters of those packed cells should be transplanted to result in the long-term freedom from the need for exogenous insulin. We will meet a lot of difficulty to implant such a large volume of cell suspension. To minimize trauma, the volume of the final implant should be minimized as possible.

Some BAP devices are mostly dead space and are therefore too large for convenient implantation. For example, when the radius of a microcapsule is three times larger than that of an isletlike cell cluster, a volume of the microcapsule becomes 27 times larger than that of the cluster because the volume of a sphere increases with third powers of the radius. The final volume of the BAP made of insulin-producing cell cluster from the stem cell is:

$$150 \text{ ml} \times 27 = 4050 \text{ ml}$$

It is unrealistic to transplant such a large volume of BAP in the clinical practice. Much effort should be exerted to enhance insulin production of cells derived from stem cells or new concepts for preparation of BAP should be introduced to minimize the BAP volume from now on.

CLINICAL TRIALS

Although serious investigative efforts were devoted to develop BAP during the last quarter of the 20th century, very few clinical trials have been done, and results of only two clinical studies can be available from journals easily accessible. Clinical trials approved by

the U.S. Food and Drug Administration were initiated in 1993 by two investigative groups. The two have used very different approaches. One implanted alginate-PLLA microcapsules containing human islets into the peritoneum of a 38-year-old diabetic male.^[34] The patient, who was immunosuppressed with cyclosporin for a preexisting kidney transplant, received a total dose of 10^6 pooled allo-islets administered. Exogenous insulin administration could be discontinued for one month. However, results are difficult to evaluate because the recipient received tight immunosuppressive therapy. The other groups designed to investigate whether immunoisolation protects islets against the autoimmune challenge of type I diabetes.^[35] This is an important issue because autoimmune rejection of transplanted islets normally destroys unencapsulated islets in a nonimmunosuppressed host. Islets were macroencapsulated in semipermeable hollow fibers made of Poly(AN-co-VC) having a nominal molecular weight cut-off of approximately 65 kDa. Capsules were implanted subcutaneously into type I and type II diabetic patients for two weeks. From histologic analysis and in-vitro glucose challenge tests, the investigators concluded that this trial demonstrated the short-term biocompatibility and acute immunoprotective capacity of the Poly(AN-co-VC) hollow fibers.

TECHNOLOGY OF BAP FOR OTHER DISEASES

Various endocrine diseases, in addition to diabetes, have had experimental and clinical recovery by implantation of bioactive molecules releasing cells enclosed in a SP membrane. The technology of BAP have been expanded to the treatment of chronic pain^[36] and neurodegenerative diseases,^[37] including Parkinson's disease, Alzheimer's, and amyotrophic lateral sclerosis (ALS). Other disorders that have been addressed with this type of therapy include hypocalcemia, dwarfism, anemia, hemophilia, and cancer.^[38]

CONCLUSION

In the past two decades, BAP has progressed from in vitro investigation through small-animal and large-animal studies into clinical investigation. Relevant mechanisms of host recognition and immunoreactivity against allograft appear cellular in nature; therefore, only the exclusion of host effector cells is required for immunoprotection of allo-islets. Membranes of this type are not difficult to prepare. The major obstacle for clinical application of BAP is a severe shortage of insulin-releasing cells. Differentiation and maturation of human insulin-releasing cells from human ES cells or

adult stem cells are not easily achieved. On the other hand, xeno-insulin releasing cells, i.e. porcine islets, are not difficult to obtain. However, tighter membranes, which not only provide protection against cellular attack but also confer protection against immune rejection mediated by antibodies and complement proteins, are required. In addition, the other hard task is to deny the possibility of the risk of transmitting novel infectious agents, such as PERV, from animal tissue into human. In these two approaches, the former seems to be promising. Human ES cells were established, and various approaches to differentiate ES cells to somatic functioning cells have been intensively studied. Human ductal fragment tissue is also another promising source for insulin-releasing cells.

Understanding the immunologic and molecular basis of host response to implanted devices and establishment of methods to differentiate stem cells to insulin-releasing cells, though incomplete, has proved adequate for successful BAP implants containing allo-cells for treatment of type I diabetic patients.

ARTICLES OF FURTHER INTEREST

Bioreactors; Insulin Delivery; Rotary Cell Culture Systems; Stem Cells

REFERENCES

- Lim, F.; Sun, A.M. Microencapsulated islets as bioartificial endocrine pancreas. *Science* **1980**, *210*, 908–910.
- de Vos, P.; Hamel, A.F.; Tatarkiewicz, K. Considerations for successful transplantation of encapsulated pancreatic islets. *Diabetologia* **2002**, *45* (2), 159–173.
- Sun, Y.; Ma, X.; Zhou, D.; Vacek, I.; Sun, A.M. Normalization of diabetes in spontaneously diabetic cynomolgus monkeys by xenografts of microencapsulated porcine islets without immunosuppression. *J. Clin. Invest.* **1996**, *98* (6), 1417–1422.
- Iwata, H.; Takagi, T.; Amemiya, H.; Shimizu, H.; Yamashita, K.; Kobayashi, K.; Akutsu, T. Agarose for a bioartificial pancreas. *J. Biomed. Mater. Res.* **1992**, *26* (7), 967–977.
- Zekorn, T.; Siebers, U.; Horcher, A.; Schnettler, R.; Klock, G.; Bretzel, R.G.; Zimmermann, U.; Federlin, K. Barium-alginate beads for immunisolated transplantation of islets of Langerhans. *Transplant. Proc.* **1992**, *24* (3), 937–939.
- Murakami, M.; Satou, H.; Kimura, T.; Kobayashi, T.; Yamaguchi, A.; Nakagawara, G.; Iwata, H. Effects of micro-encapsulation on morphology and endocrine function of cryopreserved neonatal porcine islet-like cell clusters. *Transplantation* **2000**, *70* (8), 1143–1148.
- Buschard, K. Cultivation of islets of Langerhans in Millipore chamber in vivo. *Horm. Metab. Res.* **1975**, *7*, 441.
- Ohgawara, H.; Hirotani, S.; Miyazaki, J.; Teraoka, S. Membrane immunoisolation of a diffusion chamber for bioartificial pancreas. *Artif. Organs* **1998**, *22* (9), 788–794.
- Maki, T.; Ubhi, C.S.; Sanchez-Farpon, H.; Sullivan, S.J.; Borland, K.; Muller, T.E.; Solomon, B.A.; Chick, W.L.; Monaco, A.P. Successful treatment of diabetes with the biohybrid artificial pancreas in dogs. *Transplantation* **1991**, *51* (1), 43–51.
- Prevost, P.; Flori, S.; Collier, C.; Muscat, E.; Rolland, E. Application of AN69 hydrogel to islet encapsulation. Evaluation in streptozotocin-induced diabetic rat model. *Ann. N.Y. Acad. Sci.* **1997**, *831*, 344–349.
- Chick, W.L.; Like, A.A.; Lauris, V.; Galletti, P.M.; Richardson, P.D.; Panol, G.; Mix, T.W.; Colton, C.K. A hybrid artificial pancreas. *Trans. Am. Soc. Artif. Intern. Organs* **1975**, *21*, 8–15.
- Sullivan, S.J.; Maki, T.; Borland, K.M.; Mahoney, M.D.; Solomon, B.A.; Muller, T.E.; Monaco, A.P.; Chick, W.L. Biohybrid artificial pancreas: Long-term implantation studies in diabetic, pancreatectomized dogs. *Science* **1991**, *252* (5006), 718–721.
- Iwata, H.; Morikawa, N.; Ikada, Y. Permeability of filters used for immunoisolation. *Tissue Eng.* **1996**, *2*, 289–298.
- Sun, A.M.; Vacek, I.; Sun, Y.L.; Ma, X.; Zhou, D. In vitro and in vivo evaluation of microencapsulated porcine islets. *ASAIO J.* **1992**, *38* (2), 125–127.
- Weber, C.J.; Safley, S.; Hagler, M.; Kapp, J. Evaluation of graft-host response for various tissue sources and animal models. *Ann. N.Y. Acad. Sci.* **1999**, *875*, 233–254.
- Lacy, P.E.; Hegre, O.D.; Gerasimidi-Vazeou, A.; Gentile, F.T.; Dionne, K.E. Maintenance of normoglycemia in diabetic mice by subcutaneous xenografts of encapsulated islets. *Science* **1991**, *254* (5039), 1782–1784.
- Patience, C.; Takeuchi, Y.; Weiss, R.A. Infection of human cells by an endogenous retrovirus of pigs. *Nat. Med.* **1997**, *3* (3), 282–286.
- Clemenceau, B.; Jegou, D.; Martignat, L.; Sai, P. Microchimerism and transmission of porcine endogenous retrovirus from a pig cell line or specific pathogen-free pig islets to mouse tissues and human cells during xenografts in nude mice. *Diabetologia* **2002**, *45* (6), 914–923.
- Soria, B.; Roche, E.; Berna, G.; Leon-Quinto, T.; Reig, J.A.; Martin, F. Insulin-secreting cells derived from embryonic stem cells normalize glycemia in streptozotocin-induced diabetic mice. *Diabetes* **2000**, *49* (2), 157–162.
- Lumelsky, N.; Blondel, O.; Laeng, P.; Velasco, I.; Ravin, R.; McKay, R. Differentiation of embryonic stem cells to insulin-secreting structures similar to pancreatic islets. *Science* **2001**, *292* (5520), 1389–1394.
- Hori, Y.; Rulifson, I.C.; Tsai, B.C.; Heit, J.J.; Cahoy, J.D.; Kim, S.K. Growth inhibitors promote differentiation of insulin-producing tissue from embryonic stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (25), 16105–16110.
- Blyszczuk, P.; Czyz, J.; Kania, G.; Wagner, M.; Roll, U.; St-Onge, L.; Wobus, A.M. Expression of Pax4 in embryonic stem cells promotes differentiation of

- nestin-positive progenitor and insulin-producing cells. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (3), 998–1003.
23. Assady, S.; Maor, G.; Amit, M.; Itskovitz-Eldor, J.; Skorecki, K.L.; Tzukerman, M. Insulin production by human embryonic stem cells. *Diabetes* **2001**, *50* (8), 1691–1697.
 24. Rajagopal, J.; Anderson, W.J.; Kume, S.; Martinez, O.I.; Melton, D.A. Insulin staining of ES cell progeny from insulin uptake. *Science* **2003**, *299* (5605), 363.
 25. Soria, B. In-vitro differentiation of pancreatic beta-cells. *Differentiation* **2001**, *68* (4–5), 205–219.
 26. Bonner-Weir, S.; Sharma, A. Pancreatic stem cells. *J. Pathol.* **2002**, *197* (4), 519–526. Review.
 27. Ramiya, V.K.; Maraist, M.; Arfors, K.E.; Schatz, D.A.; Peck, A.B.; Cornelius, J.G. Reversal of insulin-dependent diabetes using islets generated in vitro from pancreatic stem cells. *Nat. Med.* **2000**, *6* (3), 278–282.
 28. Bonner-Weir, S.; Taneja, M.; Weir, G.C.; Tatarkiewicz, K.; Song, K.H.; Sharma, A.; O'Neil, J.J. In vitro cultivation of human islets from expanded ductal tissue. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97* (14), 7999–8004.
 29. Chen, R.; Meseck, M.; McEvoy, R.C.; Woo, S.L. Glucose-stimulated and self-limiting insulin production by glucose 6-phosphatase promoter driven insulin expression in hepatoma cells. *Gene Ther.* **2000**, *7* (21), 1802–1809.
 30. Thule, P.M.; Liu, J.; Phillips, L.S. Glucose regulated production of human insulin in rat hepatocytes. *Gene Ther.* **2000**, *7* (3), 205–214.
 31. Cheung, A.T.; Dayanandan, B.; Lewis, J.T.; Korbitt, G.S.; Rajotte, R.V.; Bryer-Ash, M.; Boylan, M.O.; Wolfe, M.M.; Kieffer, T.J. Glucose-dependent insulin release from genetically engineered K cells. *Science* **2000**, *290* (5498), 1959–1962.
 32. Tang, S.C.; Sambanis, A. Development of genetically engineered human intestinal cells for regulated insulin secretion using rAAV-mediated gene transfer. *Biochem. Biophys. Res. Commun.* **2003**, *303* (2), 645–652.
 33. Ryan, E.A.; Lakey, R.T.; Rajotte, R.V.; Korbitt, G.S.; Kin, T.; Imes, S.; Rabinovitch, A.; Elliot, J.F.; Bigam, D.; Kneteman, N.M.; Warnock, G.L.; Larsen, I.; Shapiro, A.M.J. Clinical outcomes and insulin secretion after islet transplantation with the edmonton protocol. *Diabetes* **1999**, *50*, 710–719.
 34. Soon-Shiong, P.; Heintz, R.E.; Merideth, N.; Yao, Q.X.; Yao, Z.; Zheng, T.; Murphy, M.; Moloney, M.K.; Schmehl, M.; Harris, M.; Mendez, R.; Mendez, R.; Sandford, P.A. Insulin independence in a type 1 diabetic patient after encapsulated islet transplantation. *Lancet* **1994**, *343* (8903), 950–951.
 35. Scharp, D.W.; Swanson, C.J.; Olack, B.J.; Latta, P.P.; Hegre, O.D.; Doherty, E.J.; Gentile, F.T.; Flavin, K.S.; Ansara, M.F.; Lacy, P.E. Protection of encapsulated human islets implanted without immunosuppression in patients with type I or type II diabetes and in nondiabetic control subjects. *Diabetes* **1994**, *43* (9), 1167–1170.
 36. Buchser, E.; Goddard, M.; Heyd, B.; Joseph, J.M.; Favre, J.; de Tribolet, N.; Lysaght, M.; Aebischer, P. Immunoisolated xenogenic chromaffin cell therapy for chronic pain. Initial clinical experience. *Anesthesiology* **1996**, *85* (5), 1005–1012.
 37. Emerich, D.F.; Winn, S.R. Immunoisolation cell therapy for CNS diseases. *Crit. Rev. Ther. Drug Carr. Syst.* **2001**, *18* (3), 265–298.
 38. Zielinski, B.A.; Lysaght, M.J. Immunoisolation. In *Principles of Tissue Engineering*, 2nd Ed.; Lanza, R.P., Langer, R., Vacanti, J., Eds.; Academic Press: San Diego, CA, 2000; 321–329.

Tissue Engineering of Peripheral Nerve

Archit B. Sanghvi

Department of Biomedical Engineering, University of Texas, Austin, Texas, U.S.A.

Joy L. Murray

Department of Chemistry and Biochemistry, University of Texas, Austin, Texas, U.S.A.

Christine E. Schmidt

Department of Biomedical Engineering, University of Texas, Austin, Texas, U.S.A.

T

INTRODUCTION

Nerve damage debilitates thousands of Americans each year and leads to problems such as paralysis. In addition to the physical and emotional outcome, financial burdens amounting to nearly \$7 billion in the United States of America alone are associated with nerve injury (American Paralysis Association, 1997). In fact, nearly 50,000 peripheral nerve repair procedures were performed in 1995 (National Center for Health Statistics, 1995). Nerve injuries can occur in the spinal cord in the central nervous system or, more commonly, in nerve bundles in the peripheral nervous system. Nerve injury can result from mechanical, thermal, chemical, or pathological mechanisms.^[1]

Current clinical methods to treat peripheral nerve injury rely primarily on either end-to-end anastomosis or grafts (e.g., autologous nerve or vein grafts) to bridge the severed nerve ends. Short gap repair has shown success with end-to-end reconstruction. However, this technique is limited by the tension introduced into the nerve cable when the nerve ends are reconnected, which can inhibit nerve regeneration. The use of autologous nerve grafts addresses the problems associated with end-to-end reconstruction, but there exist some limitations including incomplete regeneration and the possibility of irreversible damage.^[1] In addition, there are other disadvantages associated with the use of autologous grafts, including loss of function at the donor site, the inability to directly match the size of the donor nerve to the injured nerve, and the need for multiple surgeries.

The shortcomings of the current clinical methods have motivated the search for new engineering techniques that stimulate nerve regeneration. One such technique is the development of a synthetic guide that provides an optimal physical and chemical environment to enhance peripheral nerve repair.

This article highlights the recent advancements in the area of peripheral nerve tissue regeneration. Collaborative efforts between disciplines ranging from biology to materials science are facilitating the development of engineering strategies to enhance peripheral nerve regeneration. A selection of the current approaches

researched for peripheral nerve tissue engineering are reviewed in this article including guidance therapies, and physicochemical- and genetic-based approaches.

NERVOUS SYSTEM ANATOMY

The nervous system regulates and controls the sensory and motor functions of the body (Fig. 1A). It is divided into two main subsystems: the central nervous system (CNS) and the peripheral nervous system (PNS). Together the CNS and PNS monitor and control every organ system through positive and negative feedback loops. Communication in the nervous system is modulated by the transmission of electrical signals. The CNS, consisting primarily of the brain and the spinal cord, is linked to the PNS via branched spinal, cranial, and autonomic nerves. The PNS, which is composed of nerves (bundles of neuronal fibers), connects the CNS to other parts of the body.^[2] The CNS is distinguished from the PNS in that it contains both the nerve cell body and their processes, whereas the PNS consists only of the processes. Dorsal root ganglia (DRG), which carry incoming sensory information, are the only exception; DRG cell bodies are contained in the PNS along with their neuronal processes.

The basic functional unit of the nervous system is the neuron, which is composed of the nerve cell body and its neuronal processes. The cell body consists of the usual organelles common to all cells. Usually the cell body is located within the CNS and serves as the focal point for outgrowth of nerve fibers. Neuronal processes (i.e., dendrites and axons) are cytoplasmic extensions from the cell body that expand in various directions. Their primary function is to communicate between the nerve cell body and other cells. Dendrites function to receive signals in the form of graded potentials that travel over short distances. Axons serve to transmit signals in the form of action potentials. Action potentials propagate through the axon and end at the terminal branches near its target, which can be another neuron, muscle fiber, or end sensory organ such as pacinian corpuscles in subcutaneous tissue.^[3,4]

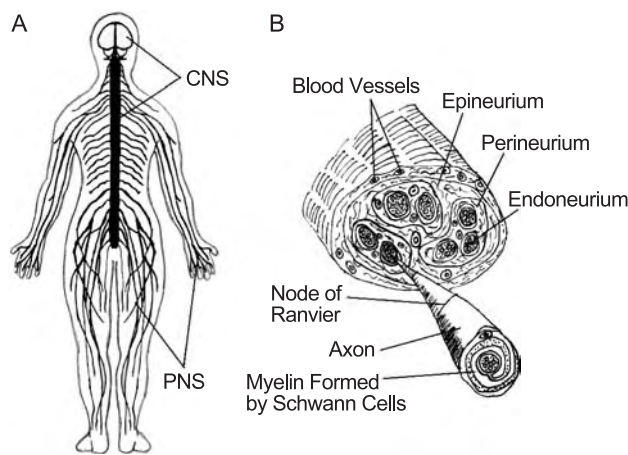


Fig. 1 The nervous system: A) The central nervous system (CNS) includes the brain and spinal cord. The peripheral nervous system (PNS) contains two main pathways (afferent and efferent) and connects the CNS to other parts of the body. B) Schematic diagram showing enlarged segment of neurons that are arranged to present the relationship of the axons to the myelin sheath formed by Schwann cells and connective tissue organization.

Glial cells, which are more abundant than neurons, play a critical role in nerve function. They provide physical and biochemical support for neurons. Specifically, Schwann cell cytoplasm insulates axons found in the PNS by creating myelin, multiple concentric layers of phospholipid membrane.^[3] In addition to being wrapped with myelin, axons are organized by a connective tissue that wraps around concentric layers, resulting in a bundle or fascicle (Fig. 1B).

NERVE TISSUE DEGENERATION

Injury and the subsequent response by nerve tissue is a well documented process in the PNS. Moments after injury, a process known as Wallerian or anterograde degeneration occurs in the distal end of the nerve stump (the end farthest from the cell body). First, the neuronal cytoskeleton begins to break down, followed shortly thereafter by the partial dissolution of the cell membrane. Fragmentation of the myelin sheath occurs and leads to the formation of axon and myelin fragments (ovoids), as illustrated in Fig. 2. The proximal end of the nerve stump (the end closest to the cell body) swells but often only undergoes minimal degradation. It is through the proximal stump that the neuron retains the potential to regenerate and reestablish nerve function.^[5]

Myelin protein is collected in ovoids, which flatten as the axon they enclose deteriorates. The loss of myelin due to neuronal injury suggests that signals from the axon are necessary to maintain both myelin

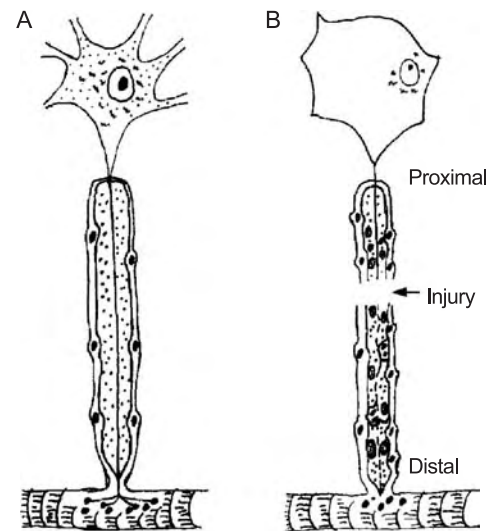


Fig. 2 Wallerian degeneration: A) Illustrated is a normal nerve fiber in its intact form innervating skeletal muscle. Represented in this image are the positions of the nucleus within the cell body and organized Schwann cells. B) When severing injury occurs, the nucleus moves to the cell periphery, and the distal end of the nerve fiber and its myelin sheath disintegrate. This is followed by phagocytosis of myelin debris by macrophages. Adapted from Ref. [5].

protein production and Schwann cell insulation. Eventually, phagocytic cells, such as macrophages, clean myelin debris; this process can take up to seven months in humans.^[6]

MODE OF AXONAL REGENERATION

Axonal regeneration requires motility of the growth cone (a specialized region on the tip of a growing neurite) and extension into damaged and degenerating nervous tissue. The mammalian PNS has a well recognized ability to support axonal regeneration, but PNS regeneration is limited by a slow rate of axonal growth that diminishes the ability of the damaged nerve to reestablish appropriate cell-to-cell signaling. Regeneration is dependent on the type of injury (e.g., crush, partial cut, complete transection) and can be limited by the size and location of the injury.

Crush injuries generally do not require surgical intervention. However, in the case of transection injury, surgical intervention is often necessary to repair and guide regenerating axons into the growth environment of the distal stump. Although distal segments of the axons degenerate, connective tissue elements (e.g., extracellular matrix proteins) of the distal stump generally remain intact. Despite the surgical efforts to utilize the remaining distal stump, repair often fails to achieve complete functional recovery. This problem is

further complicated if the injury requires regeneration over great distances. The regeneration period can be much longer than predicted from reported regeneration rates of 3–5 mm/day in humans.^[7] Therefore, it can take months and even years to achieve functional recovery.^[8]

Axonal outgrowth mediated by growth cone motility is accompanied by proliferating Schwann cells that eventually bridge the fibrin matrix that is formed post injury (Fig. 3A). Axonal sprouts emerge from the first node of Ranvier (unmyelinated region of a myelinated axon) proximal to the injury site.^[9] Axons with multiple sprouts could initially grow slowly until enough sprouts assemble to allow routing of structural material (basal lamina tubes generated by Schwann cells) to the dominant sprout, resulting in nerve fiber regeneration (Fig. 3B). If growing neurites do not enter these tubes and reestablish contact with Schwann cells, then growth is disorganized and regeneration remains incomplete even after several months (Fig. 3C). Growing neurites can form a painful neuroma if they are unable to find a path to direct their growth.^[5]

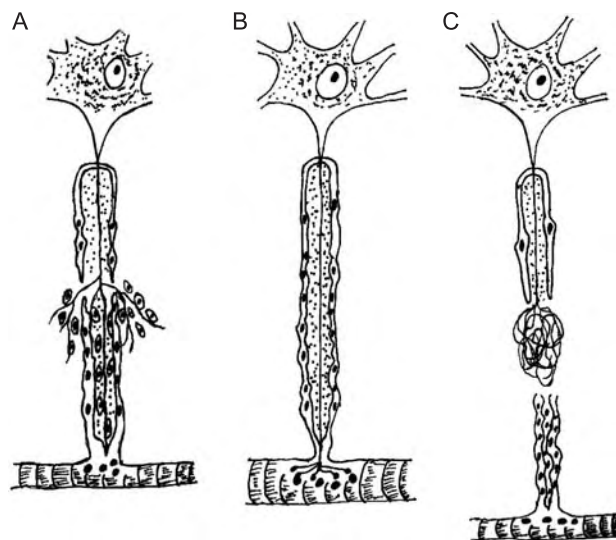


Fig. 3 Axonal regeneration: A) Following injury, Schwann cells lead neuronal outgrowth by forming a bridge between the proximal and the distal stump. This is succeeded by the formation of neurites that sprout from the proximal end. Schwann cells play multiple roles in neurite formation by serving as guides for reinnervation of the regenerating axon and by providing a chemical environment conducive to neuron survival and growth. B) Many of the neurites do not survive, but with the help of Schwann cell tubes, most of the neurites proceed to their destination successfully. C) If the neurites do not reestablish contact with the distal stump, then they become disorganized, leading to poor alignment of nerve ends, formation of a neuroma, and failed reinnervation. Adapted from Ref. [5].

ASSESSMENT OF FUNCTIONAL RECOVERY

Recovery of function in injured peripheral nerves is dependent on axonal outgrowth, reinnervation of target organs, and the subsequent establishment of functional characteristics of regenerated nerve fibers.^[10] One of the most common experimental models used is the rat sciatic nerve for evaluating motor and sensory function. Despite research efforts to develop a model system to study peripheral nerve regeneration, the rat sciatic nerve remains the choice for many neuroscientists. The rat sciatic nerve is comparable to subhuman primates because it has equivalent capacity for regeneration and for functional recovery evaluation.^[11] Several parameters are used to assess nerve regeneration including histology (axon count and diameter), electrophysiology (compound nerve action potentials and compound muscle action potentials), axoplasmic and horseradish peroxidase transport across the nerve reconstruction site, and sciatic function index (SFI) of walking tracks.^[12,13]

Histology is the most common method of monitoring nerve regeneration.^[1] Histology uses light or electron microscopy to determine the overall diameter of the regenerated nerve cable, size and number of axons, axon morphologies, presence of Schwann cells, and degree of myelination at the midpoint of the regenerated nerve cable. Histological sample analysis is subjective and does not necessarily indicate reinnervation and the subsequent functional recovery. Many studies report only histological findings, because these can be performed rapidly. Another method called retrograde labeling permits direct evaluation of reinnervation and axonal branching. In this method, a target muscle is injected with dye that is conducted along the axonal processes, using retrograde transport, back to the cell bodies. Both histological and retrograde labeling require that the animal be sacrificed, eliminating the possibility of any further measurement of regeneration in the same animal. Many studies combine structural and functional methods to give a more accurate assessment of true regeneration.^[1]

Another method of assessing nerve regeneration measures electrophysiological response by placing electrodes proximal and distal to the transection site.^[10] Electrophysiological response measures compound nerve action potential and conduction velocity of regenerated fibers. Electromyography (EMG) provides an alternative where information on reinnervation is obtained by measuring compound action potentials in the muscle after stimulation of the nerve. The main advantage of EMG over morphological studies is that it is a functional rather than structural measure of regeneration that can be monitored over the duration of the regeneration process.^[1,10]

Walking track analysis, which is used to develop sciatic function index (SFI), is an additional method for

assessing functional recovery. In 1982, De Medinaceli first reported the use of SFI to evaluate total lower-limb function.^[14] Walking track analysis is a noninvasive method that provides a quantitative means to evaluate sciatic nerve function based on several measurements of the footprints made on X-ray film. Since then investigators have modified SFI analysis by using finger paint and white paper, which eliminated radial diffusion problems commonly associated with X-ray films.^[15,16] The factors that are involved in SFI analysis include print-length factor (PLF), toe-spread factor (TSF), and intermediary toe-spread factor (ITF). These factors are then applied to the Bain-Mackinnon-Hunter (BMH) sciatic function index-formula: $SFI = -38.3 \times PLF + 109.5 \times TSF + 13.3 \times ITF - 8.8$. An SFI of 0 is normal and SFI of -100 indicates total impairment, which would result from complete transection of the sciatic nerve.^[12] Measurements are based on the footprints that are the direct outcome of coordinated function involving sensory input, motor response, and cortical integration. Therefore, many researchers believe that SFI provides a better method for evaluating functional recovery when compared to basic electrophysiology and morphometric analysis of axon growth and muscle reinnervation.^[11]

ENGINEERING STRATEGIES FOR NERVE REGENERATION

Researchers have combined knowledge from various disciplines to address the limitations of current approaches for improving nerve regeneration. Some engineering strategies currently being investigated include nerve guidance channels, acellular nerve grafts, support cells, neurotrophic factors, biomimetic strategies, stem cells, and gene therapy.

Nerve Guidance Channels

The nerve guidance channel, also called the nerve conduit, is one method used to bridge the gap between the proximal and distal ends of the nerve when nerve damage spans a distance that is too great for anastomosis. The channel allows the natural wound healing process to occur in a protected environment. In addition to providing a physical pathway, the conduit decreases the infiltration of scar tissue and increases concentrations of locally produced growth factors.

Several physical properties of the guidance channel appear to affect the extent and quality of regeneration dramatically. Some of these properties, depicted in Fig. 4, include conduit dimensions, wall porosity, surface texture of the inner luminal compartment, and inherent electrical properties.^[17] Conduits may swell *in vivo*; thus, the conduit dimensions need to be large

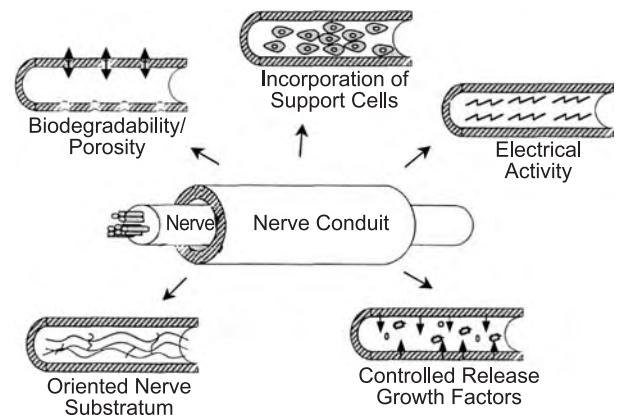


Fig. 4 Properties of a nerve conduit that are under active research include (clockwise from top-left): 1) biodegradability or a porous channel wall, or both; 2) the inclusion of support cells; 3) electrical activity; 4) the release of growth-promoting molecules; and 5) the inclusion of internally oriented matrices.

enough to allow for swelling without inhibiting axonal growth. The walls of the conduit should be sufficiently porous to allow nutrients to flow freely, but the pores should be small enough to exclude inflammatory cells. The ideal guidance channel would be composed of a material that maintains its physical integrity long enough for the nerve to fully regenerate, but eventually degrades leaving only natural tissue. It is believed that regeneration distance can be increased by the optimum combination of guidance channel materials and growth factors and/or support cells in the conduit lumen.^[17]

Many natural and synthetic materials have been explored for use as guidance channels. Some of the natural materials tested in nerve applications include veins, laminin, fibronectin, collagen, fibrin, hyaluronic acid, alginate, and agarose.^[18] Natural materials tend to have good biocompatibility and low toxicity and augment the migration of support cells. Some difficulties associated with the use of natural materials include a possible host immune response (necessitating the use of immunosuppressive drugs) and challenging manufacturing constraints (e.g., batch-to-batch variations and small sources of natural components).^[17]

As an alternative to natural materials, synthetic materials are attractive for use in guidance channels because of the control that can be obtained over the chemical and mechanical properties. Properties of synthetic polymers can be tailored to possess a wide range of physical characteristics (e.g., biocompatibility, porosity, degradation, and mechanical strength).^[19] Synthetic materials can be formed into a variety of macroscopic structures such as meshes, sponges, and bundled tubes in an attempt to mimic the natural nerve tissue structure more closely.^[18]

Initial studies of synthetic guidance channels used silicone tubes because of their flexibility and

biocompatibility. The silicone tube is a useful model for studying nerve regeneration in vivo because it is transparent and thus allows spatial and temporal examination of the regeneration process.^[20] A variety of studies have explored the effect of growth-promoting matrices (e.g., laminin and collagen) placed within the lumen of the silicon guidance channel.^[18,19] The downside of this model is that the silicone tube is not biodegradable. Fibrosis and chronic nerve compression are associated with long term use of nondegradable conduits.

To overcome limitations of nondegradable materials, Langer and coworkers have worked with a bioresorbable polymer system of poly(lactic acid) (PLA), poly(glycolic acid) (PGA), and their copolymers, poly(lactic-co-glycolic acid) (PLGA).^[21] These polymers have been FDA approved for a variety of applications and have many attractive features for tissue engineering. One can control degradation by varying molecular weight, copolymer ratios, and crystallinity. These polymers are easily processed into a variety of structures including microparticles, films, and nerve guidance conduits.^[19,22]

In addition to PLGA, a variety of other synthetic materials have been explored. Nonresorbable polymers that have durability such as poly(tetrafluoroethylene) and resorbable polymers such as poly(organo)phosphazine and poly(caprolactone) have been explored to improve the rate and quality of regeneration.^[23–25] Another material, poly(vinylidene fluoride) (PVDF), has also been explored for nerve guidance channels. PVDF is a piezoelectric material that has the ability to generate transient surface charges by mechanical deformation. This unique property of PVDF has been exploited for nerve guidance channel development and has been shown to enhance nerve regeneration.^[26] A complete listing of all the materials used to create guidance channels is beyond the scope of this article. Table 1 highlights some of the polymers that have been researched. For an excellent review of the natural and synthetic materials used in this area, see Seal et al.^[18]

Recent advancements have led to the development of a few nerve conduits that have received FDA approval. Some of these include collagen-based nerve conduits by Collagen Matrix, Inc.^[33] and Integra NeuroSciences, which markets the NeuraGen™ conduit.^[34] Another nerve conduit, available from SaluMedica, is an elastomer/hydrogel-based system that is marketed as SaluBridge.^[35] However, these conduits can be used only for small defects (<5 mm). Thus, strategies to repair larger defects are still under active research.

Acellular Grafts

An alternative to nerve guidance channels and autologous nerve grafts is the use of an acellular nerve

Table 1 Nerve conduits—a select list of materials for nerve guidance channels

	Reference
<i>Synthetic nonresorbable polymer</i>	
Poly(vinylidene fluoride) (PVDF)	[26]
Poly(tetrafluoroethylene) (PTFE)	[23]
Silicone	[20]
<i>Synthetic resorbable polymer</i>	
Poly(glycolic acid) (PGA)	[27]
Poly(lactic acid) (PLA)	[28]
Poly(ester)	[29]
Poly(organo)phosphazine	[24]
Poly(L-lactide- ω -caprolactone)	[25]
<i>Growth-promoting luminal matrices</i>	
Agarose	[30]
Collagen	[31]
Fibrin	[32]
Matrigel	[23]

graft. Acellular nerve grafts are nonautologous tissues that have been treated to eliminate all living components, while maintaining extracellular matrix (ECM) structure. Possible sources of nerve graft tissue include: 1) allografts or nerve tissue from the same species and 2) xenografts or nerve tissue from a different species. Currently, autologous grafts are the only type of graft used clinically, because cellularly intact allografts and xenografts generate a vigorous immune response in the host. However, acellular nerve grafts could provide an alternative to autologous grafts because they do not provoke an immune response.

A review by Lundborg discusses a variety of studies using acellular nerve and muscle allografts for peripheral nerve regeneration.^[35] Specifically, acellular muscle tissue grafts have been explored because of the similar structure of the basal lamina tubes found in nerve and muscle tissue. Small amounts of muscle tissue can be removed without significant loss of function.^[36] Acellular grafts composed primarily of insoluble ECM components can be used to provide a scaffold for nerve regeneration. Some methods used to develop acellular grafts are: 1) detergents; 2) alternate cycles of freezing and thawing; and 3) irradiation. These methods aim to remove the living cellular components while preserving the ECM structure.^[17,35] Biomolecules and support cells can be subsequently introduced into the acellular graft to improve graft performance.

Support Cells and Neurotropic Factors

Glial cells (e.g., Schwann cells) are necessary for optimal axon migration. These cells release bioactive factors that further enhance nerve migration,

T

although the underlying mechanism for this system is not well understood. One study showed that a nerve can regenerate across a distance of 15–20 mm in a rat model using a guidance channel seeded with Schwann cells, compared to 10 mm in a control channel without Schwann cells.^[37] Schwann cells secrete an assortment of neurotropic/neurotrophic factors such as nerve growth factor (NGF), brain-derived neurotrophic factor (BDNF), and ciliary neurotrophic factor (CNTF). Schwann cells are also a common target for gene therapy, which attempts to boost in-vivo production of regeneration-promoting factors. Gene therapy will be discussed in greater detail later in this chapter.

Soluble neurotropic factors, which influence direction of neuron outgrowth, and neurotrophic factors, which promote neuron survival, have been incorporated directly into nerve guidance channels. Some of these factors include NGF, BDNF, insulin-like growth factor (IGF-1, IGF-2), platelet-derived growth factor (PDGF), fibroblast growth factor (FGF), and CNTF.^[19] Both neurotropic and neurotrophic factors have been delivered by osmotic pumps,^[38] gels,^[39] and fibronectin mats^[40] to provide temporal release and spatial organization to guide neuronal regeneration.

Biomimetic Strategies

In spite of the recent use of various natural and synthetic materials to bridge severed nerves, no material has surpassed the current clinical choice of the autologous nerve graft. In order to develop nerve guidance channel materials that enhance nerve regeneration, researchers are focusing on modifying existing materials to change their interfacial and biological properties. Current research efforts are focused on developing biomimetic systems that are composed of and organized with biomolecules that play a significant role in the development of neural tissue.

The overall goal of biomimetic strategies is to create a microenvironment that mimics the natural growth environment. Guidance cues in the microenvironment play a critical role in directing axonal growth during embryonic development. These guidance cues can be either instructive or permissive. Instructive cues lead to growth-cone motility based on the combination of inhibitory and attractive growth factors. Permissive guidance cues (e.g., laminin, fibronectin, and type IV collagen), which are found in the basal lamina, aid in supporting neurite outgrowth.^[42] By structurally patterning these permissive guidance cues, researchers have shown the ability to enhance nerve fiber extension both in vitro and in vivo.^[43–45]

A biomimetic surface is created in a selective manner to control cell receptor-mediated interactions with

ligands (guidance cues). By spatially controlling the distribution of ligands, patterned substrates can be used to guide neurite extension. One of the first attempts in substrate patterning demonstrated that linear paths of laminin promoted the initiation and growth of neurites from dorsal root ganglia.^[46] More recent studies using microlithography and microcontact printing have shown potential in using collagen, laminin, and other extracellular matrix (ECM) proteins for enhancing nerve regeneration.^[47,48] Biomolecules that mimic laminin and fibronectin activity were immobilized on PLA surfaces with micron-scale precision.^[49] Results of this study showed spatial control of cells and directed neurite extension.

There are many lithographic techniques used to control surface chemistry for developing micropatterned substrates. A thorough review by Wheeler et al.^[50] provides an explanation of these techniques including photoresist, laser ablation, and microcontact printing. The overall goal of these techniques is to develop a micropatterned substrate that contains two different chemical components: one component with a nonadhesive property and the other with an adhesive property. By using the adhesive component, cell attachment and the subsequent activity can be studied. These types of advancements in surface patterning have provided enormous opportunities for developing micropatterned substrates for axonal guidance and ultimately for designing nerve regeneration systems.

Stem Cell-Based Approaches

The versatility of stem cells to differentiate into specialized cells has drawn tremendous attention to their potential in the treatment of neurodegenerative diseases such as Parkinson's disease, multiple sclerosis, and Huntington's disease. Recent work using adult bone-marrow stromal cells (MSCs) that were transplanted into the distal stump of rat sciatic nerve showed significant improvement as measured using walking track analysis.^[51] Because MSCs injected into a lesioned peripheral nerve can survive, migrate, differentiate into Schwann cells, and promote functional recovery, they may be an important source for cellular therapy in several neurological diseases. Such studies show great potential for the use of stem cell transplantation in neuronal regeneration, but ultimately the understanding of optimal differentiation conditions and disease pathology will be necessary for developing stem cell-based therapy. As the basic biology and the capacity of stem cell-based therapy is further elucidated, the potential for using stem cells in nerve regeneration applications will improve.

Gene Therapy–Based Approaches

Advancements in gene transfer techniques have provided new opportunities to induce expression of specific proteins in selected cells. The ideal gene delivery system should be able to transduce nondividing cells, target specific cells, be nontoxic and nonimmunogenic, have high transfection efficiency, be able to bypass or overcome intracellular endosomal mechanisms, and be able to translocate into the nucleus effectively or integrate into the cellular genome without disturbing innate gene expression.^[52]

The natural tendency of viruses to infect host cells has been utilized for its advantage in gene therapy. By deleting the virulence genes with recombinant DNA tools, viruses can be used as effective carriers for a gene of interest (Navarro). Modified adeno-associated virus (AAV) and herpes simplex virus (HSV) vectors can deliver and express foreign genes in neurons. Several studies using AAV as a method for delivering genes for neurotrophic factors such as BDNF, CNTF, glial cell line–derived neurotrophic factor (GDNF), and neurotrophin (NT-3), have been shown to prevent loss of motoneurons after axotomy in neonatal rats.^[53–55] The overall goal of these studies was to induce overexpression of neurotrophic factors within a peripheral nerve graft to promote axonal growth. While preliminary results show potential for using viral vectors for gene therapy, further characterization and optimization are necessary to develop an effective treatment.

Although viral vectors exhibit high transfection efficiency, their use in vivo has resulted in several drawbacks such as nonspecific targeting, low capacity for carrying large therapeutic genes, risk of replication, and a possible host immune reaction.^[56–58] The safety issue in using viruses as gene carriers was brought into focus by the recent death of an 18-year-old gene therapy patient.^[59] This has compelled advancements in gene therapy with nonviral vectors to achieve safe and efficient transfection (e.g., polymer microspheres and liposomes). However, to date, nonviral vectors have not been applied for peripheral nerve tissue regeneration.

Genetic engineering techniques have been used to enhance nerve regeneration by promoting the local expression of neurotrophins (e.g., NGF), cell-adhesion molecules (e.g., NgCAM and axonin-1/TAG-1), and cytoplasmic surface phosphoproteins (e.g., B-50/GAP-43).^[60] One particular study utilized genetically engineered human dermal fibroblasts (hDFBs) to release human NGF. Results from this study, both in vitro and in vivo, showed significant enhancement of NGF levels that mimic the natural environment for regenerating nerve.^[61]

Although experiments with gene therapy–based techniques to stimulate regeneration have shown some success, a greater understanding of factors that

regulate neurogenesis is necessary. Once a greater understanding of the underlying pathologies of neurodegenerative diseases is achieved, gene therapy approaches will provide the most promising and natural approach to nerve tissue regeneration.

CONCLUSION

Despite the growing sophistication of medical technology, nerve tissue regeneration remains a frequent, costly, and serious problem in health care. Limited regeneration has been accomplished by using the autologous graft, but a more effective and convenient system to overcome peripheral nerve injury is still under investigation. The strategies highlighted in this chapter have shown preliminary success. A combination of these promising approaches will provide a better treatment option for damaged nervous tissue.

ARTICLES OF FURTHER INTEREST

Extracellular Matrix Scaffolds; Gene Therapy; Nerve Guides; Neuroprostheses; Peripheral Nerve Repair and Regeneration: Historical Overview; Stem Cells; Tissue Engineering Scaffolds

REFERENCES

1. Furnish, E.J.; Schmidt, C.E. Tissue Engineering of the Peripheral Nervous System. In *Frontiers in Tissue Engineering*, 1st Ed.; Patrick, C.W., Jr., Mikos, A.G., McIntire, L.V., Eds.; Elsevier Science: New York, 1998.
2. Kandel, E.R.; Schwartz, J.H.; Jessell, T.M. *Principles of Neural Science*, 3rd Ed.; Elsevier Science: New York, 1991.
3. Silverthorn, U.D. *Human Physiology An Integrated Approach*, 2nd Ed.; Prentice-Hall, Inc.: Upper Saddle River, NJ, 2001.
4. Berne, R.M.; Levy, M.N. *Physiology*, 4th Ed.; Mosby: St. Louis, MO, 1998.
5. Ross, M.H.; Romrell, L.J.; Kaye, G.I. Nervous Tissue. In *Histology A Text and Atlas*, 3rd Ed.; Williams & Wilkins: Baltimore, MD, 1995; 256–284.
6. Chaudhry, V.; Glass, D.J.; Griffin, W.J. Wallerian degeneration in peripheral nerve disease. *Neurol. Clin.* **1992**, *10* (3), 613–627.
7. Jacobsen, S.; Guth, L. An electrophysiological study of the early stages of peripheral nerve regeneration. *Exp. Neurol.* **1965**, *11*, 48–60.
8. Kline, D.G.; Hudson, A.R. Nerve injuries: Operative results for entrapments and tumors. *J. Exp. Biol.* **1995**, *19*, 14–44.
9. Morris, J.H.; Hudson, A.R.; Weddell, G.A. Study of degeneration and regeneration in the divided rat sciatic nerve based on electron microscopy. II. The

- development of the regenerating unit. *Z. Zellforsch. Mikrosk. Anat.* **1972**, *124*, 103–130.
10. Fugleholm, K.; Schmalbruch, H.; Krarup, C. Post reinnervation maturation of myelinated nerve fibers in the cat tibial nerve: Chronic electrophysiological and morphometric studies. *J. Peripher. Nerv. Syst.* **2000**, *5*, 82–95.
 11. Varejao, A.S.P.; Meek, M.F.; Ferreira, A.J.A.; Patricio, J.A.B.; Cabrita, A.M.S. Functional evaluation of peripheral nerve regeneration in the rat: Walking track analysis. *J. Neurosci. Methods* **2001**, *108*, 1–9.
 12. Kanaya, F.; Firrell, J.C.; Breidenbach, W.C. Sciatic function index, nerve conduction tests, muscle contraction, and axon morphometry as indicators of regeneration. *Plast. Reconstr. Surg.* **1996**, *98*, 1264–1271.
 13. Mackinnon, S.E. Sciatic function index, nerve conduction tests, muscle contraction, and axon morphometry as indicators of regeneration—Discussion. *Plast. Reconstr. Surg.* **1996**, *98*, 1272–1274.
 14. De Medinaceli, L.; Freed, W.J.; Wyatt, R.J. An index of the functional condition of rat sciatic nerve based on measurements made from walking tracks. *Exp. Neurol.* **1982**, *77*, 634–643.
 15. Dijkstra, J.R.; Meek, M.F.; Robinson, P.H.; Gramsbergen, A. Methods to evaluate functional nerve recovery in adult rats: Walking track analysis, video analysis and the withdrawal reflex. *J. Neurosci. Methods* **2000**, *96*, 89–96.
 16. Johnston, R.B.; Zachary, L.; Dellon, A.L.; Seiler, W.A.; Teplica, D.M. Improved imaging of rat hind foot prints for walking track analysis. *J. Neurosci. Methods* **1991**, *38*, 111–114.
 17. Hudson, T.; Evans, G.R.D.; Schmidt, C. Engineering strategies for peripheral nerve repair. *Clin. Plast. Surg.* **1999**, *26* (4), 617–628.
 18. Seal, B.L.; Otero, T.C.; Panitch, A. Review: Polymeric biomaterials for tissue and organ regeneration. *Mater. Sci. Eng.* **2001**, *34* (4), 147–230.
 19. Evans, G.R.D. Peripheral nerve injury: A review and approach to tissue engineering constructs. *Anat. Rec.* **2001**, *263*, 396–404.
 20. Danielsen, N.; Dahlin, L.B.; Thomsen, P. Inflammatory cells and mediators in the silicone chamber model for nerve regeneration. *Biomaterials* **1993**, *14* (15), 1180–1185.
 21. Hadlock, T.; Elisseff, J.; Langer, R.; Vacanti, J.; Cheney, M. A tissue-engineered conduit for peripheral nerve repair. *Arch. Otolaryngol. Head Neck Surg.* **1998**, *124* (10), 1081–1086.
 22. Lavik, E.B.; Hrkach, J.S.; Lotan, N.; Nazarov, R.; Langer, R. A simple synthetic route to the formation of a block copolymer of poly(lactic-co-glycolic acid) polylysine for the fabrication of functionalized, degradable structures for biomedical applications. *J. Biomed. Mater. Res.* **2001**, *58* (3), 291–294.
 23. Valentini, R.F.; Sabatini, A.M.; Dario, P.; Aebischer, P. Polymer electret guidance channels enhance peripheral nerve regeneration in mice. *Brain Res.* **1989**, *48*, 300–304.
 24. Langone, F.; Lora, S.; Veronese, F.M.; Caliceti, P.; Parnigotto, P.P.; Valenti, F.; Palma, G. Peripheral nerve repair using a poly(organo)phosphazene tubular prosthesis. *Biomaterials* **1995**, *6* (5), 347–353.
 25. den Dunnen, W.F.; van der Lei, B.; Schakenraad, J.M.; Stokroos, I.; Blaauw, E.; Bartels, H.; Pennings, A.J.; Robinson, P.H. Poly(DL-lactide-epsilon-caprolactone) nerve guides perform better than autologous nerve grafts. *Microsurgery* **1996**, *17* (7), 348–357.
 26. Aebischer, P.; Valentini, R.F.; Dario, P.; Domenici, C.; Galletti, P.M. Piezoelectric guidance channels enhance regeneration in the mouse sciatic nerve after axotomy. *Brain Res.* **1987**, *435*, 165–168.
 27. Molander, H.; Engkvist, O.; Hagglund, J.; Olsson, Y.; Torebjork, E. Nerve repair using a polyglactin tube and nerve graft: An experimental study in the rabbit. *Biomaterials* **1983**, *4*, 276–280.
 28. Nyilas, E.; Chiu, T.H.; Sidman, R.L.; Henry, E.W.; Brushart, T.M.; Dikkes, P.; Madison, R. Peripheral nerve repair with bioresorbable prosthesis. *Trans. Am. Soc. Artif. Int. Organs* **1983**, *29*, 307–313.
 29. Seckel, B.R.; Chiu, T.H.; Nyilas, E.; Sidman, R.L. Nerve regeneration through synthetic biodegradable nerve guides: Regulation by the target organ. *Plast. Reconstr. Surg.* **1984**, *74* (2), 173–181.
 30. Labrador, R.O.; Buti, M.; Navarro, X. Peripheral nerve repair: Role of agarose matrix density on functional recovery. *Neuroreport* **1995**, *6* (15), 2022–2026.
 31. Madison, R.D.; Da Silva, C.F.; Dikkes, P. Peripheral nerve repair: Role of agarose matrix density on functional recovery. *Brain Res.* **1988**, *447* (2), 325–334.
 32. Williams, L.R.; Longo, F.M.; Powell, H.C.; Lundborg, G.; Varon, S. Spatial-temporal progress of peripheral nerve regeneration within a silicone chamber: Parameters for a bioassay. *J. Comp. Neurol.* **1983**, *218*, 460–470.
 33. Li, S.T. A Multi-Layered, Semipermeable Conduit for Nerve Regeneration Comprised of Type I Collagen, Its Method of Manufacture and a Method of Nerve Regeneration Using Said Conduit. U.S. Patent No. 4,963,146.
 34. Archibald, S.J.; Shefner, J.; Krarup, C.; Madison, R.D. Monkey median nerve repaired by nerve graft or collagen nerve guide tube. *J. Neurosci.* **1995**, *15* (5), 4109–4123.
 35. Lundborg, G. A 25-year perspective of peripheral nerve surgery: Evolving neuroscientific concepts and clinical significance. *J. Hand Surg. Am.* **2000**, *25* (3), 391–414.
 36. Lundborg, G.; Rosen, B.; Dahlin, L.; Danielsen, N.; Holmberg, J. Tubular versus conventional repair of median and ulnar nerves in the human forearm: Early results from a prospective, randomized, clinical study. *J. Hand Surg.* **1997**, *22A*, 99–106.
 37. Hall, S.M. Regeneration in cellular and acellular autografts in the peripheral nervous system. *Neuropathol. Appl. Neurobiol.* **1986**, *12* (1), 27–46.
 38. Heath, C.A.; Rutkowski, G.E. The development of bioartificial nerve grafts for peripheral nerve regeneration. *TIBTECH* **1998**, *16*, 163–168.
 39. Young, C.; Miller, E.; Nicklous, D.M.; Hoffman, J.R. Nerve growth factor and neurotrophin-3 affect functional recovery following peripheral nerve injury differently. *Restor. Neurol. Neurosci.* **2001**, *18* (4), 167–175.
 40. Verdu, E.; Labrador, R.O.; Rodriguez, F.J.; Ceballos, D.; Fores, J.; Navarro, X. Alignment of collagen and

- laminin-containing gels improve nerve regeneration within silicone tubes. *Restor. Neurol. Neurosci.* **2002**, *20* (5), 169–180.
41. Ahmed, Z.; Brown, R.A.; Ljungberg, C.; Wiberg, M.; Terenghi, G. Nerve growth factor enhances peripheral nerve regeneration in non-human primates. *Scand. J. Plast. Reconstr. Surg. Hand Surg.* **1999**, *33* (4), 393–401.
 42. Manthorpe, M.; Engvall, E.; Ruoslahti, E.; Longo, F.M.; Davis, G.E.; Varon, S. Laminin promotes neuritic regeneration from cultured peripheral and central neurons. *J. Cell Biol.* **1983**, *97*, 1882–1890.
 43. Ceballos, D.; Navarro, X.; Dubey, N.; Wendelschafer-Crabb, G.; Kennedy, W.R.; Tranquillo, R.T. Magnetically aligned collagen gel filling a collagen nerve guide improves peripheral nerve regeneration. *Exp. Neurol.* **1999**, *158*, 290–300.
 44. Patel, N.; Bhandari, R.; Shakesheff, K.M.; Cannizzaro, S.M.; Davies, M.C.; Langer, R.; Roberts, C.J.; Tendler, S.J.; Williams, P.M. Printing patterns of biospecifically-adsorbed protein. *J. Biomater. Sci., Polym. Ed.* **2000**, *11* (3), 319–331.
 45. Bonner, J.; O'Connor, T.P. The permissive cue laminin is essential for growth cone turning in vivo. *J. Neurosci.* **2001**, *21* (24), 9782–9791.
 46. Hammarback, J.A.; Palm, L.S.; Furcht, L.T.; Letourneau, P.C. Guidance of neurite outgrowth by pathways of substratum-adsorbed laminin. *J. Neurosci. Res.* **1985**, *13*, 213–220.
 47. Clark, P.; Britland, S.; Connolly, P. Growth cone guidance and neuron morphology on micropatterned laminin surfaces. *J. Cell Sci.* **1993**, *105*, 203–212.
 48. Saneinejad, S.; Shoichet, S.M. Patterned glass surfaces direct cell adhesion and process outgrowth of primary neurons of the central nervous system. *J. Biomed. Mater. Res.* **1998**, *42*, 13–19.
 49. Patel, N.; Padera, R.; Giles, S.H.W.; Cannizzaro, S.M.; Davies, M.C.; Langer, R.; Roberts, C.J.; Tendler, S.J.; Williams, P.M.; Shakesheff, K.M. Spatially controlled cell engineering on biodegradable polymer surfaces. *FABES J.* **1998**, *12*, 1447–1454.
 50. Wheeler, B.C.; Corey, J.M.; Brewer, G.J.; Branch, D.W. Microcontact printing for precise control of nerve cell growth in culture. *J. Biomech. Eng.* **1999**, *121*, 73–78.
 51. Cuevas, P.; Carceller, F.; Dujovny, M.; Garcia-Gomez, I.; Cuevas, B.; Gonzalez-Corrochano, R.; Diaz-Gonzalez, D.; Reimers, D. Peripheral nerve regeneration by bone marrow stromal cells. *Neurol. Res.* **2002**, *24* (7), 634–638.
 52. Navarro, J.; Oudrhiri, N.; Fabrega, S.; Lehn, P. Gene delivery systems: Bridging the gap between recombinant viruses and artificial vectors. *Adv. Drug Deliv. Rev.* **1998**, *30*, 5–11.
 53. Gravel, C.; Gotz, R.; Lorrain, A.; Sendtner, M. Adenoviral gene transfer of ciliary neurotrophic factor and brain-derived neurotrophic factor leads to long-term survival of axotomized motor neurons. *Nat. Med.* **1997**, *3*, 765–770.
 54. Baumgartner, B.J.; Shine, H.D. Neuroprotection of spinal motoneurons following targeted transduction with an adenoviral vector carrying the gene for glial cell line-derived neurotrophic factor. *Exp. Neurol.* **1998**, *153*, 102–112.
 55. Gimenez, Y.R.M.; Revah, F.; Pradier, L.; Loquet, I.; Mallet, A.P. Prevention of motoneuron death by adenovirus-mediated neurotrophic factors. *J. Neurosci. Res.* **1997**, *48*, 281–285.
 56. Lemarchand, P.; Jaffe, H.A.; Danel, C.; Cid, M.C.; Kleinman, H.K.; Stratford-Perricaudet, L.D.; Perricaudet, M.; Pavirani, A.; Lecocq, J.; Crystal, R.G. Adenovirus-mediated transfer of a recombinant human α_1 -antitrypsin cDNA to human endothelial cells. *PNAS* **1992**, *89*, 6482–6486.
 57. Harris, J.D.; Lemoine, N.R. Strategies for targeted gene therapy. *Trends Genet.* **1996**, *12*, 400–405.
 58. Miller, N.; Vile, R. Vectors for gene therapy. *FASEB J.* **1995**, 190–199.
 59. Verma, I.M.; Somia, N. Gene therapy—promises, problems and prospects. *Nature* **1997**, *389*, 239–242.
 60. Wim, T.J.; Hermens, M.C.; Verhaagen, J. Viral vectors, tools for gene transfer in the nervous system. *Prog. Neurobiol.* **1998**, *55*, 399–432.
 61. Patrick, C.W., Jr.; Zheng, B.; Wu, X.; Gurtner, G.; Barlow, M.; Koutz, C.; Chang, D.; Schmidt, M.; Evans, G.R. Muristerone A-induced nerve growth factor release from genetically engineered human dermal fibroblasts for peripheral nerve tissue engineering. *Tissue Eng.* **2001**, *7* (3), 303–311.

Tissue Engineering of Rotator Cuff Tendons

Jonathan P. Van Kleunen

McKay Orthopaedic Surgery Research Laboratory, University of Pennsylvania Health System, Philadelphia, Pennsylvania, U.S.A.

Louis J. Soslowsky

Orthopaedic Surgery, University of Pennsylvania Health System, Philadelphia, Pennsylvania, U.S.A.

David L. Glaser

The Shoulder and Elbow Service, University of Pennsylvania Health System, Philadelphia, Pennsylvania, U.S.A.

INTRODUCTION

Tears in the rotator cuff tendons are common and increase in frequency with age. These injuries are a significant cause of debilitation and economic hardship. While some injuries respond to conservative treatment, those that do not improve may require surgical intervention. Recovery from surgery takes considerable time and is complicated by adhesion formation, decreased tendon strength, and compromised function. For these reasons, the rotator cuff could benefit from healing augmentation, improved remodeling, or de novo synthesis of graft material. Tissue engineering seeks to repair or replace a variety of tissues through the delivery of implanted cells, scaffolds, genetic material, or proteins at the time of surgery. Its study combines aspects of both medicine and engineering and focuses upon advancements in cell biology and the development of new biological tissues. The integration of these components leads to the creation of new systems to augment the healing process.

Because tendons serve an important biomechanical role in the musculoskeletal system, the restoration of their mechanical properties is an important consideration in the development of an appropriate tissue engineering model. The concept of “functional tissue engineering” addresses this issue by giving increased attention to the functional properties of a given tissue and the restoration of healing tissue to an acceptable functional level. This concept demands understanding of normal tissue function, limitations of damaged tissue, definitions of a successful repair, and recognition of the appropriate determinants in reaching this goal.

This article focuses on basic principles of functional tissue engineering as it applies to the rotator cuff. We will discuss the requirements for a viable model of functional tissue engineering. As noted, these needs include an understanding of the biomechanical properties of healthy tendon, definitions of a successful repair, and recognition of physical and biological

factors that affect the utility of a potential model. We will review the pathological bases for rotator cuff disorders and how these lead to tendon injury. Several current models of functional tissue engineering will be examined including rehabilitation, cell therapy, scaffolds, and gene therapy. The mechanisms of action of these models will be described in addition to their potential uses. While these studies investigate the components important to a successful functional tissue engineering program, they also demonstrate our limited depth of knowledge regarding the true regulatory mechanisms of the healing process and foreshadow the continued advancements additional study will achieve.

BASIC PRINCIPLES OF TISSUE ENGINEERING DESIGN

The Key to Success

A successful tissue engineering program has several requirements.^[1,2] Before the healing tendon is considered, the biomechanical properties of tissue in vivo and the mechanical parameters that result in failure must be identified. Biomechanical analyses have identified multiple tensile regions in each tendonous unit of the rotator cuff. Principal units have been identified in each component of the rotator cuff. The anterior third of the supraspinatus, midsuperior, and inferior portions of the infraspinatus, and superior and midsuperior portions of the subscapularis tendons have been identified as these primary units (Table 1).^[3-5] Generally, there needs to be an understanding of the basic events in the normal healing of a rotator cuff focusing on identification of therapeutic targets. This includes defining the temporal, spatial, cellular, and molecular events involved with tendon healing as well as the origin of the responding cells. The intra- and extracellular signaling pathways used by these cells need to be elucidated in addition to the regulation of

Table 1 Tensile properties measured in human rotator cuff tendons

Tendon	Cross-sectional area (mm ²)	Ultimate load (N)	Stiffness (N/m)	Ultimate stress (MPa)	Elastic modulus (MPa)
SS anterior third	25.5	411	N.R.	17	N.R.
SS middle third	24.7	153	N.R.	6	N.R.
SS posterior third	21.9	88	N.R.	4	N.R.
IS superior portion	29.0	463	134	16	120
IS midsuperior portion	26.3	677	171	27	156
IS midinferior portion	20.8	315	97	16	111
IS inferior portion	26.2	550	149	22	140
TM	49.0	67	23	1	14
Sub. superior portion	40.3	623 (0°) 478 (60°)	147 (0°) 209 (60°)	N.R.	N.R.
Sub. midsuperior portion	N.R.	706 (0°) 598 (60°)	175 (0°) 182 (60°)	N.R.	N.R.
Sub. midinferior portion	N.R.	454 (0°) 400 (60°)	128 (0°) 130 (60°)	N.R.	N.R.
Sub. inferior portion	27.3	75 (0°) 30 (60°)	27 (0°) 10 (60°)	N.R.	N.R.

Note: Values for the infraspinatus and teres minor tendons are combined measurements for both 0° and 60° of abduction. Values are listed for the subscapularis tendon in both 0° and 60° of abduction. SS = supraspinatus. IS = infraspinatus. TM = teres minor. Sub. = subscapularis. N.R. = not recorded.

(From Refs. [3–5].)

pathway inhibitors. The investigator needs to be able to create the engineered tissue by altering the responding cells, augmenting the healing signals, blocking the inhibitor pathways, or creating de novo tissue in a bioreactor. In the case of true functional tissue engineering, stem cells placed on a scaffold need to be manipulated with the proper signaling molecules in the correct biomechanical environment to form a tendon. Additionally, there needs to be the ability to evaluate the properties of the engineered tissue quantitatively. This assessment demands established standards of successful biomechanical restoration. Prior to application, a potential model must be shown to be effective with the essential characteristics of the human clinical condition.^[2,6]

Defining the Standards of a Rotator Cuff Tendon

Before a tissue engineering model can be created, defining standards for the engineered tissue is an initial step and guides the remaining developmental process. In rotator cuff healing, this need translates to understanding the basic events involved with tendon healing at the muscle–tendon junction, along the tendon, and at the tendon–bone interface, along with the

responding cells, cytokines, chemokines and matrix proteins involved in these activities.

In addition to evaluating microscopic and molecular indicators of healing, standards must be established to characterize a successful repair. The ability of an engineered tissue to withstand forces approximating those experienced during a healthy state helps to define this goal. Biomechanical properties for a given tissue must be understood and prioritized. Stress and strain parameters should be measured in both normal and repaired tissue, and these factors must be studied to elucidate their impact on the healing process. The underlying organizational composition of the tissue principally determines the biomechanical properties in engineered tissue. Load has several effects on the normal biochemical and biomechanical qualities of tendon that vary according to the magnitude and type of force applied and are mediated by several growth factors including platelet-derived growth factor (PDGF) and insulin-like growth factor-1 (IGF-1).^[7]

Understanding the Pathological Basis of Rotator Cuff Disorders

Pathological responses to a loaded environment may be a result of lower metabolic activity in tendon associated with an inability to react to changes in load.

An increase in the proportion of collagen III, increased fibril diameter, reduced collagen turnover, increased mature cross-linking, decreased proteoglycan and water content, and reduced cellularity have been found to occur with aging. These changes produce a less compliant tissue that is more susceptible to injury and inadequately prepared to heal efficiently following pathological events. All of these changes can be addressed with tissue engineering.^[7-9]

Injury may be precipitated by direct events or may be purely a result of chronic degeneration. Initially following insult, the tissue surrounding the tendinous injury undergoes a hemorrhagic response that triggers the inflammatory pathway. Secretion of growth factors, including PDGF, transforming growth factor- α and β (TGF- α , TGF- β), basic fibroblast growth factor (bFGF), and epidermal growth factor (EGF), accompanies this response to direct the attraction and proliferation of fibroblasts and to stimulate collagen and protein synthesis. Fibroblasts play the central role in healing and remodeling tendon at the site of injury. This process takes considerable time, and even at maturation, the biomechanical properties of healed tissue are inferior to those of uninjured tendon due to a complicated system including a proportional decrease in the amount of collagen I and an increase in the amount of collagen III (Fig. 1). Again, these cytokines can be targets for tissue engineering.^[10]

Difficulty in elucidating the cellular pathways of the healing process in tendon has provided an obstacle to developing an ideal system of repair augmentation. While attempts have been made to identify the origin of cells that respond to growth factors secreted in the

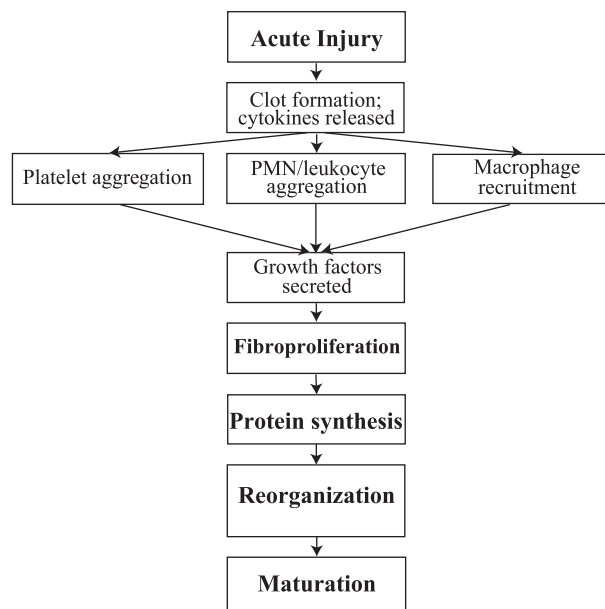


Fig. 1 Sequence of biological and mechanical events in the healing process in tendon following acute injury.

area of repair, both extrinsic and intrinsic sources have been described and no certain answer has been elucidated.^[11,12] We have recently investigated the heritage of cells involved in the healing tendon. The objective of this study was to examine the specific ancestry of cells that participate in tendon healing including myoblasts, activated satellite cells, differentiated skeletal muscle cells, vascular and nonvascular smooth muscle cells, pericytes, endothelial cells of developing and mature blood vessels, and bone marrow-derived cells. This study suggested that smooth muscle cells from mature or immature vessels or from pericytes are a major source of responding cells in the fibroproliferative stage of tendon healing and that cells of endothelial origin respond to a tendon injury by participating in neo-vascularization but do not contribute to fibroproliferation. Despite the hypothesized role satellite cells play in healing, activated satellite cells do not have a major role in healing tendon injuries. Cells of bone marrow origin contributed to the inflammatory process. However, despite recent studies suggesting bone-marrow stem cell recruitment to various sites, in this study bone marrow-derived cells did not contribute to the fibroproliferative response or develop into tendon cells. This study helped to elucidate the origins of cells involved in the various steps of the healing process.^[13,14]

CURRENT DIRECTIONS OF TISSUE ENGINEERING IN THE ROTATOR CUFF

Given the biomechanical demands placed upon tendon and the potential inability of aged or damaged tendons to meet these requirements, an ideal healing system is one that is both capable of restoring damaged tendon to its full functional ability and of promoting healing tissue to a biomechanical stability comparable to its healthy state. The time course of healing in tendon and the inferior tensile properties in healed tendon make the unaided repair process less than ideal. A successful model of tissue engineering for tendon repair in the rotator cuff is one that is capable of augmenting this process through improved remodeling and characterized by a faster return to near-normal biomechanical stability.

In addition to facilitating repair within the tendon substance, a successful model of tissue engineering must be able to achieve functional recovery along the tendon-bone interface. As the tissue properties change from soft, compressible material to hard, incompressible material, the tendon-bone interface is expected to be the most common site of failure due to stress concentrations.^[15] This interface has been categorized into four histological zones. The continuous change in composition across these zones helps to maximize the efficiency of load transfer across the tendon-bone

interface.^[16] We recently studied the genetic expression of extracellular matrix components and histological changes in the insertion site of a rat supraspinatus tendon that was acutely detached and repaired. Results indicated that a dramatic healing response occurred at the site of injury but that the healing tissue was poorly organized and was ineffective in recreating the original insertion site.^[15] A successful model of functional tissue engineering must be able to maximize the efficiency of healing along the transition from tendon to bone and promote an organized healing response. For this process to occur, a model of tissue engineering must be successfully integrated into the healing region and must help to guide healing progress. The continuous transition in tissue zones found in the healthy tendon–bone interface must be recreated to optimize mechanical function. Biomechanical properties must be restored to near normal levels to ensure interface stability. The healing process must occur in a timely, organized process to optimize early stability and tissue recovery time. These issues within both the tendon substance and the tendon–bone interface must be addressed in the development of a model of functional tissue engineering.

To achieve these goals, attempts at tissue engineering in rotator cuff tendons have taken several directions. Several of these studies have demonstrated promise as potential means of augmenting the healing response in damaged tendon. Unfortunately, no studies have combined all of these techniques into one system. In this section, we will describe several models of tissue engineering that have demonstrated improvements in tendon healing. Rehabilitation is a mechanical therapy that aims to improve tendon healing through the application of regulated mechanical stresses on the healing tendon. Cell therapy focuses upon the implantation of stem cells, or other pluripotent cells, into a site of tendon healing. These cells may serve a role in regenerating tissue and producing growth factors important in the healing process. Small intestine submucosa is an acellular biological scaffold that has come under recent examination as an organized matrix for the direction of connective tissue healing. Gene therapy is a broad field of study that is based upon the incorporation of genetic material into host tissue to augment the healing process from a genetic biochemical level. The proposed mechanisms by which each of these models serves to augment tendon healing will be discussed in addition to their appropriate roles in the repair of the rotator cuff.

Rehabilitation

Mechanical techniques focus upon the physical manipulation of repaired tendon to meet its original

needs. Because the primary functions of the rotator cuff and other tendons are as load-bearing constructs, physical rehabilitation of tendons focuses on placing healing tendon in a similar mechanically loaded environment under controlled conditions. While adequate rest following rotator cuff repair is necessary for healing to occur and recent studies have suggested that rest can lead to improved organization within tendon, extended rest may lead to shoulder stiffness, muscle atrophy, and disorganized healing.^[7] Controlled rehabilitation accentuates tendon healing in an organized manner.

Several studies have demonstrated that postrepair continuous passive motion and physical therapy increase tendon mobility and stiffness and help to prevent postoperative adhesion formation.^[17–20] In addition, several in-vivo studies have shown that increased stress in the region of repair following surgery, such as that seen in physical therapy, is associated with a greater organization of collagen fibrils and a higher rate of matrix synthesis compared to controls for both tendon and ligament.^[21,22] While this activity demonstrated a significant improvement in healing compared to control specimens with rotator cuff injuries, healing tissue was still found to be biomechanically inferior to uninjured tendon to at least 16 weeks following injury.^[21] These studies demonstrate the potential of mechanical rehabilitation to improve healing in damaged and repaired tendon and define the inferiority of healing tissue compared to healthy tendon during periods of rehabilitation.

Cell Therapy

Mesenchymal stem cells (MSCs) and other stem cells are pluripotent cells isolated from various sources that have not yet differentiated into a specific cell line. These cells have the potential to develop into a wide variety of tissues including bone, cartilage, tendon, fat, marrow stromal cells, and muscle. Studies examining the transplantation of allogeneic and xenogeneic stem cells into human tissue both in vitro and in vivo have shown that these cells maintain their pluripotency following transplantation and undergo site-specific differentiation.^[23–25] In this manner stem cells transplanted into a region of damaged tendon may differentiate into endothelial cells, osteoblasts, and fibroblasts to act as a biological patch and to augment the healing process further through the increased production of collagen and proteoglycan.^[10]

While the potential for improved healing is substantial in this model of tissue engineering, many challenges exist that must be met in order to develop a functional application of cell therapy to tendon repair. An adequate source of stem cells must exist for a

substantial transplantation of stem cells to be a reality. Autogenous donation is ideal in terms of delivering an immunologically low-risk population of cells, but the collection process may incur other morbidities. Allogenic and xenogenic sources may be easier to develop but carry the risk of potential rejection or disease transmission.^[26] Once a reliable population of stem cells is isolated, a delivery vehicle must be selected that will protect the cells but will also allow them to function normally following transplantation. Once both of these needs are met, a technique must be developed to combine the stem cells with the carrier *in vitro* and deliver them *in vivo*. This composite must be provided in such a way that makes its use feasible, safe, and efficacious.^[26]

Early studies in animal models and in humans have demonstrated successful implementation, survival, and differentiation of MSCs in tendon, but more work is required to elucidate fully the recipient response to donor cells and to determine optimal conditions for successful implementation of the technique.^[27–29] Standardization of a safe, effective delivery system will facilitate the potential use of cell therapy as an adjunct to tendon repair.

Small Intestinal Submucosa

Various methods of soft tissue grafting, including tendon and fascial autografts, allografts, xenografts, and synthetic constructs, have been examined in the repair of large tendon and ligament defects. Results from these studies have been inconsistent and complicated by adhesion formation, infection, functional failure, and immune rejection.^[30,31] Porcine small intestinal submucosa (SIS) is highly collagenous biological scaffold that has been found to promote a site-specific reconstructive response in connective tissue. This healing process is characterized by active angiogenesis and host cell migration and differentiation. The resulting new soft tissue is more representative of the surrounding host tissue than typical fibrous scar tissue.^[30]

Several studies have demonstrated the utility of SIS for tissue engineering in the repair of tendinous, ligamentous, and fascial defects in animal models.^[30–34] Examination of the use of SIS in grafts of rotator cuff defects has shown promise for the use of SIS grafting as a successful means of repairing large tendon defects.^[30] The assessment of healing at 6 months after the surgical repair of infraspinatus tendon defects using SIS grafts indicated a similar tendon size compared to sham-operated specimens and exhibited extensive signs of remodeling without adhesion formation. Histologically, the tendon–bone insertion site of SIS-regenerated tendons was similar to healthy

controls. While mechanical testing indicated that SIS-regenerated tendons were able to withstand lower failure loads than healthy tendon, these loads were similar to those in sham-operated tendons. This study supports the use of SIS in rotator cuff repair because it shows some improvements over conventional repair techniques. Of particular interest is the potential use of SIS in large gap defects precluding normal repair because it may provide the necessary scaffold to support tendon healing across extended distances.^[30]

The effectiveness of SIS appears to be determined by multiple microenvironmental factors, including vascular supply, oxygen availability, pH, and mechanical factors. While the development of new tissue along the SIS scaffold appears to depend, in part, on the quality of surrounding connective tissue in the site of a graft, stress along the grafted SIS is also important to the induction of regenerative processes. Progression of cellular infiltration and organized remodeling has been found to be greater in tendons that are mobilized postoperatively compared to those that are immobilized.^[34] This behavior is particularly advantageous for use in load-bearing tissues such as tendon. SIS is a promising model for augmented tendon healing because of its ability to serve as a scaffold for organized remodeling in damaged tissue with minimal scar formation.

Gene Transfer

While MSC transplantation and SIS grafting are models of tissue engineering that act by supplying adjunct cells or a scaffold to augment the healing progress, gene transfer therapy is a model that seeks to directly induce greater healing ability in the cells already in existence around a site of injury. Several growth factors have been implicated in the healing response of tendon including TGF β , TGF- α , bFGF, PDGF, IGF-I, and EGF.^[7,35–38] The goals of gene transfer, in regard to tendon healing, are to up-regulate the synthesis of these growth factors and to suppress production of endogenous proteins that may inhibit efficient, organized remodeling of damaged tendon.

In order for transfer to occur to host cells, an appropriate vector must be developed that will successfully deliver genetic material to the desired target and allow implementation into the host's genome. Viruses, liposomes, and gene guns are all methods of delivery that have demonstrated potential use for this purpose.^[10] Successful delivery and implementation of genetic material into the host's genome will lead to synthesis of the proteins encoded by the delivered gene or to down-regulation in the production of an

unwanted protein from a separate path. The advantages of this model of tissue engineering are that it induces a significant, local production of a substance that augments healing and that it is not complicated by morbidities associated with the insertion of donor tissue. The disadvantages are those most commonly associated with the vectors used for gene delivery. The production of certain viral vectors is a difficult process with a frequently low yield. Several viral strains carry a risk of mutagenesis that may interfere with the intended genetic result. Some strains of virus, adenovirus being the most notable, are capable of stimulating an inflammatory response and secondarily immunogenic rejection in the host. Gene guns, which involve particle bombardment of genetic material into host tissue, require specialized equipment that significantly increases the costs of production and utilization. Liposomes and naked DNA, while less immunogenic than other vectors, are not as successful as other methods in transfecting host cells. Because of these limitations, vector development remains an extremely active component of gene transfer research.^[39]

While the use of gene transfer is in its early stages, some studies have demonstrated successful incorporation of marker genes into tendon and ligament.^[40,41] Promising studies have explored the transfer of antagonists to FAK, a protein kinase linked to adhesion formation, and of genes encoding PDGF.^[10,39] In our lab, we have used both plasmid DNA and adenovirus-mediated gene transfer to demonstrate successfully reporter gene expression at the site of rotator cuff injury (Fig. 2). The success of these studies shows the potential for the use of gene transfer in the tendon. Future studies will continue to examine the insertion

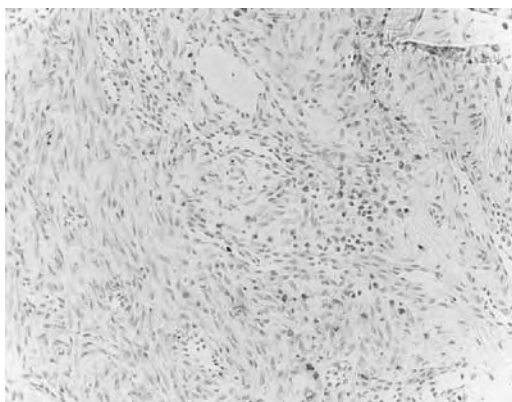


Fig. 2 Rat rotator cuff tendon at six days after transfer of a Luciferase reporter gene. A collagen carrier mixed with plasmid DNA was placed in a 2 mm defect in the rat supraspinatus tendon. Luciferase expression is demonstrated by brown staining on histological examination (200 ×).

of genes to augment the production of growth factors involved in the healing processes of tendon or of antagonists to inhibit the production of proteins found to interfere with organized remodeling.

CONCLUSIONS

As described above, the goals of tissue engineering are to increase the healing response of the body to insults in order to regenerate tissues efficiently and to produce new tissues that are comparable to healthy tissue in strength and function. In the case of tendon, this goal translates to a restoration of load-bearing function with a decreased risk of recurrent failure. The rotator cuff is a musculotendonous complex that is likely to benefit from the development of tissue engineering models for tendon repair.

Several techniques have been utilized as parts of tissue engineering strategies ranging from mechanical techniques to stem cell-based gene therapy. Mechanical techniques, such as immobilization and physical rehabilitation, seek to augment tendon healing by exposing healing tendon to low-level mechanical forces in a controlled environment. The use of MSCs is a cellular model that uses the incorporation of pluripotent cells into the healing tendon. These cells are capable of differentiating into cells resembling their surrounding environment. They help regenerate repaired tendon in a more organized manner closely resembling the original type of tissue. SIS grafts are biological scaffolds that augment repair organization and mechanical stability. Gene transfer techniques involve transfecting cells in the region of repair to increase the production of certain prohealing growth factors and decrease the production of antagonistic proteins. These methods augment the healing process in existing cells without incorporating new tissue into the repair site. While each of these techniques has limitations, they demonstrate promise for the development of repair methods and therapies to reduce the complications and morbidities of rotator cuff repair significantly. Such a development would improve the quality of life in postsurgical patients and lessen the limitations of postoperative function and economic hardship. Continued support for the study of tendon tissue engineering is vital to the achievement of these goals.

ARTICLES OF FURTHER INTEREST

Extracellular Matrix Scaffolds; Gene Therapy; Stem Cells; Tissue Engineering of Bone; Tissue Engineering of Tendons

REFERENCES

1. Carpenter, J.E.; Thomopoulos, S.; Flanagan, C.L.; DeBano, C.M.; Soslowky, L.J. Rotator cuff defect healing: A biomechanical and histologic analysis in an animal model. *J. Shoulder Elb. Surg.* **1998**, *7*, 599–605.
2. Butler, D.L.; Goldstein, S.A.; Guilak, F. Functional tissue engineering: The role of biomechanics. *J. Biomech. Eng.* **2000**, *122*, 570–575.
3. Itoi, E.; Berglund, L.J.; Grabowski, J.J.; Schultz, F.M.; Growney, E.S.; Morrey, B.F.; An, K.N. Tensile properties of the supraspinatus tendon. *J. Orthop. Res.* **1995**, *13*, 578–584.
4. Halder, A.; Zobitz, M.E.; Schultz, F.; An, K.N. Mechanical properties of the posterior rotator cuff. *Clin. Biomech.* **2000**, *15*, 456–462.
5. Halder, A.; Zobitz, M.E.; Schultz, F.; An, K.N. Structural properties of the subscapularis tendon. *J. Orthop. Res.* **2000**, *18*, 829–834.
6. Guilak, F. Functional tissue engineering: The role of biomechanics in reparative medicine. *Ann. N.Y. Acad. Sci.* **2002**, *961*, 193–195.
7. Hyman, J.; Rodeo, S.A. Injury and repair of tendons and ligaments. *Phys. Med. Rehabil. Clin. North Am.* **2000**, *11* (2), 267–288.
8. Bank, R.A.; TeKoppele, J.M.; Oostingh, G.; Hazleman, B.L.; Riley, G.P. Lysylhydroxylation and non-reducible crosslinking of human supraspinatus tendon collagen: Changes with age in chronic rotator cuff tendonitis. *Ann. Rheum. Dis.* **1999**, *58* (1), 35–41.
9. Riley, G.P.; Harrall, R.L.; Constant, C.R.; Chard, M.D.; Cawston, T.E.; Hazleman, B.L. Tendon degeneration and chronic shoulder pain: Changes in the collagen composition of the human rotator cuff tendons in rotator cuff tendonitis. *Ann. Rheum. Dis.* **1994**, *53*, 359–366.
10. Woo, S.L.Y.; Hildebrand, K.; Watanabe, N.; Fenwick, J.A.; Papageorgiou, C.D.; Wang, J.H.C. Tissue engineering of ligament and tendon healing. *Clin. Orthop.* **1999**, *367S*, S312–S323.
11. Potenza, A.D. Tendon healing within the flexor digital sheath in the dog: An experimental study. *J. Bone Jt. Surg. [Am.]* **1962**, *44*, 49–64.
12. Gelberman, R.H.; Manske, P.R.; Akeson, W.H.; Woo, S.L.Y.; Lundborg, G.; Amiel, D. Flexor tendon repair. *J. Orthop. Res.* **1986**, *4*, 119–128.
13. Glaser, D.L.; Ramachandran, R.; Shore, E.M.; Lin, T.W.; Yeh, P.C.; Beredjikian, P.; Kaplan, F.S.; Goldhammer, D.J.; Soslowky, L.J. The Origin of Cells Within a Healing Tendon. In *Transactions of the Orthopaedic Research Society*, 49th Annual Meeting, New Orleans, Feb. 1–5, 2003.
14. Lin, T.W.; Cardenas, L.; Glaser, D.L.; Soslowky, L.J. Tendon Healing in Interleukin-6 Knockout Mice. In *Transactions of the Orthopaedic Research Society*, 49th Annual Meeting, New Orleans, Feb. 1–5, 2003.
15. Thomopoulos, S.; Hattersley, G.; Rosen, V.; Mertens, M.; Galatz, L.; Williams, G.R.; Soslowky, L.J. The localized expression of extracellular matrix components in healing tendon insertion sites: An in situ hybridization study. *J. Orthop. Res.* **2002**, *20* (3), 454–463.
16. Woo, S.L.Y.; Maynard, J.; Butler, D.; Lyon, R.; Torzilli, P.; Akeson, W.; Cooper, R.; Oakes, B. Ligament, Tendon, and Joint Capsule Insertions to Bone. In *Injury and Repair of the Musculoskeletal Soft Tissues*; Woo, S.L.Y., Buckwalter, J.A., Eds.; American Academy of Orthopaedic Surgeons: Savannah, 1987; 129–166.
17. Gelberman, R.H.; Nunley, J.A., II; Osterman, A.L.; Breen, T.F.; Dimick, M.P.; Woo, S.L. Influences of the protected passive mobilization interval on flexor tendon healing. A prospective randomized clinical trial. *Clin. Orthop.* **1991**, *264*, 189–196.
18. Reichmister, J.P.; Friedman, S.L. Long-term functional results after manipulation of the frozen shoulder. *Md. Med. J.* **1999**, *48* (1), 7–11.
19. Strickland, J.W. Development of flexor tendon surgery: Twenty-five years of progress. *J. Hand Surg. [Am.]* **2000**, *25*, 214–235.
20. Lastayo, P.C.; Wright, T.; Jaffe, R.; Hartzel, J. Continuous passive motion after repair of the rotator cuff. A prospective outcome study. *J. Bone Jt. Surg. [Am.]* **1998**, *80* (7), 1002–1011.
21. Thomopoulos, S.; Williams, G.R.; Soslowky, L.J. Tendon to bone healing: Differences in biomechanical, structural, and compositional properties due to a range of activity levels. *J. Biomech. Eng.* **2003**, *125* (1), 106–113.
22. Gomez, M.A.; Woo, S.L.; Amiel, D.; Harwood, F.; Kitabayashi, L.; Matyas, J.R. The effects of increased tension on healing medial collateral ligaments. *Am. J. Sports Med.* **1991**, *19* (4), 347–354.
23. Liechty, K.W.; MacKenzie, T.C.; Shaaban, A.F.; Radu, A.; Moseley, A.B.; Deans, R.; Marshak, D.R.; Flake, A.W. Human mesenchymal stem cells engraft and demonstrate site-specific differentiation after in utero transplantation in sheep. *Nat. Med.* **2000**, *6* (11), 1282–1286.
24. Prockop, D.J. Marrow stromal cells as stem cells for nonhematopoietic tissues. *Science* **1997**, *276*, 71–74.
25. Lazarus, H.M.; Haynesworth, S.E.; Gerson, S.L.; Rosenthal, N.S.; Caplan, A.I. Ex vivo expansion and subsequent infusion of human bone marrow-derived stromal progenitor cells (mesenchymal progenitor cells): Implication for therapeutic use. *Bone Marrow Transplant.* **1995**, *16*, 557–564.
26. Butler, D.L.; Awad, H.A. Perspectives on cell and collagen composites for tendon repair. *Clin. Orthop.* **1999**, *367S*, S324–S332.
27. Awad, H.A.; Butler, D.L.; Boivin, G.P.; Smith, F.N.; Malaviya, P.; Huijbregtse, B.; Caplan, A.I. Autologous mesenchymal stem cell-mediated repair of tendon. *Tissue Eng.* **1999**, *5*, 267–277.
28. Young, R.G.; Butler, D.L.; Weber, W.; Caplan, A.I.; Gordon, S.L.; Fink, D.J. Use of mesenchymal stem cells in a collagen matrix for Achilles tendon repair. *J. Orthop. Res.* **1998**, *16*, 406–413.
29. Hildebrand, K.A.; Jia, F.; Woo, S.L. Response of donor and recipient cells after transplantation of cells to the ligament and tendon. *Microsc. Res. Tech.* **2002**, *58*, 34–38.
30. DeJardin, L.M.; Arnoczky, S.P.; Ewers, B.J.; Haut, R.C.; Clarke, R.B. Tissue-engineered rotator cuff

- tendon using porcine small intestine submucosa: Histologic and mechanical evaluation in dogs. *Am. J. Sports Med.* **2001**, *29* (2), 175–184.
31. DeJardin, L.M.; Arnoczky, S.P.; Clarke, R.B. Use of small intestinal submucosal implants for regeneration of large fascial defects: An experimental study in dogs. *J. Biomed. Mater. Res.* **1999**, *46*, 203–211.
 32. Aiken, S.W.; Badylak, S.F.; Toombs, J.P.; Shelbourne, K.D.; Hiles, M.C.; Lantz, G.C.; van Sickle, D. Small intestinal submucosa as an intra-articular ligamentous graft material: A pilot study in dogs. *Vet. Comp. Orthop. Traumatol.* **1994**, *7*, 124–128.
 33. Badylak, S.F.; Tullius, R.; Kokini, K.; Shelbourne, K.D.; Klootwyk, T.; Voytik, S.L.; Kraine, M.R.; Simmons, C. The use of xenogeneic small intestinal submucosa as a biomaterial for Achilles tendon repair in a dog model. *J. Biomed. Mater. Res.* **1995**, *29*, 977–985.
 34. Hodde, J.P.; Badylak, S.F.; Shelbourne, K.D. The effect of range of motion on remodeling of small intestinal submucosa (SIS) when used as an Achilles tendon repair material in the rabbit. *Tissue Eng.* **1997**, *3* (1), 27–37.
 35. Chang, J.; Most, D.; Stelnicki, E.; Siebert, J.W.; Longaker, M.T.; Hui, K.; Lineweaver, W.C. Gene expression of transforming growth factor beta-1 in rabbit zone II flexor tendon wound healing: Evidence for dual mechanisms of repair. *Plast. Reconstr. Surg.* **1997**, *100* (4), 937–944.
 36. Abrahamsson, S.O.; Lundborg, G.; Lohmander, L.S. Recombinant human insulin-like growth factor-I stimulates in vitro matrix synthesis and cell proliferation in rabbit flexor tendon. *J. Orthop. Res.* **1991**, *9*, 495–502.
 37. Duffy, F.J., Jr.; Seiler, J.G.; Gelberman, R.H.; Hergrueter, C.A. Growth factors and canine flexor tendon healing: Initial studies in uninjured and repair models. *J. Hand Surg. [Am.]* **1995**, *20*, 645–649.
 38. Stein, L.E. Effects of serum, fibroblast growth factor, and platelet-derived growth factor on explants of rat tail tendon: A morphological study. *Acta Anat.* **1985**, *123*, 247–252.
 39. Evans, C.H.; Robbins, P.D. Genetically augmented tissue engineering of the musculoskeletal system. *Clin. Orthop.* **1999**, *367S*, S410–S418.
 40. Gerich, T.G.; Kang, R.; Fu, F.H.; Robbins, P.D.; Evans, C.H. Gene transfer to the rabbit patellar tendon: Potential for genetic enhancement of tendon and ligament healing. *Gene Ther.* **1996**, *3*, 1089–1093.
 41. Lou, J.; Manske, P.R.; Aoki, M.; Joyce, M.E. Adenovirus-mediated gene transfer into tendon and tendon sheath. *J. Orthop. Res.* **1996**, *14*, 513–517.

Tissue Engineering Scaffolds

Eugene D. Boland

Department of Biomedical Engineering, Virginia Commonwealth University, Richmond, Virginia, U.S.A.

Paul G. Espy

Division of Urology, Virginia Commonwealth University, Richmond, Virginia, U.S.A.

Gary L. Bowlin

Department of Biomedical Engineering, Virginia Commonwealth University, Richmond, Virginia, U.S.A.

INTRODUCTION

Tissue engineering is an interdisciplinary field that blends classical engineering and the life sciences. The field, by name, is still in its infancy, yet the principles are as old as interventional surgery. The desire to restore, maintain, or improve tissue function remains the end goal, but a new set of tools and techniques are emerging. The study of structure–function relationships in both normal and pathological tissues has been coupled with the development of biologically active substitutes or engineered materials. These advances have led to a variety of approaches for the development of engineered tissue analogues. The common trait to most of these strategies begins with the creation of a porous scaffold. The focus of this article will be the processes used to create tissue engineering scaffolds that are capable of mimicking native tissue.

OVERVIEW

The synthetic scaffold is used by the tissue engineer as an analogue of the extra-cellular matrix (ECM). Ideally, the synthetic scaffold should mimic both the structural and functional profile of the native ECM. One such approach is to create a scaffold with large pores to allow cell penetration and remodelling. Additionally, scientists attempt to reproduce either the overall appearance of the ECM or mimic the materials found in the ECM. Some fabrication techniques allow for the combination of multiple approaches. Regardless of approach, a biomimicking ECM analogue is critical to successfully mimic cell–matrix interactions and induce the expression of normal cellular phenotypes.

In order to meet this criterion, tissue engineering scaffolds must be designed to conform to a specific set of requirements. The first requirement is that the material must be biocompatible and function without interrupting other physiological processes. This functionality includes the ability to promote normal cell

growth and differentiation while maintaining three-dimensional orientation/space for the cells. Secondly, the scaffold should not promote or initiate any adverse tissue reactions.^[2] In addition, for clinical and commercial success, scaffold production should be simple yet versatile enough to produce a wide array of configurations to accommodate the size, shape, strength, and other intricacies of the target tissue.^[1–7] Once implemented, *in vitro* or *in vivo*, the material should either be removed from the site through degradation and absorption or incorporated through native remodeling mechanisms that leave only native tissue.

Beyond these generalized requirements, the interactions of single native cells within the immediate surroundings of those cells must also be analyzed. Cells are not autonomous entities kept in proximity to each other by the ECM nor do cells “communicate” only with other cells. Cells and the ECM have a dynamic three-dimensional interrelationship that is kept in balance and influenced by both internal and external stimuli. Evidence indicates that part of this bidirectional communication is a result of direct mechanical connections between fibrous proteins in the ECM and the cells’ cytoskeleton via integrins.^[8] This model requires that any scaffold material must be able to interact with cells in three dimensions and facilitate communication. The goal is to create a scaffold that will duplicate all the essential intercellular reactions and promote native intracellular responses. The incorporation of tissue-inductive factors into these scaffolds may also be useful to facilitate this communication.

After materials, size is the most critical factor when designing a tissue engineering scaffold. These scaffolds must allow cell penetration and neovascularization in order to be viable. The most commonly held paradigm requires pores larger than 10 microns, and in some cases larger than 100 microns, to allow cellular ingrowth and vessel formation.^[5] Observations of native tissues reveal that the structural ECM protein fibers are one to two orders of magnitude smaller than the cell itself, thus, allowing the cell to be in direct contact

with many of the fibers defining its three-dimensional orientation.^[9] Resolution of pore size, total porosity, and ECM fiber size may be a crucial factor in determining the success or failure of a tissue engineering scaffold.

SCAFFOLD FABRICATION TECHNIQUES

With the task of producing a scaffold that meets the above conditions, scientists have employed numerous fabrication techniques that include traditional engineering methods such as casting, molding, and extrusion; textile technologies such as spinning and weaving; decellularization of donor tissue; and engineering rapid prototyping technologies. All these methods have the common goal of producing highly porous scaffolds with large surface areas. Each technique has its own strengths and weaknesses. The aim of this article is to serve as a compilation and review of these methods in a concise format.

Fiber Bonding

Developed in the early 1990s, fiber bonding is a technique that develops tissue scaffolds from a mesh of polymer fibers. This process leads to the formation of highly porous scaffolds with a large ratio of surface area to volume. Two methods to accomplish this have been described. In the first method, a poly(L-lactic acid) (PLLA)/methylene chloride solution is cast into a glass Petri dish that contains a nonwoven poly(glycolic acid) (PGA) fiber mesh. The methylene chloride is then removed through evaporation followed by 24 hours of vacuum drying. This process leaves the PGA fiber mesh intact, but coats the PGA fibers with the solvent cast PLLA. At this point, the nonbonded PGA/PLLA composite is heated beyond the melting points of these polymers. PGA fibers bond at their cross-points, and it is thought that the PLLA matrix preserves the PGA fiber structure and three-dimensional shape during heating. Scaffolds are then quenched and any remaining water vapors are removed through vacuum drying. The final step in the process is the removal of PLLA by dissolving the composite structure in methylene chloride, again leaving the PGA intact, followed by vacuum drying. Scaffolds prepared in this manner were found to have a porosity of 81%, with pore diameters up to 500 μm .^[10]

In a second processing method, described as spray casting, PLLA or poly(lactide-co-galactide) (50/50 PLGA) is dissolved in chloroform and sprayed over the surface of a tube-shaped PGA fiber mesh using a dental atomizer. Chloroform is a poor solvent for PGA and, therefore, has minimal effect on its fiber structure.

A lyophilizer is used to remove the solvent, leaving PGA fibers that are bonded at their cross-points by PLLA or PLGA. Areas of bonding occur only on the external surface of the scaffold, with fibers deep in the scaffold remaining unchanged. The PLLA or PLGA film that develops on the surface of the scaffold appears to be dependent on the concentration of the polymer solution. The surface properties of this scaffold are dictated by PLLA or PLGA, while the mechanical properties are dictated by the PGA mesh. These scaffolds show an interconnected pore structure, with pore diameters as large as 500 μm . Scaffolds demonstrate a greater resistance to compressive forces when bonded with PLLA compared to PLGA.^[11]

Fiber bonding techniques are particularly useful for creating scaffolds with interconnected pore networks and high porosities. The fiber structures also provide a large surface area-to-volume ratio. Some limitations are that pore sizes cannot be controlled easily and both methods require the use of potentially toxic solvents. Polymer selection is limited to a few polymers whose physical properties meet the requirements necessary for fiber bonding conditions. These strategies have produced flat and tubular scaffolds, but do not permit the creation of complex three-dimensional structures. Furthermore, high temperatures and organic solvents preclude the incorporation of bioactive molecules into the scaffold.

Solvent Casting and Particulate Leaching

Solvent casting and particulate leaching are based on the principle that porogens (most commonly salt particles) can be dispersed into a polymer solution and later dissolved, creating a highly porous scaffold, also known as a foam. Several different processing techniques have evolved. In one method, PLLA or PLGA is dissolved in methylene chloride or chloroform. Salt particles are added to the solution, and the solution is cast into a glass dish. The solvent is allowed to evaporate, and the scaffold is subsequently vacuum dried. Salt particles are then leached from the scaffold by immersion in water. Scaffolds are then dried, and the result is a highly porous, highly crystalline structure. The scaffold surface in contact with the glass dish typically has a different morphology than the surface that is open to air.^[12] Alternatively, before salt particles are leached, scaffolds can be subjected to a heating and compression molding process, followed by cooling in water and vacuum drying.^[13] These methods have produced scaffolds with 500 μm diameter pores and porosities of over 90% (see Fig. 1). Waxy hydrocarbons have also been utilized as porogens, and highly porous scaffolds as thick as 2.5 cm were constructed.^[14] Pentanes or hexanes, which could leave

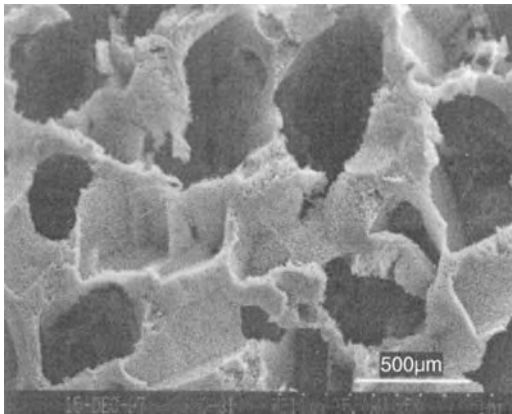


Fig. 1 Poly (L-lactide) tissue engineering scaffold formed through a combined gelation and sugar-leaching technique. The typical pore diameter is approximately 500 microns. (Reprinted with permission of Elsevier Science [Ref. [45] p. 723].)

toxic residues, are needed to remove waxy hydrocarbons from a scaffold. Poly(ethylene glycol) (PEG) can be blended with PLLA or PLGA, resulting in a more pliable scaffold, while maintaining high porosity and pore diameters over 100 μm .^[15] With all of these techniques, the pore diameter is dependent on the size of the porogen, and porosity is dependent on the porogen concentration.

Solvent casting and particulate leaching overcome one of the difficulties of fiber bonding in that they allow more precise control over pore size, total porosity, and surface-to-volume ratio. Highly porous structures with large pore diameters can be readily manufactured. In addition, processing parameters control the crystallinity of the scaffold. Like fiber bonding, this method requires potentially toxic solvents. Also, there are some concerns regarding irregularly shaped pores that may not be completely interconnected or uniformly distributed throughout the scaffold, due in part to evidence suggesting incomplete leaching of the porogens. Finally, as in fiber bonding, rigorous processing conditions generally prevent the inclusion of bioactive molecules.

Membrane Lamination

One of the limitations of the aforementioned scaffold processing methods is the inability to create complex three-dimensional structures. Membrane lamination was developed in an effort to overcome this obstacle while maintaining high porosity.^[16] First, a conceptual contour plot is created by dividing a three-dimensional shape of interest into cross sections of equal thickness. Each layer (or slice) has a specific contour according to

its location within the structure. Multiple scaffolds (membranes) are produced from PLLA or PLGA using a solvent casting and particulate leaching process. Membranes are then cut into unique shapes corresponding to each contour template. The membranes are sequentially bonded together at their adjoining surfaces using a small amount of chloroform, eventually forming a three-dimensional scaffold. Subsequent analysis demonstrates that pore morphology and average pore diameter are unchanged by the lamination process. In addition, the boundaries between individual layers cannot be distinguished with scanning electron microscopy. Membrane lamination is capable of producing scaffolds with pore diameters over 150 μm and porosities up to 90%.^[16]

An obvious advantage of this technique is the ability to create a complex three-dimensional scaffold. Layered membranes retain important physical properties of individual membranes, such as porosity and pore diameter. Drawbacks of membrane lamination are similar to those of solvent casting and particulate leaching, notably the use of toxic solvents. Also, there is no mechanism to align pores between layers; consequently, most channels are not connected through the bulk of the scaffold.

Melt Molding

Melt molding is another method that uses the principles of solvent casting and particulate leaching to manufacture complex three-dimensional scaffolds. A polymer powder is mixed with gelatin microspheres, placed into a Teflon mold, and heated beyond the polymer glass transition temperature. PLGA is generally preferred over PLLA or PGA because it can be processed at lower temperatures. Elevated temperatures can preclude the incorporation of bioactive molecules and may decrease the gelatin solubility. After heating, the composite structure is placed in water, which leaches out the water-soluble microspheres.^[17] This technique is similar to solvent casting and particulate leaching in that pore size and porosity are determined by porogen diameter and concentration, respectively. Also, like other leaching methods, pores do not have uniform diameters and incomplete porogen removal is probable. In contrast, melt molding does avoid the use of toxic solvents.^[17]

One of the highlights of melt molding is that the scaffold shape is limited only by the degree to which the Teflon mold may be shaped. Another improvement over previously discussed processing methods is the lower temperature requirement. This fact, along with the absence of toxic solvents, creates more favorable conditions for the incorporation of bioactive molecules

into scaffolds. One restriction is that this technique is typically limited to PLGA.

Extrusion

Extrusion is a technique for producing tubular scaffolds and was originally utilized in tissue engineering to facilitate guided regeneration of nervous tissue.^[18] The process begins by making PLLA or PLGA wafers by solvent casting and particulate leaching with salt particles. Dried polymer/salt wafers are placed in a piston-extrusion tool, which is mounted on a hydraulic press. The tool is heated, pressure is applied between the nozzle and the piston, and a composite structure is extruded in a tubular conformation. Tubes are then cooled, immersed in water to leach the salt, and, lastly, vacuum dried. The pressure required for extrusion is dependent on the extrusion temperature. Pores up to 300 μm in diameter and porosities of 83% are possible. Both the pore size and porosity increases as the salt weight fraction increases. Using larger salt particles also increases pore diameter. Pore size is unaffected by the extrusion process for temperatures less than 275°C.^[18]

Scaffolds created by extrusion have large pore diameters and high porosity, and the process allows for good control over these parameters. It is thought that these scaffolds may provide an opportunity for the incorporation of microspheres carrying tissue-inductive factors if they can be shielded from the high processing temperatures or the organic solvents. The use of toxic solvents may be a limiting factor for both the bulk material and any embedded microspheres. Other drawbacks associated with solvent casting and particulate leaching are mentioned in previous sections.

Gas Foaming

Gas foaming is another method for producing porous scaffolds without the use of organic solvents. This technique is based on the ability of carbon dioxide (CO_2) molecules to cluster and form nuclei (gas nucleation). There are a number of different approaches to manufacturing gas foams. With the first method, PLGA pellets are put through heating and compression molding, which yields solid polymer disks.^[19] These disks are saturated with CO_2 over a 72-hour period in a high pressure chamber. The pressure is then rapidly returned to ambient levels. This processing method has produced scaffolds with porosities of 97% and pore diameters of 100–500 μm . Porosity and pore structure are dependent on the amount of gas dissolved in the polymer, the rate and type (homogeneous or heterogeneous) of gas nucleation, and the diffusion rate of gas molecules through the polymer to the pore nuclei.^[19]

A similar process was described that combines the principles of gas foaming and particulate leaching.^[20] The primary distinction is that salt particles are incorporated into the polymer at the initial stage and later leached from the scaffold by water immersion. Another difference is that all steps are carried out at ambient temperature. The use of salt particles has allowed pores to be more interconnected, with pore diameters of 100–500 μm and porosities up to 97%. It should be noted that two types of pores form: large, open pores from the salt particles; and small, closed pores from gas nucleation.^[20]

A third method of gas foaming begins by mixing ammonium bicarbonate into a PLLA solution (in methylene chloride or chloroform).^[21] This combination forms a paste-like material, which can be readily molded into various shapes. The mold is allowed to air-dry at atmospheric pressure. At this step in the process, there are two options for creating pores. One technique involves vacuum drying the composite for two weeks, followed by immersion in warm water. The other option is to immerse the composite in hot water, followed by a cold water bath. In both cases, the scaffold is subsequently freeze dried. Scaffolds subjected to a prolonged vacuum-drying step show a dense surface skin layer and partially collapse during the process, making this method generally unacceptable. With hot water immersion, however, interconnected pores with diameters of 200–500 μm and porosities over 90% are possible.^[21]

Gas foaming was originally developed to overcome the need for toxic solvents; the first and second methods discussed maintain this quality, but the third technique does involve organic solvents. After the initial disk is formed in method one, all the aforementioned approaches avoid high temperatures in processing, a factor important for the potential use of bioactive molecules. In comparison to solvent casting and particulate leaching, gas foaming is reported to produce scaffolds with more uniform pore structure and stronger mechanical properties. Some drawbacks of the first method are that a polymer skin forms over the surface of the scaffold and gas nucleation leads to a closed pore morphology. Another drawback is that it is difficult to form a complex three-dimensional scaffold using gas foaming techniques other than the third method without sophisticated molding techniques.

Freeze Drying

Freeze drying is a method of removing solvents via sublimation. This process is commonly performed in a lyophilizer by subjecting frozen specimens to a deep vacuum, thus reducing the solvents' vapor pressure. One freeze-drying method that has been described is

emulsion freeze drying.^[22] Water is added to a PLGA/methylene chloride solution, forming two immiscible layers. The layers are homogenized, and the result is an emulsion consisting of a continuous polymer/solvent phase and a dispersed water phase. This emulsion is poured into a copper mold and quenched in liquid nitrogen. Lastly, scaffolds are freeze dried to remove the water and solvent. Scaffold structure is determined by a number of thermodynamic principles, a discussion of which is beyond the scope of this article. This technique is capable of producing scaffolds with 15–35 μm diameter pores and porosities as high as 95%.^[22]

Freeze-drying techniques have been applied to a number of different polymers, including PLLA, PGA, PLGA, and a PLGA/PPF blend.^[23] The polymers are dissolved in glacial acetic acid or benzene and frozen at -10°C , followed by lyophilization for one week. Scaffold morphology (capillary versus leaflet) depends on the type of polymer and solvent used. Furthermore, the investigators demonstrated the incorporation of the drug isoniazid into these scaffolds and concluded that the polymer type influences the drug's rate of release.^[23] No data on porosity or pore sizes was collected in this study.

Some advantages of the freeze-drying technique are that it does not require either a separate leaching step or the use of high temperatures. Scaffolds fabricated with the first method have an interconnected pore network and high porosity. One drawback is that the pores are relatively small, however. Both of the described techniques require the use of toxic solvents. Another drawback is that long processing times are necessary. While no measurements were reported in either paper, the latter method of freeze drying has shown the potential for the incorporation and release of bioactive molecules.

Phase Separation

Phase separation is another means of scaffold processing designed with the intent of incorporating bioactive molecules. A liquid–liquid phase separation technique has been illustrated to produce foams with the potential for drug delivery.^[24] PLLA and solid naphthalene are mixed in a flask, heated, and stirred to obtain a homogenous solution. The solution is then poured or sprayed (using an atomizer) into a cooled mold. Naphthalene is removed by vacuum drying for 12 hours. This process creates polymer-rich and polymer-poor phases within the solution. The specific foam morphology and pore distribution depend on the thermodynamic mechanism of phase separation (spinodal decomposition versus metastable nucleation). This technique creates a relatively uniform pore distribution with diameters of 50–100 μm . Porosities up to

87% can be achieved, depending on the polymer concentration in the solution.^[24] The investigators were able to incorporate different dyes into the scaffolds, and results showed that the distribution and rate of release were influenced by the hydrophobicity of the polymer molecules.

An alternative method, solid–liquid phase separation, has been developed to enhance the mechanical properties of scaffolds.^[25] PLLA or PLGA is dissolved in dioxane or a dioxane/water mixture, and hydroxyapatite (the mineral component of normal bone) powder is added to the solution. The mixture is cooled to solidify the solvent and create a solid–liquid phase separation, and then quenched in liquid nitrogen. Several days of freeze drying are needed to remove any remaining solvent. Scaffolds with porosities of 95% and interconnected pore diameters as large as 600 μm are possible. The porosity, pore size, and pore morphology are dependent on the polymer concentration, polymer structure, hydroxyapatite content, phase-separation temperature, and solvent.^[25]

The foremost advantage of the first method is the potential use of the scaffold as a delivery vehicle for bioactive molecules. Enhanced mechanical strength resulting from the second technique is particularly relevant for the engineering of osseous tissue. Both methods use organic solvents that may have detrimental effects on cells.

Acellular Matrix and Submucosa

All of the processing methods discussed thus far have utilized synthetic polymers to construct scaffolds. However, a number of studies have also investigated the use of natural biomaterials for tissue engineering. Virtually all body tissues are built upon an ECM that consists of structural and functional proteins, proteoglycans, and glycoproteins. Collagen (primarily type I) is the most abundant ECM protein. Cytokines and growth factors are also present in the ECM. The composition of the ECM varies according to tissue function and location. In short, decellularization involves the removal of cells from the ECM through a series of chemical and/or enzymatic steps, resulting in an acellular matrix that can serve as a tissue engineering scaffold. One such scaffold has been successfully prepared using the following technique.^[26] Whole bladders are excised and run through a series of chemical washes to lyse cells and remove cytoplasmic and nuclear material. In the final step, bladders are placed in 70% ethanol for 24 hours, washed with Sorenson's phosphate buffer, and refrigerated in Dulbecco's PBS until ready for use. In addition to the bladder, ECM can be harvested from the dermis, small intestine, pericardium, stomach, basement membrane,

and stroma of the decellularized liver, and decellularized Achilles tendon.^[27] In general, cell extraction from these sources leaves behind primarily collagen and elastin.

The submucosa (natural biomaterial) is defined as the layer of tissue beneath a mucous membrane or the layer of connective tissue beneath the tunica mucosa. Submucosa was used as early as 1989, when a small intestine submucosa was utilized as a vascular graft in a dog model.^[28] ECM and submucosa are touted as good biomaterials for tissue engineering because of the array of proteins, growth factors, and cytokines located in these tissues. These factors are believed to contribute significantly to neovascularization and nerve regeneration, as observed during *in vivo* studies.^[29] Natural biomaterials are also thought to be less immunogenic than synthetic scaffolds. This, however, varies greatly by source and processing even though collagen is highly conserved across species. A drawback of ECM and submucosa is that they are rapidly degraded *in vivo*. This is especially true of submucosa, which is relatively thin (approximately 0.1 mm thick). Scaffold shrinkage *in vivo* is also problematic with ECM. Calcification is commonly reported for both ECM and submucosa. Though the discussed methods do not require toxic chemicals, some methods have utilized chemical cross-linking agents to improve mechanical properties. It should be noted that doing so may often cause a fibrous tissue response.^[29]

Connective Tissue Capsule

When a foreign object is placed inside the body, the immune system reacts by surrounding the material with a layer of connective tissue. Investigators have taken advantage of this phenomenon to create natural cell delivery vehicles.^[30,31] A sterile silicone sheet, tube, or block of desired shape is placed in the subcutaneous tissue, and, after several weeks, a capsule forms around the block during wound healing due to the foreign body. If the foreign body is placed close to a large blood vessel, it has the potential to form a single vascular pedicle, which can be mobilized as a flap. The capsule, which is several millimeters thick, generally contains a highly vascularized layer of fibroblasts, a dense collagen network, laminin, and fibronectin. The capsular surface may act as a basal lamina upon which cells can organize and develop. Similar methods have been developed to grow vascular prosthetics in a patient's peritoneal cavity.^[32]

This technique has the advantage of eliciting a strong angiogenic response, which is often a critical factor in determining the extent of tissue development. Furthermore, a variety of shapes are possible, depending on the shape of the implanted foreign body. A

limitation of this method is that at least two procedures are required: one to insert the silicone object and one to seed cells and/or perform the surgical procedure of interest. The time required for capsule formation can be up to several weeks. Mechanical strength of this tissue is questionable.

Rapid Prototyping Techniques

The utilization of rapid prototyping techniques to build tissue engineering scaffolds is becoming more prevalent. In general, this method allows for the design and construction of thick constructs that are highly porous and possess an intricate network of interconnected pores. These techniques include stereolithography (SLA), laminated object manufacturing (LOM), selective laser sintering (SLS), fused deposition modeling (FDM), solid ground curing (SGC), and three-dimensional printing (3-DP).^[33] Target tissue data from magnetic resonance imaging (MRI) or computed tomography (CT) scans can be fed into a computer to generate a model with the same gross architecture. These models can then be used in conjunction with one of the previously mentioned rapid prototyping systems to duplicate the design as a three-dimensional structure.^[34]

One such method describes the use of FDM to fabricate scaffolds from polycaprolactone (PCL),^[35,36] as shown in Fig. 2. Once a three-dimensional object is digitally created on a CAD workstation, the file is transferred to the FDM software, which mathematically slices the model into horizontal layers of known thicknesses. These layer thicknesses are a function of the equipment and polymers used, but are typically a few hundred microns. A tool path is also generated and the parsed file is downloaded to the FDM hardware. The FDM consists of an extrusion head above a precision X- and Y-axis positioning platform that traces out each layer fed into the machine. The Z-axis is controlled by automatically lowering the platform after each slice is reproduced.^[37]

A thermoplastic polymer, in this example PCL, is fed into the temperature controlled extrusion head and heated to a semi-fluid state. The polymer is then extruded and deposited onto the base, where it quickly cools and solidifies. This rapid cooling allows for unsupported structures to be built and is one advantage FDM has over many other rapid prototyping methods.^[33] For the purpose of testing, a honeycomb-like structure was produced with pore sizes that ranged from 160 to 700 microns.^[36] Fibers between 260 and 370 microns were produced for the structure. The slice thickness was set at 254 microns. The combination of these parameters resulted in scaffolds with porosities ranging from 48% to 77%. Mechanical testing showed

T

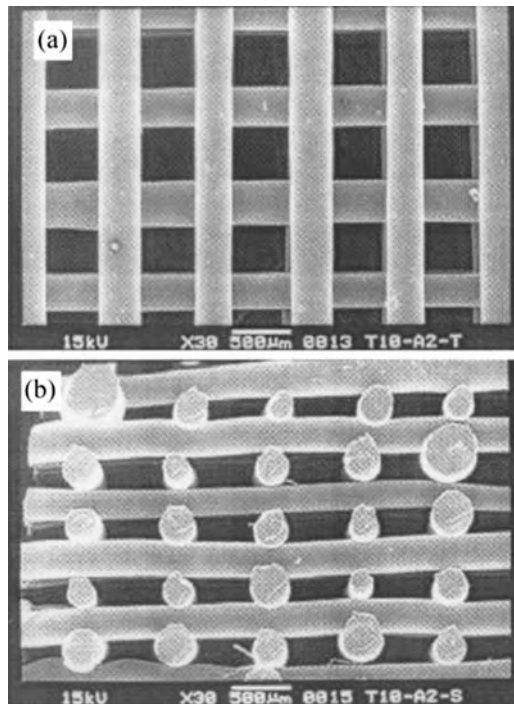


Fig. 2 Top (a) and side (b) view of a polycaprolactone tissue engineering scaffold produced via a CAD/CAM-fused deposition modeling approach. Size bar insert is 500 microns. (Reprinted with permission of Elsevier Science [Ref. [36] p. 1176].)

that the constructs possessed high compressive strength and that strength could be tailored by varying the honeycomb pattern. Scanning electron and light microscopy confirmed the highly ordered patterns and the repeatability of the structure. Tissue culturing on the scaffold demonstrated that cells penetrated the scaffold and proliferated within its three-dimensional space.^[35,36]

Another rapid prototyping technique utilized for development of tissue engineering scaffolds is 3-D printing (3-DP). Briefly, this technique uses ink jet printing heads filled with special polymers attached to a frame over a precision X-axis and Y-axis positioning table. Molten droplets of low melting point polymers are formed by the piezoelectric transducer and these droplets solidify upon contact with the positioning table or previously completed layer to form beads. Adjacent beads overlap to form lines and adjacent lines overlap to form layers. Multiple print heads allow different polymers to be used in a single layer or model. After the layer is complete, the top surface of the layer is flattened to control thickness and provide a flat surface for the next layer to be built. The platform is lowered, and the process is repeated to build the next layer, which adheres to the previous layer, until the entire model is complete.^[37]

This 3-DP technique was utilized in a two-step fabrication process to create a novel collagen scaffold.^[38] CAD software was used to design a mold possessing the negative shape of the scaffold. The CAD file is then downloaded to the 3-DP equipment and a model was printed by controlled droplet deposition of mold and support materials to form each layer. The completed mold containing support material structures was immersed into a selective solvent to remove the mold material. A collagen dispersion was then cast into the mold cavity and frozen. The support material was then dissolved in ethanol, and the collagen scaffold was subjected to critical point drying. The result was a dry collagen scaffold with controlled pore channels and porous channel walls. This technique is capable of details as fine as 181 microns, but only details of 200 microns and larger could be precisely controlled by the software in this example.^[37]

Both of these rapid prototyping techniques are capable of producing structural scaffolds with specifically designed geometries and networks of inter-connecting pores. Material choices range from bioabsorbable thermoplastics like PCL and poly(lactide-co-glycolide) to collagen. The CAD software can be coupled with MRI and CT data to create scaffolds that very closely match the target organs. One of the major drawbacks may be resolution limitations; these techniques typically have X, Y, and Z resolutions of a few hundred microns to one millimeter. These scaffolds have been shown in the lab to facilitate cellular ingrowth but the pore/layer size mismatch with cells (10 to 100 times larger than a cell) may prevent completely natural interactions. In addition, the more sophisticated methods, such as the 3-DP collagen scaffold, could leave residual toxic solvents in the scaffold.

Electrospinning

One textile manufacturing process that is gaining popularity for creation of tissue engineered scaffolds is electrospinning. Electrospinning traces its roots to electrostatic spraying, which was first described more than 100 years ago. In electrostatic spraying, charge is injection into a liquid, typically 5–30 kV, from an electrode. The charged liquid is separated some distance from a second electrode of opposite polarity to establish a static electric field. A so-called Taylor Cone forms, due to the competing forces of the static electric field and the liquid's surface tension. For liquids with a finite conductivity, charged droplets are dispersed from the tip of the Taylor Cone and are delivered to a grounded target. If the liquid consists of a polymer melt or a polymer in solution and the concentration of

that polymer is sufficiently high to cause molecular chain entanglement, a fiber, rather than a droplet, is drawn from the tip of the Taylor Cone.^[38]

A basic electrospinning system consists of a charged polymer solution (or melt) that is fed through a small opening or nozzle (usually a needle or pipette tip). Because of its charge, the solution is drawn toward a grounded collecting plate (usually a metal screen, plate, or rotating mandrel), typically 5–30 cm away, as a jet. During the jet's travel, the solvent gradually evaporates, and a charged polymer fiber is left to accumulate on the grounded target. The charge on the fibers eventually dissipates into the surrounding environment. The resulting product is a nonwoven fiber mat that is composed of tiny fibers with diameters between 50 nanometers and 10 microns.^[39] This nonwoven mat forms the foundation of the scaffold. By varying the orientation of the fibers in the scaffold, from parallel alignment to a random arrangement, the mechanical properties of the scaffold can be varied.^[39]

As Fig. 3 shows, the electrospinning process has been used to create tissue engineering scaffolds composed of PLLA, PGA, and PCL (including blends and copolymers), as well as natural polymers such as collagen, elastin, and fibrinogen.^[38–43] All the scaffolds were reported to have strengths and stiffnesses that are equivalent to native soft tissues. Uniaxial material characterization demonstrated the anisotropic nature of the aligned scaffolds, mimicking the anisotropic characteristics of most soft tissues. Both in-vivo (rat intramuscular implant) and in-vitro cell culture studies were performed to confirm biocompatibility and cellular infiltration. These scaffolds had porosities greater than 80% and pore diameters that ranged from a few hundred nanometers to approximately 10 microns. Since

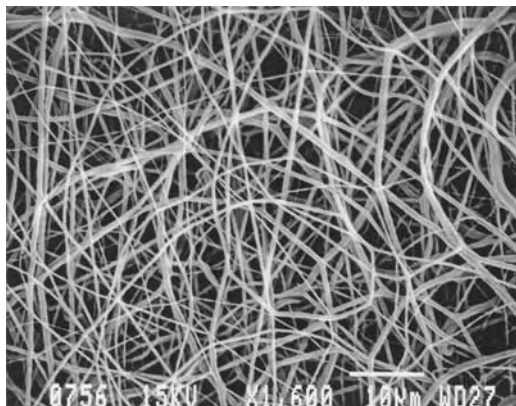


Fig. 3 Random fibrous mat of PGA produced through the process of electrospinning. The fiber diameters average 0.63 microns with pore dimensions of approximately 0.20 microns.

cells were able to penetrate all these scaffolds,^[38–44] the previously mentioned paradigm for cellular infiltration must be reevaluated. It is possible that cells respond differently to fibers that more closely mimic the size of natural ECM fibers.

A lack of mechanical strength is often touted as the principal disadvantage of most textile-based scaffold fabrication techniques. In addition, the fluorinated solvents necessary for many of the polymers to be electrospun are toxic. These solvents could adversely affect cellular response if not fully extracted from the scaffolds. The initial compatibility tests did not indicate any solvent-related problems in rat models; however, tests for residual solvents were not conducted (unpublished data). Also, cross-linking agents were used when preparing the natural polymer scaffolds. These effects must also be elucidated. Even with these possible limitations, electrospinning shows great promise in the development of biomimicking tissue engineering scaffolds.

CONCLUSION

Current research in tissue engineering is at the threshold of providing significant benefits to patients and physicians in the treatment of injury and disease. Many methods have emerged, from the simple concept of placing cells in a degradable scaffold to building native tissues either in vivo or in vitro. These advances come on the back of breakthroughs in the life sciences that provide critical information about the nature and development of tissues and diseases. Ultimately, engineers must match applications, materials, and fabrication processes to best meet the needs of the tissue they wish to build.

ARTICLES OF FURTHER INTEREST

Biodegradable Polymers, An Overview; Biofunctional Polymers; Biologic and Synthetic Apatites; Biomimetic Materials; Electrospinning; Extracellular Matrix Scaffolds; Fabrics; Fibrin; Microporous Materials; Poly(Glycolic Acid); Poly(lactic acid)s

REFERENCES

1. Bell, E. Tissue engineering: A perspective. *J. Cell. Biochem.* **1991**, *45*, 239–241.
2. Greisler, H.P.; Gosselin, C.; Ren, D.; Kang, S.S.; Kim, D.U. Biointeractive polymers and tissue engineered blood vessels. *Biomaterials* **1996**, *17*, 229–236.

3. How, T.V.; Guidoin, R.; Young, S.K. Engineering design of vascular prostheses. *Proc. Inst. Mech. Eng.* **1992**, *206*, 61–71.
4. Vacanti, J.P.; Vacanti, C.A. The Challenge of Tissue Engineering. In *Principles of Tissue Engineering*; Lanza, R., Langer, R., Chick, W., Eds.; Academic Press, 1997; 1–5.
5. Mooney, D.J.; Langer, R.S. Engineering Biomaterials for Tissue Engineering: The 10–100 Micron Size Scale. In *The Biomedical Engineering Handbook*; CRC Press, 1995; 1609–1618.
6. Starke, G.R.; Douglas, A.S.; Conway, D.J. An Integral Mathematical Approach to Tissue Engineering of Vascular Grafts. In *Tissue Engineering of Prosthetic Vascular Grafts*; Zilla, P., Greisler, H.P., Eds.; R.G. Landes Company, 1999; 441–459.
7. Hsu, S.; Kambic, H. On matching compliance between canine carotid arteries and polyurethane grafts. *Artif. Organs* **1997**, *21*, 1247–1254.
8. Hynes, R.O. Integrins: Bidirectional, allosteric signaling machines. *Cell* **2002**, *110*, 673–687.
9. Olsen, B.R. Matrix Molecules and Their Ligands. In *Principles of Tissue Engineering*; Lanza, R., Langer, R., Chick, W., Eds.; R.D. Landes Co., 1997; 47–65.
10. Mikos, A.G.; Bao, Y.; Cima, L.G.; Ingber, D.E.; Vacanti, J.P.; Langer, R.S. Preparation of poly(glycolic acid) bonded fiber structures for cell attachment and transplantation. *J. Biomed. Mater. Res.* **1993**, *27*, 183–189.
11. Mooney, D.J.; Mazzoni, C.L.; Breuer, C.; McNamara, K.; Hern, D.; Vacanti, J.P.; Langer, R.S. Stabilized polyglycolic acid fiber-based tubes for tissue engineering. *Biomaterials* **1996**, *17* (2), 115–124.
12. Mikos, A.G.; Lyman, M.D.; Freed, L.E.; Langer, R.S. Wetting of poly(L-lactic acid) and poly(DL-lactic-co-glycolic acid) foams for tissue culture. *Biomaterials* **1994**, *15* (1), 55–58.
13. Thomson, R.C.; Yaszemski, M.J.; Powers, J.M.; Mikos, A.G. Hydroxyapatite fiber reinforced poly(α -hydroxy ester) foams for bone regeneration. *Biomaterials* **1998**, *19*, 1935–1943.
14. Shastri, V.P.; Martin, I.; Langer, R.S. Macroporous polymer foams by hydrocarbon templating. *Proc. Natl. Acad. Sci.* **2000**, *97*, 1970–1975.
15. Wake, M.C.; Gupta, P.K.; Mikos, A.G. Fabrication of pliable biodegradable polymer foams to engineer soft tissues. *Cell Transplant* **1996**, *5* (4), 465–473.
16. Mikos, A.G.; Sarakinos, G.; Leite, S.M.; Vacanti, J.P.; Langer, R.S. Laminated three-dimensional biodegradable foams for use in tissue engineering. *Biomaterials* **1993**, *14* (5), 323–328.
17. Thomson, R.C.; Yaszemski, M.J.; Powers, J.M.; Mikos, A.G. Fabrication of biodegradable polymer scaffolds to engineer trabecular bone. *J. Biomater. Sci., Polym. Ed.* **1995**, *7*, 23–28.
18. Widmer, M.S.; Gupta, P.K.; Lu, L.; Meszlenyi, R.K.; Evans, G.R.D.; Brandt, K.; Savel, T.; Gurlek, A.; Patrick, C.W., Jr.; Mikos, A.G. Manufacture of porous biodegradable polymer conduits by an extrusion process for guided tissue regeneration. *Biomaterials* **1998**, *19*, 1945–1955.
19. Mooney, D.J.; Baldwin, D.F.; Suh, N.P.; Vacanti, J.P.; Langer, R.S. Novel approach to fabricate porous sponges of poly(D,L-lactic-co-glycolic acid) without the use of organic solvents. *Biomaterials* **1996**, *17*, 1417–1422.
20. Harris, L.D.; Kim, B.S.; Mooney, D.J. Open pore biodegradable matrices formed with gas foaming. *J. Biomed. Mater. Res.* **1998**, *42* (3), 396–402.
21. Nam, Y.S.; Yoon, J.J.; Park, T.G. A novel fabrication method of macroporous biodegradable polymer scaffolds using gas foaming salt as a porogen additive. *J. Biomed. Mater. Res.* **2000**, *53* (1), 1–7.
22. Whang, K.; Thomas, C.H.; Healy, K.E. A novel method to fabricate bioabsorbable scaffolds. *Polymer* **1995**, *36* (4), 837–842.
23. Hsu, Y.Y.; Gresser, J.D.; Trantolo, D.J.; Lyons, C.M.; Gangadharam, P.R.J.; Wise, D.L. Effect of polymer foam morphology and density on kinetics of in vitro controlled release of isoniazid from compressed foam matrices. *J. Biomed. Mater. Res.* **1997**, *35*, 107–116.
24. Lo, H.; Guggino, S.E.; Leong, K.W. Poly(L-lactic acid) foams with cell seeding and controlled-release capacity. *J. Biomed. Mater. Res.* **1996**, *30*, 475–484.
25. Zhang, R.; Ma, P.X. Poly(α -hydroxy acids)/hydroxyapatite porous composites for bone-tissue engineering. I. Preparation and morphology. *J. Biomed. Mater. Res.* **1999**, *44*, 446–455.
26. Reddy, P.P.; Barrieras, D.J.; Wilson, G.; Bagli, D.J.; McLorie, G.A.; Khoury, A.E.; Merguerian, P.A. Regeneration of functional bladder substitutes using large segment acellular matrix allografts in a porcine model. *J. Urol.* **2000**, *164*, 936–941.
27. Badylak, S.F. The extracellular matrix as a scaffold for tissue reconstruction. *Semin. Cell Dev. Biol.* **2002**, *13* (5), 377–383.
28. Badylak, S.F.; Lantz, G.C.; Coffey, A.; Geddes, L.A. Small intestine submucosa as a large diameter vascular graft in the dog. *J. Surg. Res.* **1989**, *47*, 74–80.
29. Yoo, J.J.; Meng, J.; Oberpenning, F.; Atala, A. Bladder augmentation using allogenic bladder submucosa seeded with cells. *Urology* **1998**, *51* (2), 221–225.
30. Lee, H.B.; Lew, D.H. De novo induction of island capsule flap by using two silastic sheets: Part 1. Generation. *Plast. Reconstr. Surg.* **1996**, *104*, 1023–1028.
31. Schoeller, T.; Lille, S.; Stenzl, A.; Ninkovic, M.; Piza, H.; Otto, A.; Russell, R.C.; Wechselberger, G. Bladder reconstruction using a prevascularized capsular tissue seeded with urothelial cells. *J. Urol.* **2001**, *165*, 980–985.
32. Campbell, J.H.; Efendy, J.L.; Campbell, G.R. Novel vascular graft grown within recipient's own peritoneal cavity. *Circ. Res.* **1999**, *85*, 1173–1178.
33. Jacobs, P.F. *Stereolithography and Other RP&M Technologies from Rapid Prototyping to Rapid Tooling*; ASME Press: New York, 1996; 1–26.
34. Sun, W.; Lal, P. Recent development on computer aided tissue engineering—A review. *Comput. Methods Programs Biomed.* **2002**, *67* (2), 85–103.
35. Hutmacher, D.W. Scaffold design and fabrication technologies for engineered tissues—State of the art and future perspectives. *J. Biomater. Sci., Polym. Ed.* **2001**, *12*, 107–124.

36. Zein, I.; Hutmacher, D.W.; Tan, K.C.; Teoh, S.H. Fused deposition modeling of novel scaffold architectures for tissue engineering applications. *Biomaterials* **2002**, *23*, 1169–1185.
37. Sachlos, E.; Reis, N.; Ainsley, C.; Derby, B.; Czer-nuszka, J.T. Novel collagen scaffolds with predefined internal morphology made by solid freeform fabri-cation. *Biomaterials* **2003**, *24*, 1487–1497.
38. Bowlin, G.L.; Pawlowski, K.J.; Stitzel, J.D.; Boland, E.D.; Simpson, D.G.; Fenn, J.B.; Wnek, G.E. Electro-spinning of Polymer Scaffolds for Tissue Engineering. In *Tissue Engineering and Biodegradable Equivalents: Scientific and Clinical Applications*; Wise, D., Eds.; The Humana Press, Inc. 2002; 65–178.
39. Boland, E.D.; Wnek, G.E.; Simpson, D.G.; Pawlowski, K.J.; Bowlin, G.L. Tailoring tissue engineering scaf-folds using electrostatic processing techniques: A study of poly(glycolic acid). *J. Macromol. Sci.* **2001**, *38*, 1231–1243.
40. Stitzel, J.D.; Pawlowski, K.J.; Wnek, G.E.; Simpson, D.G.; Bowlin, G.L. Arterial smooth muscle cell proliferation on a novel biomimicking, biodegradable vascular graft scaffold. *J. Biomater. Appl.* **2001**, *15*, 1–12.
41. Matthews, J.A.; Simpson, D.G.; Wnek, G.E.; Bowlin, G.L. Electrospinning of collagen nanofibers. *Biomacromolecules* **2002**, *3*, 232–238.
42. Matthews, J.A.; Boland, E.D.; Wnek, G.E.; Simpson, D.G.; Bowlin, G.L. Electrospinning of collagen type II: A feasibility study. *J. Bioact. Compat. Polym.* **2003**, *18*, 125–134.
43. Wnek, G.E.; Carr, M.E.; Simpson, D.G.; Bowlin, G.L. Electrospinning of nanofiber fibrinogen structures. *Nano Lett.* **2003**, *3* (2), 213–216.
44. Sanders, J.E.; Stiles, C.E.; Hayes, C.L. Tissue response to single-polymer fibers of varying diameters: Evalua-tion of fibrous encapsulation and macrophage density. *J. Biomed. Mater. Res.* **2000**, *52*, 231–237.
45. Zhang, R.; Ma, P.X. Processing of Polymer Scaffolds: Phase Separation. In *Methods of Tissue Engineering*; Atala, A., Lanza, R.C., Eds.; Academic Press 2002; 715–724.

Tissue Engineering of Skeletal Muscle

Christopher Keen

Department of Physiology, Virginia Commonwealth University, and Department of Anatomy and Neurobiology, Virginia Commonwealth University, Richmond, Virginia, U.S.A.

Daniel Newton

Department of Anatomy and Neurobiology, Virginia Commonwealth University, Richmond, Virginia, U.S.A.

Todd A. Telemeco

Shenandoah University, Winchester, Virginia, U.S.A.

Gary L. Bowlin

Department of Physiology, Virginia Commonwealth University, and Department of Biomedical Engineering, Virginia Commonwealth University, Richmond, Virginia, U.S.A.

Gary E. Wnek

Department of Chemical Engineering, Case Western Reserve, Cleveland, Ohio, U.S.A.

Clive M. Baumgarten

Department of Physiology, Biomedical Engineering, and Internal Medicine (Cardiology), Medical College of Virginia, Virginia Commonwealth University, Richmond, Virginia, U.S.A.

David G. Simpson

Department of Physiology, Virginia Commonwealth University, and Department of Anatomy and Neurobiology, Virginia Commonwealth University, Richmond, Virginia, U.S.A.

INTRODUCTION

The congenital defects and traumatic injuries that compromise the structure and function of skeletal muscle range from the functionally benign to the profoundly debilitating. In either circumstance, dysfunctional muscle can impact a patient on a number of levels. A defect to the musculature of the face may not have a significant impact on patient mortality; however, even a relatively minor cosmetic defect can have profound psychological effects and greatly degrade the quality of life. In many cases, for example, the reconstruction of cleft palate in young children, a relatively minor increase in the amount of tissue available for reconstructive surgery can greatly improve the therapeutic outcome. From a clinical prospective, we believe that congenital defects of the face might represent one of the initial applications where bioengineered muscle can have an impact on patient health care. Other potential applications include bioengineered muscle for the reconstruction of facial muscles damaged by traumatic injury and implants designed for protein or gene-mediated therapies.

STRUCTURE OF MUSCLE

To truly restore function to a damaged muscle, the architectural features of the native tissue must be substantially reconstituted in the bioengineering paradigm. The structure and mechanical properties of skeletal muscle are defined by a dynamic interaction between the active, contractile cellular elements of this tissue and the passive, structural constituents of the surrounding extracellular matrix. The active cellular elements of the prototypic skeletal muscle are distributed into dense, parallel arrays. This structural feature allows the tissue to effectively apply force along a defined axis. The multinucleated skeletal muscle cell, or myotube, exhibits a highly organized internal cytoarchitecture. The myofibrils, the contractile filaments of muscle, are tightly packed within the sarcoplasm of these cells. Each myofibril is subdivided into a series of tandem repeating structures called sarcomeres. The molecular motors of contraction, the actin thin filaments and myosin thick filaments, are organized into nearly crystalline arrays within the sarcomeres. Accessory cytoskeletal proteins stabilize the structural arrangement of the sarcomeres and link adjacent

myofibrils to one another and the surrounding extracellular environment. The myofibrils are interspersed with columns of mitochondria and are surrounded by projections of the membranous sarcoplasmic reticulum. T-tubules, which are invaginations of the plasma membrane (sarcolemma), penetrate the myotubes and interact with the sarcoplasmic reticulum to distribute the wave of depolarization initiated by the central nervous system that regulates voluntary muscle contraction.

The active cellular elements of muscle are surrounded in the extracellular space by a connective tissue scaffolding. This connective tissue scaffolding, or extracellular matrix, supports and defines the overall structural organization of the tissue. It can be arbitrarily subdivided into four interconnected compartments: the basal lamina, the endomysium, the perimysium, and the epimysium.^[1,2] The basal lamina is synthesized by the myotubes and is enriched in laminin, type IV collagen, and a variety of proteoglycans. It is intimately applied over the external surface of each individual myotube and is believed to lend structural support to the outer leaflet of the plasma membrane. The endomysium is continuous with and overlays the basal lamina and is enriched in the extracellular proteins type III collagen, laminin, tenascin, and the multifunctional protein, fibronectin. The molecular composition of the perimysium is similar to that of the endomysium; type I collagen, type III collagen, and fibronectin are present in this compartment. The relative position and structural organization of perimysium are unique and distinguish this connective tissue compartment from the investing endomysium. The perimysium surrounds several individual myotubes and functions to group several myotubes into a discrete structural unit. The largest subdivision of the connective tissue compartment is the epimysium. It is visible at gross dissection and completely encases the entire muscle. The structural integrity of this connective tissue compartment varies across different muscle groups. It is composed predominantly of type I collagen. In addition to lending mechanical integrity, the connective tissue scaffolding of muscle also provides a pathway for the distribution of nerves, blood vessels, and lymphatics. The larger arteries and veins of skeletal muscle run through the perimysium; capillaries and motor nerves are distributed throughout muscle by passing along the endomysium.

EXTERNAL CONSTRAINTS ON THE BIOENGINEERING PARADIGM

To date, the ability to fabricate muscle and other multicellular tissues *in vitro* for use in reconstructive therapy *in vivo* has been limited by a series of nearly circular biological and technical limitations. *In vivo*,

normal developmental processes take place over an extended period of time. During embryonic development, the mass and density of skeletal muscle, and other tissues, increases in parallel with the volume of capillaries present and the capacity of the cardiovascular system to deliver oxygen and provide nutrient exchange. Tissue fabricated *in vitro* lacks cardiovascular support and must be sustained by passive diffusion. This consideration limits the mass and density of tissue that can be fabricated and supported *in vitro*. Small scale constructs of muscle have long been, and routinely, prepared in conventional tissue culture.^[3-5] However, the structure of muscle prepared under conventional culture conditions restricts the degree to which the engineered tissue can tolerate manual manipulation, and, more importantly, the small size of these constructs greatly limits the contractile performance of the resulting muscle tissue. Theoretically, these mechanical and functional limitations can be overcome simply by increasing the density of the cellular elements and the overall mass of the engineered tissue. However, as cell density increases, the ability to provide adequate rates of oxygen, nutrient, and waste exchange to a bioengineered organ become limiting factors. Preparing bioengineered tissue in a bioreactor (a generic designation for a device designed to increase nutrient exchange) can, to some extent, increase the mass of tissue that can be supported *in vitro*. However, when first implanted, any tissue fabricated *in vitro* must still be supported by the passive diffusion of nutrients from the surrounding interstitium until a functional capillary network and nascent interconnections with the cardiovascular system of the host develop. These biological and technical limitations represent daunting barriers to the bioengineering paradigm for muscle and many other types of tissue.

HISTORICAL PROSPECTIVE OF SKELETAL MUSCLE ENGINEERING

Several different procedures have been used in attempts to reconstruct damaged or missing domains of skeletal muscle. For example, in perhaps the most direct approach, intact muscle can be harvested and transplanted directly into a recipient site.^[6-8] The advent and development of microsurgical techniques that make it possible to anastomose small vessels of the donor muscle and the recipient tissue has made this approach feasible, and muscle transplants are used in the laboratory setting with great success.^[9] However, tissue may be routinely, and arbitrarily, harvested from a donor animal and used for experimental transplant therapy. In the clinical setting, this does not represent a common treatment option. Donor material for organ

transplants is critically limited in supply, has a very limited window of viability, and must be matched to the recipient, or the recipient's immune system must be suppressed to manage tissue rejection. A more realistic approach is to use allograft donor material to reconstruct damaged muscle. This strategy, commonly referred to as a Free Tissue Transfer (Free Flap), has been used in cosmetic, and to some extent, in the functional reconstruction of damaged tissues. Donor tissues include the rectus abdominus, latissimus dorsi, serratus anterior, gracilis, and the extensor digitorum brevis muscles. These muscles are used in free tissue transfer procedures because they have a well-defined blood supply that makes it possible to anastomose vessels of the donor tissue to the vessels of the donor site. This procedure is often used to provide protection to the terminal ends of an amputated limb and provide cushioning between the transected bone and a mechanical prosthesis. The use of an allograft for muscle reconstruction has several obvious advantages. This type of donor tissue has the histological features, vascular circulation, and structural integrity of differentiated muscle. As an allograft, an implant prepared as a Free Flap should not invoke any immune response. The limitations to this approach include the functional deficits and the disfigurement that may occur in the donor site. Also, depending on the nature of the surgical site, the implanted tissue will, in all likelihood, not be subject to voluntary control from the central nervous system. An alternative reconstructive strategy, and the one that our laboratories have chosen to pursue, is to address the limitations that are intrinsic to the bioengineering paradigm and fabricate a tissue engineered muscle.

FABRICATION OF SKELETAL PROSTHETICS

The identification of suitable donor material represents one of the barriers that must normally be overcome in the bioengineering paradigm. Unlike many tissue types, this consideration is not necessarily a limiting factor in the fabrication of engineered skeletal muscle. Satellite muscle cells, a population of partially differentiated stem cells, reside within skeletal muscle and are scattered throughout the interstitial space.^[10] Following injury these cells are induced to divide, differentiate, and repopulate the muscle bed.^[11] The characteristic distribution and biological properties of the satellite cell make it possible to isolate this population from discrete and unobtrusive sites by conventional biopsy techniques and use it as donor material in muscle reconstruction.

Satellite muscle cells can be grown and expanded in tissue culture and, under the appropriate conditions, will undergo differentiation to form striated myotubes.

Differentiation is initiated by the formation of cell-to-cell contacts between adjacent satellite cells. The specific phenotypic properties of the resulting myotubes are determined by an interaction between the intrinsic genetic potential of the cells and the extrinsic environmental signals that shape muscle architecture. For example, the differentiation and the expression of contractile filaments, other muscle-specific cytoarchitectural features, and the basal lamina are regulated by the intrinsic potential of the cells. Extrinsic signals, such as neuronal inputs, soluble growth factors, mechanical forces, and the composition and tertiary structure of the extracellular matrix can interact to regulate striated muscle cell mass, cell shape, and cell polarity.^[12,13] Incorporating these extrinsic phenotypic signals into the bioengineering paradigm to direct the expression of the desired structural characteristics is an essential element of the fabrication process.

In the course of characterizing satellite muscle cells, a subpopulation of multipotent stem cells has been detected. This subpopulation, termed muscle-derived stem cells (MDSC), exhibits a reasonably robust capacity to differentiate into a variety of tissue types.^[14] The MDSC may represent a readily obtainable donor source of material that can be used in the reconstruction of other tissues or to augment satellite muscle cells in the reconstruction of skeletal muscle. For example, there is evidence that some MDSC cells can undergo differentiation into microvascular endothelial cells. This type of cell population could make it possible to fabricate a muscle construct that is preseeded with cells that can form a capillary network.

IN VIVO TISSUE ENGINEERING

The satellite muscle cell is a nearly ideal donor source and theoretically can be used to fabricate a skeletal muscle prosthesis as an autologous graft, limiting the risk of host vs. graft rejection. The efficacy of using satellite muscle cells as a donor source of material was demonstrated in classic experiments, even before this cell type was fully characterized. In these experiments, the epimysium of a donor muscle was incised, the muscle removed en bloc, and minced into small segments or dissociated into a single cell suspension.^[15,16] The donor material was then reimplanted into the original muscle bed and within the intact epimysium. In effect, the epimysium functioned as a natural, albeit primitive, tissue-engineering scaffolding. Over time, the differentiated muscle remaining in the donor tissue died and the surviving satellite muscle cells underwent division, repopulated the muscle bed, and subsequently differentiated into striated muscle. Skeletal muscle reconstructed with this strategy exhibits a rather hazardous cell polarity, but can achieve a very acceptable

degree of mechanical performance. In the rat, muscle reconstructed in this manner can undergo reinnervation and generate contractile forces that approximate normal, intact tissue.^[15]

The chaotic muscle structure observed in the early reconstruction experiments probably results from the complete disruption of the basal lamina, and the other compartments of the extracellular matrix, that occurs when the intact tissue is isolated en bloc to prepare the donor material. Following injury *in vivo*, and in the course of normal muscle regeneration, the basal lamina of a damaged myotube can remain substantially intact, even as the underlying cell undergoes necrosis. This extracellular ghost of the dying cell can act in concert with other components of the extracellular matrix in directing the differentiating satellite muscle cell to adopt a defined polarity.^[17] In the absence of extracellular elements, satellite muscle cells are not constrained to spread along a defined axis and adopt the polarized phenotype that characterizes intact muscle.

There are obvious limitations in the use of muscle fragments or single cell suspension therapies in the reconstruction of damaged tissue *in situ*. Signals that are intrinsic to the satellite cell allow this donor source to differentiate. However, the extrinsic signals necessary to shape muscle polarity are not fully provided in this type of natural bioengineering strategy. It represents a relatively passive approach that does not allow the surgeon to sculpt the implant to any great extent and requires a preexisting muscle bed to act as an implantation site. This is a structural feature that may be completely absent in a patient with a congenital defect or a devastating traumatic injury. In addition, the haphazard distribution of myotubes in muscle that has been reconstructed from tissue fragments or dissociated cell preparations makes this therapeutic approach unsuitable for many applications due to functional and cosmetic considerations. However, these experiments provide several essential insights. Foremost, these experiments demonstrate clear proof of concept and indicate that muscle reconstruction therapy using a satellite cell population is a feasible strategy. This approach anticipated how a local microenvironment might be used to regulate and control the expression of the differentiated skeletal muscle phenotype.

TISSUE ENGINEERING SCAFFOLDS

An amalgam of the transplant and dissociated cell strategies is to tissue-engineer a muscle implant *in vitro* for reconstructive surgery *in vivo* using satellite muscle cells or another suitable donor source of material. This engineered tissue must be sufficiently robust to

withstand manual manipulation, the physical rigors of the *in vivo* environment, and yet large enough to generate a meaningful amount of work. Conversely, the tissue cannot be so massive or dense that it cannot be supported by the passive diffusion of nutrients from the surrounding environment while cardiovascular support is established from the host. Passive oxygen influx usually is sufficient to support the metabolism of a 1-mm diameter cylinder of working muscle, a rather small prosthesis. In our design philosophy, the tissue-engineering scaffold plays a central role in overcoming the diffusion constraints that limit construct mass. We believe that a tissue-engineering scaffold must have sufficient mechanical integrity to support and define the initial structural properties of a bioengineered organ. Simultaneously, the scaffold must support rapid microvascular endothelial cell migration and the formation of nascent blood vessels. Over time, the fabrication scaffolding should be degraded and replaced by constituents of the native ECM as the bioengineered construct begins to assume the structural features of mature tissue. This latter characteristic greatly reduces the risk of host vs. graft rejection.

Considerable effort has been directed at developing proteins of the collagen family into tissue-engineering scaffolds. The collagens are highly conserved, represent the principal structural elements of the native interstitium, and are the most abundant protein species of the mammalian system. Collagen is typically isolated by acid extraction of a starting material. A hydrated gel forms when the acid soluble fraction of the extract is returned to physiological ionic concentration and neutral pH. Unfortunately, the delicate nature of these gels has limited the usefulness of this type of collagen isolate in tissue-engineering applications.

In muscle bioengineering experiments, myoblasts and satellite muscle cells have been seeded onto the exterior of preexisting collagen gels^[18] or mixed with collagen and literally enveloped within the material as it undergoes polymerization.^[19] Subsequently, the cells have been allowed to differentiate within the three-dimensional environment of the gel. The value of tissue fabricated by seeding cells onto a collagen gel is limited by the preferential distribution of differentiated myotubes on the external surfaces of the scaffold and the low overall cell density that is achieved with this approach. Experiments in which cells are directly incorporated into a collagen gel as it undergoes polymerization have yielded more densely and uniformly populated constructs.^[19] However, these constructs still remain far less dense than intact tissue and, as a result, the mechanical properties of the tissue are largely defined by the delicate fibrils of the collagen gel. Several different techniques have been used to address the underlying structural limitations of the collagen gel.

“ENGINEERING” THE TISSUE ENGINEERING SCAFFOLD

Cell-mediated bioengineering has been used in attempts to improve the mechanical properties of the conventional collagen gel.^[20] In this strategy, interstitial cells are seeded onto a collagen gel and cultured for several days prior to the addition of the donor material. The objective of this preincubation interval is to allow the interstitial cells to deposit nascent collagen, remodel and condense the scaffold, and cross-link the collagen fibrils of the existing collagen gel. Unfortunately, this technique is time consuming and appears to provide marginal benefits. Collagen gels also have been augmented with synthetic, biocompatible materials. For example, incorporating sheets of polyester mesh into a collagen gel lends considerable strength to the scaffolding while allowing the construct to retain a substantial portion of its elastic properties.^[21] However, tissue fabricated with this type of strategy is still composed of sparsely distributed cells, and nearly any nonbiodegradable material can function as a nexus for an inflammatory response, fibrotic encapsulation, and tissue rejection.^[22]

An alternative approach is to increase the structural stability of the muscle prosthesis by subjecting the tissue to mechanical stimulation to induce hypertrophic growth *in vitro*. This type of training can greatly increase the mass of individual cells and promote the deposition of ECM constituents, thereby improving the structural characteristics of the construct.^[23] Elegant experiments have been conducted with a variety of different devices designed to impart static and cyclic strains on tissue-engineered muscle.^[24,25] Very specific patterns of *in vitro* training have been demonstrated to be necessary to promote maximal growth in these experimental systems.^[26,27] However, even trained skeletal muscle remains less dense than native tissue and diffusion constraints still place severe limits on the mass (and density) of tissue that can be fabricated. In effect, tissue-engineered muscle is subject to a physiological failure mode; the muscle is too small or far too sparsely populated by striated muscle myotubes to produce a meaningful amount of contractile activity.

Mechanical loading of muscle tissue *in vitro* also can induce the expression and activation of matrix metalloproteinases (MMPs). These enzymes, which are nearly ubiquitous across all tissue types,^[28] mediate the degradation of a variety of matrix constituents. In bioengineered muscle, the premature destruction of the tissue engineering scaffolding that supports the construct by these enzymes may lead to the catastrophic failure of the prosthesis; we refer to this as an adaptive failure mode. In adaptive failure, elements of the tissue engineering scaffolding may be degraded long before

the mechanical stimulation induces cell hypertrophy or the deposition of sufficient matrix constituents to stabilize the structure of the muscle.

GROWTH AND MUSCLE ADAPTATION

From these data, it is clear that it is possible to fabricate a muscle-like construct *in vitro*. However, we conclude that diffusion constraints and other biological limitations will restrict the mass and overall contractile performance that can be achieved with this type of tissue. To fabricate a clinically relevant muscle prosthesis, the diffusion constraints that limit construct design must be addressed. Embryonic development, neonatal growth, and maturation are events that occur gradually over an extended period of time. During development, the differentiation of striated tissue is subject to regulation by neurohormonal elements, soluble peptide growth factors, and mechanical activity. These events promote myotube differentiation and hypertrophy, capillary growth, and the elaboration of the extracellular matrix. The tissue is conditioned and, over time, undergoes hypertrophy and increases in density. As a result, it is able to sustain and survive progressively more intense mechanical activity. This mechanical activity is a critical element in the regulatory events that control the growth and overall mass of skeletal muscle. Recapitulating the morphogenetic events that direct the formation of the striated muscle phenotype in an artificial environment will be difficult, at best. From a biological perspective, the structural integrity and viability of muscle tissue subjected to mechanical conditioning *in vitro* can begin to exhibit evidence of deterioration within days. This is particularly evident in tissue fabricated with traditional collagen gels. *In vitro* training regimes promote muscle hypertrophy and the expression and deposition of matrix constituents. Unfortunately, these same mechanical loads also induce the expression, deposition, and activation of MMPs. Collagen gels, by nature, are already quite delicate and are subject to rapid attack and degradation by these proteases, leading to the loss of construct integrity. Mechanical stimulation is essential for muscle growth; however, the catabolic aspects of the hypertrophic response induced by mechanical stimulation outpace and literally “outrun” the anabolic events that should lead to a more structurally stable construct.

A similar maladaptive response can be observed when tissue that is engineered *in vitro* is placed *in vivo*. With respect to an intact muscle, tissue fabricated *in vivo* is not subject to a very challenging mechanical environment. This is true even for tissue that has undergone a preconditioning training regime. At implantation, bioengineered tissue is instantaneously bombarded by a host of biochemical and mechanical

signals; these signals initiate an adaptive hypertrophic response. We should note that one reason to precondition tissue is to blunt the impact of the hypertrophic signals that are encountered when engineered muscle is implanted. Components of the adaptive hypertrophic response are highly desirable, and necessary, to promote the formation of muscle structure and the seamless integration of the engineered tissue into the host site. The expression, deposition, and activation of matrix metalloprotease enzymes are an integral element of this response. These enzymes exhibit a broad spectrum of proteolytic activity against matrix constituents and, like *in vitro*, the loss of scaffold integrity *in vivo* can lead to catastrophic failure of the implant before the muscle has fully adapted to the host site. Any inflammatory response within the implant site can be expected to further exacerbate the enzymatic attack of a bioengineered scaffold composed of native proteins; enzymes of the MMP family are expressed by a variety of circulating inflammatory cells. Clinically, it is not feasible to completely suppress MMP activity, even within the local domain of the bioengineered tissue. Enzymatic activity associated with the MMPs is necessary and essential for the migration of microvascular endothelial cells and the subsequent formation of capillary networks.^[29] Blocking MMP activity will suppress the angiogenic response that is so necessary for implant survival. Finally, the diffusion characteristics of tissue fabricated *in vitro* limit the mass of the construct, even with the advent of devices that can support progressively larger-scale constructs. When implanted, tissue engineered *in vitro* must still be supported by the passive diffusion of oxygen and nutrients from the surrounding tissue of the host. This characteristic places upper limits on the mass of tissue that can be fabricated *in vitro* and used for reconstructive surgery *in vivo*.

INTEGRATING SCAFFOLD DESIGN WITH TISSUE FABRICATION STRATEGIES: ELECTROSPINNING

In our design philosophy, we believe that a bioengineered organ should be fabricated *in vivo* at the recipient site from a stem cell population, or in the case of a striated muscle prosthesis, from a population of satellite muscle cells. These cells are then allowed to undergo differentiation *in vivo*. By using undifferentiated cells as a donor source, it may be possible to mimic aspects of the developmental process and circumvent, within limits, the diffusion constraints that ultimately restrict the bioengineering paradigm. With respect to fully differentiated muscle, undifferentiated and noncontractile satellite muscle cells (or stem cells) exhibit nominal oxygen demand. This is a desirable

characteristic for any tissue that must initially be supported by the passive diffusion of nutrients from the surrounding environment. In reality, our bioengineering strategy is a combination of events borrowed from development and the wound repair process.

Our fabrication scheme places strict constraints on the tissue-engineering scaffolds used to construct the tissue. The scaffold must define the initial structural properties of the tissue; a loose aggregate of undifferentiated cells may have nominal oxygen demand, but it lacks any structural cohesiveness. The scaffold must confine the donor cells to the implant site and provide the cues necessary to promote the expression of the proper phenotype, a densely populated muscle composed of parallel arrays of myotubes. Microvascular endothelial cells from the surrounding interstitium must freely penetrate the scaffold and form functional capillary beds as the differentiation process proceeds. Next, the scaffold must be biocompatible and, ideally, replaced by a native matrix that is expressed and deposited by the donor cell population as the implant begins to assume the structural features of differentiated muscle. In this specification, the rate or interval over which the tissue-engineering scaffold is remodeled and lost must be controlled. Finally, to reach commercial and clinical success, the scaffolding material must pass FDA scrutiny.

We have concluded that conventional scaffolds cannot meet these design constraints and require the development of a new class of tissue-engineering scaffolds. Collagen gels, synthetic materials, and conventional hybrid materials lack the biological properties necessary to meet the stated design considerations. Our laboratories believe that electrospinning technology may represent a strategy that provides the nucleus necessary to develop more physiologically relevant tissue-engineering scaffolds.^[30] A detailed description of electrospinning is beyond the scope of this discussion, but it is described in detail elsewhere in this volume. In brief, in electrospinning, a polymer solution, or melt, is prepared and injected with current to create a charge imbalance.^[31] The charged polymer solution is then placed in proximity to a grounded target, creating an electric field. At a critical voltage, the charge imbalance overcomes the surface tension of the polymer source, forming an electrically-charged jet. Within the electric field, the jet is directed toward the grounded target and solvent evaporates as it traverses the air gap (Fig. 1). Depending on reaction conditions, and the polymers used in the process, electrospinning can be utilized to produce a fine aerosol of material or a continuous nonwoven mat of fibrillar material. We have adapted electrospinning and use it to process collagen and other native proteins into sub-micron diameter fibers, a scale that approaches the diameter of the native collagen fibril. An ultrastructural examination of electrospun collagen reveals a 67 nm

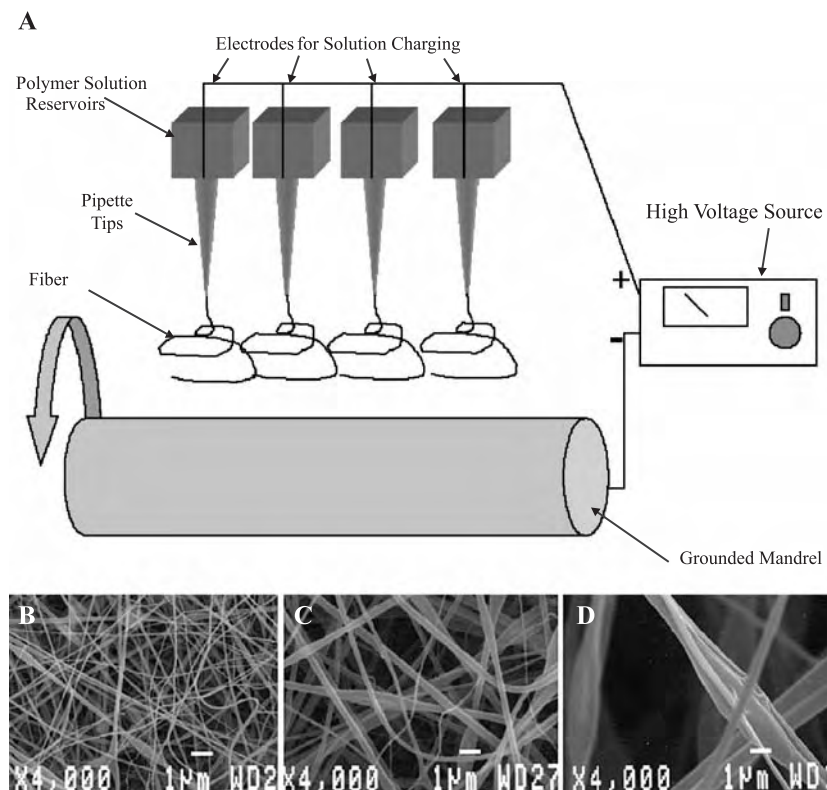


Fig. 1 (A) Schematic of electrospinning apparatus used in the fabrication of scaffoldings for tissue engineering. The key elements of the electrospinning system include a high-voltage power supply, a source reservoir for the polymer, and a grounded mandrel. Our system utilizes a cylindrical target mandrel; however, the electrospinning process can be adapted to produce much more complex shapes. Single and/or multiple polymers can be independently or simultaneously delivered to the electric field from one or more source reservoirs. Electrospinning distinct and unique polymers from separate sources in a temporal sequence can be used to produce a laminated structure. Fiber orientation can be regulated by controlling the motion of the target mandrel with respect to the source solution reservoir. The polarity of the system depicted in this image is arbitrary and, depending on the material to be processed, can be reversed. (B–D) Demonstration that fiber diameter varies as a function of the starting concentration of polymer. These images depict the average fiber diameter obtained by electrospinning collagen at 60 mg (B), 80 mg (C), or 100 mg/ml (D). Note all images were captured at the same magnification (4000 \times).

repeating banded pattern.^[30] In the native collagen fibril, this banding pattern is associated with the formation of specific binding sites that appear to facilitate cell migration. In our biocompatibility studies, electrospun collagen is nonimmunogenic. When implanted within the interstitial space of the rat, vastus lateralis muscle tissue-engineering scaffolds of electrospun collagen are infiltrated by functional capillary beds within seven days.^[32] Similar results were observed in a guinea pig model. In comparison, scaffolds of electrospun poly(glycolic acid) (PGA) initiated a pronounced immune response in these tests, despite the observation that this material exhibited similar fiber diameters and porosity values. We have concluded from these experiments that fibrils of electrospun collagen recapitulate many of the structural and functional properties of the native collagen fibril. An added advantage to electrospun collagen is the observation that collagen has FDA

approval in many fields of use, potentially reducing the obstacles that must be addressed in order to bring this type of scaffolding material to the clinical setting.

IN VIVO MUSCLE ENGINEERING: A FEASIBILITY STUDY

To test our design principles, we have fabricated a muscle prosthesis using electrospun scaffolds of collagen. In preliminary experiments, we have used electrospinning to produce a collagen-based tissue-engineering scaffold. Conditions were optimized to produce individual fibrils that ranged in diameter from 500 to 800 nm. This scaffold was deposited onto a cylindrical target mandrel, producing a cylindrically-shaped construct (4-mm I.D. \times 40 mm in length, with a wall thickness 250–300 microns). The construct was

removed from the target mandrel and hydrated. One end of the construct was sutured shut; the lumen of the cylinder was filled with a suspension of satellite muscle cells isolated from the neonatal rat.^[32] The remaining end of the cylinder was then sutured shut. This technique results in a seamless tube with satellite muscle cells within the central lumen of the cylindrical construct. The outer electrospun scaffolding functioned as a fascial sheath; the terminal ends were slightly extended beyond the enclosed cell population and then flattened to serve as a tendon. The entire bioengineered muscle was then implanted into the adult rat (150 g) in the plane subjacent to the vastus lateralis. The distal ends of the bioengineered construct were sutured to the tendons of origin and insertion for the overlaying vastus muscle. The objective of these experiments was to supplement the quadriceps muscle and produce a quadriceps muscle group with the bioengineered tissue. Within seven days, the implanted tissue exhibited evidence of differentiation and the formation of functional capillaries. After eight weeks, the bioengineered tissue displayed the histological features of differentiated muscle. Parallel arrays of myotubes that were densely packed with myofibrils were evident. The electrospun matrix was substantially remodeled and appeared to be replaced by native materials.

FUTURE DIRECTIONS

As the bioengineering paradigm matures in the fabrication of striated muscle constructs and other tissues, additional hurdles must be addressed. For a skeletal muscle prosthesis, it will be essential to identify rational strategies that can be used to promote the reinnervation, or the formation of nascent and functional nerve–muscle contacts, between the host and the implanted bioengineered muscle. Striated tissue must be placed under voluntary control to truly restore function to a domain of damaged or missing muscle. Without conscious control, the tissue represents little more than filler material. Another issue that requires attention is the development of techniques to increase the overall mass of tissue that can be fabricated on site.

Once again, developmental events and the process of wound repair may provide the insights necessary to achieve these objectives. For example, the terminal ends of a damaged peripheral nerve can resprout, grow in a directed fashion, and reinnervate muscle. This is a common occurrence following injury. The very earliest muscle fabrication experiments demonstrated that muscle bioengineered and fabricated within the fascial sheath of a preexisting muscle bed will undergo innervation.^[15,16] The formation of nascent capillary beds within an injured domain of muscle represents another essential milestone in the regeneration process.

Microvascular endothelial cells are specifically recruited from the surrounding tissue to a damaged domain to restore normal circulation. The processes of neurite outgrowth and angiogenesis are both directed by a variety of soluble factors that can be released in response to an insult that causes structural damage. Our laboratories have begun to examine the efficacy of incorporating selected peptide growth factors directly into tissue-engineering scaffolds during the electrospinning process to modulate the composition of the local microenvironment. Our objective is to use the tissue-engineering scaffold as a solid phase delivery device and establish a gradient of these peptides to direct neurite outgrowth and accelerate capillary formation.

TAILORING THE COMPOSITION OF THE TISSUE ENGINEERING SCAFFOLD

Because electrospinning is carried out at high voltages with very little current, sample heating and the attendant risk of damaging the tertiary structure of peptides is limited. For many growth factors, the solvents used in electrospinning are more likely to induce unfavorable changes in peptide structure than the electric fields used to produce the tissue-engineering scaffold. We have used nerve growth factor (NGF), a peptide that may have efficacy in promoting the directed outgrowth of peripheral neurons, to model how the electro-supplementation of tissue-engineering scaffolds might be used to direct reinnervation. NGF is a soluble member of the neurotrophin family.^[33,34] This growth factor is necessary for normal neural development in the mammalian system and may induce neurite outgrowth in the central and peripheral nervous systems of vertebrates. Receptors for NGF have been detected on motor neurons following injury and during the reinnervation of injured tissue.^[35] Analogous to NGF, vascular endothelial growth factor (VEGF) is a small peptide that plays a role in regulating the proliferation and migration of microvascular endothelial cells.^[36,37] Gradients of VEGF can be used to direct—and accelerate—the formation of capillaries in defined domains.

To characterize the basic release profiles of NGF and VEGF from electrospun collagen, we directly supplemented our electrospinning solutions with 0, 500, or 1000 ng of 7S NGF per 80 mg of collagen per 1 ml of electrospinning solvent (trifluoroethanol, TFE). The 7S NGF isolate was selected for use because this subunit contains the biologically active elements of the peptide. We utilized an intact isolate of a recombinant VEGF peptide in our experiments. Electrospun scaffolds of collagen supplemented with or without NGF were prepared, vapor cross-linked in glutaraldehyde (2%) for 15 minutes, and placed into 1 ml centrifuge

tubes containing 1 ml of sterile phosphate-buffered saline. This type of cross-linking interval stabilizes the collagen fibrils and partially immobilizes the incorporated peptides within the matrix. Aliquots of 100 μ l were taken from each experimental condition at varying intervals (0, 1, 2, 4, 24, 36, and 92 hours). Collected samples were frozen until analyzed. Finally, the amount of NGF present in the different samples was evaluated by ELISA.

In proof-of-concept experiments, the release and recovery of NGF varied as a function of the starting concentration of growth factor present within the electrospun scaffold. The release of NGF from the samples supplemented with 500 and 1000 ng NGF was linear over the initial interval of the assay and appeared to peak at 24 hours (Fig. 2A). To determine if NGF released under these conditions retains bioactivity, we next cultured PC-12 cells in the presence of the electrosupplemented scaffold. In tissue culture, cells of the PC-12 line will flatten and extend neurites in response to a critical concentration of NGF. This characteristic behavior makes this cell line a nearly ideal model to evaluate our novel peptide delivery system.

Assays with PC-12 cells indicate that NGF released from electrospun materials retains sufficient biological activity to induce neurite outgrowth (Fig. 2B–C). To examine how the structure of electrospun collagen might contribute to the sequestration of NGF, we have used electrospun gelatin as a substrate to deliver NGF. Gelatin is identical in chemical composition to acid soluble collagen (both are isolated from calf-skin). However, gelatin is heat denatured during the processes used to prepare this protein for commercial distribution (heated prior to electrospinning and partially degraded). Gelatin and collagen will both electrospin from TFE into submicron diameter fibrils; however, only electrospun collagen exhibits the 67 nm repeat typical of the native collagen fibril. Electrospun gelatin lacks this feature and individual fibers exhibit a nondescript amorphous structure in longitudinal thin sections. In controlled release experiments with electrosupplemented gelatin, very little incorporated NGF was released from this type of matrix. The underlying mechanisms that are responsible for this result remain to be determined. However, it seems clear that NGF associates with the fibrils of electrospun collagen and

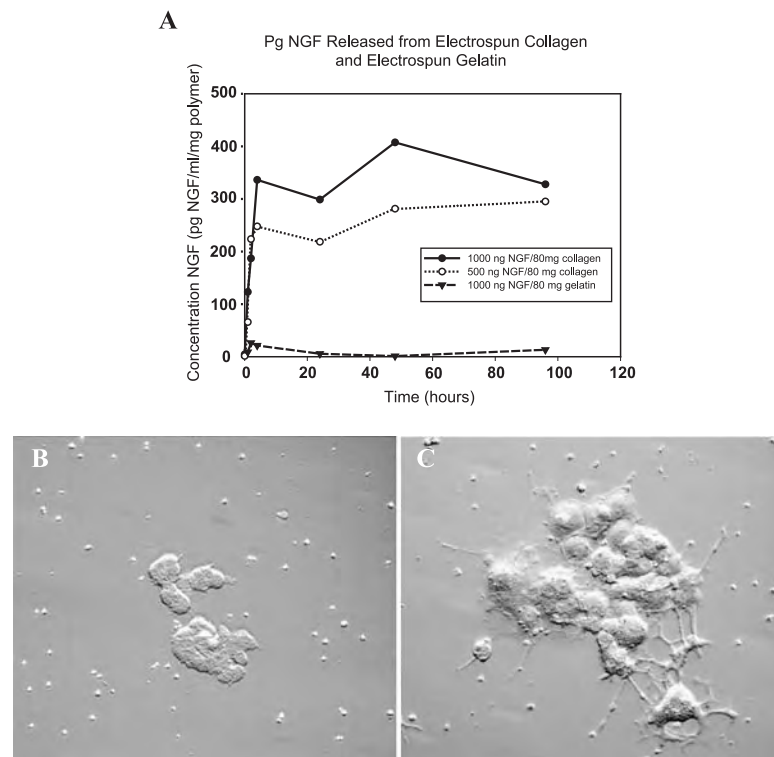


Fig. 2 (A) Representative release profile of NGF from electrospun collagen. Release kinetics are similar regardless of the starting NGF concentration; however, not surprisingly, a higher initial starting concentration results in the release of more total peptide. Panels (B) and (C) illustrate the response of PC-12 cells to NGF released from electrospun collagen. In the absence of NGF, these cells expressed a rounded cell shape in tissue culture. In contrast, after 96 hours of culture in the presence of NGF released from electrospun collagen, extensive neurite outgrowth was evident (C). These data indicate that NGF retains biological activity when delivered from electrospun collagen to the cultured PC-12 cells.

electrospun gelatin in fundamentally different ways. NGF or other neural active peptides may at some time be exploited to bring bioengineered tissues under voluntary control of the central nervous system.

One variable that is difficult to evaluate from our *in vitro* experiments concerns how the release profile might be affected by cellular activity and protease-induced remodeling of an electrospun matrix. This will clearly alter the release kinetics and the actual concentration of growth factor incorporated into the matrix necessary to modulate the microenvironment may be far less (or more) than our *in vitro* experiments might indicate. For example, our preliminary work with materials implanted *in vivo* suggests that 50 ng of VEGF incorporated per 80 mg of collagen (prepared in 1 ml of electrospinning solution) can have a dramatic impact on capillary ingrowth (Fig. 3). As noted, we lightly cross-link electrospun collagen to stabilize the

fibrils. This treatment also links incorporated peptides to the electrospun fibrils. There are risks to this processing strategy. If cross-linking is too extensive, the peptide may be damaged or never released in a biologically relevant pattern to induce microvascular endothelial cell migration. In controlled-release studies conducted *in vitro* with VEGF, a 15-minute interval of vapor fixation substantially reduced the rate of VEGF release from an electrospun matrix of collagen (Fig. 3). These data suggest that fixation intervals could be tailored to modulate the release profile.

To test the biological activity of electrospun VEGF, we have implanted samples of electrosupplemented collagen into the belly of the rat vastus lateralis muscle. Cylinders of electrospun collagen supplemented with or without 50 ng/80 mg collagen were prepared (2 mm I.D. \times 20 mm in length) and placed into a blunt dissected channel within the muscle. In controls

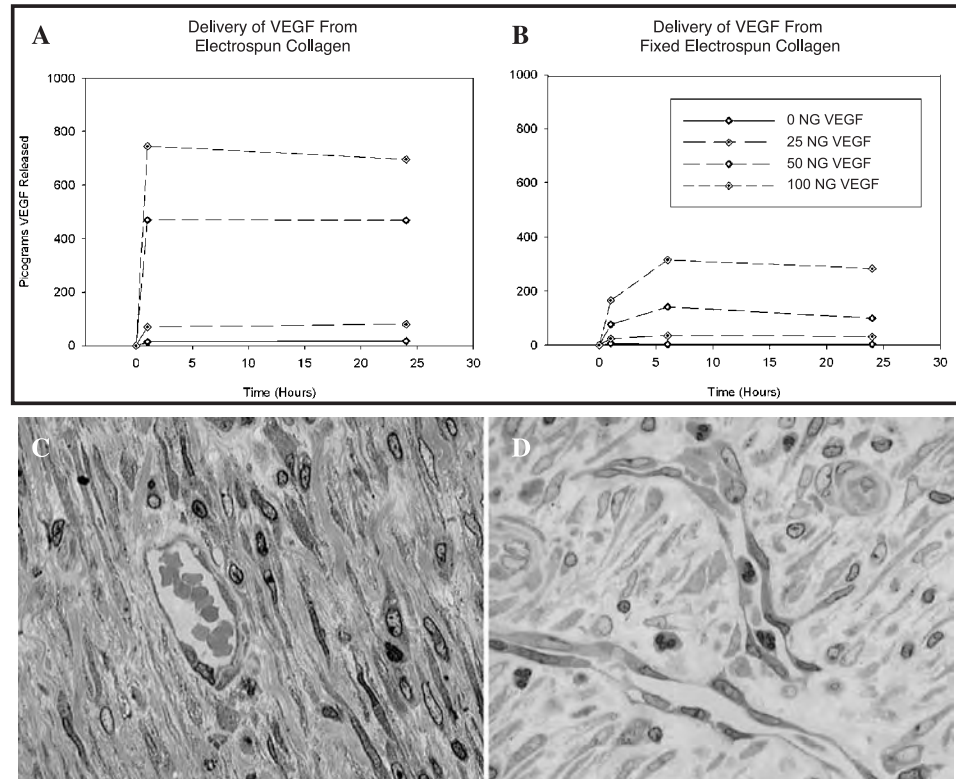


Fig. 3 Representative release profile of VEGF from unfixed (A) and fixed electrospun collagen (B). VEGF was coelectrospun at a rate of 0, 25, 50, or 100 ng/80 mg of collagen. The material was then immersed in PBS and samples were isolated and subjected to ELISA at varying intervals. A second sample of this same material was subjected to a 15-minute interval of glutaraldehyde vapor fixation prior to experimentation. Results indicate that fixation reduces the rate of release; compare relative concentration of VEGF released in panel A versus panel B per unit of time. Panels C and D illustrate the performance of electrospun collagen *in vivo*. Samples were implanted for seven days within the rat vastus lateralis and recovered for microscopic analysis. Panel C depicts electrospun collagen minus VEGF, Panel B electrospun collagen plus VEGF. Note that even unsupplemented collagen supports rapid cellular infiltration and the formation of functional capillaries (C). Panel D depicts the increased frequency of blood vessels in a longitudinal orientation. At recovery, this tissue was visibly more vascular (D).

(un-supplemented with VEGF), interstitial cells from the surrounding tissue freely migrated into the implanted scaffolding. Within seven days, the material was densely populated. Functional capillaries were evident throughout the tissue. During processing for microscopic evaluation, we remove the implanted material and cross-section the samples perpendicular to the long axis of the cylinder and through the wall of the electrospun scaffold. In this orientation, the bulk of the vessels present in control samples exhibit a round or ovoid cross-sectional shape. In contrast to these results, samples recovered from electrospun cylinders of collagen supplemented with VEGF exhibited a higher capillary density. There was also an increased frequency of vessels observed running in a longitudinal orientation with respect to the local axis of the electrospun matrix. These data indicate that supplementing electrospun collagen with VEGF has efficacy in promoting the increased growth of capillaries within a local microenvironment. Delivering effective doses of angiogenic peptides to the local microenvironment may allow larger tissue masses to be fabricated on site in vivo by accelerating the ingrowth of nascent capillaries.

We note that the delivery of VEGF to local domains by gene transfer has been explored to some extent. In these experiments, the local microenvironment of cells genetically engineered to express increased levels of VEGF exhibits increased capillary density.^[38,39] One limitation to this approach, however, concerns how the persistent elevated expression of VEGF might affect the local environment. Many growth factors must be present at a critical concentration over some interval of time to exert biological effects; however, it may not be desirable to have those same peptides permanently present at increased levels. Certainly this issue could be addressed by using inducible promoters; however, supplementing a tissue-engineering scaffold with a defined amount of VEGF protein, or other peptide, may represent a viable alternative to gene transfer.

CONCLUSIONS

The bioengineering process is complex and subject to biological and technical limitations. However, there is considerable evidence to suggest that it is feasible to bioengineer skeletal muscle. From a clinical perspective, it remains an open question where and under what circumstances this type of treatment regime might be undertaken. Until effective methods can be developed to overcome the diffusion constraints that limit the overall mass of tissue that can be fabricated in the bioengineering paradigm, we will be limited to the production of relatively small pieces of muscle. In the near term, this would appear to restrict the use of

engineered skeletal muscle to the treatment of congenital defects and/or injuries affecting the facial musculature. However, as noted, even a traumatic injury to the face is rarely life threatening. This places constraints on the number of patients that might be eligible for this type of reconstructive therapy and thereby limits the incentive to make capital investments in the bioengineering of muscle. A similar argument can be made for a variety of other tissue types. Muscle tissue also might be fabricated and designed solely for use as a platform to deliver specific gene sequences or protein products systemically or to the local microenvironment.^[40,41]

From a basic science perspective, research in the bioengineering of skeletal muscle has, and will continue, to provide valuable insights into the bioengineering process. Using a tissue to model the bioengineering process that is not absolutely essential for life has distinct advantages in this respect. We believe that many of the lessons that can be derived from the bioengineering of skeletal muscle can be applied to other organs. For example, the fabrication of skeletal muscle represents an excellent intermediate step in the construction of more clinically sensitive tissue such as the myocardium. Gene therapy applications using muscle-derived stem cells have also shown promise. These engineered cells may be used to develop treatments for such intractable injuries as lacerated tendons^[42] and nonhealing bone fractures.^[43] The biotechnology industry is poised to expand substantially over the next several decades. Before tissue-engineered organs can become an integral part of clinical practice, issues concerning the distribution of tissue to patients will become the limiting factors rather than the actual construction of the tissue.

ACKNOWLEDGMENTS

This work supported in part by the Department of Defense contract DAMD17-00-1-0512 (Simpson), NanoMatrix Inc. (Simpson, Bowlin, and Wnek), and the Whitaker Foundation (Bowlin). Simpson, Bowlin, and Wnek wish to disclose that they have financial interests in the intellectual property concerning the electrospinning technology described in this report. This electrospinning technology is subject to pending U.S. and International Patents.

ARTICLES OF FURTHER INTEREST

Artificial Muscles; Collagen; Controlled Release; Electrospinning; Matrix Metalloproteinases; Stem Cells; Tissue Engineering Scaffolds

REFERENCES

1. Sanes, J.R. The basement membrane/basal lamina of skeletal muscle. *J. Biol. Chem.* **2003**, *278* (15), 12601–12604.
2. Zacks, S.I.; Vandeburgh, H.H.; Sheff, M.F. Cytochemical and physical properties of myofiber external lamina. *J. Histochem. Cytochem.* **1973**, *21* (10), 895–901.
3. Nakai, J. Skeletal muscle in organ culture. *Exp. Cell Res.* **1965**, *40* (2), 307–315.
4. Ishikawa, H. Formation of elaborate networks of T-system tubules in cultured skeletal muscle with special reference to the T-system formation. *J. Cell Biol.* **1968**, *38* (1), 51–66.
5. Vandeburgh, H.H. Dynamic mechanical orientation of skeletal myofibers in vitro. *Dev. Biol.* **1982**, *93* (2), 438–443.
6. Markey, J.M.; Faulkner, J.A.; Carlson, B.M. Regeneration following transplantation of skeletal muscles in monkeys. *Plast. Reconstr. Surg.* **1978**, *62*, 415–422.
7. Hanson-Smith, F.M.; Carlson, B.M.; Irwin, K.L. Revascularization of the freely grafted digitorum muscle in the rat. *Am. J. Anat.* **1980**, *158*, 65–82.
8. Faulkner, J.A.; Niemeyer, J.H.; Maxwell, L.C.; White, T.P. Contractile properties of transplanted extensor digitorum muscles in cats. *Am. J. Physiol.* **1980**, *238*, C120–C126.
9. Tonken, H.P.; Zhang, F.; Sudekum, A.E.; Siko, P.P.; Newlin, L.; Partington, M.T.; Bunckle, H.J.; Sineaweaver, W.C. Microvascular transplant of the gastrocnemius muscle in the rat. *Microsurgery* **1992**, *42*, 121–124.
10. Blau, H.M.; Webster, C. Isolation and characterization of human muscle cells. *Proc. Natl. Acad. Sci.* **1981**, *78*, 5623–5627.
11. Carlson, B.M. Muscle regeneration in amphibians and mammals: Passing the torch. *Dev. Dyn.* **2003**, *226* (2), 167–181.
12. Vandeburgh, H.H. Cell shape and growth regulation in skeletal muscle: Exogenous versus endogenous factors. *J. Cell. Physiol.* **1983**, *116*, 363–371.
13. Simpson, D.G.; Terracio, L.T.; Terracio, M.; Borg, T.K. Modulation of cardiac phenotype in vitro by the composition and organization of the extracellular matrix. *J. Cell. Physiol.* **1994**, *161*, 89–105.
14. Jankowski, R.J.; Deasy, B.M.; Huard, J. Muscle-derived stem cells. *Gene Ther.* **2002**, *9* (10), 642–647.
15. Carlson, B.M.; Gutmann, E. Development of contractile properties of minced muscle regenerates in the rat. *Exp. Neurol.* **1972**, *36*, 239–249.
16. Carlson, B. Regeneration of the completely excised gastrocnemius muscle in the frog and rat from minced muscle fragments. *J. Morph.* **1968**, *12*, 447–471.
17. Vracko, R.; Benditt, E.P. Basal lamina: The scaffold for orderly cell replacement. Observations on regeneration of injured skeletal muscle fibers and capillaries. *J. Cell Biol.* **1972**, *55*, 406–419.
18. Van Wachem, P.B.; Van Luyan, M.J.A.; Ponte da Costa, M.L. Myoblast seeding in a collagen matrix evaluated in vitro. *J. Biomed. Mater. Res.* **1996**, *30*, 353–360.
19. Okano, T.; Matsuda, T. Hybrid muscular tissues: Preparation of skeletal muscle cell-incorporated collagen gels. *Cell Trans.* **1997**, *6*, 109–118.
20. Feng, Z.; Yamato, M.; Akutsu, T.; Nakamura, T.; Okano, T.; Umezumi, M. Investigation on the mechanical properties of contracted collagen gels as a scaffold for tissue engineering. *Artif. Organs* **2003**, *27* (1), 84–91.
21. Okano, T.; Matsuda, T. Tissue engineered skeletal muscle: Preparation of highly dense, highly oriented hybrid muscular tissue. *Cell Trans.* **1997**, *7*, 71–82.
22. Zilla, P.; Von Oppell, U.; Deutsch, M.J. The endothelium: A key to the future. *Card. Surg.* **1993**, *8*, 32–60.
23. Powell, C.A.; Smiley, B.L.; Mills, J.; Vandeburgh, H.H. Mechanical stimulation improves tissue-engineered human skeletal muscle. *Am. J. Physiol., Cell Physiol.* **2002**, *283* (5), C1557–C1565.
24. Vandeburgh, H.H. A computerized mechanical cell stimulator for tissue culture: Effects on skeletal muscle organogenesis. *In Vitro Cell. Dev. Biol.* **1988**, *24* (7), 609–619.
25. Vandeburgh, H.H.; Karlisch, P. Longitudinal growth of skeletal myotubes in vitro in a new horizontal mechanical cell stimulator. *In Vitro Cell. Dev. Biol.* **1989**, *25* (7), 607–616.
26. Vandeburgh, H.H.; Hatfaludy, S.; Karlisch, P.; Shansky, J. Skeletal muscle growth is stimulated by intermittent stretch-relaxation in tissue culture. *Am. J. Physiol.* **1989**, *256* (3 Pt. 1), C674–C682.
27. Simpson, D.G.; Majeski, M.; Borg, T.K.; Terracio, L. Regulation of cardiac protein turnover and myofibrillar structure in vitro by specific directions of stretch. *Circ. Res. Ultrarapid Commun.* **1999**, *85*, 59E–69E.
28. Telemeco, T.A.; Simpson, D.G. Matrix Metalloproteinases in Tissue Engineering. In *The Encyclopedia of Biomaterials and Bioengineering*, 2003. (In press, this volume).
29. Senger, D.R.; Perruzzi, C.A.; Streit, M.; Kotliansky, V.E.; de Fougerolles, A.R.; Detmar, M. The $\alpha_1\beta_1$ and $\alpha_2\beta_1$ integrins provide critical support for vascular endothelial growth factor signaling, endothelial cell migration, and tumor angiogenesis. *Am. J. Pathol.* **2002**, *160*, 195–204.
30. Matthews, J.A.; Wnek, G.E.; Simpson, D.G.; Bowlin, G.L. Electrospinning collagen nanofibers. *Biomacromolecules* **2002**, *3*, 232–238.
31. Doshi, J.; Reneker, D.H. Electrospinning process and applications of electrospun fibers. *J. Electrostat.* **1995**, *35*, 151–160.
32. Telemeco, T.A.; Bowlin, G.L.; Wnek, G.; Boland, E.; Cohen, N.; Vaida, A.; Tang, D.G.; Baumgarten, C.M.; Mathews, J.; Terracio, L.; Wise, B.; Simpson, D.G. *Electrospun Collagen as a Surgical Biomaterial: Fabrication of a Skeletal Muscle Prosthetic*, 2003. In review Tissue Engineering.
33. Lessmann, V.; Gottmann, K.; Malcangio, M. Neurotrophin secretion: Current facts and future prospects. *Prog. Neurobiol.* **2003**, *69* (5), 341–374.
34. Bradshaw, R.A. Nerve growth factor. *Ann. Rev. Biochem.* **1978**, *47*, 191–216.

35. Wood, S.J.; Pritchard, J.; Sofroniew, M.V. Re-expression of nerve growth factor receptor after axonal injury recapitulates a developmental event in motor neurons: Differential regulation when regeneration is allowed or prevented. *Eur. J. Neurosci.* **1990**, *2*, 650–657.
36. Ferrara, N.; Gerber, H.P.; LeCouter, J. The biology of VEGF and its receptors. *Nat. Med.* **2003**, *9* (6), 669–676.
37. Hopkins, S.P.; Bulgrin, J.P.; Sims, R.L.; Bowman, B.; Donovan, D.L.; Schmidt, S.P. Controlled delivery of vascular endothelial growth factor promotes neovascularization and maintains the limb function in a rabbit model of ischemia. *J. Vasc. Surg.* **1997**, *27*, 886–895.
38. Nabal, E.G. Stem cells combined with gene transfer for therapeutic vasculogenesis: Magic bullets? *Circulation* **2002**, *105*, 672–674.
39. Iwaguro, H.; Yamaguchi, J.; Kalka, C.; Murasawa, S.; Masuda, H.; Hayashi, S.; Silver, M.; Li, T.; Isner, J.M.; Asahara, T. Endothelial progenitor cell vascular endothelial growth factor gene transfer for vascular regeneration. *Circ.* **2002**, *105*, 732–738.
40. Lu, Y.; Shansky, J.; Del Tatto, M.; Ferland, P.; McGuire, S.; Marszalkowski, J.; Maish, M.; Hopkins, R.; Wang, X.; Kosnik, P.; Nackman, M.; Lee, A.; Creswick, B.; Vandeburgh, H. Therapeutic potential of implanted tissue-engineered bioartificial muscles delivering recombinant proteins to the sheep heart. *Ann. N. Y. Acad. Sci.* **2002**, *961*, 78–82.
41. Deasy, B.M.; Huard, J. Gene therapy and tissue engineering based on muscle-derived stem cells. *Curr. Opin. Mol. Ther.* **2002**, *4* (4), 382–389.
42. Pelinkovic, D.; Lee, J.Y.; Engelhardt, M.; Rodosky, M.; Cummins, J.; Fu, F.H.; Huard, J. Muscle cell-mediated gene delivery to the rotator cuff. *Tissue Eng.* **2003**, *9* (1), 143–151.
43. Young, B.H.; Peng, H.; Huard, J. Muscle-based gene therapy and tissue engineering to improve bone healing. *Clin. Orthop.* **2002**, *403* (Suppl.), S243–S251.

Tissue Engineering of Skin

Ricardo R. Brau
Ioannis V. Yannas

*Division of Biological Engineering, Massachusetts Institute of Technology,
and Department of Mechanical Engineering, Massachusetts Institute of Technology,
Cambridge, Massachusetts, U.S.A.*

T

INTRODUCTION

Skin loss can lead to death or to an unacceptable quality of life, and efforts of tissue engineers to synthesize skin have focused on that vital clinical need. The approaches used have varied from those in which an organoid that mimics skin is synthesized in vitro (cultured epidermal autograft, living skin equivalent, living dermal replacement) to those in which a biologically active scaffold, optionally seeded with keratinocytes, is implanted (dermis regeneration template) at the desired anatomical site to achieve in-vivo synthesis (regeneration). Protocols for these apparently contrasting technologies are summarized and compared. Theoretical approaches include the hypothesis that organ regeneration is induced when wound contraction at an injured anatomical site is appropriately blocked by an active scaffold. Further developments in both in-vitro and in-vivo experimental approaches are expected to derive from the additional hypothesis, based on data from two organs (skin and peripheral nerves), that the necessary and sufficient reactants for organ synthesis comprise simply a biologically active scaffold seeded with epithelial cells of the desired organ.

SKIN WOUND HEALING

The integration of the engineering disciplines with the biological sciences has led to many medical advances and new fields of study. One such field, tissue engineering, was recognized by the National Science Foundation in 1987^[1] and defined as the application of the methods of engineering and the life sciences toward the development of substitutes to restore, maintain, or improve biological function. Tissue engineering is a rapidly evolving field. Attempts to substitute organs such as skin, peripheral nerves, kidney, conjunctiva, liver, and others have led to exciting and often satisfactory, but not ideal, results. In this article we describe and compare the various methodologies developed to synthesize skin.

Structure and Function of Skin

The skin is the body's largest organ. Like many other organs in the body, it comprises three structures: epidermis, basement membrane, and dermis. Together, these structures allow skin to perform four known critical physiological roles.^[2] First, it prevents dehydration and invasion from microorganisms and protects from mechanical, chemical, thermal, and ultraviolet insults. Second, it behaves as a signal transducer as it contains receptors for touch, pressure, pain, and temperature. Third, it thermoregulates the body by sweating and vasoconstriction. Fourth, it possesses certain metabolic functions as it synthesizes vitamin D and liberates triglycerides from the subcutaneous adipose stores.^[3] Fig. 1 presents a schematic of some of the structural features of the skin.

The epidermis is the external layer, in direct contact with the environment. It has a thickness of approximately 100 μm and is mainly composed of stratified layers of keratinocyte cells (keratinocytes). In the epidermis, layers of keratinocytes are organized in a maturity gradient in which the innermost (basal) layer comprises relatively immature cells containing negligible amounts of keratin. Cells from the inner layers continuously migrate outward, becoming increasingly mature as well as increasingly filled with keratin until they reach the outer layer (stratum corneum) and are eventually sloughed off. Pigment-containing cells (melanocytes), as well as Langerhans cells that participate in the immune defense of the body, reside in the epidermis as well.

The basement membrane is located between the dermis (stroma) and epidermis (epithelia), providing an anchoring layer for the two tissues. It has a thickness of approximately 100 nm and is primarily composed of collagen types IV and VII. It also contains fibronectin, the proteoglycans heparin sulfate and chondroitin sulfate, as well as entactin, α 1-microglobulin, thrombospondin, and tenascin.^[4] The basement membrane is often characterized by electron microscopy using immunohistochemical markers to detect one or more of several specific proteins, such as laminin or type IV collagen.^[3]

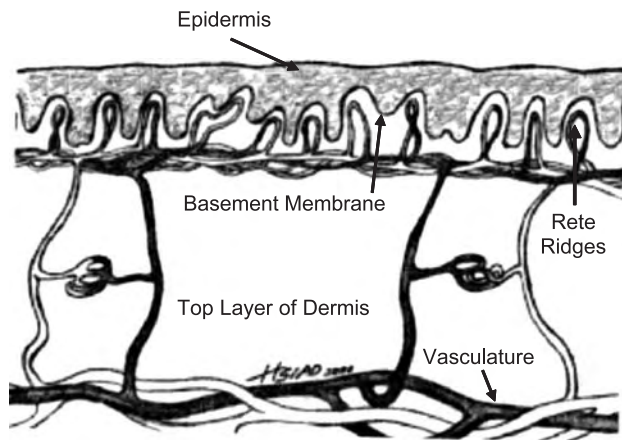


Fig. 1 Schematic diagram of the anatomy of skin. Skin is composed of three layers: epidermis, basement membrane, and dermis. The top portion of the dermis is well vascularized and provides nutrients to the epidermis and basement membrane via capillary loops in rete ridges. (Image adapted from Ref. [2].)

The dermis is the thick, innermost layer of skin, mainly comprising type I collagen and elastin. The combined contributions of these two different types of macromolecules provide robust mechanical support for the epidermis. The dermis contains a rich vascular system that feeds the avascular epidermis and basement membrane and comprises sensory nerves that terminate in the epidermis. Finally, the dermis contains several skin appendages such as hair follicles and sweat and oil-secreting glands.^[3]

Mammalian Response to Skin Injury

Adult mammals respond to injury with contraction and scar synthesis; in the fetal stage, they respond by regeneration. During wound contraction in the adult, specialized contractile cells known as myofibroblasts orchestrate closure of the open wound by exerting forces on the edges of the wound that bring the edges together toward the interior of the injury. Contraction differs from scar formation and regeneration in that there is mechanical deformation of the injured site and there is no extensive synthesis of new tissue. Scar formation is the process by which collagen fibers are synthesized by fibroblasts in the presence of the mechanical field generated by contraction, and leads to filling up of the fraction of the anatomical defect that did not close by contraction. Scar formation and contraction account for wound closure in the adult. However, the resulting tissue structure is the product of an irreversible healing process; it lacks most of the physiological function and normal appearance of normal skin. In contrast, the process of regeneration,

spontaneously available to the fetus during the early stages of gestation, restores all physiological function and appearance; it erases all, or almost all, evidence of an inflicted severe wound and results in reversible wound healing. In the general wound-healing process, closure occurs by contributions from one or more of these three processes; the simple relationship among them is known as the defect closure rule.^[2]

It is very unlikely that the three wound closure processes act independently from each other. However, the extent of their participation is dictated by the severity and timing of the injury along the path of organism development. For example, deep fetal skin wounds that have been inflicted before the third trimester of gestation heal almost flawlessly by regeneration,^[5,6] without contributions from contraction and scar synthesis, while those inflicted after that period generally close by contraction and scar formation with no contribution from regeneration. It has been suggested that fetal healing reflects lack of or suppression of an inflammatory response following injury during the early stages of development.^[7] In contrast, a strong inflammatory response characterizes the response to severe injury of the mammalian adult, where wounds heal irreversibly. The severity of injury is a critical factor as well. In the adult, injury strictly limited to the epidermis and basement membrane heals reversibly, while the dermis (stroma) heals irreversibly.^[2,8]

Need for Skin Substitutes and Current Treatments

The practice of skin replacement dates back to before 1000 B.C. Such practice was aimed at the survival of the injured and is commonly recounted in the popular stories of American Indians who replaced lost skin with that of animals. Today in the United States, more than 80,000 patients a year suffer from severe skin burns that require clinical treatment.^[9] This figure does not include individuals that suffer from venous or diabetic ulcers or those that are subjected to plastic or reconstructive surgery that typically requires the use of skin grafts. When serious enough, these injuries represent a threat to the survival of the organism because they may allow massive bacterial invasion that may lead to sepsis and may induce severe water, protein, and electrolyte loss that can lead to shock. Autografting has provided the standard method when treating full-thickness (completely excised dermis) wounds, corresponding to the severity of injury in a third-degree burn. It has the capability of restoring all functions and structures, but it is limited to fairly small wound areas and has the often significant disadvantage of leaving behind a morbid donor site. Issues of biocompatibility have prevented allogeneic grafts from becoming the

definitive treatment. Furthermore, current treatments for serious skin injuries have a focus beyond survival; they are concerned with restoring organ function as well. In addition to the need for clinical skin replacement with burn patients and those undergoing reconstructive surgery, there is a current need for in-vitro skin models useful in drug testing. In response to such a complexity of requirements, the field of tissue engineering has developed methodologies aimed at synthesis of physiological skin either in vitro or in vivo. A review of such methodologies and results obtained by their use, as well as their advantages and disadvantages, will be presented in the following section.

SKIN EQUIVALENTS AND SKIN REGENERATION TEMPLATES

Recent advances in the methodologies of scaffold synthesis and cell culture have allowed for the development of skin regeneration templates and a number of dermal and epidermal equivalents. Here we review those methodologies that have been at least minimally successful in providing cover for various types of skin wounds and that have led to the development of commercial products for the treatments of patients with severe skin wounds. We refer to these technologies by their commonly used names and acronyms: cultured epidermal autograft (CEA), living dermal replacement (LDR), living skin equivalent (LSE), and dermal regeneration template (DRT). The results of the use of such technologies have been mixed and there is need for further research and experimentation, as the goal of a complete substitution of skin remains elusive. Still, these results are promising and suggest that the ultimate goal of perfect organ replacement is not far ahead. Many of the protocols and results described hereafter have been extensively reviewed.^[2]

Cultured Epidermal Autograft

One of the first approaches used to treat massive skin injuries was that of covering wounds with epidermal grafts synthesized in vitro. These are commonly known as cultured epidermal autografts (CEAs) and have preceded the era of contemporary skin tissue engineering. CEAs were developed as a result of pioneering cell culture work in which the problems of contamination,^[10] viability,^[11] and low cell expansion numbers^[12] were addressed. Later, Rheinwald and Green^[13,14] modified these early techniques for cultivation of enzymatically isolated keratinocytes from biopsies, developing a new methodology that resulted in cultures of stratified, keratinizing keratinocytes with expansion factors of over 10,000 within approximately

three weeks. When treated with CEAs, dermis-free defects have been observed to form a highly differentiated epidermis and basement membrane.^[15]

CEAs have been moderately successful in preventing water and protein loss and protecting the body from infectious agents. The “take” of grafts based on CEAs has been reported to be very good on partial skin wounds, but questionable on full-thickness skin wounds. Two of the disadvantages of CEAs are that they require fairly lengthy cultivation periods after injury and that they are quite fragile during handling. Cultured epidermal allografts have been developed as well,^[16] but they are prone to face biocompatibility issues. In addition, there is no evidence that a normal dermis is synthesized when dermis-free defects are grafted with CEAs, and the lack of mechanical strength of the resulting integument resulting from the use of CEAs on dermis-free defects has been attributed primarily to a lack of anchoring fibrils (reviewed in Ref. [2]). Despite these shortcomings, CEAs are often used as temporary biologically active dressings for chronic open wounds.^[17]

Living Skin Equivalent

An early trend in tissue engineering was an attempt to synthesize the desired tissue structure in vitro followed by transplantation of the missing tissue. The potential of such technology is significant, especially in the preparation of skin substitutes for drug screening applications.

An early approach at synthesizing as much of the structure of skin as possible prior to grafting became known as the living skin equivalent (LSE). LSE consists of two layers that mimic the dermis and the epidermis and are synthesized sequentially in vitro.^[18–21] First, human skin fibroblasts are incorporated into a collagen solution in tissue culture medium. Second, the resulting suspension is gelled at neutral pH and allowed to incubate until the gel is contracted by the fibroblasts. Third, the top surface of the resulting collagen lattice is seeded with keratinocytes and, after several days of expansion, exposed to air to promote the synthesis of a partly keratinized epidermis. Formation of a stratum corneum was promoted in the presence of increased calcium concentrations and in the absence of epidermal growth factor. The resulting epidermis has been characterized by the presence of desmosomes, tonofilaments, and keratohyalin granules;^[22] however, it has been noted that the permeability of this layer is several times higher than that of intact skin.^[23] The cells in the LSE seem to display skinlike behavior, having been shown capable of secreting soluble growth factors such as the interleukins 1, 3, 6, and 8, transforming growth factors α and β , and basic fibroblast growth factors.^[24] However, the LSE lacks a continuous basement membrane in vitro.^[25]

Initial testing of the LSE began on full-thickness skin wounds in the rat.^[22] Both the dermal and epidermal layers underwent considerable remodeling during the early stages of integration. The LSE epidermis appeared to be fully differentiated and functional by the first week after grafting, while its dermal component was observed to be vascularized at two weeks after grafting. No dermal appendages were found, and there were no rete ridges in the dermo-epidermal junction. Grafting onto athymic mice revealed that the presence of LSE initiated synthesis of a basement membrane with occasional anchoring fibrils.^[26] Other studies failed to detect elastic fibers,^[25] although nerve fibers were observed,^[27] in the remodeled neodermis. Clear evidence of synthesis of a physiological dermis has been lacking. Parenteau and coworkers^[23] were able to observe the formation of rete ridges after grafting by lengthening the exposure period to air of the keratinocyte layer during the manufacturing of the graft.

Treatment of human skin injuries with the LSE has provided mixed results. When full-thickness skin wounds arising from burns were treated with the LSE, lysis of at least 20% of the grafted area was observed within two days.^[28] In another study, the tissue from patients who had skin tumors removed and treated with the LSE contained evidence of contraction and scar formation.^[29] However, skin ulcers treated with the LSE closed significantly faster than those treated with traditional methods.^[30]

Living Dermal Replacement

Other attempts at synthesizing skin substitutes *in vitro* have focused on using meshlike polymers as scaffolds for cells. In one such study, the cells were seeded in a synthetic biodegradable polymeric mesh comprising a copolymer of lactic acid (10% wt.) and glycolic acid (90% wt.), referred to as polyglactin-910 surgical mesh (PGL).^[31,32] The mesh was first cultured with neonatal skin fibroblasts until confluence was achieved and several components of the extracellular matrix (ECM), such as collagen, fibronectin, glycosaminoglycans, and growth factors, were secreted. The next step in the manufacturing process involved seeding the mesh with keratinocytes until confluence was achieved in this step as well. The resulting graft became known as living dermal replacement (LDR), and its properties were initially studied by grafting full-thickness skin wounds on athymic mice.

LDR has been reported to be capable of preventing wound contraction in dermis-free defects provided that keratinocytes have been seeded in the polymeric mesh. The inflammatory response to the polymeric mesh was observed to be minimal. It was further reported that,

by 20 days after grafting, the seeded scaffold had been completely degraded and the resulting epithelium had become cornified. No rete ridges had formed. The tissue that has formed underneath the epithelium was reported to contain a large number of fibroblasts and was vascularized.^[31,32] This technology has been used as a temporary wound cover for burn patients who had been previously treated by excision of necrotic tissue.^[33] The LDR has also been placed underneath a meshed autograft to treat human full-thickness skin wounds.^[34] In both experimental and control groups, basement membranes were formed and wounds were observed to close at similar speeds. However, formation of elastic fibers was not observed in the neodermal tissue one year after grafting. In addition, it was noted that the take of the graft was reduced when the polymeric mesh was utilized underneath the autograft.

Dermis Regeneration Template

The final technology that will be reviewed is that pertaining to synthesis and use of the dermis regeneration template (DRT). The DRT has been successful at regenerating most, but not all, of the structures of skin, including a highly functional dermis and epidermis, as well as rete ridges that have not been synthesized using other techniques. It is currently used clinically to treat patients with severe skin wounds as well as in reconstructive surgery applications.

The DRT is an analogue of the extracellular matrix (ECM) comprising a graft copolymer of type I collagen and a sulfated glycosaminoglycan (GAG). Although a large number of ECM analogues have been synthesized over the years, the specific collagen-GAG copolymer referred to as DRT has shown unprecedented biological activity, consisting of blocking of wound contraction and induction of partial skin regeneration. ECM analogues of this type are prepared by mechanically dispersing bovine tendon collagen and chondroitin 6-sulfate in aqueous acetic acid, a good solvent for collagen fibers. The resulting suspension is frozen to a desired temperature in the range -5 to -60°C , and the resulting polymer-containing ice is sublimated under vacuum, leaving behind a porous polymeric network with a characteristic pore volume fraction and average pore diameter. The pore structure of the resulting solid is essentially a negative replica of the ice dendritic structure formed during the freezing step, and therefore the average pore diameter depends critically on the freezing temperature selected during the preceding step. The degradation rate of the ECM analogue can be manipulated by subjecting it to processes for physical or chemical cross-linking. Detailed

methodologies for synthesizing ECM analogues have been published.^[35]

The biological activity of the ECM analogues can be controlled by varying the collagen/GAG ratio, density of covalent cross-links, pore volume fraction, average pore diameter, and pore channel orientation. Appropriate variations in matrix characteristics have been capable of successfully inducing the regeneration of skin,^[36–40] peripheral nerves,^[41,42] and the conjunctiva^[43] in adults. In particular, it has been shown that the biological activity of the DRT is lost when the chemical composition, cross-linking density, and average pore diameter deviate from their optimal levels.^[40] The DRT has been used as a bilayer device, with a thin film of poly(dimethyl siloxane) (silicone elastomer) bonded to its top surface to prevent flow of exudates out of the wound, protect the defect against bacteria, and maintain moisture permeability to nearly physiological levels.^[44] The silicone layer is either spontaneously ejected or is peeled off using negligible force several days after grafting.

A prominent characteristic of DRT is its ability to delay significantly the onset of wound contraction. The area of a standardized full-thickness skin wound in the guinea pig model contracted with a half life of eight days in the absence of the DRT graft but in 25 or more days in the presence of the biodegradable DRT.^[36,38,45] Delay of contraction of a standardized wound has since become a critical assay for biological activity of analogues of ECM.^[40] The kinetics of contraction of full-thickness skin wounds in the guinea pig are illustrated in Fig. 2. Underneath the newly synthesized epidermis in the strongly contracting guinea pig model, histological studies revealed the presence of nonvascular scar tissue in the untreated wounds and a small mass of well-vascularized neodermis in wounds treated with DRT.^[36,38,39,45] The available data from several independent sources have been reviewed, and it has been hypothesized that contraction and regeneration are antagonistically related in the adult mammal.^[2]

Seeding the DRT with keratinocytes, separated from a small skin biopsy, prior to grafting has led to a significantly different product.^[36–39,46] The keratinocyte-seeded DRT was capable of arresting, rather than simply delaying, the onset of contraction about 35 days after grafting and eventually induced synthesis of a skin that was physiological in all respects except for the absence of appendages (sweat glands, hair, etc.).^[36,38] This imperfectly regenerated skin possessed a normal epidermis, basement membrane, dermo-epidermal junction with rete ridges, well-vascularized dermis, elastin fibers, and nonmyelinated nerves.^[36–39,46]

Further studies with the DRT seeded with uncultured keratinocytes were conducted in the swine model, in which full-thickness wounds contract somewhat less than in the rodent models.^[47–50] The newly

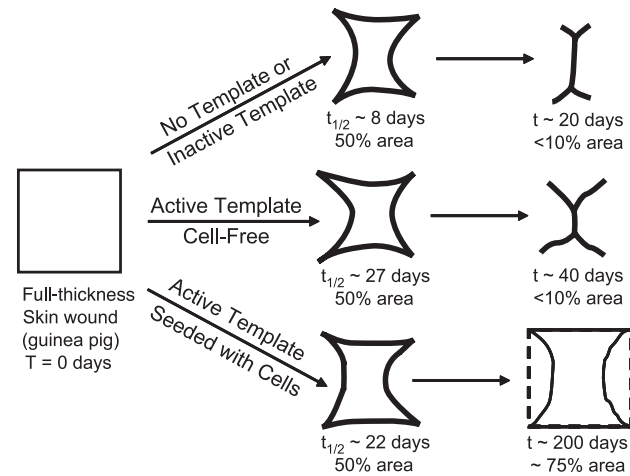


Fig. 2 The kinetics of contraction of full-thickness skin wounds in the guinea pig are significantly affected by the type of template used. A cell-free template delays contraction, while a keratinocyte-seeded one arrests it. The wound half life, $t_{1/2}$, refers to the number of days necessary to reduce the original wound area to 50%. (Image adapted from Ref. [64].)

formed epidermis exhibited fully differentiated cells and rete ridges, and the dermis was well vascularized with collagen bundles resembling those of physiological dermis. Histological staining revealed the presence of many structural components of the dermo-epidermal junction (Fig. 3). Hair follicles and sweat and oil glands were not observed.^[47–50]

Clinical trials of the cell-free DRT have been conducted with patients suffering from massive skin wounds.^[51–53] No evidence of scar was observed post-grafting in both the gross and histological levels. Full-thickness wounds proceeded to heal resembling physiological skin, although skin appendages were missing. The take of the DRT was reported to be good, and only a minimal immune response to it was detected.^[54] Long-term follow-up studies have failed to report complications with the regenerated tissue when used to treat patients with severe burns.^[55]

The DRT has also been used in an attempt to synthesize skin *in vitro* by seeding it with fibroblasts and keratinocytes, leading to a so-called “composite graft.”^[56–59] The epidermis and some of its structures, such as its stratified organization and desmosomal connections, have been successfully recreated. Histological studies of dermis-free defects in the athymic mouse treated with this composite graft provided evidence for the presence of laminin, type IV collagen, anchoring fibrils, and rete ridges in the dermo-epidermal junction. The synthesized dermis in burn patients resembled physiological dermis, although no skin appendages were observed.^[60,61]

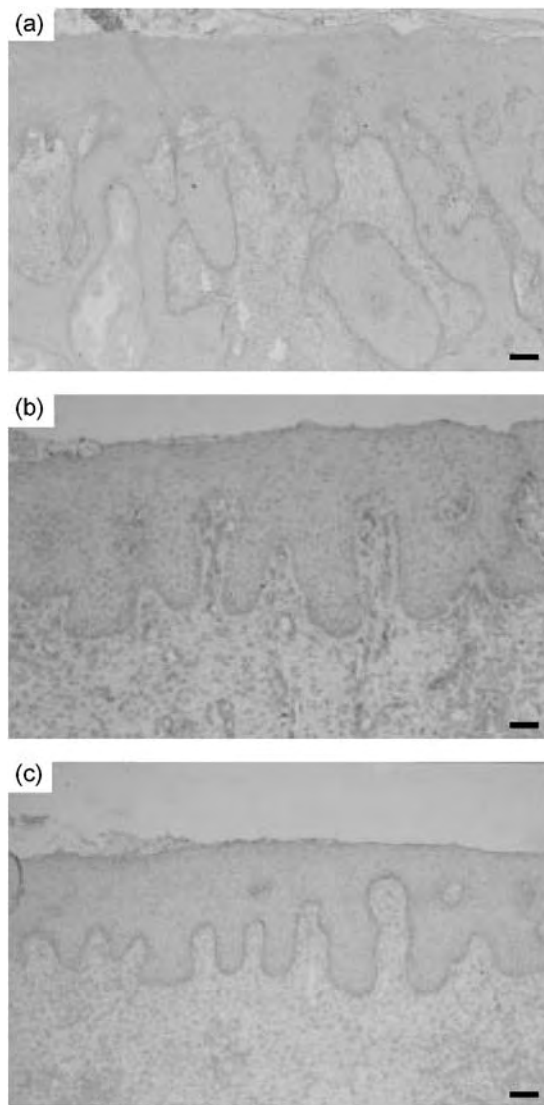


Fig. 3 Histological images of full-thickness swine skin wounds regenerated with the cell-seeded DRT containing evidence of synthesis of the dermo-epidermal junction. (a) Immunohistochemical staining for type VII collagen 12 days after grafting, a marker for anchoring fibrils in the basement membrane. Scale bar, 150 μm . (b) Vascular loops in rete ridges are identified by immunostaining for factor VIII at 35 days after grafting. Scale bar, 75 μm . (c) Hemidesmosomal staining pattern of the regenerated epidermis resembles that of normal skin. Immunostaining for the $\alpha_6\beta_4$ integrin at 35 days after grafting. Scale bar, 100 μm . (All images adapted from Ref. [49].)

THEORETICAL DIRECTIONS OF FUTURE RESEARCH

Skin was the first organ to be partially regenerated and has provided impetus for further studies along the direction of regeneration of other organs. A review of

the available data has led to the hypothesis that organ regeneration in the adult becomes possible provided that the adult processes of wound closure, i.e., contraction and scar synthesis, have been appropriately blocked.^[2] This line of reasoning is supported by several lines of independent evidence, including the observations that the biological activity of the DRT scaffold requires the presence of appropriate ligands for cell-matrix binding, as well as a minimal density of ligands and an optimal duration of the scaffold in a nondegraded form.^[62] These structural features are required for optimal blocking of contraction in standardized animal models. Tests of this hypothesis require studies with organs other than skin, peripheral nerve, and the conjunctiva. Efforts to accelerate the regeneration process and improve the quality of the final product may also involve seeding of stem cells into the scaffold.^[63]

A major question that continuously reappears in the thinking of tissue engineers refers to the choice of an in-vitro versus an in-vivo protocol. In this debate, the relative merits of synthetic protocols that make maximal use of in-vitro conditions to achieve synthesis of an organoid prior to eventual implantation at the correct anatomical site are contrasted to the advantages of using protocols that make minimal use of in-vitro processes and rely instead on in-vivo synthesis via induced regeneration. The advantages and disadvantages have been discussed, with particular emphasis on the lack of a physiological cytokine field (the spatial concentration history of several cytokines during the multiday healing process) in protocols currently conducted in vitro.^[2] In-vitro synthetic approaches make use of simpler reaction conditions but suffer from lack of a theoretical paradigm to guide the selection of reaction conditions. On the other hand, although protocols for regeneration derive systematically from a theoretical paradigm (see preceding discussion), they suffer from the relative experimental complexity encountered when a reaction is conducted in vivo. The two approaches share a required nontrivial final step, consisting of implantation and successful incorporation into the desired anatomical site of the host either of the organoid prepared in vitro or of the scaffold (with or without seeded cells) used with in-vivo protocols. This critical similarity of in-vitro and in-vivo protocols suggests strongly that in future research efforts these two approaches will coalesce, leading to a unified approach that varies in the particulars of the organ being synthesized.

Progress toward a unified synthetic approach in tissue engineering has recently become possible following systematic representation of experimental protocols for tissue and organ regeneration as examples of complex chemical synthetic processes. In this symbolic scheme, each experimental protocol used to

study regeneration is represented as a reaction diagram. Not to be confused with a chemical equation (which includes information on the stoichiometry, i.e., the relative masses of reactants and products), a reaction diagram is a much simpler symbolic representation of the synthetic process. It identifies simply the reactants added by the investigator, the reactor conditions, and the product obtained in each case.^[2] Such a representation simplifies greatly the task of comparing a large variety of experimental strategies used by different investigators and leads easily to identification of the pathway that requires the minimal number of reactants. Systematic use of this approach has led to rules that can be used to synthesize *in vivo* (regenerate) a hypothetical organ consisting of epithelia, basement membrane, and stroma; the reactants required for such a synthesis are simply a scaffold seeded with epithelial cells dissociated from the desired organ that is implanted in an anatomically well-defined site. This protocol has been successfully used to partially synthesize skin, the conjunctiva, and peripheral nerves *in vivo* (regeneration).^[2] The simple protocol suggests the elements of a systematic approach for investigators of *in-vitro* synthetic protocols as well.

ACKNOWLEDGMENTS

The authors acknowledge partial support from the NIH/NIBIB, grant 8 R01 EB 000235, as well as a grant from the Cambridge University–MIT Institute.

ARTICLES OF FURTHER INTEREST

Burn Dressing; Collagen; Extracellular Matrix Scaffolds; Tissue Engineering, Microscale; Tissue Engineering of Peripheral Nerve; Tissue Engineering Scaffolds

REFERENCES

- Nerem, R.M. Tissue engineering in the USA. *Med. Biol. Eng. Comput.* **1992**, *30* (4), CE8–CE12.
- Yannas, I.V. *Tissue and Organ Regeneration in Adults*; Springer-Verlag: New York, 2001.
- Burkitt, H.G.; Young, B.; Heath, J.W. *Wheater's Functional Histology*, 3rd Ed.; Churchill Livingstone: Edinburgh, United Kingdom, 1993.
- Rigal, C.; Pieraggi, M.-T.; Vincent, C.; Prost, C.; Bouissou, H.; Serre, G. Healing of full-thickness cutaneous wounds in the pig I: Immunohistochemical study of epidermodermal junction regeneration. *J. Invest. Dermatol.* **1991**, *96* (5), 777–785.
- Lorenz, H.P.; Longaker, M.T.; Perkocho, L.A.; Jennings, R.W.; Harrison, M.R.; Adzick, N.S. Scarless wound repair: A human fetal skin model. *Development* **1992**, *114* (1), 253–259.
- Mast, B.A.; Diegelmann, R.F.; Krummel, T.M.; Cohen, I.K. Scarless wound healing in the mammalian fetus. *Surg. Gynecol. Obstet.* **1992**, *174* (5), 441–451.
- Martin, P. Wound healing: Aiming for perfect skin regeneration. *Science* **1997**, *276* (5309), 75–81.
- Vracko, R. Basal lamina scaffold: Anatomy and significance for maintenance of orderly tissue structure. *Am. J. Pathol.* **1974**, *77* (2), 313–346.
- Brigham, P.; McLoughlin, F. Burn incidence and medical care use in the United States: Estimates, trends, and data sources. *J. Burn Care Rehabil.* **1996**, *17* (2), 95–107.
- Billingham, R.E.; Reynolds, J. Transplantation studies on sheets of pure epidermal epithelium and epidermal cell suspensions. *Br. J. Plast. Surg.* **1952**, *5*, 25–36.
- Karasek, M.A. Growth and differentiation of transplanted epithelial cell-cultures. *J. Invest. Dermatol.* **1968**, *51* (4), 247–252.
- Igel, H.J.; Freeman, A.E.; Boeckman, C.R.; Kleinfeld, K.L. A new method for covering large surface area wounds with autografts II: Surgical application of tissue culture expanded rabbit-skin autografts. *Arch. Surg.* **1974**, *108* (5), 724–729.
- Rheinwald, J.G.; Green, H. Formation of a keratinizing epithelium in culture by a cloned cell line derived from a teratoma. *Cell* **1975**, *6* (3), 317–330.
- Rheinwald, J.G.; Green, H. Serial cultivation of strains of human epidermal keratinocytes: The formation of keratinizing colonies from single cells. *Cell* **1975**, *6* (3), 331–343.
- Carver, N.; Navsaria, H.A.; Fryer, P.; Green, C.J.; Leigh, I.M. Restoration of basement membrane structure in pigs following keratinocyte autografting. *J. Plast. Surg.* **1993**, *46* (5), 384–392.
- Phillips, T.; Kehinde, O.; Green, H.; Gilchrist, B.A. Treatment of skin ulcers with cultured epidermal allografts. *J. Am. Acad. Dermatol.* **1989**, *21*, 191–199.
- Pomahac, B.; Svenson, T.; Yao, F.; Brown, H.; Eriksson, E. Tissue engineering of skin. *Crit. Rev. Oral Biol. Med.* **1998**, *9* (3), 333–344.
- Bell, E.; Ivarsson, B.; Merrill, C. Production of a tissue-like structure by contraction of collagen lattices by human fibroblasts of different proliferative potential *in vitro*. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76* (3), 1274–1278.
- Bell, E.; Ehrlich, E.P.; Buttle, D.J.; Nakatsuji, T. Living tissue formed *in vitro* and accepted as skin-equivalent tissue of full thickness. *Science* **1981**, *211* (4486), 1052–1054.
- Bell, E.; Ehrlich, E.P.; Sher, S.; Merrill, C.; Sarber, R.; Hull, B.; Nakatsuji, T.; Church, D.; Buttle, D.J. Development and use of a living skin equivalent. *Plast. Reconstr. Surg.* **1981**, *67* (3), 386–392.
- Bell, E.; Sher, S.; Hull, B.; Merrill, C.; Rosen, S.; Chamson, A.; Asselineau, D.; Dubertret, L.; Coulomb, B.; Lapiere, C.; Nusgens, B.; Neveux, Y. The reconstruction of living skin. *J. Invest. Dermatol.* **1983**, *81* (1), 2s–10s.
- Hull, B.E.; Sher, S.E.; Rosen, S.; Church, D.; Bell, E. Structural integration of skin equivalents grafted to



- Lewis and Sprague-Dawley rats. *J. Invest. Dermatol.* **1983**, *81* (5), 429–436.
23. Parenteau, N.; Sabolinski, M.; Prosky, S.; Nolte, C.; Oleson, M.; Kriwet, K.; Bilbo, P. Biological and physical factors influencing the successful engraftment of a cultured human skin substitute. *Biotechnol. Bioeng.* **1996**, *52*, 3–14.
 24. Eaglstein, W.H.; Falanga, V. Tissue engineering for skin: An update. *J. Am. Acad. Dermatol.* **1998**, *39* (6), 1007–1010.
 25. Nolte, C.F.M.; Oleson, M.A.; Hansbrough, J.F.; Morgan, J.; Greenleaf, G.; Wilkins, L. Ultrastructural features of composite skin cultured grafted on athymic mice. *J. Anat.* **1994**, *185*, 325–333.
 26. Bosca, A.R.; Tinois, E.; Faure, M.; Kanitakis, J.; Roche, P.; Thivolet, J. Epithelial differentiation of human skin equivalents after grafting onto nude mice. *J. Invest. Dermatol.* **1988**, *91* (2), 136–141.
 27. English, K.B.; Stayner, N.; Krueger, G.G.; Tuckett, R.P. Functional innervation of cultured skin grafts. *J. Invest. Dermatol.* **1992**, *99* (2), 120–128.
 28. Wasserman, D.; Slotterer, M.; Toulon, A.; Cazalet, C.; Marien, M.; Cherruau, B.; Jaffray, P. Preliminary clinical studies of biological skin equivalent in burned patients. *Burns* **1988**, *14*, 326–330.
 29. Eaglstein, W.H.; Iriondo, M.; Laszlo, K. A composite skin substitute (graftskin) for surgical wounds. *Dermatol. Surg.* **1995**, *21* (10), 839–843.
 30. Falanga, V.; Margolis, D.; Alvarez, O.; Auletta, M.; Maggiacomo, J.; Altman, M.; Jensen, J.; Sabolinski, M.; Hardin-Young, J. Rapid healing of venous ulcers and lack of clinical rejection with an allogeneic cultured human skin equivalent. *Arch. Dermatol.* **1998**, *134* (3), 293–300.
 31. Cooper, M.L.; Hansbrough, J.F.; Spielvogel, R.L.; Cohen, R.; Bartel, R.L.; Naughton, G. In vivo optimization of a living dermal substitute employing cultured human fibroblasts on a biodegradable polyglycolic acid of polyglactin mesh. *Biomaterials* **1991**, *12* (2), 243–248.
 32. Hansbrough, J.F.; Morgan, J.L.; Greenleaf, G.E.; Bartel, R. Composite grafts of human keratinocytes grown on a polyglactin mesh-cultured fibroblast dermal substitute function as a bilayer skin replacement in full-thickness wounds on athymic mice. *J. Burn Care Rehabil.* **1993**, *14* (5), 485–494.
 33. Purdue, G.F. Dermagraft-TC pivotal safety and efficacy study. *J. Burn Care Rehabil.* **1996**, *18*, S13–S14.
 34. Purdue, G.F.; Hunt, J.L.; Still, J.M., Jr.; Law, E.J.; Herndon, D.N.; Goldfarb, I.W.; Schiller, W.R.; Hansbrough, J.F.; Hickerson, W.L.; Himel, H.N.; Kealey, P.; Twomey, J.; Missavage, A.E.; Solem, L.D.; Davis, M.; Totoritis, M.; Gentzkow, G.D. A multicenter clinical trial of a biosynthetic skin replacement, Dermagraft-TC, compared with cryopreserved human cadaver skin for temporary coverage of excised burn wounds. *J. Burn Care Rehabil.* **1997**, *18*, 52–57.
 35. Chamberlain, L.J.; Yannas, I.V. Preparation of Collagen-Glycosaminoglycan Copolymers for Tissue Regeneration. In *Methods of Molecular Medicine*; Morgan, J.R., Yarmush, M.L., Eds.; Humana Press: New Jersey, 1998.
 36. Yannas, I.V.; Burke, J.F.; Warpehoski, M.; Stasikelis, P.; Skrabut, E.M.; Orgill, D.; Giard, D.J. Prompt, long-term functional replacement of skin. *Trans. Am. Soc. Artif. Intern. Organs* **1981**, *27*, 19–23.
 37. Yannas, I.V.; Burke, J.F.; Orgill, D.P.; Skrabut, E.M. Wound tissue can utilize a polymeric template to synthesize a functional extension of skin. *Science* **1982**, *215* (4529), 174–176.
 38. Yannas, I.V.; Burke, J.F.; Orgill, D.P.; Skrabut, E.M. Regeneration of skin following closure of deep wounds with a biodegradable template. *Trans. Soc. Biomater.* **1982**, *5*, 24–27.
 39. Yannas, I.V.; Orgill, D.P.; Skrabut, E.M.; Burke, J.F. Skin Regeneration with a Bioreplaceable Polymeric Template. In *Polymeric Materials and Artificial Organs*; Gebelein, C.G., Ed.; American Chemical Society: Washington, DC, 1984.
 40. Yannas, I.V.; Lee, E.; Orgill, D.P.; Skrabut, E.M.; Murphy, G.F. Synthesis and characterization of a model extracellular matrix which induces partial regeneration of adult mammalian skin. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 933–937.
 41. Chang, A.S.; Yannas, I.V.; Perutz, S.; Loree, H.; Sethi, R.R.; Krarup, C.; Norregaard, T.V.; Zervas, N.T.; Silver, J. Electrophysiological Study of Recovery of Peripheral Nerves Regenerated by a Collagen-Glycoaminoglycan Copolymer Matrix. In *Progress in Biomedical Polymers*; Gebelein, C.G., Ed.; Plenum: New York, 1990.
 42. Chamberlain, L.J.; Yannas, I.V.; Hsu, H.-P.; Strichartz, G.; Spector, M. Collagen-GAG substrate enhances the quality of nerve regeneration through collagen tubes up to level of autograft. *Exp. Neurol.* **1998**, *154* (2), 315–329.
 43. Hsu, W.-C.; Spilker, M.H.; Yannas, I.V.; Rubin, P.A.D. Inhibition of conjunctival scarring and contraction by a porous collagen-GAG implant. *Investig. Ophthalmol. Vis. Sci.* **2000**, *41* (9), 2404–2411.
 44. Yannas, I.V.; Burke, J.F. Design of an artificial skin I. Basic design principles. *J. Biomed. Mater. Res.* **1980**, *14* (1), 65–81.
 45. Yannas, I.V. Use of Artificial Skin in Wound Area Management. In *The Surgical Wound*; Dineen, P., Ed.; Lea & Febiger: Philadelphia, 1981.
 46. Yannas, I.V.; Lee, E.; Orgill, D.P.; Ferdman, A.; Skrabut, E.M.; Murphy, G.F. De novo synthesis of skin. *Proc. Am. Chem. Soc. Div. Polym. Mater.* **1987**, *57*, 28–32.
 47. Butler, C.E.; Orgill, D.P.; Yannas, I.V.; Compton, C.C. Effect of keratinocyte seeding of collagen-glycosaminoglycan membranes on the regeneration of skin in a porcine model. *Plast. Reconstr. Surg.* **1998**, *101* (6), 1572–1579.
 48. Butler, C.E.; Yannas, I.V.; Compton, C.C.; Correia, C.A.; Orgill, D.P. Comparison of cultured and uncultured keratinocytes seeded into a collagen-GAG matrix for skin replacements. *Br. J. Plast. Surg.* **1999**, *52* (2), 127–132.

49. Compton, C.C.; Butler, C.E.; Yannas, I.V.; Warland, G.; Orgill, D.P. Organized skin structure is regenerated in vivo from collagen-GAG matrices seeded with autologous keratinocytes. *J. Invest. Dermatol.* **1998**, *110* (6), 908–916.
50. Orgill, D.P.; Butler, C.; Regan, J.F.; Barlow, M.S.; Yannas, I.V.; Compton, C.C. Vascularized collagen-glycosaminoglycan matrix provides a dermal substrate and improves take of cultured epithelial autografts. *Plast. Reconstr. Surg.* **1998**, *102* (2), 423–429.
51. Burke, J.F.; Yannas, I.V.; Quinby, W.C.; Bondoc, C.C.; Jung, W.K. Successful use of a physiologically acceptable artificial skin in the treatment of extensive burn injury. *Ann. Surg.* **1981**, *194*, 413–428.
52. Heimbach, D.; Luterman, A.; Burke, J.; Cram, A.; Herndon, D.; Hunt, J.; Jordan, M.; McManus, W.; Solem, L.; Warden, G. Artificial dermis for major burns. *Ann. Surg.* **1988**, *208*, 313–320.
53. Stern, R.; McPherson, M.; Longaker, M.T. Histologic study of artificial skin used in the treatment of full-thickness thermal injury. *J. Burn Care Rehabil.* **1990**, *11* (1), 7–13.
54. Michaeli, D.; McPherson, M. Immunologic study of artificial skin used in the treatment of thermal injuries. *J. Burn Care Rehabil.* **1990**, *11* (1), 21–26.
55. Sheridan, R.L.; Hegarty, M.; Hegarty, R.G.; Tompkins, R.G.; Burke, J.F. Artificial skin in massive burns: Results to ten years. *Eur. J. Plast. Surg.* **1994**, *17*, 91–93.
56. Boyce, S.T.; Hansbrough, J.F. Biologic attachment and growth of cultured human keratinocytes onto a graftable collagen and chondroitin 6-sulfate substrate. *Surgeon* **1988**, *103* (4), 421–431.
57. Boyce, S.T.; Christianson, D.; Hansbrough, J.F. Structure of a collagen-GAG substitute optimized for cultured epidermal keratinocytes. *J. Biomed. Mater. Res.* **1988**, *22* (10), 939–957.
58. Cooper, M.L.; Hansbrough, J.F. Use of a composite skin graft composed of cultured human keratinocytes and fibroblasts and a collagen-GAG matrix to cover full-thickness skin wounds on athymic mice. *Surgery* **1991**, *109* (2), 198–207.
59. Cooper, M.L.; Andree, C.; Hansbrough, J.F.; Zapata-Sirvent, R.L.; Spielvogel, R.L. Direct comparison of a cultured composite skin substitute containing human keratinocytes and fibroblasts to an epidermal sheet graft containing human keratinocytes on athymic mice. *J. Invest. Dermatol.* **1993**, *101* (6), 811–819.
60. Hansbrough, J.F.; Boyce, S.T.; Cooper, M.L.; Foreman, T.J. Burn wound closure with cultured autologous keratinocytes attached to a collagen-glycosaminoglycan substrate. *JAMA* **1989**, *262* (15), 2125–2130.
61. Boyce, S.T.; Greenhalgh, D.G.; Kagan, R.J.; Housinger, T.; Sorrell, J.M.; Childress, C.P.; Rieman, M.; Warden, G.D. Skin anatomy and antigen after burn wound closure with composite grafts of cultured skin cells and biopolymers. *Plast. Reconstr. Surg.* **1993**, *91* (4), 632–641.
62. Yannas, I.V. Models of organ regeneration processes induced by templates. *Ann. N.Y. Acad. Sci.* **1997**, *831*, 280–293.
63. Bianco, P.; Robey, P.G. Stem cells in tissue engineering. *Nature* **2001**, *414*, 118–121.
64. Yannas, I.V. Regeneration Templates. In *The Biomedical Engineering Handbook*; Bronzino, J.D., Ed.; CRC Press: Boca Raton, FL, USA, 2000; 113-1–113-18.

Tissue Engineering of the Small Intestine

Simon M. Gabe

Richard M. Day

Biomaterials & Tissue Engineering Group, St. Mark's Hospital, London, United Kingdom

Aldo Boccaccini

Imperial College London, London, United Kingdom

INTRODUCTION

Recent developments in material science and cellular biology have enabled bioengineered tissues to be developed with applications relevant to clinical medicine. Less than a decade ago, many scientists believed that human tissue could only be replaced by a transplant from a compatible donor or with fully artificial parts made of plastic or metal, e.g., joint prosthesis or cardiac pump. Tissue engineering is an interdisciplinary field that applies the principles of engineering and the life sciences toward the development of biological substitutes that restore, maintain, or improve tissue function. The primary goal of all approaches in tissue engineering is the restoration of function through the delivery of living elements that become integrated into the patient. Although some techniques of guided tissue regeneration rely on matrices alone, and other approaches rely on cells alone, most investigators in tissue engineering use cells combined with matrices to achieve new tissue formation. The development of a tissue-engineered intestine requires overcoming obstacles in developing epithelial growth and differentiation on top of a vascularized subepithelial matrix as well as developing an innervated muscular layer.

In tissue engineering, there are two fundamentally different approaches, or philosophies, that can be taken. The first is to replicate the organ anatomically, with the expectation that the function of the engineered organ will therefore be the same. The second approach is simply to replicate its function. In general, researchers who aim to engineer intestine have adopted the anatomical approach and while this has been successful to a large degree, key problems including the development of a muscular layer and neuronal innervation are major challenges to its success. If, on the other hand, the aim is to develop an absorptive surface with neointestinal epithelium, it is possible to be more imaginative about how this can be achieved.

This article will cover the early work that led to the realization that it is possible to develop intestinal neoe epithelium. Scaffold developments are then addressed, followed by the different approaches employed

by investigators. Finally, important new developments in scaffold technology and stem cell biology are covered as well as how this may benefit patients in the future.

Why Engineer Small Intestine?

The most important drive to develop a small intestinal construct is to help patients who have short bowel syndrome (SBS) and are dependent on intravenous nutrition (IVN) in order to survive. SBS is a form of intestinal failure characterized by diarrhea, dehydration, malabsorption, and progressive malnutrition. If the fluids, electrolytes, and nutrients are not replaced, there is deterioration from dehydration and malnutrition to death. SBS occurs after extensive small bowel resection and disturbs the normal absorptive processes for nutrients and fluids. This clinical condition afflicts both adults and children in whom less than one-third of the normal jejunal-ileal length remains. IVN has improved survival, but is associated with serious complications including cholestasis, fibrotic liver disease, and venous thrombosis with progressive loss of venous access as well as catheter-related sepsis.^[1] Patient survival on home parenteral nutrition (HPN) is quoted at 86% at two years and 75% at 5 years.^[1]

In adult patients with short bowel syndrome, a small bowel length of <100 cm is highly predictive of permanent intestinal failure, and the presence of terminal ileum and/or colon in continuity enhances weaning off IVN and improves survival.

Surgical procedures are described to lengthen the small bowel and increase the absorptive surface area in children.^[2] However, these can rarely be performed in adults, as complications such as fistula formation and sepsis are much more frequent in Crohn's disease, which is the most common underlying pathology in this patient group. Possible surgery can be performed to restore continuity in cases where some intestine has been disconnected, but again this applies only to a small number of patients. Small intestinal transplantation is a promising method of treatment and survival rates have improved considerably over the past five

years. Currently, one- and three-year graft survival is 53 and 31%, respectively, and one- and three-year patient survival is 77 and 55%, respectively.^[3] These are still not as good as the survival rates on HPN and patients can suffer serious complications associated with transplantation and immunosuppression including rejection, sepsis, and lymphoproliferative disease. Currently transplantation therefore is reserved for patients who have had severe complications of IVN.^[3,4] Furthermore, in the field of transplantation, the scarcity of donor organs has become a major limitation, especially in the pediatric population. Thus, for patients suffering with intestinal failure and requiring HPN, it would be of considerable benefit to engineer an intestinal construct that could increase absorptive area for nutrient absorption without the need for potent immunosuppression. In addition, other advantages of developing an intestinal construct are that it will improve our understanding about epithelial regeneration, epithelial-mesenchyme interactions, cell-matrix interactions, and the actions of different growth factors in the intestine.

The Development of Primary Cultures of Intestinal Epithelium

In the early 1990s, the maintenance of proliferation of isolated intestinal epithelial cells *ex vivo* was a major limitation. While epithelial cell lines had been developed from colonic adenocarcinoma and from transformed cell lines, these cells had undergone changes and selection to facilitate long-term, *in vitro* growth. These cell lines were not adequate to use to develop an intact or normal epithelium. In order to develop a primary culture of intestinal epithelium, it is necessary to culture cells from the crypt bases. At each crypt base resides epithelial stem cells that divide to produce daughter cells that migrate up toward the crypt mouth and differentiate along their journey.^[5]

In 1992, Evans et al. developed a method for the preparation of intestinal epithelial cell primary cultures in the rodent.^[6] This important article describes a reproducible method for growing small intestinal epithelium derived from the suckling rat intestine in short-term, primary cultures. In this method, organoid units are obtained from neonatal rats, representing the intestinal epithelial crypts (Fig. 1). This group found that the optimal method of isolating small intestinal epithelium was by using a mixture of crude collagenase and dispase to digest the supporting connective tissue. This enzymic solution has a low level of tryptic activity and does not result in complete separation of the epithelial cells, allowing clusters of crypts and villi to be released after gentle pipetting. When cultured, the organoids can be seen to contain heterogeneous

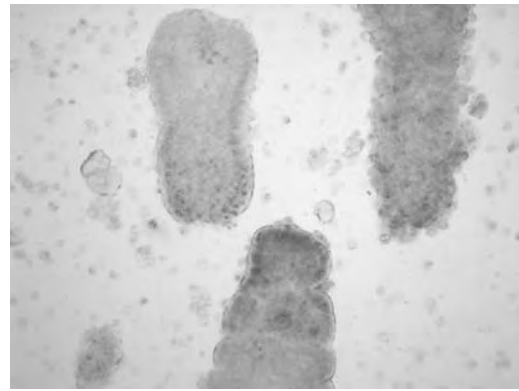


Fig. 1 Organoid units. Intestinal epithelial organoid units are defined as villus structures with overlying epithelium and a core of mesenchymal stromal cells. The average size of an organoid unit is 0.1 mm \times 0.25 mm and, on average, 40,000 organoid units can be isolated from a neonatal rat small intestine.

populations of cells, with fibroblasts and smooth musclelike cells.

In the gastrointestinal tract, multipotential stem cells are found within gastric glands and intestinal crypts. These cells are capable of producing all four main types of epithelial lineages found in the gut: columnar cells, mucin-secreting cells, endocrine cells, and Paneth cells. The isolation and use of organoid units to provide a source of epithelium for tissue-engineered small intestine neotissue has been favored to date because it guarantees the capture of the multipotential stem cells. In addition to the epithelial stem cells, the seeded organoid unit also contains the other cell lineages present within the crypt. A layer of intestinal subepithelial myofibroblasts surrounds each intestinal crypt, which is thought to play a role in epithelial-mesenchymal interactions by secreting growth factors, such as hepatocyte growth factor, and transforming growth factor- β and keratinocyte growth factor. These growth factors can bind to receptors located on the adjacent epithelial cells, allowing paracrine regulation of cell differentiation. Subepithelial myofibroblasts probably play a vital role in supporting adhesion and growth of the isolated epithelial organoid units that are seeded onto scaffolds during the creation of artificial intestinal constructs.^[7,8] The presence of mesenchymal cells is important for epithelial cell proliferation^[9,10] and differentiation.^[11]

The technique described by Evans et al. to isolate primary cultures of intestinal epithelium continues to be used with only minor modifications. However, the key development has been to implant these organoid units *in vivo* in order to develop them into a neoe-pithelium. This has been performed subcutaneously,^[12-16] on mucossectomized colon,^[17] and after

placing the units on a three-dimensional scaffold and connecting them to the native intestine.^[7,8,18–20] In addition, this method has been successfully used to isolate colonic epithelial cells from resected mucosa and biopsy specimens.^[21]

The Scaffolds

Scaffolds requirements and materials

The scientific challenge of tissue engineering encompasses understanding the cells themselves, their mass transport requirements, and biological environment as well as the development of suitable scaffold materials, usually porous, that act as templates for cell adhesion, growth, and proliferation. Several physicochemical and biological requirements have to be fulfilled by the scaffold, depending on the particular tissue under consideration, which are closely dependent upon the scaffold porosity and porous structure. Thus, the scaffold properties should be specifically tailored to enable cell penetration and tissue formation in three dimensions. Ideally, tissue engineering scaffolds should have the following characteristics:^[22] 1) three-dimensional and highly porous with an interconnected pore network for cell growth and flow transport of nutrients and metabolic waste; 2) biocompatible and bioresorbable with a controllable degradation and resorption rate to match cell/tissue growth *in vitro* and/or *in vivo*; 3) suitable surface chemistry for cell attachment, proliferation, and differentiation; and 4) mechanical properties to match those of the tissues at the site of implantation. Moreover, a suitable scaffold should be made from versatile processing techniques that can produce irregular, usually complex, shapes to match that of the defect in the tissue of the patient.

Materials used to fabricate scaffolds for soft tissue engineering are therefore selected to provide temporary structures that exhibit adequate mechanical and biological properties. The use of biodegradable scaffolds avoids the need for a permanent implant made of an engineered material remaining in the tissue. In general, materials chosen for the fabrication of tissue engineering scaffolds belong to the group known as third generation biomaterials, *i.e.*, materials that are both resorbable and bioactive.^[23] These materials are designed to stimulate specific cellular responses at the molecular level. For example, molecular modifications of degradable polymers elicit specific interactions with cell integrins and thereby direct cell proliferation, differentiation, and extracellular matrix production and organization.

Numerous bioresorbable matrices are being investigated as scaffolds for soft tissue engineering and tissue repair, including naturally occurring or

biologically derived matrices,^[24–26] synthetic polymers,^[27–30] and polymer composites.^[31]

Significant examples of natural scaffolds are decellularized biological matrices, which are either allogeneic or xenogeneic tissues that are decellularized by enzymatic or detergent methods. Examples of biologically derived scaffolds are collagen-based matrices and small intestinal submucosa. Other natural polymers considered are hydrogels like gelatin, agar, or fibrin. However, the materials most commonly used as scaffolds for mammalian cells are of synthetic origin.^[25] Advantages of using synthetic scaffolds include the ability to precisely engineer the material's composition and micro- and macrostructure. This allows tailoring of scaffold properties, thus enabling the incorporation of optimal conditions for cell survival, proliferation, and subsequent tissue formation.

Synthetic bioresorbable polymers have been increasingly used as tissue engineering scaffolds during the past 10 years, particularly polylactide (PLA), polyglycolide (PGA), and their copolymers, but also hydrogels, polyhydroxybutyrate, and polyhydroxyoctanoate. PLA- and PGA-based polymers in particular have already demonstrated success in clinical use, for example, as resorbable surgical sutures and meshes or in drug delivery systems, and they have the Food and Drug Administration (FDA) approval for clinical use. The use of bioresorbable synthetic polymers in tissue engineering applications has been reviewed in a number of publications.^[22,27–30]

Porosity and pore structure are key parameters determining the properties and the applicability of scaffolds. Figure 2 shows a summary of the different functions related to the pore structure of a tissue engineering scaffold. In general, scaffold porosity, pore morphology, and pore orientation must be tailored to the particular tissue under consideration. Porous polymeric tissue engineering scaffolds of three-dimensional (3-D) structure have been prepared by numerous techniques, including solid-liquid and liquid-liquid phase separation, solution casting, gel casting, gas saturation, gas foaming, fiber bonding, fabric forming, and related textile-based processes. Porous scaffolds have also been produced by sintering of polymeric microspheres, combined solvent casting and extrusion, emulsion freeze drying, several rapid prototyping techniques, *e.g.*, 3-D printing and fused deposition modeling, as well as by various solvent casting/particulate leaching methods using different porogen additives. Overviews regarding fabrication methods for 3-D porous polymeric scaffolds for tissue engineering are given by several authors.^[22,30,32,33] Regardless of tissue type under consideration, one of the major challenges faced by tissue-engineered devices is promoting healing in three dimensions, which is particularly relevant when complex organ shapes are required,

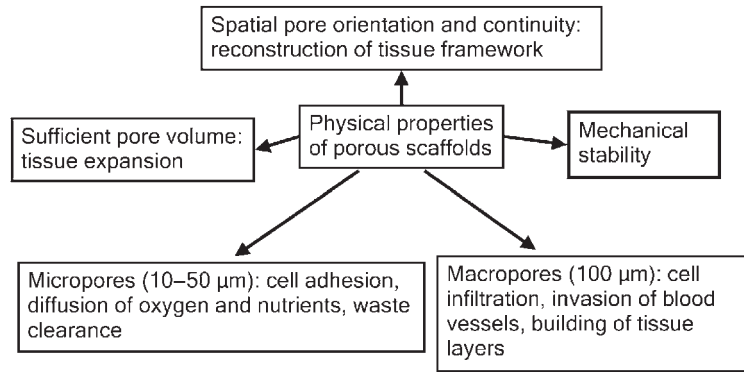


Fig. 2 Schematic diagram showing the different functions of a tissue engineering scaffold, depending on its porosity and pore structure.

T

e.g., multilayered tubular structures such as intestine. Allowing blood vessel formation (neovascularization) through the scaffold is also critical to the success of the approach.

Scaffolds for tissue-engineered intestine

Limited options are available for tissue engineering scaffolds suitable for engineering tissues containing a luminal structure. In fact, there has been relatively limited previous research on developing macrotubular, synthetic biodegradable constructs for tissue engineering applications in comparison with other simple-shaped scaffold architectures. These previous studies have focused mainly on developing scaffolds for vascular tissue engineering, but applications in other areas requiring tubular scaffolds such as peripheral nerve, long bone, and intestine^[34,35] have been considered.

There are a number of challenges related to the development of tubular scaffolds with the desired pore structure, which are related to attaining the optimum flexibility and pliability of the materials used and the maintenance of the tubular shape and structural integrity *in vivo* for the required period after implantation. Intestinal tissue is composed of a layer of smooth muscle sandwiched between a layer of collagenous vascularized support matrix and an epithelial lining. Such a structure can be built by seeding the different cell types for each layer sequentially on suitable porous, degradable scaffolds.

Successful scaffold architectures developed specifically for intestine tissue engineering are those created by Vacanti's group,^[35] which are based on biodegradable polymer tubes, 10 mm in length, with an internal diameter of 2 mm, created from sheets of nonwoven mesh of polyglycolic acid (PGA) fibers (fiber diameter, 5 μm, porosity >0.95, mean pore size, 250 μm) and sprayed on the outer surface with 5% poly-L-lactic (PLLA) acid. These tubes were coated with 200 ml of 0.1% collagen solution and seeded with intestinal epithelial organoid units (6.0104 units/polymer).

Pliable biodegradable polymer foams with potential use as scaffolds for intestinal tissue have been developed by Conley Wake et al.^[34] The materials used were poly(lactic-co-glycolic acid) (PLGA) and poly(ethylene glycol) (PEG) blends and the fabrication technique was based on solvent casting and particulate leaching. A wide range of shear moduli (0.59 to 9.55 MPa), porosities (0.79 to 0.94), and pore diameter (71 to 154 μm) was achieved by varying the processing conditions.

Another recent scaffold structure for tissue-engineered intestine has been developed by Shimizu's group in Japan,^[36,37] which is based on a collagen sponge substrate. The collagen was extracted from porcine skin by enzymatic treatment. The collagen consists of type I and type III atelocollagen. The collagen sponge scaffold was reinforced by polyglycolic acid felt applied to the base. Figure 8 represents schematically the operative procedure followed, which was mainly intended for lengthening small intestine.^[37]

Recent developments

A recent development by the authors has focused on highly porous tubular foam scaffolds fabricated from poly(D,L-lactide-co-glycolide) (PLGA) copolymers. The foams were fabricated following a thermally induced phase separation process, also termed freeze drying.^[30] Different polymer concentrations and casting volumes of solutions were assessed for their suitability to produce porous membranes for the fabrication of tubular constructs. The membrane selected for use in an *in vivo* study contained the largest pores, while retaining its structure. Tubular foam constructs were fabricated by rolling the polymer membranes into a tube. The opposing edges were joined by slowly dissolving in chloroform and pressing the edges together. This resulted in a tube with a lumen approximately 3 mm in diameter and walls approximately 1.5–2 mm thick. Figure 3 shows a scanning electron microscopy (SEM) micrograph of the typical

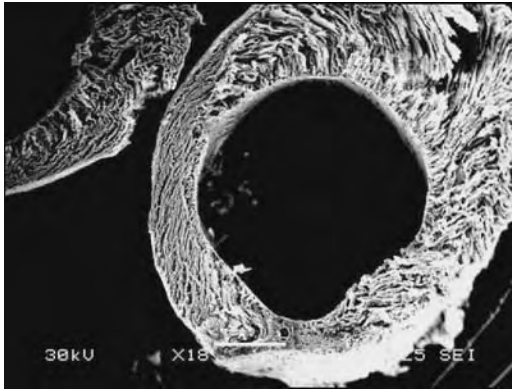


Fig. 3 Scanning electron microscopy (SEM) micrograph of a tubular porous structure formed by rolling the PLGA membranes and dissolving the opposing edges with chloroform.

microstructure of the tubular constructs used. These scaffolds demonstrate good biocompatibility when tested *in vivo* and become rapidly cellularized.

Models

The small intestinal mucosa is an intricate, precisely organized structure composed of both epithelial and mesenchymal tissues, which interact to produce a functional and regulated organ. Different groups have had diverse approaches in developing an artificial layer of small intestinal mucosa. Campbell's group has concentrated on growing organoid units in a subcutaneous plane in mice, while the Vacanti group has grown these organoid units on a three-dimensional scaffold and obtained a cystic structure lined by a neomucosa anastomosed onto the intestine. A different approach has been to lengthen the intestine using a subepithelial matrix, and other groups have focused on developing a suitable matrix that enhances neoeptithelial growth. These different approaches are described subsequently in more detail.

Organoid units implanted subcutaneously

A novel model successfully developed by Campbell's group, primarily as a way of examining small intestinal mucosal regeneration, has been to implant neonatal mouse organoid units subcutaneously. Tait et al. used this method and demonstrated that small intestinal-like structures had developed after 14 days in 39% of cellular transplants to nude mice, and 84% of cellular transplants to inbred rats.^[15] The structures that developed were composed of a circumferential layer of epithelium surrounding a central, mucin-filled lumen and formed crypts and villi. Furthermore, the epithelial

cells had differentiated to form absorptive enterocytes, goblet cells, Paneth's cells, and entero-endocrine cells. Brush border enzyme expression increased as the structures matured and by day 25, the enzyme concentrations were comparable to the control small intestine.^[16] Continuing work from the same unit confirms that subcutaneously implanted organoid units show a temporal sequence of neomucosal morphogenesis and differentiation^[12,14] (Fig. 4). Characterization of the neomucosa revealed that it maintains all the key features of the normal intact intestinal mucosa. In addition, using a mixture of transgene marked and unmarked cells, Slorach et al. demonstrated that neomucosa forms initially from single organoids, which then fuse as they increase in size to generate a chimeric mucosa.^[14]

The next extension of this model has given us a better understanding of some of the factors controlling intestinal epithelial regeneration as well as extending the cells used to human cells. Sattar and colleagues isolated organoid units (disaggregated human fetal intestinal epithelium) from human fetal small intestine and transplanted these units subcutaneously in SCID mice.^[13] The implanted organoid units induced a coordinated regeneration response with separation of cell populations at different stages of histogenesis and cytodifferentiation (Fig. 5), in keeping with the murine model previously described. In addition, growth regulatory genes were assayed and, during early regeneration, Bcl-2, Ki67, and p53 were expressed throughout the epithelial compartment. This indicates suppression of apoptosis (Bcl-2), enhanced proliferation (Ki67), and differentiation (p53). On completion of neomucosal regeneration, these genes were expressed only in crypt epithelium and were absent from the villi.^[13]

The success of this method in generating ectopic intestinal neomucosa can be put down to the cell mix within the organoid units. Patel et al. compared the growth of organoid units with cell suspensions derived from organoid units and demonstrated clearly that the implanted cell suspensions did not develop into small intestinal neomucosa.^[12] This illustrates that the preservation of cell interactions in the organoids is beneficial to the regeneration of mucosa from the crypt stem cells. Earlier work using enterocytes derived from organoid units from 20-day fetal rat intestine developed to form an intestinal-like mucosa when cocultured with fetal mesenchyme prior to implantation into the renal subcapsular space of adult nude mice.^[38] The importance of the mesenchyme in regulating and maintaining endodermal differentiation suggests that mesenchyme has a role in regulating epithelial stem cell function.

These experiments demonstrate that after extensive mucosal disaggregation, postnatal organoid units, which contain small intestinal epithelial progenitor

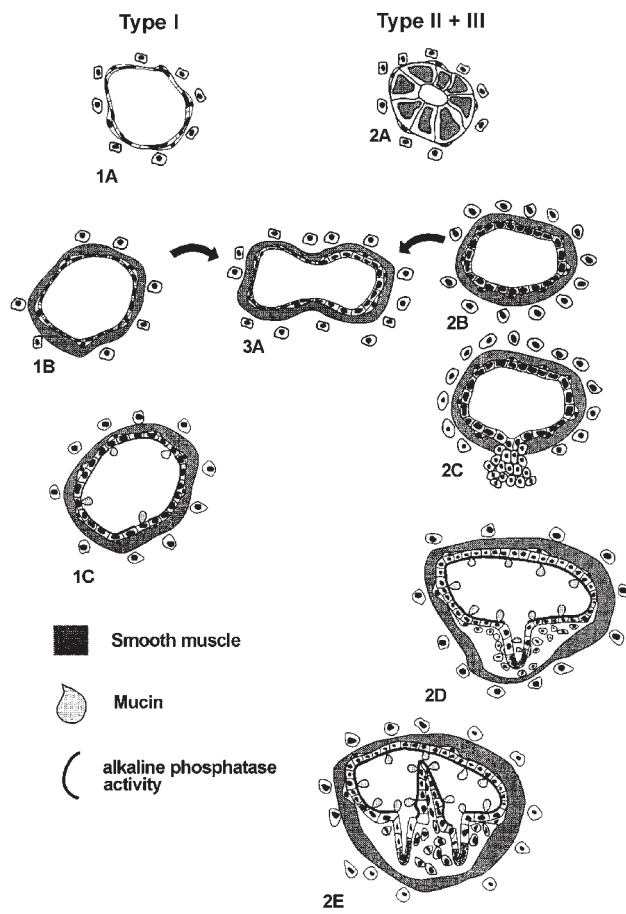


Fig. 4 Model of neomucosal development from grafted organoids. Type I, II, and III cysts are derived from single organoids containing only TA cells, only stem cells, and both TA and stem cells, respectively. (1A) Type I cysts (have a limited ability to proliferate). (1B) Type I cysts increase in size and produce an undifferentiated flattened epithelial lining. (1C) These cells take on a columnar phenotype as they differentiate into absorptive enterocytes, goblet, and entero-endocrine cells. (2A) Type II and III cysts. (2B) Undifferentiated epithelial cells proliferate to produce polarized columnar epithelial cells. (2C) At distinct sites, proliferating epithelium begins to migrate down toward the underlying mesenchyme. The mesenchyme migrates through the layers of smooth muscle cells toward the epithelium. (2D) Cell migration continues, leading to the generation of crypts with Paneth cells at their base. (2E) Regions of epithelium between the crypts protrude toward the central lumen, leading to the generation of intestinal villi and full intestinal mucosal architecture. (3A) Neighboring cysts may fuse together. (Adapted and reproduced from Ref. [14].)

cells, retain the capacity for organotypic regeneration of neomucosa when transplanted to ectopic sites in adult recipients. Importantly, these small aggregates of epithelium and stroma are capable of generating the topographical signals necessary for the three-dimensional regeneration of this tissue. Furthermore, the

multipotent generative potential of the stem cells within these cellular aggregates is maintained with production of all progeny.

Organoid units transplanted onto mucosectomized colon

In an attempt to develop a functional intestinal neomucosa, Campbell's group has seeded organoid units onto the ascending colon after surgical mucosectomy in rats.^[17] They demonstrated that a neomucosa with typical small bowel morphology developed in 76% of grafts. The neomucosa that developed had normal morphology with crypts and villi developing as well cellular differentiation with the formation of enterocytes, goblet cells, entero-endocrine cells, and Paneth cells. Control loops underwent a mucosectomy without organoid implantation and showed no epithelial regrowth.^[17]

This study demonstrates that with appropriate stromal support, transplanted small intestinal stem cells have the capacity to re-epithelialize denuded colonic muscle with small bowel neomucosa.

Organoid units on an artificial matrix

Using the principles of tissue engineering, Vacanti's group studied the transplantation of intestinal cells using synthetic biodegradable polymer scaffolds to generate new intestinal tissue. Initial efforts to tissue engineer small intestine began with a study by Vacanti and colleagues in 1988, in which cell preparations were attached to biodegradable artificial polymers and implanted into animals.^[20] This technique involved harvesting single cells and clusters of fetal and adult small intestinal cells. The cells were seeded onto biodegradable polymers of polyglactin 910, polyanhydrides, and polyorthoester. The fetal intestinal cells proliferated on the polymer scaffolds and after four days in culture, the cell-polymer scaffolds were implanted into host animals (omentum, interscapular fat pad, or mesentery). In three cases, there was successful engraftment of fetal intestinal cells when implanted into the omentum and coupled with a partial hepatectomy. This was encouraging and raised many issues to be addressed including the development of primary cultures of intestinal epithelium, three-dimensional scaffolds, and the role of growth factors.

Subsequently, pure populations of crypt stem cells isolated from adult intestine were transplanted on biodegradable scaffolds to generate stratified epithelium, reminiscent of embryonic gut development.^[18,19] More recently, the same group has shown that transplantation of fetal organoid units on biodegradable scaffolds survive, proliferate, and regenerate small cystic structures

T

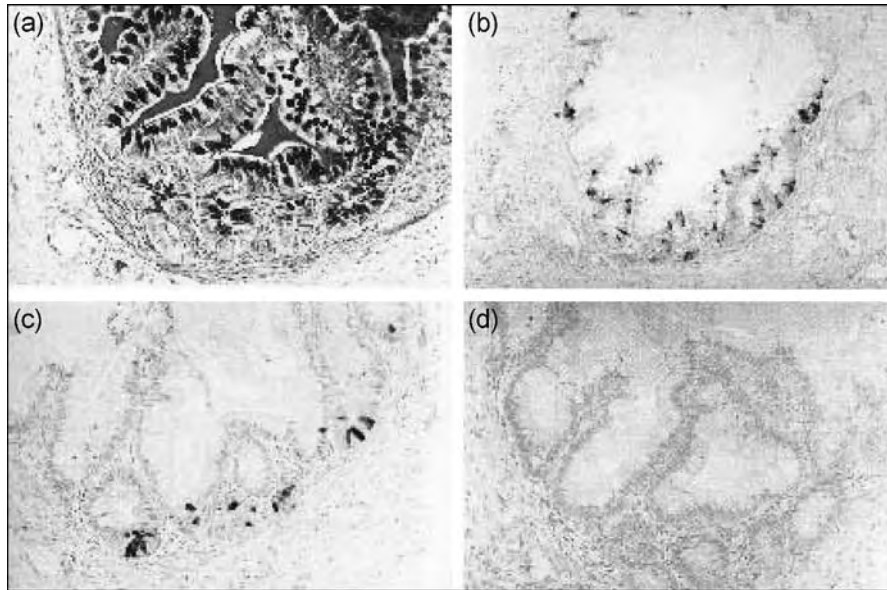


Fig. 5 Cytodifferentiation of human fetal organoid units at 50 days after subcutaneous implantation in SCID mice. (a) Goblet cells identified by PAS/Alcian blue histochemistry; (b) entero-endocrine cells identified by immunohistochemistry for chromogranin; (c) Paneth cells identified by immunohistochemistry for muramidase; (d) absorptive enterocytes identified by Vector red immunohistochemistry for alkaline phosphatase. (Reproduced with permission from Ref. [13].)

lined with intestinal mucosa.^[8] The scaffolds used were tubes created of nonwoven polyglycolic acid fibers (PGA). These allowed nutrient-waste exchange between implanted cells and the surrounding tissue, while allowing neovascularization of the developing tissue. However, PGA fiber meshes are incapable of resisting significant compressional forces. To stabilize the PGA meshes, atomized solutions of poly(L-lactic acid) (PLLA) and a 50/50 copolymer of poly(D,L-lactic-co-glycolic acid) (PLGA) were sprayed over meshes formed into hollow tubes. The PLLA and PLGA coated the PGA fibers and physically bonded adjacent fibers. This increased the compression resistance, and bonding with PLLA allowed the tubes to degrade more slowly.^[39]

Organoid units seeded onto polymer scaffolds can survive, vascularize, proliferate, and form cysticlike structures after implantation. The inner lumen of the cyst becomes lined with a well-developed neomucosal layer characterized by crypt-villous structures and surrounded by smooth muscle.^[7,8] In these studies, the neomucosa that develops within the cysts is histologically normal with cell differentiation, brush border enzymes are present, and electrophysiological testing exhibited similar transepithelial resistance to adult ileal mucosa. In further studies, cysts were anastomosed onto the native intestine after three weeks^[40–43] (Fig. 6). Histological examination demonstrated that 10 weeks after anastomosis, cysts are lined by a neomucosa in continuity to native small bowel across the anastomotic site. The effect of anastomosis increases the villus number, height, and surface length when compared to neomucosa formed in cysts that were not anastomosed onto the native intestine (Fig. 7). Other factors that will stimulate neomucosal growth include small bowel resection^[40] and, to a lesser extent,

portacaval shunting.^[44] In addition, tissue-engineered intestine can develop a mucosal immune system with an immunocyte population similar to that of native small intestine.^[35] This technique has also been applied to the colon to create an artificial pouch lined with colonic mucosa.^[45]

Intestinal extension using a subepithelial matrix

In Japan, Shimizu's group has developed a different technique that exploits the body's wound healing ability to enable tissue to be regenerated at the site of a tissue defect in vivo. They have grafted an acellular collagen sponge, which acts as a scaffold for migration of host cells and tissue,^[37] and have successfully repaired defects created in the trachea,^[46–48] esophagus,^[49–52] and stomach.^[37]

To regenerate a segment of stomach, a 4 × 4 cm square portion of the anterior stomach wall was resected and a silicon sheet used to repair the defect on the luminal side. On the serosal side, a blood-soaked collagen sponge scaffold was sutured and then wrapped with omentum.^[37] Four weeks later, the silicon sheet was removed endoscopically and at this stage, there was an inflammatory infiltrate into the sponge scaffold. After another 12 weeks, gastric mucosa had grown over the scaffold.^[37]

The same technique has been applied to lengthening the small intestine.^[36,37] After forming a jejunal bypass, a 5 cm segment of disconnected jejunum was resected and a silicone tube stent inserted from the luminal aspect. Collagen sponge scaffold was then sutured to the serosal side (Fig. 8) and wrapped with omentum. Four weeks later, the silicon tube was removed

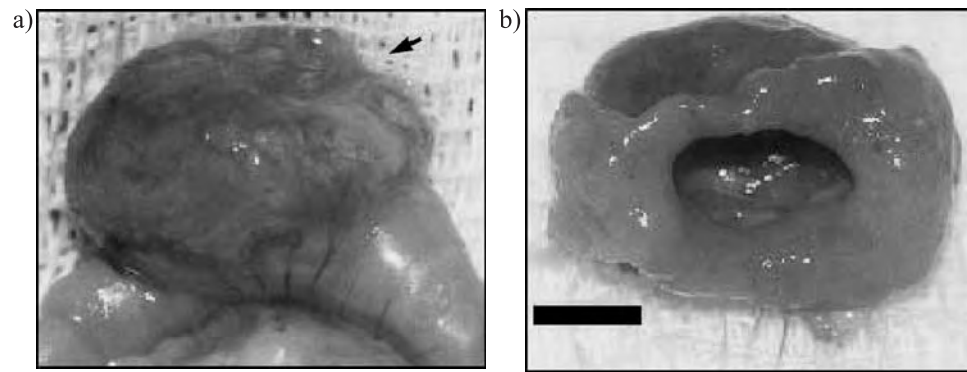


Fig. 6 a) Tissue-engineered neointestine after anastomosis to native small intestine at the time of harvest (arrow indicates tissue-engineered neointestine). b) Patent anastomosis between the tissue-engineered construct and native small intestine (bar = 5 mm). (Reproduced from Ref. [42].)

endoscopically and at this stage, regeneration was observed at the site of the scaffold graft. Twelve weeks after removal of the silicon stent, the mucosal layer covered the luminal surface of the regenerated intestine.^[36] Both of these techniques succeeded in allowing intestinal mucosa to grow over an artificial cellularized scaffold. However, the regenerated segment lacked a muscular layer and this was addressed by seeding mesenchymal stem cells into the developing graft.^[36] Unfortunately, a muscle layer was not generated in this way, as it is likely that pluripotent cells will need specific stimuli to encourage them to develop in a particular direction.

Recent Developments

Stem cells

Following intestinal mucosal irradiation, the majority of crypt cells are lost and there is evidence using microcolony assays to suggest that only a single stem cell is required to regenerate the crypt with all of the intestinal epithelial cell types.^[30,53,54] Investigations of human gastrointestinal stem cells have proved difficult because of a lack of distinct intestinal stem cell

markers. Without these markers, pure populations of stem cells cannot be harvested and used for tissue engineering purposes. It has been hypothesized that colonic stem cells might be isolated and cultured on the basis of specific $\beta 1$ integrin expression patterns.^[55] Cells located in the lower third of crypts expressed higher levels of $\beta 1$ integrin than the remainder of the crypt, a finding similar to the distribution of $\alpha 2\beta 1$ integrin in the small intestine.^[56] Cell sorting based on the increased expression of $\beta 1$ integrin in colonic crypts produced a population of cells with increased ability to form colonies, indicating $\beta 1$ integrin may be a surface marker on the clonogenic cells of the colonic crypt.^[55] Other candidate markers for intestinal stem cells include Musashi-1, a neural binding protein, and Hes1, a transcriptional factor regulated by Musashi-1.^[57,58] These molecules are coexpressed at the putative position of stem cells, in the crypt base, and cells just above the Paneth cells, suggesting they may be markers for epithelial stem and early progenitor cells.

The discovery that stem cells in some tissues are not lineage restricted suggests it may be possible to generate epithelium for tissue-engineered small intestinal constructs from nongastrointestinal tissue. Historically, hematopoietic bone marrow stem cells have been considered to give rise to a single lineage of formed

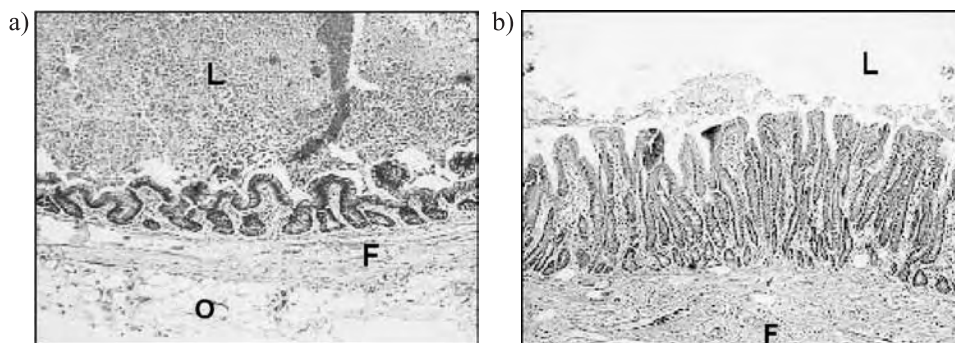


Fig. 7 Neomucosa at 10 weeks after implantation. a) Construct not anastomosed, and b) anastomosed onto the native intestine. P = polymer, L = cyst lumen, O = omentum, F = fibrovascular tissue. (Reproduced from Ref. [42].)

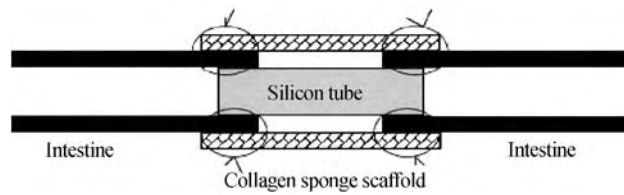


Fig. 8 Illustration of the technique to lengthen small intestine. (Reproduced from Ref. [36].)

elements found in peripheral blood. Recently, the plasticity of adult bone marrow stem cells has been demonstrated in irradiated humans who received bone marrow transplantation. In addition to the formation of new erythroid, granulocyte/macrophage, megakaryocyte, and lymphoid lineages,^[59] hematopoietic stem cells can also differentiate into many different cell types including hepatocytes,^[60] endothelial cells,^[61] renal tubular epithelial cells,^[62] and neuronal cells.^[63] More relevant to small intestinal tissue engineering is the finding that bone marrow-derived cells from male donors were identified in regenerating damaged epithelium in every part of the gastrointestinal tract of female bone marrow transplant recipients.^[64] Importantly, the number of bone marrow-derived gastrointestinal epithelial cells was markedly increased during mucosal repair and epithelial regeneration after graft versus host disease (GVHD) or ulcer formation.^[64] This phenomenon has also been observed with bone marrow-derived pericryptal myofibroblasts in the intestines of human patients suffering from GVHD following bone marrow transplantation.^[65] Stem cell plasticity therefore raises the possibility of seeding tissue-engineered constructs with an individual's own bone marrow stem cells, thus avoiding problems associated with either allograft reactions or autologous donor cell shortages that may arise from a diseased organ. It remains to be seen whether transdifferentiated bone marrow stem cells used for this purpose possess the same functional and sustained regenerative capacity as the endogenous cell types of the target tissue.^[66]

CONCLUSION

The past 15 years have brought many advances in tissue engineering of the small intestine, and different animal models have now been developed in the subcutaneous plane or onto artificial matrices, either using mucosal regeneration or by transplanting organoid units onto the matrix. These models aim to replicate both the anatomy and function of the small intestine and, while it is currently possible to develop an intestinal epithelium on a biodegradable scaffold, there is

more work to be done to increase the developed surface area to be significant in short bowel syndrome. Further developments will also be necessary to develop an innovated muscular layer. Advances in stem cell and scaffold technology and an increasing understanding of growth factors in the intestine will help bring tissue engineering of the small intestine into the clinical arena, where there is a clear clinical need.

ARTICLES OF FURTHER INTEREST

Collagen; Extracellular Matrix Scaffolds; Poly (Glycolic Acid); Poly(lactic acid)s; Stem Cells; Tissue Engineering of the Intestine; Tissue Engineering Scaffolds

REFERENCES

1. Messing, B.; Crenn, P.; Beau, P.; Boutron-Ruault, M.C.; Rambaud, J.C.; Matuchansky, C. Long-term survival and parenteral nutrition dependence in adult patients with the short bowel syndrome. *Gastroenterology* **1999**, *117*, 1043–1050.
2. Bianchi, A. Intestinal loop lengthening—A technique for increasing small intestinal length. *J. Pediatr. Surg.* **1980**, *15*, 145–151.
3. Kato, T.; Ruiz, P.; Thompson, J.F.; Eskind, L.B.; Wepler, D.; Khan, F.A.; Pinna, A.D.; Nery, J.R.; Tzakis, A.G. Intestinal and multivisceral transplantation. *World J. Surg.* **2002**, *26*, 226–237.
4. Thompson, J.S. Surgical aspects of the short-bowel syndrome. *Am. J. Surg.* **1995**, *170*, 532–536.
5. Potten, C.S.; Loeffler, M. Stem cells: Attributes, cycles, spirals, pitfalls and uncertainties. Lessons for and from the crypt. *Development* **1990**, *110*, 1001–1020.
6. Evans, G.S.; Flint, N.; Somers, A.S.; Eyden, B.; Potten, C.S. The development of a method for the preparation of rat intestinal epithelial cell primary cultures. *J. Cell. Sci.* **1992**, *101* (Pt. 1), 219–231.
7. Choi, R.S.; Riegler, M.; Pothoulakis, C.; Kim, B.S.; Mooney, D.; Vacanti, M.; Vacanti, J.P. Studies of brush border enzymes, basement membrane components, and electrophysiology of tissue-engineered neointestine. *J. Pediatr. Surg.* **1998**, *33*, 991–996.
8. Choi, R.S.; Vacanti, J.P. Preliminary studies of tissue-engineered intestine using isolated epithelial organoid units on tubular synthetic biodegradable scaffolds. *Transplant. Proc.* **1997**, *29*, 848–851.
9. Kedinger, M.; Haffen, K.; Simon-Assmann, P. Intestinal tissue and cell cultures. *Differentiation* **1987**, *36*, 71–85.
10. Kondo, Y.; Rose, I.; Young, G.P.; Whitehead, R.H. Growth and differentiation of fetal rat small intestinal epithelium in tissue culture. Relationship to fetal age. *Exp. Cell Res.* **1984**, *153*, 121–134.
11. Sanderson, I.R.; Ezzell, R.; Kedinger, M.; Erlanger, M.; Xu, Z.X.; Pringault, E.; Leon-Robine, S.; Louvard, D.;

- Walker, W.A. Human fetal enterocytes in vitro: Modulation of the phenotype by extracellular matrix. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 7717–7722.
12. Patel, H.R.; Tait, I.S.; Evans, G.S.; Campbell, F.C. Influence of cell interactions in a novel model of post-natal mucosal regeneration. *Gut* **1996**, *38*, 679–686.
 13. Sattar, A.; Robson, S.C.; Patel, H.R.; Angus, B.; Campbell, F.C. Expression of growth regulatory genes in a SCID mouse-human model of intestinal epithelial regeneration. *J. Pathol.* **1999**, *187*, 229–236.
 14. Slorach, E.M.; Campbell, F.C.; Dorin, J.R. A mouse model of intestinal stem cell function and regeneration. *J. Cell. Sci.* **1999**, *112* (Pt. 18), 3029–3038.
 15. Tait, I.S.; Flint, N.; Campbell, F.C.; Evans, G.S. Generation of neomucosa in vivo by transplantation of dissociated rat postnatal small intestinal epithelium. *Differentiation* **1994**, *56*, 91–100.
 16. Tait, I.S.; Penny, J.I.; Campbell, F.C. Does neomucosa induced by small bowel stem cell transplantation have adequate function? *Am. J. Surg.* **1995**, *169*, 120–125.
 17. Tait, I.S.; Evans, G.S.; Flint, N.; Campbell, F.C. Colonic mucosal replacement by syngeneic small intestinal stem cell transplantation. *Am. J. Surg.* **1994**, *167*, 67–72.
 18. Organ, G.M.; Mooney, D.J.; Hansen, L.K.; Schloo, B.; Vacanti, J.P. Transplantation of enterocytes utilizing polymer-cell constructs to produce a neointestine. *Transplant. Proc.* **1992**, *24*, 3009–3011.
 19. Organ, G.M.; Mooney, D.J.; Hansen, L.K.; Schloo, B.; Vacanti, J.P. Enterocyte transplantation using cell-polymer devices to create intestinal epithelial-lined tubes. *Transplant. Proc.* **1993**, *25*, 998–1001.
 20. Vacanti, J.P.; Morse, M.A.; Saltzman, W.M.; Domb, A.J.; Perez-Atayde, A.; Langer, R. Selective cell transplantation using bioabsorbable artificial polymers as matrices. *J. Pediatr. Surg.* **1988**, *23*, 3–9.
 21. Gibson, P.R.; van de, P.E.; Maxwell, L.E.; Gabriel, A.; Doe, W.F. Isolation of colonic crypts that maintain structural and metabolic viability in vitro. *Gastroenterology* **1989**, *96*, 283–291.
 22. Karp, J.M.; Dalton, P.D.; Shoichet, M.S. Scaffolds for tissue engineering. *MRS Bull.* **2003**, *4*, 301–306.
 23. Hench, L.L.; Polak, J.M. Third-generation biomedical materials. *Science* **2002**, *295*, 1014–1017.
 24. Matthew, H.W.T. Polymers for Tissue Engineering Scaffolds. In *Polymeric Biomaterials*; Dumitriu, S., Ed.; Marcel Dekker: New York, 2002; 167–186.
 25. Stock, U.A.; Vacanti, J.P. Tissue engineering: Current state and prospects. *Annu. Rev. Med.* **2001**, *52*, 443–451.
 26. Weng, J.; Wang, M. Producing chitin scaffolds with controlled pore size and interconnectivity for tissue engineering. *J. Mater. Sci. Lett.* **2001**, *20*, 1401–1403.
 27. Behrend, D.; Schmitz, K.-P.; Haubold, A. Bioresorbable polymer materials for implant technology. *Adv. Eng. Mater.* **2000**, *3*, 123–125.
 28. Chaignaud, B.E.; Langer, R.; Vacanti, J.P. The History of Tissue Engineering Using Synthetic Biodegradable Polymer Scaffolds and Cells. In *Synthetic Biodegradable Polymer Scaffolds*; Atala, A., Mooney, D.J., Eds.; Birkhauser: Boston, 1997; 1–14.
 29. Griffith, L.G. Polymeric biomaterials. *Acta Mater.* **2000**, *48*, 263–277.
 30. Maquet, V.; Jerome, R. Design of macroporous biodegradable polymer scaffold for cell transplantation. *Mat. Sci. Forum* **1997**, *250*, 15–42.
 31. Boccaccini, A.R.; Roether, J.A.; Hench, L.L.; Maquet, V.; Jerome, R. A composites approach to tissue engineering. *Ceram. Eng. Sci. Proc.* **2002**, *23*, 805–816.
 32. Thomson, R.C.; Yaszemski, M.J.; Mikos, A.G. Polymer Scaffold Processing. In *Principles of Tissue Engineering*; Lanza, R.P., Langer, R., Chick, W.L., Eds.; R.G. Landes: Austin, TX, 1997; 263–271.
 33. Yang, S.; Leong, K.-F.; Du, Z.; Chua, C.-K. The design of scaffolds for use in tissue engineering. I. Part, Traditional factors. *Tissue Eng.* **2001**, *7*, 679–689.
 34. Conley Wake, M.; Gupta, P.K.; Mikos, A.G. Fabrication of pliable biodegradable polymer foams to engineer soft tissues. *Cell Transplant* **1996**, *5*, 465–473.
 35. Perez, A.; Grikscheit, T.C.; Blumberg, R.S.; Ashley, S.W.; Vacanti, J.P.; Whang, E.E. Tissue-engineered small intestine: Ontogeny of the immune system. *Transplantation* **2002**, *74*, 619–623.
 36. Hori, Y.; Nakamura, T.; Kimura, D.; Kaino, K.; Kurokawa, Y.; Satomi, S.; Shimizu, Y. Experimental study on tissue engineering of the small intestine by mesenchymal stem cell seeding. *J. Surg. Res.* **2002**, *102*, 156–160.
 37. Hori, Y.; Nakamura, T.; Matsumoto, K.; Kurokawa, Y.; Satomi, S.; Shimizu, Y. Tissue engineering of the small intestine by acellular collagen sponge scaffold grafting. *Int. J. Artif. Organs* **2001**, 50–54.
 38. Tait, I.S.; Evans, G.S.; Kedinger, M.; Flint, N.; Potten, C.S.; Campbell, F.C. Progressive morphogenesis in vivo after transplantation of cultured small bowel epithelium. *Cell Transplant* **1994**, *3*, 33–40.
 39. Mooney, D.J.; Mazzoni, C.L.; Breuer, C.; McNamara, K.; Hern, D.; Vacanti, J.P.; Langer, R. Stabilized polyglycolic acid fibre-based tubes for tissue engineering. *Biomaterials* **1996**, *17*, 115–124.
 40. Kaihara, S.; Kim, S.S.; Benvenuto, M.; Choi, R.; Kim, B.S.; Mooney, D.; Tanaka, K.; Vacanti, J.P. Anastomosis between tissue-engineered intestine and native small bowel. *Transplant. Proc.* **1999**, *31*, 661–662.
 41. Kaihara, S.; Kim, S.S.; Benvenuto, M.; Choi, R.; Kim, B.S.; Mooney, D.; Tanaka, K.; Vacanti, J.P. Successful anastomosis between tissue-engineered intestine and native small bowel. *Transplantation* **1999**, *67*, 241–245.
 42. Kim, S.S.; Kaihara, S.; Benvenuto, M.S.; Choi, R.S.; Kim, B.S.; Mooney, D.J.; Vacanti, J.P. Effects of anastomosis of tissue-engineered neointestine to native small bowel. *J. Surg. Res.* **1999**, *87*, 6–13.
 43. Tavakkolizadeh, A.; Berger, U.V.; Stephen, A.E.; Kim, B.S.; Mooney, D.; Hediger, M.A.; Ashley, S.W.; Vacanti, J.P.; Whang, E.E. Tissue-engineered neomucosa: Morphology, enterocyte dynamics, and SGLT1 expression topography. *Transplantation* **2003**, *75*, 181–185.
 44. Kim, S.S.; Kaihara, S.; Benvenuto, M.; Choi, R.S.; Kim, B.S.; Mooney, D.J.; Taylor, G.A.; Vacanti, J.P. Regenerative signals for intestinal epithelial organoid units transplanted on biodegradable polymer scaffolds

- for tissue engineering of small intestine. *Transplantation* **1999**, *67*, 227–233.
45. Grikscheit, T.C.; Ogilvie, J.B.; Ochoa, E.R.; Alsberg, E.; Mooney, D.; Vacanti, J.P. Tissue-engineered colon exhibits function in vivo. *Surgery* **2002**, *132*, 200–204.
 46. Okumura, N.; Nakamura, T.; Shimizu, Y.; Natsume, T.; Ikada, Y. Experimental study of a new tracheal prosthesis made from collagen-grafted mesh. *ASAIO Trans.* **1991**, *37*, M317–M319.
 47. Okumura, N.; Nakamura, T.; Takimoto, Y.; Natsume, T.; Teramachi, M.; Tomihata, K.; Ikada, Y.; Shimizu, Y. A new tracheal prosthesis made from collagen grafted mesh. *ASAIO J.* **1993**, *39*, M475–M479.
 48. Yamamoto, Y.; Nakamura, T.; Shimizu, Y.; Matsumoto, K.; Takimoto, Y.; Kiyotani, T.; Sekine, T.; Ueda, H.; Liu, Y.; Tamura, N. Experimental replacement of the thoracic esophagus with a bioabsorbable collagen sponge scaffold supported by a silicone stent in dogs. *ASAIO J.* **1999**, *45*, 311–316.
 49. Natsume, T.; Ike, O.; Okada, T.; Takimoto, N.; Shimizu, Y.; Ikada, Y. Porous collagen sponge for esophageal replacement. *J. Biomed. Mater. Res.* **1993**, *27*, 867–875.
 50. Takimoto, Y.; Nakamura, T.; Yamamoto, Y.; Kiyotani, T.; Teramachi, M.; Shimizu, Y. The experimental replacement of a cervical esophageal segment with an artificial prosthesis with the use of collagen matrix and a silicone stent. *J. Thorac. Cardiovasc. Surg.* **1998**, *116*, 98–106.
 51. Takimoto, Y.; Okumura, N.; Nakamura, T.; Natsume, T.; Shimizu, Y. Long-term follow-up of the experimental replacement of the esophagus with a collagen-silicone composite tube. *ASAIO J.* **1993**, *39*, M736–M739.
 52. Yamamoto, Y.; Nakamura, T.; Shimizu, Y.; Matsumoto, K.; Takimoto, Y.; Kiyotani, T.; Sekine, T.; Ueda, H.; Liu, Y.; Tamura, N. Experimental replacement of the thoracic esophagus with a bioabsorbable collagen sponge scaffold supported by a silicone stent in dogs. *ASAIO* **1999**, *45*, 311–316.
 53. Potten, C.S.; Hendry, J.H. *Radiation and Gut*; Elsevier: Amsterdam, 1995; 45–59.
 54. Withers, H.R.; Elkind, M.M. Microcolony survival assay for cells of mouse intestinal mucosa exposed to radiation. *Int. J. Radiat. Biol. Relat. Stud. Phys. Chem. Med.* **1970**, *17*, 261–267.
 55. Fujimoto, K.; Beauchamp, R.D.; Whitehead, R.H. Identification and isolation of candidate human colonic clonogenic cells based on cell surface integrin expression. *Gastroenterology* **2002**, *123*, 1941–1948.
 56. Beaulieu, J.F. Differential expression of the VLA family of integrins along the crypt-villus axis in the human small intestine. *J. Cell. Sci.* **1992**, *102* (Pt. 3), 427–436.
 57. Kayahara, T.; Sawada, M.; Takaishi, S.; Fukui, H.; Seno, H.; Fukuzawa, H.; Suzuki, K.; Hiai, H.; Kageyama, R.; Okano, H.; Chiba, T. Candidate markers for stem and early progenitor cells, Musashi-1 and Hes1, are expressed in crypt base columnar cells of mouse small intestine. *FEBS Lett.* **2003**, *535*, 131–135.
 58. Potten, C.S.; Booth, C.; Tudor, G.L.; Booth, D.; Brady, G.; Hurley, P.; Ashton, G.; Clarke, R.; Sakakibara, S.; Okano, H. Identification of a putative intestinal stem cell and early lineage marker; musashi-1. *Differentiation* **2003**, *71*, 28–41.
 59. Bhatia, M.; Wang, J.C.; Kapp, U.; Bonnet, D.; Dick, J.E. Purification of primitive human hematopoietic cells capable of repopulating immune-deficient mice. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 5320–5325.
 60. Alison, M.R.; Poulson, R.; Jeffery, R.; Dhillon, A.P.; Quaglia, A.; Jacob, J.; Novelli, M.; Prentice, G.; Williamson, J.; Wright, N.A. Hepatocytes from non-hepatic adult stem cells. *Nature* **2000**, *406*, 257.
 61. Gao, Z.; McAlister, V.C.; Williams, G.M. Repopulation of liver endothelium by bone-marrow-derived cells. *Lancet* **2001**, *357*, 932–933.
 62. Poulson, R.; Forbes, S.J.; Hodivala-Dilke, K.; Ryan, E.; Wyles, S.; Navaratnasah, S.; Jeffery, R.; Hunt, T.; Alison, M.; Cook, T.; Pusey, C.; Wright, N.A. Bone marrow contributes to renal parenchymal turnover and regeneration. *J. Pathol.* **2001**, *195*, 229–235.
 63. Eglitis, M.A.; Mezey, E. Hematopoietic cells differentiate into both microglia and macroglia in the brains of adult mice. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 4080–4085.
 64. Okamoto, R.; Yajima, T.; Yamazaki, M.; Kanai, T.; Mukai, M.; Okamoto, S.; Ikeda, Y.; Hibi, T.; Inazawa, J.; Watanabe, M. Damaged epithelia regenerated by bone marrow-derived cells in the human gastrointestinal tract. *Nat. Med.* **2002**, *8*, 1011–1017.
 65. Brittan, M.; Hunt, T.; Jeffery, R.; Poulson, R.; Forbes, S.J.; Goldman, J.M.; Goldman, J.; Alison, M.R.; Wright, N.A. Bone marrow derivation of pericryptal myofibroblasts in the mouse and human small intestine and colon. *Gut* **2002**, *50*, 752–757.
 66. Anderson, D.J.; Gage, F.H.; Weissman, I.L. Can stem cells cross lineage boundaries? *Nat. Med.* **2001**, *7*, 393–395.

Tissue Engineering of Tendons

Dominique P. Pioletti

Center for Orthopedic Research, Swiss Federal Institute of Technology, Hôpital Orthopédique de la Suisse Romande, and Bone Bioengineering Group, STI/BIO-E/CRO, Lausanne, Switzerland

Olivier Siegrist

Pierre-Yves Zambelli

Hôpital Orthopédique de la Suisse Romande, Lausanne, Switzerland

T

INTRODUCTION

Rupture of tendons due to either high load or laceration is a common injury, especially for the hand. Tendon disorders are traditionally treated by reducing inflammation, restoring flexibility, and, if necessary, performing surgical repair.^[1] It has been estimated that hand surgeons repair nearly one-third of a million digital flexor tendon injuries per year in the United States.^[2] Despite the fact that surgery of tendons is a common procedure, its outcome is often unsatisfactory. Recovery after tendon surgery can take months. At best, the restored tendon is about half of its initial mechanical properties, so surgeons often face patients' dissatisfaction. On the other hand, if left untreated, healing tendon is not as good as original structure from a mechanical point of view. Moreover, the related pain is difficult to treat.

Several factors may explain these unsatisfactory results such as the difficulty to anchor the tendon in muscle, the high mechanical constraint on the repair structure, or the unavailability of ideal tendon grafts (currently, tendon grafts are either auto, allo, or synthetic). It has been shown that early mobilization following tendon surgery gives superior results.^[3] The repaired tendon should then be able to support immediate loading. However, suturing techniques to joint split ends of tendons, while providing sufficient mechanical strength to limit gapping, are inadequate to carry normal loads.

Based on the shortcomings of the actual tendon surgery, tendon tissue engineering may offer an alternative approach. The objective in developing an engineered tendon will then be the reconstruction of the injured structure with structural and functional properties as good as the original. Moreover, an engineered tendon could reduce the rate of secondary problems by making the surgical procedure less extensive.

This article gives first some general details on anatomy and physiological properties of tendons, followed by a list of requirements for an ideal engineered

tendon. A review of the actual state of the art in tendon replacement is then performed, and finally, some remarks on the future of tissue engineering of tendons are proposed.

ANATOMICAL AND PHYSIOLOGICAL PROPERTIES OF TENDONS

Tendons are dense, regularly organized, and by their arranged collagen fibers, allow connection between muscles and bones. The structure is composed of long arranged collagen fibers, and the function and behavior of tendons depend on of their mechanical properties.

They are capable of resisting important strength in order to transmit the muscular power to the bones and articulations. They protect the muscle (fascia) and fix the length of the muscle belly. They also play a role of damper and shock absorber.

A tendon is composed of multiple fascicles surrounded by a fascicular membrane and septa. Inside the fascicle, we find a mix of fibroblasts in relative central position and fibrils. The fibrils are composed of microfibrils, which themselves are made of tropocollagen.

The fascicular membrane is the endotenon, which is draped over by a sliding structure, the peritendon. Between the peritendon and the endotenon run vessels, nerve, and lymphatics. Depending on the localization, the organization of the surrounding layers is more complex.

A hand flexor tendon is encircled by a well-defined sheath covered by synovial cells. In this situation, the external layer (paratenon) is called tenosynovium, but if there is no synovial lining, the paratenon is called tenovagium.

At the musculo-tendinous junction, the perimysium is continuous with the endotenon. At the other extremity on the tendon–bone interface, the collagen fibers and the endotenon become continuous with the periosteum.

The tendon insertion into bone is of two types. The first one described by Cooper and Misol,^[4] shows a transition from the tendon to a layer of fibrocartilage with digitations in the periosteum. The fibrocartilage is progressively ossified and merges with bone. The second one is more complex and is characterized by having the periosteum as an intermediate layer. Superficial fibrils are fixed into the periosteum and deeper, directly into bone.

Tendon is largely vascularized, the blood flow coming from the perymysium, periosteum attachment, and surrounding tissues. In particular, the perfusion through the surrounding tissue reaches the tendon penetrating the paratenon, the mesotenon, or vincula. We can separate vascularized from avascular tendons: vascularized tendons concern those with a rich surrounding vascularization, which penetrate directly inside the endotenon. The avascular tendons are contained in a sheath, and the vessels are penetrating through thin structure (mesotendon or vincula), which reticulate and bind the tendon.

The innervation plays a sensory role through mechanoreceptors located near the musculo-tendinous junction.

REQUIREMENTS FOR TISSUE ENGINEERING OF TENDON

Several requirements need to be fulfilled in the tissue engineering of tendons. Some are common to most engineered tissues used for applications in musculo-skeletal applications, but some are specific for the tissue engineering of tendons.

Common Requirements in Tissue Engineering

1. It is usually stated that an engineered tissue is composed of three elements: scaffold, cells, and growth factors. Tissue engineering is then, by definition, multi-disciplinary. In order to bring useful and adequate solutions, it is especially important that physicians, who will be the users of the developed engineered tissue, are involved at every step of the project development. For example, surgical procedures should be anticipated since the beginning of the project to insure that the developed tissue will be manageable in the constraint environment of a surgical room.
2. As a basic rule, the engineered tissue should be easy to use and its manipulation should be as close as possible to the actual standard surgical procedures.
3. The engineered tissue will need to be approved by governmental agencies (e.g., FDA) before being available on the market. Especially if cells and/or

growth factors are used, it is important to be aware of the tests required by these governmental agencies to obtain an approval for clinical use. This would avoid the repetition of experiments not following established standards.

Specific Requirements for Tendon Tissue Engineering

1. Tendon auto-grafts are freely available during surgery. Due to economical constraints on the health care system, it would be difficult to propose an expensive engineered tendon. The only economical factors, which may favor the use of an engineered tendon from the health insurance point of view, would be the reduction of the operative time and a faster patient recovery. Based on these economical considerations, the production and stock management of the engineered tendons should be considered from the beginning of the project.
2. Tendon biomechanics is complex due to its non-linear stress-strain relationship and viscoelasticity (Fig. 1).^[5,6] In-vivo forces quantification of tendons is also a useful approach for the determination of mechanical design parameters.^[7] An engineered tendon should be able to reproduce these mechanical behaviors. It should be noted that different tendons might present different mechanical properties as has been shown for fatigue damage.^[8]

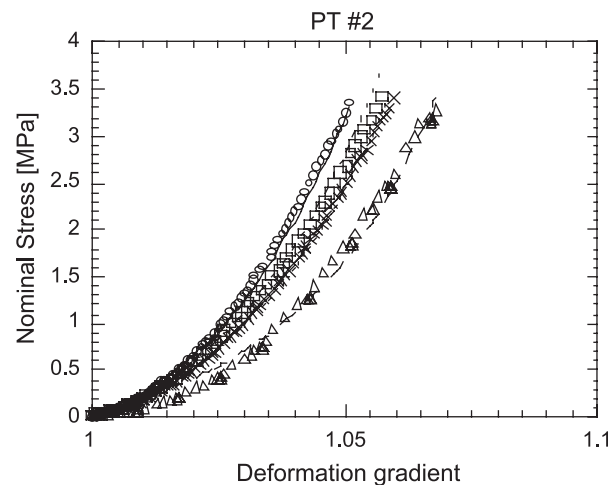


Fig. 1 Experimental and theoretical stress-strain curves obtained at four different rates of elongation for one human patellar tendon specimen. Good correlation between experimental and theoretical curves are found. The proposed constitutive law with only three parameters is then able to fit nonlinear stress-strain curves obtained at different strain rates. (Reproduced with permission from Ref. [6].)

3. Fixation on bone and muscle may be the weakness mechanical part of the engineered tendon. It is then important to incorporate this point in the design of the tendon. For example, it has to be decided early in the project if the scaffold will be sutured or glued to the muscle. In case of high mechanical constraints (early loading), it would be more advisable to suture than to glue the tendon. However, the drawback of tendon suturing is the possibility of inducing a devascularization of the tendon and a corresponding necrosis of the tissue. It should be noted that if left untreated, some tendons might spontaneously heal, suggesting that the gluing process of tendon to muscle is a normal way of healing.
4. Tendon cells must communicate to sustain growth and matrix expression.^[9] It has been shown that cells in the epitenon and internal compartment of whole tendon are connected physically to each other and express gap junctions.^[10] Successful engineered tendon replacements must include designs that allow for cell-to-cell connectivity to allow intercellular communication. Scaffolds should then be easily seeded by cells and encourage cell division and matrix expression.^[9]
5. If tendon fibroblasts are used, it has to be noted that these cells are site-dependent and may induce different results regarding proliferation and/or gene expression.^[11,12] Autologous cells' transplantation has the advantage of not inducing immune reaction and is more easily implemented from a regulatory point of view. However, an initial biopsy followed by a cellular expansion period is necessary, rendering this approach time consuming. Disadvantages of autologous cells' transplantation disappear with allogous cells' transplantation, but the benefit of no immune reaction and easier regulatory implementation are, on the other hand, lost. A careful selection of the cells used, depending on the application, should then be performed.

ACTUAL APPROACHES FOR TISSUE ENGINEERING OF TENDON

In the term tissue engineering, we consider all the engineering and biological approaches used in the development of an artificial tendon. Different solutions proposed for treating tendon rupture are presented from the scaffold, cells, and growth factor/gene therapy points of view.

Scaffold

In order to bridge tendon lesions, synthetic scaffolds have been used with relatively modest success.^[13]

Synthetic tendon made of Dacron™, nylon, or even carbon fibers has been used. One of the only synthetic scaffolds remaining on the market is the Leeds-Keio ligament made of polyethylene terephthalate (PET). Recently, it has been shown that treatment with radiofrequency (RF)-generated glow discharge increased cell attachment and proliferation on this PET implant.^[14]

Organic polymers' matrix, such as acellular collagen, have also been employed as tendon graft.^[15] These scaffolds are supposed to provide early mechanical support and promote cellular infiltration. However, achilles tendon repairs using collagen scaffold achieved only 36% of normal maximum force.^[16] Naturally occurring extracellular matrices were also proposed for tendon repair. There are limited locations within the body that can be used as the source of these extracellular matrices. One is derived from porcine small intestinal submucosa.^[17] With this tissue used as tendon scaffold, a transient and significant weakening of mechanical properties was found compared to patellar tendon in the first months after implantation in an animal study. After 12 months, no difference was found. In order to increase the mechanical properties of the scaffold, collagen scaffold with poly-L-lactic acid was developed.^[18] Mechanical properties were increased with this method. Probably one of the most promising scaffold developments from biomechanical and biocompatible points of view would be based on biomimetic strategy. In this approach, the scaffold should include polymeric collagen as the fundamental fibrous phase and being cross-linked to give the mechanical strength of the engineered tendon. To this end, derived from skate egg capsule, a collagen-based scaffold was developed.^[11] Fibers produced from pepsin-solubilized, bovine tendon type I collagen were polymerized with di-catechol nordihydroguaiaretic acid (NDGA). These NDGA cross-linked fibers showed high mechanical properties comparable to native tendon and demonstrated excellent biocompatibility properties.^[19]

Cells

As for other engineered tissues (e.g., bone, cartilage), incorporation of cells is supposed to increase the healing potential of the tissues. A combination of scaffold and cells is then developed, the scaffold playing either the role of cell delivery system alone or cell delivery system and mechanical support.

Use of autologous mesenchymal stem cells seeded in collagen gel *in vitro* and then implanted in the body has been proposed.^[20] The results indicated that delivering mesenchymal stem cell-contracted, organized collagen implants to large tendon defects could

significantly improve the biomechanics, structure, and probably the function of the tendon after injury.^[21] Mesenchymal stem cells, which have the potential to differentiate in (tendon) fibroblasts, hold the promise to be used for allogous cells' transplantation. However, if the cells are not oriented in the gel, only modest improved biomechanical properties are observed compared to natural repair of unfilled defect.^[22] The orientation of the cells in the gel, in particular due to the gel contraction mediated by cells, improved the biomechanical properties.^[23] Besides mesenchymal stem cells, it may be useful to use differentiated fibroblasts in order to speed up the healing process. Moreover, fibroblasts seeded on a scaffold and in vitro mechanically stimulated the mechanical properties of the graft.^[24] So, it has then been proposed that the delay between rupture and surgery may be used to engineer the tendon in vitro by applying mechanical stimulus. This approach may, however, be difficult to apply from an economical point of view. In order to amplify the potential of cells in the healing process, autologous mesenchymal stem cells might be genetically modified to produce growth factors useful in the healing process of tendon.^[25]

Growth Factor/Gene Therapy

Use of growth factors is related to a good understanding of the tendon healing process. This healing process could be viewed as four overlapping phases: hemorrhagic, inflammatory, proliferation, and remodeling/maturation. The second phase, inflammatory, is probably a key phase as many different growth factors and cytokines are produced and will drive the outcome process of healing. A detailed description of these phases may be found in Ref. [26]. It has been shown that treatment with growth factors such as PDGF and IGF-1 stimulate DNA synthesis of tendon fibroblasts.^[27] However, as for other tissues, issues such as the difficulty to deliver the growth factors or the supraphysiologic concentrations needed to induce an effect^[28] need to be solved. A potential solution might be the use of gene therapy. Genetically modified cells^[25,27] or gene-activated matrix^[29] might be used to target specific location (ruptured zone of the tendon) and specific biological process (healing of the tendon). This approach is still in its infancy, but promising results have already been obtained.

FUTURE DEVELOPMENT FOR TISSUE ENGINEERING OF TENDONS

As mentioned by Hildebrand,^[25] the regeneration of normal tendon has not been achieved and further refinements are necessary for these treatments. Indeed,

none of the actual solutions followed all the requirements noted earlier for tissue engineering of tendons. Especially, problems like fixations of tendons, incorporation of biomechanical properties, choice of scaffold with adequate degradation properties, or packaging of the engineered tendon^[20] still need to be solved. Nature used millions of evolution years to optimize the functionality of the tendons. It might then be worthy to use this knowledge in order to engineer tendons. The combination of biomimetism, as proposed by Koob^[11] to engineer tendon scaffold, and gene therapy, either with cells, e.g., Ref. [25], or with gene activated matrix^[29] to target specific healing tendon process, held probably the most promising development in tissue engineering of tendon.

CONCLUSION

Important progresses have been achieved recently in tissue engineering. Although no ideal solution exists for tissue engineering of tendons, it is reasonable to believe that good solutions will emerge from laboratories. In particular, the combination of biomimetism and gene therapy with mesenchymal stem cells opens new fields of possibilities from a scientific and clinical point of view. However, the development of tissue engineering of tendons can only be successful if the challenging economical constraints are satisfied. Based on the past industrial experiences in tissue engineering for bone or cartilage, it would be an illusion to believe that engineered tendon with cells may be entirely supported by private industries. It seems more reasonable to develop collaborations between industries and hospitals. The industries would be responsible for developing and managing the scaffolds, and the hospitals would be responsible for managing the cells' bank and seeding process of the scaffolds with the selected cells. With this kind of strategy, it is hoped that, in the future, surgeons will no more face the paradox of treating a ruptured tendon by collecting another tendon in the patient.

ARTICLES OF FURTHER INTEREST

Collagen; Fabrics; Tendons and Ligaments, Mechanical Testing Of; Tissue Engineering of Bone; Tissue Engineering of Ligament; Tissue Engineering of Rotator Cuff Tendons; Tissue Engineering Scaffolds

REFERENCES

1. Kuwada, G.T. Diagnosis and treatment of Achilles tendon rupture. *Clin. Podiatr. Med. Surg.* **1995**, *12* (4), 633–652.

2. Pennisi, E. Tending tender tendons. *Science* **2002**, *295* (5557), 1011.
3. Kubota, H.; Manske, P.R.; Aoki, M.; Pruitt, D.L.; Larson, B.J. Effect of motion and tension on injured flexor tendons in chickens. *J. Hand Surg. [Am.]* **1996**, *21* (3), 456–463.
4. Cooper, R.R.; Misol, S. Tendon and ligament insertion. A light and electron microscopic study. *J. Bone Jt. Surg. Am.* **1970**, *52* (1), 1–20.
5. Pioletti, D.P.; Rakotomanana, L.R. Non linear viscoelastic law for soft biological tissues. *Eur. J. Mech.* **2000**, *19*, 749–759.
6. Pioletti, D.P.; Rakotomanana, L.R.; Benvenuti, J.F.; Leyvraz, P.F. Viscoelastic constitutive law in large deformations: Application to human knee ligaments and tendons. *J. Biomech.* **1998**, *31* (8), 753–757.
7. Juncosa, N.; West, J.R.; Galloway, M.T.; Boivin, G.P.; Butler, D.L. In vivo forces used to develop design parameters for tissue engineered implants for rabbit patellar tendon repair. *J. Biomech.* **2003**, *36* (4), 483–488.
8. Ker, R.F. The implications of the adaptable fatigue quality of tendons for their construction, repair and function. *Comp. Biochem. Physiol., Part A, Mol. Integr. Physiol.* **2002**, *133* (4), 987–1000.
9. Banes, A.J.; Weinhold, P.; Yang, X.; Tsuzaki, M.; Bynum, D.; Bottlang, M.; Brown, T. Gap junctions regulate responses of tendon cells ex vivo to mechanical loading. *Clin. Orthop.* **1999**, *367 Suppl.*, S356–S370.
10. Mcneilly, C.M.; Banes, A.J.; Benjamin, M.; Ralphs, J.R. Tendon cells in vivo form a three dimensional network of cell processes linked by gap junctions. *J. Anat.* **1996**, *189* (Pt. 3), 593–600.
11. Koob, T.J. Biomimetic approaches to tendon repair. *Comp. Biochem. Physiol., Part A, Mol. Integr. Physiol.* **2002**, *133* (4), 1171–1192.
12. Vogel, K.G. Breakout session 5: Tendon and ligament. *Clin. Orthop.* **1999**, *367 Suppl.*, S371–S374.
13. Mooney, D.J.; Mikos, A.G. Growing new organs. *Sci. Am.* **1999**, *280* (4), 60–65.
14. Rowland, J.R.; Tsukazaki, S.; Kikuchi, T.; Fujikawa, K.; Kearney, J.; Lomas, R.; Wood, E.; Seedhom, B.B. Radiofrequency-generated glow discharge treatment: Potential benefits for polyester ligaments. *J. Orthop. Sci.* **2003**, *8* (2), 198–206.
15. Goldstein, J.D.; Tria, A.J.; Zawadsky, J.P.; Kato, Y.P.; Christiansen, D.; Silver, F.H. Development of a reconstituted collagen tendon prosthesis. A preliminary implantation study. *J. Bone Jt. Surg. Am.* **1989**, *71* (8), 1183–1191.
16. Kato, Y.P.; Dunn, M.G.; Zawadsky, J.P.; Tria, A.J.; Silver, F.H. Regeneration of Achilles tendon with a collagen tendon prosthesis. Results of a one-year implantation study. *J. Bone Jt. Surg. Am.* **1991**, *73* (4), 561–574.
17. Badylak, S.; Arnoczky, S.; Plouhar, P.; Haut, R.; Mendenhall, V.; Clarke, R.; Horvath, C. Naturally occurring extracellular matrix as a scaffold for musculoskeletal repair. *Clin. Orthop.* **1999**, *367 Suppl.*, S333–S343.
18. Ide, A.; Sakane, M.; Chen, G.; Shimojo, H.; Ushida, T.; Tateishi, T.; Wadano, Y.; Miyanaga, Y. Collagen hybridization with poly(L-lactic acid) braid promotes ligament cell migration. *Mater. Sci. Eng.* **2001**, *17*, 95–99.
19. Koob, T.J.; Hernandez, D.J. Material properties of polymerized NDGA–collagen composite fibers: Development of biologically based tendon constructs. *Biomaterials* **2002**, *23* (1), 203–212.
20. Butler, D.L.; Awad, H.A. Perspectives on cell and collagen composites for tendon repair. *Clin. Orthop.* **1999**, *367 Suppl.*, S324–S332.
21. Young, R.G.; Butler, D.L.; Weber, W.; Caplan, A.I.; Gordon, S.L.; Fink, D.J. Use of mesenchymal stem cells in a collagen matrix for Achilles tendon repair. *J. Orthop. Res.* **1998**, *16* (4), 406–413.
22. Awad, H.A.; Butler, D.L.; Boivin, G.P.; Smith, F.N.; Malaviya, P.; Huibregtse, B.; Caplan, A.I. Autologous mesenchymal stem cell-mediated repair of tendon. *Tissue Eng.* **1999**, *5* (3), 267–277.
23. Awad, H.A.; Butler, D.L.; Harris, M.T.; Ibrahim, R.E.; Wu, Y.; Young, R.G.; Kadiyala, S.; Boivin, G.P. In vitro characterization of mesenchymal stem cell-seeded collagen scaffolds for tendon repair: Effects of initial seeding density on contraction kinetics. *J. Biomed. Mater. Res.* **2000**, *51* (2), 233–240.
24. Goulet, F.; Rancourt, D.; Cloutier, R.; Germain, L.; Poole, A.R.; Auger, F.A. Tendon and ligament. In *Principle of Tissue Engineering*, 2nd Ed.; Lanza, R., Langer, R., Vacanti, J., Eds.; Academic Press, 2000; 711–722.
25. Hildebrand, K.A.; Jia, F.; Woo, S.L. Response of donor and recipient cells after transplantation of cells to the ligament and tendon. *Microsc. Res. Tech.* **2002**, *58* (1), 34–38.
26. Woo, S.L.; Hildebrand, K.; Watanabe, N.; Fenwick, J.A.; Papageorgiou, C.D.; Wang, J.H. Tissue engineering of ligament and tendon healing. *Clin. Orthop.* **1999**, (367 Suppl.), S312–S323.
27. Banes, A.J.; Tsuzaki, M.; Hu, P.; Brigman, B.; Brown, T.; Almekinders, L.; Lawrence, W.T.; Fischer, T. PDGF-BB, IGF-I and mechanical load stimulate DNA synthesis in avian tendon fibroblasts in vitro. *J. Biomech.* **1995**, *28* (12), 1505–1513.
28. Ripamonti, U.; Van Den Heever, B.; Sampath, T.K.; Tucker, M.M.; Rueger, D.C.; Reddi, A.H. Complete regeneration of bone in the baboon by recombinant human osteogenic protein-1 (hOP-1, bone morphogenetic protein-7). *Growth Factors* **1996**, *13* (3–4), 273–289.
29. Bonadio, J. Tissue engineering via local gene delivery: Update and future prospects for enhancing the technology. *Adv. Drug Deliv. Rev.* **2000**, *44* (2–3), 185–194.

Titanium and Its Alloys

Mitsuo Niinomi

Department of Production Systems Engineering, Toyohashi University of Technology, Tempaku-cho, Toyohashi, Japan

Tomokazu Hattori

Department of Materials Science and Engineering, Faculty of Science and Technology, Meijo University, Tempaku-ku, Nagoya, Japan

Toshihiro Kasuga

Department of Materials Science and Engineering, Nagoya Institute of Technology, Gokiso-cho Showa-ku, Nagoya, Japan

Hisao Fukui

Department of Dental Materials Science, School of Dentistry, Aichi-Gakuin University, Chikusa-ku, Nagoya, Japan

INTRODUCTION

Biomedical and dental applications of titanium and its alloys are constantly increasing because of their excellent biocompatibility, corrosion resistance, and balance of mechanical properties, being lightweight compared with the other metallic materials. Nowadays, functionalities like low rigidity, shape memory effect, super elasticity, bioactive modifications, etc. are getting much more attentions. Recent topics on representative titanium and its alloys for biomedical applications, low-rigidity biomedical titanium alloys, shape memory or superelastic biomedical titanium alloys, mechanical properties of biomedical titanium alloys, bioactive surface modification of biomedical titanium alloys, titanium alloys for dental applications, etc. are discussed in this entry.

BACKGROUND

Since the population percentage of aged people is rapidly growing, the number of aged persons demanding replacement of failed tissue with artificial instruments made of biomaterials is increasing. In particular, the use of instruments for replacing failed hard tissues such as artificial hip joints, dental implants, etc. is increasing among the aged people. Till recently, metallic biomaterials have been the most suitable for replacing failed hard tissue. The chief metallic biomaterials are stainless steels, cobalt (Co) based alloys, titanium (Ti) and its alloys. Recently, titanium alloys are getting much attention as biomaterials because they have excellent specific strength and corrosion resistance, no allergic problems, and the best biocompatibility among metallic biomaterials. Pure titanium and titanium–6 mass% aluminum (Al)–4 mass% vanadium

(V) (Ti–6Al–4V) (the numbers in the alloy name indicate mass% of the element) are still the most widely used for biomedical applications among the titanium alloys. They occupy almost all of the market for titanium biomaterials. However, these are basically developed as structural materials mainly for aerospace structures. Therefore, the development of titanium alloys targeted for biomedical applications is required. Therefore, research and development of titanium alloys composed of nontoxic elements has started, and is continuing to increase.^[1–3] The titanium (Ta) alloys composed of nontoxic elements that were developed in the early stage are mainly $\alpha + \beta$ types, whose constituents are a α phase with hexagonal close-packed (hcp) structure and β phase with body-centered cubic (bcc) structure. Nowadays, the mechanical biocompatibility of biomaterials is regarded as an important factor, and therefore the research and development of β -type titanium alloys whose constituent is metastable β phase with bcc structure at room temperature is increasing because they are advantageous in that respect.^[1,4,5]

Very recently, the possibility of the development of new titanium alloys for biomedical applications, which show superelasticity, is growing, and the research and development of shape memory titanium alloys composed of nontoxic elements for biomedical applications is also attracting attention.^[6–11] The superelastic or shape memory characteristics are expected to facilitate new applications of metallic materials not only in the medical field but also for general purposes. The superelastic characteristic of some titanium alloys appears to be difficult to understand through conventional mechanisms, and therefore they are also very interesting from the point of view of material science.^[12,13]

The research and development of bioactive surface modifications for improving the biocompatibility of titanium alloys is also increasingly being done because

Table 1 Titanium and its selected alloys for biomedical applications

Titanium and its alloys	ASTM standard	ISO standard	JIS standard	Type of alloy
Ti CP-1	ASTM F 67	ISO 5832-2	JIS T 7401-1	α
Ti CP-2	ASTM F 67	ISO 5832-2	JIS T 7401-1	α
Ti CP-3	ASTM F 67	ISO 5832-2	JIS T 7401-1	α
Ti CP-4	ASTM F 67	ISO 5832-2	JIS T 7401-1	α
Ti-6Al-4V ELI	ASTM F 136	ISO 5832-3	—	$\alpha + \beta$
Ti-6Al-4V	ASTM F 1472	ISO 5832-3	JIS T 7401-2	$\alpha + \beta$
Ti-3Al-2.5V	ASTM F-2146	—	—	$\alpha + \beta$
Ti-5Al-2.5Fe	—	ISO 5832-10	—	$\alpha + \beta$
Ti-6Al-7Nb	ASTM F 1295	ISO 5832-11	JIS T 7401-4	$\alpha + \beta$
Ti-5Al-3Mo-4Zr	—	—	—	$\alpha + \beta$
Ti-6Al-6Nb-1Ta	—	—	—	$\alpha + \beta$
Ti-6Al-2Nb-1Ta	—	—	JIS T 7401-3	$\alpha + \beta$
Ti-15Sn-4Nb-2Ta-0.2Pd	—	—	—	$\alpha + \beta$
Ti-15Zr-4Nb-2Ta-0.2Pd	—	—	JIS T 7401-5	β
Ti-12Mo-6Zr-2Fe	ASTM F 1813	—	—	β
Ti-13Nb-13Zr	ASTM F 1713	—	—	β
Ti-12Mo-5Zr-5Sn	—	—	—	β
Ti-15Mo	ASTM F 2066	—	—	β
Ti-16Nb-10Hf	—	—	—	β
Ti-15Mo-2.8Nb-0.2Si	—	—	JIS T 7401-6	β
Ti-15Mo-5Zr-3Al	—	—	—	β
Ti-30Ta	—	—	—	β
Ti-45Nb	AMS 4982	—	—	β
Ti-35Zr-10Nb	—	—	—	β
Ti-35Nb-7Zr-5Ta	Task Force F-04.12.23	—	—	β
Ti-29Nb-13Ta-4.6Zr (TNTZ)	—	—	—	β
Ti-55-8Ni	ASTM F 2063	—	—	Intermetallic

the titanium alloys are grouped into bioinert materials from the point of view of patterns of osteogenesis.^[14-16]

Furthermore, the direct or indirect evaluation of biocompatibility using animals or cells, and evaluations of mechanical performance such as fatigue, fretting fatigue, fracture toughness, etc. are also frequently being done.^[17-22]

Both newly developed alloys and mold materials are used in the dental applications of titanium alloys for making dental prostheses.

REPRESENTATIVE TITANIUM AND ITS ALLOYS FOR BIOMEDICAL APPLICATIONS

Biocompatible titanium alloys that have so far been introduced are listed in Table 1.^[23,24] This table includes both those alloys that are already licensed and implemented for use in living bodies and those at the

research and development stage, which are neither licensed nor implemented.

Pure titanium and $\alpha + \beta$ -type titanium alloys, including the Ti-6Al-4V extra-low interstitials (ELI) and

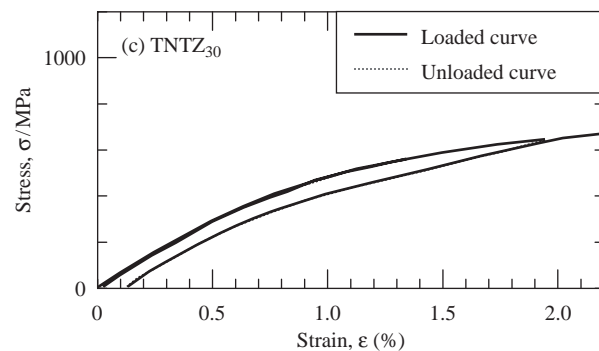


Fig. 1 Tensile loaded-unloaded stress-strain curve of Ti-30Nb-10Ta-5Zr.

Ti-6Al-4V, were originally designed for use as general structural materials, especially for aerospace structures, and were only later adopted for biomedical applications. Then, the toxicity of the β -stabilizing element of V was reported.^[25,26] Therefore, V in Ti-6Al-4V was replaced by another β -stabilizing element, iron (Fe) or niobium (Nb), which are considered to be safer for the living body as compared to V. Ti-5Al-2.5Fe and Ti-6Al-7Nb, which are also $\alpha + \beta$ -type titanium alloys, were introduced.^[27,28] Ti-5Al-2.5Fe has been registered only in ISO standardization, while Ti-6Al-7Nb has been registered in ASTM, ISO, and JIS standardizations.^[29-32] Therefore, it can be said that the development of titanium alloys especially intended for use in living bodies began with the introduction of these alloys. Based on the same concept, other $\alpha + \beta$ -type biomedical titanium alloys like Ti-6Al-6Nb-1Ta and Ti-6Al-2Nb-1Ta have been developed.^[33] The latter is supposed to have been developed mainly as a structural titanium alloy for military use. Subsequently, $\alpha + \beta$ -type titanium alloys that do not contain V or Al, such as the Ti-15zirconium (Zr) based alloys, were introduced.^[34]

The elastic moduli of biomaterials are required to be equal to that of cortical bone because if the elastic moduli of biomaterials are much greater than that of cortical bone, bone resorption occurs. The elastic modulus of the $\alpha + \beta$ -type titanium alloy Ti-6Al-4V, which is around 110 Gpa, is much lower than those of stainless steel and Co based alloys, which are around 1800 and 2100 GPa, respectively.^[35] The elastic modulus of Ti-6Al-4V is still much greater than that of cortical bone, which is around 20 GPa. The elastic

moduli of β -type titanium alloys are known to be smaller than those of α - or $\alpha + \beta$ -type titanium alloys. Therefore, thereafter, mainly β -type titanium alloys were introduced, and most of them aimed at a low modulus of elasticity.^[1,31-40] Of the β -type titanium alloys for biomedical application, Ti-13Nb-13Zr, Ti-12molibdenum (Mo)-6Zr-2Fe, and Ti-15Mo have been registered, and Ti-35Nb-7Zr-5Ta is expected to be registered in the ASTM standardization.^[36-38,40] Previously, β -type titanium alloys with low modulus were mainly developed in the United States, but in recent years they have also been developed in Japan.^[41] The Ti-Nb-Ta-Zr system alloys can be judged as the safest for biomedical applications because Nb, Ta, and Zr are nontoxic and nonallergic alloying elements. Ti-29Nb-13Ta-4.6Zr (TNTZ) has been designed using the d-electron design method based on the DV- X_α cluster theory and was developed at almost the same time when Ti-35Nb-7Zr-5Ta was developed.^[1]

Recently introduced biocompatible titanium alloys contain a large amount of Ta and Nb. These have a melting point and a specific gravity much higher than that of titanium, and have substantially different specific gravities, so the development of melting and casting process technology to obtain high-integrity ingots involved no casting segregation.^[42,43] Levitation melting or cold crucible levitation melting is advantageous to fabricate homogenous ingots that contain a large amount of Nb and Ta.

As low-modulus β -type titanium alloys for biomedical applications, the Ti-Fe-Ta, Ti-Zr-Nb-Ta, Ti-Ta, and Ti-hafnium (Hf) system alloys are also being developed in Japan.^[44-48]

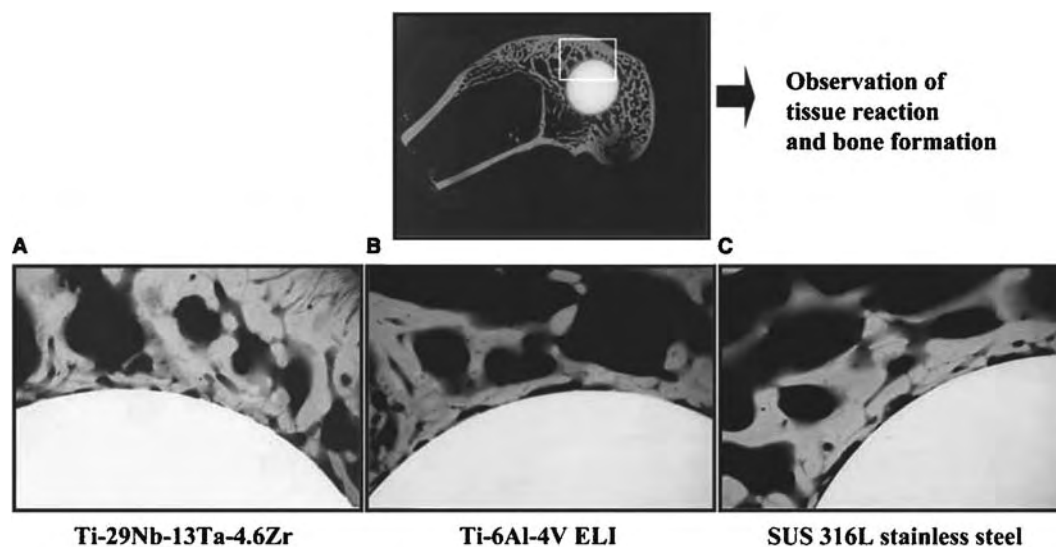


Fig. 2 Contact microradiographs of boundaries of (A) TNTZ, (B) Ti-6Al-4V ELI, and (C) SUS 316 L stainless steel and bones at 8 weeks after implantation.

NICKEL-FREE FUNCTIONAL TITANIUM ALLOYS

Nickel (Ni)-free β -type titanium alloys having not only low modulus but also functions like shape memory effect and/or superelasticity are also being developed or investigated because Ni has been reported to be a strong allergic element, although Ti-Ni shape memory alloys for biomedical and dental applications are still being actively investigated, in particular for stent, catheter, or orthodontic wire applications, and Ti-Ni is the only shape memory alloy that has practical application.^[49-52] The Ti-Nb-tin (Sn), Ti-Mo-gallium (Ga), Ti-Nb-Al, Ti-Mo-Al, Ti-Ta, Ti-Nb, Ti-scandium (Sc)-Mo, Ti-Mo-silver (Ag), Ti-Mo-Sn, and Ti-Nb-Ta-Zr system alloys are being developed as Ni-free shape memory alloys.^[53-61] Among these alloys, around 6% elastic strain has been reported to be obtained in one of the Ti-Nb-Al system alloys by controlling the texture.^[62]

Ti-29Nb-13Ta-4.6Zr also shows superelastic behavior, although shape memory effect is not observed in this alloy.^[63] Fig. 1 shows the tensile loading-unloading curve of Ti-30Nb-10Ta-5Zr powdered metal alloy, which has the simplified chemical composition of TNTZ.^[61] In the elastic region, the tensile loading-unloading stress-strain curve does not follow Hook's

law, showing nonlinear behavior. The features of stress or strain induced martensite transformation are not reflected in the stress-strain curve as shown in Fig. 1, and also in the deformed microstructure. The behavior of this alloy cannot be explained by the increase in the deformation induced martensite transformation and its reversion. In this case, maximum elastic strain is around 3%.

BIOCOMPATIBILITY OF LOW-RIGIDITY TITANIUM ALLOYS COMPOSED OF NONTOXIC AND NONALLERGIC ELEMENTS

Cytotoxicity

The cytotoxicity of low-rigidity Ti-29Nb-12Ta-4.6Zr, pure Ti, and Ti-6Al-4V has been evaluated by neutral red (NR) and 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide (MTT) methods.^[64] Regardless of the evaluation method used, the cell viability of TNTZ is nearly the same as that of pure Ti and greater than that of conventional biomedical Ti-6Al-4V in both as-extracted and filtrated extracted solutions. Therefore, the biocompatibility of TNTZ is judged to be excellent.

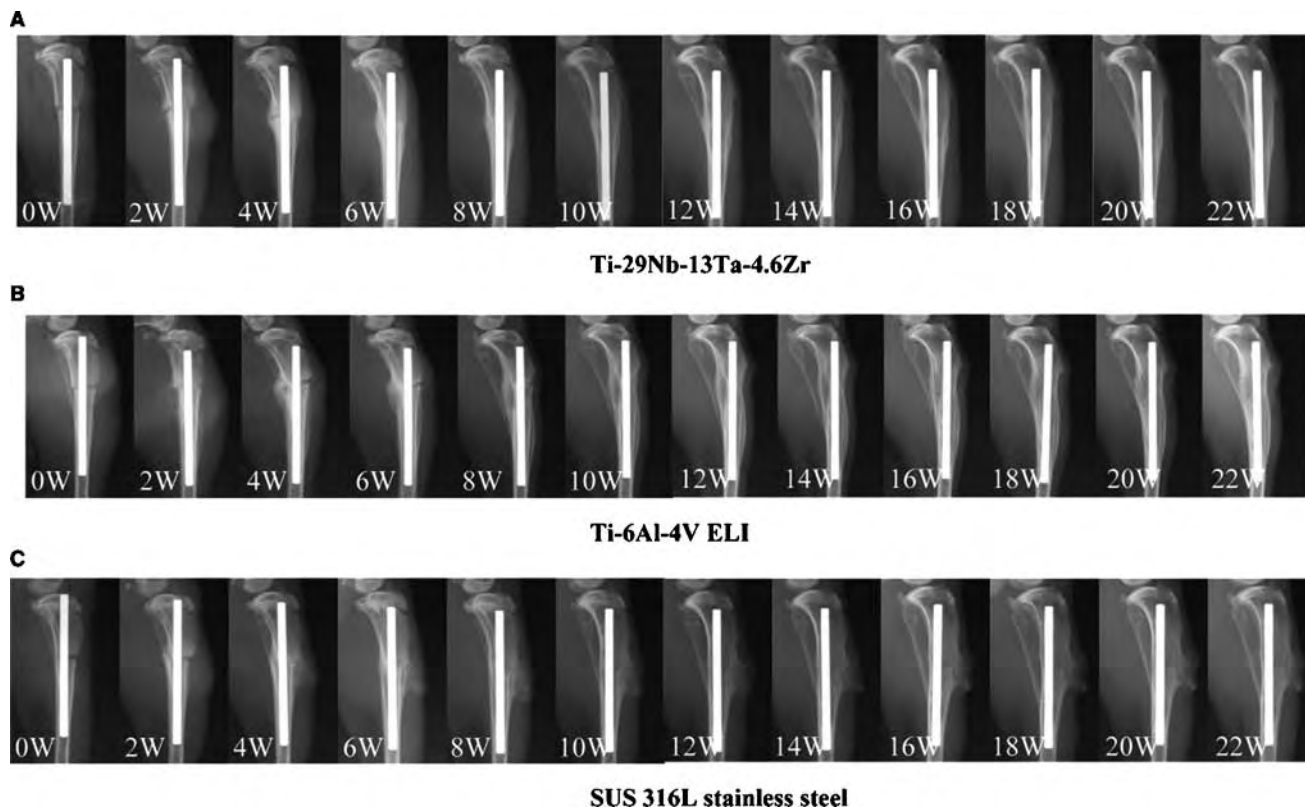


Fig. 3 Healing process of bone fracture from 0 to 22 weeks after surgery using (A) TNTZ, (B) Ti-6Al-4V ELI, and (C) SUS 316 L stainless steel.

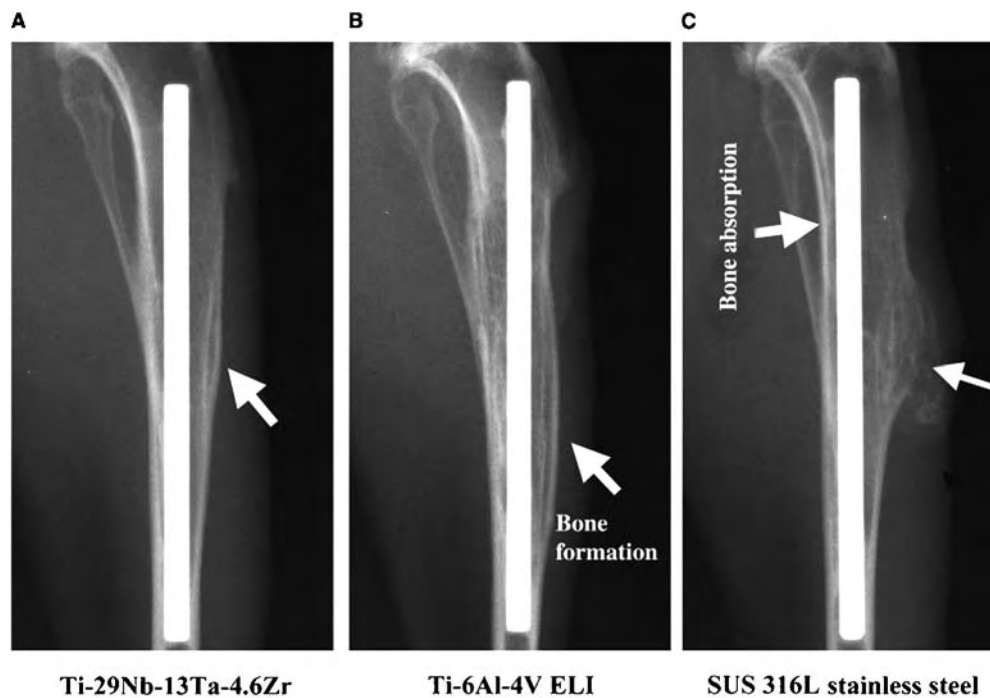


Fig. 4 Remodeling at 24 weeks after implantation for (A) TNTZ, (B) Ti-6Al-4V ELI, and (C) SUS 316 L stainless steel.

The trend of cell viability among TNTZ, pure Ti, and Ti-6Al-4V is nearly the same regardless of extracted period, although the cell viability of both as-extracted and filtrated extracted solutions decreases with increase in the extracted period. However, since the cell viability of filtrated extracted solution is greater than that of as-extracted solution in both NR and MTT methods, not only dissolved metallic ion but also fine wear debris affects the cytotoxicity. This suggests that it is necessary to suppress the process of wear as much as possible when metallic materials are used as biomaterials.

Compatibility with Bone and Remodeling

The contact microradiographs (CMR) of the boundaries of bone and low-rigidity TNTZ (in this case, as-solutionized conditions), Ti-6Al-4V ELI, or SUS (steel use stainless steel) 316 L stainless steel (Fe-18Cr-12Ni-2.5Mo-low C) implanted into the lateral femoral condyles of a rabbit is shown in Fig. 2.^[41,65] Each specimen is surrounded by newly formed bone, and the bone tissue shows direct contact partially with the specimen. However, the extent of the direct contact is greater in TNTZ as compared with Ti-6Al-4V and SUS 316 L stainless steel. Therefore, the biocompatibility of TNTZ with bone is excellent.

To confirm the advantage of low rigidity for bone healing and remodeling, using rabbits, tibial fracture

was induced experimentally in tibia by an oscillating saw just below the tibial tuberosity. An intramedullary rod made of low-rigidity TNTZ, Ti-6Al-4V ELI, or SUS 316 L stainless steel was inserted into the intramedullary canal to fix the fracture. Bone healing, remodeling, and atrophy were observed by x-ray transmission image every 2 weeks for up to 22 weeks. The results are shown in Fig. 3.^[41,65]

The outline of fracture callus was very smooth with bone remodeling in TNTZ. A similar phenomenon was observed at 8 weeks in Ti-6Al-4V ELI and SUS 316 L. In TNTZ, the fracture callus was relatively small, and gradually decreased in size from 6 weeks, and there were no traces of fracture at 10 weeks after the fixation. After 10 weeks, no changes could be observed up to 18 weeks. However, a little atrophic change was observed at the posterior tibial bone after 20 weeks. In Ti-6Al-4V ELI, the callus formation and the bone remodeling were almost similar to that in TNTZ, but slower. A slight atrophic change was observed at 18 weeks. In SUS 316 L stainless steel, a large amount of the fracture callus was observed, and remained up to the end of the succeeding period. Bone atrophy seemed to be occurring at the posterior proximal tibial bone at 10 weeks, and became apparent every 2 weeks. The posterior tibial bone became very thin at 24 weeks, as shown in Fig. 4.^[41,66] Therefore, the low-rigidity titanium alloy, TNTZ, was found to resolve the load transmission problem of the current metal implants with the high rigidity.

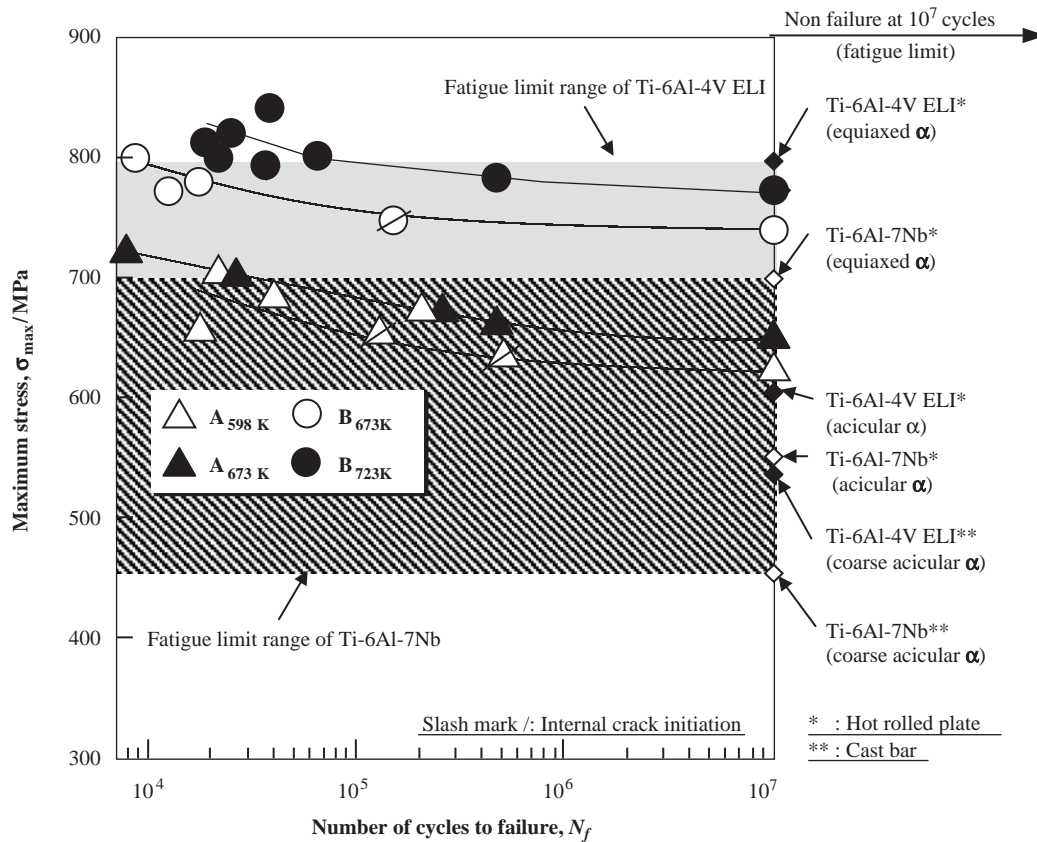


Fig. 5 *S-N* curves of TNTZ conducted with aging at 598 K for 259.2 ksec after ST (A_{598K}), aging at 673 K for 259.2 ksec after ST (A_{673K}), aging at 673 K for 259.2 ksec after CR (B_{673K}), and aging at 723 K for 259.2 ksec after CR (B_{723K}), and fatigue limit ranges of hot rolled and cast Ti-6Al-4V ELI and Ti-6Al-7Nb. CR indicates cold rolling.

MECHANICAL PROPERTIES: FATIGUE, FRETTING FATIGUE, AND WEAR CHARACTERISTICS

Fatigue and Fretting Fatigue Characteristics

The evaluation and improvement of mechanical performance, in particular plain and fretting fatigue, and wear characteristics, including surface treatment for increasing wear resistance of titanium alloys, are done in air and simulated body liquid.^[67-73] Heat treatment and thermomechanical treatment are applied to improve the mechanical properties and fatigue strength, which is important for the safe use of the implants for a long period.^[74]

In the case of β -type titanium alloys for biomedical applications, since the fatigue strength in as-solutionized conditions is low, heat treatment or thermomechanical treatment is effective to improve the fatigue strength drastically. Fig. 5 shows *S-N* (stress-number of cycles to failure) curves of TNTZ used with various thermomechanical treatments with fatigue limit ranges

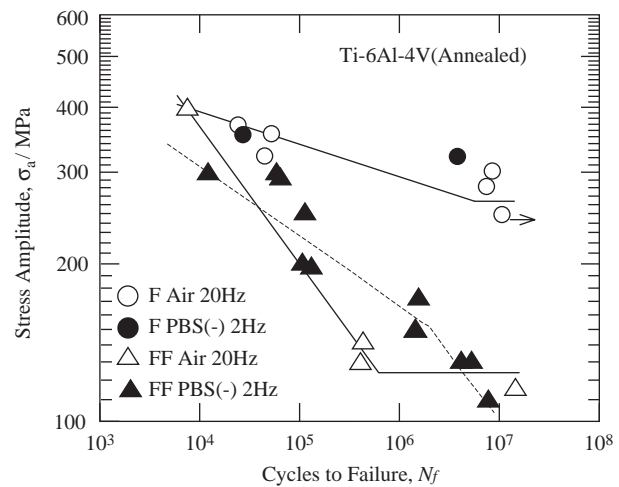


Fig. 6 *S-N* curves of Ti-6Al-4V alloy in air and PBS (-) (F, fatigue; FF, fretting fatigue).



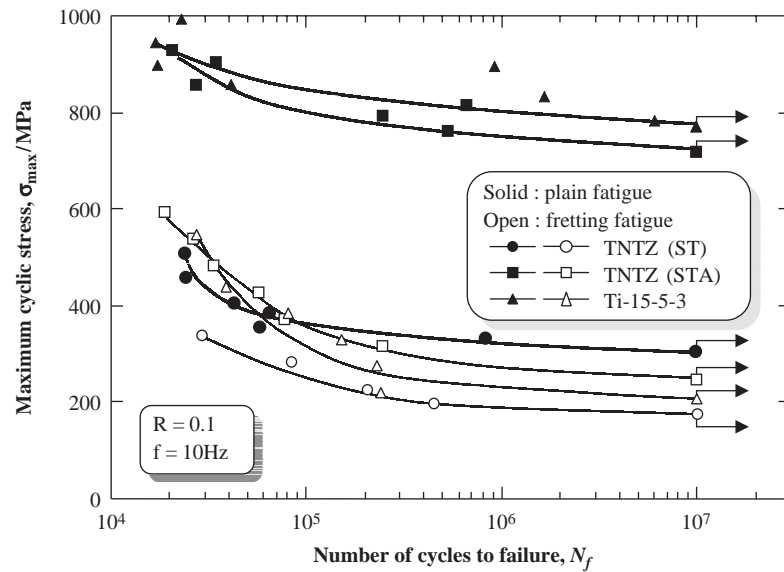


Fig. 7 S - N curves of TNTZ and Ti-15Mo-5Zr-3Al (Ti-15-3-3) conducted with each heat treatment obtained from plain fatigue and fretting fatigue tests in air. ST and STA indicate solution treatment and solution treatment followed by aging, respectively.

of hot rolled and cast Ti-6Al-4V ELI and Ti-6Al-7Nb.^[74] For example, the fatigue strength of TNTZ conducted with aging at 723 K for 259.2 ksec after cold rolling (CR), indicated by B_{723K} , is very high in the highest range of the fatigue limit of Ti-6Al-4V ELI, although the elastic modulus is a little higher, around 80 GPa, than that of as-solutionized TNTZ, which is around 60 GPa.^[74]

The fatigue strength of titanium alloys for biomedical applications is, in general, not degraded in both simulated body environment and living body environment, as in rabbit.^[75] However, the fatigue strength of Ti-6Al-4V ELI under the rotating bending condition was reported to be degraded in Ringer's solution, which is degassed by nitrogen (N_2) bubbling

to reduce the oxygen content because the oxygen content in living body liquid is very low.^[73]

When fatigue stress and fretting occur simultaneously, fatigue strength is greatly reduced, by approximately one-third to one-half. This is called fretting fatigue, which may occur, for example, between bone plate and screw. The fatigue lives of pure titanium grades 2 (N (max): 0.03, carbon (C) (max): 0.08, H (max): 0.015, Fe (max): 0.25, Ti: balance, mass%) and 3 (N (max): 0.05, C (max): 0.08, H (max): 0.015, Fe (max): 0.30, O (max): 0.35, Ti: balance, mass%) have been reported to be smaller in simulated body fluid [SBF (-)] than in air, while the fretting fatigue lives have been reported to be smaller than the fatigue lives by 20–30 %, and smaller in Ca^{2+} - and

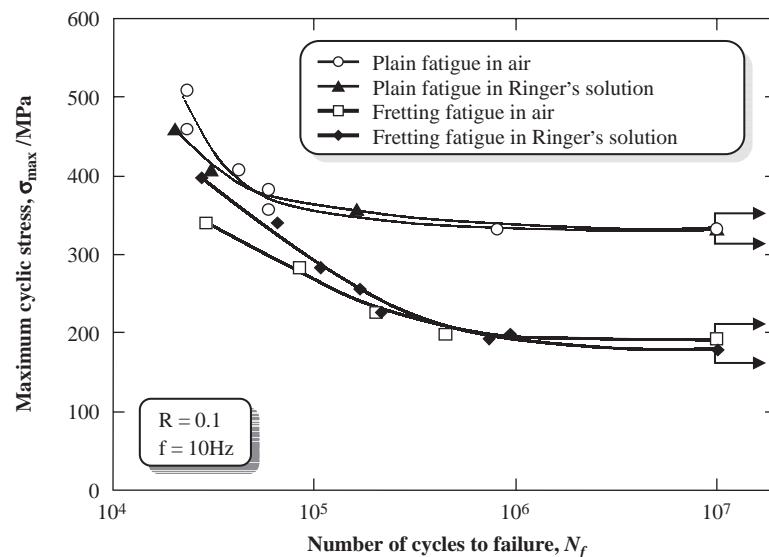


Fig. 8 S - N curves of TNTZ conducted with solution treatment obtained from plain fatigue and fretting fatigue tests in air and Ringer's solution.

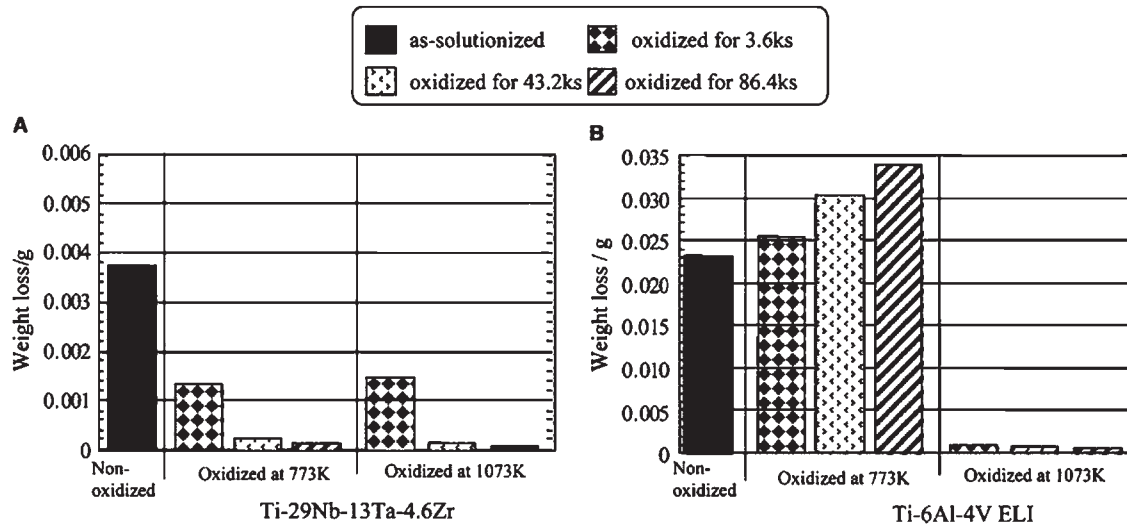


Fig. 9 Weight loss of (A) TNTZ and (B) Ti-6Al-4V ELI oxidized at 773 and 1073 K for various times after solutionizing in friction wear tests in Ringer's solution.

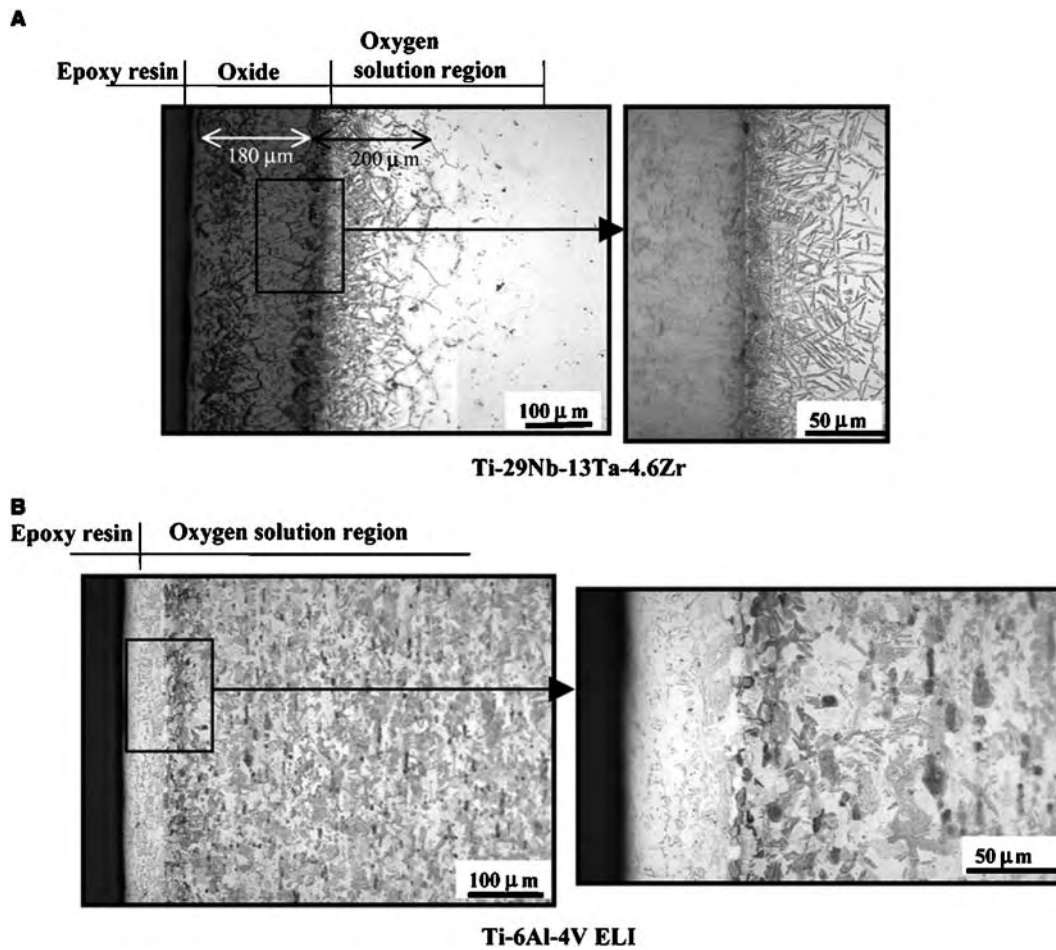


Fig. 10 Optical micrographs showing cross sections of as-solutionized (A) TNTZ and (B) Ti-6Al-4V ELI oxidized at 1073 K for 86.4 ksec.

T

Mg^{2+} -free phosphate buffered solution [PBS(-)] than in air. The fretting fatigue strength of $\alpha + \beta$ -type Ti-6Al-4V in SBF(-) is smaller in both lower- and higher-cycle fretting fatigue life regions, but greater in the intermediate-cycle fatigue life region, while fatigue strength of Ti-6Al-4V is not degraded in SBF(-), as shown in Fig. 6.^[77] $S-N$ curves obtained from fatigue and fretting fatigue tests on β -type TNTZ in as-solutionized and aged conditions, and on $\alpha + \beta$ -type Ti-15Mo-5Zr-3Al in annealed conditions in air are shown in Fig. 7.^[78] Their fretting fatigue limits are around one-third of the fatigue limits. The fatigue limit of Ti-15Mo-5Zr-3Al (TI-15-5-3) is 780 MPa and is the greatest. The second greatest fatigue limit is obtained for aged TNTZ. The as-solutionized TNTZ shows the lowest fatigue limit. Fretting fatigue limit of each material is much lower than the fatigue limit of each material. The fretting fatigue limit of as-solutionized TNTZ is about 190 MPa, which is lower than its fatigue limit by 42%, that is, 140 MPa. The difference in fretting fatigue strength between as-solutionized TNTZ and Ti-15Mo-5Zr-3Al at around 2×10^4 is about 200 MPa, but it is only 20 MPa around the fretting fatigue limit. The fretting fatigue limit of Ti-15Mo-5Zr-3Al is lower than that of aged TNTZ by around 75 MPa. Therefore, the correlation between fatigue strength and fretting fatigue strength is low when comparing fretting fatigue among different kinds

of materials. $S-N$ curves obtained from fretting fatigue tests of as-solutionized TNTZ in air and Ringer's solution are shown in Fig. 8.^[78] The fatigue strength of as-solutionized TNTZ in Ringer's solution is equal to that in air. Therefore, the fatigue strength of as-solutionized TNTZ is not degraded in Ringer's solution. On the other hand, the fretting fatigue strength of as-solutionized TNTZ is slightly greater than that in air in the low-cycle fretting fatigue life region, but a little lower than that in air in the high-cycle fretting fatigue life region. The frictional force in Ringer's solution has been reported to be lower than that in air over the entire maximum cyclic stress and is reported to be about one-third of the frictional force in air at relatively greater maximum cyclic stress.^[78] Therefore, Ringer's solution acts as a lubricant at the fretting area and lowers the frictional force. It is reported that corrosion pits are observed on the whole contact area.^[78] The fretting fatigue crack does not initiate from the pits, but from the edge of the pad. In the high-cycle fretting fatigue life region, pits were not observed on the stick area, but they were observed on the slip area. The fretting fatigue crack initiated from the pit. Therefore, the effect of Ringer's solution as a lubricant is relatively greater in the low-cycle fretting fatigue region, which leads to increase in the fretting fatigue life of as-solutionized TNTZ. However, in the high-cycle fretting fatigue life region, the effect of Ringer's solution as a corrosive agent is relatively greater, which leads to decrease in the fretting fatigue life of as-solutionized TNTZ.

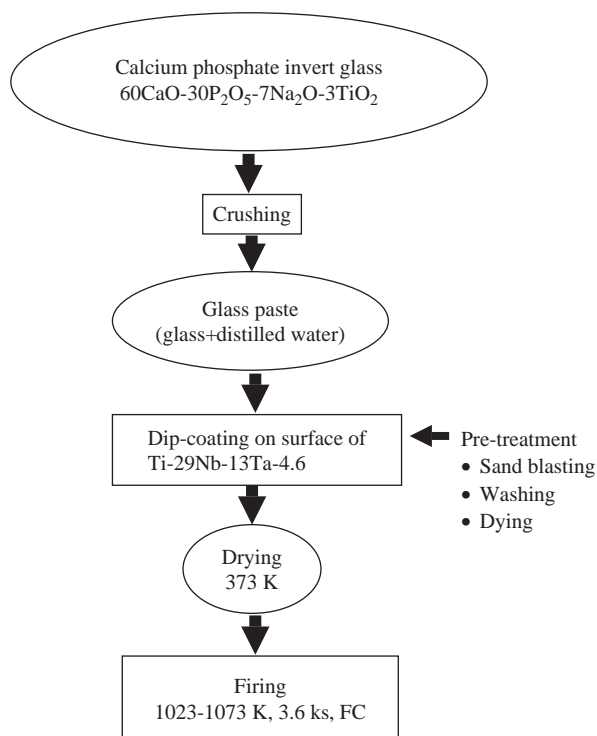


Fig. 11 Schematic drawing of coating process of CaP glass ceramic on surface of TNTZ.

Wear Characteristics

Since the wear resistance of titanium alloys is inferior to that of other metallic biomaterials such as stainless steel or Co-chromium (Cr) alloy, surface modification by oxidation, nitriding, diamond-like carbon coating,

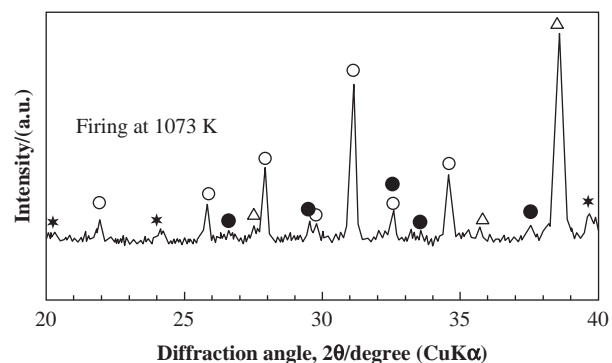


Fig. 12 X-ray diffraction profile of surface of Ti-29Nb-13Ta-4.6Zr coated with calcium phosphate invert glass followed by firing at 1073 K for 1 h; ○: TCP, ●: CPP, ▲: β -Ti; ★: unknown phase.

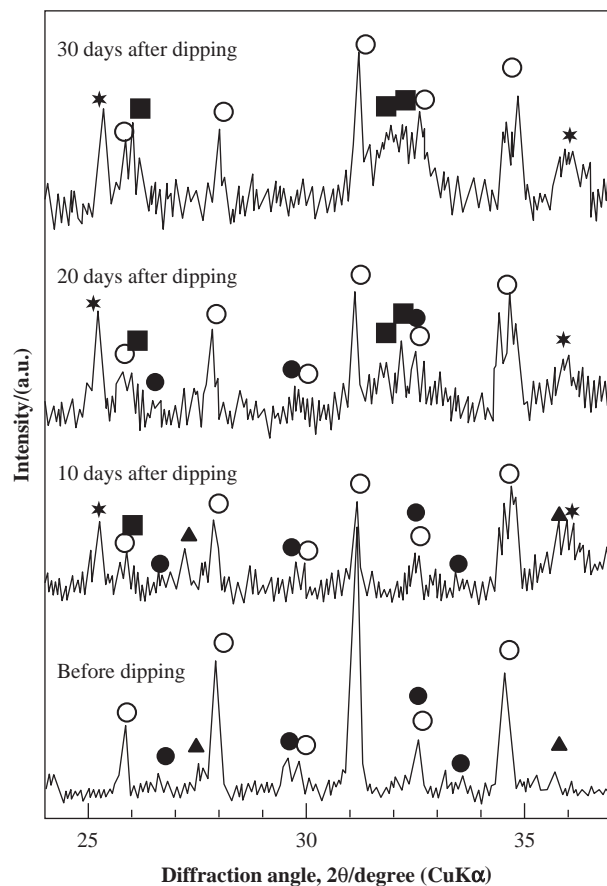


Fig. 13 X-ray diffraction profiles of surfaces of calcium phosphate glass-ceramic-coated Ti-29Nb-13Ta-4.6Zr before and after dipping into simulated body fluid, SBF, for 29Nb-13Ta-4.6Zr for 10, 20, and 30 days; ■:HAP; ○:TCP; ●:CPP; ▲:TiO₂, ★:unknown phase.

and use of a laser beam has been investigated to improve the wear resistance of titanium.^[73,79–81]

The oxidation process in air is the one of the easiest processes for improving the wear resistance of titanium

alloys for biomedical applications. Weight losses of TNTZ and Ti-6Al-4V ELI oxidized at 773 and 1073 K for various times after solutionizing in friction wear tests in Ringer's solution are shown in Fig. 9.^[79] For an oxidation treatment temperature of 773 K, weight loss of TNTZ decreases with increasing treatment time. However, weight loss of Ti-6Al-4V ELI increases in this case. For an oxidation treatment temperature of 1073 K, weight loss decreases in both TNTZ and Ti-6Al-4V ELI with increasing treatment time. It has been reported that in oxidized TNTZ, the phase precipitated below the oxide layer on the specimen surface is the main factor for improving the wear resistance of TNTZ, as shown in Fig. 10.

BIOACTIVE SURFACE MODIFICATION

Titanium alloys show the greatest biocompatibility among metallic materials for biomedical applications. However, they are grouped into bioinert materials as well as ceramics like alumina, zirconia, etc., on the basis of the pattern of osteogenesis, as already stated before, and their biocompatibility is inferior to that of calcium phosphate (CaP) or hydroxyapatite [HAP: Ca(PO₄)₃OH], which is grouped into bioactive materials. Therefore, bioactive surface treatment (bioactive surface modification) is in general applied to titanium alloys for biomedical applications to improve their biocompatibility further. In that case, CaP-type ceramics such as CaP, TCP [β-Ca₃(PO₄)₂], CCP (β-Ca₂P₂O₇), etc. and HAP are mainly coated on the surface of the titanium alloy. In general, the formation of HAP is finally targeted.

The bioactive surface treatment processes are in general divided into dry processes and wet processes. There are various dry and wet processes.^[14]

Dry processes are divided into direct HAP forming methods and indirect HAP forming methods. The

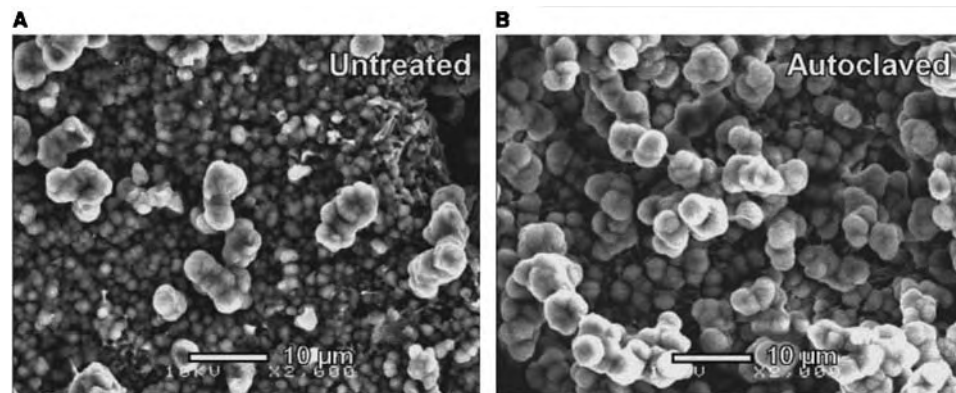


Fig. 14 Scanning electron micrographs of the (A) untreated and (B) autoclaved glass-ceramic-coated TNTZ surfaces after 10 days of soaking in SBF.

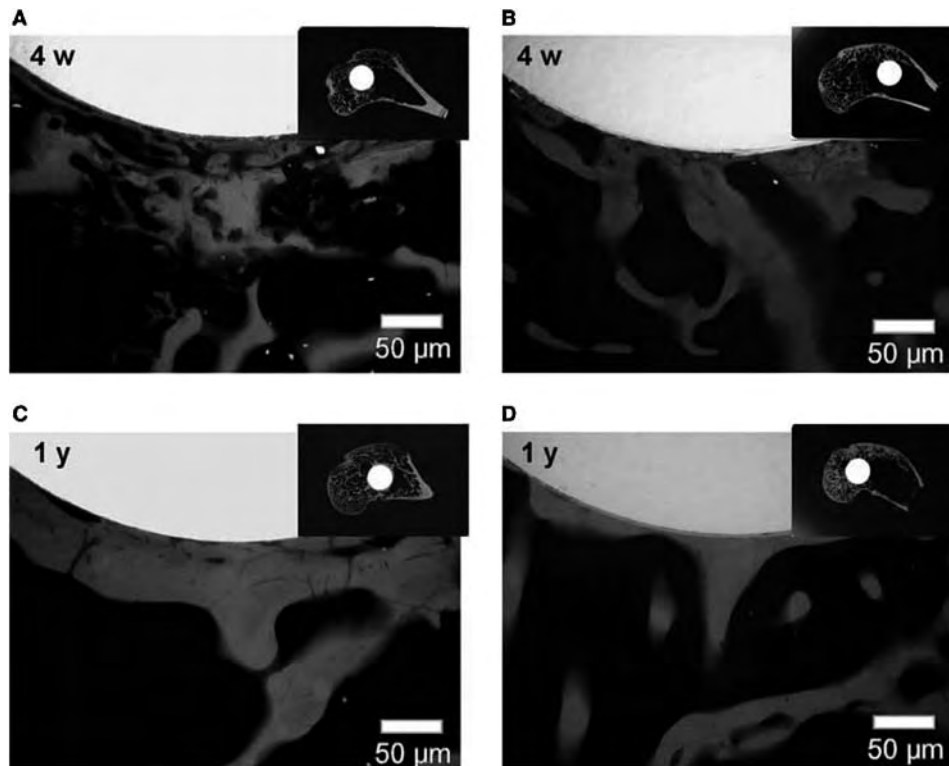


Fig. 15 Contact microradiographs of (A, C) TNTZ and (B, D) glass-ceramic-coated Ti-29Nb-13Ta-4.6Zr after implantation in the femurs of Japanese rabbits: (A, B) 4 weeks of implantation and (C, D) 1 year of implantation. Diameter of the samples is 5 mm.

former include plasma spray method, ion plating, RF magnetron sputtering, pulse laser deposition method, ion beam dynamic mixing method, superplastic joining method, etc., where HAP is formed directly on the titanium alloy surface.^[82-87] The latter include calcium ion implantation, where calcium ions are implanted into biomedical titanium alloys, calcium ion mixing method, where Ca is sputtered on the surface of biomedical titanium alloys followed by Ar ion implantation, etc.^[14,88] Calcium phosphate precipitation is enhanced on the surface of biomedical titanium alloys with these treatments when they are implanted into the living body.

Wet processes are also divided into direct HAP forming methods and indirect HAP forming methods. An example of the former is electrochemical treatment.^[89] An example of the latter is alkali treatment,

where the biomedical titanium alloy is immersed into NaOH solution and heated followed by immersing the alloy into living body liquid.^[90]

There is another interesting method where the powder of CaP invert glass mixed with distilled water is coated on the surface of the titanium alloy followed by heating at around 1073 K, and then CaP-type ceramics are precipitated (Fig. 11).^[15,91] Fig. 12 shows the x ray profiles of the surface of TNTZ obtained by this method. Bioactive β -TCP and β -CCP are formed by heating glass-coated TNTZ at 1073 K for 3.6 ksec. Furthermore, HAP is formed by immersing it in SBF, as shown in Fig. 13.^[91] This process is more advantageous for the oxidation resistant titanium alloys containing a large amount of Nb and Ta, such as TNTZ. The apatite formation by soaking in SBF is

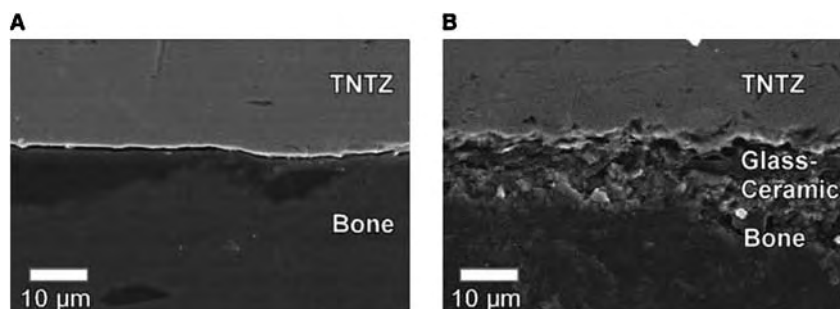


Fig. 16 Cross-sectional scanning electron micrographs of (A) TNTZ and (B) the glass-ceramic-coated TNTZ after 1 year of implantation.

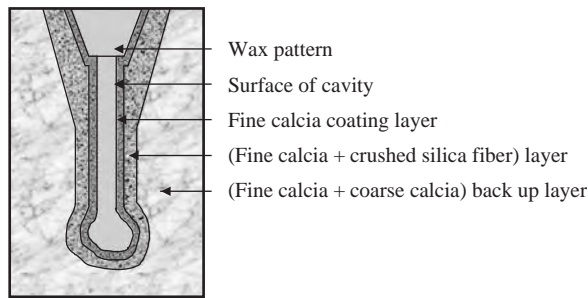


Fig. 17 Schematic drawing of cross section of duplex-coated calcia mold.

enhanced by autoclaving at between 393 and 413 K for 1 hr after heating. Fig. 14 shows scanning electron micrographs of the autoclaved and untreated glass-ceramic-coated TNTZ after 10 days of soaking in SBF.^[92] The micrographs show that apatite formation in SBF occurs on the coating surfaces. The autoclaved glass-ceramic-coated TNTZ is completely covered with apatite, while the untreated one is not completely covered and the original glass-ceramic surface is also seen in some areas.

Fig. 15 shows CMR of TNTZ and the autoclaved glass-ceramic-coated TNTZ after implantation into the femurs of Japanese rabbits. The photographs show new bone formation around TNTZ and the autoclaved glass-ceramic-coated TNTZ after 4 weeks of implantation; the bone tissue shows direct contact partially with implants. After 1 year of implantation, the direct contact of bone tissue with TNTZ and the autoclaved glass-ceramic-coated TNTZ can be seen in many areas.

Fig. 16 shows cross-sectional scanning electron micrographs after 1 year of implantation.^[92] Crack propagation occurs between TNTZ and bone, while no crack occurs between the glass-ceramic coating and bone. The glass-ceramic coating bonds directly with bone.



Fig. 18 Photograph of specimen of TNTZ cast using wax pattern formed with duplex coating of fine pure calcia and crushed silica fiber reinforced fine calcia.

DENTAL APPLICATIONS

$\alpha + \beta$ -type titanium alloys like Ti-15Zr-4Nb-4Ta are also still being actively investigated.^[93] Dental casting of Ti-15Zr-4Nb-4Ta is also being investigated, as well as the β -type titanium alloys as stated above.^[94] Recently, the use of Ti-6Al-7Nb instead of Ti-6Al-4V has been increasing in dental applications.^[95]

The precision casting process is predominant for making dental prostheses like crowns, dentures, inlays, etc. In cast titanium alloys, especially the elongation is very low although the strength is kept high. Therefore, methods for increasing elongation without reducing strength are being investigated. Hydrogenation and dehydrogenation processing, that is, thermochemical processing followed by post-heat treatment treatments like α - β solution treatment, β solution treatment, broken-up structure, etc. are effective for improving elongation without reducing strength in cast titanium alloys.^[96,97]

Titanium alloys are very reactive and have a relatively higher melting point compared with other dental alloys like gold (Au) based alloys and silver (Ag) based alloys. Especially, the recently proposed β -type titanium alloys have a very high melting point because they contain high melting point elements like Ta and Nb. Therefore, low melting point titanium alloys and low reactive mold materials are desired for dental precision castings.

In dental precision casting, alumina based and magnesia based mold materials are mainly used.^[98] Magnesia based mold material is more suitable for dental precision casting of titanium alloys than alumina based mold materials.^[99] Calcia based mold material is also reported to be suitable for dental precision casting.^[100,101] However, calcia based mold material is said to be difficult to treat. To solve this problem, the duplex coating method has been developed. In this method, the first coating of wax pattern is done using pure fine calcia slurry, and the second coating of wax pattern is done using crushed silica fiber reinforced pure fine

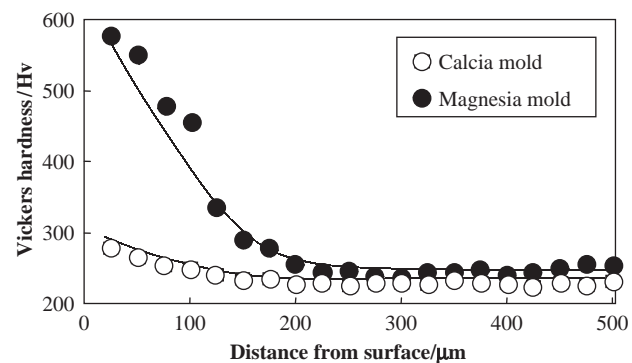


Fig. 19 Vickers hardness of specimen of TNTZ cast using calcia mold (duplex calcia coating method) and general magnesia based mold as a function of distance from surface of specimen.

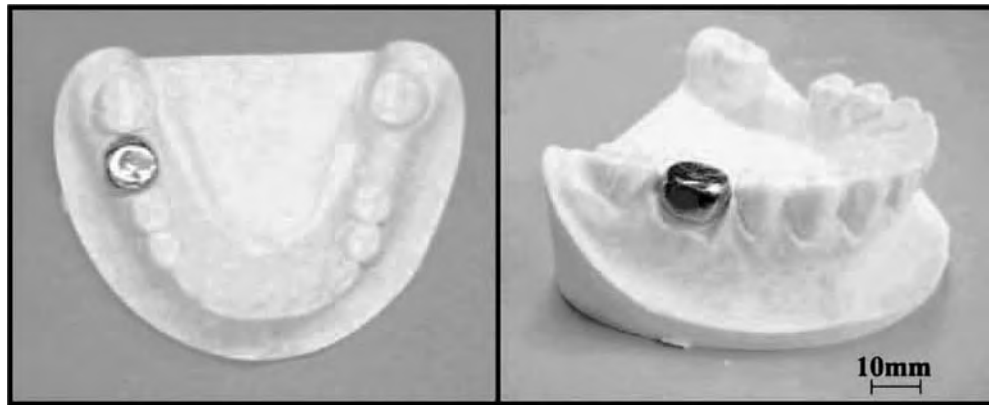


Fig. 20 Photograph of dental product, crown, made by duplex calcia coating method.

calcia slurry. The second coating is reinforced by crushed silica fiber to avoid penetration of the molten metal into the backup mold. The duplex coating mold is schematically explained in Fig. 17.^[102] The photograph of the specimen of TNTZ cast using wax pattern produced with a duplex coating of fine pure calica and crushed silica fiber reinforced fine calcia slurry is shown in Fig. 18.^[103] In this case, sandblasting treatment is not carried out. Even without sandblasting treatment, the surface condition of the specimen is very smooth without penetration. The hardness near the surface of TNTZ cast using calcia based mold is very low compared with that of TNTZ cast using magnesia based mold, as shown in Fig. 19.^[103] The photograph of the crown made of TNTZ cast using the calcia duplex coating method is shown in Fig. 20.^[103]

On the other hand, low melting point β -type titanium alloys for dental applications like Ti-29Nb-13Zr-2Cr, Ti-29Nb-15Zr-1.5Fe, Ti-29Nb-10Zr-0.5 silicon (Si), Ti-29Nb-10Zr-0.5Cr-0.5Fe, Ti-29Nb-18Zr-2Cr-0.5Si have been developed to lower the melting point of TNTZ using the DV- X_2 cluster method.^[1,104] Melting points of designed alloys fall by about 50–370 K as compared with that of TNTZ as shown in Fig. 21.^[104] Releasing of the specimen of each designed alloy from the mold that is made of commercial magnesia based investment materials is easier compared with the case of TNTZ as represented in Fig. 22, showing almost no adhesion of mold materials on the surface of the as-cast specimen surface of Ti-29Nb-13Zr-2Cr.^[104] The cytotoxicity of these newly developed alloys has been reported to be excellent.

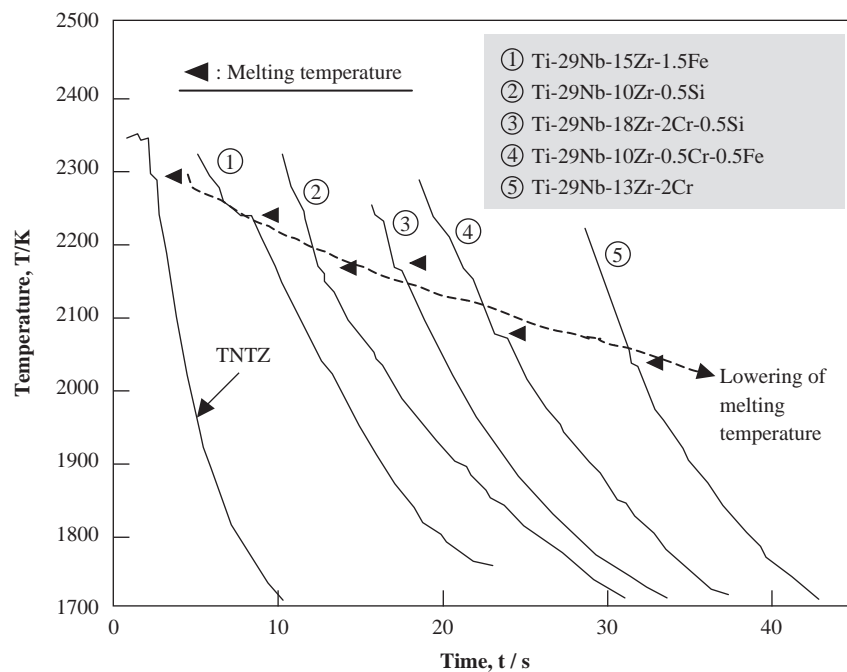


Fig. 21 Cooling curves of various designed Ti alloys and TNTZ.

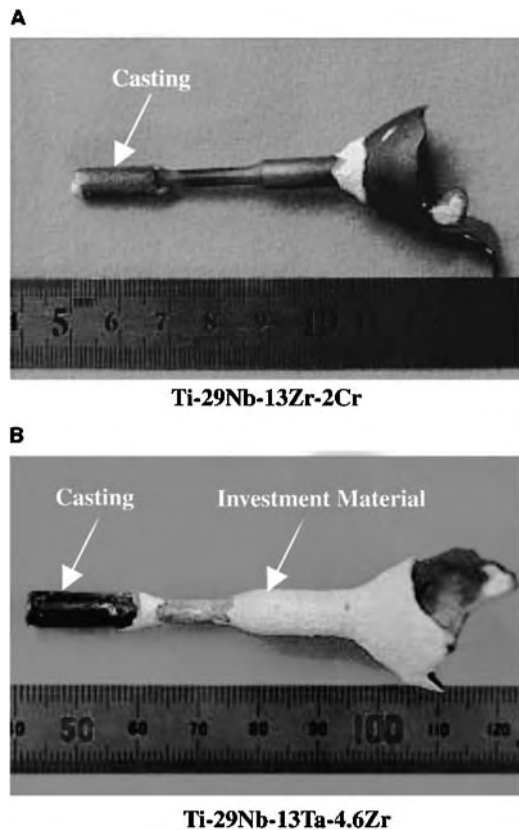


Fig. 22 Reproductions of as-cast tensile specimens of (A) Ti-29Nb-13Zr-2Cr and (B) TNTZ.

CONCLUSIONS

Titanium and titanium alloys are the most attractive metallic biomaterials. However, they are not bioactive. Therefore, bioactive ceramic, mainly HAP, surface modification is needed for achieving bonding with bone. Furthermore, to offer scaffolds for growing artificial organs, biomolecular modification of titanium and titanium alloys will be required. For solving this problem, integration of metallic, ceramic, and polymer materials is needed.

ARTICLES OF FURTHER INTEREST

Nickel and Nickel Alloys; Shape Memory Metals; Spine Biomechanics

REFERENCES

1. Kuroda, D.; Niinomi, M.; Morinaga, M.; Kato, Y.; Yashiro, T. Design and mechanical properties of new β type titanium alloys for implant materials. *Mater. Sci. Eng. A* **1998**, *A243*, 244–249.
2. Niinomi, M. Developments of β type titanium alloys for biomedical applications. *Mater. Jpn.* **1998**, *37*, 843–846.
3. Okazaki, Y.; Rao, S.; Tateishi, T.; Ito, Y. Cyto-compatibility of various metal and development of new titanium alloys for medical implants. *Mater. Sci. Eng. A* **1998**, *A243*, 250–256.
4. Kovacs, P.; Davidson, J.A. The electrochemical behavior of a new titanium alloy. In *Titanium'92*; Froes, F.H., Caplan, I., Eds.; TMS, 1993; 2705–2712.
5. Ahmed, T.; Long, M.; Silvestri, J.; Ruiz, C.; Rack, H.J. A new low modulus, biocompatible titanium alloy. In *Titanium'95*; Blenkinsop, P.A., Evans, W.J., Flower, H.M., Eds.; The Institute of Materials: London, 1996; 1760–1767.
6. Nitta, K.; Watanabe, S.; Masahashi, N.; Hosoda, H.; Hanada, S. Ni-free Ti-Nb-Sn shape memory alloys. In *Structural Biomaterials for the 21st Century*; Niinomi, M., Okabe, T., Taleff, E.M., Lesuer, D.R., Lippard, H.F., Eds.; TMS, 2001; 25–34.
7. Kim, H.Y.; Ohimatsu, Y.; Kim, J.H.; Hosoda, H.; Miyazaki, S. Mechanical properties and shape memory behavior of Ti-Mo-Ga alloys. *Mater. Trans.* **2004**, *45* (4), 1090–1095.
8. Hosoda, H.; Yamamoto, A.; Miyazaki, S. Mechanical properties of Ti-Mo-Ge shape memory alloy for biomedical applications. *Proceedings of the Fall Meeting of the Japan Institute of Metals*, 2001; 401.
9. Hiromoto, H.; Mizuno, F.; Hanawa, T.; Kuroda, C.; Hosoda, H.; Wakashima, K.; Miyazaki, S. Polarization behavior of Ti-Mo-Al shape memory alloy in simulated body liquid. *Proceedings of the Annual Meeting of the Japan Institute of Metals*, 2002; 443.
10. Ikeda, M.; Nakamura, Y.; Takahama, N. Effect of Zr contents on heat treatment behaviors and phase constitution of Ti-50 mass% Ta-Zr alloy. *Proceedings of the Annual Meeting of the Japan Institute of Metals*, 2003; 130.
11. Maeshima, T.; Eto, T.; Uchiyama, H.; Uchiyama, K.; Nishida, M. Development of Ti-Sc-Mo shape memory alloy. *Proceedings of the Annual Meeting of the Japan Institute of Metals*, 2003; 134.
12. Saito, S. *Super Elastic Titanium Alloy*. *Gum Metal*, Report of the 26th Workshop of JSPS 156 Committee, Nov 22, 2001; 7–8.
13. Niinomi, M. Recent topics of titanium alloys for biomedical applications. *Netsushori* **2003**, *42*, 409–415.
14. Hanawa, T. Surface modification of metallic biomaterials. *Mater. Jpn.* **1998**, *37*, 853–855.
15. Kasuga, T.; Nogami, M.; Niinomi, M. Calcium phosphate glass-ceramic coating on a new β -type titanium alloy. In *Structural Biomaterials for the 21st Century*; Niinomi, M., Okabe, T., Taleff, E.M., Lesuer, D.R., Lippard, H.F., Eds.; TMS, 2001; 265–272.
16. Yamamuro, Y. Patterns of osteogenesis in relation to various biomaterials. *J. Jpn. Soc. Biomater.* **1989**, *7*, 19–23.
17. Kasuga, T.; Nogami, M.; Niinomi, M.; Hattori, T.; Hench, L.L. Enhancing effect of autoclaving on bioactivity of β -titanium alloy coated with calcium phosphate glass-ceramic. *Key Eng. Mater.* **2005**, *284–286*, 243–246.

18. Okazaki, Y.; Asao, S.; Rao, S.; Tateishi, T. Effect of concentration of Zr, Sn, Nb, Ta, Pd, Mo, Co, Cr, Si, Ni, Fe on the relative growth ratios of bio-cells. *J. Jpn. Inst. Metals* **1996**, *60*, 902–906.
19. Yamamoto, A.; Kobayashi, T.; Maruyama, N.; Nakazawa, K.; Sumita, M. Fretting fatigue properties of Ti–6Al–4V alloy in pseudo-body fluid and evaluation of biocompatibility by cell culture method. *J. Jpn. Inst. Metals* **1995**, *59*, 463–470.
20. Akahori, T.; Niinomi, M.; Fukunaga, K.; Inagaki, I. Effects of microstructure on the short fatigue crack initiation and propagation characteristics of biomedical α/β titanium alloys. *Metall. Mater. Trans. A* **2000**, *31A*, 1949–1958.
21. Niinomi, M.; Akahori, T.; Ishimizu, K. Fatigue and fretting fatigue of biomaterials, Ti–29Nb–13Ta–45.6Zr, in air and simulated body environment. In *Materials Lifetime Science and Engineering*; Liaw, P.K., Buchana, R.A., Klarstrom, D.L., Wei, R.P., Harlow, D.G., Eds.; TMS, 2003; 223–230.
22. Nakazawa, K.; Sumita, M.; Maruyama, N. Fatigue and fretting fatigue of austenitic stainless steels in pseudo-body fluid. *J. Jpn. Inst. Metals* **1999**, *63*, 1600–1608.
23. Freese, H.L.; Volas, M.G.; Wood, J.R. Metallurgy and technological properties of titanium and titanium alloys. In *Titanium in Medicine*; Brunette, D.M., Tengvall, P., Textor, M., Thomsen, P., Eds.; Springer, 2001; 25–51.
24. Niinomi, M. Recent biocompatible metallic biomaterials. In *Structural Biomaterials for the 21st Century*; Niinomi, M., Okabe, T., Tallef, E.H., Lesure, D.R., Lippard, H.E., Eds.; TMS, 2001; 15–24.
25. Zwicker, R.; Buehler, K.; Beck, H.; Schmid, H.J. Mechanical properties and tissue reaction of a titanium alloy for implant material. In *Titanium'80*; Kmura, H., Izumi, O., Eds.; TMS–AIME: Warrendale, PA, 1980; 2, 505–514.
26. Niinomi, M. Mechanical properties of biomedical titanium alloys. *Mater. Sci. Eng. A* **1998**, *A243*, 231–236.
27. Browy, K.–H.; Kramer, K.–H. On the properties of a new titanium alloy (TiAl5Fe2.5) as implant material. In *Titanium Science and Technology*; Lütjering, G., Zwicker, U., Bunk, W., Eds.; Deutsche Gesellschaft für Metallkunde E. V., 1985; Vol 2, 1381–1386.
28. Semlitsch, M. Classic and new titanium alloys for production of artificial hip joints. *Titan* **1986**, 721–740.
29. *Implants for Surgery—Metallic Materials—Part 10: Wrought Titanium 5–Aluminum 2.5–Iron Alloy*; ISO 5832–10; Genève, Switzerland, 1996.
30. *Standard Specification for Wrought Titanium 6Al–7Nb Alloy for Surgical Implant*; ASTM Designation F1295–92, ASTM: Philadelphia, PA, 1994; 687–689.
31. *Implants for Surgery—Metallic Materials—Part II: Wrought Titanium 6–Aluminum 7–Niobium Alloy*; ISO 5832–11; Genève, Switzerland, 1994.
32. Okazaki, Y. Standardization and activation of industry. *Mater. Jpn.* **2004**, *43* (3), 182–185.
33. Sasaki, Y.; Doi, K.; Matsushita, T. New titanium alloys for artificial hip joints. *Kinzoku* **1996**, *66*, 812–817.
34. Okazaki, Y.; Ito, Y.; Tateishi, T.; Ito, A. Effect of heat treatment on microstructure and mechanical properties of new titanium alloys for surgical implantation. *J. Jpn. Inst. Metals* **1995**, *59*, 108–115.
35. Niinomi, M. Recent metallic materials for biomedical applications. *Metall. Mater. Trans. A* **2002**, *33A*, 477–486.
36. Mishra, A.K.; Davidson, J.A.; Poggie, R.A.; Kovacs, P.; Fitzgerald, T.J. Mechanical and tribological properties and biocompatibility of diffusion hardened Ti–13Nb–13Zr—a new titanium alloy for surgical implants. In *Medical Applications of Titanium and Its Alloy*; Brown, S.A., Lemons, J.E., Eds.; ASTM STP 1272, ASTM, West Conshohocken, PA, 1996; 96–116.
37. Wang, K.K.; Gustavson, L.J.; Dumbleton, J.H. Microstructure and properties of a new beta titanium alloy, Ti–12Mo–6Zr–2Fe, developed for surgical implants. In *Medical Applications of Titanium and Its Alloy*; Brown, S.A., Lemons, J.E., Eds.; ASTM STP 1272, ASTM, West Conshohocken, PA, 1996; 76–87.
38. Zardiackas, L.D.; Mitchell, D.W.; Disegi, J.A. Characterization of Ti–15Mo beta titanium alloy for orthopaedic implant. In *Medical Applications of Titanium and Its Alloy*; Brown, S.A., Lemons, J.E., Eds.; ASTM STP 1272, ASTM, West Conshohocken, PA, 1996; 60–75.
39. Bhamri, S.K.; Shetty, R.H.; Gilberton, L.N. Optimization of properties of Ti–15Mo–2.8Nb–3Al–0.2Si and Ti–15Mo–2.8Nb–0.2Si–0.26O beta titanium alloys for application in prosthetic implants. In *Medical Applications of Titanium and Its Alloy*; Brown, S.A., Lemons, J.E., Eds.; ASTM STP 1272, ASTM, West Conshohocken, PA, 1996; 88–95.
40. Ahmed, T.; Long, M.; Silvestri, J.; Ruiz, C.; Rack, H.J. A new low modulus, biocompatible titanium alloy. In *Titanium'95*; Blenkinsop, P.A., Evans, W.J., Flower, H.M., Eds.; The Institute of Materials: London, 1996; Vol. II, 1760–1767.
41. Niinomi, M.; Hattori, T.; Niwa, S. Biomaterials in orthopedics. In *Material Characteristics and Biocompatibility of Low Rigidity Titanium Alloys for Biomedical Applications*; Yaszemski, M.J., Trantolo, D.J., Lewandrowski, K.U., Hasirci, V., Altobelli, D.E., Wise, D.L., Eds.; Marcel Dekker, Inc., 2004; 41–62.
42. Morita, A.; Fukui, H.; Tadano, H.; Hayashi, S.; Hasegawa, J.; Niinomi, M. Alloying titanium and tantalum by cold crucible levitation melting (CCLM) furnace. *Mater. Sci. Eng. A* **2000**, *A280*, 208–213.
43. Fukui, H.; Kunii, T.; Fujishiro, Y.; Morita, A.; Niinomi, M.; Yamada, A.; Hasegawa, J. Melting method for biomedical Ti–Ta alloy with high melting point. *J. Jpn. Soc. Dent. Mater. Devices* **2000**, *19*, 49–55.
44. Kawasaki, H.; Kuroda, D.; Hanawa, T.; Kuroda, H.; Kobayashi, M.; Kobayashi, T.; Imai, H. Mechanical properties of Ti–Fe–Ta system alloys for biomedical applications. *Proceedings of the Fall Meeting of the Japan Institute of Metals*, 2003; 130.
45. Ou, A.; Chou, T.; Asami, K.; Inoue, A. Properties of Ti–Zr–Nb–Ta system bio-soft titanium alloy. *Proceedings of the Annual Meeting of the Japan Institute of Metals*, 2003; 326.
46. Zhou, Y.L.; Niinomi, M.; Akahori, T. Effect of Ta content on Young's moduli and tensile strength of

- binary Ti–Ta alloys for biomedical applications. *Mater. Sci. Eng. A* **2004**, *A371*, 283–290.
47. Zhou, Y.L.; Niinomi, M.; Akahori, T. Decomposition of martensite α during aging treatments and resulting mechanical properties of Ti–Ta alloys. *Mater. Sci. Eng. A* **2004**, *A371*, 283–290.
 48. Zhou, Y.L.; Niinomi, M.; Akahori, T. Dynamic Young's modulus and mechanical properties of Ti–Hf alloys. *Mater. Trans.* **2004**, *45* (5), 1549–1554.
 49. Uggowitz, P.J.; Bähre, W.F.; Speidel, M.O. Metal injection molding of Ni–Free stainless steels. *Proc. Adv. Powder Metall. Part. Mater.* **1997**, *3*, 18–11318–121.
 50. Yoneyama, T.; Doi, H.; Kobayashi, E.; Kasuga, J.; Fukushima, O.; Murakami, M. Mechanical properties and phase transformation of Ti–Ni super elastic alloy for dental application. *Proc. AISB4 FNB* **2004**, 241–242.
 51. Okuyama, M.; Kobayashi, J.; Take, S.; Itoi, Y. Effect of plasma spray and surface sealing treatment of Ni–Ti shape memory alloy applying to biomaterials. *Mater. Trans.* **2004**, *45* (4), 1045–1053.
 52. Ogata, Y.; Takatugu, M.; Kunimasa, T.; Kobayashi, K.F. Tensile strength and pseudo-elasticity of YAG laser spot melted Ti–Ni shape memory alloy wire. *Mater. Trans.* **2004**, *45* (4), 1070–1076.
 53. Nitta, K.; Watanabe, S.; Masahashi, N.; Hosoda, H.; Hanada, S. Ni-free Ti–Nb–Sn shape memory alloys. In *Structural Biomaterials for the 21st Century*; Niinomi, M., Okabe, T., Taleff, E.M., Lesure, D.R.M., Lippard, H.E., Eds.; TMS, 2001; 25–35.
 54. Kim, H.Y.; Ohmatsu, Y.; Kim, J.I.; Miyazaki, S.; Hosoda, H. Mechanical properties and shape memory behavior of Ti–Mo–Ga alloys. *Mater. Trans.* **2004**, *45* (4), 1090–1095.
 55. Inamura, T.; Fukui, Y.; Hosoda, H.; Wakashima, K.; Miyazaki, S. Relationship between texture and macroscopic transformation strain in severely cold-rolled Ti–Nb–Al super elastic alloy. *Mater. Trans.* **2004**, *45* (4), 1083–1089.
 56. Hosoda, H.; Hosoda, N.; Miyazaki, S. Mechanical properties of Ti–Mo–Al biomedical shape memory alloys. *Trans. MRSJ* **2001**, *26* (1), 243–246.
 57. Ikeda, M.; Komatsu, S.; Nakamura, Y. Effects of Sn and Zr additions on phase constitution and aging behavior of Ti–50 mass% Ta alloys quenched from β single phase region. *Mater. Trans.* **2004**, *45* (4), 1106–1112.
 58. Toyoshima, K.; Hirasawa, T.; Ikeda, M. Effect of Zr addition on phase constitution and heat treatment behavior of Ti–35 mass% Nb alloy. *Proceedings of the Annual Meeting of the Japan Institute of Metals*, 2003; 130.
 59. Maeshima, T.; Nishida, M. Shape memory and mechanical properties of biomedical Ti–Sc–Mo alloy. *Mater. Trans.* **2004**, *45* (4), 1101–1105.
 60. Maeshima, T.; Nishida, M. Shape memory properties of biomedical Ti–Mo–Ag and Ti–Mo–Sn alloys. *Mater. Trans.* **2004**, *45* (4), 1096–1100.
 61. Sakaguchi, N.; Niinomi, M.; Akahori, T. Deformation behaviors of Ti–Nb–Ta–Zr system alloys for biomedical applications. *Mater. Trans.* **2004**, *45* (4), 1113–1119.
 62. Hosoda, H.; Miyazaki, S. Recent topics of shape memory materials and related technology. *J. Jpn. Soc. Mech. Eng.* **2004**, *107* (1028), 509–520.
 63. Niinomi, M. Recent research and development in titanium alloys for biomedical applications and healthcare goods. *Sci. Technol. Adv. Mater.* **2003**, *4*, 445–4541.
 64. Niinomi, M. Cyto-toxicity and fatigue performance of low rigidity titanium alloy, Ti–29Nb–13Ta–4.6Zr, for biomedical applications. *Biomaterials* **2003**, *24*, 2673–2683.
 65. Niinomi, M.; Akahori, T.; Hattopri, Y.; Morikawa, K.; Kasuga, T.; Fukui, H.; Suzuki, A.; Kyo, K.; Niwqa, S. Super elastic functional β type titanium alloy with low Young's modulus for biomedical applications. *J. ASTM Int.* **2005**, *2* (6), 448–464.
 66. Hattori, T.; Morikawa, K.; Niwa, S.; Sato, K.; Niinomi, M.; Suzuki, A. Bone tissue reaction to new β titanium low rigidity alloy: rabbit study on bone healing, remodelling and atrophy in intramedullary fracture fixation. *International Conference on Biomechanics and Annual Scientific Meeting of Taiwanese Society of Biomechanics*, Nov 9–1, 2001; National Yang Ming University: Taipei, Taiwan, 2001; 28.
 67. Niinomi, M. Fatigue characteristics and microstructure of titanium alloys for biomedical applications. *Proceedings Fatigue'2002*; Blom, A.F., Ed.; EMAS, 2002; 3/5, 2073–2083.
 68. Niinomi, M. Cyto-toxicity and fatigue performance of low rigidity titanium alloy, Ti–29Nb–13Ta–4.6Zr, for biomedical applications. *Biomaterials* **2003**, *24*, 2673–2683.
 69. Niinomi, M.; Akahori, T.; Yabunaka, T.; Fukui, H.; Suzuki, A. Fretting fatigue characteristics in air and simulated body environment of newly developed β type titanium alloy for biomedical applications. *J. Iron Steel Inst. Jpn.* **2002**, *88* (9), 553–560.
 70. Maruyama, N.; Nkazawa, K.; Sumita, M.; Sato, M. Effect of stress frequency of fatigue and fretting fatigue life for commercially pure Ti and Ti–6Al–4V alloy in pseudo-body fluid. *J. Jpn. Inst. Biomater.* **2000**, *18* (1), 17–23.
 71. Li, S.J.; Yang, R.; Li, S.; Hao, Y.L.; Cui, Y.Y.; Niinomi, M.; Guo, Z.X. Wear characteristics of Ti–Nb–Ta–Zr and Ti–6Al–4V alloys for biomedical applications. *Wear* **2004**, *257*, 869–876.
 72. Ueda, K.; Narushima, T.; Ouchi, C.; Iguchi, Y. Wear loss and elution of C. P. Ti and titanium alloys in simulated body fluid. *Mater. Sci. Forum* **2005**, *475–479*, 2333–2336.
 73. Tamura, Y.; Yokoyama, A.; Watari, F.; Uo, M.; Kawasaki, T. Mechanical properties of surface nitrided titanium for abrasion resistant implant materials. *Mater. Trans.* **2002**, *43* (12), 3043–3051.
 74. Akahori, T.; Niinomi, M.; Ishimizu, K.; Fukui, H.; Suzuki, A. Effect of thermomechanical treatment on fatigue characteristics of Ti–29Nb–13Ta–4.6Zr alloy for biomedical applications. *J. Jpn. Inst. Metals* **2003**, *67* (11), 652–660.
 75. Niinomi, M. Fatigue characteristics and microstructure of titanium alloys for biomedical applications.

- Proceedings Fatigue'2002; Blom, A.F., Ed.; EMAS, 2002; Vol. 3/5, 2073–2083.
76. Niinomi, M.; Kobayashi, T.; Toriyama, O.; Kawakami, N.; Ishida, Y.; Matsuyama, Y. Fracture characteristics, microstructure and tissue reaction of Ti–5Al–2.5Fe for orthopedic surgery. *Metall. Mater. Trans. A* **1996**, *27A* (12), 3925–3935.
 77. Maruyama, N.; Nakazawa, K.; Sumita, M.; Sato, M. Effect of stress frequency on fatigue and fretting fatigue life for commercially pure Ti and Ti–6Al–4V alloy in pseudo-body fluid. *J. Jpn. Biomater. Soc.* **2000**, *18* (1), 17–23.
 78. Akahori, T.; Niinomi, M.; Fukui, H.; Suzuki, A. Fatigue, fretting fatigue and corrosion characteristics of biocompatible beta type titanium alloy conducted with various thermo-mechanical treatments. *Mater. Trans.* **2004**, *45* (5), 1540–1548.
 79. Niinomi, M.; Akahori, T.; Nakamura, S.; Fukui, H.; Suzuki, A. Friction wear of surface oxidized newly developed beta type titanium alloy for biomedical applications in simulated body environment. *J. Iron Steel Inst. Jpn.* **2002**, *88* (9), 567–574.
 80. Mano, T.; Sugiyama, O.; Shibuya, Y.; Nakayama, H.; Osamu Takai, O. Formation of diamond-like carbon based double-layer film on Ti–6Al–4V substrate by ionization deposition. *Mater. Trans.* **2004**, *45* (5), 1601–1606.
 81. Nishio, K.; Yamaguchi, T.; Era, H.; Katoh, M. Surface modification of titanium using laser beam. *Mater. Trans.* **2004**, *45* (5), 1613–1619.
 82. Yoshinari, M.; Oda, Y. Titanium surface modification with dry process for dental implants. In *Structural Biomaterials for the 21st Century*; Niinomi, M., Okabe, T., Taleff, E.M., Lesuer, D.R., Lippard, H.F., Eds.; TMS, 2001; 129–136.
 83. Yoshinari, M.; Hayakawa, T.; Wolke, J.G.C.; Nemoto, K.; Jansen, J.A. Influence of rapid heating with infrared radiation on RF magnetron-sputtered calcium phosphate coating. *J. Biomater. Res.* **1997**, *37*, 60–67.
 84. Wang, C.K.; Chern, J.H.; Ju, C.P.; Ong, H.C.; Chang, R.P. Structural characterization of pulsed laser-deposited hydroxyapatite film on titanium substrate. *Biomaterials* **1997**, *18*, 1331–1338.
 85. Yoshinari, M.; Ohtsuka, Y.; Dérand, T. Thin hydroxyapatite coating produced by the ion beam dynamic mixing method. *Biomaterials* **1994**, *15*, 529–535.
 86. Cui, F.Z.; Luo, Z.S.; Feng, Q. Highly adhesive hydroxyapatite coatings on titanium alloy formed by ion beam assisted deposition. *J. Mater. Sci. Mater. Med.* **1997**, *8*, 403–405.
 87. Nonami, T.; Naganuma, K.; Kameyama, T. Superplastic bonding of biomaterial-apatite and Ti-alloy. *Mater. Jpn.* **1998**, *37*, 856–858.
 88. Hanawa, T.; Kon, M.; Ukai, H.; Murakami, K.; Miyamoto, Y.; Asaoka, K. Surface modification of titanium in calcium-ion-containing solutions. *J. Biomater. Res.* **1997**, *34*, 273.
 89. Ban, S.; Maruno, S. Effect of temperature on electrochemical deposition of calcium phosphate coatings in a simulated body fluid. *Biomaterials* **1995**, *16*, 977–981.
 90. Kim, H.M.; Miyaji, F.; Kokubo, T. Effect of heat treatment on apatite-forming ability of Ti metal induced by alkali treatment. *J. Mater. Sci.: Mater. Med.* **1997**, *8*, 341–347.
 91. Kasuga, T.; Watanabe, M.; Nogami, M.; Niinomi, M. Calcium phosphate invert glass-ceramic coating joined by self-development of compositionally gradient layers on a titanium alloy. *Biomaterials* **2001**, *22*, 577–582.
 92. Kasuga, T.; Nogami, M.; Niinomi, M.; Hattori, T.; Hench, L.L. Enhancing effect of autoclaving on bioactivity of beta-titanium alloy coated with calcium phosphate glass-ceramic. *Key Eng. Mater.* **2005**, *284–286*, 243–246.
 93. Okazaki, Y. Implant applications of highly corrosion-resistant Ti–15Zr–4Nb–4Ta alloy. *Mater. Trans.* **2002**, *43* (12), 2936–2942.
 94. Okazaki, Y. Dental casting properties of Ti–15Zr–4Nb–4Ta alloy. *Mater. Trans.* **2002**, *43* (12), 3134–3141.
 95. Srimaneepong, V., et al. Strain and displacement of Ti–6Al–7Nb alloy casting for removable partial denture frameworks. *Proc. AISB4 FNB* **2004**, 237–238.
 96. Akahori, T.; Niinomi, M.; Isohama, R.; Suzuki, A. Improvement of mechanical performance of cast alpha/beta titanium alloys for dental applications by thermochemical processing. In *Structural Biomaterials for the 21st Century*; Niinomi, M., Okabe, T., Tallef, E.H., Lesure, D.R., Lippard, H.E., Eds.; TMS, 2001; 91–98.
 97. Eylon, D.; Newman, J.R.; Thome, J.K. Titanium and titanium alloy castings. In *Metals Handbook*, 10th Ed.; Abel, L.A., Kieppura, R.T., Thomas, P., Lampman, H.F., Wheaton, N.D., Eds.; ASM International: Materials Park, OH, 1990; Vol. 2, 634–646.
 98. Yoshitani, S.; Niinomi, M.; Fukunaga, K.; Fukui, H.; Takeuchi, T.; Katsura, S. Dental precision castability of biocompatible beta-type titanium alloy. *CAMP-ISIJ* **2000**, *13*, 1345.
 99. Kobayashi, E.; Doi, H.; Takahashi, M.; Nakano, T.; Yoneyama, T.; Hamanaka, H.J. Castability and mechanical properties of Ti–6Al–7Nb-alloy dental cast. *Jpn. Soc. Dent. Mater. Dev.* **1995**, *14*, 406–413.
 100. Miyazaki, T.; Tani, Y.; Tamaki, Y.; Suzuki, E.; Miyaji, T. Application of CaO to dental investments. *J. Jpn. Soc. Dent. Mater. Dev.* **1987**, *6*, 437–440.
 101. Miyazaki, T.; Tani, Y.; Tamaki, Y.; Suzuki, E.; Miyaji, T. Casting of titanium with calcium investment (part2). Improvement of mold treatment and finess on castings. *J. Jpn. Soc. Dent. Mater. Dev.* **1987**, *6*, 633–638.
 102. Takeuchi, T.; Niinomi, M.; Akahori, T.; Katsura, S.; Fukui, H. Development of dental precision casting process for titanium alloy with high biocompatibility. *Mater. Jpn.* **2005**, *44* (1), 68–70.
 103. Niinomi, M.; Akahori, T.; Takeuchi, T.; Katsura, S. Dental precision casting of Ti–29Nb–13Ta–4.6Zr using calcium mold. *Mater. Sci. Forum* **2005**, *475–479*, 2303–2308.
 104. Niinomi, M.; Akahori, T.; Takeuchi, T.; Katsura, S.; Fukui, H.; Toda, H. Mechanical properties and cytotoxicity of newly designed beta type titanium alloys with low melting points for dental applications. *Mater. Sci. Eng. C.* **2005**, (3), 417–425.

Transdermal Drug Delivery

Brad Phipps
Michel Cormier
Bob Gale
Bill van Osdol
Jay Audett
Rama Padmanabhan
Peter Daddona

ALZA Corporation, Mountain View, California, U.S.A.

T

INTRODUCTION

Transdermal drug delivery is based on the absorption of drug into the skin after topical application. This route of drug administration has a long history dating back to the use of plasters and poultices and was most often used for drug delivery to local skin sites or directly underlying tissue. More recently, the potential of the transdermal route for systemic drug delivery has been recognized and exploited. Transdermal delivery offers many advantages over other modes of drug administration. In contrast to orally delivered drugs, compounds entering the body through the skin escape first-pass metabolism in the liver, often resulting in higher bioavailability. In addition, transdermal delivery can be used in nauseous patients, is little affected by intake of food, and can be easily interrupted. In contrast to intravenous drug delivery, transdermal administration is noninvasive and poses little risk of infection. It is also relatively easy for patients to apply and remove a transdermal system. Because transdermal drug delivery allows continuous delivery of drugs, frequent bolus dosing of drugs with short half lives is avoided. As a result, the side effects, or variability in therapeutic effect due to peaks and troughs in plasma concentration, that are seen with bolus administration are minimized.

Transdermal drug delivery, however, involves the inherent challenge of overcoming the skin barrier. Skin protects the body from the environment very effectively and generally is only permeable to small, lipophilic drugs. Transdermal delivery systems, therefore, not only have to provide drug to the skin under stable conditions and in a format convenient to the patient, but also serve to locally increase the permeability of skin to larger, charged, or hydrophilic drug molecules while minimizing irritation.

Several different strategies have been explored to increase permeability of the skin. Permeability can be modified by chemical additives, use of an electric

field to drive charged molecule across the skin, or mechanical creation of microscopic transport channels through the skin. In this article, we will briefly describe the theory underlying each of these transdermal delivery strategies, the specific demands placed on system components and materials, and some of the applications for each type of system. In addition, we will summarize current knowledge on how to prevent or mitigate irritant skin response to transdermal drug delivery. This review focuses on systems that are either already marketed or in late stages of development. Electroporation and sonophoresis, which rely on the application of a pulsed electric field or pretreatment with ultrasound, respectively, to create transient transport pathways through the skin, are still in early developmental stages and will not be covered here.

BASICS OF SKIN STRUCTURE AND FUNCTION

The primary function of the skin is to protect the body from the environment. The highly specialized outermost layer of the skin, the stratum corneum, forms a physical barrier that prevents loss of water and other vital substances from the body and entry of foreign substances or microorganisms into the body. In addition, antigen-presenting Langerhans cells in the epidermis and dendritic cells in the dermis serve to alert the immune system to any invading microorganisms, creating a second line of defense.

Skin is composed of two fundamentally different layers, the inner dermis and the outer epidermis, which are connected through the basal lamina (Fig. 1). The dermis, which gives the skin its mechanical strength, is a largely acellular structure consisting of interwoven collagen fibers and glycoprotein filaments embedded in amorphous connective tissue. Capillaries in the dermis provide a vascular surface of approximately 1 to 2 cm²

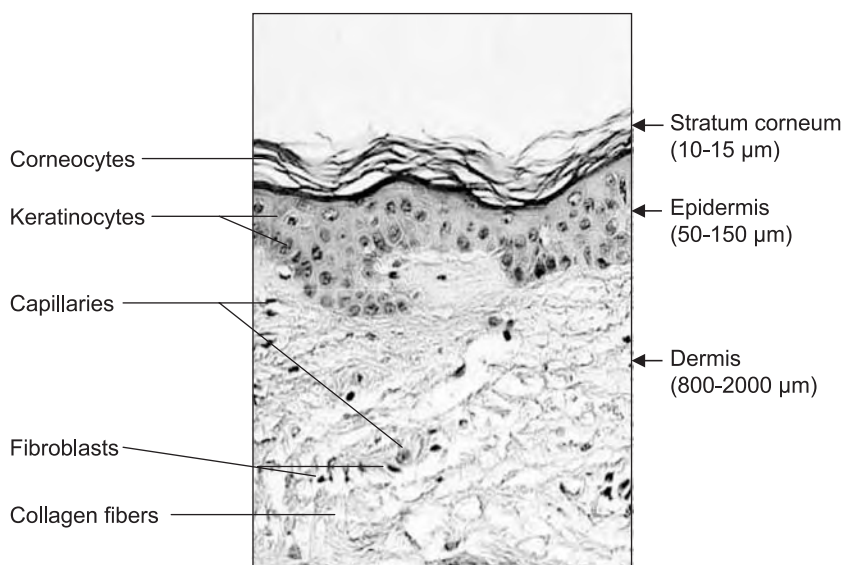


Fig. 1 Cross section of human skin (Masson trichrome staining).

per cm^2 of skin surface for exchange of substances between blood and skin.

In contrast to the dermis, the epidermis is completely avascular and entirely composed of cells. Eighty percent of the cells in the epidermis are derived from keratinocytes, which owe their name to an intracellular abundance of keratin. This filamentous protein forms a cytoskeleton that extends into a supracellular network between keratinocytes through specialized points of cell-cell adhesion (desmosomes). Keratinocytes are continuously generated by the innermost layer of the epidermis, the basal layer or stratum germinativum. As they move to the outside, they differentiate in stages that define the various strata of the epidermis. The entire epidermis renews itself, on average, every 14 days.

From the basal layer, the keratinocytes move into the stratum spinosum, where cellular organelles (lamellar granules) filled with lipids appear in the flattening cells. These granules extrude their lipid contents into the intercellular space in the next layer, the stratum granulosum. Keratinocytes reach their terminally differentiated state as corneocytes in the stratum corneum, the outermost epidermal layer. Corneocytes are devoid of all organelles and filled instead with keratin filaments enmeshed in a protein matrix. A thick band of protein is deposited on the inner surface of the plasma membrane to create the cornified cell envelope, a highly insoluble and chemically resistant barrier, and the plasma membrane is subsequently replaced by a monolayer of ω -hydroxyceramides. The flattened corneocytes are arranged into 15 to 25 cell layers in a regular, interdigitated pattern. Surrounding the cells is a lipid matrix of unusual composition; phospholipids

are virtually absent, and primarily neutral lipids form multiple stacked bilayers reminiscent of membranes. The ω -hydroxyceramide coating of the corneocytes is believed to anchor the cell bricks in this lipid mortar. Together, the corneocytes and the lipid matrix of the stratum corneum create a formidable barrier, which is only disrupted by skin appendages such as hair follicles and sebaceous and sweat glands. These appendages, however, typically cover less than 1% of the surface area and are believed to serve as a route of entry only under special circumstances, e.g., during iontophoresis.

PASSIVE TRANSDERMAL TRANSPORT WITH OR WITHOUT CHEMICAL MODIFICATION

Fundamentals of Passive Transport

Transport of a permeant through the skin is a diffusive process driven by a chemical potential gradient. It can be described by Fick's law, which states that flux of a permeant occurs down its concentration gradient and is proportional to the change of concentration over the transport distance. Flux through a given medium depends on the specific properties of both the medium and the permeant, as measured by the diffusion coefficient. Under steady-state conditions, the relationships described by Fick's laws of diffusion can be used to determine the flux of a drug through the skin, the effective transport distance, or the apparent diffusion coefficient, which expresses the combined contributions of various microscopic diffusion processes.^[1]

The skin's stratum corneum (SC) forms the main obstacle for transdermal permeation. Passive transport

is believed to occur primarily through the SC's intercellular lipid matrix, and the ability of a molecular species to cross the SC depends on both solubility and diffusivity of the permeant in the SC. One principal determinant of diffusivity is molecular volume, which can be estimated from molecular weight;^[2] molecular shape and conformational flexibility also play a role. Solubility of a permeant in the SC depends on its heat of fusion (ΔH_f) and its melting point (T_f), which are measures of the crystal lattice energy of a permeant in its solid state, and the heat of mixing, which can be calculated from the solubility parameters of the permeant and the SC lipids. This relationship is shown in Eq. 1, where x_2 is the mole fraction solubility of the permeant in the solvent (SC lipid), V_2 is the molar volume of the permeant, Φ_1 is the volume fraction of the solvent in the saturated solution, and δ_1 and δ_2 are the solubility parameters of the solvent and solute, respectively:^[3]

$$\ln x_2 = \frac{\Delta H_f}{RT} \left(\frac{T_f - T}{T_f} \right) - \frac{V_2 \phi_1^2}{RT} (\delta_1 - \delta_2)^2 \quad (1)$$

The solubility parameters are the square root of the cohesive energy densities and can be estimated by group contribution methods.^[4] Permeants such as nicotine that combine miscibility with lipid and water, low heat of fusion, and small molecular size are expected to penetrate the SC efficiently.

Permeability of the SC also depends on a number of environmental factors. Increases in the hydration state of the SC due to skin site occlusion or increases in temperature lead to greater fluidity of the SC, resulting in enhanced permeability for some compounds. Chemicals such as anionic surfactants or sulfoxide compounds can also increase the fluidity of the SC by disrupting inter- or intracellular structures. In-vitro and in-vivo methods for measuring skin permeability have been reviewed in detail by Shaw and colleagues.^[5]

System Structure and Function

The basic function of a transdermal drug delivery system is to place drug in contact with the skin under occlusive conditions. A chemical enhancer may be added to the system to raise solubility of the drug in the SC or to increase the SC's fluidity and, consequently, the drug's diffusivity. Most transdermal drug systems are multilayered assemblies minimally composed of a backing layer, a drug-in-adhesive matrix, and a peelable liner (Fig. 2A). More complex transdermal designs include additional layers such as a separate drug reservoir or a rate-controlling membrane (Fig. 2B). The amount of drug released is proportional to the area of drug-containing parts in

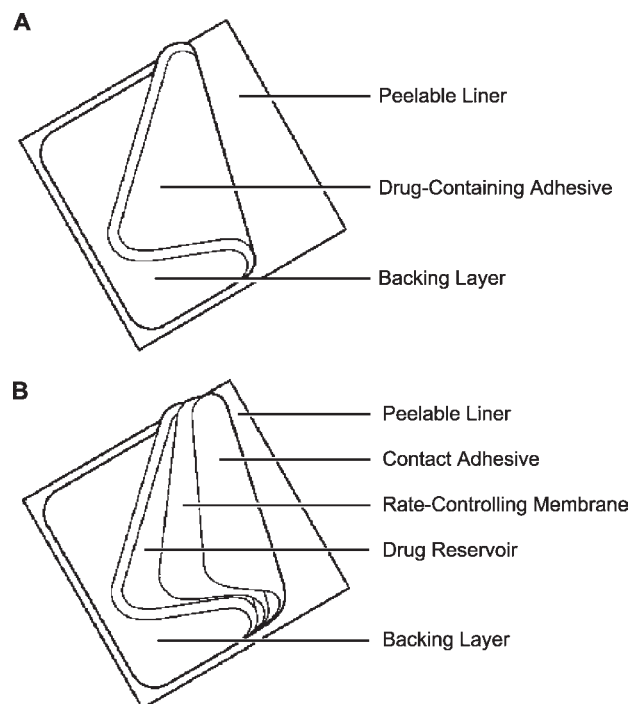


Fig. 2 Schematic diagram of two transdermal drug delivery system designs: (A) matrix and (B) membrane-controlled.

contact with the skin. In matrix systems, the rate of transdermal drug absorption depends directly on the skin/adhesive partition coefficient and the diffusion coefficient for drug in the skin and is therefore affected by inter- or intraindividual differences in skin permeability. In membrane-controlled systems, drug must pass through a rate-controlling membrane inserted between the drug reservoir and the adhesive before it contacts the skin. Diffusion of drug through the membrane restricts the rate of transdermal drug absorption in individuals or at body sites with higher skin permeability, thus limiting inter- and intraindividual variability in absorption.

System Components

Development of transdermal systems is a multidisciplinary effort requiring pharmaceutical, dermatological, physical, chemical, and materials science expertise. All system components of transdermal systems must be compatible with the skin, causing little or no irritation and sensitization, and are routinely tested for their topical irritation potential in animal models. In addition, the design of transdermal systems requires careful consideration of the solubility limits, partition and diffusion coefficients, and chemical stability of all system components to achieve desired drug delivery rates and ensure system stability.

T

Backing Materials

Backing materials must be flexible while possessing good tensile strength. Commonly used materials are polyolefins, polyesters, and elastomers in clear, pigmented, or metallized form. Elastomeric materials such as low-density polyethylene conform more readily to skin movement and provide better adhesion than less compliant materials such as polyester. Backing materials should also have low water vapor transmission rates to promote increased skin hydration and, thus, greater skin permeability. In systems containing drug within a liquid or gel, the backing material must be heat-sealable to allow fluid-tight packaging of the drug reservoir using a process known as form-fill-seal.

Drug Reservoirs

Drug reservoirs can be composed of liquids such as ethanol or silicone fluid, polymers such as ethylene–vinyl acetate copolymer, polyisobutylene, or silicone elastomer, or polymer blends such as a polyvinyl alcohol/polyvinyl pyrrolidone blend. Drug release from a polymeric reservoir can be regulated by adjusting the chemical structure of the polymer, using appropriate polymer blends, or adding plasticizers. In matrix systems, the polymeric drug reservoir doubles as the skin adhesive.

Rate-Controlling Membranes

Rate-controlling membranes can be used to regulate the diffusion of both drug and chemical penetration enhancers. Materials for rate-controlling membranes include microporous polypropylene and ethylene–vinyl acetate copolymers. Membrane permeability for drug or enhancer is adjusted through appropriate choice of vinyl acetate content or membrane thickness.

Drugs

Drugs are preferred candidates for transdermal delivery if they exhibit a short half life or are subject to extensive first-pass metabolism and require a parenteral dose of less than about 20 mg/day. Preferred drugs are further characterized by low skin toxicity, a molecular weight of less than about 500 D, a high octanol/water partition coefficient ($1 \leq \log P \leq 2.5$), and/or solubility in mineral oil and water of greater than about 1 mg/ml. Insufficient solubility of the permeant in the skin can be improved by inclusion of a biocompatible chemical enhancer, resulting in increased skin permeation.

Permeation Enhancers

Certain chemical agents enhance the permeation of drug through the stratum corneum. These agents are typically included in the drug reservoir and, like the drug, partition into and diffuse through the skin. It is therefore critical that they have a low potential for irritation or sensitization. Permeation enhancers function by increasing diffusivity or solubility of a drug in the SC, but may have similar effects in the polymeric materials of transdermal devices. Ethanol, a commonly used enhancer with excellent biocompatibility, increases the solubility of drugs in the SC. Ionic and nonionic surfactants, such as oleic acid, glyceryl laurate, and lauryl pyroglutamate, enhance drug permeation by perturbing the ordered structure of the lipid matrix within the SC, leading to higher diffusivity of the drug. Their presence in SC lipid lamellae may also increase drug solubility. Other aliphatic alcohols, fatty acids and their esters, and essential oils have also been studied for use as permeation enhancers.^[6]

Pressure-Sensitive Adhesive

The pressure-sensitive adhesive (PSA) affixes the transdermal drug delivery system firmly to the skin. Adhesive formulations must provide rapid conformation to rough skin, while limiting cold flow during storage or wear. For multiday wear systems, aggressive PSAs with low storage moduli are selected, which permit adhesion to rough skin. Adhesives must be skin-compatible, causing minimal irritation or sensitization, and removable without inflicting physical trauma or leaving residue. In addition, they must be able to dissolve drug and excipients in quantities sufficient for the desired pharmacological effect without losing their adhesive properties and skin tolerability.

PSAs used in commercially available transdermal systems include polyacrylate, polyisobutylene, and polysiloxane. Polyacrylate, the most commonly used adhesive, is a copolymer of up to three different monomers. By varying the composition of the polymer, its properties can be adjusted to a range of drug delivery requirements. In general, all acrylic adhesives are polar in character, allowing them to absorb moisture readily and to maintain adhesion to wet skin. They also dissolve most drugs well, enabling high drug loading of polyacrylate matrices. Polyisobutylenes (PIBs), in contrast, are characterized by a low solvent capacity for drugs. PIBs are often used in membrane-controlled systems where the initial burst of drug released from the adhesive layer should be limited. PIB-based adhesives are mixtures of high and low molecular weight polymers, which provide cohesion and tackiness, respectively. By adjusting the

composition of the PIB formulation, cold flow and adhesiveness can be customized for each system. Silicone adhesives are characterized by low allergenicity. Similar to PIBs, silicones dissolve most drugs poorly and regulate tackiness and cohesion through polymer size. Molecular weight of silicones, however, can be hard to control during storage of drug-adhesive formulations, since drugs containing amine groups can catalyze further polymerization in silicone adhesives retaining residual silanol groups. To address this problem, special silicones have been developed that are rendered resistant to amine-catalyzed condensation through end-capping of silanol functional groups.

Peelable Liners

Peelable liners protect the adhesive from contamination and loss of tackiness. They are usually composed of polyethylene or polyester coated with a release layer of silicone or fluorocarbon agents.

System Applications

The following examples highlight the advantages of transdermal drug delivery systems over other routes of drug administration and illustrate the respective strengths of different system designs.

Duragesic[®] (developed by ALZA Corporation, Mountain View, CA, for Janssen Pharmaceutica, Inc., Titusville, NJ) provides controlled transdermal delivery of fentanyl for patients in chronic pain who require continuous opioid analgesia and whose pain cannot be managed by lesser means. Patches are applied to the upper body for a duration of three days. Since the amount of drug released from the transdermal system is proportional to the skin contact area of the drug-containing silicone adhesive, dosage strengths can be adjusted by varying the patch size. Patches are available in four sizes ranging from 10 to 40 cm², corresponding to dosage strengths from 25 to 100 µg/hour. Ethanol gelled with hydroxypropyl cellulose is incorporated into the drug reservoir as a solubilizing agent, but also partitions into the skin and increases the flux of fentanyl through the skin. Upon application of the system to the skin, a small loading dose of fentanyl is released from the adhesive layer; subsequently, drug release is limited by the system's rate-controlling membrane, which is made of ethylene-vinyl acetate copolymer. Control of drug release from the transdermal system reduces variations in blood concentrations of fentanyl over the delivery period. Breakthrough pain, a common phenomenon with dosing via the oral or intravenous route, is thereby avoided.

Climara[®] (developed by 3M, St. Paul, MN, for Schering, Berlin, Germany) is a drug-in-adhesive matrix system for the delivery of 17β-estradiol in hormone replacement therapy. The polyacrylic adhesive layer, which is laminated to an elastomeric backing membrane, contains both the drug and a mixture of fatty acid esters serving as permeation enhancers. Patches ranging in size from 6.5 to 25 cm² are applied to the abdomen for a duration of seven days, delivering 17β-estradiol at rates from 25 to 100 µg/day. Transdermal delivery of estradiol eliminates the effects of first-pass metabolism in the liver and minimizes fluctuations in plasma levels seen after oral administration. In addition, Climara[®], which is comfortable and inconspicuous to wear, provides a convenient once-per-week dosing schedule.

Ortho EVRA[®] (Ortho-McNeil Pharmaceutical, Raritan, NJ) is a transdermal system for contraception that was introduced in 2002. The 20 cm² patch delivers norelgestromin (150 µg/day) and ethinyl estradiol (20 µg/day) over seven days. This delivery system uses a novel PIB-based adhesive with a hydrophilic polyvinyl pyrrolidone filler, which improves adhesion under moist conditions and increases solubility of drugs in the adhesive. Ortho EVRA[®] combines key advantages of implantable or injectable contraceptives with those of oral contraceptives. Similar to implantable or injectable systems, a transdermal delivery system eliminates the daily fluctuations in plasma levels of contraceptive hormones seen with oral contraceptives and offers a convenient dosing schedule leading to a high degree of user compliance. Like oral contraceptives, Ortho EVRA[®] is fully user controlled, allowing contraception to be terminated at any time.

Additional commercialized transdermal products and their compositions are summarized in Table 1.

ELECTRICALLY ASSISTED TRANSDERMAL TRANSPORT

Fundamentals of Electrotransport

Passive transdermal delivery is generally limited to lipophilic and highly potent compounds of low molecular weight. Electrically assisted transdermal delivery is used to enhance the delivery of polar, hydrophilic, and charged compounds across the skin by means of two fundamentally different approaches: electroporation and iontophoresis. Electroporation employs brief high-voltage pulses to create transient pathways for transport across the skin. Iontophoresis is the application of low-level electrical fields over long periods to drive small, charged agents across the skin.

The Nernst-Planck equation^[7] describes the transport of chemical agents across a membrane under the

Table 1 Transdermal therapeutic systems (TTS) approved in the United States

Active ingredient	Product name	System type	Drug reservoir	Enhancer	Adhesive ^a	Dosage ^b
Scopolamine	Transderm Scop [®]	Membrane controlled	Multilaminar	None	PIB/mineral oil	1.0 mg in 3 d
Clonidine	Catapres-TTS [®]	Membrane controlled	Multilaminar	None	PIB/mineral oil/ colloidal silicon dioxide	0.1, 0.2, 0.3 mg/day for 7 d
Fentanyl	Duragesic [®]	Membrane controlled	Form-fill-seal	Ethanol	Silicone	1.8, 3.6, 5.4, 7.2 mg in 3 d
Lidocaine	Lidoderm [®]	Matrix	Adhesive	Glycerin	Acrylate	~21 mg in 12 h
Testosterone	Androderm [®]	Microporous membrane	Form-fill-seal	Ethanol, glyceryl monoleate, methyl laurate, glycerin	Peripheral acrylate	2.5, 5 mg in 1 d
	Testoderm [®] TTS	Membrane controlled	Form-fill-seal	Ethanol	PIB/mineral oil	5 mg in 1 d
	Testoderm [®]	Matrix	Ethylene vinyl acetate	None	None	4, 6 mg in 1 d
	Testoderm [®] with Adhesive	Matrix	Ethylene vinyl acetate	None	Striped PIB/colloidal silicon dioxide	6 mg in 1 d
Nicotine	Habitrol [®]	Matrix	Adhesive/pad	None	Acrylate	7, 14, 21 mg in 1 d
	Nicotrol [®]	Matrix	Adhesive	None	PIB	5, 10, 15 mg in 16 h
	Nicoderm [®] -CQ [®]	Membrane controlled	Multilaminar	None	PIB	7, 14, 21 mg in 1 d
	Prostep [®]	Matrix	Form-fill-seal	None	Acrylate	11, 22 mg in 1 d
	Deponit [®]	Matrix	Adhesive	Propylene glycol monolaurate	PIB	2.4, 4.8 mg in 12 h
Nitroglycerin	Minitran [®]	Matrix	Adhesive	Fatty acid esters	Acrylate	7.2 mg in 12 h
	Nitrodisc [®]	Matrix	Microreservoirs	Polyethylene glycol, isopropyl palmitate	Peripheral acrylate	4.8, 7.2, 9.6 mg in 1 d
	Nitrodur [®]	Matrix	Adhesive	None	Acrylate	2.4, 3.6, 4.8, 7.2, 9.6 mg in 12 h
	Transderm-Nitro [®]	Membrane controlled	Form-fill seal	None	Silicone	1.2, 2.4, 4.8, 7.2 mg in 12 h

17 β -estradiol	Alora [®]	Matrix	Adhesive	Sorbitan monooleate	Acrylate	0.2, 0.3, 0.4 mg in 4 d
	Climara [®]	Matrix	Adhesive	Fatty acid esters	Acrylate	0.175, 0.35, 0.525, 0.7 mg in 7 d
	Esclim [®]	Matrix	Adhesive	Octyldodecanol, dipropylene glycol	Polyvinyl-acetate	0.1, 0.15, 0.2, 0.3, 0.4 mg in 4 d
	Estraderm [®]	Membrane controlled	Form-fill-seal	Ethanol	PIB/mineral oil	0.2, 0.4 mg in 4 d
	FemPatch [®]	Matrix	Adhesive	Monolaurate	Acrylate	0.175 mg in 7 d
	Vivelle [®]	Matrix	Adhesive	Oleic acid, dipropylene glycol	PIB/mineral oil/acrylate	0.15, 0.2, 0.3, 0.4 mg in 4 d
	Vivelle-DOT [®]	Matrix	Adhesive	Oleyl alcohol, dipropylene glycol, povidone	Silicone/acrylate	0.15, 0.2, 0.3, 0.4 mg in 4 d
	Combi Patch [®]	Matrix	Adhesive	Oleic acid, dipropylene glycol, povidone	Silicone/acrylate	0.2 mg in 4 d 0.56, 1 mg in 4 d
	Norethindrone					
	Norelgestromin ethinyloestradiol	ORTHO-EVRA [®]	Matrix	Adhesive	Crospovidone, lauryl lactate	PIB/polybutene

^aPIB: polyisobutylene.

^bd, days.

influence of an electrical field as the sum of three contributions: passive diffusion, electromigration, and electro-osmosis. For charged species, the contribution of passive diffusion is small due to the lipophilic nature of the intercellular matrix of the stratum corneum. Ionizable components of the skin, such as fatty acids, make the occurrence of electro-osmotic transport, the convective flow of solvent and species dissolved in it, possible. While electro-osmosis plays only a minor role in the transport of small ionic agents, it may be important for the transport of both neutral and charged large molecules including proteins, peptides, and oligonucleotides.^[8] For example, electro-osmosis has been used to extract glucose from the epidermis for the purpose of monitoring blood glucose levels.^[9]

For transport of small ionic agents, electromigration is the dominant process. The amount of ionic agent transported per second (N) has been shown to be linearly proportional to the applied current (I) over a wide range of currents^[10] and depends on the molecular weight (M) and the charge (z) of the agent according to Eq. 2, which is derived from Faraday's law of electrolysis:

$$N = \frac{t_d \cdot I \cdot M}{z_d \cdot F} \quad (2)$$

F denotes Faraday's constant and t_d the transport number, which expresses the charge carried by the ionic agent d as a fraction of the total transported charge. The transport number t_d is unique for each ionic agent and is a function of its mobility, charge, and concentration in the skin as compared with those of other migrating species.

System Structure and Function

Drug delivery from an electrotransport system can be tightly controlled by circuitry that continuously adjusts the magnitude of the applied voltage in order to maintain a specific current. Patient-controlled electrotransport systems are under development, and future combination with biosensors may ultimately provide systems operating with closed-loop feedback. A typical transdermal electrotransport system (Fig. 3) is composed of a source of electrical energy and two electrodes in contact with ionically conductive reservoirs, which are placed on the patient's skin and contain either the drug (donor electrode assembly) or a pharmacologically inactive electrolyte (counter electrode assembly).

Electrotransport systems can be designed as completely integrated or as modular products. In integrated systems, the drug and delivery mechanism are packaged together, making those systems simple to use

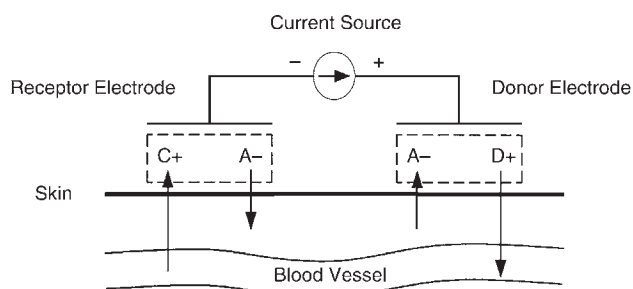


Fig. 3 Schematic diagram of an electrotransport device showing the flow of ions in response to an applied voltage.

and convenient to wear. Modular systems, in contrast, use disposable drug patches attached to a reusable power source and electronic circuitry.

System Components

Power source

Skin resistance is greatest at the start of current application and, thus, determines the system voltage required to achieve the desired current into the skin. A voltage of 10 to 20 V, easily provided by a single lithium button cell and voltage boost circuitry, is usually sufficient to achieve the desired current in a few seconds to minutes.

Control circuitry

While a simple field-effect transistor with a feedback resistor is sufficient to control the current delivered from a power source over a wide range of skin resistances, microcontrollers or customized application-specific integrated circuits (ASICs) allow precise modulation of the applied voltage to maintain the required current and also provide safety monitoring features and dosing and system maintenance information for the user.

Electrodes

The electrodes apply an electric field across the skin by converting electric current from the battery into ionic current through the reservoirs, skin, and body. The electrochemical reactions occurring at the electrodes lead to constant change in the chemical environment of the reservoirs. Long-lived nonconsumable electrodes, which are made from electrochemically inert materials such as stainless steel or platinum, lead to the accumulation of hydrolytic products and, thus, a change in pH in the reservoirs. Consumable electrodes, in

contrast, contain electroactive species such as silver or silver chloride, which are oxidized at the anode or reduced at the cathode during the passage of current,^[11] substantially reducing hydrolysis.

Formulation

The formulation includes the ingredients in the drug and counter reservoirs, which typically consist of a solvent, a matrix-forming material, and a drug salt or a biocompatible salt, respectively. Buffers, antimicrobial agents, antioxidants, and additional electrolyte salts or permeation enhancers may be added to affect the rate of delivery, biocompatibility, or product shelf life. The counter ion of the drug salt is often selected to preferentially react with electrochemically-generated species from the electrode; a chloride, for example, will react with oxidized silver ions from a silver anode.

System Applications

A number of iontophoretic systems are currently available or under development (Phoresor[®], IOMED, Salt Lake City, UT; Dupel[®] System, Empi, St. Paul, MN; lidocaine patch, Vyteris, Inc., Fair Lawn, NJ; IontoPatch[™], Birch Point Medical, St. Paul, MN; E-TRANS[®] fentanyl, ALZA Corporation, Mountain View, CA). Three of these systems are described in more detail to illustrate differences in system design (modular versus integrated) and applications (local versus systemic delivery).

The IOMED Phoresor[®] is a modular iontophoretic system, composed of a reusable controller/power source unit connected by a cable to disposable silver/silver chloride electrodes attached to hydrogel reservoirs (GelSponge[®]). The hydrogels must be hydrated with the drug solution of choice immediately before use, and a current of up to 4 mA resulting in current densities of up to 700 $\mu\text{A}/\text{cm}^2$ can be applied. While the Phoresor[®] is a highly versatile system, it is relatively costly and requires the patient to remain stationary during treatment. One of the primary uses of the Phoresor[®] is in the delivery of lidocaine for local anesthesia.

A partially integrated system for the iontophoretic delivery of lidocaine is being developed by Vyteris Inc. The Vyteris system, designed to produce topical anesthesia within 10 minutes, consists of a disposable patch prefilled with lidocaine solution and a small, reusable dose controller.

E-TRANS[®] fentanyl (ALZA Corporation, in Phase III clinical development, Fig. 4) is a fully integrated, nonreusable system for the patient-controlled transdermal delivery of fentanyl for management of acute pain. The self-adhesive system uses silver and silver

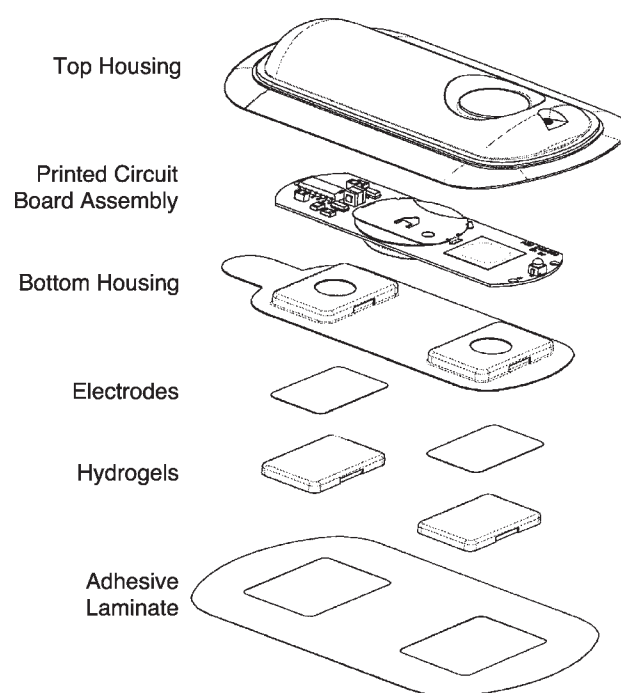


Fig. 4 Exploded view of an E-TRANS[®] system showing major components: top housing, printed circuit board assembly, a bottom housing containing reservoirs for placement of electrodes and hydrogels, and an adhesive laminate (diagram from International Patent Publication Number WO 96/39222).

chloride electrodes with polyvinyl alcohol-based hydrogels and a PIB adhesive. In clinical studies, patients experiencing pain could activate the system to have a current of 170 μA applied for 10 minutes, resulting in the systemic delivery of 40 μg of fentanyl. Systems can be worn for up to 24 hours, during which time up to 80 doses can be administered.

MECHANICALLY ASSISTED TRANSPORT

Fundamentals of Mechanically Assisted Transport

Increased skin permeability can be achieved by mechanical disruption of the stratum corneum, the skin's main permeation barrier, through abrasion by scraping or tape stripping. However, the rate of drug absorption through abraded skin is hard to control or predict. In contrast to surface abrasion methods, mechanical disruption of the stratum corneum through the formation of microchannels can be achieved in a controlled and reproducible manner. The artificially formed, superficial microchannels can serve as pathways for even very large or highly charged drug molecules, such as

proteins or nucleic acids. Since the microchannels are confined to the epidermis, which does not contain nerve endings, their formation causes little or no pain or bleeding. Microchannels through the skin typically remain open for about one to 24 h, depending on their size and whether the site is under occlusion. In contrast to surface abrasion, microchannels cause little or no irritation to the skin.

System Structure and Function

Different methods are being developed for the controlled formation of microchannels through the stratum corneum. In needleless injection systems, drug molecules in liquid (e.g., Medi-Jector™, Antares Pharma, Exton, PA, or Biojector™, Bioject, Inc., Bedminster, NJ) or dry powder form (e.g., Powder-Ject™, PowderJect Pharmaceuticals Plc, Oxford, England) are propelled toward the skin with such velocity that they are able to penetrate the skin, in essence, forming their own channels. Depending on the system, drug can be deposited in the epidermis, the dermis, subcutaneous adipose tissue, or intramuscular tissue. Other devices are based on the formation of multiple pores in the stratum corneum, through which drug can either diffuse passively or be actively driven by iontophoresis. Formation of multiple pores can be achieved either through the brief application of high temperatures to the skin (Altea Therapeutics, Atlanta, GA) or by piercing the stratum corneum with hollow or solid microprojections of uniform shape and length.^[12–15] Both methods allow the control of pore depth and diameter. Use of drug-coated solid microprojections permits drug to be directly deposited into the epidermis, achieving high delivery efficiencies.

System Components

Needleless injection systems

Needleless injection systems are syringelike devices that use mechanical compression force to propel a jet stream of drug solution from a cylinder. The force and surface area of the jet stream determine the depth of penetration (e.g., Antares Pharma). For delivery of dry powder or coated gold particles, a jet stream of an inert gas such as helium is typically used to drive the drug into the skin (e.g., Powderject Pharmaceuticals Plc). A novel mechanism currently in development is based on the rapid depression of a piston through the contraction of shape memory alloy fibers, activated by a brief electrical pulse (Norwood Abbey, Victoria, Australia).

Thermal creation of micropores

The MicroPor™ system (Altea Therapeutics) is composed of three integrated elements: an array of metallic heating filaments, placed directly against the skin, a control unit containing a power source and electrical circuitry, and a drug reservoir. An electrical pulse from the control unit activates delivery of a brief pulse of thermal energy from the heating filaments, resulting in the formation of micropores through the stratum corneum. Drug then passively diffuses from the reservoir through the micropores into the skin.

Creation of pores through microprojections

Microprojections are composed of silicon or titanium. The microprojection systems developed by Becton Dickinson^[12] and at the Georgia Institute of Technology^[13] are produced by etching a silicon substrate. Macroflux® microprojection arrays (ALZA Corporation; Refs. [14–16]) are composed of titanium and fabricated by chemical etching followed by a forming step. The shape of microprojections, which are organized into arrays of up to 1000 projections per cm², is highly uniform within the array. Microprojections are generally less than 200 μm long to avoid penetration into the dermis. They can be used to pretreat a skin site before application of a topical drug formulation or integrated with a drug reservoir for passive or iontophoretic delivery. Alternatively, microprojections can be directly coated with drug and attached to an adhesive backing. For the Macroflux® system (ALZA Corporation), use of a spring-loaded, self-actuated device to apply the microprojection array to the skin helps to ensure uniform and reproducible penetration (Fig. 5).

System Applications

At present, only needleless injection devices are commercially available in the United States for use as an alternative to conventional needle-based injection systems. Development of mechanical microchannel-based transdermal systems may help to overcome transdermal drug delivery challenges for highly charged and high molecular weight drug compounds. In particular, these technologies may enhance the delivery of commercially relevant therapeutic proteins and vaccines, which are currently administered only by parenteral injection.

STRATEGIES FOR ENHANCING BIOCOMPATIBILITY

Several commercially available transdermal systems have been reported to produce local irritation or

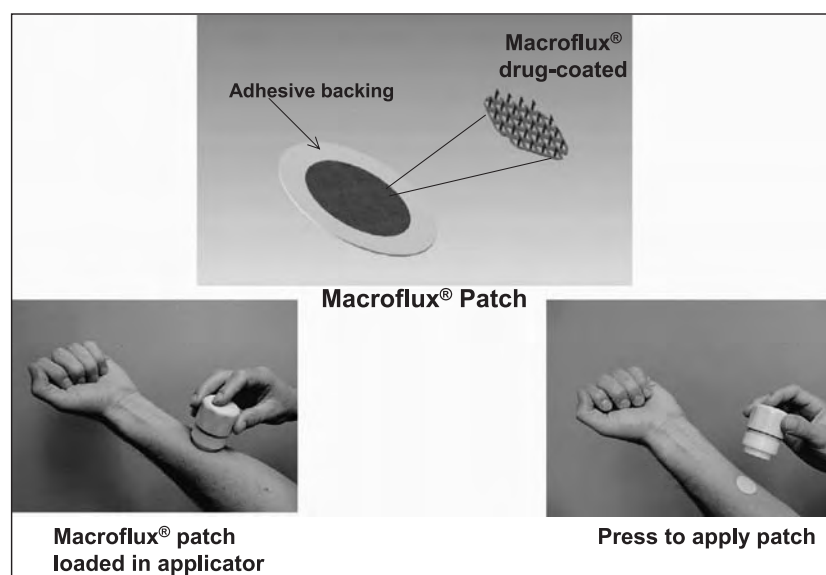


Fig. 5 A prototype Macroflux[®] transdermal system.

sensitization.^[17] Since transdermal delivery systems interface with the skin, they are inherently able to cause skin irritation. In addition, transdermal delivery leads to high local drug concentrations that can promote irritation or sensitization at the delivery site. Here, we will briefly outline different types of skin reactions to transdermal drug delivery and the various strategies that have been developed to minimize these local side effects.^[18]

System-Induced Skin Reactions

Transdermal delivery systems rely on occlusion, which causes changes in the pH and hydration state of the skin. These changes may result in increased absorption not only of drug, but also of permeation enhancers or inactive system components. Usually well-tolerated excipients such as propylene glycol and methyl cellulose have been found to cause contact dermatitis in a small fraction of the population. Similarly, ethanol can cause skin reactions when used as a permeation enhancer, particularly in individuals with aldehyde dehydrogenase deficiency. Irritation caused by passive transdermal systems, however, is typically mild and can be avoided by adequate product engineering and judicious choice of the system components, the application site, and the duration of occlusion.

In iontophoretic systems, the electrochemical processes add to the risk of skin irritation. Use of consumable electrodes that minimize the electrolysis of water prevents pH changes that can cause local irritation. All iontophoretic systems, however, cause hyperemia. While the ensuing erythema is usually mild

and self-limited, it increases with length of application and current density. High current densities in areas of low resistance, such as skin defects, can result in focal irritation and ulceration, and care should be taken to apply iontophoretic systems to normal, intact skin. The potential for skin reaction can be further reduced by uniformly lowering skin resistance through creation of artificial transport channels in the stratum corneum (see section on mechanical modification of transdermal transport) or by adjusting the pH of the electrode formulations.

Skin behaves like an ionic exchange membrane with an isoelectric point (pI) of four to five, becoming negatively or positively charged at pH values above or below the pI, respectively. Increasing the pH of the anode formulation was found to decrease skin resistance and erythema during transport of positively charged molecules from the anode. A reduction in erythema was also observed at the cathode site for transport of negative species as the pH of the cathode formulation was decreased.

Drug-Induced Skin Reactions

Even drugs with low systemic toxicity are capable of producing an irritant or allergic skin reaction when delivered transdermally. Many drugs can produce erythema and even edema through their vasoactive properties, which may become evident only at the high concentrations achieved locally by transdermal delivery. Irritation due to generation of reactive or toxic metabolites in the skin is uncommon for most transdermally delivered drugs. Most irritation seems to be

T

related to lipophilicity of a drug, since drugs with an octanol/water partition coefficient of greater than 10^4 were shown to be cytotoxic at concentrations easily achieved in any layer of the skin during transdermal transport.^[19] In addition to irritation, allergic reactions (delayed hypersensitivity) have been reported to occur upon transdermal delivery of clonidine and many other drugs from a wide range of therapeutic classes. Repeated application of transdermal systems over prolonged periods of time increases the likelihood of immune responses.

Control of Local Drug Concentrations

Irritation and, to a lesser extent, sensitization depend on local drug concentration, which is determined by drug flux into and out of the delivery site. Drug flux into the skin depends on skin permeability, which can vary up to tenfold between subjects and up to 100-fold between skin sites. Differences in skin absorption can be minimized through the use of rate-controlling membranes in passive transdermal systems or through use of iontophoresis, in which drug flux directly depends on the current applied. Removal of drug from the delivery site can be accelerated by codelivering a vasodilating agent with the drug, as shown by reduction of the irritation potential of the drug tetracaine upon codelivery of methyl nicotinate. Any decrease in the long-lasting irritation response to the drug, however, comes at the expense of a short-lasting, but visible erythema induced by the vasodilating agent. Another strategy aims at reducing the intracellular accumulation of drug in acidic organelles such as lysosomes. Codelivery of the lysosomotropic agent ammonium chloride or the ionophore monensin with the drug chloroquine inhibited intracellular accumulation of chloroquine and dramatically reduced chloroquine-induced skin irritation.

Drug Modification

Attempts have been made to manufacture a prodrug with reduced lipophilicity and, thus, decreased skin toxicity compared with the original drug. A related approach is to deliver salts of the acid or basic drug instead of its neutral form. These strategies work best for iontophoretic delivery, since passive transdermal transport depends on a drug's lipophilicity. Other approaches include camouflaging the drug by complexing it with cyclodextrins or incorporating it in micelles or liposomes. It is still uncertain, however, if transport of such aggregates is feasible.

Inhibition of Metabolism

Contact sensitization requires binding of sensitizing moieties (haptens) derived from the drug molecule to proteins, which are then processed intracellularly in acidic endosomal compartments and presented in association with major histocompatibility-complex proteins on the extracellular membrane. Since most drugs of interest for transdermal delivery are stable chemical entities, metabolism is thought to be a requisite for haptentization. In concordance with this concept, codelivery of a sensitizing agent with inhibitors of a common metabolic pathway, the monoamino oxidase pathway, reduced sensitization. There also is indication that codelivery of lysosomotropic agents, which raise the endosomal compartment pH, may inhibit sensitization responses by inhibiting intracellular processing of haptentized proteins.

Control of Inflammatory Response

If reduction of the irritation/sensitization potential of a drug or delivery system is not feasible, the skin response can be controlled by codelivery of nonspecific anti-inflammatory agents. A synthetic anti-inflammatory peptide that lacks the side effects typically associated with glucocorticoids was shown to reduce irritation caused by iontophoretic delivery of chlorpromazine. Similarly, passive codelivery of hydrocortisone with chlorpheniramine or benzoyl peroxide inhibited sensitization to these drugs. Hydrocortisone also inhibited erythema and edema resulting from iontophoretic delivery of metoclopramide. Since hydrocortisone has never been observed to cause local side effects such as skin thinning, and only a small fraction of its normal endogenous production is needed to provide local anti-inflammatory effects, this compound may be a good alternative to more potent glucocorticoids for inhibition of mild to moderate inflammatory responses.

CONCLUSION

Transdermal drug delivery is a rapidly developing field due to the advantages it provides over other routes of drug administration. The use of transdermal systems is expected to increase in the future as new methods for maximizing drug transport through skin and minimizing skin irritation are perfected. The transdermal delivery of macromolecules such as proteins, nucleic acids, and vaccines, which cannot withstand the gastrointestinal environment and currently require invasive delivery methods, is of particular research and development interest. Here, we have presented a brief

overview of the theory of transdermal drug transport and the types of systems and strategies applied to different delivery challenges. Further information on the topics presented here can be found in Refs. [20–23]. By incorporating the latest advances in materials science, electronics, microfabrication, and skin biology, transdermal systems of the future will be able to deliver a broader spectrum of pharmaceutical agents and provide more complex delivery profiles.

ARTICLES OF FURTHER INTEREST

Controlled Release; Drug Delivery, Controlled

REFERENCES

1. Crank, J.; Park, G.S. *Diffusion in Polymers*; Academic Press: London, 1968; 1–39.
2. Mitragotri, S.; Johnson, M.E.; Blankschtein, D.; Langer, R. An analysis of the size selectivity of solute partitioning, diffusion, and permeation across lipid bilayers. *Biophys. J.* **1999**, *77* (3), 1268–1283.
3. Hildebrand, J.H.; Prausnitz, J.M.; Scott, R.L. *Regular and Related Solutions*; Van Nostrand Reinhold: New York, 1970.
4. Yalkowsky, S.H.; Dannenfelser, R.-M.; Myrdal, P.; Simamore, P. Unified physical property estimation relationships (UPPER). *Chemosphere* **1994**, *28*, 1657–1673.
5. Shaw, J.E.; Prevo, M.; Gale, R.; Yum, S.I. Percutaneous Absorption. In *Biochemistry and Physiology of the Skin*; Goldsmith, L.A., Ed.; Oxford University Press: Rochester, NY, 1991; Vol. II, 1447–1479.
6. Asbill, C.S.; El-Kattan, A.F.; Michniak, B. Enhancement of transdermal drug delivery: Chemical and physical approaches. *Crit. Rev. Ther. Drug Carr. Syst.* **2000**, *17* (6), 621–658.
7. Burnette, R.J. Iontophoresis. In *Transdermal Drug Delivery: Development Issues and Research Initiatives*; Hadgraft, J., Guy, R.H., Eds.; Marcel Dekker: New York, 1989; 247–291.
8. Pikal, M.J. The role of electroosmotic flow in transdermal iontophoresis. *Adv. Drug Deliv. Rev.* **1992**, *9*, 201–237.
9. Tamada, J.A.; Garg, S.; Jovanovic, L.; Pitzer, K.R.; Fermi, S.; Potts, R.O. Noninvasive glucose monitoring. Comprehensive clinical results. *JAMA* **1999**, *282* (19), 1839–1844.
10. Phipps, J.B.; Padmanabhan, R.V.; Lattin, G.A. Iontophoretic delivery of model inorganic and drug ions. *J. Pharm. Sci.* **1989**, *78* (5), 365–369.
11. Scott, E.R.; Phipps, J.B.; Gyory, J.R.; Padmanabhan, R.V. Electrotransport Systems for Transdermal Delivery: A Practical Implementation of Iontophoresis. In *Handbook of Pharmaceutical Controlled Release Technology*; Wise, D.L., Ed.; Marcel Dekker: New York, 2000; 617–659.
12. Mikszta, J.A.; Alarcon, J.B.; Brittingham, J.M.; Sutter, D.E.; Pettis, R.J.; Harvey, N.G. Improved genetic immunization via micromechanical disruption of skin-barrier function and targeted epidermal delivery. *Nat. Med.* **2002**, *8* (4), 415–419.
13. Henry, S.; McAllister, D.V.; Allen, M.G.; Prausnitz, M.R. Microfabricated microneedles: A novel approach to transdermal drug delivery. *J. Pharm. Sci.* **1998**, *87* (8), 922–925.
14. Lin, W.Q.; Cormier, M.; Samiee, A.; Griffin, A.; Johnson, B.; Teng, C.-L.; Hardee, G.E.; Daddona, P.E. Transdermal delivery of antisense oligonucleotides with microprojection patch (Macroflux[®]) technology. *Pharm. Res.* **2001**, *18* (12), 1789–1793.
15. Matriano, J.A.; Cormier, M.; Johnson, J.; Young, W.A.; Buttery, M.; Nyam, K.; Daddona, P.E. Macroflux[®] microprojection array patch technology: A new and efficient approach for intracutaneous immunization. *Pharm. Res.* **2002**, *19* (1), 63–70.
16. Cormier, M.; Daddona, P.E. Macroflux Technology for Transdermal Delivery of Therapeutic Proteins and Vaccines. In *Modified-Release Drug Delivery Technology*; Roberts, M., Hadgraft, J., Rathbone, M.J., Eds.; Marcel Dekker: New York, 2002; 589–598.
17. Patil, S.M.; Hogan, D.J.; Maibach, H.I. Transdermal Drug Delivery Systems: Adverse Dermatologic Reactions. In *Dermatotoxicology*, 5th Ed.; Marzulli, F.N., Maibach, H.I., Eds.; Taylor & Francis: New York, 1996; 389–396.
18. Kydonieus, A.F.; Wille, J.J. *Biochemical Modulation of Skin Reactions: Transdermals, Topicals, Cosmetics*; CRC Press: New York, 2000.
19. Mize, N.K.; Johnson, J.A.; Hansch, C.; Cormier, M. Quantitative structure-activity relationship and cytotoxicity. *Curr. Probl. Dermatol.* **1995**, *23*, 224–229.
20. Bronaugh, R.L.; Maibach, H.I. *Percutaneous Absorption. Drugs–Cosmetics–Mechanism–Methodology*, 3rd Ed.; Marcel Dekker, Inc.: New York, 1999.
21. Walters, K.A. *Dermatological and Transdermal Formulations*; Marcel Dekker: New York, 2002.
22. Couvreur, P.; Duchene, D.; Green, P.; Junginger, H.E. *Minutes, Transdermal Administration, A Case Study, Iontophoresis*; Editions de Santé, Pubs.: Paris, 1997.
23. Smith, E.W.; Maibach, H.I. *Percutaneous Penetration Enhancers*; CRC Press: New York, 1995.

Transgenic and Gene-Targeted Organs

Brenda M. Ogle

Transplantation Biology Program, and the Department of Physiology, Mayo Clinic, Rochester, Minnesota, U.S.A.

Jeffrey L. Platt

Transplantation Biology Program, and the Departments of Immunology, Pediatrics, and Surgery, Mayo Clinic, Rochester, Minnesota, U.S.A.

INTRODUCTION

Today nearly 85,000 individuals await organ transplantation.^[1] Some predict that with new developments in medicine and biotechnology this number will increase substantially. This communication will consider how genetically engineered animals might be used to supply organs needed for transplantation.

One new development that may especially increase the demand for transplantation is molecular and genomic diagnosis. Advances in molecular diagnostics, genomics, and possibly proteomics may soon make it possible to predict the onset of disease long before the symptoms are manifest and the course once disease has become apparent. With this information, the physician may be tempted even more than before to recommend “preemptive” transplantation to spare patients from lethal diseases or from adding the risks of organ failure to the risks of transplantation.^[2] As only one example, molecular diagnostics and genomics are approaching the levels of specificity and sensitivity that may make it possible to diagnose cancer before it can be localized by imaging. Extending preemptive transplantation to adults with high likelihood of tumor formation could add as many as 89,000 potential recipients per year to the list of those awaiting kidney transplants. A second development that may change the demand for transplantation is increasing life expectancy. As the population ages, the potential for organ failure increases and hence the demand for organ donors increases.

Clearly, transplanting human organs will not address the ever-increasing demand for organ replacement. One potential solution to this problem is to use animal organs and tissues instead of human organs and tissues for transplantation, that is, xenotransplantation. The potential benefits of xenotransplantation extend beyond addressing the shortage of human organs for transplantation. For example, xenotransplantation might be a preferred way to treat certain diseases that afflict human but not animal organs (i.e., HIV and hepatitis). Xenotransplantation could also offer a way of dealing with acute organ failure due, for example, to new epidemics of myocarditis, viral pneumonia, or

nephritis. Xenotransplantation might also be used to enable the growth of human stem cells into organs in a surrogate animal.^[3] Finally, xenotransplantation could provide an opportunity for manipulation of the donor organ instead of the host to promote organ acceptance. Manipulation of this type might include gene delivery.^[4] Key to the use of xenotransplantation as a way of replacing organs and as a means of gene transfer is genetic engineering. How genetic engineering can be applied toward this end is the central focus of this entry.

HISTORY OF XENOTRANSPLANTATION

The first serious attempts at xenotransplantation were made in the early years of the 20th century. Princeteau and Jaboulay each tried to connect animal kidneys to human blood vessels using a newly developed technique—the vascular anastomosis.^[5–7] Animals were used as a source of organs then because using humans seemed unethical. The animal kidneys grafted to the sick patients functioned only briefly. Failure of the organs was attributed to technical difficulties associated with the surgical procedure. Xenotransplantation and subsequent attempts at allotransplantation (human to human transplantation) always failed; the reasons for failure were hotly debated. Protein incompatibility, infection, technical failure, and immunity were among the postulated causes. In the early 1940s researchers found that the immune system may be prohibiting the success of transplantation, and in the 1960s, the availability of immunosuppressive drugs brought proof that transplantations failed because of immune rejection and this problem could be overcome with immunosuppression.^[8,9]

The advent of immunosuppressive drugs made it possible to undertake organ transplantation for the treatment of disease. However, there remained the question of what source of organs might be used. Early transplants were undertaken using living related donors.^[10] Indeed, the first successful transplant was from an identical twin.^[10] However, most subjects with

renal failure did not have a living donor much less a twin available to provide a kidney. Hence, some turned once again to animals as a potential source of kidneys. In 1963, Reemtsma transplanted chimpanzee kidneys into 13 patients treated with immunosuppressive therapy. One patient survived for nearly 9 mo after receiving the organ, before dying from an electrolyte imbalance (the xenograft appeared normal).^[11–13] This relative success triggered a resurgence of interest in xenotransplantation and led to the performing of 14 kidney transplants, 8 heart transplants, and 11 liver transplants at various centers. Unfortunately, these attempts were less successful, with approximate survival averaging 15 days (Table 1).^[14,15]

Today, the transplantation of animal organs or tissues is needed more than ever but is yet to be routinely applied in the clinical setting due to several hurdles. First, there may be physiological limitations of the animal tissue in a foreign environment. Second, there exists a potential for transmission of microbial organisms that could give rise to novel infections in the

human population [e.g., porcine endogenous retrovirus (PERV)]. Third, and most crucial, the immune system of the recipient causes severe, incessant rejection of the transplant (Fig. 1).^[33,34] This entry will focus especially on the problem of rejection because this problem is the main barrier to xenotransplantation.

IMMUNOLOGIC HURDLES OF XENOTRANSPLANTATION

All xenografts elicit an immune response by the recipient of the graft; however, the impact of the immune response is dictated by the manner in which a xenograft is connected to the recipient. In the case of whole-organ transplants, the blood vessels of organ xenografts are in direct contact with recipient blood and the components of the immune system contained therein. It is primarily the interaction of these immune components with endothelium lining the blood vessels that leads to the distinct types of vascular disease associated with

T

Table 1 Historical summary of whole-organ xenotransplantation

Year	Recipients (surgeon)	Xenotransplant	Survival	Reference
1905	1 (Princeteau)	Rabbit kidney	Brief	[5]
1906	2 (Jaboulay)	Porcine/goat kidney	Brief	[6]
1909	1 (Unger)	Macaque kidney	<2 days	[16]
1914	1 (Ullman)	Canine kidney	Brief	[17]
1923	1 (Neuhof)	Sheep kidney	9 days	[18]
1963/64	13 (Reemtsma)	Chimpanzee (12), monkey (1) kidneys	One case, 9 mo	[11–13]
1964	3 (Traeger)	Chimpanzee kidney	<49 days	[19]
1964	1 (Hume)	Chimpanzee kidney	1 day	[20]
1964	6 (Starzl)	Baboon kidney	<60 days	[21]
1964	1 (Hitchcock)	Baboon kidney	5 days	[22]
1964	1 (Hardy)	Chimpanzee heart	2 hr	[23]
1965	2 (Goldsmith)	Chimpanzee kidney	4 mo	[18]
1966	1 (Starzl)	Chimpanzee liver	<1 day	[18]
1966	1 (Cortesini)	Chimpanzee kidney	31 days	[18]
1969	2 (Starzl)	Chimpanzee liver	<9 days, <2 days	[24]
1969	1 (Bertoye)	Baboon liver	<1 day	[25]
1969	1 (Marion)	Chimpanzee heart	4 hr	[26]
1970	1 (Leger)	Baboon liver	3 days	[27]
1970	1 (Marion)	Baboon liver	<1 day	[28]
1971	1 (Motin)	Baboon liver	3 days	J. Motin, as quoted in [29]
1974	1 (Starzl)	Chimpanzee liver	14 days	[30]
1977	1 (Barnard)	Baboon heart	5 hr	[31]
1977	1 (Barnard)	Chimpanzee heart	4 days	[31]
1984	1 (Bailey)	Baboon heart	20 days	[32]
1992	1 (Starzl)	Baboon liver	70 days	[15]
1993	1 (Starzl)	Baboon liver	26 days	[15]

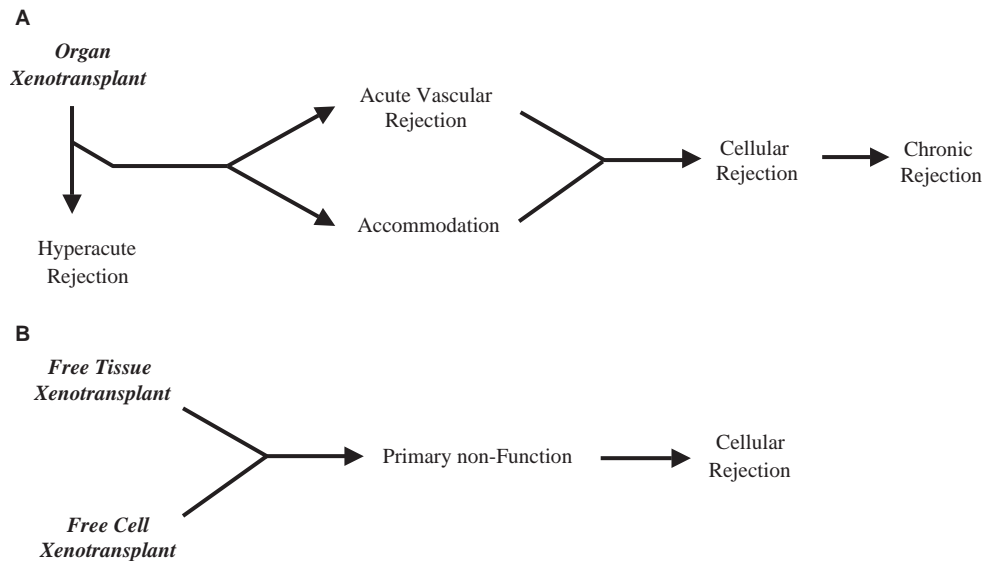


Fig. 1 Biological response to xenotransplantation.

xenogeneic organ rejection. The types of vascular disease seen in this setting are depicted in Fig. 1 and will be described below. Cell and tissue xenografts, in contrast, receive blood supply by ingrowth of blood vessels of the recipient. In this setting the blood vessels of the graft provide a barrier between immune reactants and the transplanted tissue. Below, we describe the types of rejection seen in whole-organ xenografts and how genetic engineering might be used to avert each type of rejection. Rejection of cellular and tissue xenografts is beyond the scope of this communication and so we refer the reader to a recent review.^[3]

The discussion of rejection that follows will focus specifically on the rejection of porcine xenografts. Pigs are the preferred animal for xenotransplantation. Pigs have large litters and a relatively short gestation period (114 days). In addition, porcine organs are very similar to human organs in size and structure. The large litter size and short gestation make housing in “specific pathogen-free” facilities, necessary to avert transfer of pathogens to the organ recipient, economically feasible.

Hyperacute Rejection

Hyperacute rejection is the first and most severe type of rejection seen in xenogeneic organ transplants.^[35] Hyperacute rejection is characterized by immediate bleeding into the graft and by the formation of platelet microthrombi, which in concert lead to graft destruction within several minutes to a few hours. It is very likely that the first xenografts performed in patients with renal failure were destroyed by hyperacute rejection and not by “technical factors.”

Hyperacute rejection is initiated by the binding of xenoreactive natural antibodies to the endothelial lining of donor blood vessels. Xenoreactive natural antibodies are found in the blood of all normal individuals and bind to antigens (i.e., proteins, sugars, etc.) on the surface of cells of other species. Xenoreactive natural antibodies in humans bind especially to one saccharide consisting of terminal Gal α 1-3Gal residues.^[36–38] The binding of xenoreactive antibodies to this sugar activates the complement system of the recipient. The complement system consists of 20 or more proteins in the blood. The primary function of the complement is to attach to foreign organisms and either directly kill or facilitate ingestion of the organism by white blood cells. It is the complement system acting on donor blood vessels that destroys the graft.

Complement is activated in solution and could target one’s own cells. Under normal circumstances, however, activation of complement on a person’s cells is attenuated by regulatory proteins in the plasma and on the surface of cells. These proteins protect healthy cells from inadvertent injury. However, complement regulatory proteins appear to be species specific, operating most effectively in a homologous environment.^[39,40] Thus, in the xenotransplant, complement of the recipient is not adequately controlled and so causes severe, unrestrained injury to the graft.

Several therapies have been used successfully to avert hyperacute rejection. For example, hyperacute rejection can be avoided through depletion of xenoreactive antibodies from the recipient.^[41] Depletion of xenoreactive antibodies can be accomplished by using immunoabsorbent columns in conjunction with immunosuppressive agents such as cyclophosphamide,

mycophenolate, leflunomide, or similar new agents, to prevent rapid antibody resynthesis.^[42] Another way of preventing hyperacute rejection is to inhibit the complement. Systemic inhibitors of complement such as cobra venom factor, soluble complement receptor Type 1 (SCR1), and antibodies against complement component C5 and γ -globulin have been used with relative efficacy.^[43–48] However, these treatments render the treated individual susceptible to severe infection because the complement system is generally incapacitated. A safer, more effective, and more enduring approach has been to genetically alter animals to express human complement regulatory proteins that can provide control of complement in the transplant. Human CD55, CD46 (membrane cofactor protein), and CD59 (protectin) are among the proteins that have been expressed in transgenic animals.^[49–54] Organs from these animals have been shown to control activation of complement in the xenograft. This was the first use of transgenic animals as organ donors. We describe these efforts and discuss future possibilities below.

All of the therapies discussed above—antibody depletion, systemic inhibition of complement, and use of organs transgenic for human complement regulatory proteins—successfully avert hyperacute rejection. Unfortunately, none of these approaches has allowed long-term survival of organ xenografts. In every case the grafts ultimately develop acute vascular rejection over a period of days to weeks.

Acute Vascular Rejection

When hyperacute rejection is prevented, as discussed above, a xenogeneic organ is subject to acute vascular rejection. Acute vascular rejection can destroy a xenograft over days to weeks. It is characterized by focal ischemia, intravascular fibrin deposition, and thrombosis.^[55–57] Because hyperacute rejection can be easily prevented, acute vascular rejection has been thought to be the main barrier to clinical application of xenotransplantation.

Acute vascular rejection appears to be initiated by binding of xenoreactive antibodies to the xenograft. Unlike hyperacute rejection, in which complement is activated in large amounts by these antibodies, in acute vascular rejection complement is activated only sparingly. Activated complement or antibody alone triggers profound changes in the physiology of endothelium, called endothelial activation. Activated endothelial cells express procoagulant molecules such as tissue factor and proinflammatory molecules such as E-selectin and cytokines. The pathogenesis of acute vascular rejection is summarized in Fig. 2.

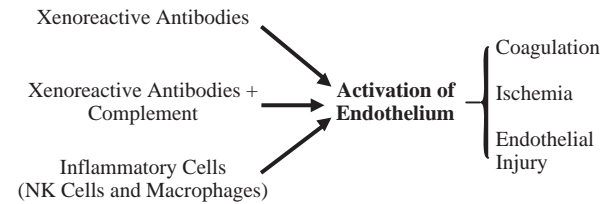


Fig. 2 Molecular mechanisms of acute vascular rejection.

Developing therapies to prevent acute vascular rejection has been a much more daunting challenge than developing therapies to prevent hyperacute rejection. Part of the challenge in developing therapies for acute vascular rejection is that very low levels of bound antibody or activated complement can initiate disease.^[3] The antibodies that initiate acute vascular rejection include anti-Gal α 1-3Gal antibodies but may also include antibodies against other foreign antigens.^[58] As a corollary, the eradication of one antigen or depletion of antibodies that recognize it may be insufficient to prevent acute vascular rejection. Still, reduction of xenograft antigen expression by genetic engineering and inhibition of the complement cascade may be a reasonable first step in addressing this problem.

Another challenge is that acute vascular rejection may be promoted by direct action of natural killer cells and monocytes on the graft. Natural killer cells and monocytes circulate in the blood and are thought to provide initial defense against some infectious organisms. These cells may act on a xenograft as if it were a foreign organism or parasite. The reason why these cells attack xenogeneic cells is not entirely known but implicated in the process are natural killer cell and monocyte-associated receptors, lectins specific for Gal α 1-3Gal and receptors (i.e., Fc) that recognize bound antibody that triggers activity of the cells. Infiltration of these inflammatory cells leads to synthesis of tissue factor and plasminogen activator inhibitor type 1, which promote coagulation and synthesis of E-selectin and cytokines, which promote inflammation. In this manner the injury triggered by xenoreactive antibodies is further exacerbated.^[59–61]

Steps to overcome acute vascular rejection include induction of tolerance. Tolerance in xenotransplantation implies cessation of the synthesis of xenoreactive antibodies. One approach to achieving this end is the transplantation of donor bone marrow, or bone marrow genetically manipulated, to bring about expression of relevant donor antigens and subsequently the deletion or suppression of B-cells of the recipient, which secrete antidonor antibodies. Bracy, Sachs, and Iacomini used this approach by expressing α 1,3galactosyltransferase in bone marrow cells by gene transfer.^[62] Administration of the Gal α 1-3Gal⁺ bone marrow cells

in α 1,3 galactosyltransferase knockout mice decreased production of anti-Gal α 1-3Gal antibodies.

Another approach for avoiding xenograft acute vascular rejection is to bring about accommodation of the graft. Accommodation is an acquired resistance to humoral injury. It was first observed in the successful transplantation of ABO-incompatible renal allografts.^[63] This demonstrated that if anti-A or anti-B antibodies from the recipient were depleted before placement of an incompatible allograft, prolonged graft survival could be achieved in spite of the presence of an intact complement system and the eventual return of anti-A or anti-B antibodies. The biological mechanisms underlying accommodation are still poorly understood and so while it is known that initial depletion of donor antibodies is necessary, the depletion duration is unclear. Some have suggested that accommodation might be initiated via expression of antiapoptotic genes such as hemoxygenase 1, A 20, and bcl2.^[64-68] If successful, this could offer a period of protection sufficient to trigger accommodation. Careful studies detailing the biochemical nature of accommodation are needed to identify the molecular nature of this process and hence the targets for genetic engineering.

T Cell-Mediated Rejection

The T cell-mediated response to a xenograft is thought to be quite potent, but because acute vascular rejection has yet to be circumvented, the impact of this response is not fully known. T cells can recognize major histocompatibility antigens as they are expressed on pig cells, similar to the T cell-mediated response to allografts.^[34,69-71] However, it is not known whether the cell-to-cell interactions that contribute to T cell recognition occur fully between disparate species. In a related manner, T cell repressor function could be defective in the xenogeneic environment (as discussed in the case of hyperacute rejection). What is known is that nearly every porcine protein can give rise to immunogenic peptides and thus the severity of the T cell-mediated response may be very high.^[58] Humoral immunity raised against the xenograft as discussed above may make cellular rejection more severe.^[72] If the cell-mediated immune response to xenografts is more severe than that to allografts, an important question is whether current immunosuppressive agents used in allotransplantation will be effective in the xenotransplant setting. Studies in experimental animals using cellular xenografts suggest that these regimens will suffice. It is conceivable that genetic engineering might be used to avert cell-mediated rejection, but first molecular targets must be more clearly defined.

GENETICALLY ENGINEERED ORGANS

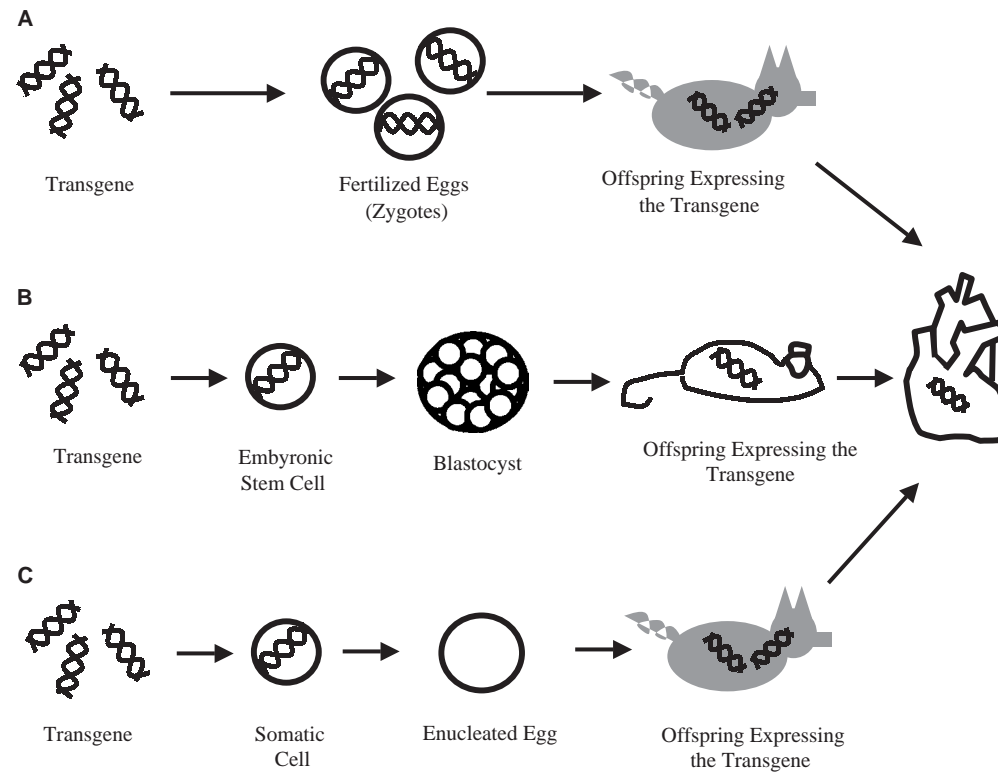
Methods

Genetic engineering provides a means to stably add, delete, or alter the expression of genes in animals. And as discussed above, the organs of transgenic animals may provide a means to avert rejection of transplanted organs. There are several means by which gene expression may be changed in animals. We describe each method below and refer the reader to two reviews of the technology necessary for genetic engineering.^[73,74]

In one approach, a DNA fragment or construct including coding and regulating sequences is micro-injected into the pronucleus of a zygote (Fig. 3A). The injected zygote develops further following its introduction into the oviduct of a pseudopregnant recipient female. The DNA integrates randomly into the chromosome and therefore germ cells may convey transgene expression allowing transmission to subsequent generations. The random nature of the insertion results in variation in the number and location of integrated copies and unpredictable control of expression. Additionally, the length of the DNA is limited.

In a second approach, genetic modification is performed in embryonic stem cells, which are then introduced into preimplantation mouse blastocysts. The genetic modification takes place *in vitro* and may include random integration or homologous recombination (Fig. 3B). The latter involves replacing a wild-type allele with a mutant allele. Embryonic stem cells exhibiting the desired mutation (gene knockout) are selected and subsequently injected into the blastocoele of a recipient blastocyst. The result is a chimeric animal with both the recipient blastocyst cells and the injected embryonic stem cells. If gametes of a chimera develop from the injected embryonic stem cells, the mutant allele present can be transmitted to subsequent generations. Unfortunately, embryonic stem cells are not currently available for pigs.

A third method involves genetic manipulation of cultured somatic cells followed by fusion of a somatic cell or nucleus to an enucleated oocyte, termed nuclear transfer (Fig. 3C). Prior to fusion, donor somatic cells are grown under special conditions to maintain quiescence. At this time it is possible to make genetic modifications (presumably to multiple genetic loci), and select cells with the desired modification. Under the correct conditions, many of which are still dimly defined and poorly understood, the introduced nucleus can lead to the formation of an embryo. This method is very inefficient as seen in the past success rates. Of over 400 manipulated embryos, 1 sheep survived, of 250 embryos, 8 calves survived, and of approximately 2500 embryos, 31 newborn mice survived.^[75-77] Along with



T

Fig. 3 Methods to genetically engineer organs.

the poor rate of success this method is very expensive. However, this is the only effective way to inactivate specific genes in large animals.

Applications

Genetic engineering of organs has been conducted to avert rejection of xenogeneic organs. One way this has been undertaken has been through the generation of lines of pigs transgenic for human complement regulatory proteins. Organs from transgenic pigs expressing human complement regulatory proteins CD55, CD46 (membrane cofactor protein), and CD59 (protectin), either singly or in combination are protected from hyperacute rejection.^[49–54] This result underscores the importance of complement regulation as a determinant of xenograft outcome.

Another way to quench hyperacute rejection using transgenic animals may be to modify expression of xenogeneic antigens. Recent reports describe the successful knockout of the α 1,3 galactosyltransferase gene in pigs using nuclear transfer.^[78–84] The piglets have yet to be extensively tested for expression of the Gal α 1-3Gal antigen. In the interim homozygous (–/–) fetal-pig fibroblast cells have been produced from heterozygous embryos using a novel selection method. These cells express low (~1%) but detectable levels of the Gal α 1-3Gal antigen.^[85] Whether cells expressing low-level antigen are as susceptible to hyperacute rejection as their wild-type counterparts remains to be seen. As cloning technology continues to advance, it is possible that much simpler and less expensive technology will be developed. One possibility proposed by Nottle et al. is the use of the Cre/loxP recombination system to create a transgene acceptor site in the porcine genome.^[86,87]

Table 2 Genetically engineered organs to avert xenograft rejection

Target mechanisms	Genetic manipulation of donor organ
Antibody–antigen interaction	Generation of α -1,3-galactosyltransferase knockout pigs Generation of pigs that overexpress an alternate glycosyltransferase (i.e., H-transferase)
Complement activation	Generation of pigs transgenic for human complement regulatory proteins (i.e., CD55, CD46, CD59)
Activation of endothelium	Suppression of procoagulant or proinflammatory genes

Cloned pigs containing the transgene acceptor site could be used as embryo donor for microinjection. Then, by injecting the transgene and Cre-recombinase construct into the pronucleus of a fertilized egg, it may be possible to achieve single-copy targeted gene insertion.

Another means by which antigen expression may be modulated is by overexpression of an alternate glycosyltransferase (H transferase) such that an H antigen resides at the terminus of some sugar chains instead of Gal α 1-3Gal. To date, these endeavors have shown *ex vivo* decrease in complement-mediated lysis of isolated porcine cells with increased H-transferase.^[88,89] Animal lines transgenic for this glycosyltransferase have not been generated and so complete elimination of hyperacute rejection *in vivo* is still uncertain.

Genetically engineered organs have been proposed and developed to avert rejection of xenografts; however, the use of genetically engineered organs could be extended in many ways. First, genetic engineering could be used to enhance the physical characteristics of an organ. For example, it has been shown that the innervation and overall morphology of the atrioventricular node in pigs is quite different from that of humans.^[90,91] Such differences could alter heart rate and contractility, thereby increasing the potential for arrhythmias. Hence, pig hearts could be genetically engineered to harbor a more similar innervation network to that of humans. Second, animal organs may be altered such that endogenous viruses are knocked out, or mutated and therefore rendered incapable of infecting human recipients.

CONCLUSIONS

The technology for producing genetically modified animals was made possible by the fusion of fundamental discoveries in several seemingly unrelated disciplines including reproductive physiology, molecular biology, and genomics. Genetically modified animals were at first viewed as a tool for investigating disease and studying new therapies. Those in the field of transplantation propose to extend the utility of genetically modified animals by using their organs to address the overwhelming demand for organ replacement. Genetically modified organs have been generated with intent to avert rejection. However, the use of transgenic organs could extend beyond this to provide therapeutic benefit to the recipient.

ARTICLES OF FURTHER INTEREST

Gene-Activated Matrix; Gene Therapy

REFERENCES

1. UNOS. Transplantation statistics. UNOS Update **1995**, *11*, 10–11.
2. Cascalho, M.; Ogle, B.M.; Platt, J.L. Emerging strategies in kidney transplantation. In *Chronic Kidney Disease, Dialysis, and Transplantation*; Pereira, B.J.G., Sayegh, M., Blake, P., Eds.; Elsevier: Philadelphia, 2005; 750–758.
3. Cascalho, M.; Platt, J.L. Xenotransplantation and other means of organ replacement. *Nat. Rev. Immunol.* **2001**, *1*, 154–160.
4. Ogle, B.M.; Platt, J.L. Genetic therapies and xenotransplantation. *Expert Opin. Biol. Ther.* **2002**, *2*, 299–310.
5. Princeteau, M. Greffe renale. *J. Med. Bord.* **1905**, *36*, 549–555.
6. Jaboulay, M. De reins au pli du coude par sutures arterielles et veineuses. *Lyon Med.* **1906**, *107*, 575–577.
7. Carrel, A. La technique opératoire des anastomoses vasculaires et la transplantation des viscères. *Lyon Med.* **1902**, *98*, 858.
8. Gibson, T.; Medawar, P.B. The fate of skin homografts in man. *J. Anat.* **1943**, *77*, 299–310.
9. Calne, R.Y. The rejection of renal homografts: inhibition in dogs by 6-mercaptopurine. *Lancet* **1960**, *1*, 417–418.
10. Merrill, J.P.; Murray, J.E.; Harrison, H.; Guild, W.R. Successful homotransplantation of the human kidney between identical twins. *J. Am. Med. Assoc.* **1958**, *160*, 277–282.
11. Reemtsma, K.; McCracken, B.H.; Schlegel, J.U.; Pearl, M.A.; Pearce, C.W.; DeWitt, C.W.; Smith, P.E.; Hewitt, R.L.; Flinner, R.L.; Creech, O. Renal heterotransplantation in man. *Ann. Surg.* **1964**, *160*, 384–410.
12. Reemtsma, K.; McCracken, B.H.; Schlegel, J.U.; Pearl, M. Heterotransplantation of the kidney: two clinical experiences. *Science* **1964**, *143*, 700–702.
13. Reemtsma, K. Renal heterotransplantation from non-human primates to man. *Ann. N.Y. Acad. Sci.* **1969**, *162*, 412–418.
14. Hardy, M.A.; Todd, G.; Reemtsma, K. Xenotransplantation. In *Bone Marrow & Organ Transplantation*; Elsevier Science Publishers: New York, 1984; 515–534.
15. Starzl, T.E.; Murase, N.; Tzakis, A.G.; Fung, J.J.; Todo, S.; Demetris, A.J.; Manes, R.; Marino, I.R.; Valdivia, L. Clinical xenotransplantation. *Xenotransplantation* **1994**, *1*, 3–7.
16. Unger, E. *Ueber Nierentransplantationen*; Wehnschr: Berlin, 1909; vol. xlvi.
17. Ullman, E. Tissue and organ transplantation. *Ann. Surg.* **1914**, *60*, 195–219.
18. Daar, A.S. Animal-to-human organ transplants—a solution or a new problem? *Bull. World Health Org.* **1999**, *77*, 54–61.
19. Traeger, J.; Gonin, A.; Delahaye, J. Pericardite urémique a evolution constrictive subaigue. A propos d'une observation anatomo-clinique chez un Brightique traite par les epurations extra renale; au long cours. *Lyon Med.* **1964**, *211*, 383–401.

20. Hume, D.M.; Magee, J.H.; Prout, G.R. Studies of renal homotransplantation in man. *Ann. N.Y. Acad. Sci.* **1964**, *120*, 578–606.
21. Starzl, T.E.; Marchioro, T.L.; Peters, G.N.; Kirkpatrick, C.H.; Wilson, W.E.C.; Porter, K.A.; Rifkind, D.; Ogden, D.A.; Hitchcock, C.R.; Waddell, W.R. Renal heterotransplantation from baboon to man: experience with 6 cases. *Transplantation* **1964**, *2*, 752–776.
22. Hitchcock, C.R.; Kiser, J.C.; Telander, R.L.; Seljeskog, E.L. Baboon renal grafts. *J. Am. Med. Assoc.* **1964**, *189*, 934–937.
23. Hardy, J.D.; Chavez, C.M.; Kurrus, F.D.; Neely, W.A.; Eraslan, S.; Turner, M.D.; Fabian, L.W.; Labecki, T.D. Heart transplantation in man: developmental studies and report of a case. *J. Am. Med. Assoc.* **1964**, *188*, 1132–1140.
24. Giles, G.R.; Boehmig, H.J.; Amemiya, H.; Halgrimson, C.G.; Starzl, T.E. Clinical heterotransplantation of the liver. *Transplant. Proc.* **1970**, *2*, 506–512.
25. Bertoye, A. Essais de traitement de certaines infirmités hépatiques graves par greffe hépatique hétérologue. *Lyon Med.* **1969**, *222*, 345.
26. Marion, P. Les transplantations cardiaques et les transplantations hépatiques. *Lyon Med.* **1969**, *222*, 585.
27. Leger, L.; Chapuis, Y.; Lenriot, J.P. Heterotopic graft of baboon liver in a patient with fulminating hepatitis. *Chirurgie* **1970**, *96*, 249–251 (English translation).
28. Marion, P.; Bertoye, A.; Mikaeloff, P.; Bolot, J.F.; Berard, P.; Rassat, J.P. Treatment of hepatic coma by auxiliary heterologous transplantation. *Chirurgie* **1970**, *96*, 152–161 (English translation).
29. Dubernard, J.M.; Bonneau, M.; Latour, M. *Heterographs in Primates*; Fondation Merieux: Villeurbanne, 1974; 42.
30. Starzl, T.E.; Putnam, C.W.; Porter, K.A.; Picache, R.; Husberg, B.S.; Halgrimson, C.G.; Schroter, G. Progress in and deterrents to orthotopic liver transplantation, with special reference to survival, resistance to hyperacute rejection, and biliary duct reconstruction. *Transplant. Proc.* **1974**, *6*, 129–139.
31. Barnard, C.N.; Wolpowitz, A.; Losman, J.G. Heterotopic cardiac transplantation with a xenograft for assistance of the left heart in cardiogenic shock after cardiopulmonary bypass. *S. Afr. Med. J.* **1977**, *52*, 1035–1038.
32. Bailey, L.L.; Nehlsen-Carnnarella, S.L.; Concepcion, W.; Jolley, W.B. Baboon-to-human cardiac xenotransplantation in a neonate. *J. Am. Med. Assoc.* **1985**, *254*, 3321–3329.
33. Cascalho, M.; Platt, J.L. The immunological barrier to xenotransplantation. *Immunity* **2001**, *14*, 437–446.
34. Dorling, A.; Lombardi, G.; Binns, R.; Lechler, R.I. Detection of primary direct and indirect human antiporcine T cell responses using a porcine dendritic cell population. *Eur. J. Immunol.* **1996**, *26*, 1378–1387.
35. Platt, J.L. *Hyperacute Xenograft Rejection*; R.G. Landes: Austin, 1995.
36. Galili, U. Interaction of the natural anti-Gal antibody with α -galactosyl epitopes: a major obstacle for xenotransplantation in humans. *Immunol. Today* **1993**, *14*, 480–482.
37. Good, A.H.; Cooper, D.K.C.; Malcolm, A.J.; Ippolito, R.M.; Koren, E.; Neethling, F.A.; Ye, Y.; Zuhdi, N.; Lamontagne, L.R. Identification of carbohydrate structures that bind human antiporcine antibodies: implications for discordant xenografting in humans. *Transplant. Proc.* **1992**, *24*, 559–562.
38. Sandrin, M.S.; Vaughan, H.A.; Dabkowski, P.L.; McKenzie, I.F.C. Anti-pig IgM antibodies in human serum react predominantly with Gal α (1,3) Gal epitopes. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 11391–11395.
39. Platt, J.L.; Vercellotti, G.M.; Dalmaso, A.P.; Matas, A.J.; Bolman, R.M.; Najarian, J.S.; Bach, F.H. Transplantation of discordant xenografts: a review of progress. *Immunol. Today* **1990**, *11*, 450–456.
40. Dalmaso, A.P.; Vercellotti, G.M.; Platt, J.L.; Bach, F.H. Inhibition of complement-mediated endothelial cell cytotoxicity by decay accelerating factor: potential for prevention of xenograft hyperacute rejection. *Transplantation* **1991**, *52*, 530–533.
41. Cooper, D.K.C.; Human, P.A.; Lexer, G.; Rose, A.G.; Rees, J.; Keraan, M.; Du Toit, E. Effects of cyclosporine and antibody adsorption on pig cardiac xenograft survival in the baboon. *J. Heart Transplant.* **1988**, *7*, 238–246.
42. Lin, S.S.; Weidner, B.C.; Byrne, G.W.; Diamond, L.E.; Lawson, J.H.; Hoopes, C.W.; Daniels, L.J.; Daggett, C.W.; Parker, W.; Harland, R.C.; Davis, R.D.; Bollinger, R.R.; Logan, J.S.; Platt, J.L. The role of antibodies in acute vascular rejection of pig-to-baboon cardiac transplants. *J. Clin. Invest.* **1998**, *101*, 1745–1756.
43. Leventhal, J.R.; Dalmaso, A.P.; Cromwell, J.W.; Platt, J.L.; Manivel, C.J.; Bolman, R.M.; Matas, A.J. Prolongation of cardiac xenograft survival by depletion of complement. *Transplantation* **1993**, *55*, 857–866.
44. Pruitt, S.K.; Kirk, A.D.; Bollinger, R.R.; Marsh, H.C., Jr.; Collins, B.H.; Levin, J.L.; Mault, J.R.; Heinle, J.S.; Ibrahim, S.; Rudolph, A.R.; Baldwin, W.M., III; Sanfilippo, F. The effect of soluble complement receptor type 1 on hyperacute rejection of porcine xenografts. *Transplantation* **1994**, *57*, 363–370.
45. Candinas, D.; Lesnikoski, B.A.; Robson, S.C.; Miyatake, T.; Scesney, S.M.; March, H.C.J.; Ryan, U.S.; Dalmaso, A.P.; Hancock, W.W.; Bach, F.H. Effects of repetitive high-dose treatment with soluble complement receptor type 1 and cobra venom factor on discordant xenograft survival. *Transplantation* **1996**, *62*, 336–342.
46. Kroshus, T.J.; Rollins, S.A.; Dalmaso, A.P.; Elliott, E.A.; Matis, L.A.; Squinto, S.P.; Bolman, R.M. Complement inhibition with an anti-C5 monoclonal antibody prevents acute cardiac tissue injury in an ex vivo model of pig-to-human xenotransplantation. *Transplantation* **1995**, *60*, 1194–1202.
47. Cotterell, A.H.; Collins, B.H.; Parker, W.; Harland, R.C.; Platt, J.L. The humoral immune response in humans following cross-perfusion of porcine organs. *Transplantation* **1995**, *60*, 861–868.
48. Wang, H.; Rollins, S.A.; Gao, Z.; Garcia, B.; Zhang, Z.; Xing, J.; Li, L.; Kellersmann, R.; Matis, L.A.; Zhong, R. Complement inhibition with an anti-C5 monoclonal antibody prevents hyperacute rejection in a

- xenograft heart transplantation model. *Transplantation* **1999**, *70*, 697–698.
49. Fodor, W.L.; Williams, B.L.; Matis, L.A.; Madri, J.A.; Rollins, S.A.; Knight, J.W.; Velander, W.; Squinto, S.P. Expression of a functional human complement inhibitor in a transgenic pig as a model for the prevention of xenogeneic hyperacute organ rejection. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 11153–11157.
 50. McCurry, K.R.; Kooyman, D.L.; Diamond, L.E.; Byrne, G.W.; Logan, J.S.; Platt, J.L. Human complement regulatory proteins in transgenic animals regulate complement activation in xenoperfused organs. *Transplant. Proc.* **1995**, *27*, 317–318.
 51. Schmoeckel, M.; Bhatti, F.N.K.; Zaidi, A.; Cozzi, E.; Waterworth, P.D.; Tolan, M.J.; Goddard, M.; Warner, R.G.; Langford, G.A.; Dunning, J.J.; Wallwork, J.; White, D.J.G. Orthotopic heart transplantation in a transgenic pig-to-primate model. *Transplantation* **1998**, *65*, 1570–1577.
 52. Zaidi, A.; Bhatti, F.; Schmoeckel, M.; Cozzi, E.; Chavez, G.; Wallwork, J.; White, D.; Friend, P. Kidneys from HDFA transgenic pigs are physiologically compatible with primates. *Transplant. Proc.* **1998**, *30*, 2465–2466.
 53. Schmoeckel, M.; Nollert, G.; Shahmohammadi, M.; Young, V.K.; Chavez, G.; Kasper-Konig, W.; White, D.J.G.; Muller-Hocker, J.; Arendt, R.M.; Wilbert-Lampen, U.; Hammer, C.; Reichart, B. Prevention of hyperacute rejection by human decay accelerating factor in xenogeneic perfused working hearts. *Transplantation* **1996**, *62*, 729–734.
 54. Byrne, G.W.; McCurry, K.R.; Martin, M.J.; McClellan, S.M.; Platt, J.L.; Logan, J.S. Transgenic pigs expressing human CD59 and decay-accelerating factor produce an intrinsic barrier to complement-mediated damage. *Transplantation* **1997**, *63*, 149–155.
 55. Leventhal, J.R.; Matas, A.J.; Sun, L.H.; Reif, S.; Bolman, R.M.; Dalmaso, A.P.; Platt, J.L. The immunopathology of cardiac xenograft rejection in the guinea pig-to-rat model. *Transplantation* **1993**, *56*, 1–8.
 56. Lin, S.S.; Platt, J.L. Immunologic barriers to xenotransplantation. *J. Heart Lung Transplant.* **1996**, *15*, 547–555.
 57. Porter, K.A. Renal transplantation. In *Pathology of the Kidney*, 4th Ed.; Little, Brown and Company: Boston, 1992; Vol. 3, 1799–1933.
 58. Gaca, J.G.; Lee, W.; Aksoy, O.; Braedehoeft, S.J.; Gonzalez-Stawinski, G.V.; Parker, W.; Davis, R.D. Evidence for polyreactive xenoreactive antibodies in the repertoire of human anti-swine antibodies: the ‘next’ humoral barrier to xenotransplantation. *Transplant. Immunol.* **2001**, *9*, 19–27.
 59. Blakely, M.L.; Van Der Werf, W.J.; Berndt, M.C.; Dalmaso, A.P.; Bach, F.H.; Hancock, W.W. Activation of intragraft endothelial and mononuclear cells during discordant xenograft rejection. *Transplantation* **1994**, *58*, 1059–1066.
 60. Malyguine, A.M.; Saadi, S.; Holzknicht, R.A.; Patte, C.R.; Sud, N.; Platt, J.L.; Dawson, J.R. Induction of procoagulant function in porcine endothelial cells by human NK cells. *J. Immunol.* **1997**, *159*, 4659–4664.
 61. Platt, J.L.; Lin, S.S.; McGregor, C.G.A. Acute vascular rejection. *Xenotransplantation* **1998**, *5*, 169–175.
 62. Bracy, J.L.; Sachs, D.H.; Iacomini, J. Inhibition of xenoreactive natural antibody production by retroviral gene therapy. *Science* **1998**, *281*, 1845–1847.
 63. Chopek, M.W.; Simmons, R.L.; Platt, J.L. ABO-incompatible renal transplantation: initial immunopathologic evaluation. *Transplant. Proc.* **1987**, *19*, 4553–4557.
 64. Bach, F.H.; Ferran, C.; Hechenleitner, P.; Mark, W.; Koyamada, N.; Miyatake, T.; Winkler, H.; Badrichani, A.; Cardinas, D.; Hancock, W.H. Accommodation of vascularized xenografts: expression of “protective genes” by donor endothelial cells in a host Th2 cytokine environment. *Nat. Med.* **1997**, *3*, 196–204.
 65. Bach, F.H.; Ferran, C.; Soares, M.; Wrighton, C.J.; Anrather, J.; Winkler, H.; Robson, S.C.; Hancock, W.W. Modification of vascular responses in xenotransplantation: inflammation and apoptosis. *Nat. Med.* **1997**, *3*, 944–948.
 66. Ferran, C.; Badrichani, A.Z.; Cooper, J.T.; Stroka, D.M.; Bach, F.H. Xenotransplantation: progress toward clinical development. *Adv. Nephrol. Necker Hosp.* **1997**, *27*, 391–420.
 67. Ferran, C.; Stroka, D.M.; Badrichani, A.Z.; Cooper, Z.T.; Bach, F.H. Adenovirus-mediated gene transfer of A20 renders endothelial cells resistant to activation: a means of evaluating the role of endothelial cell activation in xenograft rejection. *Transplant. Proc.* **1997**, *29*, 879–880.
 68. Von Albertini, M.; Ferran, C.; Brostjan, C.; Bach, F.H.; Goodman, D.J. Membrane-associated lymphotoxin on natural killer cells activates endothelial cells via an NF-kappaB-dependent pathway. *Transplantation* **1998**, *66*, 1211–1219.
 69. Murray, A.G.; Khodadoust, M.M.; Pober, J.S.; Bothwell, A.L.M. Porcine aortic endothelial cells activate human T cells: direct presentation of MHC antigens and costimulation by ligands for human CD2 and CD28. *Immunity* **1994**, *1*, 57–63.
 70. Dorling, A.; Stocker, C.; Tsao, T.; Haskard, D.O.; Lechler, R.I. In vitro accommodation of immortalized porcine endothelial cells: resistance to complement mediated lysis and down-regulation of VCAM expression induced by low concentrations of polyclonal human IgG anti-pig antibodies. *Transplantation* **1996**, *62*, 1127–1136.
 71. Kirk, A.D.; Li, R.A.; Kinch, M.S.; Abernethy, K.A.; Doyle, C.; Bollinger, R.R. The human antiporcine cellular repertoire. In vitro studies of acquired and innate cellular responsiveness. *Transplantation* **1993**, *55*, 924–931.
 72. Waldmann, H. Therapeutic approaches for transplantation. *Curr. Opin. Immunol.* **2001**, *13*, 606–610.
 73. Logan, J.S.; Martin, M.J. Transgenic swine as a recombinant production system for human hemoglobin. *Methods Enzymol.* **1994**, *231*, 435–445.
 74. Prather, R.S.; Hawley, R.J.; Carter, D.B.; Lai, L.; Greenstein, J.L. Transgenic swine for biomedicine and agriculture. *Theriogenology* **2003**, *59*, 115–125.
 75. Wilmut, I.; Schnieke, A.E.; McWhir, J.; Kind, A.J.; Campbell, K.H. Viable offspring derived from

- fetal and adult mammalian cells. *Nature* **1997**, *385*, 810–813.
76. Kato, Y.; Tani, T.; Sotomaru, Y.; Kurokawa, K.; Kato, J.; Doguchi, H. Eight calves cloned from somatic cells of a single adult. *Science* **1998**, *282*, 2095–2098.
77. Wakayama, T.; Perry, A.C.; Zuccotti, M.; Johnson, K.R.; Yanagimachi, R. Full-term development of mice from enucleated oocytes injected with cumulus cell nuclei. *Nature* **1998**, *394*, 315–316.
78. Phelps, C.J.; Koike, C.; Vaught, T.D.; Boone, J.; Wells, K.D.; Chen, S.H.; Ball, S.; Specht, S.M.; Polejaeva, I.A.; Monahan, J.A.; Jobst, P.M.; Sharma, S.B.; Lamborn, A.E.; Garst, A.S.; Moore, M.; Demetris, A.J.; Rudert, W.A.; Bottino, R.; Bertera, S.; Trucco, M.; Starzl, T.E.; Dai, Y.; Ayares, D.L. Production of alpha 1,3-galactosyltransferase-deficient pigs. *Science* **2003**, *299*, 411–414.
79. Dai, Y.; Vaught, T.D.; Boone, J.; Chen, S.; Phelps, C.J.; Ball, S.; Monahan, J.A.; Jobst, P.M.; McCreath, K.J.; Lamborn, A.E.; Cowell-Lucero, J.L.; Wells, K.D.; Colman, A.; Polejaeva, I.A.; Ayares, D.L. Targeted disruption of the alpha1,3-galactosyltransferase gene in cloned pigs. *Nat. Biotechnol.* **2002**, *20*, 251–255.
80. Lai, L.; Kolber-Simonds, D.; Park, K.W.; Cheong, H.T.; Greenstein, J.L.; Im, G.S.; Samuel, M.; Bonk, A.; Rieke, A.; Day, B.N.; Murphy, C.N.; Carter, D.B.; Hawley, R.J.; Prather, R.S. Production of α -1,3-galactosyltransferase knockout pigs by nuclear transfer cloning. *Science* **2002**, *295*, 1089–1092.
81. Lai, L.; Prather, R.S. Progress in producing knockout models for xenotransplantation by nuclear transfer. *Ann. Med.* **2002**, *34*, 501–506.
82. Onishi, A.; Iwamoto, M.; Akita, T.; Mikawa, S.; Takeda, K.; Awata, T.; Hanada, H.; Perry, A.C.F. Pig cloning by microinjection of fetal fibroblast nuclei. *Science* **2000**, *289*, 1188–1190.
83. Polejaeva, I.A.; Chen, S.; Vaught, T.D.; Page, R.L.; Mullins, J.; Ball, S.; Dai, Y.; Boone, J.; Walker, S.; Ayares, D.L.; Colman, A.; Campbell, K.H.S. Cloned piglets produced by nuclear transfer from adult somatic cells. *Nature* **2000**, *407*, 86–90.
84. Betthausen, J.; Forsberg, E.; Augenstein, M.; Childs, L.; Eilertsen, K.; Enos, J.; Forsythe, T.; Golueke, P.; Jurgella, G.; Koppang, R.; Lesmeister, T.; Mallon, K.; Mell, G.; Misica, P.; Pace, M.; Pfister-Genskow, M.; Strelchenko, N.; Voelker, G.; Watt, S.; Thompson, S.; Bishop, M. Production of cloned pigs from in vitro systems. *Nat. Biotechnol.* **2000**, *18*, 1055–1059.
85. Sharma, A.; Naziruddin, B.; Cui, C.; Martin, M.J.; Xu, H.; Wan, H.; Lei, Y.; Harrison, C.; Yin, J.; Okabe, J.; Mathews, C.; Stark, A.; Adams, C.S.; Houtz, J.; Wiseman, B.S.; Byrne, G.W.; Logan, J.S. Pig cells that lack the gene for alpha1-3 galactosyltransferase express low levels of the gal antigen. *Transplantation* **2003**, *75*, 430–436.
86. Nottle, M.B.; D'Apice, A.J.; Cowan, P.J.; Boquest, A.C.; Harrison, S.J.; Grupen, C.G. Transgenic perspectives in xenotransplantation, 2001. *Xenotransplantation* **2002**, *9*, 305–308.
87. Brunelli, J.P.; Pall, M.L. Lambda/plasmid vector construction by in vivo cre/lox-mediated recombination. *BioTechniques* **1994**, *16*, 1060–1064.
88. Sandrin, M.S.; Fodor, W.L.; Mouhtouris, E.; Osman, N.; Cohny, S.; Rollins, S.A.; Guilmette, E.R.; Setter, E.; Squinto, S.P.; McKenzie, I. Enzymatic remodeling of the carbohydrate surface of a xenogeneic cell substantially reduces human antibody binding and complement-mediated cytolysis. *Nat. Med.* **1995**, *1*, 1261–1267.
89. Sharma, A.; Okabe, J.F.; Birch, P.; Platt, J.L.; Logan, J.S. Reduction in the level of Gal (α 1,3) Gal in transgenic mice and pigs by the expression of an α (1,2) fucosyltransferase. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 7190–7195.
90. Crick, S.; Sheppard, M.N.; Ho, S.Y.; Gebstein, L.; Anderson, R.H. Anatomy of the pig heart: comparisons with normal human cardiac structure. *J. Anat.* **1998**, *193*, 105–119.
91. Crick, S.J.; Sheppard, M.N.; Ho, S.Y.; Anderson, R.H. Localisation and quantitation of autonomic innervation in the porcine heart I: conduction system. *J. Anat.* **1999**, *195*, 341–357.

Transparent Ceramics: Bioactive Materials and Tissue Engineering

Motohiro Hirose

Koji Ioku

Noriko Kotobuki

Tissue Engineering Research Group, Research Institute for Cell Engineering (RICE), National Institute of Advanced Industrial Science and Technology (AIST), Amagasaki, Hyogo, Japan

Daisuke Kawagoe

Hajime Ohgushi

Graduate School of Environmental Studies, Tohoku University, Aoba-ku, Sendai, Japan

INTRODUCTION

Tissue engineering is an interdisciplinary field that involves cell biology, materials science, and clinical research for creating functional tissues and organs. Significant advances in tissue engineering have been led by effectively combining living cells including stem cells with materials suitable for cell cultivation. Currently, many types of materials are being investigated for reconstructive bone surgery. For example, calcium phosphate ceramics such as hydroxyapatite (HA) and β -tricalcium phosphate (β -TCP) are widely used as artificial bone graft substitutes.^[1-4] When implanted into bone defects, these materials exhibit bone-bonding properties that enable them to establish chemical bonds between the host bones and the surfaces of the implants. This bone/material interface is very firm and stable, and these materials are known as “bioactive materials.”^[5,6]

Another current attractive technology has been developed for regenerating viable tissues or organs by combining cultured cells with suitable materials. Hydroxyapatite and β -TCP have high degrees of biocompatibility/bioaffinity and stimulate osteoconduction. Thus, these materials are well-known scaffold materials and play important roles in the adhesion, proliferation, and differentiation of cultured cells. We have previously demonstrated the fabrication of regenerative cultured bone tissues with cultured cells by means of tissue engineering.^[7-10] We culture expanded the number of mesenchymal stem cells (MSCs) from fresh bone marrow cells, then further cultured the cells on the HA ceramic and β -TCP ceramic samples.^[11,12] These MSCs differentiated into osteoblasts and produced bone matrices under osteogenic conditions in vitro.^[13] Moreover, the cultured osteoblast/matrix constructs (regenerative cultured bone) show immediate new bone formation capability after in vivo

implantation.^[10,14,15] Based on these results, since January 2002, we have applied the cultured bone tissues fabricated at the cell processing center of our institute to patients suffering from cases such as bone tumors (Fig. 1). We considered that the ceramic surface is very promising for supporting MSC proliferation/differentiation, which has been demonstrated by biochemical or gene expression analyses and observation by scanning electron microscopy (SEM).^[16,17] Usually, cultured cells can be observed with a light microscope when they are cultured on tissue culture polystyrene (TCPS) dishes. However, owing to the opacity of the ceramics, microscopic observation of the cells on ceramics is very difficult. Scanning electron microscopy requires cell fixation by a fixative solution, and the cells are no longer viable. Thus, consecutive observation of the same specimen during the culture periods was impossible.

If the ceramics were transparent, we would be able to observe cellular events including cell adhesion, proliferation, and differentiation cascade with an ordinary microscope. To confirm such cellular events, we attempted to fabricate novel materials for cell culture, specifically, transparent hydroxyapatite (t-HA) ceramics and transparent β -TCP (t- β -TCP) ceramics, which were made by the spark plasma sintering process (SPS).^[18] The SPS is a new process that enables sintering of materials in short periods by charging the spaces between the powder particles with electrical energy and using a high sintering pressure. The SPS systems have many advantages such as rapid sintering, sintering with fewer additives, uniform sintering, low operating expense, and easy operation compared to that of conventional systems using hot press (HP) sintering, hot isostatic pressing (HIP), or atmospheric furnaces. The SPS process can be applied to many materials including fine ceramics. Both the t-HA ceramics and the t- β -TCP ceramics made by the SPS

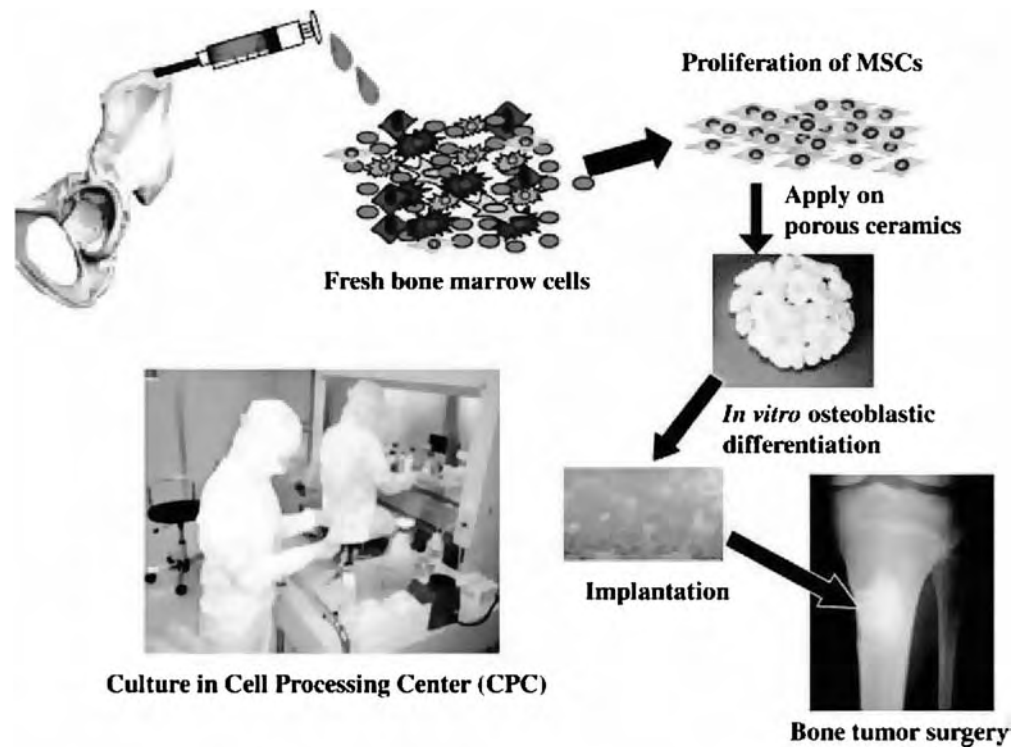


Fig. 1 Schematic representation of clinical application of MSCs to bone tumor cases. Mesenchymal stem cells from fresh bone marrow cells were allowed to proliferate and were then seeded on porous bioactive ceramics. To fabricate regenerative cultured bone, the cell/ceramic composites were cultured under an osteogenic medium. The cultured bone was used for bone tumor surgery in a hospital. The cells were cultured at the cell processing center (CPC) of our institute.

process have enabled observation of the cell behaviors of rat MSCs.^[19,20]

In this entry, we used rat MSCs derived from bone marrow as the cell source and observed cell adhesion, proliferation, and osteogenic differentiation on the t-HA ceramics as well as the t- β TCP ceramics, and conducted a comparative analysis of cell functions.

EXPERIMENTAL PROCEDURES

Preparation of Ceramics

The detailed methods of making t-HA ceramics and t- β -TCP ceramics have been previously reported.^[18] Briefly, a fine powder of HA or β -TCP was utilized as the basic material. One gram of this powder was poured into a graphite mold (inner diameter: 15 mm), then sintered by the SPS process (SPS: Dr Sinter-511S, Sumitomo Coal Mining, Tokyo, Japan). The samples were pressed uniaxially under 10 MPa, then heated at 800°C, 900°C, or 1000°C for 10 min at a rate of 25°C/min. Each ceramic sample was 5 or 20 mm in diameter and 2 mm thick. The ceramic samples were autoclaved for sterilization at 120°C for 20 min.

Surface Characterization

The sessile contact angles (SCA) of the t-HA ceramics, t- β -TCP ceramics, and TCPS dishes were determined using Milli-Q water and a goniometer (Face Contact-Angle Meter, Kyowa Kaimenkagaku Co., Ltd., Tokyo, Japan). A probability (p) value of less than 0.05 was considered significant. To evaluate the surface structures of the materials, a scanning electron microscope (SM-300, Topcon Corp., Tokyo, Japan) was used to analyze the surface of each material.

Cell Culture

Rat bone marrow cells were obtained from 7-week-old male Fischer 344 (F344) rats. Preparation and osteogenic differentiation of MSCs from rat were described by Maniatopoulos et al. and modified by us.^[7,21]

In brief, rat bone marrow cells were flushed out by a culture medium, minimum essential medium (Nacalai Tesque, Inc., Kyoto, Japan) containing 15% fetal bovine serum (JRH Biosciences, Inc., Kansas, U.S.A.), and 1% antibiotics. The bone marrow cells were cultured on T-75 TCPS dishes (Becton, Dickinson and Company, New Jersey, U.S.A.) in a humidified

atmosphere of 95% air with 5% CO₂ at 37°C. The cells adherent to the dishes containing MSCs from rat bone marrow were initially cultured up to 80% confluence and resuspended to 5×10^5 cells/mL in culture medium following harvesting using 0.05% trypsin/0.53 mM EDTA. The cell suspension was applied to the sterilized t-HA ceramic and t-β-TCP ceramic samples (5 or 20 mm in diameter and 2 mm thick), which were placed into a multiwell plate. The cells were cultured with osteogenic medium containing 10 nM dexamethasone (Dex) (Sigma-Aldrich Corporation, Missouri, U.S.A.), 10 mM β-glycerophosphate (Merck, Darmstadt, Germany), and 0.28 mM ascorbic acid 2-phosphate magnesium salt *n*-hydrate (Sigma-Aldrich Corp.) As a control, the same amount of cell suspension was also applied directly in a TCPS well with no ceramic material. The culture medium was changed two or three times a week. During the culture periods, the cell morphologies at different days of cell adhesion, proliferation, and differentiation were detected by phase-contrast (Olympus CK41SF, Olympus Optical, Tokyo, Japan) and fluorescence microscopy (Olympus IX70FL).

Alkaline Phosphatase Activity Staining

The cultured cells on the t-HA ceramics, t-β-TCP ceramics, and TCPS dishes were rinsed with phosphate buffer saline (PBS) and fixed with 4% paraformaldehyde for 10 min at 4°C. The fixed cells were soaked in 0.1% naphthol AS-MX phosphate and 0.1% fast red violet LB salt in 56 mM 2-amino-2-methyl-1, 3-propanediol (pH 9.9) for 10 min at room temperature and washed with PBS.^[7]

Calcium Staining by Alizarin Red S

The cultured cells on the t-HA ceramics, t-β-TCP ceramics, and TCPS dishes were washed with PBS and fixed with 4% paraformaldehyde for 10 min at 4°C. The fixed cells were soaked in 0.5% Alizarin red S/PBS solution for 10 min at room temperature and washed with PBS.^[7] Nonspecific reactions were removed by a washing step using 70% ethanol (EtOH). The stain was used to detect calcium in the extracellular matrix.

Visualization of Mineralized Matrix by Fluorescence Emission During Subculture

During the culture periods, the calcium in the mineralized extracellular matrix was also determined by using calcein, which is a calcium-binding fluorescent dye.^[22] Calcein (1 μg/mL) (Dojindo Laboratories, Kumamoto, Japan) was added to the culture wells

whenever the culture medium was replaced. Before the assay, the medium containing calcein was removed and the wells were washed twice with PBS. The fluorescence of the incorporated calcein was observed by using a fluorescence microscope. After this assay, the cell culture was continued in the osteogenic condition with calcein.

RESULTS

Surface Characterization of t-HA Ceramics, t-β-TCP Ceramics, and TCPS Dishes

It is well known that the surface roughness and wettability of culture substrates influence cellular activity such as cell adhesion and cell proliferation. Therefore, we studied the surface properties of the t-HA ceramic and t-β-TCP ceramic by SEM before cultivation. The surfaces of the t-HA ceramic and t-β-TCP ceramic samples were slightly rough compared with the surface of the TCPS dishes. Nevertheless, the grain structure of the surfaces of the sintered t-HA ceramic and t-β-TCP ceramic samples was fine (less than 1 μm) (data not shown). We also checked the wettability of the t-HA ceramics, t-β-TCP ceramics, and TCPS dishes, which is represented by SCA (Table 1). The SCA of the t-HA ceramics, t-β-TCP ceramics, and TCPS dishes showed no significant difference. The results indicate that surface configuration is not an important factor in comparing the cellular responses on the t-HA ceramics and t-β-TCP ceramics to that on the TCPS dishes.

Cell Adhesion and Proliferation of Rat MSCs Seeded on t-HA Ceramics, t-β-TCP Ceramics, and TCPS Dishes

Next, we observed the cell adhesion and proliferation of rat MSCs on t-HA ceramics, t-β-TCP ceramics, and TCPS dishes. The TCPS dish, which is fabricated from crystal grade polystyrene by a vacuum-gas plasma treatment, exhibits a slightly hydrophilic nature because the plasma treatment introduces carboxyl and hydroxyl groups onto the surface. Under typical

Table 1 Sessile contact angle (SCA)

Material	SCA ± SD (°)
t-HA	73.1 ± 2.36
t-β-TCP	74.7 ± 3.92
TCPS	71.0 ± 3.42

Data show the mean value of eight samples with standard deviation (SD).

(From Hirose, M., et al.)

culture conditions employing TCPS dishes, numerous types of cultured cells are able to attach and proliferate on the surface, while cells could scarcely attach to bare polystyrene because of the high degree of hydrophobicity of its surface. Therefore, TCPS is known as a gold standard substrate for cell cultivation.

Because t-HA ceramics, t- β -TCP ceramics, and TCPS dishes are transparent, cells cultured on these substrates can be observed by phase-contrast and fluorescence microscopy. The shapes of the cultured cells on both ceramics were clearly detected by phase-contrast microscopy; the resolution of the microscopy was almost the same as that on the TCPS dishes (Fig. 2). Even at a very early stage of culture (1 hr after cell seeding), many MSCs had already attached and exhibited initial spreading as evidenced by small, round cells on the t-HA ceramics and t- β -TCP ceramics (data not shown). One day after seeding, almost all the cells were able to adhere to the substrates and exhibit the morphological characteristics of mesenchymal types (spindle cell morphology) (Figs. 2A–C). At culture day 4, the cells were clearly seen proliferating well on all the substrates (Figs. 2D–F). The cascade of the cellular behaviors on the t-HA ceramics and t- β -TCP ceramics was similar to that of the cells seeded on the TCPS dishes (Fig. 2). The results suggested that rat MSCs could attach, spread, and proliferate on t-HA ceramics and t- β -TCP ceramics to the same degree as on TCPS dishes.

Osteogenic Differentiation of Rat MSCs Cultured on t-HA, t- β -TCP, and TCPS Dishes

We investigated the osteogenic differentiation of rat MSCs cultured on these substrates. To evaluate the osteogenic differentiation of rat MSCs on t-HA ceramics and t- β -TCP ceramics, the cells were cultured on both types of ceramics and on TCPS dishes in the presence or absence of Dex, which functions as an osteogenic factor, for 14 days. Cell morphology was observed by fluorescence and phase-contrast microscopy. The cells were cultured in the presence of β -glycerophosphate, ascorbic acid, and Dex [Dex (+)]. The cells differentiated into bone forming osteoblasts, which fabricate extracellular mineralized matrix nodules around the cultured cells.

Recently, we have reported a novel in situ method of monitoring for mineralization by cultured cells by using a calcium-binding fluorescent dye, calcein.^[7,8] Calcein was found to be specifically incorporated and deposited into extracellular bone matrices as evidenced by costaining with Alizarin red S. The advanced characteristics of the method included the monitoring of the mineralization of the same specimens of cultured cells in a time-dependent manner owing to continuous cultivation without the fixation processes of cell layers. To evaluate the degree of mineralization by rat MSCs cultured on the substrates, we added calcein to the culture medium and observed the cells with a

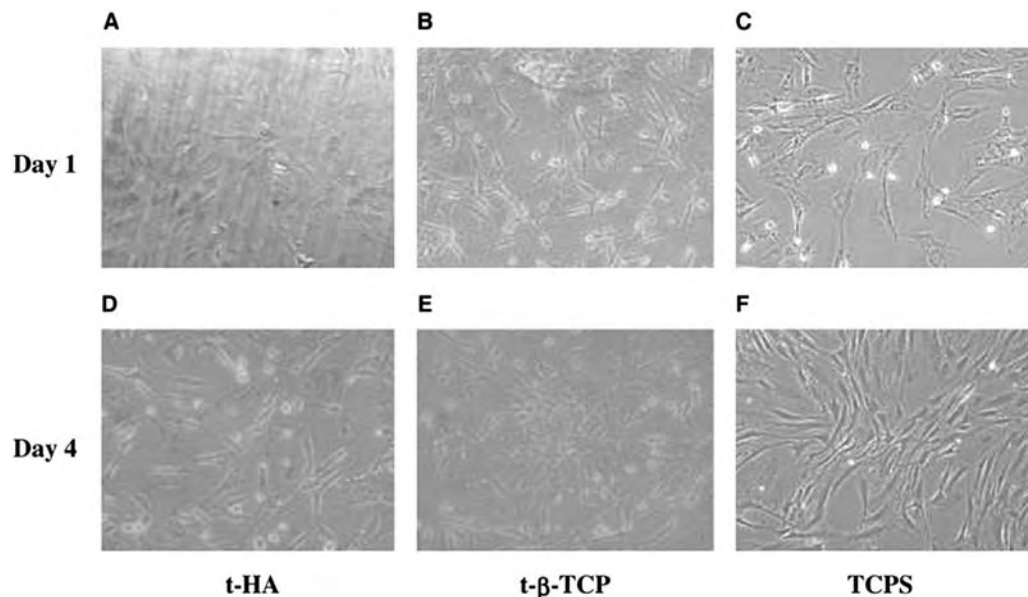


Fig. 2 Cell shape of rat MSCs on t-HA ceramics, t- β -TCP ceramics, and TCPS dishes. The cells cultured on t-HA ceramics, t- β -TCP ceramics, and TCPS dishes are shown in (A, D), (B, E), and (C, F), respectively. Cell shapes were observed by phase-contrast microscopy at culture day 1 (A–C) and day 4 (D–F) after seeding of rat MSCs on each substrate. At culture day 1, the cells had spread completely on all substrates. After 4 days, the cells had proliferated well on all substrates. Original magnification is 100 \times .

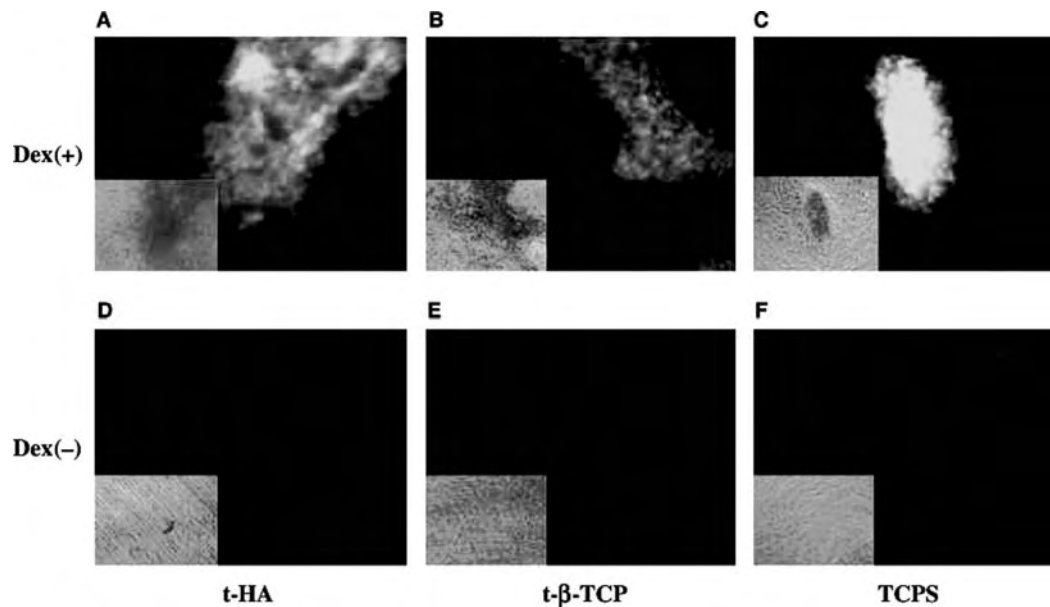


Fig. 3 Calcein uptake by rat MSCs on t-HA ceramics, t- β -TCP ceramics, and TCPS dishes. The cells cultured on t-HA ceramics, t- β -TCP ceramics, and TCPS dishes are shown in (A, D), (B, E), and (C, F), respectively. On all substrates, the cells cultured for 14 days in the presence of Dex exhibited extracellular mineralization as indicated by a green fluorescence emission of calcein (A–C). The cells cultured without Dex did not exhibit extracellular mineralization and no green fluorescence emission was observed (D–F). Each inset is a phase-contrast microscopic image of the corresponding field. Original magnification is 100 \times .

fluorescence microscope. When cultured in the osteogenic condition for 14 days, the cells exhibited the same level of mineralized bone matrix on the t-HA ceramics, t- β -TCP ceramics, and TCPS dishes (Figs. 3A–C). Fluorescence microscopy demonstrated that the cells treated with Dex [Dex (+)] could differentiate into osteoblasts and had the capability of mineralization as evidenced by calcein uptake on all of the substrates (Figs. 3A–C). Even after a profuse matrix formation on the t-HA ceramics and t- β -TCP ceramics, the cell shapes could easily be recognized under observation by phase-contrast microscopy (insets of Figs. 3A–C). In contrast, the cells cultured in the absence of Dex [Dex (–)] showed no osteoblastic cell shape but did show a fibroblastic shape (insets of Figs. 3D–F). That the calcein emission was negligible demonstrated no evidence of matrix formation by the cells without Dex (Figs. 3D–F). The findings indicate that rat MSC can easily differentiate into osteoblasts on the surface of t-HA ceramics and t- β -TCP ceramics resulting in the formation of bone matrix under conditions of osteogenic culture. Importantly, the cascade of the differentiation of rat MSCs cultured on t-HA ceramics and t- β -TCP ceramics can be equally detected on TCPS dishes.

To confirm the osteogenic differentiation of rat MSCs on t-HA ceramics and t- β -TCP ceramics, we stained the cells with ALP and calcium staining (Alizarin red S) after 14 days of cultivation by

macroscopic observation. Alkaline phosphatase (ALP), which is a cell surface protein, is known as an early marker for osteoblastic differentiation, and Alizarin red S can identify calcium, which is the major inorganic component of bone matrices. When rat MSCs were cultured with Dex, many clusters of cells were clearly stained with ALP on the t-HA ceramics, t- β -TCP ceramics, and TCPS dishes at almost the same level (Figs. 4A–C). In contrast, the cells cultured without Dex were faintly stained with ALP on all the substrates (Figs. 4D–F). Alizarin red S staining is commonly used to determine calcium deposited into bone matrices. As seen in the ALP staining (Fig. 4), the cells cultured with Dex for 14 days were strongly stained with Alizarin red S (Figs. 5A–C), while the cells cultured without Dex were scarcely stained (Figs. 5D–F). After 14 days of culture, the DNA contents from each substrate were measured. The DNA contents were similar regardless of the culture substrate (data not shown). The biochemical data showed that osteogenic differentiation of rat MSCs cultured in the presence of Dex could occur on t-HA ceramics and t- β -TCP ceramics as well as on TCPS dishes.

DISCUSSION

In this entry, cell behavior including cell adhesion, cell proliferation, and cell differentiation on t-HA ceramics

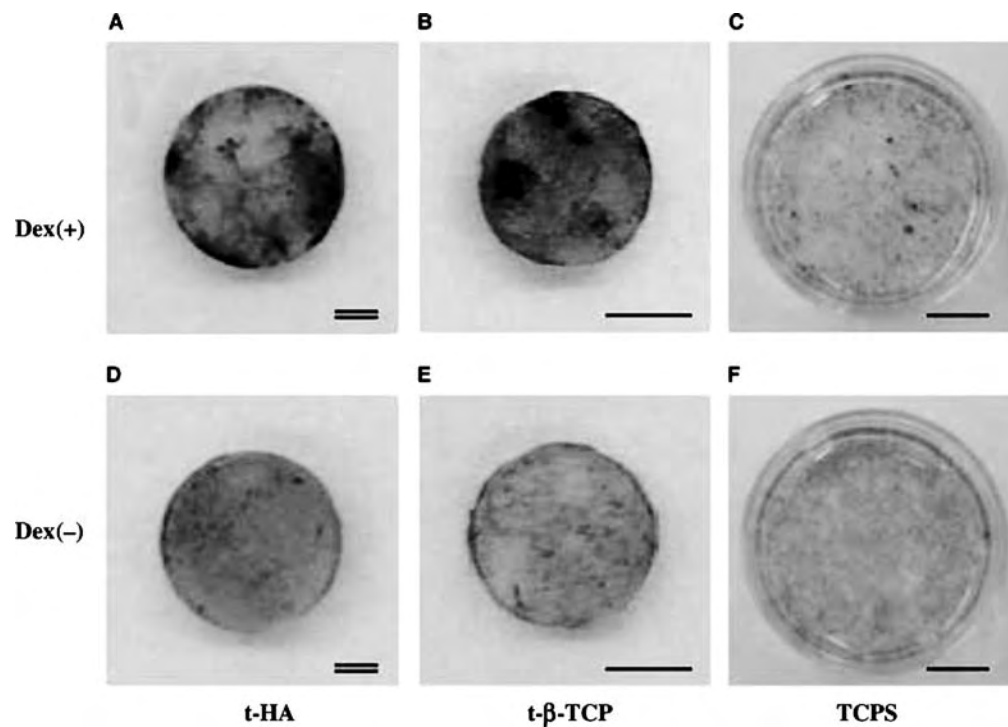


Fig. 4 Alkaline phosphatase activity staining of rat MSCs on t-HA ceramics, t- β -TCP ceramics, and TCPS dishes. The cells cultured on t-HA ceramics, t- β -TCP ceramics, and TCPS dishes are shown in (A, D), (B, E), and (C, F), respectively. On all substrates, the cells cultured for 14 days in the presence of Dex were actively stained with ALP, indicated in red (A–C). The cells cultured without Dex were faintly stained with ALP (D–F). Single bar and double bars indicate 10 and 1 mm, respectively.

and t- β -TCP ceramics was compared with that on TCPS dishes. Tissue culture polystyrene dishes are considered the gold standard substrate for cell cultivation and were therefore used in the present experiment as a positive control. In the progress of tissue engineering, extreme attention has been focused on the development of new scaffold materials suitable for cell cultivation. If cultured cells, especially stem cells, retain their biological activity within the material, they are expected to have the capability of self-renewal and differentiation into lineage-committed cells after *in vivo* implantation. Therefore, the combination of cultured cells and biomaterials is vital for the successful fabrication of regenerative cultured tissues or organs. Numerous types of materials have been used as artificial bone substitutes. Such substitutes need a high degree of biocompatibility and sufficient mechanical strength. Some kinds of ceramics have frequently been used as artificial bone substitutes and have also been used for tissue engineering as a cell scaffold resulting in good bone formation.^[23–26] Actually, for hard tissue regeneration, we have employed a variety of ceramics such as bioactive β -TCP ceramics, glass ceramics, and bioinert alumina ceramics as well as bioactive HA ceramics.^[9,11,27] Maniatopoulos et al. reported that cultured rat bone marrow cells could form mineralization *in vitro* in the presence of Dex.^[21] We then

reported that rat MSCs treated with Dex showed high levels of ALP activity and demonstrated osteocalcin mRNA expression by northern blot analysis. The localization of the gene expression was identified in the cells surrounding the mineralized area because of *in situ* hybridization.^[7] We also showed that the osteogenic differentiation of rat MSCs could occur on glass and HA ceramic surfaces.^[9,28] Furthermore, we have demonstrated that human MSCs derived from bone marrow have been found to have osteogenic capability both *in vitro* and *in vivo*.^[29–32] These previous results confirm that rat and human MSCs derived from bone marrow can differentiate into active osteoblasts, which produce bone matrices in culture.

In these situations, observation of a cascade of cell behaviors might be an attractive field of research for clarifying cell/material interactions. Usually, cell behaviors are observed using a phase-contrast microscope. Unfortunately, on many types of ceramics, cells are difficult to observe because of the opaque nature of the ceramics. We previously utilized single-crystalline transparent alumina ceramics to observe cell behaviors such as cell adhesion, cell proliferation, and osteogenic potential using human MSCs.^[29,30] However, most other types of ceramics such as HA and β -TCP are not naturally transparent and it is difficult to create single-crystal or transparent ceramics. Alumina ceramics are

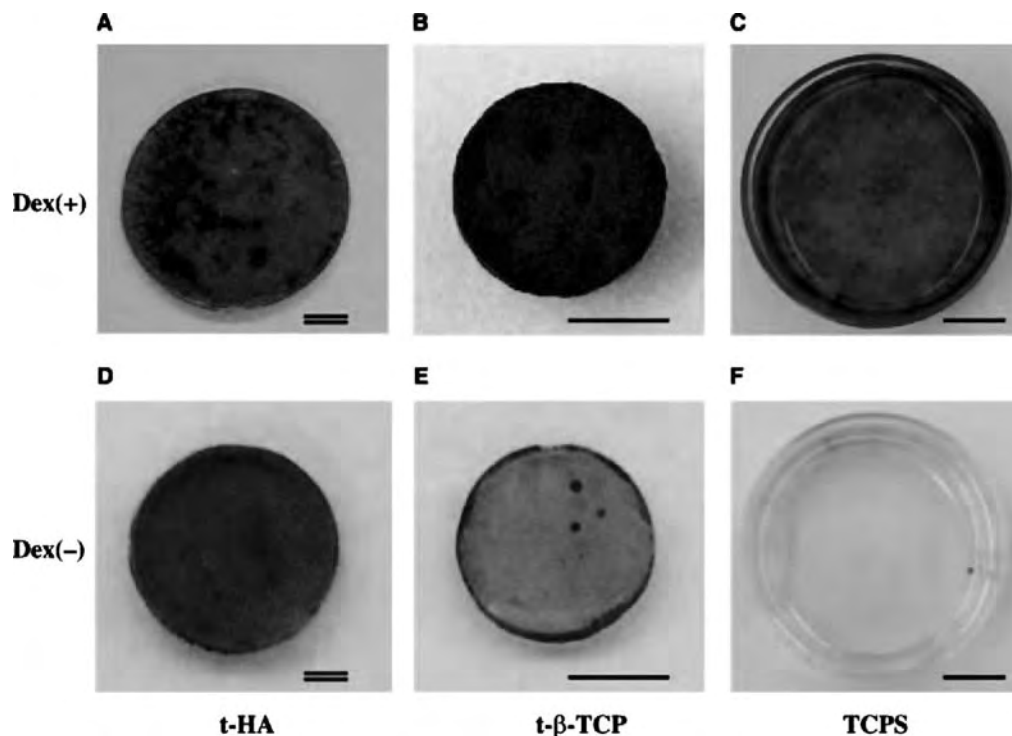


Fig. 5 Calcium staining of rat MSCs on t-HA ceramics, t- β -TCP ceramics, and TCPS dishes with Alizarin red S. The cells cultured on t-HA ceramics, t- β -TCP ceramics, and TCPS dishes are shown in (A, D), (B, E), and (C, F), respectively. On all substrates, the cells cultured for 14 days in the presence of Dex exhibited profuse extracellular mineralized matrices heavily stained with Alizarin red S (A–C). The cells cultured without Dex exhibited no obvious extracellular mineralized matrices (D–F). Single bar and double bars indicate 10 and 1 mm, respectively.

bioinert and have no bone-bonding properties. In contrast, HA ceramics and β -TCP ceramics are widely known as bioactive ceramics having bone-bonding properties and are suitable as scaffolds for culturing cells. However, owing to the opaque nature of the ceramics, cell behavior on the ceramics is difficult to observe. To overcome this shortcoming, the SPS process has recently been used for making certain types of transparent ceramics.

It is also well known that the surface characteristics of culture substrates affect cell behavior during culture periods. An appropriate level of surface hydrophilicity is needed for good cell adhesion. As seen in Table 1, the surface wettability of t-HA ceramics and t- β -TCP ceramics is similar to that of TCPS dishes. The optimal wettability for cell attachment is considered to be about 70° (defined by SCA), and the SCA of t-HA ceramics and t- β -TCP ceramics as well as TCPS dishes is nearly 70° . Additionally, the surface of t-HA and t- β -TCP ceramics is very smooth in comparison to that of TCPS dishes as observed by SEM (data not shown). These results indicate that the wettability of t-HA and t- β -TCP is suitable for cell adhesion.

As shown in Fig. 2, both t-HA ceramics and t- β -TCP ceramics exhibited cellular adhesive and proliferative characteristics comparable to those of TCPS

dishes. By using transparent materials, we can use a microscope to monitor the cellular event cascade of the same samples throughout the culture period. Cell adhesion molecules could be considered the reason the ceramics showed high cellular affinity. Although further experiments are required to clarify the mechanisms of the excellent properties of t-HA ceramics and t- β -TCP ceramics, fibronectin (FN) in serum or produced by the cells is thought to be a principal factor for promoting cellular adhesion and spreading on the substrates. In comparison with polystyrene treated for tissue culture (TCPS), the interaction of FN with bare polystyrene (as is commonly used for bacterial agar plates) is weak and about 1/10 as effective in cell binding. Therefore, a TCPS dish is an excellent substrate for cell culture and was therefore used as a positive control in this study. The degree of cell adhesion and cell proliferation on t-HA ceramics and t- β -TCP ceramics was similar to that on TCPS dishes (Fig. 2). The results suggest that the surface properties of t-HA ceramics and t- β -TCP ceramics are equivalent to those of TCPS dishes, which results in the capability of supporting cellular adhesion and proliferation.

The *in vitro* osteogenic differentiation on various types of ceramics can be evaluated by ALP staining or calcium staining with Alizarin red S after fixation, as

shown in Figs. 4 and 5. However, we have not been able to repeatedly monitor the same specimen during the culture period because of cell death caused by fixative solutions. Now, we have developed a new method for clearly monitoring the osteogenic differentiation indicated by fluorescence emission of calcein added to the culture medium (Fig. 3). The fluorescence can be repeatedly detected for the same specimen during the entire culture period. As seen in Fig. 3, we confirmed the osteogenic capability of rat MSCs cultured on t-HA ceramics and t- β -TCP ceramics in a time-dependent manner. The results also showed that the capabilities of t-HA ceramics and t- β -TCP ceramics to support in vitro osteogenic differentiation of MSCs were similar to that of TCPS dishes.

Recently, composites of various types of biomaterials and cultured cells have been used for clinical treatments. In fact, we have succeeded in fabricating regenerative cultured bone tissues by using various types of ceramic scaffolds and MSCs derived from the patients' own bone marrow (Fig. 1). The regenerative cultured bone tissues were then implanted into the same patients. Before implantation, evidence of adequate cellular functions is strongly needed to guarantee the efficacy of cell-based therapy. In the case of bone tissue regeneration, validation of the osteogenic capability of the MSCs on the ceramics is a prerequisite for clinical applications. However, it is difficult to validate this capability in cells cultured on ordinary ceramics by light microscopy because of the opacity of the ceramic material. Here, we demonstrated that transparent ceramics enable direct monitoring of living cells by light microscopy. Therefore, it is suggested that transparent biomaterials could be used as a validation system for culturing cells on such materials because they will be able to give us information such as cell morphology and differentiation. We concluded that transparent bioactive materials could be useful materials for clinical research in regenerative medicine as well as in vitro basic research in tissue engineering.

CONCLUSIONS

We fabricated new bioactive materials, t-HA ceramics and t- β -TCP ceramics as cell culture scaffolds. The ceramic materials enabled the observation of cultured cells by phase-contrast microscopy. Cell shapes on both types of transparent ceramics were clearly observed at almost the same level as that observed on TCPS dishes. In this study, rat MSCs were cultured on the substrates. Cell adhesion, proliferation, and differentiation of rat MSCs on t-HA ceramics and t- β -TCP ceramics were significantly similar to those on TCPS dishes and were confirmed by qualitative

biochemical assays. The observation of living cells on t-HA ceramics and t- β -TCP ceramics could be used as a validation system of cellular behavior in the fields of regenerative medicine and tissue engineering.

ACKNOWLEDGMENTS

This work was supported by The Three-Dimensional Tissue Module Project of METI (a Millennium Project) and a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture, Japan. We thank our colleagues at the Research Institute for Cell Engineering (RICE), National Institute of Advanced Industrial Science and Technology (AIST).

ARTICLES OF FURTHER INTEREST

Alumina; Bioactive Glass; Biologic and Synthetic Apatites; Biomimetic Materials; Tissue Engineering; Zirconia Ceramics

REFERENCES

1. Finn, R.A.; Bell, W.H.; Brammer, J.A. Interpositional grafting with autogenous bone and coralline hydroxyapatite. *J. Maxillofac. Surg.* **1980**, *8*, 217–227.
2. Gross, U.M.; Muller-Mai, C.M.; Voigt, C. The interface of calcium-phosphate and glass-ceramic in bone, a structural analysis. *Biomaterials* **1990**, *11*, 83–85.
3. Ohgushi, H.; Goldberg, V.M.; Caplan, A.I. Heterotopic osteogenesis in porous ceramics induced by marrow cells. *J. Orthop. Res.* **1989**, *7*, 568–578.
4. Ohgushi, H.; Goldberg, V.M.; Caplan, A.I. Repair of bone defects with marrow and porous ceramic (experiments in rats). *Acta. Orthop. Scand.* **1989**, *60*, 334–339.
5. Hench, L.L. Bioactive ceramics. *Ann. N.Y. Acad. Sci.* **1988**, *523*, 54–71.
6. Gross, U.; Strunz, V. The interface of various glasses and glass ceramics with a bony implantation bed. *J. Biomed. Mater. Res.* **1985**, *19*, 251–271.
7. Ohgushi, H.; Dohi, Y.; Katuda, T.; Tamai, S.; Tabata, S.; Suwa, Y. *In vitro* bone formation by rat marrow cell culture. *J. Biomed. Mater. Res.* **1996**, *32*, 333–340.
8. Yoshikawa, T.; Ohgushi, H.; Tamai, S. Immediate bone forming capability of prefabricated osteogenic hydroxyapatite. *J. Biomed. Mater. Res.* **1996**, *32*, 481–492.
9. Ohgushi, H.; Dohi, Y.; Yoshikawa, T.; Tamai, S.; Tabata, S.; Okunaga, K.; Shibuya, T. Osteogenic differentiation of cultured marrow stromal stem cells on the surface of bioactive glass ceramics. *J. Biomed. Mater. Res.* **1996**, *32*, 341–348.
10. Ohgushi, H.; Caplan, A.I. Stem cell technology and bioceramics: from cell to gene engineering. *J. Biomed. Mater. Res.* **1999**, *48*, 913–927.

11. Ohgushi, H.; Okumura, M.; Tamai, S.; Shors, E.C.; Caplan, A.I. Marrow cell induced osteogenesis in porous hydroxyapatite and tricalcium phosphate. *J. Biomed. Mater. Res.* **1990**, *24*, 1563–1570.
12. Ohgushi, H.; Okumura, M.; Yoshikawa, T.; Inoue, K.; Senpuku, N.; Tamai, S. Bone formation process in porous calcium carbonate and hydroxyapatite. *J. Biomed. Mater. Res.* **1992**, *26*, 885–895.
13. Kihara, T.; Oshima, A.; Hirose, M.; Ohgushi, H. Three-dimensional visualization analysis of *in vitro* cultured bone fabricated by rat marrow mesenchymal stem cells. *Biochem. Biophys. Res. Commun.* **2004**, *316*, 943–948.
14. Mendes, S.C.; Tibbe, J.M.; Veenhof, M.; Bakker, K.; Both, S.; Platenburg, P.P.; Oner, F.C.; De Bruijn, J.D.; Van Blitterswijk, C.A. Bone tissue-engineered implants using human bone marrow stromal cells: effect of culture conditions and donor age. *Tissue Eng.* **2002**, *8*, 911–920.
15. de Bruijn, J.D.; van den Brink, I.; Mendes, S.; Dekker, R.; Bovell, Y.P.; van Blitterswijk, C.A. Bone induction by implants coated with cultured osteogenic bone marrow cells. *Adv. Dent. Res.* **1999**, *13*, 74–81.
16. Kasten, P.; Luginbuhl, R.; van Griensven, M.; Barkhausen, T.; Krettek, C.; Bohner, M.; Bosch, U. Comparison of human bone marrow stromal cells seeded on calcium-deficient hydroxyapatite, beta-tricalcium phosphate and demineralized bone matrix. *Biomaterials* **2003**, *24*, 2593–2603.
17. Mayr-Wohlfart, U.; Fiedler, J.; Gunther, K.P.; Puhl, W.; Kessler, S. Proliferation and differentiation rates of a human osteoblasts-like cell line (SaOS-2) in contact with different bone substitute materials. *J. Biomed. Mater. Res.* **2001**, *57*, 132–139.
18. Ioku, K.; Kawagoe, D.; Toya, H.; Fujimori, H.; Goto, S.; Ishida, K.; Mikuni, A.; Mae, H. OH-designed transparent apatite ceramics prepared by spark plasma sintering. *Trans. Mater. Res. Soc. Jpn.* **2002**, *27*, 447–449.
19. Kotobuki, N.; Ioku, K.; Kawagoe, D.; Fujimori, H.; Goto, S.; Ohgushi, H. Observation of osteogenic differentiation cascade of living mesenchymal stem cells on transparent hydroxyapatite ceramics. *Biomaterials* **2005**, *26*, 779–785.
20. Kotobuki, N.; Kawagoe, D.; Nomura, D.; Kato, Y.; Muraki, K.; Fujimori, H.; Goto, S.; Ioku, K.; Ohgushi, H. Observation and quantitative analysis of rat mesenchymal stem cells cultured *in vitro* on newly formed transparent β -tricalcium phosphate. *J. Mater. Sci. in press*.
21. Maniatopoulos, C.; Sodek, J.; Melcher, A.H. Bone formation *in vitro* by stromal cells obtained from bone marrow of young adult rat. *Cell. Tissue Res.* **1988**, *254*, 317–330.
22. Uchimura, E.; Machida, H.; Kotobuki, N.; Kihara, T.; Kitamura, S.; Ikeuchi, M.; Hirose, M.; Miyake, J.; Ohgushi, H. *In-situ* visualization and quantification of mineralization of cultured osteogenic cells. *Calcif. Tissue Int.* **2003**, *73*, 575–583.
23. Ohgushi, H.; Okumura, M. Osteogenic ability of rat and human marrow cells in porous ceramics. *Acta. Orthop. Scand.* **1990**, *61*, 431–434.
24. Okumura, M.; Ohgushi, H.; Tamai, S. Bonding osteogenesis in coralline hydroxyapatite combined with bone marrow cells. *Biomaterials* **1991**, *12*, 411–416.
25. Goshima, J.; Goldberg, V.M.; Caplan, A.I. Osteogenic potential of culture-expanded rat marrow cells as assayed *in vitro* with porous calcium phosphate ceramic. *Biomaterials* **1991**, *12*, 254–258.
26. Yoshikawa, T.; Ohgushi, H.; Uemura, T. Human marrow cell-derived cultured bone in porous ceramics. *J. Biomed. Mater. Eng.* **1998**, *8*, 311–320.
27. Ohgushi, H.; Yoshikawa, T.; Nakajima, H.; Tamai, S.; Dohi, Y.; Okunaga, K. Al_2O_3 doped apatite-wollastonite containing glass ceramic provokes osteogenic differentiation of marrow stromal stem cells. *J. Biomed. Mater. Res.* **1999**, *44*, 381–388.
28. Yoshikawa, T.; Ohgushi, H.; Dohi, Y.; Davies, J.E. Viable bone formation in porous hydroxyapatite: marrow cell-derived *in vitro* bone on the surface of ceramics. *Biomed. Mater. Eng.* **1997**, *7*, 49–58.
29. Kitamura, S.; Ohgushi, H.; Hirose, M.; Funaoka, H.; Takakura, Y.; Ito, H. Osteogenic differentiation of human bone marrow-derived mesenchymal cells cultured on alumina ceramics. *Artif. Organs* **2004**, *28*, 72–82.
30. Ohgushi, H.; Kitamura, S.; Kotobuki, N.; Hirose, M.; Machida, H.; Muraki, K.; Takakura, Y. Clinical application of marrow mesenchymal stem cells for hard tissue repair. *Yonsei Med. J.* **2004**, *45*, 61–67.
31. Kotobuki, N.; Hirose, M.; Takakura, Y.; Ohgushi, H. Cultured autologous human cells for hard tissue regeneration: preparation and characterization of mesenchymal stem cells from bone marrow. *Artif. Organs* **2004**, *28*, 33–39.
32. Ohgushi, H.; Kotobuki, N.; Funaoka, H.; Machida, H.; Hirose, M.; Tanaka, Y.; Takakura, Y. Tissue engineered ceramic artificial joint -*ex vivo* osteogenic differentiation of patient mesenchymal cells on total ankle joints for treatment of osteoarthritis. *Biomaterials* **2005**, *26*, 4654–4661.

Ultra-High Molecular Weight Polyethylene (UHMWPE)

Kwideok Park

Department of Biomedical Engineering, The University of Memphis, Memphis, Tennessee, U.S.A.

Gladius Lewis

Department of Mechanical Engineering, The University of Memphis, Memphis, Tennessee, U.S.A.

Joon B. Park

Department of Biomedical Engineering, The University of Iowa, Iowa City, Iowa, U.S.A.



INTRODUCTION

Ultra-high molecular weight polyethylene (UHMWPE) is widely used in orthopedics, especially in the fabrication of articular components for arthroplasties, in particular, the acetabular cup or liner in a total hip joint replacement (THJR) and the tibial insert in a total knee joint replacement (TKJR). UHMWPE is used for these applications due to its excellent biocompatibility, high impact strength, remarkable wear resistance, and low friction coefficient. In this article on UHMWPE, we address four topics. The first section presents the procedure of how UHMWPE powder is synthesized and then processed into test specimens (for laboratory studies) and near-final and/or final net-shape components (for orthopedic uses). A key element of this synthesis is the use of a Ziegler-Natta catalyst in the chain polymerization process. The processing methods in current use are compression sintering, ram extrusion, and direct compression molding.

In the second section, the structure and properties of UHMWPE are discussed, with special reference to lamellae structure, plastic deformation of crystallites, microscopic surface morphology, and physical and mechanical properties. The lamellae (crystalline units of UHMWPE) are known to have an important influence on the mechanical properties of the polymer. UHMWPE powder particles have a porous and irregularly shaped surface morphology. However, scanning electron microscopy (SEM) observation at a high magnification reveals a complex and hierarchical structure. These morphological features play an important role in the consolidation of UHMWPE powders during the sintering process and affect the resultant bulk mechanical properties of the polymer.

The focus of the third section is description of the various methods used for sterilizing UHMWPE implants and the resultant structural changes that they produce in the polymer. The special emphasis in this section is gamma irradiation in ambient air. This sterilization method has raised many concerns in

regard to the long-term in-vitro and in-vivo stability of UHMWPE parts, because the gamma irradiation causes progressive and irreversible damage to the polymer, resulting in significant depreciation in its physical, chemical, and mechanical properties.

Some recent developments and future challenges in the field of UHMWPE are discussed briefly in the fourth section.

SYNTHESIS AND PROCESSING

In the early 1950s, Karl Ziegler of the Max Planck Institute in Muhlheim-Ruhr, Germany, developed a method for synthesizing UHMWPE powders. This method uses a Ziegler-Natta catalyst made from titanium chloride and organo-aluminum compounds (triethyl-aluminum or diethyl-aluminum chloride). Because these chemicals are very sensitive to air and moisture, all the steps in the synthesis are carried out in an inert gas environment. Pure ethylene and a suspending agent without polar impurities are important prerequisites for successful synthesis. Catalysts are produced in one vessel and are then diluted in a second vessel. The catalyst suspension is pumped into a polymerization reactor, where it reacts with the ethylene gas to form polyethylene. Polymerization takes place at a temperature of between 65°C and 85°C under a pressure of between 0.5 MPa and 2.0 MPa. The molecular weight of polyethylene can be controlled by adding very small amounts of hydrogen to the ethylene gas or by changing the polymerization temperature.^[1]

In chain polymerization, the ethylene molecules remain attached onto the active sites of the catalyst particles. The polyethylene keeps growing around these particles until the polymer finally encapsulates them. Polymer particles become much larger than the catalyst particles, but the polymerization does not proceed indefinitely. Instead, it has a limit, yielding a polymer with a molecular weight of no more than

Table 1 Nomenclature of ticona and montell UHMWPE resins

Resin	Manufacturer	Mean M.W. (10^6 g/mol)	Calcium stearate included	Maximum conc. of calcium stearate (ppm)
GUR 1150	Hoechst/Ticona	5.5~6.0	Yes	500
GUR 1050	Hoechst/Ticona	5.5~6.0	No	75
GUR 1120	Hoechst/Ticona	3.5	Yes	500
GUR 1020	Hoechst/Ticona	3.5	No	75
1900	Himont/Montell	4.4~4.9	No	NA
1900H	Himont/Montell	> 4.9	No	NA

(Note: Prior to 1998, when Ticona was known as Hoechst, the UHMWPEs that they produced in the United States were designated as GUR 4150, GUR 4050, GUR 4120, and GUR 4020.)

(From Ref. [2].)

between 6×10^6 g/mol and 7×10^6 g/mol. With the polyethylene particles that remain in suspension, a centrifuge separates the suspending agent, and the remaining solvent is removed by steam distillation. Hot nitrogen is used to remove the residual moisture. The final product is a fine white powder. Though the sizes of the powder particles are not the same, most of them are less than 300 μ m in diameter. As an additive, very small amounts of fine calcium stearate may be added to the polyethylene powders to eliminate the residual catalysts, which have the potential for corroding the conversion machines. The calcium stearate has two other functions. First, it acts as a lubricant when a part is to be fabricated using ram-extrusion of the polymer powders. Second, it helps the polyethylene powders to maintain a white color. However, there are concerns about using calcium stearate, as it may cause fusion defects in a fabricated part and, hence, lead to a depreciation of its properties. Fabricators of UHMWPE parts that are used in arthroplasties are now switching their protocols to produce UHMWPE resins with either very small amounts of calcium stearate or none at all (Table 1). The issue of the quality of UHMWPE to be used in fabricating orthopedic articular components is addressed in ASTM F 648-96 (Table 2).^[3] (It should be noted, however, that the standard applies to raw UHMWPE powders and fabricated forms of UHMWPE, and not to packaged and sterilized implants.)

The high melt-flow resistance of UHMWPE makes the processing of UHMWPE powders into bulk products very challenging. There are three fabrication methods in current use. The first is compression sintering. A charge of polymer powder is evenly distributed in a female mold cavity and then covered by a male mold. The system is then subjected to a combination of high temperature and high pressure for a certain amount of time. Then the system is cooled at a slow and uniform rate in order to minimize shrinkage and deformation of the product (the stock). The stock

is then machined into smaller blocks or cylindrical bars from which UHMWPE articular components can be machined. The second method, ram extrusion of the UHMWPE powders, produces cylindrical bar stocks ranging in diameter from 25 mm to 150 mm. In this process, the UHMWPE powder charge is fed into a chamber and then heat is applied (through electrical resistors) to the charge. A ram then compresses and extrudes the plasticized powder charge into a heated cylindrical barrel, where it is consolidated into a cylindrical bar stock. (It should be noted that as the ram moves back and forth, the powder stock in the chamber is refilled.) UHMWPE articular components may then be machined from this bar stock. In the third method, direct compression molding, the UHMWPE powder charge is consolidated into a final or semi-final bulk product using a preshaped mold. Although this process is slow and costly, orthopedic articular components made using this method have very smooth surface finishes and excellent dimensional consistency.

Fabrication method and resin type are known to be two important factors that influence the clinical

Table 2 Requirements for UHMWPE, per ASTM F 648-96

Property or feature	Requirement
Molecular weight	3×10^6 g/mol
Yield strength	19 MPa
Elongation (minimum)	250%
Density	0.927 g/cm ³ –0.944 g/cm ³
Hardness (minimum)	60 D
Impact strength	30 kJ/m ²
Contaminants	10 parts/300 g
Fusion defects	Not specified
Other	As specified

(From Ref. [3].)

performance of UHMWPE articular components. Examinations of retrieved acetabular cups revealed that directly compression-molded components (made using polymer powders that do not contain calcium stearate) showed better wear performance compared to ram-extruded and machined ones.^[4] Another study evaluated gamma-irradiated, in ambient air, and shelf-aged (2 months–16.5 years) tibial inserts made from Himont 1900, GUR 415, and GUR 412 resins.^[5] A significant difference was found in the oxidation rates between these two sets of inserts; specifically, the Himont 1900 inserts oxidized more slowly and possessed better mechanical properties than those fabricated from GUR 415 or GUR 412.

STRUCTURE AND PROPERTIES

UHMWPE is a linear homopolymer (with a repeating unit of CH_2CH_2) that has C—H and C—C covalent single bonds, with the neighboring chains interacting through van der Waals forces (Fig. 1). UHMWPE is a semicrystalline polymer, which means that it has both crystalline and amorphous phases. The regularity of the arrangement of the polymer chains gives the polymer a unique crystalline structure. In the amorphous state, the macromolecular chains are randomly oriented. In the crystalline region of polyethylene, groups of orthorhombic unit cells form a lamella, which is composed of the folded polyethylene chains (Fig. 2). The lamellae are packed together in parallel fashion and are separated from the layers of amorphous polymer chains. The thickness and length of a lamella are, typically, 5–25 nm and 10–50 μm , respectively.

In the folded-chain model, the macromolecular chains form a fold within the lamellae and reenter into the same crystallites. Several models have been proposed to explain the folding of polymer chains into the lamellae from which they emerge. However, a common feature of all these theories is the point that there are molecules that do not fold back into the same lamellae but, rather, become part of the lamellae. Such

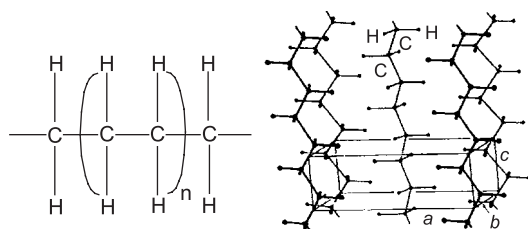


Fig. 1 Basic structural unit of polyethylene with the repeating unit in parentheses and an orthorhombic crystal structure of polyethylene with $a = 0.74$ nm, $b = 0.493$ nm and $c = 0.254$ nm. (From Ref. [6].)

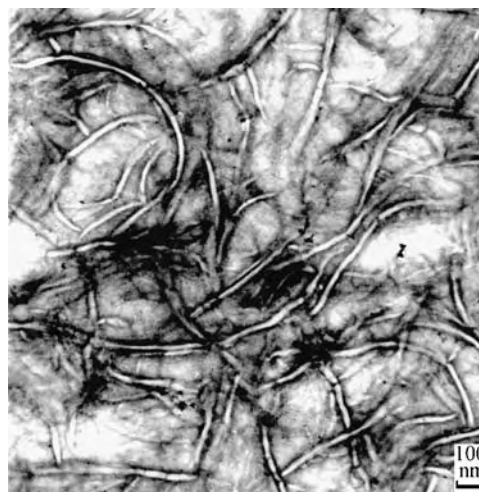


Fig. 2 Transmission electron micrograph of the lamella structure of UHMWPE. (Courtesy of Stephen Spiegelberg of Cambridge Polymer Group, Inc., Cambridge, MA.)

molecules are called tie molecules. Formation of tie molecules provides high toughness to the polymer because these molecules link the lamellae together. In the absence of these intercrystalline connections, it would be difficult for UHMWPE to undergo large plastic deformation.

UHMWPE powders have a very intriguing morphology (Fig. 3). At low magnification, the particles have a porous and irregularly shaped, rounded structure. However, at high magnification, the structure is complicated and hierarchical. A nodule is a basic structural unit and these nodules are interconnected by very fine fibrils. Aggregates of the nodules organize into a larger structural form and groups of the aggregates form a single powder particle. These highly developed structural features can contribute to a high sintering time for the UHMWPE powders; for example, 25 min at 180°C for a polymer with mean molecular weight (M_m) of 4×10^6 g/mol, and 90 min at 180°C for a polymer with M_m of 5×10^5 g/mol. It should be noted that the surface morphology of UHMWPE powder particles plays a significant role both in the sintering behavior of UHMWPE powders and in the mechanical properties of the solid polymer.^[7,8]

The deformation mechanism of UHMWPE reflects the relationship between its structure and its mechanical behavior. When a semicrystalline polymer, such as UHMWPE, is subjected to increasing stress, plastic deformation does not occur until the stress exceeds a certain limit (yield strength), at which point elastic response ends and permanent structural damages take place. In the case of UHMWPE, deformation begins with the disruption of the amorphous phase and breakdown of the crystalline phase is followed by a

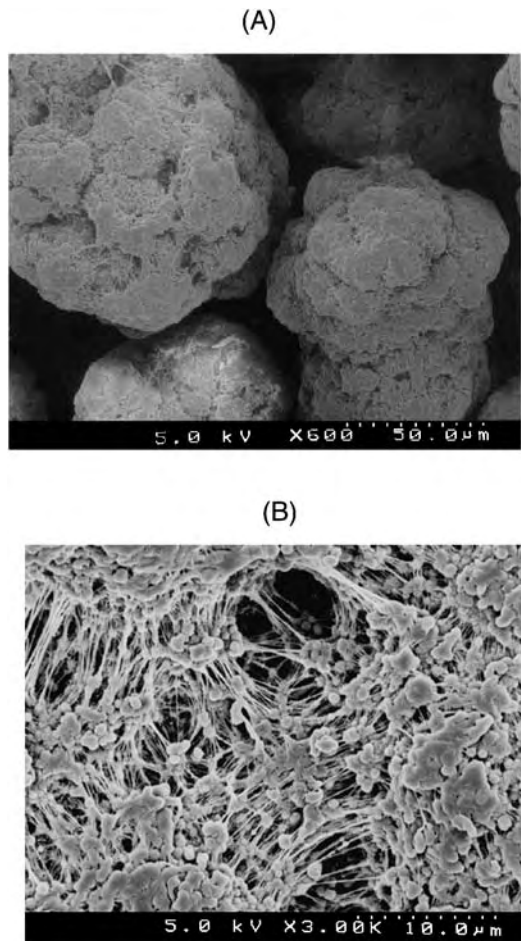


Fig. 3 Scanning electron micrographs of UHMWPE powder particles at low (A) and high (B) magnification.

large plastic strain. In some cases, the contribution of the amorphous phase is of importance in the interpretation of the viscoelastic behavior of the polymer.^[9] The predominant deformation of UHMWPE occurs through a crystallographic slip within the crystalline regions. In general, a single slip system can lead only to a simple shear deformation of the crystallites, not structural changes. Slip would take place more easily along directions where the dislocation motions require the least energy. When an UHMWPE specimen is loaded in tension, it is elongated without plastic deformation. Beyond the yield strength, however, plastic deformation begins and the polymer elongates further. At this point, more slips are produced and the resultant dislocations interact and, eventually, entangle. These events make further slip difficult, resulting in strain hardening of the polymer. More details about the plastic deformation of polyethylene have been given by many workers.^[6,10]

In some literature reports, UHMWPE is referred to as high-density polyethylene. However, there are many

Table 3 Comparison of physical and chemical properties of high-density polyethylene (HDPE) and ultra-high molecular weight polyethylene, UHMWPE

Property	HDPE	UHMWPE
Molecular weight (10^6 g/mole)	0.05–0.20	2.0–6.0
Melting point ($^{\circ}\text{C}$)	130–137	125–135
Density (g/cm^3)	0.952–0.965	0.930–0.945
Tensile yield strength (MPa)	26.2–33.1	19.3–23.0
Elongation at failure (%)	10–1,200	200–350
Tensile modulus (GPa)	0.4–4.0	0.8–1.5
Izod impact (J/m)	21–210	>1,070–no break
Shore-D hardness	66–73	60–65

(From Ref. [11].)

differences between these two polymers (Table 3). In comparison to conventional polyethylene, the long linear chains of UHMWPE compromise their ability to arrange themselves into crystalline structures such as spherulites, thus limiting the percent crystallinity (%C) and density. The density of UHMWPE is close to that of low-density polyethylene ($0.925\text{--}0.935\text{ g}/\text{cm}^3$). The melting point (T_m) and % C of UHMWPE depend on the resin type. Given the load-bearing duty of UHMWPE articular components in total joint prostheses, the polymer's fracture toughness is one of the many mechanical properties that is of particular importance. For tough polymers, such as UHMWPE, the fracture toughness is determined using the J-integral method. Lewis and Nyman^[12] reported a J-integral fracture toughness of $92.0\text{ kJ}/\text{m}^2$ for unsterilized and unaged UHMWPE and $44.3\text{ kJ}/\text{m}^2$ for gamma-irradiated, in air, unaged UHMWPE.

STERILIZATION METHODS AND STRUCTURAL MODIFICATIONS

Until recently, most orthopedic device companies sterilized UHMWPE articular components by gamma irradiating them, in ambient air. For this purpose, gamma rays from a cobalt-60 source, at a dosage of between 2.5 Mrad and 4.0 Mrad (25–40 kGy), were commonly used. However, it is now widely known that this sterilization method brings about extensive, deleterious, and irreversible structural changes in the polymer; namely, breakage of the C—C bonds (chain scission) and breakage of the C—H bonds to create free radicals, the majority of which remain highly reactive in the polymer structure for extended periods. (It is fair to point out that a few of the carbon radicals

recombine to form cross-links, a structural change that has a positive effect on some of the polymer's properties.) These free radicals can take part in oxidative reactions in the polymer by reacting with oxygen species from one or more of the following sources: 1) dissolved oxygen present at the time of the irradiation; 2) oxygen that has diffused into the component during shelf storage; and 3) for the implanted articular components, dissolved oxygen in the body fluid. There have been numerous reports of oxidative degradation of UHMWPE. In many of these reports, the change in the peak intensity of the carbonyl region (wave number of 1720 cm^{-1}) in the FTIR spectrum of UHMWPE is taken as index of the extent of oxidation (or oxidation index, OI) of the polymer. According to one study, the OI of shelf-aged (2–11 years) UHMWPE acetabular liners increased with shelf aging time and showed a maximum level at a layer positioned between 0.5 mm and 2.0 mm below the surface.^[13] It has also been reported that sterilization using ethylene oxide gas (EtO) or gas plasma leads to a substantially lower level of OI of UHMWPE compared to when gamma irradiation, in ambient air, is used.^[14]

The oxidative degradation of UHMWPE is affected by many parameters, notably sterilization method and poststerilization aging conditions (environment and time). Furthermore, it is widely agreed that the oxidative degradation of UHMWPE causes or substantially contributes to the depreciation of the in-vitro properties and in-vivo behavior of UHMWPE components that are gamma-irradiated, in ambient air. Densities of the gamma-irradiated UHMWPE specimens continuously increased up to a radiation dosage of 16 Mrad but remained unchanged thereafter.^[15] The %C showed a similar trend. The increase in %C is highly associated with rearrangement of the polymer chains. It is believed that the radiation-induced scission of tie molecules in UHMWPE contributes to: 1) an increase of %C, 2) crystallization in amorphous regions, and 3) crystal growth in the crystalline regions.^[16] Gamma irradiation, in ambient air, causes a stiffening of the material, leading to increases in the elastic modulus and the yield strength of UHMWPE. In turn, increased stiffness decreases the polymer's ductility and, consequently, influences its fracture toughness and fatigue strength.

Sutula et al.^[17] reported the side effects of gamma irradiation, in ambient air, of UHMWPE, which altered the polymer's chemical and mechanical properties over time. The results showed high subsurface oxidation, reduced ductility, and decreased strength. A subsurface white band was frequently observed in the gamma-sterilized components and this zone was also where significantly reduced levels of strength and ductility of the polymer were seen. Tensile tests of 23 UHMWPE acetabular components showed that

presence of the subsurface white band adversely affected the mechanical properties of the polyethylene.^[17] Although the ultimate tensile strength (UTS) and % elongation in the absence of gamma sterilization and subsurface white band were 45 MPa and 630%, respectively, the corresponding values for specimens that were gamma irradiated and exhibited the subsurface white band were 20 MPa and 70%, respectively. Fatigue strength of the polymer, a property that is very important for its use in fabricating articular components for arthroplasties, is also adversely affected by gamma irradiation compared to other sterilization methods.^[18] Results of rotating beam fatigue tests found the fatigue limit to be 17.5 MPa for UHMWPE gamma irradiated, in ambient air, and 31 MPa for unsterilized UHMWPE. Even though the patterns of the applied-stress-versus-number-of-cycles-to-failure curve, over the first two million cycles, were the same for these two sets of specimens, the fatigue strength of the gamma-irradiated UHMWPE began to decrease sharply and resulted in more than 40% reduction after ten million cycles. The fatigue strength (28 MPa) for EtO-sterilized UHMWPE was not significantly lower than that for the unsterilized material.

The undesirable outcomes of gamma irradiation, in ambient air, prompted the search for alternative sterilization methods, such as: 1) gamma irradiation, in a reduced oxygen, nitrogen gas, or vacuum environment; gamma radiation, in a vacuum foil package; 2) sterilization using EtO; and 3) sterilization using gas plasma. EtO sterilization of UHMWPE parts is accomplished by the diffusion of EtO gas into the near-surface areas to kill bacteria, viruses, and spores. All residual EtO is then removed from the UHMWPE part. Gas plasma sterilization is accomplished by subjecting the UHMWPE part to ionized gas (peracetic acid or hydrogen peroxide gas plasma) at a temperature below 50°C . Investigations suggest that EtO and gas plasma sterilization methods do not adversely affect the chemical, physical, and mechanical properties of UHMWPE.^[19,20] The alternative sterilization methods all led to materials with lower values of OI and higher UTS compared to the corresponding values for UHMWPE that was gamma irradiated, in ambient air. Furthermore, the clinical performance of UHMWPE articular components that are sterilized using EtO or gas plasma has been reported to be acceptable during short-term use.^[21]

RECENT DEVELOPMENTS AND FUTURE WORK

Progress continues to be made on two fronts, namely: 1) improvement in the quality of the UHMWPE stock material; and 2) use of cross-linking (and associated postcross-linking stabilization regimens) for producing



UHMWPE parts with excellent in-vitro and in-vivo wear performance.

Two examples of the commercially available improved stock materials are ArCom™ (Biomet, Inc., Warsaw, IN, U.S.A.) and Duramer™ (Wright Medical Technology, Inc., Arlington, TN, U.S.A.). ArCom™ uses a very high molecular weight resin followed by sterilizing the stock using gamma radiation, in an argon environment. After further processing, final products are placed in an argon-filled package. The combination of the sterilization and packaging methods used leads to an enhancement of the material's properties, particularly its wear resistance. ArCom™ is used to fabricate the tibial inserts in, for example, AGC and TCCK designs (Biomet, Inc., Warsaw, IN, U.S.A.) of TKJR. Duramer™ is made by an optimized ram extrusion processing of GUR 415 resin followed by sterilization of the final part using EtO.

The introduction of cross-links within the UHMWPE structure has generated great interest due to the significantly improved wear resistance of the polymer compared to the uncross-linked variant. Currently, three methods that are widely used for cross-linking UHMWPE are gamma irradiation, electron beam irradiation, and chemicals.^[22] With gamma irradiation, the dosage used ranges from 5 to 100 Mrad. As for chemical-induced cross-linking, dicumyl peroxide and 2,5-dimethyl-2,5-bis (*tert*-butylperoxy)-3-hexyne (Peroxide 130) have gained the most attention as the chemical. These peroxides, either in liquid or powder form, are typically mixed with UHMWPE powders and then consolidated by compression molding or hot isostatic pressing. The peroxides begin to decompose in the system by chemical reactions, culminating in the formation of the cross-links within polymer network. It should be pointed out that when the gamma or electron beam irradiation method is used, cross-linking occurs in the solid state of the UHMWPE, but when peroxide is used, cross-linking takes place above the polymer's melting temperature. This results in decreased density of peroxide-cross-linked UHMWPE due to a hindered crystallization of cross-linked polyethylene melt. Substantial improvement of wear resistance has been extensively reported with cross-linked UHMWPE. McKellop et al.^[23] reported that, in hip joint simulator tests, the wear resistance of UHMWPE acetabular cups cross-linked by 28 Mrad gamma radiation was excellent. In another study,^[24] it was reported that the steady-state radiographic wear rate was 0.06 mm/year for implanted 100 Mrad gamma-irradiated UHMWPE acetabular cups. This value was substantially lower than the mean of 0.29 mm/year for unirradiated cups. Oonishi et al.^[25] examined retrieved polyethylene acetabular cups that were 100 Mrad irradiated before implantation (in-vivo time of between 15 years and 16 years). Compared to

the unirradiated cups, the UTS and gel content of the retrieved UHMWPE cups did not show significant deterioration in the presence of a high free radical level.

The enhanced wear resistance of cross-linked UHMWPE is associated with the fact that cross-linking contributes to a decrease in the mobility of polymer chains. Edidin et al.^[26] have postulated that, in UHMWPE, plasticity-induced lamellar alignment is believed to be a precursor of wear generation, because this irreversible plastic deformation is a consequence of the structural deformation of the polymer followed by continuous cyclic load. The suppressed chain mobility thus minimizes the depth of the lamellar alignment layer from the articulating surface, which, in turn, produces fewer wear debris. Despite improved wear resistance, there have been many concerns about gamma-irradiation-cross-linked UHMWPE, especially with respect to its decreased ductility, fatigue crack propagation resistance, and fracture toughness.^[18,27] Investigation of the fracture toughness of the cross-linked UHMWPE found a steady decrease of resistance of crack propagation (J_{ss}) with increase in gamma irradiation dosage (116.9 kJ/m² at 0 Mrad; 98.5 kJ/m² at 5 Mrad; and 79.3 kJ/m² at 20 Mrad).^[28] Cole et al.^[29] showed that the reduction of fracture toughness is both radiation dosage- and resin type-dependent. In this study, two different resins were examined (GUR 1050 and Himont 1900). The fracture toughness values were reported to be 40.4 kJ/m² and 15.4 kJ/m² for GUR 1050 at 3.3 Mrad and 10.0 Mrad, and 21.6 kJ/m² and 19.1 kJ/m² for Himont 1900 at the same radiation dosages. Another problem with highly cross-linked UHMWPE is the large amount of free radicals produced after radiation or peroxide cross-linking. These free radicals can accelerate the oxidative degradation of an UHMWPE articular component during shelf storage as well as in vivo. In an effort to eliminate as many of these free radicals as possible, it is now standard practice to subject the cross-linked part to some stabilization treatment; for example, UHMWPE parts that are cross-linked using electron beam radiation are annealed at a temperature above the polymer's melting temperature.^[30]

Although cross-linked UHMWPE parts have been shown to have substantially improved wear resistance (relative to uncross-linked ones), other relevant material properties (such as fracture toughness and fatigue crack propagation rate) are sacrificed. Thus, as far as UHMWPE is concerned, the main challenge facing orthopedic device manufacturers is to develop cross-linking method(s) that will simultaneously enhance the polymer's wear performance with no accompanying depreciation in any of its other important material properties.

ARTICLES OF FURTHER INTEREST

Alumina; Implant, Total Hip; Knee Joint Replacement; Orthopedic Biomaterials; Polymers; Wear Debris

REFERENCES

- Birnkrant, H.W. Synthesis of UHMWPE. In *Ultra-High Molecular Weight Polyethylene as Biomaterial in Orthopaedic Surgery*; Willert, H.G., Buchhorn, G.H., Eyerer, P., Eds.; Hogrefe & Huber Publishers: New York, 1991; 3–5.
- Steven, S.M.; Kurtz, M.; Muratoglu, O.K.; Evans, M.; Edidin, A.A. Advances in the processing, sterilization, and crosslinking of ultra-high molecular weight polyethylene for total joint arthroplasty. *Biomaterials* **1999**, *20*, 1659–1688.
- American Society for Testing and Materials (ASTM). *Standard Specification for Ultra-high Molecular Weight Polyethylene Powder and Fabricated Form for Surgical Implants*; ASTM International: West Conshohocken, PA, 1996; 144–149.
- James, S.P.; Lee, K.R.; Beauregard, G.P.; Rentfrow, E.D.; McLaughlin, J.R. Clinical wear of 63 ultrahigh molecular weight polyethylene acetabular components: Effect of starting resin and forming method. *J. Biomed. Mater. Res. (Appl. Biomater.)* **1999**, *48*, 374–384.
- Currier, B.H.; Currier, J.H.; Collier, J.P.; Mayor, M.B. Effect of fabrication method and resin type on performance of tibial bearings. *J. Biomed. Mater. Res.* **2000**, *53*, 143–151.
- Lin, L.; Argon, A.S. Structure and plastic deformation of polyethylene. *J. Mater. Sci.* **1994**, *29*, 294–323.
- Siegmann, A.; Raiter, I.; Narkis, M.; Eyerer, P. Effect of UHMW-PE Powder Particle Morphology on Its Sintering Behavior. In *Ultra-high Molecular Weight Polyethylene as Biomaterial in Orthopaedic Surgery*; Willert, H.G., Buchhorn, G.H., Eyerer, P., Eds.; Hogrefe & Huber Publishers: New York, 1991; 11–14.
- Zachariades, A.E. The effect of initial morphology on the mechanical properties of ultra-high molecular weight polyethylene. *Polym. Eng. Sci.* **1986**, *26*, 658–661.
- Relaxation Phenomena in Polymers*; Matsuoka, S., Ed.; Hanser Publishers: Munich, Germany, 1992.
- Park, J.B. Strength and Strengthening Mechanisms. In *Biomaterials Science and Engineering*; Plenum Press: New York, 1984; 85–100.
- Li, S.; Burstein, A.H. Current concepts review: Ultra-high molecular weight polyethylene. *J. Bone Jt. Surg.* **1994**, *76A*, 1080–1090.
- Lewis, G.; Nyman, J.S. A new method of determining the J-integral fracture toughness of very tough polymers: Application to ultra-high-molecular-weight polyethylene. *J. Long-Term Eff. Med. Implants* **1999**, *9*, 289–301.
- Jacob, R.J.; Pienkowski, D.; Lee, K.Y.; Hamilton, D.M.; Schroeder, D.; Higgins, J. Time- and depth-dependent changes in crosslinking and oxidation of shelf-aged polyethylene acetabular liners. *J. Biomed. Mater. Res.* **2001**, *56*, 168–176.
- Goldman, M.; Lee, M.; Gronsky, R.; Pruitt, L. Oxidation of ultrahigh molecular weight polyethylene characterized by Fourier transform infrared spectrometry. *J. Biomed. Mater. Res.* **1997**, *37*, 43–50.
- Bhateja, S.K.; Andrews, E.H.; Young, R.J. Radiation-induced crystallinity changes in linear polyethylene. *J. Polym. Sci., Polym. Phys. Ed.* **1983**, *21*, 523–536.
- Shinde, A.; Salovey, R. Irradiation of ultrahigh molecular weight polyethylene. *J. Polym. Sci., Polym. Phys. Ed.* **1985**, *23*, 1681–1689.
- Sutula, L.C.; Collier, J.P.; Saum, K.A.; Currier, B.H.; Currier, J.H.; Sanford, M.W.; Mayor, M.B.; Wooding, R.E.; Sperling, D.K.; Williams, I.R.; Kasprzak, D.J.; Surprenant, V.A. Impact of gamma sterilization on clinical performance of polyethylene in the hip. *Clin. Orthop. Relat. Res.* **1995**, *319*, 28–40.
- Ries, M.D.; Weaver, K.; Rose, R.M.; Guthrie, J.; Sauer, W.; Beals, N. Fatigue strength of polyethylene after sterilization by gamma irradiation or ethylene oxide. *Clin. Orthop. Relat. Res.* **1996**, *333*, 87–95.
- Ries, M.D.; Weaver, K.; Beals, N. Safety and efficacy of ethylene oxide sterilized polyethylene in total knee arthroplasty. *Clin. Orthop. Relat. Res.* **1996**, *331*, 159–163.
- Collier, J.P.; Sutula, L.C.; Currier, B.H.; Currier, J.H.; Wooding, R.E.; Williams, I.R.; Farber, K.B.; Mayor, M.B. Overview of polyethylene as a bearing material: Comparison of sterilization methods. *Clin. Orthop. Relat. Res.* **1996**, *333*, 76–86.
- Bargmann, L.S.; Bargmann, B.C.; Collier, J.P.; Currier, B.H.; Mayor, M.B. Current sterilization and packaging methods for polyethylene. *Clin. Orthop. Relat. Res.* **1999**, *369*, 49–58.
- Lewis, G. Properties of crosslinked ultra-high molecular weight polyethylene. *Biomaterials* **2001**, *22*, 371–401.
- McKellop, H.; Shen, F.W.; Lu, B.; Campbell, P.; Salovey, R. Development of an extremely wear-resistant ultra high molecular weight polyethylene for total hip replacements. *J. Orthop. Res.* **1999**, *17*, 157–167.
- Oonishi, H.; Saito, M.; Kadoya, Y. Wear of High-Dose Gamma Irradiated Polyethylene in Total Hip Replacement: Long-Term Radiological Evaluation. *Transactions of the 44th Annual Meeting of the Orthopaedic Research Society*, , 1998; 97
- Oonishi, H.; Kadoya, Y.; Masuda, S. Gamma-irradiated cross-linked polyethylene in total hip replacements—Analysis of retrieved sockets after long-term implantation. *J. Biomed. Mater. Res. (Appl. Biomater.)* **2001**, *58*, 167–171.
- Edidin, A.A.; Pruitt, L.; Jewett, C.W.; Crane, D.J.; Roberts, D.; Kurtz, S.M. Plasticity-induced damage layer is a precursor to wear in radiation-cross-linked UHMWPE acetabular components for total hip replacement. *J. Arthroplast.* **1999**, *14* (5), 616–627.
- Pascaud, R.S.; Evans, W.T.; McCullagh, P.J.J.; FitzPatrick, D.P. Influence of gamma-irradiation sterilization and temperature on the fracture toughness of

- ultra-high molecular weight polyethylene. *Biomaterials* **1997**, *18*, 727–735.
28. Turell, M.; Bellare, A.; Gomoll, A. Fracture Toughness of Radiation Cross-Linked Ultra-high Molecular Weight Polyethylene. Transactions of the 48th Annual Meeting of the Orthopaedic Research Society, Dallas, TX, USA, , 2002; 1034
 29. Cole, J.C.; Lemons, J.E.; Eberhardt, A.W. Gamma irradiation alters fatigue-crack behavior and fracture toughness in 1900H and GUR 1050 UHMWPE. *J. Biomed. Mater. Res. (Appl. Biomater.)* **2002**, *63*, 559–566.
 30. Gsell, R.; Yao, J.Q.; Laurent, M.P.; Crowninshield, R.D. Improved Oxidation Resistance of Highly Cross-linked UHMWPE for Total Knee Arthroplasty. Transactions of the 27th Annual Meeting of the Society for Biomaterials, St. Paul, MN, USA, , 2001; 84

Ultrasound and Tissue Interaction

K. Kirk Shung

Department of Biomedical Engineering, University of Southern California, Los Angeles, California, U.S.A.

INTRODUCTION

Ultrasound has been used as a clinical imaging modality since the late 1950s. The primary form of ultrasonic imaging to date has been that of a pulse-echo mode. The principle is very similar to that of sonar and radar. In essence, following an ultrasonic pulse transmission, echoes from the medium being interrogated are detected and used to form an image. Today, ultrasound capable of both anatomic imaging and blood flow imaging based upon the Doppler principle is the second most utilized diagnostic imaging modality in medicine, next only to conventional X ray, and is a critically important diagnostic tool of any medical facility.

Ultrasound, while best known for its diagnostic capability, has also been used for therapeutic purpose. Significant progress has also been made in ultrasound hyperthermia, frequently in combination with chemotherapy or radiotherapy, for the treatment of tumors, in high-intensity focused tissue ablation, and in ultrasound-mediated drug delivery.

FUNDAMENTALS OF ACOUSTIC PROPAGATION

Ultrasound is characterized by acoustic parameters such as pressure, particle (or medium) velocity, particle displacement, density, and temperature. It is a sound wave with frequencies higher than the audible range, i.e., 20 kHz. If an ultrasound source is applied to a medium, the medium is displaced by the external energy. The distance traveled by the medium is called particle or medium displacement, which is extremely small. The velocity of the medium is called particle or medium velocity and is in the order of a few centimeters per second. The medium velocity is different from the rate at which the energy is propagating through the medium. The velocity at which the ultrasound energy propagates through the medium is defined as the phase velocity or the sound propagation velocity, c . In water, c is approximately 1500 m/s. If the displacement of the medium is in the same direction as the direction of wave propagation, this type of acoustic wave is called a longitudinal or

compressional wave. If the displacement of the medium is perpendicular to the propagation direction, this type is called a shear wave. All wave concepts can be applied to ultrasound. Ultrasound wavelength λ , frequency f , and velocity c , are related by

$$f\lambda = c \quad (1)$$

For a plane wave propagating in a medium in the z -direction, the acoustic parameters vary as a function of time and space governed by the wave equation.

$$\frac{\partial^2 W}{\partial z^2} = \frac{\rho}{v + 2\mu} \frac{\partial^2 W}{\partial t^2} \quad (2)$$

where W , ρ , and t denote, respectively, medium displacement in z -direction, density, and time. The two symbols, v and μ , are Lamé constants and μ is also called shear modulus. The solutions for this second-order differential equation has the form of $f(z \pm ct)$, where the negative sign indicates a wave traveling in the $+z$ -direction, whereas the positive sign indicates a wave traveling in the $-z$ -direction. The displacement W is in the same direction as the wave propagation direction, z . This type of wave is called a compressional or longitudinal wave. The sinusoidal solution for this equation is:

$$W^\mp(z, t) = W_0 e^{j(\omega t \pm kz)} \quad (3)$$

where W^- and W^+ denote displacements for positive- and negative-going waves, respectively, ω = angular frequency = $2\pi f$, $k = \omega/c$ is the wave number, and sound velocity, c , is given by

$$c = \sqrt{\frac{v + 2\mu}{\rho}} \quad (4)$$

For a fluid $\mu = 0$,

$$c = \sqrt{\frac{B}{\rho}} = \sqrt{\frac{1}{G\rho}} \quad (5)$$

It is worth noting from this equation that the sound velocity in a medium is determined by the density and

U

the compressibility of a medium. This means that a measurement of velocity may yield useful information about the acoustic or mechanical properties of a medium. For instance, sound velocity in air is much smaller than in water.

For a case where the displacement in x -direction U is perpendicular to the direction of propagation z , the wave equation becomes:

$$\frac{\partial^2 W}{\partial x^2} = \frac{\rho}{\mu} \frac{\partial^2 W}{\partial t^2} \quad (6)$$

The sinusoidal solution is

$$W^\mp(x, t) = W_0 e^{j(\omega t \pm k_t x)} \quad (7)$$

This type of wave expressed by Eq. (7) is called a shear or transverse wave. The wave number for the shear wave is given by $k_t = \omega/c_t$ where c_t is the shear wave propagation velocity given by

$$c_t = \sqrt{\frac{\mu}{\rho}} \quad (8)$$

It is obvious from Eq. 8 that a shear wave can only exist in a medium with nonzero shear modulus, that is, fluid cannot support the propagation of a shear wave.

Both the longitudinal and shear velocities, as apparent from Eqs. (5) and (8) are affected by the mechanical properties of a tissue. Pathological processes that alternate these properties can cause the sound velocity to change. Therefore, if the velocity can be accurately measured, the result can be used to infer or diagnose the pathology. A few ultrasonic devices in the market today for diagnosing osteoporosis are based upon this principle since osteoporosis causes a loss of bone mass.

CHARACTERISTIC IMPEDANCE

For sinusoidal plane wave acoustic propagation, the characteristic acoustic impedance of a medium is defined as the ration of pressure p to medium velocity u :

$$Z^\pm = \frac{p^\pm}{u^\pm} = \pm \rho c \quad (9)$$

The acoustic impedance has a unit of $\text{kg/m}^2\text{-sec}$ or Rayl. The positive and negative signs are for the positive- and negative-going waves, respectively. The acoustic velocity and impedance for a few common materials and biological tissues are listed in Table 1. The acoustic velocity in a medium is a sensitive function of the temperature, but its dependence on frequency is minimal over the frequency range from 1 to 15 MHz. The acoustic impedance is a very important parameter in ultrasonic imaging since it determines the amplitude of the echoes that are reflected or scattered by tissue components. These echoes are acquired by an imaging device to form an image.^[1-3]

INTENSITY

The instantaneous intensity of an ultrasonic wave is the average power carried by a wave per unit area normal to the direction of propagation over time and is given by

$$i(t) = p(t)u(t)$$

For the case of sinusoidal propagation, the average intensity I can be found by averaging $i(t)$ over a cycle.

$$I = p_0 u_0 \cdot \frac{1}{T} \int_0^T \sin^2 \omega t = \frac{1}{2} p_0 u_0 \quad (10)$$

Table 1 Acoustic properties of biological tissues and relevant materials

Material	Speed m/s at 20–25°C	Acoustic impedance, MRayl	Attenuation coefficient, neper/cm at 1 MHz	Backscattering coefficient, $\text{cm}^{-1} \text{sr}^{-1}$ at 5 MHz
Air	343	0.0004	1.38	–
Water	1480	1.48	0.00025	–
Fat	1450	1.38	0.06	–
Myocardium (perpendicular to fibers)	1550	1.62	0.35	8×10^{-4}
Blood	1550	1.61	0.02	2×10^{-5}
Liver	1570	1.65	0.11	5×10^{-3}
Skull bone	3360	6.00	1.30	–
Aluminum	6420	17.00	0.0021	–

where p_0 and u_0 denote peak pressure and medium velocity, and T is the period.

RADIATION FORCE

An acoustic wave exerts a force on any interface across which there is a decrease in ultrasonic intensity in the direction of wave propagation. For a plane wave of intensity I propagating in a medium with a sound velocity c , the radiation force per unit area or radiation pressure f_r is given by

$$f_r = a(I/c) \quad (11)$$

where a is a constant, depending on the acoustic properties of the target and its geometry.^[4] If a target is made of a material of large acoustic impedance, such as steel, and reflects almost completely the ultrasound beam propagating in water, $a=2$. If the target is a perfect absorber, $a=1$.

The radiation force has been used as a means to perturb an object remotely so that from the displacement of the object, which can be measured via ultrasonic imaging, an assessment of the elastic properties of the medium surrounding the object may be made.^[5]

REFLECTION AND REFRACTION

Reflection and refraction of a plane ultrasonic wave occurs at a flat interface between two media, I and II, of different acoustic impedance. The directions of the reflected and refracted waves are governed, just as in optics, by Snell's law. This is illustrated in Fig. 1 where the subscripts i , r , and t refer to incident, reflected and transmitted, or refracted waves, respectively. As in optics,

$$\theta_i = \theta_r, \text{ and } \sin \theta_i / \sin \theta_t = c_1 / c_2 \quad (12)$$

When $\theta_t = \pi/2$, $\sin \theta_t = 1$, and $\theta_{ic} = \sin^{-1} c_1 / c_2$ if $c_2 > c_1$. For any incident angle greater than θ_{ic} , there is

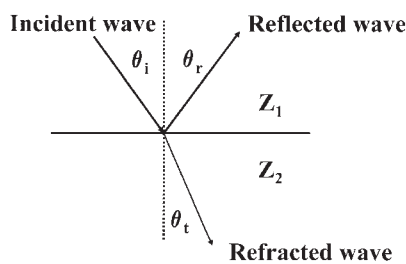


Fig. 1 A plane ultrasonic wave reflected and refracted by a flat interface between two media.

no transmission, that is, total reflection occurs. Therefore, θ_{ic} is called the critical angle.

The pressure reflection and transmission coefficients, R and T , can be easily found by using the boundary conditions that the pressure and particle velocity should be continuous across the boundary.

$$R = \frac{p_r}{p_i} = \frac{Z_2 \cos \theta_i - Z_1 \cos \theta_t}{Z_2 \cos \theta_i + Z_1 \cos \theta_t} \quad (13)$$

$$T = \frac{p_t}{p_i} = \frac{2Z_2 \cos \theta_i}{Z_2 \cos \theta_i + Z_1 \cos \theta_t} \quad (14)$$

As an acoustic wave propagates through an inhomogeneous medium such as biological tissues, part of its energy will be lost due to absorption and scattering, which will be discussed later, and part of its energy will be lost due to specular reflection at the boundary of two adjacent layers of tissues. The ultrasonic images are formed from the specularly reflected echoes due to planar interfaces as well as the diffusely scattered echoes due to small inhomogeneities in tissue parenchyma. Therefore, any change in the elastic properties of the tissues as a result of a disease may be detectable from the ultrasonic image. This has been the principal rationale behind the conventional pulse-echo ultrasonic imaging methods. Since both scattering and reflection depend on the elastic properties of the tissues, which determine the acoustic impedance of the tissue, intuitively, one may postulate that echographic visualizability of tissues is determined mostly by their connective tissue content.^[6] The acoustic impedance of connective tissues and tissues containing high concentrations of connective tissues has been found to be much higher than other types of tissue components such as fat and protein and tissues containing less connective tissues. Because of their higher impedance than surrounding tissues, tissues that contain more connective tissues are more echogenic than those that contain less. However, it must be noted that the attenuation of ultrasound in connective tissues is also much higher. Very little energy can transmit through a mass such as a solid tumor composed of mainly connective tissues. Thus, an acoustic shadow may be created behind such a mass. This has been used as one of the criteria for diagnosing tumors. The origin of the ultrasonic echoes observed in an ultrasonic image for most of the soft tissues has not yet been determined. There are, however, experimental evidences suggesting that, at least in muscle and myocardium, the collagen fibers surrounding the muscle fibers affect, to a great extent, the echogenicity of these tissues.^[7] Although the postulation that tissue echogenicity is largely determined by its connective tissue content may appear to be true, other factors such as cellular dimension and tissue complexity should also play an important role



and cannot be ignored.^[7] There has been a long-standing interest in correlating the acoustic parameters such as scattering, attenuation, and acoustic impedance to the biological composition of the tissues or ultrasonic characterization of biological tissues.^[7,8]

ATTENUATION, ABSORPTION, AND SCATTERING

When an ultrasonic wave propagates through a heterogeneous medium, its energy is reduced or attenuated as a function of distance. The energy may be diverted by reflection or scattering or absorbed by the medium and converted to heat.

The pressure of a plane monochromatic wave propagating in z -direction decreases exponentially as a function of z :

$$p(z) = p(z=0)e^{-\alpha z} \quad (15)$$

where $p(z=0)$ is the pressure at $z=0$ and α is the pressure attenuation coefficient. The attenuation coefficient has a unit of neper/cm and is sometimes expressed in units of dB/cm. Typical values of the attenuation coefficient in some materials are given in Table 1.

The relative importance of absorption and scattering to attenuation of ultrasound in biological tissues is a matter being continuously debated. Investigations to date have shown that scattering contributes little to attenuation, at most a few %, in most soft tissues.^[7] Therefore, it is safe to say that absorption is the dominant mechanism for ultrasonic attenuation in biological tissues.

The energy absorbed by the medium during an ultrasonic propagation is mostly converted to heat. The absorption mechanisms in biological tissues are quite complex and have been assumed to arise from classical absorption due to viscosity and a relaxation phenomenon. Both mechanisms depend on the frequency of the wave. In earlier developments, an ideal fluid with $\mu \sim 0$ has been assumed and this means that the absorption due to the classical viscous loss is ignored. However, in reality, this is seldom the case. Absorption of ultrasound in a lossy fluid has been shown to be proportional to $(\text{frequency})^2$. In media such as air and water, where the dimension of the inhomogeneities of the medium is much smaller than the wavelength, this $(\text{frequency})^2$ dependence of absorption is observed, but in most biological materials, experimental results on attenuation coefficient of ultrasound in biological tissues indicate an approximately linear dependence on frequencies below 15 MHz. As a result, an alternative theory is needed to explain this behavior. It has been theorized that

ultrasonic absorption in biological tissues is dominated by a relaxation process.

When an elemental volume or particle in a medium such as a molecule is pushed to a new position by a force and then released, a finite time is required for the particle to return to its neutral position. This time is called the relaxation time of the particle. For a medium that is composed of the same type of particles, the relaxation time is also the relaxation time of the medium. If the relaxation time is short compared to the period of the wave, its effect on the wave should be small. However, if the relaxation time is comparable to the period of the wave, the particle may not be able to completely return to its neutral state before a second wave arrives. When this occurs, the wave is moving in one direction and the molecules in the other. More energy is thus required to reverse the direction of the particle motion. If the frequency is increased high enough that the molecules simply cannot follow the wave motion, the relaxation effect again becomes negligible. Maximum absorption occurs when the relaxation motion of the particles is completely out of synchronization with the wave motion. Therefore, the relaxation process is characterized by a relaxation frequency where the absorption is maximal and is negligibly small for low-frequency and high-frequency regions illustrated in Fig. 2. Mathematically, the relaxation process can be represented by the following equation:

$$\alpha_r = \frac{Bf^2}{1 + (f/f_R)^2} \quad (16)$$

where α_r = component of the absorption coefficient due to the relaxation process and f_R = relaxation frequency = $1/T_R$, T_R = relaxation time, and B is a constant. In a biological tissue, there are many components, giving rise to many relaxation frequencies. The absorption coefficient in a tissue can be expressed as

$$\alpha_a/f^2 = A + \sum_i \frac{B_i}{1 + (f/f_{Ri})^2} \quad (17)$$

where A is a constant associated with classical absorption and B_i 's and f_{Ri} 's are the relaxation

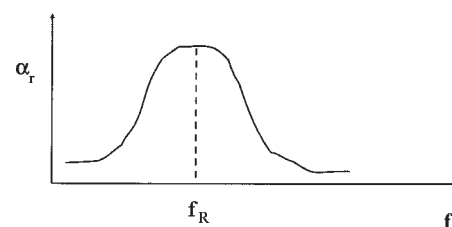


Fig. 2 Ultrasonic absorption caused by a relaxation process characterized by relaxation frequency f_R .

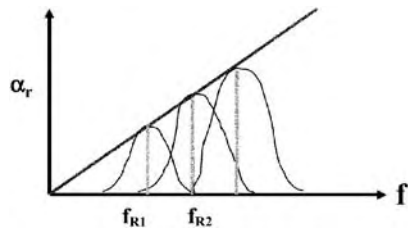


Fig. 3 Ultrasonic absorption caused by multiple relaxation processes characterized by relaxation frequency $f_{R1}, f_{R2} \dots$

constants and frequencies associated with different tissue components.^[8] A possible scenario for this equation is illustrated in Fig. 3, where many relaxation processes may overlap, resulting in a linear increase in the diagnostic ultrasound frequency range or a constant $\alpha\lambda$. Figure 4 shows the absorption of ultrasound in various biological tissues as a function of frequency. As can be seen, the absorption is more or less linearly proportional to frequency ($\alpha\lambda$ is constant) in diagnostic ultrasound frequency range. Also in this figure, it can be seen that the absorption of ultrasound in water is proportional to f^2 .

As a wave is incident on an object, as shown in Fig. 5, part of the wave will be scattered and part of the wave will be absorbed by the object. The scattering characteristics are most conveniently expressed by a term called scattering cross section.

Assume that the incident pressure is a plane monochromatic wave $p_i(r) = e^{-j\mathbf{k}\cdot\mathbf{r}}$ where $\mathbf{k} = k\mathbf{i}$ (\mathbf{i} denotes a unit vector in the incident direction) and \mathbf{r} are vectors representing the wave number and the position. The scattered wave at \mathbf{r}_s due to a scatterer at \mathbf{r}_0 is given by

$$p_s(\mathbf{r}_s) = \frac{e^{-jkR}}{R} p_i(\mathbf{r}_0) f(\mathbf{o}, \mathbf{i}) \quad (18)$$

where \mathbf{o} is a unit vector in the direction of observation, provided that the observation point is in the far field of

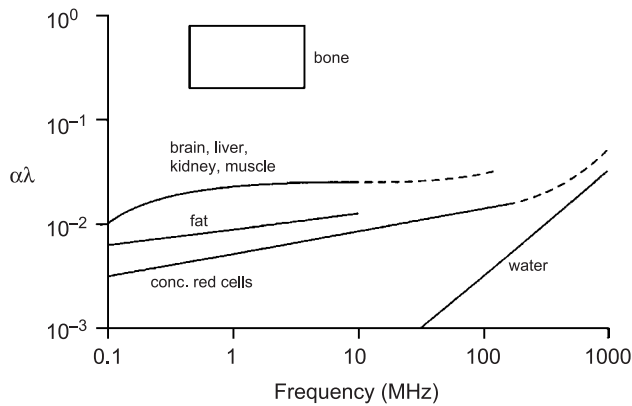


Fig. 4 Ultrasonic absorption in biological tissues as a function of frequency.

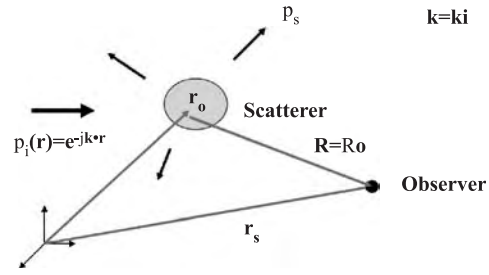


Fig. 5 A plane wave incident upon a scatterer.

the scatterer and $R = |\mathbf{r}_s - \mathbf{r}_0|$ if $kR \gg 1$. The term $f(\mathbf{o}, \mathbf{i})$ in Eq. (18) is called the scattering amplitude function, which describes the scattering properties of the object and depends on the directions of incidence and observation. By utilizing the definition of intensity, it can be shown that

$$\sigma_d(\mathbf{o}, \mathbf{i}) = |f(\mathbf{o}, \mathbf{i})|^2 = \frac{I_s R^2}{I_i} \quad (19)$$

where the term $\sigma_d(\mathbf{o}, \mathbf{i}) = |f|^2$ is defined as the differential scattering cross section, which is the power scattered in \mathbf{o} direction with the incident direction \mathbf{i} in one solid angle per unit incident intensity. When $\mathbf{o} = -\mathbf{i}$, $\sigma_d(\mathbf{i}, -\mathbf{i})$ is called the backscattering cross section. By integrating σ_d over the 4π solid angle, the scattering cross section σ_s (defined as the power scattered by the object per unit incident intensity) can be found.

$$\sigma_s = \int_{4\pi} \sigma_d d\Omega = \int_{4\pi} |f(\mathbf{o}, \mathbf{i})|^2 d\Omega$$

where $d\Omega$ is the differential solid angle. Similarly, an absorption cross section, σ_a , can be defined as the total power absorbed by the object. Thus, the attenuation in wave intensity due to the presence of the object is simply

$$2\alpha = \sigma_a + \sigma_s$$

where 2α is the intensity attenuation coefficient. If there are a number of objects, the intensity attenuation coefficient should be:

$$2\alpha = n(\sigma_a + \sigma_s)$$

where n is the object concentration per unit volume. This relation is valid only if n is small or volume concentration given by $n \cdot (\text{object volume}) < 1\%$. As n increases, multiple scattering occurs and there may be particle-particle interactions,^[7] and this relation no longer holds.

It is impossible to solve analytically for the scattering cross section of an object of arbitrary shape, although there are computer-based methods,



such as finite element analysis, which may be used. However, a number of approximations exist, which can simplify the problem considerably. One of these is the Born Approximation, which assumes that the wave inside the object is the same as the incident wave.^[9] This is a valid assumption if the size of the object is much smaller than the wavelength or if the acoustic properties of the scatterer are similar to those of the surrounding medium. By applying this approximation and using the wave equation, the scattering cross section of an object can be found. The scattering cross section for a sphere whose radius is much smaller than the wavelength is given by:^[9]

$$\sigma_s = \frac{4\pi k^4 a^6}{9} \left(\left| \frac{G_e - G}{G} \right|^2 + \frac{1}{3} \left| \frac{3\rho_e - 3\rho}{2\rho_e + \rho} \right|^2 \right) \quad (20)$$

where k is the wave number, a is the radius of the sphere, G_e and G are the adiabatic compressibilities of the particle and the surrounding medium, and ρ_e and ρ are the corresponding mass densities.

In a dense distribution of scatterers like biological tissues, for instance, human blood consisting of $5 \cdot 10^9$ red blood cells in one cubic centimeter, scatterer-to-scatterer interaction cannot be ignored. To take this into consideration, a parameter called backscattering coefficient in a unit of $(\text{cm-steradian})^{-1}$ is often used. It is also sometimes called volumetric backscattering cross section and is defined as the power backscattered by a unit volume of scatterers in one solid angle per unit incident intensity. The physical meaning of the backscattering coefficient can be easily understood simply by replacing the scatterer in Fig. 5 by a unit-scattering volume. Because the tissues are fairly inhomogeneous, the scattered signals acquired may vary greatly at any single frequency and over a band of frequencies, a parameter that is frequently used to describe scattering behavior of a tissue is the integrated backscatter. The integrated backscatter (IB) is defined as the frequency average of the backscatter over the bandwidth of the signal, which mathematically can be expressed as

$$IB = \frac{1}{2\Delta\omega} \int_{\omega_0 - \Delta\omega}^{\omega_0 + \Delta\omega} e(\omega) d\omega$$

where $2\Delta\omega$ is the bandwidth, ω_0 is the center frequency of the spectrum, and $e(\omega)$ is the backscattered signal at angular frequency ω . However, in order to eliminate the dependence of the backscattered signal from tissues on the electrical and acoustic characteristics of the experimental system, the IB is usually expressed in dB by comparing the backscattered signal to a reference signal, e.g., the echo from a flat reflector.

Therefore,

$$IB(\text{dB}) = 20 \log \left\{ \frac{1}{2\Delta\omega} \left[\int_{\omega_0 - \Delta\omega}^{\omega_0 + \Delta\omega} e(\omega) d\omega \right] \div \int_{\omega_0 - \Delta\omega}^{\omega_0 + \Delta\omega} e_r(\omega) d\omega \right\}$$

where $e_r(\omega)$ is the reflected signal from the flat reflector at angular frequency ω . The backscattering coefficient for several tissues is listed in Table 1.

The fact that the acoustic scattering characteristics of an object including angular scattering pattern depend on the shape, size, and acoustic properties of the scatterer has been known for many years. Ideally, the structure and acoustic properties of the scatterers can be deduced from measuring their scattering properties. This problem is of interest to many scientists in such diverse fields as geophysics, oceanography, and communication. It is generally termed as remote sensing or detection. Although the potential of characterizing a tissue structure from its scattering properties was realized in the biomedical ultrasound community almost three decades ago, this field still remains in its infancy, primarily due to the complex nature of biological tissues. Preliminary experimental investigations in vitro so far have shown that different tissues exhibit different angular scattering pattern and frequency dependence.^[7,8] The backscattering coefficient (defined as backscattering cross section per unit volume) of scatterers for five different types of bovine tissues as a function of frequency is shown in Fig. 6.^[7] Experimental results on scattering by red blood cells^[7] are in good agreement with Eq. 20. Scattering from myocardium has been shown to have a third power dependence on frequency.^[7] Based on this evidence and more recent results indicating that ultrasonic scattering increases as the muscle is stretched, it was suggested that the scattering from muscle may be attributed to the collagen fibers in the muscle because the scattering from cylinders whose radius is much smaller than the wavelength is proportional to f^3 . Experimental results on scattering from tissues thus far have been very scanty and inconclusive. The precise origin of ultrasonic scattering from many tissues is still unknown although there are efforts underway to shed more light on this subject.

Because pathological processes in tissues involve anatomical variations, it is likely that they will result in corresponding changes in ultrasonic backscatter. This has been demonstrated by several recent investigations. Figure 7 shows that the backscattering coefficient for regions of infarcted myocardium is substantially higher than that for normal myocardium.^[10] In addition, myocardial ultrasonic backscatter is shown to be related to the contractural state of the tissue (Fig. 8).

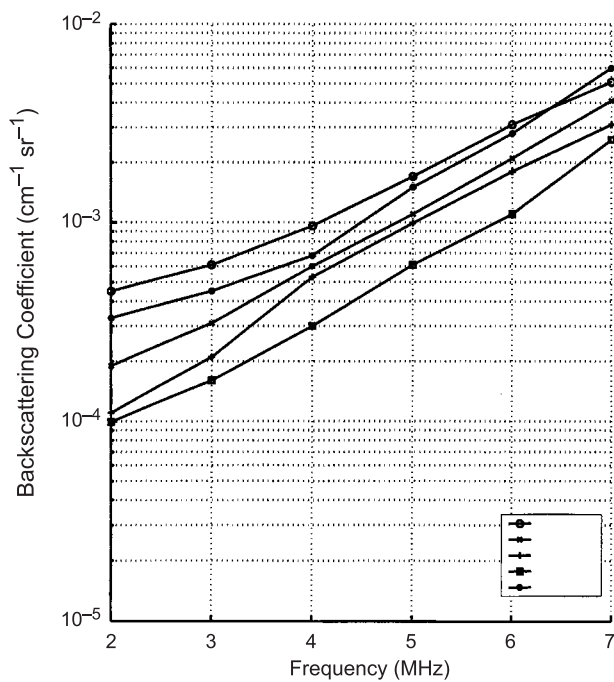


Fig. 6 Ultrasonic backscatter coefficient for several bovine soft tissues as a function of frequency.

The myocardial backscatter was found to be the highest at the end diastole and lowest at the end systole during a cardiac cycle and this cyclic behavior was blunted in ischemic heart.^[10]

In summary, the attenuation of ultrasound in biological tissues can be attributed to two mechanisms: scattering and absorption. The question remains, however, as to the relative importance of these mechanisms although absorption is believed to be the dominant mechanism. Attenuation in general is not desirable because it limits the depth of penetration ultrasound into the body. However, it may yield useful information for diagnostic purposes because it carries information about the properties of the tissues if it can be accurately estimated. Extensive work has also been carried out to demonstrate that tissue pathology can affect, to a great extent, the attenuation coefficient of tissues. Infarcted myocardium and malignant tumors in liver and breast were all shown to have an increased attenuation. Unfortunately, measuring attenuation in vivo has been proven to be quite a difficult task. Various schemes have been developed with little success.

NONLINEARITY EFFECTS AND PARAMETER B/A

Because the spatial peak temporal peak intensity I_{SPTP} of modern ultrasonic diagnostic instruments can sometimes reach the level of more than 100 W/cm^2 ,

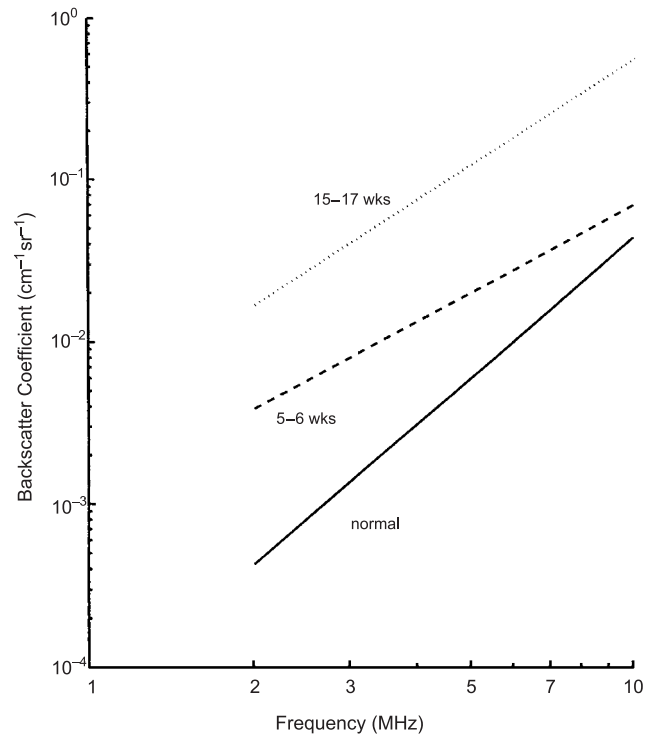


Fig. 7 Ultrasonic backscatter coefficient for canine myocardium as a function of frequency. Solid, dashed, and dotted lines represent data for normal myocardium, myocardium 5–6 weeks following infarction, and myocardium 15–17 weeks following infarction.

nonlinear acoustic phenomena may not be ignored in treating ultrasonic propagation in tissues.^[11] In fact, it has been shown that in many instances harmonic imaging that utilizes the harmonics generated due to the nonlinear propagation yields better images than conventional ultrasonic imaging.^[12] Other reasons for studying the nonlinear properties are: new tissue parameters may be derived for tissue characterization and nonlinearity can influence, to a significant extent, how the ultrasound energy is absorbed by the tissue.

The nonlinear behavior of a fluid medium can be expressed by a second-order parameter B/A . For an adiabatic process in which the entropy is constant or there is no energy flow, the relation between pressure and density can be expressed as a Taylor series expansion of pressure, p , about the point of equilibrium density, ρ_0 , and entropy, s_0 ,

$$p = p(s_0, \rho_0) + A \left(\frac{\rho - \rho_0}{\rho_0} \right) + \frac{1}{2} B \left(\frac{\rho - \rho_0}{\rho_0} \right)^2 + \dots$$

where A and B are constants and the ratio of B/A is given by

$$B/A = 2\rho_0 c_0 \left[\frac{\partial c(s_0, \rho_0)}{\partial p} \right] \tag{21}$$



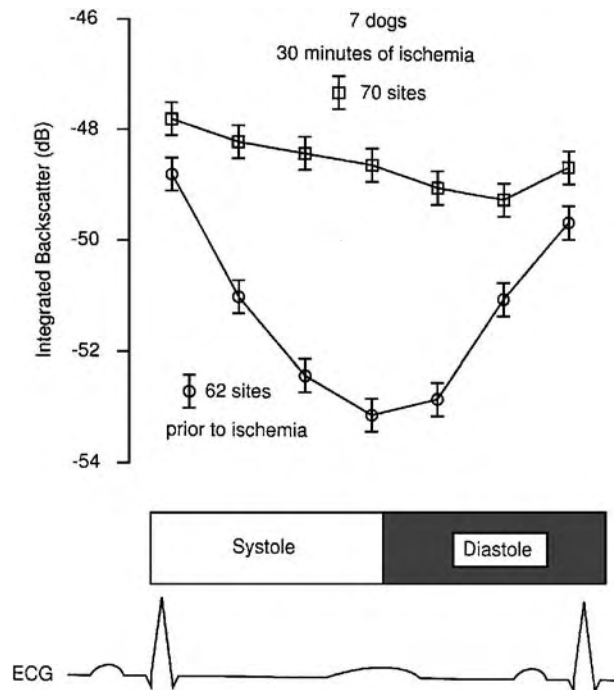


Fig. 8 Integrated backscatter from canine myocardium as a function of the cardiac cycle.

This B/A parameter is a primary indicator as to the nonlinearity of a medium and can be estimated from these quantities, which are either known or measurable. Experimental findings show that: 1) B/A is linearly proportional to solute concentration in aqueous solutions of proteins; 2) B/A is insensitive to the molecular weight of the solute at fixed concentrations; 3) B/A ranges from six to 11 for soft tissues; and 4) B/A may be dependent upon tissue structure. The B/A values for a few tissues and relevant materials are given in Table 2.

The nonlinearity of ultrasonic wave propagation can cause the distortion of waveforms as illustrated in Fig. 9, which shows that a sinusoidal waveform becomes a sawtooth waveform as a result of the generation of higher harmonics after propagating in a medium. This is plausible in that the velocity in a denser region of the medium should be greater. It is

Table 2 B/A values for various materials

Material	B/A
Water	5.0 at 20°C
Bovine serum albumin (20 g/100 ml)	6.2 at 25°C
Beef liver	7.8 at 23°C
Human breast fat	9.2 at 22°C
Porcine muscle	6.5 at 25°C
Porcine whole blood	6.2 at 30°C

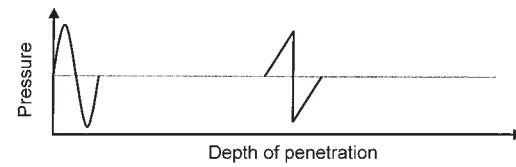


Fig. 9 Wave distortion resulted from the nonlinear interaction between ultrasound and the medium.

known that the amplitude of the fundamental frequency drops and the amplitudes of higher harmonics increase as the propagation distance increases.^[13]

In addition to heating, high-amplitude ultrasound can also induce cavitation. The term cavitation is used to describe the behavior of gas bubbles in ultrasonic fields. Two different types of cavitation may occur: transient and stable cavitation. Transient cavitation describes the phenomenon in which microbubbles suddenly grow and collapse in a liquid medium. The physical process can be described as follows: Bubbles in a medium are greatly expanded when pressure decreases rapidly. The pressure increases one-half cycle later, causing bubbles to collapse and disappear. For a very large pressure swing, the radius increases markedly, reaching a peak well past the pressure minimum, and as the pressure reaches a peak, the bubble collapses. The internal bubble pressure can become very high—up to 80,000 ATM—with a temperature approaching 10,000°K. Such high temperatures can cause decomposition of water into chemically-active acidic components, causing serious biological effects. A phenomenon called sonoluminescence, in which flashes of light with a duration less than a few picoseconds are generated, may accompany the collapsing of the bubbles.^[14] Stable cavitation, on the other hand, describes a phenomenon in which the bubbles do not collapse as just described. Under this circumstance, the behavior of such bubbles is quite stable and is known as stable cavitation. Stable cavitation is more likely to occur at lower ultrasound intensities.

Cavitation can be suppressed by degassing the liquid, increasing the viscosity of the surrounding medium, or by increasing the external pressure applied to the system. It takes a finite amount of time for the gas bubble to respond to the pressure change. Therefore, cavitation is a frequency-dependent phenomenon. If the ultrasound frequency is high enough, it should disappear.

As ultrasound propagates in a fluid, transfer of momentum to the medium via absorption causes acoustic streaming in the direction of the sound beam. If a discrete object is present in the ultrasound beam, a radiation force is exerted on the object, as discussed earlier. As the object moves, streaming of the fluid near the object may occur. If the object is an air bubble, the

oscillation of the bubble can also cause streaming of the fluid. Because acoustic streaming is related to absorption, nonlinear interaction between ultrasound and a medium may increase the acoustic streaming manifold. Acoustic streaming may induce shear stress on interfaces that border the fluid and the object. These nonlinear ultrasound effects have been used for hyperthermia for the treatment of cancer, high-intensity focused surgery,^[15] and gene transfection and drug delivery.^[3,16]

ARTICLES OF FURTHER INTEREST

Intravascular Ultrasound; Medical Imaging, 3-D; Ultrasound Doppler

REFERENCES

1. Goldberg, B.B.; Kimmelman, B.A. *Medical Diagnostic Ultrasound: A Retrospective on its 40th Anniversary*; AIUM: Laurel, MD, 1988.
2. Shung, K.K.; Smith, M.B.; Tsui, B.W.N. *Principles of Medical Imaging*; Academic Press: San Diego, 1992.
3. Miller, M.M. Gene transfection and drug delivery. *Ultrasound Med. Biol.* **2000**, *26*, S59–S62.
4. Kossoff, G. Radiation Force. In *Interaction of Ultrasound and Biological Tissues*; Reid, J.M., Sikov, M.R., Eds.; FDA: Rockville, MD, 1972; 159–161.
5. Nightingale, K.R.; Nightingale, R.W.; Palmeri, M.L.; Trahey, G.E. Finite Element Analysis of Radiation Force Induced Tissue Motion with Experimental Validation. In *Proc 1999 IEEE Ultrasonics Symp.*; Schneider, S.C., Levy, M., McVoy, B.R., Eds.; IEEE: New York, 1999; 1319–1323.
6. Fields, S.; Dunn, F. Correlation of echographic visualization of tissue with biological composition and physiological state. *J. Acoust. Soc. Am.* **1972**, *54*, 809–811.
7. Shung, K.K.; Thieme, G.A. *Ultrasonic Scattering by Biological Tissues*; CRC Press: Boca Raton, FL, 1993.
8. Greenleaf, J.A. *Tissue Characterization with Ultrasound*; CRC Press: Boca Raton, FL, 1986.
9. Morse, P.M.; Ingard, K.U. *Theoretical Acoustics*; McGraw Hill: New York, 1968.
10. Wickline, S.A.; Perez, J.E.; Miller, J.G. Cardiovascular Tissue Characterization in vivo. In *Ultrasonic Scattering in Biological Tissues*; Shung, K.K., Thieme, G.A., Eds.; CRC Press: Boca Raton, FL, 1993; 313–345.
11. Hamilton, M.F.; Blackstock, D.T. *Nonlinear Acoustics*; Academic Press: San Diego, 1998.
12. Tranquart, F.; Grenier, N.; Eder, V.; Pourcelot, L. Clinical use of ultrasound tissue harmonic imaging. *Ultrasound Med. Biol.* **1999**, *25*, 889–894.
13. Pierce, A. *Acoustics: An Introduction to Physical Principles and Applications*; McGraw Hill: New York, 1986.
14. Lighton, T.G. *The Acoustic Bubble*; Academic Press: London, 1994.
15. ter Harr, G.R. Therapeutic ultrasound—An overview. *Europ. J. Ultrasound* **1999**, *9*, 3–9.
16. Mitragotri, S.; Blankschtein, D.; Langer, R. Ultrasound-mediated transdermal protein delivery. *Science* **1995**, *269*, 850–853.



Ultrasound Doppler

Ding-Yu Fei

Department of Biomedical Engineering, Virginia Commonwealth University, Richmond, Virginia, U.S.A.

James A. Arrowood

Department of Internal Medicine, Virginia Commonwealth University, Richmond, Virginia, U.S.A.

INTRODUCTION

The Doppler effect is a physical phenomenon existing with all waves such as electromagnetic and sound waves. The effect is a change in frequency of a wave perceived by an observer, when the wave source, the observer, or both are moving with respect to the other. Many animals such as bats and dolphins use the Doppler effect to detect moving targets, enabling them to find food. The Doppler effect was discovered by the Austrian scientist Johann Christian Doppler in 1842. He derived equations relating the change of the perceived frequency to the relative speed between the wave source and the observer. Since then, the Doppler principle has been used with microwaves, light, and sound waves including ultrasound. Ultrasound is a portion of the sound spectrum with frequency higher than 20 KHz, which is beyond human hearing ability. Ultrasound Doppler has been widely used in industries as well as daily life. For example, sonar systems on ships use underwater ultrasound Doppler to detect and measure movement of fishes and submarines. Home burglar alarms use airborne ultrasound Doppler to detect moving targets in the yard and house.

In biomedical fields, ultrasound Doppler is employed to identify blood flow, measure blood flow velocity in vessels, and monitor the movement of tissues and organs. Several types of ultrasound Doppler instruments and technologies are available for a variety of medical applications. From the perspective of the ultrasonic signals transmitted, there are two types of Doppler instruments: continuous wave (CW) Doppler and pulsed wave (PW) Doppler. Ultrasound Doppler technologies are frequently used in combination with ultrasonic imaging systems or scanners to visualize the site of measurement in order to provide more useful and accurate information. Currently, the major combinations include duplex scanning, Doppler color imaging (DCI), power Doppler imaging (PDI), and tissue Doppler imaging (TDI).

Using the different technologies, the measurement of ultrasound Doppler can be carried out within a single or a few sampling volumes along a specific

ultrasound beam. The measured results may be heard with loudspeakers and/or displayed on the screen of a monitor in real time. Blood flow velocities or tissue movement may be obtained within a relatively large area by using a multiple gate procedure. In this case, color images are usually used to display the results. Signal and image processing technologies are used to extract more useful information from the detected Doppler signals. The results obtained are widely used for evaluating normal hemodynamic function and detecting abnormalities in most large- and middle-sized vessels in humans. By measuring blood flow in small vessels, ultrasound Doppler is also used to detect abnormal tissues and tumors in some organs.

DOPPLER PRINCIPLE

As mentioned previously, the Doppler effect results in a change in frequency of waves perceived by an observer under three possible cases: the source is moving, the observer is moving, or both are moving with respect to the other. When a sound wave travels in a medium, the general relationship among the frequency, f , sound speed, c , and wavelength, λ , can be expressed as

$$f = c/\lambda \quad (1)$$

Assume that a signal source emits a sound wave with a emitted frequency, f_e . An observer at a distance away from the source perceives the same frequency emitted from the source when both the source and observer are stationary. When the wave source is moving toward the observer with a velocity, v , as the wave approaches, the peaks of the sound wave are compressed toward the observer. The intervals between wave peaks or the effective wavelength is reduced to $\lambda - vT$, where T is the period of the wave, i.e., $1/f_e$. The reduction in wavelength translates into an increase in the observer perceived frequency, f_p . The difference between the perceived and the emitted frequencies is called the Doppler frequency or Doppler

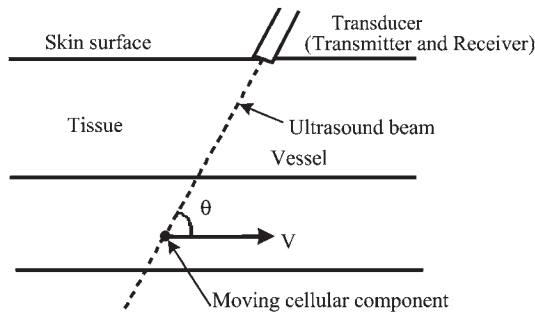


Fig. 1 Typical case using ultrasound Doppler instrument to measure blood flow velocity in a vessel.

shift, f_d , and can be expressed as

$$f_d = f_p - f_e \quad (2)$$

If the source is stationary and the observer is moving toward the source, the effective wave speed will increase to $c + v$. The increase of wave speed also translates into an increase of perceived frequency. In both of the preceding cases, the Doppler frequency is a positive value. Conversely, if the source and observer are moving away from each other, the perceived frequency will be lower than the emitted frequency and the Doppler frequency will be negative.

In the case that both the source and observer are moving toward (or away from) each other at the same velocity, v , if the velocity v is much less than the sound speed in the medium, the Doppler frequency for the combining case can be expressed as:^[1,2]

$$f_d = 2vf_e/c \quad (3)$$

where c is the sound speed in the medium.

Figure 1 shows the most popular case of ultrasound Doppler used in medical applications. The transducer consisting of the transmitter and receiver, which may be one or two separate element(s), as discussed in later sections, contacts the skin surface and the transmitted ultrasound beam passes across a vessel in the body. When the ultrasound signals arrive at the vessel, the moving cellular components of the blood pool, consisting mainly of red blood cells, will scatter the

signal in all directions. The signal scattered back (called backscattered signal) to the transducer will be detected by the receiver and the Doppler frequency will then be calculated. Sound speed varies when it travels in different medium. In humans, because the major component of all soft tissues is water, sound speeds in different soft tissues are very nearly the same.^[1] It is reasonable to assume that the sound speed travels in all soft tissues at an average speed of 1540 m/sec. On the other hand, blood velocities observed in the human body are all below a few m/sec, which is much less than the average sound speed in soft tissues. Eq. 3, therefore, is valid for this case. For the situation where the angle between the flow and beam directions is θ , the Doppler frequency perceived reflects the component of the velocity along the beam direction and Eq. 3 is modified to:

$$f_d = 2v \cos \theta f_e/c \quad (4)$$

This is the basic equation usually used for velocity determination in ultrasound Doppler systems. It can be seen from Eq. 4 that if the Doppler frequency is detected and the sound speed is known, the component of the velocity along the beam direction, i.e., $v \cos \theta$, can be calculated. If the Doppler angle, θ , is also known, the flow velocity can be determined. Normally, the range of emitted frequencies used for most medical ultrasound applications is from 2 to 8 MHz. The range of blood velocities in large- and medium-sized vessels is from several cm/sec to a few m/sec. To obtain some idea about the range of Doppler frequencies detected, Table 1 shows a few examples of the perceived Doppler frequencies used in medical ultrasound Doppler devices. In most medical applications, Doppler frequencies are within the audible sound range (below 20 KHz).

SPECIFIC TYPES OF ULTRASOUND DOPPLER

Continuous Wave (CW) Doppler

CW Doppler is one of the earliest modes of ultrasound Doppler technologies. The transducer of the device is

Table 1 Samples of ultrasound Doppler frequency in medical applications^a

Emitted frequency f_e (MHz)	Flow velocity c (cm/sec)	Doppler angle θ (degree)	Doppler frequency f_d (Hz)
3.5	100	0	4,545.5
		60	2,272.8
		70	1,554.7
2.5	50	0	2,272.8
		0	1,623.4
5.0	50	0	3,246.8

^aThe sound speed is assumed to be 1540 m/sec, which is the average of sound speeds in soft tissues.



made of two piezoelectric elements. One serves as the transmitter and emits an ultrasound signal continuously. The other serves as the receiver to detect the reflected or scattered sound signals. Because the transmitting and receiving ultrasound beams may not be the same, the detectable echo signals are the signals coming from the overlapping volume of the transmitting and receiving beams.^[2,3] Normally, the frequency, amplitude, and phase information of the echoes are processed. By comparing with the emitted frequency, the Doppler frequencies and corresponding amplitudes can be determined. A technique known as quadrature phase detection^[3] is often used to determine whether the flow direction is toward or away from the transducer. As mentioned previously, the detected Doppler frequencies usually fall in the audible sound range. In some compact devices such as simple fetal heart monitors, only loudspeakers are used to listen to the Doppler signals. Some devices may use simple methods such as zero-crossing technique to detect the frequencies. Many other devices use fast Fourier transformation (FFT)^[4] to determine Doppler frequency spectrum of the echo signals and display them on the monitor.

CW Doppler can provide velocity information on blood flow and tissue or organ movement. In the human body, Doppler signals from blood flow in vessels have much higher frequencies, but lower amplitudes compared with those from tissue movement. With CW Doppler as well as all other Doppler technologies, if the Doppler mode is designed for measurement of blood flow, a wall filter is usually used to remove the large amplitude but low-frequency Doppler signals generated by vessel wall motion. On the other hand, if the device is designed to measure tissue movement, no filter or a filter removing the high Doppler frequencies generated by the blood flow may be used.

A major advantage of CW Doppler over PW Doppler (see description later in the article) is that CW Doppler has no limitation on the maximum measurable velocity as PW Doppler systems have. The large sample volume in CW Doppler allows detection of the maximum velocity along the entire ultrasound beam. However, because the transmitter emits the ultrasonic signal continuously, the echo coming back from targets at one distance may be mixed with echoes coming back from targets at deeper positions. No distance information is included in the measured results. CW Doppler systems may also give complicated and confusing presentations if two or more different flow regions fall within the path of the ultrasound beam. For a stand-alone CW Doppler system, because no vessel visualization is provided from the instrument, the user must know the basic anatomic structure in order to detect flow in the vessel of interest.

Pulsed Wave (PW) Doppler

A PW Doppler system periodically emits short duration ultrasonic signals and then detects the reflected sound signals in the time interval between two consecutive emitted signals. Because emission and reception of the signals occur at different times, only one piezoelectric element is usually used to serve as both the transmitter and receiver. Each of the short-duration signals includes several cycles at the emitted frequency and is called a burst or pulse train. The frequency of the emitted bursts is called the pulse repetition frequency, or PRF. Using the time of the emitted signal as a reference, the time interval between the emitted and the reflected signals represents the round trip distance along the beam between the transducer and the reflector. This time interval can be used to determine the distance between two reflectors at different depths. By listening for reflected signals during a particular time interval (gate in time), PW Doppler devices can acquire reflected signals within a sample volume at a particular depth or distance from the transducer. This allows measurement of Doppler shifts (hence velocities) within a particular sample volume in space. This process is called range gating. FFT is usually used to determine the Doppler spectrum of returning echoes within the sample volume. The spectrum can be displayed on the screen in real time and the timing of the gate provides the distance information of the reflectors. For a stand-alone PW Doppler device, no image display is available. Therefore, the user must know the anatomy in order to measure flow in the vessel or location of interest.

Aliasing phenomenon is a limiting factor in PW Doppler instruments. Because the echo signals from a specified sample volume are not available continuously, PW Doppler has to determine the Doppler frequencies from timed samples of the echo signal. The sampling rate is equal to PRF of the device. According to the sampling theory,^[4,5] if the signal frequency is higher than half of the sampling rate, aliasing occurs. The calculated flow direction will be reversed and the velocity value would also be incorrect. Therefore, the highest Doppler frequency that can be detected by PW Doppler without aliasing equals one-half of the PRF and is called the Nyquist limit.

In both stand-alone CW and PW Doppler devices, information on the Doppler angle is usually not available because the vessel direction inside the body is unknown, and therefore, the actual velocities cannot be determined.

Duplex Scanning

Duplex scanning combines ultrasonic imaging and Doppler technologies to provide quantitative flow

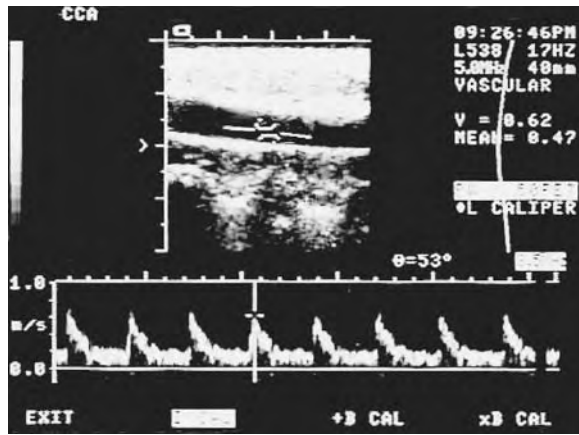


Fig. 2 Duplex ultrasound image of common carotid artery. Position of pulsed Doppler cursor in 2-D image is seen in upper part of figure. Blood flow velocity spectral pattern is seen in lower part of figure.

information in vessels. A transducer consists of one set of piezoelectric elements for two-dimensional, gray-scale Brightness or B mode imaging, and another set of elements for ultrasound Doppler, although some of the elements may be used for both B mode and Doppler. Either PW or CW Doppler can be used. From the B mode image, the user can move the transducer to find the vessel of interest and locate the ultrasound Doppler beam to the desired portion of the vessel. Figure 2 shows a typical image of the duplex scanning with PW Doppler for a common carotid artery (CCA). The acquired B mode image is shown at the top of the image. In the image, the relatively straight, dark area across the image from right to left is the CCA lumen. The vessel wall and surrounding tissues can also be seen. The position and direction of the PW Doppler beam are displayed as the dotted line in the image. There are two short, parallel lines along the Doppler beam, which represent the gate position and gate length. There is a flow angle indicating line (the broken line) between the two short lines. The orientation of this line can be adjusted around 360° and is used to indicate the flow velocity direction for angle correction. In a relatively straight segment of a vessel, it is reasonable to assume that the flow directions inside the vessel are parallel with the direction of vessel wall. When the user adjusts the angle indicating line parallel with the direction of vessel (as shown in the image), the Doppler angle can be determined by the scanner and displayed in the image (53° in this image). When the Doppler angle is known, the velocity can be calculated from the detected Doppler frequency with Eq. 4. In duplex scanning, the detected Doppler frequency spectrum is normally converted to velocity spectrum and displayed at the lower portion of the screen in real time. When the spectrum image is frozen as shown, further analysis of the image can be carried out. Using

duplex imaging, the velocity spectra at different vessels, locations, and sample volumes can be observed and analyzed.

Because duplex scanning provides additional anatomic information, the measurement site and the sampling volume can be adjusted. Analyzing velocity spectra at specified sites provides valuable information for patient monitoring and for diagnosis of a variety of cardiovascular diseases. The average velocity across a vessel obtained from duplex scanning may also be used for the estimation of blood flow rate.

Doppler Color Imaging (DCI)

DCI or Doppler color flow imaging (DCFI) uses imaging and PW Doppler technologies to acquire flow information within a relatively large area in a flow field. A same set of piezoelectric elements may be used for B mode imaging and PW Doppler flow acquisition. When acquiring flow information, the ultrasonic beam is scanned to form a linear or sector pattern. For each scan line, multiple gates are used to detect the echoes. Each gate corresponds to a small sample volume, and velocities along the scan line can be detected simultaneously. By scanning the ultrasonic beam, the velocity information within a relatively large area is obtained.

In the process of flow velocity estimation, if FFT technology is used, as the regular PW Doppler devices do, a relatively large number (perhaps up to 128) of pulse trains are required for each scan line. Because there will also be a large number of scan lines, a rather long time is needed to complete scanning one frame of the image. This will not be a realistic method for real-time acquisition of flow information. In the current color imaging systems, an autocorrelation technique or an instantaneous frequency estimation technique or combination of these two are used to make fast determination of Doppler shifts.^[6] Using these techniques, as few as three pulse trains are sufficient to determine the mean velocity at each sample volume although larger numbers of pulse trains may provide more accurate results. This makes real-time DCI possible. The mean velocities within a selected area are obtained, color coded, and superimposed on the gray-scale image to form the Doppler color image.

Figure 3 shows a typical Doppler color image of the carotid bifurcation at peak systole. A linear transducer is used. Most ultrasound scanners provide three beam directions for linear transducers. As shown in Fig. 3, the beam direction for the color imaging is toward the right from the transducer. The Doppler beam direction may be adjusted toward the left or central (vertical to the transducer). In the image, the main vessel on the right is the common carotid artery (CCA). The



Fig. 3 Doppler color image of common carotid bifurcation at peak systole using a linear array transducer.

branches on the left are internal carotid artery (ICA) (lower branch) and external carotid artery (ECA) (higher branch). Two color bars to the left of the image are the reference color scale. The bar at the top has red (lower) to yellow (higher) color to represent flow toward the transducer. The lower bar has blue (higher location represents lower velocity) to green (lower location, higher velocity) color to represent flow away from the transducer. By comparing the color in the vessel to the color scale, the measured mean flow velocity at each point can be estimated. In Fig. 3, the color at most portions of the vessel lumen is blue with higher velocity along the inner wall of the ICA. This shows that the blood flows mainly from right to left. The color in a small area at the lower and expanded portion (bulb) is red. It is a reversed flow area in the ICA, which is a normal flow pattern for most healthy persons.

In addition to normal velocity display mode, DCI systems can determine the variance of velocities within each sample volume. The results can be displayed with a variance color scale. High variance represents flow turbulence existing in the corresponding sampling location and can also be used for diagnostic purposes.

The strength of DCI lies in the fact that it can quickly provide a global view of flow in real time superimposed on anatomic structures, thus revealing the direct relationship between flow and anatomy instantaneously. It enhances the chances of determining abnormal flow patterns in a particular area and facilitates comparison of flow in different anatomic regions. Thus, DCI provides flow information faster and from a larger area, thus allowing easier detection of abnormal flow patterns.

Several limiting factors need to be considered when using DCI for flow field observation.^[6] First, the velocity obtained in DCI is the mean velocity within each sample volume. No velocity spectrum can be

determined. Second, the mean velocity is actually the component of the true velocity along the beam direction and therefore, images obtained in DCI are angle-dependent. As seen in Fig. 3, if one changes the ultrasound beam direction to the left, the components of the velocities in the flow fields along the beam direction will all change, and a flow image with totally different color will result. Generally, since blood vessels frequently exhibit curvature and branching in a relatively large flow field, the flow direction may vary significantly within the imaging region and therefore, even though the beam direction is known in linear transducer, one cannot simply use one predicted flow direction as the basis for angle correction over the entire region. If a sector transducer is used, the beam direction continuously varies from line to line and the angle dependence will be a more serious problem. The angle dependence is an important barrier against further exploiting the quantitative information and expanding the clinical applications of the DCI technique. Currently, methods based on multiple beam directions have been developed for reconstruction of angle-independent DCI.^[7-9]

Aliasing is also a limiting factor in DCI. If the color scale selected is too small, Doppler frequencies from a portion of the flow field may be higher than the Nyquist limit, and incorrect color representation of the flow field may result.

Power Doppler Imaging (PDI)

PDI can be considered as a modified mode of DCI.^[5] Similar to DCI, PDI uses PW Doppler and a multiple-gate procedure to collect flow information. Instead of finding the Doppler frequency, PDI detects the Doppler amplitude or power within each sample volume. The results are also color coded and superimposed on the gray-scale image. In other words, DCI is a frequency-coded color imaging and PDI is an amplitude- or power-coded color imaging.

Power Doppler imaging is very sensitive to the Doppler shift and can easily detect flow in small vessels such as blood in the liver and kidney. Power Doppler imaging does not need to calculate flow velocity and therefore, Doppler angle plays much less of a role in the imaging process. DPI is basically angle-independent except at a Doppler angle approaching 90° . At that angle, the Doppler frequency equals zero and no blood flow can be detected by either PDI or DCI. Also, because no velocity is calculated, aliasing will play no role in the detection of blood flow.

To further improve the sensitivity in detecting flow in small vessels, contrast-enhanced agents may be introduced.^[5,10] Studies have shown that the contrast-enhanced PDI can more easily visualize



Fig. 4 Tissue Doppler image of medial mitral valve annulus. Position of Doppler cursor on the annulus is seen in upper right of figure. Velocity spectral of the annulus is seen in lower part of figure.

small vessels and identify soft-tissue masses and abscesses.^[11,12]

Tissue Doppler Imaging (TDI)

TDI is a technique using PW Doppler to detect velocities of tissue movement.^[13] Either single-gate or multiple-gate procedures can be used. As mentioned previously, tissue and wall movements usually generate higher amplitude and lower frequency Doppler signals compared to those from blood flow. Filters to remove the high-frequency signals from blood flow may or may not be used, depending on the design of the device. Figure 4 shows a typical TD image of a normal mitral valve annulus in the heart. A sector transducer and a single-gate procedure are used. Similar to duplex scanning, the ultrasound Doppler beam and gate location are displayed in the B mode image at the upper right of the image. The detected tissue velocity spectrum within the sampling volume as a function of time is displayed at the lower portion of the image. A trace of the electrocardiogram (ECG) is also displayed above the tissue velocity spectrum to serve as a time reference. Velocities of tissue movement within a relatively large area can also be detected using multiple-gate procedure. Same as DCI, the measured tissue velocities will be color coded and superimposed on the gray-scale image.

APPLICATIONS OF ULTRASOUND DOPPLER

Ultrasound Doppler is used extensively in the peripheral vascular system, heart, and abdomen to

determine the presence or absence of blood flow and its direction, velocity, and character.

Carotid Artery Examination

In the United States, approximately 500,000 new strokes occur annually, and about 75% of these are ischemic or result from a decrease in blood flow to a particular area in the brain.^[14] Atherosclerosis, with associated vessel plaque disruption or lumen stenosis (lumen narrowing) is a common cause of stroke, particularly in the elderly. About 75% of the brain is supplied by the two internal carotid arteries. Flow characteristics in the carotid bifurcations also enhance their susceptibility to atherosclerotic plaque formation. Disruption of plaque and exposure of the underlying stroma in the vessel wall can result in thrombus formation and subsequent embolization, leading to stroke. Three large randomized studies comparing carotid endarterectomy with conservative medical therapy clearly demonstrated the superiority of endarterectomy over medical therapy for the reduction of stroke and death in patients with at least 60 to 70% internal carotid artery stenosis.^[15-17] Therefore, it is important to determine the morphology of the carotid bifurcation in many patients with acute stroke or transient ischemic attack (symptoms similar to stroke lasting less than 24 hours) or patients with carotid bruits. Two-dimensional imaging and Doppler ultrasound examination of the carotid bifurcation can provide important information about whether carotid plaque is present, its characteristics, and whether significant stenosis is present. PW Doppler and DCI remain the mainstay of carotid stenosis assessment.

Transcranial Doppler Examination

Transcranial Doppler (TCD) allows examination of the major intracerebral arteries through transorbital, transtemporal, suboccipital, and submandibular approaches.^[18] It allows detection of intracranial stenoses and arterial venous malformations in the major basal arteries of the brain. TCD is often used in conjunction with carotid ultrasound to evaluate the significance of internal carotid artery disease. Relative changes in cerebral blood flow can be measured objectively, immediately, and repeatedly, making TCD an attractive monitoring tool particularly during neurosurgical, vascular surgical, cardiac, and cerebral vascular procedures. Immediate cerebral blood flow information allows adjustments to be made that could reduce the incidence of postprocedural complications.



Arterial Disease of the Extremities

Decreased tissue perfusion in the extremities is usually caused by disease affecting the patency of arteries supplying these affected territories. The degree to which arterial disease produces symptoms and signs depends on the location and degree of arterial occlusion. If arterial disease is suspected, it is important to establish that a stenosis or occlusion is present, its location, and the nature of the underlying disease, i.e., atherosclerosis or thrombosis. Since ultrasound is noninvasive, it provides a low-risk, low-cost means for initial evaluation of patients suspected of having peripheral arterial disease. Therefore, it can help in the planning of more invasive procedures such as arteriography and intervention, both surgical and percutaneous.

One of the most important tests for assessing the patency of the arterial circulation in the lower limb is the ankle systolic pressure, which is done by placing an occluding cuff just above the ankle and assessing blood flow with CW Doppler in the posterior tibial and dorsalis pedis arteries. This is usually compared to the simultaneous brachial systolic pressure measured by brachial artery CW Doppler. The ankle-brachial index (ABI) is a ratio of the ankle systolic pressure divided by the brachial systolic pressure and is considered normal if above 0.90.^[19] The ABI has also been shown to be important in assessing risks for coronary artery disease and all-cause mortality.^[20]

Examination of the morphology of the Doppler spectral flow pattern or waveform in the main arteries of the upper or lower extremity provides additional information for assessing arterial occlusive disease. In this regard, waveforms are analyzed directly over a stenosis and proximal and distal to a stenosis, providing information about the severity of stenosis and the presence or absence of collateral flow.

Cardiac Doppler Examination

The application of ultrasound Doppler in the heart has made measurement of blood flow velocity and intracardiac pressure possible.^[21] Information that previously was only obtainable by cardiac catheterization is now obtainable noninvasively by Doppler ultrasound. This technique therefore allows serial assessment at low cost and minimal risk. A few examples follow that illustrate the application of Doppler ultrasound in cardiac imaging.

Cardiac output

Stroke volume (volume of blood ejected by the heart per beat) may be assessed by measuring the blood flow

velocity at a particular point where the cross-sectional area is known. Typically, this is the left ventricular outflow tract or the main pulmonary artery, and velocities are measured using PW Doppler. Because flow in the cardiovascular system is pulsatile, flow velocity varies during ejection, and the velocity must be integrated over the ejection period to measure the total volume of blood flow during a given ejection. This integration is called the time velocity integral (TVI) and is equal to the area enclosed by the Doppler spectral wave form. Multiplying the TVI by the cross-sectional area yields the stroke volume which, when multiplied by the heart rate, yields the cardiac output.

Valvular heart disease

Doppler ultrasound allows assessment of both valvular stenosis and valvular regurgitation (valvular incompetence). The major strength of Doppler ultrasound is that it allows accurate serial assessment of these lesions so that the hemodynamic data can be continually correlated with patient symptoms. In this way, patients can be initially treated with medical intervention and then mechanical intervention (either surgical or percutaneous valvuloplasty) as symptoms and hemodynamics dictate. Moreover, once medical therapy is instituted, further study can easily be done to assess response to therapy.

In valvular stenosis, the valve orifice is narrowed, causing an obstruction to blood flow. The assessment of the severity of valvular stenosis requires the determination of the transvalvular pressure gradient and valve area. Doppler ultrasound measures blood flow velocities and these may be converted to pressure gradient using the Bernoulli equation:

$$P_1 - P_2 = \frac{1}{2}\rho(V_2^2 - V_1^2) + \rho \int_1^2 \frac{dV}{dt} ds + R(V) \quad (5)$$

where 1 and 2 refer to an upstream (site 1) and a downstream (site 2) position, respectively, P is pressure, V is velocity, ρ is the density of the medium, and s refers to the streamline from site 1 to 2. The equation shows that the pressure drop, $P_1 - P_2$, is equal to the sum of three terms on the right-hand-side of the equation. The three terms represent the effects of convective acceleration, flow acceleration, and viscous friction loss, respectively.

In most clinical situations, flow acceleration and viscous friction can be ignored. When $\rho = 1.07 \text{ gm/cm}^3$ (the average density of human blood), the pressure change across a stenotic valve can be calculated with the simplified Bernoulli equation: $P_1 - P_2 = 4(V_2^2 - V_1^2)$, where the unit of P is mmHg and the unit of V is m/sec. Usually, flow velocity proximal (site 1) to

a narrowed valve orifice (site 2) is much smaller than the peak flow velocity at the orifice site, such that V_1 can be ignored as well. Thus, the pressure change across a stenotic valve is easily obtained by measuring the velocity across the valve with either PW Doppler or CW Doppler.

The valve area can be assessed using the continuity equation, which is an expression of conservation of flow. In a vessel, the volume of flow through a cross section at the upstream site 1 must equal the flow at the downstream site 2, i.e., $Q_1 = Q_2$ and thus, $A_1 \times V_1 = A_2 \times V_2$, where A = area. Rearranging $A_2 = (A_1 \times V_1) / V_2$. If the velocity and area at site 1 can be determined as well as the velocity at site 2, the desired stenotic area can be calculated.

In aortic stenosis, volume flow across the aortic valve at site 2 is equal to the volume flow across the left ventricular outflow tract at site 1. Velocities V_1 and V_2 can be measured by PW or CW Doppler. If it is assumed that the outflow tract is circular, the area may be calculated from the diameter measured from the B mode image, and then the stenotic valve area (A_2) can be determined.

Intracardiac pressure

Application of the Bernoulli equation to regurgitant valves allows estimation of intracardiac pressures. For example, the peak tricuspid valve regurgitant jet velocity measured usually with CW Doppler reflects the peak systolic pressure difference between the right atrium (RA) and right ventricle (RV). The RA pressure can be estimated from a physical exam of the jugular vein in the neck and in healthy individuals, usually varies between 5–10 mmHg at rest. It is elevated when pulmonary hypertension or right ventricular failure is present. Thus, the RV systolic pressure is equal to the peak systolic pressure difference between the RA and RV plus the estimated RA pressure. It follows that the pulmonary artery peak systolic pressure is equal to the peak RV systolic pressure when there is no obstruction to RV outflow. This can be assessed by measurement of Doppler velocities in the RV outflow tract and pulmonary valve. Thus, pulmonary pressures can be estimated by measuring tricuspid regurgitant jet velocities.

Abdominal Doppler Examination

Normal duplex scanning waveforms and color flow patterns have been investigated for vessels in the abdomen. These include the abdominal aorta, the inferior vena cava (IVC), and portal, hepatic, renal, splenic, and mesenteric vessels. Abnormalities can be diagnosed by comparing the spectral flow waveforms or

color flow patterns in the vessel of interest with normal patterns. Ultrasound Doppler has been used for diagnosis of abdominal aortic stenosis due to atherosclerosis, aortic aneurysms, and inferior vena cava filters.^[5,22] PDI can detect flow in small vessels in organs such as the liver, kidney, and heart, and hence assess perfusion. PDI and contrast-enhanced PDI have been used to detect tumors and masses in these organs as well.^[23]

CONCLUSION

Ultrasound Doppler is a rapidly developing technology with numerous applications in biomedical fields. It is used to quickly identify and differentiate vascular from nonvascular structures. It is employed to determine flow direction, velocity distribution, and their temporal variation. Blood flow supply to different organs may be determined. Ultrasound Doppler permits the diagnosis of a variety of abnormalities, including stenosis, aneurysms, and vessel occlusion in the cardiovascular system. Abnormal tissues, including tumor masses in different organs, may also be identified. With the development of new imaging and signal processing technologies, the applications of ultrasound Doppler will be further expanded in biomedical fields.

ARTICLES OF FURTHER INTEREST

Intravascular Ultrasound; Medical Imaging, 3-D; Ultrasound and Tissue Interaction

REFERENCES

1. Shung, K.K.; Smith, M.B.; Tsui, B.M.W. *Principles of Medical Imaging*; Academic Press, Inc.: San Diego, 1992.
2. Kremkau, F.W. *Doppler Ultrasound: Principles and Instruments*, 5th Ed.; W.B. Saunders Company: Philadelphia, 1998.
3. Fish, P.J. Doppler Methods. In *Physical Principles of Medical Ultrasonics*; Hill, C.R., Ed.; Ellis Horwood Limited: New York, 1986; 338–376.
4. Zagzebski, J.A. *Essentials of Ultrasound Physics*; Mosby-Year Book, Inc.: St. Louis, 1996.
5. *Clinical Doppler Ultrasound*; Allan, P.L., Dubbins, P.A., Pozhiak, M.A., McDicken, W.N., Eds.; Churchill Livingstone, Inc.: New York, 2000.
6. *Doppler Color Imaging*; Merritt, C.R.B., Ed.; Churchill Livingstone, Inc.: New York, 1992.
7. Fei, D.Y.; Fu, C.T.; Brewer, W.H.; Kraft, K.A. Angle independent Doppler color imaging: Determination of accuracy and a method of display. *Ultrasound Med. Biol.* **1994**, *20*, 147–155.



8. Fei, D.Y.; Fu, C.T. New method to obtain angle independent Doppler color images using a sector transducer. *Ann. Biomed. Eng.* **1999**, *27*, 187–193.
9. Hoskins, P.R. A comparison of single- and dual-beam methods for maximum velocity estimation. *Ultrasound Med. Biol.* **1999**, *25*, 583–592.
10. Arslan, H.; Sakarya, M.E.; Bozkurt, M.; Unal, O.; Dilek, O.N.; Harman, M. The role of power Doppler sonography in the evaluation of superficial soft tissue abscesses. *Eur. J. Ultrasound* **1998**, *8*, 101–106.
11. Ophir, J.; Parker, K.J. Contrast agents in diagnostic ultrasound. *Ultrasound Med. Biol.* **1989**, *15*, 319–333.
12. Mitchell, D.G.; Merton, D.A.; Liu, J.B.; Goldberg, B.B. Superficial masses with color Doppler imaging. *J. Clin. Ultrasound* **1991**, *19*, 555–560.
13. McDicken, W.N.; Sutherland, G.R.; Gordon, L.N. Color Doppler velocity imaging of the myocardium. *Ultrasound Med. Biol.* **1992**, *18*, 651–654.
14. National Advisory Neurological and Communicative Disorders and Stroke Council. *Decade of the Brain: Answers Through Scientific Research*; U.S. Dept. of Health and Human Services, Public Health Service, National Institutes of Health: Bethesda, MD, 1989. NLM Unique ID: 8912912.
15. European Carotid Surgery Trialists' Collaborative Group Interim results for symptomatic patients with severe (70–99%) or with mild (0–29%) carotid stenosis. MRC European Carotid Surgery Trial. *Lancet* **1991**, *337*, 1235–1243.
16. North American Symptomatic Carotid Endarterectomy Trial Collaborators Beneficial effect of carotid endarterectomy in symptomatic patients with high-grade carotid stenosis. *N. Engl. J. Med.* **1991**, *325*, 445–453.
17. Executive Committee for the Asymptomatic Carotid Atherosclerosis Study. Endarterectomy for asymptomatic carotid artery stenosis. *JAMA* **1995**, *273*, 1421–1428.
18. Aaslid, R.; Lindegaard, K.F. *Transcranial Doppler Sonography*; Aaslid, R., Ed.; Wien, Springer-Verlag, New York, 1986.
19. Yao, S.T. Haemodynamic studies in peripheral arterial disease. *Br. J. Surg.* **1970**, *57*, 761–766.
20. Vogt, M.T.; Cauley, J.A.; Newman, A.B.; Kuller, L.H.; Hulley, S.B. Decreased ankle/arm blood pressure index and mortality in elderly women. *JAMA* **1993**, *270*, 465–469.
21. Phillips, J.H. *Practical Quantitative Doppler Echocardiography*; CRC Press, Inc.: Boca Raton, FL, 1991.
22. White, E.M.; Choyke, P.L. Duplex Sonography of the Abdomen. In *Duplex Sonography*; Grant, E.G., White, E.M., Eds.; Springer-Verlag: New York, 1988; 129–190.
23. Sallomi, D. The use of contrast-enhanced power Doppler ultrasound in the diagnosis and follow-up of inflammatory abdominal masses associated with Crohn's disease. *Eur. J. Gastroenterol. Hepatol.* **2003**, *15*, 249–251.

Ultrasound Therapy, Bone Healing

Özgür Erdoğan

Department of Oral and Maxillofacial Surgery, Faculty of Dentistry, Çukurova University, Balcali, Adana, Turkey

INTRODUCTION

Bone healing is a complex process with several overlapping steps. Because bone is a dynamic tissue affected by systemic or local factors, bone healing is also associated with a variety of biochemical, biomechanical, cellular, hormonal, and pathologic events. Therapeutic approaches aim to accelerate bone healing in order to shorten treatment time or treat non-healing injuries caused by systemic or local diseases. Numerous therapeutic methods for bone healing have been described in the literature. These applications are systemic or local drug deliveries or physical treatments like low-intensity laser, electromagnetic fields, extracorporeal shock, mechanical stimulation, and ultrasound (US) therapy. The use of US as a therapeutic approach in bone healing has a history of more than half a century. This chapter discusses the current understanding of the molecular mechanism of US therapy, surveys the findings of the literature, and evaluates the future of the application.

Historical Perspective

The piezoelectric effect was invented by brothers Jacques and Pierre Curie in 1880 when they realized that some crystals produce acoustic waves if alternate electricity is applied in resonance frequency. The piezoelectric effect refers to the generation of an electrical response to an applied pressure. Langevin used this principle in submarine detectors in the first years of World War I. Industrial use of US began in 1928 to detect hidden flaws in materials. The biological effects of US were first discovered at that time. The therapeutic use of US in medicine was first reported by Pohlman et al. in 1939. The authors used US for the treatment of back pain, neuralgia, and myalgia. They reported improvement in the symptoms when US was applied at a frequency of 800 KHz, with US energy transmitted at 4–5 W/cm² intensity to biological tissues for 10 days. The daily US treatment lasted 10 min. Although this dose of US treatment created some increased heat, it did not cause any kind of thermal injury. Diagnostic application of US began in the late 1940s. US was used as a therapeutic tool in the treatment of many diseases until the 1950s. The general

recognition was that US caused damage and its use in bones should be avoided. Halsscheidt et al. reported in 1949 a patient with osteomyelitis in the alveolar bone who received US treatment for a lesion in the lower lip.^[1,2]

The first study concerning the effects of US on bone healing was published in 1950 by Maintz.^[2] This study was a cornerstone, because it demonstrated for the first time the positive effects, but not the harmful effects, of US on bone. In that study, the author created bilateral osteotomies in the radius of 3-month-old rabbits. One limb of the rabbit received 0.5, 1.0, 1.5, or 2.5 W/cm² US treatment. According to histological and radiological evaluations, although US treatment in high intensities could cause thermal damage in bone, lower doses led to new periosteal bone formation. A placebo controlled study by Corradi and Cozzolino^[3] demonstrated that US application stimulates new callus formation in the fracture line in the radial bones of rabbits. After this study, the authors used the application in the clinic and reported that US treatment led to periosteal callus formation in eight patients.^[4]

Shiro advocated the use of lower doses and pulsed US application to decrease thermal damage to bones.^[5] The author used pulsed US application at 0.2 W/cm² intensity and reported increased osteoblastic/chondroblastic activity in osteotomized tibiae of rabbits in the experimental group. Twenty years after Shiro's study, Dyson and Brookes created bilateral fibula osteotomy in rabbits and applied US signals at 0.5 W/cm² intensity to one side. The authors reported increased ossification in the experimental group compared with the control group.^[6] In the same year, Xavier and Duarte applied lower doses of US (30 mW/cm²) to 26 patients with lower extremity fractures with nonunion after conventional treatments.^[7] They reported 70% healing after 20 min of daily US application. After that study, Duarte demonstrated an increase of 28% in ossification in the osteotomized fibulae of rabbits that received low-intensity-pulsed-ultrasound (LIPU) application compared with the control group.^[8]

In a placebo controlled study, Pilla et al. osteotomized the tibiae of rabbits bilaterally and applied LIPU (1.5-MHz pressure wave administered in pulses of 200 msec with an average temporal and spatial intensity of 30 mW/cm²) to rabbits in the experimental group.^[9] The authors reported that the osteotomized legs in the

U

LIPU group gained the mechanical properties of the non-osteotomized legs on the 17th day, whereas the osteotomized legs in the control group gained the mechanical properties of the non-osteotomized legs on the 28th day.

In 1998, Jangushi et al. reported that LIPU positively affects all phases of endochondral ossification and the optimal frequency of US pulses is 1 Hz, which corresponds to 200-msec intervals. Another study by Azuma et al. similarly demonstrated that LIPU has positive effects on ossification in all phases of fracture healing.^[10]

After successful reports in fracture treatment, the efficiency of LIPU was studied in relation to other orthopedic surgical procedures. Positive effects of LIPU were demonstrated in spinal fusion and distraction osteogenesis procedures in rabbits.^[11,12]

As a result of these *in vitro* and *in vivo* studies, US treatment was used in clinical applications for accelerating bone healing. In a randomized, double-blind controlled study by Heckman et al., the effects of LIPU on fresh tibia fractures were evaluated. In this multicenter study consisting of 67 patients, LIPU treatment was applied to 33 patients, with 34 patients in the control group. The authors reported that LIPU application accelerates the fracture healing at a rate of 38%.^[13] After 3 years, accelerated healing of fresh dorsal radius fractures was reported by Kristiansen et al. by means of LIPU treatment. The average healing time for the fractures in the experimental group was 61 days, whereas the healing time was 98 days for the control group.^[14]

The Sonic Accelerated Fracture Healing System (SAFHS; Smith and Nephew, Exogen, Memphis, TN) is a commercially available therapeutic US device that allows home application by the patient (Fig. 1). The US signals generated by the device consist of a 1.5-MHz sine wave, administered in bursts of 200 μ sec, followed by a pause of 800 μ sec with a repetition rate of 1 KHz. The average temporal and spatial intensity is 30 mW/cm². The US is administered for 200 min daily. The SAFHS was initially cleared for marketing by the Federal Drug Administration in 1994 as a treatment for fresh, closed, posteriorly displaced distal radius fractures and fresh, closed, or grade I open tibial diaphysis fractures in

skeletally mature individuals when these fractures are orthopedically managed by closed reduction and cast immobilization. In 2000, the indications were expanded to include the treatment of established non-unions, excluding the skull and vertebrae.

Bone Healing

Bone is a dynamic, well-organized tissue that can remodel itself in response to environmental mechanical stress and hormonal activity. Mature bone consists of three major components: marrow, periosteum, and bone tissue. At the microscopic level bone can be classified into two types: lamellar and woven bone. Woven bone is usually found in growing bones, regenerating bones like fracture callus, and some specific regions such as ear ossicles and tendon or ligament attachments. The microscopic appearance of woven bone is easily distinguishable from that of lamellar bone, because the orientation of collagen fibers is not uniformly formed as in lamellar bone.

Regarding porosity and architecture, bone can be in trabecular (cancellous) or compact (cortical) forms. Cortical bone and trabecular bone have the same matrix component. However, the density is increased in cortical bone and decreased in trabecular bone because of increased marrow space.

Bone matrix is composed of organic components, inorganic components, and water. The major organic component is collagen. It provides elasticity and thus resistance to external forces. Type I collagen is predominantly present in bone tissue. The inorganic components are mainly calcium phosphate, calcium carbonate, magnesium fluoride, and sodium. These inorganic components give stiffness to the bone tissue and serve as an ion reservoir for the body.

The unique design of living parts of bone facilitates bone remodeling according to the type and direction of external forces. Osteoblasts, osteocytes, osteoclasts, and undifferentiated mesenchymal cells constitute bone tissue. The mediators present in extracellular matrix organize the transformation of hematopoietic stem cells to osteoblasts or undifferentiated mesenchymal cells to

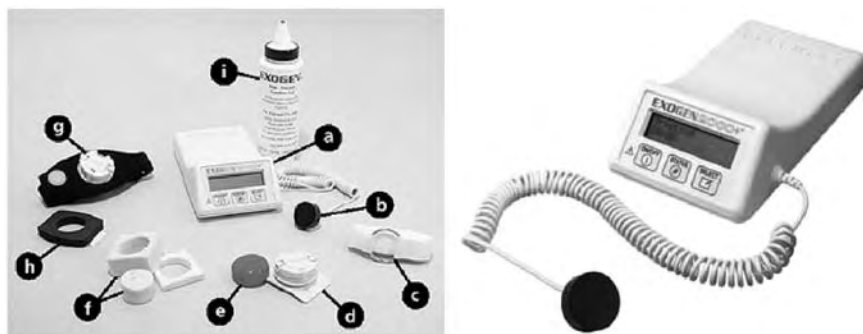


Fig. 1 Equipment of the Exogen 2000™ device.

osteoblasts. Osteocytes are the main type of bone cells in the human skeleton. They reside in and are surrounded by mineralized bone matrix. They are connected through canaliculi and communicate between themselves and the matrix. Osteoblasts build bone matrix and osteoclasts resorb bone matrix. The actions of these cells play a role in bone healing.

Bone formation may occur through transformation from cartilage tissue (endochondral ossification), a calcification surrounding the organic matrix (endochondral ossification) or formation of new bone on existing bone (appositional ossification). The three types of ossification occur throughout the formation of new bones, the regeneration process of the skeleton after injury, and surgery or formation of bone tumors.

Remodeling of bone is regulated by a combination of resorption and bone formation. The primary cells involved are osteoblasts and osteoclasts. There is a close anatomical and functional relationship between resorptive and formative cells at the remodeling sites. These cells demonstrate their action within a temporary anatomic structure called the basic multicellular unit (BMU).^[15] A mature BMU consists of a leading front of osteoclasts followed by osteoblasts, blood supplies, and the associated connective tissues. In the remodeling process the entire BMU moves forward; osteoclasts resorb bone and die by apoptosis. The resorbed bone is replaced by osteoblast, bone-forming cells, and new bone formation.

Bone metabolism is constantly regulated by hormonal activity. Major hormones that affect bone metabolism are parathyroid hormone, vitamin D, calcitonin, estrogen, glucocorticoid, and growth hormone. Parathyroid hormone regulates the calcium level in extracellular fluids by increasing the calcium flow from the intracellular compartments. Vitamin D stimulates renal calcium-binding proteins and facilitates active calcium transport. Calcitonin is secreted by parafollicular cells in the thyroid glands. It serves to decrease the levels of plasma calcium and phosphorus. Estrogen is produced by the ovaries; deficiency or absence of estrogen causes an increase in bone turnover, negatively affecting bone healing. Glucocorticoids inhibit formation of osteoblasts and synthesis of collagen by osteoblasts. Growth hormone plays a major role in the maturation of the skeletal system and increases intestinal calcium absorption.

Some proteins or growth factors released from platelets are associated with bone healing. Most well-known proteins are bone morphogenetic proteins, which induce mesenchymal stem cells that differentiate into bone cells. Other known proteins and growth factors that influence bone healing are insulin-like growth factor, transforming growth factor, platelet-derived growth factor, fibroblast growth factor, osteonectin, fibronectin, and osteocalcin.

Bone healing or fracture healing is the repair of injured bone. It has three distinct but overlapping stages: the early inflammatory stage, the repair stage, and the late remodeling stage.^[16] In the inflammatory stage, hematoma develops and inflammatory cells and fibroblasts infiltrate the injured area. A granulation tissue forms, ingrowth of vascular tissue occurs, and mesenchymal stem cells migrate in this stage. In the repair stage, fibroblasts serve as a stroma for vascular ingrowth and are responsible for synthesis of collagen matrix and angiogenesis. Collagen matrix ossifies and consequently callus tissue develops. Bone healing is completed during the remodeling stage. Woven bone is differentiated to lamellar bone and bone gains its original strength after completion of this stage.



Ultrasound Physics

Acoustic waves are vibrations that are transmitted mechanically. This means that a solid material, which consists of particles, is required for generating and transmitting the ultrasonic waves. It is not possible to transmit the ultrasonic waves in vacuumed spaces, which do not have particles. The number of impulses generated in each particulate is an important feature in determining characteristics of acoustic waves. Acoustic waves show a sinusoidal propagation when propagating in the medium.

Acoustic vibrations that are transmitted over 20-KHz frequency are called ultrasonic. Ultrasonic waves are longitudinal waves, which have compression and expansion bands. Ultrasound waves are transmitted through mechanical vibrations, unlike electromagnetic energy, which is transmitted through photons (Fig. 2).

Piezoelectric crystals are used for generating ultrasound waves. The particles in the piezoelectric crystals are stimulated and start to produce waves and transmit the waves to other particles mechanically. Piezoelectric crystals are present in the transducer of the US device. These crystals are part of the US devices and are in contact with the tissues. The crystals are deformed in response to electrical voltage applied to them and ultrasonic waves are generated on the other side of the crystals. The piezoelectric crystals used in the therapeutic US devices are usually lead–zirconium–titanate crystals. Quartz, barium–titanate, and Rochelle crystals are also used for this purpose.

The number of complete oscillations each particle undergoes in a second is called the frequency. A wave that has 100 cycles in 1 sec means that a 0.01-msec time interval exists between the cycles. The unit of frequency measurement is Hertz. Hertz defines the number of cycles per second. Higher frequencies are abbreviated as follows: 1,000,000 Hz = 1000 KHz = 1 MHz. The

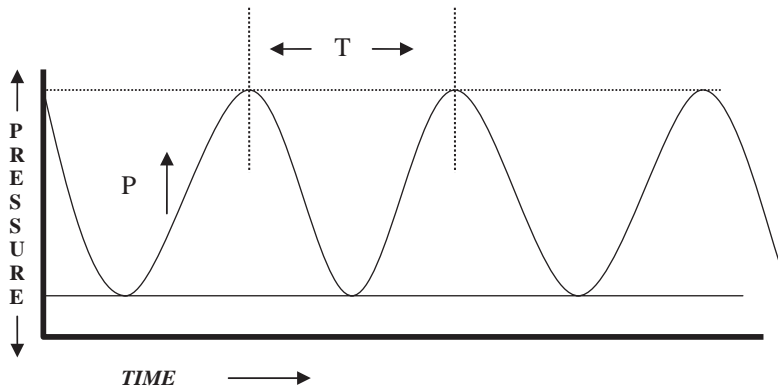


Fig. 2 A pressure/time sinusoid occurring as a result of pressure to a medium. This type of sinusoid develops when a constant vibration is administered.

frequency interval of the acoustic waves of the human hearing limit is between 20 Hz and 20 KHz, which corresponds to 10 octaves of interval. It is a considerably high interval when compared to visible light interval, which has only 2 octaves (Fig. 3). Frequencies employed in therapeutic US applications generally range from 1 MHz to 3 MHz. Higher US frequencies tend to generate narrower beam width waves.

The efficiency of piezoelectric crystals is dependent on the amount of generated ultrasonic waves, which are transformed from the electrical energy applied to them. Piezoelectric crystals are active at their particular resonance frequency. The resonance frequency of a piezoelectric crystal depends on the thickness of the crystals. A scheme of a typical US transducer is illustrated in Fig. 4.

Power and intensity are measures of the strength of US waves. In physics, power (P) is defined as the amount of energy passing through a surface per unit time and is expressed in watts. The formulation for the calculation of power is

$$P = (F \times l)/t,$$

where P is the power, F is the force, l is the distance, and t is the time.

Medical ultrasound focuses on a small area and the term "intensity" is used to define the power of ultrasound. It is the rate of energy delivered per unit area (W/cm^2). The output of the device is expressed in watts divided by the effective radiating surface area of the transducer resulting in the power. The intensity changes depending on the width of the ultrasound beam. Therefore, in focused sound beams, the intensity is greatest at the focus where the beam width is the

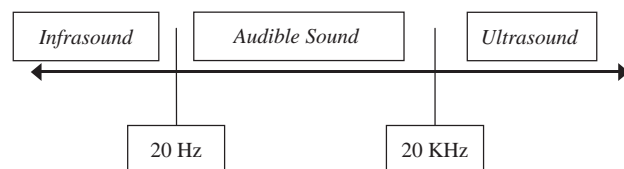


Fig. 3 Frequency intervals of acoustic waves.

narrowest. In pulsed ultrasound beams, the greatest intensity value occurs during the pulse and is zero between pulses. Therefore, acoustic intensities must be specified as to whether they are temporal-average or temporal-peak values and whether they are spatial-average or spatial-peak values. Therapeutic ultrasound may be classified as low intensity or high intensity according to the energy output. Low-intensity US has an output value as low as 20 mW/cm^2 , which is comparable to diagnostic US. It is mainly used for stimulation of the tissues in order to increase healing capacity, as in treatment of fractures. High-intensity US corresponds to intensities higher than 5 W/cm^2 and its thermal effect is utilized for physical treatments like decreasing joint stiffness following surgical procedures like ablation or excision of tissues.

Velocity is the rate of energy transmitted in a specified direction of propagation. Velocity of ultrasound waves depends on the physical properties of the medium in which the wave is transmitted. The waves are transmitted faster in solid media compared with liquids and air. For example, the velocity of

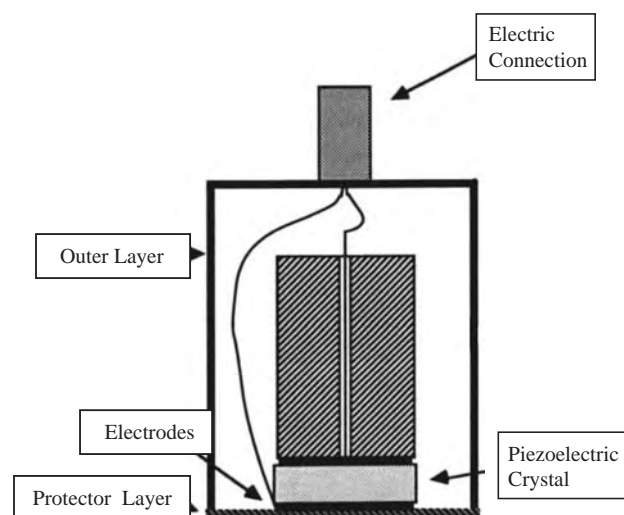


Fig. 4 A typical US transducer.

US is 1500 m/sec in saline and 350 m/sec in air. The velocity in bone is between 3000 and 5000 m/sec.^[17] Velocity of propagation does not depend on the frequency of US.

Wavelength is the distance between two corresponding points of a wave. It is dependent on frequency and velocity and formulation of calculation is

$$\text{Wavelength } (I) = \text{Velocity } (V) / \text{Frequency } (F).$$

As the US wave travels through the tissue, it continually loses energy. The decrease in energy per unit distance is called attenuation. Attenuation occurs through absorption, scattering, and transmission. Energy is lost to reflection or scattering of the sound beam when it strikes a reflecting surface or is lost to absorption as it overcomes the internal friction that exists in tissue while traveling through it. The frequency of the US beam and properties of the tissue determine the attenuation of the energy. For example, a 1-MHz US wave loses half of its energy after propagating 9 mm in muscle tissue, whereas it loses half its energy after 50 mm propagation in fat tissue.^[18] In general, the tissues comprising more protein absorb more energy compared with tissues comprising more water. The attenuation of US in bone occurs through absorption and scattering. When energy impinges on bone, some of the energy is transformed to shear, which is rapidly attenuated at the front surface of the bone.^[17]

Therapeutic ultrasound applications are performed in pulsed mode in order to reduce the thermal effect and increase the stimulation level. Pulsed ultrasound may be applied at different frequencies and pulse lengths. Intervals between the pulses may also be adjusted. In 1:1 mode, after a 200- μ sec pulse duration there is a 200- μ sec pause period. In 1:4 mode, the pause is 800 μ sec after every 200- μ sec pulse period.

In order to conduct the ultrasound waves to tissues, a coupling medium in liquid, gel, or cream form is put between the transducer and the tissue. An ideal coupling medium should fill the space, have some viscosity in order to stay in place, have similar acoustic resistance with the tissue with which it is coupled, and cause minimal energy absorption in order to transmit the waves with minimal energy loss. High water-comprising gels have been developed as coupling media for this purpose.

Interaction of Ultrasound with Bone Tissue

Despite the complex characteristics of bone healing and multiple effects of ultrasound on biological tissues, much is known today about the mechanism of action of US therapy. Several theories have been proposed to explain these effects. The combination of physical, piezoelectric, and biologic effects of US plays a role in the interaction of US with bone healing.

Physical effects

Ultrasound waves generate vibrations in all tissue components, including intracellular and extracellular fluids and cell membranes, when penetrating tissues. They cause movements or displacements of particles when transmitted in tissues. Because of these motions, US treatment produces a micromassage effect in tissues, which produces mechanical stimulation.

Acoustic vibrations have thermal and non-thermal effects on biologic tissues. The energy attenuated by tissues leads to thermal increase. US signals at high intensities create considerable heat increase. However, ultrasound therapy utilized for bone healing uses intensity as low as 20–50 mW/cm². Vibrations at these intensities generate a heat increase below 1°C, which is negligible in terms of thermal damage. However, this small increase in temperature may affect some enzymes such as matrix metalloproteinase 1.^[19] This enzyme is very sensitive to small variations in temperature. In terms of biologic effects, the thermal and non-thermal effects of US are not separable and it is assumed that non-thermal effects of US are usually accompanied by some heat increase. The observed non-thermal effects of US are stable cavitations, microstreaming, acoustic streaming, and direct effects to the cell membrane.

Stable cavitation defines the presence of gas bubbles in fluids as a result of US energy. These bubbles may be in different shapes and sizes depending on the type of US signals and lead local movements in the fluid. These movements in fluids are called microstreaming. Ter Haar et al. demonstrated that US signals of 1.5-MHz frequency and 150 mW/cm² intensity generate bubbles bigger than 10 μ m in the body fluids of pigs. Acoustic streams are localized high-velocity streams of fluid created by US signals in a liquid.^[20] These mechanisms cause movements and transfers of intracellular and extracellular ions. An increase in intercellular ion transfers leads to changes in membrane permeability, resulting in alterations of the electrophysiological properties of the cells. US application increases ion conductivity up to 22% and decreases the need for ATP consumption by sodium–potassium channels.^[21]

Piezoelectric effect

An increase in electrical potential occurs when pressure is applied to bone tissue. This effect is called the piezoelectric effect. Piezoelectricity affects the arrangement of bone resorption and bone healing. Bone as a dynamic living tissue remodels itself according to environmental forces. US application generates a piezoelectric effect on the bone, thus increasing the electric potential within the tissue. An increase in electric



potential may induce bone remodeling, thereby assisting in fracture healing.^[22]

Biologic effects

As US signals travel in bone, a biologic response develops in the tissue. There are many *in vivo* and *in vitro* studies regarding the biologic effects of US on bone healing.

In vitro studies on human or animal cell cultures demonstrated several cellular effects of US, which may be related to acceleration in bone healing or fracture repair. LIPU application stimulates aggrecan mRNA expression and proteoglycan synthesis by chondrocytes, increases prostaglandin E₂ synthesis by osteoblasts, and increases platelet-derived growth factor synthesis. US signals induce conformational changes in the cell membrane and alter ionic permeability and second messenger activity. Changes in second messenger activity lead to downstream alterations in gene expression, resulting in acceleration of the fracture–repair process by up-regulating cartilage and bone-specific genes. Others demonstrated effects of LIPU on bone healing as increased expression of transforming growth factor- β , increased efflux of potassium ions from intracellular spaces, and increased calcium ion incorporation.^[2,4]

Increased blood flow to the fracture site promotes delivery of key components, such as growth factors and cytokines necessary for the normal healing process. Capillary formation occurs in the injured area in the inflammatory phase of bone healing. US application increases the synthesis of angiogenesis-related cytokines such as interleukin-8, fibroblast growth factor (FGF), and vascular endothelial growth factor.^[23] A power Doppler sonography-based study demonstrated that low-intensity ultrasound delivered over a 10-day period stimulated a greater degree of vascularity at the site of ulnar osteotomies in dogs.^[24]

Several studies addressed the question of which phase of fracture healing is influenced by US therapy. Controversial studies have been published regarding this issue. Yang et al. reported that aggrecan gene expression was significantly higher in the LIPU-treated group on the 7th day but significantly lower on the 21st day in a rat femur fracture model.^[25] Similarly, some studies indicated that LIPU influences the early phases, consisting of inflammatory and callus formation phases of fracture healing^[9,26,27] but not the remodeling phase. Azuma et al. reported controversial results indicating that LIPU treatment affects all phases of fracture healing, including the bone remodeling phase.^[10]

According to studies concerning the mechanism of stimulation of bone healing by US, US application mainly affects endochondral bone formation. The studies reveal that US application stimulates chondrocyte

proliferation or maturation, which causes an increase in bone healing. The number of chondrocytes or osteoblasts was not increased by US application. There is no strong evidence available regarding the stimulation of intramembraneous bone healing by US therapy. From these studies it may be concluded that US therapy causes the fracture callus to gain its original strength in a shorter period of time through faster maturation of chondrocytes. Stimulation of the proliferation of chondrocytes occurs by several mechanisms such as unregulated cartilage, bone, or vascular gene expressions.

Clinical Applications

Clinical use of therapeutic US for bone healing is mostly concerned with fracture management. Most clinical studies regarding the effects of US treatment for bone healing are based on fresh fracture models. US treatment reduced the healing time of fresh fractures of tibial, distal radius, and scaphoid bones.^[13,14,28] Although fresh fractures heal without complications with standard fracture care, low-intensity ultrasound treatment may be considered for the treatment of fresh, closed fractures with closed reduction for patients with codiseases associated with bone metabolism such as diabetes, steroid therapy, osteoporosis, and smoking.

An open fracture is defined as one in which there is communication with a break in the integument, usually near the fracture site.^[29] Delayed union corresponds to a slower than expected healing process. Non-union is defined as cessation of healing for more than 3 months, as documented by serial radiographs including multiple views.^[29] Treatment of delayed unions or non-unions may be complicated because of some systemic or local conditions like osteoporosis or radiotherapy. US was beneficial in the treatment of delayed unions and non-unions. Mayr et al. utilized the application in 951 patients with delayed unions and 366 patients with non-union. They reported successful healing in more than 86% of patients in both groups.^[30] Nolte et al. applied the treatment to 29 patients with non-union after conventional therapies. The fractures were located in the tibia, femur, radius/ulna, scaphoid, humerus, metatarsal, and clavicle. They demonstrated complete healing in 86% of patients on average in 22 weeks.^[31] The efficiency of LIPU treatment on open fractures has not yet been proven by randomized clinical trials. Evidence of acceleration of bone healing is weaker compared with treatment for fresh closed fractures of long bones because of the design of the published studies concerning the treatment of delayed union or non-union. However, this patient group constitutes the main group needing US therapy.

Distraction osteogenesis is another clinical procedure that utilized US therapy. Few clinical studies

demonstrate the positive effects of LIPU on distraction callus. In a study by Tsumaki et al., 21 patients with osteoarthritis underwent bilateral tibial hemicallotaxis. One limb of every patient received ultrasound treatment for 20 min daily for 4 weeks. The authors demonstrated increased callus bone mineral density in the LIPU-treated limbs.^[32] Another study by El-Mowafi and Mohsen demonstrated accelerated healing of distraction callus of tibiae in LIPU-treated patients.^[33] Clinical studies are mainly in favor of LIPU application for distraction osteogenesis procedures. However, the number of studies and patients are limited and more clinical studies are needed in this area.

Spinal arthrodesis is another orthopedic procedure that may benefit from LIPU treatment. Although several animal studies demonstrated positive effects of LIPU treatment on spinal arthrodesis procedures, no clinical report is available so far.

Because of altered bone turnover rate, implant osseointegration is compromised in osteoporotic patients. Management of fractures is complicated in osteoporotic patients for the same reason. Although US treatment stimulated bone growth in tantalum implants in dogs, no clinical use has been reported so far.^[34]

U.S. therapy and maxillofacial bones

Maxillofacial bones differ from extremity bones in their embryologic genesis, their growth, and their histologic appearance. Most clinical and preclinical studies were conducted on long bones and the number of studies concerning the effects of LIPU on maxillofacial bones is limited and controversial. The first clinical report was published by Harris et al. in 1992.^[35] The author applied US treatment with 3-MHz frequency at 1.0 W/cm² intensity for 60 days for the treatment of osteoradionecrosis of the mandible. He reported positive progress in 10 of 21 patients.^[35] In vitro studies conducted on mandibular osteoblast and fibroblast cultures indicated that US signals increase proliferation of cells, synthesis of cytokines like interleukins and FGF, and synthesis of nitric oxide and prostaglandins, which are important mediators in angiogenesis and bone healing. The first animal study on mandible was published by El-Bialy et al.^[36] The authors demonstrated enhanced bone formation on the distraction site stimulated by US application. A clinical study reported that LIPU application provides better repair for orthodontically induced root resorption.^[37] Another clinical study by Schortinghuis et al.^[42] indicated that LIPU has no effect on ossification in the distraction gap of vertically distracted mandible. A listing of animal and clinical studies concerning maxillofacial bones is presented in Table 1.

It is mainly accepted that US treatment enhances endochondral bone formation. The fact that maxillofacial

bones have intramembraneous ossification makes the use of US therapy questionable in this particular region. The close relationship to important anatomic locations like the brain, parotid gland, and eyes, in which the effects of US are not known, is another drawback for the use of US in maxillofacial bones. Another drawback is the thicker soft tissue layer compared with tibial, ulnar, or scaphoid bone regions, which makes the penetration of US signals weaker, thus reducing the effectiveness of US treatment. The treatment of most maxillofacial fractures requires intermaxillary fixation. Long duration of intermaxillary fixations causes many problems like impairment in feeding, temporomandibular joint problems, or dental–periodontal problems. Shortening the intermaxillary fixation time may be one main indication for LIPU applications in the maxillofacial area. One animal study conducted on rabbits demonstrated increased fracture healing in a mandibular fracture model.^[38] However, more studies are needed to determine the optimal dosing and treatment modality before utilizing the procedure in clinical settings.

Complications and Side Effects

Ultrasound treatment as an aid to bone healing generates low-intensity energy, which is comparable to diagnostic US doses. Pulsed signals and low energy minimize the thermal effect of US on biological tissues. No serious complication has been reported regarding the clinical use of US therapy. Reported adverse effects as a result of US treatment include muscle spasm at the treatment side, mild erythema, and mild swelling. US signals do not interfere with metal objects and its use in sites with implanted metal biomaterial is safe. One disadvantage of the home application type of US treatment is patient dependency on the treatment. In such situations patient cooperation affects the success of the therapy. The effects of US therapy on growing bones and brain tissue are not clear. Thus its use in pediatric patients and in skull bones should be avoided. The use of US in sites of suspected neoplasia and acute infection is contraindicated because of possible accelerated progress of the disease. Patients should be evaluated by allergic reactions against the coupling gel. Patients on cardiac pacemakers should avoid US treatment because of the possible interaction with US signals.

Future of Ultrasound Therapy for Bone Healing

The minimal complication rates and adverse effects of US therapy for acceleration of bone healing make it possible to use this therapy in clinical settings. Although many clinical and animal studies demonstrate



Table 1 Clinical and animal studies concerning the effects of ultrasound therapy on maxillofacial bones

Author/year	Indication	Animal/clinical (No.)	US intensity	Conclusion
Fedotov et al., 1986 ^[39]	Mandibular fracture	Rabbits (36)	0.2–0.6 W/cm ²	Stimulated reparative process at fracture site
Harris, 1992 ^[35]	Osteoradionecrosis of mandible	Clinical (21)	1.0 W/cm ²	Effective treatment in 48% of patients
El-Bialy et al., 2002 ^[36]	Mandibular distraction osteogenesis	Rabbit (21)	30 mW/cm ²	Enhanced bone formation at the distraction site
El-Bialy et al., 2003 ^[40]	Effects on mandibular incisor teeth eruption	Rabbit (15)	30 mW/cm ²	Enhanced mandibular incisor growth and eruption
El-Bialy et al., 2003 ^[41]	Mandibular growth modification	Rabbit (8)	30 mW/cm ²	Enhanced mandibular growth
Schortinghuis et al., 2004 ^[42]	Healing of mandibular defects	Rat (72)	30 mW/cm ²	No significant benefit of US treatment
Schortinghuis et al., 2004 ^[43]	Healing of mandibular defects covered with e-PTFE membrane	Rat (72)	30 mW/cm ²	No significant benefit of US treatment
El-Bialy et al., 2004 ^[37]	Effects on orthodontic root resorption	Clinical (12)	30 mW/cm ²	Decrease in the areas of resorption
Schortinghuis et al., 2005 ^[44]	Vertical distraction osteogenesis of the mandible	Clinical (8)	30 mW/cm ²	No stimulation of bone formation at the distraction gap
Schortinghuis et al., 2005 ^[44]	Healing of mandibular defects covered with collagen membrane	Rat (64)	30 mW/cm ²	No significant benefit of US treatment
Erdogan et al., 2006 ^[38]	Mandibular fracture	Rabbits (30)	30 mW/cm ²	Stimulated reparative process at fracture site
El Bialy et al., 2006 ^[45]	Mandibular growth modification	Baboon (14)	30 mW/cm ²	Enhanced mandibular growth in growing baboons when US treatment was combined with functional appliances

the positive effects of US, controversial results also exist. Most fresh fractures or surgical approaches associated with bone tissue, such as bone distraction, heal eventually with conventional treatments. However, in cases when faster healing is desired by the patient or the physician, US therapy may be considered as an alternative approach.

There is evidence of successful treatments of non-unions and delayed unions with LIPU therapy.^[30,31] One main indication of US treatment may be this group of patients in the future. Osteoporosis is a major disease affecting bone tissue, especially in developed countries, which have an intense elderly population. Integration of biomaterials may be compromised in osteoporotic patients. Tanzer et al. demonstrated increased bone growth in porous intramedullar tantalum implants in the ulnae of dogs.^[34] One potential use of US treatment is to accelerate osseointegration of implants (in the treatment of fractures or dental implants) in osteoporotic patients. More animal and clinical studies are

needed in this area. US treatment also positively affects soft tissue wound healing.^[46] Complicated fractures that involve soft tissue and bone injury combined may benefit from US therapy.

Osteoradionecrosis is an avascular necrosis of bone, which develops as a consequence of radiotherapy. It is characterized by hypoxia and hypovascularity. The conventional treatment approach is surgical debridement with additional antibiotic and hyperbaric oxygen therapy. US treatment increases vascularity by inducing synthesis of angiogenesis-related cytokines.^[2] One study demonstrated that supportive US application helped considerably in the treatment of osteonecrosis of the jaw.^[35] US therapy may be considered a supportive treatment option for the treatment of osteonecrosis. Hypovascularity and hypoxia as a result of radiotherapy impair the success of reconstructive procedures in oncologic surgeries. One potential use of US is as a supportive procedure after reconstructive surgery in sites that receive radiotherapy.

Experimental and clinical studies are being conducted continuously to understand the mechanism of action and the effectiveness of US in clinical settings. Although the indication of use in clinical settings may be limited for a small group of patients, US therapy poses a significant treatment option in the management of bone pathologies and injuries mentioned above.

CONCLUSION

The use of US in medicine started for therapeutic but not for diagnostic purposes. Several theories explain the mechanism of action on bone healing. The studies mostly indicate that US therapy accelerates endochondral bone healing. Therapeutic US therapy in clinical settings is a non-invasive application and has no serious complications or side effects. It is an acceptable treatment of choice in fractures with delayed unions or non-unions. Although it does not seem possible to treat osteoporosis itself because of difficulties in application and a low penetration rate, its use in acceleration of osseointegration of metal biomaterials in osteoporotic patients has a promising future in orthopedic surgery and oral-maxillofacial surgery. The use of US in the treatment of fresh fractures or surgical interventions associated with bone tissue like spinal arthrodesis or distraction osteogenesis in healthy patients may not be worthwhile, because bone healing usually occurs uneventfully in these situations. Although much is known about the molecular effects of US treatment on bone healing, more studies are needed to fully understand the mechanism.

ARTICLES OF FURTHER INTEREST

Bone Fracture Fixation; Bone Remodeling; Tissue Engineering of Bone

REFERENCES

1. Halsscheidt, W.; Holfeld, A.; Rheinwald, U. Die Ultraschall-Therapie in der Zahn-, Mund- und Kieferheilkunde. *Zahnärztl Welt* **1949**, *22*, 572–279.
2. Maintz, G. Tierexperimentelle Untersuchungen über Wirkung der Ultraschallwellen auf die Knochenregenerationen. *Strahlentherapie* **1950**, *82*, 631–638.
3. Corradi, C.; Cozzolino, A. Azione degli ultrasuoni sulla evoluzione delle fratture sperimentali dei conigli. *Minerva Ortop.* **1952**, *3*, 44–45.
4. Corradi, C.; Cozzolino, A. Gli ultrasuoni (U.S.) e l'evoluzione del casso osseo nei focolai di frattura. *Arch. Ortop.* **1953**, *66*, 77–98.

5. Shiro, I. Study on the ultrasonic irradiation in orthopedic surgery. *Hirosaki Med. J.* **1964**, *16*, 242–253.
6. Dyson, M.; Brookes, M. Stimulation of bone repair by ultrasound. *Ultrasound Med. Biol.* **1983**, *2*, 61–66.
7. Xavier, C.A.M.; Duarte, L.R. Stimulation of bone callus by ultrasound. *Rev. Brasil Ortop.* **1983**, *18*, 73–80.
8. Duarte, L.R. The stimulation of bone growth by ultrasound. *Arch. Orthop. Trauma Surg.* **1983**, *101*, 153–159.
9. Pilla, A.A.; Mont, M.A.; Nasser, P.R.; Khan, S.A.; Figueiredo, M.; Kaufman, J.J.; Siffert, R.S. Non-invasive low-intensity pulsed ultrasound accelerates bone healing in the rabbit. *Orthop. Trauma* **1990**, *4*, 246–253.
10. Azuma, Y.; Ito, M.; Harada, Y.; Takagi, H.; Ohta, T.; Jingushi, S. Low-intensity pulsed ultrasound accelerates rat femoral fracture healing by acting on the various cellular reactions in the fracture callus. *J. Bone Miner. Res.* **2001**, *16*, 671–680.
11. Glazer, P.A.; Heilmann, M.R.; Lotz, J.C.; Bradford, D.S. Use of ultrasound in spinal arthrodesis. A rabbit model. *Spine* **1998**, *23*, 1142–1148.
12. Shimazaki, A.; Inui, K.; Azuma, Y.; Nishimura, N.; Yamano, Y. Low-intensity pulsed ultrasound accelerates bone maturation in distraction osteogenesis in rabbits. *J. Bone Joint Surg. Br.* **2000**, *82*, 1077–1082.
13. Heckman, J.D.; Ryaby, J.P.; McCabe, J.; Frey, J.J.; Kilcoyne, R.F. Acceleration of tibial fracture-healing by non-invasive, low-intensity pulsed ultrasound. *J. Bone Joint Surg. Am.* **1994**, *76*, 26–34.
14. Kristiansen, T.K.; Ryaby, J.P.; McCabe, J.; Frey, J.J.; Roe, L.R. Accelerated healing of distal radial fractures with the use of specific, low-intensity ultrasound. A multicenter, prospective, randomized, double-blind, placebo-controlled study. *J. Bone Joint Surg. Am.* **1997**, *79*, 961–973.
15. Parfitt, A.M. Targeted and nontargeted bone remodeling: Relationship to basic multicellular unit origination and progression. *Bone* **2002**, *30* (1), 5–7.
16. Kalfas, I.H. Principles of bone healing. *Neurosurg. Focus* **2001**, *15*, 10–14.
17. Kossoff, G. Basic physics and imaging characteristics of ultrasound. *World J. Surg.* **2000**, *24* (2), 134–142.
18. Ziskin, M.C. Fundamental physics of ultrasound and its propagation in tissue. *Radiographics* **1993**, *13* (3), 705–709.
19. Claes, L.; Willie, B. The enhancement of bone regeneration by ultrasound. *Prog. Biophys. Mol. Biol.* **2007**, *93* (1–3), 384–398.
20. Ter Haar, G.; Daniels, S.; Eastaugh, K.C.; Hill, C.R. Ultrasonically induced cavitation in vivo. *Br. J. Cancer Suppl.* **1982**, *45*, 151–155.
21. Dinno, M.A.; Dyson, M.; Young, S.R.; Mortimer, A.J.; Hart, J.; Crum, L.A. The significance of membrane changes in the safe and effective use of therapeutic and diagnostic ultrasound. *Phys. Med. Biol.* **1989**, *34*, 1543–1552.
22. Zorlu, U.; Tercan, M.; Ozyazgan, I.; Taskan, I.; Kardas, Y.; Balkar, F.; Ozturk, F. Comparative study of the effect of ultrasound and electrostimulation on bone healing in rats. *Am. J. Phys. Med. Rehabil.* **1998**, *77* (5), 427–432.



23. Reher, P.; Doan, N.; Bradnock, B.; Meghji, S.; Harris, M. Effect of ultrasound on the production of IL-8, basic FGF and VEGF. *Cytokine* **1999**, *11* (6), 416–423.
24. Rawool, N.M.; Goldberg, B.B.; Forsberg, F.; Winder, A.A.; Hume, E. Power Doppler assessment of vascular changes during fracture treatment with low-intensity ultrasound. *J. Ultrasound Med.* **2003**, *22* (2), 145–153.
25. Yang, K.H.; Parvizi, J.; Wang, S.J.; Lewallen, D.G.; Kinnick, R.R.; Greenleaf, J.F.; Bolander, M.E. Exposure to low-intensity ultrasound increases aggrecan gene expression in a rat femur fracture model. *J. Orthop. Res.* **1996**, *14* (5), 802–809.
26. Wang, S.J.; Lewallen, D.G.; Bolander, M.E.; Chao, E.Y.; Ilstrup, D.M.; Greenleaf, J.F. Low intensity ultrasound treatment increases strength in a rat femoral fracture model. *J. Orthop. Res.* **1994**, *12* (1), 40–47.
27. Hantes, M.E.; Mavrodontidis, A.N.; Zalavras, C.G.; Karantanas, A.H.; Karachalios, T.; Malizos, K.N. Low-intensity transosseous ultrasound accelerates osteotomy healing in a sheep fracture model. *J. Bone Joint Surg. Am.* **2004**, *86* (10), 2275–2282.
28. Mayr, E.; Rudzki, M.M.; Rudzki, M.; Borchardt, B.; Hausser, H.; Ruter, A. Does low intensity, pulsed ultrasound speed healing of scaphoid fractures? *Handchir. Mikrochir. Plast. Chir.* **2000**, *32*, 115–122.
29. Banken, R. Low-intensity ultrasound (Exogen™) for the treatment of fractures, Technology brief; Agence d'évaluation des technologies et des modes d'intervention en santé (AETMIS). AETMIS 03-05, Montreal, Canada, , 2004; 1–15.
30. Mayr, E.; Frankel, V.; Ruter, A. Ultrasound—An alternative healing method for nonunions? *Arch. Orthop. Trauma Surg.* **2000**, *120* (1–2), 1–8.
31. Nolte, P.A.; van der Krans, A.; Patka, P.; Janssen, I.M.; Ryaby, J.P.; Albers, G.H. Low-intensity pulsed ultrasound in the treatment of nonunions. *J. Trauma* **2001**, *51* (4), 693–702.
32. Tsumaki, N.; Kakiuchi, M.; Sasaki, J.; Ochi, T.; Yoshikawa, H. Low-intensity pulsed ultrasound accelerates maturation of callus in patients treated with opening-wedge high tibial osteotomy by hemicallotaxis. *J. Bone Joint Surg. Am.* **2004**, *86* (11), 2399–2405.
33. El-Mowafi, H.; Mohsen, M. The effect of low-intensity pulsed ultrasound on callus maturation in tibial distraction osteogenesis. *Int. Orthop.* **2005**, *29* (2), 121–124.
34. Tanzer, M.; Kantor, S.; Bobyn, J.D. Enhancement of bone growth into porous intramedullary implants using non-invasive low intensity ultrasound. *J. Orthop. Res.* **2001**, *19* (2), 195–199.
35. Harris, M. The conservative management of osteoradionecrosis of the mandible with ultrasound therapy. *Br. J. Oral Maxillofac. Surg.* **1992**, *30*, 313–318.
36. El-Bialy, T.H.; Royston, T.J.; Magin, R.L.; Evans, C.A.; Zaki, Ael-M.; Frizzell, L.A. The effect of pulsed ultrasound on mandibular distraction. *Ann. Biomed. Eng.* **2002**, *30* (10), 1251–1256.
37. El-Bialy, T.; El-Shamy, I.; Graber, T.M. Repair of orthodontically induced root resorption by ultrasound in humans. *Am. J. Orthod. Dentofacial Orthop.* **2004**, *126* (2), 186–193.
38. Erdogan, O.; Esen, E.; Ustun, Y.; Kurkcu, M.; Akova, T.; Gonlusen, G.; Uysal, H.; Cevlik, F. Effects of low-intensity pulsed ultrasound on healing of mandibular fractures: An experimental study in rabbits. *J. Oral Maxillofac. Surg.* **2006**, *64* (2), 180–188.
39. Fedotov, S.N.; Minin, E.A.; Borisov, I.N. Effect of local cooling and ultrasound on the reparative processes following mandibular fracture. *Stomatologiia* **1986**, *65* (4), 4–6.
40. El-Bialy, T.H.; el-Moneim, Z.A.; Evans, C.A. Effect of ultrasound on rabbit mandibular incisor formation and eruption after mandibular osteodistraction. *Am. J. Orthod. Dentofacial Orthop.* **2003**, *124*, 427–434.
41. El-Bialy, T.H.; El-Shamy, I.; Graber, T.M. Growth modification of the rabbit mandible using therapeutic ultrasound: Is it possible to enhance functional appliance results? *Angle Orthod.* **2003**, *73*, 631–639.
42. Schortinghuis, J.; Ruben, J.L.; Raghoobar, G.M.; Stegenga, B. Ultrasound to stimulate mandibular bone defect healing: A placebo-controlled single-blind study in rats. *J. Oral Maxillofacial Surg.* **2004**, *62*, 194–201.
43. Schortinghuis, J.; Ruben, J.L.; Raghoobar, G.M.; Stegenga, B. Therapeutic ultrasound to stimulate osteoconduction; A placebo controlled single blind study using e-PTFE membranes in rats. *Arch. Oral Biol.* **2004**, *49*, 413–420.
44. Schortinghuis, J.; Bronckers, A.L.; Stegenga, B.; Raghoobar, G.M.; de Bont, L.G. Ultrasound to stimulate early bone formation in a distraction gap: A double blind randomised clinical pilot trial in the edentulous mandible. *Arch. Oral Biol.* **2005**, *50*, 411–420.
45. El-Bialy, T.; Hassan, A.; Albaghdadi, T.; Fouad, H.A.; Maimani, A.R. Growth modification of the mandible with ultrasound in baboons: A preliminary report. *Am. J. Orthod. Dentofacial Orthop.* **2006**, *130* (4), 435–437.
46. Young, S.R.; Dyson, M. Effect of therapeutic ultrasound on the healing of full-thickness excised skin lesions. *Ultrasonics*. **1990**, *28* (3), 175–180.

Urinary Prostheses

Daniel Eberli
Anthony Atala

Wake Forest Institute for Regenerative Medicine, Winston Salem, North Carolina, U.S.A.

INTRODUCTION

Modern urologic prostheses include artificial sphincters, penile prosthesis, testicular implants, catheters, and stents. The urological use of artificial materials has a long history, starting in 3000 B.C. when Egyptians used papyrus and gold catheters to treat urinary retention. Over the next decades catheters were made from lead, treated paper, and, after the turn of the 19th century, from natural rubber. Only in the early 1970s were more sophisticated hydraulic prosthesis for the treatment of erectile dysfunction and urinary incontinence introduced.

A biomaterial used for urologic prostheses should incorporate certain features, such as biocompatibility and stability after placement, resistance to infection, chemical stability in urine, no carcinogenicity or antigenicity, ability to sterilize, and availability at an affordable price. Although not perfect, silicone and polyurethanes are the biomaterials currently used for urological implants. Both silicone and polyurethanes have excellent biocompatibility and have been successfully used in clinics for over 30 years. However, the use of hydraulic implants for the treatment of erectile dysfunction and incontinence are still considered the last treatment option if pharmacological treatment fails. Serious complications, including infection or malfunction, require instant removal or replacement of the prosthesis. On the other hand, patients with successfully implanted prostheses report the highest satisfaction scores. Careful patient selection is crucial for the success of the implanted device, since the patient has to learn how to use his prosthesis.

This chapter will critically discuss urological prosthesis in regard to biomaterials, indications, surgical treatment, and future perspectives. Furthermore, it will emphasize current clinical problems associated with the use of prostheses in the hope that upcoming research might help to overcome these obstacles.

BIOMATERIALS FOR URINARY PROSTHESES

Urological implants are made of a relatively small variety of materials including silicone and polyurethanes.

Other materials such as latex, polytetrafluoro ethylene, or hydrogels are commonly used in stents and catheters, often as coatings. Generally speaking, polyurethanes have good mechanical properties and poor stability in the body, while silicones have poor mechanical properties, but excellent stability in the body. Therefore, both materials have to be modified to meet the requirements for use as urinary prostheses. This can be achieved by chemical modification of the polymer, by mixing with additives or by using a multilayer design.

Biomaterial–Tissue Interaction

Contact of any biomaterial with tissue leads to an inflammatory reaction in varying degrees. The body reacts by activating the innate and specific immune system, the complement system, and coagulation cascades. This inflammatory response is influenced by surface characteristics of the biomaterial, the reaction to denatured proteins adsorbed on the surface of the biomaterial, and the products of degradation.

In the early phase, the adsorption of albumin, fibrinogen, and other proteins trigger the host defense. This adsorption of protein further decreases the surface tension allowing for more protein to adhere. Some urologic prostheses reduce the tissue response by utilizing a hydrophilic coating and a smooth finish to decrease the adherence of protein. Nevertheless, all biomaterials become encapsulated over time. Encapsulation is a chronic inflammatory response of the host towards any biomaterial. The fear that this chronic inflammation might trigger malignancy is discussed later in this chapter.

In summary, silicone and polyurethanes show excellent biocompatibility with minimal effect on surrounding tissue, and are therefore preferred biomaterials for urologic prostheses.

Silicone

Silicone has been chosen for the majority of urologic prosthetics because it is relatively inert, easily sterilized, maintains its strength in vivo. Silicone, or the more scientific name, Siloxane, was first discovered as a waste product during the synthesis of other polymeric



materials and has been known for over 100 years. The driving force behind the exploration was the exceptional stability in the biological environment, compared to other elastomers showing more degradation when left in vivo for long periods. The first commercial production began in 1943.^[1]

The general molecular formula for silicone polymers is R_2-SiO , where R is a carbon based side group, and the polymer chain consists of a $-(SiO)-$ backbone. The viscosity of silicone can be tailored by changing the chain length, from liquid to gum. In prostheses, silicone polymers are used in three major forms, including oils, gels, and elastomers. Silicone oils consist of mixtures of short chain polymers. Silicone gel is a mixture of silicone oil and crosslinked elastomers. Elastomers are made by combining long chains of silicone and ethyl cross-links between the chains.

Compared with other commercially available polymers, the mechanical properties of silicone are relatively poor, but the preservation of its properties in vivo has led to its widespread use. Typical tensile strength of silicone used in medical devices is 4.9–7.4 MPa, and the tensile modulus is 2.2 MPa.^[1]

Polyurethanes

The second most used material for urological prosthesis is polyurethane (Bioflex[®]). Polyurethane has a tensile strength of 52 MPa and a tensile modulus of 3.4 MPa,^[1] thus is far stronger than silicone. The elongation of silicone and polyurethane is comparable, allowing for 600% elongation before tearing.

The backbone of polyurethanes contains the urethane linkage $-OC(O)NH-$. The characteristics of the polymer are given by the functional groups such as polyethers, polyesters, ureas, silicones, and hydrocarbon groups. The exceptional mechanical properties are due to excellent secondary hydrogen bonding. Generally speaking, polyurethanes have good mechanical properties and poor stability in the body. Therefore, polyurethanes have to be modified to meet the criteria for urologic implants. The composition of polyurethanes used in medical devices (Bioflex[®]) is proprietary. Therefore, details about the formulation are not known to the public. In general terms, Mentor describes the improvement as due to two factors: a changed polyurethane chemistry and additives that optimize in vivo durability. Studies have shown that the polyurethane used by Mentor is more durable and less elastic than silicone.^[2]

Possible Reactions to Silicone and Other Thermoplastic Elastomers

The issue of the possible relationship between various diseases and silicone has been, and continues to be, the

subject of great scientific and medical debate. The legal battle is mainly focused on implants for augmentation and reconstructive mammoplasty, but since the same biomaterial is used for urologic prostheses, the problematic situation applies here as well. The medical literature has raised the possibility that there may be an association between certain immunological-based diseases and silicone implants. The diseases most commonly mentioned include scleroderma, rheumatoid arthritis, and systemic lupus erythematosus. Neurological problems have been reported in a small number of breast implant patients who also exhibit immunological symptoms. Available information does not permit precise quantification of risk. In addition, women with connective tissue disorders have reported a reduction in symptoms after their breast implants were removed. More scientific studies are needed to prove a link between the implants and immunological or neurological problems.

Even more formidable is the association of biomaterials and tumor development. The literature suggests that in vivo degradation and particle shedding of silicone elastomers may play an important role in maintaining chronic inflammatory response and encapsulation.

Since certain malignancies can be clearly linked to chronic inflammation, including gastric cancer and esophageal cancer, this problem needs serious attention. Case reports in the medical literature have associated tumors with the presence of silicone mammary implants. However, the number of tumors found in the vicinity of implants is only one to two cases globally every year.^[3] This number is very small in relation to the millions of implants performed every year. Although, the individual cancer risk is minimal, concerns among patients still remain. During the past two decades of clinical use, the medical literature generally indicates that silicone prostheses are not carcinogenic, but the debate is still ongoing.

In summary, an association between silicone and the aforementioned conditions and complications has not been verified by controlled scientific studies. Urological devices made from silicone and polyurethane should be labeled safe for clinical use. Nevertheless, patients concerns must be included in the decision-making process before surgery to ensure a favorable outcome.

PENILE PROSTHESES

Overview

The National Institute of Health estimates that 10–20 million American men are impotent. The treatment options today include oral therapy, vacuum devices,

penile injection therapy, and prosthetic surgery. Penile prosthesis is usually offered if pharmacological therapy fails. Historically, the first penile prosthesis used cartilage and bone to improve rigidity. The approach was inspired by the observation that the “penis bone” of certain species consists of cartilage. A report of 1948 Bergman et al. described the use of a rib cartilage implant placed into a tube of skin graft.^[4] Despite initially high success rates, the use of cartilage had many disadvantages, including substantial harvest site morbidity, difficulty getting the right size, and variable absorption. Therefore, rib cartilage was replaced with the first synthetic rods made of perforated acrylic or polyethylene. The material of these early artificial implants was later replaced by silicone, an inert biomaterial with the desired flexibility and rigidity. A hydraulic inflatable prosthesis allowing for adjustment of the rigidity, from flaccid to firm, was introduced in 1973.^[5] Further research and development was needed to decrease the high rate of mechanical failure of these early inflatable implants. Modern prostheses are able to resolve the rigidity problem and penetration difficulties, with no change in libido or orgasm characteristics, and there is no increase in size. These facts have to be clearly communicated to the patient to prevent an unsatisfactory outcome of the procedure. With careful patient selection and successful implantation, penile prostheses are associated with a high rate of patient satisfaction.^[6]

Types of Penile Prostheses

There are currently two types of penile prostheses available on the market, including semirigid and inflatable prostheses with one, two, or three volumes.

Semirigid Prostheses

The most commonly used semirigid implants (Fig. 1) are the Mentor Acu-Form (Mentor, Santa Barbara, CA) and the AMS 600–650 (American Medical Systems, Minnetonka, Minnesota). During surgery, a pair of these malleable silicone elastomer rods are directly placed into the corpora cavernosa. These implants consist of twisted stainless steel or silver wires wrapped in a synthetic fabric surrounded by a silicone jacket. One major drawback of this design is the problem of “spring-back” after bending. Newer models reduce the elasticity by increasing the number of cable strands reducing the “spring-back” and improving mechanical reliability. All implants are offered in various diameter and lengths to fit the anatomical specifications of each patient.

In comparison to inflatable implants, semirigid prostheses are easier to insert since there is no pump or

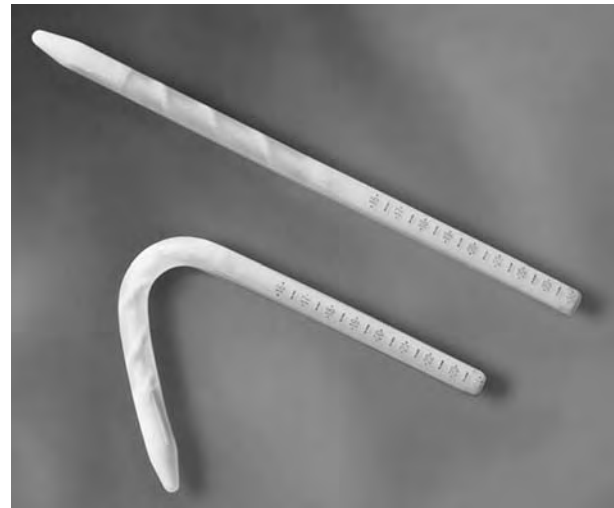


Fig. 1 Mentor genesis Semirigid implants with malleable silicone elastomer rods. (Courtesy of Mentor Corporation, Santa Barbara, CA).

reservoir to place. Furthermore, the semirigid implants are more widely accepted internationally because they are substantially less expensive than the inflatable implants.

AMS recently presented a more advanced semirigid implant (Dura-II). This device consists of a series of articulating polypropylene discs connected with a metal cable running through the center of the discs, resembling to a gooseneck lamp.

This positional memory allows the implant to be rigid during intercourse and to remain concealed when placed along the leg.

Inflatable Implants

Inflatable implants are manufactured as two- or three-piece prostheses. The two-piece implants include a pair of inflatable cylinders implanted in the corpora cavernosa and a small pump for scrotal implantation. The three piece prostheses (Fig. 2) are supplemented with an additional fluid reservoir, which is typically positioned intraperitoneally. All components are connected by tubes. Currently medical implants have nylon reinforced tubing to make kinking virtually impossible and are equipped with patented locking connectors for easy connections.

In order to increase rigidity of the penis, saline solution is moved from the reservoir to the cylinder shafts by squeezing and releasing the scrotal pump several times. Modern inflatable implants, like the AMS 700CX, are designed in a three layer fashion containing a core made of elastic silicone, a middle layer of woven monofilament (90% polyethylene and



Fig. 2 AMS 700 CX™ penile prosthesis Inflatable three piece prostheses, including a pair of inflatable cylinders implanted in the corpora cavernosa, a small pump for scrotal implantation and fluid reservoir which is typically positioned intraperitoneally. (Courtesy of American Medical Systems, Inc., Minnetonka, MN).

10% polyurethane) and an outer silicone layer for tissue interaction. The middle woven monofilament knitted layer is designed to permit controlled expansion of the cylinders. The fibers can either be woven bidirectionally, permitting length and girth expansions, or unidirectionally, permitting only girth expansion. Cylinders that expand more in girth than in length tend to straighten penile curvature better and are recommended for use in patients with significant scarring, such as that caused by priapism (prorogued erection), Peyronies disease (disorder with local corpora fibrosis) or reimplantation of a prosthesis. Finally, the inner elastic silicone core provides rigidity if filled with fluid.

The implants manufactured by Mentor are very similar in design and function to the AMS prosthesis. The main difference is that Mentor uses a polyether urea urethane elastomer called Bioflex[®], which has a higher tensile strength and is more resistant to abrasion compared to silicone.

In an effort to simplify surgical placement, both manufacturers presented two-piece inflatable prostheses that combine the reservoir with one of the other components. The AMS Ambicor prosthesis has the reservoir located at the proximal end of the cylinder. For activation, saline is pumped from the proximal reservoir to the distal inflatable chamber using the pump located in the scrotum. The Mentor Mark II has a different design and combines the pump and the reservoir. To conceal the relatively large pump reservoir, the device can be implanted between the testicles. Both models are prefilled with saline and require no tubing connections, reducing operation time and infection risk.

The original Mentor two-piece prosthesis was prone to early mechanical malfunction due to tissue

encapsulation of the pump-reservoir system and subsequent contracture of the capsule. The newer Mentor models have a textured surface added to the reservoir pump to reduce capsule formation.

In a further effort to reduce the number of volumes needed, AMS introduced a self-contained inflatable prosthesis in 1985, usually referred to as a one-piece penile implant. Unfortunately, mechanical failure rate of the original one-piece prosthesis was reported to be 79%, and the long term patient satisfaction was only 43%. The overall explantation rate of the self-contained inflatable prosthesis was as high as 55%. Studies comparing one piece prosthesis with multi-component inflatables noted a higher patient satisfaction and lower revision rate for the multi-component prosthesis. The multi-component inflatable prosthesis is the device of choice for patients requiring prosthetic surgery for erectile dysfunction.

Indications

The introduction of oral therapy (phosphodiesterase inhibitors) radically changed the management of the treatment for erectile dysfunction, leading to a reduced surgical indication. Penile prosthesis still remains the treatment of choice for patients with organic-factor related erectile dysfunction, including unsatisfactory results with other therapy modalities, contraindications to oral treatment, and rejection by the patient.

Out of the 25% nonresponders to oral pharmacotherapy, approximately 15% will be offered intracavernosal injection therapy, a modality with a high drop-out rate. Therefore, approximately 15% of patients seeking treatment for erectile dysfunction have severe and irreversible damage of the erectile mechanism and are candidates for implantation surgery.^[7] Candidates for treatment with penile prostheses should be psychologically stable, in good general health and fully informed. This is particularly important since implantation surgery is associated with significant risk of complications and is irreversible.

Success and Problems

Outcome

Among all treatment options for erectile dysfunction, penile prostheses are associated with the highest success rates. The overall published rates of satisfied patients and partners are 70%–85% and 60%–70%, respectively. A recent study compared erectile function and satisfaction rates in patients who received treatment for erectile dysfunction with sildenafil (Viagra[®]), intracavernosous prostaglandin E1 injection therapy or penile implant surgery.^[6] The Erectile Dysfunction

Inventory for Treatment Satisfaction (EDITS) and the erectile function domain score of the IIEF were used to assess the outcome. After a mean follow-up of almost two years, patients who underwent penile implant surgery had significantly higher scores than patients treated with sildenafil or intracavernous injections. Further studies using a psychosexual questionnaire found a significant alleviation of concerns about achieving and maintaining an erection during intercourse after penile implant surgery.

In summary, these results show that penile prostheses can be highly successful in improving erection, ejaculation, orgasm and sexual satisfaction.

Infections

Infection associated with the penile prosthesis is a serious event that usually necessitates the immediate removal of the device. Most infections occur as a result of contamination of the operative wound and subsequent colonization of the prostheses at the time of implantation. The most common bacteria found on infected prosthesis is *Staphylococcus epidermidis* and *Staphylococcus aureus*,^[7] both common skin inhabitants. Fortunately, the infection rates with more aggressive organisms such as methicillin-resistant *S. aureus* or *Pseudomonas aeruginosa* are relatively low, in the range of 1%–10% of cases.

Recent studies found the overall incidence of infection to be approximately 10% for primary implants. Patients with deficiency of the immune system are more prone to infection. Therefore, patients under treatment with immunosuppressive drugs or patients with poorly controlled diabetes mellitus should postpone their surgery until the underlying condition has stabilized.

Surprising results were revealed in a prospective study analyzing the infection rate of removed prostheses. Seventy percent of patients with clinically uninfected penile prostheses had culture positive bacteria; most of them were a *Staphylococcus* species.^[8]

This study shows that colonization of prostheses is far more common than clinically expected. Unfortunately, we do not yet understand why some patients develop a fulminate infection while others show no clinical signs even though the implant is colonized with bacteria.

Both manufacturers of penile prostheses made significant efforts to reduce the high rate of infected implants. AMS added a coating called InhibiZone™ containing a sustained release formula of Rifampicin and Minocycline to inhibit bacterial growth. Rifampicin shows a significant decrease after one day, while Minocycline decreases after seven days. The InhibiZone™ coating proved to be most efficient against *S. epidermidis* and *S. aureus*, both bacteria commonly found

on infected prostheses. The conclusion of a recent study including more than 4000 patients was that InhibiZone™ lowers the infection rate significantly by 60% compared to noncoated prostheses.^[9] It is important to understand that both antibiotics do not reach labeled doses for active infection, and therefore a conventional prophylactic antibiotic protocol should be administered before surgery.

Mentor has chosen a different cause of action and changed the surface structure of their penile prosthesis to reduce bacterial adherence. The extremely hydrophilic coating called Resist™ has been shown to decrease bacterial adherence in vivo.^[10] Furthermore, this coating can absorb antibiotics and release them over 24–72 h. An antibiotic-soaking solution, composed of 100,000 Units of Bacitracin and 2 g Gentamicin in 1000 ml of saline has been successfully tested and is in clinical use.

In general, it is important that the introduction of microbe-resisting coating does not result in the surgeon reducing his vigilance or sterile technique. Procedures which already have been established to reduce the incidence of infection are still of crucial importance. Therefore, administering of prophylactic antibiotics two hours prior to surgery, thorough shaving in the immediate preoperative period, and opening of the prosthesis package as late as possible should still be practiced.

Future Perspectives: Regenerative Medicine

The shortcomings of synthetic biomaterials have led to the application of regenerative medicine and tissue engineering to the field of penile reconstruction. Regenerative Medicine encompasses various areas of technology, such as tissue engineering, stem cells, and cloning. Tissue engineering, one of the major components of regenerative medicine, follows the principles of cell transplantation, materials science, and engineering towards the development of biological substitutes that can restore and maintain normal function. Tissue engineering strategies generally fall into two categories: the use of acellular matrices, which depend on the body's natural ability to regenerate for proper orientation and direction of new tissue growth, and the use of matrices with cells. Acellular tissue matrices are usually prepared by manufacturing artificial scaffolds, or by removing cellular components from tissues via mechanical and chemical manipulation to produce collagen-rich matrices. These matrices tend to slowly degrade on implantation and are generally replaced by the extracellular matrix (ECM) proteins that are secreted by the ingrowing cells.

When cells are used for tissue engineering, a small piece of donor tissue is dissociated into individual cells.



These cells are expanded in culture, attached to a support matrix, and then reimplanted into the host after expansion. The source of donor tissue can be heterologous, allogeneic, or autologous. Ideally, both structural and functional tissue replacement will occur with minimal complications. The preferred cells to use are autologous cells, where a biopsy of tissue is obtained from the host; the cells are dissociated and expanded in culture, and the expanded cells are implanted into the same host.^[11-14] Although, it may cause an inflammatory response, the use of autologous cells avoids rejection. Thus, the deleterious side effects of immunosuppressive medications can be avoided.

The first successful studies, applying the methods of tissue engineering to the treatment of erectile dysfunction, engineered human cartilage rods that could be used in patients requiring a penile prosthesis.^[15] These human cartilage rods were created by transplanting chondrocytes harvested from a small biopsy of the cartilaginous part of a human ear on polymer scaffolds. Mechanical tests showed that the engineered cartilage rods exhibited the appropriate mechanical properties required for use as natural penile prostheses. In consecutive studies the same research group started to engineer functional corpora tissue. After establishing the cell expansion technique for penile endothelial cells and smooth muscle cells^[16] a segmental replacement of corporal tissue was successfully performed in a rabbit model.^[17] Autologous endothelial cells and smooth muscle cells were seeded on acellular tissue matrices and the constructs implanted into a penile defect. This study demonstrated the possibility of engineering autologous penile corpora cavernosa tissue in a rabbit model. The engineered tissue was able to achieve adequate structural and functional parameters as proven in histology and functional measures (Cavernosometry). A study evaluating the total penile replacement is currently ongoing. Preliminary data presented at international meetings are very promising.

In summary, tissue engineering might offer a novel treatment option for patient requiring a penile prosthesis and phallic reconstruction. The use of the patient's own cells to engineer a functional corpus would clearly present a major medical breakthrough.

ARTIFICIAL URINARY SPHINCTER

Overview

Urinary incontinence is associated with impairment of quality of life, social isolation, and depressive symptoms.^[18] A conservative estimate is that urinary incontinence affects approximately 20% of women. However, the prevalence may be as high as 50% in elderly women with a peak at 47 years of age.^[19-21] In

the United States the estimated annual direct cost of caring for patients with urinary incontinence is over \$16 billion.^[22]

In women, trauma during childbirth is the main reason for developing severe urinary incontinence requiring an artificial urinary sphincter. In male patients, the main cause of incontinence is surgery for prostate disease.

In general, treatment plans are usually directed toward the specific type of urinary incontinence; the treatment option with the lowest risk for adverse complications is usually offered. Lifestyle intervention, pelvic floor muscle training, vaginal devices (pessaries), and pharmacological treatments are usually considered prior to surgical intervention.^[18] Surgical treatment options include the endoscopic injection of bulking agents, colposuspension (e.g., Burch), sling operations, traditional transvaginal needle suspension, tension-free vaginal tape (TVT), and artificial urinary sphincter.

The first non-hydraulic artificial urinary sphincter composed of silicone and stainless steel was introduced in 1973. Since then major design changes have significantly improved effectiveness and patient satisfaction. At first, the four-component model consisted of an inflatable cuff, a fluid reservoir, and two pumps: one to inflate and the other to deflate. Due to the large number of components, this first device was difficult to implant and prone to mechanical failure.^[23]

Types of Artificial Urinary Sphincters

Currently, one type of artificial urinary sphincter is dominating the market (Fig. 3). American Medical Systems (Minnetonka, Minnesota) manufactures the three-piece artificial sphincter (AMS 800) consisting of a deflation pump, a pressure regulating reservoir, and an inflatable cuff. The pump is usually placed in the scrotum, and the reservoir is placed in the lower abdomen. Finally, the cuff can be implanted either at the bladder neck or the bulbous urethra. All pieces are interconnected with reinforced silicone tubing and the system is filled with sterile saline solution. To initiate micturation, the pressure in the cuff is released by squeezing the deflation pump located in the scrotum. The pump is constructed to allow for slow return of fluid to the cuff and automatically raises the pressure around the urethra after micuration, usually within 3 min. By choosing one of the three different pressure regulating reservoirs, the pressure within the system can be adjusted. The surgery is usually done under general or spinal anesthesia and is completed within 90 min. Following surgery the device is locked in an open condition (deactivated) to allow proper healing of the wound. Four to six weeks after surgery, the device



Fig. 3 AMS 800™ artificial urinary sphincter Three-piece artificial sphincter consisting of a deflation pump, a pressure regulating reservoir and an inflatable cuff. (Courtesy of American Medical Systems, Inc., Minnetonka, MN).

is usually activated and the patient educated in the handling of the device.

Indications

The current artificial urinary sphincter is suitable for difficult incontinence in both sexes. More than 90,000 men worldwide received an AMS 800 within the last 30 years. The majority of implants were used for the treatment of incontinence after removal of the prostate due to prostate cancer. The anatomical proximity of the urinary sphincter and the prostate makes it difficult to remove the entire cancer without harming the sphincter muscle. The indication in woman is severe urinary stress incontinence with involuntary urinary leakage in a lying position. Numerous reports indicate that the artificial sphincter implantation is well tolerated and provides satisfactory results for the treatment of severe urinary incontinence in both genders.

Success and Problems

Artificial sphincters face the same problems as penile implants, including malfunction, erosion, and infection. These complications are severe since they require immediate surgical revision, often including explantation of the prosthesis. The level of mechanical malfunction of the improved AMS 800 series have been reported as 7%–16%.^[24,25] Early reports in

female patients with stress urinary incontinence described urethral erosion rate around the cuff at 10%–26%.^[26,27] More recent reports suggest that, if care is taken during implantation, the erosion rate can be minimized.^[27] The overall surgical revision rate is reported to be 29% with a follow up of 7.7 years.^[25] Therefore, patients have to be fully informed that up to one-third will need revision surgery in the years to come.

However, the artificial sphincter is associated with a high long-term success rate. A recent study evaluating men who received a prosthesis found that after seven years 80% were still satisfied with the device and achieved a social continence (requiring zero to one pad per day).^[28] Another long-term study evaluated the use of artificial sphincters for the treatment of post-prostatectomy incontinence and found that 90% of patients reported satisfaction and 96% would recommend the device to a friend.^[29]

Future Perspectives: Regenerative Medicine

Many of the current surgical treatment options are limited by the postoperative complications, including infection, graft erosion, bladder and bowel perforation, prologued catheterization, and vascular injury.^[21] In addition, 5%–10% of patients undergoing surgery for stress urinary incontinence will have worsening due to new onset of detrusor over activity, causing urge incontinence.^[21] The pelvic diaphragm prevents urinary incontinence by providing both resting urethral tone via slow-twitch fibers and rapid re-rectory contraction via fast-twitch fibers if the abdominal pressure rises. Therefore, the external striated urethral sphincter is primarily responsible for preventing stress urinary incontinence. The potential use of injectable cultured myoblasts (muscle precursor cells, MPC) for the treatment of stress urinary incontinence has been investigated and has the potential to become the first treatment to restore sphincter muscle function.^[30] Labeled myoblasts were directly injected into the proximal urethra and lateral bladder walls of nude mice with a micro-syringe in an open surgical procedure. Tissue harvested up to 35 days post-injection contained the labeled myoblasts, as well as evidence of differentiation of the labeled myoblasts into regenerative myofibers. The authors reported that a significant portion of the injected myoblast population persisted in vivo.^[31] A subsequent study demonstrated a significant improvement in leak point pressure when sciatic nerves transected rats were treated with myoblast injections.^[32] The fact that myoblasts can be labeled and survive after injection and begin the process of further myogenic differentiation supports the feasibility of using cultured cells of

muscular origin as an injectable bioimplant. Muscle precursor cells are the quiescent satellite cells found in each myofiber that proliferate to form myoblasts, and eventually myotubes and new muscle tissue. Intrinsic MPC have previously been shown to play an active role in the regeneration of injured striated urethral sphincters.^[33] In a subsequent study, autologous MPC were injected into a rat model of urethral sphincter injury, and both replacement of mature myotubes, as well as restoration of functional motor units, was noted in the regenerating sphincteric muscle tissue.^[34]

Preclinical trials for the application of this technology are currently being conducted in large animals. Sphincteric derived muscle cells have been used for the treatment of urinary incontinence in a pig model.^[35] The damaged sphincter was visualized by transurethral ultrasound and the MPC and fibroblasts injected into the damaged zone. Our institution is performing a preclinical study in dogs; animals that received a muscle precursor cell injection after damage of the sphincter showed a significant improvement in leak point and sphincter pressure compared to controls.

In summary, these results are promising and the first clinical trials are within reach. Muscle precursor cells may be a minimally invasive solution for urinary incontinence in patients with irreversible urinary sphincter muscle insufficiency.

TESTICULAR PROSTHESIS

Overview

Testicular loss occurs for various reasons, including orchiectomy for cancer, torsion, or agenesis. The scrotum plays an important psychological role in the male body image. Therefore, prosthetic devices have been developed to restore the normal appearance of the scrotum and to improve quality of life and psychological well-being. The first testicular prosthesis was implanted in 1939 using a metal alloy vitalium 1939.^[36] Over the years, glass spheres, polymerized methyl methacrylate, methacrylate, and polyurethane foams were introduced with the goal to be a cosmetically more acceptable testicular prosthesis. The use of silicone-gel filled silicone implants in the early 1970s was a major improvement,^[37] providing similar touch as a native testicle. Silicone polymer prostheses were then widely used for over 20 years. The production stopped in 1995 because of the concerns over the association of silicone implants with connective tissue disease.^[38] Although the expert panels in the United States (The National Science Panel and the Institute of Medicine) failed to find evidence for any causal linkage, the manufacturer designed a silicone-gel free testicular prosthesis. This saline-filled testicular prosthesis



Fig. 4 Mentor testicular prosthesis Saline filled testicular prosthesis with a silicone polymer shell and recessed suture tab, which allows the device to be fixed in place during surgical placement. (Courtesy of Mentor Corporation, Santa Barbara, CA).

with a silicone polymer shell (Mentor) remains the gold standard up to date (Fig. 4). The manufacturer offers four different sizes XS (5 ml), S (8 ml), M (11 ml), and L (15 ml). The silicone polymer shell has two reinforced regions. At the upper pole, the silicone shell is thicker to prevent leakage after filling the prosthesis, and at the lower pole there is a recessed suture tab that allows the device to be fixed in place during surgery. This prosthesis should not be considered a lifetime implant but can often remain without exchange or removal for decades.

Indications

Testicular implants should be offered to any patient with agenesis of a testicle or surgical removal due to cancer or torsion. An implant can restore the normal appearance of the scrotum and can improve self-esteem. It clearly remains a patient's choice whether he

wants an artificial implant or prefers to live with only one testis. In our experience, older men tend to deny the implant more often, while younger men decide to have a prosthesis implanted. Despite the excellent results and improvement of quality of life, every foreign body carries the risk of infection even years after implantation.

There are no strong contraindications except the presence of a soft tissue infection at the time of placement. In addition, the manufacturer mentions the possibility of persistent pain in patients with preexisting varicoceles, a disease where venous backflow causes dilatation of the venous plexus surrounding the testicle.

The placement can be performed on an outpatient basis. To prevent device infection the patient is given perioperative antibiotics, which are repeated for several days after surgery. Additionally, the prosthetic device is bathed in antibiotic solution and then filled with normal saline. The prosthesis is placed into the scrotal pouch and usually fixed to the surrounding fascia through the suture tab.

Success and Problems

Silicone polymer testicular prostheses have a wide acceptance and are used around the world. A recent study^[39] evaluated, prospectively, 149 patients with testicular implants. After a follow up of one year, only three patients (2%) required additional surgery due to major complications, including device extrusion and device migrations. Additionally, this study was able to show a significant improvement in self-esteem measured by a well-accepted questionnaire. The prosthesis infection rate in this group was zero and only one patient had a superficial wound infection. This might be due to the strict antibiotics regime, the low complexity of the device, and the surgical procedure. In summary, testicular prosthesis is safe and should be offered to any patient with testicular agenesis or surgical removal. Numerous investigators have reported the cosmetic benefits provided by testicular implants, which lead to psychological benefits and high levels of patient satisfaction.

CONCLUSION

Urologic prostheses range from complex hydraulic penile prostheses or artificial urinary sphincters to rather simple-saline filled testicular prosthesis. Biomaterials used for these implants are silicone and polyurethane, both materials with exceptional biocompatibility. The more complex hydraulic prostheses have a high rate of surgical revisions due to mechanical

failure, erosions, and infections. Manufacturers are constantly seeking to decrease the revision rate by improving surface texture, adding antibiotics, and improving the overall design. If implanted successfully, urological prostheses are associated with very high satisfaction rates.

ARTICLES OF FURTHER INTEREST

Silicones; Tissue Engineering of Bladder

U

REFERENCES

1. Wironen, J.; Marotta, J.; Cohen, M.; Batich, C. Materials used in urological devices. *J. Long Term Eff. Med. Implants* **1997**, *7*, 1.
2. Merrill, D.C. Mentor inflatable penile prosthesis. *Urology* **1983**, *22*, 504.
3. Seiter, H.; Schmitz, K.P.; Behrend, D. Biomaterials in urology. *Urologe A* **2000**, *39*, 463.
4. Bergman, R.T.; Howard, A.H.; Barnes, R.W. Plastic reconstruction of the Penis. *J. Urol.* **1948**, *59*, 1174.
5. Montague, D.K. Penile prostheses. An overview. *Urol. Clin. North Am.* **1989**, *16*, 7.
6. Rajpurkar, A.; Dhabuwala, C.B. Comparison of satisfaction rates and erectile function in patients treated with sildenafil, intracavernous prostaglandin E1 and penile implant surgery for erectile dysfunction in urology practice. *J. Urol.* **2003**, *170*, 159.
7. Moncada, I.; Martinez-Salamanca, J.I.; Allona, A.; Hernandez, C. Current role of penile implants for erectile dysfunction. *Curr. Opin. Urol.* **2004**, *14*, 375.
8. Henry, G.D.; Wilson, S.K.; Delk, J.R.2.; Carson, C.C.; Silverstein, A.; Cleves, M.A.; et al. Penile prosthesis cultures during revision surgery: a multicenter study. *J. Urol.* **2004**, *172*, 153.
9. Carson, C.C. Efficacy of antibiotic impregnation of inflatable penile prostheses in decreasing infection in original implants. *J. Urol.* **2004**, *171*, 1611.
10. Hellstrom, W.J.; Hyun, J.S.; Human, L.; Sanabria, J.A.; Bivalacqua, T.J.; Leungwattanakij, S. Antimicrobial activity of antibiotic-soaked, resist-coated bioflex. *Int. J. Impot. Res.* **2003**, *15*, 18.
11. Atala, A.; et al. *Methods of Tissue Engineering*; Academic Press: San Diego, CA, 2005.
12. Atala, A.; et al. *Tissue Engineering*; Birkhauser Press: Boston, MA, 2005.
13. Yoo, J.J.; Meng, J.; Oberpenning, F.; Atala, A. Bladder augmentation using allogenic bladder submucosa seeded with cells. *Urology* **1998**, *51*, 221.
14. Yoo, J.J.; Park, H.J.; Lee, I.; Atala, A. Autologous engineered cartilage rods for penile reconstruction. *J. Urol.* **1999**, *162*, 1119.
15. Kim, B.S.; Yoo, J.J.; Atala, A. Engineering of human cartilage rods: potential application for penile prostheses. *J. Urol.* **2002**, *168*, 1794.

16. Falke, G.; Yoo, J.J.; Kwon, T.G.; Moreland, R.; Atala, A. Formation of corporal tissue architecture in vivo using human cavernosal muscle and endothelial cells seeded on collagen matrices. *Tissue Eng.* **2003**, *9*, 871.
17. Kwon, T.G.; Yoo, J.J.; Atala, A. Autologous penile corpora cavernosa replacement using tissue engineering techniques. *J. Urol.* **2002**, *168*, 1754.
18. Holroyd-Leduc, J.M.; Straus, S.E. Management of urinary incontinence in women: scientific review. *JAMA* **2004**, *291*, 986.
19. Brown, J.S.; Nyberg, L.M.; Kusek, J.W.; Burgio, K.L.; Diokno, A.C.; Foldspang, A.; et al. Proceedings of the National Institute of Diabetes and Digestive and Kidney Diseases International Symposium on Epidemiologic Issues in Urinary Incontinence in Women. *Am. J. Obstet. Gynecol.* **2003**, *188*, S77.
20. Burgio, K.L.; Matthews, K.A.; Engel, B.T. Prevalence, incidence and correlates of urinary incontinence in healthy, middle-aged women. *J. Urol.* **1991**, *146*, 1255.
21. Nygaard, I.E.; Heit, M. Stress urinary incontinence. *Obstet. Gynecol.* **2004**, *104*, 607.
22. Wilson, L.; Brown, J.S.; Shin, G.P.; Luc, K.O.; Subak, L.L. Annual direct cost of urinary incontinence. *Obstet. Gynecol.* **2001**, *98*, 398.
23. Siegel, S.W. History of the prosthetic treatment of urinary incontinence. *Urol. Clin. North Am.* **1989**, *16*, 99.
24. Martins, F.E.; Boyd, S.D. Artificial urinary sphincter in patients following major pelvic surgery and/or radiotherapy: are they less favorable candidates? *J. Urol.* **1995**, *153*, 1188.
25. Gousse, A.E.; Madjar, S.; Lambert, M.M.; Fishman, I.J. Artificial urinary sphincter for post-radical prostatectomy urinary incontinence: long-term subjective results. *J. Urol.* **2001**, *166*, 1755.
26. Donovan, M.G.; Barrett, D.M.; Furlow, W.L. Use of the artificial urinary sphincter in the management of severe incontinence in females. *Surg. Gynecol. Obstet.* **1985**, *161*, 17.
27. Webster, G.D.; Perez, L.M.; Khoury, J.M.; Timmons, S.L. Management of type III stress urinary incontinence using artificial urinary sphincter. *Urology* **1992**, *39*, 499.
28. Haab, F.; Trockman, B.A.; Zimmern, P.E.; Leach, G.E. Quality of life and continence assessment of the artificial urinary sphincter in men with minimum 3.5 years of followup. *J. Urol.* **1997**, *158*, 435.
29. Litwiller, S.E.; Kim, K.B.; Fone, P.D.; White, R.W.; Stone, A.R. Post-prostatectomy incontinence and the artificial urinary sphincter: a long-term study of patient satisfaction and criteria for success. *J. Urol.* **1975**, *156*, 1996.
30. Chancellor, M.B.; Yokoyama, T.; Tirney, S.; Mattes, C.E.; Ozawa, H.; Yoshimura, N.; et al. Preliminary results of myoblast injection into the urethra and bladder wall: a possible method for the treatment of stress urinary incontinence and impaired detrusor contractility. *Neurourol. Urodyn.* **2000**, *19*, 279.
31. Cannon, T.W.; Lee, J.Y.; Somogyi, G.; Pruchnic, R.; Smith, C.P.; Huard, J.; et al. Improved sphincter contractility after allogenic muscle-derived progenitor cell injection into the denervated rat urethra. *Urology* **2003**, *62*, 958.
32. Lee, J.Y.; Cannon, T.W.; Pruchnic, R.; Fraser, M.O.; Huard, J.; Chancellor, M.B. The effects of periurethral muscle-derived stem cell injection on leak point pressure in a rat model of stress urinary incontinence. *Int. Urogynecol. J. Pelvic Floor Dysfunct.* **2003**, *14*, 31.
33. Chen, F.; Yoo, J.J.; Atala, A. Acellular collagen matrix as a possible "off the shelf" biomaterial for urethral repair. *Urology* **1999**, *54*, 407.
34. Yiou, R.; Yoo, J.J.; Atala, A. Restoration of functional motor units in a rat model of sphincter injury by muscle precursor cell autografts. *Transplantation* **2003**, *76*, 1053.
35. Strasser, H.; Berjukow, S.; Marksteiner, R.; Margreiter, E.; Hering, S.; Bartsch, G.; et al. Stem cell therapy for urinary stress incontinence. *Exp. Gerontol.* **2004**, *39*, 1259.
36. Girdansky, J.; Newman, H. Use of a vitallium testicular implant. *Am. J. Surg.* **2005**, *53*
37. Lattimer, J.K.; Vakili, B.F.; Smith, A.M.; Morishima, A. A natural-feeling testicular prosthesis. *J. Urol.* **1973**, *110*, 81.
38. Lipworth, L.; Tarone, R.E.; McLaughlin, J.K. Silicone breast implants and connective tissue disease: an updated review of the epidemiologic evidence. *Ann. Plast. Surg.* **2004**, *52*, 598.
39. Turek, P.J.; Master, V.A. Safety and effectiveness of a new saline filled testicular prosthesis. *J. Urol.* **2004**, *172*, 1427.

Vascular Fluid Dynamics

Stanley A. Berger

Departments of Mechanical Engineering and Bioengineering, University of California, Berkeley, California, U.S.A.

INTRODUCTION

The term *vascular* is defined “as relating to or containing blood vessels.” In this article, we discuss the circulatory systems whereby blood in the human body is distributed to the various organs and tissues of the body. The principal ones are the systemic and pulmonary circulations, and it is these that we discuss in most detail. The velocities in these flows are relatively low, with two immediate consequences: 1) the flows generally are incompressible and 2) the Reynolds numbers are low, ranging from much less than one to the low thousands, and not millions typical of industrial and aeronautical flows. As a result of the latter, the flows are typically laminar, at least under normal physiological conditions. The fact that the flows are laminar and incompressible presents an enormous advantage in their analysis. Complicating factors are that the flows 1) are mostly unsteady, e.g., pulsatile blood flow in the arteries; 2) occur in complex geometries, e.g., bends and junctions. Additional complications arise because the fluid, blood, is a complex suspension, and exhibits non-Newtonian properties under certain flow conditions or in parts of the circulation. Hemodynamic analyses and studies, in addition to contributing insight into the operation of one of the basic functional systems of the body, provide important information relevant to the pathogenesis of vascular diseases and malfunctions, and the means to evaluate the efficacy of palliative interventions to treat these.

BACKGROUND

Shown in Fig. 1 is a schematic of the flow in the systemic and pulmonary circulations, the major blood circulatory systems in the body. In the systemic circulation, contraction of the left ventricle ejects blood through the aortic valve into the aorta, the largest artery, whereupon the blood divides as it moves through many generations of junctions and bifurcations, passing through progressively smaller arteries and arterioles, until it finally reaches the capillaries, the smallest blood vessels in the body, across whose walls the gaseous and mass transfer processes necessary for life occur. The return circuit is through the venules, the

veins, the vena cavae, and finally into the right atrium. Deoxygenated blood is then ejected through the tricuspid valve into the right ventricle and pumped into the pulmonary arteries, ultimately reaching the pulmonary capillaries surrounding the alveoli in the lung, where it is reoxygenated, returning then to the left atrium. Ejection through the mitral valve into the left ventricle then completes one cycle of this pulsatile motion.

Table 1 lists some properties of the circulation and of the blood. While blood has about the same density as water, it is about three times more viscous, primarily as a result of the large number of suspended red blood cells (RBCs). Average Reynolds numbers are sufficiently low for the flow to be laminar throughout the circulation, and are very low in the microcirculation (arterioles, capillaries, and venules). Mean pressure falls from a relatively high value of 100 mmHg in the largest arteries to values of the order of 20 mmHg in the capillaries and even lower in the return venous circulation. To accommodate this, the arteries and arterioles are relatively thick walled, whereas the capillaries, venules, and veins are thin walled. This is schematically illustrated in Fig. 2, which also shows the mixture of the four basic tissues in each of the blood vessel types. The proportionally large amount of elastic tissue in the larger arteries means that these vessels act rather passively in response to the pressure pulse, whereas the arterioles, with a proportionally much larger fraction of smooth muscle can, by dilating or contracting, as controlled by the autonomic nervous system, actively affect the amount of blood flowing through them. This feature of the circulation provides the basis for the dominant flow control mechanism in the systemic circulation, as illustrated in Fig. 3. Note the relatively small pressure drop in the arteries, with most of the pressure drop, as much as 50 mmHg, occurring in the arterioles. It is this that makes it possible for the body to so effectively shunt blood to different organs—for digestion, physical exertion, etc.—as needed.

That changes in arteriolar diameter can so dramatically affect the pressure can be explained by using, approximately, the Poiseuille law, according to which the pressure drop varies inversely with the fourth power of vessel radius, therefore, for example, a decrease in diameter by a factor of two results in a



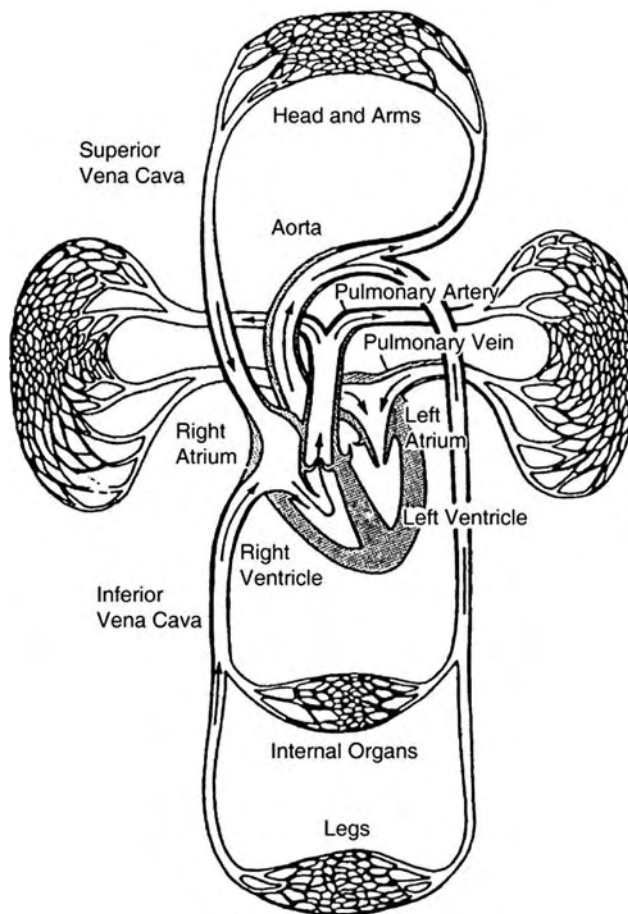


Fig. 1 Schematic of the systemic and pulmonary circulations.

16-fold increase in the resistance to flow, and vice versa. Fig. 3 shows that at the capillary level, there is little or no pulsatility in pressure; this was believed to be true with possibly some exceptions for the systemic circulation, less so for the pulmonary circulation. Recent evidence shows residual pulsatility in the systemic capillaries, with intermittent flow being common.

We now discuss in turn and in some detail each major part of the circulatory vascular system, starting with the arteries and progressing down to the capillaries.

FLOW IN THE SYSTEMIC CIRCULATION

The Arteries

The most obvious feature of blood flow in the arteries is the pulsatile nature of the flow, as evidenced by the existence of a pulse at sites remote from the heart. Our discussion of pulsatile blood flow begins with an analysis of pulse-wave propagation in the arteries.

Inviscid analysis

The analysis is based on an idealized artery, assumed to be an infinitely long, circular, cylindrical elastic tube isolated from its surroundings and subject to a pressure pulse. The blood flowing through this tube is assumed to be a homogeneous, incompressible, inviscid fluid, and the pulse-wave amplitude and tube radius are both assumed to be small compared with the wavelength. These assumptions allow the flow to be modeled as

Table 1 Some properties of the circulation and blood

Vessels	Diameter (mm)	Length (cm)	Wall thickness (mm)	Contained volume (cm ³ or ml)	Mean pressure (mmHg)	Average velocity (cm/sec)	Reynolds number	
							Average	Maximum
Aorta	25.0	40.0	2.0	100	100(av.)	40(av.)	3000	8500
Arteries	15–0.15	15.0	0.8	350	90(av.)	40–10	500	1000
Arterioles	0.14–0.01	0.2	0.02	50	60	10–0.1	0.7	—
Capillaries	0.008	0.05	0.001	300	30–20	<0.1	0.002	—
Venules	0.01–0.14	0.2	0.002	300	20	<0.3	0.01	—
Veins	0.15–15	18.0	0.6	2500	15–10	0.3–5	150	—
Vena cava	30.0	40.0	1.5	300	10–5	5–30	3000	—

Number of red blood cells (per mm³), 5×10^6 .

Number of white blood cells (per mm³), 10^4 .

Blood volume (L), 5–6.

Viscosity of whole blood (mPa sec; cP), 3–4 (in the larger vessels).

Specific gravity, 1.06.

Heart rate (per min), 60–70.

Cardiac output (L/min), 5–6.

Stroke volume (ml), 70.

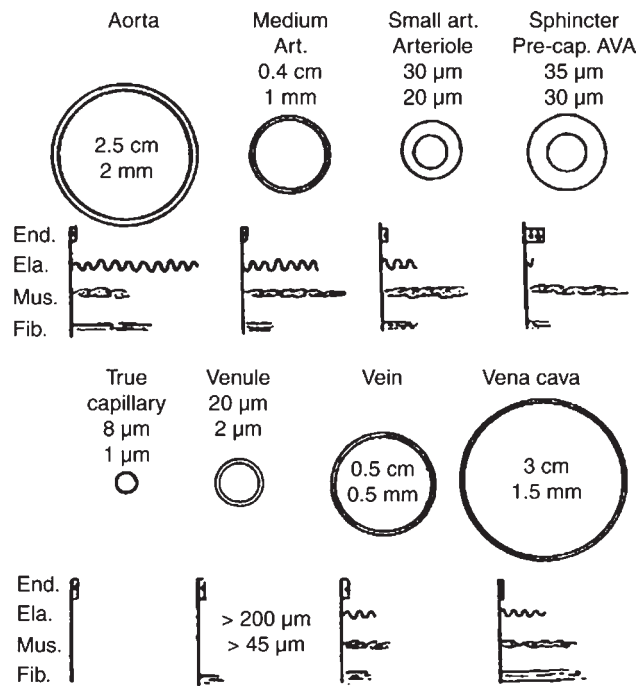


Fig. 2 Variety of sizes, thickness of wall, and admixture of the four basic tissues in the walls of different blood vessels. The figures directly under the vessel name are the lumen diameter and the wall thickness, respectively. *End.*, endothelial cells; *ela.*, elastic fibers; *mus.*, smooth muscle; *fib.*, collagenous fibers. (From Ref. [1].)

one-dimensional, with a single velocity component $u(x, t)$, where x is the coordinate in the direction of the axis of the tube, u is the velocity in this direction, and t is time.

The differential form of the mass conservation, or continuity, equation is

$$\frac{\partial A}{\partial t} + \frac{\partial(uA)}{\partial x} = 0 \quad (1)$$

where A is the cross-sectional area of the tube. Conservation of momentum is expressed by

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + \frac{1}{\rho} \frac{\partial p}{\partial x} = 0 \quad (2)$$

We need a relationship describing the dependence of the tube radius, a , on the internal pressure, p , in the tube. The simplest elastic model is a linear relation

$$da = \frac{\gamma}{2} dp, \quad (3)$$

where γ is the compliance of the tube.

The above constitute three equations for the three unknowns u , p , and $A (= \pi a^2)$. Linearizing them for small disturbances, we can then combine them to obtain the single equation

$$\frac{\partial^2 p}{\partial x^2} = \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2}, \quad \text{where } c^2 = \frac{a}{\rho \gamma} \quad (4)$$

This is the wave equation, the wave speed being c .

For blood flow in the arteries, an important example of such an elastic tube is a thin-walled tube whose material satisfies Hooke's law. Consider a small change in radius da owing to a pressure change dp in half of the cross-sectional element. The increase in the circumference is $2\pi da$, so the associated circumferential strain is $2\pi da/2\pi a = da/a$. Then, according to Hooke's law, the circumferential stress is $E(da/a)$, where E is Young's modulus of the wall material. A force balance for equilibrium on the upper half of the vessel gives, per unit axial length,

$$2adp = 2 \left(E \frac{da}{a} \right) h, \quad \text{or } da = \left(\frac{a^2}{Eh} \right) dp \quad (5)$$

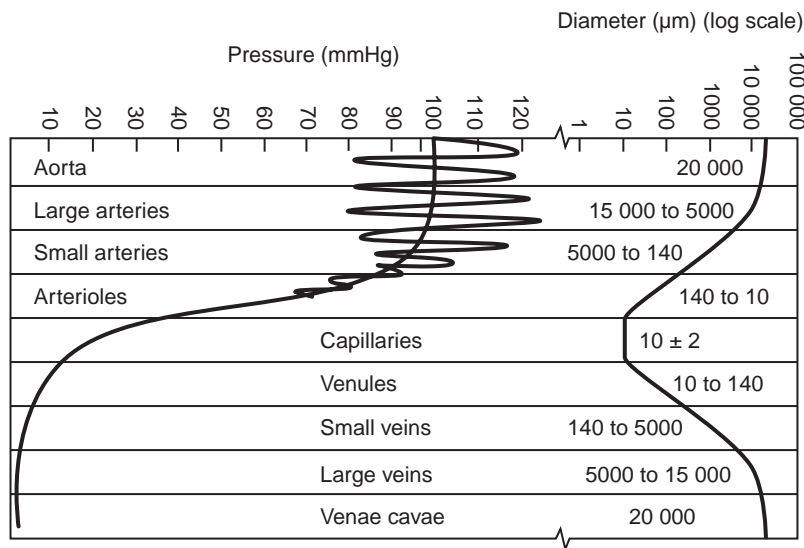


Fig. 3 Pressure variation in the systemic circulation, showing mean pressure and temporal variation about the mean in the different vessels.



where h is the tube wall thickness. Comparison of this with Eq. (3) yields the compliance of the tube as $\gamma = 2a^2/Eh$, so the pulse-wave propagation speed is, from Eq. (4),

$$c = \sqrt{\frac{a}{\rho\gamma}} = \sqrt{\frac{Eh}{2\rho a}} \quad (6)$$

This expression for the wave speed in an inviscid, incompressible fluid in a thin-walled simple elastic tube, one of the most important formulas in arterial fluid dynamics, is called the Moens–Korteweg formula. It is a good approximation to the wave speed in the larger arteries even when the viscosity of blood is taken into account.

Viscous analysis

A more realistic model of blood flow in the arteries considers wave propagation in an arbitrarily *distensible* tube containing a *viscous*, incompressible fluid. This requires more space than available here, so we simply outline the analysis. The fluid is assumed to be Newtonian, this being generally a good approximation in the arteries. The tube is again assumed to be an infinitely long, freely moving, circular cylinder. Wall displacements owing to passage of the wave are assumed to be small, and the wall material is assumed to have linear response, and behave in a purely elastic, viscoelastic, or more complicated manner. Finally, the tube is taken to be either thin- or thick walled. Based on these assumptions one obtains two partial differential equations for the axial and radial displacements of the wall, the independent variables being t and x , and the radial coordinate, r (Fig. 4). (The problem is axisymmetric, so there is no dependence on the azimuthal coordinate.)

The fluid motion is characterized by three dependent variables: the pressure and the axial and radial velocities; the three governing equations are the continuity equation and the axial and radial momentum equations (Fig. 4). For a viscous fluid, the momentum equations, the Navier–Stokes equations, are highly nonlinear because of the presence of convective acceleration terms. In the arteries, the pulse-wave speed is much larger than typical fluid velocities, so these nonlinear terms are small and are dropped.

There are five governing equations, two for the wall motion and three for the fluid motion, i.e., linear partial differential equations in three independent variables, time and the axial and radial spatial coordinates. Specification of boundary conditions completes the formulation of the problem; some of these relate to the fluid only, others require that the motion

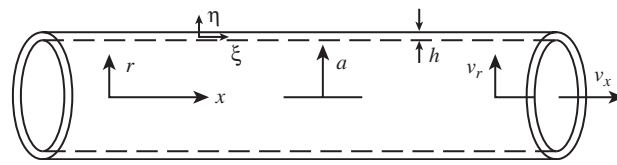


Fig. 4 Vessel geometry and coordinate system.

and state of stress in the fluid and wall match at their interface.

The solution is obtained in terms of a traveling wave solution, of the form

$$\begin{aligned} v_x &= \hat{v}_x(r) \exp[i\omega(t - x/c)] \\ v_r &= \hat{v}_r(r) \exp[i\omega(t - x/c)] \end{aligned} \quad (7)$$

where v_x and v_r are the corresponding velocity components (given by the real parts of the expressions on the right-hand side), and ω and c the frequency and wave speed of the pulse wave, respectively. The system is dissipative because the fluid is viscous, so c is complex, so also are $\hat{v}_x(r)$ and $\hat{v}_r(r)$, which turn out to be Bessel functions of complex arguments. The details of the analysis are quite elaborate; we highlight here only some features of the solution. The analysis, as we have described it, is primarily owing to Womersley.^[2]

The solution depends on the Womersley, or frequency, parameter

$$\alpha = a\sqrt{\frac{\omega}{\nu}} \quad (8)$$

where a is the vessel radius and ν is the kinematic viscosity. The Womersley parameter represents the ratio in the Navier–Stokes equations of the unsteady inertial term to the viscous term, and as such is the fundamental parameter in unsteady viscous flows. It corresponds to the Reynolds number, Re , which represents the corresponding ratio of inertial to viscous forces for a steady flow (the formal definition is $Re = UL/\nu$, where U and L are the characteristic velocity and length). Corresponding to “Reynolds” similarity, similitude modeling considerations require that if two such flows are to be “similar,” α should have the same value for both. Values of α decrease linearly with vessel radius as the flow progresses from larger to smaller arteries. For humans, α is of $O(15)$ in the aorta and $O(3)$ in the smaller femoral artery.

Among quantities of interest is the ratio c/c_0 , where c is the actual pulse-wave velocity and c_0 the Moens–Korteweg velocity for an inviscid fluid [Eq. (6)]. The variation of c/c_0 with the Womersley parameter α (Fig. 5) shows that for values $\alpha \geq 5$, appropriate for the larger arteries, the actual wave speed is very close to the inviscid wave speed (about 6 m/sec in the aorta and

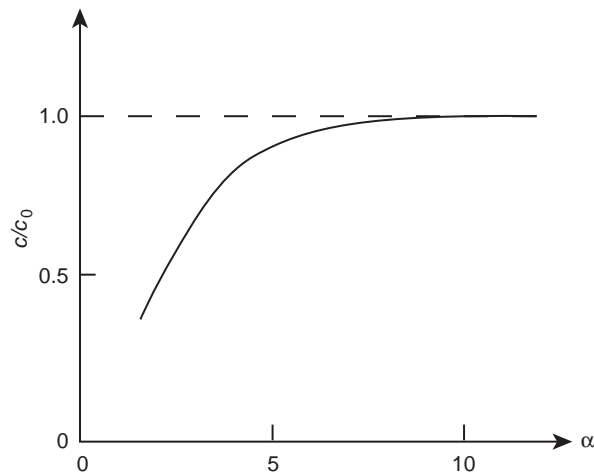


Fig. 5 Variation of wave speed c/c_0 with α .

8–10 m/sec in the femoral), justifying the inviscid analysis for the pulse-wave propagation given earlier. Attenuation of the pulse wave in one wavelength is found to be quite high, suggesting an inefficiently designed system, until one realizes that corresponding to a fundamental pulse frequency f of 1 Hz, the wavelength is $\lambda = c/f = (6 \text{ m/sec})/(1 \text{ Hz}) = 6 \text{ m}$, about four times the distance from the heart to the furthest extremity, so the actual attenuation is small (wavelengths for higher values of α are shorter, but not until the fourth harmonic of the fundamental is an entire wavelength encompassed within the body). Fig. 6 shows the ratio of the maximum flow rate in the artery to the steady Poiseuille value for the same pressure gradient.

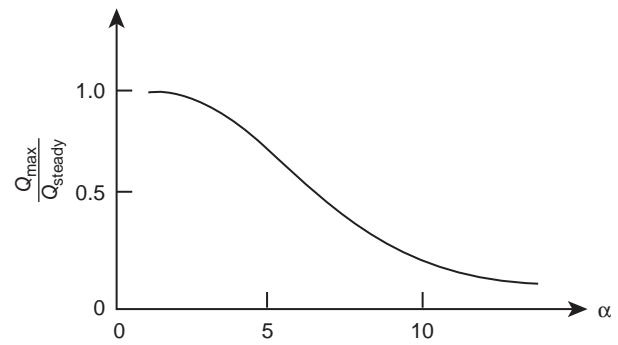


Fig. 6 Variation of flow rate ratio $Q_{\max}/Q_{\text{steady}}$ with α .

Calculated velocity profiles for a “rigid” tube at various time increments in the cycle for a sinusoidal pressure gradient for a range of values of α are shown in Fig. 7. These show that for a relatively low value of α , the flow is quasisteady, with roughly parabolic Poiseuille velocity profiles, while for larger values of α , the profiles are relatively flat, with the core flow lagging behind the rapidly responding wall (Stokes) layers—results consistent with the physical interpretation of the Womersley parameter given above.

General comments

Comparison of the results of the Womersley^[2] viscous pulsatile arterial flow analysis with canine experiments^[3] showed significant discrepancies. Womersley,^[2] accordingly, modified his original analysis,

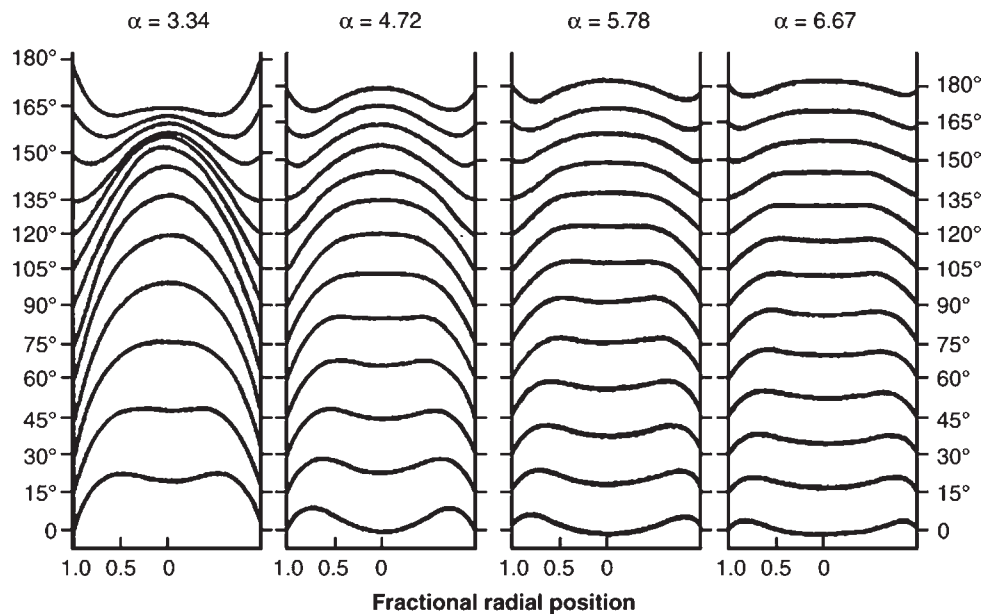


Fig. 7 Velocity profiles in a straight “rigid” tube at various points in the cycle of a sinusoidal pressure gradient for different values of α .



incorporating additional factors such as tethering—the attachment of the artery to the surrounding tissue—and allowing for internal viscous damping in the arterial wall. These substantially improved the agreement between theory and experiment. Other factors play a role in accurately modeling arterial flows, such as arterial taper, but probably none more so than incorporating more realistic arterial wall models.^[4,5] There is also nonlinearity to consider: 1) convective acceleration terms in the Navier–Stokes momentum equations; 2) finite strain of the vessel wall; and 3) nonlinear wall viscoelasticity. To account for such nonlinearities requires extensive numerical calculations, unlike for linearized theories.

Notwithstanding all these additional complications, it is interesting that the results for pulsatile flow in an infinitely long, uniform, rigid circular tube (e.g., as in Fig. 7) reasonably well describe the flow in the arteries for different values of α .

Electrical analogues of the arterial system

The arterial system is a complex multigenerational system of branching, bifurcating, distensible vessels. Modeling the flow in an entire circulatory system or some significant part of it is formidably computationally demanding. Although much progress along these lines has been made,^[6,7] the traditional approach to modeling large segments of/or the entire circulatory system has been to use the analysis for a single arterial segment in conjunction with ideas from AC circuit theory. The integrated, linearized, continuity and Navier–Stokes momentum equations are equivalent to the equations for signal transmission through a uniform cable, the “telegraph” equations, with the correspondence between fluid mechanical and electrical quantities as shown in Table 2. Human arterial anatomical atlases allow the calculation of the arterial parameters on the left-hand side of the table, such as

resistance and compliance, for each arterial segment. This determines their electrical analogues, the entries on the right-hand side. After modeling each segment of the arterial tree using this correspondence, the equations for this complex AC circuit are solved, determining the voltage V and current I in each segment, from which the corresponding values of pressure p and flow Q are determined. From these results, one can calculate the various impedances useful for understanding the gross dynamics of flow in each vessel, and most importantly, the input impedance for each vessel, a measure of the actual effective resistance the blood encounters as it enters that segment, located as it is, in a complex arterial tree.^[8]

It is important to account for the many branchings and junctions, which occur as the blood flows from larger to smaller vessels. The simplest of such branching is a simple bifurcation of a parent vessel into two daughter vessels, as shown in Fig. 8, which can be treated with a one-dimensional flow model (in the spirit of Rayleigh for acoustic waves). The total cross-sectional area of the daughter vessels at arterial tree branchings is generally slightly larger than the parent vessel.

A traveling pressure wave encountering such a bifurcation is partly transmitted and partly reflected. Continuity of volumetric flow and pressure at the junction is sufficient to determine the magnitudes of the reflected and transmitted waves. The analysis shows that reflections cause a slight increase in the pressure wave as long as the increase in total area of daughter vessel is not too great. It is the reinforcing effect of all the reflected waves from each of the multitude of bifurcations in the arterial tree that explains why, in the face of dispersive, dissipative effects of the fluid and wall, the pressure-pulse shape becomes “peakier” as it moves further from the heart into the smaller arteries.

For comprehensive discussions of the circulation, see Caro et al.,^[9] Pedley,^[4] Fung,^[10] and Berger.^[5]

Table 2 Corresponding arterial and electrical quantities

Arterial	Electrical
Pressure, p	Voltage, V
Flow, Q	Current, I
Inertance per unit length, ρ/A	Inductance per unit length, L'
Resistance per unit length, $8\mu/\pi a^4$	Resistance per unit length, R'
Compliance ^a per unit length, dA/dp	Capacitance per unit length, C'
Leakage ^b per unit length, W'	Conductance per unit length, G'

ρ = density of blood.

A = cross-sectional area of vessel of radius a .

^aOwing to distensibility.

^bOwing to lateral branches.

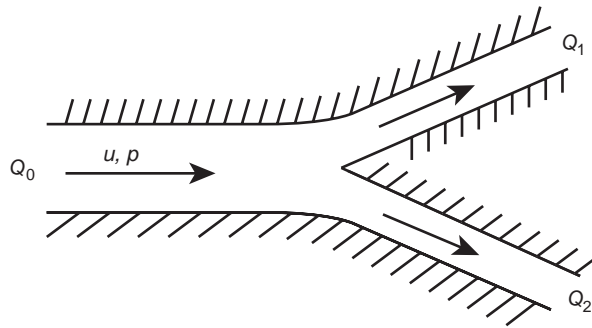


Fig. 8 Simple bifurcation of a parent vessel into two smaller vessels.

The Microcirculation

The term *microcirculation* is usually reserved for those vessels of the arterial and venous system that are smaller than roughly $180\ \mu\text{m}$, which includes the arterioles, capillaries, and venules. Arterioles are those vessels whose diameters lie between those of the smallest arteries and the true capillaries (diameter $10\ \mu\text{m}$ and smaller). They have proportionally the largest amount of smooth muscle, which makes them principal mediators of vascular control of the circulation. The capillaries have diameters $10\text{--}3\ \mu\text{m}$, and it is at their level that the important gaseous and mass transfer processes between the RBCs and the tissues occur. The venules are the vessels in the venous circulation corresponding to the arterioles. The microcirculatory bed is different in different organs and tissues, with different numbers of other microcirculatory vessels, such as metarterioles, preferential or thoroughfare channels, arteriovenous anastomoses, and precapillary sphincters, together with terminal arterioles, playing a role in the regulation of flow.^[11,12]

Pulsatility exists in the arterioles, although to a much lesser degree than in the arteries (Fig. 3), becoming more attenuated as one moves distally. Nonetheless the flow in these vessels is exceedingly hard to analyze because their diameters range from approximately 20 times greater than red cell diameters to slightly larger than red cell diameters, so it is generally not valid to treat blood as a rheologically simple, homogeneous, single-phase Newtonian fluid. Until recently, experimental observations and data were most lacking for this group of blood vessels. Non-Newtonian models have been successfully used in the microcirculation, particularly in the arterioles. This section therefore begins with a brief consideration of blood rheology, first with a discussion of rheological instruments to determine the behavior of blood, and then moves on to rheological laws to model this measured behavior.

Rheology of blood

Viscometry. There are two broad groups of instruments used to measure the viscosity of blood: 1) tube and 2) rotational instruments.

1. *Tube viscometers:* In these viscometers, also known as capillary-tube viscometers, the test fluid is made to flow from a reservoir down a narrow tube of precisely known caliber and either a) the total volume that passes within a fixed time, or b) the time necessary for a given volume to pass, under a given pressure gradient, is measured.^[13] Tubes with diameters less than $100\ \mu\text{m}$ are generally used. For a Newtonian fluid, from Poiseuille's law, the absolute (dynamic) viscosity μ is given by

$$\mu = \left(\frac{\pi d^4}{128LQ} \right) \Delta p \quad (9)$$

where d is tube diameter, L tube length, Q volume flow rate, and Δp the pressure head. The bore of the tube must be known with great precision, and corrections are needed for end effects. Viscosity values are commonly reported as relative viscosities—usually to that of water. The rate of shear is not constant in tube viscometers, so they should, strictly speaking, be used only for Newtonian fluids, although this injunction is often breached.^[13]

2. *Rotational viscometers:* These consist of two elements, separated by the liquid being tested, that rotate relative to one another about a common axis. Viscosity is determined by the relationship between the rotational velocity of the rotating element and the torque developed on the other. The two principal subgroups of instruments are: coaxial cylinder and cone-plate types (Fig. 9).

In the *coaxial cylinder* instrument, two coaxial circular cylinders, of radii R_1 and R_2 (where, $R_2 > R_1$), are made to rotate relative to one another (Fig. 9A), usually with the inner one fixed, the outer one rotating at a constant angular velocity, Ω . If the fluid in the annulus between the cylinders is Newtonian, the torque transmitted to the inner cylinder is given by

$$T = \left(\frac{4\pi R_1^2 R_2^2 h \Omega}{R_2^2 - R_1^2} \right) \mu, \quad (10)$$

determining the viscosity μ in terms of the geometry of the instrument and the measured torque T . In the *cone-plate* instrument, the space between a cone of a large apex angle and a flat surface normal to the cone axis is filled with the test liquid (Fig. 9B), and one of these



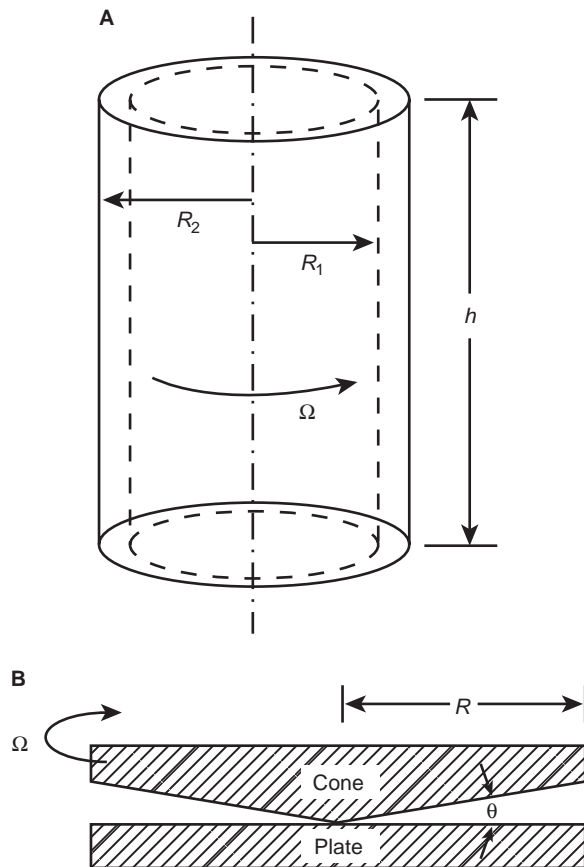


Fig. 9 Schematics of rotational viscometers: (A) coaxial cylinder; (B) cone plate.

elements is rotated, with the torque measured on the other. The rate of shear in the fluid is constant throughout the fluid.^[13]

Rotational viscometers, particularly for biomedical applications, are made with the gap between the rotating and stationary elements as narrow as possible. This has two advantages: 1) the volume of test fluid required is small; 2) the narrower the gap, the closer is the rate of shear to being constant across the gap, a necessary requirement when, using a range of rotational speeds, these instruments are used to determine the relationship between stress and rate of shear for a non-Newtonian fluid.

Stress–Shear-Rate Laws. We consider now some simple stress-rate-of-shear relations, including, wherever possible, the usual mechanical models used to represent them.

1. *Newtonian fluid:* The relationship between stress τ and shear rate D is linear,

$$\tau = \mu D \quad (11)$$

The mechanical model is a simple dashpot.

2. *Power-law fluid:* This is in the form of the empirical relationship

$$\tau = \mu D^s, \quad s = \text{const.} \quad (s = 1, \text{Newtonian}) \quad (12)$$

There is no simple mechanical model.

3. *Bingham plastic:* The mechanical model is a dashpot in parallel with a friction coupling. Friction between the sliding surfaces in the coupling makes extension impossible unless the applied force F is greater than some frictional force. If the rate of extension beyond this value is proportional to the applied stress, the material is called an ideal Bingham plastic; otherwise it is a generalized Bingham plastic. In the former case,

$$\tau - \tau_y = \mu D \quad \text{or} \quad \tau = \tau_y + \mu D \quad (13)$$

where τ_y is the yield stress.

4. *Casson plastic:* There is no mechanical model. This is an empirical extension of the law for a Bingham plastic, of the form

$$\sqrt{\tau} - \sqrt{\tau_y} = \mu \sqrt{D} \quad (14)$$

Fig. 10 shows these various laws for τ as a function of D .

Blood Viscosity; Stress–Shear-Rate Laws. Blood is a complex heterogeneous suspension. The suspended material includes white cells and platelets, but their

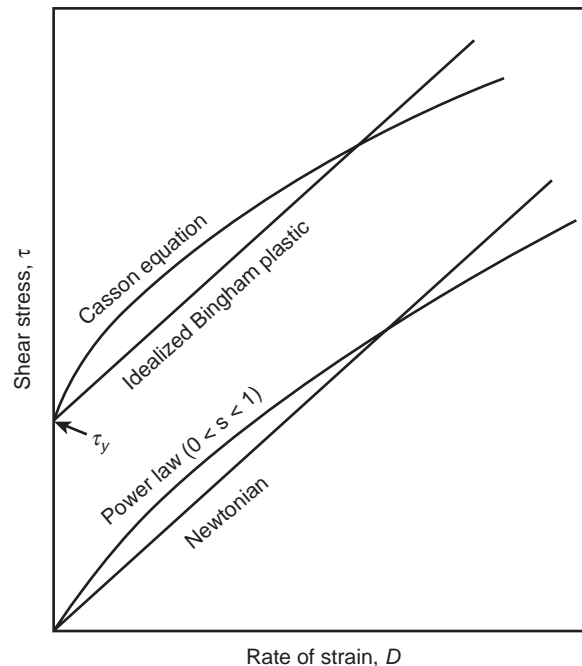


Fig. 10 Stress vs. shear rate for different rheological models.

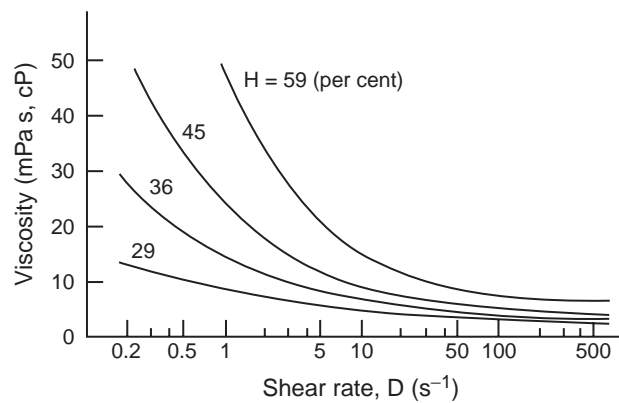


Fig. 11 Viscosity of whole blood as a function of shear rate for different hematocrits (temperature 37°C).

numbers are small. The principal contributors to the rheological behavior of blood are the erythrocytes, the RBCs, their numbers measured in terms of the hematocrit, H , the volume percentage of red cells.

Using one of the rotating viscometers described above, one obtains the results shown in Fig. 11 for the viscosity of whole blood as a function of shear rate for different hematocrits. The viscosity increases significantly as the hematocrit increases, with the curves all asymptoting to constant values of viscosity for large values of D . Blood with a normal male hematocrit of about 45% has a constant viscosity of about 3–4 mPa sec (3–4 cP) for values $D \geq 100 \text{ sec}^{-1}$. Mean shear rates in the arteries being generally much higher than 100 sec^{-1} , to a very good approximation, blood behaves in the arteries as a Newtonian fluid with a constant viscosity (This is generally only true for normal vessels; for diseased or injured arteries, or for certain geometries, there may be regions of lower shear rates owing to flow separation and recirculation so that non-Newtonian models are more appropriate. Various models have been used, often power-law models, with exponents reflecting the shear-thinning (*pseudo-plastic*) nature of blood, and including, or not, a yield stress. Popular ones are the Herschel-Bulkley, Walburn-Schneck, and Carreau models.^[14]).

Although nominal shear rates in the arterioles may be much larger than 100 sec^{-1} , the rheological behavior of blood in the arterioles is not well modeled as Newtonian. Among the models discussed above, the traditional ones used are the power-law [Eq. (12)] and the Casson equation [Eq. (14)]. For the power-law, s lying in the range $0.68 \leq s \leq 0.80$ was found to provide a good approximation for shear rates lying in the range $5\text{--}200 \text{ sec}^{-1}$ ^[13] (values of $s < 1$ reflect the shear-thinning property of blood as the shear rate increases).

The rheological formula equations [Eqs. (13) and (14)] include a yield stress, reflecting the formation at very low velocities of rouleaux, broad face-to-face

aggregations of two, and generally many more, RBCs. Their formation has been ascribed to various mechanisms. In any event, the yield stress τ_y is the (small) value of stress [$\tau_y < 0.01 \text{ Pa}$ (0.1 dyn/cm^2)] necessary to break up these aggregates before flow commences.

Rheological and hematocrit anomalies of blood^[15–18]

Effect of Tube Size^[16–18].

1. *Fåhræus–Lindqvist effect*: For tube-viscometer capillary diameters greater than 1–2 mm, the relative viscosity of whole blood remains the same. For narrower tubes, the apparent viscosity decreases with diameter, continuing to decrease until diameters reach values as low as $8 \mu\text{m}$, after which it increases sharply as the very much different capillary flow regime is entered. This Fåhræus–Lindqvist effect is one of the most significant features of microcirculatory flow. Although a number of factors contribute to the effect, the primary physical reason is the well-known tendency of RBCs to migrate away from the wall, leaving a cell-depleted, lower viscosity region near the wall, thereby reducing the overall viscosity of the blood in the tube.
2. *Fåhræus effect*: This other important characteristic of the microcirculation is related to the Fåhræus–Lindqvist effect, being a further consequence of the migration of RBCs from the tube wall. The resulting increased concentration of cells in the higher velocity part of the flow reduces on average the transit time of RBCs through the tube. As a result, the *tube hematocrit*, H_T , the hematocrit within the tube (measured, e.g., by stopping the flow and emptying the contents) is less than the hematocrit of the fluid exiting the tube, the *discharge hematocrit*, H_D .

Another important influence on hematocrit variation in the microcirculation is associated with the presence of a cell-depleted wall layer. When blood flows through a diverging microvascular bifurcation, a disproportionate fraction of RBCs flow into the branch receiving the higher flow rate, resulting in a higher hematocrit in that branch, the lower-flow branch primarily “skimming” plasma from the peripheral layer of the flow.

Cell Migration. Cell migration has very important effects on the flow and rheological behavior of blood in the vessels of the microcirculation. As the velocity increases, there is an increasing tendency of the red cells to move toward the axis of the tube, increasing the hematocrit there and lowering it near the wall. To aid



understanding of the mechanisms involved, migratory paths of rigid and deformable particles in slow viscous flow at very low Reynolds numbers, for which inertial effects are negligible, and in viscous flows at larger Reynolds numbers, for which inertial effects are not negligible, have been determined.^[15] In most cases, the particles drift to the axis or to some annular location between the axis and tube wall. Migration of neutrally buoyant rigid spheres to an annulus owing to inertial effects is called the “Segré–Silberberg effect.” These observations, based on model experiments and calculations, are only suggestive of the behavior of real RBCs, because the importance of cell flexibility and flow inertia will depend on flow state and vessel geometry. In addition to these migratory patterns, individual red cells also rotate and deform in sheared flows.

Flow in the microcirculation

Arterioles. The arteriolar network is typically organized in two general patterns: an arcade arrangement in vessels of diameter of the order of $25\ \mu\text{m}$ and larger, and a sequential branching pattern in more distal vessels. These are the vessels that control the distribution of blood to organs and tissues of the body by changing their diameters by vasodilation or vasoconstriction. A variety of methods have been used to measure velocity profiles and pressures in the arterioles and venules of various tissues. Some of these findings were discussed above. In terms of modeling flow in the arterioles, although much progress has been made, it remains presently impossible, computationally or otherwise, to analyze anything as complex as the motion of a concentrated suspension of real, deformable, red cells in narrow tubes whose diameters range from a few to 20 times the diameter of these cells, spanning the domain of particulate to continuum flow. In addition, there are structural and geometrical differences between arterioles in different tissues. In the absence of a theory of flow as simple and universal as for arteries, results for the arterioles are necessarily, in part, empirical, phenomenological, and analytical. Generally for flows in individual microvessels of the microcirculation, the Reynolds, Re , and Womersley, α , numbers are small, so inertia does not play a significant role, and the momentum balance is between pressure and viscous forces (*Stokes flow*). Complications, however, are that the viscosity, μ , cannot generally be taken as a simple constant, as for a Newtonian fluid, and that the arteriole is distensible, the diameter, d , generally depending on the pressure, $d = d(p)$, differently for different tissues. Non-Newtonian viscosity models, in the context of single and multiple-phase analyses, have been employed to deal with the former of these

complications, in formulations that are different depending on the size of the arteriole. Examples of multiple-phase analyses are two-phase continuum models, consisting of a core Newtonian viscous fluid, representing a concentrated RBC suspension, surrounded by an annulus of a less viscous Newtonian fluid, representing the cell-depleted layer, which give results in reasonable quantitative agreement with experimental data on blood apparent viscosity in tubes with diameters of $O(30\ \mu\text{m})$ and larger (for the details of these, see Refs. [16–18]).

Note, unlike the arterial system where the flow in individual arteries is often sought in its own right, individual arterioles are of interest primarily to the extent that they are part of a microvascular network, and are involved in the regulation of flow in such a network.

Capillaries. In these vessels, ranging in size from 10 down to $3\ \mu\text{m}$, the motion of individual RBCs must be taken into account. The typical normal red cell is a biconcave disk, $8\ \mu\text{m}$ in diameter and $2\ \mu\text{m}$ in thickness, consisting of a thin, flexible membrane filled with a viscous Newtonian-like fluid, hemoglobin. The flow may or may not be pulsatile. The extreme flexibility of red cells is most evident in these vessels, e.g., a red cell is readily able to move through a $3\ \mu\text{m}$ -diameter micropipette by folding about an axis through its center into a “crêpe suzette” shape. In wider, $5\text{--}7\ \mu\text{m}$, tubes, it may take on a “parachute” shape. In the largest capillaries, the cells may move down relatively undeformed, their broad face perpendicular to the flow direction. They move down in single file, one following the other.

The motion of undeformed individual red cells in the larger capillaries was early to receive attention. One such generic configuration is shown in Fig. 12. The usual analysis idealizes the cells as rigid circular disks moving steadily with constant velocity in a rigid circular tube surrounded on all sides by plasma. Because of the very small lengths and low cell velocities, the Reynolds numbers for capillary flow are exceedingly small (of the order of 10^{-3}), so these flows fall into the category of Stokes or “creeping flows.” The governing differential equations for such flows are linear and, for

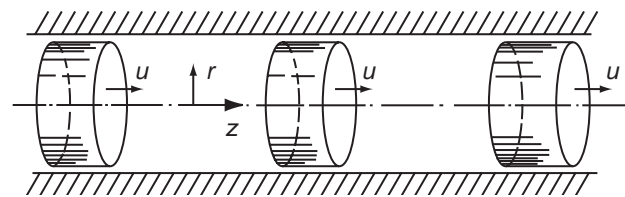


Fig. 12 Rigid disks in a tube as a model of red cell flow in a capillary.

simple configurations such as that shown in Fig. 12, may be solved in closed form.

In somewhat narrower capillaries, in which the cells must deform, e.g., into the parachute shape, to fit in the vessel, the flow in the narrow gap between the cell and the capillary wall can be analyzed as a *lubrication flow*, with due account taken of cell deformability. This is a special, difficult, creeping flow, which has been resolved by the application of sophisticated perturbation analysis.

Both of the above are axisymmetric flows. There have been advances in recent years in treating non-axisymmetric configurations, there being evidence that red cells commonly enter capillaries edge on, resulting in significant cell deformation and a “tank treading” of the cell membranes.^[10,17]

Resistance to flow depends strongly on the capillary diameter, and in the larger capillaries, it is very much greater than it would for Poiseuille flow in the same-sized tube of a fluid with the overall whole blood viscosity.

As to the nature of the flow in the plasma region between red cells, in a frame moving with the velocity of the cells, the streamlines of the plasma region follow closed circulatory paths, a pattern called “bolus flow.” Speculation, on qualitative grounds, that this recirculatory flow pattern might beneficially facilitate mass transfer of gases between the cells and the surrounding tissues has not been borne out.

The effect of the thin capillary endothelial glycocalyx layer on the motion of RBCs has been of great recent interest^[17] [the glycocalyx and its associated macromolecules are referred to, collectively, as the endothelial surface layer (ESL)].

Venules. These are postcapillary venous side vessels corresponding to the arterioles on the arterial side. They are about the same length as corresponding arterioles, but with larger lumens. The shear stress values in them are low, increasing RBC aggregation, which tends to produce a more blunted velocity profile, increasing the apparent viscosity. With aggregation, however, there is greater axial cell migration, resulting in a thicker cell-depleted layer at the wall, providing a compensatory reduction in viscosity. The blunting of the velocity profiles, which are sometimes asymmetric, may be associated with as much as a doubling of the resistance, which is also very much affected by the ESL.

Flow in venules with diameters of $O(25\ \mu\text{m})$ and larger have been modeled with two-phase continuum models, similar to those used for arteriolar flows, or with multiphase, non-Newtonian formulations. For narrower venules, there is no available theory. White blood cell (WBC)–endothelial interactions occur mainly on the venous side of the microcirculation. A

large number of experimental studies have been done, including direct observations of WBCs adhering to, rolling on, and transmigrating across venular endothelium, and in vitro and in vivo investigations into the molecular mechanisms of these interactions. There has also been a good deal of theoretical modeling and computational studies. See Ref. [18] for a discussion of this work, and, more generally, the flow in venules and the microcirculation as a whole; this review also discusses how the results for individual microvessels are integrated to model blood flow in microvascular networks.

The Veins

The veins draining a vascular bed tend to have larger lumens than the corresponding arteries supplying it. They are thin walled, with a much smaller ratio, 0.1, of wall thickness to diameter (Fig. 2), making them much more distensible than arteries, and there are greater numbers of them. These features give the venous system a much greater volumetric capacity than the arterial system, able to hold as much as 70% of the total blood volume. Their walls also contain much less smooth muscle and elastic fibers than arteries (Fig. 2), and are easily stretched. They are often elliptical, rather than circular, in cross section. Venous pressures are generally low, 12 mmHg or so, about 10% of the peak arterial systolic pressure (Fig. 3). The venous system is free of any heart-induced pulsations, the slight unsteadiness in the largest veins, near the heart, is primarily caused by respiratory pressure changes (and possibly cross talk with nearby parallel arteries). As there are no significant pressures and pressure gradients on the venous side of the circulation, blood is forced back to the heart by the squeezing action on the thin-walled veins of surrounding tissues (e.g., skeletal muscles), with the assistance of valves in the veins to prevent backflow. This motion is much less regular than arterial flows, and plays a much lesser role in the dynamics of the circulation; it has thus received much less attention.

Another very important difference between arterial and venous flows is that because the latter is a low-pressure system, the transmural pressure in the veins is sometimes negative, the result of hydraulic pressure gradients, muscular action, intrathoracic pressure changes, etc. When the pressure outside a vein exceeds that inside, there is a tendency for the vessel to collapse. This phenomenon arises in other areas of physiological flow in flexible vessels: in some circumstances for systemic arterial flow, in the pulmonary circulation, in the respiratory airways, in peristaltic transport, etc. It is part of the more general physiological problem of strong fluid–structure interaction,



and leads to important and complex physics, such as multiple geometric modes of vessel collapse, flow-induced instabilities, etc. A rich literature has emerged on flow in collapsible tubes (see Ref. [19] for an early account and Refs.^[20,21] for recent accounts).

THE PULMONARY CIRCULATION

This circulation comprises the passage of all the deoxygenated blood exiting the right ventricle through the pulmonary arteries to the alveolar capillaries and its return, oxygenated, via the pulmonary veins to the left atrium. A secondary, much smaller, circulation, the bronchial circulation, fed by the thoracic aorta, nourishes the lung itself.

The pulmonary circulation, although similar to the systemic circulation in the types of vessels, arteries, capillaries, and veins, shows important differences:^[10]

1. The pulmonary circulation is a low-pressure, low-resistance system, with a time-average pressure in the arteries of about 15 mmHg being one-sixth of that in the systemic arteries.
2. Pulmonary arteries are much thinner walled than systemic arteries, consistent with the lower pressures; they have the same Young's modulus as the latter, but are more distensible because they are thinner.
3. The pulmonary vascular bed is not regionally specialized.
4. Vasomotor control is relatively unimportant under normal conditions; unlike the systemic arteries, the pulmonary vessels do not undergo large active changes in their dimensions; the arteries and arterioles have primarily elastic tissue, but little smooth muscle.

To a great extent, the mechanics of the pulmonary circulation are affected by the mechanics of the lung and the thoracic pressure changes over the respiratory cycle. Arterial branching is very similar to that in the systemic circulation. The shape of the pressure wave in the main pulmonary arteries is like that in the aorta, but there is no significant change in shape along the pulmonary arterial system because it is so short—the total length from the beginning of the pulmonary trunk to the capillaries being only 20 cm—that the pulse occurs simultaneously along the entire length. Velocity profiles in the arteries are relatively flat because $Re \gg 1$ and $\alpha = a\sqrt{\omega/\nu} \gg 1$, and the vessels are so short that new boundary layers begin on each flow divider and never fill the vessels.

The pulmonary microcirculation has a number of features that contrast sharply with the systemic microcirculation:^[10]

1. There is less vasomotor activity in precapillary vessels (or elsewhere) than in the systemic circulation, with resistance to flow occurring mainly in the capillaries (there is some local flow regulation, and hence oxygen regulation, in the lung by vasodilation and constriction of the pulmonary arterioles).
2. The capillaries have lengths comparable to their diameters, are flattened, and are arranged like two-dimensional channels passing between pairs of alveoli, the walls of which are flexible and held apart by posts (typical thickness $\approx 4 \mu\text{m}$), so the blood flows as a sheet of fluid between almost parallel alveolar membranes held apart by frequently occurring posts of connective tissue.
3. Pulsatility of the flow is more pronounced.

As in the systemic microcirculation, $Re \ll 1$, velocities are typically about 0.005 m/sec, and $\alpha \ll 1$. Fluid inertia is therefore negligible, the force balance in the capillary sheet being between pressure and viscous forces, and the flow is quasisteady, i.e., the flow rate is in phase with the pressure gradient at all parts of the sheet at all times. In the case of a Newtonian fluid flowing in a rigid sheet, it would then follow that $Q \propto dp/dx$, i.e., the flow rate is directly proportional to the pressure gradient, but, because the sheet here is elastic, the sheet thickness, h , depends on dp/dx . This implies, as h affects the fluid resistance, that resistance depends on dp/dx . Based on these ideas and assumptions, a simple theory for flow in such a sheet can be developed.^[10] Because the alveolar pressure, p_{alv} , the pressure outside the sheet, plays a major role in determining the width of the sheet, the nature of the flow depends on the relative magnitudes of three pressures: the arterial, venous, and alveolar pressures. An increase in p_{alv} , leading to a decrease in sheet width, particularly to values as small as 2–3 μm , has a major effect on the rate of flow, for, even in a Poiseuille flow of a Newtonian fluid in such a channel, the resistance to flow increases, strongly, as h^4 , where h is the channel width, and the increase should be so much greater for whole blood with 8 μm red cells that must be deformed to flow through the sheet.

Dependence of the flow in the pulmonary microcirculation on the relative magnitudes of the arterial, p_a , venous, p_v , and alveolar pressure, p_{alv} , combined with the hydrostatic variation of pressure, leads to nonuniform perfusion over the height of the lung,^[10] as sketched in Fig. 13. The height of the lung is about 30 cm, so the total hydrostatic head between the top and bottom of the lung is 30 cm H₂O or 22.5 mmHg. The pulmonary arteries enter the lung about halfway up, and because p_a is on average about 15 mmHg, arterial pressures are about 26 mmHg at the bottom and 4 mmHg at the top of the lung. As there is a 7–8 mmHg

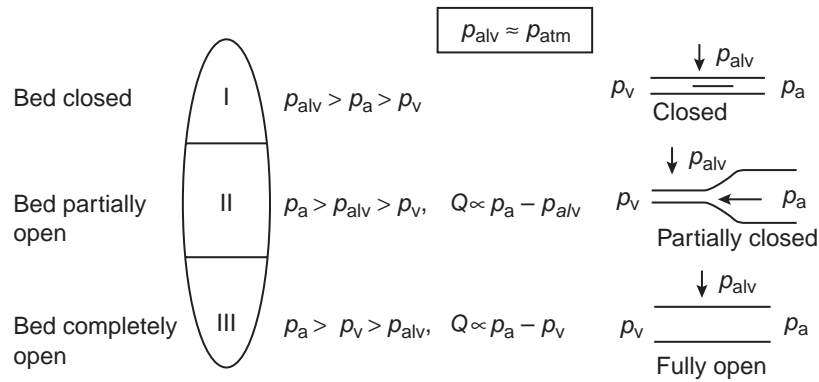


Fig. 13 Schematic of human lung showing the relationship of p_a , p_v , and p_{alv} at different heights.

drop in pressure across the pulmonary circulation, the mean pressure is normally subatmospheric in the microcirculation at the top of the lung! If the pressure in the tissue surrounding a blood vessel exceeds that in the vessel, i.e., the transmural pressure is negative, there will be a net force acting to close the vessel. Because the pressure external to the pulmonary microvessels, p_{alv} , is approximately atmospheric, the relative magnitudes of p_a , p_v , and p_{alv} at different levels in the lung will be as shown in Fig. 13, with vessels fully open near the bottom of the lung, partially open in the middle, and fully closed at the top. The flow rate, Q , in the lowest, fully open capillaries (zone III) is proportional to $p_a - p_v$, while that in zone II is proportional to $p_a - p_{alv}$ (the latter zone thus behaves as a so-called “Starling resistor,” or waterfall, because like a waterfall, the flow rate does not depend on the total height of fall, in this case the overall pressure drop $p_a - p_v$). The levels at which these zones blend into one another depend on the alveolar pressure, and therefore vary over the breathing cycle and with the arterial pressure, which increases, e.g., during exercise.^[10,21]

The above considerations suggest grossly unequal perfusion and ventilation in the lung, because air enters the lung near the top and does not normally make it to the lower recesses of the lung, whereas, from the discussion above, blood flow is greatest at the bottom and least at the top. Partly because of the highly deformable lung structure, however, there is a much closer balance of air and blood over the length of the lung than would be deduced from these considerations.

FLUID MECHANICS AND VASCULAR DISEASE

This article has thus far exclusively considered the normal vasculature. Clearly, knowledge of the normal behavior of any system or part of it is preliminary to understanding its malfunctioning. The few comments here, for lack of space, will address primarily the principal disease of the vasculature, particularly the arteries. But first, we briefly note that the patency of

arteries may be compromised for congenital reasons, by injury, or by disease. Patency may be restored or maintained, or nonpatent vessels bypassed, by the use of vascular grafts, stents, or stent grafts. They are also used for treating aneurysms. These as well as other interventions are widely used clinically. Fluid mechanics is used to determine the efficacy of various designs of these prostheses and where to site them.

The biology of, and flow in, the microcirculation was thought to be most in need of research progress, whereas the study of the arteries was regarded as a mature subject. It is therefore ironic that, increasingly, arterial flow has been a central focus of research in biofluid-dynamics, motivated by the belief that hemodynamics plays an important, perhaps critical, role in the pathogenesis and pathophysiology of atherosclerosis, the leading cause of death, via heart attack or stroke, in the developed world.^[22] Atherosclerosis is a focal disease and research has shown that atheromatous lesions are preferentially located where there are sudden changes in arterial flow geometry in unsteady flow. Among the important ways fluid mechanics affects the arteries is its effect on endothelial cell morphology and cytoskeletal organization, as well as in its helping to provide pathways and facilitate transport of low-density lipoproteins through the arterial walls, the key first step in atherogenesis. Seminal studies in recent years have elucidated many of the pathways by which mechanical forces, primarily wall shear stresses, are transduced to effect changes in the arterial wall at various levels: the cellular, subcellular, genetic, and molecular.^[22–24]

Whereas low values, or rapidly changing values, of wall shear stress are associated with atherogenesis, less is known about the role or effect of shear stresses, or fluid mechanics in general, on plaque growth and remodeling. Disruption of “older” plaque has been identified as a key pathophysiological event in coronary thrombosis and carotid thromboembolism. Factors in plaque fracture and rupture are believed to be in large part hemodynamic in origin, normal stresses, shear stresses, and long term effects such as fatigue,



perhaps all playing a role.^[24] Very much more remains to be done in this area.

CONCLUSIONS

Flows in the circulation, principally the systemic and pulmonary, are discussed. Since these flows are typically laminar, and the fluid, blood, is incompressible, they are often reasonably amenable to analytical or computational analysis, at least for the larger vessels. Complicating the flows even in these vessels are unsteadiness (pulsatility), complex geometries (bends, junctions, bifurcations, etc.) and branching networks, non-Newtonian blood rheology, vessel deformability (leading to fluid-structure interaction), etc. In the smaller vessels, those in the microcirculation, the finite size of the suspended material, especially the red blood cells, must be taken into account, making analysis of the flows in them much more difficult. The relevance of such hemodynamic analyses and studies to vascular disease and its treatment is briefly considered.

ARTICLES OF FURTHER INTEREST

Bioreactors; Mass Transfer in Tissue Function: Roles; Microcirculatory Oxygen Transport

REFERENCES

- Burton, A.C. *Physiology and Biophysics of the Circulation*; Year Book Medical Publishers: Chicago, 1965.
- Womersley, J.R. *An Elastic Tube Theory of Pulse Transmission and Oscillatory Flow in Mammalian Arteries*, WADC Technical Report TR-56-614; Wright Air Development Center: Dayton, Ohio, 1957.
- McDonald, D.A. *Blood Flow in Arteries*, 2nd Ed.; Williams & Wilkins: Baltimore, 1974.
- Pedley, T.J. *The Fluid Mechanics of Large Blood Vessels*; Cambridge University Press, 1980.
- Berger, S.A. Flow in large blood vessels. In *Contemporary Mathematics*; Cheer, A.Y., van Dam, C.P., Eds.; American Mathematical Society: Providence, RI, 1993; Vol. 141, 479–518.
- Taylor, C.A.; Draney, M.T. Experimental and computational methods in cardiovascular fluid mechanics. *Ann. Rev. Fluid Mech.* **2004**, *36*, 197–231.
- Quarteroni, A.; Formaggia, L. Mathematical modelling and numerical simulation of the cardiovascular system. In *Modelling of Living Systems; Handbook of Numerical Analysis Series*; Ayache, N., Ciarlet, P.G., Lions, J.L., Eds.; Elsevier: Amsterdam, 2003.
- Noordergraaf, A. *Circulatory System Dynamics*; Academic Press: New York, 1978.
- Caro, C.G.; Pedley, T.J.; Schroter, R.C.; Seed, W.A. *The Mechanics of the Circulation*; Oxford University Press, 1978.
- Fung, Y.C. *Biomechanics: Circulation*, 2nd Ed.; Springer-Verlag: New York, 1997.
- Zweifach, B.W. Basic mechanisms in peripheral vascular hemostasis. *Transactions of the Third Conference on Factors Regulating Blood Pressure*; Josiah Macy Foundation: New York, 1950; 13–52.
- Segal, S.S. Regulation of blood flow in the microcirculation. *Microcirculation* **2005**, *12*, 33–45.
- Whitmore, R.L. *Rheology of the Circulation*; Pergamon Press: Oxford, 1968.
- Tanner, R.I. *Engineering Rheology*, 2nd Ed.; Oxford University Press, 2000.
- Goldsmith, H.L. Red cell motions and wall interactions in tube flow. *Proc. Feder. Am. Soc. Exp. Biol.* **1971**, *30*, 1578–1588.
- Schmid-Schonbein, G.W. Biomechanics of microcirculatory blood perfusion. *Ann. Rev. Biomed. Eng.* **1999**, *1*, 73–102.
- Secomb, T.W. Mechanics of red blood cells and blood flow in narrow tubes. In *Modeling and Simulation of Capsules and Biological Cells*; Pozrikidis, C., Ed.; Chapman & Hall/CRC: Boca Raton, 2003; Chap. 4, 163–196.
- Popel, A.S. Microcirculation and hemorrheology. *Ann. Rev. Fluid Mech.* **2005**, *37*, 43–69.
- Shapiro, A.H. Steady flow in collapsible tubes. *J. Biomech. Eng. Trans. ASME*, **1977**, *99*, 126–147.
- Li, J.K.-J. *Dynamics of the Vascular System*; World Scientific: Singapore, 2004.
- Grotberg, J.B.; Jensen, O.E. Biofluid mechanics in flexible tubes. *Ann. Rev. Fluid Mech.* **2004**, *36*, 121–147.
- Wootton, D.M.; Ku, D.N. Fluid mechanics of vascular systems, diseases, and thrombosis. *Ann. Rev. Biomed. Eng.* **1999**, *1*, 299–329.
- Davies, P.J. Flow-mediated endothelial mechanotransduction. *Physiol. Rev.* **1995**, *75*, 519–560.
- Berger, S.A.; Jou, L.-D. Flows in stenotic vessels. *Ann. Rev. Fluid Mech.* **2000**, *32*, 347–382.

BIBLIOGRAPHY

- Fung, Y.C. *Biomechanics: Mechanical Properties of Living Tissues*, 2nd Ed.; Springer-Verlag: New York, 1993.
- Fung, Y.C. *Biomechanics: Motion, Flow, Stress, and Growth*; Springer-Verlag: New York, 1990.
- Milnor, W.R. *Hemodynamics*, 2nd Ed.; Williams & Wilkin: Baltimore, 1989.
- Skalak, R., Chien, S., Eds.; *Handbook of Bioengineering*; McGraw-Hill: New York, 1987.

Vascular Grafts

Deon Bezuidenhout

Peter Zilla

Cardiovascular Research Unit, University of Cape Town Medical School and Grootte Schuur Hospital, Cape Town, Western Cape, South Africa

INTRODUCTION

The integrity and function of the estimated 100,000 miles of blood vessels in the human body may be compromised by various pathological conditions and traumatic injuries. Although many cases are treated by alternative methods, severe pathologies or injuries may require the surgical replacement or bypassing of the injured vessels to restore adequate blood flow to the distal vasculature. This article will focus on the vascular grafts (substitute vessels) used in replacement and bypass surgery. It will discuss the need for vascular grafts, the different types of experimental and commercially available grafts, the healing responses elicited by grafts, the need for alternative grafts, and efforts to develop grafts able to provide good long-term outcomes in challenging, small-diameter applications.

THE NEED FOR VASCULAR GRAFTS

The total demand for vascular grafts used to augment the insufficient functioning of blood vessels has been estimated to be from over a million cases worldwide^[1] to 1.4 million in the United States alone.^[2] Figure 1 shows a breakdown of the cardiovascular procedures performed between 1990 and 1999 in the United States that required vascular grafts.^[3] The procedures are grouped into three categories based on graft diameter and listed in descending number of procedures performed:

- Coronary artery bypass grafts (CABGs, 1–4 mm diameter), used in aortocoronary bypass procedures. The National Center for Health Statistics (NCHS) reported an increase from 392,000 to 571,000 of these procedures performed between 1990 and 1999, despite the exponential increase observed in alternative treatments (coronary obstruction removal and stenting).
- Small-diameter grafts (4–6 mm), used for peripheral grafts to relieve lower-extremity ischemia and for vascular access grafts (VAGs) to treat end-stage renal disease (ESRD). The combined need for these two graft types is evident from the roughly 180,000 procedures performed each year.

- Large-diameter grafts (>7 mm), for use in the thoracic and abdominal cavities (mostly for the replacement of aortic sections), were required for a total of 40,000 procedures.

TYPES OF GRAFTS

Vascular grafts may be classified into two broad groups: bioprosthetic and synthetic. Bioprosthetic grafts are of biological origin and composed of autologous, homologous, or heterologous tissue. These tissue grafts may, in some cases, be stabilized by cross-linking with various fixatives and/or by reinforcement with synthetic materials. The second group comprises vascular prostheses fabricated entirely (or mostly) from synthetic materials.

Bioprosthetic Grafts

Autologous grafts (auto-grafts)

Fresh, viable autologous veins (especially saphenous veins) have been demonstrated to be the preferred grafts for use in relieving lower-extremity ischemia (e.g., femoropopliteal bypass grafting).^[4] Depending on its exact nature and position relative to the position of the bypass, the graft vein may either be excised and reimplanted in a reverse configuration (to prevent obstruction of flow by valves) or left in situ with only the proximal and distal ends repositioned and anastomosed to the occluded artery. The in-situ use of veins for arterial bypassing requires the removal of venous valves to allow unrestricted blood flow. Autologous veins (saphenous and jugular veins, and veins of the upper extremities) may further be used for coronary, carotid, renal, and other sites. The use of vein grafts may, however, be compromised by deterioration when exposed to arterial hemodynamic forces.^[5] In addition, the number of veins that can be transferred without causing venous disturbances is limited, and veins may be unsuitable due to preexisting conditions.

Autologous arteries may be considered the ideal arterial graft. In coronary artery bypass surgery, the



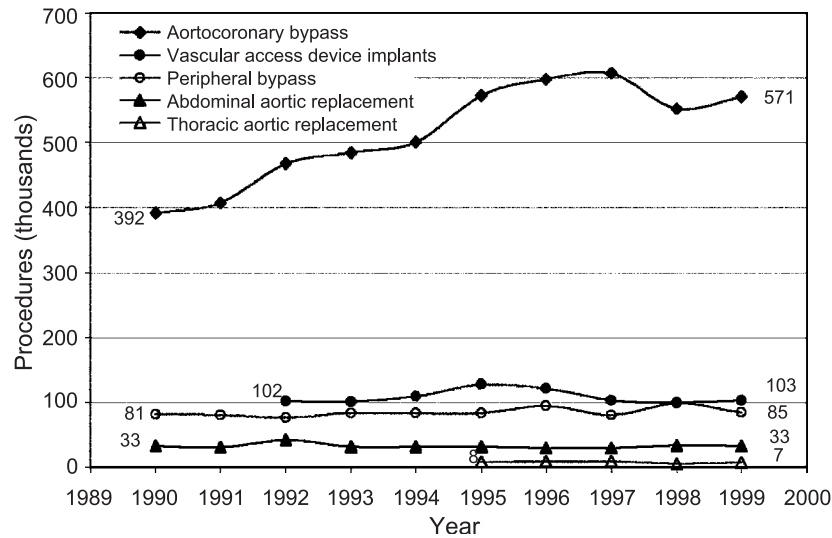


Fig. 1 Graphic representation of the number of clinical procedures involving vascular grafts performed from 1990 to 1999 in the United States. (From Ref. [3].)

pedicled left internal mammary artery (LIMA) has long been the superior conduit, but its application is generally limited to the left anterior descending (LAD) coronary artery. Depending on the anatomical position of the graft, the age of the patient, and other factors, further stenoses may be bypassed with either the pedicled or free right internal mammary artery (RIMA), radial arteries, and gastroepiploic arteries. Other specific examples of the use of autologous arteries include the application of resected autologous iliac arteries (which are subsequently replaced with a suitable synthetic prosthesis) to a more critical location, and the use of radial arteries in coronary positions of young patients. Apart from these specific severe cases, however, dispensable autologous arteries are even less readily available than dispensable veins.

Homologous grafts (allo-grafts)

Prior to the development of successful synthetic prostheses, homologous arterial tissue was the standard clinical large-bore material. This tissue type has been used in various forms: fresh, frozen, refrigerated in physiological solution, freeze-dried, glycerin preserved, alcohol preserved, dried, and preserved in formalin.^[4] Although arterial homografts are rarely used today, they might still find application as aortic arch replacements.

Glutaraldehyde-fixed homologous umbilical veins have found some application as bypass grafts for small-diameter arteries in the lower extremities. These grafts are typically externally supported with a light Dacron mesh to enhance structural integrity, and may offer a reasonable alternative to autologous veins when the latter are not available.^[6]

The shortcomings of arterial homografts, as evident from the changes that occur during the healing response, include slow wall lysis, compaction and loss of elastic tissue, ulceration, mural thrombosis, and calcification. Such changes resemble atherosclerosis^[7] and result in a poor longterm outcome. Although few nonarterial homograft tissues have been used in arterial positions, experiments evaluating homologous veins for aortic replacement show late fibrotic stenosis. Glutaraldehyde-fixed homologous umbilical veins have found some use in lower-extremity applications, but the general consensus is that they are not as reliable as autologous vein grafts.

Heterologous grafts (xeno-grafts)

Tissue fixation procedures similar to those used for the cross-linking of heterologous bioprosthetic heart valves have been used with some success to preserve heterologous arterial tissue. Typically, bovine carotid arteries were fixed in dialdehydes to reduce the antigenicity, reinforced with Dacron mesh to improve mechanical strength, and used as vascular access grafts.^[4] The use of heterologous vein grafts has not been reported to any great extent.

Heterologous arterial tissues demonstrate a decidedly accelerated tissue response, due to antigenicity, and are heavily subject to aneurysm formation and rupture. The antigenicity may be reduced by certain chemical modifications, such as cross-linking, but these grafts remain prone to aneurysmal dilation. This shortcoming has been corrected by Dacron reinforcement, and although studies have shown superior long-term fate compared with freeze-dried homografts and some textile

prostheses, the use of fixed heterologous arterial tissue is limited mostly to arteriovenous shunts.^[8]

Synthetic Grafts

The synthetic vascular graft market is dominated by two graft types: expanded polytetrafluoroethylene (ePTFE) grafts and textile polyethylene terephthalate (PET) grafts. Considerable effort has, however, been put into the development of grafts from alternative polymeric materials. Of these alternative materials, by far the most extensively reported are polyurethanes^[9] and bioresorbable polymers such as poly(glycolic acid) and poly(lactic acid).^[10] The latter type of grafts (bioresorbable) will be discussed with the alternative tissue engineering approaches.

Biostable grafts

The biostable approach to synthetic vascular grafts is based on the permanent incorporation of synthetic materials into the body. Initial investigations into the use of polymeric materials for vascular grafts concentrated on fibrillar materials that showed a high degree of in-vivo stability. These studies led to the acceptance of PET and PTFE (initially fibrillar, now expanded) as materials of sufficient stability to be used in commercially available vascular grafts. It has since been recognized that no material is completely inert in the hostile environment found in the human body, and even materials that are recognized as being extremely stable (gold, platinum, titanium) are corroded in vivo. Nonetheless, the demands placed on the polymeric materials used in durable vascular grafts require them to be chemically stable to a degree that ensures structural integrity and that prevents the generation of toxic or inflammatory breakdown products.

PET grafts

Polyethylene terephthalate (PET) is a semicrystalline polymer that is readily melt-spun and drawn into highly

crystalline filaments of high modulus and tensile strength. These filaments (10–20 μm in diameter) are typically bundled into multifilament yarns (24–108 filaments per yarn) and then woven or knitted to produce tubular or bifurcated grafts. The simplest configuration is a 1:1 or taffeta weave, in which the warp and weft of yarns are interwoven at right angles (in a one-to-one fashion), as seen in Fig. 2b. The warp yarns are aligned in the longitudinal direction of the graft (horizontal direction in Fig. 2), while the weft yarns run orthogonal to that direction. The taffeta weave results in the strongest construction, the highest fabric density, and the best dimensional stability, and is less prone to kinking than knits. Drawbacks of woven grafts include poor compliance, limited elongation (in the uncrimped form), and a tendency to fray. Woven prostheses are typically used in large-diameter emergency situations such as aortic aneurysms.

Knitted grafts contain yarns that are looped around each other as opposed to the interlacing found in woven grafts. If the yarns lie predominantly in the circumferential direction, the structure is termed a “weft knit.” The “jersey knit” is typical of this type of construction. To eliminate the dilation and unraveling found in weft-knit materials, warp knits, where the yarns lie predominantly in the longitudinal graft direction, were introduced (Fig. 2a). Many warp-knit configurations are possible; the most common ones are the two-bar warp locknit and the tricot warp knit. Warp knits have good dimensional stability and handling/suturing characteristics, and are less prone to fraying than their weft-knit counterparts. The surfaces of some knitted grafts are often roughened by the addition of textured yarns or by napping with wire brushes. This results in velour-type grafts that are reported to improve interlocking of tissue and thus improve healing. After knitting, chemical or heat treatment is used for compaction of the structure in order to decrease the water permeability of grafts to acceptable levels.^[11] PET grafts are generally sterilized by ethylene oxide gas or ionizing radiation, as dry or moist heat may further change the properties.

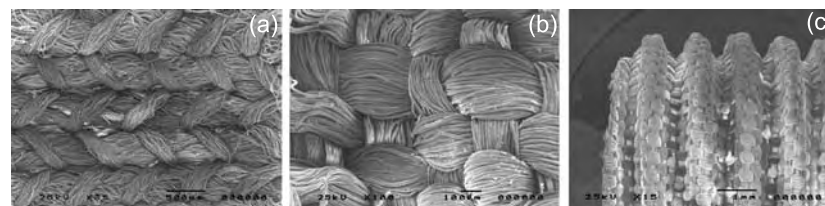


Fig. 2 Scanning electron micrographs of (a) the luminal surface of a Gelseal triaxial warp-knit Dacron vascular graft, 26 mm internal diameter, Sulzermedica (gelatin sealant removed by protease degradation for visualization of knitted structure) ($35\times$, bar = 500 μm); (b, c) the abluminal surface of a DeBakey 6010-S woven Dacron graft, 6 mm internal diameter, USCI, Division of CR Bard, Inc., with (b) showing the taffeta weave ($100\times$, bar = 100 μm) and (c) showing the crimped structure of the graft ($15\times$, bar = 1 mm).



Both woven and knitted grafts are often crimped to improve radial strength (to reduce kinking and collapse) and increase elongation potential (see Fig. 2c). To eliminate the need for surgeons to pre-clot large-diameter PET grafts, they are often sealed with collagen, albumin, or gelatin to completely eliminate blood permeability.^[12]

PET (Dacron) is the material of choice for large-diameter, high-flow vascular graft applications such as aortic replacements. The manufacturing techniques used in the production of PET grafts also lend themselves to the inclusion of side branches and bifurcations, and PET grafts are used as aortofemoral replacements. A secondary patency rate of 93% at 10 years can be achieved with large-diameter grafts of this type.^[13]

PTFE grafts

Polytetrafluoroethylene is a highly crystalline, inert material with high molecular mass, low coefficient of friction, and low surface tension. Although PTFE is not as readily processed by conventional melt and solution techniques as other thermoplastic materials, processing into fibers and subsequently into textile prostheses (similar to those obtained from PET fibers) is possible. Textile PTFE grafts were, however, withdrawn by many manufacturers due to unraveling and false aneurysm formation.

Expanded PTFE (ePTFE) grafts are produced by ram-extrusion of PTFE powder (with a lubricant), removal of the lubricant, longitudinal stretching at very high strain rates, and subsequent sintering. The resultant structure, consisting of solid nodes connected by interconnecting fibrils, may be seen in Fig. 3a. Some manufacturers use a less dense outer PTFE wrap (Fig. 3b) and either circular (Fig. 3c) or helical reinforcement (not shown). Sterilization of ePTFE grafts should be effected by ETO gas, steam, or dry heat, as ionizing radiation may cause polymer degradation.

Expanded PTFE is favored in medium-bore (6–12 mm) applications when autologous veins are unavailable or unsuitable, even though there is no

difference in the patency between ePTFE and PET of similar size.^[5] Patency rates of 40–50% after four years of implantation in femoro-popliteal reconstructions are lower than the rates achieved by using reversed saphenous veins (70–80%).^[14] ePTFE grafts fail to perform satisfactorily, however, in smaller-diameter applications (<6 mm) due to occlusion caused by thrombosis and intimal hyperplasia.

Polyurethane grafts

A wide range of segmented, thermoplastic polyurethanes (TPUs) have been developed specifically for medical applications. Their biocompatibility and elastomeric properties, combined with the fact that they may be processed by conventional melt and solution techniques, have been exploited over the past three decades in the search for a successful small-diameter vascular graft.

Polyurethane grafts may be classified into fibrillar (Fig. 4a) or foam-type (Fig. 4b and c) structures. Fibrillar grafts have been produced by conventional weaving and knitting, solution or melt spinning, electrostatic spinning, and other nonwoven techniques, whereas replamineform, thermal inversion, blowing/reticulation, gas expansion, and phase inversion have been employed in the production of foamed structures.^[9,15] Extractable porogens may be used in conjunction with phase separation techniques to introduce larger pores, with premixing of irregularly shaped porogens generally resulting in irregularly shaped pores, limited pore interconnectivity, and residual porogens remaining in the structure.^[16] A variation of the phase inversion/porogen extract method has, however, been developed to produce scaffolds with well-defined, dodecahedral pores with high interconnectivity and no residual porogen retainment (see Fig. 4c).^[15]

Although a number of polyurethane prostheses have been developed to the extent that they have been evaluated in clinical trials as A–V shunts and femoropopliteal grafts,^[17,18] they have not gained widespread acceptance or clinical use.

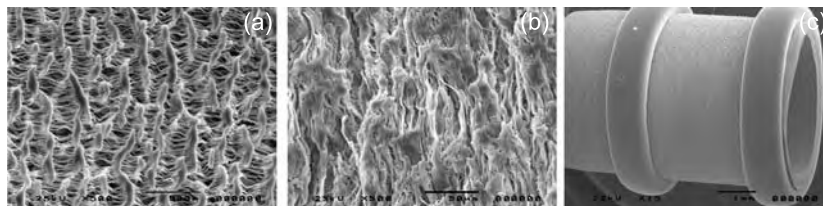


Fig. 3 Scanning electron micrographs of an ePTFE vascular graft, 4 mm internal diameter, WL Gore and Associates, showing (a) the luminal surface ($500\times$, bar = 50 μm), (b) the structure of the abluminal wrap, and (c) the external reinforcement rings.

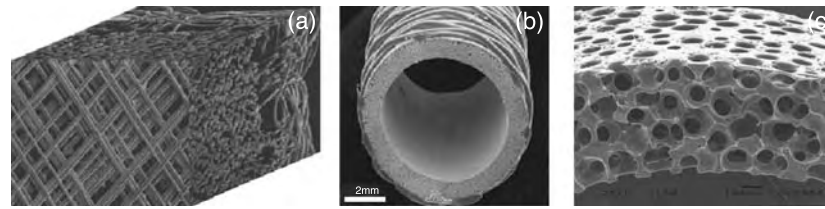


Fig. 4 Scanning electron micrographs of polyurethane vascular grafts. (a) Reconstructed 3-D image of a filament-wound Corvita vascular graft showing luminal surface and cross sections containing Dacron reinforcement ($200\times$). (b) Externally reinforced Myolink graft (bar = 2 mm). (c) Cross section of an experimental graft containing well-defined, interconnected pores produced by an adapted phase inversion technique (from Ref. [15]) ($100\times$, bar = $100\ \mu\text{m}$). [(a) Reprinted from Ref. [18] with permission from John Wiley & Sons. (b) Reprinted from Ref. [17], with permission from Elsevier Science.]

GRAFT PROPERTIES

The success of a synthetic vascular graft depends on many factors pertaining to its preimplant properties, its properties during implantation, and its ability to perform its vital function after implantation. The Association for the Advancement of Medical Instrumentation (AAMI) provides guidelines of graft standards and measurement techniques.^[19] Graft production should use cost-effective, environmentally acceptable processes that allow high and consistent quality. They should then allow for sterilization by conventional techniques and packaging/storage in a sterile state. The surgeon should be able to choose a graft of suitable diameter, wall thickness, and length and must then be able to trim the graft without undue fraying of the edges. During implantation, the graft should not only withstand the rigors of handling, trimming, and suturing, but also be of sufficient initial mechanical strength to withstand the blood pressure without excessive hemorrhage. In addition to the demand that toxic, carcinogenic, or excessively inflammatory products not be released, postimplantation properties of grafts (specifically synthetic grafts) are discussed in the following sections.

Porosity

Early investigations into the replacement of arteries with solid synthetic tubes (glass, aluminum, siliconized rubber, etc.) were unsuccessful due to thrombotic occlusion.^[20] Voorhees et al. were the first to recognize the importance of porosity when they performed a successful arterial replacement using a synthetic fabric in 1952.^[21] The importance of porosity was confirmed when a variety of synthetic materials, unsuccessful as solid-wall prostheses, were shown to be successful in porous configurations.^[22] Okoshi et al. extended the porosity principle to microporous polyurethane grafts by also showing a positive correlation between porosity and patency.^[23]

Dacron manufacturers often define the porosity of their grafts in terms of water permeability. Grafts are thus not only distinguished on the basis of the type of knit or weave, but are further characterized as being of high porosity ($1500\text{--}4000\ \text{ml}\cdot\text{cm}^{-2}\cdot\text{min}^{-1}$) or low porosity ($200\text{--}1000\ \text{ml}\cdot\text{cm}^{-2}\cdot\text{min}^{-1}$).^[24] The porosity of ePTFE grafts, on the other hand, is characterized by the average distance between the centers of adjacent nodes (internodular distance or IND). For most commercially available ePTFE grafts this value is typically in the order of $30\ \mu\text{m}$, although larger sizes have been produced. The morphology of foam-type porous scaffolds used in experimental polyurethane grafts is also generally expressed either in terms of water permeability or, where applicable, indirectly by stating the average size of the porogens used in their production. Although water permeability is a useful parameter for other reasons, it is not useful in the prediction or evaluation of the permissiveness of the material toward tissue ingrowth. Likewise, internodular distance is often a misleading dimension, as the actual space available for tissue ingrowth, namely the interfibrillar distance, is typically in the $4\text{--}5\ \mu\text{m}$ range.^[24] Thus, if one attempts to gain a better understanding of the relationship between porosity and healing, an additional parameter—average pore size available for tissue ingrowth—is required.

Mechanical Properties

The most basic function of a vascular graft, to provide a conduit for blood flow, necessitates two mechanical attributes: 1) the structural integrity to withstand internal pressure (both in the wall and at the suture line), and 2) sufficient resistance to kinking and flattening in order to retain an open lumen. Natural blood vessels exhibit further mechanical attributes, namely regulation of blood flow in response to physiological demands, and dilation and contraction in sympathy with the pulse pressure. The latter property, termed compliance, may be defined in terms of the diameter



change of a vessel between diastolic and systolic pressures, and is usually expressed as %/100 mm Hg.^[25] Equation 1 relates the diameter compliance (C_D) to the diameter (D) and pressure (P) changes, and to the radius (r), wall thickness (t), and bulk modulus of the wall (E_w):

$$C_D = \frac{(D_s - D_d)}{(P_s - P_d) \times D_d} = \frac{r}{t} \frac{1}{E_w} \quad (1)$$

where subscripts s and d denote systole and diastole, respectively. Convenient units for C_D are %/100 mmHg. In porous walls, the constitutive elements defining the wall (e.g., fibers) fill only a fraction of the wall volume. Additionally, these elements are not completely aligned in the circumferential direction, and the deformation is not purely tensile. In such cases the modulus of the vessel also depends on the frictional forces involved in the relative movement between individual elements, as well as the flexural and shear moduli of the elements themselves. It is for this reason that prostheses constructed from very stiff materials such as PET (high E_w) fibers may exhibit some, but not adequate, compliance.

As the structure and function of blood vessels change along the vascular tree, so do their mechanical properties. The larger elastic arteries are highly compliant—so much so that the ascending aorta ($C_D = 30\%/100$ mm Hg^[26]) stores approximately 40% of the cardiac output volume during systole and then releases that volume during diastole. This “windkessel” effect is responsible for the smoothing out of the pressure pulse and contributes to the diastolic pressure. The compliance decreases as one moves away from the heart, and a typical value for the compliance of a femoral artery is 7 to 14%/100 mm Hg.^[27] It is also important to note that, in addition to the variation of compliance between healthy individuals within the same age group, the compliance is a function of age and may be further influenced by disease (notably arteriosclerosis, hypertension, or diabetes). The limited compliance of ePTFE ($C_D = 1.6 \pm 0.2\%/100$ mm Hg) and Dacron ($C_D = 1.9 \pm 0.3\%/100$ mm Hg) grafts has long been indicated as a contributory factor in small-diameter graft failure. A significant correlation between patency and compliance supports this view.^[25] Increased impedance and decreased capacitance, causing disturbed flow patterns, have been cited as causes for intimal hyperplasia that results in the failure of undercompliant grafts.

Although various researchers have attempted to mimic natural vessel compliance by using elastomeric polyurethanes,^[17] the fact that very few of these grafts have progressed past clinical trials indicates that, although compliance may be an important factor, it is not a sufficient factor in the determination of graft success.

Biocompatibility

The nature of the reaction elicited by the implantation of a synthetic graft determines not only the short-term fate of the implant, but also its long-term stability and function. The sequence of events that follows implantation of medical devices—namely injury, acute and chronic inflammation, the formation of granulation tissue, foreign body reaction, and finally fibrosis and fibrous encapsulation—are described by Anderson.^[28] Complete inertness, a property, that has been shown to be unattainable with current materials, may actually be an undesirable property, as it may cause incomplete functional incorporation and healing of the device due to limited interaction with the living tissue. It is for this reason that biomaterials are now being rendered “bioactive” by a variety of surface and bulk modification techniques in order to promote constructive interaction between the implant and the host tissue.

Surface modification is a useful tool that allows the alteration or masking of the chemical nature of the implant that is exposed to the surrounding tissue, as it may be used to alter this property without undue influence on physical properties that may have been optimized for the intended application. In addition to synthetic hydrogels, modification agents include materials of biological origin that may render the material bioactive, i.e., that actively interact with the host tissue by mimicking naturally occurring materials. Other approaches rely on the immobilization of oligopeptide sequences (identified as being the active binding sites in proteins) to obtain the desired effect.

The surfaces of biomaterials may be modified for a number of reasons, all intended to improve the biocompatibility of the material. The reasons include decreasing the adhesion of proteins and/or inflammatory cells, decreasing the thrombogenicity of the material, and increasing the adhesion, proliferation, or migration of desired cell types on or in the material. Leucocyte and macrophage adhesion (and differentiation of macrophages into FBGCs) plays a pivotal role in the inflammatory response. The adhesion of these cell types has been shown to depend on several factors, including surface chemistry, hydrophilicity, and protein adsorption.^[29]

Polyethylene glycols have been shown to be very versatile polymers in surface modification applications. They have been shown not only to decrease protein and macrophage adhesion and bacterial adhesion,^[30] but also to be effective antithrombotic agents.^[31] Thrombotic events on vascular grafts have also been decreased by the immobilization of a number of active antithrombotic agents. These include acetyl salicylic acid (ASA or aspirin),^[32] hirudin,^[33] dipyridamole,^[34] phosphorylcholine,^[35] and heparin.^[36]

The improved incorporation of synthetic devices may be facilitated by the immobilization of extracellular matrix materials such as collagen.^[37] Collagen and other extracellular matrix materials contain distinct peptide sequences that have been identified as the adhesion sites for certain cells. These include the well-known RGD sequence, as well as the YIGSR and IKVAV sequences. These oligopeptides (with flanking sequences) have been covalently immobilized on biomaterials with the aim of improving tissue interaction.^[38,39]

Graft Healing

In order to understand the in-vivo response to vascular grafts, it is important to distinguish between transanastomotic and transmural healing. In the past, very little distinction was made between these modes of healing, and healing patterns were described without specific reference to the source of the tissue. Due to the differences in healing responses observed in different animal models, knowledge about the tissue source is of critical importance. A graft that may show complete surface healing (and even complete surface endothelialization) in one model due to transanastomotic outgrowth, will not show the same favorable response in clinical applications.

Transanastomotic healing

Shortly after implantation of a vascular graft, smooth muscle (and other) cells in the media of the native artery start to proliferate and migrate through the damaged internal elastic lamina into the intima.^[40] This hyperplastic tissue subsequently migrates and proliferates onto the graft (from each anastomosis) toward the midgraft area, forming a “tongue” of tissue generally referred to as the pannus.

Three major factors influence the extent of pannus growth. The first two, the porosity of the graft and the presence of tissue under the inner fibrinous capsule, may be closely interrelated. The third factor affecting transanastomotic outgrowth is the animal model. Although the proliferative capacity of human endothelial cells does not differ significantly from that of experimental animals,^[41] there is a marked difference in the extent of endothelialization observed in vivo between the human and experimental animal models. For reasons unknown, transanastomotic endothelialization stops after only a short distance from the anastomosis in humans.^[42] On standard-porosity ePTFE (30 μm), the coverage achieved in humans in 56 weeks occurs in dogs, the yellow baboon, and the chacma baboon in 3.5, 5.6, and 7.6 weeks, respectively.^[24] Thus, the chacma baboon, in which transanastomotic outgrowth

is 2.1 times slower than in the most commonly used model (the dog), seems to be the best choice for the evaluation of transanastomotic endothelialization of grafts intended for man.

Transmural healing

Peripheral vascular grafts used to alleviate lower-extremity ischemia (femoropopliteal) may be up to 60 cm long. This fact, combined with the very short distance the endothelium is able to migrate (in humans) onto the surfaces of such grafts by transanastomotic surface healing, indicates that future grafts intended for clinical use should rely more on transmural healing. If full transmural healing could be achieved, it would not only lead to the formation of a viable endothelium, but would also result in the formation of viable tissue in the graft wall, and thus to improved incorporation of the prosthetic material by the host. The mural tissue should consist, as previously indicated, not only of smooth muscle cells and accompanying extracellular matrix, but also of a fully developed vasa vasorum capable of sustaining the viability of the cells.

The ability of tissue to migrate through the graft wall is, again, dependent on both the structure of the graft and the proliferative and migratory capacity of the cells of the host. The need for sufficient porosity to accommodate cellular ingrowth has been discussed. Standard-porosity ePTFE, due to its restricted ingrowth spaces, does not allow for significant transmural ingrowth, irrespective of the animal model used. In 60 μm ePTFE, transmural ingrowth of capillary vessels leads to the surface endothelial in as little as two weeks in the yellow baboon, whereas the perigraft tissue does not even penetrate the outer two-thirds of these grafts in the chacma baboon within the same time frame.^[24]

The healing of Dacron grafts is also greatly affected by porosity. Low-porosity Dacron prostheses (woven) show similar perigraft incorporation as standard (30 μm) ePTFE during the initial period after implantation. Subsequently, some fibroblast and capillary infiltration into the narrow spaces between the Dacron yarns is observed.^[4] Even if the tissue does manage to grow through the tight interfibrillar spaces, however, it is unable to penetrate the compact inner fibrin layer deposited on the graft surface.^[43] In high-porosity Dacron (knitted), transmural tissue reaches the compact luminal fibrin layer within three to four weeks in animals and three to six months in humans.^[42] The complete endothelialization of long prostheses, comparable to that in highly porous ePTFE grafts, is occasionally seen in knitted prostheses. However, in contrast to high-porosity ePTFE grafts, it remains unclear whether this endothelium is derived from



transmural tissue ingrowth, facilitated transanastomotic outgrowth, or fall-out healing.^[44]

Alternative Approaches

In-vitro tissue engineering

Recent advances in cell culture techniques have created an opportunity for the in-vitro reconstitution of organs from living cells. Weinberg and Bell^[45] were the first to apply such techniques to produce a completely biological tissue-engineered blood vessel (TEBV) from collagen gels and cultured bovine endothelial cells (ECs), smooth muscle cells (SMCs), and fibroblasts. These grafts did not, however, display the required mechanical strength, even when reinforced with Dacron mesh. Similar results were obtained by the group of Auger with human collagen and vascular cells,^[46] and by Hirai and Matsuda,^[47] who used a canine model.

An alternative method entails the production of tissue-engineered blood vessels from cultured human cells without the use of any exogenous biological material, by allowing the cells to create their own extracellular matrix in the presence of ascorbic acid.^[48] The method involved wrapping successive layers of acellular and cellular materials onto a mandrel, removal of the mandrel, and subsequent EC seeding to produce a neo-artery. The vessels were shown to have undergone histological organization and displayed burst strengths comparable to those of natural blood vessels.

Niklason and Langer^[2] used temporary synthetic scaffolds of poly(glycolic acid) (PGA) to produce tissue-engineered blood vessels by seeding tubular nonwoven meshes of PGA with high densities of smooth muscle cells. Maturation of the cells in a bioreactor resulted in the proliferation of the cells and the degradation of the PGA mesh. Pulsatile stretch was employed to facilitate proliferation and enhance the contractile properties of the cells. Although the mechanical stretching did improve the physical strength over that obtained during static culture, the vessels were not able to withstand physiological arterial pressure.

Endothelial cell transplantation

The failure of small-diameter, synthetic, vascular grafts may be attributed mostly to surface thrombogenicity and anastomotic intimal hyperplasia. The surface thrombogenicity, in turn, can be ascribed to a persistent lack of endothelial cell coverage, even after extended implant periods. The basic concept behind endothelial cell transplantation is the lining of

autologous endothelial cells (ECs) on contemporary synthetic grafts before implantation in order to mask the thrombotic nature of the underlying synthetic material.

Single-stage venous EC seeding

This early method involved the single-staged seeding of endothelial cells harvested from subcutaneous veins onto Dacron prostheses, and the subsequent clotting of the cells into the wide-meshed porous structure of the synthetic graft.^[49] The general trend at the time toward the use of ePTFE grafts in peripheral reconstructions prompted researchers to attempt the application of this approach to an alternative material. The narrow interfibrillar spaces of the ePTFE, however, resulted in the cells being clotted onto the luminal surface of the graft instead of into the graft structure. Seeding with cell numbers that were too low for the large surface area of long grafts needed for human implantation, together with the high detachment rate of cells after the restoration of blood flow, resulted in suboptimal seeding densities. Although this probably led to partial endothelialization and moderate decrease in thrombogenicity, the procedure did not result in an improved clinical patency rate.^[50] Newer approaches to circumvent the problem of low seeding densities due to low cell numbers include the mass culture of venous endothelial cells and the mass harvest of capillary endothelial cells from fat tissue.^[51]

In-vitro endothelialization

In-vitro endothelialization was pioneered in the 1980s by a number of groups.^[52] This is essentially a two-stage method, the first stage of which is the mass culture of endothelial cells isolated from a suitable source (e.g., a section of jugular or cephalic vein). The large quantity of endothelial cells cultivated in this step is subsequently seeded at a high density onto the luminal surface of a suitable synthetic prosthesis, typically an ePTFE graft. Precoating of the graft prior to seeding improves the adhesion of the seeded cells. Post-maturation allows for the recovery from injury caused by the seeding procedure, the maturation of the cells' cytoskeletons to better withstand the shear forces exerted on it by the bloodstream, the formation of a confluent layer of cells on the luminal surface, and the reduction of the inflammatory activation of the endothelium.^[53]

Clinical trials utilizing in-vitro endothelialized grafts for femoropopliteal bypass commenced in 1989.^[54] Since then, a number of groups have reported their experience with clinical trials with follow-up periods of up to nine years. The overall success of the procedure is

evident from the excellent improvement achieved in patency rates when endothelialized grafts are compared to unendothelialized controls. Control groups typically have a five-year patency rate of 42% for above-knee ePTFE grafts, whereas patency rates in excess of 65% could be achieved when endothelialized grafts were used.^[52] This compares favorably to the 70% patency rate achieved by Veith with the use of reversed autologous saphenous veins in the same position.^[14] Multicenter clinical trials aimed at the verification of these results are currently under way in Europe.

In-vivo tissue engineering

The in-vivo approach to vascular grafts typically involves the use of degradable materials that allow tissue ingrowth and remodeling in order to produce a neo-artery. Synthetic materials used in such constructs include poly(glycolic acid) (PGA), poly(lactic acid) (PLA), PLGA copolymers, polycaprolactone (PCL), and polydioxanone (PDS). As these materials are generally nonelastomeric (and thus have limited compliance), elastic degradable materials composed of PLGA/PCL/PLGA tri-blocks, PLA/PGA blends with polyurethanes,^[55] hydrolytically degradable polyurethanes,^[56,57] and enzymatically degradable polyurethane^[58] are being developed.

Researchers in this field endeavor to effect the generation of a new vessel by the degradation of porous scaffold after ingrowth of sufficient cellular and extracellular materials (before degradation resulted in failure of the graft), followed by complete (or partial) resorption of the polymeric material.^[10,55] A major challenge in this approach is to ensure that the breakdown products are not excessively toxic, allergenic, thrombogenic, inflammatory, or carcinogenic.^[59]

Naturally occurring materials such as collagen have also been employed in the in-vivo approach. Previous use of collagenous materials as vascular grafts typically involved the use of extensive cross-linking to achieve a durable, permanent implant. In contrast, Sullivan and Brockbank^[60] describe the use of water-soluble carbodiimide (EDC) to minimally cross-link porcine intestinal collagen layers (ICLs), thereby reducing the antigenicity of the material without preventing in-vivo cell ingrowth. Explanted carotid interposition grafts in rabbits showed excellent patencies, and preliminary histology exhibited signs of remodeling of the graft by host cells. Van Wachem et al. cross-linked reconstituted type I bovine collagen to various degrees using EDC and were able to achieve decreased antigenicity and variable degradation rates in vivo.^[61]

An novel approach by Campbell et al. involved the growth of a graft within the recipient's peritoneal

cavity.^[62] This method required no synthetic or biologic matrix, apart from a piece of Silastic tubing that was implanted in the peritoneal cavity and acted as a temporary mandrel. After removal of the mandrel from the explanted mesothelium-covered myofibroblast capsule, and inversion of the construct to position the mesothelium on the luminal surface, the living grafts were successfully implanted in both the rat aortic and rabbit carotid models. More information on vascular tissue engineering may be found in an editorial by Edelman.^[63]

CONCLUSION

Autologous arteries and veins remain the best conduits for replacing or bypassing diseased and injured vessels, and are used in preference to other prostheses. Limitations on their use, caused by lack of availability of dispensable vessels and unsuitability due to preexisting conditions, necessitate the use of alternative grafts. Inherent problems associated with the use of fixed homologous or heterologous grafts imply that synthetic grafts (ePTFE and Dacron) are currently the only viable alternatives. The failure of contemporary Dacron and ePTFE grafts to perform optimally in small-diameter applications may be attributed to the poor healing response exhibited by the body toward the implanted prostheses. This lack of healing, in turn, may be attributed to a number of factors that prevent the graft from being incorporated into a viable composite of synthetic material and living tissue. Although endothelial cell transplantation onto (or into) conventional synthetic prostheses has been successful in improving the patency rates of such grafts, the negative aspects associated with both in-vitro cell culture and conventional synthetic grafts are maintained.

The in-vitro tissue engineering approaches, as well as the in-vivo approaches that require the growing of the vessel prior to graft implantation, offer the possibility of having vascular grafts "made to order" from autologous cells. Although these approaches offer invaluable insight into the conditions required to produce neo-arteries, the turnaround times from initiation to functional graft implantation may limit the application of these techniques. Thus it seems that the holy grail of vascular grafting will remain a device that is "ready to use" or "off the shelf," that would immediately be able to perform the basic functions of a graft, and that would further provide the complex and interrelated chemical, physical, and biological properties required to effect remodeling to a fully functioning neo-artery. Whether this graft will turn out to be made from synthetic or biological degradable materials, or whether durable materials may be used in vivo to form a living, composite tissue/polymer



construct, is the question that fuels research in this exiting field.

ACKNOWLEDGMENTS

The authors wish to thank Mr. Dmitri Lennert for his assistance in the preparation of the scanning electron micrographs.

ARTICLES OF FURTHER INTEREST

Fabrics; Poly(carbonate)urethanes; Polytetrafluoroethylene; Tissue Engineering of Blood Vessel; Vascular Grafts: Development Strategies; Vascular Grafts: Host Interactions; Vascularization of Engineered Constructs

REFERENCES

- Zdrahala, R. Small caliber vascular grafts. Part I: State of the art. *J. Biomater. Appl.* **1996**, *10* (4), 309–329.
- Niklason, L.E.; Langer, R.S. Advances in tissue engineering of blood vessels and other tissues. *Transpl. Immunol.* **1997**, *5* (4), 303–306.
- National Center for Health Statistics. National Hospital Discharge Survey. In *Vital Health Statistics*; Hyattsville, Massachusetts, 1990–1999.
- Wesolow, A. Biological Behaviour of Tissue and Prosthetic Grafts. In *Vascular Surgery. Principles and Techniques*; Haimovici, H., Ed.; Appleton-Century-Crofts: Norwalk, CT, 1984; 93–118.
- Bos, G.; Poot, A.; Beugeling, T.; van Aken, W.; Feijen, J. Small-diameter vascular graft prostheses: Current status. *Arch. Physiol. Biochem.* **1998**, *106* (2), 100–115.
- Dardik, H. Biological Behaviour of Glutaraldehyde—Stabilized Human Umbilical Cord Vein Grafts. In *Vascular Surgery. Principles and Techniques*; Haimovici, H., Ed.; Appleton-Century-Crofts: Norwalk, CT, 1984.
- Sauvage, L.; Wesolowski, S. Healing and fate of arterial grafts. *Surgery* **1955**, (38), 1090.
- Butt, K. Bovine Heterograft for Arteriovenous Fistula. In *Vascular Grafting*; Sawyer, P., Kaplitt, M., Eds.; Appleton-Century-Crofts: New York, 1978; 278–281.
- Zdrahala, R. Small caliber vascular grafts. Part II: Polyurethanes revisited. *J. Biomater. Appl.* **1996**, *11* (1), 37–61.
- Fox, D.; Vorp, D.; HP, G. Bioresorbable Grafts: A Counterintuitive Approach. In *Tissue Engineering of Vascular Prosthetic Grafts*; Zilla, P., Greisler, H.P., Eds.; R.G. Landes: Texas, 1999; 489–503.
- King, M.; Blais, P.; Guidoin, R.; Prowse, E.; Marcois, M.; Gosselin, C.; Noel, H. Polyethylene Terephthalate (Dacron) Vascular Prostheses—Material and Fabric Construction Aspects. In *Biocompatibility of Clinical Implant Materials*; CRC Press, 1981; 177–207.
- von Oppell, U.; Zilla, P. Tissue adhesives in cardiovascular surgery. *J. Long-Term Eff. Med. Implants* **1998**, *8* (2), 87–101.
- Sladen, J.; Gilmour, J.; Wong, R. Cumulative patency and actual palliation in patients with claudication after aortofemoral bypass. Prospective long-term follow-up of 100 patients. *Am. J. Surg.* **1986**, *152* (2), 190–195.
- Veith, F.J.; Gupta, S.K.; Ascer, E.; White-Flores, S.; Samson, R.H.; Scher, L.A.; Towne, J.B.; Bernhard, V.M.; Bonier, P.; Flinn, W.R.; Astelford, P.; Yao, J.S.T.; Bergan, J.J. Six-year prospective multicenter randomized comparison of autologous saphenous vein and expanded polytetrafluoroethylene grafts in infrainguinal arterial reconstructions. *J. Vasc. Surg.* **1986**, *3* (1), 104–114.
- Bezuidenhout, D.; Davies, N.; Zilla, P. Effect of well-defined dodecahedral porosity on inflammation and neo-vascularization. *ASAIO J.* **2002**, *48*, 465–471.
- Nam, Y.; Park, T. Porous biodegradable polymeric scaffolds prepared by thermally induced phase separation. *J. Biomed. Mater. Res.* **1999**, *47* (1), 8–17.
- Tiwari, A.; Salacinski, H.; Seifalian, A.; Hamilton, G. New prostheses for use in bypass grafts with special emphasis on polyurethanes. *Cardiovasc. Surg.* **2002**, *10* (3), 191–197.
- Eberhart, A.; Zhang, Z.; Guidoin, R.; Laroche, G.; Guay, L.; De La Faye, D.; Batt, M.; King, M.W. A new generation of polyurethane vascular prostheses: Rara avis or ignis fatuus? *J. Biomed. Mater. Res.* **1999**, *48* (4), 546–558.
- Cardiovascular Implants—Tubular Vascular Prostheses*; Association for the advancement of Medical Instrumentation, 1998.
- Haimovici, H. History of Vascular Surgery. In *Vascular Surgery. Principles and Techniques*; Haimovici, H., Ed.; Appleton-Century-Crofts: Norwalk, CT, 1984; 3–18.
- Voorhees, A.; Jaretzki, A.; AH, B. The use of tubes constructed from Vinyon “N” cloth in bridging arterial defects. *Ann. Surg.* **1952**, *135*, 332.
- Wesolowski, S. Foundations of Modern Vascular Grafts. In *Vascular Grafts*; Sawyer, P., Kaplitt, M., Eds.; Appleton Century Croft: New York, 1978.
- Okoshi, T.; Soldani, G.; Goddard, M.; Galletti, P. Penetrating micropores increase patency and achieve extensive endothelialization in small diameter polymer skin coated vascular grafts. *ASAIO J.* **1996**, *42* (5), M398–401.
- Dauids, L.; Dower, T.; Zilla, P. The Lack of Healing in Conventional Vascular Grafts. In *Tissue Engineering of Vascular Grafts*; Zilla, P., Greisler, H., Eds.; RG Landes: Austin, 1999; 3–44.
- Seifalian, A.; Giudiceandrea, A.; Schmitz-Rixen, T. Noncompliance: The Silent Acceptance of a Villain. In *Tissue Engineering of Vascular Prosthetic Grafts*; P, Z., HP, G., Eds.; R.G. Landes: Texas, 1999.
- Stefanadis, C.; Stratos, C.; Boudoulas, H.; Kourouklis, C.; Toutouzas, P. Distensibility of the ascending aorta: Comparison of invasive and non-invasive techniques in healthy men and in men with coronary artery disease. *Eur. Heart J.* **1990**, *11* (11), 990–996.

27. Kawasaki, T.; Sasayama, S.; Yagi, S.; Asakawa, T.; Hirai, T. Non-invasive assessment of the age related changes in stiffness of major branches of the human arteries. *Cardiovasc. Res.* **1987**, *21*, 678–687.
28. Anderson, J.; Gristina, A.; Hanson, S.; Harker, L.; Johnson, R.; Merritt, K.; Naylor, P.; Schoen, F. Host Reaction to Biomaterials and Their Evaluation. In *Biomaterials Science: An Introduction to Materials in Medicine*; Ratner, B., Hoffman, A., Schoen, F., Lemons, J., Eds.; Academic Press: San Diego, 1996.
29. Yun, J.; DeFife, K.; Colton, E.; Stack, S.; Azeez, A.; Cahalan, L.; Verhoeven, M.; Cahalan, P.; Anderson, J. Human monocyte/macrophage adhesion and cytokine production on surface-modified poly(tetrafluoroethylene/hexafluoropropylene) polymers with and without protein preadsorption. *J. Biomed. Mater. Res.* **1995**, *29* (2), 257–268.
30. Park, K.; Kim, Y.; Han, D.; Kim, Y.; Lee, E.; Suh, H.; Choi, K. Bacterial adhesion on PEG modified polyurethane surfaces. *Biomaterials* **1998**, *19* (7–9), 851–859.
31. Lee, J.; Ju, Y.; Lee, W.; Park, K.; Kim, Y. Platelet adhesion onto segmented polyurethane surfaces modified by PEO- and sulfonated PEO-containing block copolymer additives. *J. Biomed. Mater. Res.* **1998**, *40* (2), 314–323.
32. San Roman, J.; Bujan, J.; Bellon, J.; Gallardo, A.; Escudero, M.; Jorge, E.; de, H.J.; Alvarez, L.; Castillo, O.J. Experimental study of the antithrombogenic behavior of Dacron vascular grafts coated with hydrophilic acrylic copolymers bearing salicylic acid residues. *J. Biomed. Mater. Res.* **1996**, *32* (1), 19–27.
33. Phaneuf, M.; Berceci, S.; Bide, M.; Quist, W.; LoGerfo, F. Covalent linkage of recombinant hirudin to poly(ethylene terephthalate) (Dacron): Creation of a novel antithrombin surface. *Biomaterials* **1997**, *18* (10), 755–765.
34. Aldenhoff, Y.; van, D.V.F.; ter, W.J.; Habets, J.; Poole, W.L.; Koole, L. Performance of a polyurethane vascular prosthesis carrying a dipyridamole (Persantin) coating on its luminal surface. *J. Biomed. Mater. Res.* **2001**, *54* (2), 224–233.
35. Yoneyama, T.; Sugihara, K.; Ishihara, K.; Iwasaki, Y.; Nakabayashi, N. The vascular prosthesis without pseudointima prepared by antithrombogenic phospholipid polymer. *Biomaterials* **2002**, *23* (6), 1455–1459.
36. Esquivel, C.; Bjorck, C.; Bergentz, S.; Bergqvist, D.; Larsson, R.; Carson, S.; Dougan, P.; Nilsson, B. Reduced thrombogenic characteristics of expanded polytetrafluoroethylene and polyurethane arterial grafts after heparin bonding. *Surgery* **1984**, *95* (1), 102–107.
37. Kinoshita, Y.; Kuzuhara, T.; Kirigakubo, M.; Kobayashi, M.; Shimura, K.; Ikada, Y. Soft tissue reaction to collagen-immobilized porous polyethylene: Subcutaneous implantation in rats for 20 wk. *Biomaterials* **1993**, *14* (3), 209–215.
38. Krijgsman, B.; Seifalian, A.; Salacinski, H.; Tai, N.; Punshon, G.; Fuller, B.; Hamilton, G. An assessment of covalent grafting of RGD peptides to the surface of a compliant poly(carbonate-urea)urethane vascular conduit versus conventional biological coatings: Its role in enhancing cellular retention. *Tissue Eng.* **2002**, *8* (4), 673–680.
39. Kao, W.; Lee, D.; Schense, J.; Hubbell, J. Fibronectin modulates macrophage adhesion and FBGC formation: The role of RGD, PHSRN, and PRRARV domains. *J. Biomed. Mater. Res.* **2001**, *55* (79–88).
40. Hamdan, A.; Misare, B.; Contreras, M.; LoGerfo, F.; Quist, W. Evaluation of anastomotic hyperplasia progression using the cyclin specific antibody MIB-1. *Am. J. Surg.* **1996**, *172* (2), 168–170, discussion 170–161.
41. Zilla, P.; Fasol, R.; Dudeck, U.; Siedler, S.; Preiss, P.; Fischlein, T.; Muller, G.W.; Baitella, G.; Sanan, D.; Odell, J.; Reichart, B. In situ cannulation, microgrid follow-up and low-density plating provide first passage endothelial cell masscultures for in vitro lining. *J. Vasc. Surg.* **1990**, *12* (2), 180–189.
42. Sauvage, L.; Berger, K.; Wood, S.; Nakagawa, Y.; Mansfield, P. An external velour surface for porous arterial prostheses. *Surgery* **1971**, *70* (6), 940–953.
43. Berger, K.; Sauvage, L.; Rao, A.; Wood, S. Healing of arterial prostheses in man: Its incompleteness. *Ann. Surg.* **1972**, *175* (1), 118–127.
44. Hammond, W. Surface Population with Blood-Borne Cells. In *Tissue Engineering of Prosthetic Vascular Grafts*; Zilla, P., Greisler, H., Eds.; Landes Bioscience: Austin, 1998.
45. Weinberg, C.; Bell, E. A blood vessel model constructed from collagen and cultured vascular cells. *Science* **1986**, *231*, 397–400.
46. L'Heureux, N.; Germain, L.; Labbe, R.; Auger, F. In vitro construction of a human blood vessel from cultured vascular cells; A morphologic study. *J. Vasc. Surg.* **1993**, *17* (499–509).
47. Hirai, J.; Matsuda, T. Venous reconstruction using hybrid vascular tissue composed of vascular cells and collagen tissue regeneration process. *Cell Transplant* **1996**, *5* (93–105).
48. L'Heureux, N.; Paquet, S.; Labbe, R.; Germain, L.; Auger, F. A completely biological tissue-engineered human blood vessel. *FASEB* **1998**, *12* (47–56).
49. Herring, M.B.; Dilley, R.; Jersild, R.A.; Boxer, L.; Gardner, A.; Glover, J. Seeding arterial prostheses with vascular endothelium. The nature of the lining. *Ann. Surg.* **1979**, *190* (1), 84–90.
50. Zilla, P.; Fasol, R.; Deutsch, M.; Fischlein, T.; Minar, E.; Hammerle, A.; Krupicka, O.; Kadletz, M. Endothelial cell seeding of polytetrafluoroethylene vascular grafts in humans: A preliminary report. *J. Vasc. Surg.* **1987**, *6* (6), 535–541.
51. Jarrell, B.; Williams, S.; Stokes, G. Use of freshly isolated capillary endothelial cells for the immediate establishment of a monolayer on a vascular graft at surgery. *Surgery* **1986**, *100*, 392–399.
52. Zilla, P.; Deutsch, M.; Meinhart, J. Endothelial cell transplantation. *Semin. Vasc. Surg.* **1999**, *12* (1), 52–63.
53. Gillis-Haegerstrand, C. Adhesion Molecule Expression Following In Vitro Lining. In *Tissue Engineering of Prosthetic Vascular Grafts*; Zilla, P., Greisler, H., Eds.; RG Landes: Austin, 1999.
54. Zilla, P.; Deutsch, M.; Meinhart, J.; Puschmann, R.; Eberl, T.; Minar, E.; Dudczak, R.; Lugmaier, H.;



- Schmidt, P.; Noszian, I.; Fischlein, T. Clinical in vitro endothelialization of femoropopliteal bypass grafts: An actuarial follow-up over three years. *J. Vasc. Surg.* **1994**, *19* (3), 540–548.
55. van der Lei, B.; Wildevuur, C. Microporous, Compliant, Biodegradable Small-Calibre Vascular Grafts. In *Current Perspectives on Implantable Devices*; Williams, G., Ed.; Jai Press: London, 1990.
56. Gorna, K.; Gogolewski, S. Biodegradable polyurethanes for implants. II. In vitro degradation and calcification of materials from poly(epsilon-caprolactone)-poly(ethylene oxide) diols and various chain extenders. *J. Biomed. Mater. Res.* **2002**, *60* (4), 592–606.
57. Bruin, P.; Smedinga, J.; Pennings, A.; Jonkman, M. Biodegradable lysine diisocyanate-based poly(glycolide-co-epsilon-caprolactone)-urethane network in artificial skin. *Biomaterials* **1990**, *11* (4), 291–295.
58. Fromstein, J.D.; Woodhouse, K.A. Elastomeric biodegradable polyurethane blends for soft tissue applications. *J. Biomater. Sci., Polym. Ed.* **2002**, *13* (4), 391–406.
59. Xue, L.; Griesler, H. Blood Vessels. In *Principles of Tissue Engineering*; Lanza, R., Langer, R., Vacanti, J., Eds.; Academic Press: San Diego, 2000; 427–446.
60. Sullivan, S.; Brockbank, K.G.M. Small-Diameter Vascular Grafts. In *Principles of Tissue Engineering*; Lanza, R., Langer, R., Vacanti, J., Eds.; Academic Press: San Diego, 2000; 447–454.
61. Van Wachem, P.; Plantinga, J.; Wissink, M.; Beernink, R.; Poot, A.; Engbers, G.; Beugeling, T.; van, A.W.; Feijen, J.; van, L.M. In vivo biocompatibility of carbodiimide-crosslinked collagen matrices: Effects of crosslink density, heparin immobilization, and bFGF loading. *J. Biomed. Mater. Res.* **2001**, *55* (3), 368–378.
62. Campbell, J.; Efendy, J.; Campbell, G. Novel vascular graft grown within recipient's own peritoneal cavity. *Circ. Res.* **1999**, *85* (12), 1173–1178.
63. Edelman, E. Vascular tissue engineering: Designer arteries. *Circ. Res.* **1999**, *85* (12), 1115–1117.

Vascular Grafts: Development Strategies

James Laredo

Lian Xue

Division of Peripheral Vascular Surgery, Loyola University Medical Center, and Edward Hines Jr. Veterans Administration Hospital, Loyola University Medical Center, Maywood, Illinois, U.S.A.

Howard P. Greisler

Division of Peripheral Vascular Surgery, Loyola University Medical Center, Department of Cell Biology, Neurobiology and Anatomy, Loyola University Medical Center, and Edward Hines Jr. Veterans Administration Hospital, Loyola University Medical Center, Maywood, Illinois, U.S.A.

INTRODUCTION

Vinyon-N, developed in 1952 by Voorhees, was the first successful artificial vascular graft.^[1] Since then, polyethylene terephthalate (Dacron) and expanded polytetrafluoroethylene (ePTFE) have become the most widely used prosthetic vascular grafts. However, the development of completely nonreactive substances has been unsuccessful. Both Dacron and ePTFE react with blood components and perigraft tissues in a manner that is both beneficial and detrimental. Investigations of the vascular response in both physiological and pathological conditions, along with parallel developments in the field of biomaterials, have led to a more sophisticated strategy—namely, optimizing the tissue–biomaterial interaction to produce desirable results. Synthetic grafts can be coated with proteins, anticoagulants, and antibiotics to improve graft function. Similarly, synthetic polymers or biologically-derived structural proteins can be bonded to various growth factors to induce favorable host responses. Resorbable polymers comprise another type of graft material in which extensive in-vivo tissue ingrowth constitutes the biochemical and functional integrity of the vascular conduit. After more than twenty years of intensive research, a better understanding of the clinical value of endothelial cell (EC) seeding is beginning to emerge. Using tissue engineering techniques, a biologically viable vascular substitute is becoming feasible. In addition, genetic manipulation of vascular cells may enhance the potential function of these biologic grafts.^[2]

STRATEGIES FOR THE DEVELOPMENT OF VASCULAR GRAFTS

To improve the long-term patency of bioprosthetic grafts, the following approaches have been pursued:

- 1) modification of the blood contact surface to maximize the thromboresistance of synthetic grafts;
- 2) manipulation of in-vivo graft-host histologic responses by controlled delivery of bioactive substances;
- 3) development of bioresorbable vascular prostheses;
- 4) endothelialization of the graft blood contact surface; and
- 5) development of completely tissue-engineered vascular substitutes.

Surface Modifications

The simplest modification of a graft is to coat the inner surface with a relatively inert polymer. Carbon coating has been found to decrease surface thrombogenicity due to its negative charge and hydrophobic nature. Carbon-impregnated prosthetic grafts were found to reduce platelet deposition, demonstrate improved primary and secondary graft patency, and raise limb-salvage rates at 12 months in clinical comparative studies.^[3,4] Silicone polymer coating—which produces a smooth surface devoid of the usual ePTFE graft permeability and texture—followed by plasma glow discharge polymerization effectively abolished pannus tissue ingrowth and graft surface neointimal hyperplasia in baboon and canine models.^[5]

Protein coating was used to decrease the porosity of Dacron grafts as an alternative to blood preclotting. Degradation of the impregnated protein by the host immune system and other proteases results in tissue ingrowth.^[6] Albumin, gelatin, and collagen are most commonly used. Large-diameter, knitted Dacron prostheses coated with each of these three proteins are currently available for clinical use. Albumin coating, developed in the 1970s, was found to diminish platelet and leukocyte adhesion.^[7] In a canine thoracoabdominal aortic model, knitted Dacron grafts impregnated with albumin had less transinterstitial



blood loss at implantation, and significantly thinner inner capsules at 20 weeks.^[8]

Native collagen is intrinsically thrombogenic, whereas cross-linked collagen has shown an improved thrombogenic profile. Collagen coating establishes a good matrix for cellular ingrowth and induces neointimal formation, which theoretically should improve long-term graft patency. Promising initial results have been reported by several groups.^[9] However, a recent clinical study reported no significant improvement of graft patency with collagen impregnation of knitted Dacron aortoiliac grafts.^[6]

Gelatin is a derivative of collagen that degrades readily when utilized as a graft coating. Different coating techniques—including alkylation, plasma discharge, nonspecific cross-linking, covalent binding, and application of thin polymer films—have been developed to improve the gelatin degradation time.^[9-16]

Anticoagulants have been affixed to synthetic grafts. Heparin binding to Dacron grafts has been shown to reduce its thrombogenicity.^[17] In canine studies, heparin bound to polyurethane and polydimethylsiloxane surfaces improved graft patency rates.^[18] Thrombolytic agents such as urokinase have also been bound to bioprosthetic surfaces with promising results. Little is known about the long-term efficiency of these modifications.^[19]

Manipulation of In-Vivo Healing Process

Tissue incorporation occurs in all implanted bioprosthetic grafts. Excessive cell proliferation and extracellular matrix (ECM) deposition result in neointimal hyperplasia, leading to vascular graft failure. The ideal healing sequence would be rapid endothelialization of the blood contact surface with spatially and temporally limited subendothelial smooth muscle cell (SMC) growth, followed by phenotypic and functional maturation of cell components, as well as remodeling of the ECM. Recent developments in the study of mechanisms responsible for the migration and proliferation of endothelial cells (ECs) and SMCs provide clues to the efficacious manipulation of the healing process, thereby optimizing the microenvironments of the graft and perigraft tissue.

Fibroblast growth factors (FGFs), notably FGF-1 (acidic FGF) and FGF-2 (basic FGF), which are known to be potent mitogens, also possess potent angiogenic activities. Both factors have been experimentally applied to grafts alone and in combination with heparin. Heparin potentiates the mitogenic activity and protects FGFs from proteolytic degradation.

Achieving a controlled healing response requires a defined delivery system that can apply bioactive

substances to the graft surface and predictably release them with preserved bioactivity over an extended period of time. Greisler has evaluated the affixation of FGF-1 to synthetic graft surfaces.^[20,21] FGF-1 was applied to various synthetic grafts by sequential application of fibronectin, heparin, FGF-1, and a second heparin layer, utilizing known binding affinities between successively applied agents. Growth factor retention was 44% in Dacron grafts and 23% in polydioxanone (PDS) grafts. In addition, full growth factor bioactivity was observed. However, no significant increase in endothelialization was observed *in vivo*.^[22]

Utilizing a fibrin glue delivery system, FGF-1 and heparin-impregnated ePTFE grafts were evaluated in both canine aortoiliac and thoracoabdominal aortic models.^[23-26] A significant increase in endothelial cell proliferation was observed, along with rapid development of a confluent endothelial layer at 28 days. Extensive transinterstitial capillary ingrowth throughout the graft wall was also observed. In addition, a significant increase in subendothelial myofibroblast proliferation was observed at 1 month.^[25,26]

Coimmobilization of FGF-2 and heparin in a microporous polyurethane graft by cross-linked gelatin gel has been demonstrated to accelerate tissue regeneration. It was associated with a greater extent of endothelialization via perianastomotic and transmural capillary ingrowth in a rat aorta model.^[27] A consistent 40- μ m thick neointima with intermittent endothelialization, and an underlying layer of SMCs and fibroblasts, were observed in the midportion of the treated grafts, compared with control grafts that were covered with only a fibrin layer.^[27]

In response to exogenous stimuli, ECs, SMCs, and fibroblasts turn over and enter the cell cycle. Transformed cells divide, synthesize EMC, and produce growth factors with effects induced by autocrine and paracrine mechanisms. A number of growth factors and cell cycle proteins are involved in this process. Intimal hyperplasia is caused by excessive ECM production and cellular proliferation. Fine control of this proliferative process is the key to successful manipulation of these processes.

Bioresorbable Grafts

Currently clinically available synthetic vascular grafts—either ePTFE or Dacron—are permanent prostheses in the host after implantation. Tissue-graft reaction, ultimately resulting in graft failure at times, is inevitable as long as the graft persists. Theoretically, it may be possible to stimulate a rapid and controlled ingrowth of tissue, assuming the load-bearing material resists dilation and incorporates cellular components with

desirable physiologic characteristics, forming a neoartery. The synthetic material itself may no longer be necessary following tissue ingrowth, and may thus form the basis of another strategy in the design of vascular grafts—namely, the use of bioresorbable materials. Bioresorbable grafts' key advantage is that after tissue ingrowth has occurred, prosthetic material is no longer present to continue stimulating a foreign-body reaction or harbor bacteria leading to graft infection.

Wesolowski and Ruderman were the first to propose the concept of a slowly absorbable vascular graft that could induce a host-regenerative process producing a new functional artery.^[28,29] These early partially resorbable grafts were made of fibers composed of Dacron and catgut or Dacron and polylactide. The first report of a completely bioresorbable graft was in 1979 by Bowald, who described the use of a rolled sheet of Vicryl (a copolymer of polyglycolide and polylactide).^[30,31] Unfortunately, these early grafts were prone to aneurysmal dilation and rupture.

Woven polyglycolic acid (PGA) grafts were evaluated by Greisler in a rabbit model.^[32,33] Four weeks after implantation, the grafts contained an inner capsule composed of a confluent layer of endothelial cells and smooth muscle-like myofibroblasts amid dense collagen fibers. Similarly constructed Dacron grafts demonstrated an inner capsule containing only a fibrin coagulum with minimal cellularity. Macrophage infiltration and phagocytosis were in parallel with the resorption of the PGA, which was no longer present at 3 months. Initial experiments showed that aneurysmal dilation occurred in 10% of grafts, at the time of prosthetic material resorption and before ingrowth of tissue with adequate strength to resist hemodynamic pressures. Later studies evaluated grafts composed of PDS, a more slowly resorbed material. PDS grafts demonstrated a similar endothelialization of the regenerated luminal surface. Compared to PGA, which is completely resorbed within 3 months, PDS remains present for up to 6 months. Only 1 of 28 PDS grafts exhibited aneurysmal dilation with explant times as late as 1 year. The explanted specimens of these PDS grafts demonstrated biomechanical characteristics similar to native arteries, which were able to withstand static bursting pressures of 6000 mm Hg and mean pressures of 2000 mm Hg without fatigue.^[34]

As a requirement for clinical efficacy, a bioresorbable graft must regenerate a tissue complex of sufficient strength prior to loss of prosthetic integrity to minimize the development of aneurysmal dilation. There are several theoretical approaches to achieve this goal. One approach combines the bioresorbable material with a nonresorbable material that remains behind as a mechanical strut. Another approach involves the combination of two or more bioresorbable materials with different resorption rates such that the more

rapidly resorbed material evokes rapid tissue ingrowth whereas the second material provides temporary structural integrity to the graft. A third approach utilizes growth factors or cells that can be applied to the graft to enhance tissue ingrowth and organization.

Partially resorbable grafts have been studied by Greisler et al.^[35] Composite grafts constructed from fibers containing 69% resorbable PG910 and 31% nonresorbable polypropylene were evaluated in rabbit and canine models. Remarkably, these grafts demonstrated 100% patency with no aneurysmal dilation 1 year after implantation.^[36] In another study, completely resorbable composite grafts woven from fibers of 74% PG910 and 26% PDS also demonstrated 100% one-year patency with no aneurysms in a rabbit aorta model.^[37,38] In addition, the regenerated arteries withstood 800 mm Hg of pulsatile systolic pressure *ex vivo*. Furthermore, a confluent, functional, endothelial cell layer over circumferentially oriented smooth muscle-like myofibroblasts was present in the inner capsule of these grafts.

Wildevuur and van der Lei evaluated grafts prepared from a 95% polyurethane/5% polylactide mixture, and found that only relatively compliant grafts that induced circumferential smooth muscle development contained elastin and remained stable without dilating.^[39,40] They concluded that modifications of the graft preparation and smooth muscle cell (SMC) seeding may help enhance optimal orientation of the SMCs and may prevent aneurysm formation.^[40]

Further work and development may produce a small-diameter, completely resorbable vascular graft that may ultimately improve the currently dismal long-term patency rates of small-caliber bioprosthetic grafts.

Endothelial Cell Seeding

Endothelial cells (ECs) perform a series of important physiological functions, one of which is the maintenance of blood vessel thromboresistance. Endothelial cells possess a negative surface charge that helps prevent platelet adherence. Glycosaminoglycans—which bind antithrombin III, prostacyclins, and tissue plasminogen activator—are produced by the endothelial cell and contribute to its antithrombotic activities. From a theoretical standpoint, the presence of a confluent monolayer of endothelial cells on the luminal surface of a prosthetic graft should improve graft thromboresistance and help prevent the development of pseudointimal hyperplasia.

The first report of enhanced graft survival as a result of endothelial cell seeding came from Herring in 1978.^[41] Considerable progress has been made since that original report. Initial technical problems



associated with cell harvesting, seeding and adhesion, and prevention of desquamation have for the most part been overcome.^[42,43] The ideal seeded graft should have a confluent endothelial cell lining at the time of implantation. Endothelial cells should be able to resist shear stress after the restoration of blood flow, and retain all of their physiologic functions while present on the graft surface.

Adequate endothelial cell density is a requirement for successful graft seeding. The density of endothelial cells lining a normal artery is approximately 10^3 ECs/mm².^[44] At minimum, an attachment of at least 5×10^3 ECs/mm² is required for immediate confluent human EC coverage of a small-caliber vascular graft. In order to achieve high enough cell densities, a two-stage seeding procedure has been developed. Endothelial cells are harvested, allowed to proliferate *in vitro*, and then seeded and grown to confluence on the vascular graft prior to implantation.^[45-47] The major disadvantages of this technique are the potential for contamination and infection, alteration of the EC phenotype and function, the requirement of a 3-4-week waiting period for expansion of the cell population, and the need for two operative procedures. An alternative method utilizes microvascular endothelial cells in which small-diameter Dacron or ePTFE grafts are seeded with cells harvested from the omental microvasculature in a single stage. This technique resulted in confluent endothelial linings, a larger thrombus-free surface area, and improved patency rates at 1 year in a canine model.^[48]

Endothelial cell adhesion and retention is another key issue for prosthetic graft seeding. Endothelial cells adhere very poorly to synthetic graft materials. Therefore, adhesive proteins such as fibronectin, collagen, fibrin, and laminin have been applied to graft surfaces to improve seeding efficiency. Nitrogen-containing plasma treatment—a novel, recently developed surface modification technique—generates basic groups on the graft surface and enhances the ability of ECs to resist shear stress.^[44,49] Studies on the kinetics of EC seeding showed a 20% loss of seeded cells during the first hour, followed by continued loss up to 60% within the first 24 hours.^[50-52] Longer incubation times prior to exposure to blood flow allow maturation of the cytoskeleton and improve cell retention on the graft surface.^[53-55] Preconditioning the seeded EC monolayer with graded shear stress has been shown to promote the reorganization of the EC cytoskeleton and production of extracellular matrix, which in turn results in higher EC retention after exposure to blood flow.^[56]

Intrinsic properties of the bioprosthetic material also affect EC attachment. Dacron has a higher level of cell attachment compared to ePTFE when both types of grafts are coated with a similar matrix.^[54-57]

Polyurethane also showed better cell attachment than ePTFE.^[58] The negative surface charge present on ePTFE is an important characteristic that helps limit platelet deposition. On the other hand, the negative surface charge also prevents negatively-charged ECs from accessing the graft surface because of natural repulsive forces.

Numerous animal studies have demonstrated decreased platelet deposition and increased patency rates in EC-seeded Dacron and ePTFE grafts.^[59-64] Resistance to bacterial contamination has also been shown in seeded grafts.^[65] Endothelial cell seeding does not appear to reduce the formation of pseudointimal hyperplasia in a canine model, in which 4-mm Dacron grafts were seeded and implanted in the aortoiliac position.^[66] In fact, higher levels of PDGF and bFGF have been measured in endothelial seeded grafts, which is of particular concern given their potential role in the development of pseudointimal hyperplasia.^[67,68]

Clinical data evaluating EC-seeded synthetic grafts have been limited and have had mixed results. Magometschnigg reported a twofold increase in early graft patency and a twofold decrease in late amputation rates of femoral-tibial bypasses, using cultured endothelial cells seeded onto fibrin glue-coated ePTFE grafts.^[69] Herring reported increased 1-year patency rates for seeded ePTFE femoral-popliteal grafts.^[70] However, later studies using the same single-stage seeding technique failed to demonstrate any significant long-term advantage of seeded grafts.^[71] Utilizing the two-stage seeding technique, Zilla reported decreased platelet deposition and increased 5-year patency rates in seeded ePTFE femoral-popliteal bypass grafts over a 9-year period.^[72,73] In addition, intimal thickening and subsequent stenosis of seeded grafts have been reported.^[74,75]

Recent advances in molecular biology have allowed the seeding of genetically-manipulated endothelial cells to synthetic grafts. The first report of this approach was by Wilson et al. in 1989, in which Dacron grafts were seeded with retrovirally-transduced ECs containing the lacZ marker gene. These grafts were implanted into a canine carotid artery model, and lacZ gene expression was detected for a period of 5 weeks.^[76] Tissue plasminogen activator (tPA) has also been successfully transfected into endothelial cells, and expression of tPA activity has been detected in cells seeded onto synthetic grafts.^[77,78] Endothelial cells may be manipulated to secrete specific growth factors and inhibitors to allow optimal EC proliferation and migration and to prevent smooth muscle cell hyperplasia. Varying results have been reported regarding the proliferation, adhesion, and retention of genetically-modified ECs on the surface of synthetic grafts.^[70,79,80] Genetically-modified ECs also appear to have decreased retention on graft surfaces *in vivo*.^[81-85]

Because of these difficulties, little data have thus far been reported showing an in-vivo advantage of genetically-modified EC-seeded grafts.

The introduction of genetic engineering offers considerable potential to manipulate the function and physiology of seeded cells on graft as well as perigraft cells. These cells may in turn further augment the efficiency of EC seeding.

Tissue Engineering

It has been suggested that endothelial cells alone cannot produce a stable intima without the presence of underlying SMCs or fibroblasts.^[86] Tissue fragments containing multiple cell types—including venous tissue, adipose tissue, and bone marrow—have been seeded onto grafts and have been found to accelerate the graft-healing process.^[86] Interestingly, bone marrow cell seeding was reported to induce abundant capillary growth in the graft wall and a rapid, complete endothelialization of the inner surface without intimal hyperplasia. Multipotent stem cells in the bone marrow have the potential to differentiate in response to their microenvironment. Stem cells proliferate and secrete cytokines necessary for their survival and may represent a significant cell source for tissue engineering.

In the quest to develop the ideal vascular substitute, several attempts to construct a neoartery have been reported. Vascular cells were seeded onto a three-dimensional ECM or polymer scaffold. After implantation, the cells proliferated and produced ECM, whereas the scaffold degraded and was replaced by native tissue. The newly formed conduits were viable vessels with the ability to remodel in response to the hemodynamic environment and to maintain the cell components' normal functions.^[87] Shinoka et al. produced a graft from autologous mixed vascular cells from carotid artery explants.^[88] The cells were expanded in vitro and then seeded onto polyglactin or PGA scaffolds. After 7 days in culture, the grafts were implanted as pulmonary artery interpositions in lambs. The scaffolds were no longer present at 11 weeks, and the newly generated conduit resembled native artery architecture with ECs lining the luminal surface. In addition, the development of ECM including collagen and elastic fibers was also observed.^[88]

The first attempt at creating a totally biologic vascular conduit in vitro without a synthetic polymer was made by Weinberg and Bell.^[89] They constructed a three-layered blood vessel model with collagen matrix as a scaffold for ECs, SMCs, and fibroblasts. Similar efforts have been made by other groups.^[90–96] These early models failed to achieve sufficient mechanical strength, requiring reinforcement with Dacron mesh. This weakness is presumed to be due, at least in part, to

the lack of organization of the ECM. L'Heureux et al. generated a vessel exclusively from cultured vascular cells that produced ECM with an organization resembling that present in natural vessels.^[91] SMCs were cultured in vitro to form a cellular sheet, which was then wrapped around a tubular support to produce the media. Subsequently, a similar sheet of fibroblasts was placed around the media to compose the adventitia. After maturation, the tubular support was removed and ECs were seeded onto the luminal surface. This constructed vessel displayed a bursting strength of over 2000 mm Hg. The SMCs produced desmin, expressed circumferential and longitudinal orientations, and even demonstrated contractile responses when challenged with vasoactive agonists. Abundant ultrastructurally organized collagen and elastin fibers were present in the ECM. However, in short-term studies, after being implanted in canine femoral arteries, intramural blood infiltration was observed. Although much work will be required to conquer the technical obstacles and optimize the manufacturing system, these efforts will create possibilities to produce viable substitutes for vascular replacement.



CONCLUSION

Much progress in the development of vascular substitutes has been seen over the past several decades. Investigations into the underlying biochemical and cellular mechanisms involved in the graft-healing process have yielded a firm understanding of graft physiology and the pathological processes involved in graft failure. Utilization of numerous biochemical, molecular biological, and tissue-engineering techniques—with the goal of improving long-term graft patency—will result in the development of novel vascular substitutes and allow continuing increase in our understanding of the graft-healing process.

ARTICLES OF FURTHER INTEREST

Polytetrafluoroethylene; Tissue Engineering of Blood Vessel; Vascular Grafts; Vascular Grafts: Host Interactions; Vascularization of Engineered Constructs

REFERENCES

1. Voorhees, A.B., Jr.; Jaretzki, A.; Blakemore, A.H. The use of tubes constructed of Vinyon N cloth in bridging arterial defects. *Ann. Surg.* **1952**, *135*, 332–336.
2. Conte, M.S. The ideal small arterial substitute, a search for the Holy Grail? *FASEB J.* **1998**, *12*, 43–45.
3. Tsuchida, H.; Cameron, B.L.; Marcus, C.S.; Wilson, S.E. Modified polytetrafluoroethylene: Indium 111-labeled

- platelet deposition on carbon-lined and high porosity polytetrafluoroethylene grafts. *J. Vasc. Surg.* **1992**, *16*, 643–649.
4. Groegler, F.M.; Kapfer, X.; Meichelböck, W. Crural Prosthetic Revascularization: Randomized, Prospective, Multicentric Comparison of Standard and Carbon Impregnated ePTFE Grafts. In *Montefiore Annual Symposium on Current Critical Problems, New Horizons and Techniques in Vascular and Endovascular Surgery ('01)*; New York, NY, Nov 15–18, 2001. Abstract 01–2, I2.1–I2.3.
 5. Lumsden, A.B.; Chen, C.; Coyle, K.A.; Ofenloch, J.C.; Wang, J.H.; Yasuda, H.K.; Hanson, S.R. Nonporous silicone polymer coating of expanded polytetrafluoroethylene grafts reduces graft neointimal hyperplasia in dog and baboon models. *J. Vasc. Surg.* **1996**, *24*, 825–833.
 6. Quarmby, J.W.; Burnand, K.G.; Lockhart, S.J.M.; Donald, A.E.; Sommerville, K.M.; Jamieson, C.W.; Browse, N.L. Prospective randomized trial of woven versus collagen-impregnated knitted prosthetic Dacron grafts in aortoiliac surgery. *Br. J. Surg.* **1998**, *85*, 775–777.
 7. Phaneuf, M.D.; Quist, W.C.; Bide, M.J.; LoGerfo, F.W. Modification of polyethylene terephthalate (Dacron) via denier reduction, effects on material tensile strength, weight, and protein binding capabilities. *J. Appl. Biomater.* **1995**, *6*, 289–299.
 8. Kang, S.S.; Petsikas, D.; Murchan, P.; Cziperle, D.J.; Ren, D.; Kim, D.U.; Greisler, H.P. Effects of albumin coating of knitted Dacron grafts on transinterstitial blood loss and tissue ingrowth and incorporation. *Cardiovasc. Surg.* **1997**, *5*, 184–189.
 9. Bos, G.W.; Poot, A.A.; Beugeling, T.; van Aken, W.G.; Feijen, J. Small-diameter vascular graft prostheses, current status. *Arch. Physiol. Biochem.* **1998**, *106*, 100–115.
 10. Eberhart, R.C.; Munro, M.S.; Williams, G.B.; Kulkarni, P.V.; Shannon, W.A., Jr.; Brink, B.E.; Fry, W.J. Albumin adsorption and retention on C18-alkyl-derivatized polyurethane vascular grafts. *Artif. Organs* **1987**, *11*, 375–382.
 11. Ishikawa, Y.; Sasakawa, S.; Takase, M.; Osada, Y. Effect of albumin immobilization by plasma polymerization on platelet reactivity. *Thromb. Res.* **1984**, *35*, 193–202.
 12. Tsai, C.C.; Huo, H.H.; Kulkarni, P.V.; Eberhart, R.C. Biocompatible coatings with high albumin affinity. *Trans. Am. Soc. Artif. Intern. Organs* **1990**, *36*, M307–M310.
 13. Rumisek, J.D.; Wade, C.E.; Brooks, D.E.; Okerberg, C.V.; Barry, M.J.; Clarke, J.S. Heat-denatured albumin-coated Dacron vascular grafts: Physical characteristics and in vivo performance. *J. Vasc. Surg.* **1986**, *4*, 136–143.
 14. Kottke-Marchant, K.; Anderson, J.M.; Umemura, Y.; Marchant, R.E. Effect of albumin coating of the in vitro blood compatibility of Dacron arterial prostheses. *Biomaterials* **1989**, *10*, 147–155.
 15. Phaneuf, M.D.; Dempsey, D.J.; Bide, M.J.; Szycher, M.; Quist, W.C.; LoGerfo, F.W. Bioengineering of a novel small diameter polyurethane vascular graft with covalently bound recombinant hirudin. *ASAIO J.* **1998**, *44*, M653–M658.
 16. Dempsey, D.J.; Phaneuf, M.D.; Bide, M.J.; Szycher, M.; Quist, W.C.; Logerfo, F.W. Synthesis of a novel small diameter polyurethane vascular graft with reactive binding sites. *ASAIO J.* **1998**, *44*, M506–M510.
 17. Mohamed, M.S.; Mukherjee, M.; Kakkar, V.V. Thrombogenicity of heparin and non-heparin bound arterial prostheses, an in vitro evaluation. *J. R. Coll. Surg. Edinb.* **1998**, *43*, 155–157.
 18. Park, K.D.; Okano, T.; Nojiri, C.; Kim, S.W. Heparin immobilization onto segmented polyurethane-urea surfaces—Effect of hydrophilic spacers. *J. Biomed. Mater. Res.* **1988**, *22*, 977–992.
 19. Nojiri, C.; Kido, T.; Sugiyama, T.; Horiuchi, K.; Kijima, T.; Hagiwara, K.; Kuribayashi, E.; Nogawa, A.; Ogiwara, K.; Akutsu, T. Can heparin immobilized surfaces maintain nonthrombogenic activity during in vivo long-term implantation? *ASAIO J.* **1996**, *42*, M468–M475.
 20. Greisler, H.P.; Klosak, J.J.; Dennis, J.W.; Ellinger, J.; Kim, D.U.; Burgess, W.; Maciag, T. Endothelial cell growth factor attachment to biomaterials. *ASAIO Trans.* **1986**, *32*, 346–349.
 21. Greisler, H.P.; Klosak, J.J.; Dennis, J.W.; Karesh, S.M.; Ellinger, J.; Kim, D.U. Biomaterial pretreatment with ECGF to augment endothelial cell proliferation. *J. Vasc. Surg.* **1987**, *5*, 393–399.
 22. Greisler, H.P.; Gosselin, C.; Ren, D.; Kang, S.S.; Kim, D.U. Biointeractive polymers and tissue engineered blood vessels. *Biomaterials* **1996**, *17*, 329–336.
 23. Greisler, H.P.; Cziperle, D.J.; Kim, D.U.; Garfield, J.D.; Petsikas, D.; Murchan, P.M.; Applegren, E.O.; Drohan, W.; Burgess, W.H. Enhanced endothelialization of expanded polytetrafluoroethylene grafts by fibroblast growth factor type 1 pretreatment. *Surgery* **1992**, *112*, 244–254.
 24. Gosselin, C.; Ren, D.; Ellinger, J.; Greisler, H.P. In vivo platelet deposition on polytetra-fluoroethylene coated with fibrin glue containing fibroblast growth factor 1 and heparin in a canine model. *Am. J. Surg.* **1995**, *170*, 126–130.
 25. Zarge, J.I.; Gosselin, C.; Huang, P.; Greisler, H.P. Platelet deposition on ePTFE Grafts coated with fibrin glue with or without FGF-1 and heparin. *J. Surg. Res.* **1997**, *67*, 4–8.
 26. Gray, J.L.; Kang, S.S.; Zenni, G.C.; Kim, D.U.; Kim, P.I.; Burgess, W.H.; Drohan, W.; Winkles, J.A.; Hauderschild, C.C.; Greisler, H.P. FGF-1 affixation stimulates ePTFE endothelialization without intimal hyperplasia. *J. Surg. Res.* **1994**, *57*, 596–612.
 27. Doi, K.; Matsuda, T. Enhanced vascularization in a microporous polyurethane graft impregnated with basic fibroblast growth factor and heparin. *J. Biomed. Mater. Res.* **1997**, *34*, 361–370.
 28. Wesolowski, S.A.; Fries, C.C.; Domingo, R.T.; Liebig, W.J.; Sawyer, P.N. The compound prosthetic vascular graft. A pathologic survey. *Surgery* **1963**, *53*, 19–44.
 29. Ruderman, R.J.; Hegyeli, A.F.; Hattler, B.G.; Leonard, F. A partially biodegradable vascular prosthesis. *ASAIO Trans.* **1972**, *18*, 30–37.

30. Bowald, S.; Busch, C.; Eriksson, I. Arterial regeneration following polyglactin 910 suture mesh grafting. *Surgery* **1972**, *86*, 722–729.
31. Bowald, S.; Busch, C.; Eriksson, I. Absorbable material in vascular prostheses: A new device. *Acta Chir. Scand.* **1980**, *146*, 391–395.
32. Greisler, H.P. Arterial regeneration over absorbable prostheses. *Arch. Surg.* **1982**, *117*, 1425–1431.
33. Greisler, H.P.; Kim, D.U.; Price, J.B.; Voorhees, A.B. Arterial regenerative activity after prosthetic implantation. *Arch. Surg.* **1985**, *120*, 315–323.
34. Greisler, H.P.; Ellinger, J.; Schwarcz, T.H.; Golan, J.; Raymond, R.M.; Kim, D.U. Arterial regeneration over polydioxanone prostheses in the rabbit. *Arch. Surg.* **1987**, *122*, 715–721.
35. Greisler, H.P.; Schwarcz, T.H.; Ellinger, J.; Kim, D.U. Dacron inhibition of arterial regenerative activity. *J. Vasc. Surg.* **1986**, *3*, 747–756.
36. Greisler, H.P.; Kim, D.U.; Dennis, J.W.; Klosak, J.J.; Widerborg, K.A.; Endean, E.D.; Raymond, R.M.; Ellinger, J. Compound polyglactin 910/polypropylene small vessel prostheses. *J. Vasc. Surg.* **1987**, *5*, 572–583.
37. Greisler, H.P.; Endean, E.D.; Klosak, J.J.; Ellinger, J.; Dennis, J.W.; Buttle, K.; Kim, D.U. Polyglactin 910/polydioxanone bicomponent totally resorbable vascular prostheses. *J. Vasc. Surg.* **1988**, *7*, 697–705.
38. Galletti, P.M.; Aebischer, P.; Sasken, H.F.; Goddard, M.B.; Chiu, T.H. Experience with fully bioresorbable aortic grafts in the dog. *Surgery* **1988**, *103*, 231–241.
39. Yue, X.; van der Lei, B.; Schakenraad, J.M.; van Oene, G.H.; Kuit, J.H.; Feijen, J.; Wildevuur, C.R. Smooth muscle cell seeding in biodegradable grafts in rats: A new method to enhance the process of arterial wall regeneration. *Surgery* **1988**, *103*, 206–212.
40. van der Lei, B.; Nieuwenhuis, P.; Molenaar, I.; Wildevuur, C.R. Long-term biologic fate of neoarteries regenerated in microporous, compliant, biodegradable, small-caliber vascular grafts in rats. *Surgery* **1987**, *101*, 459–467.
41. Herring, M.; Gardner, A.; Glover, J. A single-staged technique for seeding vascular grafts with autogenous endothelium. *Surgery* **1978**, *84*, 498–504.
42. Graham, L.M.; Burkell, W.E.; Ford, J.W.; Vinter, D.W.; Kahn, R.H.; Stanley, J.C. Expanded polytetrafluoroethylene vascular prostheses seeded with enzymatically derived and cultured canine endothelial cells. *Surgery* **1982**, *91*, 550–559.
43. Jarrell, B.E.; Williams, S.K. Microvessel derived endothelial cell isolation, adherence, and monolayer formation for vascular grafts. *J. Vasc. Surg.* **1991**, *13*, 733–734.
44. Sipehia, R.; Martucci, G.; Lipscombe, J. Transplantation of human endothelial cell monolayer on artificial vascular prosthesis: The effect of growth-support surface chemistry, cell seeding density, ECM protein coating, and growth factors. *Artif. Cells Blood Substit. Immobil. Biotechnol.* **1996**, *24*, 51–63.
45. Shindo, S.; Takagi, A.; Whittemore, A.D. Improved patency of collagen-impregnated grafts after in vitro autogenous endothelial cell seeding. *J. Vasc. Surg.* **1987**, *6*, 325–332.
46. Prendiville, E.J.; Coleman, J.E.; Callow, A.D.; Gould, K.E.; Laliberte-Verdon, S.; Ramberg, K.; Connolly, R.J. Increased in vitro incubation time of endothelial cells on fibronectin-treated ePTFE increases cell retention in blood flow. *Eur. J. Vasc. Surg.* **1991**, *5*, 311–319.
47. Sentissi, J.M.; Rambner, K.; O'Donnell, T.F., Jr.; Connolly, R.J.; Callow, A.D. The effect of flow on vascular endothelial cells grown in tissue culture on polytetrafluoroethylene grafts. *Surgery* **1986**, *99*, 337–343.
48. Pasic, M.; Muller-Glauser, W.; von Segesser, L.K.; Lachat, M.; Mihaljevic, T.; Turina, M. Superior late patency of small-diameter Dacron grafts seeded with omental microvascular cells, an experimental study. *Ann. Thorac. Surg.* **1994**, *58*, 677–683.
49. Tseng, D.Y.; Edelman, E.R. Effects of amide and amine plasma-treated ePTFE vascular grafts on endothelial cell lining in an artificial circulatory system. *J. Biomed. Mater. Res.* **1998**, *42*, 188–198.
50. Schneider, P.A.; Hanson, S.R.; Price, T.M.; Harker, L.A. Preformed confluent endothelial cell monolayers prevent early platelet deposition on vascular prostheses in baboons. *J. Vasc. Surg.* **1988**, *8*, 229–235.
51. Rosenman, J.E.; Kempczinski, R.F.; Pearce, W.H.; Silberstein, E.B. Kinetics of endothelial cell seeding. *J. Vasc. Surg.* **1985**, *2*, 778–784.
52. Vohra, R.; Thomson, G.J.; Carr, H.M.; Sharma, H.; Welch, M.; Walker, M.G. In vitro adherence and kinetics studies of adult human endothelial cell seeded polytetrafluoroethylene and gelatin impregnated Dacron grafts. *Eur. J. Vasc. Surg.* **1991**, *5*, 93–103.
53. Miyata, T.; Conte, M.S.; Trudell, L.A.; Mason, D.; Wittemore, A.D.; Brinyi, L.K. Delayed exposure to pulsatile shear stress improves retention of human saphenous vein endothelial cells on seeded ePTFE grafts. *J. Surg. Res.* **1991**, *50*, 485–493.
54. Sugawara, Y.; Miyata, T.; Sato, O.; Kimura, H.; Namba, T.; Makuuchi, M. Rapid post-incubation endothelial retention by Dacron grafts. *J. Surg. Res.* **1997**, *67*, 132–136.
55. Prendiville, E.J.; Coleman, J.E.; Callow, A.D.; Gould, K.E.; Laliberte-Verdon, S.; Ramberg, K.; Connolly, R.J. Increased in vitro incubation time of endothelial cells on fibronectin-treated ePTFE increases cell retention in blood flow. *Eur. J. Vasc. Surg.* **1991**, *5*, 311–319.
56. Ott, M.J.; Ballermann, B.J. Shear stress-conditioned, endothelial cell-seeded vascular grafts, improved cell adherence in response to in vitro shear stress. *Surgery* **1995**, *117*, 334–339.
57. Vohra, R.K.; Thomson, G.J.; Sharma, H.; Carr, H.M.; Walker, M.G. Effects of shear stress on endothelial cell monolayer on expanded polytetrafluoroethylene grafts using preclot and fibronectin matrices. *Eur. J. Vasc. Surg.* **1990**, *4*, 33–41.
58. Giudiceandrea, A.; Seifalian, A.M.; Krijgsman, B.; Hamilton, G. Effect of prolonged pulsatile shear stress in vitro on endothelial cell seeded PTFE and compliant polyurethane vascular grafts. *J. Vasc. Endovasc. Surg.* **1998**, *15*, 147–154.



59. Bowlin, G.L.; Rittgers, S.E.; Milsted, A.; Schmidt, S.P. In vitro evaluation of electrostatic endothelial cell transplantation onto 4 mm interior diameter expanded polytetrafluoroethylene grafts. *J. Vasc. Surg.* **1998**, *27*, 504–511.
60. Zilla, P.; Preiss, P.; Groscurth, P.; R_emeier, F.; Deutsch, M.; Odell, J.; Heidinger, C.; Fasol, R.; von Oppell, U. In vitro-lined endothelium, initial integrity and ultrastructural events. *Surgery* **1994**, *116*, 524–534.
61. Whitehouse, W.M., Jr.; Wakefield, T.W.; Vinter, D.W.; Ford, J.W.; Swanson, D.P.; Thrall, J.H.; Froelich, J.W.; Brown, L.E.; Burkel, W.E.; Graham, L.M. Indium-111 oxide labeled platelet imaging of endothelial seeded Dacron thoracoabdominal vascular prostheses in a canine experimental model. *ASAIO Trans.* **1983**, *29*, 183–187.
62. Wakefield, T.W.; Lindblad, B.; Graham, L.M.; Whitehouse, W.M., Jr.; Ripley, S.D.; Petry, N.A.; Spaulding, S.A.; Burkel, W.E.; Stanley, J.C. Nuclide imaging of vascular graft-platelet interactions, comparison of indium excess and technetium subtraction techniques. *J. Surg. Res.* **1986**, *40*, 388–394.
63. Stanley, J.C.; Burkel, W.E.; Ford, J.W.; Vinter, D.W.; Kahn, R.H.; Whitehouse, W.M., Jr.; Graham, L.M. Enhanced patency of small-diameter, externally supported Dacron iliofemoral grafts seeded with endothelial cells. *Surgery* **1982**, *92*, 994–1005.
64. Allen, B.T.; Long, J.A.; Clark, R.E.; Sicard, G.A.; Clark, R.D.; Hopkins, K.T.; Welch, M.J. Influence of endothelial cell seeding on platelet deposition and patency in small-diameter Dacron arterial grafts. *J. Vasc. Surg.* **1984**, *1*, 224–233.
65. Graham, L.M.; Stanley, J.C.; Burkel, W.E. Improved patency of endothelial-cell-seeded, long, knitted Dacron and ePTFE vascular prostheses. *ASAIO J.* **1985**, *8*, 65–73.
66. Rosenman, J.E.; Kempczinski, R.F.; Berlatzky, Y.; Pearce, W.H.; Ramalanjaona, G.R.; Bjornson, H.S. Bacterial adherence to endothelial-seeded polytetrafluoroethylene grafts. *Surgery* **1985**, *98*, 816–823.
67. Graham, L.M.; Fox, P.L. Growth factor production following prosthetic graft implantation. *J. Vasc. Surg.* **1991**, *13*, 742–746.
68. Sapienza, P.; di Marzo, L.; Cucina, A.; Corvino, V.; Mingoli, A.; Giustiniani, Q.; Ziparo, E.; Cavallaro, A. Release of PDGF-BB and bFGF by human endothelial cells seeded on expanded polytetrafluoroethylene vascular grafts. *J. Surg. Res.* **1998**, *75*, 24–29.
69. Magometschnigg, H.; Kadletz, M.; Vodrazka, M.; Dock, W.; Grimm, M.; Grabenwoger, M.; Minar, E.; Staudacher, M.; Fenzl, G.; Wolner, E. Prospective clinical study with in vitro endothelial cell lining of expanded polytetrafluoroethylene grafts in crural repeat reconstruction. *J. Vasc. Surg.* **1992**, *15*, 527–535.
70. Herring, M.B.; Gardner, A.; Glover, J. Seeding human arterial prostheses with mechanically derived endothelium. The detrimental effect of smoking. *J. Vasc. Surg.* **1984**, *1*, 279–289.
71. Jensen, N.; Lindblad, B.; Bergovist, D. Endothelial cell seeded Dacron aortobifurcated grafts, Platelet deposition and long-term follow-up. *J. Cardiovasc. Surg.* **1994**, *35*, 425–429.
72. Zilla, P.; Deutsch, M.; Meinhart, J.; Puschmann, R.; Eberl, T.; Minar, E.; Dudczak, R.; Lugmaier, H.; Schmidt, P.; Noszian, I. Clinical in vitro endothelialization of femoropopliteal bypass grafts, an actuarial follow-up over three years. *J. Vasc. Surg.* **1994**, *19*, 540–548.
73. Deutsch, M.; Meinhart, J.; Fischlein, T.; Preiss, P.; Zilla, P. Clinical autologous in vitro endothelialization of infrainguinal ePTFE grafts in 100 patients: a 9-year experience. *Surgery* **1999**, *126*, 847–855.
74. Deutsh, M.; Meinhart, J.; Vesely, M.; Fischlein, T.; Groscurth, P.; von Oppell, U.; Zella, P. In vitro endothelialization of expanded polytetrafluoroethylene grafts, a clinical case report after 41 months of implantation. *J. Vasc. Surg.* **1997**, *25*, 757–763.
75. Park, P.K.; Jarrell, B.E.; Williams, S.K.; Carter, T.L.; Rose, D.G.; Martinez-Hernandez, A.; Carabasi, R.A., III. Thrombus free, human endothelial surface in midregion of a Dacron vascular graft in the splanchnic venous circuit—Observations after nine months of implantation. *J. Vasc. Surg.* **1990**, *11*, 468–475.
76. Wilson, J.M.; Birinyi, L.K.; Salomon, R.N.; Libby, P.; Callow, A.D.; Mulligan, R.C. Implantation of vascular grafts lined with genetically modified endothelial cells. *Science* **1989**, *244*, 1344–1346.
77. Dichek, D.A.; Nussbaum, O.; Degen, S.J.; Anderson, W.F. Enhancement of the fibrinolytic activity of sheep endothelial cells by retroviral-mediated gene transfer. *Blood* **1991**, *77*, 533–541.
78. Shayani, V.; Newman, K.D.; Dichek, D.A. Optimization of recombinant t-PA secretion from seeded vascular grafts. *J. Surg. Res.* **1994**, *57*, 495–504.
79. Brothers, T.E.; Judge, L.M.; Wilson, J.M.; Burkel, W.E.; Stanley, J.C. Effect of genetic transduction on in vitro canine endothelial cell prostanoid production and growth. *Surg. Forum* **1990**, *41*, 337–339.
80. Jaklitsh, M.T.; Biro, S.; Casscells, W.; Dichek, D.A. Transduced endothelial cells expressing high levels of tissue plasminogen activator have an unaltered phenotype in vitro. *J. Cell. Physiol.* **1993**, *154*, 207–216.
81. Huber, T.S.; Welling, T.H.; Sarkar, R.; Messina, L.M.; Stanley, J.C. Effects of retroviral-mediated tissue plasminogen activator gene transfer and expression on adherence and proliferation of canine endothelial cells seeded onto expanded polytetrafluoroethylene. *J. Vasc. Surg.* **1995**, *22*, 795–803.
82. Podrazik, R.M.; Whitehill, T.A.; Komorowski, T.A.; Karo, K.H.; Messina, L.M.; Stanley, J.C. In vivo fate of lacZ-transduced endothelial cells seeded ePTFE thoracoabdominal vascular prostheses in dogs. *Surg. Forum* **1993**, *44*, 334–337.
83. Baer, R.P.; Whitehill, T.E.; Sarkar, R.; Sarkar, M.; Messina, L.M.; Komorowski, T.A.; Stanley, J.C. Retroviral-mediated transduction of endothelial cells with the lacZ gene impairs cellular proliferation in vitro and graft endothelialization in vivo. *J. Vasc. Surg.* **1996**, *24*, 892–899.

84. Dunn, P.F.; Newmann, K.D.; Jones, M.; Yamada, I.; Shayani, V.; Virmani, R.; Dichek, D.A. Seeding of vascular grafts with genetically modified endothelial cells, secretion of recombinant tPA results in decreased seeded cell retention in vitro and in vivo. *Circulation* **1996**, *93*, 1439–1446.
85. Falk, J.; Townsend, L.E.; Vogel, L.M.; Boyer, M.; Olt, S.; Wease, G.L.; Trevor, K.T.; Seymour, M.; Glover, J.L.; Bendick, P.J. Improved adhere of genetically modified endothelial cells to small-diameter expanded polytetrafluoroethylene grafts in a canine model. *J. Vasc. Surg.* **1998**, *27*, 902–909.
86. Noishiki, Y.; Yamane, Y.; Okoshi, T.; Tomizawa, Y.; Satoh, S. Choice, isolation, and preparation of cells for bioartificial vascular grafts. *Artif. Organs* **1998**, *22*, 50–62.
87. Yue, X.; Van der Lei, B.; Schakenraad, J.M.; Van Oene, G.H.; Kuit, J.H.; Feijen, J.; Wildevuur, C.R. Smooth muscle cell seeding in biodegradable grafts in rats, a new method to enhance the process of artery wall regeneration. *Surgery* **1998**, *103*, 206–212.
88. Shinoka, T.; Shum-Tim, D.; Ma, P.X.; Tanel, R.E.; Isogai, N.; Langer, R.; Vacanti, J.P.; Mayer, J.E. Creation of viable pulmonary artery autografts through tissue engineering. *J. Thorac. Cardiovasc. Surg.* **1998**, *115*, 536–546.
89. Weinberg, C.B.; Bell, E. A blood vessel model constructed from collagen and cultured vascular cells. *Science* **1986**, *231*, 397–400.
90. L'Heurx, N.; Germain, L.; LabbJ, T.; Auger, F.A. In vitro construction of a human blood vessel from cultured vascular cells. A morphologic study. *J. Vasc. Surg.* **1993**, *17*, 499–509.
91. L'Heurx, N.; P>quet, S.; LabbJ, T.; Germain, L.; Auger, F.A. A completely biological tissue-engineered human blood vessel. *FASEB J.* **1998**, *12*, 47–56.
92. Ziegler, T.; Nerem, R.M. Tissue engineering a blood vessel, regulation of vascular biology by mechanical stresses. *J. Cell. Biochem.* **1994**, *56*, 204–209.
93. Langer, R.; Vacanti, J.P. Tissue engineering. *Science* **1993**, *260*, 920–926.
94. Fuchs, J.R.; Nasser, B.A.; Vacanti, J.P. Tissue engineering: A 21st century solution to surgical reconstruction. *Ann. Thorac. Surg.* **2001**, *72*, 577–591.
95. Moldovan, N.I.; Ferrari, M. Prospects for microtechnology and nanotechnology in bioengineering of replacement microvessels. *Arch. Pathol. Lab. Med.* **2002**, *126*, 320–324.
96. Griffith, L.G.; Naughton, G. Tissue engineering—Current challenges and expanding opportunities. *Science* **2002**, *295*, 1009–1016.



Vascular Grafts: Host Interactions

James Laredo

Lian Xue

Division of Peripheral Vascular Surgery, Department of Surgery, Loyola University Medical Center, Maywood, and Edward Hines Jr. Veterans Administration Hospital, Hines, Illinois, U.S.A.

Howard P. Greisler

Division of Peripheral Vascular Surgery, Department of Surgery, Loyola University Medical Center, Department of Cell Biology, Neurobiology and Anatomy, Loyola University Medical Center, Maywood, and Edward Hines Jr. Veterans Administration Hospital, Hines, Illinois, U.S.A.

INTRODUCTION

Approximately 550,000 vascular bypass procedures are performed annually in the United States. Autogenous vein remains the conduit of choice for infrainguinal arterial reconstruction and is the most commonly used conduit for coronary artery bypass surgery. However, in situations where autogenous vein is unsuitable or unavailable, artificial vascular grafts are utilized. Long-term patency rates of artificial vascular grafts are inferior to those observed with autogenous vein grafts; therefore, the development of vascular substitutes remains a continuing priority.

Initial graft research was directed at developing inert materials that passively transported blood and minimally interacted with blood and tissue. Vinyon-N was the first successful artificial vascular graft developed in 1952 by Voorhees.^[1] Since then, polyethylene terephthalate (Dacron) and expanded polytetrafluoroethylene (ePTFE) have become the most widely used prosthetic vascular grafts. However, the development of completely nonreactive substances has been unsuccessful.^[2] Both Dacron and ePTFE react with blood components and perigraft tissues in a manner that is both beneficial and detrimental.

The success of a vascular graft depends upon the intrinsic properties of the graft material and the hemodynamic environment in which the graft is placed. Patient variables, which include clinical indication, outflow resistance, underlying disease processes and other comorbidities, also contribute to the long-term performance of a synthetic vascular graft. A firm understanding of graft physiology and its pathological processes is critical to the development of efficacious vascular substitutes.

GRAFT-HOST INTERACTIONS

All grafts, regardless of their composition and structure, evoke complex but predictable host responses beginning at the blood-biomaterial interface. Protein adsorption begins immediately after restoration of blood flow, followed by cellular and humoral responses to synthetic materials. Host responses include plasma protein deposition, platelet activation, infiltration of neutrophils and monocytes, and migration of endothelial and smooth muscle cells. Even the most inert artificial graft materials thus far developed are still recognized as foreign. The tissue-graft-blood interfaces are highly complex microenvironments that are ultimately responsible for graft patency.

Protein Adsorption

The most abundant serum proteins, albumin, fibrinogen, and immunoglobulin G (IgG), along with other serum proteins, adsorb to the graft almost instantaneously following exposure to the systemic circulation. Redistribution of serum proteins according to the Vroman effect occurs next, where each protein's local concentration and relative biochemical and electrical affinity for the graft surface determines its distribution.^[3] Because platelets and blood cells interact predominantly with bound proteins and not the prosthetic material itself, the constitution and concentration of bound protein largely determine the interaction of blood components with the graft, and therefore affect graft survival.

Fibrinogen, laminin, fibronectin, and vitronectin all have an RGD sequence (arg-gly-asp) region that binds to the GP IIb/IIIa receptor complex on platelet membranes. Thus, platelet deposition and activation is

influenced by prior protein adsorption. The binding of the RGD sequence to β_2 integrins expressed on the cell surface also contributes to graft–leukocyte adhesion. Neutrophils bound to immobilized IgG and fibrinogen demonstrate a reduction in their bactericidal properties, which may in part be responsible for the susceptibility of artificial grafts to infections.^[4] Additional plasma proteins, including complement components, are activated by synthetic surfaces to varying degrees.^[5] The rapid accumulation of coagulation proteins such as thrombin and Factor Xa on the luminal surface after implantation contributes to graft-associated procoagulant activity.^[6]

Platelet Deposition

Early platelet deposition occurs primarily via receptor-mediated interactions with adsorbed proteins and, to a lesser extent, by direct adherence to the graft. The platelet/protein complex is mediated through von Willebrand's factor and platelet membrane glycoproteins. After adherence, platelets undergo conformational changes and degranulate, releasing variety of bioactive substances that include serotonin, epinephrine, adenosine diphosphate, and thromboxane A₂. These substances further activate additional platelets, resulting in increased thrombin generation. Growth factors platelet-derived growth factor (PDGF), epidermal growth factor (EGF), and transforming growth factor β (TGF- β) are also released by activated platelets. These factors modulate smooth muscle cell (SMC) migration and proliferation and extracellular matrix synthesis. In addition, monocyte chemoattractants platelet factor 4 and β -thromboglobulin, which mediate the recruitment of monocytes into the graft, are also released. Platelet deposition and activation continues long after graft implantation. Animal experiments have demonstrated increased levels of thromboxane and decreased systemic platelet counts one year after Dacron graft implantation.^[7] Human studies have also revealed increased platelet adhesion to grafts long after implantation.^[8–10]

The thrombogenic nature of the synthetic graft surface is a major determinant of early and late graft patency. Attenuation of platelet deposition, aggregation, and activation has long been the goal of numerous interventions aimed at improving graft patency. Antiplatelet agents acting on the platelet surface glycoprotein IIb/IIIa ($\alpha_{IIb}\beta_3$ integrin) and on functional domains of thrombin have been shown to decrease platelet accumulation on Dacron grafts.^[11,12] Techniques used to alter surface thrombogenicity without compromising the graft material's bulk properties have been developed. These include the application of antiplatelet or anticoagulant substances to

biomaterial surfaces, alteration of surface electro-negativity, and radio frequency glow discharge of plasma polymerized monomers to graft surfaces.

Neutrophil and Monocyte Infiltration

Leukocyte chemoattractants C5a and leukotriene B₄ (LTB₄) mediate the acute inflammatory response through recruitment of neutrophils to the synthetic graft surface. These cells bind to fibrinogen within the fibrin coagulum on the inner graft surface through cell surface β_2 integrins (mainly CD11b/CD18), which mediate adhesion-dependent chemotaxis and promote cell transmigration. IgG is also important for neutrophil binding to artificial graft surfaces. IgG binds to neutrophils through Fc receptors present on the neutrophil cell surface, resulting in a series of effector functions. This recruitment of phagocytic cells does not confer infection resistance to the synthetic graft surface. On the contrary, neutrophils adherent to biomaterial surfaces have been shown to have reduced phagocytic and bactericidal properties. Neutrophil binding via the Fc receptor to immobilized IgG on graft surfaces may account for this deficiency.^[4] Neutrophils also interact with other deposited proteins such as C3bi and Factor X.

Neutrophils strongly adhere to endothelial cells in the perianastomotic region through selectin and integrin-mediated mechanisms.^[13,14] Activated neutrophils release oxygen free radicals and various proteases, resulting in matrix degradation and inhibition of both endothelialization and tissue incorporation of the vascular graft.

Circulating monocytes are attracted to areas of injured or regenerating endothelium and to areas pre-activated by interleukin-1 (IL-1) and tumor necrosis factor- α (TNF- α). Additional monocyte recruitment and activating factors include LTB₄, platelet factor 4, and PDGF.^[15,16] In response to these factors, monocytes differentiate into tissue macrophages, becoming the major participant in the host's chronic inflammatory response. Like neutrophils, macrophages also release proteases and oxygen free radicals, which have long-term effects on the tissue–graft–blood interfaces.

Inflammatory cells activated by synthetic graft materials release a variety of cytokines. Macrophages grown in the presence of polygalactin 910 (a bioresorbable material readily phagocytized by macrophages) release fibroblast growth factor 2 (FGF-2), which stimulates the proliferation of quiescent SMCs, ECs, and fibroblasts.^[17,18] Monocytes and macrophages grown in the presence of Dacron or ePTFE have been shown to produce interleukin-1 β (IL-1 β), interleukin-6 (IL-6), and TNF- α .^[19] Proliferation of SMCs caused by leukocyte–biomaterial interactions is

mediated by TNF- α , whereas SMC proliferation caused by leukocyte-EC interactions is mediated by IL-1.^[20] IL-1 also induces the upregulation of insulin-like growth factor-1 expression on ECs.^[21] IL-1 β has been shown to stimulate PDGF release from human umbilical vein ECs.^[22] Because the inflammatory reaction elicits a cascade of growth responses, approaches aimed at attenuating this initial inflammatory reaction may improve the long-term graft patency.

Endothelial and Smooth Muscle Cell Ingrowth

Native uninjured blood vessels possess an endothelium that continuously releases bioactive substances that inhibit thrombosis, promote fibrinolysis, and inhibit SMC proliferation to help maintain normal blood flow. After synthetic graft implantation, tissue ingrowth originates from the cut edge of the adjacent artery. Unlike many animal models in which there is often complete endothelialization of synthetic grafts, those implanted into humans exhibit only limited EC ingrowth confined to the first 1–2 cm of the graft at the anastomotic ends. Occasional islands of ECs present in the midportion of implanted grafts at significant distances from the anastomosis have been described and suggest that other mechanisms of graft endothelialization may be important. Interstitial tissue ingrowth, accompanied by microvessels from perigraft tissue, is one proposed mechanism for graft endothelialization on high-porosity ePTFE and bioresorbable prostheses.^[23–26] Deposition and proliferation of circulating ECs or stem cells is another mechanism supported by the finding of discrete isolated islands of ECs without transmural microvessels present on the inner surface of impervious Dacron grafts in a canine model and in a 24-month human Dacron graft explant.^[27–29]

ECs growing onto prosthetic graft surfaces undergo phenotypic modulation. These phenotypically altered activated cells secrete bioactive substances that may promote thrombogenesis and SMC growth. Activated ECs increase PDGF synthesis and secretion. Subintimal SMC proliferation occurs predominantly in areas with an endothelial lining, namely, the perianastomotic area.^[30] This region is an area of chronic EC injury, with highly complex biomechanical activity and chronic inflammation. SMC proliferation begins as early as two days post implantation, which then peaks at two weeks and tapers down thereafter. On the other hand, neointimal volume continues to increase by synthesis of extracellular matrix material, which ultimately on occasion may result in the formation of a stenotic lesion.^[31] Neointima formation in the middle portion of a graft is dependent on transmural ingrowth from surrounding perigraft tissue. Like ECs, SMCs grown onto prosthetic grafts are functionally altered,

producing significantly higher amounts of PDGF, which may contribute to the development of intimal hyperplasia.^[32] In addition, inflammatory mononuclear phagocytes and foreign body giant cells produce a variety of growth-modulating substances including PDGF, FGF, and TGF- β , all of which perpetuate SMC proliferation and production of extracellular matrix components.

GRAFT CHARACTERISTICS AFFECTING VASCULAR HEALING

Chemical composition, construction parameters, and biomechanics are the intrinsic characteristics of a vascular graft that predominantly influence its interaction with its host and, in turn, determine its fate.

Porosity

A direct correlation between porosity and tissue ingrowth has been well established in textile grafts where greater tissue ingrowth is seen in high-porosity grafts such as knitted Dacron versus low-porosity grafts such as woven Dacron.

The rate of tissue ingrowth on ePTFE is dependent upon graft permeability. Clowes has shown in a baboon model enhanced tissue ingrowth with complete endothelialization in ePTFE grafts with a 60- to 90-micron internodal distance.^[23] Transinterstitial capillary ingrowth did not occur in 30-micron internodal distance ePTFE grafts. In a canine model, 90-micron internodal distance ePTFE grafts demonstrated greater endothelial cell coverage and higher patency rates compared with similarly treated 30-micron internodal distance ePTFE grafts.^[33] In contrast, a small preliminary human trial using the more widely expanded ePTFE failed to show any significant advantage in platelet deposition over standard 30-micron internodal distance ePTFE grafts.^[34] In addition, transinterstitial ingrowth may differ between biomaterials with equal porosity, e.g., PGA versus Dacron.^[35] Greisler has demonstrated significantly greater transinterstitial capillary ingrowth and subsequent endothelialization in lactide/glycolide woven grafts compared to Dacron grafts implanted in rabbit and canine models.^[26,36] High-porosity grafts have the practical disadvantage of requiring preclotting to avoid transmural blood ooze after implantation.

Compliance

Atherosclerotic arteries have a wide range of compliance characteristics. The compliance mismatch between arteries and grafts results in flow disruption,

which is thought to influence pseudointimal hyperplasia at the anastomosis.^[37] Numerous attempts at improving graft compliance have focused on increasing biomaterial flexibility. Polyurethane grafts have compliance characteristics more similar to normal arteries, compared with ePTFE.^[38] Its corrugated structure significantly reduces graft stiffness and improves compliance.^[39] Placement of a vein cuff at the distal anastomosis between prosthetic graft and native artery has been shown to improve the compliance mismatch, possibly contributing to improved long-term patency rates.^[40–42] Decreased neointimal thickness in the perianastomotic area has also been shown in animal models.^[43] Use of the distal vein cuff has gained popularity in femoropopliteal arterial reconstructions utilizing prosthetic material, especially in below-knee bypasses.

Arterial compliance variability changes over time, resulting in the development of a hypercompliant zone in the region of the anastomosis.^[44] In addition, elastic grafts lose their compliance over time due to a fibroplastic reaction that occurs following implantation.^[44]

Flow Surface Characteristics

The interfacial reactions between blood and the graft biomaterial surface are dependent on surface chemical and physical properties. These properties include surface charge, surface energy, and the degree of roughness. A negative surface charge inhibits thrombosis, whereas a positive charge and heterogeneity of charge density distribution are thought to be thrombogenic. Several approaches to modifying flow surface characteristics have been developed that stabilize the thrombotic reaction. These approaches include the modification of surface properties and incorporation of antiplatelet and/or anticoagulant substances onto the graft surface, resulting in the endothelialization of the blood contact surface.

GRAFT FAILURE AND GRAFT INFECTION

Synthetic vascular graft occlusion has several etiologies. Immediate graft failure is often the result of technical problems during implantation or of a host hypercoagulable state. Failure in the first month following graft placement is likely the result of thrombosis in the setting of high distal flow resistance. Anastomotic pseudointimal hyperplasia is the most common etiology of graft failure six months to three years post graft placement. Later graft occlusions are usually due to the progression of distal atherosclerotic disease. Thus, thrombogenicity and intimal hyperplasia represent the most common intrinsic causes of graft failure amenable to intervention by manipulation

of graft construction, composition, and/or by induced alterations in blood and tissue reactions to the implanted graft.

Clinical stenosis and patency rates reflect these intrinsic problems of synthetic grafts. Aortic reconstruction involves the placement of large-caliber Dacron or ePTFE grafts. Because of the relatively high flow rates and low outflow resistance, these grafts have an 85–95% five-year patency rate.^[45,46] Compared with smaller caliber grafts used in the infrainguinal and infrapopliteal regions, the results are less than optimal with one- and three-year patency rates for ePTFE grafts used in infrapopliteal bypasses of 43% and 30%, respectively.^[47] These small-caliber grafts are especially susceptible to anastomotic intimal hyperplasia and are more prone to early thrombosis due to lower flow rates and higher resistance outflow vessels.

Prosthetic graft infection is reported to occur in 1–5% of implanted grafts. These infections are often catastrophic, resulting in amputation rates as high as 50% and 25–75% mortality.^[48,49] Development of prosthetic materials resistant to bacterial infection remains an active area of research.^[50] Penicillins and cephalosporins have been noncovalently bound to Dacron and ePTFE surfaces with subsequent increased resistance to *Staphylococcus aureus* infection. Rifampin-bonded gelatin-sealed Dacron grafts have also been shown to provide increased protection against bacterial colonization.^[51,52]

CONCLUSION

A firm understanding of graft physiology and its pathological processes is critical to the development of efficacious vascular substitutes. Investigations into the underlying biochemical and cellular mechanisms involved in the graft-healing process has yielded a firm understanding of graft physiology and the pathological processes involved in graft failure. Utilization of numerous biochemical, molecular, biological, and tissue-engineering techniques, with the goal of improving long-term graft patency, will result in the development of novel vascular substitutes and allow a continued increase in our understanding of the graft-healing process.

ARTICLES OF FURTHER INTEREST

Antimicrobial/Antibiotic (Infection Resistance) Materials; Biomaterials Immune Response; Host Reactions; Polytetrafluoroethylene; Tissue Engineering of Blood Vessel; Vascular Grafts; Vascular Grafts: Development Strategies; Vascularization of Engineered Constructs



REFERENCES

1. Voorhees, A.B., Jr.; Jaretzki, A.; Blakemore, A.H. The use of tubes constructed of Vinyon N cloth in bridging arterial defects. *Ann. Surg.* **1952**, *135*, 332–336.
2. Conte, M.S. The ideal small arterial substitute, a search for the Holy Grail? *FASEB J.* **1998**, *12*, 43–45.
3. Vroman, L. Methods of investigating protein interaction on artificial and natural surfaces. *Ann. N.Y. Acad. Sci.* **1987**, *516*, 300–305.
4. De La Cruz, C.; Haimovich, B.; Greco, R.S. Immobilized IgG and fibrinogen differentially affect the cytoskeletal organization and bactericidal function of adherent neutrophils. *J. Surg. Res.* **1998**, *80*, 28–34.
5. Shepard, A.D.; Gelfand, J.A.; Callow, A.D.; O'Donnell, T.F., Jr. Complement activation by synthetic vascular prostheses. *J. Vasc. Surg.* **1984**, *1*, 829–838.
6. Toursarkissian, B.; Eisenberg, P.R.; Abendschein, D.R.; Rubin, B.G. Thrombogenicity of small-diameter prosthetic grafts. Relative contributions of graft-associated thrombin and factor Xa. *J. Vasc. Surg.* **1997**, *25*, 730–735.
7. Ito, R.K.; Rosenblatt, M.S.; Contreras, M.A.; Beophy, X.M.; LoGerfo, F.W. Monitoring platelet interactions with prosthetic graft implants in a canine model. *ASAIO Trans.* **1990**, *36*, M175–M178.
8. McCollum, C.N.; Kester, R.C.; Rajah, S.M.; Learoyd, P.; Pepper, M. Arterial graft maturation, the duration of thrombotic activity in Dacron aortobifemoral grafts measured by platelet and fibrinogen kinetics. *Br. J. Surg.* **1981**, *68*, 61–64.
9. Stratton, J.R.; Thiele, B.L.; Ritchie, J.L. Platelet deposition on Dacron aortic bifurcation grafts in man, quantitation with indium-111 platelet imaging. *Circulation* **1982**, *66*, 1287–1293.
10. Stratton, J.R.; Thiele, B.L.; Ritchie, J.L. Natural history of platelet deposition on Dacron aortic bifurcation grafts in the first year after implantation. *Am. J. Cardiol.* **1938**, *52*, 371–374.
11. Mazur, C.; Tschopp, J.F.; Faliakou, E.C.; Gould, K.E.; Diehl, J.T.; Pierschbacher, M.D.; Connolly, R.J. Selective α IIb β 3 receptor blockage with peptide TP9201 prevents platelet uptake on Dacron vascular grafts without significant effect on bleeding time. *J. Lab. Clin. Med.* **1994**, *124*, 589–599.
12. Kelly, A.B.; Maragamo, J.M.; Bourdon, P.; Hanson, S.R.; Harker, L.A. Antithrombotic effects of synthetic peptides targeting various functional domains of thrombin. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 6040–6044.
13. Bevilacqua, M.P.; Pober, J.S.; Wheeler, M.E.; Cotran, R.S.; Gimbrone, M.A., Jr. Interleukin-1 acts on cultured human vascular endothelium to increase the adhesion of polymorphonuclear leukocytes, monocytes, and related leukocyte cell lines. *J. Clin. Invest.* **1985**, *76*, 2003–2011.
14. Gamble, J.R.; Harlan, J.M.; Klebanoff, S.J.; Vadas, M.A. Stimulation of the adherence of neutrophils to umbilical vein endothelium by human recombinant tumor necrosis factor. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 8667–8671.
15. Ford-Hutchinson, A.W.; Bray, M.A.; Doig, M.V.; Shipley, M.E.; Smith, M.J. Leukotriene B₂, a potent chemokinetic and aggregating substance released from polymorphonuclear leukocytes. *Nature* **1980**, *286*, 264–265.
16. Deuel, T.F.; Senior, R.M.; Chang, D.; Griffin, G.L.; Heinrichson, R.L.; Kaiser, E.T. Platelet factor 4 is chemotactic for neutrophils and monocytes. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 4584–4587.
17. Greisler, H.P.; Ellinger, J.; Henderson, S.C.; Shaheen, A.M.; Burgess, W.H.; Lam, T.M. The effects of an atherogenic diet on macrophage/biomaterial interactions. *J. Vasc. Surg.* **1991**, *14*, 10–23.
18. Greisler, H.P.; Petsikas, D.; Cziperle, T.M.; Murchan, P.; Henderson, S.C.; Lam, D.J. Dacron stimulation of macrophage transforming growth factor-beta release. *Cardiovasc. Surg.* **1996**, *4*, 169–173.
19. Swartbol, P.; Truedsson, L.; Pärsson, H.; Norgren, L. Tumor necrosis factor- α and interleukin-6 release from white blood cells induced by different graft materials in vitro are affected by pentoxifylline and iloprost. *J. Biomed. Mater. Res.* **1997**, *36*, 400–406.
20. Mattana, J.; Effiong, C.; Kapasi, A.; Singhal, P.C. Leukocyte-polytetrafluoroethylene interaction enhances proliferation of vascular smooth muscle cells via tumor necrosis factor-alpha secretion. *Kidney Int.* **1997**, *52*, 1478–1485.
21. Glazebrook, H.; Hatch, T.; Brindle, N.P. Regulation of insulin-like growth factor-1 expression in vascular endothelial cells by the inflammatory cytokines interleukin-1. *J. Vasc. Res.* **1998**, *35*, 143–149.
22. Totani, L.; Cumashi, A.; Piccoli, A.; Lorenzet, R. Polymorphonuclear leukocytes induce PDGF release from IL-1beta-treated endothelial cells, role of adhesion molecules and serine proteases. *Arterioscler. Thromb. Vasc. Biol.* **1998**, *18*, 1534–1540.
23. Clowes, A.W.; Kirkman, T.R.; Reidy, M.A. Mechanisms of arterial graft healing: Rapid transmural capillary ingrowth provides a source of intimal endothelium and smooth muscle in porous PTFE prostheses. *Am. J. Pathol.* **1986**, *123*, 220–230.
24. Greisler, H.P.; Dennis, J.W.; Endean, E.D.; Ellinger, J.; Burrle, K.F.; Kim, D.U. Derivation of neointima in vascular grafts. *Circulation* **1988**, *78*, 16–112.
25. Golden, M.A.; Hanson, S.R.; Kirkman, T.R.; Schneider, P.A.; Clowes, A.W. Healing of polytetrafluoroethylene arterial grafts is influenced by graft porosity. *J. Vasc. Surg.* **1990**, *11*, 838–844.
26. Onuki, Y.; Kouchi, Y.; Yoshida, H.; Wu, M.H.D.; Shi, Q.; Wechezak, A.R.; Coan, D.; Sauvage, L.R. Early flow surface endothelialization before microvessel ingrowth in accelerated graft healing, with BrdU identification of cellular proliferation. *Ann. Vasc. Surg.* **1998**, *12*, 207–215.
27. Shi, Q.; Wu, M.H.; Hayashida, N.; Wechezak, A.R.; Clowes, A.W.; Sauvage, L.R. Proof of fall-out endothelialization of impervious Dacron grafts in the aorta and inferior vena cava of the dog. *J. Vasc. Surg.* **1994**, *20*, 546–556.
28. Kouchi, Y.; Onuki, Y.; Wu, M.H.; Shi, Q.; Ghali, R.; Wechezak, A.R.; Kaplan, S.; Walker, M.; Sauvage,

- L.R. Apparent blood stream origin of endothelial and smooth muscle cells in the neointima of long, impervious carotid-femoral grafts in the dog. *Ann. Vasc. Surg.* **1998**, *12*, 46–54.
29. Wu, M.H.; Shi, Q.; Wechezak, A.R.; Clowes, A.W.; Gordon, I.L.; Sauvage, L.R. Definitive proof of endothelialization of a Dacron arterial prosthesis in a human being. *J. Vasc. Surg.* **1995**, *21*, 862–867.
 30. Greisler, H.P. Regulation of Vascular Graft Healing by Induction of Tissue Incorporation. In *Human Biomaterials Applications*, 1st Ed.; Wise, D.L., Altobelli, D.E., Yaszemski, M.J., Greisler, H.P., Eds.; Humana Press, Inc.: Totowa, NJ, 1996; 227.
 31. Hamdan, A.D.; Misare, B.; Contreras, M.; Logerfo, F.W.; Quist, W.C. Evaluation of anastomotic hyperplasia progression using the cyclin specific antibody MIB-1. *Am. J. Surg.* **1996**, *172*, 168–171.
 32. Pitsch, R.J.; Minion, D.J.; Goman, M.L.; van Aalst, J.A.; Fox, P.L.; Graham, L.M. Platelet-derived growth factor production by cells from Dacron grafts implanted in a canine model. *J. Vasc. Surg.* **1997**, *26*, 70–78.
 33. Cameron, B.L.; Tsuchida, H.; Connall, T.P.; Nagae, T.; Furukawa, K.; Wilson, S.E. High porosity PTFE improves endothelialization of arterial grafts without increasing early thrombogenicity. *J. Cardiovasc. Surg.* **1993**, *34*, 281–285.
 34. Clowes, A.W.; Kohler, T. Graft endothelialization: The role of angiogenic mechanisms. *J. Vasc. Surg.* **1991**, *13*, 734–736.
 35. Greisler, H.P.; Kim, D.U.; Price, J.B.; Voorhees, A.B., Jr. Arterial regenerative activity after prosthetic implantation. *Arch. Surg.* **1985**, *120*, 315–323.
 36. Greisler, H.P.; Ellinger, J.; Schwarcz, T.H.; Golan, J.; Raymond, R.M.; Kim, D.U. Arterial regeneration over a polydioxanone prostheses in the rabbit. *Arch. Surg.* **1987**, *122*, 715–721.
 37. Abbott, W.M.; Cambria, R.P. Control of Physical Characteristics, Elasticity and Compliance of Vascular Grafts. In *Biological and Synthetic Vascular Prostheses*, 1st Ed.; Stanley, J.C., Ed.; Grune and Stratton: New York, 1982; 189.
 38. Stansby, G.; Berwanger, C.; Shukla, N.; Schmitz-Rixen, T.; Hamilton, G. Endothelial seeding of compliant polyurethan vascular graft material. *Br. J. Surg.* **1994**, *81*, 1286–1289.
 39. Stump, D.M.; Hart, V.G.; Tuttle, S.L. Circumferential flexible vascular grafts. *J. Biomech.* **1998**, *31*, 705–710.
 40. Benedetti-Valentini, F.; Gossetti, B.; Irace, I.; Martinnelli, O.; Gattuso, R. Composite grafts for critical ischaemia. *Cardiovasc. Surg.* **1994**, *4*, 372–379.
 41. Pappas, P.J.; Hobson, R.W., II; Meyers, M.G.; Jamil, Z.; Lee, B.C.; Silva, M.B., Jr.; Goldberg, M.C.; Padberg, F.T., Jr. Patency of infra-inguinal polytetrafluoroethylene bypass grafts with distal interposition vein cuffs. *Cardiovasc. Surg.* **1998**, *6*, 19–25.
 42. Neville, R.F.; Attinger, C.; Sidawy, A.N. Prosthetic bypass with a distal vein patch for limb salvage. *Am. J. Surg.* **1997**, *174*, 173–178.
 43. Gentile, A.T.; Mills, J.L.; Gooden, M.A.; Hagerty, R.D.; Berman, S.S.; Hughes, J.D.; Kleinhert, L.B.; Williams, S.K. Vein patching reduces neointimal thickening associated with prosthetic graft implantation. *Am. J. Surg.* **1998**, *176*, 601–607.
 44. Hasson, J.E.; Megerman, J.; Abbott, W.M. Increased compliance near vascular anastomoses. *J. Vasc. Surg.* **1985**, *2*, 419–423.
 45. Moore, W.S.; Cafferata, H.T.; Hall, A.D.; Blaisdell, F.W. In defense of grafts across the inguinal ligaments: An evaluation of early and late results of aorto-femoral bypass grafts. *Ann. Surg.* **1968**, *168*, 207–214.
 46. Lind, R.E.; Wright, C.B.; Lynch, T.G.; Lamberth, W.C., Jr.; Slaymaker, E.E.; Brandt, B., III. Aorto-femoral bypass grafting: Microvel. *Am. Surg.* **1982**, *48*, 89–92.
 47. Veith, F.J.; Gupta, S.K.; Ascher, E.; White-Flores, S.; Samson, R.H.; Scher, L.A.; Towne, J.B.; Bernhard, V.M.; Bonier, P.; Flinn, W.R. Six-year prospective multicenter randomized comparison of autologous saphenous vein and expanded polytetrafluoroethylene grafts in infra-inguinal arterial reconstructions. *J. Vasc. Surg.* **1986**, *3*, 104–114.
 48. Leikweg, W.G.; Greenfield, L.J. Vascular prostheses graft infections: Collected experiences and results of treatment. *Surgery* **1977**, *81*, 335–342.
 49. Szilagyi, D.E.; Smith, R.F.; Elliot, J.P.; Vrandecic, M.P. Infection in arterial reconstruction with synthetic vascular graft. *Ann. Surg.* **1972**, *176*, 321–333.
 50. Goldstone, J.; Moore, W.S. Infections in vascular prostheses. *Am. J. Surg.* **1974**, *128*, 225–233.
 51. Goeau-Brissoniere, O.; Mercier, F.; Nicolas, M.H.; Bacourt, F.; Coggia, M.; Lebrault, C.; Pechere, J.C. Treatment of vascular graft infection by in situ replacement with a rifampin-bonded gelatin sealed Dacron graft. *J. Vasc. Surg.* **1994**, *19*, 739–741.
 52. Vicaretti, M.; Hawthorne, W.J.; Ao, P.Y.; Fletcher, J.P. An increased concentration of rifampicin bonded to gelatin-sealed Dacron reduces the incidence of subsequent graft infections following a staphylococcal challenge. *Cardiovasc. Surg.* **1998**, *6*, 268–273.



Vascular Imaging

Tobias Saam
William S. Kerwin
Chun Yuan

Department of Radiology, University of Washington, Seattle, Washington, U.S.A.

INTRODUCTION

The term vascular imaging describes the use of different imaging modalities to visualize the arterial and venous system for the diagnosis of various diseases involving the vascular system. In 1929, Dos Santos reported the first iodine-based contrast arteriogram in a patient, approximately 34 years after the discovery of the X ray and almost simultaneously with the description of the first iodinated contrast medium. For the first time, it was possible to visualize arteries in a living human being. Since then, not only has the number of imaging methods increased, but each one continues to undergo improvement and refinement of its use in medical diagnosis. The classical angiography was aimed at identifying luminal changes due to vascular disease. These techniques, generally referred to as angiograms, do not provide extraluminal information. Newer applications like ultrasound, computer tomography, and magnetic resonance tomography are not only able to noninvasively image lumen from arteries such as the aorta (2.5 cm in diameter) as well as arteries as small as 1 mm in diameter, but also provide additional information about the arterial wall. This is especially important, as vascular disease starts in the wall (extraluminal), and only advanced vascular disease leads to severe lumen narrowing.

The need for vascular imaging is immense. In the western world, atherosclerosis is the most common cause of death, resulting in coronary heart disease, myocardial infarction, stroke, and peripheral vascular disease. Vascular imaging is able to provide important information about the extent of the disease and has become an irreplaceable tool in the clinical practice, not only in atherosclerosis imaging. Vascular imaging is routinely used before liver and renal transplantation, for surgical procedure planning of solid tumors, in the diagnosis of thrombosis or congenital heart disease, and many more. In this article, we will give an oversight of the currently used techniques in vascular imaging, their advantages and disadvantages, and future perspectives of the techniques.

ANGIOGRAPHY

Angiography was the first vascular imaging modality that entered the clinical practice. In ordinary X-ray projection images, blood vessels are hardly visible due to the very low contrast between vessels and the surrounding tissue. This contrast is enhanced by the injection of a radiopaque contrast medium into the vessels to be diagnosed. Currently, nonionic iodinated contrast agents are used. Typical reactions to contrast media include nausea and/or vomiting, scattered to extensive urticaria, bronchospastic reaction, hypotension (isolated) with compensating tachycardia, anaphylactoid reaction, vagal reaction, cardiovascular collapse, convulsion, and seizure. The rate of side effects is 0.2% with deaths occurring approximately 1 in 100,000 applications.^[1] Conventional arteriography uses small flexible catheters that are placed in the arterial system usually via puncture of the femoral artery in the groin. With the use of fluoroscopy for guidance, catheters of various sizes and shapes can be manipulated selectively into virtually every major artery. Contrast injection is performed by hand or by mechanical injector and is accompanied by a timed rapid-sequence filming or digital computer acquisition of the fluoroscopic image. The result is a timed series of images depicting the contrast flow through the artery injected and the tissues that the artery supplies. Conventional venography is performed by contrast injection of veins via distal puncture or selective catheterization.

Before the introduction of digital techniques, film angiography was the method used for all X-ray angiography. According to Mistretta,^[2] film combines excellent spatial resolution (five- or six-line pairs per millimeter) with a large available field of view (FOV), and high temporal resolution is available in cine mode.

Disadvantages of X-ray film angiography^[2] include increased imaging time due to film development requirements, the need for larger contrast agent doses than required for digital systems, and the absence of real-time subtraction capabilities so that in clinical practice, digital subtraction angiography (DSA) is the method of choice. DSA is a powerful technique for the

visualization of blood vessels in the human body and is able to depict blood vessels with a diameter of 500 μm . In DSA imaging, a sequence of images is taken to show the passage of a bolus of injected contrast material through one or more vessels of interest (see Fig. 1). The contrast distortions in these live or contrast images are largely removed by subtracting an image taken prior to the arrival of the contrast medium (referred to as the mask image). Intraarterial (IA)-DSA is used for an increasing number of X-ray angiographic procedures, e.g., stent placement or percutaneous angioplasty for the treatment of stenosis in various vascular beds, thrombolysis or mechanical thrombectomy for treatment of thrombosis, and embolotherapy to stop or prevent bleeding or to destroy/occlude neoplasm or vascular abnormalities.

The advantages of IA-DSA according to Mistretta,^[2] include reduced contrast agent requirements and smaller catheters relative to conventional angiography. It has good spatial and temporal resolution and facilitates diagnostic and interventional procedures in which the immediate availability of subtracted images and road-mapping capabilities are helpful. Selective imaging of vessels can be accomplished. The overall frequency of catheterization-related complications is about 7–9%; the risk of a major complication is about 3 to 6%.^[3] Aside from its invasiveness, the visual interpretation of angiograms exhibits significant observer variability and correlates poorly with post-mortem examination.^[4] Angiography depicts arteries as a planar silhouette of the contrast-filled lumen and any arbitrary angiographic projection can misrepresent



Fig. 1 Angiography of the aortic arch in DSA technique; bones are subtracted. The chevron shows the tip of the intraarterial catheter. The cross marks a pseudoaneurysm that developed after traumatic injury. (Courtesy Dr. E. Hoffer, Harborview Medical Center, Seattle.)

the true extent of luminal narrowing. Mechanical interventions may increase lumen irregularity, impairing the accuracy of angiography.^[5] An assessment of lesion severity requires the measurement of luminal diameter and an uninvolved normal segment. Angiography does not provide information about the arterial wall and is confounded by outward remodeling of the vessel wall, which may conceal early atherosclerosis.

COMPUTED TOMOGRAPHY ANGIOGRAPHY (CTA)

Early development of X-ray computed tomography (CT) took place in the 1960s and the first head CT scanner was introduced in 1972. For their groundbreaking work in CT, Godfrey Hounsfield of Britain and Allan Cormack of the United States earned the Nobel Prize for medicine in 1979. CT images are derived from hundreds of thousands of measurements of the X-ray attenuation through a thin (0.5–10 mm) section of the patient. A narrow, well-collimated beam of X rays is generated on one side of the patient and as the X-ray beam passes through the patient, it is attenuated by absorption and scatter. The net attenuation of the slice is then measured by sensitive X-ray detectors on the opposite side of the patient. After repeatedly measuring attenuation from many different directions by rotating the X-ray tube around the patient, CT numbers are extracted for each pixel in the image by a computer algorithm.^[6] CT numbers reflect the local attenuation coefficient and are expressed in Hounsfield units (H), named for the inventor of CT. Voxel dimensions are determined by the computer algorithm chosen for reconstruction and the thickness of the scanned slice.

In CTA, as in conventional X-ray angiography, intravenous radiopaque contrast agents are needed to demonstrate vascular anatomy and vessel patency. Currently used slice thickness depends on the vascular bed and is 0.5 mm for the coronary arteries up to 5 mm sections for the aorta. The most common indications for CTA are the evaluation of the thoracic and abdominal aorta, the diagnosis and follow-up of pulmonary embolism, and for surgical therapy planning.

A distinct benefit of CTA over DSA is the ability to obtain a three-dimensional (3D) reconstruction of vascular anatomy (see Fig. 2). Conventional (non-helical) CT obtains 3D image data one slice at a time. The patient holds his or her breath, a slice is taken, the patient breathes, the table moves, and the sequence is repeated. Such sequential acquisition and reconstruction of multiple parallel sections constituted the first truly three-dimensional (3D) medical imaging modality. This technique is, however, time-consuming, and multislice spiral CT is currently the state of the art

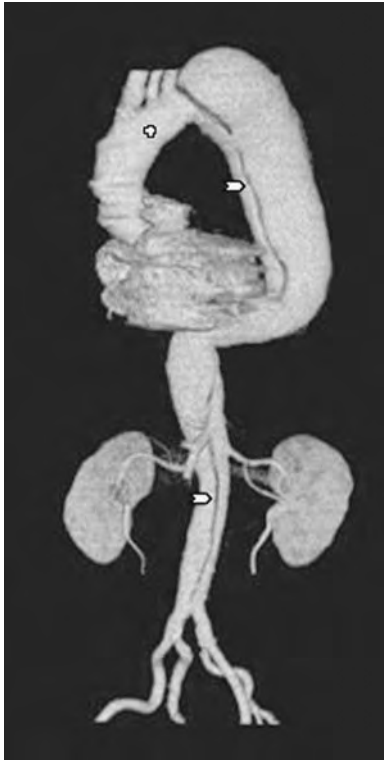


Fig. 2 Three-dimensional computer tomography angiography from the aorta and renal and iliac arteries in volume-rendering technique. The aorta (cross) has a large aneurysm and a dissection membrane (chevron) is visible.

CT unit. Helical CT, also called spiral CT, is performed by moving the patient table at a constant speed through the CT gantry while scanning continuously with an X-ray tube rotating around the patient. A continuous volume of image data is acquired during a single breath-hold. Multislice CT is based on the acquisition of more than a single slice at a time with the use of multirow detector systems. Multislice CT dramatically improves the speed of image acquisition, enables scanning during optimal contrast opacification, and eliminates artifacts caused by misregistration and variations in patient breathing.^[7] Volume acquisition enables retrospective reconstruction of multiple overlapping slices, improving visualization of small structures, and allowing high-detail three-dimensional CT angiography.^[7]

Advantages of CTA^[2] include the ability to retrospectively reproject from arbitrary angles (vs. X-ray angiography). The information obtained is anatomic and is not subject to the problems of spin saturation and intravoxel dephasing that poses difficulties in MR angiography. Scan times are shorter compared to MRI. The intravascular injection offers a reduction in invasiveness relative to IA-DSA. Spiral CT also provides extraluminal information that IA-DSA does not. Disadvantages of this technique^[2] are radiation

exposure, worse soft tissue contrast compared to MR angiography, and more side effects of the nonionic iodinated contrast agent used for CTA compared to the paramagnetic contrast agents used for MR angiography.

ULTRASOUND (US)

Vascular Ultrasound

Ultrasound (US) has come into widespread use in clinical practice for the assessment of the vascular tree since the parallel development of Doppler US and real-time two-dimensional B-mode US, beginning around 1975. Duplex US is a combination of 2-D gray-scale imaging and Doppler velocity spectral measurements obtained in a selectable region of interest (ROI) on the gray-scale image. This technique, primarily because of its noninvasiveness and low cost, is used or is under investigation for a wide variety of diagnostic screening applications, including intra- and extracranial vessels, the coronary arteries, abdominal vessels, and peripheral vessels. A duplex US examination involves two components:^[8]

(I) Real-time, B-mode imaging, which results in a two-dimensional gray-scale image that is continuously updated (hence the term real-time). The term B-mode is a shortened form of the term brightness modulation, which describes how the two-dimensional image is produced as a result of the brightness of each individual pixel in the image display (see Fig. 3). This brightness is in turn dependent on the intensity of the reflected US signal. The basic physics is well described elsewhere.^[9] Essentially, different tissues have varying acoustic properties and reflect sound differently. The US detectors in the probe receive signals reflected from these different tissues and an image is produced from this information by computer processing. Advances in gray-scale US have resulted in much improved anatomic definition, and the extent and nature of intimal thickening and plaque formation can be well seen. Calcification and air may prevent a satisfactory B-mode image from being produced, because sound energy is nearly completely absorbed at interfaces between soft tissue and bone/air, causing an acoustic shadow. Typical spatial resolution for a 3.5-to-5 MHz transducer is 0.5-to-1 mm axial and 2-to-5 mm lateral. Higher frequencies have smaller wavelengths and provide better axial resolution; however, the imaging depth is reduced. Lateral resolution is determined by the beam diameter emanating from the transducer element(s) and varies greatly with depth.

(II) Doppler assessment. The physics is well described elsewhere.^[9] Basically, moving blood generates a Doppler frequency shift in the reflected sound from

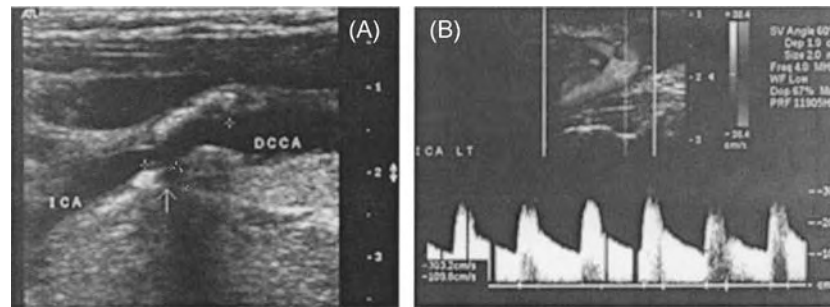


Fig. 3 (A) B-Mode ultrasound image showing a heterogeneous plaque with calcific shadowing (arrow) in the carotid artery. ICA = Internal Carotid Artery, DCCA = Distal Common Carotid Artery. (B) At the top, B-Mode ultrasound image that shows the flow in color Doppler mode of the left carotid artery, and at the bottom, corresponding duplex waveform of the internal carotid artery. The peak systolic velocity is increased to >300 cm/s, suggesting a 50–79% diameter reduction. (Courtesy of B. Mraz, J. Wigen, and J. Primozych, Diagnostic Vascular Service, University of Washington Medical Center, Seattle.)

insonated red blood cells and this frequency shift can be used to calculate the velocity of the moving blood. Low resistance vessels such as the internal carotid and renal arteries are characterized by high diastolic flow, whereas high resistance vessels such as the external carotid and lower and upper limb arteries are characterized by low or absent diastolic flow with early diastolic flow reversal. Stenosis in any vessel is characterized by both an increased range of velocities due to turbulent flow (spectral broadening), as well as an increase in systolic and diastolic velocities (see Fig. 3B). Tables have been produced and modified with experience to relate the systolic and diastolic velocities at the stenosis with that in the disease-free proximal vessel to accurately quantitate the percentage stenosis. Although there are many sources of error and artifact, the single most important technical aspect is that the interrogating angle must be <60 degrees or major errors will result.

An additional degree of convenience is provided by color Doppler flow imaging, in which average blood velocity is color-coded and superimposed on a gray-scale image. Regions of complex flow are demonstrated by irregularities in the color pattern. The most common indications of vascular ultrasound are the evaluation of the carotid, intracranial, and peripheral arteries; diagnosis of venous thrombosis; and evaluation of liver and renal transplants. Advantages of vascular US^[2] include absence of ionizing radiation and real-time noninvasive imaging that provides intra- and extraluminal information in a relatively inexpensive and widely available package. Ultrasound is accurate in diagnosing carotid artery stenosis compared to DSA,^[10] although it has been demonstrated to have lower discriminatory power compared with MRA in diagnosing 70–99% stenosis.

One of the principal limitations of US is the need for an appropriate acoustic window. Regions of

intervening air or bone prevent acquisition of diagnostic information in a substantial number of potential vascular sites. Other disadvantages include limited FOV, dependence on Doppler angle, dependence on operator skill, and less morphological information about the arterial wall compared to MRI.^[11]

Intravascular Ultrasound

In order to obtain the high spatial resolution possible at ultrasonic frequencies of 30–40 MHz (ca. 0.1 mm), the transducer is mounted at the tip of an intravascular catheter so that the blood vessel walls and associated pathologies can be examined at close range. This minimizes loss due to attenuation of ultrasound.

Nissen et al.^[4] reported on intravascular US of the coronary arteries, which they found to be a promising technique without significant risk. In a study of angiographically normal sites in coronary artery disease patients, Nissen et al. found that more than 75% had atherosclerotic abnormalities by US, emphasizing that coronary disease is often more diffuse than indicated by angiography, and undermining the need of imaging modalities that provide intra- and extraluminal information. One of the limitations of intravascular US of the coronaries (besides its invasiveness) is that the catheter size is limited to about 1 mm, which does not permit imaging of tight stenosis and precludes examination of many lesions before angioplasty.^[4]

MAGNETIC RESONANCE ANGIOGRAPHY (MRA)

Paul C. Lauterbur and Peter Mansfield earned the Nobel Prize for medicine in 2003 for their discoveries in magnetic resonance (MR) imaging in the 1970s. MR is a potent technique that produces tomographic



images by means of magnetic fields and radiowaves. Although CT evaluates only a single tissue parameter—X-ray attenuation—MR analyzes multiple tissue characteristics including hydrogen (proton) density, T1 and T2 relaxation times of tissue, and blood flow within tissue. The soft tissue contrast provided by MR is substantially better than for any other image modality. Differences in the density of protons available to contribute to the MR signal discriminate one tissue from another. The complicated physics of MR is beyond the scope of this article.

MRA was experimentally introduced in the mid-1980s to visualize major blood vessels of the neck, abdomen, and thigh. Since the introduction, three principal MRA techniques have been used in clinical imaging of vascular disease. These are time-of-flight MRA (TOF-MRA) and phase-contrast MRA (PC-MRA)—where no injection of contrast material is needed—and more recently, the contrast-enhanced MRA (CE-MRA), using intravenous injection of gadolinium contrast. These so-called bright-blood techniques enhance the signal intensity of flowing blood; thus, the lumen appears hyperintense relative to the adjacent vessel wall. Black-blood techniques refer to MR techniques that suppress the signal from flowing blood.^[12] The most common indications for MRA are the evaluation of the thoracic and abdominal aorta, as well as the evaluation of cranial and peripheral blood vessels.

Time of Flight (TOF)

TOF relies on inflow enhancement. Here, static tissue within a 2-D slice or 3-D slab gives a low signal due to the saturation effect of the long train of closely spaced (short TR) excitation pulses used in forming the image.^[13] On flowing into the imaging volume, unsaturated blood appears hyperintense relative to the static surrounding. Unfortunately, this bright-blood appearance depends on flow velocity, and the path of the vessel through the image volume (i.e., the number and strength of excitation pulses experienced). Efforts in the early 1990s focused largely on maximizing the contrast between blood and parenchyma. This involves reducing the saturation of blood by the succession of excitation pulses experienced while in the imaging volume and simultaneously maximizing the saturation of background tissue signal. This is notable in 2-D acquisition where a large flip angle minimizes the static tissue signal. Whereas this gives maximal inflow effect, segments of the vessel lying within the plane rapidly become saturated, causing signal loss. Using a lower flip angle and a 3-D acquisition, the rate at which blood becomes saturated is slowed, although the degree of background saturation is also

reduced. An intermediate design is that of a series of thin 3-D volumes (e.g., multiple overlapping thin slab angiography—MOTSA) rather than a single thick slab.

Phase-Contrast Magnetic Resonance Angiography

In PC-MRA, the contrast between flowing blood and the surrounding stationary tissue is generated by the flow-induced phase shift of the moving blood. When a magnetic spin in flowing blood is subjected to a gradient magnetic field, it acquires a phase shift that is proportional to its velocity. In the presence of a pair of bipolar gradient lobes, the stationary spins do not develop a net phase shift. Because the flow-induced phase shifts are proportional to the signal intensity, flow direction and velocity can be derived directly from PC data.

Contrast-Enhanced Magnetic Resonance Angiography

Both TOF and PC-MRA techniques are essentially noninvasive and rely on inflow enhancement or phase shift to generate contrast between flowing blood and stationary tissues. There are several disadvantages associated with these techniques. One is inplane saturation, which can be a problem with slow flowing or tortuous arteries. A second is turbulence-induced signal loss in a region of narrowing. Imaging times also tend to be long, especially for covering a large vasculature. A powerful means of maximizing blood signal intensity is through the administration of a T1 shortening contrast agent. This means the inflow pressure is not as essential as in TOF-MRA or conventional contrast arteriography. First-pass contrast-enhanced studies have become widely accepted for abdominal, thoracic, and peripheral MRA since their first demonstration by Prince in 1993^[14] (see Fig. 4). With the T1 of blood reduced to less than 50 ms (i.e., shorter than fat), very heavily T1-weighted images can be acquired in which blood will appear white because of its high signal intensity relative to the saturated surrounding tissues. As with conventional angiography, background suppression can be improved by the subtraction of a precontrast mask from the enhanced study.^[15] Following venous injection, maximal arterial enhancement is achieved by a short bolus delivery. The period of arterial enhancement before the arrival of contrast agent in neighboring veins or its uptake in surrounding tissue forms the effective time window for contrast enhanced arterial imaging.^[16] Three-dimensional CE-MRA with short echo times reduce the examination time,



Fig. 4 Contrast-enhanced magnetic resonance angiography as a multiplanar reconstruction. The cross shows the ascending aorta; the chevron shows the descending aorta. This image also shows the inferior (up arrow) and superior caval (down arrow) vein.

minimize clip artifacts and artifacts caused by movement. Breathheld data acquisition did become possible with the use of improved gradient systems permitting considerable reductions in the minimum repetition (TR) and echo times (TE). The implementation of fast 3-D gradient echo sequences on higher-performance systems permit the acquisition of complex 3-D data sets within the confines of a comfortable breathhold interval under 30 s;^[16] in some of the newer generation systems, even under 10 or 5 s. This kind of dramatic shortening of scan times has even allowed collecting temporally-resolved 3-D data sets. Temporal information regarding the transit of the paramagnetic contrast agent through the vascular system can be harvested.

Black-Blood Techniques

These sequences are ideal for plaque imaging because the conspicuity of the vessel wall is increased when adjacent to a hypointense lumen, and the echo and repetition times can be varied to optimize visualization of specific plaque components.^[17] Black-blood, fast spin-echo-based methods offer particular advantages in cardiac imaging compared to white-blood imaging,^[18] including an exclusive visualization of the coronary lumen, reduced susceptibility to turbulent flow in the regions of focal stenosis, a reduced sensitivity to artifacts induced by metallic implants, and a high contrast-to-noise ratio.

Advantages and Disadvantages of MRI

Advantages of MRA according to Mistretta include its noninvasiveness, ability to image vessels in a broad range of sites, absence of radiation (compared to CTA), ability to obtain retrospective projections (compared to DSA), and the possibility of obtaining flow and velocity information. Additionally, MRI is able to provide information about the composition and morphology of the arterial wall, information that is crucial in the assessment of atherosclerosis. Side effects of the paramagnetic contrast agent are rare (ca. 0.06%,^[11]). Paramagnetic contrast agents are non-nephrotoxic and have a low incidence of allergic reactions. Hence, they are safe for use in patients with renal insufficiency as well as patients with a history of allergic reactions to iodinated contrast media.

Disadvantages include artifacts due to surgical clips or metallic stents, patient motion, and pulsatility. Typically, trade-offs must be made because of the inability to choose imaging parameters that simultaneously maximize FOV, signal-to-noise, and spatial resolution. Contraindications are pacemaker, intracranial aneurysm clips, intraocular metal, recent stents (less than three weeks), and claustrophobia (relative contraindication).

ANGIOSCOPY

Angioscopy is an intravascular imaging method in which a fiberoptic device is placed within the arterial lumen to provide a view of the vessel lumen and internal wall. This technique relies on high-pressure flushing with saline solution to provide a temporary FOV free of blood. Although spatial resolution is high (up to 200 μm), the information is limited to the surface and the technique is invasive, highly operator-dependent, and currently not approved by the Food and Drug Administration for use in the United States.

IMAGE PROCESSING

Developing in parallel with improvements in image acquisition technology have been computer graphics hardware and software tools that enable sophisticated image processing and display. Currently, besides the axial (cross-sectional) view, four image reconstruction techniques have been used: maximum intensity projection (MIP), curved multiplanar reconstruction (MPR), shaded-surface display, and volume rendering.^[19] Postprocessing of CT or MRA image data is almost a necessity in clinical applications for demonstrating spatial positioning, connectivity of vessels, and correlative MR/CT imaging. These algorithms have

the advantage that projection angles can be arbitrary. Hence, the 3D data set can be viewed from multiple angles. The full three-dimensionality inherent to the technique can be exploited only if the data are viewed on a work station. Analysis of the images should, however, exploit the fact that truly three-dimensional data sets are acquired. Analysis should therefore not be limited to viewing maximum intensity projections or surface shaded display, but instead should include thorough evaluation of reformatted source images in all three planes.

FUTURE DIRECTIONS

Angiography

In an article recently published by Katzen,^[20] the author made the following predictions regarding catheter-based angiography:

(I) Most diagnostic angiography will be performed with noninvasive methods. MRA will be the predominant method in the peripheral circulation, with CTA having an important role in aortic imaging and coronary imaging. MRA will have increased use for elective diagnosis and in clinical emergencies.

(II) Catheter-based angiography will have an important role as an adjunct or preliminary procedure related to interventions. This is significant for planning catheter development and functionality of future angiographic equipment.

(III) The need for angiographic therapy will increase as a result of a variety of factors, including aging population, early diagnosis, and increasing acceptance of less invasive therapy.

(IV) Interventionists should be considering and planning for vascular imaging devices (MRA, CTA, and US) to be included as part of the interventional sections and the workload.

Computer Tomography Angiography

It is expected that CTA with multidetector CT scanners will result in better image quality than with the single detector CT scanners, because smaller detector elements improve resolution and faster scan rotation and multiple detectors increase the coverage and/or decrease the scan duration.^[21] The advantages of a shorter scan duration are threefold: 1) less contrast material (30–50%) can be used; 2) the injection rate can be increased with a concomitant better enhancement of the vessels; and 3) most of the data can be acquired during a determined phase (arterial phase in CTA).^[21] The latest multislice CT system^[22] offers simultaneous acquisition of up to 16 submillimeter

slices and gantry rotation times of less than 0.5 s. Large scan ranges in short examination times with reduced amounts of contrast agents can be covered by 16-slice CT systems. Owing to submillimeter slice width and optimized in-plane resolution, almost isotropic voxel sizes of $0.5 \times 0.5 \times 0.6$ mm can be realized for vascular imaging.

Duplex Ultrasound

Newer developments include the use of power Doppler, three-dimensional imaging, contrast agents, and harmonic imaging. Power Doppler is the depiction of flow based on the integrated power of the Doppler spectrum rather than the mean Doppler frequency. It is therefore more sensitive in flow detection than is standard color Doppler and permits detection of low flow. Three-dimensional imaging uses advanced computer technology to produce a three-dimensional image of the area scanned. The role of these latter techniques is not yet established.^[23] Contrast agent has recently become available and consists of biodegradable microbubbles. They enhance vascularity and may prove useful in improving accuracy in certain areas such as assessment of the renal arteries. Harmonic imaging promises to improve sonographic imaging, particularly of structures that are degraded by movement, such as the heart and proximal renal vessels. With harmonic imaging, the US machine transmits the sound beam at a certain frequency, but receives at the resonant (harmonic) frequency, which is twice that of the transmitting frequency. Consequently, the signal-to-noise ratio is improved because, in theory, only the structures being imaged will produce the resonant frequency.^[24]

Magnetic Resonance Imaging (MRI)

Of all the described noninvasive imaging modalities, MRI is the most promising tool for vascular imaging. Although similar in principle to spiral CT angiography, contrast-enhanced MRA holds considerable advantages. Beyond the absence of ionizing radiation and the ability to depict large vascular territories in 3-D imaging volumes, paramagnetic contrast agents are usually better tolerated than iodinated contrast media. It is accurate in diagnosing stenosis compared to DSA and provides extraluminal information. Whole body MRA of the entire vascular tree from the supraaortic to the pedal arteries using a moving table is feasible in 72 s.^[16] There will be further improvement in scan time and/or resolution with the new high-field scanners (3T). Currently, the focus of research is directed to: 1) MRI-contrast agents; 2) plaque imaging; and 3) interventional MRA.

MRI contrast agents

Currently, extracellular nonbinding gadolinium (Gd) chelates that have regulatory approval for use in humans have been the contrast agents of choice for MRA. Modified paramagnetic Gd-based agents with varying degrees of protein as well as superparamagnetic compounds are currently under clinical evaluation.^[25] Some new agents have higher relaxivity and different routes of excretion, but still distribute into the extracellular spaces. Blood pool or intravascular contrast agents are large enough or bind to large enough molecules when injected that they do not leak out of the capillaries, but stay within the intravascular compartment.

In view of the rapid progress of techniques based on the use of extracellular agents, the future of intravascular contrast agents for morphologic imaging of the arterial vascular tree remains doubtful. At this time, it appears likely, however, that these agents will enhance imaging of the venous system and of the coronary arteries.^[16]

Plaque imaging

Traditionally, the degree of luminal stenosis has been used as a measure of atherosclerotic disease severity. In 1988, however, Ambrose et al.^[26] demonstrated in an angiographic study that mild-to-moderate coronary stenosis may lead to acute myocardial infarction, and suggested that luminal narrowing was not the sole predictor of clinical events. Over the past decade, notable advances have been made in our understanding of vascular biology. It is now commonly accepted that factors other than the degree of luminal narrowing, such as plaque burden, plaque morphology, and plaque composition, are objective noninvasive measurements in the assessment of atherosclerotic disease. Currently, the predominant noninvasive imaging modalities being investigated for this purpose are US and MR imaging. Although carotid US is the most widely performed noninvasive test in the evaluation of a patient suspected of having carotid arterial disease, the modality is highly operator-dependent, and presently is not capable of consistent demonstration of plaque morphology.^[11] In contrast, high-resolution MRI seems to be a promising noninvasive tool to evaluate plaque burden, morphology, and composition. For instance, a protocol generating four contrast-weighted images (TOF, T1W, PDW, T2W; best-in-plane pixel size $0.25 \times 0.25 \text{ mm}^2$) has been used successfully to identify the lipid-rich/necrotic core^[27] and the status of the fibrous cap^[28] (see Fig. 5).

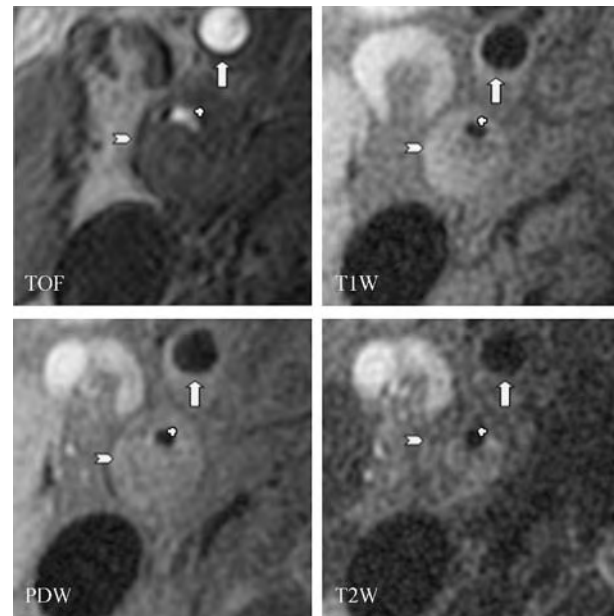


Fig. 5 High-resolution magnetic resonance plaque imaging with four contrast-weighted images from the same location. The arrow shows the normal lumen and wall from the external carotid artery; the cross shows the highly stenosed internal carotid artery; the chevron shows the arterial wall that is massively enlarged.

Interventional MRA

Magnetic resonance imaging—with its superior soft tissue contrast, multiplanar imaging capabilities, and specific vascular applications—is a promising modality in vascular interventions, such as percutaneous transluminal angioplasty, stent implantation, or arterial chemoembolization. It requires high spatial resolution with an optimized real-time imaging.^[29] Numerous technical, hardware, and software developments will be needed before interventional vascular MRA reaches clinical practice.

CONCLUSIONS

Vascular imaging is a dynamic field with a variety of imaging modalities, and each of these modalities has its value in the clinical practice. The current developments are directed at the acquisition of even more detailed intra- and extraluminal information to obtain clinically adequate information in a minimally invasive way. For each clinical indication, it should be considered which image modality is best suited to satisfy the clinical needs in a cost-effective way. Whereas some imaging techniques need only limited training and experience and can be applied almost anywhere, others require specific training and experience and can only be



performed in specialized centers. New developments in soft- and hardware will further improve our understanding of vascular disease and might open new areas for diagnosis, prevention, and treatment.

ACKNOWLEDGMENTS

The authors wish to acknowledge Marina S. Ferguson, Charlene Marino-Clark, and Stefan Keel for preparing the manuscript, and Lee M. Mitsumori for his help in acquiring the images.

ARTICLES OF FURTHER INTEREST

Intravascular Ultrasound; Medical Imaging, 3-D; Microcomputed Tomography and Its Applications; Ultrasound and Tissue Interaction; Ultrasound Doppler

REFERENCES

1. Cochran, S.T.; Bomyea, K.; Sayre, J.W. Trends in adverse events after IV administration of contrast media. *AJR Am. J. Roentgenol.* **2001**, *176*, 1385–1388.
2. Mistretta, C.A. Relative characteristics of MR angiography and competing vascular imaging modalities. *J. Magn. Reson. Imaging* **1993**, *3*, 685–698.
3. Valji, K. Standard Angiographic and Interventional Techniques. In *Vascular and Interventional Radiology*, 1st Ed.; W.B. Saunders Company: Philadelphia, 1999; 12–38.
4. Nissen, S.E.; Yock, P. Intravascular ultrasound: Novel pathophysiological insights and current clinical applications. *Circulation* **2001**, *103*, 604–616.
5. Waller, B.F. “Crackers, breakers, stretchers, drillers, scrapers, shavers, burners, welders and melters”—The future treatment of atherosclerotic coronary artery disease? A clinical-morphologic assessment. *J. Am. Coll. Cardiol.* **1989**, *13*, 969–987.
6. Barnes, J.E. Characteristics and control of contrast in CT. *Radiographics* **1992**, *12*, 825–837.
7. Fuchs, T.; Kachelriess, M.; Kalender, W.A. Technical advances in multi-slice spiral CT. *Eur. J. Radiol.* **2000**, *36*, 69–73.
8. Gareth, W.L.; Phillips, M.R.C.P. Review of arterial ultrasound. *World J. Surg.* **2000**, *24*, 232–240.
9. Stewart, J.H.; Grubb, M. Understanding vascular ultrasonography. *Mayo Clin. Proc.* **1992**, *67*, 1186–1196.
10. Nederkoorn, P.J.; van der, G.Y.; Hunink, M.G. Duplex ultrasound and magnetic resonance angiography compared with digital subtraction angiography in carotid artery stenosis: A systematic review. *Stroke* **2003**, *34*, 1324–1332.
11. Dogan, A.; Dempsey, R.J. Diagnostic modalities for carotid artery disease. *Neurosurg. Clin. N. Am.* **2000**, *11*, 205–220.
12. Finn, J.P.; Edelman, R.R. Black-blood and segmented k-space magnetic resonance angiography. *Magn. Reson. Imaging Clin. N. Am.* **1993**, *1*, 349–357.
13. Owen, R.S.; Carpenter, J.P.; Baum, R.A.; Perloff, L.J.; Cope, C. Magnetic resonance imaging of angiographically occult runoff vessels in peripheral arterial occlusive disease. *N. Engl. J. Med.* **1992**, *326*, 1577–1581.
14. Prince, M.R.; Yucel, E.K.; Kaufman, J.A.; Harrison, D.C.; Geller, S.C. Dynamic gadolinium-enhanced three-dimensional abdominal MR arteriography. *J. Magn. Reson. Imaging* **1993**, *3*, 877–881.
15. Willig, D.S.; Turski, P.A.; Frayne, R.; Graves, V.B.; Korosec, F.R.; Swan, J.S.; Mistretta, C.A.; Grist, T.M. Contrast-enhanced 3D MR DSA of the carotid artery bifurcation: Preliminary study of comparison with unenhanced 2D and 3D time-of-flight MR angiography. *Radiology* **1998**, *208*, 447–451.
16. Goyen, M.; Ruehm, S.G.; Debatin, J.F. MR-angiography: The role of contrast agents. *Eur. J. Radiol.* **2000**, *34*, 247–256.
17. Yuan, C.; Mitsumori, L.M.; Beach, K.W.; Maravilla, K.R. Carotid atherosclerotic plaque: Noninvasive MR characterization and identification of vulnerable lesions. *Radiology* **2001**, *221*, 285–299.
18. Stuber, M.; Botnar, R.M.; Kissinger, K.V.; Manning, W.J. Free-breathing black-blood coronary MR angiography: Initial results. *Radiology* **2001**, *219*, 278–283.
19. Huber, A.; Matzko, M.; Wintersperger, B.J.; Reiser, M. [Reconstruction methods in postprocessing of CT- and MR-angiography of the aorta]. *Radiologe* **2001**, *41*, 689–694.
20. Katzen, B.T. The future of catheter-based angiography: Implications for the vascular interventionalist. *Radiol. Clin. North Am.* **2002**, *40*, 689–692.
21. Cademartiri, F.; van der, L.A.; Luccichenti, G.; Pavone, P.; Krestin, G.P. Parameters affecting bolus geometry in CTA: A review. *J. Comput. Assist. Tomogr.* **2002**, *26*, 598–607.
22. Flohr, T.; Kuttner, A.; Bruder, H.; Stierstorfer, K.; Halliburton, S.S.; Schaller, S.; Ohnesorge, B.M. Performance evaluation of a multi-slice CT system with 16-slice detector and increased gantry rotation speed for isotropic submillimeter imaging of the heart. *Herz* **2003**, *28*, 7–19.
23. Murphy, K.J.; Rubin, J.M. Power Doppler: It’s a good thing. *Semin. Ultrasound CT MR* **1997**, *18*, 13–21.
24. Melany, M.L.; Grant, E.G. Clinical experience with sonographic contrast agents. *Semin. Ultrasound CT MR* **1997**, *18*, 3–12.
25. Knopp, M.V.; Tengg-Kobligh, H.; Floemer, F.; Schoenberg, S.O. Contrast agents for MRA: Future directions. *J. Magn. Reson. Imaging* **1999**, *10*, 314–316.
26. Ambrose, J.A.; Tannenbaum, M.A.; Alexopoulos, D.; Hjerdahl-Monsen, C.E.; Leavy, J.; Weiss, M.; Borrico, S.; Gorlin, R.; Fuster, V. Angiographic progression of coronary artery disease and the development of myocardial infarction. *J. Am. Coll. Cardiol.* **1988**, *12*, 56–62.
27. Yuan, C.; Mitsumori, L.M.; Ferguson, M.S.; Polissar, N.L.; Echelard, D.; Ortiz, G.; Small, R.; Davies, J.W.; Kerwin, W.S.; Hatsukami, T.S. In vivo accuracy of multispectral magnetic resonance imaging for

- identifying lipid-rich necrotic cores and intraplaque hemorrhage in advanced human carotid plaques. *Circulation* **2001**, *104*, 2051–2056.
28. Hatsukami, T.S.; Ross, R.; Polissar, N.L.; Yuan, C. Visualization of fibrous cap thickness and rupture in human atherosclerotic carotid plaque in vivo with high-resolution magnetic resonance imaging. *Circulation* **2000**, *102*, 959–964.
29. Vogl, T.J.; Balzer, J.O.; Mack, M.G.; Bett, G.; Oppelt, A. Hybrid MR interventional imaging system: Combined MR and angiography suites with single interactive table. Feasibility study in vascular liver tumor procedures. *Eur. Radiol.* **2002**, *12*, 1394–1400.

Vascularization of Engineered Constructs

Zerina Lokmic
Rejhan Idrizi
Aurora Messina
Kenneth R. Knight
Wayne Morrison
Geraldine Mitchell

Bernard O'Brien Institute of Microsurgery, and Department of Surgery, St. Vincent's Hospital, Melbourne, Victoria, Australia

INTRODUCTION

Tissue engineering is touted as the new medical alchemy. Using a combination of cells, especially stem cells, and scaffold matrices, it offers a novel approach that mimics embryological regeneration and promises not only a substitute for conventional reconstructive procedures, but for prostheses and organ transplantations. Frenzied hype and hope blur the line between science and commercial opportunism. Failed biotechnology companies have taught the market salutary lessons. At the moment, the only approved laboratory-prepared tissue engineering products are cartilage and relatively thin structures such as skin, which either do not require a blood supply or can be rapidly revascularized upon implantation.

A pivotal problem facing tissue engineers is that three-dimensional tissues require a functional vasculature. All cells must have access to a microcirculatory system perfused by circulating blood, in order to supply sufficient oxygen and nutrients to sustain their survival, proliferation, and maintenance of biological functions and to remove wastes. Few studies have investigated the production of a microcirculation in engineered tissues *in vitro*; the bulk of work addressing this issue has been undertaken *in vivo*. The current inability to vascularize large transplantable engineered tissues has resulted in two *in vivo* approaches. One approach is to vascularize constructs by the extrinsic derivation of a microcirculation, resulting from ingrowth of capillaries from the recipient vascular bed.^[1,2] The second approach favored by our Institute is to engineer constructs that develop and grow with their own independent capillary network, i.e., an intrinsic microcirculation that derives from a centrally placed macrovascular arteriovenous (AV) loop^[3,4] or flow-through pedicle^[5] placed in a chamber. An intense angiogenic sprouting occurs around the extraluminal surface of the AV loop or flow-through pedicle, probably in response to its ischemic environment. Depending on the anatomical site of the chamber and the addition of exogenous matrix, cells, or

differentiating agents, the developing product can be manipulated to recreate the full complexity of a tissue that would be appropriate for specific reconstructive or organ replacement purposes. To date, we have been able to produce living flaps of fibrous tissue,^[3,4,6] adipose tissue,^[5] skeletal muscle, and organoids of pancreas and liver (unpublished observations).

In this article we will define the major processes involved in new blood vessel growth in adults and the growth factors that control these processes. The use of these growth factors in tissue engineering applications will also be discussed. Our main aim is to describe the work that has been undertaken *in vitro* and *in vivo* in constructing vascularized tissue constructs. The *in vivo* discussion will explore both extrinsic and intrinsic modes of vascularization.

MECHANISMS OF VESSEL GROWTH

In vivo vascularization in general can proceed via four mechanisms: angiogenesis, vasculogenesis, intussusception, and inosculation, all of which are applicable to tissue-engineered constructs. When a new capillary network is required, a complex, coordinated sequence of events involving a complicated interplay of spatial and temporal signals of both genetic and microenvironmental origin guides the vascular assembly process^[7] and may involve any or several of the four mechanisms listed earlier.

The most commonly recognized process is that of angiogenesis, which involves preexisting capillaries, extracellular matrix (ECM) components, growth factors, and signaling molecules. Angiogenesis plays a vital role in a variety of physiological processes including human development, female reproduction, and wound healing, and in various pathological conditions such as tumor growth, rheumatoid arthritis, and inflammation.^[8] Angiogenesis commences with basement membrane degradation around a preexisting vessel,

often a venule, allowing the migration of endothelial cells from the existing vessel. The endothelial cells proliferate, secrete new extracellular matrix, and form solid capillary sprouts. While the sprouts are developing a lumen, the recruitment of pericytes and synthesis of a basement membrane around the endothelial cell/pericyte complex takes place. This heralds the end of the angiogenic process and formation of a fully functional microvessel.^[9]

Vasculogenesis describes the formation in the embryo of a vascular plexus from the mesoderm by differentiation of endothelial precursor cells or angioblasts, into cell clusters known as blood islands and then into endothelial cells.^[10] Vasculogenesis was previously thought to be restricted to embryogenesis, but in recent years, the existence of endothelial progenitor cells (EPC) in the adult peripheral blood, and their participation in ongoing neovascularization have been reported.^[11] Therefore, rather than being solely due to angiogenesis, adult blood vessel growth is the result of a combination of angiogenesis and vasculogenesis, and a similar range of growth factors controls both processes. In addition, two other forms of blood vessel growth—intussusceptive angiogenesis and inosculation—may contribute to vascularization of a tissue-engineered construct. Intussusceptive angiogenesis involves the division of a formed capillary by repeated insertions of new transcapillary tissue pillars (connective tissue cells and extracellular matrix) that push from outside the capillary wall, forcing it against the opposite capillary wall and dividing a single capillary in two. This process of formation of capillary meshes leads to a rapid expansion of the capillary network.^[12] Inosculation is the process by which the microvasculature of transplanted tissue or a tissue-engineered construct connects to the preexisting vasculature of the recipient site.^[13]

GROWTH FACTORS

A plethora of growth factors acting at different stages of vessel development have been identified to have a vital role in new blood vessel development. Significant vascular growth factors are:

- Vascular endothelial growth factor (VEGF)—a potent endothelial cell mitogen.
- Basic fibroblast growth factor (bFGF or FGF 2)—also a significant stimulator of endothelial and other mesenchymal cell proliferation.
- Angiopoietins 1 and 2 (Ang 1 and 2)—thought to control maturation/stabilization of growing capillaries and continued vascular sprouting or vessel regression depending on VEGF levels.^[14]
- Platelet-derived growth factor beta (PDGF β)—involved in pericyte recruitment.^[14]
- Transforming growth factor beta (TGF β)—a multifunctional growth factor with effects on proliferation, differentiation, adhesion, migration, and extracellular matrix production.^[14]

Although much is known about the roles of growth factors in normal and pathophysiological angiogenesis, their use in tissue engineering has been limited. Only VEGF and FGF 2 and, to a lesser extent, PDGF β and TGF β ^[15] have been used as angiogenic agents in tissue engineering. However, application of a single growth factor may be too simplistic an approach^[16] when endeavoring to initiate, develop, mature, and halt blood vessel growth as required in tissue engineering.

VEGF and Its Utilization in Tissue Engineering

VEGF, also known as vascular permeability factor,^[17] is a significant endothelial cell mitogen that promotes angiogenesis in three-dimensional *in vitro* models, resulting in the formation of capillarylike structures from confluent microvascular endothelial cells.^[18] In 1992, Shweiki et al. demonstrated that the expression of VEGF was upregulated by hypoxia.^[19] The examination of possible mechanisms involved in the regulation of this process resulted in elucidation of the transcriptional and posttranscriptional mechanisms of VEGF mRNA upregulation^[20] with a connection to Hypoxia-inducible factor 1 (HIF-1).

VEGF's potency in promoting vascular permeability, proliferation, and migration of endothelial cells,^[21] as well as its ability to induce the expression of interstitial collagenases,^[22] makes it an attractive target for the enhancement of vascular development in tissue-engineered constructs. Several techniques including direct injection and gene delivery of VEGF to ischemic tissues *in vivo*^[23] have been successful in promoting vascularization. There are significant issues associated with promotion of blood vessel growth using VEGF^[24] including its short half-life^[21] and physiological overproduction, leading to events such as vascular hyperpermeability, edema, and hemangioma formation. Tissue engineering applications involving VEGF have utilized slow-release delivery systems and include incorporation into alginate beads^[25] and porous degradable copolymer scaffolds.^[26] Elcin et al.,^[25] in an *in vitro* and *in vivo* study using VEGF incorporated into calcium alginate beads, found significantly increased angiogenesis in rat epigastric groin fascia at 3–8 weeks implantation. Lu et al.^[27] administered recombinant VEGF by implanting bioartificial muscle tissues (BAM), utilizing genetically modified primary skeletal myoblasts. When implanted subcutaneously in syngeneic mice, a 30-fold increase in CD31 positive capillary cells occurred within the BAM. In ischemic hindlimbs, implantation



of this construct, resulted in a two- to threefold increase in capillary density of neighboring host muscle at one week. Biologically active VEGF has been incorporated into poly(glycolic acid) (PLG) matrices containing human microvascular endothelial cells (HMVECS). VEGF was released in a sustained manner for several weeks when the construct was transplanted into severe combined immunodeficient (SCID) mice. A 260% increase in SCID mouse capillaries invaded the matrix at seven days compared to control, and HMVECS with VEGF produced a 16% increase in human-derived blood vessels at 14 days.^[26] Lee et al.^[28] used alginate hydrogels for localized delivery of VEGF and FGF 2 in subcutaneous tissue of SCID mice. Although both growth factors caused an increase in vascular density compared to control, maximal vascular density occurred with VEGF.

FGF 2 and Its Utilization in Tissue Engineering

Fibroblast growth factors have been implicated not only in angiogenesis, but also in fibroblast and smooth muscle cell proliferation and regulation of several other cytokines. The most important in wound healing and angiogenesis is FGF 2, which is released at wound sites by macrophages and injured endothelial cells^[29] and the extracellular matrix and acts as a potent angiogenic cytokine, leading to endothelial proliferation and subsequently, capillary formation.^[30] FGF 2 has also been implicated in hematopoiesis,^[31] suggesting that this growth factor could be involved in the development and/or regulation of the hemangioblast. Like VEGF, FGF 2 has a short half-life and is also considered to have the potential to cause significant side effects when administered therapeutically.^[24] Most tissue engineering applications have utilized FGF 2 using slow-release delivery techniques.

Subcutaneous plugs of Matrigel[®], a biological mouse basement membrane material, with added FGF 2 are known to significantly promote vascularization, which does not occur in the absence of FGF 2.^[32] Recently, our Institute has incorporated FGF 2 into Matrigel[®] in a vascularized flow-through mouse chamber model with small amounts of autologous fat added. Increased vascularization and adipose tissue growth compared to chambers without FGF 2 were observed (Findlay M., personal communication). Kawai et al.^[33] used artificial dermis with incorporated gelatin microspheres impregnated with FGF 2 inserted into full-thickness skin defects on the back of guinea pigs. Fibroblast proliferation and new capillary formation were increased as the dose of incorporated FGF 2 increased and was more effective than that seen with free FGF 2 at doses over 50 µg. Tabata^[34] has investigated the controlled release of FGF 2 from bioabsorbable

gelatin hydrogel microspheres. Neovascularization around the implantation site occurred when the microspheres were placed subcutaneously in mice. Neovascularization did not occur if the gelatin microspheres were added without FGF 2 or if FGF 2 was injected free. Additionally, not only does this technique promote vascularization, but also, depending on the site of FGF 2 plus gelatin microsphere implantation, can accelerate specific tissue regeneration (nerve and adipose tissue). Rinsch et al.^[35] genetically engineered myoblasts to produce FGF 2 or VEGF and encapsulated them in hollow, microporous polymer membranes, which were evaluated in an in vivo model of acute skin flap ischemia. Only encapsulated myoblasts producing FGF 2 improved flap survival. Recently, Perets et al.^[36] constructed a porous alginate scaffold that incorporates tiny (<10 µm in diameter) poly(lactic-co-glycolic acid) microspheres capable of controlled release of FGF 2. In vivo implantation of this construct in rat peritoneal mesentery produced four times more capillaries than scaffolds without growth factor.

IN VITRO VASCULARIZATION OF TISSUE-ENGINEERED CONSTRUCTS

In Vitro Creation of a Microcirculation

The properties of the individual cellular components of the circulatory system, i.e., macrovascular and microvascular endothelial cells, vascular smooth muscle cells, fibroblasts, and pericytes have been extensively studied in two- and three-dimensional culture systems. More complex coordinated functioning microcirculatory systems have rarely been created in vitro. Presumably this is because in vitro tissue-engineered constructs receive adequate nutrition via the culture system in which they are placed, be that a simple culture dish, bioreactor system growing cells on, or in three-dimensional scaffolds or cell encapsulation systems.^[37]

There are three relatively recent studies that have achieved creation of microvessels within an in vitro tissue engineering environment. The first of these was by Black et al.^[38] who demonstrated that in vitro assembly of microvessels is possible by fabricating a collagen biopolymer in which three human cell types—keratinocytes, dermal fibroblasts, and umbilical vein endothelial cells (HUVECS) were cocultured. Spontaneous formation of capillary-like structures occurred that were positively identified as endothelial by the inclusion of Weibel-Palade bodies within the capillary endothelium. Interconnected capillaries and fluid flow within the capillary-like structures were not demonstrated. Interestingly, it was noted that capillary-like structures only formed in this coculture system when proliferating fibroblasts were present, synthesizing new

extracellular matrix. Frerich et al.^[39] explored in vitro creation of a microcirculation in a soft-tissue model, in which HUVECS and human adipose tissue stromal cells were placed in a fibrin microcarrier scaffold including VEGF. Initially, a significant angiogenic response resulted, but the long-term stabilization of these vessels only occurred when Insulin-like Growth Factor-1 was added to the in vitro construct. In a totally different approach, Kaihara et al.^[40] utilized micromachining technologies on silicon and Pyrex surfaces, generating complex interconnecting trenches similar to that of a microcirculatory system. Hepatocytes and endothelial cells were cultured on these templates and then lifted as single-cell monolayers and folded into compact three-dimensional tissues. The endothelial cells formed branched vascular networks.

In Vitro Creation of Macrovascular Conduits

There is a potential need for macrovascular vessels to supply and drain blood from the microcirculation of tissue-engineered constructs and connect this engineered microcirculation to the systemic circulation, particularly in tissue engineering models that promote intrinsic development of the microcirculation within the construct. Traditionally, autologous vein grafts and more recently, autologous arterial grafts have been used for surgical vascular grafting, as they achieve high long-term patency rates (over 95%). Unfortunately, autologous grafts of the appropriate length and diameter are not always available and additionally, create an undesirable donor site scar.

The need for replacement vascular grafts has prompted research into a variety of biologically-derived or synthetic vascular grafts in the last 50 years. The vast majority have not transferred to the clinical setting due to poor patency rates in experimental in vivo settings, resulting from pathological or pathophysiological processes including aneurysm formation, allograft rejection, calcification, and most commonly, neointimal hyperplasia and/or thrombosis. Dacron and expanded polytetrafluoroethylene (ePTFE) greater than 6 mm internal diameter are the only examples of man-made synthetic vascular grafts that have been used routinely in the clinical setting, particularly as ileo-femoral bypass grafts.^[41]

Utilization of a scaffold, either biological^[42] or synthetic, into which vascular cells can be seeded to create a composite vascular graft has been a predominant focus in recent attempts to manufacture in the laboratory off-the-shelf vascular grafts. The range of biological scaffolds used is enormous and includes the connective tissue framework of intestinal submucosa^[43] and ureter,^[44] decellularized arterial and venous grafts,^[45] and mechanically conditioned tubular collagen gels

containing living smooth muscle cells (SMCs).^[17] Several studies have also combined biological and synthetic scaffold elements.^[46,47]

Initial cell seeding studies of synthetic scaffolds were onto Dacron and ePTFE. In the last few years, in vitro seedings of endothelial and/or vascular smooth muscle cells or fibroblasts have been onto biodegradable scaffolds such as PGA/polyhydroxyalkanoate,^[48] poly(L-lactic acid) (PLLA), and PGA,^[49] chitosan,^[50] and poly(carbonate-urea)urethane.^[51]

Recent in vitro advances are incorporating endothelial precursor cell seeding,^[52] retroviral transfection techniques to alter the properties and enhance the survival and function of seeded endothelial and smooth muscle cells, and adhesion materials to enhance endothelial attachment to engineered scaffolds.

In 1998, L'Heureux et al.^[53] created a completely biological tissue-engineered human blood vessel (TEBV) in vitro. Commencing with an acellular inner membrane wrapped around a tubular support, a cohesive layer of cultured human vascular SMCs followed by a fibroblast layer created the media and adventitia. After eight weeks in vitro, the tubular support was removed and endothelial cells were seeded onto the inner membrane. This TEBV produced extracellular matrix proteins including elastin, vWF, prostacyclin, incorporated acetylated low-density lipoproteins, strongly inhibited platelet adhesion, and had a burst strength of 2000 mmHg. Niklason et al. in 1999^[54] defined the concept of bioreactor production of TEBVs. Cultured bovine aortic SMCs were pipetted onto tubular biodegradable PGA scaffolds in a bioreactor and subjected to pulsatile radial stress for eight weeks prior to bovine aortic endothelial cell seeding. The SMCs produced collagen and elastin and the endothelial cells produced vWF and platelet endothelial cell adhesion molecule (PECAM). Burst strength was within physiological range at 2150 mmHg. These exciting results illustrate the potential of complete in vitro manufacture of functioning arterial-like vascular grafts. However, in both cases, vessel growth was time-consuming (2–3 months) and only very preliminary in vivo experiments have been reported.

Summary of In Vitro Vascularization of Tissue-Engineered Constructs

The number of studies addressing the creation of a microcirculatory system in vitro is limited. In vivo implantation and extensive follow-up studies to determine function, connection to the preexisting vasculature by inosculation, and long-term survival have not been undertaken. Although the tissue-engineered macrovascular blood vessels reviewed here have advanced our knowledge of the optimal in vitro conditions required to



seed, adhere, orient, and promote extracellular matrix production of/by endothelial cells, SMCs, or fibroblasts in a biological or synthetic scaffold, few researchers have placed their vascular constructs *in vivo*. Previous *in vivo* implantations have been limited to short- or medium-term studies in which patency rates fall well short of those achieved by autologous grafts. It is only under rigorous, long-term *in vivo* testing incorporating large numbers of TEBV at a variety of sites and flow conditions that properties such as the biodegradability, burst strength, elasticity,^[55] vasoactivity,^[56] compliance mismatch and consequent neointimal ingrowth,^[55] thrombogenicity, foreign body, allograft or xenograft rejection reactions, and long-term patency of TEBV will be determined.

In addition to the problem of growing a macrovascular conduit or microcirculatory system in concert with other seeded cells, laboratory-based tissue engineering must also address issues associated with placing their constructs *in vivo*. These include connection to the preexisting vasculature, immunological rejection, serum-free and animal-free media for cell culture prior to implantation, differentiation protocols, and possible cancerous transformation of stem cells, especially if genetically manipulated.

To avoid many of the problems of laboratory-based tissue engineering, tissue engineering *in vivo* has many attractions. Using our own body as the conductor and given the appropriate cues, the construct can orchestrate its own endogenous cascade of factors that will sequentially lay down the correct biological and structural matrices, adhesion molecules, growth factors, etc. to attract and differentiate key cells for tissue regeneration.

EXTRINSIC AND INTRINSIC IN VIVO VASCULARIZATION

The ultimate test of any complex, three-dimensional construct is survival and function *in vivo*. For this to occur, engineered tissues require a functional vasculature. This may be derived extrinsically by ingrowth of capillaries from the recipient vascular bed, or intrinsically from a centrally-based macrovascular conduit.

Extrinsic Vascularization

Many *in vivo* tissue engineering applications employ either synthetic or naturally occurring biomaterials to support cells when inserted *in vivo*. Synthetic or naturally occurring biomaterials can be separated into two broad groups: matrices and scaffolds that can be characterized by the presence of open pores, and the hydrogels.^[57] The most common approach to

vascularize these biomaterials is to insert thin constructs into a region where they are surrounded by vasculature. This is typically in a subcutaneous position, under the kidney capsule, or in mesenteric or adipose tissue locations. In addition, angiogenic growth factors are often added to the microenvironment to enhance vascularization.

With respect to synthetic matrices, the tissue response is generally that of encapsulation by a foreign body reaction. The capsule thus induced develops into three layers such that a macrophage layer or foreign body cell is in direct contact with the implant. A dense, relatively avascular collagenous layer (up to 100 μm thick) surrounds this layer and is in turn surrounded by a loose collagenous vascular layer. Studies carried out in the 1990s indicate that it is the microarchitecture, in particular the pore size of these matrices, that determines the nature of the encapsulation and infiltration and the extent and degree of vascular infiltration into the matrix itself. Generally, a small pore size induces production of a thick capsule with little vascular infiltration into the matrix. The capsule thickness and the proximity of capsule vessels to the implant decrease, and the degree of vascular infiltration increases as the matrix pore size increases.^[58] In the late 1990s, Sharkawy et al.^[59] characterized the vessels that encapsulate and infiltrate subcutaneous polyvinyl alcohol (PVA) implants, with respect to density, porosity, and diffusivity and showed that a PVA porosity of 60 μm was superior to both smaller and larger pore sizes. He concluded that “implants with pores large enough for cell and tissue infiltration, but not large enough to allow fibrous deposits (scar) induce optimal vascular induction.” However, Yamamoto et al. showed a significant fibrovascular ingrowth into PVA sponge only at a pore size of 250 μm , and only when incorporated with FGF 2 impregnated gelatine microspheres.^[60] These divergent reports and others not cited here indicate that pore size is important, but that all new scaffolds need to be investigated individually for specific vascularization properties.

The hydrogels do not generally induce a foreign body reaction, but do induce a granulation tissue reaction directly surrounding the gel.^[28] A fibrovascular ingrowth occurs when angiogenic growth factors are added. The most widely studied hydrogels are Matrigel[®], gelatin, and alginate.

In 1992, Passaniti et al. demonstrated an intense angiogenic response was elicited by host tissue when Matrigel[®] and FGF 2 were injected subcutaneously.^[32] Kawaguchi et al.^[61] later showed that this angiogenesis created an environment favorable to migration and differentiation of host adipogenic cells. At about the same time, Tabata's group refined a growth factor release system in which FGF 2 was incorporated into gelatin microspheres.^[34] When injected subcutaneously,

the FGF 2 incorporated into gelatin microspheres induced angiogenesis. When incorporated into Matrigel[®], the microspheres with FGF 2 enhanced both the angiogenesis and adipogenesis previously observed with FGF 2 combined in Matrigel[®] without gelatin microspheres. As an extension of these studies, prevascularization utilizing hydrogels capable of controlled release of FGF 2 from gelatin beads has been recently employed to prepare the site of cell implantation prior to cell seeding. Both pancreatic islets and hepatocytes have subsequently been transferred into prevascularization sites successfully.^[63,64] Fibrotic overgrowth was absent around these constructs. More recently, this group has investigated using the FGF 2 microspheres with a collagen sponge scaffold rather than Matrigel[®], which, being a mouse-derived extracellular matrix, cannot be used in humans.^[65]

There is also a recent trend to incorporate endothelial cells into constructs placed *in vivo*. Some of these studies have been referred to earlier.^[26] Nor et al.^[66] seeded human dermal microvascular endothelial cells onto a 1-mm-thick PLGA scaffold in Matrigel[®] in SCID mice where they differentiated into functional human vessels that anastomosed with the mouse vasculature. Recently Koike et al.^[67] seeded HUVECS and 10T1/2 mesenchymal precursor cells in a three dimensional fibronectintype 1 collagen gel which was placed in a transparent window in a mouse. The HUVECS and 10T1/2 cells initially formed long branching interconnected tubes. Connections to the mouse circulatory system were then formed and the new microcirculatory system became perfused with blood. This tissue engineered microcirculation remained stable for one year *in vivo*.

Prefabrication

A well-known technique in plastic surgery termed prefabrication enables three-dimensional tissue to be endowed with a vasculature and later undergo transfer to another part of the body as living tissue. This involves placing a source of angiogenic tissue, in the form of a vascular pedicle or loop, within the tissue to be transferred. Over a period of weeks, these constructs undergo an intense angiogenic response, which links the underlying blood vessels to the overlying tissue, producing a new vascularized flap, which can then be raised on its own pedicle and transferred to the defect site by standard microsurgical techniques.^[68] The principle of vascularization used in prefabrication has applications in tissue engineering. Of particular relevance is the advantage that the newly formed vascular bed is derived from a major vessel. Thus, the tissue can be said to have its own unique vasculature, loosely resembling the

scenario present in major organs where each is served by a main artery and vein. This feature also permits transfer of the construct.

Intrinsic Vascularization

The rationale used in prefabrication, combined with the findings of angiogenesis emanating from a surgically created arteriovenous (AV) shunt loop created in the rat groin,^[3] led us to create an *in vivo* AV loop chamber model for application in tissue engineering. In a rat and pig model, a microsurgically created blood vessel loop or pedicle is placed in the base of a plastic chamber (Fig. 1). The lid of the chamber is secured, leaving a semisealed chamber of up to 1.5 ml (rat) or 80 ml (pig) around the vessels. A small hole in the rim permits entry and exit of the artery and vein that are continuous with the systemic circulation. In a variation of this model, adapted to accommodate use of small blood vessels in a mouse (Fig. 2), the chamber is composed of a silicon tube that is wrapped around an artery, vein, and nerve, and then sealed at the ends, except where the vessels enter and exit. A semi-sealed chamber of approximately 30 μ l surrounds the vessels.

In all three animal models, in the first few days, a fibrin-rich serum seeps into the chamber with a concurrent infiltration of inflammatory cells (Fig. 3). Angiogenic sprouts are evident between 3–7 days (Fig. 4). In the next few weeks, a vigorous angiogenic process proceeds and the chambers are filled with a remarkably dense plexus of both mature and immature vessels embedded in a loose extracellular matrix (Fig. 5). The tissue matures over the ensuing weeks to produce a stable connective tissue flap, which is living on the implanted AV loop.^[4,5] This angiogenesis occurs without added exogenous growth factors in the rat chamber model.^[6] Added ECM in the form of Matrigel[®] enhances the angiogenesis considerably in the mouse model. Altering the chamber microenvironment by addition of scaffolds (Fig. 6), ECM, or growth/differentiation factors modulates the naturally occurring angiogenic process.

Application of this neovasculature to the field of tissue engineering lies in its ability to produce or support vascularized, differentiated tissue products. Using combinations of extracellular matrices with or without an added cell source, these models have been used to produce vascularized skin flaps,^[3] adipose tissue,^[5] and skeletal muscle, bone, as well as pancreatic and liver tissue (Fig. 7). These observations indicate that the spontaneous angiogenesis induced in this model is able to expand with and support added cells to form mature three-dimensional tissue. Recent experiments in the pig have shown that this technique can produce sufficient



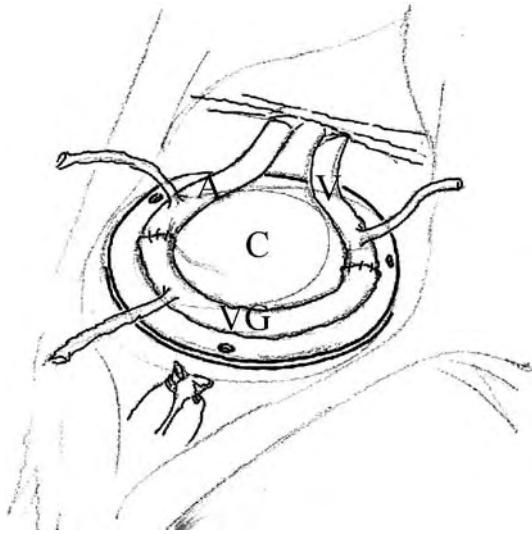


Fig. 1 Diagram of a rat arteriovenous loop, which has been placed in the base of a polycarbonate chamber (C). The femoral artery (A) and vein (V) are anastomosed to a vein graft (VG). Several cauterised branches can be seen projecting outside the chamber area. When the lid of the chamber is fitted over these branches they act as anchors stopping movement of the AV loop. (Illustration by Dr Filip Stillaert.)

three-dimensional, vascularized tissue to potentially 1) grow a standard-sized organ, or 2) repair a human breast that has undergone mastectomy. As with the pre-fabricated grafts, this flap can be transferred micro-surgically, if required.^[3]

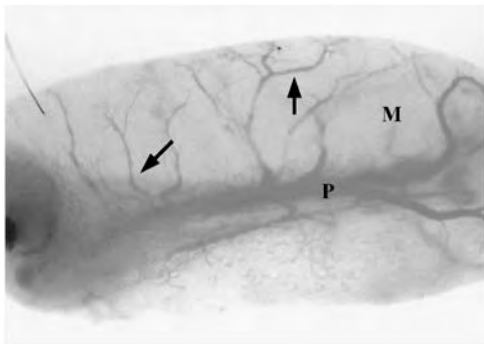


Fig. 2 A whole mount of a six week mouse flow through chamber filled with Matrigel[®]. The silicon chamber has been removed. (P) indicates the epigastric pedicle including an artery and vein which flow through the chamber and which are cleared of all surrounding connective tissue at the time of chamber creation. (M) indicates Matrigel[®] areas and the (arrows) display new blood vessel growth from the pedicle extending to the sides of the chamber. Original figure taken at 2× magnification. (Figure courtesy of Dr. Michael Findlay).

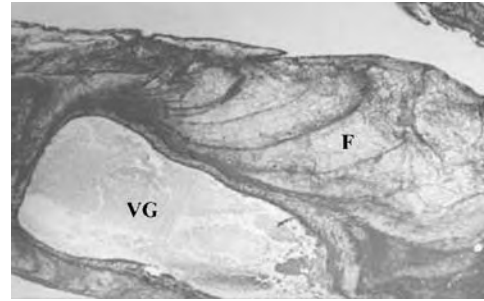


Fig. 3 Section through a rat arteriovenous loop (chamber removed) at 3 days post-insertion illustrating the vein graft (VG) of the loop surrounded by a large fibrin clot (F). Toluidine blue stained 5 μ thick paraffin section. Original figure taken at 2× magnification.

CONCLUSIONS

Although the creation of macrovascular conduits in vitro has been extensively studied, little work has been undertaken on the in vitro creation of a microcirculatory system. Even if a microcirculatory system could be readily created in vitro, it will require a minimum of several days for such a system to connect to the systemic circulation when implanted. It would appear that vascularization of tissue-engineered constructs is more easily achieved in vivo.

A number of fundamental differences between the extrinsic and intrinsic modes of in vivo vascularization of engineered tissues merit specific mention. The intrinsic model, utilizing a chamber, permits new tissue to be grown as an individual entity on its own vascular pedicle. This is an ideal therapeutic concept when

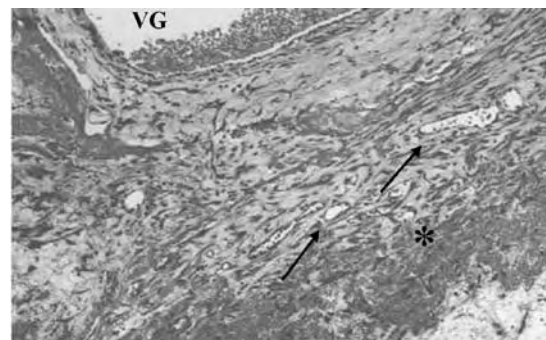


Fig. 4 Section through a rat arteriovenous loop and surrounding tissue 7 days after insertion. Adjacent to the vein graft (VG) is newly generated vascularized connective tissue. (Arrows) indicate new capillaries. Asterisk (*) indicates proliferative area where migrating endothelial and fibroblast-like cells are moving into the surrounding fibrin clot. Toluidine blue stained 5 μ thick paraffin section. Original figure taken at 10× magnification.

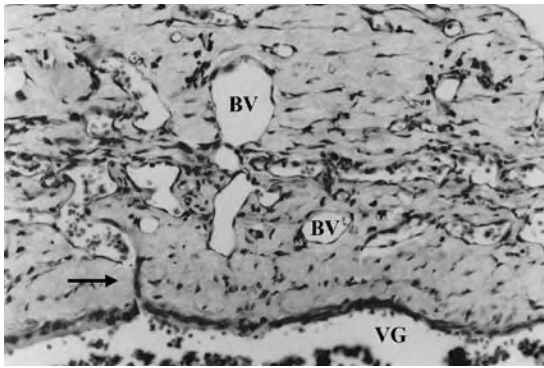


Fig. 5 Section through a rat arteriovenous loop and surrounding tissue 6 weeks after creation of the loop. (VG) indicates part of the vein graft of the loop. (BV) indicates numerous mature capillaries and small venules surrounding the vein graft supported by a collagenous connective tissue. (Arrow) indicates possible connection of capillaries to vein graft lumen. Toluidine blue stained 5 μ thick paraffin section. Original figure taken at 20 \times magnification.

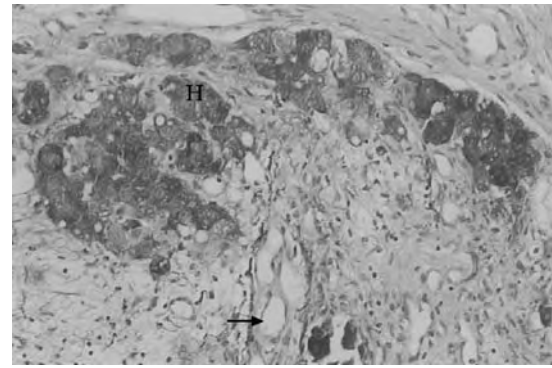


Fig. 7 Section from a rat arteriovenous loop chamber at six weeks. At the time of surgical creation of the chamber rat liver chunks were inserted around the arteriovenous loop. (H) indicates a cluster of surviving PAS positive hepatocytes surrounded by vascular connective tissue. (Arrow) indicates a blood vessel. PAS stained 5 μ thick paraffin section. Original figure taken at 40 \times magnification.

contemplating the generation of de novo organoids in vivo and the capacity for the tissue to grow over time to meet increasing demands. The use of a chamber allows the creation of a naïve space around the vessels into which the new tissue can grow unhindered. It could be expected that this mode of generating vessels might be more able to support the growth of tissues with a complex structure and high metabolic demand. The chamber-grown constructs are mobile and retrievable by virtue of their containment in a chamber and a microsurgery-friendly vascular pedicle, enabling

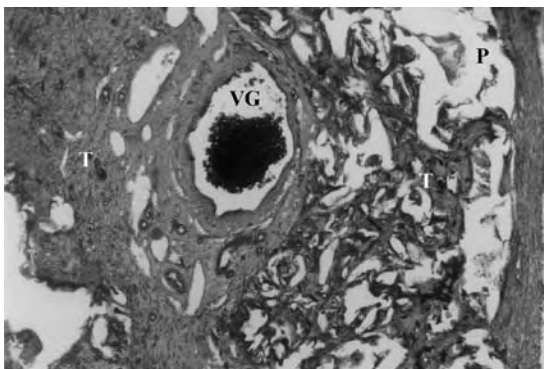


Fig. 6 Section through a rat arteriovenous loop 2 weeks following creation of the loop. This AV loop was inserted in a chamber surrounded by PLGA, which does not stain after processing for paraffin embedding and appears as empty spaces (P) in the figure. (VG) indicates a section through the vein graft of the arteriovenous loop. (T) indicates newly generated vascularized tissue derived from the loop, which has grown around the PLGA and through its pores. Haematoxylin and eosin stained 5 μ thick paraffin section. Original figure taken at 4 \times magnification.

transfer or transplantation with immediate reanastomoses to the circulatory system. In contrast, the source of extrinsic vascularization is diffuse and generally immobilizes the tissue construct at the implant location.

Both intrinsic and extrinsic models have the apparent disadvantage that it takes at least several days for new capillaries to grow into the construct. Undoubtedly, seeded cells die during this hypoxic period, but researchers using both models have demonstrated survival of seeded cells that must derive their nutrition via diffusion during this early period. Only Wang et al.^[63] and Balamurugan et al.^[64] have been able to overcome this problem, utilizing a prevascularization technique.

Both intrinsic and extrinsic vascularization require diffusion to provide nutrition in the first few days. This factor dictates that only small pieces of tissue or tissue and scaffold, or relatively small numbers of stem cells in scaffolds should be implanted. Implanted tissues or cells must grow in parallel with their supplying vasculature. It is obvious at our current state of development that tissue engineering of any body part or organ will take many weeks or months to create the desired tissue.

ACKNOWLEDGMENTS

The authors would like to thank Mr. Jason Palmer for his assistance in the final compilation of this manuscript.

ARTICLES OF FURTHER INTEREST

Tissue Engineering; Tissue Engineering of Microvascular Networks; Vascular Grafts; Vascular Grafts: Development Strategies; Vascular Grafts: Host Interactions



REFERENCES

- Patrick, C.W.; Chauvin, P.B.; Hobbey, J.; Reece, G.P. Preadipocyte seeded PLGA scaffolds for adipose tissue engineering. *Tissue Eng.* **1999**, *5* (2), 139–151.
- Ochoa, E.R.; Vacanti, J.P. An overview of the pathology and approaches to tissue engineering. *Ann. N.Y. Acad. Sci.* **2002**, *979*, 10–26. Discussion 35–38.
- Tanaka, Y.; Tsutsumi, A.; Crowe, D.M.; Tajima, S.; Morrison, W.A. Generation of an autologous tissue (matrix) flap by combining an arteriovenous shunt loop with artificial skin in rats: Preliminary report. *Br. J. Plast. Surg.* **2000**, *53* (1), 51–57.
- Mian, R.; Morrison, W.A.; Hurley, J.V.; Penington, A.J.; Romeo, R.; Tanaka, Y.; Knight, K.R. Formation of new tissue from an arteriovenous loop in the absence of added extracellular matrix. *Tissue Eng.* **2000**, *6* (6), 595–603.
- Cronin, K.J.; Messina, A.; Knight, K.R.; Penington, A.J.; Morrison, W.A. A new murine model of spontaneous autologous tissue engineering by combining arteriovenous pedicle with matrix materials. *Plast. Reconstr. Surg.* **2004**, *113*, 260–269.
- Cassell, O.C.; Morrison, W.A.; Messina, A.; Penington, A.J.; Thompson, E.W.; Stevens, G.W.; Perera, I.M.; Kleinman, H.K.; Hurley, J.V.; Romeo, R.; Knight, K.R. The influence of extracellular matrix on the generation of vascularized, engineered, transplantable tissue. *Ann. N.Y. Acad. Sci.* **2001**, *944*, 429–442.
- Hirschi, K.K.; Skalak, T.C.; Peirce, S.M.; Little, C.D. Vascular assembly in natural and engineered tissues. *Ann. N.Y. Acad. Sci.* **2002**, *961*, 223–242.
- Folkman, J. Angiogenesis in cancer, vascular, rheumatoid and other disease. *Nat. Med.* **1995**, *1* (1), 27–31.
- Risau, W. Mechanisms of angiogenesis. *Nature* **1997**, *386* (6626), 671–674.
- Larsen, W.J. Development of the vasculature. In *Human Embryology*, 3rd Ed.; Scott, W.J., Ed.; Churchill Livingstone: London, 2001; 195–234.
- Asahara, T.; Murohara, T.; Sullivan, A.; Silver, M.; van der Zee, R.; Li, T.; Witzenbichler, B.; Schatteman, G.; Isner, J.M. Isolation of putative progenitor endothelial cells for angiogenesis. *Science* **1997**, *275* (5302), 964–967.
- Burri, P.H.; Djonov, V. Intussusceptive angiogenesis—the alternative to capillary sprouting. *Mol. Aspects Med.* **2002**, *23* (Suppl. 6), 1–27.
- Converse, J.M.; Smahel, J.; Ballantyne, D.L., Jr.; Harper, A.D. Inosculation of vessels of skin graft and host bed: A fortuitous encounter. *Br. J. Plast. Surg.* **1975**, *28* (4), 274–282.
- Darland, D.C.; D'Amore, P.A. Blood vessel maturation: Vascular development comes of age. *J. Clin. Invest.* **1999**, *103* (2), 157–158.
- Stegemann, J.P.; Nerem, R.M. Phenotype modulation in vascular tissue engineering using biochemical and mechanical stimulation. *Ann. Biomed. Eng.* **2003**, *31* (4), 391–402.
- Cao, R.; Brakenhielm, E.; Pawliuk, R.; Wariaro, D.; Post, M.J.; Wahlberg, E.; Leboulch, P.; Cao, Y. Angiogenic synergism, vascular stability and improvement of hind-limb ischemia by a combination of PDGF-BB and FGF-2. *Nat. Med.* **2003**, *9* (5), 604–613.
- Dvorak, H.F.; Dvorak, A.M.; Manseau, E.J.; Wiberg, L.; Churchill, W.H. Fibrin gel investment associated with line 1 and line 10 solid tumor growth, angiogenesis, and fibroplasia in guinea pigs. Role of cellular immunity, myofibroblasts, microvascular damage, and infarction in line 1 tumor regression. *J. Natl. Cancer Inst.* **1979**, *62* (6), 1459–1472.
- Pepper, M.S.; Ferrara, N.; Orci, L.; Montesano, R. Potent synergism between vascular endothelial growth factor and basic fibroblast growth factor in the induction of angiogenesis in vitro. *Biochem. Biophys. Res. Commun.* **1992**, *189* (2), 824–831.
- Shweiki, D.; Itin, A.; Soffer, D.; Keshet, E. Vascular endothelial growth factor induced by hypoxia may mediate hypoxia-initiated angiogenesis. *Nature* **1992**, *359* (6398), 843–845.
- Shima, D.T.; Deutsch, U.; D'Amore, P.A. Hypoxic induction of vascular endothelial growth factor (VEGF) in human epithelial cells is mediated by increases in mRNA stability. *FEBS Lett.* **1995**, *370* (3), 203–208.
- Ferrara, N.; Houck, K.; Jakeman, L.; Leung, D.W. Molecular and biological properties of the vascular endothelial growth factor family of proteins. *Endocr. Rev.* **1992**, *13* (1), 18–32.
- Unemori, E.N.; Ferrara, N.; Bauer, E.A.; Amento, E.P. Vascular endothelial growth factor induces interstitial collagenase expression in human endothelial cells. *J. Cell. Physiol.* **1992**, *153* (3), 557–562.
- Iwaguro, H.; Yamaguchi, J.; Kalka, C.; Murasawa, S.; Masuda, H.; Hayashi, S.; Silver, M.; Li, T.; Isner, J.M.; Asahara, T. Endothelial progenitor cell vascular endothelial growth factor gene transfer for vascular regeneration. *Circulation* **2002**, *105* (6), 732–738.
- Epstein, S.E.; Kornowski, R.; Fuchs, S.; Dvorak, H.F. Angiogenesis therapy: Amidst the hype, the neglected potential for serious side effects. *Circulation* **2001**, *104* (1), 115–119.
- Elcin, Y.M.; Dixit, V.; Gitnick, G. Extensive in vivo angiogenesis following controlled release of human vascular endothelial cell growth factor: Implications for tissue engineering and wound healing. *Artif. Organs* **2001**, *25* (7), 558–565.
- Peters, M.C.; Polverini, P.J.; Mooney, D.J. Engineering vascular networks in porous polymer matrices. *J. Biomed. Mater. Res.* **2002**, *60* (4), 668–678.
- Lu, Y.; Shansky, J.; Del Totto, M.; Ferland, P.; Wang, X.; Vandenberg, H. Recombinant vascular endothelial growth factor secreted from tissue-engineered bioartificial muscles promotes localized angiogenesis. *Circulation* **2001**, *104* (5), 594–599.
- Lee, K.Y.; Peters, M.C.; Mooney, D.J. Comparison of vascular endothelial growth factor and basic fibroblast growth factor on angiogenesis in SCID mice. *J. Control. Release* **2003**, *87* (1–3), 49–56.
- Nicosia, R.F.; Ottinetti, A. Modulation of microvascular growth and morphogenesis by reconstituted basement membrane gel in three-dimensional cultures of rat aorta: A comparative study of angiogenesis in matrigel,

- collagen, fibrin, and plasma clot. *In Vitro Cell. Dev. Biol.* **1990**, *26* (2), 119–128.
30. Schweigener, L.; Neufeld, G.; Friedman, J.; Abraham, J.A.; Fiddes, J.C.; Gospodarowicz, D. Capillary endothelial cells express basic fibroblast growth factor, a mitogen that promotes their own growth. *Nature* **1987**, *325* (6101), 257–259.
 31. Bikfalvi, A.; Han, Z.C. Angiogenic factors are hematopoietic growth factors and vice versa. *Leukemia* **1994**, *8* (3), 523–529.
 32. Passaniti, A.; Taylor, R.M.; Pili, R.; Guo, Y.; Long, P.V.; Haney, J.A.; Pauly, R.R.; Grant, D.S.; Martin, G.R. A simple, quantitative method for assessing angiogenesis and antiangiogenic agents using reconstituted basement membrane, heparin, and fibroblast growth factor. *Lab. Invest.* **1992**, *67* (4), 519–528.
 33. Kawai, K.; Suzuki, S.; Tabata, Y.; Ikada, Y.; Nishimura, Y. Accelerated tissue regeneration through incorporation of basic fibroblast growth factor-impregnated gelatin microspheres into artificial dermis. *Biomaterials* **2000**, *21* (5), 489–499.
 34. Tabata, Y. The importance of drug delivery systems in tissue engineering. *Pharm. Sci. Technol. Today* **2000**, *3* (3), 80–89.
 35. Rinsch, C.; Quinodoz, P.; Pittet, B.; Alizadeh, N.; Baetens, D.; Montandon, D.; Aebischer, P.; Pepper, M.S. Delivery of FGF-2 but not VEGF by encapsulated genetically engineered myoblasts improves survival and vascularization in a model of acute skin flap ischemia. *Gene Ther.* **2001**, *8* (7), 523–533.
 36. Perets, A.; Baruch, Y.; Weisbuch, F.; Shoshany, G.; Neufeld, G.; Cohen, S. Enhancing the vascularization of three-dimensional porous alginate scaffolds by incorporating controlled release basic fibroblast growth factor microspheres. *J. Biomed. Mater. Res.* **2003**, *65A* (4), 489–497.
 37. Godbey, W.T.; Atala, A. In vitro systems for tissue engineering. *Ann. N.Y. Acad. Sci.* **2002**, *961*, 10–26.
 38. Black, A.F.; Berthod, F.; L'Heureux, N.; Germain, L.; Auger, F.A. In vitro reconstruction of a human capillary-like network in a tissue-engineered skin equivalent. *Faseb J.* **1998**, *12* (13), 1331–1340.
 39. Frerich, B.; Lindemann, N.; Kurtz-Hoffmann, J.; Oertel, K. In vitro model of a vascular stroma for the engineering of vascularized tissues. *Int. J. Oral Maxillofac. Surg.* **2001**, *30* (5), 414–420.
 40. Kaihara, S.; Borenstein, J.; Koka, R.; Lalan, S.; Ochoa, E.R.; Ravens, M.; Pien, H.; Cunningham, B.; Vacanti, J.P. Silicon micromachining to tissue engineer branched vascular channels for liver fabrication. *Tissue Eng.* **2000**, *6* (2), 105–117.
 41. Bos, G.W.; Poot, A.A.; Beugeling, T.; van Aken, W.G.; Feijen, J. Small-diameter vascular graft prostheses: Current status. *Arch. Physiol. Biochem.* **1998**, *106* (2), 100–115.
 42. Schmidt, C.E.; Baier, J.M. Acellular vascular tissues: Natural biomaterials for tissue repair and tissue engineering. *Biomaterials* **2000**, *21* (22), 2215–2231.
 43. Huynh, T.; Abraham, G.; Murray, J.; Brockbank, K.; Hagen, P.O.; Sullivan, S. Remodeling of an acellular collagen graft into a physiologically responsive neovessel. *Nat. Biotechnol.* **1999**, *17* (11), 1083–1086.
 44. Field, P.L. The chemically treated bovine ureter—clinical performance of a novel biological vascular prosthesis. *Cardiovasc. Surg.* **2003**, *11* (1), 30–34.
 45. Crowe, D.M.; Mitchell, G.M.; Hurley, J.V.; Olivier, T.V.; O'Loughlin, K.; Morrison, W.A. Cold stored femoral vessels as microvascular allografts: A preliminary study. *Microsurgery* **1994**, *15* (10), 712–716.
 46. Birchall, I.E.; Lee, V.W.; Ketharanathan, V. Retention of endothelium on ovine collagen biomatrix vascular conduits under physiological shear stress. *Biomaterials* **2001**, *22* (23), 3139–3144.
 47. Salacinski, H.J.; Punshon, G.; Krijgsman, B.; Hamilton, G.; Seifalian, A.M. A hybrid compliant vascular graft seeded with microvascular endothelial cells extracted from human omentum. *Artif. Organs* **2001**, *25* (12), 974–982.
 48. Shum-Tim, D.; Stock, U.; Hrkach, J.; Shinoka, T.; Lien, J.; Moses, M.A.; Stamp, A.; Taylor, G.; Moran, A.M.; Landis, W.; Langer, R.; Vacanti, J.P.; Mayer, J.E. Tissue engineering of autologous aorta using a new biodegradable polymer. *Ann. Thorac. Surg.* **1999**, *68* (6), 2298–2304. Discussion 2305.
 49. Chu, C.F.; Lu, A.; Liszkowski, M.; Sipehia, R. Enhanced growth of animal and human endothelial cells on biodegradable polymers. *Biochim. Biophys. Acta* **1999**, *1472* (3), 479–485.
 50. Chupa, J.M.; Foster, A.M.; Sumner, S.R.; Madhally, S.V.; Matthew, H.W. Vascular cell responses to polysaccharide materials: In vitro and in vivo evaluations. *Biomaterials* **2000**, *21* (22), 2315–2322.
 51. Salacinski, H.J.; Tai, N.R.; Punshon, G.; Giudiceandrea, A.; Hamilton, G.; Seifalian, A.M. Optimal endothelialisation of a new compliant poly(carbonate-urea)urethane vascular graft with effect of physiological shear stress. *Eur. J. Vasc. Endovasc Surg.* **2000**, *20* (4), 342–352.
 52. Kaushal, S.; Amiel, G.E.; Guleserian, K.J.; Shapira, O.M.; Perry, T.; Sutherland, F.W.; Rabkin, E.; Moran, A.M.; Schoen, F.J.; Atala, A.; Soker, S.; Bischoff, J.; Mayer, J.E. Functional small-diameter neovessels created using endothelial progenitor cells expanded ex vivo. *Nat. Med.* **2001**, *7* (9), 1035–1040.
 53. L'Heureux, N.; Paquet, S.; Labbe, R.; Germain, L.; Auger, F.A. A completely biological tissue-engineered human blood vessel. *Faseb J.* **1998**, *12* (1), 47–56.
 54. Niklason, L.E.; Gao, J.; Abbott, W.M.; Hirschi, K.K.; Houser, S.; Marini, R.; Langer, R. Functional arteries grown in vitro. *Science* **1999**, *284* (5413), 489–493.
 55. Tranquillo, R.T. The tissue-engineered small-diameter artery. *Ann. N.Y. Acad. Sci.* **2002**, *961*, 251–254.
 56. Nerem, R.M. Tissue engineering a blood vessel substitute: The role of biomechanics. *Yonsei Med. J.* **2000**, *41* (6), 735–739.
 57. Griffith, L.G. Emerging design principles in biomaterials and scaffolds for tissue engineering. *Ann. N.Y. Acad. Sci.* **2002**, *961*, 83–95.
 58. Sieminski, A.L.; Gooch, K.J. Biomaterial-microvasculature interactions. *Biomaterials* **2000**, *21* (22), 2232–2241.
 59. Sharkawy, A.A.; Klitzman, B.; Truskey, G.A.; Reichert, W.M. Engineering the tissue which



- encapsulates subcutaneous implants. II. Plasma-tissue exchange properties. *J. Biomed. Mater. Res.* **1998**, *40* (4), 586–597.
60. Yamamoto, M.T.Y.; Kawasaki, H.; Ikada, Y. Promotion of fibrovascular tissue ingrowth into porous sponges by fibroblast growth factor. *J. Mater. Sci. Mater. Med.* **2000**, *11* (4), 213–218.
 61. Kawaguchi, N.; Toriyama, K.; Nicodemou-Lena, E.; Inou, K.; Torii, S.; Kitagawa, Y. De novo adipogenesis in mice at the site of injection of basement membrane and basic fibroblast growth factor. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95* (3), 1062–1066.
 62. Tabata, Y.; Miyao, M.; Inamoto, T.; Ishii, T.; Hirano, Y.; Yamaoki, Y.; Ikada, Y. De novo formation of adipose tissue by controlled release of basic fibroblast growth factor. *Tissue Eng.* **2000**, *6* (3), 279–289.
 63. Wang, W.; Gu, Y.; Tabata, Y.; Miyamoto, M.; Hori, H.; Nagata, N.; Touma, M.; Balamurugan, A.N.; Kawakami, Y.; Nozawa, M.; Inoue, K. Reversal of diabetes in mice by xenotransplantation of a bioartificial pancreas in a prevascularized subcutaneous site. *Transplantation* **2002**, *73* (1), 122–129.
 64. Balamurugan, A.N.; Gu, Y.; Tabata, Y.; Miyamoto, M.; Cui, W.; Hori, H.; Satake, A.; Nagata, N.; Wang, W.; Inoue, K. Bioartificial pancreas transplantation at prevascularized intermuscular space: Effect of angiogenesis induction on islet survival. *Pancreas* **2003**, *26* (3), 279–285.
 65. Kimura, M.Y.; Ozeki, M.; Inamoto, T.; Tabata, Y. Adipose tissue engineering based on human preadipocytes combined with gelatin microspheres containing basic fibroblast growth factor. *Biomaterials* **2003**, *24* (14), 2513–2521.
 66. Nor, J.E.; Peters, M.C.; Christensen, J.B.; Sutorik, M.M.; Linn, S.; Khan, M.K.; Addison, C.L.; Mooney, D.J.; Polverini, P.J. Engineering and characterization of functional human microvessels in immunodeficient mice. *Lab. Invest.* **2001**, *81* (4), 453–463.
 67. Koike, N.; Fukumura, D.; Gralla, O.; Au, J.; Schechner, J.S.; Jain, R.K. Creation of long-lasting blood vessels. *Nature* **2004**, *428* (March 11), 138–139.
 68. Morrison, W.A.; Dvir, E.; Doi, K.; Hurley, J.V.; Hickey, M.J.; O'Brien, B.M. Prefabrication of thin transferable axial-pattern skin flaps: An experimental study in rabbits. *Br. J. Plast. Surg.* **1990**, *43* (6), 645–654.

Vascular Tissue Engineering

Alejandro Nieponice

Timothy M. Maul

Lorenzo Soletti

David A. Vorp

Departments of Surgery and Bioengineering, McGowan Institute for Regenerative Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.

INTRODUCTION

Coronary artery disease is the leading cause of death in the western world and by-pass grafting is the current therapy of choice to alleviate this disease. The development of a synthetic small diameter graft continues to be a challenge, and vascular tissue engineering appears to be a promising solution. A combination of several factors—scaffold materials, cell source, seeding methods, bioreactors, and in vivo remodeling—is necessary to achieve clinical viability of a tissue-engineered vascular graft (TEVG). This article reviews the current state-of-the-art for these factors and their role in future clinical applications.

CLINICAL NEED FOR VASCULAR TISSUE ENGINEERING

According to the American Heart Association,^[1] almost one million lives are claimed from cardiovascular disease every year in the U.S.A. alone, and more than 50% of those are caused by coronary artery disease—the leading cause of death in the western world. After Favaloro's^[2] first successful implantation of a saphenous vein graft as a coronary artery bypass graft (CABG) in 1968, the number of procedures performed annually in the U.S.A. alone to alleviate this disease has increased to over 500,000. However, almost 30% of vein grafts become occluded in the first year and one half of those occlusions occur within the first month.^[3] Moreover, only 69% of patients can reach a 10–12 yr period free of reoperations or percutaneous transluminal coronary angioplasty.^[4] Many attempts have been made to improve the results obtained with saphenous vein grafts by modifying the harvest methods, preservation, and early antithrombotic therapy. Arterial grafts are a valid alternative for CABG, but they are limited by their short size, small diameter, and availability, as many patients require multiple grafts.^[5] Synthetic grafts have had only limited success,

mainly in larger diameter vessels.^[6] It is clear that new solutions for CABG are needed.

Another related clinical problem is arteriovenous (AV) fistulae failure. The AV fistulae placement is a common procedure for vascular access for hemodialysis. These access grafts often fail within a short period of time, and this is the leading cause of hospitalization for patients suffering from end-stage renal disease. Infection and early thrombosis of synthetic grafts such as those made from expanded polytetrafluoroethylene (ePTFE) and intimal hyperplasia of AV fistulae are the common failure modes.^[7,8]

The clinical shortcomings of current vascular bypass grafts and access grafts have guided new approaches in the development of alternative solutions.^[9] However, a highly successful alternative that can remain patent for an extended period of time has yet to be discovered. A TEVG appears as a promising solution that could meet all requirements needed for that purpose.^[10,11] In the past several years, many groups have investigated different approaches to engineer the biological and biomechanical properties of TEVGs to mimic the native vascular tissue. We review in this article the key components that we believe are required to develop a successful TEVG, including scaffolds, seeding techniques, cell source, in vitro culture with bioreactors, and in vivo remodeling. We also discuss current and future perspectives for the clinical application of TEVGs.

SCAFFOLDS

One of the heaviest areas of investigation in vascular tissue engineering has been the development of an optimal tubular scaffold to foster cell adhesion, integration, and functionality, as well as to provide the necessary mechanical integrity to withstand physiologic pressures and flows.^[10] In 1954, Vorhees was the first to suggest that a synthetic material (Vinyon-N cloth) could be used to repair arterial defects. Since then, several polymeric materials including polyethylene terephthalate (Dacron) and ePTFE have been



utilized successfully in large diameter vascular replacements (>6 mm). One of the earliest approaches that can be considered an attempt to engineer a blood vessel was the *in vitro* endothelial seeding of synthetic grafts to prevent thrombus formation and fibrous deposition in small diameter constructs.^[12] Although success rates were limited, it formed the basis of most approaches in vascular tissue engineering that have relied on some form of natural or synthetic scaffold. To be clinically viable, the design of these scaffolds must meet several requirements including having mechanical properties and material strength similar to that of the replaced native vessel, as well as having sufficient biocompatibility and bioactivity. Regarding mechanical properties, suitable elastance is important to avoid compliance mismatch with the native vessels, which is a common cause of failure of current synthetic grafts.^[13,14] Because strength is critical to withstand systemic pressure, degradation rates of bioerodable scaffolds should be controllable to adapt to the *in vivo* remodeling time. The polymer and its degradation products must also be biocompatible to avoid cytotoxicity or increased inflammatory response and to support cell growth. Desirable bioactive features include low thrombogenicity, selective cell-adhesion, promotion of angiogenesis, control of cell-mediated extracellular matrix (ECM) production, cell proliferation, and incorporation of growth factors.^[15] The materials currently used as scaffolds have varied, ranging from biodegradable synthetic polymers to natural biopolymers such as collagen, elastin, fibrin, glycosaminoglycans, or decellularized grafts.^[16–20] Surface modification of these materials is an expanding research field to foster cell-material interaction.

Synthetic Biodegradable Materials

Synthetic materials have the advantage of being highly reproducible and controllable in terms of mechanical and chemical properties and degradation behavior. Porosity and pore connectivity can be controlled, allowing cell and tissue ingrowth in a short period of time. Mechanical properties can be tuned to mimic native vessels and scale-up would not be a concern for a synthetic product. The main pitfall that this approach faces from a clinical standpoint is the lack of biocompatibility of the polymer and its degradation products.^[21] Although *in vitro* assays indicate good cell survival, the *in vivo* inflammatory response cannot be avoided and that might, in itself, prevent cell survival upon implantation. The most investigated class of synthetic biodegradable biomaterials have been the polyesters poly-L-lactic acid, polyglycolic acid (PGA), polycaprolactone, and related copolymers. Recently, a highly porous biodegradable scaffold and poly(ester

urethane)urea and poly(ether ester urethane)urea have been shown to support cell adhesion and growth, and to provide mechanical and flexural properties suitable for vascular tissue engineering.^[22,23]

Natural Polymers

Natural polymers are desirable because they provide a better environment for cell growth and support, and the inflammatory response to these polymers is significantly reduced. Within this category, three different approaches for vascular tissue engineering can be identified: fashioning of materials derived from natural matrix into a tubular construct, formation of a tubular construct from decellularized ECM materials, or utilizing intact, decellularized blood vessels.

Tubular constructs fabricated from molded matrix/cell suspensions undergo cell-mediated compaction from water release within days of gelation. This phenomenon typically confers improved mechanical properties to the construct and can induce cell alignment in the presence of physical constraints such as a rigid mandrel within the tube.^[24] However, biological gel-based matrices are often limited by reduced mechanical properties and poor reproducibility.^[17,25] Type I collagen cross-linked with enzymes such as transglutaminase^[25] or temporarily supported by a silicone sleeve^[24] have shown potential for vascular tissue engineering applications but suboptimal mechanical properties still remain problematic. Hybrid constructs obtained with mixtures of collagen and fibrin gel or elastin have shown significant improvement in mechanical properties over collagen alone.^[26,27] These approaches are limited by the length of the process—a common roadblock to clinical feasibility. Using a different construction technique, mixtures of collagen and elastin have been electrospun on mandrels to obtain tubular three-layered fibrous constructs that mimic the natural morphology and structure of arteries and have demonstrated good cell integration.^[28]

Another alternative for natural polymers are those derived from decellularized tissue ECM such as small intestinal submucosa or urinary bladder matrix.^[29,30] This approach has been extensively studied in several tissue engineering applications and the ease of processing and biocompatibility have led to a wide acceptance in the clinical community.^[30] During the process of tissue regeneration, the implanted ECM-derived scaffold is gradually degraded until, eventually, only the cells and tissues that have been deposited by the host remain in place of the scaffold. The scaffold consists of a collection of both structural and functional proteins, which fundamentally alters the healing response.^[30] The ECM serves as an inductive scaffold for deposition of organized tissues, rather than

simple scar tissue formation, similar to that which occurs during fetal development. The acellular nature of the scaffold removes the antigenic stimulus that would otherwise cause an adverse immune response.^[31] Although cell loading might be an alternative to enhance regeneration, the nonporous nature of these scaffolds makes cell incorporation *in vitro* difficult. Furthermore, thrombogenic issues in small diameter vascular constructs remain a concern for this type of approach.^[30]

A third type of scaffold for vascular tissue engineering, which has been investigated, is intact decellularized veins and arteries. In this approach, a native vessel is harvested and processed to extract all the cellular components, either mechanically or by means of a detergent solution. The acellular nature of the scaffold eliminates the immune response as in the ECM-derived scaffolds but the mechanical integrity of the original vessel is greatly preserved because collagen and elastin contents remain undisturbed by the decellularizing process.^[18,32] Although different animal species can be used as sources to obtain the grafts, the procedure is labor intensive, and difficult to scale-up. Clinically relevant features for the scaffolds described above are ranked and summarized in Table 1.

Surface Modification

The surface of most scaffolds can be engineered to interact with cells at the molecular level. For example, covalent incorporation of cell adhesion-promoting oligopeptides such as RGD, YIGSR, or REDV on the surface of polymers can increase the steric availability of these ligands to cell-membrane integrin receptors.^[15] The concentration and distribution of adhesive ligands influence the strength of cell attachment, extent of cell spreading, and the speed of cell migration,^[33] and allows for the modulation and control of these parameters. Cell-specific ligands can be used, for example, to foster cell adhesion of targeted cell types in desired locations while inhibiting the adhesion of undesirable cell types.^[34]

Enhancement of adhesion can also be obtained with deposition of self-assembled films. Nanopolyelectrolyte films made of poly(D-lysine) or poly(allylamine hydrochloride) have been shown to increase adhesion of endothelial cells (ECs) in a controllable manner without protein precoating.^[35] Another desirable feature of such films is their ability to incorporate enzymatic recognition sites to give to the scaffold a specific proteolytic sensitivity that can help direct the desired remodeling of the scaffold upon implantation.^[34]

Incorporation of growth factors into the scaffold is another strategy to direct cell attachment and regulate cell proliferation, migration, differentiation, and protein expression. For example, collagen matrices covalently incorporated with heparin and loaded with vascular endothelial growth factor (VEGF) have been shown to significantly enhance angiogenesis in the collagen matrices, thereby increasing EC proliferation.^[36] More recently, polymeric gene delivery has been considered to induce the expression of functional proteins avoiding the cytotoxicity and immunogenicity issues related to the use of viral gene delivery systems.^[37] This approach might allow a controlled overexpression of anticoagulant and anti-inflammatory proteins such as thrombomodulin and/or inhibitors of procoagulant and proinflammatory proteins such as vascular cell adhesion molecule-1 (VCAM-1) and intracellular adhesion molecule-1 (ICAM-1).

SEEDING TECHNIQUES

Cell seeding constitutes a key step in those tissue engineering approaches that incorporate cells into or onto scaffolds for culture or implantation. Desirable features for any seeding technique include preserving cell viability, providing a uniform cell distribution, and attaining a high seeding efficiency, with reduced seeding time. Mechanical forces (e.g., centrifugal forces, shear stress, etc.) are often involved in seeding procedures and can be responsible for force-mediated membrane lysis or triggering of apoptotic pathways.^[38,39] It is therefore important to minimize the

Table 1 Ranking of scaffolds for four key considerations of clinical viability for use in creating a TEVG

	Mechanical properties	Cell seeding	Biocompatibility	Scale-up
Synthetic polymers	Very good	Very good	Fair	Very good
Biological gels (fibrin, collagen)	Fair	Very good	Very good	Fair
Decellularized ECM scaffolds	Good	Poor	Very good	Very good
Decellularized vessels	Very good	Poor	Very good	Fair

Mechanical properties were ranked based on published work. Cell seeding was ranked based on off-the-shelf availability. Biocompatibility was ranked taking into account cell incorporation and the inflammatory reaction triggered after implantation. Scale-up was based on the feasibility to industrialize the construction process.



mechanical stress that cells undergo during the procedure. Uniform seeding of tissue-engineered conduits is desired to achieve homogeneous tissue formation. This is particularly true for the endothelial layer, as failure to accomplish a uniform monolayer can lead to thrombosis, jeopardizing the patency of the construct.

Two common classifications regarding cell seeding exist. One-stage seeding refers to the use of freshly isolated cells, while two-stage seeding refers to cells that are previously expanded *in vitro* to obtain the desired cell number. Another distinction needs to be made between surface and bulk seeding. While surface seeding usually implies coating the luminal layer with cells, bulk seeding attempts to incorporate cells throughout the entire wall thickness. The vast majority of seeding techniques for vascular tissue engineering involve surface seeding. Bulk seeding is harder to accomplish (particularly in tubular constructs) and usually requires an active method of deployment.^[40] Surface seeding techniques on tubular scaffolds typically deliver high-density cell suspensions within the luminal side of the graft and take advantage of sedimentation/rotation, electric fields, or vacuum as driving forces to entrap the cells.^[41–44]

Even though several techniques and devices have been investigated for seeding cells into or onto scaffolds, the literature to date lacks both a complete quantitative analysis of any seeding technique and an accurate predictive model for cell distribution following seeding.

CELL SOURCE

The cellular components of a TEVG are also an important consideration. Cell sourcing is, many believe, the single most rate-limiting step in the development of a TEVG.^[10] A blood vessel comprises three primary cellular components: ECs, smooth muscle cells (SMCs), and fibroblasts. The ECs make up the inner lining of the blood vessel, and are primarily responsible for maintaining hemostasis. They play other roles such as maintaining a selectively permeable barrier to the blood and the underlying SMCs and sensing and responding to various flow environments. For example, they play a role in the initiation of the coagulation and inflammatory cascades by producing factors such as ICAM-1, VCAM-1, platelet-endothelial cell adhesion molecule, thrombomodulin, tissue plasminogen activator, and tissue factor, as well as signaling molecules to the underlying SMCs to alter vascular tone and regulate blood flow such as nitric oxide (NO) and prostaglandins (PGI₂).^[45]

The SMCs, which comprise the media of a blood vessel, are an exquisitely complex continuum of phenotypes, ranging from an active contractile

phenotype to a more passive, matrix-producing phenotype (Fig. 1).

Their phenotype can be described in terms of the specific contractile proteins they produce [e.g., α -smooth muscle actin (α -SMA), SM22 α , calponin, myosin heavy chain (MHC), smoothelin, and caldesmon], ion channels, cell junction proteins (e.g., connexin 40 and connexin 43), adhesion proteins (e.g., metavinculin, and vascular adhesion protein-1), and the types and amounts of various ECM components they produce (e.g., collagen, elastin, and proteoglycans) (Table 2).^[46–48] For in-depth information on these proteins, the reader is directed to an excellent review by Owens.^[49]

The SMC phenotype is predominantly dictated by environmental cues such as mechanical forces, inflammatory molecules, growth factors, and oxygen tension. Thus, not only do SMCs within a single vessel express a range of proteins that determine their function, but also SMCs in different blood vessels express widely different amounts of these proteins coincident with the primary function of that vessel.^[46] Fibroblasts make up most of the adventitial portion of the vessel wall, and mainly act as ECM-secreting cells. Because fibroblasts can be induced to express various markers of SMCs by mechanical and chemical stimulation, it may be possible to think of these cells as SMCs at the completely synthetic end of their spectrum.

In the design of a TEVG, there are two approaches currently available for cell source: terminally differentiated cells and progenitor cells. Both offer certain advantages over the other. The main advantage of using terminally differentiated cells is the possibility of building a TEVG with the same cellular constituents of a native artery and thus potentially increasing the chances of reproducing functionality. However, the use of terminally differentiated cells for the construction of a TEVG presents a number of pitfalls. Owing to the high number of cells required for seeding a TEVG, the isolation may be difficult to accomplish without a highly invasive procedure for the patient, in the case of an autologous approach. Nonautologous grafts, which could be derived from donor cell sources, would require immunosuppression or other intervention owing to rejection issues.

Several investigators have attempted to seed TEVGs with different types of isolated ECs to achieve a functional endothelial layer. Human umbilical vein ECs cells are the most commonly utilized cell type, though bovine and porcine aortic ECs have also been used.^[50–52] Dermal fibroblasts have been used to construct a “self-assembled” TEVG, using extensive cell culture and relying on accelerated ECM production by the cells.^[53,54] SMCs have also been utilized in several approaches to create a TEVG, including seeding on biodegradable scaffolds,^[55] culture in biological

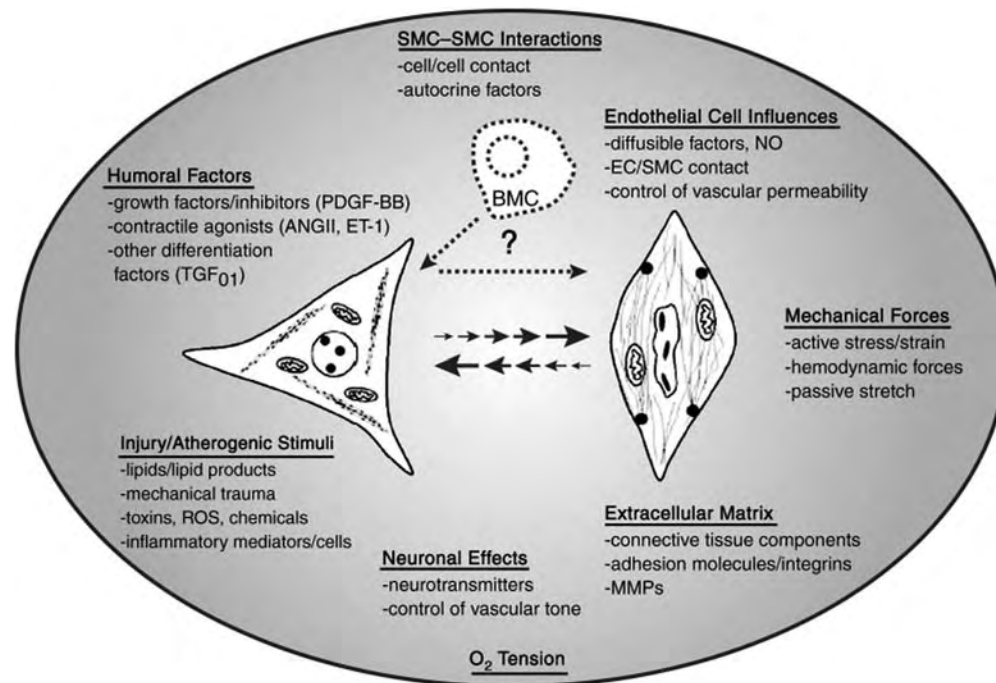


Fig. 1 SMC phenotypes exist as a continuum between the synthetic and contractile phenotypes. The particular position in this continuum is governed mostly by the external environment of the cell, and includes soluble factors (growth factors, hormones, and toxins), mechanical and neuronal stimulation, ECM, oxygen supply, and interaction with ECs. The end result of these signaling milieus will finely tune the response of the SMC to adequately adapt to changes in its environment or respond to injury. Perturbations in these signals are responsible for pathologic processes such as atherosclerosis and intimal thickening through mechanisms that have yet to be elucidated. (From Ref. [46].)

gels,^[17] or as part of cell-based constructs.^[54] Like a native vessel, engineered vessels must possess both a physical and functional component, with the physical component defined by the presence of an organized ECM, and the functional component defined by contractile SMCs.^[16] Dynamic culture of SMCs is thought to be required to avoid a phenotypic change from contractile to synthetic,^[56] though a synthetic phenotype may be a desired feature during TEVG development because it provides greater proliferation and ECM production. Therefore, modulation of SMC phenotype can be a powerful tool in vascular tissue engineering.^[57] Consequentially, it is important to be able to trace the differentiation process that SMCs undergo in different culture conditions,^[47] and this can be achieved by evaluating the levels of expression of contractile proteins, ion channels, cell junction proteins, ECM proteins, and secondary signaling proteins (Table 2).

One of the main concerns with the use of adult autologous SMCs for tissue engineering applications is their limited proliferative capability, which may be related to the enzyme telomerase.^[58] Telomerase is responsible for maintaining the length of DNA after replication.^[59] Its function is greatly diminished in adults, and research investigating the ectopic expression of telomerase has yielded cells capable of

more than 100 population doublings.^[50] Although oncological mapping was performed in these studies, clinical acceptance of this approach to develop a TEVG might be low as telomerase is not constitutively expressed in adult cells and has been shown to be active in most cancer cells.^[60]

Progenitor cells show great potential for use in tissue engineering applications and may circumvent many of the shortcomings associated with other solutions to cell sourcing. Progenitor cells are easier to harvest, and usually display a rapid, almost limitless expansion capability. They can also be viewed as being untouched by the disease state that required intervention. Extensive research has been performed on progenitor cell differentiation in vitro. Bone marrow and embryonic stem cells are the most studied, but embryonic stem cells carry with them an inherent ethical dilemma. The bone marrow contains different types of progenitor cells at various stages of differentiation. Classification of bone marrow progenitor cells (BMPCs) has been performed by specific gravity and the expression of different clusters designation markers, membrane receptors, integrins, and ECM molecules (Table 3).^[61,62]

The main division for bone marrow has been along the marker for CD34, with CD34⁺ cells being considered hematopoietic progenitor cells and the CD34⁻ cells



Table 2 Protein expression in contractile SMCs

Protein	Exclusive to SMCs	Contractile	Synthetic/developmental	Expression in other cells
SMA	N	++	++	Fibroblasts, myoblasts, inducible in ECs
H1-Calponin	Y	++	+/-	
Caldesmon	N	++	+/-	Skeletal myoblasts
Myosin heavy chain (SM-1 and SM-2)	Y	++	--	
SM22 α	Y	++	++	
γ -Vinculin	N	++	++	Cardiac
Metavinculin	N	++	++	Cardiac, skeletal muscle
α -Tropomyosin	N	++	+/-	Embryoid bodies
Vimentin	N	+/-	++	All cell types
Desmin	N	+/-	+/-	Cardiac, skeletal muscle
Angiotensin receptors (AT1 and AT2)	Y	++	Unknown	
Muscarinic receptors (M1, M2, M3)	N	++	Unknown	Neurons, fibroblasts, epithelial, cardiac
Vasopressin receptors (V1, V2)	N	++	Unknown	Neurons, epithelial, hepatocytes
Adrenergic receptors (α and β)	N	++	Unknown	Neurons, epithelial, cardiac

The ability to identify functional, contractile SMCs would be ideal for tissue engineers utilizing progenitor cells to build their constructs. However, of all the markers expressed by SMCs, very few are exclusive to SMCs, and of those that are exclusive to SMCs, only MHC (isotypes SM-1 and SM-2) appears to be limited to the contractile phenotype. The presence of MHC does not necessarily imply functionality. A functional SMC requires abundant K^+ and Ca^{2+} ion channels, as well as the receptors for vasoconstrictors and -dilators. To determine the true contractile/functional state of a presumed SMC, one needs to perform a multitude of protein assays to elucidate the phenotype and suitability for a TEVG. ++ indicates definitively positive expression; +/- indicates transient expression or conflicting reports; and -- indicates definitively negative expression.

being considered mesenchymal progenitor cells. Mesenchymal progenitor cells have high proliferative rates, and can differentiate into many different tissues including bone, fat, tendon, neural, cartilage, and muscle cells.^[63–70] Recently, a small caliber artery was reported

as being created from BMPCs that were differentiated into endothelial and SMCs in vitro.^[18] These grafts remained patent for more than eight weeks as carotid interposition grafts. Bone marrow is also thought to supply distant tissues with progenitor cells that can be

Table 3 The specific antigens, cytokines, cytokine receptors, adhesion molecules, and matrix molecules reported to be expressed or produced by BMPCs

Marker type	Designation
Specific antigens	SH2, SH3, SH4, STRO-1, α -smooth muscle actin
Cytokines and growth factors	IL- 1, 6, 7, 8, 11, 12, 14, and 15, LIF, SCF, Flt-3 ligand, GM-CSF, G-CSF, M-CSF
Cytokine and growth factor receptors	IL-1R, IL-3R, IL-4R, IL-6R, IL-7R, LIFR, SCFR, G-CSFR, IFN- γ R, TNF- α 1R, TNF- α 11R, TGF- β 1R, TGF- β 11R, bFGFR, PDGFR, EGFR
Adhesion molecules	Integrins subunits: α 1, α 2, α 3, α 4, α 5, α v, β 1, β 3, β 4, β 5 Surface adhesion molecules: ICAM-1, ICAM-2, VCAM-1, ALCAM-1, LFA-3, L-selection, endoglin, CD44
ECM	Collagen types I, III, IV, V, and VI, fibronectin, laminin, hyaluronan, proteoglycans

(From Ref. [61].)

committed for a specified purpose.^[63] Although this theory has been repeatedly challenged, it still remains a major focus of research.^[71,72] Furthermore, Reyes et al.^[67] have postulated that the blood transports these progenitor cells to their destinations, and therefore may be a more easily accessible source of progenitor cells. In fact, specific populations of progenitor cells including smooth muscle progenitors have been isolated from peripheral blood.^[73] Other sources of progenitor cells apart from the bone marrow are being investigated and may have a great potential in vascular tissue engineering. Human preadipocytes have been reported to differentiate into smooth muscle cells^[74] and ECs.^[75] Additionally, muscle-derived stem cells are being isolated from skeletal muscle, and these cells exhibit multipotentiality and self-renewal capabilities that may be exploited in vascular tissue engineering.^[76–78] While significant progress has been made in the past decade regarding the controlled differentiation of progenitor cells *in vitro*, further work is necessary to understand how progenitor cells might behave in tissue-engineered constructs.

BIOREACTORS AND IN VITRO DEVELOPMENT OF TISSUE-ENGINEERED BLOOD VESSELS

Because most tissue-engineered constructs lack the complete mechanical and functional properties necessary for a successful *in vivo* implantation immediately following seeding, bioreactors and their design have become a major focus of tissue engineering technology. The goal of a bioreactor is to replicate, as accurately as possible, the biochemical and biophysical stimuli that will ultimately guide the development of the construct into a fully functional tissue capable of *in vivo* implantation. Bioreactors have been utilized for several engineered tissues including cartilage, bone, heart valves, and blood vessels, and have varying levels of complexity and design. Here we will focus on bioreactors for TEVGs, and how biochemical and biophysical stimuli can be used to control the development of an engineered tissue.

The most basic types of bioreactors include spinner flasks and rotational culture of the vessels. These systems enhance tissue culture by improving transport of nutrients and gases. Because blood vessels exist in a dynamic environment, it is important to ensure that TEVGs are mechanically stimulated during *in vitro* development. The mechanical forces that are generally employed include shear stress and radial strain. Kanda, Matsuda, and Oka^[79] were among the first to recognize that mechanical stimulation was necessary for the development of a TEVG with sufficient mechanical properties for *in vivo* implantation. Subsequently, several bioreactor designs have centered

around a silicone tube running through the lumen of a construct to provide a passive means for imparting cyclic stretching—a key force experienced by all layers of a blood vessel.^[16,24] Such mechanical stimulation has been combined with bioactive chemicals such as ascorbic acid, copper sulfate, proline, alanine, and glycine, which are the building blocks of ECM molecules and can enhance matrix cross-linking.^[16,80] Using this approach, accelerated ECM deposition and polymer resorption can be achieved to yield TEVGs with a burst strength equivalent to that of the saphenous vein.^[16] Studies have demonstrated that both the frequency and magnitude of cyclic stretching have effects on tissue inhibitors of matrix metalloproteinase and collagen deposition, and that higher frequency stimulation on the order of fetal or neonatal heart rate (165 bpm) may enhance vessel development.^[81] There is a delicate balance between matrix metalloproteinases (MMPs) (in particular MMP-2) and matrix deposition that occurs under cyclic stretching, and this may be ultimately responsible for the increases seen in the mechanical properties of TEVGs under such stimulation.^[20,82] However, these systems are limited by media diffusion in the luminal surface because of the silicone sleeve. In addition, most systems can only provide radial distension, whereas many arteries in the body also undergo other cyclic deformations, e.g., the coronary arteries.^[83] Therefore, bioreactors that mimic these deformations may be important for TEVGs to be used in such locations.^[84]

The use of perfusion systems to provide stimulation directly to the luminal surface via pressurized culture media as well as the subsequent radial distension has been employed by several researchers to alleviate the nutrient diffusion issues, as well as to provide direct stimulation to an EC layer. Shear stress applied at physiologic arterial levels sustains the expression by ECs of vasodilators [nitric oxide synthase (NOS) and NO], SMC proliferation inhibitors [transforming growth factor (TGF- β), and antithrombotic factors (cyclooxygenase-2, thrombomodulin, and tissue factor plasminogen inhibitors)], while depressing expression of vasoconstrictors (endothelin-1), procoagulation proteins (protease activated receptor-1), and inflammatory proteins (monocyte chemoattractant protein-1).^[45] Shear stress coupled with radial strain may also play a role in increasing cell number and collagen deposition.^[85]

In addition to the potentially beneficial effects of mechanical stimulation, bioreactors impart a means for the delivery of the important growth factors necessary for the proper *in vitro* development of a TEVG such as platelet-derived growth factor (PDGF), TGF- β , basic fibroblastic growth factor (bFGF), and VEGF. Vascular cells, with the exception of ECs that do not generally express PDGF receptors, respond to each of



these molecules through receptor signaling.^[86] The beneficial effects of each of these growth factors on the *in vitro* development of a TEVG are discussed individually below.

Basic fibroblastic growth factor is a strong mitogen and morphogen for most cell types.^[86,87] The mitogenic effects on SMCs and ECs is particularly important for wound healing and angiogenesis, where these cells are needed to grow to fill a defect or void around the injured tissue.^[88] The increase in proliferation consequently causes a decrease in collagen I and III production in SMCs,^[89] but the effects of bFGF in concert with other factors, such as ascorbic acid may prove useful in tissue engineering because their concerted effects induce proliferation *and* matrix production in SMCs.^[90]

PDGF is also a mitogen for SMCs, EC progenitors, and microvascular ECs, but not quiescent ECs. PDGF induces proliferation by allowing cells to pass through G₀ into G₁ phase,^[91] and also has an effect on SMC phenotype.^[49] Several researchers have reported that a specific isoform of PDGF—PDGF-BB—not only induces SMC proliferation, but also suppresses the expression of their contractile machinery^[49,57] through the destabilization of mRNA as it leaves the nucleus.^[92] This phenomenon is also conserved in three-dimensional collagen gels, where PDGF inhibits the response of SMCs to mechanical stimulation.^[57] It has been hypothesized that the decrease in contractile machinery is related to the focus of the cell on proliferation rather than maintenance of phenotype.^[93] However, it has been reported that PDGF is sufficient to induce SMC contractile protein expression in progenitor cells.^[73]

TGF- β is involved in mitogenesis and morphogenesis throughout the body in development and adult tissue remodeling. One of the main effects of TGF- β on vascular tissue is a decrease in proliferation. However, during angiogenesis, TGF- β appears to stimulate growth through overlapping signaling mechanisms related to the current cellular phenotype and ECM composition.^[94] Other effects of TGF- β in the vasculature are increases in matrix, matricellular protein, and integrin proteins to help maintain the vessel wall integrity.^[95] Taken together, the increase in matrix production and MMP expression seems to indicate a role for TGF- β in cell migration and remodeling. TGF- β can also have varying effects on the way the cells respond to other growth factors such as PDGF. For tissue engineering, TGF- β has shown increases in α -SMA and matrix production in collagen gels exposed to both static and dynamic environments.^[57] TGF- β also can induce differentiation of BMPCs through expression of α -SMA.^[93]

VEGF is one of the most studied and characterized vascular growth factors because of its importance

during fetal development, oncogenesis, lymphangiogenesis, and angiogenesis. VEGF is a specific mitogen for ECs and progenitor cells.^[86] Additionally, stimulation of CD34⁺ progenitor cells with VEGF alone has been shown to be sufficient for EC differentiation.^[96,97] This differentiation can be enhanced by utilizing several VEGF isoforms that have specific binding affinities for matrix proteins.^[98] VEGF also stimulates the production of NOS mRNA, ultimately leading to the production of more NO.^[99] In addition to the vasoprotective NO, VEGF also stimulates the production of prostaglandins, such as PGI₂, which can also act to maintain SMCs in a quiescent state, and may prevent such pathologies as intimal hyperplasia.^[100,101]

Although growth factors and biomechanical stimulation can independently alter the behavior of the cells constituting a TEVG, it is their constant interaction that provides the precise results for the maintenance of a healthy phenotype. In addition to these two signals, relative oxygen tension, and other biochemicals such as amino acids, vitamins, and minerals have an affect on cellular behavior and require tight control. Thus, a bioreactor is the ideal means to provide this type of control *in vitro*. The bioreactor, through its design, can impart strict biomechanical regimens by controlling fluid flow that imparts shear stress and radial distension through pressure, and alter the timing and concentration of growth factor delivery. It can also ensure that the proper oxygen tension, amino acids, vitamins, and minerals are present throughout the tissue by various mixing and mass transport processes.

The ability to perfuse multiple constructs in a single system is an attractive scale-up feature necessary for the mass production of engineered tissues in a controlled manner. One such system uses air to move a silicone membrane, displacing media in a pulsatile manner, and has been utilized to simultaneously perfuse four TEVG constructs.^[102] Williams and Wick^[103] have also designed a bioreactor with modular components that can be stacked to provide single or multiple construct development.

Although bioreactors cannot control every facet of TEVG development that would exist in an *in vivo* environment, they can help us bring the raw materials of biology together, refine them, and efficiently produce a tissue-engineered product that can withstand the rigors of the body and provide a key component to regenerative medicine.

IN VIVO REMODELING

One of the hallmarks of tissue engineering is the body's natural ability to restore and regenerate functional tissues over biological substitutes.^[31] In some tissue

engineering approaches, the body works as a “bioreactor” to complete the construction of the desired graft through remodeling. It is important to note that most of the tissue engineering approaches that have been successfully carried into the clinical field have utilized the *in vivo* remodeling as their main concept.^[104]

A key factor for successful *in vivo* remodeling is the proper ECM.^[29,30] The correct architecture of the ECM has been shown to induce changes in the cellular signaling process that yield arranged and functional tissue regeneration as opposed to scar tissue formation. In tissue engineering, biomaterials have been used to replicate the biologic and mechanical function of the native ECM found in tissues in the body.^[31] Several authors, particularly in the vascular field, have successfully utilized the *in vivo* remodeling concept.^[105,106] In one such study, the infrarenal aorta was replaced with an ECM graft in a dog model resulting in functional tissue remodeling and good patency rates after eight years.^[29,30] In another study, a segment of the carotid artery was replaced with a native ECM-based scaffold and was found to have similar histology as a saphenous vein graft.^[107]

Using a different approach, an *in vivo* model was used as a bioreactor for the complete construction of the TEVG by implanting a rod in the peritoneal cavity and retrieving a tubular conduit several weeks later.^[108] Although these approaches rely completely on *in vivo* remodeling, there are several other authors who have utilized an *in vivo* model only to complete the development of the grafts.^[16,18,31,109] Rather than trying to achieve a fully functional and morphological tissue *in vitro*, the aim of the *in vivo* approach is to provide initial mechanical support through a scaffold and stimulate tissue replacement through cell signaling by loading these scaffolds with the desired cell type or growth factors. It is not certain which components of the initial graft are more relevant to guide the proper remodeling. There is enough evidence that the architectural structure of the scaffold has a significant impact on the outcome. Porosity, fiber alignment, and nanostructure have shown relevant implications on cellular integration.^[21,28,110] It is also important to consider the role that the seeded cells play in the remodeling process. Some studies have shown that part of the seeded cells are integrated in the final vessel histology with the desired phenotype and those constructs had a better outcome than the acellular controls.^[109] BMPCs that were tracked using the carboxyfluorescein diacetate marker also expressed endothelial markers and were engrafted within a TEVG eight weeks after implantation. In a separate study, scaffolds seeded with BMPCs and implanted remained patent after eight weeks whereas nonseeded grafts occluded after two weeks.^[18] The specific

mechanisms by which the cells contribute to the remodeling remain unclear, despite years of research. A potential answer to that question may emerge from the studies utilizing cell therapy to restore functionality to damaged or diseased tissues and have already shown some clinical success.^[111]

CLINICAL APPLICATION

To be clinically viable, a TEVG must meet several requirements and successful preclinical testing is only one of them. Many successful studies in animal models require long-term cell culture and expansion preventing them from being available off-the-shelf.^[16,54,108] A long manufacturing process would create a large production cost from an industrial perspective and hamper widespread clinical use. Reproducibility is also important to be able to test and predict clinical outcomes in different patient populations. For that purpose, research should be focused on a manufacturing process where all the steps involved in the construction of the TEVG are user independent. This is a very important consideration when dealing with complex bioreactors that require a high degree of skill for implementation.^[103] If off-the-shelf availability is to be combined with an industrialized approach, it is likely that the successful TEVG will rely on a bioreactor being implemented inside the operating room where all nonsurgical events have to be maintained as simple as possible for safety purposes. Moreover, biological storage and transportation between the health care institution and the manufacturing company would need to address significant requirements for regulatory approval that would hinder commercial feasibility. The existence of so many different approaches to develop a TEVG suggests the difficulty in developing the right combination of the key factors summarized in this review.

Most specialists in the field agree that we are at least a decade away from a complete clinical application of these innovative technologies, but early-stage clinical trials are already being conducted with promising results. Shin'oka and coworkers^[112–114] have successfully implanted a porous P(CL/LA) scaffold reinforced with an external sleeve of nonwoven PGA seeded manually with BMPCs isolated at the same time of the procedure to repair single ventricle anomalies using the total cavopulmonary connection approach. They reported eight patients who were asymptomatic and with no need of reoperation at an average follow-up of 13.4 mo. These encouraging results provide evidence that although widespread clinical applications of vascular tissue engineering may be a long way off, clinical translation of this emerging science is clearly feasible.



CONCLUSIONS

Despite efforts to improve the current therapeutic modalities, available vascular grafts fall short in necessary quantity and quality. Vascular tissue engineering research has undergone great advancement in the last decade with numerous successful approaches in the preclinical setting and more recently with preliminary clinical applications. Cell source, scaffold type, seeding methods, and in vitro bioreactors are factors that are key to achieving a clinically viable TEVG. Future efforts should be focused on finding the right combination of these factors to obtain off-the-shelf availability, reproducibility, and scale-up possibilities to reach a widespread clinical benefit of this evolving technology.

ACKNOWLEDGMENTS

The authors express thanks for the funding from NIH BRP #R01 HL069368-01A1 (to VorpDA), AHA Postdoctoral Fellowship 0525585U (to NieponiceA), and NIH F31 predoctoral fellowship #5F31EB004791-2 (to MaulTM).

ARTICLES OF FURTHER INTEREST

Bioreactors; Tissue Engineering; Tissue Engineering of Blood Vessel; Tissue Engineering of Microvascular Networks

REFERENCES

1. *Heart Disease and Stroke Statistics-2004 Update*; American Heart Association: Dallas, TX, 2004.
2. Favaloro, R.G. Saphenous vein autograft replacement of severe segmental coronary artery occlusion: operative technique. *Ann. Thorac. Surg.* **1968**, *5* (4), 334–339.
3. Fitzgibbon, G.M.; Leach, A.J.; Keon, W.J.; Burton, J.R.; Kalfa, H.P. Coronary bypass graft fate. Angiographic study of 1,179 vein grafts early, one year, and five years after operation. *J. Thorac. Cardiovasc. Surg.* **1986**, *91* (5), 773–778.
4. Weintraub, W.S.; Jones, E.L.; Craver, J.M.; Gyton, R.A. Frequency of repeat coronary bypass or coronary angioplasty after coronary artery bypass surgery using saphenous venous grafts. *Am. J. Cardiol.* **1994**, *73* (2), 103–112.
5. Predel, H.G.; Yang, Z.; von Segesser, L.; Turina, M.; Buhler, F.R.; Luscher, T.F. Implications of pulsatile stretch on growth of saphenous vein and mammary artery smooth muscle. *Lancet* **1992**, *340* (8824), 878–879.
6. Williams, S.K.; Rose, D.G.; Jarrell, B.E. Microvascular endothelial cell seeding of ePTFE vascular grafts: improved patency and stability of the cellular lining. *J. Biomed. Mater. Res.* **1994**, *28* (2), 203–212.
7. Brattich, M. Vascular access thrombosis: etiology and prevention. *ANNA J.* **1999**, *26* (5), 537–540.
8. Johansen, K.; Lyman, D.; Sauvage, L. Biomaterials for hemodialysis access. *Blood. Purif.* **1994**, *12* (1), 73–77.
9. Greisler, H.P. *New Biologic and Synthetic Vascular Prostheses*; R.G. Landes, 1991.
10. Nerem, R.M.; Seliktar, D. Vascular tissue engineering. *Ann. Rev. Biomed. Eng.* **2001**, *3*, 225–243.
11. Vorp, D.; Maul, T.; Nieponice, A. Molecular aspects of vascular tissue engineering. *Front. Biosci.* **2005**, *10*, 768–789.
12. Herring, M. Endothelial seeding blood flow surfaces. In *Vascular Grafting: Clinical Applications and Technique's*; PSG Inc., 1983; 275–314.
13. Abbott, W.M.; Megerman, J.; Hasson, J.E.; L'Italien, G.; Warnock, D.F. Effect of compliance mismatch on vascular graft patency. *J. Vasc. Surg.* **1987**, *5* (2), 376–382.
14. Haruguchi, H.; Teraoka, S. Intimal hyperplasia and hemodynamic factors in arterial bypass and arteriovenous grafts: a review. *J. Artif. Organs* **2003**, *6* (4), 227–235.
15. Hubbell, J.A. Bioactive biomaterials. *Curr. Opin. Biotechnol.* **1999**, *10* (2), 123–129.
16. Niklason, L.E.; Gao, J.; Abbott, W.M.; Hirschi, K.K.; Houser, S.; Marini, R.; Langer, R. Functional arteries grown in vitro. *Science* **1999**, *284* (5413), 489–493.
17. Grassl, E.D.; Oegema, T.R.; Tranquillo, R.T. Fibrin as an alternative biopolymer to type-I collagen for the fabrication of a media equivalent. *J. Biomed. Mater. Res.* **2002**, *60* (4), 607–612.
18. Cho, S.W.; Lim, S.H.; Kim, I.K.; Hong, Y.S.; Kim, S.S.; Yoo, K.J.; Park, H.Y.; Jang, Y.; Chang, B.C.; Choi, C.Y.; Hwang, K.C.; Kim, B.S. Small-diameter blood vessels engineered with bone marrow-derived cells. *Ann. Surg.* **2005**, *241* (3), 506–515.
19. Sandusky, G.E.; Badylak, S.F.; Morff, R.J.; Johnson, W.D.; Lantz, G. Histologic findings after in vivo placement of small intestine submucosal vascular grafts and saphenous vein grafts in the carotid artery in dogs. *Am. J. Pathol.* **1992**, *140* (2), 317–324.
20. Seliktar, D.; Black, R.A.; Vito, R.P.; Nerem, R.M. Dynamic mechanical conditioning of collagen-gel blood vessel constructs induces remodeling in vitro. *Ann. Biomed. Eng.* **2000**, *28* (4), 351–362.
21. Guan, J.; Sacks, M.S.; Beckman, E.J.; Wagner, W.R. Synthesis, characterization, and cytocompatibility of elastomeric, biodegradable poly(ester-urethane)ureas based on poly(caprolactone) and putrescine. *J. Biomed. Mater. Res.* **2002**, *61* (3), 493–503.
22. Guan, J.; Fujimoto, K.L.; Sacks, M.S.; Wagner, W.R. Preparation and characterization of highly porous, biodegradable polyurethane scaffolds for soft tissue applications. *Biomaterials* **2005**, *26* (18), 3961–3971.
23. Nieponice, A.S.L.; Guan, J.; Huard, J.; Wagner, W.R.; Vorp, D.A. Preliminary evaluation of a novel muscle-derived stem cell-based vascular construct. In

- ET2005: Engineering Tissues*; Hilton Head, SC, March 9–13, 2005.
24. Seliktar, D.; Nerem, R.M.; Galis, Z.S. Mechanical strain-stimulated remodeling of tissue engineered blood vessel constructs. *Tissue Eng.* **2003**, *9* (4), 657–666.
 25. Orban, J.M.; Wilson, L.B.; Kofroth, J.A.; El-Kurdi, M.S.; Maul, T.M.; Vorp, D.A. Crosslinking of collagen gels by transglutaminase. *J. Biomed. Mater. Res.* **2004**, *68A* (4), 756–762.
 26. Cummings, C.L.; Gawlitta, D.; Nerem, R.M.; Stegmann, J.P. Properties of engineered vascular constructs made from collagen, fibrin, and collagen-fibrin mixtures. *Biomaterials* **2004**, *25* (17), 3699–3706.
 27. Berglund, J.D.; Nerem, R.M.; Sambanis, A. Incorporation of intact elastin scaffolds in tissue-engineered collagen-based vascular grafts. *Tissue Eng.* **2004**, *10* (9–10), 1526–1535.
 28. Boland, E.D.; Matthews, J.A.; Pawlowski, K.J.; Simpson, D.G.; Wnek, G.E.; Bowlin, G.L. Electrospinning collagen and elastin: preliminary vascular tissue engineering. *Front. Biosci.* **2004**, *9*, 1422–1432.
 29. Badylak, S.F.; Lantz, G.C.; Coffey, A.; Geddes, L.A. Small intestinal submucosa as a large diameter vascular graft in the dog. *J. Surg. Res.* **1989**, *47* (1), 74–80.
 30. Badylak, S.F. Xenogeneic extracellular matrix as a scaffold for tissue reconstruction. *Transpl. Immunol.* **2004**, *12* (3–4), 367–377.
 31. Atala, A. Tissue engineering and regenerative medicine: concepts for clinical application. *Rejuvenation Res.* **2004**, *7* (1), 15–31.
 32. Schaner, P.J.; Martin, N.D.; Tulenko, T.N.; Shapiro, I.M.; Tarola, N.A.; Leichter, R.F.; Carabasi, R.A.; Dimuzio, P.J. Decellularized vein as a potential scaffold for vascular tissue engineering. *J. Vasc. Surg.* **2004**, *40* (1), 146–153.
 33. Palecek, S.P.; Loftus, J.C.; Ginsberg, M.H.; Lauffenburger, D.A.; Horwitz, A.F. Integrin-ligand binding properties govern cell migration speed through cell-substratum adhesiveness. *Nature* **1997**, *385* (6616), 537–540.
 34. Seliktar, D.; Zisch, A.H.; Lutolf, M.P.; Wrana, J.L.; Hubbell, J.A. MMP-2 sensitive, VEGF-bearing bioactive hydrogels for promotion of vascular healing. *J. Biomed. Mater. Res. A* **2004**, *68* (4), 704–716.
 35. Boura, C.; Menu, P.; Payan, E.; Picart, C.; Voegel, J.C.; Muller, S.; Stoltz, J.F. Endothelial cells grown on thin polyelectrolyte multilayered films: an evaluation of a new versatile surface modification. *Biomaterials* **2003**, *24* (20), 3521–3530.
 36. Steffens, G.C.; Yao, C.; Prevel, P.; Markowicz, M.; Schenck, P.; Noah, E.M.; Pallua, N. Modulation of angiogenic potential of collagen matrices by covalent incorporation of heparin and loading with vascular endothelial growth factor. *Tissue Eng.* **2004**, *10* (9–10), 1502–1509.
 37. Lee, M.; Kim, S.W. Polyethylene glycol-conjugated copolymers for plasmid DNA delivery. *Pharm. Res.* **2005**, *22* (1), 1–10.
 38. Carter, J.; Hristova, K.; Harasaki, H.; Smith, W.A. Short exposure time sensitivity of white cells to shear stress. *ASAIO J.* **2003**, *49* (6), 687–691.
 39. Dardik, A.; Chen, L.; Frattini, J.; Asada, H.; Aziz, F.; Kudo, F.A.; Sumpio, B.E. Differential effects of orbital and laminar shear stress on endothelial cells. *J. Vasc. Surg.* **2005**, *41* (5), 869–880.
 40. Soletti L, N.A.; Vorp, D.A. A new seeding device for tissue engineered tubular structures. In *ET2005: Engineering Tissues*; Hilton Head, SC, March 9–13, 2005.
 41. Shiota, T.; He, H.; Yasui, H.; Matsuda, T. Human endothelial progenitor cell-seeded hybrid graft: proliferative and antithrombogenic potentials in vitro and fabrication processing. *Tissue Eng.* **2003**, *9* (1), 127–136.
 42. Mazzucotelli, J.P.; Roudiere, J.L.; Bernex, F.; Bertrand, P.; Leandri, J.; Loisanche, D. A new device for endothelial cell seeding of a small-caliber vascular prosthesis. *Artif. Organs* **1993**, *17* (9), 787–790.
 43. Bowlin, G.L.; Rittgers, S.E. Electrostatic endothelial cell transplantation within small-diameter (<6 mm) vascular prostheses: a prototype apparatus and procedure. *Cell transplant.* **1997**, *6* (6), 631–637.
 44. van Wachem, P.B.; Stronck, J.W.; Koers-Zuideveld, R.; Dijk, F.; Wildevuur, C.R. Vacuum cell seeding: a new method for the fast application of an evenly distributed cell layer on porous vascular grafts. *Biomaterials* **1990**, *11* (8), 602–606.
 45. Chien, S.; Li, S.; Shyy, Y.J. Effects of mechanical forces on signal transduction and gene expression in endothelial cells. *Hypertension* **1998**, *31* (1 Pt 2), 162–169(Online).
 46. Owens, G.K.; Kumar, M.S.; Wamhoff, B.R. Molecular regulation of vascular smooth muscle cell differentiation in development and disease. *Physiol. Rev.* **2004**, *84* (3), 767–801.
 47. Owens, G.K. Molecular control of vascular smooth muscle cell differentiation. *Acta physiol. Scand.* **1998**, *164* (4), 623–635.
 48. Valiunas, V.; Gemel, J.; Brink, P.R.; Beyer, E.C. Gap junction channels formed by coexpressed connexin40 and connexin43. *Am. J. Physiol. Heart Circ. Physiol.* **2001**, *281* (4), H1675–H1689.
 49. Owens, G.K. Regulation of differentiation of vascular smooth muscle cells. *Physiol. Rev.* **1995**, *75* (3), 487–517.
 50. Mckee, J.A.; Banik, S.S.; Boyer, M.J.; Hamad, N.M.; Lawson, J.H.; Niklason, L.E.; Counter, C.M. Human arteries engineered in vitro. *EMBO Rep.* **2003**, *4* (6), 633–638.
 51. Fields, R.C.; Solan, A.; McDonagh, K.T.; Niklason, L.E.; Lawson, J.H. Gene therapy in tissue engineered blood vessels. *Tissue Eng.* **2003**, *9* (6), 1281–1287.
 52. Remy-Zolghadri, M.; Laganier, J.; Oligny, J.F.; Germain, L.; Auger, F.A. Endothelium properties of a tissue engineered blood vessel for small diameter vascular reconstruction. *J. Vasc. Surg.* **2004**, *39* (3), 613–620.
 53. L'Heureux, N.; Germain, L.; Labbe, R.; et al. In vitro construction of a human blood vessel from cultured vascular cells: a morphologic study. *J. Vasc. Surg.* **1993**, *17* (3), 499–509.
 54. L'Heureux, N.; Paquet, S.; Labbe, R.; et al. A completely biological tissue engineered human blood vessel. *FASEB J.* **1998**, *12* (1), 47–56.
 55. Niklason, L.E.; Langer, R.S. Advances in tissue engineering of blood vessels and other tissues. *Transplant. Immunol.* **1997**, *5* (4), 303–306.



56. Williams, B. Mechanical influences on vascular smooth muscle cell function. *J. Hypertens.* **1998**, *16* (12 Pt 2), 1921–1929.
57. Stegemann, J.P.; Nerem, R.M. Phenotype modulation in vascular tissue engineering using biochemical and mechanical stimulation. *Ann. Biomed. Eng.* **2003**, *31* (4), 391–402.
58. Bonin, L.R.; Madden, K.; Shera, K.; Ihle, J.; Matthews, C.; Aziz, S.; Perez-Reyes, N.; McDougall, J.K.; Conroy, S.C. Generation and characterization of human smooth muscle cell lines derived from atherosclerotic plaque. *Arterioscler. Thromb. Vasc. Biol.* **1999**, *19* (3), 575–587.
59. Collins, K.; Mitchell, J.R. Telomerase in the human organism. *Oncogene* **2002**, *21* (4), 564–579.
60. Shay, J.W.; Bacchetti, S. A survey of telomerase activity in human cancer. *Eur. J. Cancer* **1997**, *33* (5), 787–791.
61. Minguell, J.J.; Erices, A.; Conget, P. Mesenchymal stem cells. *Exp. Biol. Med.* **2001**, *226* (6), 507–520.
62. Jackson, K.A.; Majka, S.M.; Wulf, G.G.; Goodell, M.A. Stem cells: a mini review. *J. Cell. Biochem. Suppl.* **2002**, *38*, 1–6.
63. Caplan, A.I. Mesenchymal stem cells. *J. Orthop. Res.* **1991**, *9* (5), 641–650.
64. Aubin, J.E.; Liu, F.; Malaval, L.; Gupta, A.K. Osteoblast and chondroblast differentiation. *Bone* **1995**, *17* (2 suppl), 77S–83S.
65. Jaiswal, N.; Haynesworth, S.E.; Caplan, A.I.; Bruder, S.P. Osteogenic differentiation of purified, culture-expanded human mesenchymal stem cells in vitro. *J. Cell Biochem.* **1997**, *64* (2), 295–312.
66. Pittenger, M.F.; Mackay, A.M.; Beck, S.C.; Jaiswal, R.K.; Douglas, R.; Mosca, J.D.; Moorman, M.A.; Simonetti, D.W.; Craig, S.; Marshak, D.R. Multi-lineage potential of adult human mesenchymal stem cells. *Science* **1999**, *284* (5411), 143–147.
67. Reyes, M.; Dudek, A.; Jahagirdar, B.; Koodie, L.; Marker, P.H.; Verfaillie, C.M. Origin of endothelial progenitors in human postnatal bone marrow. *J. Clin. Invest.* **2002**, *109* (3), 337–346.
68. Reyes, M.; Verfaillie, C.M. Characterization of multipotent adult progenitor cells, a subpopulation of mesenchymal stem cells. *Ann. N. Y. Acad. Sci.* **2001**, *938*, 231–233; Discussion 233–235.
69. Wakitani, S.; Saito, T.; Caplan, A.I. Myogenic cells derived from rat bone marrow mesenchymal stem cells exposed to 5-azacytidine. *Muscle Nerve* **1995**, *18* (12), 1417–1426.
70. Young, R.G.; Butler, D.L.; Weber, W.; Caplan, A.I.; Gordon, S.L.; Fink, D.J. Use of mesenchymal stem cells in a collagen matrix for achilles tendon repair. *J. Orthop. Res.* **1998**, *16* (4), 406–413.
71. Terada, N.; Hamazaki, T.; Oka, M.; Hoki, M.; Matsu-terz, D.M.; Nakano, Y.; Meyer, E.M.; Morel, L.; Petersen, B.E.; Scott, E.W. Bone marrow cells adopt the phenotype of other cells by spontaneous cell fusion. *Nature* **2002**, *416* (6880), 542–545.
72. Heil, M.; Ziegelhoeffer, T.; Mees, B.; Schaper, W. A different outlook on the role of bone marrow stem cells in vascular growth: bone marrow delivers software not hardware. *Circ. Res.* **2004**, *94* (5), 573–574.
73. Simper, D.; Stalboerger, P.G.; Panetta, C.J.; Wang, S.; Caplice, N.M. Smooth muscle progenitor cells in human blood. *Circulation* **2002**, *106* (10), 1199–1204.
74. Abderrahim-Ferkoune, A.; Bezy, O.; Astri-Roques, S.; Elabd, C.; Ailhaud, G.; Amri, E.Z. Transdifferentiation of preadipose cells into smooth muscle-like cells: role of aortic carboxypeptidase-like protein. *Exp. Cell Res.* **2004**, *293* (2), 219–228.
75. Planat-Benard, V.; Silvestre, J.S.; Cousin, B.; Andre, M.; Nibbelink, M.; Tamarat, R.; Clergue, M.; Manneville, C.; Saillan-Barreau, C.; Duriez, M.; Tedgui, A.; Levy, B.; Penicaud, L.; Casteilla, L. Plasticity of human adipose lineage cells toward endothelial cells: physiological and therapeutic perspectives. *Circulation* **2004**, *109* (5), 656–663.
76. Deasy, B.M.; Huard, J. Gene therapy and tissue engineering based on muscle-derived stem cells. *Curr. Opin. Mol. Ther.* **2002**, *4* (4), 382–389.
77. Lee, J.Y.; Qu Petersen, Z.; Cao, B.; Kimura, S.; Jankowski, R.; Cummins, J.; Usas, A.; Gates, C.; Robbins, P.; Wernig, A.; Huard, J. Clonal isolation of muscle-derived cells capable of enhancing muscle regeneration and bone healing. *J. Cell Biol.* **2000**, *150* (5), 1085–1100.
78. Qu-Petersen, Z.; Deasy, B.; Jankowski, R.; Ikezawa, M.; Cummins, J.; Pruchnic, R.; Mytinger, J.; Cao, B.; Gates, C.; Wernig, A.; Huard, J. Identification of a novel population of muscle stem cells in mice: potential for muscle regeneration. *J. Cell Biol.* **2002**, *157* (5), 851–864.
79. Kanda, K.; Matsuda, T.; Oka, T. Mechanical stress induced cellular orientation and phenotypic modulation of 3-d cultured smooth muscle cells. *ASAIO J.* **1993**, *39*, M686–M690.
80. Niklason, L. Progress in the engineering of small caliber arterial prostheses. *Engineering Tissue Growth International Conference & Exposition, Pittsburgh, 2002.*
81. Solan, A.; Mitchell, S.; Moses, M.; Niklason, L. Effect of pulse rate on collagen deposition in the tissue engineered blood vessel. *Tissue Eng.* **2003**, *9* (4), 579–586.
82. Seliktar, D.; Nerem, R.M.; Galis, Z.S. The role of matrix metalloproteinase-2 in the remodeling of cell-seeded vascular constructs subjected to cyclic strain. *Ann. Biomed. Eng.* **2001**, *29* (11), 923–934.
83. Gross, M.F.; Friedman, M.H. Dynamics of coronary artery curvature obtained from biplane cineangiograms. *J. Biomech.* **1998**, *31* (5), 479–484.
84. Vorp, D.A.; Severyn, D.A.; Steed, D.L.; Webster, M.W. A device for the application of cyclic twist and extension on perfused vascular segments. *Am. J. Physiol.* **1996**, *270* (2 Pt 2), H787–H795.
85. Jeong, S.I.; Kim, S.H.; Kim, Y.H.; Jung, Y.; Kwon, J.H.; Kim, B.S.; Lee, Y.M. Manufacture of elastic biodegradable PLCL scaffolds for mechano-active vascular tissue engineering. *J. Biomater. Sci. Polym. Ed.* **2004**, *15* (5), 645–660.
86. Nomi, M.; Atala, A.; Coppi, P.D.; Soker, S. Principals of neovascularization for tissue engineering. *Mol. Asp. Med.* **2002**, *23* (6), 463–483.
87. Basilico, C.; Moscatelli, D. The FGF family of growth factors and oncogenes. *Adv. Cancer Res.* **1992**, *59*, 115–165.

88. Gilbert, S.F. *Developmental Biology*; Sinauer Associates, 1997.
89. Pickering, J.G.; Ford, C.M.; Tang, B.; Chow, L.H. Coordinated effects of fibroblast growth factor-2 on expression of fibrillar collagens, matrix metalloproteinases, and tissue inhibitors of matrix metalloproteinases by human vascular smooth muscle cells. Evidence for repressed collagen production and activated degradative capacity. *Arterioscler. Thromb. Vasc. Biol.* **1997**, *17* (3), 475–482.
90. Fu, P.; Sodian, R.; Luders, C.; Lemke, T.; Kraemer, L.; Hubler, M.; Weng, Y.; Hoerstrup, S.P.; Meyer, R.; Hetzer, R. Effects of basic fibroblast growth factor and transforming growth factor-beta on maturation of human pediatric aortic cell culture for tissue engineering of cardiovascular structures. *ASAIO J.* **2004**, *50* (1), 9–14.
91. Stiles, C.D.; Capone, G.T.; Scher, C.D.; Antoniadis, H.N.; Van-Wyk, J.J.; Pledger, W.J. Dual control of cell growth by somatomedins and platelet-derived growth factor. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76* (3), 1279–1283.
92. Corjay, M.H.; Blank, R.S.; Owens, G.K. Platelet-derived growth factor-induced destabilization of smooth muscle alpha-actin mRNA. *J. Cell Physiol.* **1990**, *145* (3), 391–397.
93. Kinner, B.; Zaleskas, J.M.; Spector, M. Regulation of smooth muscle actin expression and contraction in adult human mesenchymal stem cells. *Exp. Cell Res.* **2002**, *278* (1), 72–83.
94. Iruela-Arispe, M.L.; Sage, E.H. Endothelial cells exhibiting angiogenesis in vitro proliferate in response to TGF-beta 1. *J. Cell Biochem.* **1993**, *52* (4), 414–430.
95. Pepper, M.S. Transforming growth factor-beta: vasculogenesis, angiogenesis, and vessel wall integrity. *Cytokine Growth Factor Rev.* **1997**, *8* (1), 21–43.
96. Eggermann, J.; Kliche, S.; Jarmy, G.; Hoffmann, K.; Mayr Beyrle, U.; Debatin, K.M.; Waltenberger, J.; Beltinger, C. Endothelial progenitor cell culture and differentiation in vitro: a methodological comparison using human umbilical cord blood. *Cardiovasc. Res.* **2003**, *58* (2), 478–486.
97. Gehling, U.M.; Ergun, S.; Schumacher, U.; Wagener, C.; Pantel, K.; Otte, M.; Schuch, G.; Schafhausen, P.; Mende, T.; Kilic, N.; Kluge, K.; Schafer, B.; Hossfeld, D.K.; Fiedler, W. In vitro differentiation of endothelial cells from AC133-positive progenitor cells. *Blood* **2000**, *95* (10), 3106–3112.
98. Wijelath, E.S.; Rahman, S.; Murray, J.; Patel, Y.; Savidge, G.; Sobel, M. Fibronectin promotes VEGF-induced CD34 cell differentiation into endothelial cells. *J. Vasc. Surg.* **2004**, *39* (3), 655–660.
99. Hood, J.D.; Meininger, C.J.; Ziche, M.; et al. VEGF upregulates ECNOS message, protein, and no production in human endothelial cells. *Am. J. Physiol.* **1998**, *274* (3 Pt 2), H1054–H1058.
100. He, H.; Venema, V.J.; Gu, X.; Venema, R.C.; Marrero, M.B.; Caldwell, R.B. Vascular endothelial growth factor signals endothelial cell production of nitric oxide and prostacyclin through FLK-1/KDR activation of C-SRC. *J. Biol. Chem.* **1999**, *274* (35), 25130–25135.
101. Asahara, T.; Bauters, C.; Pastore, C.; Kearney, M.; Rossow, S.; Bunting, S.; Ferrara, N.; Symes, J.F.; Isner, J.M. Local delivery of vascular endothelial growth factor accelerates reendothelialization and attenuates intimal hyperplasia in balloon-injured rat carotid artery. *Circulation* **1995**, *91* (11), 2793–2801.
102. Hoerstrup, S.P.; Zund, G.; Sodian, R.; Schnell, A.M.; Grunenfelder, J.; Turina, M.I. Tissue engineering of small caliber vascular grafts. *Eur. J. Cardiothorac. Surg.* **2001**, *20* (1), 164–169.
103. Williams, C.; Wick, T.M. Perfusion bioreactor for small diameter tissue-engineered arteries. *Tissue Eng.* **2004**, *10* (5–6), 930–941.
104. Shinoka, T.; Matsumura, K.; Hibino, N.; Murata, A.; Kosaka, Y.; Kurosawa, H. Clinical practice of transplantation of regenerated blood vessels using bone marrow cells. *Nippon Naika Gakkai Zasshi* **2003**, *92* (9), 1776–1780.
105. Lantz, G.C.; Badylak, S.F.; Coffey, A.C.; Geddes, L.A.; Blevins, W.E. Small intestinal submucosa as a small-diameter arterial graft in the dog. *J. Invest. Surg.* **1990**, *3* (3), 217–227.
106. Lantz, G.C.; Badylak, S.F.; Coffey, A.C.; Geddes, L.A.; Sandusky, G.E. Small intestinal submucosa as a superior vena cava graft in the dog. *J. Surg. Res.* **1992**, *53* (2), 175–181.
107. Sandusky, G.E.; Lantz, G.C.; Badylak, S.F. Healing comparison of small intestine submucosa and ePTFE grafts in the canine carotid artery. *J. Surg. Res.* **1995**, *58* (4), 415–420.
108. Chue, W.L.; Campbell, G.R.; Caplice, N.; Muhammad, A.; Berry, C.L.; Thomas, A.C.; Bennett, M.B.; Campbell, J.H. Dog peritoneal and pleural cavities as bioreactors to grow autologous vascular grafts. *J. Vasc. Surg.* **2004**, *39* (4), 859–867.
109. Matsumura, G.; Miyagawa-Tomita, S.; Shin-Oka, T.; Ikada, Y.; Kurosawa, H. First evidence that bone marrow cells contribute to the construction of tissue-engineered vascular autografts in vivo. *Circulation* **2003**, *108* (14), 1729–1734.
110. Stankus, J.J.; Guan, J.; Wagner, W.R. Fabrication of biodegradable, elastomeric scaffolds with sub-micron morphologies. *J. Biomed. Mater. Res. A* **2004**, *70* (4), 603–614.
111. Weissberg, P.L.; Qasim, A. Stem cell therapy for myocardial repair. *Heart* **2005**, *91* (5), 696–702.
112. Isomatsu, Y.; Shin'oka, T.; Matsumura, G.; Hibino, N.; Konuma, T.; Nagatsu, M.; Kurosawa, H. Extracardiac total cavopulmonary connection using a tissue-engineered graft. *J. Thorac. Cardiovasc. Surg.* **2003**, *126* (6), 1958–1962.
113. Naito, Y.; Imai, Y.; Shin'oka, T.; Kashiwagi, J.; Aoki, M.; Watanabe, M.; Matsumura, G.; Kosaka, Y.; Konuma, T.; Hibino, N.; Murata, A.; Miyake, T.; Kurosawa, H. Successful clinical application of tissue-engineered graft for extracardiac Fontan operation. *J. Thorac. Cardiovasc. Surg.* **2003**, *125* (2), 419–420.
114. Shin'oka, T.; Imai, Y.; Ikada, Y. Transplantation of a tissue-engineered pulmonary artery. *N. Eng. J. Med.* **2001**, *344* (7), 532–533.



Vitrification

Kelvin G. M. Brockbank

John R. Walsh

Ying C. Song

Michael J. Taylor

Organ Recovery Systems, Charleston, South Carolina, U.S.A.

INTRODUCTION

The urgent and growing need for improved methods of cryopreservation for viable tissue-engineered products has stimulated discussion and debate in the literature regarding the relative merits of traditional freezing methods versus approaches involving ice-free vitrification. Prevention of freezing by vitrification means that the water in a tissue remains liquid during cooling. Vitrification is the solidification of a liquid without crystallization. As cooling proceeds, however, the molecular motions in the liquid permeating the tissue decrease. Eventually, an arrested liquid state known as a glass is achieved. It is this conversion of a liquid into a glass that is called vitrification (derived from *vitri*, the Greek word for glass). A glass is a liquid that is too cold or viscous to flow. A vitrified liquid is essentially a liquid in molecular stasis. Vitrification does not have any of the biologically-damaging effects associated with freezing because no appreciable degradation occurs over time in living matter trapped within a vitreous matrix. Vitrification is potentially applicable to all biological systems.

Vitrification has been shown to provide effective preservation for a number of cells, including monocytes, ova and early embryos, and pancreatic islets.^[2-5] We review here some of the basic physicochemical and biophysics principles of vitrification and progress on the extension of vitrification from single cells and cell aggregates to more complex structured tissues. Cryopreservation by vitrification versus conventional freezing is illustrated using articular cartilage as a model. Despite decades of research, articular cartilage has proved refractory to satisfactory cryopreservation using conventional freezing methods. Therefore, cartilage was selected as an ideal model to test the hypothesis that vitreous cryopreservation, in which the formation of extracellular ice is inhibited or prevented, will result in significantly improved preservation.

Vitrification and freezing (water crystallization) are not mutually exclusive processes; the crystalline phase and vitreous phase often coexist within a system. In fact, during conventional cryopreservation involving

controlled freezing of cells, a part of the system vitrifies. This occurs because during freezing, the concentration of solutes in the unfrozen phase increases progressively until the point is reached when the residual solution is sufficiently concentrated to vitrify in the presence of ice. Conventional cryopreservation techniques are optimized by designing protocols that avoid intracellular freezing (Fig. 1). Under these cooling conditions the cell contents actually vitrify due to the combined processes of dehydration, cooling, and the promotion of vitrification by intracellular macromolecules. However, the term, vitrification, is generally used to refer to a process in which the objective is to attempt to vitrify the whole system from the outset such that any ice formation (intracellular and extracellular) is avoided.^[6-11]

PHYSICOCHEMICAL BASIS OF VITRIFICATION

During the low-temperature preservation of biological systems, events rarely take place under true equilibrium conditions. Many interdependent factors determine whether an aqueous system, such as a biological system, approaches the thermodynamic state of lowest free energy during cooling. Metastability is thus often unavoidable, especially in concentrated systems. Such nonequilibrium states are, however, sufficiently reproducible and permanent to have been described as pseudoequilibrium states, and conversion of such metastable thermodynamic states to more stable forms may be subject to large kinetic barriers. The prevalence of so-called unfreezable or bound water in the vicinity of macromolecules is a prime example, where the expected path of thermodynamic stabilization by way of crystallization is prevented by large kinetic restraints.^[12-14] A clear understanding of the occurrence and effects of metastable states during the cooling of compartmentalized living systems is complicated by the interaction of thermodynamic and kinetic factors. Some basic principles have been established with the aid of aqueous solutions of cryoprotective solutes and other macromolecules that interact with water by

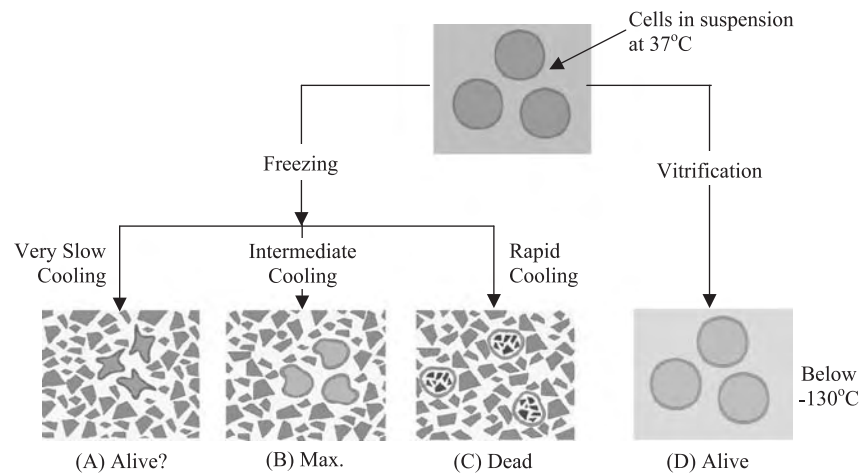


Fig. 1 Comparison of cryopreservation using vitrification or freezing strategies for cells in suspension. If the cells are cryopreserved by freezing (left side), ice forms initially in the extracellular environment and the cells undergo cooling rate-dependent shrinkage due to osmotic dehydration. The slower the cooling rate, the longer intracellular water has the opportunity to move out of the cell by osmosis due to the increasing osmolality of the extracellular environment as water is incorporated into ice crystals. The cells also become concentrated at slower cooling rates as they are pushed together by the forming ice (A,B). Maximum cell viability is usually achieved at an intermediate cooling rate (B) that balances osmotic dehydration and the risk of intracellular ice formation. Rapid cooling (C) permits intracellular ice formation and usually leads to cell death upon rewarming. Very slow cooling (A) may lead to excessive cell dehydration and cell death. In contrast, right side (D), cells cryopreserved by vitrification undergo neither ice formation nor shrinkage due to dehydration and most of the cells should be viable.^[1]

hydrogen bonding.^[15,16] Phase diagrams have proved to be a useful tool in understanding the physico-chemical relationship between temperature, concentration, and change of phase (Fig. 2). For a detailed discussion of the role and interpretation of solid-liquid state diagrams in relation to low-temperature biology, please refer to the review by Taylor.^[14] In particular, supplemented phase diagrams that combine nonequilibrium data on conventional equilibrium phase diagrams serve to depict the important transitions inherent in cooling and warming aqueous solutions of cryoprotective solutes (Fig. 2).

The equilibrium freezing curve, labeled T_m , is often described as the liquidus curve and represents points at which a solution having a particular concentration will freeze (or melt) under equilibrium conditions of temperature change. Hence, this curve represents the phase change boundary for the two-component solution as a function of temperature. Cooling a solution below the liquidus curve will result in ice formation if the conditions are favorable for nucleation, with the result that the remaining liquid phase becomes more concentrated in the solute as defined by the curve. In practice, freezing is rarely initiated at the liquidus point. Solutions tend to undercool to varying degrees before significant nucleation and ice crystal growth occur. Heterogeneous nucleation occurs in water at temperatures above -38.5°C , and it is usually catalyzed by the surfaces of particulate impurities that act

as seeds for crystal growth. Pure samples of water will self-nucleate at the homogeneous nucleation temperature (T_h), -38.5°C .^[14] Both heterogeneous and homogeneous nucleation temperatures decrease with increasing dissolved solute concentration (Fig. 2).

The phase diagram for propanediol (Fig. 2) shows that in the region of 0–35%, freezing will occur at some point 5–20°C below T_m , invariably by heterogeneous nucleation. At sufficiently high concentrations and low temperatures, the kinetics of the process become so slow that T_h is difficult to detect and any nucleated crystals that form in the region of T_h remain microscopic. As temperature is lowered further, molecular motion is slowed to the point where translational and rotational molecular motion is essentially halted and the system is trapped in a high-energy state that resembles a liquidlike configuration, or a vitreous glass.^[17] This glass transition (T_g) is associated with a marked change in physical properties such as specific heat and refractive index and certain mechanical properties such that T_g can be clearly identified. Determination of the transition temperatures that provide data for the construction of supplemented phase diagrams is usually derived from thermograms generated using differential scanning calorimetry (DSC) or the related technique, differential thermal analysis (DTA).^[14] A DSC-derived thermogram for a complex solution of cryoprotectants in water is presented in Fig. 3. The kinetic nature of these transitions



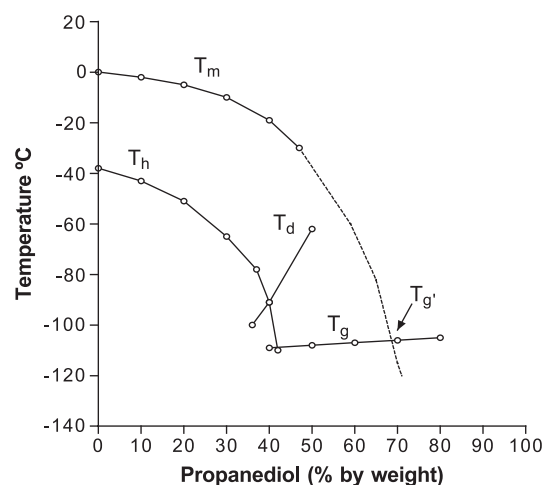


Fig. 2 Supplemented phase diagram for propanediol in water. Data are derived from differential scanning calorimetry thermograms showing the principal events and phase changes associated with cooling and heating. A supplemented phase diagram combines nonequilibrium data on a conventional equilibrium phase diagram and serves to depict the important transitions inherent in cooling and warming aqueous solutions of cryoprotective solutes. (T_m =equilibrium melting point curve, T_h =homogeneous nucleation curve, T_d =devitrification curve, T_g =glass transition curve). (Redrawn from the data of MacFarlane et al. (From Refs. [17,19,20].)

means that T_g has to be defined with reference to a particular set of experimental conditions.

Reference to Fig. 2 shows that in the region of 35–40% propanediol, it is possible to cool samples through the T_h curve without apparent freezing and form what have been referred to as doubly unstable glasses.^[18] This term reflects the high probability that

the vitreous system contains ice nuclei and if warming is not sufficiently rapid, further nucleation and crystallization will occur. This event is known as devitrification and is depicted on the supplemented phase diagram as T_d . Hence, during cooling the sample attains the glassy state, but it invariably contains ice nuclei, the growth of which is arrested along with all other molecular motions in the sample. However, upon rewarming, crystallization can be detected, either visibly or by an exothermic event in a thermogram (Fig. 3), reflecting the growth of ice by devitrification (transition of glassy to crystalline state) and recrystallization (growth of existing ice crystals).^[14] The phenomenon of crystallization on warming a glassy sample to temperatures in the vicinity of T_g is often referred to as devitrification of a doubly unstable glass since it is unstable with respect to both the liquid and solid states.^[18,19] Hence, the process by which a metastable glass, or supercooled liquid obtained by heating the glass above its glass transition temperature, forms the stable crystalline phase is generally referred to as devitrification.^[19]

In the higher concentration range of 41–50% for propanediol (Fig. 2), the T_h curve meets T_g and in this region, it is possible to slowly cool even bulk liquids directly to T_g without experiencing any detectable freezing events, and devitrification can be avoided by using moderate warming rates. Devitrification ceases to be detectable at concentrations above 50%, even at low warming rates, and the system can be regarded as stable (Fig. 2). The intersection of the melting curve and the glass transformation curve at T_g' indicates the minimum concentration of propanediol in aqueous solution that will vitrify irrespective of cooling rate. The concentration at which a glass transition occurs varies according to the nature of the solute. It appears

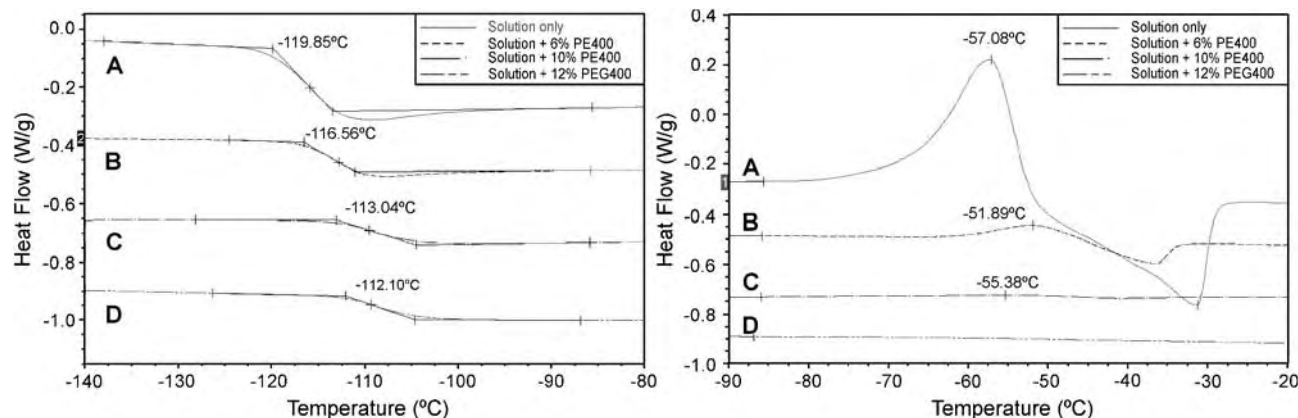


Fig. 3 Representative thermograms illustrating the influence of poly(ethylene glycol) 400 concentration on the glass transition (A) and peak devitrification (B) temperatures of EuroCollin's solution containing 3.0 M DMSO and 3.0 M propanediol. PEG400 elevated the glass transition temperature (A), reduced the energy associated with the transition (A), and reduced the risk of devitrification (B) in a concentration-dependent manner.

that those systems with the strongest solute–solvent hydrogen bonding provide the best suppression of ice nucleation and promote vitrification.^[17,20]

VITREOUS STABILITY

Stability of the vitreous state is critical for the retention of vitrified tissue integrity and viability. Comprehensive studies of vitreous stability for a variety of potentially important cryoprotective mixtures have been made.^[21] Glass stability of vitrified blood vessel samples stored in vapor phase liquid nitrogen storage with retention of smooth muscle function has been demonstrated up to four months of storage.^[22] The stability of the amorphous state has been defined empirically in terms of the critical heating rate, V_{cr} , above which there is insufficient time for a vitreous sample to crystallize before T_m is reached. The smaller the value of V_{cr} , the more stable the amorphous state. The dependence of T_d on the rate of warming can be measured and $T_m - T_d$ has been used to define the stability of the vitreous state.^[21,23] The warming rate for which $T_m - T_d$ is zero is defined as the V_{cr} for which the supercooled mixture neither devitrifies nor recrystallizes. Such studies have shown that stability of glasses formed from aqueous solutions of 1,2-propanediol are much greater, for the same water contents, than for all other solutions of commonly used cryoprotectants including glycerol, dimethyl sulfoxide (DMSO), and ethylene glycol. Unfortunately solutions of poly(alcoholic cryoprotectants) (CPAs) such as propanediol and butanediol, which show the most promise in terms of cooling rates and concentrations necessary for vitrification, also required unrealistically high heating rates to avoid devitrification. Moreover, due principally to isomeric impurities that form a hydrate at reduced temperatures, 2,3-butanediol has proved to have an unanticipated biological toxicity at concentrations below that necessary for vitrification.^[24–27]

Despite developments to devise solutions that would vitrify at practically attainable cooling rates for sizeable biological tissues, the corresponding critical warming rate necessary to avoid devitrification remains a critical challenge. Conceptually, elevated pressures,^[28] electromagnetic heating,^[29–31] the use of naturally occurring antifreeze molecules,^[32] and synthetic ice blockers^[33] have been proposed as means to tackle the problem.

TISSUE VITRIFICATION

Advances in biostabilization and low-temperature biology have produced high-viability preservation

technologies for cells and tissues in the areas of hypothermic storage, cryopreservation by freezing, and vitrification, as well as anhydrobiotic preservation.^[34] However, the development of preservation methods is not straightforward. Process development requires the optimization of chemical and thermal treatments to achieve maximal survival and stability. In a recent editorial,^[35] the need for ice-free cryopreservation methods was emphasized. The consensus opinion was that viable tissues such as blood vessels, corneas, and cartilage that have proven refractory to cryopreservation by conventional freezing methods, despite decades of intense research by many investigators, can only be successfully preserved if steps are taken to prevent or control the ice that forms during cooling and warming. Mathematical modeling may ultimately improve our ability to optimize freezing procedures for tissues,^[36] but has not yet contributed to significant advances. Our laboratory has developed a cryopreservation approach using vitrification, which thus far has demonstrated >80% preservation of smooth muscle cell viability and function in cardiovascular grafts^[37,38] (Table 1).

Avoidance of ice by vitrification can generally be achieved with one of two approaches or a combination of both. The first approach employs cooling highly concentrated solutions (typically >50% w/w) that become sufficiently viscous at low temperatures to suppress crystallization rates. Typically, a vitrified material is considered solid when the viscosity reaches 10^{15} poise^[12] Vitrification can also be achieved by selecting sufficiently high cooling rates to prevent ice crystallization in relatively dilute solutions (<50% w/w). This second approach generally produces a metastable state that is at risk of devitrification (recrystallization) during warming. Ice formation during warming is just as potentially injurious as during cooling.

The tissue vitrification technology described here employs >50% w/w cryoprotective agents. The formulation and method was licensed from the American Red Cross, where it was intended for organ preservation.^[39,40] However, even though rabbit kidneys were vitrified, they could not be rewarmed. Viability

Table 1 Vascular tissue functions after either frozen or vitrified cryopreservation and storage

Tissue type	Survival outcome	
	Frozen (%)	Vitrified (%)
Jugular vein ^[37,38]	6–22	84–87
Carotid artery ^[72]	<30	>80
Tissue-engineered blood vessels ^[73]	10.7	82.7



was lost due to ice formation upon rewarming by the aforementioned process. The rewarming of vitrified materials requires careful selection of heating rates sufficient to prevent significant thermal cracking, devitrification, and recrystallization during heating. The use of carefully designed warming protocols is necessary to maximize product viability and structural integrity. Vitrified materials, which may contain appreciable thermal stresses developed during cooling, may require an initial slow warming step to relieve residual thermal stresses. Dwell times in heating profiles above the glass transition should be brief to minimize the potential for devitrification and recrystallization phenomena. Rapid warming through these temperature regimes generally minimizes prominent effects of any ice crystal damage. It is presently not possible to rewarm organs rapidly enough due to their high volume relative to the volume of tissues. Development of optimum vitrification solutions requires selecting compounds with glass-forming tendencies and tolerable levels of toxicity at the levels required to achieve vitrification. Due to the high total solute concentration within the solution, stepwise protocols should be used at low temperatures for the addition and removal of cryoprotectants to limit excessive cell volume excursions and lower the risk of cytotoxicity. For a current comprehensive review of vitrification, see Taylor et al.^[33]

In addition to *in vitro* studies of cardiovascular tissues, transplant studies have been performed that demonstrate normal *in vivo* behavior of vitrified tissues.^[37,38,41] More recently, we have discovered that vitrification decreases the rate of calcification observed in subcutaneously implanted frozen rat heart valves,^[42] suggesting that vitrification may have other benefits for tissues in addition to increasing cell viability and tissue functions.

More recently, we have extended our vitrification studies to musculoskeletal tissues. Although fresh osteochondral allografts have proven to be effective and functional for transplantation, the limited availability of fresh allograft tissues necessitates the use of osteoarticular allograft banking for long-term storage.^[43–46] Conventional cryopreservation by means of freezing is currently a preferred method for storing tissue in general until needed; however, such protocols result in the death of 80–100% of the chondrocytes in articular cartilage plus extracellular matrix damage due to ice formation. These detrimental effects are major obstacles preventing successful clinical utilization of osteochondral allografts^[45,47,48] and commercial success of tissue-engineered cartilage constructs. Consequently, the search has continued for a better preservation method. The method we have employed for vitrification of articular cartilage is outlined subsequently as an example of a vitrification protocol.

Table 2 Cryopreservation methods

	Freezing	Vitrification
Cryoprotectants	1 M DMSO 2.5% Chondroitin Sulphate	3.1 M DMSO 2.2 M 1,2-propanediol 3.1 M Formamide
Cooling rate	−1°C/min to −80°C	−40°C/min to −100°C −3°C/min to −135°C
Storage	Vapor phase nitrogen	
Warming	Rapid	
Cryoprotectant removal	3 steps	6 steps

This method is contrasted with the method^[49] that we have employed for cryopreservation by freezing of cartilage in Fig. 1 and Table 2.

Example: Protocol for Cryopreservation of Cartilage by Vitrification

Osteochondral plugs (4.0 mm diameter) from New Zealand white rabbits' femurs were gradually infiltrated with an 8.4 M vitrification solution consisting of 3.10 M DMSO, 3.10 M formamide, and 2.21 M 1,2-propanediol in EuroCollins solution at 4°C.^[37,38] Precooled vitrification solution (4°C) was added in six sequential, 15-minute steps. After addition of the final vitrification solution, osteochondral plugs were placed in a glass scintillation vial (Dia. × H, 25 × 60 mm) containing 2 ml of the precooled vitrification solution. The top of the vitrification solution was then covered with 0.7 ml of 2-methylbutane (isopentane, freezing point: −160°C, density: 0.62) at 4°C to prevent direct air contact. A thermocouple was inserted into a separate dummy sample of the same vitrification solution and its output monitored via a digital thermometer throughout the cooling process. Samples were cooled rapidly (43°C/min) to −100°C, followed by slow cooling (3°C/min) to −135°C, and finally storage in a freezer at −135°C for a minimum of 24 hours. Vitrified plugs were rewarmed in two stages: first, slow warming to −100°C (30°C/min) and then rapid warmed to melting (225°C/min). After rewarming, the vitrification solution was removed in a stepwise manner.

CARTILAGE VITRIFICATION

Cryosubstitution studies of frozen and vitrified articular cartilage plugs revealed negligible ice in the vitrified specimens (Fig. 4), and extensive ice

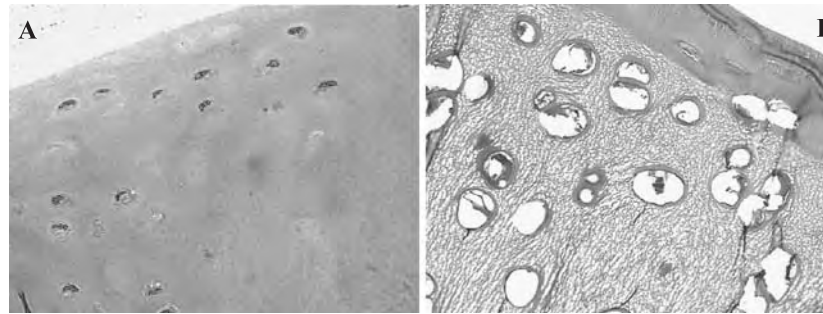


Fig. 4 Light microscopic comparison of frozen and vitrified cartilage specimens that were cryosubstituted during low-temperature storage. The vitrified specimen (A) is free of ice. In contrast, the frozen specimen (B) demonstrates extensive ice in both the lacunae and the extracellular matrix. Cells on the outermost edge (top, right side) of the frozen cartilage specimen are free of ice. Cryosubstitution is a method for demonstrating the presence or absence of ice during cryopreserved storage.^[50] Cryosubstitution was performed using chilled (-90°C) 1% osmium tetroxide in 100% methanol in high-density polyethylene scintillation vials containing cryopreserved specimens at -90°C . The tissues were dehydrated by replacing the frozen or vitrified water with cryosubstitution medium over a period of several days at -90°C . The heat-sink and vials were then placed in a -20°C freezer overnight, followed by 4°C for one hour, and then finally brought to room temperature. This gradual warming of the tissue and cryosubstitution media ensures complete osmium tetroxide tissue fixation. Finally, these tissues were transferred to 100% acetone; infiltrated with araldite resin; and polymerized, sectioned, stained, and viewed by light microscopy.

formation, both in the extracellular matrix and deeper lacunae, in frozen specimens (Fig. 4). Some cell shrinkage was observed in the lacunae of vitrified specimens (Fig. 4), which is most likely related to high cryoprotectant concentrations. The presence and absence of ice are better appreciated in representative electron micrographs of frozen (Fig. 5) and vitrified, cryosubstituted (Fig. 5) cartilage samples.

Viability assessment of fresh and cryopreserved rabbit articular cartilage specimens demonstrated that oxidation–reduction in vitrified samples was approximately 85% of fresh samples (Fig. 6). Similar values have been obtained using live/dead stains for assessment of cell viability.^[51] Transplantation studies have also been performed in rabbits to compare the performance of fresh, frozen, and vitrified specimens. These studies demonstrated that vitrified cartilage performance was not significantly different to fresh, untreated cartilage. In contrast, frozen cartilage performance was significantly different when compared to either fresh or vitrified cartilage.^[74] These studies combine to demonstrate that the vitrification process results in ice-free preservation of rabbit articular cartilage plugs and that about 85% of cellular metabolic activity is retained following rewarming. Frozen tissues contained ice within the cells and the matrix, with the exception of the articular surface, where some viable cells were observed (Fig. 4).

GENERAL DISCUSSION

We believe that the experimental data presented here makes a strong case for cryopreservation of cartilage

by vitrification for osteochondral grafts. The baseline vitrification process, which has yet to be optimized, protected the cartilage from ice formation and demonstrated retention of cell metabolism and viability.^[51] These observations support the hypothesis that vitrification may be applicable to all biological systems.

Simple freezing of cells or tissues results in nonviable, nonfunctional materials. Little advance was made in the field of cryopreservation until Polge et al.^[52] discovered the cryoprotective properties of glycerol. Subsequently, Lovelock and Bishop^[53] discovered that DMSO could also be used as a cryoprotectant. Since the discovery of these cryoprotective agents, cryoprotection during freezing and thawing of biological materials has become established and many other cryoprotectants have been identified that may play a role in cryopreservation, either by freezing or vitrification methods (Table 3). Isolated chondrocytes in suspension, in common with many other cell types, can be preserved using conventional cryopreservation methods involving freezing. In such methods, the cells may be concentrated and vitrified in channels between regions of extracellular ice (Fig. 1); however, chondrocytes embedded in their natural matrix are extremely difficult to preserve by similar freezing methods, presumably because the cells cannot move away from forming extracellular ice.

Studies using a variety of animal articular cartilage models^[44,45,54,55] and human cartilage biopsies^[47] have revealed no more than 20% chondrocyte viability following conventional cryopreservation procedures employing either DMSO or glycerol as cryoprotectants. Ohlendorf et al.^[45] used a bovine articular cartilage, osteochondral plug model to develop a clinical





Fig. 5 Transmission and scanning microscopic comparison of frozen and vitrified cartilage specimens that were cryosubstituted during low-temperature storage. Both the transmission electron and scanning electron micrographs (A & C) illustrate the absence of cellular material in the articular cartilage lacunae. In contrast, the cells are still present in the lacunae of vitrified articular cartilage specimens (B & D). Specimens were cryosubstituted as indicated in the legend for Fig. 4. Tissue for transmission electron microscopy was dehydrated in acetone, gradually infiltrated with an araldite epoxy resin, embedded, and polymerized for 18 hours at 60°C. Sections (75 nm) were cut using a diamond knife with a Reichert OMU2 ultramicrotome. The sections were then double-stained using uranyl acetate followed by lead citrate. The sections were viewed in a JEM-1210 transmission electron microscope (JEOL USA Inc., Peabody, MA) at 80 kV accelerating voltage. For scanning electron microscopy, three changes of 10 ml Hexamethyldisilazane were used in the drying process. The tissue was then oriented on a specimen stub and coated with 20 nm of gold/palladium using an ion-sputter coater. The samples were viewed using a JSM 5410 scanning electron microscope (JEOL USA Inc., Peabody, MA) operated at 10 kV accelerating voltage.

cryopreservation protocol. This protocol employed slow-rate cooling and 8% DMSO as the cryoprotectant. They observed loss of viability in all chondrocytes except those in the most superficial layer at the articular surface. Muldrew et al.^[54] previously investigated chondrocyte survival in a similar sheep model. These researchers observed cells surviving postcryopreservation close to the articular surface and deep at the bone/cartilage interface. The middle layer was devoid of viable cells. More recently, Muldrew et al. demonstrated improved results using a step-cooling cryopreservation protocol, but cell survival posttransplantation was poor and again there was significant loss of cells in the midportion of the graft.^[56] The reason for lack of cell survival deeper than the superficial layers of articular cartilage is most likely multifactorial and related principally to heat and mass transfer considerations.^[57] Surface cells freeze and thaw more rapidly than cells located deep within

the matrix. This phenomenon could result in a greater opportunity for ice to form, both within cells and in the extracellular matrix, deeper within the articular cartilage. Furthermore, typically employed concentrations of DMSO (8–20%) may not penetrate adequately to limit intracellular ice formation. Recent data from Jomha et al.^[58] demonstrated that increasing DMSO concentrations to 6 M can result in higher overall cell survival (40%) after cryopreservation. These observations suggest that use of higher DMSO concentrations results in better penetration of the DMSO into the cartilage.

We are aware that other factors, in addition to ice formation, may have biological consequences during freezing procedures. Two of these factors are the inhibitory effects of low temperatures on chemical and physical processes, and, perhaps more importantly, the physiochemical effects of rising solute concentrations as the volume of liquid water decreases

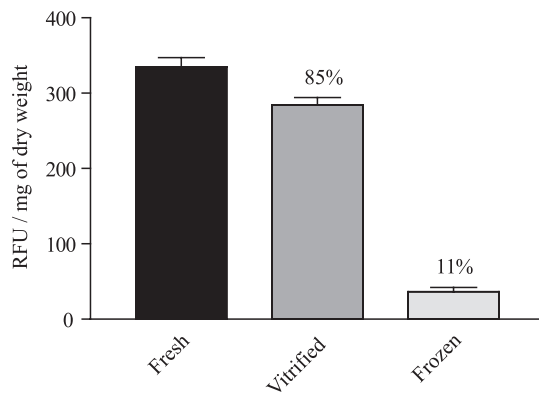


Fig. 6 Metabolic comparison of frozen and vitrified specimens after rewarming under tissue culture conditions. The alamarBlue™ assay was employed as a noncell-specific viability assay. This assay utilizes a water-soluble fluorometric viability indicator based on the detection of metabolic activity, specifically, an oxidation-reduction (REDOX) indicator, which both fluoresces and changes color in response to chemical reduction of the cell culture medium caused by cell metabolism. Aliquots of medium from tissue samples were incubated with alamarBlue™ working solution in microtiter plate wells and read on a microtiter plate spectrofluorometer at 590 nm. The data are expressed as the mean \pm 1 standard error of the mean of 10 samples of articular cartilage assayed using alamarBlue™. Viability is expressed as relative fluorescence units (RFU) relative to the dry weight of each articular cartilage sample. (Reproduced with permission from Ref. [33].)

during crystallization. This latter process results in a decrease in cell volume and the risk of solute precipitation. Several hypotheses have been published on mechanisms of freezing-induced injury based upon such factors,^[57,59] but our own experiences with mammalian tissues concur with others that the principal disadvantage of conventional cryopreservation revolves primarily around ice formation.^[37,38,40,60–62]

Furthermore, cryopreservation by vitrification of tissues offers several important advantages compared with procedures that allow or require ice formation. First, complete vitrification eliminates concerns for the known damaging effects of intra- and extracellular ice crystallization. Second, tissues cryopreserved by vitrification are exposed to less concentrated solutions of cryoprotectants for shorter time periods. For example, Rall^[9] has calculated for embryos that during a typical cryopreservation protocol involving slow freezing to -40°C or -70°C , the cells are exposed to cryoprotectant concentrations of 21.5 and 37.6 osmolal, respectively. In contrast, cells dehydrated in vitrification solutions are exposed for much shorter periods to <18 osmolal solution, although the temperature of exposure is higher. Finally, unlike conventional cryopreservation procedures that employ freezing,

Table 3 Cryoprotective Agents

Small molecular compounds	Permeable (e.g., DMSO, glycerol, 1,2-propanediol) Nonpermeable (e.g., sucrose, polyvinylpyrrolidone)
Used in slow rate freezing and vitrification	
Protect cells via several mechanisms	Displacement of freezable water Suppression of high electrolyte concentration Used to control frozen fraction of water

vitrification does not require controlled cooling and warming at optimum rates. A principal benefit of vitrification is the elimination of requisite studies to determine optimal cooling rates for tissues with multiple cell types. Successful vitrification requires that the thermal processing be rapid enough to transition regions of maximal ice crystal nucleation and growth that occur above the glass transition temperature of the solution. Thus, it is only necessary to cool solutions at rates in which a negligible fraction of the solution forms ice (typically $<0.2\%$).^[63] Vitrified materials have a similar rate requirement during heating, when samples are rewarmed for subsequent use, to limit ice formation to negligible levels (typically $<0.5\%$).^[64] A major objective of our current research is the development of nontoxic vitrification formulations that have a minimal risk of ice formation during cooling and warming of bulk samples.

REMAINING HURDLES IN THE DEVELOPMENT OF VITRIFICATION METHODS FOR BIOLOGICAL SPECIMENS

There are a number of significant challenges for deployment of vitrification methods for tissue-engineered medical products. Vitrification approaches to preservation have some of the limitations associated with conventional freezing approaches. First, both approaches require low-temperature storage and transportation conditions. Neither can be stored above its glass transition temperature for long without significant risk of product damage due to inherent instabilities leading to ice formation and growth. Both approaches employ cryoprotectants with their attendant problems and require competent technical support during rewarming and cryoprotectant elution phases. The high concentrations of cryoprotectants necessary to facilitate vitrification are potentially toxic because



the cells may be exposed to these high concentrations at higher temperatures than in freezing methods of cryopreservation. Cryoprotectants can kill cells by direct chemical toxicity, or indirectly by osmotically-induced stresses during suboptimal addition or removal. Upon completion of warming, the cells should not be exposed to temperatures above 0°C for more than a few minutes before the glass-forming cryoprotectants are removed. It is possible to employ vitrified products in highly controlled environments, such as a commercial manufacturing facility or an operating theater, but not in a doctor's outpatient office, or in third world environments. Formamide, one of the components of the 55% (v/v) vitrification solution consisting of 3.10 M DMSO, 3.10 M formamide, and 2.21 M 1,2-propanediol in EuroCollins solution at 4°C^[37,38] (known as VS55), is a known mutagen. Alternatives to formamide with less safety risks and potentially easier clinical acceptance are being sought. However, the cytotoxicity of complex cryoprotectant formulations containing formamide is surprisingly much less than the cytotoxicity of single-component formulations at the same concentrations.

Heat transfer issues are the primary hurdle for scaling up the successes in tissues to larger organs. The limits of heat and mass transfer in bulky systems result in nonuniform cooling and contribute to stresses that may initiate cracking. In fact, the higher cooling rates that facilitate vitrification will typically lead to higher mechanical stresses. Macroscopic scale instability manifested as cracking or fractures was first reported during the vitrification of glycerol solutions.^[65] Fractures may also provide an interface for nucleation that can initiate devitrification.^[66] Very little information on the material properties of vitreous aqueous solutions exists. Material properties such as thermal conductivity and fracture strength of vitreous aqueous solutions have many similarities with their inorganic analogues that exist at normal temperatures, e.g., window glass and ceramics. Any material that is unrestrained will undergo a change in size (thermal strain) when subjected to a change in temperature. Materials in general, and tissues as they are cryopreserved in particular, shrink when they are brought from physiological temperature down to lower temperatures. Extensive testing is necessary to determine how the thermal strain (shrinkage) of a material depends on the temperature and possibly on the rate at which the temperature is changed. Under these unrestrained conditions when thermal expansion or contraction is free to occur, the stress remains zero, i.e., no forces act on the material. As tissues are cryopreserved, they are externally free to shrink. However, in practice, it is impossible to cool a tissue uniformly. The outer surface decreases in temperature more rapidly than the inside. The outside of the tissue

is forced to shrink less and the inside to shrink more. The stresses that arise to accommodate the differential shrinkage are dependent on the stiffness and relaxation of the material. If these stresses are too severe, fractures may occur.

Calculations of stress in frozen biological tissues have shown that thermal stress can easily reach the yield strength of the frozen tissue, resulting in plastic deformations or fractures.^[67-69] The driving mechanism of thermal stress is the constrained contraction of the frozen or vitrified tissue. It is commonly assumed that thermal expansion of frozen biological tissues is similar to that of pure water ice crystals,^[67,70] and Rabin et al.^[71] have confirmed this experimentally. Moreover, their studies provided some preliminary insight with regard to the effect of the presence of cryoprotectants on the thermal expansion. Results of pilot expansion tests of rabbit muscle permeated by the cryoprotectants DMSO and glycerol solutions, and pig liver perfused with DMSO solution indicated that the cryoprotectants dramatically reduced the thermal expansion at higher temperatures and created a maximum value of thermal expansion within the temperature range of -70 to -100°C. Rabin (personal communication) has preliminary data indicating that tissue thermal expansion decreases with increasing DMSO concentration that may be explained by a change in physical properties associated with a glass transition. The competing needs of vitrification and minimization of mechanical damage demand a greater understanding of both vitrification and stress development. Although mechanical stress has long been recognized as an important mechanism of tissue destruction, it has received very little attention in the context of cryobiology.

Another major hurdle for deployment of vitrification methods is the development of effective rapid warming techniques for larger specimens (greater than 10 ml volume) to prevent devitrification and ice growth by recrystallization. Microwave warming has been attempted, but has never been successful due to the uneven warming of specimens and problems with thermal runaway, which results in heat-denatured tissues. In 1990, Ruggera and Fahy^[31] reported success in warming test solutions at rates of up to about 200°C/min using a novel technology based on electromagnetic techniques (essentially microwave heating). Unfortunately, unpublished results indicate that this method is also problematic due to the uneven warming of specimens and problems associated with thermal runaway. Others have taken a systematic approach to develop a dielectric heating device to achieve uniform and high rates of temperature change.^[29,30] This has been achieved in some early model systems, but application of this technology to the survival of cells and tissues has yet to be reported.

ACKNOWLEDGMENTS

Supported by U.S. Public Health Grants #HL66688, #HL59731, and AR47273. The authors would like to thank Elizabeth Greene and Kelsey K. Brockbank for their assistance in manuscript preparation and acknowledge Fred Lightfoot for preparation of the light and electron micrographs.

ARTICLES OF FURTHER INTEREST

Bioactive Glass; Cryopreservation of Living Cells

REFERENCES

1. Coger, R.; Toner, M. Preservation Techniques for Biomaterials. In *Biomedical Engineering Handbook*; CRC Press: Boca Raton, 1995; 1557–1566.
2. Takahashi, T.; Hirsh, A.G.; Erbe, E.F.; Bross, J.B.; Steere, R.L.; Williams, R.J. Vitrification of human monocytes. *Cryobiology* **1986**, *23*, 103–115.
3. Jutte, N.H.P.M.; Heyse, P.; Jansen, H.G.; Bruining, G.J.; Zeilmaker, G.H. Vitrification of mouse islets of Langerhans: Comparison with a more conventional freezing method. *Cryobiology* **1987**, *24*, 292–302.
4. Jutte, N.H.P.M.; Heyse, P.; Jansen, H.G.; Bruining, G.J.; Zeilmaker, G.H. Vitrification of human islets of Langerhans. *Cryobiology* **1987**, *24*, 403–411.
5. Van Wagtenonck-De Leeuw, A.M.; Den Daas, J.H.G.; Kruip, T.A.M.; Rall, W.F. Comparison of the efficacy of conventional slow freezing and rapid cryopreservation methods for bovine embryos. *Cryobiology* **1995**, *32*, 157–167.
6. Fahy, G.M.; MacFarlane, D.R.; Angell, C.A.; Meryman, H.T. Vitrification as an approach to cryopreservation. *Cryobiology* **1984**, *21*, 407–426.
7. Fahy, G.M. Vitrification. In *Low Temperature Biotechnology: Emerging Applications and Engineering Contributions*; McGrath, J.J., Diller, K.R., Eds.; The American Society of Mechanical Engineers: New York, 1988; 113–146.
8. Pegg, D.E.; Diaper, M.P. Freezing Versus Vitrification; Basic Principles. In *Cryopreservation and Low Temperature Biology in Blood Transfusion*; Smit Sibinga, C.T., Das, P.C., Meryman, H.T., Eds.; Kluwer Academic Publishers: Dordrecht, 1990; 55–69.
9. Rall, W.F. Factors affecting the survival of mouse embryos cryopreserved by vitrification. *Cryobiology* **1987**, *24*, 387–402.
10. Armitage, W.J.; Rich, S.J. Vitrification of organized tissues. *Cryobiology* **1990**, *27*, 483–491.
11. Fahy, G.M. Vitrification as an Approach to Organ Cryopreservation: Past, Present, and Future. In *Cryopreservation and Low Temperature Biology in Blood Transfusion*; Smit Sibinga, C.T., Das, P.C., Meryman, H.T., Eds.; Kluwer Academic Publishers: Dordrecht, 1989.
12. Franks, F. The Properties of Aqueous Solutions at Subzero Temperatures. In *Water: A Comprehensive Treatise*; Franks, F., Ed.; Plenum Press: New York, 1982; Vol. 7.
13. Franks, F. Physiological Water Stress. In *Biophysics of Water*; Franks, F., Mathias, S., Eds.; John Wiley: Chichester, 1982; 279–294.
14. Taylor, M.J. Physico-Chemical Principles in Low Temperature Biology. In *The Effects of Low Temperatures on Biological Systems*; Grout, B.W.W., Morris, G.J., Eds.; Edward Arnold: London, 1987; 3–71.
15. Franks, F. Solution and conformational effects in aqueous solutions of biopolymer analogues. *Philos. Trans. R. Soc. Lond.*, **B 1977**, *278*, 33–57.
16. MacKenzie, A.P. Non-equilibrium freezing behaviour of aqueous systems. *Philos. Trans. R. Soc. Lond.*, **B 1977**, *278*, 167–189.
17. MacFarlane, D.R.; Forsyth, M.; Barton, C.A. Vitrification and Devitrification in Cryopreservation. In *Advances in Low Temperature Biology*; Steponkus, P.L., Ed.; JAI Press: Greenwich, CT, 1992; Vol. 1, 221–278.
18. Angell, C.A.; Sara, E.J.; Donnelly, J.; MacFarlane, D.R. Homogeneous nucleation and glass transition temperatures in solutions of Li salts in D₂O and H₂O: A doubly unstable glass region. *J. Phys. Chem.* **1981**, *85*, 1461–1464.
19. MacFarlane, D.R. Devitrification in glass-forming aqueous solutions. *Cryobiology* **1986**, *23*, 230–244.
20. MacFarlane, D.R.; Forsyth, M. Recent insights on the role of cryoprotective agents in vitrification. *Cryobiology* **1990**, *27*, 345–358.
21. Boutron, P.; Mehl, P.; Kaufmann, A.; Angibaud, P. Glass-forming tendency and stability of the amorphous state in the aqueous solutions of linear polyalcohols with four carbons. *Cryobiology* **1986**, *23*, 453–469.
22. Song, Y.C.; Li, C.Y.; Lightfoot, F.G.; Boggs, J.M.; Greene, E.D.; Taylor, M.J.; Chen, Z.Z.; Brockbank, K.G.M. Stability of Vitrified Tissues in the Vapor Phase of Liquid Nitrogen. Presented at the 39th Annual Meeting of the Society for Cryobiology, 2002; 45 (3), 243.
23. Boutron, P.; Kaufmann, A. Stability of the amorphous state in the system water-1,2-propanediol. *Cryobiology* **1979**, *16*, 557–568.
24. Mehl, P.; Boutron, P. Cryoprotection of red blood cells by 1,3-butanediol and 2,3-butanediol. *Cryobiology* **1988**, *25*, 44–54.
25. Hunt, C.J.; Taylor, M.J.; Chapman, D. Exposure of rabbit corneas to multimolar concentrations of butane-2,3-diol: A transmission and scanning electron microscope study of the endothelium. *Cryobiology* **1991**, *28* (6), 561.
26. Taylor, M.J.; Foreman, J. Tolerance of isolated pancreatic islets to butane-2,3-diol at 0°C. *Cryobiology* **1991**, *28* (6), 566–567.
27. Mejean, A.; Pegg, D.E. Development of a vehicle solution for the introduction and removal of butane-2,3-diol in rabbit kidneys. *Cryobiology* **1991**, *28* (6), 518.
28. MacFarlane, D.R.; Angell, C.A. Homogenous nucleation and glass formation in cryoprotective systems at high pressures. *Cryo-Lett.* **1981**, *2*, 353–358.



29. Marsland, T.P.; Evans, S.; Pegg, D.E. Dielectric measurements for the design of an electromagnetic rewarming system. *Cryobiology* **1987**, *24*, 311–323.
30. Robinson, M.P.; Pegg, D.E. Rapid electromagnetic warming of cells and tissues. *IEEE Trans. Biomed. Eng.* **1999**, *46* (12), 1413–1425.
31. Ruggera, P.S.; Fahy, G.M. Rapid and uniform electromagnetic heating of aqueous cryoprotectant solutions from cryogenic temperatures. *Cryobiology* **1990**, *27*, 465–478.
32. DeVries, A.L. Antifreeze peptides and glycopeptides in cold-water fishes. *Annu. Rev. Physiol.* **1983**, *45*, 245–260.
33. Taylor, M.J.; Song, Y.C.; Brockbank, K.G.M. Vitrification in Tissue Preservation: New Developments. In *Life in the Frozen State*; Benson, E., Fuller, B., Lane, N., Eds.; Taylor and Francis Books: London, 2004; Chapter 22.
34. Walsh, J.R.; Taylor, M.J.; Brockbank, K.G.M. Storage and Transport Issues for Tissue Engineered Medical Products. In *Tissue Engineered Medical Products (TEMPs)*, ASTM STP 1452; Picciolo, G.L., Schutte, E., Eds.; ASTM International: West Conshohocken, PA, 2003.
35. Kaiser, J. New prospects of putting organs on ice. *Science* **2002**, *295* (5557), 1015.
36. Karlsson, J.O.M. Cryopreservation: Freezing and vitrification. *Science* **2002**, *296*, 655–656.
37. Taylor, M.J.; Song, Y.C.; Kheirabadi, B.S.; Lightfoot, F.G.; Brockbank, K.G.M. Vitrification Fulfills its Promise as an Approach to Reducing Freeze-Induced Injury in a Multicellular Tissue. In *Advances in Heat and Mass Transfer in Biotechnology*; American Society of Mechanical Engineers, 1999; 93–102. Volume number HTD-Vol. 363/BED-Vol. 44.
38. Song, Y.C.; Kheirabadi, B.S.; Lightfoot, F.G.; Brockbank, K.G.M.; Taylor, M.J. Vitreous cryopreservation maintains the function of vascular grafts. *Nat. Biotechnol.* **2000**, *18*, 296–299.
39. Fahy, G. Vitrification. In *Low Temperature Biology*; McGrathDillerEds.; ASME, 1988; 113–146.
40. Fahy, G.M.; Saur, J.; Williams, R.J. Physical problems with vitrification of large systems. *Cryobiology* **1990**, *27*, 492–510.
41. Song, Y.C.; Hagen, P.O.; Lightfoot, F.G.; Taylor, M.J.; Smith, A.C.; Brockbank, K.G.M. In vivo evaluation of the effects of a new ice-free cryopreservation process on autologous vascular grafts. *J. Invest. Surg.* **2000**, *13* (5), 279–288.
42. Brockbank, K.G.M.; Song, Y.C. Mechanisms of bioprosthetic heart valve calcification. *Transplantation* **2003**, *75* (8), 1133–1135.
43. Malinin, T.I.; Martinez, O.V.; Brown, M.D. Banking of massive osteoarticular and intercalary bone allografts—12 years' experience. *Clin. Orthop.* **1985**, *197*, 44–57.
44. Marco, F.; Leon, C.; Lopez-Oliva, F.; Perez, A.J.; Sanchez-Barba, A.; Lopez-Duran, S.L. Intact articular cartilage cryopreservation. In vivo evaluation. *Clin. Orthop.* **1992**, *283*, 11–20.
45. Ohlendorf, C.; Tomford, W.W.; Mankin, H.J. Chondrocyte survival in cryopreserved osteochondral articular cartilage. *J. Orthop. Res.* **1996**, *14*, 413–416.
46. Bakay, A.; Csonge, L.; Papp, G.; Fekete, L. Osteochondral resurfacing of the knee joint with allograft. Clinical analysis of 33 cases. *Int. Orthop.* **1998**, *22* (277), 281.
47. Stone, B.B.; Defranzo, B.E.; Dicesare, C.; Rapko, S.M.; Brockbank, K.G.M.; Wolfrum, J.M.; Wrenn, C.A.; Grossman, J.D. Cryopreservation of human articular cartilage for autologous chondrocyte transplantation. *Cryobiology* **1998**, *37* (abstract), 445–446.
48. Tomford, W.W.; Fredericks, G.R.; Mankin, H.J. Studies on cryopreservation of articular cartilage chondrocytes. *J. Bone Jt. Surg., Am.* **1984**, *66*, 253–259.
49. Brockbank, K.G.M. Method for Preserving Musculoskeletal Tissues. Patent 5131850, 1992.
50. Brockbank, K.G.M.; Lightfoot, F.G.; Song, Y.C.; Taylor, M.J. Interstitial ice formation in cryopreserved homografts: A possible cause of tissue deterioration and calcification in vivo. *J. Heart Valve Dis.* **2000**, *9* (2), 200–206.
51. Song, Y.C.; Lightfoot, F.G.; Chen, Z.; Taylor, M.J.; Brockbank, K.G.M. Vitreous preservation of rabbit articular cartilage. *Cell Preserv. Technol.* **2004**, *2* (1), 67–74.
52. Polge, C.; Smith, A.Y.; Parkes, A.S. Revival of spermatozoa after vitrification and dehydration at low temperatures. *Nature* **1949**, *164*, 666.
53. Lovelock, J.E.; Bishop, M.W.H. Prevention of freezing damage to living cells by dimethyl sulfoxide. *Nature* **1959**, *183*, 1394.
54. Muldrew, K.; Hurtig, M.; Schachar, N.; McGann, L.E. Localization of freezing injury in articular cartilage. *Cryobiology* **1994**, *31*, 31–38.
55. Wu, F.J.; Davisson, T.H.; Pegg, D.E. Preservation of tissue-engineered articular cartilage. *Cryobiology* **1998**, *37*, 410.
56. Muldrew, K.; Novak, K.; Studholme, C.; Wohl, G.; Zernicke, R.; Schachar, N.; McGann, L.E. Transplantation of articular cartilage following a step-cooling cryopreservation protocol. *Cryobiology* **2001**, *43*, 260–267.
57. Karlsson, J.O.M.; Toner, M. Long-term storage of tissues by cryopreservation: Critical issues. *Biomaterials* **1994**, *17*, 243–256.
58. Jomha, N.M.; Anoop, P.C.; Bagnall, K.; McGann, L.E. Effects of increasing concentrations of dimethyl sulfoxide during cryopreservation of porcine articular cartilage. *Cell Preserv. Technol.* **2002**, *1* (2), 111.
59. Mazur, P. Freezing of living cells: Mechanisms and implications. *Am. J. Physiol.* **1984**, *247*, 125.
60. Taylor, M.J.; Pegg, D.E. The effect of ice formation on the function of smooth muscle tissue following storage at -21°C and -60°C . *Cryobiology* **1982**, *20*, 36–40.
61. Hunt, C.J.; Taylor, M.J.; Pegg, D.E. Freeze-substitution and isothermal freeze fixation studies to elucidate the pattern of ice formation on smooth muscle at 252 K (-21°C). *J. Microsc.* **1982**, *125* (2), 177–186.
62. Taylor, M.J. Sub-Zero Preservation and the Prospect of Long-Term Storage of Multicellular Tissues and Organs. In *Transplantation Immunology-Clinical and Experimental*; Calne, R.Y., Ed.; Oxford University Press: Oxford, 1984; 360–390.

63. Boutron, P.; Mehl, P.; Kaufmann, A.; Augibaud, P. Glass forming tendency and stability of the amorphous state in aqueous solutions of polyalcohols with four carbons. I Binary systems water-polyalcohols. *Cryobiology* **1986**, *23*, 453–469.
64. Boutron, P.; Mehl, P. Theoretical predictions of devitrification tendency: Determination of critical warming rates using finite expansions. *Cryobiology* **1990**, *27*, 359–377.
65. Kroener, C.; Luyet, B. Formation of cracks during the vitrification of glycerol solutions and disappearance of the cracks during rewarming. *Biodynamica* **1966**, *10* (201), 47–52.
66. Williams, R.J. Four modes of nucleation in viscous solutions. *Cryobiology* **1989**, *26* (6), 568.
67. Rabin, Y.; Podbilewicz, B. Temperature-controlled microscopy for 4-D imaging of living cells: Apparatus, thermal analysis, and temperature dependency of embryonic elongation in *C. elegans*. *J. Microsc.* **2000**, *199* (3), 214–223.
68. Rabin, Y.; Steif, P.S. Thermal stresses in a freezing sphere and its application to cryobiology. *ASME J. Appl. Mech.* **1998**, *65* (2), 328–333.
69. Rabin, Y.; Steif, P.S. Thermal stress modeling in cryosurgery. *Int. J. Solids Struct.* **2000**, *37*, 2363–2375.
70. Rubinsky, B.; Cravalho, E.G.; Mikic, B. Thermal stress in frozen organs. *Cryobiology* **1980**, *17*, 66–73.
71. Rabin, Y.; Taylor, M.J.; Wolmark, N. Thermal expansion measurements of frozen biological tissues at cryogenic temperatures. *J. Biomech. Eng.* **1998**, *120*, 259–266.
72. Song, Y.C.; Taylor, M.J.; Brockbank, K.G.M. Ice-free cryopreservation of arterial grafts. *Cryobiology* **2001**, *41*, 370 (abstract).
73. Brockbank, K.G.M.; Song, Y.C. Development of preservation methods for cellular engineered tissues. Presented at the First World Congress on Regenerative Medicine, Leipzig, Germany. *Int. J. Artif. Organs* **2003**, *26*, 9, 818 (abstract).
74. Song, Y.C.; An, Y.H.; Kang, Q.K.; Li, C.; Bogg, J.M.; Chen, Z.; Taylor, M.J.; Brockbank, K.G.M. Vitreous preservation of articular cartilage grafts. *J. Invest. Surg.* **2004**, *17*, 1–6.



Wear Debris

John Fisher
Eileen Ingham

*Institute of Medical and Biological Engineering, School of Mechanical Engineering,
University of Leeds, Leeds, United Kingdom*

INTRODUCTION

Over two million artificial hip and knee joints are implanted in patients every year. For many patients, these provide excellent clinical performance for 10 years or longer, with relief of pain and restoration of function. Clinical studies show survivorship of 90% for prosthetic components after 10 years implantation. In the longer term, failure rates start to increase. One major cause of medium- to long-term failure is bone loss or bone lysis (osteolysis) caused by adverse reactions to the wear particles generated by the artificial joints. Billions of micron and submicron size wear particles are generated every year by the articulating interfaces in the artificial joints. These particles accumulate in periprosthetic tissues and in interface tissues where adverse cellular reactions, predominantly mediated by macrophages, lead to bone resorption and eventually loosening and failure of the prostheses. The particle/cell-mediated biological pathways that lead to osteolysis are complex. Two independent processes predominate. For most prostheses, micron and submicron particles (such as those generated by the polyethylene bearing surfaces) are taken up by macrophages, which are stimulated to produce osteolytic cytokines such as TNF alpha, which act to cause bone resorption. For chemically stable particles such as polyethylene, these are not digested by the macrophages and persist to produce a chronic inflammatory response. Additionally, other wear particles that are chemically active, such as metal, release ions into the tissues. At high concentrations, these can affect cell viability (of macrophages, fibroblasts, osteocytes) causing local tissue necrosis.

These adverse reactions to wear particles are dependent on the concentration, size, morphology, and chemistry of the wear particles generated by the artificial joints, and directly determine the longevity of the prosthesis. Additionally, the accumulation of particles within the periprosthetic tissues is dependent on the transport processes within the body, which, in turn, are determined by particle size and morphology, as well as the structure and integrity of local tissue interfaces. Larger particles, one micron and above, as well as submicron (0.1 to 1 μm) size particles have been found

to remain and accumulate in the periprosthetic tissues. Smaller nanometer size particles (1 to 100 nm) have also been found to be distributed more widely and cause adverse systemic reactions. The longevity of the prostheses and its functional biocompatibility is critically determined by the wear volume and characteristics of the wear particles generated during articulation. Traditionally, the materials used in artificial joints have been primarily selected on the basis of volumetric wear rate, and while this remains important, it is now recognized that the characteristics and biological activity of the wear debris produced is also critical to the clinical performance of the artificial joints.

OVERVIEW

In this article, an holistic approach to wear debris generation in artificial joints is taken in which the tribological processes that generate the wear particles and the wear rates, the characteristics of the wear particles, and the biological reactions and functional biocompatibility of the wear particles are considered in an integrated manner.^[1,2]

In artificial hip joints, the traditional bearing materials have been a metal femoral head articulating against an ultrahigh molecular weight polyethylene socket. Concerns about polyethylene wear debris-induced osteolysis have led to the development of improved polyethylenes, the use of ceramic femoral heads, and the introduction of hard-on-hard bearing surfaces such as metal on metal and ceramic on ceramic. The introduction of modular prosthetic designs has led to additional interfaces in which wear particles can be generated, such as the head stem junction and the acetabular insert/acetabular cup interface, as well as fixation interfaces such as the stem cement interface. Although the volume of debris generated at these non-articulating interfaces is often much less than at the articulating surfaces, the debris can have direct access to local areas of periprosthetic tissues (such as through screw holes or through the cement mantle) and, hence, produce high localized concentrations and osteolysis.

In artificial knee joints, the predominant bearing materials currently used are metal and polyethylene. Clinical experience and follow-up in knees is generally shorter than in hips and the longer-term problem of osteolysis is only just emerging. Although few materials have been used in the knee, there are a greater number of design variables that influence wear particle generation.

In this article, wear particle generation in traditional metal-on-polyethylene hips is described along with the recently introduced improved polyethylenes. Wear particle generation in hard-on-hard bearings for hips is then discussed. In the knee, polyethylene wear particle generation is described with a particular emphasis on the effect of different prosthetic designs on particle generation.

HISTORICAL POLYETHYLENE ON METAL HIP PROSTHESES

Polyethylene-on-metal bearing couples for artificial hip joints were first introduced over 40 years ago by the late Sir John Charnley. Charnley also introduced the concept of bone cement fixation. Both these approaches remain the primary bearing surface and fixation methods used in hip prostheses today. In these prostheses, the primary source of wear debris is the articulating interface between the polished metallic spherical convex head and the concave hemispherical ultrahigh-molecular-weight polyethylene acetabular cup. As the polished head slides over the polymer cup during normal activities, the polyethylene wear particles are produced by microscopic asperity interactions, causing adhesive, abrasive, and micro fatigue wear mechanisms. The polyethylene particles have been found to be in the range 0.1 to 10 μm in size with the majority being submicron in dimensions. These very small particles mean that there are over 10 billion particles produced for every one cubic millimeter of wear volume. Clinical studies^[3] have revealed wear rates of approximately 60 mm^3 per year, with substantial patient variability. Total wear volumes of over 600 mm^3 at failure are equivalent to approximately 6000 billion polyethylene wear particles released into the periprosthetic tissue.

Laboratory wear studies of the historical polyethylene have shown wear rates of approximately 40 mm^3 per million cycles,^[4] with similar wear particle characteristics to those found in vivo.^[5] Laboratory wear studies have allowed the identification of factors that can increase the acceleration of wear debris generation in the historical polyethylenes and that may contribute to the patient-to-patient variation seen clinically. Increased patient activity (number of cycles) as well as a larger head size increases the wear rate due

to increased sliding distances.^[6] Damage to the polished metallic heads, often caused by hard third body particles, can accelerate wear by a factor of two to three times,^[7] and this has been also shown in clinical studies.^[3]

The majority of the historical ultrahigh molecular weight polyethylene had a molecular weight in the range of two to four million and was sterilized with gamma irradiation in the presence of oxygen. Irradiation produces chain scission of the polyethylene, generating both cross-links and free radicals. In the presence of oxygen, the latter cause degradation of the polyethylene, both while on the shelf and also once implanted in vivo.^[8] This degradation can accelerate the wear rate^[9] and, most importantly, also reduce the size of the wear particles.^[5] Serum proteins have been shown to be absorbed by polyethylene.^[10] Other factors, such as the nature of the lubricant, affect wear, with increased lipid concentration reducing wear.^[11] The nature of the gait cycle and the kinematics of the wear path can alter wear rates in polyethylene. Pooley and Tabor^[12] showed that with linear polymers, linear motion or sliding produced strain hardening and increased wear resistance. Wang^[13] showed that for conditions in the hip in which there is multidirectional sliding, the material is strain softened in the transverse direction, allowing cross shear frictional forces found in multidirectional sliding to accelerate wear. A clinical correlation between the aspect ratio of the wear track and the wear rate has been shown by Bennett et al.^[14] Low-aspect tracks with higher cross shear increased wear, while high-aspect, more linear wear tracks produced lower wear rates, confirming earlier simulator studies.^[4,13] The concentration, size, and morphology of the polyethylene wear particles have been found to be critical in determining the stimulation of macrophages and the release of osteolytic cytokines. Increased volumetric concentrations of wear particles increased the release of osteolytic cytokines.^[15] Polyethylene particles in the size range 0.1 to 1 μm were found to be six times more reactive than particles in the range 1 to 10 μm .^[16] Direct studies of bone resorption^[17] have confirmed higher levels of osteolysis with the smaller particles in the 0.1 to 1 μm size range. These findings of the osteolytic response being dependent on wear particle size have significant impact. For example, oxidative degradation, which produces smaller wear particles,^[5] will work synergistically with increased wear volume to accelerate osteolysis and loosening. A small increase in the molecular weight of polyethylene from two to four million to four to six million has also been shown to reduce the size of the wear particles and increase the levels of osteolytic cytokines produced.^[18] Laboratory models have now been developed that combine tribological and biological studies to directly predict the functional

osteolytic potential of polyethylene-bearing materials and designs in artificial hip joints.^[19] These can be applied to investigate reductions in osteolytic potential of new and improved polyethylenes that have been introduced into clinical practice in recent years.

NEW AND IMPROVED POLYETHYLENES FOR HIP PROSTHESES

The recognition that polyethylene wear particles are a major cause of osteolysis and failure in hip prostheses, and the understanding of the role of oxidation following gamma irradiation in air leading to degradation and increased wear rates has led to two new approaches for improvement of polyethylene. Stabilized polyethylenes have been produced either through alternative sterilization methods such as ethylene oxide or gas plasma or through irradiation in an inert atmosphere such as nitrogen, vacuum, or argon. The latter is thought to have an advantage as it also introduces a moderate level of cross-linking which can reduce wear rates. Moderate cross-linking of ultrahigh-molecular-weight polyethylene (molecular weight 2 to 4 MRad) has been shown to reduce the volumetric wear rate of ultrahigh molecular weight polyethylene by 30% in hip simulator studies^[20] (from 49 to 35 mm³ per million cycles). However, the moderate cross-linking also produced a greater percentage volume of the smaller particles in the 0.1 to 1 µm size range, with the net result of little difference in the functional osteolytic potential of the two stabilized materials. However, both of these stabilized materials have substantially reduced wear rates and osteolytic potential compared to historical gamma irradiated in air and aged polyethylene.^[21]

The reduction in wear volume found by moderately cross-linking polyethylene led to intentional cross-linking with higher levels (medium or high) of irradiation, either gamma irradiation in an inert atmosphere or electron beam irradiation, followed by heat treatment to recombine free radicals and produce additional cross-links. These increased levels of cross-linking increase the wear resistance of the ultrahigh molecular weight polyethylene, but also reduce its ductility and toughness. McKellop et al. have shown an 85% reduction in wear rate with 5 MRad gamma irradiation and heat treatment.^[22] While Muratoglu et al. have shown a greater reduction of over 90% with 10 MRad of electron beam irradiation,^[23] Kurtz et al. have also shown reduced wear rates with increased levels of cross-linking.^[24] In our wear simulator studies,^[25] the reduction in wear rate with increased cross-linking was significant, but not as great as reported in other studies (see Table 1).

Table 1 Relative wear rates with irradiation cross-linking of UHMWPE

Cross-linked polyethylene	Relative wear rate
0 MRad	100%
5 MRad	80%
10 MRad	25%

Analysis of the wear particles showed a trend toward smaller particles in the 0.1 to 1.0 micrometer size range with the higher levels of cross-linking. Cell culture studies also showed that a unit volume of wear debris from the cross-linked material produced higher levels of osteolytic cytokines.^[26] This has to be considered alongside the reduction in wear volume. Analysis of the volume of wear particles in the most biologically active size range (0.1 to 1 µm) showed a reduction with cross-linking (Table 2), and this was consistent with a predicted reduction in functional osteolytic potential with cross-linking of the higher molecular weight polyethylene.

These studies of cross-linked higher molecular weight polyethylene have to be considered with the separate studies, which showed that the lower molecular weight polyethylene wear debris in its non-cross-linked form also had a substantial reduction in osteolytic potential compared to this higher molecular weight material.^[18,26] Clinical studies are now under way with the new cross-linked polyethylenes. However, longer periods of time of follow-ups of 10 to 20 years will be necessary to determine the reduction in osteolytic potential clinically.

ALTERNATIVE HARD-ON-HARD BEARING COUPLES

Ceramic-on-ceramic^[27] and metal-on-metal bearing couples^[28] have been used in very limited numbers for over 30 years. It is only in the last five years, with concerns over polyethylene wear debris-induced osteolysis, that they have been used more extensively

Table 2 Relative total volume of wear particles in critical size range 0.1 to 1 µm for higher molecular weight polyethylene

Cross-linked polyethylene	Relative total volume of wear particles in size range 0.1 to 1 µm rate
0 MRad	100%
5 MRad	95%
10 MRad	35%

clinically. Both materials are substantially harder than polyethylene with the hardness of alumina ceramic (25,000 MPa) and cobalt chrome alloy (4000 MPa) being substantially greater than ultrahigh molecular weight polyethylene (less than 100 MPa). In classical wear terms, the wear rate is inversely proportional to hardness, giving lower wear rates for the harder materials. Additionally, the smooth surface finishes of the ceramic and metal bearings (surface roughness $\sim 0.01 \mu\text{m Ra}$), compared to $0.1 \mu\text{m Ra}$ for polyethylene, provide further potential for wear reduction through fluid film lubrication.

Clinically-recorded wear rates for these historical hard-on-hard bearings have been approximately 1 mm^3 per year for metal on metal^[28] and even less for alumina ceramic on ceramic couples.^[29] These values are 100 times less than those found for historical polyethylene, which is consistent with the differences in hardness of the materials. It is not surprising that there has been little evidence of wear debris-induced osteolysis in hard-on-hard bearings. However, historical wear rates and clinical performance have been variable, and geometrical factors such as radial clearance and sphericity, anatomical factors such as positioning, fixation stability, and soft tissue tension,^[30,31] as well as historical material quality^[27] have all influenced wear rates. Only recently has the wear of these hard-on-hard bearings been investigated systematically and the wear debris and subsequent cellular and tissue reactions studied.

There have been considerable improvements in the properties of alumina ceramics over the last 20 years with smaller grain sizes and more recently, hot isostatic pressing (HIPING) producing increased fracture toughness and hardness. Under standard walking cycle hip simulator conditions, the wear rates of alumina ceramic-on-ceramic bearings are extremely low, less than 0.1 mm^3 per million cycles^[32] with a small reduction in wear with HIPED material.^[33] The wear mechanism found was relief polishing,^[32,33] leaving an extremely smooth wear surface, $Ra < 0.01 \mu\text{m}$, and extremely small wear debris, 10 nm in size.^[34] However, these wear rates, scars, mechanisms, and debris did not fully represent clinical retrievals, in which relief polishing was found on all components. Additionally, on many retrievals, more severe stripe wear and intergranular fracture were found on femoral heads, cup rim damage was identified, and higher wear rates of up to 1 mm^3 per year were found.^[29,31] Furthermore, clinical wear debris was both nanometer in size with large micron-size debris associated with stripe wear, identified in periprosthetic tissue.^[35] More recently, this has been simulated in the laboratory by introducing joint laxity and contact of the head on the rim of the acetabular cup,^[30] which produced higher wear rates up to 1 mm^3 per million cycles. The nanometer

and micrometer ceramic wear particles and the characteristic stripe wear on the head and cup rim damage^[36] were also generated. Hence, it has been found that this head cup rim contact is necessary to replicate clinical wear rates, mechanisms, and debris generated in alumina ceramic-on-ceramic bearings.^[34]

Alumina ceramic has been used extensively, as in its bulk it is a bioinert biomaterial.^[27] However, as we have seen with polyethylene, the biological response to particles has been found to be very different to that of the bulk. Recent studies have directly cultured macrophages with clinically relevant alumina ceramic wear debris and have found that at very high volumetric concentrations, it will stimulate macrophages to produce osteolytic cytokines such as $\text{TNF}\alpha$.^[37] However, for a given volumetric concentration, ceramic debris is less reactive than the more widely used polyethylene debris. Most importantly, at functional concentrations, which are much lower for alumina ceramic, the osteolytic potential wear debris of alumina ceramic bearings is much less than polyethylene bearings.^[38]

The clinical wear rates of historical cobalt chrome alloy were reviewed by Amstutz et al.^[28] and although found to be highly variable, they were found to be substantially lower than polyethylene, approximately 1 mm^3 per year. Doorn et al. studied the characteristics of cobalt chrome debris in surrounding periprosthetic tissues and generally found it to be very small in size (10 to 100 nm) with only a few particles larger than 100 nm in size.^[39] Again, the harder historical bearing was found to produce lower wear rates and have smaller particles and a smoother wear surface than polyethylene. There have been clinical concerns about the chemical activity of small cobalt chrome wear particles, and release of cobalt chrome ions into the periprosthetic tissues and systemically in the body. There is anecdotal evidence of widespread distribution of smaller metallic particles throughout the body.^[40]

The recent development and clinical introduction of second generation metal-on-metal hip prostheses has led to more widespread simulator studies and better understanding of the variables that control the generation of wear debris in cobalt chrome metal on metal hip prostheses. The carbon content of the cobalt chrome alloy has been found to be an important determinant of wear, with combinations of high carbon cobalt chrome alloys producing lower wear rates than low carbon cobalt chrome alloys.^[41,42] In both cases, nanometer size wear particles in the range 10 to 100 nm were identified. Metal-on-metal hip prostheses appear to be very sensitive to the tribological conditions of the simulator with kinematics and loading conditions changing the wear rates from 0.1 mm^3 /million cycles to 1.5 mm^3 /million cycles for the same prostheses.^[43] Chan et al. found stop start motion to

increase wear^[44] and more recently, it has been shown that both elevation of the swing phase load from 100N to 270N and also increased joint laxity can increase wear rates from 0.1 to over 1 mm³/million cycles.^[45] This is indicative that metal-on-metal hips are sensitive to the contact conditions and resulting lubricating film thicknesses.^[46] This is then supportive of findings that larger diameter metal-on-metal hips will reduce wear by improved lubrication.^[46] However, there are concerns about these arguments as the reduction of wear by lubrication is dependent on many factors such as contact conditions, loading conditions (including swing phase load), and the need for continuous motion and activity. If the fluid film breaks down, boundary lubrication will occur, friction and wear will increase, and under these conditions, the greater sliding distance of the larger diameter head will increase wear rates. Although the wear rates of metal on metal under these conditions of 1 to 2 mm³/million cycles remain low compared to polyethylene, it is the nature of the wear debris and its biochemical activity that remain a concern. Recent cell culture studies with clinically relevant nanometer-size cobalt chrome wear debris showed substantial reductions in cell viability at low concentration levels of 5 (μm)³/cell.^[47] It is interesting to compare this with the concentration of 30 to 100 (μm)³/cell of polyethylene particles that are needed to stimulate significant amounts of osteolytic cytokines.^[15,16] At the lower levels of concentration 5 (μm)³/cell, neither polyethylene nor ceramic debris has any effect on cell viability. The small cobalt chrome wear particles have a very large surface area for release of both cobalt and chromium ions, which influences cell and tissue viability at low concentration levels. Clinical evidence of local tissue necrosis (metallosis) has been found when metallic ion release has been high.^[28,34] The more recent introduction of large diameter metal-on-metal surface replacement hip prostheses has further highlighted these concerns.

ARTIFICIAL KNEE PROSTHESES

The predominant bearing surfaces in knee prostheses have been ultrahigh-molecular-weight polyethylene (UHMWPE) articulating against polished cobalt chrome femoral components. Whereas the articulating surfaces are the primary source of wear debris, other interfaces such as the backside wear with the metal tibial tray can generate debris, which may directly access the periprosthetic interfaces through screwholes. UHMWPE wear debris is the predominant concern in the knee with the potential for osteolysis as found in the hip.

Historically, as in the hip, UHMWPE has been irradiated in the presence of air. Oxidative degradation,^[8] which results in reduced toughness, ductility,

and fatigue strength has led to failures of knees by delamination fatigue^[48,49] or delamination wear. This type of wear is predominantly associated with designs that have high contact stress, by excessive gliding of femoral components over tibial trays, by misalignment, or poor surgical technique.^[50] The introduction of oxidative resistant polyethylenes and stabilized materials has been shown to substantially reduce or eliminate this type of delamination failure in laboratory tests.^[51,52]

Although methods of reducing or eliminating delamination failures in knees have been identified, concerns about surface wear and the generation of micron and submicron-size polyethylene wear particles, which can cause osteolysis, remain.^[53] Whereas excellent clinical results at 10 years in the knee have been found,^[54] with over 90% success rates, as in the hip, concerns about longer-term wear debris-induced osteolysis remain.^[53] Howling et al. have identified micron and submicron-size polyethylene wear particles in the knee from early retrievals and predict the increase and accumulation with time to produce osteolysis and late failure, as in the hip.^[55]

In the last five years, there have been several studies of the wear of polyethylene in knee prostheses in knee joint simulators. Initial studies reported low wear rates of less than 5 mm³/million cycles, which is much less than found in the hip.^[56–58] This may have been due to the application of force control to tibial tray motions, which did not realize physiological displacements^[56,57] or the absence of a critical tibial motion such as internal–external rotation^[58] in the simulation. More recent studies using displacement controlled knee simulators, which directly control the tibial tray motions to clinically relevant displacement conditions, have shown much higher wear rates of the order of 5 to 40 mm³/million cycles.^[55,59,60] These studies have also revealed a predominance of submicron and micron-size wear particles as found in the surface wear mechanisms in the hip. Wear rates have been found to be greater in gamma irradiated in air polyethylene^[61] than in more recently developed stabilized polyethylene irradiated in a vacuum.^[62] The wear rate has also been found to be critically dependent on the displacement inputs to the tibial tray. An increase in internal–external rotation and anterior–posterior displacements by a factor of two revealed a sevenfold increase in wear rate.^[59,61] The increased tibial rotation provides additional cross shear frictional forces in the transverse direction, which has been strain softened,^[13] so substantially increasing wear. This dependency of wear on kinematics is similar to that found in the hip.^[4,12–14] In the knee, the development of rotating platform mobile bearing designs, which allows rotation at the tibial platform interface, as well as articulation at the femoral condylar, produces, effectively, linear motion at both

interfaces and a substantial reduction in wear volume.^[59,62] This more unidirectional motion has the additional advantage of generating a lower percentage number of the most biologically active submicron polyethylene wear particles.^[63] Studies of different types of polyethylene materials have shown that they appear to have less influence on wear than the different kinematic conditions or the influence of design.^[59] As increasing numbers of knee replacements are being implanted in younger, more active patients, the incidence of polyethylene wear debris-induced osteolysis and long-term failure may well increase.

CONCLUSIONS

Artificial joints are implanted to provide interfaces for both load carriage and articulation. Wear and the generation of wear debris is an inevitable consequence of restoration of function. The release of wear debris and its accumulation in periprosthetic tissues can lead to adverse tissue reactions, loosening, and failure. These adverse biological reactions are dependent on the volumetric concentration, physical characteristics, size, morphology, and chemistry of the wear particles. This has led to the concept of functional biocompatibility of wear debris^[19] generated in artificial joints, which is dependent on the materials, design, tribological conditions, and surgical approach as well as individual patient specific responses.

Historical polyethylenes that have been gamma irradiated in the presence of air, degrade due to oxidation both in vitro and in vivo, leading to increased wear and increased volumes of smaller and more biologically active wear particles. This has led to osteolysis, loosening, and failure in the hip and the knee. Polyethylene wear particles in the size range 100 nm to 1 μ m have been found to be the most biologically active. Stabilized polyethylenes that do not degrade oxidatively produce lower wear volumes and debris that is less biologically active. In contrast, the use of higher-molecular-weight polyethylene produces debris that is more reactive. Increased levels of cross-linking further reduce the volumetric wear rate, but some of the advantage is lost as the debris itself is smaller and more biologically active. For all polyethylenes, variables such as femoral head damage, increased femoral head size, low aspect ratio (more circular) wear paths, and gait will accelerate wear or debris generation.

Hard-on-hard bearings produce smaller volumes of wear debris than polyethylene bearings in the hip. Additionally, alumina ceramic debris has been found to be both nanometer as well as submicron in size and produces lower levels of biological activity, resulting in substantially reduced osteolytic potential

than polyethylene bearings. Whereas metal-on-metal bearings do give rise to smaller wear volumes, the small nanometer-size particles are chemically active, releasing cobalt and chromium ions, which, even at low volume concentrations, can reduce cell viability. Designs or tribological conditions that accelerate metal-on-metal wear above the low levels of 1 mm³ per year may lead to adverse local tissue reactions and necrosis, as well as elevated systemic metal ion levels. For all the different types of wear debris, there are substantial variations in the levels of individual responses seen in different patients, and further work is needed to understand these patient-specific reactions to wear debris.

The generation of wear debris, which can lead to failure of joint replacement, is a result of complex and interactive mechanical and biological processes. Development and evaluation of new and improved materials and designs, which lead to improved functional biocompatibility and reduced osteolytic potential, need to involve an holistic, multidisciplinary biological engineering approach.



ACKNOWLEDGMENTS

A number of research studies described in this article were supported by the EPSRC and ARC.

ARTICLES OF FURTHER INTEREST

Centrifugal Blood Pumps; Implant, Total Hip; Orthopedic Biomaterials; Titanium and Its Alloys; Ultra-Molecular Weight Polyethylene (UHMWPE); Wear Debris, Bone Resorption Animal Models; Wear Debris from Joint Arthroplasties, Biological Effects Of

REFERENCES

1. Willert, H.G.; Semlitsch, M. Reactions of the articular capsule to wear products of artificial joint prostheses. *J. Biomed. Mater. Res.* **1977**, *11*, 157–164.
2. Ingham, E.; Fisher, J. Biological reactions to wear debris in total joint replacement. *Proc. Inst. Mech. Eng., H* **2000**, *214*, 21–37.
3. Tipper, J.L.; Ingham, E.; Hailey, J.L.; Besong, A.A.; Fisher, J.; Wroblewski, B.M.; Stone, M.H. Quantitative analysis of polyethylene wear debris, wear rate and head damage in retrieved Charnley hip prostheses. *Mater. Sci., Mater. Med.* **2000**, *11*, 117–124.
4. Barbour, P.S.M.; Stone, M.H.; Fisher, J. A hip joint simulator study using simplified loading and motion cycles generating physiological wear paths and rates. *Proc. Inst. Mech. Eng., H* **1999**, *214*, 455–467.

5. Besong, A.A.; Tipper, J.L.; Ingham, E.; Stone, M.H.; Wroblewski, B.M.; Fisher, J. Quantitative comparison of wear debris from UHMWPE that has and has not been sterilised by gamma irradiation. *J. Bone Jt. Surg.* **1998**, *80B*, 340–344.
6. Fisher, J.; Dowson, D. Tribology of total artificial joints. *Proc. Inst. Mech. Eng.*, H **1991**, *205*, 73–79.
7. Barbour, P.S.M.; Stone, M.H.; Fisher, J. A hip joint simulator study using new and physiologically scratched femoral heads with ultra-high molecular weight polyethylene acetabular cups. *Proc. Inst. Mech. Eng.*, H **2000**, *214*, 569–576.
8. Blunn, G.; Brach del Preva, E.M.; Costa, L.; Fisher, J.; Freeman, M.A.R. Ultrahigh molecular weight polyethylene (UHMWPE) in total knee replacement: Fabrication, sterilisation and wear. *J. Bone Jt. Surg.* **2002**, *84-B*, 946–949.
9. Fisher, J.; Chan, K.L.; Hailey, J.L.; Shaw, D.; Stone, M.H. Preliminary study of the effect of ageing following irradiation on the wear of ultrahigh-molecular-weight polyethylene. *J. Arthroplast.* **1995**, *10* (5), 689–692.
10. Blanchet, L. Serum lubrication absorption by UHMWPE. Orthopaedic bearing implants. *J. Tribol.* **2002**, *124*, 1–4.
11. Bell, J.; Tipper, J.L.; Ingham, E.; Stone, M.H.; Fisher, J. The influence of phospholipid concentration in protein-containing lubricants on the wear of ultra-high molecular weight polyethylene in artificial hip joints. *Proc. Inst. Mech. Eng.*, H **2001**, *215*, 259–263.
12. Pooley, C.M.; Tabor, D. Friction and molecular structure: The behaviour of some thermoplastics. *Proc. R. Soc. Lond.* **1972**, *329A*, 251–258.
13. Wang, A.; Stark, C.; Dumbleton, J.H. Mechanistic and morphological origins of UHMWPE wear debris in total joint replacement prostheses. *Proc. Inst. Mech. Eng.*, H **1996**, *210*, 137–141.
14. Bennett, D.; Orr, J.; Beverland, D.E.; Baker, R. The influence of shape and sliding distance of femoral head loci on the wear of acetabular cups. *Proc. Inst. Mech. Eng.*, H **2002**, *216*, 393–402.
15. Campbell, P.; Doorn, P.; Dorey, F.; Amstutz, H.C. Wear and morphology of ultra-high molecular weight polyethylene wear particles from total hip replacements. *Proc. Inst. Mech. Eng.*, H **1996**, *210*, 167–174.
16. Green, T.R.; Fisher, J.; Stone, M.H.; Wroblewski, B.M.; Ingham, E. Polyethylene particles of a 'critical size' are necessary for the induction of cytokines by macrophages in vitro. *Biomaterials* **1998**, *19*, 2297–2302.
17. Green, T.R.; Fisher, J.; Matthews, J.B.; Stone, M.H.; Ingham, E. Effect of size and dose on bone resorption activity of macrophages by in vitro clinically relevant ultra high molecular weight polyethylene particles. *J. Biomed. Mater. Res. (Appl. Biomater.)* **2000**, *53*, 490–497.
18. Ingram, J.; Matthews, J.B.; Tipper, J.L.; Stone, M.H.; Fisher, J.; Ingham, E. Comparison of the biological activity of grade GUR 1120 and GUR 415 HP UHMWPE wear debris. *Bio-Med. Mater. Eng.* **2002**, *12*, 177–188.
19. Fisher, J.; Bell, J.; Barbour, P.S.M.; Tipper, J.L.; Matthews, J.B.; Besong, A.A.; Stone, M.H.; Ingham, E. A novel method for the prediction of functional biological activity of polyethylene wear debris. *Proc. Inst. Mech. Eng.*, H **2001**, *215*, 127–132.
20. Endo, M.; Tipper, J.L.; Barton, D.C.; Stone, M.H.; Ingham, E.; Fisher, J. Comparison of wear, wear debris and functional biological activity of moderately cross-linked and non-crosslinked polyethylenes in hip prostheses. *Proc. Inst. Mech. Eng.*, H **2002**, *216*, 111–122.
21. Tipper, J.L.; Firkins, P.J.; Besong, A.A.; Barbour, P.S.M.; Nevelos, J.; Stone, M.H.; Ingham, E.; Fisher, J. Characterisation of wear debris from UHMWPE on zirconia ceramic, metal-on-metal and alumina ceramic-on-ceramic hip prostheses generated in a physiological anatomical hip joint simulator. *Wear* **2001**, *250*, 120–128.
22. McKellop, H.; Shen, F.; Lu, B.; Campbell, P.; Salovey, R. Development of extremely wear resistant UHMWPE for total hip replacement. *J. Bone Jt. Surg.* **1999**, *81A*, 157–167.
23. Muratoglu, O.K.; Bragdon, C.R.; O'Connor, D.O.; Harris, W. Unified wear model for highly crosslinked UHMWPE. *Biomaterials* **1999**, *20*, 1463–1478.
24. Kurtz, S.M.; Muratoglu, O.K.; Evan, M.; Edidin, A. Advances in processing, sterilisation and crosslinking of ultra high molecular weight polyethylene. *Biomaterials* **1999**, *81A*, 157–167.
25. Galvin, A.; Tipper, J.L.; Stone, M.H.; Ingham, E.; Fisher, J. Reduction in wear of crosslinked polyethylene under different conditions. Refining future strategies in hip replacement. *Proc. Inst. Mech. Eng.* **2002**, *C601/005*.
26. Ingram, J.H.; Fisher, J.; Stone, M.H.; Ingham, E. Effect of crosslinking on biological activity of UHMWPE wear debris. *Trans. 49th Orthop. Res. Soc.* **2003**, *28*, 1439.
27. Boutin, P.; Christel, P.S.; Dorlot, J.M.; Meunier, A.; deRoquancourt, B.D.; Herman, S.; Sedel, L.; Witvoet, J. The use of dense alumina–alumina combination in total hip replacement. *J. Biomed. Mater. Res.* **1988**, *22*, 1203–1209.
28. McKellop, H.; Park, S.H.; Chiesa, R.; Doorn, P.; Normand, B.L.P.; Grigoris, P.; Amstutz, H. In vivo wear of three types of metal on metal hip prostheses during 2 decades of use. *Clin. Orthop. Relat. Res.* **1966**, *329*, S128–S140.
29. Nevelos, J.E.; Prudhommeaux, F.; Hamadouche, M.; Doyle, C.; Ingham, E.; Meunier, A.; Nevelos, A.B.; Sedel, L.; Fisher, J. Comparative analysis of two different types of alumina–alumina hip prosthesis retrieved for aseptic loosening. *J. Bone Jt. Surg.* **1999**, *83-B* (4), 598–603.
30. Nevelos, J.; Ingham, E.; Doyle, C.; Streicher, R.; Nevelos, A.; Walter, W.; Fisher, J. Microseparation of the centers of alumina–alumina artificial hip joints during simulator testing produces clinically relevant wear and patterns. *J. Arthroplast.* **2000**, *15*, 793–795.
31. Nevelos, J.E.; Ingham, E.; Doyle, C.; Nevelos, A.B.; Fisher, J. Analysis of retrieved alumina ceramic components for Mittelmeier total hip prostheses. *Biomaterials* **1999**, *20*, 1833–1840.
32. Nevelos, J.E.; Ingham, E.; Doyle, C.; Nevelos, A.B.; Fisher, J. The influence of acetabular cup angle on the

- wear of "BIOLOX Forte" alumina ceramic bearing couples in a hip joint simulator. *J. Mater. Sci., Mater. Med.* **2001**, *12*, 141–144.
33. Nevelos, J.E.; Ingham, E.; Doyle, C.; Nevelos, A.B.; Fisher, J. Wear of HIPed and non-HIPed alumina–alumina hip joints under standard and severe simulator testing conditions. *Biomaterials* **2001**, *22*, 2191–2197.
 34. Tipper, J.L.; Hatton, A.; Nevelos, J.E.; Ingham, E.; Doyle, C.; Streicher, R.; Nevelos, A.B.; Fisher, J. Alumina–alumina artificial hip joints. Part II: Characterisation of the wear debris from in vitro hip joints simulations. *Biomaterials* **2002**, *23*, 3441–3448.
 35. Hatton, A.; Nevelos, J.E.; Nevelos, A.A.; Banks, R.E.; Fisher, J.; Ingham, E. Alumina–alumina artificial hip joints. Part I: A histological analysis and characterisation of wear debris by laser capture microdissection of tissues retrieved at revision. *Biomaterials* **2002**, *23*, 3429–3440.
 36. Stewart, T.; Tipper, J.L.; Streicher, R.; Ingham, E.; Fisher, J. Long-term wear of HIPed alumina on alumina bearings for THR under microseparation conditions. *J. Mater. Sci., Mater. Med.* **2001**, *12*, 1053–1056.
 37. Ingham, E.; Hatton, A.; Matthews, J.B.; Nevelos, J.E.; Fisher, J. Osteolytic potential of clinically relevant alumina ceramic wear debris. *Trans. 49th Orthop. Res. Soc.* **2003**, *28*, 212.
 38. Tipper, J.L.; Matthews, J.B.; Ingham, E.; Stewart, T.D.; Fisher, J.; Stone, M.H. Wear and Functional Biological Activity of Wear Debris Generated from UHMWPE-on-Zirconia Ceramic, Metal-on-Metal, and Alumina Ceramic-on-Ceramic Hip Prostheses During Hip Simulator Testing. In *Friction, Lubrication and Wear of Artificial Joints*; PEP: London; 2002; 7–28.
 39. Doorn, P.F.; Mirra, J.M.; Campbell, P.A.; Amstutz, H.C. Tissue reaction to metal on metal total hip prostheses. *Clin. Orthop.* **1996**, *329*, S187–S205.
 40. Jacobs, J.J.; Skipor, A.K.; Patterson, L.M.; Hallab, N.J.; Paprosky, W.G.; Black, J.; Galante, J.O. Metal release in patients who have had a primary total hip arthroplasty. A prospective, controlled, longitudinal study. *J. Bone Jt. Surg.* **1998**, *80A*, 1447–1458.
 41. Tipper, J.L.; Firkins, P.J.; Ingham, E.; Stone, M.H.; Farrar, R.; Fisher, J. Quantitative analysis of the wear and wear debris from low and high carbon content cobalt chrome alloys used in metal on metal total hip replacements. *J. Mater. Sci., Mater. Med.* **1999**, *10*, 355–362.
 42. Firkins, P.J.; Tipper, J.L.; Saadatizadeh, M.R.; Ingham, E.; Stone, M.H.; Farrar, R.; Fisher, J. Quantitative analysis of wear and wear debris from metal-on-metal hip prostheses tested in a physiological hip joint simulator. *Bio-Med. Mater. Eng.* **2001**, *11*, 143–157.
 43. Firkins, P.J.; Tipper, J.L.; Ingham, E.; Stone, M.H.; Farrar, R.; Fisher, J. Influence of simulator kinematics on the wear of metal-on-metal hip prostheses. *Proc. Inst. Mech. Eng., H* **2001**, *215*, 119–121.
 44. Medley, J.B.; Chan, F.W.; Krygier, J.J.; Bobyn, J.D. Comparison of alloys and designs in a hip simulator study of metal on metal implants. *Clin. Orthop. Relat. Res.* **1996**, *329*, S148–S159.
 45. Stewart, T.; Butterfield, M.; Stone, M.H.; Williams, S.; Ingham, E.; Fisher, J. Wear of metal metal and ceramic ceramic hip prostheses under swing phase microseparation. *Trans. 48th Orthop. Res. Soc.* **2002**, *48*, 128.
 46. Jin, Z.M.; Dowson, D.; Fisher, J. Analysis of fluid film lubrication in artificial hip joint replacements with surfaces of high elastic modulus. *Proc. Inst. Mech. Eng., H* **1997**, *211*, 247–256.
 47. Germain, M.A.; Hatton, A.; Williams, S.; Matthews, J.B.; Stone, M.H.; Fisher, J.; Ingham, E. Comparison of cytotoxicity of clinically relevant cobalt chrome and alumina ceramic wear particles. *Biomaterials* **2003**, *24*, 469–479.
 48. Goodfellow, J. Knee prostheses—One step forward two steps back. *J. Bone Jt. Surg.* **1992**, *74B*, 1–2.
 49. Jones, S.M.G.; Pinder, I.M.; Moran, C.A.; Malcolm, A.J. Polyethylene wear in uncemented knee replacements. *J. Bone Jt. Surg.* **1992**, *74B*, 18–22.
 50. Fisher, J. Wear of ultra high molecular weight polyethylene in total artificial joints. *Curr. Orthop.* **1994**, *8*, 164–169.
 51. Bell, C.J.; Walker, P.G.; Blunn, G.W. The effect of oxidation on the delamination of UHMWPE tibial components. *J. Arthroplast.* **1998**, *13*, 280–290.
 52. Reeves, E.A.; Barton, D.C.; FitzPatrick, D.; Fisher, J. Comparison of gas plasma and gamma irradiation in air sterilization on the delamination wear of the ultra-high molecular weight polyethylene used in knee replacements. *Proc. Inst. Mech. Eng., H* **2000**, *214*, 249–255.
 53. Fisher, J.; McEwen, H.M.J.; Barnett, P.I.; Bell, C.J.; Stewart, T.D.; Stone, M.H.; Ingham, E. Mini-symposium: Total knee replacement—Practical considerations (i) Wear of polyethylene in artificial knee joints. *Curr. Orthop.* **2001**, *15*, 399–405.
 54. Schai, P.; Thornhill, T.; Scott, R.D. Total knee arthroplasty with the PFC system. *J. Bone Jt. Surg.* **1998**, *80B*, 850–851.
 55. Howling, G.I.; Barnett, P.I.; Tipper, J.L.; Stone, M.H.; Fisher, J.; Ingham, E. Quantitative characterization of polyethylene debris isolated from periprosthetic tissue in early failure knee implants and early and late failure Charnley hip implants. *J. Biomed. Mater. Res. Appl. Biomater.* **2001**, *58*, 415–420.
 56. Walker, P.S.; Blunn, G.W.; Perry, J.P.; Bell, C.J.; Campbell, P. Methodology for long term wear testing of knee replacements. *Clin. Orthop. Relat. Res.* **2000**, *372*, 290–301.
 57. Benson, L.C.; Des Jardins, J.; Harman, M.K.; LaBerge, M. Effect of stair descent loading on UHMWPE wear in a force controller knee simulator. *Proc. Inst. Mech. Eng., H* **2002**, *216*, 409–418.
 58. Burgess, I.; Kolan, M.; Cunningham, J.; Unsworth, A. Development of a six station knee wear simulator. *Proc. Inst. Mech. Eng., H* **1997**, *211*, 37–47.
 59. McEwen, H.M.J.; McNulty, D.E.; Auger, D.D.; Farrar, R.; Liao, Y.S.; Stone, M.H.; Fisher, J. Wear—Analysis of Mobile Bearing Knee. In *LCS Mobile Bearing Knee Arthroplasty—A 25 Years Worldwide Review*; Hamelynck, S., Ed.; Springer: Amsterdam; 2002; 67–73. Chapter 8.1.

60. Barnett, P.I.; McEwen, H.M.J.; Auger, D.D.; Stone, M.H.; Ingham, E.; Fisher, J. Investigation of wear of knee prostheses in a new displacement/force-controlled simulator. *Proc. Inst. Mech. Eng., H* **2002**, *216*, 51–61.
61. Barnett, P.I.; Fisher, J.; Auger, D.D.; Stone, M.H.; Ingham, E. Comparison of wear in a total knee replacement under different kinematic conditions. *J. Mater. Sci., Mater. Med.* **2001**, *12*, 1039–1042.
62. McEwen, H.M.J.; Fisher, J.; Goldsmith, A.A.J.; Auger, D.D.; Hardaker, C.; Stone, M.H. Wear of fixed bearing and rotating platform mobile bearing knees subjected to high levels of internal and external tibial rotation. *J. Mater. Sci., Mater. Med.* **2001**, *12*, 1049–1052.
63. Besong, A.A.; Tipper, J.L.; Stone, M.H.; Ingham, E.; Fisher, J. The influence of joint kinematics on the number and morphology of polyethylene wear particles in models of hip and knee prostheses. *Trans. IMechE Conf. Knee Replace. C* **1999**, 70–73.

Wear Debris, Bone Resorption Animal Models

Xiaochao Wei
Regis J. O'Keefe
Edward M. Schwarz

*The Center for Musculoskeletal Research, University of Rochester, Rochester,
New York, U.S.A*

INTRODUCTION

Although total joint replacement is one of the most successful procedures in modern medicine, many patients will suffer from aseptic loosening of their implants. This complication is a result of periprosthetic osteolysis, which has been documented in up to 20% of patients at ten years.^[1] The osteolysis in aseptic loosening is caused by both biomechanical and biological factors.^[2,3] Motion at the bearing surfaces of the prosthetic joint results in the generation of sub-micrometer particles composed of polyethylene and metallic debris. These implant wear particles from the artificial joint accumulate in the synovium and activate cells in the periprosthetic membrane surrounding the joint. A variety of cell responses triggered by these particles results in the activation and survival of osteoclasts and leads to bone resorption. Thus, the process is initiated and propagated by wear debris particles.

PROINFLAMMATORY CELLULAR RESPONSE TO WEAR DEBRIS

Macrophage and Foreign Body Reaction

As the first line of the innate immune defense, macrophages act as a housekeeper to get rid of any abnormal or non-self entities encountered in the body. Suggested by their name, these cells take advantage of their specific function called phagocytosis, which is the engulfment of foreign particles that can be as large as 10 micrometers in diameter. Phagocytosis of inert particles leads to a foreign body reaction, which is generally regarded as a secondary response following the unsuccessful attempt of macrophages to remove the extracellular materials. The foreign body reaction is characterized by the formation of giant cells, encapsulation of the implant, and chronic inflammation. Foreign body giant cells are formed from the fusion of many individual macrophages when a large object such as an artificial joint implant is encountered. They can secrete degradative agents such as superoxides and free

radicals, causing damage to implants, while at the same time inducing the infiltration of inflammatory cells. In periprosthetic osteolysis, macrophages and foreign body giant cells tend to remain at the interfacial membrane around the implant and form granulation tissue. Initial studies in this field focused on tissues retrieved from patients with revision implant surgery and demonstrated the correlation of the increasing number of macrophages with the deteriorated clinical grade of loosening and osteolysis, suggesting the important roles of macrophages.^[4] The next line of investigation focused on the proinflammatory molecules produced by macrophages following the phagocytosis of wear debris particles. In vitro studies using peripheral blood mononuclear cells (PBMC) and different macrophages cell lines demonstrated a series of cell responses, including the production of: interleukin-1 (IL-1), prostaglandin E2 (PGE2), tumor necrosis factor- α (TNF- α), interleukin-6 (IL-6), lysosomal enzymes, nitric oxide synthase, cyclooxygenase-2 (COX-2), monocyte chemoattractant protein-1, monocyte inflammatory protein-1 α , and RANTES.^[4-9] These osteoclastogenic factors will participate in the inflammatory processes leading into bone resorption, as discussed below. The direct interaction of particles with macrophages was also examined by electron microscopy, revealing that in vitro macrophages tend to surround and attach to the surface of the larger particles and phagocytose the smaller particles within 2 hr.^[10]

In addition to the production of proinflammatory factors that induce osteolysis, a direct role in bone resorption by macrophages has also been suggested from in vitro organ (calvaria) culture studies that demonstrated the bone-resorbing ability of activated macrophages by wear particles.^[11,12] As such, the macrophage has been considered not only an initiator of inflammation mediated osteolysis, but also an important executor of bone digestion.

Tumor Necrosis Factor and Interleukins

Tumor necrosis factor- α is an important agent in human diseases.^[13] The fact that TNF- α has potent

W

and diversified biological effects on cells including apoptosis, host defense, inflammation, and development rendered it a critical factor in both physiology and pathology. The function of TNF- α depends on the interaction of this soluble ligand with its cognate cellular receptors (TNFR1 and TNFR2), activating intracellular signaling events that lead to various biological consequences depending on the circumstances.

One of the first factors investigated in the pro-inflammatory response to wear debris was TNF- α .^[8,9] First, these studies demonstrated that it is an abundant cytokine released by wear debris-activated macrophages/monocytes. Histology studies on periprosthetic membranes indicated increased TNF- α expression in patients with implant loosening with osteolysis vs. patients without osteolysis.^[14] Moreover, the importance of TNF- α in wear debris-induced osteolysis was unambiguously supported by *in vivo* experiments in the mouse calvaria model.^[15] These studies indicated that mice genetically designed to overproduce TNF- α (human TNF- α transgenic mice) demonstrated an exaggerated response, whereas mice defective in TNF- α signaling (TNF-receptor I knockout mice) had significantly less inflammation and bone resorption in response to wear debris. Further investigation using the soluble TNF- α inhibitor etanercept (Enbrel) also demonstrated that blockade of TNF- α function *in vivo* not only prevented the osteolysis induced by wear particles, but also ameliorated the established osteolysis.^[16] All together, this evidence indicates a critical role for TNF- α in aseptic loosening.

Another class of cytokines believed to be critically involved in wear debris-induced osteolysis is the interleukins. Interleukins are hormone-like proteins that mainly regulate the immune system. Normally secreted by lymphocytes, interleukins can activate receptors on cells to influence a variety of inflammation-associated cell behaviors including cell proliferation, cell-cell contact, lymphocyte activation, cell differentiation, T-cell attraction, and other cytokine production. Interleukins also play fundamental roles in many other aspects of homeostasis.^[17] Not surprisingly, interleukins are among the cytokines associated with prosthetic osteolysis reported in patients with aseptic loosening.

IL-1 was first reported in macrophages located in periprosthetic membranes obtained from patients during revised joint replacement.^[5] Subsequently, it was demonstrated that IL-1 is released from macrophages following stimulation with wear debris and mediates increased bone resorption activity *in vitro*.^[11] Similarly, a remarkable level of IL-6 was observed in synovial fluid from failed total joint replacement patients. In a canine model of aseptic loosening, IL-6 was reported in the implant interface following the addition of particles.^[18] Moreover, IL-6 was reported

in many cells challenged with particles *in vitro*,^[8,9] including macrophages, osteoblasts, and fibroblasts, suggesting a broad range of activity. Interestingly, cocultures of macrophages with fibroblasts together significantly increase the production of IL-6, particularly compared with particle-activated macrophages alone, indicating a synergistic effect in the inflammatory response from macrophage-fibroblast interaction in the periprosthetic membrane. The anti-inflammatory mediator IL-10 was also investigated in human PBMC activated by particles. IL-10 can block the production of TNF- α and IL-6 and potentially down-regulate the cytokine response to wear debris. In therapeutic animal models, the viral gene delivery of either IL-1 inhibitor (human IL-1 receptor antagonist) or IL-10 can directly reduce the local inflammation elicited by particles, indicating pivotal roles played by interleukins in the pathology of aseptic loosening.^[19]

In summary, during the foreign body reaction around orthopedic implants, phagocytosis of wear particles by macrophages can release a series of cytokines, including TNF- α , IL-1, and IL-6, to promote an inflammatory network around the implant. By activating other cells in the vicinity and attracting immune cells including T cells, this proinflammation phase can become chronic, leading to a sustained bone resorption phase and aseptic loosening.

OSTEOCLASTOGENESIS INDUCED BY WEAR DEBRIS

Osteoclasts and Rank Ligand

In adult life, bone is continuously turned over, or remodeled, by bone-forming cells (osteoblasts) and bone-resorbing cells (osteoclasts).^[20,21] An imbalance in remodeling often leads to bone defects. Osteoporosis occurs when resorption exceeds formation, whereas osteopetrosis is caused by a lack of bone resorption. In both cases, the number and activity of osteoclasts are usually the key features of the level of bone resorption and therefore are central to the pathogenesis of bone disease. Osteoclasts are derived from the monocyte-macrophage lineage of the hematopoietic stem cell. Osteoclasts form via cell fusion of mononuclear pre-osteoclasts to become multinuclear giant cells and then form specialized structures (i.e., sealing zone, ruffle border) that allow them to firmly attach to bone and secrete acid and degradative enzymes to dissolve the mineralized bone matrix.^[20,21] Osteoclast formation normally requires cell contact from stromal cells or osteoblasts *in vivo*. Macrophage clonal stimulating factor (M-CSF) and receptor activator of nuclear factor- κ B ligand (RANKL) are two main factors that drive the differentiation of osteoclasts. Whereas

M-CSF works as a growth factor for permissive proliferation of osteoclast precursors (OCP), RANKL serves as a differentiation factor specific for osteoclast formation.^[22]

RANKL is a member of the TNF/TNR receptor (TNFR) superfamily and is present primarily as a membrane protein produced by osteoblasts/stromal cells. Its receptor, receptor activator of nuclear factor- κ B (RANK), is located on macrophage/osteoclast precursors and, upon activation by RANKL, stimulates the expression of genes required for osteoclast differentiation. In order to downregulate bone resorption, stromal cells and osteoblast also produce a soluble decoy receptor of RANKL called OPG (osteoprotegerin), which specifically blocks RANK/RANKL signaling.^[20–22] Both gain- and loss-of-function experiments using genetically altered mice established RANKL/RANK as essential determinants for osteoclast formation. The absence of RANKL or overproduction of OPG in mice results in severe osteopetrosis and loss of osteoclasts, whereas the overexpression of RANKL or deletion of OPG induces osteoporosis by excessive osteoclast formation.

Various stromal cell populations express RANKL. RANKL is induced in osteoblasts by many systemic regulators such as parathyroid hormone, vitamin D₃, and the IL-6 family of cytokines and is inhibited by transforming growth factor- β (TGF- β). RANKL expression is associated with many inflammation-associated joint diseases, including rheumatoid arthritis. Synovial fibroblasts cultured from rheumatoid arthritis patients expressed RANKL in response to vitamin D₃. The fact that high levels of RANKL are present in rheumatoid synovial tissue indicates that the inflammatory cell reaction in synovial tissue is essential in the pathogenesis of bone resorption, which likely involves RANKL expression.^[23]

Overwhelming bone resorption is the key feature in the pathogenesis of aseptic loosening, indicating the pivotal role played by osteoclastogenesis. Thus, RANKL is considered an essential regulator in osteolysis induced by wear debris particles. Experiments revealed that either bone marrow cell culture or synovial fibroblasts can increase RANKL expression in response to wear particles *in vitro*^[24] and the enhancement of both RANKL expression and osteoclastogenesis is associated with peri-implant tissue of patients in whom the dramatic osteolytic region is observed.

The importance of RANKL and osteoclasts for aseptic loosening is further indicated by the results of animal experiments in which osteolysis was evaluated in mice lacking RANK/RANKL signaling.^[25–27] First, in a gene therapy model, particles stimulated a sixfold increase in osteoclastogenesis compared with controls without particles, which was efficiently prevented by

the exogenous OPG.^[25,27] Second, RANK^{-/-} mice and mice treated with RANK:Fc are completely resistant to wear debris-induced osteolysis because they have no osteoclasts.^[26] Finally, the same suppression is demonstrated in mice deficient in intracellular downstream effector of RANK (NF- κ B1 knockout mice).^[15]

Thus, preclinical studies formally demonstrated that osteoclastogenesis is a therapeutic target for the prevention of osteolysis in aseptic loosening. For example, the osteoclast inhibitor alendronate controlled the implant failure induced by osteolysis.^[28] Similarly, drugs that inhibit RANKL are also strong candidates for therapeutic interventions in aseptic loosening.

Cytokines and Osteoclast

For a long time, IL-1 and TNF- α have been considered important regulators of osteoclast formation, as well as proinflammation.^[20] Both cytokines can induce bone resorption and osteoclast-like cell formation in murine and human marrow cultures indirectly by inducing RANKL in effector cells and directly stimulating osteoclast differentiation. IL-1 can induce hypercalcemia (a symptom with high calcium release from bone indicating more bone loss) and enhance the growth and differentiation of early osteoclast precursors. IL-1 also increases the number of more committed OCP and stimulates their maturation to bone-resorbing osteoclasts.^[20] Similarly, TNF- α affects the activity of mature osteoclasts and markedly stimulates the induction of OCP in which CD11b is highly expressed.^[29] In addition, TNF- α potentiates the effects of IL-1 on osteoclast formation.^[20] IL-6 can stimulate *in vitro* bone resorption in fetal mouse calvaria and plays roles in systemic osteoclast recruitment and RANKL regulation.^[20,30] Although IL-6-deficient mice do not develop obvious bone abnormality, they are protected from bone loss after estrogen depletion, indicating that IL-6 might be involved in postmenopausal osteoporosis.^[31] As described previously, all of these cytokines are observed in the proinflammation phase of the cell response to wear particles and are implicated in the pathogenesis of aseptic loosening. Thus, it was hypothesized that this synergistic inflammatory microenvironment around bone is associated with osteoclast activation, followed by increasing bone resorption.

TGF- α is a potent inhibitory cytokine for osteoclastic bone resorption.^[20] TGF- β is released and activated during osteoclastic bone resorption, suggesting that TGF- β may be an important coupling factor that induces subsequently new bone formation after inhibiting excess osteoclast formation.^[20] The decrease in total TGF- β observed in the particle-exposed animal group therefore might indicate another mechanism



involving both increasing bone resorption and the suppression of bone formation.^[32]

Cyclooxygenase-2 and Prostaglandin E2

Among factors observed in periprosthetic membranes retrieved from patients with loose, non-septic failed hip replacements, prostaglandins represent another predominant player involved in osteolysis. PGE2 is a primary product of arachidonic acid metabolism in many cells. The COX enzymes mediate a critical step in prostaglandin biosynthesis and include two major isoforms,^[33] COX-1, which is constitutive, and COX-2, which is inducible. Non-steroidal anti-inflammatory drugs, which block both COX-1 and COX-2, have been used successfully to reduce inflammatory responses to a variety of different stimuli. More recently, specific inhibitors for COX-2 have been introduced that have the ability to selectively inhibit the COX-2 isoform and avoid potential complications associated with the blockade of COX-1-mediated events.^[33] There are also four different types of protein membrane receptors for PGE2, named EP1–4.^[34] All of them are members of the G-protein-coupled receptor family. Different species of G proteins are associated with different EP receptors, linking them with the respective downstream gene regulation.^[34] Such differences in the intracellular signaling pathway and downstream targets result in complex PGE2-induced signals in different cell types *in vivo*.

In addition to its role as a proinflammatory mediator, PGE2 is also important in bone regulation. In terms of osteoclastogenesis, PGE2 is normally thought of as a strong activator, although its mechanism of action is not fully understood and contradictory evidence exists.^[20] Analysis of periprosthetic membranes revealed that synovial-like lining cells synthesized a large amount of PGE2, which was consistent with their rheumatoid synovial tissue-like phenotype.^[6] Similarly, analyses of retrieval tissue from revision arthroplasty cases demonstrated that loose components, whether cemented or not, have statistically higher PGE2 levels in the surrounding membrane compared with the non-loose group.^[7] These all suggest that PGE2 may be associated with the bone lysis associated with prosthesis loosening.

Previous work performed in animals also implicated the importance of PGE2 in wear debris-induced osteolysis. In COX-2^{-/-} mice or in mice treated with COX-2 inhibitors, bone resorption was significantly reduced compared with wild type or COX-1^{-/-} mice, indicating that PGE2 induced by wear debris particles through COX-2 may play an essential role leading to local osteolysis.^[35] PGE2 can be produced by macrophages and fibroblasts of synovial tissue in

inflammatory joint diseases and in cells exposed to wear particles *in vitro*.^[24]

Compared to the strong evidence for the pivotal role of COX-2 in wear debris-induced osteolysis, the mechanism and function of PGE2 is poorly understood. Using various EP receptor knockout mice, it is possible to examine the physiologic role of PGE2 mediated by each receptor. EP2 and EP4 receptors mediate the bone resorption activity of PGE2, because studies reported that osteoclast formation is impaired in cells cultured from EP2- or EP4-deficient mice.^[36,37] Moreover, PGE2 can also induce RANKL expression in osteoblast and stromal cells.^[37] In aseptic loosening, synovial fibroblasts may be another cell target for PGE2 because the evidence indicated that wear particles can induce synovial fibroblasts to express RANKL in a COX-2-dependent manner.^[24]

PGE2 can also induce bone formation by driving osteoblast differentiation. The exogenous addition of PGE2 or receptor agonist can directly induce bone formation both in organ cultures and in the mouse model.^[38,39] Thus, it seems that PGE2 has opposite effects on the regulation of bone turnover. Given the fact that different prostaglandin metabolites could be synthesized through COX-2 and the controversial roles of PGE2 through different downstream receptors, the significance of PGE2 and the underlying mechanisms in osteolysis induced by wear debris must be further clarified.

In summary, following inflammation stimulated by particles, a bone resorption phase is gradually developed, involving the enhancement of osteoclastogenesis. Many factors are involved in this process including RANKL, cytokines, and COX-2. By increasing OCP numbers and promoting their proliferation, differentiation, and activation to bone-resorbing cells, these factors increase the number and activity of the osteoclasts surrounding the prosthesis to mediate osteolysis and aseptic loosening.

Animal Models of Wear Debris-Induced Osteolysis

Central to the investigation of the etiology of human disease is the development of animal models that recapitulate *in vivo* pathology. Because the cellular response to wear debris particles is believed to be the basis for the commencement and development of aseptic loosening, a variety of *in vivo* bone resorption models using dogs, rabbits, rats, and mice has been developed.^[15,16,18,19,25–28,32,35,40–54] Combined with *in vitro* studies and clinical observations, these animal models can facilitate our research to further the understanding of wear debris-induced osteolysis and evaluate novel therapeutic interventions. Here we

discuss several of these animal models and describe their advantages and disadvantages.

Animal models of aseptic loosening are divided into two groups. The first comprises larger, long-term models that contained a joint implant designed to closely resemble the clinical situation.^[18,32,40–43] These dogs suffer acute inflammation and bone resorption within 24 hr of surgery and a classic inflammatory reaction that includes macrophage infiltration and the elevation of PGE2 and IL-1 is observed. Furthermore, Shanbhag et al. were able to demonstrate the efficacy of oral bisphosphonate therapy (alendronate) to successfully inhibit wear debris-mediated bone resorption in this model.^[40]

These implant models have also been useful toward understanding other aspects of total joint replacement. For instance, a comparison between non-cemented smooth and porous implant surfaces was made^[41] and femoral–acetabular component interactions were reported.^[42] Also, an implant coated with hydroxyapatite prevented the migration of particles along the interface and demonstrated significant benefits in a canine total hip arthroplasty (THA) model.^[43] Others used a rabbit model as an alternative for the long-term effects of wear debris-mediated bone resorption.^[44,45] In this model, particles can be directly introduced through a drill hole into the proximal tibia of rabbits, or submicrometer particles of UHMWPE are placed around pressfit tibial hemiarthroplasties in rabbits to determine the biological reaction.^[45]

Unfortunately, the ultimate goal of these large animal models, to recapitulate aseptic loosening, has yet to be demonstrated. The results of these studies indicate that the >5 year postop period required for clinical manifestation of aseptic loosening appears to be a requisite for these models as well, because accelerated aseptic loosening of a well-fixed implant has yet to be demonstrated in any *in vivo* model. For example, in a canine total hip arthroplasty model in which synovial membrane biopsies were harvested after surgery, no signs of loosening were reported after 2 years.^[40] Even in a rat model in which the investigators utilized a running wheel in an attempt to accelerate the biomechanical effects, the prostheses were mechanically stable after 6 months, with no evidence of loosening.^[46] Our interpretation of these results is that the host bone healing response to injury in healthy animals is very robust, and attempts to induce loosening of an osteointegrated implant by any acute insult are doomed to failure. However, if an implant model could be devised with roughened components that would be a source of chronic wear debris, this might lead to a model of aseptic loosening.

Thus, it is easy to focus mainly on the biology of osteolysis induced by wear debris in animal

models despite the other critical mechanical and biomechanical components of aseptic loosening. Based on this, other non-implant models were developed, and the particular biology of osteolysis stimulated by wear particles was analyzed. In a rat model with a non-weight-bearing plug inserted at the distal part of the femur, injections of particles into the knee joint elicited bone resorption at this stable interface.^[47] A bone chamber model in the rabbit was also developed.^[44] A bone-harvest chamber with a transverse channel for tissue ingrowth was introduced into the proximal part of the tibia. After a 6-week period of initial incorporation, the contents of the channel were harvested for assay repeatedly. The addition of particles such as polyethylene or cobalt-chromium stimulated a chronic inflammation with the reduction of bone formation and enhancement of bone resorption.^[45] The mitigating effect by local delivery of fibroblast growth factor-2, which can increase bone growth, was also suggested in this model.

Although these large animal models are used successfully, their more general application is significantly limited by their major disadvantages, including the inconvenience of experimental handling (time consuming and complex surgery), their cost, and the scarcity of tools for the system (much less information for dog or rabbit compared with mouse, e.g., a database of genome and transgene animals). For these reasons, two models were developed in the mouse to facilitate research and incorporate a variety of exquisite molecular reagents and genetically defined animal strains. Together with highly quantitative outcome measurements of osteoclast formation and bone resorption, these mouse models can be used in delicate biological studies to pinpoint the major factors involved in wear debris-induced osteolysis.

The mouse calvaria model is perhaps the most widely used animal model in the field.^[28] Because there is only a little soft tissue around the skull bones that contains a large number of osteoclast and osteoblast precursor cells, calvaria are useful for many *in vitro* and *in vivo* studies to evaluate the biological responses of bone. For example, in a study on wear particles, an *in vitro* experiment using neonatal calvaria organ culture has been frequently used to demonstrate the ability of particles to induce bone resorption.^[11,12] In the *in vivo* mouse model, a fixed amount of Ti, polymethylmethacrylate (PMMA), or PE particles are surgically implanted, which initiates inflammatory bone loss after a short period of time (3–7 days). In this manner, the *in vivo* mechanisms responsible for wear debris-induced osteolysis can be readily elucidated. Another advantage of this model is that it can be utilized for drug screening, which often requires a large number of animals to define the effective dose and regimen.

To date, several factors associated with wear particles have been identified using the murine calvaria model. In particular, the roles of RANKL, TNF- α , and NF κ B signaling have been extensively studied in this animal model using knockout mice.^[15,16,25,26] Similarly, the importance of COX-2 and PGE2 has been suggested.^[35] A comparison between PE and Ti particles was performed, indicating that bone resorption response to particles in other materials could also be examined in this model.^[48] For clinical relevance, some drug trials such as those using pentoxifylline or alendronate can significantly reduce inflammatory osteolysis in particle implantation in mice, providing a good example for potential drug development.^[28] Moreover, the efficiency of gene therapy was further tested in this model, suggesting an alternative treatment for the disease.^[25]

The air pouch model is another useful murine model for the evaluation of inflammatory response to wear debris particles in vivo.^[49] Filtered air is injected subcutaneously into the back of mice and a section of femur or calvaria bone is surgically introduced into the established air pouch from a syngenic mouse donor in the presence of debris particles. This model was generally used to assess the cellular and cytokine effects in inflammation processes. Shown as a relevant tissue response to orthopedic materials, a series of resembled reactions including cellular infiltration and production of inflammatory mediators can be observed in the air pouch model compared with the peri-implant tissue. The analysis on harvested air pouch tissue containing wear particles may reveal the basics in the pathogenesis of peri-implant osteolysis.

In this setting of wear-induced inflammation and bone resorption, the involvement of a variety of factors is implicated. Particles in different materials (UHMWPE, PMMA, cobalt-chrome and titanium alloy) and different shapes (elongated or globular) were examined.^[50] Additionally, the gene expression profile was investigated in pouch tissue.^[51] Higher RANKL/RANK ratio, IL-1, and TNF- α were associated with tissue inflammation status and osteoclast-like cell accumulation. The protective effects of anti-inflammation cytokines (IL-10 and IL-1 receptor antagonists and soluble receptor of TNF- α) and the negative regulator for bone resorption (OPG) by gene delivery were also evaluated.^[19,27,52] On the other hand, the locations of wear debris in these models are different from that of joint replacement. Such discrepancy in the adjacent cell population and micro-environment for osteoclast function might result in significant impacts in the interpretation of data.

Apart from these established animal models, more studies and new designs for improvement and refinement are being performed.^[53,54] For example, a novel xenograft-based model was recently used in which

human periprosthetic tissues obtained from patients during revision arthroplasty were transplanted into immune competent mice [severe combined immunodeficiency (SCID) mouse].^[54] Provided that animal experimental models have been successfully used to further understanding of the pathology and developing novel therapies, we hope that animal studies will ultimately lead to an effective intervention for aseptic loosening.

CONCLUSION

Above all, wear debris-induced osteolysis is associated with complex cellular responses that involve many biological factors. These factors likely interact to form a network that strongly promotes local bone loss by enhancing osteoclast formation and activity. Because the inevitable result of wear debris-induced osteolysis is implant failure, which remains one of the major problems in orthopedics, understanding the mechanisms responsible is important. As demonstrated by the evidence from both in vitro and in vivo experiments discussed here, osteolysis induced by wear debris is derived from a chronic inflammatory process followed by an increase in osteoclast formation. TNF- α , interleukins, RANKL, and prostaglandin are the main effector molecules that have been described, and other novel mediators continue to be identified. However, the answers to several important questions remain elusive, including the identity of the initiator messenger following exposure to wear debris. It has been postulated that osteolysis is initiated by the uptake of particles by cells, although the detailed mechanisms that govern this process are not fully clear, and many new factors are being described, such as particle-associated endotoxin or reactive oxygen species as a result of cell damage. Another key question relates to the function of cytokines in the disease. An exact explanation of the role of cytokines is still needed. Researchers must continue to seek these answers from in vitro studies and experiments with animal models.

ACKNOWLEDGMENT

The authors are supported by a Public Health Service Award AR46545.

ARTICLES OF FURTHER INTEREST

Biologic and Synthetic Apatites; Bone-Implant Interface; Bone Remodeling; Ceramics; Ceramics in Dentistry; Composites; Dental Implants; Implant, Total Hip; Knee Joint Replacement; Orthopedic Biomaterials;

Wear Debris from Joint Arthroplasties, Biological Effects Of

REFERENCES

- Glassman, A.H.; Engh, C.A.; Suthers, K.E. Ten year follow-up of total hip arthroplasty extensively porous coated femoral components. *Orthop. Trans.* **1993**, *16*, 702.
- Schmalzried, T.P.; Jasty, M.; Harris, W.H. Periprosthetic bone loss in total hip-arthroplasty-polyethylene wear debris and the concept of the effective joint space. *J. Bone Joint Surg. Am.* **1992**, *74A* (6), 849–863.
- Maloney, W.J.; Smith, R.L.; Schmalzried, T.P.; Chiba, J.; Huene, D.; Rubash, H. Isolation and characterization of wear particles generated in patients who have had failure of a hip-arthroplasty without cement. *J. Bone Joint Surg. Am.* **1995**, *77* (9), 1301–1310.
- Amstutz, H.C.; Campbell, P.; Kossovsky, N.; Clarke, I.C. Mechanism and clinical significance of wear debris-induced osteolysis. *Clin. Orthop. Relat. Res.* **1992**, *276*, 7–18.
- Jiranek, W.A.; Machado, M.; Jasty, M.; Jevsevar, D.; Wolfe, H.J.; Goldring, S.R.; Goldberg, M.J.; Harris, W.H. Production of cytokines around loosened cemented acetabular components—Analysis with immunohistochemical techniques and in-situ hybridization. *J. Bone Joint Surg. Am.* **1993**, *75A* (6), 863–879.
- Goldring, S.R.; Schiller, A.L.; Roelke, M.; O'Neil, D.A.; Harris, W.H. The synovial-like membrane at the bone–cement interface in loose total hip replacement and its proposed role in bone lysis. *J. Bone Joint Surg. Am.* **1983**, *65* (5), 575–584.
- Goodman, S.B.; Chin, R.C.; Chiou, S.S.; Schurman, D.J.; Woolson, S.T.; Masada, M.P. A Clinical–pathologic–biochemical study of the membrane surrounding loosened and nonloosened total hip arthroplasties. *Clin. Orthop. Relat. Res.* **1989**, *244*, 182–187.
- Blaine, T.A.; Rosier, R.N.; Puzas, J.E.; Looney, R.J.; Reynolds, P.R.; Reynolds, S.D.; O'Keefe, R.J. Increased levels of tumor necrosis factor-alpha and interleukin-6 protein and messenger RNA in human peripheral blood monocytes due to titanium particles. *J. Bone Joint Surg. Am.* **1996**, *78A* (8), 1181–1192.
- Nakashima, Y.; Sun, D.H.; Trindade, M.C.; Maloney, W.J.; Goodman, S.B.; Schurman, D.J.; Smith, R.L. Signaling pathways for tumor necrosis factor-alpha and interleukin-6 expression in human macrophages exposed to titanium-alloy particulate debris in vitro. *J. Bone Joint Surg. Am.* **1999**, *81* (5), 603–615.
- Voronov, I.; Santerre, J.P.; Hinek, A.; Callahan, J.W.; Sandhu, J.; Boynton, E.L. Macrophage phagocytosis of polyethylene particulate in vitro. *J. Biomed. Mater. Res.* **1998**, *39* (1), 40–51.
- Glant, T.T.; Jacobs, J.J.; Molnar, G.; Shanbhag, A.S.; Valyon, M.; Galante, J.O. Bone-resorption activity of particulate-stimulated macrophages. *J. Bone Miner. Res.* **1993**, *8* (9), 1071–1079.
- Glant, T.T.; Jacobs, J.J. Response of three murine macrophage populations to particulate debris: Bone resorption in organ cultures. *J. Orthop. Res.* **1994**, *12* (5), 720–731.
- Locksley, R.M.; Killeen, N.; Lenardo, M.J. The TNF and TNF receptor superfamilies: Integrating mammalian biology. *Cell* **2001**, *104*, 487–501.
- Goodman, S.B.; Huie, P.; Song, Y.; Schurman, D.; Maloney, W.; Woolson, S.; Sibley, R. Cellular profile and cytokine production at prosthetic interfaces. Study of tissues retrieved from revised hip and knee replacements. *J. Bone Joint Surg. Br.* **1998**, *80* (3), 531–539.
- Schwarz, E.M.; Lu, A.P.; Goater, J.J.; Benz, E.B.; Kollias, G.; Rosier, R.N.; Puzas, J.E.; O'Keefe, R.J. Tumor necrosis factor-alpha/nuclear transcription factor-kappa B signaling in periprosthetic osteolysis. *J. Orthop. Res.* **2000**, *18* (3), 472–480.
- Childs, L.M.; Goater, J.J.; O'Keefe, R.J.; Schwarz, E.M. Efficacy of etanercept for wear debris-induced osteolysis. *J. Bone Miner. Res.* **2001**, *16* (2), 338–347.
- Dinarello, C.A.; Mier, J.W. Interleukins. *Annu. Rev. Med.* **1986**, *37*, 173–178.
- Jones, L.C.; Frondoza, C.; Hungerford, D.S. Effect of PMMA particles and movement on an implant interface in a canine model. *J. Bone Joint Surg. Br.* **2001**, *83* (3), 448–458.
- Yang, S.; Wu, B.; Mayton, L.; Evans, C.H.; Robbins, P.D.; Wooley, P.H. IL-1Ra and vIL-10 gene transfer using retroviral ameliorates particle-associated inflammation in the murine air pouch model. *Inflamm. Res.* **2002**, *51* (7), 342–350.
- Goodman, G.D. Cell biology of the osteoclast. *Exp. Hematol.* **1999**, *27*, 1229–1241.
- Teitelbaum, S.L. Bone resorption by osteoclasts. *Science* **2000**, *289* (5484), 1504–1508.
- Kong, Y.Y.; Yoshida, H.; Sarosi, I.; Tan, H.L.; Timms, E.; Capparelli, C.; Morony, S.; Oliveira-dos-Santos, A.J.; Van, G.; Itie, A.; Khoo, W.; Wakeham, A.; Dunstan, C.R.; Lacey, D.L.; Mak, T.W.; Boyle, W.J.; Penninger, J.M. OPG is a key regulator of osteoclastogenesis, lymphocyte development and lymph-node organogenesis. *Nature* **1999**, *28* (397), 315–323.
- Crotti, T.N.; Smith, M.D.; Weedon, H.; Ahern, M.J.; Findlay, D.M.; Kraan, M.; Tak, P.P.; Haynes, D.R. Receptor activator NF-kappaB ligand (RANKL) expression in synovial tissue from patients with rheumatoid arthritis, spondyloarthropathy, osteoarthritis, and from normal patients: semiquantitative and quantitative analysis. *Ann. Rheum. Dis.* **2002**, *61* (12), 1047–1054.
- Wei, X.; Zhang, X.; Zuscik, M.J.; Drissi, M.H.; Schwarz, E.M.; O'Keefe, R.J. Fibroblasts express RANKL and support osteoclastogenesis in a COX-2-dependent manner after stimulation with titanium particles. *J. Bone Miner. Res.* **2005**, *20* (7), 1136–1148.
- Goater, J.J.; O'Keefe, R.J.; Rosier, R.N.; Puzas, J.E.; Schwarz, E.M. Efficacy of ex vivo OPG gene therapy in preventing wear debris induced osteolysis. *J. Orthop. Res.* **2002**, *20* (2), 169–173.
- Childs, L.M.; Paschalis, E.P.; Xing, L.P.; Dougall, W.C.; Anderson, D.; Boskey, A.L.; Puzas, J.E.; Rosier,

- R.N.; O'Keefe, R.J.; Boyce, B.F.; Schwarz, E.M. In vivo RANK signaling blockade using the receptor activator of NF- κ B:Fc effectively prevents and ameliorates wear debris-induced osteolysis via osteoclast depletion without inhibiting osteogenesis. *J. Bone Miner. Res.* **2002**, *17* (2), 192–199.
27. Yang, S.Y.; Mayton, L.; Wu, B.; Goater, J.J.; Schwarz, E.M.; Wooley, P.H. Adeno-associated virus-mediated osteoprotegerin gene transfer protects against particulate polyethylene-induced osteolysis in a murine model. *Arthritis Rheum.* **2002**, *46* (9), 2514–2523.
 28. Schwarz, E.M.; Benz, E.B.; Lu, A.P.; Goater, J.J.; Mollano, A.V.; Rosier, R.N.; Puzas, J.E.; O'Keefe, R.J. Quantitative small-animal surrogate to evaluate drug efficacy in preventing wear debris-induced osteolysis. *J. Orthop. Res.* **2000**, *18* (6), 849–855.
 29. Li, P.; Schwarz, E.M.; O'Keefe, R.J.; Ma, L.; Looney, R.J.; Ritchlin, C.T.; Boyce, B.F.; Xing, L. Systemic tumor necrosis factor alpha mediates an increase in peripheral CD11bhigh osteoclast precursors in tumor necrosis factor alpha-transgenic mice. *Arthritis Rheum.* **2004**, *50* (1), 265–276.
 30. Wong, P.K.; Campbell, I.K.; Egan, P.J.; Ernst, M.; Wicks, I.P. The role of the interleukin-6 family of cytokines in inflammatory arthritis and bone turnover. *Arthritis Rheum.* **2003**, *48* (5), 1177–1189.
 31. Poli, V.; Balena, R.; Fattori, E.; Markatos, A.; Yamamoto, M.; Tanaka, H.; Ciliberto, G.; Rodan, G.A.; Costantini, F. Interleukin-6 deficient mice are protected from bone loss caused by estrogen depletion. *EMBO J.* **1994**, *13* (5), 1189–1196.
 32. Sacomen, D.; Smith, R.L.; Song, Y.; Fornasier, V.; Goodman, S.B. Effects of polyethylene particles on tissue surrounding knee arthroplasties in rabbits. *J. Biomed. Mater. Res.* **1998**, *43* (2), 123–130.
 33. Marnett, L.J.; Rowlinson, S.W.; Goodwin, D.C.; Kalgutkar, A.S.; Lanzo, C.A. Arachidonic acid oxygenation by COX-1 and COX-2. Mechanisms of catalysis and inhibition. *J. Biol. Chem.* **1999**, *274* (33), 22903–22906.
 34. Kobayashi, T.; Narumiya, S. Function of prostanoid receptors: Studies on knockout mice. *Prostaglandins Other Lipid Mediat.* **2002**, *68–69*, 557–573.
 35. Zhang, X.P.; Morham, S.G.; Langenbach, R.; Young, D.A.; Xing, L.; Boyce, B.F.; Puzas, J.E.; Rosier, R.N.; O'Keefe, R.J.; Schwarz, E.M. Evidence for a direct role of cyclo-oxygenase 2 in implant wear debris-induced osteolysis. *J. Bone Miner. Res.* **2001**, *16* (4), 660–670.
 36. Li, X.; Okada, Y.; Pilbeam, C.C.; Lorenzo, J.A.; Kennedy, C.R.; Breyer, R.M.; Raisz, L.G. Knockout of the murine prostaglandin EP2 receptor impairs osteoclastogenesis in vitro. *Endocrinology* **2000**, *141* (6), 2054–2051.
 37. Suzawa, T.; Miyaura, C.; Inada, M.; Maruyama, T.; Sugimoto, Y.; Ushikubi, F.; Ichikawa, A.; Narumiya, S.; Suda, T. The role of prostaglandin E receptor subtypes (EP1, EP2, EP3, and EP4) in bone resorption: An analysis using specific agonists for the respective EPs. *Endocrinology* **2000**, *141* (4), 1554–1559.
 38. Paralkar, V.M.; Borovecki, F.; Ke, H.Z.; Cameron, K.O.; Lefker, B.; Grasser, W.A.; Owen, T.A.; Li, M.; DaSilva-Jardine, P.; Zhou, M.; Dunn, R.L.; Dumont, F.; Korsmeyer, R.; Krasney, P.; Brown, T.A.; Plowchalk, D.; Vukicevic, S.; Thompson, D.D. An EP2 receptor-selective prostaglandin E2 agonist induces bone healing. *Proc. Natl. Acad. Sci. USA* **2003**, *100* (11), 6736–6740.
 39. Yoshida, K.; Oida, H.; Kobayashi, T.; Maruyama, T.; Tanaka, M.; Katayama, T.; Yamaguchi, K.; Segi, E.; Tsuboyama, T.; Matsushita, M.; Ito, K.; Ito, Y.; Sugimoto, Y.; Ushikubi, F.; Ohuchida, S.; Kondo, K.; Nakamura, T.; Narumiya, S. Stimulation of bone formation and prevention of bone loss by prostaglandin E EP4 receptor activation. *Proc. Natl. Acad. Sci. USA* **2002**, *99* (7), 4580–4585.
 40. Shanbhag, A.S.; Hasselman, C.T.; Rubash, H.E. Inhibition of wear debris mediated osteolysis in a canine total hip arthroplasty model. *Clin. Orthop. Relat. Res.* **1997**, *344*, 33–43.
 41. Bobyn, J.D.; Jacobs, J.J.; Tanzer, M.; Urban, R.M.; Aribindi, R.; Sumner, D.R.; Turner, T.M.; Brooks, C.E. The susceptibility of smooth implant surfaces periimplant fibrosis and migration of polyethylene wear debris. *Clin. Orthop. Relat. Res.* **1995**, *311*, 21–39.
 42. Skurla, C.P.; James, S.P. Postmortem retrieved canine THR: Femoral and acetabular component interaction. *Biomed. Sci. Instrum.* **2004**, *40*, 255–260.
 43. Rahbek, O.; Overgaard, S.; Lind, M. Sealing effect of hydroxyapatite coating on peri-implant migration of particles. An experimental study in dogs. *J. Bone Joint Surg. Br.* **2001**, *83* (3), 441–447.
 44. Goodman, S.; Aspenberg, P.; Song, Y.; Knoblich, G.; Huie, P.; Regula, D.; Lidgren, L. Tissue ingrowth and differentiation in the bone-harvest chamber in the presence of cobalt-chromium-alloy and high-density-polyethylene particles. *J. Bone Joint Surg. Am.* **1995**, *77A* (7), 1025–1035.
 45. Goodman, S.B.; Davidson, J.A.; Song, Y.; Martial, N.; Fornasier, V.L. Histomorphological reaction of bone to different concentrations of phagocytosable particles of high-density polyethylene and Ti-6Al-4V alloy in vivo. *Biomaterials* **1996**, *17* (20), 1943–1947.
 46. Pap, G.; Machner, A.; Rinnert, T.; Horler, D.; Gay, R.E.; Schwarzberg, H.; Neumann, W.; Michel, B.A.; Gay, S.; Pap, T. Development and characteristic of a synovial-like interface membrane around cemented tibial hemiarthroplasties in a novel rat model of aseptic prosthesis loosening. *Arthritis Rheum.* **2001**, *44* (4), 956–963.
 47. Howie, D.W.; Vernon-Roberts, B.; Oakeshott, R.; Manthey, B. A rat model of resorption of bone at the cement–bone interface in the presence of polyethylene wear particles. *J. Bone Joint Surg. Am.* **1988**, *70* (2), 257–263.
 48. von Knoch, M.; Jewison, D.E.; Sibonga, J.D.; Sprecher, C.; Morrey, B.F.; Loer, F.; Berry, D.J.; Scully, S.P. The effectiveness of polyethylene versus titanium particles in inducing osteolysis in vivo. *J. Orthop. Res.* **2004**, *22* (2), 237–243.
 49. Wooley, P.H.; Morren, R.; Andary, J.; Sud, S.; Yang, S.Y.; Mayton, L.; Markel, D.; Sieving, A.; Nasser, S. Inflammatory response to orthopaedic biomaterials in the murine air pouch. *Biomaterials* **2002**, *23* (2), 517–526.

50. Yang, S.Y.; Ren, W.; Park, Y.; Sieving, A.; Hsu, S.; Nasser, S.; Wooley, P.H. Diverse cellular and apoptotic responses to variant shapes of UHMWPE particles in a murine model of inflammation. *Biomaterials* **2002**, *23* (17), 3535–3543.
51. Ren, W.; Yang, S.Y.; Fang, H.W.; Hsu, S.; Wooley, P.H. Distinct gene expression of receptor activation of nuclear factor-kappaB and rank ligand in the inflammatory response to variant morphologies of UHMWPE particles. *Biomaterials* **2003**, *24* (26), 4819–4826.
52. Sud, S.; Yang, S.Y.; Evans, C.H.; Robbins, P.D.; Wooley, P.H. Effects of cytokine gene therapy on particulate-induced inflammation in the murine air pouch. *Inflammation* **2001**, *25* (6), 361–372.
53. Epstein, N.J.; Warm, B.A.; Spanogle, J.; Ma, T.; Bragg, B.; Smith, R.L.; Goodman, S.B. Interleukin-1 modulates periprosthetic tissue formation in an intramedullary model of particle-induced inflammation. *J. Orthop. Res.* **2005**, *23* (3), 501–510.
54. Yang, S.Y.; Nasser, S.; Markel, D.C.; Robbins, P.D.; Wooley, P.H. Human periprosthetic tissue implanted in severe combined immunodeficient mice respond to gene transfer of a cytokine inhibitor. *J. Bone Joint Surg. Am.* **2005**, *87* (5), 1088–1097.

Wear Debris from Joint Arthroplasties, Biological Effects Of

Ting Ma

Stuart B. Goodman

Department of Orthopedic Surgery, Stanford Medical Center, Stanford, California, U.S.A.

INTRODUCTION

Total joint replacement (TJR) is currently one of the most successful operations performed in orthopedic surgery. For the hip and knee, these operations consistently yield excellent clinical results, with 15–20 yr survivorship of 90% or more, using contemporary prostheses and techniques.^[1] These results have prompted the use of joint replacement in younger patients, who frequently lead more active lifestyles than the sedentary, elderly population for which TJR was originally designed. Furthermore, the general population is aging, and the elderly are becoming more physically active. These facts have led to increased expectations by patients undergoing TJR, and increased demands on the prosthesis construct. Although some technical advances have been associated with an increasing success rate of specific prostheses, the number of cases requiring surgical revision is still increasing. According to a recent report, a 10-year survival rate could be observed in 90% of patients older than 70, and only 72% of patients younger than 50.^[2]

Wear of TJRs and the biological effects of wear debris currently remain the most important problems in arthroplasty surgery. The wear between the primary bearing surfaces is generally a silent process with no clinical symptoms or signs in the early stages. However, when debris-associated chronic synovitis and osteolysis become more advanced, revision surgery is indicated, including thorough debridement and replacement of worn and damaged components.

In this article, the characteristics of the tissue at the bone–implant interface are reviewed, and the biological effects of wear debris from TJRs are summarized.

THE TISSUE AT THE BONE–IMPLANT INTERFACE IN HUMANS

Characteristics of Wear Debris Retrieved from Surgery or Autopsy

Any prosthesis used for total hip replacement (THR) can generate biologically responsive particles.^[3] The observation of tissue surrounding loose prostheses

provided us the firsthand information about implant wear and has substantially enhanced our understanding of the nature of the wear debris generated from the prosthesis.

Histological observations of the retrieved periprosthetic membrane first suggested that the generation of wear particles and the process of loosening were related.^[4,5] Numerous particles of bone cement, polyethylene, and metallic debris were observed in the retrieved periprosthetic tissue. The nature of the particles is different depending on the materials used for the prosthesis.^[5–7] It is believed that cementless prostheses are associated with more particles than cemented ones.^[3] McKellop et al.^[8] predicted that using a metal-on-polyethylene joint prosthesis, several hundreds of thousands of polyethylene particles can be expected during a single gait cycle. When metal-on-metal wear is involved, black metal particles can be easily found in the retrieved membrane. In cases in which massive amounts of metallic particles have been generated, the tissues may be grey or black on gross examination. Perhaps these particles are the easiest, but often the smallest, ones to identify. The size of the particles ranges from several microns in diameter to the submicron range. Many of them are located in the interstitium and within macrophages.^[3,9–14] It is believed that these metal particles are usually associated with less overall inflammatory reaction.^[15]

Particles of polymethylmethacrylate (PMMA) are often found to be surrounded by a florid foreign-body response in a fibrous stroma. Smaller particles of PMMA, and particles of the radiopaque marker in bone cement (barium sulfate or zirconium), less than approximately 10 μm are phagocytosed by macrophages; larger particles up to approximately 30 μm are found within multinucleated foreign-body giant cells. Still larger agglomerates of large and small particles are surrounded by a foreign-body granulomatous reaction.^[16,17]

Polyethylene particles are perhaps the most difficult to identify histologically in retrieved tissues.^[18] However, it is now widely accepted that ultrahigh-molecular-weight polyethylene (UHMWPE) is probably one of the dominant wear debris contributing to particle-associated osteolysis.^[19,20] Irregularly shaped polyethylene particles with the size ranging from 0.2 to 7 μm are especially bioactive, and provoke greater

inflammatory responses than other particles.^[21] Using laser diffraction wear particle analysis, Elfick et al.^[18] reported that the majority of the UHMWPE particles in periprosthetic tissue of THR are submicron in size, with the median particle diameter being 0.69 μm . No UHMWPE smaller than 0.113 μm were found. The number of particles associated with the chronic foreign-body reaction in tissue was higher than 10^9 particles per gram tissue, and can interact with cells that are then induced to release bone-resorbing mediators.^[20,22,23]

Till date, particles generated from all biomaterials clinically in use for THR are bioactive, that is, able to promote inflammatory responses, formation of foreign-body granulomas, and eventually the destruction of bone structure.^[3] Osteolysis refers to the destruction of bone seen radiographically as localized or more widespread, large radiolucent areas. Particles generated from the wearing of prosthesis have been shown to elicit a florid foreign-body reaction, eventually leading to bone destruction.^[16,24–28] The prostheses do not have to be mechanically loose to demonstrate this reaction. Cemented femoral prostheses made of titanium alloy, especially those with a titanium alloy femoral head, may be associated with higher than expected wear rates, loosening, and osteolysis.^[29–32] Ceramic-on-ceramic prosthesis may yield a relatively low number of particles, although foreign-body response and periprosthetic osteolysis are present with ceramic-bearing surfaces.^[33]

Biochemistry

Current research efforts have identified factors in the retrieved membrane that regulate inflammatory processes, bone formation, and osteolysis following THR.

The presence of proinflammatory cytokines including interleukin-1 (IL-1), interleukin-6 (IL-6), and tumor necrosis factor alpha (TNF-alpha) have been well documented.^[34,35] Interleukin-1 and TNF-alpha are potent osteoclastogenic cytokines produced by activated monocytes/macrophages. Their important role in particle-induced osteolysis has therefore received much attention.^[6] Other cytokines such as IL-8^[36] and IL-11^[37] are also identified in the membrane. Their potential role in the periprosthetic osteolysis is yet to be determined.

In addition to the cytokines, prostaglandin E₂ (PGE₂),^[38] metalloproteinases, collagenase,^[39] superoxide anions such as peroxynitrite,^[40] receptor activator of nuclear factor kappa B (NF-kappaB) ligand (RANKL),^[41] macrophage colony-stimulating factor (M-CSF),^[42] and vascular endothelial growth factor (VEGF)^[43] are released by cells associated with wear particles.^[44]

PGE₂ plays an important role in regulating osteoclastic bone resorption.^[45,46] Enzymes such as collagenase promote bone resorption by digesting the extracellular matrix to expose bone mineral for osteoclastic activities.^[47] RANKL, together with M-CSF, function as potent factors for osteoclastogenesis. Their presence in fibroblast and macrophage lineage cells in the retrieved membrane suggests the in situ involvement of RANKL in osteoclastic bone resorption occurring during the process of prosthetic osteolysis.^[48] The VEGF was detected in the synovial fluid and in the periprosthetic membrane.^[43] In association with other preangiogenic mediators present in the membrane such as TNF-alpha,^[37,49,50] granulocyte-macrophage colony-stimulating factor,^[51,52] basic fibroblast growth factor,^[53] and transforming growth factor-beta 1 and 2,^[54] endothelial leukocyte adhesion molecule-1, and vascular cell adhesion molecule-1,^[55] the process of angiogenesis has been implicated in osteolysis.

Biochemical studies of harvested tissue have added another dimension to our understanding of the events at the bone-implant interface of orthopedic prostheses. Further studies are necessary to document and fully understand the plethora of substances produced at the prosthetic interface and their role in loosening and resorption of bone.

Cellular Profiles

The retrieved periprosthetic membrane surrounding the implant contains various types of cells including macrophages, fibroblasts, and lymphocytes.^[6,38,56] In cases of loosening without osteolysis, fibroblast-like mesenchymal cells are predominant; in cases of aggressive granulomatosis, mono- and multinucleated histiocytes are common, and activated fibroblasts are more scarce.^[57,58] As to the involvement of lymphocytes in aseptic loosening of THR, research results have been contradictory. Farber et al.^[59] detected large numbers of antigen-presenting cells expressing B7-1 or B7-2, two well-characterized ligands of the costimulatory signal, and T cells using monoclonal antibodies. Although there were no statistically significant differences in expression of proteins in the B7-CD28 costimulatory pathway between the osteolytic and nonosteolytic groups, they postulated the involvement of activated immune response in process of aseptic loosening owing to the magnitude of positive staining. In a study by Li et al.^[60] the number of T cells in the membrane-like tissue was low, and none of the T cells was activated as shown by the absence of interleukin-2 receptor immunoreactivity. Sandhu et al.^[61] reported that T cells constitute approximately 10% of the total cell population and likely play an important role in modulating the foreign body and chronic

inflammatory reaction in response to wear particles by a nonspecific immune mechanism.

BIOLOGICAL EFFECTS OF PARTICLES IN VIVO AND IN VITRO

Particulate wear debris can alter the function of a variety of cell types including macrophages, fibroblasts, osteoblasts, and osteoclasts by means of phagocytosis or surface activation.^[3,21,62-64] At present, monocyte/macrophages are thought to be the most predominant cells associated with the inflammatory and foreign-body cascade.^[44]

Macrophages

Wear particles less than approximately 10 μm can be engulfed by macrophages,^[65] thus activating a cascade of events resulting in upregulation of proinflammatory mediators. These mediators in turn perpetuate the inflammation by means of autocrine and paracrine mechanisms, resulting in the differentiation and activation of more macrophages.^[43,44] To date, almost all the particles generated from wearing of prostheses, including metal particles, UHMWPE, PMMA, ceramics, and hydroxyapatite, have been shown to activate macrophages.^[28,66-70] Following the initial uptake of particles, the materials may be subsequently exposed to macrophage-derived enzymes with the potential for material degradation. To examine the impact of polymeric surface characteristics on the production of macrophage biodegradative enzymes, Santerre and coworkers^[71] cultured human peripheral blood monocytes on biomaterials with different surface characters. They found that the material surface itself is a potent activator of macrophages, and the activities of degradative enzymes they assayed showed different patterns depending on the materials added.

Activation of macrophages following particle exposure may not require phagocytosis.^[62] Particles can be associated with specific serum proteins to form a particle-protein complex. This complex interacts with cell surface receptors on macrophages and activates these cells.^[6,72] The surface characters of particles can also have a marked effect on the material's ability to elicit cellular responses in this respect.^[73,74]

The potential role of adherent endotoxin in the process of macrophage activation and cytokine production remains unclear. Greenfield and coworkers^[75] demonstrated that removal of adherent endotoxin significantly inhibited the response of murine marrow cells and human peripheral blood monocytes to titanium particles. Addition of lipopolysaccharide (LPS)

to the endotoxin-free particles restored the inflammatory cytokine release of macrophages. However, this result was challenged by the study of Clohisy et al.^[76] that showed the PMMA-induced activation of NF-kappaB, a crucial transcription factor mediating TNF signaling necessary for osteoclastogenesis, is independent of LPS because the level of NF-kappaB activation is no different between wild type mice and LPS-hyporesponsive mice.

Because of the predominant presence of macrophages at the bone-implant interface, some believe that macrophages themselves are capable of direct bone resorption.^[77,78] The enzymes capable of direct bone degradation were detected within macrophages by immunohistochemistry.^[79] The report by Quinn et al.^[80] that macrophages containing PMMA particles could directly erode bone supports the direct role of macrophages in osteolysis.

Osteoclasts

The differentiation, migration, and activation of osteoclasts has been implicated as the key step accelerating the degradation of bone.^[6] Most inflammatory mediators present at the bone-implant interface, including PGE₂, TNF-alpha, IL-1beta, and M-CSF are potent stimulators of osteoclastogenesis.^[81-83] Recent studies suggested that the essential signaling mechanism involved in osteoclastogenesis is the interaction of RANKL and RANK. Osteoprotegerin (OPG), a decoy receptor of RANKL, blocks the RANKL-RANK coupling, thus, inhibiting the osteoclastic activity. Baumann et al.^[84] studied the effect of polyethylene and TiAlV wear particles on expression of RANK, RANKL, and OPG using a human monocytic leukemic cell line. They showed that TiAlV and polyethylene particles induced upregulation of RANK mRNA expression; OPG mRNA expression was transiently upregulated, and RANKL mRNA was not detectable.^[84] This study may not completely represent the *in vivo* situation of human cells given the fact that cell lines behave differently from primary cells in many aspects. In fact, Kim et al.^[85] demonstrated that the OPG levels in the joint fluid from failed total hip arthroplasty patients were much lower than that collected from patients with osteoarthritis. RANK-RANKL coupling is required for other well-observed events of osteoclastic bone resorption. For example, it has been proved that both TNF-alpha and IL-1 signaling on osteoclastogenesis depend on RANK-RANKL interactions. In the presence of permissive levels of RANKL, TNF-alpha potentiates osteoclastogenesis.^[81] IL-1 prolongs the survival and actin ring formation (a functional marker of osteoclasts) of

osteoclasts in the presence of RANKL, and appears to act independent of TNF-alpha.^[82,86]

Osteoblasts

Osteoblasts are mainly responsible for the process of bone formation. Therefore, inhibition of osteoblast activity may be another mechanism accelerating periprosthetic bone loss. We have observed reduced new bone formation in the presence of particles using an in vivo rabbit model.^[17,87] Furthermore, osteoblasts are able to phagocytose particles in culture, which implicates the active involvement of these bone-forming cells in periprosthetic osteolysis.^[64] Vermes et al.^[63,64] treated MG-63 osteoblast-like cells with metal particles or polymeric particles. They found that a maximum of 40–60 particles can be phagocytosed by one cell, and those particles suppress the synthesis of type I collagen and increase the secretion of IL-6 and PGE₂ by activating osteoblasts.^[88] It is therefore likely that, in human patients, while the ability for new bone formation had been impaired by particulate wear debris, the process of bone resorption continues to be potentiated by factors released from osteoblasts. In addition, it has been reported that metallic particles alter osteoblast adhesion behavior, proliferation, viability, and gene expression.^[49,89–91]

Mesenchymal Stem Cells

Submicron-sized metallic wear particles have been identified in bone marrow surrounding the femoral stem and in more remote sites such as the iliac crest when implant fixation is relatively poor.^[92] Thus the behavior of bone marrow cells is continuously challenged by wear debris. Wang et al.^[93] reported that the exposure to wear particles could compromise the viability of human mesenchymal stem cells (MSCs) by direct and indirect induction of apoptosis. They also found that the osteogenic differentiation of human marrow stroma-derived MSCs was suppressed by titanium particles. These findings suggest that prolonged exposure of bone marrow cells to implant-derived particles may significantly decrease the number of viable stem cells. With the diminished number of the stem cell population and the compromised stem cell function, the process of osseointegration and prosthesis fixation can be severely jeopardized.^[94]

CONCLUSIONS

The biological response to wear debris in patients with a joint arthroplasty is a complicated process. In vitro and in vivo studies using multiple investigative

techniques, including histology, cell biology, and molecular biology have provided insightful information to our current understanding of these biological processes. Wear debris generated from implants, in particulate form, stimulates a foreign body and chronic inflammatory response that is capable of inducing the resorption of bone by osteoclasts and macrophages. The particles disrupt the formation of bone by negatively impacting the normal behavior of osteoblasts and MSC differentiation. The degree to which this reaction occurs is dependent on the location of the implantation, the specific experimental model, the characteristics of the particles, and the particle load. Because of possible adverse reactions to particles, it behooves surgeons, material scientists, and biologists to develop new strategies to prevent the generation of particulate debris from orthopedic implants and modulate the process of osteolysis at an early stage.



ARTICLES OF FURTHER INTEREST

Biomaterials Immune Response; Wear Debris; Wear Debris, Bone Resorption Animal Models

REFERENCES

1. Wright, T., Goodman, S.B., Eds.; Implant wear in total joint replacement; American Academy of Orthopaedic Surgeons: Rosemont, IL, 2001.
2. Puolakkam, T.J.S.; Pajamäki, K.J.J.; Halonen, P.J.; Pulkkinen, P.O.; Paavo Iainen, P.; Nevalainen, J.K. The finnish arthroplasty register. Report of the hip register. *Acta Orthop. Scand.* **2001**, *72*, 433–441.
3. Gallo, J.; Kaminek, P.; Ticha, V.; Rihakova, P.; Ditmar, R. Particle disease. A comprehensive theory of periprosthetic osteolysis: a review. *Biomed. Papers* **2002**, *146* (2), 21–28.
4. Goodman, S.B.; Lind, M.; Song, Y.; Smith, R.L. In vitro, in vivo, and tissue retrieval studies on particulate debris. *Clin. Orthop. Relat. Res.* **1998**, *352*, 25–34.
5. Santavirta, S.; Takagi, M.; Gomez-Barrena, E.; Nevalainen, J.; Lassus, J.; Salo, J.; Kontinen, Y.T. Studies of host response to orthopedic implants and biomaterials. *J. Long Term Effects Med. Implants.* **1999**, *9* (1–2), 67–76.
6. Goodman, S.B.; Huie, P.; Song, Y.; Schurman, D.; Maloney, W.; Woolson, S.; Sibley, R. Cellular profile and cytokine production at prosthetic interfaces. Study of tissues retrieved from revised hip and knee replacements. *J. Bone Joint Surg. Br.* **1998**, *80* (3), 531–539.
7. Goodman, S.B.; Trindade, M.; Ma, T.; Genovese, M.; Smith, R.L. Pharmacologic modulation of periprosthetic osteolysis. *Clin. Orthop. Relat. Res.* **2005**, *430*, 39–45.
8. McKellop, H.A.; Campbell, P.; Park, S.-H.; Schmalzried, T.P.; Grigoris, P.; Amstutz, H.C.; Sarmiento, A. The origin of submicron polyethylene wear debris in total hip arthroplasty. *Clin. Orthop.* **1995**, *311*, 3–20.

9. Rae, T. A study on the effects of particulate metals of orthopaedic interest on murine macrophages in vitro. *J. Bone Joint Surg.* **1975**, *57B*, 444–450.
10. Rae, T. The toxicity of metals used in orthopaedic prostheses. An experimental study using cultured human synovial fibroblasts. *J. Bone Joint Surg.* **1981**, *63B*, 435–440.
11. Rae, T. The biological response to titanium and titanium-aluminum-vanadium particles. II. Long term animal studies. *Biomaterials* **1986**, *7*, 37–40.
12. Buchert, P.K.; Bradley, K.V.; Mallory, T.M.; Engh, C.A.; Bobyn, J.D. Excessive metal release due to loosening and fretting of sintered particles on porous-coated hip prostheses. *J. Bone Joint Surg.* **1986**, *68A*, 606–609.
13. Agins, H.S.; Alcock, N.W.; Bansal, M.; Salvati, E.A.; Wilson, P.D.; Pellici, P.M.; Bullough, P.G. Metallic wear in failed titanium-alloy total hip replacements. *J. Bone Joint Surg.* **1988**, *70A*, 347–356.
14. Howie, D.W.; Vernon-Roberts, B.; Oakeshott, R.; Manthey, B. A rat model of resorption of bone at the cemented bone interface in the presence of polyethylene wear particles. *J. Bone Joint Surg.* **1988**, *70A*, 257–263.
15. Doorn, P.F.; Mirra, J.M.; Campbell, P.A.; Amstutz, H.C. Tissue reaction to metal on metal total hip prostheses. *Clin. Orthop. Relat. Res.* **1996**, *329* (suppl), S187–S205.
16. Jones, L.C.; Hungerford, D.S. Cement disease. *Clin. Orthop.* **1987**, *225*, 193–206.
17. Goodman, S.B.; Fornasier, V.L.; Kei, J. The effects of bulk versus particulate ultra-high-molecular-weight polyethylene on bone. *J. Arthroplasty* **1988**, *3* (suppl), 41–46.
18. Elfick, A.P.; Green, S.M.; Krikler, S.; Unsworth, A. The nature and dissemination of UHMWPE wear debris retrieved from periprosthetic tissue of THR. *J. Biomed. Mater. Res. A.* **2003**, *65* (1), 95–108.
19. Friedman, I.R.; Black, J.; Galante, J.O.; Jacobs, J.J.; Skinner, H.B. Current concepts in orthopaedic biomaterials and implant fixation. *J. Bone Joint Surg.* **1993**, *75A*, 1086–1109.
20. Learmonth, I.D. Biocompatibility: a biomechanical and biological concept in total hip replacement. *Surgeon* **2003**, *1* (1), 1–8.
21. Ingham, E.; Fisher, J. Biological reactions to wear debris in total joint replacement. *Proc. Instn. Mech. Engrs. Part H J. Eng. Med.* **2001**, *214*, 21–37.
22. Margevicius, K.J.; Bauer, T.W.; McMahon, I.T.; Brown, S.A.; Merritt, K. Isolation and characterisation of debris from around total joint prostheses. *J. Bone Joint Surg.* **1994**, *76A*, 1664–1675.
23. Shanbhag, A.S.; Jacobs, J.J.; Glant, T.T.; Gilbert, J.L.; Black, J.; Galante, J.O. Composition and morphology of wear debris in failed uncemented total hip replacement. *J. Bone Joint Surg.* **1994**, *76B*, 60–67.
24. Huddleston, H.D. Femoral lysis after cemented hip arthroplasty. *J. Arthroplasty* **1988**, *3* (4), 285–297.
25. Anthony, P.P.; Gie, G.A.; Howie, C.R.; Ling, R.S.M. Localized endosteal bone lysis in relation to the femoral components of cemented total hip arthroplasties. *J. Bone Joint Surg.* **1990**, *72B*, 971–979.
26. Maloney, W.J.; Jasty, M.; Rosenberg, A.; Harris, W.H. Bone lysis in well-fixed cemented femoral components. *J. Bone Joint Surg.* **1990**, *72B*, 966–970.
27. Maloney, W.J.; Peters, P.; Engh, C.A.; Chandler, H. Severe osteolysis of the pelvis in association with acetabular replacement without cement. *J. Bone Joint Surg.* **1993**, *75A*, 1627–1635.
28. Willert, H.G.; Bertram, H.; Buchhorn, G.H. Osteolysis in alloarthroplasty of the hip. The role of ultra-high molecular weight polyethylene wear particles. *Clin. Orthop.* **1990**, *258*, 95–107.
29. Lombardi, A.V.; Mallory, T.H.; Vaughn, B.K.; Drouillard, P. Aseptic loosening in total hip arthroplasty secondary to osteolysis induced by wear debris from titanium-alloy modular femoral heads. *J. Bone Joint Surg.* **1989**, *71A*, 1337–1342.
30. Witt, J.D.; Swann, M. Metal wear and tissue response in failed titanium alloy total hip replacements. *J. Bone Joint Surg.* **1991**, *73B*, 559–563.
31. Buly, R.L.; Huo, M.H.; Salvati, E.; Brien, W.; Bansal, M. Titanium wear debris in failed cemented total hip arthroplasty. An analysis of 71 cases. *J. Arthroplasty* **1992**, *7* (3), 315–323.
32. Salvati, E.A.; Betts, F.; Doty, S.B. Particulate metallic debris in cemented total hip arthroplasty. *Clin. Orthop.* **1993**, *293*, 160–173.
33. Yoon, T.R.; Rowe, S.M.; Jung, S.T.; Seon, K.J.; Maloney, W.J. Osteolysis in association with a total hip arthroplasty with ceramic bearing surfaces. *J. Bone Joint Surg. Am.* **1998**, *80* (10), 1459–1468.
34. Kim, K.J.; Chiba, J.; Rubash, H.E. In vivo and in vitro analysis of membranes from hip prostheses inserted without cement. *J. Bone Joint Surg. Am.* **1994**, *76*, 172–180.
35. Jiranek, W.A.; Machado, M.; Jasty, M. Production of cytokines around loosened cemented acetabular components: analysis with immunohistochemical techniques and in situ hybridization. *J. Bone Joint Surg. Am.* **1993**, *75*, 863–879.
36. Lassus, J.; Waris, V.; Xu, J.W.; Li, T.F.; Hao, J.; Nietosvaara, Y.; Santavirta, S.; Konttinen, Y.T. Increased interleukin-8 (IL-8) expression is related to aseptic loosening of total hip replacement. *Arch. Orthop. Trauma Surg.* **2000**, *120* (5–6), 328–332.
37. Xu, J.W.; Li, T.F.; Partsch, G.; Ceponis, A.; Santavirta, S.; Konttinen, Y.T. Interleukin-11 (IL-11) in aseptic loosening of total hip replacement (THR). *Scand. J. Rheumatol.* **1998**, *27* (5), 363–367.
38. Goldring, S.R.; Jasty, M.; Roelke, M.S.; Rourke, C.M.; Bringhurst, F.R.; Harriw, W.H. Formation of a synovial-like membrane at the bone-cement interface. Its role in bone resorption and implant loosening after total hip replacement. *Arthritis Rheum.* **1986**, *29* (7), 836–842.
39. Nakashima, Y.; Sun, D.H.; Maloney, W.J.; Goodman, S.B.; Schurman, D.J.; Smith, R.L. Induction of matrix metalloproteinase expression in human macrophages by orthopaedic particulate debris in vitro. *J. Bone Joint Surg. Br.* **1998**, *80*, 694–700.
40. Puskas, B.L.; Menke, N.E.; Huie, P.; Song, Y.; Ecklund, K.; Trindade, M.C.; Smith, R.L.; Goodman, S.B. Expression of nitric oxide, peroxy nitrite, and apoptosis in loose total hip replacements. *J. Biomed. Mater. Res. A.* **2003**, *66* (3), 541–549.
41. Mandelin, J.; Li, T.F.; Liljestrom, M.; Kroon, M.E.; Hanemaaijer, R.; Santavirta, S.; Konttinen, Y.T.

- Imbalance of RANKL/RANK/OPG system in interface tissue in loosening of total hip replacement. *J. Bone Joint Surg. Br.* **2003**, *85* (8), 1196–1201.
42. Xu, J.W.; Konttinen, Y.T.; Waris, V.; Patiala, H.; Sorsa, T.; Santavirta, S. Macrophage-colony stimulating factor (M-CSF) is increased in the synovial-like membrane of the periprosthetic tissues in the aseptic loosening of total hip replacement (THR). *Clin. Rheumatol.* **1997**, *16* (3), 243–248.
 43. Miyanishi, K.; Trindade, M.C.; Ma, T.; Goodman, S.B.; Schurman, D.J.; Smith, R.L. Periprosthetic osteolysis: induction of vascular endothelial growth factor from human monocyte/macrophages by orthopaedic biomaterial particles. *J. Bone Miner Res.* **2003**, *18* (9), 1573–1583.
 44. Santerre, J.P.; Labow, R.S.; Boynton, E.L. The role of the macrophage in periprosthetic bone loss. *Can. J. Surg.* **2000**, *43* (3), 173–179.
 45. Kobayashi, Y.; Mizoguchi, T.; Take, I.; Kurihara, S.; Udagawa, N.; Takahashi, N. Prostaglandin E2 enhances osteoclastic differentiation of precursor cells through protein kinase A-dependent phosphorylation of TAK1. *J. Biol. Chem.* **2005**, *280* (12), 11395–11403.
 46. Suzawa, T.; Miyaura, C.; Inada, M.; Maruyama, T.; Sugimoto, Y.; Ushikubi, F.; Ichikawa, A.; Narumiya, S.; Suda, T. The role of prostaglandin E receptor subtypes (EP1, EP2, EP3, and EP4) in bone resorption: an analysis using specific agonists for the respective EPs. *Endocrinology* **2000**, *141* (4), 1554–1559.
 47. Heersche, J.N.; Aubin, J.E. Regulation of cellular activity of bone forming cells. In *Bone*; Hall, B.K., Ed.; Telford Press: Caldwell, NJ, 1990; 327–349.
 48. Horiki, M.; Nakase, T.; Myoui, A.; Sugano, N.; Nishii, T.; Tomita, T.; Miyaji, T.; Yoshikawa, H. Localization of RANKL in osteolytic tissue around a loosened joint prosthesis. *J. Bone Miner Metab.* **2004**, *22* (4), 346–351.
 49. Takei, H.; Pioletti, D.P.; Kwon, S.Y.; Sung, K.L. Combined effect of titanium particles and TNF-alpha on the production of IL-6 by osteoblast-like cells. *J. Biomed. Mater. Res.* **2000**, *52*, 382–387.
 50. Xu, J.W.; Konttinen, Y.T.; Lassus, J.; Natah, S.; Ceponis, A.; Solovieva, S.; Aspenberg, P.; Santavirta, S. Tumor necrosis factor-alpha (TNF-alpha) in loosening of total hip replacement (THR). *Clin. Exp. Rheumatol.* **1996**, *14*, 643–648.
 51. Neale, S.D.; Sabokbar, A.; Howie, D.W.; Murray, D.W.; Athanasou, N.A. Macrophage colony-stimulating factor and interleukin-6 release by periprosthetic cells stimulates osteoclast formation and bone resorption. *J. Orthop. Res.* **1999**, *17* (5), 686–694.
 52. Al-Saffar, N.; Khwaja, H.A.; Kadoya, Y.; Revell, P.A. Assessment of the role of GM-CSF in the cellular transformation and the development of erosive lesions around orthopaedic implants. *Am. J. Clin. Pathol.* **1996**, *105*, 628–639.
 53. Waris, V.; Xu, J.W.; Nordsletten, L.; Sorsa, T.; Santavirta, S.; Konttinen, Y.T. Basic fibroblast growth factor (bFGF) in the synovial-like membrane around loose total hip prostheses. *Scand. J. Rheumatol.* **1996**, *25*, 257–262.
 54. Konttinen, Y.T.; Waris, V.; Xu, J.W.; Jiranek, W.A.; Sorsa, T.; Virtanen, I.; Santavirta, S. Transforming growth factor-beta 1 and 2 in the synovial-like interface membrane between implant and bone in loosening of total hip arthroplasty. *J. Rheumatol.* **1997**, *24*, 694–701.
 55. Al-Saffar, N.; Mah, J.T.L.; Kadoya, Y.; Revell, P.A. Neovascularisation and the induction of cell adhesion molecules in response to degradation products from orthopaedic implants. *Ann. Rheum. Dis.* **1995**, *54*, 201–208.
 56. Perry, M.J.; Mortuza, F.Y.; Ponsford, F.M.; Elson, C.J.; Atkins, R.M. Analysis of cell types and mediator production from tissues around loosening joint implants. *Br. J. Rheumatol.* **1995**, *34* (12), 1127–1134.
 57. Tallroth, K.; Eskola, A.; Santavirta, S.; Konttinen, Y.T.; Lindholm, T.S. Aggressive granulomatous lesions after hip arthroplasty. *J. Bone Joint Surg.* **1989**, *71B*, 571–575.
 58. Santavirta, S.; Konttinen, Y.T.; Hoikka, V.; Eskola, A. Immunopathological response to loose cementless acetabular components. *J. Bone Joint Surg.* **1991**, *73B*, 38–42.
 59. Farber, A.; Chin, R.; Song, Y.; Huie, P.; Goodman, S. Chronic antigen-specific immune-system activation may potentially be involved in the loosening of cemented acetabular components. *J. Biomed. Mater. Res.* **2001**, *55* (3), 433–441.
 60. Li, T.F.; Santavirta, S.; Waris, V.; Lassus, J.; Lindroos, L.; Xu, J.W.; Virtanen, I.; Konttinen, Y.T. No lymphokines in T-cells around loosened hip prostheses. *Acta Orthop. Scand.* **2001**, *72* (3), 241–247.
 61. Sandhu, J.; Waddell, J.; Henry, M.; Boynton, E.L. The role of T-cells in polyethylene induced inflammation. *J. Rheumatol.* **1998**, *25*, 1794–1799.
 62. Nakashima, Y.; Sun, D.H.; Trindade, M.C.D.; Maloney, W.J.; Goodman, S.B.; Schurman, D.J.; Smith, R.L. Titanium alloy particle activation of signaling pathways for tumor necrosis factor-alpha and interleukin-6 expression in human macrophages in vitro. *J. Bone Joint Surg.* **1999**, *81A*, 603–615.
 63. Vermes, C.; Chandrasekaran, R.; Jacobs, J.J.; Galante, J.O.; Roebuck, K.A.; Glant, T.T. The effects of particulate wear debris, cytokines, and growth factors on the functions of MG-63 osteoblasts. *J. Bone Joint Surg.* **2001**, *83A*, 201–211.
 64. Vermes, C.; Glant, T.T.; Hallab, N.J.; Fritz, E.A.; Roebuck, K.A.; Jacobs, J.J. The potential role of the osteoblast in the development of periprosthetic osteolysis. Review of in vitro osteoblast responses to wear debris, corrosion products, and cytokines and growth factors. *J. Arthroplasty* **2001**, *16* (suppl. 1), 95–100.
 65. Trindade, M.C.; Lind, M.; Sun, D.; Schurman, D.J.; Goodman, S.B.; Smith, R.L. In vitro reaction to orthopaedic biomaterials by macrophages and lymphocytes isolated from patients undergoing revision surgery. *Biomaterials* **2001**, *22* (3), 253–259.
 66. Goldring, S.R.; Schiller, A.L.; Roilke, M.; Rourke, C.M.; O'Neil, D.A.; Harris, W.H. The synovial-like membrane at the bone-cement interface in loose total hip replacement and its proposed role in bone lysis. *J. Bone Joint Surg. Am.* **1983**, *65* (5), 575–584.
 67. Haynes, D.R.; Rogers, S.D.; Hay, S.; Percy, M.J.; Howie, D.W. The differences in toxicity and release of bone-resorbing mediators induced by titanium and



- cobalt-chromium-alloy wear particles. *J. Bone Joint Surg Am.* **1993**, *75* (6), 825–834.
68. Horowitz, S.M.; Doty, S.B.; Lane, J.M.; Burstein, A.H. Studies of the mechanism by which the mechanical failure of polymethylmethacrylate leads to bone resorption. *J. Bone Joint Surg. Am.* **1993**, *75* (6), 802–813.
 69. Harada, Y.; Wang, J.; Doppalapudi, V.A.; Willis, A.A.; Jasty, M.; Harris, W.H. Differential effects of different forms of hydroxyapatite and hydroxyapatite/tricalcium phosphate particulates on human monocyte/macrophages in vitro. *J. Biomed. Mater. Res.* **1996**, *31*, 19–26.
 70. Wang, X.Y.; Li, S.; Salovey, R. Processing of UHMWPE. *J. Appl. Polym. Sci.* **1988**, *35*, 2165–2171.
 71. Labow, R.S.; Meek, E.; Santerre, J.P. Differential synthesis of cholesterol esterase by monocyte-derived macrophages cultured on poly(ether or ester)-based poly(urethane)s. *J. Biomed. Mater. Res.* **1998**, *39* (3), 469–477.
 72. Sun, D.H.; Trindade, M.C.; Nakashima, Y.; Maloney, W.J.; Goodman, S.B.; Schurman, D.J.; Smith, R.L. Human serum opsonization of orthopedic biomaterial particles: protein-binding and monocyte/macrophage activation in vitro. *J. Biomed. Mater. Res. A* **2003**, *65* (2), 290–298.
 73. Czajkowska, B.; Blazewica, M. Phagocytosis of chemically modified carbon materials. *Biomaterials* **1997**, *18*, 69–74.
 74. Hunt, J.A.; Flanagan, B.F.; McLaughlin, P.J.; Strickland, I.; Williams, D.F. Effect of biomaterial surface charge on the inflammatory response: evaluation of cellular infiltration and TNF alpha production. *J. Biomed. Mater. Res.* **1996**, *31*, 139–144.
 75. Bi, Y.; Seabold, J.M.; Kaar, S.G.; Ragab, A.A.; Goldberg, V.M.; Anderson, J.M.; Greenfield, E.M. Adherent endotoxin on orthopedic wear particles stimulates cytokine production and osteoclast differentiation. *J. Bone Miner Res.* **2001**, *16* (11), 2082–2091.
 76. Clohisy, J.C.; Teitelbaum, S.; Chen, S.; Erdmann, J.M.; Abu-Amer, Y. Tumor necrosis factor-alpha mediates polymethylmethacrylate particle-induced NF-kappaB activation in osteoclast precursor cells. *J. Orthop. Res.* **2002**, *20* (2), 174–181.
 77. Kadoya, Y.; Revell, P.A.; al-Saffar, N.; Kobayashi, A.; Scott, G.; Freeman, M.A. Bone formation and resorption in failed total joint arthroplasties: histomorphometric analysis with histochemical and immunohistochemical technique. *J. Orthop. Res.* **1996**, *14*, 473–482.
 78. Kadoya, Y.; Kobayashi, A.; Ohashi, H. Wear and osteolysis in total joint replacements. *Acta Orthop. Scand.* **1998**, *69* (suppl. 278), 1–14.
 79. Kadoya, Y.; al-Saffar, N.; Kobayashi, A.; Revell, P.A. The expression of osteoclast markers on foreign body giant cells. *Bone Miner* **1994**, *27* (2), 85–96.
 80. Quinn, J.; Joyner, C.; Triffitt, I.T.; Athanasou, N.A. Polymethylmethacrylate-induced inflammatory macrophages resorb bone. *J. Bone Joint Surg. Br.* **1992**, *74*, 652–658.
 81. Lam, J.; Takeshita, S.; Barker, J.E.; Kanagawa, O.; Ross, F.P.; Teitelbaum, S.L. TNF-alpha induces osteoclastogenesis by direct stimulation of macrophages exposed to permissive levels of RANK ligand. *J. Clin. Invest.* **2000**, *106*, 1481–1488.
 82. Ma, T.; Miyanishi, K.; Suen, A.; Epstein, N.J.; Tomita, T.; Smith, R.L.; Goodman, S.B. Human interleukin-1-induced murine osteoclastogenesis is dependent on RANKL, but independent of TNF-alpha. *Cytokine* **2004**, *26* (3), 138–144.
 83. Suda, K.; Udagawa, N.; Sato, N.; Takami, M.; Itoh, K.; Woo, J.T.; Takahashi, N.; Nagai, K. Suppression of osteoprotegerin expression by prostaglandin E2 is crucially involved in lipopolysaccharide-induced osteoclast formation. *J. Immunol.* **2004**, *172* (4), 2504–2510.
 84. Baumann, B.; Rader, C.P.; Seufert, J.; Noth, U.; Rolf, O.; Eulert, J.; Jakob, F. Effects of polyethylene and TiAlV wear particles on expression of RANK, RANKL and OPG mRNA. *Acta Orthop. Scand.* **2004**, *75* (3), 295–302.
 85. Kim, K.J.; Kotake, S.; Udagawa, N.; Ida, H.; Ishii, M.; Takei, I.; Kubo, T.; Takagi, M. Osteoprotegerin inhibits in vitro mouse osteoclast formation induced by joint fluid from failed total hip arthroplasty. *J. Biomed. Mater. Res.* **2001**, *58* (4), 393–400.
 86. Jimi, E.; Nakamura, K.; Duong, L.T.; Ikebe, T.; Takahashi, N.; Rodan, G.A. Interleukin-1 induces multinucleation and bone resorbing activity of osteoclasts in the absence of osteoblasts/stromal cells. *Exp. Cell Res.* **1999**, *247*, 84–93.
 87. Sacomen, D.; Smith, R.L.; Song, Y.; Fornasier, V.; Goodman, S.B. Effects of polyethylene particles on tissue surrounding knee arthroplasties in rabbits. *J. Biomed. Mater. Res.* **1998**, *43*, 123–130.
 88. Lohmann, C.H.; Dean, D.D.; Bonewald, L.F.; Schwartz, Z.; Boyan, B.D. Nitric oxide and prostaglandin E2 production in response to ultra-high molecular weight polyethylene particles depends on osteoblast maturation state. *J. Bone Joint Surg.* **2002**, *84A*, 411–419.
 89. Kwon, S.Y.; Lin, T.; Takei, H.; Ma, Q.; Wood, D.J.; O'Connor, D. Alterations in the adhesion behavior of osteoblasts by titanium particle loading: inhibition of cell function and gene expression. *Biorheology* **2001**, *38*, 161–183.
 90. Kwon, S.Y.; Takei, H.; Pioletti, D.P.; Lin, T.; Ma, Q.J.; Akeson, W.H. Titanium particles inhibit osteoblast adhesion to fibronectin-coated substrates. *J. Orthop. Res.* **2001**, *18*, 203–211.
 91. Pioletti, D.P.; Takei, H.; Kwon, S.Y.; Wood, D.; Sung, K.L. The cytotoxic effect of titanium particles phagocytosed by osteoblasts. *J. Biomed. Mater. Res.* **1999**, *46*, 399–407.
 92. Engh, C.A., Jr.; Moore, K.D.; Vinh, T.N.; Engh, G.A. Titanium prosthetic wear debris in remote bone marrow. A report of two cases. *J. Bone Joint Surg. Am.* **1997**, *79* (11), 1721–1725.
 93. Wang, M.L.; Tuli, R.; Manner, P.A.; Sharkey, P.F.; Hall, D.J.; Tuan, R.S. Direct and indirect induction of apoptosis in human mesenchymal stem cells in response to titanium particles. *J. Orthop. Res.* **2003**, *21* (4), 697–707.
 94. Wang, M.L.; Sharkey, P.F.; Tuan, R.S. Particle bio-reactivity and wear-mediated osteolysis. *J. Arthroplasty* **2004**, *19* (8), 1028–1038.

Wheeled Mobility: Wheelchairs and Personal Transportation

Jon L. Pearlman

Rory A. Cooper

Rehabilitation Science and Technology, School of Health and Rehabilitation Sciences, University of Pittsburgh, and Human Engineering Research Laboratories, VA Pittsburgh Healthcare System, Pittsburgh, Pennsylvania, U.S.A.

INTRODUCTION

There are an estimated 600 million people with disabilities (PWD) worldwide. In the U.S.A. alone, there are over two million full-time wheelchair users, roughly 150,000 of them user electric powered wheelchairs (EPWs).^[1,2] The population of part-time wheelchair (WC) users is much larger, and another subset of the population uses scooters to facilitate community and home mobility (approximately 142,000^[1,2]). These mobility devices fall into the family of devices named assistive technologies (AT), which are governed by regulations that can simultaneously ensure the products are of high quality (protecting consumers), but also restrict the types of technological innovations that are eligible for reimbursement from public and private insurance companies (the most typical payer of these devices). Like most technology, significant advancements have been made in AT quality and function over the last several decades. This has significantly impacted on the lives of PWD, as more complex disabilities can be accommodated with these advanced AT, allowing a larger population to utilize these technologies. Unfortunately, this technological advancement is slow, especially with manual WCs and EPWs. The overarching reason for this is that a large portion of society still sees WC users as unable to fully participate in society (regardless of the AT provided). There is a stigma associated with WC users, which stems from what is termed a “medical” model of disability—where PWD are judged on whether they can be “cured.” Progressive thinkers realize that a “social” model of disability is more appropriate—where technology minimizes the consequence of impairment and maximizes the opportunity for PWD to integrate and participate in society; this maximizes the likelihood that PWD will lead more fulfilling lives by integrating socially and having careers (and thus contributing to rather than being supported by society). The cascade of the “medical” model thinking has prioritized WCs under other AT, such as prosthesis, which affects both technological advancement (as WC research and development funds are not high priorities), and also reimbursement mechanisms (as insurers encourage the

lowest-cost WCs). While the medical model misconceptions are still prevalent, they are fading, which is met with advances in WC technology.

Advances in manual WCs are largely focused on setup, injury prevention, and weight—much of this progress has been driven by research showing correlations between secondary injuries (such as carpal tunnel syndrome), propulsion style (related to WC setup), and the weight of the WC and user. EPWs have become more complex; with active–passive suspension to minimize potentially injurious road vibrations, seats that can be adjusted in several degrees of freedom to both minimize the incidence of pressure ulcers (another commonly deadly secondary injury) and allow the user to reposition to maximize their interaction with the world (reach high cabinets, cook, maneuver under desks, etc.). The most advanced EPWs are mobile robots, constantly evaluating the users input (desired direction) and the physical world to maximize the safety and maneuverability of the device in all physical environments.

Electric powered scooters are similarly mobility devices for PWD, but are most suited for people who need part-time mobility accommodation. These devices are commonly used by individuals who tire or fatigue quickly or are in pain when ambulating long distances. Furthermore, they are commonly used outdoors, and can be critical to allow social integration of PWD. While PWD in Western countries have, in many cases, benefited from the advances in mobility devices, PWD in developing countries still lag far behind. Many countries have laws similar to the US Rehab Act of 1973 and the American with Disabilities Act (India, for example), but they have neither the medical reimbursement nor the regulations in place to ensure PWD have access to appropriate AT and full access to public buildings and job opportunities.

DEMOGRAPHICS OF WC USERS

The number of WC and scooter users has nearly sextupled between 1970 and 2000, and expansion is



expected to continue at a significant rate.^[2] This expansion stems from several sources. First, until after World War II, people with spinal injuries or traumatic brain injuries had short life expectancies. As medical technology advanced, so did the life expectancy, which necessitated AT that could be functional (for younger individuals) and that could last longer. Much of the technology, recreational devices and activities, and occupational rehabilitation stem from the military's response to the surviving wounded soldiers. Second, and similar to above, medical technology has increased the life expectancy of the populous as a whole. Because there is a strong positive correlation between age and disability, the population of PWD naturally increases, expanding the need for WCs and other mobility devices. Third, the quality and function of AT itself has had a dramatic impact on the number of people who can use that technology. Because of advances in EPWs, people who were previously confined to their bed can be more independent and mobile; thus, as AT improves, more people can utilize it, effectively expanding the market. Finally, stigmas that have made people hesitant to use wheeled mobility are being broken down. As the devices become more socially acceptable, they are more acceptable to potential users. For instance, individuals who have been using crutches their entire life to help them ambulate (at the cost of shoulder injuries) are switching to WCs and EPWs, which allow them to be more mobile and involved in society. Notably, the electric powered scooter is a common choice for PWD interested in socially acceptable devices. Scooters do not have the same stigma attached to them as WCs. It is common in our clinic (Center for Assistive Technology, University of Pittsburgh Medical Center) to encounter PWD who are candidates for EPW (to facilitate in-the-home mobility) to choose to order a scooter because they consider it to be more "socially acceptable," regardless of whether it is the most appropriate device.

MANUAL WCs

Surprisingly, the most common WC manufactured today is of nearly the same design as the Everest and Jennings (E&J) design of the 1950s (Fig. 1). These "depot" WCs are designed and manufactured for institutional use (hospitals, airports, etc.), but were (and unfortunately still are in many cases) accepted as the standard WC for some users. Today, these low-quality WCs are mostly manufactured in China and Taiwan and are the standard of care for developing countries. As discussed above, injured soldiers who



Fig. 1 E&J style WC.

were young and wanted to be mobile drove the advances in manual WCs. While the original E&J weigh roughly 80 lb, current titanium and carbon-fiber WCs can weigh less than 15 lb. Weight has a significant impact on maneuverability and portability, allowing WC users to accelerate and decelerate very quickly and independently load the WC into their automobile.

Research has shown that weight is an important factor in secondary injuries, including median nerve conduction, carpal tunnel syndrome, and impingement of tissue in the shoulder.^[3] A heavy WC can have a cascading negative affect on the health of the upper extremity musculoskeletal system. To accelerate heavier WCs, users must apply more push-rim force (based on purely physical arguments). These higher force requirements put increased loading on the joints, increasing the risk for musculoskeletal injury.^[3] Practical observations and engineering arguments^[4] suggest that WC fitting is an important issue in WC prescription. The E&J had a fixed axle position, and the seat was set in a 90° (seat-to-back angle) configuration. Axle position adjustment allows the users or technician to shift the weight distribution of the user with respect to the wheels. Shifting the weight can allow individuals to balance on the back tires (allowing them to go up and down curbs and maneuver through rough terrain) and reduce their rolling resistance, and also allows users to optimize their propulsion efficiency (by positioning the wheels so they can take long strokes).^[4-6] Adjusting the seat setup has allowed users to minimize the risk of deadly pressure ulcers, maximize stability, and accommodate postural deformities—all providing more safe and functional mobility for the MWC user. While adjustability is always encouraged (to allow

customized setup of the MWC to the user), once a user has found their optimal axle and seat configuration, they may choose to have a custom, non-adjustable WC built to these specifications; these customized WCs are lighter than their adjustable counterparts, because they shed the weight of the fasteners needed for the adjustable components.

Manual WC frames vary widely, but most fall into two categories—either folding or rigid. Folding WCs commonly have a cross-brace that allows them to collapse laterally and fit into a narrow cabinet. Rigid frame WCs are non-collapsible, but may have a folding backrest. When the wheels are removed and the backrest is folded down on a rigid WC frame, the WC can be easily transported. The benefits of the rigid frame are that they are lighter, and the increased stiffness of the frame tends to make them more efficient, comfortable, and maneuverable.

Other specialized manual WCs exist for both adult and pediatric applications. Adult variations typically include specialized seating systems to accommodate people with postural deformities or reduced neuromuscular function. Both adult and pediatric WCs can be highly specialized, allowing manual tilt and recline functions and specialized postural support devices. Specialized WCs to accommodate people with complex disabilities tend to be highly adjustable, and many accessories can be added so to meet the needs of the user.

EPWs

The first EPWs were developed in the early 1900s and were based on manual WC structural designs.^[6] The manual WC wheels were driven by motors retrofitted to the WC. The power and drive systems were coopted automobile parts, using the 12-V liquid-acid batteries for power, and the starter motors as the drive motors. Early controllers used solenoid or relay devices that performed simple on/off switching for each direction (rather than proportional control), making the devices hard to maneuver.^[7] Major advances in EPW did not occur until the 1960s and 1970s, when battery, controller, and structural components improved, making the EPW a viable mobility aid for PWD.

STRUCTURAL COMPONENTS

As noted above, original EPWs frames were identical to manual WCs, where the seat is integrated into the main structure of the WC (Fig. 2). These traditional designs still exist; although in most cases, they have been strengthened to support the added weight of the batteries, motors, and controller. A more recent



Fig. 2 Conventional EPW based on manual WC frame (Alero shown).

development is the power base design, where the main structural and drive components are integrated into the base unit, and the seating system is added to the base via a post or other substructure (Fig. 3). The power base design is novel, because it allows for different seating systems to be used on the same power base. Also, it allows the base to be optimized by engineers, without affecting the seating systems that can vary based on the clients needs.

POWER AND CONTROL SYSTEMS

As discussed above, the first power and control systems were coopted parts from automobiles. These systems have made steady advancement over the last decades and now allow EPW and scooter users to maneuver their WCs or scooters accurately and safely with a number of different input devices (joysticks, switches, etc.) over different terrains and for substantial distances before they need to be recharged. Original 12 V DC designs have been upgraded to 24 V DC systems (in most cases, although higher voltages are sometimes used), which increases the efficiency of the motors. Motors are typically permanent magnet with worm- or spur-gear reduction. Worm drives are most common, because the motor axle and drive axle are perpendicular, allowing the motor to be easily bolted to



Fig. 3 Power base style EPW.

the frame or suspension system. Spur-gear motors are more efficient, but can be noisier and are more difficult to design into the structural system, given the motor is oriented in-line with the drive axle. Typically, motors have solenoid deactivated magnetic fail-safe brakes fixed to the end of the motor arbor. A voltage is applied across the solenoid by the controller to release the brakes before current passes through the motors. Thus, if there is a controller or power-system malfunction, the brakes turn on to stop the device.

While the motors operate off of 24 V, the controllers are 12 V, and use an amplifier and a metal-oxide semiconductor field-effect transistor (MOSFET) H-bridge in lieu of relays. Thus, the motor speed is varied using a pulse width-modulated oscillating square-wave. The higher duty cycle, the more effective current passes through the motor, increasing the speed. Proportional control of the MOSFET circuitry is typically provided by an X-Y joystick that users manipulate.

While the joystick is most prevalent, and can be moved to be manipulated by the foot, head, or mouth, many other input devices are available to take advantage of the user's functional abilities. Examples of these

input devices include a steering wheel type input (for use with the foot or hand), a sip-and-puff device to be used by blowing or sucking on a straw-like device, or a single or multiswitch selection type of input. Individuals with severely limited sensory-motor function, who have only one reliable body movement can drive via a single switch. For these situations, a controller can be programmed to automatically scan through the different possible directions, and the user presses the switch to select the desired direction. More advanced input devices are being developed, including variable compliance joysticks, which can be tailored to read kinematic motion (like a typical joystick) combined with force production.^[8] Some individuals can produce force in certain direction, but have limited range of motion (owing to contractures, for example); this joystick could be tuned to read this input, allowing the individual to safely maneuver the EPW.^[8] Head position input devices are also being developed to derive input information from the kinematic motion of the head relative to the headrest.^[9] While head-controlled joysticks have been in use for many years, they are only effective for individuals with the appropriate range of neck motion. The new head position monitor based on kinematic tracking of magnets worn on a hat or headband will allow more individuals to safely maneuver an EPW. The forefront of input devices are those directly reading brain inputs either inter- or extracranially.^[10] For individuals with the most disabling impairments, advances in reading and filtering volitional brain signals may be the only means for them to independently and safely maneuver an EPW.^[10]

Related to the advancing input devices is the prospect of being able to tune the input device to the individual through software. The latest EPW controllers are microprocessor based, but still rely on simple control algorithms that only have approximate models or no model of the EPW and the user (the system). Programmable controllers are included on most EPWs, allowing parameters such as maximum and minimum speed and acceleration and directional sensitivity (forward/reverse) of the joystick to be changed. Furthermore, some controllers integrate feedback information from the motors and/or seating functions to adjust or limit the speed to make maneuvering more safe. While currently available controller technology has come a long way over the past decades, innovations are still be made in controller technology. For example, controllers are being developed to integrate more feedback information from sensors attached to the EPW, which can detect nearby obstacles or the inclination of the WC and adjust the controller to maximize the safety of the device.^[7] Similarly, controllers are being augmented to allow for more advanced programming of the input device. This can be useful, for example, if a user has a complex

impairment limiting their functional range of hand motion in one direction (e.g., anterior/posterior) but not in another (e.g., medial/lateral). Current advanced controllers can scale the sensitivity symmetrically (forward/reverse and/or left/right), but advancements are being made to allow complex mapping between user input and controller output (to the motors).^[11] In practice, the required mapping function would be determined by allowing the prospective EPW user to interface with a joystick attached to a “virtual” EPW driving task (much like a computer or video game). By performing computer-guided driving tasks, an optimal mapping function would be developed and loaded into the controller.^[11]

SEATING

Seating systems, especially those used in the power base style EPWs, can vary widely both in their postural support and in their functionality. Low-tech seating systems may allow for good postural support and are minimal or non-adjustable. Higher-tech seating systems can include adjustable postural supports and custom-formed cushions. Still higher-tech seating systems integrate functions that can be controlled by either the user or their assistant. Furthermore, these functions can be manually operated or powered by actuators that are integrated into the original power and control system of the WC. These seating systems can include all of the following functions: recline of the backrest, tilting of the seat, elevation of the seat, extension of the footrests, standing mechanism, and lateral tilting of the seat.

The importance of these seating functions cannot be overstated, for both clinical reasons, and to allow the individual to be more independent and integrate into society. Clinically, changing the seated posture helps to relieve high stresses transferred between the seat and the user's soft tissue, helping to reduce the incidence of pressure ulcers. Furthermore, postural changes are required to lower the risk of long-term postural deformities. While upright sitting (with the back at 90° to the horizontal) may be functional and required for some tasks (such as eating or working), it is an inherently unstable position for individuals with neuromuscular impairment. Gravitational forces acting on the torso tend to initiate and/or accelerate postural deformities (such as a kyphosis), when the spine is in a vertical position (Fig. 4). A slight posterior tilt of the seat and back will stabilize the spine by distributing gravitational forces acting on the body through the backrest, rather than vertically through the spine, stabilizing the torso.

Seating functions have a pivotal role in promoting functional independence for EPW users, in all contexts

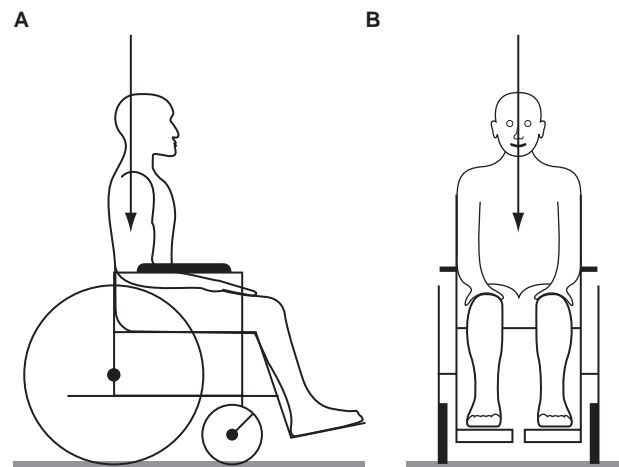


Fig. 4 Sagittal (A) and coronal (B) views of the effects of gravity on the vertical spine. When the spine is vertical, and the torso is not confined with postural devices, there is a tendency for the spine to deform either laterally or anteriorly. This can be avoided with slight posterior recline or tilt.

of their life—in the home, community, and work. For example, in the home, a power seat elevation system can allow independent and safe meal preparation (reaching the stove, into cabinets, etc.). Seat elevators can also be essential for independent transfers: By adjusting the PWC seat level, user can take advantage of gravitational forces by always transferring to a lower seat height (bed, couch, toilet, etc.). By allowing independent transfers and meal preparation [along with other activities of daily living (ADLs)], the seat elevator can decrease or eliminate the need for an assistant. However, they are almost universally denied by public and private health insurance as *not medically necessary*.

Besides providing pressure relief functions, power tilt and recline seating functions allow the user to position themselves into a comfortable posture. This improves the user's quality of life, and will also minimize the need for transfers; if a user is comfortable in their EPW for a longer period of time, then they are less likely to want to move to a more comfortable chair or couch. Importantly, we discuss the tilt and recline functions together in this entry. This is because in only very few cases should recline be used independently from tilt. Recline alone allows for improved pressure distribution when the seat back is reclined slightly. But as the seatback is moved into to an upright position after it has been reclined, the user is shifted forward in the seat, and shear stresses increase on the user's buttocks. Mechanically, this is because the pivot point of the seatback is not at the point of hip rotation. Clinically, these shear stresses can accelerate the formation of pressure ulcers, possibly negating the affect of the beneficial pressure distribution provided by the recline

function. Only users with intact sensory systems in their buttocks and the motor ability to reposition should be candidates for recline-only PWC seating systems, as they would be aware of the shear stresses on their buttocks and can then reposition to relieve these stresses.

Occupational tasks may require a person maneuvering under a desk, or reposition in a standing posture to teach a group of students. Seat elevators and standing functions on a PWC can allow for this repositioning. Changing to a standing posture may also help the user clinically. A survey study suggests that standing may be associated with increased circulation, increased range of motion of the legs (helping to avoid contractures), improved bladder function, and increased bone mass.^[12] Furthermore, because force is relieved off the buttocks, standing can help lower the risk of pressure ulcers.

SCOOTERS

Scooter technology has changed only slightly over the past decades. Structurally, these devices are typically three wheeled (although some are made with four wheels) and have a tiller-type steering system (Fig. 5). Thus, a fairly intact range of motion of the upper extremities is necessary to safely drive the scooter. The power system is 24 V DC, and scooters typically use a single permanent magnet motor that is attached to a small differential, which drives the rear wheels. Like in an automobile, the differential allows each wheel to rotate at a different speed when the scooter is turning. Because of the single motor, and the fact that the turning mechanism is mechanical, the controllers for



Fig. 5 Scooter (Rascal shown).

scooters are much simpler than those for EPWs. The control system has an H-Bridge MOSFET circuit, but it takes only forward/backward information from the input device on the tiller. The input device is typically a small lever that operates a rheostat and can be depressed either by the left or right thumb. In most cases, pressing the lever with the right thumb is the forward direction, and pressing the lever with the left thumb is reverse.

Scooters tend to be laterally unstable because of the user's high center of mass, narrow rear wheelbase, and relatively high top speed. Design changes have been made to try to minimize this instability. Four-wheel scooters have been developed, which increase lateral stability. Also, small antitipper wheels have been placed near the front wheel to help prevent the scooter from tipping completely on its side if it does begin to tip. Like power and manual WCs, scooters are subjected to the American National Standards Institute/Rehabilitation Engineering and Assistive Technology Society of North America (ANSI/RESNA) standards. Unfortunately, the static stability tests, where scooters sometimes perform quite poorly, are not pass/fail; instead, they just report the angle where stability is lost.

Scooters tend to be the best for individuals who need a part-time mobility aid because they tire easily or have joint pain. The reasons for this are several fold, including the lateral stability issues described above, the mechanical tiller, and the fact that maneuverability of the scooter is relatively poor. For example, a user would likely find maneuvering around their house difficult in a scooter because of its large turning radius owing to its long front-to-back wheelbase. An EPW is generally a much better choice for individuals who need mobility in their home. However, there are mid-wheel drive scooters designed for use in the home, which may be appropriate for some clients.

Regardless that there are drawbacks of scooters, they are widely used, and their numbers are likely to increase. One of the reasons for their current popularity is that they cost around half of the price of an EPW. Furthermore, the largest population of scooter users is that of older generations; as people age, they tend to be less able to walk long distances because of fatigue or pain. Because there is a growing population of older individuals in U.S.A. (as well as other countries), there will likely be an increased need for scooters. The prevalence of scooters in society has removed much of the stigma associated with them. Thus, scooters have been a good example of how AT can be accepted into society, which has not fully occurred with manual WCs and EPWs. On the other hand, because of their social acceptance, many individuals who may be better served with a manual WC or EPW would rather have a scooter. This poses a challenge for clinicians providing the devices, who want to provide

the device that is in the best interest of the user (both functionally and socially). Until manual and EPWs have less of a stigma associated with them, clinicians may struggle to advise their client on the technology that is best for them.

ADVANCED MOBILITY DEVICES: HYBRID DEVICES AND MOBILE ROBOTS

Aside from the advancing power and manual WCs, new, novel AT have emerged. Pushrim-activated power assist WC (PAPAW) wheels are an example of this technology.(Fig. 6).^[1] Rather than having a purely manually propelled WC, these devices incorporate a motor into the hub of a manual WC wheel, and are actuated by the push-rim. Thus, these WCs are propelled like manual WCs, but have a power assist. These devices can be used as a transition between manual and powered mobility, and are especially useful for manual WC users who cannot propel their WCs for long distances because of upper extremity pain or fatigue.

Another example of these advanced devices are termed “mobile robots.” These devices are significantly more advanced than power chairs, incorporating advanced mobility and sensor functions. One example of these mobile robots is the IBOT 3000 mobility device, from Johnson & Johnson (Fig. 7). Using redundant computers, gyroscopes, and a number of sensors, the device constantly evaluates its environment and adjusts based both on the users commands and on the sensor information. This allows the device to expand the basic functions of a manual WC by



Fig. 6 PAPAW. Quickie extender shown.



Fig. 7 IBOT 3000 mobility device from Johnson & Johnson.

allowing it to balance on two wheels and climb stairs, significantly expanding the function of the EPW.

Other mobile robots are being developed and are based on EPW and PAPAW designs. The overall goal of these devices is to use sensor information to provide feedback into the controller unit on the device.^[7] This essentially makes the device more informed about the environment and/or the user. This feedback can be used to alter how the user input is read into the controller. For example, for users with vision impairments, who are typically unable to use power mobility for safety reasons, PAPAWs are being designed with sensors to detect obstacles.^[13] Thus, user’s command to maneuver forward would be augmented to avoid an obstacle. This makes using these mobility devices more safe, allowing a broader population of PWD to use them, and ultimately affording them more independence.

Efforts are also being made to integrate more sensors into EPWs to give feedback of the position and the orientation of the device.^[14] Approximately, 70,000 WC accidents are reported per year in U.S.A. with 70% of them attributed to tips and falls.^[15] More extensive user training may be one way to reduce the incidence of these accidents, but another possibly more reliable way is to program the EPW controller to

respond to the physical environment.^[16] Thus, if the user is on a slope, the WC controller would reduce the maximum speed to lower the risk of tipping. Furthermore, there is increasing evidence that vibrations endured when driving an EPW may cause neck and back injuries.^[17] Accelerometers giving feedback to active suspension systems could be dynamically tuned to ensure that the risk of injury was minimized.

WC STANDARDS

Formal development of WC quality assurance standards did not begin until 1979, and was funded by the Department of Veterans Affairs (VA) Rehabilitation Research and Development, and the US Department of Education through the National Institutes of Disability and Rehabilitation Research (NIDRR) Rehabilitation Engineering Research Center (RERC).^[6] Formal US standards have been developed and are titled ANSI/RESNA WC Standard Volumes I and II.^[18,19] International versions of these standards are under the auspices of the International Standards Institute (ISO) and are numbered 7176.^[20] These standards have been important in improving WC quality in U.S.A. and throughout the world, and have important components that are relevant to consumers, WC designers and manufactures, service delivery providers, governmental regulatory departments, and insurance providers.

These standards are broken up into 20 sections, and additional tests are performed on powered devices (WCs and scooters) to ensure electrical safety. Sections include information disclosure; static and dynamic stability analysis; speed, acceleration, and retardation analysis; static and fatigue strength analysis; and environmental testing such as hot, cold, and wet weather and storage testing. Electric powered devices also undergo short-circuiting (of the power and charging system), battery overcharging tests, and energy consumption tests. Recently in 2000, the ANSI/RESNA^[21] and^[22,23] WC standards have been broadened to include standards, ensuring that WCs and tie-down mechanisms are appropriate for use in both public and private transportation (busses, automobiles, etc.)

Data resulting from applying the standards to select WCs has been published^[24-28] and provides comparison data for consumers and clinicians who provide these devices to select WCs. These studies have highlighted the more hazardous and less durable products that the consumer should avoid. WC standards also provide a basic guideline and a set of minimum requirements for WC designers, allowing them to avoid financially costly design iterations. These standards have also been useful as a regulatory tool, allowing governmental organizations (such as the US

Food and Drug Administration) to require that all products available to consumers meet the ANSI/RESNA standards. Similarly, third-party payers for these devices such as Medicare and other insurance companies also use the ANSI/RESNA standards both as a minimum requirement for the devices that they will reimburse for, and also as a way to distinguish between higher- and lower-rated devices that it will pay varying amounts for.^[1,6]

DISTRIBUTION AND PAYMENT METHODS

The most common purchaser of mobility devices are third-party payers and not the consumers. In U.S.A., Medicare, Medicaid, and the Department of VA pay for many of the WCs and scooters provided. In 2003, Medicare alone paid for over 230,000 EPWs, 830,000 light-weight (K0004) WCs, and 2.2 million depot WCs.^[29] In developing countries, charitable organizations typically pay for the majority of the mobility aids consumed. India and China both have AT reimbursement programs targeted toward the low-wage earners, but the low reimbursement amount (\$150 in India^[30] and \$50 in China^[31]) restricts the types of AT that can be provided to very low cost. The consequence of the third-party paying system is that it is difficult for the consumers to advocate for better technology. In a typical market system, competition and users' demands drive product development and quality. In the case of AT, third-party payers fundamentally play the pivotal role in demanding technological advances and higher quality devices; because these payers are commonly for profit (such as private insurance companies), their primary interest may be in increased profit, which typically does not translate into more advanced (and thus more costly) technology that may be in the best interest of the consumer. Thus, the role of regulatory agencies such as the FDA is important to require insurance companies to provide high-quality devices that comply with the ANSI/RESNA standards. Consumer advocacy groups such as the Paralyzed Veterans of America and United Spinal Association also can successfully lobby governmental organizations to regulate the devices and clinical care provided to PWD.

CONCLUSIONS

Mobility devices are essential to hundreds of millions of people worldwide, to allow them to participate in society, to work, and to have a more fulfilling life. As the average age of the population increases (as it is in U.S.A. and Japan, along with other countries), more mobility devices will be required. Furthermore, the more appropriate the devices are for the prospective

users, the more socially integrated and self-sufficient they will be.

We have described many of the aspects of personal mobility, highlighting the technology, the payment systems, and, in some cases, the social implications of providing more appropriate technology. While these AT are advancing, they still, in many cases, fail to optimally meet the needs of the users; this is especially true in developing countries, but is commonplace in developed countries where either the technology does not exist or the insurance system will not cover the costs of the necessary technology. Because of these issues, there is much work to be done not only advancing the state of AT, but also advocating for people to be provided technology that will optimally meet their needs.

ACKNOWLEDGMENT

We are grateful to our funding sources, including the Department of VA, University of Pittsburgh Medical Center, National Science Foundation, National Institutes of Health, and the Department of Education.

ARTICLES OF FURTHER INTEREST

Gait Analysis; Medical Robotics; Telemedicine

REFERENCES

- Cooper, R.; Cooper, R.; Boninger, M.L. Trends and Issues in Wheelchair Technologies. *Assistive Technology*, In Review **2005**.
- LaPlante, M.P. Demographics of wheeled mobility device users. In *Conference on Space Requirements for Wheeled Mobility*; Buffalo: New York, 2003.
- Boninger, M.L., et al. Pushrim biomechanics and injury prevention in spinal cord injury: recommendations based on CULP-SCI investigations. *J. Rehabil. Res. Dev.* **2005**, *42* (3 Suppl 1), 9–20.
- Brubaker, C.E. Wheelchair prescription: an analysis of factors that affect mobility and performance. *J. Rehabil. Res. Dev.* **1986**, *23* (4), 19–26.
- Boninger, M.L.; Baldwin, M.; Cooper, R.A.; Koontz, A.; Chan, L. Manual wheelchair pushrim biomechanics and axle position. *Arch. Phys. Med. Rehabil.* **2000**, *81* (5), 608–613.
- Cooper, R. *Wheelchair Selection and Configuration*, 1st Ed.; Demos Medical Publishing: New York, 1998; Vol. 1, 410.
- Cooper, R. *Rehabilitation Engineering Applied to Mobility and Manipulation*, 1st Ed.; Medical Science Series; I.o.P. Publishing Institute of Physics Publishing: Bristol and Philadelphia, 1995; Vol. 1, 516.
- Spaeth, D.M. Variable compliance joystick for control interface research. In *First Joint BMES/EMBS conference Service Humanity, Advancing Technology*; IEEE: Atlanta, GA, 1999; USA.
- Bevly, A.; Spaeth, D.M.; Cooper, R. Determining head position to assist electric-powered wheelchair operations for persons with traumatic brain injury. 28th Annual Resna Conference, Atlanta, GA, 2005.
- Moore, M.M. Real-world applications for brain-computer interface technology. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2003**, *11* (2), 162–165.
- Ding, D.; Cooper, R.; Spaeth, D.M. Optimized joystick controller. 26th Annual International Conference of the IEEE EMBS. IEEE: San Francisco, CA, USA, 2004.
- Dunn, R.B.; Walter, J.S.; Lucero, Y.; Weaver, F.; Langbein, E.; Fehr, L.; Johnson, P.; Riedy, L. Follow-up assessment of standing mobility device users. *Assist. Technol.* **1998**, *10* (2), 84–93.
- Ding, D.; Guo, S.; Hayashi, S.; Lopresti, E.F.; Simpson, R.; Cooper, R.A. Software design of the smart power assistance module for manual wheelchairs. 28th Annual RESNA Conference; RESNA: Atlanta, GA, 2005.
- Authier, E.; Wyszomierski, S.; Zipfel, E.; Puhlman, J.; Guo, S.; Cooper, R.A.; Lopresti, E.F. Smart power assistance module mechanical hardware suitable for user testing. 28th Annual RESNA Conference Proceedings; Atlanta, GA, 2005.
- Calder, C.J.; Kirby, R.L. Fatal wheelchair-related accidents in the United States. *Am. J. Phys. Med. Rehabil.* **1990**, *69* (4), 184.
- Cooper, R.A. Intelligent control of power wheelchairs. In *IEEE Engineering in Medicine and Biology Magazine*; IEEE, Ed.; 1995; 423–431.
- Cooper, R.A.; Wolf, E.; Fitzgerald, S.G.; Kellerher, A.; Ammer, W.; Boninger, M.L.; Cooper, R. Evaluation of selected sidewalk pavement surfaces for vibration experienced by users of manual and powered wheelchairs. *J. Spinal Cord Med.* **2004**, *27* (5), 468–475.
- Ansi/Resna. In *American National Standard for Wheelchairs: Additional Requirements for Wheelchairs (Including Scooters) With Electrical Systems*; Rehabilitation Engineering and Assistive Technology Society of North America: Virginia, 1998; Vol. 2.
- Ansi/Resna. In *American National Standard for Wheelchairs: Requirements and Test Methods for Wheelchairs (Including Scooters)*; Rehabilitation Engineering and Assistive Technology Society of North America: Virginia, 1998; Vol. 1.
- ISO, I.S.O., *Wheelchair Standards (7176:1–24)*. 1985–2004, International Standards Organization.
- Ansi/Resna. In *American National Standard for Wheelchairs—Wheelchairs used as seats in motor vehicles*; Rehabilitation Engineering and Assistive Technology Society of North America: Virginia, 2000; Vol. 1, 1–19.
- ISO. *Wheelchairs intended for use in Motor Vehicles (7176–19)*; International Standards Organization; 2001.
- ISO. *Wheelchair Tiedown and Occupant Restraint Systems (WTORS) (10542-1/2)*; International Standards Organization, 2001–2005.



24. Fass, M.V.; Cooper, R.A.; Fitzgerald, S.G.; Schmeler, M.; Boninger, M.L.; Algood, S.D.; Ammer, W.A.; Rentschler, A.J. Durability, value, and reliability of selected electric powered wheelchairs. *Arch. Phys. Med. Rehabil.* **2004**, *85* (5), 805.
25. Fitzgerald, S.G.; Cooper, R.A.; Boninger, M.L.; Rentschler, A.J. Comparison of fatigue life for 3 types of manual wheelchairs. *Arch. Phys. Med. Rehabil.* **1982**, *82* (10), 1484.
26. Pearlman, J.L.; Cooper, R.A.; Karnawat, J.; Cooper, R.; Boninger, M.L. Evaluation of the safety and durability of low-cost electric powered wheelchairs. *Arch. Phys. Med. Rehabil.* **2005**, *86* (12), 2361–2370.
27. Rentschler, A.J.; Cooper, R.A. A comparison of the dynamic and static stability of power wheelchairs versus scooters. Proceedings of the First Joint BMES/EMBS Conference, 1999.
28. Rentschler, A.J.; Cooper, R.A.; Fitzgerald, S.G.; Boninger, M.L.; Song-Feng, G.; Ammer, W.; Vitek, M.; Algood, D. Evaluation of selected electric-powered wheelchairs using the ANSI/RESNA standards. *Arch. Phys. Med. Rehabil.* **2004**, *85* (4), 611–619.
29. Centers for medicare and medicaid services. In *Part B Physician/Supplier Nat'l Data, CY 2003 Top 200 Level II Healthcare Common Procedure Coding System*; , 2004.
30. India, G.O. *Scheme of Assistance to Disabled Persons for Purchasing/Fitting of Aids/Appliances (ADIP Scheme)*; Ministry of Social Justice and Empowerment: New Delhi.
31. Guan, T. *Information Regarding Chinese Reimbursement for Assistive Technology*; Pearlman, J.L., Ed.; 2005, Dalian, China.

Xenografts

Brenda M. Ogle

Transplantation Biology, Mayo Clinic, and Department of Physiology, Mayo Clinic, Rochester, Minnesota, U.S.A.

Jeffrey L. Platt

Transplantation Biology, Mayo Clinic, Department of Surgery, Mayo Clinic, Department of Immunology, Mayo Clinic, and Department of Pediatrics, Mayo Clinic, Rochester, Minnesota, U.S.A.

INTRODUCTION

Among a number of medical advances at the beginning of the twentieth century was the vascular anastomosis, the connecting of the cut ends of blood vessels. First performed in 1896, this technique would launch the field of organ transplantation, which now offers the preferred means of treating chronic failure of the heart, lungs, kidney, liver, pancreas, and, most recently, small intestine. Although a medical triumph, transplantation is equally a source of frustration, as the number of organs needed far exceeds those available. The shortage of organs has spurred investigation of alternative means to treat organ failure. Among these means, the most widely studied and perhaps nearest to clinical practice is xenotransplantation, the use of living cells, tissues, and organs from a nonhuman animal source for human transplantation, implantation, or ex-vivo perfusion. The potential benefits of xenotransplantation extend beyond addressing the shortage of human organs for transplantation. For example, xenotransplantation might be a preferred way to treat certain diseases that afflict human but not animal organs (i.e., HIV and hepatitis). Xenotransplantation could also offer a way of dealing with acute organ failure owing, for example, to new epidemics of myocarditis, viral pneumonia, or nephritis. In addition, xenotransplantation could provide an opportunity for manipulation of the donor organ instead of the host to promote organ acceptance. Manipulation of this type might include expression of genes by production of transgenic animals that promote acceptance of the organ or impact new functions. Finally, xenotransplantation might be used to enable the growth of human stem cells into organs in a surrogate animal.

With the potential uses for xenotransplantation, there are equivalent hurdles to clinical application. The first hurdle is the immune reaction of the host against the graft. Significant progress has been made in understanding this reaction and in devising strategies to overcome it. The second hurdle is physiological incompatibility between the xenograft and the recipient.

The third hurdle is the transfer of zoonotic agents from the graft to the recipient and potentially to other members of society. The following section describes the historical context in which xenotransplantation arose, the potential applications and the hurdles to application, approaches to overcoming these hurdles, and a projection for clinical application.

HISTORICAL PERSPECTIVE

Xenotransplantation has a long history. Some would place its inception more than two hundred years ago with blood transfusion from animals to humans. The first serious attempts at xenotransplantation were made in the early years of the 20th century when Jaboulay^[1] and Ullman^[2] each tried to connect animal kidneys to human blood vessels. The subjects suffered kidney failure and in neither case did the xenograft kidney exhibit more than momentary function. Several other failed attempts discouraged further studies in the field of xenotransplantation until the early 1960s with the advent of immunosuppression. In 1963, Reemtsma transplanted chimpanzee kidneys into thirteen patients. One patient survived for nearly nine months after receiving the organ, before dying from an electrolyte imbalance (the xenograft appeared normal).^[3] This relative success triggered a resurgence of interest in xenotransplantation and led to the conducting of fourteen kidney transplants, eight heart transplants, and eleven liver transplants at various centers. Unfortunately, these attempts were less successful, with approximate survival averaging 15 days (Table 1).^[4]

Due to the limited clinical success of xenotransplantation and the attendant risk of transmitting infectious viruses from the animal to the recipient patient, the U.S. Food and Drug Administration has suggested that solid-organ xenotransplantation not enter clinical trials until there is a “reasonable expectation for success.”^[5] In an effort to improve clinical outcome, extensive benchtop and preclinical studies have been initiated over the past three decades

Table 1 History of whole-organ xenotransplantation

Year	Recipients (surgeon)	Xenotransplant	Survival
1910	1 (Unger)	Monkey kidney	<2 days
1913	1 (Schonstadt)		
1963/64	13 (Reemtsma)	Chimpanzee (12), monkey (1) kidneys	1 case 9 months
1964	3 (Traeger)	Chimpanzee kidney	<49 days
1964	1 (Hume)	Chimpanzee kidney	1 day
1964	6 (Starzl)	Baboon kidney	<60 days
1964	1 (Hitchcock)	Baboon kidney	5 days
1964	1 (Hardy)	Chimpanzee heart	2 hours
1965	2 (Goldsmith)	Chimpanzee kidney	4 months
1966	1 (Starzl)	Chimpanzee liver	<1 day
1966	1 (Cortesini)	Chimpanzee kidney	31 days
1969	2 (Starzl)	Chimpanzee liver	<9 days, <2 days
1969	1 (Bertoye)	Baboon liver	<1 day
1969	1 (Marion)	Chimpanzee heart	4 hours
1970	1 (Leger)	Baboon liver	3 days
1970	1 (Marion)	Baboon liver	<1 day
1971	1 (Poyet)	Baboon liver	<1 day
1971	1 (Motin)	Baboon liver	3 days
1974	1 (Starzl)	Chimpanzee liver	14 days
1977	1 (Barnard)	Baboon heart	5 hours
1977	1 (Barnard)	Chimpanzee heart	4 days
1984	1 (Bailey)	Baboon heart	20 days
1992	1 (Starzl)	Baboon liver	70 days
1993	1 (Starzl)	Baboon liver	26 days

to understand the molecular and biochemical mechanisms responsible for xenograft rejection. Such efforts are discussed in great detail below.

THE SOURCE OF XENOGRAFTS

Clearly, the organs of nonhuman primates would stand the greatest chance of survival in a human recipient, as nonhuman primates are quite similar to humans with respect to genetic composition, anatomy, physiology, immune components, hormones, and enzymes. However, relatedness also extends to receptor molecules for infectious microorganisms including filoviruses (e.g., Ebola virus and Marburg disease) and herpesviruses (e.g., B virus), thus increasing the likelihood of disease transfer from nonhuman primate to human. In addition, because monkeys and apes have long gestation periods and few offspring, they would not supply enough organs to meet the need. One possible exception may be the baboon; however, these animals are smaller than humans, and hence their usefulness to adults might be limited.

Because of the limitations to the use of nonhuman primates as a source of xenografts, many now focus on lower animals such as pigs. Pigs have several advantages over other nonprimate species for xenotransplantation. First, porcine organs are very similar to human organs in size and structure. Second, pigs have large litters and a relatively short gestation period of 114 days; thus, large numbers of organs could become available in a short time. Third, the large litter size and short gestation make housing in specific pathogen-free facilities economically feasible.

Pathogen-free facilities will reduce the risk of transferring exogenous viruses, but the question of endogenous viral transfer remains unclear. All mammalian species harbor endogenous retroviruses. These viruses enter an oocyte or sperm cells at random evolutionary time points and are passed on to progeny like normal genes. Pigs harbor at least one family of retroviruses termed porcine endogenous retroviruses (PERV). It is unclear whether PERVs are able to infect humans *in vivo* and whether they are pathogenic (see the section titled "zoonosis"). This aside, pigs offer a final advantage in that their organs may be genetically

modified to promote organ acceptance by the human host (see the following sections on rejection therapies).

PHYSIOLOGICAL CONCERNS

In using pigs for xenotransplantation, the question of whether physiological compatibility exists between pig organs and the human host must be addressed. In the case of the heart, cardiac outputs and action potentials are quite similar between pigs and humans.^[6] Nevertheless, the innervation and overall morphology of the atrioventricular node in pigs is quite different from that of humans. Such differences could alter heart rate and contractility, thereby increasing the potential for arrhythmias. In the transplantation setting, this difference has not been significantly detrimental, as porcine hearts exhibit excellent cardiac output in immunosuppressed baboons.^[7] Normal sinus rhythm without electrocardiographic ST-segment elevation was also observed and serological parameters indicated little myocardial injury. Unfortunately, early immune rejection has made long-term physiological assessment impossible.

Porcine and human kidneys are also physiologically similar. Kidneys from these two species are of similar size and structure and maintain similar maximal concentration ability (1080 mosmol/liter) and glomerular filtration rate (126–175 ml/h).^[6] Porcine kidney xenografts in primates maintain creatinine (glomerular filtration), urea, sodium, and potassium within the range of normal.^[8] The only insufficiency found relates to porcine erythropoietin. In all cases tested, primates with porcine kidney xenografts experienced severe anemia despite production of porcine erythropoietin.^[9] Primates received species-specific erythropoietin to stimulate primate red blood cell production.

The porcine lung exhibits adequate hemodynamics and gas exchange in baboons.^[10] Liver tissue also seems to function adequately in the short life of the graft, though the electrophoretic motility of the proteins change from that of the primate to that of the pig.^[11] Many of the porcine proteins and enzymes secreted may prove incompatible with a primate host. In addition, xenogeneic proteins may disturb normal metabolic and physiological functions of the recipient. For example, a transplanted liver may secrete prothrombotic or proinflammatory products into the blood of the recipient. There are a great number of proteins that may fail to be cross-species compatible and many more that may disturb normal recipient functions. The critical question is which of these defects can be overcome by the recipient and which must be repaired by pharmaceutical or genetic therapies. Only with extended xenograft survival will the answers to these questions and others related to physiological deficiencies become apparent.

IMMUNOLOGICAL CONCERNS

All xenografts elicit an immune response; however, the impact of the immune response is dictated by the manner in which a xenograft is connected to the recipient. In the case of whole-organ transplants, anastomoses of large blood vessels connect donor to recipient (Fig. 1b). Hence, the blood vessels of organ xenografts are in direct contact with recipient blood and the components of the immune system contained therein. It is primarily the interaction of these immune components with donor blood vessel endothelium that leads to the distinct types of vascular disease associated with xenogeneic organ rejection. The progression of vascular disease in this setting is depicted in Fig. 1 and will be described in detail later in this article. In contrast, cell and tissue xenografts receive blood supply by ingrowth of blood vessels of the recipient. In this setting, the blood vessels of the graft provide a barrier between immune reactants of the transplanted tissue (Fig. 1a). In the following sections we describe the types of rejection seen in whole-organ vs. cellular/tissue xenografts and approaches taken to prevent xenograft rejection.



Whole Organ Xenotransplantation

Hyperacute rejection

Hyperacute rejection is the first immunological hurdle to xenogeneic organ success. Hyperacute rejection is characterized by bleeding into the transplant and formation of platelet microthrombi, which in concert lead to graft destruction within several minutes to a few hours. Hyperacute rejection is initiated by the binding of xenoreactive natural antibodies to the endothelial lining of donor blood vessels. The main antigen recognized by xenoreactive antibodies is a saccharide consisting of terminal Gal α 1-3Gal residues.^[12] The binding of xenoreactive antibodies to this sugar activates the complement system of the recipient. Activation of complement in the xenografts is amplified because complement regulatory proteins are species-specific and, thus, relatively ineffective at controlling complement of foreign species.^[13] The result is rapid insertion of terminal complement complexes into the cell membrane of donor luminal endothelial cells and eventual severe organ injury.

Hyperacute rejection therapies

Hyperacute rejection can be averted in several ways. Hyperacute rejection can be averted by depleting recipient plasma of xenoreactive antibodies using immunoabsorption columns.^[14] This therapy must be

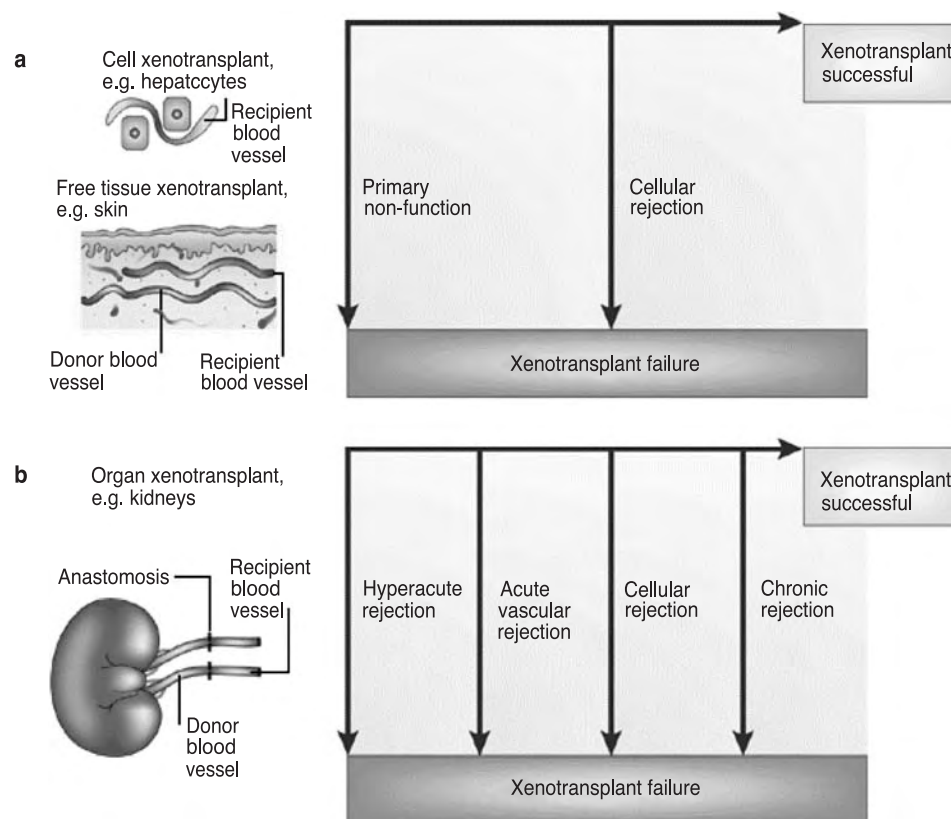


Fig. 1 Biological responses to transplantation depend on the means by which a xenograft receives its vascular supply. a) Cell and tissue xenografts are vascularized, for the most part, by recipient blood vessels. These grafts are subject to failure of engraftment (primary nonfunction) or cellular rejection, but not to vascular rejection. b) Organ xenografts are vascularized by donor blood vessels. The grafts are subject to a series of vascular responses beginning with hyperacute rejection (minutes to hours), acute vascular rejection (days to weeks), cellular rejection, and chronic rejection. (Reprinted and adapted from Ref. [39] with permission.)

used in conjunction with immunosuppressive agents such as cyclophosphamide, mycophenolate, leflunomide, or similar new agents, to prevent rapid antibody resynthesis. These agents alone, despite good success in human allotransplantation, cannot prevent hyperacute rejection. Another way of preventing hyperacute rejection is to inhibit complement. Systemic inhibitors of complement, such as cobra venom factor, soluble complement receptor Type 1 (SCR1), and antibodies against complement component C5 and γ globulin have been used with relative efficacy.^[15] However, inhibition of complement deprives the treated individual of a vital defense against extracellular pathogens and, hence, the treated individual is susceptible to severe infection.

Genetic engineering affords an opportunity to inhibit complement without abridging the defense of the transplant recipient. Specifically, lines of animals that are transgenic for human complement regulatory proteins have been proposed. Organs from these animals would presumably be able to control activation of complement in the xenograft. Organs from transgenic

pigs carrying and expressing human complement regulatory proteins CD55, CD46 (membrane cofactor protein), and CD59 (protectin), either singly or in combination, do not undergo hyperacute rejection after transplantation into nonhuman primates.^[8,16] This result underscores the importance of genetic engineering as a means of improving the outcome of xenografts.

Another way to preclude hyperacute rejection may be to modify expression of antigens recognized by xenoreactive natural antibodies. Transgenic model systems have been developed to accomplish that end. One approach has involved overexpressing an alternate glycosyltransferase (H transferase), which would compete with α 1,3 galactosyltransferase, placing H-antigen residues at the terminus of some sugar chains instead of Gal α 1-3Gal.^[17] To date, these endeavors have shown an ex-vivo decrease in complement-mediated lysis of isolated porcine cells with increased H-transferase, but complete elimination of hyperacute rejection is still uncertain.

Cloning has also been used to modify expression of antigens recognized by xenoreactive natural antibodies. Recent reports describe the successful double knockout in pigs of both alleles of the gene $\alpha 1,3$ -galactosyltransferase. Targeted disruption of one allele was accomplished by homologous recombination in vitro followed by nuclear transfer of the modified cells to generate cloned pigs with the mutant $\alpha 1,3$ -galactosyltransferase allele.^[18] A selection procedure based on a bacterial toxin was used to select for cells in which the second allele of the $\alpha 1,3$ -galactosyltransferase gene was knocked out. Sequencing analysis demonstrated that knockout of the second allele was caused by a T-to-G single point mutation at the second base of exon 9, resulting in inactivation of the $\alpha 1,3$ -galactosyltransferase protein.^[19] Three consecutive rounds of cloning produced the $\alpha 1,3$ -galactosyltransferase double-knockout piglets. The piglets have yet to be extensively tested for expression of the Gal $\alpha 1$ -3Gal antigen. In the interim, homozygous (–/–) fetal-pig fibroblast cells have been produced from heterozygous embryos using a novel selection method. These cells express low (~1%) but detectable levels of the Gal $\alpha 1$ -3Gal antigen.^[20] Whether cells expressing low-level antigen are as susceptible to hyperacute rejection as their wild-type counterparts remains to be seen. As cloning technology continues to advance, it is possible that much simpler and less expensive technology will be developed.

Acute vascular rejection

When hyperacute rejection is prevented, a xenogeneic organ is subject to acute vascular rejection, sometimes known as acute humoral or delayed xenograft rejection. The pathology of acute vascular rejection is characterized by focal ischemia, intravascular fibrin deposition, and thrombosis. Acute vascular rejection may begin within 24 hours of transplantation with graft failure ensuing over a period of days to weeks, or it may begin later and exhibit a more indolent course. Acute vascular rejection has been thought to lead inexorably to destruction of the graft and is, therefore, the main barrier to clinical xenotransplantation.

Acute vascular rejection appears to be initiated by the binding of xenoreactive antibodies to the xenograft, causing activation of endothelium in the graft by three potential mechanisms. First, endothelial cell activation might be induced by the binding of antibodies to cell surface antigens, such as integrins, which can transduce signals.^[21] Second, xenoreactive antibodies could activate complement in small amounts, leading to activation of endothelial cells.^[22] And finally, xenoreactive antibodies could serve as targets for F_c-bearing cells, which may in turn activate endothelium.^[23] Activated

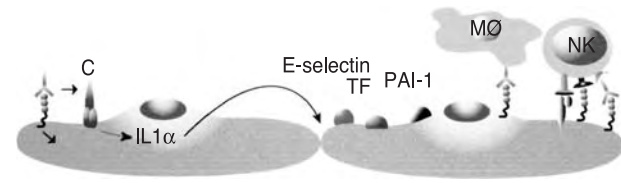


Fig. 2 Pathogenesis of acute vascular rejection. The endothelial lining of blood vessels can be activated by xenoreactive antibodies (Ab), complement (C), platelets, and perhaps inflammatory cells (natural killer (NK) cells and macrophages). Endothelial cell activation leads to expression of new pathophysiological properties. These new properties include synthesis of tissue factor and plasminogen activator inhibitor type 1 (procoagulation) and E selectin and other cytokines (proinflammation). These changes in turn cause thrombosis, ischemia, and endothelial injury, the characteristic features of acute vascular rejection. (Reprinted and adapted from Ref. [40] with permission.)

endothelial cells express procoagulant molecules such as tissue factor and proinflammatory molecules such as E-selectin and cytokines. The pathogenesis of acute vascular rejection is summarized in Fig. 2.

Developing therapies to prevent acute vascular rejection has been more challenging than developing therapies to prevent hyperacute rejection. Part of the challenge is that acute vascular rejection is triggered by low levels of antibody binding. The antibodies that initiate acute vascular rejection include anti-Gal $\alpha 1$ -3Gal antibodies but may also include other antidonor antibodies.^[24] Hence, eradication of one antigen or depletion of antibodies that recognize it may be insufficient to prevent acute vascular rejection. Nevertheless, much effort is underway to generate animals deficient in Gal $\alpha 1$ -3Gal and donor pigs transgenic for complement regulatory proteins, as discussed for hyperacute rejection, to help avert acute vascular rejection.^[8,16]

Acute vascular rejection may also be promoted by NK cells and macrophages. Infiltration of these inflammatory cells leads to synthesis of tissue factor and plasminogen activator inhibitor type 1. Tissue factor and plasminogen activator inhibitor type 1 promote coagulation and synthesis of E-selectin and cytokines, which promotes inflammation. In this manner, the injury triggered by xenoreactive antibodies is further exacerbated.^[25]

Acute vascular rejection therapies

Acute vascular rejection poses a more difficult challenge than hyperacute rejection, owing to the many immunological players involved and to the need for

Table 2 Therapeutic strategies for acute vascular xenograft rejection

Therapeutic approach	Manipulation of recipient	Manipulation of donor
Induction of immune tolerance	Pretransplant infusion with donor bone marrow or genetically engineered Gal α 1-3Gal ⁺ bone marrow	
Induction of accommodation	Transient depletion of xenoreactive antibodies	
Inhibition of endothelial cell activation	Administration of antiinflammatory agents	Suppression of procoagulant or proinflammatory genes
Reduction/elimination of xenograft antigen expression	Specific depletion of xenoreactive antibodies	Breeding pigs with inherently low levels of antigen
	Prevention of xenoreactive antibody synthesis (e.g., cyclophosphamide, leflunomide)	Generating transgenic pigs with low levels of antigen
		Cloning pigs without antigen (i.e., α 1,3 GT knockout and possibly other gene knockouts)

ongoing therapy. Therefore, a multifaceted approach might be envisioned for treating acute vascular rejection (Table 2). One approach involves induction of immunological tolerance. Tolerance in xenotransplantation implies cessation of the synthesis of xenoreactive antibodies. Such a condition might be achieved by transplantation of donor bone marrow, or genetically manipulated bone marrow, to bring about the expression of relevant donor antigens and, subsequently, the deletion or suppression of B-cells of the recipient that secrete antidonor antibodies. Bracy et al. used the latter approach by expressing α 1,3 galactosyltransferase in bone marrow cells by gene transfer. Administration of the Gal α 1-3Gal⁺ bone marrow cells in α 1,3-galactosyltransferase knockout mice decreased production of anti-Gal α 1-3Gal antibodies. One possible limitation to the induction of tolerance is that the antibodies directed against Gal α 1-3Gal are produced independently of T cells. Hence, if anti-Gal α 1-3Gal antibodies are important in causing xenograft rejection, approaches focused on T cells, such as inhibition of costimulation, may not be as effective in inducing tolerance in the xenograft setting.

Another approach to avoiding acute vascular rejection is to bring about accommodation of the graft. Accommodation is an acquired resistance to humoral injury. Accommodation was first observed in the successful transplantation of ABO-incompatible renal allografts.^[26] For this purpose, anti-blood group antibodies were depleted from the recipient before placement of an incompatible allograft. The grafts survived in spite of the return of anti-blood group antibodies. Others found that accommodation can occur in pig-to-primate xenografts, in that organ transplants from pigs can survive in primates for even one month without evidence of acute vascular rejection.^[27] The biological

mechanisms underlying accommodation are still incompletely understood, and thus although it is known that initial depletion of donor antibodies is necessary, depletion duration is unclear. Some have suggested that accommodation might be initiated via expression of antiapoptotic genes such as hemoxygenase 1, A 20 and bcl2,^[28] changes in functional properties of xenoreactive antibodies, changes in antigens, or changes in the graft that render it less susceptible to humoral injury. Careful studies to elucidate the full nature of accommodation are needed, however.

A third approach to avoiding acute vascular rejection is to decrease xenograft antigen expression (as discussed previously for hyperacute rejection). Although decreased antigen expression was initially conceived for prevention of hyperacute rejection, the main application will likely be for prevention of acute vascular rejection. There is evidence to suggest that the level of antibody binding needed to initiate acute vascular rejection is lower than the level needed to initiate hyperacute rejection. Accordingly, the antigen expression on the cells of xenogeneic organs may have to be reduced significantly to achieve therapeutic benefit for acute vascular rejection. Thus, the α 1,3-galactosyltransferase knockout animals (as discussed previously) may offer the greatest therapeutic potential. In addition to lowering or eliminating antigen expression, it is likely that expression of human complement regulatory proteins will be helpful in preventing acute vascular rejection. Cozzi and associates^[9] achieved prolonged survival of xenografts, presumably by preventing acute vascular rejection, by using transgenic pigs expressing high levels of decay-accelerating factor and cynomolgus monkeys treated with very high doses of cyclophosphamide. The immunosuppression may have precluded antidonor antibody synthesis.

Cell-mediated rejection

If hyperacute rejection and acute vascular rejection can be circumvented, a xenograft would be subject to cellular rejection. It is unclear whether and to what extent cellular immunity to xenografts differs from cellular immunity to allotransplants. Here we discuss the potential differences and similarities in cell-mediated rejection of xenografts vs. allografts.

There are four critical issues that likely bear on cell-mediated immune responses to xenotransplantation. The first is the interaction of recipient T cells with donor T cells. There is evidence to suggest that human T cells can recognize pig major histocompatibility antigens as they are expressed in pig cells, just as T cells can recognize major histocompatibility complexes of other human cells of an allograft.^[29] But is recognition of pig cells enough to elicit full T cell activation of human cells? Complete activation between species may be hindered if, for example, incompatibilities of cell adhesion or cell signaling molecules exist. The second important issue is that while in allotransplantation, the polymorphism giving rise to rejection is largely restricted to major histocompatibility complex antigens. In xenotransplantation, almost any porcine protein can give rise to immunogenic peptides. Therefore, the number of different T cell clones directed against xenogenic antigens is likely to be very large. The importance of this diversity of peptide antigens is unknown, but it may prove a great challenge to those interested in applying tolerance to the prevention of cell-mediated xenograft rejection. The third important factor relates to the role of humoral immunity. Allotransplant rejection is more likely if the recipient of an allotransplant has antibodies directed against the donor. Because virtually all xenotransplant recipients will have antibodies directed against the donor, cellular rejection of xenotransplants may be much worse. The final issue relates to the function of immunoregulatory T cells. Defects in T-cell recognition of xenogenic cells may affect the ability of regulatory T cells to control cell-mediated immunity.

Cell-mediated rejection therapies

A variety of immunosuppressive agents and immunomodulatory regimens are currently under investigation to combat cell-mediated rejection. These regimens are based on those used for allograft and include cyclosporin A, tacrolimus (FK506), azathioprin, and/or prednisolone. Because cell-mediated immune responses to xenografts may be more severe than in allografts, it is possible that currently used immunosuppressive agents will not suffice to prevent cellular rejection. On the other hand, a recent study on transplantation of

xenogenic hepatocytes suggests that disparate cells might be transplanted using even less immunosuppression than is used for allografts in humans.^[30] In addition, pharmaceutical therapies for xenotransplantation can be planned in advance (instead of the few hours available when using cadaveric organs). The extra time may allow pretreatment of the xenograft recipient in an effort to lessen the cell-mediated or humoral response. Such pretreatment may extend into the application of immunological tolerance, making certain regimens feasible for xenotransplantation that could not be considered for allotransplantation.

Aside from pharmaceutical regimens, immunological tolerance and/or limiting cell-mediated immunity may also be promoted by genetic engineering. Certainly, efforts to control the natural immune barriers to xenotransplantation may contribute to limiting the elicited cell-mediated response. Given the recognition of donor cells by recipient T cells, the xenogenic donor might be engineered to downregulate expression of corecognition molecules (i.e., CD4 and CD8) or costimulatory molecules (i.e., CD28). Alternatively, or in addition, molecules that inhibit T-cell interactions such as CD59 or Fas ligand might be overexpressed in the donor cells. Expression of inhibitory molecules is currently under investigation for allotransplantation but may prove more effective in xenotransplantation since the inhibitory genes can be introduced as transgenes and, therefore, expressed only in relevant cells in the graft.

Chronic rejection

Chronic rejection in the context of xenotransplantation is rarely discussed, given the short survival of xenografts. It is assumed that chronic rejection akin to that observed for allografts is likely in the xenotransplantation setting. Though the mechanisms of chronic rejection are not fully understood, potential triggers may include repeated acute rejection crises or ischemic injuries of the grafted tissue. Whatever the trigger, the results are ongoing inflammatory reactions and diffuse intimal proliferation in the arterial vessels eventually leading to graft arteriosclerosis. Current immunosuppression therapies are unable to combat chronic rejection in the allotransplant setting and will likely be insufficient for xenotransplants as well. However, unlike the current circumstances in the case of human transplantation, xenografts undergoing chronic rejection could be replaced.

Cellular and Tissue Xenografts

Cell and tissue transplants are of increasing interest as potential means of treating disease and regenerating

diseased organs. Transplantation of isolated islets of Langerhans has been pursued as a treatment for diabetes mellitus, transplantation of hepatocytes for genetic disease of the liver, transplantation of myoblasts for treatment of cardiac failure, and transplantation of neurons for treatment of neurological defects.^[31] Another potential application of cell and tissue transplants is as a means of gene delivery. Although in some cases it is possible to use cells from the affected individual, for some purposes, autologous cells cannot be used, and the main limitation to transplanting human cells and tissues is availability of donors. For this reason, there has been interest in the potential use of animal cells and tissues for transplantation.

Cell and tissue transplants are less susceptible than organ transplants to humoral immunity because the cell and tissue transplants are fed by recipient blood vessels and, therefore, the blood vessels are not subject to attack by xenoreactive antibodies (Fig. 1a). Cell and tissue xenografts that are in contact with recipient blood such as xenogeneic bone marrow cells injected into the systemic circulation or xenogeneic islets injected in the portal vein are subject to injury by xenoreactive antibodies, complement, and other inflammatory mediators.^[32] Cell and tissue transplants are also subject to immediate failure, called primary nonfunction. The mechanistic details of primary nonfunction are not fully understood, but it may be caused by one or more of three factors: 1) the inability of growth factors of the recipient to support newly implanted cells and/or failure of graft factors to support host angiogenesis; 2) the action of natural killer cells or recently activated T cells on the newly implanted graft; and 3) the action of complement on xenogeneic cells and tissues introduced into the blood.

Therapeutic strategies for cellular and tissue xenografts are likely to require little more than presently available immunosuppressive agents. Those cellular/tissue xenografts in contact with recipient blood will require therapies similar to those under investigation for organ transplants (as discussed previously). Another approach to promote cellular xenograft survival may be microencapsulation of the graft with a biocompatible membrane. Microcapsules could protect the xenograft from immune attack while enabling the exchange of nutrients and hormones.^[33] This approach may be particularly useful in the context of pancreatic islet and parathyroid transplantation. It may even prove useful for prevention of endogenous viral transfer.

ZOONOSIS

Recent progress in clinical application of cellular xenotransplantation has brought to the fore concerns

about the possibility of the transfer of infectious agents from the xenograft to the recipient and concern that these agents might spread more broadly in the population. Most pathogens known to cross the species barrier can be readily detected and eliminated from the xenograft source. An exception to this may be the porcine endogenous retroviruses (PERVs), which are present in the germ line of all pigs. Some would argue that since humans have existed in close proximity to pigs for years without known episodes of infection, PERV transfer is unlikely in the transplantation setting. However, PERV can infect cultured human cell lines^[34] and has been shown to infect human cells transferred to SCID mice.^[35] In a multitude of long-term clinical studies with patients exposed to porcine cells (islets and hepatocytes), however, no evidence of retroviral infection has been detected.^[36]

If a given organism such as PERV were to prove a significant hazard, novel therapeutics of various sorts could be applied. For example, introduction of genes encoding retroviral decoys might succeed in preventing or decreasing the production of infectious virus. This type of strategy has proven unsuccessful for the treatment of human retroviruses because the genes could not be delivered to stem cells. This limitation would not apply in animals where the gene of interest could be introduced into the germ line.

In addition to endogenous retroviral transmission, there is the possibility that the recipient will be subject to infections of the xenograft to which humans are not ordinarily susceptible. Thus, it may be useful to investigate the efficacy of animal-specific antimicrobial agents in humans and the efficacy of human-specific antimicrobial agents against animal microorganisms.

ETHICAL CONSIDERATIONS

Experimentation conducted in humans is subject to certain ethical principles (as outlined in the Belmont Report^[37]), including respect for persons, beneficence, and justice. Ethical consideration of xenotransplantation in humans is distinct from this principle because the risks may extend beyond the patient to society as a whole. Therefore, the Institute of Medicine and the Food and Drug Administration^[38] have endeavored to establish guidelines for clinical trials that take into account the potential impact of xenotransplantation on society. The ethical issues under scrutiny by these bodies include granting consent (when potential harm extends beyond an individual to public health), monitoring potential zoonotic infections, assessing psychological effects of receiving an

animal organ, and ensuring appropriate care and use of animal donors.

CLINICAL PROJECTIONS

Although scientific progress in the field of xenotransplantation is significant, substantial difficulties remain, particularly with respect to the immune response of the recipient against the xenograft. There is no reason to doubt that cellular xenotransplantation from pig to human could and in fact has been undertaken^[31] through the use of presently available immunosuppressive agents. However, xenogeneic organ transplantation is a different matter. Hyperacute rejection has been overcome, but the barriers to acute vascular rejection remain intact. And, in fact, the barriers seem insurmountable in light of the possibility that antibodies against antigens other than Gal α 1-3Gal might exist. Still, pharmaceutical progress in overcoming the humoral barrier to allotransplantation and the possibility of genetic engineering provides some encouragement that acute vascular rejection may be prevented or at least suppressed in some individuals. Even so, chronic rejection has yet to be addressed. Chronic rejection is the major cause for allograft loss and if, as some research suggests, chronic rejection is caused by antidonor antibodies, it is likely to be exponentially worse for xenografts. Nevertheless, one redeeming aspect of xenotransplantation validates pursuance of clinical application. If a clinical xenograft were to fail, it might be easily replaced—an option not routinely afforded when a clinical allograft fails. And, in fact, the next step to further xenotransplantation may be bridge or temporary organ transplants to be followed either by an allograft or subsequent xenograft. Bridge transplants will help address the problem of the human organ shortage and will provide opportunity for incisive analysis to overcome the remaining immunological hurdles barring whole-organ xenotransplantation. As a final note, tandem investigation of PERV and other zoonotic agents must be furthered if xenograft safety and eventual FDA clearance is to be attained.

ARTICLES OF FURTHER INTEREST

Allografts; Burn Dressing; Collagen Fixation; Liver, Bio-Artificial

REFERENCES

- Jaboulay, M. De reins au pli du coude par suture arterielles et veineuses. *Lyon Med.* **1906**, *107*, 575–577.
- Ullman, E. Tissue and organ transplantation. *Ann. Surg.* **1914**, *60*, 195–219.
- Reemtsma, K.; McCracken, B.H.; Schlegel, J.U.; Pearl, M.A.; Pearce, C.W.; DeWitt, C.W.; Smith, P.E.; Hewitt, R.L.; Flinner, R.L.; Creech, O. Renal heterotransplantation in man. *Ann. Surg.* **1964**, *160*, 384–410.
- Starzl, T.E.; Murase, N.; Tzakis, A.G.; Fung, J.J.; Todo, S.; Demetris, A.J.; Manez, R.; Marino, I.R.; Valdivia, L. Clinical xenotransplantation. *Xenotranspl.* **1994**, *1*, 3–7.
- Guidance for Industry: Source Animal, Product, Pre-clinical, and Clinical Issues Concerning the Use of Xenotransplantation Products in Humans, Final Guidance*; U.S. Department of Health and Human Services, Food and Drug Administration, Center for Biologics Evaluation and Research (CBER), 2003.
- Kirkman, R.L. Of Swine and Men: Organ Physiology in Different Species. In *Xenograft 25*; Elsevier Science Publishers BV: Amsterdam, 1989; 125–132.
- Brenner, P.; Reichenspurner, H.; Schmoeckel, M.; Wimmer, C.; Rucker, A.; Eder, V.; Meiser, B.; Hinz, M.; Felbinger, T.; Muller-Hocker, J.; Hammer, C.; Reichart, B. IG-therasorb immunoapheresis in orthotopic xenotransplantation of baboons with landrace pig hearts. *Transplantation* **2000**, *69*, 208–214.
- Zaidi, A.; Bhatti, F.; Schmoeckel, M.; Cozzi, E.; Chavez, G.; Wallwork, J.; White, D.; Friend, P. Kidneys from HDAF transgenic pigs are physiologically compatible with primates. *Transplant. Proc.* **1998**, *30*, 2465–2466.
- Cozzi, E.; Bhatti, F.; Schmoeckel, M.; Chavez, G.; Smith, K.G.; Zaidi, A.; Bradley, J.R.; Thiru, S.; Goddard, M.; Vial, C.; Ostlie, D.; Wallwork, J.; White, D.J.; Friend, P.J. Long-term survival of non-human primates receiving life-supporting transgenic porcine kidney xenografts. *Transplantation* **2000**, *70*, 15–21.
- Daggett, C.W.; Yeatman, M.; Lodge, A.J.; Chen, E.P.; Van Trigt, P.; Byrne, G.W.; Logan, J.S.; Lawson, J.H.; Platt, J.L.; Davis, R.D. Swine lungs expressing human complement-regulatory proteins are protected against acute pulmonary dysfunction in a human plasma perfusion model. *J. Thorac. Cardiovasc. Surg.* **1997**, *113*, 390–398.
- Minguela, A.; Ramirez, P.; Carrascosa, C.; Majado, M.J.; Munitiz, V.; Hernandez, Q.; Loba, M.; Munoz, A.; Chavez, R.; Alvarez-Lopez, M.R.; Sanchez-Bueno, F.; Parrilla, P. Identification of porcine proteins in baboon sera after pig liver xenotransplantation. *Transplant. Proc.* **1999**, *31*, 2635–2637.
- Sandrin, M.S.; Vaughan, H.A.; Dabkowski, P.L.; McKenzie, I.F.C. Anti-pig IgM antibodies in human serum react predominantly with Gal α (1,3)Gal epitopes. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 11391–11395.
- Dalmasso, A.P.; Vercellotti, G.M.; Platt, J.L.; Bach, F.H. Inhibition of complement-mediated endothelial cell cytotoxicity by decay accelerating factor: Potential for prevention of xenograft hyperacute rejection. *Transplantation* **1991**, *52*, 530–533.



14. Cooper, D.K.C.; Human, P.A.; Lexer, G.; Rose, A.G.; Rees, J.; Keraan, M.; Du Toit, E. Effects of cyclosporine and antibody adsorption on pig cardiac xenograft survival in the baboon. *J. Heart Transplant.* **1988**, *7*, 238–246.
15. Leventhal, J.R.; Dalmaso, A.P.; Cromwell, J.W.; Platt, J.L.; Manivel, C.J.; Bolman, R.M.; Matas, A.J. Prolongation of cardiac xenograft survival by depletion of complement. *Transplantation* **1993**, *55*, 857–866.
16. Byrne, G.W.; McCurry, K.R.; Martin, M.J.; McClellan, S.M.; Platt, J.L.; Logan, J.S. Transgenic pigs expressing human CD59 and decay-accelerating factor produce an intrinsic barrier to complement-mediated damage. *Transplantation* **1997**, *63*, 149–155.
17. Sharma, A.; Okabe, J.F.; Birch, P.; Platt, J.L.; Logan, J.S. Reduction in the level of Gal (α 1,3) Gal in transgenic mice and pigs by the expression of an α (1,2) fucosyltransferase. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 7190–7195.
18. Lai, L.; Kolber-Simonds, D.; Park, K.W.; Cheong, H.T.; Greenstein, J.L.; Im, G.S.; Samuel, M.; Bonk, A.; Rieke, A.; Day, B.N.; Murphy, C.N.; Carter, D.B.; Hawley, R.J.; Prather, R.S. Production of α -1,3-galactosyltransferase knockout pigs by nuclear transfer cloning. *Science* **2002**, *295*, 1089–1092.
19. Phelps, C.J.; Koike, C.; Vaught, T.D.; Boone, J.; Wells, K.D.; Chen, S.H.; Ball, S.; Specht, S.M.; Polejaeva, I.A.; Monahan, J.A.; Jobst, P.M.; Sharma, S.B.; Lamborn, A.E.; Garst, A.S.; Moore, M.; Demetris, A.J.; Rudert, W.A.; Bottino, R.; Bertera, S.; Trucco, M.; Starzl, T.E.; Dai, Y.; Ayares, D.L. Production of alpha 1,3-galactosyltransferase-deficient pigs. *Science* **2003**, *299*, 411–414.
20. Sharma, A.; Naziruddin, B.; Cui, C.; Martin, M.J.; Xu, H.; Wan, H.; Lei, Y.; Harrison, C.; Yin, J.; Okabe, J.; Mathews, C.; Stark, A.; Adams, C.S.; Houtz, J.; Wiseman, B.S.; Byrne, G.W.; Logan, J.S. Pig cells that lack the gene for alpha 1-3 galactosyltransferase express low levels of the gal antigen. *Transplantation* **2003**, *75*, 430–436.
21. Holzkecht, Z.E.; Platt, J.L. Identification of porcine endothelial cell membrane antigens recognized by human xenoreactive antibodies. *J. Immunol.* **1995**, *154*, 4565–4575.
22. Hattori, R.; Hamilton, K.K.; McEver, R.P.; Sims, P.J. Complement proteins C5b-9 induce secretion of high molecular weight multimers of endothelial von Willibrand factor and translocation of granule membrane protein GMP-140 to the cell surface. *J. Biol. Chem.* **1989**, *264*, 9053–9060.
23. Inverardi, L.; Samaja, M.; Motterlini, R.; Mangili, F.; Bender, J.R.; Pardi, R. Early recognition of a discordant xenogeneic organ by human circulating lymphocytes. *J. Immunol.* **1992**, *149*, 1416–1423.
24. Gaca, J.G.; Lee, W.; Aksoy, O.; Braedehoeft, S.J.; Gonzalez-Stawinski, G.V.; Parker, W.; Davis, R.D. Evidence for polyreactive xenoreactive antibodies in the repertoire of human anti-swine antibodies: The 'next' humoral barrier to xenotransplantation. *Transpl. Immunol.* **2001**, *9*, 19–27.
25. Malyguine, A.M.; Saadi, S.; Holzkecht, R.A.; Patte, C.R.; Sud, N.; Platt, J.L.; Dawson, J.R. Induction of procoagulant function in porcine endothelial cells by human NK cells. *J. Immunol.* **1997**, *159*, 4659–4664.
26. Chopek, M.W.; Simmons, R.L.; Platt, J.L. ABO-incompatible renal transplantation: Initial immunopathologic evaluation. *Transplant. Proc.* **1987**, *19*, 4553–4557.
27. Lin, Y.; Goebles, J.; Xia, G.; Ji, P.; Vandeputte, M.; Waer, M. Induction of specific transplantation tolerance across xenogeneic barriers in the T-independent immune compartment. *Nat. Med.* **1998**, *4*, 173–180.
28. Bach, F.H.; Ferran, C.; Hechenleitner, P.; Mark, W.; Koyamada, N.; Miyatake, T.; Winkler, H.; Badrichani, A.; Cardinas, D.; Hancock, W.H. Accommodation of vascularized xenografts: Expression of "protective genes" by donor endothelial cells in a host Th2 cytokine environment. *Nat. Med.* **1997**, *3*, 196–204.
29. Murray, A.G.; Khodadoust, M.M.; Pober, J.S.; Bothwell, A.L.M. Porcine aortic endothelial cells activate human T cells: Direct presentation of MHC antigens and costimulation by ligands for human CD2 and CD28. *Immunity* **1994**, *1*, 57–63.
30. Gunsalus, J.R.; Brady, D.A.; Coulter, S.M.; Gray, B.M.; Edge, A.S.B. Reduction of serum cholesterol in Watanabe rabbits by xenogeneic hepatocellular transplantation. *Nat. Med.* **1997**, *3*, 48–53.
31. Deacon, T.; Schumacher, J.; Dinsmore, J.; Thomas, C.; Palmer, P.; Kott, S.; Edge, A.; Penney, D.; Kassissieh, S.; Dempsey, P.; Isacson, O. Histological evidence of fetal pig neural cell survival after transplantation into a patient with Parkinson's disease. *Nat. Med.* **1997**, *3*, 350–353.
32. Gritsch, H.A.; Glaser, R.M.; Emery, D.W.; Lee, L.A.; Smith, C.V.; Sablinski, T.; Arn, J.S.; Sachs, D.H.; Sykes, M. The importance of nonimmune factors in reconstitution by discordant xenogeneic hematopoietic cells. *Transplantation* **1994**, *57*, 906–917.
33. Aebischer, P.; Schlupe, M.; Deglon, N.; Joseph, J.M.; Hirt, L.; Heyd, B.; Goddard, M.; Hammang, J.P.; Zurn, A.D.; Kato, A.C.; Regli, F.; Baetge, E.E. Intrathecal delivery of CNTF using encapsulated genetically modified xenogeneic cells in amyotrophic lateral sclerosis patients. *Nat. Med.* **1996**, *2*, 696–699.
34. Patience, C.; Takeuchi, Y.; Weiss, R.A. Infection of human cells by an endogenous retrovirus of pigs. *Nat. Med.* **1997**, *3*, 282–286.
35. Van der Laan, L.J.W.; Lockey, C.; Griffeth, B.C.; Frasier, F.S.; Wilson, C.A.; Onions, D.E.; Hering, B.J.; Long, Z.; Otto, E.; Torbett, B.E.; Salomon, D.R. Infection by porcine endogenous retrovirus after islet xenotransplantation in SCID mice. *Nature* **2000**, *407*, 501–504.
36. Paradis, K.; Langford, G.; Long, Z.; Heneine, W.; Sandstrom, P.; Switzer, W.M.; Chapman, L.E.; Lockey, C.; Onions, D. Search for cross-species transmission of porcine endogenous retrovirus in patients treated with living pig tissue. *Science* **1999**, *285*, 1236–1241.

37. Ryan, K.J.; Brady, J.V.; Cooke, R.E. *The Belmont Report*; U.S. Department of Health, Education and Welfare, 1979.
38. Institute of Medicine. *Xenotransplantation: Science, Ethics, and Public Policy*; National Academy Press: Washington, DC, 1996.
39. Cascalho, M.; Platt, J.L. Xenotransplantation and other means of organ replacement. *Nat. Rev., Immunol.* **2001**, *1*, 154–160.
40. Platt, J.L. New directions for organ transplantation. *Nature* **1998**, *392* (Suppl.), 11–17.

Zirconia Ceramics

Mariano A. Polack

Gainesville, Virginia, U.S.A.

INTRODUCTION

High-performance ceramic materials are increasingly used in a variety of engineering and biomechanical applications. Some significant examples include all-ceramic dental crowns to replace dental enamel, total hip replacement prostheses, and thermal barrier coatings in jet turbine and diesel engines. These ceramics are particularly appealing because of their high hardness, strength, and modulus that provide them with increased resistance to fracture and deformation under long-term cyclic conditions, among other properties. Zirconia is an example of this group of materials and is an attractive candidate for the above-mentioned applications. This chapter will identify the composition and main properties of zirconia, describe its dental applications, and discuss ways of improving its clinical behavior through an understanding of current research.

COMPOSITION AND PROPERTIES

General

Zirconia, or zirconium dioxide (ZrO_2), is a ceramic that can exist in three major phases: monoclinic, tetragonal, and cubic.^[1,2] The monoclinic phase is the low-temperature, stable phase. Pure zirconia is monoclinic at room temperature, and this phase is stable up to $1170^\circ C$ ^[3] but its structure can transition to tetragonal and cubic with increasing temperatures. Upon cooling, the phase transformation from cubic to tetragonal to monoclinic is accompanied by a 1–5% volume expansion that induces large stresses in the material. This causes cracks in pure zirconia, decimating its mechanical properties and making it useless for any of the applications mentioned previously.

Partially Stabilized Zirconia

In order to stabilize ZrO_2 , different oxides that dissolve in its crystal structure can be added. These typically include yttrium oxide (Y_2O_3), calcium oxide (CaO), and magnesium oxide (MgO). The exact mechanisms that strengthen and toughen tetragonal zirconia are beyond the scope of this chapter but, in general, they depend on

the type and amount of stabilizing agent, the grain size, and the thermal history of the ceramic. This generates two different partially stabilized zirconias (PSZ) known as TTZ (tetragonally toughened zirconia) and TZP (tetragonal-zirconia polycrystal). The former is a MgO partially stabilized zirconia, a common commercial material also known as MgPSZ or MgTTZ. The latter requires the addition of 3% Y_2O_3 (or more rarely, ceria) to ZrO_2 as a stabilizing agent and produces a material known as yttrium-tetragonal zirconia polycrystal (YTZP).^[1] This very fine grain ceramic has outstanding toughness and hardness that can be challenging to produce.

Transformation Toughening

When a stress is applied to YTZP and a crack tries to propagate through it, the stress concentration at the leading edge of the crack can induce the YTZP to transform from its tetragonal to a monoclinic configuration.^[4] The monoclinic phase is 3–5% larger than the tetragonal crystal, and therefore the phase change generates a volume expansion at the tip of the crack that places it in compression, preventing its propagation. This process, called transformation toughening, increases the fracture toughness of the ceramic (the material's ability to resist crack propagation)^[5] and makes it appealing for cutting, wear-resistant, and dental applications (e.g., replacing the tooth structure subject to load).^[4] However, the value of this mechanism for the latter application is being questioned.^[6]

Thermal and Electrical Behavior

All toughened zirconias suffer degradation of their properties with increasing temperatures,^[2] limiting their applications to below $800^\circ C$. At room temperature, zirconia is highly resistant but its conductivity increases at high temperatures, and it becomes a very good conductor above $2000^\circ C$.

The addition of sufficient CaO oxide amounts allows the cubic structure of zirconia to be maintained at room temperature, which in turns allows it to be used as a refractory material. Stabilized zirconia, because of its high ionic and low electronic conductivity, is also an important electroceramic. Some other applications for

zirconia include high-density ball and pebble mill-grinding media, rollers and guides for metal tube forming, thread and wire guides, oxygen sensors, and fuel cell membranes.

DENTAL APPLICATIONS

Where severe destruction of the tooth structure has occurred, a dental crown may be needed to replace the loss of enamel or dentin to restore anatomy, function, and esthetics. In the past, dental crowns were exclusively fabricated with a metallic core that was veneered with feldspathic porcelain to provide the esthetic component. These were known as porcelain fused to metal crowns (Fig. 1). The search for a more esthetic, biocompatible, and inert option^[7] has led, in the past few years, to a surge of all-ceramic systems that are rapidly becoming the material of choice for dental crowns.^[8]

The initial attempts at developing such systems started with a monolithic approach (i.e., a single-layer ceramic such as Dicor[®], Dentsply/Caulk, Milford, DE; or IPS Empress[®], Ivoclar, Schaan, Lichtenstein) but presented failure rates between 3% and 6% per year,^[9–11] significantly higher than the 2–3% failure estimate at 10 years of their porcelain fused to metal counterparts.^[12] In consequence, a bilayer design was utilized with an alumina core that was either glass-infiltrated (InCeram[®], Vident, Brea, CA)^[13] or densely sintered (Procera[®], Nobel Biocare, Yorba Linda, CA)^[14] to support the veneering porcelain applied for esthetic reasons. Clinical studies demonstrated a lower failure rate for the alumina systems when used for single crowns,^[15,16] but results have not been encouraging when used for replacing a missing tooth with a single all-ceramic structure (i.e., fixed-partial denture (FPD) or dental bridge) in the posterior intraoral

regions,^[17,18] where forces of mastication are greater. In fact, failure rates in some cases were up to 24% for molars.^[17]

YTZP cores were developed based on a need to improve the clinical performance of all-ceramic systems through stronger substructures with improved physical properties^[7,19] and can be used for anterior or posterior crowns and bridges (Lava[™], 3M ESPE, Minneapolis, MN; DCS Smart-Fit, Austenal, Chicago, IL; Cercon Smart Ceramics, Dentsply Ceramco, Burlington, NJ; Procera[®] AllZirkon, Nobel Biocare, Yorba Linda, CA; Vita YZ, Vident, Brea, CA).^[20,21] Fracture toughness of 9–19 MPa^[22] and a flexural strength between 900 and 1200 MPa^[22] are some of the properties that make YTZP appealing as a core material.

Core Manufacturing

The manufacturing of zirconia cores (for single crowns) or frameworks (for multiple teeth replacement with one structure) will depend on the system utilized. Frameworks for fixed-partial dentures are traditionally fabricated by machining techniques that require conventional waxing first and then scanning or a computer-aided design/computer-assisted manufacturing (CAD/CAM) protocol.^[23] Lava (3M ESPE) is an example of the latter. In order to fabricate a zirconia core or framework with this system, four main steps are needed: a) scanning of the prepared areas; b) computer-assisted design of the framework; c) CAM milling of the partially sintered zirconia block; and d) sintering. After this step, compatible feldspathic porcelain is used to veneer the framework to obtain full anatomic contour and appropriate esthetics (Figs. 2 and 3). During the fabrication process, the substructures are expanded by 20% through a computer to compensate for the sintering shrinkage that takes place for 6–8 h in a



Fig. 1 Lateral and intaglio surface of a porcelain fused to metal crown.



Fig. 2 Intaglio surface of a YTZP fixed partial denture fabricated with the Lava system that will replace the maxillary left central incisor.

Z



Fig. 3 Intraoral view of the fixed partial denture from Fig. 2.

high-temperature oven. A clinical study of posterior three-unit FPDs (i.e., replacing one tooth) fabricated with the Lava system demonstrated a success rate of 100% after a mean follow-up of 31.2 months.^[6] In this study, all assessed parameters of the restorations were rated “alpha” except for minor chipping of the veneering porcelain detected in five FPDs (“bravo” for clinical fracture resistance) and the marginal integrity of one restoration that was rated bravo at 36 months.

The DCS Smart-Fit system (Austenal) is a CAD/CAM system that mills the cores and frameworks from fully sintered blocks under hot isostatic pressure. This results in an extremely hard and dense ceramic,^[1] although milling fully sintered zirconia blocks can introduce microcracks and damage its microstructure.^[24]

Single copings and FPD frameworks fabricated with the Vita InCeram YZ blocks (Vident) and the Cercon Smart Ceramics system (Dentsply Ceramco) also undergo an enlarging and sintering process. The former system employs presintered YTZP blocks^[25] that are milled in the CEREC inLab unit (Sirona, Charlotte, NC), a specific CAD/CAM machine. The substructures are then sintered in a special furnace for 2 h at 1530°C to permit shrinkage to full density.^[26] The Cercon system utilizes partially sintered YTZP with no CAD stage. Instead, the substructure is waxed by a laboratory technician using traditional techniques that are analog to the “lost-wax” fabrication method of a metal coping for a porcelain fused to metal crown. The wax pattern is then optically scanned into a computer and virtually enlarged before the automatic milling (CAM stage) and sintering of the ceramic blocks take place. Veneering of the substructures with compatible feldspathic porcelain follows.

Procera Zirconia (Nobel Biocare) is also a CAD/CAM system that utilizes this ceramic as an option for the fabrication of single^[27] and multiple-unit restorations.^[25] In this system, the master cast is scanned and the information transferred through a modem to a

production unit. Then the dies or frameworks are virtually enlarged to balance the sintering shrinkage and milled out of a solid piece of high-purity zirconium oxide, which is then sintered. The resulting substructures are sent back to a laboratory where they are veneered for esthetics.

One other system (InCeram Zirconia, Vident) employs a glass-infiltrated 35% zirconia/65% alumina material for the fabrication of posterior crowns and FPDs,^[28] which can be slip cast or milled from presintered blocks. Its flexural strength is about 620 MPa,^[29] which is lower than that reported for densely sintered zirconia, the ceramic used in the other systems mentioned above.^[22,30] A 3-year clinical study of posterior FPDs fabricated with InCeram Zirconia demonstrated a success rate of approximately 94.5%.^[31]

Implant Abutments

YTZP also has applications in implant dentistry. When a missing tooth is replaced with a dental implant, an abutment is connected to the implant to support the crown that will be placed over it. In the past, abutments were fabricated in metal (Fig. 4) but with the introduction of high-strength ceramics, alumina and YTZP have been utilized for these purposes whenever the esthetic requirements are high (Fig. 5).^[32] Most recently, there seems to be increased interest in the development of a zirconia implant (versus the traditional ones made of titanium or titanium alloy).^[33]

BEHAVIOR AND CLINICAL CONSIDERATIONS FOR YTZP CROWNS

Radial Cracking

In clinical practice, the most reported cause of failure for all-ceramic crowns is bulk fracture.^[34] Contrary to



Fig. 4 Metal abutments connected to osseointegrated implants (implants not seen) that will support porcelain fused to metal crowns.

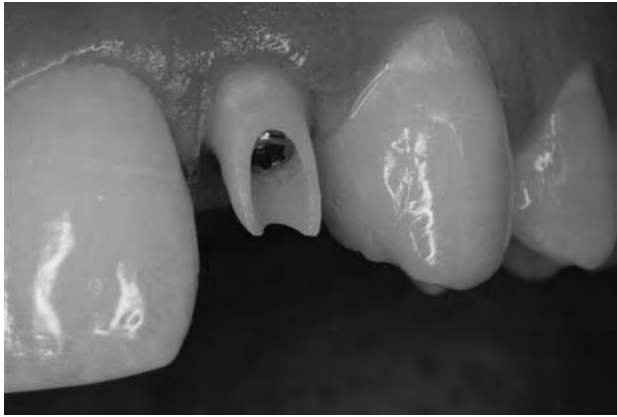


Fig. 5 Zirconia abutment connected to osseointegrated implant (implant not seen) that will support an all-ceramic crown.

intuition, this type of failure does not arise from the weaker occlusal (top) surface but from radial cracks located at the cementation interface (i.e., between the tooth and the ceramic core) generated by a high enough load (critical load).^[35] These radial cracks produced by flexure induced by loads at the top surface have been identified as the most harmful for prosthesis survival.^[19]

Lawn et al. demonstrated that once formed, radial cracks will travel and spread laterally outward as the load increases but will tend to close up, without healing, as the load diminishes.^[19] When this occurs, radial cracks become difficult to detect by top-surface inspection, making them particularly dangerous: A sufficient high load can make the existing cracks penetrate the entire thickness of the core and lead to the catastrophic failure of the restoration.

Radial cracking is highly dependent on the thickness of the ceramic layer (the combined total thickness of the core and the feldspathic porcelain veneer applied over it). This has a direct effect on crown design, where the feldspathic veneer requires a certain thickness for esthetics (approximately 1.5 mm) but reaching this thickness at the expense of the core can increase the chances of radial cracking. Therefore, sufficient reduction of tooth structure must be done when one of these all-ceramic crowns will be utilized, so that the core and the veneering porcelain will have the appropriate dimensions.

Cyclic Fatigue

It has been also questioned whether YTZP will sustain its performance when exposed to mechanical damages, as in prosthesis handling (for crowns) or insertion procedures (for total hip replacements). There have

been numerous reports of fatigue for YTZP stemming from some combination of slow crack growth and mechanical degradation.^[36] Research indicates that cyclic loading causes substantial loss of strength for YTZP, pointing to a mechanical degradation process by slow crack growth. In addition, YTZP can suffer degradation from thermal events that might be encountered during sterilization, for example. This could weaken the material by inducing transformations leading to microcracking.^[37]

In summary, YTZP seems to have some mechanical and thermodynamic instability, requiring strict control during manufacturing and handling to protect it from inadvertent, excessive degradation.^[36] This supports the application of YTZP as a core material in dental crowns, where its surface is somewhat shielded from external influences by the veneering feldspathic porcelain. However, it is worth noting that critical loads for the initiation of radial cracks tend to diminish with time from intrusion of moisture into starting flaws at the cementation interface.

Role of Moisture

There is a potentially dangerous top-surface damage mode called inner-cone crack that appears mostly under cyclic loading in water.^[38] In ceramics such as YTZP, these cone cracks can appear before radial cracks and lead to failure in bilayer structures by entering the brittle coating layer (dentin in the case of dental crowns) and causing delamination.^[39] The mechanical driving force for this phenomenon appears to be hydraulic pumping: as confined water enters and exits the crack interface under pressure during cyclic loading, it forces entrapped liquid toward the crack tip, opening surface flaws and driving the cracks deep into the sublayer.^[40] With more cycles, more water enters the crack interfaces turning into a continuous, deep driving force for the cracks. This demonstrates the deleterious role that the oral environment, and specifically saliva, can have on the longevity of zirconia restorations under these loading circumstances.

Sharp Occlusal Contacts

In addition to the above, YTZP is highly sensitive to fatigue from sharp-indenter damage (micrometer scale indentations) in cyclic flexural stressing.^[40] This means that small indentations below which radial cracking occur can generate significant strength loss in YTZP. For instance, a sharp contact with a load of 0.1 N is as damaging as a 1.98-mm-radius blunt surface contact at 300 N. Therefore, under these conditions, YTZP must be considered a somewhat brittle ceramic that must be



handled with care during fabrication and handling to avoid deleterious strength damage.

In short, sharp contacts will concentrate stress on a small number of microcracks, although much more strongly than blunt contacts. As the load increases, radial cracks develop and slow crack growth becomes the dominant fatigue mechanism.^[41] One clinical procedure that could prove beneficial to neutralize the effect of sharp contacts would be rounding off sharp cusps in the dentition opposing YTZP crowns.

Curved Surfaces

Research on all-ceramic crowns performed on flat surfaces can provide important information about the behavior of these materials. However, teeth are curved, and this could change the results obtained. The influence of surface curvature on the conditions to set off and propagate radial cracks has been studied and, in general, higher loads are needed to initiate these cracks in curved surfaces.^[42] This early inhibition of the initiation of radial cracks by curved surfaces is followed by an acceleration of the propagation of cracks toward the support base, particularly on convex surfaces. At the same time, the load needed to propagate radial cracks to the support base (failure load) diminishes as convexity increases. This study also demonstrated that tests performed on flat surfaces^[43] can provide accurate approximations of the loads needed to initiate radial cracking in more convoluted surfaces. The clinical implication is to avoid sharp convex curvatures during the construction of YTZP crowns.

Sandblasting

Sandblasting the interior surface of an all-ceramic crown is a common procedure when cementing these restorations.^[44] Sandblasting produces a roughened surface that enhances the mechanical bond with dental cements. On the other hand, sandblasting introduces defects that could affect the strength of the crown. Research indicates that under dynamic and cyclic loading, slow crack growth will degrade the strength of YTZP by a factor of 2–4.^[45] Sandblasting YTZP crowns before cementation could further reduce strength levels under certain loading conditions,^[46] outweighing its possible benefits. The recommendations from the literature indicate that the best cementation for YTZP restorations is achieved with an adhesive phosphate monomer containing resin luting cement.^[47]

CONCLUSIONS

The use of high-strength ceramics is increasing at a rapid pace. In this respect, the introduction of zirconia

for biomechanical applications has been interesting because of its enhanced mechanical properties, esthetics, and biocompatibility. The utilization of a strong and stiff core material such as YTZP in dental crowns may offer some protection against strength degradation from cyclic loading and moisture-induced cone cracks compared with other ceramics. However, to improve the life expectancy of replacement structures fabricated with zirconia, a better understanding of its behavior is needed.

Extensive research has been done in this respect, and some design strategies for dental crowns can be applied to maximize their performance, such as providing adequate crown thickness to avoid radial cracks. Additionally, eliminating sharp opposing contacts and extremely convex surfaces during crown construction can enhance the lifetime of YTZP restorations. Finally, a phosphate monomer containing resin–cement should be used for the cementation of zirconia crowns, whereas sandblasting should be avoided to prevent introducing flaws in the core structure that could reduce its strength.

ARTICLES OF FURTHER INTEREST

Biologic and Synthetic Apatites; Bone–Implant Interface; Bone Remodeling; Ceramics; Ceramics in Dentistry; Dental Implants; Implant, Total Hip; Knee Joint Replacement; Orthopedic Biomaterials; Wear Debris, Bone Resorption Animal Models

REFERENCES

1. Sorensen, J.A. The Lava system for CAD/CAM production of high-strength precision fixed prosthodontics. *Quintessence Dent. Technol.* **2003**, *26*, 57–67.
2. Papanagiotou, H.P.; Morgano, S.M.; Giordano, R.A.; Pober, R. In vitro evaluation of low-temperature aging effects and finishing procedures on the flexural strength and structural stability of Y-TZP dental ceramics. *J. Prosthet. Dent.* **2006**, *96* (3), 154–164.
3. Piconi, C.; Maccauro, G. Zirconia as a ceramic biomaterial. *Biomaterials* **1999**, *20* (1), 1–25.
4. Raigrodski, A.J. Clinical and laboratory considerations for the use of CAD/CAM Y-TZP-based restorations. *Pract. Proced. Aesthet. Dent.* **2003**, *15* (6), 469–476; quiz, 77
5. Lawn, B.R. *Fracture of Brittle Solids*, 2nd Ed.; Cambridge University Press: Cambridge, New York, 1993.
6. Raigrodski, A.J.; Chiche, G.J.; Potiket, N.; Hochstedler, J.L.; Mohamed, S.E.; Billiot, S.; Mercante, D.E. The efficacy of posterior three-unit zirconium-oxide-based ceramic fixed partial dental prostheses: A prospective clinical pilot study. *J. Prosthet. Dent.* **2006**, *96* (4), 237–244.
7. Lawn, B.R.; Deng, Y.; Thompson, V.P. Use of contact testing in the characterization and design of all-ceramic

- crownlike layer structures: A review. *J. Prosthet. Dent.* **2001**, *86* (5), 495–510.
8. Oden, A.; Andersson, M.; Krystek-Ondracek, I.; Magnusson, D. Five-year clinical evaluation of Procera AllCeram crowns. *J. Prosthet. Dent.* **1998**, *80* (4), 450–456.
 9. Malament, K.A.; Socransky, S.S. Survival of Dicor glass-ceramic dental restorations over 14 years: Part. I. Survival of Dicor complete coverage restorations and effect of internal surface acid etching, tooth position, gender, and age. *J. Prosthet. Dent.* **1999**, *81* (1), 23–32.
 10. Fradeani, M.; Aquilano, A. Clinical experience with Empress crowns. *Int. J. Prosthodont.* **1997**, *10* (3), 241–247.
 11. Sjogren, G.; Molin, M.; van Dijken, J.W. A 5-year clinical evaluation of ceramic inlays (Cerec) cemented with a dual-cured or chemically cured resin composite luting agent. *Acta Odontol. Scand.* **1998**, *56* (5), 263–267.
 12. Walton, T.R. A 10-year longitudinal study of fixed prosthodontics: Clinical characteristics and outcome of single-unit metal-ceramic crowns. *Int. J. Prosthodont.* **1999**, *12* (6), 519–526.
 13. Probst, L.; Diehl, J. Slip-casting alumina ceramics for crown and bridge restorations. *Quintessence Int.* **1992**, *23* (1), 25–31.
 14. Russell, M.M.; Andersson, M.; Dahlmo, K.; Razzoog, M.E.; Lang, B.R. A new computer-assisted method for fabrication of crowns and fixed partial dentures. *Quintessence Int.* **1995**, *26* (11), 757–763.
 15. McLaren, E.A.; White, S.N. Survival of In-Ceram crowns in a private practice: A prospective clinical trial. *J. Prosthet. Dent.* **2000**, *83* (2), 216–222.
 16. Zawta, C. Fixed partial dentures with an all-ceramic system: A case report. *Quintessence Int.* **2001**, *32* (5), 351–358.
 17. Sorensen, J.A.; Kang, S.K.; Torres, T.J.; Knode, H. In-Ceram fixed partial dentures: Three-year clinical trial results. *J. Calif. Dent. Assoc.* **1998**, *26* (3), 207–214.
 18. Zimmer, D.; Gerds, T.; Strub, J.R. [Survival rate of IPS-Empress 2 all-ceramic crowns and bridges: Three year's results]. *Schweiz. Monatsschr. Zahnmed.* **2004**, *114* (2), 115–119.
 19. Lawn, B.R.; Pajares, A.; Zhang, Y.; Deng, Y.; Polack, M.A.; Lloyd, I.K.; Rekow, E.D.; Thompson, V.P. Materials design in the performance of all-ceramic crowns. *Biomaterials* **2004**, *25* (14), 2885–2892.
 20. Filser, F.; Kocher, P.; Weibel, F.; Luthy, H.; Scharer, P.; Gauckler, L.J. Reliability and strength of all-ceramic dental restorations fabricated by direct ceramic machining (DCM). *Int. J. Comput. Dent.* **2001**, *4* (2), 89–106.
 21. Besimo, C.E.; Spielmann, H.P.; Rohner, H.P. Computer-assisted generation of all-ceramic crowns and fixed partial dentures. *Int. J. Comput. Dent.* **2001**, *4* (4), 243–262.
 22. Christel, P.; Meunier, A.; Heller, M.; Torre, J.P.; Peille, C.N. Mechanical properties and short-term in-vivo evaluation of yttrium-oxide-partially-stabilized zirconia. *J. Biomed. Mater. Res.* **1989**, *23* (1), 45–61.
 23. Raigrodski, A.J. Contemporary materials and technologies for all-ceramic fixed partial dentures: A review of the literature. *J. Prosthet. Dent.* **2004**, *92* (6), 557–562.
 24. Luthardt, R.G.; Holzner, M.S.; Rudolph, H.; Herold, V.; Walter, M.H. CAD/CAM-machining effects on Y-TZP zirconia. *Dent. Mater.* **2004**, *20* (7), 655–662.
 25. Witkowski, S. (CAD-)/CAM in dental technology. *Quintessence Dent. Technol.* **2005**, *28*, 169–184.
 26. Haraguchi, K.; Sugano, N.; Nishii, T.; Miki, H.; Oka, K.; Yoshikawa, H. Phase transformation of a zirconia ceramic head after total hip arthroplasty. *J. Bone Joint Surg. Br.* **2001**, *83* (7), 996–1000.
 27. Blatz, M.B.; Sadan, A.; Martin, J.; Lang, B. In vitro evaluation of shear bond strengths of resin to densely-sintered high-purity zirconium-oxide ceramic after long-term storage and thermal cycling. *J. Prosthet. Dent.* **2004**, *91* (4), 356–362.
 28. McLaren, E.A.; White, S.N. Glass-infiltrated zirconia/alumina-based ceramic for crown and fixed partial dentures: Clinical and laboratory guidelines. *Quintessence Dent. Technol.* **2000**, *23*, 63–76.
 29. Guazzato, M.; Albakry, M.; Swain, M.V.; Ironside, J. Mechanical properties of In-Ceram alumina and In-Ceram zirconia. *Int. J. Prosthodont.* **2002**, *15* (4), 339–346.
 30. Sorensen, J.A. The Lava system for CAD/CAM production of high-strength precision fixed prosthodontics. *Quintessence Dent. Technol.* **2003**, *26*, 57–67.
 31. Suarez, M.J.; Lozano, J.F.; Paz Salido, M.; Martinez, F. Three-year clinical evaluation of In-Ceram zirconia posterior FPDs. *Int. J. Prosthodont.* **2004**, *17* (1), 35–38.
 32. Brodbeck, U. The ZiReal Post: A new ceramic implant abutment. *J. Esthet. Restor. Dent.* **2003**, *15* (1), 10–23; discussion, 24
 33. Kohal, R.J.; Klaus, G.; Strub, J.R. Zirconia-implant-supported all-ceramic crowns withstand long-term load: A pilot investigation. *Clin. Oral Implants Res.* **2006**, *17* (5), 565–571.
 34. Thompson, J.Y.; Anusavice, K.J.; Naman, A.; Morris, H.F. Fracture surface characterization of clinically failed all-ceramic crowns. *J. Dent. Res.* **1994**, *73* (12), 1824–1832.
 35. Kelly, J.R. Ceramics in restorative and prosthetic dentistry. *Ann. Rev. Mater. Sci.* **1997**, *27*, 443–468.
 36. Zhang, Y.; Pajares, A.; Lawn, B.R. Fatigue and damage tolerance of Y-TZP ceramics in layered biomechanical systems. *J. Biomed. Mater. Res. B Appl. Biomater.* **2004**, *71* (1), 166–171.
 37. Lange, F.F.; Dunlop, G.L.; Dais, B.I. Degradation during aging of transformation-toughened ZrO_2 - Y_2O_3 materials at 250°C. *J. Am. Ceram. Soc.* **1986**, *69*, 237–240.
 38. Kim, D.K.; Jung, Y.G.; Peterson, I.M.; Lawn, B.R. Cyclic fatigue of intrinsically brittle ceramics in contact with spheres. *Acta Mater.* **1999**, *47*, 4711–4725.
 39. Davis, J.B.; Cao, H.C.; Bao, G.; Evans, A.G. The fracture energy of interfaces: An elastic indentation technique. *Acta Metall.* **1991**, *39*, 1019–1024.
 40. Zhang, Y.; Song, J.K.; Lawn, B.R. Deep-penetrating conical cracks in brittle layers from hydraulic cyclic contact. *J. Biomed. Mater. Res. B Appl. Biomater.* **2005**, *73* (1), 186–193.



41. Lathabai, S.; Mai, Y.W.; Lawn, B.R. Cyclic fatigue behavior of an alumina ceramic with crack-resistance curves. *J. Am. Ceram. Soc.* **1989**, *72*, 1760–1763.
42. Qasim, T.; Ford, C.; Bush, M.B.; Hu, X.; Malament, K.A.; Lawn, B.R. Margin failures in brittle dome structures: Relevance to failure of dental crowns. *J. Biomed. Mater. Res. B Appl. Biomater.* **January 2006**, *80* (1), 78–85.
43. Deng, Y.; Lawn, B.R.; Lloyd, I.K. Characterization of damage modes in dental ceramic bilayer structures. *J. Biomed. Mater. Res.* **2002**, *63* (2), 137–145.
44. Suliman, A.H.; Swift, E.J., Jr.; Perdigao, J. Effects of surface treatment and bonding agents on bond strength of composite resin to porcelain. *J. Prosthet. Dent.* **1993**, *70* (2), 118–120.
45. Zhang, Y.; Lawn, B. Long-term strength of ceramics for biomedical applications. *J. Biomed. Mater. Res. B Appl. Biomater.* **2004**, *69* (2), 166–172.
46. Zhang, Y.; Lawn, B.R.; Rekow, E.D.; Thompson, V.P. Effect of sandblasting on the long-term performance of dental ceramics. *J. Biomed. Mater. Res. B Appl. Biomater.* **2004**, *71* (2), 381–386.
47. Blatz, M.B. Cementation of zirconium-oxide ceramic restorations. *Pract. Proced. Aesthet. Dent.* **2004**, *16* (1), 14.

Index

A

- Abdominal aortic aneurysm, endovascular stent grafts, 2480–2482
- Abdominal examination, Doppler ultrasound imaging, 2949
- Abdominal surgery, fibrin sealant applications, 1086
- Abiocoar implantable artificial heart, 1780–1781
- Ablation. *See also* Photoablation, light-tissue interactions
- photothermal laser-tissue interactions, 1627–1628
 - plasma-induced, 1630
- Absolute moments, mass transfer and tissue function, 1754–1755
- Absolute tolerance, microelectromechanical systems, 1863–1864
- Absorption. *See also* Adsorption
- biomedical optics
 - coefficient of, 2031
 - spectroscopy, 2034–2035
 - laser-tissue interaction, 1623–1624
 - protein-based drug delivery systems, 2359–2361
 - suture classification and configuration, 2563–2568, 2575–2577
 - ultrasound-tissue interaction, 2936–2939
- Absorption, distribution, metabolism, elimination, and toxicity (ADMET) mechanisms, toxicity testing, animal surrogate systems, 51–58
- Accountability, intellectual property rights, 1546
- Accunet (Guidant) distal protection device, 875–876
- Accuracy, microcomputed tomography, 1853–1854
- Acellular grafts
- nerve tissue engineering, 2815
 - valvular prosthetics, anticalcification treatments, 81
- Acetabulum, total hip arthroplasty, reaming techniques, 1377–1379
- Acid extraction, collagen scaffolds, 642
- Acid-soluble collagen, degradation, 635
- Acoustics
- laser-tissue interactions, 1629
- masking
- deaf persons, applications for, 2616–2617
 - history of, 2611–2612
 - tinnitus
 - designs, 2612–2613
 - FM masking, 2613
 - high audio/ultrasonic bone vibration, 2614–2615
 - history, 2611–2612
 - personal stereos, 2613–2614
 - ultrasound, 2933–2941
 - bone healing, 2953–2955
- Acousto-optic deflector (AOD), confocal microscopy, 711–712
- Acquired immunodeficiency syndrome (AIDS), gene therapy, 1155
- Acrylamide polymers, hydrogels, 1446
- Acrylic polymers, hydrogels, 1446
- Actin hydrogels, basic properties, 1447
- Action potentials
- electrogastrography, 947–949
 - electromyography, 956–959
- Activation studies, positron emission tomography, 2319
- Active fiberoptic biosensors, 1067–1068
- Active gradient screening, magnetic resonance imaging, 1717
- Activities of daily living (ADL), wheeled mobility systems design, 3087–3088
- Acute insulin delivery, 1516–1517
- Acute lung injury (ALI), lung surfactants, basic properties, 1683–1684
- Acute myocardial infarction (AMI), distal protection device reduction of, 878
- Acute renal failure (ARF). *See also* Hemodialysis
- artificial kidneys
 - biocompatibility, 1583
 - cellulosic membranes, 1583–1584
 - classification, 1580–1583
 - dialysis membranes, 1583–1585
 - efficiency, 1581–1583
 - flux changes, 1581
 - future research issues, 1592
 - historical perspective, 1578–1580
 - hollow-fiber membranes, 1579–1580, 1585
 - overview, 1576
 - synthetic membranes, 1584–1585
 - uremic solute classification, 1577–1578
- renal replacement therapies, 1590–1591
- Acute respiratory distress syndrome (ARDS)
- arteriovenous carbon dioxide removal, 1673–1675
 - artificial lung systems, overview and background, 1661–1662
 - lung surfactants
 - basic properties, 1683–1684
 - dysfunction physiology, 1688–1690
 - therapeutic regime, 1690–1691
- nitric oxide inhalation therapy, 1954
- Acuvue Advance™ silicone hydrogel contact lenses, 727
- Acyl azides, collagen tissue cross-linking, 646–647
- Acyl-homoserine lactone (acyl-HSL) systems, biofilm formation, 209–211
- ADAMs (a disintegrin and metalloproteinase), 1531–1532
- Adaptation mechanisms
- blood vessel mechanics and, 399–400
 - tissue engineering, skeletal muscle scaffolds, 2842–2843
- Adaptive filtering, electrogastrography signals, 951–954
- Adaptive immunity, gene therapy, 1158
- Adaptive Resonance Theory (ART) networks, 119–120
- Adductor muscles, groin pain syndrome, 1394–1395
- Adeno-associated virus (AAV)
- gene-activated matrix, 1145
 - gene therapy, 1154–1155
- Adenosine triphosphate (ATP)
- heart valve prosthetic failure, calcium binding site inactivation, 1308–1309
 - oxygen transport, 1842–1843
- Adenovirus
- gene-activated matrix, 1145
 - gene therapy, 1156–1157
- Adhesion and adhesives. *See also* Bioadhesion
- aldehyde-gelatin film, 1190
 - bioadhesion mechanisms, 1–2
 - biofilm formation, 207–215
 - bone cements and graft substitutes, 159
 - cell culture assays, 564
 - electrical cell impedance sensing, 910–913
 - cell surface analysis, substratum topography, 2556–2557
- ceramics
- dental ceramics, 603
 - transparent ceramics, 2918–2923
- coating surfaces, 2534
- cytoadhesives, 5–6
- defined, 1, 152
- dental adhesives and cements, 159
- fibrin glue and sealants, 1081–1088, 1187
- fibrin sealants, 1081–1088
- future research issues, 160
- hard tissue
- basic properties, 2
 - gene-activated matrix scaffolds, 1146
- inherently conductive polymers, 699–701
- mechanisms of, 1–2
- medical bioadhesives
- bone cements and graft substitutes, 159
 - dental adhesives and cements, 159
 - muscle glue protein, 159–160
 - skin adhesives, 158
 - tissue sealants, 158–159
- muscle glue protein, 159–160
- natural materials
- aquatic adhesives, 240–245
 - bond strengths, 237–240
 - overview, 236–240
 - terrestrial adhesives, 245–247
 - wound healing and tissue repair, 247–249
- neutrophils, 1935–1936
- bacterial killing, 1940
 - degranulation, 1936–1937, 1940
 - evaluation, 1938–1941
 - functional changes, 1937–1938
 - modification of, 1936
 - molecule expression, 1939
 - motility, 1940
 - oxidative burst, 1936, 1940
 - phagocytosis, 1940
 - viability, 1941
- pharmaceutical bioadhesives, 152–153
- administration routes, 157–158
 - bioadhesive bond, 153–155
 - formulation factors, 156–157
 - multifunctional, 155
 - polymers, 153–154
 - second-generation bioadhesives, 155–156
- platelets, blood-material interactions, 371–372
- skin adhesives, 158
- soft-tissue adhesives, 2–5, 3–5
- thrombosis
- neutrophil adhesion and fibrin formation, 2604–2606
 - normal erythrocytes, 2607–2608
 - platelet adhesion, 2601–2602
 - red-blood-cell adhesion, 2603–2604

- Adhesion and adhesives (*cont.*)
tissue sealants, 158–159
transdermal drug delivery, pressure-sensitive adhesives, 2896
- Adipocytes, tissue engineering
breast, 2680
fat tissue, 2727–2728
- Adipogenic inducers, fat tissue engineering, 2730
- Adipose stem cells (ASCs), tissue engineering applications, 2293–2294
- Adipose tissue. *See also* Fat tissue
origin and characteristics, 2726
tissue engineering
breast, 2682
fat tissue, 2725–2734
- Adjustability parameters, ergonomics principles, 1005–1006
- Administration routes, bioadhesives, 157–158
- ADSEP extractor, electrophoretic cell separation, 1700
- Adsorption. *See also* Absorption
dialysis membranes, 1588
plasmapheresis, 386–387
protein structures
multicomponent protein structures, 276–278
nonfouling surface resistance to, 278–280
single-component protein solutions, 272–276
submolecular mechanisms, 280–281
- Advanced processing algorithms, hearing aid design, 1268–1269
- Affine transform, real-time tomographic reflection calibration, 2393
- Affinity biosensors
development of, 341–343
long-term implantation, 1180
microparticle image velocimetry, 1880–1881
- Agar materials
biocompatibility testing, 171
electrical impedance tomography, 923–924
- Agarose
articular cartilage repair, scaffold materials, 556
hydrogels, 1447
- Age-related macular degeneration (AMD)
angiogenesis inhibitors, 28
laser-tissue interactions, 1626
ocular implants, controlled drug delivery, 1981–2
- Aggregation mechanisms
blood-material interactions, 1415–1416
proteoglycans, 1197–1198
- Aging
blood vessel mechanics and, 399–400
osteoclasts, bone regeneration and, 2095–2096
osteogenic progenitor cells, mesenchymal stem cells and, 2092
stress fractures and, 2511
tendon/ligament tensile testing, 2591–2592
- Albumin
glues, 1190–1192
gradient surfaces, protein-surface interactions, 1220–1222
hemocompatibility, 1346
- Alcohols, valvular prosthetics, anticalcification treatments, 80
- Aldehyde dehydrogenase (ALDH), hematopoietic stem cell progenitors, 1338–1339
- Aldehydes
collagen tissue fixation, 644–645
cross-linked hyaluronan biomaterials, 1426
surgical glues, 1189–1192
- Alginates
adhesives, 244
artificial pancreas, 2803
basic properties, 350
biosynthesis, 350–353
burn dressings, 459
future applications, 354–355
hydrogels, 1446–1447
medical applications, 353–354
nerve guides, 1917
nonmedical applications, 354
supercritical fluid synthesis, 2527–2528
tissue engineering applications, 2292
smart materials, 2294
wound healing and tissue repair, 248
- Algorithm of pattern extraction (ALOPEX)
basic principles, 123–125
error back propagation, 123
multilayered perceptron training, 125–126
very large scale integration implementation, 126–128
- Alignment stitch, heart valve prosthetic development of, 1295
failure analysis, 1312–1313
- Aliphatic polyurethanes, 2212
- Alizarin Red S, calcium staining, ceramic materials, 2918
- Alkaline phosphatase
ceramic activity staining, 2918
valvular prosthetics, calcification, 78
- Alkali-treated collagen, degradation, 635
- Alkanethiols, self-assembled monolayers, metal-alkanethiol interactions, 2409–2411
- Allografts
aortic valve cusp microstructure, 2742
biomaterials, 8–9
bone, tissue engineering, 2652–2653
burn injuries, skin substitutes, biological and synthetic, 453–456
collagen tissue, 641–644
donor sources, 14–15
fat tissue engineering, 2725–2734
hematopoietic stem cells, 1340–1342
historical background, 8–9
insulin delivery systems, 1518–1519
ligament injury and repair, animal repair models, 1643–1647
orthopedic soft tissue, 13–14
properties of, allo–10
skin, 10–11
tissue engineering, 2283
artificial pancreas, 2802–2809
breast, 2678–2679
nerve tissue, 2815
valvular prosthetics, 11–13, 1298
degeneration mechanisms, 2738
failure analysis, 1310–1312
vascular grafts, 2986
- Alloys. *See also* Metals
cardiac pacing leads, 535
ceramic materials, 589–590
hip joint prostheses, failure analysis, 1839–1841
membrane coatings, gold-palladium alloy permeability, 2354
- nickel
chemistry, 1942
hypersensitivity reactions, 1948
nickel-chromium, 1942–1943
nickel-titanium, 1945–1948
stainless steel, 1943–1945
- orthodontic wires
beta-titanium, 2057
cobalt-chromium alloys, 2057
nickel-titanium alloys, 2057–2058
overview, 2054–2056
stainless steel, 2056–2057
strengths and limitations, 2059
- orthopedic devices
cobalt-chrome alloys, 2062
fracture fixation and joint implantation, 1826–1827
hip joint prostheses, failure analysis, 1839–1841
nickel-titanium, 2063
tantalum, 2063
titanium, 2063
- stent devices, endovascular stent grafts, 2486–2487, 2494
- titanium
basic properties, 2876–2877
biocompatibility, 2879–2880
biomedical applications, 2877–2878
bone-implant interactions, 2880
cytotoxicity, 2879–2880
dental applications, 2887–2889
fatigue and fretting fatigue, 2881–2884
mechanical properties, 2881–2885
nickel-free alloys, 2879
surface modification, 2885–2887
wear characteristics, 2884–2885
total hip arthroplasty, 1476–1478
- Alpha chains, integrin structure, 1524–1525
- Alternating current electrokinetics, tissue engineering
basic principles, 2628–2629
biomaterial cell trapping, 2633–2634
cell properties characterization, 2631–2632
cell separation, 2632
cell sorting, 2632–2633
dielectrophoresis, 2629–2630
electro-orientation, 2629
electrorotation, 2630
electrostriction, 2629
traveling wave dielectrophoresis, 2630–2631
- Altruism, biofilm stability, 213
- Alumina
applications, 19–23
bone spacers, 21–22
dental materials, 22
joint articulation, surface replacement, 19–21
prosthetics, 22–23
- Bayer extraction and purification process, 17
bone-implant interface, 425
ceramics, 589
chemical structure, 17
compatibility, 19
dental implants, 22, 597–599
future research, 23
matrix composites, 23
physical properties, 17–18
wear properties, 18–19, 3061–3063
- Aluminum oxides, ceramic implants, 588–589
- Alveolar-capillary membrane, artificial lung systems, 1661–1662
- Alveolar function, mass transfer and, 1755–1756
Fick's laws and time constants, 1757–1759
- Alveolar surfactants, structure and properties, 1684
- Alzheimer's disease
functional magnetic resonance imaging, 1125–1126
magnetic resonance imaging, 1722
- Ambulatory implantable left ventricular assist devices, 1777–1778
- American Association for the Accreditation of Laboratory Animal Care (AAALAC), 1619–1620
- Ames Salmonella test, 564
- Amine-gelatin glues, 1189–1190
- Amino-oleic acid, valvular prosthetics, anticalcification treatments, 80
- Amniotic membrane, burn injuries and grafting with, 455–456

- Amorphous polymers, piezoelectric materials, 2169–2170
- Amplatz coronary catheters, 500–501
- Amplifiers
- biopotential applications
 - basic requirements, 302–303
 - design criteria, 303–304
 - electrocardiography, 301–302, 306–307
 - electroencephalography, 302, 307
 - electromyography, 302, 307–308
 - electrooculogram, 302, 308
 - electroretinogram, 302, 308
 - interference problems, 304–306
 - overview, 301
 - electrocardiography, 942–943
 - electromyography signal conditioning, 963
 - flow cytometry systems, 1103
 - hearing aid design, 1267–1268
 - sensing amplifiers, implantable
 - cardioverter-defibrillators, 776
- Amyloid plaques, magnetic resonance microscopy, 1742–1744
- Anabolic steroids, microcomputed tomography research, 1855
- Analog-to-digital (A/D) conversion
- digital signal acquisition and processing, 843–845
 - speech processors, cochlear implants, 1509–1510
- Analytes, fiberoptic biosensors, 1064
- Anchorage-dependent survival, integrin signaling, 1533
- Aneurysms
- aortic, macrocirculatory hemodynamics, 1366–1367
 - endovascular stent-grafts, 2480–2483
- Angina, transmyocardial laser revascularization, 1632–1637
- Angiogenesis
- bone healing and, 1260
 - glucose sensor implants, 1176–1177
 - hyaluronan materials, 1423
 - inhibitors
 - acquired resistance to, 29–30
 - age-related macular degeneration, 28
 - anticoagulant therapy, 29
 - anti-vascular endothelial growth factor, 27
 - biotechnology strategies, 27–29
 - chemotherapy vs., 30–31
 - collagen tumor growth inhibition, 27
 - development strategies, 27
 - diagnostic imaging, 31
 - historical background, 26–27
 - kininostatin, 27
 - matrix metalloproteinases, 27–28
 - 2-methoxyestradiol, 28
 - multiple myeloma therapy, 28–29
 - overview, 26
 - research issues, 29–31
 - small molecule integrin antagonists, 28
 - thrombosis, 31
 - microcomputed tomography, 1855–1856
 - tissue engineering, liver, 2774–2775
 - transmyocardial laser revascularization, 1632–1637
- Angiography
- catheterization techniques, 492–493
 - intravascular ultrasound, 1561–1564
 - magnetic resonance imaging, 1723–1724
 - three-dimensional imaging, 1797
 - vascular imaging, 3012–3013
 - future research issues, 3018
- Angioguard (Cordis) distal protection devices, 873–874
- Angioplasty procedures, distal protection devices
- Accunet device, 875–876
 - advantages/disadvantages, 877
- Angioguard device, 873–874
- carotid artery, 870–871
- current research evolution, 877–878
- embolization reduction devices, 871
- failure mechanisms, 877
- Filterwire Ex, 875
- future applications, 878
- NeuroShield device, 874–875
- overview, 870
- Parodi Anti-Emboli System, 876–877
- Percutaneous guardwire, 871–873
- saphenous vein grafts, 870–871
- Triactiv (Kensey Nash), 873
- Angioscopy, vascular imaging, 3017–3018
- Angular position, small-angle light scattering, soft tissue analysis, 2454–2456
- Animal models
- arteriovenous carbon dioxide removal, 1674–1675
 - bio-artificial liver systems, 1650–1658
 - bone tissue engineering
 - basic model properties, 33
 - critical-sized defects and bony wound healing, 39–40
 - generic animal models, 34
 - gross anatomy, function and growth, 36–39
 - human clinical conditions and, 40–45
 - microarchitecture and biomechanics, 35–36
 - model appropriateness, 33–34
 - phylogenetically closer animal models, 34–35
 - cardiovascular system
 - degradation and tissue reaction, rat model, 1569
 - functional assessment, 1567
 - morphological assessment, 1567–1568
 - thrombogenicity, acute pig model, 1568–1569
 - vascular graft replacement, 1569–1571
 - in vivo models, 1566–1567
 - ceramic materials adhesion, 2918–2923
 - ethical issues
 - AAALAC international accreditation, 1619–1620
 - Animal Welfare Act, 1617–1618
 - Endangered Species Act, 1621
 - future developments, 1621
 - Good Laboratory Practices Act, 1621
 - Guide for Care and Use of Laboratory Animals, 1620–1621
 - Health Research Extension Act, 1618–1619
 - overview, 1616–1617
 - regulations, 1621
 - genetic engineering, organ transplants, 2910–2912
 - hearing mechanisms, 1271–1272, 1277–1278
 - in-vivo modeling
 - degradation and tissue reaction, rat model, 1569
 - functional assessment, 1567
 - implant site selection, 172–173
 - morphological assessment, 1567–1568
 - overview, 1566
 - techniques, 173–175
 - thrombogenicity, acute pig model, 1568–1569
 - vascular graft replacement, 1569–1570
 - in vivo models, 1566–1567
 - welfare issues, 172–173
 - ligament repair
 - anatomy, structure, and function, 1640–1641
 - biomechanics, 1645
 - dogs, 1644
 - healing responses, 1644–1645
 - histology, 1645–1646
- injury and healing models, 1641–1644
- joint laxity and osteoarthritis, 1643
- overview, 1640
- rabbits, 1644
- rats, 1644
- reconstruction, 1643–1644
- rupture/transection, 1641–1643
- sheep and goats, 1644
- vascular physiology, 1646–1647
- in vitro models, 1644
- lung surfactant dysfunction, 1688–1689
- prenatal bladder disease, tissue engineering, 2640
- rotary cell culture systems, insect cells, 2406
- spinal deformity biomechanics, 2470–2471
- tendons and ligaments
 - tensile testing, 2590–2592
 - viscoelasticity, 2596–2597
- tissue engineering
- bladder, 2636–2640
 - breast, 2683
 - cardiac muscle, 2686
 - esophagus, 2720
 - polymers, 2289–2293
 - skin, 2852
- toxicity testing
- barrier models, 53–55
 - cell culture analog systems, 57–58
 - chemical structure-related models, 51–52
 - coculture models, 55
 - computational models, 51–53
 - conditioned medium systems, 56–57
 - microscale designs, 57–58
 - multiple cell devices, 56–57
 - overview, 51
 - shared medium systems, 57
 - single cell in vitro models, 53
 - three-dimensional models, 55–56
 - whole organism models, 52–53
- transgenic/gene-targeted organs, 2906–2912
- transmyocardial laser revascularization, 1632–1637
- vitrification technology, 3051–3054
- wear debris-induced osteolysis, 3070–3072
- Animal Welfare Act, 1617–1618
- Anionic substitution, hydroxyapatite, 1452
- Anisometric contractions, electromyographic signals, 966
- Anisotropy
- articular cartilage biomechanics, 87–91
 - atrioventricular valve mechanics, 1324–1326
 - biomedical optics, 2031
 - composite continuum mechanics, 661–666
 - hard tissue elasticity and, 1241–1242
 - ligaments and tendons, multiaxial mechanical testing, 2592–2593
 - magnetic resonance imaging, 1722
- Ankle ligaments, tissue engineering, injuries and reconstruction, 2758–2759
- Ankle prostheses, alumina ceramics, 21
- Annular fiber bundles, artificial lungs, 1679–1680
- Anodic reaction
- corrosion electrochemistry, 754–757
 - electrical stimulation waveform, 1015
- Anterior cervical spine plate, failure analysis, 1832–1833
- Anterior cruciate ligament (ACL)
- anatomy and function, 1608–1609
 - animal models of repair
 - anatomy, structure, and function, 1640–1641
 - biomechanics, 1645
 - dogs, 1644
 - healing responses, 1644–1645
 - histology, 1645–1646
 - injury and healing models, 1641–1644
 - joint laxity and osteoarthritis, 1643
 - overview, 1640

- Anterior cruciate ligament (ACL) (*cont.*)
 rabbits, 1644
 rats, 1644
 reconstruction, 1643–1644
 rupture/transsection, 1641–1643
 sheep and goats, 1644
 vascular physiology, 1646–1647
 in vitro models, 1644
 bioreactor applications, 311–312
 loading properties, 2762
 prosthetic materials design, 1609–1610
 tissue engineering
 immune response, 2765–2766
 injuries and reconstruction, 2758
- Anterior knee pain, overuse injury, 1598–1599
- Anthropometry, ergonomics principles, 1005–1006
- Antiadhesive materials
 hyaluronan, 1427
 polyepoxy compounds, 2238
- Antibiotic/antimicrobial materials
 antiseptic/antibiotic coatings, 64–66
 biocatalysis and development of, 162
 isolated enzyme techniques, 165–167
 whole-cell biocatalysis, 163–165
 biomaterials-related infections, 61–63
 burn dressings, 456–457
 chitosan sustained release systems, 604–605
 coating surface properties, 2534–2535
 medical device infections, 1497–1504
 metals, basic properties, 66–67
 overview, 61
 phosphorylcholine technology, 2145–2146
 polyurethane biomaterials, 2310–2312
 recent advances, future research, 67
 surface modification, bacterial adherence prevention, 63–64
- Antibodies, biosensors and, 340–343
- Anticalcification treatments, valvular prosthetics, 79–81, 1295–1301
- Anti-cell adhesion, coating surface properties, 2534
- Anticoagulants
 angiogenesis and, 29
 vascular endothelium, heparin biocompatibility, 1368–1369
- Antiencrustation coatings, surface properties, 2533
- Antigen-presenting cells (APCs)
 host reactions, 1417
 immune response, 263–264
- Anti-inflammatory compounds
 chondroitin and dermatan sulfates, 618–626
 polymeric structure, 221
- Antimineralization. *See* Biomineralization
- Antiplatelet agents, hemocompatibility, 1346–1348
- Antisense oligonucleotides (ODNs), cardiac patch engineering, 547
- Antisthenotic agents, controlled-release systems, 887
- Anti-thrombogenicity, tissue engineering, blood vessels, 2649–2650
- Anti-vascular endothelial growth factor, 26–27
- Aortic aneurysms, macrocirculatory hemodynamics, 1366–1367
- Aortic valve
 endovascular stent grafts, 2482
 leaflets, mechanical properties, 1319–1320
 stenosis
 mechanical heart valves, blood-material interactions, 1336
 valvular prosthetics, 1293
 structure and function, 1317–1319
 valvular prosthetics
 leaflets, 1295
 microstructure/micromechanics, 2738–2742
- Apaceram[®] hydroxyapatite, 463
- Apatite. *See also* Hydroxyapatite; Hydroxyapatite (HA)
 biologic apatites, 226–229
 biphasic calcium phosphate, 359–364
 historical background, 225–226
 orthopedic devices, ceramics and glasses, 2067
 overview, 225
 synthetic apatites, 229–233
- Apheresis technology, blood purification, 391
- Apoptosis. *See* Cell death
- Apparent diffusion coefficient (ADC), magnetic resonance imaging, 1722
- Application-specific integrated circuits (ASICs), electrotransport mechanisms, transdermal drug delivery, 2900
- Applied loading, finite element analysis, 1094
- Approximate entropy, digital biomedical signal acquisition and processing, 865
- Aquaculture, alginate applications in, 354
- Aquatic adhesives, 240–245
 algal adhesives, 244
 bacterial adhesives, 243–244
 barnacle cement, 241–242
 fungal adhesives, 244
 limpet and periwinkle mucus, 243
 mussel glue, 240–241
 sea star and sea cucumber, 242–243
 stickleback spiggin, 243
- Arc lamps, optical mapping, 2020
- Areolar connective tissue, healing properties, 1253
- Arg-Gly-Asp (RGD) compounds
 adhesive properties, 218–219
 cell-material interactions, 570–571
 graft polymers, tissue engineering applications, 2293–2294
 self-assembled monolayers, 2410–2411
- Aromatic amines
 bone cements, 410
 polyamides/aramides, 2192–2193
- Aromatic polyurethanes, 2212
- Array technologies, biosensors, 341–343
- Arteries and arterioles. *See also* Blood vessels
 grafts, gene-activated matrix regeneration, 1147–1148
 mechanical properties, 397–399
 microcirculatory oxygen transport, 1845–1846
- stents
 diamond/diamond-like carbons, 828–829
 endovascular stent-grafts, 2480–2483
 suprarenal stents, 2488
 superficial femoral artery, distal protection device intervention, 878
 vascular fluid dynamics, 2972–2981
 electrical analogues, 2975
 inviscid analysis, 2972–2974
 viscous analysis, 2974–2975
 vessel wall properties
 arterial pulse waves, 1364–1365
 hemodynamics, 1358–1359
- Arteriovenous carbon dioxide removal (AVCO₂R), 1673–1675
- Arthritis. *See also* Osteoarthritis
 bioreactor applications, 311–312
 hip joint biomechanics and, 1377
- Arthroplasty procedures, wear debris, 3076–3079
- Articular cartilage. *See also* Cartilage
 allograft materials and techniques, 13–14
 alumina joint prostheses, 21
 anatomy and physiology, 552–554
 biomechanics
 basic principles, 85–91
 composition and structure, 85–86
 mathematical models, 91–93
 overview, 85
- bioreactor applications, 311–312
 finite element analysis, 1098
 hyaluronan biomaterials, 1427
 proteoglycans
 degradation and aging, 1200–1201
 structure and properties, 1196–1197
 regeneration and repair
 applied tissue engineering, 554–555
 cell constructs, 556–557
 chemical factors, 557–558
 healing potential, 1254
 in-vitro procedures, 558–559
 overview, 552
 repair response, 554
 scaffold selection, 555–556
 surface modification, 557
 synthetic polymers, 556
 tissue engineering
 bioengineering principles, 2701
 biphasic construct substrates, 2700–2701
 cell cultures, 2697–2698
 cell signaling mechanisms, 2697
 cell sources, 2701–2702
 chondrocyte subpopulations, 2703
 composition, 2694–2695
 differentiated chondrocytes, 2701–2702
 function, 2696–2697
 growth factors, 2702
 mechanical stimulation, 2702
 overview, 2694
 pre/post-transplantation assessment, 2701
 quality control and improvement, 2702–2703
 scaffold materials, 2698–2700
 stem cells, 2702
 zonal organization, 2695–2696
- Artifacts, biopotential amplifiers, 305–306
- Artificial bladder, tissue engineering, 2639–2640
- Artificial heart
 Abiocor implantable artificial heart, 1780–1781
 battery system, 1493
 centrifugal blood pump systems, 581–583
 positive displacement pumps, fluid dynamics, 95–105
 total artificial heart, structure and classification, 475–477
- Artificial joints, disc arthroplasty, 2468
- Artificial kidneys
 biocompatibility, 1583
 cellulosic membranes, 1583–1584
 classification, 1580–1583
 dialysis membranes, 1583–1585
 efficiency, 1581–1583
 flux changes, 1581
 future research issues, 1592
 historical perspective, 1578–1579
 hollow-fiber membranes, 1579–1580, 1585
 overview, 1576
 synthetic membranes, 1584–1585
 uremic solute classification, 1577–1578
- Artificial liver. *See* Bio-artificial liver systems
- Artificial lungs
 annular fiber bundles, 1679
 arteriovenous carbon dioxide removal, 1673–1675
 blood oxygenator, 1666–1667
 cardiopulmonary bypass oxygenators, 1667–1669
 CORx[™], 1679
 current research and future directions, 1673
 diffusional boundary layers, 1665
 extracorporeal membrane oxygenation, 1669–1670
 gas exchange determinants, 1663
 historical background, 1661–1662
 integrated heart-lung assisted device, 1679

- intravascular respiratory catheters, 1675–1677
- mass transfer correlations, 1665–1666
- membrane permeance, 1663–1665
- next-generation technologies, 1680
- operating principles, 1662–1666
- overview, 1661
- photolytic artificial lung, 1679–1680
- total artificial lungs, 1677–1678
- Artificial membranes, interfacial transport
- pH effect on permeability, 2351–2352
- pore size and permeability, 2353–2354
- protein permeability, 2350–2351
- Artificial muscles
- conductive polymers, 110
- dielectric elastomer electroactive polymers, 109
- electroactive polymeric materials
- conducting polymers, 926–929
- drug delivery systems, 932–933
- electro-responsive gels, 932
- ferroelectric properties, 108–109
- nonionic polymer gels, 109
- overview, 926
- piezoelectric polymers, 929–932
- polyacrylonitrile fibers, 110–111
- future research issues, 115
- ionic polymer-metal composites, 109–110, 112–113
- liquid crystal elastomer materials, 109, 113–114
- magnetically activated gels, 108
- overview, 108
- shape memory alloys and polymers, 108, 110
- tissue engineering, 2838–2844
- Artificial nerve grafts (ANG), 2127–2130
- Artificial neural networks (ANNs)
- clinical applications, 119
- neurological process modeling, 130
- neuromorphic models, 129–130
- overview, 119
- Artificial organs, interfacial transport and, 2348–2349
- Artificial oxygen carriers, microcirculatory oxygen transport, 1847
- Artificial pancreas
- insulin delivery, 1519–1520
- interstitial fluid sensors, 1175–1178
- tissue engineering, 2802–2809
- Artificial proteins, dendrimer-based drug delivery, 794–795
- Artificial silks, design and synthesis, 2446–2447
- Artificial tendons, fibroblast mechano-regulation, 1788–1789
- Artificial urinary sphincter, basic properties, 2966–2968
- Aseptic loosening phenomenon, host reactions, particulate matter, 2114–2115
- Aspartic acid target sequence, integrin-binding, 1527–1528
- Assembly process, cardiac catheters, 495
- Assessment techniques, ergonomics, 1002–1004
- Assistive devices
- robots, 1804–1805
- wheeled mobility, 3083–3092
- Astrocytes, structure and function, nanoscale neurosurgery and, 1905–1906
- Atherosclerosis. *See also* Cardiovascular disease
- macrocirculatory hemodynamics, 1366–1367
- plaque analysis, magnetic resonance microscopy, 1746–1747
- vascular prostheses, in vivo animal models, 1570
- Athletic injuries. *See also* Overuse injuries
- hip joint
- external rotator syndrome, 1397
- gluteus medius syndrome, 1397
- groin pain syndrome, 1393–1395
- hamstring syndrome, 1399–1401
- iliopsoas tendinitis/bursitis, 1395–1396
- overview, 1393
- rectus femoris tendinitis, 1396–1397
- snapping hip syndrome, 1397–1399
- knee joint
- breaststroker's knee, 1602
- epidemiology, 1597/1598
- Fabella syndrome (fabellitis), 1605
- fat pad syndrome (Hoffa disease), 1603–1604
- future research issues, 1605
- iliotibial band friction syndrome, 1600–1602
- Osgood-Schlatter disease, 1603
- patellar tendinitis/tendinosis, 1599–1600
- patellofemoral pain syndrome, 1597–1599
- pes anserinus tendinitis and bursitis, 1604–1605
- plica syndrome, 1604
- popliteal tendinitis, 1605
- semimembranosus tendinitis, 1604
- Sinding-Larsen-Johansson disease, 1603
- stress fractures, 2512–2513
- Atmospheric discharge, plasma biomaterials modification, 2175
- Atomic force microscopy (AFM)
- bone mineral composition, 2042
- bone structure and composition, 1236
- gradient surfaces characterization, 1218–1219
- hemocompatibility assessment, 1349
- nanoscale neurosurgery, tumor cell invasion and axonal extension, 1908–1909
- Atom-transfer radical polymerization (ATRP), 2279–2280
- Atrial depolarization, electrocardiography, 938
- Atrial fibrillation
- electrocardiography, 944–945
- electrophysiology, 480–481, 488
- mechanisms of, 772–773
- Atrial repolarization, electrocardiography, 939
- Atrioventricular (AV) block, cardiac pacemakers, 540
- Atrioventricular interval (AVI)
- cardiac pacemakers, 537–538
- electrocardiography, 938
- Atrioventricular node (AVN)
- cardiac conduction system, 532–533
- electrocardiography, heart block, 944
- electrophysiology, 481–482, 488
- optical mapping, 2024–2025
- Atrioventricular valves
- hemodynamics, 1323
- mechanics, 1323–1326
- structure and function, 1320–1323
- Attachment techniques, bioreactor design, 322–324
- Attentional mechanisms, eye movement research, 1034–1038
- Attenuated total reflection (ATR), gradient surface analysis, 1218–1219
- Attenuated total reflection infrared (ATR-IR) spectroscopy, hemocompatibility assessment, 1349
- Attenuation
- positron emission tomography, 2318
- ultrasound-tissue interaction, 2936–2939
- Audibility correction, hearing mechanisms, 1273–1275
- Augmented reality, real-time tomographic reflection, 2391
- Autocorrelation function, biomedical signal acquisition and processing, 849–850
- spectral analysis, 858–861
- Autofluorescence, endoscopic techniques, 994–995
- Autogenous fibro-collagenous tubes, hybrid vascular prostheses, 1434–1435
- Autologous grafts
- articular cartilage repair, 554
- burn injuries, 456
- defined, 9
- ligament injury and repair, animal repair models, 1643–1647
- tissue engineering, 2283
- bone, 2652–2653
- breast tissue, 2679
- fat tissue, 2725–2734
- nerve tissue, 2815
- skin, 2852–2854
- tendons, 2872–2873
- vascular grafts, 2985–2986
- in vivo animal models, 1567
- Automated endoscopy system for optimal placement (AESOP), 689
- Automatic external defibrillators (AEDs), basic properties, 774
- Automatic gain control (AGC)
- cochlear implants, 1508–1510
- hearing aid design, 1267–1268
- Automatic speech recognition (ASR) systems, development of, 1278
- Autonomic nervous system (ANS), disorders of, pacing therapy, 540
- Autoregressive-moving-average model (ARMA), biomedical signal acquisition and processing, 849–851
- Autoregressive time-variant models, digital biomedical signal acquisition and processing, 859–861
- Auxiliary partial liver transplantation (APLT), bio-artificial liver systems, 1651
- Averaging repetitive EEG, basic principles, 307
- Axial pumps, continuous flow ventricular assist device, 474–475
- Axis deviation (cardiac), electrocardiography, 943–944
- Axonal degeneration, nerve tissue engineering, 2812–2813
- Axonal space constant, electrical stimulation, 1014
- Axonotmesis, mechanisms of, 1912–1913
- Axons
- nanoscale neurosurgery, axonal extensions, 1908–1909
- structure and function, nanoscale neurosurgery and, 1903–1906
- Azeotropic dehydration, poly(lactide-co-glycolide), 2259–2261
- Azide chitosan lactose, glue materials, 1192
- B**
- Back pain and dysfunction, work-related risk factors, 1000–1001
- Backscattering coefficient, ultrasound-tissue interaction, 2938–2939
- Bacteria. *See also* Antibiotic/antimicrobial materials
- adhesive properties, 6
- alginate biosynthesis and, 351–355
- aquatic adhesives, 243–244
- wound healing and tissue repair, 248
- biofilm formation and, 207–208, 213–215
- host reactions, 1414–1415

- Bacteria** (*cont.*)
 medical device infections, 1498–1499
 non-antibiotic approaches, 1504
 neutrophil-biomaterial adhesion, 1940
 supercritical fluid inactivation, 2525
- Balanus* spp., adhesive properties, 241–242
- Balloon-expanded stents, endovascular stent grafts, 2486
- Balloon occlusion devices, distal embolization reduction, 871, 877
- B/A parameter, ultrasound-tissue interaction, 2939–2941
- Barbs, endovascular stent graft fixation, 2487
- Barium titanate, piezoelectric ceramics, 2166
- Barnacle cement, 241–242
- Barrett's esophagus, tissue engineering, 2719–2720
- Barrier models, toxicity testing, animal surrogate systems, 53–55
- Basal insulin replacement, insulin delivery systems, 1516–1517
- Basic fibroblast growth factor (bFGF)
 angiogenesis inhibitors, 26
 polymeric structure, 221
- Basic multicellular units (BMUs), bone remodeling, 439–440
 canalicular fluid flow, 442–443
 strain-related remodeling, 440–441
- Basic Rate Interface (BRI), real-time telemedicine systems, 2580–2581
- Batch processing, microelectromechanical systems, 1863
- Batteries**
 cardiac pacemakers, 535–536
 centrifugal blood pumps, 582–583
 implant power sources, 1485–1493
 commercial applications, 1493
 electrochemical power generation, 1485–1486
 future development issues, 1493–1495
 lithium/carbon monofluoride, 1492
 lithium/CF/SVO hybrid cathode, 1492
 lithium-iodine, 1491
 lithium-ion rechargeable, 1492–1493
 lithium/silver vanadium oxide, 1492
 lithium/thionyl chloride, 1491–1492
 mechanical design, 1489–1490
 power, capacity, and longevity, 1487
 primary vs. secondary, 1487–1489
 requirements, 1489
 voltage and current thermodynamics and kinetics, 1486–1487
- Bayer extraction process, alumina, 17
- Baylor Gyro centrifugal blood pumps, 576–583
- Baylor-Nikkiso centrifugal blood pump, 576–577
- B cells, host reactions, 1416–1418
- Beamhouse process
 tanning industry, collagen processing, 654
- Beer-Lambert law
 absorption coefficient, biomedical optics, 2031
 radiative transfer, optical detection of cancer, 2010–2011
- Beer's law, laser-tissue interaction, 1624
- Behind-the-ear (BET) hearing aid, 1266–1267
- Belmont Report, The*, biomedical ethics and, 286–293
- Benzoyl peroxide (BPO), bone cement, 403–410
- Bergman minimal model, insulin delivery, 1520
- Bernoulli equation, Doppler ultrasound imaging, valvular disease, 2948–2949
- Bessel function, macrocirculatory hemodynamics, pulsatile flow properties, 1361–1364
- β -Titanium, orthodontic wires, 2057
- β -Tricalcium phosphate. *See* Tricalcium phosphate (TCP)
- Beta chains, integrin structure, 1524–1525
- Biaxial testing, blood vessel mechanics, 394–395
- Bifurcated stent-grafts, endovascular stent grafts, 2482
- Biglycan, structural properties, 1199
- Bile acids, surfactant properties, 2337
- Bileaflet valve, structure and components, 1329–1330
- Bingham plastic, vascular fluid dynamics, 2978
- Bioabsorbable materials**
 biocomposites, 182–183
 melt spun fibers, 1822–1823
 nerve guides, 1915–1916
 orthopedic devices
 absorbable fixation, 2073
 ceramics, 2067
 composites, 2069
 plates and screws, 2073–2078
 plates and screws
 bone graft fixation, 436
 cancer surgery, 436
 clinical applications, 434–437
 experimental techniques, 434
 future research issues, 437
 orthognathic surgery, 434–435
 orthopedic devices, 2073–2078
 overview, 432–434
 pediatric surgery, 436–437
 trauma surgery, 435–436
 sutures, 2570–2574
 vascular graft development, 2998–2999
 zygomatic and midface fractures, 2398–2399
- Bioactive materials**
 bone-implant interface, 425–426
 ceramics, orthopedic devices, 2066–2068
 glasses
 bone cements and grafts, 159
 clinical applications, 138
 composition and reactivity, 135–136
 future research issues, 139–140
 mechanical properties, 137–138
 mechanisms in, 134–135
 melt spinning, 1818–1821
 microporous materials
 ceramics, 1886–1887
 composites, 1890–1891
 foaming of processes, 1888–1891
 melt-derived glasses, 1890
 overview, 1885
 scaffold applications, 1885–1886
 sol-gel glasses, 1887–1888
 tissue engineering applications, 1885
 orthopedic devices, 2066–2068
 osteoinductive substances and materials, 2104
 overview, 133
 soft tissue bonding, 136–137
 synthesis and structure, 133–134
 tissue attachment and interfacial strength, 136
 tissue engineering, 138–139, 144–150, 2916–2923
 vitrification, 3046–3054
- tissue engineering
 calcium phosphates, 143
 composites, 144–145, 149
 fabrication, 145–150
 foam coatings, 149–150
 glass-ceramics, 144–150
 microsphere sintering, 149
 overview, 142
 powder-based fabrication, 147
 scaffold structures, 142
 silicate glasses, 143–144
 sol-gel fabrication, 147–148
 solid free-form fabrication, 148–149
 solvent casting/particle leaching, 149
 thermally induced phase separation, 149
- Bio-artificial liver (BAL) systems**
 bioreactors, 1656–1657
 cell-cell interactions, 1654–1655
 cell source and therapeutic cell mass, 1651–1652
 clinical setting, 1649–1651
 extracellular matrix and cell polarity, 1653–1654
 future research issues, 1658
 hepatocyte transplantation, 1650
 hollow-fiber systems, 1656–1657
 overview, 1649
 parallel plate systems, 1657
 performance monitoring, 1657–1658
 plasma/blood effects, 1656
 porous three-dimensional matrices, 1655
 scale-up hepatocyte culture, 1656
 soluble factors, 1655–1656
 surface chemistry optimization, 1653
 temporary support system, 1651
 tissue engineering, 1652–1656
 xeno (cross-species transplantation), 1650
- Biocatalysis**
 historical background, 162–163
 isolated enzymes, 165–167
 overview, 162
 poly(lactide-*co*-glycolide) synthesis, 2260–2261
 whole-cell techniques, 163–165
- Biocompatibility**
 alumina, 19
 bone fracture fixation materials, 416
 cell culture assays, 562–563
 in-vitro testing
 future research, 566–567
 limitations, 566
 overview, 561–562
 requirements for, 565–566
 standards, 563
 defined, 169–170
 degradable polymer composites, 780–781
 dialysis membranes, 1583
 diamond/diamond-like carbons, 828
 elastomers, 904–905
 foreign body response, 1111–1116
 gradient surfaces, 1219–1225
 hard tissue-biomaterial interactions, 1229–1232
 hemocompatibility
 active materials, 1346–1348
 inert materials, 1345–1346
 overview, 1345
 platelet interaction, 1353
 protein adsorption, 1350–1353
 surface modification strategies, 1345–1348
 surface properties analysis, 1348–1350
- heparin
 Bioline coating, 1372
 Carmeda Bioactive Surface, 1371
 catalytic role, 1368
 clinical applications, 1372
 complement and inflammatory response, 1370
 contact activation, 1370
 Corline heparin surface, 1371
 Duraflo surface, 1371–1372
 enzymatic coagulation interaction, 1370
 future research, 1373
 Hepoface coating, 1372
 immobilized biological effects, 1369–1370
 overview, 1368
 Pharmacia HSM introcular lenses, 1371
 platelet adhesion/activation, 1369
 protein interactions, 1370

- T-NCVC coating, 1372
 Trillium biopassive surface, 1371
 vascular endothelial anticoagulant activity, 1368–1369
 white blood cell adhesion/activation, 1369
- host reactions, 1414–1415
 hydroxyapatite, 1459–1460
 nerve graft materials, 2127–2130
 photoinitiators, 2152–2153
 poly(carbonate)urethanes, 2211–2212
 poly(glycerol sebacate), 331–333
 poly(lactide-co-glycolide), 2265–2266
 polyurethane biomaterials, 2306
 antimicrobial/antibacterial compounds, 2310–2312
- pyrolytic carbon, 2375
 silicone purification, 2438–2439
 suture materials, 2574–2575
 synthetic apatite composites, 232–233
 testing
 future research issues, 176–177
 implant clinical trials, 175–176
 in-vitro testing, 170–172
 in-vivo testing, 172–175
 tissue engineering, microvascular networks, 2796
- titanium and titanium alloys, 2879–2880
 transdermal drug delivery, 2902–2904
 vascular grafts, 2990–2991
- Biodegradable materials.** *See also*
 Degeneration and degradation
- bioabsorbable plates and screws
 bone graft fixation, 436
 cancer surgery, 436
 clinical applications, 434–437
 experimental techniques, 434
 future research issues, 437
 orthognathic surgery, 434–435
 overview, 432–434
 pediatric surgery, 436–437
 trauma surgery, 435–436
- biphasic calcium phosphate ceramics, 363–364
- collagen tissue, 644
- elastomers
 future research issues, 905–906
 overview, 897
 polyesters, 897–901
 biodegradation mechanisms, 900–901
 cross-linked synthesis, 898
 linear synthesis, 898–899
 mechanical properties, 900
 polyglycolide, polylactide, polycaprolactone copolymers, 897–900
 polyhydroxyalkanoate (PHA) copolymers, 899
- gene-activated matrix scaffolds, 1145–1146
- hydrogels, 1447–1448
 polycaprolactone, 2198–2200
 poly(glycolic acid), 2242–2244
 poly(lactic acid), 2252–2254
- polymers
 basic properties, 195–196
 biorubber/poly(glycerol sebacate), 329–331
 composites, 779–780
 controlled release drug delivery systems, 739
 degradation properties, 201–204
 free radicals, 203–204
 glycolide, lactide, and caprolactone-based compounds, 196–199
 linear aliphatic polymers, 196
 natural materials, 199–200
 orthopedic devices, 2063–2065
 poly(ester amide)s, 200–201
- polyurethanes, 901–905, 2306–2309
 biocompatibility, 904–905
 biodegradation mechanisms, 904
 mechanical properties, 903–904
 processing and applications, 905
 synthesis, 901–904
- stents, 2497
 sutures, 2575–2577
- tissue engineering
 microvascular networks, 2796
 vascular system, 3034
- Bioelectrical signals, amplifiers of, 307–308**
- Bioengineering**
 alginates, 354–355
 elastin materials, 894–895
 finite element analysis, 1097–1098
- Biofilms**
 altruism and stability of, 213
 basic properties, 207
 biomaterials-related infections, 61–63
 burn dressings and, 458
 dental ceramics and, 593–594
 formation mechanisms, 207–213
 colonization and maturation, 209–211
 detachment and dissolution, 212–213
 early-attachment events, 207–208
 extracellular polysaccharides, 211
 gene expression, 211–212
 historical background, 207
 medical device infections, 1498–1500
 pathogens and, 213–215
- Biofouling**
 bacterial adhesion, 6
 hydrogels, soft contact lenses, 1441–1442
 protein adsorption-resistant surfaces, 278–280
- Biofunctional polymers, 218.** *See also* specific polymer compounds
 ECM-mimetic materials, 218–221
 heparin biocompatibility, 1371–1375
 nitric oxide materials, 1951–1956
 pharmacologically active materials, 220–222
 piezoelectric materials, 2170–2171
 poly(glycerol sebacate), 327–333
 poly(glycolic acid)
 applications, 2244–2245
 basic properties, 2241
 characterization, 2242
 degradation, 2242–2244
 regulatory issues, 2245–2246
 synthesis, 2241–2242
 in vivo tissue response, 2246–2247
 poly(lactic acid)
 biomedical applications, 2254–2255
 characterization, 2252
 degradation, 2252–2254
 formation mechanisms, 2250–2251
 overview, 2249
 synthesis, 2249–2250
 supercritical fluid synthesis, 2527
 tissue engineering, bladder, cell delivery vehicles, 2637
- BioImplant valve, 1313**
- Bioinert materials, bone-implant interface, 425**
- Bioline Coating™, heparin biocompatibility, 1372**
- Biological fixation, bioactive glass, 136**
- Biological fixation techniques, bone healing mechanisms, 414**
- Biological recognition elements, biosensors, 339–343**
- Biological systems**
 controlled release drug delivery, 743–744
 mimics (mimetic materials) of, 298–299
 movement control, basic principles, 729–730
- BioLung total artificial lung, 1678–1679**
- Biomechanics**
 articular cartilage repair, 553–554
 cardiac muscle, 1282–1286
 force-length curves, isolated cardiac muscle, 1282–1283
 future research issues, 1288–1289
 global/regional deformation measurement, 1286–1287
 growth and remodeling, 1287–1288
 myocardial material properties, 1284–1286
 myocardial stresses, 1283–1284
 pressure-volume curves, 1280–1282
 pump functions, 1280–1282
 structural constitutive modeling, 1286–1287
 ventricular function curves, 1280
- dental implants, 817–819
 fibrin clot structure, 1074–1077
- hip joint**
 arthritis, normal and artificial hips, 1377
 cup stability, micromotion, fixation, and osteolysis, 1380
 femoral head subluxation, 1381–1382
 finite element modeling and validation, 1381–1385
 future research issues, 1388–1390
 gait loading results, 1386
 implant fixation techniques, 1379–1380
 liner micromotion, 1385–1386
 pelvic osteolysis and adjunctive screws, 1380–1388
 post THA dislocation, 1381
 subluxation loading results, 1386–1388
 total hip replacement techniques, 1377–1379
 ligament injury and repair, animal repair models, 1645–1647
- Biomedical ethics.** *See* Ethical issues
- Biomedical modeling.** *See* Modeling, biomedical
- Biomedical signaling mechanisms**
 digital acquisition and processing
 acquisition principles, 843–845
 digital filter, 845–846
 event detection, 851–853
 nonlinear biomedical signal processing, 864–867
 overview, 843
 quantization, 845
 recent developments, 854
 sampling process, 844–845
 signal averaging, 848–849
 signal enhancement, 845–847
 spectral analysis, 849–851
 higher-order spectra, 861–864
 nonstationary signals, 858–861
 wavelet analysis, 854–858
- digital biomedical signal compression
 efficiency evaluation, 678
 electrocardiogram, 679–681
 electroencephalography, BAEP recordings, 683–684
 electromyography, 681–682
 electrooculography, 682–683
 fidelity and clinical quality, 678–679
 future applications, 684–685
 lossless, near losses, and lossy methods, 677–678
 motivation, 677
 overview, 677
- BioMedicus BP-80 centrifugal blood pump, 578–579**
- Biomimetic materials**
 biological systems, 298–299
 cell component mimics, 294–299
 membrane, 296–297
 nucleic acid, 297

- Biomimetic materials** (*cont.*)
 proteins, 294–296
 sugar materials, 297–298
 ceramics, 590–591
 elastin, 891–895
 future research, 299
 gradient surfaces interaction, 1219–1225
 cellular interactions, 1222–1225
 protein interactions, 1220–1222
 hydrogels, 1447–1448
 microelectromechanical systems, 1865–1866
 overview, 294
 scaffold structures, 218–220
 surface modification and attachment, 2545–2546
 tissue engineering
 bone, 2656
 peripheral nerve cells, 2816
 valvular prosthetics, 2743–2747
 collagen fiber bundles, 2743–2744
 cross-linked hyaluronans, 2744–2746
 elastin, 2745–2747
- Biom mineralization**
 soft tissue, 71–73
 valvular prosthetics
 anticalcification treatments, 79–81
 cells and cell remnants, 73
 collagen, 74
 elastin, 74–75
 ground substance, 75
 tissue calcification, 75–79
- Biomolecules**
 composite materials, protein adsorption modeling, 2326–2327
 corrosion of implants, 757
 hydrogels, 1446–1447
 polyelectrolyte multilayers, 2223–2226
 self-assembled monolayers, 2413–2415
 surface modification and attachment, 2545–2546
- Bionic Glove neuroprosthetic**, 1929
- Bioparticle mobility, electrophoretic cell separation**, 1701
- Biopex-R[®] ceramic**, 465
- Biopotential amplifiers**
 basic requirements, 302–303
 design criteria, 303–304
 electrocardiography, 301–302, 306–307
 electroencephalography, 302, 307
 electromyography, 302, 307–308
 electrooculogram, 302, 308
 electroretinogram, 302, 308
 interference problems, 304–306
 overview, 301
- Bioprosthetics. See Prosthetic devices; Valvular prosthetics**
- Bioreactors**
 articular cartilage repair, 559
 bio-artificial liver systems, 1656–1658
 bone tissue engineering, 2047
 cardiovascular applications, 314–320
 cardiac patch engineering, 544
 ergonomic features, 316
 integrated control electronics, 317
 material properties measurement, 317–318
 moving magnet motor, 315–316
 multistation configuration, 320
 pressure-flow measurement instrumentation, 319–320
 pulsatile pressure control, 316–317
 rotating-wall bioreactor, 546–547
 cellular characteristics, 312
 clinical applications, 311
 design issues, 313–314
 intelligent bioreactors, 311–312, 314–320
 mechanics of, 310–311
 orthopedic applications, 320–324
 cartilage repair, 559
 chamber modification, 321–322
 gripping and attachment issues, 322–324
 material/construct properties, 323–324
 measurement instrumentation, 323
 mechanical stimulation subassembly, 321
 overview, 310
 rotary cell culture systems, 2402–2410
 tissue engineering, 310, 2623
 bone marrow
 perfusion cultures, 2671–2672
 scaffold materials, 2673–2675
 static cultures, 2670–2671
 hepatic bioreactors, 2772–2773
 vascular system, 3039–3040
- Biorubber polymers**
 design criteria, 327–328
 historical background, 327
 overview, 327
- Biosensors. See also Sensors**
 biological recognition elements, 339–343
 cell-based sensors, 343–344
 components of, 338–339
 detector elements, 344–345
 electrical cell impedance sensing applications, 910–913
 future applications, 913–914
 impedance principles, 908–909
 instrumentation, 909–910
 overview, 908
 electroactive polymeric materials, 926–933
 fiberoptics
 advantages and limitations, 1064–1065
 analytes, 1064
 applications, 1065–1071
 classification, 1065
 combined fiberoptical biosensors, 1069–1071
 components, 1063–1064
 delivery/sensing functions, 1063–1064
 design principles, 1062–1063
 detector, 1064
 future research issues, 1071
 laser/light source, 1063
 optical transducer, 1064
 overview, 1062
 glucose sensors
 blood glucose sensors, 1180–1182
 clinical applications, 1174–1175
 interstitial fluid sensors, 1175–1179
 overview, 1174
 microfabrication, 345
 overview, 335–339
 response time and dynamics, 345–348
 sensitivity, 347
 silk materials, 2445–2446
- Bioseparation, magnetic nanoparticles**, 1709–1710
- Biosynthesis, glycosaminoglycans and proteoglycans**, 1199–1200
- Biotechnology, intellectual property rights and**, 1544–1545
- Bipedal walking, movement control**, 731–732
- Biphasic calcium phosphate (BCP). See also Calcium phosphate ceramics; Tricalcium phosphate (TCP)**
 articular cartilage biomechanics, 91–92
 bioactive, osteoconductive, and osteoinductive properties, 364–365
 biodegradation/biodissolution, 363–364
 as bone-substitute material, 230–232
 cell interactions, 362
 Ceratite[®], 464
 crystal properties, 361
 dissolution properties, 361
 future research, 365
 hydroxyapatite/tribasic composition, 360
 mechanical properties, 361
 microporosity/macroporosity, 360–361
 overview, 359
 post-implant cell and tissue response, 362–363
 preparation, 359–360
 tissue engineering
 articular cartilage, 2700–2701
 scaffold materials, 232–233
- Bipolar electrodes and dipoles, electrical stimulation**, 1011–1012
- Bispectrum cumulants, digital signal acquisition and processing**, 863
- Bite analysis, dental implants**, 818–819
- Black-blood techniques, vascular imaging**, 3017
- Black-box model, movement control**, 734
- Bladder**
 neuroprostheses, 1929–1930
 tissue engineering
 artificial bladder, 2639–2640
 cell transplants, 2636–2640
 future research issues, 2640–2641
 regeneration matrices, 2636
- Bleeding complications, circulatory support systems**, 1779
- Blindness. See Visual function and impairment**
- Block copolymers, hydrogels**, 1445–1446
- Blood brain barrier (BBB)**
 drug delivery, extracellular space, nanoscale structure, 1906–1908
 toxicity testing, animal surrogate systems, 54–55
- Blood composition and structure**
 blood volume, cardiac output measurement, 514–519
 heparin biocompatibility, 1369
 purification
 hemodialysis, 378–384
 hemofiltration, 384
 overview, 378
 plasmapheresis, 384–391
 viscosity, hemodynamics and, 1357–1358
- Blood glucose monitoring. See Glucose monitoring**
- Blood-material interactions**
 artificial liver systems, 1656
 artificial lung systems, diffusional boundary layers, 1665
 artificial surfaces and cardiovascular devices, 373–374
 blood-contacting materials, 374–376
 cardiac assist devices, fluid dynamics, positive displacement pumps, 95–96
 cardiac catheters, 496–498
 circulatory support systems, 1779
 complement system, 372–373
 diamond/diamond-like carbons, 829
 erythrocytes, 373
 fibrin glues, 1185–1187
 fibrin structure and biomechanics, 1072–1077
 foreign body response, 1111–1116
 gradient surfaces, 1210–1211
 hemocompatibility
 active materials, 1346–1348
 inert materials, 1345–1346
 overview, 1345
 platelet interaction, 1353
 protein adsorption, 1350–1353
 competitive adsorption, 1352
 conformation and orientation mechanisms, 1352–1353
 monolayer model, 1351–1352
 surface chemistry, 1352
 surface modification strategies, 1345–1348
 surface properties analysis, 1348–1350

- hemostasis, 367–370
- heparin
- Bioline coating, 1372
 - Carmeda Bioactive Surface, 1371
 - catalytic role, 1368
 - clinical applications, 1372
 - complement and inflammatory response, 1370
 - contact activation, 1370
 - Corline heparin surface, 1371
 - Duraflon surface, 1371–1372
 - enzymatic coagulation interaction, 1370
 - future research, 1373
 - Hepoface coating, 1372
 - immobilized biological effects, 1369–1370
 - overview, 1368
 - Pharmacia HSM intraocular lenses, 1371
 - platelet adhesion/activation, 1369
 - protein interactions, 1370
 - T-NCVC coating, 1372
 - Trillium biopassive surface, 1371
- infection, 1497–1504
- intravascular ultrasound imaging, 1559–1564
- leukocytes, 373
- lung surfactant dysfunction, 1687
- mechanisms of, 1415–1416
- phases of, 1411
- phosphorylcholine technology, 2141–2145
- polyurethane biomaterials, 2310–2312
- thrombosis, 2604–2608
- erythrocyte adhesion, 2607–2608
 - fibrin formation, 2604–2606
 - platelet-dependent neutrophil enhancement, 2606–2607
- tissue engineering, blood vessels, 2648–2650
- ultrasound-tissue interaction, 2937–2938
- valvular prosthetics
- biomineralization
 - anticalcification treatments, 79–81
 - cells and cell remnants, 73
 - collagen, 74
 - elastin, 74–75
 - ground substance, 75
 - overview, 71
 - tissue calcification, 75–79
 - mechanical heart valves, 1334–1335
 - overview, 367
 - platelets, 370–372
 - in vivo animal models, 1566–1571
- Blood oxygenation level dependent (BOLD)
- changes, functional magnetic resonance imaging, 1119
- brain activation localization, 1121
- brain mapping, 1120–1121
- pharmacological effects, 1126
- Blood oxygenators, artificial lung systems, 1666–1667
- Blood vessels. *See also* Circulatory system; Macrocirculatory hemodynamics; Microcirculatory system; Vascular grafts
- centrifugal blood pumps, 574–583
 - embolic materials, 1188–1189
 - growth mechanisms, 3022–3023
 - mass transfer and tissue function, orders of magnitude, 1759–1760
 - mathematical modeling, 1894–1896
 - mechanics of
 - biaxial and shear testing, 394–395
 - component and layer testing, 395
 - development, adaptation, and aging, 399–400
 - elastic arteries, 397–398
 - growth modeling, 397
 - morphological analysis, 395
 - muscular arteries and arterioles, 398–399
 - opening angle measurements, 395
 - overview, 392
 - pseudo-elasticity, 396
 - residual strain and layered models, 396–397
 - in situ studies, 395
 - theoretical framework, 395–396
 - uniaxial testing, 394
 - vascular disease and, 400
 - veins, 399
 - in vitro studies, 394–395
 - in vivo studies, 393–394
 - prostheses, polyepoxy compounds, 2236–2237
 - size, structure, and function, 392
 - stents, nickel alloys, 1942–1949
 - tissue engineering, 2624, 3033–3042
 - cell sources, 2645–2646
 - cellular-based scaffolds, 2647
 - clinical indications, 2643–2644
 - collagen-based scaffolds, 2646–2647
 - construct design, 2646–2647
 - design criteria, 2644–2645
 - future research issues, 2650
 - limitations, 2648–2650
 - mechanical conditioning, 2648
 - microvascular blood flow modeling, 2793–2794
 - mimicry properties, 2645
 - scaffold materials, 2647–2648
 - vascular anatomy, 2643
 - vascular fluid dynamics, 2971–2984
- Body-level mathematical modeling, cardiovascular system, 1900
- Body temperature, polyurethane softening, 2212
- Boltzmann constant
- biomaterials, 836
 - MRI thermometry, molecular diffusion coefficient, 1728–1729
 - nuclear magnetic resonance, 1961
- Boltzmann machine learning, feed-forward neural networks, 123–124
- Boltzmann superposition integral, hard materials viscoelasticity, 1243–1244
- Bombyx mori* silk, 2442–2446
- Bond formation properties, bioactive glasses, 134–135
- Bond strengths, natural adhesives, 237–240
- Bone. *See also* Hard tissue
- animal models
 - anatomy, 36
 - cell biology, 34
 - growth, 36–39
 - human diseases, 40–45
 - microarchitecture and biomechanics, 35–36
 - biocomposite comparisons with, 179–180
 - biologic apatites as, 226–229
 - elastic properties
 - anisotropy, 1241–1242
 - behavioral modeling, 1242–1243
 - future research, 1245
 - overview, 1236
 - physical principles, 1236–1241
 - Reuss orthotropic modulus, 1246
 - Reuss transverse isotropic modulus, 1245
 - Voigt orthotropic modulus, 1246
 - Voigt transverse isotropic modulus, 1245
 - formation mechanisms, 2045
 - fracture fixation
 - bioabsorbable devices, 2073
 - plates and screws, 436–437
 - external devices, 2077–2078
 - skeletal fixation, 419–420
 - internal devices, 414–416, 2073–2077
 - intramedullary nails, 416–418, 2076
 - materials, 414
 - overview, 413
 - plates, 2075–2076
 - poly(lactide-co-glycolide) biocompatibility, 2267
 - rods, pins, and wires, 2074–2075
 - screws, 2075
 - soft tissue-bone attachment, 2076–2077
- fracture mechanisms, 414
- angiogenesis, 1260
 - healing characteristics, 1255–1256
 - precursor cell identity, 1259
 - soluble signals, 1259–1260
- functional properties, 2044–2045
- healing process, 413–414
- overview, 1255
 - primary/secondary bone formation, 1258–1259
 - soft callus formation, 1256–1258
 - tissue engineering, 1260–1261
 - ultrasound therapy, 2951–2959
- hearing and conductivity of, 1272
- implants
- alumina ceramics, 21–22
 - bioactive glass, bond formation properties, 134–135, 138
 - bioadhesive substitutes, 159
 - biphasic calcium phosphate ceramics, cell and tissue response, 362–363
 - cemented implants, 1379–1380
 - cementless implants, 427–428
 - fixation techniques, 426–427
 - hydroxyapatite, 1451–1460
 - interface, 423–430
 - materials, 425–426
 - metal-hard tissue interactions, 1229–1231
 - microporous materials, 1885–1892
 - ongrowth/ingrowth factors, 428–429
 - osteointegration, 424
 - polycaprolactone, 2203
 - reconstructive biomaterials, 2401
 - supercritical fluids, 2528–2529
 - synthetic apatites, 229–231
 - titanium and titanium alloys, 2877–2878, 2880–2881
- morphogenesis, tissue engineering, 2653
- osteoclasts
- basic properties, 2081
 - degradation mechanisms, 2082–2084
 - differentiation and activation signaling, 2084–2087
 - morphology, 2081–2082
 - remodeling applications, 2088
- osteogenic progenitor cells
- aging and, 2092
 - future research issues, 2096–2097
 - growth factor differentiation and proliferation, 2094–2095
 - mesenchymal stem cell lines, 2091–2092
 - overview, 2091
 - phenotypic characteristics, 2092–2093
 - plasticity and interconversion potentials, 2093–2094
 - regeneration and tissue engineering, 2095–2096
- remodeling
- basic multicellular units, 439–440
 - finite element analysis, 1098
 - loading direction, osteocyte apoptosis, 443–444
 - osteoclasts, 2088
 - osteocyte mechanosensing, 441–442
 - secondary bone characteristics, 440–441
 - strain-induced canalicular fluid flow, 442–443
 - strain-related remodeling, 440–441
 - tissue properties, 439
 - unwanted bone loss, 444–445
 - structure and composition, 413, 1236, 2039–2041
 - microcomputed tomography research, trabecular architecture, 1855–1856

- Bone** (*cont.*)
- mineral characteristics, 2041–2042
 - organic matrix, 2042–2044
 - orthopedic device failure analysis, 1824
 - tissue engineering, 2045–2047
 - animal models, 33–45
 - bioactive materials, 144–150
 - cell properties, 2653–2654
 - future research issues, 2657
 - gene-activated matrix regeneration, 1146
 - histiogenesis, 1405
 - overview, 2652–2653
 - scaffold materials, 2654–2655
 - signaling molecules, 2656–2657
 - viscoelasticity, 1243–1245
- Bone cement**
- antiseptic/antibiotic additives in, 65
 - basic properties, 2, 403
 - bioadhesives, 159
 - bone-implant interface, 426–427
 - calcium phosphate ceramics
 - β -tricalcium phosphate, 463–464
 - biphasic calcium phosphate, 464
 - calcium orthophosphates, 461–462
 - cements, 464–466
 - composite hydroxyapatite/poly-L-lactic acid, 466
 - hydroxyapatite, 462–463
 - hydroxyapatite-collagen composite, 466–467
 - overview, 461
 - tissue engineering, 467–468
 - chemical components, 403–404
 - exotherm, 406
 - future research issues, 410
 - hydroxyapatite, calcium phosphate bone cements, 1457–1460
 - mechanical properties, 408–409
 - mixing techniques, 405–406, 409–410
 - orthopedic devices, ceramics, 2067–2068
 - overview, 403
 - physical properties, 406–408
 - polymerization reaction, 404–405
 - porosity and volumetric shrinkage, 407–408
 - residual monomer, 407
 - rheology and viscosity, 406–407
 - total knee replacement, 1612–1613
- Bonaceram[®]** hydroxyapatite, 463
- Bone-inducing growth factor (BP)**, chitosan/glycerol phosphate gels, 605
- Bone-ligament complex**
- structure and properties, 2761
 - tensile testing, 2589–2592
- Bone marrow**
- tissue engineering
 - cell culture parameters, 2662–2665
 - culture media, 2662
 - culture systems
 - perfusion cultures, 2671–2672
 - static cultures, 2670–2671
 - stirred suspension cultures, 2672
 - stromal support, 2670
 - three-dimensional cultures, 2672–2675
 - feeding schedule and culture duration, 2664
 - growth factors, 2662–2663
 - hematopoiesis, 2660–2662
 - inoculum density, 2663–2664
 - mesenchymal stem cells, 2728
 - oxygen tension, 2664–2665
 - pH levels, 2664
 - transplantation
 - gene therapy and, 1159–1160
 - hematopoietic stem cell sources, 1340–1341
 - tissue engineering, cell properties, 2653–2654
- Bone marrow-derived progenitor cells (BMPCs)**, tissue engineering, blood vessels, 2645–2646
- Bone-material interactions**
- coating surface properties, 2536
 - dental implants, 810
 - evaluation, 815–816
 - fracture fixation materials, 416, 2076–2077
 - implants
 - cementless implants, 427–428
 - fixation techniques, 426–427
 - interface, 423–430
 - materials, 425–426
 - ongrowth/ingrowth factors, 428–429
 - osteointegration, 424
 - poly(lactide-co-glycolide) biocompatibility, 2267
 - wear debris, 3076–3078
 - intramedullary nailing, 416–418
 - ligaments, tissue engineering, 2763
 - nickel alloys, 1948
 - photopolymerization, 2160–2161
 - polycaprolactone composites, 2203
 - tissue engineering, tendons, 2873
 - titanium and titanium alloys, 2880–2881
 - ultrasound therapy, 2955–2956
 - wear debris, 3058–3063
- Bone morphogenetic proteins (BMPs)**
- bone healing, 1259
 - histiogenesis, 1407–1408
 - hydroxyapatite-collagen composites, 466
 - osteogenic progenitor cell differentiation and proliferation, 2094–2095
 - osteoinductive substances and materials, 2102–2103
 - tissue engineering
 - cell properties, 2653–2654
 - signaling molecules, 2657
- Bonetite[®]** hydroxyapatite, 463
- Bone vibration masking**, tinnitus device design, 2614–2615
- Bonfil[®]** hydroxyapatite, 462–463
- Bonject[®]** composite, 466
- Born approximation**, ultrasound-tissue interaction, 2938
- Bottom-up processing**
- magnetic nanoparticles, 1705
 - microelectromechanical systems, 1866–1870
 - tissue engineering, polymer materials, 2295
- Boundary conditions**, finite element analysis, 1094
- Bovine serum albumin (BSA)**, chitosan retention, 607
- Bowman's layer**, defined, 747
- Box structures**, dendrimer-base drug delivery, 790–792
- Braided fabrics**
- cardiac catheters, 495
 - three-dimensional structure and properties, 1052
 - two-dimensional structure and properties, 1049
- Brain auditory evoked potential (BAEP)**, signal compression, 683–684
- Brain structure and function**
- cortical mapping, tinnitus device design, 2617
 - electrical stimulation, 1017
 - functional magnetic resonance imaging, 1121–1123
 - hearing mechanisms, 1277
 - magnetic resonance imaging, 1720–1723
 - mass transfer and, 1751–1752
 - nanoscale neurosurgery and, 1903–1906
 - nuclear magnetic resonance imaging, 1967–1970
- Branching parameters**
- polymer structure/property relationships, 2275–2276
 - stent-grafts, endovascular applications, 2480–2481
- Breast**, tissue engineering
- adipose tissue cellular components, 2679–2680, 2682
 - anatomy, 2679
 - animal models, 2683
 - autologous transplants, 2679
 - cell sources, 2680–2681
 - clinical indications, 2678
 - current treatment and reconstruction options, 2678–2679
 - drug delivery systems, 2683–2684
 - future research issues, 2683–2684
 - injectable composite systems, 2682–2683
 - prosthetics, 2678–2679
 - scaffold materials, 2681
 - three-dimensional in vitro systems, 2684
 - in vitro procedures, 2681–2682
- Breaststroke's knee**, overuse injuries, 1602
- Bridge systems**, real-time telemedicine, 2581
- Bridge-to-transplant technology**, centrifugal blood pump systems, 581–583
- Bridgework configuration**, dental implants, 803
- Brownian motion**
- bacterial adherence and, 63–64
 - microparticle image velocimetry, 1879–1881
- Bruxism**, dental ceramics and, 594
- Buccal mucosa insulin delivery**, 1515
- Bulk fillers**, dental ceramics, 600–602
- Bulk modulus**
- composite materials, plane stress, 667–668
 - hard tissue elasticity, 1240–1241
- Bulk wave measurements**, hard materials, elasticity properties, 1236–1241
- Burn injuries**
- dressing materials
 - antimicrobial dressings, 456–457
 - basic criteria for, 450
 - classification and definitions, 451–452
 - electrospun polymer nanofibers, 982–983
 - future research issues, 459
 - nonbiological, nonmedicated dressings, 457–459
 - overview, 448–449
 - skin substitutes, biological and synthetic, 451–456
 - pathophysiology, 449–450
 - skin structure and function, 448
 - wound healing, regeneration and repair, 450
- Bursitis**, pes anserinus tendinitis/bursitis, 1604–1605
- Buttler-Volmer equation**, implantable batteries, 1486
- C**
- Cadaver skin**, burn injuries, 454–455
- Caged-ball mechanical heart valve**, structure and components, 1329–1330
- Calcein**, ceramic materials monitoring, 2919–2923
- Calcification**. *See also* Anticalcification treatments
- articular cartilage, 2696
 - glutaraldehyde compounds, 2234–2235
 - soft tissue, 72–73
 - valvular prosthetics, 75–81, 1295–1301, 2738
 - dystrophic calcification, 1297
- Calcium aluminate**, ceramic implants, 588–589
- Calcium binding sites**, heart valve prosthetic failure and, 1308–1309

- Calcium-deficient apatites (CDAs)
 biphasic calcium phosphate preparation, 359–360
 structure and properties, 227
- Calcium phosphate ceramics (CPCs)
 β -tricalcium phosphate, 463–464
 biphasic calcium phosphate, 464
 bone cements, 464–466
 hydroxyapatite, 1457–1460
 bone-implant interface, 428
 calcium orthophosphates, 461–462
 cements, 464–466
 composite hydroxyapatite/poly-L-lactic acid, 466
 hard tissue interactions, 1231–1232
 hydroxyapatite, 462–463
 hydroxyapatite-collagen composite, 466–467
 implant devices, 588
 orthopedic devices, 2067
 overview, 461
 tissue engineering, 467–468
 articular cartilage, 2700–2701
 transparent ceramic materials, 2916–2923
- Calcium staining, ceramic materials, 2918
- Callus formation, bone healing, 1256–1258
 cells and signals, 1259
- Calmodulin, heart valve prosthetic failure, calcium binding site inactivation, 1308–1309
- Canalicular fluid flow, bone remodeling, 442–443
- Cancellous bone
 biocomposite comparisons with, 181–182
 bone-implant interface, 429
- Cancellous screws, fracture fixation, 415, 2075
- Cancer diagnosis and therapy
 angiogenesis inhibitors, 27
 chemotherapy vs., 30–31
 controlled-release drug delivery, 887–888
 gene therapy, 1150
 host response mechanisms, 1158–1159
 magnetic nanoparticles, 1708–1709
 magnetic resonance imaging, 1720
 optical detection
 confocal microscopy, 2015
 continuous-wave imaging, 2011
 diffuse optical tomography, 2013
 fluorescence method, 2014–2015
 frequency-domain measurements, 2012
 miscellaneous techniques, 2015
 optical coherence tomography, 2013–2014
 overview, 2009
 radiative transfer method, 2009–2014
 Raman spectroscopy and multiphoton techniques, 2015
 spectroscopy measurements, 2011–2012, 2014–2015
 time-resolved measurements, 2012–2013, 2015
 orthognathic surgery, bioabsorbable plates and screws, 436
 positron emission tomography, 2320
 proton resonance frequency shift, 1733–1734
 rotary cell culture systems, 2405–2406
- Candida* spp., biofilm formation and, 214–215
- Capacitive coupling, biopotential amplifiers, 305–306
- Capillaries
 mass transfer and tissue function, 1749–1751
 microcirculatory oxygen transport, 1846
 vascular fluid dynamics, 2980–2981
- Carbodiimide-mediated cross-linking
 collagen tissue, 646
 hyaluronan, 1425–1426
- Carbon. *See also* Pyrolytic carbon (PyC)
 blood-contacting materials, 375
 plasma deposition and coating, 2181
 pyrolytic carbon
 biocompatibility, 2375
 design criteria, 2375–2381
 development and applications, 2371–2374
 examples of, 2379–2380
 fatigue testing, 2377
 fracture mechanics, 2377
 future research issues, 2379
 long-term implants, 2376
 microcrack propagation/non-propagation, 2377–2378
 orthopedic applications, 2374–2375
 overview and history, 2370
 physical and mechanical properties, 2375–2376
 structural analysis, 2378–2379
- Carbonated hydroxyapatite (CHA)
 basic properties, 1452–1454
 bioactive glass, properties and applications, 134–136
 clinical applications, 1456–1460
 structure and properties, 226
- Carbon dioxide, supercritical fluids
 basic properties, 2522–2524
 sterilization of medical devices, 2507–2508
- Carbon-hydrogen-nitrogen elemental analysis, hydroxyapatite characterization, 1458
- Carbopols
 first-generation bioadhesives, 152
 soft-tissue adhesives, 3–5
- Carcinogenesis
 in-vivo biocompatibility testing, 174
 total hip arthroplasty, 1480–1481
- Cardiac arrhythmias
 defibrillator structure and function, 772–777
 electrocardiography, 944–945
 transmyocardial laser revascularization, 1636–1637
- Cardiac assist devices (CADs)
 artificial heart, 475–477
 positive displacement pumps, fluid dynamics, 95–105
 battery power sources, 1484–1495
 blood-material interactions, 373–374
 centrifugal blood pumps
 clinical applications, 581–583
 components, 575–580
 overview, 574–575
 performance evaluation, 580–581
 circulatory support systems
 Abiocr implantable artificial heart, 1780–1781
 ambulatory implantable left ventricular assist devices, 1777–1778
 complications, 1778–1780
 implantable pulsatile devices, 1781–1782
 overview, 1775–1776
 pump physiology, 1776
 clinical indications, 470–471
 complications, 477
 defibrillators
 biophysics, 773–774
 external, 774
 implantable, 774–777
 overview, 772
 ventricular fibrillation/tachycardia, 772–773
 heart anatomy and function, 470
 intraaortic balloon pumps, 471–472
 optical mapping, 2024–2025
 pediatric devices, positive displacement pumps, 97–100
 roller pumps, 471
 surface coatings, 2532
 testing and evaluation, 476–477
 ventricular assist devices, 472–475
- Cardiac catheterization
 Amplatz catheters, 500–501
 bypass grafts, 501–503
 catheter properties, 498–499
 current materials and technology, 494–498
 diagnostic applications, 492–493
 femoral artery approach, 493–494
 heparin biocompatibility, 1368–1373
 historical background, 491–492
 infection prevention, 1503
 Judkins catheters and technique, 499–500
 pigtail catheters, 503–504
 right heart techniques, 504
- Cardiac elastography
 clinical applications, 510–512
 methods, 508
 results, 508–510
 strain imaging, 507
- Cardiac electrophysiology
 atria, 480–481, 488
 atrio-ventricular node, 481–482, 488
 cardiac action potential, 482–487
 excitability recovery, 484–485
 ionic mechanisms, 483–484
 cardiac pacemakers and, 532–533
 electrocardiography and, 937–938
 future research issues, 487–489
 His-Purkinje system, 482, 488–489
 optical mapping
 artifacts, 2021–2022
 cardiac cell culture, 2025–2026
 dye pharmacological effects, 2021
 experimental applications, 2024–2025
 fluorescence principles, 2019
 fluorophore instrumentation, 2019–2020
 future research issues, 2026–2027
 light collection systems, 2020–2021
 light sources, 2020
 mechanical motion artifacts, 2021–2022
 noise artifacts, 2022
 organ mapping, 2024
 overview, 2019
 photodiode array optical system design, 2022–2024
 tissue isolation, 2024–2025
 overview, 480
 sino-atrial node, 480, 487–488
 ventricle system, 480–489, 482
- Cardiac hypertrophy, electrocardiography, 943–944
- Cardiac output measurement. *See also* Heart anatomy and function
 blood volume, 514–517
 dilution methods, 517–13, 517–528
 Doppler ultrasound imaging, 2948
 Fick method, 517–518
 impedance method, 526–530
 indicator-dilution method, 518–522
 overview, 514
 stroke volume calculation, 527–528
 thermal-dilution method, 522–526
 verification studies, 528–530
- Cardiac pacemakers. *See* Pacemakers
- Cardiac patch engineering. *See also* Distal protection devices
 biomaterial substrates, 546
 bioreactors, 544
 biorubber/poly(glycerol sebacate) degradation, 333
 cell injection, 543
 collagen-myocyte interaction, 545
 construct tissue engineering, 543–544
 electron microscopy applications, 974–975
 extracellular matrix arrangement, 545
 future research issues, 550
 historical background, 543–546
 integrins, 545

- Cardiac patch engineering (*cont.*)
 mechanical forces, 545–5
 myocyte phenotype regulation, 545
 overview, 542–543
 rotating-wall bioreactor, 546–547
 techniques for, 544–545
- Cardiac valve surgery, computer-assisted techniques, 691
- Cardiopulmonary bypass (CPB)
 artificial lung systems, 1679–1680
 oxygenator devices, 1667–1669
 coating surfaces and, 2533
 nitric oxide inhalation therapy, 1953–1954
 roller pumps, 471
- Cardiovascular disease. *See also* Congenital heart disease; Pediatric cardiology
 atherosclerosis, macrocirculatory hemodynamics, 1366–1367
 bioreactor applications, 311, 314–320
 cardiopatch engineering, 542–543
 ergonomic features, 316
 integrated control electronics, 317
 material properties measurement, 317–318
 moving magnet motor, 315–316
 multistation configuration, 320
 pressure-flow measurement instrumentation, 319–320
 pulsatile pressure control, 316–317
 blood vessel embolic materials, 1188–1189
 blood vessel mechanics and, 400
 Doppler ultrasound imaging, 2948–2949
 fat tissue engineering, 2733–2734
 fibrin sealant applications, 1086–1087
 fibrin surgical glue, 1186–1187
 finite element analysis, 1097–1098
 hyaluronan applications in, 1427–1428
 intravascular ultrasound imaging, 1559–1564
 magnetic resonance imaging, 1724
 positron emission tomography, infarct imaging, 2319–2320
 transmyocardial laser revascularization, 1632–1637
 vascular fluid dynamics, 2983–2984
- Cardiovascular system
 mathematical modeling, 1897–1900
 tissue engineering, microscale technology, 2785–2787
- Carmedia Bioactive Surface (CBAS), heparin biocompatibility, 1371
- Carotid artery
 Doppler ultrasound imaging, 2945–2947
 percutaneous intervention
 distal protection devices, 870–871
 in vivo animal models, 1570–1571
- Carotid sinus
 hypersensitivity syndrome, pacing therapy, 540
 macrocirculatory hemodynamics, 1367
- Carrageenan, hydrogels, 1447
- Cartilage. *See also* Articular cartilage biomechanics, 85–92
 hard tissue interactions, 1229–1233
 histogenesis, 1409
 polycaprolactone grafts, 2202–2203
 proteoglycan degradation and aging, 1200–1201
 regeneration and repair
 applied tissue engineering, 554–555
 bioreactor applications, 320–324
 cell constructs, 556–557
 chemical factors, 557–558
 chondroitin sulfate pharmacology, 624–625
 in-vitro culture
 hydrostatic pressure, 558
 mechanical environment, 558
 shear forces, 559
 overview, 552
 rotary cell culture systems, 2405
 scaffold selection, 555–556
 surface modification, 557
 synthetic polymers, 556
 vitrification, 3050–3054
- Cascade membrane technology, plasmapheresis, 387
- Casein hydrolysis, magnetic/electrophoretic cell separation, 1702–1703
- Casson's equation
 blood flow modeling, microvascular tissue engineering, 2793–2794
 blood viscosity, 1358
 vascular fluid dynamics, 2978
- Casting techniques
 dental ceramics, 601
 titanium and titanium alloys, dental implants, precision casting, 2887–2889
- Casts, fixation materials, 2077–2078
- Cataracts. *See also* Corneal implants
 intraocular lenses, 1977–1978
- Catechol-quinone systems, collagen cross-linking, 649
- Cathepsins, organic matrix degradation, 2083–2084
- Catheters and catheterization
 cardiac catheterization
 Amplatz catheters, 500–501
 bypass grafts, 501–503
 catheter properties, 498–499
 current materials and technology, 494–498
 diagnostic applications, 492–493
 femoral artery approach, 493–494
 historical background, 491–492
 Judkins catheters and technique, 499–500
 pigtail catheters, 503–504
 right heart techniques, 504
 central venous catheters, antiseptic/antibiotic materials and, 65
 infection
 adjunct therapies, 1503–1504
 bacterial infections, 1498–1500
 colonization vs., 1500–1501
 cost issues, 1497–1498
 detection of, 1501
 device-related factors, 1500
 host factors, 1500
 overview, 1497
 prevention techniques, 1501–1503
 interstitial fluid sensors, dialysis catheters, 1178–1180
 intravascular respiratory catheters, 1675–1677
 nanoscale neurosurgery, 1907–1908
 transmyocardial laser revascularization, 1634–1635
- Cathodic reaction, corrosion electrochemistry, 754–757
- Cations
 hydroxyapatite substitution, 1452
 integrin concentration, 1528
 valvular prosthetics, anticalcification treatments, 80
- Cauchy equation, movement control models, 733–734
- Cauchy-Green deformation, heart biomechanical modeling, 1287
- Cavitation
 mechanical heart valves, 1331, 1335–1336
 ultrasound-tissue interaction, 2940–2941
- CD4⁺ T lymphocytes, immune response, 264–265
- CD34 glycoprotein, hematopoietic stem cells enrichment, 1341
 progenitors, 1338–1339
- CD44 hyaladherin receptor, structure and properties, 1422
- Cell activation
 biphasic calcium phosphate ceramics, 362–363
 cardiac conduction system, 532
 gradient surfaces, 1222–1225
 hyaluronan and, 1422–1423
 inherently conductive polymers, 699–701
 integrin regulation, 1533–1534
 oxygen transport, 1842–1843
- Cell adhesion
 biofunctional polymers, 218–220
 cell culture assays, 565
 electrical cell impedance sensing, 911–913
 glass ceramics, 2918–2929
 histogenesis, 1405–1406
 polyelectrolyte multilayers, 2223–2226
- Cell-based biosensors, development of, 343–344
- Cell-cell interactions, bio-artificial liver systems, 1654
- Cell cultures. *See also* Whole-cell biocatalysis; specific cell types, e.g. Hepatocytes
 analogs, toxicity testing, animal surrogate systems
 microscale designs, 57–58
 three-dimensional modeling, 56
 articular cartilage regeneration, 556–557
 assays
 biocompatibility, 562
 biomaterial testing, 563–564
 bone marrow tissue engineering, 2660
 electrical cell impedance sensing, 910–913
 flow cytometry, 1104–1109
 future research, 566–567
 hematopoietic stem cells
 cell source, 1340
 enrichment, 1340–1341
 future potential of, 1342
 mesenchymal stem cells, 1341
 progenitor cell assays, 1338–1340
 in-vitro testing, 565–566
 overview, 561–562
 sorting, flow cytometry, 1104–1106
 standards for, 563
 cardiac patch engineering, 543–544
 cell culture analogs (CCA), 56–58
 electrorotation and dielectrophoresis, 2631–2632
 fabric scaffolds, 1058
 gradient surface interactions, 1222–1225
 rotary cell culture systems
 cancer, 2405–2406
 cartilage, 2405
 future applications, 2406–2407
 infectious disease, 2406
 insect cells, 2406
 liver, 2406
 neural tissue, 2406
 overview, 2402
 pancreatic islets, 2406
 theoretical principles, 2402–2403
 tissue engineering applications, 2403–2410
 tissue engineering, 2620–2621
 articular cartilage, 2697–2698
 blood vessels, 2645
 bone marrow
 perfusion cultures, 2671–2672
 static cultures, 2670–2671
 stirred suspension cultures, 2672
 stromal support, 2670
 three-dimensional cultures, 2672–2675
 breast, 2680
 cardiac muscle, 2686–2687
 liver, 2770–2771
 microvascular networks, 2798–2799
 transparent ceramics, 2917–2918

- Cell death
osteoclasts, 2084
valvular prosthetics, 73
- Cell injection techniques, cardiac patch engineering, 543
- Cell-material interactions
basic principles, 568
biomaterial classification, 568–569
cell functions, 569–570
microtopography, 571–572
natural biomaterials, 570–571
polyurethanes, 2307
stem cells, 2475
synthetic materials with RGD peptides, 571
- Cell-mediated immune response
host reactions, 1417
xenografts, 3099
- Cell migration, vascular fluid dynamics, 2979–2980
- Cell populations, aortic valve cusp microstructure, 2742
- Cell separation
applications, 1702–1703
dielectrophoresis, 2632–2633
electrophoretic separation
existing methods, 1699
heat transfer, 1701–1702
mass transfer, 1700–1701
mathematical models, 1700
multistage methods, 1699–1700
future research issues, 1703
magnetic cell separation
existing methods, 1695–1696
mathematical modeling, 1697
multistage method, 1696–1697
viscous medium model, 1698–1699
magnetic nanoparticles, 1709–1710
overview, 1695
scale-up and economic issues, 1702
- Cell sheets, tissue engineering, esophagus, 2721
- Cell signaling
articular cartilage, tissue engineering, 2697
interfacial transport, biomaterials
basic principles, 2347–2348
current applications, 2348–2349
electrolyte concentration and permeability, 2352–2353
future research issues, 2354–2355
gold-palladium alloy coatings, 2354
membrane transport, 2349–2358
molecular mass
artificial membranes, 2350–2351
permeability and, 2354
pH effects, 2351–2352
overview and history, 2346
pore size and permeability, 2353–2354
water/wastewater treatment applications, 2349
- surfactants
biomaterials applications, 2340–2343
cationic materials, 2342–2343
future research issues, 2344
hydrophilicity/hydrophobicity, 2341–2342
intracellular/extracellular mechanisms, 2339–2340
natural materials, 2337–2338
overview, 2337
pH and LDH activity, 2341–2343
semipermeable membrane transport, 2343
- Cell sources, tissue engineering
fat tissue, 2727–2728
intestinal system, 2753–2754
vascular system, 3036–3039
- Cell structures
biofilm formation, detachment and dissolution mechanisms, 212–213
bioreactors and differentiation, 310–311
cell culture assays, proliferation mechanisms, 564
confocal microscopy analysis, 711
connective tissue healing and proliferation, 1251–1252
cryopreservation
cold shock injury, below-freezing water temperatures, 767
future research, 768
mechanisms, 767–768
osmotic and potential toxic effects on, 766–767
overview, 762–765
slow freezing, 765–766
storage period, 766
vitrification, 766
warming process, 766
ligaments, 2761
magnetic resonance imaging, 1726
magnetic resonance microscopy, 1744–1747
mass transfer and tissue function, 1751
mathematical modeling, cardiovascular system, 1898
mechano-transduction, 1789–1790
mimics of components, 294–298
peripheral nerve injury, 2118–2120
tissue engineering
bladder, cell transplants, 2637
fat tissue, 2726–2727
topographical surface analysis
chemistry vs., 2552
classification, 2551–2556
discontinuous substratum, 2555
fibers, 2553
microfabrication, 2552
nanotopography vs. microtopography, 2556
overview, 2551
pits/holes, 2553
repeating grooved substrata, 2553
roughened surfaces, 2555–2556
single steps, 2552–2553
substratum influences, 2556–2558
tunnels/tubes, 2555
in vivo applications, 2558–2560
xenografts, 3099–3100
- Cellular solids. *See* Foams, polymer foams
- Cellulosic dialysis membranes, 1583–1584
- Cemented implants, total hip arthroplasty, 1379–1380
- Cementless implants
bone-implant interface, 427–428
total hip arthroplasty, 1379–1380
total knee replacement, 1612
- Center-of-mass modeling, gait analysis, 1138
- Central difference interrogation (CDI), microparticle image velocimetry, 1875–1876
- Central nervous system (CNS). *See also* Peripheral nerves
disorders, controlled-release drug delivery systems, 888
electrical stimulation
basic principles, 1009–1012
brain tissue stimulation, 1017
electrochemistry, 1015–1017
membrane-electric field interaction, 1012–1015
overview, 1009
nanoscale neurosurgery and, 1903–1906
tissue engineering, gene-activated matrix regeneration, 1147–1148
- Centrifugal blood pumps
clinical applications, 581–583
components, 575–580
continuous flow ventricular assist device, 473–474
overview, 574–575
performance evaluation, 580–581
purging vs. nonpurging systems, 576–577
- Centrifugation techniques, bone cements, 409–410
- Ceramics. *See also* Bioactive materials; Glasses and glass ceramics
alumina, 17–18
biocomposites, resorbable vs. nonresorbable components, 183–184
biomimetics, 590–591
biphasic calcium phosphate
bioactive, osteoconductive, and osteoinductive properties, 364–365
biodegradation/biodissolution, 363–364
cell interactions, 362
crystal properties, 361
dissolution properties, 361
future research, 365
hydroxyapatite/tribasic composition, 360
Japanese manufacturing of, 464
mechanical properties, 361
microporosity/macroporosity, 360–362
overview, 359
post-implant cell and tissue response, 362–363
preparation, 359–360
- calcium phosphates
 β -tricalcium phosphate, 463–464
biphasic calcium phosphate, 464
calcium orthophosphates, 461–462
cements, 464–466
composite hydroxyapatite/poly-L-lactic acid, 466
hydroxyapatite, 462–463
hydroxyapatite-collagen composite, 466–467
overview, 461
tissue engineering, 467–468
- cell-material interactions, 568–569, 571
coatings, 589–590
composites, 589
degradable polymer-ceramics, 783–785
nonporous polymer-ceramics, 784–785
- dental implants
accuracy and fit, 600
adhesion properties, 603
aesthetic criteria, 593–594
alumina-reinforced ceramics, 599
bulk fillers, 600–602
CAD-CAM techniques, 601–602
casting techniques, 601
copy milling, 601
fabrication methods, 600
feldspathic porcelain, 595–597
glass ceramics, 598
glass-infiltrated cores, 598
high-alumina porcelain, 597–598
hot pressing techniques, 601
leucite-reinforced ceramics, 598
limitations of, 595
oral environment, 594
overview, 593
porcelain-metal bonds, 597
preform inlays, 602
resin composites, 600
sintering techniques, 600–601
teeth structure and properties, 594–595
zirconia-reinforced ceramics, 599
- electroactive polymeric materials, 926–933
fracture and fatigue, 586–587
hard tissue interactions, 1231–1232
hydroxyapatite, 587–588
clinical applications, 1456–1460
implants, 588–5
microporous materials, 1886–1889
orthopedic devices, 2066–2068
osteoinductive substances, 2105
overview, 585
piezoelectric materials, 2166–2167

- Ceramics (*cont.*)
 powder-based fabrication, 147
 processing, 585–586
 shape memory metal coatings, 2423
 sol-gel fabrication, 147–148
 tissue engineering
 bone, 2655
 glass-ceramics, 144–150
 total hip arthroplasty, 1472, 1474
 transparent ceramics
 alkaline phosphatase activity staining, 2918
 calcium staining, 2918
 cell culture, 2917–2918
 mineralized matrix, 2918
 osteogenic differentiation, 2919–2920
 surface characterization, 2917–2919
 tissue engineering applications, 2916–2923
 wear debris, 3058–3063
 zirconia ceramics
 composition and properties, 3104–3105
 dental applications, 3105–3106
 YTZP crowns, 3106–3108
- Ceramics process, glass dental ceramics, 598
- Ceratitis[®] biphasic ceramic, 464
- Cervical spine plates, anterior cervical spine plate, failure analysis, 1832–1834
- Chain-growth polymerization, poly(lactic acid), 2250–2251
- Chain polymerization, poly(lactide-co-glycolide) synthesis, 2260–2261
- Characterization techniques, multilayer assemblies, polyelectrolytes, 2221–2222
- Charge-coupled device (CCD)
 endoscopic techniques, 987–993
 optical mapping, 2020–2021
 small-angle light scattering, soft tissue analysis, 2452
 wireless capsule endoscopy, 992–993
- Charge-duration curve, electrical stimulation waveform, 1015
- Chemical fiber optic biosensors, 1068–1069
- Chemical fixation techniques
 allografts, 8–9
 microscale tissue engineering, 2782
- Chemical shift, nuclear magnetic resonance, 1962
- Chemical vapor deposition (CVD), pyrolytic carbon, 2371
- Chemiluminescence, biosensors and, 339–343
- Chemokines, immune response, major histocompatibility complex, 266–267
- Chemotaxis, electrical cell impedance sensing, 913
- Chemotherapy
 angiogenesis inhibitors *vs.*, 30–31
 controlled-release drug delivery systems, 887–888
- Children. *See also* Neonatology; Pediatric cardiology
 orthognathic surgery, bioabsorbable plates and screws, 436–437
 radiofrequency electromagnetic field/radiation effects in, 257–258
 ventricular assist devices, positive displacement pumps, 97–100
- Chiral separation, multilayer assemblies, polyelectrolytes, 2223
- Chitosans
 applications, 604
 azide chitosan lactose, 1192
 chitosan/glycerol phosphate gels, applications, 605
 drug delivery systems, 604–609
 first-generation bioadhesives, 152
 food technology, 613–614
 gene therapy, 607–611
 hydrogels, 1446–1447
 inherently conductive polymer composites, 699
 orthopedic devices, 2069
 overview, 604
 tissue engineering, 611–613, 2291
- Chloramphenicol acetyltransferase (CAT)
 expression, chitosan nanoparticles, 609–610
- Chondrocytes
 articular cartilage, 558–559
 tissue engineering
 composition, 2694–2695
 differentiated chondrocytes, 2701–2702
 function, 2696–2697
 zonal organization, 2695–2696
 hyaluronan biomaterials, 1427
- Chondroitin sulfate (CS)
 biological properties, 622
 biomaterials applications, 626
 future research issues, 626
 glycosaminoglycans, 1195
 osteoarthritis efficacy, 1204–1205
 overview, 618
 pharmacological effects, 624–625
 proteoglycans, 619–622
 structure, 618–619
- Chondromalacia, patella, 1597–1598
- Christiansen/Lo three phases model, composite micromechanics, 666–667
- Chromium and chromium alloys
 cobalt alloys, orthodontic wires, 2057
 corrosion mechanics, surface analysis, 758–759
 nickel, 1942–1943
 orthopedic devices
 cobalt alloys, 2062
 fracture fixation and joint implantation, 1827
 hip prostheses, failure analysis, 1839–1841
 total hip arthroplasty, 1476, 1478
- Chromophores
 biosensor development, 344–345
 laser-tissue interactions, 1626
- Chromoscopy, endoscopic techniques, 993–994
- Chronic artificial lung (CAL), 1678–1679
- Chronic obstructive pulmonary disease (COPD)
 artificial lung systems, overview and background, 1661–1662
 nitric oxide inhalation therapy, 1954
 Chronic rejection, xenografts, 3099
- Chronic ulcer, healing deficiencies and, 1254
- Ciliary neurotrophic factor (CNTF), live cell ocular implants, 1992
- Circular dichroism (CD) spectropolarimetry, protein-surface interactions, 280–281
- Circulatory system. *See also* Blood vessels; Macrocirculatory system; Microcirculatory system; Vascular grafts
 anatomy, cardiac output and, 514–516
 blood vessel size, structure, and function, 392
 centrifugal blood pumps, 574–583
 clinical applications, 581–583
 components, 575–580
 overview, 574–575
 performance evaluation, 580–581
 fluid dynamics
 cardiac assist devices, positive displacement pumps, 95–105
 centrifugal blood pumps
 clinical applications, 581–583
 components, 575–580
 overview, 574–575
 performance evaluation, 580–581
 intravascular ultrasound imaging, 1561–1564
 ligament injury and repair, animal repair models, 1646–1647
 mechanics of
 biaxial and shear testing, 394–395
 component and layer testing, 395
 development, adaptation, and aging, 399–400
 elastic arteries, 397–398
 growth modeling, 397
 morphological analysis, 395
 muscular arteries and arterioles, 398–399
 opening angle measurements, 395
 overview, 392
 pseudo-elasticity, 396
 residual strain and layered models, 396–397
 in situ studies, 395
 theoretical framework, 395–396
 uniaxial testing, 394
 vascular disease and, 400
 veins, 399
 in vitro studies, 394–395
 in vivo studies, 393–394
 microcomputed tomography research, 1855–1856
 near-infrared absorption spectroscopy sensor, blood glucose monitoring, 1182
 support devices
 Abiocor implantable artificial heart, 1780–1781
 ambulatory implantable left ventricular assist devices, 1777–1778
 complications, 1778–1780
 implantable pulsatile devices, 1781–1782
 overview, 1775–1776
 pump physiology, 1776
 tissue engineering, 1185–1188
- Clark sensors, evolution of, 336–337
- Clausius-Clapeyron equation, cryopreservation of cells, 764–765
- Clays, ceramic implants, 588–589
- Clearance parameters, artificial kidneys, 1588–1589
- Cleavage mechanisms, proteoglycan degradation and aging, 1201
- Cleveland Clinic Cor-Aide centrifugal blood pump, 578–579
- Closed-loop control systems
 bioreactor design, 312–313
 glucose hydrogels, 1166–1171
 insulin delivery, 1521
 movement control, 732–734
- Closed-mold process, biocomposite fabrication, 189–190
- Clotting factors
 blood-material interactions, 367–370, 1415–1416
 mechanical heart valves, 1334–1335
 fibrin structure and biomechanics, 1072–1077
- Coagulase-negative staphylococci (CoNS), biomaterials-related infections, 61–62
- Coagulation cascade
 blood-material interactions, 367–370
 host reactions, 1415–1416
 cardiopulmonary bypass, 1668–1669
 fibrin sealants, 1080
 fibrin structure and biomechanics, 1073–1077
 hemocompatibility and, 1345
 heparin biocompatibility, 1370
 thrombosis, 2601–2064
- Coatings
 anticell adhesion coatings, 2534
 antiencrustation coatings, 2533
 antimicrobial coatings, 2534–2536
 antiseptic/antibiotic biomaterials, 64–65

- artificial membranes, gold-palladium alloy coating permeability, 2354
- bioactive glasses and, 137–138
- blood-contacting materials, 375–376
- cardiac catheters, 496–498
- ceramics, 589–590
- hydroxyapatite, 587–588
- dental implants, 814
- diamond/diamond-like carbons, 822–823
- biocompatibility, 828
 - ion beam conversion coatings, 824–825
 - smoothness, 826–827
- electrospun polymer nanofibers, 983
- functionality, 2533
- future research, 2538
- heparin biocompatibility, 1371–1372
- hydroxyapatite, 1456–1460
- clinical applications, 1465–1466
 - future research issues, 1467–1468
 - limitations, 1466–1467
 - overview, 1464
 - plasma spraying, 1464–1465
 - retrieval studies, 1466
 - in vivo studies, 1465
- maneuverability properties, 2531–2533
- medical devices, 2531–2538
- microcomputed tomography research, 1856–1857
- orthopedic devices, porous coatings, 2063
- osteoconductivity, 2536
- overview, 2531
- phosphorylcholine technology
- applications, 2142–2146
 - basic principles, 2136
 - blood-material interactions, 2142–2145
 - coating stability, 2139
 - drug delivery capability, 2139–2140
 - future research issues, 2146–2147
 - ocular implants, 2145–2146
 - polymer evolution and synthesis, 2136–2138
 - surface mobility and swelling properties, 2139
 - in vitro bioevaluation, 2140–2141
- plasma deposition, 2181–2182
- protein adsorption modeling
- biomaterial properties and surface analysis, 2326–2328
 - overview, 2326
 - simulation modeling, 2328–2333
- protein-based drug delivery systems, 2361–2362
- pyrolytic carbon, 2371–2374
- shape memory metal coatings
- ceramics, 2423
 - hydroxyapatite, 2423
 - polymers, 2423–2424
- silicone hydrogel contact lenses, 724–725
- stent devices, 2495
- surface modification, 2537–2538
- biomolecule attachment/biomimetic surfaces, 2545–2546
 - methods, 2450
 - plasma surface techniques, 2543–2545
 - ion implantation methods, 2543
 - polymerization, 2543–2545 - polymer grafting, 2545
 - radiation-based methods, 2540–2541
 - ion beam techniques, 2540–2541
 - laser, ultraviolet, and gamma radiation, 2541–2543 - self-assembled monolayers, 2545
 - wet chemistry, 2545–2546
- surface passivation, 2536–2537
- sutures, 2568
- synthetic hydroxyapatite, 230–231
- thromboresistant coatings, 2534
- tissue engineering
- bone, 2655
 - liver, 2771–2772
 - total hip arthroplasty, 1476–1478
 - total knee replacement fixation, 1612
 - vascular prosthetics, in vivo animal modeling, 1571
- Coaxial cylinder, vascular fluid dynamics, 2977–2978
- Cobalt and cobalt alloys
- chromium alloys
 - orthodontic wires, 2057
 - orthopedic devices, 2062 - orthopedic devices
 - chromium alloys, 2062
 - fracture fixation and joint implantation, 1827
 - hip prostheses, failure analysis, 1839–1841 - total hip arthroplasty, 1476, 1478
- Cochlea
- anatomy, 1263–1264
 - hearing loss, 1276–1277
- Cochlear implants
- basic principles, 1507–1508
 - battery system, 1493
 - clinical issues, 1512–1513
 - electroneural interface, 1511–1512
 - functional air conduction mechanisms, 1277–1278
 - hearing aid design, 1269
 - neuroprosthetic technology, 1930–1931
 - normal hearing mechanisms, 1507
 - overview, 1507
 - receiver-stimulator implementation, 1510–1511
 - speech processor implementation, 1508–1510
- Coculture models, toxicity testing, animal surrogate systems, 55
- Co-drug reservoir ocular implants, 1991
- Coefficients of thermal expansion (CTE)
- ceramics, 589–590
 - diamond/diamond-like carbons, 822–823
- Congenital defects, bladder tissue engineering, 2640
- Cold shock injury, cryopreservation and, 767
- Cole-Cole equation, tissue dielectric properties, 836
- Collagen
- aortic valve cusp microstructure, 2739–2740
 - articular cartilage
 - anatomy and function, 553
 - composition and structure, 85–86, 2695
 - scaffold materials, 555–556 - biocomposite comparisons with, 181–182
 - bone matrix composition, 2042–2044
 - cardiac patch engineering
 - bioreactor applications, 546–547
 - myocyte interaction, 545 - cell-material interactions, 570–571
 - degradation, 634–636
 - alkali- and enzyme-treated collage, 635
 - denaturation, 635–636
 - extraction, 635
 - methods, 657–658
 - neutral salt-soluble collagen, 635
 - synthetic cross-linking, 634–635 - fibrillogenesis, 639–643, 657
 - fibroblast mechano-regulation, 1783–1788
 - fixation properties
 - future research issues, 650
 - overview, 639
 - strategies and techniques, 642–649
 - tissues and materials, 639–643 - future research, 636
 - gene-activated matrix scaffolds, 1145–1146
- inherently conductive polymer composites, 699
- ligament injury and repair, 2759–2760
- animal models, 1640–1647
 - tissue engineering and, 2763–2765
- matrix metalloproteinases
- collagenases, 1766–1768
 - gelatinases, 1768–1769
 - historical background, 1763–1764
 - interstitial collagenase (MMP-1), 1766–1767
 - membrane-type, 1771
 - neutrophil collagenase, 1767–1768
 - overview, 1763
 - stromelysins, 1769–1771
 - structure and regulation, 1764–1766
 - tissue engineering applications, 1771–1772
- osteoinductive substances and materials, 2103
- overview, 628
- periodontal ligament, biomechanics, 799–802
- piezoelectric properties, 931–933
- polycaprolactone composites, skin grafts, 2202
- polyepoxy cross-linking with, 2232
- hemostats, 2238
- polyurethane biomaterials, 2311–2312
- processing
- future applications, 657
 - industrial applications, 654–656
 - laboratory applications, 656–657
 - leather goods, 652
 - overview, 652
 - recovery, 652–653
- scaffold materials
- articular cartilage, 555–556
 - extracellular matrix, 1021–1025
 - gene-activated matrix scaffolds, 1145–1146
- small-angle light scattering
- crimp phenomenon, 2458–2459
 - fiber architecture mapping, 2459
- structure and function, 628–630
- primary structure, 629
 - secondary structure, 629
 - tertiary structure, 630
- synthesis, 630–7
- cross-links, 633–634
 - fibrillogenesis, 632–633
- tissue engineering
- articular cartilage, 2695
 - bladder, 2636–2641
 - blood vessels, 2646–2647
 - cardiac muscle, 2686
 - corneal implants, 2710–2711
 - esophagus, 2721
 - fat tissue, 2729–2730
 - historical perspective, 2620–2621
 - polymer sources, 2289–2292
 - skeletal muscle scaffolds, 2841–2844
- type IV, angiogenesis inhibitors, 27
- valvular prosthetics
- calcification, 74
 - fiber bundles, 2744
 - matrix mechanics, 2742–2743
 - mechanical injury, 1299–1300
- vascular grafts
- development strategies, 2998
 - Mandrel-guided FCTs, 1434–1436
- Collagenases, structure and classification, 1766–1768
- Colloids, sol-gel bioactive glasses, microporous materials, 1887–1888
- Colon, tissue engineering, organoid transplants, 2865
- Colonization, medical device infections, 1500–1502

- Colony-forming units (CFUs)
bone marrow tissue engineering, culture media, 2662–2664
hematopoietic stem cell progenitors, 1338–1339
osteogenic progenitor cells, 2092
- Combustible materials, microporous materials, 1887
- Common-mode rejection ratio (CMRR), biopotential amplifiers, 303–304
- Compact bone, bone-implant interface, 429
- Compatibility. *See* Biocompatibility
- Competence stimulating peptide (CSP), biofilm formation, 210–211
- Competitive adsorption, hemocompatibility assessment, 1352
- Complement system
blood-material interactions, 372–373
heparin biocompatibility, 1370
host reactions, 1412–1414
thrombosis and activation of, 2604
- Complexed DNA, gene-activated matrix, 1144–1145
- Complex Motion neuroprosthetic, 1929
- Compliance mechanisms, vascular grafts, 2989–2990
host reactions, 3008–3009
- Composite materials. *See also* Hybrid materials
alumina matrix composites, 23
anisotropic stiffness, 662–666
biomolecular-material composites, protein adsorption modeling, 2326–2327
bone tissue-engineering scaffolds, 180–181
ceramics, 589
dental resin composites, 600
clinical applications, 672–673
degradable polymers, 779–780
clinical applications and limitations, 782–783
nonporous polymer-ceramics, 784–785
scaffold materials, 784
- disordered composites
micromechanics, 666–671
phase continuity and percolation, 671–672
- fabrication methods, 185–190
closed-mold processes, 189–190
fiber meshes/bonding, 190
freeze drying, 187–188
gas foaming, 186
melt molding, 186–187
phase separation, 188–189
solvent-casting particulate-leaching, 186
future research issues, 193
glasses, 137–138
historical background, 659
hydrogels, 1449
hydroxyapatite, 231–232, 1457
hydroxyapatite-collagen ceramics, 466–467
hydroxyapatite-poly-L-lactic acid, 466
- inherently conductive polymers, 699
- ionic polymer-conductor composites, artificial muscles, 112–113
- mechanical properties, 184
- melt-spun reinforcements, 1822–1823
- micromechanics, 659–661
- microporous materials, 1890–1891
- modeling, hard tissue elasticity and, 1242–1243
- nanotechnology, 193
- natural composites, 181–182
- ordered composites, micromechanics, 666–671
- orthopedic devices, 2068–2069
- overview, 179–180, 659
- polycaprolactone, 2200–2201
- rapid prototyping, 190–192
- synthetic biocomposites, 182–184
tissue engineering, 144–145, 149
breast, 2682–2683
- Compounding process, cardiac catheters, 494–495
- Compression
articular cartilage biomechanics, 88–91
direct compression, 558
dynamic compression, 416, 1829–1832
hearing aid design, 1267–1268
periodontal ligament, biomechanics, 800–802
tissue engineering, microvascular networks, 2796
- Computational models
atrioventricular valve mechanics, 1324–1326
biodegradable polymers, computational chemistry, 202–203
electromyography signal quantification, 964–965
ligament and tendon viscoelasticity, 2596–2597
toxicity testing, animal surrogate systems, 51–53
- Computed tomography angiography (CTA), vascular imaging, 3013–3014
future research issues, 3018
- Computed tomography (CT)
intravascular ultrasound imaging, 1561–1564
microcomputed tomography
future research, 1857
historical background, 1850–1851
image generation, 1851–1852
image reconstruction, 1853
osteoporosis research, 1854–1855
overview, 1850
precision and accuracy, 1853–1854
projection data collection, 1853
three-dimensional measurements, 1854
enteric coated tablets, 1856–1857
trabecular bone architecture, 1855
vessel imaging and angiogenesis, 1855–1856
X-ray and computerized tomography theory, 1851
x-ray attenuation, 1852
stress fracture diagnosis, 2515
- Computer-assisted design/computer-assisted manufacturing (CAD/CAM). *See also* Robotics
bioactive material fabrication, solid free-form techniques, 148–149
biocomposite fabrication, 190–192
dental ceramics, 598–600
microporous materials, 1887
rapid prototyping
basic principles, 2382–2383
fused deposition modeling, 2387–2388
future research issues, 2388–2389
selective laser sintering, 2384–2386
stereolithography, 2383–2388
three-dimensional printing, 2386–2387
total hip arthroplasty, implant design, 1381
zirconia ceramics, 3105–3106
- Computer-assisted surgery
cardiac valve surgery, 691
da Vinci system, 689–690
esophageal surgery, 691
evolution of, 687–688
future applications, 692–693
overview, 687
pelvic surgery, 691–692
Zeus system, 688–689
- Computer numerical controllers (CNCs), microelectromechanical systems, 1860–1864
- Concentration diffusion gradient
hemodialysis, 379–380
mass transfer and tissue function, 1756–1761
microparticle image velocimetry, 1879–1880
- Condensation cure, silicone thermosetting elastomers, 2431–2433
- Conditional entropy, digital biomedical signal acquisition and processing, 865–866
- Conditioned medium systems, toxicity testing, animal surrogate systems, 56–57
- Conditioning film deposition, medical device infections, 1500
- Conducting polymers (CPs)
artificial muscles, 110
electroactive polymeric materials, 926–929
conducting polymers, 926–929
drug delivery systems, 932–933
electro-responsive gels, 932
overview, 926
piezoelectric polymers, 929–932
electrospinning technology, 978–983
- inherently conductive polymers
biological functionality, 697–699
cellular communication, 699–701
chemical polymerization, 697
electrochemical properties, 699
electrochemical synthesis, 696–697
enzyme-catalyzed polymerization, 697
future applications, 701–702
overview, 695
photochemical polymerization, 697
structure and properties, 695–696
synthesis, 696–699
- Conductive hearing loss, mechanisms of, 1264–1265
- Conductivity, tissue engineering, 833–834
- Confidentiality, intellectual property rights and, 1543–1544
- Confocal laser scanning microscopy (CLSM), bioreactor analysis, 312
- Confocal microscopy, 705
biomedical optics, 2035–2036
cancer diagnosis and therapy, 2015
combined fiberoptic biosensors, 1070–1071
endoscopic techniques, 996–997
fast confocal microscopy, 712–713
fluorophores, 711
incident light intensity, 710–711
living cells, 712
multicolor fluorescence, 712
Nipkow disk, 713
photobleaching, 711–712
pinhole size, 710
principles and instrumentation, 705–706
procedures and applications, 707–709
resolution, 710
three-dimensional visualization, 709–710
two-photon microscopy, 713
- Congenital deformities
cardiovascular, nitric oxide inhalation therapy, 1953–1954
spinal deformity biomechanics, 2469–2470
- Congestive heart failure (CHF)
cardiac patch engineering, 542–543
circulatory support systems
Abiocor implantable artificial heart, 1780–1781
ambulatory implantable left ventricular assist devices, 1777–1778
complications, 1778–1780
implantable pulsatile devices, 1781–1782
overview, 1775–1776
pump physiology, 1776
transmyocardial laser revascularization, 1633–1637
- Conjugate Residual Method, blood flow modeling, microvascular tissue engineering, 2793–2794

- Conjugation mechanisms, dendrimer-drug delivery systems, 793
- Connective tissue. *See also* Soft tissue allograft materials and techniques, 13–14 classification, 1251 fibroblast mechano-regulation, 1787–1788 healing process in deficient healing, 1254 fibrosis, 1254 healing potential, 1253–1254 inflammation, 1251 outcomes, 1254–1255 overview, 1249 post-injury hemostasis, 1250 proliferation, 1251–1252 regeneration, 1254–1255 remodeling, 1252–1253 soft tissue healing, 1249–1250 tissue engineering, 1255 scaffold materials, 2833
- Consequentialist principles, biomedical ethics and, 289–290
- Constitutive modeling composite materials, 662–663 heart biomechanics, 1286–1287 periodontal ligament, biomechanics, 801–802 piezoelectric materials, 2165–2166
- Contact activating system, heparin biocompatibility, 1370
- Contact angle measurement, hemocompatibility assessment, 1350
- Contact factor activation, blood-material interactions, 369–370
- Contact guidance, cell surface topography, 2557–2558
- Contact lenses. *See also* Intraocular lens (IOL) diamond/diamond-like carbons, 829 hydrogels silicone chemical coatings, 725 core surface treatments, 725 current commercial sources, 725–727 evolution of, 718–719 hydrophilic components, 720 hydrophobic comonomers, 721 ion permeability and on-eye movement prediction, 722 macromer architecture, 719–720 ocular environment, 715 overview, 715 oxygen permeability, 715–717 on phases, 722–723 plasmas, 724–725 polymer phases, 721–722 recent developments, 719 research background, 717–718 surface properties, 723–725 transport properties, 721 soft lenses, 1439–1442 fouling effects, 1441–1442 historical background, 1440 materials requirements, 1442 oxygen permeability, 1441 silicone hydrogel, 1441
- Contact potential, biopotential amplifiers, 305–306
- Contact testing, tendon/ligament tensile testing, 2590–2592
- Continuous arteriovenous hemofiltration (CAVH), acute renal failure, 1590–1591
- Continuous flow ventricular assist device, 473–475
- Continuous renal replacement therapy (CRRT), acute renal failure, 1590–1591
- Continuous-time controlled hybrid movement control model, 736–737
- Continuous-wave (CW) imaging cancer diagnosis and therapy, 2011 Doppler ultrasound, 2943–2944 light propagation in tissue, 2033–2034 nuclear magnetic resonance, 1963–1964 photothermal laser-tissue interactions, 1626–1627
- Continuous Wavelet Transform (CWT), digital signal acquisition and processing, 854–858
- Continuum mechanics composites, 659–661 anisotropic stiffness, 661–666 optical coherence tomography, 1999 protein adsorption modeling, 2328–2329
- Contraction levels, electromyographic signals, 966
- Contrast agents, vascular imaging, 3019
- Contrast-enhanced magnetic resonance angiography, vascular imaging, 3016
- Control engineering movement control, 732–734 wheeled mobility systems, 3085–3087
- Controlled-release drug delivery systems advantages and purposes, 739 antirestenotic agents, 887 biological systems, 743–744 cancer treatment, 887–888 central nervous system disorders, 888 chitosan, 604 clinical applications, 887–888 degradation-erosion-based systems, 741–742, 885 diffusion controlled, 743, 884 dissolution-controlled systems, 884–885 electrospun polymer nanofibers, 982–983 fabrication and encapsulation, 881–883 formulation classifications, 883–887 future applications, 745 glucose hydrogels, 1163–1171 indications for, 880–881 insulin delivery routes, 1514–1516 acute insulin delivery, 1517 artificial pancreas, 1519–1521 basal insulin replacement, 1516–1518 industrial contributions, 1518 inhalation systems, 1515–1516 intravenous administration, 1516 ocular, nasal, and buccal mucosa delivery, 1515 oral enteric insulin, 1515 physiological insulin replacement, 1518–1519 polymeric systems, 1519 subcutaneous, 1514–1515 transcutaneous transport, 1516 ion-exchange systems, 885–886 mechanical/electrical systems, 739–741 ocular implants basic principles, 1982–1983 future research issues, 1992 matrix implants, 1983–1987 inserts, 1984 scleral plugs, 1984–1985 miscellaneous designs, 1986–1987 oculex designs, 1985–1986 overview, 1981 reservoir implants, 1987–1992 codrugs, 1991 corticosteroids, 1989–1990 cyclosporine, 1990–1991 live cells, 1991–1992 osmotic pump systems, 1992 vitasert, 1988–1989 osmotic delivery systems, 740–741, 885 overview, 739, 880 peptides and proteins, 887 polycaprolactone, 2201–2202 polymeric systems, 741–743, 881 prodrugs, 886–887 regulatory issues, 744–745 release profiles and models, 743 soft-tissue adhesives, 3–5 supercritical fluids, 2525–2526 transdermal systems backing materials, 2896 basic principles, 2893 biocompatibility strategies, 2902–2904 inflammatory response, 2904 metabolism inhibition, 2904 skin reactions, 2903–2904 clinical applications, 2897–2899 drug reservoirs, 2896 electrotransport mechanisms, 2897, 2900–2901 applications, 2901 control circuitry, 2900 electrodes, 2900–2901 power source, 2900 mechanically-assisted transport, 2901–2902 micropore creation, 2902 microprojection systems, 2902 needleless injection systems, 2902 passive transport principles, 2894–2895 peelable layers, 2897 permeation enhancers, 2896 pressure-sensitive adhesives, 2896–2897 rate-controlling membranes, 2896 skin structure and function, 2893–2894 system components, 2895
- Convection dialysis membranes, 1587–1588 mass transfer and tissue function, 1761 microcirculatory oxygen transport, 1844–1845
- Cooling rates, bioactive glass, 133–134
- Copolymers biodegradable compounds, 196–199 natural materials, 199–200 hemocompatibility, inert materials, 1346 hydrogels, 1445–1446 interpenetrating polymer networks, 1552–1554 orthopedic devices, 2064 piezoelectric materials, 2168–2169 poly(lactide-co-glycolide) basic properties, 2259 biocompatibility and tissue reactions, 2265–2267 biomedical applications, 2267 crystallinity, 2262 degradation, 2263–2265 future research issues, 2267 mechanical properties, 2262–2263 performance characteristics, 2261–2267 solubility, 2261–2262 synthesis, 2259–2261 thermal stability, 2262 polyurethane biomaterials, 2304 tissue engineering applications, 2293–2294
- Copper ions, as antimicrobial biomaterials, 66–67
- Coprecipitation, magnetic nanoparticles, 1706–1707
- Copy milling, dental ceramics, 601
- Copyright law, 1545–1546
- Corline[®] Heparin Surface (CHS), heparin biocompatibility, 1371
- Corneal implants. *See also* Keratoprostheses; Ocular implants alumina ceramics, 23 fibrin sealant applications, 1086 future research issues, 751 glossary of terms, 747–748 intracorneal implants, 1974–1975 keratoprosthesis, 750–751

- Corneal implants (*cont.*)
 overview, 747
 phosphorylcholine technology, 2145–2146
 refractive keratoplasty, 748–750
 tissue engineering
 cell sources, 2707–2710
 clinical applications, 2715–2716
 collagen gels, 2710–2711
 collagen sponges, 2711
 endothelial cells, 2710
 epithelial cells, 2708–2710
 experimental applications, 2715
 extracellular matrix expression, 2714
 fibrin gel cell cultures, 2710
 future challenges, 2714–2716
 integrin expression and wound healing, 2713–2714
 keratocytes, 2710
 reconstructive technology, 2707–2708
 self-assembly techniques, 2711–2712
 wound healing, 2712–2714
- Corneal reflection, eye movement research and, 1029–1031
- Corona discharge
 plasma biomaterials modification, 2175
 wettability gradient surfaces, 1213
- Coronary angiography, catheterization techniques, 492–493
- Coronary artery bypass graft (CABG)
 cardiac catheterization, 501–503
 tissue-engineered blood vessels, 2643–2644
- Coronary artery disease (CAD)
 stent implants, *in vivo* animal models, 1571
 two-dimensional strain imaging, 506–512
- Correction factor, cardiac output measurement, 520
- Correlation function, microparticle image velocimetry, 1881–1882
- Corrosion
 electrochemical reactions, 754–757
 electrolyte chemical analysis, 759–760
 endovascular stent-grafts, 2486–2487
 metal surface analysis, 758–759
 nickel alloys, 1943–1947
 overview, 754
 shape memory metal surfaces, 2421–2422
 total hip arthroplasty implants, 1479–1480
- Cortical bone, elasticity properties, 1237–1241
- Cortical remapping, tinnitus device design, 2617
- Cortical screws, fracture fixation, 415, 2075
- Corticosteroids, reservoir ocular implants, 1989–1990
- CORx™ artificial lung system, 1679–1680
- Counter-current distribution technology, cell separation, 1695
- Counterion polarization, biological materials, 836–837
- Covalent attachment, cell-adhesive polymers, 218–220
- Covalent chemistry, hemocompatibility assessment, 1352–1353
- Crack propagation
 orthopedic device failure analysis, 1832–1833
 pyrolytic carbon, 2377–2380
 zirconia ceramics, 3106–3108
- Craniofacial trauma fixation, reconstructive biomaterials, 2399–2400
- Cranio-maxillofacial (CMF) skeleton,
 orthognathic surgery, bioabsorbable plates and screws, 434–435
- Creep
 articular cartilage biomechanics, 87
 ligament and tendon viscoelasticity, 2596
- Crevice corrosion, implant materials, 755
- Crimp, ligament mechanics and, 2759–2760
- Critical properties, synthetic apatite composites, 232–233
- Critical-size defects (CSDs), bone tissue engineering
 animal models, 39–40
 basic principles, 33
 phylogenetically similar animal models, 34
- Critical stress intensity factor, ceramic fracture and fatigue, 587
- Crohn's disease, intestinal tissue engineering, 2751–2752
- Cross-bridge models, movement control, 734
- Cross-correlation function (CCF), biomedical signal acquisition and processing, 852–853
- Cross-linked materials
 aquatic adhesives, 240–241
 chitosan nanoparticles, 606–607
 collagen
 degradation, 634–635
 fibrillogenesis, 632–633
 inhibition of, 633–634
 scaffold materials, 641–642
 acyl azide cross-linking, 646–647
 biological compounds, 648–649
 carbodiimide cross-linking, 646
 fixation strategies, 642–644
 polyepoxy compounds, 645
 elastomers, 898–901
 electro-responsive gels, 932
 extracellular matrix, scaffold materials, 1024–1025
 hyaluronan, 1425–1426, 1428
 hydrogels
 glucose hydrogels, 1164–1165
 photopolymerization, 2155
 interpenetrating polymer networks, 1550–1551
 polyepoxy compounds
 artificial organs, 2231
 biomaterials, 2231–2232
 collagen molecules, 2232
 valvular prostheses, 2235–2236
 polymers, 2279
 protein-based drug delivery, 2358–2359
 ultra-high molecular weight polyethylene, 2929–2930
 valvular prosthetics
 failure analysis, 1292, 1306–1308
 hyaluronans, 2744–2746
- Cross-sectional geometry, intramedullary nailing, 418
- Cross-species transplantation, bio-artificial liver systems, 1650–1658
- Cryofiltration, plasmapheresis, 387–390
- Cryo-Life technique, allograft preservation heart valve allografts, 13
 orthopedic allografts, 13–14
- Cryopreservation
 allografts, 12–14
 cryoinjury prevention mechanisms, 765–768
 cryoprotective agents, 762–767
 future research, 768
 living cells, 762–765
 vitrification
 cartilage, 3050–3051
 future research issues, 3053–3054
 overview, 3046
 physiochemistry, 3046–3049
 stability properties, 3049
 tissue engineering, 3049–3050
- Cryotherapy, magnetic resonance thermometry and, 1734–1735
- Crystalline materials
 ceramics
 biphasic calcium phosphate, 361
 processing, 585–586
 degradable polymer composites, 780
 poly(lactide-*co*-glycolide), 2262
 pyrolytic carbon, 2375–2376
- Cultured epidermal autografts (CEAs)
 burn dressings, 456–457
 skin tissue engineering, 2853
- Cumulants, digital signal acquisition and processing, higher-order spectra, 862–863
- Current density
 cardiac pacemakers, 536
 dielectric and conductive media, 1010
 electrical impedance tomography, 917–918
 implantable batteries, 1486
 magnetic resonance imaging, 1719
- Cyanoacrylate
 glue biomaterials, 1188
 tissue sealants, 158–159
- Cyanophycin, 2187
- Cyclic fatigue analysis, zirconia ceramics, 3107
- Cyclic guanosine monophosphate (cGMP), nitric oxide, 1952
- Cyclooxygenase-2, wear debris, inflammatory response, 3070
- Cyclophanes, biomimetic materials, 294–296
- Cyclosporine, reservoir ocular implants, 1990–1991
- Cyclotides, 2188
- Cytoadhesives
 basic properties, 5–6
 biofunctional polymers, 218–220
- Cytokines
 articular cartilage repair, 557
 bone marrow tissue engineering, culture media, 2660–2661
 extracellular matrix, scaffold materials, 1023–1025
 fibroblast mechano-regulation, 1783–1788
 gene-activated matrix, 1142–1143
 host reactions, particulate matter, 2110–2114
 immune response, major histocompatibility complex, 266–267
 matrix metalloproteinases, 1764–1765
 osteoclast differentiation and activation signaling, 2086–2087
 proteoglycan degradation and aging, 1200–1201
 valvular prosthetic failure, pannus formation, 1300–1301
 wear debris
 inflammatory response, 3067–3068
 osteoclasts, 3069–3070
 joint arthroplasties, 3077–3079
- Cytomatrix, surfactant properties, 2338
- Cytomegalovirus, ocular implants
 controlled drug delivery, 1981–1982
 ganciclovir reservoir ocular implant, 1988–1989
- Cytoskeletons, integrin regulation, 1534
- Cytotoxicity
 biocompatibility testing, 170
 cell culture assays, 563–564
 titanium and titanium alloys, 2879–2880
- ## D
- Dark noise, optical mapping artifacts, 2022
- Data acquisition and management
 electromyography, 962–965
 gait analysis, 1129–1130, 1134–1136
 intellectual property rights, 1545–1546
 positron emission tomography, 2317–2318
- da Vinci system, computer-assisted surgery, 689–690
- DDD pacing mode, cardiac pacemakers, 537–538
- Deadhesion, integrins, 1530–1532
- Deafness
 robotics applications, 1812
 ultrasonic acoustic maskers and, 2616–2617

- Debye theory, tissue dielectric properties, dipolar relaxation, 836
- Decibel levels (Db), hearing mechanisms, 1273–1275
- Deep Reactive Ion Etching (DRIE), tissue engineering, microvascular networks, 2795–2796
- Deep vein thrombosis (DVT). *See also* Thrombosis
angiogenesis inhibitors, 29
vascular compression devices and, 2608
- Defibrillators
biophysics, 773–774
external, 774
implantable, 774–777
optical mapping, transmembrane polarization, 2024
overview, 772
ventricular fibrillation/tachycardia, 772–773
waveforms, 773–774
- Degeneration and degradation
collagen, 634–636
controlled-release drug delivery, 741–742, 884
heart valve prosthetic failure, 1296–1297, 1310–1312
hyaluronan, 1423
hydrogels, photopolymerization, 2155–2157
nerve guide technology, 1912–1913
nerve tissue engineering, 2812
osteoclasts, 2083–2084
polycaprolactone, 2198–2200
poly(lactic acid), 2252–2254
poly(lactide-co-glycolide), 2263–2265
polymers
 composite materials, 780
 tissue engineering, 2284–2285
proteoglycan degradation and aging, 1200–1201
total hip arthroplasty implants, 1478–1480
ultra-high molecular weight polyethylene, 2929–2930
valvular prosthetics, 2738
 failure analysis, 1310–1312
 inflammatory degradation, 1297–1298
vascular prosthetics, in vivo animal models, 1569–1570
- Degradation. *See* Biodegradable materials; Degeneration and degradation
- Degranulation, neutrophil-biomaterial adhesion, 1936–1937, 1940
- Degree of interaction, three-dimensional imaging, 1800
- Degrees of freedom (DOF)
computer-assisted surgery, 687–688
movement control models, 733–734
- Dehydration, collagen fixation, 647
- Dehydrothermal treatments (DHT), collagen degradation, 634–635
- Delimiting
tanning industry, collagen processing, 655
- Delimiting, collagen processing, 655
- Demineralization
bone histogenesis, 1408–1409
bone matrix, tissue engineering, bone, 2654–2655
osteoinductive substances and materials, 2104–2105
- Denacol polyepoxy compound, cross-linked biomaterials, 2231–2232
- Denaturation, collagen degradation, 635–636
- Dendrimer structures, drug delivery systems
chemical bonding, 793–794
dendrimer/DNA complexes, 792–793
dendritic box structures, 790
drug conjugates, 793
encapsulation and bonding, 790–793
glycodendrimers, 793
hydrogens, 792
overview, 789
peptide dendrimers, 793
properties, 794–795
synthesis, 789–790
unimolecular dendritic micelles, 790–792
- Denervation, transmucosal laser revascularization, 1632–1637
- De novo adipogenesis, fat tissue engineering, 2727, 2732
- Dental implants. *See also* Periodontal ligament (PDL)
abutment interface, 810
adhesives and composites
 bioactive glasses, 138
 cements, 2, 159
 apatites, 226–229
apical design criteria, 811
biomechanics, 802–805, 817–819
 configuration, 803
 future research issues, 805–806
 integration issues, 803
 mechanical testing, 803–804
 numerical analysis, 804–805
 overview, 799
bone-implant interface, 423–430
ceramics, 587–588
 adhesion properties, 603
 aesthetic criteria, 593–594
 alumina, 22
 alumina-reinforced ceramics, 599
 bulk fillers, 600–602
 feldspathic porcelain, 595–597
 glass ceramics, 598
 glass-infiltrated cores, 598
 high-alumina porcelain, 597–598
 leucite-reinforced ceramics, 598
 limitations of, 595
 oral environment, 594
 overview, 593
 porcelain-metal bonds, 597
 preform inlays, 602
 resin composites, 600
 teeth structure and properties, 594–595
 zirconia-reinforced ceramics, 599, 3105–3106
coating materials, 814
crest module, 811
design principles, 810–813
fibrin sealants, 1084–1085
finite element analysis, 813–814, 1097
force distribution analysis, 811–812
future research issues, 819–820
hard tissue interactions, 1230–1231
histomorphometry, 816
interface evaluation, 815–816
length criteria, 811
machine surfacing, 814
materials properties, 813–819
microtextured titanium, 814
nickel alloys, 1942–1943
orthodontic wires, 2054–2059
osseointegration, 809–810
overview and history, 808
shape geometry, 810
shape memory metals, 2420
surface characteristics, 812–815
surgical technique, 808–809
survival of, 816–817
threads, 810
titanium and titanium alloys, 2887–2889
zirconia ceramics, 3105–3106
- Dentin
apatite composition, 226–227
composition and formation, 2047–2048
tissue engineering, 2048
- Deontological principles, biomedical ethics and, 290–291
- Depolarization, cardiac conduction physiology, 532
- Derivatized biomaterials, hyaluronan, 1424–1425
- Dermatan sulfate (DS)
biological properties, 623–624
future research issues, 626
glycosaminoglycans, 1195
overview, 618
pharmacological applications, 625–626
proteoglycans, 619, 621–622
structure, 618–619
- Dermis regeneration template (DRT), skin tissue engineering, 2854–2856
- Descemet's membrane, defined, 747
- Detector systems. *See also* Sensors
biosensors, 337–338, 344–345
electromyography, 960
fiber optic biosensors, 1064
flow cytometry, 1103–1104
positron emission tomography design, 2315–2316
- Detrended fluctuation analysis (DFA), digital signal acquisition and processing, nonlinear biomedical signals, 864–867
- Developmental anomalies, blood vessel mechanics and, 399–400
- Dextran, hydrogels, 1446–1447
- Diabetes mellitus
pancreatic tissue engineering, 2802–2809
self-regulated insulin delivery, glucose hydrogels, 1165–1170
- Diacrylates, glue systems, 1192
- Diagnomagnetism, magnetic cell separation, 1695–1696
- Diagnostic techniques
angiogenesis, 31
cardiac catheterization, 492–493
- Dialysance, hemodialysis, 383–384
- Dialysate-side clearance equation, artificial kidneys, 1589
- Dialysis. *See* Hemodialysis
- Dialysis catheter interstitial fluid sensors, glucose monitoring, 1178–1180
- Diamond/diamond-like carbons
biomedical applications, 829
cardiovascular applications, 828–829
coating materials, 822–823
 biocompatibility, 828
 ion beam conversion coatings, 824–825
 smoothness, 826–827
deposition techniques, 823–825
direct ion beam deposition, 824
dual ion beam sputtering, 824
filtered cathodic arc deposition, 825
future research issues, 829–830
magnetron sputter coating, 826
orthopedic applications, 827–828
overview, 822
plasma deposition and coating, 2181
plasma source ion implantation and deposition, 825–826
pulsed laser ablation, 824
pyrolytic carbon
 biocompatibility, 2375
 design criteria, 2375–2381
 development and applications, 2371–2374
 examples of, 2379–2380
 fatigue testing, 2377
 fracture mechanics, 2377
 future research issues, 2380
 long-term implants, 2376
 microcrack propagation/non-propagation, 2377–2378
 orthopedic applications, 2374–2375
 overview and history, 2370
 physical and mechanical properties, 2375–2376
 structural analysis, 2378–2379

- Diamond/diamond-like carbons (*cont.*)
 radiofrequency plasma-activated chemical vapor deposition, 826
 tribological properties, 826
- Diazoniumdiolates, polymer biomaterials, 222
- Dielectric coefficient, radiative transfer, optical detection of cancer, 2009–2010
- Dielectric media
 electric stimulation, 1010
 tissue engineering, 834–838
 conductivity and permittivity, 833–834
 counterion polarization, 836–837
 data sources, 841
 dipolar relaxation, 836
 dispersion regions, 839–840
 frequency ranges, 840–841
 interfacial polarization, 837–839
 microwave frequencies, 841
 overview and history, 832–833
- Dielectric relaxation
 biomaterials, 836–837
 tissue engineering, 834–836
- Dielectrophoresis (DEP)
 microparticle image velocimetry, particle trap device, 1876–1878
 tissue engineering, 2629–2630
 cell characterization and separation, 2631–2633
 cell entrapment, 2633–2634
 traveling wave dielectrophoresis, 2630–2631
- Diepoxy cross-linking, hyaluronan, 1425
- Die tooling, polyurethanes, 2213
- Differential equations, movement control models, 733–734
- Differential pulse code modulation, ECG signal compression, 679–681
- Differential scanning calorimetry (DSC), cryopreservation of cells, 764–765
- Differential voltage recordings, electrical impedance imaging, 917–918
- Differentiation mechanisms, stem cells, 2475
- Diffuse optical tomography (DOT)
 biomedical optics, 2037
 cancer diagnosis and therapy, 2013
- Diffuse photon intensity wave, radiative transfer, optical detection of cancer, 2011
- Diffusible indicators, cardiac output measurement, 526
- Diffusional boundary layers, artificial lung systems, 1665
- Diffusion chambers, artificial pancreas, 2803–2804
- Diffusion mechanisms
 bioadhesion, 1–2
 biosensor dynamics, 345–347
 controlled-release drug delivery systems, 743, 883
 dialysis membranes, 1586–1587
 gradient surface preparation, 1211–1212
 light propagation in tissue, 2033–2034
 magnetic resonance imaging, tensor imaging and tractography, 1722
 mass transfer and tissue function
 concentration diffusion, 1756–1757
 reaction and, 1760–1761
 microparticle image velocimetry, 1880–1881
 nerve guide technology, 1918–1919
- Diffusion-weighted imaging (DWI), magnetic resonance imaging, 1722
- Digital biomedical signals
 acquisition and processing
 acquisition principles, 843–845
 digital filter, 845–846
 event detection, 851–853
 nonlinear biomedical signal processing, 864–867
 overview, 843
 quantization, 845
 recent developments, 854
 sampling process, 844–845
 signal averaging, 848–849
 signal enhancement, 845–849
 spectral analysis
 basic principles, 849–851
 higher-order spectra, 861–864
 nonstationary signals, 858–861
 wavelet analysis, 854–858
- compression methods
 efficiency evaluation, 678
 electrocardiogram, 679–681
 electroencephalography, BAEP recordings, 683–684
 electromyography, 681–682
 electrooculography, 682–683
 fidelity and clinical quality, 678–679
 future applications, 684–685
 lossless, near losses, and lossy methods, 677–678
 motivation, 677
 overview, 677
- electrical impedance imaging, 918–923
 electrogastrography, 951–954
 eye movements, 1028–1031
- Digital filter, biomedical signal acquisition and processing, 845–846
- Digital subtraction angiography (DSA), vascular imaging, 3012–3013
- Dilution curve, cardiac output measurement, 521–526
- Dilution principle, cardiac output, 516
- Dilution techniques, cardiac output measurement, 516–526
 Fick method, 517–518
 indicator-dilution method, 518–522
 thermal-dilution method, 522–526
- Dimethyl aminoethyl methacrylate (DMEAEM), glucose hydrogels, 1166–1170
- Dimethylsulfoxide (DMSO), vitrification technology, 3051–3054
- Dipalmitoyl phosphatidylcholine (DPPC), lung surfactants
 basic properties, 1683–1684
 function, 1686–1687
- Diphosphonates, valvular prosthetics, anticalcification treatments, 81
- Dipolar relaxation, biological materials, 836
- Dip-pen nanolithography (DPN), self-assembled monolayers, 2414–2415
- Dipping technique, polyelectrolyte multilayer deposition, 2219
- Direct compression, 558
- Direct contact tests, biocompatibility testing, 171
- Directed collagen gel shrinkage, valvular prosthetics, collagen fiber bundles, 2744
- Directed evolution, biocatalysis and, 166–167
- Direct fiberoptic biosensors, 1068–1069
- Direct ion beam deposition, diamond/diamond-like carbons, 824
- Direct kinematics, movement control theory, 733
- Disabilities, wheeled mobility technology and, 3083–3092
- Disc degeneration, spine biomechanics, 2467–2468
- Discontinuous substratum, topographical surface analysis, 2555
- Discrete Wavelet Transform (DWT), digital biomedical signal acquisition and processing, 856–858
- Discretization techniques, finite element analysis, 1092–1093
 error analysis, 1097
- Disease markers, hyaluronan as, 1423
- Disordered composites
 micromechanics, 666–671
 phase continuity and percolation, 671–672
- Dispersion regions, tissue dielectric properties, 839–841
- Dissolution properties
 biphasic calcium phosphate ceramics, 361
 controlled-release drug delivery systems, 883–884
- Distal protection devices
 AccUNET device, 875–876
 advantages/disadvantages, 877
 Angioguard device, 873–874
 carotid artery, 870–871
 current research evolution, 877–878
 embolization reduction devices, 871
 failure mechanisms, 877
 Filterwire Ex, 875
 future applications, 878
 NeuroShield device, 874–875
 overview, 870
 Parodi Anti-Emboli System, 876–877
 Percutaneous guardwire, 871–873
 saphenous vein grafts, 870–871
 Triactiv (Kensey Nash), 873
- Distance learning, telemedicine systems, 2585
- Distillation, silicone purification, 2437
- Distribution centroid, small-angle light scattering, soft tissue analysis, 2453–2454
- Divinyl sulfone cross-linking, hyaluronan biomaterials, 1426
- DNA chip technology, biosensors and, 341–343
- DNA encapsulation, chitosan nanoparticles, 609–610
- DNA-mediated assembly, microelectromechanical systems, 1869–1870
- DNA shuffling, whole-cell biocatalysis, 165
- DNA structures, dendrimer-base drug delivery, dendrimer/DNA complexes, 792–794
- Documentation requirements, intellectual property and patent management, 1541–1542
- Donnan effect, articular cartilage biomechanics, 553–554
- Donnan ionic equilibrium theory, polyacrylonitrile fibers, artificial muscles, 112
- Dopant anions, inherently conductive polymers, 698
- Doppler ultrasound. *See also* Ultrasound
 arterial disease, 2948
 basic principles, 2942–2943
 cardiac examination, 2948–2949
 carotid artery examination, 2947
 color imaging, 2945–2946
 continuous wave Doppler, 2943–2944
 duplex scanning, 2944–2945
 optical coherence tomography, 2002–2003
 power Doppler imaging, 2946–2947
 pulsed wave Doppler, 2944
 tissue imaging, 2947
 transcranial examination, 2947
 vascular imaging, 3014–3015
- Dreaming research, eye movements and, 1035–1036
- Driven-right-leg approach, biopotential amplifiers, 306
- Drop foot simulator, development of, 1926–1927
- Drug delivery systems. *See also* Controlled-release drug delivery systems
 biofunctional polymers, 220–222
 chitosan, 604–606
 coatings, 2531
 surface properties, 2536–2538

- dendrimer structures
 chemical bonding, 793–794
 dendrimer/DNA complexes, 792–793
 dendritic box structures, 790
 encapsulation and bonding, 790–793
 hydrogens, 792
 overview, 789
 properties, 794–795
 synthesis, 789–790
 unimolecular dendritic micelles,
 790–792
- electroactive polymeric materials, 932–933
- endovascular stent graft diameter,
 2489–2490
- extracellular space, nanoscale structure,
 1906–1908
- fibrin glues, 1187
- hydrogels, glucose hydrogels, 1163–1164
- insulin delivery routes, 1514–1516
 acute insulin delivery, 1517
 artificial pancreas, 1519–1520
 basal insulin replacement, 1517–1518
 industrial contributions, 1518
 inhalation systems, 1515–1516
 intravenous administration, 1516
 ocular, nasal, and buccal mucosa
 delivery, 1515
 oral enteric insulin, 1515
 physiological insulin replacement,
 1518–1519
 polymeric systems, 1519
 subcutaneous, 1514–1515
 transcutaneous transport, 1516
- magnetic/electrophoretic cell separation,
 1702–1703
- ocular implants, 1979
 basic principles, 1982–1983
 future research issues, 1992
 matrix implants, 1983–1987
 inserts, 1984
 scleral plugs, 1984–1985
 miscellaneous designs, 1986–1987
 oculex designs, 1985–1986
 overview, 1981
 reservoir implants, 1987–1992
 codrugs, 1991
 corticosteroids, 1989–1990
 cyclosporine, 1990–1991
 live cells, 1991–1992
 osmotic pump systems, 1992
 vitrasert, 1988–1989
- phosphorylcholine technology
 applications, 2142–2146
 basic principles, 2136
 blood-material interactions, 2142–2145
 coating stability, 2139
 drug delivery capability, 2139–2140
 future research issues, 2146–2147
 ocular implants, 2145–2146
 polymer evolution and synthesis,
 2136–2138
 surface mobility and swelling properties,
 2139
 in vitro bioevaluation, 2140–2141
- polycaprolactone, 2201–2202
- poly(glycolic acid), 2244–2245
- poly(lactic acid), 2254–2255
- polyurethane biomaterials, 2308–2309
- proteins
 absorption and formulation issues,
 2359–2360
 alternate insulin delivery routes, 2366
 enzyme and pH absorption effects,
 2360–2361
 future research issues, 2367
 internal architecture and design,
 2357–2359
 oral insulin delivery, 2364–2366
 overview, 2357
- pH-responsive hydrogels, 2363–2364
 proteolytic activity elimination, 2362
- smart polymers, 2294–2295
- stent devices, drug-eluting stents, 2497
- supercritical fluids, 2525–2526
- surfactants, 2338
- tissue engineering
 breast, 2683–2684
 scaffold materials and, 2657
- transdermal drug delivery
 backing materials, 2896
 basic principles, 2893
 biocompatibility strategies, 2902–2904
 clinical applications, 2897–2899
 drug reservoirs, 2896
 electrotransport mechanisms, 2897,
 2900–2901
 mechanically-assisted transport,
 2901–2902
 passive transport principles, 2894–2895
 peelable layers, 2897
 permeation enhancers, 2896
 pressure-sensitive adhesives, 2896–2897
 rate-controlling membranes, 2896
 skin structure and function, 2893–2894
 system components, 2895
- Drug interactions
 alginate biosynthesis and, 354
 positron emission tomography imaging,
 2321
- Drug reservoirs, transdermal drug delivery,
 2896
- Drug resistance, angiogenesis inhibitors,
 29–31
- Dry mouth syndrome, hard tissue elasticity,
 1245
- Dry table formulas, bioadhesives, 157–158
- Dual-chamber cardiac pacemaker,
 components, 534–535
- Dual ion beam sputtering, diamond/diamond-
 like carbons, 824
- Duplex ultrasound
 basic principles, 2944–2945
 vascular imaging, 3018
- Durability testing, poly(carbonate)urethanes,
 2209–2211
- Duraflon II coating, heparin biocompatibility,
 1371–1372
- Duties, legal principle of, biomedical ethics
 and, 290
- Dynamic compression
 bone fracture fixation, 416
 plate failure analysis, 1829–1832
- Dynamic interlocking fixation technique, bone
 fracture fixation, 418
- Dynamic mechanical analysis (DMA)
 bioreactors, 312–313
 orthopedic applications, 323–324
 movement control models, 733–734
- E**
- Ear. *See also* Cochlea
 anatomy
 hearing aid design and, 1263–1264
 hearing mechanisms and, 1272–1273
 nerve function and, 1275–1276
 implants, bioactive glass, bond formation
 properties, 138
- Ear, nose and throat (ENT) procedures,
 alumina ceramics in, 18, 22–23
- Early-attachment events, biofilm formation
 and, 208–209
- Echocardiography
 basic principles, 506–507
 robotic tele-echocardiography, 1812
 semilunar aortic valve structure,
 1318–1319
- Echo coatings, surface properties, 2532
- Echo-planar imaging (EPI)
 magnetic resonance imaging, 1717
 MRI thermometry, 1729
- Economic issues
 nerve guide technology, 1920
 wheeled mobility devices, 3090
- Ectopic calcification, soft tissue, 73
- Ectopic transcription factor, tissue
 engineering, pancreas, 2806–2807
- Effectiveness factors, mass transfer and tissue
 function, orders of magnitude, 1760
- Effective orifice area (EOA) measurements,
 mechanical heart valve fluid
 mechanics, 1331–1332
- Ehler Danlos syndrome, glycosaminoglycan
 dysfunction, 1202–1203
- Elasticity
 blood vessel mechanics, 396–399
 hard tissue
 anisotropy, 1241–1242
 behavioral modeling, 1242–1243
 future research, 1245
 overview, 1236
 physical principles, 1236–1241
 Reuss orthotropic modulus, 1246
 Reuss transverse isotropic modulus, 1245
 Voigt orthotropic modulus, 1246
 Voigt transverse isotropic modulus, 1245
 suture materials, 2570
- Elastic scattering, biomedical optics, 2031
- Elastin
 aortic valve cusp microstructure, 2740–2742
 basic properties, 891
 bioengineering applications, 894–895
 gene structure, 891
 protein structure, 891–893
 tropoelastin assembly and ultrastructure,
 893–894
 valvular prosthetics
 biomimetic constructs, 2745–2747
 calcification, 74–75
 matrix mechanics, 2742–2743
- Elastohydrodynamic lubrication, articular
 cartilage biomechanics, 92
- Elastomers. *See also* specific polymers
 biodegradability, 897, 905–906
 poly(carbonate)urethanes, 2208–2214
 polyesters, 897–901
 polyglycolide, polylactide,
 polycaprolactone copolymers, 897–900
 polyhydroxyalkanoate (PHA) copolymers,
 899
 polytetrafluoroethylene, 2302
 polyurethanes, 901–905
 silicones, 2430–2433
 transdermal drug delivery, backing
 materials, 2896
 urinary prostheses, 2962
- Electrical cell impedance sensing (ECIS)
 applications, 910–913
 future applications, 913–914
 impedance principles, 908–909
 instrumentation, 909–910
 overview, 908
- Electrical impedance tomography (EIT)
 hardware requirements, 916–918
 overview, 915–916
 safety requirements, 918–919
 sample images, 923–924
 software requirements, 919–923
 three-dimensional imaging, 1797
- Electrically charged fluid jet technology,
 electrospun polymer nanofibers,
 978–981
- Electrical stimulation
 cardiac tissue, 1017–1018
 central nervous system
 basic principles, 1009–1012
 brain tissue stimulation, 1017

- Electrical stimulation (*cont.*)
 electrochemistry, 1015–1017
 membrane-electric field interaction, 1012–1015
 overview, 1009
 muscle tissue, 1017
- Electrical systems, controlled-release drug delivery, 739–740, 886–887
- Electric-powered wheelchairs (EPW), components and technology, 3085
- Electric total artificial heart, 475–476
- Electric ventricular assist device, 473
- Electroactive polymers (EAPs)
 artificial muscles
 dielectric elastomers, 109
 electronic/ferroelectric polymers, 108–109
 polyacrylonitrile fibers, 110–112
 conducting polymers, 926–929
 drug delivery systems, 932–933
 electro-responsive gels, 932
 electrospun polymer nanofibers, 980–981
 overview, 926
 piezoelectric polymers, 929–932
- Electrocardiography (ECG/EKG)
 abnormalities, 943–945
 amplifier components, 942–943
 arrhythmias, 944–945
 atrial depolarization and P wave, 938
 atrial repolarization, 939
 atrio-ventricular delay, 938
 atrioventricular node, 944
 biopotential amplifiers, 301–302, 306–307
 cardiac cellular electrophysiology, 937–938
 cardiac conduction physiology, 533
 components, 938–5
 digital signal acquisition and processing, 843–845
 quadratic time-frequency representation, 860–861
 signal enhancement, 845–849
 wavelet transform, 857–858
 digital signal compression, 679–681
 heart block, 944
 heart mechanical activity, 940
 historical background, 936–937
 lead system, 940–941
 measurement of, 941–942
 paper recording, 942
 safety issues, 943
 S-T segment, 939
 U wave, 940
 ventricular depolarization and QRS complex, 938–939
 ventricular repolarization and T wave, 939
- Electrochemical reactions
 battery systems, 1485–1486
 corrosion of implants, 754–757
 electrical stimulation, 1015–1016
 inherently conductive polymers, 696–697
- Electrochemistry, shape memory metal surfaces, 2422–2423
- Electrodes
 electrical cell impedance sensing, 909–910
 electrical impedance tomography, 916–918
 electrical stimulation techniques, 1012–1013
 damage studies, 1016
 tissue interface, 1014
 electrocardiography, 940–941
 electrogastrography, 950–951
 electromyography, 960–961
 electrotransport mechanisms, transdermal drug delivery, 2900–2901
- Electro discharge machining (EDM), 1860–1861
- Electrodynamics, radiative transfer, optical detection of cancer, 2009–2011
- Electroencephalography (EEG)
 biopotential amplifiers, 302, 307
 digital signal acquisition and processing, 843–845
 quadratic time-frequency representation, 860–861
 signal enhancement, 845–849
 wavelet transform, 857–858
 signal compression, 683–684
- Electrogastrography (EGG)
 digital signal acquisition and processing, spectral analysis, 861
 fasting EGG patterns, 954
 future research issues, 955
 overview and historical background, 947
 parameters, 953–954
 physiology, 947–949
 postprandial EGG patterns, 954–955
 recording procedures and equipment, 949–954
- Electrolytes
 corrosion chemistry, 759–760
 interfacial transport, protein permeability, artificial membranes, 2352–2353
- Electromyography (EMG)
 basic principles, 956
 biopotential amplifiers, 302, 307–308
 clinical applications, 966–967
 digital signal acquisition and processing
 biomedical signals, 852–853
 signal compression, 681–682
 spectral analysis, 861
 electrode classification and placement, 960–692
 electrophysiology and, 965–966
 elementary signals, 959
 graphical representation, 965
 motor unit, 956–959
 nerve tissue engineering assessment, 2813–2814
 quantification and data processing, 964–965
 signal changes, 965
 signal conditioning and data recording, 962–964
 signaling mechanisms, 959–960
 signaling processing chain, 960
 spine biomechanics, 2465–2466
 surface signals, 959
- Electron-beam machining (EBM), 1861
- Electroneural interface, cochlear implants, 1511–1512
- Electronic theory
 bioadhesion, 1–2
 flow cytometry, 1103
- Electron microscopy
 cardiovascular tissue engineering, 974–975
 hard biomaterials and polymer specimens, 972–973
 instrumentation, 968–969
 medical applications, 974
 overview, 968
 scanning vs. transmission techniques, 969–970
 SEM specimen preparation, 970–972
 TEM specimen preparation, 973–974
- Electron spectroscopy chemical analysis (ESCA), gradient surfaces
 characterization, 1217–1218
 hydroxy groups, 1213–1214
 protein-surface interactions, 1220–1222
- Electron spin resonance (ESR), magnetic resonance imaging, 1720
- Electrooculography (EOG)
 biopotential amplifiers, 302, 308
 signal compression, 682–683
- Electro-orientation, tissue engineering, 2629
- Electroosmotic flow, polyelectrolyte multilayer microfluidics, 2226
- Electrophoretic cell separation
 applications, 1702–1703
 existing methods, 1699
 future research, 1703
 heat transfer, 1701–1702
 mass transfer, 1700–1701
 mathematical models, 1700
 multistage methods, 1699–1700
 scale-up and economic issues, 1702
- Electrophoretic countercurrent extractor (ECCE), electrophoretic cell separation, 1700
- Electrophysiological signals,
 electromyographic signal association with, 965–966
- Electroporation, gene therapy, 1153
- Electro-responsive gels, basic properties and applications, 932
- Electroretinogram (ERG), biopotential amplifiers, 302, 308
- Electrorotation, tissue engineering, 2630
- Electrospinning
 polycaprolactone tissue engineering, 2203–2204
 polymer nanofibers, 978–983
 polyurethane biomaterials, 2309–2310
 tissue engineering
 scaffold fabrication, 2834–2835
 skeletal muscle scaffolds, 2843–2844
 yarns and fabrics, 1053–1054
- Electrostatic interactions, multilayer assemblies, polyelectrolytes, 2221
- Electrostriction, tissue engineering, 2629
- Electrotransport mechanisms, transdermal drug delivery, 2897, 2900–2901
- Elementary electromyography, signaling mechanisms, 959
- Eligiloy, endovascular stent-grafts, 2486–2487
- Ellipsometry, gradient surfaces, 1217
- Elution tests, biocompatibility testing, 170–171
- Embolization
 cyanoacrylate glue, 1188–1189
 distal protection devices
 AccUNET device, 875–876
 advantages/disadvantages, 877
 Angioguard device, 873–874
 current research evolution, 877–878
 embolization reduction devices, 871
 failure mechanisms, 877
 Filterwire Ex, 875
 future applications, 878
 NeuroShield device, 874–875
 overview, 870
 Parodi Anti-Emboli System, 876–877
 Percutaneous guardwire, 871–873
 saphenous vein grafts, 870–871
 Triactiv (Kensey Nash), 873
- Embryonic stem cells (ESCs). *See also* Fetal tissue
 development, extracellular matrix, 1024–1025
 histiogenesis, 1405–1406
 phenotype, 2473
 sources, 2474
 tissue engineering
 bone, 2653–3
 fat tissue, 2728
 pancreas, 2806–2608
- Emulsion technique, polymer particulates, controlled-release drug delivery systems, 882
- Enamel (tooth)
 apatite composition, 226–227
 composition and formation, 2048–2049
 tissue engineering, 2049
- Encapsulation technology
 controlled-release drug delivery systems, 881–883

- dendrimer-base drug delivery, 790–793
 polyelectrolyte multilayers, 2223–2226
- Endangered Species Act, 1621
- End-diastolic pressure-volume relationship (EDPVR), heart biomechanics, 1282–1283
- Endocardial excursion, echocardiographic imaging, 507
- Endocrine disease, artificial pancreas tissue engineering, 2807
- Endocytosis, basic principles, 994
- Endogenous pathways, immune response, 263–264
- Endoglin expression, hematopoietic stem cell progenitors, 1338–1339
- Endomucosal resection (EMR), esophageal tissue engineering, 2719–2720
- Endoscopic retrograde cholangiopancreatography (ERCP), basic principles, 990
- Endoscopy
 autofluorescence, 994–995
 chromoscopy, 993–994
 emerging technologies, 992–997
 endocytosis, 994
 endoscopic retrograde cholangiopancreatography, 990
 fibrin sealant applications, 1085
 fluorescence imaging, 994
 future applications, 997
 high-magnification endoscopy, 992–993
 image detection, 989–990
 instrumentation and components, 986–989
 narrow band imaging, 993
 non-gastrointestinal applications, 989
 OCT imaging technique, 995–996
 overview, 986
 photodynamic imaging, 995
 ultrasound and, 990–991
 in vivo confocal microscopy, 996–997
 wireless capsule techniques, 991–992
- Endosseous ridge maintenance implant (ERMI), bioactive glass, 138
- Endothelial cells (ECs)
 angiogenesis inhibitors and, 26–27
 corneal implants, 747
 gradient surfaces, 1222–1225
 hemocompatibility, 1347–1348
 impedance sensing, barrier function, 911–913
 tissue engineering
 blood vessels, 2646–2647
 corneal implants, 2710
 microvascular networks, 2798
 vascular grafts, 2992–2993
 development strategies, 2998–2999
 host reactions, 3008
 seeding, 1433–1434
- Endothelial-derived relaxing factor (EDRF), nitric oxide and, 1951
- End stage renal disease (ESRD), artificial kidneys
 biocompatibility, 1583
 cellulosic membranes, 1583–1584
 classification, 1580–1583
 dialysis membranes, 1583–1585
 efficiency, 1581–1583
 flux changes, 1581
 historical perspective, 1578–1579
 hollow-fiber membranes, 1579–1580, 1585
 overview, 1576
 synthetic membranes, 1584–1585
 uremic solute classification, 1577–1578
- End-systolic pressure-volume relationship (ESPVR), heart biomechanics, 1282–1283
- Energy absorption, radiofrequency electromagnetic field/radiation (RFR) effects, 254–255
- Enteric coatings
 microcomputed tomography research, 1856–1857
 protein-based drug delivery systems, 2361–2362
- Entropy, interfacial transport, 2349
- Environmental conditions
 glucose hydrogels, 1164–1165
 interpenetrating polymer networks, 1550–1551
- Environmental stress cracking (ESC), polyurethane biodegradation, 2306–2307
- Enzyme-based electrochemical sensors
 blood glucose monitoring, 1181–1182
 long-term implantation and telemetry applications, 1180
- Enzymes
 biocatalysis
 historical background, 162–163
 inherently conductive polymers, 697
 isolated enzymes, 165–167
 overview, 162
 polycaprolactone, 2197–2198
 whole-cell techniques, 163–165
 biosensors and, 339–343
 coagulation system, heparin biocompatibility, 1370
 collagen degradation, 635
 glycosaminoglycan disorders, replacement therapies, 1203–1204
 hyaluronan degradation, 1423
 inhibitors, blood-material interactions, 370
 mimics (mimetic materials) of, 298–299
 polycaprolactone degradation, 2197–2199
 polyurethane degradation, 2307
 protein-based drug delivery systems, 2360–2361
 suture biocompatibility and, 2574–2575
 valvular prosthetics, calcification, 78
- Epidermal growth factor (EGF)
 integrin structure, 1525
 polymeric structure, 221
- Epikeratoplasty, techniques and materials, 749
- Epineural nerve repair, 2124–2125
- Epithelial cells (ECs)
 burn autografts, 456
 tissue engineering
 corneal implants, 2708–2710
 intestinal epithelium, 2861–2862
- Equation of motion, magnetic cell separation, viscous medium, 1698
- Ergometers, electromyographic signal association with, 966
- Ergonomics. *See also* Gait analysis
 assessment tools, 1002–1005
 bioreactor design, moving magnetic motor, 316–317
 checklists and surveys, 1004–1005
 current research, standards and guidelines, 999–1000
 eye movement research, 1028–1031
 complex task analysis, 1036–1037
 office conditions, 1005–1006
 overview, 999
 program elements, 1002
 wheeled mobility systems, 3087–3088
 work-related musculoskeletal disorders, 999
 low back, 1000–1001
 upper extremity, 1001
- Erosion mechanisms
 controlled release drug delivery, 742
 degradable polymer composites, 780
 endovascular stent grafts, 2485
- Error analysis, finite element analysis, 1096–1097
- Error back propagation, neural networks, 121–123
- Erythrocytes
 blood-material interactions, 373
 thrombosis, adhesion, 2603–2604
- Escherichia coli* species, whole-cell biocatalysis, 164–165
- Esophagus
 computer-assisted surgery, 691
 tissue engineering
 animal models, 2720
 cell sheets, 2721
 clinical applications, 2719–2720
 collagen type 1 scaffolds, 2721
 future research, 2722
 naturally occurring scaffolds, 2721
 synthetic scaffolds, 2720–2721
- Essential nutrient limitation, medical device infections, 1504
- Ester derivatives, hyaluronan, 1425
- ϵ -Caprolactones. *See* Polycaprolactones
- Etching process
 plasma biomaterials modification, 2176
 shape memory metal surfaces, 2421–2422
 tissue engineering, microvascular networks, 2795–2796
- Ethical issues
 animal testing
 AAALAC international accreditation, 1619–1620
 Animal Welfare Act, 1617–1618
 Endangered Species Act, 1621
 future developments, 1621
 Good Laboratory Practices Act, 1621
 Guide for Care and Use of Laboratory Animals, 1620–1621
 Health Research Extension Act, 1618–1619
 overview, 1616–1617
 regulations, 1621
 applied/practical ethics, 287
 autonomy principles and, 290–291
 consequentialist principles, 289–290
 future issues, 292–293
 historical evolution of, 285–286
 justice principles and, 291–292
 legal duties and, 290
 National Research Act of 1974, 286–287
 overview, 285
 principlist principles, 289
 standards, principles, and rules, 287–289
 valvular bioprosthetic development, 1292–1294
 xenografts, 3100
- Ethylene oxide, device sterilization, 2504
- Eukaryotic cells, mass transfer and tissue function, intracellular organelles, 1751–1752
- Euler's equation, spinal deformity biomechanics, 2469–2470
- Evanescence wave fiberoptic biosensors, 1066
- Event detection, biomedical signal acquisition and processing, 851–852
- Evoked potentials (EP), digital signal acquisition and processing, 858
 signal enhancement, 845–849
- Excitation-emission matrix (EEM), optical detection of cancer, 2014–2015
- Exogenous pathway, immune response, 263–264
- Exothermic reaction, bone cement, 406
- Expanded polytetrafluoroethylene (ePTFE), vascular prostheses
 hybrid materials, 1432–1437
 in vivo animal models, 1567–1571
- Explosive surface evaporation, laser-tissue interactions, 1629–1630
- Exposure frequency and duration studies, radiofrequency electromagnetic field/radiation effects, 256

- Ex-situ tissue formation, tissue engineering, bladder, 2637–2638
- External fracture fixation devices, 419–420, 2077–2078
- External hip rotator syndrome, overuse injury, 1397
- Exteroceptors, movement control, 730–731
- Extracellular matrix (ECM). *See also* Integrins
- articular cartilage
 - anatomy and function, 552
 - composition and structure, 85–86, 2694–2695
 - artificial liver systems, 1652–1654
 - biofunctional polymers, 218–222
 - bioreubber/poly(glycerol sebacate), 327
 - bone, 439
 - formation mechanisms, 2045
 - functional properties, 2044–2045
 - mechanosensing, osteocytes, 441–442
 - mineral composition, 2041–2042
 - organic matrix composition, 2042–2044
 - osteocyte apoptosis, 445
 - structure and function, 2039–2041
 - tissue engineering, 2045–2047
 - cardiac patch engineering
 - heart anatomy and function, 545
 - myocyte phenotype regulation, 545
 - pattern, composition, and arrangement, 544–545
 - scaffold constructs, 543–544
 - connective tissue, 1249
 - proliferation mechanisms, 1251–1252
 - fibroblast mechano-regulation, 1783
 - gene expression, 1783–1788
 - histiogenesis, 1406
 - hyaluronan, structure and properties, 1421–1422
 - matrix metalloproteinases, 1763–1764
 - collagenases, 1766–1768
 - gelatinases, 1768–1769
 - historical background, 1763–1764
 - interstitial collagenase (MMP-1), 1766–1767
 - membrane-type, 1771
 - neutrophil collagenase, 1767–1768
 - overview, 1763
 - stromelysins, 1769–1771
 - structure and regulation, 1764–1766
 - tissue engineering applications, 1771–1772
 - medical device infections, 1504
 - nerve guide hybrids, 1916–1917
 - scaffold materials
 - component properties, 1022–1024
 - host response, 1024
 - overview, 1021
 - sources, 1024
 - tissue engineering, 1024–1025, 2828–2829
 - stem cell-biomaterial interactions, 2475
 - surfactant properties, 2338
 - protein interactions, 2339–2340
 - structure and properties, 2337–2338
 - teeth composition and formation, 2047–2049
 - tissue engineering
 - blood vessels, 2646–2647
 - bone, 2655–2656
 - cardiac muscle, 2688–2692
 - corneal implants, 2714
 - current technologies, 2620–2624
 - fat tissue, 2729–2730
 - liver, 2770–2775
 - polymer sources, 2283–2284
 - skeletal muscle scaffolds, 2842
 - skin, 2853–2857
 - vascular system, 3040–3041
 - vascular grafts, 2998
- Extracellular polysaccharides (EPS)
- bacterial adhesives, 243–244
 - biofilm formation, 211
- Extracellular space (ECS)
- drug delivery through, 1906–1908
 - nanoscale neurosurgery and, 1906
- Extracorporeal membrane oxygenation (ECMO)
- artificial lung systems
 - blood oxygenators, 1666–1667
 - cardiopulmonary bypass, 1667–1669
 - venoarterial and venovenous systems, 1669–1670
 - heparin biocompatibility, 1372–1373
 - nitric oxide infusion, 1955
 - roller pumps, 471
- Extraction techniques, collagen degradation, 635
- Extraneous noise, optical mapping artifacts, 2022
- Extremozymes, biocatalysis and, 166
- Extrinsic calcification, soft tissue, 73
- Extrinsic factors
- integrin regulation, 1527–1528
 - vascularization, 3026–3028
- Extrusion process
- cardiac catheters, 495
 - melt spinning of polymers, 1817
 - polyurethanes, 2213
 - tissue engineering scaffolds, 2831
- Eye movement
- anatomy, ocular implant drug delivery, 1982–1983
 - biomedical applications, overview, 1033
 - characteristics, 1028
 - complex tasks, 1036–1037
 - future research, 1038
 - interactive applications, 1037–1038
 - memory, imagery, and dreaming, 1035–1036
 - reading applications, 1034–1035
 - research background, 1028–1031
 - silicone hydrogel contact lenses, on-eye movement prediction, 722
 - spatial attention and scene perception, 1033–1034
 - visual-social language context for, 1035
- F**
- Fabella syndrome (Fabellitis), 1605
- Fabrication techniques
- controlled-release drug delivery systems, 881–883
 - microelectromechanical systems
 - biomimetics, 1865–1866
 - bottom-up process, 1866–1870
 - historical background, 1860–1862
 - overview, 1860
 - top-down processing, 1862–1865
 - tissue engineering
 - acellular matrix and submucosa, 2832–2833
 - blood vessels, 2645–2646
 - connective tissue capsule, 2833
 - electrospinning, 2834–2835
 - extrusion, 2831
 - fiber bonding, 2829
 - freeze drying, 2831–2832
 - gas foaming, 2831
 - melt molding, 2830–2831
 - membrane lamination, 2830
 - phase separation, 2832
 - rapid prototyping, 2833–2834
 - scaffolds, 2829–2835
 - solvent casting and particulate leaching, 2829–2830
 - ultra-high molecular weight polyethylene, 2926–2927
 - vascularization, 3027
 - zirconia ceramics, 3105–3106
- Fabrics
- braided three-dimensional fabrics, 1052–1053
 - electrospun yarns and fabrics, 1053–1054
 - endovascular stent grafts, 2483–2484
 - implantation properties, 1043–1044
 - melt spinning, 1821–1822
 - multiaxial warp, 1051–1052
 - performance characteristics, 1055–1056
 - polymer foams, 2274
 - scaffold materials, tissue engineering, 1056–1059
 - structural components, 1044–1054
 - fibers, 1046–1048
 - textile units, 1044–1046
 - yarn-to-fabric structures, 6–7
 - three-dimensional structural geometry, 1049–1053
 - two-dimensional structural geometry, 1048–1049
 - woven fabrics, 1048–1050
- Facial skeletal fixation, reconstructive biomaterials, 2396. *See also* Maxillofacial bones, ultrasound therapy
- craniofacial trauma, 2399–2400
 - fronto-orbital fractures, 2399
 - zygomatic and midface fractures, 2398–2399
- Factor proteins
- blood-material interactions, 1415–1416
 - coagulation cascade, thrombosis, 2601
 - complement system, host reactions, 1413–1414
 - fibrin structure and biomechanics, 1072–1077
- Fåhræus effect, vascular fluid dynamics, 2979
- Fåhræus-Lindqvist effect, vascular fluid dynamics, 2979
- Failure analysis
- circulatory support systems, 1780
 - composite materials, 660
 - dental implants, 816–817
 - distal protection devices, 877–878
 - orthopedic devices
 - anterior cervical spine plate, 1832–1834
 - dynamic compression plate, 1829–1832
 - fracture fixation, 1827–1839
 - intramedullary nails, 1835–1837
 - overview, 1824
 - prostheses, 1839–1841
 - tibial tray, 1841
 - titanium screws, 1834–1835
 - valvular prosthetics, 1306–1314
 - bioprosthetic heart valves, 1306–1314
 - mechanical heart valves, 1329–1331
 - mechanisms for, 2737–2738
 - vascular grafts, host reactions, 3009
- FAK-related nonkinase (FRNK), integrin deadhesion, 1531
- Fan beam technology, microcomputed tomography, 1853
- Faraday's law, electrotransport mechanisms, transdermal drug delivery, 2900
- Fascicular nerve repair, 2124–2125
- Fast confocal microscopy, 711–712
- Fast Fourier Transform (FFT)
- electrogastrography signals, 951
 - electromyography signal quantification, 964–965
 - microparticle image velocimetry, 1875–1876
 - speech processors, cochlear implants, 1509–1510
- Fast imaging with steady state free precession (FISP), magnetic resonance imaging, 1724
- Fasting EGG patterns, properties, 954

- Fasting whole-body metabolism, mass transfer and tissue function, 1750–1751
- Fast spoiled gradient-echo (FSPGR), temperature measurement, 1732
- Fatigue testing
ceramics, 586–587
mechanical heart valves, 1330–1331
muscle fatigue, electromyographic signals, 966
pyrolytic carbon, 2377
titanium and titanium alloys, 2881–2884
zirconia ceramics, 3107
- Fat pad syndrome, overuse injuries, 1603–1604
- Fat tissue, tissue engineering
cell sources, 2727–2728
clinical applications, 2725
de novo adipogenesis, 2732
growth factors and adipogenic inducers, 2730
origin and characteristics, 2726
research applications, 2733–2734
scaffold materials, 2729–2730
strategies, 2726–2727
surgical techniques and limitations, 2725–2726
synthetic hollow support structures, 2732–2733
vascularization, 2730–2731
in vitro studies, 2731
in vivo studies, 2731–2732
- Feedback suppression, hearing aid design, 1267–1268
- Feedback systems, glucose hydrogels, 1169–1170
- Feed-forward networks
basic properties, 120–121
training methods, 123–124
- Feldspathic porcelain, dental ceramics, 595–597
- Femoral artery
cardiac catheterization through, 493–494
superficial femoral artery, distal protection device intervention, 878
- Femoral head prosthetics
morphology and component rotation, 1612
total hip arthroplasty, 1471–1472
dislocation and subluxation, 1381–1382
total knee replacement, 1609–1610
- Femur, stress fracture, 2518–2519
- Femur-ACL-tibia (FACT) complex, tensile testing, 2591–2592
- Fenestrated stent-grafts, endovascular systems, 2483
- Ferrogels
artificial muscle material, 108
electroactive polymeric materials, 927–929
- Ferromagnetism, magnetic cell separation, 1695–1696
- Fetal tissue. *See also* Embryonic stem cells (ESCs)
bladder tissue engineering, 2640
collagen scaffolds, 642
- Fiber optics, optical coherence tomography, 1996–1998
- Fibers. *See also* Sutures
artificial lungs
annular fiber bundles, 1679–1680
I-VOX intravascular respiratory catheters, 1675–1677
biocomposite fabrication, meshes/bonding, 190
biosensor applications
advantages and limitations, 1064–1065
analytes, 1064
applications, 1065–1071
classification, 1065
combined fiberoptic biosensors, 1069–1071
components, 1063–1064
delivery/sensing functions, 1063–1064
design principles, 1062–1063
detector, 1064
future research issues, 1071
laser/light source, 1063
optical transducer, 1064
overview, 1062
collagen materials, 641–644
degradable polymer composites, 782
delivery/sensing applications, 1063–1064
melt spinning
bioactive glass fibers, 1818–1821
biomedical applications, 1821–1823
braided, 1822
knitted, 1821–1822
polymers, 1817–1818
woven and non-woven, 1821
movement control models, 734
polyamides, 2186–2194
scaffold materials, bonding properties, 2829
silks
Bombyx mori silk, 2442–2446
spider silks, 2448
structural properties, 2442
wild silkworm silks, 2447–2448
small-angle light scattering (SALS) analysis, 2450–2461
structural properties and applications, 1046–1048
topographical surface analysis, 2553
Fibrillogenesis, collagen processing, 632–633, 657
- Fibrils
bone matrices, 2047
collagen materials, 642
valvular prosthetic failure, 1299–1300
- Fibrin
biomatrix, tissue regeneration, 1075–1077
clinical applications, 1080–1081
clot structure and biomechanical properties, 1074–1077
glue materials, 1185–1189
hemocompatibility, 1347–1348
hydrogels, basic properties, 1447
overview, 1072
research background, 1072–1073
sealants
abdominal surgery, 1086
adhesive production and composition, 1082–1084
cardiovascular surgery, 1086–1087
clinical applications, 1080–1081
coagulation pathway, 1080
commercially available products, 1081–1082
endoscopic applications, 1085
fibrin structure and biomechanics, 1072–1077
future research issues, 1087–1088
hepatic/pancreatic surgery, 1086
limitations, 1084
neurosurgical applications, 1084
ophthalmology, 1086
oral surgery applications, 1084–1085
overview, 1080
skin grafts, 1085–1086
thoracic surgery, 1087
tissue repair, 158–159
urologic applications, 1086
thrombosis, 2603–2606
tissue engineering applications, 2290
- Fibrinogen
biological activities, 1076–1077
fibrin structure and biomechanics, 1072–1077
platelet adhesion and aggregation, thrombosis, 2601–2602
protein adsorption behavior, 277–278
- Fibroblast growth factor (FGF), vascular grafts, 3024
- Fibroblast-populated collagen gels (FPCGs), mechano-regulation mechanisms, 1784–1788
- Fibroblasts
foreign body response, 1114
mechano-regulation
cellular transduction, 1789–1790
gene expression, 1783–1788
overview, 1783
tissue engineering, 1788–1789
- Fibro-collagenous tubes (FCTs), hybrid vascular prostheses, 1432–1437
- Fibronectin
cell-material interactions, 571
chondroitin sulfate interactions, 622
extracellular matrix, scaffold materials, 1023–1025
fibroblast mechano-regulation, 1783–1784
hemocompatibility, platelet interaction and, 1353
tissue engineering, polymer sources, 2283–2284
- Fibrosa, aortic valve cusp microstructure, 2738–2742
- Fibrosis, tissue healing and, 1254
- Fick's law
cardiac output measurement, 517–518
controlled release drug delivery systems, 743
dialysis membrane diffusion, 1587
mass transfer and tissue function, concentration diffusion, 1757–1759
- Fidelity properties, digital biomedical signal compression, 678–679
- Field-assisted cell separation, overview, 1695
- Filler materials, silicone thermosetting elastomers, 2433
- Filtered cathodic arc deposition, diamond/diamond-like carbons, 825
- Filterwire EX (Boston Scientific) distal protection device, 875
- Filtration techniques
distal embolization reduction, 871, 877
electrogastrography signals, 951–952
electromyography signal conditioning, 963
silicone purification, 2437
- Finite element analysis (FEA)
articular cartilage biomechanics, 92–93
atrioventricular valve mechanics, 1324–1326
bioengineering applications, 1097–1098
bone remodeling, canalicular fluid flow, 442–443
cardiac biomechanical modeling, 1287
cardiovascular system, 1097–1098
dental implants, 805
dental implants, 1097
design and distribution, 810–811
material properties, 814
electrical impedance imaging, 919–923
error analysis, 1096–1097
future research issues, 1098
musculoskeletal system, 1098
numerical analysis, 1094–5
overview, 1091
periodontal ligament, biomechanics, 802
postprocessing, 1095–1096
preprocessing, 1091–1094
research background, 1091
silicone hydrogel contact lenses, phase morphology, 722–723
tissue dielectric properties, 837–839
total hip arthroplasty, 1382–1385
results analysis, 1388–1390
- Finite impulse response (FIR) filters, biomedical signal acquisition and processing, 847
- First-generation bioadhesives, 152

- Fixation devices, endovascular stent grafts, 2487
- Fixed charge density (FCD), articular cartilage composition and structure, 85–86
- Flammacerium, burn dressings, 457
- Flatbed perfusion systems, bone marrow tissue engineering, 2672
- Flexible ceramics, orthopedic devices, 2066–2067
- Flory-Huggins parameter, interpenetrating polymer networks, 1551
- Flow cytometry. *See also* Cell separation applications, 1106–1109
calibration issues, 1108–1109
cell biology, 1106–1108
cell sorting, 1104–1106
detectors, 1103–1104
electronic systems, 1103–1104
fluidic systems, 1101–1102
microbiology, 1108
optical systems, 1102–1103
overview, 1100–1101
pharmaceutical applications, 1108
plant and animal science, 1108
reproductive medicine applications, 1108
spectral compensation, 1104
stem cell analysis, 2473–2474
toxicity testing, animal surrogate systems, 53–55
- Flow field measurement
mechanical heart valves, in-vitro measurements, 1332–1334
microparticle image velocimetry, 1876–1879
- Flow reversal devices, distal embolization reduction, 871, 877
- Flow surface properties, vascular grafts, host reactions, 3009
- Flow visualization techniques, centrifugal blood pump testing, 580–583
- Fluid dynamics
artificial heart systems, positive displacement pumps, 95–105
flow analysis, rotary cell culture systems, 2402–2410
macrocirculatory hemodynamics, 1359–1366
mass transfer and tissue function, 1752–1756
mechanical heart valves, 1331–1332
blood-material interactions, 1336
microparticle image velocimetry
basic principles, 1873–1876
correlation function, 1881–1882
flow field measurement, 1876–1879
DEP particle trap device, 1876–1878
microviscometry, 1878–1879
overview, 1873
particle detection, 1879–1881
microvascular networks, tissue engineering, 2793–2794
polyurethane biomaterials, 2312
- Fluidic systems. *See also* Microfluidics
flow cytometry, 1101–1102
- Fluidized bed reactor, pyrolytic carbon processing, 2370–2374
- Fluid mechanics, vascular fluid dynamics, 2983–2984
- Fluorescence imaging
biomedical optics, 2035–2036
cancer diagnosis and therapy, 2014–2015
ceramic materials, mineralized matrix visualization, 2918
confocal microscopy, 705–706
endoscopic techniques, 994
light-tissue interactions, 2032–2033
microscopy, 707
optical detection of cancer, spectroscopic techniques, 2011–2012, 2014–2015
optical mapping, 2019
- Fluorescence lifetime imaging microscopy (FLIM), biomedical optics, 2036
- Fluorescence photobleaching recovery (FPR), confocal microscopy, 711
- Fluorescence recovery after photobleaching (FRAP), confocal microscopy, 711
- Fluorinated polymers,
polytetrafluoroethylene, 2302
- Fluoroapatite (FAp)
orthopedic devices, 2067
structure and properties, 226–228
- Fluorophores
confocal microscopy, 710
optical mapping, 2019–2020
- FM acoustic masking, tinnitus, 2613
- Foam materials
bioactive scaffold fabrication, 149–150
burn dressings, 458–459
gas foaming
biocomposite fabrication, 186
polymer foams, 2271
supercritical fluid synthesis, 2527
tissue engineering scaffolds, 2831
microporous materials, 1887–1890
orthopedic devices, ceramics, 2068
polymer foams
cellular solids, 2270
fabrication, 2271–2274
fugitive phase removal, 2271–2273
gas foaming, 2271
overview, 2270
phase separation, 2273
pore formation, 2272–2273
sintering, 2273–2274
structure, morphology, and mechanical properties, 2271
textile processing, 2274
- Focal adhesion kinase (FAK), integrin deadhesion, 1530–1532
- Focused ultrasound therapy (FUS), magnetic resonance imaging, temperature measurement, 1730–1731
- FOCUS Night&Day silicone hydrogel contact lenses, 727
- Food technology, chitosan, 613–614
- Foot modeling, gait analysis, 1137
- Footwear, stress fractures, 2513
- Force feedback, robotics-based surgical technology, 1808–1809
- Forcefield molecular mechanics, protein adsorption modeling, 2330–2332
- Force generation, dental implants, 818
- Foreign body giant cells (FBGCs), chronic inflammation, 1411–1412
- Foreign body response. *See also* Host reactions; Immune response
fibroblasts, 1114
giant cell response, 1114
host reactions, 1411–1419
interstitial fluid sensors, capsule formation, 1175–1179
mechanisms of, 1115–1116
medical device infections, 1497–1498
overview, 1111–1112
phagocytosis, 1112–1114
poly(lactide-co-glycolide) biocompatibility, 2265–2267
wear debris, 3067–3068
- Formaldehyde, sterilization of medical devices, 2506
- Formulation factors, bioadhesives, 156–157
- Forward simulation, gait analysis, 1138
- Fourier Slice Theorem, microcomputed tomography, 1853
- Fourier transform (FT)
digital signal acquisition and processing, wavelet analysis, 854–5
nuclear magnetic resonance, pulsed Fourier transform spectroscopy, 1963–1964
- Fourier transform infrared (FTIR) spectroscopy
bone mineral composition, 2042
hydroxyapatite characterization, 1459
- Fracture fixation
bioabsorbable plates and screws
bone graft fixation, 436
cancer surgery, 436
clinical applications, 434–437
experimental techniques, 434
future research issues, 437
orthognathic surgery, 434–435
overview, 432–434
pediatric surgery, 436–437
trauma surgery, 435–436
bone healing mechanisms and, 414
corrosion analysis, 755–757
degradable polymer composites, 779–780
external skeletal fixation, 419–420
fracture mechanisms and, 414
internal fixation, 414–418
intramedullary nailing, 416–418
materials, 414
orthopedic devices, 1824–1827
absorbable devices, 2073
dynamic compression plate, 1829–1832
external devices, 2077–2078
failure analysis, 1827–1839
internal devices, 2073–2077
intramedullary nails, 2076
plates, 2075–2076
rods, pins, and wires, 2074–2075
screws, 2075
soft tissue-bone attachment, 2076–2077
overview, 413
poly(lactide-co-glycolide) biocompatibility, 2265–2267
reconstructive biomaterials
craniofacial trauma, 2399–2400
facial skeleton, 2396
fronto-orbital fractures, 2399
plate selection, 2398
zygomatic and midface fractures, 2398–2399
spinal fracture biomechanics, 2468–2469
- Fracture mechanics
ceramics, 586–587
pyrolytic carbon, 2376–2377
stress fractures
diagnostics, 2514–2516
differential diagnosis, 2516
epidemiology, 2511–2512
femur, 2518–2519
lumbar spine, 2519–2520
metatarsals, 2517–2518
pathophysiology, 2513–2514
predisposing factors, 2512–2513
properties of, 2511
tibia, 2517
treatment, 2516–2517
ultrasound therapy, 2952–2959
- Frames, orthopedic fixation, 2077–2078
- Free electrophoresis, electrophoretic cell separation, 1699
- Free energy adsorption, single-component protein solutions, 274–276
- Freehand neuroprosthetic system, 1927–1928
- Free induction decay (FID), nuclear magnetic resonance, 1963–1964
- Free radicals
biodegradable polymers, 203–204
photopolymerization mechanisms, 2151–2152

- Freeze drying
 biocomposite fabrication, 187–188
 microporous materials, 1887
 tissue engineering scaffolds, 2831–2832
- Freeze-gelation, microporous materials, sol-gel bioactive glasses, 1888
- Freeze spray atomization, polymer particulates, controlled-release drug delivery systems, 882
- Frequency-domain analysis
 cancer diagnosis and therapy, 2012
 light propagation in tissue, 2033–2034
- Frequency ranges
 electrical impedance imaging, 917–918
 electrogastrography imaging, 953–954
 magnetic resonance imaging, 1716
 radiofrequency electromagnetic field/radiation (RFR) effects, 254–255
 tissue dielectric properties, 839–841
- Frequency response analyzers (FRA), electrical cell impedance sensing, 909–910
- Frequency transposition, hearing mechanisms and, 1278
- Fresnel's equation, tissue optics, 1623–1625
- Fretting corrosion, implant materials, 755
- Friction properties
 alumina, 18–19
 suture materials, 2570–2574
- Frog glue, 245
 wound healing and tissue repair, 248–249
- Fronto-orbital fractures, reconstructive biomaterials, 2399
- Fruit fly, larval glue, 246
- Fugitive phase removal, polymer foam processing, 2271–2273
- Full-width-at-the-half-maximum (FWHM), optical coherence tomography, 1998–1999
- Functional electrical stimulation (FES), neuroprostheses, 1924
- Functional magnetic resonance imaging (fMRI)
 brain activation localization, 1121
 brain mapping, 1119–1120
 brain organization and function, 1121–1123
 clinical applications, 1121–1124
 disease and treatment-related functional changes, 1125–1126
 mental disorders diagnosis, 1123
 neurosurgical planning, 1123–1125
 pharmacological effects measurement, 1126
 development, 1119
 disease-related functional changes, 1125–1126
 future applications and developments, 1126–1127
 integrated MRI, 1126–1127
 mental disorders analysis, 1123
 neurological applications, 1121–1127
 neurosurgery planning, 1123–1125
 overview, 1119
 pharmacological effects analysis, 1126
 quantitative analysis, 1126
 very high-field MRI scanner, 1126
- Functional secretory domain, osteoclasts, 2082
- Functional testing
 organ systems, 1751
 photopolymerized biomaterials, 2157–2158
 in vivo biocompatibility testing, 173
 vascular prosthetics, animal models, 1567
- Fungal adhesives
 aquatic sources, 244
 wound healing and tissue repair, 248
- Fung exponential, heart biomechanical modeling, 1287
- Fused deposition modeling (FDM)
 biocomposite fabrication, 191
 rapid prototyping, 2387–2388
- G**
- Gait analysis. *See also* Ergonomics; Movement control
 data acquisition and analysis, 1129–1130, 1134–1136
 ergonomics, 1000–1005
 kinematics, 1131–1133
 kinetics, 1133–1134
 modeling optimization, 1137–1138
 movement control, 731–732
 nerve tissue engineering, 2813–2814
 neuroprostheses, 1926–1927
 overview, 1129
 temporal-spatial parameters, 1130–1131
 terminology, 1129–1130
 testing protocols, 1134
 total hip arthroplasty results, 1386
- Galactosylated chitosan-graft-dextran (GCD), structure and properties, 611
- Galerkin equations, electrical impedance imaging, 920–923
- Galvanic corrosion, implant materials, 755–756
- Galvanometry, electrocardiography and, 936–937
- Galyfilcon silicone hydrogel contact lenses, 726–727
- Gamma radiation
 surface modification, 2541–2543
 ultra-high molecular weight polyethylene, 2925–2930
- Gamma-structure, collagen, 628–629
- Ganciclovir reservoir ocular implant, 1988–1989
- Gantry assembly, positron emission tomography, 2316
- Gap guidance, cell surface topography, 2557–2558
- Gas antisolvent (GAS), basic principles, 2323–2324
- Gas diffusion, gradient surfaces, 1211–1212
- Gas-exchange parameters, artificial lung systems, 1661–1663
- Gas fiber optic biosensors, 1069
- Gas foaming
 biocomposite fabrication, 186
 polymer foams, 2271
 supercritical fluid synthesis, 2527
 tissue engineering scaffolds, 2831
- Gastric pacemaker potentials, electrogastrography and, 947–949
- Gastrointestinal physiology, electrogastrography and, 947–949
- Gaussian function, microparticle image velocimetry, particle detection, 1879–1881
- Gauss quadrature, finite element analysis, 1095
- Gaze-contingent display paradigms, eye movement research, 1031
- GDP-mannuronic acid, in alginates, 350–351
- Gelatinases, structure and function, 1768–1769
- Gelatin-resorcinol-formaldehyde (GRF) glues
 albumin glue, 1190
 NHS-PLGA compounds, 1191
 overview, 1185
 structure and properties, 1189–1190
- Gel injection adjustable keratoplasty (GIAK), corneal implants, 750
- Gellan, hydrogels, 1446–1447
- Gels. *See also* Hydrogels
 burn dressings, 458
 chitosan, 605–606
 controlled-release drug delivery systems, 882–883
 electro-responsive gels, 932
 fibroblast-populated collagen gels, mechano-regulation mechanisms, 1784–1788
 interpenetrating polymer networks, 1551–1553
 microporous materials, 1889–1890
 tissue engineering, skeletal muscle scaffolds, 2841–2842
 valvular prosthetics, collagen fiber bundles, 2744
- Gender, stress fractures and, 2511–2512
- Gene-activated matrix (GAM)
 applications, 1142–1143
 basic properties, 1142
 biodegradable scaffolds, 1145–1146
 tissue regeneration, 1146–1148
 vector systems, 1144–1145
 wound healing, 1143
- Gene expression
 alginate biosynthesis, 351–353
 biofilm formation, 211–212
 cell surface topography, contact guidance, 2558
 elastin materials, 891
 fibroblast mechano-regulation, 1783–1788
 magnetic resonance microscopy, 1744–1745
- Gene gun delivery system, gene therapy, 1151–1153
- Gene therapy
 applications, 1150
 host response mechanisms, 1159
 chitosan, 607–611
 disease domains, 1150–1151
 gene-activated matrix
 applications, 1142–1143
 basic properties, 1142
 biodegradable scaffolds, 1145–1146
 tissue regeneration, 1146–1148
 vector systems, 1144–1145
 wound healing, 1143
 host reactions, 1157–1160
 adaptive immunity, 1158
 cancer, 1158–1159
 clinical trials, 1158
 genetic disease, 1159–1160
 immune responses, 1157
 innate immunity, 1157–1158
 laser-tissue interactions, 1626
 nonviral vectors, 1151–1154
 electroporation, 1153
 gene gun delivery, 1151–1153
 liposomes, 1153
 naked DNA, 1151
 synthetics, 1153–1154
 overview, 1150
 polycaprolactones, 2203–2204
 tissue engineering, 2623
 peripheral nerve cells, 2817
 tendons, 2874
 viral vectors, 1154–1157
 adeno-associated virus, 1154–1155
 adenovirus, 1156–1157
 herpes virus, 1155–1156
 retrovirus, 1155
- Genetics
 cell modification, tissue engineering, bone, 2654
 engineering
 insulin releasing cells, artificial pancreas, 2806–2807
 organ transplants, 2910–2912
 magnetic resonance imaging, phenotyping, 1726

- Genetics (*cont.*)
 osteoclasts, bone regeneration, 2095–2096
 radiofrequency electromagnetic field/
 radiation effects and, 257–258
- Gene transfer, nitric oxide materials, 1956
- Genipin compounds, collagen cross-linking,
 648–649
- Genomics, chondroitin sulfate research and,
 626
- Genotoxicity, biocompatibility testing,
 172–174
- Gentamicin, as PMMA cement additive, 65
- Geometric properties
 atrioventricular valve mechanics, 1324–1326
 finite element analysis, 1091–1092
 real-time tomographic reflection
 calibration, 2392–2392
 three-dimensional imaging techniques,
 1799–1800
- Giant cell mechanisms, foreign body response,
 1113–1114
- Gibbs equation, interfacial transport, 2349
- Glasses and glass ceramics. *See also* Ceramics
 bioactive materials
 clinical applications, 138
 composition and reactivity, 135–136
 future research issues, 139–140
 mechanical properties, 137–138
 mechanisms in, 134–135
 overview, 133
 soft tissue bonding, 136–137
 synthesis and structure, 133–134
 tissue attachment and interfacial
 strength, 136
 tissue engineering scaffolds, 138–139,
 144–150
- dental materials
 Dicor/Dicor MCG, 598
 glass-infiltrated cores, 598
 leucite-reinforced ceramics, 598
- melt spinning, 1818–1821
 bioabsorbable composite reinforcement,
 1822–1823
 biomedical applications, 1822–1823
 compositional dependence, 1818–1819
 fiber drawing, 1820
 glass fiber drawing, 1820
 machinery, 1820
 porous melt-derived glasses, 1891
 sintered porous scaffolds, 1822
 structure and strength, 1820–1821
 multilayer assemblies, polyelectrolytes, 2221
- orthopedic devices, 2066–2068
 composites, 2069
 osteoinductive substances and materials,
 2104
- processing, 585–586
- vitrification
 cartilage, 3050–3051
 ceramics, 585–586
 cryoinjury prevention, 766
 future research issues, 3053–3054
 overview, 3046
 physiochemistry, 3046–3049
 stability properties, 3049
 tissue engineering, 3049–3050
- Glass transition temperature, polymer
 structure/property relationships, 2276
- Glaucoma filtration implants (GFIs),
 1976–1977
- Glial tissue
 hyaluronan applications, 1429
 nerve tissue engineering, 2815–2816
- Glioblastoma multiforme (GBM), nanoscale
 neurosurgery, 1903
- Global flow mapping, positive displacement
 pumps, 102–103
- Global stiffness matrix, composites,
 continuum mechanics, 665–666
- Glucosamine, osteoarthritis efficacy,
 1204–1205
- Glucose monitoring
 biosensor technology, 335–337
 sensors
 blood glucose sensors, 1180–1182
 clinical applications, 1174–1175
 insulin delivery systems and, 1518
 interstitial fluid sensors, 1175–1179
 affinity chemical/optical sensor, 1180
 dialysis catheter monitoring system,
 1178–1180
 enzyme-based telemetry sensor, 1180
 needle-type, enzyme-based
 electrochemical biocatalytic sensors,
 1175–1178
 overview, 1174
- Glucose oxidase (GOD)
 hydrogel materials, 1163–1172
 silk biosensors, 2445–2446
- Glucose-responsive hydrogels
 drug delivery systems, 1163–1164
 environmentally-sensitive compounds,
 1164–1165
 future research issues, 1170–1171
 insulin delivery systems, 1165–1170
 overview, 1163
- Glues. *See also* Sealants
 albumin glue, 1190
 aldehyde systems, 1189–1190
 azide chitosan lactose, 1192
 cyanoacrylate, 1187–1189
 fibrin glue, 1185–1187
 future research issues, 1192–1193
 gelatin-resorcinol-formaldehyde formula,
 1189
 photocurable systems, 1191–1192
 surgical applications, 1185
 WSC and NHS system, 1190–1191
- Glutamate dehydrogenase (GDH)
 interfacial transport, 2350–2358
 surfactant interactions, 2340–2343
- Glutamic acid target sequence, integrin-
 binding, 1527–1528
- Glutaraldehyde (GA)
 artificial organ cross-linking, 2231
 collagen tissue fixation, 644–645
 polyepoxy compounds *vs.*, 2232–2235
 sterilization of medical devices, 2506–2507
 valvular prosthetic failure
 calcification, 76
 cross-link chemistry, 1306–1308
 historical review, 1294–1295
 inflammatory degradation, 1298–1299
- Gluteus medius syndrome, overuse injury,
 1397
- Glycation, collagen fixation, 647–648
- Glycoconjugates, hydrogels, 1447
- Glycodendrimers, dendrimer-drug delivery
 systems, 793
- Glycolic acid and glycolides
 biodegradable polymers, 196–199
 copolymers, orthopedic devices, 2064
- Glycosaminoglycans (GAGs)
 aortic valvular cusps, microstructure/
 micromechanics, 2738–2742
 articular cartilage repair, 559
 biosynthesis, 1199–1200
 chondroitin/sulfate/dermatan sulfate
 compounds, 618–619
 content measurement, 1201–1202
 disorders, 1202–1204
 extracellular matrix, scaffold materials,
 1021–1025
 future research issues, 1205
 magnetic resonance imaging, 1202
 osteoarthritis, 1204–1205
 overview, 1195
 semilunar aortic valve structure, 1318–1319
- structural properties, 1195
 tissue engineering, 2291–2292
 liver, 2772
 skin, 2853–2857
 valvular prosthetics
 allografts, 11–13
 calcification, 75
 matrix mechanics, 2742–2743
- Goal-directed movement, mechanics of,
 731–732
- Gold alloys
 dental ceramics, porcelain-bonded metal,
 597
 gold-palladium alloy, membrane coatings,
 permeability, 2354
 ions, antimicrobial biomaterials, 66–67
 multilayer assemblies, polyelectrolytes, 2221
 self-assembled monolayers, metal-
 alkanethiol interactions, 2409–2411
- Good Laboratory Practices Act, 1621
- Graded-index (GRIN) lens, optical coherence
 tomography, 2001–2002
- Gradient-based learning, feed-forward neural
 networks, 123–124
- Gradient surfaces
 biofunctional groups, 1213–1216
 biomolecular interactions, 1219–1225
 cellular interactions, 1222–1225
 characterization, 1216–1219
 future research issues, 1225–1226
 interpenetrating polymer networks, 1549
 magnetic resonance imaging, 1716
 overview, 1210–1211
 protein interactions, 1220–1222
 wettability surfaces, preparation, 1211–1213
- Graft materials and techniques
 acellular grafts
 nerve tissue engineering, 2815
 valvular prosthetics, anticalcification
 treatments, 81
 articular cartilage repair, 554
 bioabsorbable plates and screws, 436–437
 burn injuries, skin substitutes, biological
 and synthetic, 451–456
- collagen fixation
 overview, 639
 tissue grafts, 641–644
- copolymers, 2293–2294
- distal protection devices
 AccUNET device, 875–876
 advantages/disadvantages, 877
 Angioguard device, 873–874
 carotid artery, 870–871
 current research evolution, 877–878
 embolization reduction devices, 871
 failure mechanisms, 877
 Filterwire Ex, 875
 future applications, 878
 NeuroShield device, 874–875
 overview, 870
 Parodi Anti-Emboli System, 876–877
 Percutaneous guardwire, 871–874
 saphenous vein grafts, 870
- endovascular stent grafts
 balloon-expanded stents, 2486
 bifurcated grafts, 2482
 branching techniques, 2480–2482
 breakage issues, 2486–2487
 delivery system diameter, 2489–2490
 design innovation, 2484
 fabrics, 2484–2485
 fenestrated grafts, 2483
 fixation techniques, 2487–2488
 future designs, 2487
 medical applications, 2480
 modularity, 2483–2484
 multibranched grafts, 2482–2483
 sealing techniques, 2488–2489
 self-expanding stents, 2486

- support systems, 2489
 - unbranched grafts, 2480–2482
 - fat tissue engineering, 2725–2734
 - fibrin sealants, 1085–1086
 - gene-activated matrix regeneration, 1146–1148
 - glucose hydrogels, 1167–1170
 - hemocompatibility, 1347–1348
 - hybrid vascular prostheses, 1432–1437
 - hydroxyapatite, 1451–1460
 - insulin delivery systems, 1518–1519
 - ligament injury and repair, animal repair models, 1643–1647
 - nerve repair, 2125–2126
 - alternative materials, 2127–2128
 - donor site morbidity, 2126–2127
 - phosphorylcholine technology, 2142–2145
 - plasma modification, 2182–2183
 - polycaprolactone, 2202–2203
 - polymers, surface modification with, 2545
 - supercritical fluids, 2528–2529
 - tissue engineering, cardiac muscle, 2688–2692
 - vascular grafts, 2997–2998
 - basic properties, 2989–2993
 - biocompatibility, 2990–2991
 - bioprosthetics, 2985–2987
 - clinical applications, 2985
 - endothelial cell techniques, 2992–2993
 - healing mechanisms, 2991–2992
 - host reactions, 3006–3009
 - mechanical properties, 2989–2990
 - overview, 2985
 - porosity, 2989
 - single-stage venous EC seeding, 2992
 - synthetic grafts, 2987–2989
 - tissue engineering, 3033–3042
 - in vitro tissue engineering, 2992–2993
 - in vivo tissue engineering, 2993
 - xenografts
 - cellular and tissue grafts, 3099–3100
 - ethical issues, 3100
 - future research projections, xeno–9
 - historical background, 3093–3094
 - immune response, 3095–3100
 - cell-mediated rejection, 3099
 - chronic rejection, 3099
 - organ transplants, 3095–3097
 - vascular grafts, 3097–3098
 - overview, 3093
 - physiology, 3095
 - sources, 3094–3095
 - zoonosis, 3100
 - Graft *versus* host disease (GVHD)
 - allografts, 9
 - tissue engineering
 - blood vessels, 2649–2650
 - small intestine, 2759
 - toxicity and hypersensitivity reactions, 1418
 - Gramicidin, biosensor development, 342–343
 - Gram-negative bacteria, biofilm formation, 209–210
 - Gram-positive bacteria, biofilm formation, 210–211
 - Granular implants, hydroxyapatite, 1456–1460
 - Granulation tissue, host reactions, 1412
 - Granulocytes, neutrophil formation, 1934
 - Graphical analysis, electromyographic signals, 965
 - Grasping tasks
 - mechanics of, 731–732
 - neuroprostheses, 1927–1928
 - Griffith's flaws, dental ceramics, 596
 - porcelain-bonded metal, 597
 - Gripping mechanisms, bioreactor design, 322–323
 - Groin pain syndrome, overuse injuries, 1393–1395
 - Grooved-bed perfusion systems, bone marrow tissue engineering, 2672
 - Ground loop elimination, biopotential amplifiers, 306
 - Growth factors
 - articular cartilage repair, 557
 - bone healing, fracture callus signaling, 1259
 - connective tissue hemostasis, 1250–1251
 - extracellular matrix, scaffold materials, 1023–1025
 - fibrin structure and biomechanics, 1076
 - fibroblast mechano-regulation mechanisms, 1789–1790
 - gene-activated matrix, 1142–1143
 - histiogenesis, 1407
 - matrix metalloproteinases, 1764–1765
 - nerve guide technology, 1917
 - osteogenic progenitor cell differentiation and proliferation, 2094–2095
 - pharmacological activity, 221
 - tissue engineering
 - articular cartilage, 2702–2703
 - blood vessels, 2646
 - bone, 2656–2657
 - bone marrow, 2662–2663
 - fat tissue, 2730
 - liver, 2772
 - skeletal muscles, 2845–2846
 - tendons, 2874
 - vascular system, 3040–3041
 - valvular prosthetics, pannus formation, 1300–1301
 - vascular graft development, 2998
 - tissue engineering, 3023–3024
 - wear debris, joint arthroplasties, 3077–3079
 - Growth modeling
 - blood vessel mechanics, 397
 - cardiac biomechanics, 1287–1288
 - tissue engineering, skeletal muscle scaffolds, 2842–2843
 - Growth plate formation, bone healing, 1258–1259
 - Gruneisen coefficient, laser-tissue interactions, 1629
 - Guided systems, surgical robotics, 1809–1811
 - Guided tissue regeneration (GTR), histiogenesis, 1407–1408
 - Guide to the Care and Use of Experimental Animals*, 1619–1620
 - Gyromagnetic ratio, MRI thermometry, molecular diffusion coefficient, 1728–1729
- ## H
- Haber-Weiss reaction, biodegradable polymers, free radical reactions, 203–204
 - Hagen-Poiseuille equation
 - dialysis membranes, 1585
 - macrocirculatory hemodynamics, steady flow properties, 1359–1361
 - Hair cells, inner ear, 1273
 - Half-threaded pins, external fracture fixation, 420
 - Hamstring syndrome, overuse injuries, 1399–1400
 - Handling properties, suture materials, 2570–2574
 - Handmaster neuroprosthetic system, 1928–1929
 - Haptics, surgical training robotics, 1811
 - Hardness properties, pyrolytic carbon, 2376
 - Hard tissue. *See also* Bone adhesives
 - basic properties, 2
 - gene-activated matrix scaffolds, 1146
 - biomaterial interactions
 - bioceramics, 1231–1232
 - future research, 1233
 - metals, 1229–1231
 - overview, 1229
 - polymers, 1232
 - wear debris, 3058–3063
 - elastic properties
 - anisotropy, 1241–1242
 - behavioral modeling, 1242–1243
 - future research, 1245
 - overview, 1236
 - physical principles, 1236–1241
 - Reuss orthotropic modulus, 1246
 - Reuss transverse isotropic modulus, 1245
 - Voigt orthotropic modulus, 1246
 - Voigt transverse isotropic modulus, 1245
 - optical coherence tomography, 2004–2005
 - photopolymerized biomaterials, 2159–2160
 - Hardware requirements
 - electrical impedance tomography, 916–918
 - telemedicine systems, 2583–2584
 - Hashin's composite sphere model, composite micromechanics, 666–667
 - Head-mounted displays (HMD), real-time tomographic reflection, 2391
 - Healing mechanisms
 - bone healing process, 413–414
 - overview, 1255
 - primary/secondary bone formation, 1258–1259
 - soft callus formation, 1256–1258
 - tissue engineering, 1260–1261
 - ultrasound therapy, 2151–2159
 - burn dressings, 450
 - connective tissues, 1253–1254
 - ligament injury and repair, animal repair models, 1644–1647
 - vascular grafts, 2991–2992
 - host reactions, 3008–3009
 - Health Research Extension Act, 1618–1619
 - Hearing aids. *See also* Cochlear implants; Inner ear implants
 - advanced processing algorithms, 1268–1269
 - basic properties, 1263
 - battery system, 1493–1495
 - cochlear implants
 - basic principles, 1507–1508
 - clinical issues, 1512–1513
 - electroneural interface, 1511–1512
 - normal hearing mechanisms, 1507
 - overview, 1507
 - receiver-stimulator implementation, 1510–1511
 - speech processor implementation, 1508–1510
 - current designs and styles, 1266–1267
 - future research issues, 1269
 - hearing loss, 1264–1265
 - hearing mechanisms and, overview, 1271–1278
 - historical background, 1265–1266
 - linear amplification, compression, and feedback suppression, 1267–1268
 - normal hearing mechanisms, 1263–1264
 - usage guidelines, 1269
 - Hearing mechanisms. *See also* Conductive hearing loss; Sensorineural hearing loss audibility or loudness correction, 1273–1275
 - basic principles, hear mech–1–3
 - brain aids to, 1277
 - cochlear implants, 1507
 - ear anatomy, hear mech–2–3
 - future research on, 1277–1278
 - hearing aid design and, 1263–1264
 - inner-ear biology and neurology, 1275–1277
 - robotics applications, 1812
 - robotics systems and, 1812–1813

- Hearing mechanisms (*cont.*)
 thresholds, 1264–1265
 tinnitus
 acoustic masking, 2610–2615
 neural plasticity and cortical mapping, 2617
 pathology, 2610
 transdermal electrical stimulation, 2612
 ultrasonic hearing, 2615–2616
 ultrasonic hearing, 2615–2616
- Heart anatomy and function. *See also* Cardiovascular disease; Valvular anatomy and function
 animal models, 1566–1571
 atria, 480–481
 electrophysiology, 488
 atrio-ventricular node, 481–482
 electrophysiology, 488
 biomechanics
 cardiac muscle mechanics, 1282–1286
 force-length curves, isolated cardiac muscle, 1282–1283
 future research issues, 1288–1289
 global/regional deformation measurement, 1286–1287
 growth and remodeling, 1287–1288
 myocardial material properties, 1284–1286
 myocardial stresses, 1283–1284
 pressure-volume curves, 1280–1282
 pump functions, 1280–1282
 structural constitutive modeling, 1286–1287
 ventricular function curves, 1280
 cardiac action potential, 482–487
 cardiac assist devices, 470
 cardiac conduction physiology, 532–533
 cardiac muscle mechanics, 1282–1284
 cardiac output measurement
 blood volume, 514–518
 dilution methods, 517–518
 distribution of, 517
 Fick method, 517–518
 human output ranges, 516
 impedance method, 526–530
 indicator-dilution method, 518–522
 overview, 514
 stroke volume calculation, 527–528
 thermal-dilution method, 522–526
 verification studies, 528–530
 cardiac patch engineering, 542
 cardiovascular data, 516
 circulatory system, 515
 electrical tissue stimulation, 1017–1018
 electrocardiography
 heart block, 944
 mechanical activity, 940
 electrophysiology, 487–489
 cellular activation, 937–939
 overview, 480
 extracellular matrix structural and functional arrangement, 545
 fibroblast mechano-regulation, 1784–1787
 His-Purkinje system, 482
 electrophysiology, 488–489
 mathematical modeling, 1895–1901
 optical mapping, 2025–2026
 overview, 1280
 oxygen transport, 1843–1844
 pulmonary circuit, 517–518
 sino-atrial node, 480
 electrophysiology, 487–488
 tissue engineering
 cell properties, 2689–2690
 cell seeding, 2686–2687
 clinical applications, 2691–2692
 functional requirements, 2690–2691
 liquid cell matrices, 2687
 methods, 2686–2688
 monolayer stacking, 2687–2688
 polycaprolactone, 2203–2204
 scaffold materials, 2688–2689
 solid matrices, 2686–2687
 in vivo applications, 2691
 two-dimensional strain imaging, 506–512
 ventricle system, 482
 electrophysiology, 480–489
- HeartMate III centrifugal blood pump, 579
- HeartMate Implantable Pump, 1777–1779
- HeartMate Vented Electric (VE) ambulatory implantable left ventricular assist device, 1777–1779
- Heart rate (HR)
 cardiac output measurement, 514
 digital signal acquisition and processing, 860–861
 detrended fluctuation analysis, 864–865
- Heart valves. *See* Valvular anatomy and function; Valvular prosthetics
- Heat transfer
 cardiac output measurement, 525–526
 electrophoretic cell separation, 1701–1702
 vitrification technology, 3054
- Hematocrit
 cardiac output measurement, 516
 microcirculatory oxygen transport, 1845
- Hematopoietic stem cells (HSCs)
 biomaterials and research on, 2475–2476
 bone marrow tissue engineering
 cell culture parameters, 2662–2665
 culture media, 2662
 culture systems
 perfusion cultures, 2671–2672
 static cultures, 2670–2671
 stirred suspension cultures, 2672
 stromal support, 2670
 three-dimensional cultures, 2672–2675
 feeding schedule and culture duration, 2664
 growth factors, 2662–2663
 hematopoiesis, 2660–2662
 inoculum density, 2663–2664
 oxygen tension, 2664–2665
 pH levels, 2664
 cell source, 1340
 enrichment, 1340–1342
 flow cytometry, 1104–1109
 future potential of, 1342
 histogenesis, 1407–1408
 mesenchymal stem cells, 1341
 phenotype, 2473
 progenitor cell assays, 1338–1340
- Hemocompatibility
 active materials, 1346–1348
 inert materials, 1345–1346
 overview, 1345
 phosphorylcholine technology, 2141–2145
 platelet interaction, 1353
 protein adsorption, 1350–1353
 surface modification strategies, 1345–1348
 surface properties analysis, 1348–1350
 testing for, 171–172
- Hemodialysis, 378–384
 biocompatibility, 1583
 body chemistry and, 380–381
 cellulosic membranes, 1583–1584
 classification, 1580–1583
 concentration gradient, 379–380
 dialysis membranes, 1583–1585
 efficiency, 1581–1583
 flux changes, 1581
 historical perspective, 1578–1579
 hollow-fiber membranes, 1579–1580, 1585
 mass transfer and tissue function, 1754–1756
 pharmacokinetic approximations, 1755–1756
- membrane biocompatibility, 1583–1585
 molecular permeability, 380
 overview, 1576
 performance evaluation, 382–384
 principles of, 378–379
 synthetic membranes, 1584–1585
 uremic solute classification, 1577–1578
 waste product removal, 379
 water removal, 381–382
- Hemodynamics
 cardiac overload, 1287–1288
 macrocirculatory
 arterial pulse waves, 1364–1365
 blood viscosity, 1357–1359
 circulation properties and mechanics, 1356–2
 fluid dynamics, 1359–1366
 overview, 1356
 pathology, 1366–1367
 pulsatile flow mechanics, 1361–1364
 steady flow mechanics, 1359–1361
 venous hemodynamics, 1365–1366
 vessel wall properties, 1359–1360
 mechanical heart valves, 1332–1334
 positron emission tomography, 2319–2320
 valvular structure and function
 atrioventricular valve, 1323
 semilunar valve, 1319
- Hemofiltration, basic principles, 384
- Hemoglobin-based oxygen carriers (HBOCs), microcirculatory oxygen transport, 1847
- Hemolysis testing
 cardiac assist devices, positive displacement pumps, 96
 centrifugal blood pumps, 579–583
- Hemophilia B, gene therapy, 1154–1155, 1159
- Hemostasis
 blood-material interactions, 367–370
 connective tissue healing, 1250–1251
 fibrin glue, 1186–1187
 fibrin structure and biomechanics, 1073–1077
 polyepoxy compounds, 2238
- Henry's law constant, mass transfer and tissue function, 1756
- Henry-Greenstein phase function, biomedical optics, 2031
- Hepaface coating, heparin biocompatibility, 1372
- Heparan sulfates, glycosaminoglycans, 1195
- Heparin
 artificial lung systems, cardopulmonary bypass, 1668–1669
 biocompatibility
 Bioline coating, 1372
 Carmeda Bioactive Surface, 1371
 catalytic role, 1368
 clinical applications, 1372
 complement and inflammatory response, 1370
 contact activation, 1370
 Corline heparin surface, 1371
 Duraflo surface, 1371–1372
 enzymatic coagulation interaction, 1370
 future research, 1373
 Hepoface coating, 1372
 immobilized biological effects, 1369–1370
 overview, 1368
 Pharmacia HSM intracular lenses, 1371
 platelet adhesion/activation, 1369
 protein interactions, 1370
 T-NCVC coating, 1372
 Trillium biopassive surface, 1371
 vascular endothelial anticoagulant activity, 1368–1369
 white blood cell adhesion/activation, 1369

- cardiac catheterization coatings, 497–498
 dermatan sulfate interaction, 623–624
 hemocompatibility, 1346–1348
 as pharmacological active polysaccharide, 220
 polyepoxy compounds, vascular prostheses, 2235–2236
- Hepatic surgery, fibrin sealant applications, 1086
- Hepatitis C virus (HCV), tissue engineering applications, 2776
- Hepatocytes
 artificial liver systems
 cell-cell interactions, 1654
 cell source and replication, 1651–1652
 culture media, 1655–1656
 extracellular matrix and cell polarity, 1653–1654
 optimization and preservation, 1652
 scale-up cultures, 1656
 tissue engineering principles, 1652–1654
 transplantation techniques, 1650
 tissue engineering, liver
 interactions, 2770
 sources, 2770–2771
 structure and function, 2769–2770
- Herpes simplex viruses (HSV 1/2)
 gene-activated matrix, 1145
 gene therapy, 1155–1156
- Hessian matrices, feed-forward neural networks, 123
- Heterogeneous mechanics, composites, 660–661
- Heteronuclei, magnetic resonance thermometry and, 1734–1735
- Hierarchical hybrid movement control model, 735–737
- High Aspect Ratio Vessel (HARV), rotary cell culture systems, 2404
- High audio bone vibration maskers, tinnitus, 2614–2615
- High-efficiency hemodialyzers, 1581–1582
- High-flux hemodialyzers, 1581
- High gradient magnetic separations (HGMSs), magnetic cell separation, 1695–1696
- High-magnification endoscopy, basic principles, 993–994
- High-molecular weight (HMW) copolymers. *See also* Ultrahigh molecular weight polyethylene (UHMWPE)
 poly(lactide-*co*-glycolide), 2259–2261
- High molecular weight kininogen (HMWK), adsorption behavior, 277–278
- High-temperature superconducting (HTS) coils, magnetic resonance microscopy, 1740–1741
- High-throughput technology
 biofilm formation, 215
 flow cytometry, 1108
- Hip joint. *See also* Total hip arthroplasty (THA)/total hip replacement (THR)
 biomechanics
 arthritis, normal and artificial hips, 1377
 cup stability, micromotion, fixation, and osteolysis, 1380
 femoral head subluxation, 1381–1382
 finite element modeling and validation, 1381–1385
 future research issues, 1388–1390
 gait loading results, 1386
 implant fixation techniques, 1379–1380
 liner micromotion, 1385–1386
 pelvic osteolysis and adjunctive screws, 1380–1388
 post THA dislocation, 1381
 subluxation loading results, 1386–1388
 total hip replacement techniques, 1377–1379
 normal anatomy, 1377
 overuse injuries
 external rotator syndrome, 1397
 gluteus medius syndrome, 1397
 groin pain syndrome, 1393–1395
 hamstring syndrome, 1399–1401
 iliopsoas tendinitis/bursitis, 1395–1396
 overview, 1393
 rectus femoris tendinitis, 1396–1397
 snapping hip syndrome, 1397–1399
 prosthetic devices
 corrosion, 755–756
 failure analysis, 1839–1841
 wear debris, 3059–3060
- Hirudin
 hemocompatibility, 1346–1348
 pharmacological activity, 220
- His-Purkinje system, electrophysiology, 482, 488–489
- Histiogenesis
 basic principles, 1403
 embryonic, 1405–1406
 engineered procedures, 1407–1408
 hierarchy, 1403–1404
 postnatal, 1406–1407
 terminology, 1403
 tissue complexity and variability, 1404–1405
- Histology
 ligament injury and repair, animal repair models, 1645–1647
 magnetic resonance microscopy, 1742–1744
 optical coherence tomography, 2004–2005
 reconstructive biomaterials, 2398
- Histomorphometry, dental implants, 816
- Hoffa disease, overuse injuries, 1603–1604
- Hollow-fiber systems
 artificial lung systems, 1662–1666
 bio-artificial liver systems, 1656–1657
 dialysis membranes, 1579–1580, 1585
- Hollow support structures, fat tissue engineering, 2732–2733
- Holmium-YAG laser, transmyocardial laser revascularization, 1633–1637
- Homogeneous precipitation, magnetic nanoparticles, 1706
- Homogenization, collagen processing, 656
- Homografts. *See* Allografts
- Hooke's law
 elastic properties, 1236–1241
 hard tissue viscoelasticity, 1243–1244
 vascular fluid dynamics, 2973–2975
- Hormones
 breast tissue engineering, 2684
 fibroblast mechano-regulation, 1783–1788
 protein-based drug delivery systems, 2362–2363
- Host/guest supramolecular gels, basic properties, 1448–1449
- Host reactions. *See also* Foreign body response; Immune response
 bacterial vs. tissue colonization, 1414–1415
 blood-material interactions, 1415–1416
 complement system, 1412–1414
 extracellular matrix, scaffold materials, 1023–1025
 foreign body response, 1412
 fibroblasts, 1114
 giant cell response, 1114
 mechanisms of, 1115–1116
 overview, 1111–1112
 phagocytosis, 1112–1114
 gene therapy, 1157–1160
 granulation tissue, 1412
 immune system, 1416–1418
 inflammation, 1411–1412
 intravascular ultrasound imaging, 1559–1564
 medical device infections, 1499–1500
 orthopedic devices, particulate matter
 aseptic loosening, 2110–2111
 future research issues, 2114–2115
 macrophage detectors, 2111–2114
 osteoclast/osteoblast effector cells, 2114
 overview, 2110
 overview, 1411
 poly(lactide-*co*-glycolide) biocompatibility, 2265–2267
 remodeling, 1412
 suture biocompatibility and, 2574–2575
 systemic toxicity and hypersensitivity reactions, 1418
 tumorigenesis, 1419
 urinary prostheses, 2961–2962
 valvular prosthetics, calcification, 77–79
 vascular grafts, 3006–3009
- Hot pressing techniques
 alumina joint prostheses, 20–21
 dental ceramics, 601
 wear debris, 3061–3063
- Hub and spoke topology, telemedicine applications, 2582–2584
- Human embryonic kidney cells (HEK), chitosan gene expression, 610–611
- Human-machine interface
 neuroprostheses, 1924–1925
 surgical robotics, 1807–1809
- Human umbilical vein endothelial cell (HUVEC)
 electrical cell impedance sensing, 912–913
 in vitro vascular constructs, 3024–3025
- Humoral immune response
 artificial pancreas, membrane immunoisolation efficacy, 2804–2805
 host reactions, 1417
- Hyaladherins
 cardiovascular applications, 1428
 hyaluronan role in, 1422
- Hyaline membrane disease
 lung surfactant dysfunction, 1689–1690
 therapeutic regime, 1690–1691
- Hyaluronan
 basic properties, 1421
 biomaterials applications, 1426–1429
 antiadhesion, 1427
 cardiovascular, 1427–1428
 drug delivery systems, 1429
 neural and glial applications, 1429
 ophthalmology, 1428
 orthopedic, 1427
 tissue engineering and surgery, 1429
 wound healing and dermatology, 1428
 cell degradation and, 1423
 cell migration and, 1422–1423
 cross-linked biomaterials, 1425–1426
 derivatized biomaterials, 1424–1425
 extracellular matrix and hyaladherin interactions, 1422
 future research issues, 1429
 histiogenesis, 1406
 osteoarthritis efficacy, 1204–1205
 pathological processes and, 1423–1424
 physicochemical and structural properties, 1421–1422
 valvular prosthetics, cross-linked compounds, 2744–2746
 wound repair and angiogenesis, 1423
- Hyaluronic acid (HA)
 cell-material interactions, 571
 chondroitin sulfate proteoglycans interaction, 620
 glue materials, 1192
 glycosaminoglycans, 1195
 proteoglycans, 1197–1198
- Hybrid materials. *See also* Composite materials
 hydrogels, 1447–1448

- Hybrid materials (*cont.*)
 nerve guides, 1916–1917
 polymer processing, 2281
- Hybrid mobility devices, 3089–3090
- Hydration, tendon/ligament tensile testing, 2591–2592
- Hydrocephalus, shunt infections, antiseptic/antibiotic materials and reduction of, 66
- Hydrocolloids
 burn dressings, 459
 skin adhesives, 158
- Hydrogels
 artificial pancreas, 2803–2805
 bioadhesive properties, 4–5
 biomimetic compounds, 1447–1449
 biomolecules, 1446–1447
 burn dressings, 458–459
 chemical gels, 1439–1444
 chitosan sustained release systems, 604–605
 corneal implants, 748–749
 dendrimer-base drug delivery, 793
 electroactive polymeric materials, electroact–7
 glucose-responsive hydrogels
 drug delivery systems, 1163–1164
 environmentally-sensitive compounds, 1164–1165
 future research issues, 1170–1171
 insulin delivery systems, 1165–1170
 overview, 1163
 hemocompatibility, inert materials, 1345–1346
 hyaluronan cross-linked biomaterials, 1425–1426, 1428
 interpenetrating polymer networks, 1550–1554
 nerve guides, 1917
 nitric oxide materials, 1955
 ocular implants, 1975
 orthopedic devices, 2065
 overview, 1439
 phenylboronic acid properties, 1442–1443
 photopolymerization, 2154–2158
 physical gels, 1444–1449
 polymer biomaterials, 222
 polyurethane biomaterials, 2308–2309
 protein-based drug delivery systems, 2362–2363
 quick response compounds, 1444
 silicone contact lenses
 chemical coatings, 725
 core surface treatments, 725
 current commercial sources, 725–727
 evolution of, 718–719
 hydrophilic components, 720
 hydrophobic comonomers, 721
 ion permeability and on-eye movement prediction, 722
 ocular environment, 715
 overview, 715
 oxygen permeability, 715–717
 on phases, 722–723
 plasmas, 724–725
 polymer phases, 721–722
 recent developments, 719
 research background, 717–718
 surface properties, 723–725
 transport properties, 721
 soft contact lenses, 1439–1442
 fouling effects, 1441–1442
 historical background, 1440
 materials requirements, 1442
 oxygen permeability, 1441
 silicone hydrogel, 1441
 supercritical fluid synthesis, 2527–2528
 synthetic polymers, 1445–1446
- tissue engineering
 corneal implants, 2710–2711
 photopolymerization, 2158–2160
 smart polymers, 2294
 volume phase transition, 1442–1444
 reentrant volume phase transition, 1444
- Hydrolysis
 hydroxyapatite synthesis, 1455
 polycaprolactone degradation, 2198–2199
 poly(glycolic acid) degradation, 2242–2244
 polyurethane degradation, 2306
- Hydron film, burn dressings and, 458–459
- Hydrophilicity
 adsorption behavior, single-component protein solutions, 274–276
 dialysis membranes, 1584–1585
 polyepoxy compounds, 2234
 polymers, blood-contacting materials, 375
 protein-surfactant interactions, 2341–2343
 silicone hydrogel contact lenses, 720
- Hydrophobicity
 adsorption behavior, single-component protein solutions, 274–246
 protein-surfactant interactions, 2341–2343
 silicone hydrogel contact lenses, 721
- Hydrosilation cure, silicone thermosetting elastomers, 2430
- Hydrostatic pressure, articular cartilage repair, 558
- Hydrothermal transformation, hydroxyapatite synthesis, 1455
- Hydroxyapatite (HA)
 anionic substitutions, 1452–1454
 aqueous techniques, 1455
 basic properties, 1451
 bioactive glasses, 134
 biocomposite comparisons with, 181–182
 biomedical coatings, 1456–1457
 biphasic calcium phosphate preparation, 360
 bone
 response to, 1459–1460
 tissue engineering, 2047
 bone-implant interactions, 1229–1231
 calcium phosphate bone cements, 1457
 carbonate substitutions, 1452–1454
 cationic substitutions, 1452
 ceramic implants, 1457–1459
 ceramics, 587–588
 calcium orthophosphates, 461–462
 coating materials, 589–590
 collagen composites, 466–467
 composites, 589–590
 hydroxyapatite-poly-L-lactic acid composite, 466
 Japanese production of, 462–463
 transparent ceramics, 2916–2923
 chemical synthesis, 1454–1455
 chondroitin sulfate interaction, 626
 coating materials, 1464–1468
 composites
 degradable polymer-ceramics, 783–784
 fillers, 1457
 crystallography, 1451–1452
 dense monolithic ceramics, 1456
 dental implants, surface analysis, 814
 granular implants, 1456
 hydrolysis, 1455
 hydrothermal transformations, 1455
 microporous materials, 1886–1887
 mineralized tissues, 2039
 orthopedic devices, 2067–2069
 osteoinductive substances and materials, 2105
 porous scaffolds, 1456
 precipitate, 1455
 shape memory metal coatings, 2423
 silicate substitutions, 1454
 sintering methods, 1455–1456
- solid-state reactions, 1455
 structure and properties, 225–226
 substituted apatites, 1451–1452
 supercritical fluid synthesis, 2528–2529
 synthetic composites, 229–233
 tissue engineering
 bioactive materials, 144–150
 bone, 2655
 scaffold materials, 232–233
 tooth enamel, 2048–2049
 total knee replacement, 1612–1613
- Hydroxy biofunctional groups, gradient surfaces, 1213–1214
- Hydroxylases, collagen synthesis, 632
- Hydroxylated surface, self-assembled monolayers, organosilicon derivatives, 2411–2412
- Hydroxyl radicals, biodegradable polymers, 203–204
- Hyperacute rejection, xenografts, 2908–2909, 3095–3097
- Hyperopia/hypermotropia, defined, 747
- Hypersensitivity
 biocompatibility testing, 172
 host reactions, 1418
 nickel alloys, 1948
 total hip arthroplasty materials, 1480–1481
- Hyperthermia therapy
 magnetic nanoparticles, 1708–1709
 proton resonance frequency shift, 1733–1734
- Hysteresis
 ligament and tendon viscoelasticity, 2595–2596
 piezoelectric materials, 2163–2164
- I**
- Iduronic acid, dermatan sulfate proteoglycans, 621–622
- If-Then rules, movement control models, 734
- Iliac artery, bifurcated stent-grafts, endovascular systems, 2482–2483
- Iliopsoas tendinitis/bursitis, overuse injuries, 1395–1396
- Iliotibial band
 friction syndrome (ITBFS), 1600–1602
 snapping hip syndrome, 1398–1399
- Image registration, three-dimensional imaging techniques, 1798–1799
- Image weighting, magnetic resonance imaging, 1718
- Imaging probes, optical coherence tomography, 2001–2002
- Imaging techniques. *See also* Optics, biomedical; Three-dimensional imaging techniques
 cancer diagnosis and therapy
 confocal microscopy, 2015
 continuous-wave imaging studies, 2011
 diffuse optical tomography, 2013
 fluorescence method, 2014–2015
 frequency-domain measurements, 2012
 optical coherence tomography, 2013–2014
 overview, 2009
 radiative transfer method, 2009–2014
 Raman spectroscopy and multiphoton techniques, 2015
 spectroscopy measurements, 2011–2012, 2014–2015
 time-resolved measurements, 2012–2013, 2015
- confocal microscopy, 705
 fast confocal microscopy, 712–713
 fluorophores, 711
 incident light intensity, 710–711
 living cells, 712
 multicolor fluorescence, 712

- Nipkow disk, 713
 photobleaching, 711–712
 pinhole size, 710
 principles and instrumentation, 705–706
 procedures and applications, 707–709
 resolution, 710
 three-dimensional visualization, 709–710
 two-photon microscopy, 713
- electrical impedance imaging
 hardware requirements, 916–918
 overview, 915–916
 safety requirements, 918–919
 sample images, 923–924
 software requirements, 919–923
- electrogastrography
 fasting EGG patterns, 954
 future research issues, 955
 overview and historical background, 947
 parameters, 953–954
 physiology, 947–949
 postprandial EGG patterns, 954–955
 recording procedures and equipment, 949–954
- electron microscopy
 cardiovascular tissue engineering, 974–975
 hard biomaterials and polymer specimens, 972–973
 instrumentation, 968–969
 medical applications, 974
 overview, 968
 scanning vs. transmission techniques, 969–970
 SEM specimen preparation, 970–972
 TEM specimen preparation, 973–974
- endoscopy
 autofluorescence, 994–995
 chromoscopy, 993–994
 emerging technologies, 992–997
 endocytoscopy, 994
 endoscopic retrograde cholangiopancreatography, 990
 fluorescence imaging, 994
 future applications, 997
 high-magnification endoscopy, 992–993
 image detection, 989–990
 instrumentation and components, 986–989
 narrow band imaging, 993
 non-gastrointestinal applications, 989
 OCT imaging technique, 995–996
 overview, 986
 photodynamic imaging, 995
 ultrasound and, 990–991
 in vivo confocal microscopy, 996–997
 wireless capsule techniques, 991–992
- functional magnetic resonance, 1119–1127
- intravascular ultrasound
 clinical settings and treatments, 1559–1560, 1562
 future research issues, 1562–1564
 imaging studies and modalities, 1560–1562
 overview, 1559
- magnetic resonance imaging, 1713–1727
- microparticle image velocimetry
 basic principles, 1873–1876
 correlation function, 1881–1882
 flow field measurement, 1876–1879
 overview, 1873
 particle detection, 1879–1881
- optical coherence tomography
 applications, 2003–2004
 future research issues, 2004–2005
 imaging probes, 2001–2002
 imaging resolution, 1998–2000
 imaging speed, 2000–2001
 operating principles, 1996–1998
 overview, 1996
 spectroscopic, Doppler, and polarization techniques, 2002–2003
- positron emission tomography, 2315–2321
 real-time tomographic reflection, 2391–2392
- telemedicine applications, 2581–2582
- ultrasound-tissue interaction
 acoustic propagation principles, 2933–2934
 attenuation, absorption, and scattering, 2936–2939
 characteristic impedance, 2934
 intensity, 2934–2935
 linearity effects, 2939–2941
 radiation force, 2935
 reflection/refraction, 2935–2936
- vascular system
 angiography, 3012–3013
 angioscopy, 3017
 computed tomography angiography, 3013–3014
 future research issues, 3018–3020
 magnetic resonance angiography, 3015–3017
 overview, 3012
 processing protocols, 3017–3018
 ultrasound, 3014–3015
- Immobilized biocatalysts, techniques for, 165–166
- Immune response
 biomaterials
 CD4⁺ lymphocyte subsets, 264–265
 lymphocytes, 264–265
 macrophages, 262–264
 major histocompatibility complex, 265–267
 overview, 262
 gene therapy, 1157–1160
 host reactions, 1416–1418
- tissue engineering
 ligaments, 2765–2766
 liver, 2772
 microscale technology, 2782–2783
- valvular prosthetics
 calcification, 79
 failure analysis, 1298, 1311–1312
 xenografts, 2907–2910, 3095–3100
- Immunoassays, biosensors and, 339–343
- Immunoglobulin G (IgG) hydrogel, soft contact lenses, 1442–1443
- Immunoglobulins
 heart valve prosthetic failure, 1310–1312
 host reactions, 1416–1418
- Immunosuppressive drugs, allografts, 9
- Impedance principles
 cardiac output measurement, 526–530
 contraindications, 530
 stroke volume calculation, 527–528
 verification studies, 528–530
- electrical cell impedance sensing, 908–910
- electrical impedance tomography
 hardware requirements, 916–918
 overview, 915–916
 safety requirements, 918–919
 sample images, 923–924
 software requirements, 919–923
- ultrasound-tissue interaction, 2934
- Implant devices and materials. *See also* Transplant devices and materials
- animal models, cardiovascular system, 1566–1571
- battery power sources, 1485–1493
- commercial applications, 1493
- electrochemical power generation, 1485–1486
- future development issues, 1493–1495
- lithium/carbon monofluoride, 1492
- lithium/CF/SVO hybrid cathode, 1492
- lithium-iodine, 1491
- lithium-ion rechargeable, 1492–1493
- lithium/silver vanadium oxide, 1492
- lithium/thionyl chloride, 1491–1492
- mechanical design, 1489–1490
- power, capacity, and longevity, 1487
- primary vs. secondary, 1487–1489
- requirements, 1489
- voltage and current thermodynamics and kinetics, 1486–1487
- biomaterials-related infection incidence, 62
- bone-implant interface
 cemented implants, total hip arthroplasty, 1379–1380
 cementless implants, 427–428
 fixation techniques, 426–427
 interface, 423–430
 materials, 425–426
 ongrowth/ingrowth factors, 428–429
 osteointegration, 424
- cardioverter-defibrillators
 design, 775–776
 history, 774–775
- ceramics, 588–589
- circulatory support systems, 1775–1782
- clinical trials, biocompatibility testing, 175–176
- coating surfaces and, 2533
- cochlear implants, 1507–1513
 functional air conduction mechanisms, 1277–1278
 hearing aid design, 1269
- corneal implants
 future research issues, 751
 glossary of terms, 747–748
 keratoprosthesis, 750–751
 overview, 747
 refractive keratoplasty, 748–750
 tissue engineering, 2707–2714
- corrosion analysis, 754–760
- degradable polymer composites, 779–780
- dental implants
 biomechanics, 817–819
 biomechanics testing, 802–805
 design, 810–813
 future research issues, 819–820
 interface evaluation, 815–816
 materials, 813–815
 osseointegration, 809–810
 overview, 808–809
 surgical technique, 808–809
 survival, 816–817
- fabric structure and classification
 braided three-dimensional fabrics, 1052–1053
 electrospun yarns and fabrics, 1053–1054
 implantation properties, 1043–1044
 multiaxial warp, 1051–1052
 orthogonal three-dimensional fabrics, 1050–1051
 performance characteristics, 1055–1056
 scaffold materials, tissue engineering, 1056–1059
 cell proliferation, 1058
 cell-scaffold architecture interaction, 1058–1059
 structural components, 1044–1054
 fibers, 1046–1048
 textile units, 1044–1046
 three-dimensional fabrics, 1049–1053
 two-dimensional structural geometry, 1048–1049
 woven fabrics, 1048–1050
- foreign body response, 1111–1116
- hydroxyapatite, 1456–1460
 coatings, plasma spraying, 1464–1468
- interstitial fluid sensors (ISF), 1175–1179
- affinity chemical/optical sensor, 1180
- dialysis catheter monitoring system, 1178–1180

- Implant devices and materials (*cont.*)
- enzyme-based telemetry sensor, 1180
 - needle-type, enzyme-based electrochemical biocatalytic sensors, 1175–1178
 - microporous materials, 1886
 - nerve guide technology, 1919–1920
 - nickel alloys, 1942–1949
 - nitric oxide, 1955
 - ocular implants
 - drug delivery systems, 1979
 - glaucoma filtration implants, 1976–1977
 - intraocular lenses, 1977–1978
 - keratoprostheses and intracorneal implants, 1974–1975
 - orbital implants, 1974
 - scleral buckles, 1976
 - silicone oil, 1979
 - viscoelastic agents, 1979
 - orthopedic devices, 1824–1827
 - biodegradable polymers, 2063–2065
 - failure analysis, 1827–1839
 - stable polymers, 2066
 - plasma modification
 - corona and atmospheric discharge, 2175
 - deposition and coating, 2181–2182
 - discharge techniques, 2174–2175
 - implantation, 2176–2177
 - ion implantation, 2175
 - metallic ion implantation, 2177–2180
 - overview, 2174
 - polymer ion implantation, 2180–2181
 - radiofrequency discharge, 2174–2175
 - spraying, 2183–2184
 - sputtering and etching, 2176
 - surface interactions, 2175–2176
 - thin film deposition, 2175–2176
 - vacuum arc, 2175
 - power sources
 - future technologies, 1493–1495
 - historical background, 1484–1485
 - overview, 1484
 - pulsatile devices, 1781–1782
 - pyrolytic carbon, 2376–2380
 - design analysis, 2379–2380
 - spinal biomechanics
 - disc arthroplasty, 2468
 - disc degeneration, 2467–2468
 - spinal deformity biomechanics, 2470–2471
 - spinal fusion, 2467–2468
 - sterilization, 2501–2508
 - supercritical fluids, 2528
 - tissue engineering
 - breast, 2678–2679
 - liver, 2773–2775
 - organoid units, small intestine, 2864–2866
 - titanium and titanium alloys, 2876–2887
 - total hip arthroplasty, 1472–1473
 - carcinogenesis, 1480
 - ceramics, 1474
 - cobalt-chromium alloys, 1476–1478
 - current technology, 1473–1478
 - degradation, wear and corrosion, 1478–1480
 - degradation mechanisms, 1478–1480
 - femoral head prostheses/short-stem prostheses, 1471–1472
 - fixation procedures, 1379–1380
 - future research issues, 1480–1482
 - historical background, 1470–1473
 - hypersensitivity responses, 1480
 - limitations and complications, 1478–1480
 - long-stem prostheses, 1472
 - metal ion release, 1480
 - metals, 1474
 - mold arthroplasty, 1470–1471
 - new alloys and surface coatings, 1476–1478
 - overview, 1470
 - particle distribution, 1480
 - polymers, 1474
 - refractory metals, 1478
 - stainless steel alloys, 1474–1476, 1478
 - titanium alloys, 1476–1478
 - ultra-high molecular weight polyethylene, 2925–2930
 - urinary prostheses, 2961–2969
 - vascular grafts, in vivo animal models, 1566–1571
 - zirconia ceramics, 3105–3106
 - Impulse response (IR) filter, biomedical signal acquisition and processing, 846–847
 - Incident light intensity, confocal microscopy, 709–710
 - Indicator-dilution method, cardiac output measurement, 518–522
 - Indicator fiberoptic biosensors, 1068
 - Induced acceleration analysis, gait analysis, 1138
 - Inductive coupling, biopotential amplifiers, 305–306
 - Inductively coupled plasma spectroscopy (ICP)
 - hydroxyapatite characterization, 1458
 - orthopedic device failure analysis, 1827–1830
 - Industrial design
 - intellectual property rights and, 1541–1543
 - magnetic/electrophoretic cell separation, 1702–1703
 - Industrial imaging, magnetic resonance imaging, 1726–1727
 - Inert materials
 - alumina, 18–19
 - bioactive glasses, 134
 - bone-implant interface, 425
 - hemocompatibility, 1345–1346
 - Infection. *See also* Immune response biomaterials and, 61–62
 - catheterization and adjunct therapies, 1503–1504
 - bacterial infections, 1498–1500
 - colonization vs., 1500–1501
 - cost issues, 1497–1498
 - detection of, 1501
 - device-related factors, 1500
 - host factors, 1500
 - overview, 1497
 - prevention techniques, 1501–1503
 - circulatory support systems, 1779
 - gene therapy for, 1150
 - host reactions, 1414–1415
 - penile prostheses, 2965
 - rotary cell culture systems, 2406
 - valvular prosthetics, calcification, 79
 - vascular grafts, host reactions, 3009
 - Infinity impulse response (IIR) filters, biomedical signal acquisition and processing, 847
 - Inflammation
 - connective tissue healing, 1251
 - foreign body response, 1111–1116
 - heart valve prosthetic degradation, 1296–1298
 - heparin biocompatibility, 1370
 - host reactions, 1411–1412
 - tissue engineering, ligaments, 2765
 - transdermal drug delivery, 2904
 - Inflatable implants, penile prostheses, 2963–2964
 - Infrared (IR) spectroscopy
 - bone mineral composition, 2042
 - gradient surfaces characterization, 1218
 - hemocompatibility assessment, 1349
 - Infusion pumps, controlled release drug delivery, 739–740
 - Inhalant medications, insulin inhalation, 1515–1516
 - Inherently conductive polymers (ICPs)
 - biological functionality, 697–699
 - cellular communication, 699–701
 - chemical polymerization, 697
 - electrochemical properties, 699
 - electrochemical synthesis, 696–697
 - enzyme-catalyzed polymerization, 697
 - future applications, 701–702
 - overview, 695
 - photochemical polymerization, 697
 - structure and properties, 695–696
 - synthesis, 696–699
 - Innate immunity, gene therapy, 1157–1158
 - Inner ear
 - biology, hearing aid design and, 1275–1276
 - implants
 - basic principles, 1507–1508
 - clinical issues, 1512–1513
 - electroneural interface, 1511–1512
 - normal hearing mechanisms, 1507
 - overview, 1507
 - receiver-stimulator implementation, 1510–1511
 - speech processor implementation, 1508–1510
 - Inner hair cells (IHCs), hearing mechanisms and, 1273–1275
 - Inoculum density, bone marrow tissue engineering, culture media, 2663–2664
 - Inorganic materials
 - matrix structure and classification, 2039–2049
 - shells, 2049
 - stent coatings, 2495
 - Insect cells, rotary cell culture systems, 2406
 - Insect mating plugs, adhesive properties, 246
 - Inside-out signaling, integrin regulation, 1534
 - In situ studies
 - blood vessel mechanics, 395
 - electrical cell impedance sensing, 910–911
 - tissue engineering, bladder, 2638
 - Instantaneous axis of rotation (IAR), spine biomechanics, kinematics, 2464–2465
 - Insulin
 - basic properties, 1514
 - bio-artificial liver systems, 1655–1656
 - controlled release delivery systems, 739–741, 744–745
 - glucose hydrogels, 1165–1170
 - glucose sensors, 1180–1182
 - delivery routes, 1514–1516
 - acute insulin delivery, 1517
 - artificial pancreas, 1519–1520
 - basal insulin replacement, 1517–1518
 - industrial contributions, 1518
 - inhalation systems, 1515–1516
 - intravenous administration, 1516
 - ocular, nasal, and buccal mucosa delivery, 1515
 - oral enteric insulin, 1515
 - physiological insulin replacement, 1518–1519
 - polymeric systems, 1519
 - subcutaneous, 1514–1515
 - transcutaneous transport, 1516
 - formulations, 1516–1517
 - protein-based drug delivery systems, 2364–2366
 - pump and pen systems, 2366
 - tissue engineering, pancreas, 2805–2807
 - Integrated circuits, electrotransport mechanisms, transdermal drug delivery, 2900

- Integrated discrete multiple organ cell culture (IdMOC), toxicity testing, animal surrogate systems, 56
- Integrated functional magnetic resonance imaging, 1126–1127
- Integrated Heart-Lung Assisted Device (IHLAD), 1679
- Integrated Services Digital Network (ISDN), real-time telemedicine systems, 2580–2581
- Integrin linked kinase (ILK), integrin postadhesion, 1529–1530
- Integrins. *See also* Scaffold materials
 alpha and beta chains, 1525–1526
 angiogenesis inhibitors, small molecule integrin antagonists, 28
 aspartic acid/glutamic acid-based target sequences, 1527–1528
 basic properties, 1524
 cardiac patch engineering, collagen-myocyte interactions, 546
 chondroitin sulfate interactions, 622
 closed and open states, 1527
 deadhesion, 1530–1531
 extracellular events, 1531–1532
 extrinsic factors, 1527–1528
 fine structure, 1524–1525
 intrinsic factors, 1525–1526
 ligand accessibility, 1528
 ligand binding regulation, 1525
 local cation concentration, 1528
 osteoclast morphology, 2081–2082
 postadhesion events, 1528–1530
 receptor activation state, 1526–1527
 signaling mechanisms, 1532–1534
 structure and function, 1524–1525
 synergy sites, 1528
 tissue engineering
 corneal implants, 2713–2714
 polymer sources, 2283–2284
- Intellectual property management
 biotechnological rights and property varieties, 1544–1545
 confidential information, 1543–1544
 copyright, 1545–1546
 industrial designs, 1542–1543
 patents, 1540–1542
 research institutes, 1546–1548
 trademarks, 1543
- Intelligent materials, polymers, tissue engineering applications, 2294
- Intelligent Tissue Engineering via Mechanical Stimulation (ITEMS)
 cardiovascular applications, 314–320
 ergonomics, 316
 integrated control electronics, 317
 material properties, 317–318
 moving magnetic motor, 315–316
 multistation configuration, 320
 orthopedic applications, 320–324
 pressure/flow measurement, 319–320
 pulsatile pressure control, 316–317
- Intensity parameters
 radiofrequency electromagnetic field/radiation (RFR) effects, 254–255
 ultrasound therapy, bone healing, 2954–2955
 ultrasound-tissue interaction, 2934–2941
- Interactive research, eye movements and, 1037–1038
- Interbody fusion constructs, spinal fusion biomechanics, 2468
- Interconversion potentials, osteogenic progenitor cells, 2093–2094
- Interfacial transport
 biological material polarization, 837–839
 protein and cell signaling in biomaterials
 basic principles, 2347–2348
 current applications, 2348–2349
- electrolyte concentration and permeability, 2352–2353
 future research issues, 2354–2355
 gold-palladium alloy coatings, 2354
 membrane transport, 2349–2358
 molecular mass, 2350–2352
 permeability and, 2354
 pH effects, 2351–2352
 overview and history, 2346
 pore size and permeability, 2353–2354
 water/wastewater treatment applications, 2349
- Interference problems, biopotential amplifiers, 304–306
- Interleukins. *See also* Cytokines
 osteoclast differentiation and activation signaling, 2086–2087
 wear debris, inflammatory response, 3067–3068
- Intermittent renal replacement therapy (IRRT), acute renal failure, 1590–1591
- Internal fracture fixation, 414–418, 2073–2077
 reconstructive biomaterials, facial skeletal fixation, 2396
- Internal load modeling, gait analysis, 1137
- International Patent Classification (IPC), intellectual property management, 1542
- Interpenetrating polymer networks (IPNs)
 environmentally sensitive hydrogels, 1550–1551
 hydrogels, 1550–1553
 overview, 1549
 phase separation and mechanical behavior, 1549–1550
 temperature-sensitive hydrogels, 1551–1554
- Interpenetration theory, bioadhesion, 1–2
- Interstitial cells of Cajal (ICC),
 electrogastrography and, 947–949
- Interstitial collagenase, 1766–1767
- Interstitial fluid sensors (IFS), 1175–1179
 affinity chemical/optical sensor, 1180
 dialysis catheter monitoring system, 1178–1180
 enzyme-based telemetry sensor, 1180
 needle-type, enzyme-based electrochemical biocatalytic sensors, 1175–1178
- Intervertebral disc (IVD) degeneration, spine biomechanics
 kinematics, 2464–2465
 pathology of, 2467
- Intestine, tissue engineering
 feasibility studies, 2752–2753
 overview, 2751–2752
 scaffold materials, 2754
 small intestinal submucosa, bladder regeneration, 2636
- In-the-ear (ITE) hearing aid, 1267–1268
- Intraaortic balloon pumps (IABP),
 components and applications, 471–472
- Intracardiac pressure, Doppler ultrasound imaging, 2949
- Intracellular adhesion molecules (ICAMs)
 fibroblast mechano-regulation mechanisms, 1789–1790
 mass transfer and tissue function, organelle structures, 1751–1752
 complex mechanisms, 1761
 nickel alloys, 1948–1949
 surfactant-protein interactions, 2338–2340
- Intracellular ice formation (IIF),
 cryopreservation of cells, 764–765
- Intracorneal implants
 ring implants, 749–750
 structure and properties, 1974–1975
- Intramedullary nails
 bone fracture fixation, 416–419
 failure analysis, 1835–1836
 fracture fixation, 2076
- Intranasal drug delivery, insulin administration, 2366
- Intraocular lens (IOL). *See also* Ocular implants
 defined, 747
 heparin biocompatibility, 1371–1372
 phosphorylcholine technology, 2145–2146
 structure and properties, 1977–1978
- Intrathoracic Artificial Lung (ITAL), 1677–1678
- Intravascular capillary unit (IVC), artificial pancreas, 2804
- Intravascular lung assist device (ILAC), 1676–1677
- Intravascular respiratory catheters, 1675–1677
- Intravascular sensors, blood glucose monitoring, 1181–1182
- Intravascular ultrasound (IVUS)
 basic principles, 3015
 clinical settings and treatments, 1559–1560, 1562
 Doppler principles, 2942–2949
 future research issues, 1562–1564
 imaging studies and modalities, 1560–1562
 optical coherence tomography, 2004
 overview, 1559
 tissue interaction, 2933–2941
- Intravenous administration systems
 insulin delivery, 1516
 nutritional administration, small intestine tissue engineering, 2860–2861
- Intravenous membrane oxygenator (IMO), 1676–1677
- Intrinsic factors
 integrin structure, 1525–1526
 vascularization, 3026–3028
- Intrinsic viscosity, reconstructive biomaterials, facial skeletal fixation, 2398
- Invention, intellectual property rights and, 1540
- Inverse kinematics, movement control theory, 733
- Inviscid analysis, vascular fluid dynamics, 2972–2974
- In vitro procedures
 articular cartilage repair, 558–559
 artificial pancreas, 2805
 biocompatibility testing, 170–172
 biorubber/poly(glycerol sebacate), 331–333
 biphasic calcium phosphate ceramics, 361–362
 blood vessel mechanics, 394–395
 cardiac assist devices, 476
 positive displacement pumps, 96–105
- cell culture assays
 future research, 566–567
 limitations, 566
 overview, 561–562
 requirements for, 565–566
 standards, 563
- hematopoietic stem cell progenitors, 1338–1339
- ligament injury and repair, animal repair models, 1644
- phosphorylcholine technology, 2140–2141
- radiofrequency electromagnetic field/radiation effects, 257
- tissue engineering
 bladder, 2636
 breast, 2681–2682, 2684
 fat tissue, 2731
 historical perspective, 2621–2622
 microscale technology, 2789
 skin, 2856–2857
 vascular constructs, 3024–3026, 3039–3040

- In vitro procedures (*cont.*)
 toxicology
 animal surrogate systems, 53
 electrical cell impedance sensing, 912–913
 valvular prosthetics, mechanical heart valves, 97–98
 flow field measurements, 1332–1334
 vascular grafts, 2992–2993
 wear debris, joint arthroplasties, 3078–3079
- In vivo procedures. *See also* Animal models
 biocompatibility testing, 172–175
 animal tests, 173–175
 animal welfare issues, 172–173
 carcinogenicity testing, 174
 functional tests, 173
 genotoxicity testing, 173–174
 implant site selection, 173
 irritation tests, 174
 nonfunctional tests, 173
 sensitization tests, 174
 systemic effects, 174–175
 blood vessel mechanics, 393–394
 cardiac assist devices, 476–477
 cardiovascular animal models
 degradation and tissue reaction, rat model, 1569
 functional assessment, 1567
 implantation techniques, 1566–1567
 morphological assessment, 1567–1568
 overview, 1566
 thrombogenicity, acute pig model, 1568–1569
 vascular graft replacement, 1569–1570
 cell-topography interactions, 2558–2560
 dental implants, 803–804
 endoscopic techniques, 996–997
 hydroxyapatite coatings, 1465
 nuclear magnetic resonance, brain imaging, 1967–1969
 optical detection of cancer, spectroscopic techniques, 2011–2012
 periodontal ligament, biomechanics, 800–802
 poly(glycerol sebacate) degradation, 329–331
 poly(glycolic acid), tissue response, 2246–2247
 poly(lactic acid) degradation, 2253–2254
 poly(lactide-co-glycolide) degradation, 2263–2265
 stem cell therapy, 2476–2477
 tissue engineering
 cardiac muscle, 2691
 fat tissue, 2731
 microscale technology, 2789
 skeletal muscles, 2840–2841, 2844–2845
 skin, 2856–2857
 vascularization, 3026–3028
 vascular system, 3040–3041
 vascular grafts, 2993, 2998
 wear debris, joint arthroplasties, 3078–3079
- Ion beam-assisted deposition (IBAD)
 diamond/diamond-like carbons, 823–824
 surface modification, 2540–2541
- Ion beam conversion coatings, diamond/diamond-like carbons, 824–825
- Ion channels
 atria, 488
 atrio-ventricular node, 488
 biosensor development, 342–343
 cardiac action potential, 483–486
 cardiac conduction system, 532
 controlled-release drug delivery systems, 884–885
 electrocardiography, 937–938
 His-Purkinje system, 488–489
 mathematical modeling, 1896–1897
 sino-atrial node, 487–488
 ventricular function, 480–489
- Ionescu-Shiley heart valve prosthetic, 1312–1313
- Ion fiber optic biosensors, 1069
- Ionic electroactive polymers/ionic polymer gels (IPGs), artificial muscles, 109
- Ionic functional groups, gradient surfaces, 1214–1215
- Ionic hydrogels, 1163–1164
- Ionic polymer-conductor composites (IPCCs), artificial muscles, 112–113
- Ionic polymer-metal composites (IPMCs), artificial muscles, 109–110
- Ion implantation, plasma biomaterials, 2175
 metallic materials, 2177–2178
 polymers, 2180–2181
 surface modification, 2543
- Ionizing radiation, sterilization of medical devices, 2505–2506
- Ionomers, orthopedic devices, glass composites, 2069
- Ionophoretic systems, transdermal drug delivery, 2901
- Ion permeability, silicone hydrogel contact lenses, 722–723
- Iridoid compounds, collagen cross-linking, 648–649
- Irradiance, tissue optics, 1624–1625
- Irritation testing, in-vivo biocompatibility testing, 174
- Islet cells, rotary cell culture systems, 2406
- Isocyanates
 collagen tissue fixation, 645–646
 polyurethane segmentation, 2304–2305
- Isomeric structures, degradable polymer composites, 780
- Isometric contractions, electromyographic signals, 966
- Isotope analysis, nuclear magnetic resonance, 1966
- Isotropic reactive ion etching, tissue engineering, microvascular networks, 2795–2796
- i-STAT device, microparticle image velocimetry, 1873
- IVOX intravascular respiratory catheters, 1675–1677
- J**
- Jade, implant devices, 587–588
- Johnson noise, magnetic resonance microscopy, 1739–1740
- Joint properties
 articular cartilage biomechanics, 92–93
 chondroitin sulfate pharmacology, 624–625
 gait analysis, 1131–1134
 hip joint biomechanics
 arthritis, normal and artificial hips, 1377
 cup stability, micromotion, fixation, and osteolysis, 1380
 femoral head subluxation, 1381–1382
 finite element modeling and validation, 1381–1385
 future research issues, 1388–1390
 gait loading results, 1386
 implant fixation techniques, 1379–1380
 liner micromotion, 1385–1386
 pelvic osteolysis and adjunctive screws, 1380–1388
 post THA dislocation, 1381
 subluxation loading results, 1386–1388
 total hip replacement techniques, 1377–1379
 ligament guidance, 2761–2762
 movement control models, 732–734
 protein-surfactant interactions, 2337
 wear debris, arthroplasty procedures, 3076–3079
- Joint prostheses
 alumina, 17
 articulating surfaces, 19–21
 bioadhesive bonds and, 155
 diamond/diamond-like carbons, 827–828
 failure analysis, 1824–1837
 anterior cervical spine plate, 1832–1834
 dynamic compression plate, 1829–1832
 fracture fixation, 1827–1839
 intramedullary nails, 1835–1837
 overview, 1824
 prostheses, 1839–1841
 tibial tray, 1841
 titanium screws, 1834–1835
 hyaluronan biomaterials, 1427
 orthopedic soft tissue, allografts, 13–14
 pyrolytic carbon, 2374–2375
 total hip arthroplasty, 1470–1481
 total knee replacement, 1607–1613
 ultra-high molecular weight polyethylene, 2925–2930
 wear debris, arthroplasty procedures, 3076–3079
- Judkins cardiac catheters, 499–500
- Justice, biomedical ethics and, 291–292
- K**
- Karhunen-Loeve transform, ECG signal compression, 680–681
- Keratinocytes, burn injuries and, 449
- Keratin sulfates (KSSs), glycosaminoglycans, 1195–1198
- Keratocytes, tissue engineering, corneal implants, 2710
- Keratophakia, materials and techniques, 748–749
- Keratoprosthesis (KPro)
 defined, 747
 porous skirts, 751
 structure and properties, 1974–1975
 techniques and materials, 750–751
- Kevlar[®], basic properties, 2189–2190
- Kidneys
 anatomy and function, 1576–1577
 artificial
 biocompatibility, 1583
 cellulosic membranes, 1583–1584
 classification, 1580–1583
 dialysis membranes, 1583–1585
 efficiency, 1581–1583
 flux changes, 1581
 future research issues, 1592
 historical perspective, 1578–1579
 hollow-fiber membranes, 1579–1580, 1585
 overview, 1576
 synthetic membranes, 1584–1585
 uremic solute classification, 1577–1578
- Kinematics
 gait analysis, 1131–1133
 spine biomechanics, 2464–2465
- Kinetics
 gait analysis, 1133–1134
 implantable batteries, 1486
 protein-based drug delivery stability, 2358–2359
 spine biomechanics, 2465–2466
- Kininostatin, current research, 27
- Kirchoff-Love hypothesis, composite materials, plane stress, 665–666
- Kirschner wires (K-wires), fracture fixation, 2074–2075
- Knee joint. *See also* Total knee replacement
 allografts, materials and techniques, 13–14
 alumina joint prostheses, 21
 bioreactor applications, 311–312

- overuse injuries
 breaststroke's knee, 1602
 epidemiology, 1597–1598
 Fabella syndrome (fabellitis), 1605
 fat pad syndrome (Hoffa disease), 1603–1604
 future research issues, 1605
 iliotibial band friction syndrome, 1600–1602
 Osgood-Schlatter disease, 1603
 patellar tendinitis/tendinosis, 1599–1600
 patellofemoral pain syndrome, 1597–1599
 pes anserinus tendinitis and bursitis, 1604–1605
 plica syndrome, 1604
 popliteal tendinitis, 1605
 semimembranosus tendinitis, 1604
 Sinding-Larsen-Johansson disease, 1603
- replacement techniques
 anatomical properties, 1608–1609
 bioreactor applications, 311–312
 component fixation, 1612–1613
 contraindications, 1608
 indications for, 1607–1608
 overview and history, 1607
 postoperative function, 1613
 prosthetic design factors, 1609–1612
 wear debris, 3062–3063
- Knitted fabrics
 three-dimensional structure and properties, 1051
 two-dimensional structure and properties, 1049
 vascular grafts, 2987–2988
- Kramers-Kronig relations, permittivity and conductivity, 834
- Krogh cylinder model, oxygen transport, 1844–1845
- Kyphoplasty, bone cement, 403
- L**
- Lactate dehydrogenase (LDH)
 interfacial transport, 2350–2358
 surfactant interactions, 2340–2343
- Lactic acids and lactides. *See also* Poly(lactic acid)/poly(lactide) (PLA); Poly(lactic/lactide-co-glycolide/glycolic acid) (PLGA)
 articular cartilage repair, 556
 biodegradable polymers, 196–199
 copolymers, orthopedic devices, 2064
 poly(lactic acid) chain-growth polymerization, 2250–2251
 reconstructive biomaterials, 2996–2997
- Lag screws, fracture fixation, 415, 2075
- Lamé constants, acoustic propagation, ultrasound, 2933–2934
- Laminate theory, composite materials, 663
- Laminin, extracellular matrix, scaffold materials, 1023–1025
- Langmuir-Wilhelmy surface balance, lung surfactants, 1686
- Language, eye movements and visual-social context for, 1035
- Laplace equation
 biomedical optics, 2029–2030
 composite properties, 659–661
 dielectric and conductive media, 1010
 lung surfactant structure and properties, 1684–1685
- Larmor frequency
 magnetic resonance imaging, 1715
 nuclear magnetic resonance, 1961
- Laser Doppler anemometry (LDA)
 cardiac assist devices, positive displacement pumps, 96–97, 100–101
 centrifugal blood pump evaluation, 580–583
- Laser Doppler Velocimetry (LDV), valvular structure and function, semilunar valve, 1319
- Laser in situ keratomileusis (LASIK), wound healing, 2712–2714
- Laser technology
 acoustics, laser-tissue interactions, 1629
 fiberoptic biosensors, 1063
 flow cytometry, 1102–1103
 laser beam machining, 1861
 laser in situ keratomileusis (LASIK), 2033
 laser interstitial thermotherapy, 1728
 laser-tissue interaction
 biological effects, 1630
 laser tissue welding, 2032
 light-tissue interactions, 1625–1626
 overview, 1623
 photochemical interactions, 1626
 photomechanical interaction, 1628–1630
 photothermal interactions, 1626–1627
 tissue optics, 1623–1625
 magnetic resonance imaging, temperature measurement, 1730
 microparticle image velocimetry, 1874–1876
 optical mapping, 2020–2021
 rapid prototyping, selective laser sintering, 2384–2386
 small-angle light scattering, soft tissue analysis, 2450–2452
 surface modification, 2541–2543
 tendon/ligament tensile testing, 2591–2592
 transmyocardial laser revascularization, 1632–1637
- Late-onset auditory deprivation, 1277
- Latex interpenetrating polymer networks, 1549
- Lead materials
 cardiac pacemakers, 534–535
 electrocardiography, 940–941
 implantable cardioverter-defibrillators, 776
 lead zirconate titanate, piezoelectric ceramics, 2167
- Leaflet loading mechanics, atrioventricular valves, 1324–1326
- Leather goods, collagen processing, 652–656
- Lectin-mediated gels, basic properties, 1447
- Lectins
 bioadhesive formulation factors, 156–157
 as second-generation bioadhesives, 156
- Left ventricular assist devices (LVADs). *See also* Ventricular assist devices
 Abiocoar implantable artificial heart, 1780–1781
 ambulatory implantable left ventricular assist devices, 1777–1778
 battery system, 1493
 complications, 1778–1780
 echocardiographic imaging, 506–507
 heart biomechanics, 1288
 HeartMate implantable pump, 1777–1778
 HeartMate Vented Electric (VE) device, 1777
 implantable pulsatile devices, 1781–1782
 nitric oxide materials, 1955–1956
 Novacor LVAD, 1778
 overview, 1775–1776
 positive displacement pumps, 95–96
 pump physiology, 1776
- Legal principles, biomedical ethics and, 285–292
- Leucite-reinforced ceramics, dental applications, 598
- Leukocyte filters, plasmapheresis, 390–391
- Leukocytes, blood-material interactions, 373
- LexO₂con, cardiac output measurement, 518
- Lifshitz-van der Waals forces, bacterial adherence and, 63–64
- Ligaments
 fibroblast mechano-regulation, 1783–1788
 healing potential, 1253–1254
 multiaxial mechanical testing
 anisotropic properties, 2592–2593
 bidirectional testing, 2593–2594
 shear loading, 2594–2595
 structural framework, 2588–2589
 tissue engineering
 ankle injuries, 2758–2759
 anterior cruciate ligament injuries and reconstructions, 2758
 bone-tissue integration, 2763
 cell structure, 2761
 composition and structure, 2759
 design issues, 2762–2768
 immunological acceptability, 2765–2766
 inflammation, avoidance, 2765
 joint guidance, 2761–2762
 ligament-bone insertions, 2761
 loading mechanisms, 2762
 mechanical properties, 2759–2760
 proprioception, 2762
 scaffold materials, 2759
 shoulder injuries, 2759
 strength maintenance over time, 2763–2765
 vascularization and nutrient transport, 2765
 uniaxial tensile testing, 2589–2592
 viscoelasticity, 2595–2597
 in vitro analysis, 2591–2592
- Ligand binding, integrin regulation, 1525
 accessibility, 1528
 closed and open states, 1527
- Light emitting diodes (LEDs)
 optical mapping, 2020
 polyelectrolyte multilayers, 2222–2223
- Light microscopy, orthopedic device failure analysis, 1828–1830
- Light-scattering spectroscopy
 biomedical optics, 2035
 small-angle light scattering, soft tissue analysis, 2453–2454
- Light sources
 fiberoptic biosensors, 1063
 optical mapping, 2020
 positron emission tomography, 2314–2315
- Light-tissue interactions
 basic principles, 1625–1626
 laser technology
 basic principles, 1625–1626
 biological effects, 1630
 overview, 1623
 photochemical interactions, 1626
 photomechanical interaction, 1628–1630
 photothermal interactions, 1626–1627
 tissue optics, 1623–1625
 optical properties
 applications, 2032–2033
 imaging applications, 2035–2037
 light propagation in tissue, 2033–2034
 light properties, 2029–2030
 overview, 2029
 spectroscopic techniques, 2034–2035
 tissue properties, 2030–2031
- Light transport equation, tissue optics, 1625
- Limpet mucus, adhesive properties, 243
- Linear, time-invariant (LTI) filters, biomedical signal acquisition and processing, 845–847
- Linear amplification, hearing aid design, 1267–1268
- Linear array scanning, endoscopic techniques, 991
- Linearized images, electrical impedance tomography, 919

- Linear polymers, biodegradable compounds, 196
 elastomers, 898–899
 natural materials, 199–200
- Linear variable differential transducers (LVDT), total hip arthroplasty, 1382, 1388–1390
- Liner micromotion, total hip arthroplasty, finite element modeling, 1385–1386, 1388–1390
- LionHeart™ Left Ventricular Assist System, structure and components, 95–96
- Lipids
 biomimetic materials, 296–297
 nuclear magnetic resonance, brain imaging and suppression of, 1969
- Lipopolysaccharides, biofilm formation, 208–209
- Liposomes, gene therapy, 1153
- Liquid cell matrix cultures, tissue engineering, cardiac muscle, 2686–2687
- Liquid crystal elastomer (LCE) materials, artificial muscles, 109, 113–114
- Lithium batteries
 cardiac pacemakers, 535–536
 classification and characteristics, 1491–1493
- Lithography. *See also* Soft lithography
 tissue engineering
 microscale technology, 2780–2781
 microvascular networks, 2794–2795
- Lithotripsy, light-tissue interactions, 2033
- Live cell ocular implants, drug delivery, 1991–1992
- Liver
 bio-artificial liver systems
 bioreactors, 1656–1657
 cell-cell interactions, 1654–1655
 cell source and therapeutic cell mass, 1651–1652
 clinical setting, 1649–1651
 extracellular matrix and cell polarity, 1653–1654
 future research issues, 1658
 hepatocyte transplantation, 1650
 hollow-fiber systems, 1656–1657
 overview, 1649
 parallel plate systems, 1657
 performance monitoring, 1657–1658
 plasma/blood effects, 1656
 porous three-dimensional matrices, 1655
 scale-up hepatocyte culture, 1656
 soluble factors, 1655–1656
 surface chemistry optimization, 1653
 temporary support system, 1651
 tissue engineering, 1652–1656
 xeno (cross-species transplantation), 1650
- rotary cell culture systems, 2406
- tissue engineering
 bioreactors, 2772–2773
 coatings, 2771–2772
 future research issues, 2775–2776
 hepatic environment, 2769–2771
 immune response, 2772
 liver-assist devices, 2773–2774
 liver transplants, 2773–2776
 microscale technology, 2784–2785
 overview, 2769
 polycaprolactone, 2203
 poly(glycerol sebacate), 333
 scaffold materials, 2771
- Liver-assist devices (LADs), tissue engineering, 2773
- Living dermal replacement (LDR), skin tissue engineering, 2854
- Living skin equivalent (LSE), skin tissue engineering, 2853–2854
- Loading properties
 ligaments, 2762
 load-elongation curve, tendon/ligament tensile testing, 2589–2592
- Locking capacity, intramedullary nailing, bone fracture fixation, 418–419
- Long bones, healing mechanisms, 1255–1256
- Long-stem prostheses, total hip arthroplasty, 1472
- Long-term culture-initiating cell (LTC-IC) assay, development of, 1339
- Long-term repopulating assay (LTRA), development of, 1339
- Lorentzian functions, nuclear magnetic resonance, 1963–1964
- Lossless compression, digital biomedical signals, 677–678
- Loss modulus, hard tissue viscoelasticity, 1244
- Lossy compression, digital biomedical signals, 677–678
 electrocardiography, 681
 electromyography, 681–682
 EOG signals, 682–683
- Lotraflicon silicone hydrogel contact lenses, 726–727
- Loudness correction, hearing mechanisms, 1273–1275
- Loudness recruitment, hearing loss and, 1266
- Low back pain and dysfunction, ergonomics guidelines and, 1000–1001
- Lower critical solution temperature (LCST) glucose hydrogels, 1164–1165
 interpenetrating polymer networks, 1550–1551
- Lower extremities
 neuroprostheses, 1926–1927
 robotics, 1807
- Low-flux hemodialyzers, 1581
- Low-frequency dielectric properties, biomaterials, 840
- Low-intensity-pulsed-ultrasound (LIPU), bone healing, 2951–2959
- Low-molecular-weight fluids
 cationic surfactant, protein-surfactant interactions, 2342–2343
 silicones, 2429
- Low-molecular-weight heparin (LMWH), angiogenesis inhibitors and, 29
- Lubricants
 coating materials, 2532
 surfactants, protein interactions, 2337
- Lumbar spine, stress fracture, 2518–2519
- Lungs
 artificial
 annular fiber bundles, 1679
 arterovenous carbon dioxide removal, 1673–1675
 blood oxygenator, 1666–1667
 cardiopulmonary bypass oxygenators, 1667–1669
 CORx™, 1679
 current research and future directions, 1673
 diffusional boundary layers, 1665
 extracorporeal membrane oxygenation, 1669–1670
 gas exchange determinants, 1663
 historical background, 1661–1662
 integrated heart-lung assisted device, 1679
 intravascular respiratory catheters, 1675–1677
 mass transfer correlations, 1665–1666
 membrane permeance, 1663–1665
 next-generation technologies, 1680
 operating principles, 1662–1666
 overview, 1661
 photolytic artificial lung, 1679–1680
 total artificial lungs, 1677–1678
- surfactant properties
 clinical exogenous surfactant therapy, 1690–1691
 diseases of, 1689–1690
 drug inhibition resistance and activity, 1691
 endogenous surfactant activity, 1683–1684
 endogenous surfactant deficiency and dysfunction, 1687
 future research issues, 1691–1692
 inhibitor-induced dysfunction, 1687–6
 overview, 1683
 physiological correlates of dysfunction, 1688–1689
 protein interactions, 2337
 specific component properties, 1686–1687
 surface activity measurement, 1685–1686
- Luteinizing hormone-releasing hormone (LHRH), magnetic nanoparticles, 1710
- Lymphocytes
 host reactions, particulate matter, 2113–2114
 immune response, 264–265
- Lysosomal storage disease, glycosaminoglycan dysfunction, 1203
- Lysozymes, protein adsorption behavior, 278
- ## M
- Macrocirculatory hemodynamics
 arterial pulse waves, 1364–1365
 blood viscosity, 1357–1359
 circulation properties and mechanics, 1356–1357
 fluid dynamics, 1359–1366
 overview, 1356
 pathology, 1366–1367
 pulsatile flow mechanics, 1361–1364
 steady flow mechanics, 1359–1361
 vascular tissue engineering, 3025
 venous hemodynamics, 1365–1366
 vessel wall properties, 1359–1360
- Macrocytic compounds, biomimetic materials, 294–296
- Macroencapsulation, insulin delivery systems, 1519
- Macrogeometry, alumina, 19
- Macromer technology, silicone hydrogel contact lenses
 architecture, 719–720
 evolution of, 718–719
- Macromonomers, silicones, 2428–2429
- Macrophage colony stimulating factor (M-CSF), osteoclast differentiation and activation signaling, 2086
- Macrophages
 cell-topography interactions, 2558–2560
 foreign body response, 1113
 immune response mechanisms, 262–264
 interstitial fluid sensors, 1177–1178
 particulate matter, host reactions, 2111–2114
 polycaprolactone degradation, 2199–2200
 valvular prosthetic failure, pannus formation, 1300–1301
 wear debris
 inflammatory response, 3067–3068
 joint arthroplasties, 3078
- Macroporosity/microporosity, biphasic calcium phosphate ceramics, 360–361
- Macroscopic viscoelastic models, movement control, 734
- Madaline Rule II (MRII), error back propagation, 122–123
- Magnesium, in biologic apatites, 226–229
- Magnetic cell separation
 applications, 1702–1703
 existing methods, 1695–1696

- future research, 1703
 mathematical modeling, 1697
 multistage method, 1696–1697
 scale-up and economic issues, 1702
 viscous medium model, 1698–1699
- Magnetic nanoparticles**
 bioseparation, 1709–1710
 coprecipitation, 1706–1707
 disease therapy, 1710–1711
 homogeneous precipitation, 1706
 magnetic resonance imaging, 1709
 microemulsion formation, 1710
 nonaqueous colloidal precipitation, 1707
 overview, 1705
 pyrolysis, 1710
 superparamagnetism, 1705–1706
- Magnetic resonance angiography (MRA)**, 1723–1724
 basic principles, 3015–3017
 interventional techniques, 3019
- Magnetic resonance imaging (MRI)**.
See also Functional magnetic resonance imaging; Nuclear magnetic resonance spectroscopy
 basic principles, 1713
 biomedical applications, 1725–1726
 brain imaging, 1967–1970
 clinical applications, 1720–1725
 contrast sources, 1718–1719
 current density, 1719
 glycosaminoglycans, 1202
 gradients and imaging, 1716
 historical background, 1713–1714
 image weighting, 1718
 industrial/nonmedical applications, 1726–1727
 intravascular ultrasound, 1561–1564
 magnetic nanoparticles, 1709
 motion-induced contrast, 1718–1719
 myocardial strain, 507
 physical principles, 1714–1715
 polarization transfer, 1720
 proton density, 1718
 radiofrequency excitation and signal detection, 1715
 rapid imaging, 1716–1717
 relaxation, 1715–1716
 relaxation times, 1718
 signals and magnetization, 1714–1715
 spectral information, 1719–1720
 stress fracture diagnosis, 2515–2516
 technology, 1717–1718
 temperature measurement
 clinical applications, 1730–1731, 1735
 future research issues, 1734
 molecular diffusion coefficient, 1728–1729
 overview, 1728
 proton resonance frequency shift, 1731–1734
 spin-lattice relaxation time, 1729–1730
 thermometry methods, 1734–1735
 three-dimensional imaging, 1794
 vascular imaging, 3018–3019
- Magnetic resonance microscopy**
 atherosclerotic plaques, 1746–1747
 biological cell visualization, 1745–1746
 cells and gene expression, 1744–1745
 defined, 1738
 future research issues, 1747
 high-temperature superconduction coils, 1740–1741
 receive coils, 1741
 histology and amyloid plaque characterization, 1742–1744
 molecular imaging, 1746
 mouse phenotyping, 1741–1742
 signal-to-noise ratio, 1738–1741
- Magnetic shape memory (MSM) materials**, artificial muscles, 108
- Magnetic stimulation**
 artificial muscle, 108
 controlled-release drug delivery systems, 886
 excitable tissue, 1018
- Magnetization principles**, magnetic resonance imaging, 1714–1715
- Magnetostatic potential energy**, magnetic cell separation, mathematical modeling, 1697
- Magnetostrictive materials**, artificial muscles, 108
- Magnetron sputter coating**, diamond/diamond-like carbons, 826
- Major histocompatibility complex (MHC)**
 antigen presentation, 263–264
 heart valve prosthetics, remnant tissue immunogenicity and failure of, 1310–1312
 immune response, 265–267
 toxicity and hypersensitivity reactions, 1418
- Malate dehydrogenase (MDH)**
 interfacial transport, 2350–2358
 surfactant interactions, 2340–2343
- Malgaigne's sign**, groin pain syndrome, 1394–1395
- Mandibular reconstruction plate**, failure analysis, 1837–1839
- Mandrel-guided fibro-collagenous tubes (FCTs)**, hybrid vascular prostheses, 1434–1436
- Man-made movement control**, basic principles, 732–733
- Mannich-type reactions**, heart valve prosthetic failure, 1307
- Mannose-binding lectin (MBL)**, complement system, host reactions, 1413
- Mantid oothecal protein**, adhesive properties, 246
- Manual wheelchairs**, components and technology, 3084–3085
- Mapping techniques**, electrical impedance tomography, 916–918
- Mass transfer**
 artificial lung systems, 1665–1666
 dialysis membranes, 1586–1588
 electrophoretic cell separation, 1700–1701
 rotary cell culture systems, 2403
 tissue function
 complex mechanisms, 1761
 concentration diffusion, 1756–1761
 flow-dominated systems, 1752–1756
 intracellular organelles, 1751–1752
 major organ systems, 1749–1751
 molecular complexes, 1752
 organ function, 1751
 overview, 1749
 spatio-temporal hierarchies, 1749–1752
 whole cells, 1751
- Mast cells**, connective tissue healing, 1251
- Matched filters (MF)**, biomedical signal acquisition and processing, 852–853
- Material properties analysis**, finite element analysis, 1094
- Mathematical modeling**
 applications, 1896–1897
 articular cartilage biomechanics, 91–92
 basic principles, 1894
 blood flow, microvascular tissue engineering, 2793–2794
 cardiovascular system, 1897–1900
 body-level modeling, 1900
 cell-level modeling, 1898
 molecular modeling, 1897–1898
 organ-level modeling, 1899–1900
 tissue-level modeling, 1898–1899
- electrical impedance tomography, 919–923
 electrophoretic cell separation, 1700–1701
 finite element analysis, 1091–1098
 historical background, 1894–1896
 insulin delivery, 1520
 magnetic cell separation, 1697–1698
 movement control, 732–733
 oxygen transport, 1844–1845
 periodontal ligament, 801–802
 physiome project, 1900–1901
 theory and process, 1896
- Matrices**. *See* Extracellular matrix; Matrix metalloproteinases
- Matrix composites**
 alumina, 23
 controlled release drug delivery
 ocular implants, 1983–1987
 inserts, 1984
 miscellaneous designs, 1986–1987
 scleral plugs, 1984–1985
 polymers, 741–742
 tissue engineering, small intestine, 2865–2866
 valvular prosthetics, mechanics, 2742–2743
- Matrix metalloproteinases (MMPs)**
 cancer metastasis, 27–28
 collagenases, 1766–1768
 collagen degradation, 634
 gelatinases, 1768–1769
 heart valve prosthetic failure, 1298
 historical background, 1763–1764
 integrin adhesion, 1531
 membrane-type, 1771
 nanoscale neurosurgery, tumor cell invasion and axonal extension, 1908–1909
 ocular angiogenesis, 28
 organic matrix degradation, 2083–2084
 overview, 1763
 proteoglycans cleavage, 1201
 stromelysins, 1769–1771
 structure and regulation, 1764–1766
 tissue engineering
 clinical applications, 1771–1772
 skeletal muscle scaffolds, 2842
 valvular prosthetics, calcification, 74, 78–81
- Maxillofacial bones**, ultrasound therapy, 2957–2957
- Maximum surface tension**, lung surfactants, 1686
- Maxwell elements**
 composite micromechanics, 666–671
 magnetic cell separation, stress tensors, 1697
- Maxwell's equation**
 electrical stimulation, 1010
 optical detection of cancer, 2009–2010
 tissue optics, 1623–1625
- Maxwell-Wagner effect**
 interfacial polarization, biological materials, 837–839
 mid-range frequencies, 841
- Mechanical/electrical systems**, controlled release drug delivery, 739–740
- Mechanical heart valves (MHV)**
 blood-material interactions, 1334–1336
 cavitation, 1331
 classification, 1329
 durability, 1329–1330
 fluid mechanics applications, 1331–1332
 overview, 1329
 in vitro flow field measurements, 1332–1334
 wear and fatigue failure, 1330–1331
- Mechanical precision machining**, microelectromechanical systems vs., 1863
- Mechanical properties**. *See also* Biomechanics
 bioactive glasses, 137–138
 biphasic calcium phosphate ceramics, 361
 bone cements, 408–409

- Mechanical properties (*cont.*)
 dental implants, 802–805
 fibroblast mechano-regulation mechanisms, 1787–1790
 loading functions, 1787–1788
 stretching functions, 1783–1785
 hard materials, elasticity, 1236–1241
 interpenetrating polymer networks, 1549–1550
 periodontal ligament, 800–802
 sutures, 2569–2570
 tissue engineering, blood vessels, 2647
 transdermal drug delivery, 2901–2902
 micropore creation, 2902
 microprojection systems, 2902
 needleless injection systems, 2902
 valvular prosthetics
 calcification, 77–78
 failure, 1296, 1299–1301
 valvular structure and function
 atrioventricular valves, 1323–1326
 semilunar valve, 1319–1320
- Mechanical stimulation
 bioreactors for, 321
 bone healing and, 1259
 cardiac patch engineering, 545–546
 tissue engineering, 2623
 articular cartilage, 2702
- Medial collateral ligaments (MCLs)
 animal repair models, 1640–1647
 anatomy, structure, and function, 1640–1641
 injury and healing mechanisms, 1641–1643
 crimp mechanics, 2760
- Medical education, robotic surgical training
 simulators and haptics, 1811–1812
- Medical informatics, telemedicine systems, 2584
- Megabalanus rosa*, adhesive properties, 242
- Melting points
 corneal degradation, 747
 polymer structure/property relationships, 2275–2276
- Melt molding, tissue engineering scaffolds, 2830–2831
- Melt processing
 biocomposite fabrication, 186–187
 poly(lactide-co-glycolide), 2259–2261
 polyurethanes, 2212–2213
- Melt spinning
 basic principles, 1816
 bioactive glasses, 1818–1821
 biomedical applications, 1822–1823
 compositional dependence, 1818–1819
 fiber drawing, 1820
 structure and strength, 1820–1821
 biomedical applications, 1821–1823
 polymers, 1816–1818
 biomedical applications, 1821–1822
 extrusion, 1817
 fiber drawing, 1817–1818
 fiber quenching, 1817
 properties, 1818
- Membrane attack complex (MAC), complement system, host reactions, 1413
- Membrane structures
 artificial lung systems, 1663–1665
 biomimetic materials, 296–297
 biosensors, 338–339
 electrical stimulation, 1012–1016
 hemodialysis, 379
 interstitial fluid sensors, 1177–1178
 membrane-covered stents, basic properties, 2495–2496
 on-line filtration, plasma treatment, 386–387
 protein-surfactant interactions, 2343
 tissue engineering scaffolds, laminations, 2829–2830
- Membrane transport, interfacial transport, 2349–2358
- Membrane-type matrix metalloproteinases (MT-MMP), structure and function, 1771
- Memory, eye movements and, 1035–1036
- Memory units, implantable cardioverter-defibrillators, 776
- Meningitis, cochlear implants and risk of, 1512–1513
- Meniscal replacements. *See* Articular cartilage; Knee joint; Total knee replacement
 allografts, materials and techniques, 13–14
 alumina joint prostheses, 21
- Mental disorders diagnosis, functional magnetic resonance imaging, 1123
- Mesenchymal stem cells (MSCs)
 articular cartilage repair, 556–557
 biomaterials and research on, 2476
 ceramic materials adhesion, 2916–2923
 chitosan gene expression, 610–611
 host reactions, particulate matter, 2112–2114
 osteogenic progenitor cells, 2091–2092
 phenotype, 2473
 photopolymerized biomaterials, 2159–2160
 properties and applications, 1341–1342
 tissue engineering
 bone, 2653–2654
 fat tissue, 2726, 2728
 intelligent polymers, 2294–2295
 tendons, 2873–2874
 wear debris, joint arthroplasties, 3079
- Messenger RNA (mRNA), collagen synthesis, 630–632
- Metabolic changes
 light-tissue interactions, 2032–2033
 positron emission tomography imaging, 2320–2321
 transdermal drug delivery, inhibition of, 2904
- Metabolic engineering, whole-cell biocatalysis, 165
- Metabolite clearance (MC), hemodialysis, 382–383
- Metacarpophalangeal (MCP) joints, pyrolytic carbon prostheses, 2374–2375
- Metal ion-dependent adhesion site (MIDAS), integrin structure, 1525
- Metals. *See also* Alloys
 as antimicrobial biomaterials, 66–67
 blood-contacting materials, 375
 cardiac pacing leads, 535
 cell-material interactions, 568–569
 corrosion analysis
 electrochemical reactions, 754–757
 crevice corrosion, 755
 fretting corrosion, 755
 galvanic corrosion, 755–756
 stress corrosion cracking, 756–757
 electrolyte chemical analysis, 759–760
 metal surface analysis, 758–759
 overview, 754
 dental implants, 813–814
 porcelain-bonded metal, 597
 fiberoptic biosensors, 1063–1064
 hard tissue-biomaterial interactions, 1229–1231
 nickel and nickel alloys
 chemistry, 1942
 hypersensitivity reactions, 1948
 nickel-chromium, 1942–1943
 nickel-titanium, 1945–1948
 stainless steel, 1943–1945
 orthodontic wires
 beta-titanium, 2057
 cobalt-chromium alloys, 2057
 nickel-titanium alloys, 2057–2059
 overview, 2054–2056
 stainless steel, 2056–2057
 strengths and limitations, 2059
 orthopedic devices, 2061
 anterior cervical spine plate, 1832–1834
 bone biology, structure and properties, 1824
 dynamic compression plate, 1829–1832
 failure analysis, 1824
 fracture fixation failure analysis, 1824–1839
 intramedullary nails, 1835–1837
 joint replacement, 1824–1827
 prostheses failure analysis, 1839–1841
 tibial tray, 1841
 titanium screws, 1834–1835
 plasma biomaterials modification, ion implantation, 2177–2180
 self-assembled monolayers, metal-alkanethiol interactions, 2409–2411
 shape memory metals
 basic properties, 2418–2419
 biomedical applications, 2419–2420
 overview, 2418
 surface modifications, 2420–2422
 ceramic coatings, 2423
 chemical and electrochemical treatments, 2422–2423
 hydroxyapatite coatings, 2423
 oxidation, 2422
 polishing, ultrasonic bath, and physical etching, 2421–2422
 polymer coatings, 2423–2424
 sterilization, 2424
 stent devices, endovascular stents, 2494
 titanium and titanium alloys
 basic properties, 2876–2877
 biocompatibility, 2879–2880
 biomedical applications, 2877–2878
 bone-implant interactions, 2880
 cytotoxicity, 2879–2880
 dental applications, 2887–2889
 fatigue and fretting fatigue, 2881–2885
 mechanical properties, 2881–2885
 nickel-free alloys, 2879
 surface modification, 2885–2887
 wear characteristics, 2884–2885
 total hip arthroplasty, 1472–1474, 1478
 hypersensitivity response, 1480–1481
 ion release, 1480
 wear debris, 3058–3063
 Metastases, electrical cell impedance sensing, 912–913
 Metatarsals, stress fracture, 2517–2518
 Methacrylic acid (MAA), interpenetrating polymer networks, 1550–1551
 2-Methoxyestradiol (2ME), angiogenesis inhibitors, 28
 Methyl methacrylate monomer residuals, bone cement polymers, 407
 Michaelis-Menten model, mass transfer and tissue function, 1760–1761
 Microassembly techniques, microelectromechanical systems, 1866–1870
 Microbial compounds
 adhesion
 biomaterials-related infections, 61–62
 surface modification for prevention of, 63–64
 biomaterials-related infections, 61–62
 magnetic/electrophoretic cell separation, 1702–1703
 surfactants, 2338

- Microbial surface components recognizing adhesive matrix molecules (MSCRAMMs), biofilm formation, 209
- Microbiology, flow cytometry, 1108
- Microcirculatory system. *See also* Blood vessels
- basic properties, 1844
 - microparticle image velocimetry, 1878–1879
 - oxygen transport
 - arterioles, 1845–1846
 - artificial oxygen carriers, 1847
 - cardiopulmonary system, 1843–1844
 - carrier systems, 1843
 - cellular metabolism, 1842–1843
 - experimental methods, 1845
 - future research issues, 1847–1848
 - nitric oxide, 1847
 - overview, 1842
 - pathophysiology, 1847–1848
 - research issues, 1844–1845
 - tissue properties, 1845–1847
 - venules, 1846–1847
 - vascular fluid dynamics, 2977–2978
 - in vitro vascular constructs, 3024–3025
- Microcomputed tomography
- future research, 1857
 - historical background, 1850–1851
 - image generation, 1851–1852
 - image reconstruction, 1853
 - osteoporosis research, 1854–1855
 - overview, 1850
 - precision and accuracy, 1853–1854
 - projection data collection, 1853
 - three-dimensional measurements, 1854
 - enteric coated tablets, 1856–1857
 - trabecular bone architecture, 1855
 - vessel imaging and angiogenesis, 1855–1856
- X-ray and computerized tomography
- theory, 1851
 - x-ray attenuation, 1852
- Microcrack propagation/non-propagation, pyrolytic carbon, 2377–2380
- Microelectromechanical systems (MEMS). *See also* Microparticle image velocimetry
- future applications, 1870
 - manufacturing applications
 - biomimetics, 1865–1866
 - bottom-up process, 1866–1870
 - historical background, 1860–1862
 - overview, 1860
 - top-down processing, 1862–1865
 - tissue engineering
 - liver, 2769
 - microvascular networks, 2794–2798
- Microemulsion formation, magnetic nanoparticles, 1708
- Microencapsulation
- artificial pancreas, 2803
 - insulin delivery systems, 1518–1519
- Microfabrication
- biosensors, 345
 - cell-material interactions, 572
 - microvascular tissue engineering, scaffold fabrication, 2794–2798
 - topographical surface analysis, 2552
- Microfluidics, polyelectrolyte multilayer deposition, 2221
- flow control, 2226
- Microinjection techniques, transdermal drug delivery, 2902
- Micromechanics
- composites, 659–661
 - ordered and disordered composites, 666–671
 - tissue engineering, 2780
- Microorganisms
- biofilm development, 207
 - biosensor development, 344–345
 - magnetic/electrophoretic cell separation, 1702–1703
- Microparticle image velocimetry
- basic principles, 1873–1876
 - correlation function, 1881–1882
 - flow field measurement, 1876–1879
 - DEP particle trap device, 1876–1878
 - microviscosimetry, 1878–1879
 - overview, 1873
 - particle detection, 1879–1881
 - concentration measurement, 1879–1880
 - diffusion measurement, 1880–1881
- Microporous materials
- artificial pancreas, 2803–2804
 - bioactive materials, 1886
 - ceramics, 1886–1887
 - composites, 1890–1891
 - foaming of ceramics and glasses, 1888–1889
 - melt-derived bioactive glasses, 1890
 - overview, 1885
 - polymer foams, 2270–2274
 - scaffold applications, 1885–1886
 - sol-gel bioactive glasses, 1887–1888
 - tissue engineering applications, 1885
 - transdermal drug delivery, 2902
- Microprocessing units, implantable
- cardioverter-defibrillators, 776
- Micropumps, controlled release drug delivery, 741
- Microscale tissue engineering
- cardiovascular system, 2785–2787
 - chemical control, 2782
 - control systems, 2781–2782
 - future research issues, 2788–2789
 - immunological control, 2782–2783
 - liver, 2784–2785
 - micromachining technology, 2780
 - multidimensional techniques, 2789
 - nerve tissue, 2787–2788
 - overview and historical background, 2779–2780
 - pancreatic tissue, 2783–2784
 - soft lithography, 2780–2781
 - spatial control, 2782
 - surface engineering, 2780
 - topographical control, 2781–2783
 - toxicity testing, animal surrogate systems, 57–58
 - in vitro/in vivo procedures, 2789
- Microscopy
- confocal microscopy, 705
 - fast confocal microscopy, 712–713
 - fluorophores, 711
 - incident light intensity, 710–711
 - living cells, 712
 - multicolor fluorescence, 712
 - Nipkow disk, 713
 - photobleaching, 711–712
 - pinhole size, 710
 - principles and instrumentation, 705–706
 - procedures and applications, 707–709
 - resolution, 710
 - three-dimensional visualization, 709–710
 - two-photon microscopy, 713
 - electron microscopy
 - cardiovascular tissue engineering, 974–975
 - hard biomaterials and polymer specimens, 972–973
 - instrumentation, 968–969
 - medical applications, 974
 - overview, 968
 - scanning vs. transmission techniques, 969–970
 - SEM specimen preparation, 970–972
 - TEM specimen preparation, 973–974
 - fluorescence microscopy, 706
 - magnetic resonance
 - atherosclerotic plaques, 1746–1747
 - biological cell visualization, 1745–1746
 - cells and gene expression, 1744–1745
 - defined, 1738
 - future research issues, 1747
 - high-temperature superconduction coils, 1740–1741
 - receive coils, 1741
 - histology and amyloid plaque characterization, 1742–1744
 - molecular imaging, 1746
 - mouse phenotyping, 1741–1742
 - signal-to-noise ratio, 1738–1741
 - two-photon microscopy, 712–713
 - Microsphere sintering, bioactive scaffold fabrication, 149
 - Microstructure analysis
 - aortic valve cusps, 2738–2742
 - dental implants, 811–812
 - Microtextured titanium (MTX), dental implants, 814
 - Microthermal effects, radiofrequency electromagnetic field/radiation, 257
 - Microtopography, cell-material interactions, 571–572
 - surface analysis, 2556
 - Microtubules, mass transfer and tissue function, 1752–1754
 - Microvascular tissue engineering
 - assembly techniques, 2796–2797
 - biocompatible polymer molding, 2796
 - biodegradable polymer fabrication, 2796
 - endothelial cell seeding, 2798
 - lithographic techniques, 2794–2795
 - magnetic resonance microscopy, 1745–1746
 - microcirculatory oxygen transport, 1842–1848
 - organ and tissue properties, 2792–2793
 - plasma etching, 2795–2796
 - poly(glycerol sebacate), 333
 - polymer replica molding, 2796
 - porous devices, 2797–2798
 - scaffold fabrication, 2794–2798
 - vascular fluid dynamics, 2793–2794
 - vascularization requirements, 2792
 - whole organ engineering, 2798–2799
 - Microviscosimetry, microparticle image velocimetry, 1878–1879
 - Microwave frequencies
 - proton resonance frequency shift, 1733–1734
 - tissue dielectric properties, 841
 - Migration, telemedicine systems, 2584–2585
 - Migration rates and mechanisms, endovascular stent graft fixation, 2488
 - Mimicry
 - natural bone healing, 1261
 - tissue engineering, blood vessels, 2645
 - Mineralization
 - bone, 2041–2042
 - healing process, 1258–1259
 - histiogenesis, 1408–1409
 - crystal dissolution, osteoclasts, 2082–2083
 - tissue structure and classification, 2039–2049
 - Miniaturization. *See* Microelectromechanical systems (MEMS)
 - Minimally invasive surgery (MIS)
 - computer-assisted surgery, 687–693
 - robotics applications, 1807–1808
 - Minimum discernable width (MDW), electrical impedance imaging, 920–923
 - Mitochondrial transport, mass transfer and tissue function, 1754–1755
 - Mitral valve
 - defects, prosthetic heart valves, 1293
 - hemodynamics, 1323

- Mitral valve (*cont.*)
 mechanical properties, 1323–1326
 structure and function, 1320–1323
- Mobile-bearing insert components, total knee replacement, 1611–1612
- Mobility assistance devices, robotics, 1804
- Mobilized peripheral blood (MPB), as hematopoietic stem cell source, 1338–1341
- Modality, three-dimensional imaging techniques, 1800
- Modeling, biomedical
 animal models
 bone tissue engineering, 33–45
 toxicity testing
 barrier models, 53–55
 cell culture analog systems, 57–58
 chemical structure-related models, 51–52
 coculture models, 55
 computational models, 51–53
 conditioned medium systems, 56–57
 microscale designs, 57–58
 multiple cell devices, 56–57
 overview, 51
 shared medium systems, 57
 single cell *in vitro* models, 53
 three-dimensional models, 55–56
 whole organism models, 52–53
 elastin structural properties, 891–893
 finite element analysis, 1091–1098
 hard materials, elastic behavior modeling, 1242–1243
 insulin delivery, 1520
 predictive control algorithms, 1521
 movement control, 732–734
 protein adsorption modeling
 biomaterial properties and surface analysis, 2326–2328
 biomolecular-material composites, 2326–2327
 molecular modeling, 2327
 physicochemical phenomena, 2327–2328
 overview, 2326
 simulation modeling, 2328–2333
 continuum modeling, 2328–2329
 forcefield-based molecular mechanics and dynamics, 2330–2333
 quantum mechanics, 2330
 solvent properties, 2329–2330
 rapid prototyping
 basic principles, 2382–2383
 fused deposition modeling, 2387–2388
 future research issues, 2388–2389
 selective laser sintering, 2384–2386
 stereolithography, 2383–2388
 three-dimensional printing, 2386–2387
- Moens-Korteweg velocity
 macrocirculatory hemodynamics, arterial pulse waves, 1364–1365
 vascular fluid dynamics, 2974–2975
- Moisture effects, zirconia ceramics, 3106–3108
- Molar mass, polymer materials, 2281
- Mold arthroplasty, total hip replacement, 1470–1471
- Mold release agents, alginate applications in, 354
- Molecular absorption/extinction coefficient, tissue optics, 1624
- Molecular complexes
 cardiovascular system, mathematical modeling, 1897–1898
 mass transfer and tissue function, 1752
 supercritical fluid processing, 2524–2525
- Molecular diagnostics, evolution of, 2906
- Molecular diffusion coefficient, MRI thermometry, 1728–1729
- Molecular dynamics (MD), protein adsorption modeling, 2330–2333
- Molecular imaging
 magnetic resonance imaging, 1725
 magnetic resonance microscopy, 1746
 positron emission tomography, 2321
- Molecular mass
 interfacial transport, biomaterials artificial membranes, 2350–2351
 permeability and, 2354
 pH effects, 2351–2352
 protein permeability and, 2354
 protein-surfactant interactions, 2341–2343
- Molecular modeling, protein adsorption modeling, 2327
- Molecular weight distribution (MWD)
 degradable polymer composites, 780
 polymer synthesis, 2279–2280
- Molecule Equivalents of Soluble Fluorescein (MESF) units, flow cytometry, 1108–1109
- Moment arm modeling, gait analysis, 1137
- Moments, digital signal acquisition and processing, higher-order spectra, 862–863
- Monitoring systems, interstitial fluid sensors, 1175–1178
- Monoclonal antibodies, osteogenic progenitor cells, 2092–2093
- Monocyte infiltration, vascular grafts, host reactions, 3007–3008
- Monolayer assembly
 hemocompatibility assessment, 1351–1352
 tissue engineering, cardiac muscle, 2687–2688
- Monomers
 bone cement, 403
 residual monomers, 407
 cyanoacrylate glues, 1188
 inherently conductive polymers, 697–698
 poly(lactide-*co*-glycolide) degradation, 2264–2265
 polytetrafluoroethylene, 2301
 silicone hydrogel contact lenses, 715–717
 hydrophobic comonomers, 721
 silicones, 2427–2429
- Mononuclear leukocytes, acute inflammation, 1411
- Monopole point source, dielectric and conductive media, 1010–1011
- Mono-substituted ethylenes, structure/property relationships, 2276–2278
- Monte Carlo simulation, optical detection of cancer, 2010–2011
- Moore's law, mathematical modeling, 1895–1896
- Morphological analysis
 blood vessel mechanics, 395
 orthopedic device failure analysis, 1832–1837
 vascular prosthetics, *in vivo* animal models, 1567–1568
- Morphological fixation, bioactive glass, 136
- Morphometrics, tissue engineering, 2623
- Morphotropic phase boundary, piezoelectric ceramics, 2167
- Motility assessment, neutrophil-biomaterial adhesion, 1940
- Motion artifacts, optical mapping, 2021–2022
- Motion-induced contrast, magnetic resonance imaging, 1718–1719
- Motor unit action potential (MUAP), electromyographic signal, 959
 quantification and processing, 964–965
- Motor unit (MU), electromyographic signals, 956–959
- Movement control
 biological control, 729–733
 future research issues, 737–738
- hierarchical hybrid control, 735–737
 man-made control, 732–734
 overview, 729
 peripheral nerve injury and repair, 2122
- MTT test, cell culture assays, 564–565
- Mucoadhesive biomaterials, basic properties, 3–5
- Mucopolysaccharidoses (MPS), glycosaminoglycan disorders, 1202–1204
- Mucous membranes, bioadhesive bond, 153–155
- Multiaxial mechanical testing, tendons and ligaments
 anisotropic properties, 2592–2593
 bidirectional testing, 2593–2594
 shear loading, 2594–2595
- Multiaxial warp (MWK) fabric, structure and applications, 1051–1052
- Multicolor fluorescence, confocal microscopy, 711
- Multicomponent protein solutions, adsorption behavior, 276–278
- Multidimensional tissue engineering, microscale technology, 2789
- Multifunctional bioadhesives, 155
- Multilayer assemblies, polyelectrolytes biomolecules, cell adhesion/encapsulation, 2223–2226
 characterization, 2221–2222
 deposition techniques, 2219–2221
 formation mechanisms, 2218–2219
 light emitting diodes, 2222–2223
 materials, 2219
 microfluidics flow control, 2226
 overview, 2217
 separations, 2223
 structural properties, 2217–2218
- Multilayered perceptrons (MLPs)
 algorithm of pattern extraction training, 125–126
 feed forward networks, 122–123
 neural networks, 119
- Multiparallelism, microelectromechanical systems, 1866–1867
- Multiphoton imaging
 biomedical optics, 2036
 cancer diagnosis and therapy, 2015
- Multiple cell models, toxicity testing, animal surrogate systems, 56
- Multiple-disk centrifugal blood pump, 577–578
- Multiple myeloma, angiogenesis inhibitors, 28–29
- Multiplet patterns, nuclear magnetic resonance, 1962–1963
- Multistage cell separation
 electrophoretic methods, 1699–1700
 magnetic cell separation, 1696–1697
 overview, 1695
- Muscle activation
 cardiac muscle mechanics
 force-length curves, isolated muscle, 1282–1283
 myocardial stress analysis, 1283–1284
 electrical stimulation, 1017
 electromyographic signals, 956–967
 fiber potential, 959
 ergonomics and work-related disorders, 999–1001
 gait analysis, 1137
 motor unit, 956–959
 movement control, 729–733
 models of, 734
- Muscle-derived stem cells (MDSC), tissue engineering, skeletal muscles, 2840
- Muscle glue protein, basic properties, 159–160
- Muscles. *See* Artificial muscles

- Muscle tissue
 histogenesis, 1404–1405
 stress fractures and, 2513
- Musculoskeletal system
 ergonomics and work-related disorders, 999
 low back, 1000–1001
 upper extremity, 1001
 finite element analysis, 1098
- Mussel glue, 240–241
 wound healing and tissue repair, 247–248
- Mutagenicity, biocompatibility testing, in-vitro testing, 172
- Myelinated axon activation, electrical stimulation, 1013–1014
- Myocardial infarction (MI)
 distal protection device reduction of, 878
 electrocardiography, 942
 positron emission tomography imaging, 2320
- Myocardium
 material properties, 1284–1286
 stress analysis in, 1283–1284
- Myocytes
 cardiac patch engineering
 collagen interaction, 545
 mechanical signals, 545–546
 rotating-wall bioreactor, 547–548
 phenotype regulation, cardiac patch engineering, 545
- Myopia, defined, 747–748
- N**
- N-acetyl aspartate (NAA), brain imaging, nuclear magnetic resonance, 1967–1970
- Naked DNA
 gene-activated matrix, 1144–1145
 gene therapy and, 1151
- “Nanoclinics,” magnetic nanoparticles, 1708–1709
- Nanotechnology
 biocomposites, 192–193
 chitosan
 controlled release systems, 606
 gene expression, 609–611
 composite materials
 nanomechanics, 660–661
 polymer composites, 785
 electrospun fibers
 coated nanofibers, 983
 future research issues, 983–984
 medical applications, 982–983
 nanofiber fabrication, 978–982
 nonwoven nanofiber sheets, 979–980
 overview, 978
 yarns and fabrics, 1053–1054
 fiber optic nanobiosensors, 1068
 hydroxyapatite coatings, 1467–1468
 magnetic nanoparticles
 bioseparation, 1709–1710
 coprecipitation, 1706–1707
 disease therapy, 1710–1711
 homogeneous preparation, 1706
 magnetic resonance imaging, 1709
 microemulsion formation, 1710
 nonaqueous colloidal precipitation, 1707
 overview, 1705
 pyrolysis, 1710
 superparamagnetism, 1705–1706
 microelectromechanical systems, 1866–1870
 movement control applications, 737–738
 multilayer assemblies, polyelectrolytes, 2217–2226
 neurosurgery
 brain submicron structure, 1903–1906
 extracellular space, drug delivery, 1906–1908
 future research issues, 1909
 overview, 1903
 tumor cell invasion and axonal cell extension, 1908–1909
 polymer processing, 2281
 supercritical fluids, drug nanoparticle preparation, 2526
 tissue engineering, polymer sources, 2295
 topographical surface analysis, 2556
 Narrow band imaging (NBI), basic principles, 993
- Nasal insulin delivery, 1515
- NASPE/BPEG coding system, cardiac pacemakers, 536–538
- National Institute for Occupational Safety and Health (NIOSH), lifting equation, ergonomics guidelines and, 1002–1003
- National Research Act of 1974, biomedical ethics and, 286–287
- Natural killer cells, host reactions, 1417
- Natural materials
 apatites, 226–229
 articular cartilage repair, 555–556
 bioadhesives
 aquatic adhesives, 240–245
 bond strengths, 237–240
 overview, 236
 terrestrial adhesives, 245–247
 wound healing and tissue repair, 247–249
 biocomposites, 181–182
 biofilm formation, 207–215
 burn dressings, 451–452, 457–458
 cell-material interactions, 570–571
 collagen fibers
 cross-linking compounds, 648–649
 fixation applications, 640–642
 composite compounds, 672–673
 microelectromechanical systems, 1865–1866
 polyamides, 2186–2188
 polymers, tissue engineering applications, 2289–2293
 silks
Bombyx mori silk
 biosensors, 2445–2446
 chemical properties, 2445
 fiber formation, 2443–2444
 mechanical properties, 2445
 structural properties, 2442–2443
 sutures, 2445
 tissue engineering scaffolds, 2446
 spider silks, 2448
 structural properties, 2442
 wild silkworm silks, 2447–2448
 tissue engineering
 esophagus, 2721
 fat tissue, 2729–2730
 vascular system, 3034–3035
 valvular prosthetics, calcification, 75–76, 81
 Navier-Stokes equation
 macrocirculatory hemodynamics, pulsatile flow properties, 1361–1364
 vascular fluid dynamics, 2974–2975
 NB hydrogel, soft contact lenses, 1442–1443
 Near-infrared (NIR) absorption spectroscopy sensor, blood glucose monitoring, 1182
 Near lossless compression, digital biomedical signals, 677–678
 Needleless injection systems, transdermal drug delivery, 2902
 Needles, sutures and, 2562–2563
 Needle-type, enzyme-based electrochemical biocatalytic sensors, properties and applications, 1175–1178
 Neighboring rights, intellectual property, 1545
 Neobone[®] hydroxyapatite, 463
 Neonatology, nitric oxide inhalation therapy, 1953–1954
 Nernst equation, mass transfer and tissue function, 1761
- Nerve growth factor (NGF), tissue engineering, skeletal muscles, 2845–2848
- Nerve guides
 basic properties, 1911
 biohybrid guides, 1916–1918
 alginate hydrogels, 1917
 extracellular matrix, 1916–1917
 growth factors, 1917
 support cells, 1917–1918
 degeneration and regeneration, 1912–1913
 economic impact, 1920
 future developments, 1920–1921
 human applications, 1919–1920
 nerve tissue engineering, 2814–2815
 peripheral nervous system, 1911–1912, 2127–2130
 physical properties, 1918–1919
 principles and techniques, 1914
 resorbable guides, 1915–1916
 silicone tube model/nonresorbable guides, 1914–1915
 sutures and transplantations, 1913–1914
- Nerve tissue. *See also* Peripheral nerves
 gene-activated matrix regeneration, 1147–1148
 hyaluronan applications, 1429
 mathematical modeling, historical background, 1895–1896
 plasticity, tinnitus device design, 2617
 polycaprolactone grafts, 2202
 rotary cell culture systems, 2406
 tissue engineering, 2622
 acellular grafts, 2815
 axonal degeneration, 2812–2813
 biomimetic strategies, 2816
 functional recovery assessment, 2813–2814
 gene therapy, 2817
 microscale technology, 2787–2788
 nerve guidance channels, 2814–2815
 nerve regeneration strategies, 2814–2816
 nerve tissue degeneration, 2812
 nervous system anatomy, 2811–2812
 overview, 2811
 stem cells, 2816
 support cells and neurotrophic factors, 2815–2816
- Net positive suction head (NPSH), centrifugal blood pumps, 574–575
- Network systems, telemedicine applications, 2582–2584
- Neural networks (NNs)
 ALOPEX project, 124–125
 multilayer perceptron training, 125–126
 VLSI implementation, 126–128
 basic properties, 119–120
 ECG signal compression, 681
 error back propagation, 121–123
 feed-forward networks, 120–121
 training methods, 123–124
 universal approximation, 126
- Neural stem cells (NSCs), biomaterials and research on, 2476
- Neurapraxia, mechanisms of, 1912–1913
- Neurodegenerative disease, artificial pancreas tissue engineering, 2807
- Neuroglia, structure and function, nanoscale neurosurgery and, 1905–1906
- Neurolinguistic programming (NLP), eye movement research and, 1036
- Neurological function
 eye movements and, 1035–1036
 functional magnetic resonance, 1119–1127
 magnetic resonance imaging, 1720–1722
 movement control, 729–733
 neuromorphic modeling, 130
- Neuromorphic modeling, neural networks, 129–130

- Neuromuscular junction (NMJ), electromyography, 957–959
- Neurons, structure and function, nanoscale neurosurgery and, 1903–1906
- Neuroprostheses
 basic properties, 1924
 bladder management, 1929–1930
 cochlear implants, 1930–1931
 electrical stimulation, 1009
 historical background, 1926
 physiology, 1924
 reaching and grasping applications, 1927–1929
 technology, 1924–1926
 walking applications, 1926–1927
- Neuroshield (Mednova) distal protection device, 874–875
- Neurosurgery
 fibrin sealant applications in, 1084
 functional magnetic resonance imaging, 1123–1125
 nanoscale techniques
 brain submicron structure, 1903–1906
 extracellular space, drug delivery, 1906–1908
 future research issues, 1909
 overview, 1903
 tumor cell invasion and axonal cell extension, 1908–1909
- Neurotmesis, mechanisms of, 1912–1913
- Neurotrophic factors, nerve tissue engineering, 2815–2816
- Neutral hydrogels, drug delivery systems, 1163–1164
- Neutral salt-soluble collagen, degradation, 635
- Neutrophils
 adherent protein structure, 1935
 adhesion mechanisms, 1935–1936
 bacterial killing, 1940
 degranulation, 1936–1937, 1940
 evaluation, 1938–1941
 functional changes, 1937–1938
 modification of, 1936
 molecule expression, 1939
 motility, 1940
 oxidative burst, 1936, 1940
 phagocytosis, 1940
 viability, 1941
 biomaterial interactions, 1934–1935
 collagenase, 1767–1768
 electrical cell impedance sensing, 912–913
 flow cytometry, 1107–1108
 foreign body response, 1113
 granulocyte formation, 1934
 lifespan, 1938
 thrombosis
 fibrin formation and adhesion of, 2604–2605
 function, 2602
 homotypic interactions, 2602–2603
 platelet-dependent coagulation, 2606–2607
 platelet-neutrophil heterotypic interactions, 2603
 vascular grafts, host reactions, 3007–3008
- Newington model, gait analysis, 1131–1133
- Newtonian equation, magnetic nanoparticles, 1709–1710
- Newtonian fluids
 blood viscosity mechanics, 1357–1358
 vascular fluid dynamics, 2977–2978
- N*-hydroxysuccinimide (NHS) coupling agents, surgical glues, 1190–1191
- Nickel and nickel alloys
 chemistry, 1942
 hypersensitivity reactions, 1948
 nickel-chromium, 1942–1943
 nickel-titanium, 1945–1948
- stainless steel, 1943–1945
 titanium alloys, orthodontic wires, 2057–2058
- 96-well plate assay, toxicity testing, animal surrogate systems, 53–55
- Nipkow disk, confocal microscopy, 712
- Nitinol, endovascular stent-grafts, 2486–2487
- Nitric oxide (NO)
 in adults, 1954
 biomedical engineering applications, 1955
 bone remodeling, osteocyte apoptosis, 444
 cellular toxicity, 1953
 gas infusion, 1955
 inhalant physiology, 1952
 lung surfactant dysfunction, 1688
 mechanism of action, 1951–1952
 microcirculatory oxygen transport, 1847
 neonatology applications, 1953–1954
 no-generating biomaterials, 1955
 in pediatric patients, 1954
 polymer biomaterials, 221–222
 pulmonary vasodilation, 1952–1953
 synthesis and chemistry, 1951
 tissue engineering, 1955–1956
- Nitrogen, plasma implantation, 2177–2178
- N,N*-dimethyl-*p*-toluidine (DMPT), bone cement, 403–410
- Noble's test, iliotibial band friction syndrome, 1601–1602
- Noise artifacts, optical mapping, 2022
- Nonaqueous colloidal preparation, magnetic nanoparticles, 1707
- Nonaqueous enzymology, biocatalysis and, 166
- Non-collagenous proteins (NCPs)
 bone matrix composition, 2044
 valvular prosthetics, calcification, 78
- Noncontact testing, tendon/ligament tensile testing, 2590–2592
- Nondegradable polymers, tissue engineering, 2284–2285
- Nonfouling surfaces, protein adsorption resistance, 278–280
- Nonfunctional testing, in-vivo biocompatibility testing, 173
- Nonionic polymer gels/electroactive polymers, artificial muscles, 109
- Nonlinear biomedical signals
 digital signal acquisition and processing, 864–867
 applications, 866–867
 approximate entropy, 865
 detrended fluctuation analysis, 864–865
 regularity index, 865–866
 electrical impedance tomography, 919
 ultrasound-tissue interaction, 2939–2941
- Nonlinear function approximation, neural networks, 123
- Nonparametric analysis, biomedical signal acquisition and processing, 849
- Nonporous polymer-ceramic composites, structure and properties, 784–785
- Nonresorbable materials
 biocomposites, 182–184
 nerve guides, 1914–1915
 suture classification and configuration, 2563–2568
- Nonstationary signals, digital biomedical signal acquisition and processing, spectral analysis, 858–861
- Nonthermal effects, radiofrequency electromagnetic field/radiation exposure, 256–257
- Nonwoven fabrics, structure and properties, 1052–1053
- Nordihydroguaiaretic acid (NDGA), collagen cross-linking, 649
- No-reflow phenomenon, distal protection devices, 870
- Normative principles, biomedical ethics and, 287
- Nosocomial infection, biofilm formation and, 207–208
- Novacor left ventricular assist device, 1778–1779
- Novelty, intellectual property rights and, 1540
- Nuclear magnetic resonance (NMR)
 biological studies, 1966–1967
 brain imaging, 1967–1970
 proton spectra information, 1969–1970
 in vivo spatial localization, 1967–1969
 water and lipid suppression, 1969
 chemical shift, 1962
 historical background, 1960
 intravascular ultrasound imaging, 1561–1564
 pulsed Fourier transform spectroscopy, 1963–1964
 rotating frame experiments, 1964–1966
 relaxation times, 1964–1965
 signal-to-noise ratio, 1961
 spectrum characteristics, 1962–1963
 spin echoes, 1965–1966
 spin-spin coupling, 1962–1963
 theoretical principles, 1960–1961
 three-dimensional imaging, 1794
- Nuclear matrix, surfactant properties, 2338
- Nuclear tomographic imaging, 1794–1797
- Nucleic acids, mimetic materials, 297
- Nucleus Spectral Maxima (SPEAK) sound processor, cochlear implants, 1509–1510
- Numerical analysis
 dental implants, 804–805
 finite element analysis, 1094–5
 periodontal ligament, biomechanics, 801–802
- Nutrient transport, ligaments, tissue engineering, 2765
- Nylon
 basic properties, 2188–2189
 cardiac catheters, 496
 piezoelectric materials, 2169
- Nyquist frequency, positron emission tomography, 2317
- O**
- O₂Optix™ silicone hydrogel contact lenses, 727
- Oasys™ silicone hydrogel contact lenses, 727
- Ober's test
 iliotibial band friction syndrome, 1602
 snapping hip syndrome, 1398–1399
- Occlusal loading, dental implant biomechanics, 803–804
 zirconia ceramics, 3107–3108
- Occupational Safety and Health Administration (OSHA), ergonomics guidelines, 999–1000
- Ocular angiogenesis, angiogenesis inhibitors, 28
- Ocular implants. *See also* Contact lenses; Corneal implants
 corneal implants
 future research issues, 751
 glossary of terms, 747–748
 keratoprosthesis, 750–751
 overview, 747
 refractive keratoplasty, 748–750
 drug delivery systems, 1979
 basic principles, 1982–1983
 bioadhesives, 157–158
 future research issues, 1992
 insulin delivery, 1515
 matrix implants, 1983–1987
 inserts, 1984
 scleral plugs, 1984–1985

- miscellaneous designs, 1986–1987
 oculex designs, 1985–1986
 overview, 1981
 reservoir implants, 1987–1992
 codrugs, 1991
 corticosteroids, 1989–1990
 cyclosporine, 1990–1991
 live cells, 1991–1992
 osmotic pump systems, 1992
 vitrasert, 1988–1989
 glaucoma filtration implants, 1976–1977
 intraocular lenses, 1977–1978
 keratoprosthesis and intracorneal implants, 1974–1975
 orbital implants, 1974
 phosphorylcholine technology, 2145–2146
 scleral buckles, 1976
 silicone oil, 1979
 viscoelastic agents, 1979
 Oculex matrix implants, drug delivery systems, 1985–1986
 Oculometry, eye movement research, 1030–1031
 Ohm's law
 electrical impedance tomography, 915–916
 electrophoretic cell separation, 1702
 Oil-in-water technique, supercritical fluid synthesis, 2527–2528
 Oligodendrocytes, brain structure, nanoscale neurosurgery, 1906
 Oligo(ethylene glycol) (OEG), adsorption resistance of, 278–280
 Omniflow Vascular Prosthesis (OVP), basic properties, 1434–1436
 On-eye movement prediction, silicone hydrogel contact lenses, 722
 Onsager relationship, interfacial transport, 2350
 Open-loop controllers, movement control, 732–734
 Ophthalmology
 diamond/diamond-like carbons, 829
 fibrin sealant applications, 1086
 hyaluronan applications, 1428
 Opsonization
 acute inflammation, 1411
 valvular prosthetic failure, 1298–1299
 Optical coherence tomography (OCT)
 applications, 2003–2004
 biomedical optics, 2036–2037
 cancer detection, 2013–2014
 endoscopic techniques
 basic principles, 986
 image resolution, 989–990
 medical applications, 995–996
 future research issues, 2004–2005
 imaging probes, 2001–2002
 imaging resolution, 1998–2000
 imaging speed, 2000–2001
 operating principles, 1996–1998
 overview, 1996
 spectroscopic, Doppler, and polarization techniques, 2002–2003
 Optical mapping
 biological electrical activity
 artifacts, 2021–2022
 dye pharmacological effects, 2021
 fluorescence principles, 2019
 fluorophore instrumentation, 2019–2020
 light collection systems, 2020–2021
 light sources, 2020
 mechanical motion artifacts, 2021–2022
 noise artifacts, 2022
 overview, 2019
 cardiac cell culture, 2025–2026
 experimental applications, 2024–2026
 eye movements, basic principles, 1028–1031
 future research issues, 2026–2027
 organ mapping, 2024
 photodiode array optical system design, 2022–2024
 tissue isolation, 2024–2025
 Optical transducers, fiberoptic biosensors, 1064
 Optics, biomedical
 biosensors, 344–345
 biological recognition elements, 339–343
 detector elements, 344–345
 cancer diagnosis and therapy
 confocal microscopy, 2015
 continuous-wave imaging studies, 2011
 diffuse optical tomography, 2013
 fluorescence method, 2014–2015
 frequency-domain measurements, 2012
 optical coherence tomography, 2013–2014
 overview, 2009
 radiative transfer method, 2009–2014
 Raman spectroscopy and multiphoton techniques, 2015
 spectroscopy measurements, 2011–2012, 2014–2015
 time-resolved measurements, 2012–2013, 2015
 eye movement research, 1028–1031, 1033–1038
 fiberoptic biosensors
 advantages and limitations, 1064–1065
 analytes, 1064
 applications, 1065–1071
 classification, 1065
 combined fiberoptic biosensors, 1069–1071
 components, 1063–1064
 delivery/sensing functions, 1063–1064
 design principles, 1062–1063
 detector, 1064
 future research issues, 1071
 laser/light source, 1063
 optical transducer, 1064
 overview, 1062
 flow cytometry
 applications, 1106–1109
 calibration issues, 1108–1109
 cell biology, 1106–1108
 cell sorting, 1104–1106
 detectors, 1103–1104
 electronic systems, 1103–1104
 fluidic systems, 1101–1102
 microbiology, 1108
 optical systems, 1102–1103
 overview, 1100–1101
 pharmaceutical applications, 1108
 plant and animal science, 1108
 reproductive medicine applications, 1108
 spectral compensation, 1104
 imaging applications, 2035–2037
 laser-tissue interaction, 1623–1625
 biological effects, 1630
 light-tissue interaction, 1625–1626, 2029–2030
 overview, 1623
 photochemical interactions, 1626
 photomechanical interaction, 1628–1630
 photothermal interactions, 1626–1628
 tissue optics, 1623–1625
 light propagation in tissue, 2033–2034
 light properties, 2029–2030
 optical coherence tomography, 1996–2005
 optical mapping, 2019–2025
 overview, 2029
 silicone hydrogel contact lenses,
 contact1–1–13
 spectroscopic techniques, 2034–2035
 tissue properties, 2030–2031
 transmynocardial laser revascularization, 1632–1637
 visual impairment robotics, 1812–1813
 Optimal frame configuration, external fracture fixation, 421
 Optimization techniques
 gait analysis, 1137–1139
 three-dimensional imaging, 1800
 Optrode technology, optical mapping, 2021
 Oral drug delivery
 bioadhesives, 157–158
 enteric insulin delivery, 1515
 protein-based drug delivery systems
 hydrogel oral delivery systems, 2363–2364
 insulin delivery, 2364–2366
 Oral environment
 dentistry ceramics, 594
 fibrin sealant applications in, 1084
 Orbital implants
 matrix ocular implants, 1984
 structure and function, 1974
 Ordered composites, micromechanics, 666–671
 Orders of magnitude, mass transfer and tissue function, 1759–1760
 Organic matrices
 bone
 formation mechanisms, 2045
 functional properties, 2044–2045
 mineral composition, 2041–2042
 organic matrix composition, 2042–2044
 structure and function, 2039–2041
 tissue engineering, 2045–2047
 osteoclast degradation, 2083
 teeth
 dentin composition and formation, 2047–2048
 dentin tissue engineering, 2048
 enamel composition and formation, 2048
 enamel tissue engineering, 2049
 Organoid units, small intestine tissue engineering, 2864–2866
 Organosilicon derivatives, self-assembled monolayers, 2411–2412
 Organ systems. *See also* Artificial organs
 allografts, donation, 14–15
 functional units, 1751
 mass transfer and tissue function, 1749–1750
 mathematical modeling, 1896
 cardiovascular system, 1899–1900
 optical mapping, 2024
 Organ transplants
 cell-material interactions, 568
 microvascular tissue engineering, 2792–2799
 polyepoxy compounds, 2231
 transgenic/gene-targeted organs
 acute vascular rejection, 2909–2910
 genetic engineering, 2910–2912
 hyperacute rejection, 2908–2909
 immune response, 2907–2910
 overview, 2906
 T cell-mediated rejection, 2910
 xenotransplantation, 2906–2907
 xenografts, 3095–3100
 Orientation mechanisms
 cell surface analysis, 2557
 polymers, 2279
 Orthodontic devices
 nickel alloys, 1942–1949
 shape memory metals, 2420
 wire systems
 beta-titanium, 2057
 cobalt-chromium alloys, 2057
 nickel-titanium alloys, 2057–2059
 overview, 2054–2056
 stainless steel, 2056–2057
 strengths and limitations, 2059
 Orthognathic surgery, bioabsorbable plates and screws, 434–435

- Orthogonal three-dimensional fabrics, 1050–1051
- Orthopedic devices and materials. *See also* Organic matrices
- biomaterials
 - ceramics and glasses, 2066–2068
 - cobalt-chrome alloys, 2062
 - composites, 2068–2069
 - metals, 2061
 - polymers, 2063–2066
 - porous coatings, 2063
 - stainless steels, 2061–2062
 - tantalum, 2063
 - titanium alloys, 2063
 - titanium-nickel alloys, 2063
 - bioreactors, 311, 320–324
 - chamber modification, 321–322
 - gripping and attachment issues, 322–323
 - material/construct properties, 323–324
 - measurement instrumentation, 323
 - mechanical stimulation subassembly, bio311r–11–12
 - diamond/diamond-like carbons, 827–828
 - failure analysis
 - anterior cervical spine plate, 1832–1834
 - dynamic compression plate, 1829–1832
 - intramedullary nails, 1835–1836
 - mandibular reconstruction plate, 1837–1839
 - prostheses, 1839–1841
 - tibial tray, 1841
 - titanium screws, 1834–1835
 - fracture fixation
 - absorbable devices, 2073
 - external devices, 2077–2078
 - internal devices, 2073–2077
 - intramedullary nails, 2076
 - plates, 2075–2076
 - rods, pins, and wires, 2074–2075
 - screws, 2075
 - soft tissue-bone attachment, 2076–2077
 - hyaluronan, 1427
 - metallic materials
 - bone biology, structure and properties, 1824
 - failure analysis, 1824
 - fracture fixation and joint replacement, 1824–1827
 - fracture fixation failure analysis, 1827–1839
 - prostheses failure analysis, 1839–1841
 - nickel alloys, 1942–1949
 - particulate matter and host reactions
 - aseptic loosening, 2110–2111
 - future research issues, 2114–2115
 - macrophage detectors, 2111–2114
 - osteoclast/osteoblast effector cells, 2114
 - overview, 2110
 - poly(glycolic acid), 2244–2245
 - pyrolytic carbon prostheses, 2374–2375
 - shape memory metals, 2420
 - soft tissue allografts, 13–14
 - total hip arthroplasty, 1470–1481
 - ultra-high molecular weight polyethylene, 2925–2930
- Orthotics, robotics, 1805
- Orthotopic liver transplantation (OLT), clinical setting, 1649–1650
- Osferion[®] tricalcium phosphate, 463–464
- Osgood-Schlatter disease, overuse injuries, 1603
- Osmosis
 - controlled-release drug delivery, 740–741, 884
 - hemodialysis, 381–382
- Osmotic pump systems, reservoir ocular implants, 1992
- Osseointegration
 - bone-implant interface, 424
 - dental implants, 803, 809–810
- Ossification
 - biomineralization, 72
 - bone formation and, 2045
 - bone healing, soft callus formation, 1256–1258
- Osteitis pubis, overuse injuries, 1395
- Osteoarthritis
 - fibroblast mechano-regulation mechanisms, 1783–1790
 - glycosaminoglycans efficacy, 1204–1205
 - ligament injury and repair, animal repair models, 1643
- Osteoblasts
 - host reactions, particulate matter, 2114
 - wear debris, joint arthroplasties, 3078
- Osteoclasts
 - basic properties, 2081
 - degradation mechanisms, 2082–2084
 - differentiation and activation signaling, 2084–2087
 - functional secretory domain, 2082
 - host reactions, particulate matter, 2114
 - inactivation/apoptosis, 2084
 - mineral crystal dissolution, 2082–2083
 - morphology, 2081–2082
 - remodeling applications, 2088
 - ruffled border, 2082
 - sealing zone, 2081–2082
 - wear debris, 3068–3072, 3078
- Osteoconductivity/osteoinductivity
 - apatite composites, 232–233
 - biphasic calcium phosphate, 364–365
 - bone-implant interface, 428
 - bone tissue engineering, animal models, 33–34
 - coating surface properties, 2536
 - reconstructive biomaterials, 2401
 - substances and materials
 - biodegradable ceramics, 2105
 - bioglasses, 2104
 - bone morphogenetic protein, 2102–2103
 - bone proteins, 2103
 - demineralized bone, 2104–2105
 - phosphorhorn, 2103–2104
 - porous tricalcium phosphate tissue engineering, 2105–2106
- Osteocytes, bone remodeling
 - loading analysis, apoptosis and, 443–444
 - mechanosensing, 441–442
 - unwanted bone loss, apoptosis and, 444–445
- Osteogenic differentiation, ceramic materials, 2919–2923
- Osteogenic progenitor cells
 - aging and, 2092
 - bone-implant interface, 428–429
 - bone tissue engineering
 - animal models, 34
 - regeneration and, 2095–2096
 - future research issues, 2096–2097
 - growth factor differentiation and proliferation, 2094–2095
 - mesenchymal stem cell lines, 2091–2092
 - overview, 2091
 - phenotypic characteristics, 2092–2093
 - plasticity and interconversion potentials, 2093–2094
- Osteolysis
 - total hip arthroplasty, 1380–1381
 - wear debris, animal models, 3070–3072
- Osteo-odonto-kerato-prosthesis (OOKP), ocular implants, 1975
- Osteopontin, osteoinductive substances and materials, 2103
- Osteoporosis, microcomputed tomography research, 1854–1855
- Osteopointegrin (OPG)
 - bone remodeling, 2088
 - differentiation and activation signaling, 2084–2085
 - formation mechanism, 2081
- Outer hair cells (OHCs), hearing mechanisms and, 1273–1275
- Outside-in signaling, integrins, 1532–1533
- Overuse injuries. *See also* Athletic injuries
- hip joint
 - external rotator syndrome, 1397
 - gluteus medius syndrome, 1397
 - groin pain syndrome, 1393–1395
 - hamstring syndrome, 1399–1401
 - iliopsoas tendinitis/bursitis, 1395–1396
 - overview, 1393
 - rectus femoris tendinitis, 1396–1397
 - snapping hip syndrome, 1397–1399
 - knee joint
 - breastroker's knee, 1602
 - epidemiology, 1597–1598
 - Fabella syndrome (fabellitis), 1605
 - fat pad syndrome (Hoffa disease), 1603–1604
 - future research issues, 1605
 - iliotibial band friction syndrome, 1600–1602
 - Osgood-Schlatter disease, 1603
 - patellar tendinitis/tendinosis, 1599–1600
 - patellofemoral pain syndrome, 1597–1599
 - pes anserinus tendinitis and bursitis, 1604–1605
 - plica syndrome, 1604
 - popliteal tendinitis, 1605
 - semimembranosus tendinitis, 1604
 - Sinding-Larsen-Johansson disease, 1603
 - stress fractures
 - diagnostics, 2514–2516
 - differential diagnosis, 2516
 - epidemiology, 2511–2512
 - femur, 2518–2519
 - lumbar spine, 2519–2520
 - metatarsals, 2517–2518
 - pathophysiology, 2513–2514
 - predisposing factors, 2512–2513
 - properties of, 2511
 - tibia, 2517
 - treatment, 2516–2517
- Ownership principles, intellectual property rights, 1546–1547
- Oxidation
 - corrosion electrochemistry, 754–757
 - dental implants, 811–812
 - polyurethane biodegradation, 2306
 - shape memory metal surfaces, 2422
- Oxidative burst, neutrophil-biomaterial adhesion, 1936, 1940
- Oxygen electrode, biosensors, 335–337
- Oxygen permeability
 - hydrogels
 - silicone contact lenses, 715–717
 - soft contact lenses, 1441
 - plasma implantation, 2178
- Oxygen tension, bone marrow tissue engineering, 2664–2665
- Oxygen transport
 - cardiac output measurement, Fick method, 517–519
 - mass transfer and tissue function, 1749–1751
 - microcirculatory system
 - arterioles, 1845–1846
 - artificial oxygen carriers, 1847
 - cardiopulmonary system, 1843–1844
 - carrier systems, 1843
 - cellular metabolism, 1842–1843
 - experimental methods, 1845
 - future research issues, 1847–1848

- nitric oxide, 1847
 overview, 1842
 pathophysiology, 1847–1848
 research issues, 1844–1845
 tissue properties, 1845–1847
 venules, 1846–1847
- Oxy-hydroxyapatite (OHA)
 clinical applications, 1457–1460
 sintering, 1456–1457
- P**
- Pacemakers, cardiac
 basic components, 533–534
 battery systems, 1485–1495
 clinical indications for, 539–540
 future research issues, 540–541
 implantable cardioverter-defibrillators, 776
 pacing leads, 534–535
 pacing modes, 536–539
 power sources, 1484–1485
 pulse generator, 535–536
- Packaging materials, silicone purification, 2437–2438
- Packed-cell volume (PCV), cardiac output measurement, 516–517
- Pain management, burn injuries, dressing materials and, 450
- Pair-addition mechanism, poly(lactic acid) characterization, 2252
- Pancreas. *See also* Artificial pancreas
 islet cells, rotary cell culture systems, 2406
 surgery
 fibrin sealant applications, 1086
 tissue engineering
 artificial pancreas, 2802–2807
 cell sources, 2806–2608
 diabetes, 2802
 endocrine disease, 2807
 microscale technology, 2783–2784
- Pannus formation, valvular prosthetics, 1296, 1300–1301
- Paper industry, alginate applications in, 354
- Papillary muscles, atrioventricular valve structure, 1320–1323
- Paracellular pathways, protein-based drug delivery systems, 2360
- Paracorporeal total artificial lung, 1678–1679
- Parallel beam technology, microcomputed tomography, 1853
- Parallel plate systems, bio-artificial liver systems, 1657
- Parallel processing
 microassembly, 1868–1869
 microelectromechanical systems, microassembly techniques, 1868–1869
- Paramagnetism, magnetic cell separation, 1695–1696
- Parametric analysis, biomedical signal acquisition and processing, 849–851
- Paraplegics, neuroprostheses for, 1927
- Parasitic organisms, aquatic adhesives, 244–245
- Parenteral administration
 intestinal tissue engineering, 2751–2752
 protein-based drug delivery systems, 2359–2360
 small intestine tissue engineering, 2860–2861
- Paris law fatigue crack propagation, pyrolytic carbon design, 2379–2380
- Parkinson's disease, therapy, controlled release delivery systems, 739–740, 744–745
- Parodi Anti-Emboli System (PAES), 876–877
- Partially parallel imaging, magnetic resonance imaging, 1717
- Particle detection, microparticle image velocimetry, 1879–1881
 concentration measurement, 1879–1880
 diffusion measurement, 1880–1881
- Particle distribution, total hip arthroplasty, 1480
- Particle image velocimetry (PIV). *See also* Microparticle image velocimetry
 centrifugal blood pump evaluation, 580–583
 positive displacement pumps, 95, 100–103
- Particle physics, biomedical optics, 2030
- Particle size, supercritical fluids, 2526
- Particulate matter
 composite materials, micromechanics, 668–671
 host reactions, orthopedic devices
 aseptic loosening, 2110–2111
 future research issues, 2114–2115
 macrophage detectors, 2111–2114
 osteoclast/osteoblast effector cells, 2114
 overview, 2110
- Passivation mechanisms, coating surfaces, 2536–2538
- Passive fiberoptic biosensors, basic properties, 1066–1067
- Passive transport mechanisms, transdermal drug delivery, 2894–2895
- Patella
 malalignment, 1597
 tendinitis/tendinosis, 1599–1600
- Patellofemoral joint. *See also* Knee joint
 anatomy and function, 1608–1609
- Patellofemoral pain syndrome, epidemiology, 1597–1599
- Patents, intellectual property management, 1540–1542
- Patents Cooperation Treaty (PCT), 1542
- Pathogens, biofilm mechanisms and, 213–214
- Patterning technique, polyelectrolyte multilayer deposition, 2219–2221
- Pectin, hydrogels, 1446–1447
- Pediatric cardiology
 nitric oxide inhalation therapy, 1954
 ventricular assist devices, positive displacement pumps, 97–102
- Pelvic surgery, computer-assisted techniques, 691–692
- Penetrating keratoplasty, defined, 747–748
- Penetration depth, tissue optics, 1624
- Penicillin, biocatalysis and development of, 162
- Penile prostheses, 2962–2966
- PENSIL intravascular respiratory catheter, 1676–1677
- Pen systems, insulin delivery, 2366
- Pepsin-solubilized collagen, collagen scaffolds, 642
- Peptides
 articular cartilage repair and regeneration, 557
 controlled-release drug delivery, 887
 dendrimer-drug delivery systems, 793
 hydrogels, 1447
 internal architecture, 2357–2359
 nucleic acids, 2193–2194
 protein-based drug delivery systems
 absorption mechanisms, 2360–2361
 hydrogel oral delivery systems, 2363–2364
- Peptidomimetics, 2194
- Perception, eye movements, 1033–1034
- Perceptron convergence, feed-forward networks, 120
- Percolation, disordered composites, 671–672
- Percutaneous Guardwire (Medtronic) device, 871–873
- Percutaneous myocardial revascularization (PMR), results of, 1633
- Percutaneous transluminal coronary angioplasty (PTCA)
 intravascular ultrasound imaging, 1560–1564
 phosphorylcholine technology, 2142–2145
 in vivo animal models, 1571
- Perfluorocarbon emulsions (PFCs), microcirculatory oxygen transport, 1847
- Perfluorooctanoic acid, structure and properties, 2302
- Performance evaluation
 bio-artificial liver systems, 1657–1658
 digital biomedical signal compression, 678–679
 fabrics and fibers, 1055–1056
- Perfusion bioreactors, bone marrow tissue engineering, 2671–2672
- Pericardial tissues, valvular prosthetics, 1295
- Periodontal ligament (PDL), biomechanics analysis and testing, 800–802
 constitutive modeling and numerical analysis, 801–802
 overview, 799
 structural and mechanical properties, 799–800
- Peripheral devices, telemedicine systems, 2583–2584
- Peripheral nerves
 anatomy and function, 2118–2120
 electrical stimulation, 1017
 injury classification, 2120–2121
 nerve guide technology, 1911–1912
 repair and regeneration
 artificial nerve graft biomaterials, 2127–2130
 donor site morbidity, 2126–2127
 epineural vs. fascicular repair, 2125
 graft materials and techniques, 2125–2126
 guidance tubes, 2127
 motor assessment, 2122
 outcomes assessment, 2121–2123
 overview, 2118
 sensory assessment, 2121–2122
 surgical techniques, 2123–2127
- Schwann cells, 2120
 tissue engineering
 acellular grafts, 2815
 axonal degeneration, 2812–2813
 basic principles, 2128
 biomimetic strategies, 2816
 functional recovery assessment, 2813–2814
 gene therapy, 2817
 nerve guidance channels, 2814–2815
 nerve regeneration strategies, 2814–6
 nerve tissue degeneration, 2812
 nervous system anatomy, 2811–2812
 overview, 2811
 stem cells, 2816
 support cells and neurotrophic factors, 2815–2816
- Peripheral vascular disease (PVD), hybrid vascular prostheses, 1432
- Peristaltic pumps, controlled release drug delivery, 740
- Peritoneal cells, tissue engineering, blood vessels, 2645
- Peritoneal-grown autologous graft, hybrid vascular prostheses, 1435
- Periwinkle mucus, adhesive properties, 243
- Perlecan, structural properties, 1199
- Permeability
 articular cartilage biomechanics, 92–93
 blood molecules, 380
 enhancers, transdermal drug delivery, 2896
- Permittivity, tissue engineering, 833–834

- Perovskite structure, piezoelectric materials, 2163–2164
- Peroxide cure, silicone thermosetting elastomers, 2430–2431
- Peroxisome proliferator-activated receptor gamma (PPAR γ), fat tissue engineering, 2730–2731
- Persistent pulmonary hypertension of the newborn (PPHN), nitric oxide inhalation therapy, 1953–1954
- Personal stereos, tinnitus, acoustic masking, 2613–2614
- Pes anserinus tendinitis/bursitis, 1604–1605
- Phagocytosis
 - foreign body response, 1112–1113
 - immune response, 263–264
 - neutrophil-biomaterial adhesion, 1940
- Pharmaceutical bioadhesives, 152–153
 - administration routes, 157–158
 - bioadhesive bond, 153–155
 - formulation factors, 156–157
 - multifunctional, 155
 - polymers, 153–154
 - second-generation bioadhesives, 155–156
- Pharmacia HSM Intraocular Lens, heparin biocompatibility, 1371
- Pharmacodynamics
 - functional magnetic resonance imaging, 1126
 - toxicity testing, animal surrogate systems, 52–53
- Pharmacokinetics
 - mass transfer and tissue function, 1749–1750
 - approximations, 1755–1756
 - toxicity testing, animal surrogate systems, 51–53
- Pharmacology
 - functional magnetic resonance imaging, 1126
 - polymer structure and classification, 220–222
- Phase continuity, disordered composites, 671–672
- Phase-contrast magnetic resonance angiography (PC-MRA), vascular imaging, 3016
- Phased array coils, magnetic resonance imaging, 1717
- Phase encoding, magnetic resonance imaging, 1716
- Phase explosion, laser-tissue interactions, 1629–1630
- Phase function, biomedical optics, 2031
- Phase morphology, silicone hydrogel contact lenses, 721–722
- Phase separation
 - biocomposite fabrication, 188–189
 - interpenetrating polymer networks, 1549–1550
 - microporous materials, sol-gel bioactive glasses, 1888
 - polymer foam processing, 2273
 - polyurethane biomaterials, 2310
 - tissue engineering scaffolds, 2832–2833
- Phenomenologic constitutive modeling, heart biomechanics, 1286–1287
- Phenotyping
 - genetics, magnetic resonance imaging, 1726
 - magnetic resonance microscopy, 1741–1742
 - myocyte regulation, 545
 - osteogenic progenitor cells, 2092–2093
- Phonocardiogram, biomedical signal acquisition and processing, 853
- Phosphatases, valvular prosthetics, calcification, 78
- Phospholipids
 - gradient surfaces, 1215–1216
 - hydrogels, 1447–1448
- Phosphophoryn, osteoinductive substances and materials, 2103–2104
- Phosphorylcholine (PC)
 - applications, 2142–2146
 - basic principles, 2136
 - blood-material interactions, 2142–2145
 - coating stability, 2139
 - drug delivery capability, 2139–2140
 - future research issues, 2146–2147
 - ocular implants, 2145–2146
 - polymer evolution and synthesis, 2136–2138
 - surface mobility and swelling properties, 2139
 - in vitro bioevaluation, 2140–2141
- Photoablation, light-tissue interactions, 2033
- Photobleaching, confocal microscopy, 710–711
- Photochemistry
 - inherently conductive polymers, 697
 - laser-tissue interactions, 1626, 2032
- Photocrosslinking
 - hyaluronan biomaterials, 1426
 - hydrogels, 2154–2158
- Photocurable systems, glue materials, 1191–1192
- Photodiode arrays (PDAs), optical mapping, 2020–2021
 - design criteria, 2022–2024
- Photodisruption, light-tissue interactions, 2033
- Photodynamic imaging, endoscopic techniques, 995
- Photodynamic therapy (PDT)
 - esophageal tissue engineering, 2719–2720
 - laser-tissue interactions, 1626, 2032
- Photofix heart valve prosthetics, 1313
- Photografting, photopolymerized biomaterials, 2158
- Photoinitiators
 - biocompatibility, 2152–2153
 - photopolymerization mechanisms, 2150–2152
- Photolithography, tissue engineering, microvascular networks, 2794–2795
- Photolytic artificial lung, 1679–1680
- Photomasking, tissue engineering, microvascular networks, 2795
- Photomechanical interactions, laser-tissue interactions, 1628–1629
- Photomultiplier tubes (PMTs), flow cytometry, 1103–1104
- Photon characteristics
 - biomedical optics, 2030
 - fibers, optical coherence tomography, 1999
- Photopolymerization
 - biocompatible photoinitiators, 2152–2153
 - biomaterials, 2154–2158
 - hydrogels, 2154–2158
 - mechanisms, 2150–2152
 - overview, 2150
 - photografting techniques, 2158
 - tissue engineering applications, 2158–2160
- Photoreactivity, glue materials, 1192
- Photorefractive keratectomy (PRK), wound healing, 2712–2714
- Photostimulation systems, controlled-release drug delivery, 886
- Photothermal interactions, 1626–1627
- pH sensitivity
 - bone marrow tissue engineering, 2664
 - cell culture assays, 565
 - electrodes, biosensor technology, 335
 - glucose hydrogels, 1166–1170
 - protein-based drug delivery systems, 2360–2361
 - hydrogels, 2363–2364
 - protein-surfactant interactions, 2341–2343
- Physical gels (pseudo gels), basic properties, 1444–1446
- Physical properties
 - ultrasound therapy, bone healing, 2953–2955
 - vitrification, 3046–3049
- Physical therapy, robotics-assisted rehabilitation, 1805–1806
- Physiological insulin replacement, 1518–1519
- Physiome project, mathematical modeling, 1900–1901
- Piezoelectric materials
 - applications, 2171
 - basic principles, 2163
 - ceramics, 2166–2167
 - classification, 2164–2165
 - constitutive equations, 2165–2166
 - electroactive polymeric materials, 929–932
 - artificial muscles, 108–109
 - future research issues, 2171–2172
 - history, 2163–2164
 - polymers, 2167–2170
 - amorphous polymers, 2169–2170
 - biomaterials, 2170
 - copolymers, 2168–2169
 - nylons, 2169
 - synthetic polymers, 2168
 - ultrasound therapy, bone healing, 2953–2956
- Pigtail catheters, cardiac applications, 503–504
- Pinhole size, confocal microscopy, 709
- Pitting and holes, topographical surface analysis, 2553
- Pitting resistance equivalent (PRE), nickel alloys, 1944–1945
- Planck's constant
 - biomedical optics, 2030
 - nuclear magnetic resonance, 1961
- Plane stress, composite materials, 663–665
- Plant breeders rights (PBRs), 1544–1545
- Plant materials
 - flow cytometry, 1108
 - tissue engineering applications, 2292
- Plaque imaging techniques, vascular imaging, 3019
- Plasma filtration and adsorption, plasmapheresis, 386–387
- Plasma-induced ablation, laser-tissue interactions, 1630, 2033
- Plasma materials
 - bio-artificial liver systems, 1656
 - lung surfactant dysfunction, 1687
 - silicone hydrogel contact lenses, 724–725
 - surface modification, 2543–2545
- Plasmapheresis
 - membrane plasma separation, 384–387
 - overview, 384
- Plasma source ion implantation, diamond/diamond-like carbons, 825–826
- Plasma-spray coating
 - ceramics, 590
 - dental implants, 814
 - hydroxyapatite, 1464–1465
- Plasma techniques
 - biomaterials modification
 - deposition and coating, 2181–2182
 - implantation, 2176–2177
 - metallic ion implantation, 2177–2180
 - overview, 2174
 - polymer ion implantation, 2180–2181
 - spraying, 2183–2184
 - sputtering and etching, 2176
 - corona and atmospheric discharge, 2175
 - discharge techniques, 2174–2175
 - gradient surface preparation, 1212–1213
 - ion implantation, 2175
 - radiofrequency discharge, 2174–2175
 - sterilization of medical devices, 2504–2505, 2507
 - surface interactions, 2175–2176

- thin film deposition, 2175–2176
tissue engineering, microvascular networks, 2795–2796
vacuum arc, 2175
- Plasmid DNA
gene-activated matrix scaffolds, hard tissue regeneration, 1146
gene therapy and, 1151
supercritical fluid processing, 2524–2525
- Plasminogen activators, pharmacological activity, 220–221
- Plasticity, osteogenic progenitor cells, 2093–2094
- Plastics, fiberoptic biosensors, 1063–1064
- Plate fixation
bioabsorbable plates and screws, 432–437
bone fractures, 415–416
failure analysis
anterior cervical spine plate, 1832–1834
dynamic compression plate, failure analysis, 1829–1832
mandibular reconstruction plate, 1837–1839
orthopedic devices, 2075–2076
reconstructive biomaterials, 2398
- Platelet-rich-plasma (PRP), histiogenesis, 1407–1408
- Platelets
adhesion
coating surfaces and, 2533
heparin biocompatibility, 1369
mechanisms of, 6
vascular grafts, host reactions, 3007
blood-material interactions, 370–372
hemocompatibility assessment, 1353
host reactions, 1415–1416
mechanical heart valves, 1334–1335
thrombosis
coagulation and neutrophil enhancement, 2606–2607
function and adhesion, 2601–2602
neutrophil heterotypic interactions, 2603
- Plexin semaphorin integrin (PSI), integrin structure, 1525
- Plica syndrome, 1604
- P(MMA-g-EG) networks, glucose hydrogels, 1168–1170
- Pneumatic total artificial heart, 475
- Pneumatic ventricular assist device, 472–473
- Point RESolved Spectroscopy (PRESS), brain imaging, 1967–1969
- Poiseuille's law
macrocirculatory hemodynamics, steady flow properties, 1359–1361
vascular fluid dynamics, 2975
- Poisson equation, electrical stimulation, 1010
- Poisson's ratio
articular cartilage biomechanics, 553–554
composite micromechanics, 667–671
- Polarity effects, electrical stimulation, 1014
- Polarization optical coherence tomography, 2002–2003
- Polarization transfer
biomedical optics, 2029–2030
magnetic resonance imaging, 1720
- Polarons, electroactivity, 927
- Polishing, shape memory metal surfaces, 2421–2422
- Poly(2-hydroxyethyl methacrylate) (PHEMA), keratophakia implants, 748–749
- Poly(2-methacryloyloxyethyl phosphorylcholine) (Poly(MPC)), soft contact lenses, 1442
- Poly(3-hydroxy-butyrate-co-3-hydroxyvalerate) (PHBV), tissue engineering applications, 2292–2293
- Poly(3-hydroxybutyrate) (P3HB), biodegradable compounds, 199–200
- Poly(3-hydroxyvalerate) (PHV), biodegradable compounds, 199–200
- Poly(acrylic acid)s/carbomers, first-generation bioadhesives, 152
- Polyacrylonitrile (PAN), artificial muscles, electroactive fibers, 110–112
- Poly(α -hydroxyesters), tissue engineering applications, 2285–2287
- Polyalkene cements, orthopedic devices, 2069
- Polyamides
basic properties, 2186
classification, 2186
hemodialyzers, 1579–1580
natural materials, 2186–2188
cyclotides, 2188
proteins, 2816–2817
silk, 2187–2188
wool, 2188
synthetic materials, 2188–2194
aromatic polyamides/aramids, 2192–2193
Kevlar[®], 2189–2190
nylon, 2188–2189
peptide nucleic acids, 2193–2194
peptidomimetics, 2193
pyrrole-imidazole, 2191–2192
sodium polyaspartate, 2191
Versamid[®], 2190–2191
- Polyamidoamine (PAMAM) dendrimer, drug delivery systems, 789–794
- Poly(amino acids)
polyamides, 2186–2187
tissue engineering applications, 2288
- Polyanhydrides
orthopedic devices, 2065
tissue engineering applications, 2287–2288
- Poly(β -malic acid) (PMA), biodegradable compounds, 200
- Polycaprolactone (PCL)
biocomposites, 183
biodegradable polymers, 196–199
bone/cartilage interface, 2203
bone grafts, 2203
cartilage repair, 2202–2203
composite materials, 2200–2201
degradation, 2198–2200
drug delivery systems, 2201–2202
elastomers, 897–900
gene therapy, 2204
heart tissue, 2203–2204
liver repair, 2203
nerve grafts, 2202
orthopedic devices, 2064–2065
skin grafts, 2202
synthesis and structure, 2197–2198
tissue engineering, 2202
tissue engineering applications, 2287
- Polycarbonates, structure/property relationships, 2278–2279
- Poly(carbonate)urethanes
basic properties, 2208–2209
biocompatibility testing, 2211–2212
biodegradability testing, 2209–2211
melt processing, 2212–2213
regulatory status, 2213–2214
- Polycarboxylate adhesives, basic properties, 2
- Polycations, multilayer assemblies, polyelectrolytes, 2219–2221
- Polycondensation, poly(lactide-co-glycolide), 2259–2261
- Poly(dimethylsiloxane) (PDMS)
bio-artificial liver systems, 1654–1655
hydrogel contact lenses, 716–727
silicone structure, 2427
tissue engineering, microvascular networks, 2796
- Polydioxanones
biodegradability, 197–198
orthopedic devices, 2065
- Polyelectrolytes, multilayer assemblies
biomolecules, cell adhesion/encapsulation, 2223–2226
characterization, 2221–2222
deposition techniques, 2219–2221
formation mechanisms, 2218–2219
light emitting diodes, 2222–2223
materials, 2219
microfluidics flow control, 2226
overview, 2217
separations, 2223
structural properties, 2217–2218
- Polyepoxy compounds
antiadhesive membranes, 2237–2238
blood vessel prostheses, 2236
collagen tissue fixation, 645
cross-linking mechanisms
artificial organs, 2231
biomaterials, 2231–2232
collagen molecules, 2232
glutaraldehyde vs., 2232–2235
growable vascular prosthesis, 2236–2237
heart valve applications, 2235–2236
hemostat applications, 2237
- Poly(ester amide)s (PEAs), biodegradable compounds, 200–201
- Polyesters
biodegradable elastomers, 897–901
biodegradation mechanisms, 900–901
cross-linked synthesis, 898
linear synthesis, 898–899
mechanical properties, 900
endovascular stent grafts, 2483
natural origins, tissue engineering applications, 2292–2293
structure/property relationships, 2278–2279
- Poly(ethylene glycol) monomethacrylate (PEGMA), self-regulated insulin delivery, 1168–1170
- Polyethylene glycol (PEG)
adsorption resistance of, 278
bioadhesive properties, 4–5
cell-adhesive properties, 219–220
chitosan, tissue engineering, 611–613
dendrimer-base drug delivery, unimolecular dendritic micelles, 790–793
glue materials, 1191
diacrylate, 1192
hemocompatibility, 1346
NHS-PEG compounds, 1191
protein-based drug delivery systems, hydrogels, 2363–2364
self-regulated insulin delivery, glucose hydrogels, 1165–1170
tissue engineering applications, 2289
- Polyethyleneimine (PEI)
gradient surface preparation, 1215
heparin biocompatibility, Carmedia Bioactive Surface (CBAS), 1371
nitric oxides, 222
- Poly(ethylene oxide) (PEO)
biodegradability, 199
gradient surfaces, 1215
tissue engineering applications, 2289
- Polyethylene (PE). *See also* Ultra-high molecular weight polyethylene (UHMWPE)
cardiac catheters, 496
structure/property relationships, 2276–2278
wear debris, prosthetic devices, 3058–3063
- Poly(ethylene terephthalate) (PET)
structure/property relationships, 2278–2279
vascular grafts, 2987–2988
- Polyfumarates, tissue engineering applications, 2288

- Poly(glycerol sebacate) (PGS)
 applications, 333
 biocompatibility, 331–333
 characterization, 328–329
 degradation, 329–331
 design criteria, 327–328
 overview, 327
 synthesis, 328
- Poly(glycolic acid)/poly(glycolide) (PGA)
 applications, 2244–2245
 articular cartilage repair, 556
 basic properties, 2241
 biocomposites, 183
 biodegradable polymers, 196–199, 779–780
 elastomers, 897–900
 nonporous polymer-ceramic composites, 784–785
 characterization, 2242
 degradation, 2242–2244
 orthopedic devices, 2064
 regulatory issues, 2245–2246
 structure/property relationships, 2278–2279
 synthesis, 2241–2242
 tissue engineering
 clinical applications, 2285–2287
 small intestine scaffold materials, 2865–2866
 vascular grafts, development strategies, 2999
 in vivo tissue response, 2246–2247
- Polyhydroxyalkanoates (PHAs)
 elastomers, biodegradability, 899
 tissue engineering applications, 2292–2293
- Poly(hydroxybutyrate) (PHB), tissue engineering applications, 2292–2293
- Poly(lactic acid)/poly(lactide) (PLA)
 articular cartilage repair, 556
 biocomposites, 183
 biodegradable compounds, 779–780
 elastomers, 897–900
 nonporous polymer-ceramics, 784–785
 biomedical applications, 2254–2255
 characterization, 2252
 degradation, 2252–2253
 in vivo characteristics, 2253–2254
 foreign body response, 1115–1116
 formation mechanisms, 2250–2251
 orthopedic devices, 2064
 overview, 2249
 synthesis, 2249–2250
 tissue engineering, applications, 2285–2287
- Polylactideglycolic acid (PLAGA), fabric scaffolds, 1056–1059
- Poly(lactic/lactide-co-glycolide/glycolic acid) (PLGA)
 basic properties, 2259
 biocompatibility and tissue reactions, 2265–2267
 biocomposites, 183
 fabrication techniques, 186–189
 biomedical applications, 2267
 crystallinity, 2262
 degradation, 2263–2265
 future research issues, 2267
 mechanical properties, 2262–2263
 performance characteristics, 2261–2267
 reconstructive biomaterials, 2396–2397
 solubility, 2261–2262
 structure/property relationships, 2278–2279
 synthesis, 2259–2261
 thermal stability, 2262
 tissue engineering, small intestine scaffold materials, 2866
- Poly(L/DL) lactide (PLDLA), bioabsorbable plates and screws, 433–437
- Poly(L-glutamic acid) (PLGA), gelatin-based glues, 1190–1191
- Poly(L-lactic acid/poly-L-lactide) (PLLA)
 bioabsorbable plates and screws, 433–437
 biodegradability, 198–199
 free radical reactions, 203–204
 tissue engineering, small intestine scaffold materials, 2866
- Poly(L-lactide-co-D,L-lactide) (PDLLA)
 biodegradability, free radical reactions, 203–204
 facial skeletal fixation, 2996–2997
- Polymer foams
 bioactive scaffold fabrication, 149–150
 burn dressings, 458–459
 microporous materials, 1887
- Polymerization, inherently conductive polymers, 697–699
- Polymers. *See also* Elastomers;
 Photopolymerization; specific polymer compounds
 antimicrobial/antibiotic materials, 61–62
 articular cartilage repair
 natural materials, 555–556
 synthetic polymers, 557
 artificial liver systems, 1654–1655
 artificial lung systems
 extracorporeal membrane oxygenation, 1669–1670
 membrane permeance, 1663–1665
 artificial pancreas, 2803–2805
 bioadhesive properties, 4–5, 154
 formulation factors, 156–157
 biodegradable polymers
 basic properties, 195–196
 degradation properties, 201–204
 free radicals, 203–204
 glycolide, lactide, and caprolactone-based compounds, 196–199
 linear aliphatic polymers, 196
 natural materials, 199–200
 poly(ester amide)s, 200–201
 bifunctional applications, 218
 ECM-mimetic materials, 218–221
 pharmacologically active materials, 220–222
 biomimetic materials, 295–296
 biorubbers, 327–333
 blood-material interactions, 374–376
 bone cement, 404–406
 bone matrices, 2047
 osteoclasts and regeneration, 2096
 cardiac catheters, 494–498
 cardiac pacing leads, 535
 cell-material interactions, 568–572
 coatings, 2531–2532
 collagen
 polyepoxy compounds, 645
 scaffold fixation strategies, 643–644
 composites
 degradable compounds, 779–785
 clinical applications and limitations, 782–783
 nonporous polymer-ceramics, 784–785
 scaffold materials, 784
 resorbable vs. nonresorbable components, 182–184
 controlled-release drug delivery systems, 741–742
 fabrication and encapsulation technologies, 881–883
 films and rods, 882
 formulations, 881
 gel materials, 882–883
 particulates, 882–883
 release profiles and models, 743
 corneal implants
 epikeratoplasty, 749
 intraocular ring implants, 749–750
 keratophakia, 748–749
 refractive keratoplasty, 748–750
 cross-linking and orientation, 2279
 dendrimer drug delivery systems
 chemical bonding, 793–794
 dendrimer/DNA complexes, 792–794
 dendritic box structures, 790
 encapsulation and bonding, 790–793
 hydrogens, 792
 overview, 789
 properties, 794–795
 synthesis, 789–794
 unimolecular dendritic micelles, 790–792
 elastin structural modeling, 894–895
 electroactive materials
 conducting polymers, 926–929
 drug delivery systems, 932–933
 electro-responsive gels, 932
 overview, 926
 piezoelectric polymers, 929–932
 electrospinning
 coated nanofibers, 983
 future research issues, 983–984
 medical applications, 982–983
 nanofiber fabrication, 978–982
 overview, 978
 foams
 cellular solids, 2270
 fabrication, 2271–2274
 fugitive phase removal, 2271–2273
 gas foaming, 2271
 overview, 2270
 phase separation, 2273
 pore formation, 2272–2273
 sintering, 2273–2274
 structure, morphology, and mechanical properties, 2271
 textile processing, 2274
 future applications and research, 2280–2281
 gene-activated matrix scaffolds, 1145–1146
 glass transition temperature, 2276
 glucose-responsive hydrogels
 drug delivery systems, 1163–1164
 environmentally-sensitive compounds, 1164–1165
 future research issues, 1170–1171
 insulin delivery systems, 1165–1170
 overview, 1163
 gradient surfaces
 cellular interactions, 1222–1225
 hydroxy groups, 1213–1214
 ionic functional groups, 1214–1215
 phospholipid polymers, 1215–1216
 protein-surface interactions, 1220–1222
 radio-frequency plasma discharge, 1212–1213
 water contact angle, 1216–1217
 hard tissue interactions, 1232–1233
 hemocompatibility, inert materials, 1345–1346
 heparin biocompatibility, 1371–1373
 historical evolution of, 2275
 hyaluronan, structure and properties, 1421–1422
 hybrid materials systems, 2281
 hydrogels
 biomimetic materials, 1447–1448
 soft contact lenses, 1439–1442
 inherently conductive polymers
 biological functionality, 697–699
 cellular communication, 699–701
 chemical polymerization, 697
 electrochemical properties, 699
 electrochemical synthesis, 696–697
 enzyme-catalyzed polymerization, 697
 future applications, 701–702
 overview, 695
 photochemical polymerization, 697
 structure and properties, 695–696
 synthesis, 696–699

- insulin delivery systems, 1519
- interpenetrating networks
 environmentally sensitive hydrogels, 1550–1551
 hydrogels, 1550–1553
 overview, 1549
 phase separation and mechanical behavior, 1549–1550
 temperature-sensitive hydrogels, 1551–1554
- interstitial fluid sensors, 1175–1179
- low molar mass materials, 2281
- mechanical heart valves, durability, 1329–1330
- medical device infection, 1503
- melt spinning, 1816–1818
 biomedical applications, 1821–1822
 braided method, 1822
 extrusion, 1817
 fiber drawing, 1817–1818
 fiber quenching, 1817
 knitting method, 1821–1822
 properties, 1818
 woven and non-woven, 1821
- microporous materials, foam replication, 1887
- nanotechnology, 2281
- nerve guide technology, 1914
 biohybrids, 1915–1917
 nerve tissue engineering, 2815
 physical properties, 1918–1919
 resorbable nerve guides, 1915–1916
- nitric oxide materials, 1951–1956
- ocular implants, 1974–1975
 intraocular lenses, 1977–1978
 matrix ocular implants, 1983–1987
 reservoir ocular implants, 1987–1992
 scleral buckles, 1976
- orthopedic devices, 2063–2066
 biodegradable polymers, 2063–2064
 bone cements, 2067–2068
 clear UHMWPE, 2065
 composites, 2068–2069
 fracture fixation, 2073
 hydrogels, 2065
 stable implants, 2066
 synthetic biodegradable polymers, 2064–2065
- phosphorylcholine technology
 applications, 2142–2146
 basic principles, 2136
 blood-material interactions, 2142–2145
 coating stability, 2139
 drug delivery capability, 2139–2140
 evolution and synthesis, 2136–2138
 future research issues, 2146–2147
 ocular implants, 2145–2146
 surface mobility and swelling properties, 2139
 in vitro bioevaluation, 2140–2141
- piezoelectric materials, 2167–2170
 amorphous polymers, 2169–2170
 biomaterials, 2170
 copolymers, 2168–2169
 nylons, 2169
 synthetic polymers, 2168
- plasma biomaterials modification
 grafting materials, 2182–2183
 ion implantation, 2180–2181
- plasma surface modification, 2543–2545
- polyepoxy compounds
 antiadhesive membranes, 2237–2238
 artificial organ cross-linking, 2231
 biomaterial cross-linking, 2231–2232
 blood vessel prostheses, 2236
 collagen molecule cross-linking, 2232
 glutaraldehyde *vs.*, 2232–2235
 growable vascular prosthesis, 2236–2237
- heart valve applications, 2235–2236
 hemostat applications, 2237
- polyesters, 2278–2279
- polyethylene/mono-substituted ethylenes, 2276–2278
- protein-based drug delivery systems, 2362–2363
 hydrogels, 2363–2364
- scaffold materials, bioactive fabrication, 148–150
- scanning electron microscopy analysis, 972–973
- seamless synthesis and processing, 2280
 sequence control, 2280–2281
- shape memory metal coatings, 2423–2424
- shape memory polymers, artificial muscle materials, 108
- silicone hydrogel contact lenses
 chemical coatings, 725
 core surface treatments, 725
 current commercial sources, 725–727
 evolution of, 718–719
 hydrophilic components, 720
 hydrophobic comonomers, 721
 ion permeability and on-eye movement prediction, 722
 ocular environment, 715
 overview, 715
 oxygen permeability, 715–717
 on phases, 722–723
 plasmas, 724–725
 polymer phases, 721–722
 recent developments, 719
 research background, 717–718
 surface properties, 723–725
 transport properties, 721
- small organic molecules, 2275–2276
 structure/property relationships, 2275–2280
- sutures, 2563–2568
- synthesis strategies and research, 2279–2280
- tissue engineering
 biomaterials properties, 2284–2285
 bladder, cell delivery vehicles, 2637
 blood vessels, 2644–2647
 bone, 2655–2656
 cardiac muscle, 2686, 2688
 cell carrier properties, 2284
 degradable, nondegradable, and stable materials, 2284–2285
 engineering technology, 2282–3
 extracellular matrix and scaffold properties, 2283–2284
 fat tissue, 2729–2730
 future trends and research issues, 2295
 intelligent polymers, 2294–2295
 liver, 2772–2775
 microvascular networks, 2796–2798
 natural polymers, 2289–2293
 natural/synthetic compounds, 2293–2294
 overview, 2282
 skeletal muscle, 2843–2844
 sources, 2285–2295
 synthetic polymers, 2285–2289
- total hip arthroplasty, 1473–1474
- transdermal drug delivery, backing materials, 2896
- transmission electron microscopy analysis, 973–974
- vascular grafts, 2987–2989
 development strategies, 2999–3001
 hybrid prostheses, 1432–1437
 in vivo animal models, 1569–1570
- Polymer sponge technique, ceramics and glass fabrication, 147
- Poly(methacrylic acid) (PMA)
 bioadhesive properties, 4–5
 glucose hydrogels, 1163–1172
- Poly(methyl methacrylate) (PMMA)
 antiseptic/antibiotic additives, 65
 bone cement, 403–409
 bone-implant interface, 426–427
 chemical components, 403–404
 grafting techniques, 159
 polymerization reaction, 404–405
 porosity and volumetric shrinkage, 407–408
- corneal implants
 intracorneal ring implants, 749–750
 keratoprosthesis, 750–751
- interpenetrating polymer networks, 1550–1551
- intracorneal implants, 1974–1975
- intraocular lenses, 1977–1978
- orthopedic devices, ceramic composites, 2069
 plasma ion implantation, 2180–2181
 protein-based drug delivery systems, 2363–2364
 reconstructive biomaterials, 2400
 total hip replacement, 19–21
 total knee replacement, 1612–1613
- Polymorphonuclear leukocytes (PMNs), acute inflammation, host reactions, 1411
- Poly(*N*-isopropylacrylamide) (NIPAAAM)
 interpenetrating polymer networks, 1551–1552
 tissue engineering applications, 2294
- Poly(ortho esters)
 orthopedic devices, 2065
 tissue engineering applications, 2288–2289
- Polypropylene
 plasma ion implantation, 2180
 structure/property relationships, 2277–2278
- Polypropylene fumarate (PPF),
 biocomposites, 183
- Polypyrrole, electroactive polymeric materials, 927–929
- Polysaccharides
 aldehyde surgical glues, 1189–1190
 mimetic materials, 297–298
 pharmacological activity, 220
 tissue engineering applications, 2290–2292
- Polystyrene
 plasma ion implantation, 2180
 structure/property relationships, 2278
- Poly(tetrafluoroethylene) (PTFE)
 chemical properties, 2300–2301
 commercial fluorinated polymers, 2302
 endovascular stent grafts, 2483–2484
 future research issues, 2302
 monomer and polymer production, 2301
 origins, 2300
 plasma ion implantation, 2181
 toxicity, 2301–2302
 vascular grafts, 2988–2992
 development strategies, 2999–3001
- Poly(thiophene), electroactive polymeric materials, 927–929
- Polyurethanes
 aromatic *vs.* aliphatic compounds, 2212
 biodegradable elastomers, 901–905
 biocompatibility, 904–905
 biodegradation mechanisms, 904
 mechanical properties, 903–904
 processing and applications, 905
 synthesis, 901–904
- biomaterials
 antibacterial modifications, 2310
 biocompatibility and toxicity, 2306
 biodegradation mechanisms, 2306–2309
 biostable compounds, 2307–2308
 blood-material interfaces, 2310–2311
 fluid exposures, 2312
 medical applications, 2305–2306
 overview, 2304
 processing techniques, 2309–2310

- Polyurethanes (*cont.*)
 segmented structures, 2304–2305
 tissue-material interactions, 2311–2312
 cardiac catheters, 496
 cardiac pacing leads, 535
 polycarbonates
 basic properties, 2208–2209
 biocompatibility testing, 2211–2212
 biodurability testing, 2209–2211
 melt processing, 2212–2213
 regulatory status, 2213–2214
 reconstructive biomaterials, polyurethane terephthalate, 2401
 silicone copolymers, 2433–2434
 urinary prostheses, 2962
 vascular grafts, 2988–2989
- Poly(vinyl alcohol) (PVA)
 bioadhesive properties, 4–5
 corneal implants, epikeratoplasty, 749
 hydrogels, 1445
 vascularization, 3026–3028
- Polyvinylchloride, cardiac catheters, 496–498
- Poly(vinylidene fluoride) (PVDF)
 electroactive polymeric materials, 930–932
 piezoelectric materials, 2168–2169
- Poly(vinyl pyrrolidone) (PPV),
 hemocompatibility, 1345–1346
- Popliteal tendinitis, 1605
- Porcelain ceramics
 dental implants
 feldspathic porcelain, 595–597
 high-alumina porcelain, 597–598
 metal-bonded porcelain, 597
 implants, 588–589
 processing, 585–586
- Porcine endogenous retroviruses (PERVs)
 artificial pancreas, 2805–2806
 xenografts, 3100
- Porcine valves
 human valve structure and, 1322–1323
 valvular prosthetics, failure analysis, 1296–1301
- Porosity
 bone cement, 407–408
 dialysis membranes, 1585–11
 endovascular stent grafts, 2484
 fabric scaffolds, 1057–1058
 hydroxyapatite scaffolds, 1456–1460
 interfacial transport, protein permeability, artificial membranes, 2353–2354
 nerve guide technology, 1919
 polymer foams
 basic characteristics, 2270
 processing, 2272–2273
 synthetic apatite composites, 232–233
 vascular grafts, 2989
 host reactions, 3008
- Porous devices
 corneal implant skirts, keratoprosthesis, 751
 melt-derived bioactive glass, microporous materials, 1890
 three-dimensional matrices, bio-artificial liver systems, 1655
 tissue engineering, microvascular networks, 2797–2798
- Portable devices, telemedicine systems, 2584
- Positive displacement pumps
 artificial heart fluid dynamics, 95–105
 controlled release drug delivery, 741
- Positron emission tomography (PET), basic principles, 1796
- Postadhesion events, integrin regulation, 1528–1530
- Posterior capsule opacification (PCO), intraocular lenses, 1977–1978
- Posterior cruciate ligament (PCL)
 anatomy and function, 1608–1609
 animal models of repair, anatomy, structure, and function, 1640–1641
 prosthetic materials design, 1609–1610
- Post-mastectomy radiation therapy (PMRT), breast tissue engineering, 2684
- Postnatal tissue, histiogenesis, 1406–1407
- Postprandial EGG patterns, properties, 954–955
- Post-traumatic stress syndrome, eye movements and, 1036
- Postural control, mechanics of, 730
- Powder-based fabrication technique, bioactive ceramics, 147
- Power Doppler imaging (PDI), basic principles, 2946–2947
- Power law
 ceramics fracture/fatigue analysis, 587
 vascular fluid dynamics, 2978
- Power sources and systems
 cardiac pacemakers, 535–536
 centrifugal blood pumps, 582–583
 neuroprostheses, 1924–1926
 transdermal drug delivery, 2900
 ultrasound therapy, bone healing, 2954–2955
 wheeled mobility devices, 3085–3087
- Power spectral density (PSD)
 biomedical signal acquisition and processing, 849–850
 digital signal acquisition and processing, higher-order spectra, 861–863
- Preadipocytes
 biomaterials and research on, 2476
 tissue engineering
 breast, 2680
 fat tissue, 2728
- Preamplification, electromyography signal conditioning, 962
- Precession rate, magnetic resonance microscopy, 1738–1739
- Precipitation techniques, hydroxyapatite synthesis, 1455
- Precision casting, titanium and titanium alloys, dental implants, 2887–2889
- Precursor cell identity, bone healing, callus formation, 1259
- Preform inlays, dental ceramics, 602
- Prehension, mechanics of, 731–732
- Prenatal bladder disease, tissue engineering and, 2640
- Pressure sensitive adhesives
 basic properties, 2–5
 transdermal drug delivery, 2896
- Pressure-volume curves, heart biomechanics, 1280–1282
- Presterilization procedures, biomedical devices, 2502–2503
- Prima facie duties, biomedical ethics and, 290
- Primary Rate Interface (PRI), real-time telemedicine systems, 2580–2581
- Primary/secondary bone formation, bone healing, 1258–1259
- Principal component analysis (PCA), optical detection of cancer, 2014–2015
- Principlist approach to biomedical ethics, 285–293
- Probability density function (PDF), biomedical signal acquisition and processing, 845
- Probiotics, medical device infections, 1504
- Procollagen
 gene expression, fibroblast mechano-regulation, 1784–1787
 structure and properties, 630–632
- Prodrug polymers, controlled-release drug delivery systems, 885–886
- Profile method, tendon/ligament tensile testing, 2591–2592
- Progenitor cells, fat tissue engineering, adipocyte-derived cell lines, 2728
- Proinflammatory cytokines
 proteoglycan degradation and aging, 1200–1201
 wear debris, 3067–3068
- Proline, bio-artificial liver systems, 1655–1656
- 1,3 Propanediol, whole-cell biocatalysis, 165
- Proportionality constant, electrophoretic cell separation, 1701
- Proprioception, ligaments, 2762
- Prostaglandins
 osteogenic progenitor cell differentiation and proliferation, 2095
 wear debris, inflammatory response, 3070
- Prostate cancer, chemotherapy, controlled-release drug delivery systems, 887–888
- Prosthetic devices
 ankle prostheses, alumina ceramics, 21
 artificial lungs
 blood oxygenator, 1666–1667
 cardiopulmonary bypass oxygenators, 1667–1669
 diffusional boundary layers, 1665
 extracorporeal membrane oxygenation, 1669–1670
 gas exchange determinants, 1663
 historical background, 1661–1662
 mass transfer correlations, 1665–1666
 membrane permeance, 1663–1665
 operating principles, 1662–1666
 overview, 1661
 centrifugal blood pumps, 574–583
 failure analysis, orthopedic metal devices, 1829–1841
 hip joint biomechanics
 arthritis, normal and artificial hips, 1377
 cup stability, micromotion, fixation, and osteolysis, 1380
 femoral head subluxation, 1381–1382
 finite element modeling and validation, 1381–1385
 future research issues, 1388–1390
 gait loading results, 1386
 implant fixation techniques, 1379–1380
 liner micromotion, 1385–1386
 pelvic osteolysis and adjunctive screws, 1380–1388
 post THA dislocation, 1381
 subluxation loading results, 1386–1388
 total hip replacement techniques, 1377–1379
 immune response, major histocompatibility complex, 266–267
 mechanical heart valves, positive displacement pumps, 96–98
 microelectromechanical systems, 1860–1862
 neuroprostheses
 basic properties, 1924
 bladder management, 1929–1930
 cochlear implants, 1930–1931
 historical background, 1926
 physiology, 1924
 reaching and grasping applications, 1927–1929
 technology, 1924–1926
 walking applications, 1926–1927
 nitric oxide materials, 1955–1956
 ocular implants
 drug delivery systems, 1979
 glaucoma filtration implants, 1976–1977
 intraocular lenses, 1977–1978
 keratophoresis, 750–751
 keratoprosthesis and intracorneal implants, 1974–1975
 keratoprosthesis, 750–751
 orbital implants, 1974

- scleral buckles, 1976
 silicone oil, 1979
 viscoelastic agents, 1979
 orthopedic devices, failure analysis, 1829–1841
 polyepoxy compounds, 2231–2237
 pyrolytic carbon
 orthopedic devices, 2374–2375
 valvular prostheses, 2370–2374
 robotics, 1805
 skin substitutes, burn dressings, 451–452
 soft tissue prostheses
 alumina, 17
 corneal transplants, 23
 valvular prosthetics, biomineralization, 72–81
 sterilization procedures, 2501–2508
 tissue engineering
 breast, 2678–2679
 skeletal muscles, 2840–2845
 valvular prosthetics, 2737–2747
 titanium and titanium alloys
 bone implants, 2876–2887
 dental implants, 2887–2889
 total artificial lungs, 1677–1678
 total hip arthroplasty
 alumina joint prostheses, 19–21
 femoral head short-stem prostheses, 1471–1472
 long-stem prostheses, 1472
 mold arthroplasty, 1470–1471
 total knee replacement, 1609–1612
 alumina joint prostheses, 21
 urinary prostheses
 artificial sphincter, 2966–2968
 basic properties, 2961
 biomaterial-tissue interaction, 2961
 penile prostheses, 2962–2966
 clinical indications, 2964
 inflatable implants, 2963–2964
 outcomes and complications, 2964–2965
 regenerative medicine techniques, 2965–2966
 semirigid devices, 2963
 polyurethanes, 2962
 silicone, 2961–2962
 testicular prostheses, 2968–2969
 tissue interactions, 2962
 valvular prosthetics
 calcium binding site inactivation, 1308–1309
 design failures, 1312–1313
 future research issues, 1313–1314
 glutaraldehyde cross-link chemistry, 1306–1308
 overview, 1306
 remnant tissue immunogenicity, 1309–1312
 vascular grafts, 2985–2987
 hybrid materials
 biological sealants, 1432–1433
 endothelial seeding, 1433–1434
 mandrel-guided FCTs, 1434–1436
 overview, 1432
 tissue engineered substitutes, 1436–1437
 wear debris
 future research issues, 3063
 hard-on-hard bearing couples, 3060–3062
 hip prostheses, 3059–3060
 knee prostheses, 3062–3063
 overview, 3058–3059
 Prosthokaratoplasty, defined, 748
 Protein adsorption modeling
 biomaterial properties and surface analysis, 2326–2328
 biomolecular-material composites, 2326–2327
 molecular modeling, 2327
 physicochemical phenomena, 2327–2328
 overview, 2326
 simulation modeling, 2328–2333
 continuum modeling, 2328–2329
 forcefield-based molecular mechanics and dynamics, 2330–2333
 quantum mechanics, 2330
 solvent properties, 2329–2330
 Proteinases, valvular prosthetics, calcification, 78
 Proteins
 adsorption
 bacterial adherence reduction through, 64
 hemocompatibility, 1350–1353
 competitive adsorption, 1352
 conformation and orientation mechanisms, 1352–1353
 monolayer model, 1351–1352
 surface chemistry, 1352
 vascular grafts, host reactions, 3006–3007
 aquatic adhesives, 240–241
 articular cartilage repair and regeneration, 557
 biocomposites, resorbable vs. nonresorbable components, 183
 biomaterials, interaction mechanisms, 2346–2347
 biomimetic materials, 294–296
 bone matrix composition, 2044
 cell surface topography, 2558
 controlled-release drug delivery systems, 887
 drug delivery systems
 absorption and formulation issues, 2359–2360
 alternate insulin delivery routes, 2366
 enzyme and pH absorption effects, 2360–2361
 future research issues, 2367
 internal architecture and design, 2357–3
 oral insulin delivery, 2364–2366
 overview, 2357
 pH-responsive hydrogels, 2363–2364
 proteolytic activity elimination, 2362
 elastin, structural properties, 891–893
 fibrin structure and biomechanics, 1072–1077
 hydrogels, 1447
 interfacial transport
 basic principles, 2347–2348
 current applications, 2348–2349
 electrolyte concentration and permeability, 2352–2353
 future research issues, 2354–2355
 gold-palladium alloy coatings, 2354
 membrane transport, 2349–2358
 molecular mass
 artificial membranes, 2350–2351
 permeability and, 2354
 pH effects, 2351–2352
 overview and history, 2346
 pore size and permeability, 2353–2354
 water/wastewater treatment applications, 2349
 lung surfactants (SP-A/B/C)
 basic properties, 1683–1685
 disease and deficiency in, 1690
 multilayer assemblies, polyelectrolytes, 2223–2224
 osteoinductive substances and materials, 2103
 polyamides, 2186–2187
 polymers, tissue engineering applications, 2289
 structural properties, 271–272
 protein-surface interactions, 271–272
 supercritical fluid processing, 2524–2525
 dry powder formulations, 2526
 surfactants
 biomaterials applications, 2340–2343
 cationic materials, 2342–2343
 future research issues, 2344
 hydrophilicity/hydrophobicity, 2341–2342
 intracellular/extracellular mechanisms, 2339–2340
 natural materials, 2337–2338
 overview, 2337
 pH and LDH activity, 2341–2343
 semipermeable membrane transport, 2343
 tissue engineering, 2338
 vascular grafts, 2998, 3006–3007
 Protein-surface interactions
 adsorption-resistant surfaces, 278–280
 biodegradable elastomers
 future research issues, 905–906
 overview, 897
 polyesters, 897–901
 biodegradation mechanisms, 900–901
 cross-linked synthesis, 898
 linear synthesis, 898–899
 mechanical properties, 900
 polyglycolide, polylactide, polycaprolactone copolymers, 897–900
 polyhydroxyalkanoate (PHA) copolymers, 899
 polyurethanes, 901–905
 biocompatibility, 904–905
 biodegradation mechanisms, 904
 mechanical properties, 903–904
 processing and applications, 905
 synthesis, 901–904
 blood-material interactions, 373–374
 fabric scaffolds, 1058–1059
 future research issues, 281–282
 gradient surfaces, 1220–1222
 heparin biocompatibility, 1370
 neutrophils, 1935–1941
 overview, 270–271
 protein adsorption
 multicomponent solutions, 276–278
 single-component solutions, 272–276
 protein structure, 271–272
 submolecular adsorption mechanisms, 271, 280–281
 Proteoglycans (PGs)
 aggregating properties, 1197–1198
 articular cartilage
 composition and structure, 85–86, 552
 tissue engineering, 2695
 biosynthesis, 1199–1200
 cartilage degradation and aging, 1200–1201
 chondroitin sulfates, 619–622
 hyaluronic acid aggregation, 620
 large nonaggregating structures, 620–621
 small nonaggregating structures, 621
 dermatan sulfate, 619, 621–622
 small structures, 1198–1199
 structural properties, 1196–1197
 valvular prosthetics, calcification, 75
 Proteolytic enzymes
 burn dressings, 450
 polymer degradation, 219–220
 protein-based drug delivery systems, 2362–2363
 Proteomics, transgenic and gene-targeted organs, 2906
 Proton-electron double-resonance imaging (PEDRI), magnetic resonance imaging, 1720

- Proton magnetic resonance spectroscopy
brain imaging, 1969–1970
proton density, 1716–1718
- Proton resonance frequency shift (PRFS)
clinical applications, 1733–1734
temperature measurement, 1731–1733
- Prototype development, biocomposite
fabrication, 190–192
- Pseudo-diffusion coefficient, magnetic
resonance imaging, 1719
- Pseudoelasticity
blood vessel mechanics, 396
nickel-titanium alloys, 1945–1947
- Pseudophakia, defined, 748
- Psoas syndrome, overuse injuries, 1395–1396
- Public/private cooperation, intellectual
property rights, 1546
- Pull-out measurement, dental implants,
815–816
- Pulmonary circulation, vascular fluid
dynamics, 2982–2983
- Pulmonary function
cardiac output measurement, Fick method,
517–518
oxygen transport, 1843–1844
toxicity testing, animal surrogate systems,
54
- Pulmonary vasodilation, nitric oxide,
1952–1953
- Pulsatile flow properties
circulatory support systems, 1781–1782
macrocirculatory hemodynamics,
1361–1364
transmyocardial laser revascularization,
1635–1636
ventricular assist devices, 472–474
- Pulsed electric field, sterilization of medical
devices, 2507–2508
- Pulsed Fourier transform spectroscopy,
nuclear magnetic resonance, 1963–1964
- Pulsed laser radiation/ablation
diamond/diamond-like carbons, 824
laser-tissue interactions, 1626–1628
- Pulsed wave (PW) Doppler ultrasound, basic
principles, 2944
- Pulse generators, cardiac pacemakers,
535–536
- Pulse modulation, radiofrequency
electromagnetic field/radiation effects,
255
- Pulse sequence, magnetic resonance imaging,
1716
- Pumping systems
artificial heart fluid dynamics, 95–105
circulatory support systems, 1776
controlled release drug delivery, 739–741
HeartMate implantable pump, 1777–1778
insulin delivery, 1518, 2366
- Pure tone average (PTA), hearing loss and,
1265
- Purification requirements, blood composition
and structure
hemodialysis, 378–384
hemofiltration, 384
overview, 378
plasmapheresis, 384–391
- Purity requirements, silicones, 2435–2438
distillation/stripping, 2437
filtration, 2437
in-process purification, 2437
packaging materials, 2437–2438
processing environment, 2436–2437
raw material purity control, 2436
- Purkinje image eye tracker, eye movement
research, 1030–1031
- Push-out test, dental implants, 815–816
- Pushrim-activated power assist wheelchairs
(PAWPAW), 3089–3090
- P wave, electrocardiography, 938
- Pyrolysis, magnetic nanoparticles, 1708
- Pyrolytic carbon (PyC)
biocompatibility, 2375
design criteria, 2375–2381
development and applications,
2371–2374
examples of, 2379–2380
fatigue testing, 2377
fracture mechanics, 2377
future research issues, 2379
long-term implants, 2376
mechanical heart valves
durability, 1329–1330
fatigue analysis, 1330–1331
microcrack propagation/non-propagation,
2377–2378
orthopedic applications, 2374–2375
overview and history, 2370
physical and mechanical properties,
2375–2376
structural analysis, 2378–2379
- Pyrrrole-imidazole polyamides, 2191–2192
- ## Q
- QRS waves
biomedical signal acquisition and
processing, 851–852
cardiac pacemakers, 536–538
electrocardiography, 533, 938–939
- Quadratic time-frequency representations,
digital biomedical signal acquisition
and processing, 858–861
- Quantification
electromyography data, 964–965
functional magnetic resonance imaging,
1126
positron emission tomography, 2318
- Quantitative functional magnetic resonance
imaging, 1126
- Quantitative structure-activity relationships
(QSARs), toxicity testing, animal
surrogate systems, 51–52
- Quantization techniques, digital signal
acquisition and processing, 845
- Quantum mechanics, protein adsorption
modeling, 2330
- Quantum theory, biomedical optics, 2030
- Quasilinear viscoelastic theory (QLV),
ligament and tendon viscoelasticity,
2596–2597
- Quasistatic formulation, electrical
stimulation, 1010
- Quenching, melt spinning, polymer fibers,
1817–1818
- Quick response hydrogel, soft contact lenses,
1443
- Quorum sensing, biofilm formation, 209–211
- ## R
- Race and ethnicity, stress fractures and,
2512
- Radial sector scanning, endoscopic
techniques, 991
- Radiant exposure, light-tissue interactions,
1625
- Radiation
biomedical optics, 2030
sterilization of medical devices, 2505–2506
surface modification methods, 2540–2543
ion beam techniques, 2540–2541
laser, ultraviolet, and gamma radiation,
2541–2543
ultrasound systems, 2935
- Radiative transfer optical techniques, cancer
diagnosis and therapy, 2009–2010
- Radiative transport equation, light
propagation in tissue, 2033
- Radioactive coatings
diamond/diamond-like carbons, 829
stents, 2496–2497
- Radiofrequency electromagnetic field/
radiation (RFR)
basic properties, 254
exposure frequency and duration, 256
frequency, intensity and energy absorption
patterns, 254–255
future research issues, 258
research issues, 257–258
thermal/nonthermal effects, 256–257
- Radiofrequency plasma-activated chemical
vapor deposition, diamond/diamond-
like carbons, 826
- Radiofrequency plasma discharge, gradient
surfaces, 1212–1213
- Radiofrequency (RF) signals
implantable cardioverter-defibrillators,
776
magnetic resonance imaging, 1715
plasma discharge techniques, 2174–2175
speech processors, cochlear implants,
1509–1510
- Radiography, stress fracture diagnosis, 2514
- Radiopacity, coating materials, 2532
- Raman spectroscopy
biomedical optics, 2035
optical detection of cancer, 2015
- Raoult's law, cryopreservation of cells,
764–765
- Rapid acquisition with relaxation
enhancement (RARE), magnetic
resonance imaging, 1717
- Rapid expansion of supercritical solutions
(RESS), 2523
- Rapid prototyping (RP)
basic principles, 2382–2383
bioactive material fabrication, 148–149
fused deposition modeling, 2387–2388
future research issues, 2388–2389
microporous materials, 1887
selective laser sintering, 2384–2386
stereolithography, 2383–2388
three-dimensional printing, 2386–2387
tissue engineering scaffolds, 2833–2834
- Rapid-scanning all-reflective phase delay
(RSAD), optical coherence
tomography, 2000–2001
- Rapid-scanning optical phase delay (RSOD),
optical coherence tomography,
2000–2001
- Ras homology protein (RHO), integrin
postadhesion, 1529–1530
- Ras-related c3 botulinum toxin substrate
(RAC), integrin postadhesion,
1529–1530
- Rate-controlling membranes, transdermal
drug delivery, 2896
- Rayleigh's rectangular array, composite
micromechanics, 666–667
- Reaching tasks
mechanics of, 731–732
neuroprostheses, 1927–1928
- Reaction kinetics, mass transfer and tissue
function, diffusion, 1760–1761
- Reading, eye movements during, 1034–1035
- Real-time telemedicine systems, components
of, 2580–2581
- Real-time tomographic reflection (RTTR)
augmented reality and head-mounted
displays, 2391
calibration
affine transform, 2393
geometric transform, 2392–2393
current and future applications, 2393–2394
image overlay, 2391–2392
overview, 2391
techniques and procedures, 2392–2393

- Reaming techniques, total hip arthroplasty, 1377–1379
- “Receding horizon” strategy, insulin delivery, 1521
- Receiver-stimulator systems, cochlear implants, 1510–1511
- Receptor activator of NF- κ B ligand (RANKL)/receptor activator of NF- κ B (RANK)
- osteoclasts
 - basic properties, 2081
 - bone remodeling, 2088
 - differentiation and activation signaling, 2084–2085
 - wear debris, osteoclasts and, 3068–3069
- Receptor for HA-Mediated Mobility (RHAMM), structure and properties, 1422
- Receptor function
- integrins, 1526–1527
 - positron emission tomography imaging, 2320
- Recombinant DNA technology, biocatalysis and, 162
- Recombinant proteins, artificial silk design and synthesis, 2446–2447
- Recommended weight limit (RWL), ergonomics guidelines, 1002–1003
- Reconstructive biomaterials
- bone substitutes, 2401
 - chemical properties, 2397
 - evolution, 2400
 - facial skeletal fixation, 2396
 - resorbable fixation, 2398–2399
 - fracture fixation plates, 2398
 - histology, 2398
 - mechanical properties, 2398
 - medical applications, 2395
 - physical properties, 2396–2397
 - poly(methyl methacrylate), 2400
 - polyurethane terephthalate, 2401
 - resorbable fixation, 2398–2399
 - safety issues, 2395–2396
 - silicones, 2400–2401
 - tissue engineering, breast, 2678–2685
 - toxicology, 2397–2398
- Reconstructive surgery
- fat tissue engineering, 2725–2734
 - ligament injury and repair, animal repair models, 1643–1647
- Recording procedures and equipment
- electrogastrography, 949–951
 - electromyography signal conditioning, 963
- Recruitment
- of fibers, electrical stimulation, 1014
 - osteoclast differentiation and activation signaling, 2087
- Rectus femoris muscle, tendinitis, 1396–1397
- Recursive least squares (RLS), digital biomedical signal acquisition and processing, 860–861
- Red blood cells (RBCs)
- oxygen transport, 1843
 - thrombosis
 - adhesion, 2603–2604
 - normal adhesion, 2607–2608
- Reduced optical properties, tissue optics, 1625
- Reduction reaction, corrosion electrochemistry, 754–757
- Reentrant volume phase transition, soft contact lenses, 1444
- Reflectance optical imaging, biomedical optics, 2037
- Reflection, ultrasound, 2935–2936
- Refractive index
- biomedical optics, 2030–2031
 - radiative transfer, cancer diagnosis and therapy, 2010
 - ultrasound, 2935–2936
- Refractive keratoplasty, ocular implants, 748–749
- epikeratoplasty, 749
 - intracorneal ring, 749–750
 - keratophakia, 748–749
- Refractory metals, total hip arthroplasty, 1478
- Regeneration mechanisms
- articular cartilage, 554–555
 - bone healing and, 1260–1261
 - burn injuries, 449–450
 - dressing materials, 459
 - fibrin biomatrix, 1075–1077
 - hematopoietic stem cells, 1340–1342
 - nerve guide technology, 1912–1913
 - nerve tissue engineering, 2814–2816
 - penile prostheses, 2965–2966
 - polyurethane biomaterials, 2308–2309
 - skin tissue engineering, 2853–2854
 - tissue healing, 1254–1255
 - urinary prostheses, 2967–2968
 - in vivo stem cell therapy, 2476–2477
- Registration, three-dimensional imaging, 1798–1800
- Registration procedures, intellectual property and patent management, 1541
- Regularity index, digital biomedical signal acquisition and processing, 865–866
- Regulatory issues. *See also* Standards
- animal testing ethics, 1616–1621
 - controlled release drug delivery systems, 744–745
 - poly(glycolic acid), 2245–2246
 - polyurethanes, 2213–2214
 - telemedicine systems, 2584–2585
- Regurgitant flow, mechanical heart valve fluid mechanics, 1332
- Rehabilitation, robotics technology, 1804–1807
- assistive robots, 1804–1805
 - lower-limb devices, 1807
 - mobility assist devices, 1804
 - orthotocis, 1805
 - prosthetics, 1805
 - upper-limb devices, 1806
 - vocational assistance devices, 1804–1804
- Rejection syndromes. *See also* Hyperacute rejection, xenografts
- allografts, 9
 - xenografts, 3095–3099
- Relative tolerance, microelectromechanical systems, 1863
- Relaxation parameters
- dielectric relaxation, tissues, 834–838
 - magnetic resonance imaging, 1715–1718
 - nuclear magnetic resonance, 1964–1965
 - spin-echo measurements, 1965–1966
- Remnant tissue immunogenicity, heart valve bioprosthetic failure, 1295, 1309–1310
- Renal artery, distal protection device intervention, 878
- Repeating grooved substrate, topographical surface analysis, 2553
- Repetition delay, nuclear magnetic resonance, 1965
- Representative volume elements (RVEs), composite properties, 659
- micromechanics, 668–669
- Reproductive medicine, flow cytometry, 1108–1109
- Research institutes, intellectual property management, 1545–1548
- Research liaison office (RLO), intellectual property rights management, 1547
- Reservoir drug delivery systems
- ocular implants, 1987–1992
 - codrugs, 1991
 - corticosteroids, 1989–1990
 - cyclosporine, 1990–1991
 - live cells, 1991–1992
 - osmotic pump systems, 1992
 - vitrasert, 1988–1989
 - polymer systems, controlled release drug delivery, 741–742
- Resins
- composites, dental ceramics, 600
 - dental adhesive and cements, 159
- Resistive coupling, biopotential amplifiers, 305–306
- Resolution parameters
- confocal microscopy, 709
 - positron emission tomography, 2316–2317
- Resorbable materials. *See* Bioabsorbable materials
- Respiratory distress syndrome (RDS)
- lung surfactants
 - basic properties, 1683–1684
 - dysfunction physiology, 1688–1690
 - therapeutic regimes, 1690–1691
 - nitric oxide inhalation therapy, 1953–1954
- Response time, biosensor dynamics, 347–348
- Restenosis, stent complications, 2494
- Retinal detachment, scleral buckles, 1976
- Retrieval studies, hydroxyapatite coatings, 1467
- Retroviruses
- gene-activated matrix, 1144–1145
 - gene therapy and, 1155
- Reuss orthotopic modulus, elastic properties, 1246
- Reuss transverse isotropic modulus, hard materials elasticity, 1245
- Revascularization, allografts, 9
- Reversed synthesis trigonometric (RST) excitation, electrical impedance imaging, 923–924
- Reverse fabrication techniques, biocomposite fabrication, 191–192
- Reversible addition-fragmentation polymerization (RAFT), 2279–2280
- Reynolds number
- artificial lung systems, mass transfer correlations, 1666
 - cardiac assist devices, positive displacement pump fluid dynamics, 97–105
 - flow cytometry, fluidic systems, 1101–1102
 - vascular fluid dynamics, 2974–2975, 2980–2984
- Rheology
- bone cement, 406–407
 - bone cement polymers, 406–407
 - melt spinning, polymers, 1816–1817
 - vascular fluid dynamics, 2977–2979
- Right heart catheterization, techniques and materials, 504
- Rigid transformation models, three-dimensional imaging techniques, 1800
- Ring opening polymerization (ROP)
- poly(lactic acid), 2250–2251
 - poly(lactide-*co*-glycolide), 2261
- Risk factors, stress fractures, 2512–2513
- Robotics
- computer-assisted surgery
 - cardiac valve surgery, 691
 - da Vinci system, 689–690
 - esophageal surgery, 691
 - evolution of, 687–688
 - future applications, 692–693
 - overview, 687
 - pelvic surgery, 691–692
 - Zeus system, 688–689
 - eye movement tracking research, 1028–1031
 - medical applications
 - educational training tools, 1812
 - future research issues, medrobo–10
 - overview, 1804
 - rehabilitation, 1804–1807
 - surgery, 1807–1812

- Robotics (*cont.*)
 microelectromechanical systems, 1866–1870
 movement control models, 732–734
 tele-echography, 1812
 telemedicine and, 2580
 wheeled mobility devices, 3089–3090
- Rods (orthopedic), fracture fixation, 2074–2075
- Roller pumps, cardiac assistance, 471
- Root mean-square value (RMS), biomedical signal acquisition and processing, 853
- Rotary cell culture systems (RCCS)
 cancer, 2405–2406
 cardiac patch engineering, 548–549
 cartilage, 2405
 future applications, 2406–2407
 infectious disease, 2406
 insect cells, 2406
 liver, 2406
 neural tissue, 2406
 overview, 2402
 pancreatic islets, 2406
 theoretical principles, 2402–2403
 tissue engineering applications, 2403–2410
- Rotating frame experiments, nuclear magnetic resonance, 1964
- Rotating-wall bioreactor
 articular cartilage repair, 559
 cardiac patch engineering, 546–549
- Roughness properties
 diamond/diamond-like carbons, 822–823
 topographical surface analysis, 2555–2556
- Rugophilia, cell-material interactions, 569–570
- Rule-based controllers (RBCs), movement control, 734–737
 hybrid models, 736–737
- Running spectral analysis (RSA), electrogastrography signals, 951–953
- S**
- Saccadic movement, eye movements, 1028
 perception and, 1033–1034
- Safety issues
 cell culture assays, 566
 electrical impedance imaging, 918–923
 electrocardiography, 943
 electrospun polymer nanofibers, 981–982
 reconstructive biomaterials, 2395–2396
- Salamander glue, 245
- Sampling theory
 digital signal acquisition and processing, 844–845
 positron emission tomography, 2317
- Sandblasting, zirconia ceramics, 3108
- Saphenous vein grafts
 distal protection devices, 870
 tissue engineering, 2643
- Saturated steam sterilization, 2504
- Scaffold materials
 articular cartilage repair, 555–556
 chemical factors, 557–558
 selection criteria, 555–556
 artificial liver systems, 1649–1658
 bioactive materials
 calcium phosphates, 143
 composites, 144–145, 149
 fabrication, 145–150
 foam coatings, 149–150
 glass-ceramics, 138–139, 144–150
 glass scaffolds, 138–139
 microsphere sintering, 149
 overview, 142
 powder-based fabrication, 147
 scaffold structures, 142
 silicate glasses, 143–144
 sol-gel fabrication, 147–148
 solid free-form fabrication, 148–149
 solvent casting/particle leaching, 149
 thermally induced phase separation, 149
- biocomposites, 180–181
 nanotechnology, 192–193
 biofunctional polymers, 218–222
 bone matrices, 2044–2047
 cardiac patch engineering, 543–544
 biomaterial substrates, 546
 cell-material interactions, 569–570
- ceramics
 hard tissue interactions, 1231–1232
 transparent ceramics, 2916–2923
- chitosan, 611–613
- collagen
 fibril formation in, 639–643
 fixation properties
 future research issues, 650
 overview, 639
 strategies and techniques, 642–649
 tissues and materials, 639–643
- degradable polymer-ceramics, 784
- elastin
 basic properties, 891
 bioengineering applications, 894–895
 gene structure, 891
 protein structure, 891–893
 tropoelastin assembly and ultrastructure, 893–894
- electroactive polymeric materials, 926–933
- electron microscopy analysis, 974–975
- electrospun polymer nanofibers, 982–983
- extracellular matrix
 component properties, 1022–1024
 host response, 1024
 overview, 1021
 sources, 1024
 tissue repair, 1024–1025
- fabric scaffolds, 1056–1059
 pore size distribution, 1057–1058
- fibrin structure and biomechanics, 1076
- fibroblast mechano-regulation mechanisms, 1788–1789
- gene-activated matrix
 applications, 1142–1143
 basic properties, 1142
 biodegradable scaffolds, 1145–1146
 tissue regeneration, 1146–1148
 vector systems, 1144–1145
 wound healing, 1143
- hard tissue interactions, 1229–1233
- histogenesis, 1407–1408
- hyaluronan
 applications, 1426–1429
 antiadhesion, 1427
 cardiovascular, 1427–1428
 drug delivery systems, 1429
 neural and glial applications, 1429
 ophthalmology, 1428
 orthopedic, 1427
 tissue engineering and surgery, 1429
 wound healing and dermatology, 1428
- basic properties, 1421
 cell degradation and, 1423
 cell migration and, 1422–1423
 cross-linked biomaterials, 1425–1426
 derivatized biomaterials, 1424–1425
 extracellular matrix and hyaladherin interactions, 1422
 future research issues, 1429
 pathological processes and, 1423–1424
 physicochemical and structural properties, 1421–1422
 wound repair and angiogenesis, 1423
- hydroxyapatites, 1456–1460
- matrix metalloproteinases
 collagenases, 1766–1768
 gelatinases, 1768–1769
 historical background, 1763–1764
 membrane-type, 1771
 overview, 1763
 stromelysins, 1769–1771
 structure and regulation, 1764–1765
 tissue engineering applications, 1771–1772
- melt-spun porous scaffolds, 1822
- microporous materials
 applications, 1885–1886
 bioactive materials, 1886
 ceramics, 1886–1887
 composites, 1890–1891
 foaming of ceramics and glasses, 1888–1891
 melt-derived bioactive glasses, 1890
 overview, 1885
 sol-gel bioactive glasses, 1887–1888
 tissue engineering applications, 1885
- nerve guide technology, 1911–1921
- nitric oxide, 1955–1956
- photopolymerization
 biocompatible photoinitiators, 2152–2153
 biomaterials, 2154–2158
 hydrogels, 2154–2158
 mechanisms, 2150–2152
 overview, 2150
 photografting techniques, 2158
 tissue engineering applications, 2158–2160
- poly(glycerol sebacate), 333
- poly(glycolic acid)
 applications, 2244–2245
 basic properties, 2241
 characterization, 2242
 degradation, 2242–2244
 regulatory issues, 2245–2246
 synthesis, 2241–2242
 in vivo tissue response, 2246–2247
- poly(lactic acid)
 biomedical applications, 2254–2255
 characterization, 2252
 degradation, 2252–2253
 in vivo characteristics, 2253–2254
 formation mechanisms, 2250–2251
 overview, 2249
 synthesis, 2249–2250
- poly(lactide-co-glycolide), 2259–2267
- polymer foams, 2271–2274
- polymers
 ideal properties, 2284
 tissue engineering applications, 2283–2284
- polyurethanes, 2304–2312
- rapid prototyping
 basic principles, 2382–2383
 fused deposition modeling, 2387–2388
 future research issues, 2388–2389
 selective laser sintering, 2384–2386
 stereolithography, 2383–2388
 three-dimensional printing, 2386–2387
- scanning electron microscopy analysis, 972–973
- silks, 2446
- synthetic apatite composites, 232
- tissue engineering
 articular cartilage, 2698–2700
 basic principles, 2828–2829
 bladder, 2636–2641
 blood vessels, 2643–2647
 bone, 2654–2655
 bone marrow, 2660–2665
 culture systems, 2670–2675
 breast, 2681
 cardiac muscle, 2688–2692
 esophagus, 2720–2721
 fabrication techniques, 2829–2835
 acellular matrix and submucosa, 2832–2833
 connective tissue capsule, 2833

- electrospinning, 2834–2835
 extrusion, 2831
 fiber bonding, 2829
 freeze drying, 2831–2832
 gas foaming, 2831
 melt molding, 2830–2831
 membrane lamination, 2830
 phase separation, 2832
 rapid prototyping, 2833–2834
 solvent casting and particulate leaching, 2829–2830
 fat tissue, 2729–2730
 gastrointestinal system, 2754
 historical perspective, 2621
 intestinal epithelium, 2862–2864
 ligaments, 2759
 liver, 2771–2776
 microvascular networks, 2794–2798
 polymer ideal properties, 2284
 skeletal muscles, 2841–2845
 tendons, 2873
 vascular system, 3033–3035
 valvular prosthetics
 biomimetic constructs, 2743–2744
 collagen fiber bundles, 2744
 cross-linked hyaluronans, 2744–2746
 elastin, 2745–2747
 matrix mechanics, 2742–2743
 vascular grafts, 1436–1437
 Scale-up technique
 bio-artificial liver systems, 1656
 magnetic/electrophoretic cell separation, 1702
 Scaling function
 digital signal acquisition and processing, 856–858
 microelectromechanical systems, 1867
 Scan-along polygonal approximation (SAPA), ECG signal compression, 680–681
 Scanning electron microscopy (SEM)
 bioreactor analysis, 312
 hemocompatibility assessment, 1349
 instrumentation, 968–969
 orthopedic devices, failure analysis, 1828–1839
 overview, 968
 self-assembled monolayers, 2412–2413
 specimen preparation, 971–972
 TEM *vs.*, 969–970
 Scanning probe microscopy (SPM), self-assembled monolayers, 2412–2413
 Scattering theory
 biomedical optics, coefficient, 2031
 laser-tissue interaction, 1624–1625
 positron emission tomography, 2318
 radiative transfer, cancer diagnosis and therapy, 2010
 tissue optics, 1624–1625
 ultrasound-tissue interaction, 2936–2939
 Scene perception, eye movements and, 1033–1034
 Schmidt number, artificial lung systems, mass transfer correlations, 1666
 Schwann cells
 nerve guide technology, 1914
 peripheral nerve injuries, 2120
 graft materials and techniques, 2128
 Schwarz's theory, biomaterials, 837
 Sciatic function index (SFI), nerve tissue engineering, 2813–2814
 Scintigraphy, stress fracture diagnosis, 2514–2515
 Scleral buckles and plugs
 matrix ocular implants, 1984–1985
 structure and properties, 1976
 Scoliosis, spinal deformity biomechanics, 2469–2470
 Scooter systems, design and components, 3088–3089
 Screening tests, cell culture assays, 563–564
 Screw fixation, 2074–2075
 bioabsorbable plates and screws, 432–437
 bone fractures, 414–415
 calcium phosphate ceramics, 466
 failure analysis
 dynamic compression plate, 1829–1832
 titanium screws, 1834–1835
 interference screws, 2077
 total hip arthroplasty, pelvic osteolysis and, 1380–1381
 Sealants
 endovascular stent graft fixation, 2488–2489
 fibrin glue, 1186–1187
 fibrin structure and biomechanics, 1072–1077
 hybrid vascular prostheses, 1432–1433
 osteoclast morphology, 2081–2082
 Seamless synthesis and processing, polymer materials, 2280
 Sea star/sea cucumber, adhesive properties, 242–243
 Seating design, wheeled mobility systems, 3087–3088
 Secondary ion mass spectrometry (SIMS), hemocompatibility assessment, 1349–1350
 Second-generation bioadhesives, 155–156
 Second messenger systems, nitric oxide, 1952
 Security issues, telemedicine systems, 2585–2586
 Seeding techniques, tissue engineering, vascular system, 3035–3036
 Segmentation techniques, three-dimensional imaging processing and analysis, 1797–1798
 Segmented polyurethane biomaterials, structure and properties, 2304–2305
 Segré-Silberberg effect, vascular fluid dynamics, 2980
 Self-assembled monolayers (SAMs). *See also* Multilayer assemblies
 basic properties, 2409
 biotechnological applications, 2412–2415
 bone matrices, 2407
 formation mechanisms, 2409–2412
 metal-alkanethiol interactions, 2409–2411
 hemocompatibility, inert materials, 1346
 light emitting diode assembly, 2222–2223
 microelectromechanical systems
 DNA assembly, 1869–1870
 historical background, 1861–1862
 polymers, tissue engineering, 2295
 protein-surface interactions, 278–281
 surface modification, 2545
 tissue engineering, corneal implants, 2711–2712
 Self-consistent formulations, composite micromechanics, 668–669
 Self-expanding stents, endovascular stent grafts, 2486
 Self-generating technologies, tissue engineering, 2621–2622
 Self-propagating high-temperature synthesis (SHS), microporous materials, 1887
 Self-regulated insulin delivery, glucose hydrogels, 1165–1170
 Self-reinforced composite materials
 bioabsorbable plates and screws, 433–437
 degradable polymer composites, 781–783
 Semi-autonomous surgical robotics, 1809–1811
Semibalanus balanoides, adhesive properties, 242
 Semilunar heart valves
 hemodynamics, 1319
 mechanics, 1319–1320
 structure and function, 1317–1319
 Semimembranosus tendinitis, 1604
 Semipermeable membrane
 artificial pancreas, 2802–2804
 humoral immunity efficacy, 2804–2805
 protein-surfactant transport, 2343
 Senofilcon silicone hydrogel contact lenses, 727
 Sensing amplifiers, implantable cardioverter-defibrillators, 776
 Sensitivity parameters
 biosensors, 336, 347
 positron emission tomography, 2317
 small-angle light scattering, soft tissue analysis, 2454–2456
 Sensitization testing, in-vivo biocompatibility, 174
 Sensorineural hearing loss, mechanisms of, 1264–1265
 Sensors
 biosensors
 biological recognition elements, 339–343
 cell-based sensors, 343–344
 components of, 338–339
 detector elements, 344–345
 microfabrication, 345
 overview, 335–339
 response time and dynamics, 345–348
 sensitivity, 347
 glucose sensors
 blood glucose sensors, 1180–1182
 clinical applications, 1174–1175
 interstitial fluid sensors, 1175–1179
 affinity chemical/optical sensor, 1180
 dialysis catheter monitoring system, 1178–1180
 enzyme-based telemetry sensor, 1180
 needle-type, enzyme-based
 electrochemical biocatalytic sensors, 1175–1178
 overview, 1174
 telemedicine applications, 2579–2780
 Sensory assessment, peripheral nerve injury and repair, 2121–2122
 Separation technology
 interfacial transport, 2350
 polyelectrolyte multilayers, 2223
 Sequence control, polymer materials, 2280–2281
 Serial processing
 microassembly, 1867–1868
 microelectromechanical systems, 1863
 microassembly techniques, 1867–1870
 Shadow amplitude testing, tendon/ligament tensile testing, 2591–2592
 Shape forming, cardiac catheters, 495–496
 Shape memory alloys (SMAs). *See also* Distal protection devices
 artificial muscle materials, 108, 110
 metals
 basic properties, 2418–2419
 biomedical applications, 2419–2420
 nickel-titanium alloys, 1945–1947, 2063
 overview, 2418
 surface modifications, 2420–2422
 ceramic coatings, 2423
 chemical and electrochemical treatments, 2422–2423
 hydroxyapatite coatings, 2423
 oxidation, 2422
 polishing, ultrasonic bath, and physical etching, 2421–2422
 polymer coatings, 2423–2424
 sterilization, 2424
 titanium and titanium alloys, 2879
 orthopedic devices, 2063
 polymers, artificial muscle materials, 108
 Shared medium systems, toxicity testing, animal surrogate systems, 57

- Shear forces
 articular cartilage repair, 559
 blood vessel mechanics, 394–395
 composite micromechanics, 667–671
 ligaments and tendons, multiaxial
 mechanical testing, 2594–2595
 valvular prosthetics
 failure analysis, 1299
 mechanical heart valves, 1334
- Shear modulus, hard tissue elasticity, 1240–1241
- Shell components, composition and tissue engineering, 2049
- Sherwood number, artificial lung systems, mass transfer correlations, 1666
- Short bowel syndrome (SBS)
 intestinal tissue engineering, 2751–2752
 small intestine tissue engineering, 2860–2861
- Short-stem prostheses, total hip arthroplasty, 1471–1472
- Short Time Fourier Transform (STFT), digital biomedical signal acquisition and processing, 858–861
- Shot noise, optical mapping artifacts, 2022
- Shoulder ligaments, tissue engineering, injuries and reconstruction, 2759
- Shrinkage properties, bone cement polymers, 407–408
- Sickle cell disease (SCD), nitric oxide inhalation therapy, 1954
- Sieving coefficient, artificial kidney performance, 1589–1590
- Signaling mechanisms. *See also* Cell signaling
 biomedical signal compression
 efficiency evaluation, 678
 electrocardiogram, 679–681
 electroencephalography, BAEP recordings, 683–684
 electromyography, 681–682
 electrooculography, 682–683
 fidelity and clinical quality, 678–679
 future applications, 684–685
 lossless, near losses, and lossy methods, 677–678
 motivation, 677
 overview, 677
 digital acquisition and processing, 845–849
 acquisition principles, 843–845
 digital filter, 845–846
 event detection, 851–853
 overview, 843
 quantization, 845
 sampling process, 844–845
 signal averaging, 848–849
 signal enhancement, 845–847
 spectral analysis, 849–851
 electrogastronomy imaging, 951–954
 electromyography, 959–960
 conditioning and data recording, 962–963
 integrins, 1532–1533
 magnetic resonance imaging, 1714–1715
 stem cell differentiation, 2475
 surfactants, 2337–2344
 tissue engineering, intestinal system, 2574–2755
- Signaling molecules, tissue engineering, bone, 2656–2657
- Signal-to-noise ratio (SNR)
 digital signal acquisition and processing, 845
 hearing loss and, 1265
 magnetic resonance imaging, 1717
 magnetic resonance microscopy, 1738–1741
 nuclear magnetic resonance, 1961
- Signal transduction
 host reactions, particulate matter, 2112–2114
 osteoclast differentiation and activation signaling, 2087
- Silicate glasses
 bioactive materials, tissue engineering, 143–144
 fiberoptic biosensors, 1063–1064
- Silicate-substituted hydroxyapatite (SiHA)
 basic properties, 1454
 clinical applications, 1456–1460
- Silicone
 basic properties, 2427
 biocompatibility and bi durability, 2438–2439
 breast implants, 2678–2679
 cardiac catheters, 496
 cardiac pacing leads, 535
 hydrogels, contact lenses
 chemical coatings, 725
 core surface treatments, 725
 current commercial sources, 725–727
 evolution of, 718–719
 hydrophilic components, 720
 hydrophobic comonomers, 721
 ion permeability and on-eye movement prediction, 722
 ocular environment, 715
 overview, 715
 oxygen permeability, 715–717
 on phases, 722–723
 plasmas, 724–725
 polymer phases, 721–722
 recent developments, 719
 research background, 717–718
 soft lenses, 1441
 surface properties, 723–725
 transport properties, 721
- low-molecular-weight fluids, 2429
- macromonomers, 2427–2429
- multilayer assemblies, polyelectrolytes, 2221
- nerve guides, 1914–1915
- nomenclature and classification, 2427
- oil, ocular implants, 1979
- polydimethylsiloxane structure, 2427
- polyurethanes and ureas, 2433–2434
- purity requirements, 2435–2438
 distillation/stripping, 2437
 filtration, 2437
 in-process purification, 2437
 packaging materials, 2437–2438
 processing environment, 2436–2437
 raw material purity control, 2436
- reconstructive biomaterials, 2400–2401
- self-assembled monolayers
 hydroxylated surface, 2411–2412
 organosilicon derivatives, hydroxylated surfaces, 2411–2414
- shunts, antiseptic/antibiotic materials, 66
- surface analysis, 2434–2435
- thermosetting elastomers, 2430–2433
 condensation cure, 2431–2433
 fillers, 2433
 hydrosilation cure, 2430
 peroxide cure, 2430–2431
- urinary prostheses, 2961–2962
- Silk
 artificial materials design and synthesis, 2446–2447
- Bombyx mori* silk
 biosensors, 2445–2446
 chemical properties, 2445
 fiber formation, 2443–2444
 mechanical properties, 2445
 structural properties, 2442–2443
 sutures, 2445
 tissue engineering scaffolds, 2446
- dope, adhesive properties, 247
 polyamides, 2187–2188
 sericins, adhesive properties, 246
 spider silks, 2448
 structural properties, 2442
 tissue engineering applications, 2290
 wild silkworm silks, 2447–2448
- Silver compounds
 as antimicrobial biomaterials, 66–67
 burn dressings, 456–457
- Silver sulphadiazine cream, burn dressings, 457
- Simulation modeling
 protein adsorption modeling, 2328–2333
 continuum modeling, 2328–2329
 forcefield-based molecular mechanics and dynamics, 2330–2333
 quantum mechanics, 2330
 solvent properties, 2329–2330
 robotics and, 1811
- Sinding-Larsen-Johansson disease, overuse injuries, 1603
- Single cell in vitro models, toxicity testing, animal surrogate systems, 53
- Single-chamber pacing, cardiac pacemakers, 536–537
- Single-component protein solutions, adsorption behavior, 272–276
- Single-microphone noise reduction algorithms, hearing aid design, 1268
- Single photon emission computed tomography (SPECT), basic principles, 1795–1796
- Single-slit diffraction theory, small-angle light scattering, soft tissue analysis, 2453–2454
- Single-stage venous EC seeding, vascular grafts, 2992–2993
- Single time constant relaxation, tissue dielectric properties, 834–836
- Sinoatrial node (SAN)
 cardiac conduction system, 532–533
 cardiac electrophysiology, 480, 487–488
 optical mapping, atrioventricular node conduction, 2024–2025
- Sintering
 dental ceramics, 600–601
 hydroxyapatites, 1455–1456
 microporous materials
 HA and TCP ceramics, 1886
 sol-gel bioactive glasses, 1888
 microsphere sintering, bioactive scaffold fabrication, 149
 polymer foam processing, 2273–2274
 porous hydroxyapatite, tissue engineering, bone, 2655
 rapid prototyping, selective laser sintering, 2384–2386
 scaffold materials, melt-spun porous scaffolds, 1822
- Sinuses of Valsalva, 1317–1319
- Sinus node dysfunction, cardiac pacemakers, 536–539
- Sit and reach test, hamstring syndrome, 1399–1400
- Skeletal muscle
 movement control, 729–733
 tissue engineering
 electrospinning strategies, 2843–2844
 future directions, 2845
 growth and muscle adaptation, 2842–2843
 historical perspective, 2839–2840
 limitations, 2839
 prosthetics fabrication, 2840
 scaffold materials, 2841–2842, 2845–2848
 structure and function, 2838–2839
 in vivo techniques, 2840–2841, 2844–2845

- Skin anatomy and function
 optical coherence tomography, 2004–2005
 tissue engineering, 2851–2852
 transdermal drug delivery, 2893–2894
 reaction mechanisms, 2903–2904
- Skin grafts
 adhesives, basic properties, 158
 allografts, 10–11
 anatomy and function, 2851–2852
 burn injuries and, 448
 cadaver skin grafts, 451–456
 skin substitutes, biological and synthetic, 451–456
 fibrin sealant applications, 1085–1086
 hyaluronan applications, 1428
 polycaprolactone, 2202
 tissue engineering
 current treatments, 2852–2853
 dermis regeneration template, 2854–2856
 epidermal autografts, 2853
 equivalents and regeneration templates, 2853–2857
 future research, 2856–2857
 injury response, 2852
 living dermal replacement, 2854
 microscale technology, 2788
 structure and function, 2851–2852
 wound healing, 2851–2852
- Skin penetration
 glucose sensor implants, 1175–1177
 tissue engineering, gene-activated matrix regeneration, 1146–1147
 toxicity testing, animal surrogate systems, 53–54
- Slice selection, magnetic resonance imaging, 1716
- Sliding filament theory, mathematical modeling, 1895–1896
- Slow Turning Lateral Vessel (STLV), rotary cell culture systems, 2404
- Small-angle light scattering (SALS), soft tissue analysis
 angular position accuracy and multiple population sensitivity, 2454–2456
 beam diameter and spatial resolution, 2456–2458
 calibration and sensitivity, 2454–2459
 collagen crimp effects, 2458–2459
 collagen fiber mapping, 2459
 experimental techniques, 2451–2454
 fiber diameter and interfiber spacings, 2458
 future research issues, 2461
 hardware requirements, 2452
 light scattering patterns, 2453–2454
 overview, 2450–2451
 tissue engineering applications, 2459–2461
 tissue thickness effects, 2456
- Small intestine
 submucosa, blood vessel scaffolds, 2647
 tissue engineering
 animal models, 2755–2758
 clinical applications, 2860–2861
 epithelial cell culture, 2861–2862
 evolution of, 2860
 organoid unit models, 2864–2866
 recent developments, 2863–2864
 scaffold materials, 2753–2755, 2862–2863
 stem cells, 2758–2759, 2867–2868
 subepithelial matrix extension, 2866–2867
- Small leucine-rich repeats (SLRPs), proteoglycans, 1196–1197
 degradation and aging, 1201
 small proteoglycans, 1198–1199
- Small molecule integrin antagonists, angiogenesis inhibitors, 28
- Small organic molecules, polymer structure/property relationships, 2275–2276
- Smart polymers
 insulin delivery, 1519
 protein-based drug delivery systems, 2362–2363
 tissue engineering
 clinical applications, 2294
 liver, 2772
- Smart tissue-activated fiberoptic biosensors, 1066–1067
- Smith Model-Collagen Microfibril Arrangement, collagen synthesis, 632–633
- Smooth muscle cells (SMCs)
 gastrointestinal tissue engineering, 2753–2754
 tissue engineering, blood vessels, 2646–2647
 valvular prosthetics, collagen fiber bundles, 2744
 vascular grafts, 2998
 host reactions, 3008
- Snapping hip syndrome, overuse injury, 1397–1399
- Snell's law, ultrasound reflection/refraction, 2935–2936
- S-nitrosohemoglobin, nitric oxide, 1951–1952
- Snook & Ciriello's Liberty Mutual maximum acceptable weights and forces tables, ergonomics guidelines, 1003
- Sodium polyaspartate, 2191
- Soft lithography
 microelectromechanical systems, 1862–1863
 tissue engineering, microscale technology, 2780–2781
- Soft tissue
 adhesive bonding implants, 2–5
 alumina implants, 17
 corneal transplants, 23
 articular cartilage biomechanics, 85–91
 bioactive glass, 136–137
 biomineralization, 71–73
 bone healing, soft callus formation, 1256–1258
 chondroitin sulfates
 biological properties, 622
 biomaterials applications, 626
 future research issues, 626
 overview, 618
 pharmacological effects, 624–625
 proteoglycans, 619–622
 hyaluronic acid aggregation, 620
 large nonaggregating structures, 620–621
 small nonaggregating structures, 621
 structure, 618–619
 dermatan sulfate
 biological properties, 623–624
 future research issues, 626
 overview, 618
 pharmacological applications, 625–626
 proteoglycans, 619, 621–622
 structure, 618–619
 fat tissue, tissue engineering, 2725–2734
 fibroblast mechano-regulation, 1783–1790
 gene-activated matrix regeneration, 1146–1147
 healing properties, 1249–1250
 histiogenesis, 1407
 immune response, major histocompatibility complex, 266–267
 orthopedic devices
 allografts, 13–14
 bone interface, 2076–2077
 periodontal ligament, biomechanics, 799–800
 poly(lactide-co-glycolide) biocompatibility, 2265–2267
- small-angle light scattering analysis
 angular position accuracy and multiple population sensitivity, 2454–2456
 beam diameter and spatial resolution, 2456–2458
 calibration and sensitivity, 2454–2459
 collagen crimp effects, 2458–2459
 collagen fiber mapping, 2459
 experimental techniques, 2451–2454
 fiber diameter and interfiber spacings, 2458
 future research issues, 2461
 hardware requirements, 2452
 light scattering patterns, 2453–2454
 overview, 2450–2451
 tissue engineering applications, 2459–2461
 tissue thickness effects, 2456
 spine biomechanics, 2466–2467
 tendons and ligaments
 multiaxial mechanical testing
 anisotropic properties, 2592–2593
 bidirectional testing, 2593–2594
 shear loading, 2594–2595
 structural framework, 2588–2589
 uniaxial tensile testing, 2589–2592
 viscoelasticity, 2595–2597
 in vitro analysis, 2591–2592
 tissue engineering, 1255
 ligaments, 2758–2766
- Software requirements, electrical impedance tomography, 919
- Sol-gel technology
 ceramics and glasses, 1470148
 microporous materials
 bioactive glasses, 1887–1888
 foaming technology, 1888–1890
- Solid free-form fabrication techniques, ceramics and glasses, 148–149
- Solid matrices, tissue engineering, cardiac muscle, 2686–2687
- Solid state nuclear magnetic resonance (ssNMR)
 basic principles, 1966–1967
 protein-surface interactions, 280–281
- Solid-state reactions, hydroxyapatite synthesis, 1455
- Solubility parameters
 poly(lactide-co-glycolide), 2261–2262
 transdermal drug delivery systems, 2894–2895
- Soluble signals, bone healing, 1259–1260
- Solvent casting/particulate leaching (SC/PL)
 bioactive scaffold fabrication, 149
 biocomposite fabrication, 186
 tissue engineering scaffolds, 2829–2830
- Solvents
 diffusion, gradient surfaces, 1211–1212
 exogenous lung surfactants, 1691–1692
 protein adsorption modeling, 2329–2330
 resistance, polyurethanes, 2212
- Sonic Accelerated Fracture Healing System, 2952–2953
- Sorbent systems, plasmapheresis, 390–391
- Sound levels, hearing mechanisms and, 1273–1274
- Sound pressure levels (SPLs), hearing aid design and, 1264–1267
- Sound production, hearing aid design and, 1263–1264
- Spark plasma sintering (SPS), transparent ceramics, 2916–2917
- Sparse Gaussian elimination, blood flow modeling, microvascular tissue engineering, 2793–2794
- Spatial parameters
 eye movement perception, 1033–1034
 magnetic resonance imaging, 1716

- Spatial parameters (*cont.*)
 microscale tissue engineering, 2782
 small-angle light scattering, beam diameter and, 2456–2457
- Spatio-temporal parameters
 gait analysis, 1130–1131
 mass transfer and tissue function, 1749–1752
- Special education robotics, 1812
- Specific absorption rate (SAR),
 radiofrequency electromagnetic field/
 radiation effects, 255
- Specimen preparation, electron microscopy,
 970–972
- Spectral analysis. *See also* Imaging techniques
 biomedical optics, 2034–2035
 cancer diagnosis and therapy
 fluorescence imaging, 2014–2015
 radiative transfer principles, 2011–2012
 digital signal acquisition and processing
 basic principles, 849–850
 higher-order spectra, 861–864
 nonstationary signals, 858–861
 electrogastrography signals, 951–954
 flow cytometry, 1104
 magnetic resonance imaging, 1719–1720
 optical coherence tomography, 2002–2003
- Speech processor systems, cochlear implants,
 1508–1510
- Speech recognition. *See also* Automatic
 speech recognition (ASR) systems
 hearing mechanisms and, 1275–1276, 1278
- Sperm analysis, flow cytometry, 1108–1109
- Spermatozoon, cryopreservation techniques,
 767–768
- Spider silks, structure and properties, 2448
- Spinal cord injury, functional magnetic
 resonance imaging, 1125–1126
- Spinal fusion, biomechanics of, 2467–2468
- Spine biomechanics
 deformations, 2469–2470
 disc degeneration, 2467–2468
 arthroplasty, 2468
 spinal fusion, 2467–2468
 fractures, 2468–2469
 future research issues, 2471
 healthy spine
 kinematics, 2464–2465
 kinetics, 2465–2466
 tissue stiffness and strength, 2466–2467
 research overview, 2464
 stress fractures, 2518–2519
- Spin-echo measurements, nuclear magnetic
 resonance, 1965–1966
- Spin-lattice/spin relaxation, magnetic
 resonance imaging, 1716
 temperature measurement, 1729–1730
- Spin-spin couplings, nuclear magnetic
 resonance, 1962–1963
- Spongiosa, aortic valve cusp microstructure,
 2738–2742
- Spray drying, polymer particulates,
 controlled-release drug delivery
 systems, 882
- Spraying process, plasma biomaterials
 modification, 2183–2814
- Sputtering process, plasma biomaterials
 modification, 2176
- Squared coherence spectrum (SCS),
 biomedical signal acquisition and
 processing, 852–853
- Stability properties
 biostable ceramics, orthopedic devices, 2066
 polymer degradability/nondegradability,
 2284–2285
 polyurethane biomaterials, 2307–2308
 protein-based drug delivery systems,
 2358–2359
 vascular grafts, 2987
- vitrification and, 3049–3050
 zirconia ceramics, 3104–3105
- Stainless steel
 alloys
 nickel alloys, 1943–1945
 total hip arthroplasty, 1474–1478
 endovascular stent grafts, 2486–2487
 fracture fixation devices, bone, 414
 implant devices, galvanic corrosion,
 755–756
 metal orthopedic devices, fracture fixation
 and joint implantation, 1826–1827
 orthodontic wires, 2056–2057
 orthopedic devices, 2061–2062
 total hip arthroplasty, 1474–1476
- Stamping technique, polyelectrolyte
 multilayer deposition, 2219–2221
- Standards. *See also* Regulatory issues
 biomedical ethics and, 287–289
 cell culture assays, 563
 ergonomics guidelines, 1000–1006
 wheeled mobility devices, 3089–3090
- Staples, soft tissue-bone interface, 2077
- Starches, tissue engineering applications,
 2292
- Starling resistor, venous hemodynamics,
 1366
- Static cell cultures, bone marrow tissue
 engineering, 2670–2671
- Static interlocking fixation, intramedullary
 nailing, bone fracture fixation, 418–419
- Statistical noise, positron emission
 tomography, 2317–2318
- Steady flow models, macrocirculatory
 hemodynamics, 1359–1361
- Steady-state wear equation, total hip
 arthroplasty implants, 1479
- Stem cells. *See also* specific types of stem cells,
 e.g. Embryonic stem cells
 adipose stem cells, tissue engineering
 applications, 2293–2294
 analysis of, 2473–2474
 articular cartilage repair, 556–557
 basic properties and classification, 2473
 biomaterials and research on, 2475–2476
 cardiac patch engineering, 543–544
 cell-biomaterial interactions, 2475
 differentiation mechanisms, 2475
 future research issues, 2477
 hematopoietic stem cells, 1338–1342
 mesenchymal stem cells, 1341–1342,
 2091–2092
 osteoclasts, bone regeneration, 2095–2096
 phenotype, 2473
 sources of, 2474
 tissue engineering
 articular cartilage, 2702
 blood vessels, 2645
 bone, 2653–2654
 ligaments, 2766
 pancreas, 2806–2608
 peripheral nerve cells, 2816
 skeletal muscles, 2840
 small intestine, 2867–2868
 in vivo therapy and, 2476–2477
- Stent coatings and devices
 basic properties, 2492
 biodegradable, 2497
 biomechanics, 2494–2495
 clinical applications, 2492–2493
 coatings
 inorganic coatings, 2496
 thromboresistant coatings, 2496
 complications, 2493–2494
 in-stent restenosis, 2494
 subacute thrombosis, 2493
 controlled-release systems, 887
 diamond/diamond-like carbons, 829
 drug-eluting, 2497–2498
- endovascular grafts
 balloon-expanded stents, 2486
 bifurcated grafts, 2482
 branching techniques, 2480–2481
 breakage issues, 2486–2487
 column strength, 2487
 delivery system diameter, 2489–2490
 design innovation, 2484
 fabrics, 2484–2485
 fenestrated grafts, 2483
 fixation techniques, 2487–2488
 future designs, 2487
 medical applications, 2480
 metal alloys, 2494
 modularity, 2483–2484
 multibranched grafts, 2482–2483
 sealing techniques, 2488–2489
 self-expanding stents, 2486
 support systems, 2489
 unbranched grafts, 2480–2482
 future research issues, 2498
 hemocompatibility, 1347–1348
 heparin biocompatibility, 1371–1372
 historical perspectives, 2492
 membrane-covered devices, 2496–2497
 nickel alloys, 1942–1949
 titanium-nickel, 1945–1947
 phosphorylcholine technology, 2143–2145
 radioactive, 2497
 shape memory metals, 2419–2420
 surface treatments, 2495–2496
 in vivo animal models, 1570
- Stentless heart valve prosthesis
 design criteria, 1313
 historical review, 1295
- Step-growth polymerization, poly(lactic acid),
 2250
- Stepped fatigue test, pyrolytic carbon, 2377
- Stereohindrance, chitin microspheres,
 607–608
- Stereolithography, rapid prototyping,
 2383–2388
 fused deposition modeling, 2387–2388
 selective laser sintering, 2384–2386
 three-dimensional printing, 2386–2387
- Sterility assurance level (SAL), 2502
- Sterilization
 biomedical materials
 current techniques, 2503–2507
 device design, 2508–2509
 dry heat, 2504
 ethylene oxide, 2504
 experimental techniques, 2507–2508
 glutaraldehyde, 2506–2507
 ionizing radiation, 2505–2506
 non-conventional agents, 2508
 overview, 2501
 plasmas, 2504–2505, 2507
 presterilization procedures, 2502–2503
 procedures, 2501–2052
 pulsed electric field technique, 2508
 saturated steam, 2504
 steam formaldehyde, 2506
 supercritical carbon dioxide, 2507–2508
 shape memory metal surfaces, 2424
 ultra-high molecular weight polyethylene,
 2929–2930
- Stickleback spiggen, adhesive properties, 243
- Stiffness properties
 composite materials, 662–663
 elasticity properties, 1239–1241
 spine biomechanics, 2466–2467
 suture materials, 2569–2570
- Stimulated Echo Acquisition Mode
 (STEAM), brain imaging, 1967–1969
- Stirred culture systems, bone marrow tissue
 engineering, 2672
- Stochastic learning, feed-forward neural
 networks, 123–124

- Stokes-Einstein equation, dialysis membrane diffusion, 1587
- Stokes flow, vascular fluid dynamics, 2980–2984
- Stoke's law, biomaterials, 836
- Stomach anatomy and function, electrogastrography imaging, 947–955
- Store-and-forward systems, telemedicine applications, 2581–2582
- Strain analysis
 articular cartilage biomechanics, 85–91
 blood vessel mechanics, 396–397
 bone remodeling, canalicular fluid flow, 442–443
 cardiac elastography, 507–512
 echocardiography, 507
 tendon/ligament tensile testing, 2591–2592
- Strain energy function, heart biomechanical modeling, 1287
- Strain Index, ergonomics guidelines, 1003
- Strain injuries, groin pain syndrome, 1393–1395
- Strength-duration curve, electrical stimulation waveform, 1014–1015
- Strength testing
 endovascular stent graft fixation, 2487
 ligaments, tissue engineering, 2763–2765
 spine biomechanics, 2466–2467
- Stress analysis
 articular cartilage biomechanics, 85–91
 atrioventricular valve mechanics, 1324–1326
 composite materials, 662–663
 corrosion cracking, implant materials, 756
 laser-tissue interactions, 1629
 periodontal ligament, biomechanics, 800–802
 valvular prosthetics, calcification, 77–78
- Stress cracking, polyurethane biodegradation, 2306–2307
- Stress fractures
 diagnostics, 2514–2516
 differential diagnosis, 2516
 epidemiology, 2511–2512
 femur, 2518–2519
 lumbar spine, 2519–2520
 metatarsals, 2517–2518
 pathophysiology, 2513–2514
 predisposing factors, 2512–2513
 properties of, 2511
 tibia, 2517
 treatment, 2516–2517
- Stress intensity factor, pyrolytic carbon, fracture mechanics, 2377
- Stress-shear-rate laws, vascular fluid dynamics, 2978
- Stress-strain analysis
 composites, continuum mechanics, 661–663
 fabric fibers, 1044–1045
 hard tissue viscoelasticity, 1244
 tendons, tissue engineering, 2872–2873
- Stripping process, silicone purification, 2437
- Stroke
 magnetic resonance imaging, 1723
 positron emission tomography imaging, 2319
- Stroke volume calculation, cardiac output measurement, 514, 527–528
- Stroma
 defined, 748
 tissue engineering
 bone marrow, 2670
 corneal implants, 2710
- Stromelysins, structure and function, 1769–1771
- Strouhal number, positive displacement pumps, 97–102
- Structural modeling
 heart biomechanics, 1287
 pyrolytic carbon, 2378–2380
- S-T segment, electrocardiography, 939
- Subcutaneous insulin delivery, 1514–1515
- Subepithelial matrix, tissue engineering, small intestine, 2866–2867
- Subluxation loading, total hip arthroplasty results, 1386–1388
- Submicron scale, brain structure, nanoscale neurosurgery, 1903–1904
- Submolecular mechanisms, protein-surface interactions, 280–281
- Substrate properties
 gradient surfaces, 1210–1211
 polyelectrolyte multilayer deposition, 2221
- Substratum topography, cell surface analysis, 2556–2557
- Sugars, mimetic materials, 297–298
- Sulfated derivatives, hyaluronan, 1425
- Supercritical antisolvent (SAS), 2324
- Supercritical fluids
 antisolvent, 2524
 basic properties, 2522–2524
 biomaterials applications, 2524–2529
 bacterial inactivation and enzyme denaturation, 2525
 bioactive substances extraction, 2525
 bone grafts, 2528–2529
 controlled-release drug delivery, 2525–2526
 implant devices, 2528
 molecular processing, 2524–2525
 nanoparticles and particle size reduction, 2526
 pharmaceuticals, 2525–2526
 protein stabilization, 2526
 tissue engineering, 2526–2528
- carbon dioxide, sterilization of medical devices, 2507–2508
- gas antisolvent, 2523–2524
 processing, overview, 2522
 rapid expansion of supercritical solutions, 2523
- Superficial femoral artery (SFA), distal protection device intervention, 878
- SuperFIXSORB30[®] composite, 466
- Super luminescence diodes (SLDs), optical coherence tomography, 1999
- Superoxides, biodegradable polymers, 203–204
- Superparamagnetism, magnetic nanoparticles, 1705–1706
- Suprarenal artery, endovascular stent graft fixation, 2488
- Surface analysis
 cell topography
 chemistry vs., 2552
 classification, 2551–2556
 discontinuous substratum, 2555
 fibers, 2553
 microfabrication, 2552
 nanotopography vs. microtopography, 2556
 overview, 2551
 pits/holes, 2553
 repeating grooved substrata, 2553
 roughened surfaces, 2555–2556
 single steps, 2552–2553
 substratum influences, 2556–2558
 tunnels/tubes, 2555
 in vivo applications, 2558–2560
- ceramic materials, transparent ceramics, 2918–2923
- coatings
 anti cell adhesion coatings, 2534
 antiencrustation coatings, 2533
 antimicrobial coatings, 2534–2536
 functionality, 2533
 future research, 2538
 maneuverability properties, 2531–2533
 medical devices, 2531–2538
- osteconductivity, 2536
 overview, 2531
 surface modification, 2537–2538
 surface passivation, 2536–2537
 thromboresistant coatings, 2534
- corrosion mechanics, metal implants, 758–759
- dental implants, 811–814
 hydroxyapatite coating, 814
 machined surfaces, 814
 microtextured titanium, 814
 plasma-spray coating, 814
- gradient surfaces
 bifunctional groups, 1213–1216
 hydroxy groups, 1213–1214
 ionic groups, 1214–1215
 phospholipid polymers, 1215–1216
 polyethylene oxides, 1215
 biomolecular interactions, 1219–1225
 cellular interactions, 1222–1225
 characterization, 1216–1219
 atomic force microscopy, 1218–1219
 ellipsometry, 1217
 infrared spectroscopy, 1218
 water contact angle, 1216–1217
 x-ray photoelectron spectroscopy, 1217–1218
- future research issues, 1225–1226
 overview, 1210–1211
 protein interactions, 1220–1222
 wettability surfaces, preparation, 1211–1213
 gas/solvent diffusion, 1211–1212
 radio-frequency plasma discharge, 1212–1213
- hemocompatibility
 assessment, 1349–1352
 inert materials, 1345–1350
- lung surfactants, 1685–1686
- medical device infection prevention, 1502–1503
- orthopedic device failure analysis, 1832–1837
- phosphorylcholine technology, 2139
- plasma modification of biomaterials
 corona and atmospheric discharge, 2175
 deposition and coating, 2181–2182
 discharge techniques, 2174–2175
 implantation, 2176–2177
 ion implantation, 2175
 metallic ion implantation, 2177–2180
 overview, 2174
 polymer ion implantation, 2180–2181
 radiofrequency discharge, 2174–2175
 spraying, 2183–2184
 sputtering and etching, 2176
 surface interactions, 2175–2176
 thin film deposition, 2175–2176
 vacuum arc, 2175
- protein adsorption modeling, 278–280
 biomaterial properties and surface analysis, 2326–2328
 biomolecular-material composites, 2326–2327
 molecular modeling, 2327
 physicochemical phenomena, 2327–2328
 overview, 2326
- simulation modeling, 2328–2333
 continuum modeling, 2328–2329
 forcefield-based molecular mechanics and dynamics, 2330–2333
 quantum mechanics, 2330
 solvent properties, 2329–2330
- silicone hydrogel contact lenses, 723–725
- stent devices, 2494–2495
- synthetic hydroxyapatite implants, 230–231

- Surface analysis (*cont.*)
 vascular grafts, host reactions, 3009
 zirconia ceramics, 3108
- Surface electromyography (SEMG)
 data acquisition and signal conditioning, 963–964
 data quantification, 964–965
 electrode placement, 961–962
 signaling mechanisms, 959
- Surface modification
 articular cartilage repair, 557
 bio-artificial liver systems, optimization, 1653
 biomolecule attachment/biomimetic surfaces, 2545–2546
 biorubber/poly(glycerol sebacate) degradation, 330–331
 cell-material interactions, 569–570
 coating surfaces, 2537–2538
 gradient surfaces, 1216–1217
 hemocompatibility, inert materials, 1346
 methods, 2450
 plasma surface techniques, 2543–2545
 ion implantation methods, 2543
 polymerization, 2543–2545
 polymer grafting, 2545
 polyurethane biomaterials, 2311
 radiation-based methods, 2450–2543
 ion beam techniques, 2540–2541
 laser, ultraviolet, and gamma radiation, 2541–2543
 self-assembled monolayers, 2545
 organosilicon derivatives, hydroxylated surfaces, 2411–2414
 shape memory metals
 basic properties, 2418–2419
 biomedical applications, 2419–2420
 overview, 2418
 surface modifications, 2420–2422
 ceramic coatings, 2423
 chemical and electrochemical treatments, 2422–2423
 hydroxyapatite coatings, 2423
 oxidation, 2422
 polishing, ultrasonic bath, and physical etching, 2421–2422
 polymer coatings, 2423–2424
 sterilization, 2424
 silicones, 2434–2435
 tissue engineering
 microscale technology, 2780
 vascular system, 3035
 titanium and titanium alloys, 2885–2887
 vascular grafts, 2990–2991, 2997–2998
 wet chemistry, 2545–2546
- Surface plasmon resonance, biosensor development, 344–345
- Surfactants
 biomedical and bioengineering applications, 2338
 biomimetic materials, micelles, 295–296
 industrial formulations and applications, 2338
- lungs
 clinical exogenous surfactant therapy, 1690–1691
 diseases of, 1689–1690
 drug inhibition resistance and activity, 1691
 endogenous surfactant activity, 1683–1684
 endogenous surfactant deficiency and dysfunction, 1687
 future research issues, 1691–1692
 inhibitor-induced dysfunction, 1687–6
 overview, 1683
 physiological correlates of dysfunction, 1688–1689
 specific component properties, 1686–1687
 surface activity measurement, 1685–1686
- protein and cell signaling
 biomaterials applications, 2340–2343
 cationic materials, 2342–2343
 future research issues, 2344
 hydrophilicity/hydrophobicity, 2341–2342
 intracellular/extracellular mechanisms, 2339–2340
 natural materials, 2337–2338
 overview, 2337
 pH and LDH activity, 2341–2343
 semipermeable membrane transport, 2343
- valvular prosthetics, anticalcification treatments, 80
- Surgical techniques
 computer-assisted surgery
 cardiac valve surgery, 691
 da Vinci system, 689–690
 esophageal surgery, 691
 evolution of, 687–688
 future applications, 692–693
 overview, 687
 pelvic surgery, 691–692
 Zeus system, 688–689
 fibrin sealants, 1080–1088
 glue materials, 1185–1193
 hyaluronan, 1429
 intravascular ultrasound, 1560–1564
 neurosurgery, functional magnetic resonance imaging, 1123–1124
 peripheral nerve injury and repair, 2123–2124
 robotics technology, 1807–1812
 future research, 1812
 human-machine interface, 1807–1809
 surgical robots, 1809–1811
 training simulators and haptics, 1811
- sutures
 absorption properties, 2575–2577
 biocompatibility, 2574–2575
 biodegradability, 2575–2577
 commercial size and configuration, 2563–2568
 handling properties, 2570–2574
 materials properties and classification, 2568–2577
 needles, 2562–2563
 overview, 2562
 physical and mechanical properties, 2569–2570
 size classifications, 2568
- Susceptibility difference, magnetic cell separation, viscous medium, 1698–1699
- Sustained low-efficiency dialysis (SLED), acute renal failure, 1591–1592
- Suture materials
 absorption properties, 2575–2577
 biocompatibility, 2574–2575
 biodegradability, 2575–2577
 biodegradable polymers, 197–199
 commercial size and configuration, 2563–2568
 glue materials as, 1185–1193
 handling properties, 2570–2574
 materials properties and classification, 2568–2577
 needles, 2562–2563
 overview, 2562
 physical and mechanical properties, 2569–2570
 poly(glycolic acid), 2244–2245
 poly(lactic acid), 2254–2255
 poly(lactide-co-glycolide) biocompatibility, 2266–2267
 silk materials, *Bombyx mori* silk, 2445
 size classifications, 2568
 soft tissue-bone interface, 2076–2077
- Swan-Ganz catheter, basic properties, 504
- Swelling properties
 glucose-sensitive hydrogels, 1163–1172
 interpenetrating polymer networks, 1550–1551
 phosphorylcholine technology, 2139
- Symmetry, polymer structure/property relationships, 2277–2278
- Synchromed EL drug infusion system, 740
- Synergy sites, integrin regulation, 1528
- Synthetically modified cellulose (SMC)
 dialysis membranes, 1584–1585
- Synthetic materials
 apatites, 229–231
 cell-material interactions, 571
 molecular networks, gene therapy, 1153–1155
 polymers
 biodegradable compounds, 201–204
 tissue engineering, 2285–2289
 polyurethane biomaterials, 2305–2306
 tissue engineering, esophagus, 2720–2721
- Systemic circulation, vascular fluid dynamics, 2972–2981
- Systemic effects, in-vivo biocompatibility testing, 174–175
- T**
- Tacks, soft tissue-bone interface, 2077
- Tanning industry, collagen processing, 652–656
- Tantalum, orthopedic devices, 2063
- Target cells, magnetic cell separation, 1695–1696
- Tau scaling parameter, ceramics fracture/fatigue analysis, 587
- Taylor series, ultrasound-tissue interaction, 2939–2941
- T cells
 host reactions and, 1416–1418
 xenotransplantation, cell-mediated rejection, 2910
- TeCP-DCPA cement, 465–466
- Teeth. *See also* Dental implants
 biological properties, 594–595
 ceramic interaction with, 595
 dentin composition and formation, 2047–2048
 dentin tissue engineering, 2048
 enamel composition and formation, 2048
 enamel tissue engineering, 2049
 mobility, dental implants *vs.*, 817
 physical properties, 594
 structure and aesthetic properties, 593–594
- Teflon
 cardiac catheters, 496
 polytetrafluoroethylene copolymer, 2302
- Tegaderm, burn dressings, 458
- Tele-echography, 1812
- Telemedicine, 2582–2583
 clinical applications, 2579–2780
 hardware requirements, 2583–2584
 hub-and-spoke topology, 2582–2583
 medical informatics, 2584
 migration and standards adherence, 2584–2585
 overview, 2579
 peripheral equipment, 2583–2584
 portable devices, 2584
 real time systems, 2580–2581
 robotics technology, 1809
 security issues, 2585–2586
 store-and-forward systems, 2581–2582
 system architecture, 2584
 training and distance learning applications, 2585
- Telemetry systems, enzyme-based electrochemical sensors, 1180

- Teleoperation systems, surgical robotics, 1809–1811
- Temperature measurement
interpenetrating polymer networks, 1551–1554
magnetic resonance imaging
clinical applications, 1730–1731
future research issues, 1734
molecular diffusion coefficient, 1728–1729
overview, 1728
proton resonance frequency shift, 1731–1734
spin-lattice relaxation time, 1729–1730
thermometry methods, 1734–1735
melt spinning, polymers, 1816–1817
- Templating
bone tissue engineering, 2045–2047
microporous materials, sol-gel bioactive glasses, 1888
multilayer assemblies, 2217–2226
- Temporal moments, mass transfer and tissue function, 1754–1755
- Tendinitis/tendinosis
iliopsoas overuse injury, 1395–1396
patellar, 1599–1600
pes anserinus tendinitis/bursitis, 1604–1605
popliteal tendinitis, 1605
rectus femoris overuse injury, 1396–1397
semimembranosus tendinitis, 1604
- Tendons
fibroblast mechano-regulation, 1788
healing potential, 1253–1254
multiaxial mechanical testing
anisotropic properties, 2592–2593
bidirectional testing, 2593–2594
shear loading, 2594–2595
structural framework, 2588
tissue engineering
anatomy and physiology, 2871–2872
cell sources, 2873–2874
clinical indications and requirements, 2872–2873
future research issues, 2874
growth factor/gene therapy, 2874
injury mechanisms, 2871
scaffold materials, 2873
uniaxial tensile testing, 2589–2592
viscoelasticity, 2595–2597
in vitro analysis, 2591–2592
- Tensile properties
periodontal ligament, biomechanics, 800–802
suture materials, 2569
tendons and ligaments, 2589–2592
- Tensiometry, bioadhesion measurements, 5
- Tension band wiring, bone fixation, 414
- Tensorial stiffness, composite materials, 662–663
- Terrestrial adhesives, structure and classification, 245–247
- Testicular prostheses, basic properties, 2968–2969
- Tetraethylene glycol dimethacrylate (TEGDMA), self-regulated insulin delivery, 1166–1170
- Tetrapolar impedance cardiography, cardiac output measurement, 526–530
- Therapeutic cell mass, bio-artificial liver systems, 1651
- Therapeutic monitoring, magnetic resonance imaging, 1724–1725
- Thermal-dilution method, cardiac output measurement, 522–526
diffusible indicators, 526
verification studies, 524–525
- Thermal effects
laser tissue welding, 2032–2033
poly(lactide-co-glycolide), 2262
radiofrequency electromagnetic field/
radiation exposure, 256–257
zirconia ceramics, 3104
- Thermally induced phase separation (TIPS)
bioactive composite scaffolds, 149
polyurethane biomaterials, 2309–2310
- Thermodynamics
hemocompatibility assessment, 1350
implantable batteries, 1486
interfacial transport, 2349–2350
polymer structure/property relationships, 2275–2276
protein-based drug delivery systems, 2358–2359
- Thermoelastic expansion, laser-tissue interactions, 1629
- Thermofiltration, plasmapheresis, 387–391
- Thermoplastics, urinary prostheses, 2962
- Thermosets, silicone elastomers, 2430–2433
condensation cure, 2431–2433
fillers, 2433
hydrosilation cure, 2430
peroxide cure, 2430–2431
- Thin film deposition. *See also* Biofilms
plasma biomaterials modification, 2175–2176
carbon-based films, 2181
self-assembled monolayers
basic properties, 2409
biotechnological applications, 2412–2415
formation mechanisms, 2409–2412
metal-alkanethiol interactions, 2409–2411
- Thoracic aorta, endovascular stent grafts, 2481–2482
- Thoracic surgery, fibrin sealant applications, 1087
- Thoracoabdominal aorta, endovascular stent grafts, 2482
- Three-dimensional assembly, tissue engineering
bone marrow, 2672–2673
microvascular networks, 2797
- Three-dimensional imaging
biomedical optics, 2029–2035
confocal microscopy, 708–709
Doppler ultrasound, 2942–2943
arterial disease, 2948
cardiac examination, 2948–2949
carotid artery examination, 2947
color imaging, 2945–2946
continuous wave Doppler, 2943–2944
duplex scanning, 2944–2945
power Doppler imaging, 2946–2947
pulsed wave Doppler, 2944
tissue imaging, 2947
transcranial examination, 2947
electromyographic signals, 965
fluorescence spectroscopy, 2035–2036
intravascular ultrasound
clinical settings and treatments, 1559–1560, 1562
future research issues, 1562–1564
imaging studies and modalities, 1560–1562
overview, 1559
magnetic resonance imaging, 1794
basic principles, 1713
biomedical applications, 1725–1726
clinical applications, 1720–1725
contrast sources, 1718–1720
current density, 1719
gradients and imaging, 1716
historical background, 1713–1714
image weighting, 1718
industrial/nonmedical applications, 1726–1727
motion-induced contrast, 1718–1719
physical principles, 1714–1715
polarization transfer, 1720
proton density, 1718
radiofrequency excitation and signal detection, 1715
rapid imaging, 1716–1717
relaxation, 1715–1716
relaxation times, 1718
signals and magnetization, 1714–1715
spectral information, 1719–1720
technology, 1717–1718
temperature measurement
clinical applications, 1730–1731, 1735
future research issues, 1734
molecular diffusion coefficient, 1728–1729
overview, 1728
proton resonance frequency shift, 1731–1734
spin-lattice relaxation time, 1729–1730
thermometry methods, 1734–1735
medical image processing and analysis, 1797
microcomputed tomography, 1854
enteric coated tablets, 1856–1857
image reconstruction, 1853
nuclear magnetic resonance spectroscopy
biological studies, 1966–1967
brain imaging, 1967–1970
proton spectra information, 1969–1970
in vivo spatial localization, 1967–1969
water and lipid suppression, 1969
chemical shift, 1962
historical background, 1960
pulsed Fourier transform spectroscopy, 1963–1964
rotating frame experiments, 1964–1966
relaxation times, 1964–1965
signal-to-noise ratio, 1961
spectrum characteristics, 1962–1963
spin echoes, 1965–1966
spin-spin coupling, 1962–1963
theoretical principles, 1960–1961
nuclear tomographic imaging, 1795–1796
optical coherence tomography
applications, 2003–2004
future research issues, 2004–2005
imaging probes, 2001–2002
imaging resolution, 1998–2000
imaging speed, 2000–2001
operating principles, 1996–1998
overview, 1996
positron emission tomography, 1796
attenuation, 2318
biological applications, 2318–2321
cancer detection, 2320
data acquisition, 2317–2318
detector design, 2315–2316
drug development and, 2321
emission and annihilation technology, 2314
gantry assembly, 2316
hemodynamic imaging, 2318–2320
image reconstruction, 2318
light detection, 2314–2315
metabolic imaging, 2320–2321
molecular mechanisms, 2321
overview, 2314
quantification, 2318
receptor function imaging, 2321
resolution parameters, 2316–2317
sampling procedures, 2317
scatter, 2318
sensitivity parameters, 2317
statistical noise, 2317–2318
real-time tomographic reflection
augmented reality and head-mounted displays, 2391
calibration
affine transform, 2393
geometric transform, 2392–2393

- Three-dimensional imaging (*cont.*)
 current and future applications, 2393–2394
 image overlay, 2391–2392
 overview, 2391
 techniques and procedures, 2392–2393
 registration, 1798–1799
 segmentation, 1797–1798
 single photon emission computed tomography, 1795–1796
 telemedicine applications, 2581–2582
 ultrasound, tissue interaction
 acoustic propagation principles, 2933–2934
 attenuation, absorption, and scattering, 2936–2939
 characteristic impedance, 2934
 intensity, 2934–2935
 linearity effects, 2939–2941
 radiation force, 2935
 reflection/refraction, 2935–2936
 vascular system
 angiography, 3012–3013
 angioscopy, 3017
 computed tomography angiography, 3013–3014
 future research issues, 3018–3020
 magnetic resonance angiography, 3015–3017
 overview, 3012
 processing protocols, 3017–3018
 ultrasound, 3014–3015
 visualization, 1800–1801
 x-ray computed tomography, 1794–1795
- Three-dimensional modeling
 microelectromechanical systems, 1863
 toxicity testing, animal surrogate systems, 55–56
- Three-dimensional printing (3DP) techniques
 biocomposite fabrication, 190–191
 rapid prototyping, 2386–2387
- Three-Dimensional Static Strength Prediction Program (3DSSPP), ergonomics guidelines, 1004
- Thrombin inhibitors
 hemocompatibility, 1346–1348
 heparin biocompatibility, 1370
- Thrombin-thrombomodulin pathway, blood-material interactions, 370
 mechanical heart valves, 1334–1335
- Thrombogenicity
 cardiac catheterization, 497
 in vivo animal models, 1568–1569
- Thromboresistant coatings
 stent devices, 2495
 surface properties, 2534
- Thrombosis. *See also* Deep vein thrombosis (DVT)
 angiogenesis inhibitors and, 31
 basic principles, 2601
 blood-material interactions, 2604–2608
 erythrocyte adhesion, 2607–2608
 fibrin formation, 2604–2606
 platelet-dependent neutrophil enhancement, 2606–2607
 cardiopulmonary bypass, 1668–1669
 centrifugal blood pump testing, 579–583
 circulatory support systems, 1778–1779
 coagulation biology, 2601–2604
 complement system/activation, 2604
 neutrophil function, 2602–2604
 platelet function, 2601–2602
 reaction cascade, 2601
 dermatan sulfate pharmacology, 625–626
 stent complications, 2493
- Tibia
 prosthetic devices
 total knee replacement, 1609–1610
 tray prosthesis, failure analysis, 1841
 stress fracture, 2517
- Tick salivary cement, 245–246
- Tilting disc mechanical heart valve, structure and components, 1329–1330
- Time constants, mass transfer and tissue function, concentration diffusion, 1757–1759
- Time-domain analysis, light propagation in tissue, 2033–2034
- Time-of-flight (ToF) microscopy
 flow cytometry, 1103
 vascular imaging, 3016
- Time-of-flight secondary ion mass spectrometry (ToF-SIMS)
 gradient surfaces characterization, 1219
 protein-surface interactions, 280–281
- Time-resolved measurements
 cancer diagnosis and therapy
 fluorescence imaging, 2015
 radiative transfer principles, 2012–2013
 optical detection of cancer, spectroscopic techniques, 2012–2013
- Tinnitus
 acoustic masking
 designs, 2612–2613
 FM masking, 2613
 high audio/ultrasonic bone vibration, 2614–2615
 history, 2611–2612
 personal stereos, 2613–2614
 vibration masking, 2614
 neural plasticity and cortical mapping, 2617
 pathology, 2610
 transdermal electrical stimulation, 2612
 ultrasonic hearing, 2615–2616
- Tissue colonization, host reactions, 1414–1415
- Tissue-culture polystyrene (TCPS), cell-material interactions, 569–570
- Tissue culture polystyrene (TCPS) dishes, transparent ceramics, 2916–2923
- Tissue damage, electrical stimulation techniques, 1016–1017
- Tissue donation, allografts, 14–15
- Tissue Doppler imaging (TDI)
 basic principles, 2947
 clinical applications, 506–507
- Tissue engineering. *See also* Scaffold materials
 alternating current electrokinetics
 basic principles, 2628–2629
 biomaterial cell trapping, 2633–2634
 cell properties characterization, 2631–2633
 cell separation, 2632
 cell sorting, 2632–2633
 dielectrophoresis, 2629–2630
 electro-orientation, 2629
 electrorotation, 2630
 electrostriction, 2629
 traveling wave dielectrophoresis, 2630–2631
- articular cartilage
 bioengineering principles, 2701
 biphasic construct substrates, 2700–2701
 cell cultures, 2697–2698
 cell signaling mechanisms, 2697
 cell sources, 2701–2702
 chondrocyte subpopulations, 2703
 composition, 2694–2695
 differentiated chondrocytes, 2701–2702
 function, 2696–2697
 growth factors, 2702
 mechanical stimulation, 2702
 overview, 2694
 pre/post-transplantation assessment, 2701
 quality control and improvement, 2702–2703
 repair, 554–555
 scaffold materials, 2698–2700
- stem cells, 2702
 zonal organization, 2695–2696
- bioactive materials
 calcium phosphates, 143
 composites, 144–145, 149
 development, 2623
 fabrication, 145–150
 foam coatings, 149–150
 glass-ceramics, 138–139, 144–150, 1822–1823
 microsphere sintering, 149
 overview, 142
 powder-based fabrication, 147
 scaffold structures, 142
 silicate glasses, 143–144
 sol-gel fabrication, 147–148
 solid free-form fabrication, 148–149
 solvent casting/particle leaching, 149
 thermally induced phase separation, 149
- biodegradable elastomers, 897–906
- bioreactors, 310–311, 2623
- bladder
 cell transplants, 2636–2640
 future research issues, 2640–2641
 regeneration matrices, 2636
- blood vessels
 cell sources, 2645–2646
 cellular-based scaffolds, 2647
 clinical indications, 2643–2644
 collagen-based scaffolds, 2646–2647
 construct design, 2646–2647
 design criteria, 2644–2645
 future research issues, 2650
 limitations, 2648–2650
 mechanical conditioning, 2648
 mimicry properties, 2645
 scaffold materials, 2647–2648
 vascular anatomy, 2643
 vascularized engineered constructs, 3024–3028
- bone, 2044–2047
 animal models, 33–45
 bone-implant interface, 423–430
 cell properties, 2653–2654
 future research issues, 2657
 healing and regeneration, 1260–1261
 osteoclast regeneration, 2095–2096
 overview, 2652–2653
 scaffold materials, 2654–2655
 signaling molecules, 2656–2657
- bone marrow
 cell culture parameters, 2662–2665
 culture media, 2662
 culture systems
 perfusion cultures, 2671–2672
 static cultures, 2670–2671
 stirred suspension cultures, 2672
 stromal support, 2670
 three-dimensional cultures, 2672–2675
- feeding schedule and culture duration, 2664
 growth factors, 2662–2663
 hematopoiesis, 2660–2662
 inoculum density, 2663–2664
 oxygen tension, 2664–2665
 pH levels, 2664
- breast
 adipose tissue cellular components, 2679–2680, 2682
 anatomy, 2679
 animal models, 2683
 autologous transplants, 2679
 cell sources, 2680–2681
 clinical indications, 2678
 current treatment and reconstruction options, 2678–2679
 drug delivery systems, 2683–2684
 future research issues, 2683–2684

- injectable composite systems, 2682–2683
- prosthetics, 2678–2679
- scaffold materials, 2681
- three-dimensional in vitro systems, 2684
- in vitro procedures, 2681–2682
- burn injuries, skin substitutes, biological and synthetic, 451–456
- calcium phosphate ceramics, 467–468
- cardiac muscle
 - cell properties, 2689–2690
 - cell seeding, 2686–2687
 - clinical applications, 2691–2692
 - functional requirements, 2690–2691
 - liquid cell matrices, 2687
 - methods, 2686–2688
 - monolayer stacking, 2687–2688
 - scaffold materials, 2688–2689
 - solid matrices, 2686–2687
 - in vivo applications, 2691
- cardiac patch
 - biomaterial substrates, 546
 - bioreactors, 544
 - cell injection, 543
 - collagen-myocyte interaction, 545
 - construct tissue engineering, 543–544
 - extracellular matrix arrangement, 545
 - future research issues, 550
 - historical background, 543–546
 - integrins, 545
 - mechanical forces, 545–546
 - myocyte phenotype regulation, 545
 - overview, 542–543
 - rotating-wall bioreactor, 546–549
 - techniques for, 544–545
- cardiovascular system, microscale technology, 2785–2787
- cell sources, 2283
- chitosan, 611–613
- controlled release drug delivery systems, 743–744
- cornea
 - cell sources, 2707–2710
 - clinical applications, 2715–2716
 - collagen gels, 2710–2711
 - collagen sponges, 2711
 - endothelial cells, 2710
 - epithelial cells, 2708–2710
 - experimental applications, 2715
 - extracellular matrix expression, 2714
 - fibrin gel cell cultures, 2710
 - future challenges, 2714–2716
 - integrin expression and wound healing, 2713–2714
 - keratocytes, 2710
 - reconstructive technology, 2707–2708
 - self-assembly techniques, 2711–2712
 - wound healing, 2712–2714
- critical integrative technology, 2622–2623
- current approaches, 2620–2622
- dielectric properties, 834–838
 - conductivity and permittivity, 833–834
 - counterion polarization, 836–837
 - data sources, 841
 - dipolar relaxation, 836
 - dispersion regions, 839–840
 - frequency ranges, 840–841
 - interfacial polarization, 837–839
 - microwave frequencies, 841
 - overview and history, 832–833
- elastin
 - basic properties, 891
 - bioengineering applications, 894–895
 - gene structure, 891
 - protein structure, 891–893
 - tropoelastin assembly and ultrastructure, 893–894
- electroactive polymeric materials, 926–933
- electron microscopy
 - applications, 974–975
- SEM specimen preparation, 972–973
- TEM specimen preparation, 973–974
- electrospun polymer nanofibers, 982–983
- esophagus
 - animal models, 2720
 - cell sheets, 2721
 - clinical applications, 2719–2720
 - collagen type 1 scaffolds, 2721
 - future research, 2722
 - naturally occurring scaffolds, 2721
 - synthetic scaffolds, 2720–2721
- extracellular matrix
 - component properties, 1022–1024
 - host response, 1024
 - overview, 1021
 - sources, 1024
 - tissue repair, 1024–1025
- fabric structure and classification
 - braided three-dimensional fabrics, 1052–1053
 - electrospun yarns and fabrics, 1053–1054
 - fibers, 1046–1048
 - implantation properties, 1043–1044
 - multiaxial warp, 1051–1052
 - orthogonal three-dimensional fabrics, 1050–1051
 - performance characteristics, 1055–1056
 - scaffold materials, tissue engineering, 1056–1059
 - cell proliferation, 1058
 - cell-scaffold architecture interaction, 1058–1059
 - structural components, 1044–1054
 - textile units, 1044–1046
 - three-dimensional fabrics, 1049–1053
 - two-dimensional structural geometry, 1048–1049
 - woven fabrics, 1048–1050
- fat tissue
 - cell sources, 2727–2728
 - clinical applications, 2725
 - de novo adipogenesis, 2732
 - growth factors and adipogenic inducers, 2730
 - origin and characteristics, 2726
 - research applications, 2733–2734
 - scaffold materials, 2729–2730
 - strategies, 2726–2727
 - surgical techniques and limitations, 2725–2726
 - synthetic hollow support structures, 2732–2733
 - vascularization, 2730–2731
 - in vitro studies, 2731
 - in vivo studies, 2731–2732
- fibrin structure and biomechanics, 1075–1077
- fibroblast mechano-regulation, gene expression, 1783–1788
- finite element analysis, 1091–1098
- foreign body response, 1111–1116
 - fibroblasts, 1114
 - giant cell response, 1114
 - mechanisms of, 1115–1116
 - phagocytosis, 1112–1114
- future research issues, 2624–2625
- gene-activated matrix
 - applications, 1142–1143
 - basic properties, 1142
 - biodegradable scaffolds, 1145–1146
 - tissue regeneration, 1146–1148
 - vector systems, 1144–1145
 - wound healing, 1143
- gene therapy, 2623
- histiogenesis
 - basic principles, 1403
 - embryonic, 1405–1406
 - engineered procedures, 1407–1408
 - hierarchy, 1403–1404
- postnatal, 1406–1407
- terminology, 1403
- tissue complexity and variability, 1404–1405
- historical perspective, 2620
- host reactions, 1414–1415
- hyaluronan
 - cardiovascular applications, 1428
 - general applications, 1429
- insulin delivery systems, 1518–1519
- interfacial transport and, 2348–2349
- intestine
 - cell sources, 2753–2754
 - feasibility, 2752–2753
 - overview, 2751–2752
 - scaffold materials, 2754
- laser-tissue interaction, 1623–1624
 - biological effects, 1630
 - light-tissue interactions, 1625–1626
 - overview, 1623
 - photochemical interactions, 1626
 - photomechanical interaction, 1628–1630
 - photothermal interactions, 1626–1627
 - tissue optics, 1623–1625
- ligaments
 - ankle injuries, 2758–2759
 - anterior cruciate ligament injuries and reconstructions, 2758
 - bone-tissue integration, 2763
 - cell structure, 2761
 - composition and structure, 2759
 - design issues, 2762–2768
 - immunological acceptability, 2765–2766
 - inflammation, avoidance, 2765
 - joint guidance, 2761–2762
 - ligament-bone insertions, 2761
 - loading mechanisms, 2762
 - mechanical properties, 2759–2760
 - proprioception, 2762
 - scaffold materials, 2759
 - shoulder injuries, 2759
 - strength maintenance over time, 2763–2765
 - vascularization and nutrient transport, 2765
- light propagation, 2033–2035
- liver
 - bio-artificial liver systems
 - bioreactors, 1656–1657
 - cell-cell interactions, 1654–1655
 - cell source and therapeutic cell mass, 1651–1652
 - clinical setting, 1649–1651
 - extracellular matrix and cell polarity, 1653–1654
 - future research issues, 1658
 - hepatocyte transplantation, 1650
 - hollow-fiber systems, 1656–1657
 - overview, 1649
 - parallel plate systems, 1657
 - performance monitoring, 1657–1658
 - plasma/blood effects, 1656
 - porous three-dimensional matrices, 1655
 - scale-up hepatocyte culture, 1656
 - soluble factors, 1655–1656
 - surface chemistry optimization, 1653
 - temporary support system, 1651
 - tissue engineering, 1652–1656
 - xeno (cross-species transplantation), 1650
 - bioreactors, 2772–2773
 - coatings, 2771–2772
 - future research issues, 2775–2776
 - hepatic environment, 2769–2771
 - immune response, 2772
 - liver-assist devices, 2773–2774
 - microscale technology, 2784–2785
 - overview, 2769

- Tissue engineering (*cont.*)
- scaffold materials, 2771
 - transplants, 2773–2775
 - mathematical modeling, cardiovascular system, 1898–1899
 - matrix metalloproteinases, 1771–1772
 - matrix properties and definitions, 2039
 - mechanical stimulation, 2623
 - melt-spun fibers, 1822–1823
 - microelectromechanical systems, 1860–11
 - microporous materials, 1885–1892
 - microscale
 - cardiovascular system, 2785–2787
 - chemical control, 2782
 - control systems, 2781–2782
 - future research issues, 2788–2789
 - immunological control, 2782–2783
 - liver, 2784–2785
 - micromachining technology, 2780
 - multidimensional techniques, 2789
 - nerve tissue, 2787–2788
 - overview and historical background, 2779–2780
 - pancreatic tissue, 2783–2784
 - skin, 2788
 - soft lithography, 2780–2781
 - spatial control, 2782
 - surface engineering, 2780
 - topographical control, 2781–2783
 - in vitro/in vivo procedures, 2789
 - microvascular networks
 - assembly techniques, 2796–2797
 - biocompatible polymer molding, 2796
 - biodegradable polymer fabrication, 2796
 - endothelial cell seeding, 2798
 - fluid dynamics, 2793–2794
 - lithographic techniques, 2794–2795
 - organ and tissue properties, 2792–2793
 - plasma etching, 2795–2796
 - polymer replica molding, 2796
 - porous devices, 2797–2798
 - scaffold fabrication, 2794–2798
 - vascular fluid dynamics, 2793–2794
 - vascularization requirements, 2792
 - whole organ engineering, 2798–2799
 - morphometrics, 2623
 - multidisciplinary approach to, 2623–2624
 - nerve guide technology, 1911–1921, 2814–2815
 - nerve tissue
 - acellular grafts, 2815
 - axonal degeneration, 2812–2813
 - biomimetic strategies, 2816
 - functional recovery assessment, 2813–2814
 - gene therapy, 2817
 - microscale technology, 2787–2788
 - nerve guidance channels, 2814–2815
 - nerve regeneration strategies, 2814–2816
 - nerve tissue degeneration, 2812
 - nervous system anatomy, 2811–2812
 - overview, 2811
 - stem cells, 2816
 - support cells and neurotrophic factors, 2815–2816
 - neuroprostheses, 1924–1931
 - nitric oxide materials, 1955–1956
 - osteoinductive substances and materials, 2105–2106
 - overview, 2620
 - pancreas
 - artificial pancreas, 2802–2806
 - cell sources, 2806–2608
 - diabetes, 2802
 - endocrine disease, 2807
 - microscale technology, 2783–2784
 - peripheral nerves, 2128
 - photopolymerized biomaterials, 2158–2160
 - polycaprolactone, 2202–2203
 - poly(glycolic acid), 2244–2245
 - in vivo response, 2246–2247
 - poly(lactic acid), 2254–2255
 - polymers
 - biomaterials properties, 2284–2285
 - cell carrier properties, 2284
 - degradable, nondegradable, and stable materials, 2284–2285
 - engineering technology, 2282–3
 - extracellular matrix and scaffold properties, 2283–2284
 - foams, 2271–2274
 - future trends and research issues, 2295
 - intelligent polymers, 2294–2295
 - natural polymers, 2289–2293
 - natural/synthetic compounds, 2293–2294
 - overview, 2282
 - sources, 2285–2295
 - synthetic polymers, 2285–2287
 - polyurethane biomaterials, 2308–2312
 - rotary cell culture systems
 - cancer, 2405–2406
 - cartilage, 2405
 - future applications, 2406–2407
 - infectious disease, 2406
 - insect cells, 2406
 - liver, 2406
 - neural tissue, 2406
 - overview, 2402
 - pancreatic islets, 2406
 - theoretical principles, 2402–2403
 - tissue engineering applications, 2403–2410
 - scaffold materials
 - basic principles, 2828–2829
 - fabrication techniques, 2829–2835
 - acellular matrix and submucosa, 2832–2833
 - connective tissue capsule, 2833
 - electrospinning, 2834–2835
 - extrusion, 2831
 - fiber bonding, 2829
 - freeze drying, 2831–2832
 - gas foaming, 2831
 - melt molding, 2830–2831
 - membrane lamination, 2830
 - phase separation, 2832
 - rapid prototyping, 2833–2834
 - solvent casting and particulate leaching, 2829–2830
 - shell components, 2049
 - silk scaffolds for, 2446
 - skeletal muscle
 - electrospinning strategies, 2843–2844
 - future directions, 2845
 - growth and muscle adaptation, 2842–2843
 - historical perspective, 2839–2840
 - limitations, 2839
 - prosthetics fabrication, 2840
 - scaffold materials, 2841–2842, 2845–2848
 - structure and function, 2838–2839
 - in vivo techniques, 2840–2841, 2844–2845
 - skin
 - current treatments, 2852–2853
 - dermis regeneration template, 2854–2856
 - epidermal autografts, 2853
 - equivalents and regeneration templates, 2853–2857
 - future research, 2856–2857
 - injury response, 2852
 - living dermal replacement, 2854
 - microscale technology, 2788
 - structure and function, 2851–2852
 - wound healing, 2851–2852
 - small-angle light scattering (SALS), soft tissue analysis, 2459–2461
 - small intestine
 - clinical applications, 2860–2861
 - epithelial cell culture, 2861–2862
 - evolution of, 2860
 - organoid unit models, 2864–2866
 - recent developments, 2863–2864
 - scaffold materials, 2862–2863
 - stem cells, 2758–2759, 2867–2868
 - subepithelial matrix extension, 2866–2867
 - supercritical fluids, basic principles, 2526–2528
 - teeth
 - dentin, 2048
 - enamel, 2049
 - tendons
 - anatomy and physiology, 2871–2872
 - cell sources, 2873–2874
 - clinical indications and requirements, 2872–2873
 - future research issues, 2874
 - growth factor/gene therapy, 2874
 - injury mechanisms, 2871
 - scaffold materials, 2873
 - transparent ceramics
 - alkaline phosphatase activity staining, 2918
 - calcium staining, 2918
 - cell culture, 2917–2918
 - mineralized matrix, 2918
 - osteogenic differentiation, 2919–2920
 - surface characterization, 2917–2919
 - tissue engineering applications, 2916–2923
 - valvular prosthetics
 - aortic valve cusp microstructure and micromechanics, 2738–2742
 - biomimetic constructs, 2743–2747
 - bioprosthetic valve failure mechanisms, 2737–2738
 - calcification and mechanical damage, 2738
 - cellular population, 2742
 - clinical indications for, 1292–1294
 - collagen, 2739–2740, 2744
 - cross-linked hyaluronan, 2744–2746
 - degeneration mechanisms, 2738
 - elastin, 2740–2742, 2745–2747
 - failure analysis
 - calcium binding site inactivation, 1308–1309
 - design failures, 1312–1313
 - future research issues, 1313–1314
 - glutaraldehyde cross-link chemistry, 1306–1308
 - overview, 1295–1301, 1306
 - remnant tissue immunogenicity, 1309–1312
 - historical review, 1294–1295
 - matrix mechanics, 2742–2743
 - mechanical heart valves
 - blood-material interactions, 1334–1336
 - cavitation, 1331
 - classification, 1329
 - durability, 1329–1330
 - fluid mechanics applications, 1331–1332
 - overview, 1329
 - in vitro flow field measurements, 1332–1334
 - wear and fatigue failure, 1330–1331
 - natural adhesives, 247–249
 - overview, 1292, 2737
 - poly(glycerol sebacate), 333
 - scaffold materials
 - biocomposites, 180–181
 - collagen scaffolds, 641–642
 - basic properties, 639–640

- smart tissue-activated fiberoptic biosensors, 1066–1067
- soft connective tissue, 1255
- synthetic apatite composites, 232–233
- vascular grafts, 2992–2993
 - bioreactors and in vitro procedures, 3039–3040
- cell sources, 3036–3039
 - clinical applications, 3033, vasctiss–9
 - development strategies, 3001
 - engineered constructs, 3022–3028
 - future research issues, 3042
- hybrid materials
 - biological sealants, 1432–1433
 - endothelial seeding, 1433–1434
 - mandrel-guided FCTs, 1434–5
 - overview, 1432
 - tissue engineered substitutes, 1436–1437
- scaffold materials, 3033–3035
- seeding techniques, 3035–3036
- in vivo procedures
 - animal models, 1567–1569
 - remodeling, 3040–3041
- vitrification, 3049–3054
- xenografts, 3099–3100
- Tissue factors, mechanical heart valves, 1334–1335
- Tissue heating, radiofrequency
 - electromagnetic field/radiation exposure, 256–257
- Tissue inhibitors of metalloproteinases (TIMPs)
 - activities of, 1764–1766
 - connective tissue healing, 1252
- Tissue-material interactions
 - bioactive glass, 136
 - polyurethane biomaterials, 2311–2312
- Tissue-processing methods, allografts, 8–9
- Tissue properties
 - mass transfer
 - complex mechanisms, 1761
 - concentration diffusion, 1756–1761
 - flow-dominated systems, 1752–1756
 - intracellular organelles, 1751–1752
 - major organ systems, 1749–1751
 - molecular complexes, 1752
 - organ function, 1751
 - overview, 1749
 - spatio-temporal hierarchies, 1749–1752
 - whole cells, 1751
 - mineralization, bone healing, 1258–1259
 - tendon/ligament tensile testing, 2590–2592
 - thickness effects, small-angle light scattering, 2456
- Tissue sealants
 - basic properties, 158–159
 - fibrin glue, 1187–1188
- Titanium and titanium alloys
 - basic properties, 2876–2877
 - β -titanium, orthodontic wires, 2057
 - biocompatibility, 2879–2880
 - biomedical applications, 2877–2878
 - bone-implant interactions, 1230–1231, 2880
 - ceramic coatings, 590
 - corrosion mechanics
 - stress corrosion cracking, 756–757
 - surface analysis, 758–759
 - cytotoxicity, 2879–2880
 - dental implants, 813, 2887–2889
 - microtextured titanium, 814
 - fatigue and fretting fatigue, 2881–2884
 - fracture fixation, 414
 - hydroxyapatite coatings, 1464–1468
 - mechanical properties, 2881–2885
 - nickel alloys, 1945–1947
 - orthodontic wires, 2057–2059
 - orthopedic devices, 2063
 - nickel-free alloys, 2879
- orthopedic devices, 2063
 - fracture fixation and joint implantation, 1826–1827
 - screws, failure analysis, 1834–1835
 - plasma deposition and coating, 2181–2183
 - stents, in vivo animal models, 1570
 - surface modification, 2885–2887
 - total hip arthroplasty, 1476–1478
 - wear characteristics, 2884–2885
- T-NCVC coating, heparin biocompatibility, 1372
- Toe region, articular cartilage biomechanics, 87–91
- Tolerance mechanisms,
 - microelectromechanical systems, 1863–1864
- Top-down processes
 - magnetic nanoparticles, 1705
 - microelectromechanical systems, 1862–1865
- Topographical analysis
 - cell structure and behavior
 - chemistry vs., 2552
 - classification, 2551–2556
 - discontinuous substratum, 2555
 - fibers, 2553
 - microfabrication, 2552
 - nanotopography vs. microtopography, 2556
 - overview, 2551
 - pits/holes, 2553
 - repeating grooved substrata, 2553
 - roughened surfaces, 2555–2556
 - single steps, 2552–2553
 - substratum influences, 2556–2558
 - tunnels/tubes, 2555
 - in vivo applications, 2558–2560
 - tissue engineering, microscale technology, 2781–2782
- Torque evaluation, dental implants, 816
- Total artificial heart (TAH)
 - centrifugal blood pumps, 581–583
 - structure and classification, 475–477
- Total artificial lungs (TALs), 1661–1678
- Total disc replacement (TDR)
 - arthroplasty techniques, 2469
 - spine biomechanics, 2465
- Total hip arthroplasty (THA)/total hip replacement (THR)
 - alloys and surface coatings, 1476–1478
 - alumina joint prostheses, 20–21
 - alumina prostheses, 19–21
 - biomaterials selection, 1379–1380
 - carcinogenesis, 1480
 - ceramics, 1474
 - cobalt-chromium alloys, 1476, 1478
 - current technology, 1473–1478
 - degradation problems, 1478–1480
 - corrosion, 1479–1480
 - metal ion release, 1480
 - particle distribution, 1480
 - wear, 1478–1479
 - design evolution, 1472–1473
 - dislocation following, 1381
 - femoral head short-stem prostheses, 1471–1472
 - femoral head subluxation and dislocation, 1381–1382
 - future research issues, 1480–1482
 - hip joint biomechanics
 - overview, 1377
 - surgical techniques, 1377–1378
 - hypersensitivity, 1480
 - immune response, major histocompatibility complex, 266–267
 - long-stem prostheses, 1472
 - metals, 1474–1478
 - mold arthroplasty, 1470–1471
 - overview and history, 1470–1473
 - polymers, 1474
- refractory metals, 1478
- stainless steel alloys, 1474–1476
- techniques and materials, 1472–1473
 - carcinogenesis, 1480
 - ceramics, 1474
 - cobalt-chromium alloys, 1476–1478
 - current technology, 1473–1478
 - degradation, wear and corrosion, 1478–1480
 - femoral head prostheses/short-stem prostheses, 1471–1472
 - future research issues, 1480–1482
 - historical background, 1470–1473
 - hypersensitivity responses, 1480
 - limitations and complications, 1478–1480
 - long-stem prostheses, 1472
 - metal ion release, 1480
 - metals, 1474
 - mold arthroplasty, 1470–1471
 - new alloys and surface coatings, 1476–1478
 - overview, 1470
 - particle distribution, 1480
 - polymers, 1474
 - refractory metals, 1478
 - stainless steel alloys, 1474–1476, 1478
 - titanium alloys, 1476–1478
 - titanium alloys, 1476, 1478
- Total internal reflection fluorescence (TIRF),
 - gradient surfaces characterization, 1219
- Total knee replacement (TKR)
 - alumina joint prostheses, 21
 - anatomical properties, 1608–1609
 - component fixation, 1612–1613
 - contraindications, 1608
 - indications for, 1607–1608
 - overview and history, 1607
 - postoperative function, 1613
 - prosthetic design factors, 1609–1610
- Total surface immersion, polyelectrolyte multilayer deposition, 2219
- Toxicity and toxicology
 - animal surrogate testing
 - barrier models, 53–55
 - cell culture analog systems, 57–58
 - chemical structure-related models, 51–52
 - coculture models, 55
 - computational models, 51–53
 - conditioned medium systems, 56–57
 - microscale designs, 57–58
 - multiple cell devices, 56–57
 - overview, 51
 - shared medium systems, 57
 - single cell in vitro models, 53
 - three-dimensional models, 55–56
 - whole organism models, 52–53
 - electrical cell impedance sensing, 912–913
 - host reactions, 1418
 - nitric oxide, cellular toxicity, 1953
 - photopolymerization mechanisms, 2150–2153
 - polytetrafluoroethylene, 2301–2302
 - polyurethane biomaterials, 2306
 - reconstructive biomaterials, 2397–2398
- Trademarks, intellectual property rights and, 1543–1544
- Trade-Related aspects of Intellectual Property Rights (TRIPS) agreement, 1540–1548
- Trade secrets, intellectual property rights and, 1543–1544
- Training programs
 - errors, stress fractures and, 2513
 - simulators
 - future applications, 1812
 - surgical robotics, 1811
 - telemedicine systems, 2585
- Transanastomotic healing, vascular grafts, 2991

- Transcellular pathways, protein-based drug delivery systems, 2360
- Transcranial Doppler (TCD) examination, basic principles, 2947
- Transcriptional factors, alginate biosynthesis, 353
- Transcription regulation, integrin signaling, 1533
- Transcutaneous energy transmission system (TETS), electric ventricular assist device, 473
- Transcutaneous insulin transport, 1516
- Transdermal drug delivery
backing materials, 2896
basic principles, 2893
biocompatibility strategies, 2902–2904
inflammatory response, 2904
metabolism inhibition, 2904
skin reactions, 2903–2904
clinical applications, 2897–2899
drug reservoirs, 2896
electrotransport mechanisms, 2897, 2900–2901
applications, 2901
control circuitry, 2900
electrodes, 2900–2901
power source, 2900
insulin administration, 2366
mechanically-assisted transport, 2901–2902
micropore creation, 2902
microprojection systems, 2902
needleless injection systems, 2902
passive transport principles, 2894–2895
peelable layers, 2897
permeation enhancers, 2896
pressure-sensitive adhesives, 2896–2897
rate-controlling membranes, 2896
skin structure and function, 2893–2894
system components, 2895
- Transdermal electrical stimulation, tinnitus, 2612
- Transduction process
cellular mechano-transduction, 1789–1790
hearing aid design and, 1263–1264
- Transesophageal echocardiography (TEE), transmyocardial laser revascularization, 1634–1637
- Transfection, chitosan agents, 609–611
- Transformation-base compression, ECG signal compression, 680–681
- Transformation-toughened and platelet-reinforced (TPPA) alumina, properties of, 23
- Transforming growth factor beta (TGF- β)
osteogenic progenitor cells, 2094–2095
polymeric structure, 221
- Transgenic organs, xenograft techniques, 2906–2912
- Transmembrane action potential
cardiac electrophysiology, 482–487
ionic currents, 483–486
optical mapping, 2019
- Transmission electron microscopy (TEM)
bioreactor analysis, 312
instrumentation, 968–969
overview, 968
SEM vs., 969–970
specimen preparation, 973–974
- Transmucosal drug delivery, insulin administration, 2366
- Transmural healing, vascular grafts, 2991–2992
- Transmyocardial laser revascularization (TMR)
clinical results, 1637
overview, 1632–1633
procedures and techniques, 1633–1636
- Transplant devices and materials
artificial lungs
blood oxygenator, 1666–1667
cardiopulmonary bypass oxygenators, 1667–1669
diffusional boundary layers, 1665
extracorporeal membrane oxygenation, 1669–1670
gas exchange determinants, 1663
historical background, 1661–1662
mass transfer correlations, 1665–1666
membrane permeance, 1663–1665
operating principles, 1662–1666
overview, 1661
total artificial lungs, 1677–1678
- artificial pancreas, 2802–2809
- bio-artificial liver systems, 1649–1658
- centrifugal blood pump systems, 581–583
- histogenesis, 1409
- liver, tissue engineering
bioreactors, 2772–2773
coatings, 2771–2772
future research issues, 2775–2776
hepatic environment, 2769–2771
immune response, 2772
liver-assist devices, 2773–2774
liver transplants, 2773–2775
overview, 2769
scaffold materials, 2771
- nerve guide technology, 1913–1914
- small intestine tissue engineering, 2860–2861
- tissue engineering
articular cartilage, 2701
bladder cell transplants, 2636–2640
microvascular networks, 2792–2799
- transgenic/gene-targeted organs
acute vascular rejection, 2909–2910
genetic engineering, 2910–2912
hyperacute rejection, 2908–2909
immune response, 2907–2910
overview, 2906
T cell-mediated rejection, 2910
xenotransplantation, 2906–2912
- in vivo stem cell therapy, 2476–2477
- xenografts
cellular and tissue grafts, 3099–3100
ethical issues, 3100
future research projections, xeno–9
historical background, 3093–3094
immune response, 3095–3100
cell-mediated rejection, 3099
chronic rejection, 3099
organ transplants, 3095–3097
vascular grafts, 3097–3098
overview, 3093
physiology, 3095
sources, 3094–3095
zoonosis, 3100
- Transport theory
light propagation in tissue, 2033
tissue optics, 1625
- Transverse isotropic composites, micromechanics, 668–13
- Trauma, orthognathic surgery, bioabsorbable plates and screws, 435–436
- Traveling wave dielectrophoresis (TWD), tissue engineering, 2630–2631
- Triactiv (Kensey Nash) distal protection device, 873
- Tribological structures, diamond/diamond-like carbons, 826
- Tricalcium phosphate (TCP)
 β -tricalcium phosphate, 463–464, 612–613
biphasic calcium phosphate preparation, 360
ceramics
implant devices, 587–588
orthophosphates, 461–462
transparent ceramics, 2916–2923
- chitosan scaffolds, 611–613
composites
biocomposites, 183–184
degradable polymer-ceramics, 783–784
microporous materials, 1886–1887
osteoinductive substances and materials, 2105–2106
reconstructive biomaterials, facial skeletal fixation, 2398
synthetic apatites, 230–231
tissue engineering, bone, 2655
- Trichloroethene (TCE), cardiac catheterization coatings, 497–498
- Tricuspid valve
hemodynamics, 1323
mechanical properties, 1323–1326
structure and function, 1320–1323
- Trigonometric excitation, electrical impedance imaging, 923–924
- Trillium™ Biopassive Surface, 1371
- Trimethyl chitosan chloride (TMC), testing of, 606–607
- Tropoelastin assembly, elastin ultrastructure, 893–894
- Tube hematocrit, vascular fluid dynamics, 2979
- Tube viscometers, vascular fluid dynamics, 2977–2978
- Tubular structures, tissue engineering, intestinal scaffold materials, 2754
- Tumor growth
angiogenesis inhibitors, collagen type IV, 27
host reactions, 1419
magnetic resonance microscopy, 1745–1746
nanoscale neurosurgery, 1908–1909
- Tumor necrosis factor-alpha (TNF- α)
host reactions, 2111–2114
osteoclast differentiation and activation signaling, 2086
- Tumor necrosis factor receptor associated factors (TRAF), osteoclast differentiation and activation signaling, 2087
- Tumor necrosis factor (TNF), wear debris, inflammatory response, 3067–3068
- Tungsten lamps, optical mapping, 2020
- Tunnel/tube structure, topographical surface analysis, 2555
- T wave, electrocardiography, 939
- Twitch tension/resting tension (TT/RT) ratio, cardiac muscle tissue engineering, 2690–2692
- Two-dimensional engineering
microelectromechanical systems, 1863
microvascular networks, 2797
- Two-dimensional imaging, electromyographic signals, 965
- Two-dimensional submillisecond mapping, transmembrane polarization, heart defibrillation, 2024
- Two-photon microscopy, 712–713
- Tyrosinase, aquatic adhesives, 240–241
- ## U
- Ulcers
healing deficiencies and, 1254
tissue engineering, small intestine, 2759
- Ultrafiltration rate, dialysis membrane convection, 1587–1588
- Ultra-high molecular weight polyethylene (UHMWPE)
alumina inertness and, 17
future applications, 23
joint replacements, 19–21
future research issues, 2929–2930
meniscal insert materials, 1610–1611
orthopedic devices, 2065
overview, 2925

- particulate matter, host reactions, 2110–2111
- sterilization and structural modification, 2927–2929
- structure and properties, 2927–2929
- synthesis and processing, 2925–2927
- total hip arthroplasty, 1379–1380
- current technology, 1743–1474
- finite element modeling, 1383–1385
- wear debris, joint arthroplasties, 3076–3079
- Ultrasonic hearing, tinnitus, 2615–2616
- Ultrasound
- bone healing
- biologic effects, 2956
- clinical applications, 2956–2957
- complications and side effects, 2957
- future research issues, 2957–2959
- historical perspective, 2951–2953
- maxillofacial bones, 2957
- physics, 2953–2955
- piezoelectric effect, 2955–2956
- tissue interaction, 2955–2956
- cardiac elastography, 507–512
- controlled-release drug delivery systems, 886
- Doppler principles, 2942–2943
- arterial disease, 2948
- cardiac examination, 2948–2949
- carotid artery examination, 2947
- color imaging, 2945–2946
- continuous wave Doppler, 2943–2944
- duplex scanning, 2944–2945
- power Doppler imaging, 2946–2947
- pulsed wave Doppler, 2944
- tissue imaging, 2947
- transcranial examination, 2947
- endoscopic ultrasound, 990–991
- hard materials, elasticity properties, 1236–1241
- intravascular ultrasound
- clinical settings and treatments, 1559–1560, 1562
- future research issues, 1562–1564
- imaging studies and modalities, 1560–1562
- overview, 1559
- medical device infection prevention, 1503
- real-time tomographic reflection, 2392–2393
- shape memory metal surfaces, 2421–2422
- three-dimensional imaging and, 1797
- tinnitus devices, bone vibration maskers, 2614–2615
- tissue interaction
- acoustic propagation principles, 2933–2934
- attenuation, absorption, and scattering, 2936–2939
- characteristic impedance, 2934
- intensity, 2934–2935
- linearity effects, 2939–2941
- radiation force, 2935
- reflection/refraction, 2935–2936
- vascular imaging, 3014–3015
- Ultraviolet (UV) radiation
- collagen
- degradation, 634–635
- fixation and cross-linking, 647–648
- glycation, 647–648
- surface modification, 2541–2543
- Umbilical cord blood, as hematopoietic stem cell source, 1340–1342
- Uniaxial testing, blood vessel mechanics, 394
- Unimolecular dendritic micelles, dendrimer-base drug delivery, 790–793
- Union for the Protection of New Varieties of Plants (UPOV), 1544–1545
- Universal approximation, neural networks, 126
- Upper extremities
- disorders of, 1806
- ergonomics guidelines and, 1002
- gait analysis, 1137
- neuroprostheses for, 1927–1929
- robotics, 1806
- Ureas, silicone copolymers, 2433–2434
- Uremic solutes, 1577–1578
- Urinary catheterization, infection prevention, 1503
- Urinary incontinence, artificial urinary sphincter, 2966–2968
- Urinary prostheses
- artificial sphincter, 2966–2968
- basic properties, 2961
- biomaterial-tissue interaction, 2961
- penile prostheses, 2962–2966
- clinical indications, 2964
- inflatable implants, 2963–2964
- outcomes and complications, 2964–2965
- regenerative medicine techniques, 2965–2966
- semirigid devices, 2963
- polyurethanes, 2962
- silicone, 2961–2962
- testicular prostheses, 2968–2969
- tissue interactions, 2962
- Urinary tract infections (UTI), incidence and prevalence, 1497
- Urologic procedures, fibrin sealants, 1086
- Utilitarianism, biomedical ethics and, 289–290
- U wave, electrocardiography, 940
- V**
- Vacuum arc technique, plasma discharge, 2175
- Vacuum mixing, bone cements, 409
- Vacuum ultraviolet (VUV) radiation, surface modification, 2541–2543
- Validation, three-dimensional imaging techniques, 1801–1802
- Valvular anatomy and function
- atrioventricular valves
- hemodynamics, 1323
- mechanics, 1323–1326
- structure and function, 1320–1323
- biomechanics, 1282–1286
- force-length curves, isolated cardiac muscle, 1282–1283
- future research issues, 1288–1289
- global/regional deformation measurement, 1286–1287
- growth and remodeling, 1287–1288
- myocardial material properties, 1284–1286
- myocardial stresses, 1283–1284
- pressure-volume curves, 1280–1282
- pump functions, 1280–1282
- structural constitutive modeling, 1286–1287
- ventricular function curves, 1280
- cardiac assist devices, positive displacement pumps, 96
- Doppler ultrasound imaging, 2948–2949
- semilunar valves
- hemodynamics, 1319
- mechanics, 1319–1320
- structure and function, 1317–1319
- structural properties, 1317
- Valvular prosthetics
- allografts, 11–13
- biomineralization
- anticalcification treatments, 79–81
- cells and cell remnants, 73
- collagen, 74
- elastin, 74–75
- ground substance, 75
- overview, 71
- tissue calcification, 75–79
- diamond/diamond-like carbons, 828–829
- failure analysis
- calcium binding site inactivation, 1308–1309
- design failures, 1312–1313
- future research issues, 1313–1314
- glutaraldehyde cross-link chemistry, 1306–1308
- overview, 1295–1301, 1306
- pannus formation, 1300–1301
- remnant tissue immunogenicity, 1309–1310
- historical review, 1294–1295
- limitations of, 1291–1292
- mechanical
- blood-material interactions, 1334–1336
- cavitation, 1331
- classification, 1329
- durability, 1329–1330
- fluid mechanics applications, 1331–1332
- overview, 1329
- positive displacement pumps, 96
- in vitro flow field measurements, 1332–1334
- in vitro testing, 97–98
- wear and fatigue failure, 1330–1331
- polyepoxy compounds, 2235–2236
- pyrolytic carbon, 2370–2374
- tissue engineering
- aortic valve cusp microstructure and micromechanics, 2738–2742
- biomimetic constructs, 2743–2747
- bioprosthetic valve failure mechanisms, 2737–2738
- calcification and mechanical damage, 2738
- cellular population, 2742
- collagen, 2739–2740, 2744
- cross-linked hyaluronan, 2744–2746
- degeneration mechanisms, 2738
- elastin, 2740–2742, 2745–2747
- matrix mechanics, 2742–2743
- overview, 2737
- Valvular surgery, computer-assisted techniques, 691
- VanSlyke manometric apparatus, cardiac output measurement, Fick method, 518
- Vapor deposition polymerization (VPD), coating materials, 2532
- Vascular disease. *See* Cardiovascular disease
- Vascular endothelial growth factor (VEGF) angiogenesis inhibitors, research background, 26–27
- bone formation, 2045–2047
- nitric oxide, 1952
- ocular angiogenesis, 28
- tissue engineering
- skeletal muscles, 2845–2848
- vascular grafts, 3023–3024
- Vascular endothelium, anticoagulant activity, heparin biocompatibility, 1368–1369
- Vascular grafts. *See also* Fabrics
- basic properties, 2989–2993
- biocompatibility, 2990–2991
- bioprosthetics, 2985–2987
- blood-materials interactions, 375–376
- cardiac catheterization, 501–503
- clinical applications, 2985
- development strategies
- bioresorbable materials, 2999
- endothelial cell seeding, 2999–3001
- overview, 2997
- surface modifications, 2997–2998
- tissue engineering, 3001
- in vivo healing manipulation, 2998

Vascular grafts (*cont.*)

distal protection devices

- AccUNET device, 875–876
- advantages/disadvantages, 877
- Angioguard device, 873–874
- carotid artery, 870–871
- current research evolution, 877–878
- embolization reduction devices, 871
- failure mechanisms, 877
- Filterwire Ex, 875
- future applications, 878
- NeuroShield device, 874–875
- overview, 870
- Parodi Anti-Emboli System, 876–877
- Percutaneous guardwire, 871–874
- saphenous vein grafts, 870

endothelial cell techniques, 2992–2993

endovascular stents

- balloon-expanded stents, 2486
- bifurcated grafts, 2482
- branching techniques, 2480–2482
- breakage issues, 2486–2487
- delivery system diameter, 2489–2490
- design innovation, 2484
- fabrics, 2484–2485
- fenestrated grafts, 2483
- fixation techniques, 2487–2488
- future designs, 2487
- historical perspectives, 2492
- medical applications, 2480
- modularity, 2483–2484
- multibranching grafts, 2482–2483
- sealing techniques, 2488–2489
- self-expanding stents, 2486
- support systems, 2489
- unbranched grafts, 2480–2482

gene-activated matrix regeneration, arterioles, 1147–1148

healing mechanisms, 2991–2992

host reactions

- compliance, 3008–3009
- endothelial/smooth muscle cell ingrowth, 3008
- failure and infection, 3009
- flow surface properties, 3009
- healing characteristics, 3008–3009
- neutrophil-monocyte infiltration, 3007–3008
- overview, 3006
- platelet deposition, 3007
- porosity, 3008
- protein adsorption, 3006–3007

hybrid materials

- biological sealants, 1432–1433
- endothelial seeding, 1433–1434
- mandrel-guided FCTs, 1434–1436
- overview, 1432
- tissue engineered substitutes, 1436–1437

intravascular ultrasound, 1559–1564

mechanical properties, 2989–2990

microvascular tissue engineering

- assembly techniques, 2796–2797
- biocompatible polymer molding, 2796
- biodegradable polymer fabrication, 2796
- endothelial cell seeding, 2798
- fluid dynamics, 2793–2794
- lithographic techniques, 2794–2795
- organ and tissue properties, 2792–2793
- plasma etching, 2795–2796
- polymer replica molding, 2796
- porous devices, 2797–2798
- scaffold fabrication, 2794–2798
- vascular fluid dynamics, 2793–2794
- vascularization requirements, 2792
- whole organ engineering, 2798–2799

overview, 2985

phosphorylcholine technology, 2142–2145

polyepoxy compounds, 2237–2238

growable compounds, 2237–2238

porosity, 2989

single-stage venous EC seeding, 2992

synthetic grafts, 2987–2989

tissue engineering, 2622, 3033–3042

- cell sources, 2645–2646
- cellular-based scaffolds, 2647
- clinical indications, 2643–2644
- collagen-based scaffolds, 2646–2647
- construct design, 2646–2647
- design criteria, 2644–2647
- fat tissue, 2730–2731
- fibrin structure and biomechanics, 1072–1077
- future research issues, 2650
- ligaments, 2765
- limitations, 2648–2650
- liver, 2773–2775
- mechanical conditioning, 2648
- mimicry properties, 2645
- scaffold materials, 2647–2648
- vascular anatomy, 2643
- in vitro procedures, 2992–2993
- in vivo procedures, 2993
- in vivo animal models, 1566–1571
- xenografts, acute rejection, 2909–2910, 3097–3098

Vascular system. *See also* Blood vessels;

Circulatory system

Doppler color imaging, 2945–2946

fluid dynamics

- arteries, 2972–2981
- electric analogues, arterial system, 2976–2977
- inviscid analysis, 2972–2974
- microcirculation, 2977–2981
- overview, 2971–2972
- pulmonary circulation, 2982–2983
- systemic circulatory flow, 2972–2982
- vascular disease states, 2983–14
- veins, 2981–2982
- viscous analysis, 2974–2975

imaging techniques

- angiography, 3012–3013
- angiосcopy, 3017
- computed tomography angiography, 3013–3014
- future research issues, 3018–3020
- magnetic resonance angiography, 3015–3017
- overview, 3012
- processing protocols, 3017–3018
- ultrasound, 3014–3015

tissue engineering

- bioreactors and in vitro procedures, 3039–3040
- cell sources, 3036–3039
- clinical applications, 3033, vasctiss–9
- future research issues, 3042
- scaffold materials, 3033–3035
- seeding techniques, 3035–3036
- in vivo remodeling, 3040–3041

Vasovagal syncope, pacing therapy, 540

Vectors

- gene-activated matrix, 1144–1145
- gene therapy
 - nonviral vectors, 1151–1154
 - electroporation, 1153
 - gene gun delivery, 1151–1153
 - liposomes, 1153
 - naked DNA, 1151
 - synthetics, 1153–1155
 - viral vectors, 1154–1157
 - adeno-associated virus, 1154–1155
 - adenovirus, 1156–1157
 - herpes virus, 1155–1156
 - retrovirus, 1155

Veins and venules

- macrocirculatory hemodynamics, 1365–1366
- mechanical properties, 399

vascular fluid dynamics, 2981–2982

venules, microcirculatory oxygen

transport, 1846–1847

vessel wall properties, 1358–1359

Velocity mapping, cardiac assist devices,

positive displacement pumps, 97–105

Velocity profiles, vascular fluid dynamics,

2975–2976

Venous catheterization, techniques and

materials, 501–503

Ventricular assist devices (VADs), 472–475.

See also Left ventricular assist devices

(LVADs)

continuous flow, 473–475

defibrillators

biophysics, 773–774

external, 774

implantable, 774–777

overview, 772

ventricular fibrillation/tachycardia,

772–773

positive displacement pumps, 95–105

pediatric devices, 97–100

pulsatile flow, 472–473

Ventricular fibrillation, basic principles, 772–773

Ventricular function curves, heart

biomechanics, 1280

Ventricularis, aortic valve cusp

microstructure, 2738–2742

Ventricular late potential (VLP), digital signal

acquisition and processing

biomedical signals, 845–849

wavelet transform, 857–858

Ventricular structure and function

cardiac action potential, 485–486

electrocardiography

depolarization, 938–939

repolarization, 939

electrophysiology, 480–489

Ventricular tachycardia, basic principles,

772–773

Verification studies, cardiac output

measurement, 527–530

Versamid, basic properties, 2190–2191

Vertebral fractures, biomechanics, 2468–2469

Vertebroplasty, bone cement, 403

Very high-field magnetic resonance imaging

scanner, functional magnetic resonance

imaging, 1126

Very large scale integration (VLSI), neural

networks, ALOPEX algorithm, 126–128

Viability assessment, neutrophil-biomaterial

adhesion, 1941

Vibration acoustic masking, tinnitus, 2614

Vibrational spectroscopy, hydroxyapatite

characterization, 1458

Vickers microhardness measurements,

orthopedic device failure analysis,

1831

Video-teleconferencing (VTC)

clinical applications, 2579–2780

real-time telemedicine systems, 2580–2581

Viral vectors, gene-activated matrix,

1144–1145

Viscoelasticity

articular cartilage biomechanics, 87–91

hard tissue, 1243–1244

movement control models, 734

ocular implant agents, 1979

tendons and ligaments, 2595–2597

tissue engineering, 2872–2873

Viscometry, vascular fluid dynamics,

2977–2978

Viscosity

blood composition and structure,

macrocirculatory hemodynamics,

1357–1358

blood flow modeling, microvascular tissue

engineering, 2793–2794

- bone cement, 406–407
 magnetic cell separation, 1698–1699
 vascular fluid dynamics, 2974–2975
 stress-shear-rate laws, 2978–2979
- Visual function and impairment
 eye movement research, 1034–1038
 robotics applications, 1812
- Visualization, three-dimensional imaging
 techniques, 1800–1801
- Visual similarity, digital biomedical signal
 compression, 679
- Vitrasert reservoir ocular implants, 1988–1989
- Vitrification
 cartilage, 3050–3051
 ceramics, 585–586
 cryoinjury prevention, 766
 cryopreservation, 766
 future research issues, 3053–3054
 overview, 3046
 physiochemistry, 3046–3049
 stability properties, 3049
 tissue engineering, 3049–3050
- Vitronectin, hemocompatibility, platelet
 interaction and, 1353
- Vocational assistance devices, robotics,
 1804–1805
- Voight-Kelvin elements, composite
 micromechanics, 666–671
- Voigt orthotropic modulus, elastic properties,
 1246
- Voigt transverse isotropic modulus, hard
 materials elasticity, 1245
- Voltage properties
 electrical impedance tomography, 917–918
 implantable batteries, 1486
- Voltage-sensitive dyes, optical mapping, 2019
 fluorophore design, 2019–2020
 pharmacological artifacts, 2021
- Volume conductors, electrical stimulation,
 1010–1012
- Volume rendering, three-dimensional imaging
 techniques, 1801
- Volumetric shrinkage, bone cement, 407–408
- von Willebrand factor (vWF),
 hemocompatibility, platelet interaction
 and, 1353
- Vorticity mapping, cardiac assist devices,
 positive displacement pumps, 102–105
- Vroman effect, multicomponent protein
 adsorption, 276–278
- W**
- Walking, movement control mechanisms,
 731–732
- Wallace's test, hamstring syndrome, 1400
- Wallerian degeneration, mechanisms of, 1913
- Washington State Appendix B, ergonomics
 guidelines, 1003–1004
- Waste removal, hemodialysis, 379
- Water contact angle, gradient surfaces,
 1210–1211, 1216–1217
- Water removal, hemodialysis, 381–382
- Water-soluble carbodiimide (WSC), GRF
 glues, 1190–1191
- Water suppression, nuclear magnetic
 resonance, brain imaging, 1969
- Water/wastewater treatment, interfacial
 transport, 2349
- Wave equation, acoustic propagation,
 ultrasound, 2933–2934
- Waveform characteristics
 biomedical optics, 2029–2030
 defibrillation systems, 773–774
 electrical stimulation, 1014–1016
 radiofrequency electromagnetic field/
 radiation effects, 255
- Wavelet transform (WT), digital biomedical
 signals
- acquisition and processing, 854–858
 ECG signal compression, 680–681
 EMG signal compression, 681–682
 spectral analysis, 858–861
- Wear analysis
 alumina, 18–19
 bone resorption animal models
 cyclooxygenase-2/prostaglandin E2, 3070
 cytokines, 3069–3070
 osteoclastogenesis, 3068–3072
 osteolysis, 3070–3072
 proinflammatory cellular response,
 3067–3068
 Rank ligand, 3068–3069
 debris properties
 future research issues, 3063
 hard-on-hard bearing couples, 3060–3062
 hip prostheses, polyethylene debris,
 3059–3060
 knee prostheses, 3062–3063
 overview, 3058–3059
 joint arthroplasties
 bone-implant interface, 3076–3078
 macrophages, 3078
 mesenchymal stem cells, 3079
 osteoblasts, 3079
 osteoclasts, 3078–3079
 overview, 3076
 mechanical heart valves, 1330–1331
 titanium and titanium alloys, 2884–2885
 total hip arthroplasty implants, 1478–1479
 ultra-high molecular weight polyethylene,
 2930
- Wear particles. *See* Particulate matter
- Weibull function, ceramic fracture and
 fatigue, 586–587
- Welding, alginate applications in, 354
- Well's test, hamstring syndrome, 1400
- Wet chemistry, surface modification, 2545–2547
- Wettability
 cell-material interactions, 569–570
 gradient surfaces, 1211–1213
 gas/solvent diffusion, 1211–1212
 radio-frequency plasma discharge,
 1212–1213
- Wetting theory, bioadhesion, 1–2
- Wheeled mobility
 demographic overview, 3083–3084
 distribution and payment issues, 3090–3091
 electric-powered wheelchairs, 3085
 future research issues, 3091–3092
 hybrid devices, 3089–3090
 manual wheelchairs, 3084–3085
 power and control systems, 3085–3087
 robotics, 3089–3090
 scooters, 3088–3089
 seating design, 3087–3088
 standards, 3090
 structural components, 3085
- White blood cells, adhesion/activation,
 heparin biocompatibility, 1369
- White room procedures
 tanning industry, collagen processing,
 654–655
- Whole-cell biocatalysis, techniques and
 applications, 163–165
- Whole-muscle-macroscopic models,
 movement control, 734
- Whole organism models, toxicity testing,
 animal surrogate systems, 52–53
- Whole organ transplants, allografts and, 9–10
- Wide dynamic range compression (WDRC)
 hearing aid design, 1268–1269
 hearing loss and, 1275–1277
- Wiener-Khinchin theorem, biomedical signal
 acquisition and processing, 849–850
- Wigner-Ville distribution, digital biomedical
 signal acquisition and processing,
 spectral analysis, 858–861
- Wilhelmy plate measurement, bioadhesion
 measurements, 5
- Windkessel theory, macrocirculatory
 hemodynamics, arterial pulse waves,
 1364–1365
- Wire fixation devices
 external fracture fixation techniques,
 420–421
 fracture fixation, 414, 2074–2075
 orthodontic
 beta-titanium, 2057
 cobalt-chromium alloys, 2057
 nickel-titanium alloys, 2057–2059
 overview, 2054–2056
 stainless steel, 2056–2057
 strengths and limitations, 2059
- Wireless capsule endoscopy, 992–993
- Wireless technology
 radiofrequency electromagnetic
 field/radiation (RFR) effects,
 254–258
 telemedicine applications, 2582–2584
- Wolff's Law, bone healing, 1259–1260
- Womersley models
 macrocirculatory hemodynamics
 arterial pulse waves, 1364–1365
 pulsatile flow properties, 1361–1364
 vascular fluid dynamics, 2974–2975
- Wool fibers, 2188
- Workplace design
 ergonomics analysis, 1005–1006
 eye movement research and, 1036–1037
- Work-related musculoskeletal disorders
 (WMSDs), ergonomics principles,
 999
 low back, 1000–1001
 office review, 1005–1006
 upper extremity, 1001
- Wound healing
 alginate applications in, 353–354
 bone tissue engineering, animal models,
 39–40
 burn injuries, 449–450
 chitosan and, 604
 cyanoacrylate glues, 1188
 electrospun polymer nanofibers, 983
 fibrin structure and biomechanics,
 1072–1077
 gene-activated matrix, 1143
 hyaluronan and, 1423, 1428
 laser-tissue interactions, 1630
 natural adhesives, 247–249
 sutures
 absorption properties, 2575–2577
 biocompatibility, 2574–2575
 biodegradability, 2575–2577
 commercial size and configuration,
 2563–2568
 handling properties, 2570–2574
 materials properties and classification,
 2568–2577
 needles, 2562–2563
 overview, 2562
 physical and mechanical properties,
 2569–2570
 size classifications, 2568
 tissue engineering
 corneal implants, 2712–2714
 skin, 2851–2854
- Woven fabric
 three-dimensional structure, 1050
 two-dimensional structure, 1048–1049
 vascular grafts, 2987–2988
- X**
- Xenografts
 aortic valve cusp microstructure, 2742
 artificial liver, 1650–1658

Xenografts (*cont.*)

artificial pancreas, 2802–2807
 burn injuries, 455–456
 cell sources, 2283
 cellular and tissue grafts, 3099–3100
 collagen scaffolds, 641–644
 defined, 9
 ethical issues, 3100
 fat tissue engineering, 2725–2734
 future research projections, xeno–9
 historical background, 3093–3094
 immune response, 3095–3100
 cell-mediated rejection, 3099
 chronic rejection, 3099
 organ transplants, 3095–3097
 vascular grafts, 3097–3098
 nerve tissue engineering, 2815
 overview, 3093
 peripheral nerves, 2127–2130
 physiology, 3095
 sources, 3094–3095
 transgenic/gene-targeted organs
 acute vascular rejection, 2909–2910
 genetic engineering, 2910–2912
 hyperacute rejection, 2908–2909
 immune response, 2907–2910
 overview, 2906
 T cell-mediated rejection, 2910
 xenotransplantation, 2906–2912
 valvular prosthetics
 calcium binding site inactivation,
 1308–1309
 design failures, 1312–1313
 future research issues, 1313–1314
 glutaraldehyde cross-link chemistry,
 1306–1308
 inflammatory degradation, 1298–1299
 matrix mechanics, 2742–2743
 overview, 1306
 remnant tissue immunogenicity, 1292,
 1309–1310

vascular grafts, 2986–2987
 hybrid materials
 biological sealants, 1432–1433
 endothelial seeding, 1433–1434
 mandrel-guided FCTs, 1434–1436
 overview, 1432
 tissue engineered substitutes,
 1436–1437
 zoonosis, 3100
 X-linked severe combined immunodeficiency
 (X-SCID), gene therapy, 1159–1160
 X-ray angiography, intravascular ultrasound,
 1561–1564
 X-ray computed tomography (X-ray CT),
 three-dimensional imaging,
 1794–1795
 X-ray diffraction (XRD)
 bone mineral composition, 2041–2042
 hydroxyapatite characterization, 1458
 X-ray fluorescence (XRF), hydroxyapatite
 characterization, 1458
 X-ray lithography, microelectromechanical
 systems, 1861
 X-ray photoelectron spectroscopy (XPS)
 gradient surfaces characterization,
 1217–1218
 hemocompatibility assessment,
 1349–1350
 X-rays, microcomputed tomography
 attenuation, 1852
 image generation, 1851–1852
 theoretical background, 1851
 Xylosyltransferase, proteoglycan/
 glycosaminoglycan biosynthesis,
 1199–1200

Y

Yarns, structure and properties, 1047–1048
 Yellowing properties, polyurethane softening,
 2212

Young's modulus
 alumina ceramics, 21
 brain structure, nanoscale neurosurgery,
 1903–1904
 composite micromechanics, 667–671
 finite element analysis, 1094
 hard materials
 elasticity properties, 1240–1241
 modeling techniques, 1242–1244
 pyrolytic carbon, 2376
 vascular fluid dynamics, 2973–2975
 Yttrium-tetragonal zirconia polycrystal
 (YZTP) compounds,
 3104–3108

Z

Zeeman effect
 magnetic resonance imaging, 1715
 nuclear magnetic resonance, 1961
 Zero-crossing rate (ZCR), biomedical signal
 acquisition and processing, 853
 Zero-order drug release, controlled release
 systems, 740–742
 Zeroth moments, mass transfer and tissue
 function, 1753–1755
 Zeus system, computer-assisted surgery,
 688–689
 Zinc phosphate cements, dental adhesives,
 2, 159
 Zirconia
 alumina matrix composites, 23
 ceramics
 composition and properties, 3104–3105
 dental applications, 3105–3106
 YTZP crowns, 3106–3108
 dental applications, 599–600, 3105–3106
 Zonula occludens toxin (ZOT), protein-based
 drug delivery systems, 2361–2362
 Zoonosis, xenografts, 3100

Biomedical Engineering

"...The volumes are well organized and each section consistently present[s] **the most important concepts of each topic...**
A very good first reference for researchers and students."

—*Pharmaceutical Research*

"...The encyclopedia is up-to-date and covers topics in depth...**Recommended.**"

—*Choice*

"...This work will be **of considerable value** to scientists, clinicians, and others needing a conceptual framework, data, and suggested approaches to research problems and product development."

—*Dr. Frederick J. Schoen, Professor and Executive Vice Chairman, Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA*

"...Comprehensive coverage of the field...Authors include **world-renowned clinicians, engineers and biomedical scientists from all relevant disciplines...**"

—*Dr. Harvey S. Borovetz, Professor and Chair, Department of Bioengineering, University of Pittsburgh, Pennsylvania, USA*

about the book...

Written by more than 400 subject experts representing diverse academic and applied domains, this multidisciplinary resource surveys the vanguard of biomaterials and biomedical engineering technologies utilizing biomaterials that lead to quality-of-life improvements. Building on traditional engineering principles, it serves to bridge advances in materials science, life sciences, nanotechnology, and cell biology to innovations in solving medical problems with applications in tissue engineering, prosthetics, drug delivery, biosensors, and medical devices.

In nearly 300 entries, this four-volume ***Encyclopedia of Biomaterials and Biomedical Engineering, Second Edition***, covers:

- essential topics integral to tissue engineering research: bioreactors, scaffolding materials and fabrication, tissue mechanics, cellular interaction, and development of major tissues and organs being attempted by researchers worldwide
- artificial lungs and muscles, bio-artificial livers, and corneal, dental, inner ear, and total hip implants
- tissue engineering of blood vessels, heart valves, ligaments, microvascular networks, skeletal muscle, and skin
- bone remodeling, bone cement, and bioabsorbable bone plates and screws
- controlled drug delivery, insulin delivery, and transdermal and ocular implant-based drug delivery
- endovascular stent grafts, vascular grafts, and xenografts
- 3-D medical imaging, electrical impedance imaging, and intravascular ultrasound
- biomedical, protein adsorption, and *in vivo* cardiovascular modeling
- polymer foams, biofunctional and conductive polymers, and electroactive polymeric materials
- blood–material interactions, the bone–implant interface, host reactions, and foreign body responses
- and much more

about the editors...

GARY E. WNEK is the Joseph F. Toot, Jr. Professor of Engineering, Professor and Chair of the Department of Macromolecular Science and Engineering, and Faculty Director of The Institute for Management and Engineering, Case Western Reserve University, Cleveland, Ohio. He is the author or coauthor of over 130 journal articles and 19 patents, and the coeditor of four books, including the *Handbook of Pharmaceutical Controlled Release Technology*, *Photonic Polymer Systems*, and *Electrical and Optical Polymer Systems* (Informa Press). His research interests include the development of polymeric materials for applications in medicine, microfluidics, and energy conversion and storage. A member of the American Chemical Society, the Materials Research Society, and the American Institute of Chemical Engineers, Dr. Wnek received the B.S. degree in chemical engineering from Worcester Polytechnic Institute, Massachusetts, and the Ph.D. degree in polymer science and engineering from the University of Massachusetts, Amherst.

GARY L. BOWLIN is an Associate Professor and the Louis and Ruth Harris Exceptional Scholar Professorship of Biomedical Engineering at Virginia Commonwealth University, Richmond, Virginia. He has extensive experience in cardiovascular materials and the field of tissue engineering. As a researcher, he has authored or coauthored over 50 journal articles in the areas of vascular tissue engineering and biopolymer processing for tissue engineering scaffolds. He also holds five U.S. patents and is a Co-Founding inventor of NanoMatrix, Inc., based on electrospinning tissue engineering scaffold technology and TraumaCure, Inc., regarding the development of hemostatic products. Dr. Bowlin is a member of the Biomedical Engineering Society, American Institute of Chemical Engineers, Society for Biomaterials, Tissue Engineering and Regenerative Medicine International Society, and the International Society of Applied Cardiovascular Biology. Dr. Bowlin received the B.S. in chemical engineering from Youngstown State University, Ohio, the M.S. in chemical engineering from the University of New Hampshire, Durham, and the Ph.D. degree in biomedical engineering from The University of Akron, Ohio.

Printed in the United States of America

H7956

informa

healthcare

www.informahhealthcare.com

52 Vanderbilt Avenue
New York, NY 10017
Telephone House
69-77 Paul Street
London EC2A 4LQ, UK

ISBN 978-142007956-2



9 0000

9 781420 079562